

UNIVERSIDAD DE EL SALVADOR  
FACULTAD DE INGENIERÍA Y ARQUITECTURA  
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS



**DESARROLLO DE UN “DATA WAREHOUSE” PARA EL  
PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL  
CONSUMIDOR.**

PRESENTADO POR:

**RENÉ ROLANDO CUCHILLAS CAÑAS**

**OSCAR ATILIO HERNÁNDEZ PEÑA**

**YURI HILDEBRANDO MEJÍA MERLOS**

**HÉCTOR ATILIO SILVA GARCÍA**

PARA OPTAR AL TÍTULO DE:

**INGENIERO DE SISTEMAS INFORMÁTICOS**

CIUDAD UNIVERSITARIA, FEBRERO DE 2010

**UNIVERSIDAD DE EL SALVADOR**

RECTOR :

**MSC. RUFINO ANTONIO QUEZADA SÁNCHEZ**

SECRETARIO GENERAL :

**LIC. DOUGLAS VLADIMIR ALFARO CHÁVEZ**

**FACULTAD DE INGENIERÍA Y ARQUITECTURA**

DECANO :

**ING. MARIO ROBERTO NIETO LOVO**

SECRETARIO :

**ING. OSCAR EDUARDO MARROQUÍN HERNÁNDEZ**

**ESCUELA DE INGENIERIA DE SISTEMAS INFORMÁTICOS**

DIRECTOR :

**MSC. CARLOS ERNESTO GARCÍA GARCÍA**

UNIVERSIDAD DE EL SALVADOR  
FACULTAD DE INGENIERÍA Y ARQUITECTURA  
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

Trabajo de Graduación previo a la opción al Grado de:

**INGENIERO DE SISTEMAS INFORMÁTICOS**

Titulo :

**DESARROLLO DE UN “DATA WAREHOUSE” PARA EL  
PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL  
CONSUMIDOR.**

Presentado por :

**RENÉ ROLANDO CUCHILLAS CAÑAS**

**OSCAR ATILIO HERNÁNDEZ PEÑA**

**YURI HILDEBRANDO MEJÍA MERLOS**

**HÉCTOR ATILIO SILVA GARCÍA**

Trabajo de Graduación Aprobado por :

Docente Director

**ING. ELMER ARTURO CARBALLO RUIZ**

San Salvador, febrero de 2010

Trabajo de Graduación Aprobado por:

Docente Director :

**ING. ELMER ARTURO CARBALLO RUIZ**

## AGRADECIMIENTOS

A Dios todo poderoso, por bendecirme, darme vida y fortaleza que a pesar de los obstáculos que se presentaron siempre estuvo junto a mi lado para ayudarme a superarlos.

A mis padres que con mucho esfuerzo me ayudaron a emprender este largo y duro camino, gracias por su apoyo, consejos y comprensión que en cualquier situación que se me presento. Este triunfo es fruto de sus esfuerzos se los dedico gracias por haber confiado en mí, espero que Dios los bendiga y los llene de felicidad. Mamá, gracias por cada uno de los consejos que siempre me brindaste cuando yo mas lo necesite, te quiero decir que te amo y que se que en ti puedo encontrar un consejo sabio y la amistad que muchas veces necesito. Papá fuiste mi fuerza, no tienes idea de cuánto me inspiraste, jamás salió de mi mente tu esfuerzo, cuando me sentí más cansado, siempre me recordaba que tu confiaste en mi, papá este triunfo jamás hubiese sido posible sin ti, gracias por ser el mejor papá del mundo te amo. Gracias Padres por ayudarme a hacer realidad este sueño.

A mi hermano Mauricio que se desvelo casi siempre porque no lo dejaba dormir con mi ruido, por cuidarme y apoyarme cuando más lo necesite, eres un gran hermano y aunque no parezca, te quiero mucho. A mi hermano José, que a pesar de estar lejos siempre sentí su apoyo, y mientras estuve cerca fuiste un gran ejemplo, también te quiero hermano.

A mi segunda mamá Rosita, por quererme tanto, creo que no lo merezco pero tú fuiste mi otra mama la cual me ayudo a crecer espiritualmente y a tratar de que fuera un hombre de bien, gracias mamita, te puedes sentir orgullosa por haber contribuido a que obtuviera este triunfo; también a Alonso y Lissete, por quererme como su hermanito menor, los quiero mucho.

A mi familia, tíos, tías, primos y primas, abuelo y abuelas que aunque sea molestándome siempre me apoyaban y sabía que podía contar con ustedes, gracias a todos. Son muchos y aunque no lo crean al escribir esto sus recuerdos vienen a mi memoria, gracias por su apoyo.

A mis compañeros de tesis, Yuri, Oscar y René que dejaron de ser compañeros para convertirse en amigos y que al culminar este proyecto, son considerados mis hermanos, gracias por aguantarme y por poder mantener la calma mientras los desesperaba. Gracias Amigos no duden que siempre estarán en mi memoria.

A todos mis amigos que compartieron un poquito de tiempo, desde que iniciamos la U, ya que el lograr este triunfo es el fruto de la perseverancia que tuvimos desde que iniciamos la universidad. Gracias a la Cipotada FC, que convirtió a muchos de nosotros de conocidos a un grupo de amigos, Cipotada más que un club... Gracias a mis amig@s: Karla, Carmen, Karen Marcela, Lourdes, Celina, Claudia Lopez, Walo, Eduardo, LenDj, Juanito Jans, Willie, Pipo, Mama osa(Ricardo), el chele Carlos, Daniel Coello, Nefta , Jaime , javo, schumy, Nelson Gonzales, Miguel Sanchez, Mario (chito), Oscar(timo), a Gerson Molina a quien le agradezco mucho por haber contribuido en ACC, y a todos y cada uno con los que compartimos recuerdos de las experiencias que solo la Universidad de El Salvador nos puede brindar. Además quiero agradecer por la gran celebración que nos habían preparado Ligia, Johanna, Geraldine que sin saber si pasaríamos la ultima defensa, pero ellas ya habían preparado con mucho cariño nuestra celebración, Gracias por su incondicional apoyo, de igual forma les pido disculpas si alguna vez respondí pesado, pero fue por el gran stress que generaba este proyecto cada vez que se acercaba una entrega, gracias por su apoyo y cariño que me brindaron a lo largo de este proceso gracias por hacerme sentir que no importara el resultado ustedes estarían junto a mi celebrando el resultado.

A la familia Della Valle, que me adoptaron como hijo, muchas gracias por hacerme sentir parte de su familia, gracias por tanto cariño que me brindaron, gracias don Ricardo, donna Stephy, Dulce, Andrea y Ricardo, gracias por compartir la felicidad que genera su bonita familia.

A mi Asesor Elmer Carballo, que tuvo la paciencia para aguantarnos y que nos ayudo a que fuéramos responsables y que a pesar que al principio le costó coordinarnos, terminamos siendo un verdadero grupo de trabajo que concluimos este gran proyecto que se convierte en un gran triunfo en mi vida, gracias por compartir su experiencia y tiempo para este proyecto.

A los docentes por su tiempo y enseñanzas que contribuyeron a mi formación académica. A nuestro observador Pedro Peñate por el tiempo y dedicación que tuvo al evaluar este proyecto.

A la defensoría del consumidor, que gracias a la contribución de todos las personas que forman esta organización este proyecto puedo culminarse exitosamente.

A mis compañeros de trabajo Emerson, Ibeth, Canton, Alejandro, Mauricio Sánchez, Graciela de López, Ayme de Aguilar, René Alfaro, Silvia Solano, Carina Hernández, Wendy Castellanos y a todos los compañeros que estuvieron pendientes del proceso de tesis.

Gracias a TODOS  
**Héctor Silva**

## AGRADECIMIENTOS

Ahora, restando hacer mis agradecimientos como parte final de mi tesis y de mi carrera, me permito dedicarle estas palabras a Dios todo poderoso a quien le agradezco grandemente por permitirme llegar hasta este punto culminante de mi carrera, terminando con un exitoso proyecto de graduación.

Agradezco a mi familia, a mi padre José Elías Mejía quien se esforzó por apoyarme, alentarme e inspirarme para perseverar en mis estudios y alcanzar este logro, a mis hermanos Elías Alexander Mejía Merlos, a Edward Ernesto Mejía Hernández y Jeffrey Geovani Mejía Hernández por ser todos ellos parte incondicional de mi vida y de todo mi estudio durante mi carrera. No tengo las palabras suficientes para expresar la grandeza de mi agradecimiento a mi familia, por el apoyo y la fortaleza que me han dado durante los buenos momentos y los malos momentos que hemos pasado juntos, superando cada dificultad hasta hoy. A mi madre Tiburcia de Jesús Merlos por su apoyo a pesar de la distancia y dificultades por estar a mi lado.

Agradezco a mis amigos y compañeros de tesis Héctor Atilio García Silva, René Rolando Cuchillas Cañas y Oscar Atilio Hernández Peña por haber integrado junto a mí, el grupo de trabajo de graduación, aceptando un gran reto como lo fue nuestro proyecto desarrollado y culminado con éxito. Por todas esas noches de desvelo, por todos esos días que no hubieron noches, si no, días sin dormir. Por soportar y sobre llevar cada dificultad que se nos presentaba, esforzándonos juntos. Por haber soportado mi persona, mi carácter, por haberme apoyado en los momentos difíciles y por haberme alentado para esforzarnos como grupo en desarrollar y culminar un trabajo de excelencia. A Héctor Silva alias teto por haber ofrecido sin peros, su casa, su apoyo, su amistad y todo lo que necesitáramos con el fin de mantenernos siempre unidos y apoyándonos mutuamente como grupo, por haber puesto antes de su empleado nuestro trabajo de graduación. A René Cuchillas alias Robot, quien fue parte fundamental como todos los demás durante todo el trabajo de graduación, por sacrificar días importantes como su cumpleaños, por haber pues primero el trabajo de graduación antes de su empleado. Ha Oscar Hernández alias Herba, amigo y compañero, que fue desde el principio parte importante, ante poniendo siempre el trabajo de graduación a su empleo.

A nuestro asesor, director de proyecto y amigo Ing. Elmer Arturo Carballo, por guiarnos y alentarnos en todo el proceso, por haber creído en mí y mi grupo de graduación para tomar y desarrollar un trabajo de graduación único y de gran valor, demostrándonos que es más que un trabajo de graduación, es un valor agregado a nuestro conocimiento, un desafío superado y un premio ganado como profesionales. Por habernos exigidos un esfuerzo mas allá de lo que jamás habíamos dado en toda la carrera, por hacernos descubrir el valor que existe en el buen trabajo, dedicación y en el esfuerzo extra para superar las expectativas de nosotros mismo y los demás. A nuestro asesor Ing. Pedro Peñate, quien fue indiscutiblemente parte importante en el proceso de trabajo de graduación.

A todos mis amigos a quienes conocí durante toda mi carrera, a Carmen Gutiérrez, Karla Rodríguez, Enrique Rodríguez, Fátima Montoya, Gerson Molina, mi gran amiga Irma Vides, Kary Guzmán, Leni Ricardo, Mama Osa, El Chele, Ligia Lorena López, Geraldine Gonzales, Mirna Rivas, Mirna Martínez, Johanna Ramírez, Orlando Oswaldo, Wilmer Aguilar, Willie Vásquez, Nelson Gonzales, Claudia Merino, Karla Verence, Carolina Melara, Claudia Arabia, a mi gran amiga Mary Montoya, Violeta Menjivar, Sandra Aragón, Wilmer López, Iliana Isabel Ponce, Ada Lovo, Edwin Ramírez. A la CIPOTADA F.C.; Miguel Sánchez, Christian Martinez colochon, Eduardo Barrera, Roberto Cienfuegos, Mónico Alvarado, Schummy, Oscar Mendoza Timo, Javier Seguro, Jaime Beltran Cazano, Elvis Ramírez, Cheyo y Jorge el pelón. Y a todos aquellos con quien compartí durante toda mi carrera en la Universidad Nacional, a todos

los llevare en mi memoria y en mi corazón, por haber hecho de mi carrera y vida un mundo único y agradable.

También quiero agradecer de una manera muy especial a Cecilia Yescas y a Juan José Rivas, Gerente y Jefe de Desarrollo respectivamente de la Gerencia de Sistemas de Información de la Defensoría del Consumidor, por su apoyo constante e incondicional, para lograr desarrollar y culminar mi trabajo de graduación.

Agradecer a mis compañeros de trabajos Susana Rodríguez, Silvia Solano, Ayme de Aguilar, Carina Hernández, Wendy Castellanos, Ibeth Peña, Iliana Monterrosa, Emerson Reyes y Martita Mendoza.

Gracias a todas las personas que han sido mis amigos y compañeros que omito por olvido, cada uno de mis familiares, por estar de una u otra manera dentro de mi vida y haber sido parte de mi carrera. Esta victoria está dedicada a todos ellos y al todo poderoso ante todo.

**Yuri Hildebrando Mejía Merlos**

## AGRADECIMIENTOS

A Dios todopoderoso quien otorga sabiduría, paciencia, y siembra las fuerzas para culminar las metas propuestas, permitiendo aquellas situaciones que generan alegría, tristeza, emoción, aflicción y un sinfín de emociones las cuales puedo afirmar que las hemos vivido en todo este proceso. Gracias Dios por ser el mejor Arquitecto de Sistemas, creador de millones de sistemas dentro sistemas mayores en todo lo que existe, en lo que vemos y aun en lo que no vemos, aplicando a la perfección la sinergia en todo.

Muchas gracias a Oscar Antonio Hernández, y Ana Silvia Peña de Hernández, mis padres, por su apoyo incondicional para alcanzar esta meta y brindarme su ayuda y soporte para finalizar esta etapa de mi vida a quienes dedico este logro retribuyendo su esmero por hacer de mi una persona de bien para la sociedad, a mi hermana Silvia Isabel, por estar ahí apoyándome y aconsejándome sabiamente. A mi abuelita, Isabel de Leiva, por sus oraciones y apoyo hacia mi vida y mi grupo de trabajo, se que Dios escucho esas oraciones y por eso estoy muy agradecido. Reina Hernández, quien me brindo su apoyo incondicional cuando lo necesité, realmente estoy muy agradecido por todo lo que ha hecho y por su amor y cariño. A toda mi familia en general por ser parte de este sueño y este anhelo de alcanzar un logro profesional y por apoyarme siempre. Un agradecimiento especial a Geraldine González, por todo su apoyo y ayuda durante este trayecto, por brindarme una amistad incondicional y sus consejos, asimismo a Marcia Alfaro por sus oraciones y sus palabras de apoyo para seguir adelante.

Agradezco públicamente a mi grupo de trabajo, René Cuchillas, Yuri Mejía y Héctor Silva, quienes fueron un apoyo importante para culminar este proyecto, su trabajo, esfuerzo responsabilidad y motivación fue parte fundamental para alcanzar los objetivos propuestos, además de su amistad en todo momento, su cariño y sus palabras de apoyo fueron de gran ayuda en los momentos más difíciles, Gracias a ustedes, por permitirme formar parte de su equipo y por colaborar junto conmigo para finalizar lo que en un inicio nos propusimos.

A nuestro asesor Ing. Elmer Arturo Carballo por la paciencia, consejos y por todo su apoyo en todo este proceso, quien con esfuerzo nos brindo su máximo apoyo hasta el último día, muchas gracias por todo lo que hizo por nosotros y por creer que éramos capaces de dar más cada día. A la Defensoría del Consumidor por haber permitido desarrollar este proyecto y haber confiado en nuestras capacidades como Analistas de Negocios, Cecilia Yescas y Juan Rivas.

A mis amigos quienes estuvieron pendientes del proceso de este trabajo: Edgardo, Flor, Violeta, Jess, kary, Mellanie, Osa, Johanna, Ligia, Taba, Rene, Ricardo, Marcela, Chito, Armando, y todos mis amigos que no menciono pero saben que están en mi corazón, realmente no tengo palabras para agradecer lo que han hecho por mí y el grupo en general en este proceso.

A la Universidad de El Salvador, que me permitió ser parte de una nueva historia en sus páginas y por ser quien me impulso como una persona de éxito preparada para afrontar los retos de una nueva etapa en mi vida. Queridos docentes, asesores y amigos que formaron parte de toda esta carrera...  
¡Muchas gracias!

**Oscar Atilio Hernández Peña**

## **AGRADECIMIENTOS**

En primer lugar quiero agradecer a Dios por haberme guiado a lo largo de toda la carrera universitaria, el trabajo de graduación y en toda mi vida. Además por haberme brindado la fortaleza necesaria para superar las adversidades que se fueron presentando para poder conseguir este logro tan importante.

También quiero agradecer a mis padres, René Cuchillas y Rosa Aminta Cañas, por haberme apoyado incondicionalmente en esta etapa de mi vida, por el sacrificio que han realizado durante todos estos años y por la comprensión que han tenido durante todo este tiempo. Agradezco a mi hermana Karen por haberme escuchado y aconsejado en los momentos difíciles que se presentaron y a toda mi familia que estuvo pendiente y siempre mostrando su apoyo.

Un agradecimiento muy especial a mis amigos y compañeros de trabajo de graduación, Oscar Hernández, Yuri Mejía y Héctor Silva, por haber tenido la paciencia, comprensión, apoyo y por haber dado todo el esfuerzo para lograr este objetivo, que en conjunto nos habíamos trazado. Además por haber compartido alegrías, tristezas, viajes, desvelos y enojos durante todo este tiempo.

De igual forma, le agradezco a Johanna Ramírez, Geraldine González y Ligia López por habernos ayudado y apoyado durante todo este año, especialmente en la organización de las defensas de las diferentes etapas.

Agradezco también a Edgardo, Jessica, Flor, Raúl, Roberto, Eduardo, Marcela, Armando, Chito, José, Ricardo, Emerson, Enrique y todos los mostró su apoyo para poder terminar la carrera.

Al asesor del Trabajo de Graduación, el Ing. Elmer Arturo Carballo, por haber creído en el proyecto y por haber transmitido su experiencia, por el tiempo dedicado al grupo y por todas las exigencias para que este proyecto saliera adelante y que tuviera un valor agregado.

A la Defensoría del Consumidor, en especial a Cecilia Yescas y Juan José Rivas, por haber permitido la realización de este proyecto y habernos proporcionado la información necesaria, su tiempo y confianza.

**René Rolando Cuchillas Cañas**

## ÍNDICE

INTRODUCCIÓN.....	V
OBJETIVOS .....	VII
OBJETIVO GENERAL.....	VII
OBJETIVOS ESPECÍFICOS .....	VII
ALCANCE Y LIMITACIONES .....	VIII
ALCANCE.....	VIII
LIMITACIONES.....	VIII
JUSTIFICACIÓN .....	1
IMPORTANCIA.....	5
1. CAPITULO I: GENERALIDADES.....	6
1.1. LISTADO DE ACRÓNIMOS .....	7
1.2. MARCO TEÓRICO .....	8
1.2.1. <i>Business Intelligence</i> .....	8
1.2.2. <i>Sistemas Transaccionales (OLTP)</i> .....	10
1.2.3. <i>Sistemas OLAP</i> .....	11
1.2.4. <i>Data Warehouse</i> .....	12
1.2.5. <i>Data Marts</i> .....	14
1.2.6. <i>Dashboards</i> .....	14
1.2.7. <i>Introducción a Metodologías de Business Intelligence</i> .....	14
1.3. ESTÁNDARES Y DESCRIPCIÓN DE NOMENCLATURA .....	20
1.3.1. <i>Análisis de la situación actual</i> .....	20
1.3.2. <i>Nomenclatura de casos de uso</i> .....	21
1.3.3. <i>Descripción de Estereotipos utilizados para el Diseño del Data Warehouse</i> .....	22
1.3.4. <i>Diseño del proceso ETL (Extract, Transform and Load)</i> .....	24
1.3.5. <i>Estándar de nomenclatura de tablas y campos del data warehouse</i> .....	26
1.4. FORMATOS DE ESPECIFICACIÓN PARA EL ANÁLISIS DE REQUERIMIENTOS. ....	27
1.4.1. <i>Formato de especificación de requerimientos de usuario</i> .....	27
1.4.2. <i>Formato de especificación de requerimientos de sistema.</i> .....	28
1.4.3. <i>Metodología de obtención de requerimientos</i> .....	29
1.5. ANTECEDENTES .....	30
1.5.1. <i>La Defensoría del Consumidor</i> .....	31
1.5.2. <i>Servicios</i> .....	34
2. CAPITULO II: ANÁLISIS .....	36
2.1. ANALISIS DE LA SITUACION ACTUAL.....	37
2.1.1. <i>Resultados de la investigación actual</i> .....	37
2.1.2. <i>Análisis de la situación actual</i> .....	40
2.2. PLANTEAMIENTO DEL PROBLEMA .....	43
2.2.1. <i>Análisis de la Situación Problemática</i> .....	43
2.2.2. <i>Definición del problema</i> .....	46
2.2.3. <i>Formulación del problema</i> .....	46
2.3. SOLUCIÓN PROPUESTA.....	48
2.3.1. <i>Identificación de la solución</i> .....	48
2.3.2. <i>Representación de la solución</i> .....	49
2.4. ENFOQUE DE SISTEMAS DE LOS REQUERIMIENTOS. ....	50
2.5. CONTENIDO TEMÁTICO PROPUESTO .....	53
2.6. FACTIBILIDADES .....	57

2.6.1. Factibilidad Técnica .....	57
2.6.2. Factibilidad Operativa .....	66
2.6.3. Factibilidad Económica.....	71
2.7. PLANEACIÓN DE RECURSOS .....	75
2.7.1. Calculo para el Presupuesto .....	75
2.7.2. Resumen de Gastos .....	79
2.8. METODOLOGÍA.....	80
2.8.1. Herramientas para Recolección de Datos.....	80
2.8.2. Herramientas para el Diagnostico y Planteamiento del Problema.....	82
2.8.3. Modelo UML de los Componentes Bodega de Datos .....	84
2.8.4. Ciclo de vida .....	84
2.8.5. Metodología del Proyecto.....	87
2.8.6. Metodología de Kimball .....	92
2.9. RESULTADOS ESPERADOS .....	96
2.9.1. Resultados Esperados.....	96
2.9.2. Producto a Entregar.....	97
3. CAPITULO III: DISEÑO Y REFINAMIENTO .....	99
3.1. DISEÑO GENERAL.....	100
3.1.1. Modelado del Data Warehouse usando UML .....	100
3.1.2. Diseño Físico del DW (Ejemplo).....	103
3.2. DISEÑO ESPECÍFICO DE LA SOLUCIÓN PROPUESTA .....	107
3.2.1. Base de Datos Transaccional.....	107
3.2.2. ETL (Extract, Transform and Load).....	107
3.2.3. Data Warehouse.....	108
3.2.4. Cubo OLAP .....	109
3.2.5. Interfaz de Reportes .....	109
3.3. DIAGRAMA DE SECUENCIA.....	110
3.3.1. Validar usuario.....	110
3.3.2. Gestionar Roles.....	111
3.3.3. Gestionar Usuario.....	112
3.3.4. Gestionar Reportes.....	113
3.3.5. Gestionar Análisis.....	114
3.3.6. Editar Contenido.....	115
3.3.7. Compartir Contenido.....	116
3.3.8. Gestionar tareas programas.....	117
3.4. REQUERIMIENTOS Y ESPECIFICACIÓN DE USUARIOS.....	118
3.4.1. Listado de Requerimientos.....	118
3.4.2. Requerimientos Técnicos .....	118
3.4.3. Requerimientos Operativos .....	119
3.4.4. Requerimientos de Desarrollo .....	120
3.4.5. Requerimientos No Funcionales .....	121
3.4.6. Requerimientos No Funcionales Técnicos .....	123
3.4.7. Descripción de los perfiles de Usuario .....	124
3.5. ITERACIÓN 1: CENTRO DE SOLUCIÓN DE CONTROVERSIAS.....	125
3.5.1. Requerimientos Funcionales Para Centro de Solución de Controversias .....	125
3.5.2. Especificación de Requerimientos .....	128
3.5.3. Modelado Iteración 1 Centro de Solución de Controversias .....	137

3.5.4. Diagrama UML del DATAWAREHOUSE – CENTRO DE SOLUCION DE CONTROVERSIAS .....	143
3.5.5. Diagrama del proceso ETL .....	144
3.6. ITERACIÓN 2: VERIFICACIÓN Y VIGILANCIA .....	152
3.6.1. Requerimientos Funcionales .....	152
3.6.2. Especificación de Requerimientos .....	154
3.6.3. Modelado Iteración 2 Modulo de Verificación y Vigilancia .....	161
3.6.4. Diagrama UML del DATAWAREHOUSE – VERIFICACION Y VIGILANCIA.....	166
3.6.5. Diagrama del proceso ETL.....	167
3.7. ITERACIÓN 3: TRIBUNAL SANCIONADOR.....	175
3.7.1. Requerimientos Funcionales .....	175
3.7.2. Especificación De Requerimientos.....	177
3.7.3. Modelado Iteración 3 Tribunal Sancionador.....	184
3.7.4. Diagrama UML del DATAWAREHOUSE – TRIBUNAL SANCIONADOR .....	189
3.8. ITERACIÓN 4: ANFITRIÓN.....	198
3.8.1. Requerimientos Funcionales .....	198
3.8.2. Especificación de Requerimientos .....	199
3.8.3. Modelado Iteración 4: Anfitrión.....	206
3.8.4. Diagrama UML del DATAWAREHOUSE – ANFITRION.....	211
3.8.5. Diagrama del proceso ETL.....	212
3.9. ITERACCION 5: CALL CENTER .....	218
3.9.1. Requerimientos Funcionales Para Call Center .....	218
3.9.2. Especificación de Requerimientos .....	220
3.9.3. Modelado Iteración 5 Call Center.....	227
3.9.4. Diagrama UML del DATAWAREHOUSE – CALL CENTER .....	233
3.9.5. Diagrama del proceso ETL.....	234
3.10. DISEÑO DE SALIDAS .....	243
3.11. DISEÑO DE SEGURIDAD.....	245
3.11.1. Autorización.....	245
3.11.2. Seguridad Usuario/Rol.....	245
3.11.3. Seguridad por Action Sequence.....	245
3.11.4. Seguridad en Mondrian .....	245
3.11.5. Seguridad en reportes Ad Hoc .....	246
3.12. REFINAMIENTO DEL PROTOTIPO .....	248
4. CAPITULO IV: DOCUMENTACIÓN.....	250
4.1. MANUAL DE INSTALACIÓN .....	251
4.1.1. Instalación y Configuración de Pentaho BI Server.....	251
4.1.2. Instalación y Configuración de Pentaho Metadata Editor .....	265
4.1.3. Instalación Y Configuración De Java SE Development Kit (Jdk) .....	269
4.1.4. Instalación de Mysql.....	275
4.1.5. Configuración de SQL Server.....	284
4.1.6. Configuración de Superficie de SQL Server 2005 .....	288
4.1.7. Acceder a Configuración de Superficie de SQL Server 2005.....	288
4.1.8. Configuración de conexiones locales y remotas .....	288
4.2. MANUAL TÉCNICO.....	296
4.2.1. Pentaho BI Server .....	296
4.2.2. Manual Schema Workbench .....	304
4.2.3. Pentaho Metadata Editor.....	316

4.2.4. <i>Publicación en Pentaho BI Server</i> .....	331
4.2.5. <i>Integration Services</i> .....	333
4.3. MANUAL DE USUARIO.....	366
4.3.1. <i>Configuración Básica del Equipo de los Usuarios</i> .....	366
4.3.2. <i>Navegación del Sistema</i> .....	369
4.3.3. <i>Creación de Reportes</i> .....	378
4.3.4. <i>Creación de Vista de Análisis</i> .....	392
5. CAPITULO V: PLAN DE IMPLANTACIÓN.....	409
5.1. ELEGIR LA ARQUITECTURA.....	410
5.1.1. <i>Recurso Tecnológico para la Implementación</i> .....	410
5.1.2. <i>Recurso Humano</i> .....	411
5.1.3. <i>Descripción de la Arquitectura del Sistema</i> .....	414
5.1.4. <i>Herramientas para el Desarrollo del Proyecto</i> .....	417
5.2. METODOLOGÍA DE LA IMPLANTACIÓN.....	420
5.3. PLANIFICACIÓN DE ACTIVIDADES .....	421
5.3.1. <i>Planeación de Implantación</i> .....	421
5.3.2. <i>Organización Implantación</i> .....	421
5.3.3. <i>Ejecución de la Implantación</i> .....	422
5.4. CRONOGRAMA DE ACTIVIDADES .....	434
5.5. SISTEMA DE CONTROL.....	436
5.5.1. <i>Formularios para el Control</i> .....	436
5.5.2. <i>Estrategias De Control</i> .....	443
CONCLUSIONES.....	444
RECOMENDACIONES .....	446
BIBLIOGRAFÍA.....	447
LIBROS.....	447
TESIS.....	447
SITIOS WEB.....	447
GLOSARIO TÉCNICO.....	449
ANEXOS.....	452
ANEXO 1: MATRIZ DE EVALUACIÓN DE CICLO DE VIDA .....	453
ANEXO 2: FORMULARIO PARA EVALUACIÓN DEL PROTOTIPO .....	454
ANEXO 3: CALCULO DE LA MUESTRA DE LA POBLACIÓN .....	455
ANEXO 4: RESULTADOS DE LAS ENCUESTAS .....	457
ANEXO 5: EVALUACIONES TÉCNICAS .....	468
ANEXO 6: INTERFAZ DE INGRESO DE DATOS.....	472
ANEXO 7: DESCRIPCIÓN DEL MODELO UML DE LOS COMPONENTES BODEGA DE DATOS.....	474
ANEXO 8: DIAGRAMAS ENTIDAD RELACION .....	481
ANEXO 9: MINERÍA DE DATOS.....	483
ANEXO 10: IMPLEMENTACIÓN DEL LOG EN LOS PAQUETES.....	499
ANEXO 11: ACEPTADOS POR DEFENSORIA DEL CONSUMIDOR .....	500
ANEXO 12: DIAGRAMAS DE PLAN DE IMPLANTACIÓN .....	517

## INTRODUCCIÓN

Con el continuo desarrollo en el área social y económico de nuestro país las actividades económicas de venta, compra de inmuebles, prestaciones de servicios y otros aumentan considerablemente, razón por la cual la Defensoría de Consumidor entra a protagonizar un rol de acción importante dentro de este ambiente, protegiendo efectivamente los derechos de los consumidores y consumidoras, facilitando la solución de conflictos en materia de consumo, acercando a los servicios, profundizando la vigilancia de los proveedores, promoviendo la educación y participación ciudadana, para contribuir a un mejor funcionamiento del mercado.

Como parte del desarrollo organizacional para hacer frente a esta demanda la Defensoría del Consumidor se ve en la necesidad de adquirir tecnologías informáticas que brinden una solución a la administración de la información de tal manera que este recurso pueda tenerse justo a tiempo cuando se necesite tomar decisiones, dar informes y tomar acciones inmediatas. René Rohrbeck menciona en la primera conferencia internacional sobre estrategias innovadoras y creación futura:

“Aunque la detección y anticipación no son particularmente difíciles, construir una estructura organizativa que facilite una respuesta efectiva puede probar un verdadero desafío”.

En la presente se expone el ciclo completo de desarrollo de un proyecto de Inteligencia de Negocios orientado al proceso de denuncias de la Defensoría del Consumidor. Inicialmente se presenta el Anteproyecto en el cual se realiza el planteamiento del problema, análisis del mismo, identificación de la solución, evaluación de las alternativas que brinden la solución y la selección de la alternativa y sus demás elementos inmersos en la misma, como lo es la metodología de desarrollo, herramientas a utilizar y la planeación de los recursos necesarios para llevar a cabo el proyecto. Dentro de la especificación de las herramientas a utilizar, se detalla que la solución a implementar deberá que contar con tecnología bajo licencias freeware, siendo este punto parte de la estrategia organizacional del desarrollo de soluciones con bajos costos y con un gran aporte, en combinación con la tecnología ya existente dentro de la organización.

Seguido a las generalidades como documento guía, se presenta la ejecución del proyecto; “Desarrollo de un Data Warehouse para el proceso de denuncias de la Defensoría del Consumidor”, basado en la metodología de ciclo de vida prototipado, el cual se apega a la metodología de Kimball para el desarrollo de Almacenes de Datos. Cada Iteración realizada dentro del desarrollo del proyecto estará compuesta por el levantamiento de los requerimientos de los usuarios, creación de casos de uso, modelado del esquema solución para el respectivo modulo (definiendo los tres niveles; definición de modelo, definición del esquema estrella y definición de Dimensiones/Hecho) y el modelado del proceso ETL (siglas de ingles de Extraer, Transformas y Cargar). La ejecución del proyecto estará compuesto por 5 iteraciones, las cuales son parte del alcance del proyecto, abarcando 5 módulos que actualmente están en funcionamiento en el sistema SARA de la Defensoría del Consumidor, los cuales son: Centro de Solución de Controversias, Tribunal Sancionador, Call Center, Verificación y Vigilancia y Anfitrión. La tecnología a utilizar para la implementación de la solución se contempla Pentaho BI Suite la plataforma de Business Intelligence en combinación con SQL Server 2005 y las demás herramientas; Analysis Services e Integration Services.

Con el objetivo de superar las expectativas de la solución propuesta, el proyecto contempla como parte complementaria que se brindara a la Defensoría del Consumidor, la realización de análisis de información mediante pronósticos y proyecciones, utilizando Analysis Services y el algoritmo de regresión lineal. Con este punto se estará brindando una solución integral con la cual se podrá analizar la información del pasado, presente y futuro.

La documentación esta compuesta por los manuales de Configuración e Instalación, Técnico y de Usuario y el Plan de Implantación de la Solución, aspectos que contribuyen al mantenimiento del sistema y operatividad del mismo.

El desarrollo y la culminación exitosa del presente proyecto dejan como precedente la prueba fehaciente de la potencialidad que existe en las herramientas freeware para la implementación de una solución basada en el trato y manejo eficiente e inteligente de la información. Así mismo mostrando la disponibilidad de las organizaciones gubernamentales para acoger tecnologías que impliquen un bajo costo y que brinden una solución integra que llene y supere las expectativas para las unidades organizacionales y usuarios finales. Se espera que este documento contribuya al desarrollo de proyectos de vanguardia, como es hoy en día la Inteligencia de Negocios y motive al lector inmerso en tecnologías de punta el uso de alternativas de Software Libre.

## OBJETIVOS

### Objetivo General

- Desarrollar un “Data Warehouse” para el proceso de denuncias de la defensoría del consumidor, fundamentado en los requerimientos del usuario cumpliendo con los módulos en previamente definidos.

### Objetivos Específicos

- Analizar la información almacenada y generada por los módulos de SARA y el sistema ESPACIO, para conocer el comportamiento y relación de estos.
- Establecer una metodología que permita fijar lineamientos al proyecto, para facilitar el desarrollo del data warehouse.
- Establecer los módulos de estudio en base a los requerimientos del usuario identificando sus relaciones y definiendo los identificadores para cada uno de ellos.
- Aplicar técnicas de recolección de información para elaborar el análisis de los requerimientos del “Data Warehouse” para el proceso realizado en los módulos de Centro de solución de Controversias, verificación y vigilancia, anfitrión, tribunal Sancionador y Call Center.
- Analizar los procesos e interacción entre estos y el recurso humano participante en los módulos de Centro de solución de controversias, verificación y vigilancia, Anfitrión, Tribunal Sancionador y Call Center a fin de determinar las variables, establecer las dimensiones y medidas del Cubo de información.
- Definir la arquitectura<sup>1</sup> del “Data Warehouse” por medio de análisis y diseño orientado a objetos a través de diagramas UML.
- Diseñar la estructura de base de datos del Data Warehouse para establecer la base de la información que se manejara en el prototipo de cada uno de los módulos.
- Desarrollar un Cubo de información OLAP que apoye a cada una de las áreas de la Defensoría del Consumidor en la generación de reportes para análisis y mejora en la toma de decisiones.
- Realizar la documentación concerniente al data warehouse, Cubo de información y el diseño y creación de interfaces de consulta mediante la herramienta de software libre, para facilitar el mantenimiento del sistema

---

<sup>1</sup> **Arquitectura:** Conjunto de patrones y abstracciones coherentes que proporcionan el marco de referencia necesario para guiar la construcción del software para un sistema informático. Establece los fundamentos para que analistas, diseñadores, programadores, etc. trabajen en una línea común que permita alcanzar los objetivos del sistema de información, cubriendo todas las necesidades.

## **ALCANCE Y LIMITACIONES**

### **Alcance**

- Un sistema de análisis basado en la integración de la información generada por los sistemas SARA/ESPACIO con sus módulos: Centro de Solución de Controversias, Verificación y vigilancia, Anfitrión, Tribunal Sancionador y Call Center, que permita tener acceso a información detallada haciendo uso de un Data Warehouse.
- La creación de procesos ETL (son las siglas en inglés de Extraer, Transformar y Cargar) que permitan la extracción de datos contenidos en los sistemas SARA/ESPACIO logrando cumplir con las 3 operaciones básicas del ETL las cuales son: extracción, transformación y carga.
- El plan de implantación especificando los pasos y requisitos necesarios para que la información del Data Warehouse se encuentre en producción, ya que el grupo no será responsable de esta etapa.
- Elaborar los manuales de instalación y configuración, técnicos y de usuario para que faciliten la comprensión del sistema.

### **Limitaciones**

- Debido a que el proyecto contará con el apoyo total de la Defensoría del Consumidor, brindando el hardware, software y el acceso a la información necesaria, no existen limitantes para la realización del proyecto.

## JUSTIFICACIÓN

La Defensoría del Consumidor genera información desde que un consumidor inicia el proceso de apoyo en determinado caso de abuso por parte de algún comercio, servicio o empresa que trate de violar los derechos que se poseen como consumidor. Desde una simple asesoría hasta llegar a tribunal sancionador cada caso es monitoreado para llevar un control en el estado y cumplimiento de las etapas que envuelve el proceso establecido por la Dirección General.

El departamento de informática es el encargado de expandir los módulos del sistema SARA el cuales tienen como principal objetivo automatizar los procesos que se llevaban manualmente y también brindar una herramienta no solo a los trabajadores de la Defensoría sino que a aquellas personas que posean acceso a una computadora y conexión a internet para lograr beneficiarse de servicios en línea que proporcione en un futuro dicha institución.

Existen doce módulos que componen el sistema SARA de los cuales cinco de ellos han sido implementados en las tres regionales y los diferentes centros de asesorías con los que cuentan en San Salvador. De los siete módulos restantes existen cuatro que están completos pero no implementados aun, y tres que están en etapas finales de desarrollo. El trabajo del área de informática se ve interrumpido cuando en determinado momento se solicita un reporte del cual probablemente solo se utilice una tan sola vez.

El contenido de estos reportes es utilizado en dos escenarios, uno como análisis interno para las gerencias inmediatas dependiendo del área que lo solicita, dicha información puede ser vital para el nivel estratégico y táctico para toma de decisiones, así como para análisis del trabajo que se desarrolla como institución. El otro uso que se le puede dar a la información de los reportes es para brindar a la población en general un estado de los casos como tal, muchas veces en alguna conferencia o entrevistas la Gerencia General solicita datos específicos de casos a los que se les da seguimiento de la institución, para lo que se obtienen cifras significativas de las denuncias generadas por la población y con esta información se puede motivar a la población económicamente activa específicamente a los consumidores a que hagan uso del apoyo que la defensoría brinda como institución.

La creación de un “Data Warehouse” para el proceso de denuncias de la Defensoría del Consumidor, traerá un gran valor agregado para el análisis de la información y la toma de decisiones, ya que permitirá acceder a grandes volúmenes de información para realizar estudios, evaluar procesos, reduciendo tiempos de desarrollo informático y supliendo la necesidad de poseer una herramienta que facilite la generación de reportes en general. Asimismo se logra incentivar al nivel estratégico a que implemente la inteligencia de negocios, llevando a la defensoría a un nivel más maduro en cuanto a informática se refiere, utilizando como una capa de presentación una aplicación tan conocida por la mayoría de usuarios como es un navegador WEB.

Estableciendo una muestra significativa (ver ANEXO 3: Calculo de la muestra de población) para brindar un fundamento real para el desarrollo del data warehouse para la defensoría, se llevo a cabo una investigación, haciendo uso de dos entrevistas las cuales fueron lanzadas a usuarios claves dentro de la organización. Considerando como usuarios claves a todos aquellos que en al menos una ocasión han solicitado la elaboración de un reporte dentro del periodo de tiempo que se han encontrado trabajado en dicha institución. Los usuarios encuestados se clasifican en:

- 8 Usuarios de Nivel Estratégico.
- 5 Usuarios de Nivel Táctico.

16 Usuarios de Nivel Operativo.

La investigación refuerza la idea que el “DESARROLLO DE UN DATA WAREHOUSE PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR” tiene un apoyo del 100% por parte de las gerencias y se logra identificar como necesidad actual de una herramienta que proporcione acceso sencillo para la creación de reportes de información relevante en el proceso de denuncias. (Véase ANEXO 4: Resultados de las encuestas, pregunta 1)

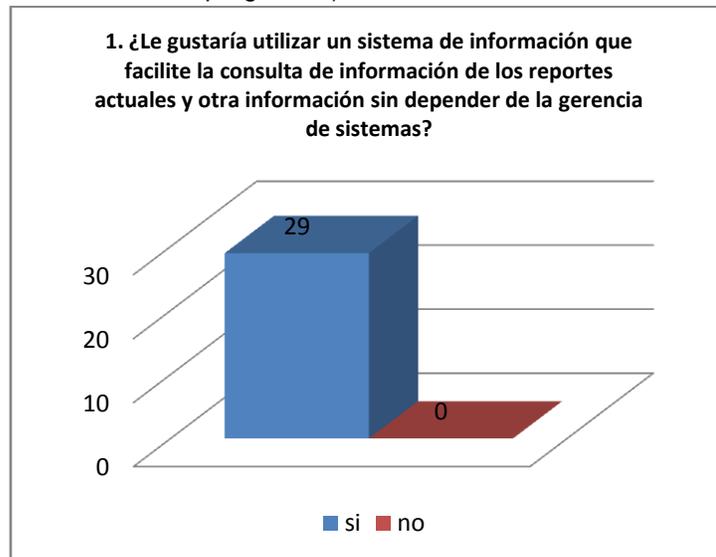


Grafico 1: Aceptación del Data warehouse para el proceso de denuncias de la DC

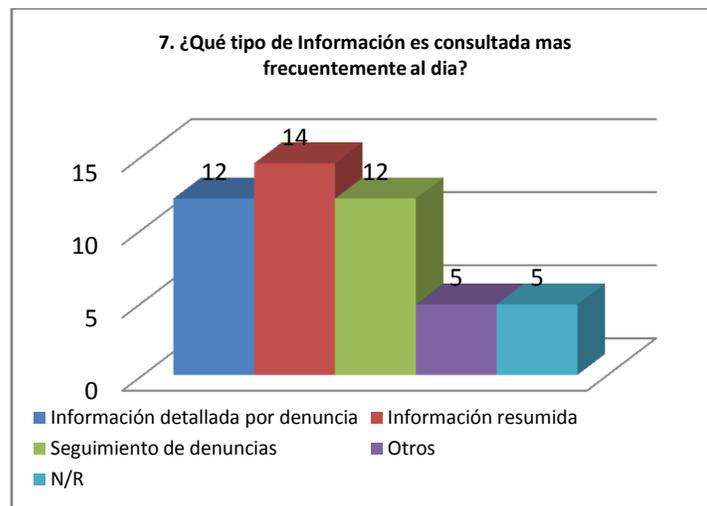
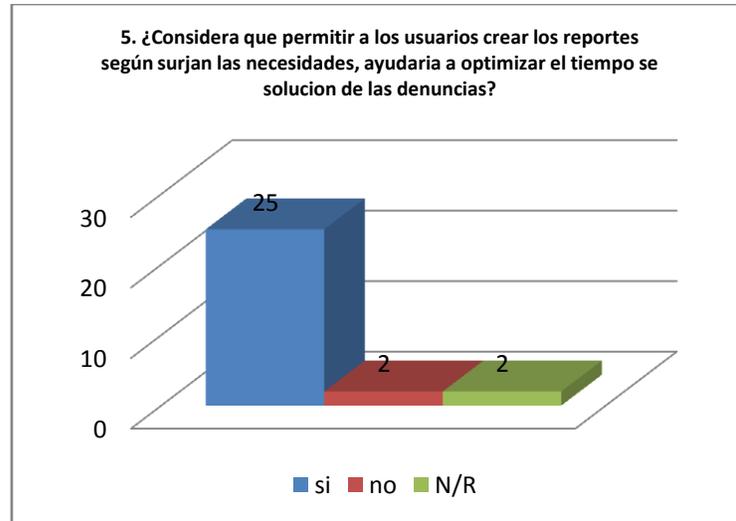


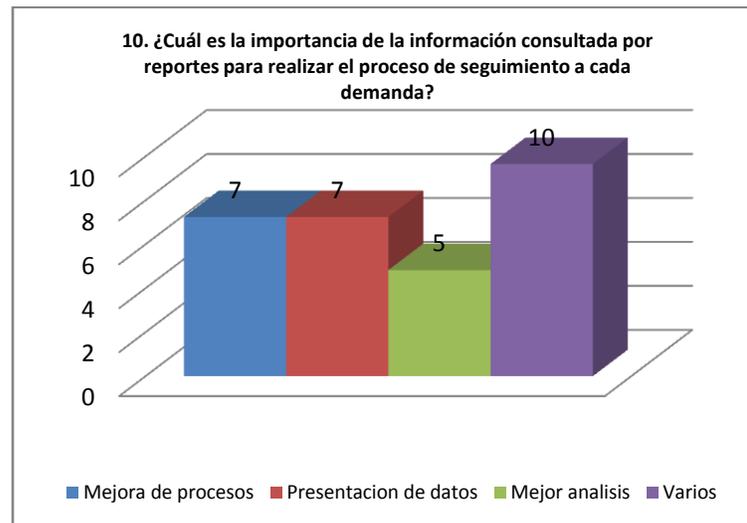
Grafico 2: Tipo de información a la que los usuarios prefieren acceder

Es importante mencionar que los usuarios consideran que el acceso a la información es de gran importancia (GRAFICO 2), el hecho de generar reportes por ellos mismos lo consideran como una mejora en el rendimiento de su trabajo diario, el 86% de los usuarios están de acuerdo que sería bueno utilizar una herramienta que facilite el acceso a los datos y que les permita generar reportes sin necesidad de solicitarlo a la Gerencia de Informática como se hace hasta la fecha. El 14% de los usuarios no dependen del sistema de consulta actual. (GRAFICO 3)



**Grafico 3: Optimización en el tiempo de solución de las denuncias**

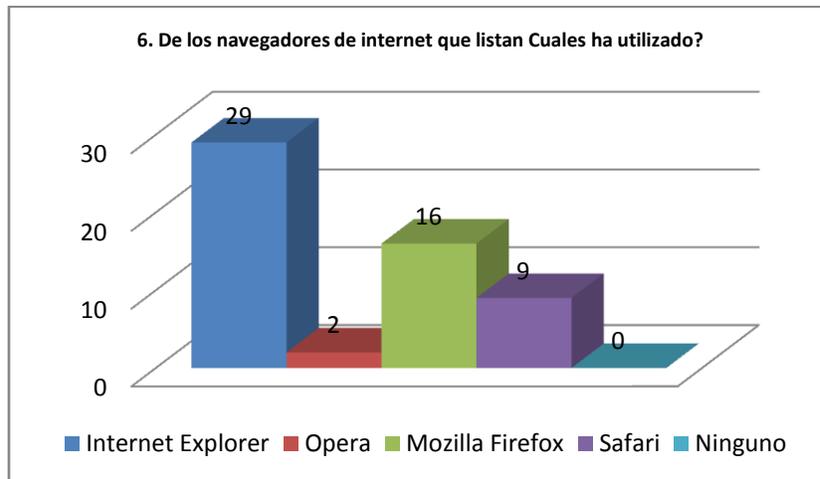
Desde los niveles de la alta gerencia hasta nivel operativo se considera que el resultado final de este proyecto mejoraría los procesos y la presentación de la información optimizando el tiempo de solución para las denuncias y teniendo acceso a la información detallada. Los niveles gerenciales y de estudios internos de la organización consideran que la herramienta ayudaría a mejorar el análisis de los procesos de las denuncias. (GRAFICO 4)



**Grafico 4: Importancia de la información consultada**

Ya que la propuesta de solución involucra como capa de presentación el uso de un navegador WEB que permita el enlace entre el usuario final y el “Data Warehouse” es necesario evaluar conocimiento de uso de navegadores WEB así como herramientas ofimáticas más conocidas, esto con el propósito de considerar los resultados obtenidos al momento de capacitar al personal.

Dentro del conocimiento técnico operativo se logra identificar que el 100% de los usuarios ha utilizado o utiliza un navegador WEB, como índice de frecuencia domina el uso de Internet Explorer considerando que es el navegador del sistema operativo de mas uso a nivel comercial, sin embargo los usuarios han tenido la oportunidad de utilizar al menos 1 navegador mas. (Ver ANEXO 4: Encuesta de conocimiento técnico operativo, pregunta 6).



**Grafico 5: Uso de navegadores WEB**

Se considera que la aceptación y el apoyo de la Gerencia de Informática está presente en el desarrollo de dicho proyecto, sin embargo es necesario demostrar que dicha solución generará una reducción de costos y de tiempo a la hora de ser implementado. Se considera que echar a andar el proyecto en el presente genera un costo de \$42,287.98 sin embargo este valor se ve con un ahorro de \$929.80 si se lleva a valor presente el trabajo como se ha ido realizando hasta ahora en un periodo de 5 años. Teniendo una inversión inicial de \$34,790.1 para el desarrollo e implementación del proyecto en el presente, generaría ahorro de hasta un 21% en los costos anuales respecto a cómo se trabaja actualmente. Es decir que la propuesta del “DESARROLLO DE UN DATA WAREHOUSE PARA EL PROCESO DE LAS CONSULTAS DE LA DEFENSORÍA DEL CONSUMIDOR”, es Justificable desde el punto de vista económico como social, ya que no solo beneficiará económicamente a la institución sino que agilizará los procesos para servir mejor a la población en general.

## **IMPORTANCIA**

La realización del proyecto “DESARROLLO DE UN DATA WAREHOUSE AL PROCESO DE DENUNCIAS PARA LA DEFENSORÍA DEL CONSUMIDOR”, con lleva a la creación de un sistema capaz de brindar un beneficio significativo para la Defensoría del Consumidor en general.

Los siguientes elementos resaltan la importancia que conlleva el realizar el proyecto:

- Un mejor análisis interno de la información al poseer un acceso rápido y fácil a los contenidos transaccionales.
- Creación de reportes a la medida de acuerdo a las necesidades del usuario final, contemplando en todo momento el grado de libertad para acceder a la información, en base al perfil del usuario.
- Todos los procesos de solicitud, adquisición, búsqueda, administración y confirmación de información estarán automatizados, de tal forma que la realización de este proyecto y finalizarlo exitosamente, abrirá camino a la creación de nuevos proyectos en beneficio de la población en general.
- El beneficio en particular para la Gerencia de Sistemas de Información se verá reflejado en la obtención de horas/hombre disponible para su utilización.
- Crear una cultura en el personal de la Defensoría del Consumidor enfocada a la utilización de herramientas informáticas para la agilización de los procesos operativos internos, lo que aportara un valor agregado al desarrollo profesional del recurso humano en general.

Por lo antes expuesto podemos asegurar que la trascendencia que tendrá el desarrollo y finalización del proyecto, tendrá efectos y consecuencias directamente sensibles para la institución en cuanto a la agilización de los procesos internos mediante la automatización de la gestión de información y para la sociedad a través del mejoramiento del servicio de solución de controversias y la generación de nuevos proyectos a mediano y largo plazo.

# Capítulo I



## 1.1. LISTADO DE ACRÓNIMOS

**BI:** Business Intelligence (Inteligencia de Negocios)

**CSC:** Centro de Solución de Controversias

**DBMS:** DataBase Management System (Sistemas de Gestión de Base de Datos)

**DE:** Direcciones Estudios Planeamiento y Desarrollo

**DF:** Defensoría del Consumidor

**DO:** Direcciones Organizacionales y Presidencia

**DW:** Data Warehouse (Almacén de datos)

**ER:** Diagrama o modelo Entidad-Relación

**ETL:** Extract, transform and load (Extraer, transformar y cargar)

**ESPACIO:** sistema brinda apoyo al Sistema Nacional de Protección al Consumidor (SNPC), permitiendo a cada institución que forma parte del sistema, obtener las herramientas para el registro y seguimiento de denuncias de los consumidores en las áreas que les compete

**JDBC:** Java Database Connectivity (estándar de Java para conectarse con Bases de Datos)

**LDAP:** Lightweight Directory Access Protocol (Protocolo Ligero de Acceso a Directorios)

**MOLAP:** Multidimensional Online Analytical Processing (Procesamiento analítico multidimensional en línea)

**OLAP:** On-Line Analytical Processing (Procesamiento Analítico en Línea)

**OLTP:** OnLine Transaction Processing (Procesamiento de Transacciones En Línea)

**SARA:** Sistema de Atención a Reclamos y Denuncias

**SSIS:** SQL Integration Services

**UML:** Lenguaje Unificado de Modelado (Unified Modeling Language)

## 1.2. MARCO TEÓRICO

### 1.2.1. Business Intelligence

#### 1.2.1.1. ¿Inteligencia de negocios?

Es la combinación de tecnología, herramientas y procesos que me permiten transformar mis datos almacenados en información, esta información en conocimiento y este conocimiento dirigido a un plan o una estrategia comercial. La inteligencia de negocios debe que ser parte de la estrategia empresarial, esta le permite optimizar la utilización de recursos, monitorear el cumplimiento de los objetivos de la empresa y la capacidad de tomar buenas decisiones para así obtener mejores resultados”<sup>2</sup>.

En 1958, en un artículo realizado por el investigador de IBM, Hans Peter Luhn, utilizo el término “business intelligence”. Definió la inteligencia como: “la habilidad de comprender la interrelación de factores presentes, de una manera para encaminar acciones para alcanzar metas deseadas”<sup>3</sup>.

En 1989 Howard Dresner (posteriormente analista de Gartner Group) propuso “business intelligence” como una sombría de términos para describir; “conceptos y métodos para mejorar la toma de decisión en negocios mediante la utilización de factores basados en sistemas de soporte”<sup>4</sup>.

Durante el transcurso de la historia ha existido el concepto de inteligencia de negocio, pero es hasta hoy en día que ha comenzado el auge de este tema. La inteligencia de negocio brinda apoyo a los diferentes niveles organizacionales, siendo estos niveles los mostrados en la figura 1.1.

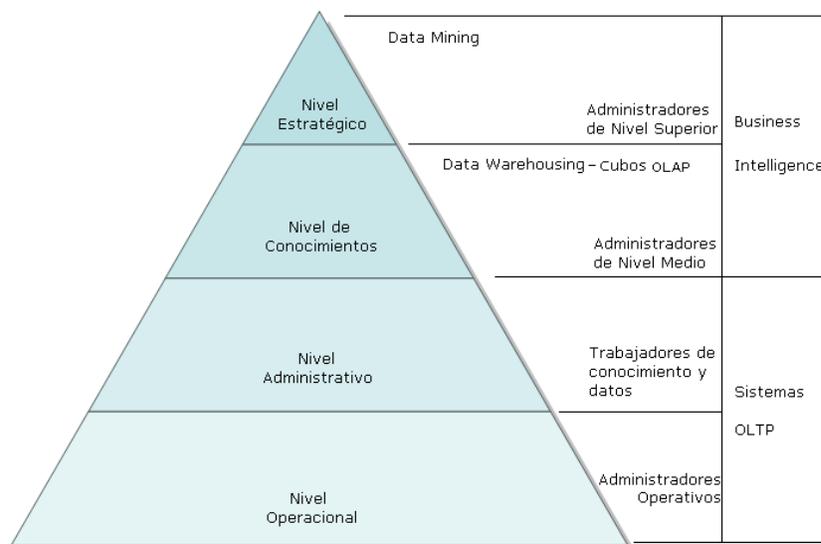


Figura 1.1: Niveles organizacionales

- **Nivel operacional:** Se utilizan sistemas de información que monitorean las actividades y transacciones elementales de la organización. Son sistemas que han cobrado un auge

<sup>2</sup> ORACLE, [http://www.oracle.com/technology/global/lad-es/documentation/bi\\_doc.html#bi](http://www.oracle.com/technology/global/lad-es/documentation/bi_doc.html#bi)

<sup>3</sup> IBM, <http://www.research.ibm.com/journal/rd/024/ibmrd0204H.pdf>

<sup>4</sup> DSSResources, <http://dssresources.com/history/dsshistory.html>

importante en la última década a consecuencia de un desarrollo organizacional orientado al mercado global.

- **Nivel de conocimiento:** En este nivel se encuentran los trabajadores de conocimiento y datos, cubriendo el núcleo de operaciones tradicionales de captura masiva de datos y servicios básicos de tratamiento de datos con tareas predefinidas
- **Nivel de administración:** Se realizan tareas de administradores de nivel intermedio apoyando las actividades de análisis, de seguimiento, de control y toma de decisiones, realizando consultas sobre información almacenada en el sistema, proporcionando informes y facilitando la gestión de la información por parte de los niveles intermedios.
- **Nivel estratégico:** Tiene como objetivo realizar las actividades de planificación a largo plazo, tanto del nivel de administración como de los objetivos que la empresa posee.

### 1.2.1.2. ¿Cuáles son los beneficios de inteligencia de negocios?

Dentro del marco de beneficio que presenta una solución de inteligencia de negocios podemos mencionar que esta nos permite:

- **Manejar el crecimiento.** El reto para las empresas es evolucionar, es crecer y esto significa “cambio”. Conocer que tan ágiles son sus procesos para enfrentar los cambios y necesidades puntuales de la empresa.
- **Control de costos.** El manejo de costos es lo que fuerza en la mayoría de casos a considerar una solución de inteligencia de negocios, para tener la capacidad de medir gastos y ver esto a un nivel de detalle que identifique la línea de negocio, producto, centro de costos, etc.
- **Entender mejor a los clientes.** Las empresas almacenan alto volumen de información y variada, relacionada a sus clientes. El reto es transformar esta información en conocimiento y este conocimiento dirigido a una gestión comercial que represente algún tipo de ganancia para la empresa.
- **Indicadores de gestión.** Los indicadores de desempeño me permiten representar medidas enfocadas al desempeño organizacional con la capacidad de representar la estrategia organización en objetivos, métricas, iniciativas y tareas dirigidas a un grupo de individuos en la organización.

Bajo el nombre de Business Intelligence (Inteligencia de Negocios) se agrupan diferentes acrónimos, herramientas y disciplinas que apuntan a dar soporte a la tarea de toma de decisiones.

Fundamentalmente podemos mencionar:

- **Data warehousing:** Su traducción en castellano “almacenes de datos”. Se basan en estructuras multidimensionales en las que se almacena la información calculando previamente todas las combinaciones de todos los niveles de todas las aperturas de análisis.
- **Data mart:** Constituye una parte de un data warehouse. Si un data warehouse está formado por todos los procesos de la organización, un Data mart constituye un determinado proceso. Por ejemplo podríamos tener un Data mart para finanzas, otro para logística. Pueden ser elaborados independiente o no de un data warehouse.
- **Data mining:** Asociado al nivel organizacional estratégico y tiene por objetivo eliminar los errores cometidos por las personas al analizar los datos debido a prejuicios y dejar que sean los datos que demuestren hechos reales.
- **Tecnología OLAP (online analytical process):** Es la tecnología que permite aprovechar como está estructurada la información de un Data mart o un Data warehouse. Fundamentalmente es una tecnología que permite analizar información dinámicamente a los

niveles tácticos y estratégicos basados en Cubos que contienen las medidas y las dimensiones.

### 1.2.2. Sistemas Transaccionales (OLTP)

Los sistemas de OLTP (On-Line Transaction Processing) son los sistemas operacionales que capturan las transacciones de un negocio y las almacenan en estructuras relacionales (bases de datos). Las características principales de los sistemas OLTP son:

- Realiza transacciones en tiempo real del proceso de un negocio, con lo cual los datos almacenados cambian continuamente.
- Los sistemas OLTP son responsables del mantenimiento de los datos (adición, actualización o eliminación).
- Las bases de datos deben estar optimizadas para validar la entrada de datos y rechazarlos si no cumplen con determinadas reglas de negocio.
- Para tomar decisiones proporcionan capacidad limitada, ya que no es su objetivo, por lo tanto no es prioridad en su diseño.
- La consulta de información histórica relativa al negocio en un sistema OLTP produce un impacto negativo en el funcionamiento del sistema.

Las organizaciones se ven entonces en la necesidad de registrar las transacciones que ocurren durante sus procesos operacionales, para su control y posterior consulta. Un sistema OLTP es utilizado en diferentes áreas (ver Figura 1.2):

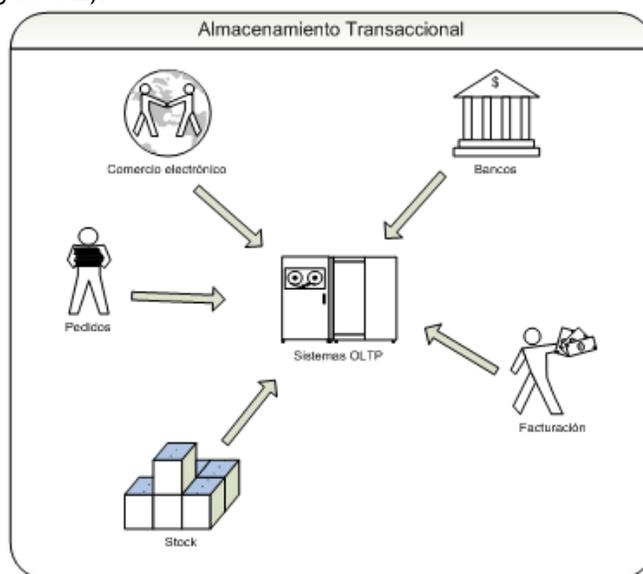


Figura 1.2: uso de los OLTP

Usos comunes de un sistema OLTP:

- Sistemas bancarios.
- Procesamiento de pedidos.
- Comercio electrónico.
- Sistemas de facturación.
- Sistemas de stock.

El modelado de estos sistemas es producido usualmente a través de un proceso de modelado entidad – relación (ER). Un diagrama entidad relación (DER) es representado a través de:

- Entidades.

- Atributos.
- Relaciones.

Este proceso de modelado para los sistemas OLTP encamina a una estructura normalizada en el que cada entidad posee su propia tabla por separado, relacionada con otras. La normalización de modelo de datos encaja muy bien para los requerimientos de los sistemas OLTP, ya que asegura que la información sea almacenada en un solo lugar, la cual puede ser actualizada fácil y eficientemente.

### 1.2.3. Sistemas OLAP

#### 1.2.3.1. ¿Qué es OLAP?

OLAP es la respuesta más próxima al análisis de consultas de naturaleza dimensional. Es una parte de la amplia categoría de inteligencia de negocios, que también incluye ETL, reportes relacionales y minería de datos. El concepto de OLAP también puede ser descrito como el análisis rápido de información multidimensional compartida (FASMI). Posee aspectos de bases de datos relacionales en navegación y jerarquía, optimizan el procesamiento más que un sistema de tipo relacional. Las Bases de datos OLAP emplean modelos de datos multidimensionales, brindando la realización de análisis complejos y ad hoc queries (consultas con fines específicos) con una ejecución rápida en tiempo. Características a resaltar de un sistema OLAP:

- **Rapidez:** Significa que los sistemas están enfocados a brindar respuesta a usuarios entre aproximadamente 5 segundos, con análisis simples no más de 1 segundo y muy pocos más de 20 segundos.
- **Análisis:** El sistema puede hacer frente a cualquier lógica de negocios y análisis estático que es relevante para la aplicación y los usuarios, manteniendo la simpleza suficiente para los usuarios finales.
- **Compartido:** Se debe implementar toda la seguridad requerida para proveer confidencialidad y si se necesita múltiple acceso para escritura y actualización concurrente es necesario establecer la seguridad a un apropiado nivel.
- **Multidimensional:** El sistema debe proveer una vista conceptual multidimensional de los datos, incluyendo soporte completo para jerarquización y múltiple jerarquización, ya que es la manera más lógica de analizar el negocio y la organización.
- **Información:** Incluye todos los datos e información derivada, de donde sea y cuan relevante sea para la aplicación.

#### 1.2.3.2. Tipos de OLAP

Existen 3 tipos de OLAP. Cada tipo posee ciertos beneficios y así mismo desventajas.

- **OLAP Multidimensional:** MOLAP es la forma clásica de OLAP y es a veces referido solo como OLAP. MOLAP utiliza bases de datos estructuradas que son generalmente optimizadas para recuperación de datos. A diferencia de la base de datos relacional, esta forma de almacenamiento esta optimizada para la aceleración de cálculos. Frecuentemente son optimizadas para la recuperación de patrones a lo largo del acceso jerarquizado. La dimensión de cada cubo son típicamente atributos como periodos de tiempo, ubicación y productos o código de cuenta. MOLAP trabaja mucho mejor en sistemas de datos pequeños, pues es más fácil calcular las agregaciones y retornar respuesta utilizando menos espacio de almacenamiento.

- **OLAP Relacional:** ROLAP trabaja directamente con bases de datos relacionales. Sin embargo, en organizaciones con procesamiento de un alto volumen de datos es difícil de implementar, frecuentemente este tipo de OLAP es ignorado.
- **OLAP Híbrido:** No existe un acuerdo formal sobre lo que constituye un OLAP híbrido, exceptuando que la base de datos divide los datos entre el relacional y almacenamiento especializado. HOLAP se encuentra entre los otros dos tipos de OLAP, mas sin embargo puede realizar procesamientos rápidos y escalables.
- Para el proyecto tratado, este se ubica dentro del tipo de MOLAP, siendo planteado como un Data Warehouse el cual es diseñado con estructuras dedicadas a la recuperación de datos.

### 1.2.3.3. Usos comunes de OLAP

Los sistemas OLAP son utilizados por las empresas para conocer la historia del negocio y poder realizar la toma de decisiones. Entre las áreas en que están más difundidos los sistemas OLAP son:

- **Sistemas de información a ejecutivos:** Los usuarios y los administradores generalmente de mandos altos y medios, reciben la información sobre los indicadores de funcionamiento dominantes del negocio y de las excepciones o las variaciones según sea de patrones y de estándares preestablecidos. Los sistemas de información a ejecutivos presentan típicamente datos multidimensionales en formatos gráficos.
- **Aplicaciones financieras:** Para diversos usos de tipo financiero se utilizan las bases de datos OLAP; para comunicar, planear y analizar. Los ejemplos de uso financiero incluyen la comunicación, análisis del mes de cierre, análisis de los beneficios del producto, los presupuestos y pronóstico.
- **Ventas y aplicaciones de marketing:** Existen diferentes formas de llegar a los clientes para alcanzar los objetivos de venta y de comercialización propuestos. Por esto es aconsejable la utilización de sistemas OLAP donde es importante contar con información organizada de manera rápida.

### 1.2.4. Data Warehouse

Un data warehouse o almacén de datos, es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se trata, sobre todo, de un expediente completo de una organización, más allá de la información transaccional y operacional, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de datos (especialmente OLAP, procesamiento analítico en línea).

Un data warehouse puede ser considerado como una copia de la información transaccional estructurada específicamente para realizar consultas y reportes.

Existen diferencias bien marcadas entre un data warehouse y una base de datos transaccional, entre las cuales se puede mencionar:

- Una base de datos transaccional almacena las actividades diarias de una organización, mientras que un data warehouse ayuda a la planificación de la organización.
- La base de datos transaccional se enfoca en el detalle, al contrario de el data warehouse que se enfoca en agregados de alto nivel.
- Un data warehouse integra datos de diferentes orígenes, en cambio, una base de datos transaccional es diseñada específicamente para una aplicación.
- La información almacenada en una base de datos es volátil, es decir, su información cambia constantemente con las transacciones diarias de la organización. Un data warehouse es no

volátil, los datos que son almacenados no sufren ninguna actualización solo son incrementados.

- Una base transaccional debe proporcionar una rápida extracción y actualización de información detallada. El data warehouse debe proporcionar la extracción rápida de información altamente resumida para la toma de decisiones.

#### 1.2.4.1. Métodos de implementación de un Data Warehouse

Existen muchas formas de construir e implementar un data warehouse, pero existen ciertos métodos que son los principales.

- **Top-Down:** este método utiliza un solo data warehouse centralizado, el cual contiene los datos resumidos y detallados en pequeños data marts que obtienen todos sus datos del data warehouse.
- **Bottom-Up:** El fin que persigue este método es entregar un valor agregado al negocio al implementar data marts dimensionales tan rápido como sea posible. Este método es flexible y amigable con el usuario.
- **Federado:** Este método no cuenta con una arquitectura definida, pero es una teoría que permite integrar los orígenes de datos, por los medios que sean necesarios, para resolver las necesidades y responder a condiciones dinámicas. Lleva a las organizaciones a compartir las métricas de “valor agregado”, dimensiones y medidas cuando y donde sea posible.

Luego de conocer las diferencias existentes entre una base de datos transaccional y un data warehouse, surge la pregunta ¿Cómo se traslada la información de la base de datos transaccional al data warehouse? La respuesta es simple, utilizando **ETL (Extract-Transform-Load)**.

#### 1.2.4.2. ETL (Extract, Transform and Load)

Extraer, transformar y cargar (ETL por sus siglas en inglés), es un proceso en data warehouse que se encarga de extraer la información de las fuentes de datos externas, transformarlas para que se adapten a las necesidades del negocio y por último cargarla en el data warehouse.

- **Extraer:** Es la primera etapa del proceso ETL, donde la información es extraída de todos los sistemas que almacenan información, la cual puede encontrarse en una base de datos relacional, hojas de Excel, archivos planos, etc. Cada uno de estos sistemas puede tener diferentes tipos de formatos de información, por lo que la etapa de transformación se vuelve crucial en este proceso.
- **Transformar:** La etapa de transformación aplica una serie de reglas o funciones a los datos extraídos para que la información pueda ser cargada. Ciertos orígenes de datos requieren poca manipulación de los datos, aunque, en otros casos, cualquier combinación de los siguientes tipos de transformación puede ser necesario:

Seleccionar únicamente las columnas a cargar. Como ejemplo, se puede decidir no cargar las columnas que tengan valores Nulos.

Traducir los códigos de los datos. Por ejemplo el sexo de una persona se puede almacenar en un sistema con las letras 'M' y 'F', y en otro se puede encontrar con '1' y '2'. Es necesario llevar estos códigos a un solo estándar en el cual serán almacenados en el data warehouse.

Establecer los campos calculados. Existen campos de los cuales solo se necesita el valor resultante de su cálculo, por ejemplo (cantidad de productos \* precio unitario).

Determinar las unidades de medidas de los atributos. Las unidades pueden tener distintas unidades de medidas, según el origen del sistema transaccional. Un ejemplo es hablar de litro, centímetros cúbicos o hectolitros.

Formatos. Otro ejemplo son los formatos de fecha que se encuentra en los diferentes sistemas operacionales. Las fechas pueden estar almacenadas como yyyy/mm/dd ó dd/mm/yyyy.

- **Cargar:** En esta fase los datos previamente transformados son cargados en el data warehouse.

### 1.2.5. Data Marts

Un data mart es una versión especializada de un data warehouse. Al igual que los data warehouse, los data marts contienen una copia de los datos transaccionales que ayudan a la gente de negocios a establecer estrategias basados en el análisis de amenazas y experiencias pasadas.

La diferencia entre estos dos componentes radica en que el data mart está enfocado a una necesidad específica y predefinida para cierto tipo de información. Existen dos tipos de Data Marts:

- **Dependientes:** Los datos son extraídos del data warehouse.
- **Independientes:** Los datos son extraídos de los sistemas transaccionales.

### 1.2.6. Dashboards

El nombre Dashboard se refiere al tablero de un automóvil, el cual ofrece al conductor información permanente sobre el estado del vehículo. El mundo de los negocios toma la palabra con un sentido similar pero en lugar de aplicarlo a los automóviles lo refiere a la empresa.

Así Dashboard es una página desarrollada en base a tecnología web mediante la cual se despliega en tiempo real información de la empresa extraída de varias fuentes o bases de datos. Su característica de tiempo real otorga a los usuarios un conocimiento completo sobre la marcha de la empresa y permite hacer análisis instantáneos e inteligencia de negocios.

### 1.2.7. Introducción a Metodologías de Business Intelligence<sup>5</sup>

En nuestros días, muchas organizaciones necesitan crear un data warehouse que le brinde soporte a la toma de decisiones. Estas organizaciones cuentan con una gran variedad productos que pueden utilizar, tanto en software como en herramientas de desarrollo. Para lograr el éxito se requiere el conocimiento de los dos modelos principales de data warehouse: el modelo de Inmon y el de Kimball.

Bill Inmon defiende un desarrollo top-down, el cual adapta las herramientas tradicionales de una base de datos relacional con las necesidades de desarrollo de un data warehouse empresarial. Desde el data warehouse son desarrolladas las bases de datos departamentales para poder cubrir con las necesidades de la toma de decisiones.

Por el otro lado Ralph Kimball, sugiere un desarrollo bottom-up utilizando un modelado dimensional, el cual es único para los data warehouse. En lugar de construir una sola base de datos empresarial, Kimball sugiere crear una base de datos (o data mart) por proceso empresarial. La cohesión empresarial es lograda utilizando una innovación de Kimball, un bus de datos estándar.

---

<sup>5</sup> TWDI, <http://www.tdwi.org/Publications/display.aspx?Id=6991>,

### 1.2.7.1. El modelo de Inmon

La arquitectura de Inmon consiste en todos los sistemas de información y sus respectivas bases de datos de una determinada organización. Él llama a este gigante “Fabrica de Información Organizacional”, GIF por sus siglas en inglés (Corporate Information Factory).

Inmon divide el ambiente de bases de datos organizacionales en cuatro niveles:

- Operacional.
- Data warehouse atómico.
- Departamental.
- Individual.

Los últimos tres niveles conforman el data warehouse. El primer nivel contiene datos de sistemas heredados y otros sistemas transaccionales. Este nivel es el que da soporte a las operaciones diarias de la organización, en otras palabras, el primer nivel soporta todo el procesamiento transaccional. En los sistemas transaccionales, la información es manipulada considerablemente y luego es movida al data warehouse atómico.

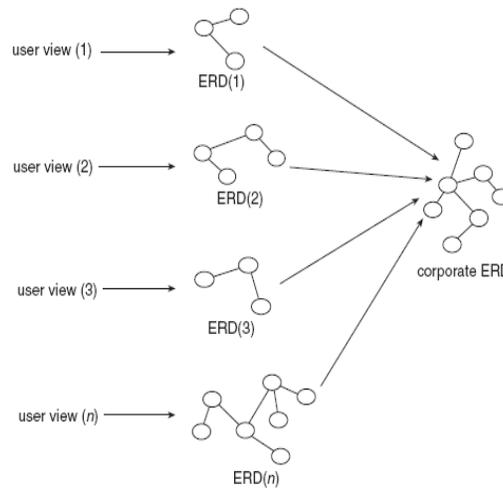
Inmon utiliza un ejemplo para mostrar la diferencia entre los datos operacionales y los datos almacenados en el data warehouse atómico. En el ejemplo, la entidad es un cliente y sus atributos de mas interés de un cliente es su record crediticio. La base de datos operacional contiene el actual record crediticio e información relacionada a este (como balances de préstamos, direcciones, etc) en un solo registro. En contraste, el data warehouse atómico, contiene el histórico del record crediticio para este cliente, totalizado por año, con un registro por año.

Los usuarios individuales crean el cuarto y último nivel de la arquitectura cuando crean sets de datos heurísticos y ad hoc como parte de soporte a la toma de decisiones. Este nivel tiende a ser temporal y a ser almacenado en la computadora personal del usuario. Por ejemplo, un usuario que trabaja en el departamento de créditos puede pedir un reporte para visualizar todas las cuentas que han sido robadas en los últimos tres años.

Si la base de datos del departamento no cuenta con la información detallada histórica, es posible realizar la consulta al data warehouse atómico. Las consultas al data warehouse atómico generalmente se realizan a través del departamento de TI. Inmon argumenta que es compensado el esfuerzo inicial requerido para construir el data warehouse atómico, ya que permite la creación de múltiples bases de datos departamentales sin correr el riesgo de generar datos inconsistentes entre ellos. Esto se logra utilizando un modelo de datos de tres niveles.

#### **Modelo de Datos de Tres Niveles**

Inmon propone el modelado de datos de tres niveles. El primero lo conforman los DER (Diagrama Entidad Relación). Al igual que en el desarrollo de bases de datos operacionales, el DER es utilizado para determinar y refinar las entidades, sus atributos y relaciones entre ellas. El equipo de desarrollo crea un conjunto de DERs por cada departamento que se pretende cubrir con el Data Warehouse. El DER corporativo es la suma de todos los DER departamentales, como se muestra en la figura 1.3.

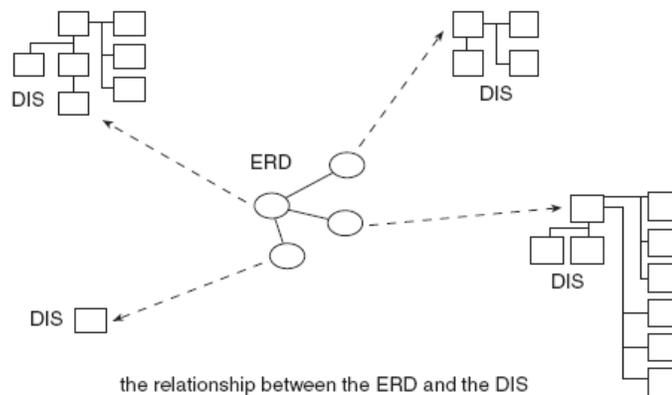


**Figura 1.3: Representación gráfica del primer nivel del modelo de datos desde las diferentes vistas de usuario según Inmon**

El segundo nivel, establece una serie de datos (DIS, data ítem set) para cada departamento. Al igual que el primer nivel, la suma de los DIS departamentales conforma el DIS empresarial. Este nivel cuenta con cuatro construcciones:

- Una agrupación de datos primaria.
- Una agrupación de datos secundaria.
- Un conector, que representa las relaciones entre los principales datos de cada área grande.
- El tipo de datos.

Un aspecto crítico de este nivel es que la agrupación de datos primaria existe una sola vez para cada área grande. Esto significa que un DER creado en el modelo de datos del primer nivel es la base para un DIS en el segundo nivel. La figura 1.4 ilustra la relación DER-DIS para una determinada vista de un usuario. También muestra como las diferentes vistas de los usuarios son combinadas en un DER y DIS organizacional. En el DIS cada rectángulo representa una tabla lógica ya sea en el DIS departamental o en el organizacional. Las relaciones entre estas tablas son las mismas que relacionan las entidades en los DER. Los rectángulos a la derecha de un determinado DIS representan la agrupación de datos secundaria.



**Figura 1.4: Cada entidad del DER es definida por su propio DIS**

Para aclarar esto Inmon utiliza un ejemplo. En un banco, la entidad “cliente” genera una agrupación de datos primaria con la cuenta. La “cuenta” puede tener diferentes formas, como préstamos y ahorros

(grupo secundario). Las relaciones muestran que un cliente puede tener diferentes cuentas. Por último, cada cuenta puede tener datos generados por actividades similares, como depósitos en ATM, depósitos bancarios, retiros en ATM y agencias bancarias. Estos son ejemplos del tipo de datos.

El último nivel del modelo de Inmon es el físico. “El modelo físico es creado desde el segundo nivel del modelo de datos simplemente extendiéndolo, incluyendo las llaves y características físicas del modelo. En este punto, el nivel físico del modelo de datos se asemeja a una serie de tablas, muchas veces llamadas tablas relacionadas”<sup>6</sup>. Inmon explica varias técnicas para optimizar el rendimiento del data warehouse en los niveles departamentales y atómicos.

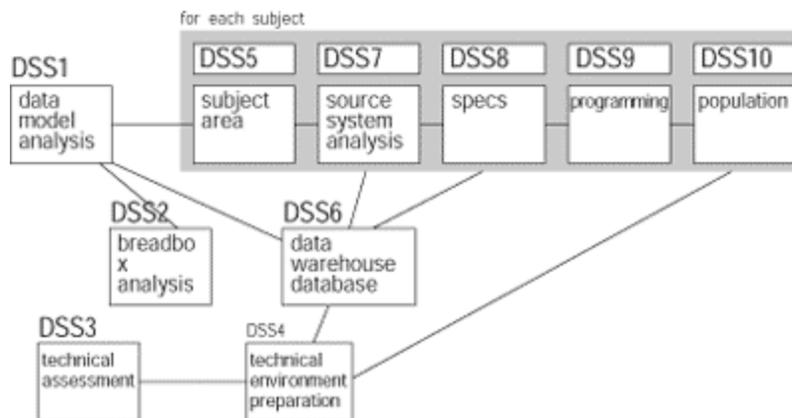
Aunque las técnicas no son muy familiares su propósito es optimizar el rendimiento de E/S al igual que en los sistemas de base de datos transaccionales. La mayoría de las técnicas involucran la desnormalización de las tablas.

Existen muchas razones por las cuales se deben de desnormalizar las tablas en el nivel físico. Por ejemplo, registros en el data warehouse atómico son rara vez actualizados puesto que son datos históricos. Esto hace posible colocar los datos físicamente en formas que no serían funcionales para la base de datos transaccional ya que su actualización es constante.

Una vez se ha completado el modelo de datos de tres niveles, el desarrollo del data warehouse comienza.

**Propuesta de Desarrollo en Espiral (Meth2)**

Un modelo de datos de tres niveles completo es el único prerequisite para utilizar la adaptación especial de la metodología de desarrollo en espiral creada por Inmon, a la cual llama *Meth2* (Meth1 es para el desarrollo de sistemas transaccionales; Meth3 es para mejorar el rendimiento de un data warehouse existente). Inmon llama el proceso de modelado DSSI (del inglés Decision Support 1, que significa Soporte a la toma de Decisiones 1). Este proceso tiene nueve pasos más como se muestra en la figura 1.5.



**Figura 1.5: Proceso de modelado DSSI**

Utilizando el modelo de datos de tres niveles como la entrada del proceso, el siguiente paso es realizar un análisis de granularidad (DSS2 en la Figura 1.5). La granularidad es la forma de medir el nivel de detalle de los datos. Por ejemplo, los datos transaccionales tienen el nivel más bajo de granularidad ya que

<sup>6</sup> Building the Data Warehouse, Third Edition, W.h. Inmon, Wiley Computer Publishing Pag98.

se necesita un gran nivel de detalle. Si el volumen de datos que se va a manejar es grande, entonces el equipo de desarrollo debe de considerar múltiples niveles de granularidad.

Una vez se han resuelto los problemas con la granularidad, se debe seleccionar el primer tema (DSS5). Esto se convertirá en la primera base de datos departamental. El equipo analiza los sistemas de origen del primer tema (DSS7), escribe las especificaciones (DSS8), codifica los programas (DSS9) y llena la base de datos (DSS10). El diseño del data warehouse atómico comienza concurrentemente (DSS6). Cuando ya se tiene suficiente información, el equipo comienza la evaluación técnica (DSS3). Esta evaluación asegura que la información del data warehouse sea accesible y manejable.

A medida que el equipo desarrolla exitosamente cada base de datos departamental, se da un impacto en el data warehouse atómico. La figura 1.5 muestra la iteración del modelo por medio de las líneas que conectan varios pasos. Existen líneas que conectan los sistemas de origen (DSS7) y las especificaciones (DSS8) con el diseño del data warehouse atómico (DSS6), indicando que el diseño del data warehouse será considerado nuevamente cada vez que se desarrolla una base de datos departamental. La línea que conecta el llenado de la base de datos (DSS10) con la preparación del ambiente técnico (DSS4) también muestra la naturaleza iterativa del Meth2. Cuando se habla de la preparación del ambiente técnico (DSS4), Inmon se refiere a asegurarse que la red del data warehouse, hardware de almacenamiento, sistema operativo, todas las interfaces de acceso y software involucrado estén listos para recibir los datos.

El manejo de los datos es un aspecto esencial de la metodología en espiral de Inmon. “Uno de los aspectos sobresalientes de una metodología basada en el manejo de los datos es que, se construye en esfuerzos previos utilizando el código y procesos que han sido desarrollados”. Esto quiere decir que el tiempo que tomaría en pasar del DSS1 al DSS2 en el desarrollo de una segunda base de datos departamental, sería menor al utilizado en el desarrollo de la primera.

### ***La filosofía de Inmon: Evolutivo, no Revolucionaria.***

Inmon ve el data warehouse como una parte integral de la Fabrica de Información Organizacional. Esto significa que el data warehouse y las bases de datos operacionales son parte de un todo. Esta percepción ayuda a explicar porque el data warehouse de Inmon se debe adherir a los estándares de los sistemas operacionales. Partiendo de esta premisa, es fácil entender como el método evolutivo de Inmon se adhiere a los métodos de desarrollo de las bases de datos y software transaccional. Los aspectos presentados como el modelo de datos de tres niveles y el desarrollo en espiral son consistentes con las prácticas utilizadas en el diseño y desarrollo de sistemas relacionales. Está basado en principios y prácticas que han sido utilizadas en el mundo de bases de datos transaccionales por lo menos una década antes de los primeros esfuerzos de realizar un data warehouse. Desde este punto de vista, el modelo de Inmon es más evolutivo que revolucionario.

### **1.2.7.2. EL modelo de Kimball<sup>7</sup>**

Difiere un poco respecto al enfoque tradicional de bases de datos relacional, posee su propia metodología de modelado a la hora de crear las bases de datos. Además tiene como objeto principal que su arquitectura general se integre a múltiples bases de datos para que sean interoperables. La interoperabilidad es una propiedad que puede predicarse de sistemas de naturaleza muy diferente, como

<sup>7</sup> <http://www.tdwi.org/Publications/display.aspx?id=6991>

pueden ser los sistemas informáticos (en cuyo caso se le suele denotar también como neutralidad tecnológica).

### ***El modelado de datos.***

También puede parecer extraño, pero el modelado de la información inicia con el diseño de las tablas en lugar del análisis entidad-relación. Las tablas son tablas de hecho o de dimensión, las tablas de hecho contienen las métricas, mientras que las tablas de dimensión contienen los atributos de las métricas en las tablas de hecho.

Las tablas de dimensión rompen la regla de normalización al crear grupos repetidos, sin embargo viola las reglas de normalización con el propósito de lograr un alto rendimiento en el almacén de datos, brindando acceso a los usuarios finales.

Las tablas de hecho poseen muchas filas pero pocas columnas para facilitar el uso y mejorar el rendimiento de las consultas. En cambio las tablas de dimensión poseen miles, millones de filas y a diferencia de las tablas de hecho pueden poseer miles de columnas o más, esto se debe a que posee los atributos de las tablas de hecho en formas altamente desnormalizadas.

### ***Cuatro Pasos del proceso para el diseño dimensional***

Kimball recomienda una metodología de desarrollo que es única para el almacenamiento de datos. Se trata en un enfoque ascendente, que en el caso de almacenes de datos significa construir el DataMart a la vez. Las cuatro fases del diseño dimensional son:

- Seleccionar el proceso de negocio
- Establecer la granularidad
- Elegir las dimensiones
- Identificar los hechos

Procesos de negocio: se refiere a la metodología que se utiliza para llevar a cabo el proceso integral del rubro de la empresa. Punto de Ventas (POS), ventas al por menos, inventarios. En orden de selección para el proceso de negocio se debe escoger aquel que genere mayor impacto. Debe responder a las necesidades más importantes del negocio y obtener los datos de manera rápida y sencilla.

La Granularidad: Se refiere al nivel de detalle que tendrá el almacén de datos. El menor nivel de granularidad se llama atómico y se refiere a que no puede subdividirse, el nivel atómico es el nivel deseado ya que los usuarios pueden agregar los datos que deseen. La elección de las dimensiones van acorde al nivel de atomicidad que se elija, entre mayor nivel de atomicidad mayor cantidad de detalles en las dimensiones que se tengan para cada hecho que se defina. El último paso es definir los hechos que se incluirán en las tablas del mismo nombre.

### ***Cuando Utilizar el modelo de Kimball<sup>8</sup>***

La metodología de Kimball de separar el mundo de BI en hechos y dimensiones es muy efectiva y lleva a una solución completa en un lapso de tiempo relativamente corto. Esta metodología es buena cuando se están realizando los primeros pasos en BI. Es perfecta cuando la complejidad del data warehouse no es muy grande y cuando la infraestructura de BI maneja datos de un número limitado de orígenes. Cuando el data warehouse se está volviendo complicado, será peligroso utilizar únicamente la metodología de Kimball. En estos casos se recomienda utilizar la metodología propuesta por Inmon.

---

<sup>8</sup> Introduction to the sqlbi methodology, Draft 1.0, Alberto Ferrari, 20 Septiembre 2008

### 1.3. ESTÁNDARES Y DESCRIPCIÓN DE NOMENCLATURA

#### 1.3.1. Análisis de la situación actual

Con el fin de obtener una visión amplia de la solución propuesta, a continuación se muestra en la Tabla 1.1. La nomenclatura utilizada para la representación de los componentes en los diagramas utilizados en la representación grafica de la situación actual y la solución propuesta.

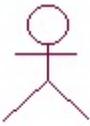
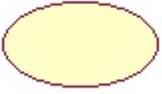
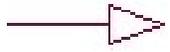
**Nomenclatura:**

Descripción	Figura
Representación de las direcciones establecidas dentro del organigrama de la defensoría del consumidor.	
Representación de técnico perteneciente a la Dirección de Sistemas de Información.	
Entrega o confirmación de un requerimiento solicitado de una dirección a otra dirección organizacional.	
Solicitud de un requerimiento de una dirección a otra dirección organizacional.	
Representación de un punto de terminal de trabajo.	
Representación de un servidor de bases de datos.	
Representación de un servidor de aplicaciones.	

**Tabla 1.1: Nomenclatura situación actual y solución propuesta**

### 1.3.2. Nomenclatura de casos de uso<sup>9</sup>

En la Tabla 1.2 se muestra la nomenclatura para definir los casos de uso.

Figura	Descripción
	<p><b>Actor:</b> es un rol que un usuario juega con respecto al sistema. Es importante destacar el uso de la palabra rol, pues con esto se especifica que un Actor no necesariamente representa a una persona en particular, sino más bien la labor que realiza frente al sistema.</p>
	<p><b>Caso de Uso:</b> Es una operación/tarea específica que se realiza tras una orden de algún agente externo, sea desde una petición de un actor o bien desde la invocación desde otro caso de uso.</p>
	<p><b>Asociación:</b> Es el tipo de relación más básica que indica la invocación desde un actor o caso de uso a otra operación (caso de uso). Dicha relación se denota con una flecha simple.</p>
	<p><b>Dependencia o Instanciación:</b> Es una forma muy particular de relación entre clases, en la cual una clase depende de otra, es decir, se instancia (se crea). Dicha relación se denota con una flecha punteada.</p>
	<p><b>Generalización:</b> Este tipo de relación es uno de los más utilizados, cumple una doble función dependiendo de su estereotipo, que puede ser de Uso (&lt;&lt;uses&gt;&gt;) o de Herencia (&lt;&lt;extends&gt;&gt;).</p> <p>Este tipo de relación está orientado exclusivamente para casos de uso (y no para actores).</p> <p><i>Extends:</i> Se recomienda utilizar cuando un caso de uso es similar a otro (características).</p> <p><i>Uses:</i> Se recomienda utilizar cuando se tiene un conjunto de características que son similares en más de un caso de uso y no se desea mantener copiada la descripción de la característica.</p>

**Tabla 1.2: Nomenclatura para definición de casos de uso**

<sup>9</sup> <http://www.dcc.uchile.cl/~psalinas/uml/casosuso.html>

### 1.3.3. Descripción de Estereotipos utilizados para el Diseño del Data Warehouse

Los estereotipos son presentados dependiendo de la clase base en la cual se especializa: Paquete, Clase y Clase de Asociación.

#### Estereotipos de Paquetes

Se han definido tres tipos de estereotipos del modelo de Paquetes, los cuales se detallan en la Tabla 1.3.

Nombre	Descripción	Representación
Paquete estrella (StarPackage)	Los paquetes de este estereotipo representan el esquema estrella en una base de datos multidimensional.	
Paquete dimensión (DimensionPackage)	Los paquetes de este estereotipo representan las dimensiones en una base de datos multidimensional.	
Paquete hechos (FactPackage)	Los paquetes de este estereotipo representan los hechos en una base de datos multidimensional.	

Tabla 1.3: Especificación de los estereotipos de paquetes

#### Estereotipos de Clases

Se han definido cuatro tipos de estereotipos partiendo del modelo de Clases, los cuales son detallados en la Tabla 1.4.

Nombre	Descripción	Representación
Hecho (Fact)	Las clases de este estereotipo representan hechos en un modelo de base de datos multidimensional.	
Dimensión (Dimension)	Las clases de este estereotipo representan las dimensiones en una base de datos multidimensional.	
Base (Base)	Las clases de este estereotipo representan los niveles jerárquicos de las dimensiones de un modelo de base de datos multidimensional.	
Hecho degenerado (DegenerateFact)	Las clases de este estereotipo representan la degeneración de hechos en una base de datos multidimensional.	

Tabla 1.4: Especificación de los estereotipos de clases

#### Estereotipos de Asociación de Clases

Un solo tipo de estereotipo de atributo es definido partiendo del modelo de Clases de Asociación. El modelo se detalla en la Tabla 1.5.

Nombre	Descripción	Representación
Hechos degenerada (DegenerateFact)	Los atributos de este estereotipo representan Hechos degenerados.	<b>MD</b>

Tabla 1.5: Especificación de estereotipo de asociación de clases

**Estereotipos de Atributo**

Se han definido cinco tipos de estereotipos del modelo de Atributos, los cuales son detallados en la Tabla 1.6.

<b>Nombre</b>	<b>Descripción</b>	<b>Representación</b>
Dimensión degenerada (DegenerateDimension)	Los atributos de este estereotipo representan dimensiones degeneradas.	<b>DD</b>
Atributo de hecho (FactAttribute)	Los atributos de este estereotipo representan atributos de Hechos o de clases de Hechos Degenerados en un modelo de base de datos multidimensional.	<b>FA</b>
OID	Los atributos de este estereotipo representan atributos OID de clases bases en un modelo de base de datos multidimensional.	<b>OID</b>
Descriptor (Descriptor)	Los atributos de este estereotipo representan el descriptor de un atributo de clases bases en un modelo de base de datos multidimensional.	<b>D</b>
Atributo de Dimensión (DimensionAttribute)	Los atributos de este estereotipo representan los atributos de clases bases en un modelo de base de datos multidimensional.	<b>DA</b>

**Tabla 1.6: Especificación de los estereotipos de atributo**

### 1.3.4. Diseño del proceso ETL (Extract, Transform and Load)

Para el diseño de los procesos ETL hemos utilizado una extensión de UML la cual permite el modelado conceptual de los procesos ETL, dicha extensión proporciona los mecanismos necesarios para especificar las operaciones típicas de los procesos ETL, como la integración de distintas fuentes de datos, transformación de atributos, generación de claves substitutas, etc.

El proceso ETL se define combinando los distintos mecanismos que se detallan adelante. Se ha seleccionado un conjunto reducido pero potente de mecanismos ETL, con el fin de reducir la complejidad del modelado y facilitar el uso.

En la Tabla 1.7 se definen los mecanismos más comunes en el diseño de los procesos ETL, los cuales se relacionan entre sí por medio de dependencia de UML. Además a cada mecanismo se le puede añadir una nota de UML para explicar el funcionamiento y definir el mapeo entre los atributos en el origen y el destino.

Para este diseño se considera fundamental explicar tres tipos de mecanismo: JOIN, LOADER y SURROGATE, los cuales describiremos a continuación:

#### 1.3.4.1. Mecanismo JOIN

Este mecanismo es utilizado para unificar dos fuentes de datos relacionados con uno o varios atributos en común. Para el modelado es importante tener claro la implementación de join en el lenguaje sql, ya que se aplica bajo el mismo concepto, teniendo para el caso:

*Join (condición): los cuales incluyen solos los registros que cumplen la condición left join (condición), right join (condición) of full join (condición): son usados para incluir los registros inclusive si no cumplen la condición fijada, por lo que su sintaxis se implementa dejando al atributo a compara seguido de un signo de interrogación antes de igualar el valor a comparar ejemplo: atributo\_a\_comparar ?= valor*

#### 1.3.4.2. Mecanismo LOADER

Este mecanismo permite realizar la carga de datos de un origen en el proceso ETL hacia una dimensión o la tabla de hechos, este tipo de mecanismo se puede implementar de dos maneras: Realizando cargas libres de datos en las cuales no se realiza ninguna verificación de datos; y la otra que en la cual la carga de datos se realiza por medio de la ayuda de verificadores definiendo el atributo que debe de existir en común en los datos de origen y destino para realizar este proceso. En el loader se puede definir de cuatro tipos los cuales son mutuamente excluyentes y se mencionan a continuación:

*Append: el destino no debe de estar vacío antes de realizar la carga de datos nuevos, la data nueva es agregada y la que ya se encontraba en el destino es actualizada.*

*Delete: el destino debe de estar vacío antes de realizar la carga de datos.*

*Insert: los datos nuevos son agregados al destino, dejando los datos que ya se encontraban sin realizarles cambios, aunque estos hayan sufrido cambios.*

*Update: solo la data existente es actualizada, dejando los nuevos datos sin cargarse al destino.*

**1.3.4.3. Mecanismo Surrogate**

Este mecanismo permite definir dentro del Data Warehouse una llave primaria la cual permita identificar, ordenar y reemplazar las llaves primarias del origen de datos con una llave uniforme dentro del Data Warehouse.

**Simbología utilizada en el diseño del proceso ETL**

Mecanismo ETL	Descripción	Icono
Aggregation	Agrega los datos (SUM, AVG, MAX / MIN, COUNT, etc.) en base a algún criterio	
Conversion	Cambia los tipos de datos, el formato o calcula nuevos datos (atributos derivados) a partir de los datos existentes	A → B
Filter	Filtra los datos no deseados y verifica la calidad de los datos en base a restricciones	
Incorrect	Redirige los registros incorrectos o descartados a un destino separado para su posterior verificación; solo se puede usar con un filter, loader o wrapper	
Join	Une dos fuentes de datos entre sí a través de uno o varios atributos	
Loader	Carga los datos en el destino de un proceso ETL(en un hecho o dimensión del almacén de datos)	
Log	Controla y registra la actividad de otros mecanismos ETL, con el fin de auditar las transformaciones realizadas	
Merge	Integra los datos provenientes de dos o más fuentes de datos con atributos compatibles.	
Surrogate	Genera una clave sustituta única, que se emplea para reemplazar la clave empleada en las fuentes de datos.	123 →
Wrapper	Transforma una fuente de datos nativa en una fuente de datos basada en registros.	

**Tabla 1.7: Mecanismos ETL y su representación en UML**

### 1.3.5. Estándar de nomenclatura de tablas y campos del data warehouse

A continuación se detallan los estándares utilizados para nombrar campos y tablas de utilizados dentro del data Warehouse.

Prefijo	Significado	Ejemplo
Fact	Será el prefijo de las tablas de hechos y a continuación se detallara uno nombre que describa.	FactCaso
Dim	Prefijo utilizado para identificar las tablas de dimensión, luego de este prefijo se escribirá un nombre que describa el contenido.	DimError
Id	Prefijo utilizado para denominar las llaves primarias utilizadas dentro del Data Warehouse.	IdError
PK	Prefijo utilizado para hacer referencia al llave primaria de las tablas de la base de datos OLTP, y poder tener información del origen de los datos.	PkError
Prc	Prefijo utilizado para nombrar procedimientos almacenados en la base de datos	Prc_prueba
Fnc	Prefijo utilizado para nombrar funciones almacenadas en la base de datos.	Fnc_prueba

**Tabla 1.8: Estándares para nombrar la tablas, procedimientos, funciones y campos en la base de datos**

## 1.4. FORMATOS DE ESPECIFICACIÓN PARA EL ANÁLISIS DE REQUERIMIENTOS.

### 1.4.1. Formato de especificación de requerimientos de usuario

Notación de especificación de requerimientos de usuario.

1. Utilizar el formulario RQU-XX.
2. Especificarlos en lenguaje natural auxiliándose de cuadros y diagramas comprensibles para los involucrados en el sistema. No utilizar términos técnicos, excepto los del dominio.
3. Organizarlos en un máximo de tres niveles jerárquicos, numerándolos desde 01 en cada nivel; separando la numeración de cada nivel con un punto (01.01.01). El primer nivel debe referirse a una categoría de requerimiento. El segundo nivel puede referirse a una sub categoría de requerimiento o a un requerimiento específico. Siempre habrá al menos dos niveles. El tercer nivel sólo ocurre cuando el segundo nivel es una sub categoría; y se refiere a un requerimiento específico.
4. Si un requerimiento específico tiene varias especificaciones, identificarlas con un literal: a, b, c, etc.
5. Agregar en la columna especificación de cada requerimiento específico el fundamento o explicación de tal requerimiento.
6. Especificar la clase de requerimiento con una letra mayúscula, así:
  - F → funcional.
  - N → no funcional.
- Completar la especificación del tipo de requerimiento no funcional agregando una letra minúscula a la derecha de la letra N según el tipo de requerimiento (No, Ni, Nd, etc.), así:
  - o → operativo.
  - i → implantación.
  - d → desarrollo.
  - t → técnico.
- Especificar el tipo de requerimiento con una letra mayúscula así:
  - O: Obligatorio → *debe*
  - D: Deseable → *debería*

En la Tabla 1.9 se presenta la notación de la hoja de especificación de requerimientos de usuario.

TEMA		PaginaX de y	
Autor :		Fecha:	
Req. Usu.	Especificación	prioridad	Tipo
01			
01.01			
01.01.01			
02			
02.01			
02.02			

Tabla 1.9: Hoja de especificación de requerimientos de usuario (RQU-01)

**1.4.2. Formato de especificación de requerimientos de sistema.**

Notación de especificación de requerimientos de sistema.

1. Utilizar el formulario REQS-XX.
2. Especificar los requerimientos utilizando lenguaje natural estructurado.
3. Para cada requerimiento especificar el requerimiento de usuario que amplía y numerarlos desde 01 dentro de tal requerimiento. La identificación de un requerimiento de sistema es el número de requerimiento de usuario seguido del correlativo, separados por un guión. Por ejemplo, 01.01.03.
4. Especificar la clase de requerimiento con una letra mayúscula, así:  
 F → funcional.  
 N → no funcional.
5. Completar la especificación del tipo de requerimiento no funcional agregando una letra minúscula a la derecha de la letra N según el tipo de requerimiento, así:  
 o → operativo.  
 i → implantación.  
 d → desarrollo.  
 t → técnico.
6. Especificar el tipo de requerimiento con una letra mayúscula así:  
 O: Obligatorio → *debe*.  
 D: Deseable → *debería*.

En la Tabla 1.10 se presenta la notación de la hoja de especificación de requerimientos de sistema.

TEMA		PaginaX de y	
Autor :		Fecha:	
Req. Usu.	Especificación	prioridad	Tipo
01			
01.01			
01.01.01			
02			
02.01			
02.02			

**Tabla 1.10: Hoja de especificación de requerimientos de sistema (RQS-01)**

En la Tabla 1.11 se presenta la notación de especificación de requerimientos de sistema utilizando una forma estándar. Esta notación permitirá describir de una forma más completa los requerimientos del sistema.

TEMA		PaginaX de y	
<b>Autor :</b>		<b>Fecha:</b>	
Req. Usu.	Especificación	prioridad	Tipo
<b>01</b>			
<b>01.01</b>			
<b>01.01.01</b>			
<b>02</b>			
<b>02.01</b>			
<b>02.02</b>			

Tabla 1.11: Especificación de requerimientos de sistema utilizando una forma estándar

### 1.4.3. Metodología de obtención de requerimientos

Dependiendo del tipo de usuario, se diseña una serie de preguntas abiertas para desarrollar la entrevista. Cada pregunta está enfocada en determinar cuáles son las necesidades que se tienen, a partir del puesto de trabajo en el que el usuario se desempeña. Además de querer conocer como se realizan los procesos actuales y como les gustaría que fueran con el Data Warehouse. Al final de la entrevista, se filtran las necesidades y se ordenan.

Una vez terminadas las entrevistas con todos los usuarios, se resumen las necesidades y se agrupan. Se realiza una nueva entrevista con el usuario, donde se presentan los requerimientos resumidos y se pide las clasifique por orden de importancia, según tabla siguiente.

Importancia del requerimiento	Significado
<b>Alta</b>	Son los requerimientos más importantes que se desean resolver.
<b>Media</b>	Son requerimientos que se quieren resolver pero los usuarios pueden esperar a que se realicen.
<b>Baja</b>	Son requerimientos que brindan un valor agregado a la solución, pero no influyen en la satisfacción que los usuarios tengan de la misma.

Tabla 1.12: Tabla de prioridades

## 1.5. ANTECEDENTES

La defensoría del consumidor nace con el objetivo de “Proteger el bolsillo de los Salvadoreños” palabras ofrecidas por el Señor Presidente de la República Elías Antonio Saca en su discurso durante la pronunciación de su cargo como Presidente en el 2004. Como objetivo primordial se considera proteger los derechos de los consumidores velando por la economía familiar dado que se habían identificado abusos por parte de comerciantes y negocios.

Asignada como representante legal a la comisionada Licenciada Ana Evelyn Yacir de Lovo quien tenía como misión principal encontrar herramientas legales e institucionales para proteger al consumidor salvadoreño. La Licenciada Evelyn Yacir coordinó la creación de una nueva ley de Protección al Consumidor para después crear la Defensoría del Consumidor.

La necesidad de controlar e identificar los abusos a los consumidores viene como resultado de un sistema económico globalizado, el cual permite amplios niveles de competencia dentro de diferentes rubros y servicios, comercio internacional y ahora en nuestra época la explotación del comercio electrónico, entre otras actividades económicas. Aunado a este preámbulo cabe mencionar que en nuestra Constitución de la República también se menciona la protección de los consumidores por parte del Estado. (Art. 101).

Como consecuencia a la relación entre los consumidores y proveedores surge la necesidad de crear el Sistema Nacional de Protección al Consumidor (SNPC) el cual brinda un soporte importante a la ley del de Protección al Consumidor, integrado por las dependencias del Órgano Ejecutivo y otras instituciones del Estado los cuales tienen como tarea primordial velar sectorialmente por los derechos de los consumidores o vigilar a empresas que operan con el público. Las empresas que son parte del Sistema Nacional de Protección al Consumidor (SNPC) son:

<b>DC:</b>	Defensoría del Consumidor
<b>SSF:</b>	Superintendencia del Sistema Financiero
<b>Valores:</b>	Superintendencia de Valores
<b>Pensiones:</b>	Superintendencia de Pensiones
<b>MSPAS:</b>	Ministerio de Salud Pública y Asistencia Social
<b>CSSP:</b>	Consejo Superior de Salud Pública
<b>MAG:</b>	Ministerio de Agricultura y Ganadería
<b>MINEC:</b>	Ministerio de Economía
<b>MARN</b>	Ministerio del Medio Ambiente y Recursos Naturales
<b>SIGET</b>	Superintendencia General de Electricidad y Telecomunicaciones
<b>ANDA:</b>	Administración Nacional de Acueductos y Alcantarillados
<b>VMVDU:</b>	Vice- Ministerio de Vivienda y Desarrollo Urbano
<b>Competencia:</b>	Superintendencia de Competencia

En 2007 se crea la PNPC Policía Nacional de Protección al Consumidor, la cual tenía como objeto velar por aspectos vitales para la población en general.

- Agua
- Alimentos
- Electricidad
- Hidrocarburos
- Medicamentos
- Publicidad
- Sistema Financiero
- Telecomunicaciones
- Vivienda

Con la PNPC la Defensoría no asume competencias de las otras instituciones en relación a la protección sectorial del consumidor, ni mucho menos suplanta su rango de acción, sino que con la ayuda de las normativas de cada institución ejerce la coordinación del mismo.

**1.5.1. La Defensoría del Consumidor**

Hasta la fecha cuenta con cuatro oficinas a nivel nacional, dos en San Salvador, una en San Miguel y otra en Santa Ana, para lograr cubrir de manera general al área occidental, central y oriente del país. El método para resolver diferencias entre consumidores y comerciantes se basan en el consenso y dialogo basándose siempre en las leyes establecidas desde un inicio, es así el éxito de esta metodología aplicada que aproximadamente el 96% de los casos se resuelven en la etapa de advenimiento.

**1.5.1.1. MISIÓN**

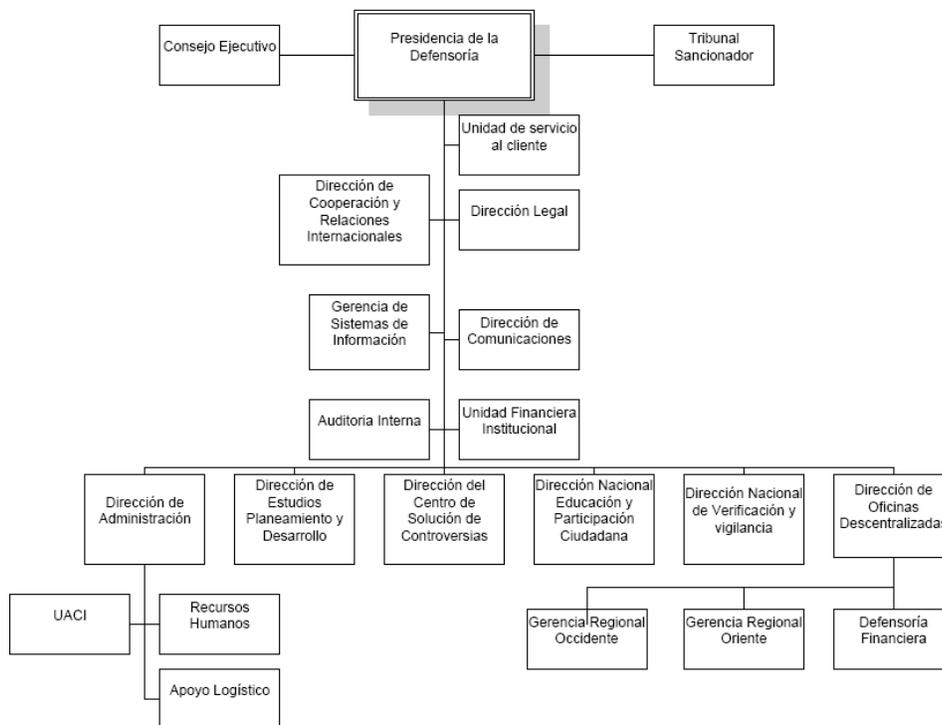
Promocionar, proteger, vigilar y hacer valer los derechos de los consumidores, y coordinar el Sistema Nacional de Protección al Consumidor, mejorando el funcionamiento del mercado, la convivencia social y el ejercicio de ciudadanía.

**1.5.1.2. VISIÓN**

Ser una institución confiable para todos los sectores y reconocida a nivel nacional e internacional por proteger los derechos de los consumidores, procurando el equilibrio, certeza y seguridad en sus relaciones con los proveedores.

**1.5.1.3. ORGANIZACIÓN:**

La organización de la Defensoría del Consumidor está definida de la siguiente manera:



**Figura 1.6: Estructura organizativa de la DC**

#### 1.5.1.4. Funciones

##### ***Consejo Consultivo***

Órgano técnico asesor de la Defensoría del Consumidor.

##### ***Presidencia de la Defensoría***

Máxima autoridad de la institución. Encargada de administrar y supervisar las actividades en general de la Defensoría.

##### ***Tribunal Sancionador***

Basados en la Ley de Protección al Consumidor le corresponde el ejercicio de la potestad sancionatoria tomando en cuenta los lineamientos sancionatorios y basándose en los derechos y la ley correspondiente.

##### ***Dirección de cooperación y relaciones internacionales***

Brinda apoyo a la defensoría en negociación y obtención de programas o proyectos de cooperación financiera no reembolsable y asistencia técnica con fuentes bilaterales y multilaterales a nivel internacional.

##### ***Dirección Legal***

Se encarga en velar que las acciones que se ejecuten dentro y fuera de la defensoría estén basadas en la constitución de la república, leyes internas, reglamentos, velar por elementos jurídicos y la coordinación de la creación de documentos y decretos legislativos y ejecutivos que se relacionen con la institución.

##### ***Gerencia de Sistemas de Información.***

Busca la modernización de la Defensoría del Consumidor por medio de la automatización integral de procesos en las direcciones en las que está constituida la defensoría, ofreciendo servicios de análisis y desarrollo que generen un beneficio cuantificable que logre solidificar elementos Tecnológicos de la institución. Soporte técnico, capacitaciones y estudios de proyecciones a futuro son elementos que se deben de considerar para buscar cumplir objetivos gerenciales.

##### ***Dirección de Comunicaciones***

Su tarea principal es obtener consolidar y distribuir la información sobre el rendimiento en el trabajo que ejerce la Defensoría, de tal manera de lograr informar a los sectores que conforman la opinión pública, así como también de lograr establecer una relación adecuada con los medios de comunicación en general y con todas aquellas instituciones relacionadas.

##### ***Auditoría Interna***

Tiene como objetivos principales evaluar el grado de cumplimiento y eficacia de los sistemas de operación, administración e información, así como de los procedimientos de control interno incorporados a ellos. Además, le compete determinar la confiabilidad de los registros, a través de exámenes de componentes de los resultados financieros; así como facilitar el análisis, evaluaciones y recomendaciones, que contribuyan al mejoramiento de los controles internos.

***Unidad Financiera Institucional***

Tiene como principal objetivo cumplir con la gestión financiera institucional, a través de las etapas del proceso administrativo financiero en las áreas de presupuesto, tesorería y contabilidad gubernamental, las que deben desarrollarse en forma integrada a través de sistemas mecanizados, con eficiencia y eficacia.

***Dirección de administración***

Tiene como principal objetivo contribuir a que las unidades que integran la Defensoría del Consumidor funcionen eficientemente, proporcionándoles de manera oportuna los servicios administrativos de apoyo necesarios; asimismo, le compete velar por la correcta aplicación de políticas y estrategias administrativas, considerando los lineamientos emanados de la Presidencia de la institución y las normativas legales aplicables.

***UACI***

A la unidad de adquisición y contratación institucional (UACI) le compete desarrollar los procesos especializados de compras en la institución, cumpliendo con la ley respectiva.

***Recursos Humanos***

Le compete coordinar eficientemente los procesos de recurso humano de la Defensoría, dotar, administrar, desarrollar y controlar el recurso humano, dando cumplimiento a la normativa legal vigente y las normas y procedimientos que establezca la Defensoría. También desarrolla procesos para el reclutamiento y selección de personal, coordinación de capacitaciones, elaboración de planillas, entre otros.

***Apoyo Logístico***

Integra oportuna y con eficiencia a los procesos, coordinando los bienes y servicios, apoyar logísticamente a las dependencias y gestionar ante las áreas de gestión pertinente la dotación oportuna de dichos recursos.

***Dirección nacional de educación y participación ciudadana.***

Tiene como principal objetivo realizar las acciones que posibiliten el derecho de los consumidores a ser educados e informados en materia de consumo, así el derecho a agruparse en asociaciones de consumidores para la protección de sus intereses y a participar en el Sistema Nacional de Protección al Consumidor (art 38 literal g, y 58 literales e y n).

***Dirección nacional de estudios, verificación y vigilancia***

Tiene como principal objetivo dar soporte a las funciones de vigilancia e inspección que establece el artículo 58 literal f de la ley de protección al consumidor. Para tales efectos, realizará inspecciones o investigaciones a fin de verificar el cumplimiento de las disposiciones de la Ley y algunas Normas Salvadoreñas Obligatorias. (Art. 58 literales j y k). En el área de estudios se cuenta con una subdirección de estudios.

***Subdirección de estudios***

Dirige la elaboración de estudios e investigaciones en materia de consumo en cumplimiento a las funciones de la Ley y Reglamento de Protección al Consumidor.

***Dirección del Centro de Solución de Controversias del Área Central***

Tiene como principal función resolver los conflictos entre proveedores y consumidores, a través de medios alternos de solución de controversias, de manera simple, breve, gratuita y confidencial. (Art. 108) El CSC cuenta con órganos internos como el centro de atención de llamadas (call center).

***Dirección de oficinas descentralizadas y desarrollo organizacional***

Tiene como principal objetivo la descentralización de las funciones de la institución, a fin de propiciar un acercamiento a los consumidores de las herramientas para hacer valer sus derechos. Cuenta con la oficina regional de occidente, la oficina regional de oriente y la oficina de atención financiera.

***Gerencia Regional Occidente***

Brinda asesoría y atiende reclamos de los consumidores de la zona occidental del país.

***Gerencia Regional de Oriente.***

Brinda asesoría y atiende reclamos de los consumidores en la zona oriental del país.

**1.5.2. Servicios****1.5.2.1. Asesoría:**

Análisis de procedencia del reclamo hecho por el consumidor. En caso que no proceda el reclamo por razones de fondo, puede recomendarse al denunciante el uso de vías administrativas o judiciales idóneas para la satisfacción a su pretensión. (Art. 110 L. P. C.).

**1.5.2.2. Avenimiento:**

Mecanismo basado en equidad y justicia, en el que se aplica cuanto medio sea necesario para la comunicación directa con el proveedor para buscar solución expedita a la pretensión del consumidor. (Art. 110 L. P. C.).

**1.5.2.3. Mediación:**

Método alternativo de solución de controversias que procede cuando existe petición expresa del consumidor; en el cual imperan los principios de confidencialidad e imparcialidad; el mediador será un facilitador de la comunicación entre las partes y podrá celebrar audiencias conjuntas o privadas con cada uno de los interesados, previa comunicación y consentimiento de la otra parte. El acuerdo total o parcial tomado en Mediación, producirá los efectos de la transacción y su certificación tendrá fuerza ejecutiva. (Arts. 115, 116 y 117 L. P. C.).

**1.5.2.4. Conciliación:**

Método alternativo de solución de controversias que procede cuando existe petición expresa del consumidor para proceder directamente a ello, o si una vez intentado el avenimiento sin ningún resultado satisfactorio, el consumidor ratifica la denuncia y ninguna de las partes solicitan mediación o arbitraje; en este método el funcionario designado actúa como un moderador quien oír a ambas partes y podrá fin al debate en el momento que considere oportuno; los acuerdos tomados en conciliación producirán los efectos de la transacción, y la certificación del acta tendrá fuerza ejecutiva. (Arts. 111 y 113 L. P. C.).

**1.5.2.5. Arbitraje:**

Método alternativo de solución de controversias que procede ante la falla de la conciliación y la mediación, o bien ante la petición expresa de su realización o designación contractual previa. Puede ser en equidad, técnico o de derecho; siendo el primero el también llamado de amigables compondores en el cual los árbitros proceden con entera libertad y deciden según sea más conveniente al interés de las partes, sin atender más que a su conciencia, la verdad y la buena fe; el arbitraje técnico es el dado en razón de específicos conocimientos en una determinada ciencia, arte u oficio; y el último es aquel en que los árbitros fundan sus decisiones en la ley. La certificación del laudo arbitral firme tiene la misma fuerza y validez de una sentencia ejecutoriada y pasada en autoridad de cosa juzgada. (Arts. 119, 120 y 142 L. P. C.).

**1.5.2.6. Tribunal sancionador:**

Ente permanente y colegiado de la Defensoría del Consumidor con las atribuciones de instruir los procedimientos sancionatorios en materia de protección al consumidor; imponer las sanciones y conocer de los demás asuntos que tiene atribuidos por ley. (Arts. 79 y 83 L. P. C.).

# Capítulo II



## 2.1. ANALISIS DE LA SITUACION ACTUAL

### 2.1.1. Resultados de la investigación actual

#### 2.1.1.1. Metodología utilizada para la investigación

La metodología seguida para la investigación de la situación actual está basada en herramientas de recolección de información. La utilización de estas herramientas será en coordinación conjunta con la DC para la programación de actividades que requieran algún tipo de participación dentro del proceso investigativo.

**Entrevistas:** Las entrevistas fueron realizadas con la programación de reuniones de aproximadamente 1 horas de duración. La estructura de las entrevistas fueron modificadas de acuerdo a los entrevistados, manteniendo el mismo objetivo según lo propuesto por Kimball<sup>10</sup> Los entrevistados fueron:

- Gerencia de Sistemas de Información
  - a. Ana Cecilia Yescas de Serrano (Gerente de Sistemas de Información)
  - b. Juan José Rivas (Coordinador de Desarrollo de Sistemas)
- Dirección de Estudios, Planeamiento y Desarrollo
  - a. Leily Melany Mendoza Valladares (Directora de Estudios)
  - b. Ada Aurora Marquina Arce (Analista de Planeamiento)
  - c. Manuel Ortiz Pedroza (Técnico de Estadísticas)
- Dirección del Centro de Solución de Controversias
  - a. Sandra Evelyn Liborio (Directora del Centro de Solución de Controversias)
  - b. María del Transito Elena de Santos (Técnico Conciliador)

**Observación:** Las observación realizada para la investigación de la situación actual estuvo dirigida a la apreciación del flujo de información dentro de la defensoría del consumidor. Confirmando los diferentes flujos existentes hacia las diferentes direcciones organizacionales, confirmando la delicadez de mantener constante estos flujos.

**Cuestionario:** La recolección de la información específica y puntual se realizo a través de esta herramienta. Se obtuvo la información siguiente:

- Existencia de Servidor de Base de Datos de Desarrollo
- Existencia de Servidor de Base de Datos de Producción
- Existencia de Servidor de Aplicaciones de Desarrollo
- Existencia de Servidor de Aplicaciones de Producción
- Existencia de Servidor para Almacenar Aplicación a Desarrollar<sup>11</sup>

#### 2.1.1.2. Procesos de situación actual

En base a la información recolectada mediante todas las herramientas y actividades mencionadas anteriormente, se puede definir los procesos principales llevados a cabo, que han sido detectados como principales puntos de atención para identificar los problemas.

A continuación se presentan los procesos más importantes a tomar en cuenta para el análisis de la situación actual. En la Figura 2.1 se presenta como funciona cotidianamente el flujo de información entre

<sup>10</sup> The Data Warehouse Lifecycle Toolkit, Ralph Kimball – Capítulo 4 “Recolección de Requerimientos”

<sup>11</sup> Desarrollo de un ‘Data Warehouse’ para el proceso de denuncias de la Defensoría del Consumidor, Anteproyecto – Capítulo 11.2 “Factibilidad Operativa”

la Dirección de Estudios, Planeamiento y Desarrollo (DE), y las demás Direcciones Organizacionales (DO). En el flujo se puede observar que existen dos puntos de caminos o acciones condicionales:

“**información adicional?**”: En caso que alguna de las DO necesite información adicional a la que fluye actualmente, esta deberá notificar a la DE sobre esta necesidad. De no necesitar información adicional, las DO seguirán realizando sus procesos diarios.

“**Reporte nuevo?**”: De existir la necesidad de contar con nueva información con un reporte nuevo o con la modificación de un reporte existente, la DE hace una solicitud de información a la Gerencia de Sistemas de Información (GS), esta se encarga de dicha solicitud, modificando o creando un nuevo reporte. Posteriormente se notifica la existencia del nuevo reporte para que pueda ser utilizado por la DE. De no existir la necesidad de un reporte nuevo o la modificación de un reporte existente, el flujo de información continua normalmente con la generación de reportes y su envío a la diferentes DO.

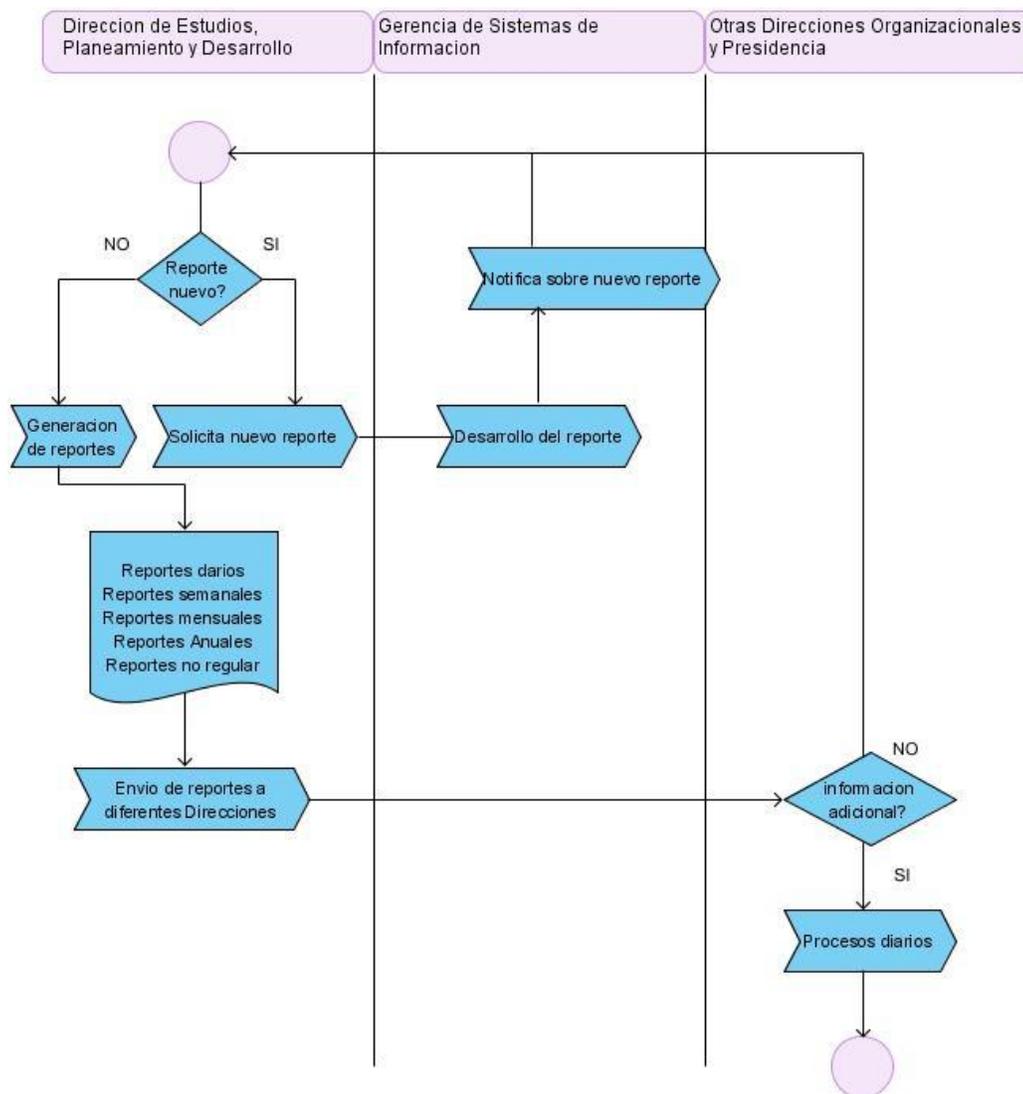


Figura 2.1: Flujo de información entre DE, DO y GS

En el caso de la DE, tiene la función de informar a las demás DO de los estudios, planeaciones y desarrollos que se realizan, por lo que se transforma en un punto de partida de flujo de información. Mas sin embargo cada unidad posee sus propios procesos y análisis con información generada por ellos o por los diferentes sistemas.

En la Figura 2.2 se muestra el comportamiento del flujo de información en una DO y su relación con la GS. En el flujo se puede observar que existe un punto de camino o acción condicional:

**”Información adicional?”**: En caso que la DO necesite información adicional a la que fluye actualmente, esta deberá notificar a la GS sobre esta necesidad, haciendo una solicitud de un nuevo reporte o la modificación de un reporte existente. La GS se encarga de dicha solicitud, modificando o creando un nuevo reporte. Posteriormente se notifica la existencia del nuevo reporte para que pueda ser utilizado por la DO. De no existir la necesidad de información adicional, el flujo de información sigue su flujo normal.



Figura 2.2: Flujo de información ente DO y GS

## 2.1.2. Análisis de la situación actual

### 2.1.2.1. Representación de la Situación Actual

La estructura organizacional de la DC se puede observar en la Figura 2.3, donde se presenta el organigrama de la entidad. Se observa que existe la “Dirección Estudios Planteamiento y Desarrollo”, la cual está encargada de realizar estudios e investigaciones en materia de consumo y desempeño, dicha actividad es la mayor generadora de información ya procesada, generando reportes gerenciales, estratégicos y operativos, los cuales son repartidos a las demás DO, tal como se representa en la Figura 2.1.

Los flujos de información también son locales a las DO, tal como se presento en la Figura 2.2. El factor común a estos dos escenarios, es la participación de la GS en el momento que el flujo de información se ve interrumpido por falta de información, no existente en los reportes actuales.

Al flujo de información mencionado anteriormente, se tiene actualmente como problema identificado (estado A);

“Existen un proceso intermedio para la obtención de información no contemplada en los reportes y sistemas actuales, lo que implica la inversión no adecuada de recurso humano, tiempo y equipo, que con lleva a atrasos para el análisis, seguimiento, control y toma de decisiones para la gestión de operaciones”.

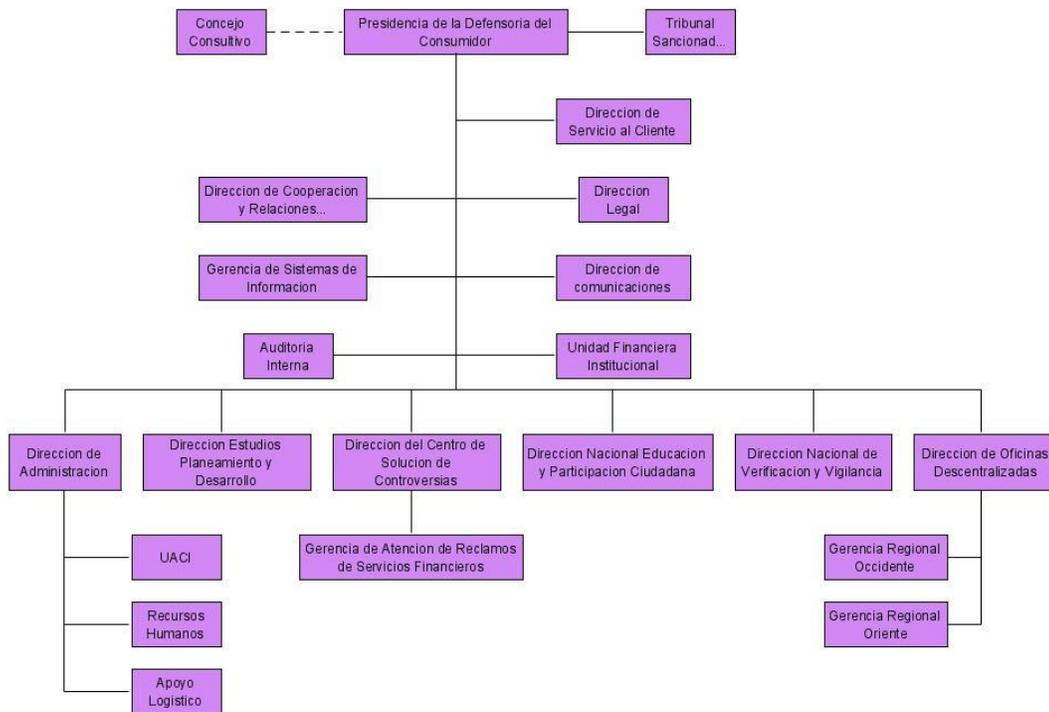


Figura 2.3: Organigrama

Este problema puede visualizarse en la Figura 2.4, en el cual se encuentran los siguientes elementos:

- **Gerencia de Sistemas de Información:** Esta encargada de recibir toda solicitud de cambios en los reportes actuales o creación de reportes nuevos para las direcciones organizacionales. Las solicitudes nacen de la necesidad de información adicional a los reportes existentes o la generación de nuevos reportes, que los sistemas actuales no son capaces de brindar. La gerencia recibe la solicitud y esta es formulada como un requerimiento a desarrollarse,

asignando recurso (técnico) para su realización. El técnico debe analizar, desarrollar, hacer pruebas e implementar los requerimientos. Una vez implementado el requerimiento, es notificado a la dirección correspondiente, para que puedan hacer uso de la nueva información.

- **Dirección Estudios Planeamiento y Desarrollo:** Está encargada de realizar estudios e investigaciones en materia de consumo y desempeño. Es una de las principales direcciones que poseen una gran necesidad de información, convirtiéndose en la dirección solicita nuevos requerimientos. Esta dirección brinda información a las diferentes direcciones organizacionales y a la presidencia directamente, siendo un foco de flujo de información, pero dependiendo de la información capaz de brindar los diferentes sistemas.
- **Direcciones organizacionales y Presidencia:** Las demás direcciones son capaces de hacer solicitudes de información adicional a los reportes existentes o la generación de nuevos reportes, esto a la “Gerencia de Sistemas de Información”.

Se puede identificar que para poder acceder a nueva información, es necesario pasar un proceso intermedio, el cual implica la solicitud de información nueva, dicha solicitud se transforma en un requerimiento, realizando análisis, desarrollo, pruebas e implementación del mismo en el ambiente de producción. Esto puede resultar en atrasos de estudios, investigación o en el desempeño del trabajo diario de las direcciones organizacionales.

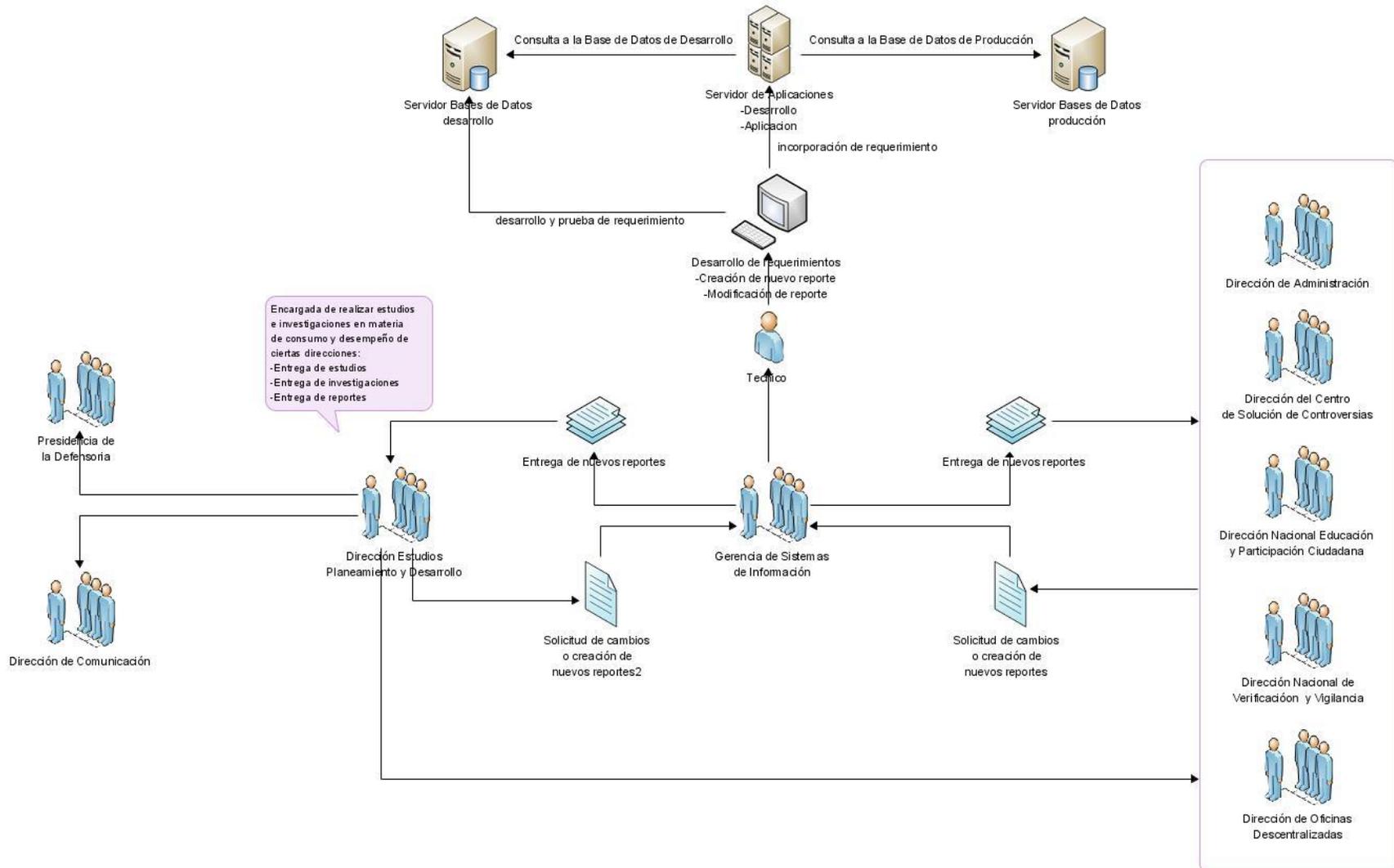


Figura 2.4 Situación actual

## 2.2. PLANTEAMIENTO DEL PROBLEMA

Actualmente la Defensoría del Consumidor posee dos sistemas de información para la administración y control de las actividades de las diferentes direcciones organizacionales, dichos sistemas no brindan reportes con un nivel de detalle que cumpla con las necesidades cambiantes de los usuarios finales.

Es por esto que surge la necesidad de crear un sistema dedicado a brindar información operacional, administrativa y gerencial, con niveles de detalles y flexibilidad que permita la exploración, extracción, análisis e impresión de reportes acoplado a lo solicitado por el usuario. La institución posee como necesidad actual y de gran importancia, el implementar un sistema capaz de:

- Realizar exploración de información.
- Realizar búsqueda de datos filtrada por diferentes criterios.
- Crear reportes de nivel operativo con el nivel de detalle de información que cumpla con la necesidad del usuario final, de acuerdo a los criterios de análisis.
- Crear reportes de nivel administrativo con el fin de apoyar las actividades de análisis, seguimiento, control y toma de decisiones.
- Crear reportes de nivel estratégico con el fin de realizar las actividades de planificación a mediano y largo plazo a nivel administrativo.

### 2.2.1. Análisis de la Situación Problemática

Para realizar un análisis de la problemática a tratar, se utilizará el Diagrama de Causa-Efecto el cual permite apreciar con claridad las relaciones entre un problema y las posibles causas que contribuyen para que este ocurra.

Para detectar las causas que generan el problema se han tomado en cuenta las siguientes categorías: Recurso humano, tiempo, Recurso informático, Procesos y Medio ambiente, los cuales se describen a continuación:

#### 2.2.1.1. Recurso humano:

- El recurso humano de la Gerencia de Sistemas es constantemente reasignada a actividades no planificadas, para dar atención a la solicitud de reportes de información no mostrada por los actuales sistemas.
- Los empleados de las diferentes direcciones organizacionales no pueden aumentar su productividad, ya que no depende directamente de ellos el evitar estancamiento de casos al no poseer información suficiente para el análisis, seguimiento, control y toma de decisiones.

#### 2.2.1.2. Recurso Información:

- No es posible realizar exploración ágil de datos históricos y actuales.
- Los sistemas actuales no son capaces de generar reportes dinámicamente, con el nivel de detalle necesario para realizar las actividades de análisis, seguimiento, control y toma de decisiones.
- Los sistemas actuales no muestran reportes de información histórica resumida para la realización de actividades de planificación a mediano y largo plazo, a nivel administrativo.

**2.2.1.3. Recurso informático:**

- La información almacenada y procesada por los sistemas de la Defensoría del Consumidor, se encuentra alojada dentro de una base de datos relacional, ya que los sistemas son de tipo OLTP, los cuales no posee las características y funcionalidades necesarias para la proporción de información para la toma de decisiones.

**2.2.1.4. Tiempo:**

- Las tareas de los empleados de las diferentes direcciones organizacionales tienden a alargarse proporcional al tiempo de respuesta de la gerencia de sistemas en los casos en que se necesite información más detallada o adicional que no es mostrada por los sistemas y reportes actuales.
- El tiempo planificado para el recurso humano del área de la gerencia de sistemas se ve alterado, perdiendo aproximadamente 5 horas por cada solicitud de nuevos reportes.
- La disminución del tiempo para la solución o estudios de casos no puede ser reducida mientras no exista información que se entregue justo a tiempo y que llene las necesidades de lo requerido por el usuario final para dichos estudios.

**2.2.1.5. Procesos:**

- Dificultad en la búsqueda de información necesaria para el estudio de casos.
- Dificultad para obtener información justo a tiempo, en los casos que no existan reportes con la información necesaria.
- No existe la posibilidad de realizar análisis de tendencias y comportamiento de datos históricos y actuales.

**2.2.1.6. Medio ambiente:**

- No existe un sistema de información actualmente que automatice la generación de reportes dinámicos y exploración de datos, lo cual evita la creación de nuevos proyectos en beneficio de la población en general.
- La población en general exige continuamente el mejoramiento de los servicios brindados por la Defensoría del Consumidor, para hacer valer sus derechos y solventar con mayor eficiencia la resolución de casos.
- Toda dirección organizacional ha expresado la necesidad de realizar con mayor eficiencia la obtención de información y suprimir en la medida de lo posible el retraso del procesamiento de denuncias.

Con respecto a las causas que originan la problemática, se detecta el siguiente efecto:

“Existen un proceso intermedio para la obtención de información no contemplada en los reportes y sistemas actuales, lo que implica la inversión no adecuada de recurso humano, tiempo y equipo, que con lleva a atrasos para el análisis, seguimiento, control y toma de decisiones para la gestión de operaciones”

Para el diagnostico que determine el problema se ha utilizado el diagrama causa – efecto, ya que permite observar los principales elementos o causas que contribuyen al efecto o al problema. En la figura 2.5 se puede observar el diagrama causa – efecto en donde se visualizan las causas que están ocasionando el problema para la proporción de información para el análisis, seguimiento, control y toma de decisiones sobre los casos.

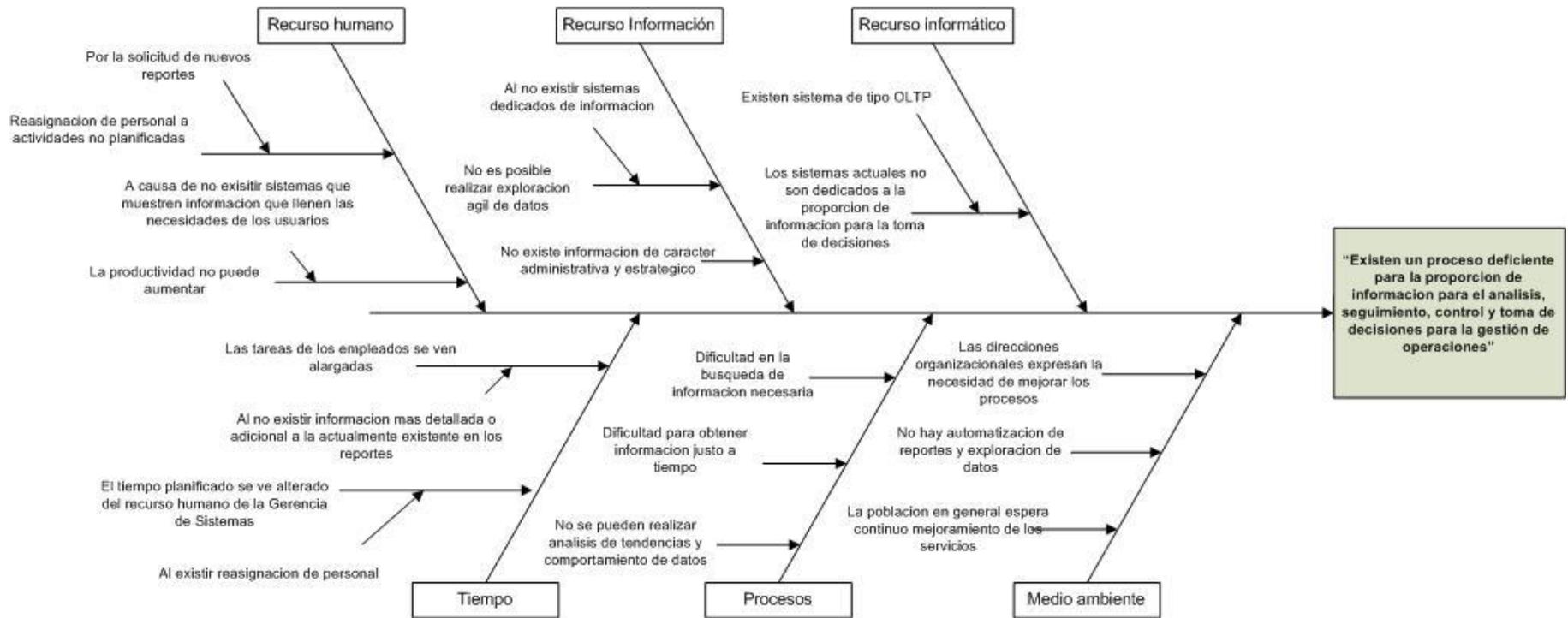


Figura 2.5: Diagrama causa- efecto para la obtención del problema

### 2.2.2. Definición del problema

En base al diagnóstico realizado mediante el diagrama causa - efecto, se ha determinado que el problema en la Gerencia de Sistema de Información es:

¿En qué medida el Desarrollo de un data warehouse para el proceso de denuncias de la Defensoría del Consumidor mejorará la gestión de operaciones, administración de las direcciones y la toma de decisiones?

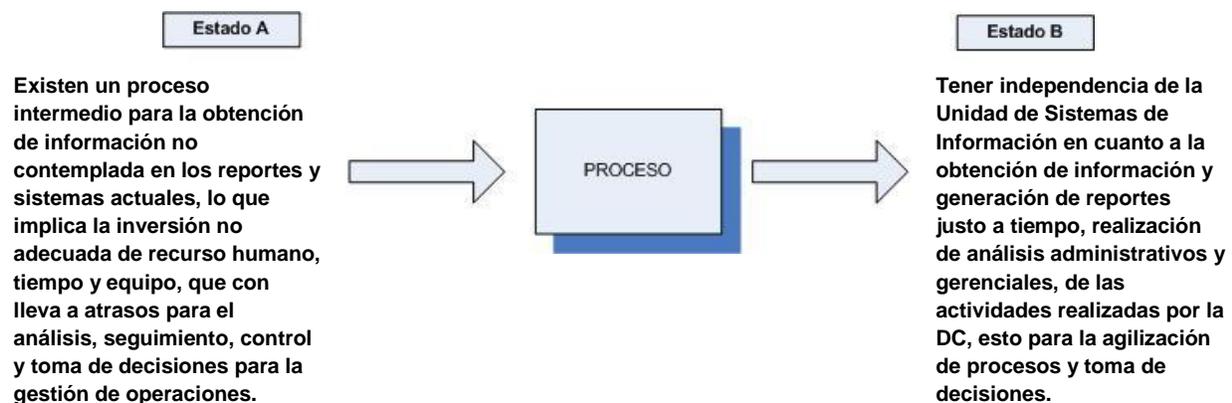
### 2.2.3. Formulación del problema

El método a utilizar para la formulación del problema es el Diagrama de la Caja Negra, mediante este se permitirá tener una visión general del problema, el cual parte de una situación actual (Estado A) hasta una situación deseada (Estado B), pasando a través de un proceso de transformación para obtener una solución. En la figura 2.6 se muestra la representación grafica.



Figura 2.6: Diagrama de la caja negra

A continuación se presenta la formulación del problema que se presenta en la Gerencia de Sistemas de Información de la Defensoría del Consumidor:



A1. No existe un sistema que suprima la dependencia de información con la Gerencia de Sistema de Información para evitar reasignación de tareas del recurso humano y no permitiendo mejorar la productividad de las direcciones organizacionales.

B1. Eliminación de la dependencia de las direcciones organizacionales de la Defensoría del Consumidor con la Gerencia de Sistemas de Información, aumentando la productividad en sus operaciones.

A2. No existe la exploración ágil de datos y no hay flexibilidad y facilidad para obtener reportes con detalle de acuerdo a las necesidades de las direcciones organizacionales.

A3. Sistemas OLTP no capaces de mostrar información para la toma de decisiones.

A4. Estancamiento de estudios y seguimiento de casos al no poseer la información necesaria justo a tiempo.

A5. No existen formas de búsqueda de información ágil e información con un nivel de detalle de acuerdo a lo necesitado en el momento para obtener datos justo a tiempo para la toma de decisiones y agilización de los procesos.

A6. No es posible la realización de nuevos proyectos en beneficio de la población al no existir la automatización de procesos en la proporción de información, reportes y exploración de datos que puedan acaecer y realizar las direcciones organizacionales.

B2. Obtención de información operativa, administrativa y estratégica de manera fácil y detallada de acuerdo a las necesidades de las direcciones organizacionales.

B3. Sistema OLAP dedicado para brindar información justo a tiempo para el análisis, seguimiento, control y toma de decisiones.

B4. Agilización de los procesos para estudios, resolución de casos, mejorando la gestión de operaciones y evitando el estancamiento por falta de información.

B5. Realización de búsqueda de información de manera fácil y ágil, obteniendo información con un nivel de detalle de acuerdo a las necesidades particulares para la toma de decisiones y agilización de procesos.

B6. Automatización de la generación y adquisición de información por parte de las direcciones organizacionales.

## 2.3. SOLUCIÓN PROPUESTA

### 2.3.1. Identificación de la solución

La solución es la respuesta a un problema o a una situación difícil. En la sección anterior se determino el problema a atacar, bajo una representación grafica, se determino la existencia de un proceso intermedio que nace en medio del flujo de información.

Utilizando el método de la caja negra<sup>12</sup>, la solución propuesta se puede definir como el estado al que se desea llegar (estado B), partiendo de la situación actual (estado A). La solución propuesta es:

*“Tener independencia de la Unidad de Sistemas de Información en cuanto a la obtención de información y generación de reportes justo a tiempo, realización de análisis administrativos y gerenciales, de las actividades realizadas por la Defensoría del Consumidor, esto para la agilización de procesos y toma de decisiones”.*

La transición de un estado a otro se puede representar a través del método de la caja negra, como se muestra en la Figura 2.7.

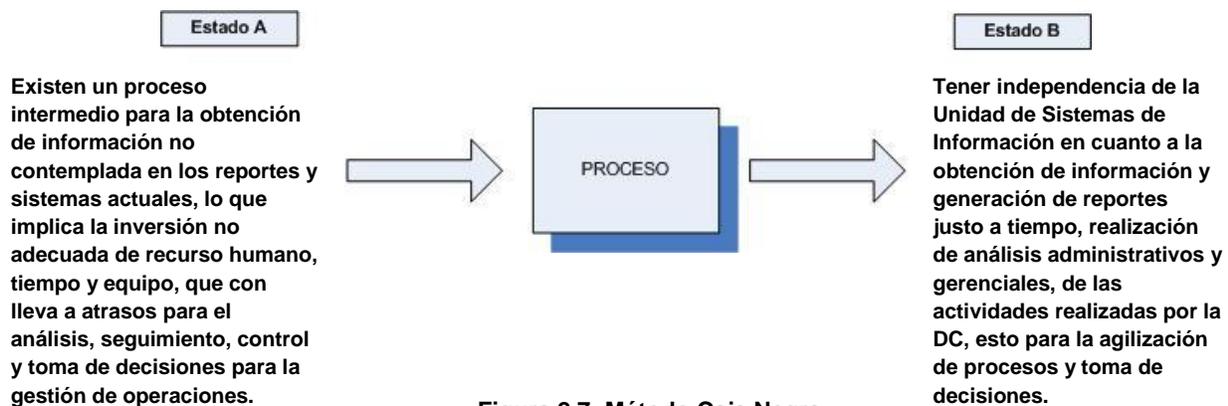


Figura 2.7: Método Caja Negra

La representación grafica de la solución propuesta se puede visualizar en la Figura 2.8, en el que se puede observar que la única relación que se posee con la “Gerencia de Sistemas de Información” es la solicitud de cambios en permisos de los usuarios en el nuevo sistema. Cada una de las direcciones organizacionales tendrá a su disposición el acceso de la aplicación, cuya funcionalidad es la presentación de la información contenida dentro del DW.

A comparación de la situación actual, no puede existir un atraso en las operaciones a causa de la necesidad de información y el flujo de información entre direcciones organizacionales, ya que se tendrá disponible la generación de reportes por acción del mismo usuario que lo necesite, en cualquier dirección organizacional.

**Solicitud de permisos de usuario:** Este proceso se refiere a la habilitación de permisos a un usuario en específico dentro de la aplicación. Los permisos se refieren a la restricción de accesos a ciertas funcionalidades del sistema, como lo son la creación de reportes, consulta de información, etc.

<sup>12</sup> Wikipedia.org, Se denomina caja negra a aquel elemento que es estudiado desde el punto de vista de las entradas que recibe y las salidas o respuestas que produce, sin tener en cuenta su funcionamiento interno.

2.3.2. Representación de la solución

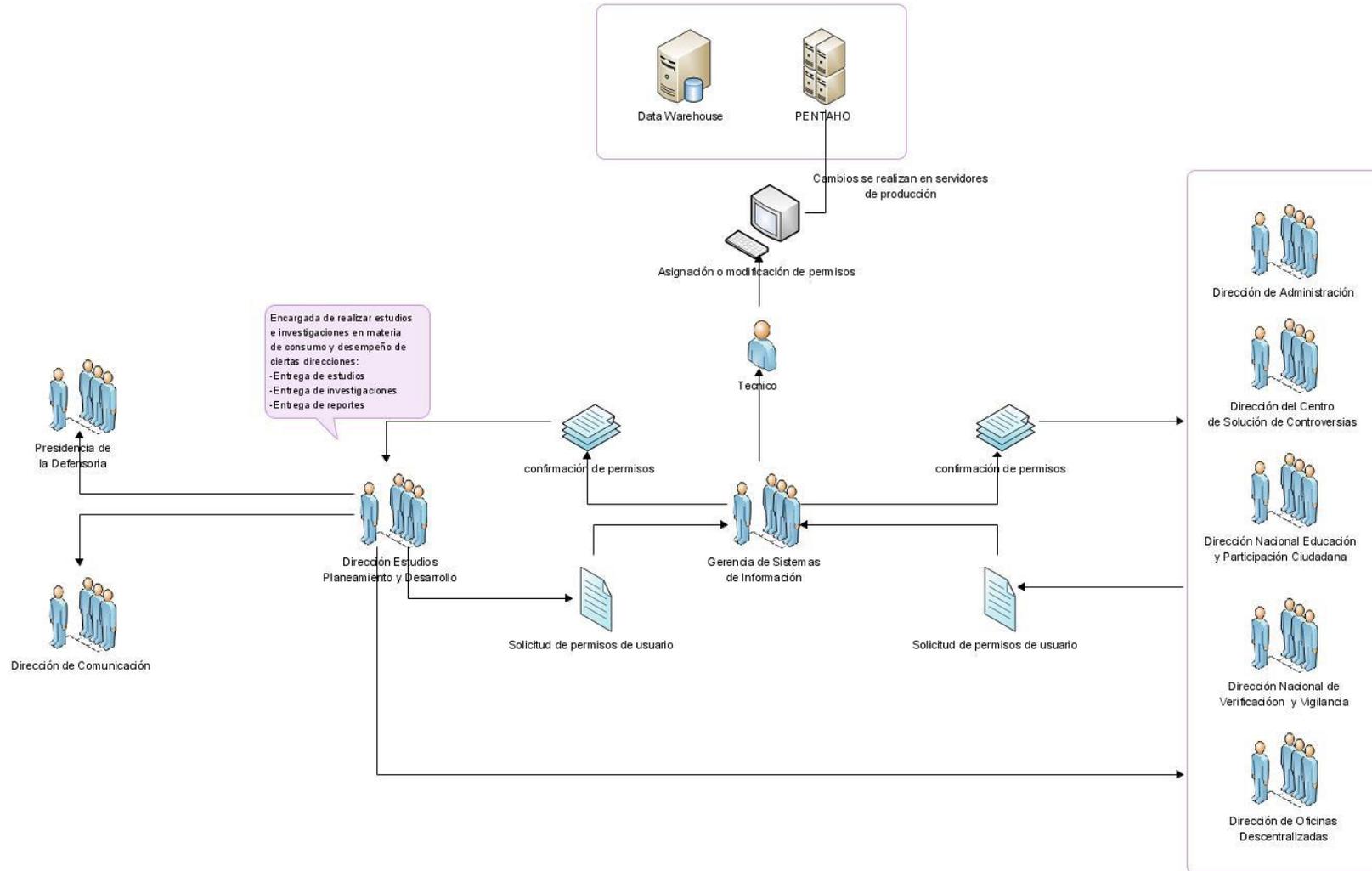


Figura 2.8: Representación de la solución propuesta

## 2.4. ENFOQUE DE SISTEMAS DE LOS REQUERIMIENTOS.

### DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.

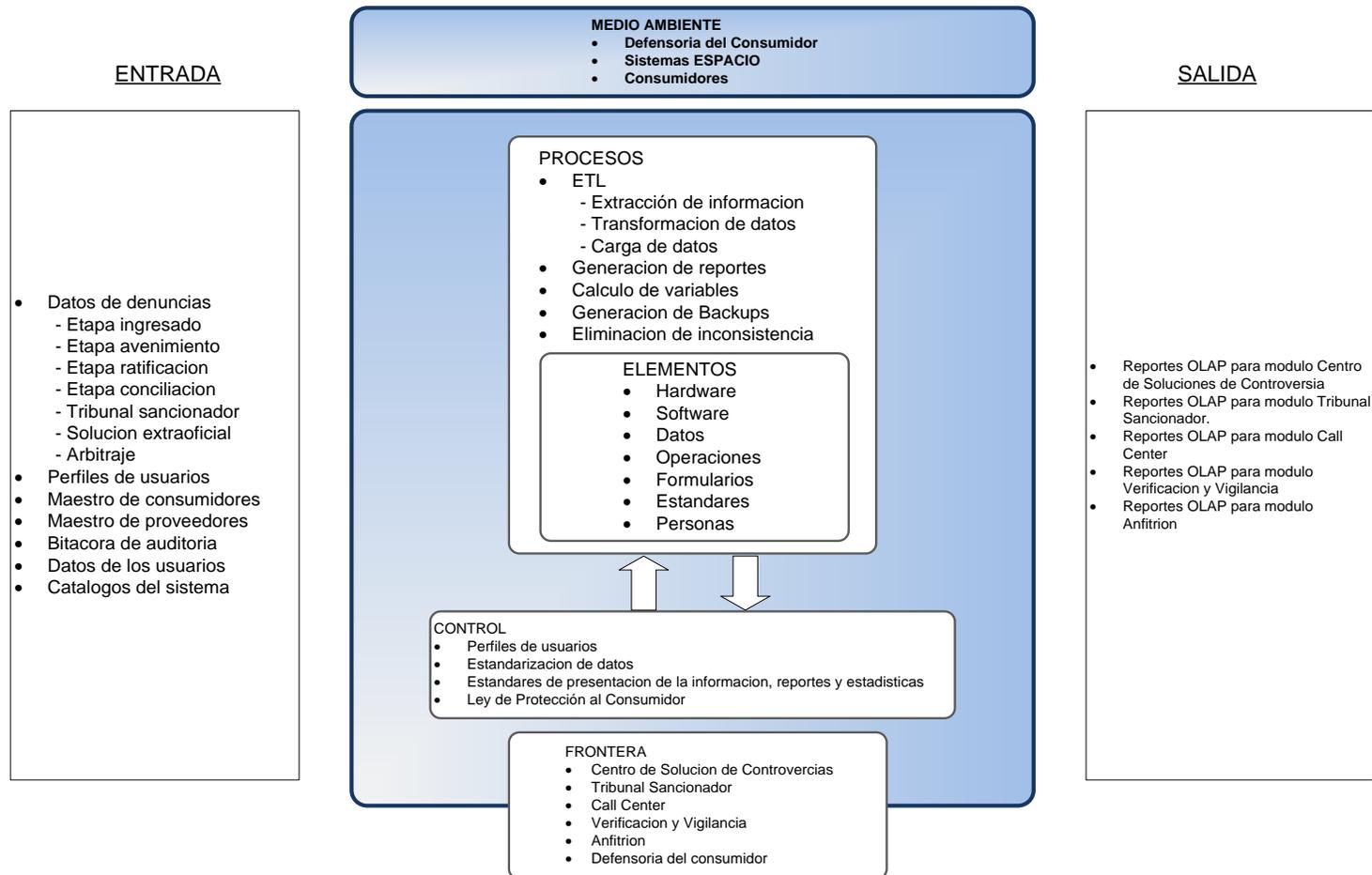


Figura 2.9: Enfoque de sistemas

El enfoque de sistema es la representación gráfica (observar figura 2.9) del comportamiento que tendrá el sistema en estudio, en este caso será el sistema de la solución propuesta. Este enfoque nos ayudara a determinar los componentes de cada uno de los elementos del enfoque de sistema. Elementos del enfoque de sistemas:

**Entrada:** Contiene todos los insumos necesarios para el procesamiento. Se define como entradas toda información concerniente a los 5 módulos que tiene como alcance el proyecto. Los insumos de entrada serán la información de los casos y demás relacionada al caso, la cual es almacenada dentro de la base de datos transaccional.

- Datos de denuncias.
- Perfiles de usuarios.
- Maestro de consumidores.
- Maestro de proveedores.
- Bitácora de auditoría.
- Datos de los usuarios.
- Catálogos del sistema.

**Proceso:** El proceso es la actividad transformadora de insumos, convirtiéndolos en las salidas deseadas. El proyecto posee los siguientes elementos que conforman el elemento procesador del enfoque:

- ETL, es el proceso que permite mover datos desde la base de datos transaccional, reformatearlos y limpiarlos, y cargarlos en otra base de datos, Data Mart o Data Warehouse para analizar y apoyar el proceso de negocio.
- Cálculo de variables, es la definición de las variables de interés para la organización, las cuales se convierten en las dimensiones de un cubo de información de estos se generan los reportes y análisis OLAP. La selección de variables se basa en los requerimientos del sistema, es decir, las salidas que deberá tener.
- Generación de reportes, existirá un proceso encargado de la generación de reportes OLAP, llamados también reportes ad hoc, permiten la creación personalizada de reportes, con las especificaciones necesarias. Por el potencial de estos reportes para degradar el rendimiento de sistemas transaccionales, usualmente son proveídos sobre un Data Warehouse.
- Generación de backups, por la importancia de resguardar la información, se considerara la posibilidad de almacenar respaldos periódicamente.
- Elementos, todos los demás componentes necesarios para procesar los insumos; software a implementar, hardware necesario para la utilización del software y soportar las salidas, generación de datos como parte del procesamiento de información, formularios a presentar a los usuarios finales, estándares necesarios para el procesamiento y generación de salidas, y las personas involucradas en el proceso.

**Salida:** Serán los resultados esperados por el sistema. La presentación de la información consolidada y sintetizada, para la toma de decisiones. La información estará dividida por módulo contemplado.

- Reportes OLAP para módulo Centro de Soluciones de Controversia.
- Reportes OLAP para módulo Tribunal Sancionador.
- Reportes OLAP para módulo Call Center.
- Reportes OLAP para módulo Verificación y Vigilancia.
- Reportes OLAP para módulo Anfitrión.

**Medio Ambiente:** Es el medio con el que interactuara el sistema, el mismo que proporcionara la información de entrada, los recursos necesarios para procesamiento y la información de salida. El medio ambiente está compuesto por:

- Defensoría del Consumidor.
- Sistema ESPACIO.
- Consumidor.

**Control:** son todos aquellos componentes a considerar sobre los cuales se deberá regir el sistema, cumpliendo con las restricciones, estándares, leyes u otros para el procesamiento y generación de las salidas. Los elementos que rigen el sistema son:

- Perfiles de usuario.
- Estandarización de datos.
- Estándares de presentación de la información, reportes y estadísticas.
- Ley de protección al consumidor.

**Frontera:** Cada sistema tiene una frontera que lo separa de su ambiente. La frontera delinea al sistema; lo que está adentro de la frontera es el sistema y los que están afuera es el ambiente en que se desenvuelve el sistema. La frontera está compuesta por:

- Modulo Centro de Solución de Controversias.
- Modulo Tribunal Sancionador.
- Modulo Call Center.
- Modulo Verificación y Vigilancia.
- Modulo Anfitrión.

## 2.5. CONTENIDO TEMÁTICO PROPUESTO

INTRODUCCIÓN

OBJETIVOS

OBJETIVO GENERAL

OBJETIVOS ESPECÍFICOS

ALCANCE Y LIMITACIONES

ALCANCE

LIMITACIONES

JUSTIFICACIÓN

IMPORTANCIA

### 1. CAPITULO I: GENERALIDADES

#### 1.1. LISTADO DE ACRÓNIMOS

#### 1.2. MARCO TEÓRICO

*1.2.1. Business Intelligence*

*1.2.2. Sistemas Transaccionales (OLTP)*

*1.2.3. Sistemas OLAP*

*1.2.4. Data Warehouse*

*1.2.5. Data Marts*

*1.2.6. Dashboards*

*1.2.7. Introducción a Metodologías de Business Intelligence*

#### 1.3. ESTÁNDARES Y DESCRIPCIÓN DE NOMENCLATURA

*1.3.1. Análisis de la situación actual*

*1.3.2. Nomenclatura de casos de uso*

*1.3.3. Descripción de Estereotipos utilizados para el Diseño del Data Warehouse*

*1.3.4. Diseño del proceso ETL (Extract, Transform and Load)*

*1.3.5. Estándar de nomenclatura de tablas y campos del data warehouse*

#### 1.4. FORMATOS DE ESPECIFICACIÓN PARA EL ANÁLISIS DE REQUERIMIENTOS.

*1.4.1. Formato de especificación de requerimientos de usuario*

*1.4.2. Formato de especificación de requerimientos de sistema.*

*1.4.3. Metodología de obtención de requerimientos*

#### 1.5. ANTECEDENTES

*1.5.1. La Defensoría del Consumidor*

*1.5.2. Servicios*

### 2. CAPITULO II: ANÁLISIS

#### 2.1. ANALISIS DE LA SITUACION ACTUAL

*2.1.1. Resultados de la investigación actual*

*2.1.2. Análisis de la situación actual*

#### 2.2. PLANTEAMIENTO DEL PROBLEMA

*2.2.1. Análisis de la Situación Problemática*

*2.2.2. Definición del problema*

*2.2.3. Formulación del problema*

#### 2.3. SOLUCIÓN PROPUESTA

- 2.3.1. *Identificación de la solución*
- 2.3.2. *Representación de la solución*
- 2.4. ENFOQUE DE SISTEMAS DE LOS REQUERIMIENTOS.
- 2.5. CONTENIDO TEMÁTICO PROPUESTO
- 2.6. FACTIBILIDADES
  - 2.6.1. *Factibilidad Técnica*
  - 2.6.2. *Factibilidad Operativa*
  - 2.6.3. *Factibilidad Económica*
- 2.7. PLANEACIÓN DE RECURSOS
  - 2.7.1. *Calculo para el Presupuesto*
  - 2.7.2. *Resumen de Gastos*
- 2.8. METODOLOGÍA
  - 2.8.1. *Herramientas para Recolección de Datos*
  - 2.8.2. *Herramientas para el Diagnostico y Planteamiento del Problema*
  - 2.8.3. *Modelo UML de los Componentes Bodega de Datos*
  - 2.8.4. *Ciclo de vida*
  - 2.8.5. *Metodología del Proyecto*
  - 2.8.6. *Metodología de Kimball*
- 2.9. RESULTADOS ESPERADOS
  - 2.9.1. *Resultados Esperados*
  - 2.9.2. *Producto a Entregar*
- 3. CAPITULO III: DISEÑO Y REFINAMIENTO
  - 3.1. DISEÑO GENERAL
    - 3.1.1. *Modelado del Data Warehouse usando UML*
    - 3.1.2. *Diseño Físico del DW (Ejemplo).*
  - 3.2. DISEÑO ESPECÍFICO DE LA SOLUCIÓN PROPUESTA
    - 3.2.1. *Base de Datos Transaccional*
    - 3.2.2. *ETL (Extract, Transform and Load)*
    - 3.2.3. *Data Warehouse*
    - 3.2.4. *Cubo OLAP*
    - 3.2.5. *Interfaz de Reportes*
  - 3.3. DIAGRAMA DE SECUENCIA
    - 3.3.1. *Validar usuario*
    - 3.3.2. *Gestionar Roles*
    - 3.3.3. *Gestionar Usuario*
    - 3.3.4. *Gestionar Reportes*
    - 3.3.5. *Gestionar Análisis*
    - 3.3.6. *Editar Contenido*
    - 3.3.7. *Compartir Contenido*
    - 3.3.8. *Gestionar tareas programas*
  - 3.4. REQUERIMIENTOS Y ESPECIFICACIÓN DE USUARIOS
    - 3.4.1. *Listado de Requerimientos.*
    - 3.4.2. *Requerimientos Técnicos*
    - 3.4.3. *Requerimientos Operativos*

- 3.4.4. *Requerimientos de Desarrollo*
  - 3.4.5. *Requerimientos No Funcionales*
  - 3.4.6. *Requerimientos No Funcionales Técnicos*
  - 3.4.7. *Descripción de los perfiles de Usuario*
  - 3.5. ITERACIÓN 1: CENTRO DE SOLUCIÓN DE CONTROVERSIAS
    - 3.5.1. *Requerimientos Funcionales Para Centro de Solución de Controversias*
    - 3.5.2. *Especificación de Requerimientos*
    - 3.5.3. *Modelado Iteración 1 Centro de Solución de Controversias*
    - 3.5.4. *Diagrama UML del DATAWAREHOUSE – CENTRO DE SOLUCION DE CONTROVERSIAS*
    - 3.5.5. *Diagrama del proceso ETL*
  - 3.6. ITERACIÓN 2: VERIFICACIÓN Y VIGILANCIA
    - 3.6.1. *Requerimientos Funcionales*
    - 3.6.2. *Especificación de Requerimientos*
    - 3.6.3. *Modelado Iteración 2 Modulo de Verificación y Vigilancia*
    - 3.6.4. *Diagrama UML del DATAWAREHOUSE – VERIFICACION Y VIGILANCIA*
    - 3.6.5. *Diagrama del proceso ETL*
  - 3.7. ITERACIÓN 3: TRIBUNAL SANCIONADOR
    - 3.7.1. *Requerimientos Funcionales*
    - 3.7.2. *Especificación De Requerimientos*
    - 3.7.3. *Modelado Iteración 3 Tribunal Sancionador*
    - 3.7.4. *Diagrama UML del DATAWAREHOUSE – TRIBUNAL SANCIONADOR*
  - 3.8. ITERACIÓN 4: ANFITRIÓN
    - 3.8.1. *Requerimientos Funcionales*
    - 3.8.2. *Especificación de Requerimientos*
    - 3.8.3. *Modelado Iteración 4: Anfitrión*
    - 3.8.4. *Diagrama UML del DATAWAREHOUSE – ANFITRION*
    - 3.8.5. *Diagrama del proceso ETL*
  - 3.9. ITERACION 5: CALL CENTER
    - 3.9.1. *Requerimientos Funcionales Para Call Center*
    - 3.9.2. *Especificación de Requerimientos*
    - 3.9.3. *Modelado Iteración 5 Call Center*
    - 3.9.4. *Diagrama UML del DATAWAREHOUSE – CALL CENTER*
    - 3.9.5. *Diagrama del proceso ETL*
  - 3.10. DISEÑO DE SALIDAS
  - 3.11. DISEÑO DE SEGURIDAD
    - 3.11.1. *Autorización*
    - 3.11.2. *Seguridad Usuario/Rol*
    - 3.11.3. *Seguridad por Action Sequence*
    - 3.11.4. *Seguridad en Mondrian*
    - 3.11.5. *Seguridad en reportes Ad Hoc*
  - 3.12. REFINAMIENTO DEL PROTOTIPO
4. CAPITULO IV: DOCUMENTACIÓN
- 4.1. MANUAL DE INSTALACIÓN

- 4.1.1. *Instalación y Configuración de Pentaho BI Server*
- 4.1.2. *Instalación y Configuración de Pentaho Metadata Editor*
- 4.1.3. *Instalación Y Configuración De Java SE Development Kit (Jdk)*
- 4.1.4. *Instalación de Mysql*
- 4.1.5. *Configuración de SQL Server*
- 4.1.6. *Configuración de Superficie de SQL Server 2005*
- 4.1.7. *Acceder a Configuración de Superficie de SQL Server 2005*
- 4.1.8. *Configuración de conexiones locales y remotas*
- 4.2. MANUAL TÉCNICO
  - 4.2.1. *Manual de Instalación*
  - 4.2.2. *Manual Schema Workbench*
  - 4.2.3. *Pentaho Metadata Editor*
  - 4.2.4. *Publicación en Pentaho BI Server*
  - 4.2.5. *Integration Services*
- 4.3. MANUAL DE USUARIO
  - 4.3.1. *Configuración Básica del Equipo de los Usuarios*
  - 4.3.2. *Navegación del Sistema*
  - 4.3.3. *Creación de Reportes*
  - 4.3.4. *Creación de Vista de Análisis*
- 5. CAPITULO V: PLAN DE IMPLANTACIÓN
  - 5.1. ELEGIR LA ARQUITECTURA
    - 5.1.1. *Recurso Tecnológico para la Implementación*
    - 5.1.2. *Recurso Humano*
    - 5.1.3. *Descripción de la Arquitectura del Sistema*
    - 5.1.4. *Herramientas para el Desarrollo del Proyecto*
  - 5.2. METODOLOGÍA DE LA IMPLANTACIÓN
  - 5.3. PLANIFICACIÓN DE ACTIVIDADES
    - 5.3.1. *Planeación de Implantación*
    - 5.3.2. *Organización Implantación*
    - 5.3.3. *Ejecución de la Implantación*
  - 5.4. CRONOGRAMA DE ACTIVIDADES
  - 5.5. SISTEMA DE CONTROL
    - 5.5.1. *Formularios para el Control*
    - 5.5.2. *Estrategias De Control*
- CONCLUSIONES
- RECOMENDACIONES
- BIBLIOGRAFÍA
- LIBROS
- TESIS
- SITIOS WEB
- GLOSARIO TÉCNICO

## **ANEXOS**

## 2.6. FACTIBILIDADES

### 2.6.1. Factibilidad Técnica

Después de definir la problemática presente y establecer las causas que ameritan el nuevo sistema, es pertinente realizar un estudio de factibilidad técnica la cual permita determinar los recursos tecnológicos necesarios para desarrollar el proyecto desde el punto de vista técnico. Se presentan los recursos técnicos con los que se cuenta actualmente y los requerimientos mínimos necesarios para el desarrollo e implementación del proyecto.

La factibilidad técnica consiste en realizar una evaluación de la tecnología existente en la organización, y verificar si cumple con la tecnología necesaria para el desarrollo e implantación de este proyecto, este estudio se enfocó en recolectar información de los componentes técnicos necesarios para el desarrollo y los que poseen en la organización.

De acuerdo a la tecnología necesaria para el desarrollo del “data warehouse” para el proceso de denuncias de la defensoría del consumidor, se evaluó bajo dos enfoques principales: Hardware y Software.

#### 2.6.1.1. Recurso tecnológico para el desarrollo

En este apartado se presenta el hardware y software que se considera necesario para desarrollar el proyecto, y que es imprescindible para tener éxito en el proceso de desarrollo. Es importante aclarar que los recursos que se utilizaran para desarrollar el proyecto son de la propiedad del grupo desarrollador.

En la tabla 2.1 se presenta el hardware que se posee para desarrollar el proyecto, incluyendo el equipo de red. En cuando al software, se utilizara aquel que se ajuste a las necesidades y al hardware que se posean, es por ello que se realizará una evaluación de alternativas de software en la cual se incluye el sistema operativo, servidor web, base de datos y lenguajes de programación dicha evaluación nos brindará las herramientas que la elección de las herramientas se apegue mejor al proyecto. Al tener las herramientas se podrá evaluar si el recurso tecnológico con el que cuenta la organización cumple con los requerimientos de este proyecto.

Equipos	Características
PC1	Procesador Intel Dual Core, velocidad de 1.8 GHZ, HD de 160 GB, RAM 2 GB.
PC2	Procesador Centrino Duo, velocidad 1.7 GHZ HD de 120 GB, RAM 2 GB
PC3	Procesador Turion X2, velocidad 1.6 Ghz, HD de 80 GB, RAM 2GB
PC4	Procesador Athlon 64 X2, velocidad 1.9 Ghz HD de 160 GB, RAM 2GB
PC5	Procesador Pentium 4 HT, velocidad 3.0 GHZ HD de 410 GB, RAM 2GB
Router	DLINK 5 Puertos
Cable UTP	50 metros categoría 5e
Conectores	50 RJ45

**Tabla 2.1: Tabla de especificación de Hardware para el desarrollo**

La metodología a seguir para las evaluaciones del software, será a través de criterios que como grupo se consideran primordiales, estos criterios fueron elegidos luego de una investigación bibliográfica acerca de las características de cada elemento las cuales se presentan en las tablas de evaluación, la siguiente tabla muestra la descripción según el peso que se utilizará en la evaluación de cada elemento:

Calificaciones	Pesos	Descripción
Excelente	5	Se considera que cumple con el objetivo del criterio evaluado completamente.
Muy Bueno	4	Cumple con los aspectos primordiales del criterio evaluado.
Bueno	3	Es considerado como aceptable, aunque no cumple con todas las consideraciones del criterio evaluado.
Regular	2	Presenta cierto rango de insatisfacción en el criterio evaluado.
Malo	1	No cumple con el criterio evaluado.

**Tabla 2.2: Calificaciones y pesos a utilizar para evaluar los criterios**

Para poder determinar los productos aceptables, bajo el acuerdo del grupo, deberán cumplir como mínimo 65% de los puntos totales de la evaluación de cada aspecto técnico.

**Sistemas Operativos**

Especificaciones	Windows Server 2003 <sup>13</sup>	Red Hat 9 <sup>14</sup>	Suse Enterprise 11 <sup>15</sup>
Procesador	133 MHz	400 MHz	500 MHz
RAM	256 MB	128 MB	512 MB
Espacio en Disco	2 GB – 3GB	500 MB	750 MB
Unidad Óptica	CD-ROM o DVD-ROM	CD-ROM o DVD-ROM	CD-ROM o DVD-ROM

**Tabla 2.3: Requisitos mínimos de instalación de los sistemas operativos**

En la tabla 2.4 se muestra la comparación entre los sistemas operativos, los cuales son evaluados en base a criterios, los cuales se detallan más adelante (Ver Anexo 5: Evaluaciones técnicas).

<sup>13</sup> <http://technet.microsoft.com/en-us/library/cc739866.aspx>

<sup>14</sup> <http://www.redhat.com/docs/manuals/linux/RHL-9-Manual/release-notes/x86/>

<sup>15</sup> <http://www.novell.com/products/server/techspecs.html>

<b>Criterios de Evaluación<sup>16</sup></b>	<b>Windows Server 2003</b>	<b>Red Hat Linux 9</b>	<b>Suse Enterprise 11</b>
<b>Sistema Operativo</b>			
Fiabilidad	4	3	3
Seguridad	3	5	5
Escalabilidad	3	4	4
Manejabilidad	5	3	3
Adaptabilidad	4	3	5
Asequibilidad	5	4	4
<b>TOTAL</b>	<b>24</b>	<b>22</b>	<b>24</b>

Tabla 2.4: Evaluación de los sistemas operativos

Para la evaluación del sistema operativo se ha tomado como referencia que los sistemas que cumplan con más de 20 puntos, pero debido a que todos cumplen con esta puntuación, se elijara el sistema que tenga mayor compatibilidad con el sistema actual.

**Servidores Web**

<b>Especificaciones</b>	<b>Microsoft Internet Information Services IIS version 6<sup>17</sup></b>	<b>Apache 2.0<sup>18</sup></b>	<b>Tomcat 6<sup>19</sup></b>
Sistema Operativo	Microsoft Windows XP, Microsoft Windows Server 2003	Windows, Unix (BSD, GNU Linux, etc), MacOS	Windows, Linux, Mac OS X, Solaris y FreeBSD
Espacio Requerido	40 MB	50 MB	110 MB
Memoria RAM	128 MB	128 MB	128 MB
Licencia	Propietaria	Gratuita	Gratuita
Software previo o adicional	Windows XP SP2 o Windows Server SP1		JRE 5 o superior

Tabla 2.5: Requisitos mínimos de instalación de los servidores

<b>Criterio</b>	<b>IIS 6</b>	<b>Apache 2</b>	<b>Apache Tomcat</b>
Disponibilidad en el mercado	3	4	4
Seguridad	3	4	5
Escalabilidad	4	3	3
Manejabilidad	3	4	4
Documentación actualizada	3	4	4
Uso equilibrado de recursos	2	4	5
Soporte técnico	3	4	4
Costo	3	5	5
<b>TOTAL</b>	<b>24</b>	<b>32</b>	<b>34</b>

Tabla 2.6: Evaluación de los servidores web

<sup>16</sup> Tomado de tesis "Comparison of Selection Criteria of Open Source DBMS Against Proprietary DBMS" en dirección <http://www.pafis.shh.fi/graduates/sohkha02.pdf>

<sup>17</sup> <http://www.microsoft.com/technet/prodtechnol/WindowsServer2003/Library/IIS/ad56540c-2323-4316-b981-7ebb70352baa.mspx?mfr=true>

<sup>18</sup> <http://httpd.apache.org/docs/2.0/es/install.html#requirements>

<sup>19</sup> <http://tomcat.apache.org/tomcat-6.0-doc/setup.html>

Por la cantidad de criterios evaluados para la elección del servidor web, se determino que para ser considerado una alternativa para este proyecto debe de cumplir con por lo menos 26 puntos en total. Debido a la puntuación reflejada por apache tomcat, se utilizara ese servidor para el desarrollo del proyecto.

**Sistemas Gestores de Bases de Datos**

<b>Criterios</b>	<b>MySql 5<sup>20</sup></b>	<b>Oracle 11g<sup>21</sup></b>	<b>PostgreSQL 8<sup>22</sup></b>	<b>Sql Server 2005<sup>23</sup></b>
Sistema Operativo	Linux, Solaris, Unix, Windows, FreeBSD, NetWare, Mac OS X	Linux, Solaris, Unix, Windows	Linux, Windows, FreeBSD, Max OS X, Unix, Solaris	Windows Server 2003 SP2
RAM	128 MB	1 GB	128 MB	512 MB
Espacio en Disco Duro	200 MB	3.47 GB	259 MB	280 MB

**Tabla 2.7: Requisitos mínimos de instalación de SGBD**

En la tabla 2.8 se detalla la comparación de diferentes gestores de base de datos, para lo que se han definido una serie de criterios cuyos elementos son detallados más adelante (Ver Anexo 5: Evaluaciones Técnicas)

<b>Criterios de Evaluación<sup>24</sup></b> <b>Base de Datos</b>	<b>MySQL 5</b>	<b>Oracle 11g</b>	<b>PostgreSQL 8</b>	<b>Sql Server 2005</b>
Costo	5	2	4	3
Rendimiento	4	5	4	5
Fiabilidad	3	5	3	4
Escalabilidad	3	5	3	5
Seguridad	3	5	3	5
Multiplataforma	5	5	5	3
Administración	3	5	2	3
Madurez	4	5	3	4
Documentación	4	4	3	5
Herramientas Integradas	2	5	1	5
<b>TOTAL</b>	<b>33</b>	<b>46</b>	<b>31</b>	<b>41</b>

**Tabla 2.8: Evaluación de los Gestores de Base de Datos**

<sup>20</sup> <http://dev.mysql.com/doc/refman/5.0/es/installing.html>  
<sup>21</sup> [http://download.oracle.com/docs/cd/B28359\\_01/install.111/b32002.pdf](http://download.oracle.com/docs/cd/B28359_01/install.111/b32002.pdf)  
<sup>22</sup> <http://www.postgresql.org/docs/8.3/static/index.html>  
<sup>23</sup> <http://msdn.microsoft.com/en-us/library/ms143506.aspx>  
<sup>24</sup> Tomado de tesis "Comparison of Selection Criteria of Open Source DBMS Against Proprietary DBMS" en dirección <http://www.pafis.shh.fi/graduates/sohkha02.pdf>

La evaluación del gestor de base de datos es un elemento imprescindible para este tipo de proyectos, por lo que se ha establecido que para tener consideración un sistema como alternativa aceptable tendrá que cumplir con una puntuación arriba de los 33 puntos. Para el caso los gestores Oracle y Sql Server cumplen con esta evaluación por lo que se elige SQL server por ser el gestor de base de datos con el que trabaja sistema transaccional y el costo de adquisición y migración de datos a otros gestor de base de datos sería muy elevado trayendo limitaciones y la necesidad de un mayor esfuerzo en el desarrollo del proyecto.

**Herramienta a Utilizar para Desarrollo de Data Warehouse**

Para la elección de la herramienta que se utilizará para modelar los datos y poder construir el data warehouse, se evaluaron dos alternativas las cuales se consideran herramientas de mucha aceptabilidad y compatibilidad con el sistema transaccional, dichas alternativas se evaluaron con la ayuda de una matriz de evaluación que es planteada por kimball para elegir las herramientas adecuadas para trabajar en proyectos de inteligencia de negocios, dichos criterios se han agrupado en cuatro áreas importantes las cuales se mencionan a continuación:

- Características de almacenamiento
- Control y programación de trabajos
- Estándares y Metadatos
- Características de los vendedores

Esta matriz se puede consultar en el apartado Anexo 5: Evaluaciones Técnicas El resultado de dicha evaluación permitió elegir a Sql server BI suite que se encuentra integrada por SQL Server Integration Services (SSIS), SQL Server 2005 Analysis Services (SSAS). Dichas características se presentan a continuación:

<b>Especificaciones</b>	<b>Pentaho BI suite<sup>25</sup></b>	<b>Sql Server Bussines Intelligence<sup>26</sup></b>
Sistema Operativo	Microsoft Windows Xp,windows 2003 server, Linux	Windows Server 2003, Windows XP,Windows 2000
Procesador	1.8 gigahertz	1 gigahertz
Espacio requerido en Disco duro	400 MB	390
Memoria Ram	2 GB	1 GB
Licencia	Gratuita	Propietaria

**Tabla 2.9: Requisitos mínimos de instalación de Herramientas para desarrollo del proyecto**

<sup>25</sup> <http://softlayer.dl.sourceforge.net/sourceforge/pentaho/2.0.0-RC1-readme.htm#Requirements>  
<sup>26</sup> <http://www.microsoft.com/spain/sql/productinfo/sysreqs/default.msp>

<b>Criterios</b>	<b>Pentaho BI suit</b>	<b>Sql Server Bussines Intelligence</b>
Disponibilidad en el mercado	2	5
Instalación y uso manejable	4	4
Documentación actualizada	4	4
Soporte en línea	2	4
Consumo de recursos	2	4
Seguridad	3	3
Precio	5	2
<b>Puntaje</b>	<b>22</b>	<b>26</b>

**Tabla 2.10: Evaluación de la herramienta para el desarrollo**

El puntaje mínimo que debe de cumplir la herramienta de inteligencia de negocios es de 23, lo cual nos brinda como herramienta de desarrollo a Sql server Bussines.

**Antivirus**

<b>Especificaciones</b>	<b>Symantec Protection Suite Enterprise Edition<sup>27</sup></b>	<b>Mcafee VirusScan 8.7<sup>28</sup></b>	<b>Kaspersky Internet Security 2009<sup>29</sup></b>
Sistema Operativo	Microsoft Windows 2008, Server 2003,xp	Microsoft Windows 2008, Server 2003,xp	Microsoft Windows 2008, Server 2003,xp
Procesador	Intel Pentium or compatible	Intel Pentium or compatible	Intel Pentium processor 800 MHz o mas
Espacio requerido en Disco duro	2 GB	150 MB	400 MB
Memoria Ram	1 GB	1 GB	1 GB
Licencia	Propietaria	Propietaria	Propietaria

**Tabla 2.11: especificaciones de los antivirus**

<b>Criterios</b>	<b>Norton Internet Security 2009</b>	<b>Mcafee VirusScan 8.7</b>	<b>Kaspersky Internet Security 2009</b>
Disponibilidad en el mercado	4	4	4
Instalación y uso manejable	2	3	3
Documentación actualizada	4	4	4

<sup>27</sup> [http://www.symantec.com/business/products/sysreq.jsp?pcid=pcat\\_security&pvid=prot\\_suite\\_ee\\_1](http://www.symantec.com/business/products/sysreq.jsp?pcid=pcat_security&pvid=prot_suite_ee_1)

<sup>28</sup> [http://www.mcafee.com/us/enterprise/downloads/beta/beta\\_mcafee/vse87i/vse87i.html](http://www.mcafee.com/us/enterprise/downloads/beta/beta_mcafee/vse87i/vse87i.html)

<sup>29</sup> [http://www.kaspersky.com/koss-1\\_system\\_requirements#comp1](http://www.kaspersky.com/koss-1_system_requirements#comp1)

Soporte en línea	4	4	4
Consumo de recursos	1	3	2
Seguridad	3	4	4
Precio	2	4	3
<b>Puntaje</b>	<b>20</b>	<b>26</b>	<b>24</b>

**Tabla 2.12: Evaluación de los antivirus**

La seguridad que debe de contar el sistema debido a la confidencialidad de los datos es de suma importancia por lo que el puntaje mínimo para esta evaluación será de 23, dejando como resultado ganador a McAfee VirusScan 8.7, debido a que obtuvo un mejor puntaje.

**Herramienta para la Interfaz de Usuarios**

La elección de la herramienta que será la interfaz de cara al usuario es de vital importancia, ya que por medio de esta el usuario podrá comprobar la funcionalidad del sistema a desarrollar, por lo que se decidió evaluar 2 herramientas de software libre, y la otra que es la que viene integrada con SQL server. A continuación se muestran sus características y los resultados de la evaluación.

<b>Especificaciones</b>	<b>Pentaho Reports<sup>30</sup></b>	<b>Reporting Services<sup>31</sup></b>	<b>OpenI<sup>32</sup></b>
Sistema Operativo	Independiente	Microsoft Windows 2003, 2000, XP profesional	Independiente
Procesador	Dual-core	Intel Pentium o compatible 500 MHz	Intel Pentium procesador 500 MHz o mas
Espacio requerido en Disco duro	1 GB	150 MB	130 MB
Memoria RAM	2 GB	1 GB	1 GB
Licencia	Gratuita	Propietaria	Gratuita

**Tabla 2.13: especificaciones de la interfaz de usuario**

<b>Criterios</b>	<b>Pentaho Reports</b>	<b>Reporting Services</b>	<b>OpenI</b>
Disponibilidad en el mercado	3	5	3
Instalación y uso manejable	4	3	3

<sup>30</sup> [http://www.abdn.ac.uk/~csc228/teaching/CS5551/practicals/community\\_user\\_guide.pdf](http://www.abdn.ac.uk/~csc228/teaching/CS5551/practicals/community_user_guide.pdf)

<sup>31</sup> <http://www.microsoft.com/sql/prodinfo/previousversions/rs/system-requirements.mspx>

<sup>32</sup> [http://www.openi.org/docs/INSTALL.html#Detailed\\_System\\_Requirements](http://www.openi.org/docs/INSTALL.html#Detailed_System_Requirements)

Documentación actualizada	4	4	3
Soporte en línea	4	4	4
Consumo de recursos	2	2	4
Seguridad	4	3	3
Precio	5	2	5
<b>Puntaje</b>	<b>26</b>	<b>23</b>	<b>25</b>

**Tabla 2.14: Evaluación de la interfaz de usuario**

Un objetivo del proyecto es realizar la interfaz que utilizará el usuario implementando herramientas de código libre. La puntuación mínima requerida que debe de cumplir para considerarse como aceptable es de 23 puntos en total. Por lo que se eligió a pentaho reports como la herramienta que ayudara a diseñar la capa de presentación (lo que el usuario visualizará).

**Software Utilitario**

Además del software especificado anteriormente, se ha considerado utilizar el siguiente software utilitario el cual brindara apoyo a la fase de desarrollo del proyecto. En su mayoría se utilizan programas gratuitos o ya se poseen, y que cumplen con las especificaciones de hardware y software evaluado.

<b>Utilitarios</b>	<b>Descripción</b>	<b>Espacio requerido en disco (MB)</b>
Nero Express 6.0	Programa para grabar CDs y DVDs.	150
Microsoft Office 2003	Paquete de oficina de usos múltiples (procesador de texto, hoja electrónica, etc.)	300
Acrobat Reader 9.0	Programa para leer archivos PDF.	150
Power Designer	Programa para modelar diagramas de bases de datos de diferentes gestores	200
Mozilla firefox 3.0	Programa que permite navegar en internet	30
<b>Total de espacio en disco</b>		<b>830</b>

**Tabla 2.15: Software utilitario disponible para desarrollo e implementación del proyecto**

**2.6.1.2. Recurso Tecnológico para La Implementación**

La organización dispone de un completo recurso tecnológico que permita la implementación de dicho proyecto, dicha institución puso a disponibilidad toda la arquitectura cliente servidor que se describe a continuación.

**Hardware**

El hardware para llevar a cabo la instalación y configuración del sistema es el siguiente.

Elemento	Servidor	Usuario 1	Usuario 2
RAM	1 GB	512 MB	1.5 GB
Discos duros	175 GB	200 GB	200 GB
Procesador	Xeon 3.17 Ghz	Pentium IV 3.06 Ghz	Pentium IV 3.06 Ghz

**Tabla 2.16:- Descripción de HW para la implementación**

**Software**

El software disponible para la instalación, configuración y administración del sistema propuesto es el siguiente:

Elemento	Servidor	Usuario 1	Usuario 2
Sistemas Operativos	Windows 2003 server	Windows Xp	Windows Xp
Bases de datos	SQL Server 2005 Enterprise		SQL Server 2005 express
Antivirus	Mcafee VirusScan 8.7	Mcafee VisusScan 8.7	Mcafee VisusScan 8.7
herramientas de desarrollo	Reporting Services		Visual Studio .net 2005 Crystal Reports 10
Herramientas de Ofimática		Microsoft office 2003	Microsoft office 2003
Otras Herramientas		Navegadores web:Mozilla firefox, Internet Explorer 6	Navegadores web:Mozilla firefox, Internet Explorer 6

**Tabla 2.17: Descripción del SW Actual**

El motivo de mostrar dos tipos de usuarios es debido a que se encuentran dos tipos de configuraciones en las maquinas distribuidas en la organización.

**2.6.1.3. Conclusión de la Factibilidad Técnica**

Desde el punto de vista técnico el proyecto es factible de ser desarrollado e implantado, ya que para el desarrollo se poseen los recursos tecnológicos requeridos, y con respecto al software utilizado es compatible totalmente con el sistema transaccional actual, lo cual facilitará la parte de extracción de datos; en cuanto a la implantación se cuenta con el hardware, software y conocimiento técnico por parte de los usuarios actuales de la organización por lo que permite la implementación de la solución sea factible.

### 2.6.2. Factibilidad Operativa

Para el desarrollo del proyecto y creación de un sistema exitoso, es necesario que exista personal capacitado para llevar a cabo el proyecto y así mismo, deben existir usuarios finales dispuestos a emplear el producto generado.

Para determinar la factibilidad del desarrollo de un data warehouse al proceso de denuncias para la Defensoría del Consumidor se debe que considerar los siguientes elementos:

- El apoyo institucional al desarrollo del proyecto.
- Aceptación del sistema por parte de los usuarios.
- Personal para la administración y utilización de la solución.
- Aspectos técnicos necesarios en los equipos.

#### 2.6.2.1. Apoyo institucional del proyecto.

Las autoridades de la Defensoría del Consumidor así como la Gerencia de Sistemas de Información se comprometen a facilitar la información que sea necesaria en cada etapa para llevar a cabo el proyecto. Dicho apoyo institucional se hace constar en una carta de compromiso por parte de la máxima autoridad.

#### 2.6.2.2. Aceptación de usuarios finales

El éxito del proyecto en su etapa de culminación depende en gran parte de del grado de aceptación por parte de los usuarios finales, para lo cual se realizo una encuesta con el fin de medir el grado de necesidad del proyecto y de la aceptación por parte de los usuarios finales.

El 100% de los entrevistados afirman estar de acuerdo en la necesidad de poseer un sistema de información que elimine la dependencia con la gerencia de sistemas de Información, esto al momento de obtener algún reporte o información necesaria. En la figura 2.10 se puede ver el grafico correspondiente a la pregunta, otros resultados de la encuestan pueden ser observado en el apartado Anexo 4: Resultados de las encuestas.

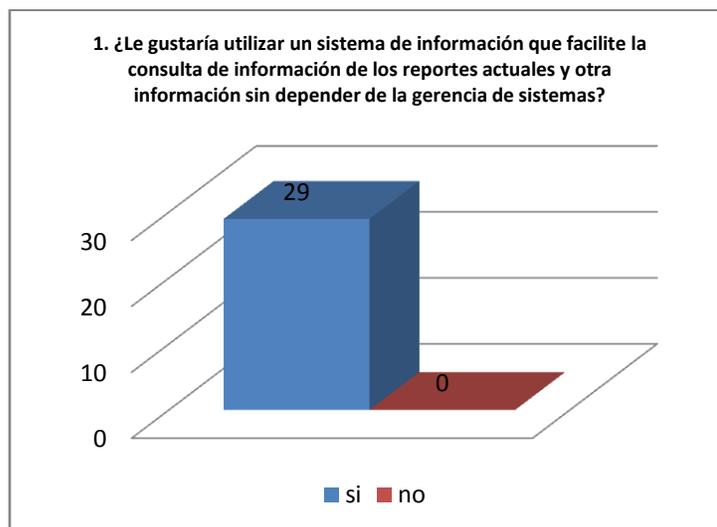
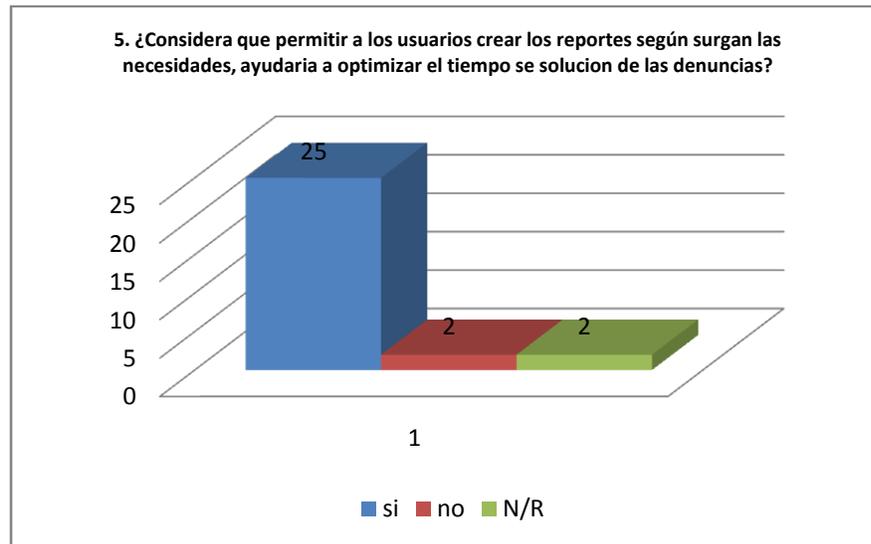


Figura 2.10: Aceptación del sistema por parte del usuario

Reforzando la necesidad un sistemas de información, se pudo conocer en la encuesta realizada, que de 29 personas encuestadas, 25 contestaron que la implementación de un nuevo sistemas que permita a

los usuarios crear reportes según surjan las necesidad, ayudara a optimizar el tiempo de solución de las denuncias, esto se refleja en la figura 2.11.



**Figura 2.11: Aceptación del sistema por parte del usuario**

Con lo anteriormente mencionado podemos concluir que el usuario final tiene diversas necesidades de información que no son cubiertas por los actuales sistemas, lo que genera ambiente de aceptación para la implementación de un sistema que solvete dichas necesidades. Cabe mencionar que para el sistema SARA, se realizó una implementación con un alto grado de resistencia al cambio en los procesos, mas sin embargo actualmente SARA es el sistemas básico para el funcionamiento de la Defensoría del Consumidor, lo que nos da un campo más optimista para el desarrollo del proyecto.

**2.6.2.3. Personal para la administración y utilización de la solución.**

Para la correcta implementación del sistema es necesario contar con el personal necesario, que cumpla con los requisitos en conocimiento para la administración y operación del sistema. Por lo anterior, se ha definido dos tipos de perfil; para el personal a administrar el sistema y el personal que operará con el sistema.

Administrador del sistema:

- Conocimiento de programación en Lenguaje SQL.
- Dominio y comprensión de modelado multidimensional
- Dominio de inglés técnico.
- Desarrollo de procesos ETL.
- Manejo de Herramientas ETL.
- Capacidad de supervisión, manejo de grupos técnicos y trabajo en equipo.
- Instalación y configuración de la herramienta Pentaho.

Operador del sistema:

- Conocimiento de Windows XP profesional (SP1 o SP2) o Vista Enterprise.
- Conocimiento en la utilización de navegador Web, puede ser;
- Internet Explorer (versión 6 o superior).
- Mozilla FireFox (versión 2.0 o superior).
- Capacitación en los diferentes módulos y funciones del sistema a desarrollar.

**2.6.2.4. VOLÚMENES DE DATOS**

El análisis de Factibilidad Operativa incluye la estimación de dimensión de la cantidad de información que se manipulara y almacenara en el sistema transaccional, desde el cual se extraerá hacia el Data Warehouse, así determinar de una forma aproximada el espacio en disco duro que necesitara para operar en los siguientes 10 años, que se establece como vida útil de un sistema (según informe de Tetsuo Tamai y Yohsuke Torimitsu)<sup>33</sup>.

La tabla padre del modulo “Centro de Solución de Controversias CSC”, modulo principal de SARA, se medirá los registros en base a los tipos de datos almacenados que se muestran en la interfaz de ingreso de Denuncias (ver Anexo 6: Interfaz de Ingreso de datos)

A continuación se presenta los tipos de datos almacenados y su tamaño en bytes:

<b>Campo</b>	<b>Tipo de datos</b>	<b>Cantidad máxima de caracteres</b>	<b>Bytes</b>
Nombre del consumidor	Varchar	30	30
Apellido del consumidor	Varchar	30	30
Sexo	Char	1	1
Numero de DUI	Varchar	15	15
Fecha de nacimiento	Date		8
Profesión	Varchar	30	30
Dirección	Varchar	250	250
Municipio	Varchar	30	30
Departamento	Varchar	30	30
Email	Varchar	30	30
Teléfono	Varchar	15	15
Fax	Varchar	15	15
Celular	Varchar	15	15
Numero de comprobante	Varchar	30	30
Sector proveedor	Varchar	30	30
Proveedor	Varchar	30	30
Motivo de denuncia	Varchar	30	30
Monto reclamado	Money		8
Número de cuenta con el proveedor	Varchar	30	30
Pregunta abierta 1	Varchar	500	500
Pregunta abierta 2	Varchar	500	500
Pregunta abierta 3	Varchar	500	500
Pregunta abierta 4	Varchar	500	500
Pregunta abierta 5	Varchar	500	500
Pregunta abierta 6	Varchar	500	500
Documento que aporta	Varchar	30	30
Otros documentos	Varchar	500	500
<b>TOTAL</b>		<b>4171</b>	<b>4187</b>

**Tabla 2.18: Tamaños del los campos del formulario de ingreso**

De acuerdo a los datos proporcionados por la Defensoría del Consumidor, el espacio en disco de la base de datos, en crecimiento del 2008 al 2009 fue de 2 GB aproximadamente, bajo sus proyecciones, tiene previsto crecer durante el 2009 4 GB.

<sup>33</sup> Tetsuo Tamai, Software Lifetime and its Evolution Process over Generations; 1998

De lo anterior, el modulo principal de SARA “Centro de Solución de Controversias CSC”, posee el registro principal, en el cual se almacena la mayoría de información de los casos, para el cual actualmente existen 21,000 registros.

Actualmente el porcentaje de espacio utilizado por las diferentes tablas dentro de la base de datos transaccional se muestra en la tabla 2.19, siendo el 75%, ya que no se ha tomado en cuenta tablas de bitácoras y campos de auditoría.

Modulo	Porcentaje en Disco duro (%)
Modulo de Centros de Solución de Controversias (CSC)	30
Modulo de Call Center	10
Modulo de Aseguramiento de Calidad	10
Modulo de Educación	5
Modulo de Tribunal Sancionador	20

**Tabla 2.19: Porcentaje de espacio utilizado según modulo**

En base a los datos presentados anteriormente, se muestra a continuación la proyección a 10 años:  
 Formula:

$$\text{Crecimiento Porcentual} = \text{Espacio en Disco} ( 1 + \text{porcentaje crecimiento} ) ^n$$

Donde:

- Crecimiento porcentual: El valor proyectado a crecer en el “n” año.
- Espacio en Disco: El espacio actual ocupado.
- Porcentaje crecimiento: El porcentaje estimado a crecer en cada año.
- n: Año a proyectar

Calculo para:

$$E = (10 * \text{porcentaje en disco del modulo}) \text{ GB}$$

$$P = 20 \%$$

$$n = 10 \text{ años}$$

Modulo	C = E * ( 1 + P ) ^n
Modulo de Centros de Solución de Controversias (CSC)	18.57520927 GB
Modulo de Call Center	6.191736422 GB
Modulo de Aseguramiento de Calidad	6.191736422 GB
Modulo de Educación	3.095868211 GB
Modulo de Tribunal Sancionador	9.287604634 GB
<b>TOTAL</b>	<b>43.342155 GB</b>

**Tabla 2.20: Porcentaje de crecimiento según modulo**

En base a la tabla 2.20., se espera que el modulo “Centro de Solución de Controversias CSC”, tenga en 10 años:

Espacio en disco por registro: 4187 Bytes

$$\text{Registros dentro de 10 años:} = \frac{43.342155 \text{ GB}}{4,187 \text{ Bytes}} = \frac{(1,073,741,824 \frac{\text{Bytes}}{1\text{GB}}) * (43.342155 \text{ GB})}{4,187 \text{ Bytes}}$$

$$= 4,763,548.86 \text{ registros}$$

Actualmente existen 6 servidores disponibles, de los cuales ha sido brindado uno para ambiente de desarrollo y posterior en producción. El espacio actual disponible en disco es:

Unidad C: 100.0 GB Disponible: 100.0 GB

Unidad D: 900.0 GB Disponible: 900.0 GB

#### **2.6.2.5. Conclusión factibilidad Operativa**

La creación de un Data Warehouse al proceso de las denuncias para la Defensoría del Consumidor traería beneficios tangibles e intangibles, cuantificables y cualitativos, sin embargo es necesario considerar la aceptación de los usuarios finales, el apoyo de la contraparte, y más importante si existe un entorno tecnológico en el cual se pueda echar a andar dicho proyecto, mediante el cual se logre tanto el desarrollo como la implantación del mismo. Por lo anterior se puede afirmar que el proyecto es operativamente factible para ser realizado.

### 2.6.3. Factibilidad Económica

El objetivo de la factibilidad económica es determinar si es conveniente para la organización el desarrollo del proyecto en términos monetarios. El estudio de factibilidad se realiza evaluando diferentes alternativas para saber cual es económicamente factible. Las alternativas a evaluar son las siguientes:

- **Alternativa A:** Continuar con el sistema actual para el desarrollo de reportes.
- **Alternativa B:** Desarrollo del data warehouse con las interfaces para obtención de los reportes.

#### 2.6.3.1. Alternativa A

Los costos que presenta el sistema actual para la creación de reportes en la Unidad Informática de la Defensoría del Consumidor se han tomado de acuerdo a las actividades que realiza esta unidad para la creación de un nuevo reporte. Estos costos se han calculado basándose en el tiempo que toma la realización de una actividad, el número de personas involucradas en la realización de esta, el promedio de salario de estas personas y la frecuencia con la que se realizan anualmente.

Para el cálculo de estos costos se consideran los salarios<sup>34</sup> de la tabla 2.21

Puesto	Salario Promedio Mensual	Salario Promedio por Hora
Analista Programador	\$ 875	\$ 5.47
Administrador de Sistemas	\$ 1100	\$ 6.88
Usuario Técnico	\$ 600	\$ 3.75

**Tabla 2.21: Salario promedio por hora tomando en cuenta que el día laboral es de 8 horas**

En la tabla 2.22 se presentan los costos de operación que se tienen con el sistema actual:

Actividad	Persona Involucrada	Cantidad de Horas	Frecuencia Promedio Mensual	Costo Mensual	Costo Total
Solicitud y Revisión de reporte	Usuario Técnico	1	20	\$ 75.00	\$900
Análisis del Requerimiento	Administrador de Sistemas	1	20	\$ 137.50	\$ 1650
Desarrollo y pruebas	Analista Programador	4	20	\$ 437.50	\$ 5250
<b>TOTAL</b>					<b>\$ 7800</b>

**Tabla 2.22: Costos de operación del sistema actual**

#### 2.6.3.2. Alternativa B

Los costos que se tendrán con el desarrollo del data warehouse y de las interfaces son los siguientes:

##### Costos de Operación

Son los costos que tendrá la administración del sistema al ser implementado. Para el cálculo de este costo se han considerado los siguientes criterios:

- La organización no contratara a una persona específica para realizar esta tarea. Un administrador de sistemas será el encargado de brindarle el mantenimiento necesario a la aplicación.

<sup>34</sup> Sueldos proporcionados por la Defensoría del Consumidor

- La persona encargada le dedicara una hora diaria para la administración y mantenimiento del sistema.
- Tomando el salario promedio por hora de la tabla 2.21 para el administrador de sistemas y tomando en cuenta el criterio anterior se tiene:
- 1 hora x 20 días al mes = 20 horas mensuales.  
20 horas mensuales x \$ 6.88 /hora = \$ 137.60 mensuales.  
\$137.60 mensuales x 12 meses = **\$ 1651.20 anuales.**

**Costos de Desarrollo**

Los costos en los que se incurren para el desarrollo del proyecto ascienden a un total de **\$ 36027.95** en concepto de consumo de recursos, tal como se detalla en la sección 18: Planeación de Recursos.

**Costos de Capacitación**

Son los costos en los que incurrirá la organización para capacitar al personal en el manejo y administración del sistema. Estos costos son detallados en la Tabla 2.23<sup>35</sup>

Cantidad de personas	Horas de Capacitación	Costo por Hora	Total
10	16	25	\$ 400

**Tabla 2.23: Costos de capacitación de alternativa B**

En la tabla 17.24 se resume el costo total de la alternativa B durante el primer año.

Costos	Total Anual
Costos de Operación	\$ 1651.20
Costos de Desarrollo	\$ 34390.31
Costos de Capacitación	\$ 400.00
<b>Total</b>	<b>\$ 34790.31</b>

**Tabla 2.24: Resumen de costos de alternativa B**

Adicionalmente la implantación del proyecto proporcionara beneficios intangibles, los cuales son detallados a continuación:

**Beneficios Intangibles**

Son todos aquellos beneficios que se tendrán en la Defensoría del Consumidor y que son muy difíciles de cuantificar, pero que son de una gran importancia para el proyecto. Dentro de estos tenemos:

- Mejora en la calidad de la información que es presentada a los usuarios y personas encargadas de la toma de decisiones.
- Satisfacción de los usuarios del sistema ya que se puede acceder fácilmente a la información y generar los reportes necesarios con mayor facilidad.
- Mejor atención a los consumidores ya que se les podrá proporcionar información más detallada y exacta.

<sup>35</sup> Precios basados en datos de capacitaciones en <http://www.pentaho.com/services/training/#onlinecourses>

**2.6.3.3. EVALUACIÓN DE ALTERNATIVAS**

**Análisis de Valor Presente**

El método de valor presente (VP) de evaluación de alternativas es muy popular debido a que los gastos o los ingresos futuros se transforman en dólares equivalentes de ahora. Es decir, todos los flujos futuros de efectivo asociados con una alternativa se convierten en dólares presentes. En esta forma, es muy fácil, aun para una persona que no está familiarizada con el análisis económico, ver la ventaja económica de una alternativa sobre otra.<sup>36</sup>

Este análisis sirve de apoyo para dar a conocer a los tomadores de decisiones de las instituciones, el valor de la inversión en el tiempo. El valor presente es la manera de valorar todos los recursos económicos gastados y generados a lo largo de la vida útil del Sistema y de comparar los costos actuales con los costos futuros, y beneficios actuales con los beneficios futuros.

El valor presente es el costo o beneficio medido a valor actual del efectivo y depende del costo del dinero, este es el costo de oportunidad o la tasa que podría obtenerse si el capital comprometido en el sistema propuesto se invirtiera en otro proyecto que sea relativamente seguro.

El valor presente a una tasa de interés *i*, se determina con el siguiente cociente:

$$P = S \frac{1}{(1 + i)^n}$$

Donde:

*n*= numero de periodos.

*i*= 3.3 %. = tasa de inflación anual proyectada para el 2009 según el Banco Central de Reserva de El Salvador.

*S*= Costos de Operación.

En la tabla 2.25 se muestra el cálculo del valor presente para las alternativas.

<b>Año</b>	<b>Valor Presente Alternativa A</b>	<b>Valor Presente Alternativa B</b>
0	\$7,800.00	\$34,790.31
1	\$7,550.82	\$1,598.45
2	\$7,309.61	\$1,547.39
3	\$7,076.09	\$1,497.95
4	\$6,850.04	\$1,450.10
5	\$6,631.21	\$1,403.78
<b>TOTAL</b>	<b>\$43,217.78</b>	<b>\$42,287.98</b>

**Tabla 2.25: Cálculo del Valor Presente para las alternativas**

En la figura 2.12 se muestra la comparación de los costos del valor presente, en el cual se puede observar que los costos del sistema actual son más elevados que los que presentaría el sistema propuesto en el tiempo.

<sup>36</sup> Ingeniería Económica, Leland T.Blank & Anhony J. Tarquin,Cuarta Edicion, McGraw Hill.

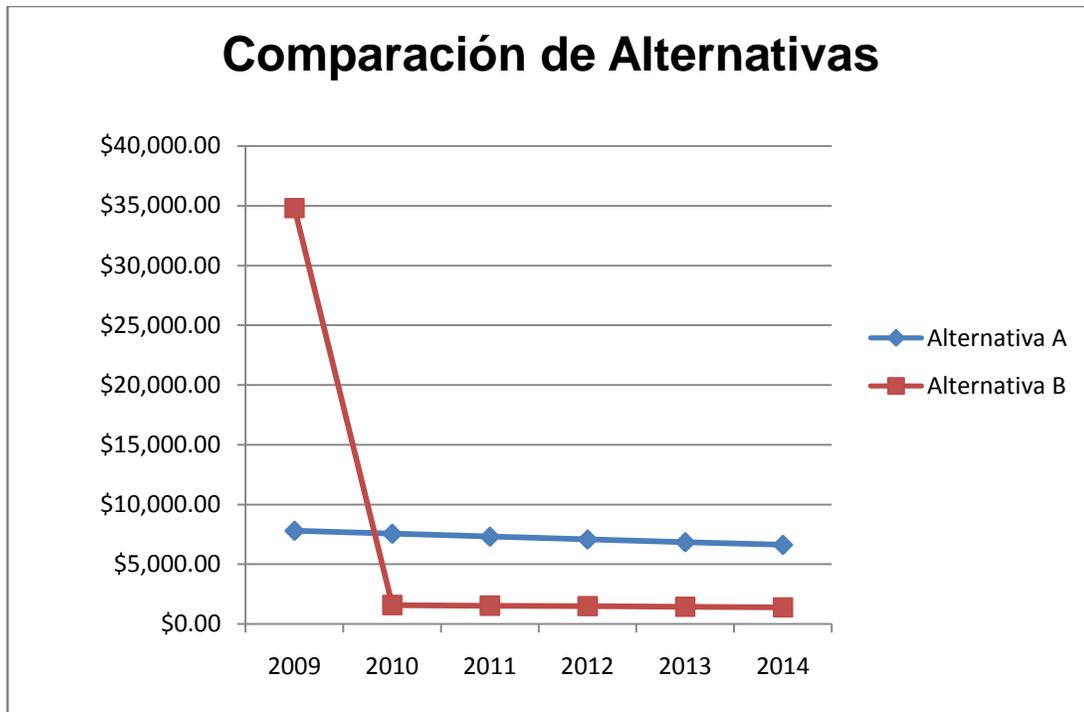


Figura 2.12: Comparación de alternativas en el tiempo

**2.6.3.4. Conclusión de la Factibilidad Económica**

Partiendo de los resultados obtenidos con la técnica del valor presente, se puede concluir que el desarrollo del proyecto es factible económicamente ya que su valor presente es de \$ 42,287.98 lo que representa un ahorro de \$ 929.80 en los primeros 5 años de uso del sistema propuesto y también es necesario resaltar los beneficios intangibles con los cuales se vería favorecida la organización.

## 2.7. PLANEACIÓN DE RECURSOS

### 2.7.1. Calculo para el Presupuesto

#### 2.7.1.1. Recurso Humano

La determinación del costo del recurso humano a utilizar en el desarrollo del proyecto se realiza por medio de un estudio de mercado de salarios para un recién graduado, así como también de salarios ya conocidos por los integrantes del grupo de trabajo.

#### **Cálculo del Salario de Equipo de Desarrollo (Analista/Programador)**

El Equipo de Desarrollo trabajará 4 horas diarias, 6 días a la semana, haciendo un total de 24 horas a la semana y 24 días al mes.

El salario promedio es en base al salario de 4 fuentes diferentes, tomando en cuenta un estudio de mercado en puestos similares. Ver tabla 2.26

Fuentes	Monto
Libro: Gerencia Informática <sup>37</sup>	\$ 605.00
www.computrabajo.com	\$ 500.00
Consulta a Analistas Programadores.	\$ 650.00
<b>Salario Promedio</b>	<b>\$ 585.00</b>

Tabla 2.26: Salario Promedio mensual de Equipo de Desarrollo

El salario mensual para las cuatro Analistas/programadores será de **\$2,540.00**, los cálculos se muestran en la Tabla 2.27

Salario	Monto
Salario Mensual por un Analistas/Programador	\$ 585.00
<b>Salario Mensual de los 4 Analistas/Programadores</b>	<b>\$ 2,340.00</b>

Tabla 2.27: Salario mensual de Equipo de Desarrollo

#### **Cálculo del Salario de Persona encargada de Soporte al Desarrollo**

Para el cálculo del salario del administrador del proyecto, en este caso el docente director, se tomo en cuenta el salario promedio de un Ingeniero Informático con 3 años de experiencia. Los datos se muestran en Tabla 2.28.

Fuentes	Monto
Libro: Gerencia Informática	\$ 934.00
Consulta a Analistas Programadores.	\$ 1000.00
<b>Salario Promedio</b>	<b>\$ 967.00</b>

Tabla 2.28: Salario mensual del Administrador de Proyecto

Partiendo de los datos obtenidos en las tablas anteriores se calcula el costo del recurso humano implicado en el desarrollo del proyecto.

<sup>37</sup> Gerencia Informática, Carlos Ernesto García, San Salvador, El Salvador. 2008

Recurso	Tiempo en Meses	Costo Mensual	Total
Equipo de desarrollo	8	\$ 2340	\$ 18720
Administrador de Proyecto	8	\$ 967	\$ 7736
<b>TOTAL</b>			<b>\$ 26456</b>

Tabla 2.29: Costo Total en Recurso Humano

### 2.7.1.2. Recursos Tecnológicos

A continuación se detalla todo el recurso tecnológico que será utilizado por el equipo de trabajo como computadoras, impresor y routers.

#### Costo de Computadoras

La Tabla 2.30 contiene los costos de las computadoras a utilizar por los 4 integrantes del grupo de desarrollo del proyecto.

Equipo	Características	Precio \$
PC1	Procesador Intel Dual Core Velocidad de 1.8 GHZ HD de 160 GB RAM 2 GB	\$ 450
PC2	Procesador Centrino Duo Velocidad 1.7 GHZ HD de 120 GB RAM 2 GB	\$ 700
PC3	Procesador Turion X2 Velocidad 1.6 Ghz HD de 80 GB RAM 2GB	\$ 700
PC4	Procesador Athlon 64 X2 Velocidad 1.9 Ghz HD de 160 GB RAM 2GB	\$ 750
PC5	Procesador Pentium 4 HT Velocidad 3.0 GHZ HD de 410 GB RAM 2GB	\$ 300
<b>TOTAL</b>		<b>\$ 2900</b>

Tabla 2.30: Costo de las computadoras a utilizar en el proyecto

#### Costo de Red Local

El equipo de red que se utilizara en el desarrollo del proyecto se detalla en la Tabla 2.31, en la cual se incluyen los costos de cada uno de los componentes.

Equipo	Cantidad	Precio Unitario	Total
Router DLINK 5 Puertos	1	\$ 50.00	\$ 50
Cable UTP categoría 5e	50 metros	\$ 0.45	\$ 22.5
Conectores RJ45	50	\$ 0.04	\$ 2.00
<b>TOTAL</b>			<b>\$ 74.5</b>

Tabla 2.31: Detalle del equipo de red local con sus respectivos costos

Otro componente importantísimo dentro de la red es el Internet, el cual es necesario para buscar información, enviar correos electrónicos, video conferencias, etc. La tabla 2.32 muestra los costos del servicio contratado.

Servicio Contratado	Costo mensual	Tiempo	Total
Internet Integra 1 Mbps	\$39.35	8 Meses	\$314.8
<b>TOTAL</b>			<b>\$314.8</b>

Tabla 2.32: Costo del servicio de Internet en el desarrollo del proyecto

Sumando los totales del costo del equipo de Red y el de servicio de Internet se tiene:

Costo equipo de Red	\$ 74.5
Costo de servicio de Internet	\$ 314.8
<b>Total</b>	<b>\$ 389.3</b>

Tabla 2.33: Total de costos de red local

### Costo de Periféricos

En este apartado se detallan otros periféricos que serán utilizados para la realización del proyecto. Los componentes son detallados en la tabla 2.34.

Periférico	Cantidad	Precio
Impresor Canon IP2500	1	\$ 39
Impresor Multifuncional Canon Mp140	1	\$ 56
<b>Total</b>		<b>\$ 95</b>

Tabla 2.34: Detalle de costos a incurrir en otros periféricos

### Costo de Medios de Almacenamiento

Se determino que se utilizaran medios de almacenamiento tales como cd y dvd, en los cuales se grabará el software necesario para el proyecto, los entregables a la Universidad y copias de respaldo. Los costos se detallan en la Tabla 2.35.

Medio de Almacenamiento	Cantidad	Precio Unitario	Total
Torre de 50 CDs	1	\$ 9.00	\$ 9.00
Torre de 50 DVDs	1	\$18.00	\$18.00
<b>Total</b>			<b>\$ 27</b>

Tabla 2.35: Costo de medios de almacenamiento externo a utilizar

El gasto total a incurrir en recurso tecnológico durante el desarrollo del proyecto es de **\$3411.3** el cual se detalla en la Tabla 2.36

Recurso	Total
Computadoras	\$ 2900
Red Local	\$ 389.3
Periféricos	\$ 95
Medios de Almacenamiento	\$ 27
<b>Total</b>	<b>\$ 3411.3</b>

Tabla 2.36: Costo total en recursos tecnológicos

### 2.7.1.3. Recursos de Operación

El equipo de desarrollo cuenta con un local dedicado exclusivamente al desarrollo del proyecto. Los gastos en los que se incurre mensualmente son: energía eléctrica, agua, alquiler y teléfono.

La tabla 2.37 muestra los gastos que se incurrirán en concepto de energía eléctrica, basándose en las tarifas publicadas en marzo de 2009, las cuales se encontraran vigentes a partir del 12 de abril de 2009 para la empresa CAESS.<sup>38</sup> El consumo promedio de Kilowatt Hora (KWH) que se tiene en el local es de **127 KWH mensual**. Partiendo de este promedio se tiene:

Tarifas Eléctricas	Precio	Cantidad KWH	Total
Primeros 99 KWH	0.154601	99	15.305499
Siguientes 100 KWH	0.15373	28	4.30444
Los restantes	0.153096	0	0
Cargo de Comercialización	0.7669	1	0.7669
Cargo de Distribución	0.046303	127	5.880481
Cargo por tasa municipal	0.11	1	0.11
<b>Total promedio mensual</b>			<b>26.36732</b>

Tabla 2.37: Calculo de gasto mensual por energía eléctrica

En la tabla 2.38 se detallan los costos mensuales de los otros recursos de operación que serán utilizados.

Gastos Fijos	Costo Mensual
Alquiler del local	\$ 125
Agua Potable	\$ 2.82
Telefonía Fija	\$ 14.85
Telefonía Móvil (para los 4 integrantes)	\$ 40.00
<b>Total promedio mensual</b>	<b>\$ 182.87</b>

Tabla 2.38: Calculo de gastos mensuales de otros recursos de operación

El costo total de los recursos de operación se obtiene de la siguiente forma:

Total recursos de operación = Suma totales de recursos de operación \* Tiempo de duración del proyecto.

Total recursos de operación = (26.37 + 182.87) \* 8 meses.

Total recursos de operación = 1637.92

### 2.7.1.4. Recursos Consumibles

A continuación se presenta el costo de los materiales consumibles que se utilizarán durante el desarrollo de todo el proyecto.

Detalle	Precio Unitario (\$)	Cantidad	Total (\$)
Resma de papel bond tamaño carta	\$5.00	13	65.00
Cartucho negro Canon IP 2500	\$22.00	14	308.00
Cartucho color Canon IP 2500	\$25.00	6	150

<sup>38</sup> [http://www.aeselsalvador.com/2009/Portals/0/SitioAES/PLIEGO\\_TARIFARIO\\_CAESS.jpg](http://www.aeselsalvador.com/2009/Portals/0/SitioAES/PLIEGO_TARIFARIO_CAESS.jpg), Consultada en abril de 2009.

Folder	\$0.05	50	2.50
Lapiceros	\$0.18	12	2.16
Lápices	\$0.25	12	3.00
Borradores	\$0.90	8	7.20
Sacapuntas	\$0.90	4	3.60
Empastados	\$15.00	10	150.00
Anillados	\$1.20	10	12.00
Copias	\$0.02	500	10.00
Renta de proyector	\$6.00/hora	9	54.00
Consumo mensual de gasolina	\$20.00	8	160.00
Viáticos para equipo de desarrollo (mensual)	\$40.00	8	320.00
<b>Total</b>			<b>1247.46</b>

Tabla 2.39: Resumen de recursos consumibles a utilizar en el desarrollo del proyecto<sup>39</sup>

### 2.7.2. Resumen de Gastos

Recurso	Total
Recurso Humano	\$ 26456.00
Recursos Tecnológicos	\$ 3411.30
Recursos de Operación	\$ 1637.92
Recursos Consumibles	\$ 1247.46
Sub Total	\$ 32752.68
Imprevistos 10%	\$ 3275.27
<b>Total</b>	<b>\$ 36027.95</b>

Tabla 2.40: Resumen de recursos

<sup>39</sup> Costos de productos obtenidos de sitio web de office depot.

## 2.8. METODOLOGÍA

### 2.8.1. Herramientas para Recolección de Datos

#### 2.8.1.1. ENTREVISTA

Será utilizada para recabar información en forma verbal, a través de preguntas propuestas. Las personas entrevistadas serán directores y responsables de las unidades, personal de desarrollo de software y usuarios operativos de la defensoría del consumidor, los cuales mantienen relación con el sistema actual, por lo tanto proporcionarán datos que serán fundamentales para la construcción de una solución que cumpla las necesidades. Estas entrevistas se realizarán al personal de forma grupal por los integrantes del trabajo de graduación acompañado de otras técnicas que se estudiarán más adelante.

##### ***Preparación de la Entrevista***

- a) Determinar la posición que ocupa de la organización el futuro entrevistado, sus responsabilidades básicas, actividades, etc.
- b) Definir el tipo de preguntas, ya sea abiertas o cerradas.
- c) Preparar las preguntas que van a plantearse, y los documentos necesarios.
- d) Fijar un límite de tiempo y preparar la agenda para la entrevista.
- e) Elegir un lugar donde se puede conducir la entrevista con la mayor comodidad.
- f) Hacer la cita con la debida anticipación.

##### ***Gestión de la Entrevista***

- a) Explicar con toda amplitud el propósito y alcance del estudio
- b) Explicar la función propietaria como analista y la función que se espera conferir al entrevistado.
- c) Hacer preguntas específicas para obtener respuestas cuantitativas.
- d) Evitar las preguntas que exijan opiniones interesadas, subjetividad y actitudes similares.
- e) Evitar el cuchicheo y las frases carentes de sentido.
- f) Ser cortés y comedido, absteniéndose de emitir juicios de valores.
- g) Conservar el control de la entrevista, evitando las divagaciones y los comentarios al margen de la cuestión.
- h) Escuchar atentamente lo que se dice, guardándose de anticiparse a las respuestas.

##### ***Resultado de la Entrevista***

- a) Escribir los resultados.
- b) Entregar una copia al entrevistado, solicitando su conformación, correcciones o adiciones.
- c) Archivar los resultados de la entrevista para referencia y análisis posteriores.

#### 2.8.1.2. OBSERVACIÓN

Otra técnica de investigación que será utilizada para la recolección de datos es la observación que consiste en ver a las personas cuando efectúan su trabajo, a fin de conocer el flujo de información y las relaciones entre los elementos que interactúan en el proceso de solución de una denuncia. Esta técnica permite estudiar a las personas en sus actividades de grupo y como miembros de la organización. El propósito es múltiple: determinar que se está haciendo, como se está haciendo, quien lo hace, cuando se lleva a cabo, cuánto tiempo toma, dónde se hace y por qué se hace.

***Preparación para la observación***

- a) Determinar e identificar a la persona que va a observarse.
- b) Estimar el tiempo necesario de observación.
- c) Obtener la autorización de la gerencia para llevar a cabo la observación.
- d) Explicar a las personas que van a ser observadas lo que se va a hacer y las razones para ello.

***Conducción de la observación***

- a) Familiarizarse con los componentes físicos del área inmediata de observación.
- b) Mientras se observa, medir el tiempo en forma periódica.
- c) Anotar lo que se observa lo más específicamente posible, evitando las generalidades y las descripciones vagas.
- d) Si se está en contacto con las personas observadas, es necesario abstenerse de hacer comentarios cualitativos o que impliquen un juicio de valores.
- e) Observar las reglas de cortesía y seguridad.

***Resultado de la observación***

- a) Documentar y organizar formalmente las notas.
- b) Revisar los resultados y conclusiones junto con la persona observada.

**2.8.1.3. CUESTIONARIO**

Se utilizara para poder recolectar datos de un gran número de personas y así conocer varios aspectos del sistema de una forma más confiable.

- a) **Cuestionario cerrado:** Es aquel que solicita respuestas breves, específicas y delimitadas.
- b) **Cuestionario abierto:** Las preguntas abiertas no delimitan de antemano las alternativas de respuesta.

**2.8.1.4. INVESTIGACIÓN BIBLIOGRÁFICA**

Por medio de material bibliográfico, se recolectará información acerca de los antecedentes, características y elementos que el proceso de resolución de las denuncias realizadas por los consumidores. Además por medio de Internet, se recolectará información relevante acerca de control modelado de información y construcción del data warehouse, sobre bases de datos, plataformas, desarrolladores y otras herramientas que ayuden al desarrollo del proyecto.

## 2.8.2. Herramientas para el Diagnóstico y Planteamiento del Problema

### 2.8.2.1. Diagrama Causa – Efecto (Ishikawa)

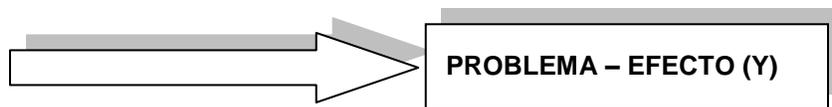
El Diagrama Causa – Efecto también conocido como “Diagrama Espina de Pescado”, es una técnica ampliamente utilizada, que permite apreciar con claridad las relaciones entre un tema o problema y las posibles causas que puedan estar contribuyendo para que él ocurra. Construido con la apariencia de una espina de pescado, esta herramienta fue aplicada por primera vez en 1953, en el Japón, por el profesor de la Universidad de Tokio, Kaoru Ishikawa, para sintetizar las opiniones de los ingenieros de una fábrica, cuando discutían problemas de calidad.

Campos de utilización:

- Visualizar en equipo, las causas principales y secundarias de un problema.
- Ampliar la visión de las posibles causas de un problema, enriqueciendo su análisis y la identificación de soluciones.
- Analizar procesos en búsqueda de mejoras.
- Conduce a modificar procedimientos, métodos, costumbres, actitudes o hábitos, con soluciones (muchas veces) sencillas y baratas.
- Educa sobre la comprensión de un problema.
- Sirve de guía objetiva para la discusión y motiva.
- Muestra el nivel de conocimiento técnico que existe en la empresa sobre un determinado problema.
- Prevee los problemas y ayuda a controlarlos, no solo al final, sino durante cada etapa del proceso.

Para la construcción se realizan los siguientes pasos:

- Establezca claramente el problema (Efecto) que va a ser analizado..
- Diseñe una flecha horizontal apuntando a la derecha y escriba el problema al interior de un rectángulo localizado en la punta de la flecha.



**Figura 2.13: Inicio de proceso para realización del diagrama Ishikawa**

- Haga una lluvia de ideas “Brainstorming” para identificar el mayor número posible de causas que puedan estar contribuyendo para generar el problema, preguntando “¿Porque está sucediendo?”.
- Agrupe las causas en categorías. Una forma muy utilizada de agrupamiento es la conocida como 6m: maquinaria, mano de obra, método, materiales, mediciones y medio ambiente.
- Para comprender mejor el problema, busque las sub causas o haga otros diagramas de Causa – Efecto para cada una de las causas encontradas.
- Para las causas pregunte ¿Porque? 5 veces o al menos 2 veces, hasta que ya no tenga más respuestas.
- Escriba cada categoría dentro de los rectángulos paralelos a la flecha principal. Los rectángulos quedaran entonces, unidos por las líneas inclinadas que convergen hacia la flecha principal.
- Se puede añadir las causas y sub causas de cada categoría a lo largo de su línea inclinada, si es necesario.

Ejemplo:

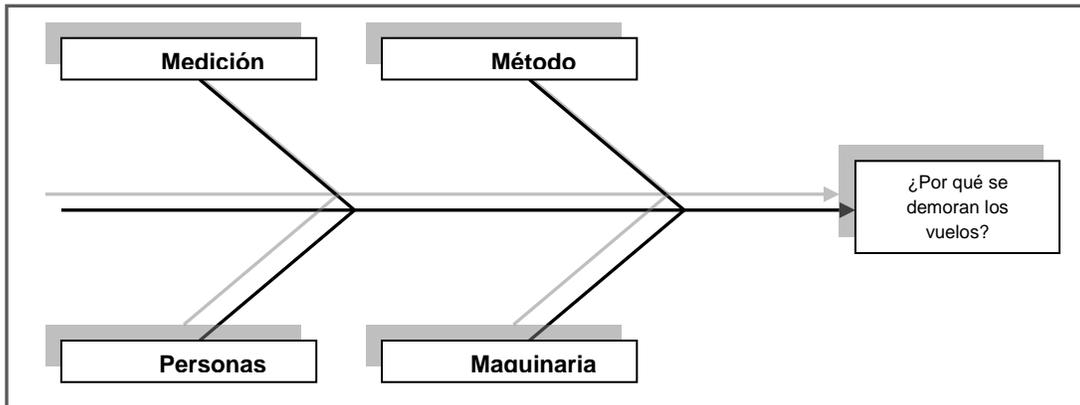


Figura 2.14: Ejemplo de diagrama Ishikawa

### 2.8.2.2. Método de la caja Negra

Esta herramienta se centra constantemente en sus objetivos totales, por consiguiente es necesario definir en primer lugar los objetivos del sistema y examinarlos continuamente. Para comprender el comportamiento y finalidad del sistema

La caja negra es un esquema metodológico que sirve como guía para la solución de problemas, en especial hacia aquellos que surgen en la dirección o administración de un sistema, al existir una discrepancia entre lo que se tiene y lo que se desea, su problemática, sus componentes y su solución.

La primicia del planteamiento del problema inicia con el planteamiento de las entradas de la caja negra, como el estado "A", siendo este el estado actual que se posee. Luego establece el estado "B", el cual define el estado final o resultado que se desea obtener. Dentro de la caja se encuentra metafóricamente los procesos o actividades necesarias para realizar el proceso de transformación del estado "A" hacia el estado "B", como se muestra en la figura 2.15.

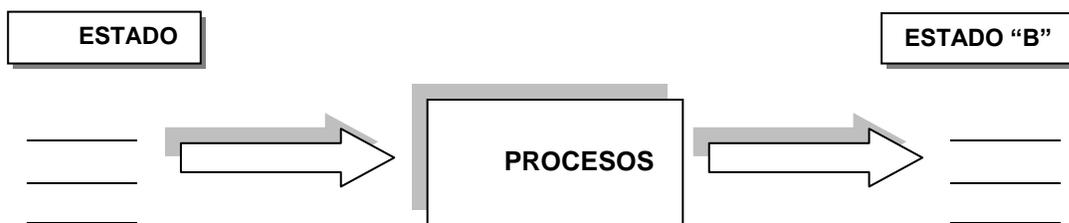


Figura 2.15: Ejemplo de caja negra

### 2.8.3. Modelo UML de los Componentes Bodega de Datos

Este componente de proceso de inteligencia de negocios, el cual se destacan Bill Inmon y Ralph Kimball. El primero lo define: “Una Bodega de Datos es una colección de datos integrados orientados a temas, integrados, no-volátiles y variables en el tiempo, organizados para soportar necesidades empresariales” y para Kimball “la Bodega de Datos es un colección de datos en forma de una base de datos que guarda y ordena información que se extrae directamente de los sistemas operacionales (ventas, producción, finanzas, marketing, etc.) y de datos externos”. Ambos puntos de vista han tenido amplia aceptación, pero el más utilizado es el Kimball. Para consultar la información completa del modelado ver apartado Anexo 7: Descripción del modelo UML de los componentes de bodega de datos.

### 2.8.4. Ciclo de vida

La metodología para el desarrollo de software es un modo sistemático de realizar, gestionar y administrar un proyecto para llevarlo a cabo con altas posibilidades de éxito. Esta sistematización nos indica como dividiremos un gran proyecto en módulos más pequeños llamados etapas y las acciones que corresponden en cada una de ellas. Nos ayuda a definir entradas y salidas para cada una de las etapas y sobre todo normaliza el modo en que administraremos el proyecto.<sup>40</sup>

Es por ello que este proyecto se realizara implementando el ciclo de vida por prototipos ya que es una alternativa que permite la definición de los requerimientos de forma tal que al capturar el conjunto inicial de necesidades se puedan implementar rápidamente con la intención de expandirlas y refinarlas iterativamente, así al ir aumentando la comprensión del sistema por parte de los usuarios estos pueden expresar de mejor manera sus necesidades a fin de que se cumplan sus expectativas, este modelo nos permite tratar mejor la incertidumbre, la ambigüedad y la volubilidad de nuestro proyecto.

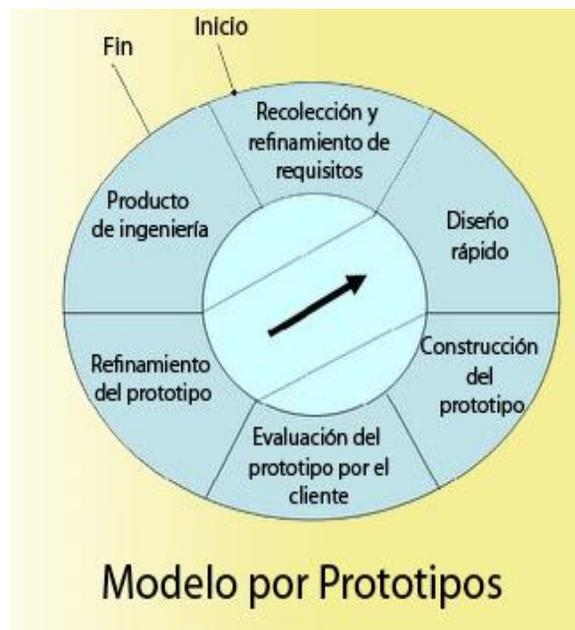


Figura 2.16: Ciclo de Vida por Prototipos

<sup>40</sup> <http://img.redusers.com/imagenes/libros/lpcu097/capitulogratís.pdf> pag 28

La principal ventaja de este modelo de ciclo de vida es que se reduce el riesgo de construir productos que no satisfagan o cumplan con las exigencias del usuario, ya que por medio de la retroalimentación permite que los nuevos prototipos consideren las observaciones ya realizadas en el prototipo anterior y por ende un mejor producto. Este modelo nos brinda uno de los mejores métodos para obtener la retroalimentación por parte del usuario.

Entre las ventajas y desventajas que ofrece este modelo son las siguientes:

Ventajas	Desventajas
Reduce el riesgo de construir productos que no satisfagan las necesidades de los usuarios.	El desarrollador puede caer en la tentación de ampliar el prototipo para construir el sistema final sin tener en cuenta los compromisos de calidad y de mantenimiento que tiene con el cliente.
Es útil cuando el cliente conoce los objetivos generales para el software, pero no identifica los requisitos detallados de entrada, procesamiento o salida.	El usuario tiende a crearse unas expectativas cuando ve el prototipo de cara al sistema final.
Ofrece un mejor enfoque cuando el responsable del desarrollo del software está inseguro de la eficacia de un algoritmo.	En aras de desarrollar rápidamente el prototipo, el desarrollador suele tomar algunas decisiones de implementación poco convenientes

**Tabla 2.41: Evaluación de ventajas y desventajas del modelo elegido**

#### 2.8.4.1. Descripción de las etapas<sup>41</sup>

- **Recolección y refinamiento de requisitos:**  
Los requerimientos del sistema son definidos con el mayor detalle posible. Para esta etapa es común hacer uso de entrevistas a un número de usuarios representativo, a fin de obtener la mayor información posible de lo que el usuario espera al terminar el proyecto.
- **Diseño rápido**  
Se diseña un prototipo preliminar, el cual será nuestro primer prototipo, frecuentemente es un modelo general el cual se irá profundizando a medida que se realice las revisiones por parte de los usuarios. Este primer modelo es a baja escala, el cual pretende ser la primera aproximación de las características del producto final.
- **Construcción del prototipo**  
En esta fase el diseño del prototipo se convierte en una extensión valiosa sobre la manera en que debe realizarse el resto del proyecto, mostrando a los usuarios muy pronto en el proceso como van a actuar las partes del sistema. La elaboración rápida de prototipos previene el exceso de asignación de recursos a un proyecto que puede eventualmente llegar a ser inmanejable.
- **Evaluación del prototipo por el cliente**  
El usuario evalúa a fondo el primer prototipo, tomando en cuenta las fortalezas y las debilidades, los aspectos que pueden ser agregados y que aspectos deben de quitarse o no incluirse. Los desarrolladores recolectan y analizan las observaciones que hizo el usuario.

<sup>41</sup> [http://searchcio-midmarket.techtarget.com/sDefinition/0,,sid183\\_gci755441,00.html](http://searchcio-midmarket.techtarget.com/sDefinition/0,,sid183_gci755441,00.html)

- Refinamiento del prototipo  
Al primer prototipo se le realizan modificaciones según los comentarios proporcionado por los usuarios y se crea un segundo prototipo el cual será evaluado de la misma manera que el primero. Estas etapas se realizan tantas veces como lo fuese necesario a fin de que el usuario quede satisfecho con el prototipo que represente al producto final deseado.
- Producto de ingeniería.  
El sistema final es construido basándose en el prototipo final, este sistema es revisado y probado a fondo. Además se le realizan mantenimientos de rutina se llevan a cabo de manera continua para evitar los fracasos a gran escala y reducir el tiempo de inactividad.

#### 2.8.4.2. Lineamientos para realizar un prototipo

Al trabajar con prototipos es básico considerar los siguientes cuatro lineamientos:

- Trabajar en módulos manejables.  
Un modulo manejable es un modulo que permite al usuario interactuar con sus características claves pero que se puede construir de forma separada de otros módulos del sistema.
- Construir el prototipo rápidamente.  
La preparación de un prototipo operacional permite al analista comprender mejor como desarrollar el resto del proyecto. Al mostrar a los usuarios en las primeras etapas del proceso como se ejecutan en la realidad algunas partes del sistema, evita que se dediquen demasiados recursos a requerimientos que a la larga podría ser imposible de concretar.
- Modificar el prototipo en iteraciones sucesivas.  
Hacer modificable un prototipo significa crearlo en módulos que no sean demasiado interdependientes. El estar conscientes de que el prototipo no es un sistema terminado y que el prototipo requerirá modificación, es una actitud positiva que demuestra a los usuarios cuan necesaria es su retroalimentación para mejorar el sistema.
- Enfatizar la interfaz de usuario.  
Puesto que en realidad el principal objetivo es conseguir que los usuarios expresen mucho mejor sus requerimientos de información. Estos deben interactuar fácilmente con el prototipo del sistema, por lo que la interfaz de usuario se debe desarrollar lo mejor posible para permitir a los usuarios una rápida comprensión del sistema y no sentirse desorientados.

### **2.8.5. Metodología del Proyecto**

Se define de forma clara y específica con el propósito de saber de qué manera se ejecutará el proyecto. Es necesario tener claro como se desarrollaran cada una de las etapas del proyecto, para evitar pérdidas de tiempo y además utilizar una herramienta o método específico de implementación para llevarlo a cabo.

#### **2.8.5.1. Fase 1: Recolección y Refinamiento de Requerimientos:**

Dentro de la metodología para Recolectar y refinar requerimientos de prototipos utilizaremos:

- Establecer una día a la semana en el cual se logre obtener un contacto con la contraparte, de esta manera se determinaran los requerimientos y se aclararan dudas que surjan. En dichas reuniones se buscara una retroalimentación constante verificando e identificando aquellos puntos en los cuales se basará el análisis de la solución.
- Identificar el área de impacto de los usuarios que beneficiara el prototipo en desarrollo haciendo uso de la información brindada por la gerencia informática referente a la incidencia de la solicitud de reportes en un área determinada. De esta manera se identificará el grado de importancia de la fase que se desarrollará.
- Realizar tabulación de datos obtenidos para brindar una mejor comprensión del impacto que genera la solución.
- Proceder a una investigación teórica de los aspectos sobresalientes que tienen relación con la solución y la inteligencia de negocios, identificando y preparando un camino sencillo del desarrollo de la solución haciendo uso de un análisis técnico.
- Haciendo uso de las incidencias de solicitud y sus históricos, plantear la problemática a la cual se busca dar solución, para tener una idea clara de lo que se pretende construir.
- Evitar al máximo dudas que se generen, el desarrollo de elementos dudosos puede ocasionar pérdidas de tiempo, la concepción de la idea original y la identificación clara de los requerimientos haciendo el uso de herramientas como encuestas y dialogo con el usuario final agilizan todo el proceso.

#### **2.8.5.2. Fase 2: Diseño Rápido**

El objetivo de esta fase es construir un modelo de información que refleje el esquema conceptual del prototipo. Es muy importante que este modelo este lo más ajustado posible a la realidad para que el diseño posterior pueda cumplir con sus objetivos.

Se realizaran entrevistas a los usuarios y se estudiara y diseñara el primer prototipo operativo, determinando sus puntos fuertes y puntos débiles, y se documentaran todas las funcionalidades. Se utiliza toda la información recopilada en la fase de Recolección y Refinamiento de Requerimientos y se elabora un diseño lógico del sistema de información. Luego se realizan las especificaciones formales, que implica diseñar procedimientos precisos de captura de datos, accesos efectivos al sistema, la interfaz con el usuario, una base de datos eficiente, etc. La dinámica que cumple dentro de las iteraciones se ve reflejada en la Figura 2.17

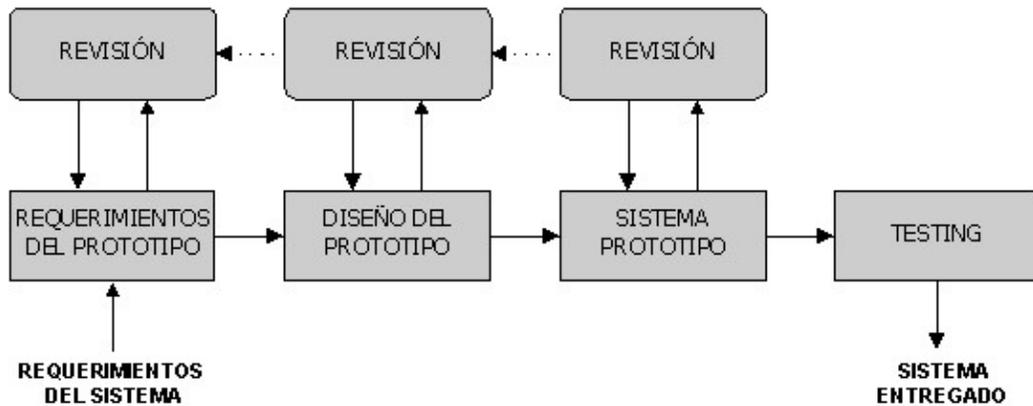


Figura 2.17: Dinámica en la Fase “Diseño Rápido”

La fase de Diseño Rápido posee los siguientes propósitos:

- Completar las funciones de los procesos de ingeniería del negocio a ser soportadas.
- Completar la selección de componentes.
- Desarrollo del modelo de datos y procesos.
- Refinar requerimientos funcionales y del sistema, que no son fácilmente expresados en el modelo de datos y procesos.

Las tareas a realizar dentro de esta fase:

- Consolidar y reafirmar necesidades del negocio
- Revisión de planes, para modificar el alcance del proyecto, incluyendo cambios en el negocio, calendarización o cronograma de actividades y requerimientos técnicos.
- Revisión de planes, para modificar los cambios en los recursos disponibles, incluyendo presupuesto, personal y entrenamiento.
- Actualización del listado de componentes reutilizables y componentes completamente reutilizables.
- Complementar requerimientos de administración de los medios electrónicos de almacenamiento.
- Identificar la adquisición de datos, conversión de estos y estrategias de validación.
- Refinar la arquitectura técnica y construir la estructura arquitectónica del prototipo.

### 2.8.5.3. Fase 3: Construcción del Prototipo

Como su nombre lo indica, esta fase tiene por objetivo la construcción del prototipo operativo de la aplicación de la “N” - enésima iteración. Esta fase consta de dos actividades principales:

El Desarrollo técnico propiamente dicho, el objetivo es traducir las especificaciones del diseño en un código de programación o de estructuración.

La construcción del prototipo debe tomar en cuenta la utilización del lenguaje de cuarta generación de BD lenguajes orientados a resolver problemas, diseñados para la reducción de tiempo, coste y esfuerzo.

#### Lenguaje de cuarta generación no procedimental.

Herramientas CASE: diversas aplicaciones destinadas a aumentar la productividad en el desarrollo de software, reduciendo el coste de las mismas en términos de tiempo y dinero.

Diccionario de datos: especificación del conjunto de metadatos que contiene las características lógicas de los datos que se van a utilizar en el sistema.

Desarrollar la documentación asociada, con la incorporación de las especificaciones al diseño del prototipo, el desarrollo realiza las modificaciones pertinentes para incorporar las funcionalidades de los nuevos requerimientos, con lo que puede modificar la documentación concerniente a la descripción del funcionamiento del sistema.

Lineamientos para el desarrollo:

Trabajar en módulos manejables, cuando se realiza el prototipo de alguna de las características de un sistema para obtener un modelo funcional, es imperativo que los analistas trabajen en módulos manejables, permitiendo la interacción con sus características principales, pero todavía puede ser construido por separado de otros módulos del sistema.

Construir el primero prototipo rápidamente, la velocidad es esencial para la elaboración satisfactoria de un prototipo en un sistema de información.

Modificar el prototipo, un tercer lineamiento para el desarrollo del prototipo enfatiza en que su construcción debe dar soporte a las modificaciones plasmadas en la etapa de diseño. Los cambios al prototipo deben mover al sistema más cerca a lo que los usuarios dicen que es importante.

Enfatizar en la interfaz de usuario, la interfaz del usuario con el prototipo es muy importante, debido a que lo que se está tratando realmente de lograr con el prototipo es hacer que los usuarios muestren cada vez más sus requerimientos de información.

Al finalizar esta fase se dispondrá de un prototipo totalmente funcional y operativo, soportando las funcionalidades solicitadas en la “N” – énesima iteración realizada, siendo sometido a múltiples pruebas en la siguiente fase para comprobar su validez.

#### **2.8.5.4. Fase 4: Evaluación del prototipo por el cliente, prueba y revisión.**

El desarrollador del modulo será el responsable de preparar y presentar el plan de pruebas o guión de pruebas que se le debe realizar al sistema, para asegurar su correcto funcionamiento de acuerdo a los requerimientos definidos por el usuario. Para realizar esta labor el desarrollador puede solicitar la ayuda del usuario del sistema y la contraparte técnica como apoyo.

El desarrollador será el responsable de preparar el ambiente necesario para la realización de las pruebas, esto incluye la preparación de los datos, su inclusión en la base de datos y la documentación de las mismas. Esta documentación o plan de pruebas, debe ser revisada por los usuarios del sistema y la contraparte técnica. Este plan de pruebas debe contener una descripción de cada una de las pruebas a realizar al sistema. Cada prueba debe contener como mínimo los siguientes apartados.

- Nombre del sistema
- Nombre del módulo
- Número de prueba
- Fecha de la prueba
- Hora de inicio de la prueba
- Hora de finalización de la prueba
- Descripción de la prueba
- Acciones a ejecutar

- Resultados esperados
- Responsables de ejecutar la prueba
- Detalle de hallazgos
- Resultados obtenidos.
- Firma de los ejecutores de la prueba

### **Informe de pruebas realizadas**

En la elaboración este informe se debe incluir el resultado de las pruebas realizadas al sistema, de tal forma que se pueda establecer lo siguiente:

- Del guión de pruebas aplicadas al sistema, cual fue el resultado que se obtuvo del mismo.
- Lista de errores detectados durante la realización de las pruebas, con el nivel adecuado de detalle que se requiere para ubicar los errores. Se debe indicar cuál fue la prueba realizada para ubicar fácilmente el error detectado. Para este efecto se debe adjuntar a dicho informe la documentación obtenida producto de la ejecución de las pruebas.

### **Aprobación del sistema**

Acta de aprobación preparada por el usuario donde manifiesta su conformidad con los productos resultantes de la etapa de Construcción, en cuanto a funcionalidad y operativa se refiere de acuerdo a los requerimientos externados inicialmente. Un formulario similar al mostrado en el numeral Anexo 4 se deberá presentar junto a la aprobación del prototipo para poseer un respaldo de la evaluación por parte de los evaluadores.

#### **2.8.5.5. Fase 5: Refinamiento del prototipo<sup>42</sup>**

Esta fase implica la realización de las modificaciones en el prototipo en base a las observaciones que el usuario ha realizado en la evaluación, es así como permite que el producto que se esté construyendo satisfaga las necesidades del usuario, por ello que el prototipo debe de ser flexible para futuras modificaciones, para ello se tendrá que elegir módulos en los que podamos trabajar de la manera menos interdependiente posible.

Por lo general el prototipo es modificado varias veces hasta que el usuario se encuentre de acuerdo con los resultados presentados. Estos cambios harán que el sistema se acerque más a los que el usuario considere más importante, cada modificación necesita que de otra evaluación por parte del usuario, esta evaluación se realizara velozmente a fin de no perder la objetividad de los cambios que se han pedido realizar. La rapidez con que se realice esta evaluación dependerá de que tan rápido realice la evaluación el usuario. La experiencia con el sistema bajo condiciones reales permite la familiaridad indispensable para determinar los cambios o mejoras que sean necesarios, o también la eliminación de características innecesarias. Para facilitar el desarrollo de esta fase se ha considerado elegir actividades primordiales las cuales se detallan a continuación:

- Se enumerara las revisiones comenzando desde el prototipo inicial como una versión 1.0, la cual irá cambiando en base a las observaciones y correcciones que se realicen al prototipo inicial. Con la incorporación de observaciones realizadas por los usuarios en la fase de evaluación, se generara la versión 1.01 y posteriores de acuerdo a la N iteración del prototipo.

---

<sup>42</sup> Metodología para Desarrollo de Sistemas de Información,

- Revisión de minuta de tareas, estas minutas servirán para el análisis de las correcciones de aspectos tales como, fuentes de datos, conversión de datos, aspectos genéricos del sistema, etc. Con las observaciones realizadas se harán retoque tanto en la interfaz como en la arquitectura de la base de datos.
- Para realizar las correcciones utilizaremos las técnicas definidas en el diseño rápido del prototipo, y procederemos a realizar las pruebas, generando un plan de pruebas considerando principalmente los cambios realizados y los módulos que tienen relación con el cambio realizado. Para crear el plan de prueba se debe de acordar con el usuario a fin de replicar un ambiente similar al que se tendrá cuando el prototipo este terminado. Cada prueba deberá tener como mínimo los siguientes apartados:
  - a. Nombre del modulo
  - b. Fecha de prueba
  - c. Descripción de prueba
  - d. Acciones a ejecutar
  - e. Resultados esperados
  - f. Resultados obtenidos
  - g. Detalle de hallazgos.
- Teniendo corregido el modulo y las observaciones del usuario han sido satisfechas, se pactara en una carta de aceptación del usuario, estableciendo los puntos que han sido evaluados y la aceptación del mismo.

#### **2.8.5.6. Fase 6: Producto de ingeniería**

Para el desarrollo de esta etapa, se buscará la integración de los módulos en los que se ha ido dividiendo el sistema, y construyendo cada prototipo. Para poder integrar un modulo, se tendrá que tener la aceptación de parte del usuario que dicho modulo cumple con las exigencias y necesidades actuales. Por lo que se han considerado las siguientes actividades como las principales que ayudaran al desarrollo de esta fase.

- Tomando como base que el diseño de cada modulo, se realiza cumpliendo estándares previamente establecidos, la integración de cada modulo a lo que será el sistema final, será un proceso de verificación de operatividad, ya que se probara que el sistema pueda seguir funcionando cuando se encuentre terminado cada modulo.
- Preparar plan de implantación. En base a cada modulo que se vaya aprobando se irá incluyendo al plan de implantación los requerimientos y descripciones de pasos necesarios para que al finalizar el ultimo prototipo. Para ello se detallara información acerca de:
  - a. Captura inicial de datos.
  - b. Instalación de hardware y software.
  - c. Recursos, tareas.
  - d. Carga esperada sobre el equipo.
  - e. Requerimientos de almacenamiento.
- Completar la documentación del sistema y actualizar la documentación de las etapas anteriores (si corresponde). Para ello se tendrá que verificar todo el sistema integrado y validar que la documentación de cada modulo este actualizada en base al producto de ingeniería que se ha construido hasta el momento. Este paso debe incluir:
  - a. Documentación etapa de análisis actualizada
  - b. Documentación etapa de construcción.
  - c. Documentación de aprobación de fin de cada etapa.
  - d. Documentación de usuario.

- Manual de usuario, para ello utilizaremos de insumo el producto, en el cual se revisaran las opciones que el usuario utilizara y se le brindara una guía completa del sistema.
- Manual técnico del sistema, se tomaran todas las consideraciones tecnológicas para establecer en el documento de forma específica, condiciones optimas para la operatividad del sistema.

## 2.8.6. Metodología de Kimball<sup>43</sup>

### 2.8.6.1. Hechos y Dimensiones

El corazón de la metodología de Kimball se basa en separar la base de datos OLAP en dos entidades diferentes:

- **Dimensión:** Una dimensión es un objeto analítico en el espacio de BI. Una dimensión puede ser un listado de productos, rubros de empresas, el tiempo o cualquier otra entidad utilizada para analizar los números. Las dimensiones tienen atributos, que son las características con las que cuenta la dimensión. Un ejemplo de atributo es el color del producto, su fabricante o peso.
- **Hecho:** Un hecho es algo que ha pasado y/o se ha medido. Un hecho puede ser las ventas de un producto a un cliente determinado o el total de ventas de un artículo en un mes. Desde el punto de vista del analista de BI, un Hecho es un numero que el usuario utilizaría constantemente y de diferentes formas para generar un reporte.

### 2.8.6.2. Esquema de Estrella y Esquema de Copo de Nieve.

Cuando se definen las dimensiones y se crean las relaciones entre las tablas de dimensiones y de hechos se obtiene un esquema de estrella. Esta estructura está compuesta por una tabla central - **tabla de hechos** - y un conjunto de tablas organizadas alrededor de ésta - **tablas de dimensiones**.

En las puntas de la estrella se encuentran las tablas de dimensión que contienen los atributos de las aperturas que interesan al negocio que se pueden utilizar como criterios de filtro y son relativamente pequeñas. Cada tabla de dimensión se vincula con la tabla de hechos por un identificador.

Las características de un esquema de estrella son:

- El centro de la estrella es la tabla de hecho.
- Los puntos de la estrella son las tablas de dimensiones.
- Cada esquema está compuesto por una sola tabla de hechos
- Generalmente es un esquema totalmente desnormalizado, pudiendo estar parcialmente normalizado en las tablas de dimensiones.

---

<sup>43</sup> Introduction to the sqlbi methodology, Draft 1.0, Alberto Ferrari, 20 Septiembre 2008

La figura 19.6. Muestra gráficamente un esquema de estrella

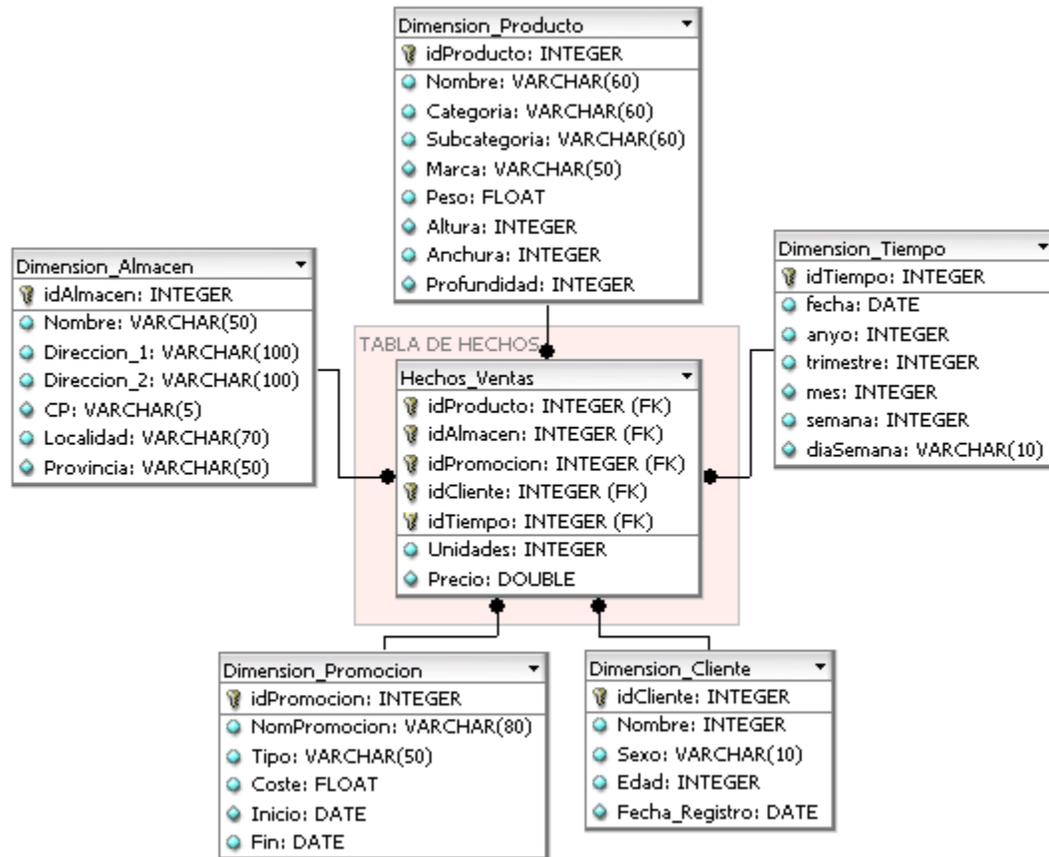


Figura 2.18: Ejemplo de un esquema de estrella

El **esquema copo de nieve** es una variación del esquema estrella donde alguna punta de la estrella se explota en más tablas. El nombre del esquema se debe a que el diagrama se asemeja a un copo de nieve. En este esquema, las tablas de dimensión copo de nieve se encuentran normalizadas para eliminar redundancia de datos. A diferencia del esquema estrella, los datos de las dimensiones se reparten en múltiples tablas. Como ventaja del esquema se destaca el ahorro de espacio de almacenamiento en disco, pero en perjuicio de un aumento en la cantidad de tablas. (Figura 19.7).

Los siguientes son las características de un copo de nieve:

- La dimensión está normalizada
- Los distintos niveles se encuentran almacenados en tablas separadas
- Se argumenta ahorro de espacio

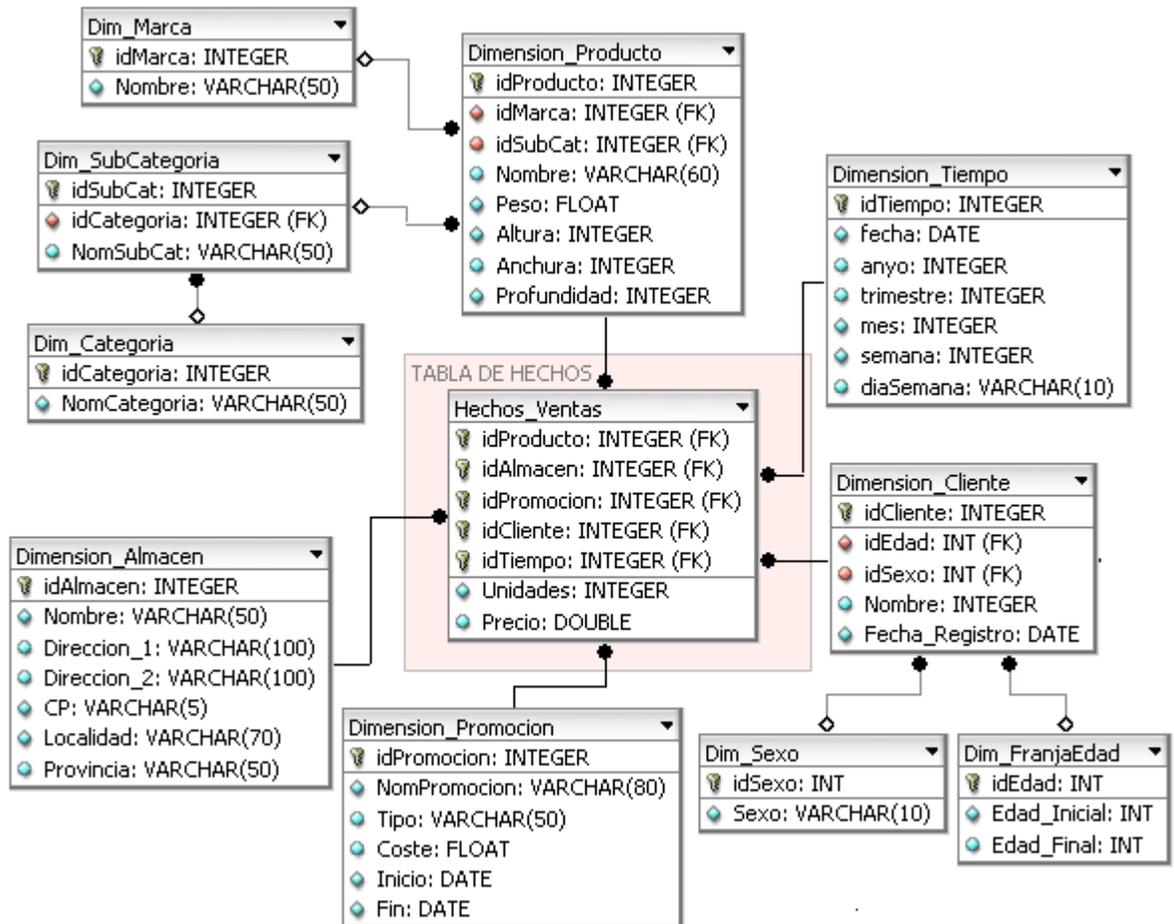


Figura 2.19: Ejemplo de esquema en copo de nieve

### 2.8.6.3. Dimensiones Chatarra

Al finalizar el proceso del modelado, se pueden encontrar atributos que no pertenecen a una dimensión específica. Normalmente estos atributos tienen un rango de valores limitados (de 3 a 4 valores cada uno, aunque muchas veces más) y no parecen ser dimensiones tan importantes como para ser consideradas. En el caso de encontrarse con estas situaciones se pueden considerar dos opciones:

- Crear una dimensión simple para cada uno de los atributos, elevando el rango de atributo a dimensión manteniendo la finalidad del proyecto. Esto conlleva a tener un gran número de dimensiones en la solución final, lo cual no será agradable para el usuario ya que el cubo se vuelve demasiado complejo.
- Unir todos los atributos en una “dimensión chatarra”. Una dimensión chatarra es simplemente una dimensión que unifica todos los atributos que no se pueden colocar en otro lugar y comparten características de tener pocos valores diferentes.

Las dimensiones chatarras son muy útiles por razones de desempeño, ya que él se limita el número de campos en la tabla de hechos. Sin embargo, el analista de BI no debe abusar de ellas. A continuación detallaremos las ventajas y desventajas de estas dimensiones.

Las razones principales para utilizarlas son:

- Si se unen varias dimensiones chatarras pequeñas en una sola, se tendrán menos campos en la tabla de hechos. Para millones de filas, esto puede ser una mejora significativa en espacio y tiempo para el proceso del cubo.
- Reduciendo el número de dimensiones hará que el motor OLAP se desempeñe mejor durante el diseño y procesamiento del cubo, mejorando el rendimiento del servidor y la experiencia del usuario final.
- Al usuario final no le gustan los cubos con demasiadas dimensiones. Su usabilidad y navegabilidad se vuelven compleja.

Sin embargo, existe una gran desventaja al utilizar dimensiones chatarra: cuando se relacionan distintos valores, se establece claramente que estos valores **nunca** tendrán el rango de una dimensión específica.

## 2.9. RESULTADOS ESPERADOS

Basado en las necesidades generales y específicas expresadas en cada una de las reuniones sostenidas con la Gerencia de Sistemas de Información de la Defensoría del Consumidor, se llegó al consenso definitivo de mutua aceptación por ambas partes, que el producto final resultante del desarrollo del proyecto serán los elementos mencionados a continuación.

A continuación se presenta los resultados del DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR, clasificados por resultados esperados producto a entregar:

### 2.9.1. Resultados Esperados

#### 2.9.1.1. Nivel Operativo

Brindando acceso a la información en el momento, dando el nivel de detalle necesario para optimizar las tareas operativas, agilizando el proceso de resolución de denuncias. Funcionalidad específicas a brindar:

- Consulta de datos históricos de solución de reclamos y asesorías presentadas por los consumidores.
- Consulta de datos históricos de llamadas atendidas dentro de Call Center, sobre atención a reclamos que presentan los consumidores, seguimiento de casos.
- Consulta de datos sobre el control de calidad sobre la atención que las oficinas de la Defensoría del Consumidor brinda a los consumidores.
- Consulta de información histórica de los procesos para la resolución de casi en el Tribunal Sancionador.

#### 2.9.1.2. Nivel Administrativo

Realizando tareas de administración de nivel intermedio apoyando las actividades de análisis, seguimiento, control y toma de decisiones, mediante la consulta de información almacenada, proporcionando informes y facilitando la gestión de la información por parte de los niveles intermedios. Funcionalidades en específicos a brindar:

- Muestra de información estadística sobre la solución de los reclamos y asesorías prestadas por los consumidores.
- Muestra de información estadística sobre el registro de llamadas de atención a reclamos que presentan los consumidores. Seguimiento y análisis de casos mediante la consulta de información particular de información transaccional almacenada.
- Muestra de información estadística referente a los controles de calidad sobre la atención que las oficinas de la Defensoría del Consumidor brinda a los consumidores.

#### 2.9.1.3. Nivel Estratégico

Planificando actividades a mediano y largo plazo, tanto de nivel administrativo como de los objetivos que la empresa posee, teniendo a cualquier momento a disposición el desempeño de la defensoría como tal, tendencias y demás comportamientos que pudieran reflejar los datos.

- Datos gráficos históricos de reclamos presentan los consumidores y asesorías realizadas.
- Datos gráficos históricos de campañas realizadas, con detalle de lugares, tema de campaña y fechas en que se realizaron.

- Brindara un fortalecimiento en el análisis a la Dirección de Educación sobre las campanas realizadas por medio de la generación de estadísticas y reportes en tiempo oportuno.

## 2.9.2. Producto a Entregar

### 2.9.2.1. Entregables a Universidad de El Salvador

- Documento de trabajo de graduación impreso (4 ejemplares)
- Disco compacto con documento de trabajo de graduación, documentación del sistema, instaladores y script de la base y procedimiento ETL.
- Proporcionar un beneficio social significativo que ponga en alto el nombre de la Universidad de El Salvador, con un gran impacto en el mejoramiento de los servicios ofrecidos a la población por parte de la Defensoría del Consumidor.

### 2.9.2.2. Entregable a Defensoría del Consumidor

Sistema a entregar	
Elemento	Descripción
Data Warehouse	La creación de un Data Warehouse, una copia de las transacciones de datos específicamente estructurada para la consulta y análisis, contemplando los módulos: <ul style="list-style-type: none"> <li>• Modulo de Centros de Solución de Controversias (CSC).</li> <li>• Modulo de Call Center.</li> <li>• Modulo de Anfitrión.</li> <li>• Modulo de Verificación y Vigilancia.</li> <li>• Modulo de Tribunal Sancionador.</li> </ul>
Cubo	Creación de un Cubo OLAP, creación de una estructura multidimensional que contenga datos resumidos de la base transaccional de SARA, el cual estará dimensionado con las variables de importancia del Data Warehouse del cual se levantara. Resaltando como principal característica la rapidez de consultas.
Front-end	Es la parte del Software que interactúa con el o los usuarios, siendo la parte responsable de recolectar los datos de entrada del usuario, que pueda ser de muchas y variadas formas, y procesarlas de una manera conforme a la especificación que el back-end pueda usar. La creación del front-end se realizara sobre una herramienta de Inteligencia de Negocios de comercialización gratuita (open source).

**Tabla 2.42: Descripción del sistema a entregar**

El objetivo principal que se persigue con el desarrollo del sistema, es el apoyo a los diferentes niveles organizacionales de la Defensoría del Consumidor.

- Nivel Operativo; brindando acceso a la información en el momento, dando el nivel de detalle necesario para optimizar las tareas operativas, agilizando el proceso de resolución de denuncias.
- Nivel Administrativo; realizando tareas de administración de nivel intermedio apoyando las actividades de análisis, seguimiento, control y toma de decisiones, mediante la consulta sobre información almacenada, proporcionando informes y facilitando la gestión de la información por parte de los niveles intermedios.

- Nivel Estratégico; planificando actividades a mediano y largo plazo, tanto de nivel administrativo como de los objetivos que la empresa posee, teniendo a cualquier momento a disposición el desempeño de la defensoría como tal, tendencias y demás comportamientos que pudieran reflejar los datos.

**2.9.2.3. Elementos adicionales inherentes al sistema como resultado a entregar**

<b>Manuales del Sistema</b>	
<b>Elemento</b>	<b>Descripción</b>
Manual de Instalación y Configuración	Proporcionara los pasos a seguir detallando las especificaciones y características técnicas de los servicios y programas que deben ser instalados y configurados en el servidor donde estará el sistema. Incluye la configuración necesaria para el correcto funcionamiento.
Manual Operativo	Documento guía para el usuario sobre la operatividad del sistema y las funciones que este ofrece a cada uno de ellos, describe cada una de las opciones que integran el sistema y el inicio de sesión a este.
Manual Técnico	Documentación que contendrá toda la descripción muy bien detallada sobre las características físicas y técnicas de cada elemento de los equipos necesarios para el funcionamiento del sistema, describiendo cada una de las características de cada recurso.
Documentación del diseño del Sistema	Documentación de todos los elementos planificados en la etapa de diseño del sistema y realizados en la etapa de desarrollo, incluirá: <ul style="list-style-type: none"> <li>• Definición de estándares</li> <li>• Diagramas de especificación del sistema:</li> <li>• Modelo conceptual</li> <li>• Diagrama de secuencia</li> <li>• Diagrama de estado</li> <li>• Modelado dimensional</li> </ul>
<b>Implantación del Sistema</b>	
Plan de Implantación	Documentación de las actividades necesarias a realizar para la correcta implantación del sistemas, incluirá: <ul style="list-style-type: none"> <li>• Establecimiento del plan de implementación</li> <li>• Formación necesaria para la implantación</li> <li>• Incorporación del sistema al entorno de operación</li> <li>• Pruebas de implantación del sistema</li> <li>• Pruebas de aceptación del sistema</li> <li>• Presentación y aprobación del sistema</li> </ul>

**Tabla 2.43: Elementos adicionales a entregar**

# Capítulo III



### 3.1. DISEÑO GENERAL

#### 3.1.1. Modelado del Data Warehouse usando UML

Modelado de datos y las técnicas tradicionales, como el diagrama Entidad-Relación (ER) y las diferentes extensiones del ER, no son apropiadas para el diseño de una DW, con relación a la complejidad de la correspondencia de dichos modelos. Por ejemplo, Ralph Kimball expone:

*“El modelo de datos Entidad-Relación es un desastre para las consultas (querying) porque no pueden ser entendidas por los usuarios y no pueden ser exploradas fácilmente por Software Administradores de Bases de Datos (DBMS). El modelo Entidad-Relación no puede ser utilizado como la base para un Data Warehouse de una empresa”*

Para solventar este problema, se ha propuesto la definición de un nuevo lenguaje de modelo, utilizando UML. Este combina elementos de los tres más conocidos diseñadores especializados en el modelado orientado a objetos: Rumbaugh’s Object-modeling Technique, Booch’s Object-Oriented y Jacobson’s Objectory.

El lenguaje que se utilizará, es una extensión de UML, llamado también UML Extendido. Es extendido con el fin de proporcionar los mecanismos para introducir los nuevos elementos para especificar el dominio si es necesario, tales como: las aplicaciones web, aplicaciones de base de datos, modelado del negocio, proceso de desarrollo de software, etc.

En la figura 3.1 se muestra la simbología de diagramado para resumir o dar un aproximado y la relación entre los diferentes diagramas.

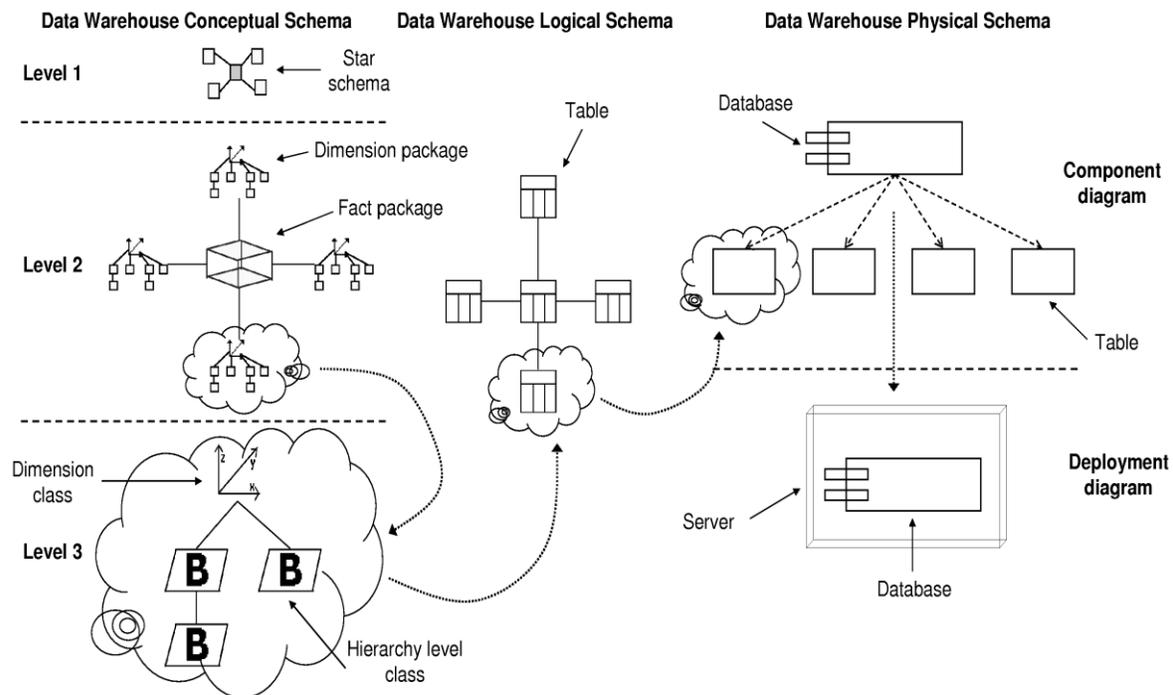


Figura 3.1: Diseño general de un Data Warehouse

**Mecanismos de extensión UML**

- UML es un lenguaje de propósito general para el modelado visual de los sistemas.
- Los diseñadores de UML se dieron cuenta que no era posible diseñar un lenguaje de modelado universal que pudiera satisfacer las necesidades presentes y futuras de todos, por lo que UML incorpora tres simples mecanismos de extensibilidad:
  - Estereotipos -> Nuevos elementos de construcción.
  - Valores etiquetados -> Nuevas Propiedades.
  - Limitaciones (Constraints) -> Nuevas Semánticas.

**Estereotipos**

En un diagrama UML, existen cuatro posibles representaciones de un elemento estereotipado: icono (este estereotipo es desplegado en lugar de la representación normal del elemento), decoración (el estereotipo es desplegado dentro del elemento), etiqueta (el nombre del estereotipo es desplegado y aparece dentro de <<>>) y ninguno (el estereotipo no es indicado), observar figura 3.2.

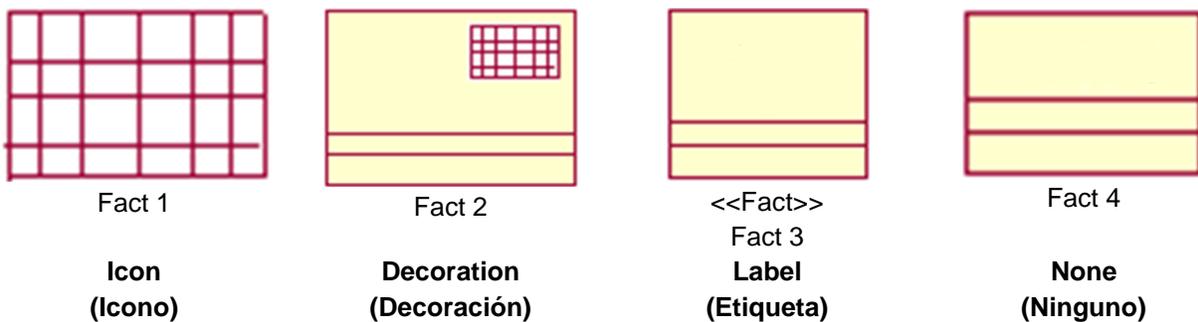
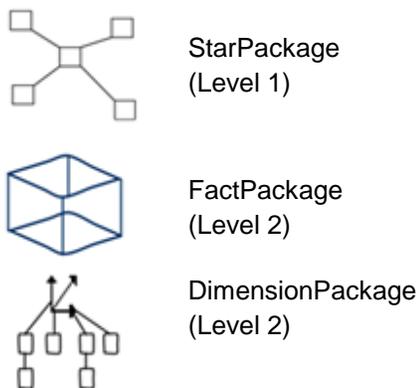


Figura 3.2: Representación de un elemento estereotipado

**Package stereotypes**



**Class stereotypes**

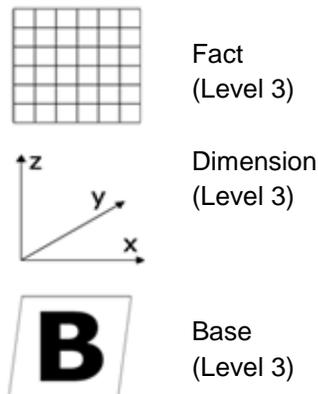


Figura 3.3: Representación de estereotipos de paquetes y de clases

**Framework de Diseño para DW**

El desarrollo de un DW puede ser estructurado en un framework integrado:

- Cinco Etapas (orígenes, integración, Data Warehouse, personalización, cliente)
- Tres Niveles (conceptual, lógico y físico)

Los diagramas se dispersan a través de las cinco etapas y los tres niveles. Cada uno de los diagramas utiliza diferentes formalismos (diagrama de clases, de componentes, etc.). Distintos perfiles UML han sido propuestos:

- Perfil Multidimensional.
- Perfil ETL
- Perfil de Mapeo de Datos
- Perfil de Implementación de Base de Datos

	<b>Fuente</b> OLTP, Datos externos, etc	<b>Integración</b>	<b>Datawarehouse</b>	<b>Personalización</b>	<b>Cliente</b> OLAP Datamining
<b>Conceptual</b>	<b>SCS</b> Diagrama de Clase <i>UML estandar</i>	<b>DM</b> Diagrama de Clase <i>Mapeo de Datos de perfil</i>	<b>DWCS</b> Diagrama de Clase <i>UML estandar perfiles Multidimensionales</i>	<b>DM</b> Diagrama de Clase <i>Mapeo de Datos de perfil</i>	<b>CSS</b> Diagrama de Clase <i>UML estandar Perfiles multidimensionales</i>
<b>Logico</b>	<b>SLS</b> Diagrama de Clase <i>Modelado de diferentes perfiles</i>	<b>ETL</b> Diagrama de Clase <i>perfil ETL</i>	<b>DWLS</b> Diagrama de Clase <i>Modelado de diferentes perfiles</i>	<b>Proceso de Exportación</b> Diagrama de Clase <i>perfil ETL</i>	<b>CLS</b> Diagrama de Clase <i>Modelado de diferentes perfiles</i>
<b>Fisico</b>	<b>SPS</b> Componentes y dist. diagramas distribución de perfiles de la BD	<b>Diagrama de transportacion</b> Diagrama de distribución, distribución de perfiles de la BD	<b>DWPS</b> Componentes y dist. diagramas distribución de perfiles de la BD	<b>Diagrama de transportacion</b> Diagrama de distribución, distribución de perfiles de la BD	<b>CPS</b> Componentes y dist. diagramas distribución de perfiles de la BD

Figura 3.4: Etapas y niveles de un DW

**Etapas**

- Orígenes: Datos de origen (OLTP, orígenes de datos externos, etc.)
- Integración: mapeo entre los orígenes de datos y el data warehouse.
- Data Warehouse: Estructura del Data Warehouse.
- Personalización: mapeo entre el Data Warehouse y las estructuras del cliente.
- Cliente: Estructuras usadas por los clientes para acceder al DW (data marts, aplicaciones OLAP, etc).

**Niveles**

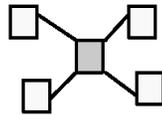
Para cada etapa hay diferentes niveles

- Conceptual
- Lógico
- Físico

No todos los proyectos de DW necesitan los quince diagramas. Los diferentes diagramas del mismo DW no son independientes pero se superponen (Mecanismo de importación de UML).

### 3.1.2. Diseño Físico del DW (Ejemplo).

#### Nivel 1: Definición del modelo



Auto-sales schema

Figura 3.5: Definición del modelo

Este es el nivel 1 del tipo de esquema conceptual en el que se realiza un análisis de los componentes que se incluirán en el Data Warehouse, figura 3.5; Cada uno de estos componentes formaran un elemento importante dentro de un Warehouse conocido como DATA MART, el cual es una especialización de un área del negocio que brinda soporte para toma de decisiones<sup>44</sup>. El Data Mart es un elemento que se accede haciendo uso de una herramienta OLAP, el contenido de un DATA MART puede estar relacionado con otros DATA MART compartiendo datos de una dimensión, cuando el esquema utilizado de un DATA MART es tipo estrella y estos se relacionan entre si se considera una estructura de tipo Constelación.

#### Nivel 2: Definición del Esquema Estrella

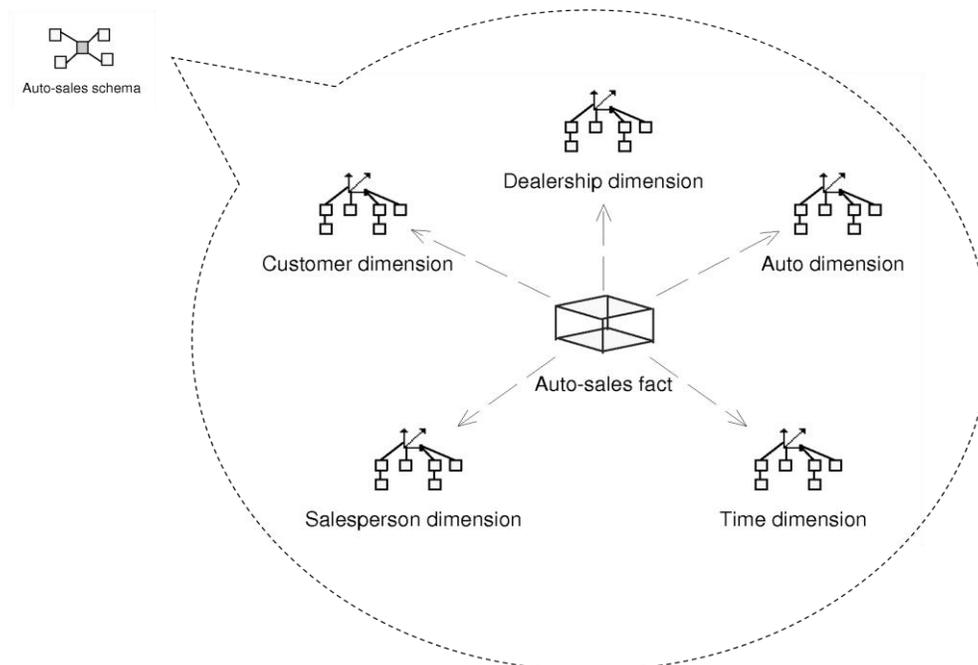


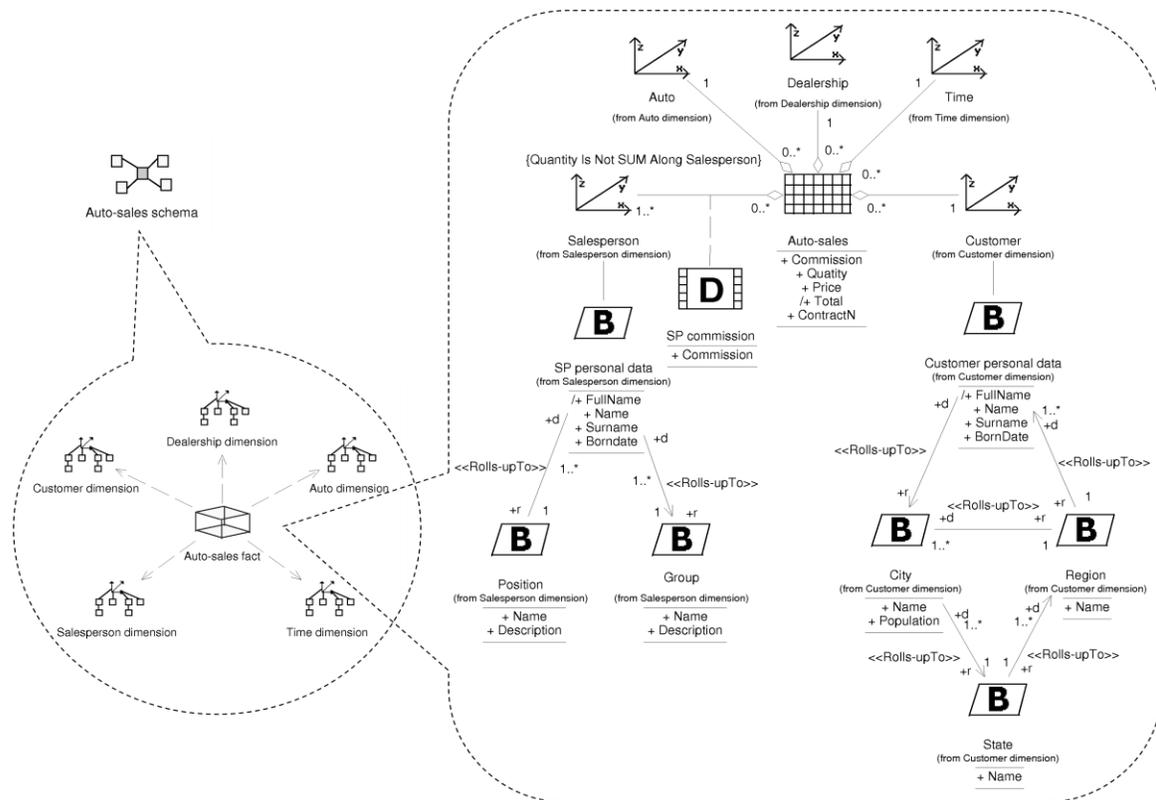
Figura 3.6: Definición del esquema estrella

El centro de un esquema de tipo estrella es la tabla de hecho, FACT TABLE que es básicamente aquella tabla que contiene los valores de las medidas de negocio. Cada medida depende de cada una de

<sup>44</sup> [http://es.wikipedia.org/wiki/Data\\_mart](http://es.wikipedia.org/wiki/Data_mart)

las dimensiones utilizadas, dichas dimensiones sirven como intersección con las medidas que se poseen y se relaciona directamente con la tabla de hechos. En la tabla de hechos se determinan elementos importantes como la granularidad de los datos, que no es otra cosa más que el nivel de detalle que se posee para análisis, la granularidad de la tabla de hechos representa el nivel atómico del Warehouse, ver figura 3.6. Es muy importante realizar un análisis del esquema estrella para lograr determinar aquellas medidas que son parte de un hecho en particular. Un hecho dentro de un negocio a diferencia de un sistema transaccional se refiere a eso, un HECHO, ventas, compras, datos de empleados, etc. En un sistema transaccional común se analiza requerimientos y se les da una solución específica, los hechos nos ayudan a realizar análisis más especializado de la información, es por eso que un DATA WAREHOUSE casi nunca puede tener solamente UN HECHO, sino más bien cada actividad que genere información para la toma de decisiones o que se considere crítico para el negocio puede ser observado como un HECHO, y nos puede facilitar el análisis en el momento de generar nuestro nivel 2 de esquema conceptual que es la descripción más detallada de los modelos definidos en el nivel 1.

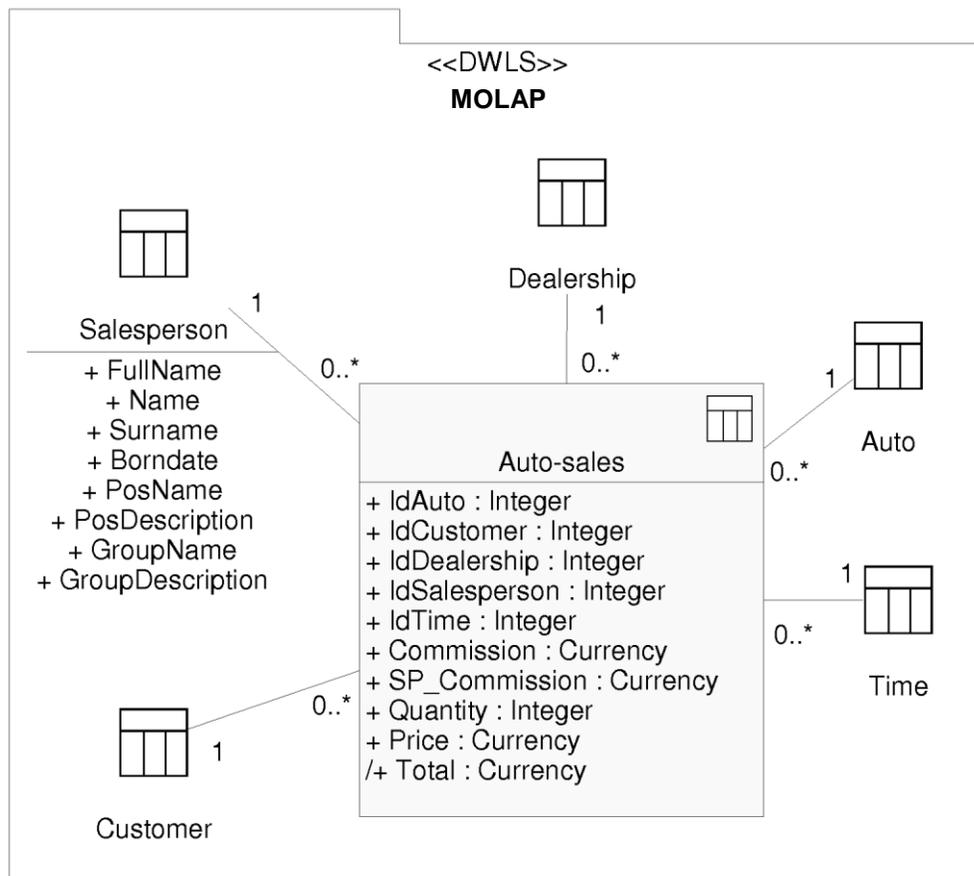
**Nivel 3: Definición de Dimensiones/Hechos**



**Figura 3.7: Definición de las dimensiones y hechos**

Una vez definido el hecho y las dimensiones se procede al detalle de cada una de las medidas de la tabla hecho y el nivel jerárquico de las dimensiones que se poseen y su nivel de dependencia entre ellas si es que las hay, ver figura 3.7. Cada Hecho relacionado con sus dimensiones es modelado realizando el análisis considerando si es permitido un análisis drill down o roll up (especialización, generalización o agrupación) con el que se lograra mostrar los datos contenidos. La cardinalidad que se posee entre la TABLA de HECHO y las dimensiones es importante ya que define el tipo de relación que se posee entre las clases consideradas.

**Diseño Físico del DW: Esquema Lógico MOLAP**



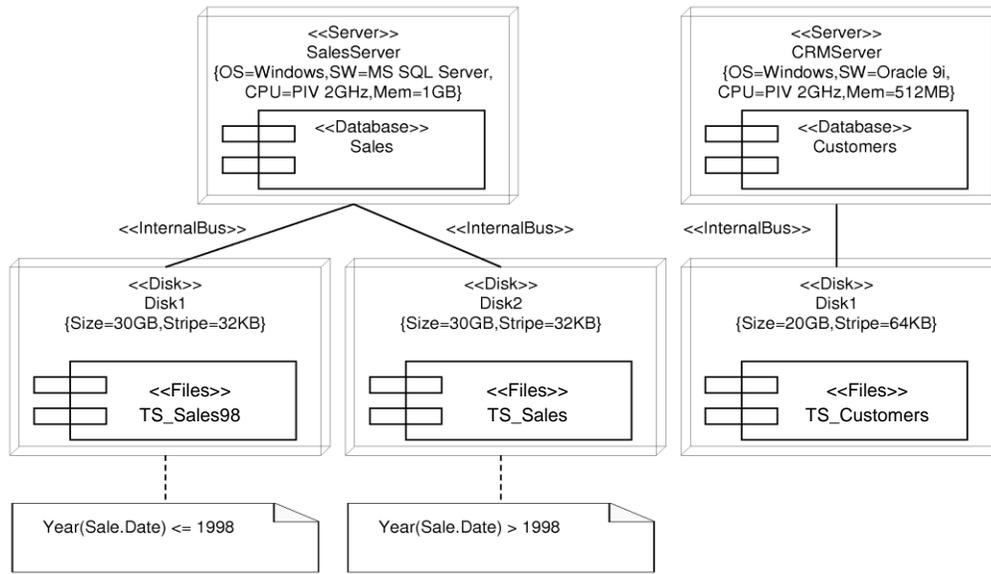
**Figura 3.8: Modelo MOLAP**

La creación de un esquema lógico se basa en el análisis conceptual elaborado anteriormente, en el esquema lógico cada dimensión considerada se convierte en una tabla dentro del DATA MART y la tabla de hecho será la que posea aquellos indicadores de relación respecto a las dimensiones dependiendo del análisis drill down o roll up que se haya considerado, ver figura 3.8. En este momento se deben considerar todos aquellos atributos que forman parte del DATA MART, en este punto no consideramos ningún método de ETL sino que nos interesa construir el esquema del DATA MART con el cual posteriormente se trabajara para el llenado de los datos (proceso ETL).

Es muy importante destacar que en este momento debemos de manejar muy bien el diccionario de datos de la base de datos transaccional, o en el caso se debería conocer el tipo de datos fuente para que concuerde con el tipo de dato destino. Este punto es de gran importancia a la hora de crear los atributos dado que se debe de poseer consistencia en la información generada en el transaccional como en el data mart.

El otro elemento importante es la cardinalidad, es de gran importancia para obtener los resultados esperado, se define la cardinalidad de las relaciones que se poseen entre la tabla de hecho con las tablas dimensiones, auxiliándose siempre con el análisis del esquema conceptual elaborado anteriormente.

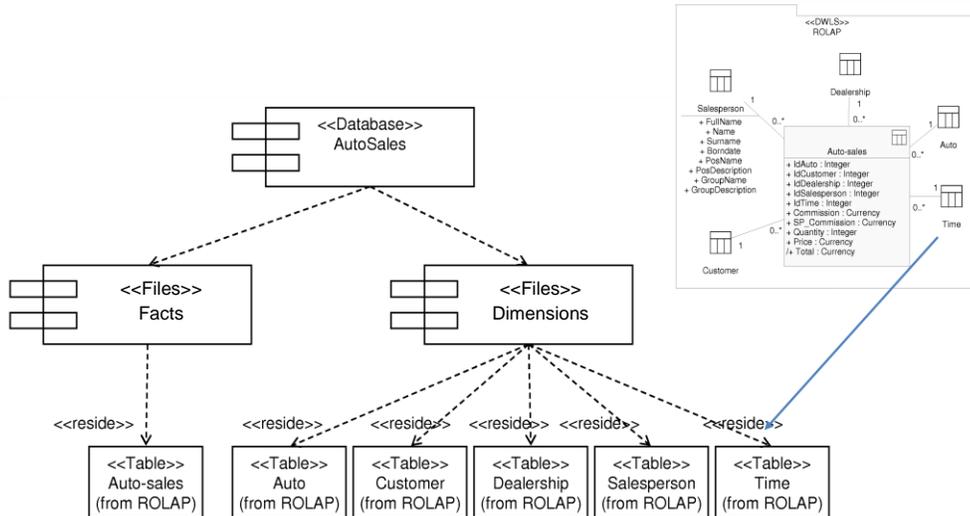
**Esquema Físico de Origen: Diagrama de Implementación.**



**Figura 3.9: Diagrama de implementación**

En la figura 3.9, se observa el diagrama de implementación, el cual nos brinda la información detallada del entorno en el cual se implementara la solución, se deben considerar elementos como tipos de servidores con sus respectivas características, así como también aquellos sistemas operativos y DBMS que se poseen en cada equipo, esto se hace para tener claro los recursos que se poseen y desde donde se está obteniendo la información, se deben detallar el tipo de arreglo de discos duros y la función principal que cumple el servidor respecto a la solución que se posee.

**DW Esquema Físico: Diagrama de Componentes.**



**Figura 3.10: Diagrama de componentes**

El diagrama de componentes nos brinda un panorama general de la estructura del data mart dentro de la base de datos que se posee, es una vista general de las tablas de hecho y dimensiones que se poseen, dicha información es de carácter general y nos muestra los elementos que poseemos dentro del esquema físico de la base de datos.

### 3.2. DISEÑO ESPECÍFICO DE LA SOLUCIÓN PROPUESTA

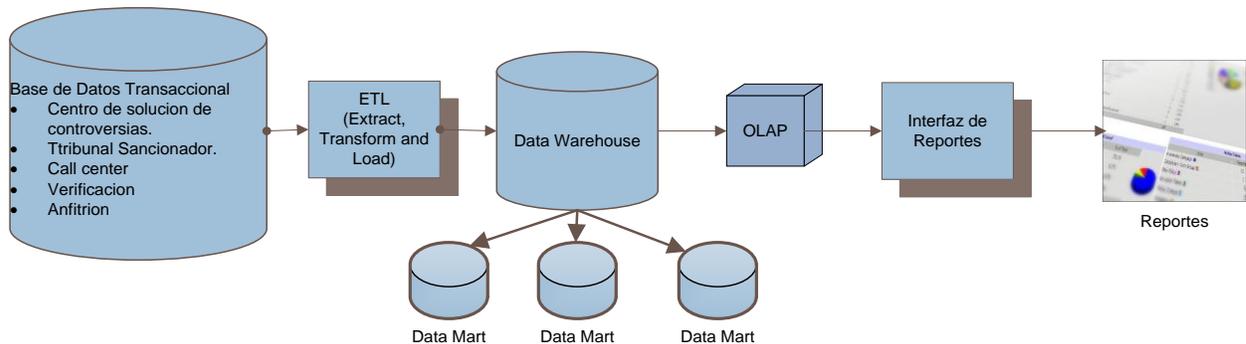


Figura 3.11: Diseño general

#### 3.2.1. Base de Datos Transaccional

Para la creación de la solución se parte de una base de datos transaccional, la cual almacena las transacciones diarias de los diferentes sistemas de la Defensoría del Consumidor. Los principales sistemas que aloja esta base de datos son los siguientes:

- Centro de Solución de Controversias
- Tribunal Sancionador
- Call Center
- Verificación
- Anfitrión

El listado de tablas pertenecientes cada módulo podrá observarse en el desarrollo de este documento, dado que cada módulo se encuentra representado por los capítulos II al capítulo VI. Esta información almacenada es el insumo principal del Data Warehouse, la cual debe ser extraída, transformada y cargada utilizando el proceso de ETL, que se detalla más adelante.

#### 3.2.2. ETL (Extract, Transform and Load)

Es el proceso que permite realizar la transferencia de datos desde la base de datos transaccional de la Defensoría del Consumidor, reformatearlos y limpiarlos para cargarlos en el Data Warehouse. Ambas bases de datos se encuentran en Sql Server. La herramienta que se utiliza para realizar este proceso es Integration Services, de Microsoft el cual es parte del paquete de Sql Server 2005 Standard Edition.

##### Extraer

Es la primera parte del proceso ETL y consiste en extraer los datos de la base de datos transaccional de SARA. Para realizar este proceso se deben escoger los orígenes de datos del módulo a desarrollar y realizar las conexiones necesarias a estos para proceder a transformarlos y luego cargarlos en el Data Warehouse.

### **Transformar**

La fase de transformación aplica una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. Algunas fuentes de datos requerirán alguna pequeña manipulación de los datos<sup>45</sup>. Esta manipulación se realiza para crear uniformidad en los datos que serán almacenados en el Data Warehouse, ya que es probable que los datos provenientes de los distintos módulos que conforman SARA no contengan la uniformidad de datos que se requiere. Este proceso intermedio se realiza en Integration Services por medio de consultas o procedimientos almacenados antes de ser almacenados en el Data Warehouse.

### **Cargar**

La fase de carga es el momento en que los datos de la fase anterior son cargados en el Data Warehouse. Esta fase interactúa directamente con la base de datos destino, en la cual se aplicarán todas las restricciones de integridad y triggers que se hayan definido en esta. Todas estas restricciones y triggers contribuyen a que se garantice la calidad del tratamiento de los datos del proceso ETL. Los datos de cada módulo de SARA son almacenados en su respectivo Data Mart. Esta carga se realiza por medio de tareas programadas en la base de datos, por lo que se define una hora específica para que se realice este proceso.

### **3.2.3. Data Warehouse**

Un Data warehouse o almacén de datos, es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se trata, sobre todo, de un expediente completo de una organización, más allá de la información transaccional y operacional, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de datos (especialmente OLAP, procesamiento analítico en línea). Un Data Warehouse puede ser considerado como una copia de la información transaccional estructurada específicamente para realizar consultas y reportes.

El propósito del Data Warehouse de la Defensoría del Consumidor es proporcionar la información necesaria para realizar análisis de información y reportes. Además mejora la calidad de información y reduce la carga de trabajo para los encargados del área de Sistemas de la Institución.

Otra de las funciones que proporciona es la integración de los datos históricos de la institución para poder realizar análisis a través del tiempo. La frecuencia de actualización de los datos en un Data Warehouse se realiza periódicamente y su rendimiento permite la obtención rápida de la información.

#### **Data Marts**

Un data mart es una versión especializada de un Data Warehouse. Al igual que los Data Warehouse, los data marts contienen una copia de los datos transaccionales que ayudan a la gente de negocios a establecer estrategias basados en el análisis de amenazas y experiencias pasadas. La diferencia entre estos dos componentes radica en que el data mart está enfocado a una necesidad específica y predefinida para cierto tipo de información.

---

<sup>45</sup> <http://es.wikipedia.org/wiki/ETL>

El data warehouse estará conformado por 5 data marts, los cuales son especializaciones de cada uno de los módulos que están involucrados en el proyecto, lo que facilitara la administración y la carga de datos, así como también la obtención de los datos para análisis.

#### 3.2.4. Cubo OLAP <sup>46</sup>

OLAP o Proceso analítico en línea es el nombre formal para el análisis de cubos multidimensionales, lo que permite una forma más intuitiva de ver la información empresarial. Con OLAP se puede ver un conjunto de datos del negocio de muchas diversas formas sin mucho esfuerzo. Los archivos OLAP o cubos modelan los datos en dimensiones. Una dimensión es una clasificación de alguna actividad en una organización por la cual se puede medir su éxito.

Hay dos clases de dimensiones que se pueden utilizar, dimensiones regulares y dimensiones de medida.

**Dimensiones Regulares:** son aquellos datos que se quieren medir, por ejemplo, si se desea seguir el control de denuncias por sector, se puede utilizar:

- **Ubicación:** Para determinar las ubicaciones de las denuncias impuestas.
- **Tiempo:** Como se han aumentado o reducido en el último año o mes.
- **Dimensiones de medida:** son los números que aparecen en el análisis dependiendo de los elementos seleccionados en las dimensiones regulares. Por ejemplo, cantidad de denuncias, montos reclamados, etc.

Existen dos operaciones básicas que se pueden realizar con un cubo OLAP:

**Rotar y Rebanar:** Se pueden cambiar las dimensiones del cubo que se está viendo y obtener una nueva vista de información. Por ejemplo, 'Casos por centro de solución ' puede cambiarse fácilmente a 'Casos por Municipio'. Rebanar es cambiar el valor de una dimensión por otro valor, por ejemplo, de los casos de enero a los casos de febrero. Rotar es aventar el cubo como si fuera un dado para obtener una nueva cara del cubo.

**Taladrar o Drilling:** Los datos de las dimensiones se pueden abrir para obtener más detalle. Una especie de taladro que se hunde más en la información. Si se visualiza información geográfica, se puede pasar de una Zona a un departamento y luego a un municipio.

Toda la información presentada por el cubo OLAP se obtiene directamente de los datos almacenados en el Data Warehouse, cuya estructura facilita el uso de todas las capacidades proporcionadas por la herramienta OLAP.

#### 3.2.5. Interfaz de Reportes

La interfaz de reportes es la que permite acceder a la información contenida en el Data Warehouse y manejarla por medio del cubo OLAP, realizando las operaciones básicas que permite. Además permite al usuario de negocio crear sus reportes de acuerdo a las necesidades de información que posea.

Una característica importante con la que cuenta una interfaz de reportes es la posibilidad de generar gráficos que faciliten el análisis de información, así como también la exportación de los datos obtenidos a herramientas como Excel y a formato PDF.

---

<sup>46</sup> <http://www.olapxsoftware.com/es/WhatsOlap.asp>

### 3.3. DIAGRAMA DE SECUENCIA

Los diagramas de secuencias Son una representación gráfica que destaca la ordenación temporal de los mensajes, además ofrece al lector una señal visual clara del flujo de control a lo largo del tiempo. Para construir un diagrama de secuencia, se coloca al a izquierda el objeto que inicia la interacción, y los objetos subordinados a la derecha, luego se colocan los mensajes que se envían y reciben a lo largo del eje Y.

#### 3.3.1. Validar usuario

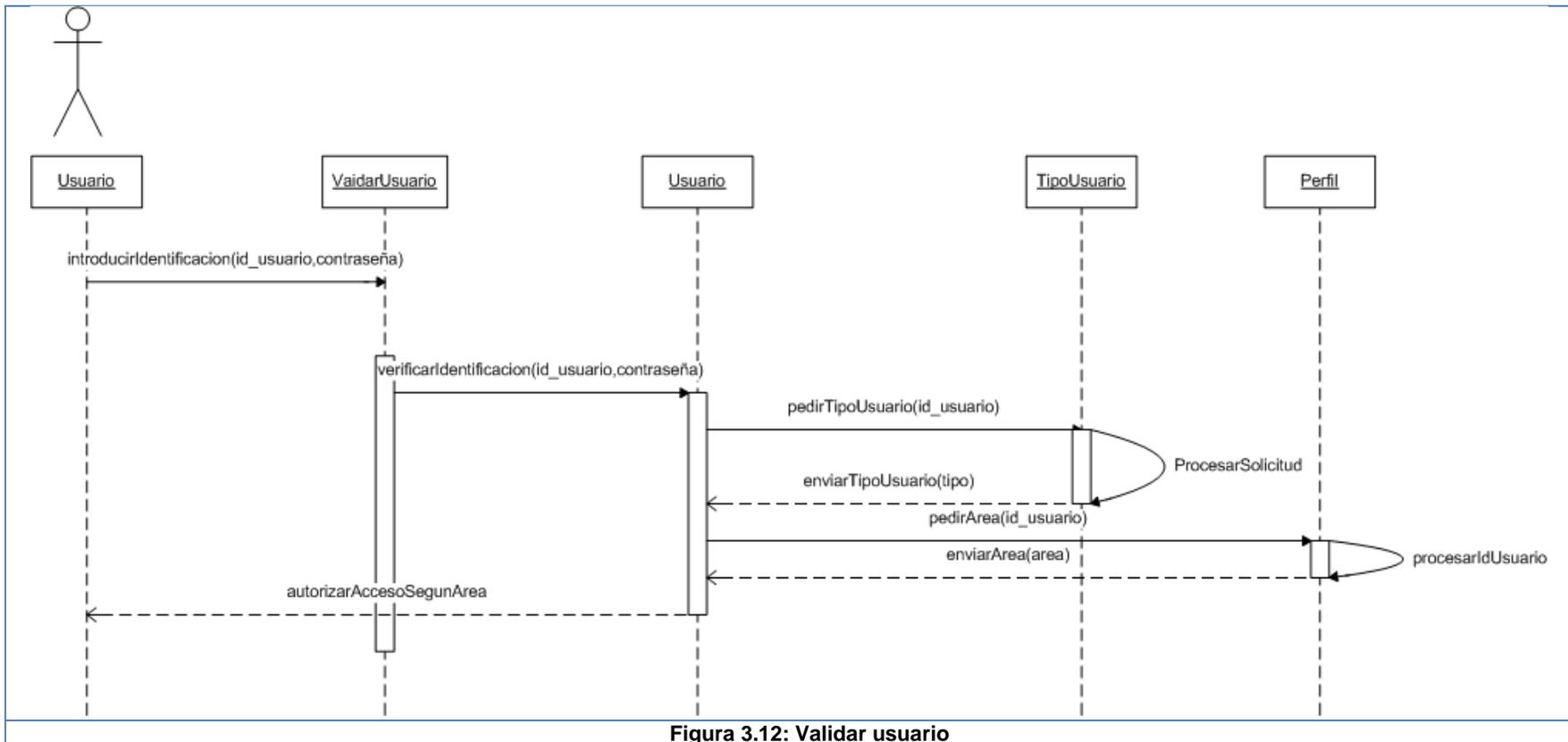


Figura 3.12: Validar usuario

### 3.3.2. Gestionar Roles

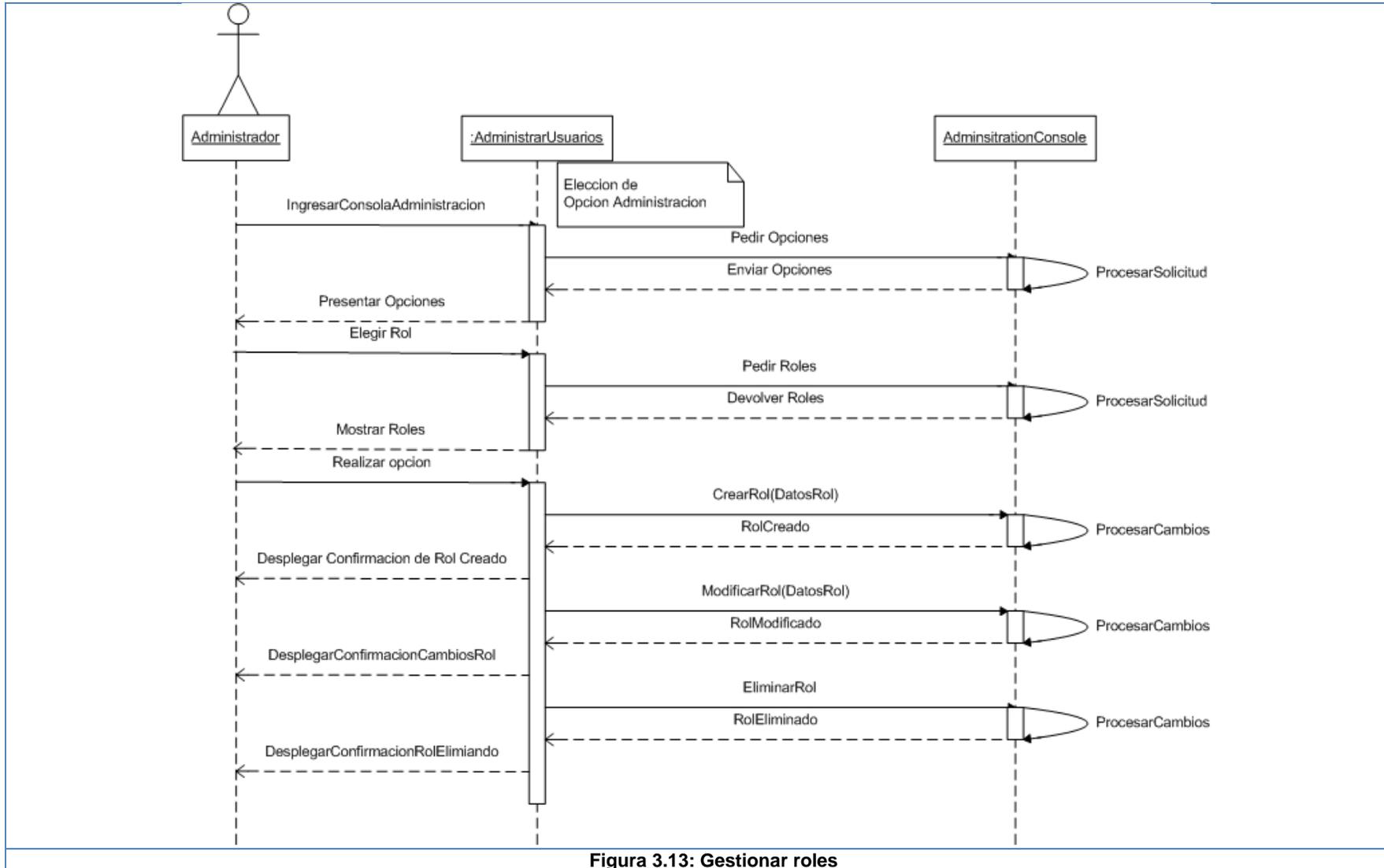
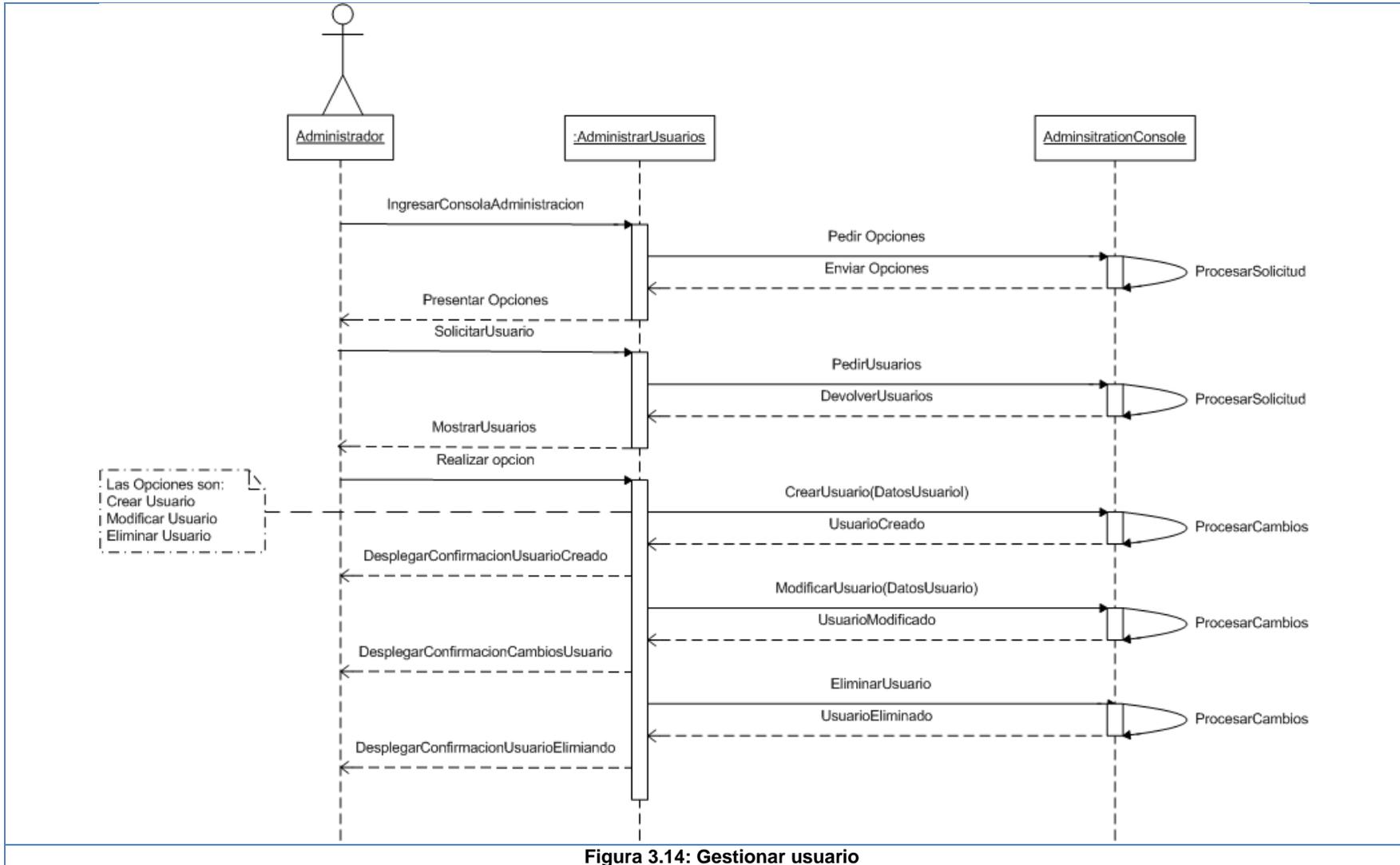


Figura 3.13: Gestionar roles

3.3.3. Gestionar Usuario



### 3.3.4. Gestionar Reportes

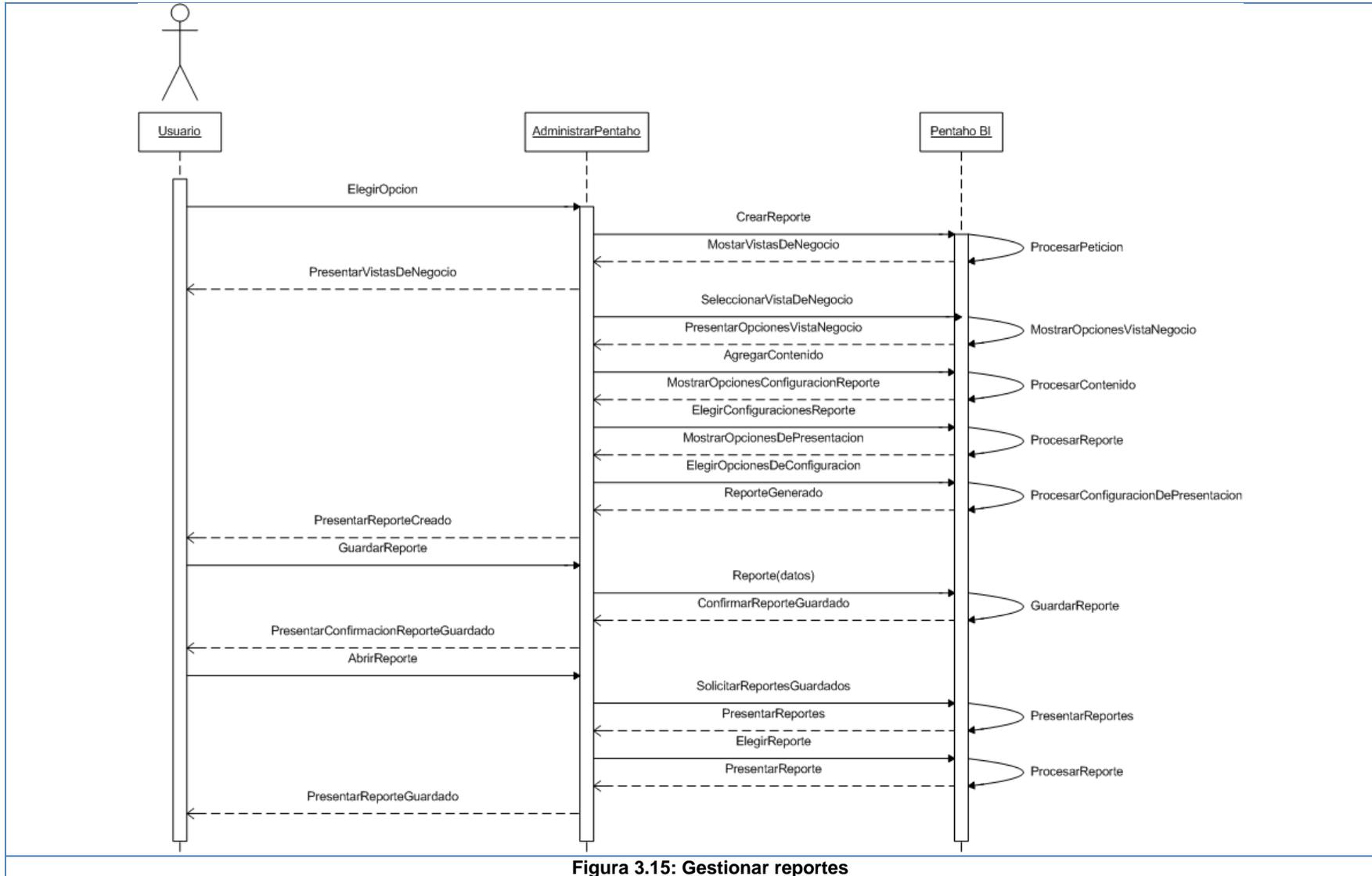


Figura 3.15: Gestionar reportes

3.3.5. Gestionar Análisis

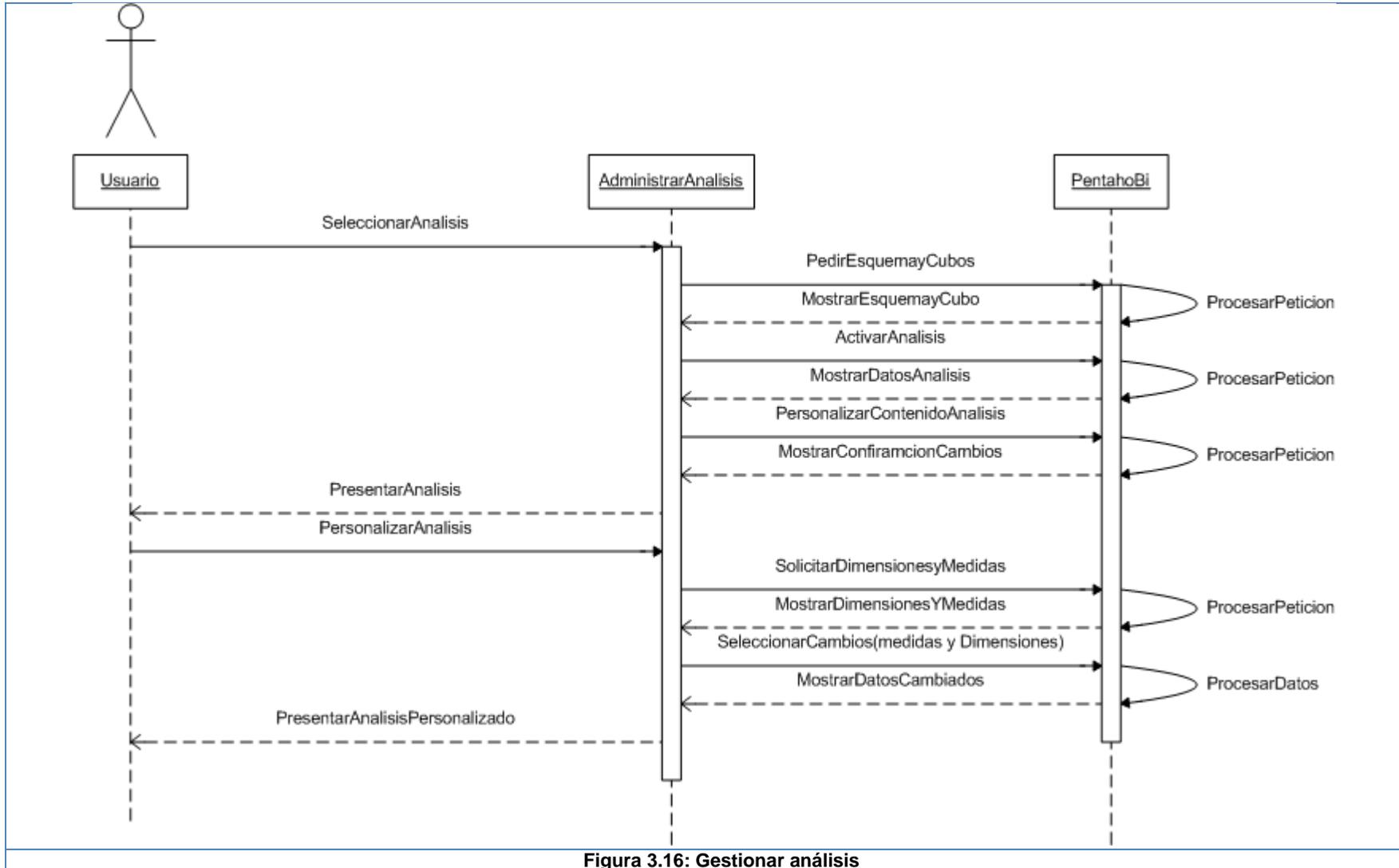


Figura 3.16: Gestionar análisis

3.3.6. Editar Contenido

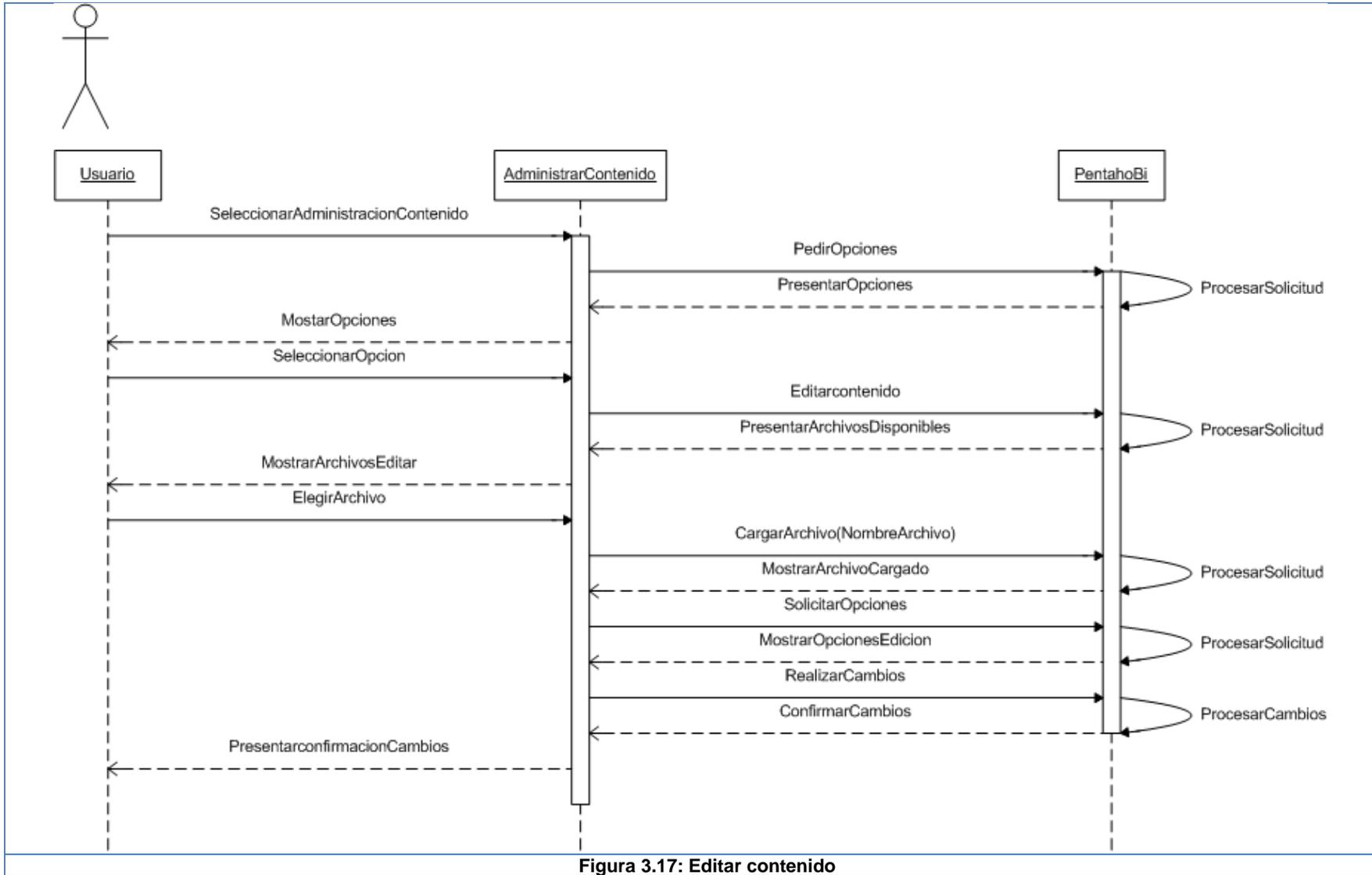


Figura 3.17: Editar contenido

### 3.3.7. Compartir Contenido

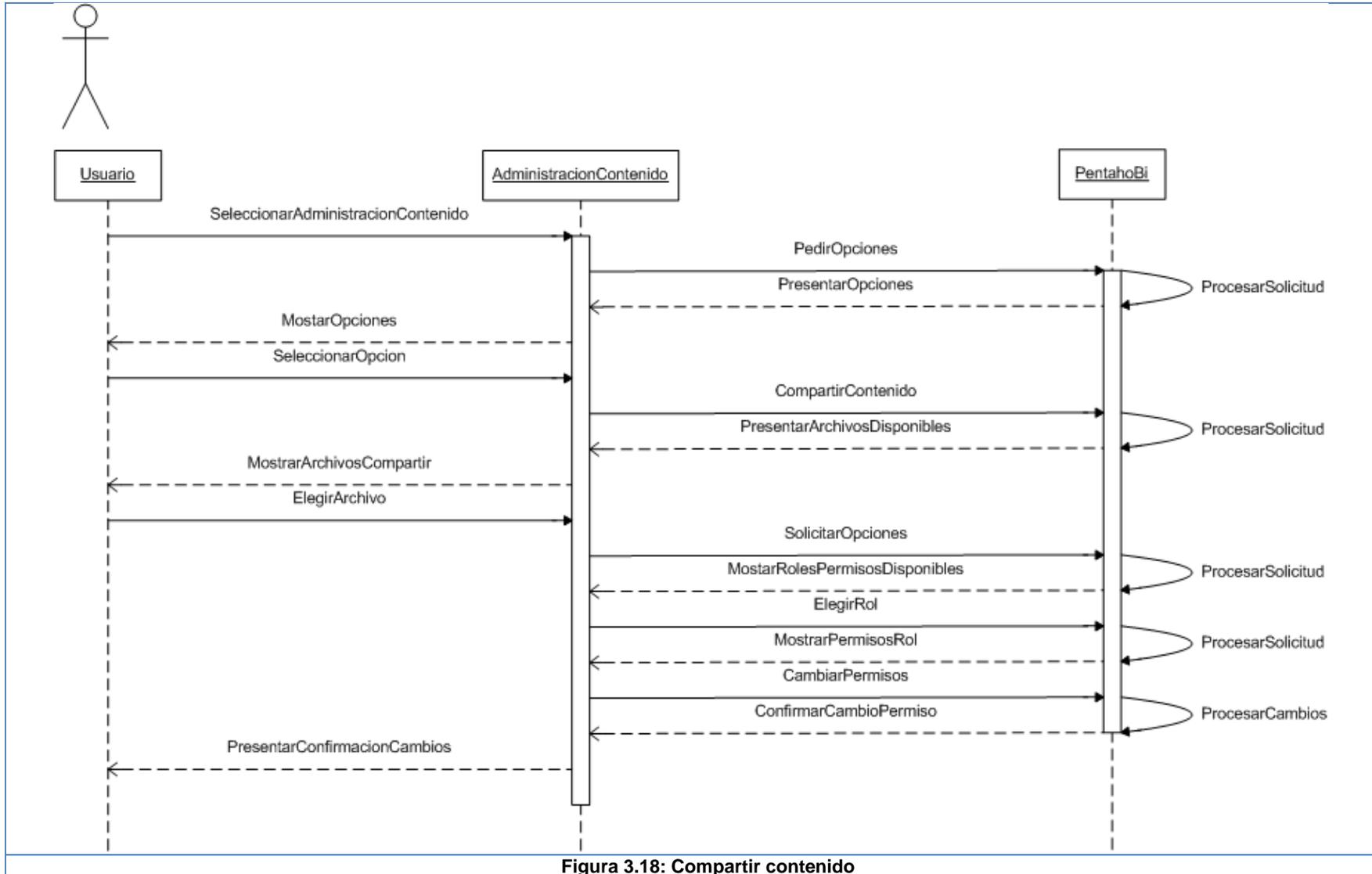


Figura 3.18: Compartir contenido

3.3.8. Gestionar tareas programadas

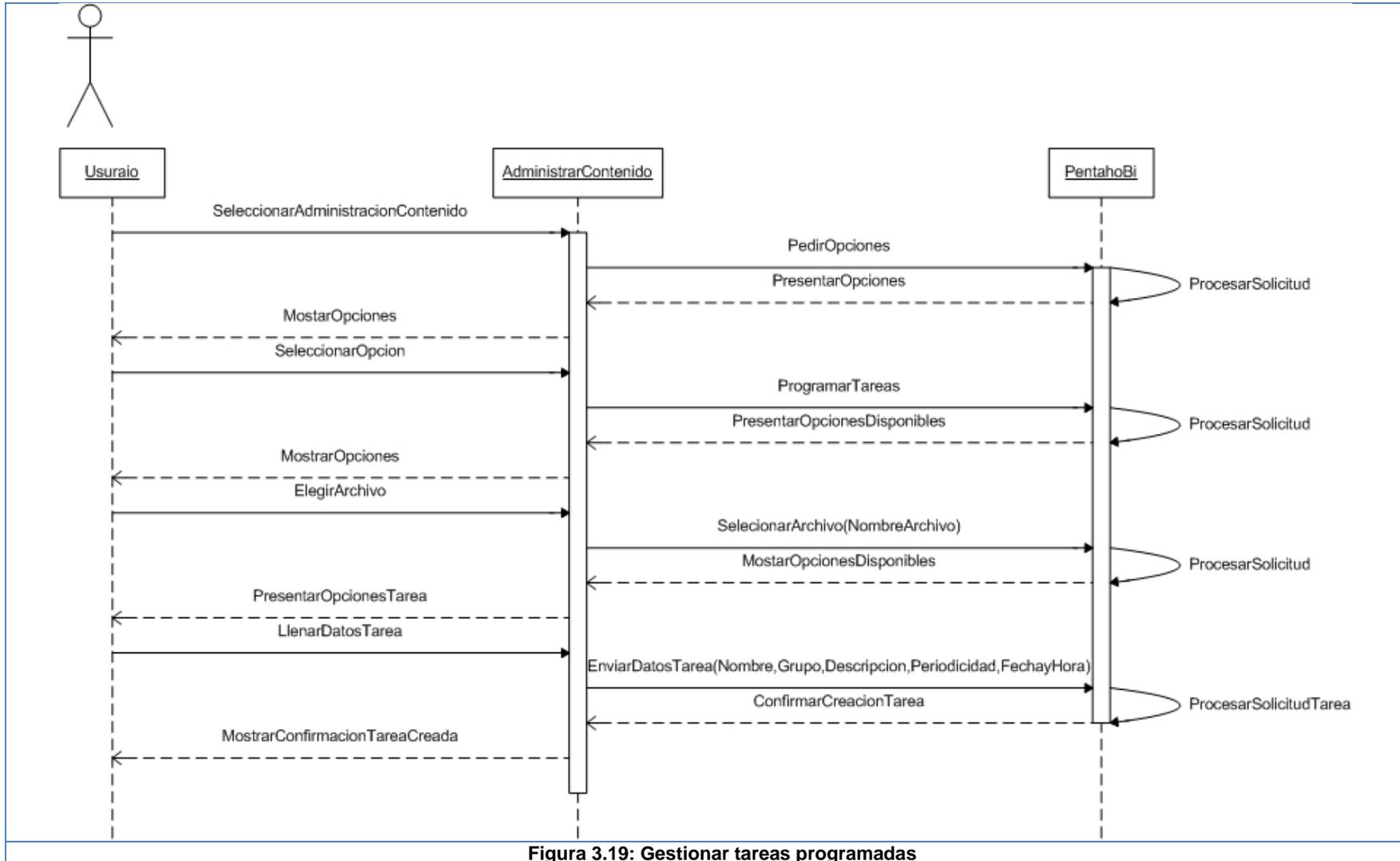


Figura 3.19: Gestionar tareas programadas

### 3.4. REQUERIMIENTOS Y ESPECIFICACIÓN DE USUARIOS

#### 3.4.1. Listado de Requerimientos.

Es de gran importancia la definición de requerimientos funcionales y no funcionales cuando se construye un data warehouse para asegurarse que el sistema será un instrumento de ayuda para alcanzar los objetivos del negocio. Los Requerimientos Funcionales definen que es lo que el sistema hace. Además contienen las características que debería poseer el data warehouse. Los Requerimientos No Funcionales guían y restringen la arquitectura.

Para definir los requerimientos funcionales se debe de comprender las operaciones del negocio, ya que dichos requerimientos son básicamente lo que el usuario espera del Data warehouse. Además de los requerimientos funcionales tenemos los requerimientos no funcionales, los cuales definen elementos como la seguridad, disponibilidad, y rendimiento. Un ejemplo de requerimiento no funcional (para este proyecto es la disponibilidad) podría ser que se espera que el data warehouse se encuentre “arriba y corriendo” siempre en menos de 1 hora al mes.

A continuación se presenta una pequeña descripción de los tipos de requerimientos con los que se trataran en este proyecto:

**Requerimientos Funcionales:** Es la declaración de los servicios que debe proporcionar el sistema de la manera en que este debe reaccionar a entradas particulares y de cómo se debe comportar en situaciones particulares.

**Requerimientos no Funcionales:** Son restricciones de los servicios o funciones ofrecidas por el sistema, incluye restricciones de tiempo, sobre el proceso de desarrollo, estándares, especificaciones técnicas, etc.

**Requerimientos Técnicos:** Es la especificación de todos aquellos aspecto técnicos necesario para la puesta en marcha del sistema, el conocimiento necesario para el recurso humano, como lo es el administrador del sistemas, además de las necesidades de espacio en disco proyectado, características técnicas de los equipos y otros.

**Requerimientos Operativos:** Para que el sistema a desarrollar pueda operar de forma correcta es necesario que cumpla con requisitos de operación que permitan la implementación correcta de este, capacidades operativas necesarias para la operación del sistema, especificaciones técnicas para su operación, etc.

#### 3.4.2. Requerimientos Técnicos

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Página 1 de y	
Autor : Grupo 14		Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo
01	<b>Características de los equipos de desarrollo</b>	Alta	FtD
01.01	El cumplimiento de dichos requerimientos facilitara el desarrollo por parte del equipo de trabajo.		
01.01.01	Procesador Doble Núcleo 1.7Ghz o Superior 2GB de RAM 80GB de Disco Duro o Superior.		

<b>01.01.02</b>	Conexión de Red categoría de cable de red 5e, 802.11g para redes inalámbricas.		
<b>02</b>	<b>Sistemas Operativos de los equipos de desarrollo</b>		
<b>02.01</b>	Es necesario poseer un sistema operativo que facilite el desarrollo del proyecto tomando en cuenta aquellas aplicaciones que se utilizaran para realizar dicha tarea.	<b>Alta</b>	<b>FtD</b>
<b>02.01.01</b>	Windows XP SP2, Windows XP SP3, Windows Vista Home Premium, Windows Vista Bussiness. Se debe considerar sistemas operativos de 64 bits en el caso de direccionar mayor cantidad de memoria.		
<b>03</b>	<b>Características del servidor</b>		
<b>03.01</b>	Características mínimas para el servidor de alojamiento de la solución.	<b>Alta</b>	<b>FtO</b>
<b>03.02</b>	Procesador Xeon 1.7 GHz 1GB de Memoria RAM, 175GB de Espacio en Disco Duro.		
<b>03.03</b>	Sistema Operativo Windows Server 2003.		

### 3.4.3. Requerimientos Operativos

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Página 1 de y	
Autor : Grupo 14		Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo
<b>01</b>	<b>Características de los equipos de los usuarios finales</b>		
<b>01.01</b>	El cumplimiento de dichos requerimientos facilitara la implementación por parte del equipo de trabajo.	<b>Alta</b>	<b>FoD</b>
<b>01.01.01</b>	Procesador Pentium IV 3.06Ghz o Superior 512 MB de RAM 80GB de Disco Duro o Superior.		
<b>01.01.02</b>	Conexión de Red categoría de cable de red 5e, 802.11g para redes inalámbricas.		
<b>02</b>	<b>Sistemas Operativos de los equipos de usuario final</b>		
<b>02.01</b>	Es necesario poseer un sistema operativo que cumplan los requerimientos de acceso a la solución tomando en cuenta aquellas aplicaciones que se utilizaran para realizar dicha tarea.	<b>Alta</b>	<b>FoD</b>
<b>02.01.01</b>	Windows XP, Windows Vista, Linux, MacOS.		
<b>03</b>	<b>Aplicaciones Necesarias en los equipos de usuario final</b>		
<b>03.01</b>	Dado que la solución es basada en WEB, se determina a continuación los navegadores recomendados para acceder al FRONEND del DW	<b>Alta</b>	<b>FoD</b>
<b>03.02</b>	Internet Explorer 6 o superior, Mozilla Firefox, Opera, Safari, Google Chrome.		

**3.4.4. Requerimientos de Desarrollo**

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Pagina 1 de y	
<b>Autor :</b> Grupo 14		<b>Fecha:</b> 29/04/10	
Req.	Especificación	prioridad	Tipo
<b>01</b>	<b>Software utilitario para Implantar la solución en el servidor.</b>	<b>Alta</b>	<b>FdO</b>
<b>01.01</b>	Todo el software utilitario para realizar el análisis y construcción del Data Warehouse.		
<b>01.01.01</b>	SQL Server 2005 Estándar, Pentaho Bi-Server, Apache Tomcat, Hibernate, Quartz, MySQL.		
<b>02</b>	<b>Software utilitario para desarrollar la solución en equipos de trabajo</b>	<b>Alta</b>	<b>FdD</b>
<b>02.01</b>	Software utilizado por el equipo de trabajo para construir y analizar elementos que forman parte de la solución, así como para realizar pruebas y verificar consistencia de la información.		
<b>02.01.01</b>	SQL Server 2005 Estándar, Pentaho Bi-Server, Visual Paradim for UML, Schema Workbench 3.1, MetadataEditor, Apache Tomcat, Hibernate, Quartz, MySQL.		

3.4.5. Requerimientos No Funcionales

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Pagina 1 de y	
Autor : Grupo 14		Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo
<b>01</b>	<b>Gráficos en los reportes</b>	<b>Baja</b>	<b>NoD</b>
<b>01.01</b>	Los gráficos facilitan el análisis de información y ayuda a comprender el comportamiento de los datos, resumir información y brindar reportes más atractivos. El sistema debe brindar la opción de generar gráficos de datos.		
<b>01.01.01</b>	Existirá una limitación en la variedad de gráficos disponibles. Por ser herramienta de software libre se tendrá dichas limitantes.		
<b>01.01.02</b>	Exportar datos a hojas de cálculo. Con los datos el usuario podrá realizar sus propios gráficos según lo necesite.		
<b>02</b>	<b>Tecnología del Data Warehouse</b>	<b>Alta</b>	<b>NtO</b>
<b>02.01</b>	El tipo de tecnología de data warehouse será basada en la web. Combinando la base de datos sql server con la interface de presentación Pentaho.		
<b>02.01.01</b>	Debe permitir la conexión vía JDBC para su relación directa con la interface de presentación Pentaho. Dicha conexión debe ser por medio de JDBC ya que Pentaho es un software desarrollado con java y montado sobre un servidor web APACHE TOMCAT.		
<b>03</b>	<b>Seguridad del Data Warehouse</b>	<b>Alta</b>	<b>NtO</b>
<b>03.01</b>	Seguridad del Data Warehouse debe que ser prioridad ya que es información confidencial.		
<b>03.01.01</b>	Control de acceso al Data Warehouse Modos de autenticación Privilegios administrativos Políticas de password		
<b>03.01.02</b>	Deben existir políticas de seguridad al acceso físico al Data Warehouse, siendo estas implementadas por la Dirección de Sistemas de Información.		
<b>03.02</b>	El detalle y tipo de información a la que tendrá acceso el usuario depende del perfil que este tiene asignado.		
<b>04</b>	<b>Estandarización de salidas</b>	<b>Media</b>	<b>NoD</b>
<b>04.01</b>	La presentación de los datos debe de incluir un estándar en el tipo de letra, permitiendo variaciones en casos en los que se quiera resaltar cierta información.		
<b>04.02</b>	La presentación de plantillas para los diferentes reportes, deberán ser de acuerdo al estándar utilizado dentro de la DC.		

<b>05</b>	<b>Accesos de usuarios avanzados</b>		
<b>05.01</b>	Los administradores del sistema podrán acceder al data warehouse por medio de consultas SQL y MDX.	<b>Media</b>	<b>NiD</b>
<b>05.02</b>	Se deberá proporcionar la herramienta para realizar este tipo de acceso.		
<b>06</b>	<b>Tiempo de Espera</b>		
<b>06.01</b>	El tiempo de respuesta para cualquier tipo de consulta y reportes debe tomar como máximo 60 segundos.	<b>Baja</b>	<b>NiD</b>
<b>07</b>	<b>Usuarios estimados</b>		
<b>07.01</b>	El soporte para los usuarios estimados es de 100 a 120 de estos el 20% son considerados usuarios frecuentes los demás son usuarios ocasionales.	<b>Baja</b>	<b>NoD</b>
<b>07.01.01</b>	El sistema debe soportar la carga de transacciones de por lo menos 20 usuarios realizando análisis.		
<b>07.01.02</b>	El 80% de los usuarios realizaran consulta de de información de forma general no llegando a niveles profundos de detalle		
<b>08</b>	<b>Disponibilidad</b>		
<b>08.01</b>	El data warehouse deberá estar disponible en horas hábiles para poder ser consultado por los usuarios.	<b>Alta</b>	<b>NoD</b>
<b>08.02</b>	Los periodos de tiempos inactivos no deberán sobrepasar 1 hora al semanal.		
<b>09</b>	<b>Consistencia de datos e información</b>		
<b>09.01</b>	Es preferible que no se realicen cambios bruscos en la aplicación ni en el esquema de la base de datos, incluyendo triggers o procedimientos almacenados.	<b>Alta</b>	<b>NiO</b>
<b>10</b>	<b>Software requerido</b>		
<b>10.01</b>	Se debe de utilizar Microsoft SQL server para construir el data warehouse.	<b>Media</b>	<b>NtD</b>
<b>10.02</b>	En la capa de presentación sugieren el uso de una herramienta de software libre.		
<b>11</b>	<b>Control de acceso de información</b>		
<b>11.01</b>	Se debe de auditar el acceso y las actividades críticas que realicen los usuarios a fin de reducir la fuga de información.	<b>Media</b>	<b>NoD</b>
<b>11.01.01</b>	El control de acceso debe de quedar registrado almacenando los intentos de información no autorizada.		
<b>11.01.02</b>	El control de los accesos al data warehouse como usuario administrador debe estar controlado por medio de una bitácora que almacenes las actividades realizadas a iniciar sesión		

### 3.4.6. Requerimientos No Funcionales Técnicos

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Página 1 de y	
<b>01</b>	<b>Accesibilidad externa al sistema.</b>		
<b>01.01</b>	El sistema encargado de la presentación del data warehouse no deberá ser accesible desde fuera de la red de la defensoría del consumidor.	<b>Baja</b>	<b>NtD</b>
<b>01.01.01</b>	Deberá permitir la configuración de la seguridad en cuanto a accesos locales y remotos.		
<b>01.01.02</b>	Se deberá llevar un registro histórico de accesos de usuario al sistema.		
<b>02</b>	<b>Requerimientos técnicos mínimos</b>		
<b>02.01</b>	Los criterios o características mínimas de los equipos de los usuarios finales para que funcione adecuadamente la solución.	<b>Media</b>	<b>NtD</b>
<b>02.02</b>	Sistema operativo: Windows XP		
<b>02.03</b>	Navegador: IE Explorer 6 o superior		
<b>02.04</b>	Memoria RAM: 512 MB		
<b>02.05</b>	Procesador: Intel Pentium D 820 o móvil AMD equivalente		
<b>02.06</b>	Resolución de pantalla: 1024x768		
<b>03</b>	<b>Respaldo de la información</b>		
<b>03.01</b>	El respaldo de la información del data warehouse actual deberá realizarse antes de que se inicie el proceso ETL.	<b>Alta</b>	<b>NtO</b>
<b>03.01.01</b>	Los backups deben estar almacenados en el servidor por lo menos de una semana, y los registros anteriores pueden almacenarse en medios externos.		
<b>03.01.02</b>	El proceso ETL debe de iniciar estrictamente hasta que termine de realizar el backup actual		
<b>04</b>	<b>Acceso remoto para el desarrollo</b>		
<b>04.01</b>	Creación de una VPN para acceder desde fuera de las las instalaciones de la defensoría del consumidor, este recurso será necesario para la realizar pruebas en el servidor de desarrollo.	<b>Baja</b>	<b>NtD</b>
<b>04.02</b>	Se debe de poseer disponibilidad de equipo local para realizar pruebas durante el desarrollo.		

### 3.4.7. Descripción de los perfiles de Usuario

A continuación se detallan los perfiles por default que se definirán para el DW, estos perfiles pueden ser modificados eliminados y agregados según decisiones de la organización, pero el compromiso como grupo será el definir los pasos para el mantenimiento de los perfiles.

No. Perfil	Nombre del perfil	Observaciones
1	Técnico	Encargado de realizar tareas como la generación de reportes y consulta de información de casos.
2	Coordinador	Encargado de la verificación del desempeño del grupo a fin de distribuir la carga de trabajo y coordinar a grupos de trabajo.
3	Director/Gerente	Su principal función es analizar información y coordinar las direcciones por las que se encuentra formada la institución.
4	Administrador	Encargado directo del manejo del sistema compuesto para el funcionamiento del Data Warehouse.
5	Presidencia	Perfil cuya principal función es la consulta de información clave; en su con estructuras ya definidas que ayudan a tomar decisiones.
6	Director Regionales	Perfil General el cual se encarga del control de los incidentes y análisis de información generada por parte del personal operativo de cada Dirección.
7	Directora Nacional de CSC	Encargado del control del desempeño de todas las oficinas de de la Defensoría del Consumidor.
8	Gerencia Informática	Perfil creado para los integrantes de la gerencia informática los cuales tendrán acceso a toda la aplicación debido a la función que se desempeñan como soporte a las decisiones informáticas.
9	Anfitrión	Perfil de usuario en el cual se incluyen las funciones necesarias para realizar análisis en el modulo de Anfitrión.
10	Call Center	Perfil de usuario en el cual se incluyen las funciones necesarias para realizar análisis en el modulo de Call Center.
11	CSC	Perfil de usuario en el cual se incluyen las funciones necesarias para realizar análisis en el modulo de CSC.
12	Tribunal Sancionador	Perfil de usuario en el cual se incluyen las funciones necesarias para realizar análisis en el modulo de Tribunal Sancionador.
13	Verificación y Vigilancia	Perfil de usuario en el cual se incluyen las funciones necesarias para realizar análisis en el modulo de Verificación y Vigilancia.

### 3.5. ITERACIÓN 1: CENTRO DE SOLUCIÓN DE CONTROVERSIAS

Se encarga de recibir denuncias de consumidores.

#### 3.5.1. Requerimientos Funcionales Para Centro de Solución de Controversias

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Página 1 de y	
Autor : Grupo 14		Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo
<b>01</b>	<b>Analizar proceso de denuncias</b>	Alta	FoO
<b>01.01</b>	Permitir al usuario analizar el proceso de denuncias por centro de solución, edad de consumidor y tipo de denuncia.		
<b>01.01.01</b>	El conocer la recepción por centro de solución le permite a la defensoría conocer la carga de trabajo que se genera en las sucursales y así poder medir su desempeño.		
<b>01.01.02</b>	Los rangos para el análisis vendrán establecidos desde el sistema transaccional para poder facilitar este análisis		
<b>01.01.03</b>	Los tipos de denuncias se agruparan por sectores o cualidades que permitan relacionar los tipos de denuncias por grupos.		
<b>02</b>	<b>Facilitar el análisis los casos etapas</b>	Alta	FoO
<b>02.01</b>	Analizar cuantos casos se encuentran en cada etapa y poder conocer las sub-etapas.		
<b>02.01.01</b>	Se le debe presentar al usuario la cantidad de denuncias en cada etapa a fin de que puedan profundizar en su análisis.		
<b>02.01.02</b>	Al tener los casos clasificados por etapas, se debe de brindar un desglose de dichos casos llegando a hasta las sub etapas.		
<b>03</b>	<b>Clasificar a los consumidores</b>	Media	FoD
<b>03.01</b>	El usuario necesita consultar la información de los consumidores por clases comunes en las que se pueda analizar cada clase según el sector al que pertenezca.		
<b>03.01.01</b>	La clasificación del consumidor va a depender de la necesidad del usuario pero en términos generales se debe clasificar ubicación geográfica, sexo del consumidor.		
<b>04</b>	<b>Calcular valores estadísticos</b>	Media	FoD
<b>04.01</b>	El usuario necesita valores calculados, por lo cual solicita la capacidad de crear sus propias formulas.		
<b>04.01.01</b>	Para algunos casos se debe permitir al usuario crear campos calculados a fin de que complementar la información brindada por los reportes.		
<b>04.01.02</b>	El cálculo de variables comunes se debe de realizar previamente en el data warehouse.		

<b>05</b>	<b>Almacenamiento de datos</b>	<b>Media</b>	<b>FoD</b>
<b>05.01</b>	Se requiere la disponibilidad completa de las transacciones recibidas en el sistema actual.(SARA)		
<b>05.01.01</b>	La extracción de información para crear el data warehouse se realizará principalmente con las transacciones que involucren a los módulos contemplados.		
<b>05.01.02</b>	Los datos contenidos en el data warehouse deben de tener lo máximo 2 días de desfase con respecto a la fecha de las transacciones más recientes del sistema transaccional		
<b>06</b>	<b>Creación de perfiles de usuario.</b>	<b>Alta</b>	<b>FoO</b>
<b>06.01</b>	El sistema permitirá a los usuarios visualizar solo el área que compete. Este acceso a la información será manejado según lo especificado en el sistema actual.		
<b>06.01.01</b>	Los reportes deben estar filtrados por oficina.		
<b>06.01.02</b>	La granularidad de la información dependerá del perfil de usuario con el que se encuentre trabajando en el sistema.		
<b>06.01.03</b>	Los usuarios que tendrán acceso estarán regidos por los perfiles de usuario establecidos para el sistema actual.		
<b>07</b>	<b>Facilitar la consulta de información</b>	<b>Alta</b>	<b>FoO</b>
<b>07.01</b>	Reportes OLAP que permitan realizar análisis profundo utilizando “drill-down” con los siguientes ítems principales:		
<b>7.01.01</b>	Clasificación de casos por fecha Casos anuales Casos mensuales Casos diarios Casos cerrados en 1 día		
<b>7.01.02</b>	Clasificación de los casos: Sectores Motivos Tipo Denuncia Asesoría Gestión Derivación Información Orientación		
<b>7.01.03</b>	Clasificación por sector geográfico Municipios Departamento		
<b>7.01.04</b>	Clasificación de Montos Montos reclamados Montos Recuperados		
<b>7.01.05</b>	Clasificación de casos por proveedores		
<b>7.01.06</b>	Tipo de solución que se ha brindado, considerando las fechas de las etapas para evaluar en qué momento llego al fin el caso.		

<b>7.01.07</b>	Razón de cierre de cada caso, ¿Por qué razón se cerró el caso?		
<b>08</b>	<b>Interfaz de usuario final versátil</b>		
<b>08.01</b>	La interfaz debe permitir almacenar los reportes generados para poderlos usar posteriormente.	<b>Media</b>	<b>FoD</b>
<b>08.01.01</b>	La interfaz debe proporcionar facilidad para exportar los reportes generados a otros tipos de archivos como PDF y Hojas de Excel, así como también la impresión de los reportes directamente desde la interfaz.		
<b>08.01.02</b>	Debe contar con un asistente que facilite la selección de los datos a incluir en el reporte para la creación del mismo.		
<b>08.01.03</b>	Permitir la navegación y exploración de los datos.		
<b>09</b>	<b>Facilitar el análisis</b>		
<b>09.01</b>	El sistema podrá mostrar figuras y diagramas los cuales podrán ser impresos, la mayoría de usuarios coinciden que necesitan gráficos y tablas, sin embargo si esta opción no está disponible se podrá exportar la información a archivos de Excel donde podrán hacer sus propios gráficos. Opción para exportar a Excel (o cvs).	<b>Alta</b>	<b>FoD</b>
<b>09.01.01</b>	Realizar Drill-down sobre los datos almacenados.		
<b>09.01.02</b>	Debe facilitar el cruce de la información almacenada.		
<b>09.01.03</b>	Facilitar el análisis de los datos por medio de gráficos.		
<b>09.01.04</b>	Presentar los datos multidimensionales y dejar a elección del usuario la selección de medidas y dimensiones a explorar para realizar el análisis.		
<b>10</b>	<b>Permitir el análisis según centro de solución</b>		
<b>10.01</b>	Se deberá considerar este análisis según rango de tiempo	<b>Alta</b>	<b>FoO</b>
<b>10.01.01</b>	Conteo de casos recibidos por centro de solución.		
<b>11</b>	<b>Análisis de caso según estado</b>		
<b>11.01</b>	Según el estado se deben de agrupar cada etapa.	<b>Media</b>	<b>FoD</b>
<b>11.01.01</b>	Según el estado en que se encuentre cada etapa deberá permitir detallar la cantidad de casos por sub-etapa.		
<b>12</b>	<b>Análisis del caso según Institución de derivación</b>		
<b>12.01</b>	Se debe de incluir solamente los casos que han sido ingresados por las derivaciones.	<b>Media</b>	<b>FoD</b>
<b>13</b>	<b>Análisis por cruce de fechas</b>		
<b>13.01</b>	El sistema deberá brindar la consulta de información en las que sus variables dependan plenamente de las fechas.	<b>Media</b>	<b>FoD</b>
<b>13.01.01</b>	Se deberá analizar las fechas de inicio y cierre por cada etapa del caso.		
<b>14</b>	<b>Análisis de Desempeño por etapa y sub etapa</b>		
<b>14.01</b>	Análisis de desempeño por etapa.	<b>Medio</b>	<b>FoD</b>
<b>14.02</b>	Análisis de desempeño por sub etapa.		

### 3.5.2. Especificación de Requerimientos

#### 3.5.2.1. Crear Reporte Analítico



Figura 3.20: Diagrama de caso de uso para reportes analíticos

<p><b>Caso de uso 1 Crear Reporte Analítico</b></p> <p><b>Personal Involucrado e interés:</b></p> <p>Usuario: Técnico, Director/Gerente, Administrador, Presidencia, Técnico Web, Estudios, Director regional, Director nacional de CSC y Gerencia informática.</p> <p>Usuario: Elaborar un análisis detallado del centro de solución de controversias para su posterior publicación o almacenamiento.</p> <p><b>Precondiciones:</b> Haber iniciado sesión en el sistema</p>
--

<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona medidas.</li> <li>6. El Sistema muestra datos de centro de solución, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A,B,C,D,E,F.</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis</li> <li>10. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>11. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Realizar análisis utilizando MDX (Usuarios Expertos)</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción MDX.</li> <li>2. El sistema muestra opción de generación de consulta MDX.</li> <li>3. Usuario construye la consulta, utilizando consulta valida.</li> <li>4. El sistema realiza la construcción del espacio de trabajo en base a la consulta generada.</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>B. Configuración de tabla OLAP</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Configurar tabla OLAP.</li> <li>2. El sistema muestra ventana de configuración OLAP</li> <li>3. El usuario selecciona las opciones de configuración</li> <li>4. El sistema aplica configuraciones</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>C. Seleccionar opciones de las vistas de los datos</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Mostrar Padres.</li> <li>2. El sistema muestra las referencias fuentes de los datos mostrados.</li> <li>3. Sino el usuario selecciona la opción ocultar repeticiones.</li> <li>4. El sistema muestra solo los datos únicos (unique) sin repeticiones.</li> <li>5. Sino el usuario el usuario selecciona mostrar propiedades.</li> <li>6. El sistema muestra propiedades de los datos.</li> <li>7. Sino el usuario selecciona suprimir columnas vacías.</li> <li>8. El sistema elimina de la consulta aquellos datos que contengan valores nulos.</li> <li>9. Sino el usuario selecciona la opción intercambiar ejes.</li> <li>10. El sistema genera un análisis roll-up.</li> <li>11. El usuario analiza los datos.</li> </ol>

**D. Generación y configuración de gráfico**

1. El usuario selecciona la opción configuración de gráfico.
2. El sistema muestra propiedades de gráfico.
3. El usuario configura opciones de gráfico, colores, leyendas, orientación y tipo de gráfico.
4. El sistema adecua el gráfico a los valores que se muestran en el espacio de trabajo.
5. El usuario selecciona opción de generar gráfico.
6. El sistema muestra gráfico.
7. El usuario analiza los datos.

**E. Seleccionar el Destino del análisis**

1. El usuario selecciona opción de configuración de impresión.
2. El sistema muestra panel de configuración de impresión.
3. El usuario selecciona opciones de configuración tales como tema de reporte, orientación, tamaño del papel.
4. El sistema configura opciones de reporte analítico.
5. El usuario publica el reporte en formato PDF, EXCEL, CSV o lo almacena en el sistema.
6. El sistema procesa el reporte y lo publica dependiendo el método seleccionado.

**F. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.5.2.2. Crear Reporte AD HOC

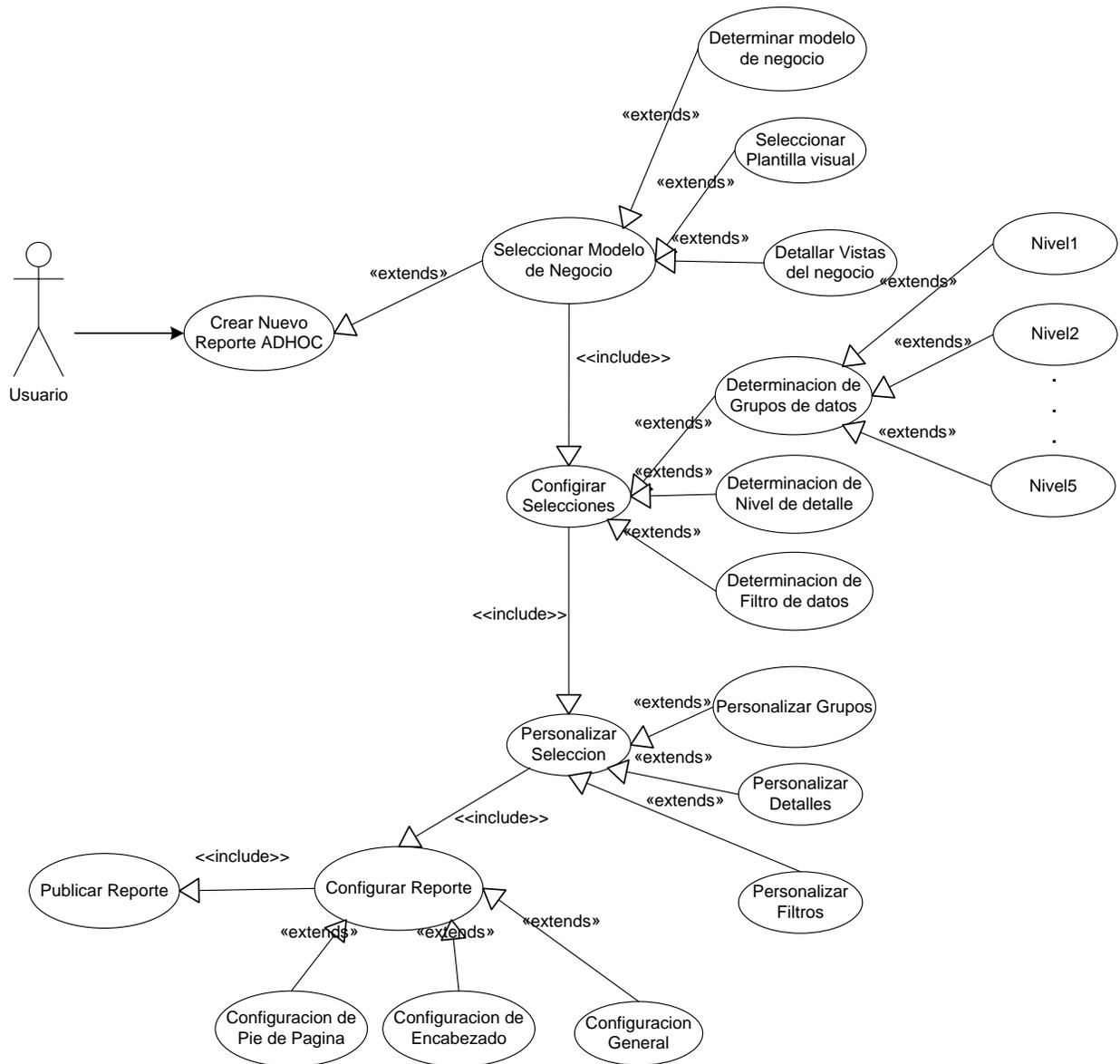


Figura 3.21: Diagrama de caso de uso para reportes AD HOC

<b>Caso de uso 2 Crear Reporte AD HOC</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Director/Gerente, Administrador, Presidencia, Técnico Web, Estudios, Director regional, Director nacional de CSC y Gerencia informática.  Usuario: Elaborar de reportes temporales para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 11, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14

**Flujo Básico**

1. El Usuario selecciona "Nuevo Reporte".
2. El sistema muestra los modelos de negocio disponibles.
3. Usuario selecciona el modelo de negocio con el que desea trabajar.
4. Sistema muestra las vistas de negocio disponibles para el modelo seleccionado.
5. El usuario procede a la siguiente fase de selección de reporte
6. El sistema muestra variables con las que se puede agrupar aplicar filtros y seleccionar niveles de detalle de los grupos definidos
7. El usuario selecciona los grupos, filtros y niveles de detalle con las que desea realizar el reporte
8. El usuario procede a la siguiente fase de personalización de selecciones
9. El sistema muestra los valores de los niveles de detalle, grupos y filtros con los que se desea generar el reporte.
10. El sistema almacena la configuración de los datos.
11. El usuario procede a la siguiente fase de configuración de reporte
12. El sistema muestra panel de configuración de reporte.
13. El usuario selecciona las opciones generales del reporte, tipo de papel orientación encabezado y pie de página.
14. El sistema almacena la configuración
15. El usuario almacena el reporte ad hoc generado para futuras consultas
16. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.

**Flujo Alternativo****A. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.5.2.3. Análisis de Casos por CSC

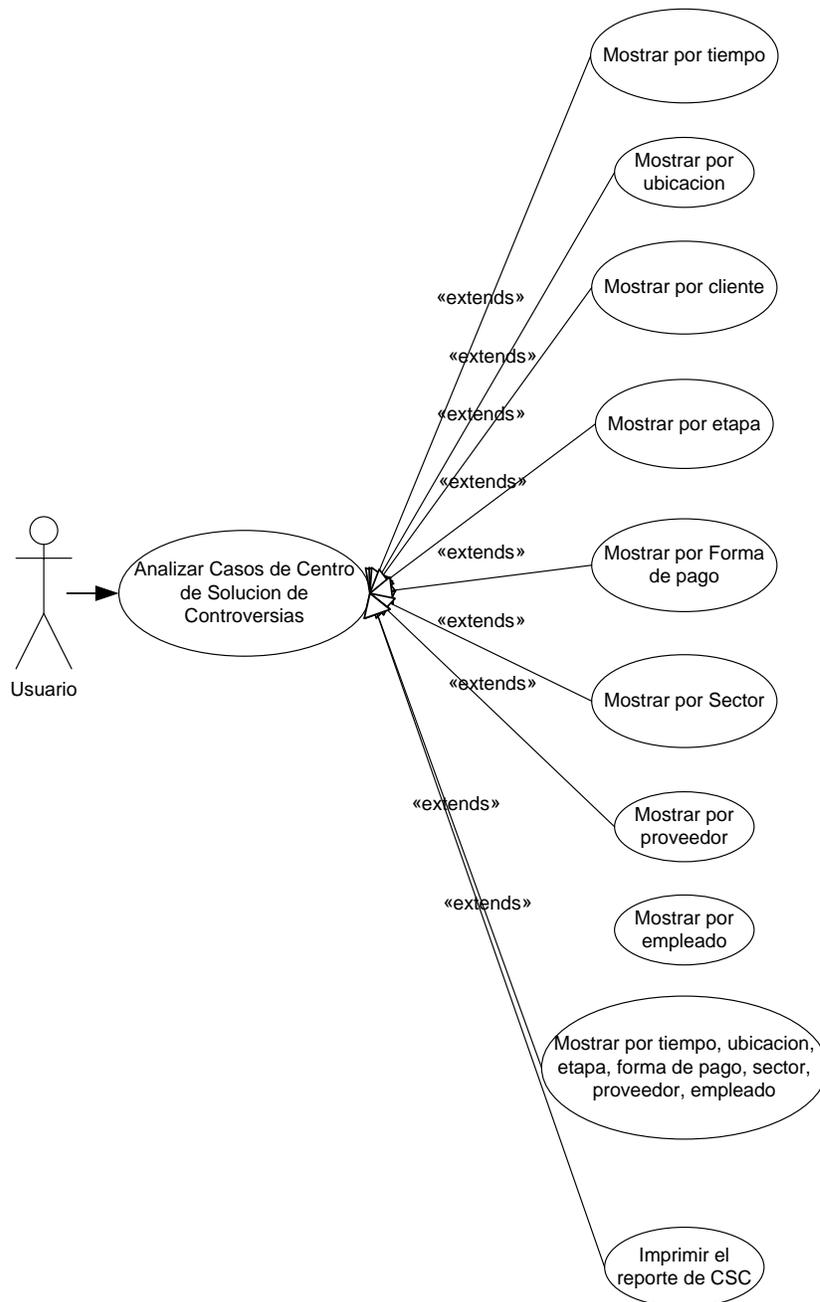


Figura 3.22: Diagrama de caso de uso para análisis de los casos en el CSC

<b>Caso de uso 3 Análisis de casos por CSC</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Director/Gerente, Administrador, Presidencia, Técnico Web, Estudios, Director regional, Director nacional de CSC y Gerencia informática.  Usuario: Elaborar un análisis detallado del centro de solución de controversias para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema

<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona medida centrosolucion.</li> <li>6. El Sistema muestra datos de centro de solución, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o el caso alterno A.</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis con medidas tales como: cantidad de caso, monto recuperado, y valor del producto. Y dimensiones como: tiempo, ubicación, etapa, forma de pago sector, proveedor, empleado.</li> <li>10. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>11. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Cancelar la operación</b></p> <ol style="list-style-type: none"> <li>1. El usuario Cierra el sistema</li> <li>2. El sistema cancela las operaciones y finaliza sesión de usuario.</li> </ol>

3.5.2.4. Exploración de Datos

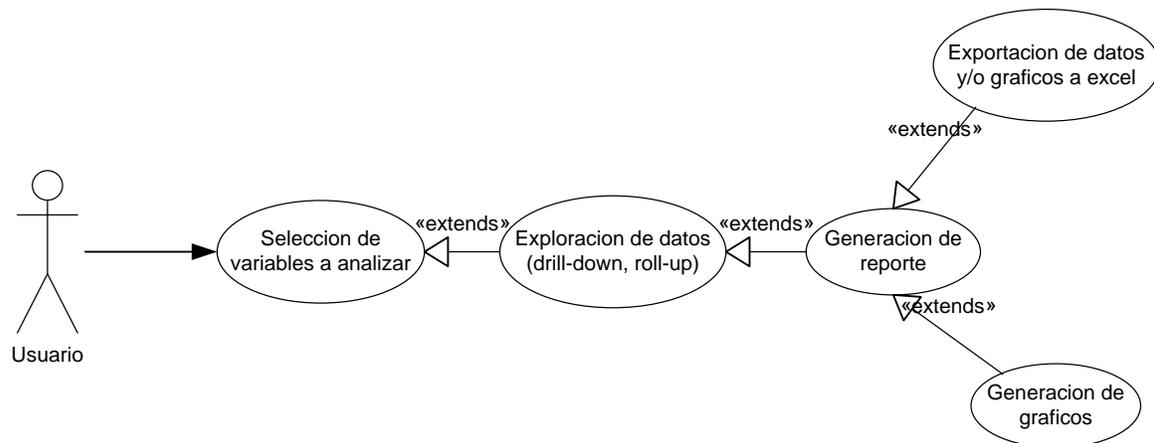
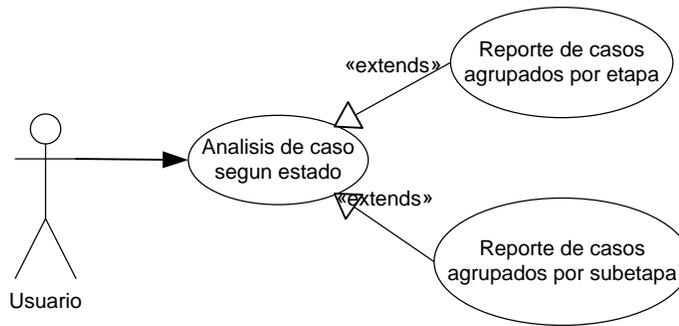


Figura 3.23: Diagrama de caso de uso para exploración de datos

<b>Caso de uso 4 Exploración de datos</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Director/Gerente, Administrador, Presidencia, Técnico Web, Estudios, Director regional, Director nacional de CSC y Gerencia informática.  Usuario: Elaborar un análisis detallado del centro de solución de controversias para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 09
<b>Flujo Básico</b> <ol style="list-style-type: none"> <li>1. Flujo básico de caso de uso 1 al paso 6.</li> <li>2. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis con medidas tales como: cantidad de caso, monto recuperado, y valor del producto. Y dimensiones como: tiempo, ubicación, etapa, forma de pago sector, proveedor, empleado.</li> <li>3. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>4. El usuario realiza una exploración haciendo uso de la técnica drill-down o roll-up</li> <li>5. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>6. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso</li> </ol>
<b>Flujo Alternativo</b> <b>A. Cancelar la operación</b> <ol style="list-style-type: none"> <li>1. El usuario Cierra el sistema</li> <li>2. El sistema cancela las operaciones y finaliza sesión de usuario.</li> </ol>

**3.5.2.5. Análisis según Estado del Caso**



**Figura 3.24: Diagrama de caso de uso para análisis de casos según estado del caso**

<b>Caso de uso 5 Análisis según estado del caso</b>
<p><b>Personal Involucrado e interés:</b></p> <p>Usuario: Técnico, Director/Gerente, Administrador, Presidencia, Técnico Web, Estudios, Director regional, Director nacional de CSC y Gerencia informática.</p> <p>Usuario: Elaborar un análisis detallado del centro de solución de controversias para su posterior publicación o almacenamiento.</p>
<b>Precondiciones:</b> Haber iniciado sesión en el sistema
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<p><b>Referencias Cruzadas</b></p> <p><b>Requerimientos: 11</b></p>
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. Flujo básico de caso de uso 1 al paso 6.</li> <li>2. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis con medidas referente a etapa y subetapa en la que se encuentra el caso.</li> <li>3. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>4. El usuario realiza una exploración haciendo uso de la técnica drill-down o roll-up</li> <li>5. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>6. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Cancelar la operación</b></p> <ol style="list-style-type: none"> <li>1. El usuario Cierra el sistema</li> <li>2. El sistema cancela las operaciones y finaliza sesión de usuario.</li> </ol>

### 3.5.3. Modelado Iteración 1 Centro de Solución de Controversias

#### 3.5.3.1. Esquema Conceptual del Data Warehouse

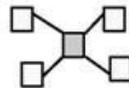
En Tabla 1.3 se encuentra la definición de la nomenclatura UML utilizada para la definición de los estereotipos del modelado conceptual.

##### Nivel 1. Definición del Modelo

Centro de Solución de Controversias:

Especialización del Negocio para análisis de los casos.

**Hecho principal:** El Estudio del Caso



Centro de Soluciones de Controversias

Figura 3.25: Definición modelo “Centro de Soluciones de Controversias”

##### Nivel 2 Definición del esquema estrella

Paquete de Hecho: Caso

Paquetes Dimensionales

Consumidor: Datos relevantes del consumidor, nombres, sexo, ubicación.

Categoría: Se refiere al tipo de denuncia por el cual el consumidor solicita ayuda.

Proveedor: Generalmente considerado como el ente al cual se le demanda algo.

Etapas: El proceso el cual lleva la denuncia,

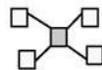
Ubicación: Contiene la ubicación geográfica del país.

Empleado: Catalogo de empleados de la DC

Forma de Pago: Forma de cómo se le remunera al consumidor su demanda.

Tiempo: Control de tiempo en años, días, cuartiles, meses, semanas.

Centro de Solución: Nombre del centro de solución en el que se recibe la denuncia.



Centro de Soluciones de Controversias

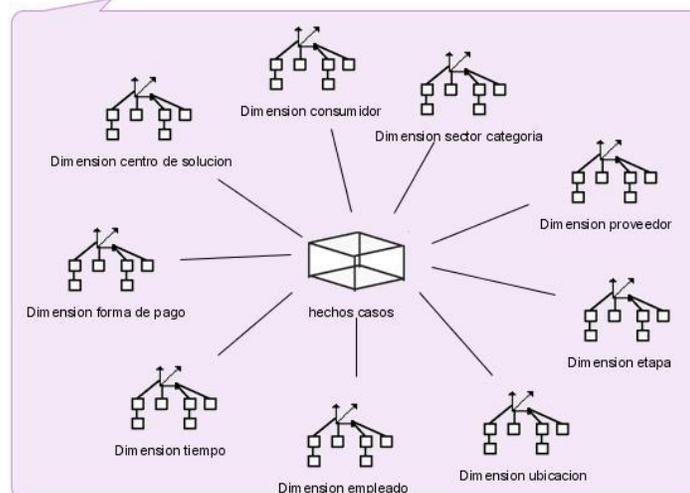


Figura 3.26: Definición del esquema estrella para “Centro de Soluciones de Controversias”

**Nivel 3: Definición de Dimensiones/Hecho.**

Especificación de atributos.

Dimensiones candidatas a roll up y drill-down: tiempo y ubicación.

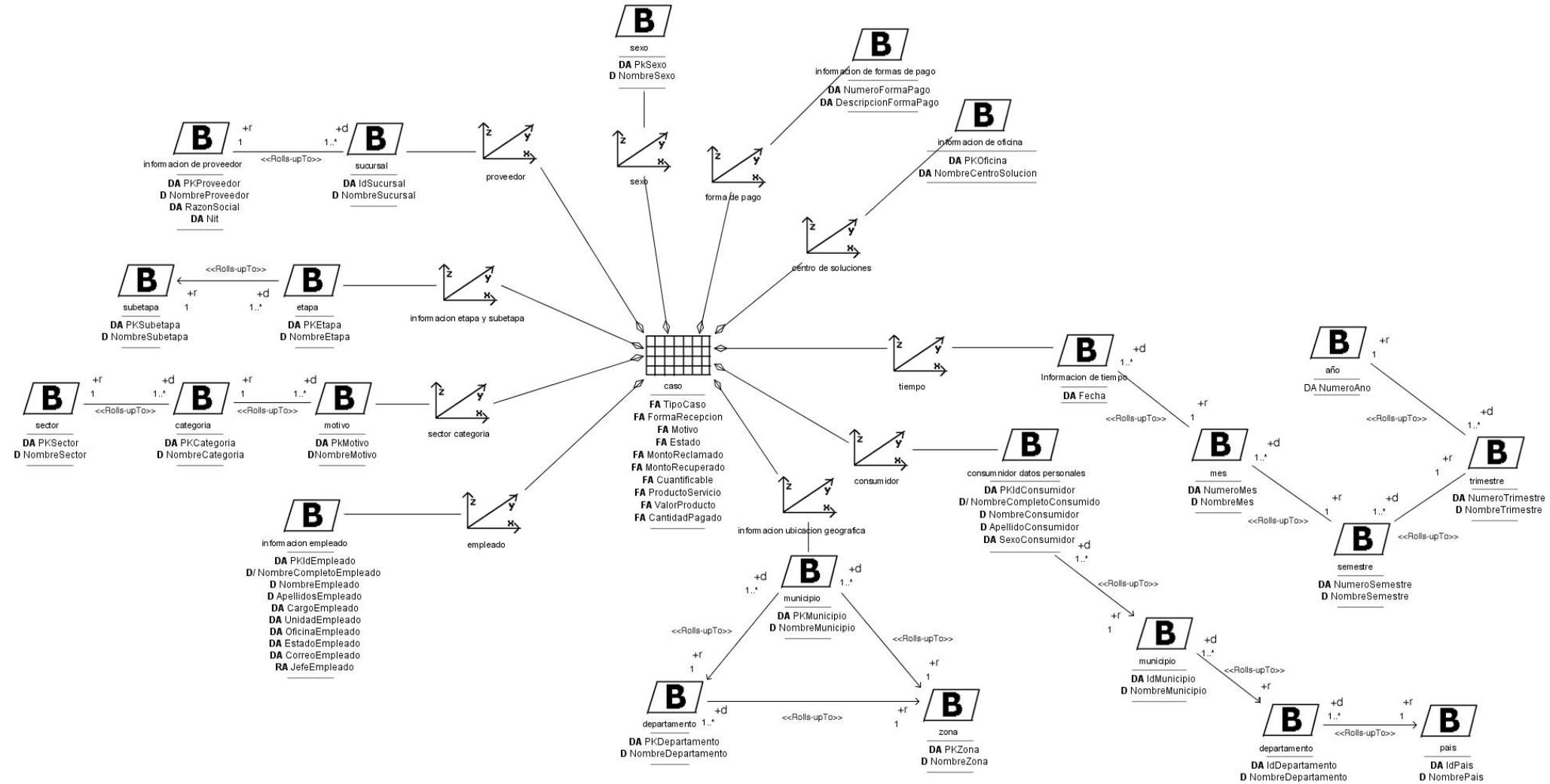


Figura 3.27: Definición de las dimensiones y hechos “Centro de Soluciones de Controversia”

3.5.3.2. Esquema lógico del Data Warehouse

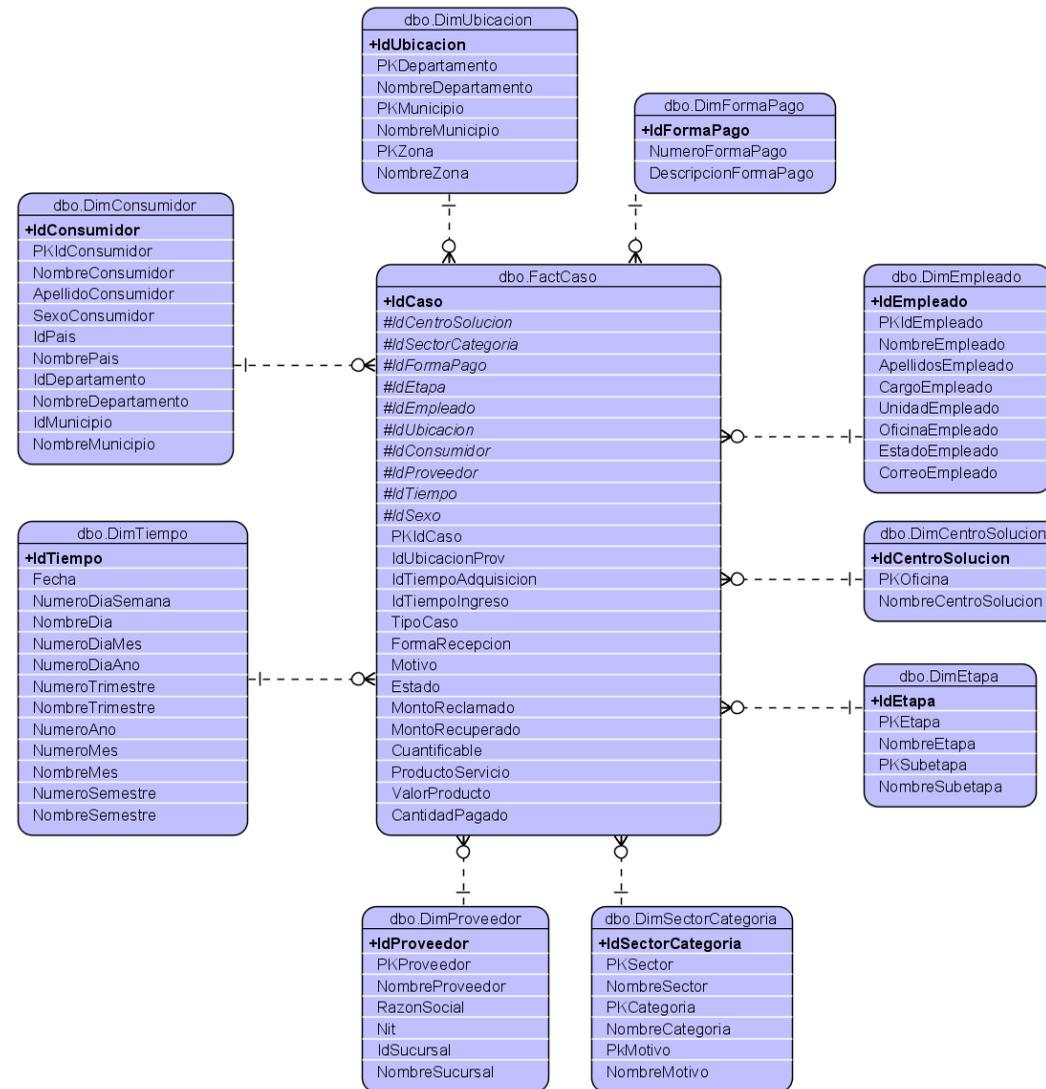


Figura 3.28: Estructura lógica del esquema estrella para el Centro de Solución de Controversias

### 3.5.3.3. Diccionario de datos de las dimensiones y tablas de hecho

DIMCONSUMIDOR		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDCONSUMIDOR	INT	Identificador del consumidor en el DW
NOMBRECONSUMIDOR	VARCHAR(50)	Nombre del consumidor
APELLIDOCONSUMIDOR	VARCHAR(50)	Apellidos del consumidor
SEXOCONSUMIDOR	CHAR(10)	Sexo del consumidor F o M
IDPAIS	VARCHAR(4)	Identificador del país del Consumidor
NOMBREPAIS	VARCHAR(50)	Nombre del país del consumidor
IDDEPARTAMENTO	VARCHAR(3)	Identificador del departamento del Consumidor
NOMBREDEPARTAMENTO	VARCHAR(50)	Nombre del departamento
IDMUNICIPIO	VARCHAR(4)	Identificador del municipio
NOMBREMUNICIPIO	VARCHAR(50)	Nombre del municipio
PKIDCONSUMIDOR	INT	Identificador del consumidor de la base de datos FUENTE

DIMEMPLEADO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDEMPLEADO	INT	Identificador del empleado en el DW
NOMBREEMPLEADO	VARCHAR(50)	Nombre del empleado
APELLIDOSEMPLEADO	VARCHAR(50)	Apellido de empleado
CARGOEMPLEADO	VARCHAR(50)	Cargo del empleado
UNIDAEMLPEADO	VARCHAR(50)	Unidad a la que pertenece el empleado
OFICINAEMPLEADO	VARCHAR(100)	Oficina a la que forma parte el empleado
ESTADOEMPLEADO	VARCHAR(50)	Estado del empleado Activo o Desactivo
CORREOEMPLEADO	VARCHAR(100)	Correo electrónico del empleado
PKIDEMPLEADO	INT	Identificador del empleado de la base de datos FUENTE

DIMPROVEEDOR		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDPROVEEDOR	INT	Identificador del proveedor en el DW
NOMBREPROVEEDOR	VARCHAR(200)	Nombre de proveedor
RAZONSOCIAL	VARCHAR(200)	Razón social del proveedor
NIT	VARCHAR(25)	NIT del proveedor
IDSUCURSAL	INT	Identificador de sucursal
NOMBRESUCURSAL	VARCHAR(100)	Nombre de la sucursal
PKPROVEEDOR	INT	Identificador del proveedor de la base de datos FUENTE

DIMCENTROSOLUCION		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDCENTROSOLUCION	INT	Identificador del centro de solución en el DW
NOMBRECENTROSOLUCION	VARCHAR(100)	Nombre de centro de solución
PKOFICINA	VARCHAR(2)	Identificador de la oficina de la base de datos FUENTE

DIMFORMADEPAGO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDFORMADEPAGO	INT	Identificador de la forma de pago en el DW
NUMEROFORMAPAGO	INT	Identificador de la forma de pago de la base FUENTE
DESCRIPCIONFORMAPAGO	VARCHAR(50)	Descripción de la forma de pago

DIMUBICACION		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDUBICACION	INT	Identificador de la ubicación en el DW
PKDEPARTAMENTO	VARCHAR(3)	Identificador del departamento de la base FUENTE
NOMBREDEPARTAMENTO	VARCHAR(50)	Nombre del departamento
PKMUNICIPIO	VARCHAR(4)	Identificador del municipio de la base FUENTE
NOMBREMUNICIPIO	VARCHAR(50)	Nombre del municipio
PKZONA	VARCHAR(3)	Identificador de la zona de la base FUENTE
NOMBREZONA	VARCHAR(50)	Nombre de la zona

DIMETAPA		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDETAPA	INT	Identificador de la Etapa en el DW
PKETAPA	VARCHAR(3)	Identificador de la etapa de la base FUENTE
NOMBREETAPA	VARCHAR(50)	Nombre de la etapa
PKSUBETAPA	VARCHAR(3)	Identificador de la sub etapa de la base FUENTE
NOMBRESUBETAPA	VARCHAR(100)	Nombre de la sub etapa

DIMSECTORCATEGORIA		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDSECTORCATEGORIA	INT	Identificador del sector categoría en el DW
NOMBRECATEGORIA	VARCHAR(100)	Nombre de la categoría
PKSECTOR	VARCHAR(3)	Identificador del sector de la base FUENTE
NOMBRESECTOR	VARCHAR(100)	Nombre del sector
PKCATEGORIA	VARCHAR(4)	Identificador de la categoría de la base FUENTE

DIMTIEMPO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDTIEMPO	INT	Identificador de tiempo del DW
FECHA	DATETIME	Fecha de ingreso de la denuncia
NUMERODIASSEMANA	INT	Numero del día en la semana
NOMBREDIA	VARCHAR(50)	Nombre del día de la semana
NUMERODIAMES	INT	Numero del día en el mes
NUMERODIAANO	INT	Numero del día en el año
NUMEROTRIMESTRE	INT	Numero del trimestre del año
NOMBRETRIMESTRE	VARCHAR(50)	Nombre del trimestre del año
NUMEROANO	INT	Numero del año
NUMEROMES	INT	Numero del mes
NOMBREMES	VARCHAR(50)	Nombre del mes
NUMEROSEMESTRE	INT	Numero del semestre
NOMBRESEMESTRE	VARCHAR(50)	Nombre del semestre

FACTCASO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDCASO	INT	Identificador del caso en el DW
IDCENTROSOLUCION	INT	Identificador del centro de solución en el DW llave foránea
IDSECTORCATEGORIA	INT	Identificador de sector categoría en el DW llave foránea
IDFORMAPAGO	INT	Identificador de forma de pago en el DW llave foránea
IDETAPA	INT	Identificador de la etapa en el DW llave foránea
IDEMPLADO	INT	Identificador del empleado en el DW llave foránea
IDUBICACION	INT	Identificador de la ubicación en el DW llave foránea
IDCONSUMIDOR	INT	Identificador del consumidor en el DW llave foránea

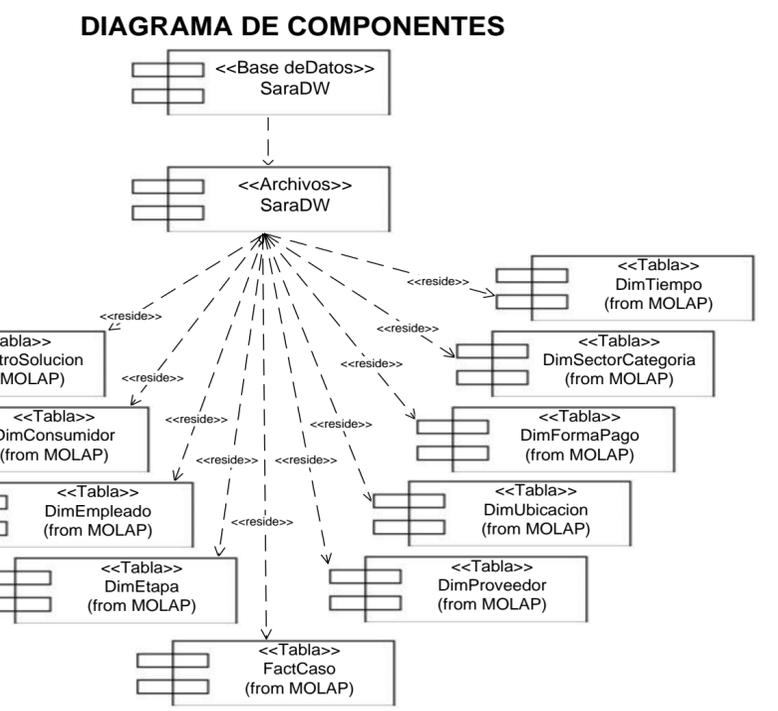
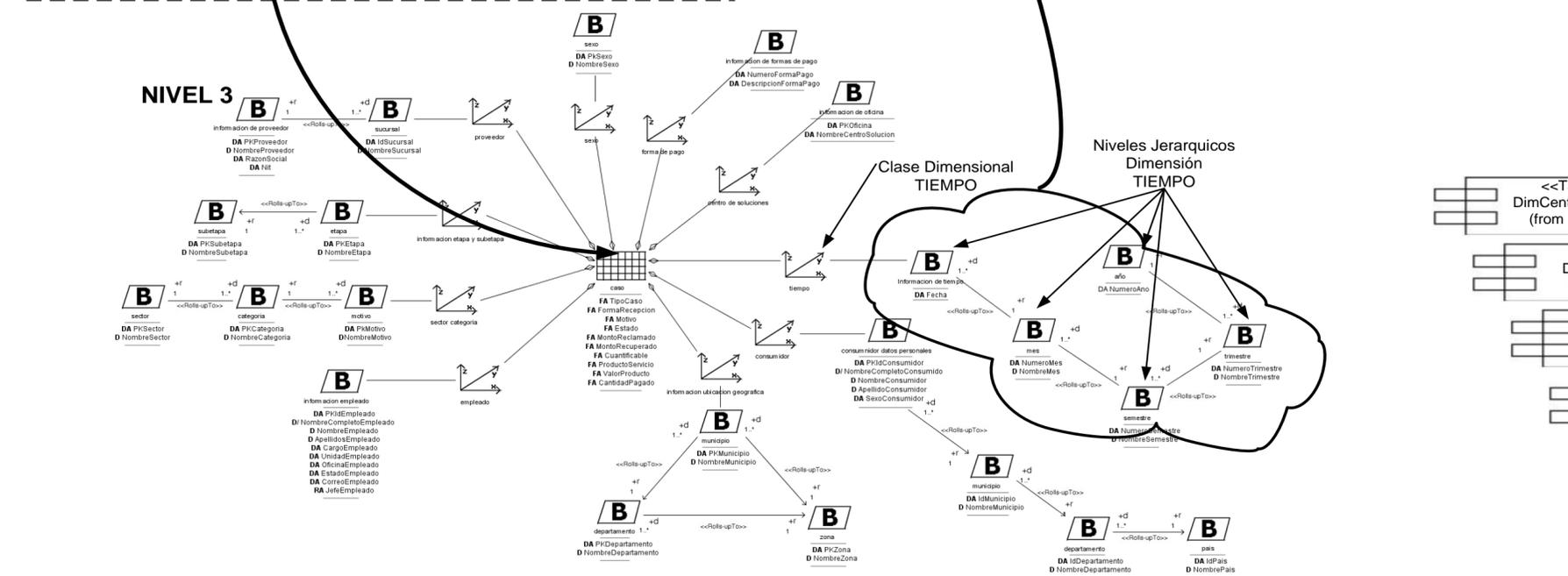
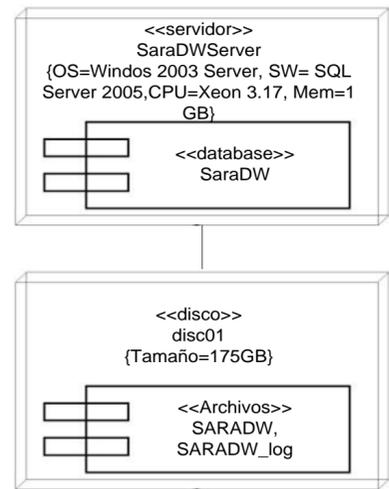
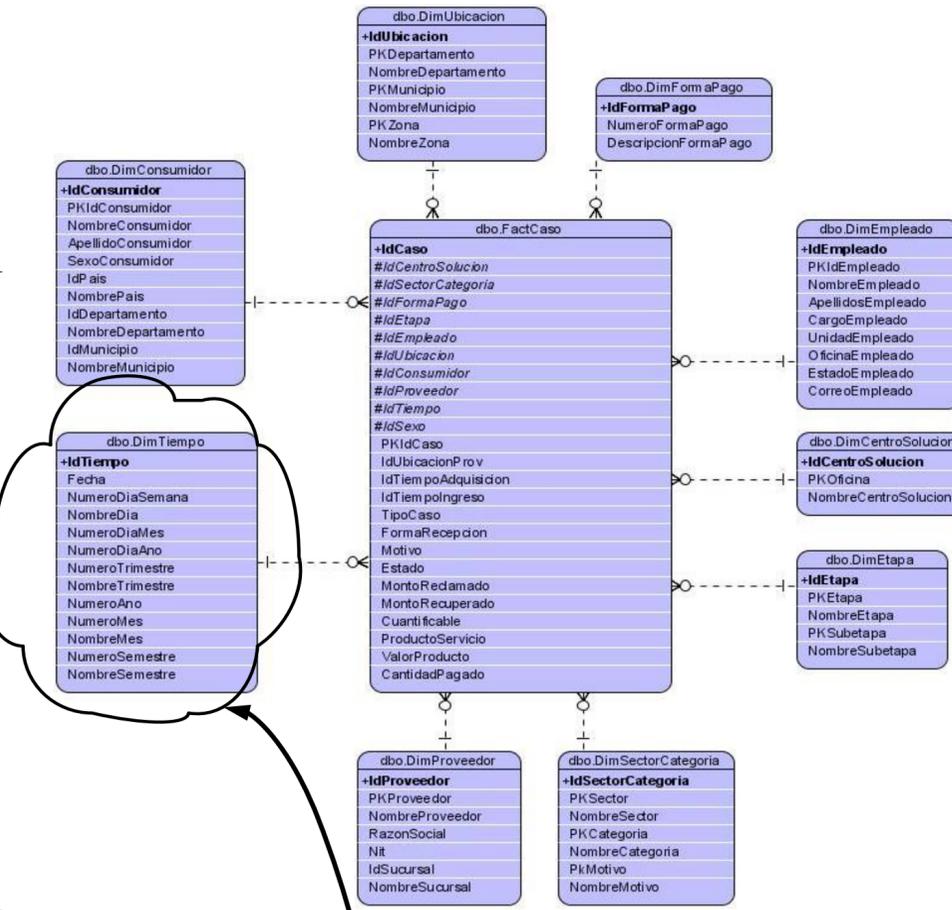
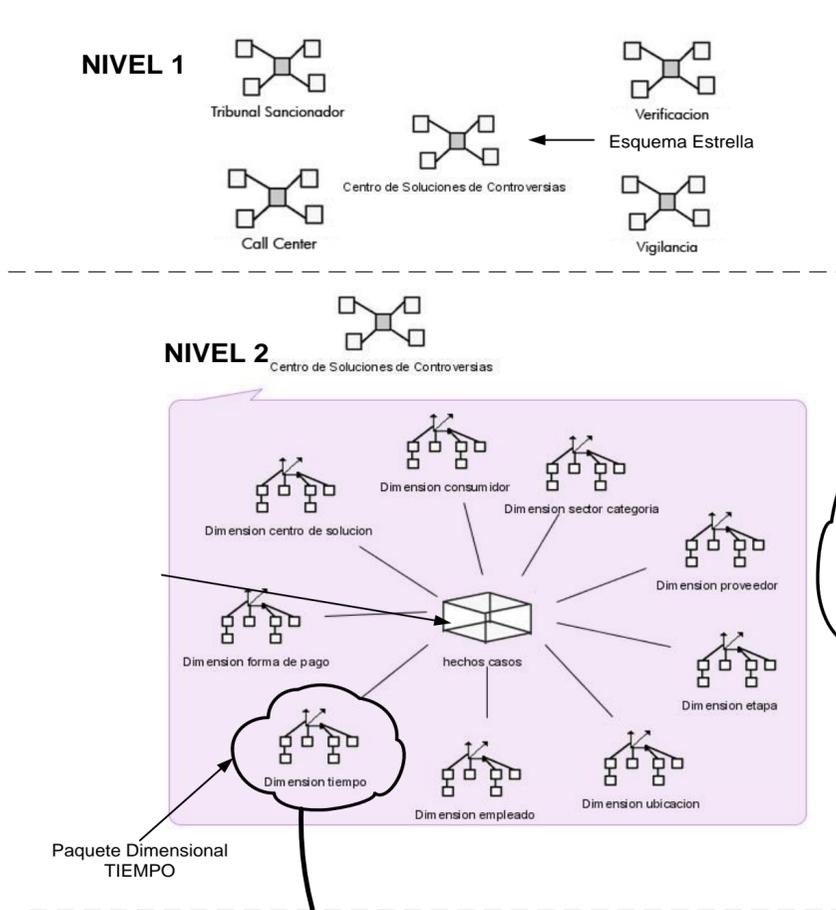
IDPROVEEDOR	INT	Identificador del proveedor en el DW llave foránea
IDTIEMPO	INT	Identificador del tiempo en el DW llave foránea
PKIDCASO	INT	Llave primaria del caso en la base de datos FUENTE
IDUBICACIONPROV	INT	Identificador de la ubicación del proveedor
IDTIEMPOADQUISICION	INT	Identificador del tiempo de adquisición
IDTIEMPOINGRESO	INT	Identificador del tiempo de ingreso del caso
TIPOCASO	VARCHAR(50)	Tipo del caso de la denuncia
FORMARECEPCION	VARCHAR(50)	Forma de recepción del caso en centro o por teléfono etc.
MOTIVO	VARCHAR(50)	Motivo de la denuncia
ESTADO	VARCHAR(50)	Estado del caso
MONTORECLAMADO	VARCHAR(50)	Monto reclamado por el cliente
MONTORECUPERADO	NUMERIC(10,2)	Monto recuperado del caso
CUANTIFICABLE	CHAR(1)	Si el reclamo es cuantificable S o N
PRODUCTOSERVICIO	CHAR(1)	Tipo de denuncia por producto o servicio
VALORPRODUCTO	NUMERIC(10,2)	Valor del producto
CANTIDADPAGADO	NUMERIC(10,2)	Cantidad pagada por el consumidor

3.5.4. Diagrama UML del DATAWAREHOUSE – CENTRO DE SOLUCION DE CONTROVERSIAS

DIAGRAMA UML DEL DATA WAREHOUSE- CENTRO DE SOLUCION DE CONTROVERSIAS

ESQUEMA CONCEPTUAL DEL DATA WAREHOUSE

ESQUEMA LÓGICO DEL DATA WAREHOUSE



3.5.5. Diagrama del proceso ETL

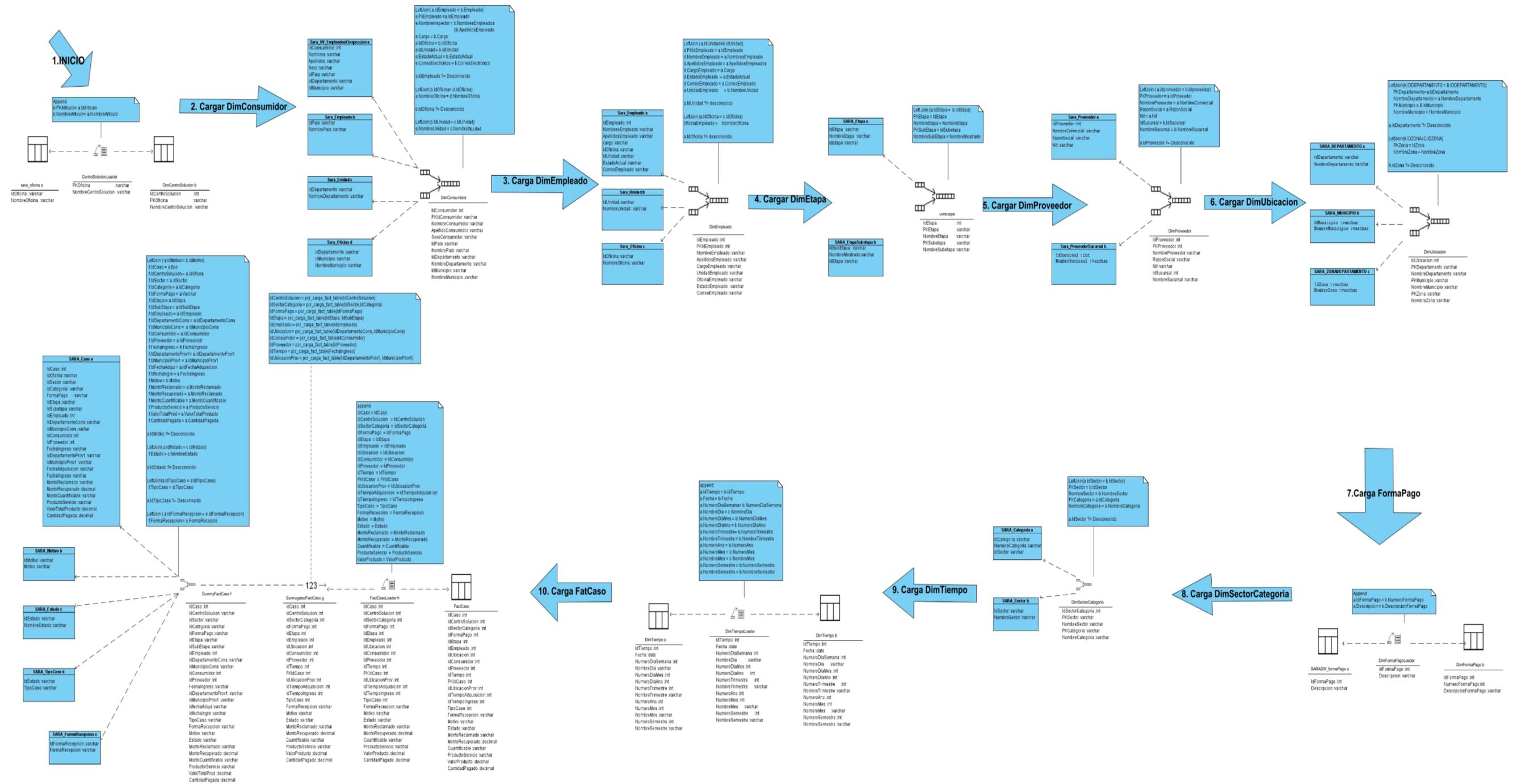
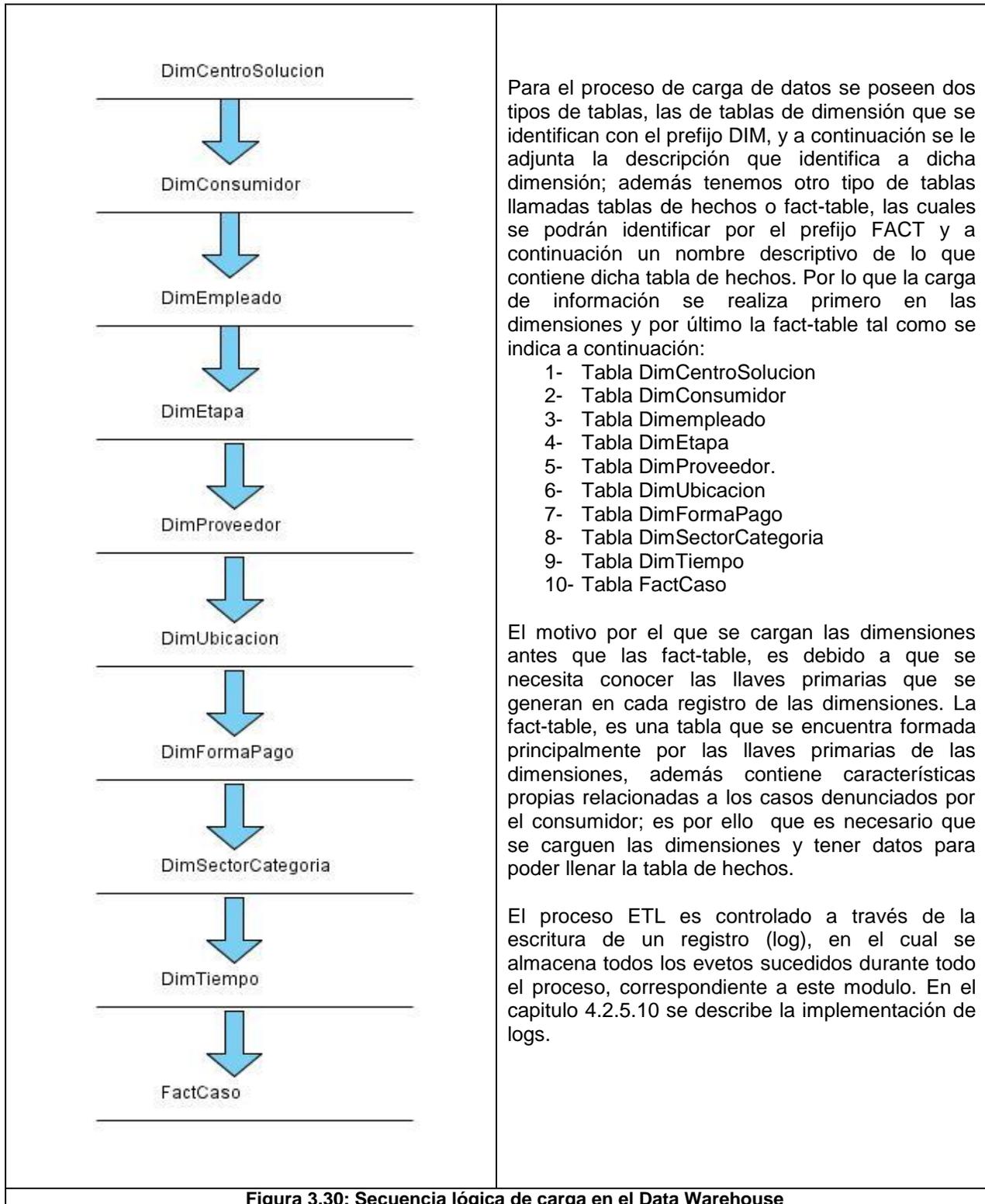


Figura 3.29: Diseño del proceso ETL para el Data Warehouse

3.5.5.1. Secuencia de carga del proceso ETL



3.5.5.2. Diagrama de carga para la dimensión Centro Solución

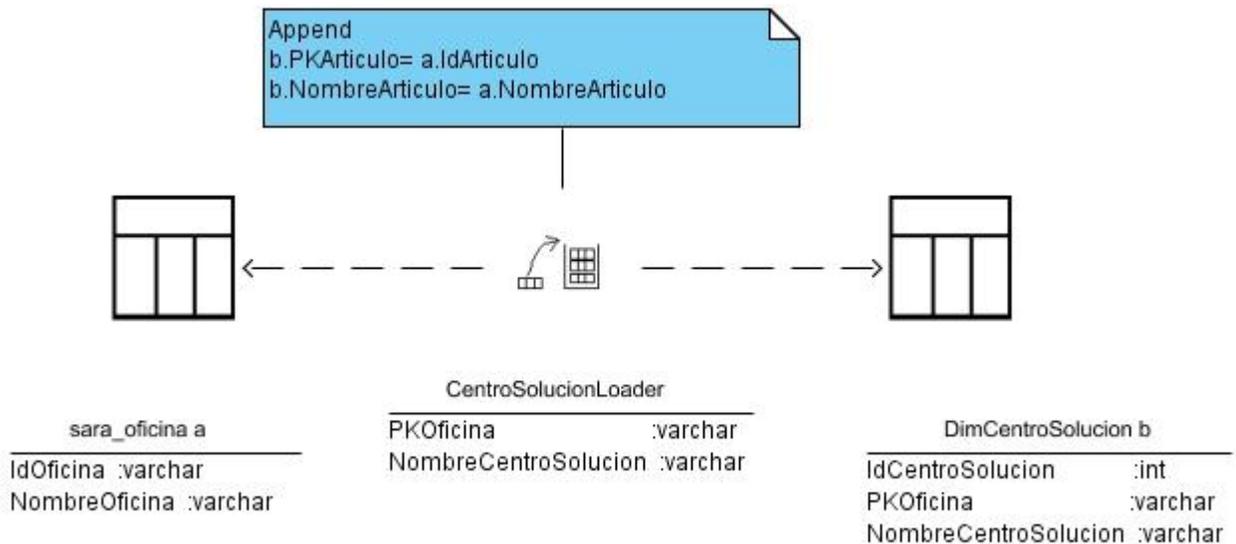


Figura 3.31: Representación de carga de datos en la tabla DimCentroSolucion utilizando mecanismo Loader

3.5.5.3. Diagrama de carga para la dimensión Etapa

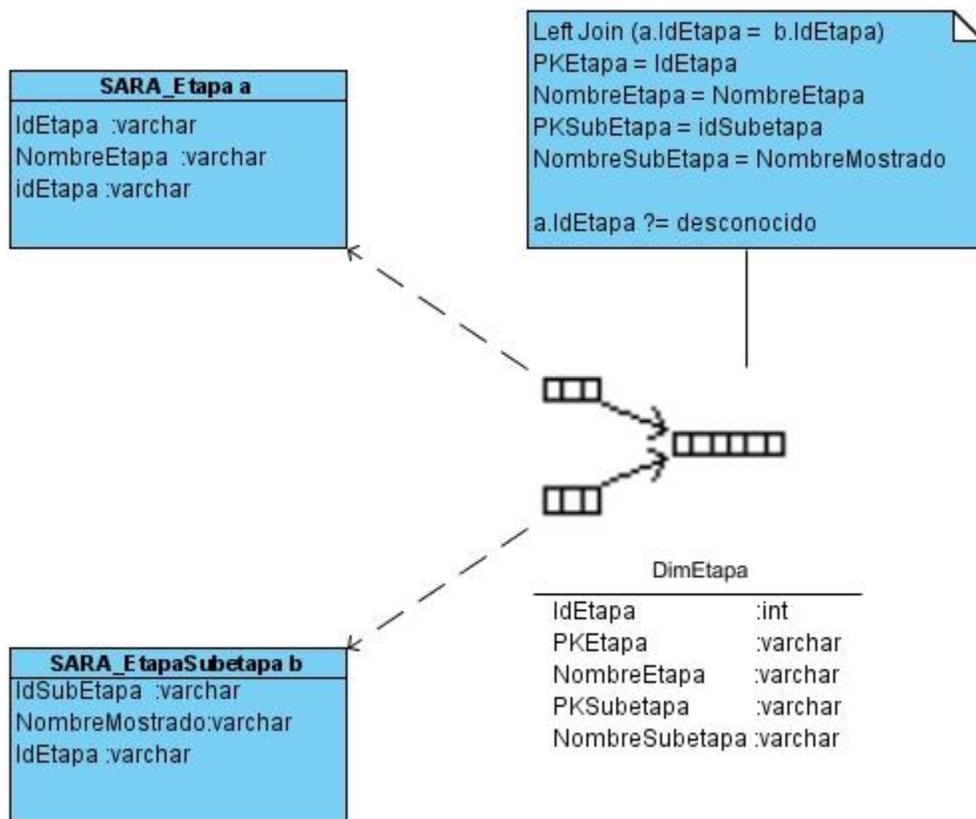


Figura 3.32: Representación de carga de datos en la tabla DimEtapa utilizando mecanismo Join

3.5.5.4. Diagrama de carga para la dimensión Proveedor

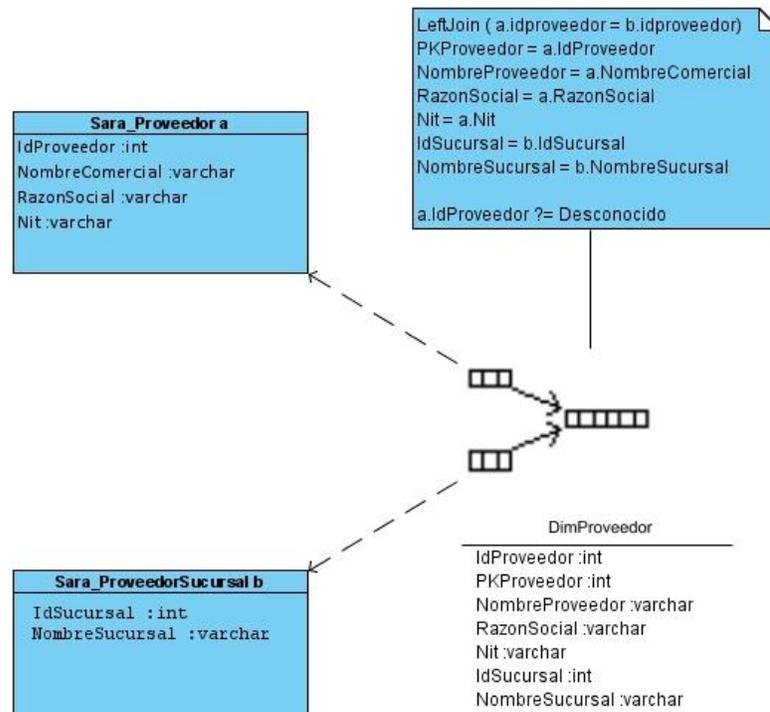


Figura 3.33: Representación de carga de datos en la tabla DimProveedor utilizando mecanismo Join

3.5.5.5. Diagrama de carga para la dimensión Ubicación

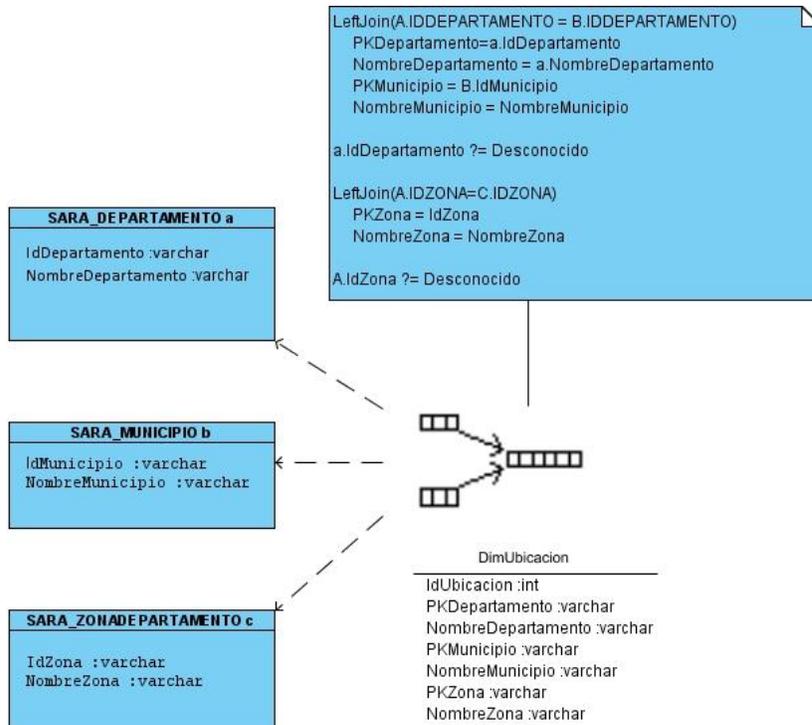


Figura 3.34: Representación de carga de datos en la tabla DimUbicacion utilizando mecanismo Join

3.5.5.6. Diagrama de carga para la dimensión consumidor

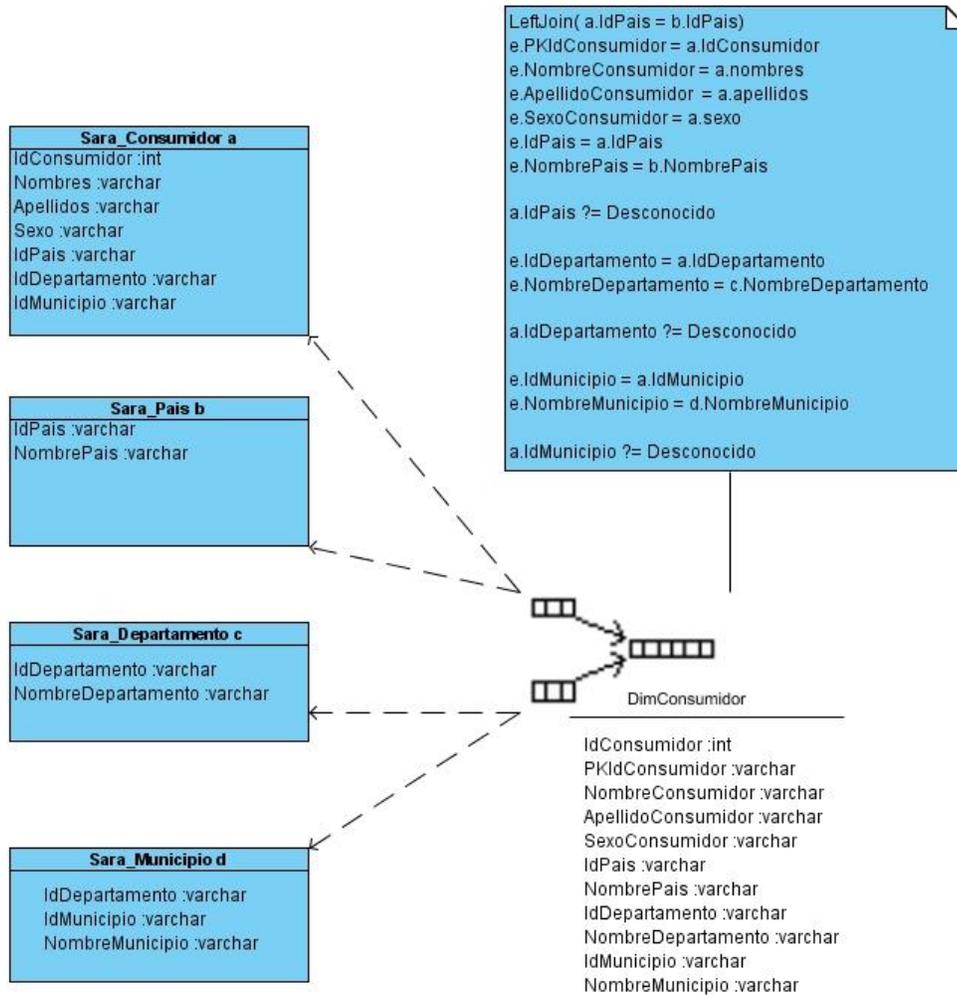


Figura 3.35: Representación de carga de datos en la tabla DimConsumidor utilizando mecanismo Join

3.5.5.7. Diagrama de carga para la dimensión forma de pago

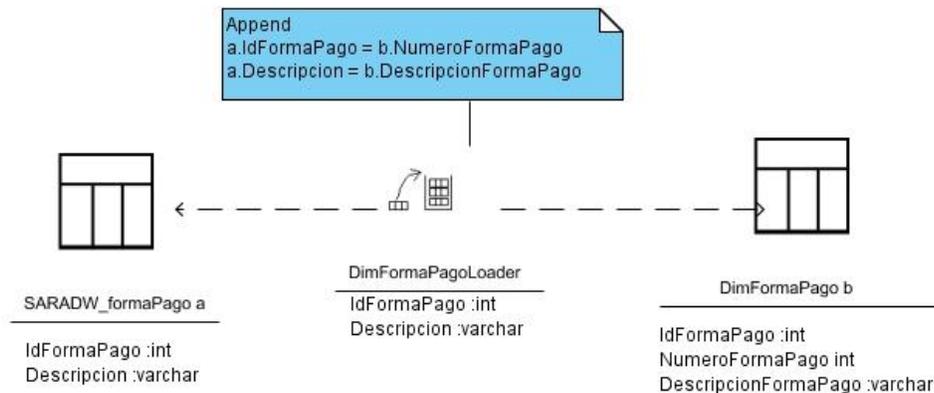


Figura 3.36: Representación de carga de datos en la tabla DimFormaPago utilizando mecanismo Loader

3.5.5.8. Diagrama de carga para la dimensión Empleado

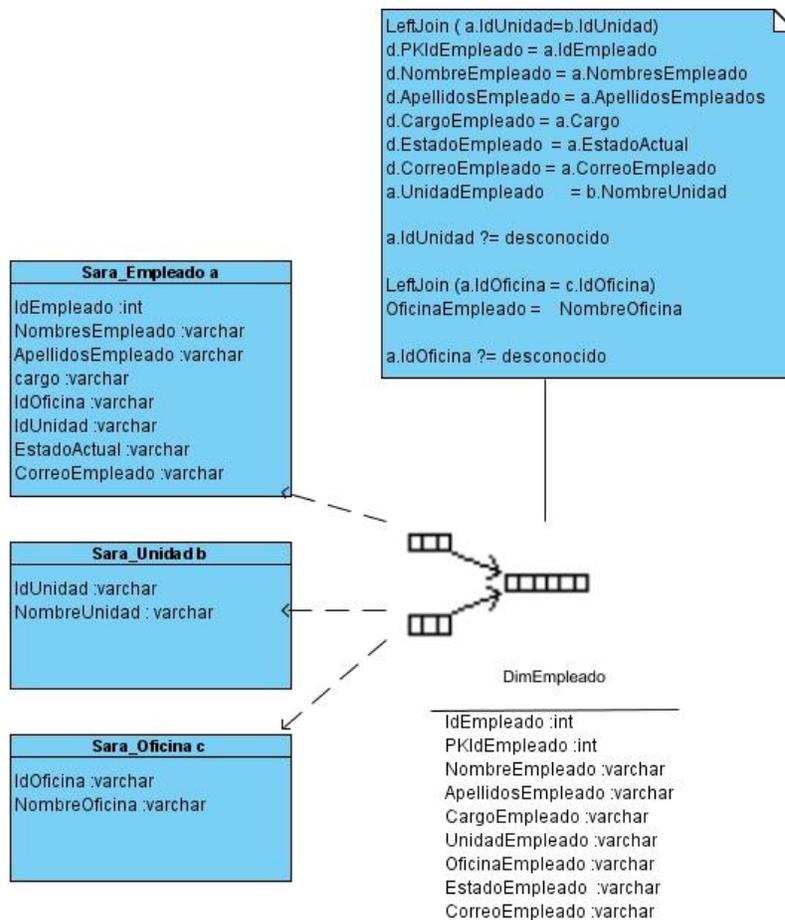


Figura 3.37: Representación de carga de datos en la tabla DimEmpleado utilizando mecanismo Join

3.5.5.9. Diagrama de carga para la dimensión Categoría

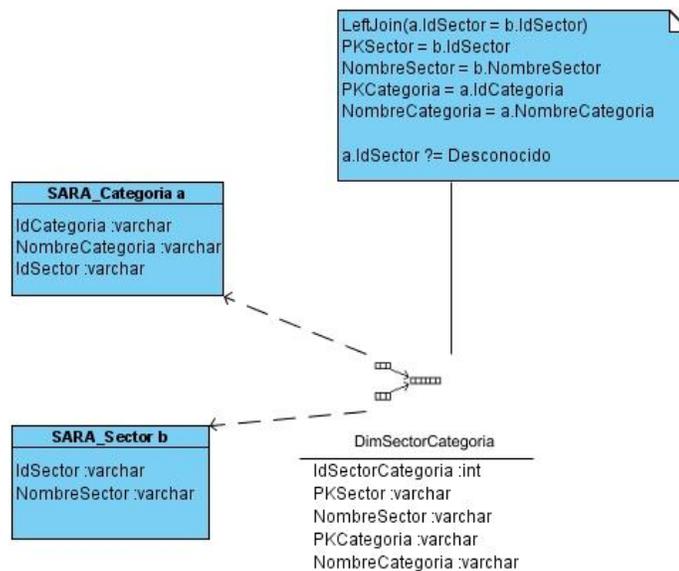


Figura 3.38: Representación de carga de datos en la tabla DimSectorCategoria utilizando mecanismo Join

3.5.5.10. Diagrama de carga para la dimensión tiempo

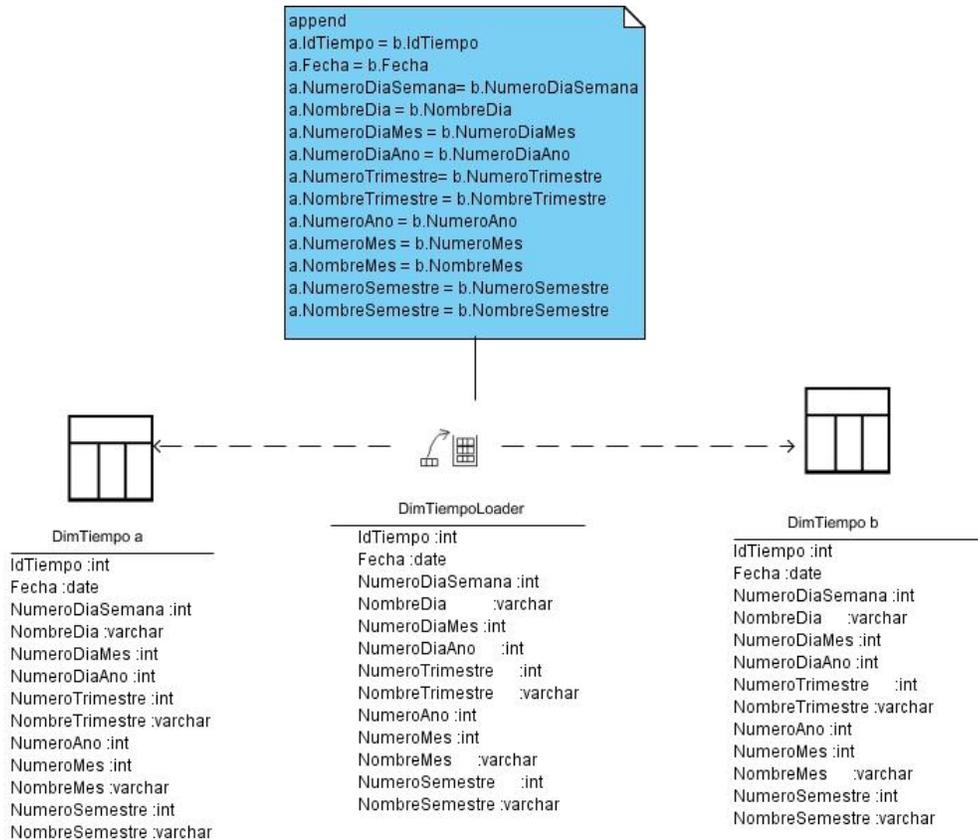


Figura 3.39: Representación de carga de datos en la tabla DimTiempo utilizando mecanismo Loader

3.5.5.11. Diagrama de carga para la factable FatCaso

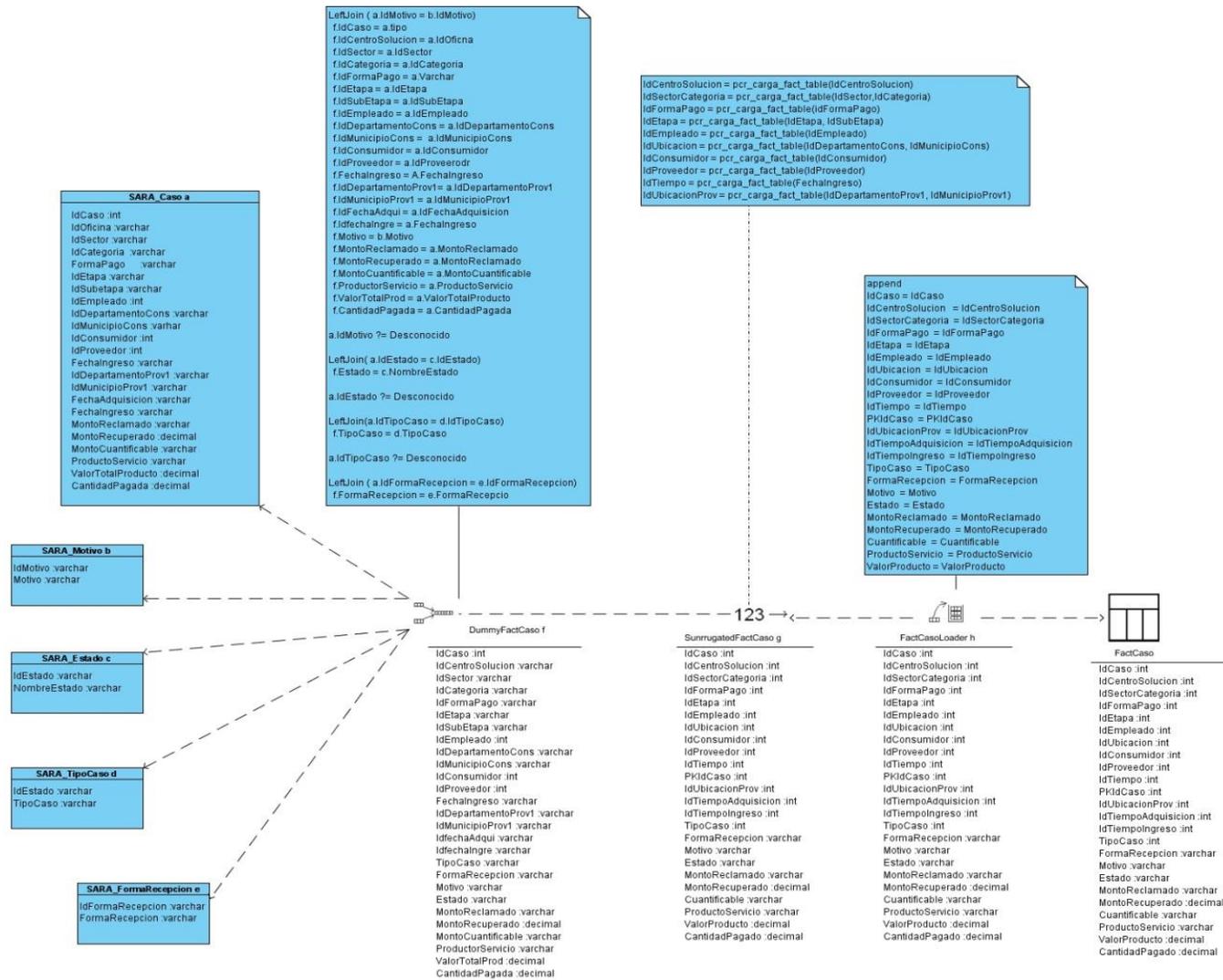


Figura 3.40: Representación de carga de datos en la tabla FactCaso utilizando mecanismo Join, Surrogate y Loader

### 3.6. ITERACIÓN 2: VERIFICACIÓN Y VIGILANCIA

Se encarga de realizar inspecciones a los proveedores.

#### 3.6.1. Requerimientos Funcionales

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Página 1 de y	
Autor : Grupo 14		Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo
<b>01</b>	<b>Analizar proceso de inspecciones</b>	Alta	FoO
<b>01.01</b>	Permitir al usuario analizar el proceso de denuncias determinando la cantidad de inspecciones y hallazgos determinados en las inspecciones		
<b>01.01.01</b>	Analizar la cantidad de inspecciones permite determinar patrones los cuales generan hallazgos a fin de poder tener mayor control.		
<b>01.01.02</b>	Los hallazgos dependen plenamente de la cantidad de artículos de ley que fueron infringidos por los proveedores según determinación del inspector encargado		
<b>01.01.03</b>	Una inspección puede tener más de un hallazgo por lo que la determinación de los hallazgos dependen plenamente de los artículos infringidos en cada inspección.		
<b>02</b>	<b>Analizar las inspecciones según fecha</b>	Alta	FoO
<b>02.01</b>	El análisis de las inspecciones según la fecha permite determinar la cantidad de trabajo que se realiza según la escala de tiempo o planes de trabajos fijados.		
<b>02.02</b>	El combinar las fecha con otras variables como los rubros ayuda a identificar las áreas en las que se ha trabajado a lo largo del tiempo		
<b>03</b>	<b>Analizar inspecciones por rubro</b>	Alta	FoO
<b>03.01</b>	Identificar los rubros en los que se encuentran la mayor cantidad de hallazgos facilita la creación de planes de trabajo los cuales se enfocan a dichos rubros a fin de mejorar el servicio al cliente, y evitar nuevas incidencias por parte de los proveedores.		
<b>03.02</b>	La combinación de las inspecciones por rubro con la fecha facilita la identificación de las temporadas en la que más problemas se han tenido en un rubro específico.		
<b>04</b>	<b>Detalle de artículos de ley relacionados por inspecciones</b>	Alta	FoO
<b>04.01</b>	El análisis de artículos relacionados es de vital importancia debido a que estos artículos son reglas las cuales tienen una base jurídica. Una inspección puede tener varios artículos relacionados		

<b>04.02</b>	Los artículos infringidos por cada inspección pueden relacionarse con otras variables, tales como tiempo y rubro, las cuales dependerán plenamente de la inspección. Por lo que es importante conocer la cantidad de inspecciones por artículo y los hallazgos por artículo.		
<b>04.03</b>	Los hallazgos dependerán si en verdad se infringe uno o varios artículos de ley en cada inspección		
<b>05</b>	<b>Análisis de inspecciones por establecimiento</b>		
<b>05.01</b>	Una inspección se realiza a cada establecimiento. Por lo que la información del establecimiento inspeccionado es muy importante para determinar la realización de análisis. Es importante aclarar que un establecimiento se refiere a una sucursal de un proveedor X.	Media	<b>FoO</b>
<b>05.01.1</b>	Mostrar la cantidad de inspecciones realizadas según la variable que se encuentre combinada con el establecimiento		
<b>05.02</b>	Mostrar la cantidad de hallazgos encontrados por establecimiento.		
<b>06</b>	<b>inspecciones realizadas por delegado</b>		
<b>06.01</b>	Mostrar la cantidad de inspecciones realizadas por cada delegado	Baja	<b>FoD</b>
<b>06.02</b>	Mostrara la cantidad de hallazgos determinados por cada delegado		
<b>07</b>	<b>Reporte adhoc de la información analizada</b>		
<b>07.01</b>	Mostrar la información detallada de para poder constatar la información que es mostrada en los análisis.	Baja	<b>Fo</b>
<b>07.02</b>	Mostrar la información dando al usuario la libertad de elegir la información que desee que incluya el usuario		
<b>08</b>	<b>Inspecciones Realizadas por Centro de solución</b>		
<b>08.01</b>	Mostrar la cantidad e inspecciones coordinadas por cada centro de solución para determinar la cantidad de trabajo desarrollado para un tiempo especificado.	Alta	<b>FoO</b>
<b>08.02</b>	Mostrar la cantidad e hallazgos encontrados en las inspecciones coordinadas por cada centro de solución para determinar el trabajo realizado en un periodo de tiempo		
<b>09</b>	<b>Análisis de inspecciones por proveedor</b>		
<b>09.01</b>	Presentar la cantidad de inspecciones realizadas a los proveedores, ya sea para un rango de tiempo o el total de inspecciones que se han realizado.	Alta	<b>FoD</b>
<b>09.02</b>	Presentar la cantidad de hallazgos encontrados por proveedor en un periodo de tiempo o en su totalidad.		

### 3.6.2. Especificación de Requerimientos

#### 3.6.2.1. Crear Reporte Analítico



Figura 3.41: Diagrama de caso de uso para reportes analíticos

<p><b>Caso de uso 1 Crear Reporte Analítico</b></p> <p><b>Personal Involucrado e interés:</b></p> <p>Usuario: Técnico, Técnico Web, Dirección de estudios, Dirección Verificación y Vigilancia, Gerencia informática.</p> <p>Usuario: Elaborar un análisis detallado del proceso realizado en la dirección de verificación y vigilancia para su posterior publicación o almacenamiento.</p> <p><b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo.</p>
--

<p><b>Postcondiciones:</b> Reporte analítico publicado , almacenado o impreso</p>
<p><b>Referencias Cruzadas</b>  <b>Requerimientos:</b> 01,02,03,04,05,06</p>
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona medidas.</li> <li>6. El Sistema muestra datos de verificación y vigilancia, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A,B,C,D,E,F.</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis</li> <li>10. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>11. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Realizar análisis utilizando MDX (Usuarios Expertos)</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción MDX.</li> <li>2. El sistema muestra opción de generación de consulta MDX.</li> <li>3. Usuario construye la consulta, utilizando consulta valida.</li> <li>4. El sistema realiza la construcción del espacio de trabajo en base a la consulta generada.</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>B. Configuración de tabla OLAP</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Configurar tabla OLAP.</li> <li>2. El sistema muestra ventana de configuración OLAP</li> <li>3. El usuario selecciona las opciones de configuración</li> <li>4. El sistema aplica configuraciones</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>C. Seleccionar opciones de las vistas de los datos</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Mostrar Padres.</li> <li>2. El sistema muestra las referencias fuentes de los datos mostrados.</li> <li>3. Sino el usuario selecciona la opción ocultar repeticiones.</li> <li>4. El sistema muestra solo los datos únicos (unique) sin repeticiones.</li> <li>5. Sino el usuario el usuario selecciona mostrar propiedades.</li> <li>6. El sistema muestra propiedades de los datos.</li> <li>7. Sino el usuario selecciona suprimir columnas vacías.</li> <li>8. El sistema elimina de la consulta aquellos datos que contengan valores nulos.</li> <li>9. Sino el usuario selecciona la opción intercambiar ejes.</li> <li>10. El sistema genera un análisis roll-up.</li> <li>11. El usuario analiza los datos.</li> </ol>

**D. Generación y configuración de gráfico**

1. El usuario selecciona la opción configuración de gráfico.
2. El sistema muestra propiedades de gráfico.
3. El usuario configura opciones de gráfico, colores, leyendas, orientación y tipo de gráfico.
4. El sistema adecua el gráfico a los valores que se muestran en el espacio de trabajo.
5. El usuario selecciona opción de generar gráfico.
6. El sistema muestra gráfico.
7. El usuario analiza los datos.

**E. Seleccionar el Destino del análisis**

1. El usuario selecciona opción de configuración de impresión.
2. El sistema muestra panel de configuración de impresión.
3. El usuario selecciona opciones de configuración tales como tema de reporte, orientación, tamaño del papel.
4. El sistema configura opciones de reporte analítico.
5. El usuario publica el reporte en formato PDF, EXCEL, CSV o lo almacena en el sistema.
6. El sistema procesa el reporte y lo publica dependiendo el método seleccionado.

**F. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.6.2.2. Crear Reporte AD HOC

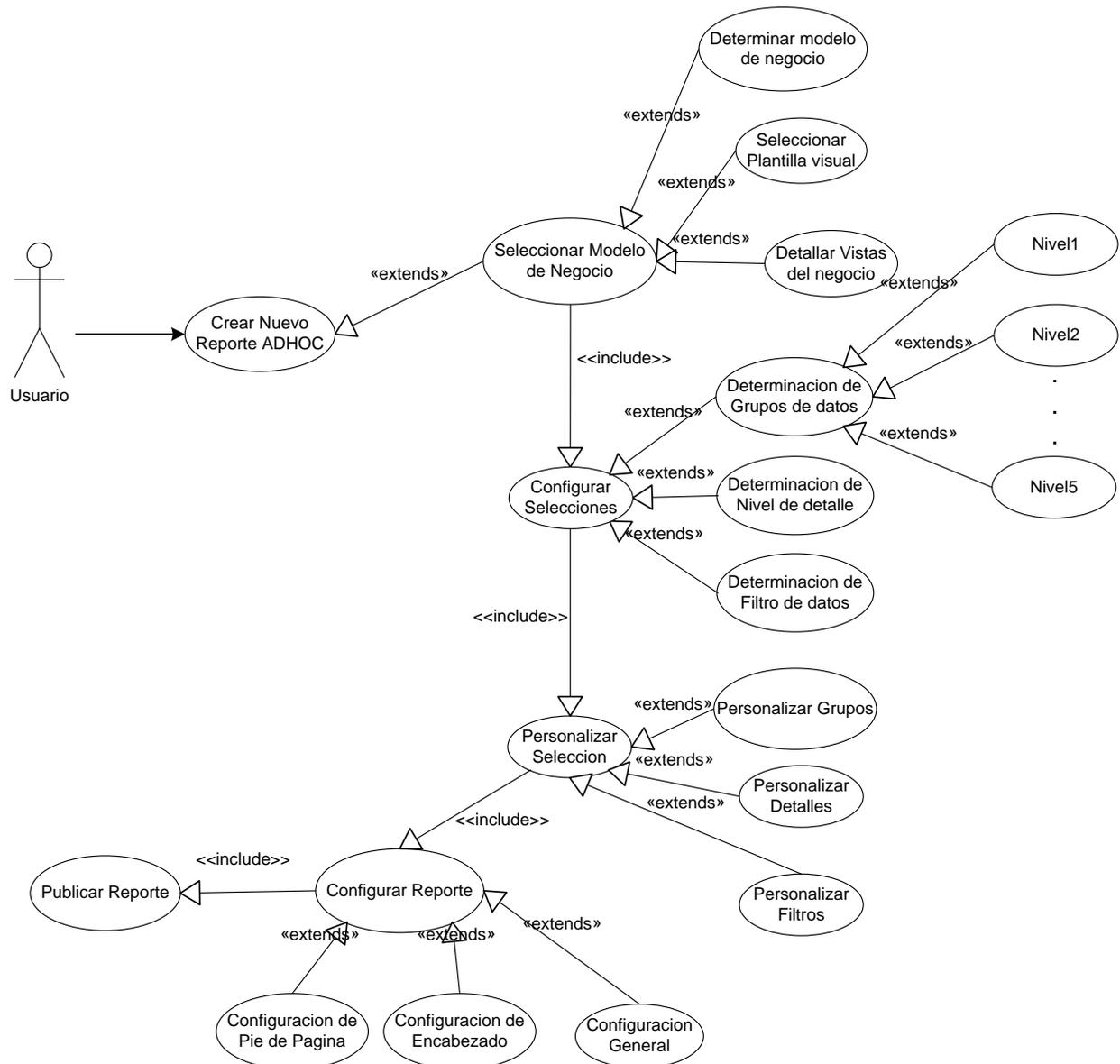


Figura 3.42: Diagrama de caso de uso para reportes AD HOC

<b>Caso de uso 2 Crear Reporte AD HOC</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Técnico Web, Dirección de Estudios, Gerencia informática.  Usuario: Elaborar de reportes temporales para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema poseer acceso a la vista de negocios de Verificación y vigilancia
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> Requerimientos: 07

**Flujo Básico**

1. El Usuario selecciona "Nuevo Reporte".
2. El sistema muestra los modelos de negocio disponibles.
3. Usuario selecciona el modelo de negocio con el que desea trabajar.
4. Sistema muestra las vistas de negocio disponibles para el modelo seleccionado.
5. El usuario procede a la siguiente fase de selección de reporte
6. El sistema muestra variables con las que se puede agrupar aplicar filtros y seleccionar niveles de detalle de los grupos definidos
7. El usuario selecciona los grupos, filtros y niveles de detalle con las que desea realizar el reporte
8. El usuario procede a la siguiente fase de personalización de selecciones
9. El sistema muestra los valores de los niveles de detalle, grupos y filtros con los que se desea generar el reporte.
10. El sistema almacena la configuración de los datos.
11. El usuario procede a la siguiente fase de configuración de reporte
12. El sistema muestra panel de configuración de reporte.
13. El usuario selecciona las opciones generales del reporte, tipo de papel orientación encabezado y pie de página.
14. El sistema almacena la configuración
15. El usuario almacena el reporte ad hoc generado para futuras consultas
16. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.

**Flujo Alternativo****A. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.6.2.3. Análisis de Inspecciones para Verificación y Vigilancia

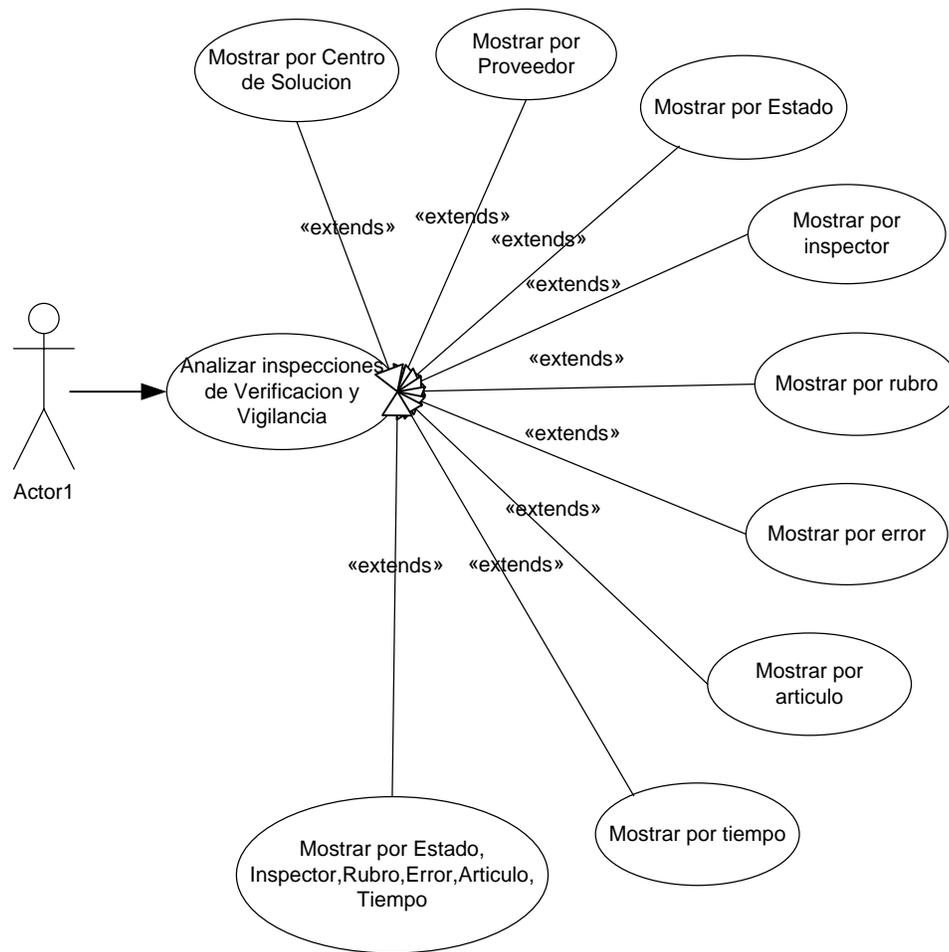


Figura 3.43: Diagrama de caso de uso para análisis de los casos Verificación y Vigilancia

<b>Caso de uso 3 Análisis de Inspecciones para Verificación y Vigilancia</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Técnico Web, Dirección de Estudios, Gerencia informática.  Usuario: Elaborar un análisis detallado del centro de solución de controversias para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo de Verificación y Vigilancia.
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 01,02,03,04,05,06
<b>Flujo Básico</b> 1. El Usuario selecciona “Nueva Vista de Análisis”. 2. El Sistema muestra esquemas y cubos disponibles. 3. El Usuario selecciona esquema y cubo de verificación y vigilancia para análisis. 4. El Sistema muestra dimensiones y medidas de análisis. 5. El Usuario selecciona las medidas. 6. El Sistema muestra datos de Verificación y Vigilancia, con diferentes opciones de

análisis.

7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A
8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas
9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis con medidas tales como: cantidad de inspecciones y hallazgos; combinando dimensiones como: tiempo, Artículos, Errores, Inspectores Centro de solución, proveedor y rubros.
10. El sistema ordena la información por filas, columnas y filtros aplicados
11. El usuario almacena el reporte analítico generado para futuras consultas
12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso

#### **Flujo Alternativo**

##### **A. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

### 3.6.3. Modelado Iteración 2 Modulo de Verificación y Vigilancia

#### 3.6.3.1. Esquema Conceptual del Data Warehouse

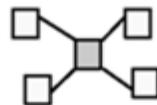
En Tabla 1.3 se encuentra la definición de la nomenclatura UML utilizada para la definición de los estereotipos del modelado conceptual.

##### Nivel 1. Definición del Modelo

Modulo de Verificación y Vigilancia:

Especialización del Negocio para análisis de los casos.

**Hecho principal:** El Estudio de las inspecciones



Verificación y Vigilancia

Figura 3.44: Definición modelo de “Verificación y Vigilancia”

##### Nivel 2 Definición del esquema estrella

Paquete de Hecho: Inspecciones

Paquetes Dimensionales

Rubros: son clasificaciones generales basadas en el rubro del negocio.

Artículos: son reglas fijadas con bases legales.

Estados: es la situación actual que se encuentra la inspección.

Tiempo: Control de tiempo en años, días, cuartiles, meses, semanas.

Inspectores: se refiere a los encargados de realizar las inspecciones.

Errores: Son los fallas que se encuentran al realizar las inspecciones.

Centro Solución: es la oficina en origen a la que hace referencia la Inspección.

Proveedor: es la entidad sometida a inspección.

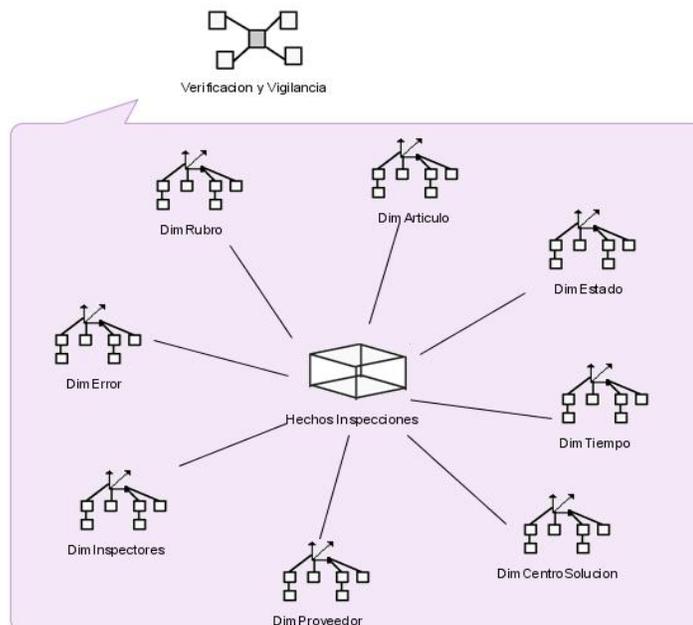


Figura 3.45: Definición del esquema estrella para modulo de “Verificación y Vigilancia”

**Nivel 3: Definición de Dimensiones/Hecho.**

Especificación de atributos.

Dimensiones candidatas a roll up y drill-down: tiempo.

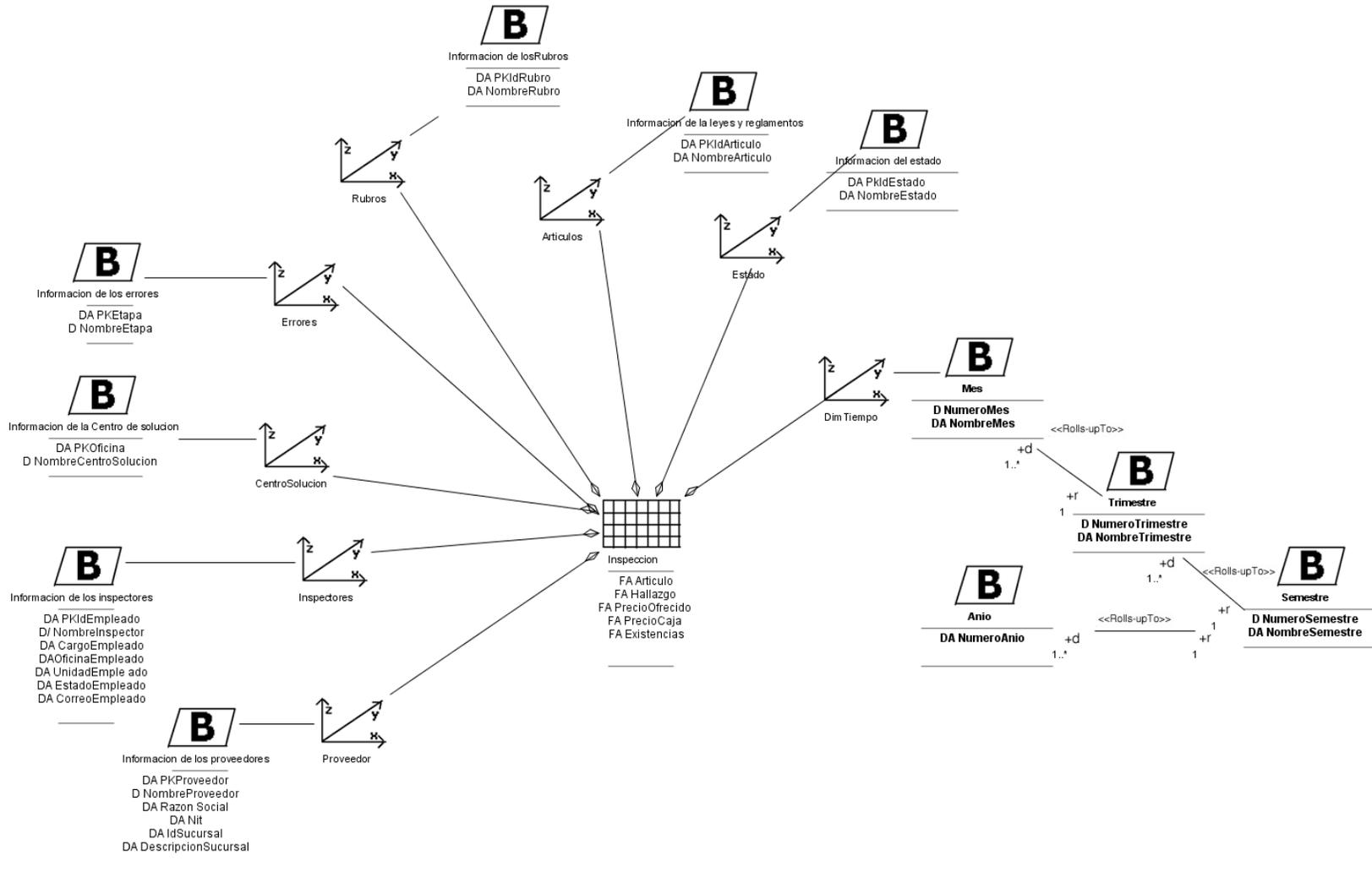


Figura 3.46: Definición de las dimensiones y hechos “Verificación y Vigilancia”

3.6.3.2. Esquema lógico del Data Warehouse

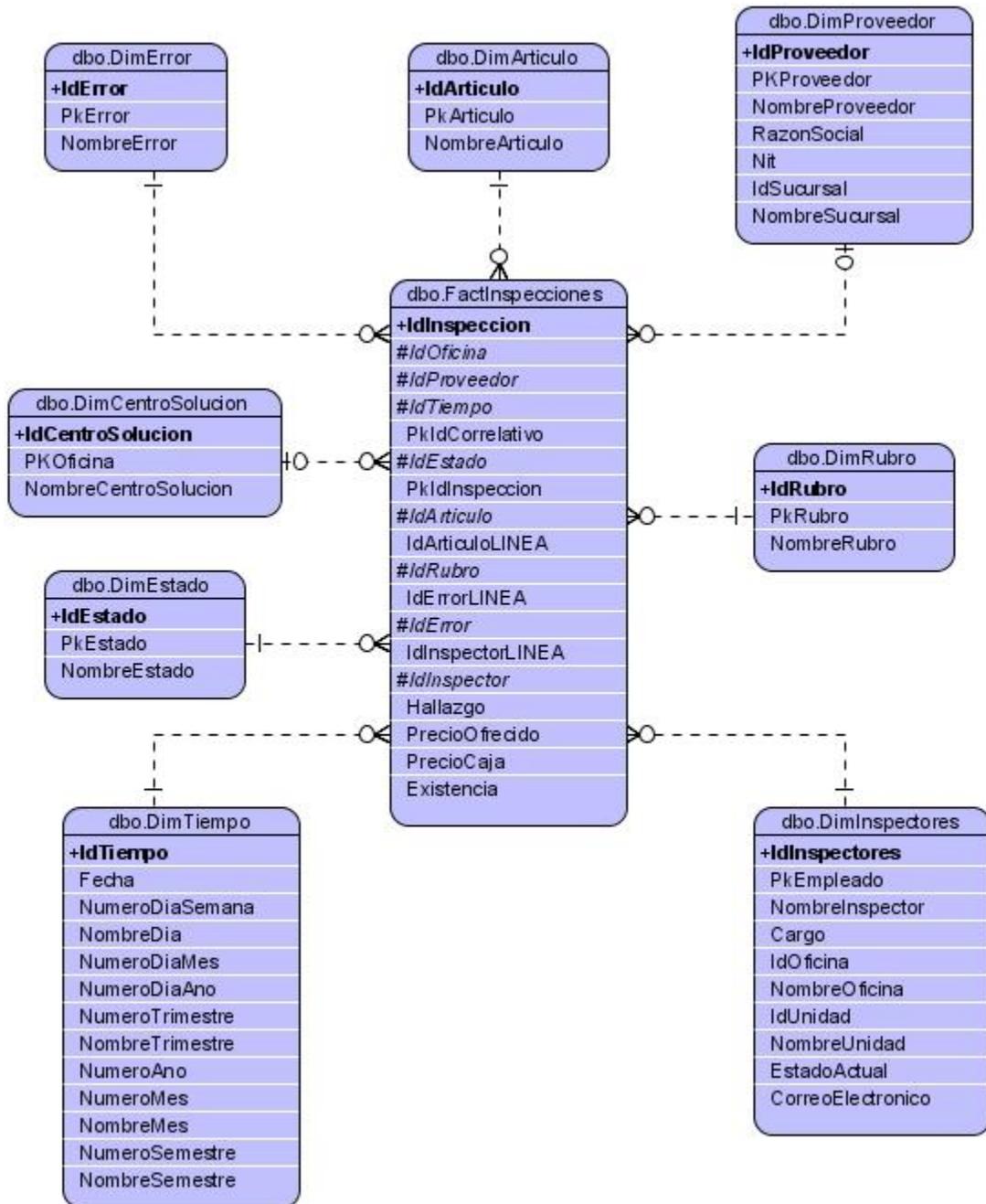


Figura 3.47: Estructura lógica del esquema estrella para el módulo de Verificación y Vigilancia

**3.6.3.3. Diccionario de datos de las dimensiones y tablas de hecho**

DIMERROR		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDERROR	INT	Identificador de la dimensión error del DW.
PKERROR	INT	Identificador de los errores de la base de origen
NOMBREERROR	VARCHAR(50)	Descripción del error

DIMRUBRO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDRUBRO	INT	Identificador de la dimensión rubro del DW.
PKRUBRO	INT	Identificador de los rubros de la base de origen
NOMBRERUBRO	VARCHAR(150)	Descripción del rubro

DIMARTICULO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDARTICULO	INT	Identificador de la dimensión articulo del DW.
PKARTICULO	INT	Identificador de los artículos de la base de origen
NOMBREARTICULO	VARCHAR(50)	Descripción del Articulo

DIMESTADO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDESTADO	INT	Identificador de la dimensión estado del DW.
PKESTADO	INT	Identificador de los estados de la base de origen
NOMBREESTADO	VARCHAR(50)	Descripción del Estado

DIMINSPECTORES		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDINSPECTORES	INT	Identificador de la dimensión inspectores del DW.
PKEMPLEADO	INT	Identificador de los inspectores de la base de origen.
NOMBREINSPECTOR	VARCHAR(200)	Nombre y apellido del inspector
CARGO	VARCHAR(50)	Puesto que desempeña dentro de la organización
IDOFICINA	VARCHAR(2)	Identificación de la oficina a la que pertenece el Inspector
NOMBREOFICINA	VARCHAR(100)	Nombre de la oficina a la que pertenece el inspector.
IDUNIDAD	VARCHAR(3)	Identificación de la unidad a la que pertenece el Inspector
NOMBREUNIDAD	VARCHAR(50)	Nombre de la unidad a la que pertenece el inspector.
ESTADOACTUAL	VARCHAR(50)	Descripción del estado actual del inspector
CORREOELECTRONICO	VARCHAR(100)	Dirección de correo electrónico del inspector

DIMTIEMPO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDTIEMPO	INT	Identificador de tiempo del DW
FECHA	DATETIME	Fecha de ingreso de la denuncia

NUMERODIASEMANA	INT	Numero del día en la semana
NOMBREDIA	VARCHAR(50)	Nombre del día de la semana
NUMERODIAMES	INT	Numero del día en el mes
NUMERODIAANO	INT	Numero del día en el año
NUMEROTRIMESTRE	INT	Numero del trimestre del año
NOMBRETRIMESTRE	VARCHAR(50)	Nombre del trimestre del año
NUMEROANO	INT	Numero del año
NUMEROMES	INT	Numero del mes
NOMBREMES	VARCHAR(50)	Nombre del mes
NUMEROSEMESTRE	INT	Numero del semestre
NOMBRESEMESTRE	VARCHAR(50)	Nombre del semestre

DIMCENTROSOLUCION		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDCENTROSOLUCION	INT	Identificador del centro de solución en el DW
NOMBRECENTROSOLUCION	VARCHAR(100)	Nombre de centro de solución
PKOFICINA	VARCHAR(2)	Identificador de la oficina de la base de datos FUENTE

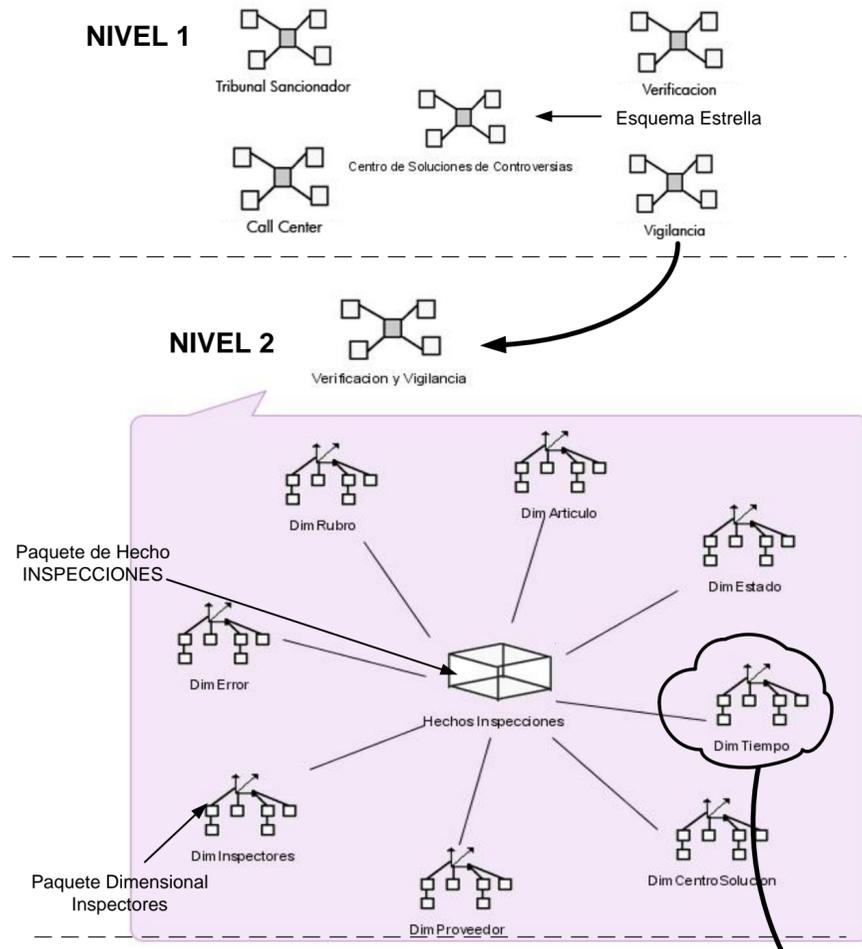
DIMPROVEEDOR		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDPROVEEDOR	INT	Identificador del proveedor en el DW
NOMBREPROVEEDOR	VARCHAR(200)	Nombre de proveedor
RAZONSOCIAL	VARCHAR(200)	Razón social del proveedor
NIT	VARCHAR(25)	NIT del proveedor
IDSUCURSAL	INT	Identificador de sucursal
NOMBRESUCURSAL	VARCHAR(100)	Nombre de la sucursal
PKPROVEEDOR	INT	Identificador del proveedor de la base de datos FUENTE

FACTINSPECCIONES		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDINSPECCION	INT	Identificador del caso en el DW
PKIDCORRELATIVO	INT	Identificador del correlativo de inspecciones de la base de origen.
PKIDINSPECCION	INT	Identificador de inspecciones de la base de origen.
IDARTICULOLINEA	INT	Identificador de la línea de artículo en el DW
IDERRORLINEA	INT	Identificador de la línea de error en el DW
IDINSPECTORLINEA	INT	Identificador de la línea de inspector en el DW
IDTIEMPO	INT	Identificador de tiempo del DW. Llave foranea
IDESTADO	INT	Identificador de estado del DW. Llave foranea
IDARTICULO	INT	Identificador de artículo del DW. Llave foranea
IDRUBRO	INT	Identificador de rubro del DW. Llave foranea
IDERROR	INT	Identificador de error del DW. Llave foranea
IDINSPECTOR	INT	Identificador de inspector del DW. Llave foranea
HALLAZGO	INT	Cantidad de fallas que se encontraron al inspeccionar un proveedor
PRECIOOFRECIDO	DECIMAL(38,2)	Precio al que se presentaba los productos
PRECIOCAJA	DECIMAL(38,2)	Precio al que se cobraba los productos
EXISTENCIA	INT	Cantidad de productos verificados en la inspección.

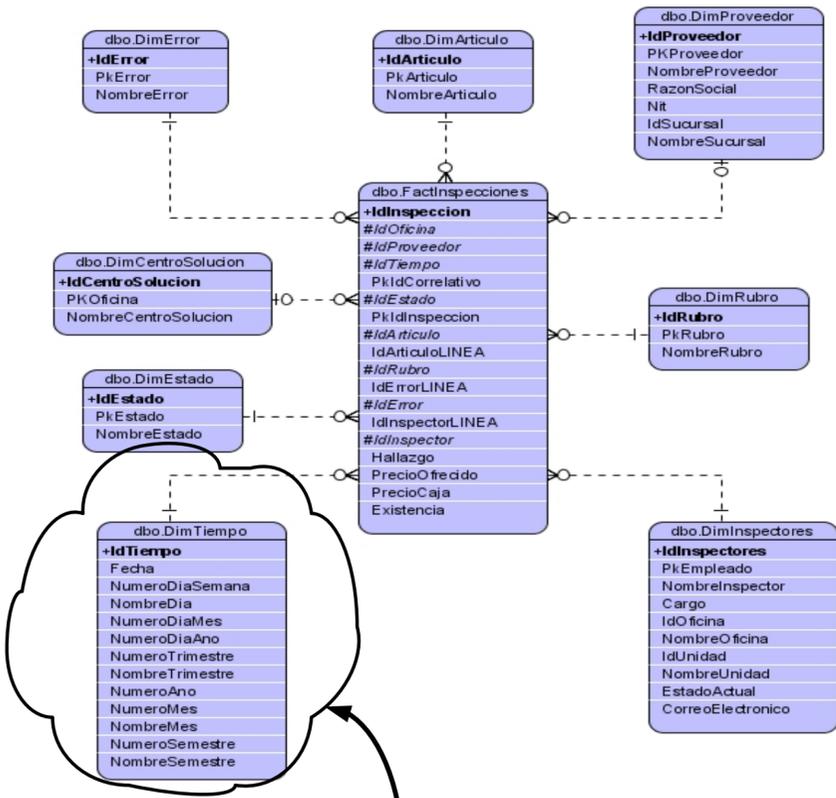
3.6.4. Diagrama UML del DATA WAREHOUSE – VERIFICACION Y VIGILANCIA

# DIAGRAMA UML DEL DATA WAREHOUSE- VERIFICACION Y VIGILANCIA

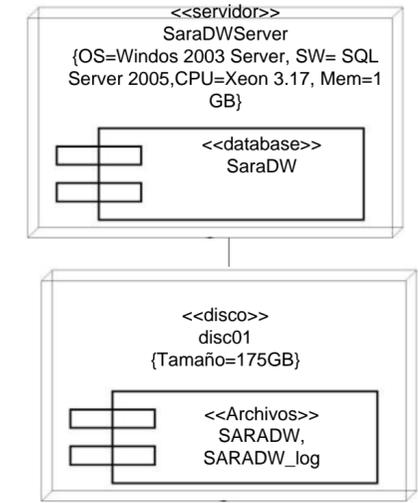
## ESQUEMA CONCEPTUAL DEL DATA WAREHOUSE



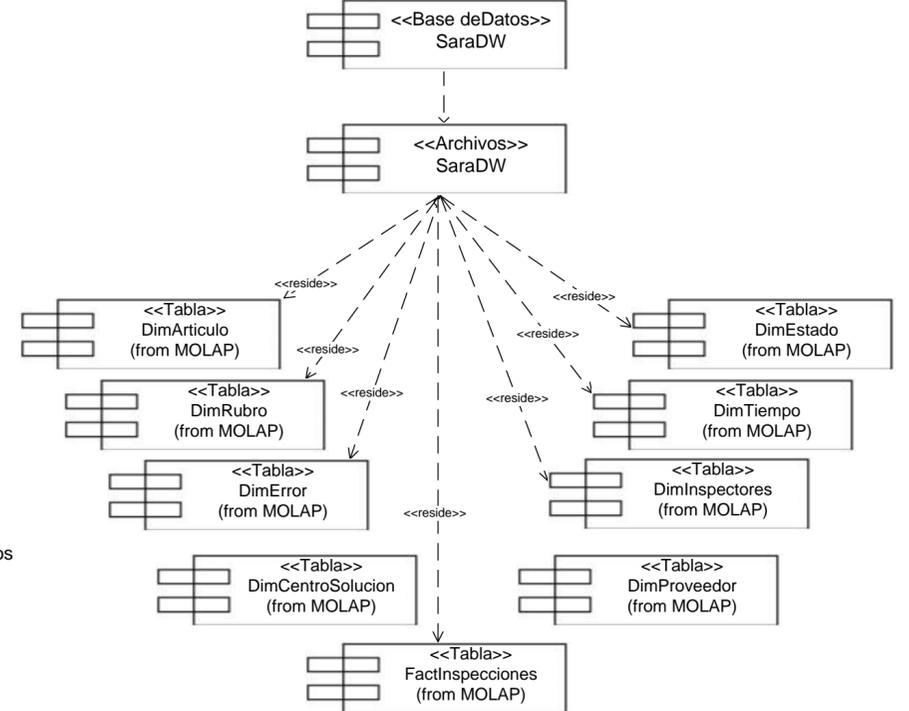
## ESQUEMA LÓGICO DEL DATA WAREHOUSE



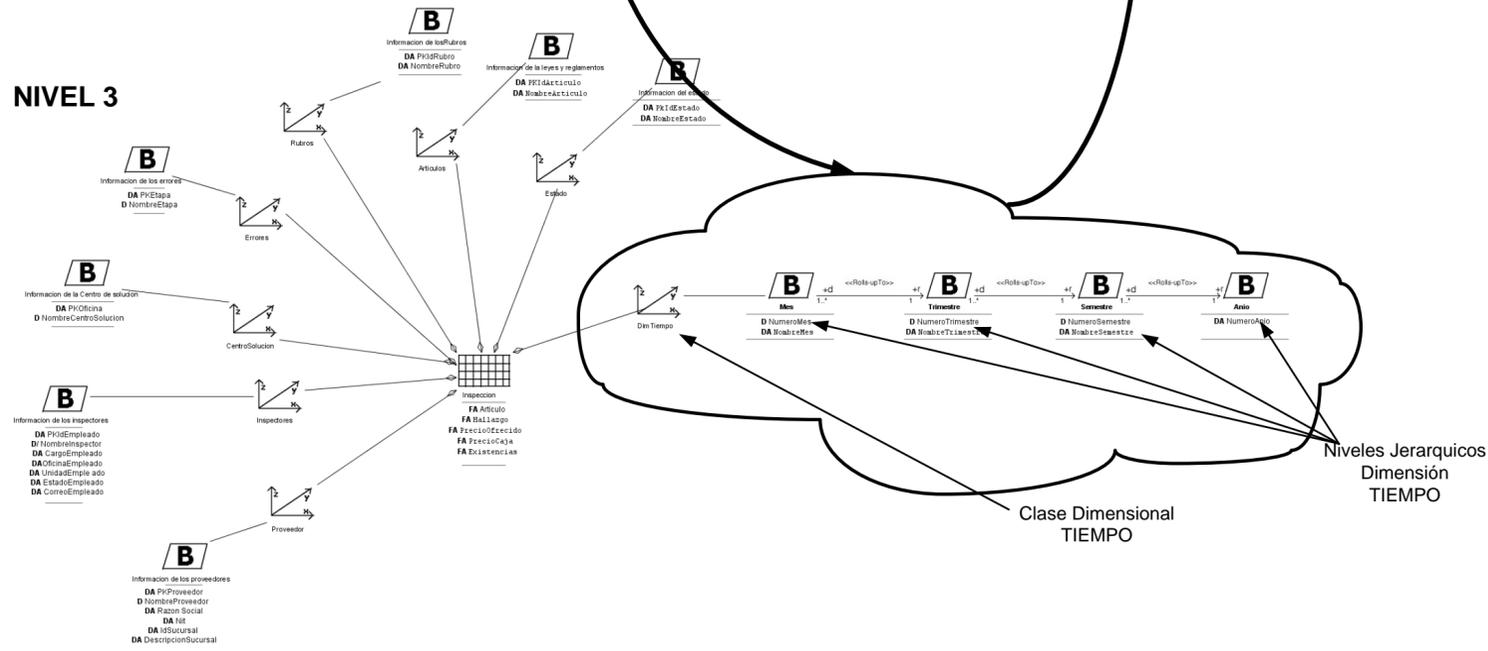
## ESQUEMA FISICO DEL DATA WAREHOUSE



## DIAGRAMA DE COMPONENTES

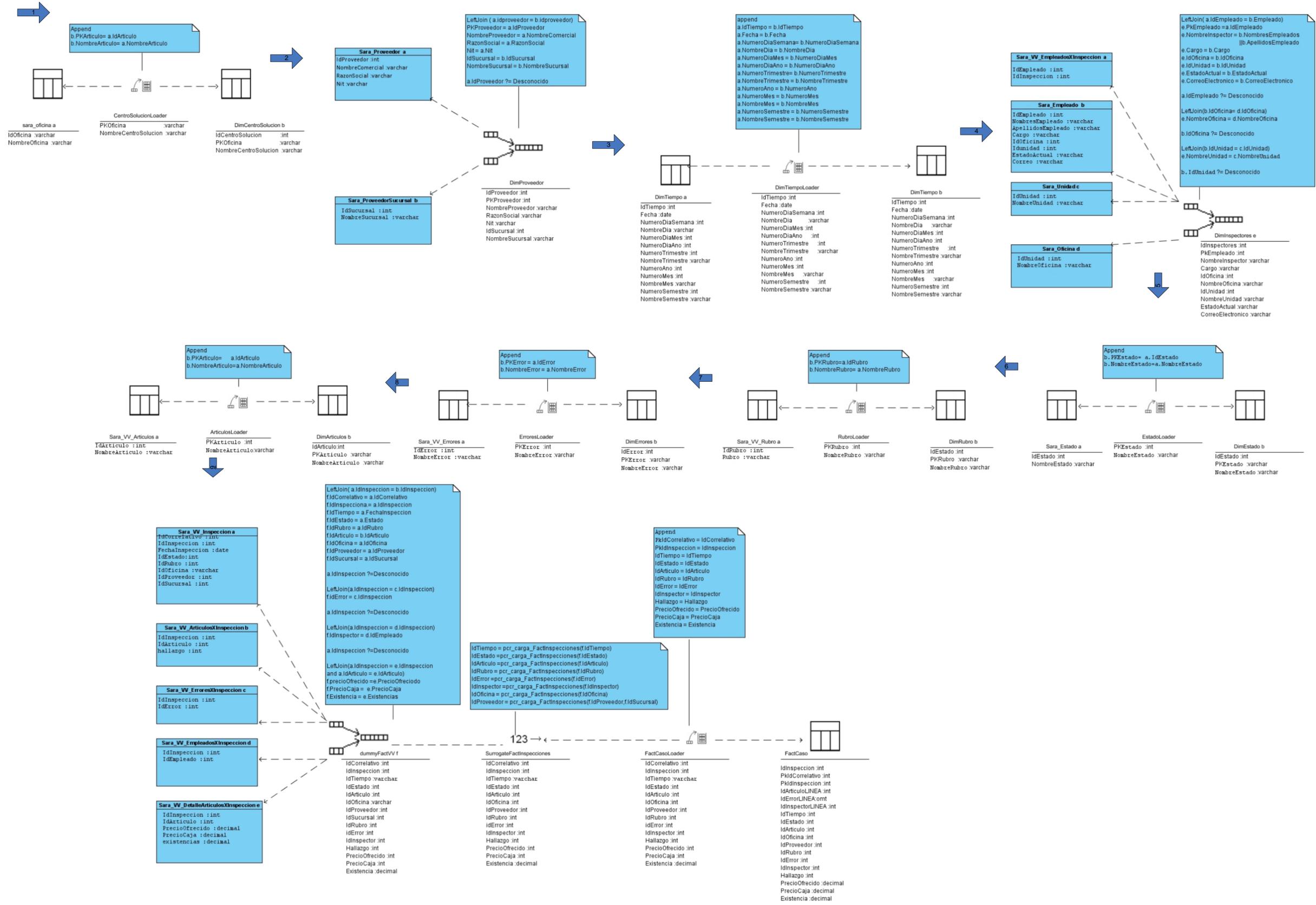


## NIVEL 3



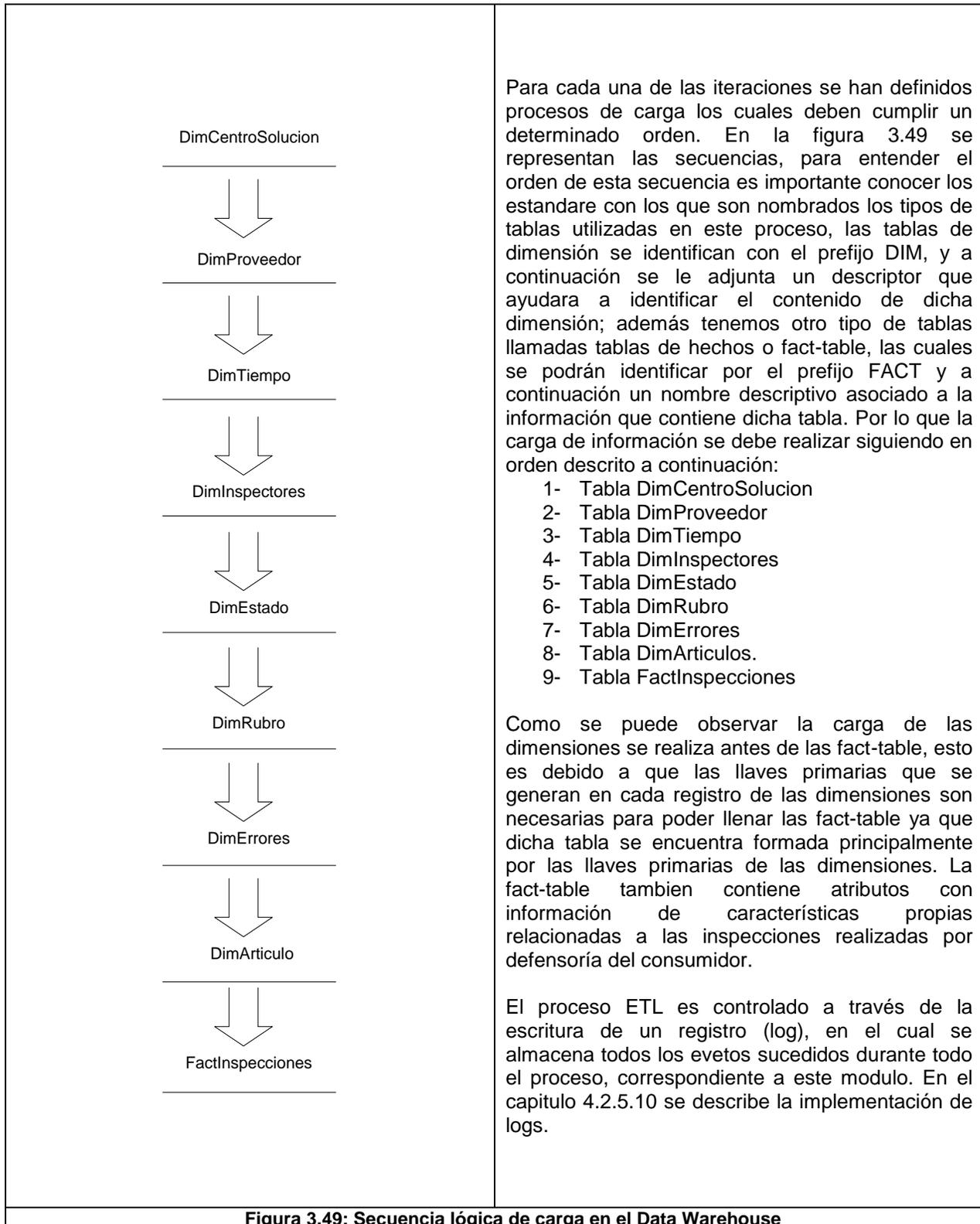
3.6.5. Diagrama del proceso ETL

### DIAGRAMA ETL – VERIFICACION Y VIGILANCIA



5  
Figura 3.48: Diseño del proceso ETL para el Data Warehouse

**3.6.5.1. Secuencia de carga del proceso ETL**



3.6.5.2. Diagrama de carga para la dimensión Centro Solución

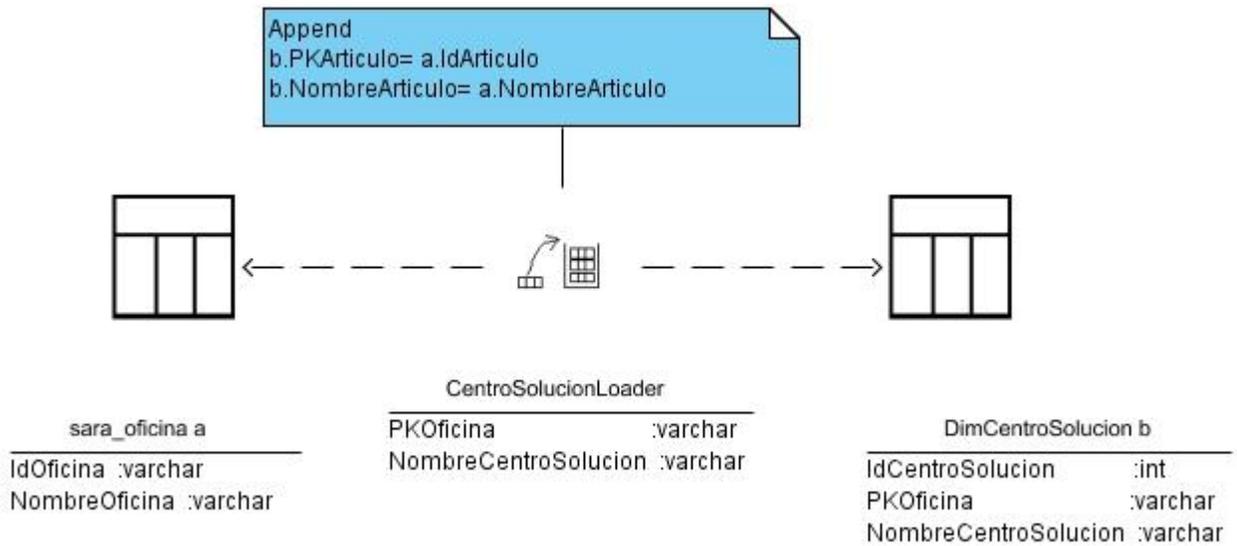


Figura 3.50: Representación de carga de datos en la tabla DimCentroSolucion utilizando mecanismo Loader

3.6.5.3. Diagrama de carga para la dimensión Proveedor

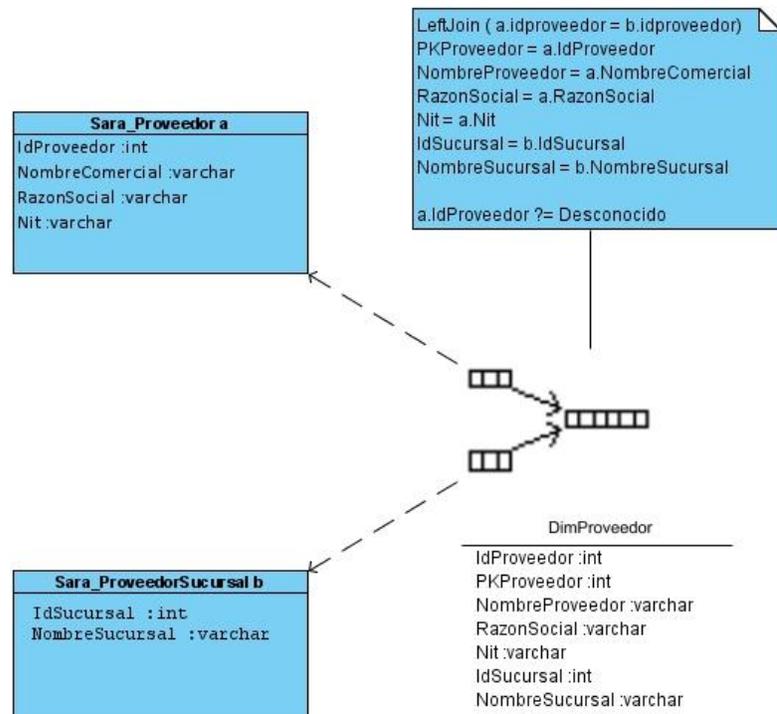


Figura 3.51: Representación de carga de datos en la tabla DimProveedor utilizando mecanismo Join

3.6.5.4. Diagrama de carga para la dimensión Tiempo

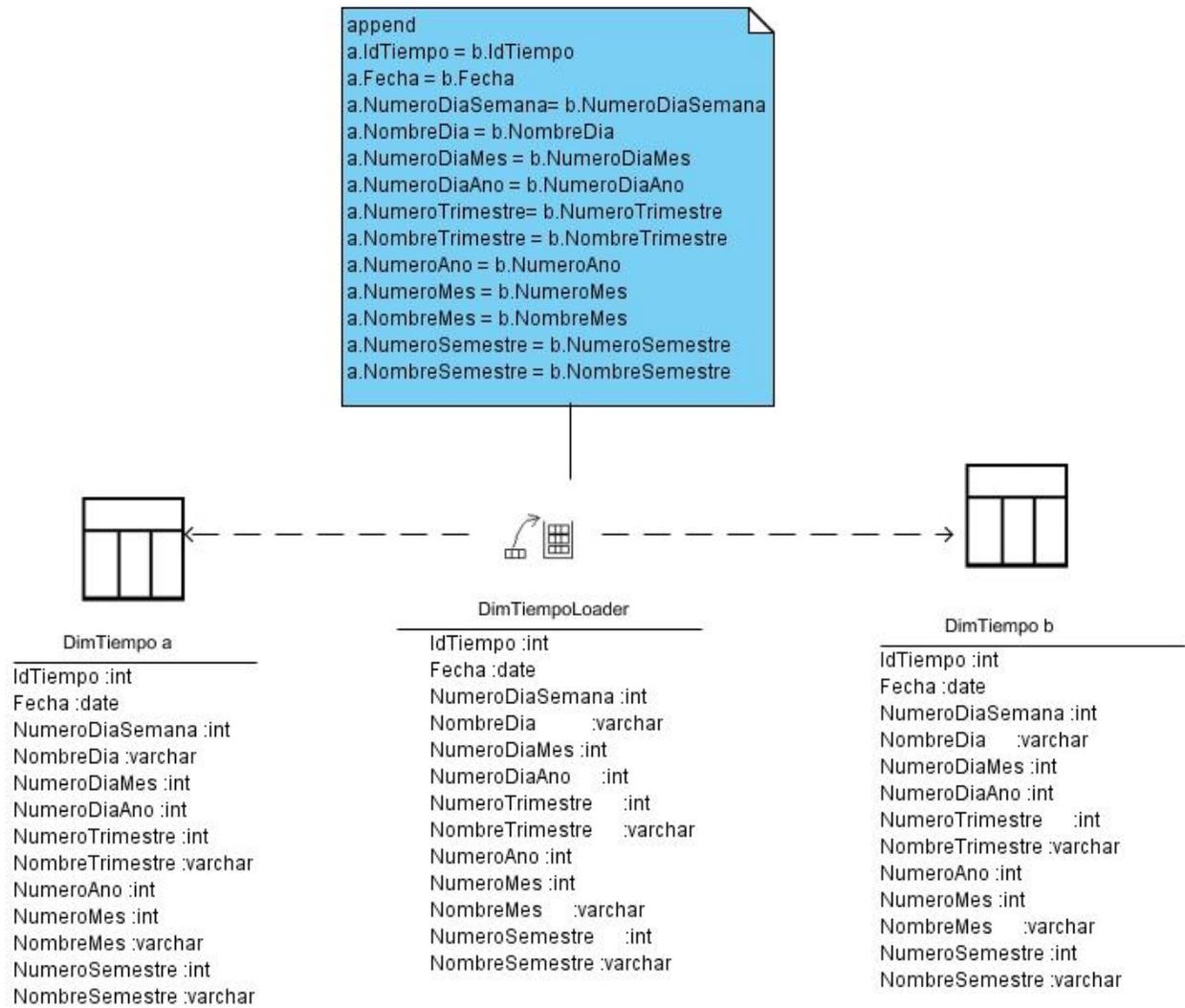


Figura 3.52: Representación de carga de datos en la tabla DimTiempo utilizando mecanismo Loader

3.6.5.5. Diagrama de carga para la dimensión Inspectores

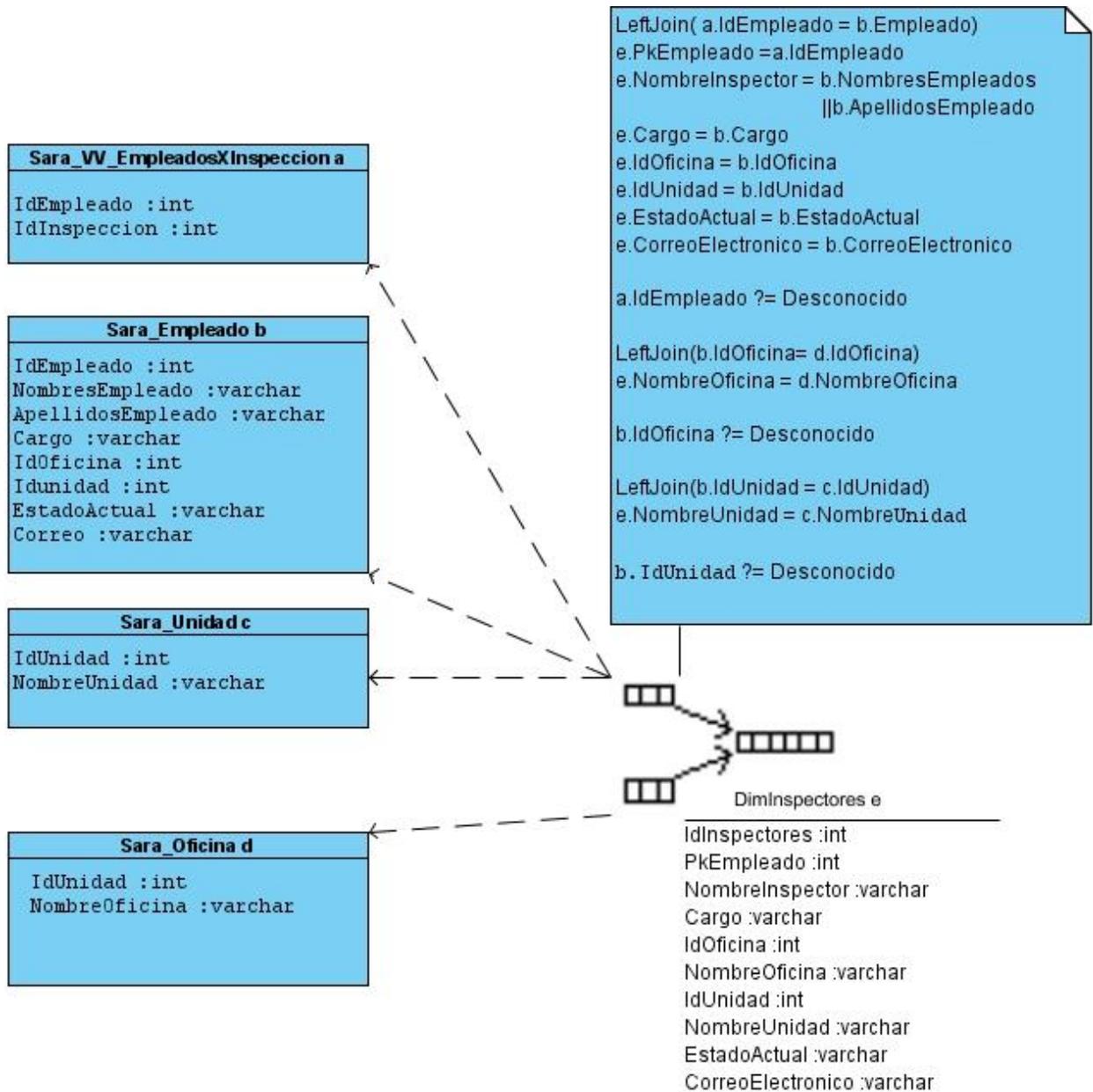


Figura 3.53: Representación de carga de datos en la tabla DimInspectores utilizando mecanismo Join

3.6.5.6. Diagrama de carga para la dimensión Estado

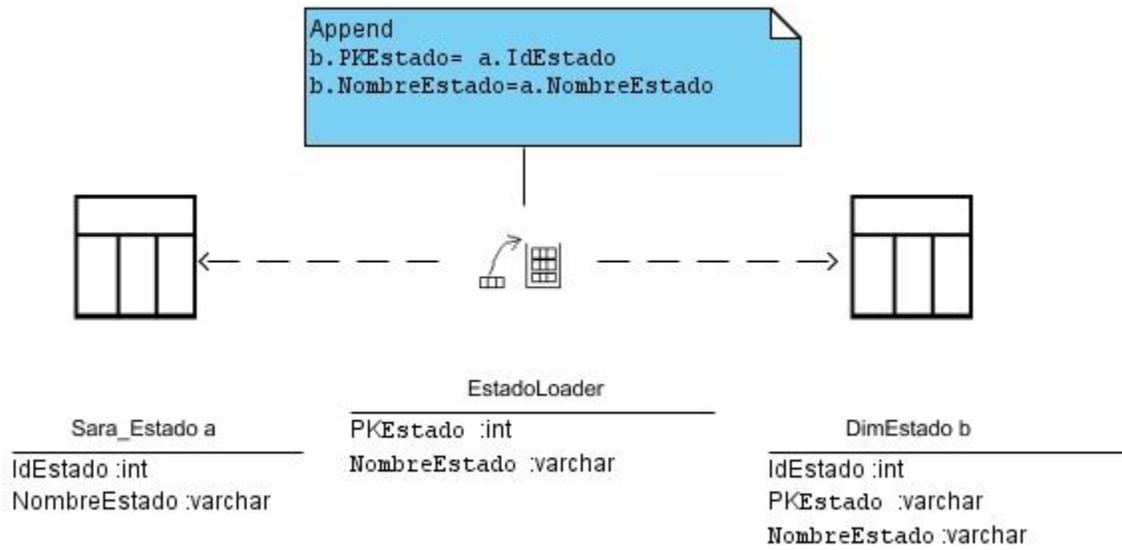


Figura 3.54: Representación de carga de datos en la tabla DimEstado utilizando mecanismo Loader

3.6.5.7. Diagrama de carga para la dimensión Rubro

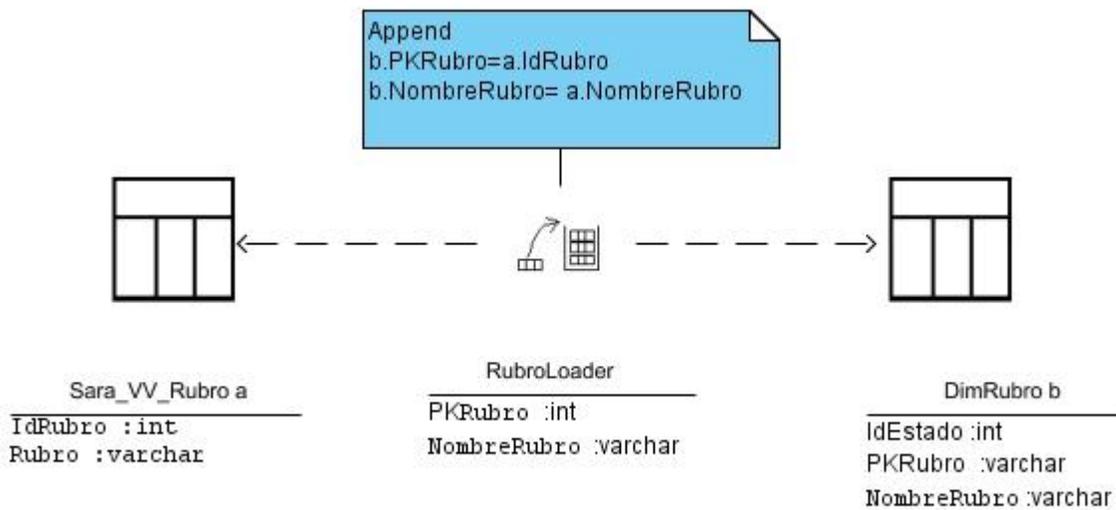


Figura 3.55: Representación de carga de datos en la tabla DimRubro utilizando mecanismo Loader

3.6.5.8. Diagrama de carga para la dimensión Error

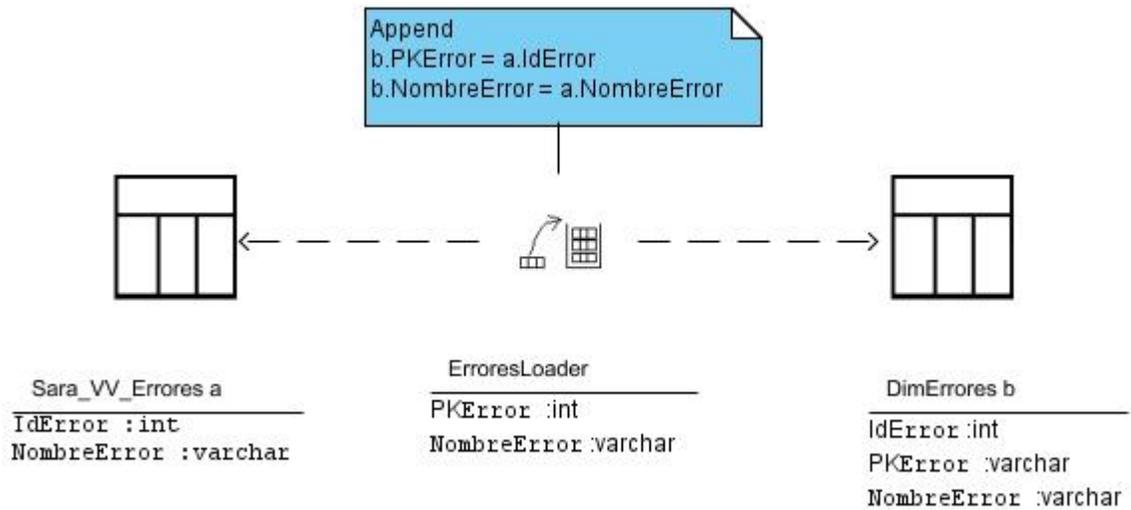


Figura 3.56: Representación de carga de datos en la tabla DimError utilizando mecanismo Loader

3.6.5.9. Diagrama de carga para la dimensión Artículo

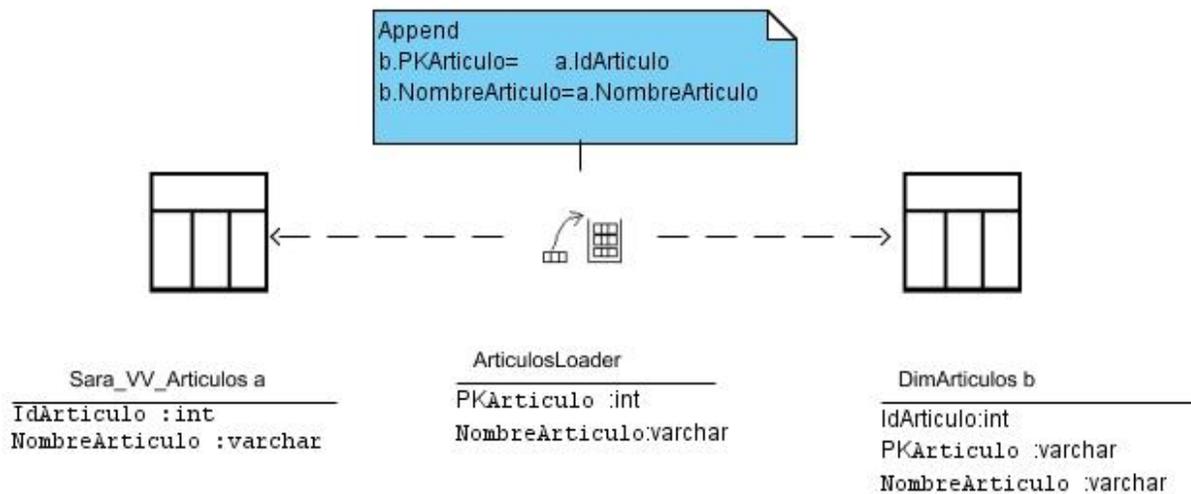


Figura 3.57: Representación de carga de datos en la tabla DimArticulo utilizando mecanismo Loader

### 3.6.5.10. Diagrama de carga para la FactInspecciones

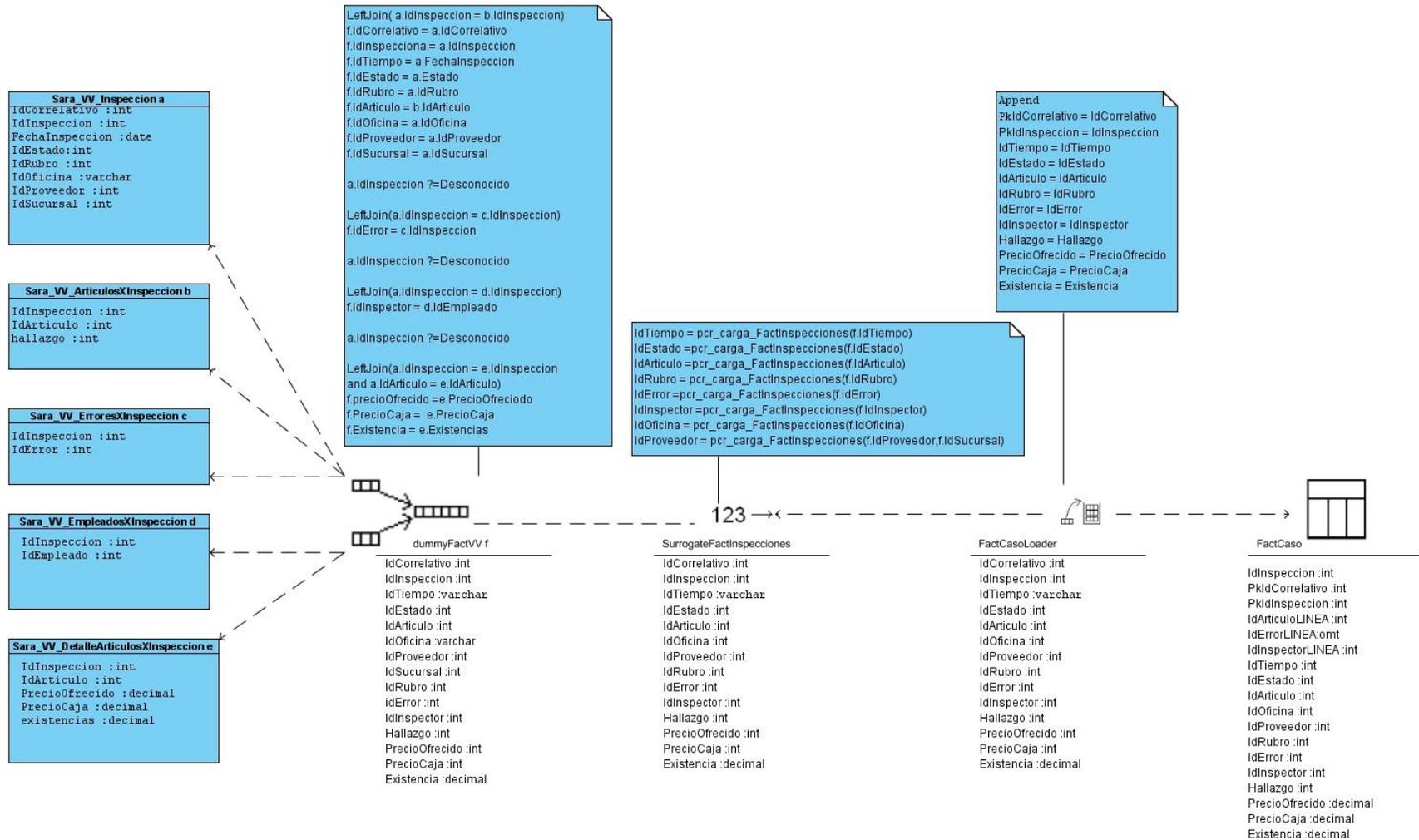


Figura 3.58: Representación de carga de datos en la tabla FactInspecciones utilizando mecanismo Join, Surrogate y Join

### 3.7. ITERACIÓN 3: TRIBUNAL SANCIONADOR

Se encarga de recibir dar seguimiento a los casos que pasan al área jurídica.

#### 3.7.1. Requerimientos Funcionales

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Página 1 de y	
Autor : Grupo 14		Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo
<b>01</b>	<b>Analizar proceso de Tribunal Sancionador</b>	Alta	FoO
<b>01.01</b>	Permitir al usuario analizar el proceso de los casos en Tribunal Sancionador determinando el tiempo, ubicación, proveedores y otros elementos de interés para la dirección de estudios.		
<b>01.01.01</b>	Analizar la cantidad de casos en Tribunal Sancionador permite determinar indicadores en los que se le deja a la ley como apoyo a la solución de casos.		
<b>01.01.02</b>	Los casos son tratados de manera similar al centro de solución de controversias pero se le da un énfasis mayor a los artículos de ley infringidos en los casos.		
<b>01.01.03</b>	Los montos recuperados también son un elemento importante por sector y/o categoría.		
<b>02</b>	<b>Analizar casos de Tribunal Sancionador según fecha</b>	Alta	FoO
<b>02.01</b>	El análisis de las inspecciones según la fecha permite determinar la cantidad casos que entran a Tribunal Sancionador en función del tiempo.		
<b>02.02</b>	El combinar las fecha con otras variables como la ubicación ayuda a identificar en que zona geográfica se genera la mayor cantidad de incidencias de este tipo de evento.		
<b>03</b>	<b>Analizar Casos involucrados en TS por sector</b>	Alta	FoO
<b>03.01</b>	Tener acceso a los sectores que generan mayor índice de violación de artículos de ley, facilita a la DC el hecho de orientar tanto a consumidores como proveedores en un tipo de negocio específico.		
<b>03.02</b>	Determinar los casos por sector y tiempo es un apoyo para identificar las temporadas en las cuales se generan mayor número de incidencias, esto también permite a la DC organizar campañas de control de las áreas o sectores más susceptibles a fallos.		
<b>04</b>	<b>Detalle de artículos de ley infringidos por los casos de que entran a TS</b>	Alta	FoO
<b>04.01</b>	El análisis de artículos relacionados es de vital importancia debido a que estos artículos son reglas las cuales tienen una base jurídica. Una inspección puede tener varios artículos relacionados		
<b>04.02</b>	Los artículos infringidos por cada caso pueden relacionarse con otras variables, tales como tiempo y sector, las cuales dependerán plenamente de la inspección		

<b>04.03</b>	Cada infracción de artículo de ley, sirve como insumo para el análisis de incidencias de incumplimiento de ley, además como respaldo para el consumidor o proveedor en el momento de aclarar detalles de infracción.		
<b>05</b>	<b>Reporte adhoc de la información analizada</b>		
<b>05.01</b>	Mostrar la información detallada de para poder constatar la información que es mostrada en los análisis.	Baja	<b>Fo</b>
<b>05.02</b>	Mostrar la información dando al usuario la libertad de elegir la información que desee que incluya el usuario		

### 3.7.2. Especificación De Requerimientos

#### 3.7.2.1. Crear Reporte Analítico



Figura 3.59: Diagrama de caso de uso para reportes analíticos

<p><b>Caso de uso 1 Crear Reporte Analítico</b></p> <p><b>Personal Involucrado e interés:</b></p> <p>Usuario: Técnico, Técnico Web, Dirección de estudios, Dirección Tribunal Sancionador, Gerencia informática.</p> <p>Usuario: Elaborar un análisis detallado del proceso realizado en la dirección de verificación y vigilancia para su posterior publicación o almacenamiento.</p> <p><b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo.</p>
---

<b>Postcondiciones:</b> Reporte analítico publicado , almacenado o impreso
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 01,02,03,04
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona medidas.</li> <li>6. El Sistema muestra datos de Tribunal Sancionador, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A,B,C,D,E,F.</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis</li> <li>10. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>11. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Realizar análisis utilizando MDX (Usuarios Expertos)</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción MDX.</li> <li>2. El sistema muestra opción de generación de consulta MDX.</li> <li>3. Usuario construye la consulta, utilizando consulta valida.</li> <li>4. El sistema realiza la construcción del espacio de trabajo en base a la consulta generada.</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>B. Configuración de tabla OLAP</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Configurar tabla OLAP.</li> <li>2. El sistema muestra ventana de configuración OLAP</li> <li>3. El usuario selecciona las opciones de configuración</li> <li>4. El sistema aplica configuraciones</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>C. Seleccionar opciones de las vistas de los datos</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Mostrar Padres.</li> <li>2. El sistema muestra las referencias fuentes de los datos mostrados.</li> <li>3. Sino el usuario selecciona la opción ocultar repeticiones.</li> <li>4. El sistema muestra solo los datos únicos (unique) sin repeticiones.</li> <li>5. Sino el usuario el usuario selecciona mostrar propiedades.</li> <li>6. El sistema muestra propiedades de los datos.</li> <li>7. Sino el usuario selecciona suprimir columnas vacías.</li> <li>8. El sistema elimina de la consulta aquellos datos que contengan valores nulos.</li> <li>9. Sino el usuario selecciona la opción intercambiar ejes.</li> <li>10. El sistema genera un análisis roll-up.</li> <li>11. El usuario analiza los datos.</li> </ol>

**D. Generación y configuración de gráfico**

1. El usuario selecciona la opción configuración de gráfico.
2. El sistema muestra propiedades de gráfico.
3. El usuario configura opciones de gráfico, colores, leyendas, orientación y tipo de gráfico.
4. El sistema adecua el gráfico a los valores que se muestran en el espacio de trabajo.
5. El usuario selecciona opción de generar gráfico.
6. El sistema muestra gráfico.
7. El usuario analiza los datos.

**E. Seleccionar el Destino del análisis**

1. El usuario selecciona opción de configuración de impresión.
2. El sistema muestra panel de configuración de impresión.
3. El usuario selecciona opciones de configuración tales como tema de reporte, orientación, tamaño del papel.
4. El sistema configura opciones de reporte analítico.
5. El usuario publica el reporte en formato PDF, EXCEL, CSV o lo almacena en el sistema.
6. El sistema procesa el reporte y lo publica dependiendo el método seleccionado.

**F. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.7.2.2. Crear Reporte AD HOC

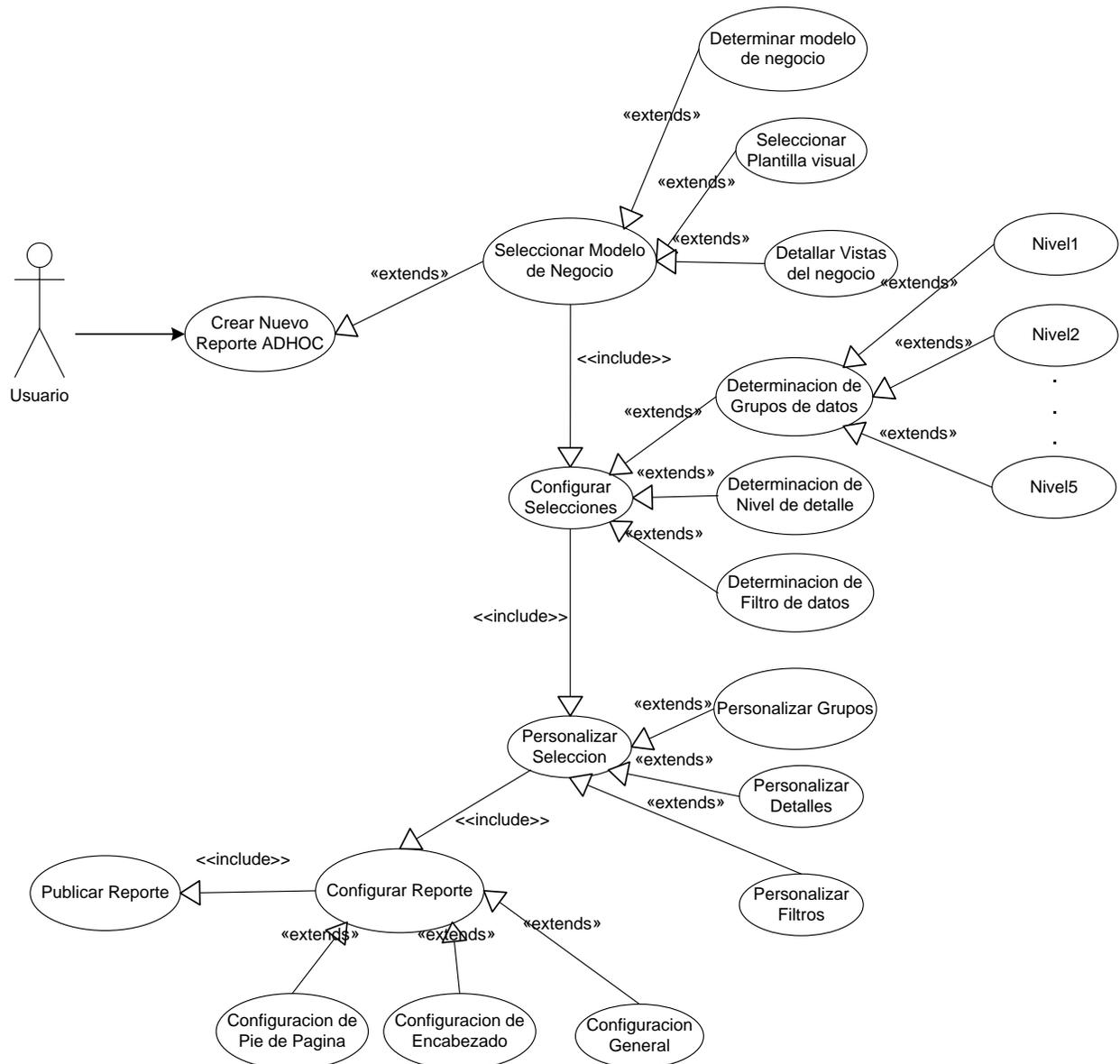


Figura 3.60: Diagrama de caso de uso para reportes AD HOC

<b>Caso de uso 2 Crear Reporte AD HOC</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Técnico Web, Dirección de Estudios, Gerencia informática.  Usuario: Elaborar de reportes temporales para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema poseer acceso a la vista de negocios de Tribunal Sancionador
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> Requerimientos: 05

**Flujo Básico**

1. El Usuario selecciona "Nuevo Reporte".
2. El sistema muestra los modelos de negocio disponibles.
3. Usuario selecciona el modelo de negocio con el que desea trabajar.
4. Sistema muestra las vistas de negocio disponibles para el modelo seleccionado.
5. El usuario procede a la siguiente fase de selección de reporte
6. El sistema muestra variables con las que se puede agrupar aplicar filtros y seleccionar niveles de detalle de los grupos definidos
7. El usuario selecciona los grupos, filtros y niveles de detalle con las que desea realizar el reporte
8. El usuario procede a la siguiente fase de personalización de selecciones
9. El sistema muestra los valores de los niveles de detalle, grupos y filtros con los que se desea generar el reporte.
10. El sistema almacena la configuración de los datos.
11. El usuario procede a la siguiente fase de configuración de reporte
12. El sistema muestra panel de configuración de reporte.
13. El usuario selecciona las opciones generales del reporte, tipo de papel orientación encabezado y pie de página.
14. El sistema almacena la configuración
15. El usuario almacena el reporte ad hoc generado para futuras consultas
16. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.

**Flujo Alternativo****A. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.7.2.3. Análisis de Inspecciones para Tribunal Sancionador



Figura 3.61: Diagrama de caso de uso para análisis de los casos en el TS

Caso de uso 3 Análisis de Inspecciones para Tribunal Sancionador
<p><b>Personal Involucrado e interés:</b></p> <p>Usuario: Técnico, Técnico Web, Dirección de Estudios, Gerencia informática.</p> <p>Usuario: Elaborar un análisis detallado del centro de solución de controversias para su posterior publicación o almacenamiento.</p>

<p><b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo de Tribunal Sancionador.</p>
<p><b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.</p>
<p><b>Referencias Cruzadas</b>  <b>Requerimientos:</b> 01,02,03,04,05</p>
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo de verificación y vigilancia para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona las medidas.</li> <li>6. El Sistema muestra datos de Tribunal Sancionador, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis con medidas tales como: cantidad de infracciones en el tiempo; combinando dimensiones como: tiempo e Infracción.</li> <li>10. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>11. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Cancelar la operación</b></p> <ol style="list-style-type: none"> <li>1. El usuario Cierra el sistema</li> <li>2. El sistema cancela las operaciones y finaliza sesión de usuario.</li> </ol>

### 3.7.3. Modelado Iteración 3 Tribunal Sancionador

#### 3.7.3.1. Esquema Conceptual del Data Warehouse

En Tabla 1.3 se encuentra la definición de la nomenclatura UML utilizada para la definición de los estereotipos del modelado conceptual.

##### Nivel 1. Definición del Modelo

Tribunal Sancionador:

Especialización del Negocio para análisis de los casos.

**Hecho principal:** El Estudio de los casos que ingresan a Tribunal Sancionador



Figura 3.62: Definición modelo “Tribunal Sancionador”

##### Nivel 2 Definición del esquema estrella

Paquete de Hecho: Casos

Paquetes Dimensionales

Infracción: Datos relevantes de los artículos infringidos en los casos que entran a TS.

Categoría: Se refiere al tipo de denuncia por el cual el consumidor solicita ayuda.

Proveedor: Generalmente considerado como el ente al cual se le demanda algo.

ProcesoSubproceso: El proceso el cual lleva la denuncia,

Ubicación: Contiene la ubicación geográfica del país respecto a los proveedores.

Empleado: Catalogo de empleados de la DC

Tiempo: Control de tiempo en años, días, cuartiles, meses, semanas.

Estado: Estado del caso.

Centro de Solución: Nombre del centro de solución en el que se recibe la denuncia.

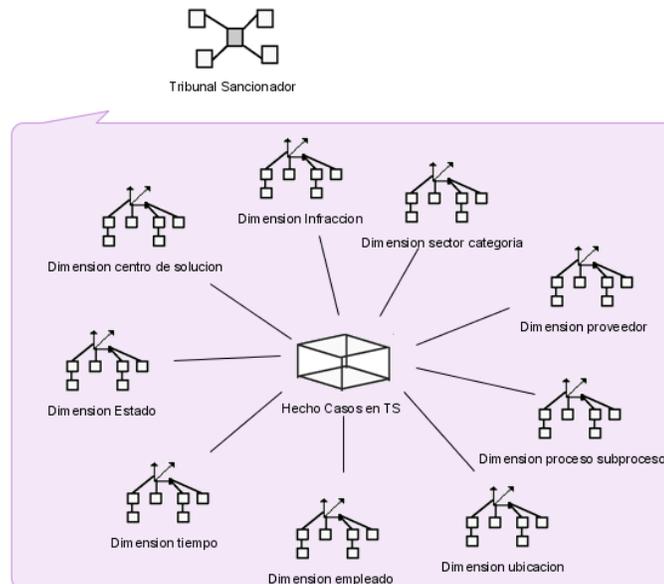
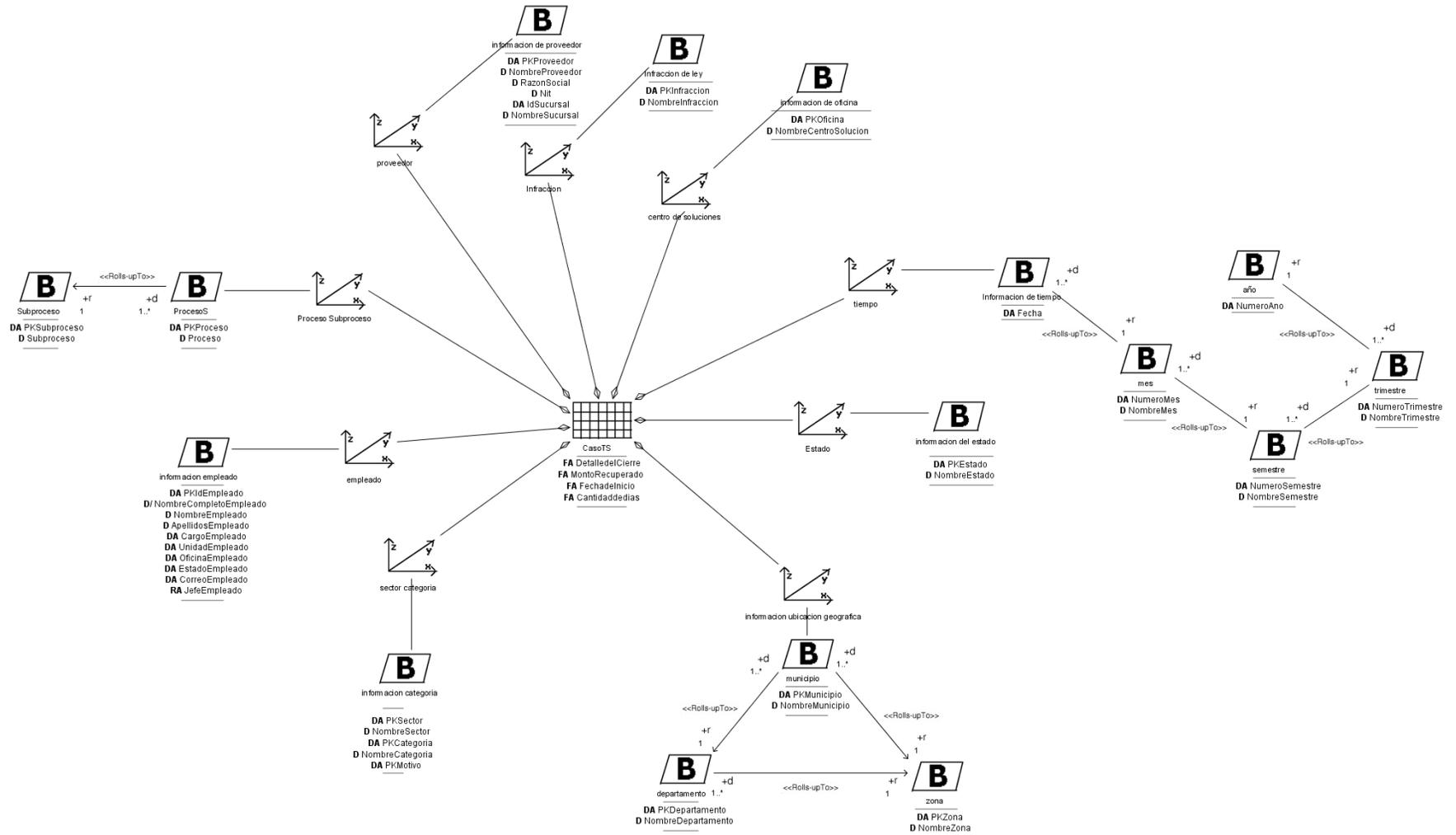


Figura 3.63: Definición del esquema estrella para “Trib”

**Nivel 3: Definición de Dimensiones/Hecho.**

Especificación de atributos.

Dimensiones candidatas a roll up y drill-down: tiempo y ubicación.



**Figura 3.64: Definición de las dimensiones y hechos “Tribunal Sancionador”**

3.7.3.2. Esquema lógico del Data Warehouse

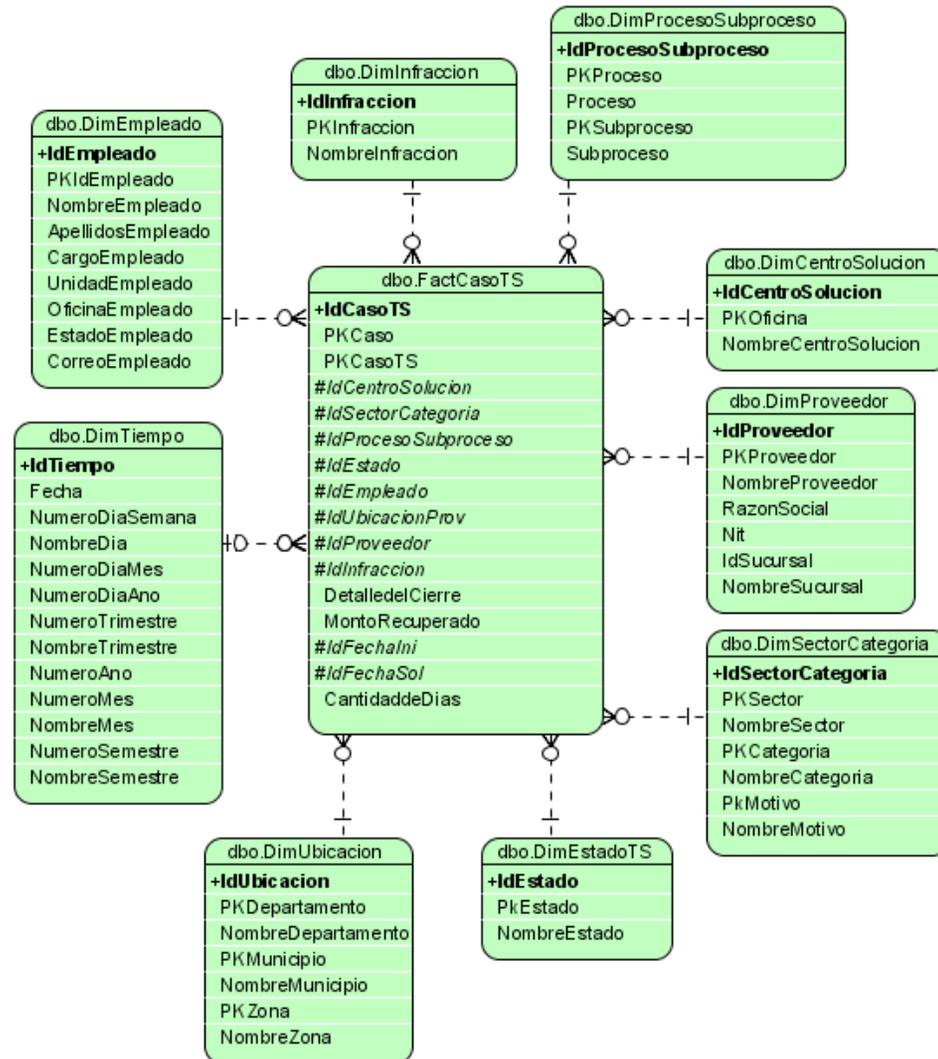


Figura 3.65: Estructura lógica del esquema estrella para el Tribunal Sancionador

**3.7.3.3. Diccionario de datos de las dimensiones y tablas de hecho**

DIMINFRACCION		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDINFRACCION	INT	Identificador de la infraccion en el DW
PKINFRACCION	INT	Identificador de la infraccion de la base de datos FUENTE
NOMBREINFRACCION	VARCHAR(100)	Descriptor de la Infraccion nombre del articulo.

DIMSECTORCATEGORIA		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDSECTORCATEGORIA	INT	Identificador del sector categoría en el DW
NOMBRECATEGORIA	VARCHAR(100)	Nombre de la categoría
PKSECTOR	VARCHAR(3)	Identificador del sector de la base FUENTE
NOMBRESECTOR	VARCHAR(100)	Nombre del sector
PKCATEGORIA	VARCHAR(4)	Identificador de la categoría de la base FUENTE

DIMPROVEEDOR		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDPROVEEDOR	INT	Identificador del proveedor en el DW
NOMBREPROVEEDOR	VARCHAR(200)	Nombre de proveedor
RAZONSOCIAL	VARCHAR(200)	Razón social del proveedor
NIT	VARCHAR(25)	NIT del proveedor
IDSUCURSAL	INT	Identificador de sucursal
NOMBRESUCURSAL	VARCHAR(100)	Nombre de la sucursal
PKPROVEEDOR	INT	Identificador del proveedor de la base de datos FUENTE

DIMPROCESOSUBPROCESO		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDPROCESOSUBPROCESO	INT	Identificador de la forma de pago en el DW
PKPROCESO	VARCHAR(3)	Identificador del proceso de la base FUENTE
PROCESO	VARCHAR(MAX)	Descripción del proceso
PKSUBPROCESO	VARCHAR(4)	Identificador del subproceso de la base FUENTE
SUBPROCESO	VARCHAR(MAX)	Descripción del subproceso

DIMUBICACION		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDUBICACION	INT	Identificador de la ubicación en el DW
PKDEPARTAMENTO	VARCHAR(3)	Identificador del departamento de la base FUENTE
NOMBREDEPARTAMENTO	VARCHAR(50)	Nombre del departamento
PKMUNICIPIO	VARCHAR(4)	Identificador del municipio de la base FUENTE
NOMBREMUNICIPIO	VARCHAR(50)	Nombre del municipio
PKZONA	VARCHAR(3)	Identificador de la zona de la base FUENTE
NOMBREZONA	VARCHAR(50)	Nombre de la zona

DIMESTADO		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDESTADO	INT	Identificador de la dimensión estado del DW.
PKESTADO	INT	Identificador de los estados de la base de origen
NOMBREESTADO	VARCHAR(50)	Descripción del Estado

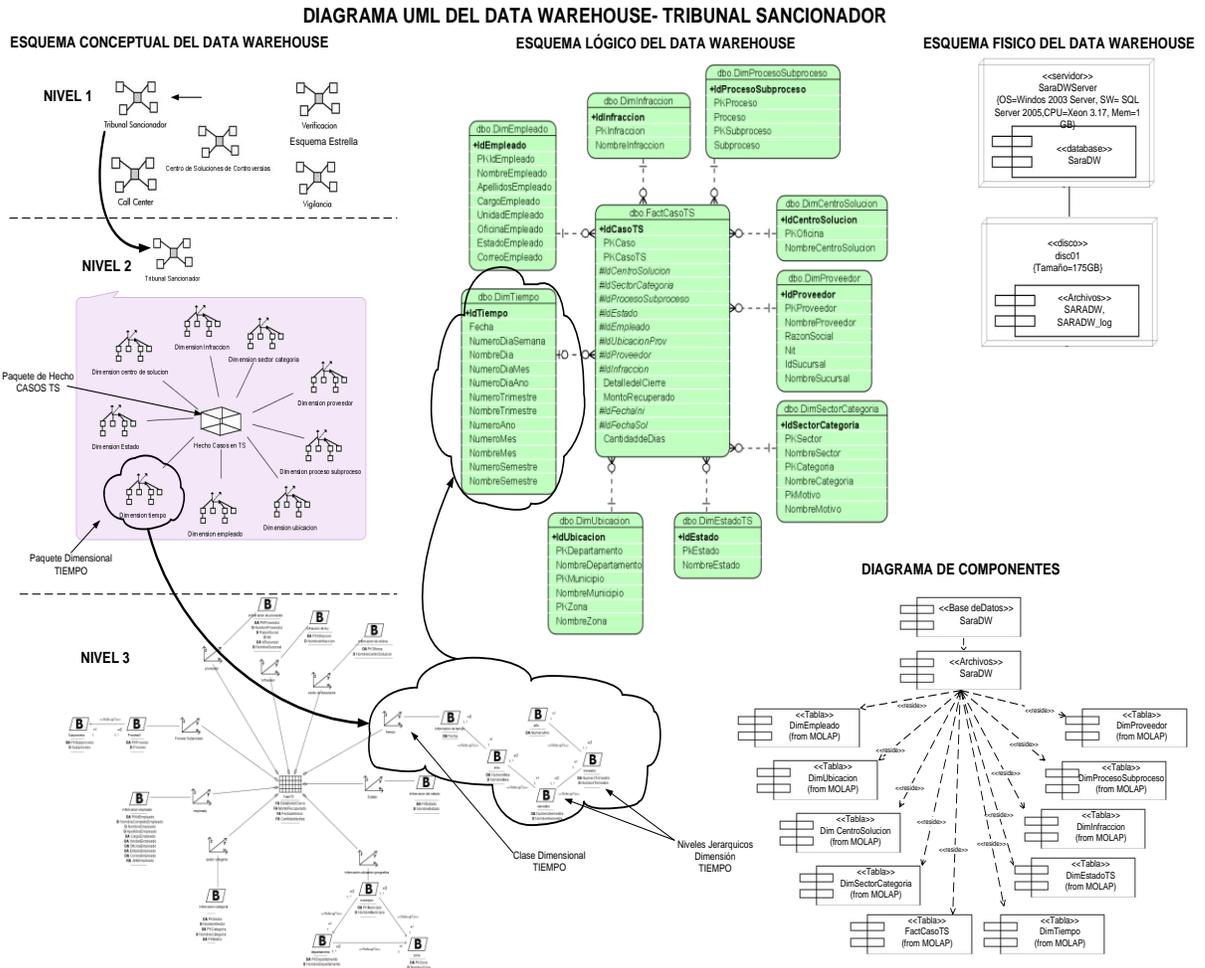
DIMEMPLEADO		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDEMPLEADO	INT	Identificador del empleado en el DW
NOMBREEMPLEADO	VARCHAR(50)	Nombre del empleado
APELLIDOSEMPLEADO	VARCHAR(50)	Apellido de empleado
CARGOEMPLEADO	VARCHAR(50)	Cargo del empleado
UNIDADEMPLEADO	VARCHAR(50)	Unidad a la que pertenece el empleado
OFICINAEMPLEADO	VARCHAR(100)	Oficina a la que forma parte el empleado
ESTADOEMPLEADO	VARCHAR(50)	Estado del empleado Activo o Desactivo
CORREOEMPLEADO	VARCHAR(100)	Correo electrónico del empleado
PKIDEMPLEADO	INT	Identificador del empleado de la base de datos FUENTE

DIMTIEMPO		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDTIEMPO	INT	Identificador de tiempo del DW
FECHA	DATETIME	Fecha de ingreso de la denuncia
NUMERODIASSEMANA	INT	Numero del día en la semana
NOMBREDIA	VARCHAR(50)	Nombre del día de la semana
NUMERODIAMES	INT	Numero del día en el mes
NUMERODIAANO	INT	Numero del día en el año
NUMEROTRIMESTRE	INT	Numero del trimestre del año
NOMBRETRIMESTRE	VARCHAR(50)	Nombre del trimestre del año
NUMEROANO	INT	Numero del año
NUMEROMES	INT	Numero del mes
NOMBREMES	VARCHAR(50)	Nombre del mes
NUMEROSEMESTRE	INT	Numero del semestre
NOMBRESEMESTRE	VARCHAR(50)	Nombre del semestre

DIMCENTROSOLUCION		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDCENTROSOLUCION	INT	Identificador del centro de solución en el DW
NOMBRECENTROSOLUCION	VARCHAR(100)	Nombre de centro de solución
PKOFICINA	VARCHAR(2)	Identificador de la oficina de la base de datos FUENTE

FACTCASOTS		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDCASO	INT	Identificador del caso en el DW
IDCENTROSOLUCION	INT	Identificador del centro de solución en el DW llave foránea
IDSECTORCATEGORIA	INT	Identificador de sector categoría en el DW llave foránea
IDFORMAPAGO	INT	Identificador de forma de pago en el DW llave foránea
IDETAPA	INT	Identificador de la etapa en el DW llave foránea
IDEMPLEADO	INT	Identificador del empleado en el DW llave foránea
IDUBICACION	INT	Identificador de la ubicación en el DW llave foránea
IDCONSUMIDOR	INT	Identificador del consumidor en el DW llave foránea
IDPROVEEDOR	INT	Identificador del proveedor en el DW llave foránea
IDTIEMPO	INT	Identificador del tiempo en el DW llave foránea
PKIDCASO	INT	Llave primaria del caso en la base de datos FUENTE
IDUBICACIONPROV	INT	Identificador de la ubicación del proveedor
IDTIEMPOADQUISICION	INT	Identificador del tiempo de adquisición
IDTIEMPOINGRESO	INT	Identificador del tiempo de ingreso del caso
TIPOCASO	VARCHAR(50)	Tipo del caso de la denuncia
FORMARECEPCION	VARCHAR(50)	Forma de recepción del caso en centro o por teléfono etc.

### 3.7.4. Diagrama UML del DATA WAREHOUSE – TRIBUNAL SANCIONADOR



3.7.4.1. Diagrama del proceso ETL

DIAGRAMA ETL – TRIBUNAL SANCIONADOR

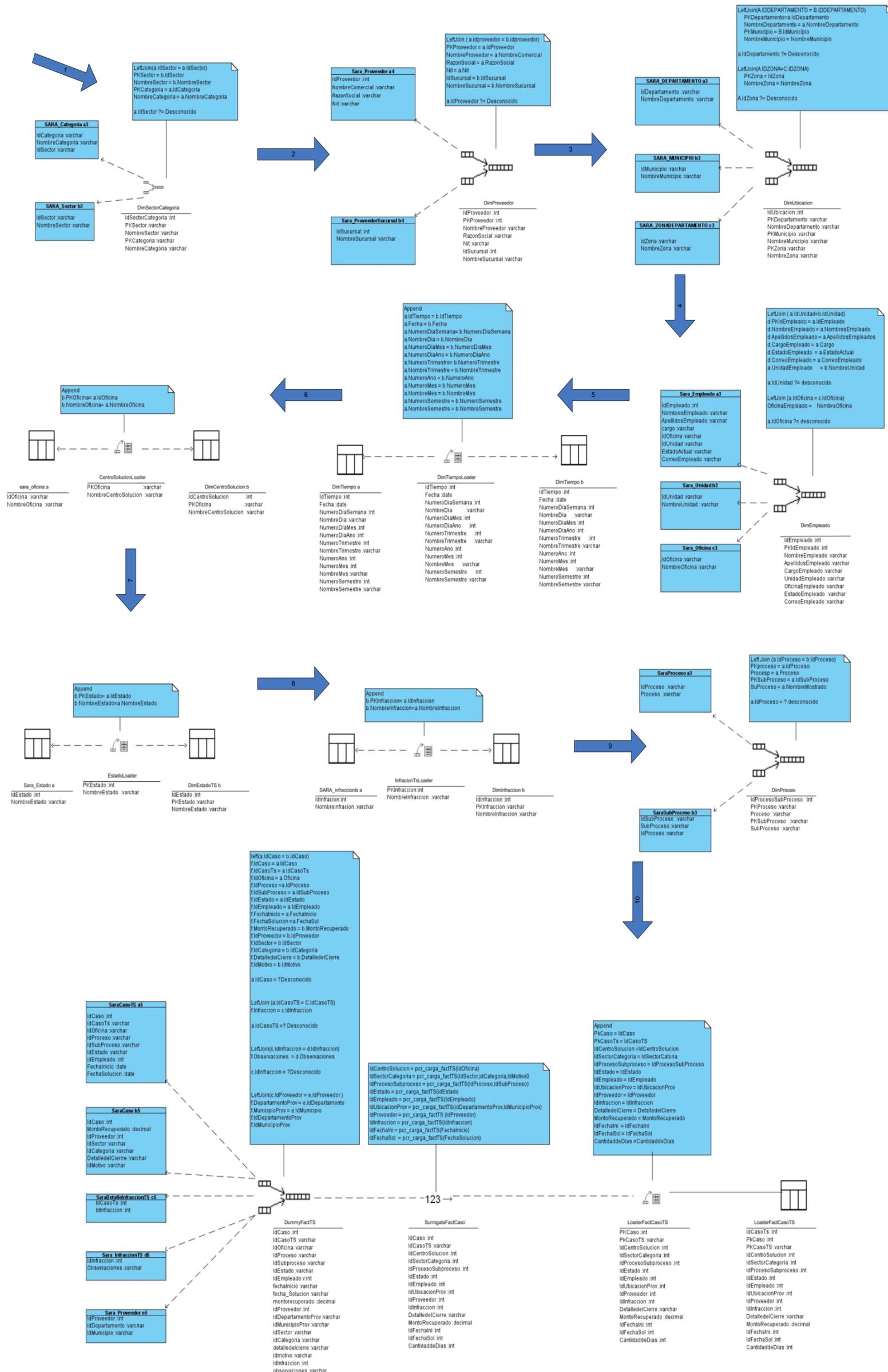
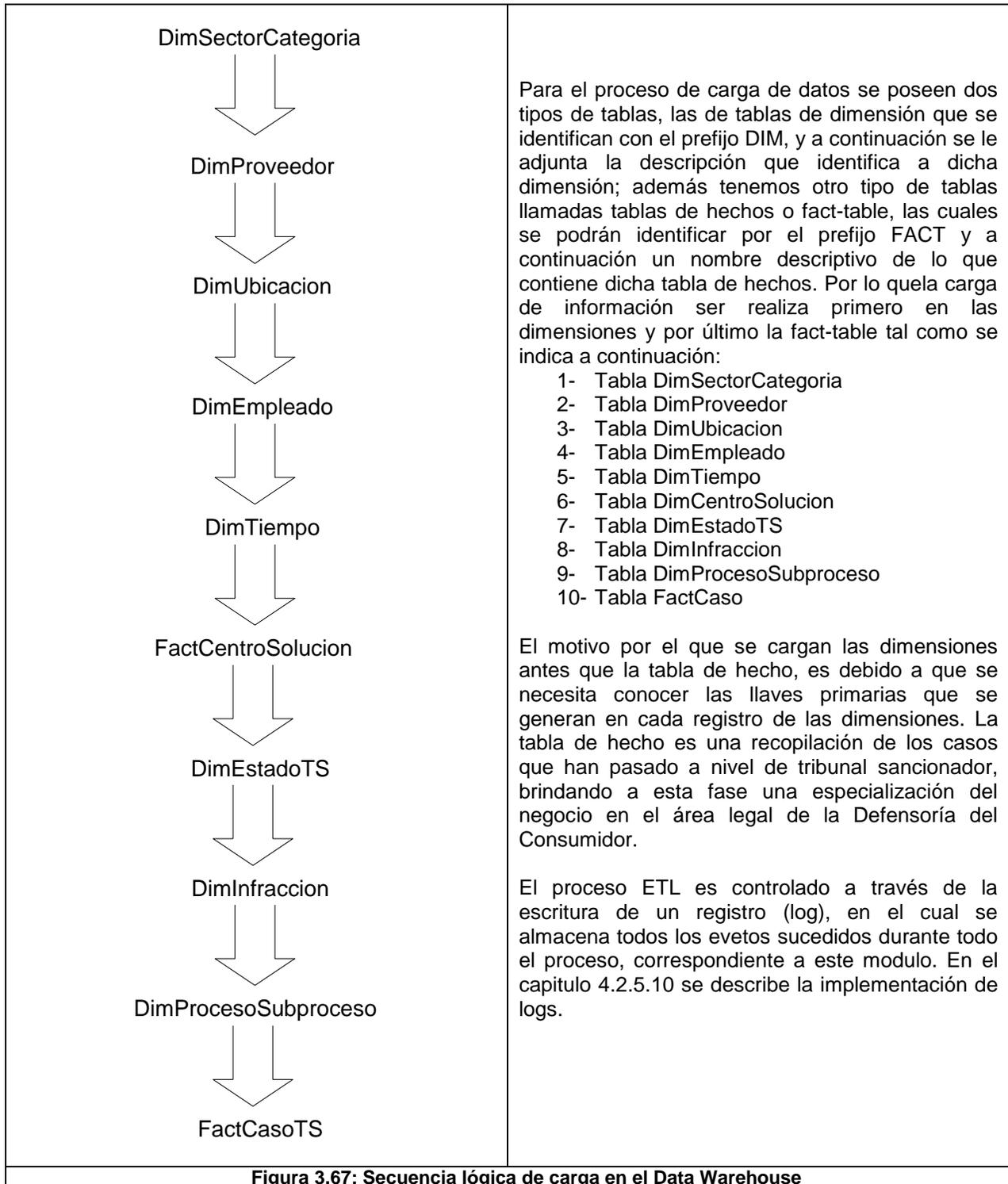


Figura 3.66: Diseño del proceso ETL para el Data Warehouse

### 3.7.4.2. Secuencia de carga del proceso ETL



### 3.7.4.3. Diagrama de carga para la dimensión EstadoTS

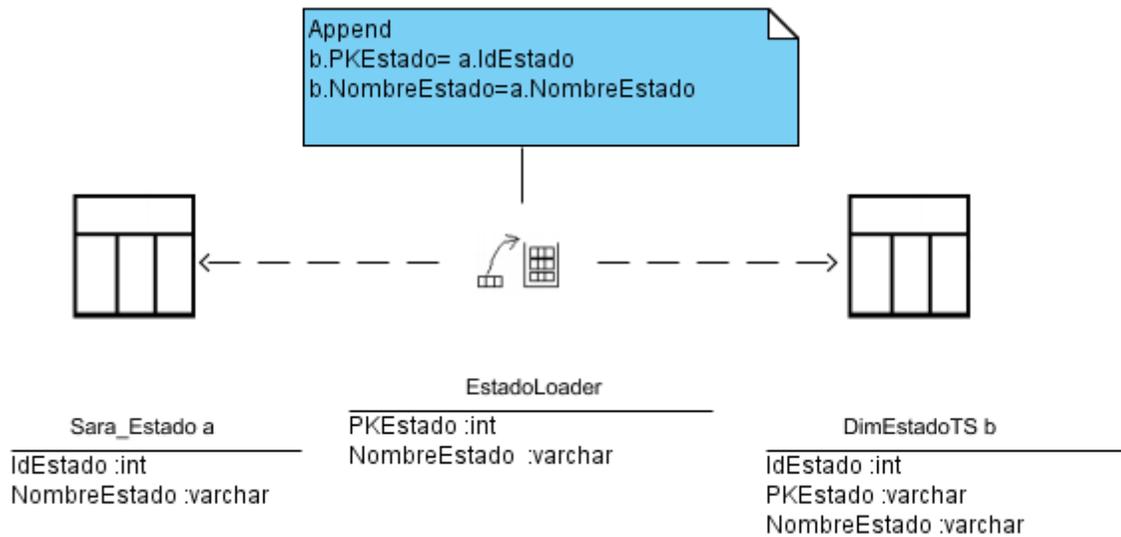


Figura 3.68: Representación de carga de datos en la tabla DimEstadoTS utilizando mecanismo Loader

### 3.7.4.4. Diagrama de carga para la dimensión Infracción

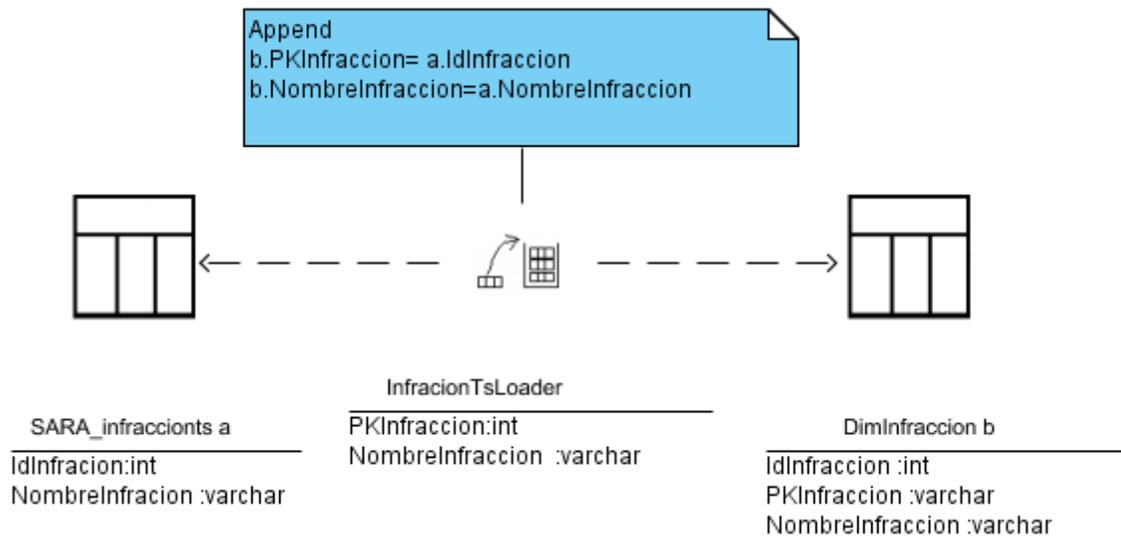


Figura 3.69: Representación de carga de datos en la tabla DimInfraccion utilizando mecanismo Loader

### 3.7.4.5. Diagrama de carga para la dimensión Proceso Subproceso

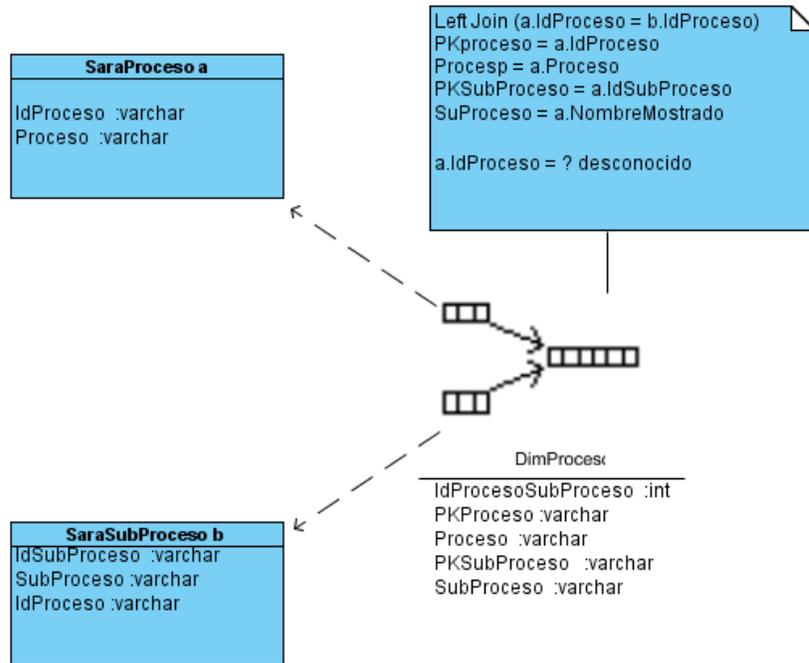


Figura 3.70: Representación de carga de datos en la tabla DimProcesoSubproceso utilizando mecanismo Join

### 3.7.4.6. Diagrama de carga para la dimensión Tiempo

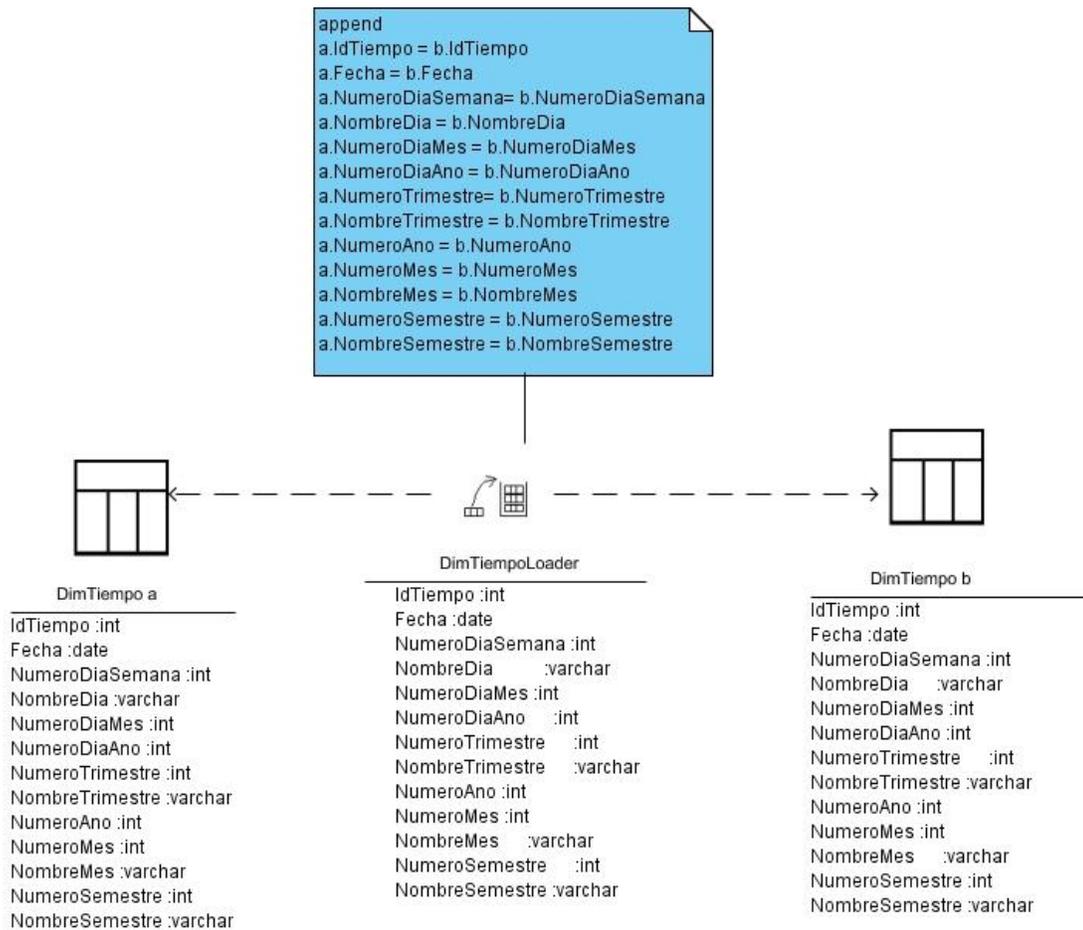


Figura 3.71: Representación de carga de datos en la tabla DimTiempo utilizando mecanismo Loader

### 3.7.4.7. Diagrama de carga para la dimensión Centro Solución

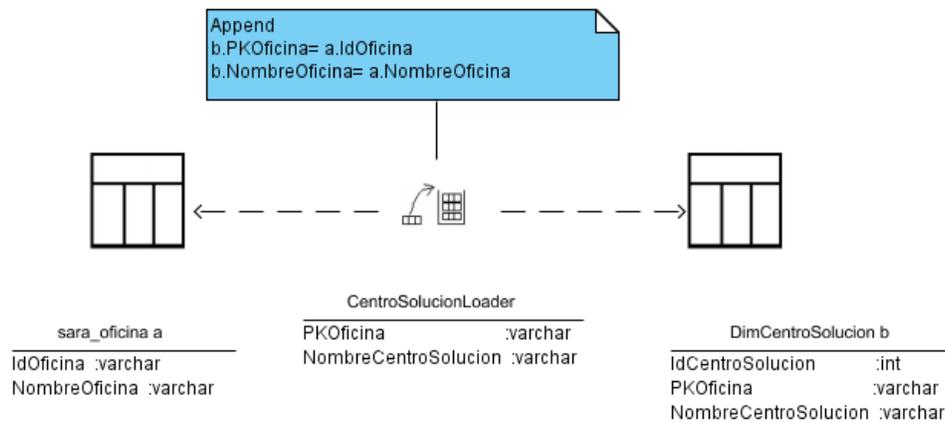


Figura 3.72: Representación de carga de datos en la tabla DimCentroSolucion utilizando mecanismo Loader

### 3.7.4.8. Diagrama de carga para la dimensión Proveedor

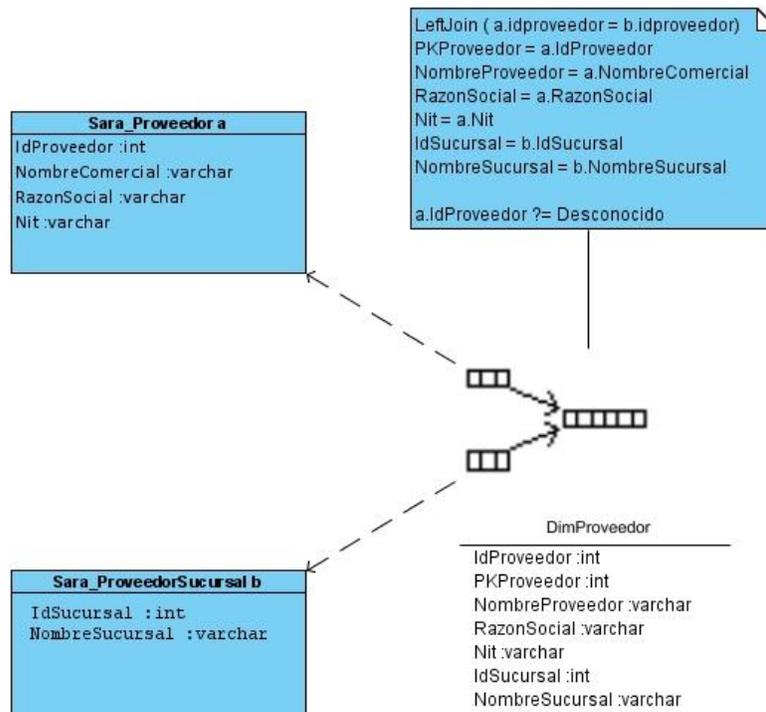


Figura 3.73: Representación de carga de datos en la tabla DimProveedor utilizando mecanismo Join

### 3.7.4.9. Diagrama de carga para la dimensión Ubicación

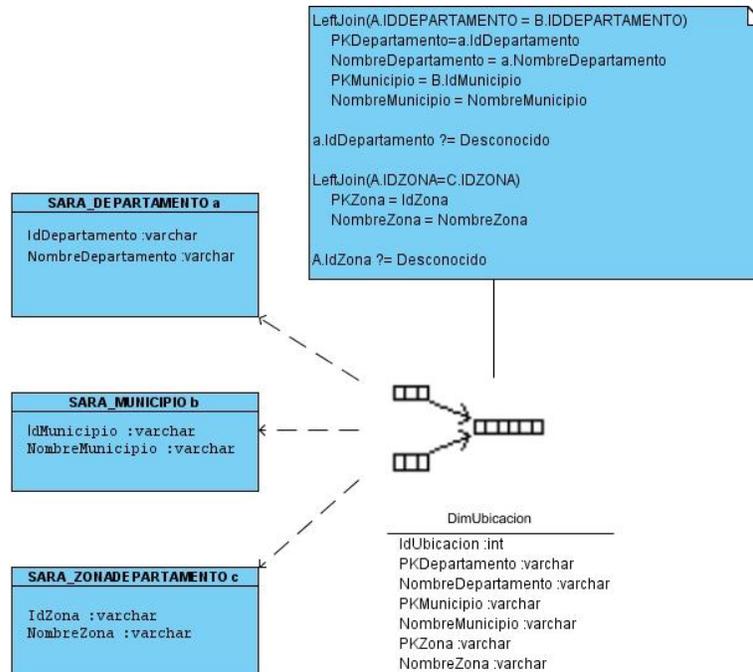


Figura 3.74: Representación de carga de datos en la tabla DimUbicacion utilizando mecanismo Join

### 3.7.4.10. Diagrama de carga para la dimensión Empleado

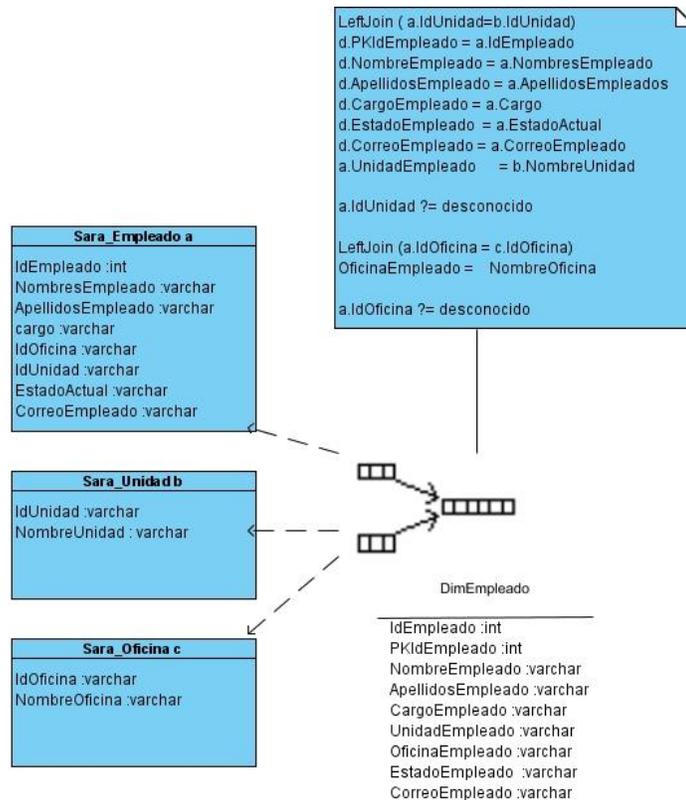


Figura 3.75: Representación de carga de datos en la tabla DimEmpleado utilizando mecanismo Join

### 3.7.4.11. Diagrama de carga para la dimensión Categoría

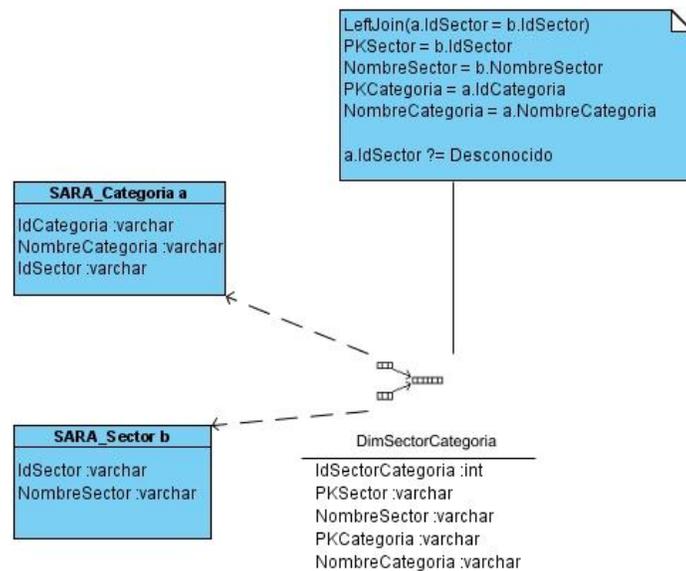


Figura 3.76: Representación de carga de datos en la tabla DimSectorCategoria utilizando mecanismo Join

### 3.7.4.12. Diagrama de carga para la FactCasoTS

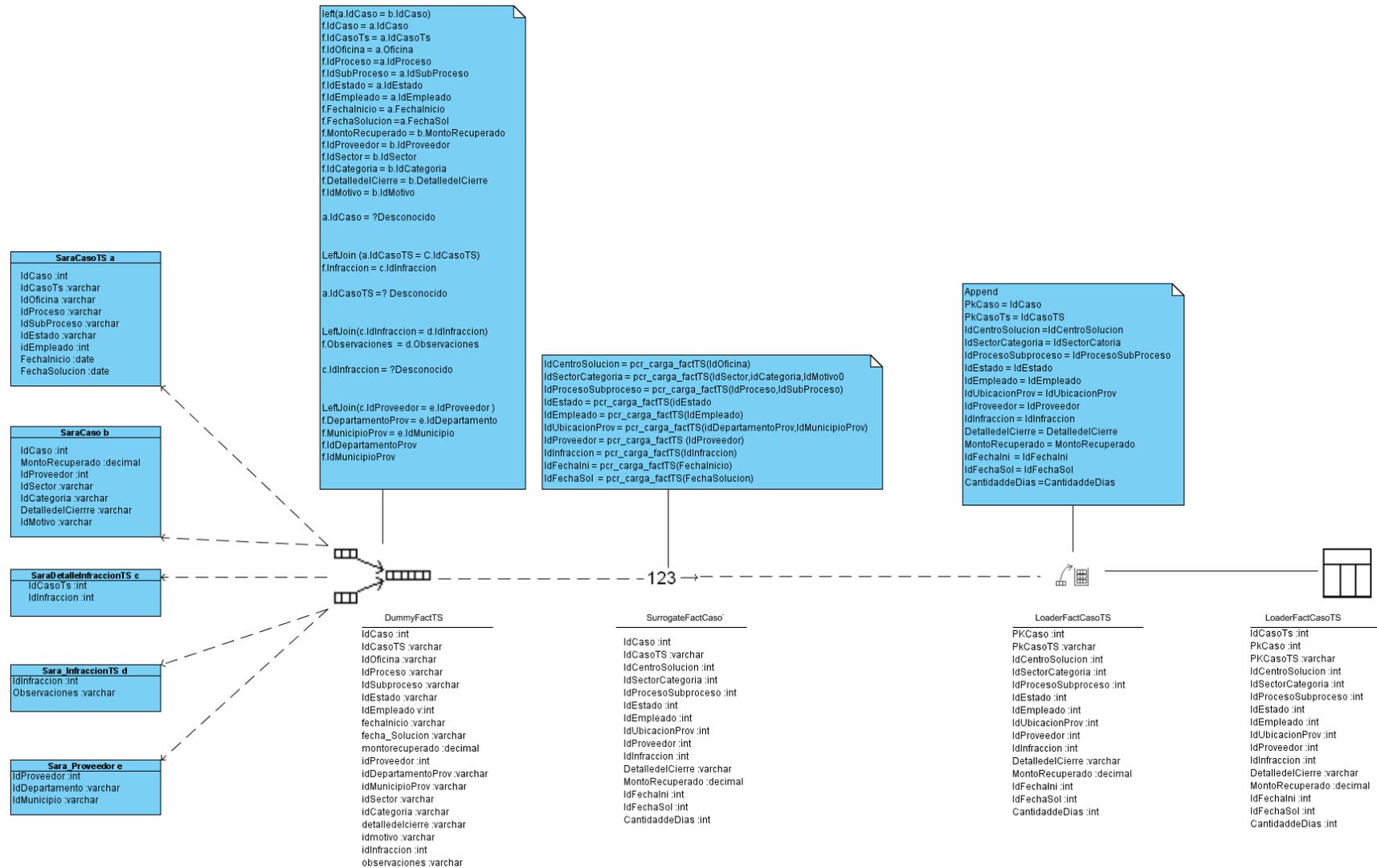


Figura 3.77: Representación de carga de datos en la tabla FacCasoTS utilizando mecanismo Join, Surrogate y Loader

### 3.8. ITERACIÓN 4: ANFITRIÓN

Se encarga del manejo de visitas de consumidores en los centros de atención.

#### 3.8.1. Requerimientos Funcionales

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.			Página 1 de y	
Autor : Grupo 14			Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo	
<b>01</b>	<b>Analizar Visitas por Motivo de la Visita</b>	Alta	FoO	
<b>01.01</b>	Permitir determinar la cantidad de visitas que se han recibido por cada uno de los motivos de visita disponibles.			
<b>01.01.01</b>	El análisis de la cantidad de visitas por motivo permite establecer las principales razones por las que los consumidores se acercan a los centros de atención.			
<b>02</b>	<b>Analizar las visitas por tipo de contacto.</b>	Alta	FoO	
<b>02.01</b>	Identificar el medio por el cual los consumidores se han informado de los servicios prestados por la DC.			
<b>02.02</b>	Combinar el tipo de contacto en rangos de tiempo para analizar el cambio de cantidad de visitas a través del mismo.			
<b>03</b>	<b>Análisis de Visitas por Técnico</b>	Alta	FoO	
<b>03.01</b>	Determinar el tiempo promedio de atención en minutos y horas por técnico en distintos periodos de tiempo.			
<b>03.02</b>	Mostrar la cantidad de visitas recibidas por técnico.			
<b>03.03</b>	Conocer el promedio de visitas por día que recibe cada técnico.			
<b>04</b>	<b>Análisis por Consumidor</b>	Alta	FoO	
<b>04.01</b>	Permitir realizar análisis de las visitas recibidas en los centros de atención por consumidor, utilizando rangos de fecha.			
<b>05</b>	<b>Reporte adhoc de la información analizada</b>	Baja	FoD	
<b>05.01</b>	Mostrar la información detallada de para poder constatar la información que es mostrada en los análisis.			
<b>05.02</b>	Mostrar la información dando al usuario la libertad de elegir la información que desee que incluya el usuario			

### 3.8.2. Especificación de Requerimientos

#### 3.8.2.1. Crear Reporte Analítico



Figura 3.78: Diagrama de caso de uso para reportes analíticos

<b>Caso de uso 1 Crear Reporte Analítico</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Técnico Web, Dirección de estudios, Gerencia informática.
Usuario: Elaborar un análisis detallado del proceso realizado en los centros de atención al consumidor para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo.
<b>Postcondiciones:</b> Reporte analítico publicado , almacenado o impreso

**Referencias Cruzadas****Requerimientos:** 01,02,03,04,05**Flujo Básico**

1. El Usuario selecciona “Nueva Vista de Análisis”.
2. El Sistema muestra esquemas y cubos disponibles.
3. El Usuario selecciona esquema y cubo para análisis.
4. El Sistema muestra dimensiones y medidas de análisis.
5. El Usuario selecciona medidas.
6. El Sistema muestra datos del modulo Anfitrión, con diferentes opciones de análisis.
7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A,B,C,D,E,F.
8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas
9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis
10. El sistema ordena la información por filas, columnas y filtros aplicados
11. El usuario almacena el reporte analítico generado para futuras consultas
12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.

**Flujo Alternativo****A. Realizar análisis utilizando MDX (Usuarios Expertos)**

1. El usuario selecciona la opción MDX.
2. El sistema muestra opción de generación de consulta MDX.
3. Usuario construye la consulta, utilizando consulta valida.
4. El sistema realiza la construcción del espacio de trabajo en base a la consulta generada.
5. El usuario realiza el análisis

**B. Configuración de tabla OLAP**

1. El usuario selecciona la opción Configurar tabla OLAP.
2. El sistema muestra ventana de configuración OLAP
3. El usuario selecciona las opciones de configuración
4. El sistema aplica configuraciones
5. El usuario realiza el análisis

**C. Seleccionar opciones de las vistas de los datos**

1. El usuario selecciona la opción Mostrar Padres.
2. El sistema muestra las referencias fuentes de los datos mostrados.
3. Sino el usuario selecciona la opción ocultar repeticiones.
4. El sistema muestra solo los datos únicos (unique) sin repeticiones.
5. Sino el usuario el usuario selecciona mostrar propiedades.
6. El sistema muestra propiedades de los datos.
7. Sino el usuario selecciona suprimir columnas vacías.
8. El sistema elimina de la consulta aquellos datos que contengan valores nulos.
9. Sino el usuario selecciona la opción intercambiar ejes.
10. El sistema genera un análisis roll-up.
11. El usuario analiza los datos.

**D. Generación y configuración de grafico**

1. El usuario selecciona la opción configuración de gráfico.
2. El sistema muestra propiedades de gráfico.
3. El usuario configura opciones de grafico, colores, leyendas, orientación y tipo de gráfico.
4. El sistema adecua el gráfico a los valores que se muestran en el espacio de trabajo.
5. El usuario selecciona opción de generar gráfico.
6. El sistema muestra gráfico.
7. El usuario analiza los datos.

**E. Seleccionar el Destino del análisis**

1. El usuario selecciona opción de configuración de impresión.
2. El sistema muestra panel de configuración de impresión.
3. El usuario selecciona opciones de configuración tales como tema de reporte, orientación, tamaño del papel.
4. El sistema configura opciones de reporte analítico.
5. El usuario publica el reporte en formato PDF, EXCEL, CSV o lo almacena en el sistema.
6. El sistema procesa el reporte y lo publica dependiendo el método seleccionado.

**F. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.8.2.2. Crear Reporte AD HOC

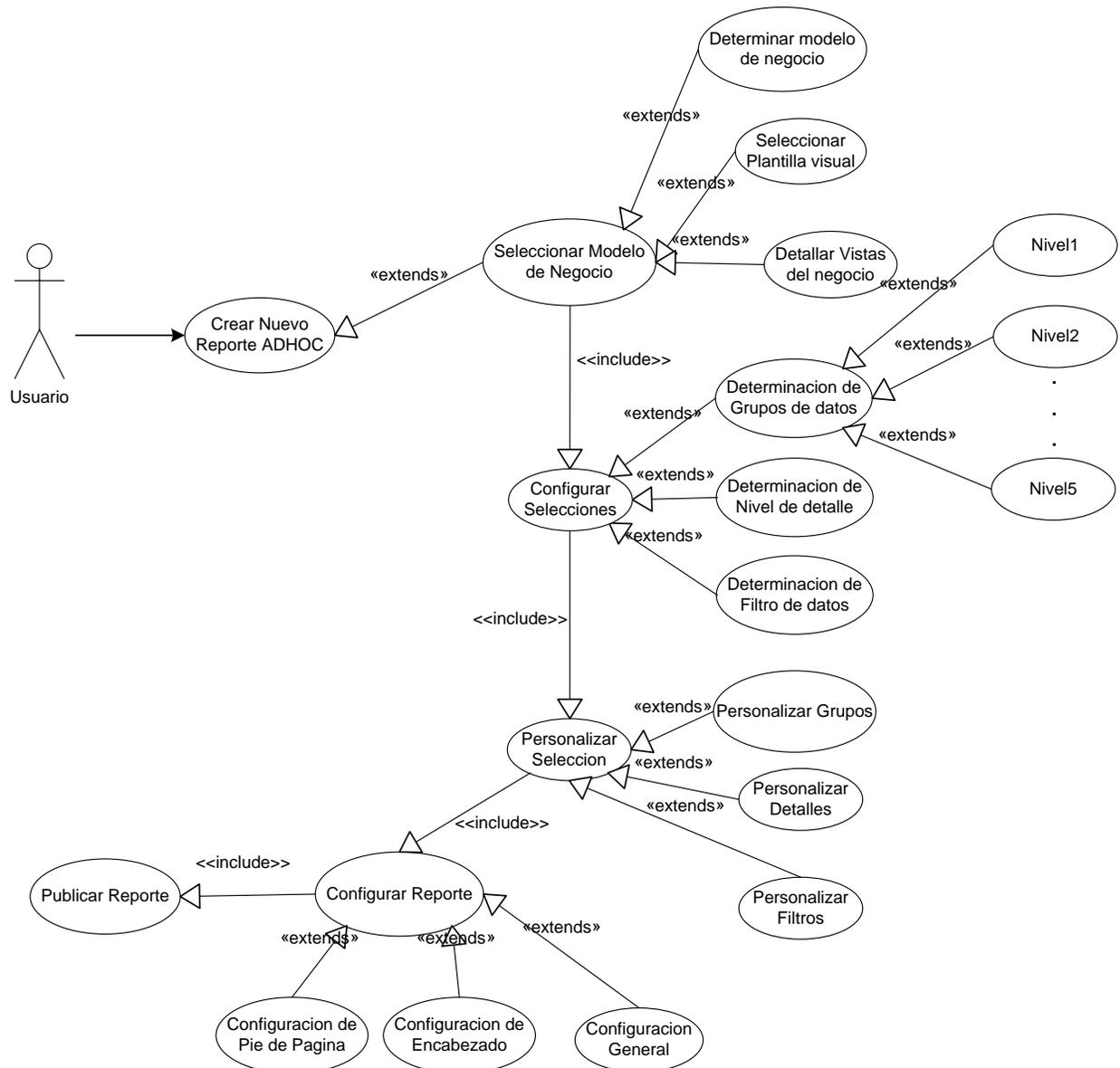


Figura 3.79: Diagrama de caso de uso para reportes AD HOC

<b>Caso de uso 2 Crear Reporte AD HOC</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Técnico Web, Dirección de Estudios, Gerencia informática.  Usuario: Elaborar de reportes temporales para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema poseer acceso a la vista de negocios de Anfitrión.
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> Requerimientos: 05

**Flujo Básico**

1. El Usuario selecciona "Nuevo Reporte".
2. El sistema muestra los modelos de negocio disponibles.
3. Usuario selecciona el modelo de negocio con el que desea trabajar.
4. Sistema muestra las vistas de negocio disponibles para el modelo seleccionado.
5. El usuario procede a la siguiente fase de selección de reporte
6. El sistema muestra variables con las que se puede agrupar aplicar filtros y seleccionar niveles de detalle de los grupos definidos
7. El usuario selecciona los grupos, filtros y niveles de detalle con las que desea realizar el reporte
8. El usuario procede a la siguiente fase de personalización de selecciones
9. El sistema muestra los valores de los niveles de detalle, grupos y filtros con los que se desea generar el reporte.
10. El sistema almacena la configuración de los datos.
11. El usuario procede a la siguiente fase de configuración de reporte
12. El sistema muestra panel de configuración de reporte.
13. El usuario selecciona las opciones generales del reporte, tipo de papel orientación encabezado y pie de página.
14. El sistema almacena la configuración
15. El usuario almacena el reporte ad hoc generado para futuras consultas
16. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.

**Flujo Alternativo****A. Cancelar la operación**

1. El usuario cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.8.2.3. Análisis Vistas Anfitrión

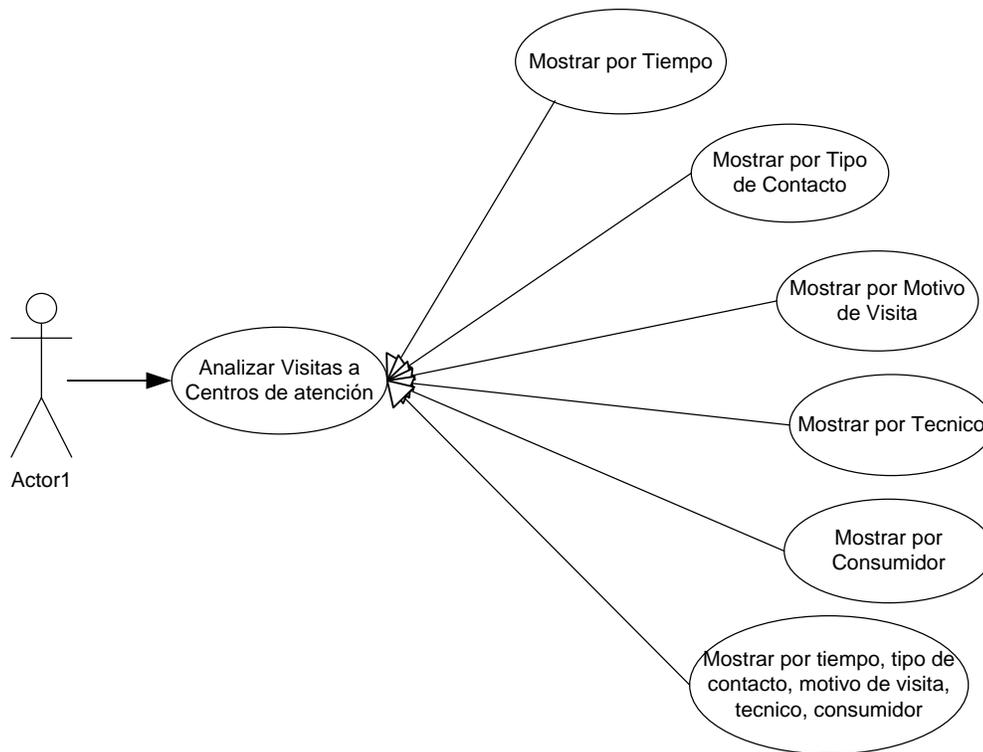


Figura 3.80: Diagrama de caso de uso para análisis de las visitas de Anfitrión

<b>Caso de uso 3 Análisis de visitas para Anfitrión.</b>
<b>Personal Involucrado e interés:</b> Usuario: Técnico, Técnico Web, Dirección de Estudios, Gerencia informática.  Usuario: Elaborar un análisis detallado del modulo Anfitrión para su posterior publicación o almacenamiento.
<b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo de anfitrión.
<b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 01,02,03,04
<b>Flujo Básico</b> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo anfitrión para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona las medidas.</li> <li>6. El Sistema muestra datos del modulo anfitrión, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis con medidas tales como: cantidad de visitas, promedio de atención en horas,</li> </ol>

minutos y segundos, minutos totales y promedio de visitas; combinando dimensiones como: tiempo, Tipo de Contacto, Técnico, Consumidor y Motivo de Visita.

10. El sistema ordena la información por filas, columnas y filtros aplicados
11. El usuario almacena el reporte analítico generado para futuras consultas
12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso

**Flujo Alternativo****A. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

### 3.8.3. Modelado Iteración 4: Anfitrión

#### 3.8.3.1. Esquema Conceptual del Data Warehouse

En Tabla 1.3 se encuentra la definición de la nomenclatura UML utilizada para la definición de los estereotipos del modelado conceptual.

##### Nivel 1. Definición del Modelo

Anfitrión:

**Hecho principal:** El Estudio de las visitas realizadas a los centros de atención.

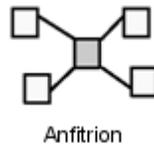


Figura 3.81: Definición modelo “Anfitrión”

##### Nivel 2 Definición del esquema estrella

**Paquete de Hecho:** Visitas

##### Paquetes Dimensionales

Técnico: Personas que atienden a los consumidores que visitan los centros de atención.

Consumidor: Datos relevantes del consumidor, nombres, sexo, ubicación

Contacto: Medio por el cual el consumidor se entero del servicio.

Motivo Visita: Razón por la cual el consumidor se acerca al centro de atención.

Tiempo: Control de tiempo en años, días, cuartiles, meses, semanas.

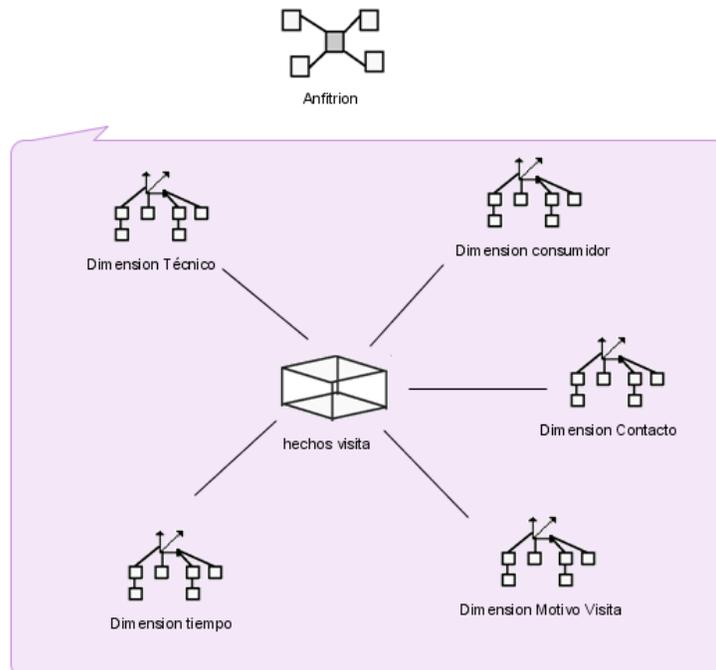


Figura 3.82: Definición del esquema estrella para “Anfitrión”

Nivel 3: Definición de Dimensiones/Hecho.

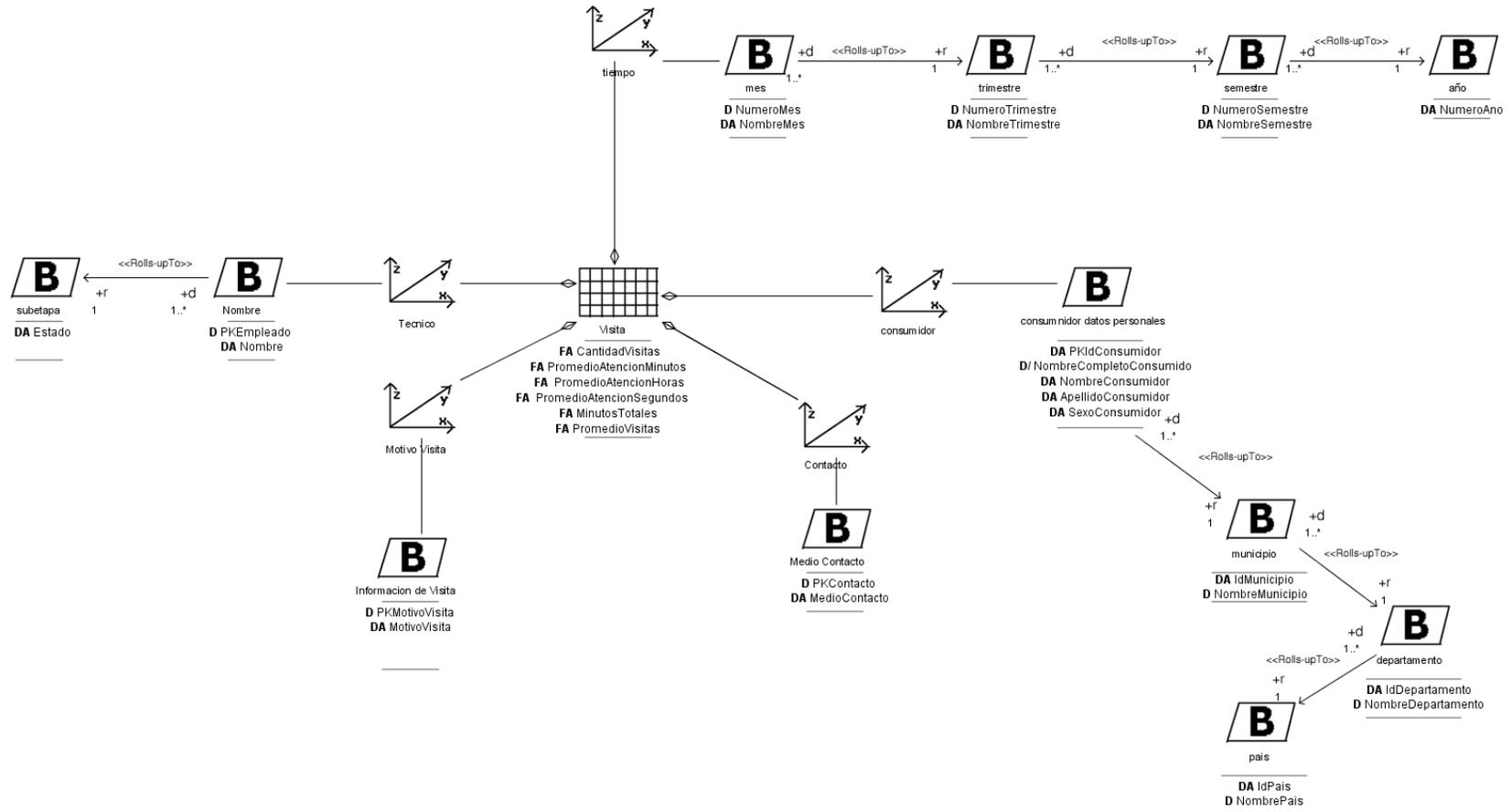


Figura 3.83: Dimensiones Candidatas para Roll-up y Drill-Down

3.8.3.2. Esquema lógico del Data Warehouse



Figura 3.84: Estructura lógica del esquema estrella para el módulo Anfitrión.

**3.8.3.3. Diccionario de datos de las dimensiones y tablas de hecho**

DIMTIEMPO		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDTIEMPO	int	Identificador de tiempo del DW
FECHA	datetime	Fecha de ingreso de la denuncia
NUMERODIASEMANA	int	Numero del día en la semana
NOMBREDIA	varchar(50)	Nombre del día de la semana
NUMERODIAMES	int	Numero del día en el mes
NUMERODIAANO	int	Numero del día en el año
NUMEROTRIMESTRE	int	Numero del trimestre del año
NOMBRETRIMESTRE	varchar(50)	Nombre del trimestre del año
NUMEROANO	int	Numero del año
NUMEROMES	int	Numero del mes
NOMBREMES	varchar(50)	Nombre del mes
NUMEROSEMESTRE	int	Numero del semestre
NOMBRESEMESTRE	varchar(50)	Nombre del semestre

DIMCONSUMIDOR		
ATRIBUTO	TIPO DE DATO	DESCRIPCION
IDCONSUMIDOR	int	Identificador del consumidor en el DW
NOMBRECONSUMIDOR	varchar(50)	Nombre del consumidor
APELLIDOCONSUMIDOR	varchar(50)	Apellidos del consumidor
SEXOCONSUMIDOR	char(10)	Sexo del consumidor F o M
IDPAIS	varchar(4)	Identificador del país del Consumidor
NOMBREPAIS	varchar(50)	Nombre del país del consumidor
IDDEPARTAMENTO	varchar(3)	Identificador del departamento del Consumidor
NOMBREDEPARTAMENTO	varchar(50)	Nombre del departamento
IDMUNICIPIO	varchar(4)	Identificador del municipio
NOMBREMUNICIPIO	varchar(50)	Nombre del municipio
PKIDCONSUMIDOR	int	Identificador del consumidor de la base de datos FUENTE

DIMCONTACTO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDCONTACTO	int	Identificador de la dimensión Contacto del DW.
PKCONTACTO	int	Identificador de los tipos de contacto en el sistema transaccional.
MEDIOCONTACTO	Varchar(50)	Descripción del medio de contacto.

DIMMOTIVOVISITA		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDMOTIVOVISITA	int	Identificador de la dimensión MotivoVisita del DW.
PKMOTIVOVISITA	int	Identificador del motivo de la visita en el sistema transaccional.
MOTIVOVISITA	Varchar(50)	Descripción del motivo de la visita.

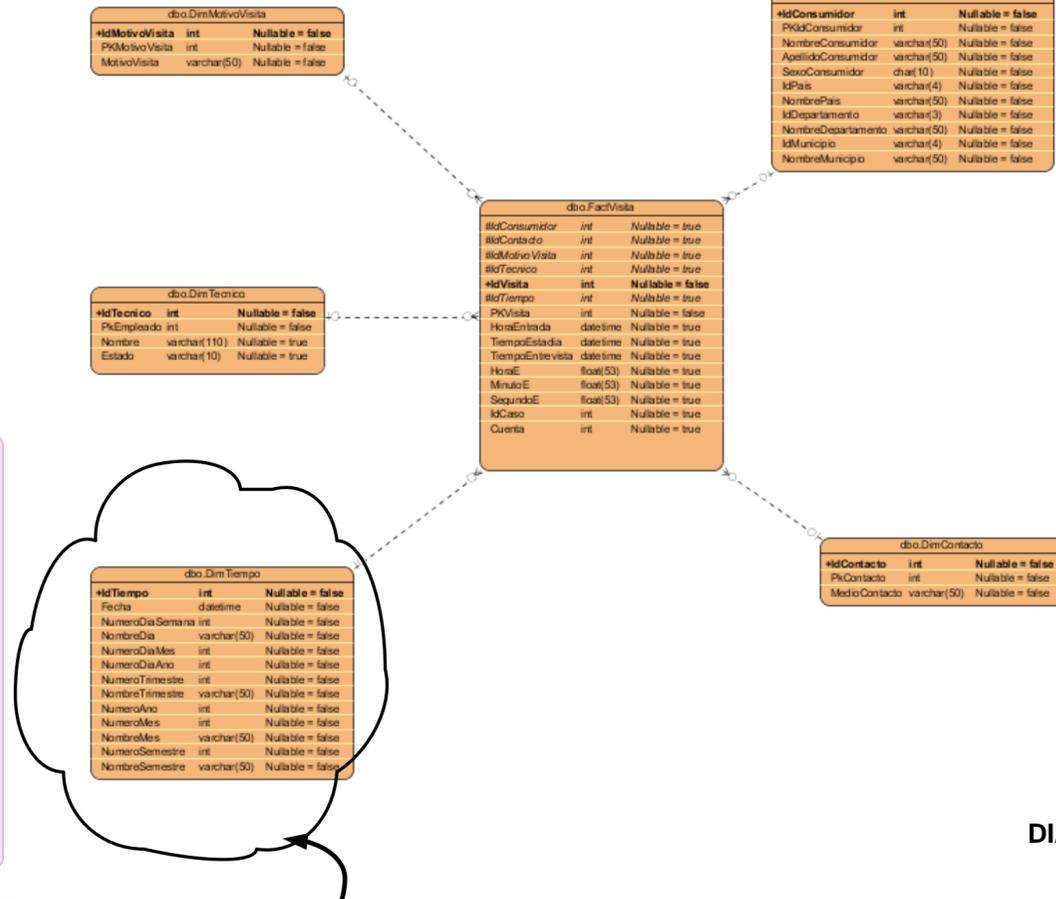
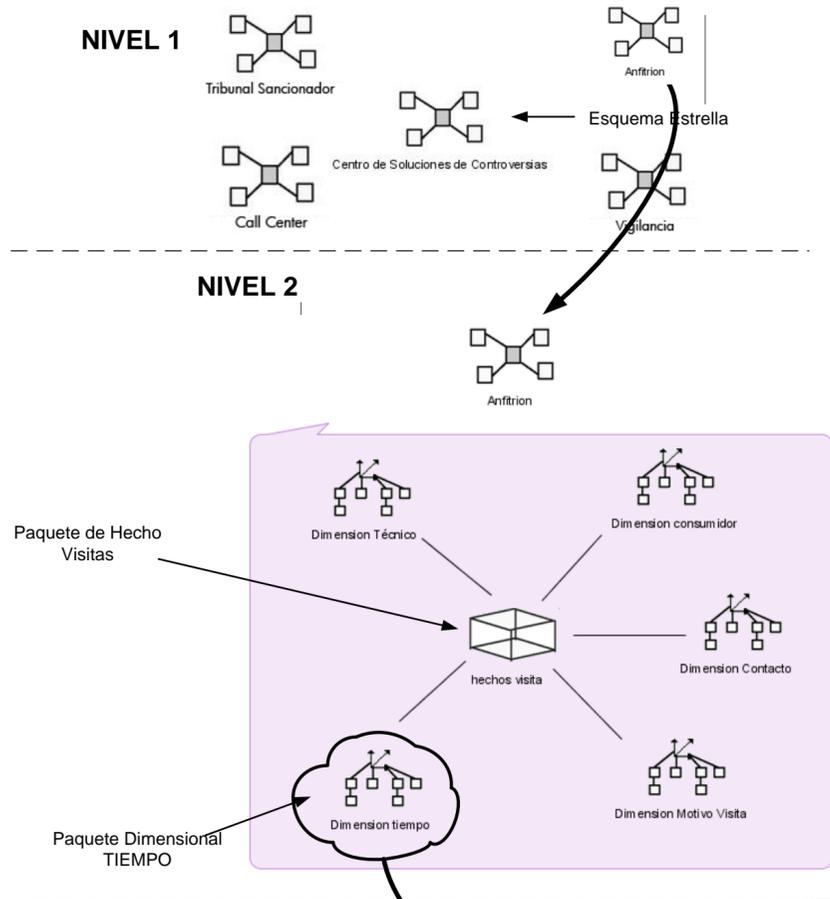
DIMTECNICO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDTECNICO	int	Identificador de la dimensión Técnico del DW.
PKEMPLEADO	int	Identificador del empleado en la base de datos origen.
NOMBRE	Varchar(110)	Nombre del empleado.
ESTADO	Varchar(10)	Estatus del empleado. (Activo, Inactivo)

FACTVISITA		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDVISITA	int	Identificador de la tabla de hechos.
PKVISITA	int	Identificador de la visita en el la base de datos origen.
IDCONSUMIDOR	int	Identificador del consumidor en el DW.
IDCONTACTO	int	Identificador del tipo de contacto en el DW.
IDMOTIVOVISITA	int	Identificador del Motivo de la Visita en el DW.
IDTECNICO	int	Identificador del Técnico en el DW.
IDTIEMPO	int	Identificador del Tiempo en el DW.
HORAENTRADA	datetime	Hora de entrada de la visita.
TIEMPOESTADIA	datetime	Tiempo de estadía de la visita.
TIEMPOENTREVISTA	datetime	Tiempo de la entrevista.
HORAE	float	Tiempo en horas de duración de la estadía.
MINUTOE	float	Tiempo en minutos de la duración de la estadía.
SEGUNDOE	float	Tiempo en segundos de la duración de la estadía.

3.8.4. Diagrama UML del DATA WAREHOUSE – ANFITRION

DIAGRAMA UML DEL DATA WAREHOUSE- ANFITRION  
ESQUEMA LÓGICO DEL DATA WAREHOUSE

ESQUEMA CONCEPTUAL DEL DATA WAREHOUSE



ESQUEMA FISICO DEL DATA WAREHOUSE

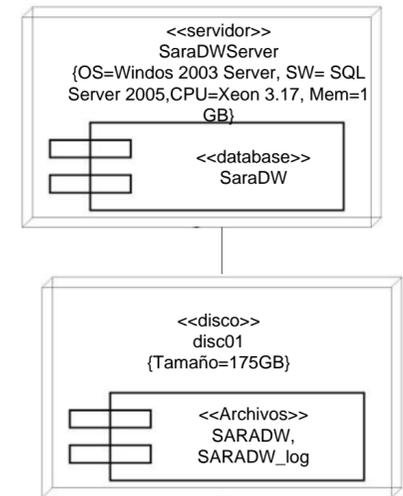
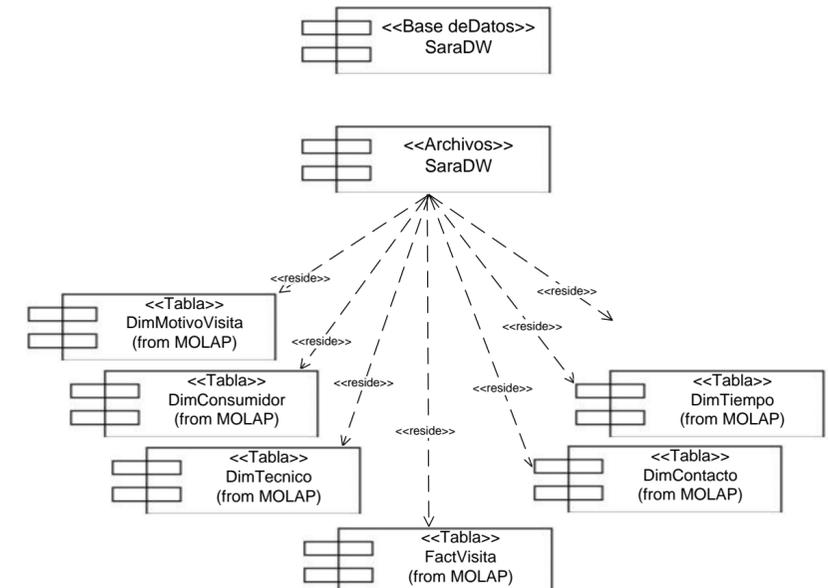
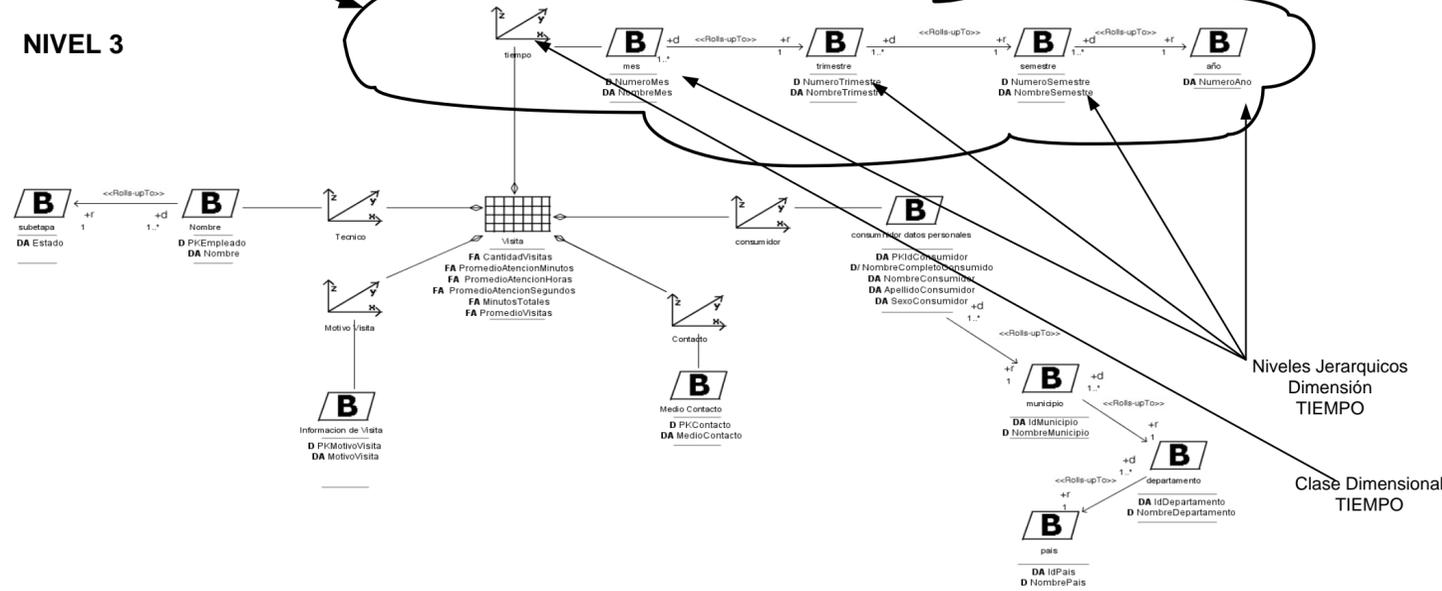


DIAGRAMA DE COMPONENTES



NIVEL 3



3.8.5. Diagrama del proceso ETL

DIAGRAMA ETL – ANFITRION

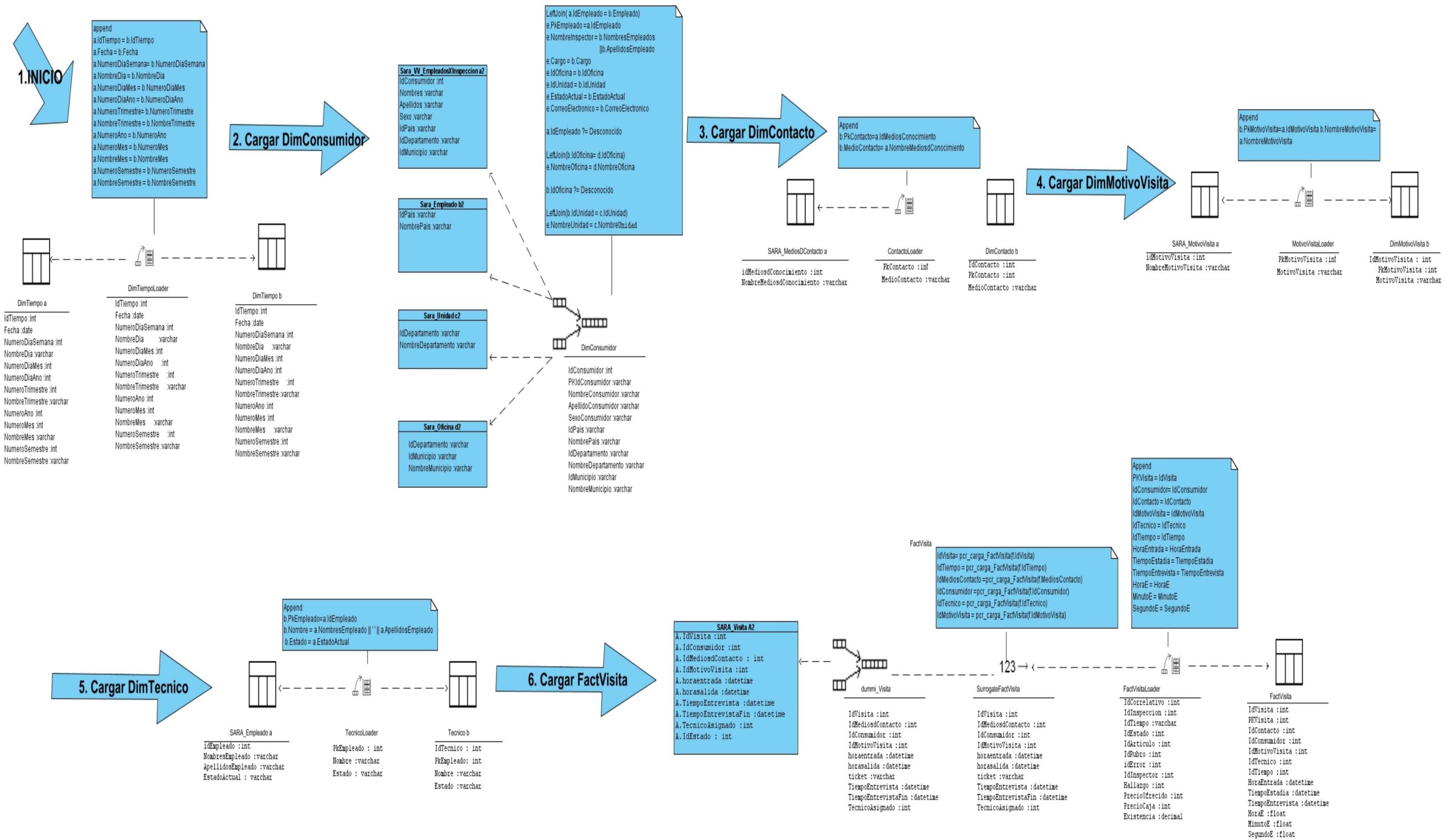
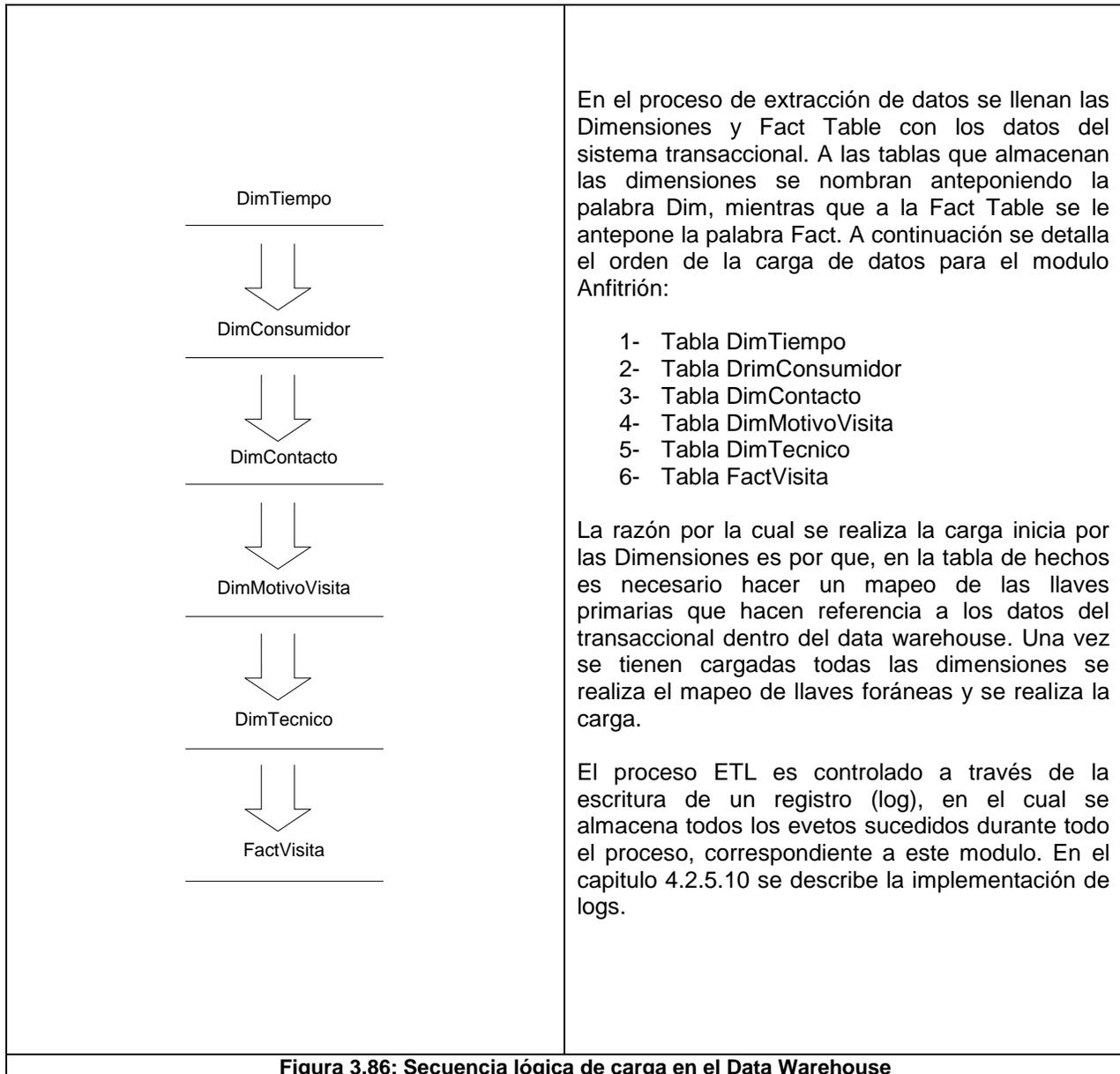


Figura 3.85: Diseño del proceso ETL para el Data Warehouse

**3.8.5.1. Secuencia de carga del proceso ETL**



3.8.5.2. Diagrama de carga para la dimensión Consumidor

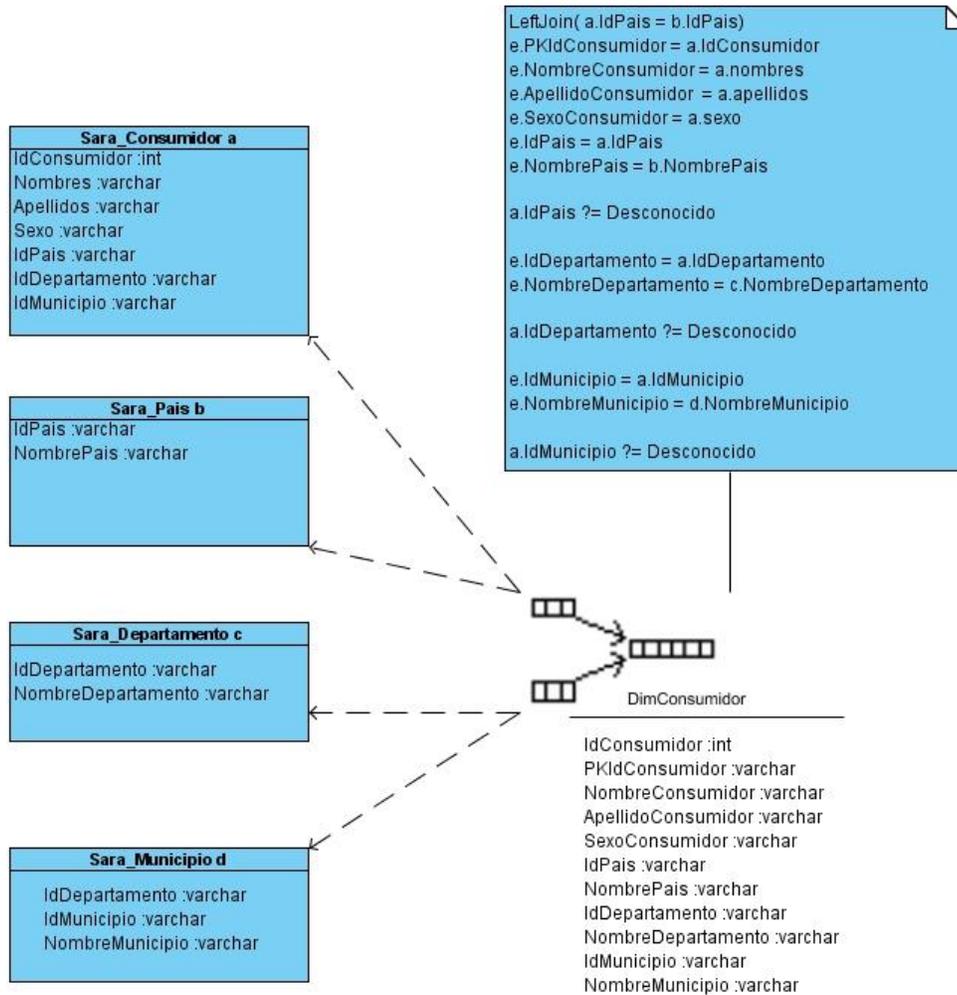


Figura 3.87: Representación de carga de datos en la tabla DimConsumidor utilizando mecanismo Join

3.8.5.3. Diagrama de carga para la dimensión tiempo

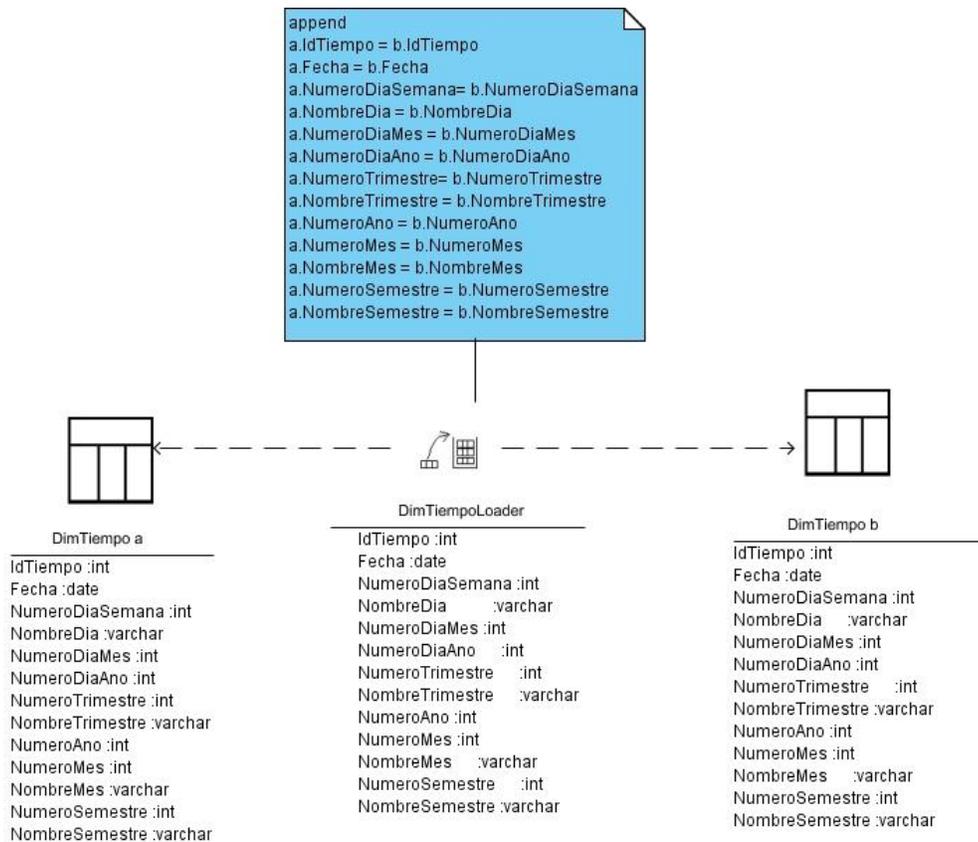


Figura 3.88: Representación de carga de datos en la tabla DimTiempo utilizando mecanismo Loader

3.8.5.4. Diagrama de carga para la dimensión Contacto

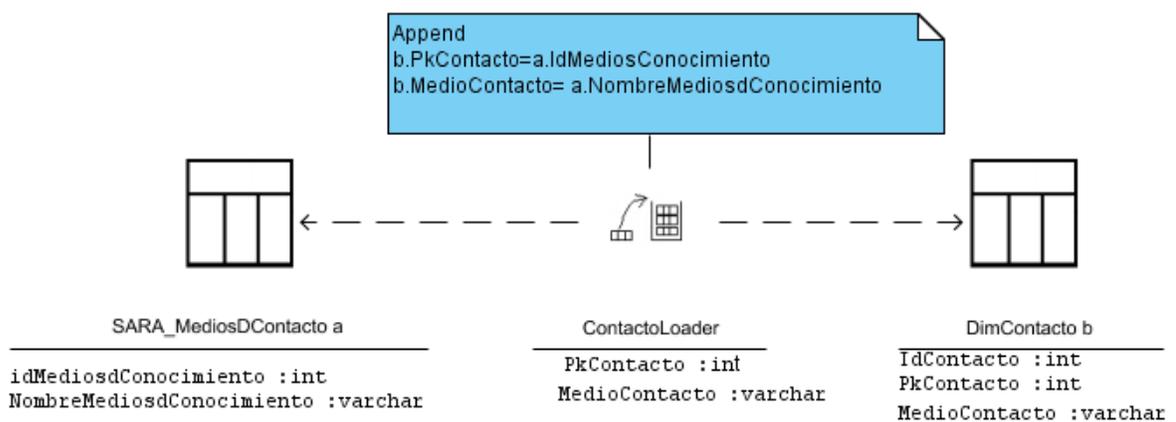


Figura 3.89: Representación de carga de datos en DimContacto utilizando mecanismo Loader

3.8.5.5. Diagrama de carga para la dimensión MotivoVisita

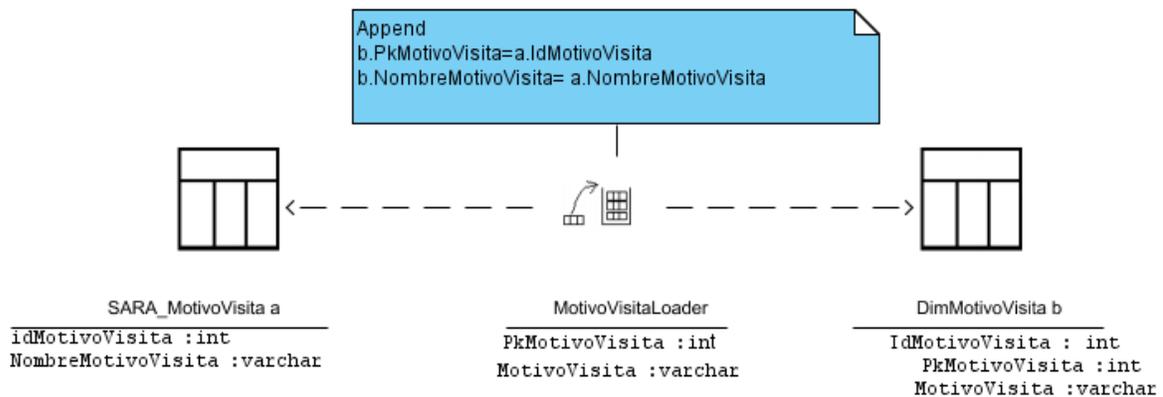


Figura 3.90: Representación de carga de datos en DimMotivoVisita utilizando mecanismo Loader

3.8.5.6. Diagrama de carga para la dimensión Técnico

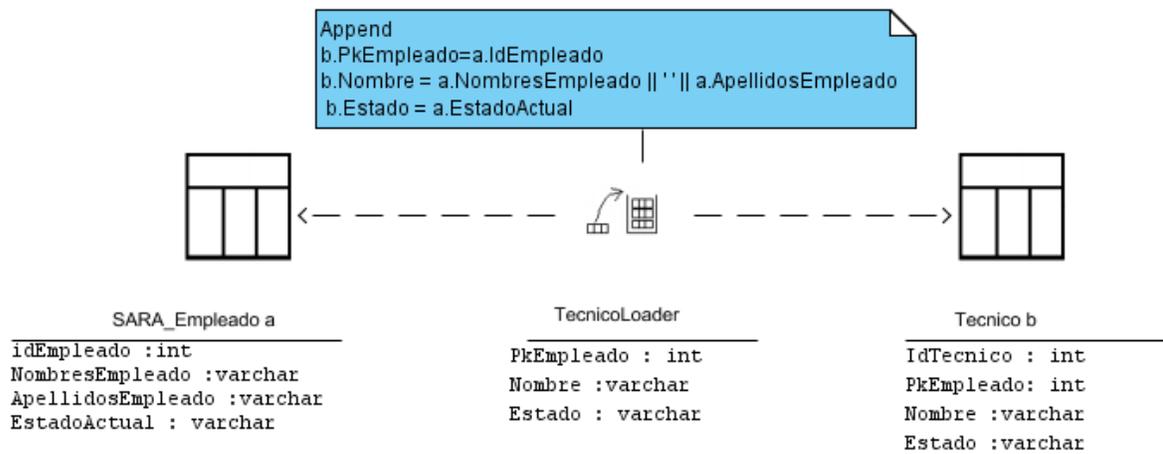


Figura 3.91: Representación de carga de datos en DimTecnico por medio de un mecanismo Loader

3.8.5.7. Diagrama de carga para la fact table Visita

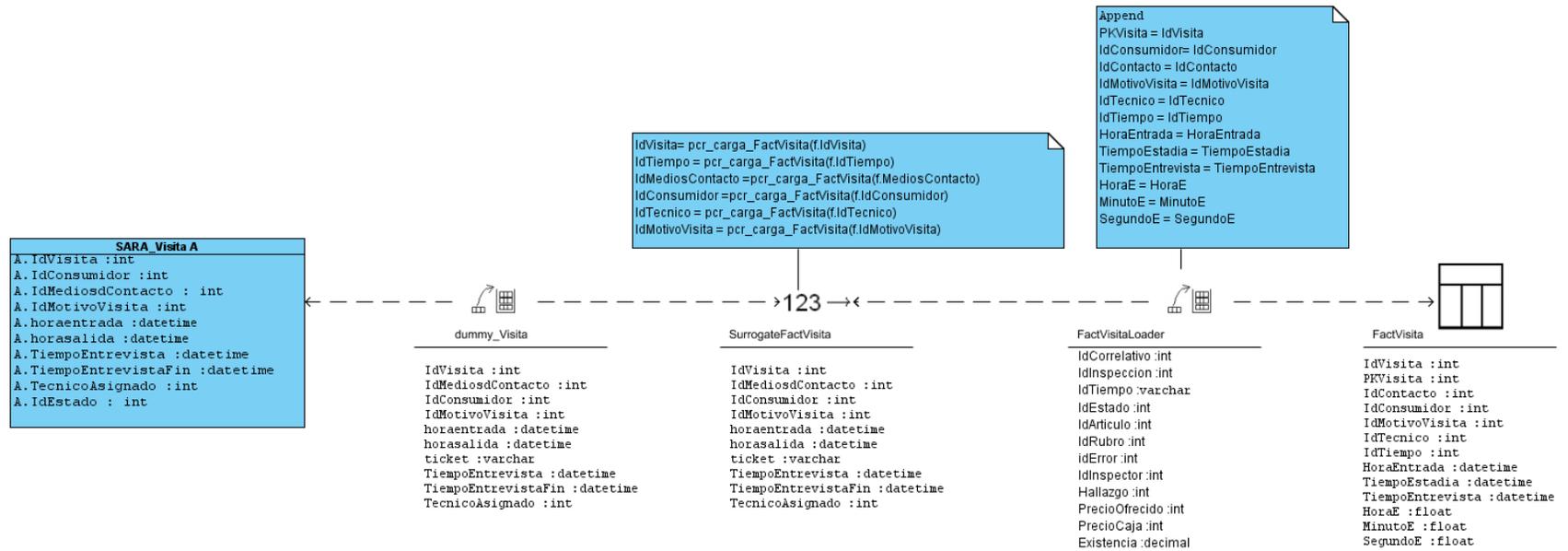


Figura 3.92: Representación de carga de datos en FactVisita utilizando mecanismo Loader y Surrogate

### 3.9. ITERACION 5: CALL CENTER

Se encarga de recibir diferentes tipos de llamadas por parte de consumidores.

#### 3.9.1. Requerimientos Funcionales Para Call Center

DESARROLLO DE UN “DATA WAREHOUSE” PARA EL PROCESO DE DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR.		Página 1 de y	
Autor : Grupo 14		Fecha: 29/04/10	
Req.	Especificación	prioridad	Tipo
<b>01</b>	<b>Analizar Llamadas Recibidas en Call Center</b>	<b>Alta</b>	<b>FoO</b>
<b>01.01</b>	Analizar la información de las llamadas recibidas por tipo de atención brindada por el técnico de call center.		
<b>01.01.01</b>	El análisis se realizara para las siguientes tipos de atención: <ul style="list-style-type: none"> <li>○ Asesorías</li> <li>○ Consulta caso</li> <li>○ Denuncias no personales</li> <li>○ Descontento sector Transporte</li> <li>○ Gestiones</li> <li>○ Información</li> <li>○ Llamadas Inútiles</li> <li>○ Orientaciones de consumo</li> <li>○ Quejas consumidor</li> </ul>		
<b>01.01.02</b>	A cada llamada recibida en call center se tipifica con un tipo de atención por lo que es necesario determinar la cantidad de llamadas recibidas por tipo de atención.		
<b>01.02</b>	El análisis de las llamadas se podrá realizar por tema de consulta para cada tipo de consulta que exista.		
<b>01.02.01</b>	Cada tipo de consulta tiene asociados varios temas de consulta para efecto de clasificar de una mejor manera las llamadas del consumidor, determinar la cantidad de llamadas. Recibidas por tema de consulta.		
<b>02</b>	<b>Analizar los montos recuperados y reclamados</b>		
<b>02.01</b>	Permitir el análisis de monto recuperado y reclamado en las llamadas de tipo Gestión.		
<b>02.02</b>	Analizar la información de los montos por sector, consumidor y fecha en la que se realiza la llamada.		
<b>03</b>	<b>Analizar las llamadas por Fecha de Ingreso</b>	<b>Alta</b>	<b>FoO</b>
<b>03.01</b>	Realizar análisis de las llamadas recibidas por año, semestre, cuarto y mes en que la llamada fue recibida en call center.		
<b>04</b>	<b>Permitir en análisis de llamadas por forma de recepción.</b>	<b>Media</b>	<b>FoO</b>

<b>04.01</b>	Los casos y denuncias llegan a los técnicos de call center a través de distintos medios, para los cuales es necesario analizar la cantidad de casos recibidos para los tipos de llamada especificados en el requerimiento 01.01.01		
<b>05</b>	<b>Realizar análisis de los casos recibidos en call center por sector y categoría.</b>	<b>Alta</b>	<b>FoO</b>
<b>05.01</b>	Realizar análisis por sector y categoría para el tipo de llamada Asesoría.		
<b>05.02</b>	Realizar análisis por sector y categoría para el tipo de llamada Denuncia no personal.		
<b>05.03</b>	Realizar análisis por sector y categoría para el tipo de llamada de Gestión.		
<b>05.04</b>	Realizar análisis por sector y categoría para el tipo de llamada de Información		
<b>05.05</b>	Realizar análisis por sector y categoría para el tipo de llamada Orientación de consumo.		
<b>06</b>	<b>Mostrar los casos por forma de archivo, es decir, si el caso es favorable o desfavorable al consumidor.</b>	<b>Media</b>	<b>FoO</b>
<b>06.01</b>	Realizar análisis por forma de archivo para el tipo de llamada Gestión.		
<b>07</b>	<b>Mostrar la información por Proveedor al que se le demanda el caso.</b>	<b>Alta</b>	<b>FoO</b>
<b>07.01</b>	Realiza análisis por Proveedor al que se le acusa o consulta por llamada.		
<b>08</b>	<b>Realizar análisis de información de la llamada recibida por ubicación geográfica.</b>	<b>Alta</b>	<b>FoO</b>
<b>08.01</b>	Realizar análisis de información por ubicación geográfica para los distintos tipos de llamadas.		
<b>09</b>	<b>Reporte adhoc de la información analizada</b>	<b>Baja</b>	<b>Fo</b>
<b>09.01</b>	Mostrar la información detallada de para poder constatar la información que es mostrada en los análisis.		
<b>09.02</b>	Mostrar la información dando al usuario la libertad de elegir la información que desee que incluya el usuario		

### 3.9.2. Especificación de Requerimientos

#### 3.9.2.1. Crear Reporte Analítico



Figura 3.93: Diagrama de caso de uso para reportes analíticos

Caso de uso 1 Crear Reporte Analítico
<p><b>Personal Involucrado e interés:</b></p> <p>Usuario: Técnico, Técnico Web, Dirección de estudios, Dirección de Call Center, Gerencia informática.</p> <p>Usuario: Elaborar un análisis detallado del proceso realizado en la dirección de Call Center para su posterior publicación o almacenamiento.</p>

<b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo.
<b>Postcondiciones:</b> Reporte analítico publicado , almacenado o impreso
<b>Referencias Cruzadas</b> <b>Requerimientos:</b> 01, 02, 03, 04, 05, 06, 07, 08
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona medidas.</li> <li>6. El Sistema muestra datos de Call Center, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A,B,C,D,E,F.</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis</li> <li>10. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>11. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Realizar análisis utilizando MDX (Usuarios Expertos)</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción MDX.</li> <li>2. El sistema muestra opción de generación de consulta MDX.</li> <li>3. Usuario construye la consulta, utilizando consulta valida.</li> <li>4. El sistema realiza la construcción del espacio de trabajo en base a la consulta generada.</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>B. Configuración de tabla OLAP</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Configurar tabla OLAP.</li> <li>2. El sistema muestra ventana de configuración OLAP</li> <li>3. El usuario selecciona las opciones de configuración</li> <li>4. El sistema aplica configuraciones</li> <li>5. El usuario realiza el análisis</li> </ol> <p><b>C. Seleccionar opciones de las vistas de los datos</b></p> <ol style="list-style-type: none"> <li>1. El usuario selecciona la opción Mostrar Padres.</li> <li>2. El sistema muestra las referencias fuentes de los datos mostrados.</li> <li>3. Sino el usuario selecciona la opción ocultar repeticiones.</li> <li>4. El sistema muestra solo los datos únicos (unique) sin repeticiones.</li> <li>5. Sino el usuario el usuario selecciona mostrar propiedades.</li> <li>6. El sistema muestra propiedades de los datos.</li> <li>7. Sino el usuario selecciona suprimir columnas vacías.</li> <li>8. El sistema elimina de la consulta aquellos datos que contengan valores nulos.</li> <li>9. Sino el usuario selecciona la opción intercambiar ejes.</li> <li>10. El sistema genera un análisis roll-up.</li> <li>11. El usuario analiza los datos.</li> </ol>

**D. Generación y configuración de gráfico**

1. El usuario selecciona la opción configuración de gráfico.
2. El sistema muestra propiedades de gráfico.
3. El usuario configura opciones de gráfico, colores, leyendas, orientación y tipo de gráfico.
4. El sistema adecua el gráfico a los valores que se muestran en el espacio de trabajo.
5. El usuario selecciona opción de generar gráfico.
6. El sistema muestra gráfico.
7. El usuario analiza los datos.

**E. Seleccionar el Destino del análisis**

1. El usuario selecciona opción de configuración de impresión.
2. El sistema muestra panel de configuración de impresión.
3. El usuario selecciona opciones de configuración tales como tema de reporte, orientación, tamaño del papel.
4. El sistema configura opciones de reporte analítico.
5. El usuario publica el reporte en formato PDF, EXCEL, CSV o lo almacena en el sistema.
6. El sistema procesa el reporte y lo publica dependiendo el método seleccionado.

**F. Cancelar la operación**

1. El usuario Cierra el sistema
2. El sistema cancela las operaciones y finaliza sesión de usuario.

3.9.2.2. Crear Reporte AD HOC

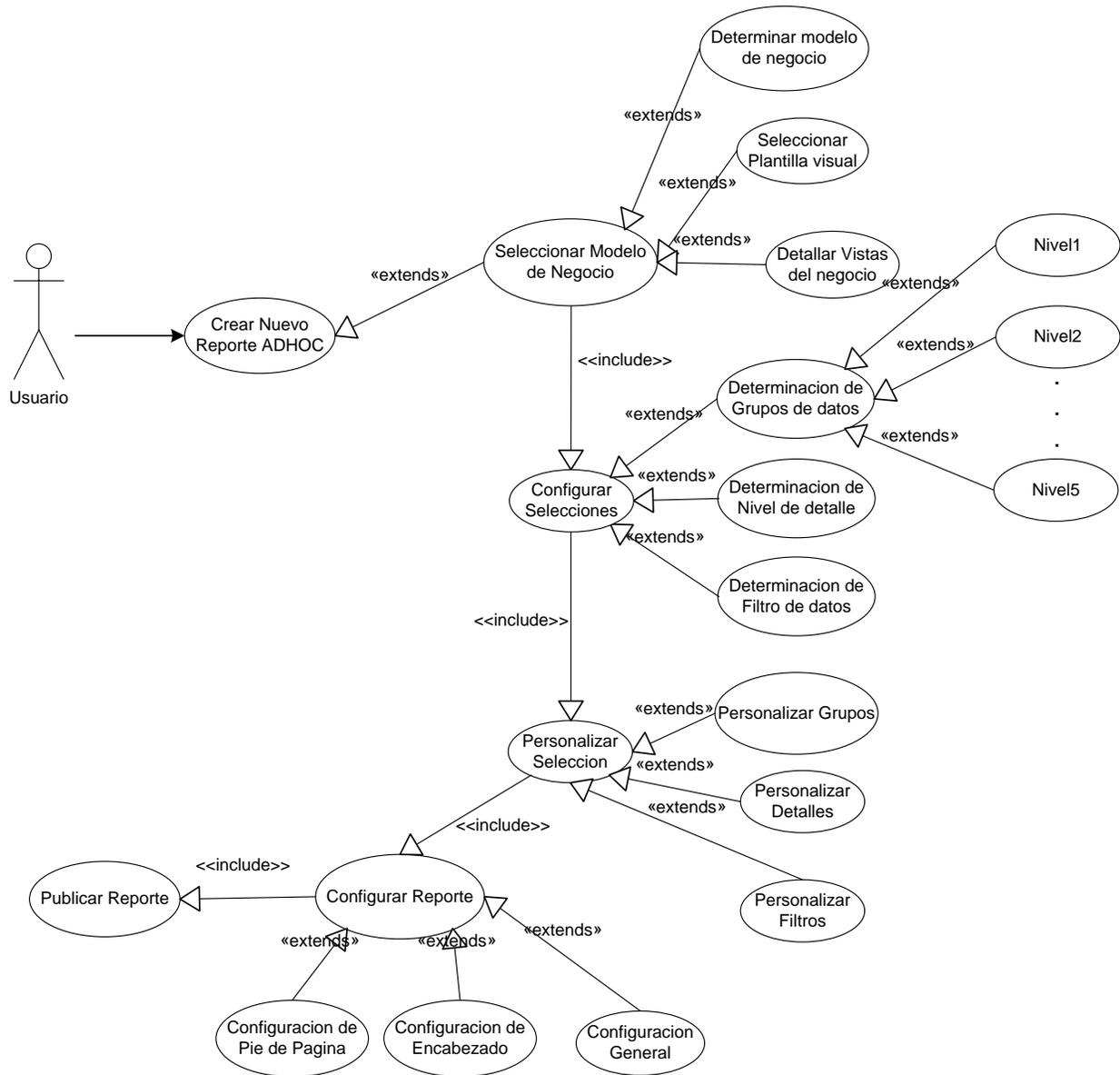


Figura 3.94: Diagrama de caso de uso para reportes AD HOC

<p><b>Caso de uso 2 Crear Reporte AD HOC</b></p> <p><b>Personal Involucrado e interés:</b>                  Usuario: Técnico, Técnico Web, Dirección de Estudios, Dirección de Call Center, Gerencia informática.</p> <p>Usuario: Elaborar de reportes temporales para su posterior publicación o almacenamiento.</p> <p><b>Precondiciones:</b> Haber iniciado sesión en el sistema poseer acceso a la vista de negocios de Verificación y vigilancia</p>
---

<p><b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.</p>
<p><b>Referencias Cruzadas</b>  <b>Requerimientos:</b> 09</p>
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nuevo Reporte”.</li> <li>2. El sistema muestra los modelos de negocio disponibles.</li> <li>3. Usuario selecciona el modelo de negocio con el que desea trabajar.</li> <li>4. Sistema muestra las vistas de negocio disponibles para el modelo seleccionado.</li> <li>5. El usuario procede a la siguiente fase de selección de reporte</li> <li>6. El sistema muestra variables con las que se puede agrupar aplicar filtros y seleccionar niveles de detalle de los grupos definidos</li> <li>7. El usuario selecciona los grupos, filtros y niveles de detalle con las que desea realizar el reporte</li> <li>8. El usuario procede a la siguiente fase de personalización de selecciones</li> <li>9. El sistema muestra los valores de los niveles de detalle, grupos y filtros con los que se desea generar el reporte.</li> <li>10. El sistema almacena la configuración de los datos.</li> <li>11. El usuario procede a la siguiente fase de configuración de reporte</li> <li>12. El sistema muestra panel de configuración de reporte.</li> <li>13. El usuario selecciona las opciones generales del reporte, tipo de papel orientación encabezado y pie de página.</li> <li>14. El sistema almacena la configuración</li> <li>15. El usuario almacena el reporte ad hoc generado para futuras consultas</li> <li>16. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso.</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Cancelar la operación</b></p> <ol style="list-style-type: none"> <li>1. El usuario Cierra el sistema</li> <li>2. El sistema cancela las operaciones y finaliza sesión de usuario.</li> </ol>

3.9.2.3. Análisis de Inspecciones para Call Center

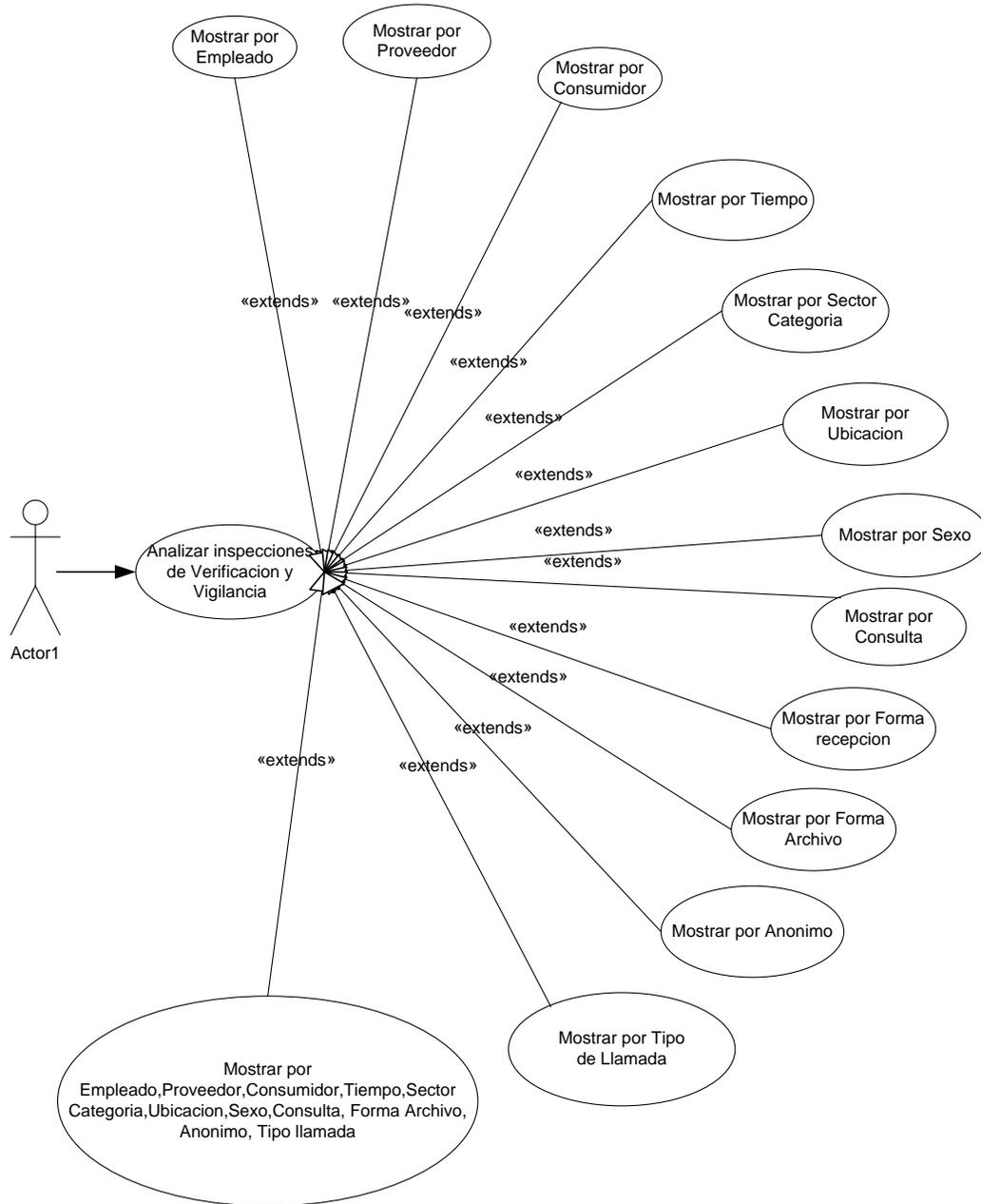


Figura 3.95: Diagrama de caso de uso para análisis de los en el modulo de Call center

Caso de uso 3 Análisis de Llamadas para Call Center
<p><b>Personal Involucrado e interés:</b></p> <p>Usuario: Técnico, Técnico Web, Dirección de Estudios, Dirección de Call Center, Gerencia informática.</p> <p>Usuario: Elaborar un análisis detallado de las llamadas para su posterior publicación o almacenamiento.</p>
<p><b>Precondiciones:</b> Haber iniciado sesión en el sistema, tener permisos de acceso al cubo de Call Center.</p>

<p><b>Postcondiciones:</b> Reporte analítico publicado y/o almacenado.</p>
<p><b>Referencias Cruzadas</b>  <b>Requerimientos:</b> 01,02,03,04,05,06</p>
<p><b>Flujo Básico</b></p> <ol style="list-style-type: none"> <li>1. El Usuario selecciona “Nueva Vista de Análisis”.</li> <li>2. El Sistema muestra esquemas y cubos disponibles.</li> <li>3. El Usuario selecciona esquema y cubo de Call Center para análisis.</li> <li>4. El Sistema muestra dimensiones y medidas de análisis.</li> <li>5. El Usuario selecciona las medidas.</li> <li>6. El Sistema muestra datos de Call Center, con diferentes opciones de análisis.</li> <li>7. El usuario selecciona Vista utilizando navegador OLAP o casos alternos A</li> <li>8. El sistema muestra variables con las que se puede aplicar filtros o seleccionar orden en filas y columnas</li> <li>9. El usuario selecciona los filtros y orden en filas y columnas con las que desea realizar el análisis con medidas tales como: cantidad de llamadas, monto recuperado y sus respectivos promedios; combinando dimensiones como: Empleado, Proveedor, Consumidor, Tiempo, Sector Categoría, Ubicación, Sexo, Consulta, Forma Archivo, Anónimo y Tipo llamada.</li> <li>10. El sistema ordena la información por filas, columnas y filtros aplicados</li> <li>11. El usuario almacena el reporte analítico generado para futuras consultas</li> <li>12. El sistema almacena el reporte y lo publica para que se encuentre disponible a los usuarios que se les permite el acceso</li> </ol>
<p><b>Flujo Alternativo</b></p> <p><b>A. Cancelar la operación</b></p> <ol style="list-style-type: none"> <li>1. El usuario Cierra el sistema</li> <li>2. El sistema cancela las operaciones y finaliza sesión de usuario.</li> </ol>

### 3.9.3. Modelado Iteración 5 Call Center

#### 3.9.3.1. Esquema Conceptual del Data Warehouse

En tabla 1.3 se encuentra la definición de la nomenclatura UML utilizada para la definición de los estereotipos del modelado conceptual.

##### Nivel1. Definición del Modelo

Call Center:

Especialización del Negocio para análisis de llamadas.

**Hecho Principal:** El Estudio de la Llamada

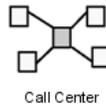


Figura 3.96: Definición modelo “Call Center”

##### Nivel 2. Definición del esquema estrella

Paquete de hecho: Llamada

Paquetes Dimensionales

Tiempo: Control de tiempo en años, días, cuartiles, meses, semanas.

Consumidor: Datos relevantes del consumidor, nombres, sexo, ubicación.

Consulta: Tipo y tema de consulta para llamadas de asesoría, consulta e información.

Sexo: Descripción sexo del empleado.

Forma Recepción: Clasificación de la forma de recepción.

Sector Categoría: Se refiere al tipo de denuncia por el cual el consumidor solicita ayuda.

Empleado: Catalogo de empleados de la DC.

Anónimo: Indica si la llamada fue anónima.

Tipo Llamada: Tipo de llamada registrada.

Forma Archivo: Denota si la gestión termina favorable o desfavorable.

Proveedor: Generalmente considerado como el ente al cual se le demanda algo.

Ubicación: Contiene la ubicación geográfica del país.

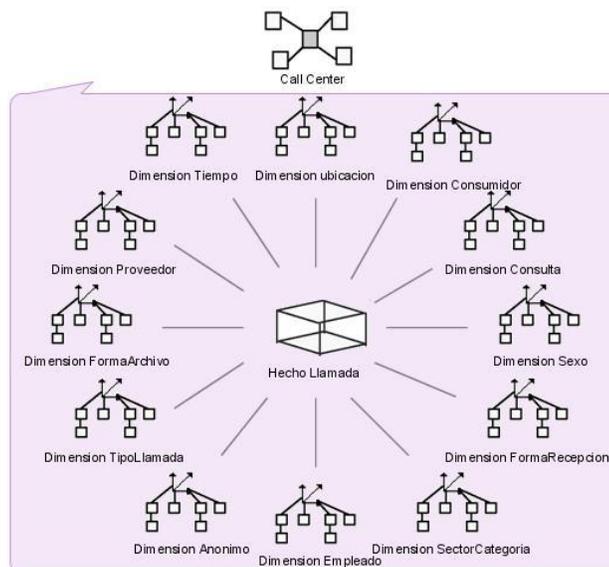


Figura 3.97: Definición del esquema estrella para “Call Center”

**Nivel 3: Definición de Dimensiones/Hechos**

Especificación de atributos

Dimensiones candidatas a roll-up y drill-down tiempo y ubicación.

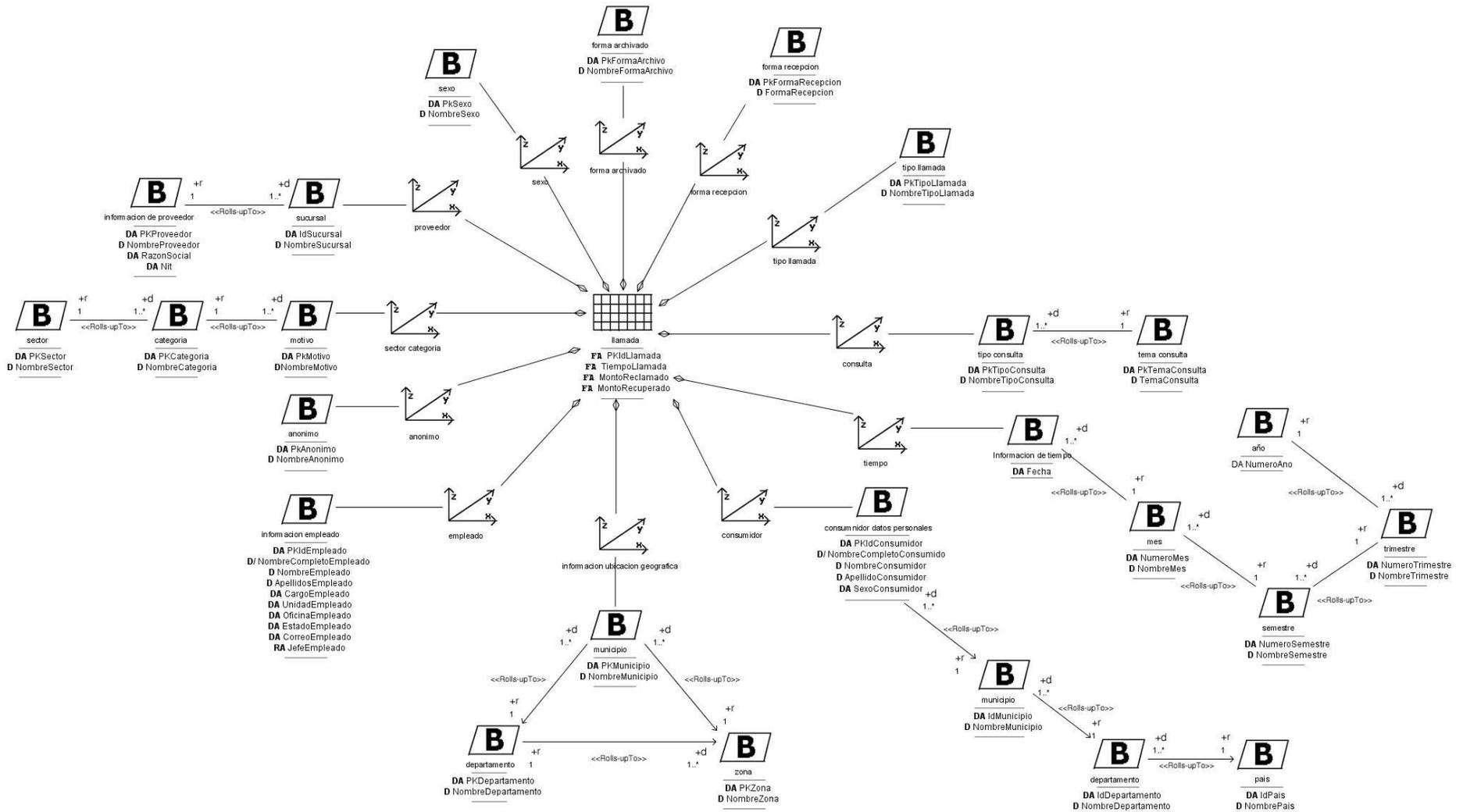


Figura 3.98: Definición de las dimensiones y hechos "Call Center"

3.9.3.2. Esquema lógico del Data Warehouse

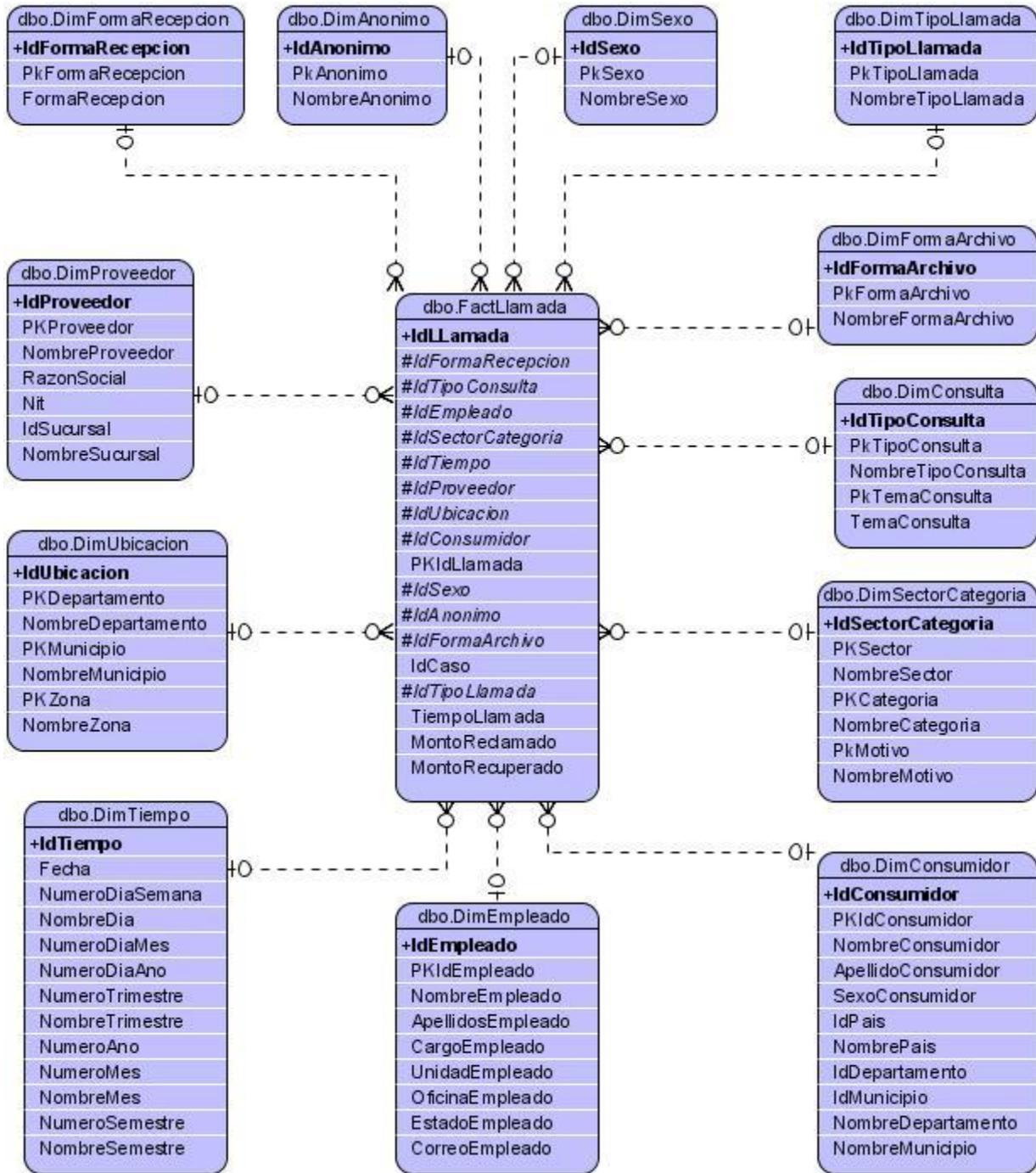


Figura 3.99: Estructura lógica del esquema estrella para el módulo de “Call Center”

### 3.9.3.3. Diccionario de datos de las dimensiones y tablas de hecho

DIM UBICACION		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDUBICACION	INT	Identificador de la ubicación en el DW
PKDEPARTAMENTO	VARCHAR(3)	Identificador del departamento de la base FUENTE
NOMBREDEPARTAMENTO	VARCHAR(50)	Nombre del departamento
PKMUNICIPIO	VARCHAR(4)	Identificador del municipio de la base FUENTE
NOMBREMUNICIPIO	VARCHAR(50)	Nombre del municipio
PKZONA	VARCHAR(3)	Identificador de la zona de la base FUENTE
NOMBREZONA	VARCHAR(50)	Nombre de la zona

DIM PROVEEDOR		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDPROVEEDOR	INT	Identificador del proveedor en el DW
NOMBREPROVEEDOR	VARCHAR(200)	Nombre de proveedor
RAZONSOCIAL	VARCHAR(200)	Razón social del proveedor
NIT	VARCHAR(25)	NIT del proveedor
IDSUCURSAL	INT	Identificador de sucursal
NOMBRESUCURSAL	VARCHAR(100)	Nombre de la sucursal
PKPROVEEDOR	INT	Identificador del proveedor de la base de datos FUENTE

DIM CONSUMIDOR		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDCONSUMIDOR	INT	Identificador del consumidor en el DW
NOMBRECONSUMIDOR	VARCHAR(50)	Nombre del consumidor
APELLIDOCONSUMIDOR	VARCHAR(50)	Apellidos del consumidor
SEXOCONSUMIDOR	CHAR(10)	Sexo del consumidor F o M
IDPAIS	VARCHAR(4)	Identificador del país del Consumidor
NOMBREPAIS	VARCHAR(50)	Nombre del país del consumidor
IDDEPARTAMENTO	VARCHAR(3)	Identificador del departamento del Consumidor
NOMBREDEPARTAMENTO	VARCHAR(50)	Nombre del departamento
IDMUNICIPIO	VARCHAR(4)	Identificador del municipio
NOMBREMUNICIPIO	VARCHAR(50)	Nombre del municipio
PKIDCONSUMIDOR	INT	Identificador del consumidor de la base de datos FUENTE

DIM SECTORCATEGORIA		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDSECTORCATEGORIA	INT	Identificador del sector categoría en el DW
NOMBRECATEGORIA	VARCHAR(100)	Nombre de la categoría
PKSECTOR	VARCHAR(3)	Identificador del sector de la base FUENTE
NOMBRESECTOR	VARCHAR(100)	Nombre del sector
PKCATEGORIA	VARCHAR(4)	Identificador de la categoría de la base FUENTE
IDSECTORCATEGORIA	INT	Identificador del sector categoría en el DW
NOMBRECATEGORIA	VARCHAR(100)	Nombre de la categoría

DIM TIEMPO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDTIEMPO	INT	Identificador de tiempo del DW
FECHA	DATETIME	Fecha de ingreso de la denuncia
NUMERODIASEMANA	INT	Numero del día en la semana
NOMBREDIA	VARCHAR(50)	Nombre del día de la semana
NUMERODIAMES	INT	Numero del día en el mes
NUMERODIAANO	INT	Numero del día en el año
NUMEROTRIMESTRE	INT	Numero del trimestre del año
NOMBRETRIMESTRE	VARCHAR(50)	Nombre del trimestre del año
NUMEROANO	INT	Numero del año
NUMEROMES	INT	Numero del mes
NOMBREMES	VARCHAR(50)	Nombre del mes
NUMEROSEMESTRE	INT	Numero del semestre
NOMBRESEMESTRE	VARCHAR(50)	Nombre del semestre

DIM SEXO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDSEXO	INT	Identificador de sexo del empleado del DW
PKSEXO	CHAR(1)	Identificador de sexo de la base FUENTE
NOMBRESEXO	VARCHAR(100)	Nombre sexo

DIM FORMARECEPCION		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDFORMARECEPCION	INT	Identificador de forma de recepción del DW
PKFORMARECEPCION	VARCHAR(4)	Identificador de forma de recepción de la base FUENTE
FORMARECEPCION	VARCHAR(100)	Nombre de la forma de recepción

DIM ANONIMO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDANONIMO	INT	Identificador de la denominación anónimo en el DW
PKANONIMO	INT	Identificador de denominación anónimo en base FUENTE
NOMBREANONIMO	VARCHAR(20)	Nombre del tipo de anónimo.

DIM EMPLEADO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDEMPLEADO	INT	Identificador del empleado en el DW
NOMBREEMPLEADO	VARCHAR(50)	Nombre del empleado
APELLIDOSEMPLEADO	VARCHAR(50)	Apellido de empleado
CARGOEMPLEADO	VARCHAR(50)	Cargo del empleado
UNIDAEMLPEADO	VARCHAR(50)	Unidad a la que pertenece el empleado
OFICINAEMPLEADO	VARCHAR(100)	Oficina a la que forma parte el empleado
ESTADOEMPLEADO	VARCHAR(50)	Estado del empleado Activo o Desactivo
CORREOEMPLEADO	VARCHAR(100)	Correo electrónico del empleado
PKIDEMPLEADO	INT	Identificador del empleado de la base de datos FUENTE

DIM TIPOLLAMADA		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDTIPOLLAMADA	INT	Identificador de tipo de llamada del DW
PKTIPOLLAMADA	VARCHAR(5)	Identificador de tipo de llamada de la base FUENTE
NOMBRETIPOLLAMADA	VARCHAR(50)	Nombre del tipo de llamada

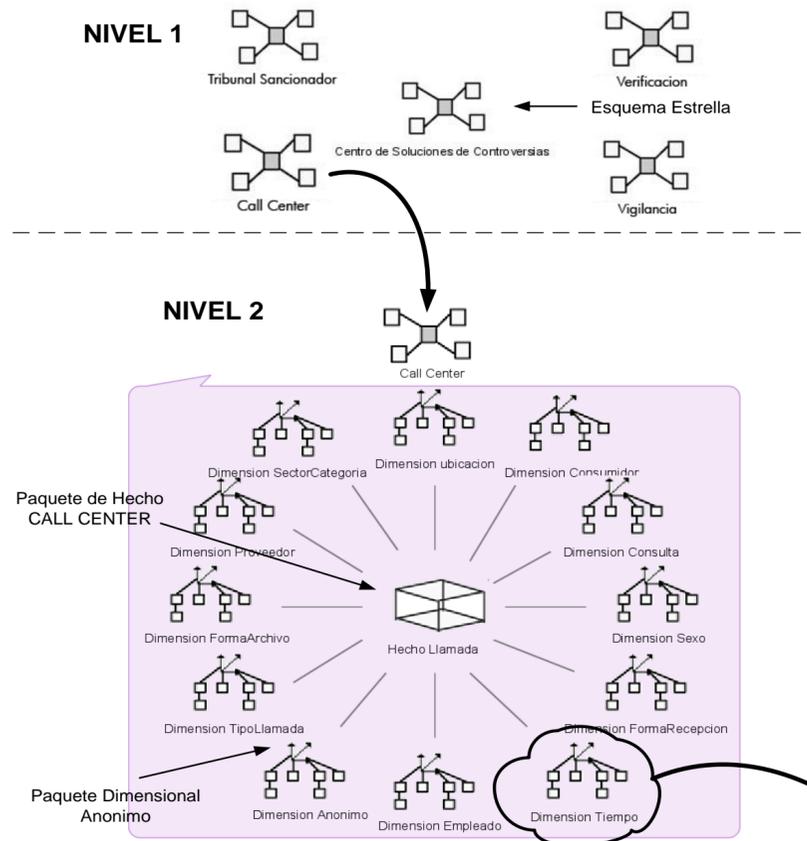
DIM FORMAARCHIVO		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDFORMAARCHIVO	INT	Identificador de forma archivo del DW
PKFORMAARCHIVO	INT	Identificador de forma archivo de la base FUENTE
NOMBREFORMAARCHIVO	VARCHAR(20)	Nombre de la forma archivo

DIM CONSULTA		
ATRIBUTO	TIPO DE DATO	DESCRIPCIÓN
IDTIPOCONSULTA	INT	Identificador del tipo de consulta del DW
PKTIPOCONSULTA	VARCHAR(3)	Identificador del tipo de consulta de la base FUENTE
NOMBRETIPOCONSULTA	VARCHAR(100)	Nombre del tipo de consulta
PKTEMACONSULTA	VARCHAR(4)	Identificador del tema consulta de la base FUENTE
TEMACONSULTA	VARCHAR(100)	Nombre del tema consulta

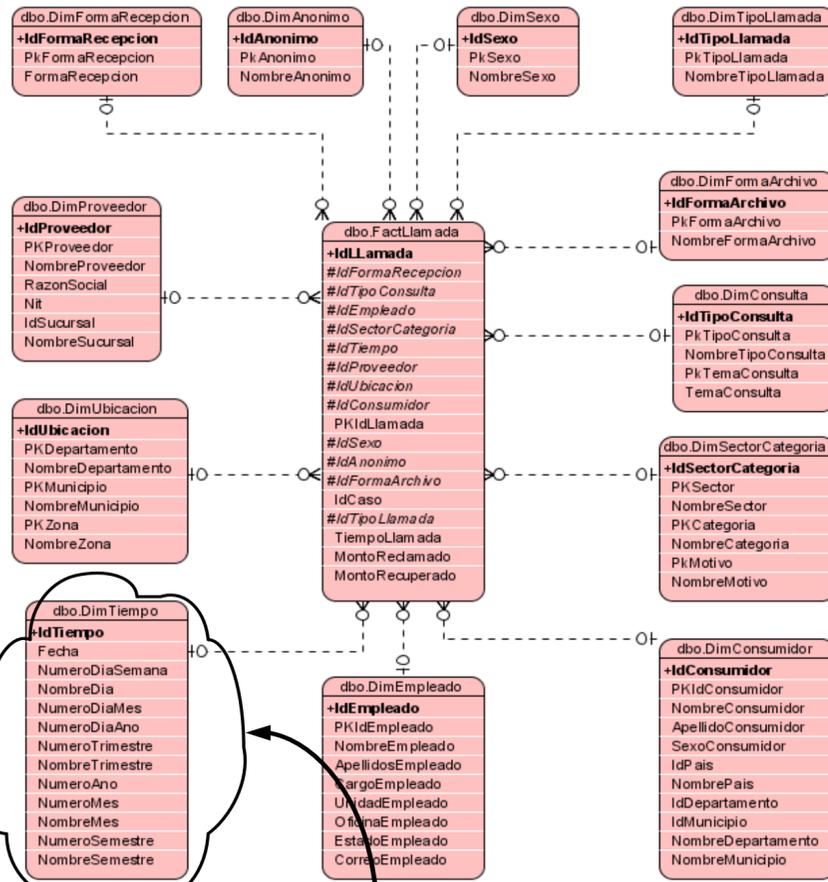
3.9.4. Diagrama UML del DATA WAREHOUSE – CALL CENTER

DIAGRAMA UML DEL DATA WAREHOUSE- CALL CENTER

ESQUEMA CONCEPTUAL DEL DATA WAREHOUSE



ESQUEMA LÓGICO DEL DATA WAREHOUSE



ESQUEMA FISICO DEL DATA WAREHOUSE

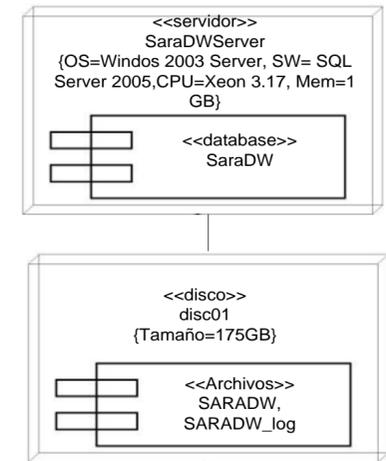
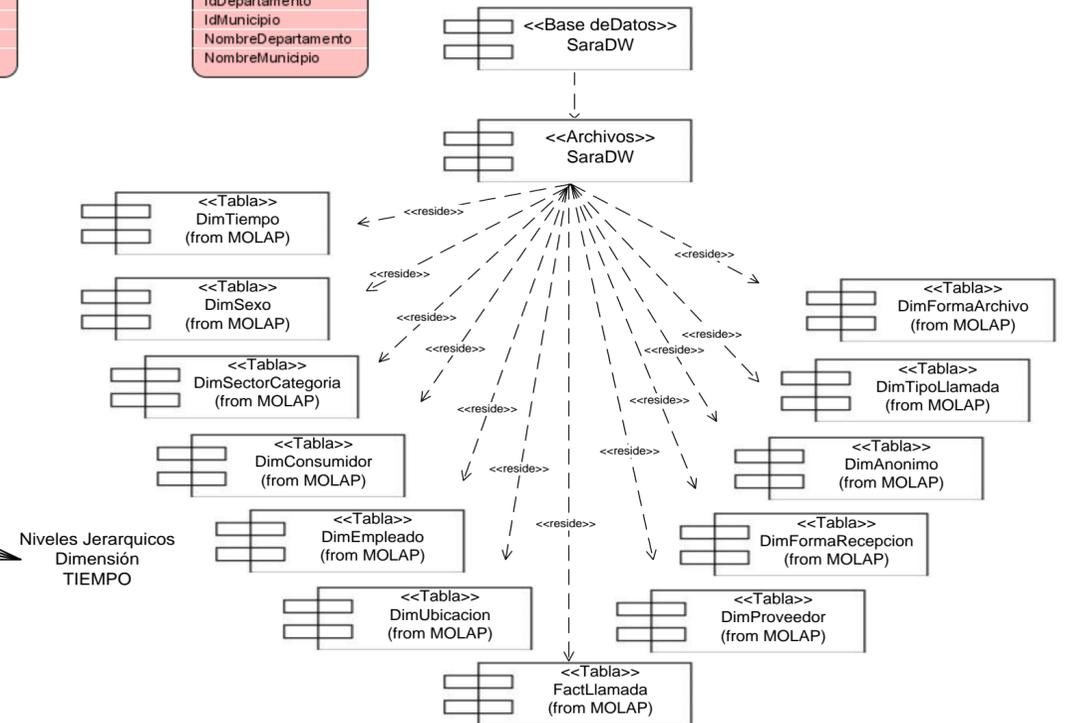


DIAGRAMA DE COMPONENTES



Clase Dimensional TIEMPO

Niveles Jerarquicos Dimensión TIEMPO

# DIAGRAMA ETL – CALL CENTER

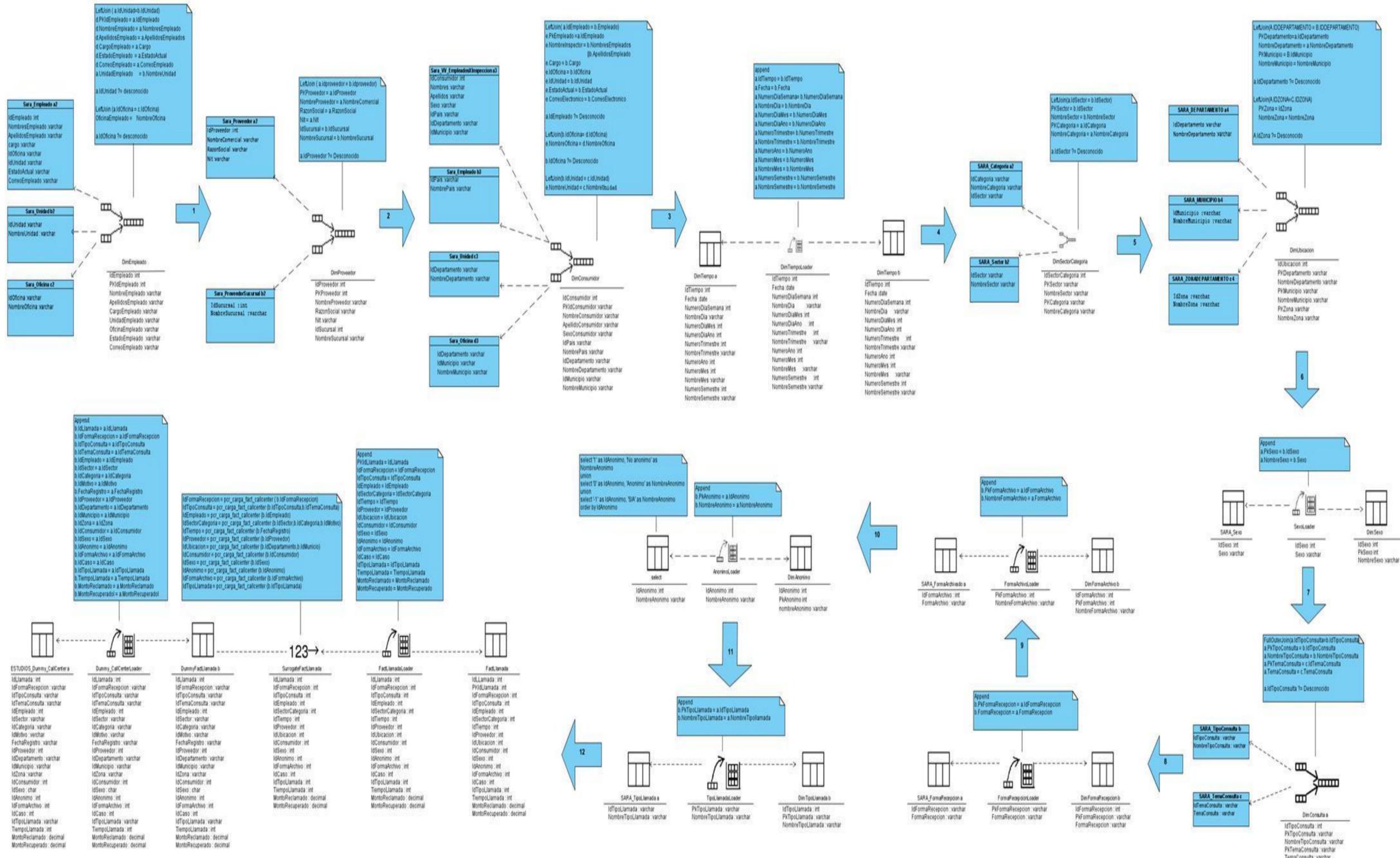
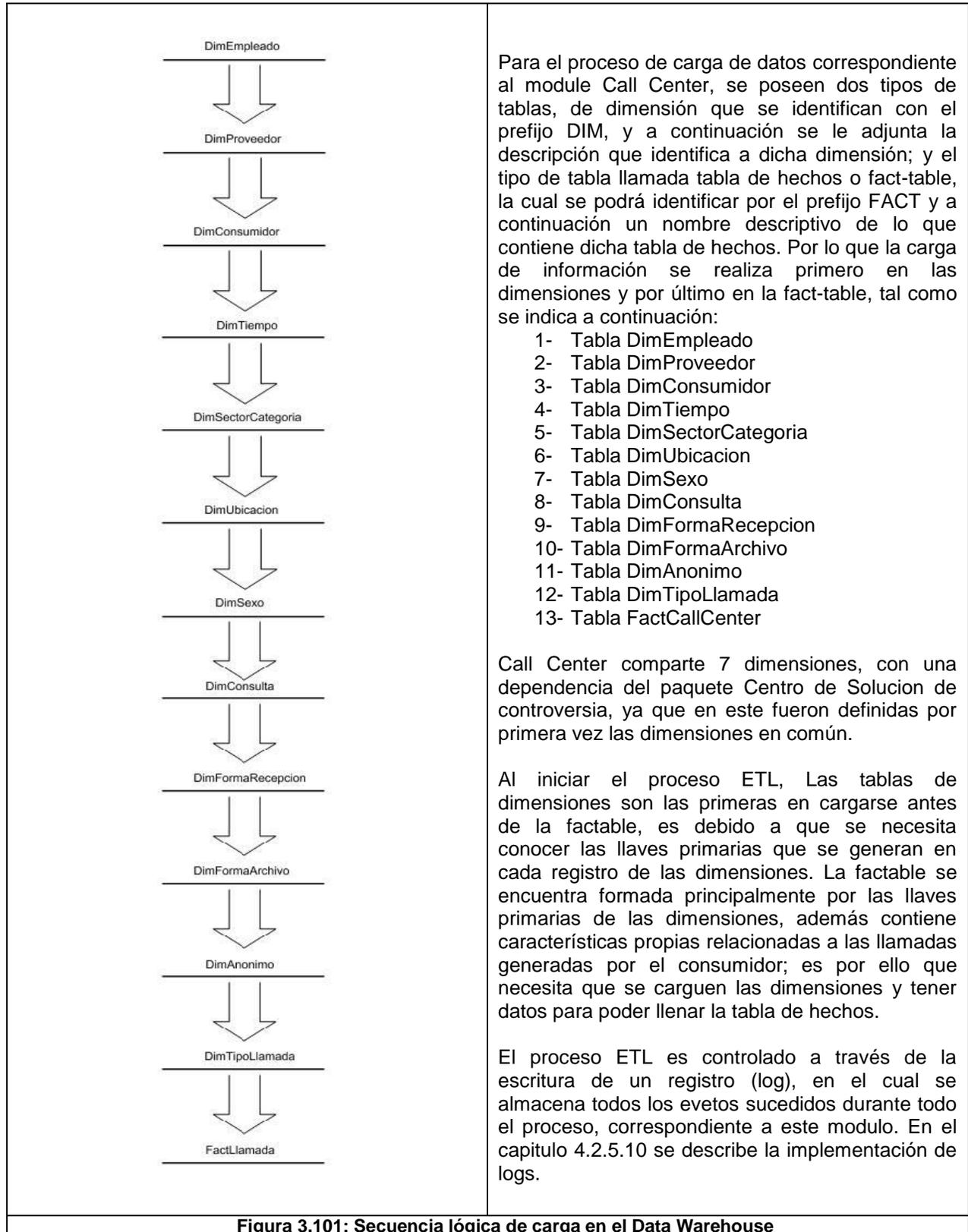


Figura 3.100: Diseño del proceso ETL para el Data Warehouse

**3.9.5.1. Secuencia de carga del proceso ETL**



**Figura 3.101: Secuencia lógica de carga en el Data Warehouse**

3.9.5.2. Diagrama de carga para la dimensión Empleado

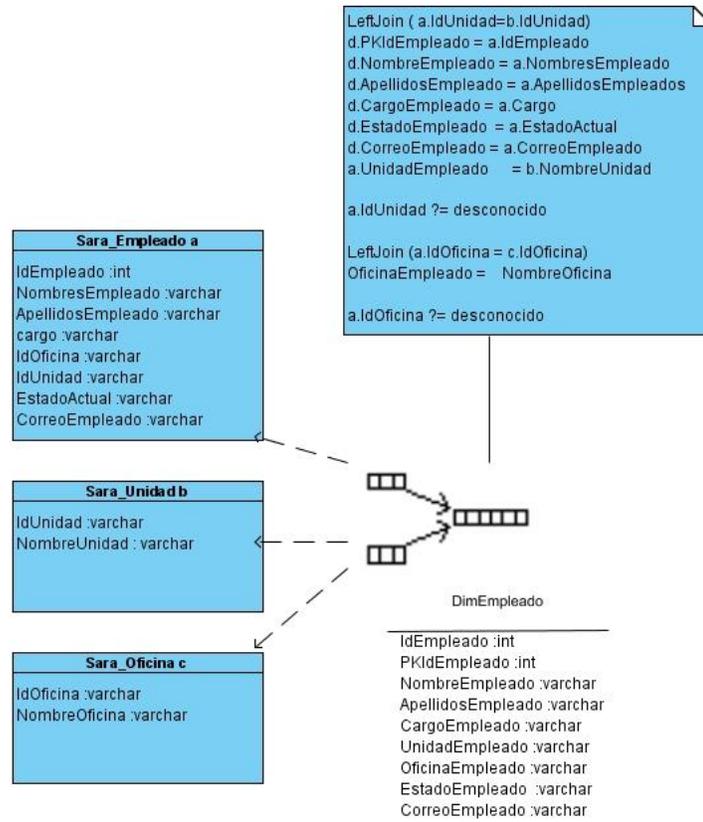


Figura 3.102: Representación de carga de datos en la tabla DimProveedor utilizando mecanismo Join

3.9.5.3. Diagrama de carga para la dimensión Proveedor

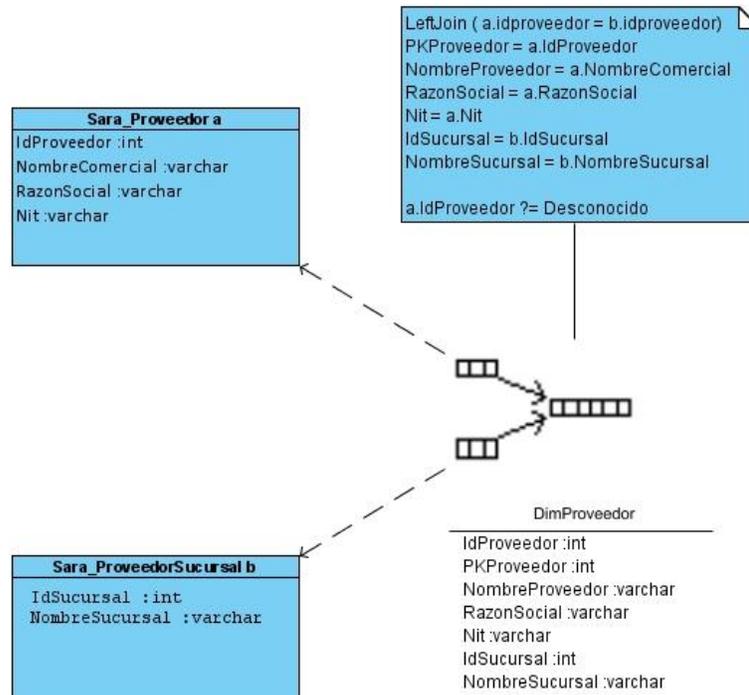


Figura 3.103: Representación de carga de datos en la tabla DimProveedor utilizando mecanismo Join

3.9.5.4. Diagrama de carga para la dimensión Consumidor

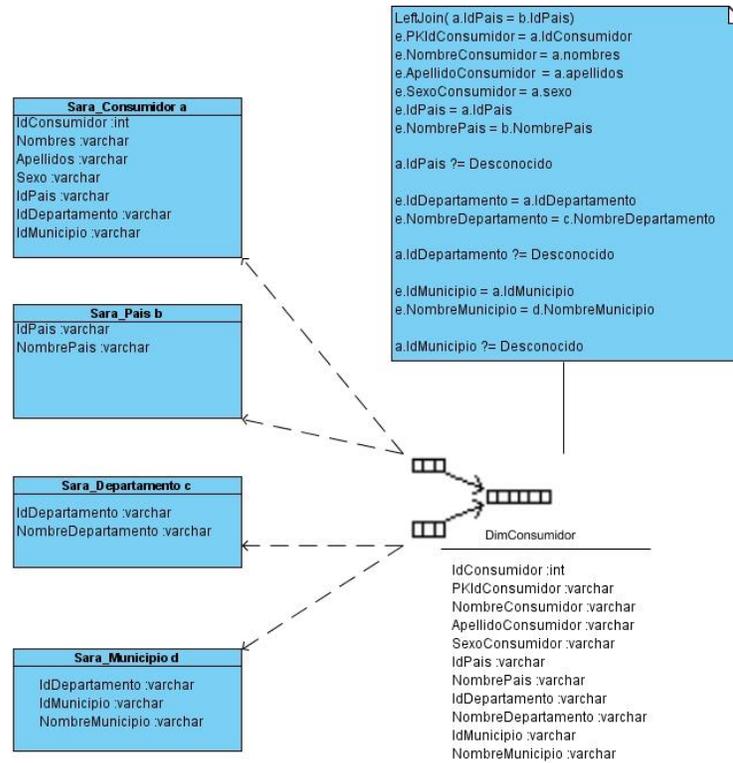


Figura 3.104: Representación de carga de datos en la tabla DimConsumidor utilizando mecanismo Join

3.9.5.5. Diagrama de carga para la dimensión Tiempo

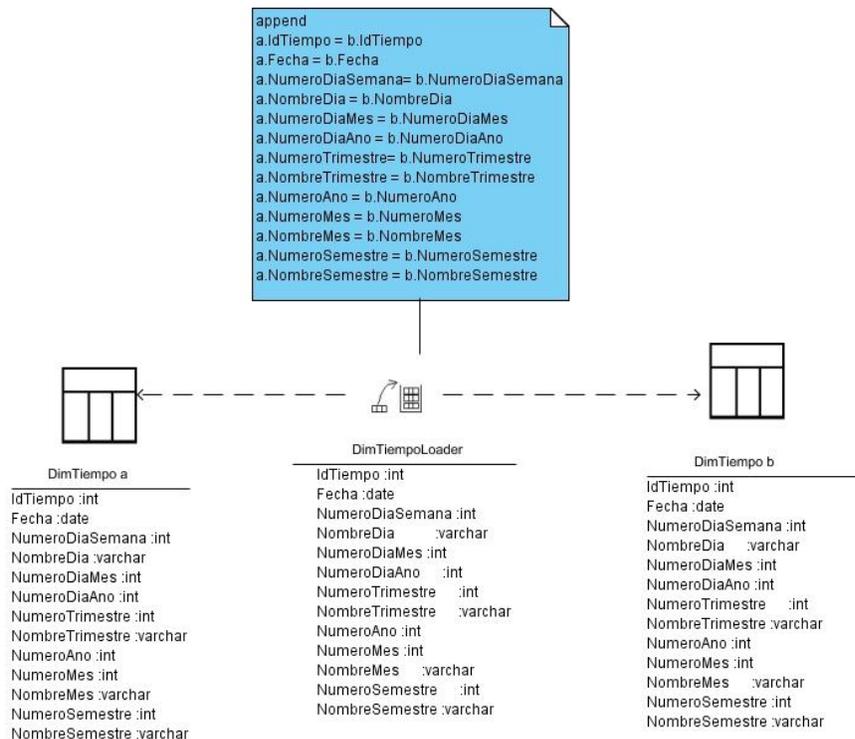


Figura 3.105: Representación de carga de datos en la tabla DimTiempo utilizando mecanismo Loader

3.9.5.6. Diagrama de carga para la dimensión SectorCategoría

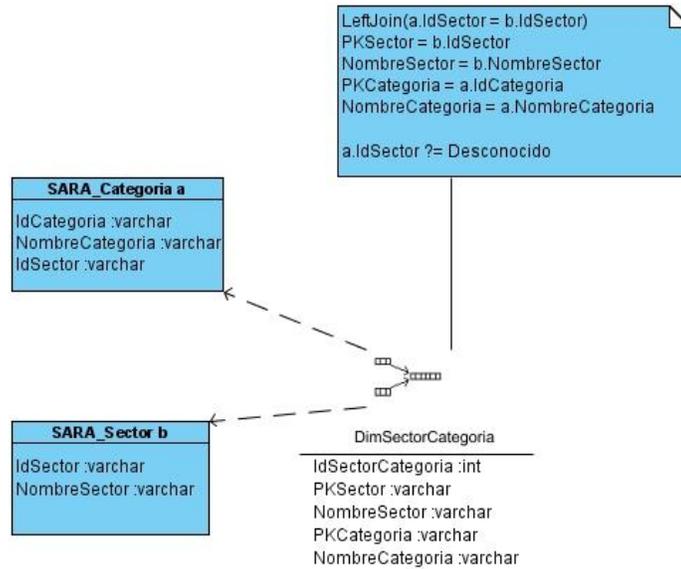


Figura 3.106: Representación de carga de datos en la tabla `DimSectorCategoría` utilizando mecanismo `Join`

3.9.5.7. Diagrama de carga para la dimensión Ubicación

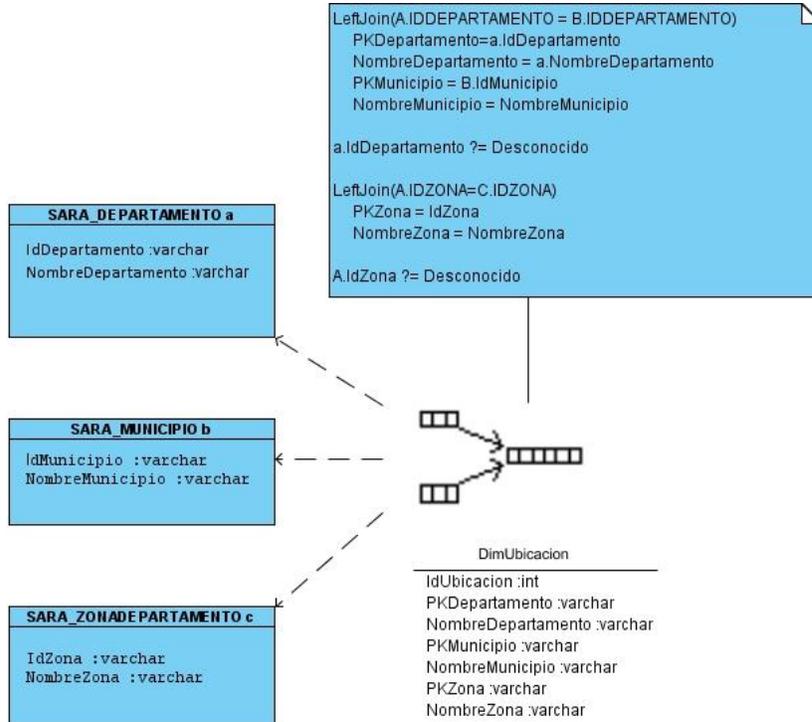


Figura 3.107: Representación de carga de datos en la tabla `DimUbicación` utilizando mecanismo `Join`

3.9.5.8. Diagrama de carga para la dimensión Consulta

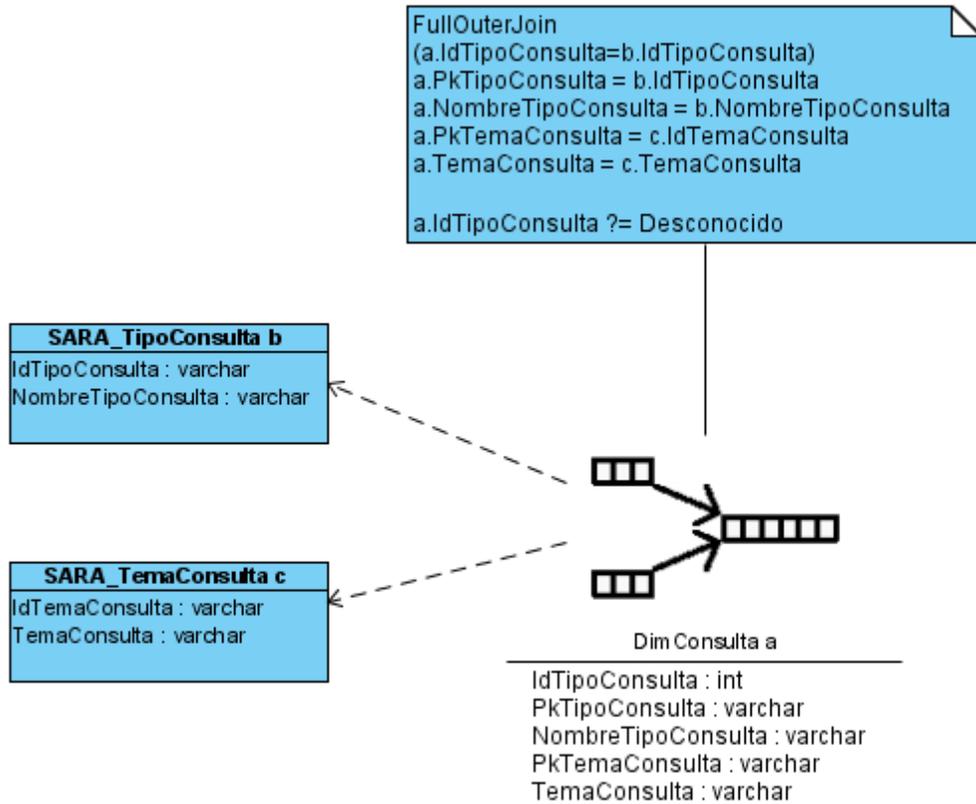


Figura 3.108: Representación de carga de datos en la tabla DimConsulta utilizando mecanismo Join

3.9.5.9. Diagrama de carga para la dimensión FormaRecepcion

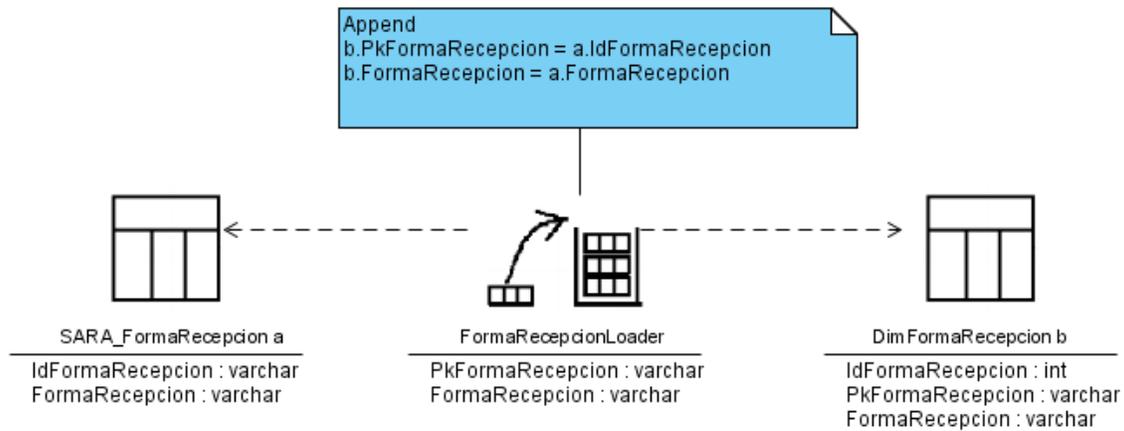


Figura 3.109: Representación de carga de datos en la tabla DimRecepcion utilizando mecanismo Loader

3.9.5.10. Diagrama de carga para la dimensión FormaArchivo

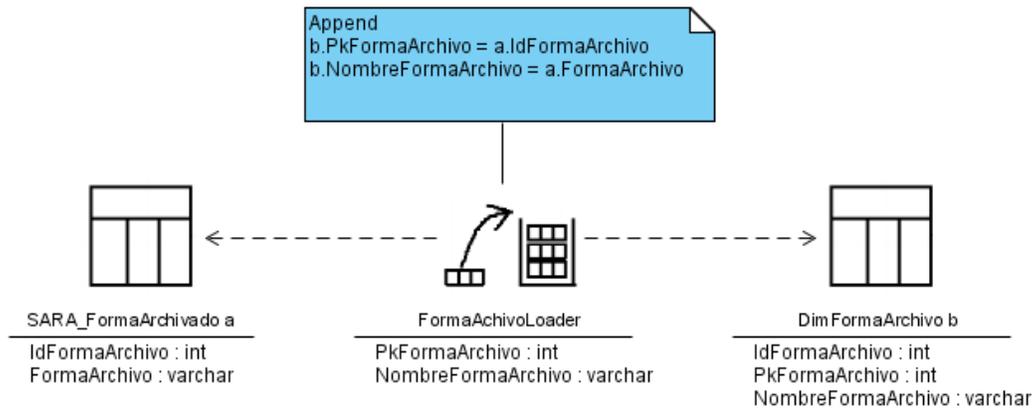


Figura 3.110: Representación de carga de datos en la tabla DimFormaArchivo utilizando mecanismo Loader

3.9.5.11. Diagrama de carga para la dimensión Anonimo

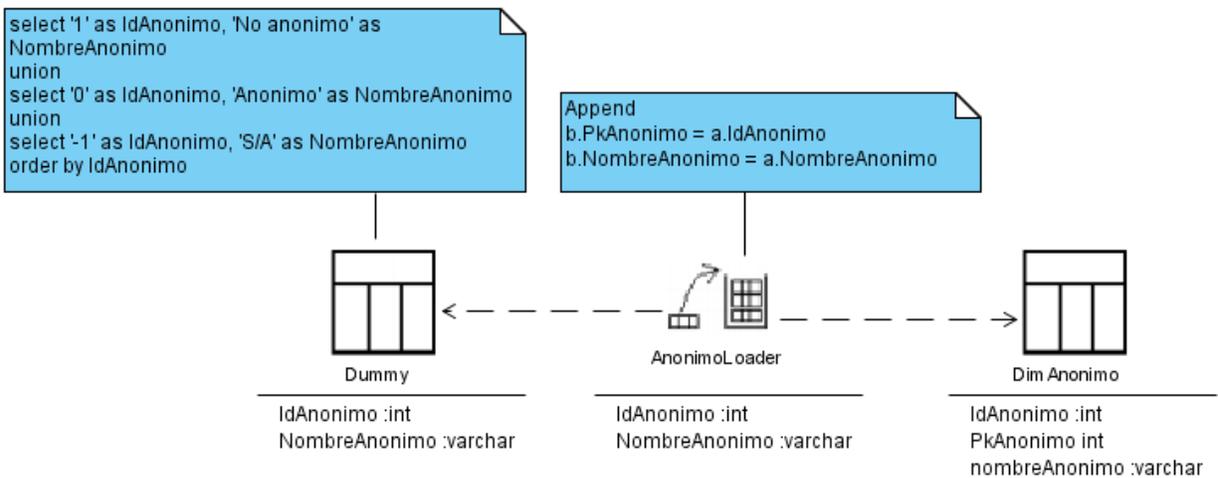


Figura 3.111: Representación de carga de datos en la tabla DimAnonimo utilizando mecanismo Loader

3.9.5.12. Diagrama de carga para la dimensión TipoLlamada

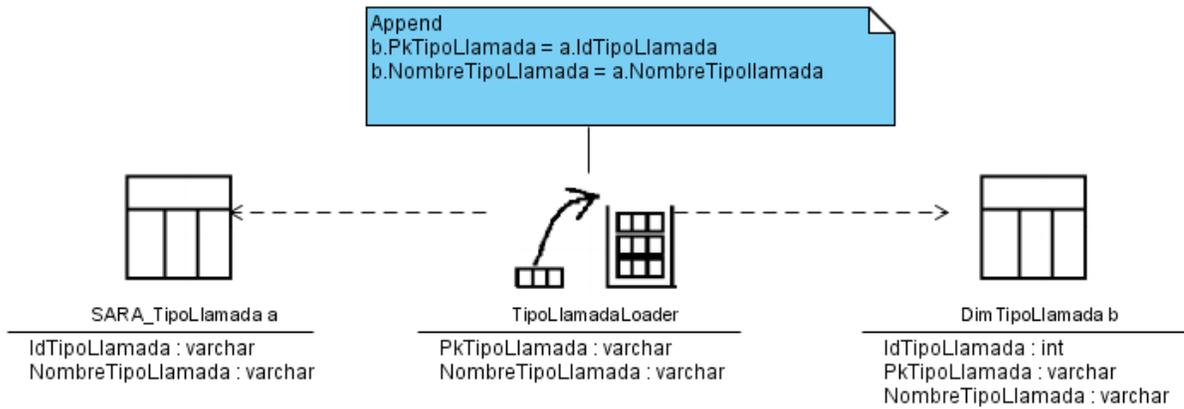


Figura 3.112: Representación de carga de datos en la tabla DimTipoLlamada utilizando mecanismo Loader

### 3.9.5.13. Diagrama de carga para la dimensión FactCallCenter

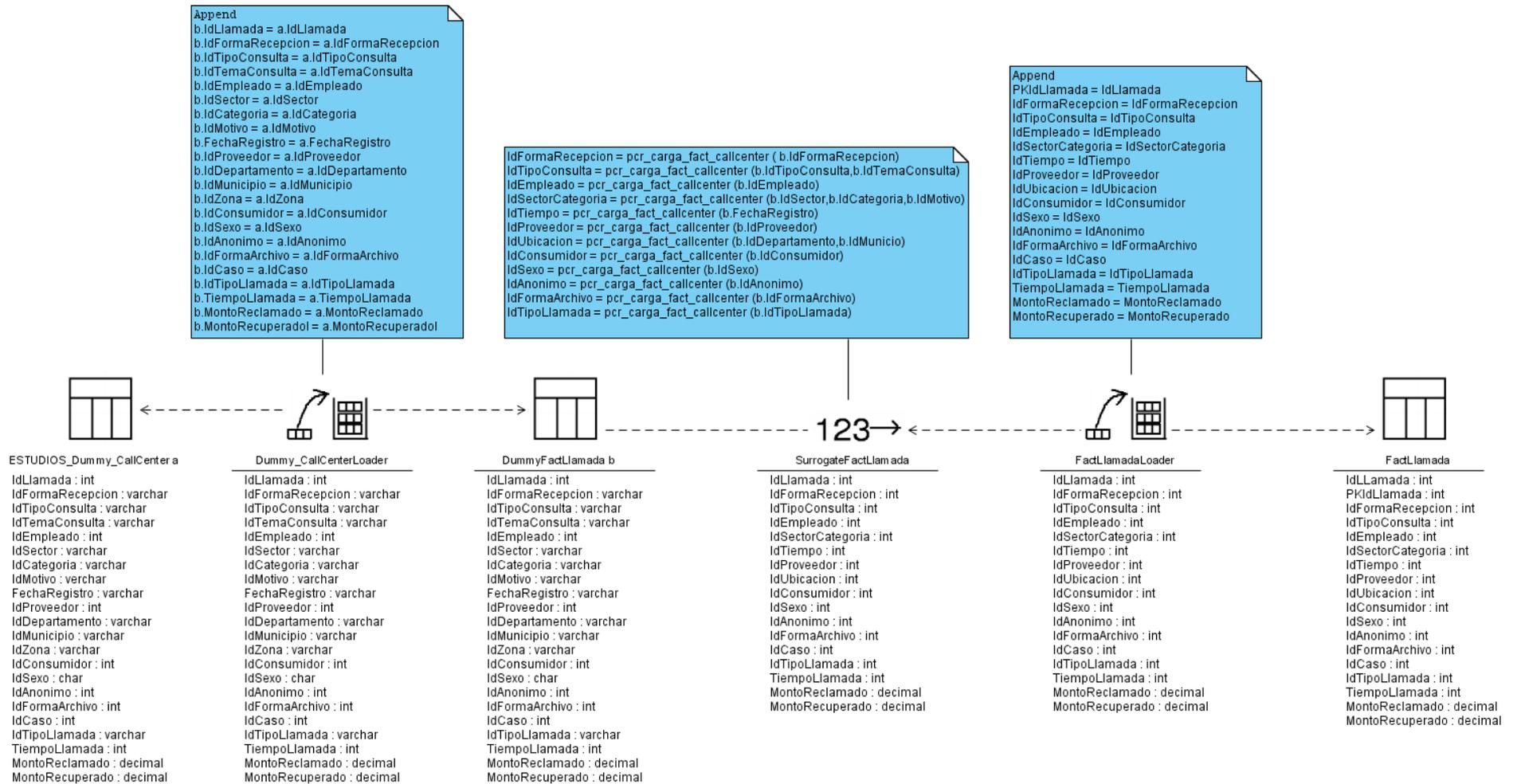


Figura 3.113: Representación de carga de datos en la tabla FactCallCenter utilizando mecanismo Loader y Surrogate

### 3.10. DISEÑO DE SALIDAS

#### 3.10.1.1. Diseño de interfaces de salida para reportes de Vista de Análisis

Descripción General de la Salida	
<b>Salida</b>	Reportes de Vista de Análisis
<b>Descripción</b>	Detalle de las partes que contiene la salida.
<b>Tecnología de Salida</b>	Pantalla, Archivo.
<b>Propósito de la Salida</b>	Presentar al usuario información correspondiente al modulo del CSC.

Tabla 3.1: Elementos de interfaces de salida reportes vista de análisis

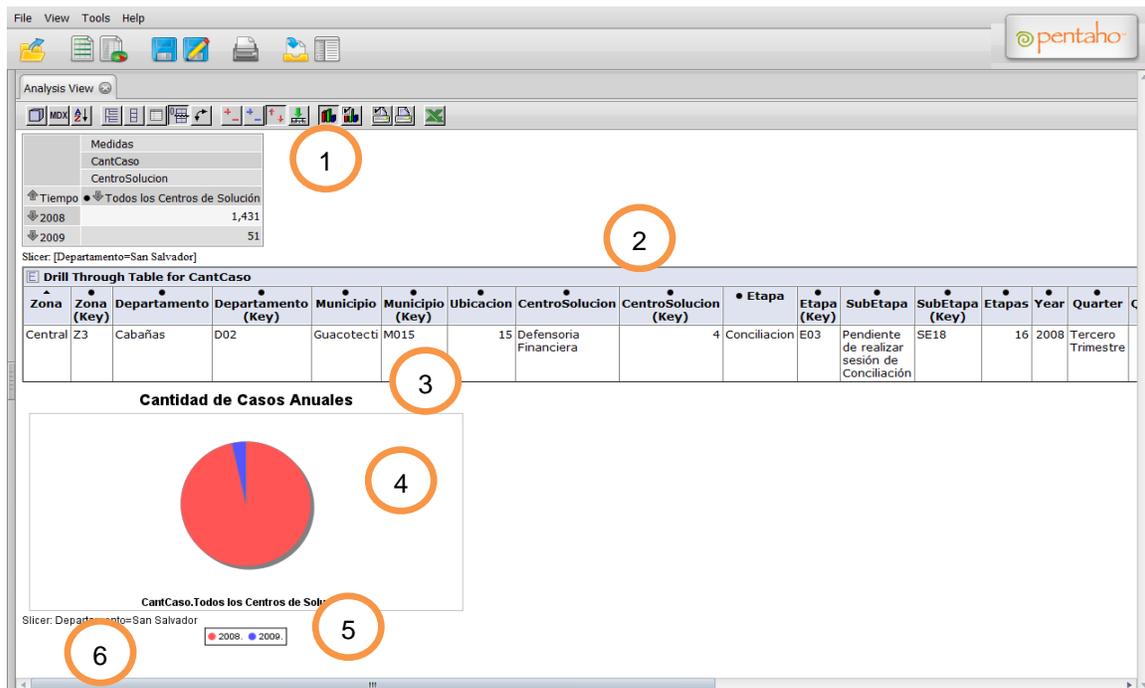


Figura 3.114: Salida de Reporte de Análisis

En la Figura 3.114 se detalla la salida para un reporte de análisis, en la tabla 3.2 se especifican cada uno de los puntos marcados en la figura.

Número	Descripción	Mostrado por defecto
1	En esta área se muestran las dimensiones y medidas seleccionadas para la realización del análisis. Además se muestran los resultados de las consultas de análisis realizadas. Elemento mostrado por defecto.	Si
2	Área en la que se muestra el detalle de los datos de los cuales se obtiene la información resumida para el análisis de la información.	No
3	Título del gráfico a desplegar.	No
4	Gráfico que representa los datos mostrados en el área 1.	No
5	Leyenda del gráfico.	No

6	En esta área se especifican los filtros utilizados para la realización del análisis.	No
---	--	----

Tabla 3.2: Descripción de componentes de reportes analíticos

3.10.1.2. Diseño de interfaces de salida para reportes Ad Hoc

Descripción General de la Salida	
Salida	Reportes
Descripción	Detalle de las partes que contiene la salida.
Tecnología de Salida	Pantalla, Archivo.
Propósito de la Salida	Presentar al usuario información correspondiente al modulo del CSC con los parámetros seleccionados.

Tabla 3.3: Descripción de elementos de interfaces de salida

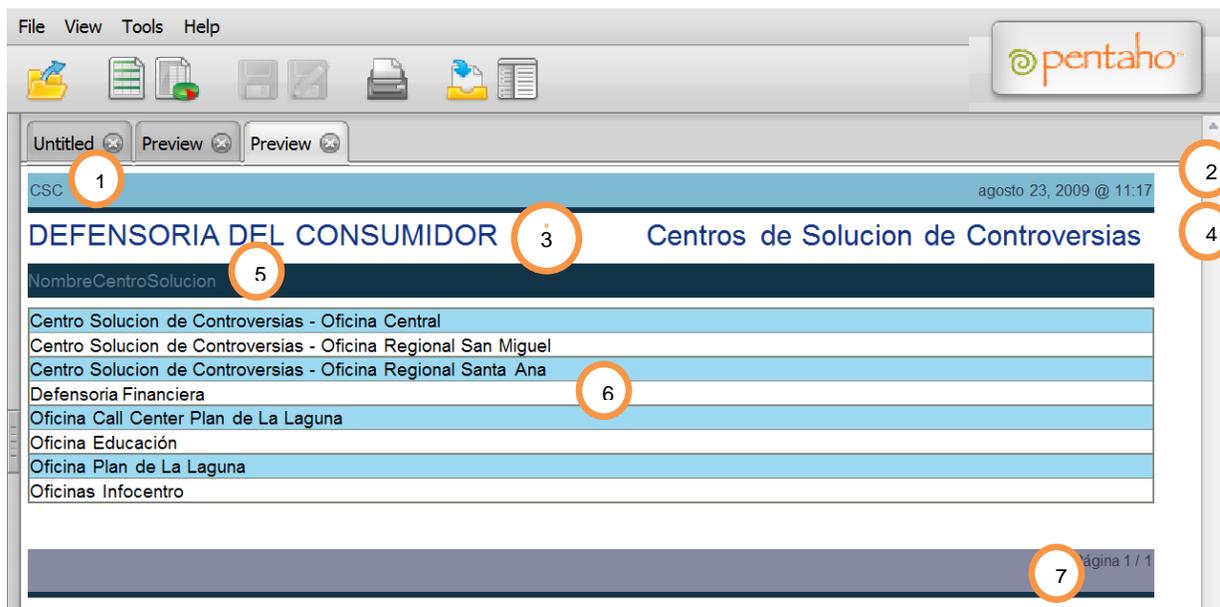


Figura 3.115: Salida para reportes Ad Hoc

En la Figura 3.115 se detalla la salida para un reporte de análisis, en la tabla 3.4 se especifican cada uno de los puntos marcados en la figura.

Número	Descripción	Mostrado por defecto
1	Área del encabezado de página. El asistente permite al usuario digitar la leyenda que desee.	No
2	Área en la que se muestra fecha y hora de creación del reporte.	Si
3	Título de la Organización.	Si
4	Encabezado del reporte. El asistente permite al usuario digitar la leyenda que desee.	No
5	Encabezado del detalle de datos mostrados.	Si
6	Área en la que se muestra el detalle del reporte.	Si
7	Numero de página/Total de Páginas	Si

Tabla 3.4: Detalle de componentes de reportes Ad Hoc

### 3.11. DISEÑO DE SEGURIDAD

Uno de los aspectos que toda aplicación debe considerar es la seguridad, entendiendo como tal la necesidad que el usuario es quien dice ser (autenticación), y permite el acceso sólo a aquellos recursos necesarios (autorización). La seguridad que proporciona la plataforma BI de Pentaho está basada en “Acegi Security Sistem for Spring”, el cual es un framework Java/Java EE, el cual provee autenticación, autorización y otras funciones de seguridad para aplicaciones.

#### 3.11.1. Autorización

Pentaho utiliza dos tipos de autorización:

- Autorización de recursos web: Habilitada por defecto, los usuarios deben iniciar sesión para poder acceder a la interfaz Web de Pentaho.
- Autorización basada en objetos de dominio: Deshabilitada por defecto, permite la ejecución de action secuencias (archivos xaction) pueden ser ejecutados directamente sin permiso alguno.

#### 3.11.2. Seguridad Usuario/Rol

Existen tres formas de implementar la seguridad basada en usuarios y roles en Pentaho:

- Memoria: La instalación por defecto de Pentaho utiliza la este tipo de seguridad por defecto.
- JDBC: Utiliza una base de datos relacional para manejar la autenticación.
- LDAP: Utiliza un directorio para manejar la autenticación. Se puede utilizar el Active Directory como servidor LDAP.

La seguridad que será utilizada en el proyecto es la JDBC ya que provee integración con todas las herramientas de desarrollo y las bases de datos que soportan este tipo de conexión. Este tipo de seguridad permite la publicación de contenido en la Suite BI de Pentaho. Otra ventaja es que todos los motores de base de datos tienen un cliente ya probado que permite una implementación más fácil.

#### 3.11.3. Seguridad por Action Sequence

Una vez se ha definido la seguridad usuario/rol, se pueden asignar permisos a action sequences individuales utilizando listas de control de acceso (ACLs). También se puede utilizar la seguridad en la propia action sequence, basándose en las credenciales del usuario logeado.

Además se puede quitar la seguridad de cualquier action sequence para accesos anónimos.

#### 3.11.4. Seguridad en Mondrian

Mondrian es el motor OLAP que viene integrado con la Suite BI de Pentaho. Para el manejo de la seguridad en Mondrian, se pueden definir Roles como parte del esquema y verificar el rol cuando se intenta realizar una conexión. Las propiedades de seguridad que permite mondrian son las siguientes:

- **SchemaGrant:** Define el acceso por defecto para los objetos de un esquema. El atributo `access` puede ser “all” (todos) o “none” (ninguno); esta propiedad puede que no sea tomada en cuenta para objetos específicos.
- **CubeGrant:** Define el acceso a un cubo en particular. Al igual que el `SchemaGrant`, posee el atributo `access` el cual puede no ser tomado en cuenta para sub-objetos particulares del cubo.
- **HierarchyGrant:** Define el acceso a una jerarquía. Este atributo puede ser “all” (todos), donde todos los miembros son visibles; “none” (ninguno), donde la jerarquía no es mostrada al usuario y “custom” (personalizado), en la cual se puede utilizar el atributo “`topLevel`” (nivel máximo) para definir el nivel máximo al cual pueden acceder los usuarios. Esta propiedad se puede utilizar para que los usuarios no puedan visualizar datos a nivel de país, permitiéndoles ver datos únicamente a nivel de departamento. Además, existe el atributo “`bottomLevel`” (nivel inferior) para definir el nivel más bajo que será visible.
- **MemberGrant:** Permite el acceso a ciertos miembros definiendo elementos de este tipo. Únicamente se pueden definir elementos `MemberGrant` si la propiedad `HierarchyGrant` tiene acceso de tipo “custom”. Los accesos por miembro otorgan acceso a un miembro basándose en cuatro reglas, las cuales son:

**Los miembros heredan el acceso de sus padres.** Por ejemplo, si niega el acceso al departamento de San Salvador, no se podrá observar los datos del municipio de Mejicanos.

**Los accesos son dependientes del orden en el cual son asignados.** Si se da acceso al departamento de San Salvador y posteriormente se quita el acceso a Mejicanos, entonces, no se podrá observar Mejicanos; pero si se niega el acceso a Mejicanos y luego se le da acceso al departamento de San Salvador, se podrá observar todo lo correspondiente al departamento.

**Un miembro es visible si cualquiera de sus hijos son visibles.** Suponiendo que se niega el acceso al departamento de San Salvador y luego se le otorga acceso al municipio de Mejicanos, se podrá observar el departamento de San Salvador y Mejicanos, pero no se podrá ver el municipio de San Salvador. Sin embargo, los totales correspondientes al departamento de San Salvador reflejarán el de todos los municipios.

**Los accesos de miembros no anulan los accesos otorgados a las jerarquías.** Si se define la propiedad `topLevel` a nivel de Centro de Solución de Controversias, y se le da acceso al departamento de San Salvador, no se podrán observar los datos correspondientes al país.

Otro método de seguridad utilizado por Mondrian son los llamados “**Rollup policy**” o políticas de Roll-up. La cual determina como será calculado el total de un miembro si el rol no puede ver todos los hijos del miembro. La política por defecto que se asigna es llamada “full” (completa), con la cual el cálculo total incluye los datos correspondientes a los hijos que no son visibles. Por ejemplo, si un usuario pertenece a un rol que puede visualizar Mejicanos, pero no puede ver el municipio de San Salvador, los datos totales incluirán los contenidos en el municipio de San Salvador. Las políticas que se pueden definir son las que se detallan a continuación:

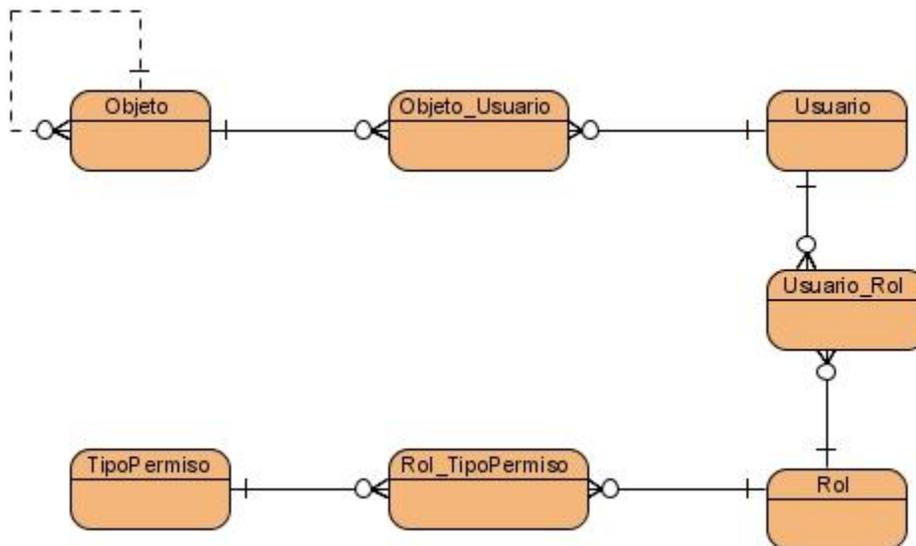
- Full: El total de un miembro incluye todos sus hijos.
- Partial: El total de un miembro incluye únicamente a los hijos que tiene acceso.
- Hidden: Si cualquiera de los hijos es inaccesible, el total no es mostrado.

### 3.11.5. Seguridad en reportes Ad Hoc

Los reportes ad hoc se manejan por medio de Pentaho Metadata, el cual facilita a los usuarios el acceso a la información en términos de negocio. Los niveles de seguridad proporcionados son:

- Seguridad a nivel de columna: Permite controlar los resultados que son devueltos basándose en la seguridad a nivel de usuario. Se puede especificar las columnas a las cuales puede acceder un determinado usuario o rol.
- Restricciones basadas en Roles: Si se utiliza este tipo de seguridad, el motor de Metadatos determina que restricciones se aplican al usuario y se las aplica inmediatamente a la consulta que está realizando el usuario.
- Seguridad a nivel de Vista de Negocio: Permite el acceso por medio de roles a las vistas de negocio definidas en el modelo de metadatos. Si el rol no tiene acceso a una vista de negocio, esta no será elegible para realizar reportes ad hoc con la información contenida en ella.

De lo anterior se puede crear una estructura que representa la seguridad proporcionada por Pentaho BI Suite. La figura 3.116 muestra dicha estructura.



**Figura 3.116: Esquema lógico de la seguridad proporcionada por Pentaho**

En la tabla 3.5 se detalla la descripción de cada una de las entidades de la figura 3.116

Entidad	Descripción
Objeto	Almacenamiento y relación padre-hijo entre todos los objetos de la plataforma BI.
Usuario	Datos de inicio de sesión del usuario.
Rol	Los distintos roles que pueden ser asignados a los usuarios previamente definidos en la entidad Usuario.
TipoPermiso	Se definen los distintos tipos de permiso dentro de los elementos de Pentaho.
Objeto_Usuario	Objetos a los que tiene acceso un usuario determinado.
Usuario_Rol	Los roles asignados a cada uno de los usuarios.
Rol_TipoPermiso	Permisos proporcionados a los roles.

**Tabla 3.5: Descripción de las entidades**

### 3.12. REFINAMIENTO DEL PROTOTIPO

Con el fin de garantizar un sistema como producto final funcional, es necesario aplicar control de calidad a cada una de las iteraciones realizadas, por lo cual la metodología seleccionada incluye el refinamiento del prototipo creado por iteración.

La dinámica que se presentara implica que al primer prototipo se le realizaran modificaciones según los comentarios proporcionado por los usuarios y se creara un segundo prototipo el cual será evaluado de la misma manera que el primero. Estas etapas se realizan tantas veces como lo sea necesario a fin de que el usuario quede satisfecho con el prototipo que represente al producto final deseado.

A continuación se presenta el formulario a utilizar para las pruebas a realizar en el refinamiento de los prototipos, junto a la descripción de cada campo que compone el formulario.

Campos.

1. Nombre del Sistema: Nombre del sistema a probar.
2. Nombre del Modulo: Nombre del modulo del sistema a probar
3. Número de Prueba: Numero correlativo de la prueba realizada
4. Fecha de la Prueba: Fecha en que se realizo la prueba
5. Hora Inicio: Hora de inicio de la prueba.
6. Hora Finalización: Hora de finalización de la prueba
7. Descripción: Descripción de la(s) pruebas realizadas.
8. Referencia Casos de Uso: Caso en uso en que se basa la prueba realizada.
9. Acción a Ejecutar: Acciones o secuencias a probar contempladas por el caso de uso.
10. Responsables: Persona encargada en realizar la prueba
11. Detalle de Hallazgos: Secuencia de errores (bugs) encontrados.
12. Resultados Obtenidos: Resultados encontrados de a causa de errores (bugs).
13. Firma Responsable: Firma de la persona responsable de la prueba.
14. Firma Ejecutor: Firma del ejecutor de la prueba.

<b>FORMULARIO DE PRUEBAS                      DESARROLLO DE UN DATAWAREHOUSE PARA EL PROCESO DE                      DENUNCIAS DE LA DEFENSORIA DEL CONSUMIDOR</b>			
Nombre del Sistema:	(1)		
Nombre del Modulo	(2)		
Numero de prueba	(3)	Fecha de la Prueba	(4)
Hora Inicio	(5)	Hora Finalización	(6)
Descripción	(7)		
Referencia Caso de Uso	(8)		
Acciones a ejecutar	(9)		
Responsables	(10)		
Detalle de Hallazgos	(11)		
Resultados Obtenidos	(12)		
Firma Responsable	(13)	Firma Ejecutor	(14)

# Capítulo IV



## 4.1. MANUAL DE INSTALACIÓN

### 4.1.1. Instalación y Configuración de Pentaho BI Server

#### 4.1.1.1. Preparación

Se debe tener instalado y configurado previamente la última versión del JDK de Java, JRE y MySql Server. (Los pasos para instalar el JDK de Java y parametrizar el path se encuentran en el capítulo 4.1.3 de este manual)

#### 4.1.1.2. Java JDK and JRE

Se debe verificar si se tiene instalado el JDK y JRE de Java desde una consola utilizando el comando `java -version`. El resultado debe parecerse al siguiente:

```
C:\>java -version
java version "1.6.0_14"
Java(TM) SE Runtime Environment (build 1.6.0_14-b08)
Java HotSpot(TM) Client VM (build 14.0-b16, mixed mode, sharing)
```

Además se debe verificar si se tiene configurada la variable de entorno correctamente, nuevamente en la consola ejecutar el comando `set path`. En el resultado deberá tener una línea como la siguiente:

```
C:\>set path
C:\Archivos de programa\Java\jdk1.6.0_14\bin;
```

#### 4.1.1.3. MySQL Server

Para verificar si se tiene instalado correctamente MySQL se debe verificar accediendo al servidor por medio de la consola utilizando el comando `mysql -u root -p*****`. (Reemplazar los \* por la contraseña).

```
C:\>mysql -u root -p*****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 1
Server version: 5.0.67-community-nt MySQL Community Edition (GPL)
Type 'help;' or '\h' for help. Type '\c' to clear the buffer.
mysql>
```

#### 4.1.1.4. SQL Server 2005

Es necesario tener instalada una instancia del SQL Server 2005 en su versión Standard, Developer o Enterprise.

#### 4.1.1.5. Instalación y Configuración de PENTAHO

#### 4.1.1.6. Definiciones

A continuación se muestra la lista de paquetes que son relevantes en la configuración y administración de Pentaho BI Suite como herramienta de Business Intelligence.

**Apache Tomcat**

**Descripción:** Es el servidor de aplicaciones web que provee la plataforma para la capa de presentación de la Suite BI.

**Puerto:** 8080 (Por Defecto).

**Acceso por URL:** <http://localhost:8080/>

**Sitio Web:** <http://www.apache-tomcat.com>

**Hibernate**

**Descripción:** Es un paquete Java que facilita el trabajo entre la base de datos relacional y pentaho.

**Sitio Web:** <http://hibernate.org>

**Jetty**

**Descripción:** Servidor web que provee la plataforma para la consola de Administración.

**Puerto:** 8099 (Por Defecto).

**Acceso por URL:** <http://localhost:8099/>

**Sitio Web:** <http://www.ajetty-server.com/>

**MySQL**

**Descripción:** Servidor de base de datos que utiliza quartz, hibernate y pentaho.

**Puerto:** 3306 (Por defecto)

**URL de acceso por JAVA:** `jdbc:mysql://localhost/nombre_de_base_de_datos`

**Sitio Web:** <http://www.mysql.com/>

**Quartz**

**Descripción:** La herramienta para tareas de reportes que utiliza Pentaho.

**Sitio Web:** <http://www.opensymphony.com/quartz/>

**4.1.1.7. Instalación****Paso 1**

Inserte el cd con viñeta “Data Warehouse DC” en la unidad de CD-ROM, en este momento se iniciara el autoejecutable con el que se instalara Pentaho BI.

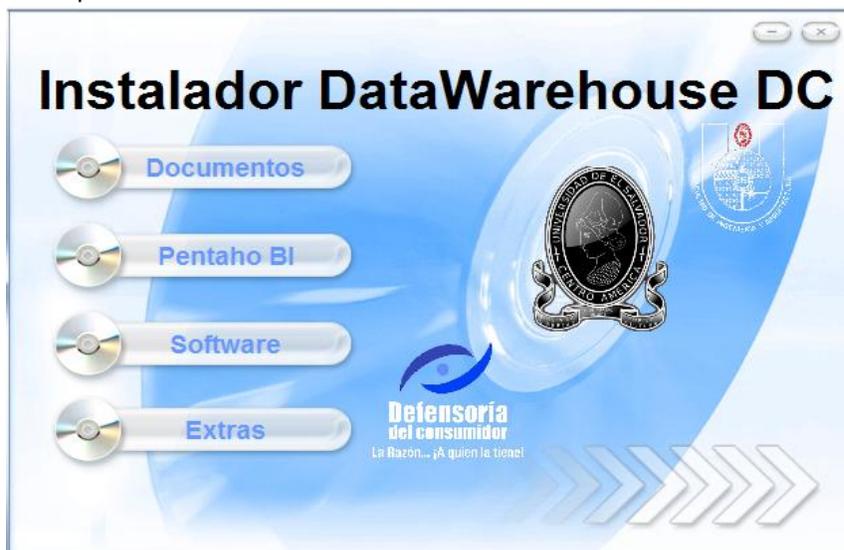
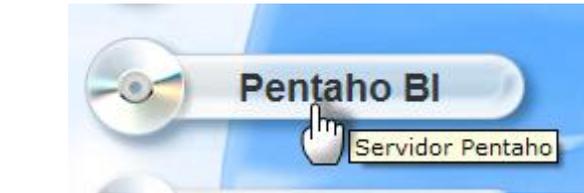


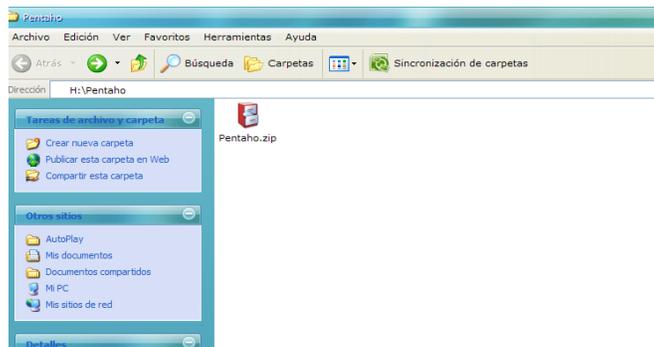
Figura 4.1: Autoejecutable del CD de Instalación

Una vez Iniciado el ejecutable, seleccionar Pentaho BI haciendo clic sobre la imagen.



**Figura 4.2: Abrir carpeta contenedora del servidor BI**

El explorador de Windows se abrirá mostrando el contenido de la carpeta en la que se almacena el servidor Pentaho BI.

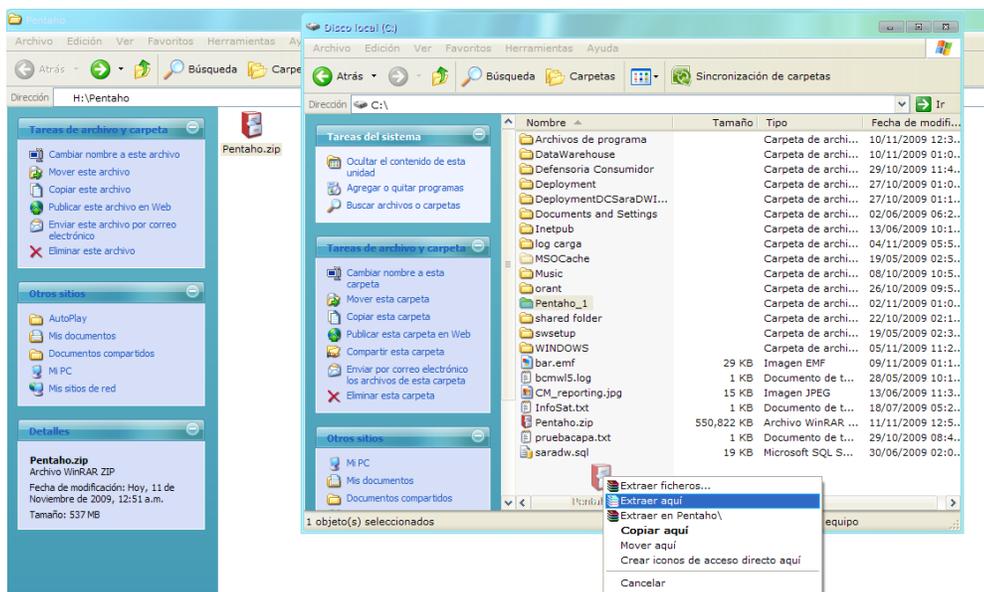


**Figura 4.3 Carpeta contenedora del servidor BI**

El contenido de esta carpeta es un archivo con el nombre PENTAHO.ZIP, el cual contiene toda la configuración del servidor de pentaho, así como los esquemas de los cubos desarrollados.

La instalación de Pentaho a partir de este momento es sencilla, simplemente haciendo uso de un programa para descomprimir archivos se deberá extraer el contenido en la carpeta que se desea almacenar el servidor, en el caso se extraerá en la raíz del disco duro local C:

**Paso 2**



**Figura 4.4: Extracción de Pentaho BI**

Para el ejemplo anterior haciendo uso de WINRAR, simplemente arrastramos el archivo PENTAHO.zip haciendo clic derecho sobre el mismo y desplazándolo sobre un espacio libre en el disco local C:, aparecerá un menú contextual y seleccionamos la opción Extraer aquí...

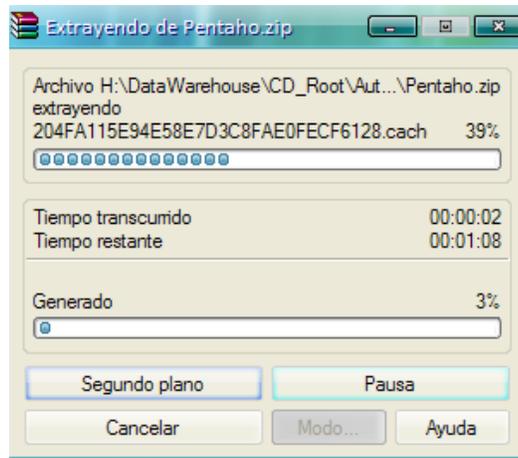


Figura 4.5: Extracción de Pentaho BI utilizando WINRAR

Una vez extraído todos los ficheros tendremos una carpeta llamada pentaho que contiene los elementos más importantes de la Suite de Pentaho BI como son Biserccer-ce, schema-workbench y Metadata Editor.

**Paso 3**

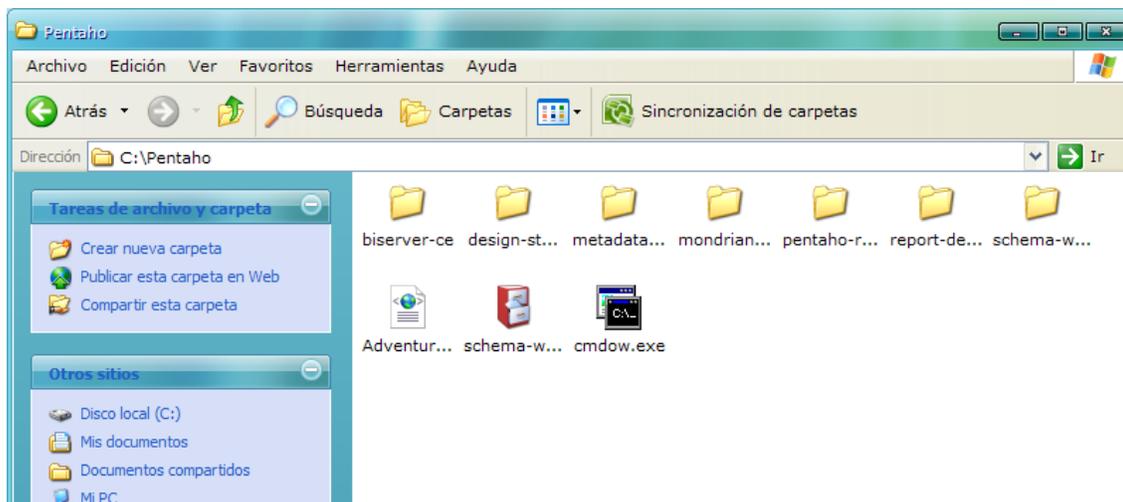


Figura 4.6: Contenido de Pentaho Bi Server

Con esta simple extracción el servidor Pentaho se encuentra instalado en el equipo.

**4.1.1.8. Carga de Base de datos del data warehouse**

En el menú del cd seguir los siguientes pasos.

**Paso 1**

Elegir la opción extras.

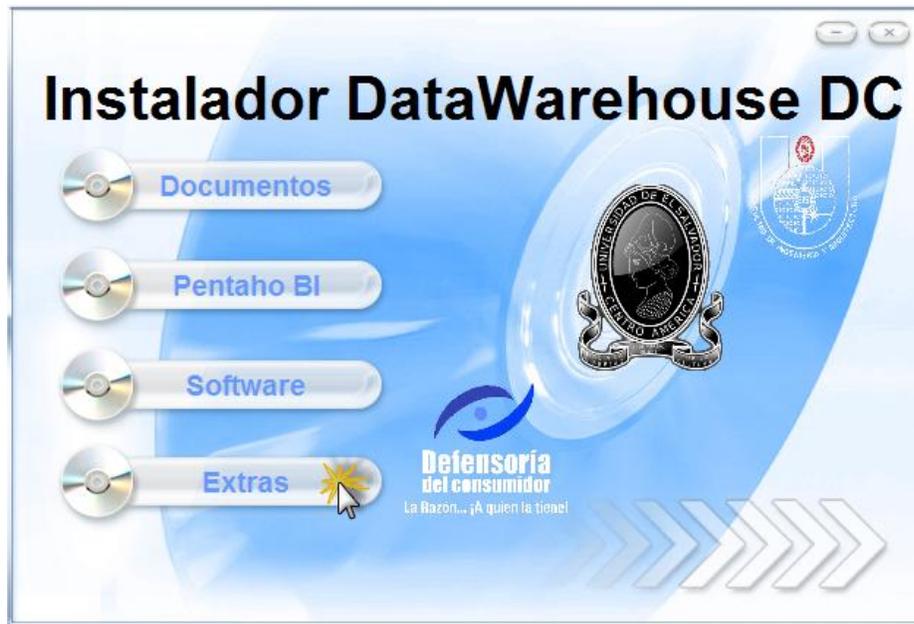


Figura 4.7: Eligiendo la opción extra

**Paso 2**

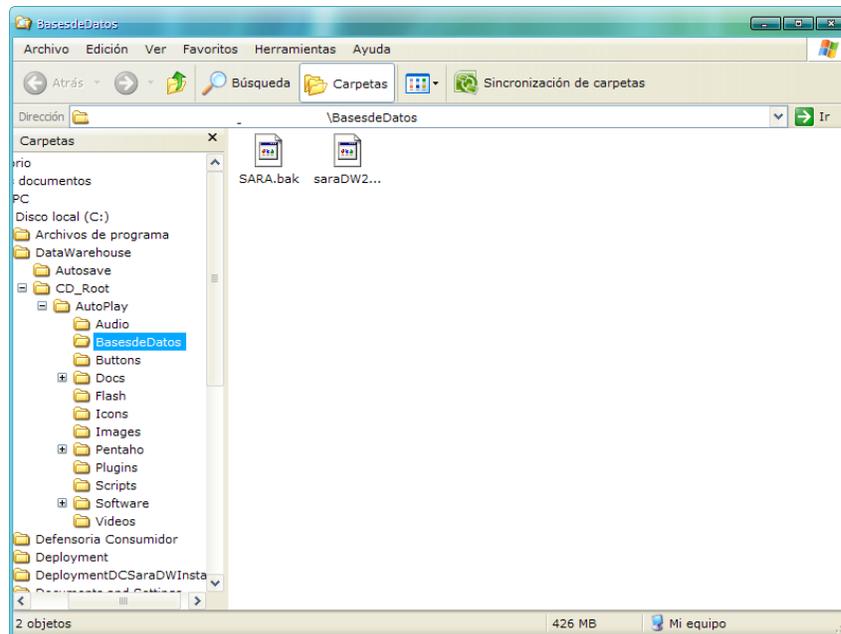
Elegir la opción Respaldo de bases



Figura 4.8: Eligiendo la opción extra

**Paso 3**

Verificar que en el explorador de Windows aparezca los archivo “.Bak” los cuales son los respaldos de las bases de datos transaccionales y del Data Warehouse.



**Figura 4.9: Carpeta contenedora de los respaldos de las bases de datos**

**Paso 4**

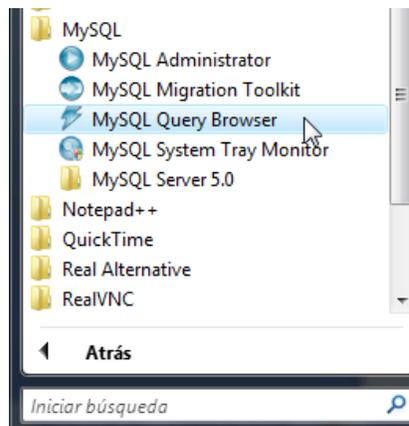
Se procede a la restauración de las bases de datos transaccional (SARA) y de data warehouse (SARADW). Si no se conocen los pasos necesarios para realizar dicha restauración, ver Capítulo 4.1.8.1 (Restaurar una copia de seguridad completa de la base de datos).

**4.1.1.9. Carga de Bases de Datos Iniciales de Pentaho y Scripts de Ejemplo**

Para que Pentaho funcione correctamente se deben ejecutar 5 scripts en MySQL, los cuales se encuentran en el directorio biserver-ce\data\mysql5. Los pasos para ejecutarlos son los siguientes:

**Paso 1**

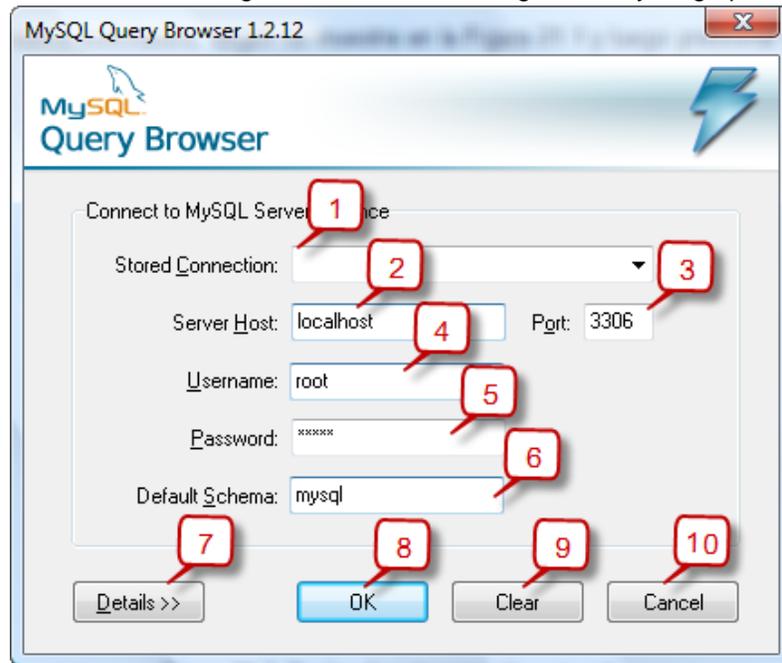
Abrir la herramienta MySQL Query Browser, a la cual se puede acceder desde Inicio -> Todos los Programas -> MySQL -> MySQL Query Browser. (Figura 4.10). (El instalador se encuentra en el cd del proyecto).



**Figura 4.10: Acceso a MySQL Query Browser**

**Paso 2**

Proporcionar los datos solicitados, según se muestra en la Figura 4.11 y luego presionar OK.



**Figura 4.11** Parámetros de conexión a mysql

Número	Nombre	Descripción
1	Stored Connection	Lista despegable que permite seleccionar las configuraciones de conexión almacenadas.
2	Server Host	Servidor o IP de la base de datos.
3	Port	Puerto de conexión a MySQL.
4	Username	Usuario de conexión a la base de datos
5	Password	Contraseña de conexión a MySQL.
6	Default Schema	Esquema por defecto de la sesión a crear.
7	Details	Detalles de la conexión.
8	Ok	Botón para realizar la conexión con los parámetros proporcionados.
9	Clear	Limpiar de la pantalla todos los parámetros ingresados.
10	Cancel	Cancela la operación de conexión y cierra el programa.

**Tabla 4.1:** Parámetros de conexión a MySQL

**Paso 3**

En el Query Browser ir a File -> Open Script (Figura 4.12). En la pantalla de selección del Script ir a la carpeta biserver-ce\data\mysql5 y abrir el archivo create\_repository\_mysql.sql (Figura 4.13). Una vez

abierto hacer clic en el botón  para ejecutar el script.

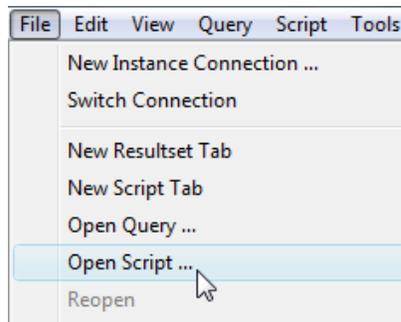


Figura 4.12: Selección de menú para abrir un script

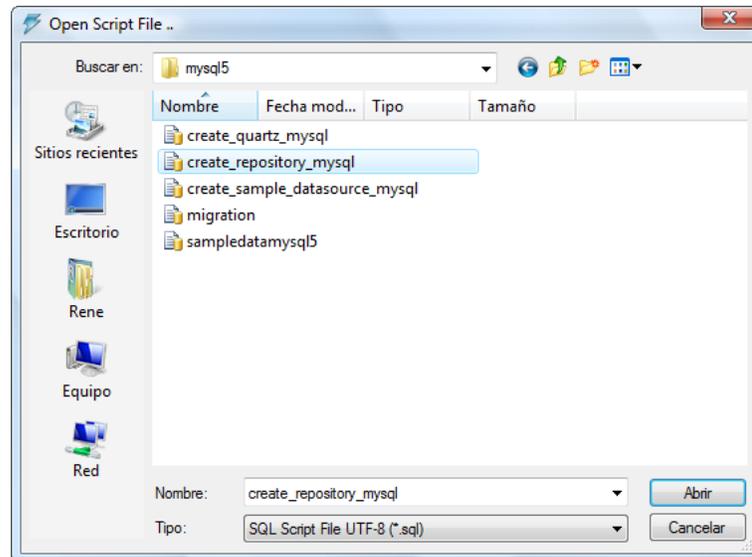


Figura 4.13: Ventana mostrada para abrir script

**Paso 4**

Repetir el paso 3 con los scripts en el siguiente orden.

1. create\_quartz\_mysql.sql
2. create\_repository\_mysql.sql
3. create\_sample\_datasource\_mysql.sql
4. sampledatabmysql5.sql

**4.1.1.10. Configurar conexión a MySQL desde Pentaho**

Para que el servidor BI se pueda conectar a las bases de datos y esquemas de Quartz e hibernate definidos en MySQL, es necesario modificar los parámetros de conexión en distintos archivos xml, para utilizar las credenciales del servidor.

**Paso 1**

El primer archivo a modificar es **applicationContext-acegi-security-jdbc.xml** el cual se encarga de establecer la autenticación por JDBC para el sistema de seguridad Spring Security utilizado en Pentaho BI Server.

Este archivo se encuentra en la ruta `/biserver-ce/pentaho-solutions/system/` y con un editor de texto abrir el archivo **applicationContext-acegi-security-jdbc.xml**. Buscar el siguiente texto:

```
<bean id="dataSource"
class="org.springframework.jdbc.datasource.DriverManagerDataSource">
<property name="driverClassName" value="com.mysql.jdbc.Driver" />
<property name="url" value="jdbc:mysql://localhost:3306/hibernate" />
<property name="username" value="root" />
<property name="password" value="admin" />
</bean>
```

En donde:

Property name	Value	Descripción
url	jdbc:mysql://localhost:3306/hibernate	url jdbc del servidor MySQL
Username	root	Usuario de base de datos con el que se realiza la conexión.
Password	admin	Contraseña del usuario de base de datos.

**Tabla 4.2: Parámetros a modificar para conexión Spring Security**

**Paso 2**

En este paso se procede a modificar el archivo xml que se encarga de permitir la conexión entre Spring security y la base de datos Hibernate.

El archivo a modificar se llama **applicationContext-acegi-security-hibernate.properties** y se encuentra ubicado en: biserver-ce/pentaho-solutions/system. Una vez abierto el archivo con un editor de texto, buscar lo siguiente:

```
jdbc.driver=com.mysql.jdbc.Driver
jdbc.url=jdbc:mysql://localhost:3306/hibernate
jdbc.username=root
jdbc.password=admin
hibernate.dialect=org.hibernate.dialect.MySQLDialect
```

En donde:

Propiedad	Valor	Descripción
jdbc.url	jdbc:mysql://localhost:3306/hibernate	url jdbc del servidor MySQL
jdbc.username	root	Usuario de base de datos con el que se realiza la conexión.
jdbc.password	admin	Contraseña del usuario de base de datos.

**Tabla 4.3: Parámetros de conexión entre Hibernate y Spring Security**

**Paso 3**

En este paso se procede a configurar el archivo de conexión entre Hibernate y MySQL para que se permitan conexiones administrables en el servidor BI.

El archivo a modificar se llama **mysql5.hibernate.cfg.xml** y se encuentra en la ruta: biserver-ce/pentaho-solutions/system/hibernate. Es necesario abrir el archivo xml con un editor de texto y buscar lo siguiente:

```
<property name="connection.driver_class">com.mysql.jdbc.Driver</property>
<property
name="connection.url">jdbc:mysql://localhost:3306/hibernate</property>
<property
name="dialect">org.pentaho.platform.repository.hibernate.MySQL5InnoDBDialect</property>
<property name="connection.username">root</property>
<property name="connection.password">admin</property>
<property name="connection.pool_size">10</property>
<property name="show_sql">>false</property>
<property name="hibernate.jdbc.use_streams_for_binary">>true</property>
```

En donde:

Property name	Valor	Descripción
connection.url	jdbc:mysql://localhost:3306/hibernate	url jdbc del servidor MySQL
connection.username	root	Usuario de base de datos con el que se realiza la conexión.
connection.password	admin	Contraseña del usuario de base de datos.

**Tabla 4.4: Parámetros de conexión entre Hibernate y MySQL**

**Paso 4**

El último archivo a modificar corresponde a la configuración de Hibernate y Quartz en tomcat, para lo cual se debe dirigir a la ruta /biserverce/ tomcat/webapps/pentaho/META-INF y con un editor de texto modificar el archivo: **context.xml**

```
<Context path="/pentaho" docbase="webapps/pentaho/">
  <Resource name="jdbc/Hibernate" auth="Container" type="javax.sql.DataSource"
  factory="org.apache.commons.dbcp.BasicDataSourceFactory" maxActive="20" maxIdle="5"
  maxWait="10000" username="root" password="admin" driverClassName="com.mysql.jdbc.Driver"
  url="jdbc:mysql://localhost/hibernate" validationQuery="select 1" />
  <Resource name="jdbc/Quartz" auth="Container" type="javax.sql.DataSource"
  factory="org.apache.commons.dbcp.BasicDataSourceFactory" maxActive="20" maxIdle="5"
  maxWait="10000" username="root" password="admin" driverClassName="com.mysql.jdbc.Driver"
  url="jdbc:mysql://localhost/quartz" validationQuery="select 1" />
  <Resource name="jdbc/SARADW" auth="Container" type="javax.sql.DataSource"
  factory="org.apache.commons.dbcp.BasicDataSourceFactory" maxActive="20" maxIdle="5"
  maxWait="10000" username="sa" password="grupo14"
  driverClassName="com.microsoft.sqlserver.jdbc.SQLServerDriver"
  url="jdbc:sqlserver://localhost;databaseName=SARADW;" validationQuery="select 1" />
</Context>
```

En donde:

Propiedad	Valor	Descripción
url	jdbc:mysql://localhost:3306/hibernate	url jdbc del servidor MySQL
username	root	Usuario de base de datos con el que se realiza la conexión.
password	admin	Contraseña del usuario de base de datos.

**Tabla 4.5: Parámetros de configuración para Hibernate y Quartz**

**4.1.1.11. Iniciar el Servidor BI y la Consola de Administración**

El paquete del Servidor BI de Pentaho provee los scripts de inicialización de tomcat, la herramienta BI y la consola de administración por medio de archivos bat para sistemas operativos basados en Windows y sh para los basados en Unix.

Para iniciar el servidor BI se deben realizar los siguientes pasos:

**Paso 1**

En la carpeta biserver-ce ejecutar el archivo start-pentaho.bat. Este se encarga de establecer los parámetros de inicialización de Tomcat, así como también de iniciar el servidor web y las aplicaciones como hibernate y quartz.



**Paso 2**

Para determinar si el inicio de Tomcat fue exitoso, abrir la dirección <http://localhost:8080/> en un explorador. Si el arranque de todos los servicios fue correcto, en el explorador se deberá visualizar la pantalla de inicio de sesión de Pentaho, como se muestra en la Figura 4.14.



**Figura 4.14: Pantalla de inicio de sesión de Pentaho**

**Paso 3**

Luego de iniciar pentaho, es necesario iniciar la consola de administración. Ir al directorio `biserver-ce/administration-console`, y ejecutar el archivo `startup.bat` que se encuentra en el directorio.



**Paso 4**

En un explorador ir a la dirección <http://localhost:8099/> y se pedirán los datos de autenticación para ingresar a la consola, tal como se muestra en la Figura 4.15

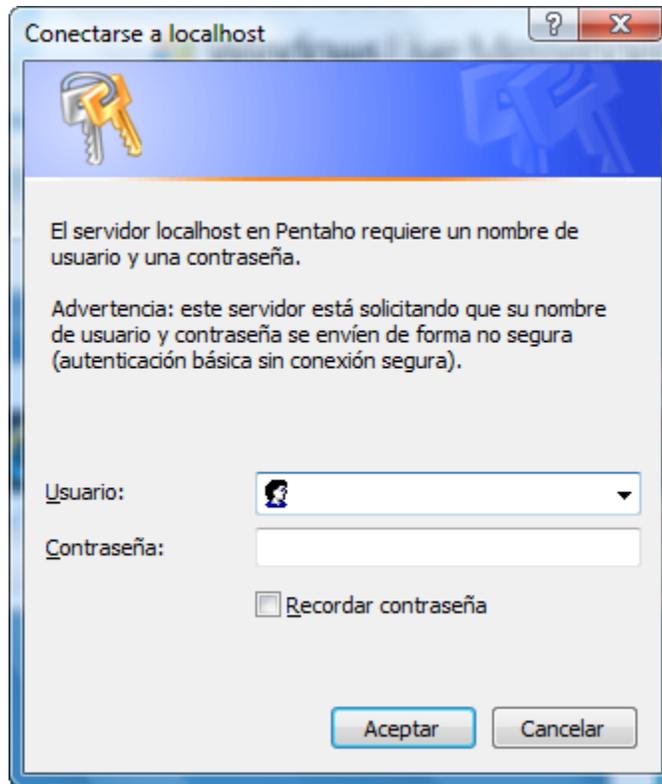


Figura 4.15: Pantalla de autenticación de usuario para la Consola de Administración

**Paso 5**

Una vez se ha realizado la autenticación se muestra la consola de administración.

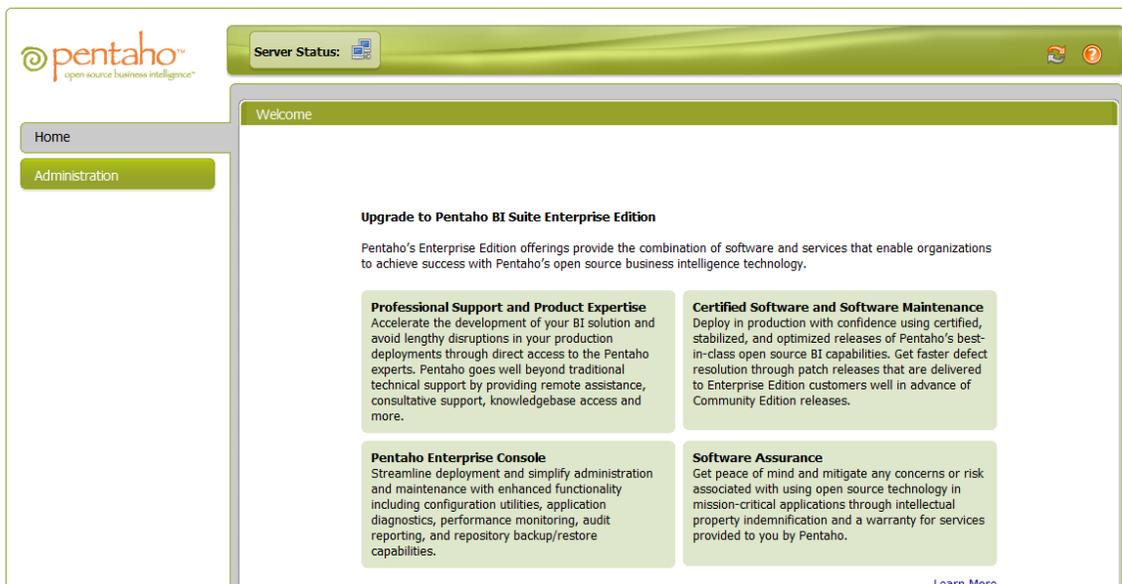


Figura 4.16: Pantalla inicial de la Consola de Administración

#### 4.1.1.12. Crear Conexión a SqlServer

##### Paso 1

Pentaho utiliza JDBC para realizar la conexión a los diferentes DBMS. Para SqlServer es necesario copiar el archivo sqljdbc.jar en los siguientes directorios:

- biserver-ce\administration-console\jdbc
- biserver-ce\administration-console\lib
- biserver-ce\data\lib
- biserver-ce\jre\lib
- biserver-ce\tomcat\common\lib
- biserver-ce\tomcat\server\lib
- biserver-ce\tomcat\shared\lib
- biserver-ce\tomcat\webapps\pentaho\WEB-INF\lib
- biserver-ce\tomcat\webapps\ROOT\WEB-INF\lib
- 

##### Paso 2

Una vez copiado el driver JDBC, ingresar a la consola de administración desde un explorador de internet, utilizando la dirección <http://localhost:8099/> y autenticarse.

##### Paso 3

Una vez Autenticado hacer clic en la pestaña “Administración” y luego en “Data sources”, como se muestra en la Figura 4.17.

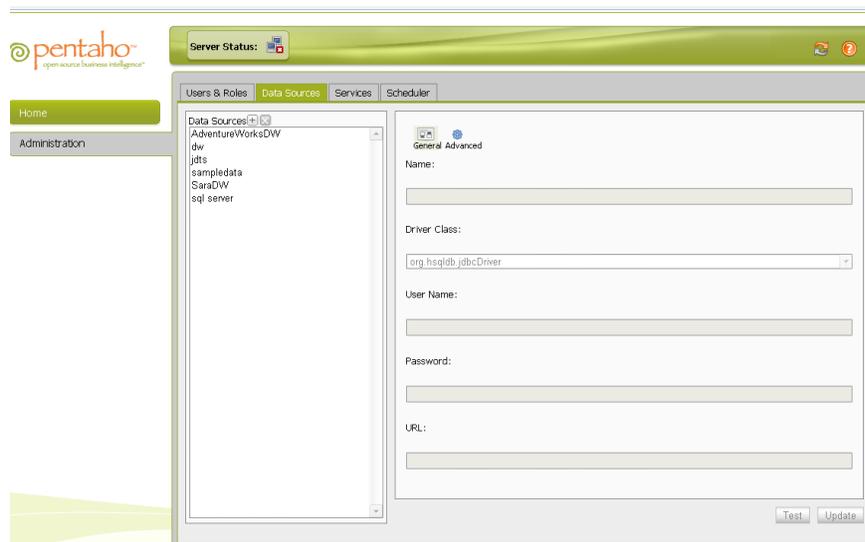


Figura 4.17: Pantalla administración

##### Paso 4

Una vez dentro de los “Data Sources” hacer clic en el botón  para agregar un nuevo origen de datos, ingresando los datos solicitados por medio de la pantalla que se muestra en la Figura 4.18.

Para agregar el data source que se conectara al data warehouse, es necesario tener habilitadas las conexiones para el usuario sa de sql server (para habilitarlo ver capítulo 4.1.5) y permitir las conexiones remotas (para habilitar las conexiones remotas en sql server 2005 ver capítulo 4.1.8)

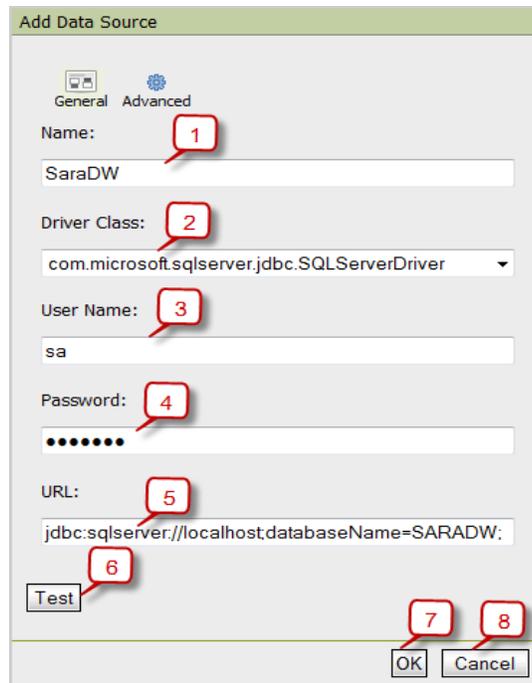


Figura 4.18: Ingreso de parámetros para creación de un nuevo Data Source

En la Tabla 4.6 se describe cada uno de los parámetros que son solicitados para crear un nuevo “Data Source”, así como los datos que deben ser ingresados para conectarse a una base de datos SqlServer 2005.

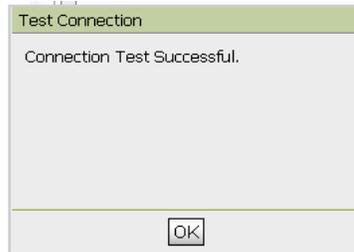
Número	Atributo	Descripción	Datos
1	Name	Nombre que tendrá el Data Source, el cual es utilizado por otras herramientas de la suite. Para este proyecto se debe crear el data source con el nombre <b>SARADW</b> .	SaraDW
2	Driver Class	El driver JDBC que será utilizado para conectarse a la base de datos.	Com.microsoft.sqlserver.jdbc.SQLServerDriver
3	User Name	Usuario que se utilizara para conectarse a la base de datos.	sa
4	Password	Contraseña de acceso para el usuario de la base de datos.	*****
5	URL	Url de conexión del driver jdbc para establecer la conexión con la base de datos.	jdbc:sqlserver://localhost;databaseName=SARADW;
6	Test	Botón que se utiliza para probar la conexión a la base de datos con los parámetros	

		proporcionados.	
7	Ok	Botón para almacenar la conexión.	
8	Cancel	Cerrar la ventana sin guardar la conexión.	

**Tabla 4.6:** Muestra los datos a ingresar para crear una conexión a una base de datos Sql Server

**Paso 5**

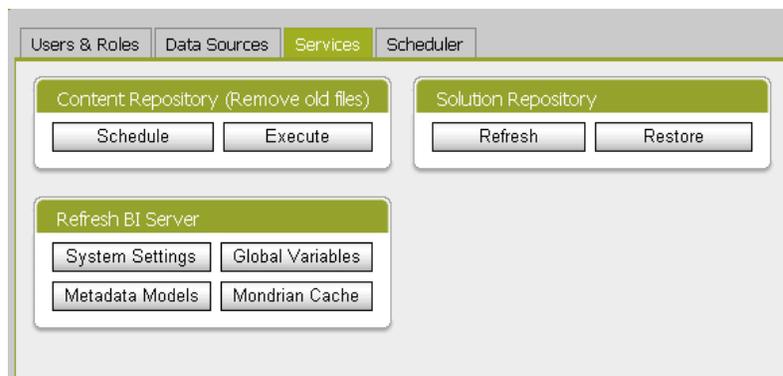
Al presionar el botón Test la Consola de Administración mostrara un mensaje, indicando que la conexión se ha realizado con éxito, como se muestra en la Figura 4.19



**Figura 4.19:** Conexión realizada con éxito

**Paso 6**

En la consola de administración, seleccionar la pestaña **Services**. En el contenido de la pestaña (ver Figura 4.20), en la sección “Refresh BI Server” hacer clic en los botones: System Settings, Global Variables, Metadata Models, Mondrian Cache, los cuales actualizaran todos los repositorios de información del data warehouse.



**Figura 4.20:** Pantalla de opciones en la pestaña Services

**Paso 7**

Acceder al sistema por medio de la dirección <http://localhost:8080/> y seguir los pasos del manual de usuario.

**4.1.2. Instalación y Configuración de Pentaho Metadata Editor**

**4.1.2.1. Descarga de Pentaho Metadata Editor**

Pentaho Metadata Editor (PME) es una herramienta Freeware y Open Source, es parte de la plataforma completa de Inteligencia de negocios Pentaho. Es posible descargar sus archivos fuentes y

ejecutables desde su sitio de descarga. La URL donde se pueden encontrar dichos archivos es: <http://sourceforge.net/projects/pentaho/files/>

Dentro del sitio Source Forge se puede descargar las diferentes versiones de PME. El versionado de las diferentes aplicaciones utiliza las siguientes abreviaciones:

- stable: versión estable de la aplicación.
- RC1: (Release candidate 1) candidata a liberar 1.
- RC2: (Release candidate 2) candidata a liberar 2.

Para nuestro interés utilizaremos la última versión estable de PME, como se observa en la figura 4.21, la versión 3.0.0 es la última versión estable liberada.

- pme-ce-3.0.0-STABLE.zip: Archivos para su instalación en Windows.
- pme-ce-3.0.0-STABLE.tar.gz: Archivos para su instalación en Linux.

File/Folder Name	Platform	Size	Date	Downloads	Notes/Subscribe
Subdirectory ( <a href="#">view all files</a> )					
▼ Pentaho Metadata					 
▶ 1.6.0-stable					 
▶ 1.7.0-stable					 
▶ 1.7.1-stable					 
▶ 2.0.0-stable					 
▼ 3.0.0-stable					 
<a href="#">pme-ce-3.0.0-STABLE-javadoc.zip</a>		699.8 KIB	Wed Aug 12 2009 04:55	104	
<a href="#">pme-ce-3.0.0-STABLE-sources.zip</a>		244.3 KIB	Wed Aug 12 2009 04:55	85	
<a href="#">pme-ce-3.0.0-STABLE.tar.gz</a>		41.3 MiB	Wed Aug 12 2009 04:57	193	
<a href="#">pme-ce-3.0.0-STABLE.zip</a>		41.3 MiB	Wed Aug 12 2009 04:59	532	
▶ 3.5.0-RC1					 
▶ 3.5.0-RC2					 

Figura 4.21: Ventana de descarga de Pentaho Bi-Server

El archivo descargado será un archivo comprimado, con extensión ZIP:



#### 4.1.2.2. Instalación y configuración Pentaho Metadata Editor

##### Requerimientos

Requerimientos necesarios para la instalación de PME:

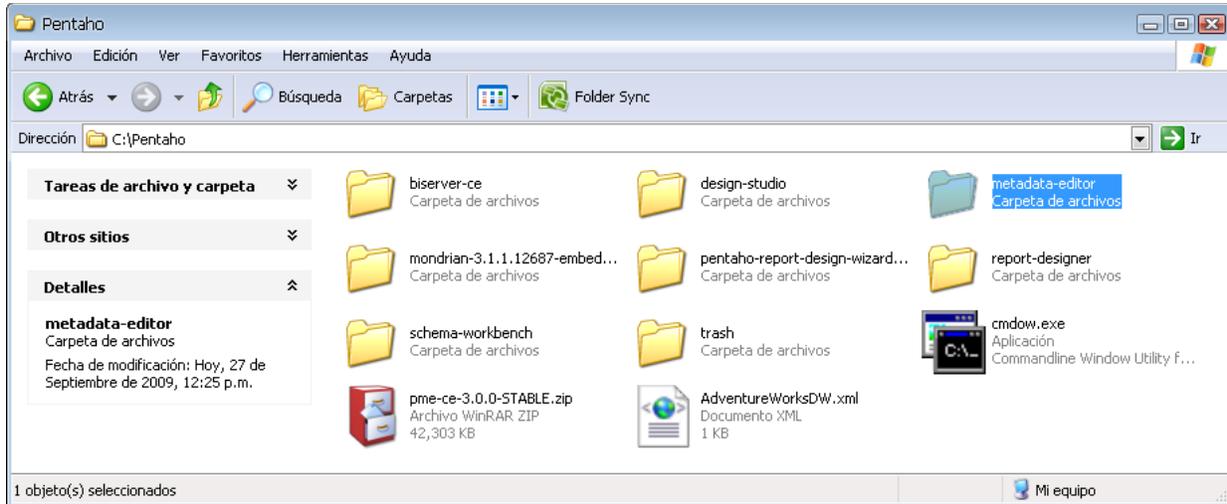
- JRE (Java Runtime Environment), versión 1.5 o superior.
- Pentaho BI Server debe estar disponible, el tener acceso a Pentaho BI Server es únicamente requerido en el caso que se desee configurar la seguridad dentro de un modelo metadata, o si se desea publicar dominio metadata dentro del servidor. De no utilizar lo mencionado, es posible trabajar con todas las demás funcionalidades de Pentaho Metadata Editor sin tener disponible Pentaho BI Server.

- Para realizar conexiones JNDI es necesario tener disponible Microsoft SQL Server JDBC Driver, para su utilización. Puede ser descargado desde : <http://msdn.microsoft.com/en-us/data/aa937724.aspx>

**Instalación**

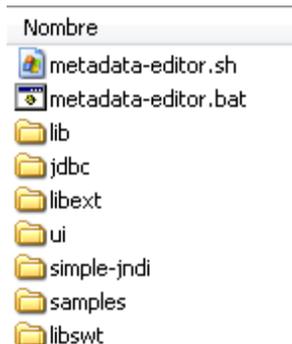
Si la aplicación fue descargada del sitio de Source Forge, deberá ser ubicado en el directorio de Pentaho (R representa la unidad de almacenamiento): R:/Pentaho/

El archivo comprimido deberá ser extraído, creando una carpeta que contendrá PME, como se muestra en la Figura 4.22.



**Figura 4.22: Contenido de la carpeta de Pentaho**

La carpeta metadata-editor contendrá:



Dentro del directorio se encontraran 2 script ejecutables, el que utilizaremos será el archivo por lotes MetaEditor.bat:

- MetaEditor.bat (para Windows)
- Metaeditor.sh (otras plataformas)

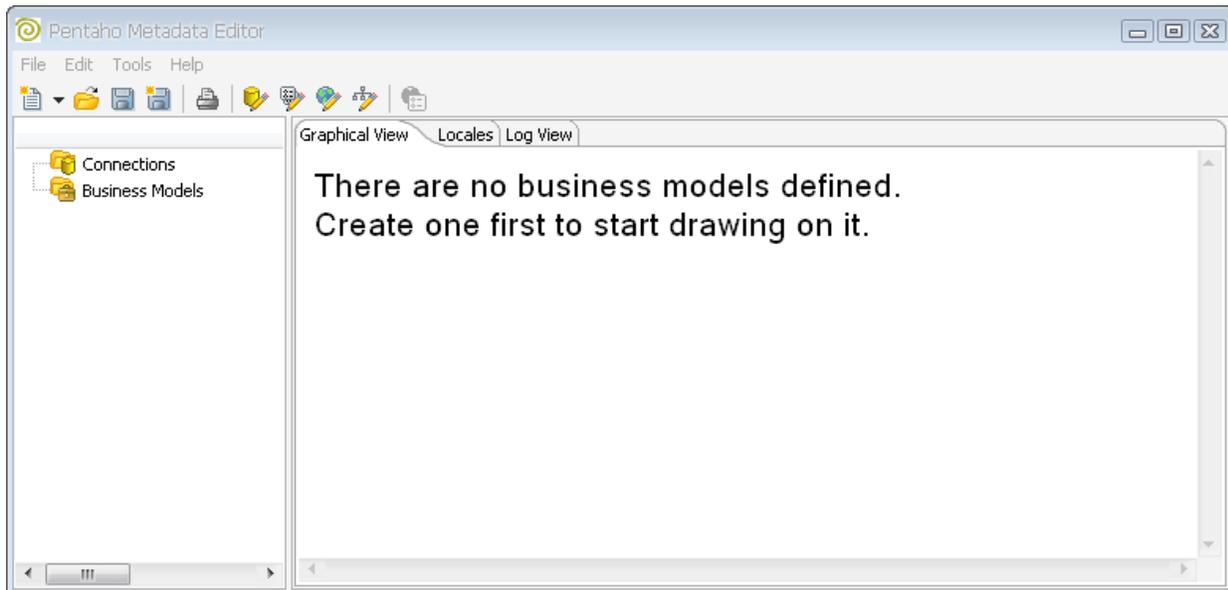
**4.1.2.3. Iniciando Pentaho Metadata Editor**

El inicio de la herramienta lo realiza el archivo por lotes MetaEditor.bat, al hacer doble clic sobre él deberá iniciar la aplicación apareciendo la pantalla de carga como se muestra en la Figura 4.23.



**Figura 4.23: Ventana de Inicio de Pentaho Metadata Editor**

Posterior a la pantalla de carga se presentara la aplicación Pentaho Metadata Editor 3.0.0, Figura 4.24.



**Figura 4.24: Espacio de trabajo de Metadata Editor**

#### 4.1.2.4. Desinstalación Pentaho Metadata Editor

Ya que PME es una aplicación sin instalador, este contiene dentro de la carpeta metadata-editor todos los archivos que utiliza. Para su desinstalación es suficiente el remover dicho directorio: R:/Pentaho/metadara-editor

### 4.1.3. Instalación Y Configuración De Java SE Development Kit (Jdk)

#### Paso 1



Hacer doble clic sobre el icono `jtk-6u14-wind`. En la Figura 4.25 y 4.26 se muestran las pantallas de carga del instalador.

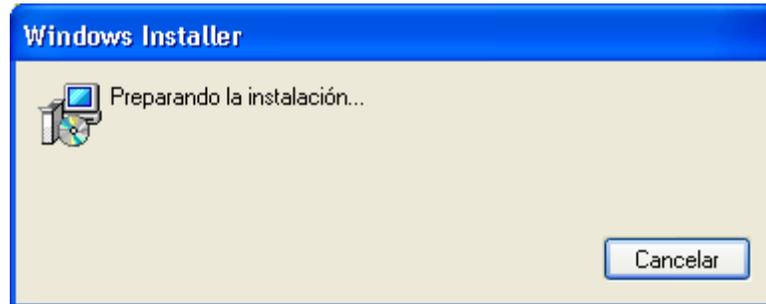


Figura 4.25: Preparación del instalador de Java JDK



Figura 4.26: Pantalla de preparación del asistente de instalación

#### Paso 2

El asistente de instalación solicita aceptar el acuerdo de licencia para proceder con la instalación. Si está de acuerdo con el contrato haga clic en el botón "Accept>", de lo contrario haga clic en el botón "Decline" con lo que la instalación será cancelada.

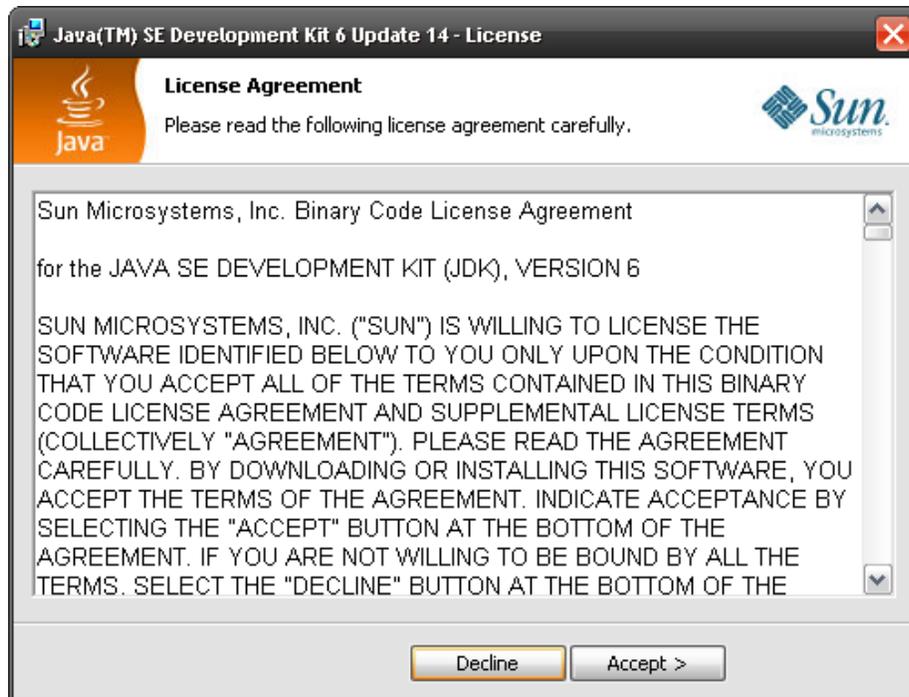


Figura 4.27: Confirmación del acuerdo de contrato de licencia

**Paso 3**

Se deberán seleccionar los componentes a instalar, como mínimo debe estar seleccionada la opción “Development Tools”. Además se puede cambiar la ubicación de instalación del JDK. En la Figura 4.28 se muestra la pantalla con la ruta por defecto que es “C:\Archivos de Programa\Java\jdk1.6.0\_14”, donde los últimos números corresponden a la versión que se esta instalando.

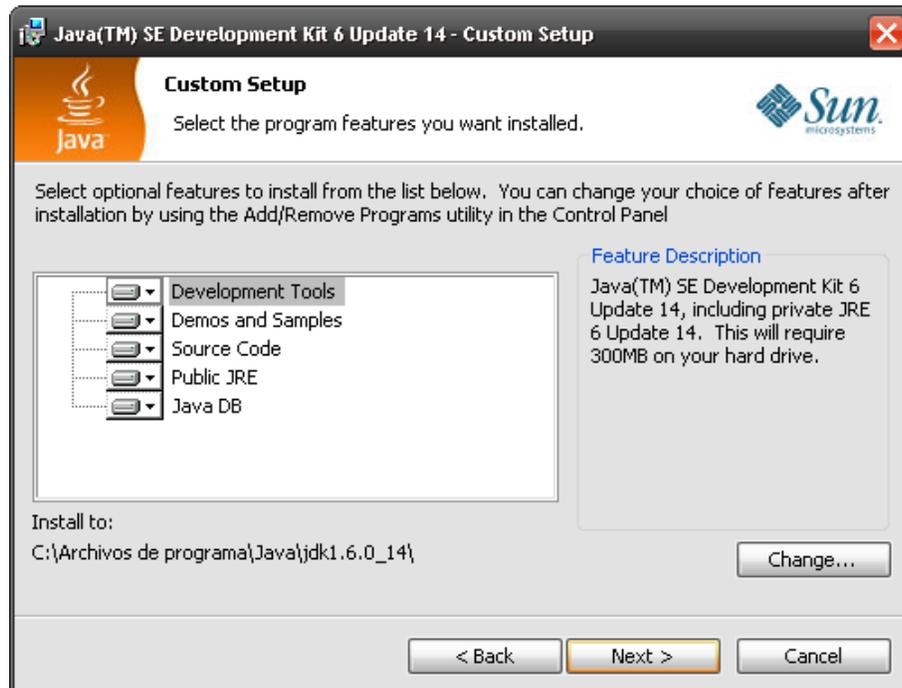


Figura 4.28: Selección de ubicación de instalación y componentes a instalar

Para proceder con la instalación hacer clic en “Next”, con lo que el asistente inicia la copia de los archivos (ver Figura 4.29). Este proceso puede durar varios minutos.

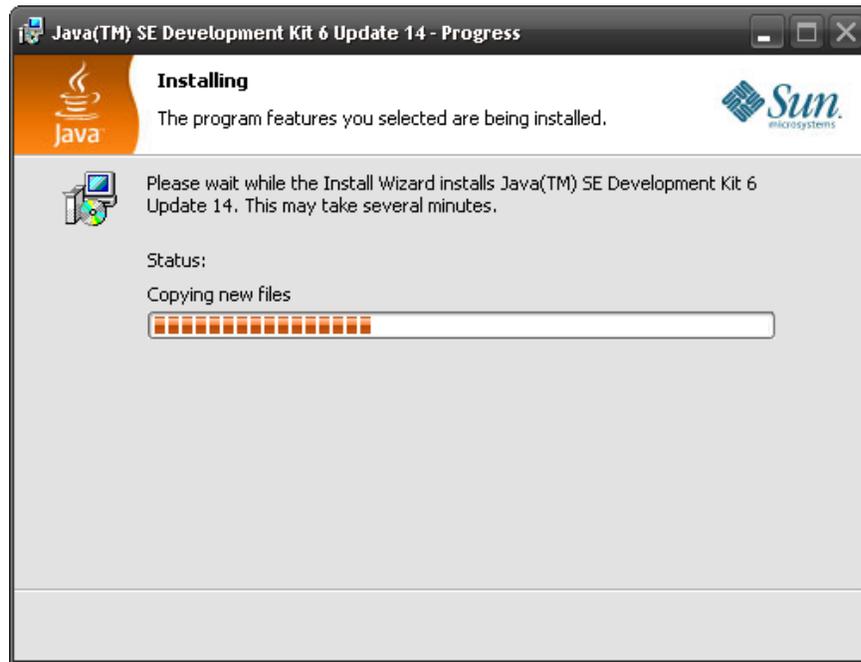


Figura 4.29: Copia de archivos

**Paso 5**

Una vez copiados los archivos el asistente de instalación procede a instalar el JRE6, tal como se muestra en la Figura 4.30. En esta pantalla se puede cambiar la ubicación de la instalación del JRE haciendo clic en el botón “Change”, para proceder con la instalación hacer clic en “Next>”. El asistente procederá a copiar los archivos (ver Figura 4.31).

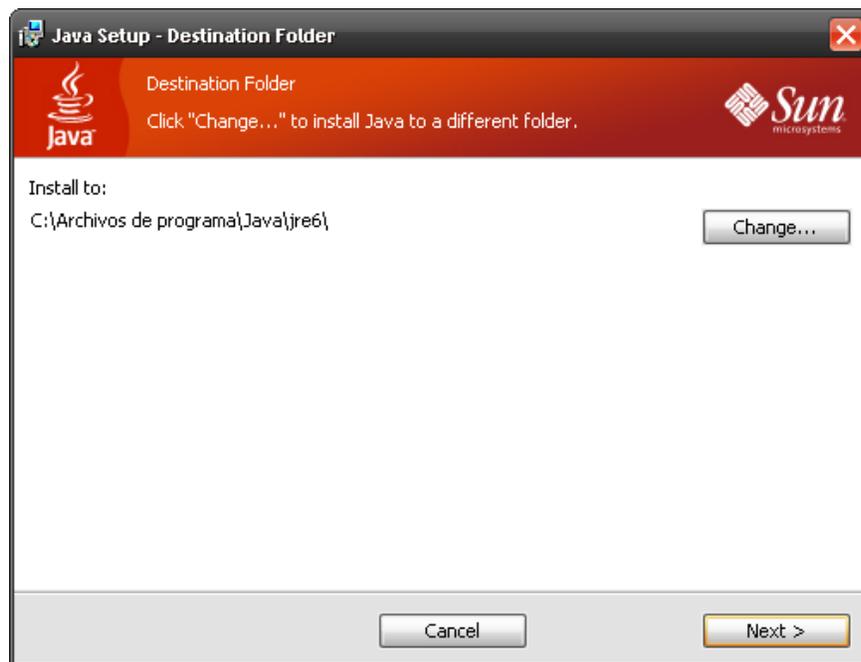


Figura 4.30: Instalación de JRE



Figura 4.31: Copia de archivos de JRE al disco local

**Paso 6**

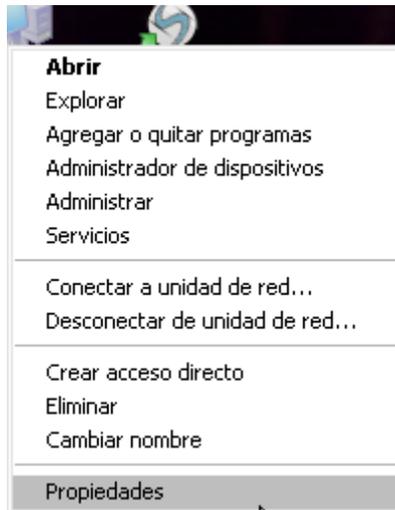
Cuando el asistente ha terminado de copiar los archivos se muestra la pantalla de la Figura 4.32. Para finalizar con la instalación hacer clic en “Finish”



Figura 4.32: Pantalla mostrada al finalizar la instalación

**Paso 7**

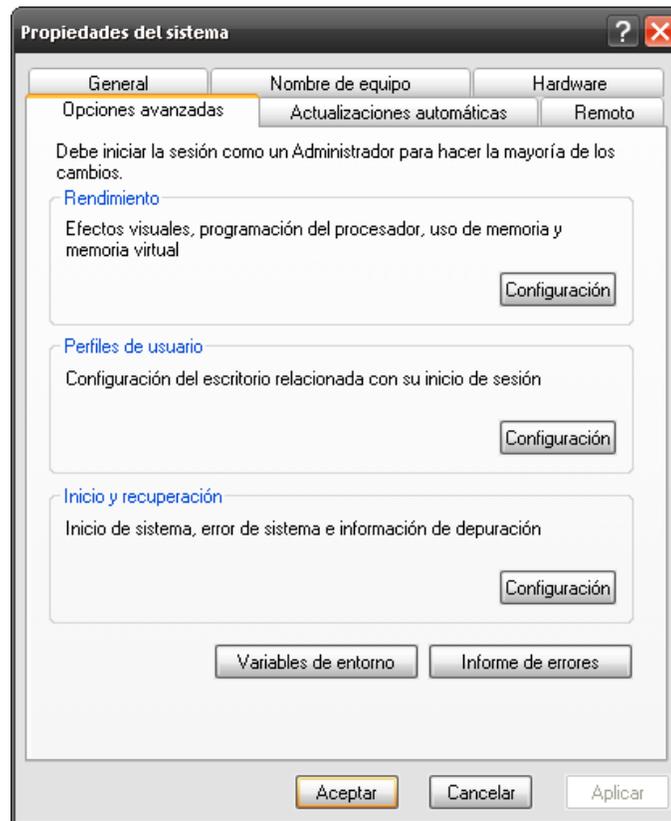
Se procede a configurar la variable de entorno de Java, para lo cual se debe hacer clic derecho en “Mi PC” y luego en “Propiedades” tal como se muestra en la figura 4.33.



**Figura 4.33: Selección de “Propiedades” en menú emergente**

**Paso 8**

En la ventana de propiedades del equipo, seleccionar la pestaña “Opciones Avanzadas” (Figura 4.34).

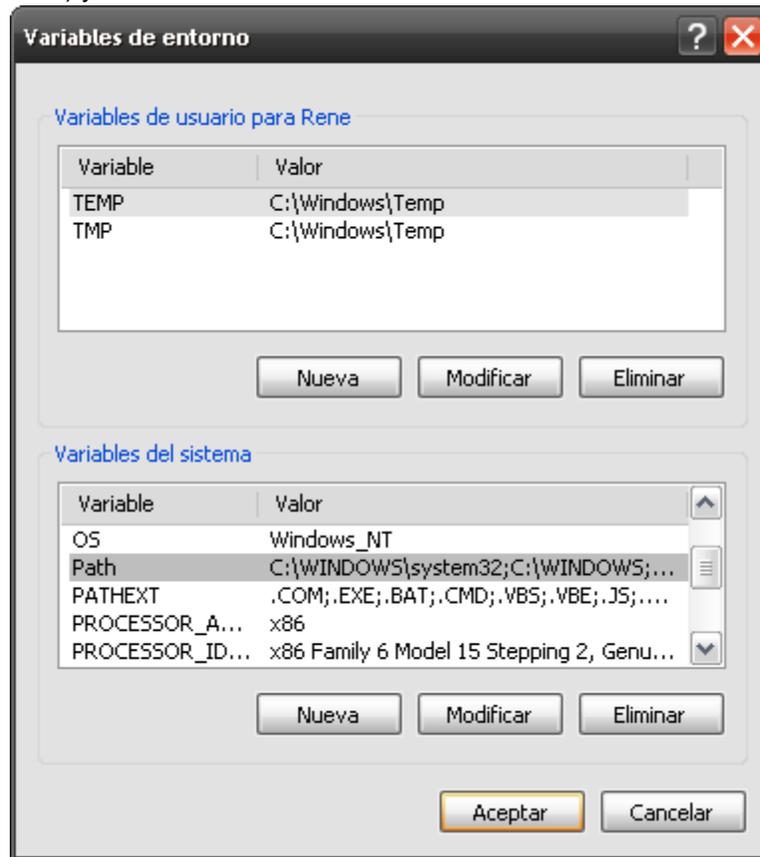


**Figura 4.34: Pantalla de opciones avanzadas de Sistema**

En la ventana, hacer clic en el botón “Variables de Entorno”.

**Paso 9**

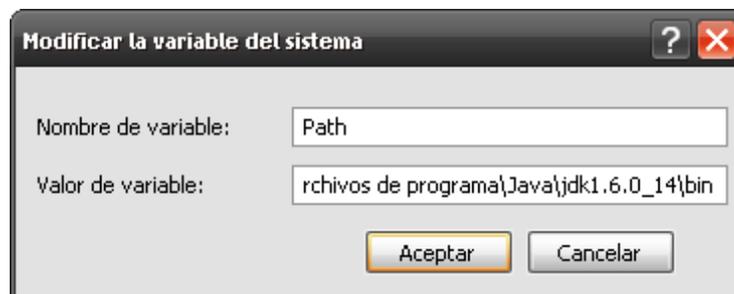
Se mostrará la ventana de la figura 4.35. En el recuadro “Variables de Sistema” seleccionar la variable “Path” (Figura 4.34) y hacer clic en el botón “Modificar”.



**Figura 4.35: Pantalla de configuración de variables de entorno del sistema**

**Paso 10**

Posicionarse en el campo “Valor de la variable” y posicionarse en el final de la línea. Agregar un “;” y luego escribir la ruta de instalación del JDK, para el ejemplo se ha utilizado la ruta por defecto. “C:\Archivos de Programa\Java\jdk1.6.0\_14\bin”. A la ruta de instalación se le debe añadir el directorio “bin”. Luego hacer clic en “Aceptar” (figura 4.36).



**Figura 4.36: Modificación de variable de sistema “Path”**

Una vez realizado este paso ya se tiene configurado correctamente Java JDK.

#### 4.1.4. Instalación de Mysql

##### Paso 1

Ejecutar el archivo Setup.exe, se mostrara la pantalla de la Figura 4.37 mientras se cargan los archivos necesarios para la instalación.

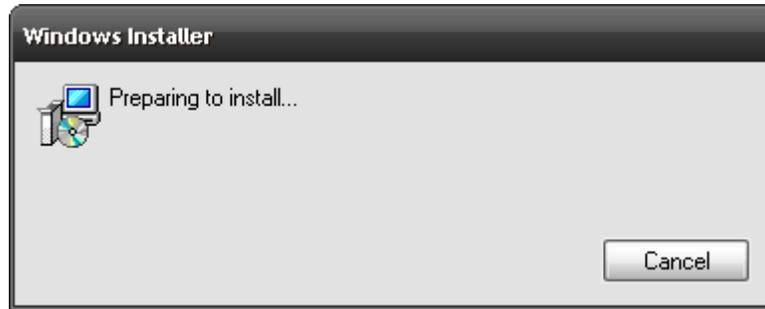


Figura 4.37: Preparación de la instalación

##### Paso 2

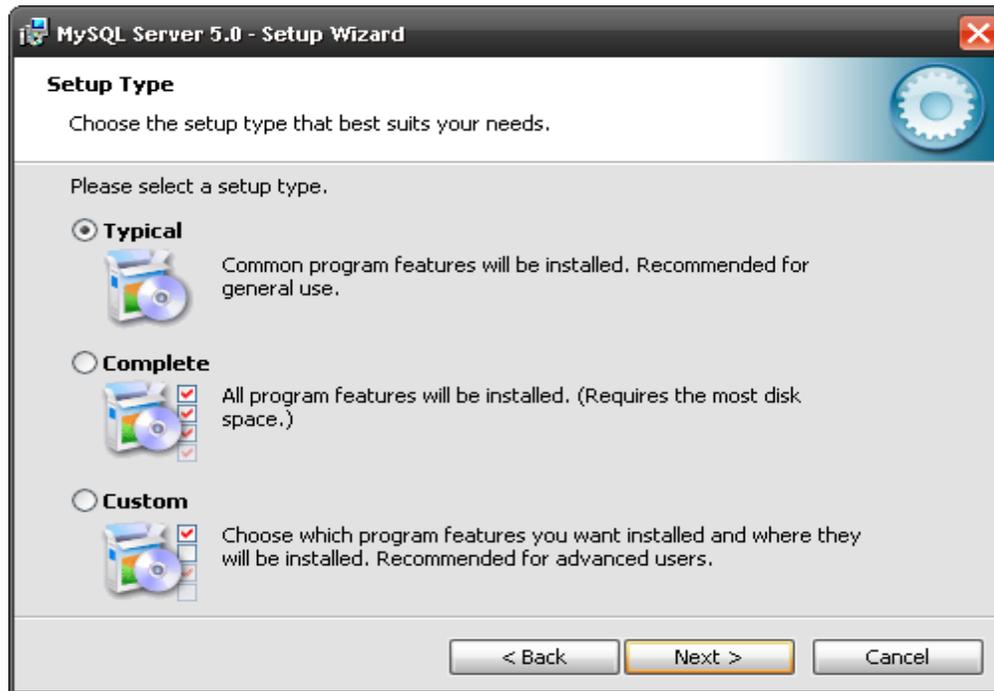
En la pantalla de la figura 4.38 se presiona el botón **NEXT** para proceder a con la instalación.



Figura 4.38: Pantalla de inicio de la instalación

**Paso 3**

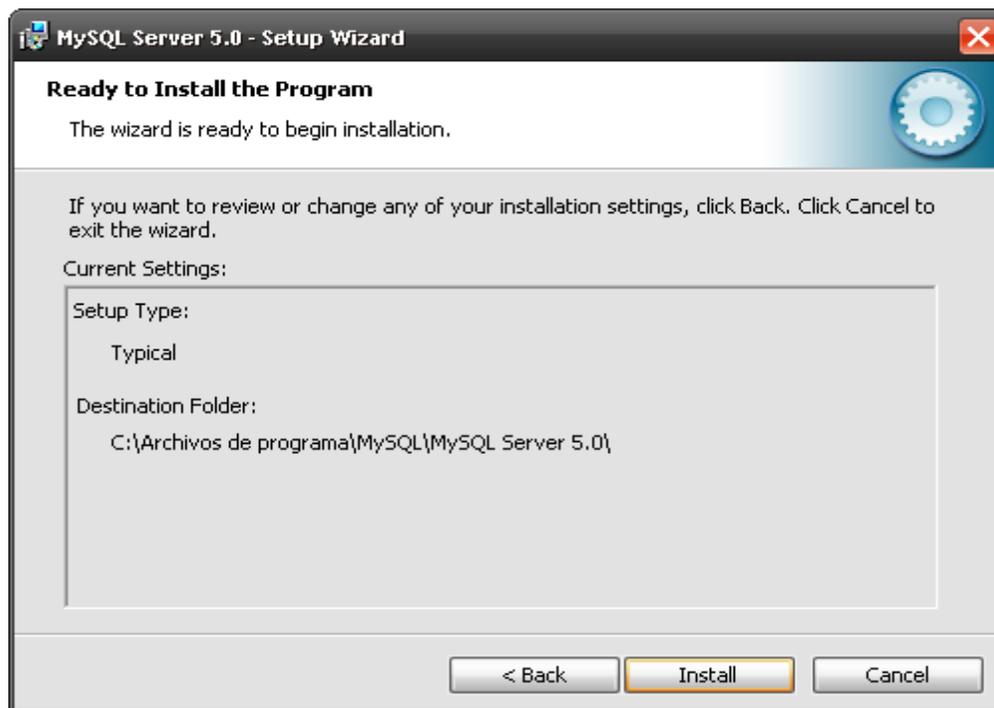
Seleccionar el tipo de instalación que se desea realizar, como se muestra en la Figura 4.39 y luego presionar **Next**.



**Figura 4.39: Tipo de instalación a realizar**

**Paso 4**

En la pantalla de la Figura 4.40 se muestra la ubicación en la cual serán copiados los archivos necesarios para la instalación, para continuar hacer click en **Install**.



**Figura 4.40: Confirmación de la configuración de instalación seleccionada**

**Paso 5**

Se inicia la copia de archivos (Figura 4.41):



Figura 4.41: Inicio de copia de archivos

**Paso 6**

En la pantalla de la Figura 4.42 hacer clic en **Next**:

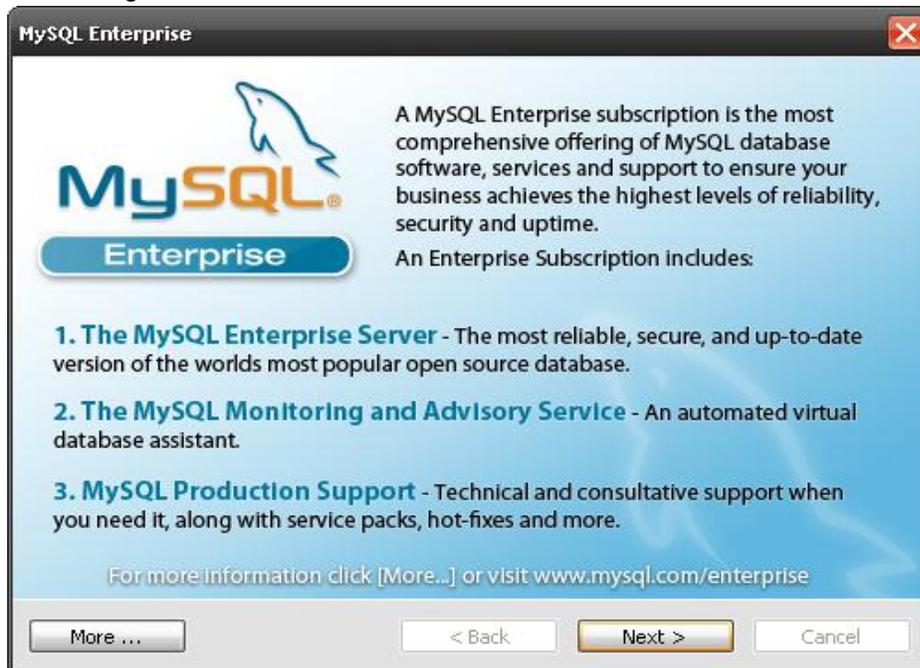


Figura 4.42: Anuncio de MySQL Enterprise

**Paso 7**

Como se muestra en Figura 4.43 presionar **Finish** para continuar con la configuración del servidor MySQL.



Figura 4.43: Finalización de la instalación de MySQL Server 5.0

**Paso 8**

Presionar Next para iniciar el asistente de configuración del servidor MySQL, como se indica en la Figura 4.44.

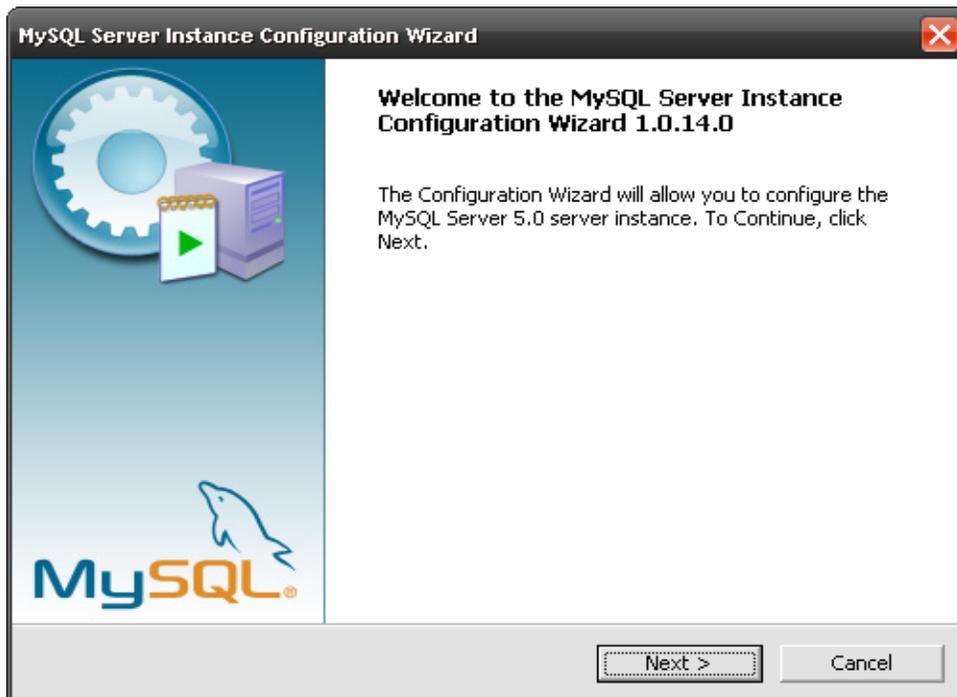


Figura 4.44: Bienvenida al asistente de configuración de MySQL Server Instance

**Paso 9**

Se elige el tipo de configuración detallada para la nueva instancia, se selecciona **Detailed Configuration**; presionar **NEXT**, como se muestra en la figura 4.45



Figura 4.45: Tipo de configuración a realizar

**Paso 10**

En la pantalla de la Figura 4.46 seleccionar **Server Machine** y presionar **Next**



Figura 4.46: Tipo de servidor a usar

**Paso 11**

A continuación como muestra la Figura 4.47, se elige el tipo de base de datos que será usada, se selecciona **Multifunctional Database** y se presionar **NEXT**.



Figura 4.47: Tipo de base de datos a utilizar

**Paso 12**

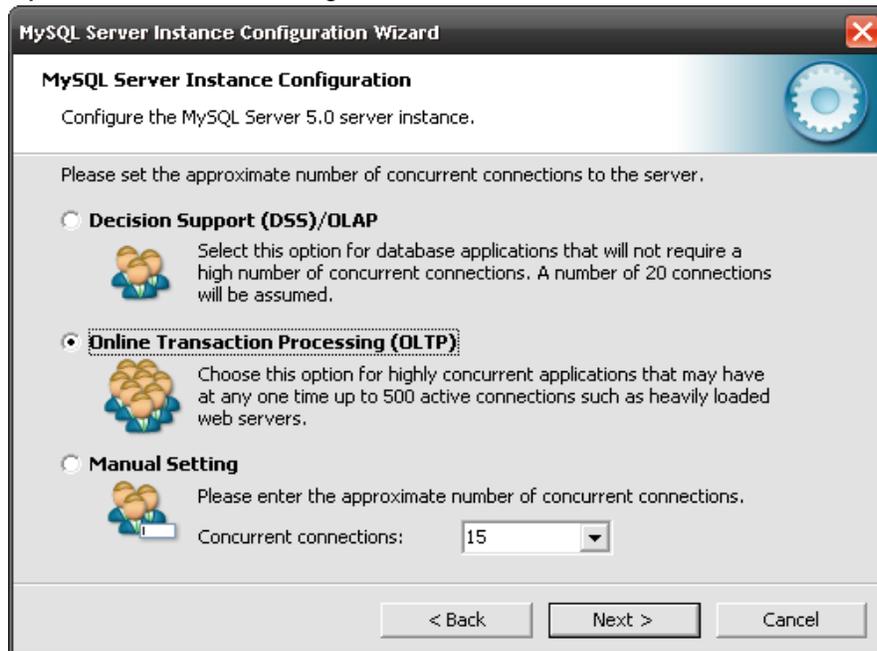
Al igual que la Figura 4.48, seleccionamos los valores por defecto; presionar **NEXT**



Figura 4.48: Ubicación del tablespace

**Paso 13**

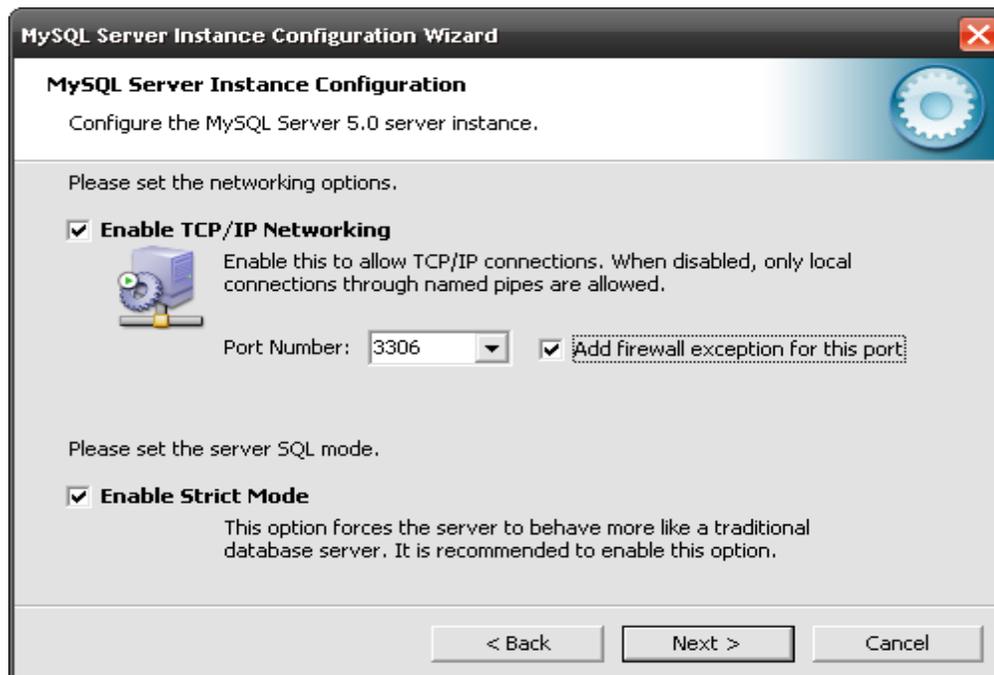
Seleccionar el número concurrente de conexiones al servidor, para este caso **Online Transaction Processing (OLTP)**, como se indica en la Figura 4.49.



**Figura 4.49: Selección del número de conexiones a realizar**

**Paso 14**

Como la pantalla de la Figura 4.50 se presentan las opciones de red, por defecto es TCP/IP y el número de puerto 3306, además se habilita modo estricto y las excepciones de firewall; presionar el botón **Next**



**Figura 4.50: Opciones de red a usar**

**Paso 15**

Se selecciona la opción por defecto de caracteres estándar mostrado en la Figura 4.51:



**Figura 4.51: Conjunto de caracteres a utilizar**

**Paso 16**

Se instala el nuevo servicio y su respectivo nombre presionando **Next**, como se indica en la Figura 4.52



**Figura 4.52: Opciones de instalación del nuevo servicio**

**Paso 17**

La pantalla de la Figura 4.53, muestra las opciones de seguridad, para establecer una contraseña, elegimos la opción **Modify Security Settings** y se escribe la contraseña y su confirmación, en este caso se utilizara la contraseña por defecto con el que están configuradas todas las conexiones a pentaho es "admin".



Figura 4.53: Opciones de seguridad

**Paso 18**

Iniciar el servicio presionando **Execute** como se indica en la Figura 4.54

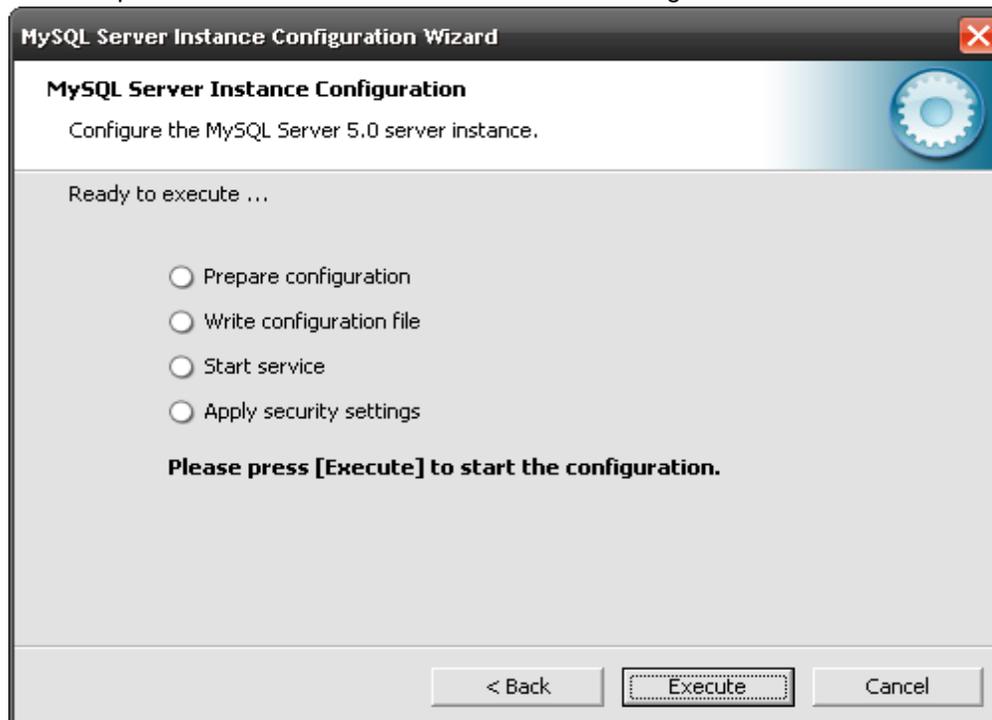
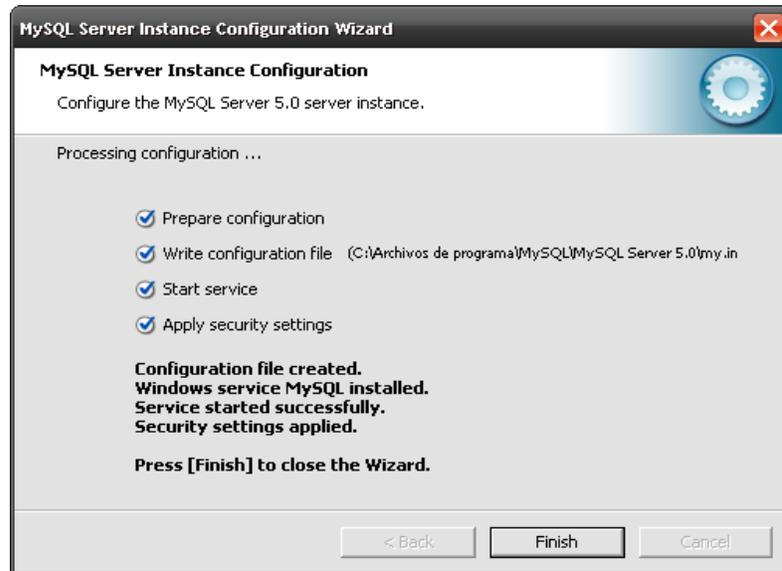


Figura 4.54: Iniciación del servicio

**Paso 19**

Finalizar el asistente de configuración del servidor MySQL, presionando **Finish** como muestra la Figura 4.55



**Figura 4.55: Finalización del asistente de configuración**

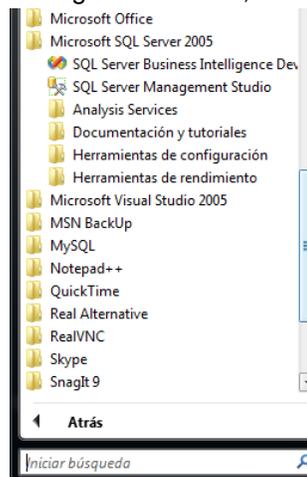
**4.1.5. Configuración de SQL Server**

**4.1.5.1. Habilitar conexiones con usuario sa**

Para poder habilitar las conexiones al servidor de base de datos con el usuario sa, es necesario habilitar el usuario y asignar la contraseña de inicio de sesión.

**Paso 1**

Abrir el programa “Sql Server Management Studio” haciendo clic en Inicio-> Todos los Programas-> Microsoft SQL Server 2005 -> Sql Server Management Studio, como se muestra en la figura 4.56.



**Figura 4.56: Acceso a Sql Server Management Studio**

Luego se mostrara el inicio de sesión al servidor (ver Figura 4.57). Hacer clic en el botón “conectar”



Figura 4.57: Pantalla de inicio de sesión en el servidor de SQL Server

**Paso 2**

El siguiente paso consiste en habilitar el modo de autenticación de servidor en la base de datos, para lo cual se debe hacer clic derecho el servidor **localhost (SQL Server 9.0.1399 - ReneLenovo\Rene)** y en el menú contextual hacer clic sobre “Propiedades” tal como se muestra en la Figura 4.58.

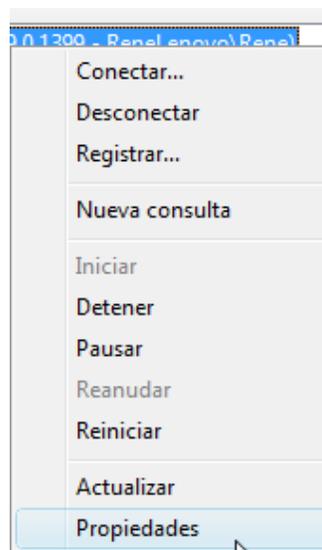


Figura 4.58: Acceso a las propiedades del servidor de base de datos

**Paso 3**

En la ventana seleccionar la pagina Seguridad y en la ficha de autenticación del servidor seleccionar “Modo de autenticación de Windows y SQL Server” como se observa en la Figura 4.59.

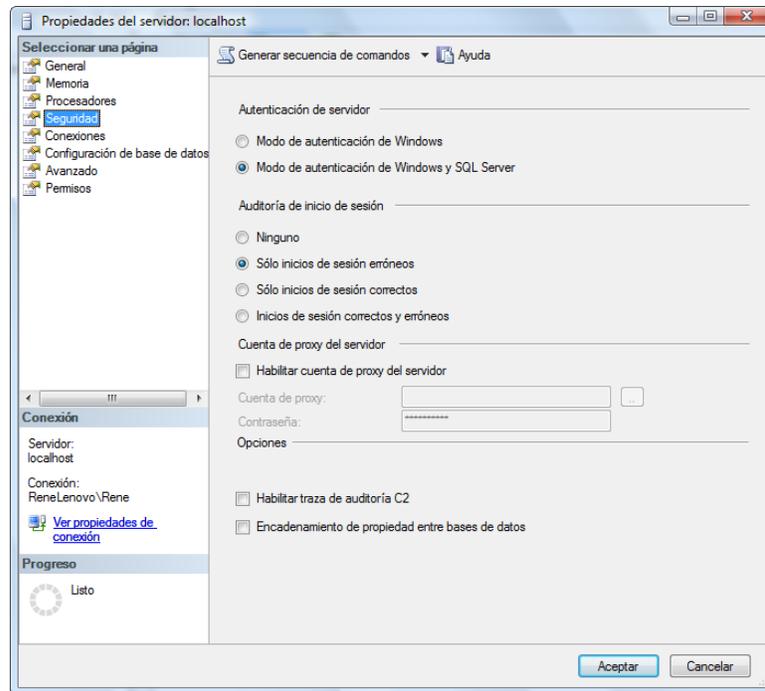


Figura 4.59: Permitir método de autenticación SQL Server

**Paso 4**

En la misma ventana seleccionar la pagina “Conexiones” y seleccionar la opción “Permitir conexiones remotas con este servidor” (ver Figura 4.60) y hacer clic en “Aceptar” y luego es necesario reiniciar el servicio de la base de datos.

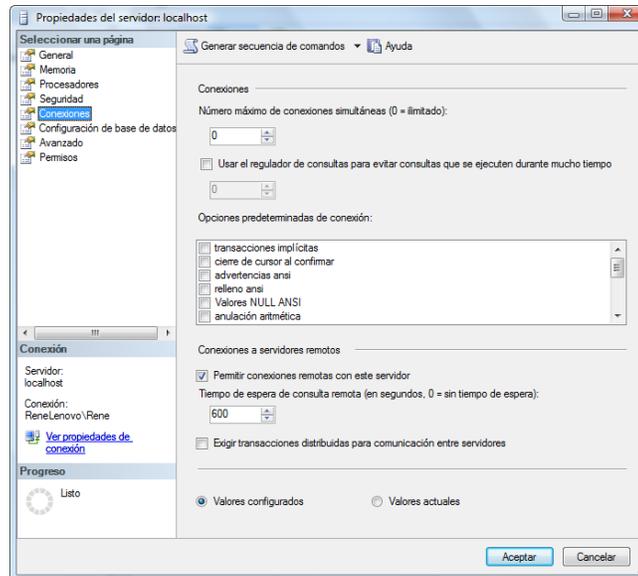


Figura 4.60: Habilitar conexiones remotas hacia el servidor de base de datos

**Paso 5**

El siguiente paso consiste en habilitar la cuenta de usuario sa, para lo cual se debe seleccionar el usuario, como se muestra en la Figura 4.61.

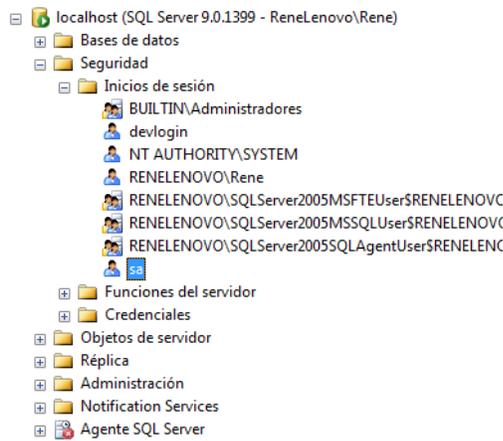


Figura 4.61: Localización del usuario sa

Luego hacer clic derecho sobre el usuario y seleccionar la opción “Propiedades” (Figura 4.62)

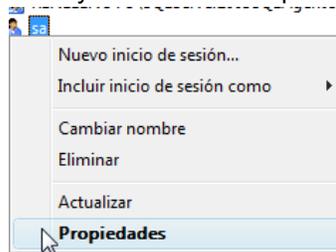


Figura 4.62: Acceso a las propiedades del usuario sa

**Paso 6**

En la ventana de propiedades del usuario sa, seleccionar la pagina “General” y asignar la contraseña con la cual tendrá acceso al servidor (Figura 4.63).

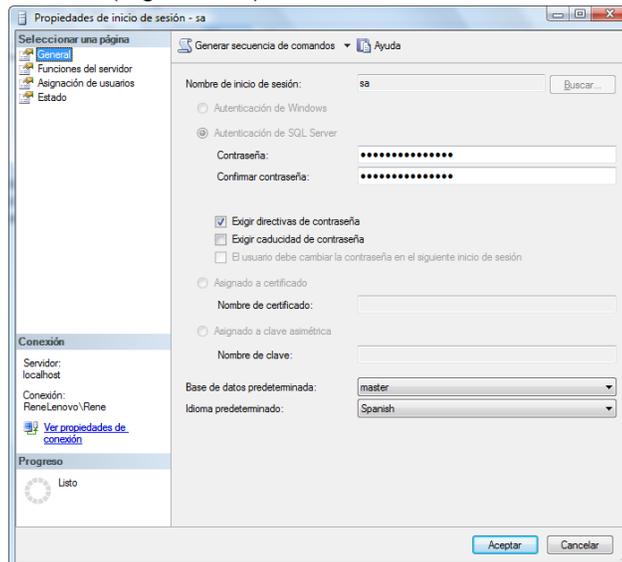


Figura 4.63: Asignación de clave de acceso al servidor

Luego seleccionar la pagina “Estado” y seleccionar las opciones tal como se muestra en la Figura 4.64 y luego hacer clic en aceptar.

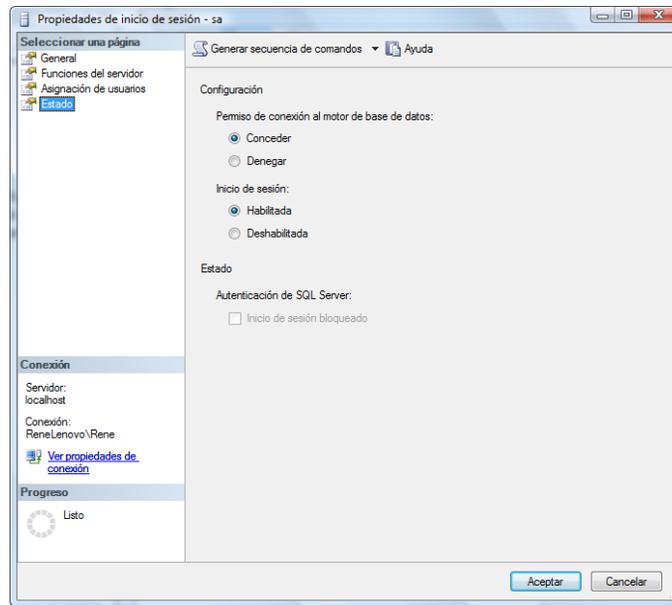


Figura 4.64: Selección de permisos de conexión e inicio de sesión al servidor para el usuario sa

#### 4.1.6. Configuración de Superficie de SQL Server 2005

Superficie de SQL Server 2005 es una herramienta que se instala por defecto con el SQL Server 2005, permite configurar varios aspectos de SQL Server. Básicamente permite:

- Habilitar o deshabilitar servicios que en algún momento no se utilizan de SQL Server 2005, como protocolos de red para conexiones remotas.
- Habilitar o deshabilitar características de los componentes que en algún momento no se utilizan de SQL Server.

#### 4.1.7. Acceder a Configuración de Superficie de SQL Server 2005

Ingresar a Configuración de Superficie de SQL Server 2005, desde inicio de Windows, como se muestra en la Figura 4.65.

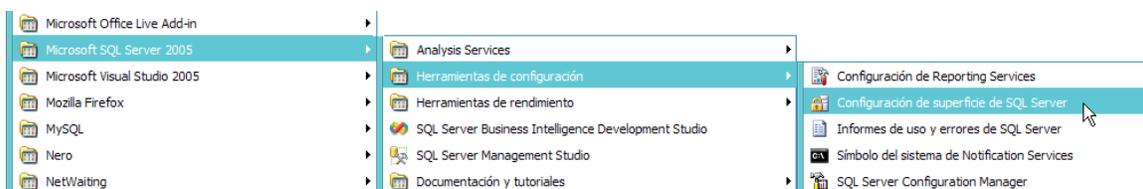


Figura 4.65 Acceder a Configuración de Superficie de SQL Server 2005

#### 4.1.8. Configuración de conexiones locales y remotas

En la Figura 4.66 se presenta la pantalla inicio de Configuración de Superficie de SQL Server 2005. Se selecciona “Configuración de superficie para servicios y conexiones”.



Figura 4.66: Pantalla de inicio Configuración de Superficie de SQL Server 2005

Luego aparecerá la pantalla Configuración de Superficie para servicios y conexiones – localhost, como se muestra en la Figura 4.67. Luego se deberá acceder a Conexiones remotas (MSSQLSERVER>Motor de base de datos>Conexiones remotas).

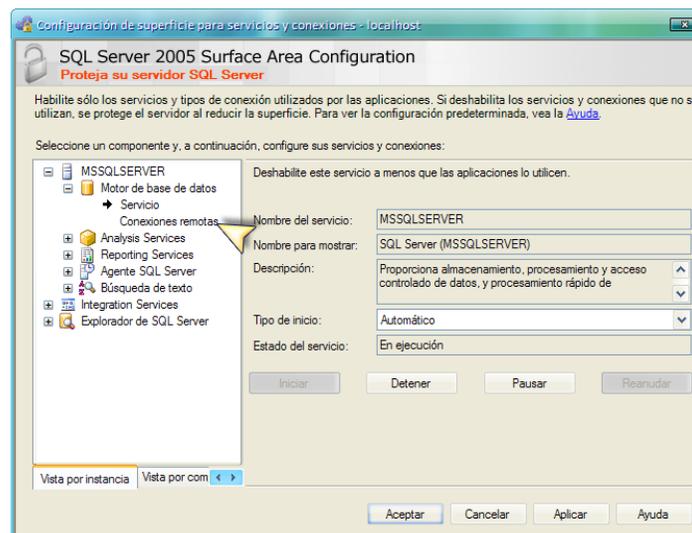
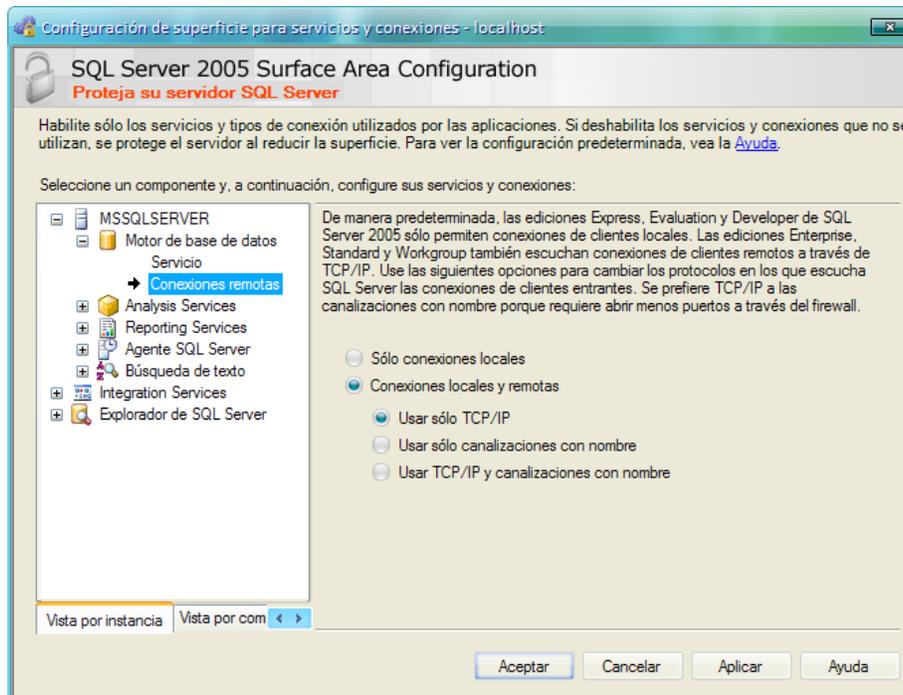


Figura 4.67: Pantalla Configuración de superficie para servicios y conexiones - localhost

Dentro de Conexiones remotas, seleccionaremos dentro de “**Conexiones locales y remotas**”, la opción “**Usar solo TCP/IP**”, como se muestra en la Figura 4.68. Y para finalizar seleccionamos aceptar. Luego es necesario reiniciar el servicio de la base de datos.



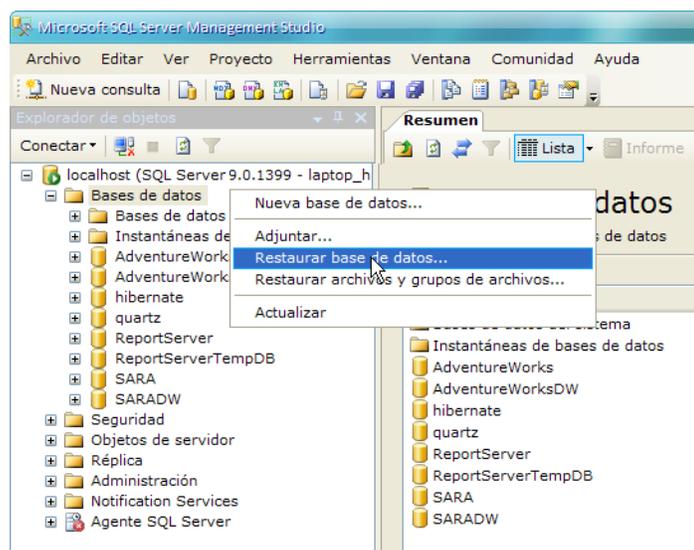
**Figura 4.68: Configurando Conexiones locales y remotas**

Este tipo de configuración permitirá el acceso al servidor SQL Server 2005, únicamente mediante conexiones TCP/IP. Esta modificación se realiza debido a que las versiones de SQL Server Edición Estándar, viene configurado predeterminadamente con la opción de “Solo conexiones locales”.

#### 4.1.8.1. Restaurar una copia de seguridad completa de la base de datos

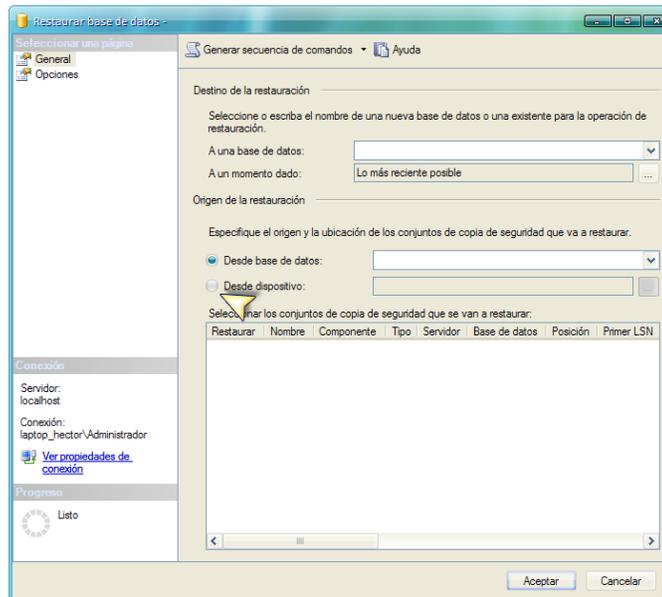
Después de conectarse a la instancia adecuada de SQL Server Database Engine (Motor de base de datos de SQL Server) de Microsoft, en el Explorador de objetos.

**Paso 1:** Hacer clic en el nombre del servidor y se mostrara un menú emergente tal como se muestra en la siguiente figura:



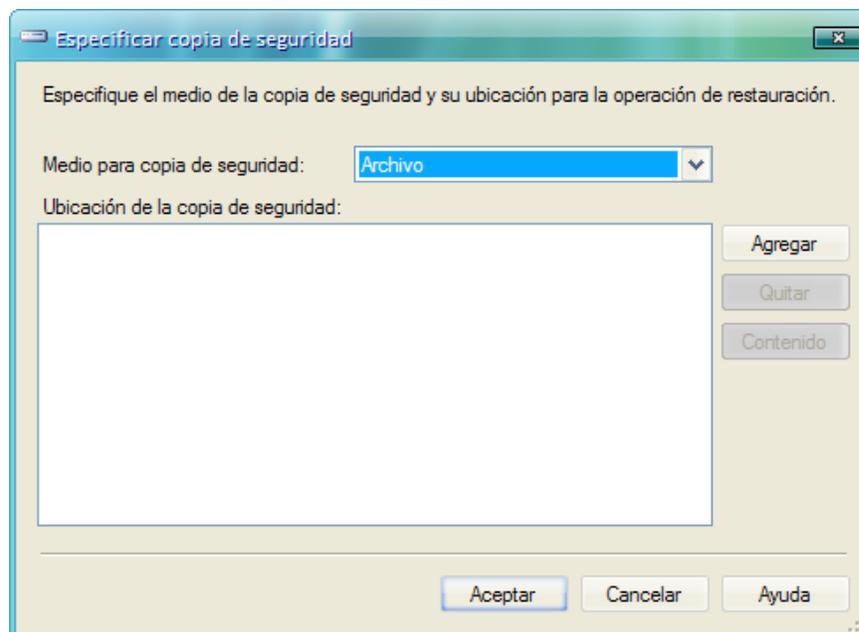
**Figura 4.69: Opción para restaurar base de datos**

**Paso 2:** Elegir en nuestro origen de la restauración la opción Desde Dispositivo, tal como señala el indicador amarillo de la figura siguiente:



**Figura 4.70: Ventana de opciones para restaurar una base de datos**

**Paso 3:** Al estar situado en esa opción hacer clic sobre el botón con tres puntos que aparece en la parte derecha de la opción. Y se presentara un cuadro de dialogo como el siguiente:



**Figura 4.71: Cuadro de dialogo para especificar la ruta del respaldo de la base**

**Paso 4:** Para continuar debemos elegir la ubicación del archivo a restaurar. Para ellos se debe hacer clic en el botón Agregar y se presentara otro cuadro de dialogo como el siguiente. Para finalizar hacer clic en aceptar

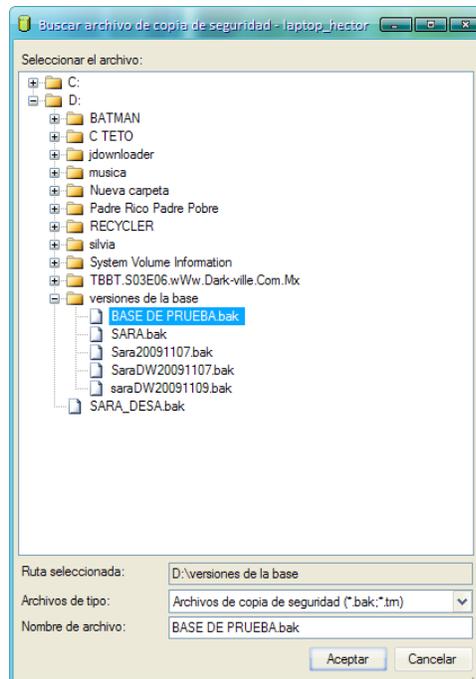


Figura 4.72: Cuadro de dialogo elegir la ubicación de la base de datos

**Paso 5:** Hacer clic de nuevo para confirmar el archivo a restaurar. Para finalizar hacer clic en aceptar.

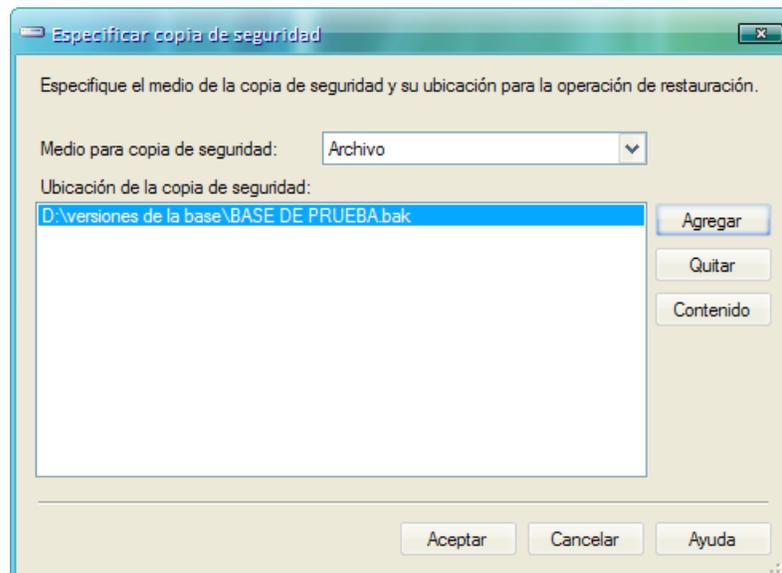


Figura 4.73: Cuadro de dialogo con la ubicación fijada

**Paso 6:** En caso de no conocer el nombre de la base se puede elegir haciendo clic sobre la lista desplegable que aparece junto a la opción “A una base de datos” y erigiremos el nombre de la base de datos que se restaurará.

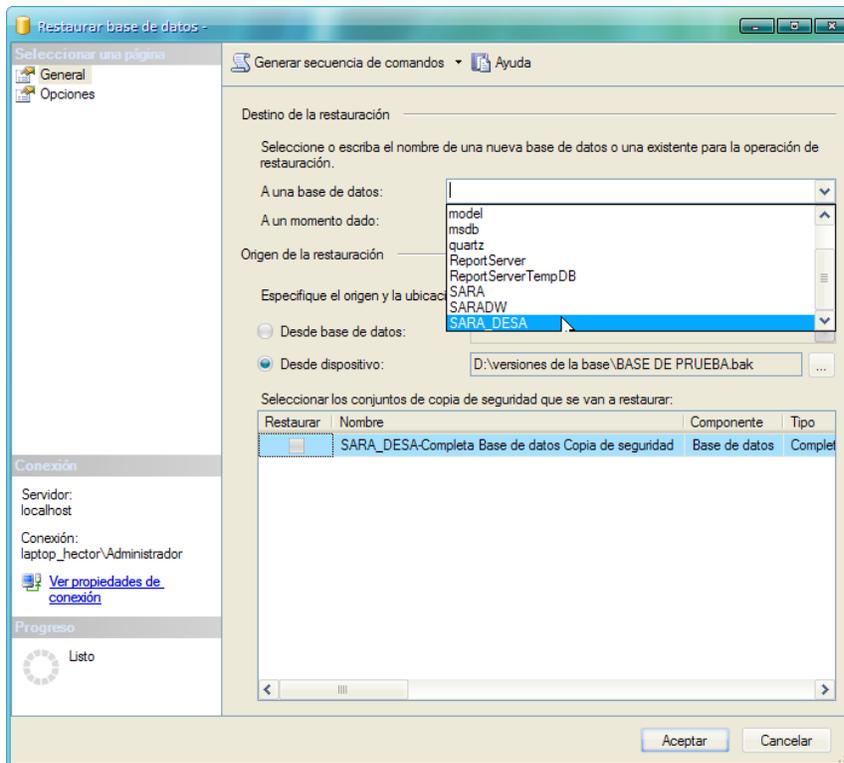


Figura 4.74: Elijiendo la base de datos a restaurar

**Paso 7:** Seleccionar la base de datos a restaurar. Y hacer clic sobre “Opciones” situado en el selector de página.

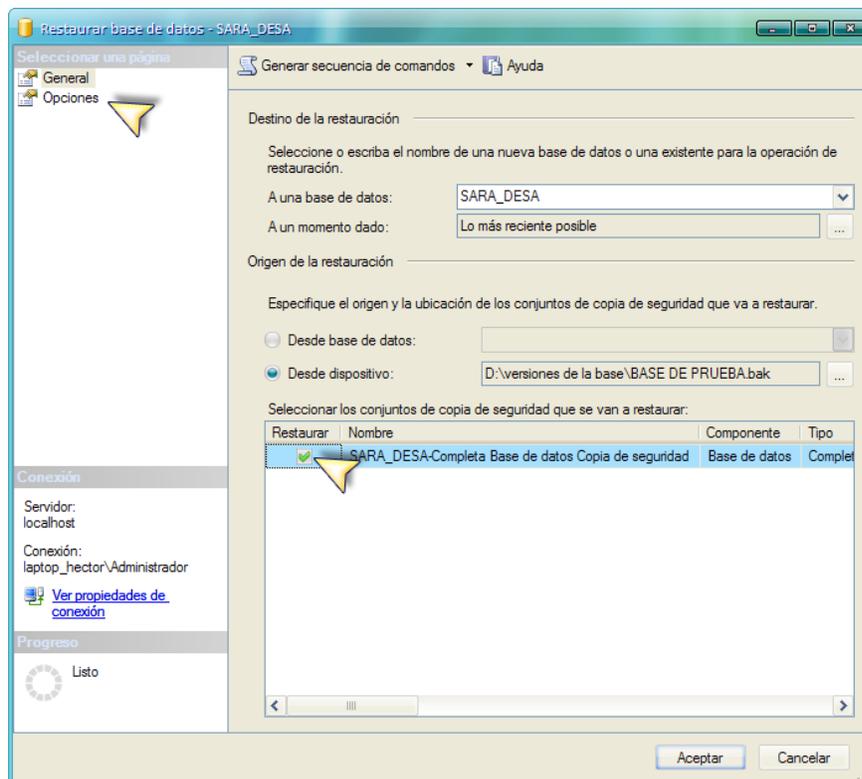
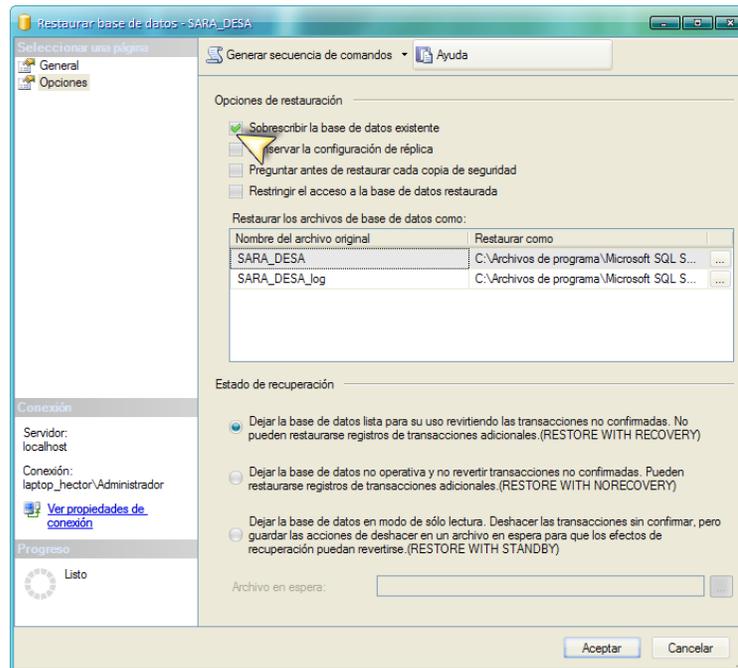


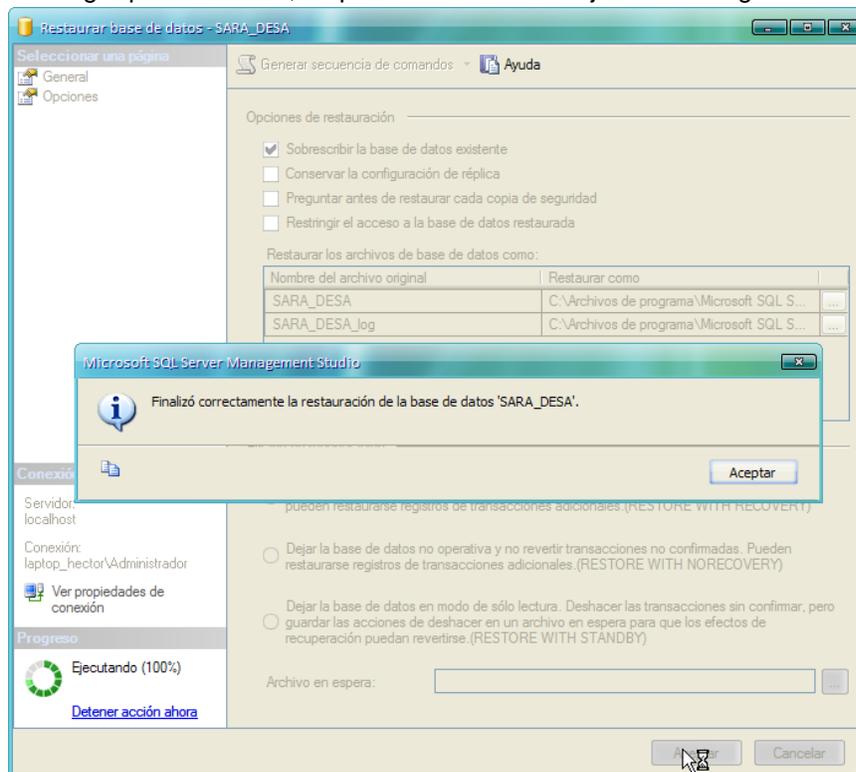
Figura 4.75: Elegir opciones para determinar las características de la restauración

**Paso 8:** Al seleccionar las opciones, se mostraran las siguientes opciones, se debe de seleccionar la opción “Sobrescribir la base de datos existente”. Y para finalizar hacer clic en aceptar.



**Figura 4.76: Elegir opciones para la restauración**

**Paso 9:** Si todo se logro procesar bien, se presentara un mensaje como el siguiente.



**Figura 4.77: Confirmación de la restauración realizada correctamente**

**Paso 10:** Solo queda verificar que la base se encuentre restaurada correctamente.



**Figura 4.78:** Confirmación de la restauración realizada correctamente

## 4.2. MANUAL TÉCNICO

### 4.2.1. Pentaho BI Server

#### 4.2.1.1. Creación de Roles

**Paso 1**

Ingresa a la consola de administración.

**Paso 2**

Hacer clic en la pestaña “Administration”.



**Paso 3**

En la pestaña “Users & Roles” hacer clic en la figura . Se muestra la información correspondiente a los roles creados previamente tal como se muestra en la Figura 4.79.

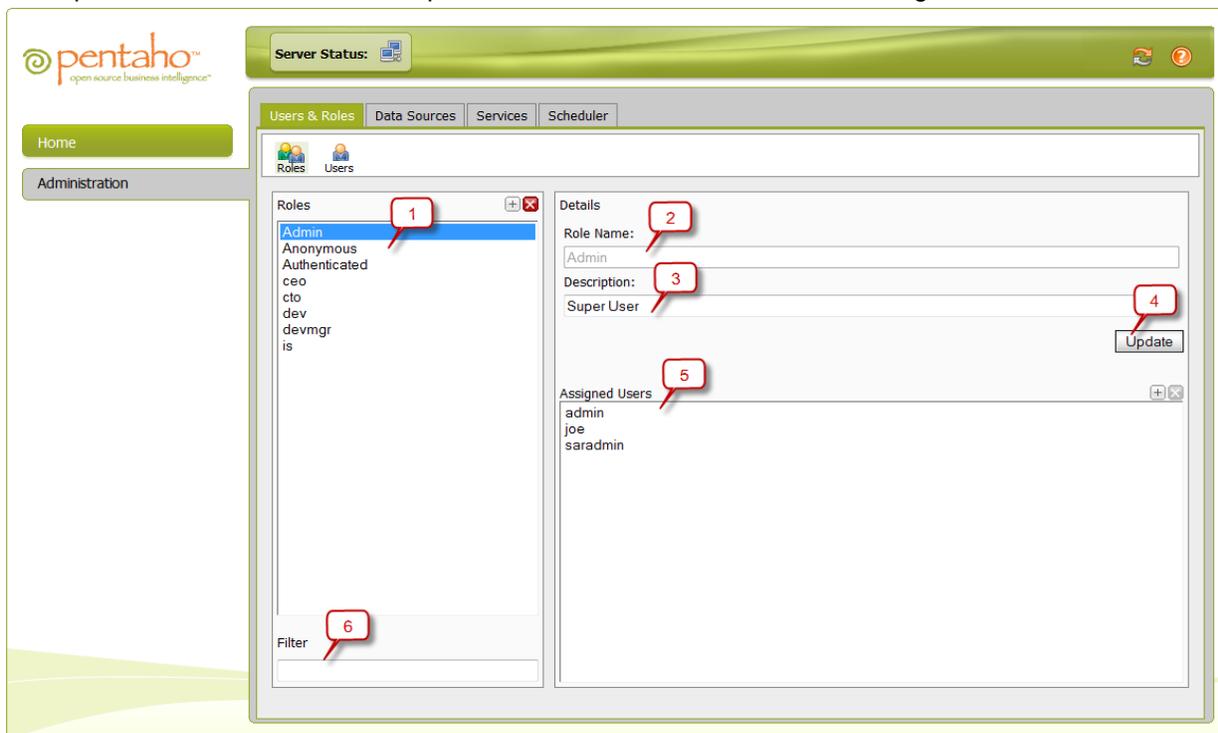


Figura 4.79: Consola de administración de roles

Número	Descripción
1	Roles existentes en el sistema.
2	Nombre del Rol.
3	Descripción del Rol. Puede ser modificado.
4	Botón de actualización de cambios realizados a la descripción del Rol seleccionado.
5	Usuarios asignados al Rol.
6	Filtro que permite la búsqueda de roles.

**Paso 4**

Hacer clic en el botón mostrado en la Figura 4.80 para agregar un nuevo rol.

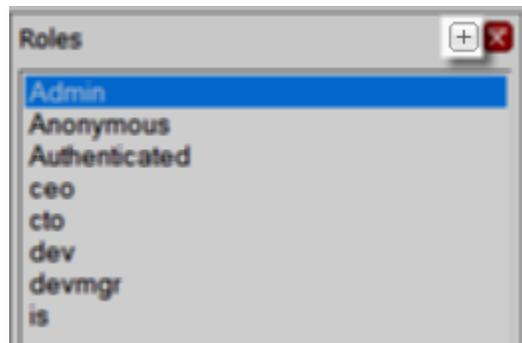


Figura 4.80: Botón para adicionar un nuevo rol

**Paso 5**

En la figura se muestra la Figura 4.81 se muestra la pantalla para ingresar los datos correspondientes al nombre y descripción del Rol.

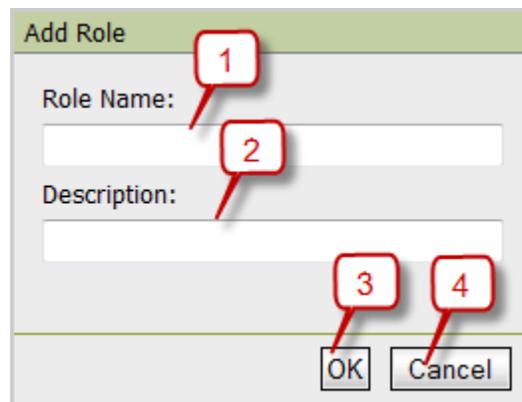


Figura 4.81: Cuadro de dialogo solicitando información del nuevo Rol

Número	Descripción
1	Campo para ingresar el nombre del Rol.
2	Campo para ingresar la descripción del Rol.
3	Guardar el Rol.
4	Cancelar los cambios.

Luego de presionar el botón “ok” el rol creado es mostrado en el listado de roles disponibles para asignar.

### 4.2.1.2. Asignar Rol a Usuario

**Paso 1**

Repetir pasos del 1 al 3 de la creación de Roles.

**Paso 2**

Hacer clic en el símbolo “+” tal como se muestra en la Figura 4.82



Figura 4.82: Botón de asignación de usuario a rol

**Paso 3**

Seleccionar el usuario al que se le desea asignar el rol de la lista de usuarios disponibles, tal como se muestra en la Figura 4.83.

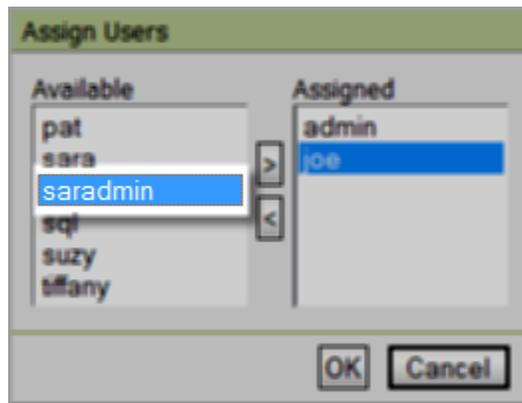


Figura 4.83: Usuario seleccionado para asignación de rol

**Paso 4**

Hacer clic en el botón “>” mostrado en la Figura 4.84

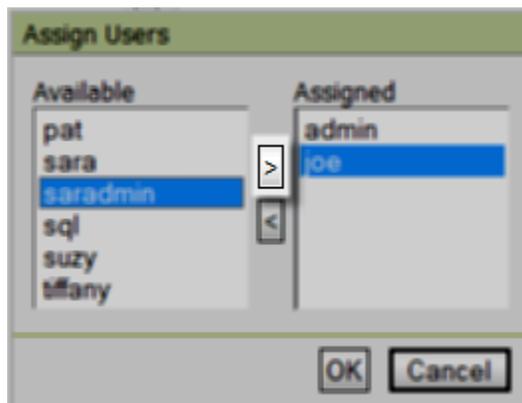


Figura 4.84: Botón de asignación de rol a usuario

**Paso 5**

El usuario es mostrado en la lista de Asignados tal como se muestra en la Figura 4.85. Para guardar los cambios realizados hacer clic en “OK” de lo contrario en “Cancel”.

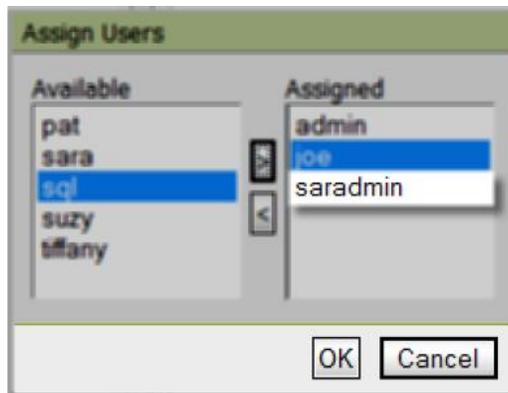


Figura 4.85: Usuario asignado a rol seleccionado

#### 4.2.1.3. Quitar Rol a Usuario

**Paso 1**

Realizar paso 1, 2,3 de la Creación de Roles.

**Paso 2**

Seleccionar el Rol que se desea quitar al usuario, como se muestra en la Figura 4.86

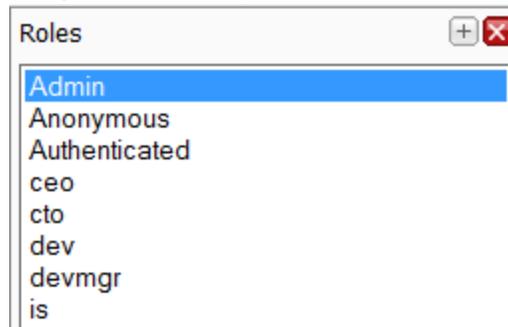


Figura 4.86: Selección de rol

**Paso 3**

En el recuadro de “Usuarios asignados” seleccionar el usuario al que se le desea quitar el Rol, como se muestra en la Figura 4.87.

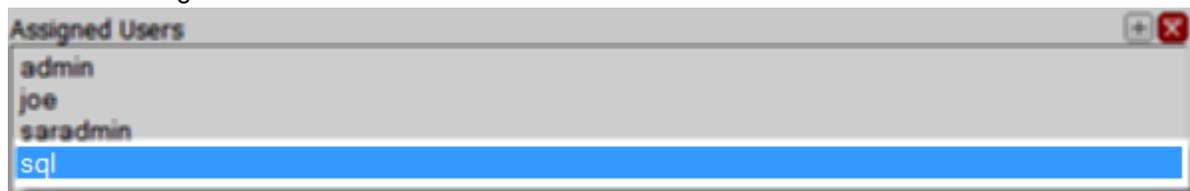


Figura 4.87: Selección de usuario a quitar rol

**Paso 4**

Hacer clic en la “X” mostrada en la Figura 4.88

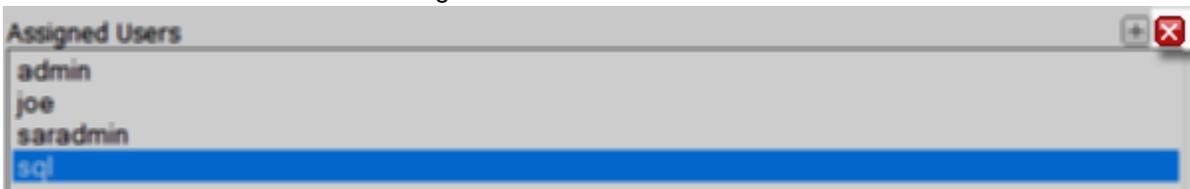


Figura 4.88: Botón para quitar rol a usuario

**Paso 5**

Se muestra la ventana de la Figura 4.89, en la cual se pide la confirmación de la des asignación del rol al usuario. Para completar con el proceso hacer clic en “OK” de lo contrario hacer clic en “Cancel”.

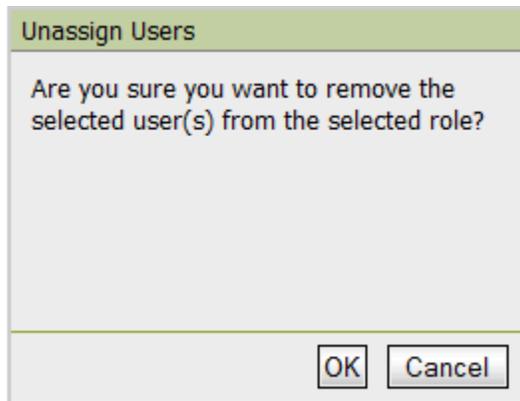


Figura 4.89: Cuadro de confirmación para quitar rol a usuario

#### 4.2.1.4. Eliminar un Rol

**Paso 1**

Realizar paso 1, 2,3 de la Creación de Roles

**Paso 2**

Seleccionar los Roles que se desea eliminar.

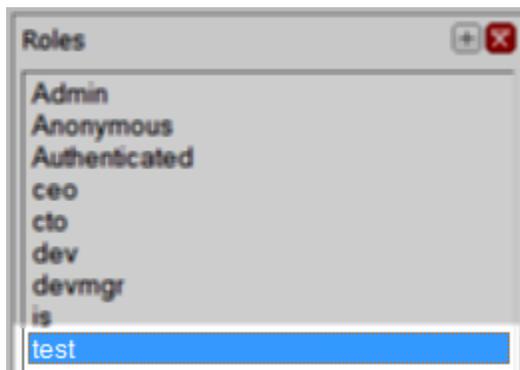


Figura 4.90: Selección de rol

**Paso 3**

Hacer clic en la "X" mostrada en la Figura 4.91

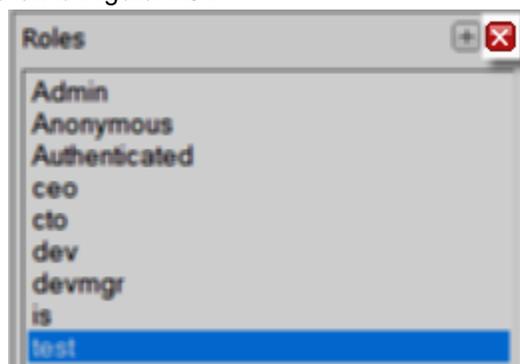


Figura 4.91: Botón de eliminación de rol

**Paso 4**

Se mostrara le cuadro de confirmación de la Figura 4.92. Si se desea proceder con la eliminación del Rol, hacer clic en "OK" de lo contrario hacer clic en "Cancel".

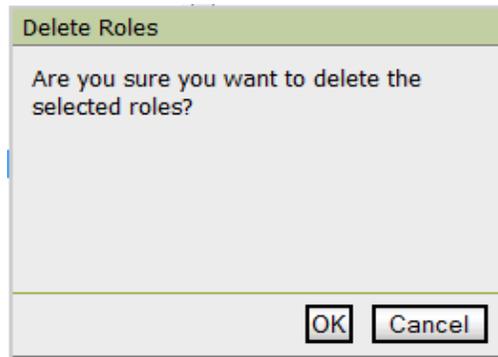


Figura 4.92: Cuadro de confirmación de eliminación de rol

#### 4.2.1.5. Creación de Usuarios

**Paso 1**

Ingresar a la consola de administración.

**Paso 2**

Hacer clic en la pestaña "Administration".



**Paso 3**



En la pestaña "Users & Roles" hacer clic en la figura . Se muestra la información correspondiente a los usuarios creados previamente tal como se muestra en la Figura 4.93.

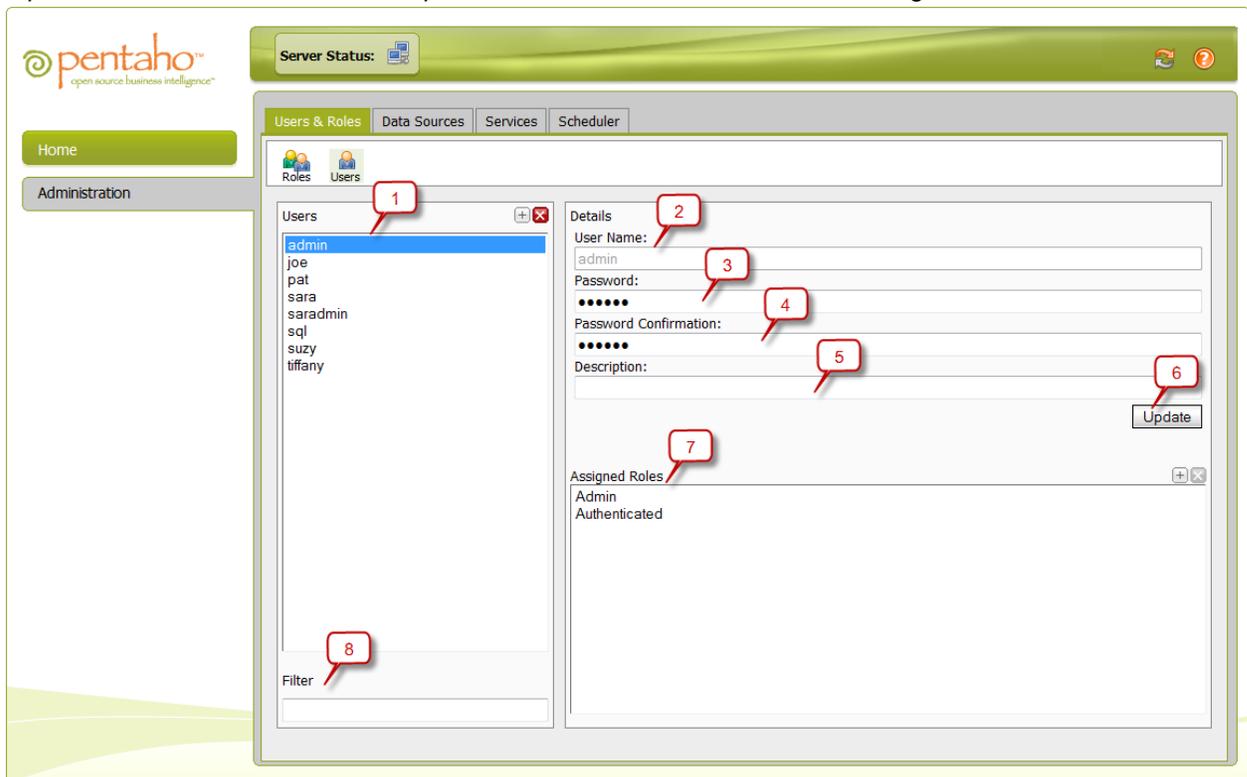


Figura 4.93: Consola de administración de usuarios

Número	Descripción
1	Usuarios existentes en el sistema.
2	Nombre del usuario seleccionado.
3	Campo para realizar cambio de contraseña del usuario seleccionado.
4	Confirmación del cambio de contraseña.
5	Descripción del usuario
6	Botón de actualización de cambios.
7	Roles asignados al usuario seleccionado
8	Filtro de búsqueda de usuarios por nombre.

**Paso 4**

Hacer clic en el botón mostrado en la Figura 4.94 para agregar un nuevo usuario.

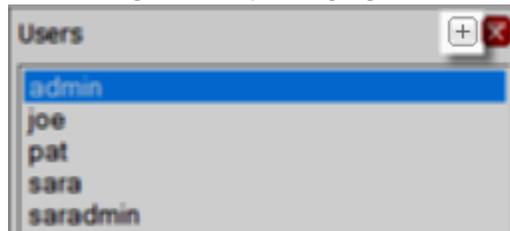


Figura 4.94: Botón de adición de usuarios

**Paso 5**

En la figura se muestra la Figura 4.95 se muestra la pantalla para ingresar los datos requeridos para crear un nuevo usuario.

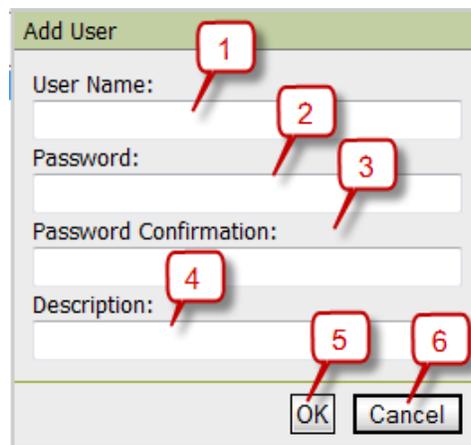


Figura 4.95: Pantalla de ingreso de datos de nuevo usuario

Número	Descripción
1	Campo para ingresar el nombre del usuario.
2	Campo para ingresar la contraseña del usuario.
3	Campo para confirmar la contraseña del usuario.
4	Campo para ingresar la descripción del usuario.
5	Guardar el Usuario.
6	Descartar los cambios realizados.

Luego de presionar el botón “ok” el usuario creado es mostrado en el listado de usuarios disponibles en el sistema.

#### 4.2.1.6. Eliminar un Usuario

**Paso 1**

Realizar paso 1, 2,3 de la Creación de Usuarios.

**Paso 2**

Seleccionar los Usuarios que se desea eliminar.

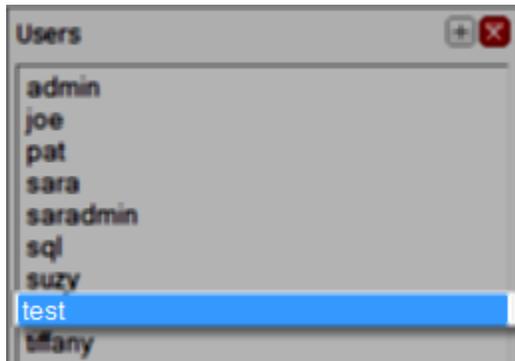


Figura 4.96: Selección de usuario a eliminar

**Paso 3**

Hacer clic en la “X” mostrada en la Figura 4.97

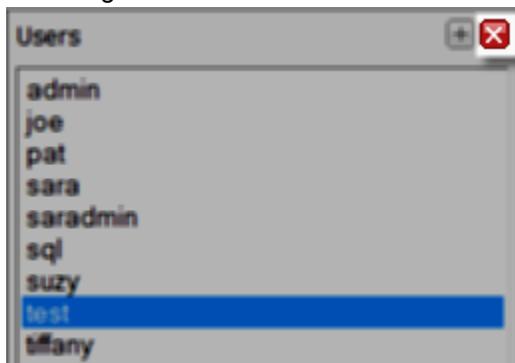


Figura 4.97: Botón para eliminación de usuario

**Paso 4**

Se mostrara le cuadro de confirmación de la Figura 4.98. Si se desea proceder con la eliminación del Usuario, hacer clic en “OK” de lo contrario hacer clic en “Cancel”.

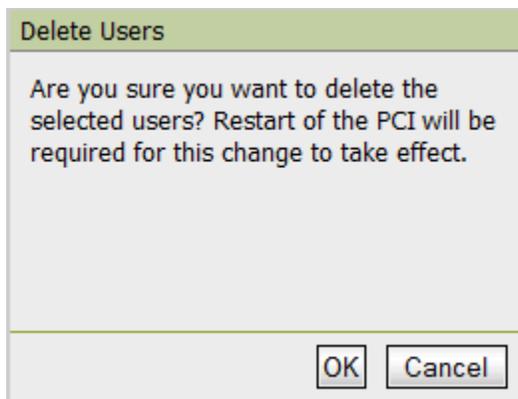


Figura 4.98: Pantalla de confirmación de eliminación de usuario

El usuario será eliminado y quitado de la lista de usuarios disponibles.

### 4.2.2. Manual Schema Workbench

Mondrian Schema Workbench es una interfaz de diseño que permite crear y probar de forma visual esquemas OLAP para Mondrian. El motor de Mondrian procesa las consultas MDX hacia los modelos OLAP. Estos esquemas son archivos modelos de metadatos XML que son creados con una estructura específica por el motor de Mondrian. Estos modelos XML pueden ser considerados como estructuras parecidas a los cubos, los cuales utilizan Hechos existentes y tablas de Dimensión en el RDBMS.

#### Preparación

Antes de iniciar schema workbench es necesario verificar si se encuentra instalado el driver jdbc para conectarse al Data Warehouse en SQL Server 2005. En la carpeta <directorio\_instalacion>\schema-workbench\drivers verificar que se encuentre el archivo *sqljdbc4*.

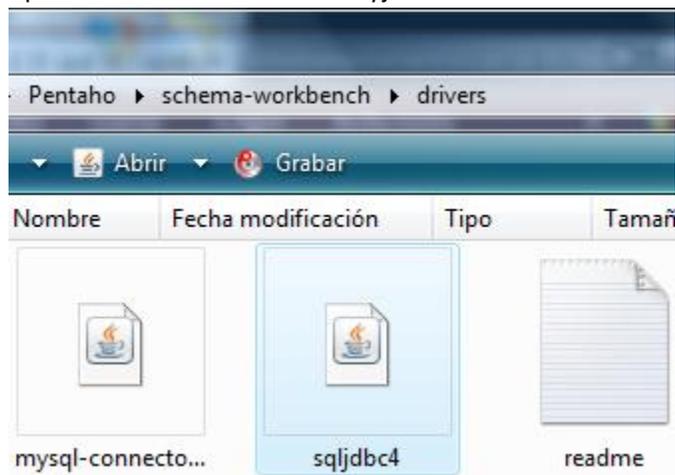


Figura 4.99: Comprobación del driver jdbc

#### 4.2.2.1. Configuración de conexión al DW desde Schema Workbench

##### Paso 1

El primer paso consiste en abrir el programa. Para esto es necesario dirigirse al directorio <directorio\_instalacion>\schema-workbench\ y ejecutar el archivo *workbench.bat*.



Figura 4.100: Icono del archivo .bat que inicia la herramienta Schema Workbench

Una vez iniciado se mostrara la ventana principal (Figura 4.101)

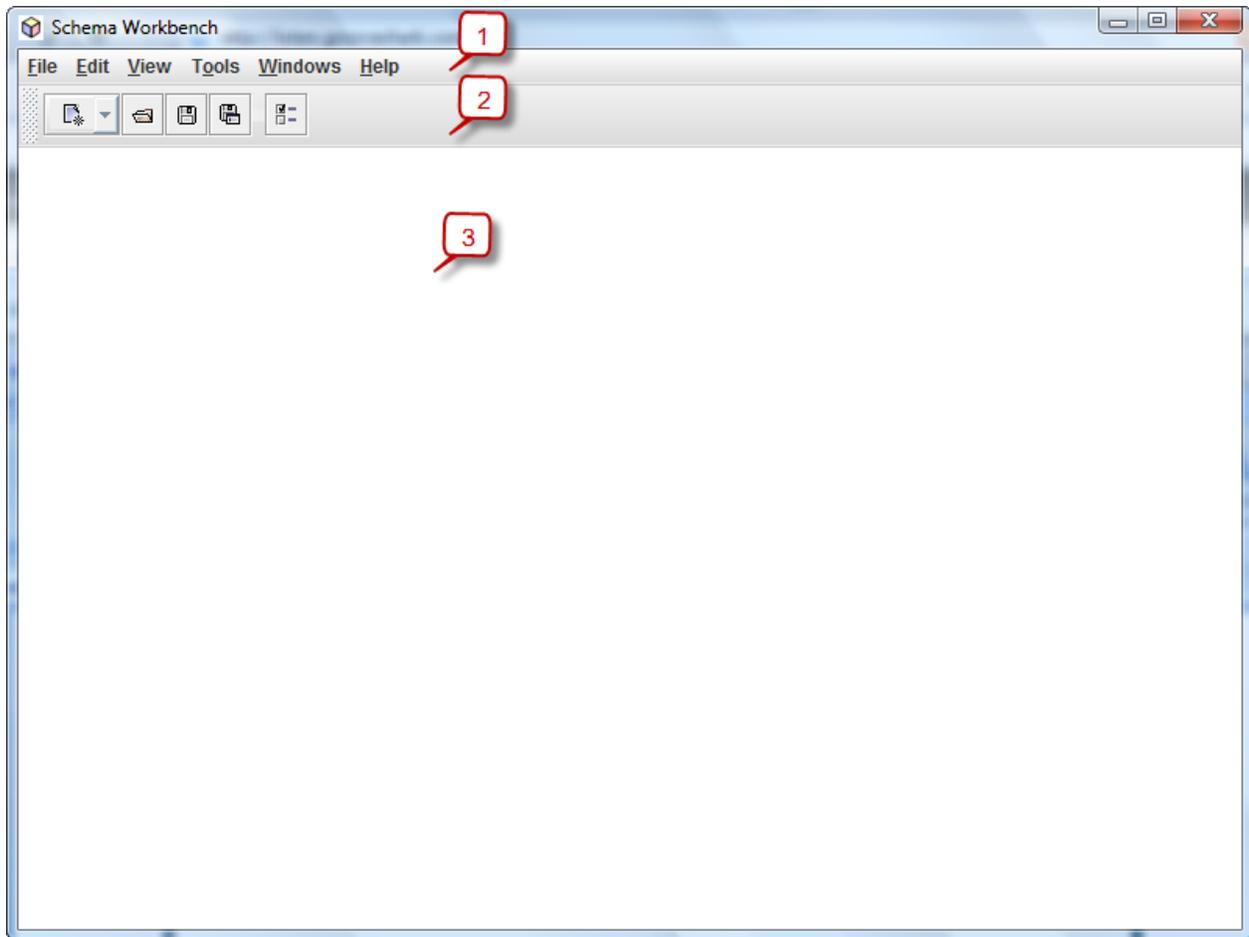


Figura 4.101: Ventana principal de Schema Workbench

Número	Nombre	Descripción
1	Barra de menú	Disponibles todos los comandos de schema workbench.
2	Barra de funciones	En esta barra se encuentran botones como nuevo, abrir, guardar, etc.
3	Espacio de Trabajo	En este espacio se abren nuevas ventanas sobre las cuales se realizan las tareas en schema workbench.

**Paso 2**

El siguiente paso consiste en configurar la conexión al DW. (Figura 4.102)

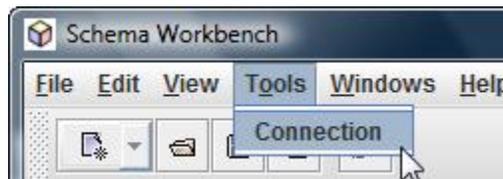
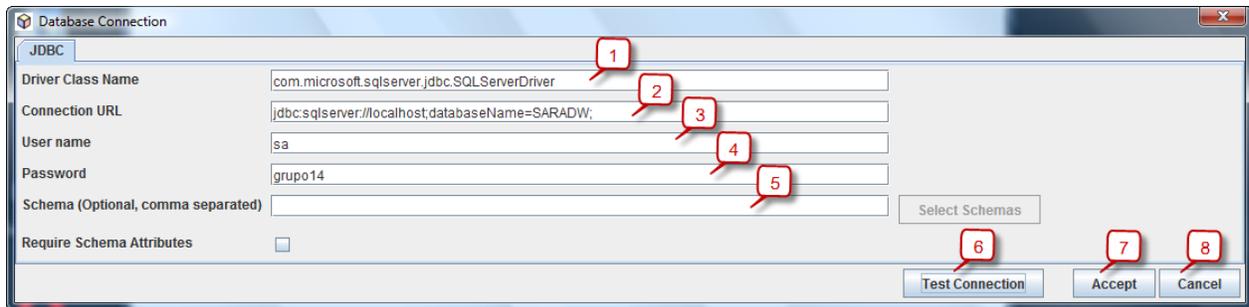


Figura 4.102: Selección de menú para acceder a configuración de conexión

**Paso 3**

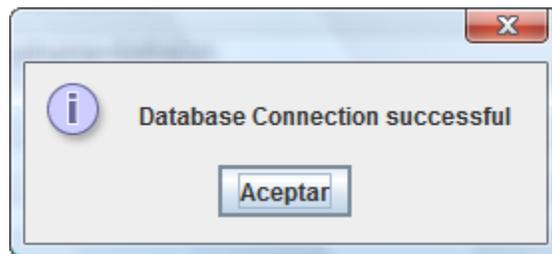
Se despliega la ventana de la Figura 4.103, la cual solicita la información necesaria para realizar la conexión por medio de JDBC al DW.



**Figura 4.103: Pantalla de parámetros para conexión al DW**

Número	Nombre	Descripción
1	Driver Class Name	Nombre de la clase de conexión a la base de datos.
2	Connection URL	URL de conexión a la base de datos.
3	User Name	Nombre de usuario de la base de datos.
4	Password	Contraseña del usuario de conexión a la base de datos.
5	Schema	Esquema de la base de datos al cual se desea conectar (parámetro opcional).
6	Test Connection	Botón que realiza una prueba de conexión utilizando los parámetros digitados.
7	Accept	Botón que confirma la utilización de los parámetros de conexión.
8	Cancel	Deshacer los cambios realizados a la conexión.

Antes de hacer clic en el botón Aceptar es recomendable probar los parámetros de conexión haciendo clic en el botón Test Connection. Si la prueba es realizada con éxito se mostrara el mensaje de la Figura 4.104.



**Figura 4.104: Mensaje de éxito en la prueba de conexión a la base de datos**

Una vez se ha confirmado la conexión con el DW hacer clic en el botón Aceptar.

#### 4.2.2.2. Creación de un cubo

##### Paso 1

En la barra de menú hacer clic en File -> New -> Schema (Ver Figura 4.105)

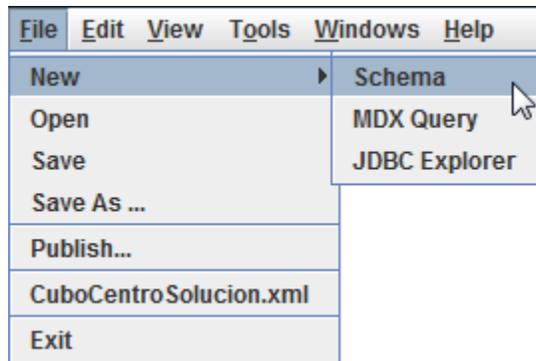


Figura 4.105: Menú de acceso para creación de un nuevo cubo

La ventana de la Figura 4.106 es mostrada en el espacio de trabajo.

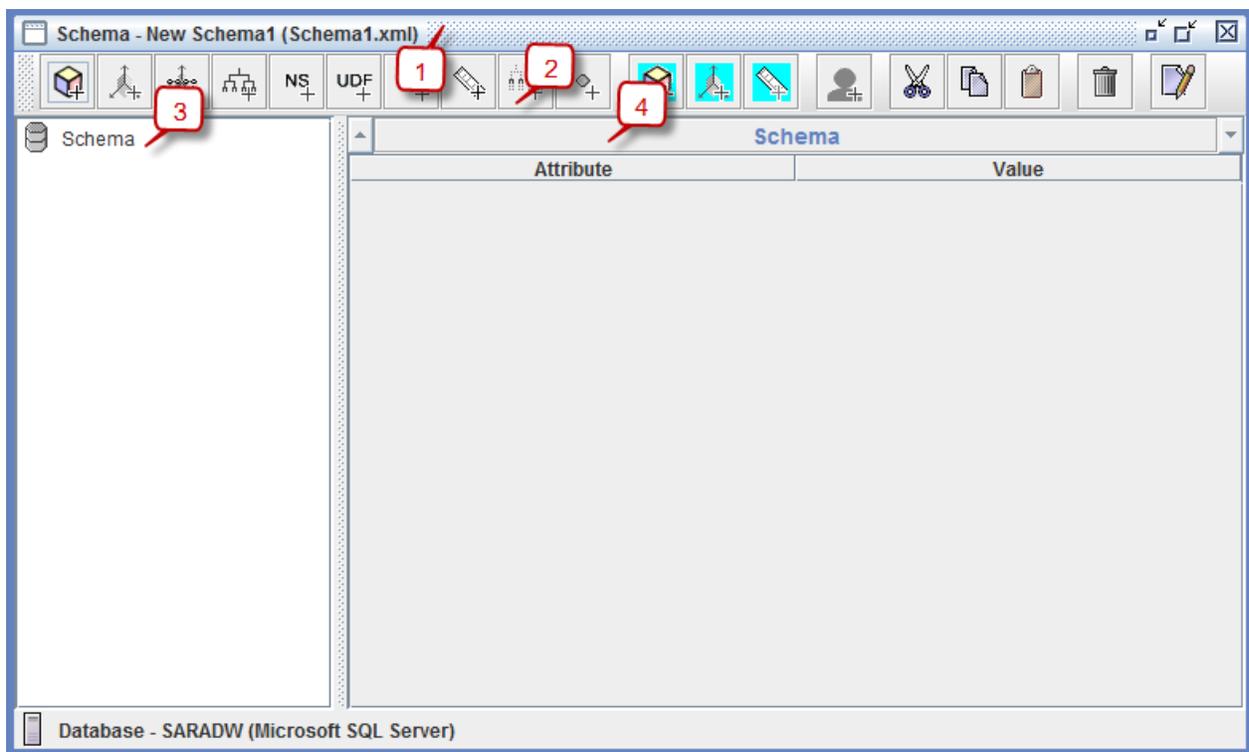


Figura 4.106: Espacio de trabajo para el manejo de cubos

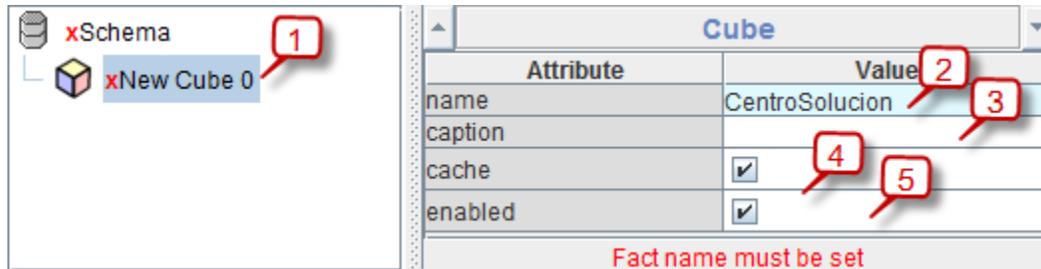
Número	Nombre	Descripción
1	Barra de Titulo	Muestra el nombre del esquema en el que se está trabajando.
2	Barra de Herramientas	Cuenta con las opciones tales como agregar dimensión, medida, etc.
3	Árbol de navegación.	Muestra cada uno de los elementos correspondientes al esquema del cubo.
4	Ventana de Propiedades	En esta se muestran las propiedades de los elementos seleccionados en el menú de árbol.

Icono	Nombre	Descripción
	Add Cube	Adiciona un cubo a un esquema.
	Add Dimension	Agrega una dimensión a un cubo.
	Add Dimension Usage	Establece el uso de dimensión para un Hecho.
	Add Hierarchy	Agrega una jerarquía dentro de una dimensión del cubo.
	Named Set	Define conjuntos con nombre. Un conjunto con nombre es una expresión que devuelve un conjunto de miembros de la dimensión.
	User Defined Function	Añade el uso de funciones creadas por el usuario, las cuales deben estar dentro del class path de mondrian.
	Calculated Member	Define miembros calculados para las medidas del cubo.
	Add Measure	Añade una medida a un Hecho
	Add Level	Agrega un nivel a la jerarquía de la dimensión
	Add Property	Permite agregar propiedades de miembro a una jerarquía. Estos son atributos de de una dimensión que proporcionan información adicional a los usuarios del cubo.
	Add Virtual Cube	Añade un cubo virtual, el cual es un cubo lógico basado en uno o más cubos regulares o relacionados. Su concepto es similar al de una vista en un RDBMS.
	Add Virtual Cube Dimension	Agrega una Dimension Virtual, que hace referencia a dimensiones compartidas y añadir las al cubo virtual.
	Add Virtual Cube Measure	Añade una medida virtual al cubo, para utilizar las medidas definidas en otros cubos.
	Add Role	Agrega un Rol al cubo para otorgar accesos a los elementos del cubo.
	Cut	Corta un elemento
	Copy	Copia un elemento.
	Paste	Pega el elemento cortado o copiado previamente.
	Delete	Elimina un elemento del cubo.
	Edit Mode	Cambia la ventana de propiedades por el código XML de la definición del objeto seleccionado.

Tabla 4.7: Elementos de la barra de herramientas

**Paso 2**

Para agregar un cubo al esquema se debe hacer clic en el icono  y luego llenar las propiedades del cubo. (Figura 4.107)

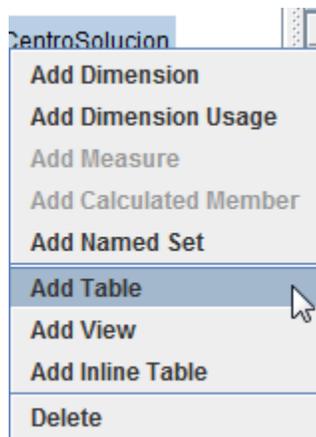


**Figura 4.107: Propiedades del cubo añadido**

Número	Nombre	Descripción
1	Cubo agregado	Icono que representa el cubo agregado.
2	Name	Nombre del cubo.
3	Caption	Leyenda del cubo a utilizar
4	Cache	Si el cubo utilizara cache.
5	Enabled	El cubo está disponible para su utilización.

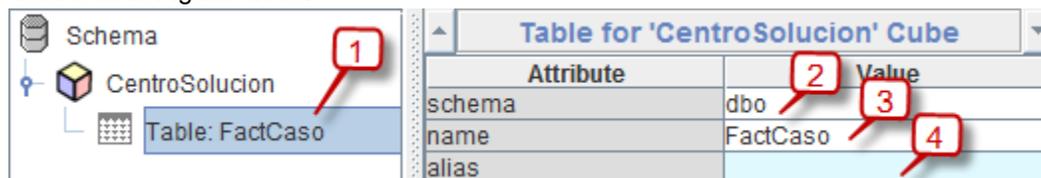
**Paso 3**

El siguiente paso es agregar la tabla de hechos al cubo, para lo cual se debe hacer clic derecho sobre el cubo creado y en el menú desplegable de la Figura 4.108. Seleccionar la opción Add Table «Agregar Tabla», la cual asocia una tabla del DW a un objeto del cubo.



**Figura 4.108: Menú desplegable al hacer clic derecho sobre el cubo**

Luego se mostrara una ventana donde se deben llenar los datos de la tabla de hechos del cubo tal como se muestra en la Figura 4.109.



**Figura 4.109: Datos de la tabla de hechos**

Número	Nombre	Descripción
1	Table: FactCaso	Tabla agregada.
2	Schema	Esquema al cual pertenece la tabla.
3	Name	Nombre de la tabla.
4	Alias	Alias a utilizar para hacer referencia a la tabla.

**Paso 4**

El siguiente paso es agregar una dimensión al esquema, para lo cual es necesario hacer clic en el botón  de la barra de herramientas. En el menú de navegación del esquema se añadirá la dimensión y una jerarquía a las cuales es necesario ingresar las propiedades (Figura 4.110).

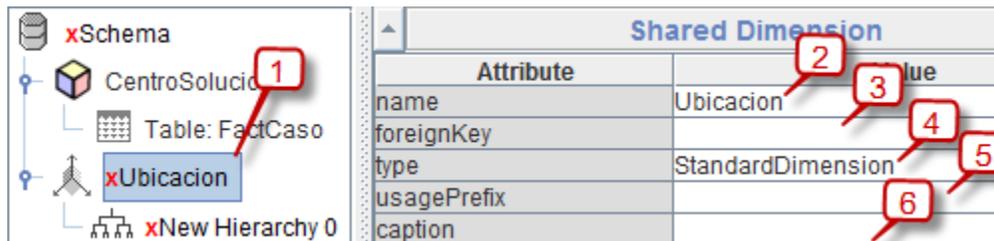


Figura 4.110: Propiedades de la dimensión

Número	Nombre	Descripción
1	Ubicación	Nombre asignado a la dimensión.
2	Name	Nombre que se le desea asignar a la dimensión.
3	foreignKey	Llave foránea de la dimensión a la tabla de hechos.
4	Type	El tipo de dimensión si es Standard o de Tiempo.
5	usagePrefix	Prefijo que antecederá a la dimensión antes de llamarla.
6	Caption	Nombre con el que será visualizada la dimensión.

**Paso 5**

Luego de configurar la dimensión se procede a configurar la jerarquía, para lo cual es necesario agregar la tabla del DW correspondiente a la dimensión, seleccionando la jerarquía tal como se muestra en el Paso 3.

Ingresar las propiedades de la jerarquía. (Figura 4.111).

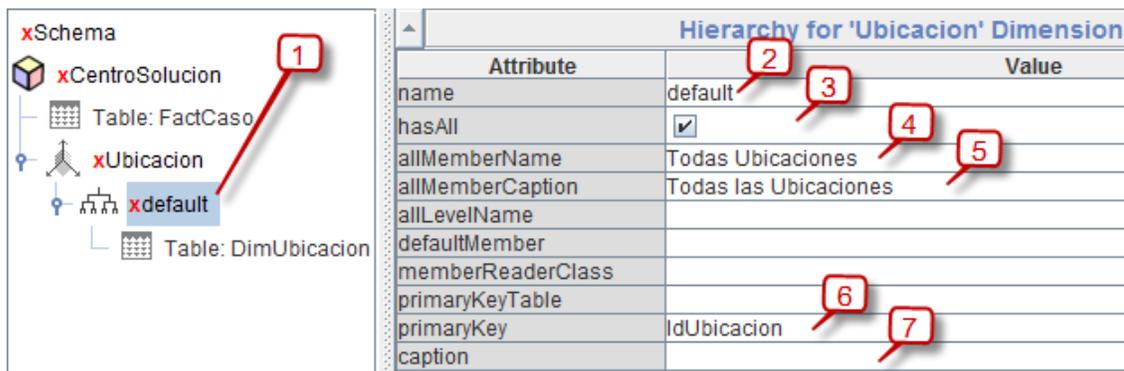


Figura 4.111: Propiedades de la jerarquía

Número	Nombre	Descripción
1	Default	Objeto jerarquía en el árbol de navegación.
2	name	Nombre asignado a la jerarquía.
3	hasAll	Seleccionado si la jerarquía posee miembros "all". «todos»
4	allMemberName	Nombre de la propiedad de los miembros "all".
5	allMemberCaption	Nombre con el que se visualizara los miembros "all"
6	usagePrefix	Prefijo que antecederá a la dimensión antes de llamarla.
7	Caption	Nombre con el que será visualizada la jerarquía. Si no se especifica toma el nombre de la jerarquía.

**Paso 6**

Una vez configurada la jerarquía se añaden los niveles, para lo cual es necesario hacer clic derecho en la jerarquía y luego seleccionar Add Level «Añadir nivel» (Figura 4.112).

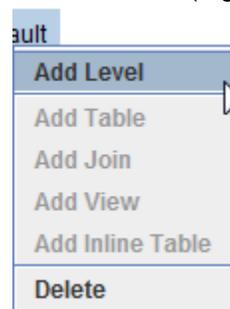


Figura 4.112: Selección de menú para crear un nuevo nivel

Un objeto de nivel es creado en el árbol de navegación, el cual debe ser configurado tal como muestra la Figura 4.113.

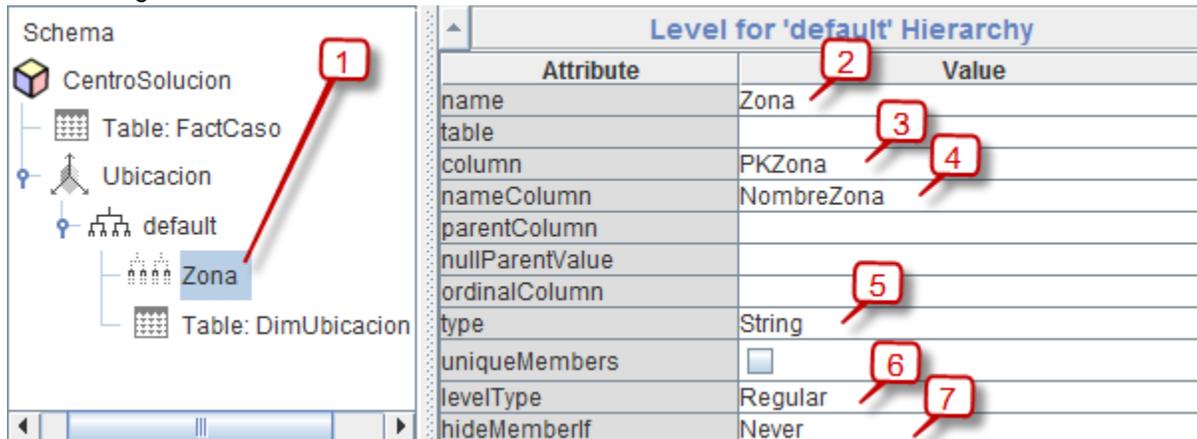


Figura 4.113: Propiedades de un nivel

Número	Nombre	Descripción
1	Zona	Objeto nivel en el árbol de navegación
2	name	Nombre asignado al nivel.
3	Column	Columna de la tabla que será utilizada como identificador del nivel.
4	nameColumn	Descripción asociada al valor de la columna.
5	Type	Tipo de datos de la columna correspondiente a la jerarquía.

6	levelType	Si el nivel es de tipo regular o alguna medida de tiempo.
7	hideMemberlf	Bajo qué condiciones se deberá ocultar el miembro.

Repetir este último pasó para añadir más niveles a la jerarquía.

**Paso 7**

En este paso se agrega el uso de la dimensión a la Tabla de Hechos para que pueda ser utilizada por

las medidas, para lo cual se selecciona el cubo en el árbol de navegación y se hace clic en el icono  , el cual añade un Dimension Usage «Uso de dimensión». La configuración de las propiedades se muestra en la Figura 4.114.

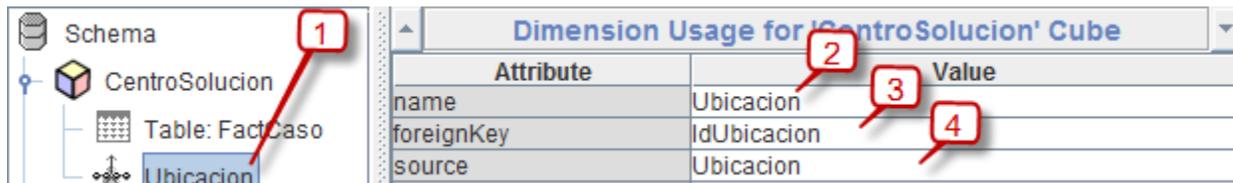


Figura 4.114: Propiedades de un uso de dimensión

Número	Nombre	Descripción
1	Ubicación	Objeto nivel en el árbol de navegación que representa el uso de dimensión.
2	name	Nombre asignado al uso de dimensión.
3	foreignKey	Llave foránea que hace referencia a la dimensión.
4	Source	Dimension que servirá como origen para el uso de dimensión.

Este paso deberá repetirse para cada una de las dimensiones con las cuales se desea relacionar las medidas.

**Paso 8**

El último paso en la creación de un cubo es agregar las medidas que serán utilizadas para realizar los distintos análisis.

Para agregar una medida, hacer clic en el botón  y configurar la medida tal como se muestra en la Figura 4.115

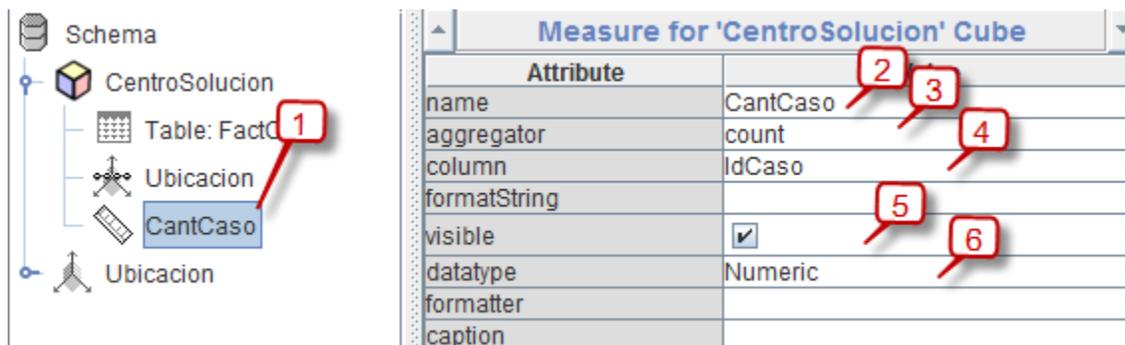


Figura 4.115: Propiedades de una medida

Número	Nombre	Descripción
1	CantCaso	Objeto nivel en el árbol de navegación que representa una medida.
2	name	Nombre asignado a la medida.
3	aggregator	Funcion de agregación que se utilizara en la medida. Ej: count, max, min, sum, etc
4	Column	Columna de la tabla de hechos a la cual le será aplicada la función de agregación.
5	Visible	Si la medida es visible para trabajar con ella.
6	Datatype	Tipo de datos de la columna a utilizar en la función de agregación.

#### 4.2.2.3. Publicación de un Cubo a Pentaho Bi Suite

Para que un cubo esté disponible para su uso en Pentaho es necesario publicarlo siguiendo los pasos mostrados a continuación.

##### Paso 1

Antes de publicar cualquier elemento a Pentaho, este solicita una contraseña de publicación la cual debe ser configurada la ruta **Pentaho\bi\server-ce\pentaho-solutions\system** por medio del archivo **publisher\_config.xml**.

Una vez abierto el archivo xml en un editor de texto, digitar la contraseña de publicación en el tag `<publisher-password>`, como se muestra en la Figura 4.116. La contraseña proporcionada por defecto es **admin**.

```

1  <?xml version="1.0"?>
2  <!--
3      This file must be configured before client tools
4      will be able to publish onto the server. Publishing
5      involves accepting a file (action sequence, xml document,
6      anything) and placing it into the solution. This
7      action should only be allowed if the tool is configured
8      with a server/deployment specific password. The
9      RepositoryFilePublisher is now coded to expect an MD5
10     hash of the password entered here before it will allow
11     publish into the server.
12
13     This is a critical security measure designed to prevent
14     malicious actions from being published into your server.
15     Please take care to make the password sufficiently
16     difficult to guess. Also take care that the contents of
17     this file are not generally available
18 -->
19 <publisher-config>
20     <publisher-password>admin</publisher-password>
21 </publisher-config>
22

```

Figura 4.116: Archivo XML con contraseña de publicación

##### Paso 2

En Schema Workbench seleccionar en la barra de menú File->Publish (Figura 4.117)

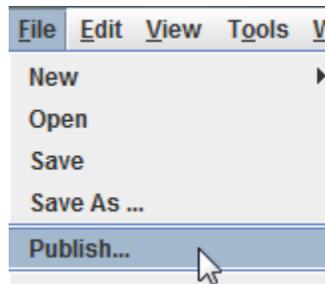


Figura 4.117: Menú para publicar

**Paso 3**

Se despliega la ventana de la Figura 4.118 solicitando las credenciales necesarias para la publicación. Es importante mencionar que el Servidor BI de Pentaho debe estar iniciado.

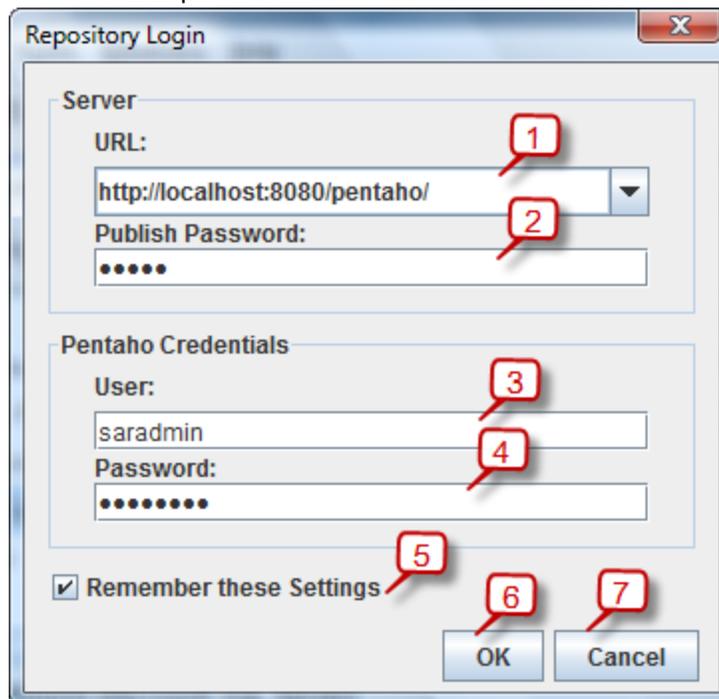


Figura 4.118: Datos solicitados para la publicación

Número	Nombre	Descripción
1	URL	URL del servidor BI de Pentaho
2	Publish Password	Contraseña de publicación.
3	User	Usuario con rol Administrator de Pentaho
4	Password	Contraseña del usuario.
5	Remember these Settings	Si desea recordar los datos ingresados
6	OK	Boton para proceder con la publicación.
7	Cancel	Cancelar la operación.

**Paso 4**

Luego de presionar el botón OK, Schema Workbench se conecta a Pentaho con las credenciales proporcionadas en el paso anterior (Figura 4.119) y posteriormente pide la ubicación y conexión a utilizar para el nuevo archivo.

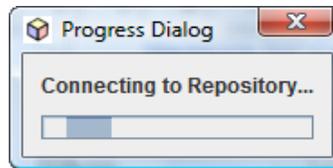


Figura 4.119: Conexión a Pentaho

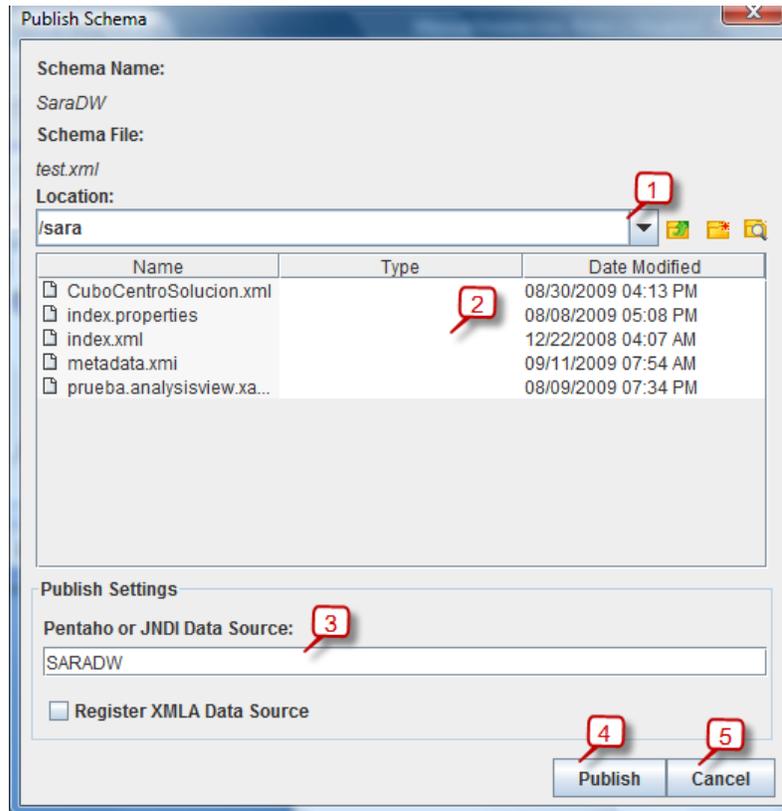


Figura 4.120: Selección de la ubicación de publicación

Número	Nombre	Descripción
1	Location	Carpeta de ubicación del archivo dentro del servidor.
2	Navegador	Archivos en la carpeta.
3	Pentaho or JNDI Data Source	Data Source a utilizar para publicar.
4	Publish	Publicar el archivo.
5	Cancel	Cancelar la operación.

Si la publicación se realiza de forma exitosa es mostrado el mensaje de la Figura 4.121 y el cubo puede ser cargado desde la BI Suite de Pentaho.

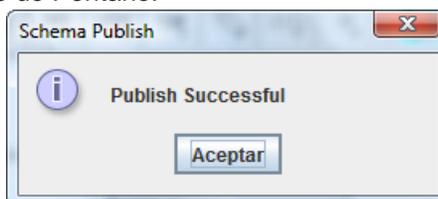


Figura 4.121: Mensaje mostrado al realizar la publicación sin inconvenientes

### 4.2.3. Pentaho Metadata Editor

#### 4.2.3.1. Introducción

Pentaho Metadata Editor (PME) es una herramienta utilizada para construir dominios y modelos. PME, mapea la estructura física de una base de datos en un modelo de negocio lógico. Este mapeo se almacena en un repositorio metadata centralizado el cual puede ser administrado para:

- Definir un lenguaje de negocio para tablas de bases de datos complejas o encriptados.
- Decrementar el costo e impacto asociado al bajo nivel de cambios de una base de datos.
- Configurar parámetros de seguridad, limitando a los usuarios el acceso de datos presentado en reportes.
- Manejar el formato de texto, días, numeración de la presentación de datos en reportes y su mantenimiento.
- Localizar información dependiendo de la configuración regional del usuario.

El metadata modelo de negocio es actualmente el componente superior dentro del dominio de Pentaho metadata. El dominio encapsula ambos modelos; la descripción física de los objetos de la base de datos y el modelo lógico (modelo de negocio) que es la representación abstracta de la base de datos.

El siguiente diagrama (figura 4.122) representa los objetos del negocio y su relación con el dominio de Pentaho metadata.

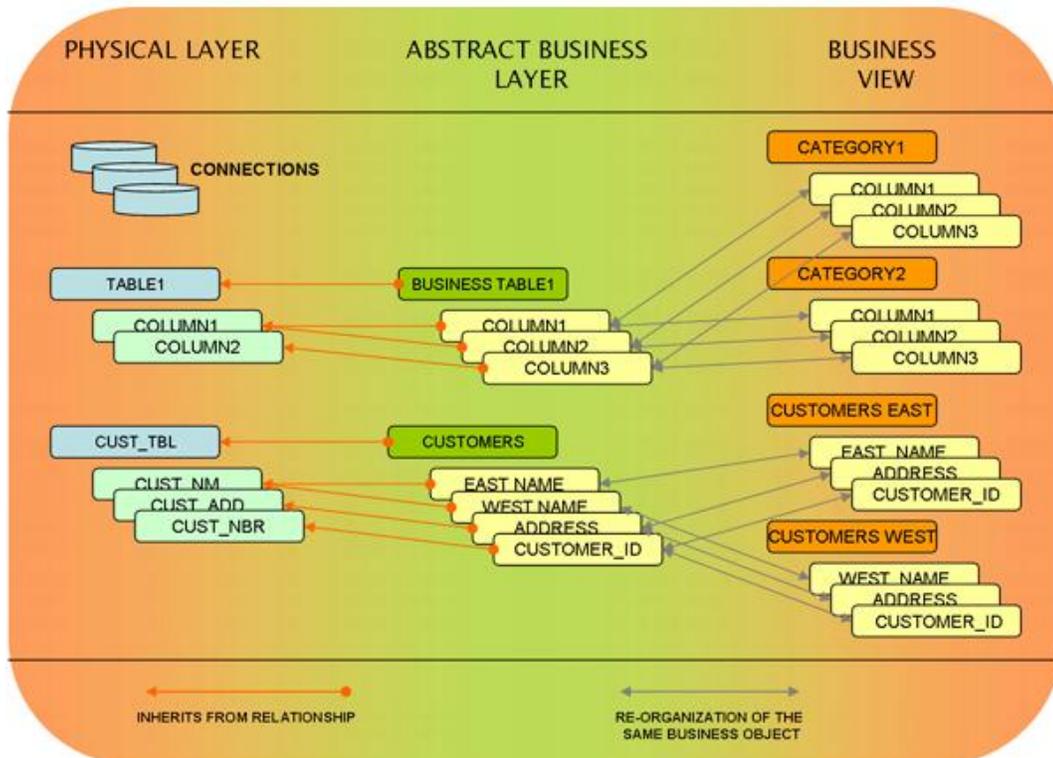
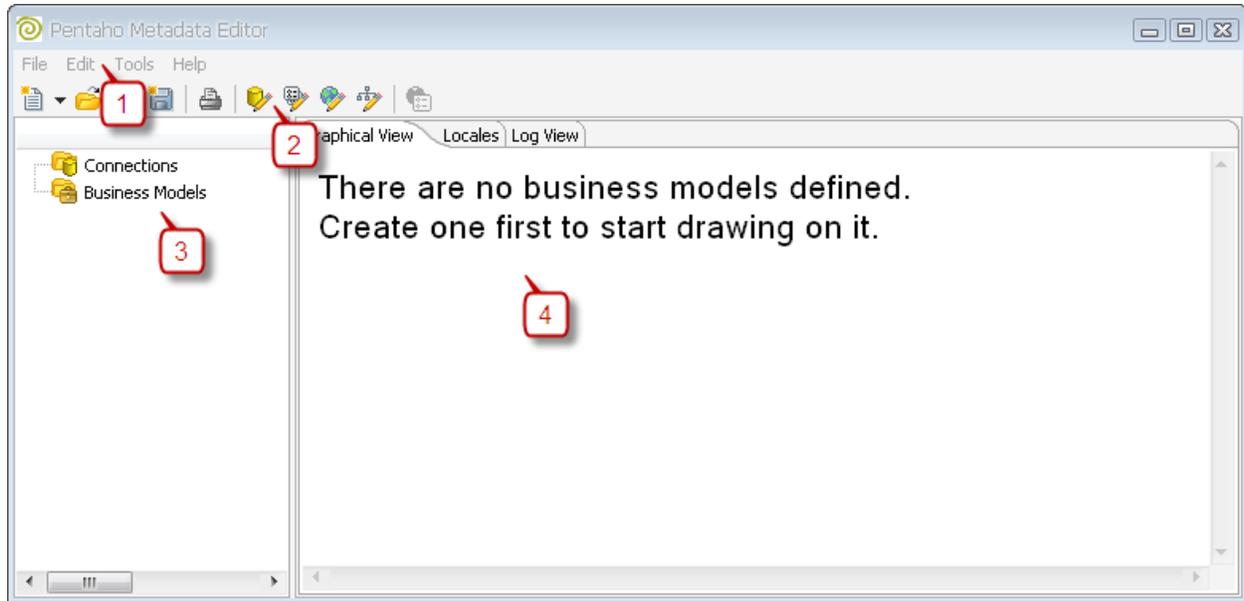


Figura 4.122: Esquema de los Modelos de Negocio

### 4.2.3.2. Iniciando Pentaho Metadata Editor

Al iniciar Pentaho Metadata Editor, podrá observar el ambiente principal de trabajo, tal como se muestra en la Figura 4.123.



**Figura 4.123: Espacio de trabajo**

Los elementos que componen la pantalla principal de PME se describen en la tabla 4.8.

Numero	Nombre	Descripción
1	Barra de Menú	Barra para la administración de las tareas que se realizan dentro de la aplicación, contiene los menús: <ul style="list-style-type: none"> <li>• File</li> <li>• Edit</li> <li>• Tool</li> <li>• Help</li> </ul>
2	Barra de Funciones	Barra que posee los botones que disparan funcionalidades de la aplicación: guardar, abrir, query editor, etc...
3	Navigator Tree «Árbol de Navegación»	Es la ventana que se utiliza para la navegación entre los elementos que posee PME, elementos que se agregan o que son creados.
4	Editor Graph «Editor Grafico»	Espacio de trabajo en el cual se reflejan las tablas de negocios, relaciones entre tablas y los modelos de vista de negocio. Otro elementos que puede reflejar son:

**Tabla 4.8: Elementos pantalla PME**

Pentaho Metadata Editor es una aplicación creada bajo idioma inglés, por lo cual todos los nombres de sus elementos, descripciones y demás están únicamente en inglés. Este manual maneja en inglés como referencia, especificando al mismo tiempo su traducción al castellano.

### 4.2.3.3. Crear Dominio en Pentaho Metadata Editor

#### Crear Dominio Nuevo

Al crear un nuevo dominio es necesario iniciar en blanco las conexiones y los modelos de negocios, para lo cual deberá crear un nuevo dominio mediante File > New > Domain File, tal como se muestra en la Figura 4.124.

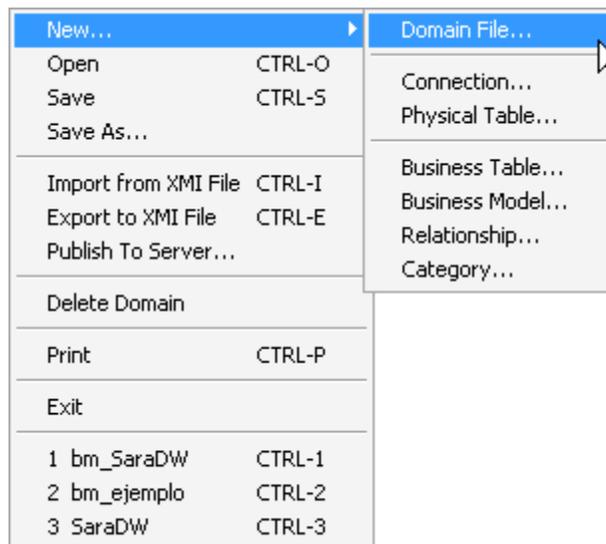


Figura 4.124: File>New>Domain File

#### Crear Conexión

Pentaho Metadata Editor soporta una vasta cantidad de fuentes de datos, que son la base para los modelos de negocios. Antes de definir un modelo de negocio, es necesario describir la base de datos o la fuente de datos que se desea modelar, esto se realiza definiendo una o más conexiones.

Para definir una conexión se abrirá el cuadro de dialogo **Database Connection** «Conexión a Base de Datos», haciendo clic derecho sobre la rama **Connections**, ubicado dentro del **Tree Navigator**, seleccionando **New Connection...**, desde el menú desplegable, Figura 4.125.

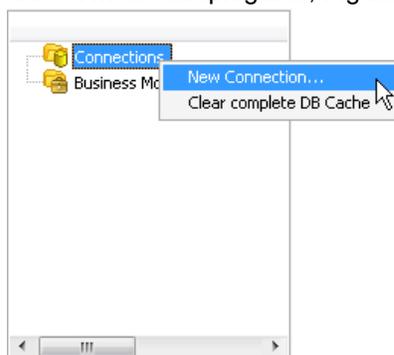


Figura 4.125: Connections>New Connection...

El cuadro de dialogo se mostrara como en la Figura 4.126, este permite definir la conexión con la base de datos en detalle.

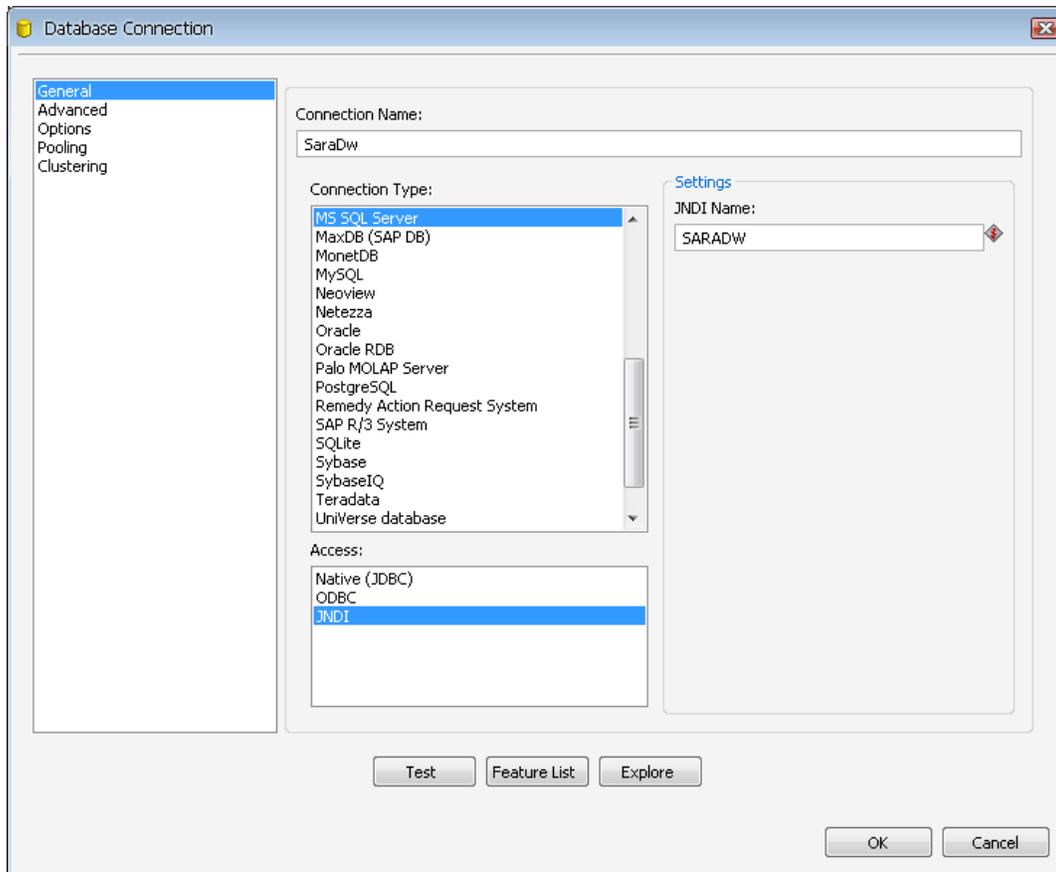


Figura 4.126: Conexión a Base de Datos

En nuestro caso los parámetros de conexión serán para SQL Server, se deberá ingresar la siguiente información en los campos mencionados:

- **Connection Name:** SARADW
- **Connection Type:** MS SQL Server
- **JNDI Name:** SARADW
- **Access:** JNDI (Java Naming and Directory Interface)

Para completar la conexión es necesario realizar los siguientes pasos:

1. Es debe agregar en la ruta **<directorio\_instalacion>\libext\JDBC\**, el driver de SQL Server (Microsoft SQL Server JDBC Driver) para la conexión. Como se muestra en la Figura 4.127, deberá existir el driver **sqljdbc4.jar**.
2. Con el fin que Pentaho Metadata Editor sepa que significa realmente el nombre de la conexión JNDI es necesario configurar el archivo de propiedades JNDI. Abrir con un editor de texto el archivo: **<directorio\_instalacion>\simple-jndi/jdbc.properties**. Ingresando el siguiente texto (Figura 4.128) al final del archivo:
 

```
SARADW/type=javax.sql.DataSource
SARADW/driver=com.microsoft.sqlserver.jdbc.SQLServerDriver
SARADW/url=jdbc:sqlserver://localhost:1433;databaseName=SARADW
SARADW/user=sa
SARADW/password=password
```

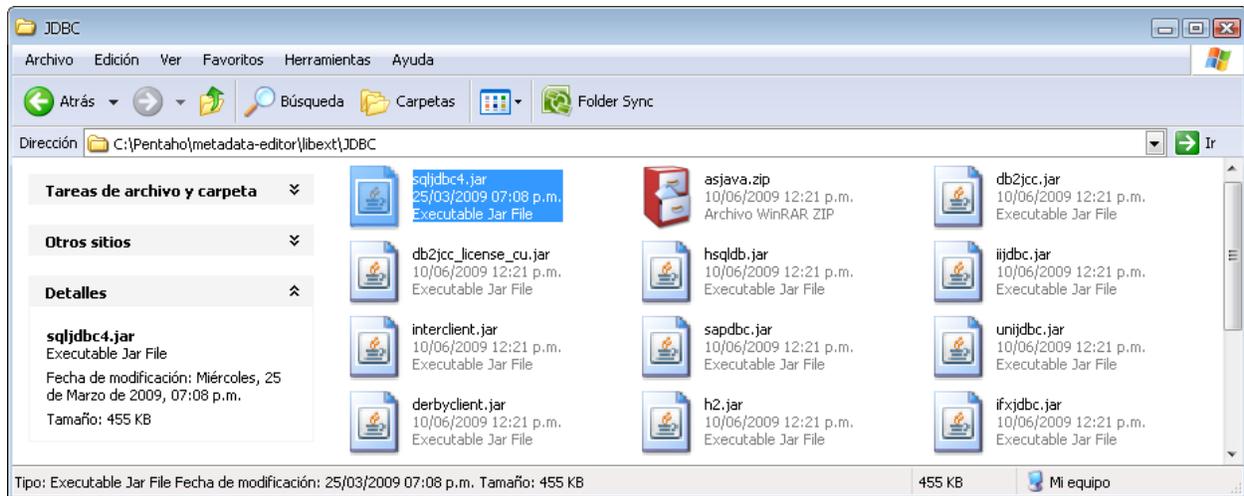


Figura 4.127: Adición Driver Microsoft SQL Server

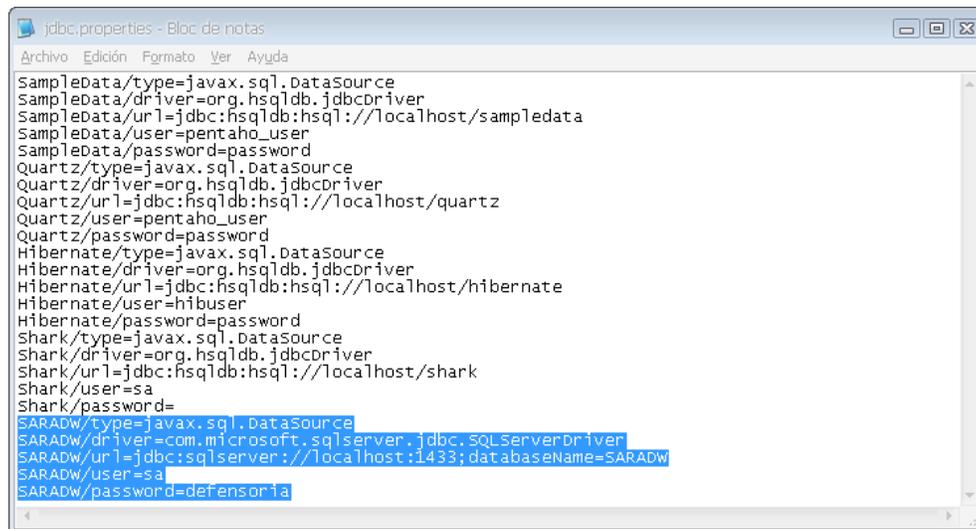


Figura 4.128: Configuración archivo jdbc.properties

Para que Pentaho Metadata Editor reconozca los cambios hechos sobre el archivo de propiedades JNDI será necesario reiniciar la aplicación. Una vez reiniciado Pentaho Metadata Editor, podemos realizar la prueba de conexión y verificar que se ha hecho correctamente la configuración. Para realizar la prueba de conexión con la Base de Datos, hacer clic en el botón **Test** que aparece en la pantalla presentada en la Figura 4.126, si la conexión se establece con éxito, deberá aparecer un cuadro de diálogo como se presenta en la Figura 4.129.

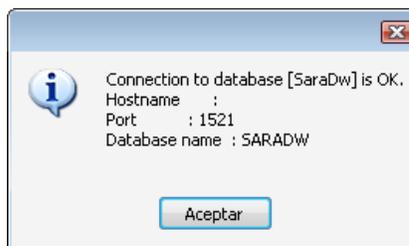


Figura 4.129: Conexión establecida con éxito

#### 4.2.3.4. Importando Tablas Físicas

Dentro del **Tree Navigator**, hacemos clic derecho sobre el nodo SaraDW de la ramificación. Seleccionamos **Import Tables** del menú desplegable, apareciendo el cuadro de dialogo **Import Tables** «Importar Tablas», Figura 4.130.

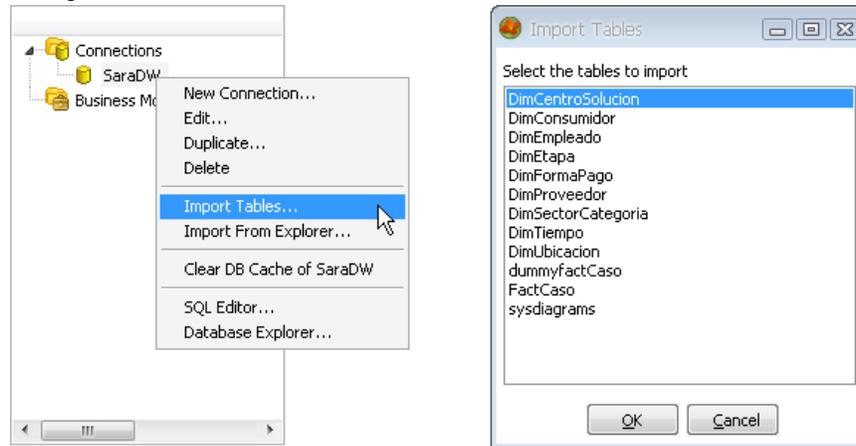


Figura 4.130: Abriendo tabla Import Tables

En el cuadro de Dialogo **Import Tables** aparecerán todas las tablas de la Base de Datos a la cual se realizo la conexión. Se selecciona las tablas que se desea incluir y se presiona OK. Aparecerán abajo del nodo SaraDW las tablas agregadas, como se muestra en la figura 4.131.

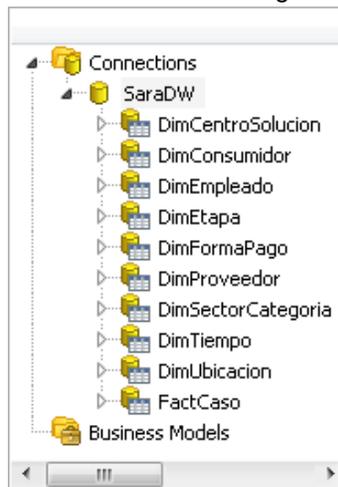


Figura 4.131: Tree Navigator «Árbol de Navegación»

#### 4.2.3.5. Creación Modelo de Negocio

Creando un modelo de negocio:

1. Clic derecho sobre nodo **Business Models**, selecciona **New Business Model...**, del menú desplegable, como se muestra en la Figura 4.132.

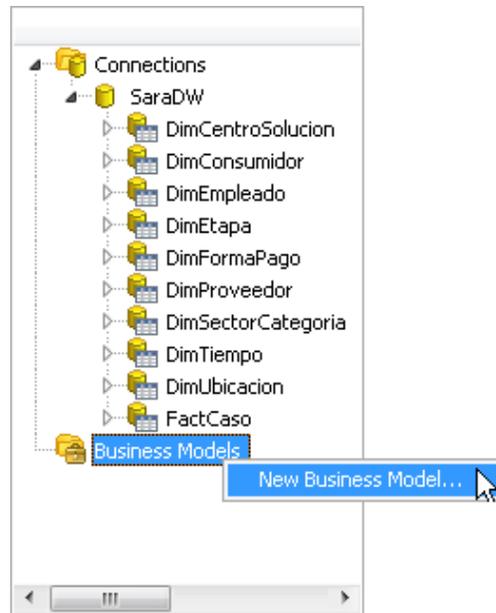


Figura 4.132: Creando un nuevo Modelo de Negocio

2. Se desplegará el cuadro de diálogo **Business Model Properties**, Figura 4.133.

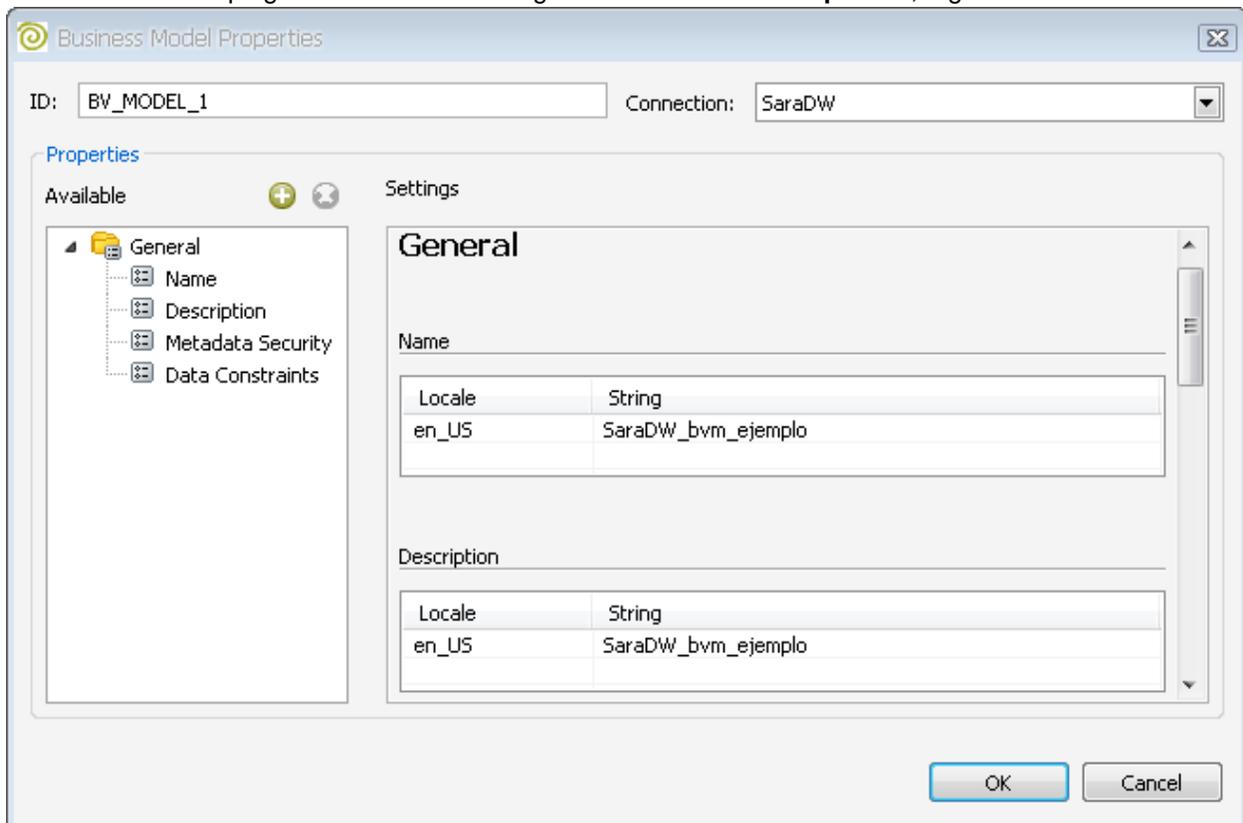


Figura 4.133: Propiedades del Nuevo Modelo de Negocio

3. Al tope del cuadro de diálogo **Business Model Properties** se encuentra un campo para el **ID** del modelo de negocio. El cuadro de dialogo se llena automáticamente con valores. Se

recomienda que se acepten los valor predeterminados, como el **ID**, ya que este valor tiene que ser único entre todos los modelos que se definirán.

4. Para nombrar el nuevo modelo, ingresar el valor en la propiedad **Name**, al lado derecho del cuadro de dialogo.
5. El listado de valores a ingresar para el ejemplo son:
  - ID:** BV\_MODEL\_1
  - Connection:** SaraDW
  - Name:** SaraDW\_bvm\_ejemplo
  - Description:** Ejemplo para DW SARA
6. Al finalizar el ingreso de valores de los parámetros, se presionara **OK** y se mostrar el espacio de trabajo como se muestra en la Figura 4.134. Aparecerá abajo del nodo de **Business Models** el nuevo Modelo de negocio creado.

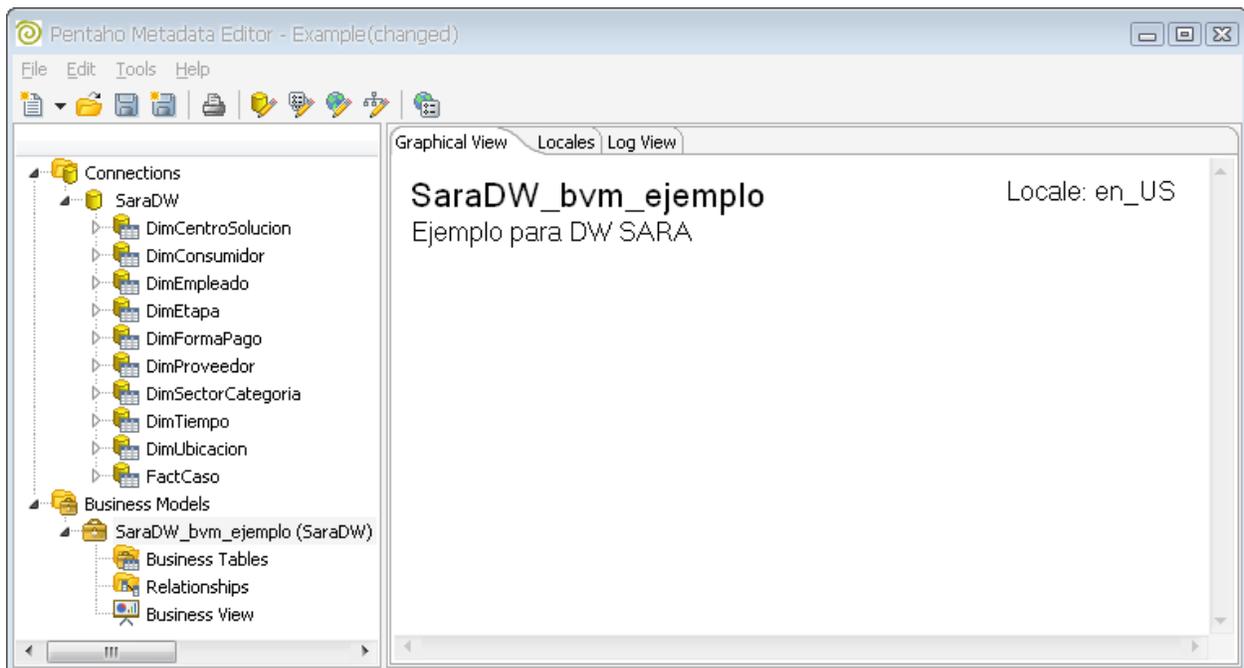


Figura 4.134: Espacio de trabajo Modelo de Negocio

#### 4.2.3.6. Creación de Tablas y Columnas de Negocio

Creando una nueva tabla de negocio:

1. Se deberá arrastrar la tabla que se desee desde el árbol de navegación hacia el **Editor Graph**, haciendo clic sobre la tabla, sosteniéndolo, dirigirse hacia el espacio de trabajo y soltar. Tal como se muestra en la Figura 4.135.
2. Al soltar la tabla sobre **Editor Graph**, se desplegara un cuadro de dialogo de las propiedades de la tabla de negocio, Figura 4.136.

Dentro del cuadro de dialogo **Business Table Properties**, los campos **Name/ID** son predefinidos con valores iniciales, este campo tiene que ser único, por lo que se recomienda aceptar los valores predeterminados para este.

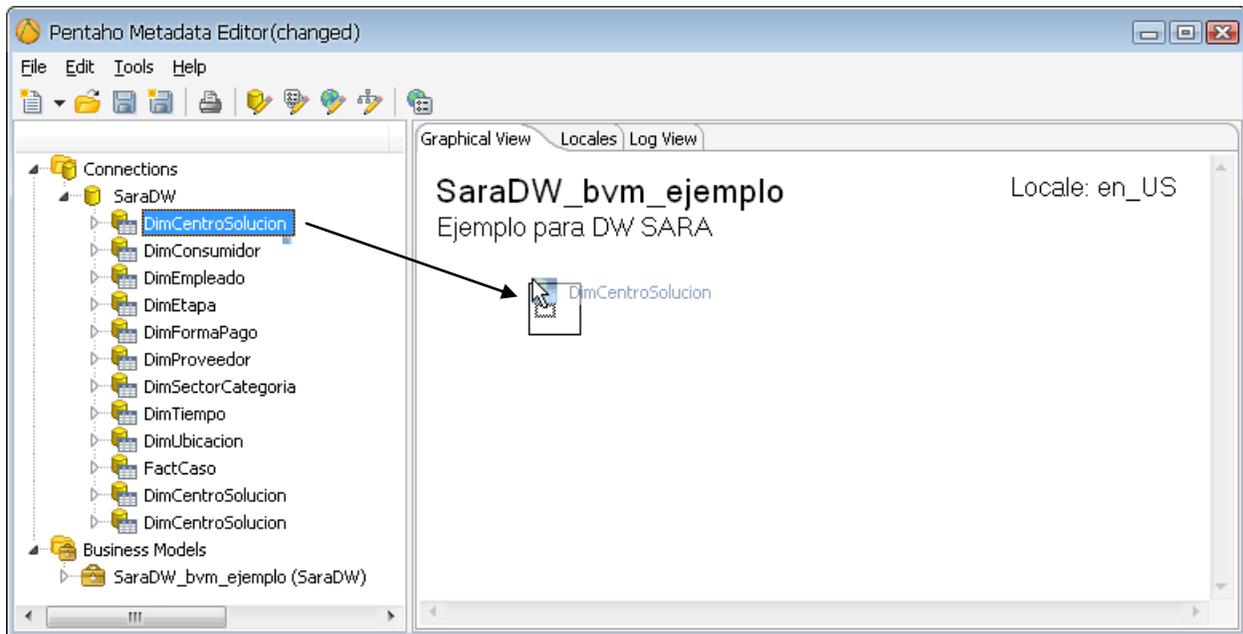


Figura 4.135: Arrastrando una Tabla hacia Editor Graph

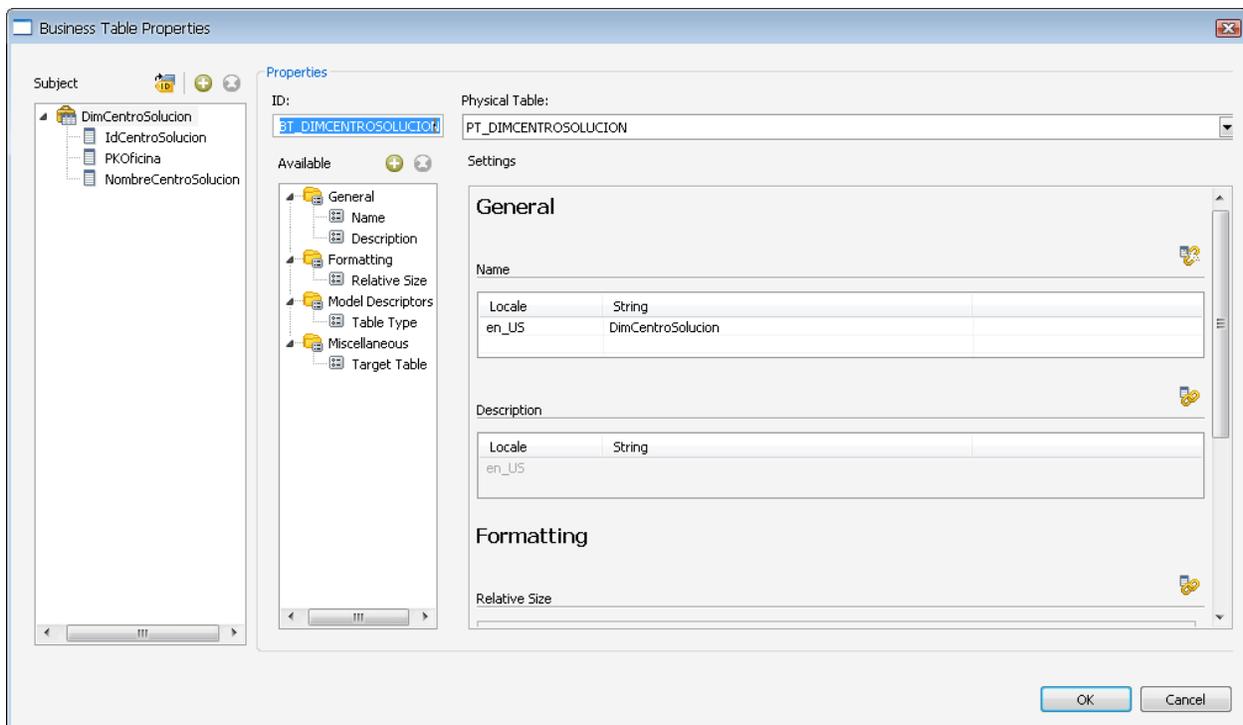


Figura 4.136: Tabla de propiedades de Tablas de negocio

Ajustar columnas de negocio:

1. En el cuadro de dialogo **Business Table Properties**, Figura 4.136, en la parte izquierda dentro de **Subject**, están detalladas las columnas de las tablas, se selecciona la columna que se desea eliminar.
2. Se presiona el botón **Delete Column**  , ubicado en la parte superior de **Subject**.

3. Si se desea agregar una columna eliminada, se presiona el botón **Add New Column** , ubicado en la parte superior de **Subject**.
4. Repetir los pasos anteriores según se desee la información para la tabla de negocio.

#### 4.2.3.7. Creación de Relaciones entre Tablas de Negocio

Una vez se tengan todas las tablas de negocio creadas, será necesario crear relaciones entre las tablas, para que el generador de consultas y el generador de SQL, que vienen incorporados en Pentaho Metadata Editor, puedan crear las consultas de datos correctamente.

Es muy parecido a la diagramación de un diagrama entidad - relación, para mostrar las relaciones de las llaves primarias y llaves foráneas. Para poder crear una relación es importante saber lo siguiente:

1. Cuales dos tablas se desean asociar con la relación.
2. Que columnas en la tabla de negocio identifica la relación.
3. Qué tipo de relación es, una a una, una a muchas, muchas a una, etc.

En la creación de una nueva relación entre tablas de negocios, primero se debe asegurar el seleccionar el modelo al cual se quiere agregar esta relación y el nodo **Relationships** este visible:

1. Clic derecho sobre la rama **Relationships**.
2. Se selecciona New **Relationship...**, del menú desplegable, Figura 4.137.

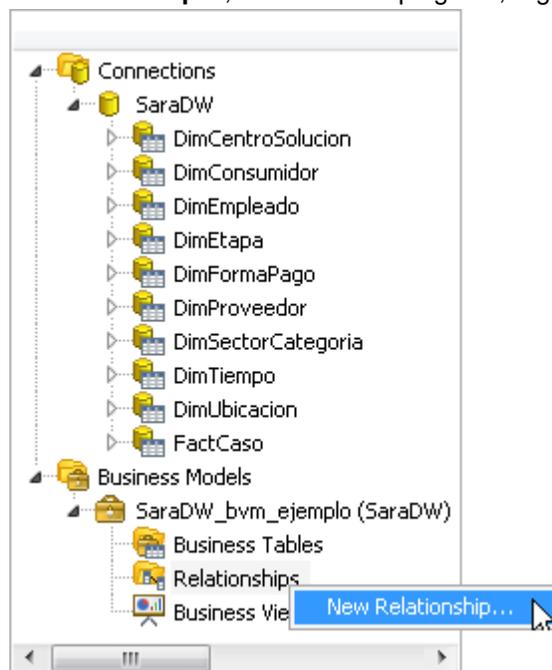


Figura 4.137: Creación de una nueva Relación entre tablas

3. El cuadro de dialogo **Relationship Properties** aparecerá como en la figura 4.138.
4. Se selecciona de la lista **From Table / Field** la tabla de negocio donde inicia la relación.
5. Se selecciona de la lista **To Table / Field** la tabla donde finaliza la relación.
6. Se deberá especificar las columnas de negocio que identifican esta relación de la lista adyacente, para cada una de las tablas de negocio. Si los nombres de columnas son similares, basta con presionar el botón **Guess Matching Field**, dejando que el cuadro de dialogo determine las columnas.
7. Siguiente paso es definir la relación desde la lista desplegable **Relationship**.



Figura 4.138: Propiedades de una Relación

8. Si la relación requiere joins complejos, se selecciona **complex join** en el checkbox, luego se introduce el SQL dentro del recuadro de texto.
9. Clic **OK** para cerrar el cuadro de dialogo. Se visualizara dentro de **Editor Graph** una nueva línea, relacionando dos tablas, y se agrego la relación bajo el nodo **Relationship**, figura 4.139.

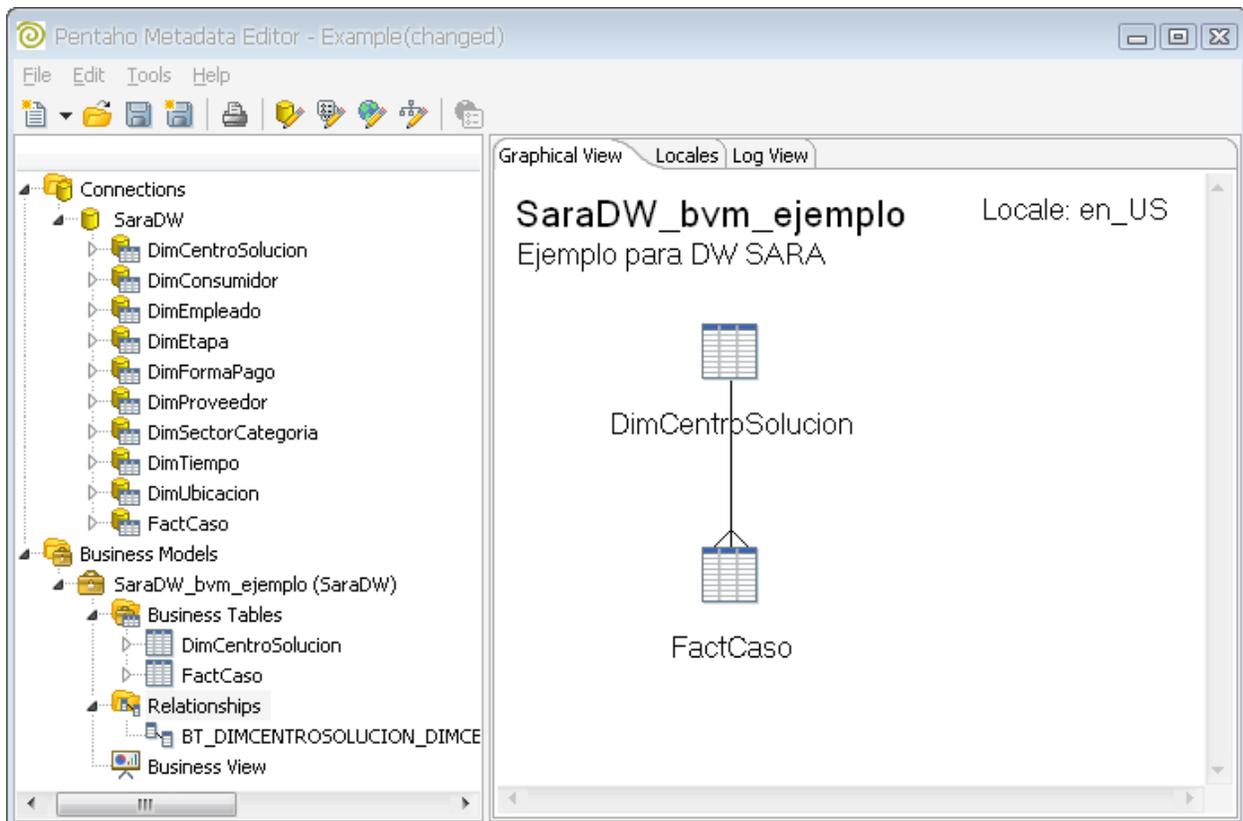


Figura 4.139: Relación entre dos Tablas

#### 4.2.3.8. Construyendo Vista de Negocio

La vista de negocio es la colección de categorías de negocio que representan la vista de un modelo que comúnmente es el entendido por los usuarios finales. Las vistas de negocio son abstracciones lógicas de la organización del negocio.

Una categoría de negocio es un contenedor con nombre con el fin de agrupar o reagrupar columnas de negocio dentro de él. Pueden imitar una tabla de negocio. Las categorías no poseen metadata asociado a ellos, no tiene ninguna atadura hacia las tablas de negocio, teniendo un simple propósito de permitir almacenar las columnas de negocio en un modelo lo mas intuitivo posible para usuarios finales.

La construcción de las vistas de negocio consiste en crear las categorías, luego mover las columnas de negocio desde las tablas de negocio hacia las categorías. Es posible mover columnas desde diferentes tablas de negocio hacia la misma categoría, incluso duplicar una misma columna de negocio dentro de diferentes categorías. **Editor Graph** únicamente representa la parte de las tablas de negocio del modelo, únicamente se utilizara el árbol de navegación de elemento y el editor de categorías para crear las vistas.

Crear una nueva categoría:

1. Clic derecho sobre la rama **Business View**.
2. Seleccionar la opción **New Category...**, del menú desplegable, figura 4.140.

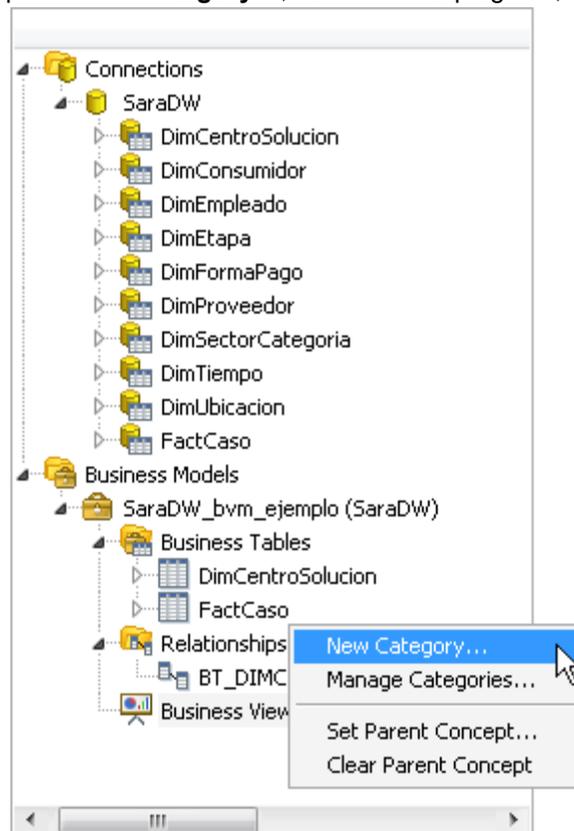


Figura 4.140: Creando una Nueva Categoría

3. Aparecerá el cuadro de dialogo **Category Properties**, figura 4.141.
4. Se ingresa un nombre a la categoría en el campo **Name/ID**.
5. Se selecciona **OK**.

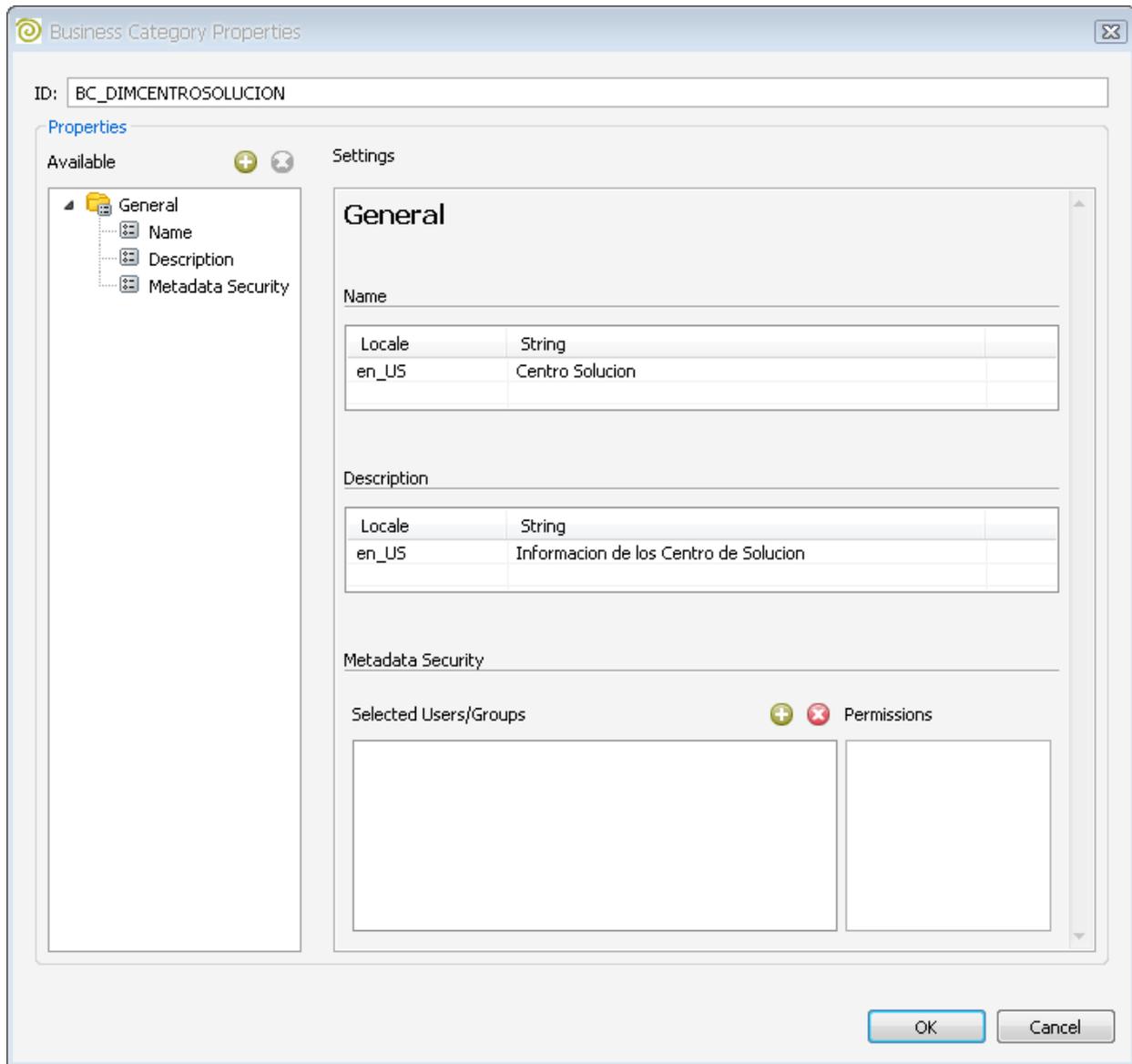


Figura 4.141: Propiedades de Categoría

#### 4.2.3.9. Agregando Seguridad a Metadata Objetos de Negocio

Pentaho MetaData provee información de seguridad en sus propiedades que permite definir seguridad a nivel de tablas o columnas que Pentaho BI Server. Antes de utilizar estas propiedades es necesario hacer saber a Pentaho Metadata Editor sobre Pentaho BI Server, así podrá extraer la lista de usuarios, roles y listas de acceso controlado necesarios.

La información de seguridad a utilizar en los modelos, serán extraídos desde Pentaho BI Server. Es necesario tener disponible los siguientes elementos:

- La URL base de Pentaho BI Server.
- Una credencial de administrador para el servidor.

Configuración del servicio de seguridad:

1. menú principal se selecciona la opción **Security Service**, del menu Tools, figura 4.142.

2. Se visualizará el cuadro de diálogo **Security Service**, como se muestra en la figura 4.143.

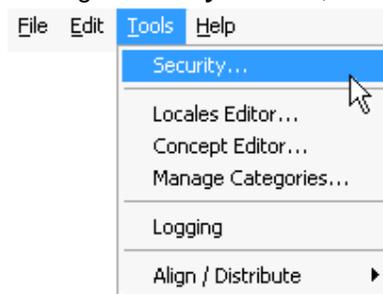


Figura 4.142: Abriendo cuadro de diálogo Security Service

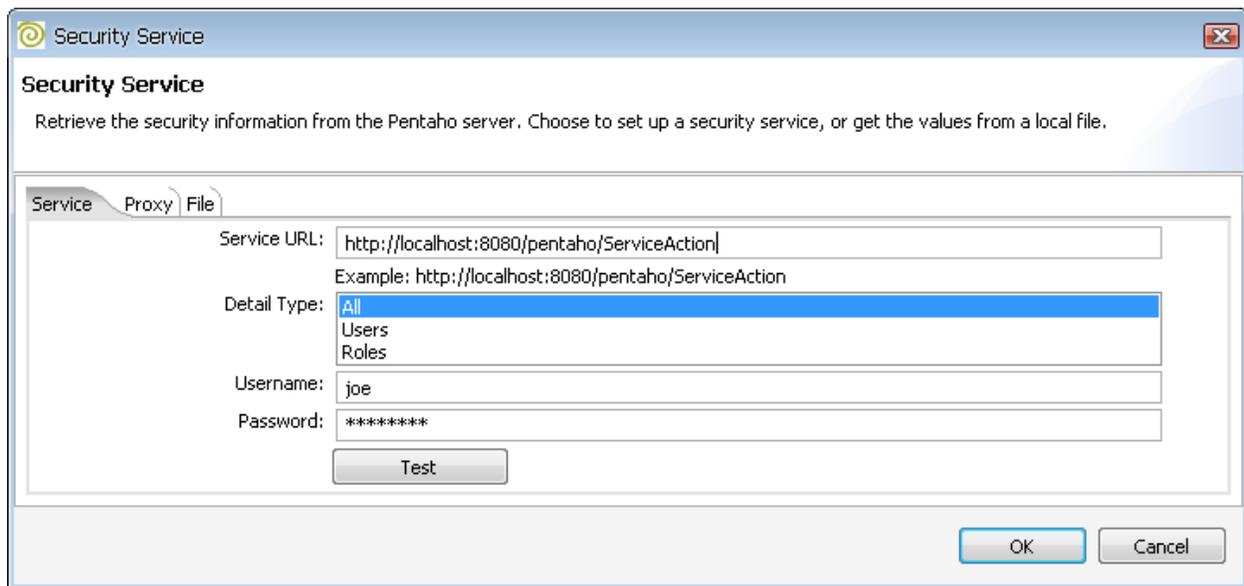


Figura 4.143: Cuadro de diálogo Security Service

3. Los elementos que componen el cuadro de diálogo **Security Service** son:
- **Service URL:** Es la ubicación del servicio a utilizar, está compuesto por la URL base y el nombre del servicio.
  - **Detail Type:** Detalla los diferentes tipos de detalle que se podrán extraer de datos, los cuales pueden ser:
    - **All** (toda la información; Users y Roles).
    - **Users** (únicamente usuarios).
    - **Roles** (únicamente roles).
  - **Username:** El nombre de usuario a utilizar para la autenticación de que solicita Pentaho BI Server. Este usuario deberá poseer privilegios de administrador dentro del Servidor.
  - **Password:** Clave del usuario que se utilizará para el logeo a Pentaho BI server.

Los valores de los diferentes campos serán:

- **Service URL:** http://localhost:8080/pentaho/ServiceAction
  - **Detail Type:** All
  - **Username:** joe
  - **Password:** password
4. Realizar la prueba de configuración haciendo clic sobre **Test**, se deberá visualizar un mensaje similar al presentado en la figura 4.144.

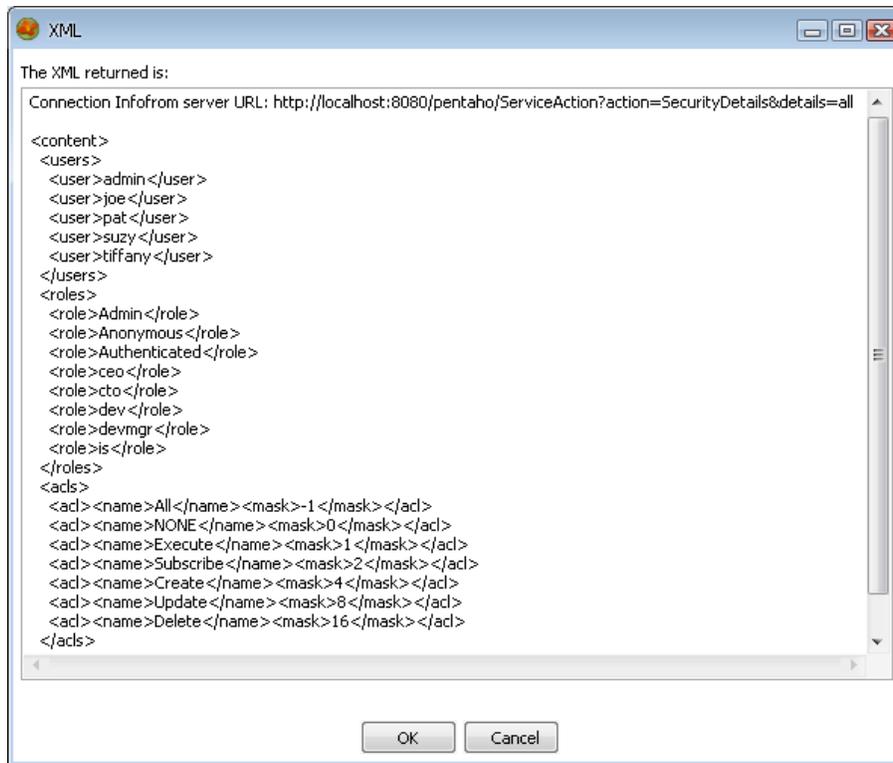


Figura 4.144: Comprobación de la conexión a Pentaho BI Server

#### 4.2.3.10. Modificando Security Constraints

Para agregar un nuevo **security constraints** especificando una tabla o columna de negocio:

1. Clic derecho sobre la tabla de negocio, se selecciona la opción **Edit..**, del menú desplegable, figura 4.145.

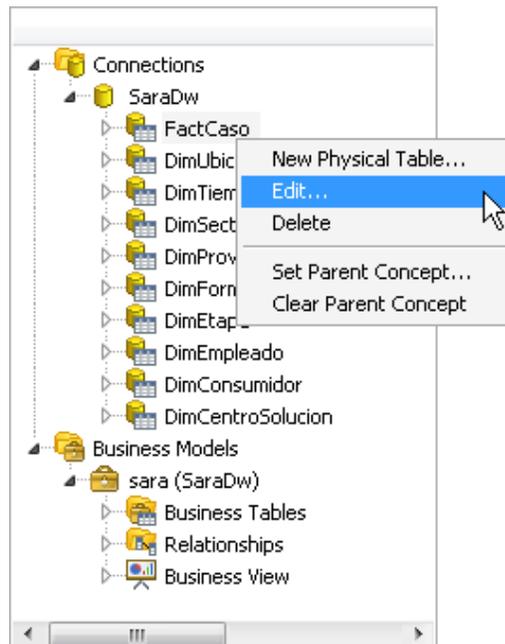


Figura 4.145: Abriendo el cuadro de dialogo Business Table Properties

2. Aparecerá el cuadro de dialogo **Business Table Properties**, figura 4.146.

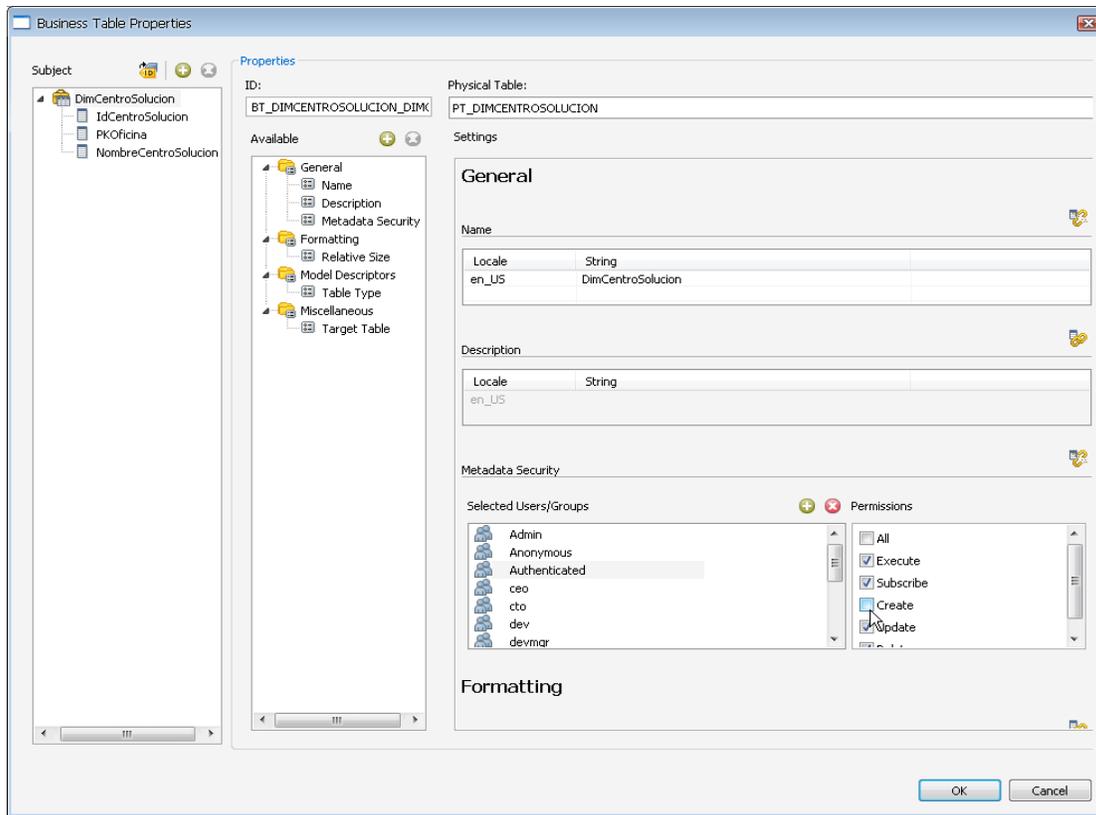


Figura 4.146: Cuadro de dialogo Business Table Properties

3. Dentro del cuadro de dialogo se presiona el botón **override**  para poder modificar los valores de **Metadata Security**.
4. Se seleccionan o se deseleccionan los permisos que se desean asociar a un rol o usuario en específico.
5. Se presiona **OK** para guardar los cambios y finalizar.

#### 4.2.4. Publicación en Pentaho BI Server

Del menú principal, ingresar a **File > Publish to Server...**, como se muestra en la figura 4.147.

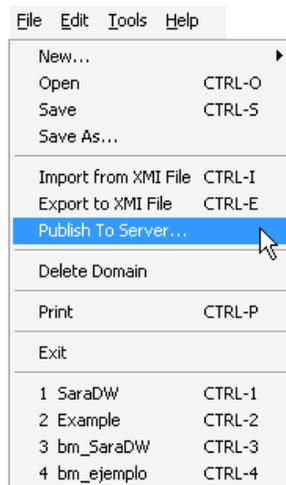
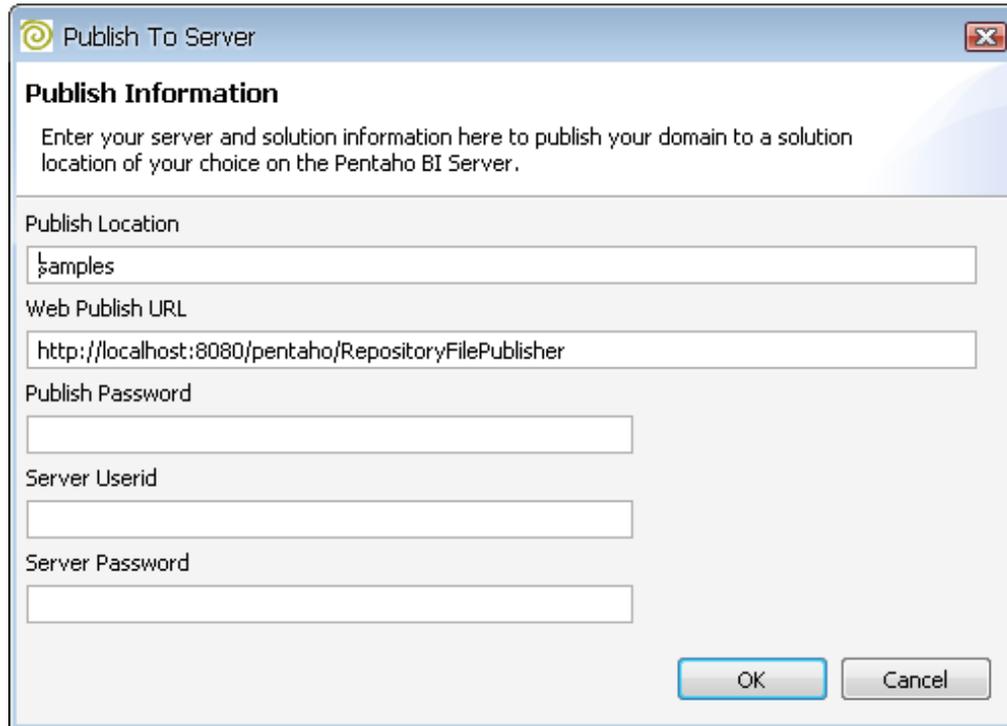


Figura 4.147: Publicación de Dominio

Al seleccionar **Publish...**, aparecerá el cuadro de dialogo Publish To Server, como se muestra en la figura 4.148,



**Figura 4.148: Cuadro de dialogo Publish To Server**

Los datos de logeo son:

**Publish Location**, Nombre de la solución: SARA

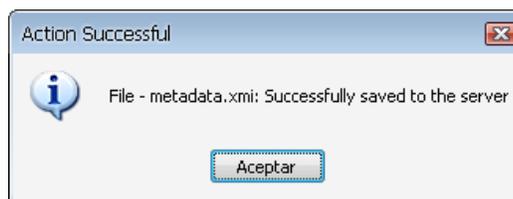
**Web Publish URL**, URL base de Pentaho BI Server, concatenado con el nombre del servicio de publicación: http://localhost:8080/pentaho/RepositoryFilePublisher

**Publish Password**, Password establecida para permitir publicar: admin

**Server Uderid**, Usuario con rol de administrador: joe

**Server Password**, Password de usuario: password

Luego de ingresar la información al cuadro de dialogo Publish To Server, se presionara el botón **OK**, luego se desplegara un cuadro que confirmara que la publicación se realizo con éxito, como se muestra en la figura 4.149.



**Figura 4.149: Notificación publicación con éxito**

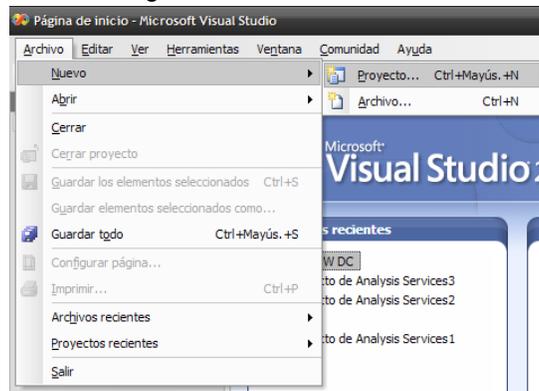
## 4.2.5. Integration Services

### 4.2.5.1. Creación de un paquete Etl con Integration Services 2005

Para iniciar proceso es necesario tener instalado el SQL Server Business Intelligence, por lo que iniciaremos dicho proceso partiendo del punto que ya se tienen configuradas las bases tanto la de origen como la de destino.

**Paso 1:** Crear un Nuevo proyecto.

Se debe de elegir la opción según se muestra en pantalla, ingresando al menu Archivo -> Nuevo -> Proyecto. Esta opción se muestra en la Figura 4.150



**Figura 4.150: Pantalla de inicio de SQL Server Business Intelligence Studio**

Al ingresar a esa opción nos presentara el siguiente cuadro de dialogo en el cual definiremos la información básica del proyecto que crearemos.

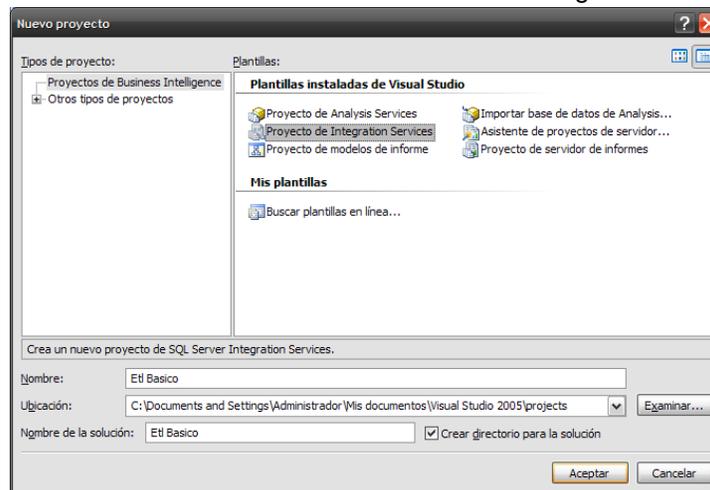
Debemos de elegir la plantilla “Proyecto de Integration Services”, y a continuación definir la información general del proyecto; para el caso se nos presenta las opciones para elegir el nombre del proyecto, ubicación y nombre de la solución. Estas opciones serán llenadas con la información presentada a continuación:

Nombre del proyecto: ETL Basico.

Ubicación: cualquiera.

Nombre de la Solución: ETL Basico.

Esta información puede ser verificada tal como se muestra en la Figura 4.151:



**Figura 4.151: Cuadro de dialogo de nuevo proyecto**

Al tener lleno este cuadro de diálogo se puede proceder a hacer clic en botón aceptar. Luego nos presentará el área de trabajo tal como se muestra en la Figura 4.152:

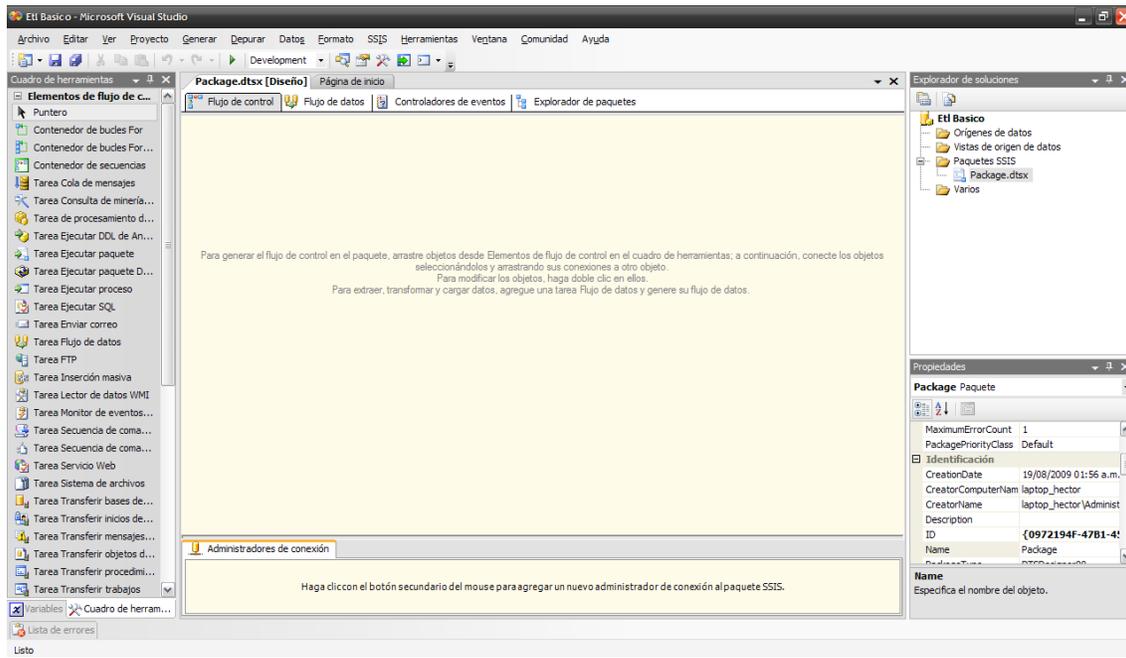


Figura 4.152: Entorno de trabajo SQL Integration Services

**Paso 2:** Creación de un paquete con el asistente.

La forma más sencilla de realizar un paquete es utilizando el asistente. Es por ello que lo realizaremos de esta forma, a fin de que se facilite la creación de paquetes. Para iniciar el asistente debemos acceder al menú “Proyecto”, y buscar la opción “Asistente para importación y exportación de SSIS”, esta opción se puede ver en la figura 4.153.

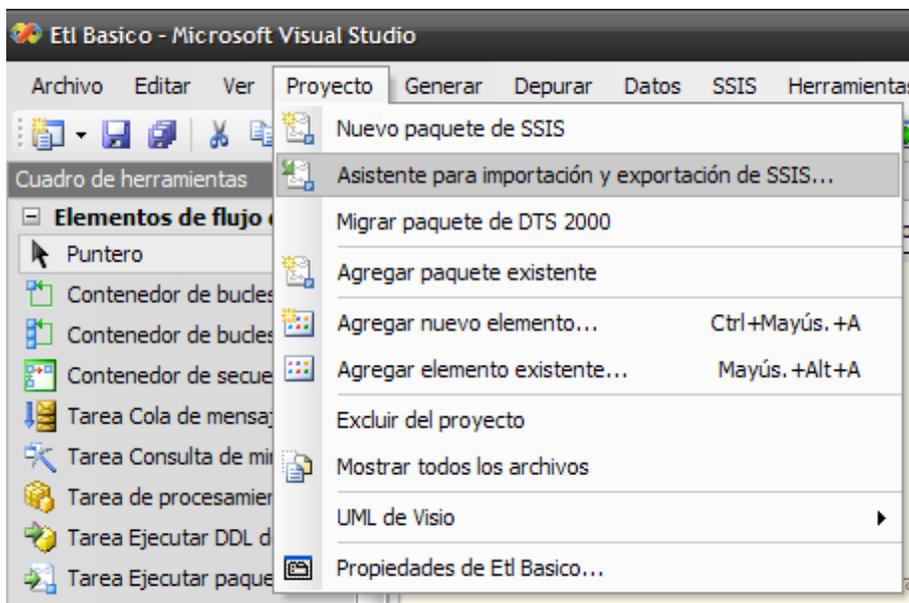


Figura 4.153: Opción para crear el paquete con el asistente

Al ingresar a esa opción, nos aparecerá un cuadro de dialogo como el siguiente:



Figura 4.154: Cuadro de Dialogo Inicial del Asistente de Importación y exportación

Al hacer clic en el botón siguiente se muestra un cuadro de dialogo como el de la Figura 4.155:

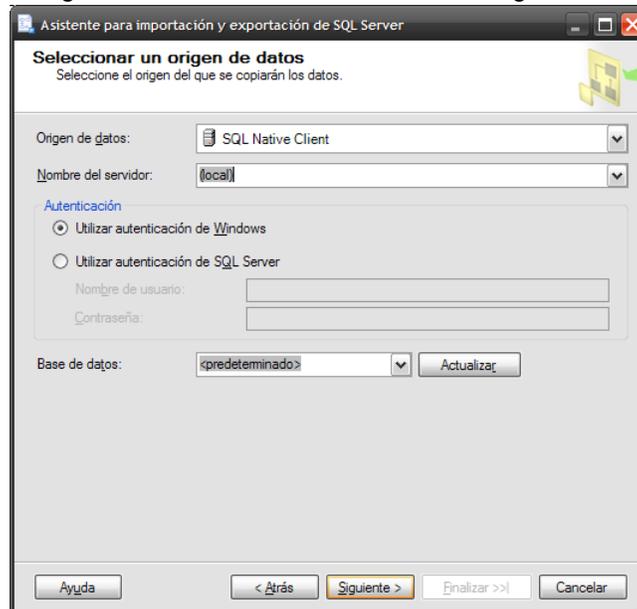


Figura 4.155: Cuadro de dialogo para definir la conexión de origen

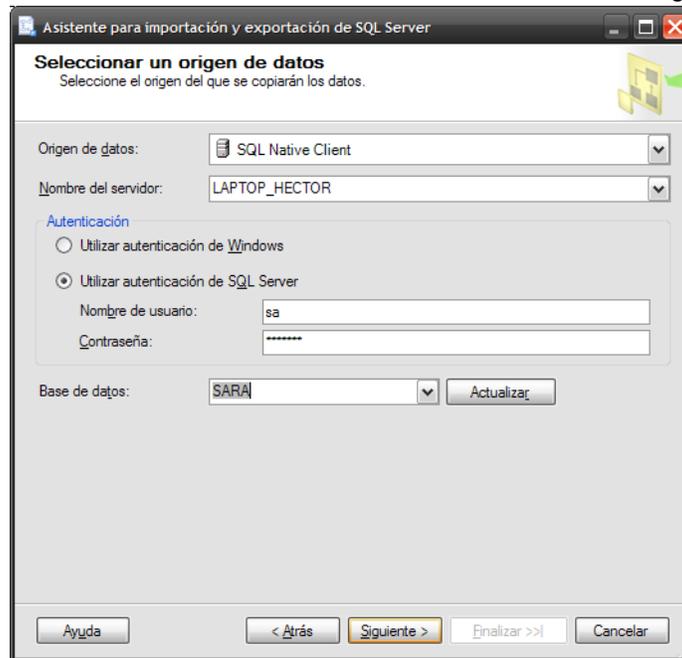
Se debe de ingresar la información del servidor y la base de origen. Como ejemplo tenemos la siguiente información para el servidor de origen:

**Nombre del Servidor:** Laptop\_Hector

**Autenticación:** Utilizar Autenticación de SQL Server

**Usuario:** sa  
**Contraseña:** La contraseña de SA  
**Base de datos:** Sara

Por lo que la pantalla queda completa con la información, tal como se muestra en la siguiente Figura 4.156. Al terminar de llenar esta información se debe de hacer clic en el botón siguiente.



**Figura 4.156: Datos de la Conexión de origen**

A continuación se nos muestra el cuadro de dialogo en el cual se debe de definir la información del servidor destino. Se llenara de la misma forma como se hizo con el servidor de origen. Por lo que se debe de tomar en cuenta la siguiente información:

**Nombre del Servidor:** Laptop\_Hector  
**Autenticación:** Utilizar Autenticación de SQL Server  
**Usuario:** sa  
**Contraseña:** La contraseña de SA  
**Base de datos:** SaraDW

La información es similar ya que por ser una prueba se encuentra montadas las bases de origen y destino en el mismo equipo. La información debe de quedar tal como se muestra en la Figura 4.157.

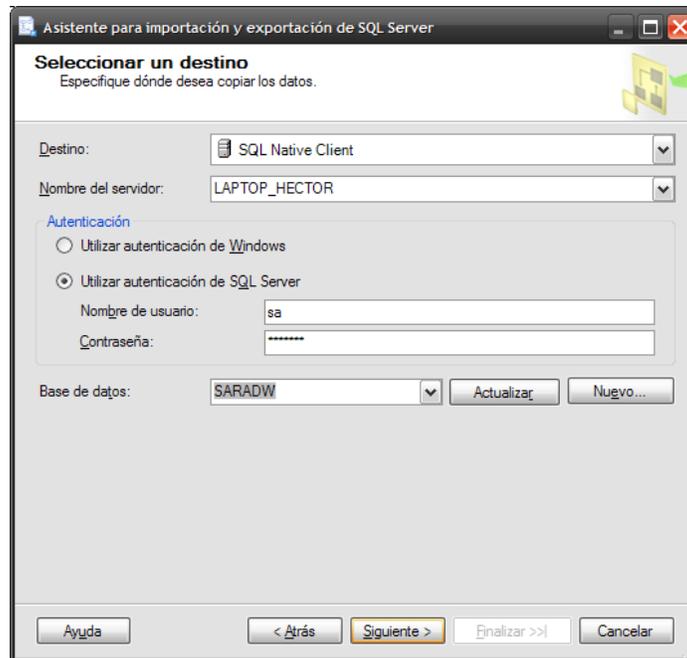


Figura 4.157: Datos para la conexión destino

Al llenar los datos y hacer clic en siguiente se muestra un cuadro de dialogo como el de la Figura 4.158.

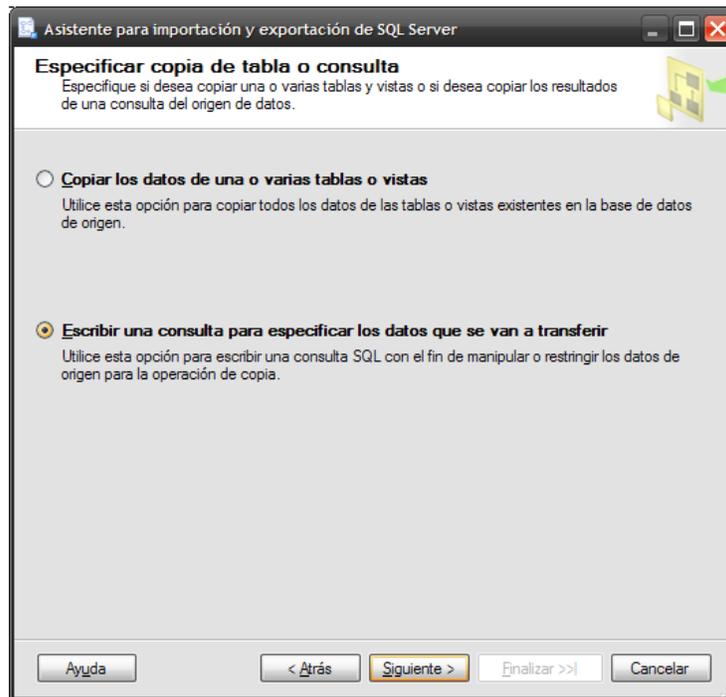
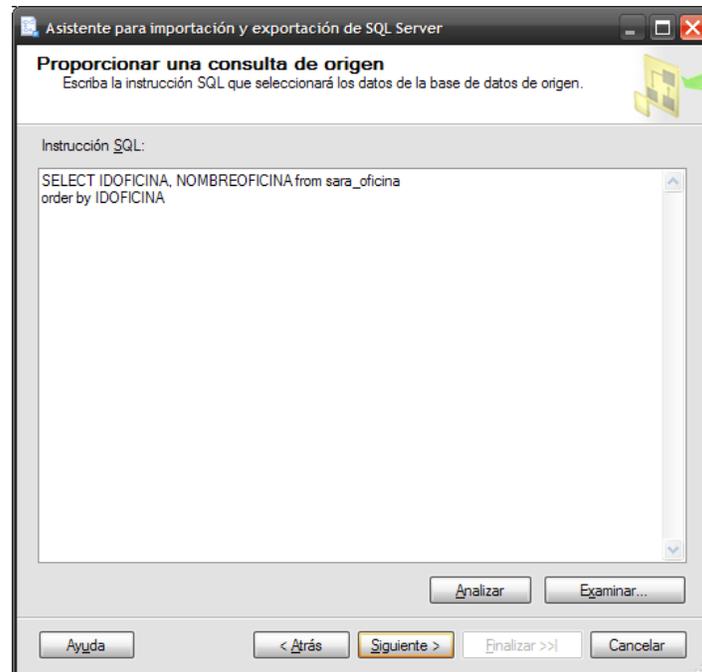


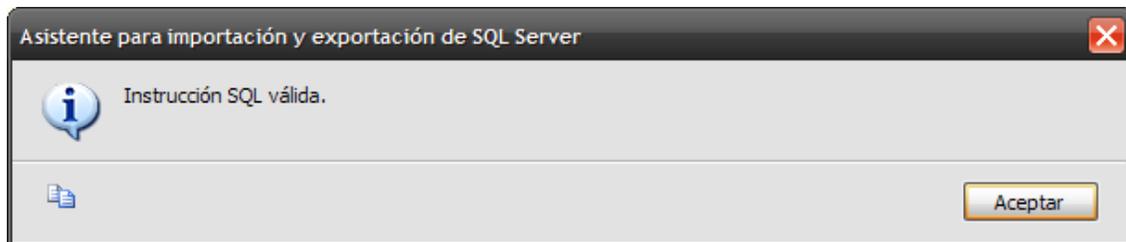
Figura 4.158: Especificación del tipo de acción a realizar

Hacer clic en siguiente, para especificar la consulta que será nuestra fuente de datos, tendremos que llenarla con un SQL, como se muestra en la Figura 4.159



**Figura 4.159: Cuadro de dialogo para especificar consulta de extracción de datos**

Teniendo escrita la consulta, podemos verificar que este correcta en el botón “Analizar”. Si esta correcta nos debe de presentar un mensaje como el de la Figura 4.160 para avanzar se debe hacer clic en el botón aceptar:



**Figura 4.160: Cuadro de dialogo de verificación de consulta**

Esta cuadro de dialogo es muy importante ya que se definen la tabla de origen, que para este caso es una consulta SQL, y el destino que será una dimensión del DW; para realizar bien este proceso es necesario hacer elegir el nombre de la tabla destino, como ejemplo se debe utilizar la tabla “[SARADW].[dbo].[DimCentroSolucion] “. Este proceso se muestra claramente la Figura 4.161.

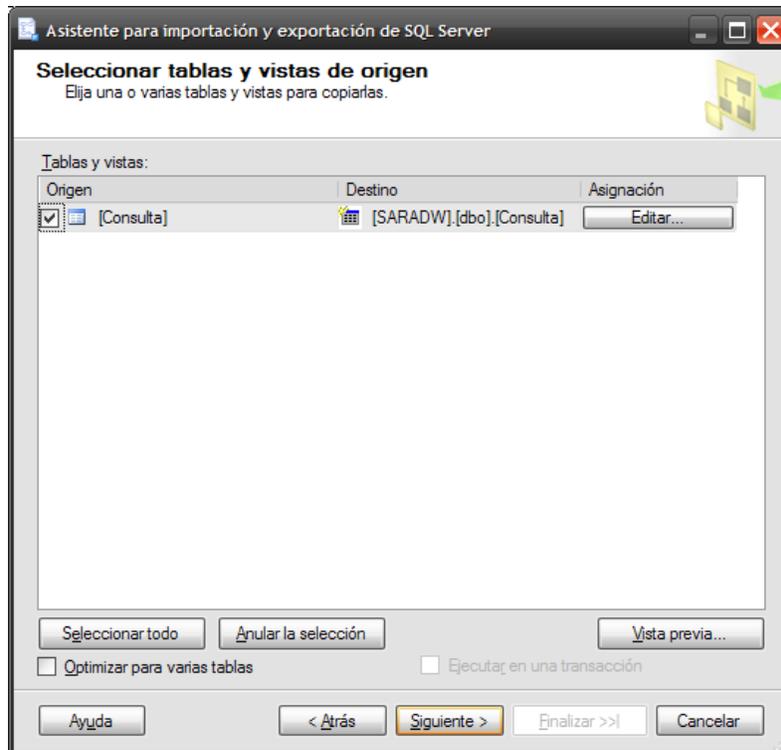


Figura 4.161: Pantalla para elegir el mapeo de los datos de origen con el destino

A continuación debemos hacer clic en editar, y nos presentara una pantalla como la que se muestra en la Figura 4.162:

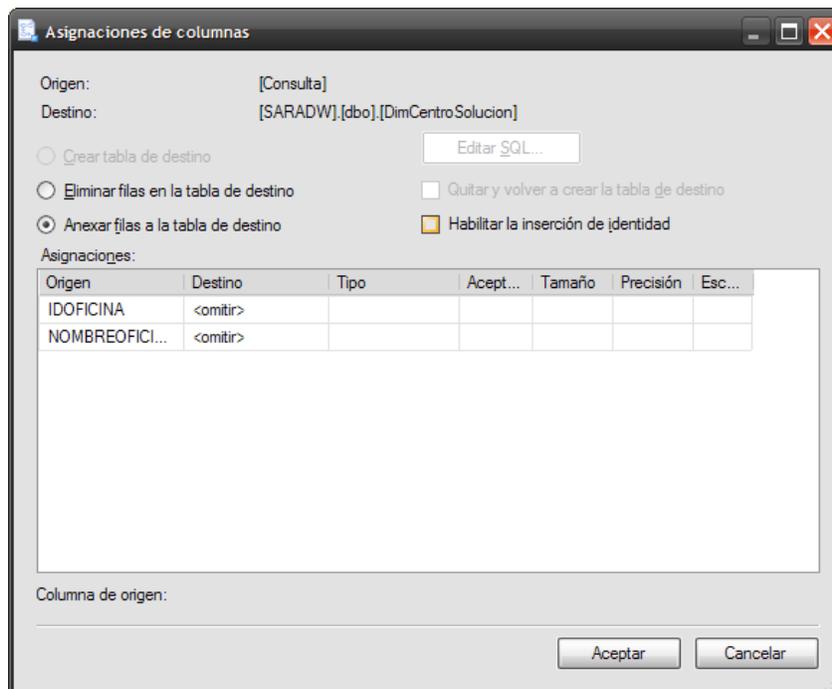
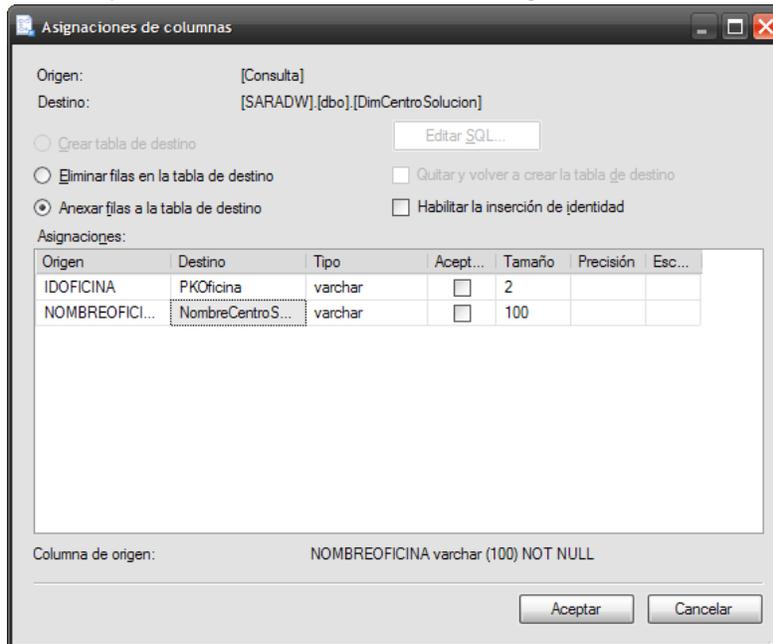


Figura 4.162: Definición de relaciones entre columnas de origen y destino

En muchas ocasiones es necesario definir el nombre de las columnas destino en las que se recibirá la información, además tenemos la opción de definir si la información a ser cargada en la tabla destino, será agregada o se eliminara la que ya existe y se sustituirá por la nueva, esta acción se define en las opciones :

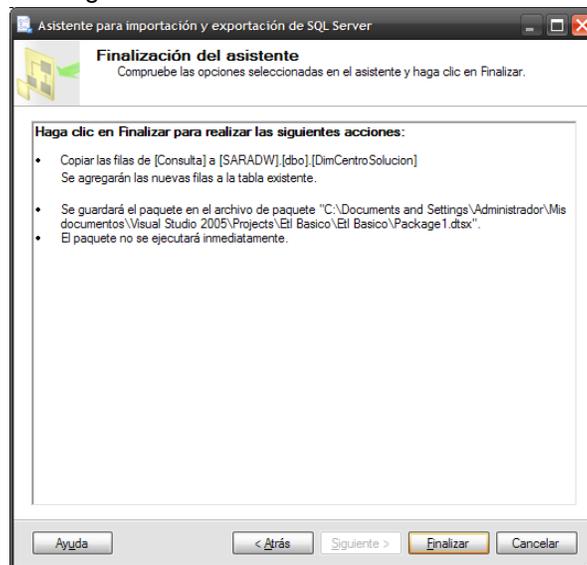
- Eliminar filas en la tabla de destino
- Anexar filas a la tabla de destino

Por lo que nos debe de quedar tal como se muestra en la Figura 4.163:



**Figura 4.163: Columnas de origen y destino relacionadas**

Con esto tenemos listo nuestro primer paquete básico de ETL, por lo que se debe de mostrar una cuadro de dialogo como el de la Figura 4.164:



**Figura 4.164: Resumen de cambios en el Paquete**

Es necesario tomar en cuenta que al crear el paquete las acciones que se muestran en la Figura 4.165 deberán reflejarse con un estado “Correcto”:

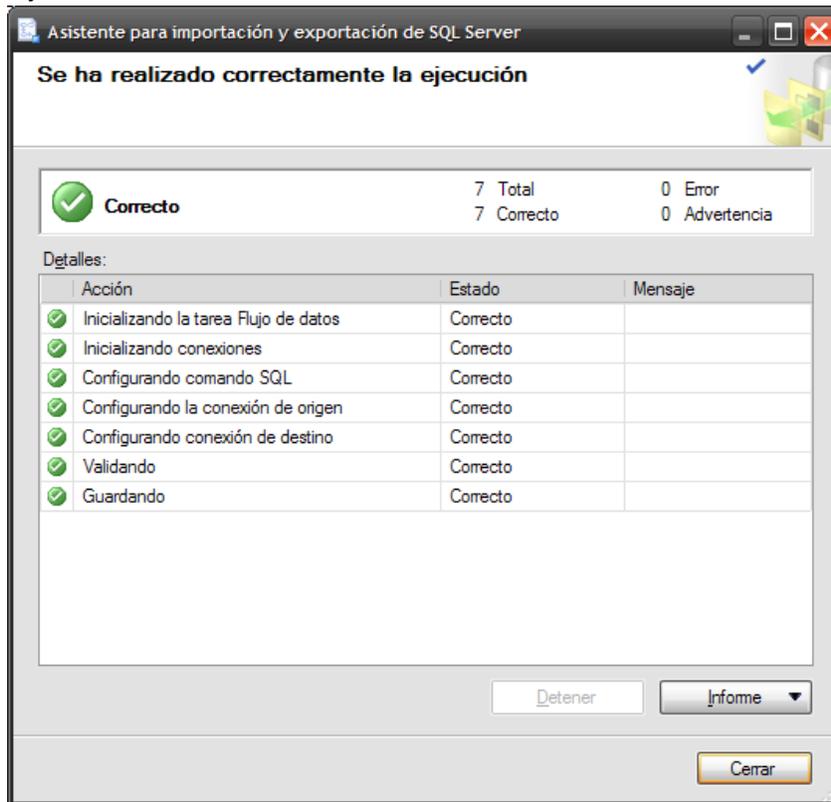


Figura 4.165: Resumen de acciones luego de la creación

Al hacer clic en cerrar nos mostrara una imagen como la que se puede observar en la Figura 4.166, solo nos queda probarlo, realizando una carga de datos.

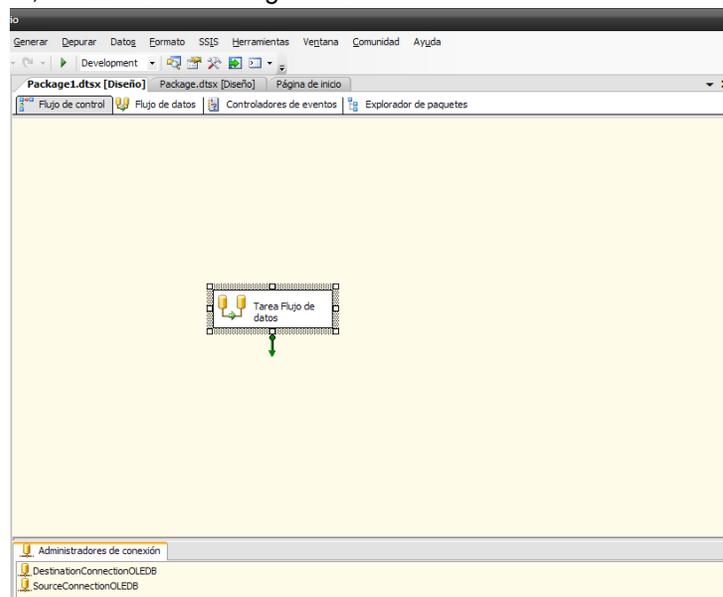
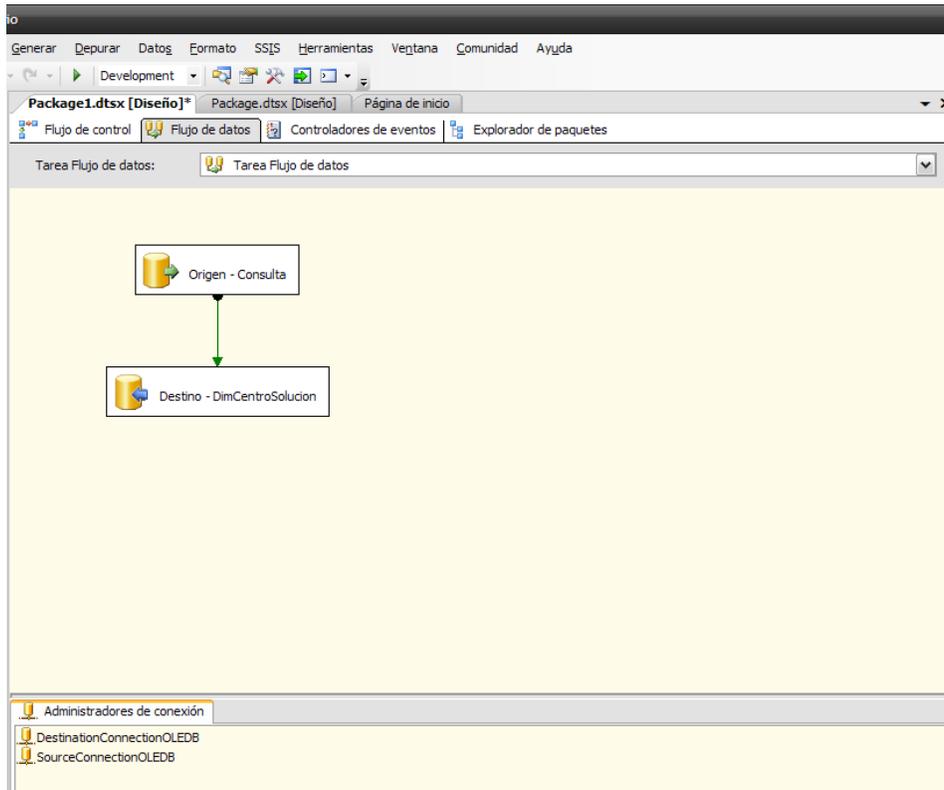


Figura 4.166: Tarea de flujo de datos en el entorno de trabajo

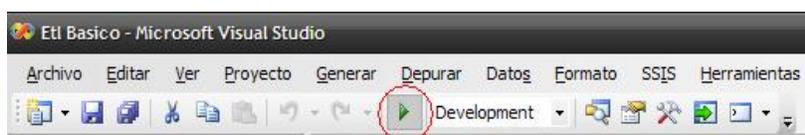
En la opción Flujo de datos se muestran el detalle de las operaciones realizadas en el paquete recién definido, este detalle se puede observar en la Figura 4.167.



**Figura 4.167: Detalle del flujo de datos**

**Paso 3:** Ejecutar el paquete ETL.

Teniendo creado el paquete solo es necesario hacer clic en el botón ejecutar mostrado en la Figura 4.168.



**Figura 4.168: Botón ejecutar Paquete**

Al ejecutar esta opción, se muestra el resumen de las transacciones que han sido cargadas al destino fijado, este se muestra en la Figura 4.169.

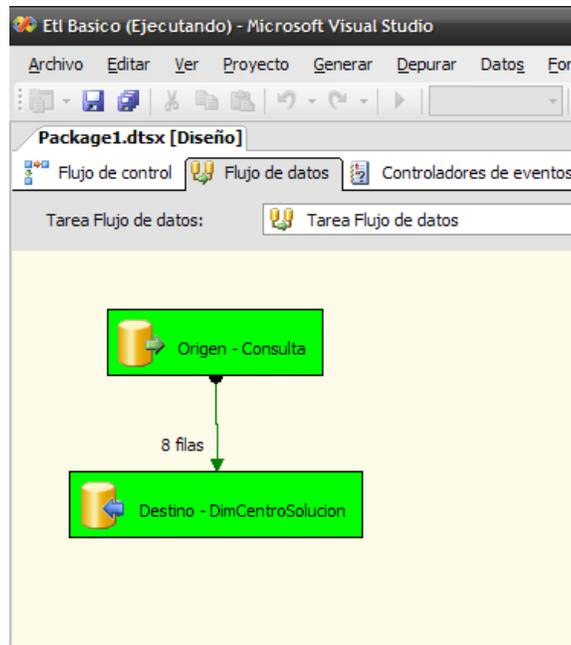


Figura 4.169: Resultado de ejecutar el paquete

Además nos brinda la opción de verificar el log, del resultado de la carga realizada tal como se muestra en la Figura 4.170 Y con esto se termina la creación de un paquete sencillo en SQL Integration Services 2005.

Figura 4.170: Registro de la ejecución del proceso

### 4.2.5.2. Generar paquete de Implementación

En el explorador de soluciones de Integration Services hacer clic derecho sobre el area marcada a continuacion.

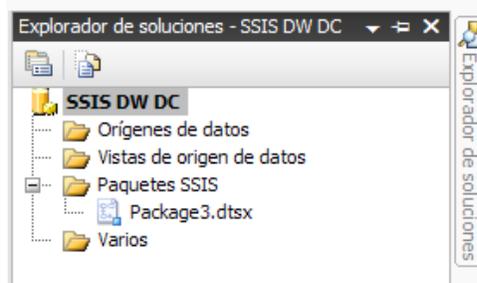


Figura 4.171: explorador de soluciones

Al hacer clic derecho nos mostrara el siguiente menú contextual, en el cual elijeremos la opción propiedades.

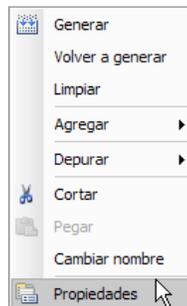


Figura 4.172: Menú de propiedades de la solución

A continuación se muestra una ventana en la cual se muestran las opciones disponibles para el proyecto.

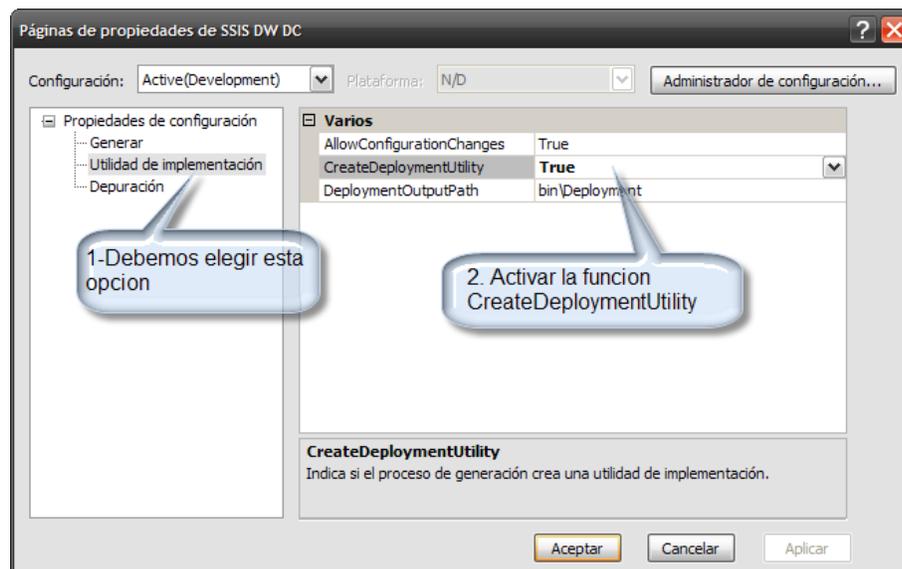


Figura 4.173: Ventana de propiedades del proyecto

A continuación debemos generar de nuevo el proyecto. Tal como se muestra en la siguiente figura.

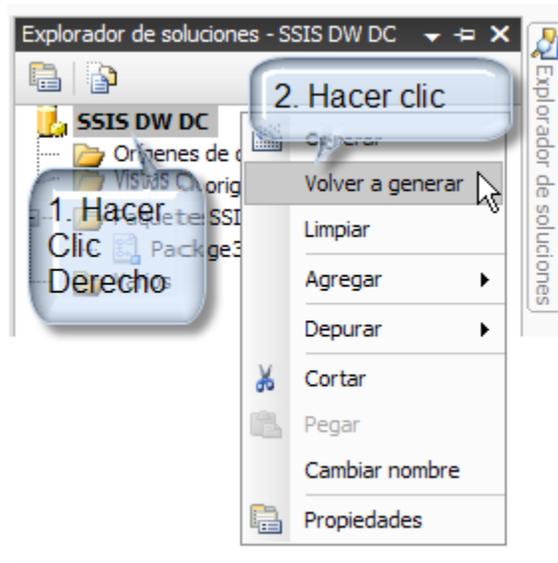


Figura 4.174: Explorador de soluciones

Si todo sale bien, en la parte inferior izquierda de la pantalla nos mostrara un mensaje en el cual nos indica que todo finalizo correctamente. Dicho mensaje se muestra a continuación:

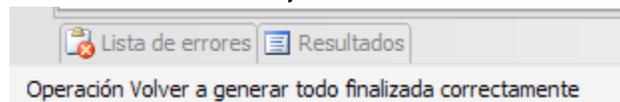


Figura 4.175: Barra de estado

Si ha usado la ubicación predeterminada, el paquete de implementación está en la carpeta Bin\Deployment dentro de la carpeta Deployment de la solución que pasa el caso se llama SSIS DW DC.

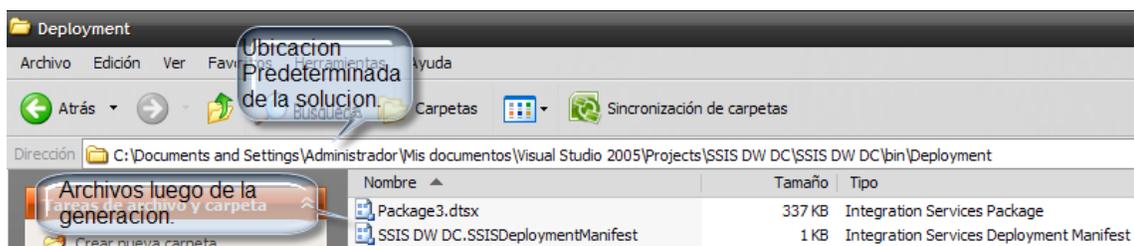


Figura 4.176: Ubicación del proyecto

Con esto tendremos el archivo listo para poder implementarse.

#### 4.2.5.3. Ejecutar paquete para la carga de datos.

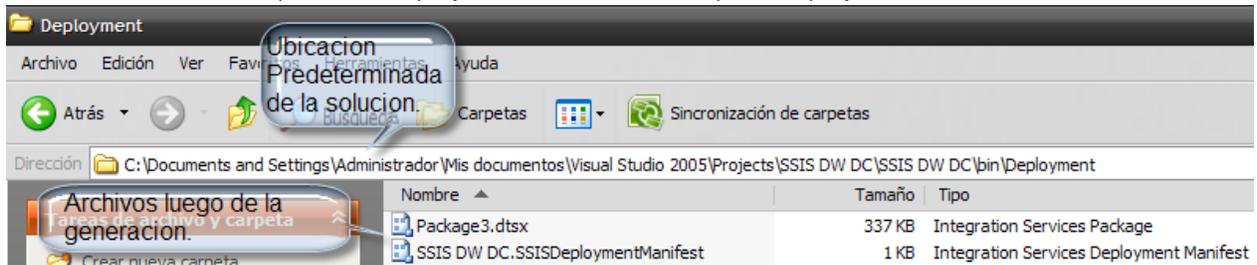
Para realizar este proceso es necesario realizar 3 tareas fundamentales:

- Tarea 1: Copiar el paquete de implementación
- Tarea 2: Ejecutar el Asistente para la instalación de paquetes
- Tarea 3: Probar los paquetes implementados

A continuación se detalla cada una de las tareas.

**4.2.5.4. Tarea 1: Copiar el paquete de implementación.**

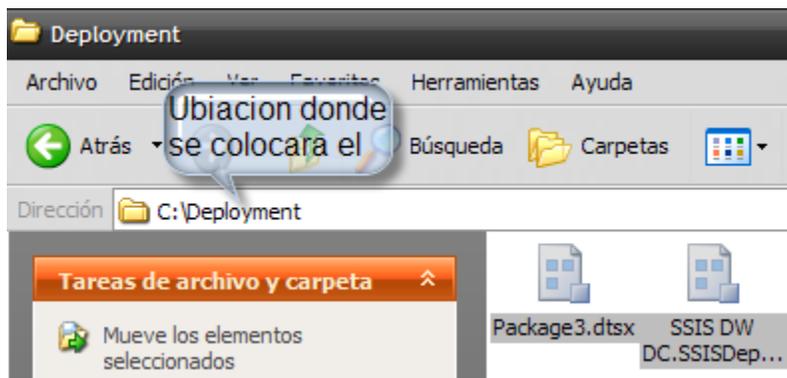
1. Busque el paquete de implementación en el equipo donde se genero el paquete.
  - Si ha usado la ubicación predeterminada, el paquete de implementación está en la carpeta Bin\Deployment dentro de la carpeta Deployment Tutorial.



**Figura 4.177: ubicación del archivo para la ejecución**

Haga clic con el botón secundario en la carpeta Deployment y, a continuación, en Copiar.

Busque el recurso donde desea copiar la carpeta en el equipo de destino y haga clic en Pegar.



**Figura 4.178: Ubicación destino lo lugar donde se correrá el proceso**

**4.2.5.5. Tarea 2: Ejecutar el Asistente para la instalación de paquetes**

2. En la carpeta Deployment, haga doble clic en el archivo de manifiesto: SSIS DW DC.SSISDeploymentManifest.
 

Luego del doble clic se mostrara la pantalla de bienvenida del Asistente para la instalación de paquetes, haga clic en **Siguiente**.

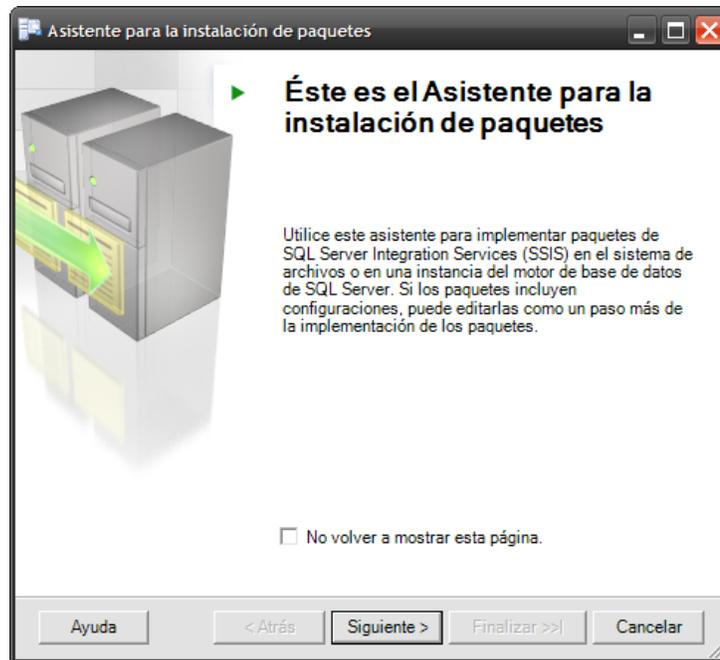


Figura 4.179: ubicación del archivo para la ejecución

3. En la página Implementar paquetes SSIS, seleccione la opción **Implementación en SQL Server**, active la casilla de verificación **Validar los paquetes después de la instalación** y, a continuación, haga clic en **Siguiente**.

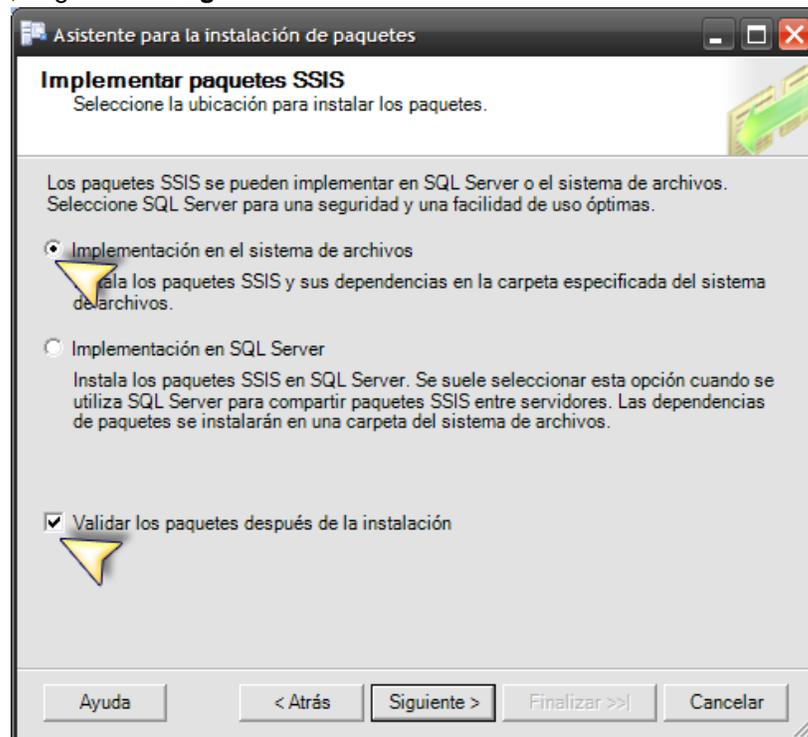
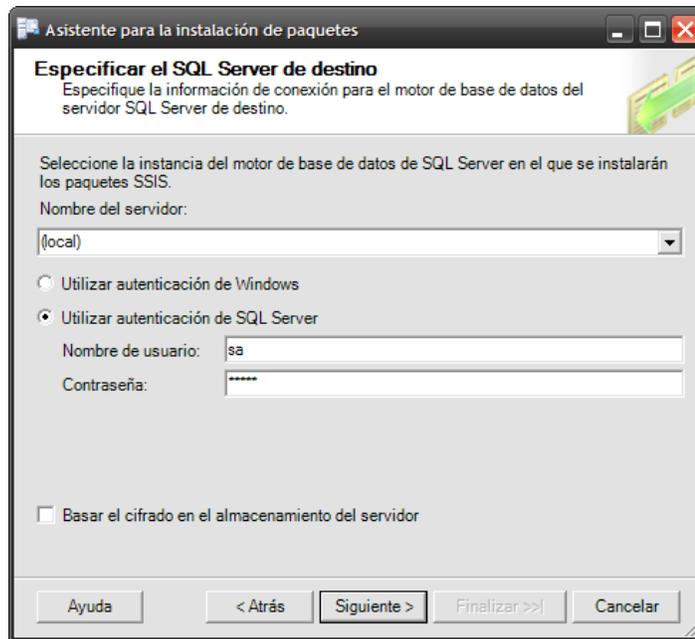


Figura 4.180: Asistente para la instalación del paquete

4. En la página Especificar el SQL Server de destino, especifique (local) en el cuadro Nombre del servidor.

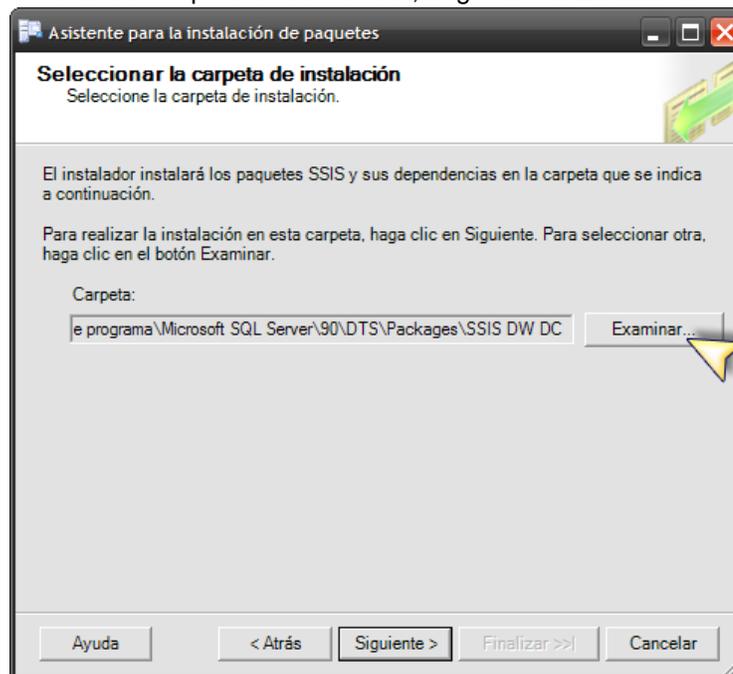


**Figura 4.181: Configuraciones para especificar el destino del proceso**

Si la instancia de SQL Server admite la autenticación de Windows, seleccione Utilizar autenticación de Windows; en caso contrario, seleccione Utilizar autenticación de SQL Server y proporcione un nombre de usuario y una contraseña.

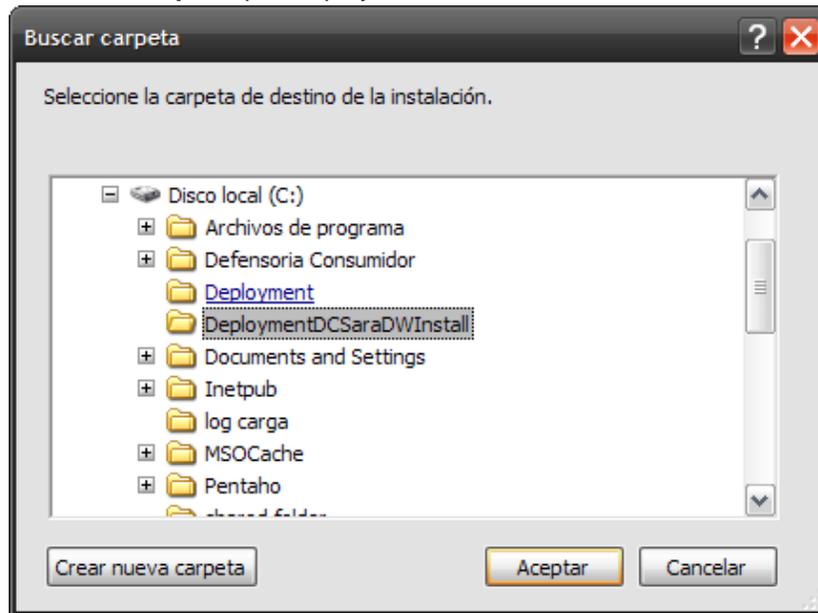
Compruebe que la casilla de verificación Basar el cifrado en el almacenamiento del servidor está desactivada.

5. En la página Seleccionar la carpeta de instalación, haga clic en Examinar.



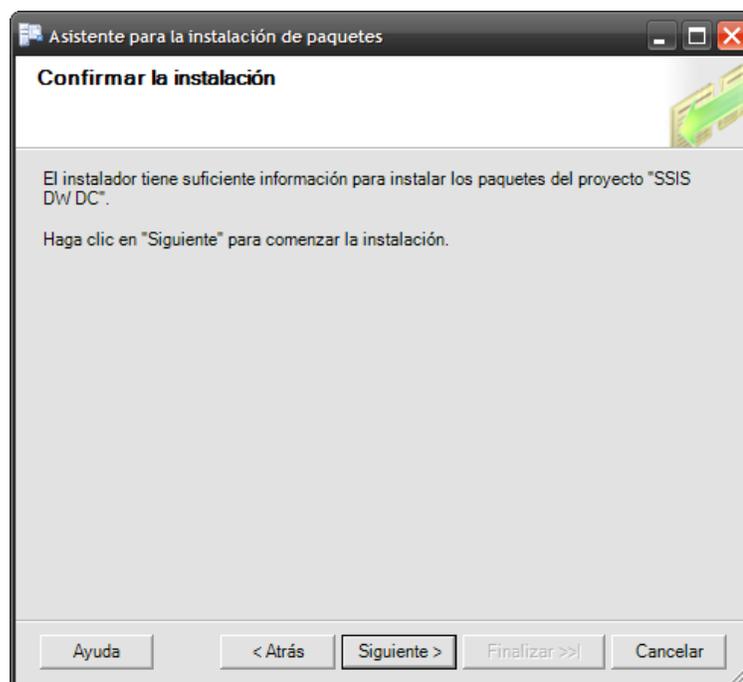
**Figura 4.182: Ubicación destino**

En el cuadro de diálogo Buscar carpeta, expanda Mi PC y, a continuación, haga clic en Disco local (C:). Haga clic en Crear nueva carpeta y reemplace el nombre predeterminado de la nueva carpeta, **Nueva carpeta**, por DeploymentDCSaraDWInstall.



**Figura 4.183: Configuraciones para especificar el destino del proceso**

6. En la página Seleccionar la carpeta de instalación, compruebe que el cuadro Carpeta contiene C:\ DeploymentDCSaraDWInstall y, a continuación, haga clic en Siguiente.



**Figura 4.184: Ventana para confirmar la instalación**

- En la página Confirmar la instalación, haga clic en **Siguiente**. El asistente instala los paquetes. Una vez finalizada la instalación, se abre la página Configurar paquetes.

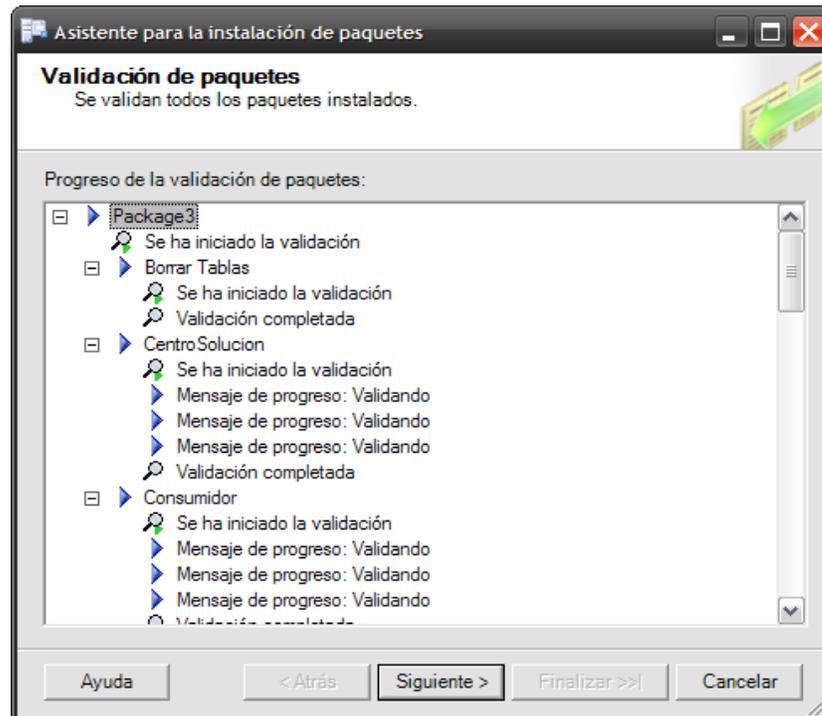


Figura 4.185: Verificación de ejecución del paquete

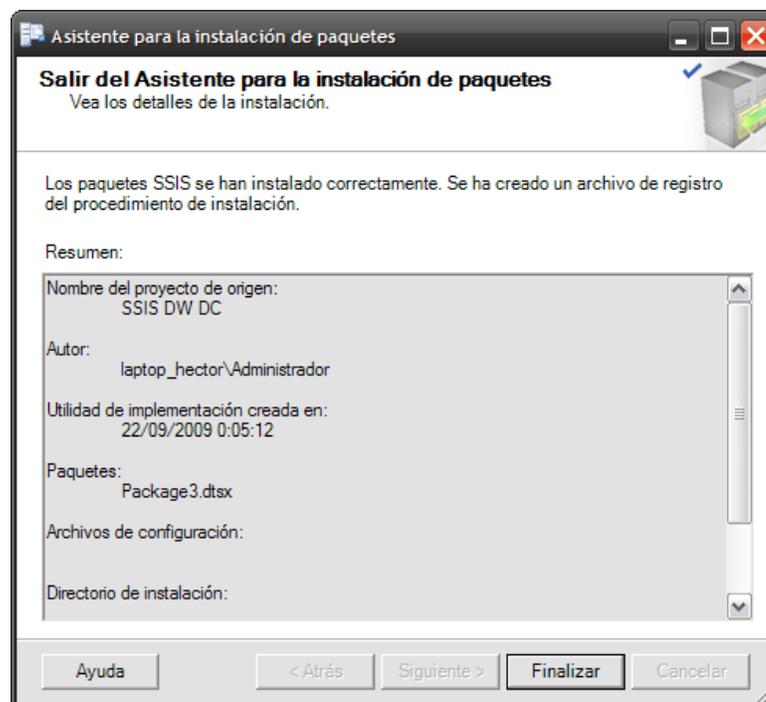


Figura 4.186: Verificación de la ejecución del paquete

- Para terminar hacer clic en **Finalizar**.

### 4.2.5.6. Tarea 3: Probar los paquetes implementados

Para ejecutar el paquete es necesario tener en una ubicación conocida la carpeta deployment.

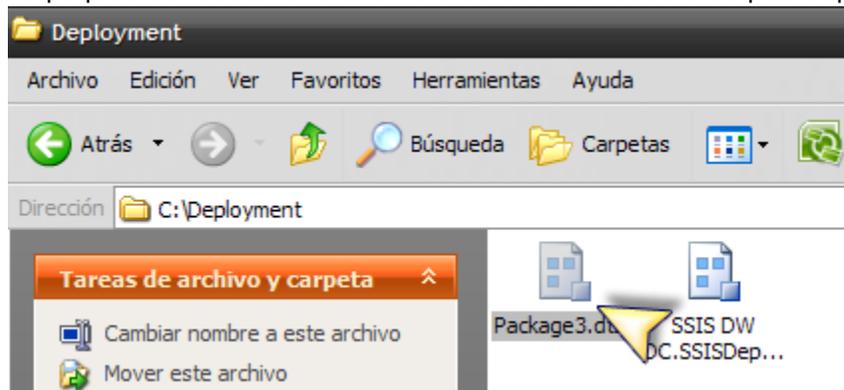


Figura 4.187: Ubicación de paquete generado

A continuación haremos doble clic sobre el paquete generado. Y llenamos la información según se muestra en la figura siguiente:

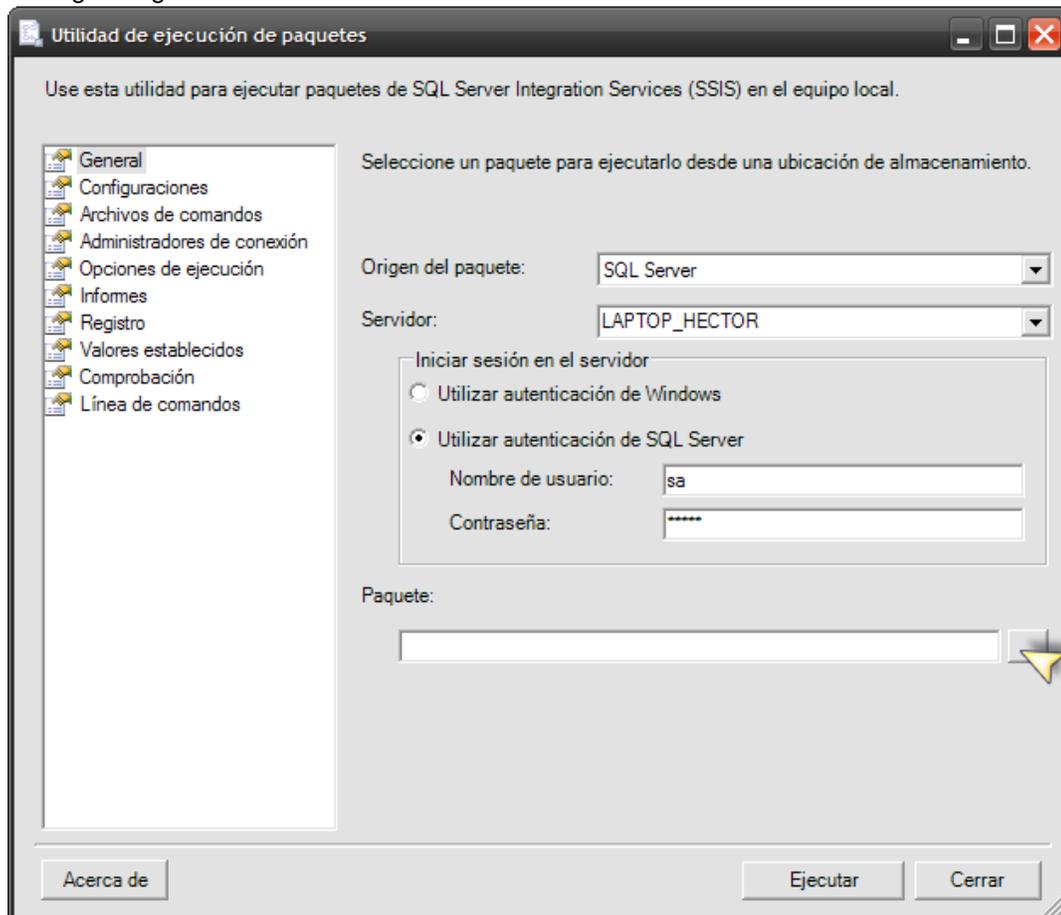


Figura 4.188: Configuración del paquete

Hacer clic en la fecha amarilla según se muestra en la figura anterior, a continuación se nos mostrara la siguiente pantalla donde podremos elegir que paquete se ejecutara, luego hacer clic en **Aceptar**.

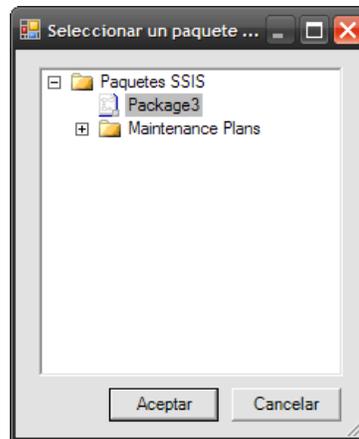


Figura 4.189: Elección del paquete a ejecutar

Para ejecutar el paquete hacer clic en **ejecutar** según se muestra en la siguiente figura.

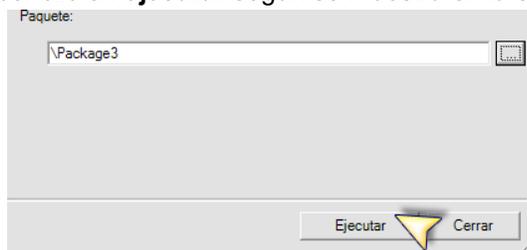


Figura 4.190: Confirmación de la ejecución del paquete

Al ejecutar el paquete nos muestra a detalle el recorrido y el tiempo que se tarda en realizar la carga de la información.

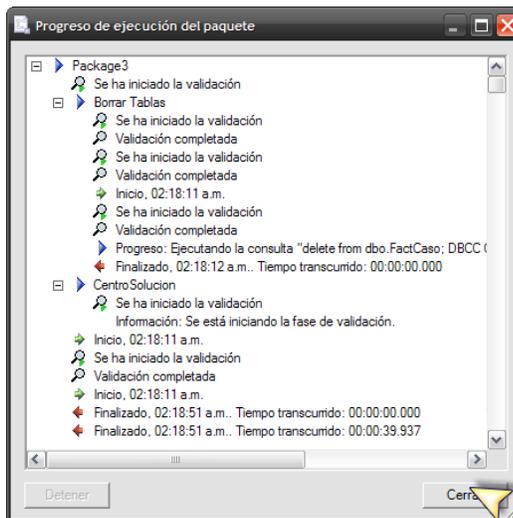


Figura 4.191: Ubicación de paquete generado

Para finalizar, hacer clic en cerrar en las 2 ventanas.

#### 4.2.5.7. Ejecución del paquete desde SQL Agent

Para poder realizar este proceso es necesario haber terminado la **Tarea 2: Ejecutar el Asistente para la instalación de paquetes**. Y seguir las instrucciones que se presentan a continuación:

Ingresar a SQL Server Management Studio, y abrir la instancia SQL server que se desea crear un trabajo, esto se podrá ver en la barra de Explorador de objetos

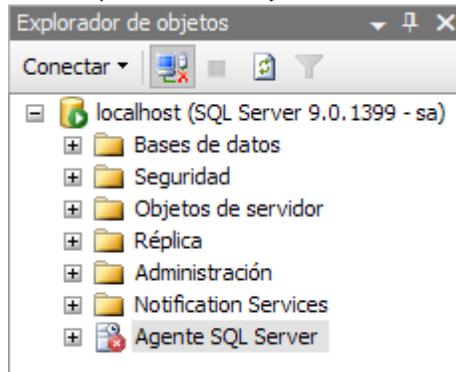


Figura 4.192: Ubicación de paquete generado

Si el agente no ejecutándose, hacer clic derecho sobre **Agente SQL Server** y luego elegir la opción iniciar, tal como se muestra en la figura siguiente:

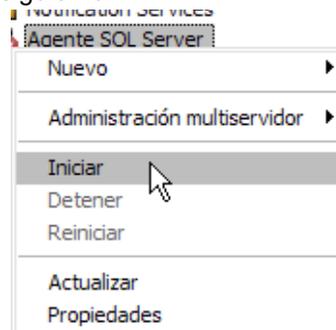


Figura 4.193: Opción para iniciar el Agente SQL

A continuación se presenta un cuadro de diálogo en el que se pide que se confirme que se desea iniciar la ejecución de este servicio, al elegir la opción Si, se muestra otro cuadro de diálogo con el proceso de inicio de este servicio.

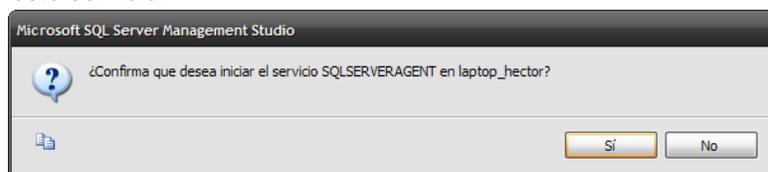
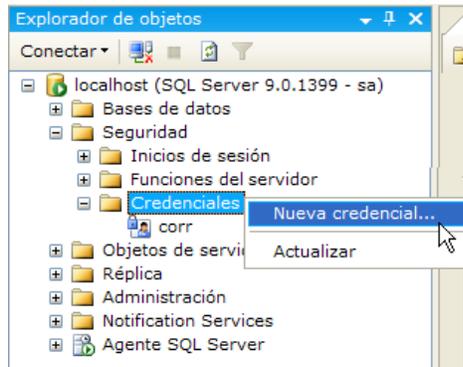


Figura 4.194: Confirmación de inicio del Agente



Figura 4.195: barra de progreso de inicio del Agente SQL

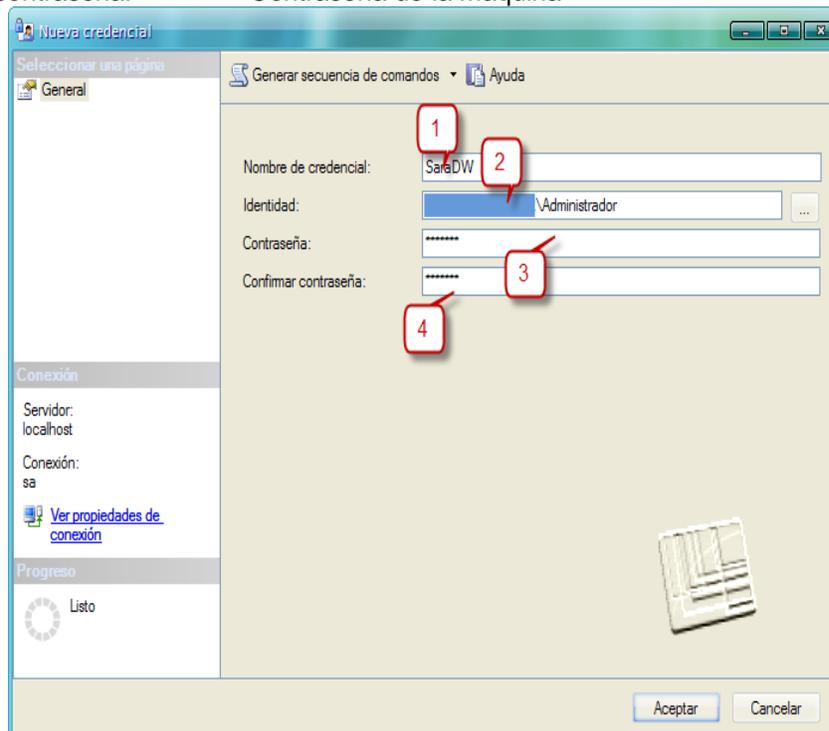
Para poder programar el paquete desde el Agente de SQL, es necesario establecer configuraciones de seguridad propias de la base. A continuación se detallan los pasos a seguir para establecer el ambiente necesario. Para ello es necesario crear una credencial de seguridad en la base de datos, tal como se muestra en la siguiente figura.



**Figura 4.196: Opciones para la creación de una credencial**

A continuación se nos muestra una pantalla como la siguiente en donde se tienen que definir información como:

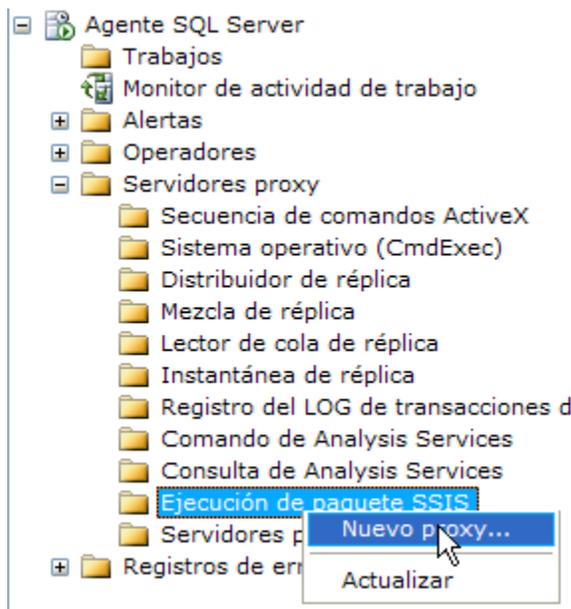
- 1- Nombre de credencial: SaraDW
- 2- Identidad: LAPTOP\_HECTOR\Administrador
- 3- Contraseña: Contraseña de la maquina
- 4- Confirmar contraseña: Contraseña de la maquina



**Figura 4.197: Datos Generales de la credencial**

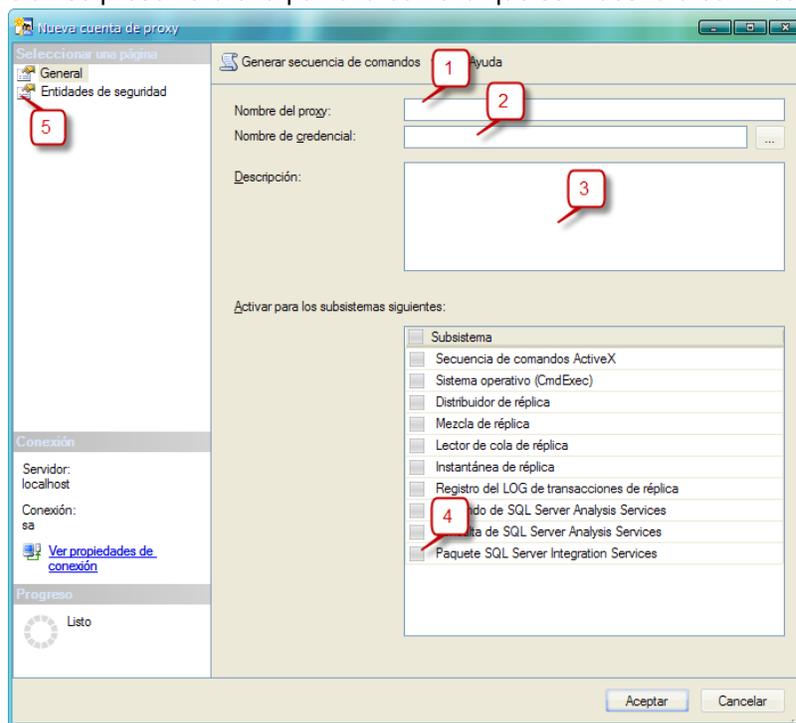
Teniendo creada la credencial es necesario definir una cuenta la cual será utilizada como proxy para la ejecución del paquete que se va a ejecutar. Esta cuenta se crea en el agente de SQL Server haciendo

clic derecho sobre servidores proxy, y eligiendo la opción **Nuevo Proxy**, la información se llena tal como se muestra a continuación:



**Figura 4.198: Creación de un nuevo proxy para ejecutar paquete**

Al elegir dicha opción se presentara una pantalla como la que se muestra a continuación:



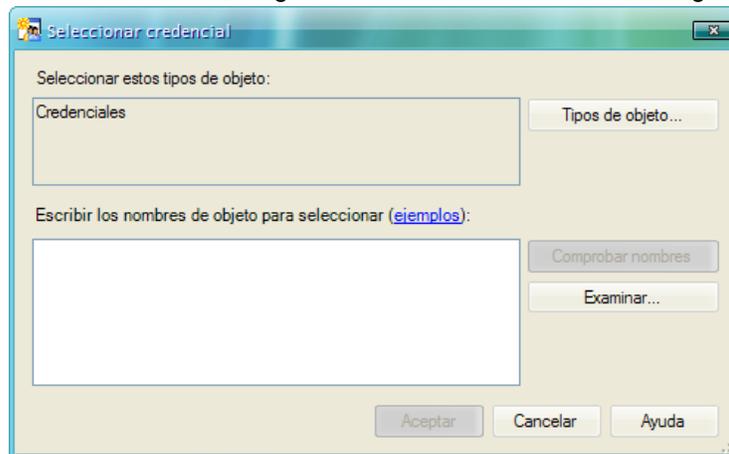
**Figura 4.199: Descripción de la pantalla de la creación del proxy**

Cada uno de los puntos se describe a continuación:

- 1- Nombre del proxy: se utiliza para definir un nombre que describa al proxy. Ejemplo: SaraProxy

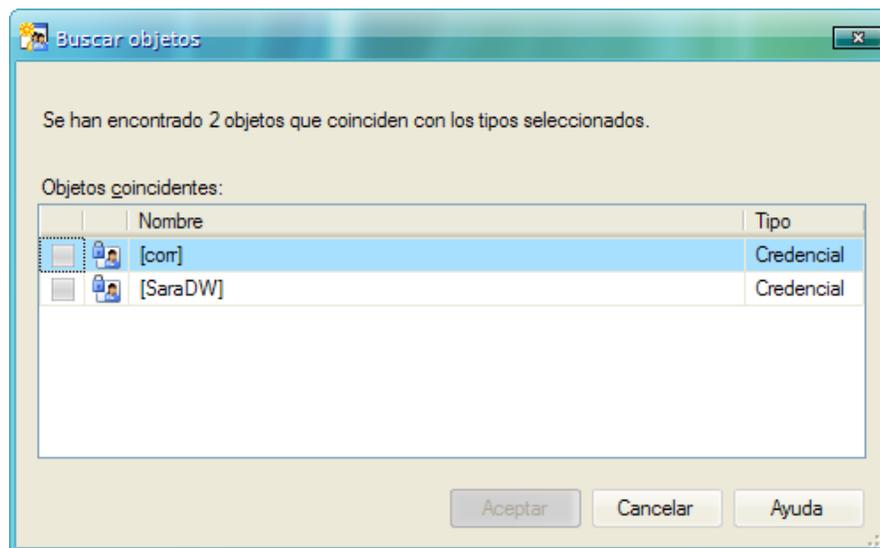
- 2- Nombre de credencial: *esta credencial debe de estar creada en la base de datos, es por ello que se explican los pasos anteriores. Debido a eso elijeremos la credencia: SaraDW*
- 3- Descripción: *un descriptor de proxy sirve para especificar información adicional acerca del paquete.*
- 4- Activar para los subsistemas siguientes: *para nuestro caso solo nos interesa que se encuentre marcada la opción de **Paquete SQL Server Integration Services.***

En la opción de nombre credencial nos permite buscar las credenciales creadas, haciendo clic en el botón derecho del nombre de la credencia, luego se mostrara una cuadro de dialogo como el siguiente:



**Figura 4.200: Cuadro de dialogo para la especificación de credenciales**

En el cuadro de dialogo anterior podemos visualizar todas las credenciales creadas haciendo clic en el botón examinar. Y presenta otro cuadro de dialogo como el siguiente.



**Figura 4.201: Descripción del todas las credenciales**

Este cuadro de dialogo nos debe de quedar llena similar a la pantalla que se muestra a continuación:

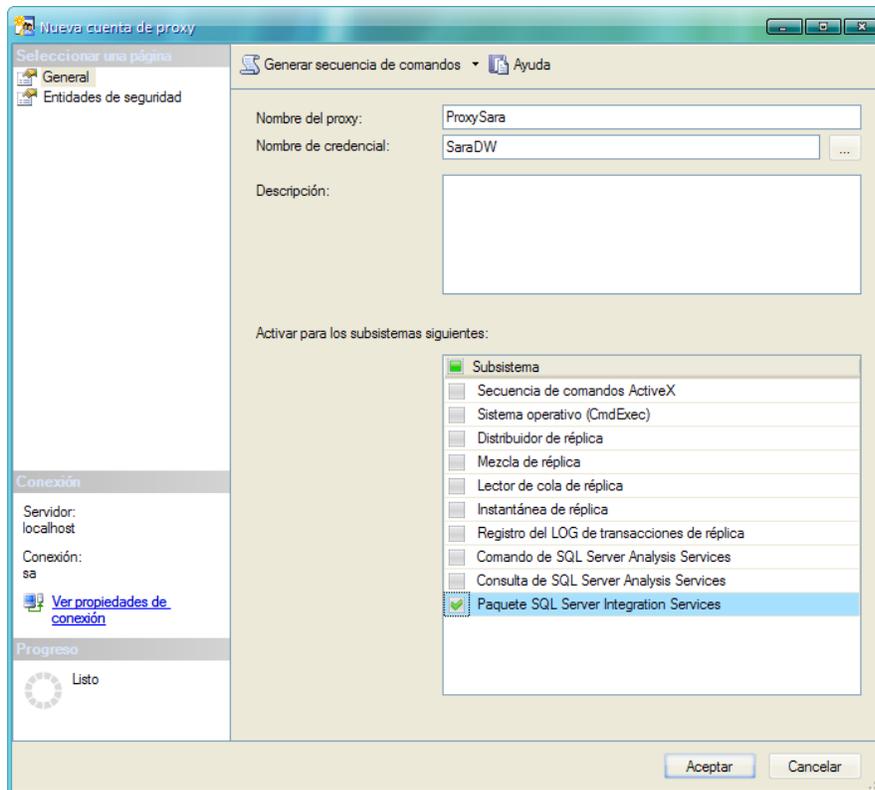


Figura 4.202: Forma correcta del llenado

Para finalizar debemos de elegir la página **Entidades de seguridad**, y se presentara una ventana como la siguiente:

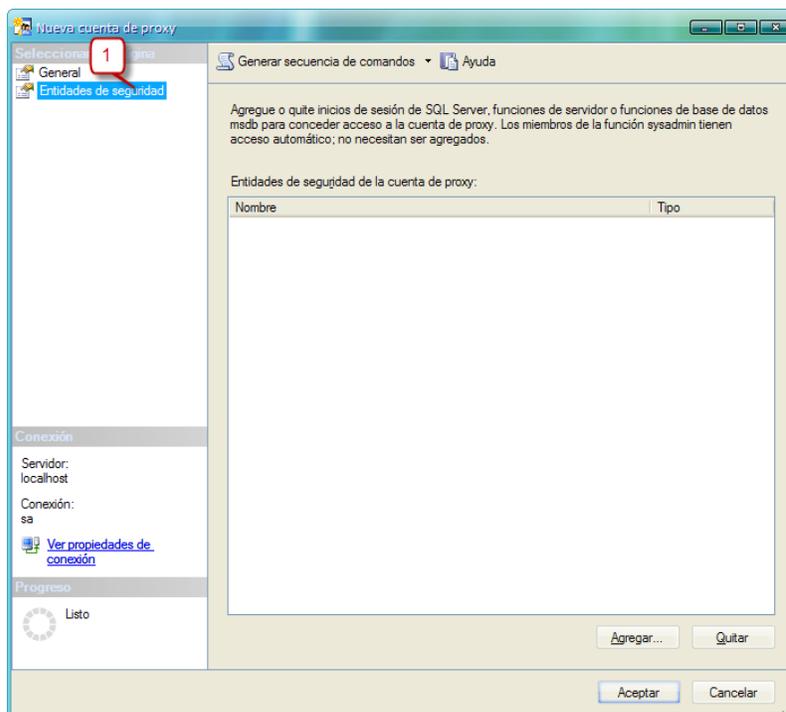
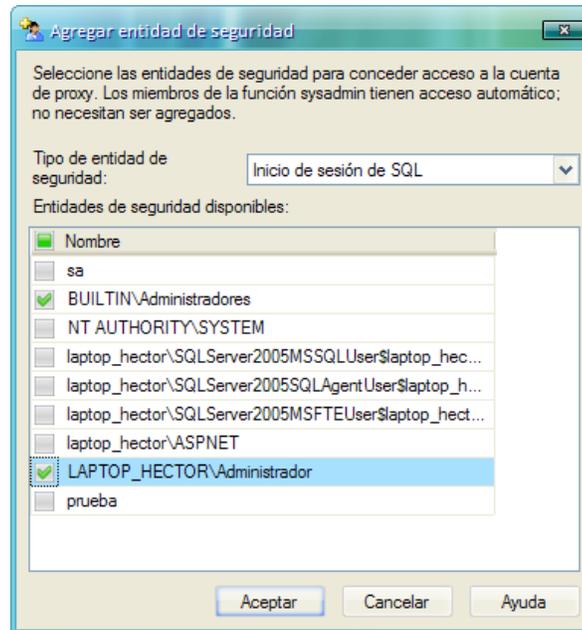


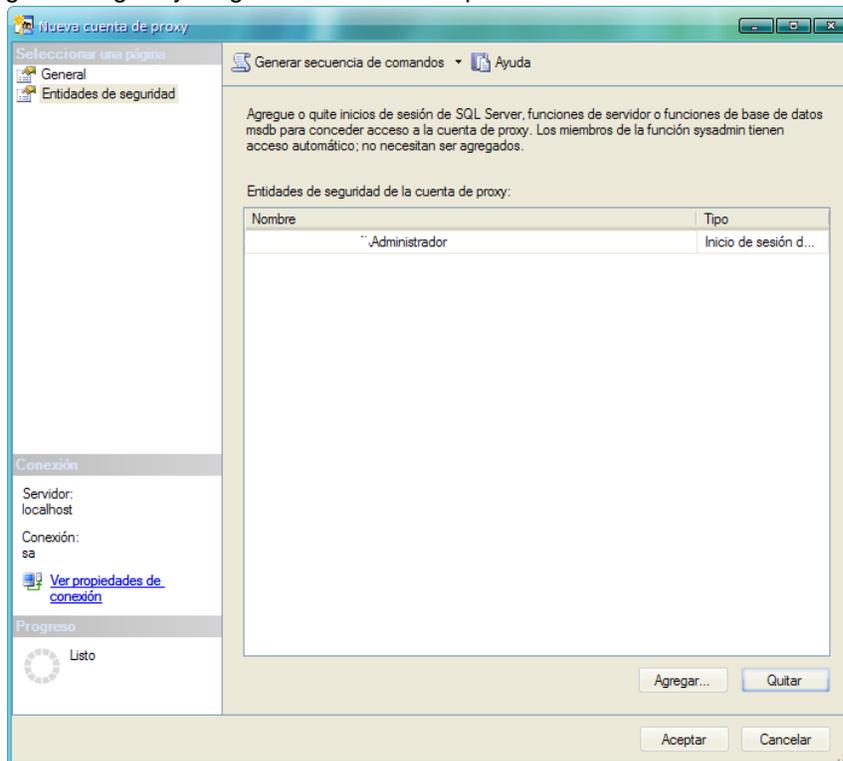
Figura 4.203: Pagina de Entidades de seguridad

Esta pantalla se debe especificar con que usuario se estará el nombre de usuario con el que se ejecutaran los paquetes, para ello es necesario hacer clic en agregar y se mostrará el siguiente cuadro de dialogo y se deben elegir los inicios de sesión tanto de SQL como los de inicios de Windows.



**Figura 4.204: Listados de los inicios de sesión**

Para finalizar se debe de verificar que se tiene la llena la pantalla con la información similar a la presentada en la siguiente figura y luego hacer clic en aceptar:



**Figura 4.205: Listados de entidades de la cuenta proxy**

Con esto tendremos una cuenta configurada para la ejecución de paquetes.

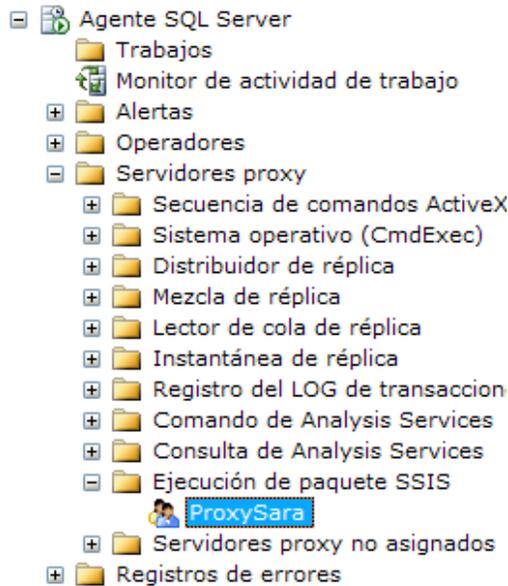


Figura 4.206: Listados de proxys creados

#### 4.2.5.8. Creación y programación de un paquete

Ahora debemos expandir el Agente SQL Server y hacer clic con el botón derecho en **Trabajos** y, a continuación, hacer clic en **Nuevo**

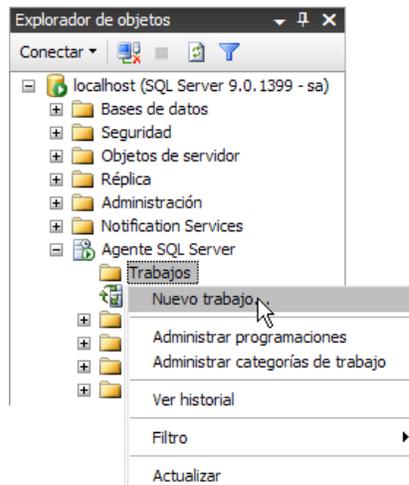
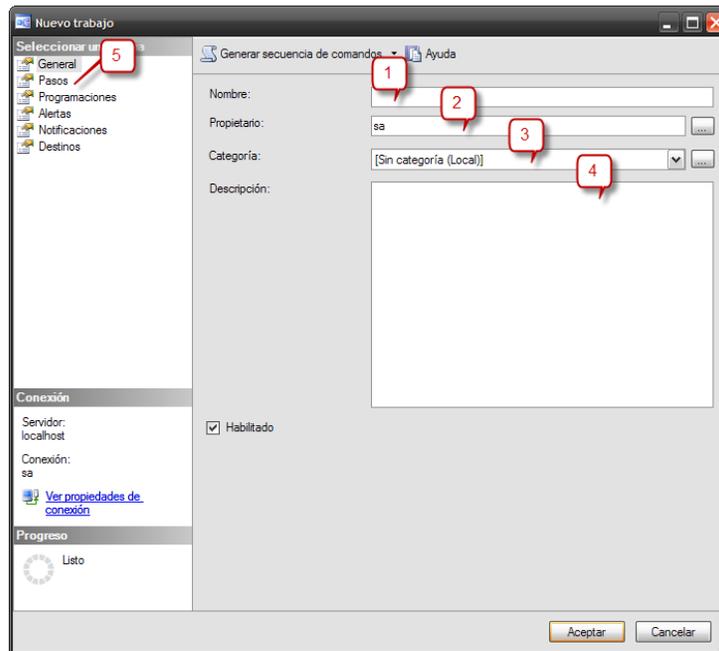


Figura 4.207: Opción para crea un nuevo trabajo

A continuación se presenta una ventana, donde se especificarán los datos generales para la el trabajo. En esta ventana se definen los siguientes datos:

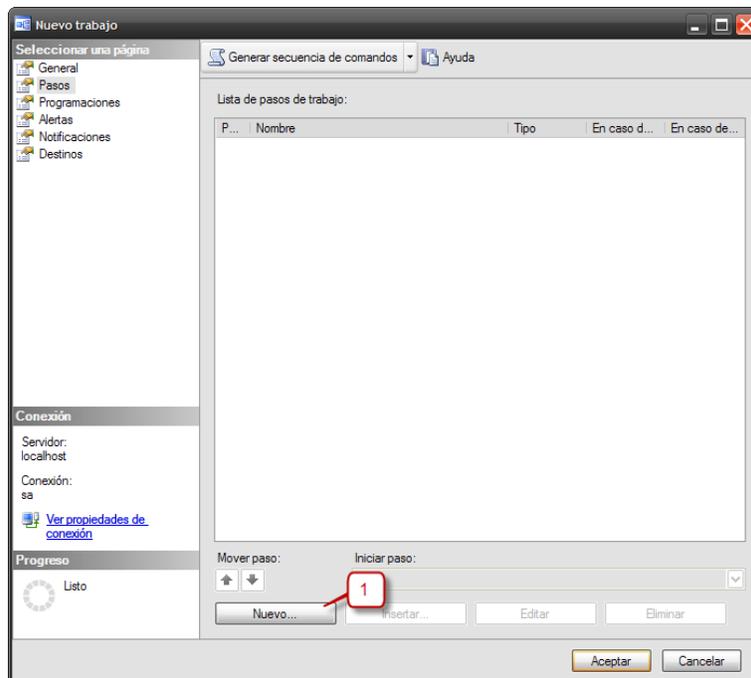
- 1- Nombre: se especifica el nombre del nuevo paquete.
- 2- Propietario: es el usuario con el que se está creando el trabajo.
- 3- Categoría: para especificar qué tipo de trabajo es, esto es opcional pero sirve para clasificar.

- 4- Descripción: esta es opcional y se utiliza para poder especificar las actividades que realiza este paquete.



**Figura 4.208 : Pestaña para los datos generales del trabajo**

Al llenar todos los campos hacer clic sobre la opción Pasos, tal señala el indicador número 5 de la figura anterior. Al hacer clic sobre esta opción se nos presentara una ventana como la que se muestra a continuación:



**Figura 4.209 : Pantalla para creación de los pasos**

En esta ventana se debe de elegir la opción nuevo, según señala el indicador numero 1 de la figura anterior, la cual mostrara otra ventana como la siguiente.

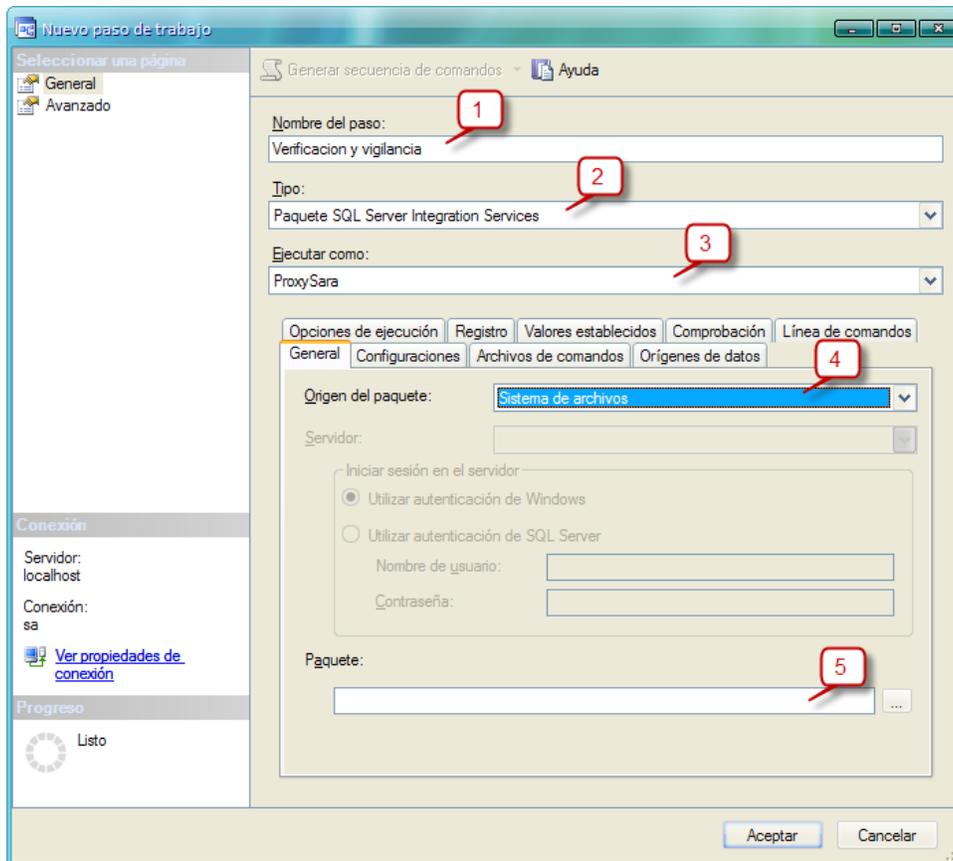


Figura 4.210 : Pantalla para creación de los pasos

La información obligatoria que se debe proporcionar en esta figura son los cinco elementos que se encuentran señalado en la figura anterior, a continuación se detalla la descripción de cada campo:

- 1- Nombre del paso: se utiliza para definir un nombre representativo para el paso.
- 2- Tipo: se pueden elegir dependiendo de la tarea que se está programando, para nuestro caso debe estar elegido: Paquete SQL Server Integration Services.
- 3- Ejecutar como: se utiliza para definir la cuenta de proxy que se ha definido para la ejecución de los paquetes de SQL Server Integracion Services (SSIS).
- 4- General (pestaña): puede ser de 3 tipos: Sistema de Archivos, SQL Server y Almacén de paquetes SSIS, para este caso debe utilizar Sistema de archivos.
- 5- Paquete: este ítem se utiliza para definir la ubicación específica del archivo generado por el paquete, este dependerá de la ubicación que se encuentre definido el proyecto de integración services. Como ejemplo si la ubicación del proyecto es c:\ETI\ la ubicación de los archivos generados por los paquetes será: c:\ETL\ETL\bin\Deployment.

Y para finalizar se debe hacer clic en aceptar.

#### 4.2.5.9. Probando la ejecución del paquete

La ejecución se puede realizar ya sea programando una hora específica para su ejecución o haciendo clic derecho sobre el trabajo recién creado y eligiendo la opción **Iniciar trabajo**. Para la verificación de la ejecución de la prueba, se realizara iniciando la el trabajo manualmente.

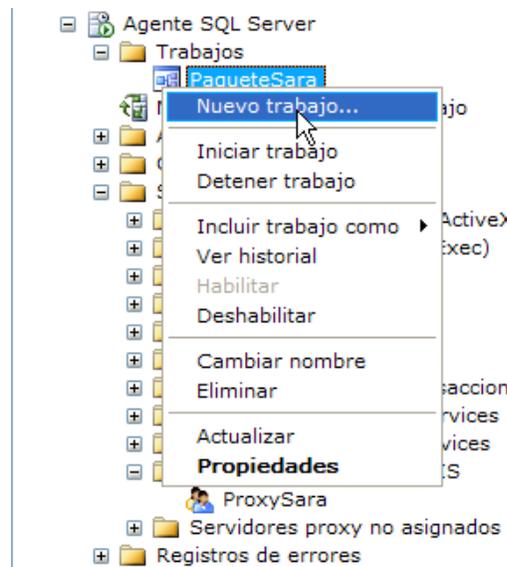


Figura 4.211: Pantalla para creación de los pasos

A continuación se mostrara un cuadro de dialogo indicando el progreso de la ejecución:

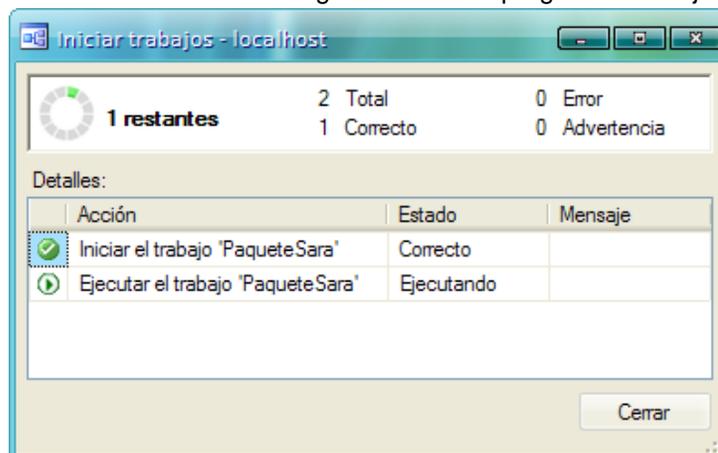


Figura 4.212: Progreso de la ejecución del trabajo

Cuando termine de ejecutar el paquete se presentará un cuadro de dialogo en el que se definen los pasos del trabajo que se han ejecutado correctamente, dicha pantalla se muestra a continuación:

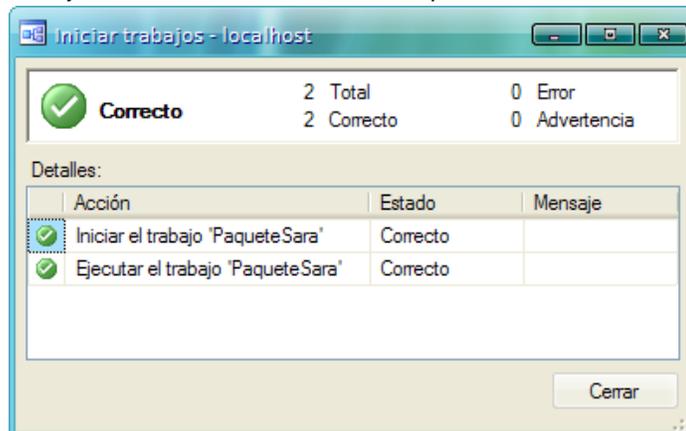


Figura 4.213 : Finalización exitosa de la ejecución del trabajo

Con esto se tiene programado la ejecución del paquete utilizado el agente de SQL server.

#### 4.2.5.10. Implementando un log en el paquete ETL

Estando ubicados en la pantalla de trabajo de SQL Server Integrations Services, se debe de elegir el menú “SSIS” tal como se muestra a continuación:

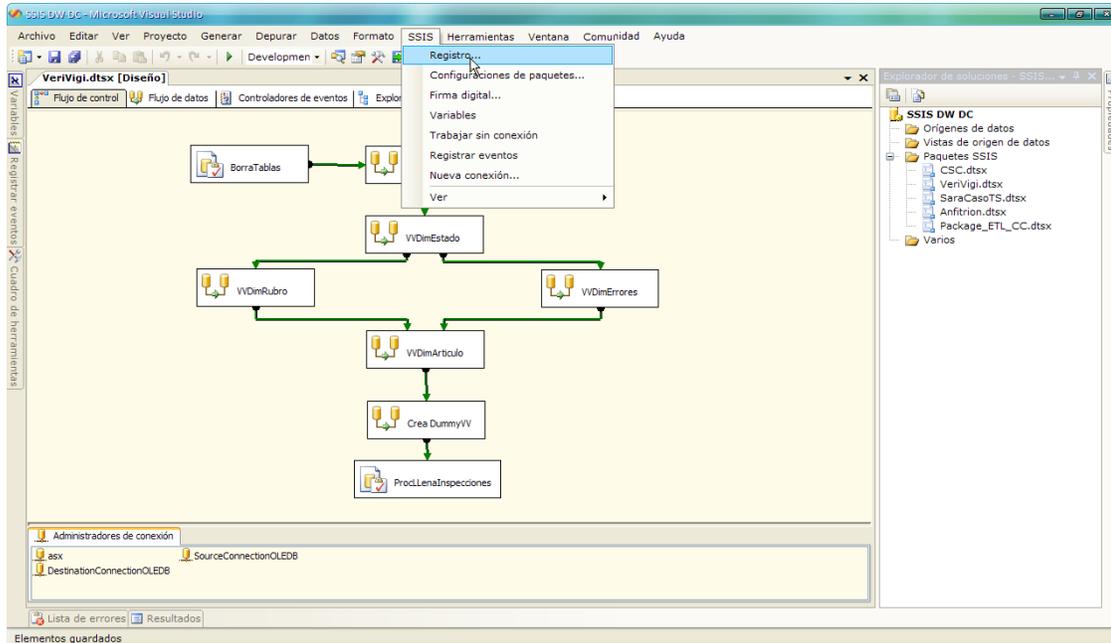


Figura 4.214: flujo de trabajo de Integrations Services

Al elegir dicha opción se presentará un cuadro de dialogo como el siguiente, en donde tendremos que elegir el paquete que se está trabajando y al que se le asociara un log

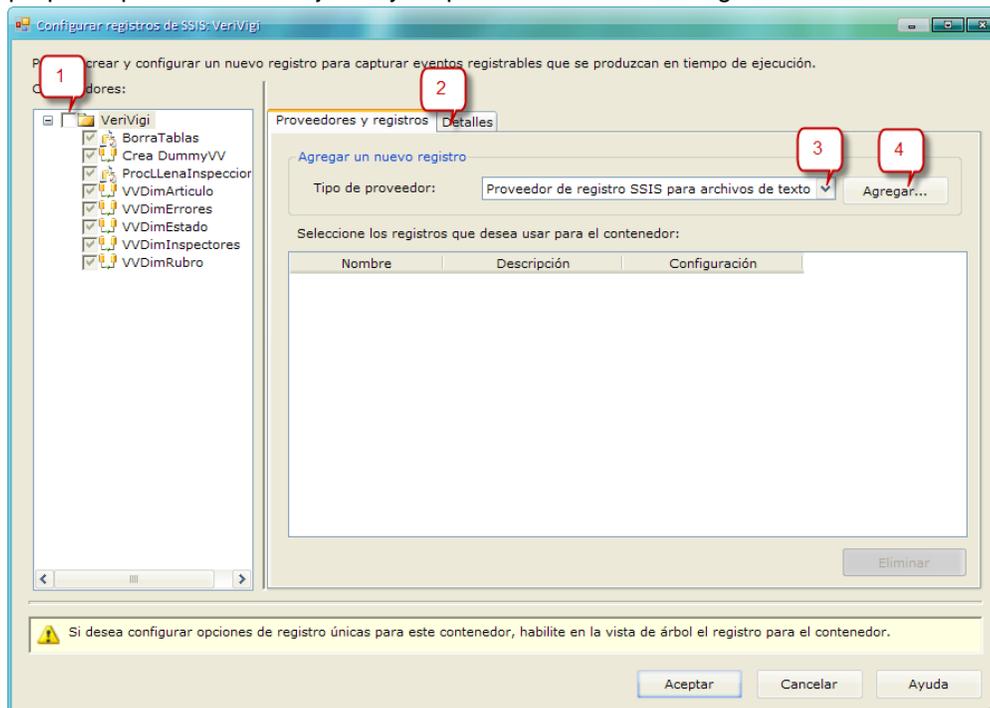
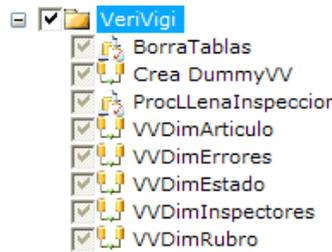


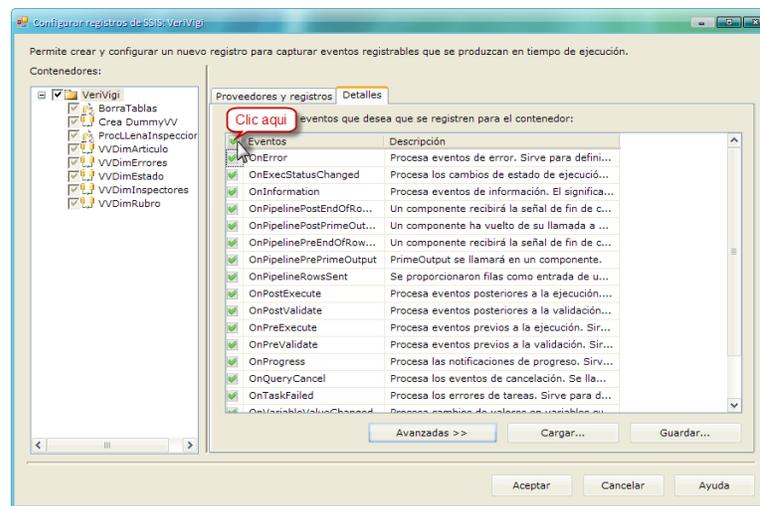
Figura 4.215: Flujo de trabajo de Integrations Services

Luego debemos elegir el paquete en la parte superior donde se muestra con el número 1. Esto se muestra de forma clara a continuación:



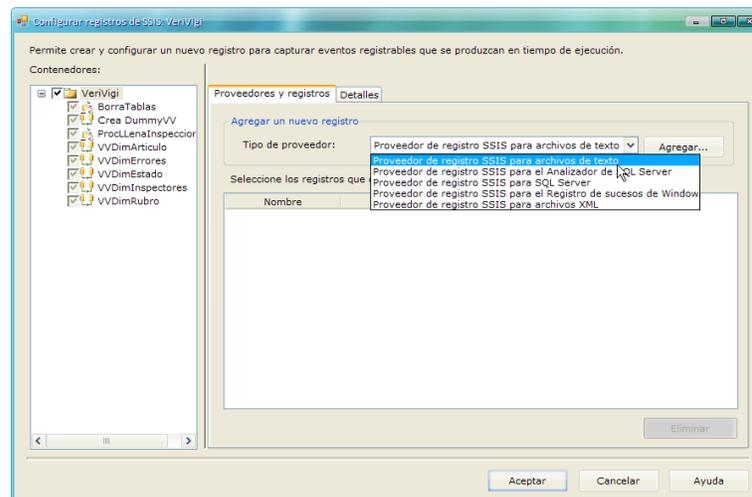
**Figura 4.216: Elegir un paquete para asociarle un log**

Al tener elegido el paquete debemos hacer clic sobre la pestaña según indica el número 2 de la Figura 4.217, y se mostrara el listado de eventos que deseamos almacenar en el log. Hacer clic donde indica el puntero para almacenar todos.



**Figura 4.217: Eventos disponibles para realizar guardar en los logs**

A continuación deberemos regresar a la pestaña “Proveedores y registros” para poder elegir el tipo de registro que almacenaremos. Para ello deberemos elegir entre el listado que se muestra en la siguiente figura:



**Figura 4.218: listado de proveedores de logs**

Teniendo elegido al proveedor que para nuestro caso es un archivo de texto deberemos definir una conexión para que se ejecute el log.

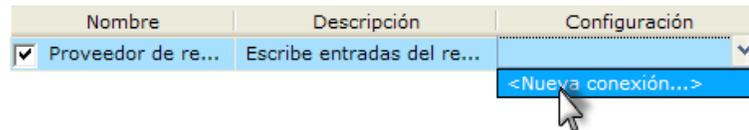


Figura 4.219: Creación de una nueva conexión para el log

Al elegir la nueva conexión se muestra un cuadro de dialogo como el siguiente, en este se podrá definir los tipos de archivo que se generaran.

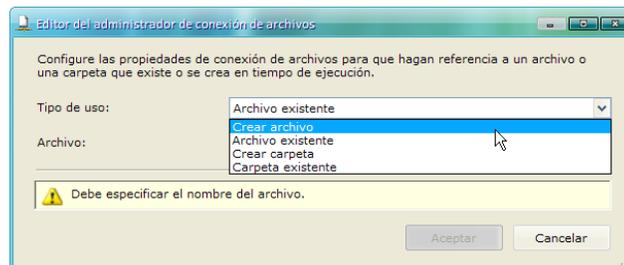


Figura 4.220: Eligiendo donde se almacenarán los registros

Al tener elegido deberemos especificar la ubicación del log, este se mostrará al hacer clic sobre la ventana examinar. La ubicación del archivo se ha definido en: "c:\log carga\SaraDWCarga.log "

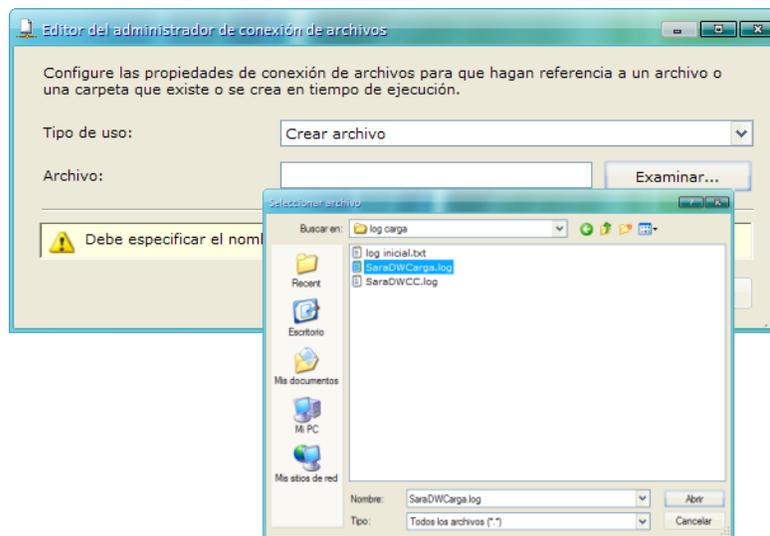


Figura 4.221: Eligiendo la ubicación del archivo log

Para finalizar hacer clic en finalizar en los cuadros de dialogo y se mostrar una un indicador de log como el que se puede apreciar en la figura siguiente.

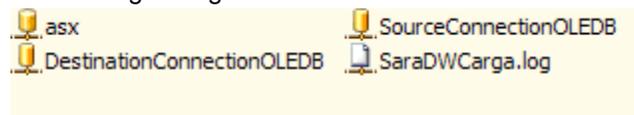


Figura 4.222: Verificando la creación de la conexión para que se ejecute el log

Con esto tenemos listo la implementación del log.

## 4.3. MANUAL DE USUARIO

### 4.3.1. Configuración Básica del Equipo de los Usuarios

#### 4.3.1.1. Reconociendo el ambiente de trabajo

Bienvenido al manual del usuario, como elementos principales es necesario tener presente los siguientes elementos:

- El uso de un navegador de internet.
- Resolución mínima de pantalla.
- Conexión a la red interna de trabajo de la Defensoría del Consumidor.
- La dirección electrónica de la aplicación.

**El navegador de internet:** Se refiere al programa de aplicación que usualmente utiliza para conectarse a internet, dentro de los más comunes tenemos: Internet Explorer 7 o superior, Mozilla Firefox 3 o superior, Google Chrome 2 superior, Opera 9 o superior, Safari 3 o superior. Es recomendable que se utilice Internet Explorer en el caso de las computadoras con sistemas Windows y Safari en el caso de computadoras con sistemas Mac Os. En el caso de instalar otro navegador en su computadora consulte en los sitios oficiales de cada navegador para considerar los requerimientos mínimos que debe poseer su computadora para ejecutarse sin ningún problema.

Recomendamos visitar para mayores referencias:

	Internet Explorer	<a href="http://www.microsoft.com/spain/windows/products/winfamily/ie/default.mspx">http://www.microsoft.com/spain/windows/products/winfamily/ie/default.mspx</a>
	Mozilla Firefox	<a href="http://www.mozilla-europe.org/es/firefox/">http://www.mozilla-europe.org/es/firefox/</a>
	Google Chrome	<a href="http://www.google.com/chrome/index.html?hl=es">http://www.google.com/chrome/index.html?hl=es</a>
	Opera	<a href="http://www.opera.com/">http://www.opera.com/</a>
	Safari	<a href="http://www.apple.com/es/safari/">http://www.apple.com/es/safari/</a>

**Resolución de pantalla:** Generalmente en las computadoras actuales poseen una resolución mayor a la de 1024x768, sin embargo se considera importante mencionar este punto con el propósito que usted pueda disfrutar de una navegación más cómoda teniendo acceso completo a la mayor cantidad de opciones que brinda dicho sistema.

#### Windows XP

1. En el menú **Inicio**, haga clic en **Panel de control**.
2. Abra el panel de control de Pantalla.
3. Haga clic en la pestaña Ajustes y localice el regulador de resolución de pantalla.
4. Seleccione una resolución mayor de 1024 x 640 admitida por la pantalla ajustando el regulador.

5. Haga clic en Aceptar.
6. Es posible que la pantalla parpadee brevemente y que aparezca un mensaje que le pida que confirme la configuración. Haga clic en Sí **si la imagen de la pantalla tiene un aspecto normal**.  
Nota: si la imagen de la pantalla se ve oscura o distorsionada, espere 15 segundos a que la pantalla vuelva a su estado anterior.

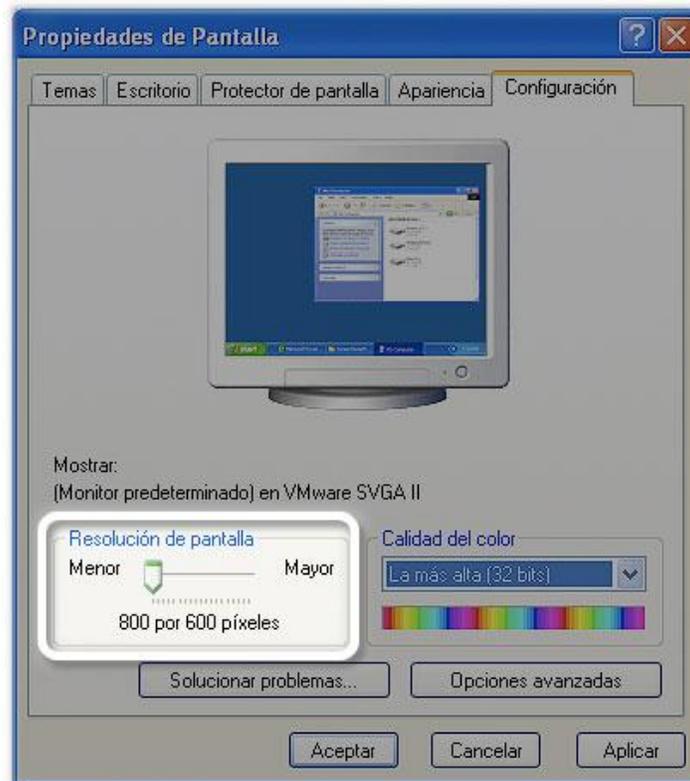


Figura 4.223: Propiedades de pantalla en Windows XP

### Windows Vista

1. En el menú **Inicio**, haga clic en **Panel de control**.
2. Haga clic en Hardware y sonido en la ventana principal del Panel de control y haga clic en Personalización.

**También es posible hacer clic en Personalización en la vista clásica del Panel de control.**

3. Haga clic en Configuración de pantalla.
4. Seleccione una resolución mayor de 1024 x 640 admitida por la pantalla ajustando el regulador.
5. Haga clic en Aceptar.
6. Es posible que la pantalla parpadee brevemente y que aparezca un mensaje que le pida que confirme la configuración. Haga clic en Sí **si la imagen de la pantalla tiene un aspecto normal**.  
Nota: si la imagen de la pantalla se ve oscura o distorsionada, espere 15 segundos a que la pantalla vuelva a su estado anterior.

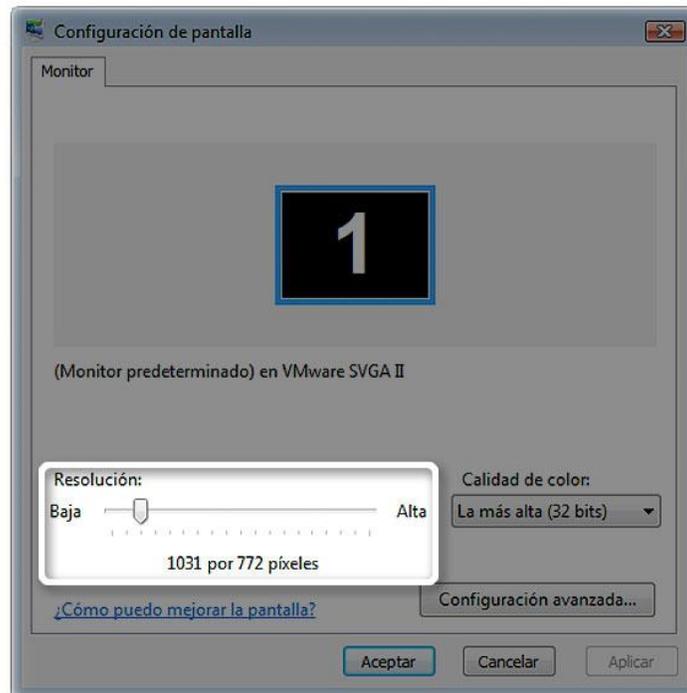


Figura 4.224: Configuración de pantalla en Windows Vista

### Mac OS X

1. Elija menú **Apple > Preferencias del Sistema**
2. Seleccione **Ver > Pantallas**.
3. Haga clic en la ficha Pantalla.
4. Seleccione una resolución mayor de 1024 x 640 admitida por la pantalla ajustando el regulador.

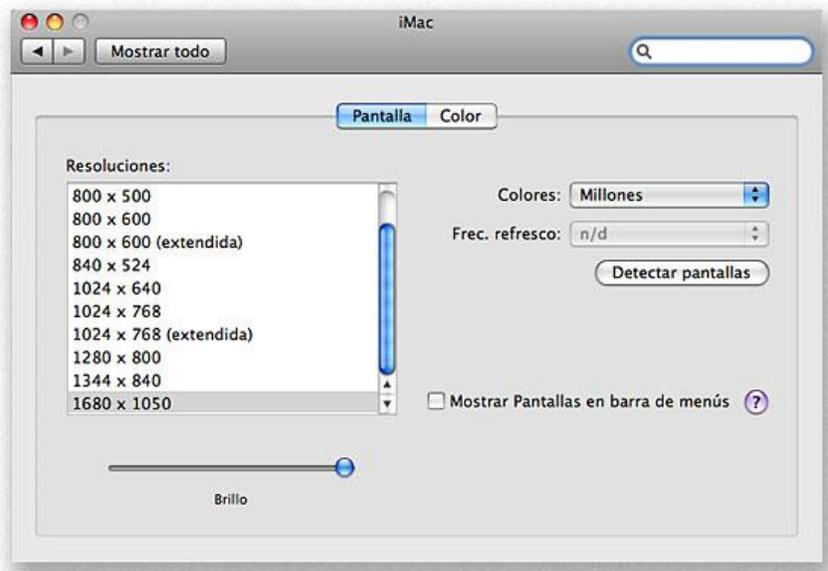


Figura 4.225: Resoluciones disponibles

Si tiene alguna duda de cómo realizar esta configuración, solicite ayuda a una persona experimentada en el área de configuración de sistemas operativos o al departamento de sistemas de su organización.

**Conectado a la red interna de la DC:** Esta aplicación solo está disponible si usted se encuentra físicamente dentro de alguna delegación de la DC.

**Dirección electrónica de la aplicación:** Para ingresar a Pentaho digite la siguiente dirección electrónica en su navegador: <http://localhost:8080>

Desde aquí podrá ingresar al sistema y realizar las tareas que se detallaran más adelante, dependiendo de los accesos que posea de acuerdo a su nombre de usuario.

## 4.3.2. Navegación del Sistema

### 4.3.2.1. Reconociendo el Ambiente de Trabajo

**Pantalla Inicial:** Es la pantalla principal del sistema de navegación en el cual se presenta la opción de iniciar sesión, el sistema se encuentra en inglés, por esa razón es muy importante identificar el ambiente de trabajo con el que se cuenta, para sacar el máximo provecho de la herramienta y realizar análisis y reportes sin ningún inconveniente.



Figura 4.226: Pantalla de inicio de sesión de Pentaho Bi Server

**Pentaho User Console Login «Consola de Usuario de Pentaho, Inicio de Sesión»:** Es utilizado para iniciar sesión al sistema, al hacer clic sobre dicho botón lograremos acceder a la ventana de inicio de sesión.

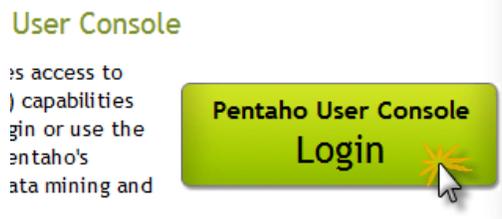


Figura 4.227: Botón para activar ventana de inicio de sesión

**Ventana de Inicio de Sesión:** «Ingresar» Se utiliza para ingresar el nombre de usuario con el que se desea ingresar al sistema, dicho usuario determina el acceso que ha sido asignado previamente para que se logre visualizar la información correcta y relacionada con el área de trabajo del usuario, con el propósito de facilitar el análisis y generación de reportes para cada departamento de la organización.

La ventana de inicio de sesión tiene el siguiente aspecto:

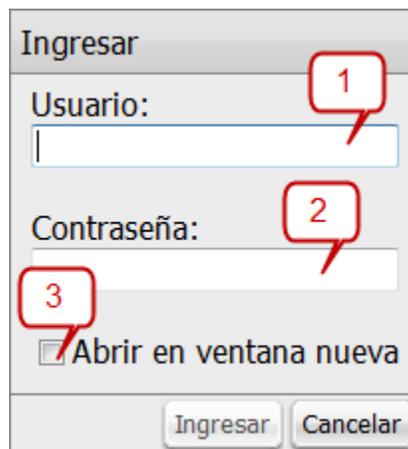


Figura 4.228: Ventana de Inicio de Sesión

1	Ingreso de Nombre Manual
2	Ingreso de Contraseña Manual
3	Abrir Navegación en Nueva Ventana

**Inicio de sesión:** debe de ingresar su nombre y contraseña para tener acceso al sistema, el nombre de usuario y contraseña será proporcionado por el administrador.

**Abrir en una nueva Ventana:** esta opción es de gran utilidad en el caso de tener múltiples sesiones de acceso o si se posee otras páginas abiertas en el navegador, para activar esta opción haga clic en el checkbox (caja de selección) para activar esta funcionalidad.

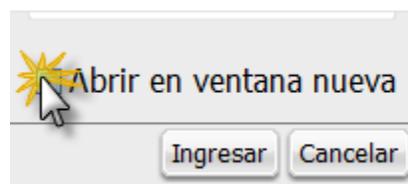


Figura 4.229: CheckBox para activar ventana nueva o emergente

Dependiendo de la configuración del navegador y la funcionalidad del mismo lograremos ingresar haciendo uso de pestañas o de ventanas emergentes

**Pestañas:** un aspecto de navegación haciendo uso de pestañas es el siguiente...

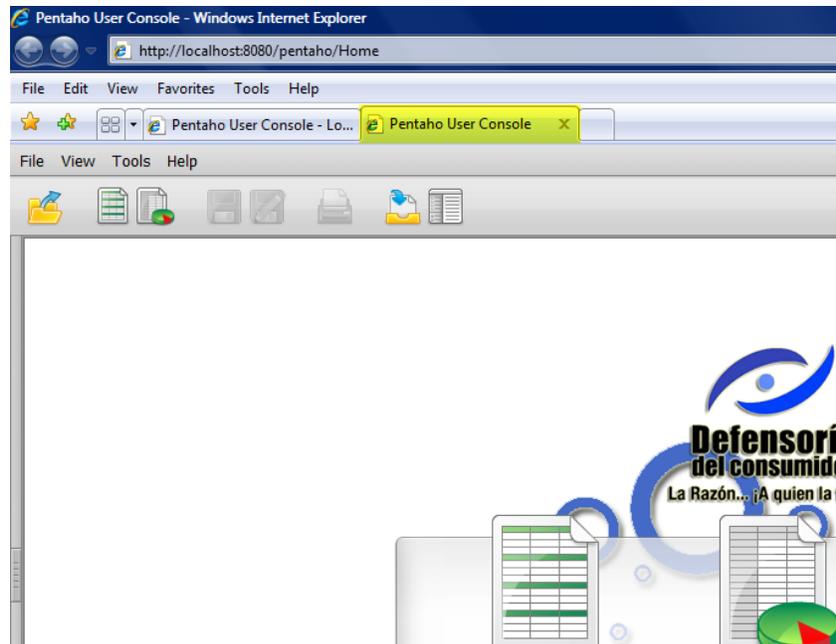


Figura 4.230: Navegación del sistema por pestañas

Como se puede apreciar se posee la pestaña de inicio de sesión junto a la pestaña de la consola propia del usuario en la que se muestran las opciones principales de trabajo.

**Uso de Ventanas Emergentes:**

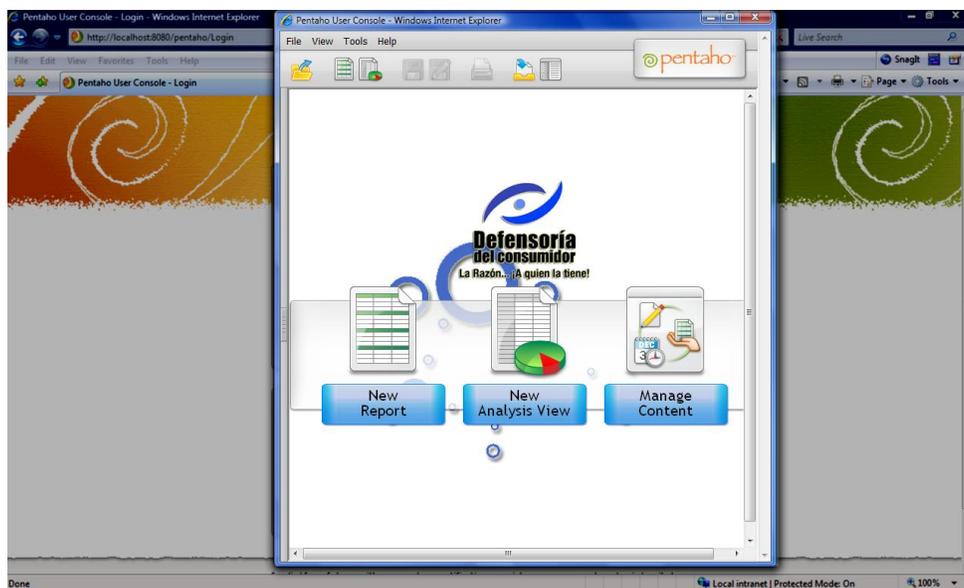


Figura 4.231: Navegación del sistema por ventanas emergentes

Se accede al sistema en una ventana distinta a la actual, en una ventana que se activa independiente a la que nos encontramos actualmente.

**OJO:** El concepto de ventanas emergentes y pestañas dependerá del navegador que se utiliza, esta opción se encuentra activa en Internet Explorer 7 o superior, consulte la página oficial de su navegador proporcionada anteriormente para verificar la configuración de pestañas y ventanas emergentes.

Una vez seleccionado el usuario, haber ingresado la contraseña y haber determinado el acceso, se procede a dar clic en el botón INGRESAR

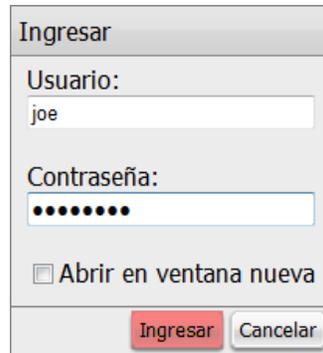


Figura 4.232: Ventana de Inicio de Sesión

Una vez se ha ingresado al sistema se muestra el espacio de trabajo para el usuario que inicio sesión.

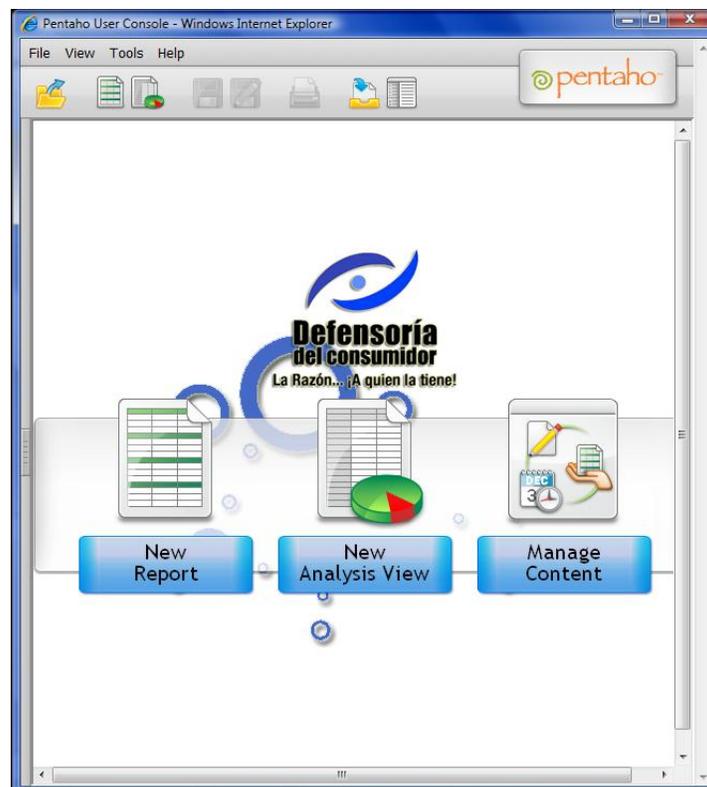


Figura 4.233: Ventana Principal del sistema Pentaho

### Entorno del Área de Trabajo

**Barra de Menú:** Muestra la navegación principal de las acciones con las que se puede trabajar con el sistema



Figura 4.234: Barra de Menú y Barra de Funciones

1	Barra de Menús
2	Barra de Funciones

### BARRA DE MENÚS



Figura 4.235: Barra de Menús

Dentro de las Opciones de la Barra de Menú se poseen File «*Archivo*», View «*Ver*» y Help «*Ayuda*».

**Menú File «*Menu Archivo*»:** Presenta las opciones generales:

Nuevo -> Reporte, Vista de Análisis

Abrir

Administrar -> Editar, Compartir, Programar Tarea

Cerrar Sesión

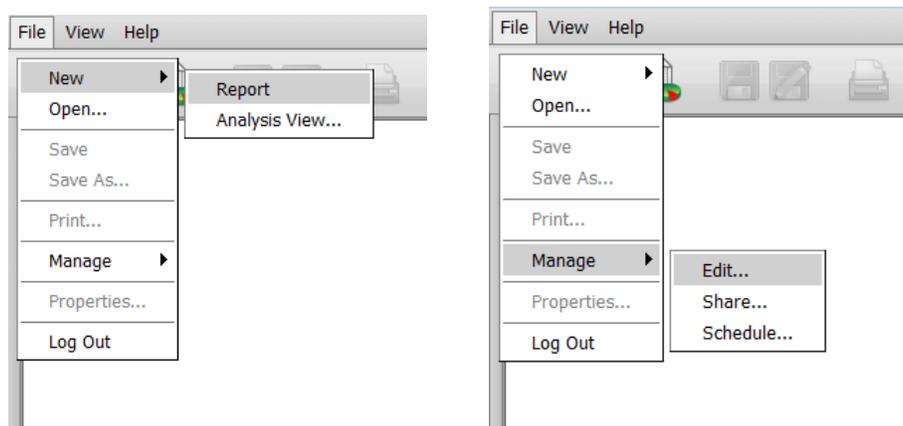


Figura 4.236: Contenido del Menú File «*Menú Archivo*»

Algunos elementos del Menú Archivo se encuentran en el panel principal o en la barra de funciones, estos elementos se identificarán en su debido momento. A continuación se detallará los elementos que se poseen en el **Menú File «*Menú Archivo*»**.

Primeramente identificamos al iniciar sesión que se encuentran elementos bloqueados (sombreados con color gris) que son inaccesibles, estos elementos se habilitan en el caso que se realice un reporte analítico, reporte estático o se ingrese al menú de manejo de contenido.

**New «Nuevo»:** Es utilizado para generar 2 tipos de reportes:

**Analysis View «Vista de Análisis»**, que es utilizado para visualizar en un nivel gerencial datos relacionados a un área específica, esta función generalmente se utiliza para hacer un análisis más profundo de la información que se posee y poder tomar decisiones basándose en datos reales

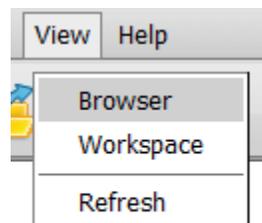
**Report «Reporte»:** Se utiliza para generar reportes estáticos, al igual que los reportes analíticos dichos reportes pueden ser utilizados para compartirlos con otros usuarios, así como para realizar análisis periódicos sin necesidad de crearlos cada vez que se desea visualizar determinada información, los reportes estáticos se basan en vistas de negocio con los cuales se pueden realizar agrupaciones que faciliten la comprensión de las tendencias o eventos que son parte del día a día de la organización.

**Open «Abrir»:** Se utiliza para abrir reportes, vistas de análisis, dashboards que se hayan almacenado con anterioridad, ciertos usuarios brindaran acceso a otros usuarios para acceder a sus reportes con el propósito de compartir información importante para un área del negocio específica, con la opción Open «Abrir» se lograra acceder al contenido compartido por otros usuarios, así como a elementos que se hayan almacenado o programado previamente.

**Manage «Administrar»:** Función que permite a los usuarios editar, compartir y programar elementos relacionados o creados por su cuenta de sesión del sistema. Ciertos usuarios pueden almacenar sus propios reportes, la Función **Edit «Editar»** es utilizada para modificar reportes creados con anterioridad, en el caso que se desee alterar una vista de análisis o que se desee agregar algún elemento dentro de un reporte estático.

**Logout «Cerrar Sesión»:** Función utilizada para salir del sistema.

**Menú View «Ver»:** Presenta como opciones Generales:



**Figura 4.237: Contenido del Menú View «Menú ver»**

**Browser «Explorador»:** Facilita la navegación del sistema en la búsqueda de reportes almacenados, compartidos o publicados, al activarse la opción del explorador aparece un panel al lado izquierdo del espacio de trabajo:

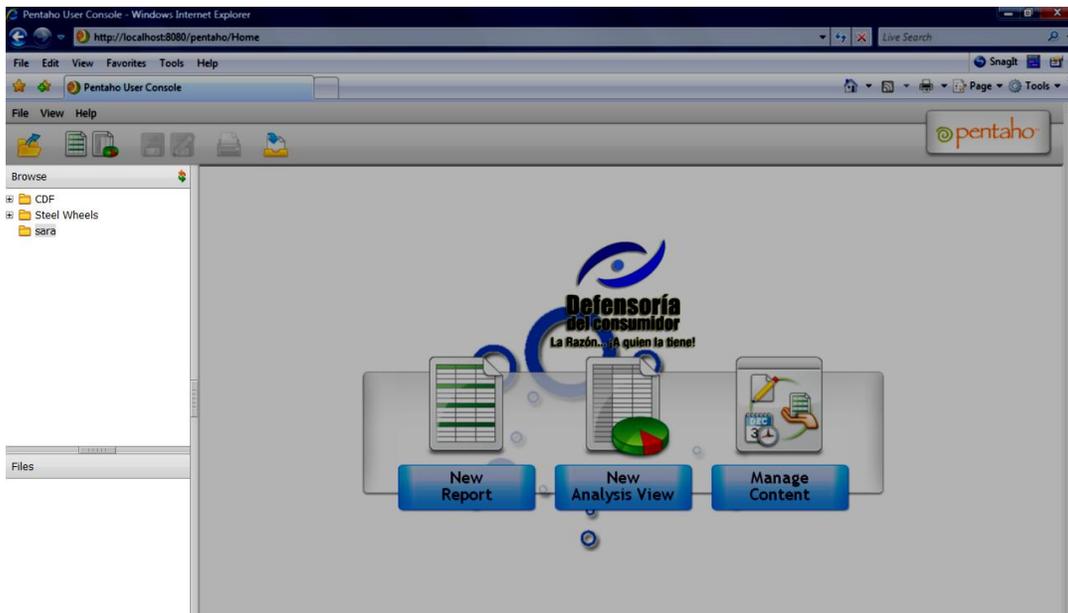


Figura 4.238: Espacio de trabajo para el Browser «Navegador»

En la parte superior del navegador se muestra el árbol de directorio que se encuentra disponible para el almacenamiento de los reportes, en la parte inferior se encuentran los archivos que contiene cada carpeta, en un principio todos los usuarios poseen acceso a los reportes, mas el contenido de los reportes depende del nivel de acceso que posea determinado usuario.

**Workspace «Espacio de Trabajo»:** Es un espacio en el que se muestra los reportes que se han enviado al servidor para que se ejecuten en determinado momento o fecha (existen reportes que pueden ser programados para que se ejecuten en determinada hora o fecha específica) aquí se pueden cancelar aquellos que no se han ejecutado aun, y además se puede visualizar o eliminar aquellos que se han ejecutado satisfactoriamente.

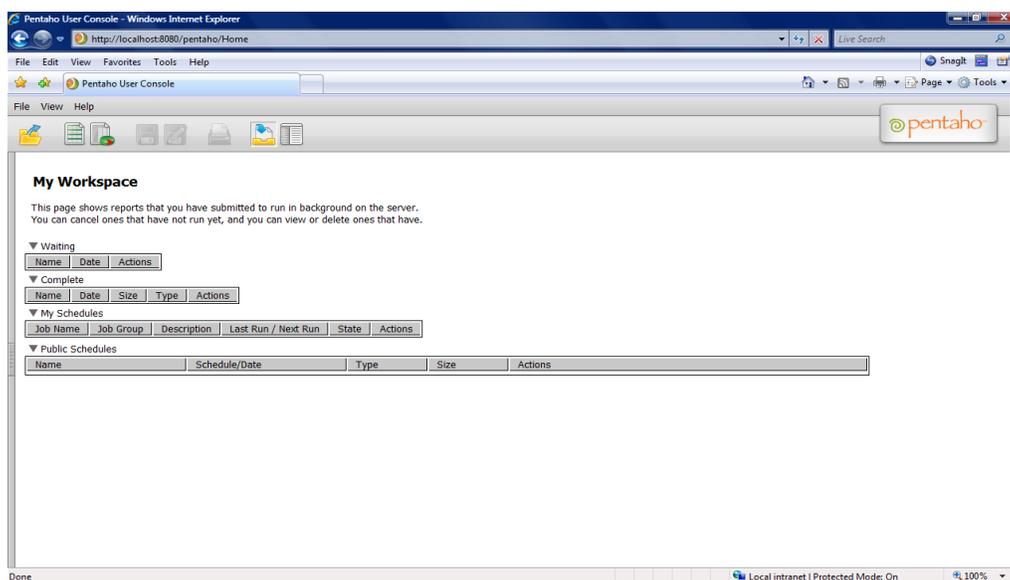


Figura 4.239: WorkSpace «Espacio de Trabajo»

**Refresh** «Actualizar»: Se utiliza para actualizar o refrescar el contenido de la página activa.

**Menú Help** «Ayuda»: Presenta como opciones Generales:

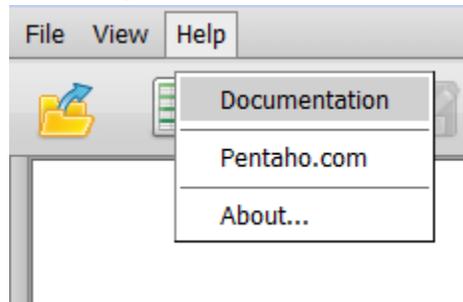


Figura 4.240: Documentación de ayuda

**Documentation** «Documentación»: Muestra un documento en formato PDF en el cual se detallan generalidades y funciones que posee la herramienta Pentaho BI.

**Pentaho.com** «Pentaho»: Hipervínculo (link) hacia el sitio oficial de Pentaho.

**About...** «Acerca de...»: Muestra la versión que se está ejecutando de la herramienta Pentaho.

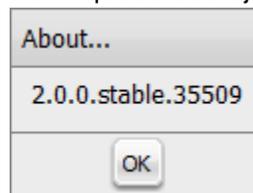


Figura 4.241: Versión de Pentaho

## BARRA DE FUNCIONES



Figura 4.242: Barra de funciones

La barra de funciones posee acceso a las funciones similares a las funciones que encontramos en la **Barra de Menús** pero con la característica que estos botones de acceso son los que más se utilizan y por esa razón se encuentran más accesibles.

Al igual que la barra de Menús en la Barra de Funciones encontramos elementos que no se encuentran disponibles solo cuando se accede a realizar una nueva vista analítica o nuevo reporte.

Las funciones disponibles que tenemos son:

Open... «Abrir»

Report... «Reporte»

Analysis View... «Vista de Análisis»

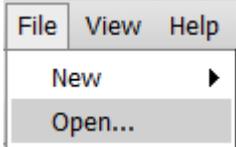
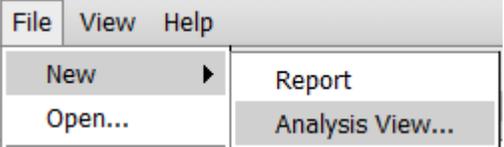
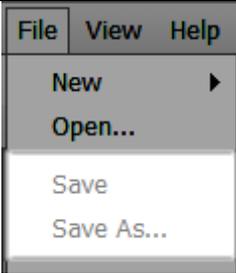
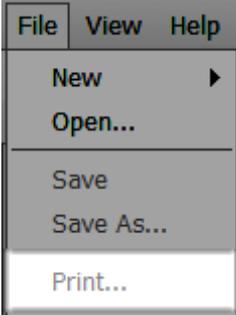
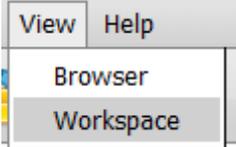
Save... & Save As... «*Guardar y Guardar como...*» (No disponibles en la pantalla inicial del espacio de trabajo)

Print «*Imprimir*»

WorkSpace «*Espacio de Trabajo*»

Toogle Browser «*Activar Explorador*»

Los equivalentes entre menú y botones son:

Ubicación en Barra de Menús	Elemento en Barra de Función	Descripción
		<p><b>Open...</b> «<i>Abrir</i>»</p>
		<p><b>Report...</b> «<i>Reporte</i>»</p>
		<p><b>Analysis View...</b> «<i>Vista de Análisis</i>»</p>
		<p><b>Save... &amp; Save As...</b> «<i>Guardar y Guardar como...</i>»</p>
		<p><b>Print</b> «<i>Imprimir</i>»</p>
		<p><b>WorkSpace</b> «<i>Espacio de Trabajo</i>»</p>



### 4.3.3. Creación de Reportes

#### 4.3.3.1. Report «Reporte»

La generación de reportes es una de las principales funciones de Pentaho para iniciar el asistente de nuevo reporte tenemos 3 opciones las cuales son:

**En el Menú File «Archivo»**



Figura 4.243: Nuevo reporte desde el menú archivo

**Desde la barra de Función:**



Figura 4.244: Icono de nuevo reporte

**En la ventana principal de navegación:**

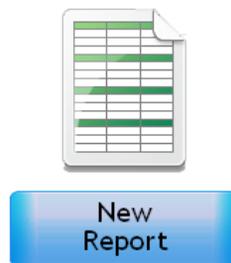


Figura 4.245: Icono nuevo reporte, desde ventana de navegación

La generación de reportes es una herramienta con la cual se debe de tener claro que es lo que se desea visualizar, tenemos a disposición un motor generador de reportes partiendo de las vistas de negocios que se han creado en el proceso de desarrollo del proyecto, estas vistas de negocio dependen de la necesidad propia de los usuarios finales y algunas de estas especializaciones de negocio son detalles enfocados en un área específica del negocio. Las vistas del negocio disponibles son el resultado del trabajo entre el área de IT y los usuarios finales, los cuales han solicitado dicha información, cabe mencionar que las vistas de negocio pueden ser modificadas solo por los administradores de la

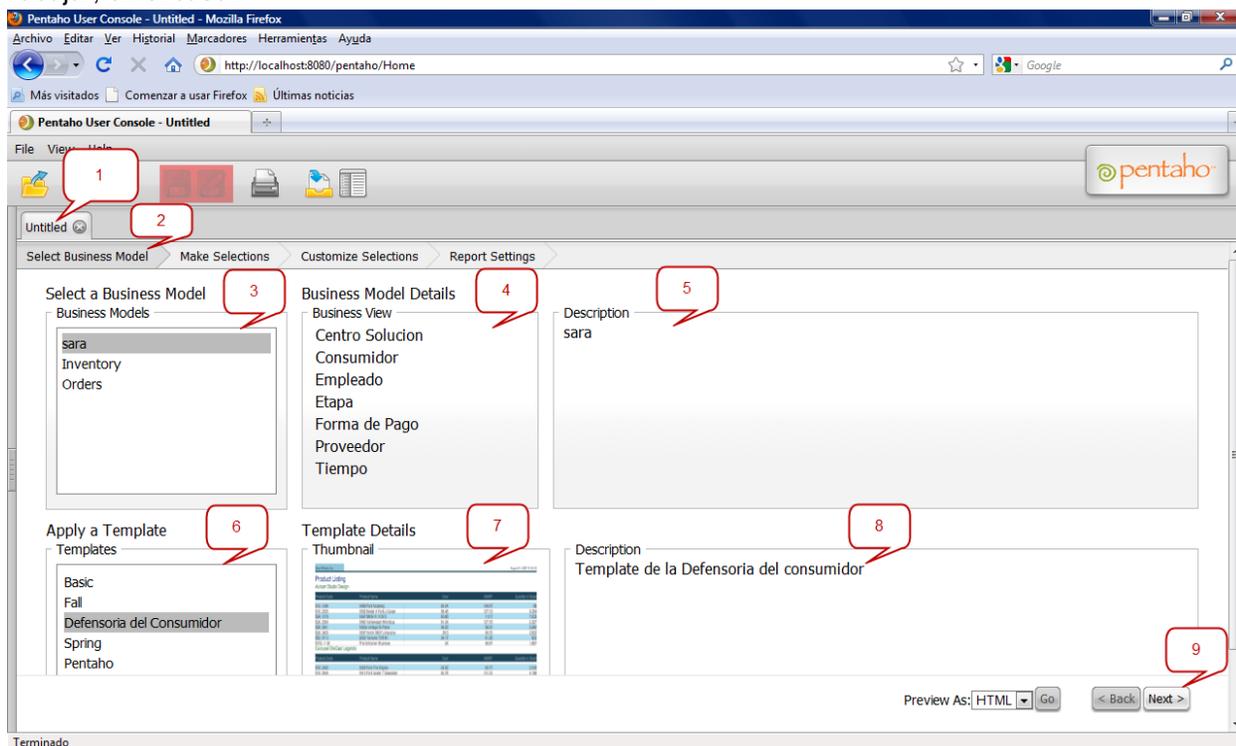
herramienta mas no por los usuarios finales. Esta es una aclaración importante dado que puede surgir la inquietud de crear una vista de negocio nueva. De ahí se denota la importancia de trabajar de la mano con los desarrolladores de la herramienta y los usuarios finales para identificar los elementos desde el inicio del proyecto.

**OJO: No quiere decir que se no se pueden cambiar las vistas de negocio, es muy importante tratar de mejorar la herramienta, como usuario final, usted tiene una gran importancia en el correcto funcionamiento del sistema, sin embargo los cambios en los esquemas presentados o cambios en las vistas de negocio deben ser solicitadas al área de sistemas, los cuales deberán analizar, identificar elementos, y realizar los cambios.**

Al seleccionar cualquier acceso se iniciara un proceso de creación de reporte mediante un asistente sencillo y práctico para crear nuestro reporte a la medida.

**Select Business Model** «*Seleccionar Modelo de Negocio*»

Es el primer paso para la creación de reportes, se debe seleccionar el modelo con el que deseamos trabajar, en el caso.



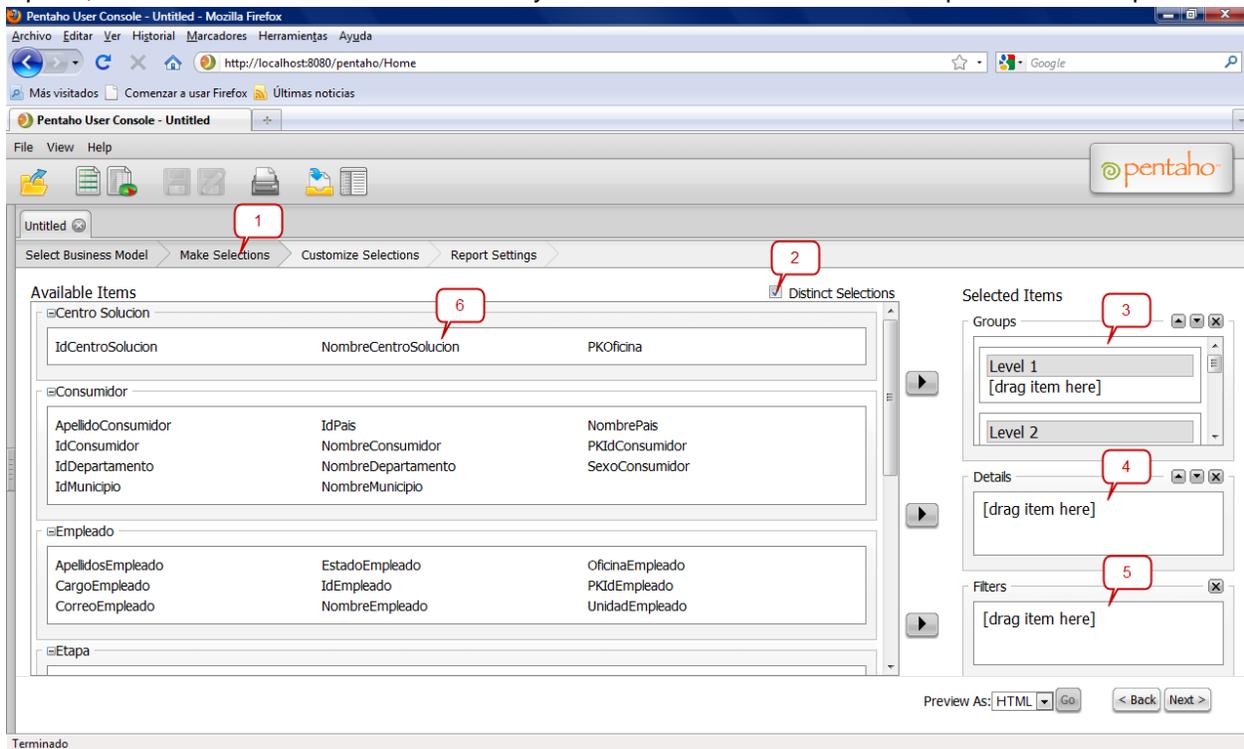
**Figura 4.246: Descripción de la vista de negocios**

1	Nos notifica que el documento no se ha guardado por defecto se coloca el nombre <b>UNTITLED</b> «Sin Titulo»
2	<b>Barra de Secuencia.</b> Muestra el proceso de generación del reporte.
3	<b>Select a Business Model</b> « <i>Selección del Modelo de Negocio</i> » se debe seleccionar de que área de negocio se desea crear el reporte
4	<b>Business Model Details</b> « <i>Detalle del Modelo de Negocio</i> » detalla los elementos que contiene el modelo de negocio seleccionado.

5	<b>Description</b> « <i>Descripción</i> » Una breve descripción del modelo seleccionado, esta información es colocada cuando se crea el modelo en la herramienta de desarrollo Meta-data-Editor
6	<b>Templates</b> « <i>Plantillas</i> » Selección de la plantilla con la que deseamos crear el reporte, este paso es para decidir cómo se visualizara el reporte, no afecta en el contenido de la información.
7	<b>Template Details</b> (Thumbnail) « <i>Detalle de Plantillas (Vista en miniatura)</i> » permite tener una vista previa de la plantilla que se ha seleccionado
8	<b>Description</b> « <i>Descripción</i> » una descripción referente a la plantilla seleccionada
9	<b>Next</b> « <i>Siguiente</i> » Una vez configurado se debe dar clic en siguiente para pasar a la siguiente fase de creación del reporte.
10	El área sombreada indica que aun <b>no se puede salvar</b> ya que hay que finalizar el proceso de creación de reportes.

**Make Selections** «*Realizar Selecciones*»

El segundo paso de la creación de reportes es determinar los grupos en los que se desea enfocar el reporte, seleccionar los detalles de los datos y finalmente seleccionar los filtros que se desean aplicar.



**Figura 4.247: Selección de información para la creación del reporte**

1	Segundo paso de creación de reportes.
2	Seleccionar datos sin duplicados.
3	Área de Agrupamiento, se agrupara el reporte por elementos arrastrados a esta área.
4	Nivel de detalle, se especifica en función de que se desean visualizar los datos agrupados.
5	Se especifica el filtro que se desea aplicar, puede ser año, cantidades recuperadas los operadores son mayor que, menos que igual

Es muy importante tener claro que se desea visualizar ya que depende de eso el resultado del reporte final. La estructura con la que se defina el reporte será clave para lograr obtener los datos

correctos. El nivel de agrupamiento como su nombre permite visualizar cierto tipo de información, el nivel detalle define que deseamos visualizar del grupo definido y los filtros delimitan la información a la que se despliega.

Para determinar los grupos en los que se desea trabajar se debe seleccionar del panel de **Available Items** «*Artículos Disponibles*» aquellos elementos por los que se desea agrupar el reporte. Para hacer eso simplemente selecciones el ítem con el mouse y arrastre al nivel deseado del panel **Groups** «*Grupos*» se pueden definir 5 niveles de agrupación con los que se limitarán los datos

**Customize Selections** «*Personalizar Selección*» Como siguiente paso se posee la personalización de las selecciones, en este momento se determinan las características visuales que deben poseer los datos, si son numéricos, mostrar cantidades en miles, cantidades de puntos decimales restricciones o filtros, alineación de los elementos

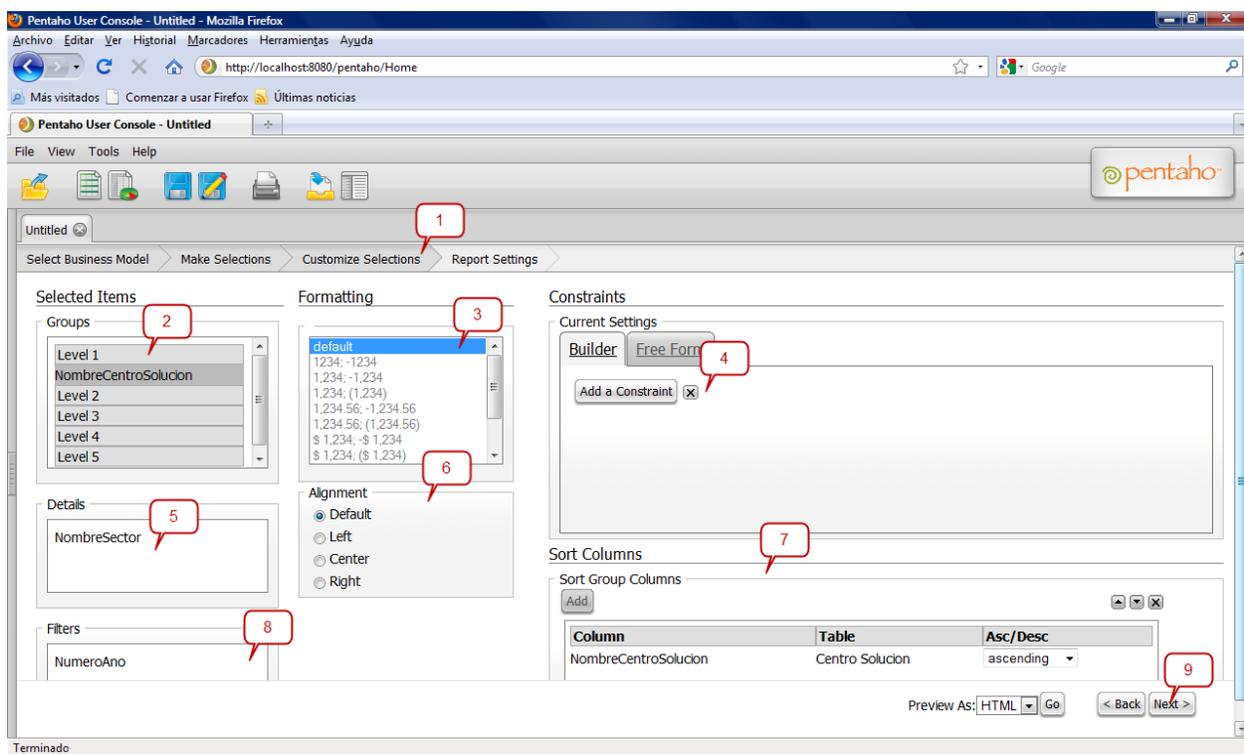


Figura 4.248: Personalización del formato del reporte

1	<b>Customize Selections</b> « <i>Personalizar Selección</i> » Tercera fase de la creación de reportes, en la que se define elementos de formato y restricciones de contenido.
2	<b>Groups</b> « <i>Grupos</i> » Al hacer clic sobre algún elemento de los grupos se habilitara los elementos de formato (3) y alineación (6), dependiendo del grupo se podrá definir o configurar dichos elementos.
3	<b>Formatting</b> « <i>Formato</i> » Se define el tipo de dato que se mostrara en el caso de ser tipo moneda, defecto, positivos, negativos o ambos y la forma en que se representaran dichos valores.
4	<b>Constraints</b> « <i>Restricciones</i> » Existen varias formas de utilizar las restricciones o CONSTRAINTS cada elemento (GRUPO, DETALLE o FILTRO) puede tener su propia restricción (constraint) sin embargo es necesario saber que se desea poner en cada restricción, el propósito principal es limitar el espacio de los datos mostrados para lograr tener un panorama más específico de la

	información.
5	<b>Detail</b> « <i>Detalle</i> » Elementos definidos para el detalle los cuales pueden ser configurados de igual manera a los grupos, agregar formato, agregar restricción y configurar el alineamiento.
6	<b>Alingment</b> « <i>Alineamiento</i> » Se utiliza para configurar la presentación de los datos en cuanto a formato de párrafos o grupos de datos si se desea normal, alineado a la izquierda, alineado a la derecha o centrado.
7	<b>Sort Columns</b> « <i>Configurar los datos de las columnas</i> » si se desea visualizar de manera ascendente o descendente.
8	<b>Filters</b> « <i>Filtros</i> » por medio de que hemos filtrado el reporte se pueden definir varios filtros.
9	<b>NEXT</b> para continuar al siguiente paso.

**Report Settings** «*Propiedades del Reporte*» Es el último paso para generar un reporte en el cual se definen elementos como tamaño de papel, orientación de la pagina, encabezado y pie de página

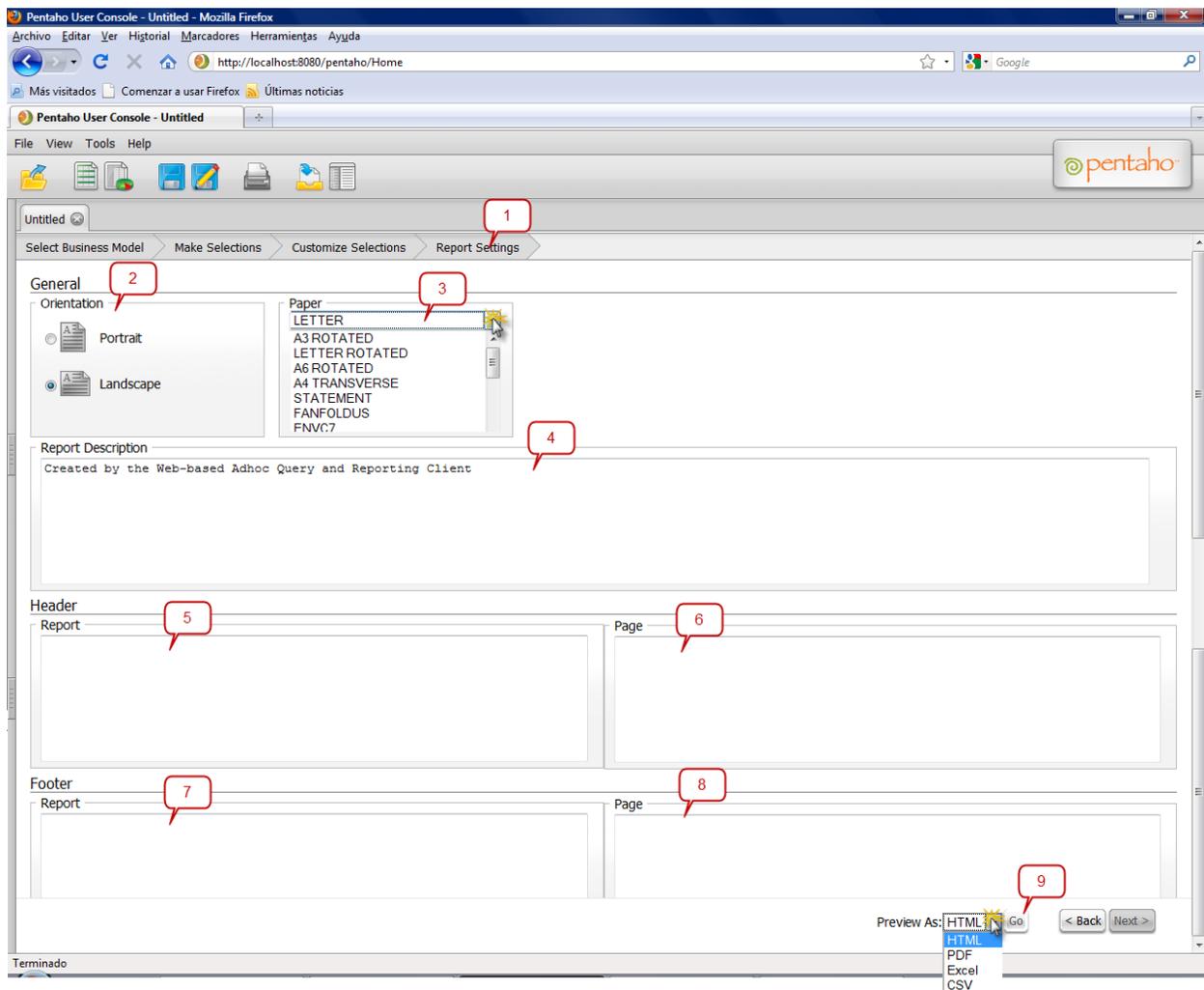


Figura 4.249: Configuraciones generales del reporte

1	<b>Report Settings</b> « <i>Propiedades del Reporte</i> » Ultimo paso de generación de reporte en el que se definen elementos de formato final.
2	<b>Orientation</b> « <i>Orientación</i> » Como deseamos visualizar la pagina, horizontal o verticalmente.
3	<b>Paper</b> « <i>Papel</i> » Se define el tamaño de la hoja de papel en la que se desea generar el reporte.

4	<b>Description</b> « <i>Descripción</i> » Se puede adicionar una breve descripción para futuras referencias.
5	<b>Header Report</b> « <i>Encabezado de reporte</i> » Permite definir un encabezado para el reporte.
6	<b>Header Page</b> « <i>Encabezado de Página</i> » Permite definir un encabezado para las paginas.
7	<b>Footer Report</b> « <i>Pie de página de reporte</i> » Permite definir un pie de página al reporte.
8	<b>Footer Page</b> « <i>Pie de página</i> » Permite definir un pie de página
9	<b>Go</b> « <i>Ir...</i> » Antes de generar el reporte se puede visualizar para luego poder publicarlo, las vistas preliminares pueden ser una nueva ventana, en archivos Excel, PDF o CSV.

Una vez configurado ya se habilita la opción de guardar y guardar como, esta opción se utiliza para almacenar el reporte en el área de trabajo que ha sido destinada para el usuario activo.

De igual manera se habilita el botón de impresión, desde aquí podemos imprimir el reporte hacia un impresor que tengamos configurado en nuestra maquina.

**EJEMPLO DE REPORTE SENCILLO:**

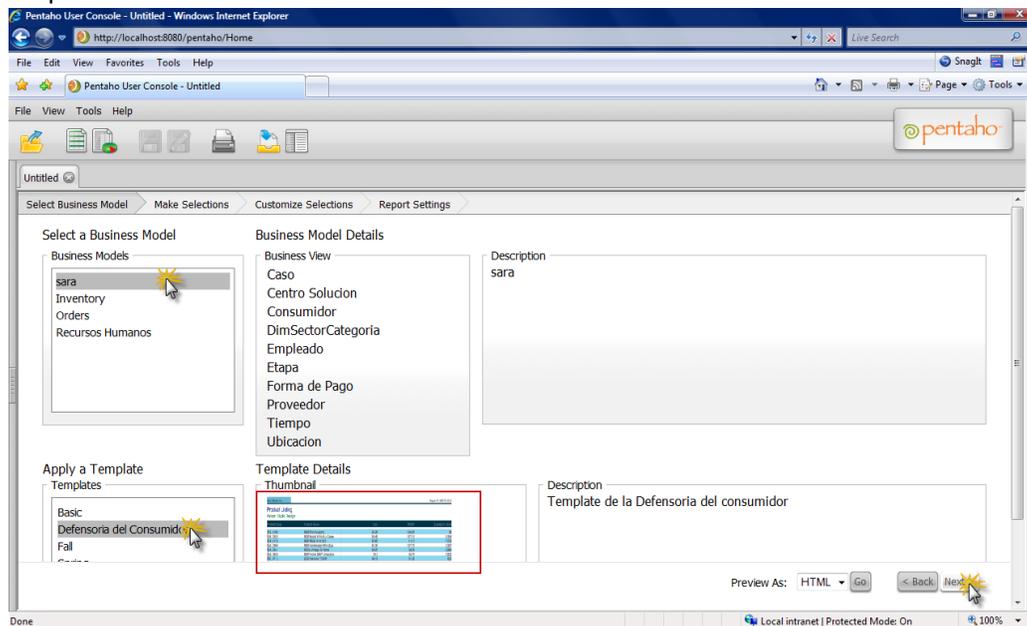
**Ojo: los datos que se mostraran en este ejemplo no coinciden con los datos reales esta es una prueba para demostrar el uso de generador de reportes, y facilitar la comprensión para los lectores de este documento.**

Habiendo iniciado sesión y al estar en la pantalla de inicio de sistema hacer clic en **NEW REPORT**



**Figura 4.250: Icono nuevo reporte, desde ventana de navegación**

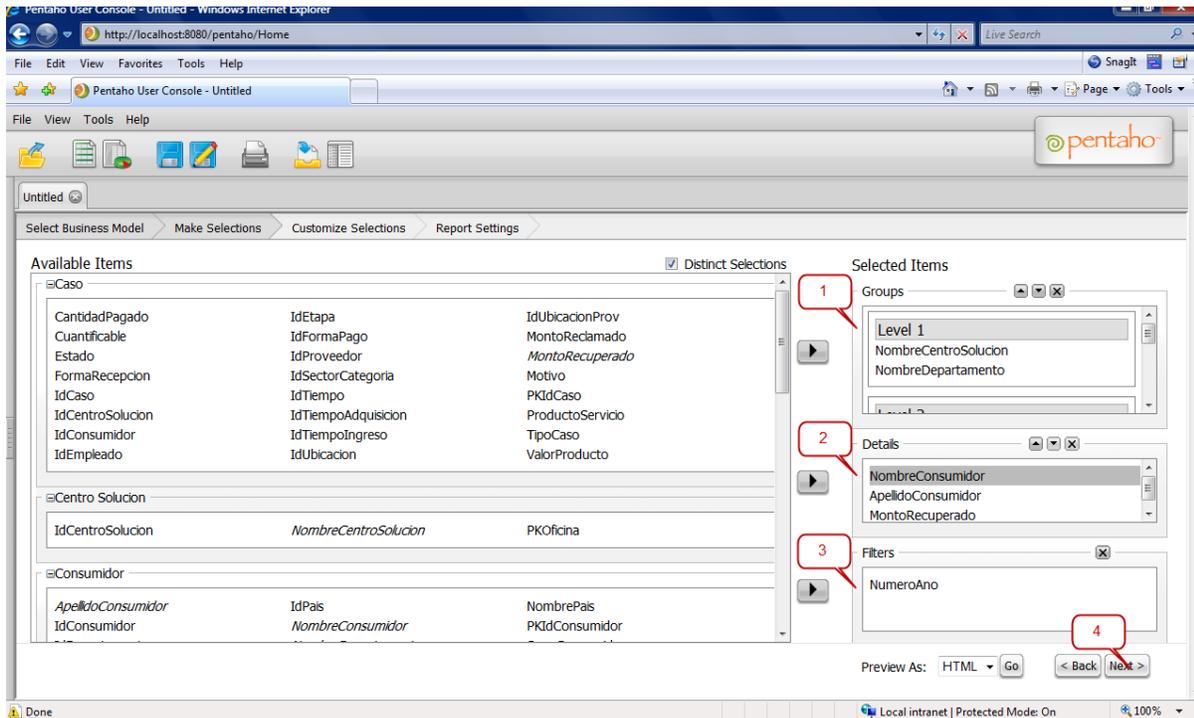
En la pantalla de Select bussines model seleccione la opción adecuada para el modelo que desea generar el reporte en nuestro caso el modelo le hemos llamado **SARA**



**Figura 4.251: Selección de Plantilla para el reporte**

Como plantilla de reporte seleccione **DEFENSORÍA DEL CONSUMIDOR** y se mostrara en ese instante el **THUMBNAIL** (vista preliminar) de la plantilla seleccionada finalmente hacer clic en botón **NEXT**

En la ventana **MAKE SELECTIONS** tendremos la siguiente configuración:



**Figura 4.252: Personalización de los elementos de datos del menú**

1	En <b>GROUPS</b> : <i>Nombre del centro de solución</i> y <i>Nombre de departamento</i>
2	En <b>DETAILS</b> : <i>Nombre Consumidor</i> , <i>Apellido Consumidor</i> y <i>Monto Recuperado</i>
3	En <b>FILTERS</b> : <b>Numero Año</b>
4	Una vez seleccionado proseguimos con un clic en <b>NEXT</b>

Lo que pretendemos mostrar es el monto recuperado por consumidores en los casos demandados, hay casos en los que no se logra obtener beneficio monetario aquí deseamos visualizar cual es el monto de recuperación de los casos agrupados por el centro de solución de las denuncias relacionadas con el departamento de Santa Ana para el año 2008, trataremos de visualizar aquellos elementos en los que el monto de recuperación sea mayor a cero y el tipo de dato moneda para los montos.

Toda esta configuración lo logramos hacer en el siguiente paso, en el cual definimos las restricciones (**CONSTRAINTS**), formatos de datos y agrupaciones.

Vamos a generar las restricciones paso a paso para identificar el uso de la herramienta de búsqueda que poseemos

En la ventana **CUSTOMIZE SELECTIONS** clic sobre **NombreDepartamento** ahora clic **ADD A CONSTRAINT**

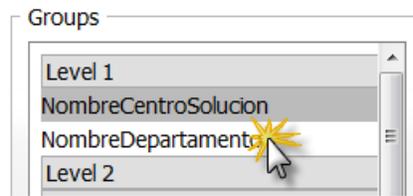


Figura 4.253: Personalización de los grupos

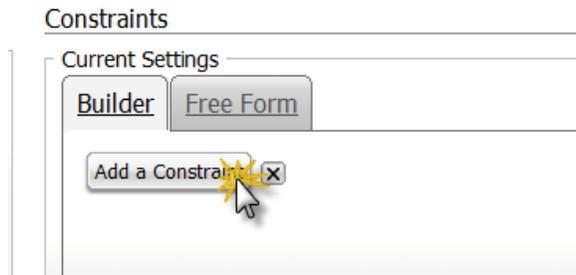


Figura 4.254: Personalización de las restricciones para el reporte

Al hacer clic inmediatamente aparecerá una línea con la restricción que estamos agregando, ahora utilizaremos la herramienta de búsqueda (representado por una lupa), inmediatamente aparecerá un cuadro de búsqueda de datos

**Constraints**

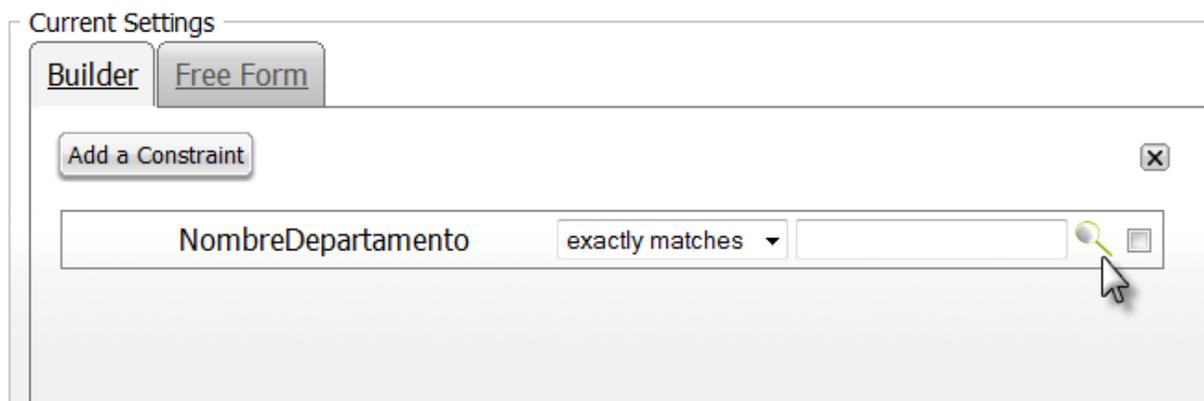


Figura 4.255: Modificación de las restricciones para el reporte

Es muy importante recalcar que debemos saber que dato es el que deseamos encontrar, dado que le estamos diciendo que nos muestre datos exactos, en el caso escribiremos **Santa Ana** y seguidamente clic en el botón **SEARCH**

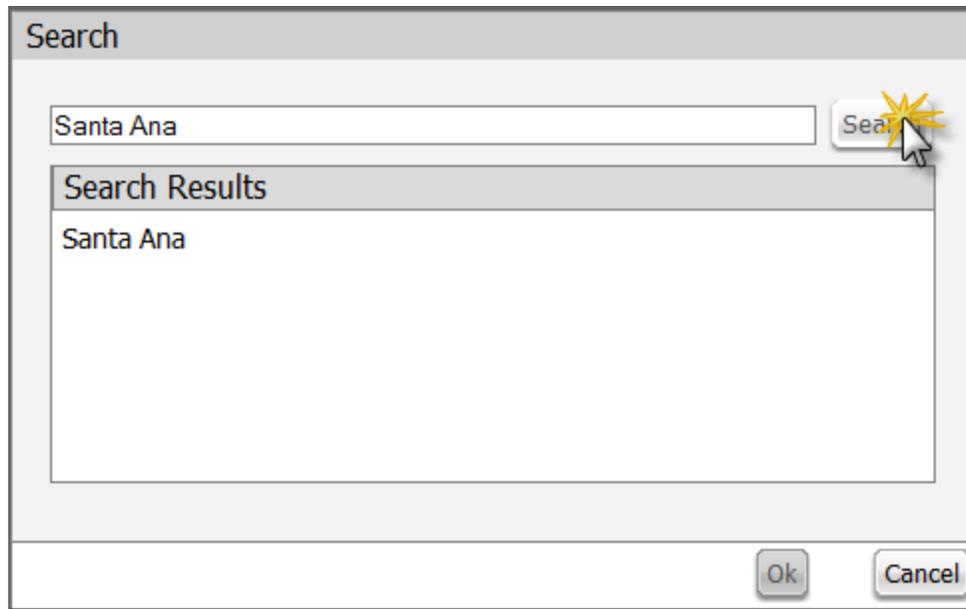


Figura 4.256: Búsqueda de datos para las restricciones

Una vez mostrados los resultados **SELECCIONAMOS** el deseado y luego el botón **OK**

Inmediatamente aparecerá el valor seleccionado en el campo de la restricción al que deseamos limitar.

### Constraints

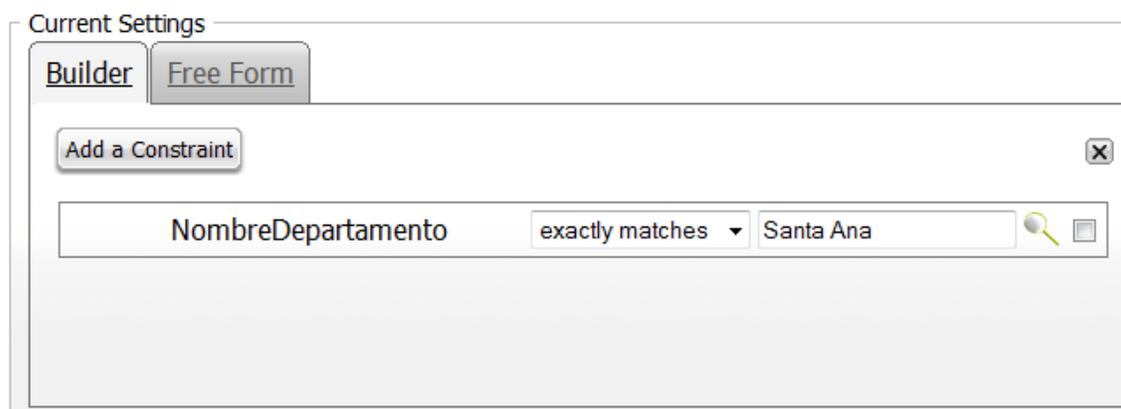


Figura 4.257: Finalización de la construcción de las restricciones para reporte

Ahora configuramos el monto recuperado para que nos muestre el símbolo de dólar y valores mayores a cero.

Primero seleccionamos el detalle **MONTORECUPERADO (1)**

Luego en **NUMERIC FORMAT** el tipo de datos **\$1,234: -\$1,234 (2)**

Seguidamente agregamos la restricción **(3)**

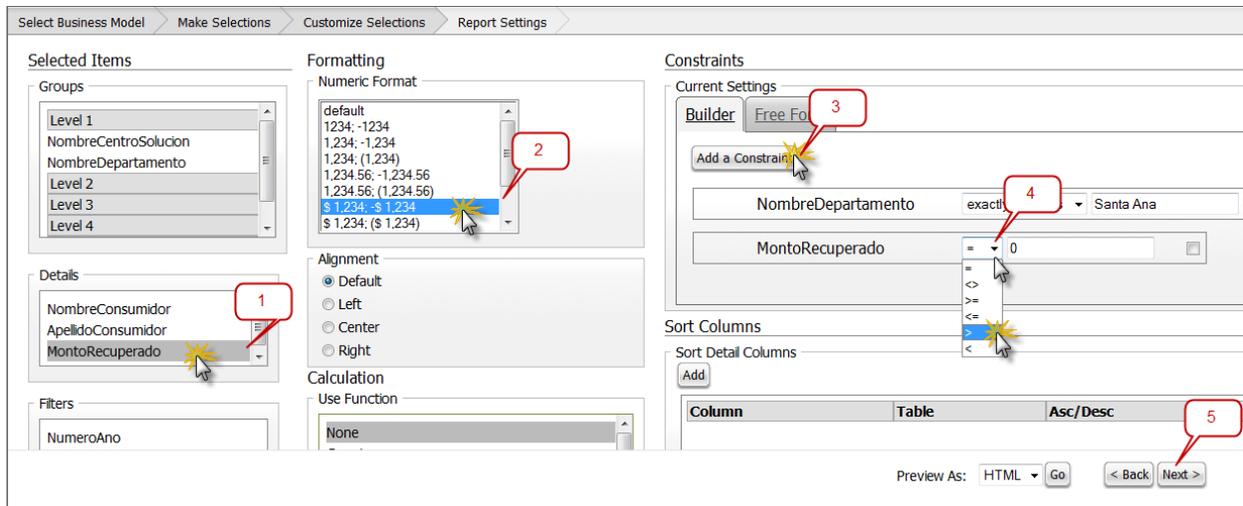


Figura 4.258: Configuración del formato para una columna

En la restricción nos aparece una lista desplegable de operadores de valor **(4)** Ahí seleccionamos el símbolo **MAYOR QUE >** y finalmente en el campo escribimos **0** para limitar los datos a valores mayores que cero.

Como paso final de la personalización de las selecciones agregaremos la restricción para el año que sea igual a 2008, similar al caso del monto recuperado seleccionamos el filtro **NUMEROANO** y luego **AGREGAMOS LA RESTRICCIÓN** seleccionando el **OPERADOR IGUAL (=)** y escribimos **2008**.

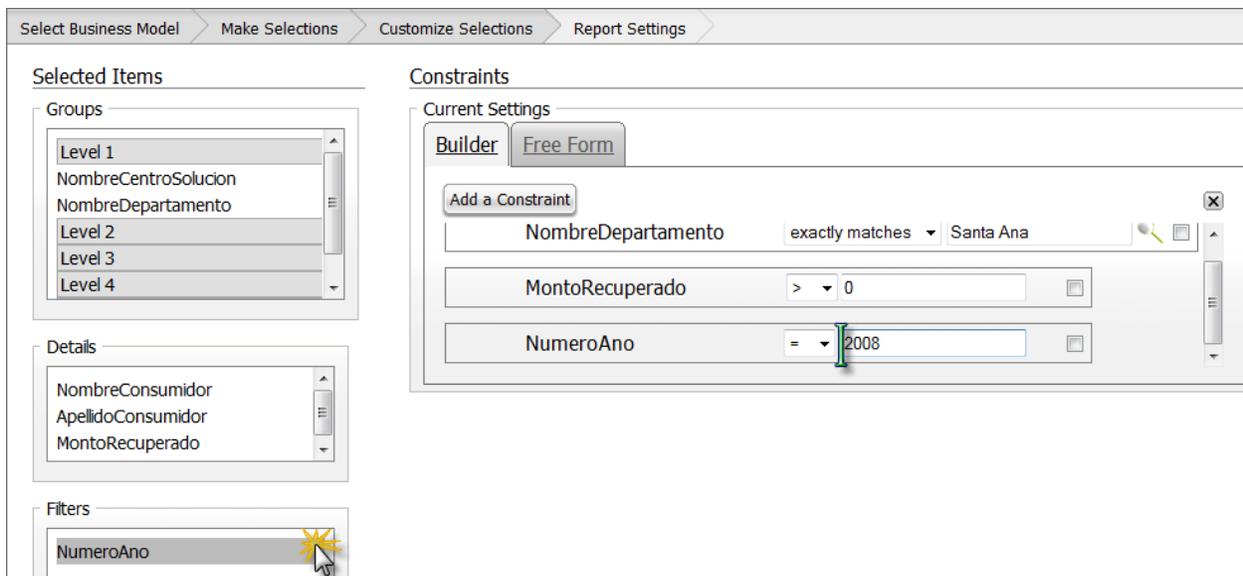


Figura 4.259: Agregando otra restricción al reporte

El reporte está configurado con los datos que mostrara ahora no queda más que darle clic al botón **NEXT** para definir el formato de los elementos extras que llevara el reporte así como la orientación de la página.

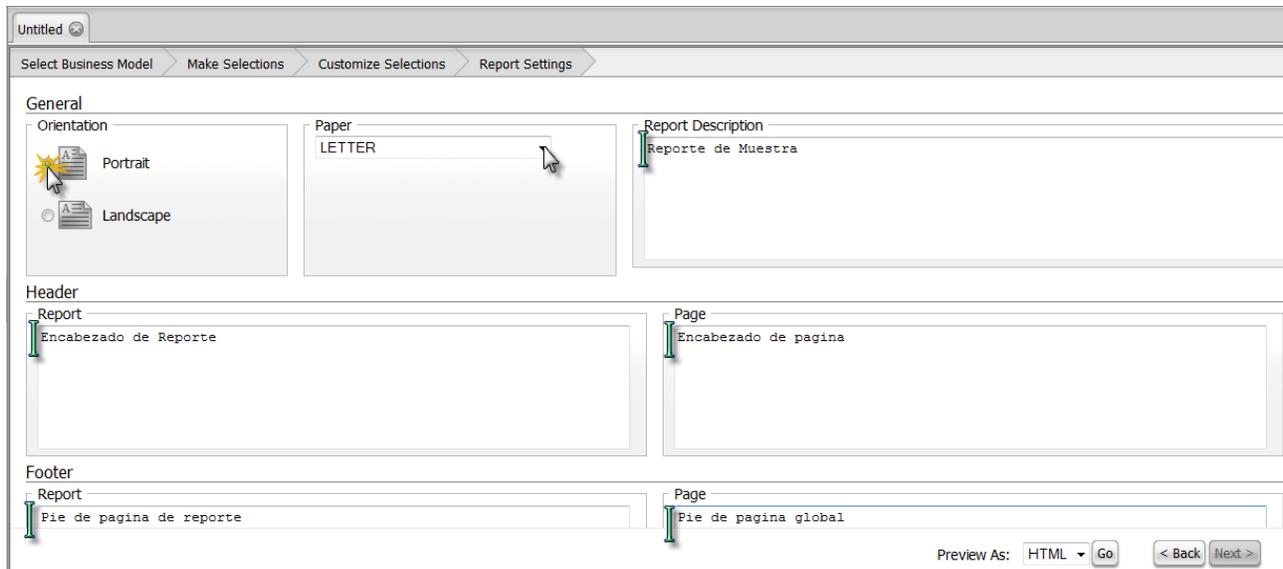


Figura 4.260: Configuraciones generales del reporte

Seleccionamos la orientación, el tamaño del papel **CARTA (LETTER)**, la descripción de reporte, los encabezados de página y de reporte así como los pie de página del reporte como de las páginas.

Ahora tenemos 4 tipos de formas de visualizar el resultado final.

- En la web
- En PDF
- En Excel o
- En CSV

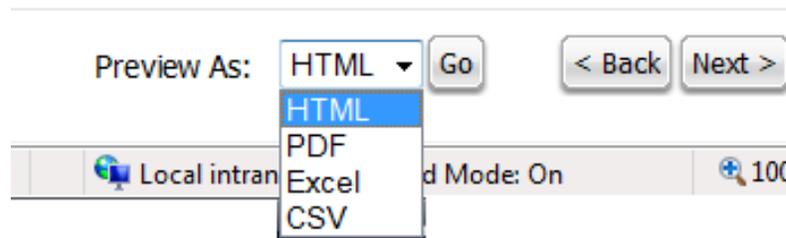


Figura 4.261: Formatos a los que se puede realizar los reportes

Seleccionamos la opción HTML para visualizarlo en el navegador actual

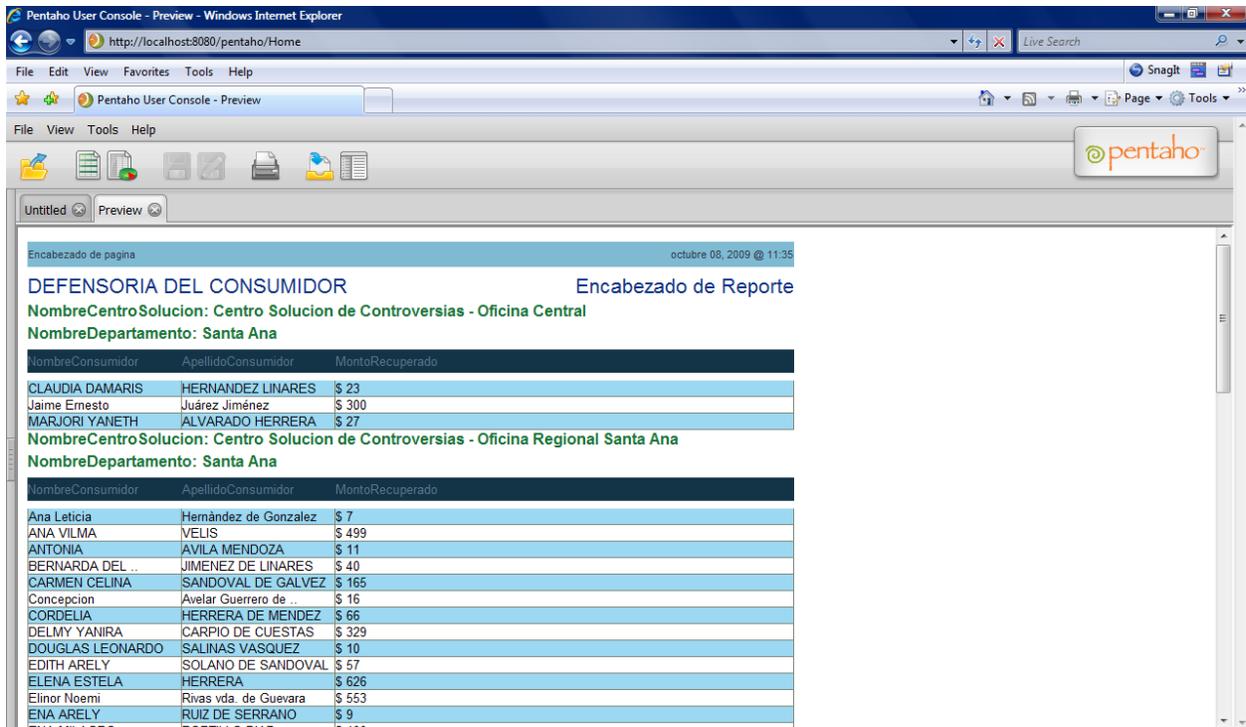


Figura 4.262: Muestra del reporte generado en html

Si se necesita visualizar el pdf se selecciona del menú desplegable la opción PDF



Figura 4.263: Muestra del reporte generado en pdf

En el caso de querer guardar el documento en pdf seleccionar la opción guardar del navegador y seleccionar la carpeta destino en la que se desea almacenar el documento.

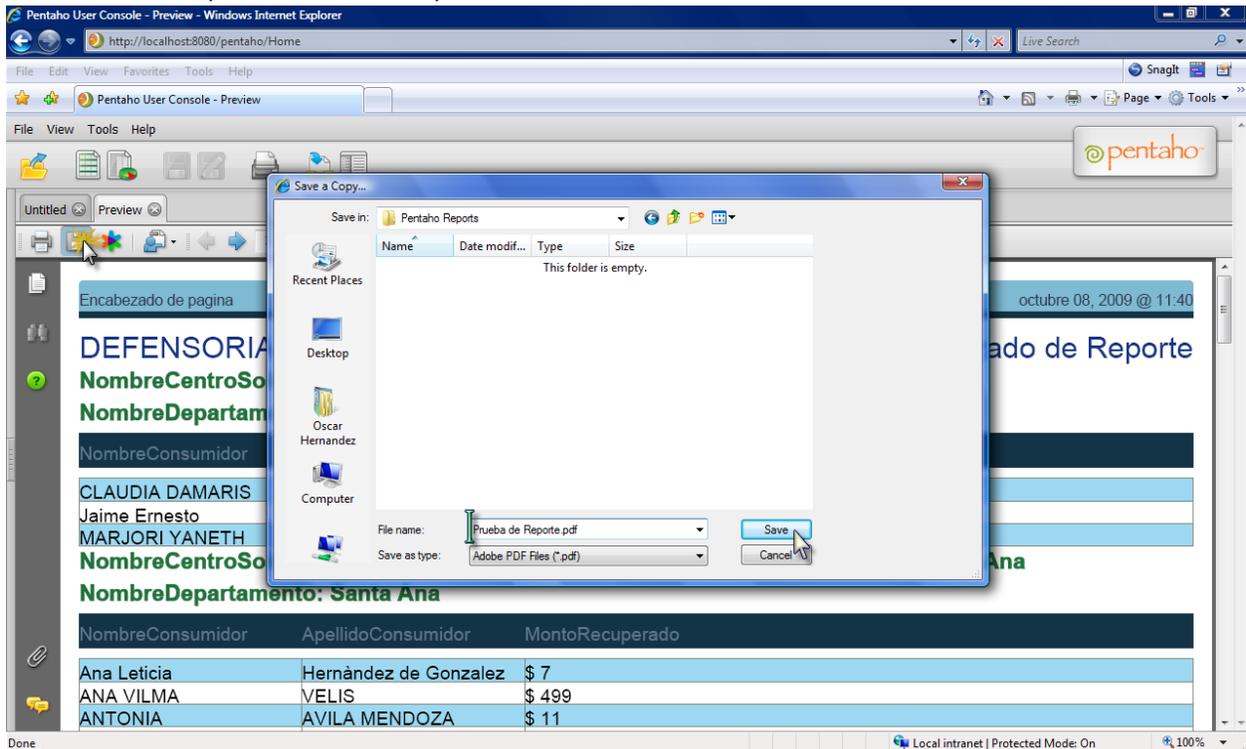


Figura 4.264: Guardando un reporte generado en pdf

Asimismo con el formato EXCEL nos muestra una pantalla con la que nos permite decidir que deseamos hacer si guardar, abrir o cancelar, esta opción crea un archivo de formato XLS

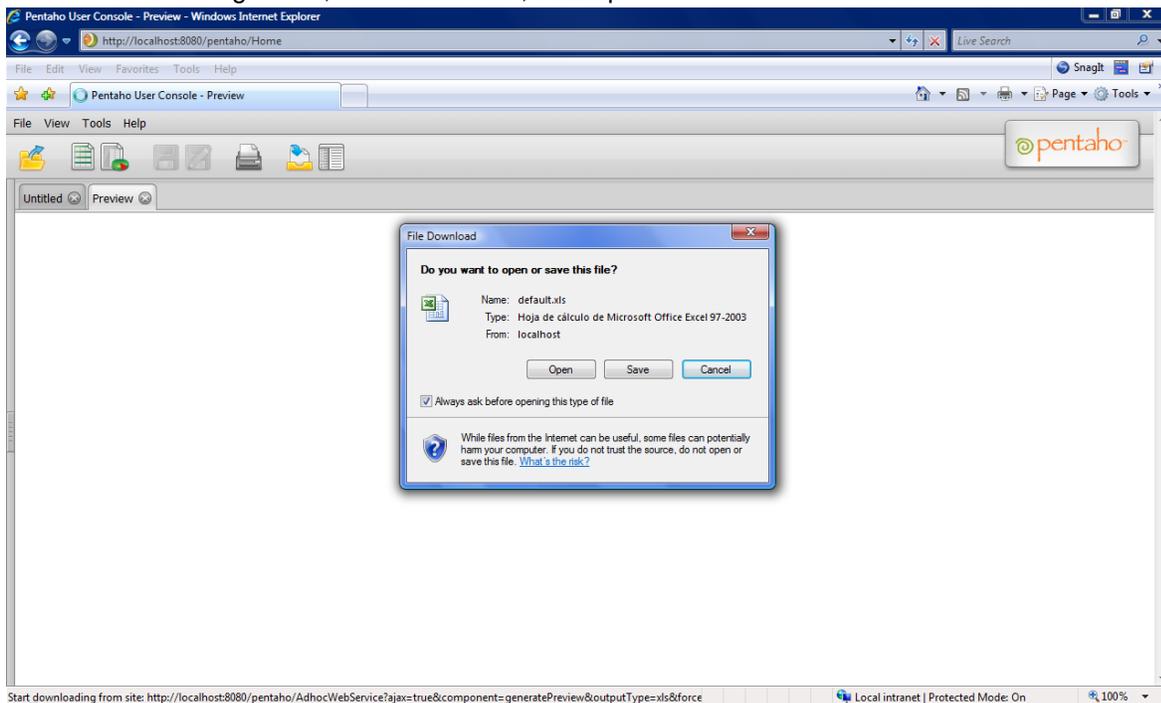


Figura 4.265: Muestra del reporte generado en xls

Como última opción la generación de archivos **CSV** el cual puede ser abierto con **EXCEL** de igual manera que el anterior.

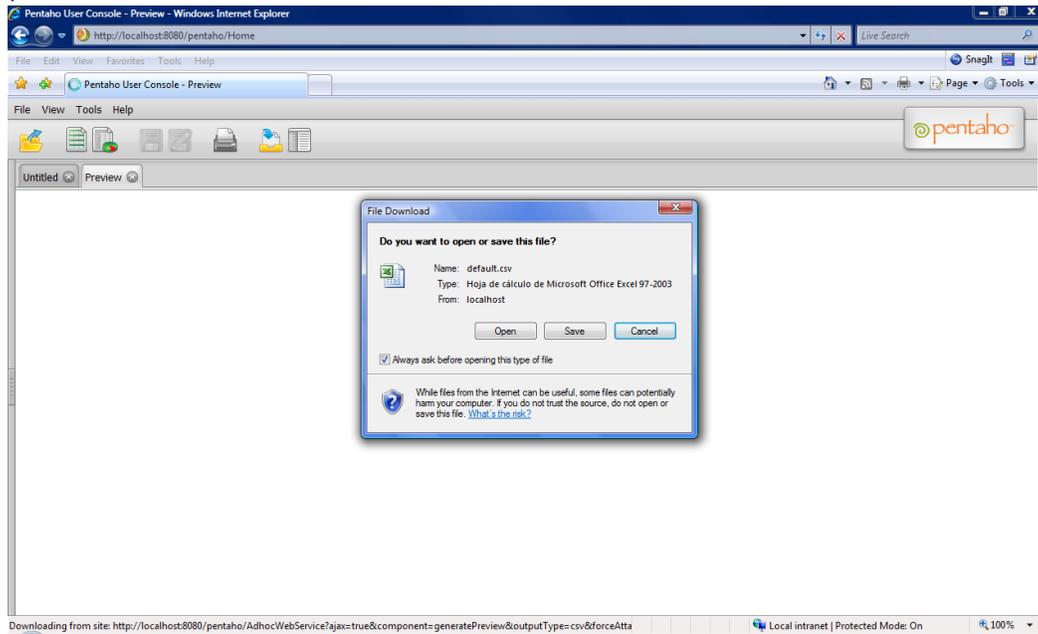


Figura 4.266: Muestra del reporte generado en csv

Como opción adicional se encuentran habilitadas las opciones de almacenar o de imprimir sin necesidad de visualizarlo de las cuatro maneras mencionadas anteriormente estas opciones son útiles cuando el reporte es generado con frecuencia ya que se almacena y cuando se abre se logra visualizar los datos creados en un determinado momento, solo especificamos el nombre del reporte.

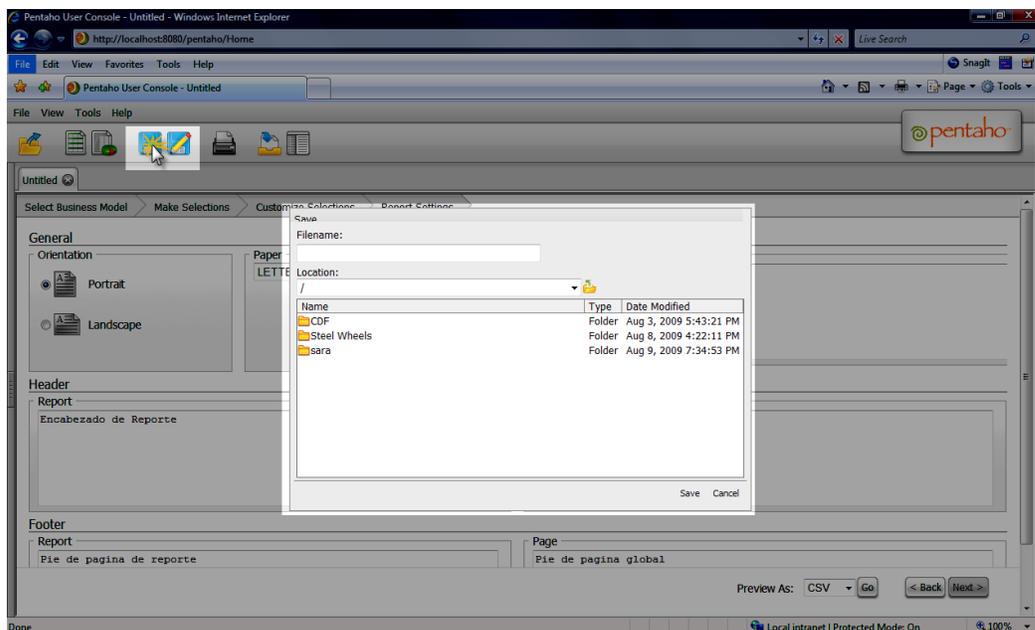


Figura 4.267: Guardando reporte generado en la aplicación

Y la opción de imprimir nos permite enviar el documento a algún impresor que se tenga configurado en el ordenador.

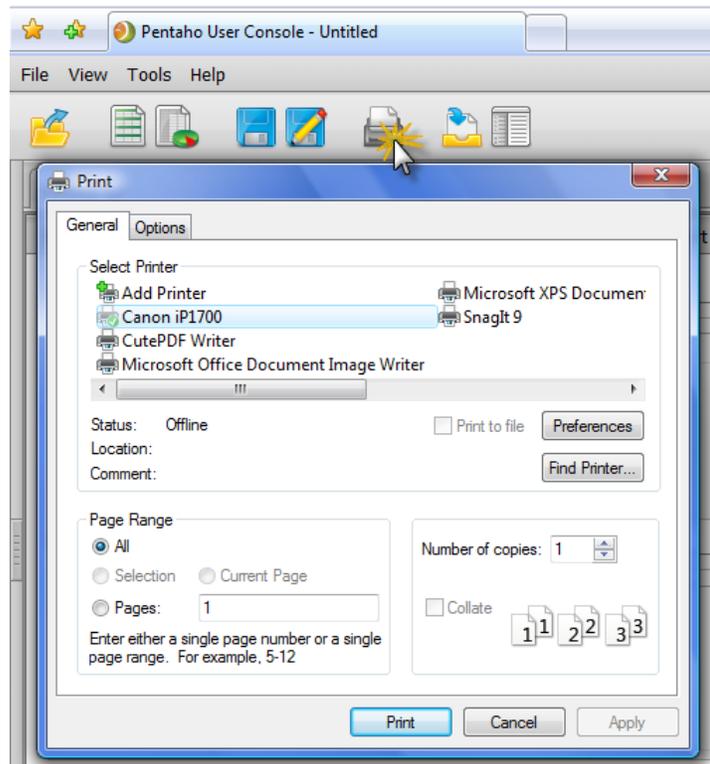


Figura 4.268: Imprimiendo reporte generado

Básicamente estas son las opciones que tenemos con la creación de reportes que brinda la herramienta cualquier duda o inquietud se recomienda ir al sitio <http://wiki.pentaho.com> en el que se brinda ayuda y herramienta de consulta mediante foros (generalmente en inglés)

#### 4.3.4. Creación de Vista de Análisis

##### 4.3.4.1. New Analysis View «Nueva vista de análisis»



Figura 4.269: Icono de creación de un nuevo análisis

La vista de análisis permite generar análisis que facilita la toma de decisiones, brindando un acceso inmediato a la información y a las áreas de negocio a las que se le ha realizado un estudio y especialización con la creación de **CUBOS DE INFORMACIÓN**, los cubos disponibles proporcionan datos y vistas analíticas que

Se inicia el generador de vistas de análisis e inmediatamente se solicita la selección del cubo con el que se desea trabajar.

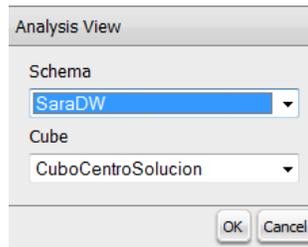


Figura 4.270: Eligiendo vista de análisis y cubo para análisis

Una vez se selecciona el cubo se procede a dar ok para iniciar el espacio de trabajo del cubo elegido, donde se cargaran las vistas dimensionales, las configuraciones de las medidas con las que se podrá trabajar.

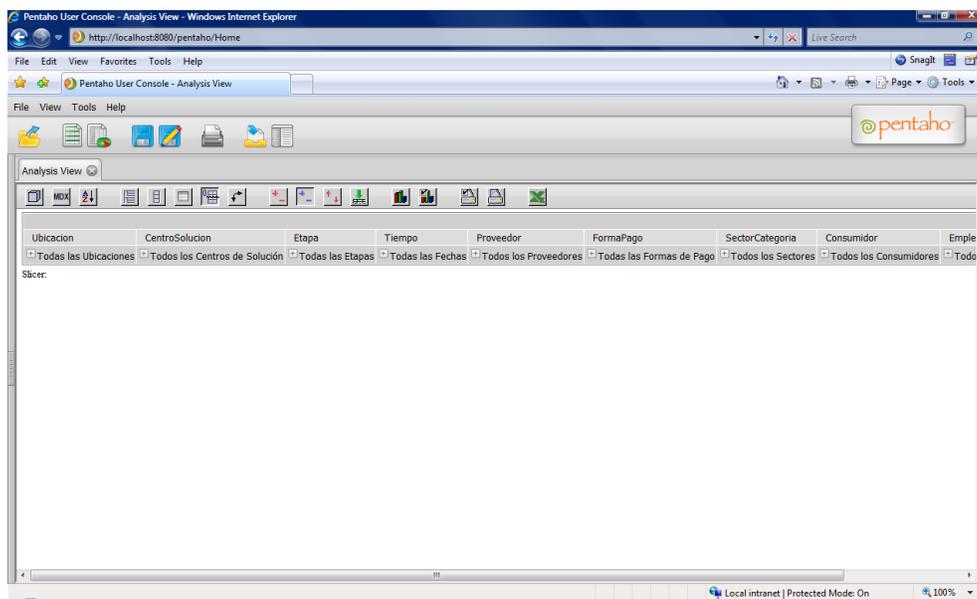


Figura 4.271: Presentando dimensiones disponibles para análisis

El espacio de trabajo contiene elementos comunes con el proceso de creación de reportes, sin embargo se tiene una barra de herramientas la cual facilita ciertas tareas para la generación de vistas.



Figura 4.272: Barra de herramientas para los análisis

El otro elemento sobre el cual la barra de herramientas acciona es la **HOJA DE DATOS** o **DATA GRID**, la cual es una cuadrícula estructurada con variables dimensionales y medidas con las cuales se realiza el análisis de la información.

			Medidas
Tiempo	CentroSolucion	SectorCategoria	▼ CantCaso
+ Todas las Fechas	- Todos los Centros de Solución	+ Todos los Sectores	5,789

Slicer:

Figura 4.273: Muestra de un análisis

Con el data grid se pueden realizar muchas vistas y configuraciones tales como filtrar ordenar por columnas o filas agregar medidas, configurar las medidas, eliminar datos repetidos entre otros, los cuales serán explicados y ejemplificados en el presente documento.

Cada una de las funciones será descrita brevemente a continuación:



**Abrir navegador OLAP:** Se utiliza para hacer SLICE a los datos. Imaginémonos un pastel el slice como su nombre lo indica en ingles es hacer un corte de una porción de un todo, es prácticamente reducir a un grupo o segmento del universo de los datos al que se tiene acceso. Aquí se definen los filtros y la estructura de los datos si se desea visualizar en filas o columnas, también se establecen las medidas con las que se desea trabajar. Es muy importante aclarar que todos los datos con los que se puede trabajar y a los que se puede acceder es el producto del trabajo realizado en conjunto con los usuarios, agregar elementos a estas estructuras es posible sin embargo lleva un análisis y cambios en procesos propios de los desarrolladores de dicho proyecto.

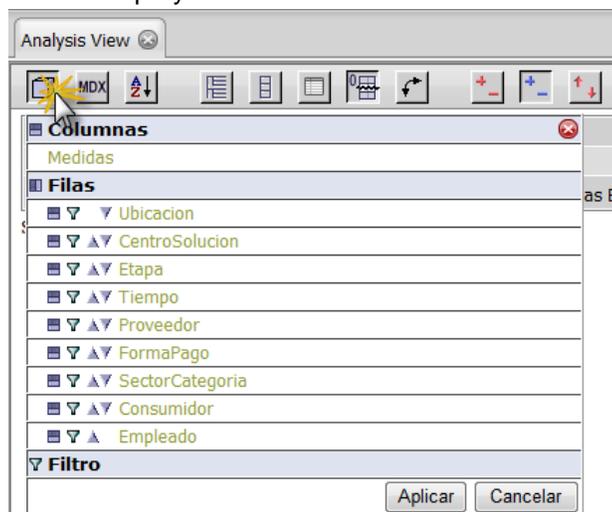


Figura 4.274: Navegador OLAP

Si se desea quitar del análisis algún elemento se debe seleccionar el botón de filtro del elemento que se desea suprimir , de esta manera quitamos de nuestra vista de análisis los datos que se muestran por defecto, en el caso que se desee realizar un análisis más específico de los datos. Los botones de ordenamiento nos facilitan distribuir los datos dependiendo de la vista que se desee crear para tener la información distribuida de una manera específica.  Para cambiar de posición a filas o columnas se utilizan los botones  para pasar a columnas o  para pasar a filas.

Icono	Descripción
	Mover Dimensiones al área de columnas
	Mover dimensión al área de filas
	Mover dimensión al área de filtros
	Mover dimensión hacia arriba o hacia abajo en el área actual

Con las vistas no existe estándar de creación de documentos dependerá de cómo se desean visualizar los datos por parte de los usuarios, el uso de la herramienta facilita la comprensión del acceso que se posee.

**Explorar las dimensiones:** se puede especializar las medidas y dimensiones que se posee esta función es sencilla para realizar un filtrado más específico de los datos, es un método de agrupación gráfico interactivo y sencillo.

A continuación un ejemplo de este sencillo método, haciendo uso de la dimensión tiempo.



Figura 4.275: Configuración de las dimensiones

Al hacer clic en la dimensión tiempo se logra ingresar a los datos más detallados de esta dimensión como lo vemos a continuación:

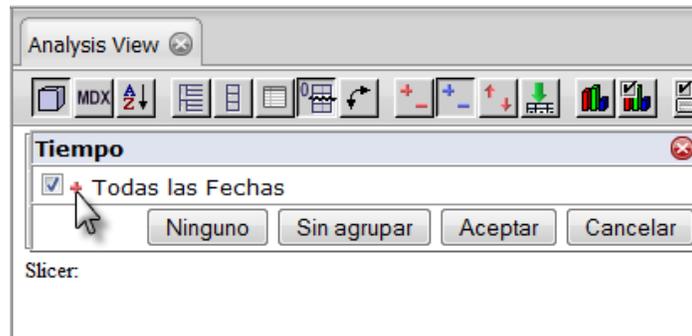


Figura 4.276: Configuración de los filtros para las dimensiones

Se puede seleccionar más detalle de fechas al hacer clic en el símbolo +.

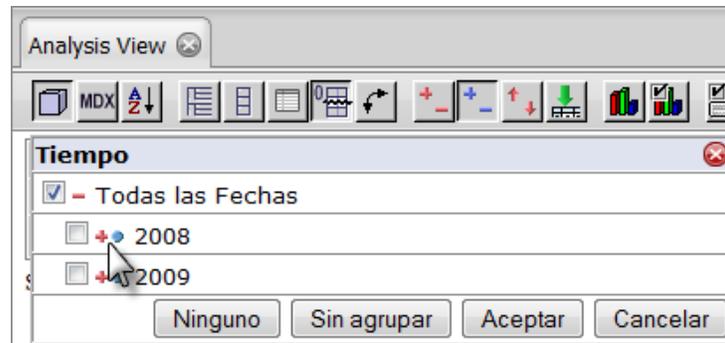


Figura 4.277: Muestra de Información disponible en las dimensiones

El primer nivel que tenemos disponibles es las fechas por nivel de años.

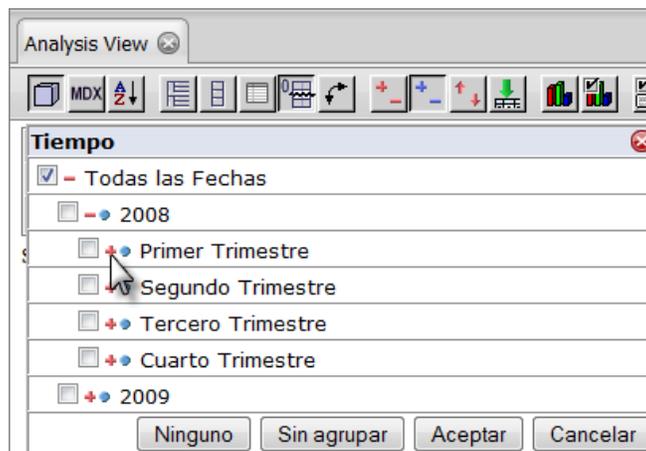


Figura 4.278: Exploración de las dimensiones desde el navegador OLAP

Dentro de cada año tenemos el análisis por trimestres.

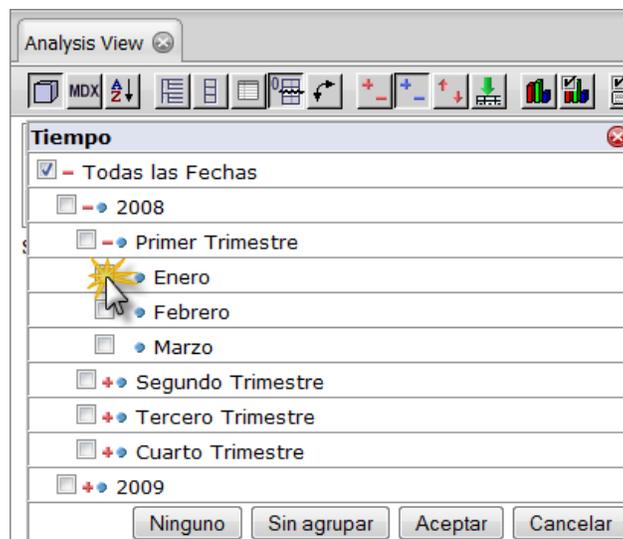


Figura 4.279: Eligiendo un filtro para la dimensión

Como ultimo nivel podemos limitar las vistas a nivel mensuales. Es importante aclarar que se pueden limitar diversos elementos de tipo tiempo para un solo análisis, esto dependerá de la comprensión completa de la herramienta y la capacidad de determinar lo que se desea visualizar.



**Editor MDX:** se posee de una herramienta de creación de vistas haciendo uso de lenguaje MDX, esta función puede ser utilizada por usuarios más experimentados y que poseen la capacidad de generar vistas haciendo uso de dicho lenguaje. Las bondades del MDX es que se posee suficiente información con la cual se puede dominar la técnica de creación de dichos reportes para generar reportes en base a las necesidades.

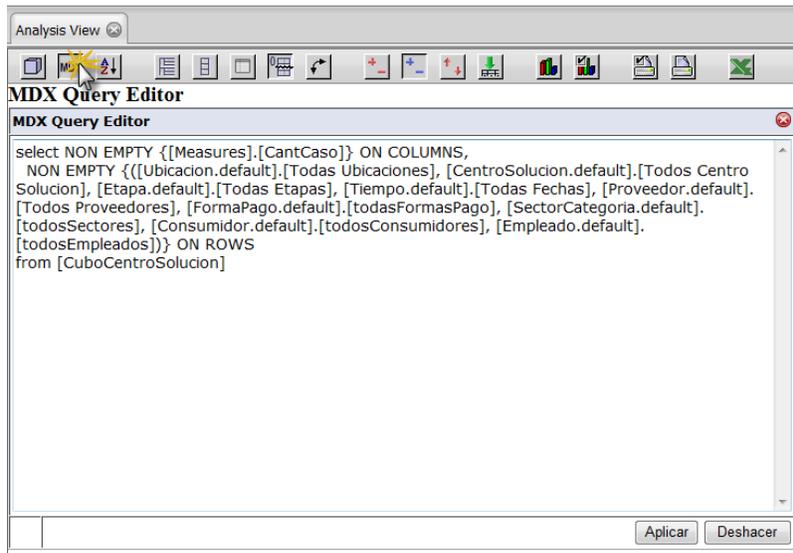


Figura 4.280: Editor de Consulta MDX



**Configurar tabla OLAP:** se utiliza para configurar la ordenación de de los datos que se desplegarán al hacer clic tendremos una ventana como la siguiente:

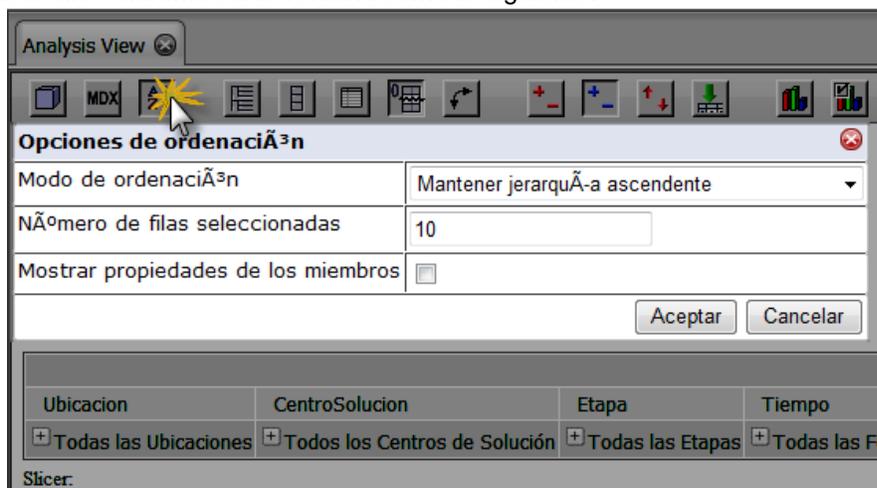


Figura 4.281: Configuración de tabla OLAP

El Número de Filas seleccionadas se utiliza para mostrar los mayores valores de una medida específica, es necesario ordenar ya sea ascendentemente o descendentemente por una medida específica para luego



**Mostrar Padres:** Muestra los indicadores de las vistas dimensionales que se poseen, esta función es de gran ayuda para tener una mejor guía de la vista, facilita determinar los niveles de detalles a los que se está accediendo.

Ubicación	CentroSolucion	Etapa	SubEtapa	Tiempo	Proveedor	
(All)	(All)	(All)		(All)	(All)	
+ Todas las Ubicaciones	- Todos los Centros de Solución	+ Todas las Etapas		+ Todas las Fechas	+ Todos los Proveedores	
		+ Todas las Etapas	- Ingresado	+ Todas las Fechas	+ Todos los Proveedores	
			Ingresado	Ingresado en CSCa	+ Todas las Fechas	+ Todos los Proveedores
				Ingresado en Web	+ Todas las Fechas	+ Todos los Proveedores
				Pendiente de archivar	+ Todas las Fechas	+ Todos los Proveedores
				Pendiente de remitir a Tribunal Sancionador	+ Todas las Fechas	+ Todos los Proveedores
				Avenimiento	+ Todas las Fechas	+ Todos los Proveedores

Figura 4.282: Mostrar padres en los análisis



**Ocultar Repeticiones:** Es utilizado para prevenir la repetición de encabezados duplicados cuando múltiples dimensiones están anidadas en una sola tabla de análisis.

CentroSolucion	Tiempo	SectorCategoria
- Todos los Centros de Solución	+ Todas las Fechas	+ Todos los Sectores
Centro Solucion de Controversias - Oficina Central	+ Todas las Fechas	+ Todos los Sectores
Centro Solucion de Controversias - Oficina Central	+ 2008	+ Todos los Sectores
Centro Solucion de Controversias - Oficina Central	+ 2009	+ Todos los Sectores
Centro Solucion de Controversias - Oficina Regional San Miguel	+ Todas las Fechas	+ Todos los Sectores
Centro Solucion de Controversias - Oficina Regional San Miguel	+ 2008	+ Todos los Sectores
Centro Solucion de Controversias - Oficina Regional San Miguel	+ 2009	+ Todos los Sectores
Centro Solucion de Controversias - Oficina Regional Santa Ana	+ Todas las Fechas	+ Todos los Sectores
Defensoria Financiera	+ Todas las Fechas	+ Todos los Sectores
Oficina Call Center Plan de La Laguna	+ Todas las Fechas	+ Todos los Sectores
Oficina Plan de La Laguna	+ Todas las Fechas	+ Todos los Sectores
Oficina Educación	+ Todas las Fechas	+ Todos los Sectores

Slicer:

Figura 4.283: Tabla de datos repetidos

CentroSolucion	Tiempo	SectorCategoria	Medidas
[-] Todos los Centros de Solución	[+] Todas las Fechas	[+] Todos los Sectores	
Centro Solucion de Controversias - Oficina Central	[-] Todas las Fechas	[+] Todos los Sectores	
	[+] 2008	[+] Todos los Sectores	
	[+] 2009	[+] Todos los Sectores	
Centro Solucion de Controversias - Oficina Regional San Miguel	[-] Todas las Fechas	[+] Todos los Sectores	
	[+] 2008	[+] Todos los Sectores	
	[+] 2009	[+] Todos los Sectores	
Centro Solucion de Controversias - Oficina Regional Santa Ana	[+] Todas las Fechas	[+] Todos los Sectores	
Defensoria Financiera	[+] Todas las Fechas	[+] Todos los Sectores	
Oficina Call Center Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	
Oficina Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	
Oficina Educación	[+] Todas las Fechas	[+] Todos los Sectores	

Slicer:

Figura 4.284: Tabla de datos sin repetición



**Mostrar propiedades:** Permite mostrar u ocultar las propiedades de los miembros definidos en el modelo OLAP definido. Para mayor información de cómo definir las propiedades de los miembros, por favor refiérase a la documentación de Mondrian en:

<http://mondrian.pentaho.org/documentation/schema.php>



**Suprimir filas/Columnas vacías:** Hay momentos en los que en los reportes se desea visualizar todos los datos, pero generalmente los elementos que poseen resultados ceros no son relevantes, este botón realiza la función de discriminar este tipo de resultados, con el propósito de mostrar datos que puedan ser de utilidad para los analistas.

CentroSolucion	Tiempo	SectorCategoria	Medidas
[-] Todos los Centros de Solución	[+] Todas las Fechas	[+] Todos los Sectores	5.789
Centro Solucion de Controversias - Oficina Central	[-] Todas las Fechas	[+] Todos los Sectores	
Centro Solucion de Controversias - Oficina Regional San Miguel	[-] Todas las Fechas	[+] Todos los Sectores	
Centro Solucion de Controversias - Oficina Regional Santa Ana	[+] Todas las Fechas	[+] Todos los Sectores	
Defensoria Financiera	[+] Todas las Fechas	[+] Todos los Sectores	
Oficina Call Center Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	
Oficina Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	
Oficina Educación	[+] Todas las Fechas	[+] Todos los Sectores	
Oficinas Infocentro	[+] Todas las Fechas	[+] Todos los Sectores	
Centro Solucion de Controversias - Oficina Central	[+] 2008	[+] Todos los Sectores	3.930
	[+] 2009	[+] Todos los Sectores	155
Centro Solucion de Controversias - Oficina Regional San Miguel	[+] 2008	[+] Todos los Sectores	804
	[+] 2009	[+] Todos los Sectores	64

Figura 4.285: Tabla de datos incluyendo cero

Se logra visualizar datos con cero en este reporte.

CentroSolucion	Tiempo	SectorCategoria	Medidas
[-] Todos los Centros de Solución	[+] Todas las Fechas	[+] Todos los Sectores	5.789
Centro Solucion de Controversias - Oficina Central	[-] Todas las Fechas	[+] Todos los Sectores	4.085
	[+] 2008	[+] Todos los Sectores	3.930
	[+] 2009	[+] Todos los Sectores	155
Centro Solucion de Controversias - Oficina Regional San Miguel	[-] Todas las Fechas	[+] Todos los Sectores	868
	[+] 2008	[+] Todos los Sectores	804
	[+] 2009	[+] Todos los Sectores	64
Centro Solucion de Controversias - Oficina Regional Santa Ana	[+] Todas las Fechas	[+] Todos los Sectores	690
Defensoria Financiera	[+] Todas las Fechas	[+] Todos los Sectores	76
Oficina Call Center Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	18
Oficina Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	46
Oficina Educación	[+] Todas las Fechas	[+] Todos los Sectores	6

Figura 4.286: Tabla de datos sin incluir datos con cero

Reporte arreglado y configurado para que no muestre datos nulos o con valores cero.



**Intercambiar Ejes:** Esta es una de las funciones más poderosas de las vistas de análisis de Pentaho. Permite utilizar de pivote los datos, intercambiar las dimensiones miembro de filas con las dimensiones miembro de las columnas del área visual.

Vista de ejemplo con las medidas en la posición tradicional.

CentroSolucion	Tiempo	SectorCategoria	Medidas
[-] Todos los Centros de Solución	[+] Todas las Fechas	[+] Todos los Sectores	5.789
Centro Solucion de Controversias - Oficina Central	[-] Todas las Fechas	[+] Todos los Sectores	4.085
	[+] 2008	[+] Todos los Sectores	3.930
	[+] 2009	[+] Todos los Sectores	155
Centro Solucion de Controversias - Oficina Regional San Miguel	[-] Todas las Fechas	[+] Todos los Sectores	868
	[+] 2008	[+] Todos los Sectores	804
	[+] 2009	[+] Todos los Sectores	64
Centro Solucion de Controversias - Oficina Regional Santa Ana	[+] Todas las Fechas	[+] Todos los Sectores	690
Defensoria Financiera	[+] Todas las Fechas	[+] Todos los Sectores	76
Oficina Call Center Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	18
Oficina Plan de La Laguna	[+] Todas las Fechas	[+] Todos los Sectores	46
Oficina Educación	[+] Todas las Fechas	[+] Todos los Sectores	6

Figura 4.287: Ejemplo con las medidas en la posición tradicional

Al hacer uso del intercambio de ejes tenemos el siguiente resultado tomando los datos como elemento de pivote.

CentroSolucion					
Todos los Centros de Solución					
Centro Solucion de Controversias - Oficina Central		Centro Solucion de C			
Tiempo					
Todas las Fechas		2008		2009	
SectorCategoria					
Todos los Sectores		Todos los Sectores		Todos los Sectores	
CantCaso	5.789	4.085	3.930	155	868

Figura 4.288: Ejemplo con las medidas intercambiando ejes



Botones de especialización: Existen tres diferentes tipos de especialización con los que se puede controlar el comportamiento de las consultas cuando el analista accede a un miembro en particular.

Icono	Modo de especialización	Descripción
	Por Miembro	En este modo las instancias de esos miembros son expandidos(múltiples instancias de un miembro será mostrado cuando se poseen dimensiones anidadas en un eje)
	Por Posición	En este modo, solo la instancia del miembro a la que se le hace clic será accedida, aun si existieran otras instancias del mismo miembro.
	Por Reemplazo	En este modo, el miembro al que es accedido, será <i>reemplazado</i> por sus miembros hijo.

A continuación un ejemplo de accesos por miembro vs acceso por posición. En ambos casos tenemos la opción de hacer un análisis por todos los años:

CentroSolucion	Tiempo	SectorCategoria	Medid
Todos los Centros de Solución			
Centro Solucion de Controversias - Oficina Central	Todas las Fechas	Todos los Sectores	Cant
Centro Solucion de Controversias - Oficina Regional San Miguel	Todas las Fechas	Todos los Sectores	Cant
Centro Solucion de Controversias - Oficina Regional Santa Ana	Todas las Fechas	Todos los Sectores	Cant
Defensoria Financiera	Todas las Fechas	Todos los Sectores	Cant
Oficina Call Center Plan de La Laguna	Todas las Fechas	Todos los Sectores	Cant
Oficina Plan de La Laguna	Todas las Fechas	Todos los Sectores	Cant
Oficina Educación	Todas las Fechas	Todos los Sectores	Cant

Figura 4.289: Ejemplo de accesos por miembro vs acceso por posición

Teniendo la opción de hacer un análisis por año (DRILL DOWN) si utilizamos el **ACCESO POR MIEMBRO** quiere decir que se accederá a todos los miembros del mismo tipo y en el mismo nivel.

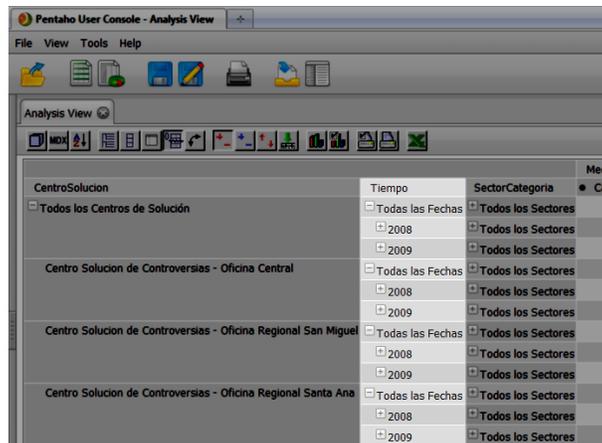


Figura 4.290: Ejemplo de accesos por miembro

Si se utiliza el acceso por posición se accede solamente al miembro al que se le hace clic.

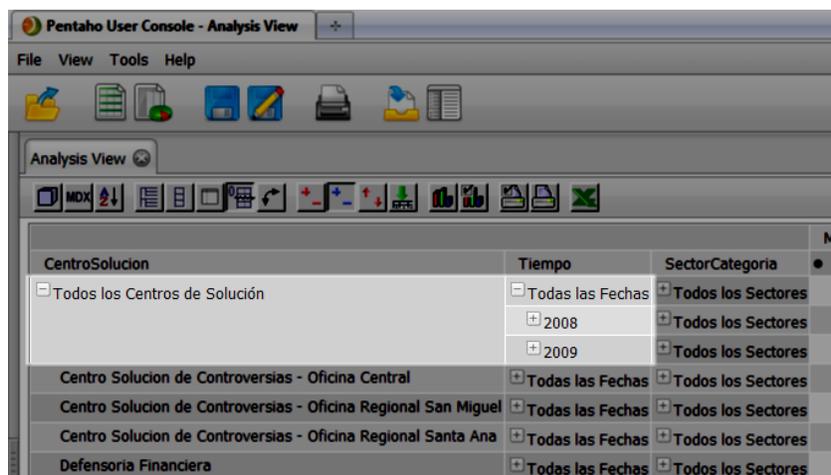


Figura 4.291: Ejemplo de accesos por miembro

En el ejemplo anterior solo al hacer clic en las órdenes en proceso para todos los años se logra ingresar a los años disponibles para dicho elemento.

El Acceso por Reemplazo se utiliza para entrar a un nivel de detalle mayor dentro de elementos comunes.

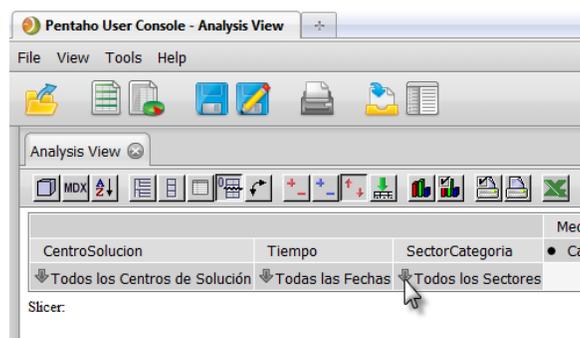


Figura 4.292: Ejemplo acceso por Reemplazo

Al entrar en el detalle de Todos los sectores podemos ver los datos disponibles por sector

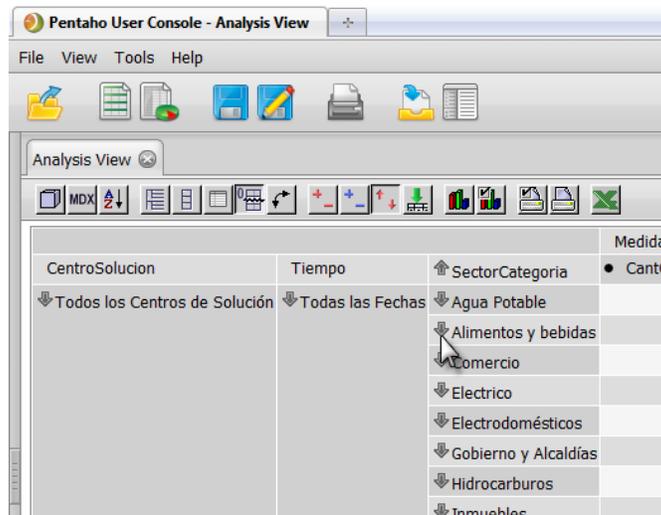


Figura 4.293: Entrando en detalle de todos los sectores

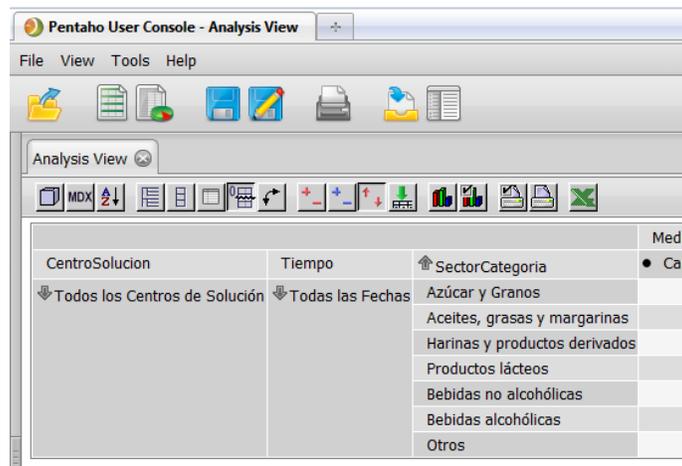


Figura 4.294: Entrando en sector alimentos y bebidas

Finalmente se llega a los datos de cada categoría por sector, por centro de solución este tipo de reporte es útil para visualizar datos relevantes dentro de los grupos que se poseen.



**Mostrar datos de Origen:** Una función muy útil para entrar en detalle de los datos que se visualizan, aquí podemos ver las relaciones de origen destino que posee el reporte con esto garantizamos la integridad de los datos y de la información que se maneja.

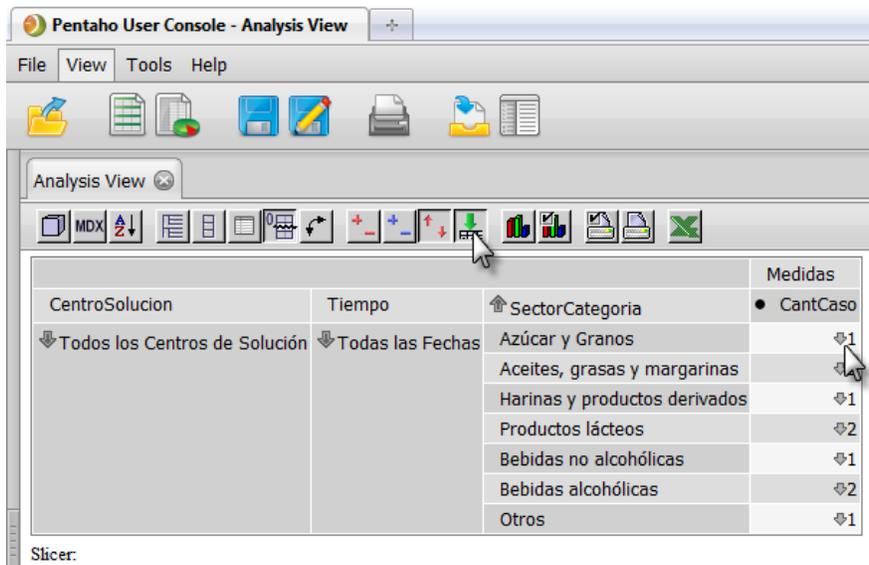


Figura 4.295: Icono para mostrar datos de origen

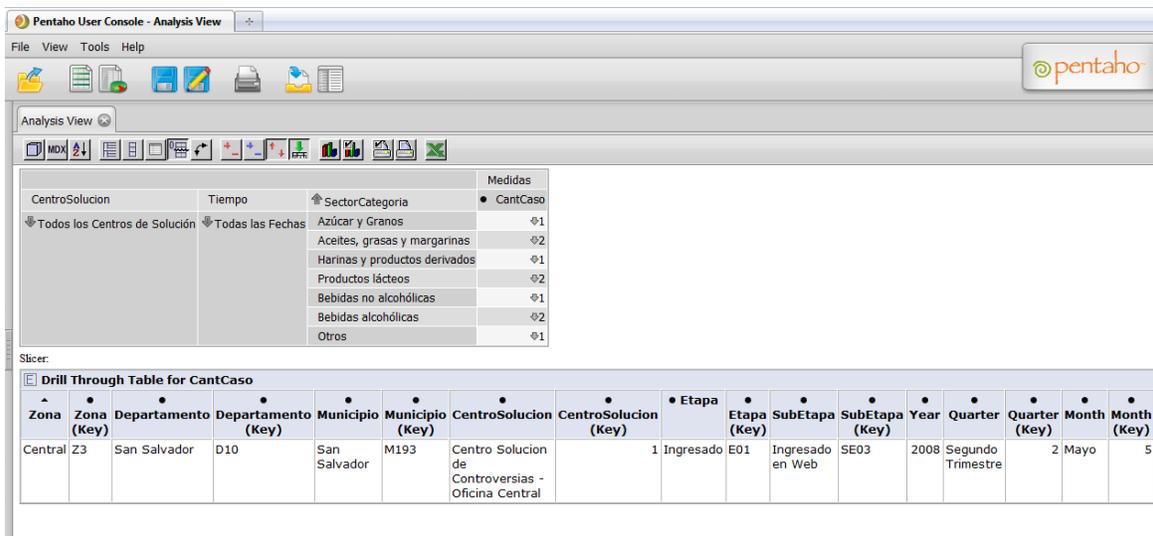


Figura 4.296: Información de datos de origen

Con la tabla de detalle se pueden visualizar elementos de interés que dan valor agregado a los reportes que se elaboran con la vista de análisis.



**Mostrar Grafico/Configurar Grafico:** son 2 opciones que están relacionadas ya que se debe configurar primero el grafico para luego visualizarlo como se desee.

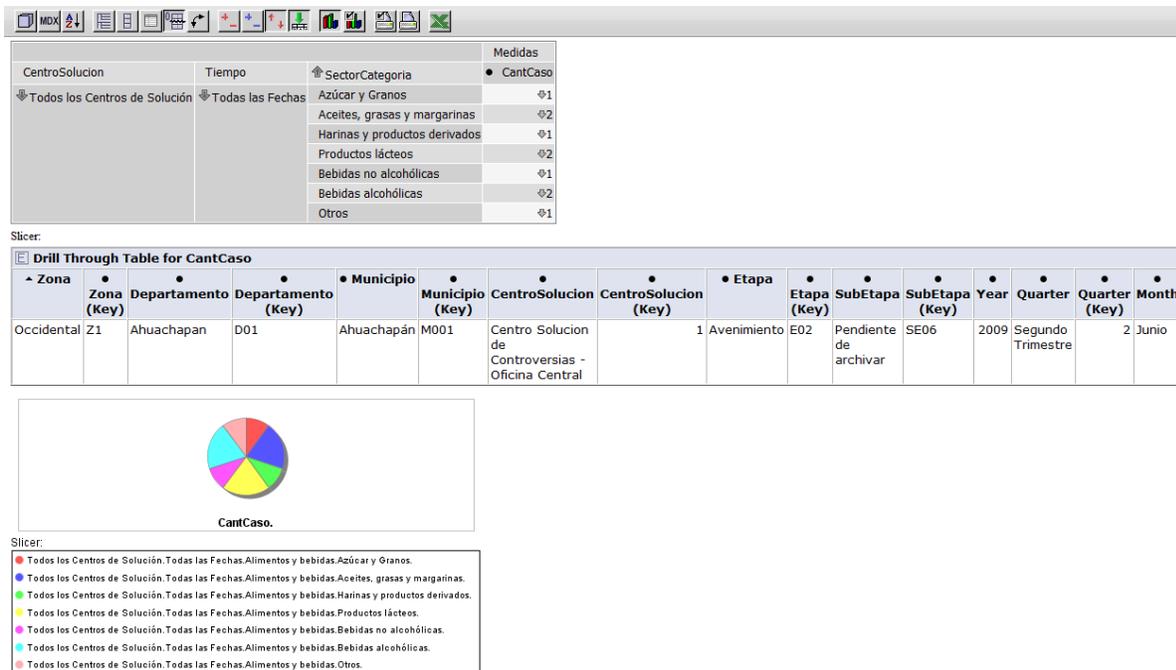


Figura 4.297: Mostrando gráficos generados

**Configurar Grafico:** se utiliza para configurar las opciones de grafico que se mostrara, el tipo de grafico, titulo, tipo de letra que se mostrara, los encabezados de los ejes y si se puede hacer DRILL TROUGH esta opción es similar a entrar en detalle que se mostro anteriormente solo que a diferencia aquí se entra en detalle desde la grafica, lo que facilita el análisis de los datos.

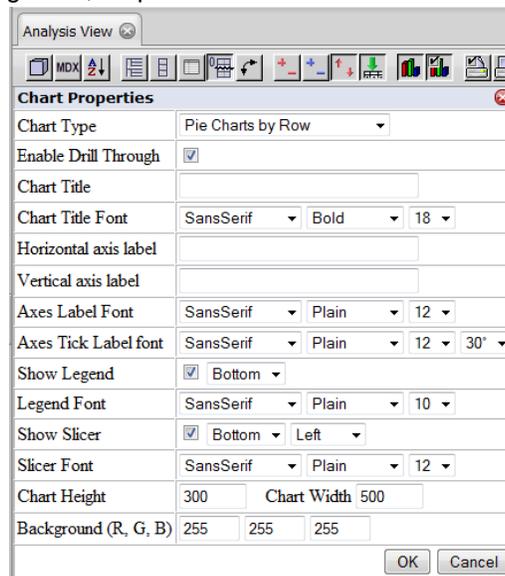


Figura 4.298: Configuración de Gráficos

Se pueden seleccionar diferentes tipos de gráficos dependiendo de la vista que se desea generar.



Figura 4.299: Tipos de gráficos

Entre ellos grafico de barra, de pastel, linear, y de diversas posiciones horizontales o verticales, en todos se puede activar la opción de navegar para ver los datos de origen.



**Configurar Impresión/Imprimir:**

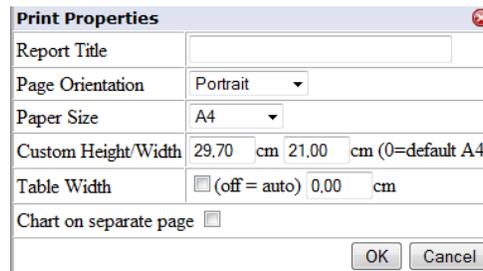


Figura 4.300: Configuración para impresión de gráficos

Con la opción de configurar impresor se puede determinar elementos como el titulo de reporte, la orientación del papel, tamaño de papel, ancho de tabla y los gráficos separados en páginas. El icono de impresión se utiliza para exportar a PDF con las configuraciones antes seleccionadas.

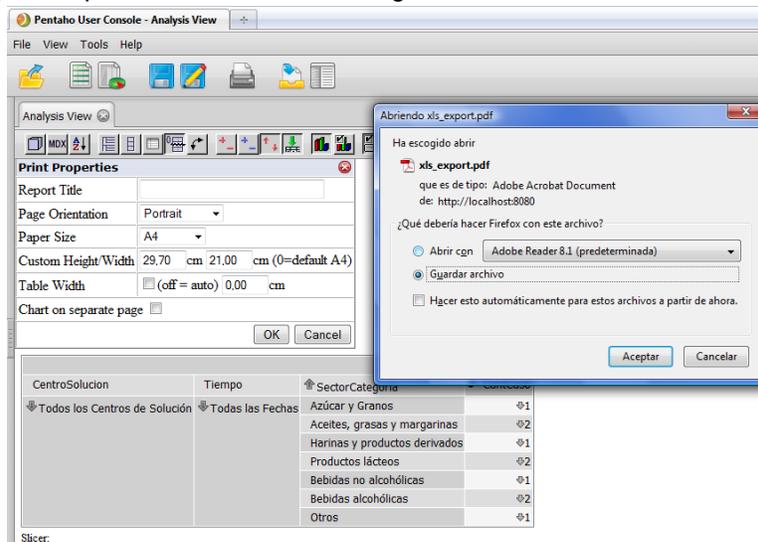


Figura 4.301: Configuración para exportar a pdf



Exportar a EXCEL: con esta opción se logra enviar los datos que se poseen en la ventana actual para que puedan ser manipulados en hojas de cálculo, esta opción es muy útil para las empresas ya que automatiza el proceso de exportación de los datos a una herramienta conocida por los usuarios.

**EJEMPLO DE ANALYSIS VIEW.**

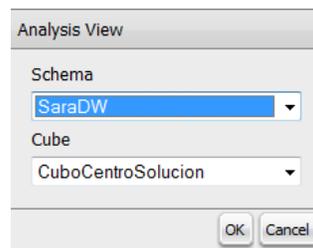
Visualizaremos para un ejemplo sencillo los casos de enero 2009 recibidos en oficinas centrales para el servicio eléctrico en específico para categorías de suministro de energía eléctrica.

En la pantalla principal seleccionamos NEW ANALYSIS VIEW



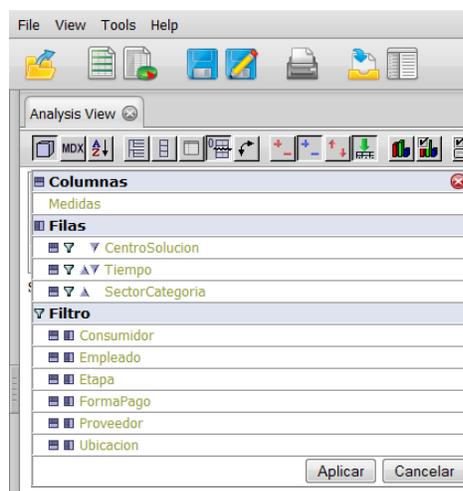
**Figura 4.302: Icono para nuevo análisis**

Luego seleccionamos el cubo con el que deseamos trabajar para el caso el CentrodeSolucion



**Figura 4.303: Elegir a que esquema y cubo se va a configurar**

En la pantalla de bienvenida seleccionamos el cubo OLAP para realizar el filtrado de los datos y en las filas dejamos solo CentroSolucion, Tiempo, SectorCategoria



**Figura 4.304: Elegir a medidas disponibles para el cubo**

Finalmente podemos hacer una navegación por Centro de Solución para todas las fechas en 2009, Primer trimestre, enero, Todos los sectores, Eléctrico y finalmente llegamos a suministro de energía, como se muestra a continuación:

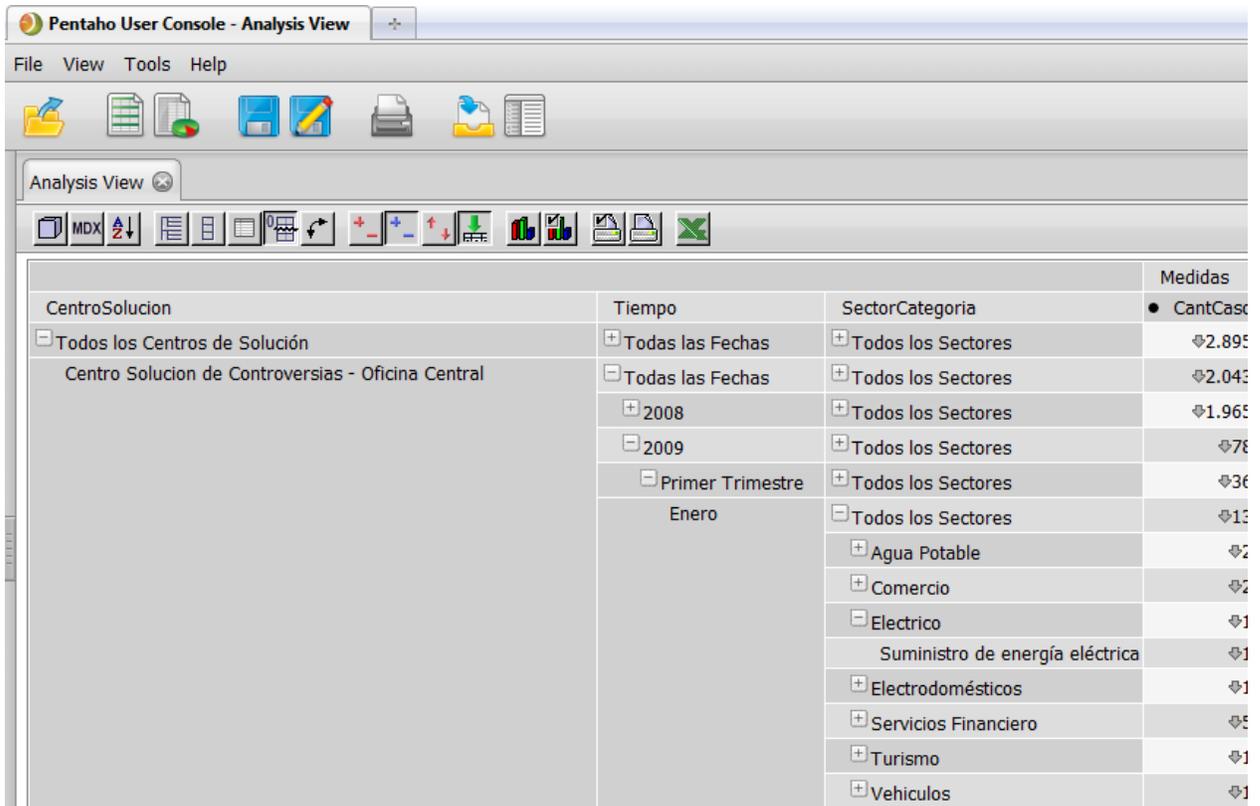


Figura 4.305: Explorando los sectores y fechas

En las medidas se puede agregar las cantidades de casos el monto recuperado, valor del producto y cantidad pagada, depende del análisis que se desea realizar.

# Capítulo V



## 5.1. ELEGIR LA ARQUITECTURA

La arquitectura necesaria para el óptimo desempeño del Data Warehouse es un punto clave que debe definirse a fin de evitar problemas cuando el sistema se encuentre totalmente en funcionamiento. Es por ello que se ha tratado de brindar como sugerencia las características del Hardware, Software necesarios para la implantación.

### 5.1.1. Recurso Tecnológico para la Implementación

La organización dispone de un completo recurso tecnológico que permita la implementación de dicho proyecto, dicha institución puso a disponibilidad toda la arquitectura cliente servidor que se describe a continuación.

#### 5.1.1.1. Hardware

El hardware para llevar a cabo la instalación y configuración del sistema se ha determinado que debe de cumplir con gran exigencia tanto por parte del servidor como la para los usuarios. Es por ello que se definen las siguientes características, estas características son extraídas en base a la investigación preliminar que se realizó en el anteproyecto.

Elemento	Servidor	Usuario 1	Usuario 2
RAM	1 GB	512 MB	1.5 GB
Discos duros	175 GB	200 GB	200 GB
Procesador	Xeon 3.17 Ghz	Pentium IV 3.06 Ghz	Pentium IV 3.06 Ghz
Tarjetas de RED	10/100	10/100	10/100

Tabla 5.1: Descripción de HW para la implementación

#### 5.1.1.2. Software

El software disponible para la instalación, configuración y administración del sistema propuesto se encuentra fundamentado en la investigación preliminar y se puede constatar en el anteproyecto.

Elemento	Servidor	Usuario 1	Usuario 2
Sistemas Operativos	Windows 2003 server	Windows Xp	Windows Xp
Bases de datos	SQL Server 2005 Enterprise		SQL Server 2005 express
Antivirus	Mcafee VirusScan 8.7	Mcafee VisusScan 8.7	Mcafee VisusScan 8.7
herramientas de desarrollo	Reporting Services		Visual Studio .net 2005 Crystal Reports 10
Herramientas de Ofimática		Microsoft office 2003	Microsoft office 2003
Otras Herramientas		Navegadores web:Mozilla firefox, Internet Explorer 6	Navegadores web:Mozilla firefox, Internet Explorer 6

Tabla 5.2: Descripción del SW Actual

El motivo de mostrar dos tipos de usuarios es debido a que se encuentran dos tipos de configuraciones en las maquinas distribuidas en la organización.

### 5.1.2. Recurso Humano

Para la correcta implementación del sistema es necesario contar con el personal necesario, que cumpla con los requisitos en conocimiento para la administración y operación del sistema. Por lo anterior, se ha definido 3 tipos de perfiles.

#### 5.1.2.1. Perfil de Analista de Negocio

##### ***Requisitos Personales:***

- Capacidad para trabajar en equipo.
- Acostumbrado a trabajar bajo presión.
- Altos valores morales
- Capacidad analítica y de resolución de problemas.
- Compromiso con la visión y valores de la compañía
- Iniciativa para enfrentar situaciones desconocidas

##### ***Requisitos Técnicos:***

##### ***Área de desarrollo y administración de Datawarehouse en entorno Microsoft/SQLServer 3 años de experiencia deseable.***

- Experto en Integration Services 2005 (2008 deseable)
- Analysis Services 2005 (2008 deseable)
- Desarrollador de cubos de Analysis Services 2005 (2008 deseable)
- Implementación de cubos, dimensiones, medidas, campos calculados, KPIs, etc.
- Administración de Integration Services:
- Desarrollo de consultas MDX: Soporte a los desarrolladores. Implementación de consultas complejas.
- Generación y administración de procesos de extracción y transformación de datos utilizando Integration Services.

**Área de desarrollo en Herramientas de Modelado de Negocio PENTAHO BI Server****Conocimientos deseables:**

- Pentaho Bi-Server 2
- MetaData Editor
- SchemaWorkbench
- Configuración de Servidor WEB, conocimiento de servicios basados en Apache Web Server.
- Hibernate
- Spring
- Administración de aplicaciones J2EE
- MySQL

**5.1.2.2. Perfil del Capacitador de Personal****Requisitos Personales**

- Capacidad para la toma de decisiones oportunas
- Maneja estrategias para el trabajo en equipo
- Posee habilidades comunicativas que le facilitan interactuar en diferentes contextos
- Dominio de estrategias e instrumentos para evaluar los aprendizajes
- Dominio de estrategias para la investigación e innovación educativa

**Conocimientos Técnicos**

El Capacitador de personal debe de conocer y dominar la Herramienta de uso final. PENTAHO BI server 2, dominando las siguientes áreas:

**- Entorno general de la herramienta:**

- Comprender la arquitectura y las herramientas de Pentaho Bi server así como las herramientas disponibles para generación de reportes
- Construir reportes de distintas fuentes de datos.
- Realizar reportes visibles haciendo uso de la herramienta.

**- Diseño de los Reportes:**

- Incluir y configurar los gráficos en reportes.
- Utilizar técnicas de agrupamiento, dependiendo del área al que se le está dando seguimiento.
- Generación de gráficos complejos.
- Uso de parámetros de cálculo, funciones y expresiones en los reportes.

### 5.1.2.3. Perfil del Administrador de Aplicaciones

**Requisitos Personales:**

- Capacidad para trabajar en equipo.
- Capacidad para la toma de decisiones oportunas
- Iniciativa para enfrentar situaciones desconocidas

**Requisitos Técnicos:****Conocimiento deseable (Tiempo de experiencia 1 año en el área de soporte técnico):**

El Administrador de aplicaciones brinda soporte técnico en el área de configuración de la herramienta, en el caso que se genere alguna falla en el desarrollo de las capacitaciones.

**- Instalación de herramientas básicas en terminales remotas:**

- Instalación y configuración de JDK en terminales.
- Configuración y mantenimiento de servidores web (APACHE).
- Uso de ADMINISTRATION CONSOLE de Pentaho para refrescar repositorios, variables.

Es necesario que el administrador de aplicaciones tenga conocimientos de la herramienta de modelado de negocio, así como el uso de la herramienta final, con el fin de brindar un soporte integral y pueda brindar soluciones prácticas y efectivas en el caso que se presente alguna falla.

### 5.1.3. Descripción de la Arquitectura del Sistema

La arquitectura del sistema se encuentra compuesta por los elementos presentados en la figura 5.1 en dicha figura se muestran los enlaces de comunicación que tienen entre estos. Por el tipo de proyecto se tiene una arquitectura similar a la llamada de tres capas, con un agregado que es la interacción con la base de datos transaccional y el "Data Warehouse". Esta relación se mantiene ya que los datos son cargados al "Data Warehouse" a partir de los datos generadas en los sistemas transaccionales.

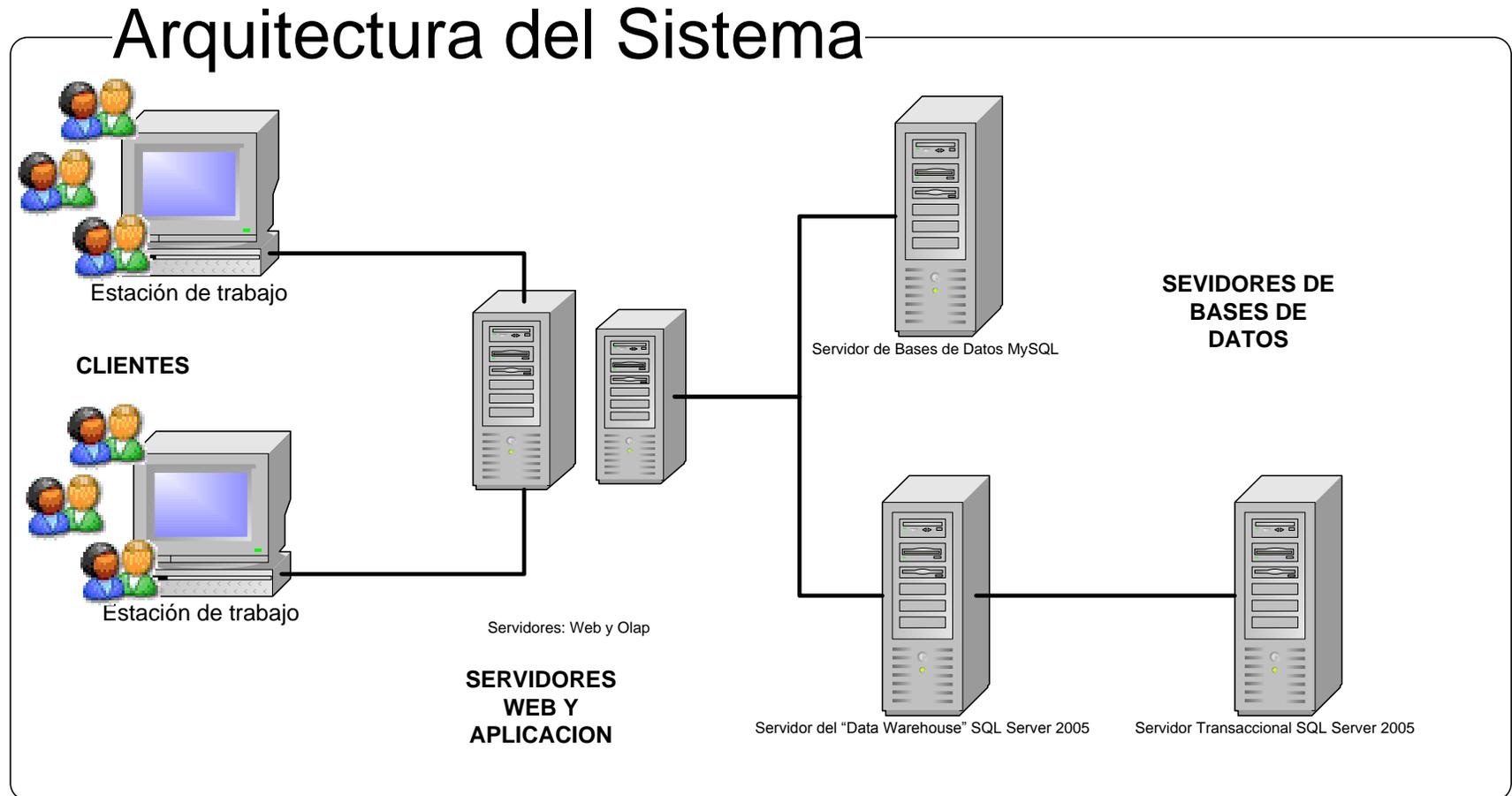
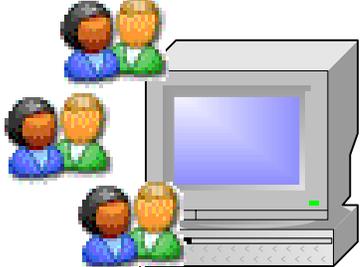
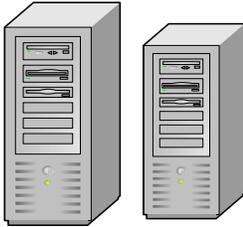


Figura 5.1: Enlaces de comunicación de los elementos del sistema

**5.1.3.1. Descripción de las partes de la Arquitectura del Sistema**

A continuación se describen cada uno de los elementos que interactúan en la arquitectura del sistema.

 <p style="text-align: center;">Estación de trabajo</p>	<p>Los clientes estarán conectados a través de estaciones de trabajo, que ya se encuentran distribuidas en la organización. El acceso al “Data Warehouse” se realizará por medio de un navegador web, el cual permitirá ingresar a una aplicación web que tiene por objetivo facilitar el análisis de la información disponible en el cubo.</p>
 <p style="text-align: center;">Servidores: Web y Olap</p>	<p>El servidor web y OLAP, son de los elementos más importantes ya que será el negociador entre la data y las aplicaciones en las que permitirán el análisis. Este servidor se encuentra compuesto por:</p> <ul style="list-style-type: none"> <li>• Un servidor web, Apache Tomcat, tiene soporte a aplicaciones java y funciona en cualquier sistema operativo que disponga de la maquina virtual Java.</li> <li>• Un servidor OLAP, Mondrian: siendo este un servidor OLAP escrito en Java, que permite analizar grandes cantidades de datos almacenados en las bases de datos de una forma interactiva, sin necesidad de escribir las sentencia SQL que serian necesarias para ello.</li> </ul> <p>Además contienen un conjunto de aplicaciones como:</p> <ul style="list-style-type: none"> <li>• PENTAHO BI server la cual es una aplicación más primordial ya que brinda la interfaz grafica al usuario para que realice sus análisis y reportes</li> <li>• Schema workbench que se utiliza para realizar el diseño y mantenimiento del cubo de información.</li> <li>• Consola de administración de PENTAHO que se utiliza para la gestión total de las configuraciones de PENTAHO BI server.</li> </ul>

 <p>Servidor de Bases de Datos MySQL</p>	<p>El servidor de bases de datos de MySQL se utiliza plenamente para manejar aspectos de seguridad y configuración de PENTAHO. Ya que PENTAHO y MySQL son software libre, y debido a la facilidad que brinda en la configuración de la seguridad al mezclar estos dos elementos se decidió incluir en la arquitectura esta base de datos.</p>
 <p>Servidor del "Data Warehouse" SQL Server 2005</p>	<p>Este servidor es el encargado de almacenar los datos que se cargan de los sistemas transaccionales, por lo que la tiene una gran interacción con el servidor web y OLAP.</p>
 <p>Servidor Transaccional SQL Server 2005</p>	<p>Este servidor se encarga de almacenar todas las transacciones generadas en la organización. Por lo que se hace necesario realizar un proceso ETL, para transportar los datos de forma que puedan ser aceptados en el servidor del "Data Warehouse".</p>

**Tabla 5.3: Descripción de los elementos de la arquitectura del sistema**

## **5.1.4. Herramientas para el Desarrollo del Proyecto**

### **5.1.4.1. SQL Server 2005 Estándar**

Una de las principales herramientas con las que se cuenta para el desarrollo de un Data Warehouse para el proceso de denuncias de la Defensoría del Consumidor es SQL SERVER 2005 Estándar, el propósito del uso de dicha herramienta es brindar una solución efectiva que no genere costos adicionales en la adquisición de nuevas herramientas considerando el uso de las herramientas que se poseen actualmente. La Defensoría del Consumidor cuenta con el Gestor de Bases de datos SQL Server 2005 para administrar el flujo de información generada por sus aplicaciones transaccionales. La información almacenada por el sistema SARA es el insumo principal del Data Warehouse el cual se considero generarlo haciendo uso de él mismo gestor de bases de datos. SQL Server Estandar Edition, brinda características de alta disponibilidad y de Business Intelligence lo cual facilita explotar las capacidades básicas que se necesita para llevar un control mejorado de las transacciones dentro de la Defensoría.

### **5.1.4.2. SQL Management Studio**

Es una herramienta gráfica de administración gratuita y fácil de usar para SQL Server 2005. Viene integrado en el paquete de SQL Server y es utilizado para administrar, configurar y personalizar los componentes que posee SQL Server. La herramienta tanto el editor de scripts como la parte grafica que trabaja con objetos y las características propias del servidor. Un elemento importante del SQL Management Studio es el Object Explorer que es un navegador que permite seleccionar y realizar determinadas acciones sobre cualquier objeto dentro del Servidor. En el desarrollo del proyecto es utilizado para navegar tanto en los elementos que posee la base de datos transaccional, como los elementos que posee el Data Warehouse.

### **5.1.4.3. SQL Integration Services 2005**

SQL Server 2005 Estándar posee herramientas propias para la administración de procesos orientadas a la inteligencia de negocios (Business Intelligence), una de las herramientas utilizadas en el desarrollo de este proyecto en particular es el Integration Services, esta herramienta es capaz de auxiliar en el proceso de extracción, transformación y carga, (ETL por sus siglas en ingles), SSIS (SQL Integration Services) logra cumplir dicho proceso dependiendo de la ubicación en la que se encuentran almacenados los datos fuente. El proceso de ETL es básicamente la unificación de diversas fuentes de información para lograr estandarizar la información y poder “hablar de un solo lenguaje”. Es muy importante tener un estándar de datos que unifique el conjunto de datos fuentes, con el fin único de poder manipular la información teniendo certeza que se están considerando las mismas unidades, medidas, métricas. Los archivos fuentes pueden ser datos almacenados de bases de datos propias de SQLServer u otro gestor de bases de datos, hojas de cálculo u otro tipo de dato fuente, en el caso del desarrollo del Data Warehouse para la DC, los archivos fuente son los datos almacenados en la base de datos transaccional del sistema SARA.

#### 5.1.4.4. Pentaho Bi-Server

El Pentaho Bi-Server (BI Suite) es un servidor orientado a soluciones de BI (Business Intelligence) el cual posee junto con él un conjunto de aplicaciones que son fundamentales para el desarrollo del proyecto, **Apache Tomcat** como servidor de aplicaciones WEB que es el que provee la capa de presentación de la BI Suite, **Hibernate** que es un paquete elaborado en Java que facilita la comunicación entre la base de datos relacional y Pentaho, **Jetty** es un servidor basado en tecnología WEB que proporciona acceso a la consola de Administración **Quartz** el cual es utilizado para las tareas de reportes que utiliza Pentaho BI. A simple vista Pentaho es el que presenta la capa de aplicación del Data Warehouse, la ventaja de ser una herramienta de distribución libre proporciona un valor de peso en el desarrollo de este proyecto, generalmente Pentaho BI se puede configurar con SQL SERVER para realizar la conexión que brinde control y referencia administrativa a la BI Suite sin embargo por razones de administración y considerando la carga de datos y funcionamiento independiente se ha decidido utilizar el servidor por defecto con el que funciona Pentaho, el cual es MySQL. Además Pentaho BI, incluye un servidor OLAP, llamado Mondrian que permite analizar grandes cantidades de datos almacenados en las bases de datos de una forma interactiva, sin necesidad de escribir las sentencias que serían necesarias para ello.

#### 5.1.4.5. MySQL

MySQL Server es un gestor de bases de datos que posee la característica especial de ser de distribución libre, es un motor que soporta multiusuario y multi hilo. En la solución la función principal de MySQL es brindar el control administrativo de las herramientas que posee pentaho BI Server, configuraciones importantes de Quartz, el Datasource, elementos de migración y datos de prueba es lo que almacena el gestor de MySQL, es importante mencionar que Pentaho fue desarrollado para funcionar con MySQL 5 el cual proporciona grandes capacidades para enfrentar los requerimientos que exige un BI escalable<sup>47</sup>. Pentaho es una herramienta de BI Open-Source el cual posee una alianza especial con MySQL es por esa razón que todos los módulos desarrollados para Pentaho son compatibles para ejecutarse sobre MySQL como servidor de aplicación.

#### 5.1.4.6. MySQL Administrator

Es una herramienta que permite administrar de manera grafica Servidores de Bases de Datos MySQL. Es un software multiplataforma, que cuenta con un entorno grafico sencillo y guiado para los administradores de MySQL que permite realizar tareas como:

- La configuración de las opciones de inicio de los servidores.
- Inicio y detención de servidores.
- Monitorización de conexiones al servidor.
- Administración de usuarios.
- Monitorización del estado del servidor, incluyendo estadísticas de uso.
- Visualización de los logs de servidor.
- Gestión de copias de seguridad y recuperaciones.

---

<sup>47</sup> [http://dev.mysql.com/tech-resources/articles/mysql\\_5.0\\_pentaho.pdf](http://dev.mysql.com/tech-resources/articles/mysql_5.0_pentaho.pdf)

#### 5.1.4.7. Schema Workbench 3.1

Componente que forma parte de Pentaho que permite crear y probar esquemas de cubos OLAP mediante una interfaz gráfica. Los archivos de los esquemas son metadata<sup>48</sup> en formato XML los cuales son creados con una estructura específica por el motor de Mondrian; Dichos modelos pueden ser considerados estructuras tipo Cubos los cuales contienen las tablas de Hechos y dimensiones que se encuentran en el Sistema Administrador de Bases de Datos Relacionales que se posee<sup>49</sup>.

#### 5.1.4.8. Metadata-Editor

Es una herramienta que permite generar el modelado del Negocio, permite crear de manera sencilla el modelo de metadata de un origen de datos relacional con características complejas.

#### 5.1.4.9. Pentaho Design Studio

Pentaho Design Studio es una colección de editores, visores y módulos de administración integrados a una sola aplicación que provee un entorno gráfico para la creación y prueba de documentos de secuencias de acción (Documentos XML acerca del rendimiento de las actividades de PENTAHO BI), y reportes JFree. Design Studio está basado en el Framework Eclipse, software que gracias a su robustez permite ejecutar herramientas de desarrollo de manera independiente pero integrada a la vez permitiendo reutilizar componentes existentes reduciendo considerablemente el tiempo de desarrollo de otras aplicaciones.

#### 5.1.4.10. Pentaho Report Design Wizard

Parte de Pentaho Design Studio el cual permite crear reportes, si los usuarios convencionales no han tenido la oportunidad de crear reportes propios y personalizados, estarán muy agradecidos de tener esta herramienta a la mano, y en caso que el usuario sea de un nivel más experimentado también se sentirá agradecido por el tiempo que se ahorra para crear sus reportes gracias a esta herramienta. Básicamente el Reports Designer define el esquema de distribución que contendrá el reporte, especificando el encabezado y el pie de página, así como el cuerpo del reporte y la sección de ítems o elementos que contiene las columnas que pertenecen a determinado grupo, es donde aparecerán los valores deseados de la consulta.

---

<sup>48</sup> <http://es.wikipedia.org/wiki/Metadata>

<sup>49</sup> <http://mondrian.pentaho.org/documentation/workbench.php>

## 5.2. METODOLOGÍA DE LA IMPLANTACIÓN

La definición de la metodología es trascendental para lograr una implantación del sistema, tomando en cuenta las dos alternativas más conocidas y utilizadas para este fin: la implantación directa e implantación en paralelo. Para ello determinamos las ventajas y desventajas para cada una de ellas

<b>METODOLOGÍA</b>	<b>VENTAJAS</b>	<b>DESVENTAJAS</b>
Implantación directa	Algunos recursos no son compartidos	Ausencia de respaldo en caso de que falle el nuevo sistema
	Reduce el tiempo para llevar a cabo la implantación	Aumenta las posibilidades de resistencia al cambio
	Los costos de implantación se reducen	Reduce la promoción del nuevo sistema
Implantación en Paralelo	Se cuenta con un respaldo si el sistema propuesto falla	Algunos recursos pueden ser compartidos
	Permite una mejor adaptación al nuevo sistema	El tiempo de implantación tiende a aumentar.
	Promueve el nuevo sistema a través del actual	Los costos pueden aumentar considerablemente.

**Tabla 5.4: Ventajas y desventajas de las metodologías de implantación**

Al evaluar ventajas y desventajas de cada una de las metodologías expuesta, se puede determinar que dependiendo de la situación en que se encuentre cada sistema a implantar así se tendrá una elección de una opción sobre otra. No existe una mejor que otra sino que estas dependen directamente del estado en que se encuentre la organización, considerando principalmente la disponibilidad de recursos. Para nuestro caso por ser un sistema que servirá principalmente para el análisis de datos y apoyo a la toma de decisiones, es de vital importancia que mientras se encuentre en el periodo de implantación se pueda contar un sistema contra el que se comparen los datos y se puedan constatar la validez de la información presentada.

Debido a eso es que se concluye que la metodología más conveniente para este proyecto es la implantación en paralelo, ya que permitirán que los usuarios puedan adaptarse al sistema, dando la opción de poder comparar la información del sistema actual contra este nuevo sistema. Así mismo en caso de que falle el nuevo sistema, no se afectaran las operaciones de la organización.

Habiendo seleccionado la metodología de implantación, es necesario definir un punto de convergencia en que el sistema actual y el sistema propuesto trabajaran de forma paralela para que se pueda verificar la funcionalidad del nuevo sistema y lograr la aceptación por parte de los usuarios. Para lograr esta evaluación, se recomienda que el periodo sea de tres a seis meses aunque esto dependerá del coordinador del proyecto.

### 5.3. PLANIFICACIÓN DE ACTIVIDADES

El proceso de implantación del sistema desarrollado seguirá el siguiente proceso administrativo, basado en macro actividades y desglosadas según cantidad de procesos a realizar.

#### 5.3.1. Planeación de Implantación

<b>Objetivo</b>	
Definir los elementos necesarios que compone el ambiente en que se ejecutara la implementación del sistema desarrollado.	
<b>Actividades</b>	
<ul style="list-style-type: none"> <li>• Definir estrategias y metas. Como inicio de ejecución del plan de implantación y puesta en marcha del sistema desarrollado.</li> <li>• Gestión de los recursos disponibles y limitantes. Identificar cambios sorpresivos a causa de recursos no disponibles o limitados, previniendo la readecuación del presupuesto y planificación cronológica.</li> <li>• Preparación del ambiente organizacional. Anunciando fechas planificadas para cada una de las direcciones que se verán involucradas y la disponibilidad del recurso humano, preparando el ambiente organizacional para la puesta en marcha del sistema desarrollado.</li> <li>• Envié de memo a las direcciones involucradas.</li> <li>• Reunión con Directores para brindar una descripción del proyecto y del plan de implantación.</li> </ul>	
Tiempo	3 días
Horas Hombre	41.33
Costo	\$406.67

#### 5.3.2. Organización Implantación

<b>Objetivo</b>	
Organizar y planificar el recurso humano necesario para la implementación del sistema desarrollado.	
<b>Actividades</b>	
<ul style="list-style-type: none"> <li>• Definición del equipo ejecutor. Establecer el recurso humano encargado de ejecutar el plan de implantación, estableciendo la estructura organizativa definiendo las líneas de mando y los medios de comunicación entre los participantes del proceso.</li> <li>• Estudio de perfiles y funciones del equipo ejecutor. Revisión de cada uno de los perfiles que integraran el equipo ejecutor del plan de implantación, verificando con concordancia de funciones de acuerdo al perfil.</li> <li>• Definición de responsabilidades. Asignar actividades a cada uno de los miembros del equipo ejecutor y su respectiva responsabilidad dentro de esta.</li> </ul>	
Tiempo	12.5 días
Horas Hombre	100 horas
Costo	\$654.67

### 5.3.3. Ejecución de la Implantación

<p><b>Objetivo</b></p> <p>Ejecutar cada una de las actividades necesarias para la implantación del proyecto desarrollado bajo el título “Desarrollo de un Data Warehouse para el proceso de denuncias de la Defensoría del Consumidor”.</p>
<p><b>Actividades</b></p> <p><b>Instalación y configuración de equipo informático.</b> Los equipos a utilizar estarán bajo la estructura cliente - servidor, para lo cual:</p> <ul style="list-style-type: none"> <li>• Servidor: El equipo informático que se utilizara como servidor para el aplicativo esta detallado en el apartado “Recurso Tecnológico para la Implementación”. Este equipo deberá estar incorporado a la red local de trabajo de la DC, recomendando la verificación de los estándares recomendados por IEEE para la conectividad de la red. La configuración del Servidor esta detallada en:             <ul style="list-style-type: none"> <li>○ Manual de Instalación y Configuración. Dicta la guía para la instalación y configuración de la aplicación.</li> <li>○ Manual Técnico. Guía en el uso del software que compone la aplicación, para la modificaciones pertinentes que sean necesarias y puedan surgir durante o posterior la implantación del sistema.</li> <li>○ Cliente: Todo equipo informático a utilizar como cliente se recomienda que se configure de acuerdo a las recomendaciones realizadas en el Manual de Usuario.</li> </ul> </li> </ul> <p><b>Ejecución y control de ETL inicial.</b> En coordinación con el equipo de implantación, desarrolladores y DBA de la DC, controlar la correcta ejecución y carga de datos del repositorio origen hacia el DW, vigilando la consistencia e integridad de datos:</p> <ul style="list-style-type: none"> <li>• Revisión de LOGs escritos posterior a la terminación del proceso ETL en busca de posibles errores que surjan en el proceso.</li> <li>• Cada tabla de hecho posea la misma cantidad de registros extraídos del repositorio origen.</li> <li>• Cada dimensión posea la misma cantidad de registros de catálogos, en concordancia a la lógica del negocio y los requerimientos del proyecto.</li> </ul> <p><b>Prueba de aplicativo (control de calidad).</b> Establecimiento de la estructura que organiza evaluaciones, inspecciones, auditorias y revisiones que aseguren que se cumpla las responsabilidades asignadas, se utilicen eficientemente los recursos y se logre el cumplimiento de los objetivos del aplicativo. Este equipo deberá mantener bajo control un proceso y eliminar las causas de los defectos. Se recomiendan la realización de los siguientes puntos:</p> <ul style="list-style-type: none"> <li>• Análisis del proyecto.</li> <li>• Identificar escenarios de prueba.</li> <li>• Verificación de estándares y requerimientos.</li> </ul>

- Ejecución y documentación de pruebas (casos de pruebas).
- Ingresar, clasificar los errores y/o defectos (bugs) encontrados a la bitácora de errores.
- Revalidar las correcciones.
- “Regression Test”.
- Realizar reportes estadísticos al final del proyecto.

**Preparación de plan de capacitación.** Preparación del personal que estará asignado al proceso de capacitación y adiestramiento. El temario de capacitación está definido en el manual de usuario, temas:

- Objetivo del proyecto.
- Funcionalidades de la herramienta.
- Configuración básica del equipo.
- Reconocimiento del ambiente de trabajo:
  - Microsoft Windows XP (todas las versiones).
  - Microsoft Windows Vista (todas las versiones).
  - Mac OS X.
- Navegación del sistema Modulo “Centro de Solución de Controversias”:
  - Reconocimiento del ambiente dentro de la aplicación BI.
  - Creación de reportes.
  - Creación de vista de análisis.
- Navegación del sistema Modulo “Tribunal Sancionador”:
  - Reconocimiento del ambiente dentro de la aplicación BI.
  - Creación de reportes.
  - Creación de vista de análisis.
- Navegación del sistema Modulo “Call Center”:
  - Reconocimiento del ambiente dentro de la aplicación BI.
  - Creación de reportes.
  - Creación de vista de análisis.
- Navegación del sistema Modulo “Verificación y Vigilancia”:
  - Reconocimiento del ambiente dentro de la aplicación BI.
  - Creación de reportes.
  - Creación de vista de análisis.

- Navegación del sistema Modulo “Anfitrión”:
  - Reconocimiento del ambiente dentro de la aplicación BI.
  - Creación de reportes.
  - Creación de vista de análisis.
- Vías de Soporte.

**Presentación y Capacitación.** A continuación se definen las actividades a realizar con los usuarios:

- Notificación de usuarios a capacitar:
  - Dirección de Estudios, Planeamiento y Desarrollo;  
Leily Melany Mendoza Valladares (Directora de Estudios).  
Ada Aurora Marquina Arce (Analista de Planeamiento).  
Manuel Ortiz Pedroza (Técnico de Estadísticas).
  - Dirección del Centro de Solución de Controversias;  
Sandra Evelyn Liborio (Directora del Centro de Solución de Controversias).  
María del Transito Elena de Santos (Técnico Conciliador) Gerencia de Sistemas de Información.
  - Director de Call Center.
  - Director de Tribunal Sancionador.
  - Director de Verificación y Vigilancia
  - Directores de Centros de Soluciones.  
Centro Solución de Controversias - Oficina Central.  
Centro Solución de Controversias - Oficina Regional San Miguel.  
Centro Solución de Controversias - Oficina Regional Santa Ana.
- Preparar Lugar de capacitación para usuarios. La sala de capacitación deberá ser lo suficientemente amplia para ubicar cada usuario. Cada usuario deberá tener a su disposición durante la capacitación:
  - Terminal por cada usuario.
  - La terminal de cada usuario deberá estar conectada a la red local de la Defensoría del Consumidor, con acceso a la aplicación desarrollada.
  - Manual de Usuario para cada usuario.
  - Libreta para anotaciones para cada usuario.
  - Terminal para el Capacitador.

<ul style="list-style-type: none"> <li>○ Proyector para demostraciones.</li> <li>• Capacitación de usuarios, dicha organización se presenta en las tabla 5.5 (“Abreviación de direcciones y jefaturas”) y tabla 5.6 (“Organización de capacitación”).</li> </ul>	
Tiempo	70.67 días
Horas Hombre	573.33 horas
Costo	\$3,141.33

Grupo	Identificador
Dirección de Estudios, Planeamiento y Desarrollo	DEPD
Director de Call Center	DCC
Director de Tribunal Sancionador	DTS
Director de Verificación y Vigilancia	DVV
Dirección del Centro de Solución de Controversias	DCSC
Centro Solución de Controversias - Oficina Central	CSC-OC
Centro Solución de Controversias - Oficina Regional San Miguel	CSC-ORM
Centro Solución de Controversias - Oficina Regional Santa Ana	CSC-ORA

**Tabla 5.5:** “Abreviación de direcciones y jefaturas”

Actividad	Grupo	Contenido	Horas
1	DEPD DCC DTS DVV DCSC	Objetivo del proyecto Funcionalidad de la herramienta Configuración básica del equipo	2
2	DEPD DCSC	Modulo “Centro de Solución de Controversias”: <ul style="list-style-type: none"> <li>• Reconocimiento del ambiente dentro de la aplicación BI.</li> <li>• Creación de reportes:                             <ul style="list-style-type: none"> <li>○ Report</li> <li>○ Ejemplo de reporte</li> </ul> </li> <li>• Creación de vista de análisis.                             <ul style="list-style-type: none"> <li>○ New Analisis View (Nueva vista de análisis)</li> <li>○ Ejemplo de vista de análisis</li> </ul> </li> </ul>	6
3	DEPD DCSC	Modulo “Centro de Solución de Controversias”: <ul style="list-style-type: none"> <li>• Hecho de análisis: Caso en Centro de Solución de Controversias.</li> <li>• Dimensiones de análisis:                             <ul style="list-style-type: none"> <li>1- Dimensión Centro de Solución</li> <li>2- Dimensión Consumidor</li> <li>3- Dimensión Empleado</li> <li>4- Dimensión Etapa</li> <li>5- Dimensión Proveedor.</li> <li>6- Dimensión Ubicación</li> </ul> </li> </ul>	12

		<p>7- Dimensión Forma de Pago                  8- Dimensión Sector Categoría                  9- Dimensión Tiempo</p>	
4	DTS	<p>Modulo "Tribunal Sancionador":</p> <ul style="list-style-type: none"> <li>• Reconocimiento del ambiente dentro de la aplicación BI.</li> <li>• Creación de reportes:                         <ul style="list-style-type: none"> <li>○ Report</li> <li>○ Ejemplo de reporte</li> </ul> </li> <li>• Creación de vista de análisis.                         <ul style="list-style-type: none"> <li>○ New Analisis View (Nueva vista de análisis)</li> <li>○ Ejemplo de vista de análisis</li> </ul> </li> </ul>	6
5	DEPD DTS	<p>Modulo "Tribunal Sancionador":</p> <ul style="list-style-type: none"> <li>• Hecho de análisis: Caso en Tribunal Sancionador.</li> <li>• Dimensiones de análisis:                         <ul style="list-style-type: none"> <li>1- Dimensión Sector Categoría</li> <li>2- Dimensión Proveedor</li> <li>3- Dimensión Ubicación</li> <li>4- Dimensión Empleado</li> <li>5- Dimensión Tiempo</li> <li>6- Dimensión Centro de Solución</li> <li>7- Dimensión Estado Tribunal Sancionador</li> <li>8- Dimensión Infracción</li> <li>9- Dimensión Proceso y Subproceso</li> </ul> </li> </ul>	10
6	DCC	<p>Modulo "Call Center":</p> <ul style="list-style-type: none"> <li>• Reconocimiento del ambiente dentro de la aplicación BI.</li> <li>• Creación de reportes:                         <ul style="list-style-type: none"> <li>○ Report</li> <li>○ Ejemplo de reporte</li> </ul> </li> <li>• Creación de vista de análisis.                         <ul style="list-style-type: none"> <li>○ New Analisis View (Nueva vista de análisis)</li> <li>○ Ejemplo de vista de análisis</li> </ul> </li> </ul>	6
7	DEPD DCC	<p>Modulo "Call Center":</p> <ul style="list-style-type: none"> <li>• Hecho de análisis: Llamada realizada</li> </ul>	12

		<ul style="list-style-type: none"> <li>• Dimensiones de análisis:                             <ol style="list-style-type: none"> <li>1- Dimensión Empleado</li> <li>2- Dimensión Proveedor</li> <li>3- Dimensión Consumidor</li> <li>4- Dimensión Tiempo</li> <li>5- Dimensión Sector Categoría</li> <li>6- Dimensión Ubicación</li> <li>7- Dimensión Sexo</li> <li>8- Dimensión Consulta</li> <li>9- Dimensión Forma de Recepción</li> <li>10- Dimensión Forma de Archivo</li> <li>11- Dimensión Anónimo</li> <li>12- Dimensión Tipo de Llamada</li> </ol> </li> </ul>	
<b>8</b>	DVV	Modulo "Verificación y Vigilancia": <ul style="list-style-type: none"> <li>• Reconocimiento del ambiente dentro de la aplicación BI.</li> <li>• Creación de reportes:                             <ul style="list-style-type: none"> <li>○ Report</li> <li>○ Ejemplo de reporte</li> </ul> </li> <li>• Creación de vista de análisis.                             <ul style="list-style-type: none"> <li>○ New Analisis View (Nueva vista de análisis)</li> <li>○ Ejemplo de vista de análisis</li> </ul> </li> </ul>	<b>6</b>
<b>9</b>	DEPD DVV	Modulo "Verificación y Vigilancia": <ul style="list-style-type: none"> <li>• Hecho de análisis: Inspección realizada</li> <li>• Dimensiones de análisis:                             <ol style="list-style-type: none"> <li>1- Dimensión Centro de Solución</li> <li>2- Dimensión Proveedor</li> <li>3- Dimensión Tiempo</li> <li>4- Dimensión Inspectores</li> <li>5- Dimensión Estado</li> <li>6- Dimensión Rubro</li> <li>7- Dimensión Errores</li> <li>8- Dimensión Artículos.</li> </ol> </li> </ul>	<b>10</b>
<b>10</b>	DEPD	Modulo "Anfitrión": <ul style="list-style-type: none"> <li>• Hecho de análisis: Visita del consumidor</li> </ul>	<b>4</b>

		<ul style="list-style-type: none"> <li>• Dimensiones de análisis:                             <ol style="list-style-type: none"> <li>1- Dimensión Tiempo</li> <li>2- Dimensión Consumidor</li> <li>3- Dimensión Contacto</li> <li>4- Dimensión Motivo de Visita</li> <li>5- Dimensión Técnico</li> </ol> </li> </ul>	
11	CSC-OC CSC-ORM CSC-ORA	Modulo “Centro de Solución de Controversias”: <ul style="list-style-type: none"> <li>• Reconocimiento del ambiente dentro de la aplicación BI.</li> <li>• Creación de reportes:                             <ul style="list-style-type: none"> <li>○ Report</li> <li>○ Ejemplo de reporte</li> </ul> </li> <li>• Creación de vista de análisis.                             <ul style="list-style-type: none"> <li>○ New Analisis View (Nueva vista de análisis)</li> <li>○ Ejemplo de vista de análisis</li> </ul> </li> </ul>	6
12	CSC-OC CSC-ORM CSC-ORA	Modulo “Centro de Solución de Controversias”: <ul style="list-style-type: none"> <li>• Hecho de análisis: Caso en Centro de Solución de Controversias.</li> <li>• Dimensiones de análisis:                             <ol style="list-style-type: none"> <li>1- Dimensión Centro de Solución</li> <li>2- Dimensión Consumidor</li> <li>3- Dimensión Empleado</li> <li>4- Dimensión Etapa</li> <li>5- Dimensión Proveedor.</li> <li>6- Dimensión Ubicación</li> <li>7- Dimensión Forma de Pago</li> <li>8- Dimensión Sector Categoría</li> <li>9- Dimensión Tiempo</li> </ol> </li> </ul>	12

**Tabla 5.6: “Organización de capacitación”**

**5.3.3.1. Calculo de las horas de capacitación:**

*Promedio de horas de iniciación: 2 horas.*

**Fuente:** TRAINING COURSE Getting Started<sup>50</sup> Sitio Oficial en el que se calcula 3 horas de consumo sin embargo consideran elementos de arquitectura básica los cuales se imparten en una segunda fase donde se explica el modelado de cada iteración, en el caso se ha propuesto la configuración básica del equipo y la funcionalidad general de la herramienta, por esa razón se considera que se puede impartir en 2 horas.

<sup>50</sup> [http://www.pentaho.com/services/training/prdt\\_session1.php](http://www.pentaho.com/services/training/prdt_session1.php)

*Promedio de horas para reconocimiento del área de trabajo y generación de reportes: 6 horas.*

**Fuente:** TRINING COURSE Designing Reports, Dinamic Reports. Sitio Oficial se calcula 6 horas, tiempo suficiente para el diseño y la generación de reportes dinámicos y reportes estáticos.

*Promedio de horas para modelado del área de negocio: 12 horas (MÁXIMO ESTIMADO).*

**Fuente:** TRINING COURSE Business Intelligence with Pentaho BI Suite. Sitio Oficial se calcula 9 horas, sin embargo teniendo una extensión de tiempo (3 horas) dado que la herramienta está desarrollada en inglés puede considerarse que lleva un poco más de tiempo adaptarse a las funciones de la misma.

### 5.3.3.2. Recurso a Utilizar

La ejecución de las actividades en la capacitación y adiestramiento de usuarios, se requiere de recurso técnico, humano y materiales que den soporte y completen dichas actividades, razón suficiente para tener dentro de los elementos de planificación el listado de recursos necesarios clasificados por su tipo. En las tablas 5.7, 5.8 y 5.9 se presenta el Recurso Técnico, Humano y Materiales, respectivamente a utilizar.

Recurso Técnico		
Recurso	Unidades	Descripción
<b>Computadoras Terminales</b>	11	Una computadora para 10 para un grupo de 10 usuarios máximo por capacitación y una computadora para el capacitador.
<b>Direcciones IP</b>	11	Direcciones IP de acceso a la red local y a la aplicación, para 11 terminales.
<b>Proyector</b>	1	Un proyector para las presentaciones y ejemplificación del manejo de la herramienta.
<b>Energía en la Sala de capacitación</b>	Según horas de capacitación	Energía eléctrica disponible en el local que se utilizara para las capacitaciones
<b>Sala de capacitación</b>	1	La sala de capacitación debe tener una capacidad de espacio para ubicar 11 terminales, 10 usuarios y el capacitador, dejando un margen de libertad de movimiento del 50% del espacio.

Tabla 5.7: “Tabla de Recursos Técnicos a utilizar”

Recurso Humano		
Recurso	Unidades	Descripción
<b>Capitador de personal</b>	1	Persona que brindara las capacitaciones a los usuarios sobre la nueva herramienta.
<b>Administrador de aplicaciones</b>	1	Encargado de configurar las computadoras con acceso a la red local y a la aplicación.

Tabla 5.8: “Tabla de Recurso Humano a necesitar”

<b>Materiales</b>		
<b>Recurso</b>	<b>Unidades</b>	<b>Descripción</b>
<b>Manual de usuario</b>	8	Guía de usuario para el uso de la herramienta, soporte físico para el proceso de capacitación.
<b>Libreta de anotaciones</b>	20	Soporte físico para anotación.
<b>Lápiz</b>	24	Para escribir.

**Tabla 5.9: “Tabla de Materiales a Utilizar”**

**5.3.3.3. Control de la Implantación**

<b>Objetivo</b>	
Asegurar la calidad de la implementación del aplicativo construido según lo planificado y programado.	
<b>Actividades</b>	
<ul style="list-style-type: none"> <li>• Verificación de cumplimiento de control. Verificar que la implementación del sistema se realice según lo planificado; para que de esta manera logre establecer la trayectoria a seguir ante posibles variaciones; creando así, una actividad de auditoría con lo que se persigue, verificar que la implementación se ejecuta según los elementos de control establecidos.</li> <li>• Elaboración de informes de avance y control. Elaborar los informes sobre el seguimiento del plan de implementación del sistema y de los inconvenientes detectados durante las actividades de verificación realizadas como parte del control • de avance tanto financiero como de recurso humano y tiempo</li> </ul>	
Tiempo	2.17 días
Horas Hombre	26.67 horas
Costo	\$186.67

**5.3.3.4. Costos basados en horas hombre.**

Actividad	días	Fecha de Inicio	Fecha de Finalización	Costo	Horas
<b>Planeación de Implantación</b>	<b>3</b>	<b>Lunes, 03 de Mayo de 2010</b>	<b>Miércoles, 05 de Mayo de 2010</b>	<b>\$406.67</b>	<b>41.33</b>
Definir estrategias y metas	1	Lunes, 03 de Mayo de 2010	Lunes, 03 de Mayo de 2010	\$56.00	8
Gestión de los recursos disponibles y limitantes	1.17	Martes, 04 de Mayo de 2010	Miércoles, 05 de Mayo de 2010	\$65.33	9.33
Preparación del ambiente organizacional	1	Martes, 04 de Mayo de 2010	Martes, 04 de Mayo de 2010	\$120.00	8
Envío de memo a las direcciones involucradas	0.83	Martes, 04 de Mayo de 2010	Martes, 04 de Mayo de 2010	\$100.00	6.67
Reunión con Directores	1.17	Martes, 04 de Mayo de 2010	Miércoles, 05 de Mayo de 2010	\$65.33	9.33
<b>Organización Implantación</b>	<b>12.5</b>	<b>Jueves, 06 de Mayo de 2010</b>	<b>Lunes, 24 de Mayo de 2010</b>	<b>\$654.67</b>	<b>100</b>
Definición del equipo ejecutor	7.83	Jueves, 06 de Mayo de 2010	Lunes, 17 de Mayo de 2010	\$438.67	62.67
Estudio de perfiles y funciones del equipo ejecutor	2.83	Lunes, 17 de Mayo de 2010	Jueves, 20 de Mayo de 2010	\$113.33	22.67
Definición de responsabilidades	1.83	Jueves, 20 de Mayo de 2010	Lunes, 24 de Mayo de 2010	\$102.67	14.67
<b>Ejecución de la Implantación</b>	<b>70.67</b>	<b>Jueves, 27 de Mayo de 2010</b>	<b>Jueves, 02 de Septiembre de 2010</b>	<b>\$3,141.33</b>	<b>573.33</b>
Instalación y configuración de equipo informático	3	Jueves, 27 de Mayo de 2010	Lunes, 31 de Mayo de 2010	\$120.00	24
Ejecución y control de ETL inicial	1.17	Martes, 01 de Junio de 2010	Miércoles, 02 de Junio de 2010	\$46.67	9.33
Prueba de aplicativo (control de calidad).	32	Miércoles, 02 de Junio de 2010	Viernes, 16 de Julio de 2010	\$1,280.00	256
Análisis del proyecto	2	Miércoles, 02 de Junio de 2010	Viernes, 04 de Junio de 2010	\$80.00	16
Identificar escenarios de prueba	3.17	Viernes, 04 de Junio de 2010	Miércoles, 09 de Junio de 2010	\$126.67	25.33
Verificación de estándares y requerimientos	1.17	Miércoles, 09 de Junio de 2010	Jueves, 10 de Junio de 2010	\$46.67	9.33
Ejecución y documentación de pruebas	2.17	Jueves, 10 de Junio de 2010	Lunes, 14 de Junio de 2010	\$86.67	17.33
Ingresar, clasificar los errores y/o defectos	1	Lunes, 14 de Junio de 2010	Martes, 15 de Junio de 2010	\$40.00	8
Revalidar las correcciones	1.17	Martes, 15 de Junio de 2010	Miércoles, 16 de Junio de 2010	\$46.67	9.33
Regression Test	18.5	Miércoles, 16 de Junio de 2010	Martes, 13 de Julio de 2010	\$740.00	148
Realizar reportes estadísticos a final de proyecto	2.83	Martes, 13 de Julio de 2010	Viernes, 16 de Julio de 2010	\$113.33	22.67
Preparación de plan de capacitación	3.33	Viernes, 16 de Julio de 2010	Miércoles, 21 de Julio de 2010	\$160.00	26.67
Presentación y Capacitación	31.17	Miércoles, 21 de Julio de 2010	Jueves, 02 de Septiembre de 2010	\$1,534.67	257.33
Notificación de usuarios a capacitar	1	Miércoles, 21 de Julio de 2010	Jueves, 22 de Julio de 2010	\$48.00	8
Preparar Lugar de capacitación para usuarios	1.17	Miércoles, 21 de Julio de 2010	Jueves, 22 de Julio de 2010	\$46.67	9.33

<b>Actividad</b>	<b>días</b>	<b>Fecha de Inicio</b>	<b>Fecha de Finalización</b>	<b>Costo</b>	<b>Horas</b>
Capacitación	30	Jueves, 22 de Julio de 2010	Jueves, 02 de Septiembre de 2010	\$1,440.00	240
actividad - 1	1	Jueves, 22 de Julio de 2010	Viernes, 23 de Julio de 2010	\$48.00	8
actividad - 2	2	Viernes, 23 de Julio de 2010	Martes, 27 de Julio de 2010	\$96.00	16
actividad – 3	4	Martes, 27 de Julio de 2010	Lunes, 02 de Agosto de 2010	\$192.00	32
actividad – 4	2	Lunes, 02 de Agosto de 2010	Miércoles, 04 de Agosto de 2010	\$96.00	16
actividad – 5	3	Miércoles, 04 de Agosto de 2010	Lunes, 09 de Agosto de 2010	\$144.00	24
actividad – 6	2	Lunes, 09 de Agosto de 2010	Miércoles, 11 de Agosto de 2010	\$96.00	16
actividad – 7	4	Miércoles, 11 de Agosto de 2010	Martes, 17 de Agosto de 2010	\$192.00	32
actividad – 8	2	Martes, 17 de Agosto de 2010	Jueves, 19 de Agosto de 2010	\$96.00	16
actividad – 9	3	Jueves, 19 de Agosto de 2010	Martes, 24 de Agosto de 2010	\$144.00	24
actividad – 10	1	Martes, 24 de Agosto de 2010	Miércoles, 25 de Agosto de 2010	\$48.00	8
actividad – 11	2	Miércoles, 25 de Agosto de 2010	Viernes, 27 de Agosto de 2010	\$96.00	16
actividad – 12	4	Viernes, 27 de Agosto de 2010	Jueves, 02 de Septiembre de 2010	\$192.00	32
<b>Control de la Implantación</b>	<b>2.17</b>	<b>Jueves, 02 de Septiembre de 2010</b>	<b>Lunes, 06 de Septiembre de 2010</b>	<b>\$186.67</b>	<b>26.67</b>
Verificación de cumplimiento de control	1.17	Jueves, 02 de Septiembre de 2010	Viernes, 03 de Septiembre de 2010	\$65.33	9.33
Elaboración de informes de avance y control	2.17	Jueves, 02 de Septiembre de 2010	Lunes, 06 de Septiembre de 2010	\$121.33	17.33
<b>Totales</b>	<b>88.34</b>			<b>\$4,389.34</b>	<b>4741.33</b>

**5.3.3.5. Resumen de Costos basados en Horas Hombre**

<b>MACROACTIVIDAD</b>	<b>COSTO</b>
Planeación de Implantación	\$406.67
Organización Implantación	\$654.67
Ejecución de la Implantación	\$3,141.33
Control de la Implantación	\$186.67
<b>TOTAL</b>	<b>\$4,389.34</b>

**5.3.3.6. Costo por Materiales**

<b>Materiales</b>			
<b>Recurso</b>	<b>Unidades</b>	<b>Costo Unitario</b>	<b>Costo Total</b>
Manual de usuario	8	\$6.00	\$48.00
Libreta de anotaciones	20	\$1.50	\$30.00
Lápiz	24	\$0.12	\$2.88
<b>TOTALES</b>			<b>\$80.88</b>

**5.3.3.7. Total de Costos**

Resumen de Costos en los generados por las horas/hombre y los materiales relacionados con la implementación del proyecto.

<b>DETALLE DE COSTOS</b>	<b>TOTAL</b>
Total costos por horas hombre:	\$4,389.34
Total de costos por materiales:	\$80.88
Total final de costos:	<b>\$4,470.22</b>

### 5.4. CRONOGRAMA DE ACTIVIDADES

No	Actividad	Días	Predecesora	Recurso Asignado
1	<b>Planeación de Implantación</b>	3		
2	Definir estrategias y metas	1		Project Manager
3	Gestión de los recursos disponibles y limitantes	1.17	2	Project Manager
4	Preparación del ambiente organizacional	1		Gerente de Sistemas
5	Envío de memo a las direcciones involucradas	0.83		Gerente de Sistemas
6	Reunión con Directores	1.17	5	Project Manager
7	<b>Organización Implantación</b>	12.5	1	
8	Definición del equipo ejecutor	7.83		Project Manager
9	Estudio de perfiles y funciones del equipo ejecutor	2.83	8	Tecnico de RRHH
10	Definición de responsabilidades	1.83	9	Project Manager
11	<b>Ejecución de la Implantación</b>	70.67	7	
12	Instalación y configuración de equipo informático	3		Administrador de Aplicación 1
13	Ejecución y control de ETL inicial	1.17	12	Administrador de Aplicación 2
14	Prueba de aplicativo (control de calidad).	32	13	
15	Análisis del proyecto	2		Administrador de Aplicación 1
16	Identificar escenarios de prueba	3.17	15	Administrador de Aplicación 1
17	Verificación de estándares y requerimientos	1.17	16	Administrador de Aplicación 1
18	Ejecución y documentación de pruebas	2.17	17	Administrador de Aplicación 1
19	Ingresar, clasificar los errores y/o defectos	1	18	Administrador de Aplicación 2
20	Revalidar las correcciones	1.17	19	Administrador de Aplicación 2
21	Regression Test	18.5	20	Administrador de Aplicación 2
22	Realizar reportes estadísticos al final del proyecto	2.83	21	Administrador de Aplicación 2
23	Preparación de plan de capacitación	3.33	14	Capacitador
24	Presentación y Capacitación	31.17	23	
25	Notificación de usuarios a capacitar	1		Capacitador
26	Preparar Lugar de capacitación para usuarios	1.17		Administrador de Aplicación 1

No	Actividad	Días	Predecesora	Recurso Asignado
27	Capacitación	30	26	
28	actividad - 1	1		Capacitador
29	actividad - 2	2	28	Capacitador
30	actividad - 3	4	29	Capacitador
31	actividad - 4	2	30	Capacitador
32	actividad - 5	3	31	Capacitador
33	actividad - 6	2	32	Capacitador
34	actividad - 7	4	33	Capacitador
35	actividad - 8	2	34	Capacitador
36	actividad - 9	3	35	Capacitador
37	actividad - 10	1	36	Capacitador
38	actividad - 11	2	37	Capacitador
39	actividad - 12	4	38	Capacitador
40	<b>Control de la Implantación</b>	2.17	11	
41	Verificación de cumplimiento de control	1.17		Project Manager
42	Elaboración de informes de avance y control	2.17		Project Manager

En el cronograma anterior se ha listado todos los elementos terminales del proyecto, con los días previstos a tomar por cada actividad. Un diagrama Gantt en el Anexo 12.1 proporciona una representación grafica y total del proyecto, siendo la guía a tomar para llevar a cabo la implantación del sistema desarrollado.

Añadido a la presentación del diagrama Gantt se puede observar el análisis Pert en el Anexo 12.2 los cuales poseen el análisis y representación grafica respectivamente. Esto ayudara a analizar las tareas involucradas para completar el proyecto.

## 5.5. SISTEMA DE CONTROL

Con el objetivo de ejecutar el plan de implantación de una manera eficiente, se propone un plan de control el cual se puede apoyar con técnicas de medición basadas en resultados de formularios, los cuales brindan información del progreso de la implantación proyectando un seguimiento en función del tiempo.

La técnica basada en formularios se ejecuta principalmente por el encargado de la actividad, el cual debe velar que la ejecución de la tarea sea completada respecto a lo programado con anterioridad, haciendo uso de las herramientas, recursos y tiempo previamente establecidos, en el caso de surgir alguna variante, se realiza la corrección o adaptación que sea necesaria para la correcta ejecución de las actividades.

Una vez ejecutada la actividad y completado los formularios de control deberán archivarse para tener una bitácora de ejecución, y además deben considerarse los cambios efectuados en el proceso mismo de implantación, para tener un control más minucioso y completo de las actividades.

### 5.5.1. Formularios para el Control

La documentación de actividades genera un beneficio al momento de estudiar fallas, éxitos, deficiencias o mejoras que se hayan presentado en las actividades o tareas desarrolladas, además se puede llevar un mejor control de gastos en los que se haya incurrido, con el propósito de mejorar futuras actividades para lograr reducir costos a la organización. Por esa razón, los formularios propuestos para utilizar como herramientas de control son los siguientes:

- Formulario de control de Actividades
- Formulario de control de tareas
- Formulario de gastos

#### 5.5.1.1. Formulario De Control De Actividades

Este formulario se deberá utilizar en el caso de llevar un control de todas las actividades programadas para la Implantación, de esta manera se posee un panorama más amplio del progreso de Implantación, además se lleva un control de las actividades con mayor prioridad y las actividades que se deben de ejecutar antes que otras.

- Elementos Importantes del Formulario:
- Elaborado por: Coordinador de Proyecto.
- Orientado a: Gerente de Sistemas
- Frecuencia Recomendada: Periodos semanales de control

Instructivo de Llenado de Formulario:

1. Detallar la persona que llena el Formulario, con nombres y apellidos.
2. Agregar la fecha de elaboración del formulario (dd/mm/aaaa).
3. Agregar la fecha de inicio del periodo del informe (dd/mm/aaaa).
4. Agregar la fecha fin del periodo del informe (dd/mm/aaaa).
5. Detallar el nombre de la persona responsable de la actividad.
6. Anotar el correlativo de la actividad.
7. Escribir el nombre de la actividad.
8. Verificar si la actividad ha sido ejecutada.
9. Verificar si la actividad está pendiente.
10. Colocar el número de actividades esperadas al momento.
11. Calcular el porcentaje de actividades esperadas respecto al total de actividades al momento.
12. Colocar el número de actividades realizadas al momento.
13. Calcular el porcentaje de tareas realizadas.
14. Anotar número de actividades pendientes.
15. Anotar porcentaje de tareas pendientes.
16. Anotar total de tareas.
17. Anotar porcentaje total.
18. Agregar observaciones referentes a las actividades ejecutadas.
19. Agregar nombre de la persona que revisa el formulario.
20. Agregar fecha de revisión de formulario (dd/mm/aaaa).
21. Anotar el número de página del formulario.
22. Anotar el número total de páginas de formulario.



### 5.5.1.2. Formulario De Control De Tareas

Este formulario se deberá utilizar en el caso de llevar un control de todas las tareas programadas para la Implantación.

- Elementos Importantes del Formulario:
- Elaborado por: Coordinador de Proyecto.
- Orientado a: Gerente de Sistemas
- Frecuencia Recomendada: Periodos semanales de control

Instructivo de Llenado de Formulario:

1. Detallar la persona que llena el Formulario, con nombres y apellidos.
2. Agregar la fecha de elaboración del formulario (dd/mm/aaaa).
3. Agregar la fecha de inicio del periodo del informe (dd/mm/aaaa).
4. Agregar la fecha fin del periodo del informe (dd/mm/aaaa).
5. Detallar el nombre de la persona responsable de la tarea.
6. Anotar el correlativo de la tarea.
7. Escribir el nombre de la tarea.
8. Verificar porcentaje esperado de la tarea a la fecha.
9. Verificar porcentaje real de avance de la tarea.
10. Agregar observaciones referentes a las tareas ejecutadas.
11. Agregar nombre de la persona que revisa el formulario.
12. Agregar fecha de revisión de formulario (dd/mm/aaaa).
13. Anotar el número de página del formulario.
14. Anotar el número total de páginas de formulario.

IMPLANTACIÓN DEL PROYECTO DESARROLLO DE UN DATAWAREHOUSE PARA EL PROCESO DE LAS DENUNCIAS DE LA DEFENSORÍA DEL CONSUMIDOR				
FORMULARIO DE CONTROL DE TAREAS				No. 2
Elaborado por:	(1)		Fecha elaboración:	(2)
Periodo:	Fecha Inicio:	(3)	Fecha Fin:	(4)
Responsable de la Tarea:	(5)			
No.	Nombre de la Tarea		Porcentaje esperado	Porcentaje de Avance
(6)	(7)		(8)	(9)
Observaciones	(10)			
Revisado por:	(11)	Fecha de revisión:	(12)	
No. Pagina	(13)	Total Paginas:	(14)	

Figura 5.3: Formulario de control de tareas

### 5.5.1.3. Formulario De Control De Gastos

Este formulario se deberá utilizar en el caso de llevar un control de todos los gastos estimados y los gastos no previstos de la Implantación.

- Elementos Importantes del Formulario:
- Elaborado por: Coordinador de Proyecto.
- Orientado a: Gerente de Sistemas
- Frecuencia Recomendada: Periodos semanales de control

Instructivo de Llenado de Formulario:

1. Detallar la persona que llena el Formulario, con nombres y apellidos.
2. Agregar la fecha de elaboración del formulario (dd/mm/aaaa).
3. Agregar la fecha de inicio del periodo del informe (dd/mm/aaaa).
4. Agregar la fecha fin del periodo del informe (dd/mm/aaaa).
5. Detallar el nombre de la persona responsable de la actividad.
6. Anotar el correlativo de la actividad.
7. Escribir el nombre de la actividad.
8. Colocar la fecha en la que se ejecuto la actividad (dd/mm/aaaa).
9. Detallar el costo total incurrido para dicha actividad.
10. Realiza la suma total de todas las actividades que se han realizado.
11. Agregar observaciones referentes a las actividades ejecutadas.
12. Agregar nombre de la persona que revisa el formulario.
13. Agregar fecha de revisión de formulario (dd/mm/aaaa).
14. Anotar el número de página del formulario.
15. Anotar el número total de páginas de formulario.



### 5.5.2. Estrategias De Control

Se recomienda además de hacer uso de las herramientas de control propuesta, considerar las siguientes estrategias para mejorar en la calidad de la implantación propuesta:

Realizar el control personal de parte de los encargados de las actividades y tareas, para tener un cronograma histórico del plan de implantación, logrando documentar cada fase, cada tarea y cada gasto de una manera constante, realizando las evaluaciones de una manera sincera y efectiva para lograr evaluar el desempeño del personal teniendo el grado de eficiencia y calidad con el que se ha desarrollado cada fase.

Evaluar los resultados de los formularios con cierta frecuencia, para identificar aquellos elementos positivos y negativos que se han presentado, de tal manera que se pueda corregir aquello que genera atraso o mayor costo, con esto se logra realizar una mejor ejecución del plan propuesto y se mejoran considerablemente las actividades y tareas, minimizando los tiempos y gastos.

Proponer un sistema de control alternativo en el que se pueda evaluar de manera indirecta a las personas involucradas para tener un mecanismo de control en el que se pueda llevar un seguimiento de todo el plan, esto en el caso que no se lleve un control estricto por parte de los encargados siempre se podrán con otro método de control de respaldo.

Los elementos identificados como posibles mejoras se deberán no solo detallar en un reporte, sino mas bien tratar de agregarlos al plan de implantación para que se pongan en marcha, con el propósito de mejorar procesos y disminuir costos de tiempo y otros elementos en los que se pueda generar atrasos de ejecución.

## CONCLUSIONES

- La definición y utilización de estándares conlleva a desarrollar de forma ordenada y uniforme, las actividades en las iteraciones del ciclo de vida, proporcionando a los desarrolladores y usuario una fácil comprensión de los elementos del sistema, lo que significa una notable mejora en el mantenimiento que se aplicará al sistema una vez finalizado por el hecho de que todos los sistemas con el tiempo tienen que cambiar para adaptarse a las necesidades del negocio, con el uso de estándares estas ventajas pueden ser aprovechadas enormemente con el fin de brindar un mejor servicio al usuario.
- El uso de técnicas tales como la entrevista, cuestionarios y observación directa, ayudan a la realización del diagnóstico de la situación actual, permitiendo formular la problemática de forma precisa y objetiva, estableciendo las causas que la originan y facilitando presentar una propuesta de solución acorde a las necesidades que se tienen.
- El análisis de factibilidades permitió demostrar que efectivamente era factible realizar el presente proyecto, haciendo constar que los aspectos técnicos necesarios están disponibles para el desarrollo del proyecto, teniendo como respaldo el análisis económico que demostró el retorno de inversión en 5 años y finalmente se plasma operativamente que es posible la implementación del proyecto y la aceptación de este por parte del usuario final.
- A partir de la investigación preliminar de las metodologías se logró obtener cual era la que mejor se adecuaba a las necesidades del proyecto apoyándose en la tecnología con la que cuenta la organización, gracias a dicha investigación se logró identificar los elementos que se debían considerar dentro del análisis y diseño sin perder el enfoque principal del proyecto.
- Modelar la arquitectura del sistema permite determinar y comprender el funcionamiento de los distintos elementos que lo conforman en cada una de las fases. Esta arquitectura es el conjunto de elementos que proporcionan un marco de referencia para guiar la construcción del sistema. Con esta arquitectura se puede establecer los fundamentos para que los analistas, diseñadores, programadores, etc. trabajen en una línea en común que permita alcanzar los objetivos del sistema de información.
- El diseño de las seguridades tanto físicas como lógicas permiten establecer reglas para proteger los datos de cualquier amenaza que pueda presentarse, permitiendo el acceso a la información únicamente a los usuarios autorizados a utilizar el sistema, basándose en perfiles de usuario previamente definidos. La seguridad física describe todos los aspectos tangibles que ayudan a salvaguardar los datos; en cambio la seguridad lógica se consideran todos los mecanismos intangibles de protección. Estas dos seguridades se complementan y deben estar lo mejor definidas especialmente para este proyecto en el cual se tienen accesos a datos confidenciales.
- La utilización de UML como herramienta de apoyo para el modelado, contribuyó enormemente a iteraciones posteriores, facilitando la integración de prototipos y a realizar cambios según las observaciones del usuario; debido a que se especifican los elementos según la interacción que realizan, la documentación de dichos procesos es mucho más comprensible y fácil de realizar.

- La integración de cada una de las iteraciones a lo que se convertiría en producto final se facilitó grandemente gracias a tener definida una metodología y establecer estándares desde la primera iteración, esto permitió poder brindarle más tiempo a la evaluación por parte del usuario y poder cumplir con las necesidades que el usuario planteaba.
- El apoyo completo de la organización es necesario para el desarrollo e implantación de este tipo de proyectos, ya que por ser un área en la que recientemente las empresas nacionales están invirtiendo la resistencia al cambio por parte de los usuarios se convierte en un completo desafío que solo se puede sobrellevar definiendo claramente un plan de implantación.
- Respecto a la barrera de recursos económicos con la que se enfrentan las instituciones públicas del país, la investigación y explotación de herramientas de licencia gratuita, da una oportunidad para que proyectos de esta índole puedan encontrar un papel protagónico y contribuir al desarrollo social y tecnológico.

## RECOMENDACIONES

- Se recomienda considerar las limitantes que se presentan como la documentación y soporte para las herramientas de software libre que se han utilizado.
- Se recomienda encarecidamente el uso de estándares para desarrollar las diferentes actividades que exige el mantenimiento de este sistema, ya que de esta manera se mejoraría significativamente la comunicación y la comprensión de todos aquellos elementos involucrados con el sistema.
- El paradigma orientado a objetos provee una infinidad de elementos para modelar una solución de software en las diferentes etapas del ciclo de vida de desarrollo de sistemas, se recomienda utilizar aquellos que se apeguen a la necesidad del negocio, ya que existen algunos componentes que describen de mejor forma una situación en particular.
- Reconociendo la amplitud y variedad de conocimientos requeridos para implementar la inteligencia de negocios es recomendable promover cursos, seminarios y congresos orientados a impulsar las nuevas Tecnologías de Información y Comunicación, a fin de familiarizar con dicho concepto al personal de la organización.
- Para el desarrollo de este proyecto se proporciono la ultima versión de la base de datos transaccional del sistema SARA, posterior al desarrollo completo, dicha versión fue sustituida por nueva versiones por desarrollos de otros modulos, razón por la cual se recomienda confrontar contra la ultima versión de la base transaccional, con el objetivo de encontrar posibles cambios que puedan afectar de alguna manera la solución desarrollada.

## BIBLIOGRAFÍA

### Libros

- The Data Warehouse Lifecycle Toolkit, Ralph Kimball, Laua Reeves, Margy Ros, 2<sup>nd</sup> Edition, Wiley, 2009, ISBN 8126516895, 9788126516896
- Building the data warehouse, William H. Inmon, 4<sup>th</sup> Edition, 2005, Wiley, 2005, ISBN 0764599445, 9780764599446

### Tesis

- Data Warehouse Design with UML, Sergio Lujan, Juan Trujillo, Department of software and computing systems, University of Alicante.
- Desarrollo de un portal web para dar soporte organizacional al programa olimpiadas especiales El Salvador, Danuel Carranza, Nilson Nuñez, Universidad de El Salvador.

### Sitios Web

- Pentaho Community Forums
- <http://forums.pentaho.org/archive/index.php>
- <http://wiki.bizcubed.com.au:8080/xwiki/bin/view/Pentaho+Tutorial/Security>
- <http://wiki.pentaho.com/display/ServerDoc1x/Adhoc+Reporting+Templates>
- <http://wiki.pentaho.com/display/analysis/Publishing+an+Analysis+Schema+Using+Schema+Workbench>
- <http://wiki.pentaho.com/display/COM/Evaluating+Pentaho+Analysis>
- <http://wiki.pentaho.com/display/ServerDoc2x/Configuring+JNDI+Datasources+for+use+in+the+Pentaho+BI+Platform>
- <http://wiki.pentaho.com/display/ServerDoc2x/Security>
- [http://mondrian.pentaho.org/documentation/schema.php#Access\\_control](http://mondrian.pentaho.org/documentation/schema.php#Access_control)
- Comparación entre Red Hat Enterprise y Suse Enterprise  
[http://www.novell.com/rc/docrepository/public/37/basedocument.2009-04-01.6574230940/SLES\\_vs\\_Red\\_Hat\\_Enterprise\\_Linux\\_White\\_Paper\\_en.pdf](http://www.novell.com/rc/docrepository/public/37/basedocument.2009-04-01.6574230940/SLES_vs_Red_Hat_Enterprise_Linux_White_Paper_en.pdf)
- Artículo sobre la comparación de la fiabilidad entre Windows server y Linux  
<http://www.prnewswire.co.uk/cgi/news/release?id=158688>
- Comparación de la seguridad entre Windows server y Red Hat  
[http://www.theregister.co.uk/2004/10/22/security\\_report\\_windows\\_vs\\_linux/#comparison](http://www.theregister.co.uk/2004/10/22/security_report_windows_vs_linux/#comparison)
- Artículo de Microsoft de comparación con Red Hat  
<http://www.microsoft.com/windowsserver/compare/default.mspx>
- El uso del prototipo en el ciclo de desarrollo de sistema  
Ing. Germán Pereyra, <http://www.monografias.com/trabajos12/proto/proto.shtml>
- Fundamentos de ingeniería de Software

Juan Antonio *López Quesada*, Departamento de Informática y Sistemas Facultad de Informática Campus Universitario de Espinardo – Murcia,  
[http://dis.um.es/~lopezquesada/documentos/FIS\\_0708/curso/UT3.pdf](http://dis.um.es/~lopezquesada/documentos/FIS_0708/curso/UT3.pdf)

- Prototyping and the Software Development Cycle  
Diane Wilson, Firelily Designs,  
<http://www.firelily.com/opinions/cycle.html>
- MPR: Metodología de Prototipado Rápido, <http://www.scaridad.com/files/MPR.pdf>

## GLOSARIO TÉCNICO

### A

**AD HOC:** Permite crear a los usuarios consultas específicas y personalizadas, utilizando una interfaz grafica amigable sin la necesidad de tener el conocimiento sobre SQL o el esquema de la base de datos.

**ADAPTABILIDAD:** Es la medida de que tan bien el sistema operativo puede interactuar con otros, ya que en nuestros días se cuenta con una gran variedad de ambientes como las aplicaciones en el cliente, terminales, etc.

**AGREGACIONES:** son resúmenes de datos precalculados que mejoran el tiempo de respuesta por el simple hecho de tener preparadas las respuestas antes de que se planteen las consultas.

**ASEQUIBILIDAD:** Se mide desde dos puntos de vista, precio y soporte. Es importante no ser engañado por un precio bajo si el costo del soporte es demasiado elevado.

### B

**BUSINESS INTELLIGENCE (BI)** es una disciplina que, junto con sus correspondientes herramientas, hacen centro en el análisis de la información para la correcta toma de decisiones que le permita a la organización cumplir con los objetivos de negocio.

### C

**Cronograma de actividades:** Se describe en forma progresiva las actividades a realizar en el tiempo, para asegurar el cumplimiento de las metas.

### D

**DASHBOARD:** es una página desarrollada en base a tecnología web mediante la cual se despliega en tiempo real información de la empresa extraída de varias fuentes o bases de datos. Su característica de tiempo real otorga a los usuarios un conocimiento completo sobre la marcha de la empresa y permite hacer análisis instantáneos e inteligencia de negocios.

**DATOS:** se refiere a medidas, cifras, nombres o hechos.

**DATA MART:** Son almacenes de datos con información de interés particular para un determinado sector de la empresa. El almacén de datos de un hecho en particular se denomina Data Mart (DM).

**DATA WAREHOUSING:** Los Data Warehouses o Almacenes de Datos se basan en estructuras multidimensionales (cubos) en las que se almacena la información calculando previamente todas las combinaciones de todos los niveles de todas las aperturas de análisis.

**DICE:** Al seleccionar varios miembros de varias dimensiones se forma sub-cubo, cubo más pequeño o dado (dice). Tanto Slice como Dice son formas particulares de Filtrado.

**DIMENSIONES:** Las dimensiones organizan los datos en función de un área de interés para los usuarios.

**DRILL-DOWN:** OLAP puede moverse en la dirección contraria y presentar automáticamente datos detallados que abarcan datos consolidados

**E**

**ENTIDAD:** Entidad: Representa una "cosa" u "objeto" del mundo real con existencia independiente, es decir, se diferencia únicamente de cualquier otro objeto o cosa, incluso siendo del mismo tipo.

**ESCALABILIDAD:** Numero de procesadores soportados, tamaño máximo de archivos soportado y tipos de sistema de archivo soportados.

**ETL:** Extraer, Transformar y Cargar (Extract, Transform and Load), llamado así al proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, data mart o data warehouse para analizar, o en otro sistema operacional para apoyar un proceso de negocio.

**F**

**FASMI:** Análisis rápido de información multidimensional compartida (Fast Analysis of Shared Multidimensional Information).

**FIABILIDAD:** una medida de que tan bien corre la computadora las aplicaciones y los servicios.

**I**

**IMPLANTAR:** Establecer y poner en ejecución nuevas doctrinas, instituciones, prácticas o costumbres.

**IMPLEMENTAR:** Poner en funcionamiento, aplicar métodos, medidas, etc., para llevar algo a cabo.

**INFORMACIÓN:** son datos procesados u organizados.

**L**

**LOG:** es un término anglosajón, equivalente a la palabra bitácora en lengua castellana.

**M**

**MANEJABILIDAD:** Representa el conjunto de funciones, que automatizan la operación del servidor, así como también las habilidades necesarias del personal responsable de la administración del sistema operativo y el servidor.

**MDX:** Expresiones Multidimensionales (acrónimo de Multidimensional Expressions)

**MEDIDAS:** Las medidas son los valores de datos que se analizan.

**METODOLOGÍA:** hace referencia al conjunto de procedimientos basados en principios lógicos, utilizados para alcanzar una gama de objetivos que rigen en una investigación científica o en una exposición doctrinal.

**MINERÍA DE DATOS:** es la tecnología cuyo objetivo es descubrir conocimiento oculto en un conjunto de datos y expresarlo de forma inteligible.

**MODELO PROTOTIPADO:** Modelo de ciclo de vida del tipo Evolutivo, basado en iteraciones compuestas en ciclos de vida tradicionales, cada iteración incluye refinamiento de los requerimientos y producto desarrollado.

MOLAP: Procesamiento Analítico Multidimensional en Línea (Multidimensional OnLine Analytical Processing).

## O

OLAP: Procesamiento Analítico en Línea (On Line Analytical Processing).

OLTP: es la sigla en inglés de Procesamiento de Transacciones En Línea (OnLine Transaction Processing) es un tipo de sistemas que facilitan y administran aplicaciones transaccionales, usualmente para entrada de datos y recuperación y procesamiento de transacciones.

ROLAP: Procesamiento Analítico Online Relacional, se trata de sistemas y herramientas OLAP construidas sobre una base de datos relacional.

## P

POLÍTICA DE SEGURIDAD: Una política de seguridad en el ámbito de la criptografía de clave pública o PKI es un plan de acción para afrontar riesgos de seguridad, o un conjunto de reglas para el mantenimiento de cierto nivel de seguridad. Pueden cubrir cualquier cosa desde buenas prácticas para la seguridad de un solo ordenador, reglas de una empresa o edificio, hasta las directrices de seguridad de un país entero.

PROCESO: Es la representación de un proceso de negocio de forma que admita manipulación y automatización. Está formado por un conjunto de actividades y relaciones entre ellas para indicar el comienzo y el fin del proceso.

## R

RELACIÓN: Describe cierta dependencia entre entidades o permite la asociación de las mismas.

ROL: función que alguien o algo cumple, papel de un actor

ROLL-UP: Calcula las medidas en función de agrupamientos, realiza el re-cálculo de la medida de acuerdo a los ajustes de escala.

ROTACIÓN: Selecciona el orden de visualización de las dimensiones, rota o gira el cubo según sus dimensiones.

## S

SARA: Sistema de Atención a Reclamos y Denuncias, sistema actualmente implementado y bajo desarrollo dentro de la Institución Gubernamental Defensoría del Consumidor.

SEGURIDAD: es la habilidad del sistema de protegerse de un ataque de denegación de servicios, así como también de intrusiones.

SLICE: Al seleccionar un miembro en particular de una dimensión se forma una especie de rebanada (slice) del cubo original.

## U

UML: Lenguaje de Modelado Unificado. Lenguaje gráfico para visualizar, especificar, construir y documentar un sistema.



**ANEXO 1: Matriz de Evaluación de Ciclo de Vida**

No	Requerimientos	Cascada	Incremental	Evolutivo	Prototipado	Espiral
	Los requerimientos son fácilmente definidos y/o bien conocidos?	SI	SI	NO	NO	NO
	Pueden ser los requerimientos definidos temprano durante el ciclo de vida?	SI	SI	SI	NO	NO
	Cambiarán los requerimientos frecuentemente?	NO	NO	NO	NO	SI
	Hay la necesidad de demostrar los requerimientos para alcanzar la definición?	NO	NO	NO	SI	SI
	Hay la necesidad de demostrar la validez de concepto para el sistema?	NO	NO	NO	SI	SI
	Especifican los requerimientos un sistema complejo?	NO	SI	SI	SI	SI
	Es uno de los requerimientos la funcionalidad temprana del sistema?	NO	SI	SI	SI	SI
Otros Factores						
	El gerente de proyecto supervisará cercanamente el progreso del proyecto?	SI	SI	SI	NO	SI
	Es una prioridad la facilidad de asignación de recursos?	SI	SI	SI	NO	NO

## ANEXO 2: Formulario para Evaluación del Prototipo

En la siguiente matriz se permite llevar el control de la aceptación por parte del usuario.

Forma de evaluación de Prototipo				
Nombre del observador:			Fecha:	
Nombre del sistema o proyecto:			Compañía o ubicación:	
Nombre o N° de Programa			Versión	
	Usuario 1	Usuario 2	Usuario 3	Usuario 4
Nombre de Usuario				
Periodo Observado				
Reacciones del usuario				
Sugerencias del Usuario				
Innovaciones				
Planes de Revisión				

### ANEXO 3: Calculo de la Muestra de la Población<sup>51</sup>

El cálculo del tamaño de la muestra es uno de los aspectos importantes que ayudan a concretar el grado de credibilidad que concederemos a los resultados de la investigación.

Para realizar dicho calculo se decidió implementar una fórmula que orienta sobre el cálculo del tamaño de la muestra para datos globales, dicha fórmula se detalla a continuación.

$$n = \frac{k^2 * p * q * N}{(e^2 * (N - 1)) + k^2 * p * q}$$

Donde:

**N:** es el tamaño de la población o universo (número total de posibles encuestados).

**k:** es una constante que depende del nivel de confianza que asignemos. El nivel de confianza indica la probabilidad de que los resultados de nuestra investigación sean ciertos: un 95,5 % de confianza es lo mismo que decir que nos podemos equivocar con una probabilidad del 4,5%.

Los valores k más utilizados y sus niveles de confianza son:							
k	1,15	1,28	1,44	1,65	1,96	2	2,58
Nivel de confianza	75%	80%	85%	90%	95%	95.5%	99%

**e:** es el error muestral deseado. El error muestral es la diferencia que puede haber entre el resultado que obtenemos preguntando a una muestra de la población y el que obtendríamos si preguntáramos al total de ella

**p:** es la proporción de individuos que poseen en la población la característica de estudio. Este dato es generalmente desconocido y se suele suponer que p=q=0.5 que es la opción más segura.

**q:** es la proporción de individuos que no poseen esa característica, es decir, es 1-p.

**n:** es el tamaño de la muestra (número de encuestas que vamos a hacer).

A continuación se muestra los datos que utilizamos para calcular el espacio muestral:

N=	137	Cantidad de usuarios
p=	0.23	Porcentaje de usuarios que utilizan reportes
q=	0.77	Porcentaje de usuarios que dependen de los reportes
K=	1.44	Para nivel de confianza del 85 %
e=	0.1	Diferencia que puede haber entre el resultado de encuestar a todos o solo a la muestra

Sustituyendo en la ecuación:

$$n = \frac{(1.44)^2 * 0.23 * 0.77 * 137}{(0.1^2 * (137 - 1)) + 1.44^2 * 0.23 * 0.77} = 29.1281 \cong 2$$

<sup>51</sup> <http://www.feedbacknetworks.com>, tema: La experiencia de preguntar

Por lo que se determina que el tamaño adecuado para la muestra según la población es de 29 personas, en las cuales se distribuirá con ayuda de la gerencia de sistemas, debido a que los usuarios del sistema se encuentran distribuidos según la ubicación de los centros de solución de controversias y unidades que conforman la defensoría del consumidor.

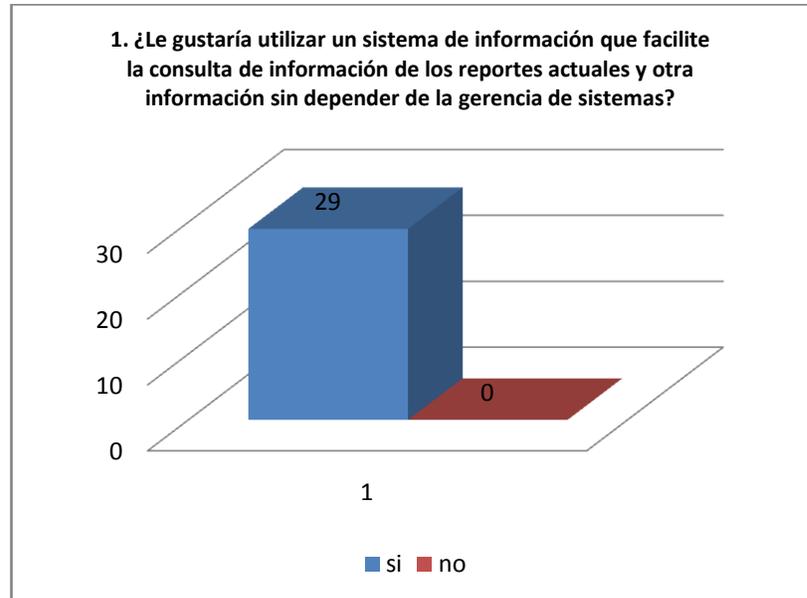
La obtención de los datos a implementar han sido calculados en base a la información brindada en las reuniones y entrevistas con los usuarios de la defensoría del consumidor obteniendo los siguientes resultados.

## ANEXO 4: Resultados de las Encuestas

### 4.1 ENCUESTA DE ACEPTACIÓN DE USUARIO

1. ¿Le gustaría utilizar un sistema de información que facilite la consulta de información de los reportes actuales y otra información sin depender de la gerencia de sistemas?

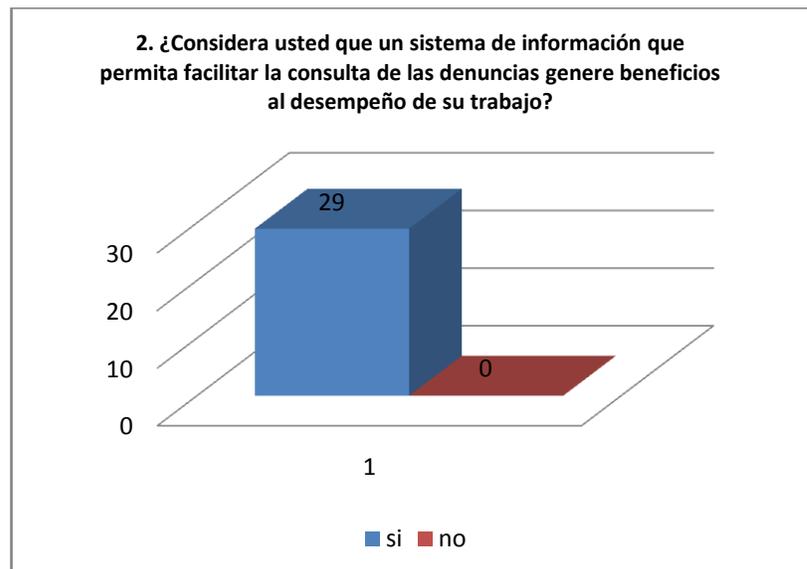
Respuesta	Cantidad
Si	29
No	0
<b>TOTAL</b>	<b>29</b>



El 100% de las personas encuestadas consideran que es necesario contar con un sistema de información que sea capaz de permitir analizar información y generar reportes siendo independientes de la gerencia de sistemas.

2. ¿Considera usted que un sistema de información que permita facilitar la consulta de las denuncias genere beneficios al desempeño de su trabajo?

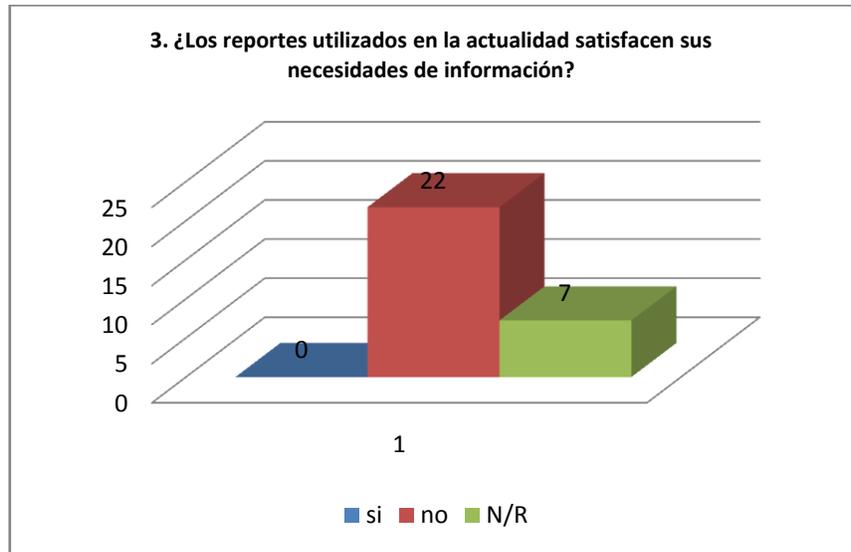
Respuesta	Cantidad
si	29
no	0
<b>TOTAL</b>	<b>29</b>



Todas las personas encuestadas consideran que un sistema capaz de facilitar el análisis y consulta de información mejorara el desempeño en sus labores.

3. ¿Los reportes utilizados en la actualidad satisfacen sus necesidades de información?

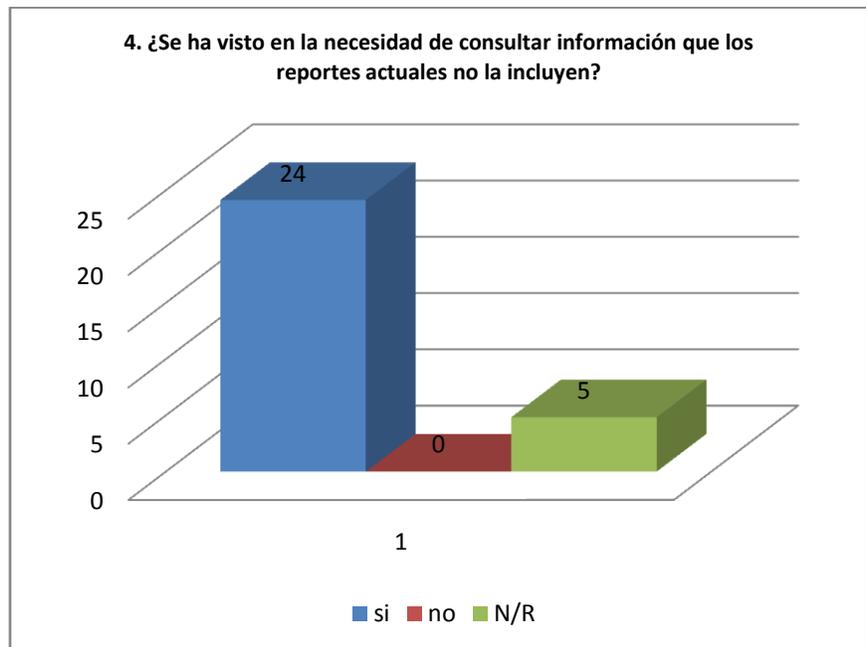
Respuesta	Cantidad
si	0
no	22
N/R	7
<b>TOTAL</b>	<b>29</b>



Un 75.86 % de los encuestados considera que los reportes generados actualmente por el sistema no satisfacen completamente sus necesidades de información

4. ¿Se ha visto en la necesidad de consultar información que los reportes actuales no la incluyen?

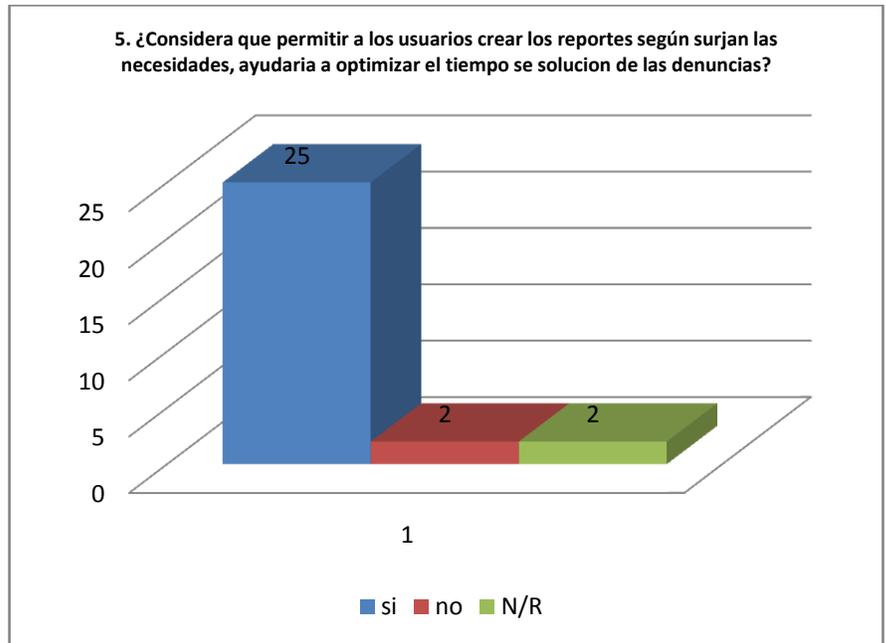
Respuesta	Cantidad
si	24
no	0
N/R	5
<b>TOTAL</b>	<b>29</b>



El 82.76% de los encuestado han tenido la necesidad de analizar y visualizar información que no se encuentra disponible en los reportes generados actualmente por los sistemas

5. ¿Considera que permitir a los usuarios crear los reportes según surjan las necesidades, ayudaría a optimizar el tiempo se solución de las denuncias?

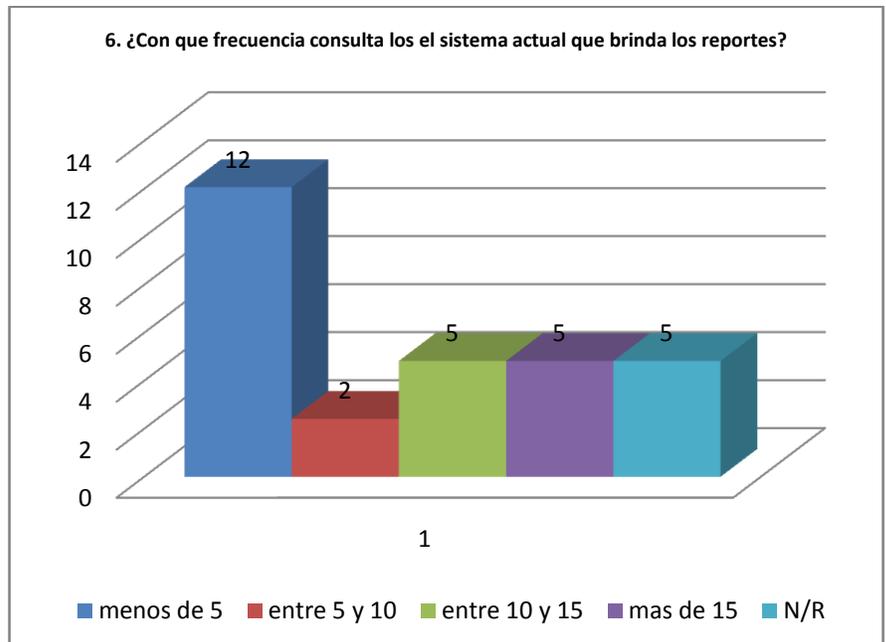
Respuesta	Cantidad
si	25
no	2
N/R	2
<b>TOTAL</b>	<b>29</b>



6.90% de las personas encuestadas, consideran que no se optimizaría el tiempo para resolver las denuncias al contar con un sistema que provea información para análisis y reportes que se ajusten a sus necesidades. En cambio un 86.2 % consideran que el tiempo de resolución de las denuncias sería acortado al contar con un sistema de este tipo.

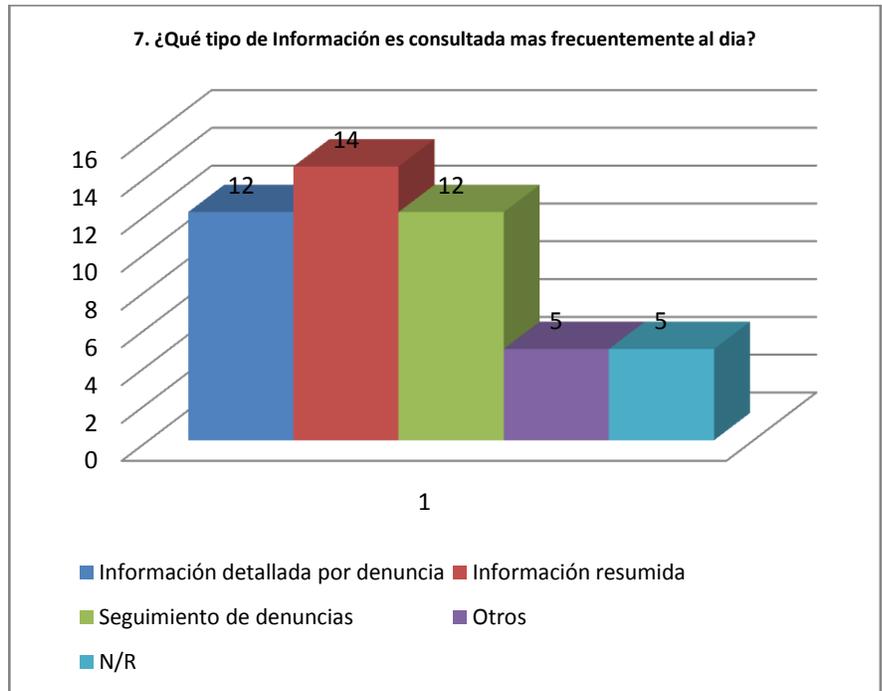
6. ¿Con que frecuencia consulta los el sistema actual que brinda los reportes?

Respuesta	Cantidad
menos de 5	12
entre 5 y 10	2
entre 10 y 15	5
mas de 15	5
N/R	5
<b>TOTAL</b>	<b>29</b>



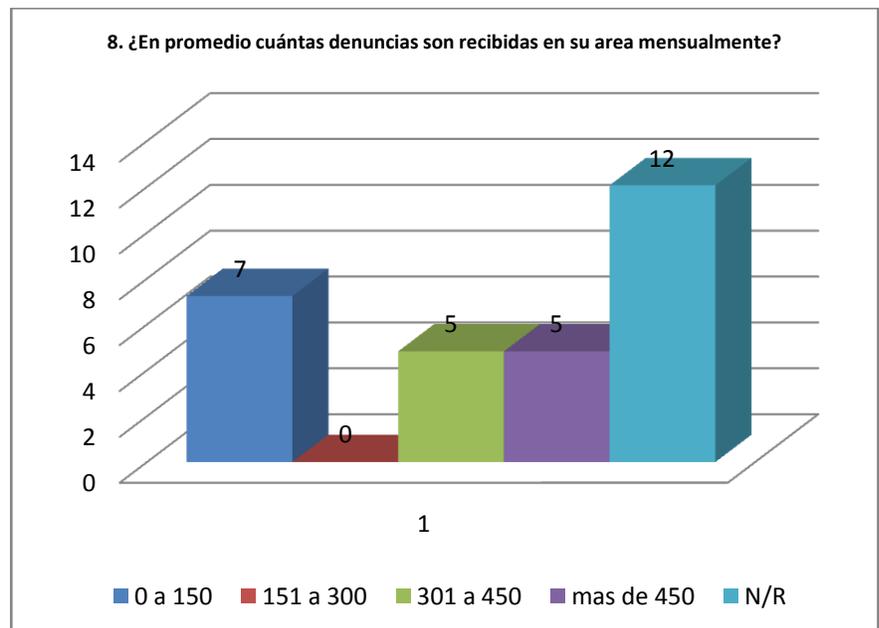
Se puede concluir que la importancia del sistema actual para el proceso de denuncias es de vital importancia, ya que más del 80% del personal realiza por lo menos una consulta de información por medio de reportes la cual le facilita la el análisis de la información.

7. ¿Qué tipo de Información es consultada más frecuentemente al día?	
Respuesta	Cantidad
Información detallada por denuncia	12
Información resumida	14
Seguimiento de denuncias	12
Otros	5
N/R	5
<b>TOTAL</b>	<b>48</b>



Un 48.27 % de las personas encuestadas tiene necesidades de información resumida, con la cual se puede realizar un mejor análisis de la información de los casos. Este tipo de información puede ser proporcionada por un Data Warehouse.

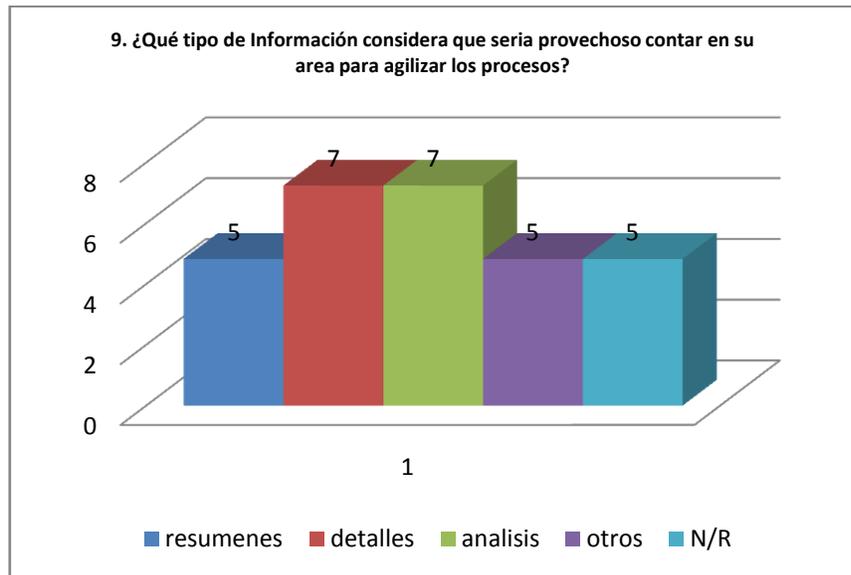
8. ¿En promedio cuántas denuncias son recibidas en su área mensualmente?	
Respuesta	Cantidad
0 a 150	7
151 a 300	0
301 a 450	5
mas de 450	5
N/R	12
<b>TOTAL</b>	<b>29</b>



La cantidad de información y reportes solicitados por los diferentes niveles de la Defensoría del Consumidor es significativa, por lo que es necesario un sistema que permita realizar con mayor facilidad un mejor análisis de la información.

9. ¿Qué tipo de Información considera que sería provechoso contar en su área para agilizar los procesos?

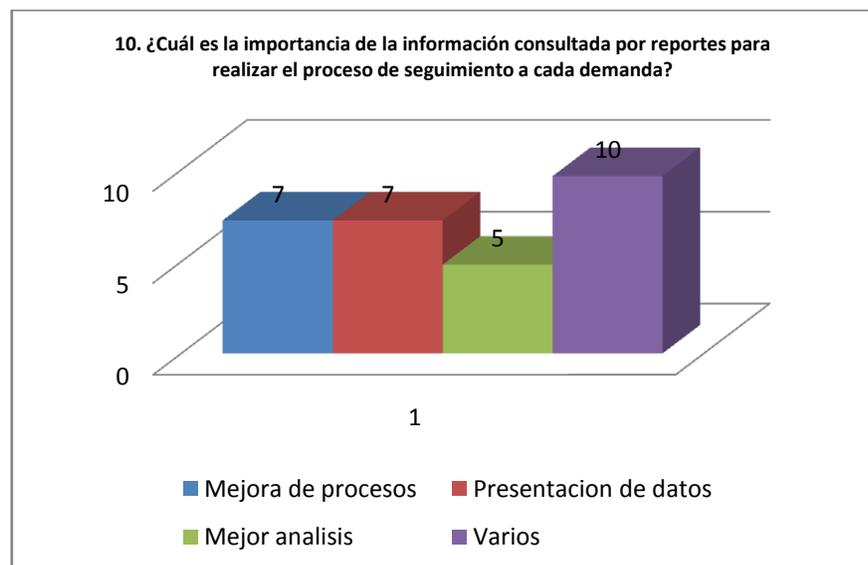
Respuesta	Cantidad
resúmenes	5
detalles	7
análisis	7
otros	5
N/R	5
<b>TOTAL</b>	<b>29</b>



De las personas encuestadas un 24.14% considera necesario contar con información que permita realizar un análisis más profundo de la información de las denuncias. El mismo porcentaje de personas respondió que necesita información detallada y un 17.24% que necesita información resumida. Esta variedad de formas de presentación de información puede ser fácilmente proporcionada por un data warehouse.

10. ¿Cuál es la importancia de la información consultada por reportes para realizar el proceso de seguimiento a cada demanda?

Respuesta	Cantidad
Mejora de procesos	7
Presentación de datos	7
Mejor análisis	5
Varios	10
<b>TOTAL</b>	<b>29</b>



Una gran parte de las personas encuestadas coincide en que la información que es consultada tiene una gran importancia para mejorar los procesos y la presentación de los datos de la organización. También que esta información es útil para realizar mejores análisis a los procesos de denuncias.

## 4.2 ENCUESTA DE ASPECTOS TÉCNICOS



UNIVERSIDAD DE EL SALVADOR  
 FACULTAD DE INGENIERÍA Y ARQUITECTURA  
 ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

Fecha 21 /04 /2009

**TRABAJO DE GRADUACIÓN:** DESARROLLO DE UN DATA WAREHOUSE AL PROCESO DE DENUNCIAS, PARA LA DEFENSORÍA DEL CONSUMIDOR.

**Introducción:** La siguiente encuesta es realizada con el fin de desarrollar un data warehouse que permita facilitar la consulta de información a la DC, el cual nosotros como estudiantes de la Universidad de El Salvador desarrollaremos para optar al título de Ingeniero de Sistemas Informáticos.

**Objetivo:** Justificar la situación actual de la DC y medir el grado de aceptación y aspectos técnicos para desarrollar un Sistema de información, que facilite la consulta de información.

**Indicación:** Responda la pregunta en el espacio asignado.

1. ¿Cuáles son las especificaciones técnicas del Hardware con las que se cuenta para el desarrollo del proyecto?

Servidores	Usuarios	Desarrollo
Base de datos: - Procesador: Xeon 3.16 Ghz, 3.17 Ghz - RAM: 1 GB - Disco Duro: 30 GB y 175 GB	Procesador: Pentium IV 3.06 Ghz RAM: 512 Mb Disco Duro: 200 Gb	Procesador: Pentium IV 3.06 Ghz RAM: 1.5 Gb Disco Duro: 200 Gb

2. ¿Cuáles son las especificaciones técnicas del Software con las que se cuenta para el desarrollo del proyecto?

Software	Servidores	Usuarios	Desarrollo
Sistema operativo	Windows 2003	Windows XP	Windows XP
Base de datos	SQL Server 2005 Enterprise		SQL Server 2005 Express
Antivirus	MCafee VirusScan 8.7	MCafee VirusScan 8.7	MCafee VirusScan 8.7
Firewall	Alteon 2208		
Herr. De Diseño	Reporting Services		Photoshop
Herr. De Desarrollo			Visual Studio .net 2005 Crystal Reports 10

3. ¿A partir de qué fecha se tiene registro histórico de denuncias recibidas? (según cada modulo)

Nombre Modulo	Fecha de inicio	Nombre Modulo	Fecha de inicio
SIPRODEN	2004	Sistema de Atención a Reclamos y Denuncias	15/11/2007
SIDECON	10/10/2005		

4. ¿Se utiliza algún tipo de estándar en el área informática para el desarrollo de sistemas?

Para el desarrollo de software se ha implementado el siguiente estándar:

Desarrollo:

Se ha implementado el desarrollo de 3 capas, a continuación se describe cada una de las capas:

Capa de presentación	Componentes	dbHelpers
wfNombrePagina.aspx	cNombreComponente.vb	dbNombre.vb

En la base de datos se ha implementado el siguiente estandar:

Tablas	Vistas	Procedimiento almacenado
SARA_NombreTabla	V_NombreVista	SP_NombreProcedimiento

5. ¿Cuál es la experiencia técnica del recurso humano del área operativa?

Todos los usuarios cumplen con el aspecto técnico necesario para el manejo del sistema actual, es por ello que se les brinda una pequeña capacitación para el uso del sistema

6. ¿Se cuenta con herramientas para el apoyo del diseño de elementos de Inteligencia de Negocios?

- Herramientas de SQL Server 2005
- Analysis Services
- Integration Services
- Reporting Services
- Data Mining

7. ¿Cómo se encuentran clasificados los reportes?

Actualmente la clasificación de reportes se ha establecido por perfil de usuarios, a cada perfil se ha habilitado los reportes necesarios para cada puesto.

8. ¿Cuántos reportes utiliza en promedio los usuarios (de cada división)?

A continuación se detalla por perfil la cantidad de reportes disponibles por perfil:

Técnico: 2 reportes  
 Coordinador: 44 reportes  
 Gerente: 44 reportes  
 Estudios: 39 reportes

Se cuenta con un generador de reportes que permite a los usuarios obtener información utilizando algunas variables para realizar el reportes (es muy limitado)

9. ¿Le gustaría que se utilizara un sistema de información que facilite la consulta de información de los reportes actuales y otra información sin depender de la gerencia de sistemas?

SI  X  NO:  \_\_\_

10. ¿Cuál es la importancia de la creación de un sistema que facilite la consulta de información para la gerencia de sistemas?

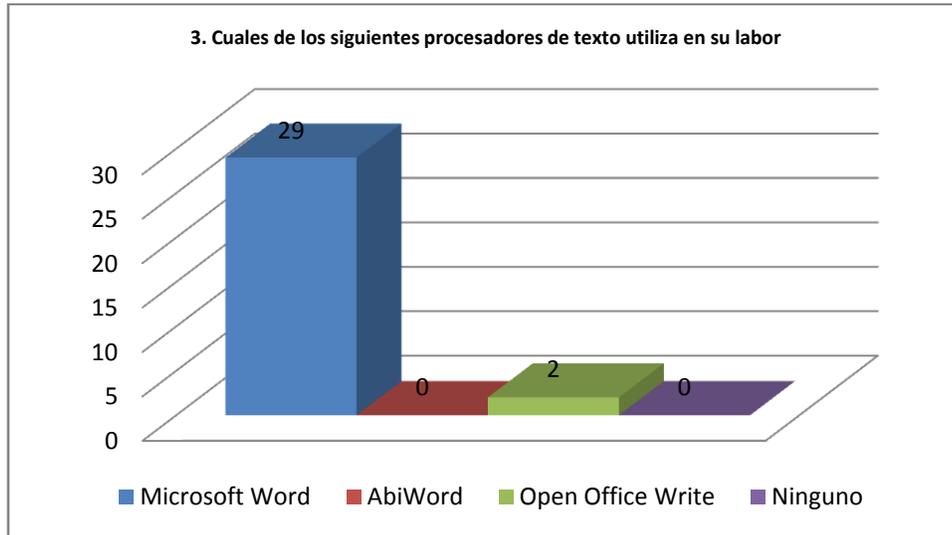
Es importante porque permitirá a los usuarios realizar reportes variados y de acuerdo a las necesidades de cada área, además facilitará la labor en el área de desarrollo puesto que no se dedicará tiempo planificado para otras actividades en hacer reportes

11. ¿Cuántas denuncias se reciben mensualmente en promedio?  
1440 aprox
12. ¿Cuántos usuarios se tienen que utilizan el sistema actual?  
70 usuarios
13. ¿Cuántos usuarios se forman parte de la gerencia de sistemas?  
7 usuarios
14. ¿En promedio cuantos equipos de computo se conectan para utilizar el sistema actual?  
77 equipos
15. ¿Qué tipo de Información considera que seria provechoso considerar para agilizar los procesos de las denuncias?
  - Agilizar el trámite de las denuncias
  - Reportar atrasos en seguimiento de casos
  - Analizar tendencias de motivos de denuncias

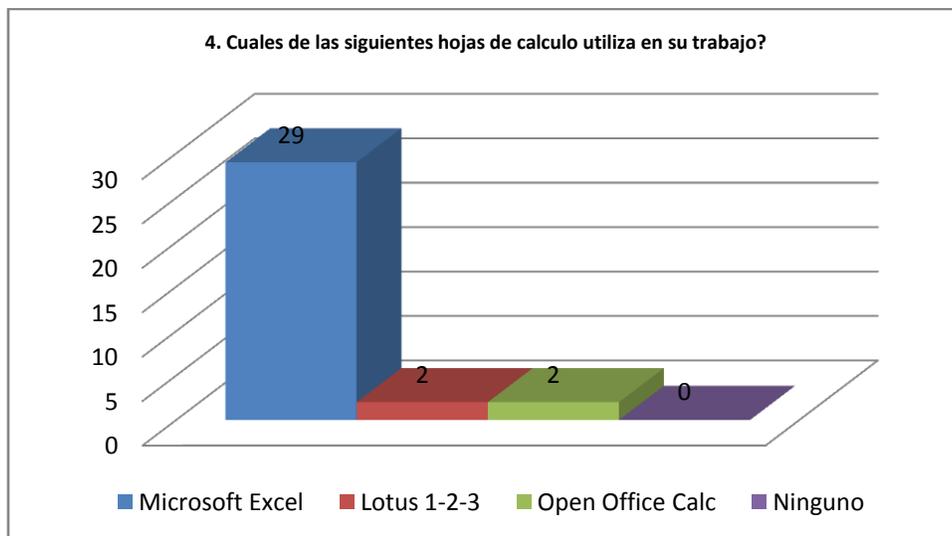
### 4.3 ENCUESTA DE CONOCIMIENTO TÉCNICO OPERATIVO

Encuesta para determinar el grado de conocimiento de herramientas WEB y ofimáticas en los miembros de la Defensoría del Consumidor.

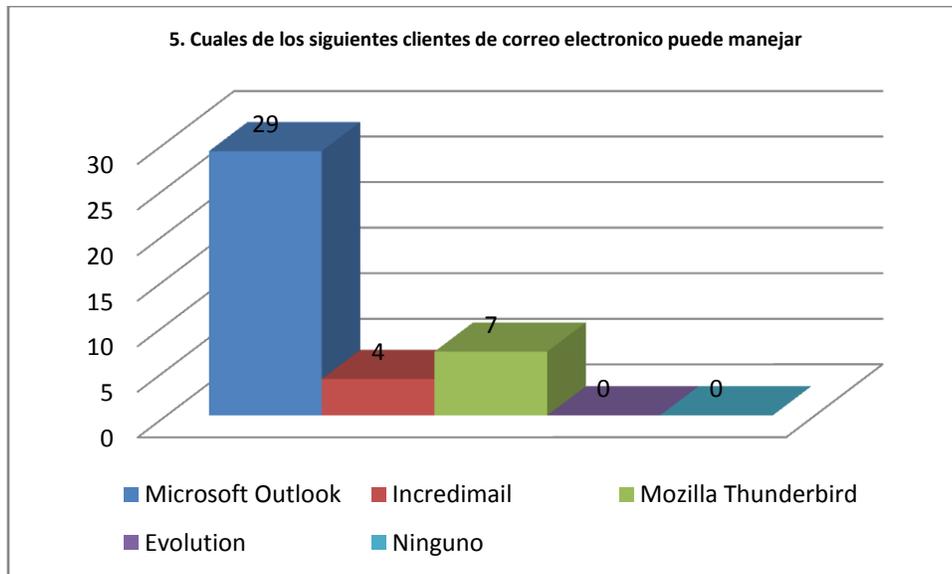
Personas de Nivel Estratégico:	<b>8</b>
Personas de Nivel Táctico:	<b>5</b>
Personas de Nivel Operativo:	<b>16</b>



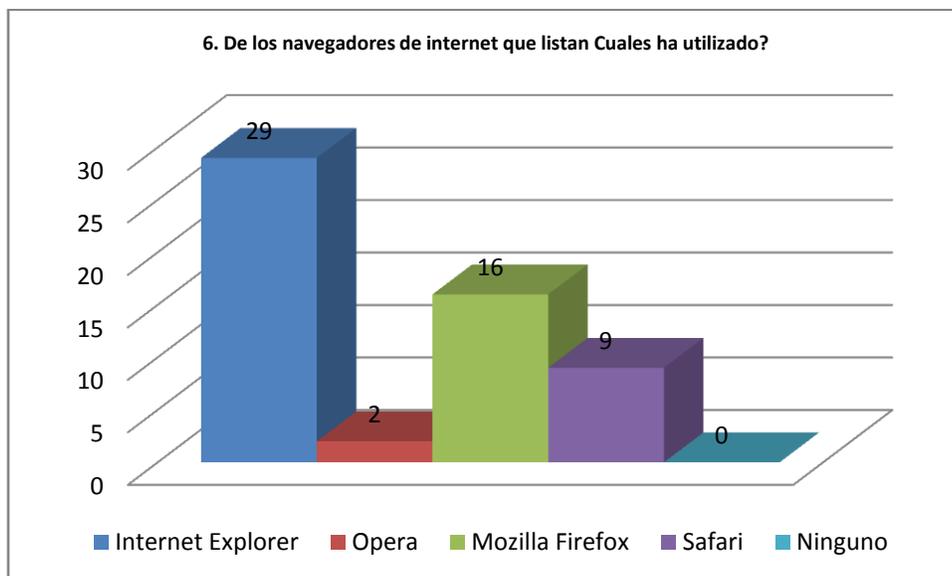
Todas las personas de la organización se encuentran familiarizadas con el uso de procesadores de texto. Los resultados de la encuesta muestran que un 100% puede utilizar Microsoft Word, mientras que un 6.89% utiliza dos procesadores de texto para la realización de sus labores.



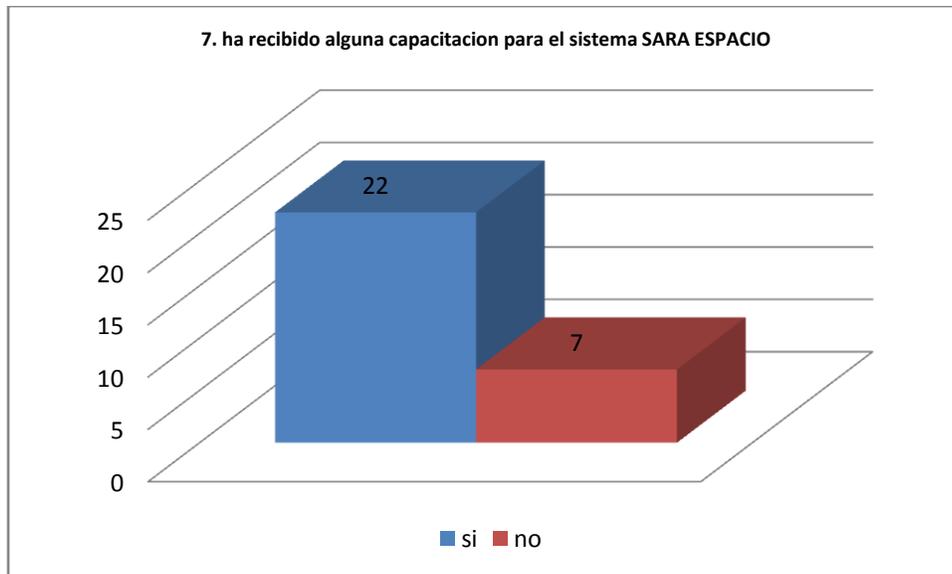
El 100% de los encuestados utiliza hojas de cálculo para la realización de sus labores. La más utilizada dentro de la organización es Microsoft Excel, mientras que otras herramientas como Open Office Calc y Lotus son utilizados por un 6.89 %.



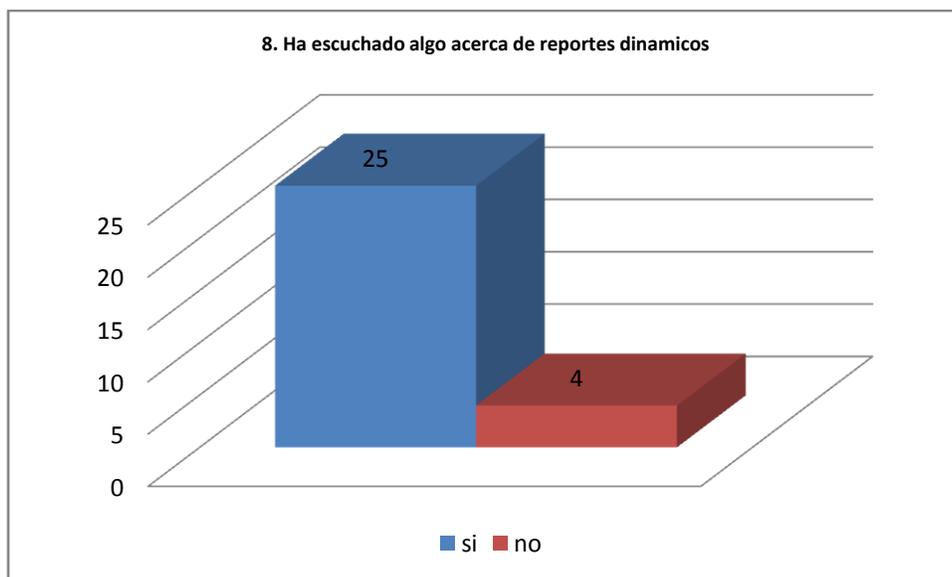
De las personas encuestadas el 100% maneja clientes de correo electrónico, siendo Microsoft Outlook el más utilizado.



Todas las personas encuestadas se encuentran familiarizadas con el uso de navegadores web, el más utilizado es Internet Explorer, teniendo una aceptación del 100 % de los usuarios. El manejo de un navegador web es importante para la realización de proyecto, ya que esta es la herramienta que se utiliza para acceder al los datos proporcionados por el Data Warehouse.



El 75.86% ha recibido capacitación por parte de la organización para el manejo de los sistemas de información que esta posee, por lo que se puede concluir que la Defensoría del Consumidor se preocupar por capacitar al sus usuarios en el manejo de los sistemas de información que esta posee.



El 86% de los encuestados ha escuchado y conoce un poco sobre los reportes dinamicos u OLAP que permiten un mejor análisis de la información, lo que facilitaría en cierta forma la implantación del proyecto.

## ANEXO 5: Evaluaciones Técnicas

### 5.1 CRITERIOS PARA ESCOGER EL SISTEMA OPERATIVO<sup>52</sup>

- **Fiabilidad:**  
La fiabilidad es una medida de que tan bien corre la computadora las aplicaciones y los servicios. La integridad de los datos en memoria, así como también la escritura en disco.
- **Seguridad:**  
La seguridad es la habilidad del sistema de protegerse de un ataque de denegación de servicios, así como también de intrusiones.
- **Escalabilidad:**  
Numero de procesadores soportados, tamaño máximo de archivos soportado y tipos de sistema de archivo soportados.
- **Manejabilidad:**  
Representa el conjunto de funciones, que automatizan la operación del servidor, así como también las habilidades necesarias del personal responsable de la administración del sistema operativo y el servidor. Cuando se considera un sistema operativo también se debe buscar el manejo del ambiente, facilidad de crecimiento, facilidad de integración con lo diferentes elementos del ambiente y la seguridad.
- **Adaptabilidad:**  
Es la medida de que tan bien el sistema operativo puede interactuar con otros, ya que en nuestros días se cuenta con una gran variedad de ambientes como las aplicaciones en el cliente, terminales, etc.
- **Asequibilidad:**  
Se mide desde dos puntos de vista, precio y soporte. Es importante no ser engañado por un precio bajo si el costo del soporte es demasiado elevado.

<sup>52</sup> [http://irrigation.up.nic.in/MIS/pact/Final\\_Ph0\\_PDF/Appendix%20G\\_OS%20Eval%20Criteria.pdf](http://irrigation.up.nic.in/MIS/pact/Final_Ph0_PDF/Appendix%20G_OS%20Eval%20Criteria.pdf)

## 5.2 CRITERIOS PARA LA SELECCIÓN DEL SISTEMA GESTOR DE BASE DE DATOS

**Costo:**

Costo de la licencia  
Costos de administración y soporte  
Costo de capacitación

**Rendimiento:**

Velocidad

**Fiabilidad:**

Tiempos de baja del servidor.  
Rapidez de respuesta a bugs.

**Escalabilidad:**

Habilidad para expandirse en otros servidores a medida va creciendo la base de datos.

**Seguridad:**

No existen o son pocas las vulnerabilidades conocidas.  
Validación independiente de la seguridad.

**Plataformas:**

Disponibilidad de instalación en diferentes sistemas operativos.  
Soporte Multi-Lenguaje

**Administración:**

Facilidad de Administración

**Madurez:**

Madurez del producto  
Fiabilidad de la versión actual  
Facilidad de Instalación  
Información disponible sobre mejores practicas  
Certificación disponible

**Documentación:**

Existe buena documentación  
Conferencias de Usuarios  
Entrenadores Certificados

**Herramientas Integradas:**

Posee un conjunto de herramientas integradas que permite el manejo de los datos y acceso a estos.

### 5.3 Criterios de evaluación del software

Para la elección del software que utilizaremos para el desarrollo del data warehouse se decidió realizar una investigación comparando las herramientas que brindaban soporte a la base transaccional que se maneja actualmente, para tomar tan importante elección se realizó una matriz de evaluación de herramienta de desarrollo para sistemas de inteligencia de negocios. Dicha matriz se encuentra estructurada en 4 grandes áreas las cuales se consideran que son imprescindibles para cumplir con la solución propuesta, estas áreas son las siguientes

- Características de almacenamiento
- Control y programación de trabajos
- Estándares y Metadatos
- Características de los vendedores

Todos los requisitos deben tener prioridad, tanto los de la organización como los técnicos, esta prioridad permite asignarle factores de peso a los diferentes requerimientos por lo que se asignan valores numéricos típicamente entre 1 y 5, en caso de tener un empate se reasignan los pesos por cada factor.

#### Características de almacenamiento

Este área de evaluación se enfoca en la capacidad que tiene la herramienta para poder manejar las fuentes de datos, considerando la compatibilidad, y funciones principales que utilizaremos para la extracción de datos.

#### Control y programación de trabajos

Los criterios incluidos en esta área están enfocados al soporte que posee la herramienta para facilitar las tareas de rutina como lo son la carga de datos a las tablas de hechos, realizar copias de seguridad, el monitoreo y recuperación ante errores.

#### Estándares y metadatos

Esta área permite evaluar la herramienta a fin de verificar la compatibilidad que se posee con otras herramientas de desarrollo. Es por ello que se evalúa el intercambio de metadata con otros productos, y los medios de comunicación que utiliza para realizar tareas de mensajería.

#### Características de los Vendedores

En la mayoría de los casos son llamados Partner, y son los encargados de la distribución y venta de las herramientas. Este aspecto es muy importante ya que el producto a elegir no es un producto independiente, por lo que se debe tener el mayor apoyo de por parte de los distribuidores en cuanto al soporte técnico. En esta área se evalúa un aspecto que en la mayoría de organizaciones es el principal criterio de elección, el cual es el costo de la herramienta para el desarrollo.

**5.3.1 Matriz de evaluación de productos**

Matriz de evaluación de productos			
CARACTERÍSTICAS	PESO	PENTAHO Suit BI	SQLServer BI suit
<b>CARACTERÍSTICAS BÁSICAS DE ALMACENAMIENTO</b>			
Soporta acceso a múltiples fuentes de información como plataformas de warehouse , Bases de datos relacional, IDMS,WKS, DB2, VSAM	85	5	3
Provee rapidez en las copias de seguridad y replicaciones	85	4	5
Permite la Compresión y descompresión de datos	50	2	4
Soporta funciones de transformación de validación, integración, verificación de integridad, calculo, asignación.	85	3	5
Permite funciones de carga de datos vía reemplazo, inserción , actualización en RDMS.	85	4	4
<b>CONTROL Y PROGRAMACIÓN DE TRABAJOS</b>			
Soporta tareas programadas basadas en tiempo y eventos (JOB) .	75	4	4
Soporta monitoreo, recuperación informes de tareas incompletas.	75	2	4
Soporta almacenamiento y restauración.	25		5
Soporta controles de seguridad para el acceso y cifrado.	25	2	5
Permite sincronización de actualizaciones en línea	25	2	4
<b>ESTÁNDARES Y METADATOS</b>			
Soporta el intercambio de metadato con otros productos.	85	3	3
Soporta estándares de mensajería, incluyendo COM, DCE, COBRA.	25	3	4
Soporta niveles de transporte incluyendo TCP/IP , FTP.	25	4	4
<b>Características de los vendedores</b>			
Costo.	85	5	2
Soporte Técnico.	85	2	5
Documentación.	50	2	5
Disponibilidad y calidad de las Capacitaciones.	25	2	5
Disponibilidad y calidad de las Consultorías.	50	2	4
<b>RESULTADO TOTAL (basado en 5225 para escalas de 1 a 5)</b>		<b>3285</b>	<b>4220</b>

**ANEXO 6: Interfaz de Ingreso de Datos**



# Defensoría del consumidor



Boletín Informativo

---

Inicio | ¿Quiénes Somos? | Prensa | Documentos | **Área del Consumidor** | Calcule su Presupuesto | Estudios y Verificación

**Consultas y Denuncias**

Sondeo de Precios

Denuncia en Línea

Consumidor Informado

Oficinas Regionales

Su Opinión Cuenta

Consumidor Joven

Cartas de Derechos

## Denuncias en línea

**Datos personales:**

Nombre consumidor: *	<input type="text"/>
Apellidos consumidor: *	<input type="text"/>
Sexo: *	<input type="text"/>
Número DUI: *	<input type="text"/>
Fecha de nacimiento: *	<input type="text"/> / <input type="text"/> / 19 <input type="text"/>
Profesión: *	<input type="text"/>
Dirección: *	<input style="height: 40px;" type="text"/>
Municipio: *	<input type="text"/>
Departamento: *	<input type="text"/>
Email: *	<input type="text"/>
Teléfono: *	<input type="text"/>
Fax:	<input type="text"/>
Celular:	<input type="text"/>

Sondeos

Estudios

Inspecciones

**Datos del proveedor:**

Núm. comprobante:	<input type="text"/>
Sector proveedor:	<input type="text"/>
Proveedor: *	<input type="text"/>
Motivo de denuncia:	<input type="text"/>
Monto reclamado:	<input type="text"/>
Número de cuenta con el proveedor (si lo hay):	<input type="text"/>

**Denuncia:**

¿Qué reclama? ¿Cuál es el acto o hecho que denuncia? \*

¿Dónde ocurrió?

¿Cuándo ocurrió?

¿Qué le había ofrecido el proveedor en el momento de la compra? (si aplica)

Si ya presentó su reclamo ante el proveedor, ¿cuál fue la respuesta?

¿Cuál es su pretensión? ¿Qué es lo que desea? \*

**Documentos que aporta:**

- |  |   |
|--|---|
| <input type="checkbox"/> Factura compra    | <input type="checkbox"/> Comprobante depósito |
| <input type="checkbox"/> Factura servicios | <input type="checkbox"/> Comprobante pago     |
| <input type="checkbox"/> Factura consumo   | <input type="checkbox"/> Copia contrato       |
| <input type="checkbox"/> Detalle consumo   | <input type="checkbox"/> Copia garantía       |

**Otros documentos: (especificar)**

\* campos requeridos


<b>Oficinas Administrativas</b> Calle Circunvalación # 20, Plan de La Laguna, Antiguo Cuscatlán. El Salvador Tel: 2526-9000

## ANEXO 7: Descripción del Modelo UML de los Componentes Bodega de Datos

Este componente de proceso de inteligencia de negocios, el cual se destacan Bill Inmon y Ralph Kimball. El primero lo define: “Una Bodega de Dato es una colección de datos integrados orientados a temas, integrados, no-volátiles y variables en el tiempo, organizados para soportar necesidades empresariales” y para Kimball “la Bodega de Datos es un colección de datos en forma de una base de datos que guarda y ordena información que se extrae directamente de los sistemas operacionales (ventas, producción, finanzas, marketing,etc.) y de datos externos”. Ambos punto de vista han tenido amplia aceptación, pero el más utilizado es el de Kimball.

### 7.1 Componente de una bodega de datos

En la Figura 7.1 se puede observar los componentes de una bodega de datos dados por Kimbal.

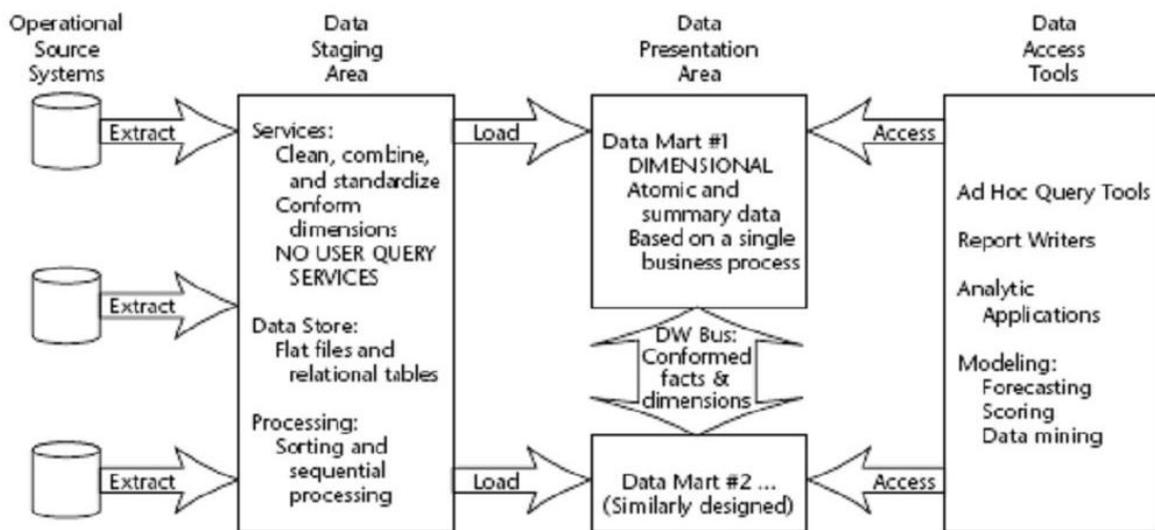


Figura 0.1: Componentes básicos de una Bodega de Datos, Kimball, R. & Ross, M. “The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling”

#### Sistemas fuente operaciones:

Son los que guardan las transacciones del negocio, las cuales son la información vital para los procesos que se llevan a cargo en la organización. Las características de estos sistemas son desempeño del procesamiento y la disponibilidad.

#### Área de depuración de datos:

Es la zona donde se encuentra cuales de los datos que forman parte en la parte transaccional se pueden usar en la bodega de datos, Es esta parte aparece ETL proceso de extracción, carga y transformación de los datos, este proceso se lleva a cabo por medio de diversas herramientas las cuales permite tener integrados muchos datos de diferentes Fuentes.

#### Área de presentación de datos:

Esta área se pone los datos en espacio de almacenamiento de la bodega de datos, los datos son almacenados en las dimensiones y en las tablas de hecho donde se pueden formar diversos datamart .

### Herramienta de acceso de datos:

Es la parte donde el usuario puede tomar y realizar las diversas consultas para formar los famosos cubos dimensionales, los cuales pueden ser usados para la generación de reportes e indicadores cuando ocurre un proceso OLAP, o se le puede aplicar un proceso de minería de datos para descubrir conocimiento, el cual la organización no conocía.

## 7.2 Niveles para el modelamiento de los datos

“El modelamiento de los datos es una técnica que almacena los inventarios, formas, tamaños, contenidos, y reglas de los elementos de los datos usados en el alcance en proceso del negocio”. El resultado del modelamiento de los datos es un tipo de mapa que se describe en uso del dato en el proceso.

Tradicionalmente, se encuentran tres niveles para el modelamiento de los datos en las bodegas de datos que son: conceptual, lógico y el físico.

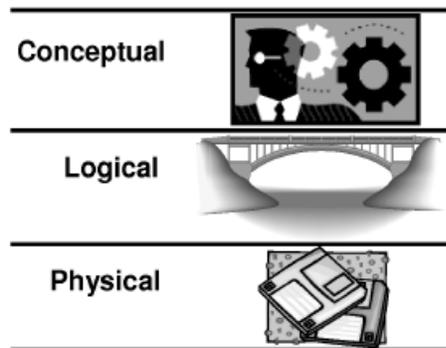


Figura 0.2: Niveles Para el modelamiento de los datos.

### 7.2.1 Modelos Conceptuales

En la actualidad se encuentran vigentes varios modelos de los cuales se pueden destacar:

Modelo Hecho-Dimensional (Golfarelli)

Modelo Multidimensional/ER (Sapia)

Modelo Estrella/ER (Tryfona)

Modelo GOLD (Trujillo)

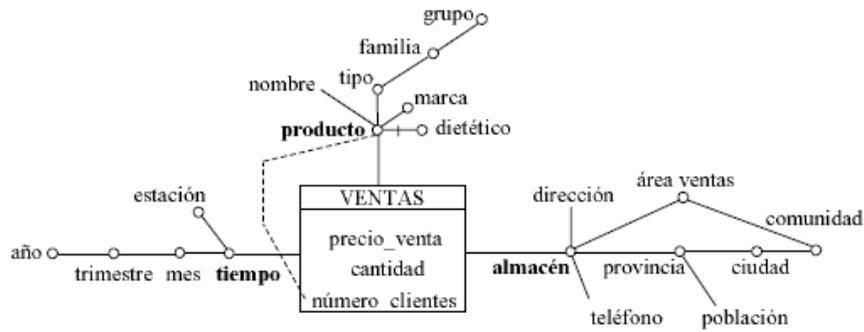
Modelo Husemann

Modelo YAM2

Para ver cada modelo se revisaran sus partes estructurales (Hechos, dimensiones) y dinámicas y su notación gráfica

#### Modelo Hecho-Dimensional

Es un modelo donde en las tablas de hechos permite la cardinalidad de muchos a uno con las dimensiones, sus atributos son atómicos. En las dimensiones se representan como un grafo dirigido acíclico dirigido, permite jerarquías múltiples y caminos alternos, asume cardinalidad muchos a uno entre niveles de jerarquías. En la Figura 7.3 se puede ver la notación gráfica.



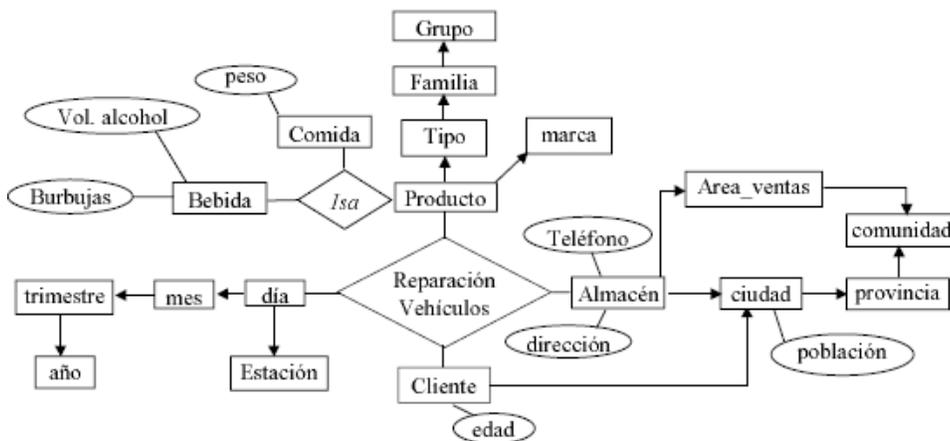
**Figura 0.3: Representación gráfica modelo Hecho dimensional**

En su parte dinámica hace una representación formal de de los requerimientos iniciales de usuario, no propone operaciones OLAP para aplicar sobre los requerimientos iniciales de usuarios y no modela el comportamiento de los requerimientos. No permite filtros definidos.

En su parte lógica es una transformación al esquema estrella de Kimball, a demás este esquema no se puede ejecutar directamente sobre una aplicación OLAP.

**Modelo Multidimensional /ER**

En este modelo se define la entidad “fact” para representar los hechos, cardinalidad muchos a uno con cada dimensión en particular, los atributos son atómicos. Define la entidad “Level” para representar los niveles de jerarquía, Define la relación “roll-up” para representar una relación binaria entre dos entidades “level”. Estas entidades (levels) deben formar un grafo acíclico dirigido (G.A.D). Permite jerarquías múltiples y caminos alternativos. Permite categorización de las dimensiones relaciones “is-a” propias del modelo ER. En la Figura 7.4 se puede ver su notación gráfica



**Figura 0.4: Representación gráfica modelo multidimensional / ER**

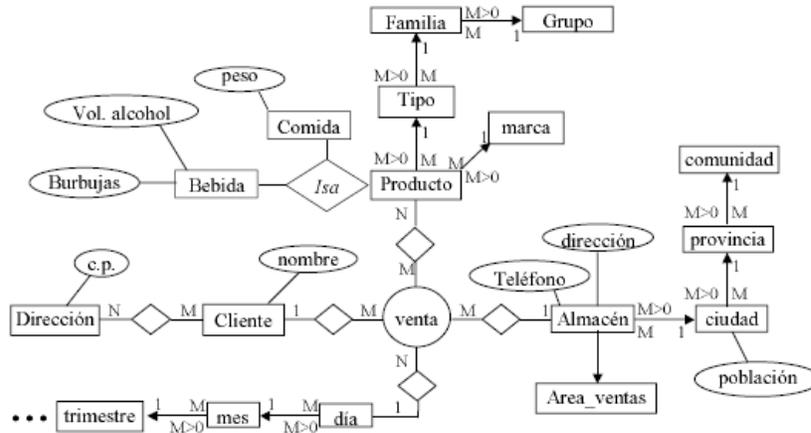
En su parte dinámica presenta este modelo un patrón de requerimiento como las medidas, dimensiones de filtros, y dimensiones de condición de agrupamiento, los filtros se especifican en valores exactos, permite operaciones OLAP (roll-up, drill - down, rotate, Focus change).

En el esquema lógico puede hacer transformaciones en los sistemas ROLAP y MOLAP, se permite definir el esquema lógico en herramientas OLAP comerciales.

**Modelo Estrella/ER**

En la tabla de hechos se define como una relación de asociación que representa una relación muchos a muchos entre con todas las dimensiones, permite representar la cardinalidad muchos a muchos con cada dimensión en particular, no se puede definir atributos derivados. Define una relación de asociación para relacionar los distintos niveles de jerarquía, permite categorización de dimensiones con relaciones "is-a", utilice las relaciones del modelo ER como la agregación par representar otras relaciones entre entidades además de las clásicas jerarquías de clasificación.

En la Figura 7.5 se observa la forma de representación de este modelo



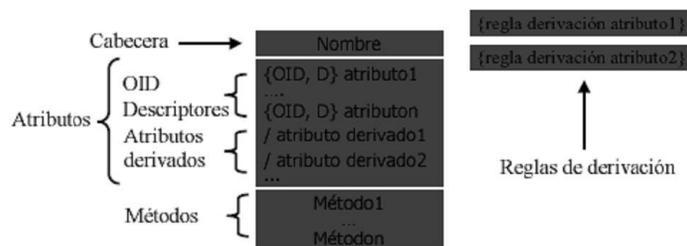
**Figura 0.5: Representación gráfica modelo estrella /ER**

No proporciona operaciones OLAP, no hay ningún mecanismo para la representación inicial de usuarios. Es fácil de generar un modelo estrella a partir de este modelo.

**Modelo GOLD**

Este modelo se soporta sobre una estandarización de UML, define las tablas de hechos y de dimensiones como clases y sus relaciones como asociaciones y agregaciones. Permite realizar las OLAP (roll-up, drill - down, combine, divide, rotate) y ROLAP, en su parte dinámica se puede modelar por medios de diagram de estados, interacciones. En la parte lógica permite generar información estructural, generación de información dinámica y transformaciones para sistemas ROLAP y MOLAP.

Su notación grafica se basa en UML en las Figura 7.6 y Figura 7.7 se puede ver.



**Figura 0.6: Representación gráfica modelo GOLD de una clase**

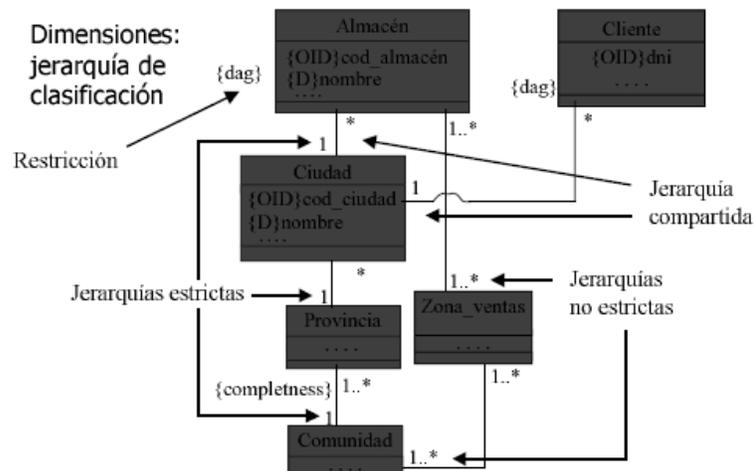


Figura 0.7: Representación gráfica modelo GOLD Dimensiones con Jerarquías.

**Modelo Husemann**

Este modelo presenta su parte de hechos atributos atómicos, en la parte de dimensiones proporciona jerarquía y categorización en las dimensiones, no presenta una parte dinámica ni lógica por lo tanto no permite operaciones OLAP, pero posee una notación gráfica que se puede ver en la Figura 7.8.

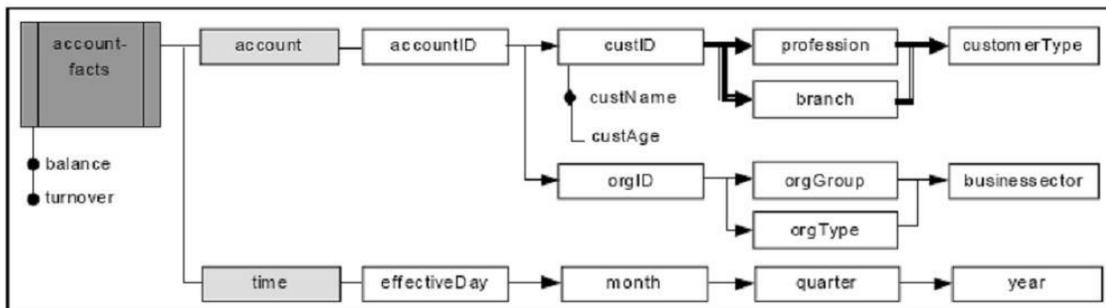


Figura 0.8: Representación gráfica modelo Hussemann.

**Modelo YAM 2**

Este modelo consiste en mostrar varios datamart en forma de estrellas en los diagramas de clases (UML), este modelo se puede definir como una extensión del UML, la estructura de este modelo se pueden definir en metaclasses que son una especialización de la metaclasses del UML Finalmente se muestra un conjunto de operaciones algebraica en los cubos que se puede definir en términos operaciones basadas en funciones matemáticas.

Permite las relaciones muchos a muchos con dimensiones en particular permite el manejo de multi estrellas, medidas atómicas y derivadas, pueden compartir dimensiones permite clasificación por jerarquía y categorización por dimensiones, en la parte dinámica permite OLAP. No está implementado en ninguna herramienta comercial OLAP.

Su notación gráfica se puede ver en la Figura 7.9.



- Verificar todas las diferentes parte en la bodega de datos (Dimensiones, hechos, proceso ETL, herramientas OLAP, etc)

Verificado lo anterior ya se puede pensar en una Buena implantación de la bodega de datos, revisando sobre que plataforma se va instalar, como será el óptimo de ubicación de los datos (Bases de datos distribuidas, centralizadas), distribución de información en varios disco (RAID), que motor de bases de datos se va usar, sistemas operativos, híbridos.

#### 7.2.4 COMPONENTE MINERÍA DE DATOS

La minería de datos ofrece un componente importante en la inteligencia de negocios. La minería de datos consiste en técnica enmarcada en conseguir conocimiento partiendo de un conjunto de datos a los cuales se le aplica unos algoritmo para sacar patrones o tendencia nuevas las cuales los dueños de esos datos no tenían conocimiento.

En la minería de datos se reúnen varias áreas como son la estadística, la inteligencia artificial, la computación grafica, las bases de datos,, y el procesamiento masivo.

Una definición forma es dada por Fayyad “Un proceso no trivial de identificación válida, novedosa, potencialmente útil y entendible de patrones comprensibles que se encuentran ocultos en los datos.”(1996) y desde el punto de vista empresaria lo define Molina como “La integración de un conjunto de áreas que tienen como propósito la identificación de un conocimiento obtenido a partir de las bases de datos que aporten un sesgo hacia la toma de decisión” (2001).

La minería de datos es una técnica compuesta por etapas que integran varias áreas.

A. Principales características y objetivos en la minería de datos:

- Explorar los datos se encuentran en las bases de datos y bodega de datos.
- En el proceso se encuentran a menudo descubrimientos de resultados valiosos e inesperados.
- Debido a la gran cantidad de datos se pueden usar procesamientos paralelos.
- Se producen cinco tipos de información: Asociaciones, secuencias, clasificaciones, agrupamientos y pronósticos.

B. Facetas para el desarrollo de un proceso de minería de datos:

Para el proceso de minería de datos aplican cuatro facetas

- Filtrado de datos
- Selección de variables
- Extracción de conocimiento
- Interpretación y evaluación.





## ANEXO 9: Minería de Datos

### 9.1 ¿Qué es la minería de datos?

La minería de datos es un miembro clave en la familia de los productos de Inteligencia de negocios, bajo este concepto se engloba todo un conjunto de técnicas encaminada a la extracción de conocimiento procesable, implícito en las bases de datos; según la el concepto publicado en wikipedia, la minería de datos (DM, Data Mining) consiste en la extracción no trivial de información que reside de manera implícita en los datos. Dicha información era previamente desconocida y podrá resultar útil para algún proceso. En otras palabras, la minería de datos prepara, sondea y explora los datos para sacar la información oculta en ellos.

Durante la década pasada grandes volúmenes de información eran acumulados y almacenados en la bases de datos, muchos de estos datos provenían de software propios del negocio como aplicaciones contables, sistemas de planificación de recursos empresariales (ERP: Enterprise Resource Management), Sistemas de gestión de las relaciones con los clientes (CRM: Customer Relationship Management). El resultado de este almacenamiento de datos es que las organizaciones poseen gran cantidad de datos y se obtiene poco conocimiento de estos, el objetivo principal de la minería de datos es la extracción de patrones a partir de datos que se poseen a fin de aumentar el valor intrínseco y crear conocimientos en base a los datos. Para entender mejor este concepto se puede ejemplificar tomando como base la figura siguiente:

Gender	ParentIncom	IQ	ParentEncouragement	CollegePlans
Male	46580	100	Not Encouraged	No
Male	39687	121	Not Encouraged	No
Male	63482	102	Encouraged	Yes
Female	40454	129	Not Encouraged	No
Male	7333	86	Not Encouraged	No
Female	17617	105	Not Encouraged	No
Male	33540	110	Not Encouraged	No
Male	48171	102	Not Encouraged	Yes
Male	33356	79	Not Encouraged	No
Male	73325	120	Encouraged	Yes
Male	33153	112	Not Encouraged	No
Male	10331	94	Not Encouraged	No
Female	33505	106	Not Encouraged	Yes
Female	30052	76	Encouraged	Yes
Male	24579	105	Not Encouraged	No
Male	37497	72	Not Encouraged	No
Male	31572	98	Encouraged	No
Female	41979	138	Not Encouraged	No
Female	11151	61	Not Encouraged	No
Female	9532	86	Encouraged	No
Male	73580	124	Encouraged	Yes
Female	70149	104	Encouraged	No
Male	44318	122	Encouraged	Yes
Male	14915	100	Not Encouraged	No
Male	52417	68	Encouraged	No

Figura 9.1: Listado de graduados de la escuela secundaria.

La figura anterior muestra una tabla relacionan que contiene un listado de graduados de la escuela secundaria, dicha tabla almacena registro con los detalles del genero, el coeficiente intelectual (IQ), el nivel de estimulo de los padres, el ingreso de los padre y la intención del estudiante en continuar superándose. Con esta información se podrían responder algunas preguntas tales como ¿Qué impulsa a los graduados de secundaria a ir a la universidad?, para responder esta inquietud, podríamos escribir una

consulta SQL, para conocer cuántos estudiantes del sexo masculino versus el sexo femenino asistirán a la universidad. También se puede escribir otras consultas SQL para determinar el impacto que tiene el estímulo de los padres basándose en la columna de dicha tabla, o la cantidad de estudiantes del sexo femenino que no son estimuladas por los padres. Para ello tendríamos que escribir cientos de estas consultas para poder cubrir todas las combinaciones posibles.

En contraste para el enfoque de la minería de datos estas preguntas son muy simples de responder. Todo lo que se necesita hacer es seleccionar el algoritmo y especificar el uso de la columna, es decir las columnas de entrada y las de predicción (que son el objetivo de este análisis). Un modelo de árbol de decisión puede trabajar correctamente para determinar la importancia del estímulo de los padres en la decisión de continuar con la universidad de los estudiantes. Para ello se puede seleccionar las columnas IQ, genero, ingreso de los padres y estímulo de los padres como columnas de entrada y dejando la columna de Continuar con la universidad como la columna a predecir.

El algoritmo de árbol de decisión analiza los datos y analiza el impacto de cada atributo de entrada con respecto al objetivo y se selecciona el más importante para realizar la división del atributo. Cada división separa un conjunto de datos en dos subconjuntos de modo que la distribución del valor de los planes de ir a la Universidad es tan diferente como sea posible entre estos dos subconjuntos. Este proceso se repite recursivamente en cada subconjunto hasta que el árbol esta completamente distribuido. Una vez que el proceso de formación es completado, se pueden ver los patrones que se descubrieron navegando por el árbol.

La siguiente figura muestra un árbol de decisión basado en el conjunto de datos de los graduados de la escuela secundaria. Cada ruta de acceso desde el nodo raíz a un nodo hoja constituye una regla. Ahora podemos decir que los estudiantes con un coeficiente intelectual (IQ) superior a los 100 y que son alentados por los sus padres tienen una probabilidad del 94% de asistir a la universidad. Este es un gran ejemplo de cómo podemos extraer conocimiento a partir de los datos

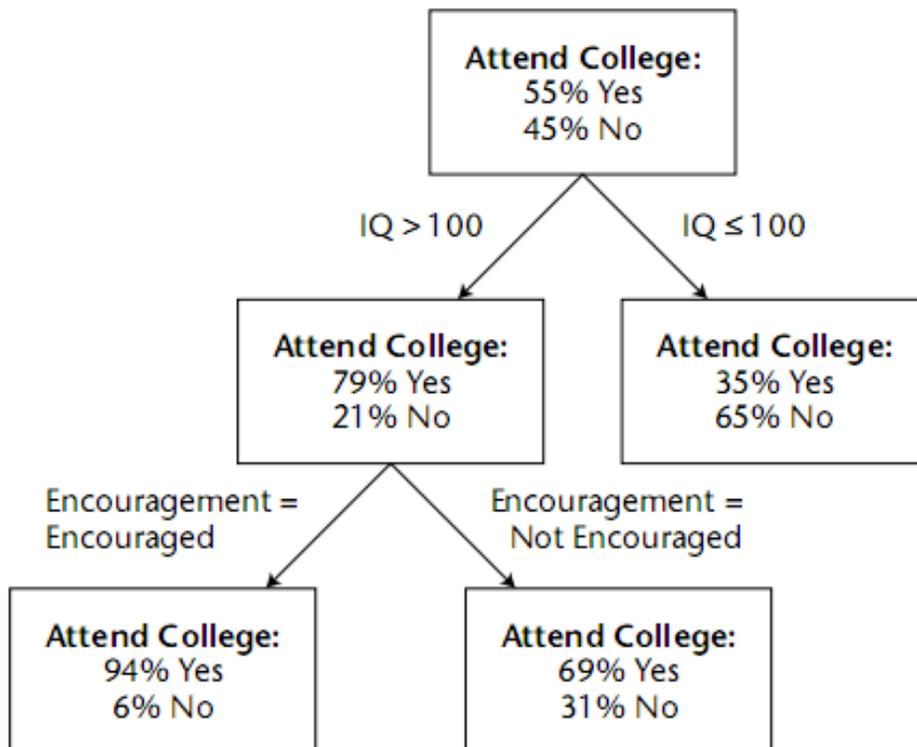


Figura 9.2: Árbol de decisión basado en los datos de recién graduados.

La minería de datos proporciona gran mucho valor agregado a las empresas, pero ¿porque hasta ahora estamos tan interesados en la minería de datos? La respuesta puede fundamentarse tomando en cuenta las siguientes razones.

- Gran cantidad de datos disponibles: en la última década el precio del hardware, especialmente el espacio en disco duro, ha decaído enormemente, a esto se le incluye que las empresas han reunido grandes cantidades de datos a través de diversas aplicaciones, con todos estos datos para explorar, las empresas desean ser capaz de encontrar patrones ocultos para mejorar sus estrategias de negocio.
- Aumento de la competencia: actualmente la competencia es muy alta, debido a la comercialización moderna y canales de distribución como internet y las telecomunicaciones, por lo que las empresas se enfrentan a la competencia en todo el mundo, por lo que el éxito empresarial es la capacidad de retener a los clientes existentes y adquirir nuevos. La minería de datos contiene las tecnologías que permiten a las empresas analizar los factores que afectan a estas situaciones.
- Tecnologías preparadas: las tecnologías de minería de datos anteriormente existían solo en el ámbito académico, pero gracias a la madurez que han presentado muchas de estas tecnologías en la actualidad se encuentran listas para sean aplicadas a la industria.

## 9.2 Problemas de las empresas tratados por la minería de datos

Las técnicas de minería de datos pueden tener muchas aplicaciones, respondiendo a varios tipos de preguntas que se generan en los negocios. La siguiente lista ilustra unos cuantos problemas típicos que se solucionan utilizando minería de datos.

- Análisis Rotativo: ¿Qué clientes son más propensos a cambiar a un competidor? Las telecomunicaciones, la banca, los seguros y las industrias se enfrentan a gran cantidad de competencia en estos días. Este análisis puede ayudar a los gerentes de marketing entender la razón de la rotación de clientes, mejorar las relaciones con los clientes y, finalmente, aumentar la lealtad del cliente.
- La venta cruzada: ¿Qué productos son los clientes dispuestos a comprar? La venta de la cruzada es un reto empresarial importante para los minoristas. Muchos minoristas, sobre todo los minoristas en línea, suelen utilizar esta característica para aumentar sus ventas. Por ejemplo, si usted va a las librerías en línea como Amazon.com o BarnesandNoble.com para comprar un libro, puede observar que el sitio Web le da una serie de recomendaciones acerca de libros relacionados. Estas recomendaciones pueden ser derivadas del análisis de minería de datos.
- Detección del fraude: Las compañías de seguros proceso de miles de reclamaciones al día. Es imposible para ellos para investigar cada caso. La minería de datos puede ayudar a identificar los créditos que tienen más probabilidades de ser falso.
- Gestión del riesgo: ¿El préstamo se aprobó para un cliente? Esta es la pregunta más común en el escenario de la banca. Técnicas de minería de datos puede ser utilizadas para marcar el nivel de riesgo del cliente, ayudando al gerente a tomar una decisión adecuada para cada aplicación.
- Previsión de ventas: ¿Cuántos tipos de vinos voy a vender la próxima semana en una tienda? ¿Cuál será el nivel de inventario que en un mes? Técnicas de previsión de la minería de datos pueden ser utilizada para responder a este tipo de preguntas relacionadas con el tiempo.

### 9.3 Tareas de minería de datos

La minería de datos puede ser utilizado para resolver cientos de problemas que surgen para los negocios. Basándose en la naturaleza de estos problemas, podemos agruparlos en las siguientes tareas de minería de datos.

#### 9.3.1 Clasificación

La clasificación es una de las más populares de la minería de datos. Problemas como la rotación de análisis, gestión de riesgos son parte de la clasificación.

La clasificación se refiere a la asignación de casos en categorías basadas en un atributo de predicción. Cada caso contiene un conjunto de atributos, uno de los cuales es el atributo de clase (llamado también atributo de predicción). Para este tipo de tareas se requiere encontrar un modelo que describa el atributo de clase en función de los atributos de entrada. En el conjunto de datos de los graduados de secundaria anteriormente descritos, el atributo de clase es planes universitarios (CollegePLans) con dos estados definidos: Si o no.

Para preparar la tarea de clasificación, se necesitan conocer los tipos de valores de entrada del conjunto de datos definido, que usualmente son valores históricos. Son parte de esta tarea los algoritmos como árbol de decisión, redes neuronales y el de Naïve Bayes.

#### 9.3.2 Clustering o segmentación

Esta tarea es usada para identificar agrupaciones naturales de casos basados en un conjunto de atributos, los casos del mismo grupo poseen más o menos atributos con valores similares. En la figura siguiente se muestra un conjunto de datos muy sencillo, con dos atributos, la edad y los ingresos. Los algoritmos para este tipo de tareas han sido clasificados en tres segmentos basados en los dos atributos mencionados anteriormente. El grupo uno (clúster 1) contiene la población joven con bajos ingresos. El grupo dos (clúster 2) contiene personas de edad media y con altos ingresos. Y el grupo tres (clúster 3) es un grupo que contiene señores con un ingreso relativamente bajo.

Clustering es una tarea de supervisión de minería de datos. No solo el atributo se utiliza para guiar el proceso de formación. Todos los atributos de entrada son tratados por igual. La mayoría de los algoritmos de este tipo construyen el modelo a través de un número de iteraciones y se detendrá hasta que el modelo converja, es decir, cuando se estabilizan las fronteras de estos segmentos.

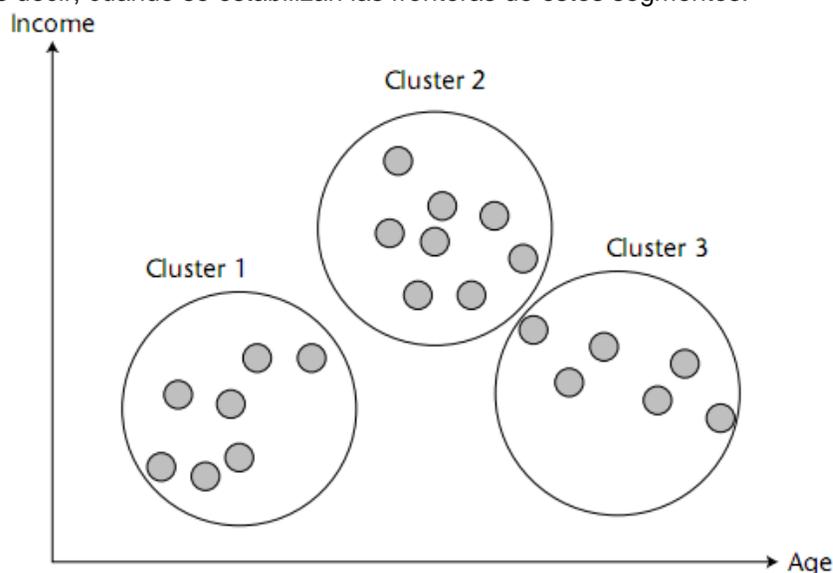


Figura 9.3: Agrupaciones formadas por algoritmos de Clustering.

### 9.3.3 Asociación

Es otra técnica popular de Minería de Datos. La asociación también es conocida como análisis de bolsa de mercado. Un problema común para utilizar la técnica de *Asociación* es analizar una tabla de ventas e identificar aquellos artículos vendidos comúnmente en los carritos de compra de la tienda. El uso común de la asociación es identificar grupos comunes de elementos aplicando reglas con el propósito de analizar combinaciones de ventas

En términos de asociación, cada producto, o más generalmente, cada atributo/pares de valores son considerados un solo elemento. La tarea de la asociación busca 2 fines, encontrar grupos de elementos frecuentes y encontrar reglas de asociación para dichos elementos.

La mayoría de los algoritmos de asociación encuentran los grupos de elementos comunes mediante la exploración de los conjuntos de datos varias veces. El umbral de frecuencia (soporte) es definido por el usuario antes de ejecutar y procesar el modelo. Por ejemplo, soporte de 2% significa que el modelo analiza solo elementos que aparecen en al menos un 2% de los carros de compra. Un conjunto de elementos frecuentes se visualizara como [Producto="Pepsi", Producto="Papas fritas", Producto="Jugo"]. Cada grupo de elementos tiene un tamaño, el cual es el número de elementos que contiene. El tamaño del grupo de elementos anterior es 3.

Ademas de buscar e identificar los grupos de elementos frecuentes basados en un umbral de frecuencia, el algoritmo busca reglas. Una regla de asociación posee la forma  $A, B \Rightarrow C$  con una probabilidad, donde A, B y C son grupos de elementos frecuentes. Asimismo la probabilidad se refiere al nivel de **confianza** en la minería de datos. La probabilidad es un umbral que se debe definir por el usuario antes de la formación del modelo. Por ejemplo, la siguiente es una regla típica para este modelo: Producto="Pepsi", Producto="Papas fritas", Producto="Jugo" con un 80% de probabilidad. La interpretación de dicha regla es sencilla, si un cliente compra Pepsi y papas fritas existe un 80% que también compre jugo. La figura 9.4 muestra patrones de asociación. Cada nodo en la figura representa un producto, cada línea representa una relación, la dirección de la punta de flecha representa la dirección de la predicción. Por ejemplo la relación entre Lecha y Queso indica que aquellos que compran Leche también pueden comprar Queso.

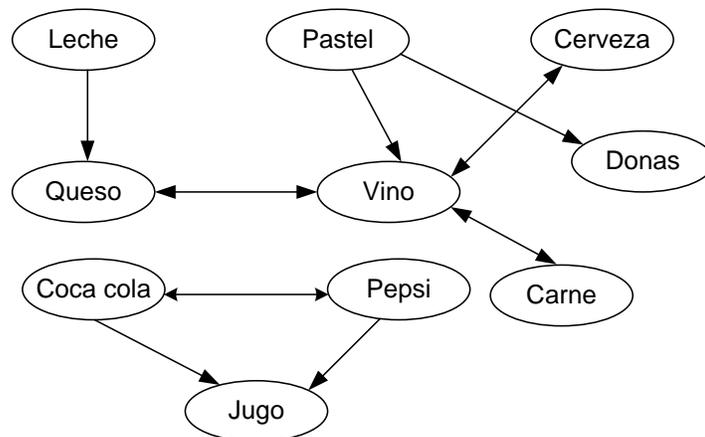


Figura 9.4: Asociación de Productos

### 9.3.4 Regresión

La regresión posee elementos similares a la clasificación. La principal diferencia es que el atributo predecible es un número constante. La técnica de la regresión, ha sido muy estudiada durante siglos en el campo de la estadística. La regresión lineal y la regresión lógica son los métodos más populares. Otras Técnicas de regresión incluyen arboles de decisión y redes neuronales.

La principal tarea de la regresión es resolver múltiples problemas de negocio. Por ejemplo, la regresión puede ser utilizada para predecir tasas de recuperación basados en valor nominal, distribuciones de métodos y volúmenes de distribución, o para predecir velocidad de los vientos basados en temperaturas, presión del aire y humedad.

### 9.3.5 Previsión

La previsión es otra tarea importante en la minería de datos. ¿Cuál será el valor de Microsoft mañana?, ¿Cuál será el ingreso neto de Pepsi para el otro mes?, la previsión puede ayudar a contestar estas preguntas. Por lo general toma como entrada grupos de datos basados en series de tiempo, por ejemplo una secuencia de números con un atributo representan el tiempo. Las series de tiempo generalmente poseen observaciones adyacentes, que son orden-dependientes. La técnica de previsión encaja con la mayoría de tendencias, periodos y filtros que se establecen. El momento más popular de las series de tiempo es el Modelo de medición autoregresivo móvil e integrado.

La figura 9.5 contiene 2 curvas. La línea sólida es la línea de tiempo basado en precio actual en la tienda Microsoft, mientras que la línea punteada es una serie de tiempo basado en mover el promedio utilizando la técnica de previsión.

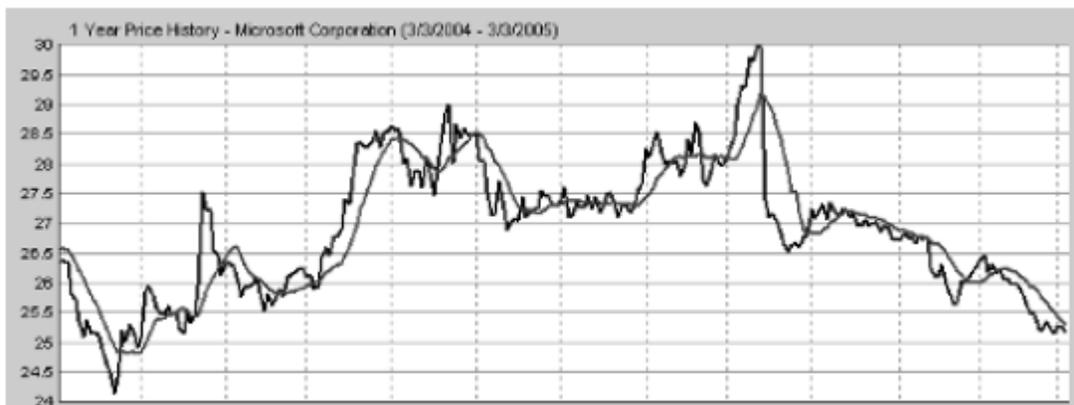


Figura 9.5: previsión moviendo promedio de precio en el tiempo.

### 9.3.6 Análisis de secuencia

Es utilizada para busca patrones es series discretas. Una secuencia está compuesta por una serie de valores discretos (o estados). Por ejemplo, la secuencia del ADN es una larga serie compuesta por cuatro diferentes estados: A, G, C, y T. Un sitio web de hipervínculos contiene una serie de URLs en las que se puede navegar. Las compras de un cliente también pueden ser modeladas como datos secuenciales. Por ejemplo un cliente primero compra una computadora, luego los parlantes, y finalmente una cámara web. Tanto las secuencias como las series de tiempo contienen vistas adyacentes que son dependientes. La diferencia es que las secuencias contienen datos discretos, en cambio las series de tiempo números constantes.

La Secuencias y Asociación son similares en el sentido que cada caso individual contiene un grupo de elementos o estados. La diferencia entre las secuencias y asociaciones, es que las secuencias analizan la transición de los estados, mientras que la asociación considera que cada elemento en una compra sea igual e independiente de otros. Con el modelo de secuencias, comprar una computadora antes de comprar parlantes es una secuencia diferente que comprar parlantes y luego una computadora. Con el algoritmo de asociación se considera que se ha comprado el mismo grupo de elementos.

La Figura 9.6 muestra una secuencia de clics dentro de una WEB. Cada nodo es una categoría en una URL, cada línea posee una dirección, representando la transición entre dos direcciones. Cada transición está asociada a un peso, representando la probabilidad de la transición ente una URL y otra.

El análisis de secuencia es una tarea de minería de datos relativamente nueva. Se ha hecho cada vez más importante principal en el uso de dos tipos de aplicaciones: Análisis de registros WEB y el análisis de ADN. Existen varios análisis de secuencia diferentes hoy en día como lo son las cadenas de Markov, además los investigadores siguen explorando activamente nuevos algoritmos en dicho campo.

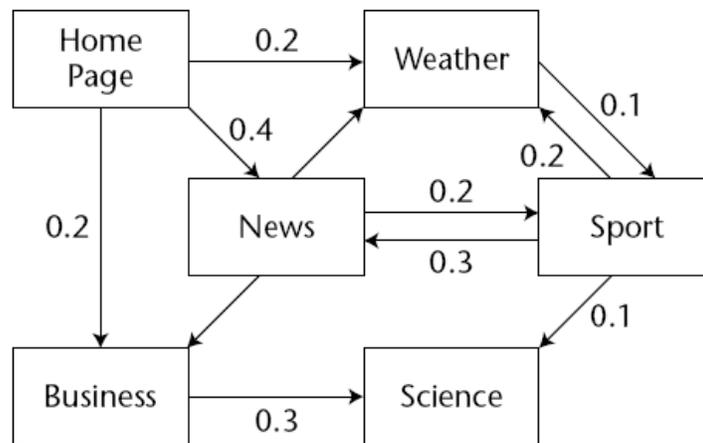


Figura 9.6: análisis de secuencia de navegación de un sitio WEB

### 9.3.7 Análisis de desviación

El análisis de desviación es utilizado para encontrar esos casos raros en el que el comportamiento es diferente a todos los anteriores. También es conocido como detección de valores atípicos, que se refiere a la detección de cambios significativos en comportamientos previamente observados. El análisis de desviación puede ser utilizado en muchas aplicaciones, el caso más común es la detección de fraudes en las tarjetas de crédito. Para identificar casos anormales en millones de transacciones es una tarea muy difícil. Otra aplicación incluye la detección de intrusos en la red, análisis de errores de fabricación y así sucesivamente.

No existe una técnica estándar para la desviación, es un tema de investigación activo en la actualidad. Generalmente los analistas emplean versiones modificadas de arboles de decisión, agrupamiento o algoritmos de redes neuronales para esta tarea. Con el fin de generar normas importantes, los analistas deben sobrepoblar los casos anómalos en el conjunto de datos de muestra.

### 9.4 Técnicas de Minería de Datos

Aunque minería de datos como concepto es relativamente nuevo, la mayoría de las técnicas de minería de datos han existido por años. Si se observan las raíces de los algoritmos más populares para la

minería de datos, se observa que, en su mayoría, son derivados de 3 campos: estadística, aprendizaje automático<sup>53</sup> (machine learning) y bases de datos.

La mayoría de las tareas de minería de datos mencionadas anteriormente, ha surgido en la comunidad de la estadística, como la regresión, series de tiempo y arboles de decisión. La técnica de la regresión ha existido por cientos de años. Los algoritmos de serie de tiempo han sido estudiados por décadas, es uno de las técnicas más recientes que data de mediados de los 80.

La minería de datos se enfoca en el descubrimiento automático o semiautomático de patrones. Muchos de los algoritmos de aprendizaje automático han sido aplicados a la minería de datos. Las redes neuronales es una de esas técnicas y son excelentes para clasificación y regresión, especialmente cuando el atributo de las relaciones es no lineal. El algoritmo genético es otro de los algoritmos de aprendizaje automático, el cual simula el proceso de evolución natural trabajando con un conjunto de candidatos y una función de sobrevivencia. La función de sobrevivencia selecciona rápidamente a los candidatos más idóneos para la próxima generación. Los algoritmos genéticos pueden ser usados para clasificación y tareas de clúster. Además pueden ser utilizados en conjunto con otros algoritmos, por ejemplo, ayudar a una red neuronal a encontrar el mejor set de pesos entre todas las neuronas.

Una base de datos es el tercer recurso para la minería de datos. Las estadísticas tradicionales asumen que todos los datos pueden ser cargados en la memoria para el análisis estadístico, desafortunadamente, este no es siempre el caso en el mundo moderno. Los expertos en bases de datos saben cómo manejar grandes cantidad de datos que no caben en la memoria, por ejemplo, encontrando reglas de asociación en una tabla de hechos conteniendo millones de registros de ventas, de hecho, la asociación más efectiva de algoritmos provienen de la comunidad de investigación de bases de datos. Existen ciertas versiones escalables de algoritmos de clasificación y clúster que utilizan técnicas de bases de datos.

### 9.5 Flujo de Datos

La minería de datos es uno de los miembros claves en la familia del data warehouse. La figura 48.16 ilustra el flujo normal de una empresa en el cual puede ser aplicado la minería de datos.

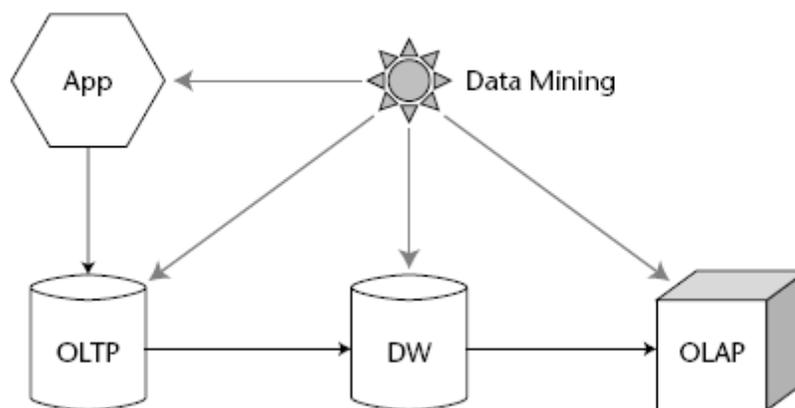


Figura 9.7: Flujo de datos de la minería de datos.

Pero surge la pregunta, ¿Cómo puede darle un valor agregado la minería de datos a este flujo? Primeramente, la minería de datos puede ser aplicada a un data warehouse donde los datos ya se encuentran depurados. Los patrones de descubrimiento por los modelos de minería pueden ser

<sup>53</sup> Machine Learning: disciplina científica que diseña y desarrolla algoritmos que le permiten a las computadoras cambiar su comportamiento basándose en datos.

presentados a los miembros del equipo de mercadeo a través de reportes. Usualmente en empresas pequeñas no poseen un data warehouse, por consiguiente, la minería se hace directamente a las tablas del OLTP, usualmente, haciendo una copia de las tablas en una base de datos separada.

La minería de datos puede tener un vínculo directo con las aplicaciones de negocio, más comúnmente a través de las predicciones. En un escenario de comercio electrónico, cuando un cliente añade un ítem al carrito de compras, se hace una consulta de predicción de minería de datos para obtener la lista de productos recomendados basados en análisis de asociación.

Además, la minería de datos, puede ser aplicada para analizar cubos OLAP. Un cubo es una base de datos multidimensional con medidas y dimensiones. La cantidad de registros en un cubo es exponencial al número de dimensiones y miembros en una dimensión, por lo cual se vuelve más difícil encontrar patrones interesantes manualmente. Las técnicas de minería de datos pueden ser aplicadas para descubrir patrones ocultos en un cubo. Por ejemplo, un algoritmo de asociación puede ser aplicado a un cubo de ventas, analizando los patrones de compras de los clientes para una región y periodo de tiempo específico. Se pueden aplicar las técnicas a predicciones de las medidas como ventas y ganancias.

## 9.6 Ciclo de vida de un Proyecto de Minería de Datos.

¿Cuál es el ciclo de vida de un proyecto de minería de datos? ¿Cuáles son los pasos desafiantes? ¿Quiénes deberían estar involucrados en un proyecto de minería de datos? Para responder a estas preguntas, a continuación se detallan los pasos de un proyecto típico de minería de datos.

### Paso 1: Recopilación de datos

El primer paso de la minería de datos es usualmente la recopilación de los datos. Los datos del negocio se encuentran almacenados en diferentes sistemas en toda la organización. El primer paso es obtener los datos relevantes a una base de datos o un data mart donde será aplicado el análisis. Por ejemplo si se trata de una empresa que analiza los clics que hacen los usuarios en internet, el primer paso sería bajar los logs web de cada uno de los servidores.

Si el data warehouse ya existe, puede que la información contenida en el no sea lo suficientemente rica y se necesitara obtener datos de otras fuentes. Suponiendo que ya se tienen el data warehouse que contiene la información relacionada a los clics rastreados por la empresa, se tiene información básica acerca de los patrones de navegación de los clientes, pero, debido a que no existe suficiente información demográfica acerca de los visitantes al sitio web, sería necesario comprar o obtener cierta información de este tipo desde otros orígenes de datos para poder crear un modelo más acertado.

Luego que los datos han sido recopilados, se pueden escoger la muestra de datos para reducir el volumen del set de datos, ya que en muchas ocasiones, los patrones que se muestran en 50000 clientes son los mismos que en 1 millón.

### Paso 2: Limpieza de datos y Transformación

La limpieza y transformación es la etapa con mayor uso intensivo de recursos en un proyecto de minería de datos. El propósito de la limpieza de datos es eliminar información irrelevante y ruido del set de datos. El propósito de la transformación de los datos es modificar el origen de los datos en diferentes formatos en términos de tipos de datos y valores. Existen varias técnicas que se pueden aplicar para realizar esta fase, incluyendo:

- Transformación de tipo de datos: Esta es la transformación más simple. Un ejemplo de este tipo de transformación, es pasar un campo de tipo Boolean a Entero. La razón de esta

transformación se da debido a que ciertos algoritmos de minería de datos tienen mejor desempeño sobre datos enteros, mientras que otros se desempeñan de mejor manera sobre los booleanos.

- Transformación continua de columnas: Para datos continuos como la que se encuentra en columnas de ingresos y edades, una transformación típica en estos casos es almacenar estos datos en contenedores. Por ejemplo, se puede desear almacenar los datos de las edades en 5 grupos predefinidos de edades. Además del almacenamiento, existen técnicas populares para transformación de datos continuos como la normalización, ya que esta mapea todos los valores numéricos en un número entre 0 y 1 (o de -1 a 1) para asegurarse que los números grandes no dominen los números pequeños durante el análisis.
- Agrupamiento: En ciertas ocasiones existen demasiados valores distintos (estados) para una columna discreta, por lo que es necesario agrupar estos valores en pocos grupos para reducir la complejidad del modelo. Por ejemplo, la columna Profesión puede tener varios valores diferentes como Ingeniero de Software, Ingeniero en Telecomunicaciones, Consultor, etc., en este caso se pueden agrupar varias profesiones de Ingeniería utilizando un solo valor: Ingeniería. El agrupamiento facilita la interpretación del modelo.
- Agregación: Esta técnica es otra transformación importante. Suponiendo que existe una tabla que contiene los registros de detalle de llamada (CDR) para cada cliente, y la meta es segmentar los clientes basados en su uso mensual de telefonía, es necesario realizar una agregación de todas las llamadas en ciertos atributos derivados como el número total de llamadas y el promedio de duración, ya que un CDR contiene información demasiado detallada de una llamada. Estos atributos derivados pueden ser usados en el modelo más adelante.
- Manejo de valores perdidos: La mayoría de los sets de datos contienen valores perdidos. Existen varias causas por las cuales se da este fenómeno, por ejemplo, se pueden tener 2 tablas de clientes que vienen de dos bases de datos OLTP distintas. Al realizar la unión de estas tablas se puede dar valores perdidos, ya que la definición de las tablas no es exactamente la misma. Otro ejemplo se da cuando la información demográfica de un cliente tiene una columna de edad, pero los clientes no siempre proporcionan esta información cuando se están registrando. Existen ciertas formas de manejar este problema, uno de los más populares es reemplazar los valores perdidos con una constante. Si no se conoce la edad de un cliente, esta se puede reemplazar con el promedio de edad de todos los clientes. Cuando una fila contiene demasiados valores perdidos se puede optar por quitarla. En situaciones más avanzadas, se puede construir un modelo de minería utilizando los casos completos y luego aplicar el modelo para predecir el valor más probable de cada columna perdida.
- Eliminación de valores extremos: Los valores extremos son casos anormales en un set de datos, los cuales pueden afectar la calidad del modelo. Por ejemplo, si se quiere crear un modelo de segmentación de clientes basado en el uso telefónico del cliente (duración promedio, número total de llamadas, factura mensual, llamadas internacionales, etc.), existen clientes (0.5%) que se comportan de una forma muy distinta, por ejemplo, utilizando roaming todo el tiempo. Si se incluyen estos casos anormales en el modelo, se terminara con un modelo que contiene a la mayoría de los clientes en un segmento y pequeños segmentos conteniendo estos casos anormales. La mejor forma de tratar estos valores extremos es simplemente eliminándolos antes del análisis, esta eliminación se puede realizar basándose en atributos individuales, por ejemplo, eliminando el 0.5% de los clientes con mayor y menor facturación. Para realizar esto se puede utilizar un algoritmo de clúster.

### **Paso 3: Construcción del Modelo**

Una vez se ha realizado la limpieza de los datos y las variables han sido transformadas, se puede proseguir a la construcción de los modelos. Antes de construir cualquier modelo, es necesario comprender el objetivo de un proyecto de minería de datos y el tipo de tarea de minería de datos. ¿Es este proyecto una tarea de clasificación, de asociación o de segmentación? En esta etapa, es necesario reunirse con los analistas de negocio que posean conocimiento y dominio del tema, por ejemplo, en una minería a las telecomunicaciones, se debería incorporar a los empleados del departamento de mercadeo quienes conocen del negocio de las telecomunicaciones.

La construcción del modelo es el núcleo de la minería de datos, aunque no es tan intensiva en el uso de recursos como la transformación de los datos. Una vez se comprende la tarea de minería de datos, es relativamente fácil escoger el algoritmo adecuado, ya que para cada tarea de minería de datos, existen varios algoritmos elegibles. En muchas ocasiones, no se conoce cuál es el mejor algoritmo que se puede aplicar antes de la capacitación del modelo. La precisión del algoritmo depende de la naturaleza de los datos como el número de estados del atributo de predicción, el valor de distribución de cada atributo, las relaciones entre los atributos, etc.

Por ejemplo, si la relación entre los atributos de entrada y los atributos predecibles fuese lineal, el algoritmo de árbol de decisión sería una buena opción. Si la relación entre los atributos es más complicada, entonces el algoritmo de red neuronal debería ser considerado.

El enfoque correcto es construir múltiples modelos utilizando diferentes algoritmos y luego compararlos con la exactitud de estos modelos utilizando una herramienta, como el gráfico de elevación, el cual es descrito en el siguiente paso. Aun para el mismo algoritmo, puede ser necesario construir múltiples modelos utilizando diferentes parámetros para poder afinar la precisión del modelo.

### **Paso 4: Evaluación del Modelo (Model Assessment)**

Durante el estado de construcción del modelo, se construye un conjunto de modelos utilizando ajuste con diferentes parámetros y algoritmos. ¿Cuál es el mejor modelo en términos de exactitud?, ¿cómo se evalúan estos modelos? Existen pocas herramientas populares para la evaluación de la calidad de estos modelos, la mejor conocida es Lift Chart. Lift Chart utiliza un modelo muy bien probado para predecir los valores de un grupo de datos de prueba. Basado en el valor predicho y su probabilidad, despliega el modelo en un gráfico descriptivo.

Durante el estado de evaluación del modelo, no únicamente se utilizan herramientas para evaluar la exactitud del modelo, también es necesario discutir el significado de encontrar patrones con análisis de negocio. Por ejemplo, si se construye un modelo de asociación sobre un grupo de datos, seguramente se encontrarán reglas como:

Relación = Esposo => Genero => Masculino con 100% de confianza.

Ciertas ocasiones el modelo no contiene patrones útiles. Esto puede ocurrir por pocas razones. Primera razón, es porque los datos son completamente aleatorios. Mientras sea posible tener datos aleatorios, en la mayoría de casos, grupos reales de datos si contienen información valiosa. La Segunda razón, y más lógica, en caso que el ajuste de variable en el modelo no es el mejor para utilizar, en este caso generalmente se necesita repetir la limpieza de datos y su transformación, con el objetivo de derivar, variables más significativas. La Minería de datos es un proceso cíclico; usualmente toma pocas iteraciones para encontrar el modelo correcto.

## **Paso 5: Reporteo (Reporting)**

El reporte es un importante canal de entrega para la búsqueda de Minería de Datos. En muchas organizaciones, la meta de la Minería de Datos es entregar los reportes a los ejecutivos de Marketing. La mayoría de herramientas de minería de datos poseen características de reporte que permiten a los usuarios generar reportes predefinidos desde el modelo de minería, junto a salidas graficas y/o textuales. Existen dos tipos de reportes: Reportes acerca de patrones encontrados y Reportes acerca de predicciones o pronóstico.

## **Paso 6: Predicción (Scoring)**

En muchos proyectos de Minería de Datos, encontrar patrones es solo la mitad del trabajo; la meta final es utilizar estos modelos para predicciones. La predicción es también llamada scoring en terminología de Minería de Datos. Para brindar predicciones, se necesita poseer un modelo lo suficientemente probado y un grupo de casos nuevos. Considerar el escenario bancario en el cual se deba que construir un modelo acerca de la evaluación del riesgo de préstamos. Cada día existen miles de nuevas aplicaciones para préstamos. Se puede utilizar el modelo de evaluación de riesgos para predecir el riesgo potencial para cada una de las aplicaciones de préstamos.

## **Paso 7: Integración con Aplicaciones**

Incorporar Minería de datos dentro de las aplicaciones de negocio, significa aplicar nuevamente inteligencia de negocios, lo cual es muy cercano a un lazo de análisis (analysis loop). De acuerdo a Gartner Research, en los próximos años, cada vez más y más, las aplicaciones de negocios cubrirán los componentes de Minería de Datos como un valor agregado. Por ejemplo, las aplicaciones CRM (sigla del término en inglés "Customer Relationship Management") podrían poseer características de Minería de Datos que agrupe clientes en segmentos. Las aplicaciones ERP (siglas en inglés "Enterprise resource planning") podrían poseer características de Minería de Datos para la predicción de producción.

Una tienda de libros en línea, puede brindar a sus clientes recomendaciones sobre libros en tiempo real. La integración de características de Minería de Datos, especialmente en componentes de predicciones en tiempo real dentro de las aplicaciones, es un paso muy importante para los proyectos de Minería de Datos. Este es el paso clave incorporar el uso en masa de la Minería de Datos

## **Paso 8: Administración del Modelo**

Es todo un desafío el mantenimiento del estatus del Modelo de Minería. Cada Modelo de Minería posee un ciclo de vida. En algunos negocios, los patrones son relativamente estables y los modelos no requieren frecuentemente reevaluación. Pero en muchos patrones de negocios varia frecuentemente. Por ejemplo, en la tienda de libros en línea, nuevos libros aparecen cada día. Esto significa que nuevas reglas de asociación aparecen cada día. La duración de un Modelo de Minería es limitado. Una nueva versión de un modelo debe ser creado frecuentemente. Últimamente, determinar la exactitud del modelo y la creación de una nueva versión del modelo deberá ser realizado utilizando procesos automatizados.

Como cualquier dato, el Modelo de Minería posee problemas de seguridad. Los Modelos de Minería poseen patrones, muchos de estos patrones son las sumatoria de datos sensitivos. Es necesario mantener los privilegios de lectura, escritura y predicción para diferentes perfiles de usuarios.

## 9.7 Minería de datos y El Mercadeo Actual

En esta sección se brinda una vista amplia sobre el mercadeo de la Minería de Datos actual y algunos vendedores dentro de este campo.

### 9.7.1 Tamaño del Mercadeo de la Minería de Datos

Giga investigaciones estiman el tamaño del mercado para la Minería de Datos, ha sobre pasado la marca de billón de dorales, incluyendo software y servicios (consultoría y servicios de oficina). Otra investigación de una organización difiere y realiza estimaciones más conservadoras de su tamaño de mercado, desde \$200 hasta \$700 millones. Sin embargo, una conclusión de investigación es compartida por varios analistas: el mercado de la Minería de Datos es el componente de inteligencia de negocios que más rápido crece (reporteo, OLAP, empaclado de Data Marts, etc.). La Minería de Datos actualmente representa cerca del 15% del mercado de la inteligencia de negocios.

### 9.7.2 Principales Vendedores y Productos

Existen cientos de compañías que brindan productos y consultorías dentro de la Minería de Datos. KD Nuggets (kdnuggets.com) ha extendido una lista de la mayoría de estas compañías y de sus productos en el campo de Minería de Datos. A continuación se listan los principales productos de Minería de Datos de compañías.

- **SAS:** SAS es probablemente el producto de Minería de Datos más amplio en términos de mercado compartido. SAS ha permanecido en el campo estadístico por décadas. SAS basa su contenido en un rico conjunto de funciones estadísticas que pueden ser utilizadas para cualquier tipo de análisis. También posee un poderoso lenguaje script llamado SAS Script.
- **SPSS:** SPSS es otra de las principales compañías estadísticas. Poseer un número productos de Minería de Datos incluyendo SPSS basado en respuestas y en árbol de decisiones.
- **IBM:** IBM posee el producto de Minería de Datos llamado Intelligence Miner, desarrollado por la subsidiaria Alemania IBM. Intelligence Miner contiene un grupo de algoritmos y herramientas visuales. Intelligence Miner exporta Modelos de Minería en Predictive Modeling Markup Language (PMML), el cual fue definido por Data Mining Group (DMG), una organización de la industria.
- **Microsoft Corporation:** Microsoft fue el primer principal vendedor de bases de datos que incluían características de Minería de Datos en bases de datos relacionales. SQL Server 2000 liberado en Septiembre del 2000, contiene 2 algoritmos patentes de Minería de Datos: Microsoft Decisión Trees y Microsoft Clustering. A parte de estos algoritmos, la característica más importante de Minería de Datos es la implementación de OLE DB para Minería de Datos.
- **ORACLE:** Oracle 9i en el 2000, contenía algunos algoritmos de Minería de Datos, basados en asociaciones y Naïve Bayes. Oracle 10g incluía muchas más herramientas y algoritmos para la Minería de Datos. Oracle también incorporo Java Data Mining API, el cual es un paquete de Java para tareas de Minería de Datos.
- **Angoss:** Angoss KnowlfeigesSTUDIO es una herramienta de Minería de Datos que incluye la capacidad de poder construir arboles de decisión, cluster analysis y una gran cantidad de modelos predictivos, permitiendo al usuarios poseer y entender los datos desde diferentes perspectivas. Incluye una poderosa herramienta visual para el soporte y explicación de descubrimientos.

- KXEN: KXEN es un software de Minería de Datos, proveído desde su base en Francia. Posee un numero de algoritmos para la Minería de Datos que incluye SVM, regresión, Series en tiempo, segmentación, etc. También provee soluciones de Minería de Datos para cubos OLAP. Desarrollo un add-in para Excel que permite a los usuarios realizar Minería de Datos en un ambiente muy familia como Excel.

### 9.7.3 Problemas Actuales y Desafíos.

Durante los últimos años, la Minería de Datos ha sido muy pronunciada, pero sigue siendo un mercado muy pequeño relativamente. La mayoría de usuarios de la Minería de Datos, son los analistas de grandes negocios en los sectores de finanzas, telecomunicaciones, y seguranzas. La Minería de Datos todavía se considera como una opción muy alta, por parecer muy sofisticada para la mayoría de desarrolladores el entender, muy pocas aplicaciones incorporan las características de Minería de Datos.

La Minería de Datos no es una de las principales líneas de tecnología, mas sin embargo posee el potencial para brindar un valor agregado para casi cualquier tipo de aplicación de negocios. Estos son algunos pocos desafíos a sobre pasar antes que la Minería de Datos se convierta una tecnología de masas.

- Paquetes Horizontales propietarios sin estadares API.
- Orientación a análisis en lugar de orientación a desarrollo.
- Educación limitada de usuarios.
- Características limitadas de algoritmos.

### 9.8 Estándares de Minería de Datos

La Minería de Datos es un campo relativamente nuevo. Se puede comparar los días del mercado de la Minería de Datos contra el mercado de las bases de datos de hace 20 años, cuando no existían conceptos sobre bases de datos relacionales, ni SQL. Cada vendedor tenía su propio formato de almacenamiento, y no era muy sencillo consultar diferentes fuentes de datos. Durante los últimos años, los vendedores han empezado a reconocer problemas muy similares en el campo de la Minería de Datos y han realizado un esfuerzo por estandarizar la Metadata de la Minería de Datos, APIs y formato de contenido. Este esfuerzo fue principalmente llevado por cuerpos industriales o vendedores de software de Minería de Datos independientes.

### 9.9 Aplicación de la Minería de Datos al Data Warehouse de la Defensoría del Consumidor.

Para la aplicación de la minería de dato en el proyecto se utiliza el servidor Microsfot Analysis Servicess 2005, el cual permite la aplicación de diferentes algoritmos para la minería de datos. Para la aplicación sobre el Data Warehouse se utiliza el algoritmo de Series de tiempo (time series) , el cual permite realizar proyecciones basado en una serie de tiempo.

#### 9.9.1 Algoritmo de series de tiempo

Una serie de tiempo consste en una serie de datos coleccionados en varios incrementos de tiempo o cualquier otro indicador o secuencia. Una secuencia de valores es una variable que, a travez del tiempo, forma una serie de tiempo, como puede ser las ventas mensuales de Pepsi en una serie de tiempo. Generalmente, en una serie de tiempo, un valor en un determinado tiempo es correlativo a los valores en los tiempos que le preceden. Por ejemplo, las ventas de Pepsi el 10 de mayo están relacionadas

fuertemente con las ventas del 9 y 8 de mayo. El tiempo incrementa en una serie de tiempo puede ser discretos o continuos, pero en el caso de la minería de datos, se consideran únicamente las series de tiempo en las cuales los incrementos son discretos. Además, los valores observados en una serie de tiempo pueden ser discretos o continuos, por ejemplo, el valor de existencias de inventario, ventas e ingresos de una organización tienen observaciones continuas, en cambio, una serie de tiempo de las predicciones meteorológicas como soleado, nublado o lluvia tienen observaciones discretas.

El principal propósito de recolectar la información en series de tiempo es para hacer predicciones de valores futuros.

### 9.9.2 Preparación de los datos

Para aplicar este algoritmo de minería de datos se creó una vista en el data warehouse llamada "PROYECCION\_CASO\_TimeSeries", en la cual se tiene la serie de tiempo de los casos por centro de solución, con el fin de obtener una proyección de los casos que serán recibidos en el futuro por centro de solución de controversias.

En la tabla 9.8 se muestran los campos de la vista creada.

Nombre	Tipo	Descripción
Fecha	Entero	Contiene el año y el mes en que se ha ingresado un caso. Llave de tiempo para la aplicación del algoritmo.
NombreCentroSolucion	Varchar	Contiene el nombre del centro de solución de controversias que recibió la denuncia
cantidad	Int	Cantidad de visitas recibidas en la serie de tiempo. Variable de proyección.

Tabla 9.8: Tipos de dato y descripción de los campos de la vista.

### 9.9.3 Salida mostrada por minería de datos

A continuación se muestra la interfaz de Microsoft Visual Studio BI 2005, en la cual se puede ver los resultados de la aplicación un algoritmo de series de tiempo para realizar proyecciones. La pantalla se muestra en la figura 9.9.

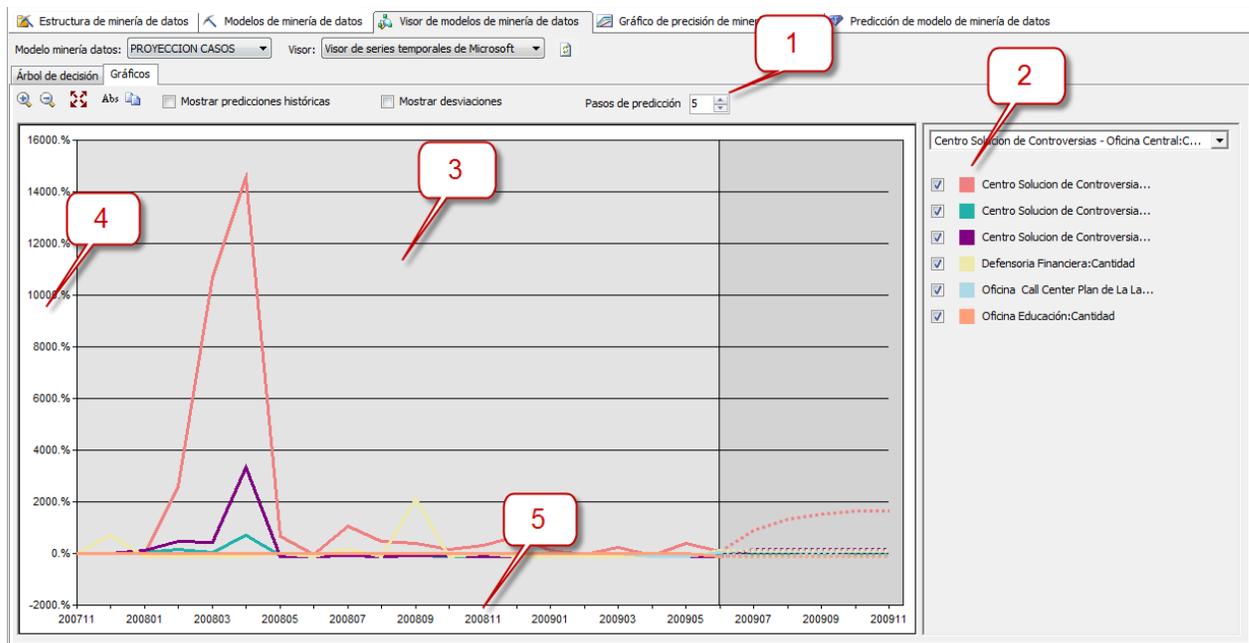


Figura 9.9: Salida para proyecciones de minería de datos.

<b>Número</b>	<b>Descripción</b>
1	Pasos de predicción futura que se desean visualizar en el área del grafico.
2	Listado de Centro de Solución de Controversia que desean ser visualizados en el grafico
3	Área de grafico de proyección.
4	Eje Y, donde se muestran las ocurrencias
5	Eje X, donde se muestran las series de tiempo.

**Tabla 9.1: Detalle de componentes de visualizador de Proyecciones (serie de tiempo).**

## ANEXO 10: Implementación del log en los paquetes

Log es un nombre que se le da a un registro oficial de eventos durante un periodo de tiempo en particular, para este caso se ha configurado un log que permite detectar la ejecución errónea de los paquetes definidos en SQL Server Integrations Services (SSIS), este log queda registrado en una ubicación específica, esta ubicación se ha definida en la carpeta C:\log carga. La configuración de este log se encuentra definida en el manual de instalación.

A continuación se muestra una figura que representa la interacción de cada paquete con el archivo de log definido previamente.

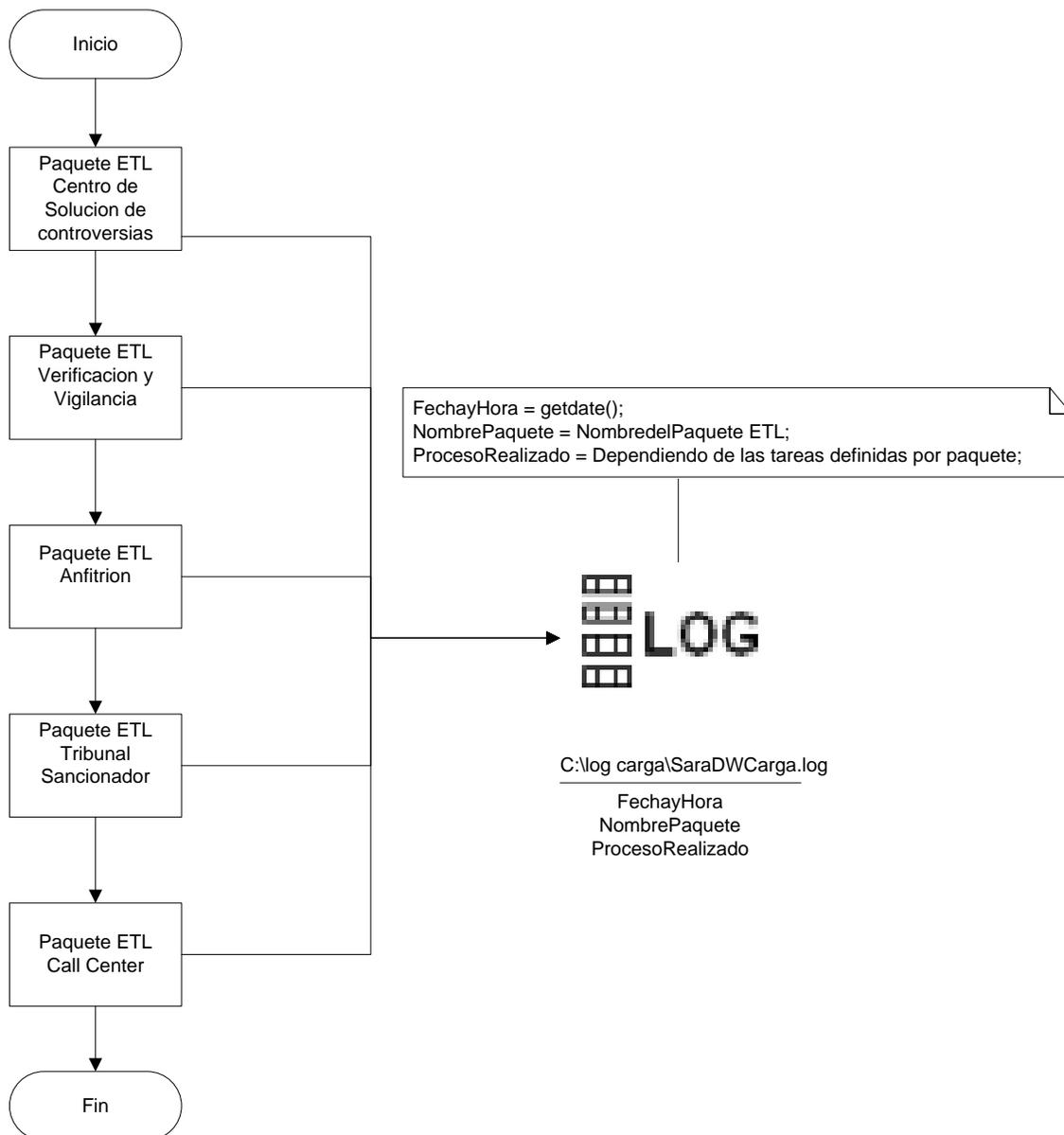


Figura 10.1: Flujo que muestra la interacción con el archivo del log.

## ANEXO 11: Aceptados por Defensoría del Consumidor

### 11.1 Requerimientos para Centro de Solución de Controversias

*Dirección de Estudios, Planeamiento y Desarrollo*  
Requerimientos para el SARA.



#### Requerimientos para el nuevo módulo de reportes

##### **Información usada comúnmente dentro de los reportes**

Los reportes incluyen generalmente dos tipos de información que corresponde a los campos por los que es agrupado. Los datos son:

- Conteo de casos
  - Total de casos
  
- Monto reclamado
  - Total del monto reclamado
  - Promedio del monto reclamado por las variables que se agrupe

##### **Campos usados comúnmente para agrupar los casos**

A pesar de haber una serie de reportes regulares que son solicitados al departamento de estudios, la variedad de solicitudes de información es grande, y se ha determinado que lo más conveniente es que se cuente con categorías por las cuales se pueda agrupar la información para ser colocada en columnas o filas en varios niveles (dos o más categorías en filas o columnas)

- Tipo de caso
  - Denuncia: existen denuncias personales y no personales en la base, debe ser posible mostrarlas juntas como solo denuncias y separadas en sus dos categorías
  - Asesoría
  - Gestión
  - Derivación
  - Información: este tipo existe solamente para call center y debe incluir el tipo de información del que se trata
  - Orientación de consumo: solamente para call center y debe incluir el tema de consulta
- Centro de solución receptor: es el centro que recibe el caso en su primera oportunidad
  - Call Center
  - San Salvador
  - Santa Ana
  - San Miguel
  - Financiera
- Fecha de ingreso
  - Día de ingreso
  - Semana de ingreso
  - Mes de ingreso
  - Año de ingreso

*Dirección de Estudios, Planeamiento y Desarrollo*  
Requerimientos para el SARA.



- Tipo de solución  
La solución de un caso es determinada de la manera siguiente:
  1. Si el estado de un caso es “Abierto” su solución es “Abierto”
  2. Si no se cumple lo anterior y su subetapa es “Cerrado por desistimiento” la solución es “Cerrado por desistimiento”
  3. Si no se cumple lo anterior y su etapa es “Avenimiento” la solución es “Avenimiento”
  4. Si no se cumple lo anterior y su etapa es “Conciliación” la solución es “Conciliación”
  5. Si no se cumple lo anterior y su etapa es “Tribunal sancionador” la solución es “Tribunal sancionador”
  6. Si no se cumple lo anterior y su subetapa es “Caso registrado con prevención” la solución es “Prevención”
  7. Si no se cumple lo anterior y su subetapa es “Falta de ratificación” la solución es “Falta de ratificación”
  
- Sector
  
- Categoría: normalmente agrupado dentro de su sector correspondiente
  
- Motivo: normalmente agrupado dentro de su sector correspondiente o por proveedor.
  
- Se requiere que los motivos puedan ser agrupados en categorías más reducidas que serían definidas por el departamento de estudios
  
- Proveedor; Se requiere que los proveedores puedan ser agrupados en categorías más reducidas que serían definidas por el departamento de estudios
  
- Técnico receptor
  
- Estado
  
- Etapa: agrupados dentro de su estado
  
- Subetapa: agrupada dentro de la etapa
  
- Departamento (de El Salvador): generalmente en conjunto con sector
  
- Municipio: generalmente agrupado dentro del departamento al que corresponde
  
- Institución de derivación: únicamente para las derivaciones
  
- Sexo del consumidor

Dirección de Estudios, Planeamiento y Desarrollo  
 Requerimientos para el SARA.



## Ejemplos de los resultados de los cruces

A continuación, algunos ejemplos del tipo de información

### Asesorías por sector y motivo

Motivo Asesoría	Total
<b>COBROS Y CARGOS</b>	
COBRO INDEBIDO	
COBRO EXCESIVO POR ALTO CONSUMO	
CONSUMO EXCESIVO	
COBRO POR CAMBIO DE TARIFA	
CARGO INDEBIDO	
COBRO POR CONSUMO ESTIMADO	
COBRO POR PÉRDIDA DE SUBSIDIO	
COBRO POR RECONEXIÓN	
RECARGO POR MORA	
COBRO POR ERROR DE LECTURA	
COBRO DE AGUA POTABLE NO REGISTRADO	
COBRO POR CONSUMO ACUMULADO	
COBRO POR PROMEDIOS INDEBIDOS	
COBRO DIFERIDO	
COBRO POR DEFICIENCIA DE SERVICIO	
COBRO POR CRUCE DE CUENTA	
SALDO INCORRECTO	
<b>CALIDAD DEL SERVICIO</b>	
MAL SERVICIO	
MALA CALIDAD	
<b>MEDIDOR</b>	
COBRO POR MEDIDOR DEFECTUOSO	
COBRO POR MEDIDOR DESTRUIDO	
<b>VARIOS</b>	
MOTIVOS VARIOS	
PLAN DE PAGOS	
EMBARGO	
INCUMPLIMIENTO DE CONTRATO	
INCUMPLIMIENTO DE OFERTA	
INCUMPLIMIENTO DE GARANTIA	
MALA LECTURA	
ENTREGA INCOMPLETA	
PUBLICIDAD ENGAÑOSA	
ENGAÑO	
<b>Total general</b>	

*Dirección de Estudios, Planeamiento y Desarrollo*  
 Requerimientos para el SARA.



**Casos recibidos por centro**

CSC	Denuncia	Gestión	Derivación	Total
San Salvador				
Santa Ana				
San Miguel				
Financiera				
<b>Total</b>				

**Montos reclamados por año y sector**

	2007	2008	2009
Agua Potable	\$1	\$1	\$1
Alimentos	\$1	\$1	\$1
Comercio	\$1	\$1	\$1
Electrodomésticos	\$1	\$1	\$1
Energía Eléctrica	\$1	\$1	\$1
Gobierno	\$1	\$1	\$1
Hidrocarburos	\$1	\$1	\$1
Inmuebles	\$1	\$1	\$1
Libros	\$1	\$1	\$1
Medicamentos	\$1	\$1	\$1
Muebles	\$1	\$1	\$1
Publicidad	\$1	\$1	\$1
Servicios	\$1	\$1	\$1
Sistema Financiero	\$1	\$1	\$1
Telecomunicaciones	\$1	\$1	\$1
Turismo	\$1	\$1	\$1
Varios	\$1	\$1	\$1
Vehículos	\$1	\$1	\$1
<b>Total</b>	\$1	\$1	\$1

## 11.2 Requerimientos para Verificación y Vigilancia



### Requerimientos para módulo de Verificación y Vigilancia

#### Información usada comúnmente dentro de los reportes

Los reportes incluyen generalmente varios tipos de información que corresponde a los campos por los que es agrupado. Los datos son:

- Conteo de inspecciones
  - Total de inspecciones
  - Total inspecciones con hallazgos
  - Inspecciones por rubro
  - Inspecciones por rubro con hallazgos
  - Inspecciones por fechas

Ejemplo:

#### Inspecciones realizadas por oficina regional

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Total general
San Salvador	26	226	174	102	97	86	43	64	15				833
Santa Ana	23	101	81	62	61	76	23	23	17				467
San Miguel	28	105	74	46	47	48	30	24					402
<b>Total</b>	<b>77</b>	<b>432</b>	<b>329</b>	<b>210</b>	<b>205</b>	<b>210</b>	<b>96</b>	<b>111</b>	<b>32</b>				<b>1702</b>

#### Inspecciones realizadas por oficina regional con hallazgos

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Total general
San Salvador	17	26	51	33	29	33	20	29	6				244
Santa Ana	9	48	45	22	22	35	16	18	9				224
San Miguel	1	3	22	22	14	24	11	5					102
<b>Total</b>	<b>27</b>	<b>77</b>	<b>118</b>	<b>77</b>	<b>65</b>	<b>92</b>	<b>47</b>	<b>52</b>	<b>15</b>				<b>570</b>

#### Inspecciones por rubros

Rubro	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Total general
SUPERMERCADO	20	9	27	34	10	16	64	74	13	267
OTROS	7	112				3		3		125
ALMACEN DE ROPA, ZAPATOS Y ACCESORIOS	34	262		3		2			1	302
ALMACENES DE MUEBLES Y ELECTRODOMÉSTICOS	2	1		15						18
OTROS ALMACENES	2	23								25
OTROS PROVEEDORES DE ALIMENTOS		1	6	1		1		5		14
HOTELES Y CENTROS RECREATIVOS			4	28						32
VENTA DE VEHÍCULOS Y REPUESTOS								6		6
CAFETIN								1		1
<b>Total general</b>	<b>65</b>	<b>408</b>	<b>37</b>	<b>81</b>	<b>10</b>	<b>22</b>	<b>64</b>	<b>89</b>	<b>14</b>	<b>790</b>



**Inspecciones por rubros con hallazgos**

Rubro	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Total general
SUPERMERCADO	14	9	19	21	2	7	36	37	6	151
OTROS	3	9				0		2		14
ALMACEN DE ROPA, ZAPATOS Y ACCESORIOS	5	42		0		2			0	49
ALMACENES DE MUEBLES Y ELECTRODOMÉSTICOS	0	0		4						4
OTROS ALMACENES	0	6								6
OTROS PROVEEDORES DE ALIMENTOS		0	5	0		0		4		9
HOTELES Y CENTROS RECREATIVOS			3	13						16
VENTA DE VEHÍCULOS Y REPUESTOS								2		2
CAFETIN								1		1
<b>Total general</b>	<b>22</b>	<b>66</b>	<b>27</b>	<b>38</b>	<b>2</b>	<b>9</b>	<b>36</b>	<b>46</b>	<b>6</b>	<b>252</b>

**Inspecciones por rubros con hallazgos y sin hallazgos**

Rubro	Ene			Feb			Total general
	Sin hallazgos	Con hallazgos	Total	Sin hallazgos	Con hallazgos	Total	
SUPERMERCADO	14	9	23	19	21	50	73
OTROS	3	9	12				12
<b>Total general</b>	<b>17</b>	<b>18</b>	<b>35</b>	<b>19</b>	<b>21</b>	<b>50</b>	<b>85</b>

- Detalle de artículos de la Ley relacionados en inspecciones
  - Por artículos. Que puedan verse 1 o más tipos de artículos y que estos puedan relacionarse por rubro o por establecimiento o proveedor.
  - Por hallazgos

Ejemplo:

**Por artículo verificado**

Tipo de inspección	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Total general
Art. 14 Fechas de vencimiento	9	4	120	59	64	53	15	34	10	36
Art. 14 Fechas de vencimiento con hallazgos	10	5	42	27	25	22	13	11	4	15
<b>Total</b>	<b>19</b>	<b>9</b>	<b>162</b>	<b>86</b>	<b>89</b>	<b>75</b>	<b>28</b>	<b>45</b>	<b>14</b>	<b>52</b>

**Por cantidad de productos por cada artículo verificado**

Tipo de inspección	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Total general
Suma de Cantidad productos Art. 14	181	104	590	327	1343	443	167	115	41	331
Suma de Cantidad productos Art. 27	10	199	153	77	82	63	20	29	11	64
Suma de Cantidad de productos Art. 43 B	6	1	2	6		2	6	6	1	3
Suma de Cantidad de productos Art. 27										
<b>Total</b>	<b>197</b>	<b>304</b>	<b>745</b>	<b>410</b>	<b>1425</b>	<b>508</b>	<b>193</b>	<b>150</b>	<b>53</b>	<b>398</b>



**Por rubro y artículo verificado**

Rubro	Art. 14	Art. 14 hallazgos	Cantidad productos	Art. 31	Art. 31 hallazgos	Art. 27 Precios a la vista	Art. 27 Precios a la vista hallazgos	Cantidad de productos
SUPERMERCADO	127	91	1325	161	25	141	88	
OTROS	1	1	7	1		107	12	
ALMACEN DE ROPA, ZAPATOS Y ACCESORIOS					2	250	47	
ALMACENES DE MUEBLES Y ELECTRODOMÉSTICOS						2		
OTROS ALMACENES						19	6	
OTROS PROVEEDORES DE ALIMENTOS	3	1	46	1	5	3		
HOTELES Y CENTROS RECREATIVOS	16	16	149			32		
VENTA DE VEHÍCULOS Y REPUESTOS								
<b>Total general</b>	<b>147</b>	<b>109</b>	<b>1527</b>	<b>163</b>	<b>32</b>	<b>554</b>	<b>153</b>	

- Información de los inspectores
    - Cantidad de inspecciones realizadas por delegado, por mes y totales
    - Cantidad de inspecciones realizadas por delegado con hallazgos
    - Información de los errores: con detalle del delegado y la cantidad de errores. (Pendiente cuadro ejemplo)
- Ejemplo:

**Cantidad de inspecciones realizadas por delegado**

Inspectores	Total
Alicia de Gómez	113
Arnoldo Reyes	61
Candelaria Carrillo	1
CLAUDIA REYES	5
DORA ALICIA ORELLANA	45
Edgardo Prieto	62
Eduardo Hidalgo	44
FREDIS HERNÁNDEZ	4
Guillermo Funes	101
Gustavo Ernesto Barahona Rojas	82
Jaime Ernesto Peña Urrutia	10
Manuel Menjivar	48
Mauricio Brizuela	30
Nelson Galeas	13
Oscar Montes	70
Romeo Elías Barrera	73
Sifredo Romero	4
Silvia Elizabeth Blanco	140
VALMORE GÓMEZ	66
Zoila Angélica Vásquez de Peñate	201
<b>Total general</b>	<b>1221</b>



Dirección Nacional de Verificación y Vigilancia

- Detalle del proveedor
  - Por nombre del establecimiento y por sucursal.
  - Por hallazgos

Que la información sea opcional, puede o no incluirse toda la información de los campos.

Ejemplo: Con detalle solo del art. 14

**Por nombre del establecimiento y artículo verificado**

Nombre del Establecimiento	Propietario	OFI	Zona o Sucursal	Dirección del Establecimiento	Art. 14 Fechas de vencimiento	Art. 14 Fechas de vencimiento hallazgos
SUPER SELECTOS	CALLEJA S.A. DE C.V.	SS	AGUILARES	AVENIDA CENTRAL NORTE Y FINAL SEGUNDA CALLE ORIENTE		1
			ARCE	CALLE ARCE NÚMERO 470	2	
			BEETHOVEN	PASEO GENERAL ESCALÓN Y AVENIDA ALTAMIRA	1	
			CARIBE	EDIFICIO RIALTO, COLONIA ESCALÓN, FRENTE A EL SALVADOR DEL MUNDO	1	
			CENTRO LIBERTAD	CUARTA CALLE OREINTE, ENTRE SEGUNDA Y CUARTA AVENIDA SUR		2
			CIUDAD DELGADO	AVENIDA PALECA Y CALLE LA JOYA, NÚMERO 14 Y 16		
			COJUTEPEQUE	PRIMERA AVENIDA SUR Y CUARTA CALLE PONIENTE		1
			EL FARO	CENTRO COMERCIAL EL FARO, ANCLA "B-1", KILOMETRO VEINTITRES Y MEDIO	1	1
			EL PASEO	PASEO GENERAL ESCALÓN ENTRE 77 Y 79 AVENIDA SUR		1
			GIGANTE	AVENIDA OLÍMPICA Y 59 AVENIDA SUR	1	1
Total general					23	24

### 11.3 Requerimientos para Tribunal Sancionador

*Dirección de Estudios, Planeamiento y Desarrollo*  
 Requerimientos para el SARA.



#### Requerimientos Cubo para el módulo del Tribunal Sancionador

##### Información usada comúnmente dentro de los reportes

Los reportes incluyen generalmente varios tipos de información que corresponde a los campos por los que son agrupados. Los datos son:

- Conteo de casos
  - Total de casos

Ejemplo:

##### Denuncias Recibidas por centro

	Ene-06	Feb-06	Mar-06	Abr-06	May-06	Jun-06	Jul-06	Ago-06	Sep-06	Oct-06	Nov-06	Dic-06
Presidencia	45	64	12	77	45	44	6	8	45	30	22	81
San Salvador	273	210	202	181	280	244	260	246	269	331	337	211
Santa Ana	85	56	96	50	72	140	105	107	147	427	395	252
San Miguel	55	42	62	30	45	37	63	44	100	170	155	125
Financiera	38	38	60	31	56	27	28	33	28	34	31	16
<b>Total</b>	<b>451</b>	<b>346</b>	<b>420</b>	<b>292</b>	<b>453</b>	<b>448</b>	<b>456</b>	<b>430</b>	<b>544</b>	<b>962</b>	<b>918</b>	<b>604</b>

- Montos recuperados
  - Total de montos impuestas agrupadas por distintas variables
  - Promedio de los montos impuestas por las variables que se agrupe

Ejemplo:

##### Monto por año y sector

	2007	2008	2009
Agua Potable	\$1	\$1	\$1
Alimentos	\$1	\$1	\$1
Comercio	\$1	\$1	\$1
Varios	\$1	\$1	\$1
Vehículos	\$1	\$1	\$1
<b>Total</b>	<b>\$1</b>	<b>\$1</b>	<b>\$1</b>

*Dirección de Estudios, Planeamiento y Desarrollo*  
Requerimientos para el SARA.



### **Campos usados comúnmente para agrupar los casos**

A pesar de haber una serie de reportes regulares que son solicitados a la Dirección de Estudios, la variedad de solicitudes de información es grande, y se ha determinado que lo más conveniente es que se cuente con categorías por las cuales se pueda agrupar la información para ser colocada en columnas o filas en varios niveles (dos o más categorías en filas o columnas)

- Número de expediente
  
- Centro de Solución de origen: es el centro que recibe el caso en su primera oportunidad
  - San Salvador
  - Santa Ana
  - San Miguel
  - Financiera
  - Presidencia: aplica a todas las denuncias no personales no recibidas por los centros
  
- Fecha de ingreso al Tribunal Sancionador
  
- Sector\*
  
- Infracción
  
- Estado y/o etapa
  
- Categoría: normalmente agrupado dentro de su sector correspondiente\*
  
- Motivo: normalmente agrupado dentro de su sector correspondiente o por proveedor.

*Se requiere que los motivos puedan ser agrupados en categorías más reducidas que serían definidas por el departamento de estudios*

- Proveedor (nombre comercial y legal): Se requiere que los proveedores puedan ser agrupados en categorías más reducidas que serían definidas por el departamento de estudios
  
- Departamento (de El Salvador) del proveedor denunciado: generalmente en conjunto con sector

---

\* Es posible que estos campos no se encuentren actualmente disponibles en el módulo SARA.

*Dirección de Estudios, Planeamiento y Desarrollo*  
 Requerimientos para el SARA.



- Municipio del proveedor denunciado: generalmente agrupado dentro del departamento al que corresponde

### Casos por infracción

Las infracciones son un buen indicador de la problemática de los consumidores por lo que resulta relevante conocer cuáles son las causas más comunes.

Ejemplo: Casos absueltos en 2008

Infracción	Total de Casos
Artículo 43	48
Artículo 115	132
Artículo 142	252
<b>Total</b>	<b>432</b>

### Casos por sector a lo largo del tiempo

Las alertas para este tipo de análisis deberían ser un poco diferentes, pues lo que se busca resaltar es el crecimiento de los valores a lo largo del tiempo.

	2005_09	2005_10	2005_11	2005_12	2006_01	2006_02	2006_03	2006_04	2006_05	2006_06	2006_07	2006_08	2006_09	2006_10	2006_11	2006_12
Agua potable	10	2	2	4	16	12	10	2	5	4	6	7	7	6	1	5

## 11.4 Requerimientos para Anfitrión

*Dirección de Estudios, Planeamiento y Desarrollo*  
 Requerimientos para el SARA.



### Requerimientos Cubo para el Módulo anfitrión

#### **Información usada dentro de los reportes**

Los reportes incluyen generalmente varios tipos de información que corresponde a los campos por los que son agrupados. Los datos son:

- Tiempo de estadía: entendido como el tiempo entre la llegada y salida del consumidor
  - Promedio de acuerdo a la variable por la que se esté agrupando

#### **Tiempo de estadía por tipo de caso**

	Ene-06	Feb-06	Mar-06	Abr-06	May-06	Jun-06	Jul-06	Ago-06	Sep-06	Oct-06	Nov-06	Dic-06
Asesorías	273	210	202	181	280	244	260	246	269	331	337	211
Derivaciones	85	56	96	50	72	140	105	107	147	427	395	252
Gestiones	55	42	62	30	45	37	63	44	100	170	155	125
Denuncias	38	38	60	31	56	27	28	33	28	34	31	16
<b>Total</b>	<b>451</b>	<b>346</b>	<b>420</b>	<b>292</b>	<b>453</b>	<b>448</b>	<b>456</b>	<b>430</b>	<b>544</b>	<b>962</b>	<b>918</b>	<b>604</b>

- Tiempo de entrevista: entendido como la duración de la entrevista con el consumidor
  - Promedio de acuerdo a la variable por la que se esté agrupando

#### **Campos usados para agrupar los casos**

A pesar de haber una serie de reportes regulares que son solicitados a la Dirección de Estudios, la variedad de solicitudes de información es grande, y se ha determinado que lo más conveniente es que se cuente con categorías por las cuales se pueda agrupar la información para ser colocada en columnas o filas en varios niveles (dos o más categorías en filas o columnas)

- Centro de solución de Controversias: es el centro que recibe la visita
- Fecha de ingreso del caso
- ID visita
- Ticket: ayudaría a identificar cada caso individualmente
- Técnico receptor
- Tiempos de espera (hora de salida – hora de entrada)

## 11.5 Requerimientos para Call Center

*Dirección de Estudios, Planeamiento y Desarrollo*  
 Requerimientos para el SARA



### Requerimientos Cubo para módulo del Call Center

#### Información usada comúnmente dentro de los reportes

Los reportes incluyen generalmente varios tipos de información que corresponde a los campos por los que son agrupados. Los datos son:

- Conteo de casos
  - Total de casos

Ejemplo:

#### Tipos de Atención

	Ene-06	Feb-06	Mar-06	Abr-06	May-06	Jun-06	Jul-06	Ago-06	Sep-06	Oct-06	Nov-06	Dic-06
Asesorías	273	210	202	181	280	244	260	246	269	331	337	211
Orientaciones de consumo	85	56	96	50	72	140	105	107	147	427	395	252
Información	55	42	62	30	45	37	63	44	100	170	155	125
Llamadas inútiles	38	38	60	31	56	27	28	33	28	34	31	16
<b>Total</b>	<b>451</b>	<b>346</b>	<b>420</b>	<b>292</b>	<b>453</b>	<b>448</b>	<b>456</b>	<b>430</b>	<b>544</b>	<b>962</b>	<b>918</b>	<b>604</b>

- Monto reclamado
  - Total del monto reclamado por variables
  - Promedio del monto reclamado por las variables que se agrupe
- Monto recuperado
  - Total del monto recuperado por variables
  - Promedio del monto recuperado por las variables que se agrupe

Ejemplo:

#### Montos reclamados por año y sector

	2007	2008	2009
Agua Potable	\$1	\$1	\$1
Alimentos	\$1	\$1	\$1
Comercio	\$1	\$1	\$1
Electrodomésticos	\$1	\$1	\$1
Energía Eléctrica	\$1	\$1	\$1
Gobierno	\$1	\$1	\$1

*Dirección de Estudios, Planeamiento y Desarrollo*  
 Requerimientos para el SARA



Hidrocarburos	\$1	\$1	\$1
Inmuebles	\$1	\$1	\$1
Libros	\$1	\$1	\$1
Medicamentos	\$1	\$1	\$1
Muebles	\$1	\$1	\$1
Publicidad	\$1	\$1	\$1
Servicios	\$1	\$1	\$1
Sistema Financiero	\$1	\$1	\$1
Telecomunicaciones	\$1	\$1	\$1
Turismo	\$1	\$1	\$1
Varios	\$1	\$1	\$1
Vehículos	\$1	\$1	\$1
<b>Total</b>	\$1	\$1	\$1

**Campos usados comúnmente para agrupar los casos**

A pesar de haber una serie de reportes regulares que son solicitados a la Dirección de Estudios, la variedad de solicitudes de información es grande, y se ha determinado que lo más conveniente es que se cuente con categorías por las cuales se pueda agrupar la información para ser colocada en columnas o filas en varios niveles (dos o más categorías en filas o columnas)

- Código de caso
- Tipo de caso
  - Asesorías
  - Orientaciones de consumo
  - Denuncias no personales
  - Gestiones
  - Consultas de caso
  - Información
  - Llamadas inútiles
  - Quejas del Call Center
  - Quejas del sector Transporte
  - Llamadas no finalizadas
- Subtipo de caso: algunos tipos de caso cuentan con subtipos
  - Orientaciones de consumo
    - Precios Gasolina
    - Precios Canasta Básica
    - Tasas de Interés
    - Útiles Escolares

*Dirección de Estudios, Planeamiento y Desarrollo*  
Requerimientos para el SARA



- Paquetes Vacacionales
- Información
  - Ley de Protección al Consumidor
  - Información de Oficinas Defensoria
  - Requisitos para interponer denuncia
  - Asesorías otras Leyes e Instituciones
  - Llamadas interna y/o transferencias
  - Información de otras Instituciones
  - Gestiones 910
  - Comentarios
  - Voice Mail
- Fecha de ingreso
  - Día de ingreso
  - Semana de ingreso
  - Mes de ingreso
  - Año de ingreso
- Forma de recepción:
  - Buzón de voz
  - Centro de Mediación-PGR
  - Escrita
  - Telefónica
  - Verbal
- Sector
- Forma de Archivo: en el caso de las gestiones se distingue si fueron archivadas favorable o desfavorablemente
- En el caso de las gestiones, procedencia como denuncia: las gestiones tienen la cualidad de poder convertirse en denuncias bajo ciertas condiciones y esta columna lo indica
- Con los casos consultados que requieran respuesta por parte del técnico del CSC, es conveniente contar con los campos: fecha de registro.
- Categoría: normalmente agrupado dentro de su sector correspondiente
- Motivo: normalmente agrupado dentro de su sector correspondiente o por proveedor.

*Se requiere que los motivos puedan ser agrupados en categorías más reducidas que serían definidas por el área de estudios*

*Dirección de Estudios, Planeamiento y Desarrollo*  
Requerimientos para el SARA



- Proveedor: Se requiere que los proveedores puedan ser agrupados en categorías más reducidas que serían definidas por el departamento de estudios
- Técnico receptor
- País
- Departamento (de El Salvador): generalmente en conjunto con sector
- Municipio: generalmente agrupado dentro del departamento al que corresponde
- País
- Sexo del consumidor

## **Ejemplos de los resultados de los cruces**

A continuación, algunos ejemplos del tipo de información

### **Análisis de la problemática del consumidor**

Este tipo de análisis tiene como propósito dar una alerta temprana del comportamiento de los casos recibidos en la Defensoría y resumir la problemática más reciente. Por ello, este tipo de reportes es el que exige la mayor flexibilidad en cuanto a los filtros y maneras de agrupar la información, pues se trata de la parte más dinámica de la labor del área de Estudios.

Este tipo de reportes, idealmente, deben de permitir que se resuma los datos al agruparlos de acuerdo a varias categorías. Por ejemplo, es posible que en algunos casos sea prioritario el observar la tendencia histórica de un cierto tipo de casos, en otras, podría necesitarse determinar si una zona geográfica recibe una mayor proporción de denuncias.

### **Casos por sector y motivo**

Este tipo de reporte es el que utiliza el “agrupamiento arbitrario” que propone el área de estudios. Una peculiaridad de la manera en la que se reciben los casos, es que los consumidores exigen que se sea muy específico en cuanto al motivo por el que se ha presentado un caso a la Defensoría. Sin embargo, se desea poder agrupar algunos motivos en categorías más grandes pues poseen características que los hacen muy similares.

Lamentablemente, no todos los motivos han sido clasificados en grupos coherentes, y de acuerdo con el propósito con el que estos se estudien, estos grupos están sujetos a cambios conforme se incorporen nuevos motivos o que estos comiencen a ganar mayor importancia.

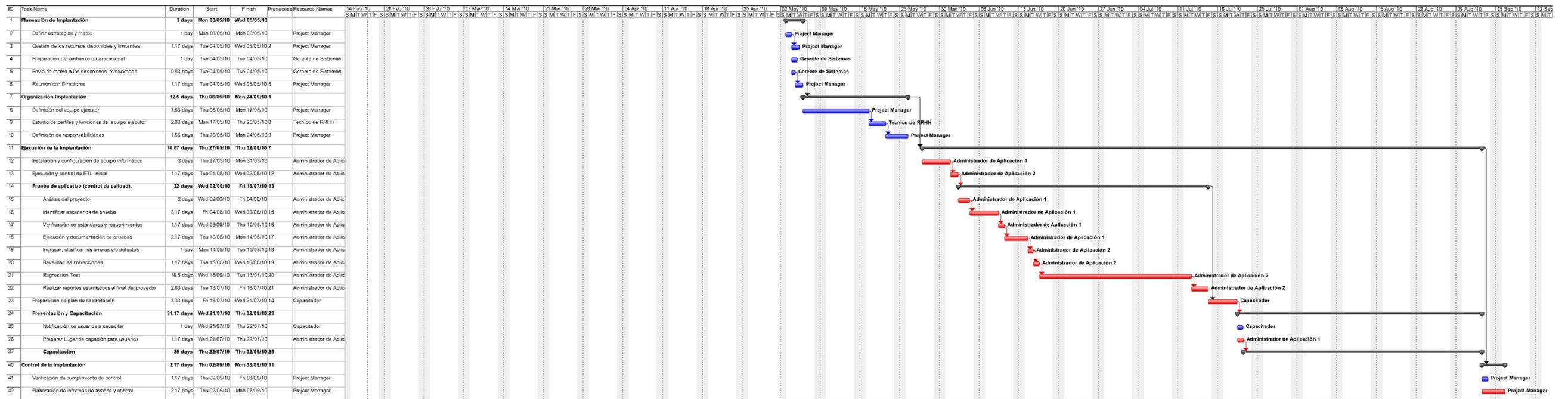
*Dirección de Estudios, Planeamiento y Desarrollo*  
 Requerimientos para el SARA



Motivo Asesoría	Total
<b>COBROS Y CARGOS</b>	
COBRO INDEBIDO	
COBRO EXCESIVO POR ALTO CONSUMO	
CONSUMO EXCESIVO	
COBRO POR CAMBIO DE TARIFA	
CARGO INDEBIDO	
COBRO POR CONSUMO ESTIMADO	
COBRO POR PÉRDIDA DE SUBSIDIO	
COBRO POR RECONEXIÓN	
<b>CALIDAD DEL SERVICIO</b>	
MAL SERVICIO	
MALA CALIDAD	
<b>MEDIDOR</b>	
COBRO POR MEDIDOR DEFECTUOSO	
COBRO POR MEDIDOR DESTRUIDO	
<b>VARIOS</b>	
MOTIVOS VARIOS	
PLAN DE PAGOS	
EMBARGO	
INCUMPLIMIENTO DE CONTRATO	
PUBLICIDAD ENGAÑOSA	
ENGAÑO	
<b>Total general</b>	

ANEXO 12: Diagramas de Plan de Implantación

12.1 Diagrama de Gantt



12. 2 Diagrama PERT

