

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



DESARROLLO DE UN MODELO MULTIDIMENSIONAL PARA EL SISTEMA DE
ATENCIÓN A RECLAMOS Y ASESORÍAS DE LA DEFENSORÍA DEL
CONSUMIDOR

PRESENTADO POR:

IVIS YOSSELIN NAVARRO RAMÍREZ
JORGE ROBERTO IRAHETA VÁSQUEZ
OFILIO ESAÚ SERPAS JIMÉNEZ
SAMUEL ALEXANDER NUÑEZ FLORES

PARA OPTAR AL TITULO DE:

INGENIERO DE SISTEMAS INFORMÁTICOS

DOCENTE ASESOR:

MSC. ELMER ARTURO CARBALLO RUIZ

CIUDAD UNIVERSITARIA, SEPTIEMBRE DE 2019

UNIVERSIDAD DE EL SALVADOR

RECTOR:

MSc. ROGER ARMANDO ARIAS ALVARADO

SECRETARIO GENERAL:

MSc. CRISTOBAL HERNAN RIOS BENITEZ

FACULTAD DE INGENIERIA Y ARQUITECTURA

DECANO:

ING. FRANCISCO ANTONIO ALARCON SANDOVAL

SECRETARIO:

ING. JULIO ALBERTO PORTILLO

ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

DIRECTOR:

ING. JOSÉ MARÍA SÁNCHEZ CORNEJO

DOCENTE ASESOR:

MSC. ELMER ARTURO CARBALLO RUIZ

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERIA Y ARQUITECTURA
ESCUELA DE INGENIERIA DE SISTEMAS INFORMATICOS

Trabajo de Graduación previo a la opción al Grado de:
INGENIERO DE SISTEMAS INFORMATICOS

Título :

DESARROLLO DE UN MODELO MULTIDIMENSIONAL PARA EL SISTEMA DE
ATENCIÓN A RECLAMOS Y ASESORÍAS DE LA DEFENSORÍA DEL CONSUMIDOR

Presentado por:

JORGE ROBERTO IRAHETA VÁSQUEZ

IVIS YOSSELIN NAVARRO RAMÍREZ

SAMUEL ALEXANDER NUÑEZ FLORES

OFILIO ESAÚ SERPAS JIMÉNEZ

Trabajo de Graduación Aprobado por:

Docente Asesor:

MSC. ELMER ARTURO CARBALLO RUIZ

SAN SALVADOR, SEPTIEMBRE DE 2019

Trabajo de Graduación Aprobado por:

Docente Asesor:

MSC. ELMER ARTURO CARBALLO RUIZ

AGRADECIMIENTOS

Al pensar en cada uno de los días atrás, todo el esfuerzo realizado hasta el día de hoy, los logros importantes de mi vida y este trabajo de tesis definitivamente pienso en que no hubieran podido ser realizados o hubiese sido mucho más difícil sin el apoyo de muchas personas importantes. Hoy deseo darles las gracias a todos ellos(as) por haberme ayudado tanto directa o indirectamente, consciente o inconscientemente a este gran logro profesional en mi vida.

Quiero agradecer primeramente a Dios, porque en los momentos más difíciles estuvo conmigo, en esos momentos incluso en los que casi me doy por vencido, fuiste tu quien movió todo para que yo llegue a donde estoy, gracias por haberme enseñado que nada en esta vida es pura casualidad todo tiene un propósito y es fruto de nuestros actos, esfuerzos, pensamientos y emociones.

También quiero agradecer a mi padre Rosendo Nuñez Canales, a quien admiro mucho como persona por su perseverancia y esfuerzo, le agradezco por haber trabajado fuertemente para que yo pueda convertirme en lo que soy. Te admiro porque, aunque no tuviste la oportunidad de tener estudio, hiciste todo lo posible por brindarme dicha oportunidad. Por lo que considero que este mérito y fruto no es solo mío, sino más bien de ambos. Has sido mi fuente de inspiración y el principal motivo por el cual ahora estoy escribiendo estas palabras, y en este documento.

Agradezco a mi madre Ana Margarita Flores Portillo por apoyarme en los momentos que más necesite su ayuda, sus consejos y por sus oraciones que sin duda alguna fueron escuchadas. Por lo que también pienso que este mérito es de ambos. Gracias también a mi familia, Mauricio Flores, Nancy Nuñez, Jaqueline Raymundo, Liseth Iraheta.

Agradezco a todos mis amigos, en especial a Delmy Stefany Menjívar Sandoval y José Manuel Díaz Hernández que de forma directa e indirecta fueron de gran apoyo intelectual y emocional a lo largo de mi carrera y de quienes he aprendido muchas cosas de la vida que me han ayudado a crecer como persona.

Gracias a mis compañeros de estudio y amigos Ivis Navarro, Jorge Iraheta, Esaú Serpas, Juan Diego, Ruth Torrento, Elvis Ticas, Fernando Bolaños, Nelson Rubio, Carlos Hernández con quienes compartí buenos momentos en la Universidad y fuera de ella.

Gracias todos los miembros de la Defensoría del Consumidor que nos apoyaron en el proceso de graduación. Gracias al Licenciado Jorge Salvador Pocasangre, actual gerente de Informática de la Defensoría del Consumidor quien nos dio seguimiento y apoyo durante todo el trabajo de tesis, le agradezco también por el conocimiento de vida y conocimiento técnico compartido mientras fue mi jefe. Gracias a Alexander Rodríguez, quien también nos apoyó activamente con conocimientos técnicos y consejos para el desarrollo de este proyecto.

Gracias al Ingeniero Elmer Arturo Carballo docente de la Universidad de El Salvador y asesor de este trabajo de graduación. Le agradezco primeramente como docente, ya que sin duda alguna lo admiro por el conocimiento técnico brindado lo cual me ha ayudado a crecer profesionalmente y también le agradezco por su apoyo técnico brindado durante nuestro trabajo, gracias por darnos ánimos, antes, durante y después de nuestro trabajo de grado y por ser el mediador de todo el equipo.

Samuel Nuñez

AGRADECIMIENTOS

«Un camino de mil millas comienza con un paso» Benjamín Franklin.

Agradecer primero a Dios y la Virgencita María por llegar al final de este primer paso profesional, por permitirme alcanzar mi título de Ingeniero de Sistemas Informáticos.

Agradecer a mis padres, **Olimpia Vásquez de Iraheta** y **Jorge Roberto Iraheta**, por su apoyo incondicional durante toda mi etapa de estudiante. Por sus palabras, apoyo, regaños y alientos a seguir, especialmente a mi mejor hermana que tengo, **Beatriz Iraheta**, quién es una parte fundamental en mi vida y a quién debo todo mi esfuerzo y entrega por ser un ejemplo para ella. ¡Los quiero mi familia!

Agradezco a muchos de mis compañeros con quienes tuvimos que sufrir, llorar y en muchas veces reír y ahora, definitivamente, celebrar que ya se ha alcanzado la meta, amigos que se convirtieron en familia como la Inga. **Ingrid Aguilar** y la Inga. **Vanessa Campos**, quienes fueron mis compañeras, mis amigas y mis hermanas en muchos tramos de la carrera. Así mismo, a **Carolina Alvarado**, **María Bernal**, Ing. **Irvin Leiva**, Ing. **Miguel Mártir**, **Sixto Rosales**, etc. Una mención especial al Milagro del grupo, **Denisse Mejía**, por su apoyo y su cercanía con mi familia.

Agradezco a la persona que me conoce sentimentalmente, a quién me ha ayudado a ser una mejor persona, quién me ha apoyado en mis decisiones desde que la conozco, para siempre tener cautela y los pies sobre la tierra. Le agradezco **Diana Linares** por haberme aguantado en toda esta etapa de tesis, sé que no ha sido fácil para ambos.

Al asesor el Ing. **Elmer Carballo**, por sus consejos profesionales, por su visión en el desarrollo de este trabajo y sus sugerencias en todo momento, manteniendo en todo momento unido este grupo de tesis.

Agradezco a mis compañeros en el desarrollo de este trabajo de tesis, **Ivis Navarro**, una gran amiga, compañera en muchos ámbitos y persona excepcional, a quién admiro mucho y sé que este será uno de muchos pasos en su carrera profesional. **Samuel Núñez**, una persona que pasamos de ser completamente desconocidos a grandes amigos, persona dedicada y profesional, el principal culpable que tomáramos este tema de tesis. **Esaú Serpas**, persona visionaria, con muchas capacidades y virtudes, enfocado en muchos proyectos con la realización de este trabajo de tesis.

En fin, agradezco a cada uno de todos los antes mencionados, por su apoyo incondicional, principalmente por creer y siempre apoyarme. ¡Infinitas Gracias a todos!

Jorge Iraheta

AGRADECIMIENTOS

«Todo ser humano, si se lo propone, puede ser escultor de su propio cerebro» Santiago Ramón y Cajal

Agradezco a Dios y la Virgen María, por el gran amor que tienen por todos, este logro no es más que una muestra más de su amor y misericordia.

A mis padres: mi madre Dinora del Carmen Ramírez, a quien amo y admiro tanto, por su fortaleza, paciencia, amor y apoyo incondicional. Y mi padre José Manuel Navarro, a quien amo y agradezco por su esfuerzo, apoyo y consejos. A ambos gracias por sus sacrificios y darme todo lo mejor que han podido.

A mi Abuela María Felipa, a mis hermanas Nancy y Mariela y mi sobrinito Santi, por su amor y por ser un apoyo y motivación para mí.

A Julio Pacheco, porque en tantos años compartidos se convirtió en alguien importante, agradezco por su tiempo, consejos, amor y por ser mi compañero y apoyo durante tantos años.

A mis amigos y equipo de trabajo de graduación: Jorge Iraheta, a quien admiro por su dedicación y buena voluntad, gracias por su amistad, apoyo y cariño.

Samuel Núñez, por su amistad, apoyo, paciencia, consejos y compartir sus conocimientos, siempre he admirado su disciplina y confianza.

Esaú Serpas, a quien admiro por su espíritu emprendedor, gracias por su amistad, consejos y apoyo. A los tres, gracias por todo el esfuerzo, dedicación y desvelos compartidos gracias a ello hoy finalmente podemos decir: ¡Lo logramos!

A mis amigos y compañeros de trabajo y estudios: Manuel Calderón, Josué Alemán, Marisol Franco, Néstor Guzmán, Luis Marroquín y Jairo Hernández por su apoyo en el trabajo y la carrera y por darme ánimos en días difíciles, gracias.

A mis docentes en la universidad: Ing. Elmer Carballo, quien fue asesor de este trabajo, a quien admiro como profesional y como persona, agradezco su guía y apoyo a lo largo del proyecto.

Ing. Damián Morales, por su apoyo en el trabajo, lo admiro por su disciplina y dedicación que lo hacen ser un excelente profesional.

Lic. Teresa Bautista a quien admiro por su inteligencia intelectual y emocional, por su apoyo, por su admirable labor y por darme herramientas tan valiosas para la vida como aprender que podemos manejar nuestros pensamientos para cambiar la forma en que nos sentimos ¡Gracias!

A la Defensoría del consumidor, en especial al Lic. Jorge Pocasangre, Ing. Alex Rodríguez, Ing. Juan José Rivas y todos los directores de las diferentes unidades, gracias por su apoyo y buena disposición a lo largo del proyecto.

Agradezco a cada uno de ellos por influir en mi vida, animarme y confiar en que era capaz de alcanzar mis metas aún cuando muchas veces yo no tuve confianza en mí. Gracias por ayudarme a saber que soy capaz.

Ivis Navarro

AGRADECIMIENTOS

«Si he visto más lejos es porque estoy sentado sobre hombros de gigantes», Isaac Newton

Desde que me embarqué en esta etapa de mi vida académica universitaria y culminando con este proyecto, muchas personas han contribuido de muchas maneras para que yo esté cerrando este ciclo y me permito estas líneas en nuestro trabajo final para agradecerles y desearles que la vida les dé mil veces lo que me han dado.

A mi papá Ofilio de Jesus Serpas por ser esa voz incansable que me ha alentado a seguir y culminar las metas que me propongo en la vida, tu has sido mi primer maestro y aun sigo aprendiendo de ti, gracias papá.

A mi mamá Ana Miladis Jimenez por que aun en la distancia has estado ahí para apoyarme en los momentos que más necesite.

A mi familia Lorena, Aurora y Oliver, por ser mi inspiración y motor para seguir cumpliendo metas, mis logros son para ustedes, Marielos, Diego, Victoria, por brindarme su apoyo, gracias.

A mis amigos René Martínez por motivarme y ayudarme para iniciar mis estudios y Celina Machado por el apoyo incondicional y creer en mí en cada momento, gracias.

A mis compañeros: Oscar, Javier, Rafael, Ulises, Walter, Delmy, Paty, Hector, Nelson, Catalina, Jorge, Mirna, Milton, Daniel, Carlos, Noel, Teresa, Samuel, Julisa, Karina, Aaron, Cesar y muchos más que compartimos cientos de horas en la biblioteca, clases, exámenes, laboratorios, proyectos sin duda aprendí mucho de cada uno, chicos gracias.

También a todos mis maestros, a los que se esforzaban por enseñarnos y daban esa milla más, muchas gracias, también a los que no, por que desarrolle otras habilidades de aprendizaje, gracias.

Y finalmente quiero agradecer a todos los involucrados en este proyecto:

A la Defensoría del consumidor por darnos la oportunidad de desarrollar este proyecto, a cada uno de los gerentes y directores que nos brindaron parte de su tiempo para hacer realidad este proyecto, al Lic. Salvador Pocasangre Gerente de informática por apoyarnos constantemente en el desarrollo del proyecto, a nuestro asesor y director de proyecto Ing. Elmer Carballo por guiarnos y apoyarnos oportunamente en cada etapa, a mis compañeros: Ivis, Jorge y Samuel el equipo de ensueño, esto no sería posible sin el trabajo de cada uno de ustedes, ha sido un privilegio trabajar con ustedes y a los que han estado tras bambalinas nuestras familias a todos mil gracias.

Esau Serpas

CONTENIDO

1	INTRODUCCIÓN.....	1
2	OBJETIVOS.....	3
2.1	Objetivo general.....	3
2.2	Objetivos específicos	3
3	JUSTIFICACIÓN.....	4
4	IMPORTANCIA	6
5	MARCO TEÓRICO.....	7
5.1	Data warehouse	7
5.1.1	Definición.....	7
5.1.2	Características	7
5.1.3	Data Marts	10
5.1.4	Enfoques de implementación de Data Warehouse.....	10
5.1.5	Arquitectura de Data Warehouses.....	14
5.1.6	Componentes de un Data Warehouse.....	16
5.2	Extracción, Transformación y Carga (ETL)	20
5.2.1	Extracción de Datos (Data Extraction).....	21
5.2.2	Transformación de Datos (Data Transformation)	22
5.2.3	Carga de Datos (Data Loading).....	23
5.3	Diseño multidimensional	23
5.3.1	Características del modelo multidimensional.....	24
5.3.2	Jerarquías y Niveles.....	26
5.4	OLTP.....	26
5.5	OLAP.....	26
5.5.1	Persistencia MOLAP, ROLAP, HOLAP	26
5.6	Dashboards.....	29
6	ANTECEDENTES.....	30
6.1	Defensoría del Consumidor	30
6.2	Unidades del Negocio	31
6.2.1	Dirección Jurídica.....	31
6.2.2	Dirección de Ciudadanía y Consumo.....	32
6.2.3	Tribunal sancionador.....	32
6.2.4	Centro de Solución de Controversias.....	33
6.2.5	Dirección de descentralización.....	33

6.2.6	Unidad de Análisis y Consumo de Mercados	33
7	SITUACIÓN ACTUAL	34
7.1	Defensoría del Consumidor	34
7.2	Unidades del Negocio	34
7.2.1	Dirección Jurídica.....	34
7.2.2	Dirección de Ciudadanía y Consumo.....	35
7.2.3	Tribunal Sancionador	35
7.2.4	Dirección del Centro de Solución de Controversias	37
7.2.5	Dirección de Descentralización.....	38
7.2.6	Unidad de Análisis y Consumo de Mercados	39
8	SOLUCIÓN PROPUESTA.....	41
8.1	Visión del Producto.....	41
8.2	Arquitectura propuesta.....	41
8.3	Orígenes de datos	42
8.4	Procesos ETL	42
8.5	Tratamiento de datos	42
8.6	Tecnología OLAP	42
8.7	MDX	42
8.8	SQL	42
9	METODOLOGÍA	43
9.1	Actividades de Desarrollo por Iteración (Sprint).....	43
9.2	Herramientas de Análisis y Diseño.....	43
9.3	Nomenclaturas	44
9.3.1	Nomenclatura de Paquetes de información.....	44
9.3.2	Nomenclatura de Diagramas mutidimensionales.....	44
9.3.3	Nomenclatura de Diagramas de Procesos ETL.....	45
9.3.4	Nomenclatura para Nombres de Procesos ETL	46
9.3.5	Nomenclatura para Nombres de JOBS ETL	47
10	PLANIFICACIÓN	49
10.1	Estructura organizacional del proyecto.....	49
10.2	Plan de Lanzamiento	50
11	SPRINT 0	51
11.1	Sprint Backlog.....	51
11.2	Descripción Historias de usuario	51
11.3	Arquitectura de Servidores Desplegada.....	52

11.4	Consolas de configuración	54
11.4.1	Servidor Pentaho.....	54
11.4.2	Servidor Staging	55
11.4.3	Servidor Windows	56
11.5	Diagrama de Componentes	57
12	SPRINT 1	58
12.1	Sprint Backlog.....	58
12.2	Descripción de Historias de Usuario	58
12.3	Refinamiento del Requerimiento de Información.....	62
12.3.1	Proceso BPMN de la Unidad.....	63
12.3.2	Paquetes de información.....	64
12.4	Especificación de Requerimientos	65
12.4.1	Casos de Uso Análisis de actas finalizadas en la Dirección Jurídica	65
12.4.2	Caso de Uso Análisis de casos finalizados en la Dirección Jurídica.....	65
12.5	Diseño Multidimensional.....	66
12.5.1	Diseño Conceptual del Data Mart.....	66
12.5.2	Diseño Físico Base de Datos Data mart	69
12.5.3	Diseño Físico Base de Datos Staging Area	70
12.6	Diseño de Procesos ETL	71
12.6.1	Procesos ETL Staging.....	71
12.6.2	Procesos ETL Data Mart	72
12.7	Pruebas	73
12.7.1	Pruebas de Procesos ETL	73
12.7.2	Pruebas Cubos OLAP	77
12.7.3	Pruebas Dashboard	79
13	SPRINT 2	81
13.1	Sprint Backlog.....	81
13.2	Descripción de Historias de Usuario	81
13.3	Refinamiento del Requerimiento de Información.....	85
13.3.1	Proceso BPMN de la Unidad.....	86
13.3.2	Paquetes de información.....	87
13.4	Especificación de Requerimientos	87
13.4.1	Caso de uso de cantidad de atenciones recibidas en la defensoría del consumidor. ...	87
13.5	Diseño Multidimensional.....	88
13.5.1	Diseño Conceptual del Data Mart.....	88

13.5.2	Diseño Físico Base de Datos Data mart	91
13.5.3	Diseño Físico Base de Datos Staging Area	92
13.6	Diseño de Procesos ETL	93
13.6.1	Procesos ETL Staging.....	93
13.6.2	Procesos ETL Data Mart	94
13.7	Pruebas	94
13.7.1	Pruebas de Procesos ETL	94
13.7.2	Pruebas Cubos OLAP	98
13.7.3	Pruebas Dashboard	99
14	SPRINT 3	101
14.1	Sprint Backlog.....	101
14.2	Descripción de Historias de Usuario	101
14.3	Refinamiento del Requerimiento de Información.....	106
14.3.1	Proceso BPMN de la Unidad.....	106
14.3.2	Paquetes de información.....	108
14.4	Especificación de Requerimientos	110
14.4.1	Caso de Uso: análisis de denuncias ingresadas en Tribunal Sancionador.....	110
14.4.2	Caso de Uso: análisis de resoluciones emitidas en Tribunal Sancionador	111
14.4.3	Caso de Uso: análisis de Sanciones impuestas en Tribunal Sancionador	112
14.5	Diseño Multidimensional.....	112
14.5.1	Diseño Conceptual del Data Mart.....	112
14.5.2	Diseño Físico Base de Datos Data mart	120
14.5.3	Diseño Físico Base de Datos Staging Area	122
14.6	Diseño de Procesos ETL	123
14.6.1	Procesos ETL Staging.....	123
14.6.2	Procesos ETL Data Mart	124
14.7	Pruebas	124
14.7.1	Pruebas de Procesos ETL	124
14.7.2	Pruebas Cubos OLAP	128
14.7.3	Pruebas Dashboard	130
15	SPRINT 4	132
15.1	Sprint Backlog.....	132
15.2	Descripción de Historias de Usuario	132
15.3	Refinamiento del Requerimiento de Información.....	137
15.3.1	Proceso BPMN de la Unidad.....	138

15.3.2	Paquetes de información.....	139
15.4	Especificación de Requerimientos.....	142
15.4.1	Casos de Uso Análisi de actas finalizadas en el Centro de Solución de Controversias 142	
15.4.2	Caso de uso cantidad de casos cerrados	143
15.4.3	Caso de uso de cantidades de casos reaberturados	144
15.5	Diseño Multidimensional.....	145
15.5.1	Diseño Conceptual del Data Mart.....	145
15.5.2	Diseño Físico Base de Datos Data mart	154
15.5.3	Diseño Físico Base de Datos Staging Area	157
15.6	Diseño de Procesos ETL.....	158
15.6.1	Procesos ETL Staging.....	158
15.6.2	Procesos ETL Data Mart	159
15.7	Pruebas.....	159
15.7.1	Pruebas de Procesos ETL.	159
15.7.2	Pruebas Cubos OLAP	163
15.7.3	Pruebas Dashboard	166
16	SPRINT 5	168
16.1	Sprint Backlog.....	168
16.2	Descripción de Historias de Usuario	168
16.3	Refinamiento del Requerimiento de Información.....	173
16.3.1	Proceso BPMN de la Unidad.....	174
16.3.2	Paquetes de información.....	175
16.4	Especificación de Requerimientos.....	177
16.4.1	Caso de uso cantidad de atenciones recibidas	177
16.4.2	Caso de uso cantidad de casos cerradas.....	178
16.5	Diseño Multidimensional.....	179
16.5.1	Diseño Conceptual del Data Mart.....	179
16.5.2	Diseño Físico Base de Datos Data mart	184
16.5.3	Diseño Físico Base de Datos Staging Area	186
16.6	Diseño de Procesos ETL.....	187
16.6.1	Procesos ETL Staging.....	187
16.6.2	Procesos ETL Data Mart	187
16.7	Pruebas.....	189
16.7.1	Pruebas de Procesos ETL	189

16.7.2	Pruebas Cubos OLAP	191
16.7.3	Pruebas Dashboard	193
17	SPRINT 6	195
17.1	Sprint Backlog.....	195
17.2	Descripción de Historias de Usuario	195
17.3	Refinamiento del Requerimiento de Información.....	200
17.3.1	Proceso BPMN de la Unidad.....	201
17.3.2	Paquetes de información.....	203
17.4	Especificación de Requerimientos	206
17.4.1	Caso de uso cantidad de atenciones recibidas	206
17.4.2	Caso de uso análisis de cantidad y montos de los casos.....	207
17.5	Diseño Multidimensional.....	208
17.5.1	Diseño Conceptual del Data Mart (UML).....	208
17.5.2	Diseño Físico Base de Datos Data mart	212
17.5.3	Diseño Físico Base de Datos Staging Area	214
17.6	Diseño de Procesos ETL.....	215
17.6.1	Procesos ETL Staging.....	215
17.6.2	Procesos ETL Data Mart	216
17.7	Pruebas.....	218
17.7.1	Pruebas de Procesos ETL	218
17.7.2	Pruebas Cubos OLAP	220
17.7.3	Pruebas Dashboard	221
18	ENTREGABLES.....	223
19	CONCLUSIONES.....	224
20	RECOMENDACIONES	225
21	BIBLIOGRAFÍA.....	226
22	GLOSARIO.....	228
23	ANEXOS.....	229
23.1	Anexo 1: Visión del producto.	229
23.2	Anexo 2: Documentación de Excepciones.....	230

1 INTRODUCCIÓN

Ante un mundo que evoluciona día a día, la necesidad de información precisa, clara y concisa se vuelve vital para las empresas. El continuo desarrollo en el área social de nuestro país, las actividades económicas de venta, compra, 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, así mismo, acercando los servicios que brindan a la comunidad, profundizando la vigilancia de los proveedores y los mercados, promoviendo así, la educación y participación ciudadana, para contribuir a un mejor funcionamiento del mercado.

Actualmente, los servicios de asesoría, información y recepción de denuncias de la Defensoría del Consumidor cubren gran parte de la población, posee 25 centros de atención a lo largo del país, además ofrecen sus servicios de recepción de denuncias en ventanillas móviles y poseen una plataforma de defensoría en línea. Esta cobertura de atención a las quejas de los salvadoreños en materia de consumo, significa altas cifras de atenciones, denuncias y asesorías.

Debido a este gran crecimiento en el volumen de información manejada en la institución, y como parte de su visión estratégica organizacional a futuro, la Defensoría se ve en la necesidad de una solución informática que permita sacar el máximo provecho a la información que se genera en materia de resolución de casos dentro de la institución, logrando satisfacer las necesidades de información de las Direcciones, en la toma de decisiones de calidad en el momento oportuno, con beneficio de la institución y en general de los consumidores salvadoreños.

El presente documento describe, el proyecto denominado “**Desarrollo de un modelo multidimensional para el sistema de atención a reclamos y asesorías de la defensoría del consumidor**”. En el desarrollo de este proyecto se hizo uso de una Metodología Ágil, la cual divide el desarrollo en diferentes lanzamientos o entregables y estos a su vez se dividen en Sprints. El proyecto se desarrolló en dos lanzamientos, cada uno con tres Sprints, uno por cada unidad de las que se trabajó, siendo estas: Dirección Jurídica, Ciudadanía y Consumo, Tribunal Sancionador, Centro de Solución de controversias, Descentralización y Unidad de Análisis y Consumo de Mercados.

La entrevista fue una de las principales técnicas de investigación, con la ayuda de herramientas y técnicas como la elaboración de paquetes de información, y el modelado de procesos BPMN, se logra identificar las necesidades de información de cada una de las unidades involucradas. Por cada unidad se ha construido uno o más paquetes de información, los cuales describen la información relevante para cada una de las Unidades de la Defensoría.

Identificadas las necesidades de información por cada unidad mediante los paquetes de Información, se da paso a los diagramas físicos de base de datos de las tablas fuentes, que dan respuesta a dichas necesidades. Para ello se realizaron las consultas pertinentes a las bases de datos fuentes, con la ayuda de la unidad de Informática, logrando así, la claridad en la búsqueda de los datos en el Sistema de Atención a Reclamos y Asesorías (SARA) de la Defensoría del Consumidor.

Asimismo, se presenta todo el proceso de elaboración de los almacenes de datos para cada unidad, siguiendo con la metodología predefinida donde se retoman las actividades de: Identificación de datos fuentes, diseño y desarrollo de los procesos ETL, diseño y desarrollo de cubos de información, diseño y desarrollo de vistas (Dashboards) e integración y despliegue.

Con la integración y despliegue de las unidades presentadas se completa el modelo multidimensional, con el cual se dará solución a las necesidades de información, sirviendo de apoyo a las jefaturas de la Defensoría del Consumidor para sus procesos de toma de decisiones.

2 OBJETIVOS

2.1 Objetivo general

Desarrollar un Modelo Multidimensional sobre el Sistema de Atención a Reclamos y Asesorías de la Defensoría del Consumidor para integrar, depurar y consultar la información generada por las diferentes unidades.

2.2 Objetivos específicos

- Analizar las necesidades de información basado en preguntas de negocio para identificar la solución a las áreas de negocio de la Defensoría del Consumidor.
- Diseñar el Modelo Multidimensional para Dirección Jurídica (DJ), Dirección de Ciudadanía y Consumo (CYC), Tribunal Sancionador (TS), Dirección del Centro de Solución de Controversias (CSC), Dirección de Descentralización (DD) y la Unidad de Análisis y Consumo de Mercados (UACM).
- Construir el proceso de extracción, transformación y carga (ETL) para Dirección Jurídica (DJ), Dirección de Ciudadanía y Consumo (CYC), Tribunal Sancionador (TS), Dirección del Centro de Solución de Controversias (CSC), Dirección de Descentralización (DD) y la Unidad de Análisis y Consumo de Mercados (UACM).
- Diseñar un esquema de infraestructura que soporte el modelo multidimensional para el despliegue y operación del sistema.
- Unificar un repositorio de datos de valor accesible para los tomadores de decisiones de las diferentes unidades de negocio.
- Implementar los cubos de datos para la generación de la reportería OLAP que contribuya a la toma de decisiones de las jefaturas y direcciones de la Defensoría del Consumidor.

3 JUSTIFICACIÓN

La Defensoría del Consumidor, en su labor de proteger y promover los derechos de los consumidores dentro de la República de El Salvador, ofrece sus servicios a través de diferentes canales y puntos de servicio distribuidos en todo el país, por su naturaleza y gracias a los mecanismos de protección del consumidor, sus servicios generan alta demanda, en consecuencia se apertura muchos casos para el seguimiento correspondiente y a lo largo del proceso se genera mucha información relevante, tanto de los casos y sus involucrados, como del proceso en sí.

Actualmente el principal sistema informático denominado SARA, es una plataforma modular utilizada en los procesos internos de la defensoría del consumidor, desplegado en todas sus puntos de servicio y conectando con los diferentes departamentos de la organización, SARA es la principal herramienta de recolección y procesamiento de información de acuerdo a los servicios y a la operatividad de la institución, almacenando de forma centralizada toda la información operativa generada por cada unidad.

La información que se genera de todos los servicios y procesos operativos es de vital importancia para la institución y la alta gerencia, sobre todo porque permite generar métricas que evalúen los resultados generados y también genera información relevante para la toma de decisiones.

Actualmente es la unidad de informática la que se encarga de solventar las necesidades de información, generando reportes para todas las áreas, estos reportes se generan a demanda de las diferentes gerencias y se manejan como solicitudes de desarrollos, estas llevan asociado un tiempo de desarrollo, por lo que la información no se obtiene de inmediato, además de presentarla en un formato complejo que requiere una experticia en tablas dinámicas para extraer información relevante para análisis.

Las 6 áreas que solicitan información para toma de decisiones con mayor frecuencia a la Unidad de Informática son:

- Dirección Jurídica
- Dirección de Ciudadanía y Consumo
- Tribunal Sancionador
- Centro de Solución de Controversias
- Dirección de Descentralización
- Unidad de Análisis de Consumo de Mercados

Todas las áreas manifiestan interés y apoyo total a una solución para esta problemática, puesto que la necesidad de tener una herramienta que permita la integración, depuración de la información y el análisis desde diferentes perspectivas es evidente para todos.

De las entrevistas realizadas en la investigación se obtuvo que las seis unidades encuestadas califican como ineficiente la elaboración actual de informes.

Las 6 unidades indican se emplea un promedio de 2 a 3 días (una carga de 20 horas laborales) para la obtención de los datos, preparación y elaboración de reporte.

Así mismo todas las unidades encuestadas respondieron que calificaban en total acuerdo a que se buscara una mejora en el proceso de obtención de informes para brindar un apoyo eficiente en la toma de decisiones.

Es por ello que el desarrollo de un modelo multidimensional es una solución adecuada para la institución, ya que no solo cumple con la necesidad de integración y depuración de la información, sino que además se prevé la reducción en los tiempos de obtención de información, lo que a su vez se convierte en una ventaja competitiva que permitirá agilizar los procesos de toma de decisiones, generando un impacto positivo en la institución.

4 IMPORTANCIA

Un proyecto de inteligencia de negocios, como un modelo multidimensional, aportará grandes beneficios a la Defensoría del Consumidor facilitando los procesos de toma de decisiones, ya que se logrará tener un repositorio de datos actualizados, integrados y accesibles para los usuarios tanto a nivel agregado como en detalle según sus necesidades, logrando mayor flexibilidad en el análisis de datos. El Modelo Multidimensional ayudará al proceso de toma de decisiones, logrando reducir considerablemente los tiempos en que se obtiene información para análisis y permitiendo que los usuarios se sirvan su propia información sin depender completamente del departamento de informática, optimizando, de esta forma, el proceso de generación de reportes.

5 MARCO TEÓRICO

5.1 Data warehouse

5.1.1 Definición

El término Data **Warehouse** (DW), en español **Almacén de Datos**, fue acuñado por primera vez por *Bill Inmon* en el año de 1990. *Bill Inmon* define un Data Warehouse como: “*una colección de datos orientada a un tema específico, integrado, variante en el tiempo y no volátil, que soporta el proceso de toma de decisiones*”. (Power Data, n.d.).

Sean Kelly, otro líder en las prácticas de Data Warehousing, define un Data Warehouse en base a las características de sus datos (Ponniiah, 2010).

- Aislada.
- Disponible.
- Integrada.
- Con marca de tiempo (Time Stamped).
- Orientado a un tema.
- No volátil.
- Accesible.

Por otro lado, *Ralph Kimball*, autor de la metodología bottom-up para la implementación de Data Warehouses, define un Data Warehouse como: “*La conglomeración de las áreas de preparación, almacenamiento y presentación donde los datos operacionales de una organización están específicamente estructurados para consultas, análisis, rendimiento y facilidad de uso*”.

Genéricamente un Data Warehouse es una colección de métodos, técnicas, y herramientas usadas para dar soporte a trabajadores del conocimiento (Gerentes Senior, Directores, Gerentes, y Analistas) en la realización de análisis de datos que ayuden a realizar procesos de toma de decisiones y mejorar los recursos de información. (Kimball & Ross, 2002)

Por tanto, un Data Warehouse es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta.

Para tener una visión más técnica y detallada de este concepto, en los siguientes párrafos se describirán, las características, componentes, y las arquitecturas existentes en los Data Warehouses.

5.1.2 Características

Algunas de las características claves de los Data Warehouses que los describen a *Warehouses* y que hacen que se diferencie de los sistemas de almacenamiento transaccionales tradicionales son:

- Datos orientados a uno o varios temas.
- Datos Integrados.
- Datos Variantes en el tiempo.
- Datos No Volátiles.
- Granularidad en los datos.

En los siguientes apartados se profundizará en cada una de las características que se han mencionado.

5.1.2.1 Datos orientados a uno o varios temas

En todas las industrias, los conjuntos de datos se organizan en torno a aplicaciones transaccionales individuales. Estos conjuntos de datos individuales tienen que proporcionar datos para las aplicaciones específicas para realizar las funciones de manera eficiente. Por lo tanto, los conjuntos de datos para cada aplicación deben organizarse en torno a esa aplicación específica. A diferencia de este enfoque, los Data Warehouses son organizados en base a uno o varios temas de interés de la compañía. La Figura 5.1 muestra gráficamente esta diferencia.

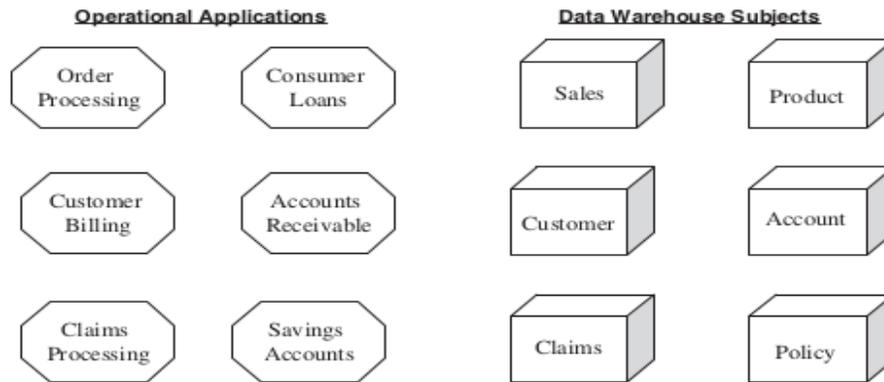


Figura 5.1: Diferencia entre las aplicaciones operacionales y los Data Warehouses

En un Data Warehouse los datos son almacenados por sujetos o eventos comerciales del mundo real, no por aplicaciones.

5.1.2.2 Datos integrados

Para una toma de decisiones adecuada, un Data Warehouse debe reunir todos los datos relevantes de las distintas aplicaciones. Los datos en el almacén de datos pueden provenir de varios sistemas operacionales. Los datos de origen normalmente residen en diferentes bases de datos, archivos y segmentos de datos. Las fuentes de datos normalmente son aplicaciones dispares, por lo que las plataformas operativas y los sistemas operacionales podrían ser diferentes. Los diseños de archivos, las representaciones de códigos de caracteres y las convenciones de nombres de campo podrían ser diferentes también. La Figura 5.2 muestra esta característica de forma gráfica.

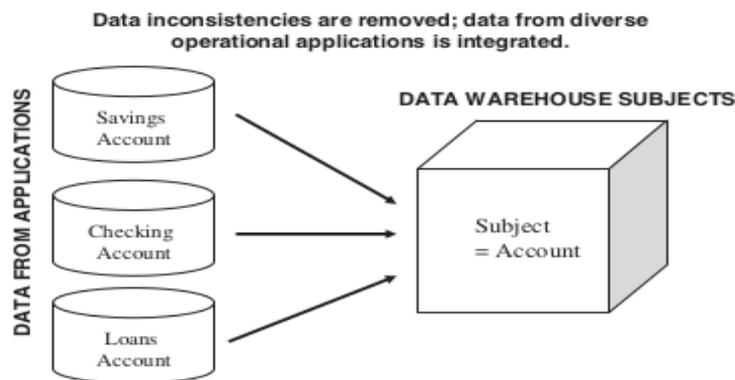


Figura 5.2: Característica de integración de los datos en un Data Warehouse

5.1.2.3 Datos variantes en el tiempo

Para un sistema operacional, los datos almacenados contienen los valores actuales. Por supuesto, algunas transacciones del pasado también son almacenados en sistemas operacionales, pero, esencialmente, los sistemas operacionales reflejan la **información actual** porque estos sistemas respaldan las operaciones corrientes **diarias**.

Por otro lado, los datos en el almacén de datos están destinados para el análisis y la toma de decisiones. Un almacén de datos, debido a la naturaleza misma de su propósito, debe contener **datos históricos**, no solo valores actuales. Los datos se almacenan como instantáneas sobre períodos **pasados y actuales**. Los cambios en los datos se rastrean y registran de manera que, si es necesario, se pueden generar informes para mostrar los cambios a lo largo del tiempo. Cada estructura de datos en el almacén de datos contiene el elemento de tiempo. Este aspecto del almacén de datos es bastante significativo tanto para el diseño como para las fases de implementación.

La naturaleza variante en el tiempo de los datos en un almacén de datos permite:

- El análisis del pasado
- Relacionar la información con el presente
- Previsiones para el futuro

5.1.2.4 Datos no volátiles

Los datos extraídos de los diferentes sistemas operacionales y los datos obtenidos de otras fuentes son transformados, integrados y almacenados en el Data Warehouse. Los datos en el Data Warehouse **no** están diseñados para el funcionamiento diario del negocio. En el almacén de datos, mantiene los datos de estado de stock extraídos como instantáneas a lo largo del tiempo. Los datos en el almacén de datos **no** se actualizan cada vez que ocurre una transacción.

Los datos de los sistemas operacionales se trasladan al almacén de datos a intervalos específicos. Según los requisitos del negocio, estos movimientos de datos se pueden realizar por ejemplo dos veces al día, una vez al día, una vez a la semana o una vez cada dos semanas, dependiendo de los requerimientos del negocio. De hecho, en un almacén de datos típico, los movimientos de datos a diferentes conjuntos de datos pueden tener lugar en diferentes tiempos.

La Figura 5.3 demuestra la característica de no volatilidad en un Data Warehouse:

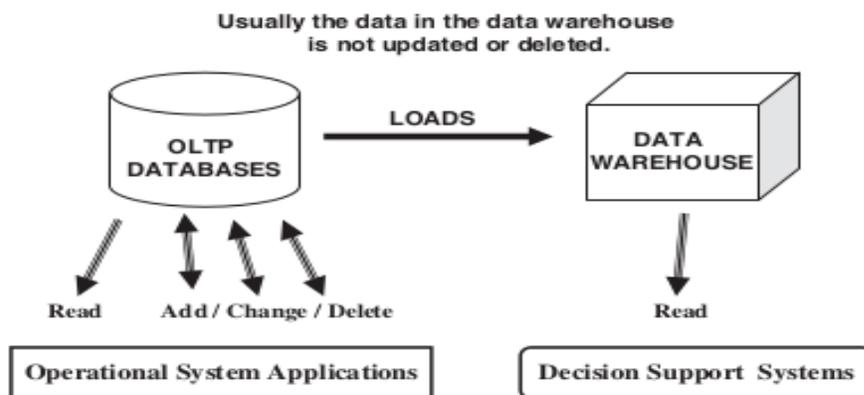


Figura 5.3: Característica de no volatilidad de los Data Warehouses.

Cada transacción del negocio **no actualiza** los datos en el almacén de datos. Las transacciones del negocio actualizan las bases de datos del sistema **operacional** en tiempo real. Se agregan, cambian o eliminan datos de un sistema operacional a medida que ocurre cada transacción, pero generalmente no se actualizan los datos en el almacén de datos.

En el almacén de datos, **no** se eliminan los datos en tiempo real. Una vez que los datos son capturados y confirmados en el almacén de datos, no se ejecutan transacciones individuales para cambiar los datos allí. Las actualizaciones de datos son comunes en una base de datos **transaccional**; no es así en un almacén de datos. Los datos en un almacén de datos no son tan **volátiles** como los datos en una base de datos transaccional. Los datos en un almacén de datos son principalmente para **consultas y análisis**.

5.1.2.5 Granularidad en los datos

En un sistema transaccional, los datos generalmente se mantienen en el nivel más bajo de detalle. Normalmente, en un sistema transaccional, no se guardan datos resumidos.

Por otro lado, en un almacén de datos, resulta eficiente **mantener los datos resumidos en diferentes niveles**. Dependiendo de la consulta, puede ir un nivel particular de detalle y satisfacer la consulta. **La granularidad de los datos en un almacén de datos se refiere al nivel de detalle**. Cuanto **menor** sea el nivel de detalle, más **fina** será la granularidad de los datos. Por supuesto, si se desea mantener los datos en el nivel más bajo de detalle, se deben almacenar una gran cantidad de datos en el almacén de datos. Tendrá que decidirse sobre los niveles de granularidad en función de los tipos de datos y el rendimiento esperado del sistema para las consultas.

5.1.3 Data Marts

Un Data Mart es una base de datos **departamental**, especializada en el almacenamiento de los datos de **un área de negocio específica**. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento (Sinnexus, 2016).

Una Data Mart es una versión especial de almacén de datos (Data Warehouse). Son subconjuntos de datos con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones. Los datos existentes en este contexto pueden ser agrupados, explorados y propagados de múltiples formas para que diversos grupos de usuarios realicen la explotación de los mismos de la forma más conveniente según sus necesidades (Wikipedia, 2017)

Los Data Marts son sistemas orientados a la consulta, en el que **se producen procesos batch de carga de datos (altas) con una frecuencia baja y conocida. Es consultado mediante herramientas OLAP** (On line Analytical Processing - Procesamiento Analítico en Línea) que ofrecen una visión multidimensional de la información.

En síntesis, se puede decir que los Data Marts son pequeños Data Warehouses centrados en un tema o un área de negocio específico dentro de una organización

5.1.4 Enfoques de implementación de Data Warehouse

Existen dos enfoques de implementación de los Data Warehouses, uno es Top-Down (de arriba hacia abajo) y el otro es Botton-Up (de abajo hacia arriba), (Ponniah, 2010).

Antes de comenzar todo proceso de construcción de Data Warehouse se debe responder a las siguientes preguntas:

ENFOQUE TOP-DOWN	ENFOQUE BOTTOM-UP
¿Debería ver el panorama general de la organización, adoptar un enfoque descendente y construir un enorme almacén de datos?	¿Debería adoptar un enfoque de abajo hacia arriba, ver los requerimientos departamentales y construir un Data Mart a la medida de cada departamento?
¿Debería construir un gran almacén de datos y luego dejar que ese repositorio alimente datos en los Data Marts departamentales?	¿Debería crear Data Marts locales individuales y combinarlos para formar un almacén de datos en general?
¿Deberían los Data Marts depender del almacén de datos general para la alimentación de datos en estos?	¿Deberían estos Data Marts ser independientes el uno del otro?

Tabla 5.1: Preguntas según enfoque de implementación de un Data Warehouse

5.1.4.1 Enfoque Top-Down

Bill Inmon es uno de los principales defensores del enfoque de arriba hacia abajo. El define un almacén de datos como un repositorio centralizado para toda la empresa. Con este enfoque los datos en el almacén de datos son almacenados al más bajo nivel de granularidad basado en un modelo de datos normalizado.

Bill Inmon ve la necesidad de transferir la información de los diferentes OLTP (Sistemas Transaccionales) de las organizaciones a un lugar centralizado donde los datos puedan ser utilizados para el análisis (sería **el CIF o Corporate Information Factory**). Insiste además en que ha de tener las siguientes características:

- **Orientado a temas.** Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.
- **Integrado.** La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.
- **No volátil.** La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo lectura, y se mantiene para futuras consultas.
- **Variante en el tiempo.** Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.

La información ha de estar a los máximos niveles de detalle. Los DW departamentales o Data Marts son tratados como subconjuntos de este DW corporativo, que son construidos para cubrir las necesidades individuales de análisis de cada departamento, y siempre a partir de este DW central (del que también se pueden construir los ODS (Operational Data Stores) o similares).

Según el punto de vista de Inmon, **el almacén de datos se encuentra en el centro de la "Fabrica de Información Corporativa" (Corporate Information Factory - CIF)** y proporciona el marco lógico para entregar inteligencia empresarial a la empresa. Las operaciones del negocio proporcionan datos para conducir el CIF. El almacén de datos centralizado alimentaría los Data Marts que pueden diseñarse basándose en un **modelo de datos dimensional**.

Las ventajas y desventajas de este enfoque se consolidan en la siguiente tabla:

ENFOQUE TOP-DOWN	
VENTAJAS	DESVENTAJAS
Un verdadero esfuerzo corporativo, una visión empresarial de los datos.	Toma más tiempo construir incluso con un método iterativo.
Intrínsecamente diseñado, no es una unión de Data Marts con datos dispares.	Alta exposición al riesgo de fallo.
Almacenamiento único y central de datos sobre el contenido.	Necesita un alto nivel de habilidades interfuncionales.
Control y reglas centralizadas.	Elevado gasto sin justificación de concepto.
Se puede ver resultados rápidos si se implementan con iteraciones	

Tabla 5.2: Ventajas y desventajas del enfoque Top-Down de un DW

Este es un enfoque de visión amplia en el que se construye un almacén de datos **global**, para toda la empresa. Aquí no se tiene una colección de **islas** fragmentadas de información. El almacén de datos es **grande e integrado**. Sin embargo, este enfoque lleva más tiempo en la construcción de un Data Warehouse y tiene un alto riesgo de falla. Si no se tienen profesionales con experiencia, este enfoque podría ser **peligroso**. Además, es un poco más difícil vender este enfoque a la gerencia superior y los patrocinadores, ya que no es probable que vean los resultados lo suficientemente pronto.

5.1.4.2 Enfoque Botton-Up

Ralph Kimball, otro autor principal y experto en el almacenamiento de datos, es un defensor del enfoque que se ha dado a conocer como el enfoque ascendente. Kimball, el cual, visualiza el almacén de datos corporativos como una colección conformada por Data Marts. En este enfoque, los Data Marts se crean primero para proporcionar capacidades analíticas y de informes para temas del negocio específicos basados en el modelo de datos dimensionales.

Los Data Marts contienen datos en el nivel más bajo de granularidad y también como resúmenes según las necesidades de análisis. Estos Data Marts son unidos de forma conjunta conformando las dimensiones. Las ventajas y desventajas de este enfoque se consolidan en la siguiente tabla:

ENFOQUE BOTTOM-UP	
VENTAJAS	DESVENTAJAS
Implementación más rápida y fácil de piezas manejables.	Cada Data Mart tiene su propia visión estrecha de los datos.
Rendimiento favorable de la inversión y prueba del concepto de costo.	Permea datos redundantes en cada Data Mart.
Menor riesgo de fracaso.	Perpetúa datos inconsistentes e irreconciliables.
Intrínsecamente incremental; puede programar Data Marts importantes primero.	Proliferan las interfaces inmanejables.
Permite que el equipo del proyecto aprenda y crezca.	

Tabla 5.3: Ventajas y desventajas del enfoque Bottom-Up de un DW

En el enfoque de abajo hacia arriba, se construyen Data Marts departamentales uno por uno, se crea un esquema de prioridades para determinar cuál Data Mart se debe construir primero. **El inconveniente más grave** de este enfoque es la fragmentación de datos. Cada centro de datos independiente estará ciego a los requisitos generales de toda la organización.

Según Kimball, el Data Warehouse es un conglomerado de todos los Data Marts dentro de una empresa, siendo una copia de los datos transaccionales estructurados de una forma especial para el análisis, de acuerdo al **Modelo Dimensional (no normalizado)**, que incluye las dimensiones de análisis y sus atributos, su organización jerárquica, así como los diferentes hechos de negocio que se quieren analizar. Por un lado, se tiene las tablas para representar las dimensiones y por otro lado las tablas para los hechos (las facts tables). Los diferentes Data Marts están conectados entre sí por la llamada bus structure, que contiene los elementos anteriormente citados a través de las dimensiones conformadas (que permiten que los usuarios puedan realizar consultas conjuntas sobre las diferentes Datas Marts, pues este bus contiene los elementos en común que los comunican).

Una dimensión conformada puede ser, por ejemplo, la dimensión cliente, que incluye todos los atributos o elementos de análisis referentes a los clientes y que puede ser compartida por diferentes Datas Marts (ventas, pedidos, gestión de cobros, etc.).

Además de proponer una arquitectura específica de Data Warehouse, Kimball, también provee un marco de referencia para la construcción de un Data Warehouse que puede ser utilizado como guía durante el desarrollo de este último, la Figura 5.4 muestra gráficamente este marco de trabajo. En este enfoque, luego de definir los requerimientos del negocio Kimball propone enfocar el proyecto en tres líneas (tracks):

- Tecnología. (cuadros amarillos) trata sobre la selección de las tecnologías a utilizar en el proyecto (Hardware, DBMS, Herramientas ETL, Herramientas de Consulta, Herramientas de Reportes, etc.).
- Datos. (cuadros naranja) trata sobre el modelado dimensional, modelo físico y de ETL, así como también su implementación.
- Aplicaciones. (cuadros verdes) trata sobre el diseño y/o desarrollo de las aplicaciones de BI e interfaces de navegación la búsqueda de aplicaciones candidatas etc.

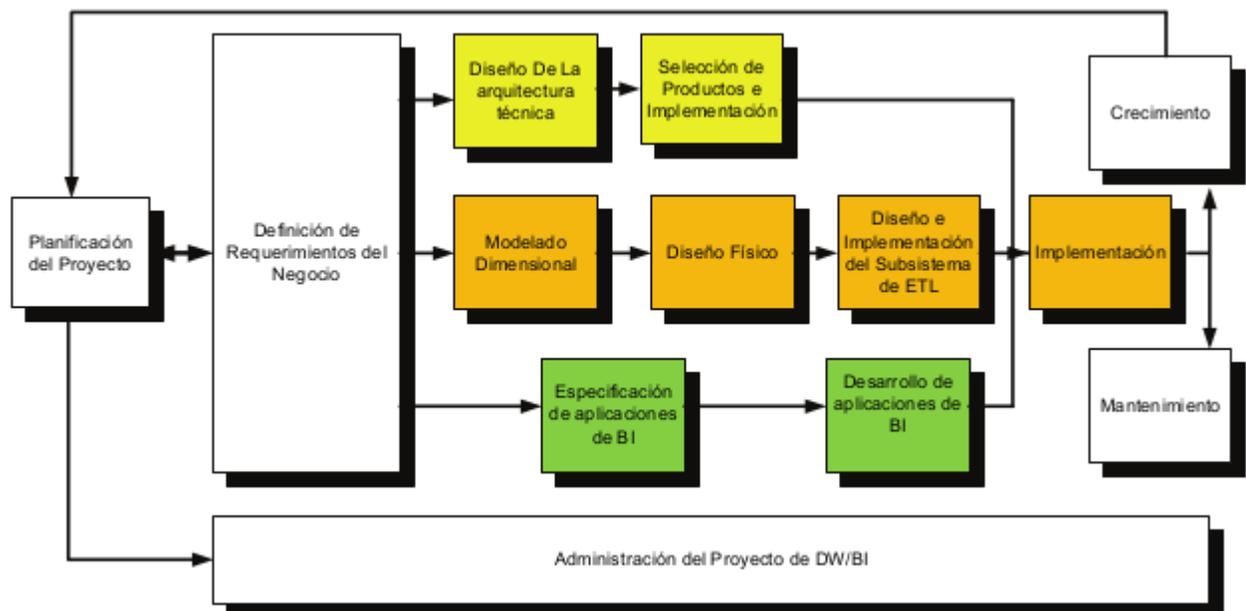


Figura 5.4: Framework propuesto por Kimball

5.1.5 Arquitectura de Data Warehouses

Como cualquier sistema, los Data Warehouses están compuestos por subsistemas o componentes. Con el paso de los años, en la práctica, varios arreglos de Data Warehouses y Data Marts han evolucionado en función de los requisitos individuales de las organizaciones. Cada uno de los tipos de arquitectura hace énfasis en la relación entre los Data Warehouses y los Data Marts.

5.1.5.1 Data Warehouses centralizados

Este tipo arquitectónico tiene en cuenta los requisitos de información a nivel empresarial. Se establece una infraestructura general. Datos atómicos al nivel más bajo de granularidad son almacenado en la tercera forma normal¹ (Véase Tabla 5.4). Ocasionalmente, se incluyen algunos datos resumidos. Las consultas y aplicaciones acceden a los datos normalizados en el almacén de datos central. No hay Data Marts separados. Figura 5.5 muestra gráficamente esta arquitectura:

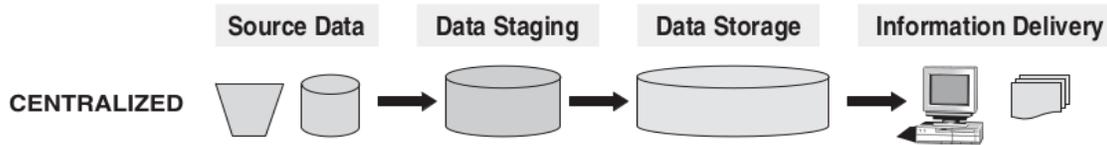


Figura 5.5 Arquitectura Centralizada de un Data Warehouse

REGLA	DESCRIPCIÓN
Primera Forma Normal (1FN)	Incluye la eliminación de todos los grupos de columnas y/o filas repetidas.
Segunda Forma Normal (2FN)	Asegura que todas las columnas que no son llave sean completamente dependientes de la llave primaria (Todas las dependencias parciales, las cuales son datos que no dependen de la llave primaria, son eliminadas y separadas dentro de sus propias tablas)
Tercera Forma Normal (3FN)	Elimina cualquier dependencia transitiva. Una dependencia transitiva es aquella en la cual las columnas que no son llave son dependientes de otras columnas que tampoco son llave.

Tabla 5.4: Descripción 1FN, 2FN y 3FN en una base de datos.

5.1.5.2 Data Marts independientes

Este tipo de arquitectura evoluciona en compañías donde las unidades organizativas desarrollan sus propios centros de datos (Data Marts) para sus propios fines específicos. Aunque cada centro de datos sirve a la unidad organizativa en particular, estos mercados de datos separados no proporcionan "una sola versión de la verdad". Los Data Marts son independientes entre sí. Como resultado, es probable que estos Data Marts diferentes tengan inconsistentes es y estándares. Tales variaciones impiden el análisis de datos a lo largo de los Data Marts. Por ejemplo, si hay dos Data Marts independientes, uno para ventas y otro para envíos, aunque las ventas y los envíos son temas relacionados, los Data Marts independientes dificultarían el análisis de ventas y envíos juntos. La Figura 5.6 muestra gráficamente esta arquitectura.

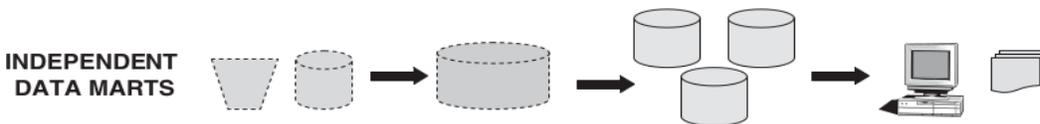


Figura 5.6: Arquitectura de Data Marts Independientes de un Data Warehouse

¹ En la teoría de bases de datos relacionales, las formas normales (NF) proporcionan los criterios para determinar el grado de vulnerabilidad de una tabla a inconsistencias y anomalías lógicas. Cuanta más alta sea la forma normal aplicable a una tabla, menos vulnerable será a inconsistencias y anomalías.

5.1.5.3 Federado

Algunas empresas ingresan al almacenamiento de datos con un legado existente de una variedad de estructuras de soporte de decisiones en forma de sistemas operacionales, conjuntos de datos extraídos, centros de datos primitivos, etc. Para tales compañías, puede no ser prudente descartar toda esa gran inversión y comenzar de cero.

La solución práctica es un tipo de arquitectura federada donde los datos pueden integrarse física o lógicamente a través de campos clave compartidos, metadatos globales generales, consultas distribuidas y otros métodos. En este tipo de arquitectura, no hay un almacén de datos general.



Tabla 5.5: Arquitectura Federada de un Data Warehouse.

5.1.5.4 Hub-and-Spoke

Este es el enfoque de Factoría de Información Corporativa de Inmon. De manera similar a la arquitectura centralizada del almacén de datos, aquí también se encuentra un almacén de datos general para toda la empresa. Los datos atómicos en el almacén de datos centralizado, se almacenan en la tercera forma normal. La principal y útil diferencia es la presencia de centros de datos dependientes en este tipo arquitectónico. Los Data Marts dependientes obtienen datos del almacén de datos centralizado.

El almacén de datos centralizado forma el centro para alimentar los datos a los Data Marts en los alrededores. Los Data Marts dependientes se pueden desarrollar para una variedad de propósitos: necesidades analíticas departamentales, consultas especializadas, extracción de datos, etc. Cada Data Mart dependiente puede tener estructuras de datos normalizadas, des normalizadas, resumidas o dimensionales basadas en requisitos individuales. La mayoría de las consultas se dirigen a los Data Marts dependientes, aunque el almacén de datos centralizado puede utilizarse para consultar.

Este tipo de arquitectura resulta de la adopción de un enfoque de arriba hacia abajo para el desarrollo del almacén de datos. La Figura 5.7 muestra gráficamente esta arquitectura.

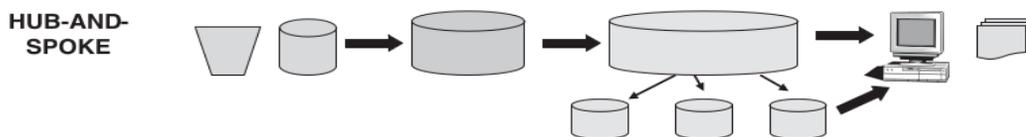


Figura 5.7: Arquitectura Hub and Spoke de un Data Warehouse.

5.1.5.5 Data Mart Bus

Este es el enfoque de la conformación de Supermarts de Kimball. Comienza por analizar los requisitos para un tema de negocio específico, como pedidos, envíos, facturación, reclamos de seguro, alquiler de automóviles, etc.

Se construye la primera plataforma de datos (supermart) usando dimensiones y métricas del negocio. Estas dimensiones del negocio serán compartidas por los Data Marts futuros. La noción principal es que

al conformar las dimensiones entre los diferentes Data Marts, el resultado sería supermarts lógicamente integrados que proporcionarían una visión empresarial de los datos.

Los mercados de datos contienen datos atómicos organizados como un modelo de datos dimensionales. Este tipo de arquitectura resulta de la adopción de un enfoque ascendente mejorado para el desarrollo de una Data Warehouse.

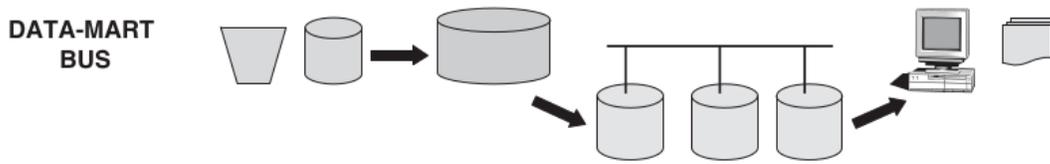


Figura 5.8: Arquitectura Data-Mart Bus de un Data Warehouse.

En la Figura 5.9 se muestra una representación gráfica de los diferentes tipos de arquitecturas en los Data Warehouses.

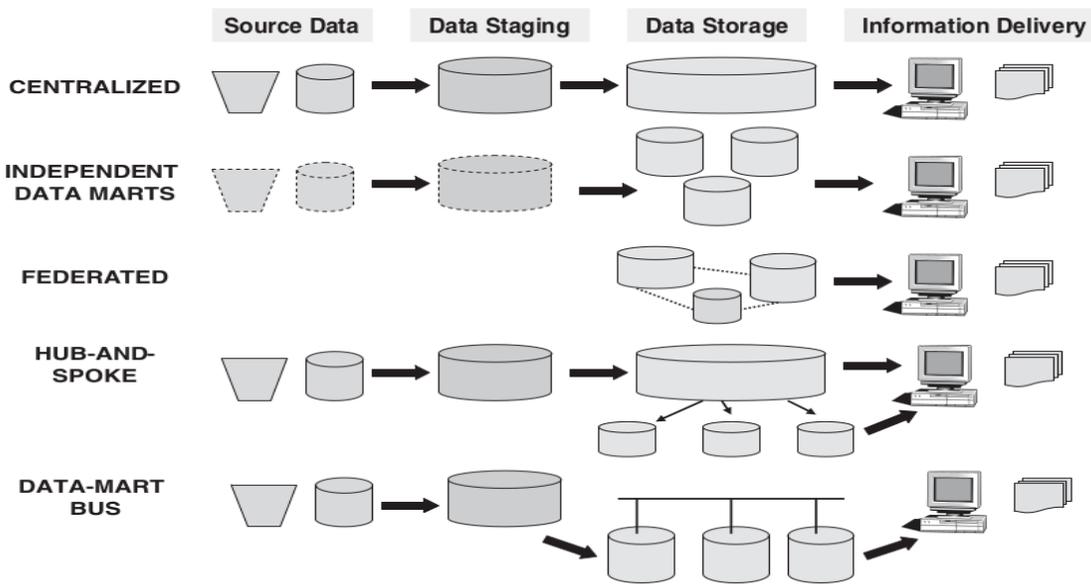


Figura 5.9: Tipos de arquitecturas en los Data Warehouses.

5.1.6 Componentes de un Data Warehouse

La arquitectura de un sistema, es la disposición correcta de los componentes. Un almacén de datos está compuesto por componentes de software y hardware. Para adaptarse a los requisitos de la organización, los componentes se organizan de una determinada manera para obtener el máximo beneficio. Es posible que se desee poner especial énfasis en un componente específico; Es posible que se desee reforzar otro componente con herramientas y servicios adicionales. Todo esto depende de las circunstancias de la organización. A continuación, se muestra gráficamente los componentes de un Data Warehouse:

Architecture is the proper arrangement of the components.

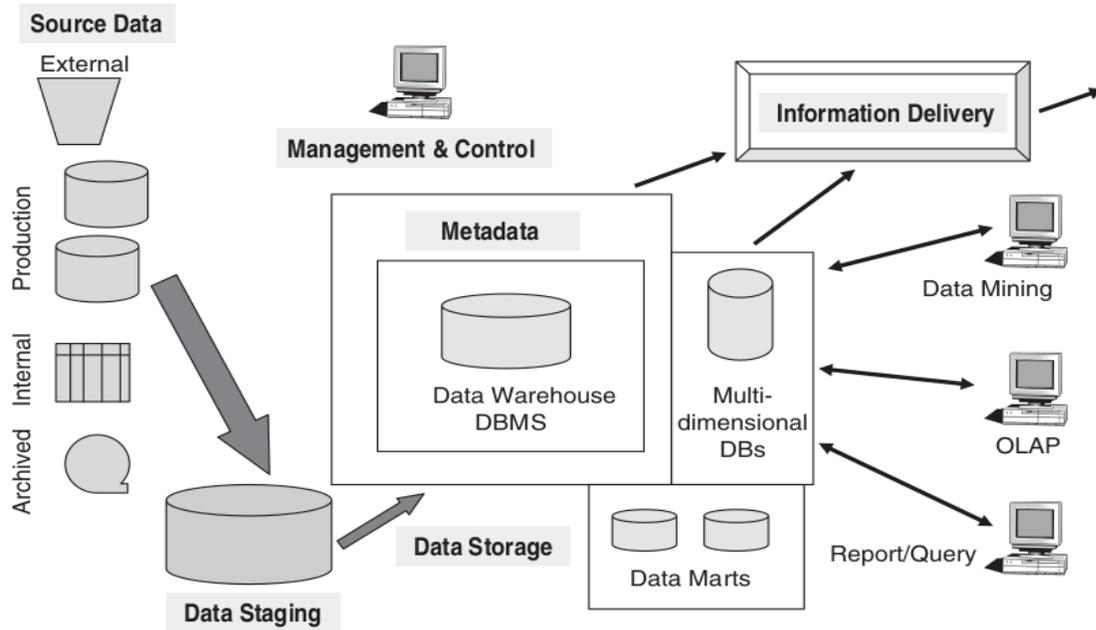


Figura 5.10: Componentes de un Data Warehouse

En los siguientes apartados, se explican brevemente en qué consisten estos componentes.

5.1.6.1 Fuente de datos (Source Data)

Los datos de origen que ingresan al almacén de datos se pueden agrupar en cuatro categorías amplias, como se explica aquí.

Datos de Producción (Production Data):

Esta categoría de datos proviene de los diversos sistemas transaccionales de la empresa. Estos normalmente incluyen sistemas financieros, sistemas de manufactura, sistemas a lo largo de la cadena de suministro y sistemas de administración de relaciones con los clientes. Según los requisitos de información en el almacén de datos, se eligen segmentos de datos de los diferentes sistemas transaccionales. Al tratar con estos datos, normalmente se encuentran con muchas variaciones en los formatos de datos. También se observa que los datos residen en diferentes plataformas de hardware. Además, los datos son compatibles con diferentes sistemas de bases de datos y sistemas operativos.

Datos Internos (Internal Data):

En esta categoría los datos pueden ser hojas de cálculo, documentos, perfiles de clientes, e incluso bases de datos departamentales. Se trata de los datos internos que podrían ser útiles en un almacén de datos.

Los datos internos agregan complejidad adicional al proceso de transformación e integración de los datos antes de que puedan almacenarse en el almacén de datos. Se deben determinar estrategias para recopilar datos de hojas de cálculo, encontrar formas de tomar datos de documentos textuales y vincularlos a bases de datos departamentales para recopilar datos pertinentes de esas fuentes. De nuevo, es posible que se desee programar la adquisición de datos internos. Inicialmente, es posible que se desee limitar a solo algunas porciones importantes antes de comenzar con su primer Data Mart.

Datos Archivados (Archived Data):

Los sistemas transaccionales están destinados principalmente a ejecutar el negocio actual. En cada sistema transaccional, periódicamente se toman los datos antiguos y se almacenan en archivos como datos históricos.

Como se mencionó anteriormente, un almacén de datos guarda instantáneas históricas de los datos. Básicamente, se necesitan datos históricos para analizarlos a lo largo del tiempo. Para obtener esa información histórica, se busca en los conjuntos de datos históricos. Según los requisitos del almacén de datos, así será la cantidad de datos históricos que se incluyan. Este tipo de datos es útil para **discernir patrones y analizar tendencias**.

Datos Externos (External Data):

En muchas organizaciones, los empleados necesitan de datos externos, por ejemplo, el almacén de datos de una empresa de alquiler de vehículos contiene datos sobre los programas de producción actuales de los principales fabricantes de automóviles. Esta información externa en el almacén de datos ayuda al plan de la compañía de alquiler de automóviles a gestionar su flota. Para detectar las tendencias de la industria y comparar el rendimiento con otras organizaciones, se necesitan datos de fuentes externas.

5.1.6.2 Componentes de preparación de datos (Data Staging Component)

Después de haber extraído datos de varios sistemas transaccionales y de fuentes externas, se deben preparar los datos para almacenarlos en el almacén de datos. Los datos extraídos que provienen de varias fuentes diferentes deben modificarse, convertirse y prepararse en un formato que sea adecuado para almacenar y consultar.

Se deben realizar tres funciones principales para preparar los datos. Debe extraer los datos, transformar los datos y luego cargar los datos en el almacenamiento del almacén de datos. Estas tres funciones principales de extracción, transformación y preparación para la carga tienen lugar en un área de preparación. **El componente de preparación de datos consiste en un banco de trabajo para estas funciones.**

La organización de datos proporciona un lugar y un área con un conjunto de funciones para limpiar, cambiar, combinar, convertir, duplicar y preparar datos de origen para su almacenamiento y uso en el almacén de datos.

Las tres principales funciones de este componente será tratado un apartado específico para profundizar un poco más en ello. Como puede observarse en la Figura 5.11 **es básicamente una base de datos donde se preparan y almacenan temporalmente los datos antes de ser cargados al Data Warehouse**, este componente se denomina muchas veces como ODS (Operational Data Store), y es común mente utilizado en los Data Warehouses centralizados. Básicamente, **un ODS es una base de datos diseñada para integrar datos de múltiples fuentes para operaciones adicionales en los datos, para informes, controles y soporte de decisiones operativas**. A diferencia de un almacén de datos maestro de producción, los datos no se pasan a los sistemas operativos. Se puede pasar para más operaciones y al almacén de datos para informes (Wikipedia, 2018). Así mismo en el área de preparación se lleva a cabo el proceso de ETL (Extracción, Transformación y Carga), el cual será explicado con mayor detalle en una sección por aparte.

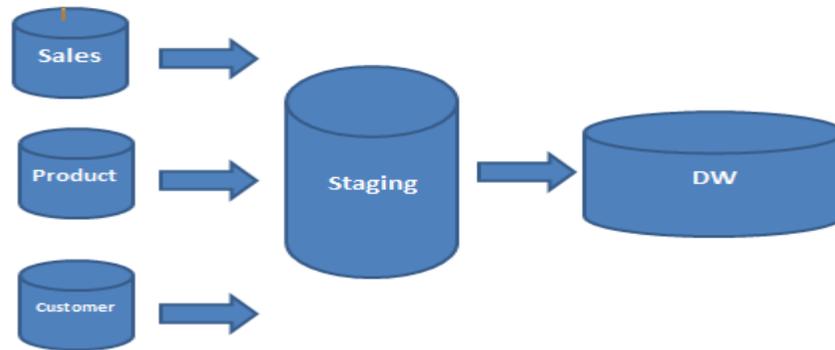


Figura 5.11: Componente de Staging área de un Data Warehouse.

5.1.6.3 Componente de almacenamientos de datos (Data Storage Component)

El almacenamiento de datos para el almacén de datos es un repositorio separado. Los sistemas transaccionales de las empresas respaldan las operaciones cotidianas. Estas son aplicaciones de procesamiento de transacciones en línea. Los repositorios de datos para los sistemas transaccionales suelen contener solo los datos actuales. Además, estos repositorios de datos contienen los datos estructurados en formatos altamente normalizados para un procesamiento rápido y eficiente. Por el contrario, la base de datos para un almacén de datos, debe mantener grandes volúmenes de datos históricos para su análisis. Además, debe mantener los datos en el almacén de datos en estructuras adecuadas para el análisis, y no para la recuperación rápida de piezas de información individuales. Por lo tanto, el repositorio de datos para el almacén de datos se mantiene separado del repositorio de datos para los sistemas transaccionales.

Generalmente, la base de datos en el almacén de datos debe estar abierta. Dependiendo de los requisitos, es probable que se usen herramientas de múltiples proveedores. El almacén de datos debe estar abierto a diferentes herramientas. La mayoría de los almacenes de datos emplean sistemas de gestión de bases de datos relacionales. Muchos almacenes de datos también emplean sistemas de gestión de bases de datos multidimensionales. Los datos extraídos del almacenamiento del almacén de datos se agregan de muchas maneras y los datos de resumen se guardan en las bases de datos multidimensionales (MDDB). Dichos sistemas de bases de datos multidimensionales son generalmente productos patentados.

5.1.6.4 Componente de entrega de información (Information Delivery Component)

Con el fin de proporcionar información a la amplia comunidad de usuarios del almacenamiento de datos, el componente de entrega de información incluye diferentes métodos de entrega de información. La siguiente figura se muestra los diferentes métodos de entrega de información. Los informes ad hoc son informes predefinidos destinados principalmente a usuarios principiantes y ocasionales. La provisión para consultas complejas, análisis multidimensionales (MD) y análisis estadístico atienden las necesidades de los analistas de negocios y usuarios avanzados. La información alimentada en los sistemas de información ejecutiva (EIS) está dirigida a altos ejecutivos y gerentes de alto nivel. Algunos almacenes de datos también proporcionan datos a aplicaciones de minería de datos. Las aplicaciones de minería de datos son sistemas de descubrimiento de conocimiento donde los algoritmos de minería lo ayudan a descubrir tendencias y patrones a partir del uso de sus datos.

En un almacén de datos, se pueden incluir varios mecanismos de entrega de información. Lo más común es que proporcione consultas e informes en línea. Los usuarios ingresarán sus solicitudes en línea y recibirán los resultados en línea.

Se puede configurar la entrega de informes programados por correo electrónico o se puede hacer un uso adecuado de la intranet de la organización para la entrega de información. Recientemente, la entrega de información a través de Internet ha ido ganando terreno.

5.1.6.5 Componentes de metadato (Metadata Component)

Los metadatos en un almacén de datos son similares al diccionario de datos o al catálogo de datos en un sistema de gestión de bases de datos. En el diccionario de datos, conserva la información sobre las estructuras de datos lógicas, la información sobre los archivos y las direcciones, la información sobre los índices, etc. El diccionario de datos contiene datos sobre los datos en la base de datos. Del mismo modo, el componente de metadatos son los datos sobre los datos en el almacén de datos. Esta definición es una definición comúnmente utilizada. Los metadatos en un almacén de datos son similares a un diccionario de datos, pero mucho más que un diccionario de datos.

5.1.6.6 Componente de gestión y control (Management and Control Component)

Este componente de la arquitectura del almacén de datos se encuentra sobre todos los demás componentes.

El componente de gestión y control coordina los servicios y actividades dentro del almacén de datos. Este componente controla la transformación de datos y la transferencia de datos en el almacenamiento del almacén de datos. Por otro lado, modera la entrega de información a los usuarios.

Funciona con los sistemas de gestión de bases de datos y permite que los datos se almacenen correctamente en los repositorios. Supervisa el movimiento de datos en el área de ensayo y, desde allí, en el almacenamiento del almacén de datos.

El componente de gestión y control interactúa con el componente de metadatos para realizar las funciones de gestión y control. Como el componente de metadatos contiene información sobre el almacén de datos en sí, los metadatos son la fuente de información para el módulo de gestión.

5.2 Extracción, Transformación y Carga (ETL)

En el contexto de la inteligencia de negocio, las herramientas ETL han sido la opción usual para alimentar el Data Warehouse (José Curto). En el proceso de ETL los datos son extraídos desde una base de datos OLTP, transformada para que coincida con el esquema de datos del Data Warehouse y cargados en la base de datos del almacén de datos.

Muchos Data Warehouses también incorporan datos de sistemas no OLTP como, por ejemplo, archivos de texto, sistemas heredados y hojas de cálculo, y estos datos también requieren extracción transformación y carga.

El ETL a menudo es una combinación compleja de procesos y tecnologías que consumen una parte significativa del esfuerzo en el desarrollo de un Data Warehouse. Las operaciones de un ETL deberían ser realizadas en un servidor de base de datos relacional separado de las bases de datos fuentes y la base de datos propia del Data Warehouse (Priya & Sultana.A, n.d.).

Crear una separación lógica y física entre las fuentes de datos y el Data Warehouse minimiza el impacto de la intensa actividad periódica del ETL en las bases de datos de los sistemas fuentes y la del Data Warehouse.

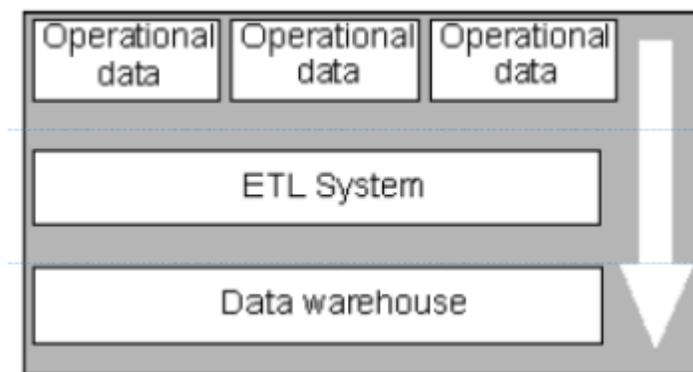


Figura 5.12: Representación gráfica del subsistema ETL

En Data Warehouse es imprescindible tener de un componente de preparación de datos por separado. En los siguientes apartados, se explicarán brevemente las tres funciones principales que tienen lugar en el área de preparación de datos de un Data Warehouse.

5.2.1 Extracción de Datos (Data Extraction)

Esta función tiene que tratar con numerosas fuentes de datos. Se tiene que emplear la técnica apropiada para cada fuente de datos. Los datos de origen pueden ser de diferentes sistemas fuente en diversos formatos de datos. Parte de los datos fuente puede estar en una base de datos relacional. Algunos datos pueden estar en otra red heredada y modelos de datos jerárquicos. En muchos casos los datos de las fuentes aún pueden estar en archivos planos.

Es posible que desee incluir datos de hojas de cálculo y locales o conjuntos de datos departamentales. La extracción de datos puede volverse bastante compleja en esta fase.

Existen herramientas en el mercado para la extracción de datos. Se debe considerar el uso o no de herramientas externas diseñadas para ciertos tipos de fuentes. Para las otras fuentes de datos, es posible desarrollar programas para realizar la extracción. La compra de herramientas externas puede implicar altos costos inicialmente. Los programas desarrollados en casa, por otro lado, significan costos dirigidos a desarrollo y mantenimiento.

Usualmente el equipo de implementación de Data Warehouses extrae la fuente en un entorno físico separado desde el cual sería más fácil mover los datos al almacén de datos. En este entorno separado, se puede extraer los datos de origen en un grupo de archivos planos, o una base de datos relacional de preparación de datos, o una combinación de ambos.

Existen tres técnicas principales de extracción de datos. El tipo de necesidad de la organización es lo que, normalmente, determinará la elección de una u otra forma (Power Data, 2013).

- a) **Full Extract o extracción total:** esta modalidad consiste en extraer la totalidad de datos. En este caso, se barren tablas completas que pueden llegar a tener millones de registros.
- b) **Incremental Extract o extracción incremental:** se va procesando por lotes únicamente lo que fue modificado o agregado. También puede haber filas que se borren por estar duplicadas, tratarse de datos erróneos, etc.
- c) **Update Notification o notificación de actualizaciones:** en este caso, solo se van extrayendo los datos a medida que se produce una actualización (por ejemplo, un insert).

5.2.2 Transformación de Datos (Data Transformation)

En toda implementación de sistema, la conversión de los datos es una función importante. Si la extracción de los datos en un Data Warehouse representa grandes retos, la transformación de ellos plantea retos aún más grandes. Otro factor en el almacén de datos es que la alimentación de datos no es solo una carga inicial. Tendrá que seguir recogiendo los cambios en curso de los sistemas fuente. Cualquier tarea de transformación que configure para la carga inicial se adaptará también para las cargas en curso.

En la **transformación** de los datos se realizan ciertas tareas específicas. Primero, se limpian los datos extraídos desde cada fuente. La limpieza puede ser solo corrección de errores ortográficos, o pueden incluir la resolución de conflictos entre códigos de estado y códigos postales en los datos de origen, o pueden tratar de proporcionar valores predeterminados para elementos de datos faltantes, o la eliminación de duplicados se extrae los mismos datos de múltiples sistemas de origen.

La estandarización de los elementos de datos forma una gran parte de la transformación de datos. Se deben estandarizar los tipos de datos y las longitudes de campo para los mismos elementos de datos recuperados de las diversas fuentes. La estandarización semántica es otra tarea importante. Aquí se resuelven sinónimos y homónimos. Cuando dos o más términos de diferentes sistemas fuente significan lo mismo, se resuelven estos los sinónimos.

Cuando un único término significa muchas cosas diferentes en diferentes sistemas de origen, se resuelven estos homónimos.

Las técnicas o métodos que se pueden aplicar durante el proceso de limpieza de datos son (Power Data, 2013):

- **Aplicación de reglas de unificación de datos.** Por ejemplo, poner en la fila correspondiente al sexo la misma letra identificativa, como podría ser “M” para masculino y “F” para femenino. En este caso, también se tendrían que identificar o corregir posibles errores, como que algún usuario haya puesto la “M” como mujer.
- **Validaciones de completitud.** Por ejemplo, comprobar que en todos los registros de datos de los clientes de un banco esté introducida la dirección postal completa, saltando una alarma si falta alguno.
- **Estandarización de datos.** El objetivo es que todos los datos del mismo tipo estén introducidos de idéntica forma. Un ejemplo sería el DNI con la letra final de identificación fiscal junto a los números y sin guion de separación.

La transformación de datos implica muchas formas de combinar datos de las diferentes fuentes. Este proceso también **implica purgar datos de origen** que no son útiles y separar los registros de origen en nuevas combinaciones. La clasificación y fusión de datos tiene lugar a gran escala en el área de preparación de datos. En muchos casos, las claves elegidas para los sistemas transaccionales son valores de campo con significados integrados.

Por ejemplo, el valor de la clave del producto puede ser una combinación de caracteres que indiquen la categoría del producto, el código del almacén donde se almacena el producto y algún código para mostrar el lote de producción. **Las claves primarias en el almacén de datos no pueden tener significados incorporados.** La **transformación** de datos también incluye la asignación de claves sustitutas derivadas de las claves primarias del sistema de origen.

Un sistema operacional en el punto de venta de la cadena de supermercados mantiene los montos de ventas e ingresos unitarios por transacciones individuales en el mostrador de salida de cada tienda. Pero en el almacén de datos, puede que no sea necesario mantener los datos en este nivel detallado. Es posible que desee resumir los totales por producto en cada tienda para un día determinado y mantener los totales resumidos de las unidades de venta y los ingresos en el almacenamiento del almacén de datos. En tales casos, la función de transformación de datos incluiría un resumen apropiado.

5.2.3 Carga de Datos (Data Loading)

Cuando finaliza la función de transformación de datos, se tiene una colección de datos integrados limpia, estandarizada y resumida. En ese punto se tienen los datos listos para cargar en cada conjunto de datos en el almacén de datos.

Dos grupos distintos de tareas forman la función de carga de datos. Cuando se completa el diseño y la construcción del almacén de datos y este comienza a funcionar por primera vez, se realiza la carga inicial de los datos en el almacenamiento del almacén de datos. La carga inicial mueve grandes volúmenes de datos usando grandes cantidades de tiempo.

A medida que el almacén de datos comienza a funcionar, continúa extrayendo los cambios en los datos de origen, transformando las revisiones de datos y alimentando las revisiones incrementales de datos de forma continua.

La Figura 5.13 ilustra los tipos comunes de movimientos de datos desde el área de preparación hasta el almacenamiento del almacén de datos.

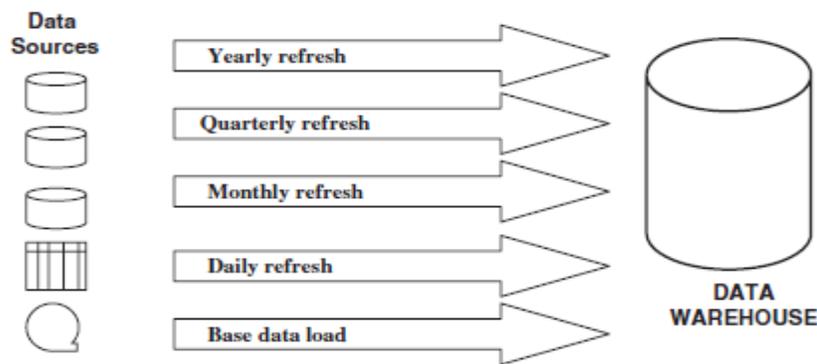


Figura 5.13: Frecuencia de movimientos en un Data Warehouse.

5.3 Diseño multidimensional

El diseño multidimensional es una solución en el manejo de altos volúmenes de información orientado a la analítica de datos, consiste en una adaptación de los diseños relacionales de los sistemas operacionales a uno más cercano a la visión de los usuarios finales, lo que facilita la obtención de datos y mejora el rendimiento, el resultado de estructurar los datos en modelos multidimensionales se conoce como hipercubo de información o cubos multidimensionales.

En la Figura 5.14 se muestra una representación de un cubo de tres dimensiones, si hacemos analogía a los ejes **x, y, z** para coordenadas espaciales, podemos señalar una ubicación específica conociendo tres puntos, en el modelo multidimensional cada eje representa una dimensión y de igual forma que las coordenadas espaciales, un cubo de información de tres dimensiones podemos ubicar un dato en particular señalando un punto de intersección de las tres dimensiones.

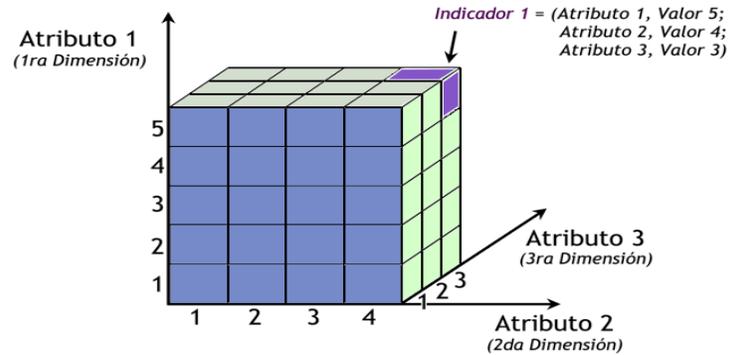


Figura 5.14: Cubo tridimensional

5.3.1 Características del modelo multidimensional

En general, la estructura básica de un Data Warehouse para el modelo multidimensional está definida por dos elementos: **esquemas y tablas**. En una base de datos multidimensional se definen dos tipos de tablas, las tablas de hecho y las tablas de dimensión.

Tablas de hecho (o tablas fact): es la tabla central de un esquema dimensional y representa al elemento de negocio del que se desea obtener indicadores usualmente son los procesos de negocio.

Tablas de dimensión (o tablas lock up): son las tablas que contienen los detalles asociados a las tablas de hechos, por lo general contienen, una llave simple y un conjunto de atributos que describen la dimensión. En dependencia del esquema multidimensional que se siga, pueden existir atributos que representen llaves foráneas de otras tablas de dimensión, es decir, que establecen una relación de esta tabla con otra dimensión.

A la colección de tablas que representan los cubos de información se les conoce como esquemas multidimensional, de los cuales se destacan tres, el esquema estrella, el esquema copo de nieve y el esquema constelación.

Esquema estrella: Es un paradigma de modelado que tiene un solo objeto en el medio conectado con varios objetos de manera radial, el objeto en el centro de la estrella es la tabla de hechos y los objetos conectados a ella son las tablas de dimensiones ver.

Esquema Estrella

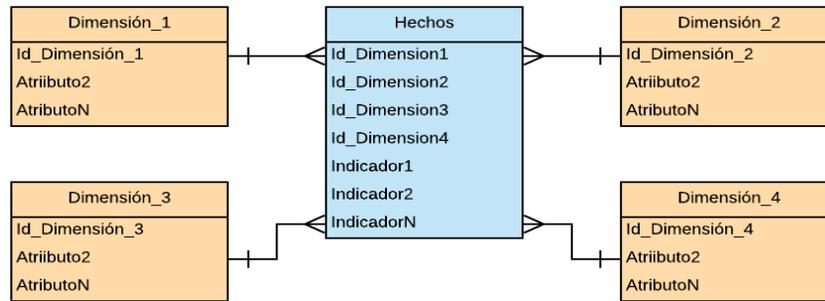


Figura 5.15: Esquema estrella de un modelo multidimensional.

Esquema copo de nieve: El esquema copo de nieve es una extensión del esquema estrella en donde cada una de las puntas de la estrella puede dividirse en más puntas. En esta forma de esquema, las tablas de dimensión pueden tener relación con otras tablas de dimensión. La Figura 5.16 muestra una representación gráfica de este esquema.

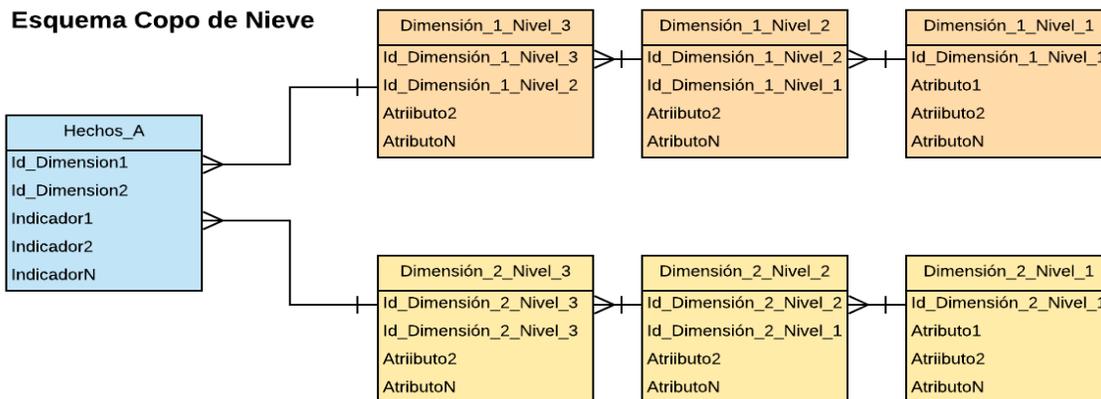


Figura 5.16: Esquema copo de nieve de un modelo multidimensional

Esquema constelación: Se compone de una serie de esquemas estrella, y como se muestra en la Figura 5.17, está formado por una tabla de hechos principal y por una o más tablas de hechos auxiliares, las cuales pueden ser sumalizaciones de la principal. Dichas tablas yacen en el centro del esquema y se relacionan con sus respectivas tablas de dimensiones.

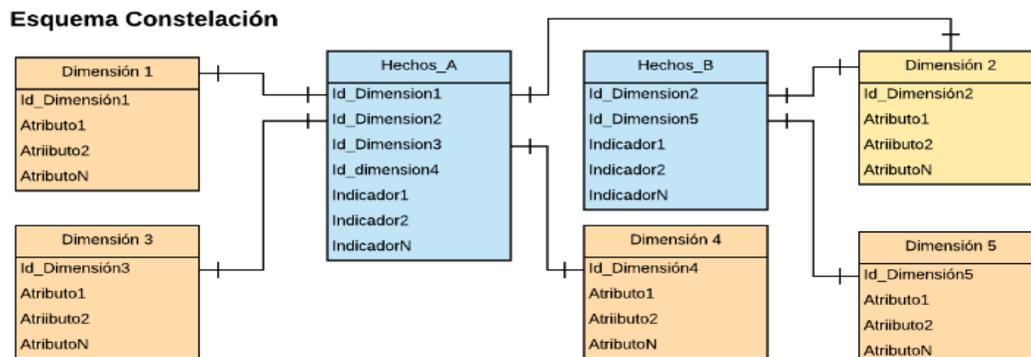


Figura 5.17: Esquema constelación de un modelo multidimensional.

5.3.2 Jerarquías y Niveles

Un nivel se corresponde con un atributo dentro de una dimensión, este representa un nivel particular de agregación.

A nivel de dimensiones es posible definir jerarquías, las cuales son grupos de atributos que siguen un orden preestablecido. Una jerarquía implica una organización de niveles dentro de una dimensión, con cada nivel representando el total agregado de los datos del nivel inferior. Las jerarquías definen cómo los datos son sumariados desde los niveles más bajos hacia los más altos. Una dimensión típica soporta una o más jerarquías naturales. Una jerarquía puede, pero no exige contener todos los valores existentes en la dimensión.

5.4 OLTP

Los sistemas OLTP que significa por sus siglas en inglés, On-Line Transactional Processing, son bases de datos orientadas al procesamiento de transacciones. Una transacción genera un proceso atómico, que debe ser validada por un commit o un rollback y que puede envolver operaciones de insertar, modificar y borrar a los datos. El proceso transaccional es típico de las bases de datos operacionales.

Más concretamente OLTP es un programa que facilita y gestiona aplicaciones orientadas a transacciones, típicamente para ingreso de datos y retención de transacciones en varias industrias en donde se trabaja con una nutrida cantidad de datos. Los procesos de transacciones en línea de hoy en día requiere cada vez más el apoyo que abarquen una red y puedan más áreas de una compañía. (Sinergia e Inteligencia de Negocio S.L., 2016).

5.5 OLAP

Los sistemas OLAP que significan por sus siglas en inglés, On-Line Analytical Processing, son bases de datos orientadas al procesamiento analítico. Este análisis suele implicar, la lectura de grandes cantidades de datos para llegar a extraer información útil: tendencias de ventas, patrones de comportamiento de clientes, elaboración de informes, etc. Este sistema es típico de los Data Marts. (Sinergia e Inteligencia de Negocio S.L., 2016).

5.5.1 Persistencia MOLAP, ROLAP, HOLAP

Los cubos, las dimensiones y las jerarquías son la esencia de la navegación multidimensional del OLAP. Al describir y representar la información de esta manera, los usuarios pueden navegar intuitivamente en un conjunto de datos. Aunque, el solo describir el modelo de datos en una forma más intuitiva, hace muy poco para ayudar a entregar la información al usuario más rápidamente.

La Figura 5.18 muestra una representación gráfica de un cubo de información con las dimensiones de producto, territorio y tiempo.

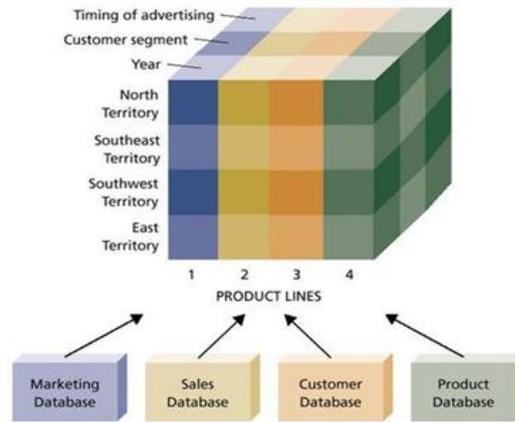


Figura 5.18: Ejemplo de un cubo de información.

Un principio clave del OLAP es que los usuarios deberían obtener tiempos de respuestas consistentes para cada vista de datos que requieran, ya que la información se colecta en el nivel de detalle solamente, el resumen de la información es usualmente calculado por adelantado, siendo estos valores pre calculados, la base de las ganancias de desempeño del OLAP.

En un principio la tecnología OLAP, se asumía que su utilidad era única en modelos de almacenamiento no relacional. Con el tiempo se descubrió que a través del uso de estructuras de base de datos (esquema de estrella y de copo de nieve), índices y el almacenamiento de agregados, se podrían utilizar sistemas de administración de bases de datos relaciones para el OLAP.

Así nació la tecnología OLAP Relacional (ROLAP) y la tecnología OLAP Multidimensional (MOLAP). Las implementaciones MOLAP normalmente se desempeñan mejor que las tecnologías ROLAP, pero tienen problemas de escalabilidad. Por otro lado, las implementaciones ROLAP son más escalables y son frecuentemente atractivas a los clientes debido a que aprovechan las inversiones en tecnologías de bases de datos relacionales preexistentes. (Sinergia e Inteligencia de Negocio S.L., 2016).

5.5.1.1 MOLAP

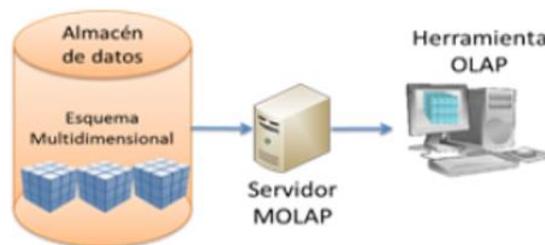


Figura 5.19. Estructura de una arquitectura MOLAP

MOLAP usa una base de datos multidimensionales para proporcionar el análisis, y ser visualizada en varias dimensiones. Utiliza una arquitectura en dos niveles: la base de datos multidimensional (encargada del manejo, acceso y obtención del dato) y el motor analítico. El nivel de aplicación es el responsable de la ejecución de los requerimientos y el nivel de presentación se integra con el de aplicación y proporciona una interfaz a través del cual los usuarios finales visualizan los análisis OLAP.

La información procedente de los sistemas operacionales se carga, mediante una serie de rutinas de lotes, una vez cargada la base de datos multidimensionales (MDDDB) se realizan una serie de cálculos por lotes, para calcular datos agregados, a través de las dimensiones de negocio, rellendo la estructura MDDDB. Tras rellenar esta estructura, se generan índices y algoritmos de tablas hash para mejorar los tiempos de acceso a las consultas. Una vez concluya este proceso y la MDDDB esté lista, los usuarios pueden solicitar informes a través de la interfaz y la lógica de aplicación de la MDDDB obtiene el dato. La arquitectura MOLAP requiere unos cálculos intensivos de compilación. (Sinergia e Inteligencia de Negocio S.L., 2016).

5.5.1.2 ROLAP

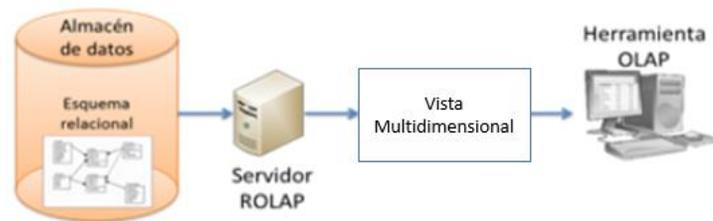


Figura 5.20 Estructura de una arquitectura ROLAP

La arquitectura ROLAP accede a los datos almacenados en un Data Warehouse para proporcionar los análisis OLAP. Las capacidades OLAP se soportan mejor contra las bases de datos relacionales.

Utiliza una arquitectura de tres niveles: La base de datos relacional maneja los requerimientos de almacenamiento de datos, y el motor ROLAP proporciona la funcionalidad analítica. El nivel de base de datos usa bases de datos relacionales para el manejo, acceso y obtención del dato. El nivel de aplicación es el motor que ejecuta las consultas multidimensionales de los usuarios.

El motor ROLAP se integra con niveles de presentación, con los que se realizan los análisis OLAP. Después de que el modelo de datos para el Data Warehouse se ha definido, los datos se cargan desde el sistema operacional, ejecutándose rutinas de bases de datos para agregar el dato, si es requerido por el modelo de datos, creándose así los índices para optimizar los tiempos de acceso a las consultas. Los usuarios ejecutan sus análisis multidimensionales, a través del motor ROLAP, que transforma dinámicamente sus consultas a consultas SQL, ejecutándose estas últimas en las bases de datos relacionales y así, sus resultados se relacionan mediante tablas cruzadas y conjuntos multidimensionales para devolver los resultados requeridos.

La arquitectura ROLAP es capaz de usar datos pre calculados si estos están disponibles, o generarlos si es preciso. Esta arquitectura accede directamente a los datos del Data Warehouse y soporta técnicas de optimización de accesos para acelerar las consultas. Estas optimizaciones son, por ejemplo, particionado de los datos a nivel de aplicación, soporte a la desnormalización y joins múltiples. (Sinergia e Inteligencia de Negocio S.L., 2016)

5.5.1.3 HOLAP

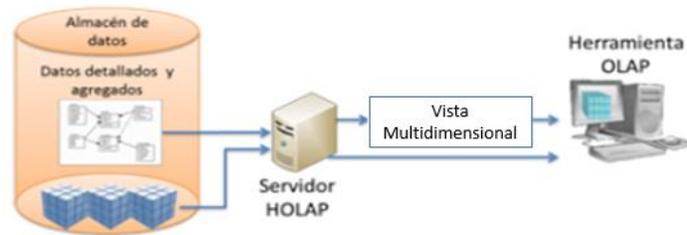


Figura 5.21: Arquitectura HOLAP

La arquitectura HOLAP es la denominada OLAP HÍBRIDA, la cual combina las arquitecturas ROLAP y MOLAP para brindar una solución con características de ambas, como el desempeño superior y gran escalabilidad. Un tipo HOLAP mantiene los registros de detalle (los volúmenes más grandes) en la base de datos relacional, mientras que mantiene las agregaciones en un almacén MOLAP separado. (Sinergia e Inteligencia de Negocio S.L., 2016).

5.6 Dashboards

Los Dashboards (Cuadros de mando) son una herramienta muy popular dado que permiten entender muy rápidamente la situación de negocio y son muy atractivos visualmente. Por ello, todas las soluciones del mercado los incluyen. La oferta se diferencia principalmente en el nivel de madurez del proceso de creación del cuadro de mando, en las opciones disponibles de visualización, y en la capacidad de trabajar con flujos continuos de datos y el reflejo de dichos cambios en tiempo real. Una de las últimas tendencias, en el contexto de cuadros de mando, es la inclusión de elementos gráficos que permiten el análisis de grandes cantidades de información. Es lo que se conoce como “visual analytics”. No todas las herramientas del mercado incluyen esta tendencia: sólo se encuentra en algunos productos innovadores ².

² Curto, Díaz, Josep, and Díaz, Josep Curto. Introducción al Business Intelligence, Editorial UOC, 2010.

6 ANTECEDENTES

6.1 Defensoría del Consumidor

La Defensoría del Consumidor es la institución pública que tiene por misión proteger y promover efectivamente los derechos de las personas consumidoras, facilitando la solución de controversias de consumo con calidez e innovación. (Marco Institucional Defensoría del Consumidor, s.f.). La Defensoría como institución está compuesta por las unidades que se muestran en la Figura 6.1: Organigrama de la Defensoría del Consumidor.

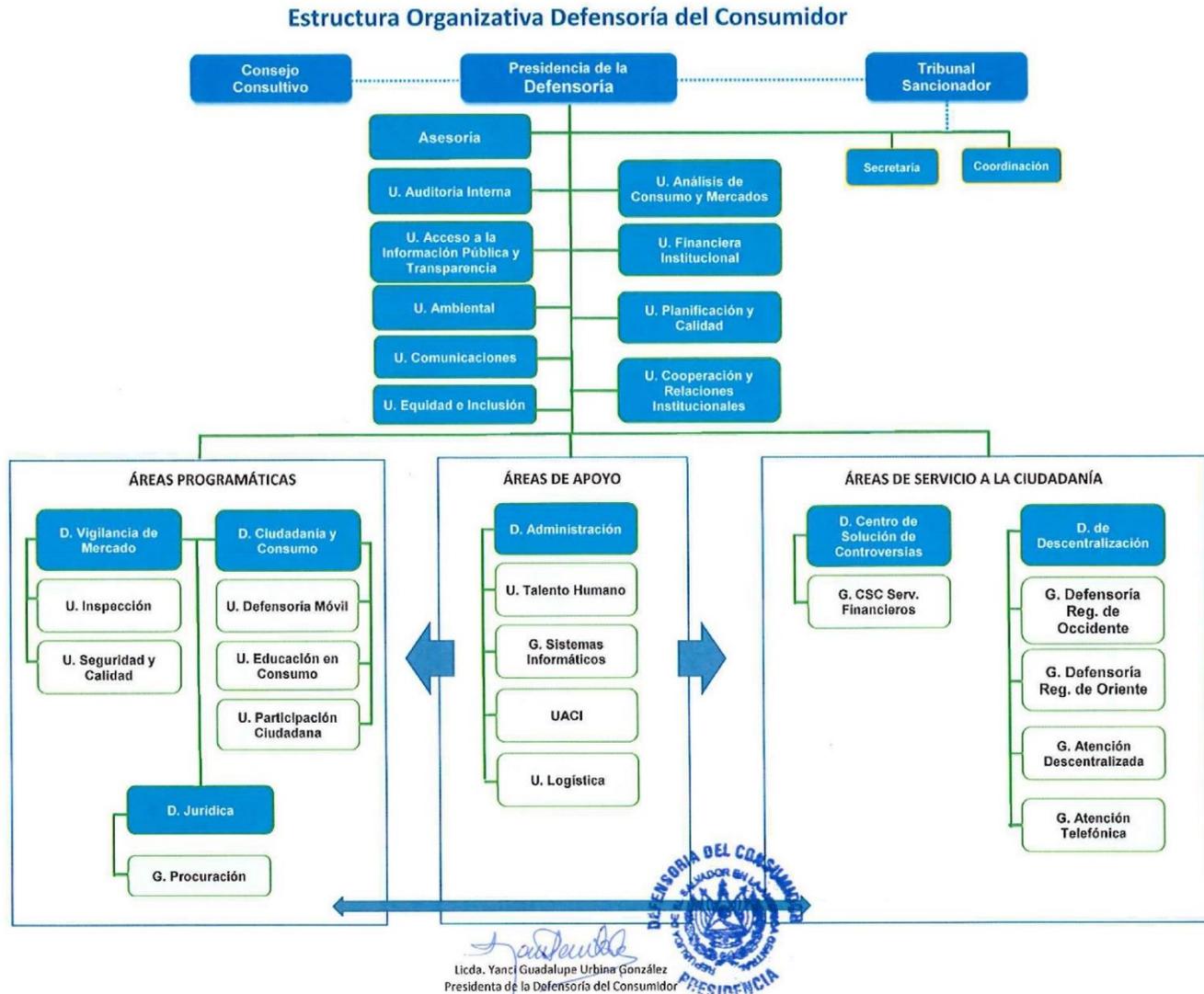


Figura 6.1: Organigrama de la Defensoría del Consumidor

La función operativa principal de la Defensoría es atender y dar seguimiento a los diferentes casos presentados por los consumidores en los diferentes medios que la institución pone a disposición del público.

Los tipos de casos que la defensoría recibe en su sistema por parte de un ciudadano pueden ser identificados con los siguientes ejemplos:

Acción	Ejemplo
<i>Asesoría</i>	“Juan Pérez solicita conocer, ¿Qué puede hacer? en el caso que compre un jugo enlatado en un supermercado, pero la compra esté vencida” La defensoría realiza un asesoramiento para Juan, ante los pasos a seguir para presentar una denuncia formal.
<i>Denuncia</i>	“Juan Pérez verifica que el producto enlatado está vencido y decide acudir a la defensoría a interponer una demanda contra el supermercado” La defensoría recibe los datos y Juan puede realizar su denuncia.
<i>Derivación</i>	“Juan Pérez también tiene problemas en su casa, ya que la luz eléctrica se ausenta de 7 a 8 pm”. Sin embargo, esta acción no le compete a la defensoría, por tanto, Juan Pérez es registrado como una derivación. La defensoría le informa que no puede administrar ese tipo de denuncias y que será remitido a las autoridades competente
<i>Denuncia no personal</i>	“Juan Pérez realiza una denuncia por teléfono, ya que quiere omitir su nombre, él sabe que, si no brinda los suficientes detalles sobre la denuncia, probablemente desestimen el caso” La defensoría permite realizar denuncias anónimas, sin embargo, deben estar correctamente argumentadas.
<i>Denuncia colectiva</i>	“Juan Pérez y toda su comunidad realizan una denuncia ya que los precios en el tambo de gas propano se han incrementado, arbitrariamente, con respecto a otras zonas del país”. La defensoría toma nota de las denuncias y actúa según corresponda.

Tabla 6.1: Ejemplo de casos que resuelve la Defensoría del Consumidor.

Cabe recalcar que una denuncia puede llegar a ser de prevención, cuándo el consumidor no presenta la documentación completa; Así mismo, una denuncia puede cerrarse o ser inadmisibles cuándo no se presenta toda la documentación.

6.2 Unidades del Negocio

6.2.1 Dirección Jurídica

La Dirección Jurídica de la Defensoría del consumidor es la encargada de la elaboración de denuncias colectivas e individuales para ser presentadas al Tribunal Sancionador, así como demandas judiciales individuales y colectivas en representación de los consumidores afectados.

Durante el proceso de elaboración de denuncias la Dirección Jurídica procesa actas de inspección que provienen de la Dirección de Vigilancia de Mercado, las cuales relatan y evidencian hallazgos de posibles incumplimientos de la ley del consumidor, son los técnicos de la unidad los que se encargan de la elaboración de las denuncias de acuerdo a los conocimientos del reglamento y ley de protección al consumidor para los casos que apliquen.

Finalmente, elaboradas las denuncias estas son enviadas al Tribunal Sancionador el cual determinan las acciones a tomar (Resoluciones).

6.2.2 Dirección de Ciudadanía y Consumo

La dirección de Ciudadanía y Consumo es la unidad de la defensoría encargada del monitoreo a nivel nacional en 3 actividades diferentes: Defensorías Móviles, Educación Ciudadana y Participación Ciudadana.

En el área de Educación Ciudadana se brinda información de interés a la ciudadanía, mediante la entrega de volantes, talleres, diplomados y charlas a un grupo específico de ciudadanos o comunidades que hayan sido identificadas como afectadas en la prestación de un servicio o la recepción de un bien. Esta actividad no se registra en SARA.

En la Participación Ciudadana se provee ayuda orientada a la asociación de consumidores, en donde se promueven derechos mediante asambleas ciudadanas, o mesas temáticas, ferias y festivales con el fin de capacitar a un sector de una comunidad, para que luego sean estos quienes transmitan la información aprendida. Esta actividad no se registra en SARA.

Las defensorías móviles son registradas en SARA, las cuales se subclasifican en 4 tipos: 52 Municipios, 210 Municipio, Defensorías a solicitud y Jornadas Móviles.

- 52 Municipios: Muestras los 52 municipios más importantes y que tienen en común desconformidad con algún producto o servicio y que este sea de carácter importante o urgente.
- 210 Municipios: toman los 210 municipios donde se han registrado más denuncias está a diferencia de la primera, se tratan desconformidades en general de la población.
- Defensoría a solicitud: cuando la defensoría asiste a un llamado de una institución para proveer sus servicios
- Jornadas Móviles: cuando la defensoría planifica salidas de campo y van asistiendo a todo ciudadano que desea realizar una asistencia, consulta o denuncia.

6.2.3 Tribunal sancionador

La dirección de Tribunal Sancionador, es la unidad de la defensoría encargada en dictar las resoluciones a las denuncias impuestas. En el proceso de resolución de una denuncia, puede conllevar varios pasos a seguir, algunos hasta repetitivos o cíclicos. Una denuncia puede llevar inicialmente las etapas de: Ingreso, Apertura/Prueba y Resolución Final. Dependiendo del caso puede llegar a necesitar las otras etapas como las siguientes:

- Diligencia que es un paso intermedio entre Apertura/Prueba y Resolución Final
- Etapa Complementaria
- Etapa Revocatoria

El proceso inicia para cada denuncia con el ingreso a SARA por parte de la recepcionista, en este punto se da el rotulado de la denuncia según el número que le asigne SARA, luego pasa al secretario de la unidad el cual da el visto bueno y con ello finaliza la etapa de Ingreso y da parte a la etapa de Apertura/Prueba. La denuncia pasa en manos de un técnico quien es el encargado de elaborar la resolución de la denuncia. Una vez el documento de resolución está elaborado es revisado por el Coordinador Jurídico, en el caso que este correctamente elaborado, pasa a un cargo identificado como “Miembro del tribunal” quienes realizan una segunda revisión. Al no haber observaciones se regresa al secretario para realizar la notificación y que esta sea entregada a las partes involucradas. Este proceso es realizado por un notificador y es validado por un administrador.

Cuando todo ha sido completado exitosamente, es devuelto al secretario para concluir con la etapa de Apertura/Prueba y dar paso (Si no hay procesos intermedios) al dictamen de la resolución. Cabe mencionar que este proceso no tiene un tiempo estimado y dependerá en gran medida de las dimensiones de la denuncia. Existen 2 vías de ingresos, ya sea por Presidencia (que comprende a la dirección de Vigilancia y Mercado) o por Dirección Jurídica. En el segundo caso de ingreso puede ser por la unidad de Centro de Solución de controversias.

6.2.4 Centro de Solución de Controversias.

Centro de solución de controversias (CSC) es la encargada de llevar los procesos alternos de solución de conflicto, como avenimiento, conciliación y arbitraje. Los técnicos del CSC, son los encargados de recibir y clasificar las quejas de los consumidores, así como también, darles soluciones a estas quejas en cualquier de las soluciones alternas.

6.2.5 Dirección de descentralización

La dirección de descentralización, son los encargados de brindar la atención a la población en general por medio de las formas de recepción siguientes: Por teléfono, sitio web, ventanillas, e incluso presenciales, sub dividiendo estas formas de recepción en defensoría-móvil, medios electrónicos, teléfono directo y/o ventanillas descentralizadas y ciudad mujer.

6.2.6 Unidad de Análisis y Consumo de Mercados

La unidad de análisis y consumo de mercados, se encarga del estudio e investigación del fenómeno de consumo para generar información útil en la elaboración de propuestas de política pública que fortalezcan la protección efectiva y eficiente de los derechos e intereses de las personas consumidoras.

Lleva a cabo la administración y divulgación de información institucional que sea socialmente útil en materia de consumo para la protección de las personas consumidoras.

7 SITUACIÓN ACTUAL

7.1 Defensoría del Consumidor

La Defensoría del Consumidor es una institución que está al servicio de la población salvadoreña, su objetivo de velar por los derechos de los consumidores lo llevan a estar en constante necesidad de mejora, sin embargo, como cualquier otra institución del estado, trabaja con recursos limitados, lo que hace mucho más relevante la tarea de identificar los puntos vitales en la operación, para así, focalizar esfuerzos y mejorar sus servicios. Las gerencias y direcciones tienen como reto tomar las mejores decisiones amparándose en la información recopilada del trabajo realizado en períodos anteriores de cada unidad, inclusive, necesitando información de un día anterior. Esta información en su gran mayoría, se registra en el Sistema de Atenciones a Reclamos y Asesorías (SARA), de donde se extrae todos los informes necesarios. SARA es un sistema implementado en 2008, desde su lanzamiento se han desarrollado reportes de acuerdo a las necesidades registradas, en su momento, algunas de estas necesidades han venido evolucionando y al día de hoy se requieren otro tipo de informes.

Actualmente las diferentes unidades de la defensoría del consumidor cuentan con tres opciones para obtener información:

- Apoyarse de los informes generados automáticamente por SARA.
- Autogenerar informes utilizando hojas de cálculo tomando como base un informe master multipropósito que trata de incluir de forma genérica las necesidades de todas las unidades.
- Solicitar desarrollos de informes nuevos a la unidad de informática.

De lo anterior ninguna de estas opciones satisface de manera eficiente las necesidades, ya que si bien se puede obtener la información en su gran mayoría implica tratamiento manual de los datos o tiempo de espera muy prolongado.

7.2 Unidades del Negocio

7.2.1 Dirección Jurídica

7.2.1.1 Situación Actual

La Dirección Jurídica cuenta con módulos del sistema transaccional SARA para el procesamiento y seguimiento de actas y casos de posibles denuncias, sin embargo, carece de un módulo o una vista que les permita acceder a información estadística sobre rendimiento de la unidad, visualizar las cantidades de actas o denuncias que son atendidas en determinado intervalo de tiempo. Actualmente para obtener esta información es necesario llevar controles manuales o solicitar apoyo de la unidad de informática para generarlos.

7.2.1.2 Necesidades del Negocio

Durante las entrevistas realizadas en la Dirección Jurídica, el Gerente de Dirección Jurídica, identificó que existe la necesidad de un mecanismo para obtener y analizar la información de denuncias (actas y casos) desde diferentes perspectivas:

Las actas finalizadas y tiempos invertidos en actas deben poder analizarse desde las siguientes perspectivas:

- Por fecha
- Por empleado

- Por proveedor
- Por etapa

Los casos y tiempos invertidos en ellos que deben poder analizarse desde las siguientes perspectivas:

- Por fecha
- Por empleado
- Por proveedor
- Por etapa
- Por tipo de caso

7.2.2 Dirección de Ciudadanía y Consumo

7.2.2.1 Situación Actual

Ciudadanía y Consumo registra los datos de cada defensoría móvil en hojas llamadas “Formatos Contingenciales”, luego de realizar las defensorías móviles, la unidad de Centro de Solución de Controversias realiza el volcado de información a SARA. Una vez los datos están en el sistema, Ciudadanía y Consumo descarga los datos en formato de hoja de cálculo para su procesamiento y análisis de la unidad.

7.2.2.2 Necesidades del Negocio

En la Dirección de Ciudadanía y Consumo se identificó, en base a las entrevistas realizadas al Director de Ciudadanía y Consumo representante de dicha unidad, la necesidad de obtener información sobre las atenciones que se realizan en los diferentes puntos del país, mediante las campañas móviles y los casos recibidos mediante las ventanillas de atención de la defensoría con el fin de evaluar los resultados de las diferentes campañas de atención y asesoría realizadas por la Defensoría en las diferentes temporadas.

La información obtenida de las atenciones debe poderse visualizar desde las siguientes perspectivas:

- Clasificación (sector, categoría)
- Motivo
- Tipo de atención
- Tiempo (temporada, mensual)
- Forma de recepción
- Consumidor (dirección)
- Proveedor
- Lugar móvil
- Oficina

7.2.3 Tribunal Sancionador

7.2.3.1 Situación Actual

Actualmente Tribunal Sancionador se apoya de la unidad de Sistemas Informáticos para el análisis de los datos, por medio de la solicitud de un archivo de Excel que provea de una captura de la situación actual de todos los casos, las resoluciones, los estados, las etapas y las subetapas que se hayan registrado a lo largo del tiempo. Esta hoja de cálculo es trabajada por medio de filtros y fórmulas menores (Sumas y promedios) gracias a las habilidades de la unidad.

7.2.3.2 Necesidades del Negocio

En entrevistas realizadas a la presidencia del Tribunal Sancionador, se encontró que requieren obtener información sobre las denuncias que ingresan a la unidad por medio del Centro de Solución de Controversias o Presidencia de la Defensoría del consumidor, para así, darle seguimiento a los tramites que se van realizando en cuanto a las denuncias y tener mayor visibilidad de la información los casos activos, cerrados, archivados y pendientes de notificar.

La información obtenida de las cantidades de denuncias ingresadas debe poderse visualizar desde las siguientes perspectivas:

- Tiempo
- Proveedor
- Consumidor
- Empleado
- Oficina Recepción
- Tipo Establecimiento
- Procedencia
- Motivo
- Clasificación
- Ubicación

La información obtenida de las cantidades de resoluciones finales en la unidad desde las siguientes perspectivas:

- Tiempo
- Proveedor
- Consumidor
- Empleado
- Oficina Recepción
- Tipo Resolución
- Tipo Establecimiento
- Procedencia
- Motivo
- Clasificación
- Ubicación

La información obtenida de las cantidades de sanciones impuestas y el monto recuperado por la unidad desde las siguientes perspectivas:

- Tiempo
- Proveedor
- Consumidor
- Empleado
- Oficina Recepción
- Infracción
- Tipo Establecimiento

- Motivo
- Clasificación
- Ubicación

7.2.4 Dirección del Centro de Solución de Controversias

7.2.4.1 Situación Actual

El centro de solución de controversias cuenta con un módulo de SARA, para brindar el seguimiento y encontrar una solución a los casos; Sin embargo, no cuentan con un módulo que les permita acceder a la información estadística, ni que les permita visualizar la cantidad de atenciones que se han ingresado, la cantidad de casos cerrados o la cantidad de casos reaperturados.

7.2.4.2 Necesidades del Negocio

Durante las entrevistas realizadas al director del Centro de Solución de Controversias, se identificó que existe la necesidad de un mecanismo para obtener y analizar la información de la cantidad de atenciones recibidas, los casos que alcanzaron la etapa de cerrados y la cantidad de casos que han sido reaperturados:

- La cantidad de atenciones recibidas deben visualizarse bajo las siguientes perspectivas:

- Por fecha
- Por lugar móvil
- Por tipo atención
- Por forma de recepción
- Por oficina de recepción
- Por proveedor
- Por consumidor
- Por clasificación (Sector y Categoría)
- Por motivo
- Por campaña
- Por ubicación del proveedor
- Por tipo bien
- Por institución
- Por empleado
- Por etapa

La cantidad de casos cerrados junto a la cantidad de tiempo ingresado, tiempo de avenimiento, tiempo de ratificación y el tiempo de conciliación, si lo hubiera, debe permitir analizarse desde las siguientes perspectivas:

- Por fecha
- Por proveedor
- Por ubicación del proveedor
- Por consumidor
- Por empleado
- Por oficina de recepción
- Por clasificación (Sector y Categoría)

- Por motivo
- Por forma de recepción
- Por lugar móvil
- Por campaña
- Por tipo atención
- Por etapa Cierre
- Por tipo bien
- Por institución que toma la denuncia.

- La cantidad de casos reaperturados, debe permitir analizarse desde las siguientes perspectivas:

- Por fecha
- Por proveedor
- Por ubicación del proveedor
- Por consumidor
- Por empleado
- Por oficina de recepción
- Por clasificación (Sector y Categoría)
- Por motivo
- Por forma de recepción
- Por lugar móvil
- Por campaña
- Por tipo atención
- Por etapa Cierre
- Por tipo bien
- Por institución que toma la denuncia.

7.2.5 Dirección de Descentralización

7.2.5.1 Situación Actual

La dirección de descentralización cuenta con un módulo de SARA para brindar el ingreso al sistema en diferentes puntos de atención, sin embargo, esta unidad no cuenta con un módulo que les permita acceder a la información estadística de la unidad, pudiendo visualizar la cantidad de atenciones que se han ingresado y/o la cantidad de casos cerrados.

7.2.5.2 Necesidades del Negocio

Durante las entrevistas realizadas a la directora de la Dirección de Descentralización, se identificó que existe la necesidad de un mecanismo para analizar la información de atenciones recibas y las atenciones finalizadas:

- La cantidad de atenciones recibidas debe permitir analizarse bajo las siguientes perspectivas:

- Por fecha
- Por lugar móvil
- Por tipo atención
- Por forma de recepción
- Por oficina de recepción

- Por proveedor
- Por consumidor
- Por clasificación (Sector y Categoría)
- Por motivo
- Por campaña
- Por ubicación del proveedor
- Por tipo bien
- Por institución
- Por empleado
- Por etapa

- La cantidad de casos cerrados junto al monto recuperado, debe permitir analizarse desde las siguientes perspectivas:

- Por fecha
- Por proveedor
- Por ubicación del proveedor
- Por consumidor
- Por empleado
- Por oficina de recepción
- Por clasificación (Sector y Categoría)
- Por motivo
- Por forma de recepción
- Por lugar móvil
- Por campaña
- Por tipo atención
- Por etapa Cierre
- Por tipo bien
- Por institución que toma la denuncia.

7.2.6 Unidad de Análisis y Consumo de Mercados

7.2.6.1 Situación Actual

La unidad de análisis y consumo de mercados cuenta con un módulo de SARA para brindar el análisis oportuno del desarrollo de actividades de la defensoría del consumidor en general, en este módulo el técnico de la unidad descarga y realiza “limpieza de datos”, mediante el análisis de hojas de cálculo mes a mes. Sin embargo, esta unidad no cuenta con un módulo amigable que les permita acceder a la información estadística de la defensoría del consumidor, pudiendo visualizar la cantidad de atenciones que se han ingresado, la cantidad de casos cerrados, el monto recuperado y el monto reclamado.

7.2.6.2 Necesidades del Negocio

Durante las entrevistas realizadas a la jefa de la Unidad de Análisis y Consumo de Mercados, se identificó que existe la necesidad de un mecanismo para analizar la información, por ejemplo, la cantidad de atenciones ingresadas, la cantidad de casos finalizados y reclamados registrados en la defensoría del consumidor. A continuación se detallan aún más las necesidades de negocio.

- La cantidad de atenciones recibidas debe permitir analizarse bajo las siguientes perspectivas:

- Por fecha
- Por lugar móvil
- Por tipo atención
- Por forma de recepción
- Por oficina de recepción
- Por proveedor
- Por consumidor
- Por clasificación (Sector y Categoría)
- Por motivo
- Por campaña
- Por ubicación del proveedor
- Por tipo bien
- Por institución
- Por empleado
- Por etapa

- La cantidad de casos cerrados debe permitir analizar los montos recuperados y montos reclamados, todas estas métricas deben permitir analizarse desde las siguientes perspectivas:

- Por fecha
- Por proveedor
- Por ubicación del proveedor
- Por consumidor
- Por empleado
- Por oficina de recepción
- Por clasificación (Sector y Categoría)
- Por motivo
- Por forma de recepción
- Por lugar móvil
- Por campaña
- Por tipo atención
- Por etapa Cierre
- Por tipo bien
- Por institución que toma la denuncia

8 SOLUCIÓN PROPUESTA

8.1 Visión del Producto

La Visión del producto proporciona un enfoque claro al proyecto, sirve como base para el desarrollo de las historias épicas para la primera versión del Product Backlog.

En el Anexo 1 se muestra la Visión del producto.

En Tabla 8.1 Se muestra la visión del producto en base a la plantilla de Geoffrey Moore como se menciona en el apartado de Metodología en este mismo documento.

Para	Las direcciones y gerencias de la Defensoría del Consumidor
Quienes	Necesitan mejorar el proceso de obtención de información para la toma de decisiones
El	Modelo multidimensional
Es	Una solución de inteligencia de negocios
Que	Permite la generación de reportes personalizados, análisis y divulgación eficiente de la información Potenciará el proceso de obtención de información, gracias a su diseño orientado a la analítica de datos, utilizando la información generada por el sistema de atención a reclamos y asesorías
A diferencia de	Los reportes estáticos generados por el sistema SARA y las hojas de cálculo generadas desde la base de datos periódicamente.
Nuestro producto	Proveerá de forma fácil, eficiente y confiable la información útil para la toma de decisiones, facilitando la analítica mediante el cruce y personalización de la información

Tabla 8.1: Declaración Visión del Producto.

8.2 Arquitectura propuesta

La solución propuesta consta de cinco elementos fundamentales:

- **Orígenes de datos:**
Son las bases de datos del Sistemas transaccional de la defensoría del consumidor.
- **Procesos ETL:**
Los datos se trasladan de la fuente hacia un área de almacenamiento intermedio y luego a la base de datos con los esquemas multidimensionales, pero durante este proceso se realizan tareas de extracción, transformación y/o limpieza y carga de los datos.
- **Almacenamiento intermedio:**
El Stagin Área, es una base de datos intermedia que se utiliza como apoyo para el procesamiento de los datos antes de ser trasladados a la base del modelo multidimensional
- **Data Marts**
Son las Bases de datos con una estructura multidimensional que facilitan el análisis de información
- **Presentación**
Es la interfaz gráfica mediante la que se accederá de forma fácil a los datos, los cuales se podrán presentar en tablas, gráficos o consultas personalizadas.

8.3 Orígenes de datos

Los orígenes de los datos provienen de las bases utilizadas en la Defensoría del consumidor, del sistema denominado SARA (Sistema de Atención a Reclamos y Asesorías), contienen las tablas transaccionales de todas las unidades involucradas en este proyecto. Del sistema SARA se utiliza SARA_DL (Dirección Legal, ahora esta unidad se llama Dirección Jurídica) y SARA_CSC (Centro de solución de controversias) ambas en el motor SQL Server 2008.

8.4 Procesos ETL

Los procesos ETL son las transformaciones que se le dan a los orígenes de datos, para este proyecto se utilizará Pentaho Data Integration (PDI) de la Suite Pentaho, conocido también como Pentaho Kettle, el cual incluye diferentes herramientas como “SPOON” (interfaz gráfica para desarrollo y pruebas de procesos ETL), “KITCHEN” (herramienta de consola para la ejecución de Jobs de PDI), “PAN” (herramienta de consola para la ejecución de Transformaciones de PDI). En general, con este conjunto de herramientas se realiza la extracción de datos, la transformación o limpieza y la carga de la nueva data a las bases de datos que luego alimentarán los Dashboards.

8.5 Tratamiento de datos

Una vez esté identificado el origen de los datos, se realiza la limpieza de ellos (este paso consiste en verificar la validez de los datos, es decir que cumplan criterios particulares de la unidad) por medio de Pentaho con la herramienta Spoon, en el Stagin Area, con lo cual se tienen listos los datos para la siguiente fase, que consiste en una segunda base de datos en donde se crean los Data Mart que en conjunto forman al Data Warehouse.

Para este proyecto el origen de los datos se implementará en el sistema gestor de base de datos SQL Server, que luego será transformado por medio de Pentaho hacia las bases de datos del *staging* y *Data Warehouse* que estarán implementadas en el sistema gestor de base de datos PostgreSQL 9.6.

8.6 Tecnología OLAP

En este proyecto se implementarán Dashboards como herramientas de visualización, esta herramienta consume los datos ubicados en los Data Marts correspondientes. Los dashboards se construyen mediante la elaboración de consulta en DMX (cubos OLAP) o SQL (Base de datos relacional). La tecnología OLAP, son bases de datos orientados al procesamiento analítico: análisis de tendencias según una categoría, patrones de preferencia de compra según el sexo de la persona, elaboración de informes, entre otros.

8.7 MDX

Es el acrónimo de (Multi Dimension eXpressions), el cuál es un lenguaje de consulta para bases de datos multidimensionales sobre cubos OLAP, es muy utilizado en inteligencia empresarial en la generación de informes para la toma de decisiones basadas en los datos históricos, con la posibilidad de cambiar la estructura o “rotación” del cubo, el cuál se refiere a poder visualizar el cubo bajo diferentes dimensiones.

8.8 SQL

Además de poder consultar la información con consultas de MDX, será posible utilizar SQL (Structured Query Language). Al ser una base relacional la que soporta la estructura multidimensional del Data Warehouse a realizar, será posible consultar la información mediante este tipo de lenguajes.

9 METODOLOGÍA

9.1 Actividades de Desarrollo por Iteración (Sprint)

Se han identificado las principales tareas a desarrollar para conseguir los incrementos en cada iteración las cuales se listan a continuación:

No.	Actividades
1	Refinamiento del requerimiento de información
2	Poblar la base de datos de desarrollo SARA
3	Diseño del modelo multidimensional
4	Diseño BD del Stagin Area
5	Diseño del proceso ETL
6	Diseño de vistas
7	Creación de Script SQL modelo multidimensional
8	Creación de Script SQL del Stagin Area
9	Desarrollo de los procesos ETL
10	Desarrollo de cubos OLAP
11	Desarrollo de vistas
12	Implementación e integración de componentes
13	Pruebas
14	Documentación

Tabla 9.1: Actividades de Desarrollo identificadas.

9.2 Herramientas de Análisis y Diseño

Una de las herramientas de análisis de requerimientos para un Data Warehouse son los paquetes de Información. Los paquetes de información son una matriz que sirven para la identificación y la captura de requisitos de información. Por cada tema de negocio específico a analizar se debe realizar un paquete de información. La Figura 9.1 muestra la estructura de un diagrama de paquete de información. En los encabezados se colocan las variables por las cuales desea ser analizada la información; por cada variable se deben colocar sus jerarquías o categorías si existen; en el pie se deben colocar las medidas o métricas que se desean visualizar según las variables mostradas en el encabezado y según el tema específico del paquete de información.

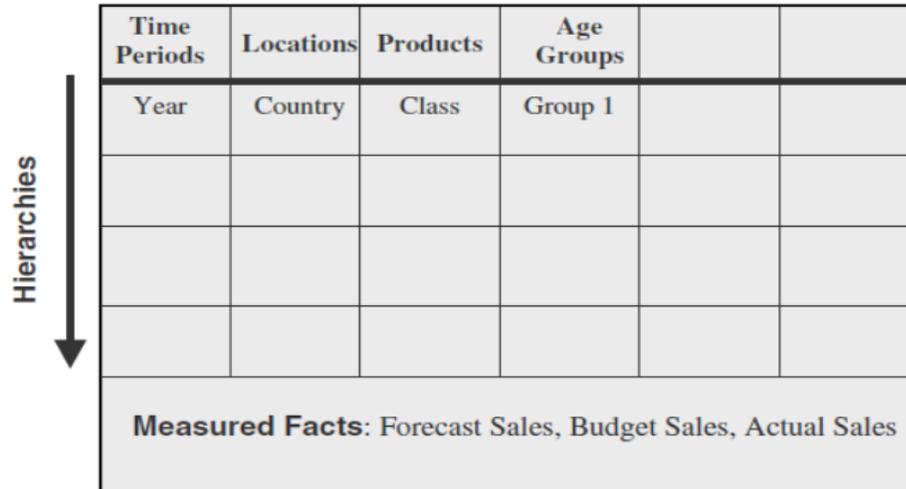


Figura 9.1: Estructura de un paquete de información

9.3 Nomenclaturas

En las siguientes subsecciones, se presenta la nomenclatura para los paquetes de información, diagramas multidimensionales, diagramas de proceso ETL, nombres de transformaciones ETL y nombres de Jobs ETL, utilizadas como convención para el proyecto.

9.3.1 Nomenclatura de Paquetes de información

El formato o nomenclatura a utilizar para la herramienta principal de toma de requerimientos es el formato de los paquetes de información.

Cuando se está en la fase de refinamiento y toma de requerimientos de información para un Data Warehouse, los usuarios normalmente no pueden describir completamente lo que esperan ver en el almacén de datos, es difícil determinar que piezas de información se almacenarán en el Data Warehouse y no se puede determinar claramente como cada tipo de usuarios va a usar el Data Warehouse. Por lo tanto, cuando los requisitos no se pueden determinar completamente, se necesita un concepto diferente para reunir y registrar los requisitos. (PONNIAH, 2010).

Es por ello que, en este proyecto, durante la fase de **análisis y toma de requerimientos** se ha optado por utilizar la herramienta denominada “**Paquete de Información**”, de la cual ya se ha hablado en la sección 9.2 y en donde se puede ver la estructura que posee.

9.3.2 Nomenclatura de Diagramas mutidimensionales

La nomenclatura a utilizar para el modelado multidimensional es UML. El perfil de UML se encuentra definido formalmente mediante reglas expresadas con Object Constraint Language (OLC) que definen el correcto uso de los nuevos elementos de modelado utilizando los paquetes UML. Se establecen 3 niveles en el proceso de diseño multidimensional:

Nivel 1: Definición del modelo. Un paquete representa un esquema estrella de un modelo multidimensional. En este nivel, una dependencia entre dos paquetes indica que los esquemas estrella comparten al menos una dimensión.

Nivel 2: Definición de un esquema estrella. Un paquete representa un hecho o una dimensión de un esquema estrella. En este nivel, una dependencia entre dos paquetes de dimensión indica que las dimensiones comparten al menos un nivel en sus correspondientes jerarquías.

Nivel 3: Definición de un hecho o dimensión. Se compone de un conjunto de clases que representan los niveles jerárquicos en un paquete de dimensión o el esquema estrella completo en el caso de un paquete de hecho.

En las siguientes subsecciones, se muestra la nomenclatura de estereotipos para el diseño del Data Warehouse.

9.3.2.1 Estereotipos de paquetes

Se definen 3 tipos de estereotipos de paquetes que se detalla en la siguiente tabla:

NOMBRE	DESCRIPCIÓN	REPRESENTACIÓN
PAQUETE ESTRELLA (STARPACKAGE)	Los paquetes de este estereotipo representan el esquema estrella en una base de datos multidimensional. Se emplea en el nivel 1.	
PAQUETE DIMENSIÓN (DIMENSIONPACKAGE)	Los paquetes de este estereotipo representan las dimensiones en una base de datos multidimensionales. Se emplea en el nivel 2.	
PAQUETE HECHOS (FACTPACKAGE)	Los paquetes de este estereotipo representan los hechos en una base de datos multidimensionales. Se emplea en el nivel 2.	

Tabla 9.2: Estereotipo de paquetes

9.3.2.2 Estereotipos de clases

Existen 3 tipos de estereotipos partiendo del modelo de clases los cuales se detallan a continuación:

NOMBRE	DESCRIPCIÓN	REPRESENTACIÓN
HECHO (FACT)	Las clases de este estereotipo representan hechos en un modelo de base de datos multidimensional. Se emplean en el nivel 3.	
DIMENSIÓN (DIMENSION)	Las clases de este estereotipo representan las dimensiones en una base de datos multidimensional. Se emplean en el nivel 3.	
BASE (BASE)	Las clases de este estereotipo representan los niveles jerárquicos de las dimensiones de un modelo de base de datos multidimensional, se emplea en el nivel 3.	

Tabla 9.3: Estereotipo de clases

9.3.3 Nomenclatura de Diagramas de Procesos ETL

Al igual que el modelado multidimensional, se utilizará UML para el diseño de procesos ETL. Específicamente se hará uso de una extensión de UML para adaptarlo al dominio del modelado específico del Data Warehouse.

La extensión de UML permite el modelado conceptual de los procesos ETL, y proporciona mecanismos necesarios para especificar las operaciones típicas de los procesos ETL, como la integración de distintas fuentes de datos, la transformación de los atributos, la generación de claves substitutas (surrogate keys), etc.

En la siguiente tabla se muestra los mecanismos ETL con su representación:

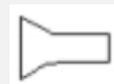
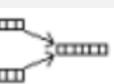
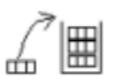
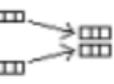
Mecanismo ETL	Descripción	Representación
Aggregation	Agrega los datos (SUM, AVG, MAX / MIN, COUNT, etc.) en base a algún criterio.	
Conversion	Filtra los datos no deseados y verifica la calidad de los datos en base a restricciones.	A → B
Filter	Filtra los datos no deseados y verifica la calidad de los 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 9.4: Nomenclatura de diagramas de procesos ETL

9.3.4 Nomenclatura para Nombres de Procesos ETL

Como equipo se ha establecido un estándar para el nombramiento de los procesos ETL. El nombre de los procesos ETL, específicamente las transformaciones ETL tendrán la siguiente nomenclatura

Prefijo_TipoCarga_TipoTabla_IdentificadorMnemotecnicoN.ktr

Tabla 9.5: Nombre de las transformaciones ETL

A continuación, se describe cada una de las partes de dicha nomenclatura.

Porción del Nombre	Valores Posibles	Significado
<i>Prefijo</i>	st	Indica que es una transformación que mueve datos desde las fuentes datos hacia el staging área.
	dw	Indica que es una transformación que mueve datos desde el staging área hacia el data warehouse.
<i>TipoCarga</i>	unico	Indica que es una transformación que implica un tipo de carga full refresh ³ y que dicha transformación será usada tanto en Jobs iniciales e incrementales de PDI ⁴ .
	inicial	Indica que es una transformación que implica un tipo de carga inicial, es decir para realizar la carga inicial del data warehouse. También indica que dicha transformación será usada únicamente en Jobs iniciales de PDI.
	incremental	Indica que es una transformación que implica un tipo de carga incremental, es decir para realizar las cargas consecutivas e incremental del data warehouse. También indica que dicha transformación será usada únicamente en Jobs incrementales de PDI, periódicamente después de la carga inicial.
<i>TipoTabla</i>	dim	Indica que el tipo de tabla final que se pretende cargar con dicha transformación es una dimensión independientemente si mueve datos al staging área o al data warehouse.
	fact	Indica que el tipo de tabla final que se pretende cargar con dicha transformación es una fact table independientemente si mueve datos al staging área o al data warehouse.
<i>IdentificadorMnemotecnico</i>	<nombre>	Es un valor mnemotécnico que indica la dimensión o hecho del que trata dicha transformación.
<i>N</i>	1,2,3,4,5...	Es un correlativo e indica el orden de ejecución de aquellas transformaciones que por su complejidad fueron divididas en dos o más archivos.
<i>.ktr</i>	.ktr	Se trata de la extensión del archivo de transformación y será siempre el mismo para todas las transformaciones.

Tabla 9.6: Descripción de los elementos del nombre de transformaciones ETL

9.3.5 Nomenclatura para Nombres de JOBS ETL

Los Jobs (PDI) en términos generales no son más que procesos ETL que ejecutan una serie de transformaciones (.ktr) en un orden predefinido. Para este trabajo se ha establecido que se distinguirán dos tipos de Jobs según la clasificación siguiente:

- **Jobs de Primer Nivel (Nivel 1):** Jobs que ejecutan únicamente transformaciones.
- **Jobs de Segundo Nivel (Nivel 2):** Jobs que ejecutan Jobs (los cuales a su vez ejecutan transformaciones).

9.3.5.1 Nomenclatura para Jobs de Primer Nivel

El nombre de los Jobs de primer nivel está compuesto por las siguientes partes:

Prefijo_TipoCarga_TipoTabla_IdentificadorMnemotecnico.kjb

Tabla 9.7: Nombre de los Jobs ETL de primer nivel

³ Full refresh (Recarga completa) hace referencia a cargar los datos en un solo paso. Trunca y vuelve a llenar la tabla.

⁴ PDI Pentaho Data Integration

A continuación, se describe cada una de las partes de dicha nomenclatura.

Porción del Nombre	Valores Posibles	Significado
<i>Prefijo</i>	job	Es un valor fijo que identificar fácilmente mediante el nombre del archivo que se trata de un job.
<i>TipoCarga</i>	unico	Indica que es un job que incluye o ejecuta únicamente transformaciones con el tipo de carga único .
	inicial	Indica que es un job que incluye o ejecuta únicamente transformaciones con el tipo de carga inicial .
	incremental	Indica que es un job que incluye o ejecuta únicamente transformaciones con el tipo de carga incremental .
<i>TipoTabla</i>	dim	Indica que es un job que incluye o ejecuta únicamente transformaciones correspondientes a una sola dimensión en específico. Es decir, todas las transformaciones para una dimensión incluyendo las que mueven los datos hacia staging área, así como también hacia data warehouse.
	fact	Indica que es un job que incluye o ejecuta únicamente transformaciones correspondientes a una sola tabla de hechos en específico. Es decir, todas las transformaciones para una tabla de hechos incluyendo las que mueven los datos hacia staging área, así como también hacia data warehouse.
<i>IdentificadorMnemotecnico</i>	<nombre>	Identificador mnemotécnico de la dimensión o tabla de hechos target del job.
<i>.kjb</i>	.kjb	Se trata de la extensión del archivo de job y será siempre el mismo para todos los jobs.

Tabla 9.8: Descripción de los elementos que componen el nombre de los Jobs ETL de primer nivel.

9.3.5.2 Nomenclatura para Jobs de Segundo Nivel

Los nombres de Jobs ETL de segundo nivel se componen por las siguientes partes.

Prefijo_TipoCarga_UnidadCorporativa.kjb
--

Tabla 9.9: Nombre para los Jobs ETL de segundo nivel.

A continuación, se describe cada una de las partes de dicha nomenclatura.

Porción del Nombre	Valores Posibles	Significado
<i>Prefijo</i>	job	Es un valor fijo que identificar fácilmente mediante el nombre del archivo que se trata de un job.
<i>TipoCarga</i>	inicial	Indica que es un job que incluye o ejecuta únicamente jobs de primer nivel con el tipo de carga único y/o inicial .
	incremental	Indica que es un job que incluye o ejecuta únicamente jobs de primer nivel con el tipo de carga único y/o incremental .
<i>UnidadCorporativa</i>	<nombre>	Identificador mnemotécnico de la unidad dueña del data mart.
<i>.kjb</i>	.kjb	Se trata de la extensión del archivo de job y será siempre el mismo para todos los jobs.

Tabla 9.10: Descripción de los elementos que componen el nombre de los Jobs ETL de segundo nivel.

10 PLANIFICACIÓN

10.1 Estructura organizacional del proyecto

A continuación, se muestra la estructura organizacional del proyecto donde se muestran los roles de cada uno de los que participan en la elaboración del proyecto.

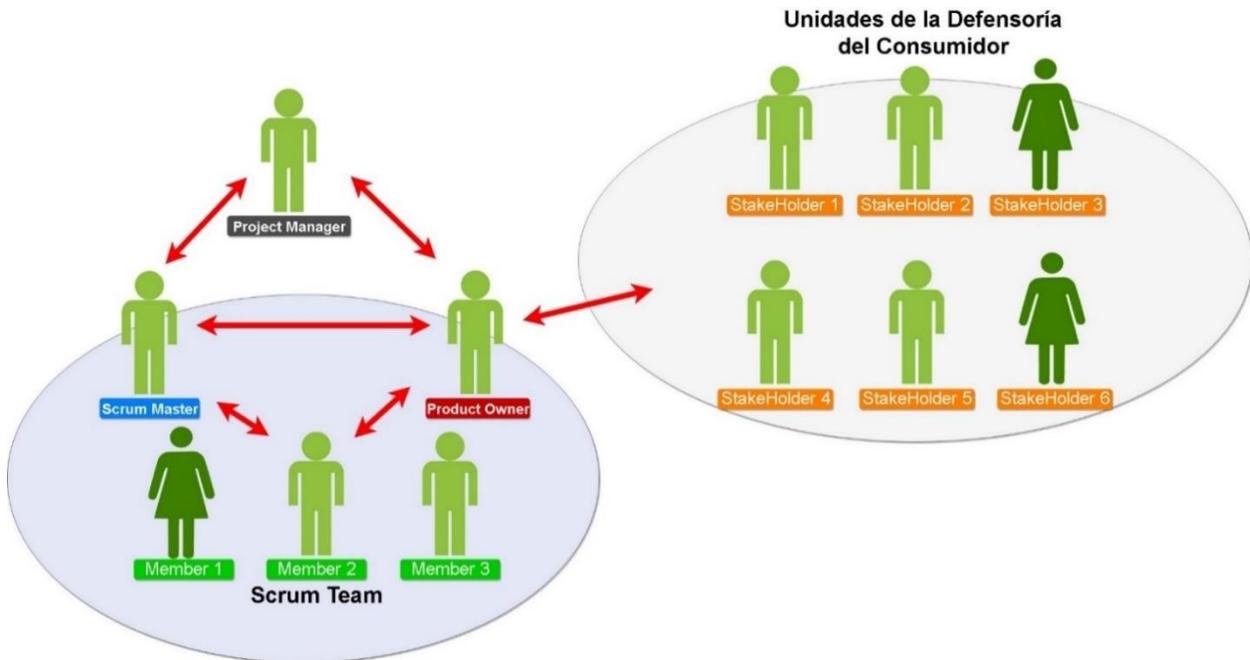


Figura 10.1: Estructura organizacional del proyecto

Los nombres específicos de la organización o unidad a la que pertenecen los participantes en el proyecto se muestran en la matriz de la estructura organizacional.

Rol	Estructura Organizacional.
Project Manager	UES (Ing. Elmer Arturo Carballo)
Product Owner	Director de la Gerencia de Informática (Lic. Salvador Pocasangre)
Scrum Master	UES (Esau Serpas)
Miembro 1	UES (Ivis Navarro)
Miembro 2	UES (Jorge Iraheta)
Miembro 3	UES (Samuel Nuñez)
Stakeholder 1	Director de Dirección Jurídica
Stakeholder 2	Director de la Unidad de Ciudadanía y Consumo.
Stakeholder 3	Directora de Tribunal Sancionador
Stakeholder 4	Analista de Información.
Stakeholder 5	Director del Centro de Solución de Controversias.
Stakeholder 6	Directora de la Unidad de Descentralización.

Tabla 10.1: Matriz de la estructura organizacional.

10.2 Plan de Lanzamiento

Para la ejecución del proyecto se ha decidido realizar un Sprint por unidad. En la Tabla 10.2 se muestra el orden de los sprints las duraciones esperadas de cada Sprint y las Unidades para las cuales se trabajarán en cada Sprint.

RELEASE	SPRINT	NOMBRE DE UNIDAD	DURACIÓN ESTIMADA DEL SPRINT (Semanas)
Lanzamiento 1	Sprint 0	(Investigación y preparación de servidores)	4
	Sprint 1	Dirección Jurídica	4
	Sprint 2	Dirección de ciudadanía y consumo	4
	Sprint 3	Tribunal Sancionador	4
Lanzamiento 2	Sprint 4	Dirección Centro de Solución de Controversias.	4
	Sprint 5	Dirección Descentralización.	4
	Sprint6	Unidad de Análisis de Consumo y Mercados	4

Tabla 10.2: Duración estimada de cada Sprint en Días según unidad.

11 SPRINT 0

11.1 Sprint Backlog

A continuación, se presenta el Sprint Backlog, que refleja las historias de usuarios, su estimación y asignación para este Sprint 0.

#	Historia	Punto Estimado	Asignado
1.1.	Plantilla para redacción de refinamiento del requerimiento	5	Ivis
1.2.	Diseño de la arquitectura	13	Esaú
1.3.	Despliegue de la arquitectura	40	Samuel
1.4.	Implementar BD del desarrollo de SARA	20	Esaú – Samuel

Tabla 11.1: Sprint Backlog para Sprint 0

11.2 Descripción Historias de usuario

A continuación, se presenta la descripción de las historias de usuario detallando: el responsable, una breve descripción de historia, criterios de aceptación y puntos de estimación. La primera historia que se presenta es la historia épica del Sprint 0.

<i>Historia Épica #1</i>	Implementación de entorno de trabajo y arquitectura
<i>Responsable</i>	Samuel Núñez/Esaú Serpas/Ivis Navarro/ Jorge Iraheta Sprint: 0
<i>Descripción</i>	Como Administrador del Data Mart necesito: Diseñar la arquitectura de servidores a utilizar para el desarrollo de los Data Marts de las 6 unidades que forman parte del alcance del proyecto. Implementar dicho diseño para que el equipo pueda tener acceso a los recursos de dicha arquitectura, así como también Investigar y desarrollar una herramienta de refinación de requerimientos para usarla en la fase inicial de cada Sprint para refinar requerimientos.
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> - Deberá existir un Servidor separado para Fuente de Datos, Staging Area, y Data Warehouse. - Los servidores deberán tener 1 dirección IP pública. - Los servidores deberán poderse acceder de forma remota (Remote Desktop o SSH). - Los servidores deberán poderse comunicar entre SI. - Los servidores deberán tener los gestores de bases de datos instalados y configurados correctamente de forma que los servicios a utilizar puedan ser accedidos de forma remota a través de internet. - Los servidores deberán configurarse para que puedan ser accedidos únicamente según los rangos de IP de los miembros del equipo de tesis. - Para Staging Area deberá existir dos bases de datos en el mismo servidor, una de pruebas y otra que será definitiva para Staging Area. - Para Data Warehouse deberá existir dos bases de datos en el mismo servidor, una de pruebas y otra que será definitiva para Data Warehouse. - En cuanto a la herramienta de refinación de requerimientos deberán poderse usar de forma genérica para todas las Unidades.
<i>Puntos estimados</i>	78

<i>Historia de usuario #1</i>	Plantilla para redacción de refinamiento de requerimientos
<i>Responsable</i>	Ivis Navarro Sprint: 0
<i>Descripción</i>	Como Equipo Scrum se necesitamos la redacción de la plantilla para el refinamiento de requerimientos para realizar el refinamiento de requisitos para cada unidad.
<i>Criterios de aceptación</i>	Cumplir con el visto bueno del equipo y el contenido mínimo siguiente: <ol style="list-style-type: none"> 1. Identificación del área 2. Prosa de la historia épica 3. Propósito

	<ol style="list-style-type: none"> 4. Definir Temática 5. Definir dimensiones <ul style="list-style-type: none"> ▪ Descripción ▪ Tipo de dato preliminar ▪ Valores ▪ Descripción de valores 6. Apartado de aceptación (firma)
Puntos estimados	5

Historia de usuario #2	Diseño de la arquitectura
Responsable	Esau Serpas Sprint: 0
	Como Administrador de Data Mart necesito: Diseñar la propuesta de arquitectura Para brindar un despliegue del modelo multidimensional.
Criterios de aceptación	-Cumplir con los requisitos mínimos de hardware. -Incluir las direcciones IP de servidores. -Incluir los nombres de las tecnologías.
Puntos estimados	13

Historia de usuario #3	Despliegue de la arquitectura
Responsable	Samuel Núñez Sprint: 0
Descripción	Como Administrador de Data Mart necesito: Realizar configuración de los servidores virtuales que se utilizaran Para poder implementar el proyecto.
Criterios de aceptación	-Creación de Virtuales -Asignación de recursos de almacenamiento, procesamiento y memoria -Configuración de servidores -Instalación de SGBD: PostgreSql -Instalación de Pentaho Server -Instalación de Pentaho DI -Instalación de Pentaho BA -Configuración de red virtual
Puntos estimados	40

Historia de usuario #4	Implementar base de datos de desarrollo SARA
Responsable	Samuel Nuñez/Esau Serpas Sprint: 0
Descripción	Como Administrador de Data Mart necesito: Contar con una réplica de las tablas y los datos del SARA, Para implementar el sistema SARA
Criterios de aceptación	-Configurar servicio base de datos SQL Server -Desplegar una copia de la base de datos SARA -Habilitar conexión remota y realizar -Prueba de conexión local y remota.
Puntos estimados	20

11.3 Arquitectura de Servidores Desplegada

A continuación, se presenta la arquitectura de servidores desplegada haciendo uso del lenguaje de modelado unificado (UML) para mostrar la configuración de hardware empleada para las fuentes de datos, el Staging Area, el Datawarehouse y Pentaho Server para reportería.

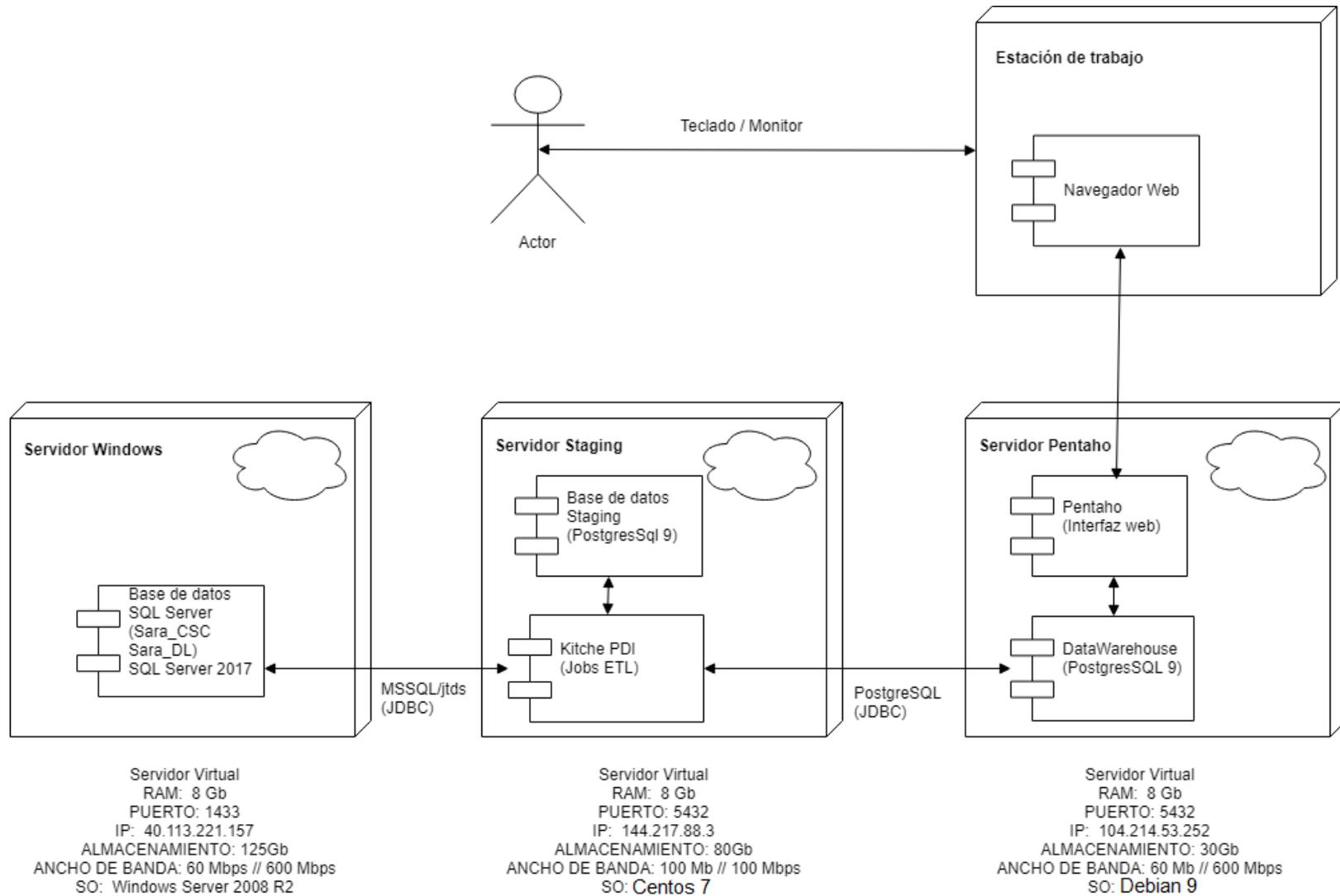


Figura 11.1: Diagrama de despliegue

11.4 Consolas de configuración

En las siguientes subsecciones se muestran las consolas de configuración de cada uno de los servidores configurados. Todos los servidores se encuentran en la nube con el proveedor de Google Cloud. Esta consola de administración permite la realización de actividades tales como, apagar o encender el servidor, configurar sus recursos de hardware, monitorear los recursos de hardware, configurar direcciones IP publicas e internas, moverlos de zona geográfica entre otras.

11.4.1 Servidor Pentaho

A continuación, se muestra la Consola del Servidor Pentaho Server (Debian 9). Se muestra parte del comportamiento de los recursos en el tiempo.

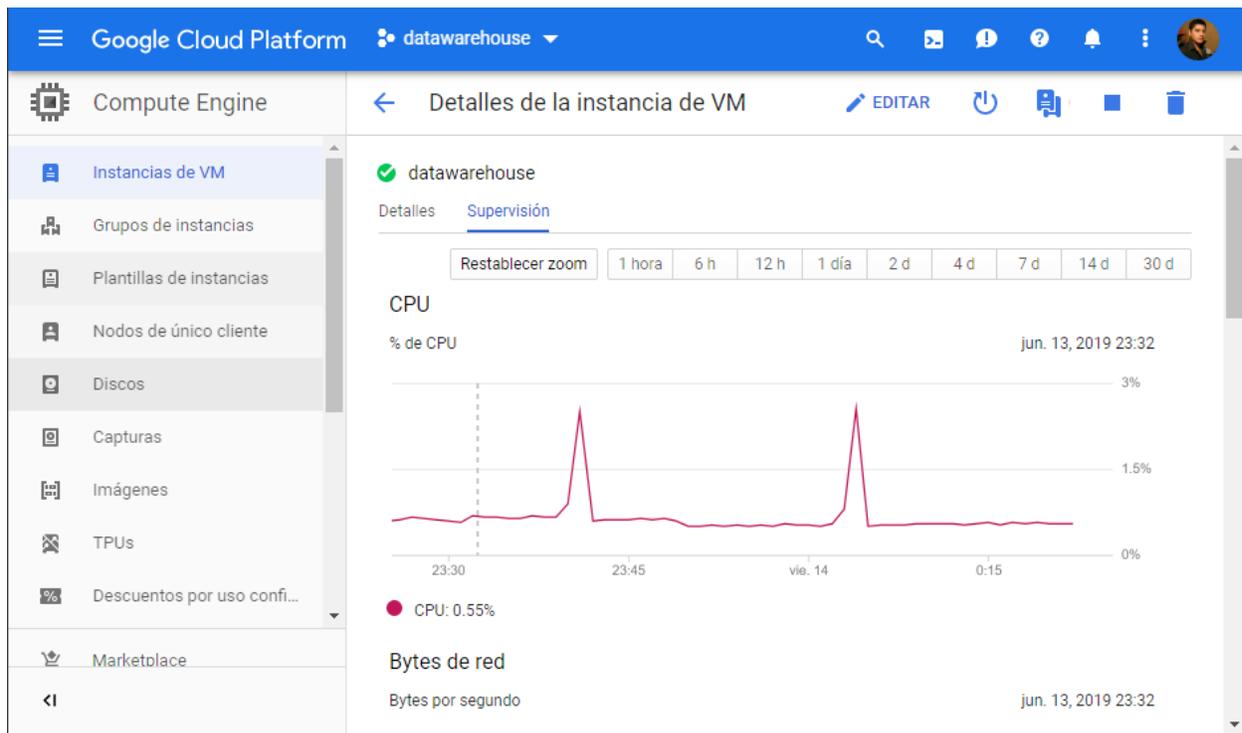


Figura 11.2: Consola de administración para Pentaho Server (Data warehouse)

11.4.2 Servidor Staging

A continuación, se muestra la consola de Administración del Servidor de Staging Area. Su sistema operativo es CentOS 7. Se muestran sus características y recursos asignados actualmente.

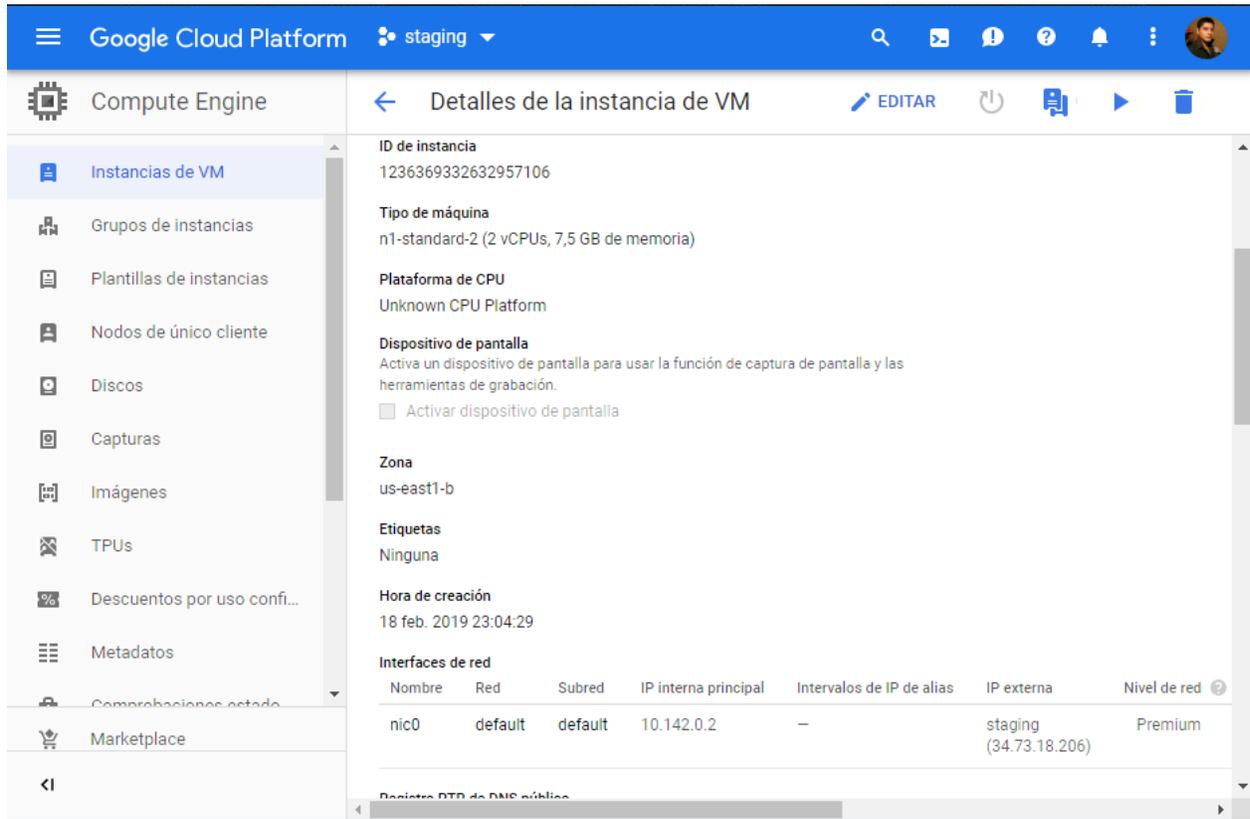


Figura 11.3: Consola de administración para staging area server

11.4.3 Servidor Windows

A continuación, se muestra la consola de administración del servidor de Fuentes de Datos SARA. Se trata de un servidor Windows Server 2008. Se muestran sus características y recursos asignados actualmente.

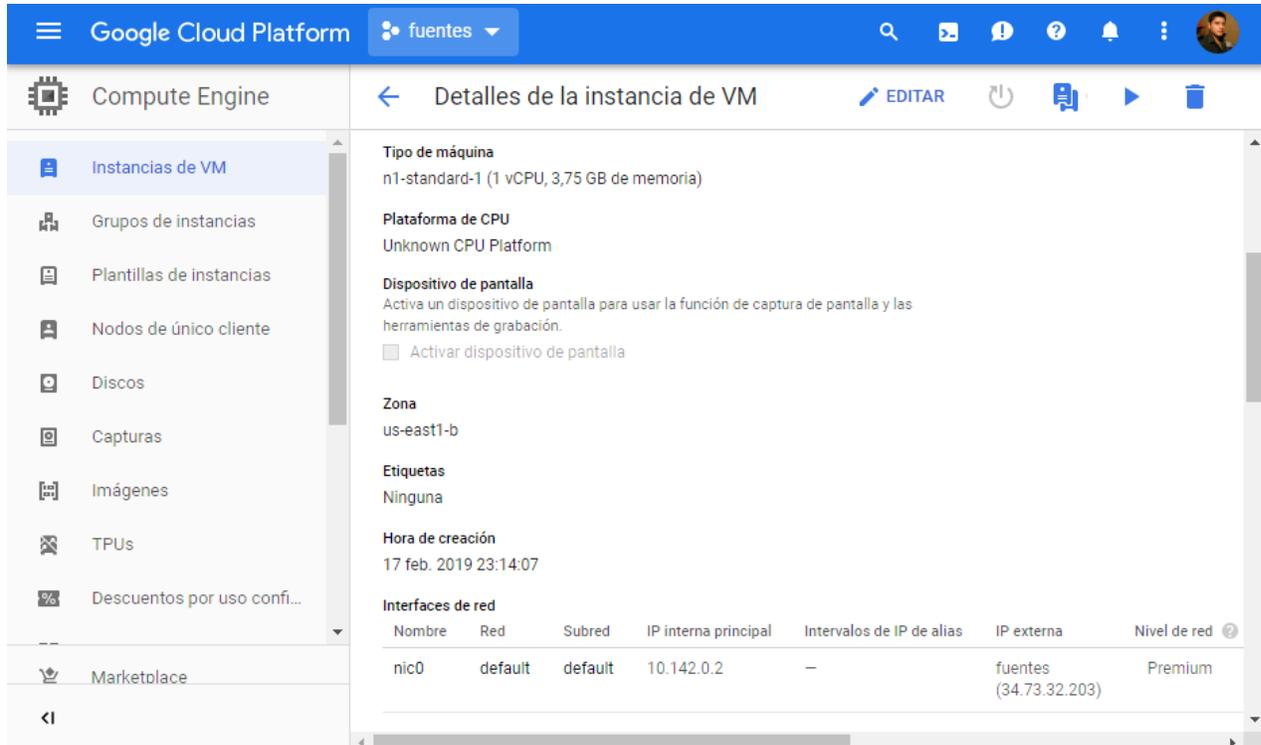


Figura 11.4: Consola de administración para Windows Server

11.5 Diagrama de Componentes

A continuación, se muestra un diagrama que muestra los diferentes componentes que constituyen el Modelo Multidimensional.

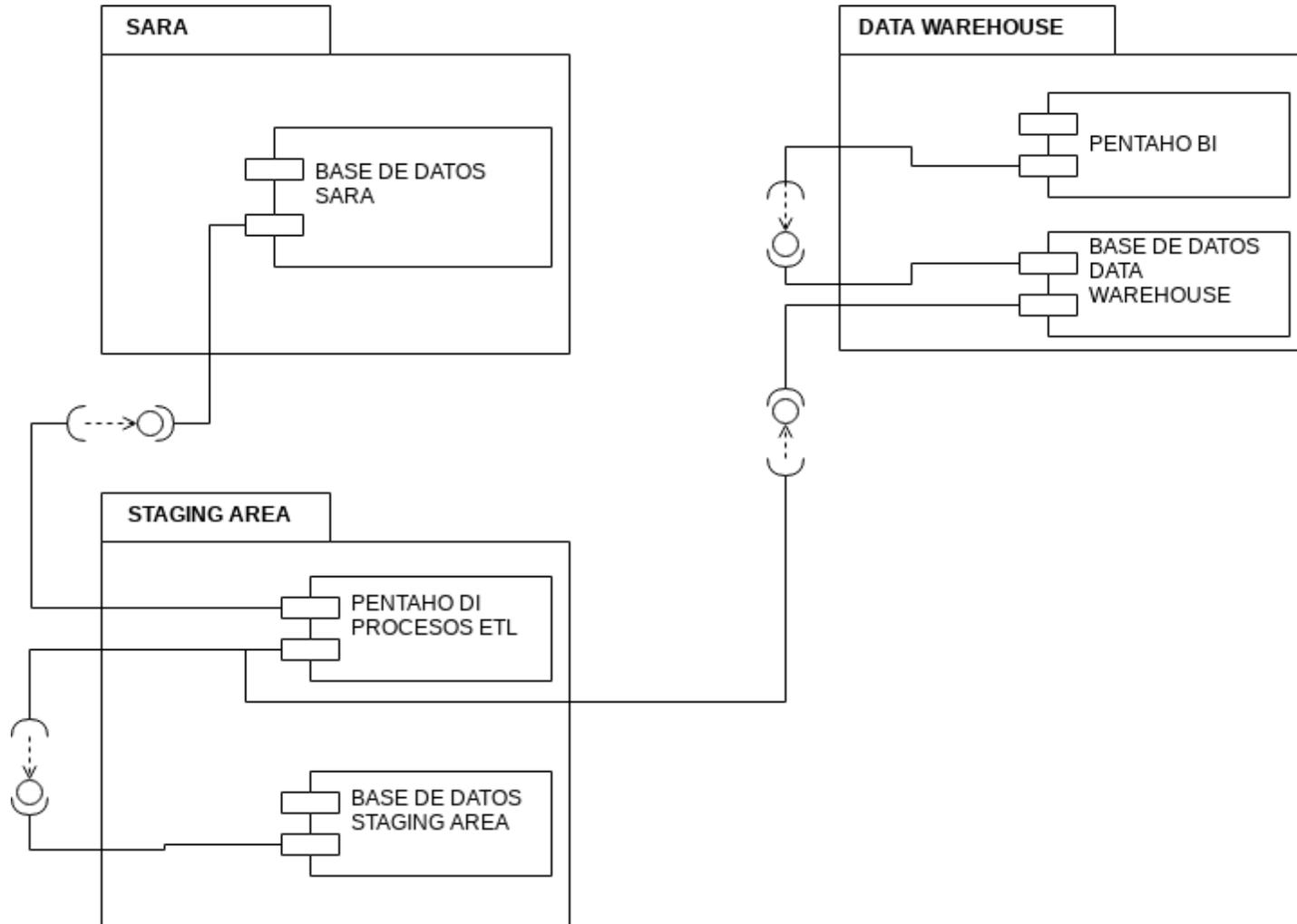


Figura 11.5: Diagrama de componentes de la solución (Modelo multidimensional)

12 SPRINT 1

12.1 Sprint Backlog

A continuación, se presenta el Sprint Backlog, que refleja las historias de usuarios, su estimación y asignación para este Sprint.

#	Historia	Punto Estimado	Asignado
1.	Redacción de la plantilla para el refinamiento de requerimientos	20	Jorge – Esaú – Ivis
2.	Poblar la base de datos de desarrollo SARA	13	Jorge – Ivis – Samuel
3.	Diseño del modelo multidimensional	8	Ivis
4.	Diseño BD del Staging Area	3	Jorge – Samuel
5.	Diseño del proceso ETL	8	Ivis – Samuel
6.	Diseño de vistas	3	Esaú - Ivis
7.	Creación de Script SQL del Staging Area	13	Jorge – Samuel
8.	Creación de Script SQL modelo multidimensional	13	Jorge – Samuel
9.	Desarrollo de los procesos ETL	100	Jorge – Samuel
10.	Desarrollo de cubos OLAP	40	Esaú – Ivis
11.	Desarrollo de vistas	13	Esaú – Ivis
12.	Implementación e integración de componentes	40	Todos
13.	Pruebas	20	Samuel
14.	Documentación	40	Todos

Tabla 12.1: Sprint Backlog, Sprint 1

12.2 Descripción de Historias de Usuario

A continuación, se presenta la descripción de las historias de usuario detallando: el responsable, al sprint al que pertenece, una breve descripción, sus criterios de aceptación y los puntos de estimación. Previamente a las historias de usuario se presenta la historia épica del Sprint 1

<i>Historia Épica #2</i>	Data Mart Dirección Jurídica.
<i>Responsable</i>	Jorge Iraheta/Esaú Serpas/Ivis Navarro/Samuel Núñez Sprint: 1
<i>Descripción</i>	Como Dirección Jurídica (jefatura) Necesito: Un mecanismo para analizar la información de denuncias (actas y casos) desde diferentes perspectivas. Para: realizar evaluaciones comparativas de rendimiento del área técnica.
<i>Criterios de aceptación</i>	- Las actas finalizadas y tiempos invertidos en actas deben poder analizarse desde las siguientes perspectivas: <ul style="list-style-type: none"> • Por fecha • Por empleado • Por proveedor • Por etapa - Los casos y tiempos invertidos en casos deben poder analizarse desde las siguientes perspectivas: <ul style="list-style-type: none"> • Por fecha • Por empleado • Por proveedor • Por etapa • Por tipo de caso - El tiempo invertido en actas y casos deberá ser medido en días.
<i>Puntos estimados</i>	334

<i>Historia de usuario #1</i>	Redacción de la plantilla para el refinamiento de requerimientos
<i>Responsable</i>	Jorge Iraheta/Esau Serpas/ Ivis Navarro Sprint: 1
<i>Descripción</i>	<p>Como Equipo Scrum necesitamos el refinamiento del requerimiento de información de Dirección Jurídica: Realizar reuniones con los interesados y especificar los elementos de información requeridos para la Dirección jurídica.</p> <ul style="list-style-type: none"> • Llenar la herramienta de recolección. • Revisar y desglosar los requerimientos solicitados • Firmar documento por el usuario como aceptación de los de los requerimientos. <p>Para tener claridad de la información que desean almacenar, visualizar y analizar mediante el Data Mart o modelo Multidimensional para ésta unidad.</p>
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> - Tener llena la plantilla de refinamiento de requerimientos de información. - Incluir únicamente requerimientos de información con naturaleza histórica y no información transaccional. - Tener un documento firmado de aceptación de requerimientos. - Identificar (mediante un listado) las tablas fuentes que servirán para cumplir con los requerimientos de información solicitados.
<i>Puntos estimados</i>	20
<i>Historia de usuario #2</i>	Poblar base de datos de desarrollo SARA
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/ Samuel Núñez Sprint: 1
<i>Descripción</i>	<p>Como Product Owner necesitamos: migrar los datos necesarios según las tablas fuentes de datos que se identifiquen durante el proceso de toma de requerimientos. Para poseer datos de prueba para el desarrollo del Data Mart.</p>
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> - Se deben poseer Scripts de INSERT para las tablas Fuente. - Todas las tablas fuentes identificadas en el refinamiento de requerimientos deben poseer datos. - Debe existir al menos una prueba de conexión y acceso remoto exitoso a los datos. - Datos cargados en la Base de datos.
<i>Puntos estimados</i>	13
<i>Historia de usuario #3</i>	Diseño del modelo multidimensional
<i>Responsable</i>	Ivis Navarro Sprint: 1
<i>Descripción</i>	<p>Cómo Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de Dirección Jurídica. Para tener la base (diagrama) para el desarrollo de Scripts de Base de Datos SQL para Data Mart.</p>
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> -Esquema del modelo multidimensional de Dirección Jurídica Finalizado. -El esquema debe cumplir las reglas de un modelo de tipo Estrella. -El esquema debe poseer las relaciones necesarias para la búsqueda según los criterios de búsqueda definidos por la Unidad Dirección Jurídica. -El esquema debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
<i>Puntos estimados</i>	8
<i>Historia de usuario #4</i>	Diseño BD del Staging Area.
<i>Responsable</i>	Jorge Iraheta/Samuel Núñez Sprint: 1
<i>Descripción</i>	<p>Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de Dirección Jurídica. Para tener el insumo para la creación de Scripts de Base de Datos SQL para Staging Area.</p>
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> -Deberá existir una o más tablas por cada dimensión. -Deberá existir una o más tablas de hecho. -El diseño debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
<i>Puntos estimados</i>	3

<i>Historia de usuario #5</i>	Diseño del proceso ETL
<i>Responsable</i>	Samuel Núñez/Ivis Navarro Sprint: 1
<i>Descripción</i>	Como Equipo Scrum necesito: Los Diseños de los Proceso ETL (Extracción, Transformación, Carga) Para cumplir con el Diseño del proceso ETL y tenerlo como insumo a la hora de desarrollar dichos procesos ETL.
<i>Criterios de aceptación</i>	-Se deberá utilizar nomenclatura UML para los pasos en cada proceso ETL. -Deberá incluir todas las tablas dimensiones y tablas de hechos. -Los nombres de las tablas o etiquetas deben poderse visualizar fácilmente.
<i>Puntos estimados</i>	8

<i>Historia de usuario #6</i>	Diseño de vistas
<i>Responsable</i>	Ivis Navarro/Esau Serpas Sprint: 1
<i>Descripción</i>	Como Administrador de Data Warehouse necesito: Realizar el diseño de vistas necesarias para visualizar la información del Data Mart de dirección Jurídica. Para Diseñar y digitalizar la vista de presentación de datos.
<i>Criterios de aceptación</i>	-Deberá existir un diseño de vista o dashboard para cada hecho a medir. -Deberá explicarse cada sección del dashboard o vista. -Etiquetas y textos deben ser legibles a simple vista.
<i>Puntos estimados</i>	3

<i>Historia de usuario #7</i>	Creación de Script SQL del Staging Area
<i>Responsable</i>	Samuel Núñez/Jorge Iraheta Sprint: 1
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será procesada, calculada y/o limpiada la información.
<i>Criterios de aceptación</i>	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Staging Area. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
<i>Puntos estimados</i>	13

<i>Historia de usuario #8</i>	Creación de Script SQL Modelo Multidimensional
<i>Responsable</i>	Samuel Núñez/Jorge Iraheta Sprint: 1
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el esquema Estrella en el SGBD. Para la implementación del esquema donde será almacenada la información.
<i>Criterios de aceptación</i>	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Modelo Multidimensional. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
<i>Puntos estimados</i>	13

<i>Historia de usuario #9</i>	Desarrollo de los procesos ETL.
Responsable	Samuel Núñez/Jorge Iraheta Sprint: 1
Descripción	Como Equipo Scrum necesito: Desarrollar los procesos ETL correspondientes Para obtener, trasladar y transformar los datos que serán almacenados en el Data Mart de Dirección Jurídica.
Criterios de aceptación	<ul style="list-style-type: none"> -Los nombres de los procesos ETL deben poseer la nomenclatura definida por el equipo. -Los procesos ETL deben poseer la documentación necesaria en forma de comentarios dentro de la herramienta PDI. -Las Transformaciones ETL deben haber sido probados en desarrollo (Equipo Local) usando la base de datos pública de pruebas. -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de pruebas. -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de final. -Las transformaciones correspondientes a una dimensión o hecho deben ser incorporados un Job de PDI. -Deberá crearse un Job que incorpore los otros Jobs correspondientes a cada dimensión y hecho. -Los Jobs de PDI deben haber sido probados con las bases de datos de prueba y base de datos final. -Las bitácoras de los procesos ETL deben almacenarse en una base de datos.
Puntos estimados	100

<i>Historia de usuario #10</i>	Desarrollo de Cubos OLAP
Responsable	Esaú Serpas/Ivis Navarro Sprint: 1
Descripción	Como Equipo Scrum necesito: una investigación sobre la herramienta para el desarrollo de cubos OLAP e implementar dicha herramienta para la creación de cubos OLAP. Para implementar el cubo correspondientes a la Dirección Jurídica.
Criterios de aceptación	<ul style="list-style-type: none"> -El cubo deberá poderse visualizar correctamente y sin errores. -Deberá existir como mínimo un cubo OLAP por Hecho. -Deberán poderse realizar las operaciones como mínimo de: Roll Up y Drill Down.
Puntos estimados	40

<i>Historia de usuario #11</i>	Desarrollo de vistas
Responsable	Esaú Serpas/Ivis Navarro Sprint: 1
Descripción	Como Equipo Scrum necesito: Desarrollar vistas o dashboards según el diseño previamente elaborado. Para Presentar información relevante hacia la unidad para la toma de decisiones.
Criterios de aceptación	<ul style="list-style-type: none"> -Vista o dashboard funcionando sin errores técnicos. -Debe coincidir con el diseño realizado previamente. -Debe existir al menos un dashboar por tabla de hechos.
Puntos estimados	13

<i>Historia de usuario #12</i>	Implementación e integración de componentes
Responsable	Ivis Navarro/Esau Serpas/Samuel Núñez/Jorge Iraheta Sprint: 1
Descripción	<p>Como Equipo Scrum necesito: La integración de todos los elementos del modelo multidimensional</p> <ul style="list-style-type: none"> - ETL - Fuentes - Stagin Área - BD multidimensional - Pentaho Server (Vistas) <p>Para el desarrollo del modelo multidimensional de forma integral y probar el funcionamiento correctamente de la totalidad del Data Mart.</p>
Criterios de aceptación	<p>- Se deberán modificar los procesos ETL para que utilicen las bases de datos finales (no las de pruebas).</p> <p>- Se deberán realizar pruebas en la visualización de los dashboards.</p> <p>- Todo el entorno (desde los procesos ETLs, hasta los dashboards o vistas) deberán tener el visto bueno de todo el equipo.</p>
Puntos estimados	40

<i>Historia de usuario #13</i>	Pruebas
Responsable	Samuel Nuñez Sprint: 1
Descripción	Como Equipo Scrum necesito: Realizar pruebas al modelo multidimensional terminado. Para verificar que funciona de la manera que se espera.
Criterios de aceptación	Pruebas exitosas en pruebas de aceptación del usuario.
Puntos estimados	20

<i>Historia de usuario #14</i>	Documentación
Responsable	Samuel Nuñez/Jorge Iraheta/Ivis Navarro/Esau Serpas Sprint: 1
Descripción	Como Equipo Scrum necesito: Documentación completa con historias, nomenclatura y diagramas empleados. Para verificar cómo se llevó a cabo el desarrollo de los cubos de la primera unidad (Dirección Jurídica).
Criterios de aceptación	Entrega de Manuales de creación de Data Mart y documentos completos
Puntos estimados	40

12.3 Refinamiento del Requerimiento de Información

Para la realización del refinamiento de los requerimientos de información se partió con la investigación del proceso general que sigue la Dirección Jurídica dentro la Defensoría del Consumidor. Además de establecer el proceso general de la unidad se hizo uso de los paquetes de información para refinar el requerimiento de información.

12.3.1 Proceso BPMN de la Unidad

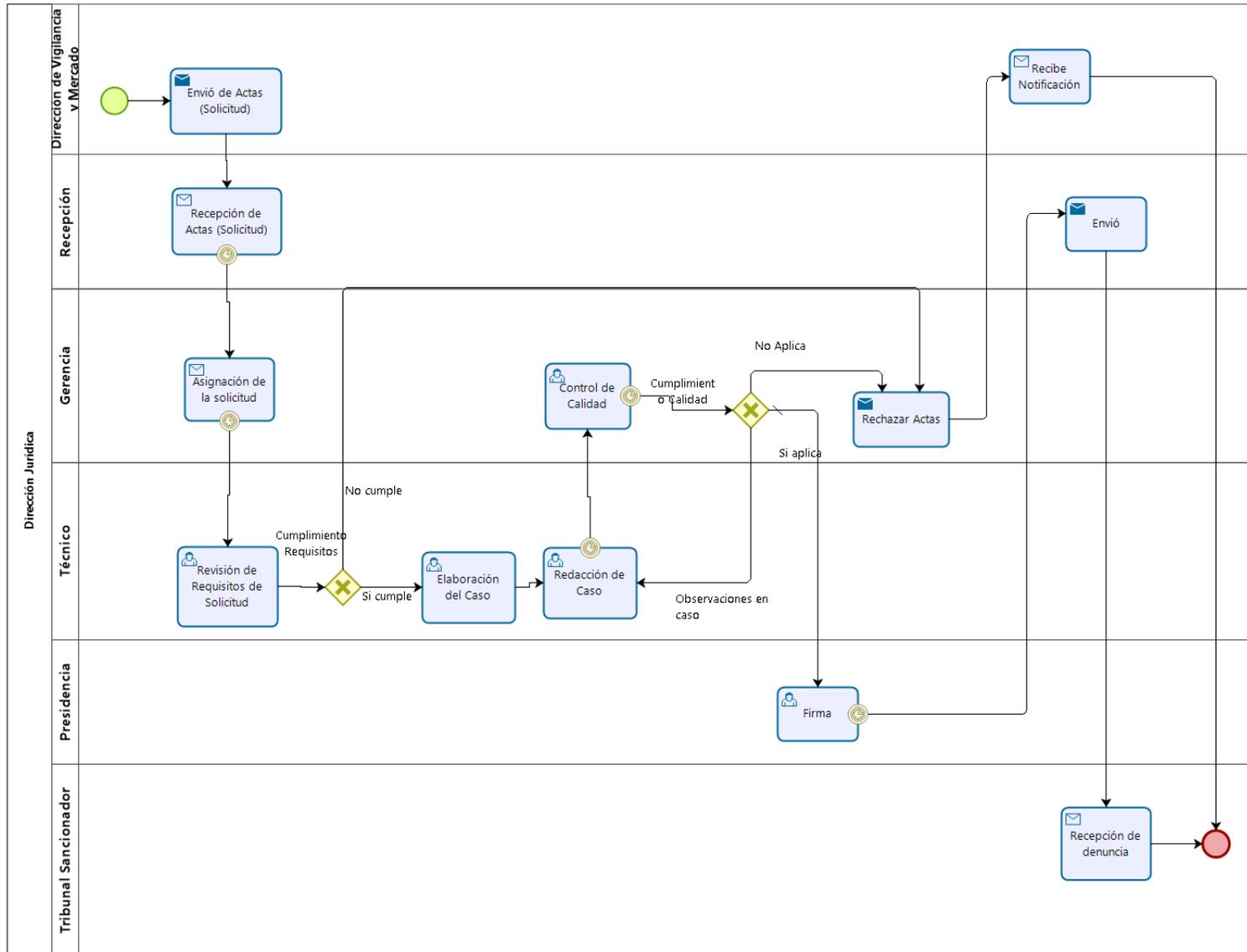


Figura 12.1: Proceso BPMN de la Dirección Jurídica

12.3.2 Paquetes de información

Para la Unidad de Dirección Jurídica de la Defensoría del Consumidor (DC) se obtuvieron dos paquetes de información para los siguientes temas:

- Actas finalizadas por la Unidad.
- Casos finalizados por la Unidad.

La Figura 12.2: Paquete de Información actas finalizadas y Figura 12.3: Paquete de información casos finalizados muestran cada paquete de información.

Tema:		Actas finalizadas por la Unidad Dirección Jurídica.		
JERARQUIAS	Tiempo	Proveedor	Etapas Finalización	Empleado
	Año	Nombre Sector	Nombre Etapa	Zona Departamento
	Semestre	Nombre Comercial	Nombre Subetapa	Departamento de la Oficina
	Trimestre	Departamento Sucursal		Municipio de la Oficina
	Mes	Municipio Sucursal		Oficina del Empleado
	Quincena	Nombre Sucursal		Cargo del Empleado
	Semana			Nombre Empleado
	Día			
	Hechos Medidos:		Cantidad de actas finalizadas, Dias invertidos en dichas actas	

Figura 12.2: Paquete de Información actas finalizadas

Tema:		Casos finalizados por la Unidad Dirección Jurídica.			
JERARQUIAS	Tiempo	Etapas Finalización	Proveedor	Empleado	Tipo de Caso
	Año	Nombre Etapa	Nombre Sector	Zona Departamento	Nombre del tipo de caso
	Semestre	Nombre Subetapa	Nombre Comercial	Departamento de la Oficina	
	Trimestre		Departamento Sucursal	Municipio de la Oficina	
	Mes		Municipio Sucursal	Oficina del Empleado	
	Quincena		Nombre Sucursal	Cargo del Empleado	
	Semana			Nombre Empleado	
	Día				
	Hechos Medidos:		Cantidad de casos finalizadas, Dias invertidos en dichos casos		

Figura 12.3: Paquete de información casos finalizados

12.4 Especificación de Requerimientos

12.4.1 Casos de Uso Análisis de actas finalizadas en la Dirección Jurídica

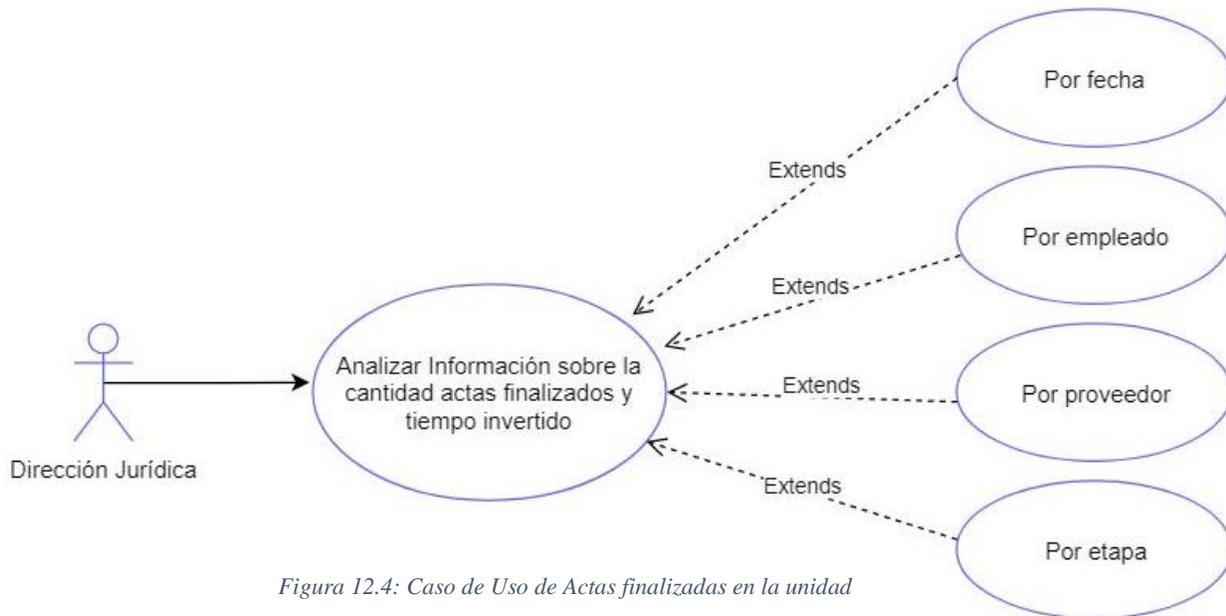


Figura 12.4: Caso de Uso de Actas finalizadas en la unidad

12.4.2 Caso de Uso Análisis de casos finalizados en la Dirección Jurídica

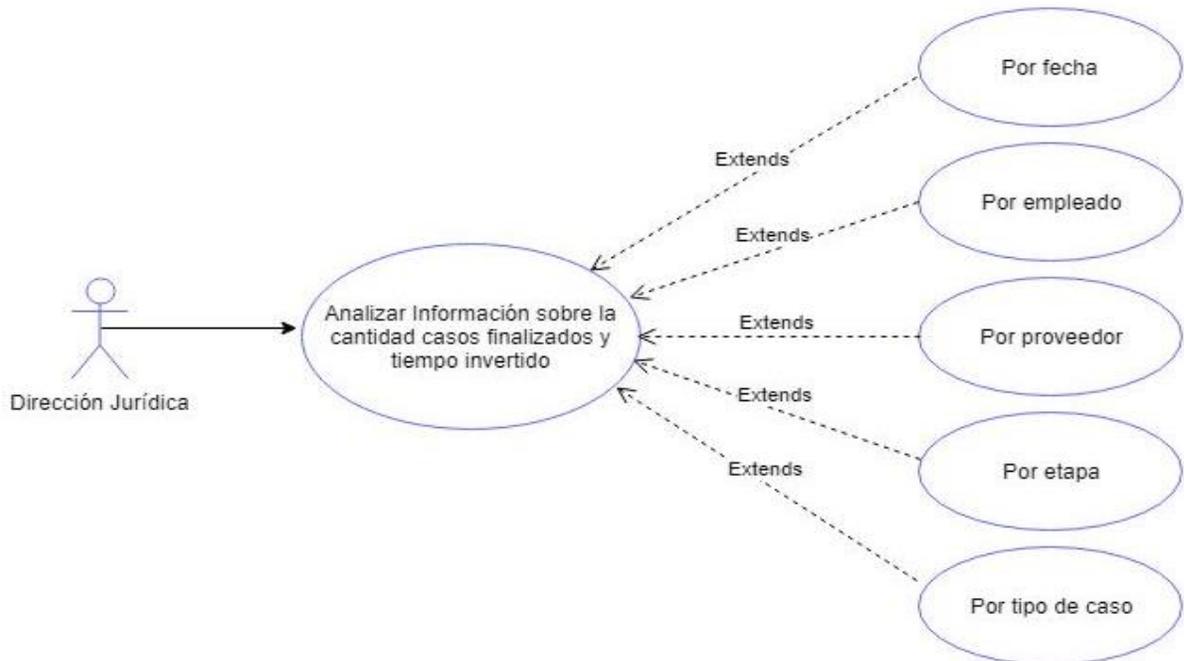


Figura 12.5: Caso de Uso de Casos finalizados en la unidad

12.5 Diseño Multidimensional

12.5.1 Diseño Conceptual del Data Mart

Especialización: Procesar las actas y casos

Hechos principales: Cantidad de actas finalizadas y Cantidad de casos atendidos

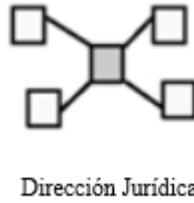


Figura 12.6: Nivel 1 Diagrama Conceptual Data Mart Dirección Jurídica

12.5.1.1 Dimensiones y Hechos

Fact-Acta:

Paquete de Hecho: Actas

Paquetes Dimensionales

1. Empleado
2. Etapa
3. Proveedor
4. Tiempo

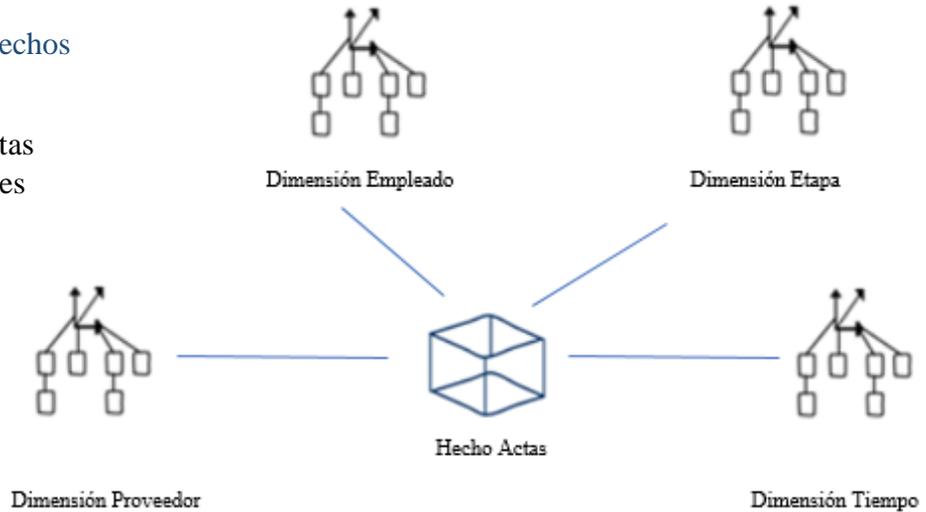


Figura 12.7: Nivel 2 Diagrama Conceptual Data Mart (Fact-Acta)

Fact-Caso:

Paquete de Hecho: Casos

Paquetes Dimensionales:

1. Empleado
2. Etapa
3. Proveedor
4. Tipo Caso
5. Tiempo

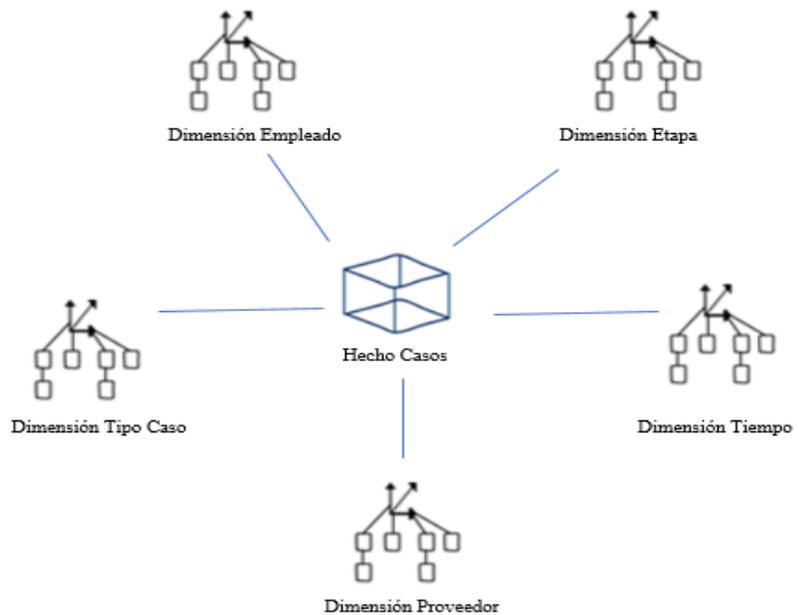


Figura 12.8: Nivel 2 Diagrama Conceptual Data Mart (Fact-Casos)

12.5.1.1.1 Fact Caso

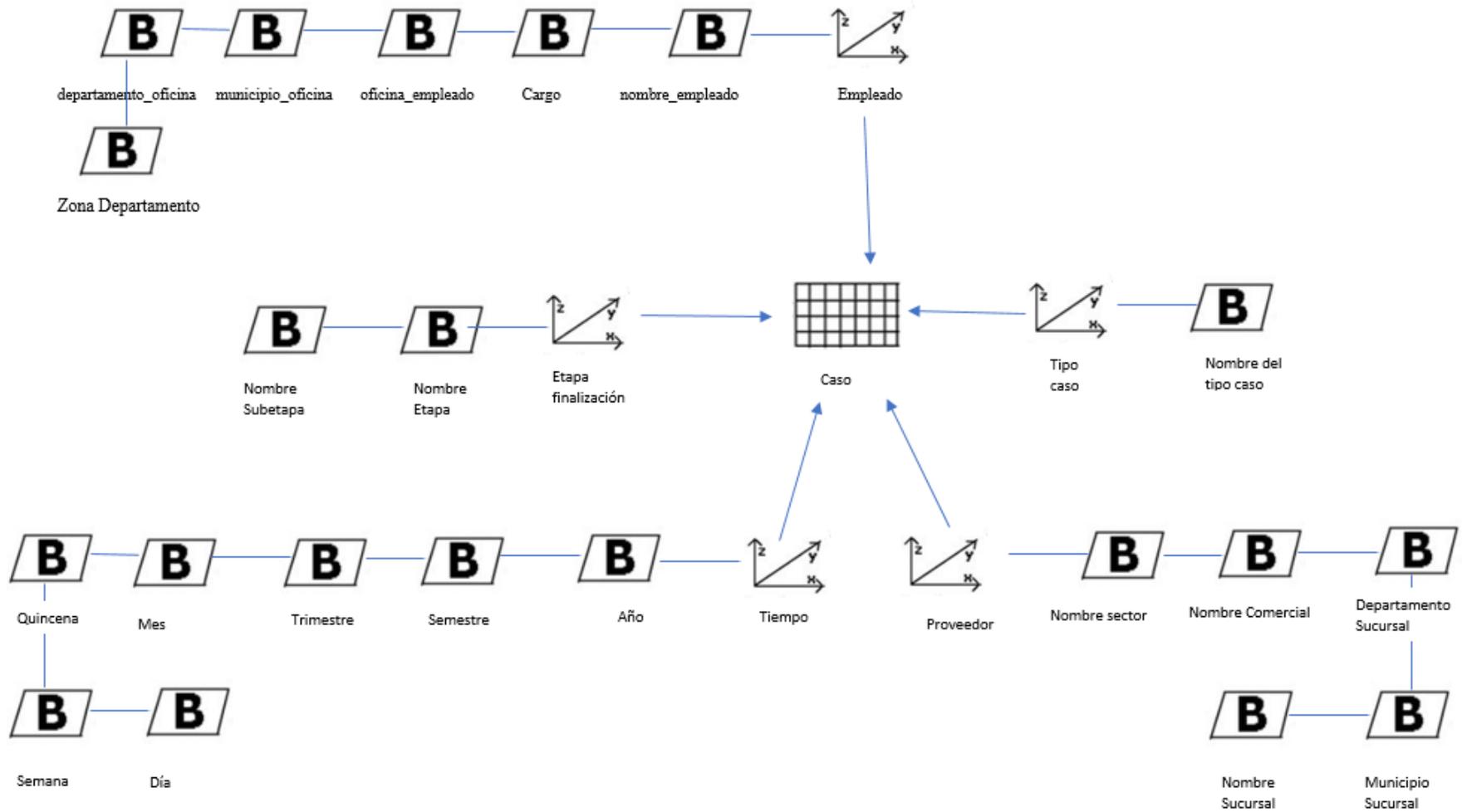


Figura 12.9: Diagrama conceptual Data Mart (Hecho: casos finalizados)

12.5.1.1.2 Fact Acta

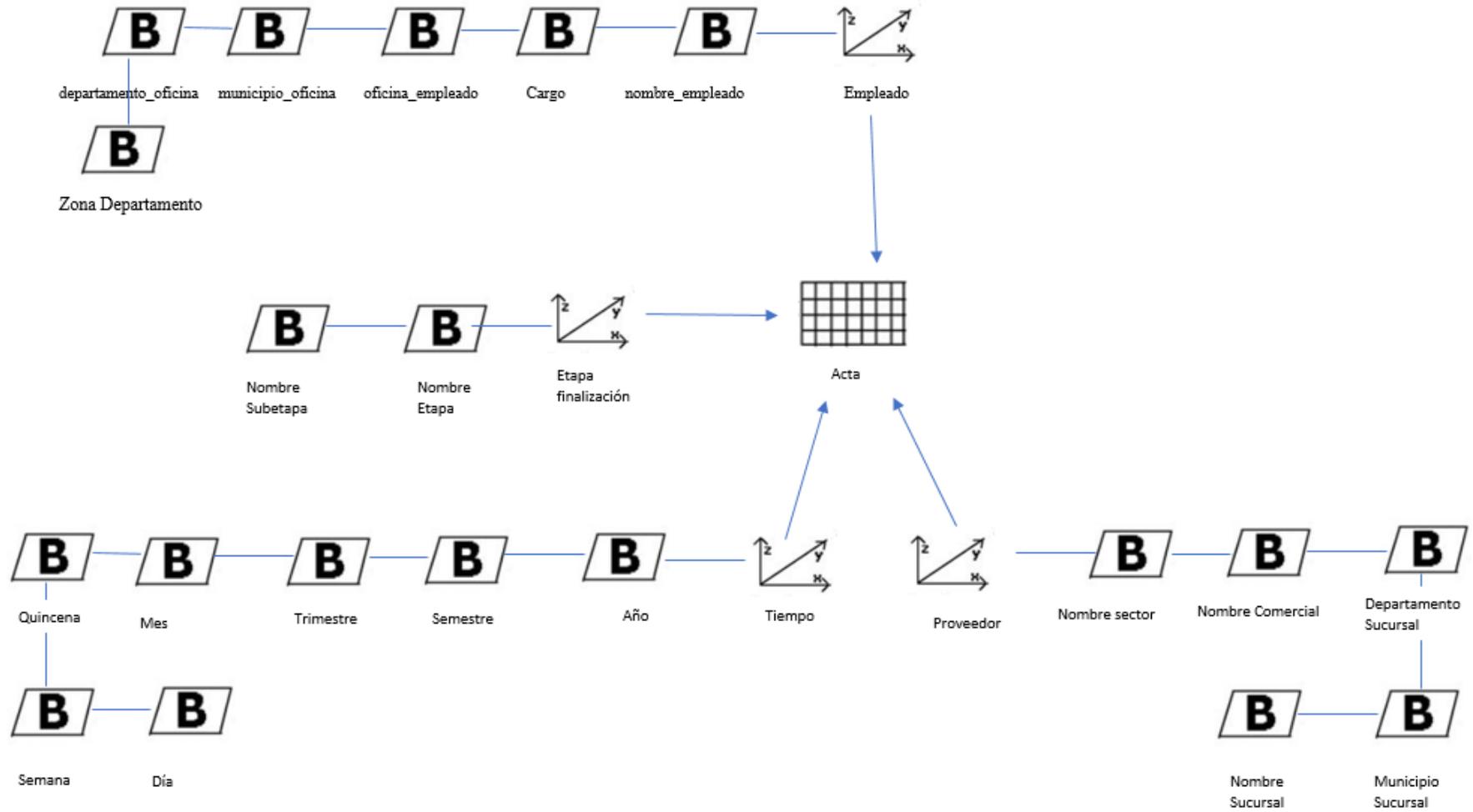


Figura 12.10: Diagrama conceptual Data Mart (Hecho: casos finalizados)

12.5.2 Diseño Físico Base de Datos Data mart



Figura 12.11: Diagrama Físico Base de Datos Datama

12.6 Diseño de Procesos ETL

En el contexto de la inteligencia de negocio, las herramientas ETL han sido la opción usual para alimentar el data Warehouse (Jose Curto). En el proceso de ETL los datos son extraídos desde una base de datos OLTP, transformada para que coincida con el esquema de datos del Data Warehouse y cargados en la base de datos del almacén de datos.

A continuación, se presenta el diseño ETL elaborado para el modelo multidimensional de la unidad de Dirección Jurídica.

12.6.1 Procesos ETL Staging

Los procesos ETL se han dividido en dos grandes grupos, aquellos que mueven datos de las tablas fuentes hacia el Staging Area, los cuales llamaremos ‘Procesos ETL Staging’; así mismo se tienen los aquellos procesos ETL que mueven datos del Staging Area hacia el Data Mart los cuales llamaremos ‘Procesos ETL Data Mart’. En el siguiente apartado se listan los procesos ETL de Staging, seguido de un diagrama representativo para esta unidad y los diseños de los procesos restantes se colocarán en el CD que acompaña al documento en la carpeta [CD-EISI\Diseños de Procesos ETL\dl \(Direccion Juridica\)](#)

Diagramas desarrollados:

St_dimension_tipo_caso
St_dimension_proveedor1
St_fact_caso2

St_dimension_etapa
St_dimension_proveedor2
St_fact_acta1

St_dimension_empleado
St_fact_caso1
St_fact_acta2

12.6.2 Procesos ETL Data Mart

A continuación, se listan aquellos procesos ETL diagramados que mueven datos del Staging Area hacia el Data Mart (Procesos ETL Data Mart) desarrollados para esta unidad, seguido de un ejemplo para ilustrar un proceso de ETL.

Diagramas desarrollados:

Dw_dim_tipo_caso
Dw_dim_proveedor

Dw_dim_etapa
Dw_fact_caso

Dw_dim_empleado
Dw_fact_acta

A continuación, se ilustra uno de los diseños del proceso de ETL desarrollados en la Figura 12.13: Proceso ETL dw_fact_caso

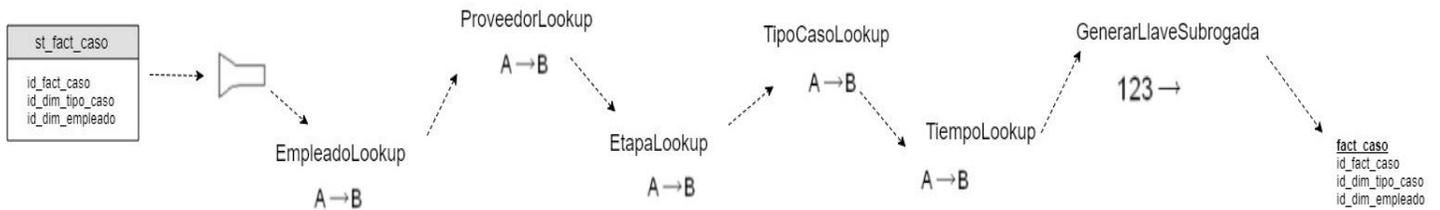


Figura 12.13: Proceso ETL dw_fact_caso

12.7 Pruebas

En las siguientes secciones se presenta las pruebas para cubos, dashboards y procesos ETL.

12.7.1 Pruebas de Procesos ETL

La realización de las pruebas consistió en ejecutar los Jobs almacenados en el servidor de Staging Área. La estructura de directorios que se tiene en dicho servidor dentro del directorio **/root/** se muestra en la imagen Figura 12.14.

```
[root@vps204427 ~]# ls -l
total 12
drwxrwxr-x 18 root root 4096 Sep 20 22:09 data-integration
drwxr-xr-x 3 root root 4096 Sep 21 09:10 jobs
drwxr-xr-x 3 root root 4096 Sep 21 09:11 transformations
[root@vps204427 ~]#
```

Figura 12.14: Directorios Existentes en Staging Área

El contenido de cada directorio se muestra en la tabla Tabla 12.2.

Directorio	Contenido
data-integration	Contiene los archivos ejecutables de Pentaho Data Integration. Aquí se encuentran los archivos necesarios para correr transformaciones y Jobs de PDI.
jobs	Contiene los Jobs de Pentaho Data Integration, los cuales ejecutan otros Jobs y Transformaciones de PDI.
transformations	Contiene las transformaciones de PDI. Estas transformaciones están siendo referenciadas dentro de los Jobs.

Tabla 12.2: Contenido Directorios Existentes en Staging Area

Las transformaciones finales para el Data Mart de Dirección Jurídica se muestran en la Tabla 12.3

No	Transformación	Descripción
1	dw_incremental_dim_empleado.ktr	Transformación para cargas incrementales desde st_dim_empleado (staging area) hacia dim_empleado (data warehouse).
2	dw_incremental_dim_proveedor.ktr	Transformación para cargas incrementales desde st_dim_proveedor (staging area) hacia dim_proveedor (data warehouse).
3	dw_incremental_fact_acta.ktr	Transformación para cargas incrementales desde st_fact_acta (staging area) hacia fact_acta (data warehouse).
4	dw_incremental_fact_caso.ktr	Transformación para cargas incrementales desde st_fact_caso (staging area) hacia fact_caso (data warehouse).
5	dw_inicial_dim_empleado.ktr	Transformación para carga inicial del data warehouse desde st_dim_empleado (staging area) hacia dim_empleado (data warehouse).
6	dw_inicial_dim_proveedor1.ktr	Transformación para carga inicial del data warehouse desde st_dim_proveedor (staging area) hacia dim_proveedor (data warehouse).
7	st_inicial_dim_proveedor2.ktr	Transformación para ingresar un registro adicional por cada registro ingresado en st_inicial_dim_proveedor1.ktr colocando la sucursal en N/A (No Aplica). Es decir cada proveedor tiene su correspondiente registro con una sucursal no especificada.
8	dw_inicial_fact_acta.ktr	Transformación para carga inicial del data warehouse desde st_fact_acta (staging area) hacia fact_acta (data warehouse).
9	dw_inicial_fact_caso.ktr	Transformación para carga inicial del data warehouse desde st_fact_caso (staging area) hacia fact_caso (data warehouse).

10	dw_unico_dim_etapa.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_dim_etapa (staging area) hacia dim_etapa (data warehouse).
11	dw_unico_dim_tipo_caso.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_dim_tipo_caso (staging area) hacia dim_tipo_caso (data warehouse).
12	st_incremental_dim_empleado.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_dim_empleado (staging area).
13	st_incremental_dim_proveedor1.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_dim_proveedor (staging area).
14	st_incremental_dim_proveedor2.ktr	Transformación para ingresar un registro adicional por cada registro ingresado en st_incremental_dim_proveedor1.ktr colocando la sucursal en N/A (No Aplica). Es decir cada proveedor tiene su correspondiente registro con una sucursal no especificada.
15	st_incremental_fact_acta1.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_fact_acta (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta primera parte llena los datos principales requeridos, datos de actas, proveedores, etapa, etc.
16	st_incremental_fact_acta2.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_fact_acta (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta segunda parte del proceso extrae los datos del empleado asociado a las actas y las carga a st_fact_acta.
17	st_incremental_fact_caso1.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_fact_caso (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta primera parte llena los datos principales requeridos, datos de actas, proveedores, etapa, tipo caso, etc.
18	st_incremental_fact_caso2.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_fact_caso (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta segunda parte del proceso extrae los datos del empleado asociado a las actas y las carga a st_fact_acta.
19	st_inicial_dim_empleado.ktr	Transformación para carga inicial del data warehouse desde las tablas fuentes hacia st_dim_empleado (staging area).
20	st_inicial_dim_proveedor.ktr	Transformación para carga inicial del data warehouse desde las tablas fuentes hacia st_dim_proveedor (staging area).
21	st_inicial_fact_acta1.ktr	Transformación para la carga inicial del data warehouse desde tablas fuentes hacia st_fact_acta (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta primera parte llena los datos principales requeridos, datos de actas, proveedores, etapa, etc.
22	st_inicial_fact_acta2.ktr	Transformación para la carga inicial del data warehouse desde tablas fuentes hacia st_fact_acta (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta primera parte llena los datos principales requeridos, datos de actas, proveedores, etapa, etc. Esta segunda parte del proceso extrae los datos del empleado asociado a las actas y las carga a st_fact_acta.
23	st_inicial_fact_caso1.ktr	Transformación para la carga inicial del data warehouse desde tablas fuentes hacia st_fact_caso (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta

		primera parte llena los datos principales requeridos, datos de actas, proveedores, etapa, tipo caso, etc.
24	st_inicial_fact_caso2.ktr	Transformación para la carga inicial del data warehouse desde tablas fuentes hacia st_fact_caso (staging area). Se ha dividido el proceso en dos archivos de transformación diferente debido a su complejidad. Esta primera parte llena los datos principales requeridos, datos de actas, proveedores, etapa, tipo caso, etc. Esta segunda parte del proceso extrae los datos del empleado asociado a las actas y las carga a st_fact_caso.
25	st_unico_dim_etapa.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_etapa (staging area).
26	st_unico_dim_tipo_caso.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_tipo_caso (staging area).

Tabla 12.3: Transformaciones para Data Mart Dirección Jurídica

Los Jobs finales de Primer Nivel para Data Mart de Dirección Jurídica se muestran la Tabla 12.4. También muestra el orden en que ejecutan las transformaciones de cada uno de los Jobs, sin embargo, el orden en que aparecen es el orden de ejecución.

No	Job Primer Nivel	Transformación	
		No	Nombre Transformación
1	job_unico_dim_tipo_caso.kjb	1	st_unico_dim_tipo_caso.ktr
		2	dw_unico_dim_tipo_caso.ktr
2	job_unico_dim_etapa.kjb	1	st_unico_dim_etapa.ktr
		2	dw_unico_dim_etapa.ktr
3	job_inicial_dim_empleado.kjb	1	st_inicial_dim_empleado.ktr
		2	dw_inicial_dim_empleado.ktr
4	job_inicial_dim_proveedor.kjb	1	st_inicial_dim_proveedor1.ktr
		2	st_inicial_dim_proveedor2.ktr
		3	dw_inicial_dim_proveedor.ktr
5	job_incremental_dim_empleado.kjb	1	st_incremental_dim_empleado.ktr
		2	dw_incremental_dim_empleado.ktr
6	job_incremental_dim_proveedor.kjb	1	st_incremental_dim_proveedor1.ktr
		2	st_incremental_dim_proveedor2.ktr
		3	dw_incremental_dim_proveedor.ktr
7	job_inicial_fact_acta.kjb	1	st_inicial_fact_acta1.ktr
		2	st_inicial_fact_acta2.ktr
		3	dw_inicial_fact_acta.ktr
8	job_inicial_fact_caso.kjb	1	st_inicial_fact_caso1.ktr
		2	st_inicial_fact_caso2.ktr
		3	dw_inicial_fact_caso.ktr
9	job_incremental_fact_acta.kjb	1	st_incremental_fact_acta1.ktr
		2	st_incremental_fact_acta2.ktr
		3	dw_incremental_fact_acta.ktr
10	job_incremental_fact_caso.kjb	1	st_incremental_fact_caso1.ktr
		2	st_incremental_fact_caso2.ktr
		3	dw_incremental_fact_caso.ktr

Tabla 12.4: Jobs de Nivel 1 para Data Mart Dirección Jurídica

Los Jobs finales de Segundo Nivel para Data Mart de Dirección Jurídica se muestran la Tabla 12.5. La tabla también muestra el orden en que ejecutan las transformaciones cada uno de los Jobs.

No	Jobs Segundo Nivel	Jobs primer nivel ejecutados	
		No	Nombre Job
1	job_inicial_dl.kjb	1	job_unico_dim_tipo_caso.kjb
		2	job_unico_dim_etapa.kjb
		3	job_inicial_dim_empleado.kjb
		4	job_inicial_dim_proveedor.kjb
		5	job_inicial_fact_caso.kjb
		6	job_inicial_fact_acta.kjb
2	job_incremental_dl.kjb	1	job_unico_dim_tipo_caso.kjb
		2	job_unico_dim_etapa.kjb
		3	job_incremental_dim_empleado.kjb
		4	job_incremental_dim_proveedor.kjb
		5	job_incremental_fact_caso.kjb
		6	job_incremental_fact_acta.kjb

Tabla 12.5: Jobs de Nivel 2 para Data Mart Dirección Jurídica

La figura Figura 12.15 muestra la secuencia del job que realiza la carga inicial, mientras que la Figura 12.16 muestra la secuencia del job que servirá para realizar las cargas incrementales.

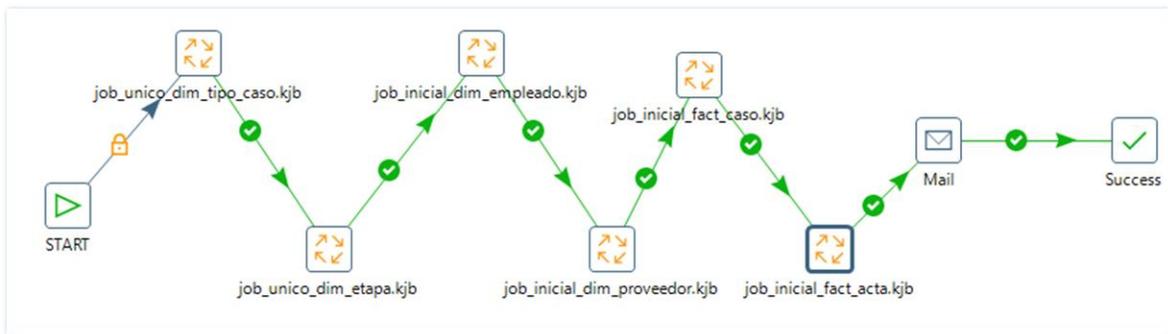


Figura 12.15: Job ETL para la carga inicial del Data Mart Dirección Jurídica

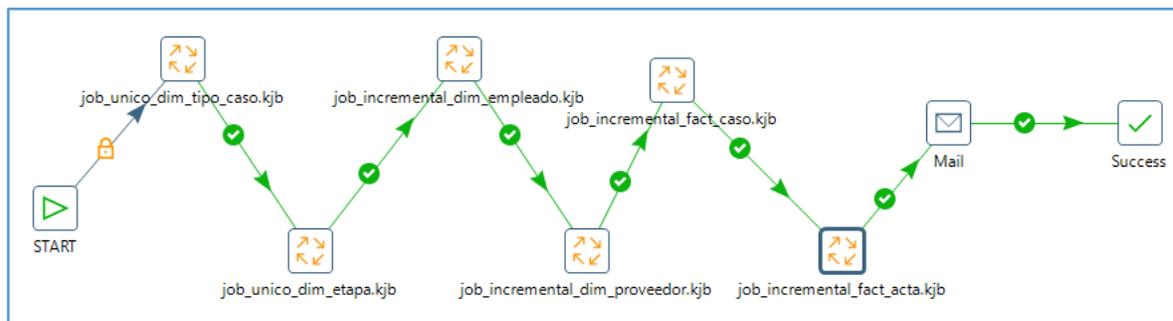


Figura 12.16: Job ETL para la carga incremental del Data Mart Dirección Jurídica

12.7.1.1 Ejecución de Procesos ETL iniciales

El comando para ejecutar el proceso ETL se muestra a continuación:

```
#/root/data-integration/kitchen.sh -file=/root/jobs/dl/job_inicial_dl.kjb -level=Debug
```

Cantidad de transformación	Tiempo	Registros
15	01:41:36.755	100,060

La finalización del proceso ETL se notifica mediante correo electrónico como se muestra en la Figura 12.17.

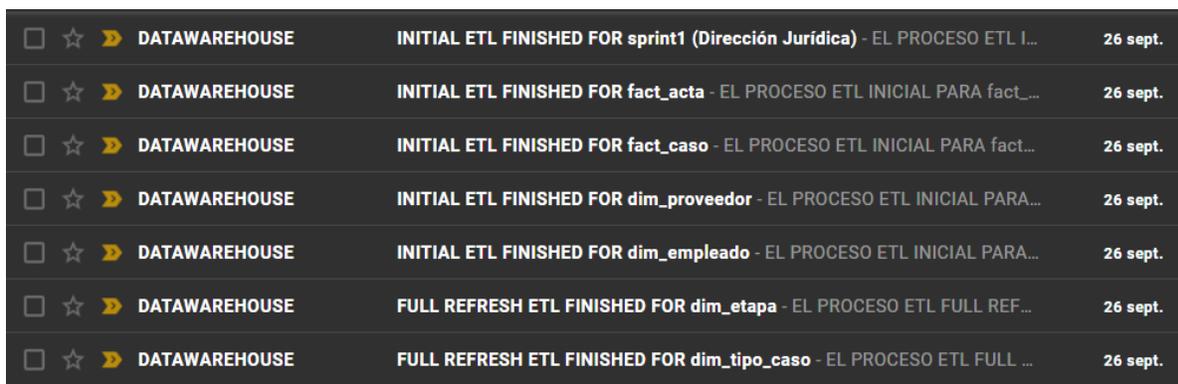


Figura 12.17: Notificación de finalización de los procesos ETL (Carga Inicial)

12.7.2 Pruebas Cubos OLAP

Dado que para Dirección Jurídica se identificaron dos hechos que se desean analizar, se han realizado dos cubos para dicha unidad. A continuación, se muestran los cubos realizados para el Data Mart de Dirección Jurídica.

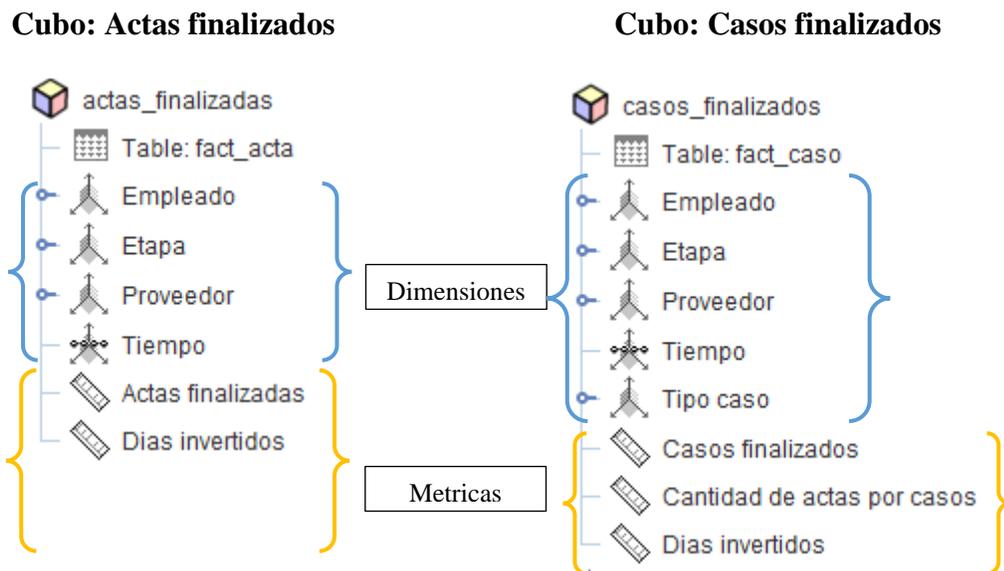


Tabla 12.6: Cubos para Dirección Jurídica

Por cada uno de ellos se ha publicado un archivo XML en el Servidor de Pentaho Server BI.

Con el uso del Plugin JPivot para Pentaho Server se logra visualizar la información según la Figura 12.18 y la Figura 12.19.

Empleado	Etapa	Proveedor	Tiempo	Medidas
▣ Todas las zonas	▣ Todas las etapas	▣ Todos los sectores	▣ Todos los años	• Actas finalizadas
			▣ 2010	4.144
			▣ Segundo semestre	82
			▣ Cuarto trimestre	82
			▣ Diciembre	48
			▣ Noviembre	18
			▣ Primera quincena	29
			▣ Segunda quincena	1
			▣ Octubre	28
			▣ Tercer trimestre	1
			▣ 2011	34
			▣ 2012	801
			▣ 2013	1.062
			▣ 2014	938
			▣ 2015	1.105
			▣ 2017	131
			▣ 2018	24
				1

Figura 12.18: Cubo Actas Finalizadas Dirección Jurídica

Empleados	Etapa	Proveedores	Tiempo	Tipo Caso	Medidas
▣ Todas las zonas	▣ Todas las etapas	▣ Todos los sectores	▣ Todos los años	▣ Todos los tipos	21
				Oficio	20
				Oficio con medida cautelar	1
			▣ 2010	▣ Todos los tipos	1
				Oficio	1
			▣ Segundo semestre	▣ Todos los tipos	1
				Oficio	1
			▣ Tercer trimestre	▣ Todos los tipos	1
				Oficio	1
			▣ Septiembre	▣ Todos los tipos	1
				Oficio	1
			▣ 2011	▣ Todos los tipos	7
				Oficio	6
				Oficio con medida cautelar	1
			▣ 2012	▣ Todos los tipos	6
				Oficio	6
			▣ 2013	▣ Todos los tipos	7
				Oficio	7

Figura 12.19: Cubo Casos Finalizados Dirección Jurídica

12.7.3 Pruebas Dashboard

A continuación, se muestra uno de los Dashboards para Dirección Jurídica.

12.7.3.1 Dashboard Dirección Jurídica

Cantidad de Actas Finalizadas por Empleado

Fecha Inicio: 01/06/2010 Fecha Fin: 17/06/2014 Exportar

Show 10 entries Search:

Empleado	Cantidad Actas	Días Promedio / Acta
MARCELA DIAZ	197	33
CAROL IVETTE ESTRADA RODEZNO	445	59
CINTHIA MELISSA CANAS GARCIA	14	59
PRACTICA2DJ PRACTICA2DJ	494	108
JOSE EMILIANO AREVALO	688	113
ERNESTO URRUTIA	173	204
ERICK GAVIDIA	45	37
PRACTICA1DJ PRACTICA1DJ	297	218
ADA BRACAMONTE	53	51
SIN ASIGNAR	2	77

Showing 1 to 10 of 13 entries Previous 1 2 Next

Actas finalizadas por empleado

Cantidad de Actas Finalizadas por Sub Etapa

Fecha Inicio: 01/06/2010 Fecha Fin: 17/06/2014 Exportar

Show 10 entries Search:

Subetapa Finalización	Cantidad Actas
ELABORACION DE CASO	3307
DOCUMENTO NO CUMPLE REQUISITOS	58

Showing 1 to 2 of 2 entries Previous 1 Next

Actas finalizadas por etapa

13 SPRINT 2

13.1 Sprint Backlog

A continuación, se presenta el Sprint Backlog, que refleja las historias de usuarios, su estimación y asignación para este Segundo Sprint.

#	Historia	Punto Estimado	Asignado
1.	Refinamiento del requerimiento de información	20	Jorge – Ivis – Samuel – Esaú
2.	Poblar la base de datos de desarrollo SARA	13	Jorge – Ivis – Samuel
3.	Diseño del modelo multidimensional	8	Ivis
4.	Diseño BD del Stagin Area	3	Jorge – Samuel
5.	Diseño del proceso ETL	8	Ivis – Samuel
6.	Diseño de vistas	3	Ivis
7.	Creación de Script SQL modelo multidimensional	13	Jorge – Samuel
8.	Creación de Script SQL del Stagin Area	13	Jorge – Samuel
9.	Desarrollo de los procesos ETL	100	Jorge – Samuel – Ivis
10.	Desarrollo de cubos OLAP	40	Ivis
11.	Desarrollo de vistas	13	Ivis
12.	Implementación e integración de componentes	40	Jorge – Samuel – Ivis
13.	Pruebas	20	Samuel
14.	Documentación	40	Jorge – Samuel – Ivis – Esaú

Tabla 13.1: Sprint Backlog, Sprint 2

13.2 Descripción de Historias de Usuario

A continuación, se presenta la descripción de las historias de usuario detallando: el responsable, al sprint al que pertenece, una breve descripción, sus criterios de validación y los puntos de estimación. Previamente a las historias de usuario se presenta la historia épica del Sprint 2

<i>Historia Épica #2</i>	Data Mart Dirección de Ciudadanía y Consumo.
<i>Responsable</i>	Jorge Iraheta/Esaú Serpas/Ivis Navarro/Samuel Núñez Sprint: 2
<i>Descripción</i>	Como Unidad de Ciudadanía y Consumo Necesito: Un mecanismo para analizar la información de atenciones (asesorías, denuncias, denuncias no personales, gestiones, etc) desde diferentes perspectivas. Para: la evaluación de resultados de dichas atenciones y asesorías realizadas por la Defensoría en diferentes temporadas o períodos de tiempo o campañas móviles.
<i>Criterios de aceptación</i>	- Las atenciones recibidas deben poder ser analizadas desde las siguientes perspectivas: <ul style="list-style-type: none"> • Por fecha • Por lugar Móvil • Por tipo de atención (asesorías, gestiones, denuncias, etc.) • Por forma de recepción. • Oficina Recepción. • Oficina Recepción (según empleado). • Proveedor • Consumidor • Clasificación (Sector y Categoría) • Motivo
<i>Puntos estimados</i>	334

<i>Historia de usuario #1</i>	Refinamiento del requerimiento de información
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/ Samuel Núñez /Esaú Serpas Sprint: 2
<i>Descripción</i>	<p>Como Equipo Scrum necesito el refinamiento del requerimiento de información de Unidad de Ciudadanía y Consumo: Realizar reuniones con los interesados y especificar los elementos de información requeridos para la Unidad de Ciudadanía y Consumo.</p> <ul style="list-style-type: none"> • Llenar la herramienta de recolección. • Revisar y desglosar los requerimientos solicitados. • Firmar documento por el usuario como aceptación de los de los requerimientos. <p>Para tener claridad de la información que desean almacenar, visualizar y analizar mediante el Data Mart o modelo Multidimensional para ésta unidad.</p>
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> - Tener llena la Plantilla de refinamiento de requerimientos de información. - Incluir únicamente requerimientos de información con naturaleza histórica y no información transaccional. - Tener un documento firmado de aceptación de requerimientos. - Identificar (mediante un listado) las tablas fuentes que servirán para cumplir con los requerimientos de información solicitados.
<i>Puntos estimados</i>	20

<i>Historia de usuario #2</i>	Poblar base de datos de desarrollo SARA
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/ Samuel Núñez Sprint: 2
<i>Descripción</i>	<p>Como Product Owner necesito: migrar los datos necesarios según las tablas fuentes de datos que se identifiquen durante el proceso de toma de requerimientos. Para poseer datos de prueba para el desarrollo del Data Mart.</p>
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> - Se deben poseer Scripts de INSERT para las tablas Fuente. - Todas las tablas fuentes identificadas en el refinamiento de requerimientos deben poseer datos. - Debe existir al menos una prueba de conexión y acceso remoto exitoso a los datos. - Datos cargados en la Base de datos
<i>Puntos estimados</i>	13

<i>Historia de usuario #3</i>	Diseño del modelo multidimensional
<i>Responsable</i>	Ivis Navarro Sprint: 2
<i>Descripción</i>	<p>Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de la Unidad de Ciudadanía y Consumo. Para tener la base (diagrama) para el desarrollo de Scripts de Base de Datos SQL para Data Mart.</p>
<i>Criterios de aceptación</i>	<ul style="list-style-type: none"> -Esquema del modelo multidimensional de la Unidad de Ciudadanía y Consumo Finalizado. -El esquema debe cumplir las reglas de un modelo de tipo Estrella. -El esquema debe poseer las relaciones necesarias para la búsqueda según los criterios de búsqueda definidos por la Unidad de Ciudadanía y Consumo. -El esquema debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
<i>Puntos estimados</i>	8

<i>Historia de usuario #4</i>	Diseño BD del Staging Area.
Responsable	Jorge Iraheta/Samuel Núñez Sprint: 2
Descripción	Como Equipo Scrum Necesito: el diseño del modelo multidimensional para el Data Mart de la Unidad de Ciudadanía y Consumo. Para tener el insumo para la creación de Scripts de Base de Datos SQL para Staging Area.
Criterios de aceptación	-Deberá existir una o más tablas por cada dimensión. -Deberá existir una o más tablas de hecho. -El diseño debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
Puntos estimados	3

<i>Historia de usuario #5</i>	Diseño del proceso ETL
Responsable	Ivis Navarro/Samuel Núñez Sprint: 2
Descripción	Como Equipo Scrum necesito: Diseñar los Proceso ETL <ul style="list-style-type: none"> • Extracción • Transformación • Carga Para cumplir con el Diseño del proceso ETL y tenerlo como insumo a la hora de desarrollar dichos procesos ETL.
Criterios de aceptación	-Se deberá utilizar nomenclatura UML para los pasos en cada proceso ETL. -Deberá incluir todas las tablas dimensiones y tablas de hechos. -Los nombres de las tablas o etiquetas deben poderse visualizar fácilmente.
Puntos estimados	8

<i>Historia de usuario #6</i>	Diseño de vistas
Responsable	Ivis Navarro Sprint: 2
Descripción	Como Administrador de Data Warehouse necesito: Realizar el diseño de vistas necesarias para visualizar la información del Data Mart de dirección Jurídica. Para Diseñar y digitalizar la vista de presentación de datos.
Criterios de aceptación	-Deberá existir un diseño de vista o dashboard para cada hecho a medir. -Deberá explicarse cada sección del dashboard o vista. -Etiquetas y textos deben ser legibles a simple vista.
Puntos estimados	3

<i>Historia de usuario #7</i>	Creación de Script SQL Modelo Multidimensional
Responsable	Samuel Núñez/Jorge Iraheta Sprint: 2
Descripción	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será almacenada la información.
Criterios de aceptación	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Modelo Multidimensional. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
Puntos estimados	13

<i>Historia de usuario #8</i>	Creación de Script SQL del Staging Area
<i>Responsable</i>	Samuel Núñez/Jorge Iraheta Sprint: 2
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será procesada, calculada y/o limpiada la información.
<i>Criterios de aceptación</i>	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Staging Area. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
<i>Puntos estimados</i>	13

<i>Historia de usuario #9</i>	Desarrollo de los procesos ETL.
<i>Responsable</i>	Samuel Núñez/Jorge Iraheta/ Ivis Navarro Sprint: 2
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar los procesos ETL correspondientes Para obtener, trasladar y transformar los datos que serán almacenados en el Data Mart de la Unidad de Ciudadanía y Consumo.
<i>Criterios de aceptación</i>	-Los nombres de los procesos ETL deben poseer la nomenclatura definida por el equipo. -Los procesos ETL deben poseer la documentación necesaria en forma de comentarios dentro de la herramienta PDI. -Las Transformaciones ETL deben haber sido probados en desarrollo (Equipo Local) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de final . -Las transformaciones correspondientes a una dimensión o hecho deben ser incorporados un Job de PDI. -Deberá crearse un Job que incorpore los otros Jobs correspondientes a cada dimensión y hecho. -Los Jobs de PDI deben haber sido probados con las bases de datos de prueba y base de datos final. -Las bitácoras de los procesos ETL deben almacenarse en una base de datos.
<i>Puntos estimados</i>	100

<i>Historia de usuario #10</i>	Desarrollo de Cubos OLAP
<i>Responsable</i>	Ivis Navarro Sprint: 2
<i>Descripción</i>	Como Equipo Scrum necesito: investigar sobre la herramienta para el desarrollo de cubos OLAP e implementar dicha herramienta para la creación de cubos OLAP. Para implementar el cubo correspondiente a ciudadanía y consumo.
<i>Criterios de aceptación</i>	-El cubo deberá poderse visualizar correctamente y sin errores. -Deberá existir como mínimo un cubo OLAP por Hecho. -Deberán poderse realizar las operaciones como mínimo de: Roll Up y Drill Down.
<i>Puntos estimados</i>	40

<i>Historia de usuario #11</i>	Desarrollo de vistas
Responsable	Ivis Navarro Sprint: 2
Descripción	Como Equipo Scrum necesito: Desarrollar vistas o dashboards según el diseño previamente elaborado. Para Presentar información relevante hacia la unidad para la toma de decisiones.
Criterios de aceptación	-Vista o dashboard funcionando sin errores técnicos. -Debe coincidir con el diseño realizado previamente. -Debe existir al menos un dashboar por tabla de hechos.
Puntos estimados	13

<i>Historia de usuario #12</i>	Implementación e integración de componentes
Responsable	Ivis Navarro/Samuel Núñez/Jorge Iraheta Sprint: 2
Descripción	Como Equipo Scrum necesito: La integración de todos los elementos del modelo multidimensional <ul style="list-style-type: none"> - ETL - Fuentes - Stagin Área - BD multidimensional - Pentaho Server (Vistas) Para el desarrollo del modelo multidimensional de forma integral y probar el funcionamiento correctamente de la totalidad del Data Mart.
Criterios de aceptación	- Se deberán modificar los procesos ETL para que utilicen las bases de datos finales (no las de pruebas). - Se deberán realizar pruebas en la visualización de los dashboards. - Todo el entorno (desde los procesos ETLs, hasta los dashboards o vistas) deberán tener el visto bueno de todo el equipo.
Puntos estimados	40

<i>Historia de usuario #13</i>	Pruebas
Responsable	Samuel Nuñez Sprint: 2
Descripción	Como Equipo Scrum necesito: Realizar pruebas al modelo multidimensional terminado. Para verificar que funciona de la manera que se espera.
Criterios de aceptación	Pruebas exitosa en pruebas de aceptación del usuario.
Puntos estimados	20

<i>Historia de usuario #14</i>	Documentación
Responsable	Samuel Nuñez/Jorge Iraheta/Ivis Navarro/Esau Serpas Sprint: 2
Descripción	Como Equipo Scrum necesito: Documentación completa con historias, nomenclatura y diagramas empleados. Para verificar cómo se llevó a cabo el desarrollo de los cubos de la segunda unidad (Unidad de Ciudadanía y Consumo).
Criterios de aceptación	Documentos completos de análisis y diseño para modelo multidimensional
Puntos estimados	40

13.3 Refinamiento del Requerimiento de Información

Para la realización del refinamiento de los requerimientos de información se partió con la investigación del proceso general que sigue la Dirección de Ciudadanía y Consumo dentro de la Defensoría del Consumidor. Además de establecer el proceso general de la unidad se hizo uso de los paquetes de información para refinar el requerimiento de información.

13.3.1 Proceso BPMN de la Unidad

A continuación, presentamos el proceso de Dirección de Ciudadanía y Consumo expresado en BPNM

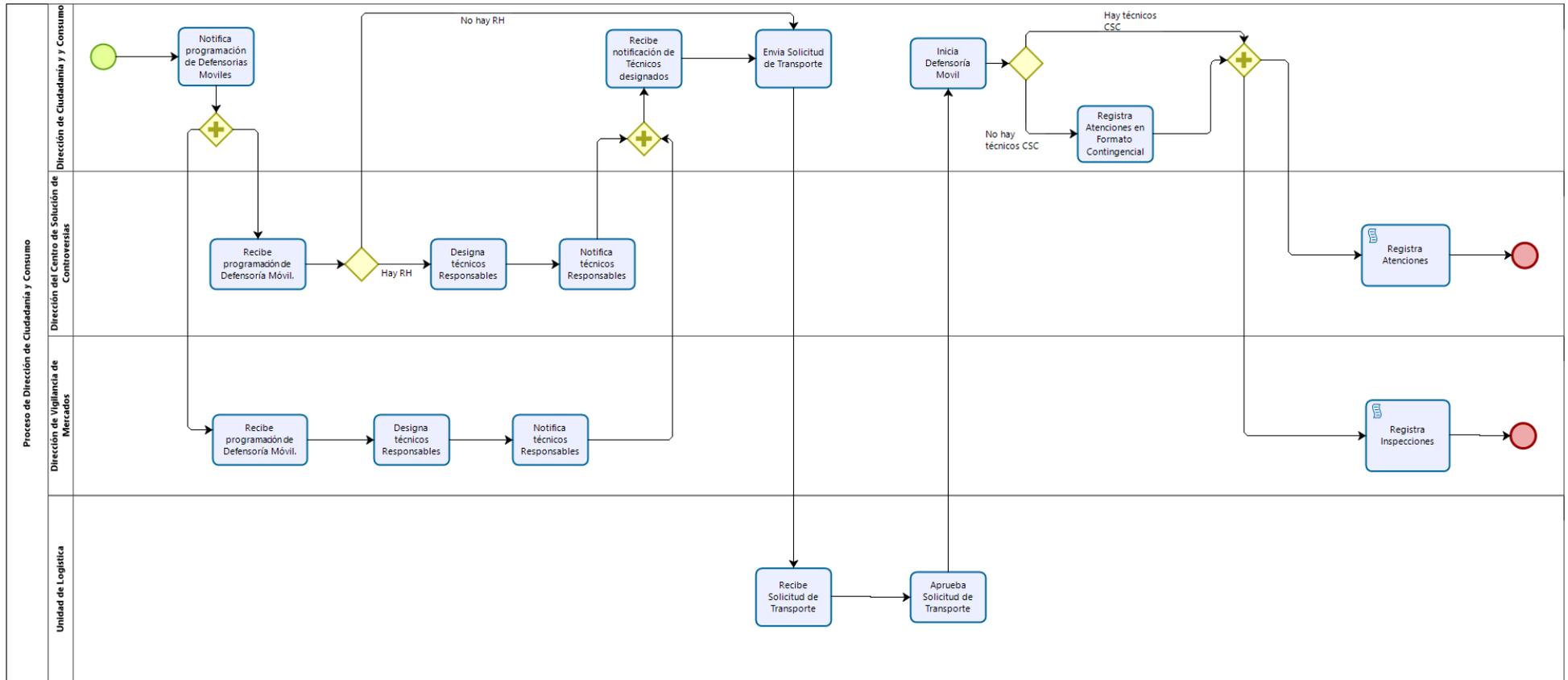


Figura 13.1 Proceso de Dirección de Ciudadanía y Consumo

13.3.2 Paquetes de información

A continuación, se muestra el paquete de información correspondiente a Ciudadanía y Consumo:

Tema:		Cantidad de atenciones recibidas en la Defensoría del consumidor.							
	Tiempo	Lugar Movil	Tipo Atencion	Forma de Recepción	Oficina Recepción	Proveedor	Consumidor	Clasificación	Motivo
J E R A R Q U I A S	Año	Nombre Lugar	Nombre Sector	Zona Departamento	Zona	NombreSector	Sexo	Nombre sector	Nombre Motivo
	Semestre		Nombre Comercial	Departamento de la Oficina	Departamento	NombreComercial	Zona	Nombre categoría	
	Trimestre		Departamento Sucursal	Municipio de la Oficina	Municipio	RazonSocial	Departamento		
	Mes		Municipio Sucursal	Oficina del Empleado	Oficina	NIT	Municipio		
	Quincena		Nombre Sucursal	Cargo del Empleado		DepartamentoSucursal			
	Semana			Nombre Empleado		MunicipioSucursal			
	Día					NombreSucursal			
	Hechos Medidos:	Cantidad de atenciones recibidas							

Figura 13.2: Paquete de información para cantidad de atenciones recibidas en la DC

13.4 Especificación de Requerimientos

13.4.1 Caso de uso de cantidad de atenciones recibidas en la defensoría del consumidor.

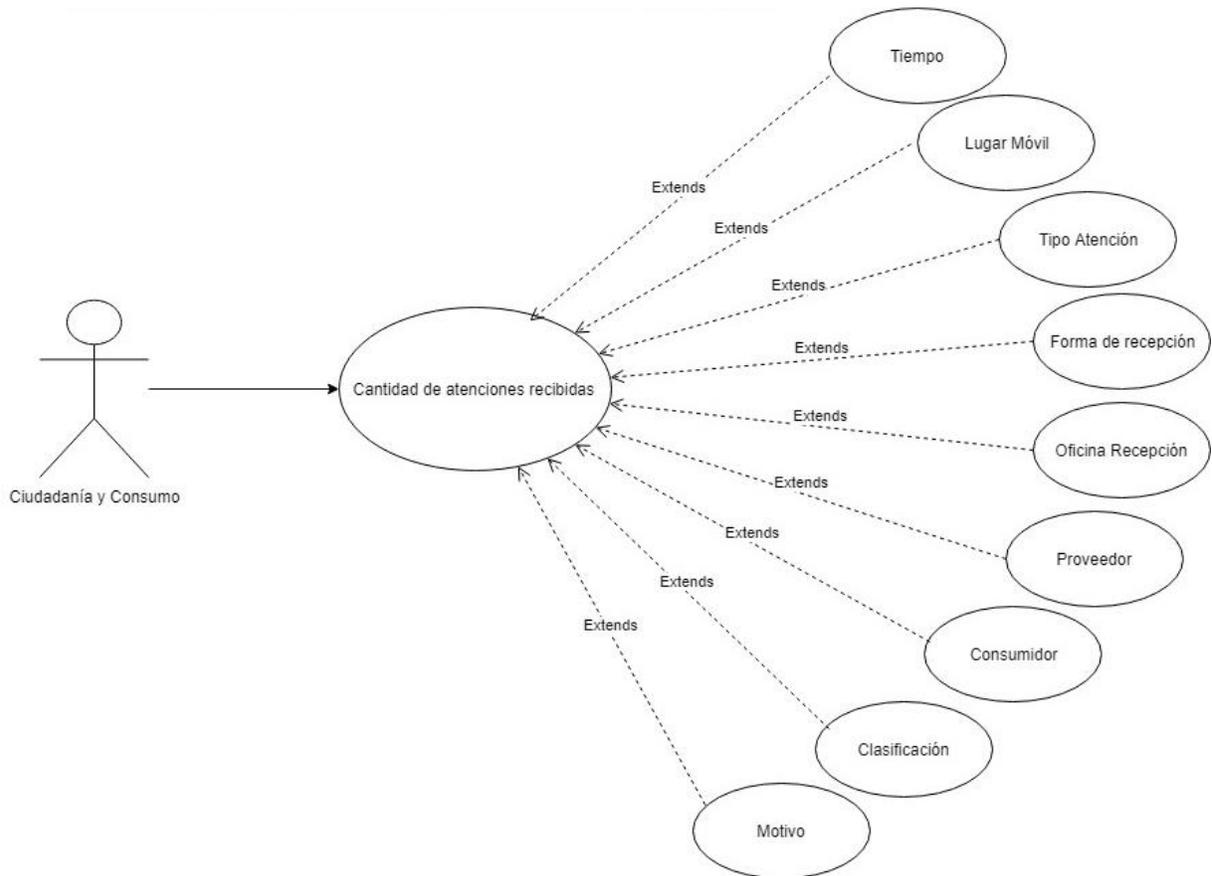


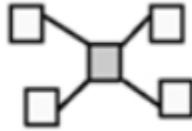
Figura 13.3: Caso de Uso de Dirección de Ciudadanía y Consumo

13.5 Diseño Multidimensional

13.5.1 Diseño Conceptual del Data Mart

Especialización: Analizar las atenciones recibidas

Hechos principales: Cantidad de atenciones recibidas



Ciudadanía y Consumo

Figura 13.4: Nivel 1 Diagrama conceptual Data Mart Ciudadanía y Consumo

13.5.1.1 Dimensiones y Hechos

Fact-Atenciones:

Paquete de Hecho: Cantidad de atenciones recibidas

Paquetes Dimensionales:

1. Tiempo
2. Lugar móvil
3. Tipo atención
4. Forma recepción
5. Oficina recepción
6. Proveedor
7. Consumidor
8. Clasificación
9. Motivo

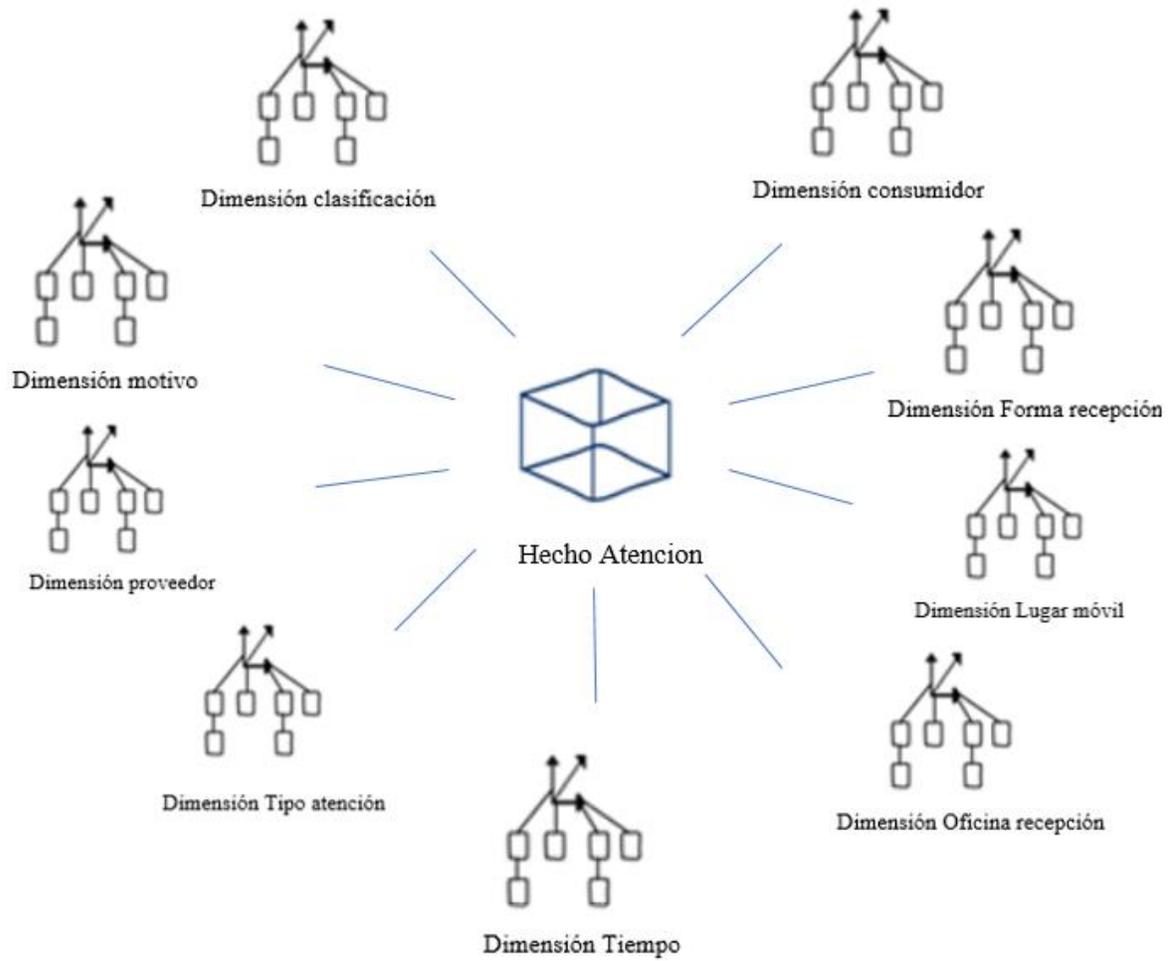


Figura 13.5: Nivel 2 Diagrama Conceptual Data Mart (Fact-Atencion)

13.5.1.1.1 Fact Atenciones

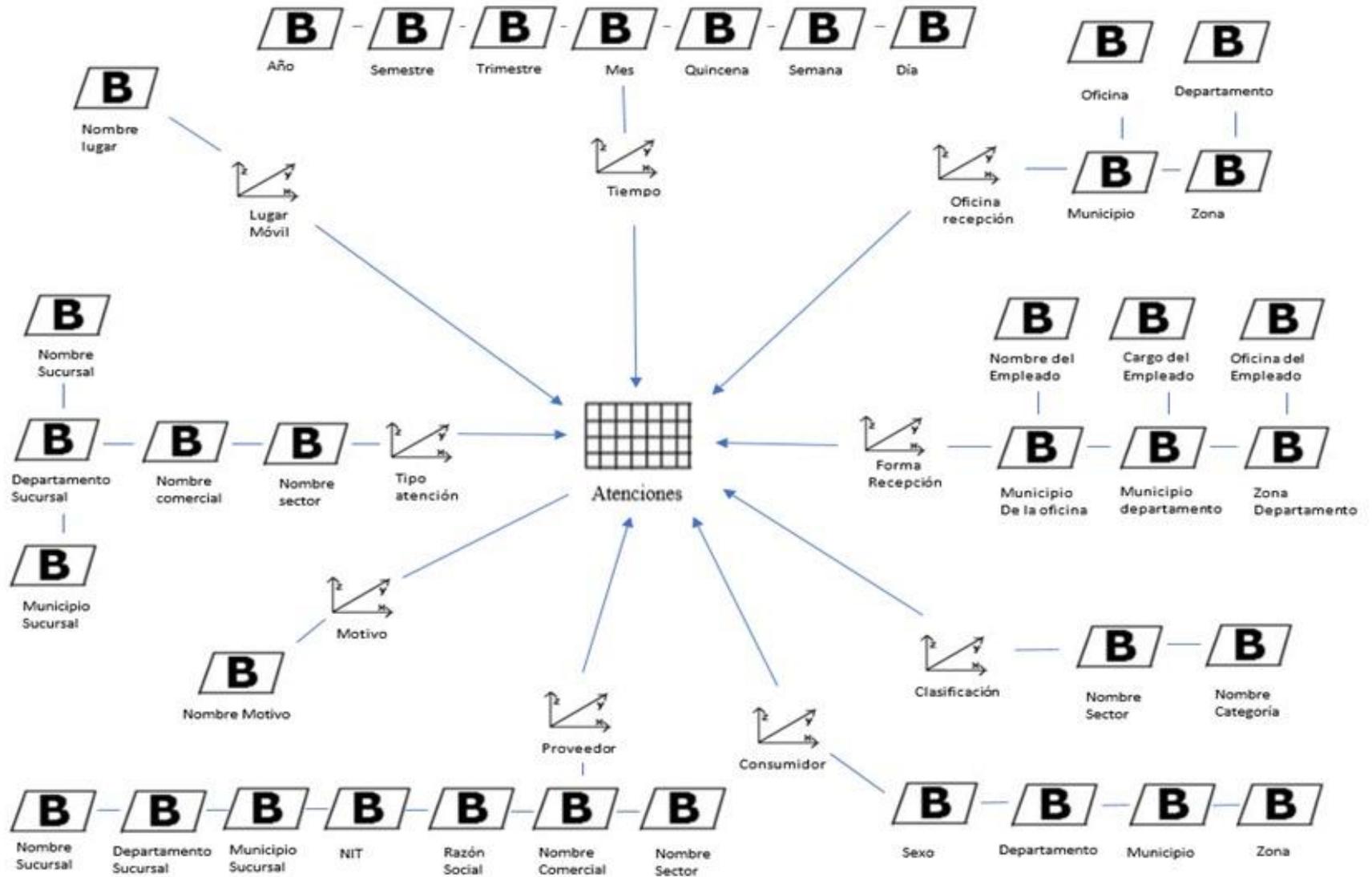


Figura 13.6: Nivel 3 Diagrama conceptual Data Mart Ciudadanía y Consumo

13.5.2 Diseño Físico Base de Datos Data mart

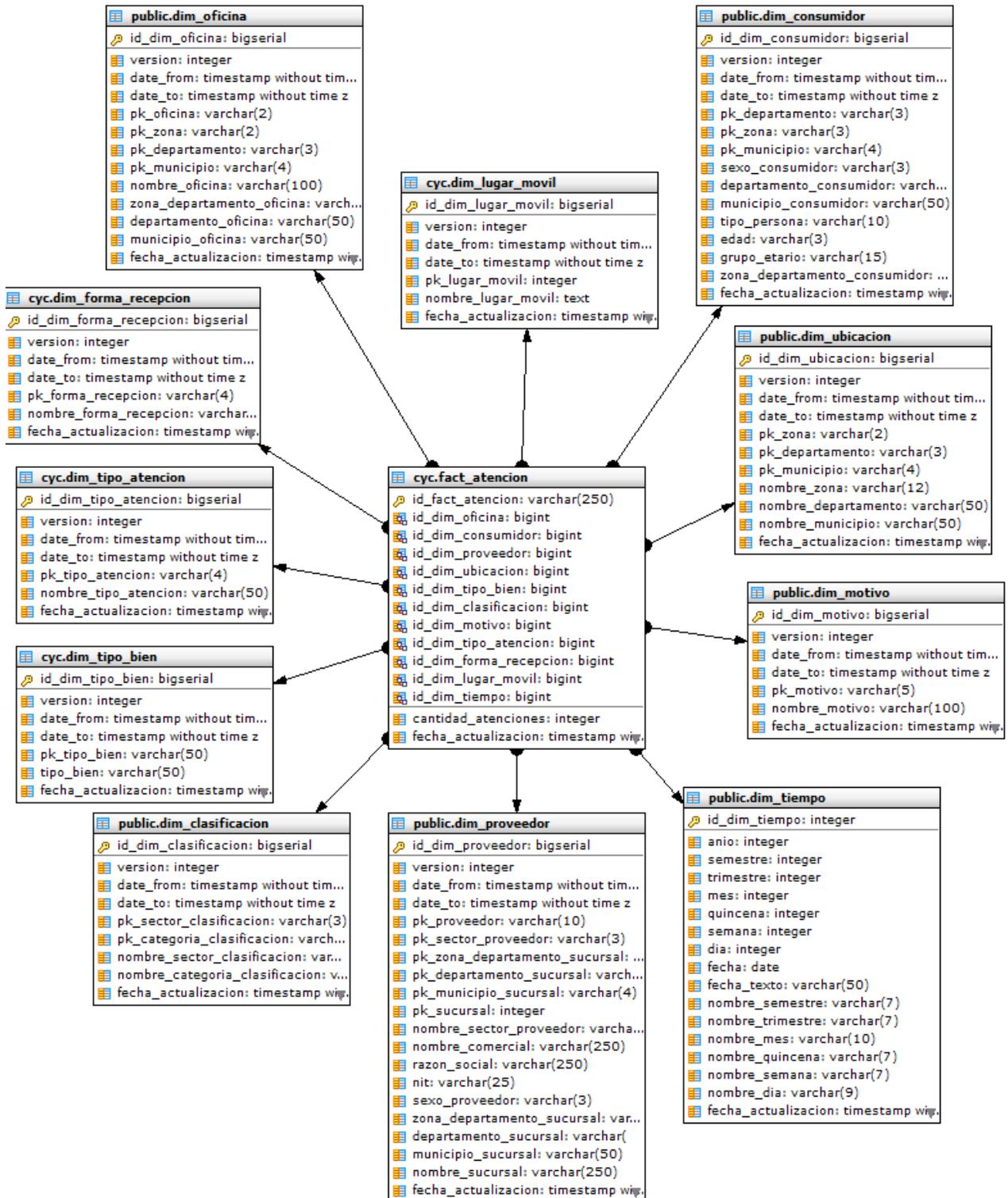


Figura 13.7: Diseño físico de base de datos Dirección de Ciudadanía y Consumo (Data Mart)

13.5.3 Diseño Físico Base de Datos Staging Area

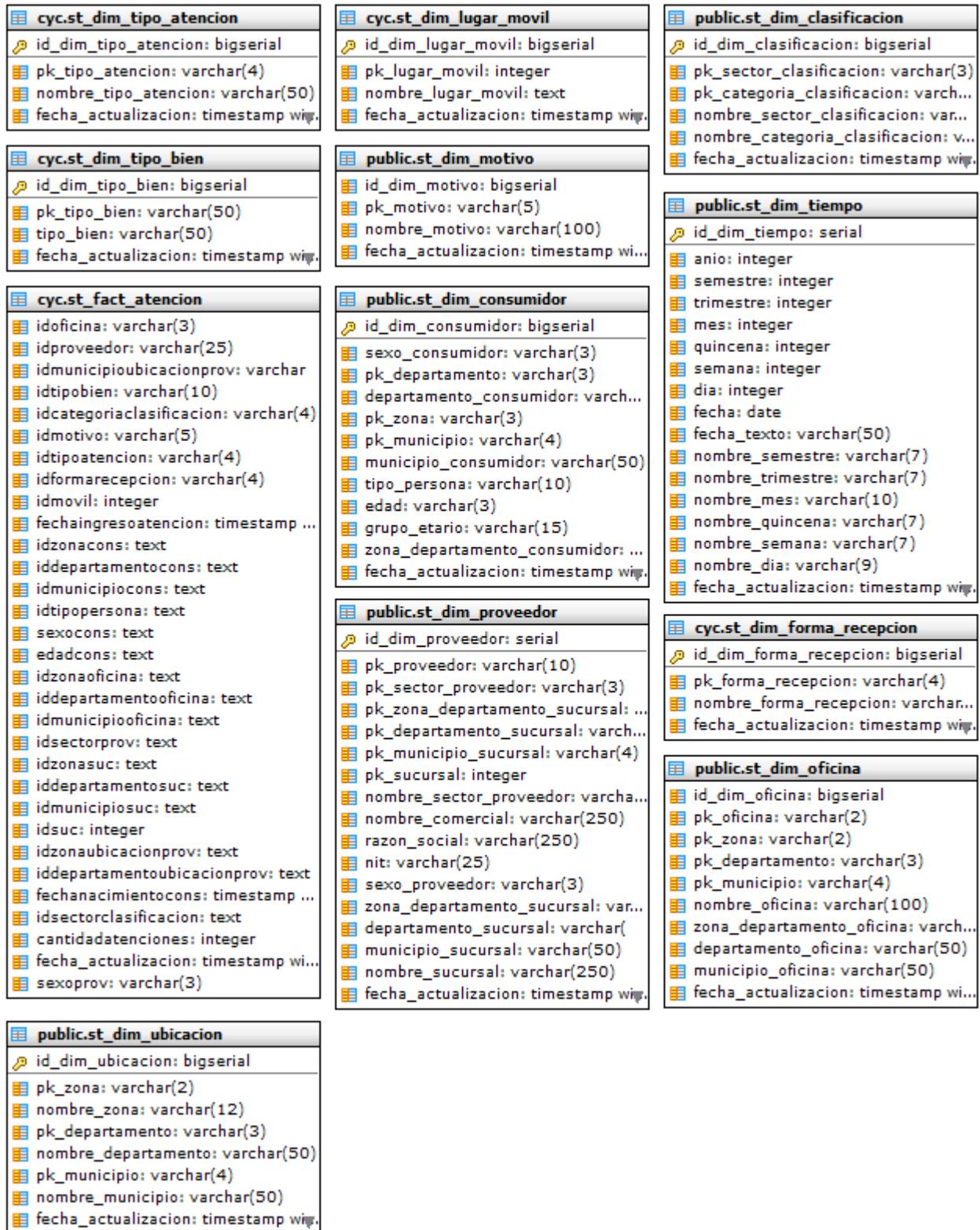


Figura 13.8: Diseño físico de base de datos Dirección de Ciudadanía y Consumo (Staging Area)

13.6 Diseño de Procesos ETL

Los procesos ETL se han dividido en dos grandes grupos, aquellos que mueven datos de las tablas fuentes hacia el Staging Area, los cuales llamaremos ‘Procesos ETL Staging’; así mismo se tienen los aquellos procesos ETL que mueven datos del Staging Area hacia el Data Mart los cuales llamaremos ‘Procesos ETL Data Mart. A continuación, se listan los diagramas de Procesos ETL de ambos grupos.

13.6.1 Procesos ETL Staging

En los siguientes apartados se listan los procesos ETL de Staging, seguido de un diagrama representativo para esta unidad y los diseños de los procesos restantes se colocarán en el CD que acompaña al documento en la carpeta [CD-EISI\Diseños de Procesos ETL\cyc \(Direccion de Ciudadania y Consumo\)](#)

Diagramas desarrollados

St_dim_lugar_movil

St_dim_oficina

St_dim_consumidor

St_fact_atencion

St_dim_tipo_atencion

St_dim_tipo_bien

St_dim_clasificacion

St_dim_forma_recepcion

St_dim_ubicacion

St_dim_motivo

13.6.2 Procesos ETL Data Mart

Los diagramas de los procesos ETL que mueven datos de las tablas del Staging Area hacia el Data Mart (Procesos ETL Data Mart) se describen en los subsiguientes apartados.

Diagramas desarrollados

Dw_dim_lugar_movil	Dw_dim_tipo_atencion	Dw_dim_forma_recepcion
Dw_dim_oficina	Dw_dim_tipo_bien	Dw_dim_ubicacion
Dw_dim_consumidor	Dw_dim_clasificacion	Dw_dim_motivo
Dw_fact_atencion		

A continuación, se ilustra uno de los procesos ETL. Se muestra el diseño del proceso de ETL de Figura 13.9: Proceso ETL dw_dim_tipo_bien

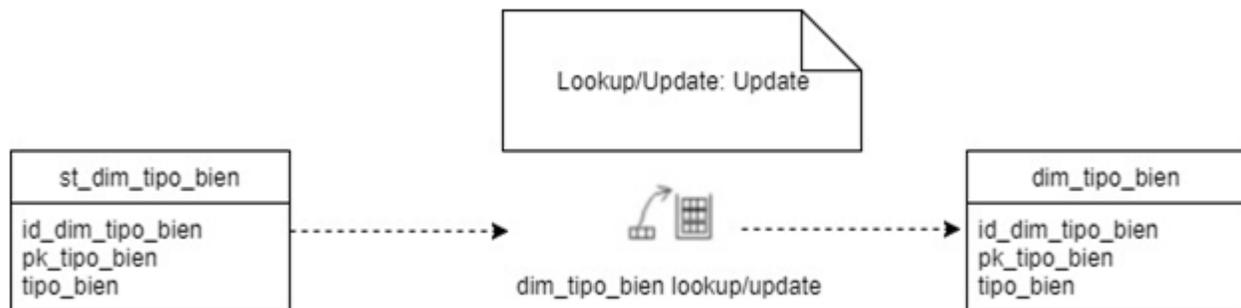


Figura 13.9: Proceso ETL dw_dim_tipo_bien

13.7 Pruebas

Al igual que en Sprint 1 las pruebas para el Data Mart de de Dirección de Ciudadanía y Consumo consisten en las pruebas de los procesos ETL y dashboards las cuales se muestran a continuación:

13.7.1 Pruebas de Procesos ETL

La realización de las pruebas consistió en ejecutar los Jobs almacenados en el servidor de Staging Área. Las transformaciones finales se muestran en la Tabla 13.2.

No	Transformación	Descripción
1	dw_incremental_fact_atencion.ktr	Transformación para cargas incrementales desde st_fact_atencion (staging area) hacia dim_fact_atencion (data warehouse).
2	dw_inicial_fact_atencion.ktr	Transformación para carga inicial desde st_fact_atencion (staging area) hacia dim_fact_atencion (data warehouse).
3	dw_unico_dim_clasificacion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_dim_clasificacion (staging area) hacia dim_clasificacion (data warehouse).
4	dw_unico_dim_consumidor.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_dim_consumidor (staging area) hacia dim_consumidor (data warehouse).
5	dw_unico_dim_forma_recepcion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_forma_recepcion (staging area) hacia dim_forma_recepcion (data warehouse).
6	dw_unico_dim_lugar_movil.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_lugar_movil (staging area) hacia dim_lugar_movil (data warehouse).

7	dw_unico_dim_motivo.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_motivo (staging area) hacia dim_motivo (data warehouse).
8	dw_unico_dim_oficina.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_oficina (staging area) hacia dim_oficina (data warehouse).
9	dw_unico_dim_tipo_atencion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_tipo_atencion (staging area) hacia dim_tipo_atencion (data warehouse).
10	dw_unico_dim_tipo_bien.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_tipo_bien (staging area) hacia dim_tipo_bien (data warehouse).
11	dw_unico_dim_ubicacion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_ubicacion (staging area) hacia dim_ubicacion (data warehouse).
12	st_incremental_fact_atencion.ktr	Transformación para carga inicial del data warehouse desde las tablas fuentes hacia st_fact_atencion (staging area).
13	st_inicial_fact_atencion.ktr	Transformación para cargas incrementales del data warehouse desde las tablas fuentes hacia st_fact_atencion (staging area).
14	st_unico_dim_clasificacion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_clasificacion (staging area).
15	st_unico_dim_consumidor.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_consumidor (staging area).
16	st_unico_dim_forma_recepcion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_forma_recepcion (staging area).
17	st_unico_dim_lugar_movil.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_lugar_movil (staging area).
18	st_unico_dim_motivo.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_motivo (staging area).
19	st_unico_dim_oficina.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_oficina (staging area).
20	st_unico_dim_tipo_atencion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_tipo_atencion (staging area).
21	st_unico_dim_tipo_bien.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_tipo_bien (staging area).
22	st_unico_dim_ubicacion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_ubicacion (staging area).

Tabla 13.2: Transformaciones para Data Mart Ciudadanía y Consumo

Los Jobs finales de Primer Nivel se muestran la también muestra el orden en que ejecutan las transformaciones de cada uno de los Jobs.

No	Job Primer Nivel	Transformación	
		No	Nombre Transformación
1	job_unico_dim_consumidor.kjb	1	st_unico_dim_consumidor.ktr
		2	dw_unico_dim_consumidor.ktr
2	job_inicial_fact_atencion.kjb	1	st_inicial_fact_atencion.ktr
		2	dw_inicial_fact_atencion.ktr
3	job_incremental_fact_atencion.kjb	1	st_incremental_fact_atencion.ktr
		2	dw_incremental_fact_atencion.ktr
4	job_unico_dim_ubicacion.kjb	1	st_unico_dim_ubicacion.ktr
		2	dw_unico_dim_ubicacion.ktr
5	job_unico_tipo_bien.kjb	1	st_inicial_tipo_bien.ktr
		2	dw_inicial_tipo_bien.ktr
6	job_unico_tipo_atencion.kjb	1	st_unico_tipo_atencion.ktr
		2	dw_unico_tipo_atencion.ktr
7	job_unico_dim_oficina.kjb	1	st_unico_dim_oficina.ktr
		2	dw_unico_dim_oficina.ktr
8	job_unico_dim_lugar_movil.kjb	1	st_unico_dim_lugar_movil.ktr
		2	dw_unico_dim_lugar_movil.ktr
9	job_unico_dim_motivo.kjb	1	st_unico_dim_lugar_movil.ktr
		2	dw_unico_dim_lugar_movil.ktr
10	job_unico_dim_forma_recepcion.kjb	1	st_unico_dim_forma_recepcion.ktr
		2	dw_unico_dim_forma_recepcion.ktr
11	job_unico_dim_clasificacion.kjb	1	st_unico_dim_lugar_clasificacion.ktr
		2	dw_unico_dim_lugar_clasificacion.ktr

Tabla 13.3: Jobs de Nivel 1 para Ciudadanía y Consumo

Los Jobs finales de Segundo Nivel se muestran en la tabla Tabla 13.4. La tabla también muestra el orden en que ejecutan las transformaciones cada uno de los Jobs

No	Jobs Segundo Nivel	Jobs primer nivel ejecutados	
		No	Nombre Job
1	job_inicial_cyc.kjb	1	job_unico_dim_ubicacion.kjb
		2	job_unico_dim_tipo_bien.kjb
		3	job_unico_dim_tipo_atencion.kjb
		4	job_unico_dim_oficina.kjb
		5	job_unico_fact_motivo.kjb
		6	job_unico_lugar_movil.kjb
		7	job_unico_forma_recepcion.kjb
		8	job_único_lugar_clasificacion.kjb
		9	job_unico_consumidor.kjb
		10	job_único_fact_atencion.kjb
	job_incremental_cyc.kjb	1	job_unico_dim_tipo_ubicacion.kjb
		2	job_unico_dim_tipo_bien.kjb
		3	job_unico_dim_tipo_atencion.kjb
		4	job_unico_dim_oficina.kjb
		5	job_unico_fact_motivo.kjb
		6	job_unico_fact_lugar_movil.kjb
		7	job_unico_fact_forma_recepcion.kjb

	8	job_unico_fact_clasificacion.kjb
	9	Job_incremental_fact_atencion.kjb

Tabla 13.4: Jobs de Nivel 2 Ciudadanía y Consumo

Las figuras siguientes muestran los Jobs de segundo nivel, los cuales son los Jobs principales que son ejecutados en el servidor. Figura 13.10: Job ETL para la carga inicial del Data Mart Ciudadanía y Consumo. La figura muestra la secuencia del job que realiza la carga inicial, mientras que la Figura 13.11 muestra la secuencia del job que servirá para realizar las cargas incrementales.



Figura 13.10: Job ETL para la carga inicial del Data Mart Ciudadanía y Consumo



Figura 13.11: Job ETL para la carga incremental del Data Mart Ciudadanía y Consumo

13.7.1.1 Ejecución de Procesos ETL iniciales

El comando para ejecutar el proceso ETL se muestra a continuación:

```
#/root/data-integration/kitchen.sh -file=/root/jobs/cyc/job_inicial_cyc.kjb -level=Debug
```

El tiempo estimado para la ejecución de los procesos ETL es el siguiente:

Cantidad de transformación	Tiempo	Registros
20	18:47:22.664	997,611

La finalización del proceso ETL se notifica mediante correo electrónico como se muestra en la Figura 13.2.

<input type="checkbox"/>	☆	DATAWAREHOUSE	INCREMENTAL ETL FINISHED FOR sprint2 (Dirección de Ciudadanía y Consumo) - EL PROCES...
<input type="checkbox"/>	☆	DATAWAREHOUSE	INITIAL ETL FINISHED FOR fact_atencion - EL PROCESO ETL INICIAL PARA fact_atencion HA FL...
<input type="checkbox"/>	☆	DATAWAREHOUSE	INITIAL ETL FINISHED FOR dim_consumidor - EL PROCESO ETL INICIAL PARA dim_consumidor...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_clasificacion - EL PROCESO ETL FULL REFRESH PARA ...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_forma_recepcion - EL PROCESO ETL FULL REFRESH PA...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_lugar_movil - EL PROCESO ETL FULL REFRESH PARA di...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_motivo - EL PROCESO ETL FULL REFRESH PARA dim_m...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_oficina - EL PROCESO ETL FULL REFRESH PARA dim_o...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_tipo_atencion - EL PROCESO ETL FULL REFRESH PARA ...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_tipo_bien - EL PROCESO ETL FULL REFRESH PARA dim...
<input type="checkbox"/>	☆	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_ubicacion - EL PROCESO ETL FULL REFRESH PARA dim...

Figura 13.12: Notificación de finalización de los procesos ETL (Carga Inicial)

13.7.2 Pruebas Cubos OLAP

Dado que para Dirección de Ciudadanía y Consumo se identificaron un solo hecho a medir, se realizó un único cubo para dar respuesta a lo solicitado por esta unidad. A continuación, se muestra el cubo realizado para el Data Mart de Dirección de Ciudadanía y Consumo.

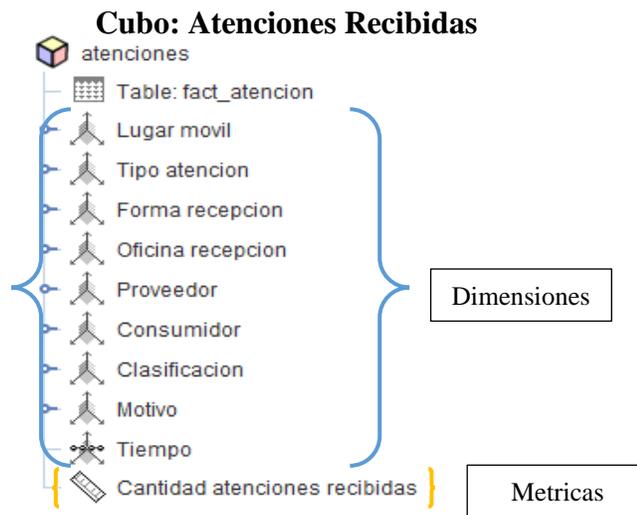


Tabla 13.5: Cubos para Dirección de Ciudadanía y Consumo

Para este cubo se ha generado y publicado su correspondiente archivo XML en el Servidor de Pentaho Server BI para su posterior visualización.

Con el uso del Plugin JPivot para Pentaho Server se logra visualizar las dimensiones que están ubicadas en las columnas, y bajo ellas las diferentes opciones de análisis para una métrica, la información según en la Figura 13.13.

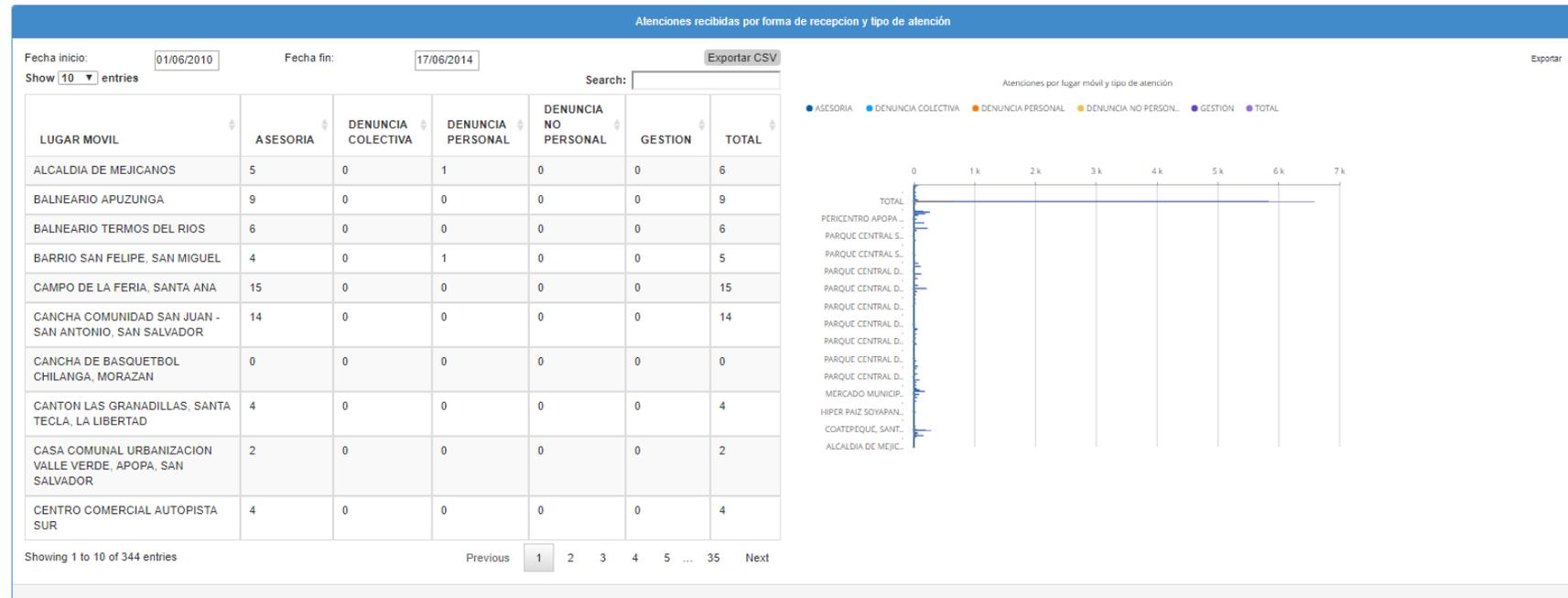
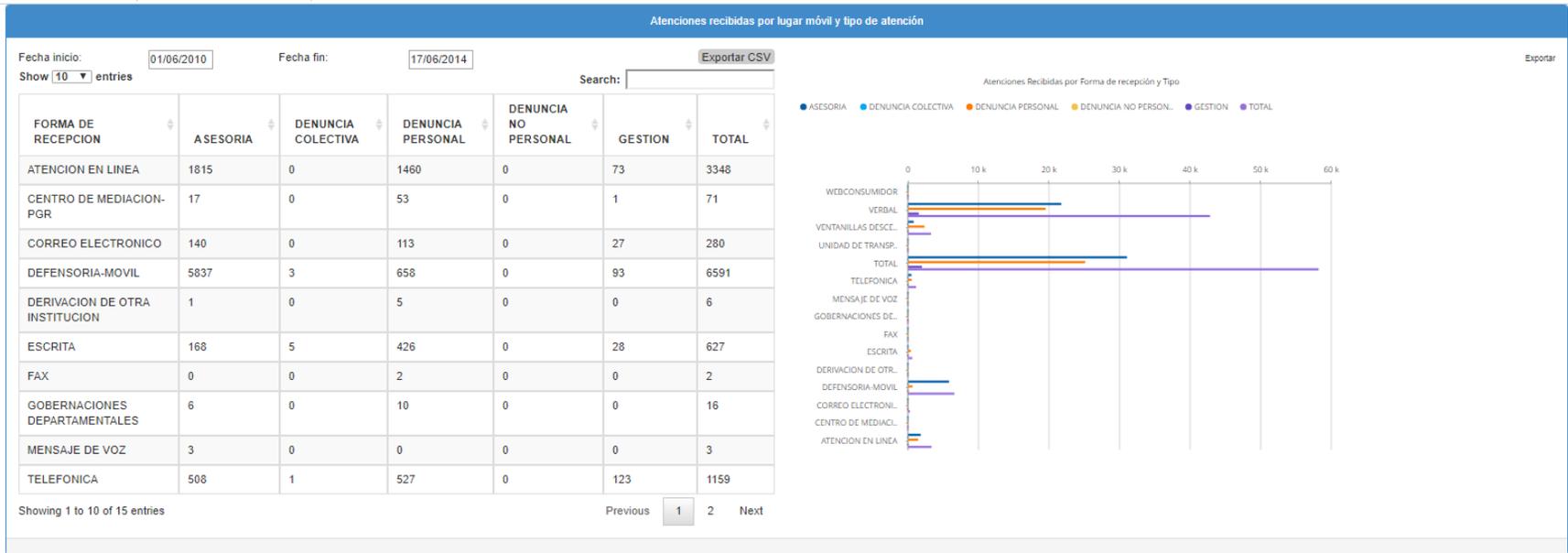
Tipo atencion	forma recepcion	Oficina recepcion	Nombre sector	Consumidor	Clasificacion	Motivo	Tiempo	Cantidad atenciones recibidas
<input type="checkbox"/> Todas las atenciones	<input type="checkbox"/> Todas las formas de recepcion	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	24.114
Asesoría	<input type="checkbox"/> Todas las formas de recepcion	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	8.258
Denuncia Colectiva	<input type="checkbox"/> Todas las formas de recepcion	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	33
Denuncia Personal	<input type="checkbox"/> Todas las formas de recepcion	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	10.780
	Atención en línea	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	158
	Centro de Mediación-PGR	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	52
	Correo Electronico	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	1.570
	Defensoria-Movil	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	274
		<input type="checkbox"/> Primer semestre						15
		<input type="checkbox"/> Segundo semestre						259
	Escrita	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	342
		<input type="checkbox"/> Primer semestre						140
		<input type="checkbox"/> Segundo semestre						202
Fax	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	1	
Telefónica	<input type="checkbox"/> Todas las oficinas	<input type="checkbox"/> Todos los sectores	<input type="checkbox"/> Todos los consumidores	<input type="checkbox"/> Todas las clasificaciones	<input type="checkbox"/> Todos los motivos	<input type="checkbox"/> 2009	172	
	<input type="checkbox"/> Primer semestre						57	

Figura 13.13: Visualización del Cubo Atenciones Recibidas por la Defensoria del Consumidor (Dirección de Ciudadanía y Consumo)

13.7.3 Pruebas Dashboard

El Dashboard realizado para Dirección de Ciudadanía y Consumo se muestra en la siguiente página.

13.7.3.1 Dashboard Dirección de Ciudadanía y Consumo



14 SPRINT 3

14.1 Sprint Backlog

A continuación, se presenta el Sprint Backlog, que refleja las historias de usuarios, su estimación y asignación para este Sprint.

#	Historia	Punto Estimado	Asignado
1.	Refinamiento del requerimiento de información	20	Jorge – Samuel – Ivis
2.	Poblar la base de datos de desarrollo SARA	13	Samuel
3.	Diseño del modelo multidimensional	8	Jorge – Samuel – Ivis
4.	Diseño BD del Stagin Area	3	Jorge – Samuel – Ivis
5.	Diseño del proceso ETL	8	Jorge – Samuel – Ivis – Esaú
6.	Diseño de vistas	3	Ivis – Esaú
7.	Creación de Script SQL modelo multidimensional	13	Jorge – Samuel
8.	Creación de Script SQL del Stagin Area	13	Jorge – Samuel
9.	Desarrollo de los procesos ETL	100	Jorge – Samuel – Ivis
10.	Desarrollo de cubos OLAP	40	Ivis – Esaú
11.	Desarrollo de vistas	13	Ivis – Esaú
12.	Implementación e integración de componentes	40	Todos
13.	Pruebas	20	Todos
14.	Documentación	40	Todos

Tabla 14.1: Sprint Backlog Sprint 3

14.2 Descripción de Historias de Usuario

A continuación, se presenta la descripción de las historias de usuario detallando: el responsable, el sprint que pertenece la historia, una breve descripción, sus criterios de validación y los puntos de estimación. Previamente a las historias de usuario se presenta la historia épica del Sprint 3.

<i>Historia Épica #4</i>	Data Mart Tribunal Sancionador.
Responsable	Jorge Iraheta/Esaú Serpas/Ivis Navarro/Samuel Núñez Sprint: 3
Descripción	<p>Como Tribunal Sancionador (jefatura) Necesito: Un mecanismo para analizar la información (Cantidad de casos ingresados, Cantida de casos finalizados con resolución y Cantidad de casos con sanciones y el monto de multas impuestas) proveniente de Centro de solución de controversias y de presidencia (Dirección de Vigilancia de mercado). Para: visibilizar el trabajo que se desempeña como unidad, optimizando la carga de trabajo de la unidad.</p>
Criterios de aceptación	<p>- La cantidad de denuncias ingresadas deben poderse analizar desde las siguientes perspectivas:</p> <ul style="list-style-type: none"> • Tiempo • Proveedor • Consumidor • Empleado • Oficina Recepción • Tipo Establecimiento • Procedencia • Motivo • Clasificación • Ubicación <p>- La cantidad de casos con resouciones finales, deben poderse analizar desde las siguientes perspectivas:</p>

	<ul style="list-style-type: none"> • Tiempo • Proveedor • Consumidor • Empleado • Infracción • Oficina recepción • Tipo de resolución • Tipo de Establecimiento • Procedencia • Motivo • Clasificación • Ubicación <p>- La cantidad de sanciones impuestas y el monto recuperado de estas sanciones por la unidad, deben poderse analizar desde las siguientes perspectivas:</p> <ul style="list-style-type: none"> • Tiempo • Proveedor • Consumidor • Empleado • Oficina Recepción • Infracción • Tipo de Establecimiento • Motivo • Clasificación • Ubicación
Puntos estimados	334

<i>Historia de usuario #1</i>	Refinamiento del requerimiento de información
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Nuñez Sprint: 3
Descripción	<p>Como Equipo Scrum necesito el refinamiento del requerimiento de información de Tribunal Sancionador: Realizar reuniones con los interesados y especificar los elementos de información requeridos para Tribunal Sancionador.</p> <ul style="list-style-type: none"> • Llenar la herramienta de recolección. • Revisar y desglosar los requerimientos solicitados • Firmar documento por el usuario como aceptación de los de los requerimientos. <p>Para tener claridad de la información que desean almacenar, visualizar y analizar mediante el Data Mart o modelo Multidimensional para ésta unidad.</p>
Criterios de aceptación	<ul style="list-style-type: none"> - Tener llena la plantilla de refinamiento de requerimientos de información. - Incluir únicamente requerimientos de información con naturaleza histórica y no información transaccional. - Tener un documento firmado de aceptación de requerimientos. - Identificar (mediante un listado) las tablas fuentes que servirán para cumplir con los requerimientos de información solicitados.
Puntos estimados	20

<i>Historia de usuario #2</i>	Poblar base de datos de desarrollo SARA
<i>Responsable</i>	Samuel Núñez Sprint: 3
<i>Descripción</i>	Como Product Owner necesito: migrar los datos necesarios según las tablas fuentes de datos que se identifiquen durante el proceso de toma de requerimientos. Para poseer datos de prueba para el desarrollo del Data Mart.
<i>Criterios de aceptación</i>	- Se deben poseer Scripts de INSERT para las tablas Fuente. - Todas las tablas fuentes identificadas en el refinamiento de requerimientos deben poseer datos. - Debe existir al menos una prueba de conexión y acceso remoto exitoso a los datos. - Datos cargados en la base de datos
<i>Puntos estimados</i>	13

<i>Historia de usuario #3</i>	Diseño del modelo multidimensional
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 3
<i>Descripción</i>	Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de Tribunal Sancionador. Para tener la base (diagrama) para el desarrollo de Scripts de Base de Datos SQL para Data Mart.
<i>Criterios de aceptación</i>	-Esquema del modelo multidimensional de Tribunal Sancionador. -El esquema debe cumplir las reglas de un modelo de tipo Estrella. -El esquema debe poseer las relaciones necesarias para la búsqueda según los criterios de búsqueda definidos por la Unidad de Tribunal Sancionador. -El esquema debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
<i>Puntos estimados</i>	8

<i>Historia de usuario #4</i>	Diseño BD del Staging Area.
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 3
<i>Descripción</i>	Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de la Unidad de Tribunal Sancionador. Para tener el insumo para la creación de Scripts de Base de Datos SQL para Staging Area.
<i>Criterios de aceptación</i>	-Deberá existir una o más tablas por cada dimensión. -Deberá existir una o más tablas de hecho. -El diseño debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
<i>Puntos estimados</i>	3

<i>Historia de usuario #5</i>	Diseño del proceso ETL
<i>Responsable</i>	Jorge Iraheta/Esau Serpas/Ivis Navarro/Samuel Núñez Sprint: 3
<i>Descripción</i>	Como Equipo Scrum necesito: Diseñar los Proceso ETL <ul style="list-style-type: none"> • Extracción • Transformación • Carga Para cumplir con el Diseño del proceso ETL y tenerlo como insumo a la hora de desarrollar dichos procesos ETL.
<i>Criterios de aceptación</i>	-Se deberá utilizar nomenclatura UML para los pasos en cada proceso ETL. -Deberá incluir todas las tablas dimensiones y tablas de hechos. -Los nombres de las tablas o etiquetas deben poderse visualizar fácilmente.
<i>Puntos estimados</i>	8

<i>Historia de usuario #6</i>	Diseño de vistas
<i>Responsable</i>	Ivis Navarro/Esau Serpas Sprint: 3
<i>Descripción</i>	Como Administrador de Data Warehouse necesito: Realizar el diseño de vistas necesarias para visualizar la información del Data Mart de Tribunal Sancionador. Para Diseñar y digitalizar la vista de presentación de datos.
<i>Criterios de aceptación</i>	-Deberá existir un diseño de vista o dashboard para cada hecho a medir. -Deberá explicarse cada sección del dashboard o vista. -Etiquetas y textos deben ser legibles a simple vista.
<i>Puntos estimados</i>	3

<i>Historia de usuario #7</i>	Creación de Script SQL Modelo Multidimensional
<i>Responsable</i>	Samuel Núñez/Jorge Iraheta Sprint: 3
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será almacenada la información.
<i>Criterios de aceptación</i>	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Modelo Multidimensional. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
<i>Puntos estimados</i>	13

<i>Historia de usuario #8</i>	Creación de Script SQL del Staging Area
<i>Responsable</i>	Samuel Núñez/Jorge Iraheta Sprint: 3
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será procesada, calculada y/o limpiada la información.
<i>Criterios de aceptación</i>	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Staging Area. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
<i>Puntos estimados</i>	13

<i>Historia de usuario #9</i>	Desarrollo de los procesos ETL.
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 3
Descripción	Como Equipo Scrum necesito: Desarrollar los procesos ETL correspondientes Para obtener, trasladar y transformar los datos que serán almacenados en el Data Mart de Tribunal Sancionador.
Criterios de aceptación	<ul style="list-style-type: none"> -Los nombres de los procesos ETL deben poseer la nomenclatura definida por el equipo. -Los procesos ETL deben poseer la documentación necesaria en forma de comentarios dentro de la herramienta PDI. -Las Transformaciones ETL deben haber sido probados en desarrollo (Equipo Local) usando la base de datos pública de pruebas. -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de pruebas. -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de final. -Las transformaciones correspondientes a una dimensión o hecho deben ser incorporados un Job de PDI. -Deberá crearse un Job que incorpore los otros Jobs correspondientes a cada dimensión y hecho. -Los Jobs de PDI deben haber sido probados con las bases de datos de prueba y base de datos final. -Las bitácoras de los procesos ETL deben almacenarse en una base de datos.
Puntos estimados	100

<i>Historia de usuario #10</i>	Desarrollo de Cubos OLAP
Responsable	Esaú Serpas/Ivis Navarro Sprint: 3
Descripción	Como Equipo Scrum necesito: una investigación sobre la herramienta para el desarrollo de cubos OLAP e implementar dicha herramienta para la creación de cubos OLAP. Para implementar el cubo correspondiente a tribunal sancionador.
Criterios de aceptación	<ul style="list-style-type: none"> -El cubo deberá poderse visualizar correctamente y sin errores. -Deberá existir como mínimo un cubo OLAP por Hecho. -Deberán poderse realizar las operaciones como mínimo de: Roll Up y Drill Down.
Puntos estimados	40

<i>Historia de usuario #11</i>	Desarrollo de vistas
Responsable	Esaú Serpas/Ivis Navarro Sprint: 3
Descripción	Como Equipo Scrum necesito: Desarrollar vistas o dashboards según el diseño previamente elaborado. Para Presentar información relevante hacia la unidad para la toma de decisiones.
Criterios de aceptación	<ul style="list-style-type: none"> -Vista o dashboard funcionando sin errores técnicos. -Debe coincidir con el diseño realizado previamente. -Debe existir al menos un dashboar por tabla de hechos.
Puntos estimados	13

<i>Historia de usuario #12</i>	Implementación e integración de componentes
Responsable	Ivis Navarro/Esau Serpas/Samuel Núñez/Jorge Iraheta Sprint: 3
Descripción	<p>Como Equipo Scrum necesito: La integración de todos los elementos del modelo multidimensional</p> <ul style="list-style-type: none"> - ETL - Fuentes - Stagin Área - BD multidimensional - Pentaho Server (Vistas) <p>Para el desarrollo del modelo multidimensional de forma integral y probar el funcionamiento correctamente de la totalidad del Data Mart.</p>
Criterios de aceptación	<p>- Se deberán modificar los procesos ETL para que utilicen las bases de datos finales (no las de pruebas).</p> <p>- Se deberán realizar pruebas en la visualización de los dashboards.</p> <p>- Todo el entorno (desde los procesos ETLs, hasta los dashboards o vistas) deberán tener el visto bueno de todo el equipo.</p>
Puntos estimados	40

<i>Historia de usuario #13</i>	Pruebas
Responsable	Ivis Navarro/Esau Serpas/Samuel Núñez/Jorge Iraheta Sprint: 3
Descripción	Como Equipo Scrum necesito: Realizar pruebas al modelo multidimensional terminado. Para verificar que funciona de la manera que se espera.
Criterios de aceptación	Pruebas exitosas en pruebas de aceptación del usuario
Puntos estimados	20

<i>Historia de usuario #14</i>	Documentación
Responsable	Samuel Nuñez/Jorge Iraheta/Ivis Navarro/Esau Serpas Sprint: 3
Descripción	Como Equipo Scrum necesito: Documentación completa con historias, nomenclatura y diagramas empleados. Para verificar el óptimo desarrollo de los cubos (Tribunal Sancionador).
Criterios de aceptación	Documentos completos de análisis y diseño para modelo multidimensionales
Puntos estimados	40

14.3 Refinamiento del Requerimiento de Información

Para la realización del refinamiento de los requerimientos de información se partió con la investigación del proceso general que sigue el Tribunal Sancionador dentro la Defensoría del Consumidor. Además de establecer el proceso general de la unidad se hizo uso de los paquetes de información para refinar el requerimiento de información.

14.3.1 Proceso BPMN de la Unidad

A continuación, presentamos el diagrama BPNM que engloba las actividades de la unidad de Tribunal Sancionador.

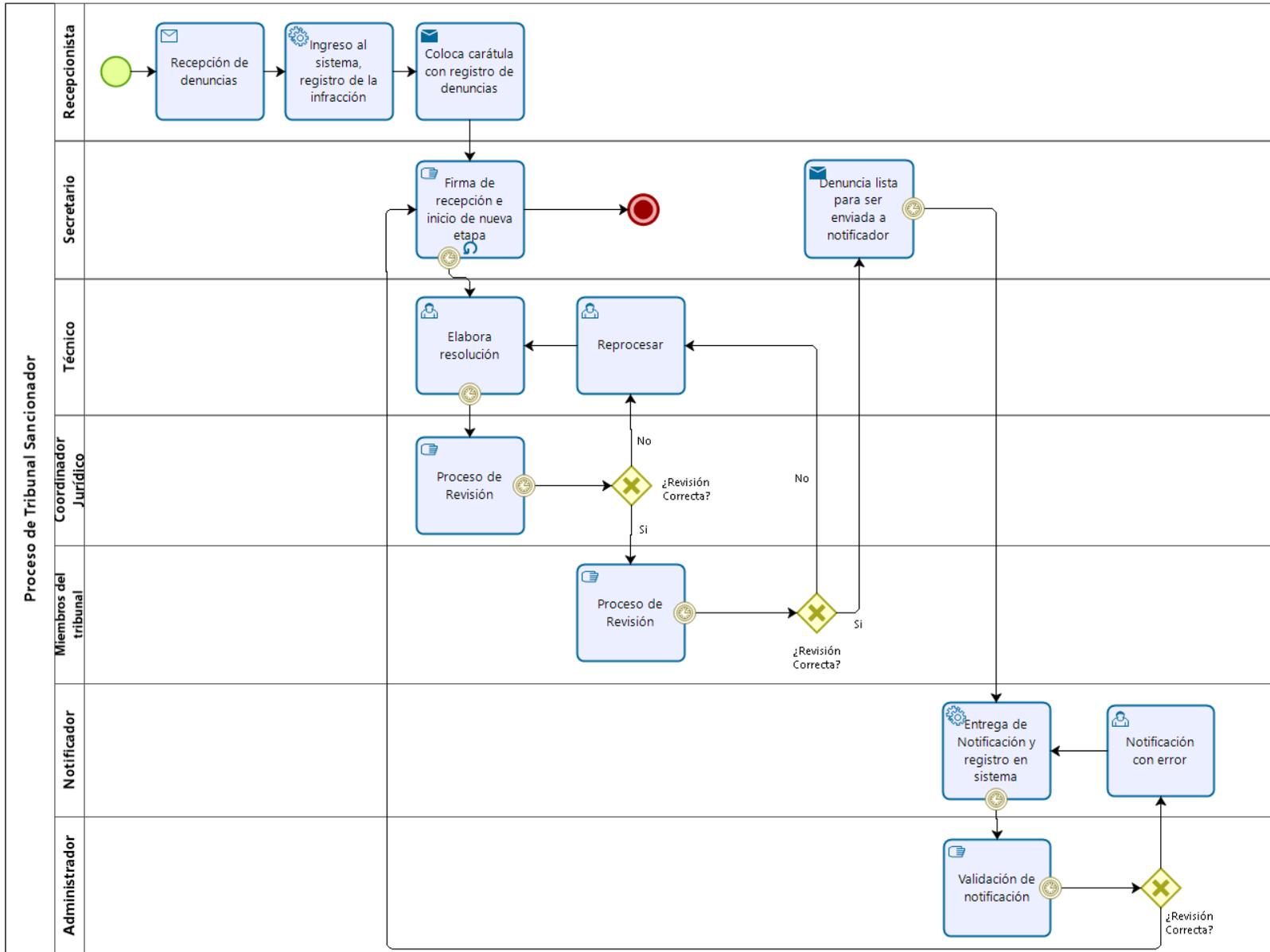


Figura 14.1: Proceso BPMN Tribunal Sancionador

14.3.2 Paquetes de información

A continuación, se presenta los paquetes de información para el Tribunal Sancionador.

Para el Tribunal Sancionador de la Defensoría del Consumidor se identificaron tres paquetes de información para los siguientes temas:

- Tabla 14.2: Denuncias ingresadas en Tribunal Sancionador
- Tabla 14.3: Resoluciones finales en Tribunal Sancionador.
- Tabla 14.4: Sanciones impuestas por Tribunal Sancionador

Tema:		Denuncias ingresadas en Tribunal Sancionador								
JERARQUIAS	Tiempo	Proveedor	Consumidor	Empleado	Oficina Recepción	Tipo Establecimiento	Procedencia	Motivo	Clasificación	Ubicación
	Año	Nombre Sector	Sexo	Zona Departamento	Zona	Nombre del tipo Establecimiento	Nombre de la procedencia	Nombre Motivo	nombre sector clasificación	Zona
	Semestre	Nombre Comercial	Departamento	Departamento de la Oficina	Departamento				nombre categoría clasificación	Departamento
	Trimestre	Departamento Sucursal	Municipio	Municipio de la Oficina	Municipio					Municipio
	Mes	Municipio Sucursal		Oficina del Empleado	Oficina					
	Quincena	Nombre Sucursal		Cargo del Empleado						
	Semana			Nombre Empleado						
	Día									
	Hechos Medidos:	Cantidad de casos ingresadas en Tribunal sancionador								

Tabla 14.2: Denuncias ingresadas en Tribunal Sancionador

Tema:		Resoluciones finales en Tribunal Sancionador									
JERARQUIAS	Tiempo	Proveedor	Consumidor	Empleado	Oficina Recepción	Tipo Resolución	Tipo Establecimiento	Procedencia	Motivo	clasificación	ubicacion
	Año	Nombre Sector	Sexo	Zona Departamento	Zona	nombre de tipo de resolución	Nombre del tipo Establecimiento	Nombre de la procedencia	Nombre Motivo	nombre sector clasificación	Zona
	Semestre	Nombre Comercial	Departamento	Departamento de la Oficina	Departamento					nombre categoría clasificación	Departamento
	Trimestre	Departamento Sucursal	Municipio	Municipio de la Oficina	Municipio						Municipio
	Mes	Municipio Sucursal		Oficina del Empleado	Oficina						
	Quincena	Nombre Sucursal		Cargo del Empleado							
	Semana			Nombre Empleado							
	Día										
	Hechos Medidos:	Cantidad de resoluciones finales									

Tabla 14.3: Resoluciones finales en Tribunal Sancionador

Tema:		Sanciones impuestas por la Unidad Dirección Jurídica.								
JERARQUIAS	Tiempo	Proveedor	Consumidor	Empleado	Oficina Recepción	Infracción	Tipo Establecimiento	Motivo	clasificación	ubicacion
	Año	Nombre Sector	Sexo	Zona Departamento	Zona	Nombre ley	Nombre del tipo Establecimiento	Nombre Motivo	nombre sector clasificación	Zona
	Semestre	Nombre Comercial	Departamento	Departamento de la Oficina	Departamento	Artículo			nombre categoría clasificación	Departamento
	Trimestre	Departamento Sucursal	Municipio	Municipio de la Oficina	Municipio	Literal				Municipio
	Mes	Municipio Sucursal		Oficina del Empleado	Oficina	Detalle				
	Quincena	Nombre Sucursal		Cargo del Empleado		Tipo Infracción				
	Semana			Nombre Empleado						
	Día									
	Hechos Medidos:	<ul style="list-style-type: none"> • Cantidad de sanciones impuestas por la unidad de Tribunal Sancionador, • El monto recuperado por estas sanciones 								

Tabla 14.4: Sanciones impuestas por Tribunal Sancionador

14.4 Especificación de Requerimientos

14.4.1 Caso de Uso: análisis de denuncias ingresadas en Tribunal Sancionador

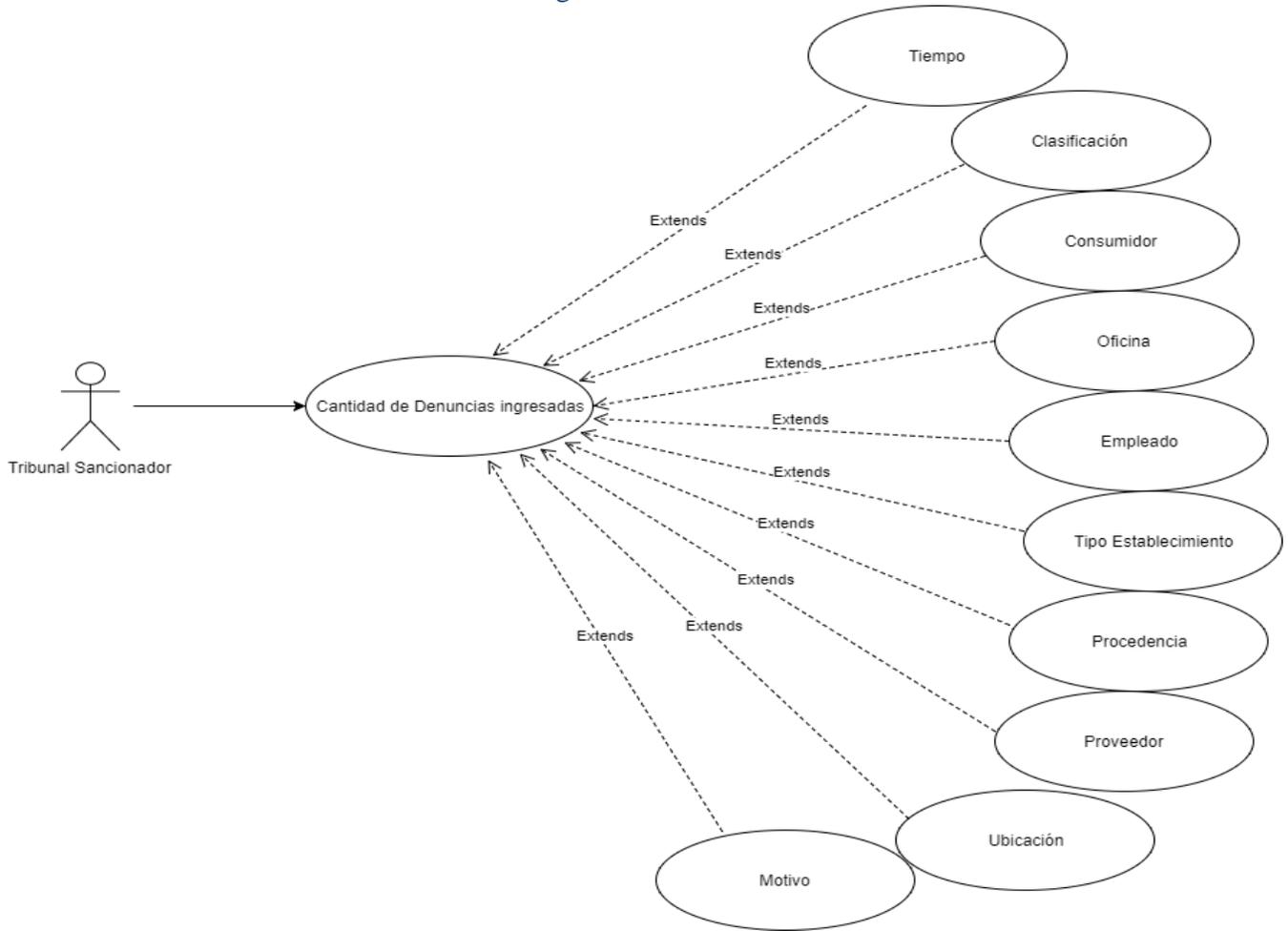


Figura 14.2: Caso de uso de denuncias ingresados al Tribunal Sancionador

14.4.2 Caso de Uso: análisis de resoluciones emitidas en Tribunal Sancionador

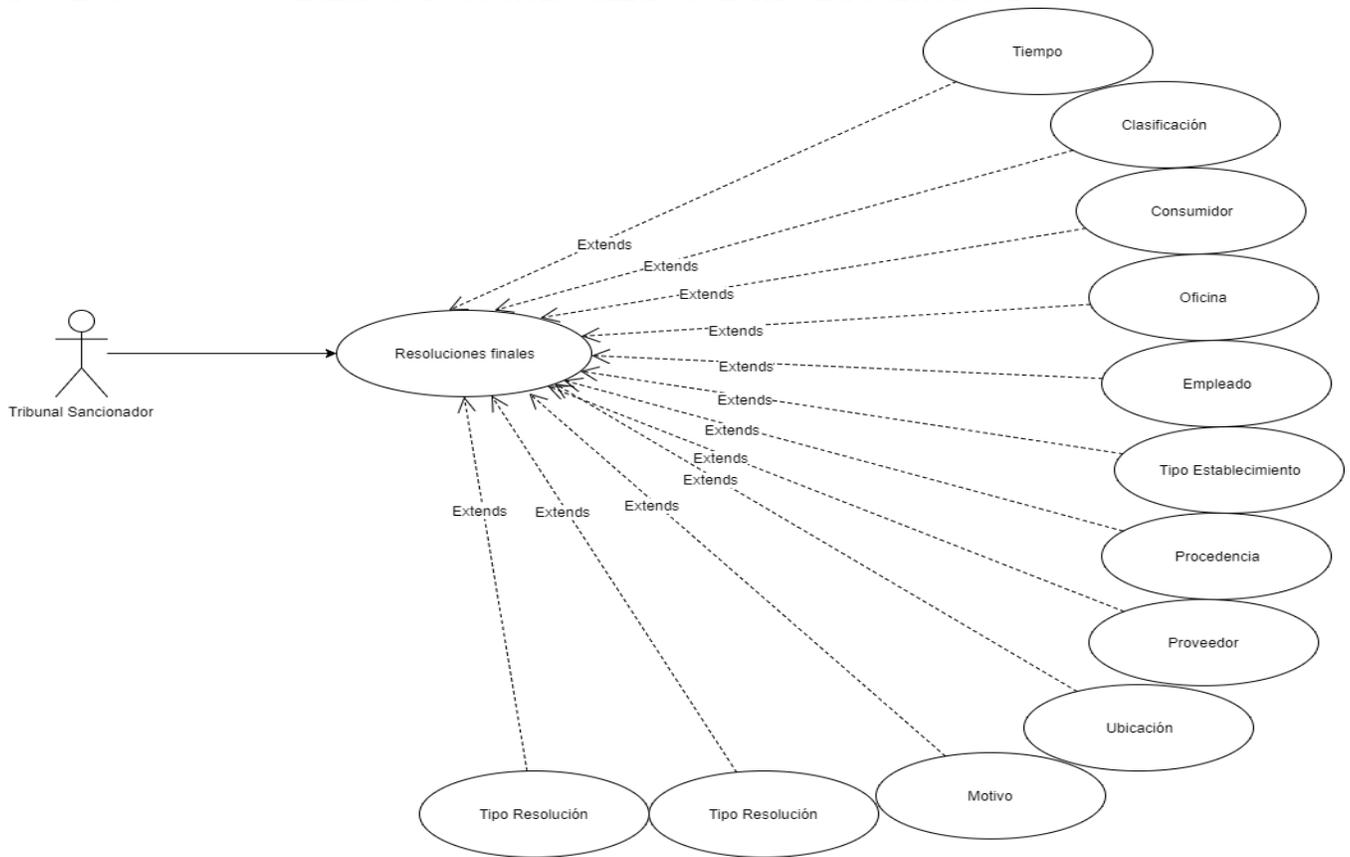


Figura 14.3: Caso de uso de resoluciones emitidas en Tribunal Sancionador

14.4.3 Caso de Uso: análisis de Sanciones impuestas en Tribunal Sancionador

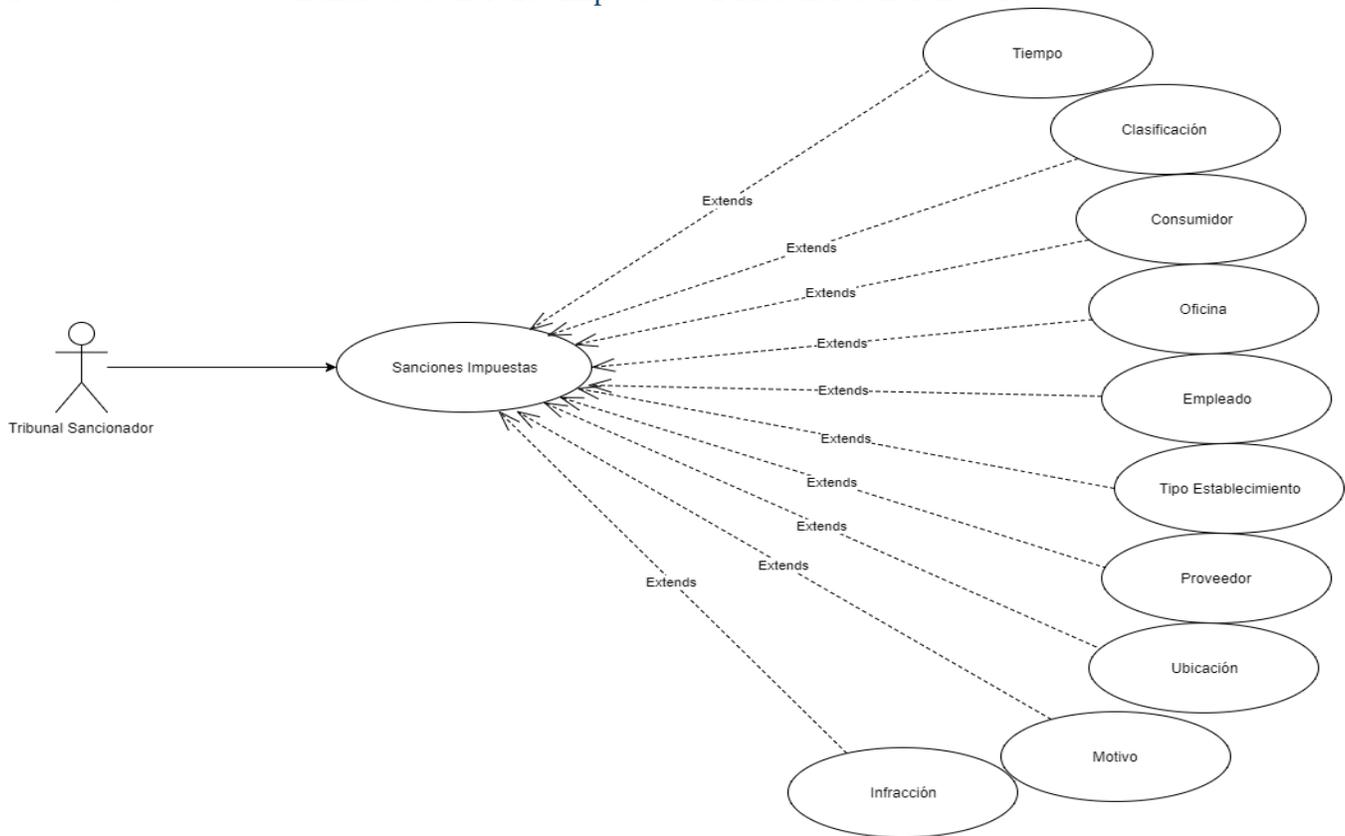


Figura 14.4: Caso de uso de Sanciones en la unidad de Tribunal Sancionador

14.5 Diseño Multidimensional

14.5.1 Diseño Conceptual del Data Mart

14.5.1.1 Nivel 1: Definición del modelo

Nivel 1:

Unidad: Tribunal Sancionador

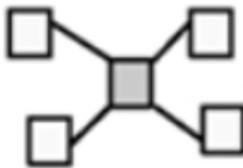
Especialización: Procesar las denuncias para dar una resolución según proceda.

Hechos principales: Cantidad de casos ingresados

Cantidad de resoluciones finales

Cantidad de sanciones impuestas

Monto recuperado



Tribunal Sancionador

Figura 14.5: Diagrama Conceptual Data Mart Tribunal Sancionador

14.5.1.2 Nivel 2: Definición del esquema estrella

Paquete de Hecho: Cantidad de casos ingresados

Paquetes Dimensionales:

1. Tiempo
2. Empleado
3. Ubicacion
4. Tipo establecimiento
5. Procedencia
6. Oficina recepción
7. Proveedor
8. Consumidor
9. Clasificación
10. Motivo

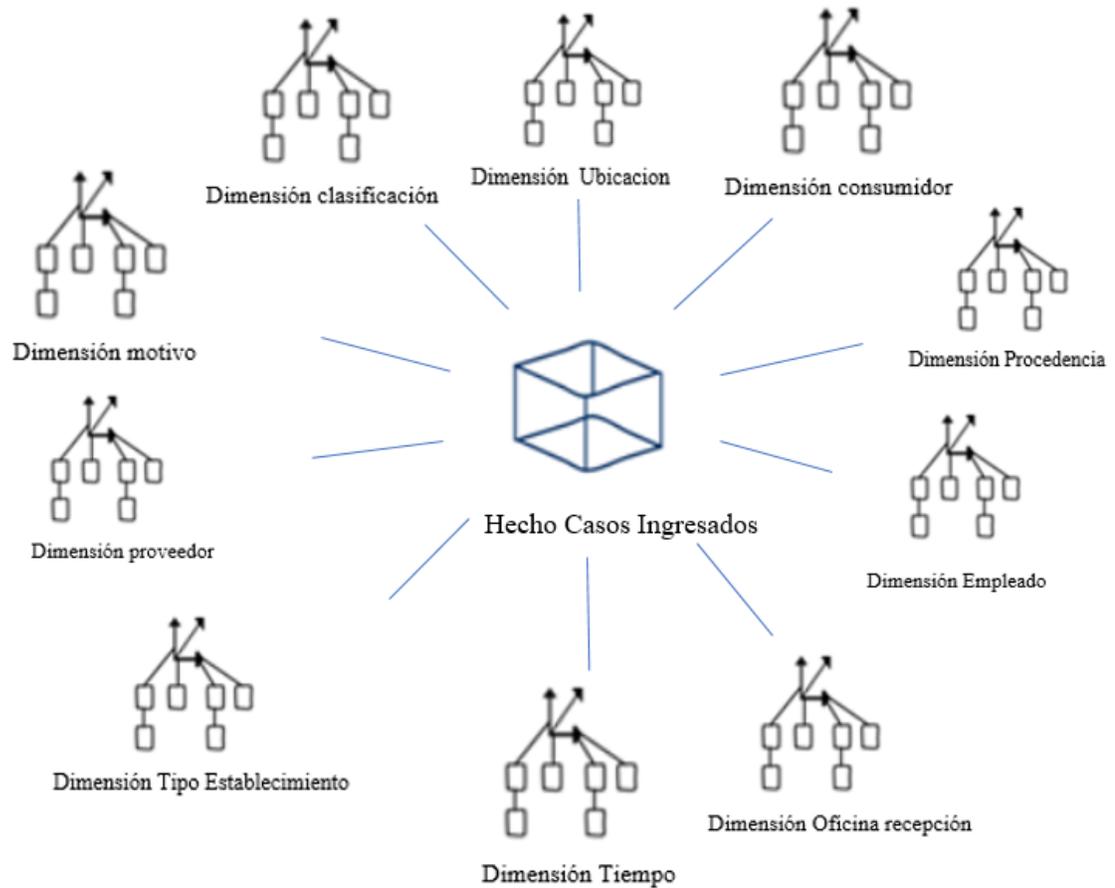


Figura 14.6: Nivel 2 Diagrama Conceptual Data Mart (Fact Casos ingresados)

Paquete de Hecho: Cantidad de resoluciones finales

Paquetes Dimensionales:

1. Tiempo
2. Empleado
3. Ubicación
4. Tipo establecimiento
5. Procedencia
6. Tipo resolución
7. Oficina recepción
8. Proveedor
9. Consumidor
10. Clasificación
11. Motivo
12. Infracción

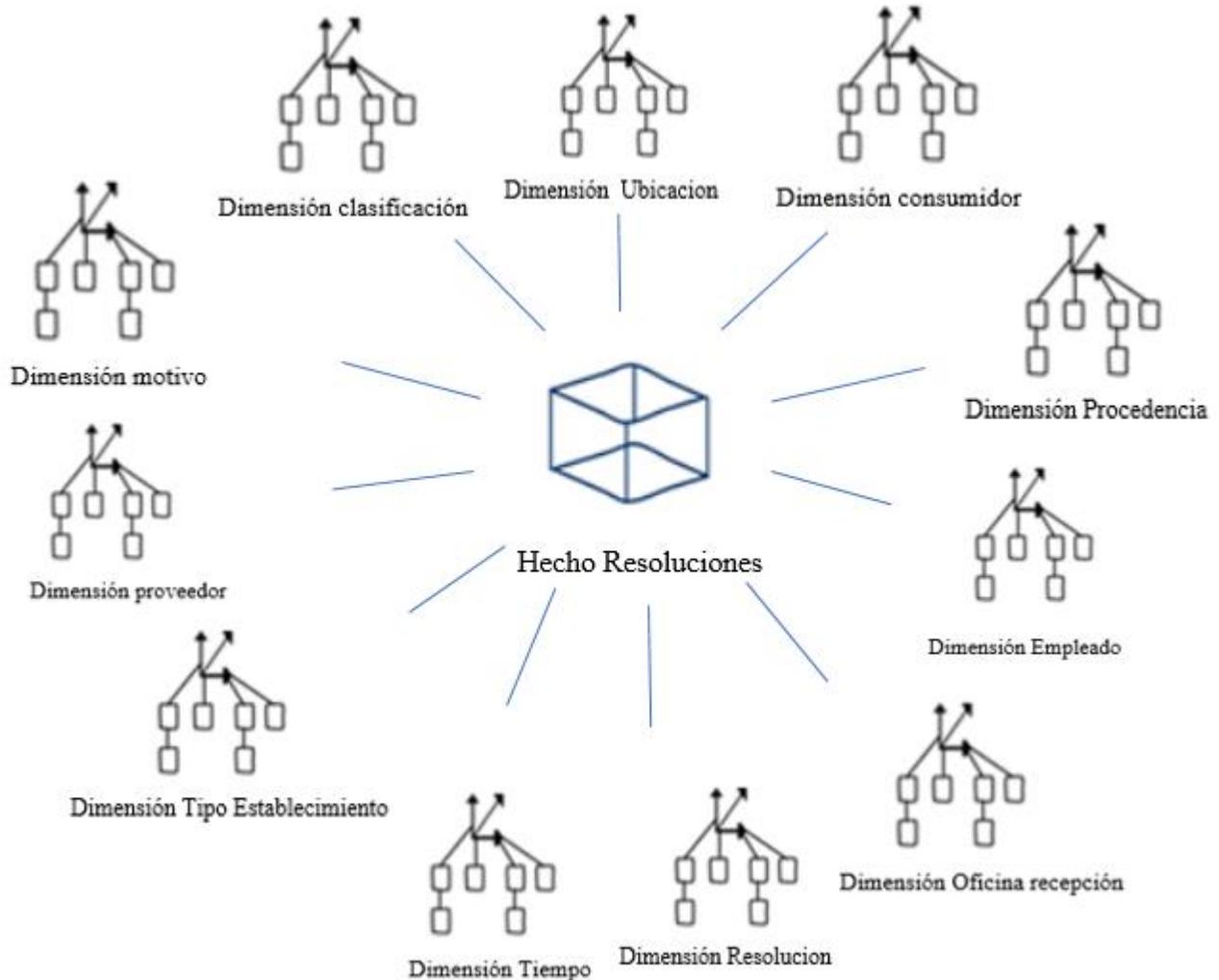


Figura 14.7: Nivel 2 Diagrama Conceptual Data Mart (Fact resoluciones)

Paquete de Hecho: Cantidad de sanciones impuestas

Paquetes Dimensionales:

1. Tiempo
2. Empleado
3. Ubicación
4. Tipo establecimiento
5. Oficina recepción
6. Proveedor
7. Consumidor
8. Clasificación
9. Motivo
10. Infracción

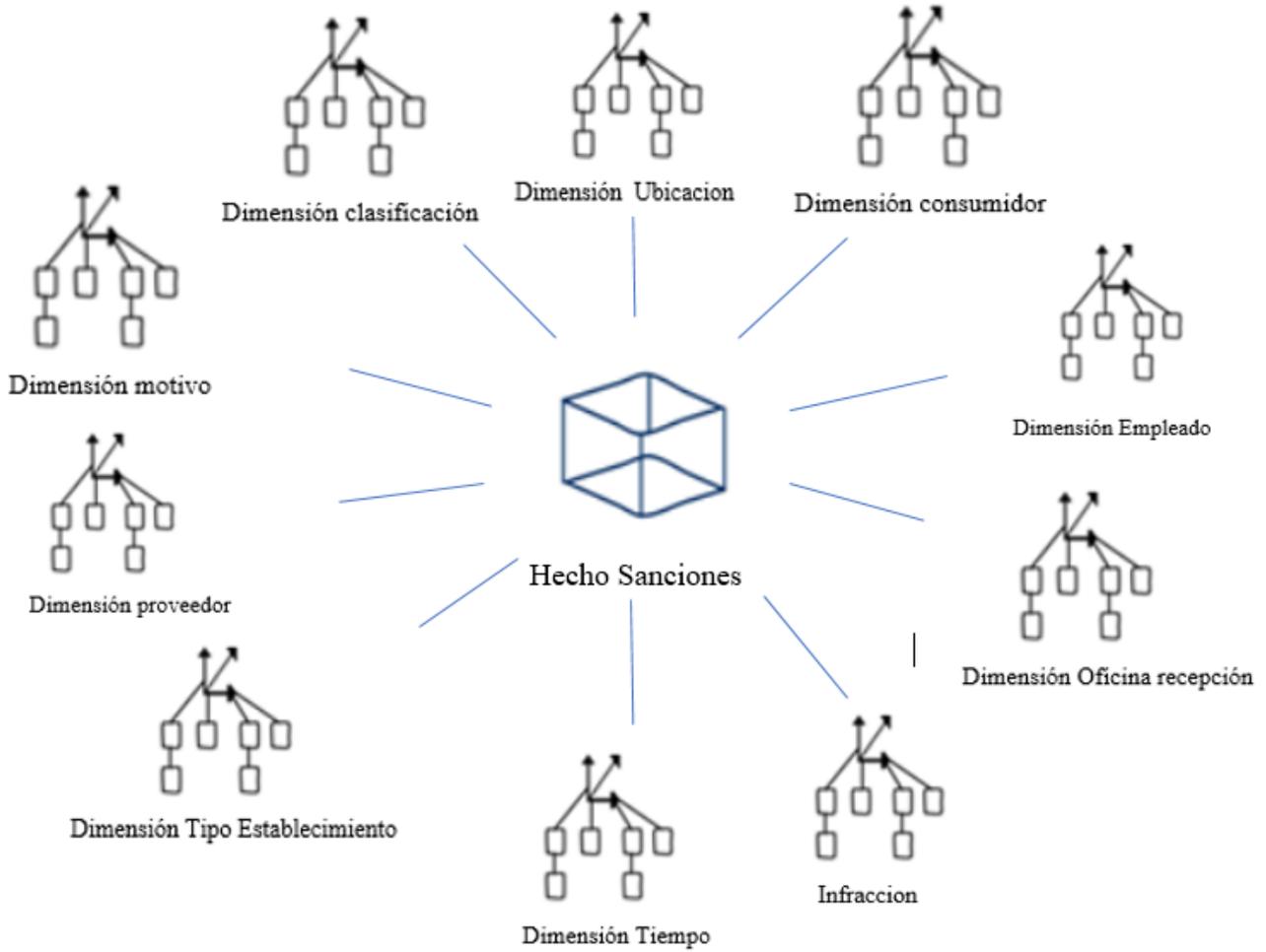


Figura 14.8: Nivel 2 Diagrama Conceptual Data Mart (Fact sanciones)

14.5.1.3 Dimensiones y Hechos

14.5.1.3.1 Fact denuncias ingresadas

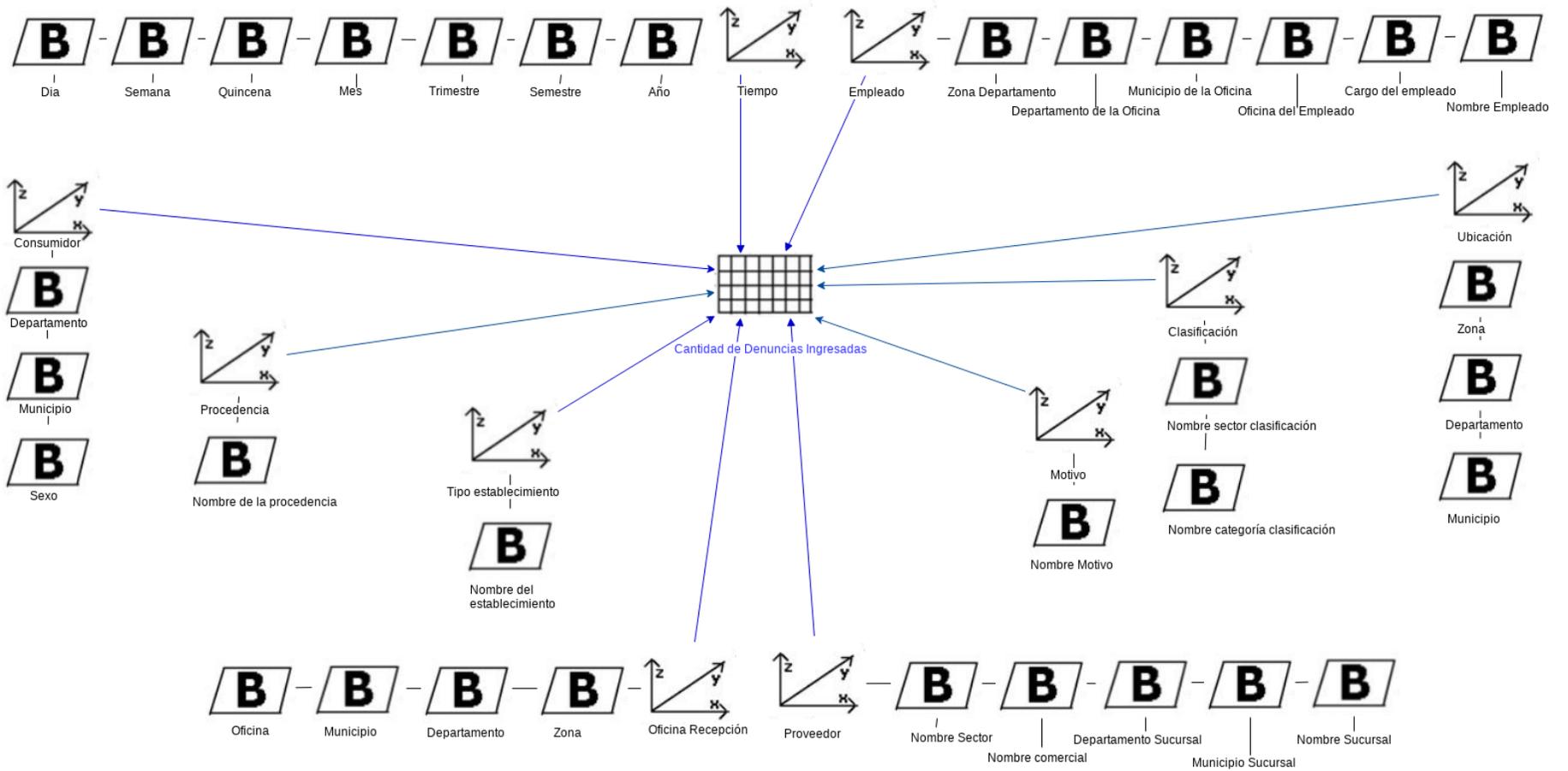


Figura 14.9: Nivel 3 Diagrama conceptual Data Mart Tribunal Sancionador

14.5.1.3.2 Fact resoluciones

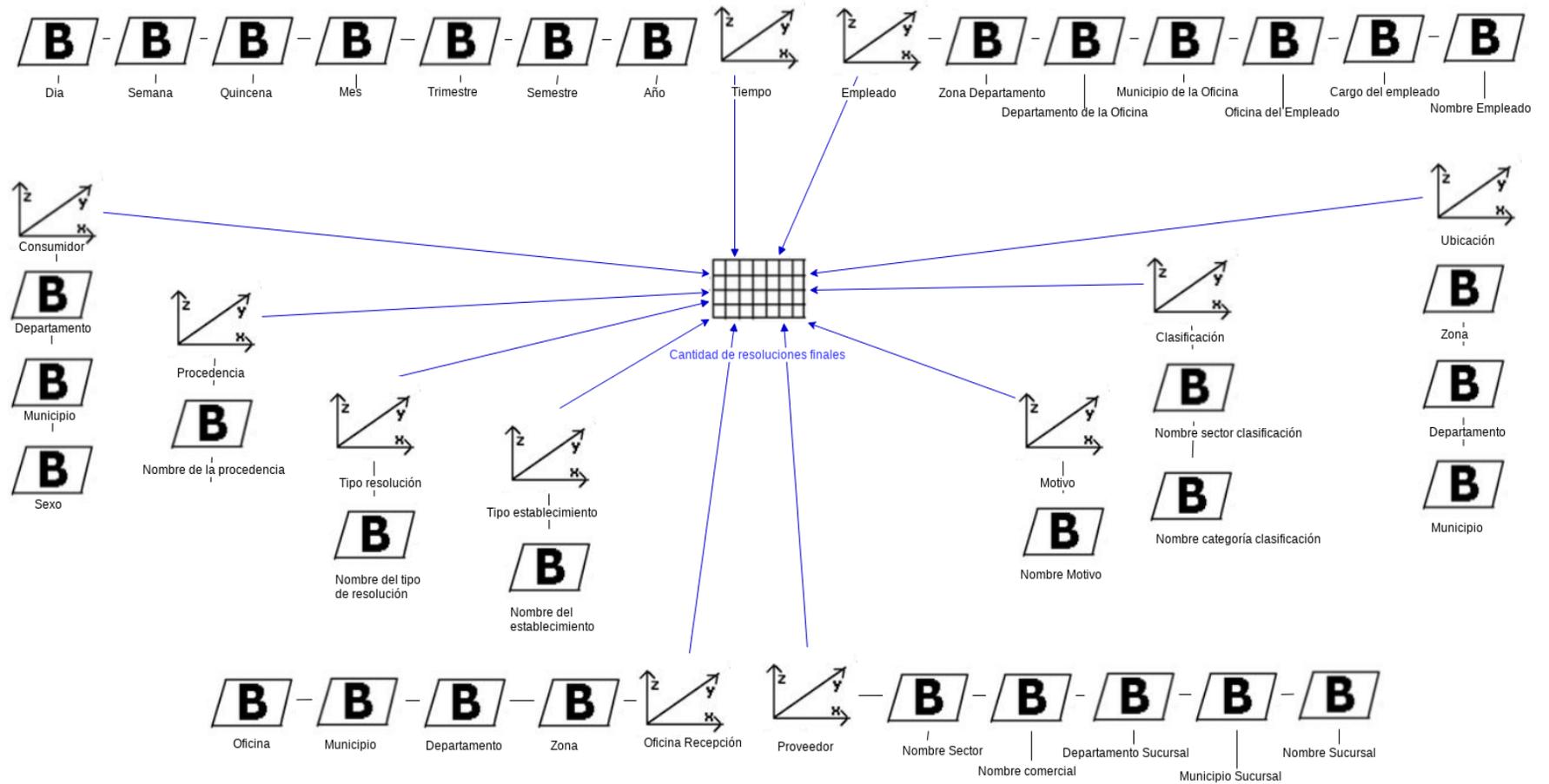


Figura 14.10: Nivel 3 Diagrama conceptual Data Mart Tribunal Sancionador

14.5.1.3.3 Fact sanciones

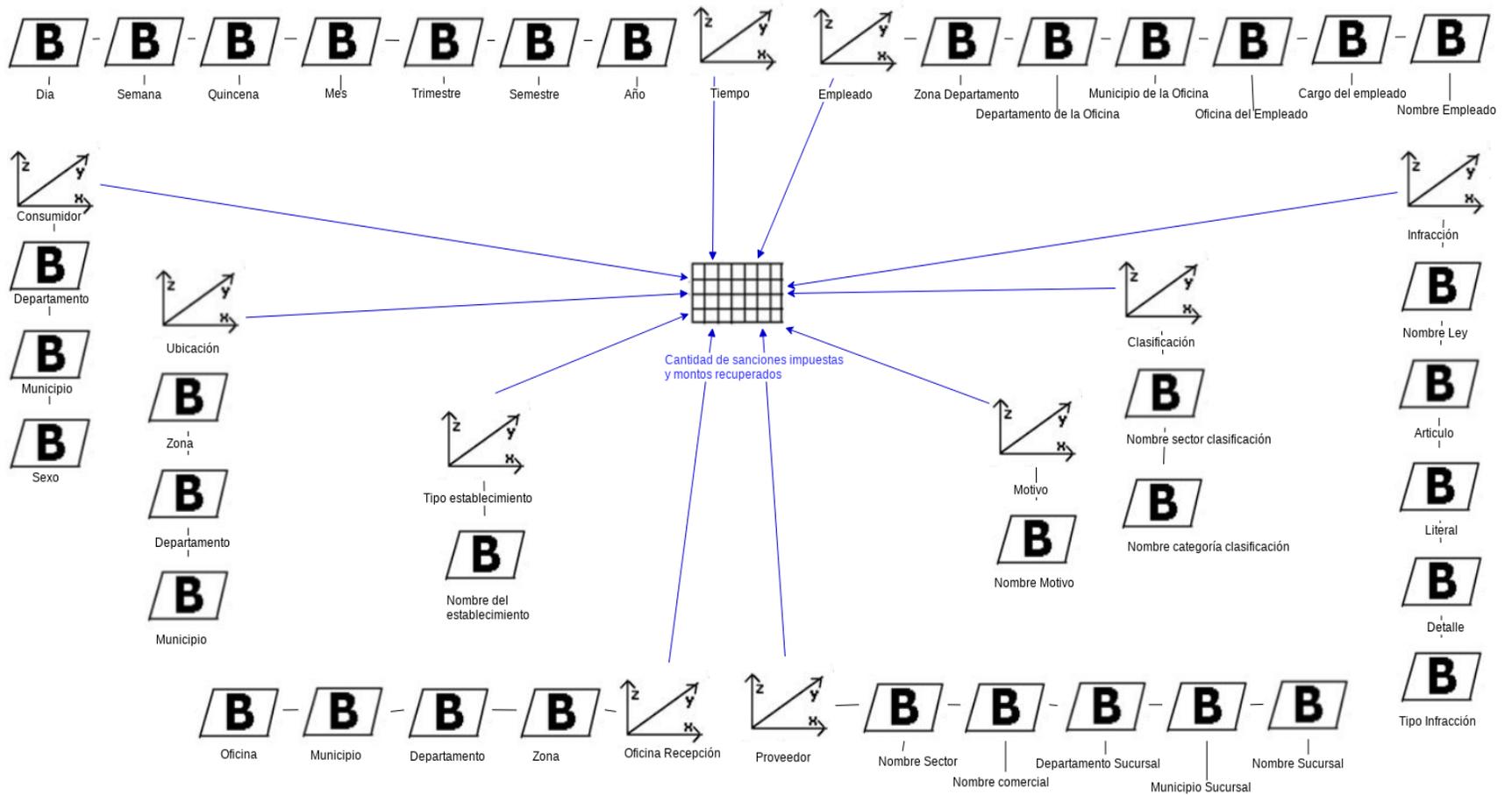


Figura 14.11: Nivel 3 Diagrama conceptual Data Mart Tribunal Sancionador

14.5.2 Diseño Físico Base de Datos Data mart

14.5.2.1 Fact Denuncias

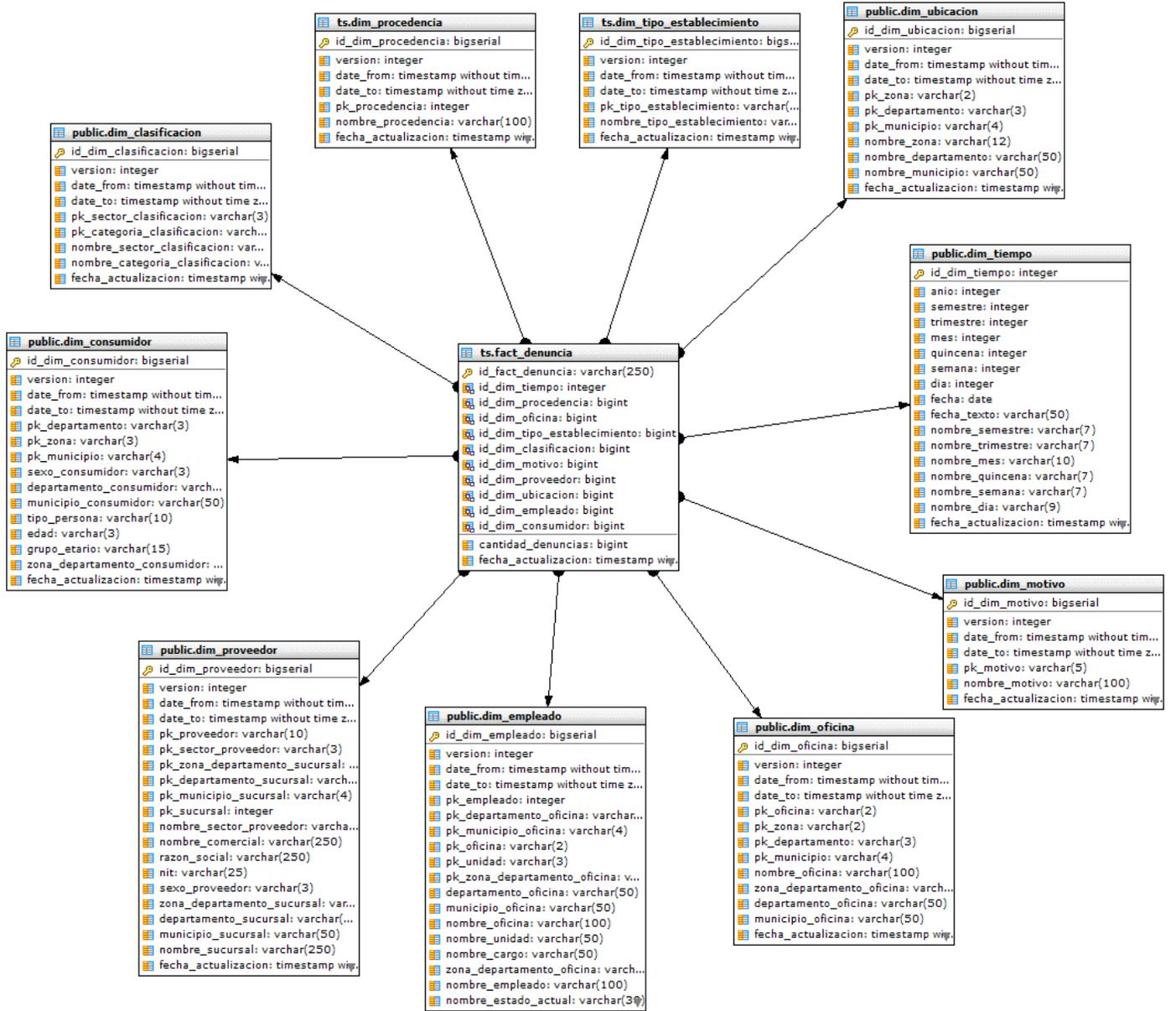


Figura 14.12: Diseño físico Data Mart - Fact Denuncias

14.5.2.2 Fact Resoluciones

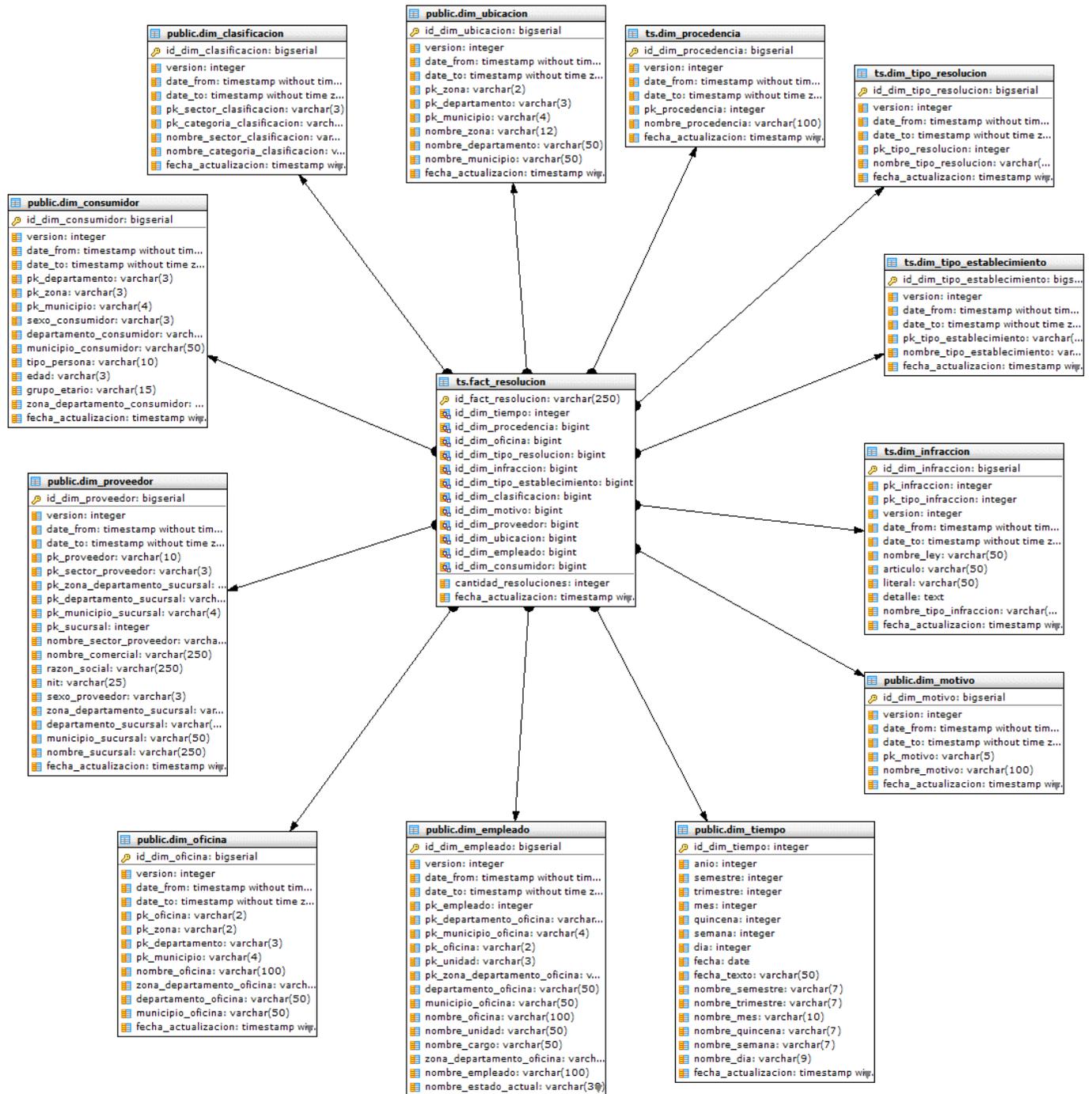


Figura 14.13: Diseño Físico Data Mart - Fact Resoluciones

14.5.3 Diseño Físico Base de Datos Staging Area

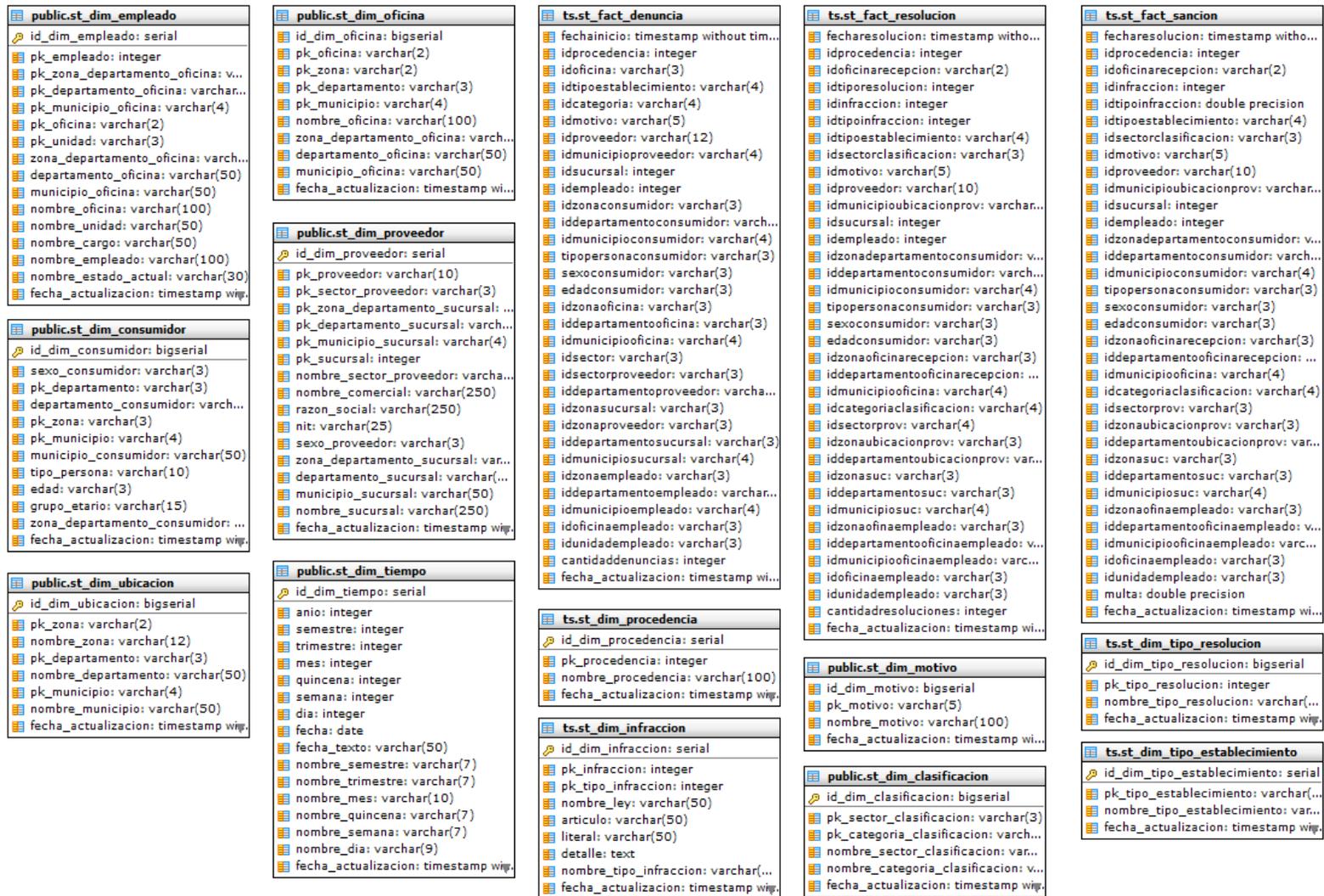


Figura 14.14: Diseño físico de Staging Area

14.6 Diseño de Procesos ETL

En el contexto de la inteligencia de negocio, las herramientas ETL han sido la opción usual para alimentar el Data Warehouse (Jose Curto). En el proceso de ETL los datos son extraídos desde una base de datos OLTP, transformada para que coincida con el esquema de datos del Data Warehouse y cargados en la base de datos del almacén de datos.

A continuación, se presenta el diseño ETL elaborado para el modelo multidimensional de la unidad de Tribunal Sancionador.

14.6.1 Procesos ETL Staging

En los siguientes apartados se listan los procesos ETL de ambos grupos de diagramas, seguido de un diagrama representativo para esta unidad y los diseños de los procesos restantes se colocarán en el CD que acompaña al documento en la carpeta [CD-EISI\Diseños de Procesos ETL\ts \(Tribunal Sancionador\)](#)

Diagramas desarrollados:

St_dim_tipo_establecimiento	St_dim_procedencia
St_dim_infraccion	St_dim_tipo_resolucion
St_fact_resolucion	St_fact_sancion
	St_fact_denuncia

14.6.2 Procesos ETL Data Mart

A continuación, se listan los procesos ETL que mueven datos del Staging Area hacia el Data Mart desarrollados para esta unidad, seguido de un ejemplo (diagrama) para ilustrar uno de los procesos ETL.

Diagramas desarrollados:

Dw_dim_tipo_establecimiento	Dw_dim_consumidor	Dim_dim_ubicacion
Dw_dim_tipo_resolucion	Dw_dim_oficina	Dw_fact_denuncia
Dw_dim_infraccion	Dw_empleado	Dw_fact_resolucion
Dw_dim_procedencia	Dw_dim_proveedor	Dw_fact_sancion
Dw_dim_clasificacion	Dw_dim_motivo	

A continuación, se ilustra el diseño del proceso de Figura 14.15: Proceso ETL dw_dim_procedencia

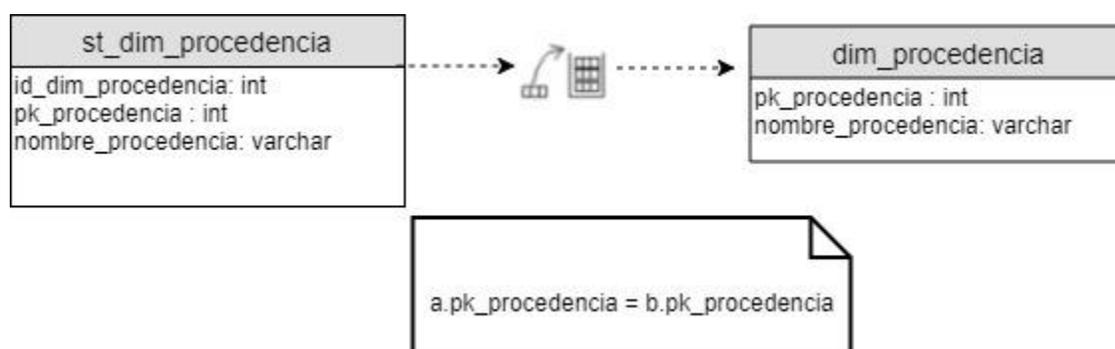


Figura 14.15: Proceso ETL dw_dim_procedencia

14.7 Pruebas

14.7.1 Pruebas de Procesos ETL

La realización de las pruebas consistió en ejecutar los Jobs almacenados en el servidor de Staging Área. Las transformaciones finales para el Data Mart de Tribunal Sancionador se muestran en la Tabla 14.5.

No	Transformación	Descripción
1	dw_incremental_fact_denuncia.ktr	Transformación para cargas incrementales desde st_fact_denuncia (staging area) hacia fact_denuncia (data warehouse).
2	dw_incremental_fact_resolucion.ktr	Transformación para cargas incrementales desde st_fact_resolucion (staging area) hacia fact_resolucion (data warehouse).
3	dw_incremental_fact_sancion.ktr	Transformación para cargas incrementales desde st_fact_sancion (staging area) hacia fact_sancion (data warehouse).
4	dw_inicial_fact_denuncia.ktr	Transformación para carga inicial desde st_fact_denuncia (staging area) hacia fact_denuncia (data warehouse).
5	dw_inicial_fact_resolucion.ktr	Transformación para carga inicial del data warehouse desde st_fact_resolucion (staging area) hacia fact_resolucion (data warehouse).
6	dw_inicial_fact_sancion.ktr	Transformación para carga inicial del data warehouse desde st_fact_sancion (staging area) hacia fact_sancion (data warehouse).

7	dw_unico_dim_infraccion.ktr	Transformación para carga unica (full-refresh) del data warehouse desde st_dim_infraccion (staging area) hacia dim_infraccion (data warehouse).
8	dw_unico_dim_procedencia.ktr	Transformación para carga unica (full-refresh) del data warehouse desde st_dim_procedencia (staging area) hacia dim_procedencia (data warehouse).
9	dw_unico_dim_tipo_establecimiento.ktr	Transformación para carga unica (full-refresh) del data warehouse desde st_dim_tipo_establecimiento (staging area) hacia dim_tipo_establecimiento (data warehouse).
10	dw_unico_dim_tipo_resolucion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde st_dim_tipo_resolucion (staging area) hacia dim_tipo_resolucion (data warehouse).
11	st_incremental_fact_denuncia.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_fact_denuncia (staging area).
12	st_incremental_fact_resolucion.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_fact_resolucion (staging area).
13	st_incremental_fact_sancion.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia st_fact_sancion (staging area).
14	st_inicial_fact_denuncia.ktr	Transformación para carga inicial del data warehouse desde st_fact_denuncia (staging area) hacia fact_denuncia (data warehouse).
15	st_inicial_fact_resolucion.ktr	Transformación para carga inicial del data warehouse desde st_fact_resolucion (staging area) hacia fact_resolucion (data warehouse).
16	st_inicial_fact_sancion.ktr	Transformación para carga inicial del data warehouse desde st_fact_sancion (staging area) hacia fact_sancion (data warehouse).
17	st_unico_dim_infraccion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_infraccion (staging área).
18	st_unico_dim_procedencia.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia st_dim_procedencia (staging área).
19	st_unico_dim_tipo_establecimiento.ktr	Transformación para carga única o full refresh del data warehouse desde las tablas fuentes hacia st_dim_tipo_establecimiento (staging area).
20	st_unico_dim_tipo_resolucion.ktr	Transformación para carga única o full refresh del data warehouse desde las tablas fuentes hacia st_dim_tipo_resolucion (staging area).

Tabla 14.5: Transformaciones finales para Data Mart Tribunal Sancionador

Los Jobs finales de Primer Nivel para Data Mart de Tribunal Sancionador se muestran la Tabla 14.6. La tabla también muestra el orden en que ejecutan las transformaciones de cada uno de los Jobs.

No	Job Primer Nivel	Transformación	
		No	Nombre Transformación
1	job_incremental_fact_denuncia.kjb	1	st_incremental_fact_denuncia.ktr
		2	dw_incremental_fact_denuncia.ktr
2	job_incremental_fact_resolucion.kjb	1	st_incremental_fact_resolucion.ktr
		2	dw_incremental_fact_resolucion.ktr
3	job_incremental_fact_sancion.kjb	1	st_incremental_fact_sancion.ktr
		2	dw_incremental_fact_sancion.ktr

4	job_inicial_fact_denuncia.kjb	1	st_inicial_fact_denuncia.ktr
		2	dw_inicial_fact_denuncia.ktr
5	job_inicial_fact_resolucion.kjb	1	st_inicial_fact_resolucion.ktr
		2	dw_inicial_fact_resolucion.ktr
6	job_inicial_fact_sancion.kjb	1	st_inicial_fact_sancion.ktr
		2	dw_inicial_fact_sancion.ktr
7	job_unico_dim_infraccion.kjb	1	st_unico_dim_infraccion.ktr
		2	dw_unico_dim_infraccion.ktr
8	job_unico_dim_procedencia.kjb	1	st_unico_dim_procedencia.ktr
		2	dw_unico_dim_procedencia.ktr
9	job_unico_dim_tipo_establecimiento.kjb	1	st_unico_dim_tipo_establecimiento.ktr
		2	dw_unico_dim_tipo_establecimiento.ktr
10	job_unico_dim_tipo_resolucion.kjb	1	st_unico_dim_tipo_resolucion.ktr
		2	dw_unico_dim_tipo_resolucion.ktr

Tabla 14.6: Jobs de Nivel 1 para Data Mart Tribunal Sancionador

Los Jobs finales de Segundo Nivel para Data Mart de Tribunal Sancionador se muestran la Tabla 14.7. La tabla también muestra el orden en que ejecutan las transformaciones cada uno de los Jobs, los Jobs presentados a continuación reflejan el orden en que deberán ser ejecutados.

No	Jobs Segundo Nivel	Jobs primer nivel ejecutados	
		No	Nombre Job
1	job_inicial_ts.kjb	1	job_unico_dim_infraccion.kjb
		2	job_unico_dim_procedencia.kjb
		3	job_unico_dim_tipo_establecimiento.kjb
		4	job_unico_dim_tipo_resolucion.kjb
		5	job_inicial_fact_sancion.kjb
		6	job_inicial_fact_resolucion.kjb
		7	job_inicial_fact_denuncia.kjb
2	job_inicial_ts.kjb	1	job_unico_dim_infraccion.kjb
		2	job_unico_dim_procedencia.kjb
		3	job_unico_dim_tipo_establecimiento.kjb
		4	job_unico_dim_tipo_resolucion.kjb
		5	job_incremental_fact_sancion.kjb
		6	job_incremental_fact_resolucion.kjb
		7	job_incremental_fact_denuncia.kjb

Tabla 14.7: Jobs de Nivel 2 para Data Mart Tribunal Sancionador

La Figura 14.16 y Figura 14.17 muestra los Jobs de segundo nivel los cuales son los Jobs principales que son ejecutados en el servidor. La figura Figura 14.16 muestra la secuencia del job que realiza la carga inicial, mientras que la Figura 14.17 muestra la secuencia del job que servirá para realizar las cargas incrementales.

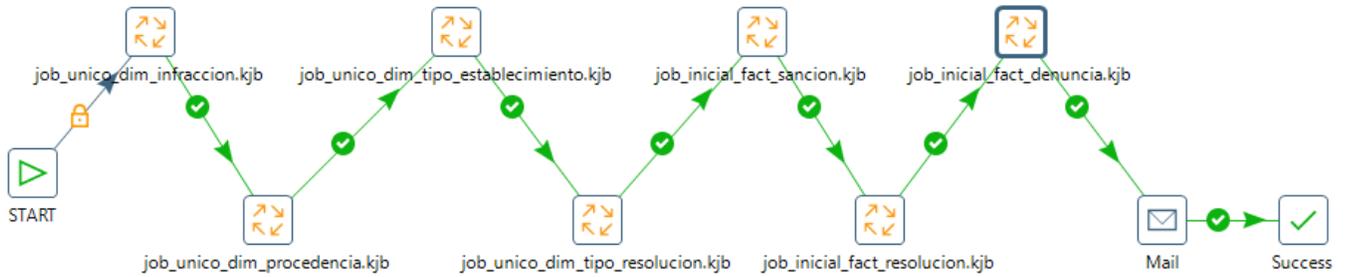


Figura 14.16: Job ETL para la carga inicial del Data Mart Tribunal Sancionador

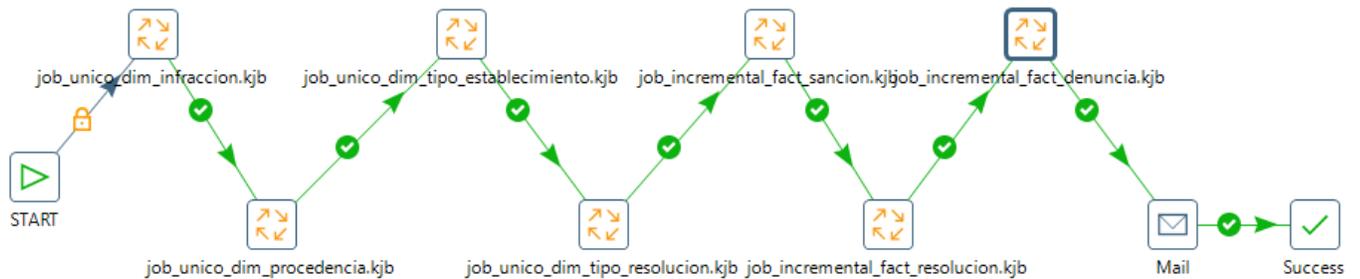


Figura 14.17: Job ETL para la carga incremental del Data Mart Tribunal Sancionador

14.7.1.1 Ejecución de Procesos ETL iniciales

El comando para ejecutar el proceso ETL se muestra a continuación:

```
#/root/data-integration/kitchen.sh -file=/root/jobs/ts/job_inicial_ts.kjb -level=Debug
```

El tiempo estimado para la ejecución de los procesos ETL es el siguiente:

Cantidad de Tranformaciones	Tiempo	Registros
14	00:12:45.186	26,731

La finalización del proceso ETL se notifica mediante correo electrónico como se muestra en la Figura 14.18

☐ ☆ ▷	DATAWAREHOUSE	INITIAL ETL FINISHED FOR sprint3 (Tribunal Sancionador) - EL PROCESO ETL INICIAL PARA sp...
☐ ☆ ▷	DATAWAREHOUSE	INITIAL ETL FINISHED FOR fact_denuncia - EL PROCESO ETL INICIAL PARA fact_denuncia HA ...
☐ ☆ ▷	DATAWAREHOUSE	INITIAL ETL FINISHED FOR fact_resolucion - EL PROCESO ETL INICIAL PARA fact_resolucion H...
☐ ☆ ▷	DATAWAREHOUSE	INITIAL ETL FINISHED FOR fact_sancion - EL PROCESO ETL INICIAL PARA fact_sancion HA FIN...
☐ ☆ ▷	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_tipo_resolucion - EL PROCESO ETL FULL REFRESH PAR...
☐ ☆ ▷	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_tipo_establecimiento - EL PROCESO ETL FULL REFRES...
☐ ☆ ▷	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_procedencia - EL PROCESO ETL FULL REFRESH PARA d...
☐ ☆ ▷	DATAWAREHOUSE	FULL REFRESH ETL FINISHED FOR dim_infraccion - EL PROCESO ETL FULL REFRESH PARA di...

Figura 14.18:Notificación de finalización de los procesos ETL (Carga Inicial)

14.7.2 Pruebas Cubos OLAP

Dado que para Tribunal Sancionador se identificaron tres hechos de los cuales se desea analizar información, se han realizado tres cubos para dicha unidad. A continuación, se muestran los cubos realizados para el Data Mart de Tribunal Sancionador.

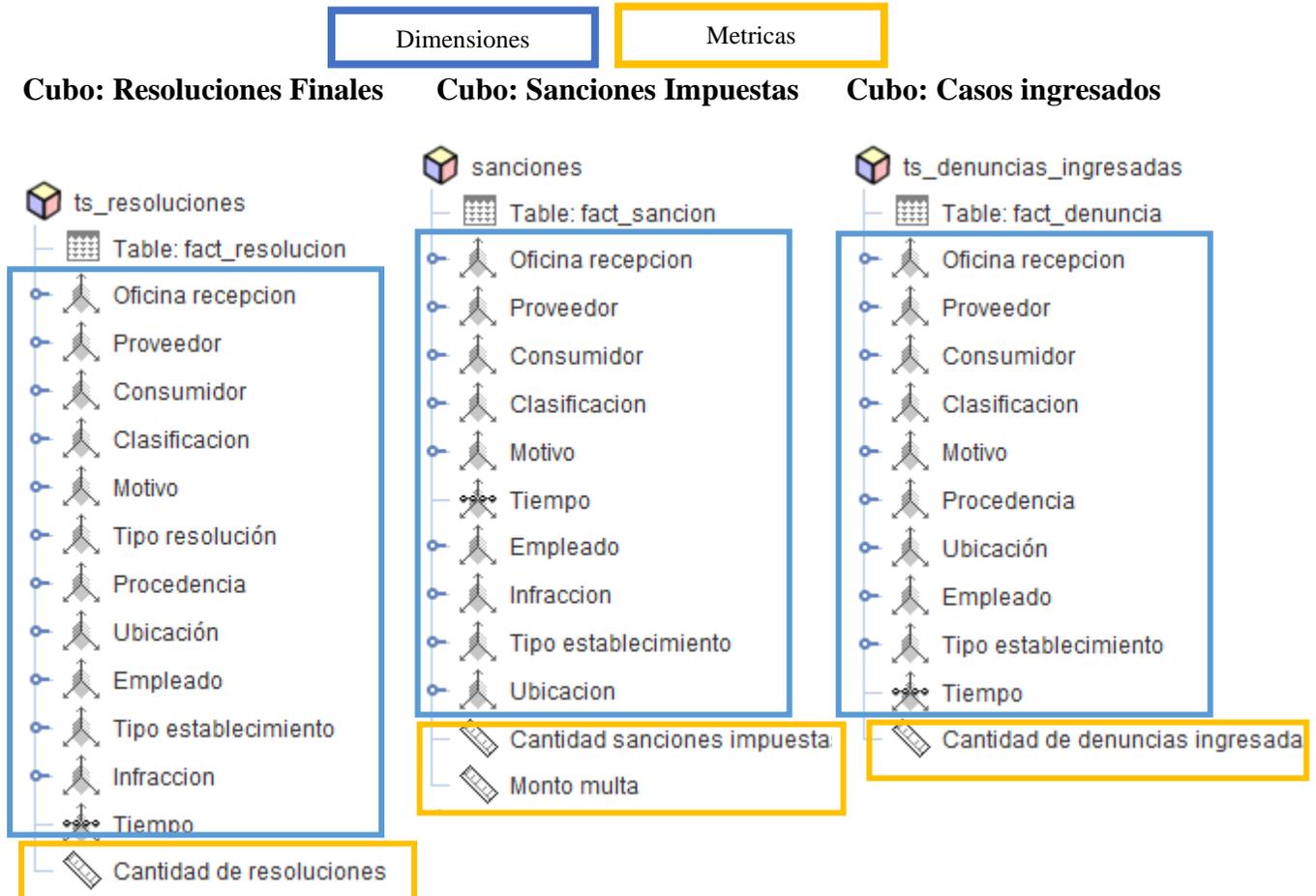


Figura 14.19: Cubos para Tribunal Sancionador

Por cada uno de los cubos de la Figura 14.19, se ha publicado un archivo XML en el Servidor de Pentaho Server BI.

Con el uso del Plugin JPivot para Pentaho Server se logra visualizar la información según la información en Figura 14.20, Figura 14.21 y Figura 14.22.

Clasificación	Motivo	Tipo de resolución	Procedencia	Ubicación Proveedor	Empleados	Tipo de establecimiento	Infracciones	Tiempo	Medidas
las clasificaciones	Todos los motivos	Todos los tipos de resolución	Todas las procedencias	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todas las leyes	Todos los años	4.213
							LPC	Todos los años	4.209
							LSTC	Todos los años	4
		Absueltos	Todas las procedencias	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todas las leyes	Todos los años	182
							LPC	Todos los años	181
							Art. 42	Todos los años	45
								2010	12
								2011	6
								2012	11
								2013	11
								2014	5
							Art. 43	Todos los años	30
								2010	28
								2012	1
								2013	1

Figura 14.20: Visualización del Cubo Resoluciones Finales dictadas por Tribunal Sancionador de la Defensoria del Consumidor

Clasificación	Motivo	Tiempo	Empleados	Infracciones	Tipo_establecimiento	Ubicación Proveedor	Medidas
Consumidores	Todas las clasificaciones	Todos los años	Todas las zonas	Todas las infracciones	Todos los establecimientos	Todas las ubicaciones	2.815 2.574.359,95
		2010	Todas las zonas	Todas las infracciones	Todos los establecimientos	Todas las ubicaciones	326 99.905
				LPC	Todos los establecimientos	Todas las ubicaciones	326 99.905
				Art. 42	Todos los establecimientos	Todas las ubicaciones	148 46.195
				e)	Todos los establecimientos	Todas las ubicaciones	148 46.195
				Farmacias	Todas las ubicaciones	Todas las ubicaciones	97 27.585
				N/A	Todas las ubicaciones	Todas las ubicaciones	2 2.000
				Operadores de servicios de telecomunicaciones	Todas las ubicaciones	Todas las ubicaciones	1 175
				Otros	Todas las ubicaciones	Todas las ubicaciones	5 610
				Otros proveedores de alimentos	Todas las ubicaciones	Todas las ubicaciones	3 1.075
				Proveedores de servicios de salud y estética	Todas las ubicaciones	Todas las ubicaciones	6 4.725
				Restaurantes y hoteles	Todas las ubicaciones	Todas las ubicaciones	13 5.775
				Tiendas	Todas las ubicaciones	Todas las ubicaciones	13 2.575
				Tiendas de Conveniencia	Todas las ubicaciones	Todas las ubicaciones	8 1.675
				Art. 43	Todos los establecimientos	Todas las ubicaciones	7 1.225

Figura 14.21: Visualización del Cubo Sanciones Impuestas por Tribunal Sancionador

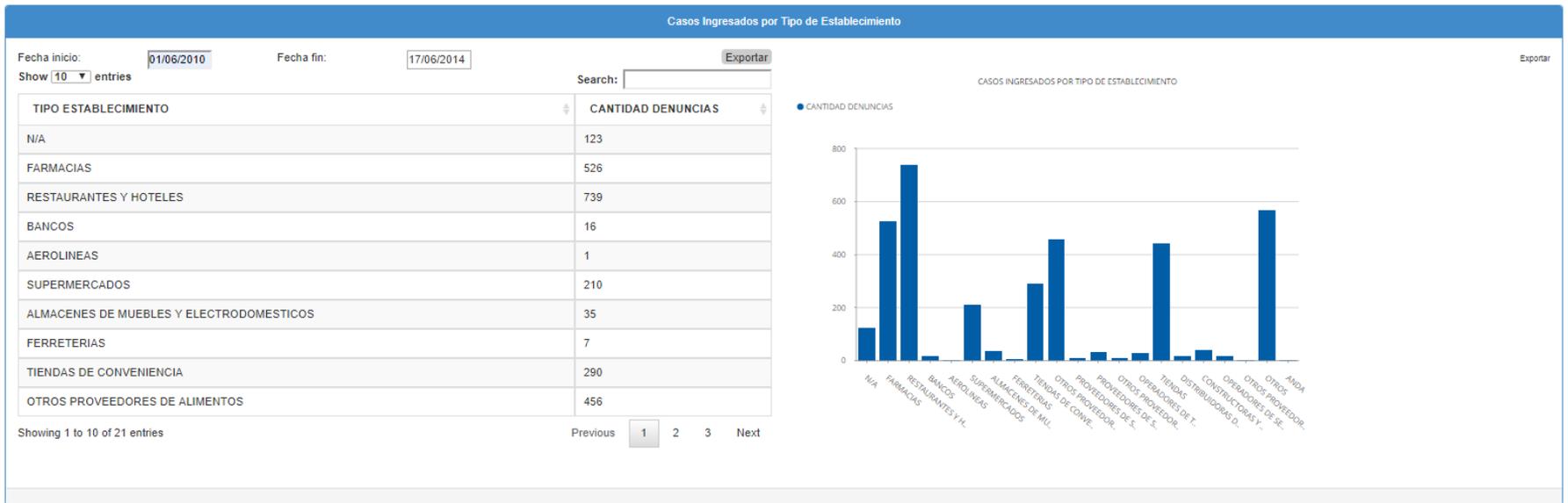
Actor	Consumidor	Clasificación	Motivo	Procedencia	Ubicación Proveedor	Empleados	Tipo de establecimiento	Tiempo	Medidas						
Actores	Todos los consumidores	Todas las clasificaciones	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	4.541						
		Agua Potable	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	3						
		Agua potable	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	68						
		Alimentos y bebidas	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	3						
		Comercio	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	78						
		Electrodomésticos	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	24					
										CSC	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	4
										DVM - Unidad de Inspecciones	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	20
		Eléctrico	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	22					
		Gobierno y alcaldías	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	5					
		Inmuebles	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	10					
		Libros	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	7					
		Medicamentos	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	1					
Muebles	Todos los motivos	Todas las procedencias	Todas las zonas	Todas las zonas	Todas las zonas	Todos los Tipos de establecimiento	Todos los años	11							

Figura 14.22: Visualización del Cubo Denuncias o casos Ingresados a Tribunal Sancionador

14.7.3 Pruebas Dashboard

Uno de los Dashboards realizado para Tribunal Sancionador se muestra en la siguiente página.

14.7.3.1 Dashboard Tribunal Sancionador



15 SPRINT 4

15.1 Sprint Backlog

A continuación, se presenta el Sprint Backlog, que refleja las historias de usuarios, su estimación y asignación para este Sprint.

#	Historia	Punto Estimado	Asignado
	Redacción de la plantilla para el refinamiento de requerimientos	20	Todos
2.	Poblar la base de datos de desarrollo SARA	13	Samuel
3.	Diseño del modelo multidimensional	8	Samuel – Ivis
4.	Diseño BD del Staging Area	3	Jorge – Samuel
5.	Diseño del proceso ETL	8	Jorge – Ivis – Samuel
6.	Diseño de vistas	3	Jorge - Ivis
7.	Creación de Script SQL del Staging Area	13	Jorge – Ivis – Samuel
8.	Creación de Script SQL modelo multidimensional	13	Jorge – Ivis – Samuel
9.	Desarrollo de los procesos ETL	100	Jorge – Ivis – Samuel
10.	Desarrollo de cubos OLAP	40	Jorge – Ivis – Samuel
11.	Desarrollo de vistas	13	Jorge – Ivis – Samuel
12.	Implementación e integración de componentes	40	Jorge – Ivis – Samuel
13.	Pruebas	20	Jorge – Ivis – Samuel
14.	Documentación	40	Todos

Tabla 15.1: Sprint backlog Sprint 4

15.2 Descripción de Historias de Usuario

A continuación, se presenta la descripción de las historias de usuario detallando: el responsable, el sprint al que pertenece dicha historia, una breve descripción, sus criterios de aceptación y los puntos de estimación. Previamente a las historias de usuario se presenta la historia épica del Sprint 4

<i>Historia Épica #5</i>	Data Mart de Centro de Solución de Controversias.
<i>Responsable</i>	Jorge Iraheta/Esau Serpas/Ivis Navarro/Samuel Núñez Sprint: 4
<i>Descripción</i>	Como Dirección de Centro de Solución de Controversias (jefatura) Necesito: Un mecanismo para analizar la cantidad de atenciones recibidas, los casos que alcanzaron la etapa de cerrados y la cantidad de casos que han sido reaperturados. Para: realizar evaluaciones comparativas de rendimiento del área técnica.
<i>Criterios de aceptación</i>	- La cantidad de atenciones recibidas deben poder ser visualizar por las siguientes perspectivas: <ul style="list-style-type: none"> • Por fecha • Por lugar móvil • Por tipo atención • Por forma de recepción • Por oficina de recepción • Por proveedor • Por consumidor • Por clasificación (Sector y Categoría) • Por motivo • Por campaña • Por ubicación del proveedor • Por tipo bien • Por institución • Por empleado

	<ul style="list-style-type: none"> • Por etapa <p>- La cantidad de casos cerrados junto a la cantidad de tiempo ingresado, tiempo de avenimiento, tiempo de ratificación y el tiempo de conciliación si lo hubiera, deben poder analizarse desde las siguientes perspectivas:</p> <ul style="list-style-type: none"> • Por fecha • Por proveedor • Por ubicación del proveedor • Por consumidor • Por empleado • Por oficina de recepción • Por clasificación (Sector y Categoría) • Por motivo • Por forma de recepción • Por lugar móvil • Por campaña • Por tipo atención • Por etapa Cierre • Por tipo bien • Por institución que toma la denuncia. <p>- La cantidad de casos reaperturados, deben poder analizarse desde las siguientes perspectivas:</p> <ul style="list-style-type: none"> • Por fecha • Por proveedor • Por ubicación del proveedor • Por consumidor • Por empleado • Por oficina de recepción • Por clasificación (Sector y Categoría) • Por motivo • Por forma de recepción • Por lugar móvil • Por campaña • Por tipo atención • Por etapa Cierre • Por tipo bien • Por institución que toma la denuncia.
Puntos estimados	334

Historia de usuario #1	Redacción de la plantilla para el refinamiento de requerimientos
Responsable	Jorge Iraheta/Esau Serpas/Ivis Navarro/Samuel Núñez Sprint: 4
Descripción	<p>Como Equipo Scrum necesito el refinamiento del requerimiento de información de Centro de Solución de Controversias: Realizar reuniones con los interesados y especificar los elementos de información requeridos para el Centro de Solución de Controversias.</p> <ul style="list-style-type: none"> • Llenar la herramienta de recolección. • Revisar y desglosar los requerimientos solicitados • Firmar documento por el usuario como aceptación de los de los requerimientos. <p>Para tener claridad de la información que desean almacenar, visualizar y analizar mediante el Data Mart o modelo Multidimensional para ésta unidad.</p>
Criterios de aceptación	<p>- Tener llena la plantilla de refinamiento de requerimientos de información.</p> <p>- Incluir únicamente requerimientos de información con naturaleza histórica y no información transaccional.</p>

	<ul style="list-style-type: none"> - Tener un documento firmado de aceptación de requerimientos. - Identificar (mediante un listado) las tablas fuentes que servirán para cumplir con los requerimientos de información solicitados.
Puntos estimados	20

Historia de usuario #2	Poblar base de datos de desarrollo SARA
Responsable	Samuel Núñez Sprint: 4
Descripción	Como Product Owner necesito: migrar los datos necesarios según las tablas fuentes de datos que se identifiquen durante el proceso de toma de requerimientos. Para poseer datos de prueba para el desarrollo del Data Mart.
Criterios de aceptación	<ul style="list-style-type: none"> - Se deben poseer Scripts de INSERT para las tablas Fuente. - Todas las tablas fuentes identificadas en el refinamiento de requerimientos deben poseer datos. - Debe existir al menos una prueba de conexión y acceso remoto exitoso a los datos. - Datos cargados en la Base de datos.
Puntos estimados	13

Historia de usuario #3	Diseño del modelo multidimensional
Responsable	Samuel Núñez/Ivis Navarro Sprint: 4
Descripción	Cómo Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart del Centro de Solución de Controversias. Para tener la base (diagrama) para el desarrollo de Scripts de Base de Datos SQL para Data Mart.
Criterios de aceptación	<ul style="list-style-type: none"> -Esquema del modelo multidimensional de Centro de Solución de Controversias finalizado. -El esquema debe cumplir las reglas de un modelo de tipo Estrella. -El esquema debe poseer las relaciones necesarias para la búsqueda según los criterios de búsqueda definidos por el Centro de Solución de Controversias. -El esquema debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
Puntos estimados	8

Historia de usuario #4	Diseño BD del Staging Area.
Responsable	Jorge Iraheta/Samuel Núñez Sprint: 4
Descripción	Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart del Centro de Solución de Controversias. Para tener el insumo para la creación de Scripts de Base de Datos SQL para Staging Area.
Criterios de aceptación	<ul style="list-style-type: none"> -Deberá existir una o más tablas por cada dimensión. -Deberá existir una o más tablas de hecho. -El diseño debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
Puntos estimados	3

Historia de usuario #5	Diseño del proceso ETL
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
Descripción	Como Equipo Scrum necesito: Los diseños de los proceso ETL (Extracción, Transformación, Carga) Para cumplir con el Diseño del proceso ETL y tenerlo como insumo a la hora de desarrollar dichos procesos ETL.
Criterios de aceptación	<ul style="list-style-type: none"> -Se deberá utilizar nomenclatura UML para los pasos en cada proceso ETL. -Deberá incluir todas las tablas dimensiones y tablas de hechos. -Los nombres de las tablas o etiquetas deben poderse visualizar fácilmente.

<i>Puntos estimados</i>	8
<i>Historia de usuario #6</i>	Diseño de vistas
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro Sprint: 4
<i>Descripción</i>	Como Administrador de Data Warehouse necesito: Realizar el diseño de vistas necesarias para visualizar la información del Data Mart de dirección Jurídica. Para Diseñar y digitalizar la vista de presentación de datos.
<i>Criterios de aceptación</i>	-Deberá existir un diseño de vista o dashboard para cada hecho a medir. -Deberá explicarse cada sección del dashboard o vista. -Etiquetas y textos deben ser legibles a simple vista.
<i>Puntos estimados</i>	3
<i>Historia de usuario #7</i>	Creación de Script SQL del Staging Area
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será procesada, calculada y/o limpiada la información.
<i>Criterios de aceptación</i>	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Staging Area. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
<i>Puntos estimados</i>	13
<i>Historia de usuario #8</i>	Creación de Script SQL Modelo Multidimensional
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el esquema Estrella en el SGBD. Para la implementación del esquema donde será almacenada la información.
<i>Criterios de aceptación</i>	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Modelo Multidimensional. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
<i>Puntos estimados</i>	13
<i>Historia de usuario #9</i>	Desarrollo de los procesos ETL.
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
<i>Descripción</i>	Como Equipo Scrum necesito: Desarrollar los procesos ETL correspondientes Para obtener, trasladar y transformar los datos que serán almacenados en el Data Mart del Centro de Solución de Controversias.
<i>Criterios de aceptación</i>	-Los nombres de los procesos ETL deben poseer la nomenclatura definida por el equipo. -Los procesos ETL deben poseer la documentación necesaria en forma de comentarios dentro de la herramienta PDI. -Las Transformaciones ETL deben haber sido probados en desarrollo (Equipo Local) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de final . -Las transformaciones correspondientes a una dimensión o hecho deben ser incorporados un Job de PDI.

	-Deberá crearse un Job que incorpore los otros Jobs correspondientes a cada dimensión y hecho. -Los Jobs de PDI deben haber sido probados con las bases de datos de prueba y base de datos final. -Las bitácoras de los procesos ETL deben almacenarse en una base de datos.
Puntos estimados	100

Historia de usuario #10	Desarrollo de Cubos OLAP
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
Descripción	Como Equipo Scrum necesito: una investigación sobre la herramienta para el desarrollo de cubos OLAP e implementar dicha herramienta para la creación de cubos OLAP. Para implementar el cubo correspondientes al Centro de Solución de Controversias.
Criterios de aceptación	-El cubo deberá poderse visualizar correctamente y sin errores. -Deberá existir como mínimo un cubo OLAP por Hecho. -Deberán poderse realizar las operaciones como mínimo de: Roll Up y Drill Down.
Puntos estimados	40

Historia de usuario #11	Desarrollo de vistas
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
Descripción	Como Equipo Scrum necesito: Desarrollar vistas o dashboards según el diseño previamente elaborado. Para Presentar información relevante hacia la unidad para la toma de decisiones.
Criterios de aceptación	-Vista o dashboard funcionando sin errores técnicos. -Debe coincidir con el diseño realizado previamente. -Debe existir al menos un dashboar por tabla de hechos.
Puntos estimados	13

Historia de usuario #12	Implementación e integración de componentes
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
Descripción	Como Equipo Scrum necesito: La integración de todos los elementos del modelo multidimensional <ul style="list-style-type: none"> - ETL - Fuentes - Stagin Área - BD multidimensional - Pentaho Server (Vistas) Para el desarrollo del modelo multidimensional de forma integral y probar el funcionamiento correctamente de la totalidad del Data Mart.
Criterios de aceptación	- Se deberán modificar los procesos ETL para que utilicen las bases de datos finales (no las de pruebas). - Se deberán realizar pruebas en la visualización de los dashboards. - Todo el entorno (desde los procesos ETLs, hasta los dashboards o vistas) deberán tener el visto bueno de todo el equipo.
Puntos estimados	40

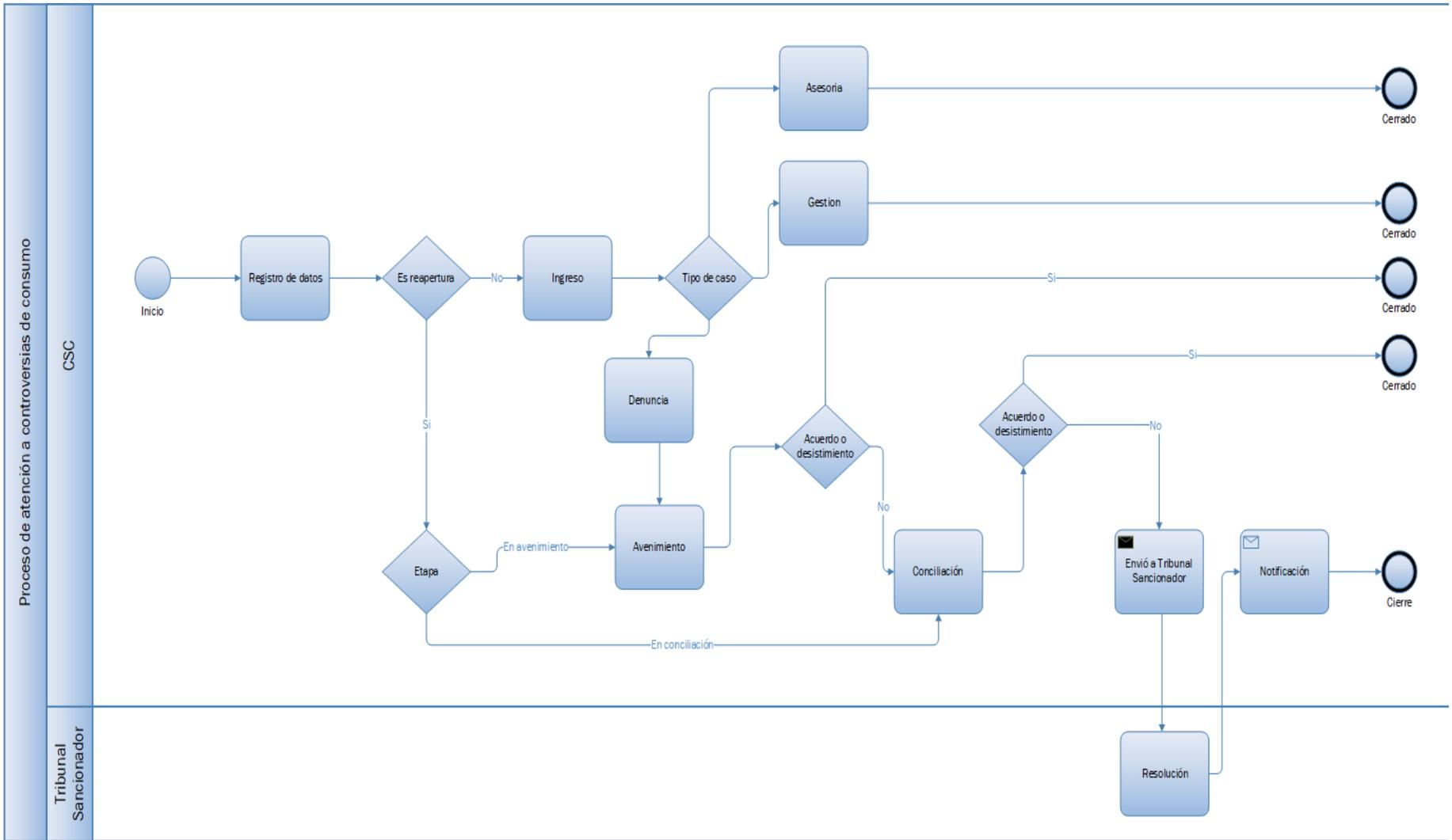
<i>Historia de usuario #13</i>	Pruebas
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 4
<i>Descripción</i>	Como Equipo Scrum necesito: Realizar pruebas al modelo multidimensional terminado. Para verificar que funciona de la manera que se espera.
<i>Criterios de aceptación</i>	Pruebas exitosas en pruebas de aceptación del usuario.
<i>Puntos estimados</i>	20

<i>Historia de usuario #14</i>	Documentación
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Nuñez/Esaú Serpas Sprint: 4
<i>Descripción</i>	Como Equipo Scrum necesito: Documentación completa con historias, nomenclatura y diagramas empleados. Para verificar cómo se llevó a cabo el desarrollo de los cubos de la cuarta unidad (Centro de Solución de Controversias).
<i>Criterios de aceptación</i>	Entrega de Manuales de creación de Data Mart y documentos completos
<i>Puntos estimados</i>	40

15.3 Refinamiento del Requerimiento de Información

Para la realización del refinamiento de los requerimientos de información se partió con la obtención del proceso general que sigue el Centro de Solución de Controversias dentro la Defensoría del Consumidor. Además de establecer el proceso general de la unidad se hizo uso de los paquetes de información para refinar el requerimiento de información.

15.3.1 Proceso BPMN de la Unidad



15.3.2 Paquetes de información

Para la Dirección del Centro de Solución de Controversias de la Defensoría del Consumidor (CSC) se obtuvieron dos paquetes de información para los siguientes temas:

- Cantidad de atenciones recibidas.
- Cantidad de casos cerrados.

Cantidad de casos reaperturados.

Tema:		Cantidad de atenciones recibidas en la Defensoría del consumidor.												
J E R A R Q U I A S	Tiempo	Lugar Movil	Tipo Atencion	Forma de Recepción	Oficina Recepción	Proveedor	Consumidor	Clasificación	Motivo	Campaña	Ubicacion proveedor	Tipo Bien	Institucion que toma la denuncia	Empleado
	Año	Nombre Lugar	Nombre tipo atención	Nombre forma recepción	Zona	NombreSector	Sexo	Nombre sector	Nombre Motivo	Zona Campaña	Zona	Nombre Tipo Bien	Nombre de la institucion	Zona Departamento
	Semestre				Departamento	NombreComercial	Zona	Nombre categoría		Departamento	Departamento		Sector de la ventanilla	Departamento de la Oficina
	Trimestre				Municipio	RazonSocial	Departamento			Municipio	Municipio		Nombre ventanilla oficina	Municipio de la Oficina
	Mes				Oficina	NIT	Municipio			Nombre campaña				Oficina del Empleado
	Quincena					DepartamentoSucursal								Cargo del Empleado
	Semana					MunicipioSucursal								Nombre Empleado
	Dia					NombreSucursal								
	Hechos Medidos:		Cantidad de atenciones recibidas											

Figura 15.2: Paquete de Información: atenciones recibidas en la Defensoría del Consumidor.

Tema:		Casos Cerrados por Dirección del Centro de Solución de Controversias													
J E R A R Q U I A S	Tiempo	Proveedor	Ubicación del Proveedor	Consumidor	Empleado	Oficina Recepción	Clasificación	Motivo	Forma Recepción	Lugar Movil	Campaña	Tipo Atención	Etapa Cierre	Tipo Bien	Institucion que toma la denuncia
	Año	Nombre Sector	Zona	Sexo	Zona Departamento	Zona	Nombre sector clasificación	Nombre Motivo	Nombre forma recepcion	Nombre Lugar	Zona	Nombre Tipo Atención	Nombre Etapa	Nombre Tipo Bien	Nombre de la institucion
	Semestre	Nombre Comercial	Departamento	Zona	Departamento de la Oficina	Departamento	Nombre categoría clasificación				Departamento		Nombre Subetapa		
	Trimestre	Razon Social	Municipio	Departamento	Municipio de la Oficina	Municipio					Municipio				
	Mes	NIT		Municipio	Oficina del Empleado	Oficina					Nombre Campaña				
	Quincena	Departamento Sucursal			Cargo del Empleado										
	Semana	Municipio Sucursal			Nombre Empleado										
	Dia	Nombre Sucursal													
Hechos Medidos:		Tiempo Ingresado, Tiempo Avenimiento, Tiempo Ratificación, Tiempo Conciliación, Tiempo Total (Tiempo de Solución), Monto Reclamado, Monto Recuperado, Cantidad de Casos (Medida calculada)													

Figura 15.3: Paquete de Información: Casos cerrados por CSC

Tema:		Casos Reaperturados por Dirección del Centro de Solución de Controversias													
J E R A R Q U I A S	Tiempo	Proveedor	Ubicación del Proveedor	Consumidor	Empleado	Oficina Recepción	Clasificación	Motivo	Forma Recepción	Lugar Movil	Campaña	Tipo Atención	Etapa Cierre	Tipo Bien	Institucion que toma la denuncia
	Año	Nombre Sector	Zona	Sexo	Zona Departamento	Zona	Nombre sector clasificación	Nombre Motivo	Nombre forma recepcion	Nombre Lugar	Zona	Nombre Tipo Atención	Nombre Etapa	Nombre Tipo Bien	Nombre de la institucion
	Semestre	Nombre Comercial	Departamento	Zona	Departamento de la Oficina	Departamento	Nombre categoría clasificación				Departamento		Nombre Subetapa		Sector de la ventanilla
	Trimestre	Razon Social	Municipio	Departamento	Municipio de la Oficina	Municipio					Municipio				Nombre ventanilla oficina
	Mes	NIT		Municipio	Oficina del Empleado	Oficina					Campaña				
	Quincena	Departamento Sucursal			Cargo del Empleado										
	Semana	Municipio Sucursal			Nombre Empleado										
	Día	Nombre Sucursal													
	Hechos Medidos:		Cantidad de Casos Reaperturados												

Figura 15.4: Paquete de Información: casos reaperturados por CSC

15.4 Especificación de Requerimientos

15.4.1 Casos de Uso Análisi de actas finalizadas en el Centro de Solución de Controversias

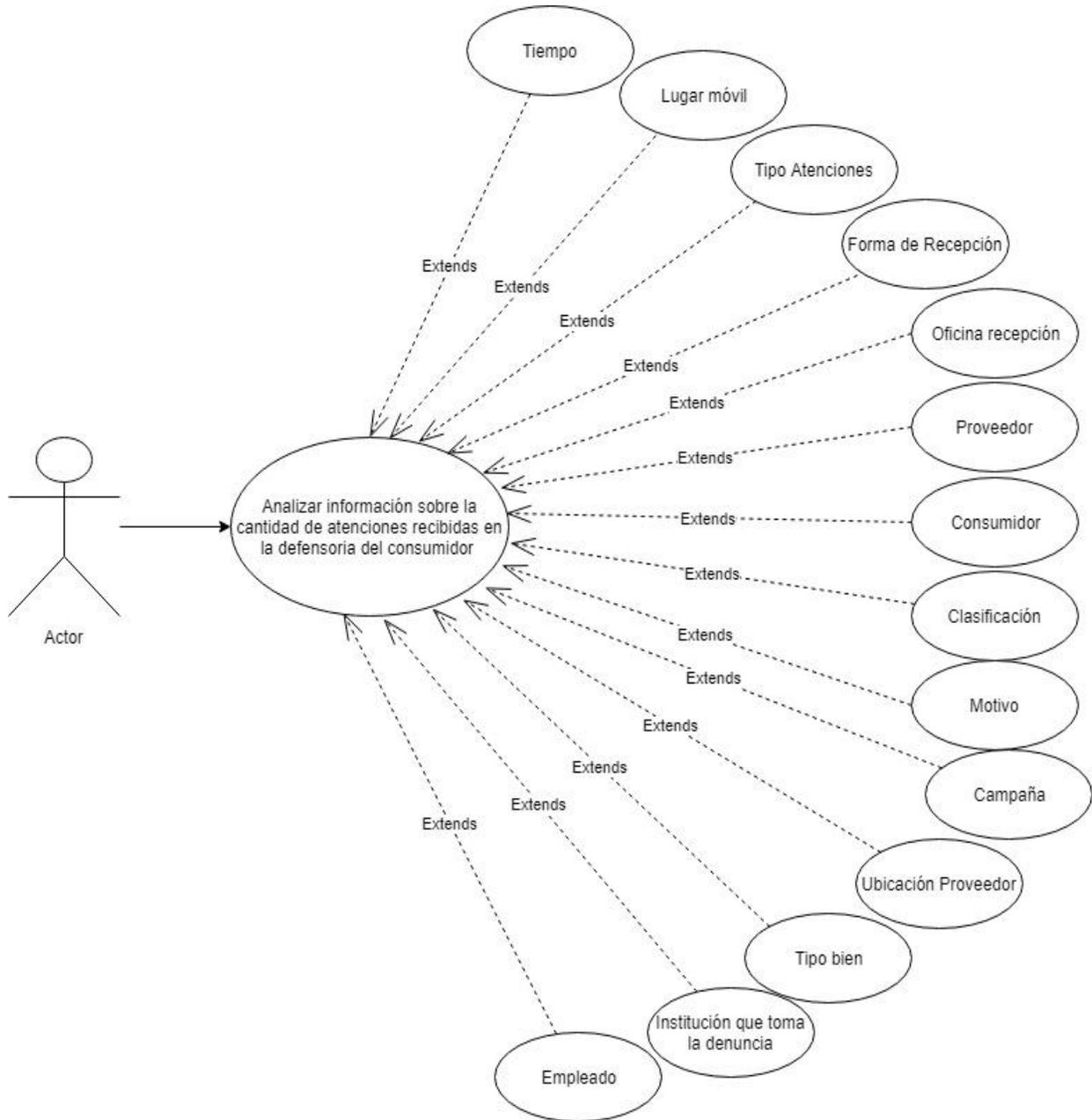


Figura 15.5: Caso de Uso de cantidad de atenciones recibida

15.4.2 Caso de uso cantidad de casos cerrados

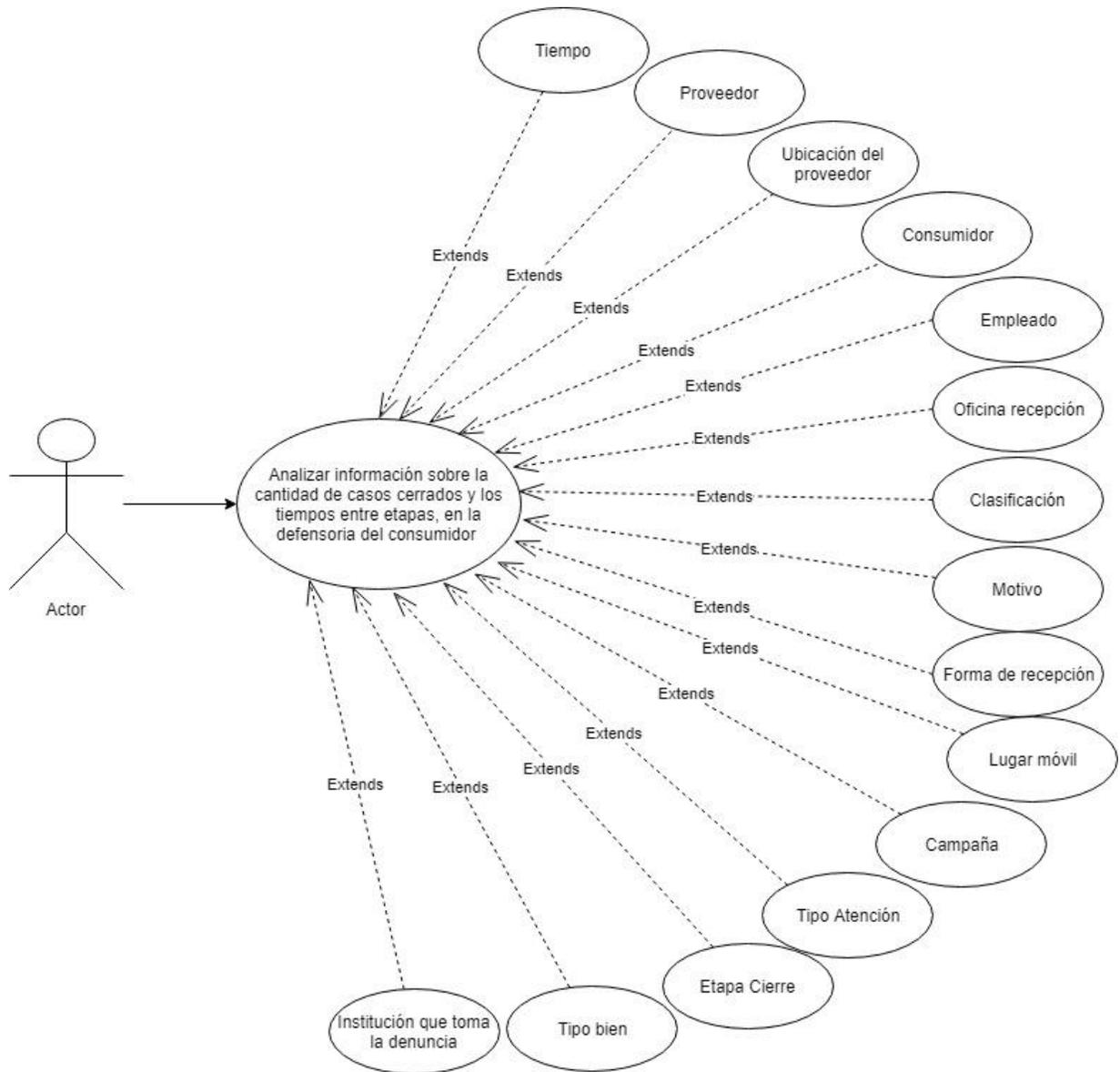


Figura 15.6: Caso de Uso de Cantidad de casos cerrados

15.4.3 Caso de uso de cantidades de casos reaperturados

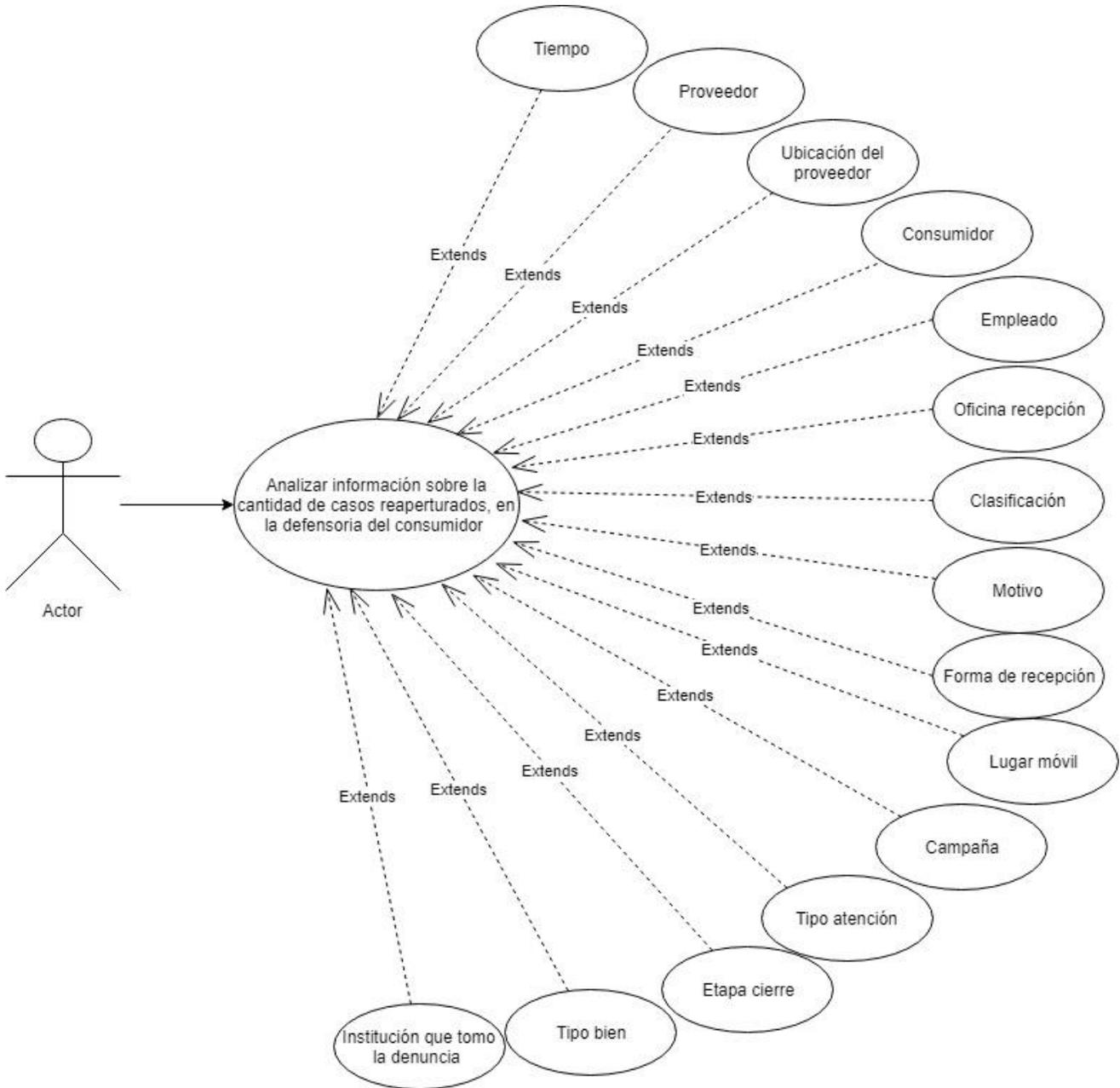


Figura 15.7: Caso de Uso de Cantidad de casos reaperturados

15.5 Diseño Multidimensional

15.5.1 Diseño Conceptual del Data Mart

15.5.1.1 Nivel 1: Definición del modelo

Unidad: Centro de Solución de Controversias

Especialización: Procesar atenciones realizadas

Hechos principales: Cantidad de atenciones recibidas, Cantidad de casos cerrados y cantidad de casos reaperturados

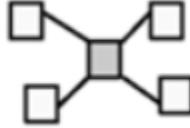


Figura 15.8: Nivel 1 Diagrama Conceptual Data Mart Centro de solución de Controversias

15.5.1.2 Nivel 2: Definición del esquema estrella

Atenciones Recibidas:

Paquete de Hecho: Atenciones recibidas

Paquetes Dimensionales

- | | | |
|-----------------|---------------------|-------------------|
| 1. Oficina | 15. Consumidor | 16. Proveedor |
| 4. Ubicación | 5. Tipo bien | 6. Empleado |
| 7. Institución | 8. Campana | 9. Tiempo |
| 10. Lugar Movil | 11. Forma recepción | 12. Tipo atención |
| 13. Motivo | 14. Clasificación | |

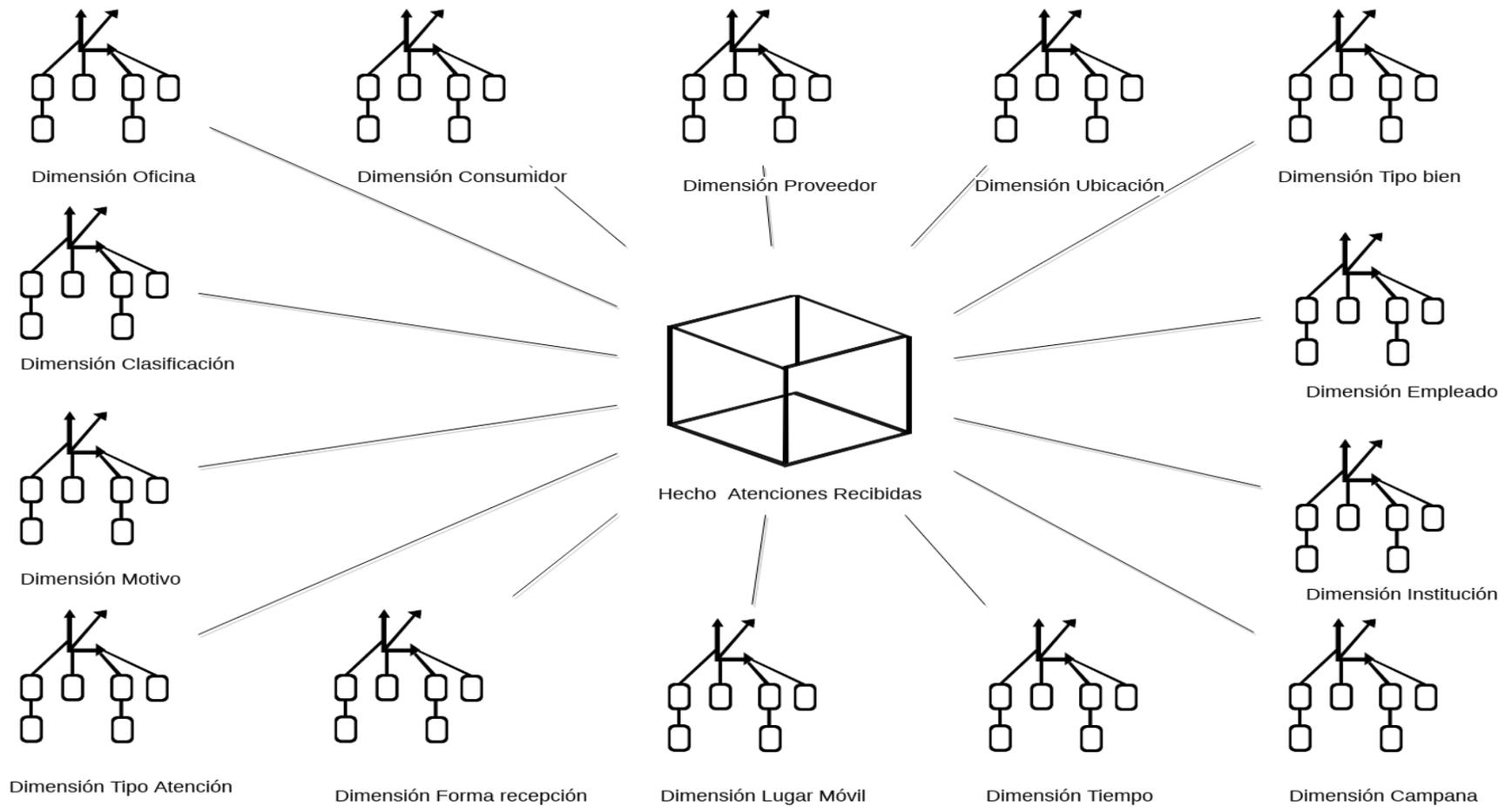


Figura 15.9: Nivel 2 Diagrama Conceptual Data Mart (Atenciones recibidas)- Centro de Solución de Controversias

Casos cerrados:

Paquete de Hecho: Casos cerrados

Paquetes Dimensionales

- | | | |
|-------------------|-----------------|------------------------|
| 1. Oficina | 2. Consumidor | 3. Proveedor |
| 4. Ubicación | 5. Tipo bien | 6. Campana |
| 7. Institución | 8. Empleado | 9. Etapa |
| 10. Tiempo | 11. Lugar móvil | 12. Forma de recepción |
| 13. Tipo atención | 14. Motivo | 15. Clasificación |

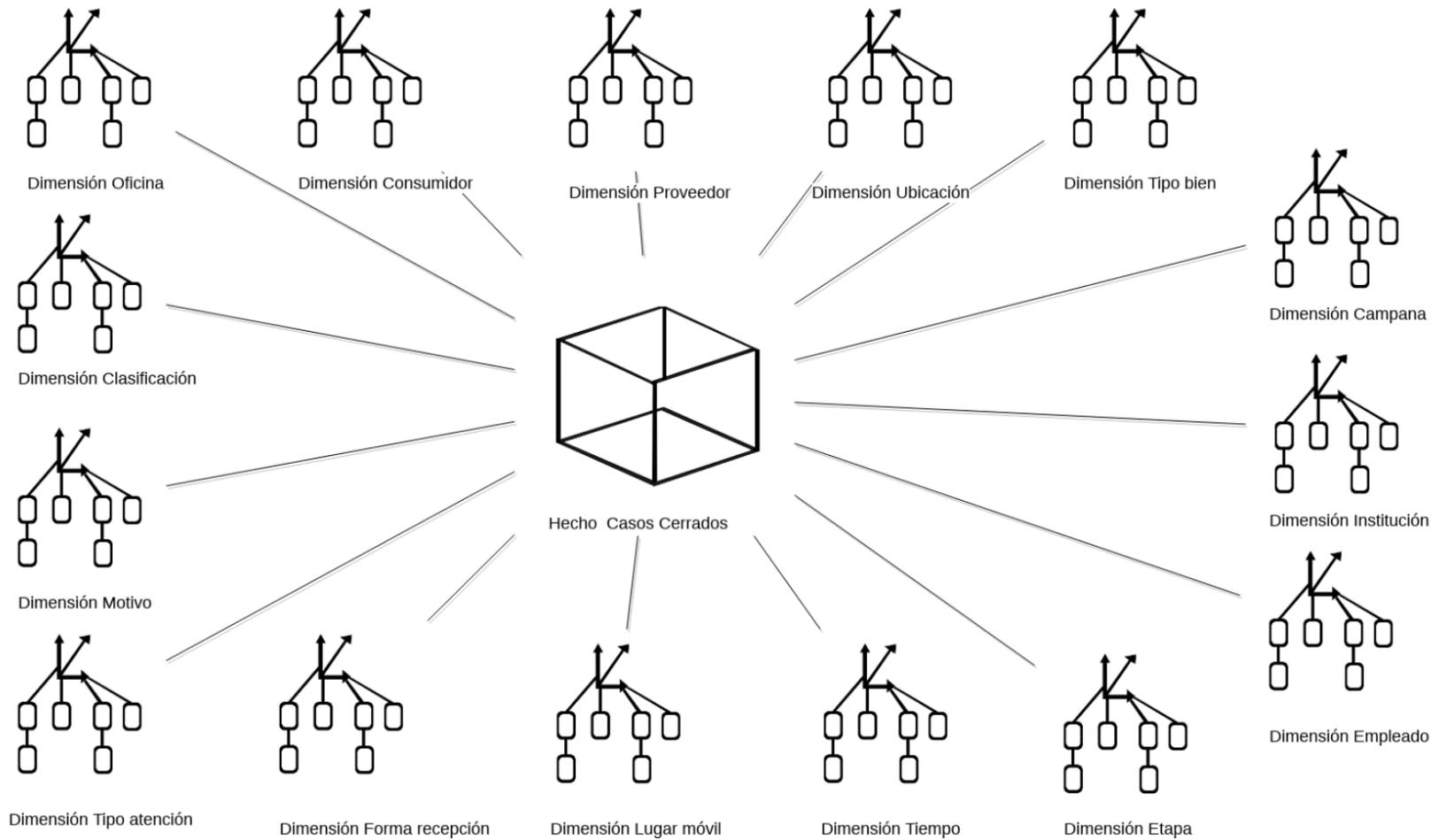


Figura 15.10: : Nivel 2 Diagrama Conceptual Data Mart (Casos cerrados)-Centro de Solución de Controversias

Casos reaperturados:

Paquete de Hecho: Casos reaperturados

Paquetes Dimensionales

- | | | | | | |
|-----|---------------|-----|-------------|-----|--------------------|
| 1. | Oficina | 2. | Consumidor | 3. | Proveedor |
| 4. | Ubicación | 5. | Tipo bien | 6. | Campana |
| 7. | Institución | 8. | Empleado | 9. | Etapa |
| 10. | Tiempo | 11. | Lugar móvil | 12. | Forma de recepción |
| 13. | Tipo atención | 14. | Motivo | 15. | Clasificación |

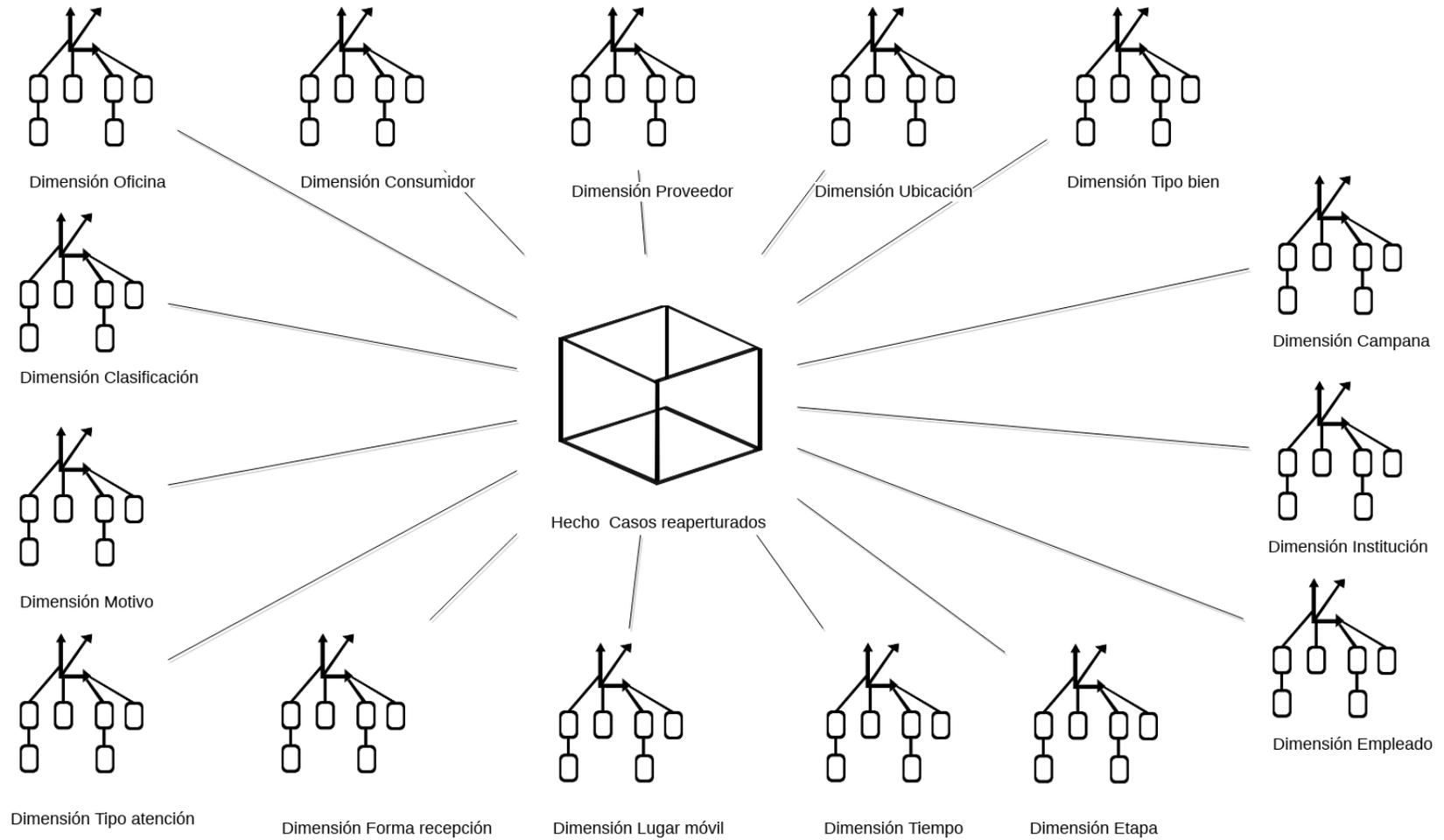


Figura 15.11: Nivel 2 Diagrama Conceptual Data Mart (Casos reaperturados)-Centro de Solución de Controversias.

15.5.1.3 Nivel 3: Dimensiones/Hechos

15.5.1.3.1 Atenciones recibidas

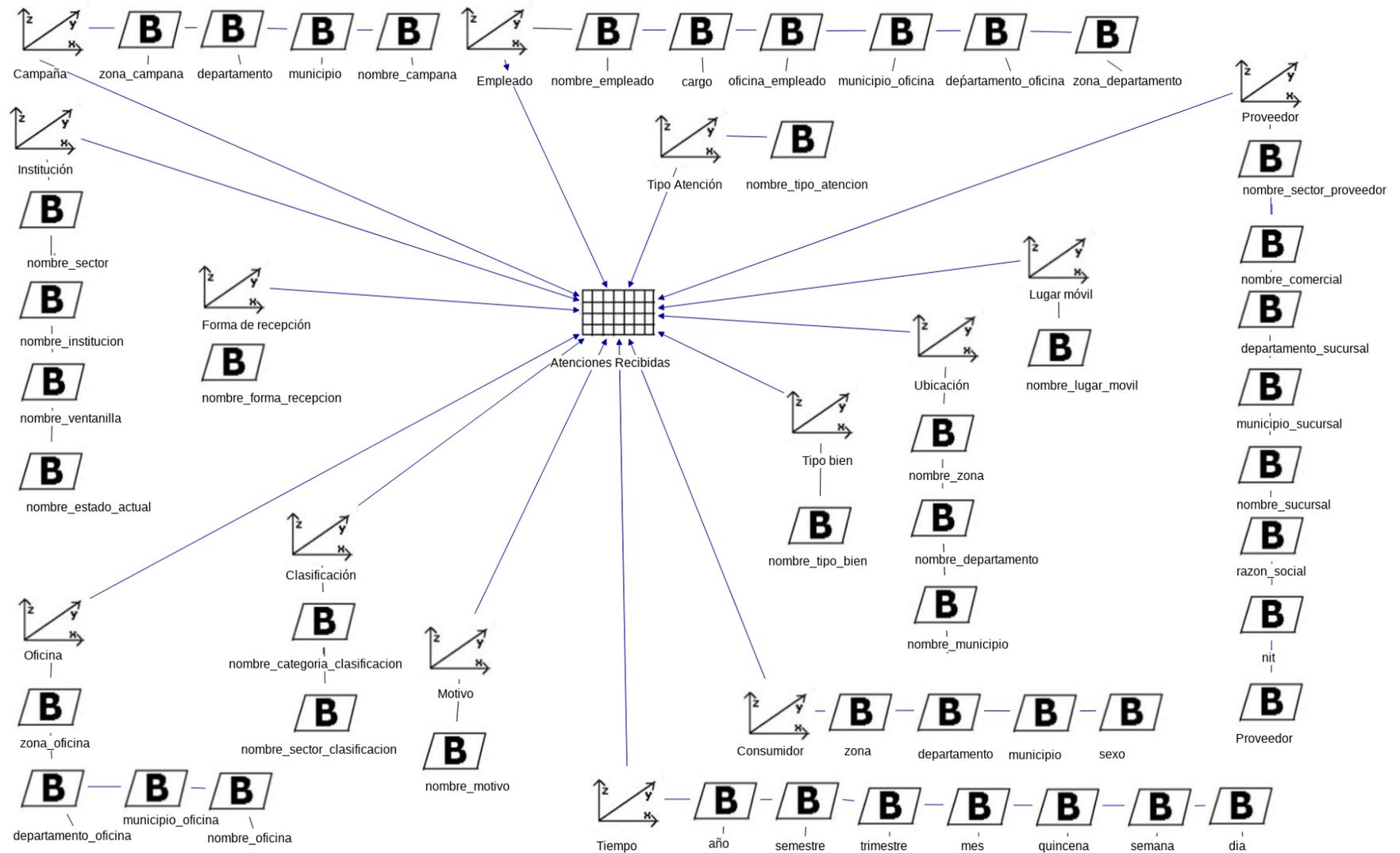


Figura 15.12: Nivel 3 Diagrama conceptual Data Mart (Hechos: Atenciones recibidas)

15.5.1.3.2 Casos cerrados

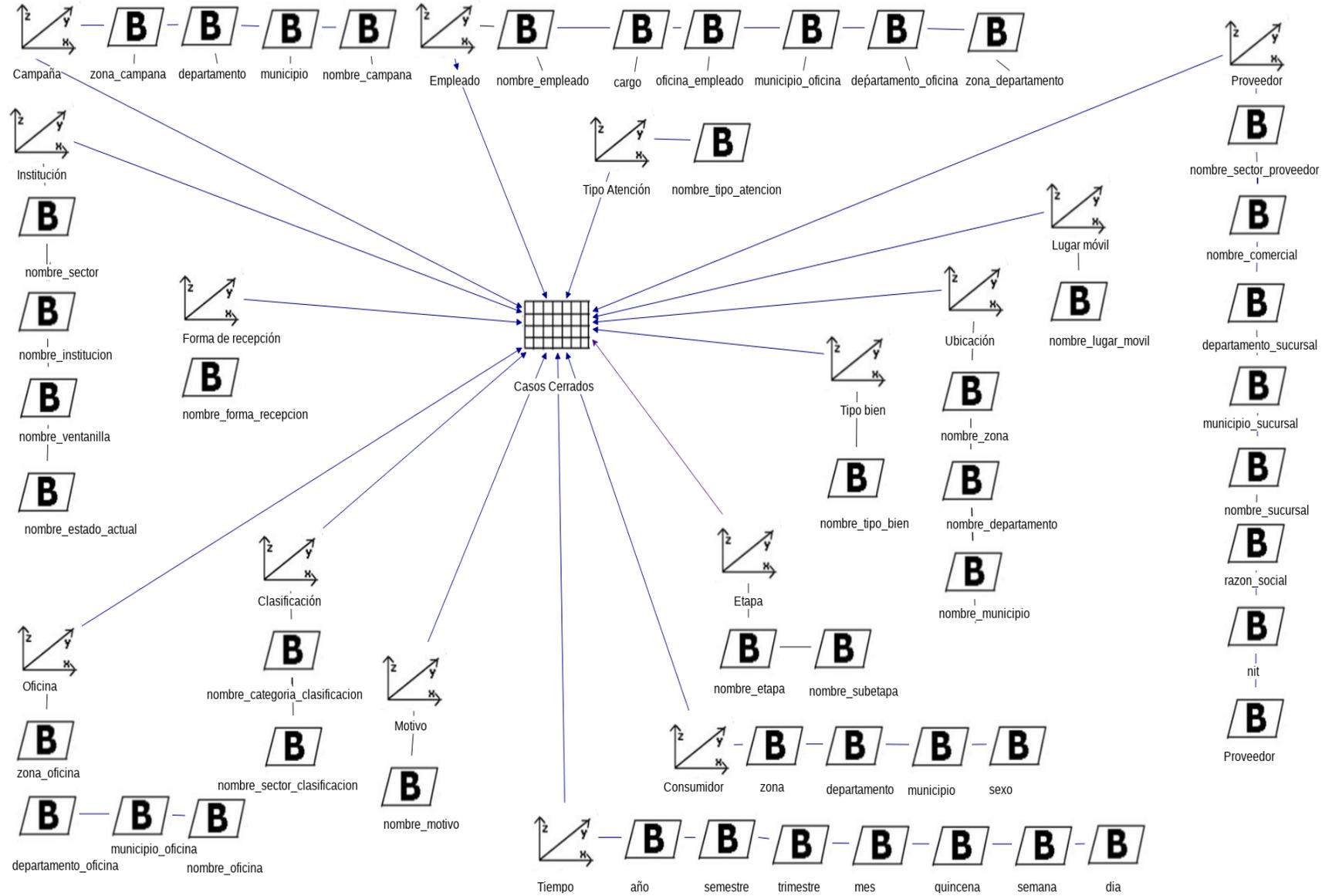


Figura 15.13: Diagrama conceptual Data Mart (Hecho: casos cerrados)

15.5.1.3.3 Casos reaperturadas

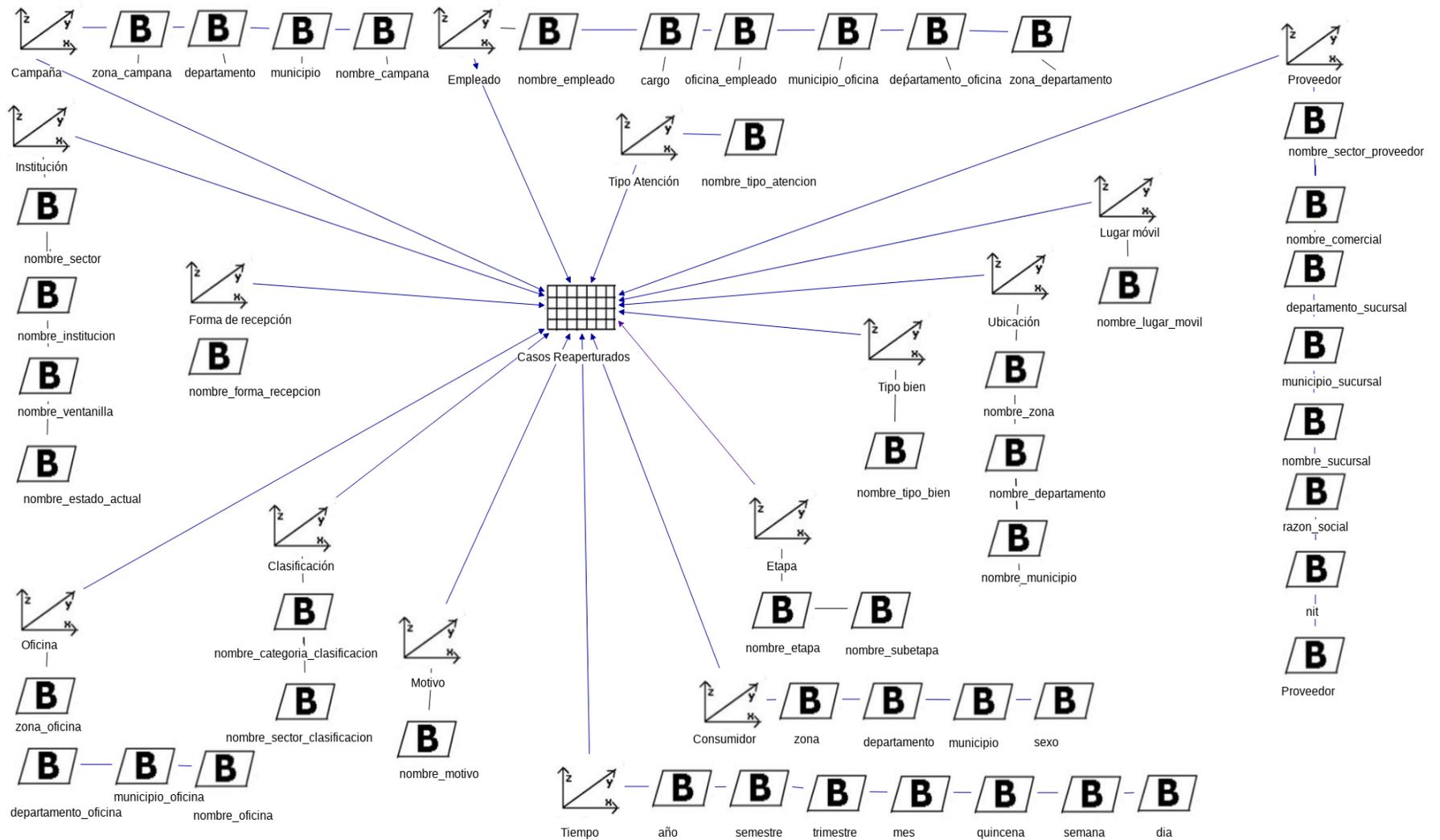


Figura 15.14: Diagrama conceptual Data Mart (Hecho: Casos reaperturados)

15.5.2 Diseño Físico Base de Datos Data mart

Dado que se tienen 3 tablas de hechos (fact tables) se han desarrollado se ha realizado un diseño para cada una de ellas. La Figura 15.15 muestra el diseño físico para la fact table de atenciones recibidas, la Figura 15.16 muestra el diseño físico para la fact table casos cerrados, y la Figura 15.17 muestra el diseño físico de base de datos para la fact table de casos reaperturados.

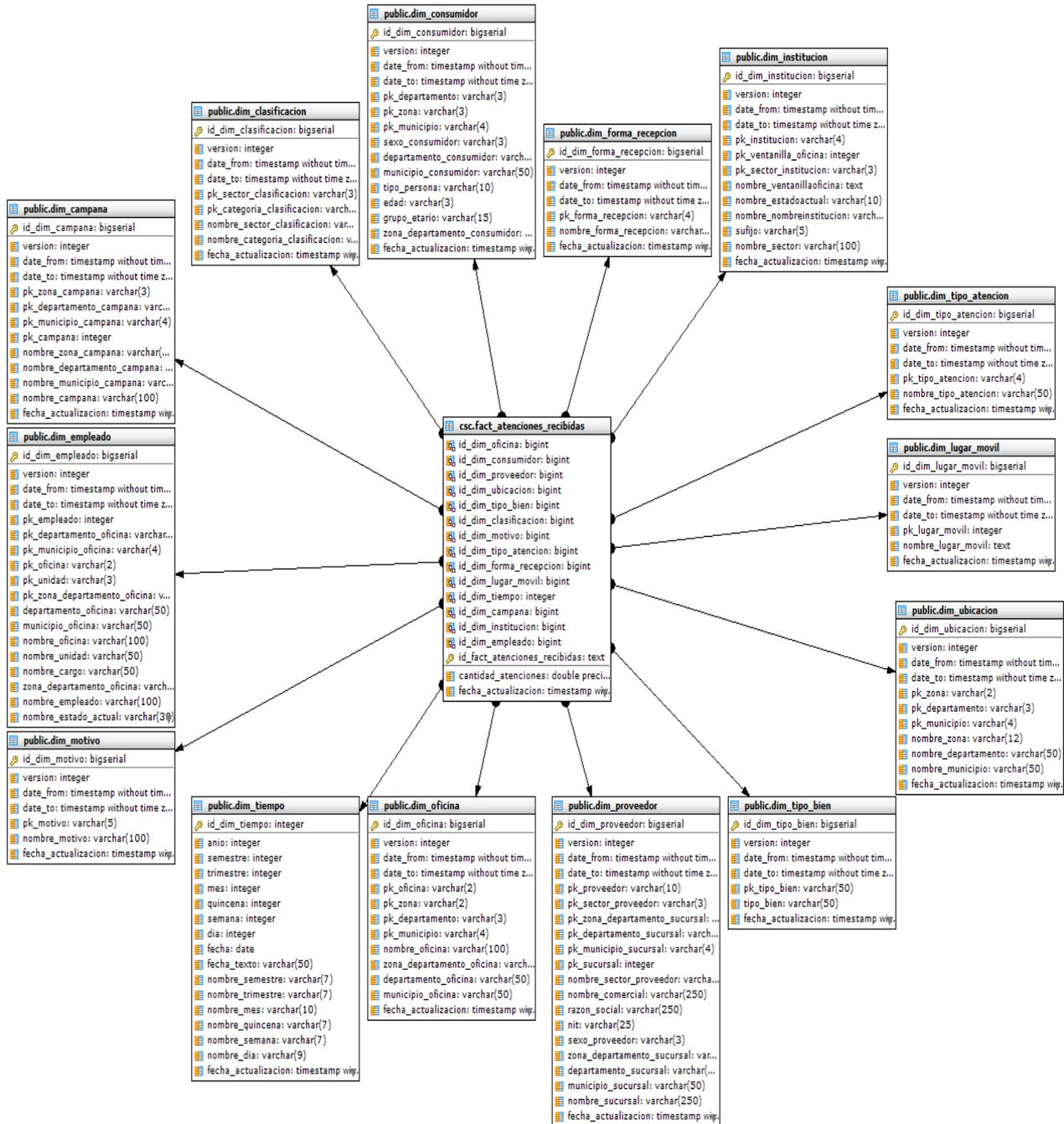


Figura 15.15: Diseño Físico para fact table Atenciones Recibidas, Centro de Solución de Controversias

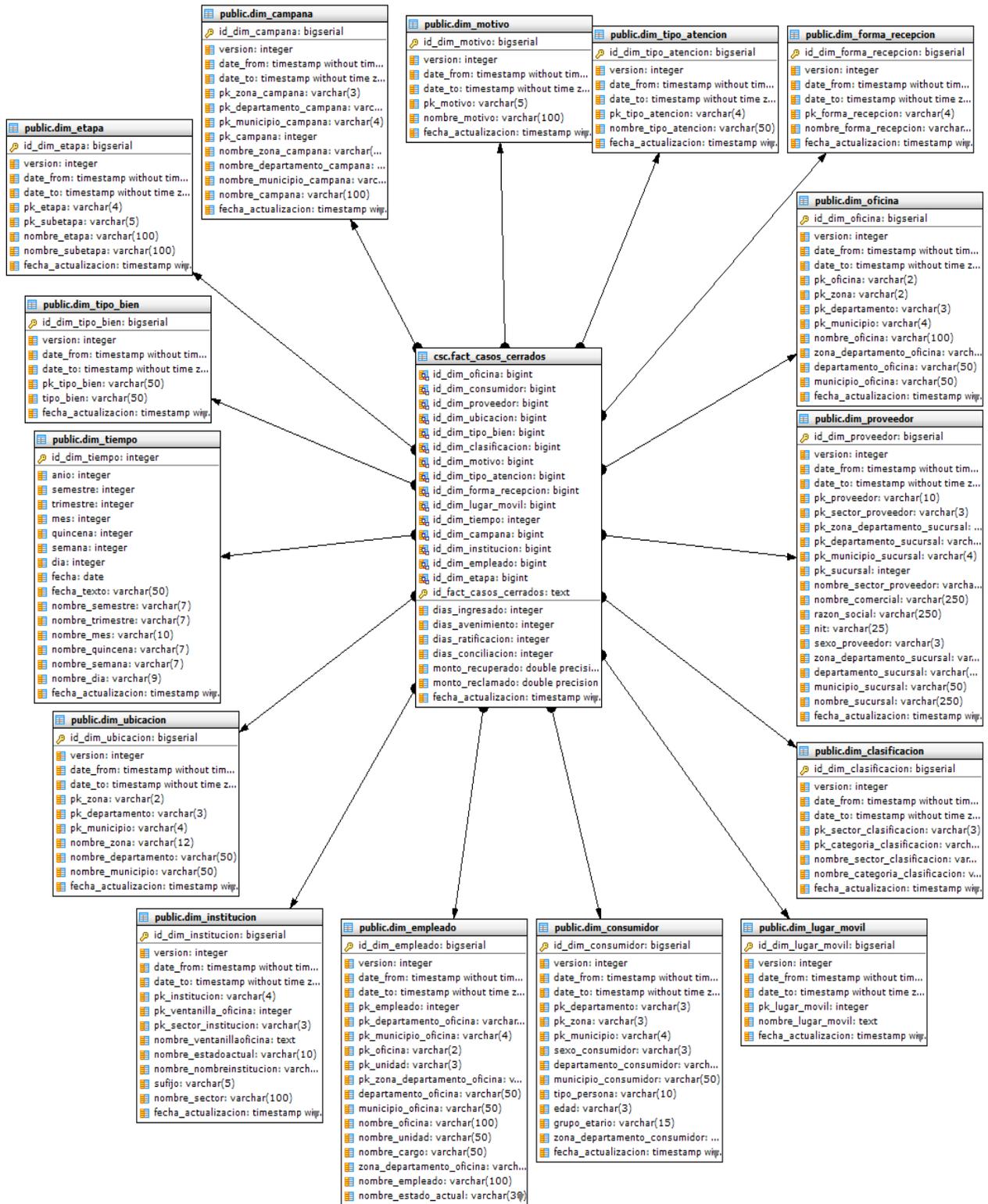


Figura 15.16: Diseño físico de base de datos, fact table Casos Cerrados, Centro de Solución de Controversias

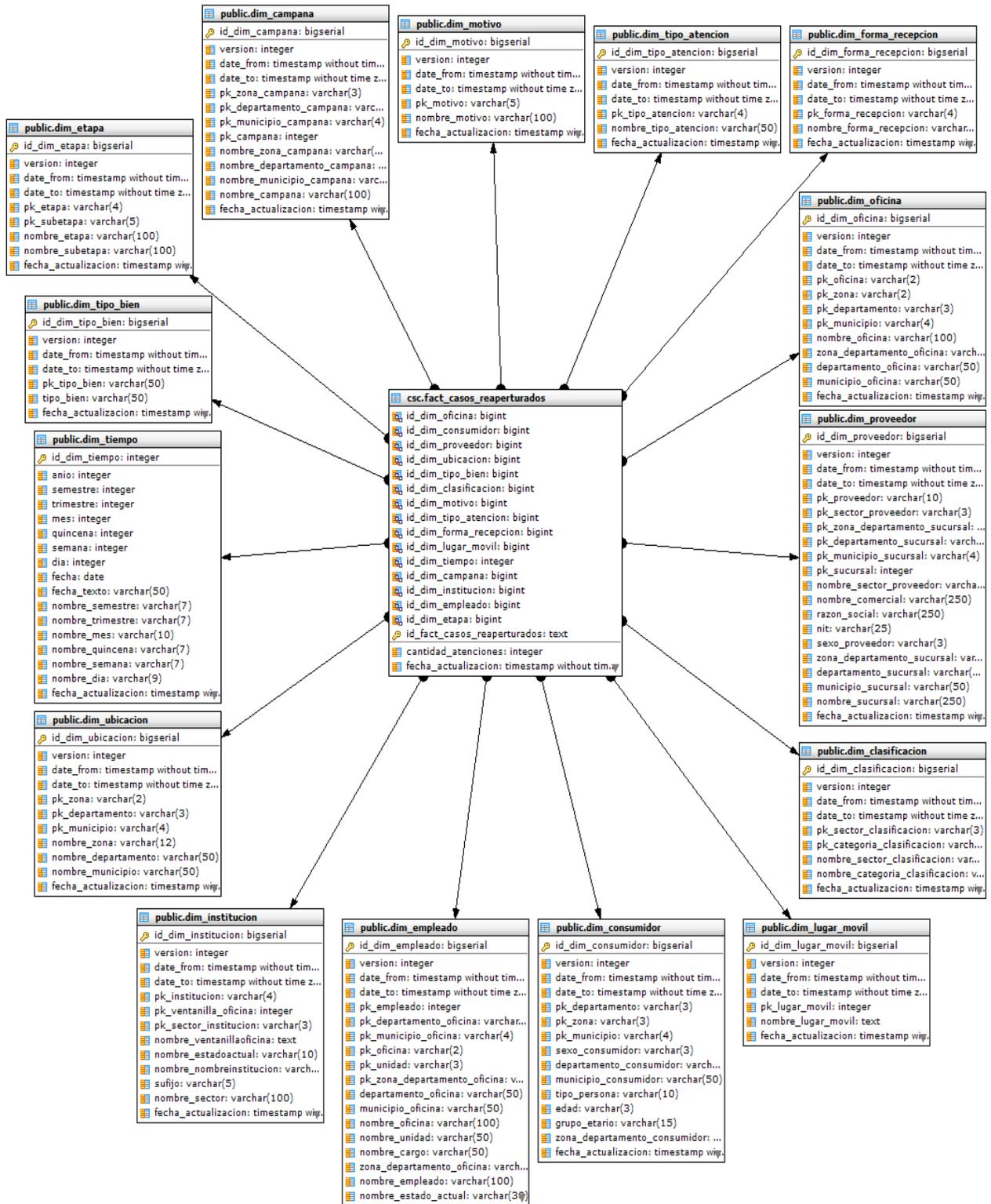


Figura 15.17: Diseño físico de base de datos, fact table Casos Reaperturados, Centro de Solución de Controversias

15.5.3 Diseño Físico Base de Datos Staging Area

A continuación, se muestra el diseño físico de base de datos del staging área para el data mart del Centro de Solución de Controversias. Las tablas se han colocado en un solo diagrama

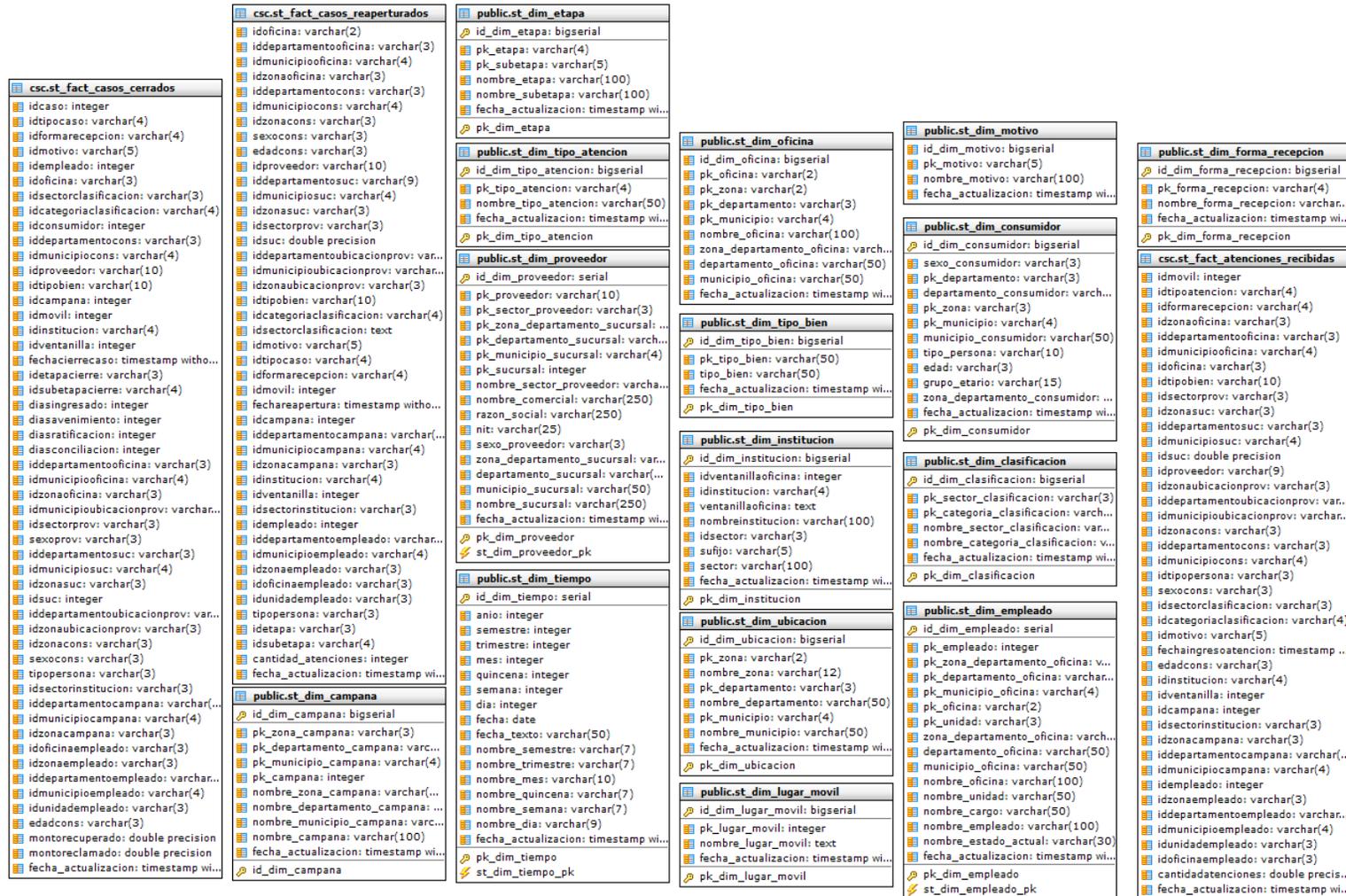


Figura 15.18: Diseño físico de base de datos, Staging Centro de Solución de Controversia

15.6 Diseño de Procesos ETL

15.6.1 Procesos ETL Staging

En los siguientes apartados se listan los procesos ETL de Staging, seguido de un diagrama representativo para esta unidad y los diseños de los procesos restantes se colocará en el CD que acompaña al documento en la carpeta `CD-EISI\Diseños de Procesos ETL\ csc (Centro de Solucion de Controversias)`

Diagramas desarrollados:

St_dim_institución

St_dim_campana

St_dim_etapa

St_Fact_atenciones_recibidas

St_fact_casos_cerrados

St_fact_casos_reaperturados

15.6.2 Procesos ETL Data Mart

A continuación, se listan los procesos ETL que mueven datos del Staging Area hacia el Data Mart (Procesos ETL Data Mart) desarrollados para esta unidad, seguido de un ejemplo (diagrama) para ilustrar uno de los procesos ETL.

Diagramas desarrollados:

- Dw_dim_institución
- Dw_dim_campana
- Dw_dim_etapa
- Dw_Fact_atenciones_recibidas
- Dw_fact_casos_cerrados
- Dw_fact_casos_reaperturados

A continuación, se ilustra uno de los diseños de procesos ETL (Figura 15.19), el cual corresponde al proceso ETL Dw_dim_etapa

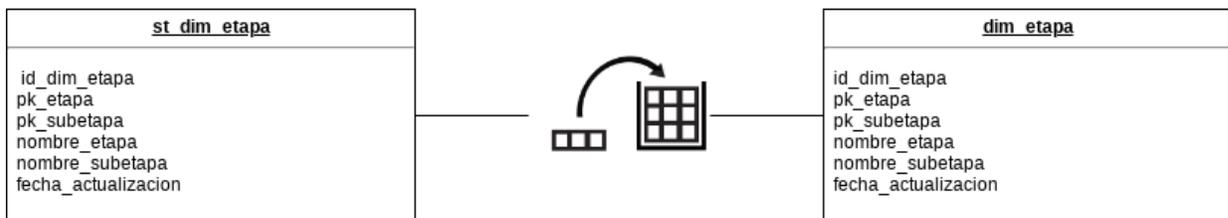


Figura 15.19: Proceso ETL de la dimensión Etapa (Data warehouse)

15.7 Pruebas

En las siguientes secciones se presenta las pruebas para cubos, dashboards y procesos ETL de la Dirección del Centro de Solución de Controversias.

15.7.1 Pruebas de Procesos ETL.

La realización de las pruebas consistió en ejecutar los Jobs almacenados en el servidor de Staging Área. Las transformaciones finales para el Data Mart de Centro de Solución de Controversias se muestran en la Tabla 15.2.

No	Transformación	Descripción
1	dw_incremental_fact_atenciones_recibidas.ktr	Transformación para cargas incrementales desde tabla st_fact_atenciones_recibidas (staging area) hacia tabla fact_atenciones_recibidas (data warehouse).
2	dw_incremental_fact_casos_cerrados.ktr	Transformación para cargas incrementales desde tabla st_fact_casos_cerrados (staging area) hacia tabla fact_casos_cerrados (data warehouse).

3	dw_incremental_fact_casos_reaperturados.ktr	Transformación para cargas incrementales desde tabla st_fact_casos_reaperturados (staging area) hacia tabla fact_casos_reaperturados (data warehouse).
4	dw_inicial_fact_atenciones_recibidas.ktr	Transformación para carga inicial desde tabla st_fact_atenciones_recibidas (staging area) hacia tabla fact_atenciones_recibidas (data warehouse).
5	dw_inicial_fact_casos_cerrados.ktr	Transformación para carga inicial del data warehouse desde tabla st_fact_casos_cerrados (staging area) hacia tabla fact_casos_cerrados (data warehouse).
6	dw_inicial_fact_casos_reaperturados.ktr	Transformación para carga inicial del data warehouse desde tabla st_fact_casos_reaperturados (staging area) hacia tabla fact_casos_reaperturados (data warehouse).
7	dw_unico_dim_campana.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tabla st_dim_campana (staging area) hacia tabla dim_campana (data warehouse).
8	dw_unico_dim_etapa.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tabla st_dim_etapa (staging area) hacia tabla dim_etapa (data warehouse).
9	dw_unico_dim_institucion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tabla st_dim_institucion (staging area) hacia tabla dim_institucion (data warehouse).
10	st_incremental_fact_atenciones_recibidas.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia tabla st_fact_atenciones_recibidas (staging area).
11	st_incremental_fact_casos_cerrados.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia tabla st_fact_casos_cerrados (staging area).
12	st_incremental_fact_casos_reaperturados.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia tabla st_fact_casos_reaperturados (staging area).
13	st_inicial_fact_atenciones_recibidas.ktr	Transformación para carga inicial del data warehouse desde tablas fuentes hacia tabla st_fact_atenciones_recibidas (staging area).
14	st_inicial_fact_casos_cerrados.ktr	Transformación para carga inicial del data warehouse desde tablas fuentes hacia tabla st_fact_casos_cerrados (staging area).
15	st_inicial_fact_casos_reaperturados.ktr	Transformación para carga inicial del data warehouse desde tablas fuentes hacia tabla st_fact_casos_reaperturados (staging area).
16	st_unico_dim_campana.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia tabla st_dim_campana (staging area).
17	st_unico_dim_etapa.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia tabla st_dim_etapa (staging area).

18	st_unico_dim_institucion.ktr	Transformación para carga unica o full refresh (inicial e incremental) del data warehouse desde tablas fuentes hacia tabla st_dim_institucion (staging area).
----	------------------------------	---

Tabla 15.2: Transformaciones para Data Mart Centro de Solución de Controversias

Los Jobs finales de Primer Nivel para Data Mart de Centro de Solución de Controversias se muestran la Tabla 15.3. También muestra el orden en que ejecutan las transformaciones de cada uno de los Jobs.

No	Job Primer Nivel	Transformación	
		No	Nombre Transformación
1	job_unico_dim_campana.kjb	1	st_unico_dim_campana.ktr
		2	dw_unico_dim_campana.ktr
2	job_unico_dim_etapa.kjb	1	st_unico_dim_etapa.ktr
		2	dw_unico_dim_etapa.ktr
3	job_unico_dim_institucion.kjb	1	st_unico_dim_institucion.ktr
		2	dw_unico_dim_institucion.ktr
4	job_inicial_fact_casos_reaperturados.kjb	1	st_inicial_fact_casos_reaperturados.ktr
		2	dw_inicial_fact_casos_reaperturados.ktr
5	job_inicial_fact_casos_cerrados.kjb	1	st_inicial_fact_casos_cerrados.ktr
		2	dw_inicial_fact_casos_cerrados.ktr
6	job_inicial_fact_atenciones_recibidas.kjb	1	st_inicial_fact_atenciones_recibidas.ktr
		2	dw_inicial_fact_atenciones_recibidas.ktr

Tabla 15.3: Jobs de Nivel 1 para Data Mart Centro de Solución de Controversias

Los Jobs finales de Segundo Nivel para Data Mart de Centro de Solución de Controversias se muestran la Tabla 15.4. Dicha tabla también muestra el orden en que ejecutan las transformaciones de cada uno de los Jobs. El orden en que aparecen los jobs en la tabla siguiente, es el orden de ejecución que se debe seguir.

No	Jobs Segundo Nivel	Jobs primer nivel ejecutados	
		No	Nombre Job
1	job_inicial_csc.kjb	1	job_unico_dim_campana.kjb
		2	job_unico_dim_etapa.kjb
		3	job_unico_dim_institucion.kjb
		4	job_inicial_fact_casos_reaperturados.kjb
		5	job_inicial_fact_casos_cerrados.kjb
		6	job_inicial_fact_atenciones_recibidas.kjb
2	job_incremental_csc.kjb	1	job_unico_dim_campana.kjb
		2	job_unico_dim_etapa.kjb
		3	job_unico_dim_institucion.kjb
		4	job_incremental_fact_casos_reaperturados.kjb
		5	job_incremental_fact_casos_cerrados.kjb
		6	job_incremental_fact_atenciones_recibidas.kjb

Tabla 15.4: Jobs de Nivel 2 para Data Mart Centro de Solución de Controversias

La Figura 15.20 y Figura 15.21 muestra los Jobs de segundo nivel, los cuales son Jobs principales que son ejecutados en el servidor. La figura Figura 15.20 muestra la secuencia del job que realiza la carga

inicial, mientras que la Figura 15.21 muestra la secuencia del job que servirá para realizar las cargas incrementales.

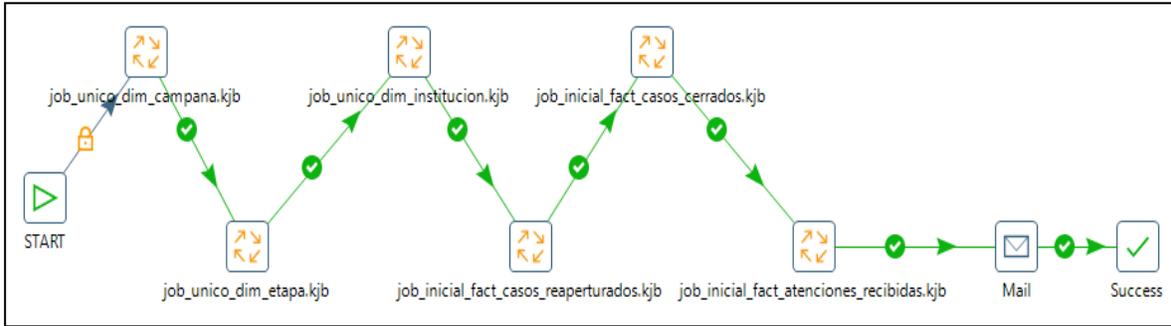


Figura 15.20: Job ETL para la carga inicial del Data Mart Centro de Solución de controversias

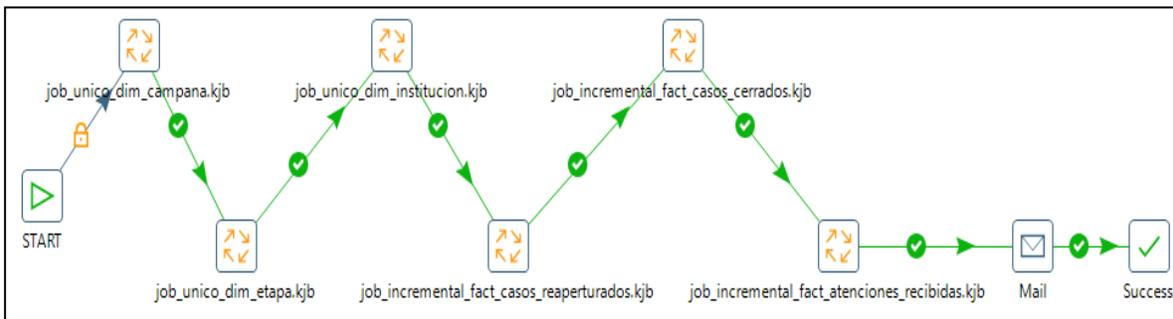


Figura 15.21: Job ETL para la carga incremental del Data Mart Centro de Solución de controversias

15.7.1.1 Ejecución de Procesos ETL Iniciales

El comando para ejecutar el proceso ETL se muestra a continuación (No se debe cerrar la ventana de comandos hasta la finalización de ejecución de procesos ETL):

```
#/root/data-integration/kitchen.sh -file=/root/jobs/csc/job_inicial_csc.kjb -level=Debug
```

El tiempo estimado para la ejecución de los procesos ETL es el siguiente:

Cantidad de transformaciones	Tiempo	Registros
18	0:14:11	365,327

La finalización del proceso ETL se notifica mediante correo electrónico como se muestra en la Figura 15.22

```

Previous results:
-----
Job entry Nr      : 6
Errors           : 0
Lines read       : 0
Lines written    : 0
Lines input      : 0
Lines output     : 0
Lines updated    : 0
Lines rejected   : 0
Script exist status : 0
Result          : true

Path to this job entry:
-----
job_inicial_csc
job_inicial_csc : : start : Start of job execution (2019/03/17 22:28:00.203)
job_inicial_csc : : START : start : Start of job execution (2019/03/17 22:28:00.209)
job_inicial_csc : : START : [nr=0, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 22:28:00.215)
job_inicial_csc : : job_unico_dim_campana.kjb : Followed unconditional link : Start of job execution (2019/03/17 22:28:00.220)
job_unico_dim_campana
job_unico_dim_campana : : START : Start of job entry : Start of job execution (2019/03/17 22:28:00.710)
job_unico_dim_campana : : START : [nr=1, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 22:28:00.711)
job_unico_dim_campana : : st_unico_dim_campana.ktr : Followed unconditional link : Start of job execution (2019/03/17 22:28:00.715)
    
```

Figura 15.22: Contenido del correo de notificación de finalización del proceso ETL (se muestra parte del contenido)

15.7.2 Pruebas Cubos OLAP

Dado que para Centro de Solución de Controversias se han identificado tres hechos que se desean analizar, se han realizado tres cubos para dicha unidad. A continuación, se muestran los cubos realizados para el Data Mart de Centro de Solución de controversias.



Figura 15.23: Cubos para Centro de Solución de Controversias

Por cada uno de ellos se ha publicado un archivo XML en el Servidor de Pentaho Server BI.

Con el uso del Plugin JPivot para Pentaho Server se logra visualizar la información según se muestra en la Figura 15.24, Figura 15.25 y Figura 15.26.

Clasificación	Motivo	Tiempo	Ubicación proveedor	Empleado	Campaña	Tipo bien	Institución	Cantidad atenciones recibidas
Todas las clasificaciones	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	133.665
Agua Potable	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	4.264
Agua potable	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	4
Agua potable	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	32.962
Alimentos	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	1
Alimentos y bebidas	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	374
Artículos del hogar	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	1
Artículos infantiles	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	2
Comercio	Todos los motivos	Todos los años	Todas las ubicaciones	Todos los empleados	Todas las campañas	Tipos de bien	Instituciones	11.372
				Central	Todas las campañas	Tipos de bien	Instituciones	7.843
						N/A	Instituciones	3.999
						Producto	Instituciones	3.037
						Servicio	Instituciones	807
				Occidental	Todas las campañas	Tipos de bien	Instituciones	1.686

Figura 15.24: Visualización del Cubo Atenciones Recibidas (Centro de Solución de Controversias)

Etapa cierre	Tipo bien	Institución	Medidas						
			Dias ingresado	Dias en avenimiento	Dias en ratificación	Dias en conciliación	cantidad de casos	Monto Recuperado	Monto reclamado
Todas las etapas	Tipos de bien	Instituciones	11,599	120,893	6,394	37,935	58.686	12.150.836,03	247.358.633,6
Todas las etapas	Tipos de bien	Instituciones	94,845	1.113,752	39,479	279,235	4.084	889.794,29	2.371.645,5
Avenimiento	Tipos de bien	Instituciones	83,486	1.414,956	0,077	0	3.208	457.683,71	1.277.133,5
Conciliación	Tipos de bien	Instituciones	91,114	11,907	2,042	1.551,556	735	432.110,58	1.019.824,2
Ingresado	Tipos de bien	Instituciones	1.799,125	0	0	0	24	0	28.260,8
Prevención	Tipos de bien	Instituciones	49,217	0	0	0	23	0	12.427,2
Ratificación	Tipos de bien	Instituciones	87,734	6,723	1.696,628	0	94	0	33.999,6
Todas las etapas	Tipos de bien	Instituciones	7,694	106,342	7,033	57,855	11.089	2.618.453,85	6.361.637,8
Todas las etapas	Tipos de bien	Instituciones	8,157	45,339	2,993	10,797	16.547	2.431.573,93	14.751.509,5
Todas las etapas	Tipos de bien	Instituciones	5,164	30,817	1,342	5,207	6.420	1.419.136,55	2.695.158,3
Todas las etapas	Tipos de bien	Instituciones	3,732	34,959	3,532	14,15	1.843	278.461,12	718.620,0
Todas las etapas	Tipos de bien	Instituciones	2,276	22,394	3,147	10,653	5.801	1.309.582,81	103.126.391,7
Todas las etapas	Tipos de bien	Instituciones	2,288	24,925	5,651	16,292	7.852	1.995.672,59	115.005.905,7
Todas las etapas	Tipos de bien	Instituciones	0,373	5,551	1,737	3,263	5.040	1.208.149,89	2.322.813,7

Figura 15.25: Visualización del Cubo Casos Cerrados (Centro de Solución de Controversias)

Motivo	Tiempo	Ubicacion proveedor	Empleado	Campana	Etapa cierre	Tipo bien	Institucion	Medidas
es <input type="checkbox"/> Todos los motivos	<input type="checkbox"/> Todos los años	<input type="checkbox"/> Todas las ubicaciones	<input type="checkbox"/> Todos los empleados	<input type="checkbox"/> Todas las campanas	<input type="checkbox"/> Todas las etapas	<input type="checkbox"/> Tipos de bien	<input type="checkbox"/> Instituciones	801
						Producto	<input type="checkbox"/> Instituciones	225
						Servicio	<input type="checkbox"/> Instituciones	576
			<input type="checkbox"/> Central	<input type="checkbox"/> Todas las campanas	<input type="checkbox"/> Todas las etapas	<input type="checkbox"/> Tipos de bien	<input type="checkbox"/> Instituciones	453
						Producto	<input type="checkbox"/> Instituciones	157
						Servicio	<input type="checkbox"/> Instituciones	296
			<input type="checkbox"/> Occidental	<input type="checkbox"/> Todas las campanas	<input type="checkbox"/> Todas las etapas	<input type="checkbox"/> Tipos de bien	<input type="checkbox"/> Instituciones	319
						Producto	<input type="checkbox"/> Instituciones	56
						Servicio	<input type="checkbox"/> Instituciones	263
			<input type="checkbox"/> Oriental	<input type="checkbox"/> Todas las campanas	<input type="checkbox"/> Todas las etapas	<input type="checkbox"/> Tipos de bien	<input type="checkbox"/> Instituciones	29
						Producto	<input type="checkbox"/> Instituciones	12
						Servicio	<input type="checkbox"/> Instituciones	17
	<input type="checkbox"/> 2008	<input type="checkbox"/> Todas las ubicaciones	<input type="checkbox"/> Todos los empleados	<input type="checkbox"/> Todas las campanas	<input type="checkbox"/> Todas las etapas	<input type="checkbox"/> Tipos de bien	<input type="checkbox"/> Instituciones	37

Transfiriendo datos desde 34.73.27.65...

Figura 15.26: Visualización del Cubo Casos Reaperturados (Centro de Solución de Controversias)

15.7.3 Pruebas Dashboard

Uno de los Dashboards desarrollados para la Dirección del Centro de Solución de Controversias se muestra en la siguiente página.

15.7.3.1 Dashboard Centro de Solución de Controversias

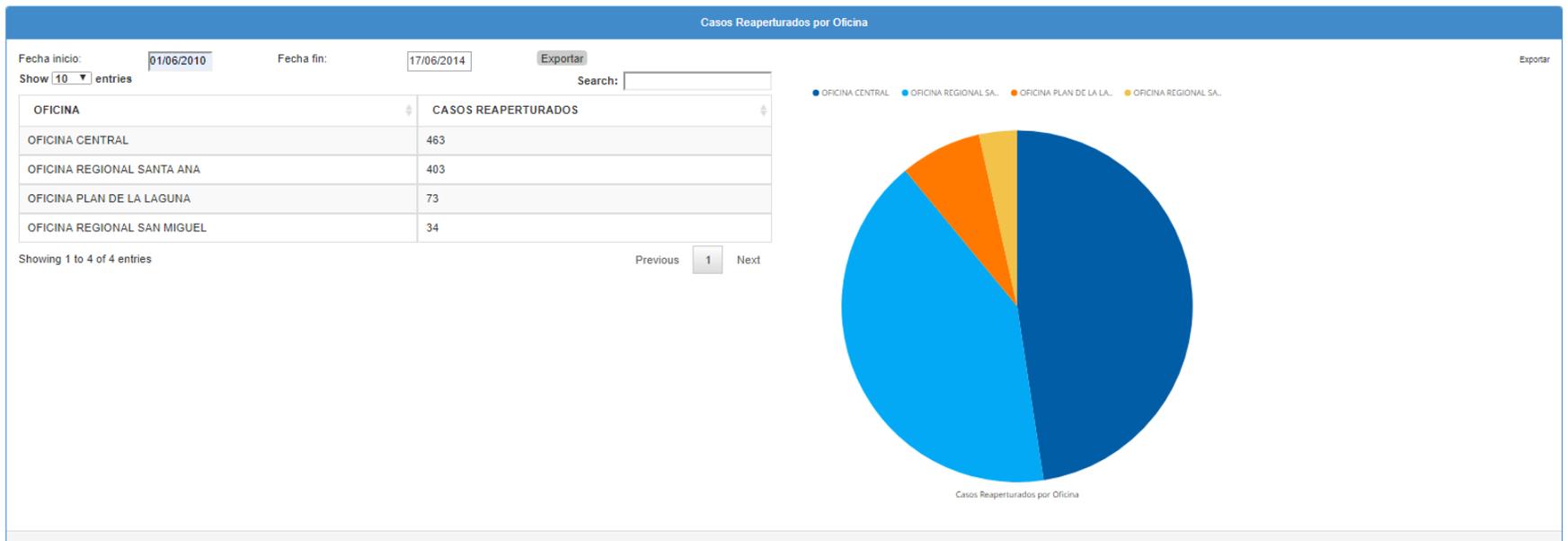
Casos Reaperturados por Proveedor

Fecha inicio: Fecha fin: [Exportar](#) Exportar

Show entries Search:

PROVEEDOR	CASOS REAPERTURADOS
ADMINISTRACION NACIONAL DE ACUEDUCTOS Y ALCANTARILLADOS	290
ANDA	71
TELEMOVIL EL SALVADOR, S.A.	49
CTE TELECOM	46
BANCO CITIBANK DE EL SALVADOR S.A.	29
JONATHAN ALBERTO FLORES GUANDIQUE	20
CTE TELECOM PERSONAL, S.A. DE C.V.	20
BANCO AGRICOLA	16
GMG SERVICIOS EL SALVADOR, S.A. DE C.V.	13
AVAL CARD	13

Showing 1 to 10 of 246 entries Previous 2 3 4 5 ... 25 Next



16 SPRINT 5

16.1 Sprint Backlog

A continuación, se presenta el Sprint Backlog, que refleja las historias de usuarios, su estimación y asignación para este Segundo Sprint.

#	Historia	Punto Estimado	Asignado
1.	Redacción de la plantilla para el refinamiento de requerimientos	20	Todos
2.	Poblar la base de datos de desarrollo SARA	13	Ivis
3.	Diseño del modelo multidimensional	8	Samuel – Ivis
4.	Diseño BD del Staging Area	3	Jorge – Samuel
5.	Diseño del proceso ETL	8	Todos
6.	Diseño de vistas	3	Ivis-Samuel -Esau
7.	Creación de Script SQL del Staging Area	13	Jorge – Ivis – Samuel
8.	Creación de Script SQL modelo multidimensional	13	Jorge – Ivis – Samuel
9.	Desarrollo de los procesos ETL	100	Jorge – Ivis – Samuel
10.	Desarrollo de cubos OLAP	40	Jorge – Ivis – Samuel
11.	Desarrollo de vistas	13	Ivis – Samuel-Esau
12.	Implementación e integración de componentes	40	Jorge – Ivis – Samuel
13.	Pruebas	20	Jorge – Ivis – Samuel
14.	Documentación	40	Todos

Tabla 16.1: Sprint Backlog, Sprint 5

16.2 Descripción de Historias de Usuario

A continuación, se presenta la descripción de las historias de usuario detallando: el responsable, al sprint al que pertenece, una breve descripción, sus criterios de validación y los puntos de estimación. Previamente a las historias de usuario se presenta la historia épica del Sprint 5

<i>Historia Épica #6</i>	Data Mart Dirección de descentralización.
Responsable	Jorge Iraheta/Esau Serpas/Ivis Navarro/Samuel Núñez Sprint: 5
Descripción	Como Dirección de Descentralización Necesito: Un mecanismo para analizar la información de atenciones recibas y las atenciones finalizadas. Para: la evaluación de resultados de dichas atenciones por la Defensoría en diferentes temporadas o períodos de tiempo.
Criterios de aceptación	- La cantidad de atenciones recibidas deben poderser visualizar por las siguientes perspectivas: <ul style="list-style-type: none"> • Por fecha • Por lugar móvil • Por tipo atención • Por forma de recepción • Por oficina de recepción • Por proveedor • Por consumidor • Por clasificación (Sector y Categoría) • Por motivo • Por campaña • Por ubicación del proveedor • Por tipo bien

	<ul style="list-style-type: none"> • Por institución • Por empleado • Por etapa <p>- La cantidad de casos cerrados junto al monto recuperado, deben poder analizarse desde las siguientes perspectivas:</p> <ul style="list-style-type: none"> • Por fecha • Por proveedor • Por ubicación del proveedor • Por consumidor • Por empleado • Por oficina de recepción • Por clasificación (Sector y Categoría) • Por motivo • Por forma de recepción • Por lugar móvil • Por campaña • Por tipo atención • Por etapa Cierre • Por tipo bien • Por institución que toma la denuncia.
Puntos estimados	334
Historia de usuario #1	Refinamiento del requerimiento de información
Responsable	Ivis Navarro Sprint: 5
Descripción	<p>Como Equipo Scrum necesito el refinamiento del requerimiento de información de la Dirección de descentralización: Realizar reuniones con los interesados y especificar los elementos de información requeridos para la Dirección de descentralización.</p> <ul style="list-style-type: none"> • Llenar la herramienta de recolección. • Revisar y desglosar los requerimientos solicitados. • Firmar documento por el usuario como aceptación de los de los requerimientos. <p>Para tener claridad de la información que desean almacenar, visualizar y analizar mediante el Data Mart o modelo Multidimensional para ésta unidad.</p>
Criterios de aceptación	<ul style="list-style-type: none"> - Tener llena la Plantilla de refinamiento de requerimientos de información. - Incluir únicamente requerimientos de información con naturaleza histórica y no información transaccional. - Tener un documento firmado de aceptación de requerimientos. - Identificar (mediante un listado) las tablas fuentes que servirán para cumplir con los requerimientos de información solicitados.
Puntos estimados	20
Historia de usuario #2	Poblar base de datos de desarrollo SARA
Responsable	Samuel Núñez Sprint: 5
Descripción	<p>Como Product Owner necesito: migrar los datos necesarios según las tablas fuentes de datos que se identifiquen durante el proceso de toma de requerimientos. Para poseer datos de prueba para el desarrollo del Data Mart.</p>
Criterios de aceptación	<ul style="list-style-type: none"> - Se deben poseer Scripts de INSERT para las tablas Fuente. - Todas las tablas fuentes identificadas en el refinamiento de requerimientos deben poseer datos. - Debe existir al menos una prueba de conexión y acceso remoto exitoso a los datos. - Datos cargados en la Base de datos

Puntos estimados	13
Historia de usuario #3	Diseño del modelo multidimensional
Responsable	Samuel Núñez/Ivis Navarro Sprint: 5
Descripción	Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de la Dirección de descentralización. Para tener la base (diagrama) para el desarrollo de Scripts de Base de Datos SQL para Data Mart.
Criterios de aceptación	-Esquema del modelo multidimensional de la Dirección de descentralización Finalizado. -El esquema debe cumplir las reglas de un modelo de tipo Estrella. -El esquema debe poseer las relaciones necesarias para la búsqueda según los criterios de búsqueda definidos por la Dirección de descentralización. -El esquema debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
Puntos estimados	8

Historia de usuario #4	Diseño BD del Staging Area.
Responsable	Jorge Iraheta/Samuel Núñez Sprint: 5
Descripción	Como Equipo Scrum Necesito: el diseño del modelo multidimensional para el Data Mart de la Dirección de Descentralización. Para tener el insumo para la creación de Scripts de Base de Datos SQL para Staging Area.
Criterios de aceptación	-Deberá existir una o más tablas por cada dimensión. -Deberá existir una o más tablas de hecho. -El diseño debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
Puntos estimados	3

Historia de usuario #5	Diseño del proceso ETL
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez/Esau Serpas Sprint: 5
Descripción	Como Equipo Scrum necesito: Diseñar los Proceso ETL <ul style="list-style-type: none"> • Extracción • Transformación • Carga Para cumplir con el Diseño del proceso ETL y tenerlo como insumo a la hora de desarrollar dichos procesos ETL.
Criterios de aceptación	-Se deberá utilizar nomenclatura UML para los pasos en cada proceso ETL. -Deberá incluir todas las tablas dimensiones y tablas de hechos. -Los nombres de las tablas o etiquetas deben poderse visualizar fácilmente.
Puntos estimados	8

Historia de usuario #6	Diseño de vistas
Responsable	Ivis Navarro/Esau Serpas/Samuel Nuñez Sprint: 5
Descripción	Como Administrador de Data Warehouse necesito: Realizar el diseño de vistas necesarias para visualizar la información del Data Mart de la Dirección de descentralización. Para Diseñar y digitalizar la vista de presentación de datos.
Criterios de aceptación	-Deberá existir un diseño de vista o dashboard para cada hecho a medir. -Deberá explicarse cada sección del dashboard o vista. -Etiquetas y textos deben ser legibles a simple vista.
Puntos estimados	3

<i>Historia de usuario #7</i>	Creación de Script SQL del Staging Area
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 5
Descripción	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será procesada, calculada y/o limpiada la información.
Criterios de aceptación	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Staging Area. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
Puntos estimados	13

<i>Historia de usuario #8</i>	Creación de Script SQL Modelo Multidimensional
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 5
Descripción	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será almacenada la información.
Criterios de aceptación	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Modelo Multidimensional. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
Puntos estimados	13

<i>Historia de usuario #9</i>	Desarrollo de los procesos ETL.
Responsable	Jorge Iraheta/ Ivis Navarro/Samuel Nuñez Sprint: 5
Descripción	Como Equipo Scrum necesito: Desarrollar los procesos ETL correspondientes Para obtener, trasladar y transformar los datos que serán almacenados en el Data Mart de la Dirección de descentralización.
Criterios de aceptación	-Los nombres de los procesos ETL deben poseer la nomenclatura definida por el equipo. -Los procesos ETL deben poseer la documentación necesaria en forma de comentarios dentro de la herramienta PDI. -Las Transformaciones ETL deben haber sido probados en desarrollo (Equipo Local) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de final . -Las transformaciones correspondientes a una dimensión o hecho deben ser incorporados un Job de PDI. -Deberá crearse un Job que incorpore los otros Jobs correspondientes a cada dimensión y hecho. -Los Jobs de PDI deben haber sido probados con las bases de datos de prueba y base de datos final. -Las bitácoras de los procesos ETL deben almacenarse en una base de datos.
Puntos estimados	100

<i>Historia de usuario #10</i>	Desarrollo de Cubos OLAP
Responsable	Jorge Iraheta/ Ivis Navarro/Samuel Nuñez Sprint: 5
Descripción	Como Equipo Scrum necesito: investigar sobre la herramienta para el desarrollo de cubos OLAP e implementar dicha herramienta para la creación de cubos OLAP. Para implementar el cubo correspondiente a la dirección de descentralización
Criterios de aceptación	-El cubo deberá poderse visualizar correctamente y sin errores. -Deberá existir como mínimo un cubo OLAP por Hecho. -Deberán poderse realizar las operaciones como mínimo de: Roll Up y Drill Down.
Puntos estimados	40

<i>Historia de usuario #11</i>	Desarrollo de vistas
Responsable	Ivis Navarro/Samuel Nuñez/Esau Serpas Sprint: 5
Descripción	Como Equipo Scrum necesito: Desarrollar vistas o dashboards según el diseño previamente elaborado. Para Presentar información relevante hacia la unidad para la toma de decisiones.
Criterios de aceptación	-Vista o dashboard funcionando sin errores técnicos. -Debe coincidir con el diseño realizado previamente. -Debe existir al menos un dashboar por tabla de hechos.
Puntos estimados	13

<i>Historia de usuario #12</i>	Implementación e integración de componentes
Responsable	Jorge Iraheta/ Ivis Navarro/Samuel Nuñez Sprint: 5
Descripción	Como Equipo Scrum necesito: La integración de todos los elementos del modelo multidimensional <ul style="list-style-type: none"> - ETL - Fuentes - Stagin Área - BD multidimensional - Pentaho Server (Vistas) Para el desarrollo del modelo multidimensional de forma integral y probar el funcionamiento correctamente de la totalidad del Data Mart.
Criterios de aceptación	- Se deberán modificar los procesos ETL para que utilicen las bases de datos finales (no las de pruebas). - Se deberán realizar pruebas en la visualización de los dashboards. - Todo el entorno (desde los procesos ETLs, hasta los dashboards o vistas) deberán tener el visto bueno de todo el equipo.
Puntos estimados	40

<i>Historia de usuario #13</i>	Pruebas
Responsable	Jorge Iraheta/ Ivis Navarro/Samuel Nuñez Sprint: 5
Descripción	Como Equipo Scrum necesito: Realizar pruebas al modelo multidimensional terminado. Para verificar que funciona de la manera que se espera.
Criterios de aceptación	Pruebas exitosa en pruebas de aceptación del usuario.
Puntos estimados	20

<i>Historia de usuario #14</i>	Documentación
<i>Responsable</i>	Samuel Nuñez/Jorge Iraheta/Ivis Navarro/Esau Serpas Sprint: 2
<i>Descripción</i>	Como Equipo Scrum necesito: Documentación completa con historias, nomenclatura y diagramas empleados. Para verificar cómo se llevó a cabo el desarrollo de los cubos de la segunda unidad (Dirección de descentralización).
<i>Criterios de aceptación</i>	Documentos completos de análisis y diseño para modelo multidimensional
<i>Puntos estimados</i>	40

16.3 Refinamiento del Requerimiento de Información

Para la realización del refinamiento de los requerimientos de información se partió con la investigación del proceso general que sigue la Dirección de Descentralización dentro de la Defensoría del Consumidor. Además de establecer el proceso general de la unidad se hizo uso de los paquetes de información para refinar el requerimiento de información.

16.3.1 Proceso BPMN de la Unidad

A continuación, presentamos el proceso de Dirección de Descentralización expresado en BPNM

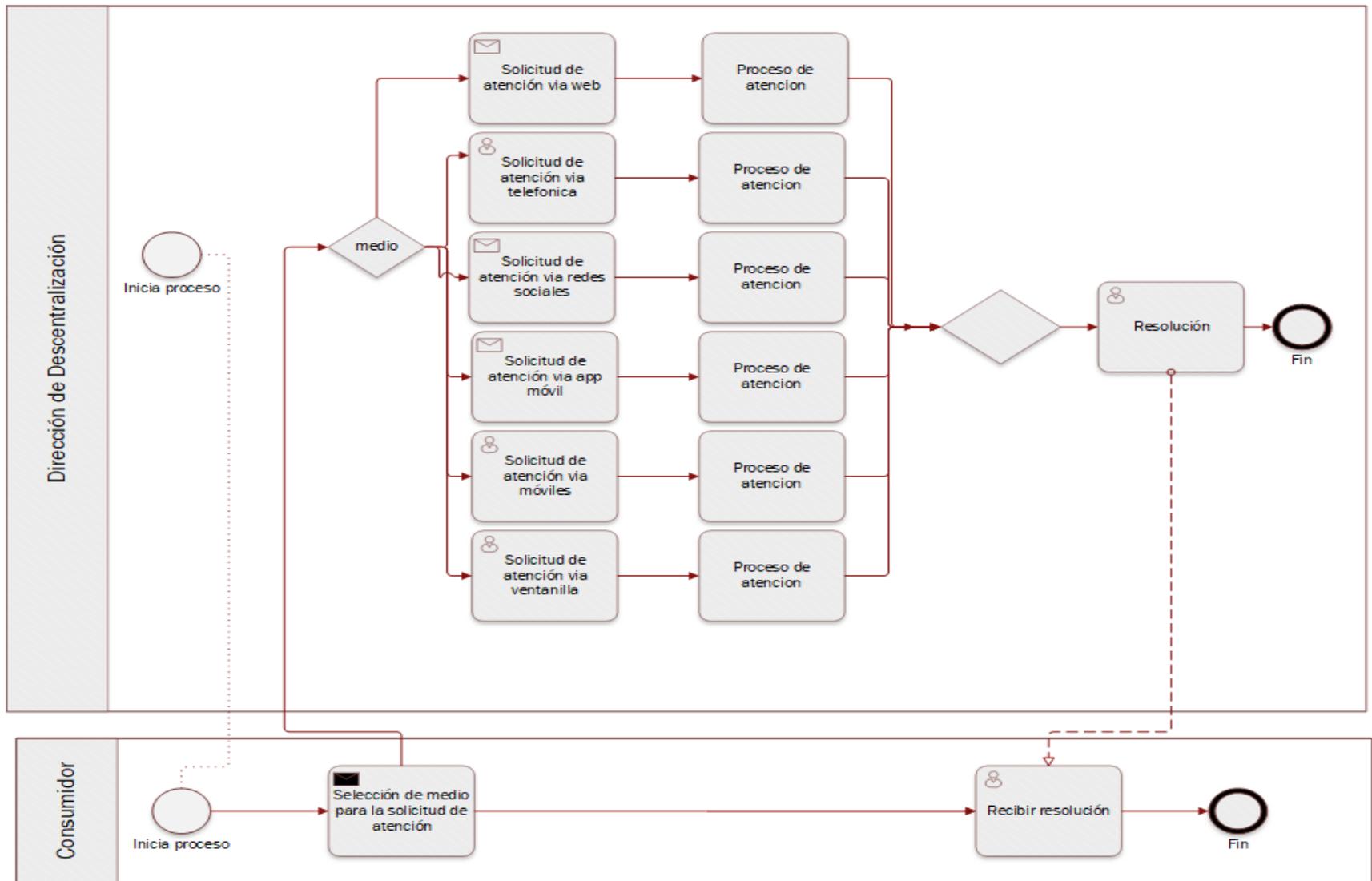


Figura 16.1: Proceso BPMN Dirección de Descentralización.

16.3.2 Paquetes de información

A continuación, se muestran los paquetes de información correspondiente a la Dirección de Descentralización:

Tema:		Atenciones recibidas en la Defensoría del consumidor.												
J E R A R Q U I N C E N A S E M A N A D I A	Tiempo	Lugar Movil	Tipo Atencion	Forma de Recepción	Oficina Recepción	Proveedor	Empleado	Consumidor	Clasificación	Motivo	ubicacion proveedor	Institucion que toma la denuncia	Campaña	Tipo Bien
	Año	Nombre Lugar	Nombre tipo atencion	Nombre forma recepcion	Zona	NombreSector	Zona Departamento	Sexo	Nombre sector	Nombre Motivo	Zona	Nombre institución	Zona campaña	Nombre Tipo Bien
	Semestre				Departamento	NombreComercial	Departamento de la Oficina		Zona	Nombre categoría		Departamento	Nombre sector	Departamento
	Trimestre				Municipio	RazonSocial	Municipio de la Oficina		Departamento			Municipio		Municipio
	Mes				Oficina	NIT	Oficina del Empleado	Municipio						Nombre campaña
	Quincena					DepartamentoSucursal	Cargo del Empleado							
	Semana					MunicipioSucursal	Nombre Empleado							
	Día					NombreSucursal								
	Hechos Medidos:	Cantidad de atenciones recibidas												

Figura 16.2: Paquete de información para cantidad de atenciones recibidas en la DC

Tema:		Casos finalizadas en la Defensoría del consumidor.												
J E R A R Q U I A S	Tiempo	Lugar Movil	Tipo Atencion	Forma de Recepción	Oficina Recepción	Proveedor	Empleado	Consumidor	Clasificación	Motivo	ubicacion proveedor	Institucion que toma la denuncia	Campaña	Tipo Bien
	Año	Nombre Lugar	Nombre tipo atencion	Nombre forma recepcion	Zona	NombreSector	Zona Departamento	Sexo	Nombre sector	Nombre Motivo	Zona	Nombre institución	Zona	Nombre Tipo Bien
	Semestre				Departamento	NombreComercial	Departamento de la Oficina	Zona	Nombre categoría		Departamento	Nombre sector	Departamento	
	Trimestre				Municipio	RazonSocial	Municipio de la Oficina	Departamento			Municipio		Municipio	
	Mes				Oficina	NIT	Oficina del Empleado	Municipio					Nombre campaña	
	Quincena					DepartamentoSucursal	Cargo del Empleado							
	Semana					MunicipioSucursal	Nombre Empleado							
	Día					NombreSucursal								
	Hechos Medidos:		Cantidad de Casos, Monto recuperado											

Figura 16.3: Paquete de información para cantidad de casos cerradas y monto recuperado en la DC

16.4 Especificación de Requerimientos

16.4.1 Caso de uso cantidad de atenciones recibidas

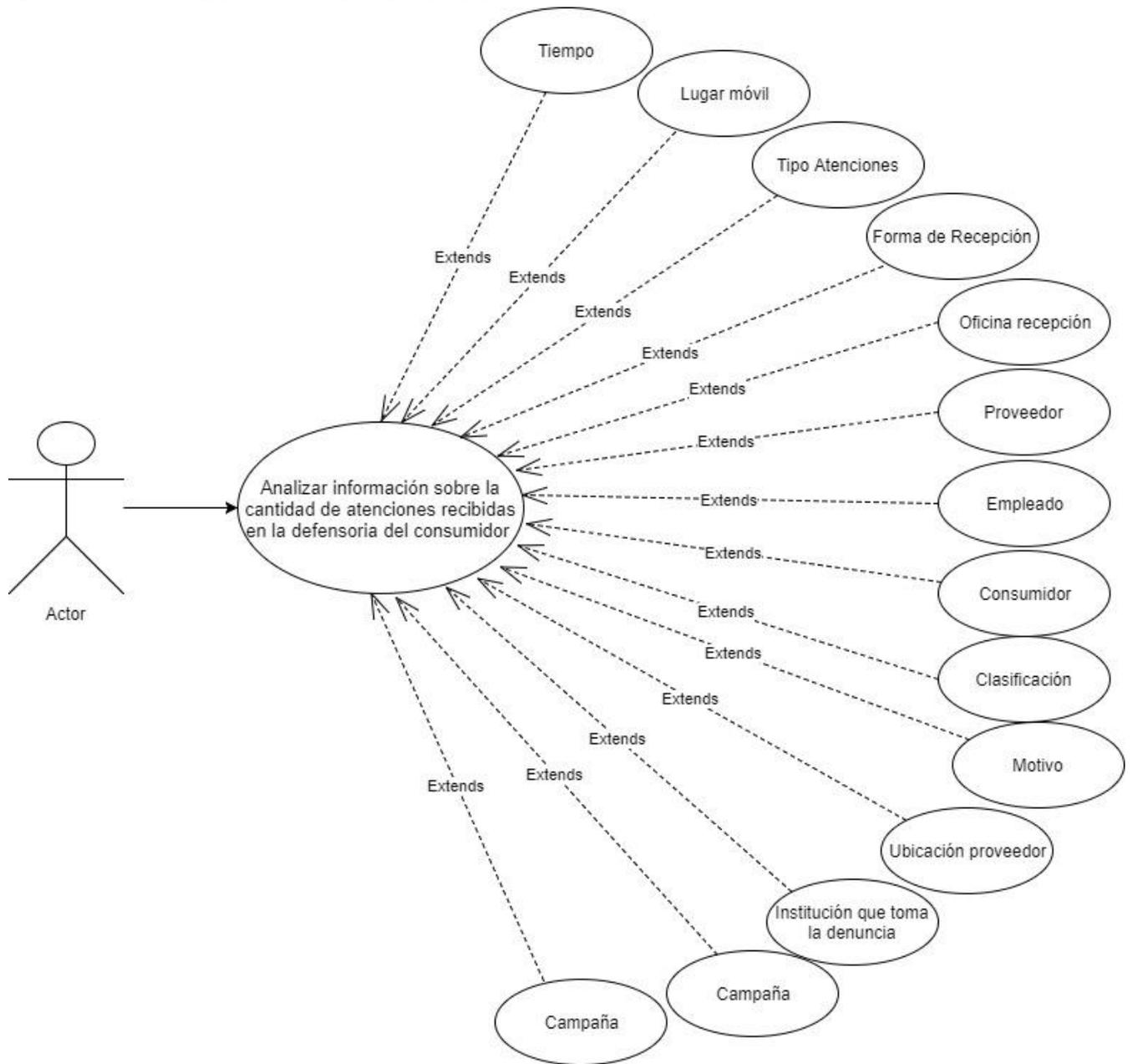


Figura 16.4: Caso de Uso de la cantidad de atenciones recibidas

16.4.2 Caso de uso cantidad de casos cerradas

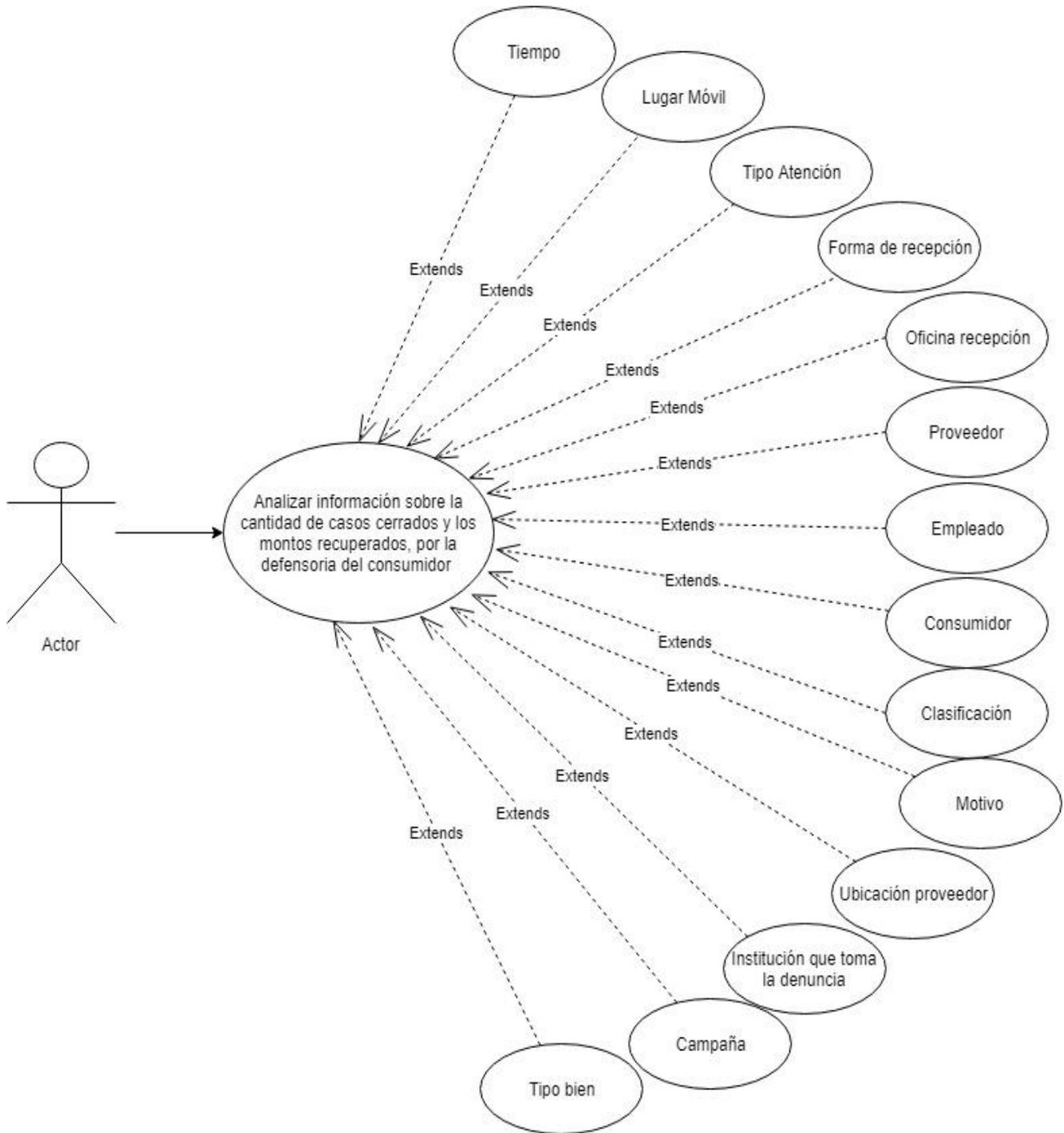


Figura 16.5: Caso de Uso de la cantidad de casos cerrados y el monto recuperado

16.5 Diseño Multidimensional

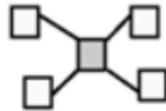
16.5.1 Diseño Conceptual del Data Mart

16.5.1.1 Nivel 1: Definición del modelo

Unidad: **Dirección de descentralización**

Especialización: Analizar las atenciones

Hechos principales: Cantidad de atenciones recibidas, Cantidad de casos cerrados y monto recuperado



Dirección de Descentralización

Figura 16.6: Nivel 1 Diagrama conceptual Data Mart Dirección de descentralización

16.5.1.2 Nivel 2: Definición del esquema estrella

Atenciones Recibidas:

Paquete de Hecho: Atenciones recibidas

Paquetes Dimensionales

- | | | |
|-----------------|---------------------|-------------------|
| 1. Oficina | 2. Consumidor | 15. Proveedor |
| 4. Ubicación | 5. Tipo bien | 6. Empleado |
| 7. Institución | 8. Campana | 9. Tiempo |
| 10. Lugar móvil | 11. Forma recepción | 12. Tipo Atención |
| 13. Motivo | 14. Clasificación | |

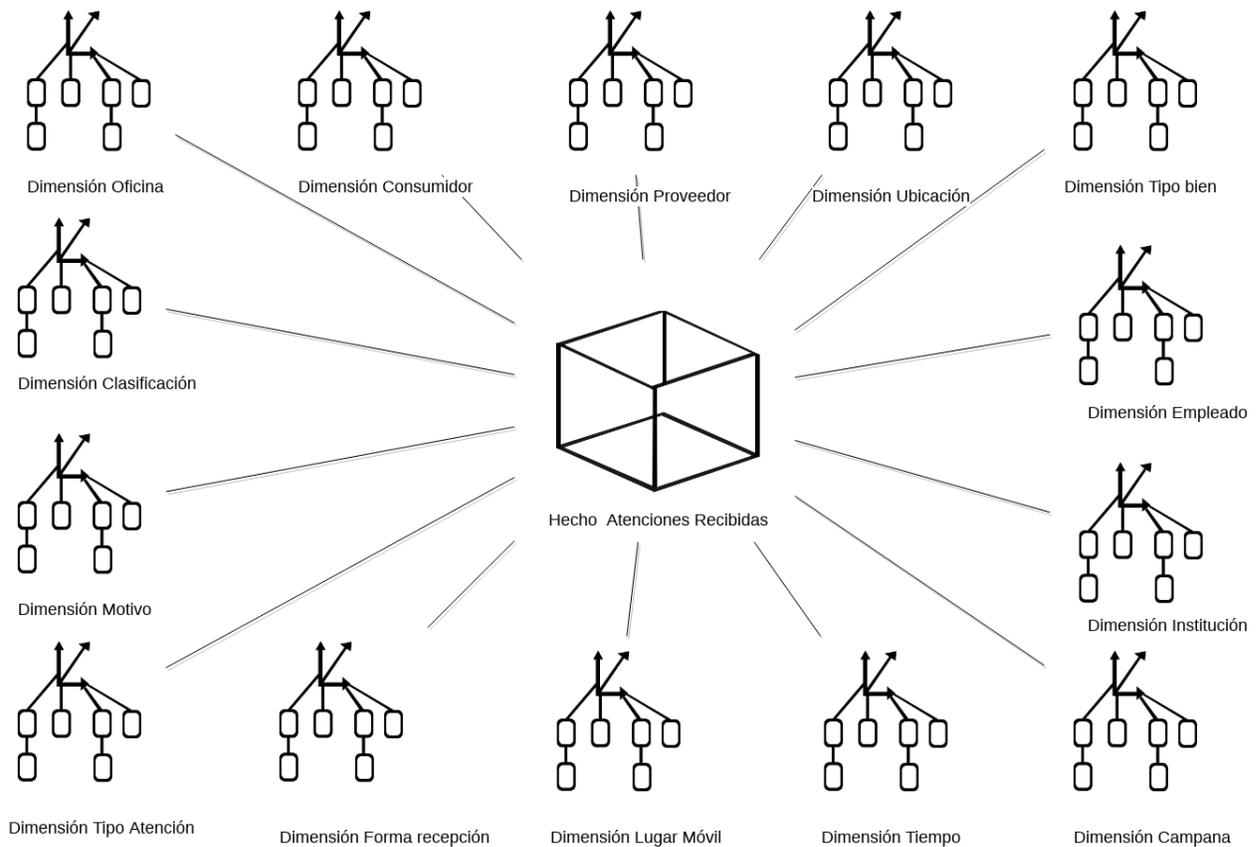


Figura 16.7: Diagrama Conceptual Data Mart (Atenciones recibidas)

Casos cerrados:

Paquete de Hecho: Cantidad de casos cerrados y monto recuperado

Paquetes Dimensionales:

- | | | |
|-------------------|-----------------|---------------------|
| 1. Oficina | 2. Consumidor | 3. Proveedor |
| 4. Ubicación | 5. Tipo bien | 6. Campaña |
| 7. Institución | 8. Empleado | 9. Etapa |
| 10. Tiempo | 11. Lugar Móvil | 12. Forma Recepción |
| 13. Tipo Atención | 14. Motivo | 15. Clasificación |

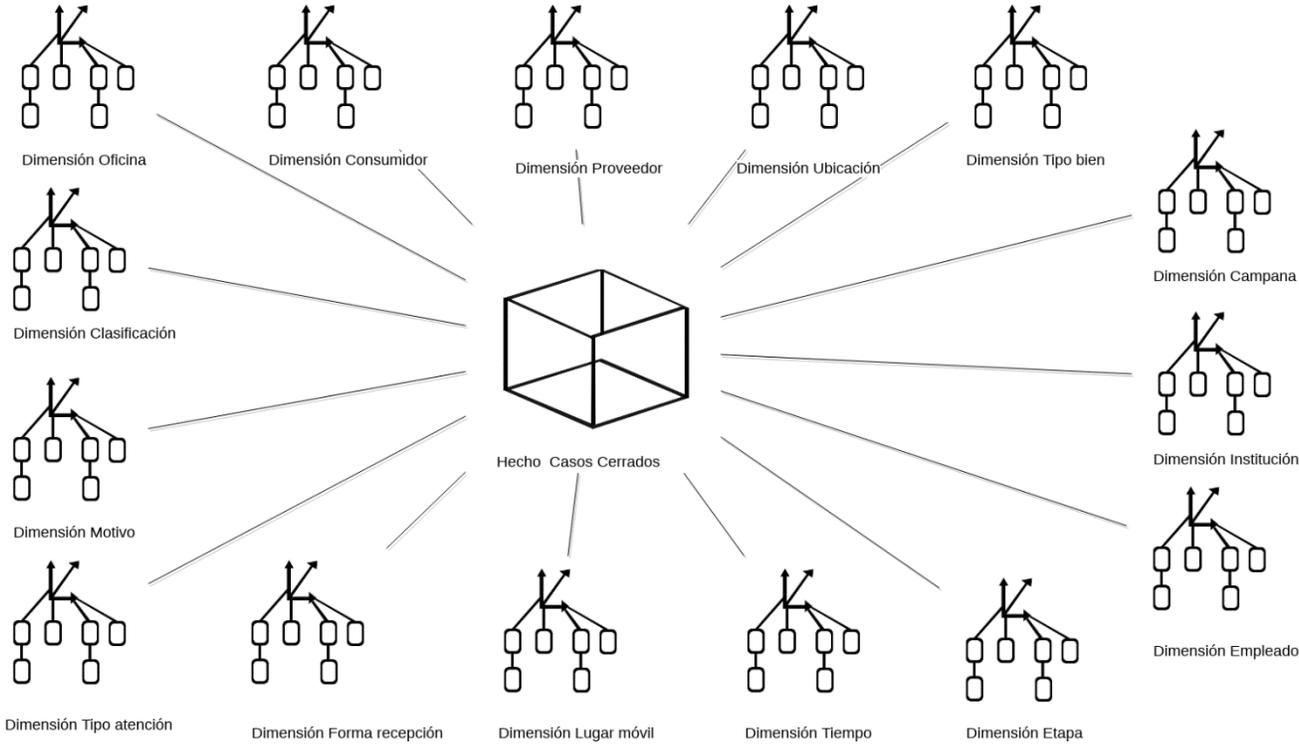


Figura 16.8: Nivel 2 Diagrama Conceptual Data Mart Casos cerrados)

16.5.1.3 Nivel 3: Dimensiones/Hechos

16.5.1.3.1 Fact Atenciones Recibidas

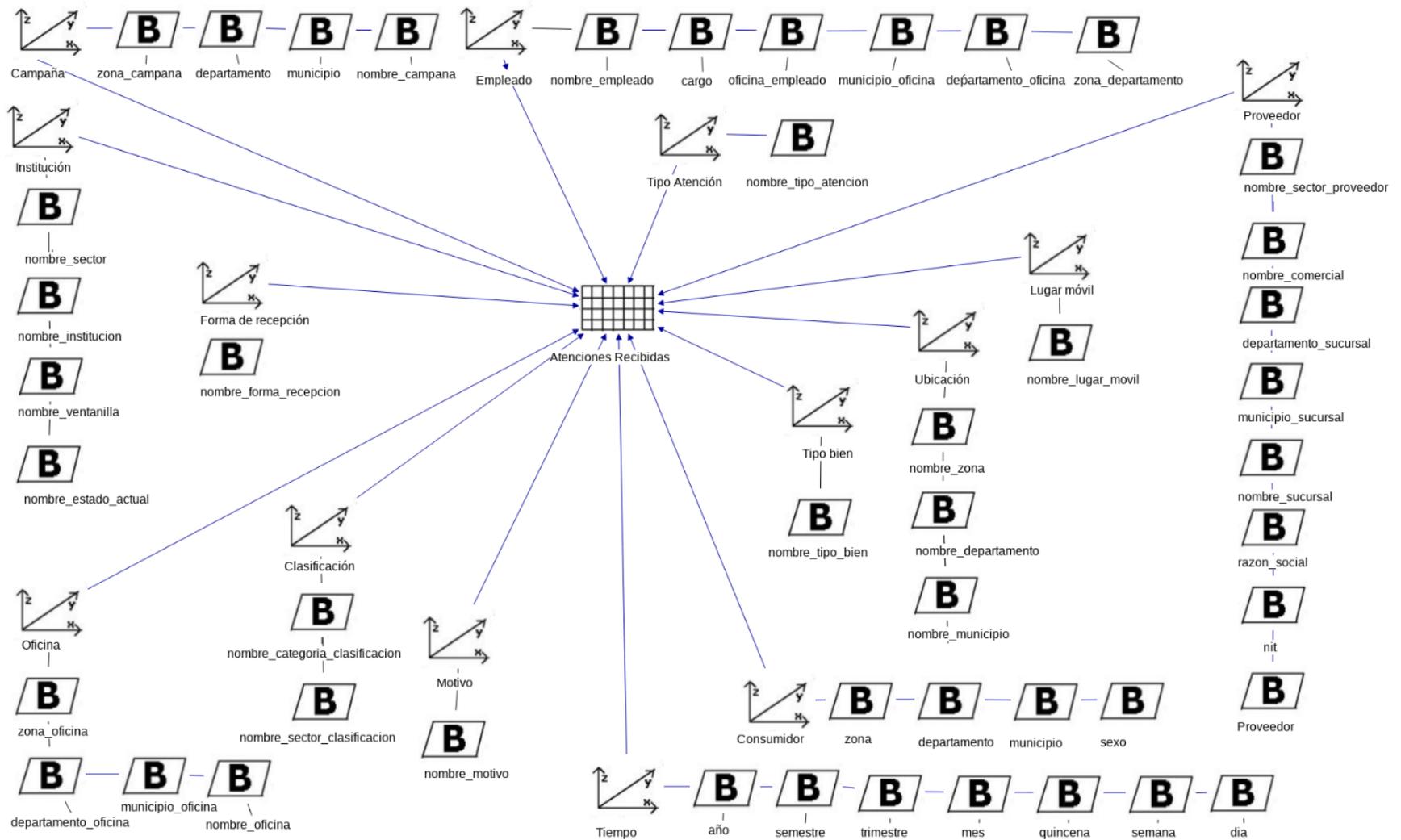


Figura 16.9: Nivel 3 Diagrama conceptual Data Mart Dirección de Descentralización (Hecho: Atenciones Recibidas)

16.5.1.3.2 Fact Casos Cerrados

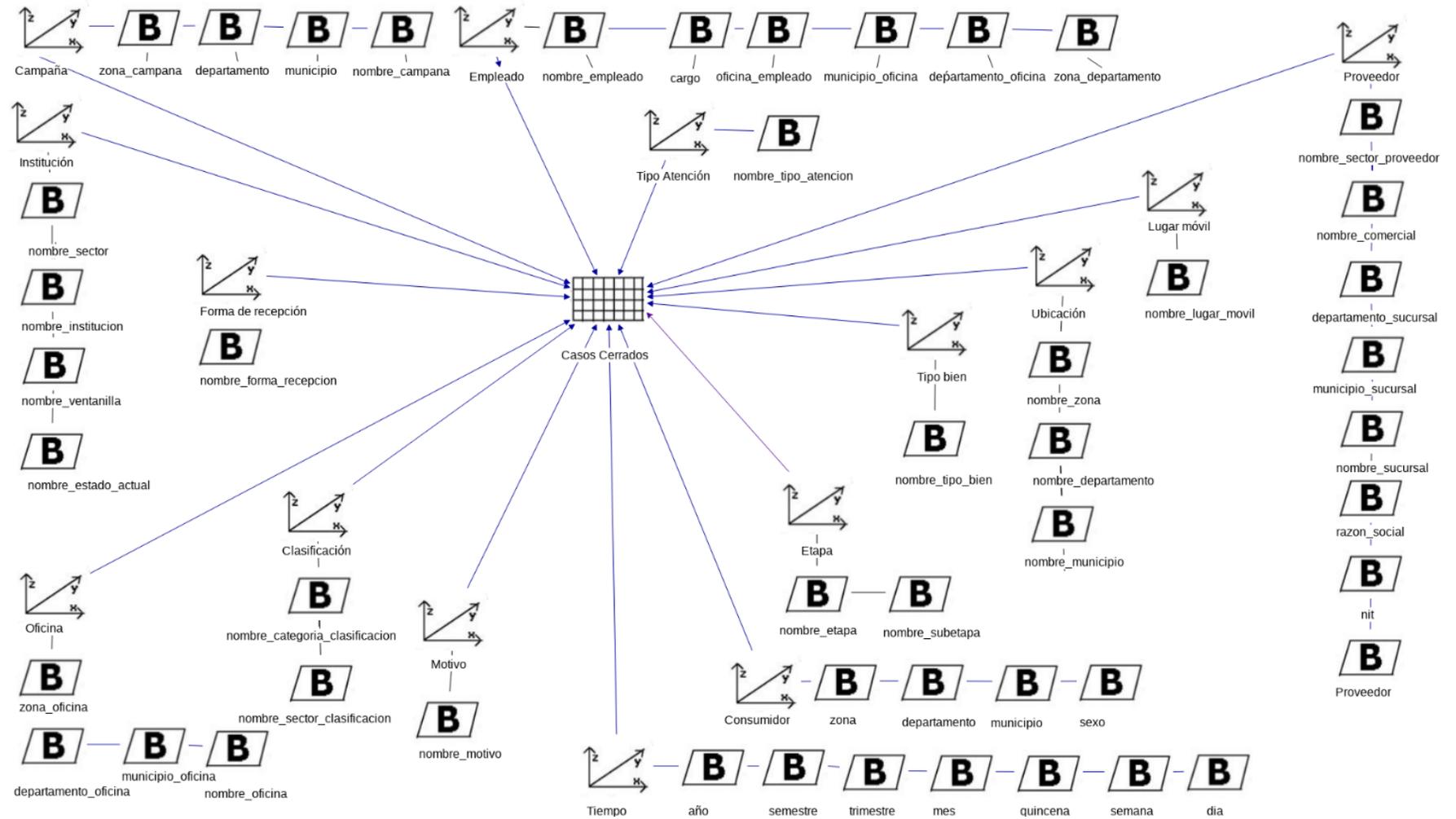


Figura 16.10: Nivel 3 Diagrama conceptual Data Mart Dirección de Descentralización (Hecho: Casos Cerrados)

16.5.2 Diseño Físico Base de Datos Data mart

A continuación se muestran los diagramas del diseño físico de base de datos del datamart para Dirección de Descentralización. Dado que se tienen dos tablas de hechos para esta unidad organizativa, se muestran dos diagramas diferentes (uno por cada tabla de hechos). La Figura 16.11 muestra el diseño físico para la tabla de hechos de atenciones recibidas, y la Figura 16.12 muestra el diseño físico para la tabla de hechos de casos cerrados.

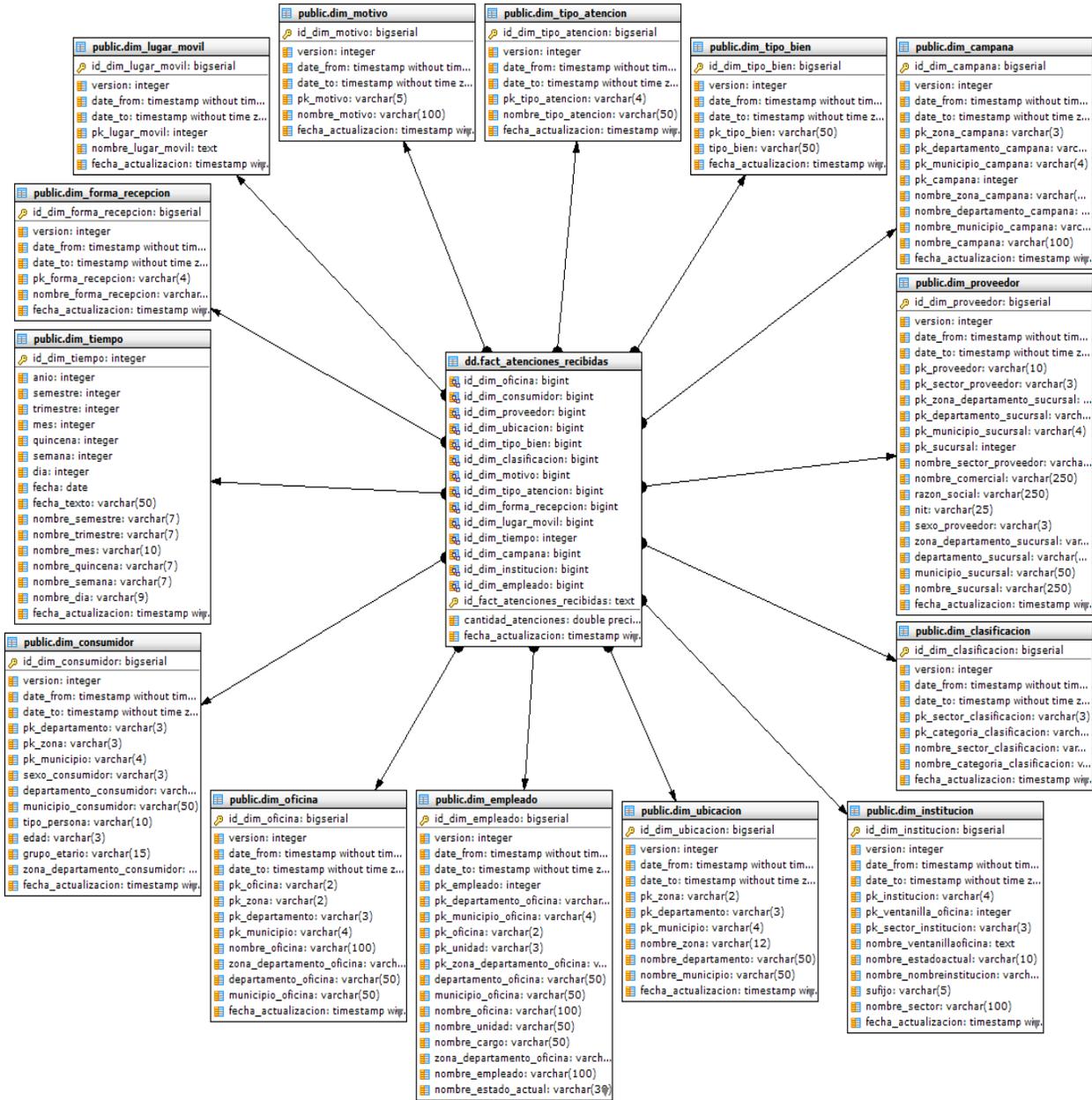


Figura 16.11: Diseño físico de base de datos, fact table atenciones recibidas, Dirección de Descentralización

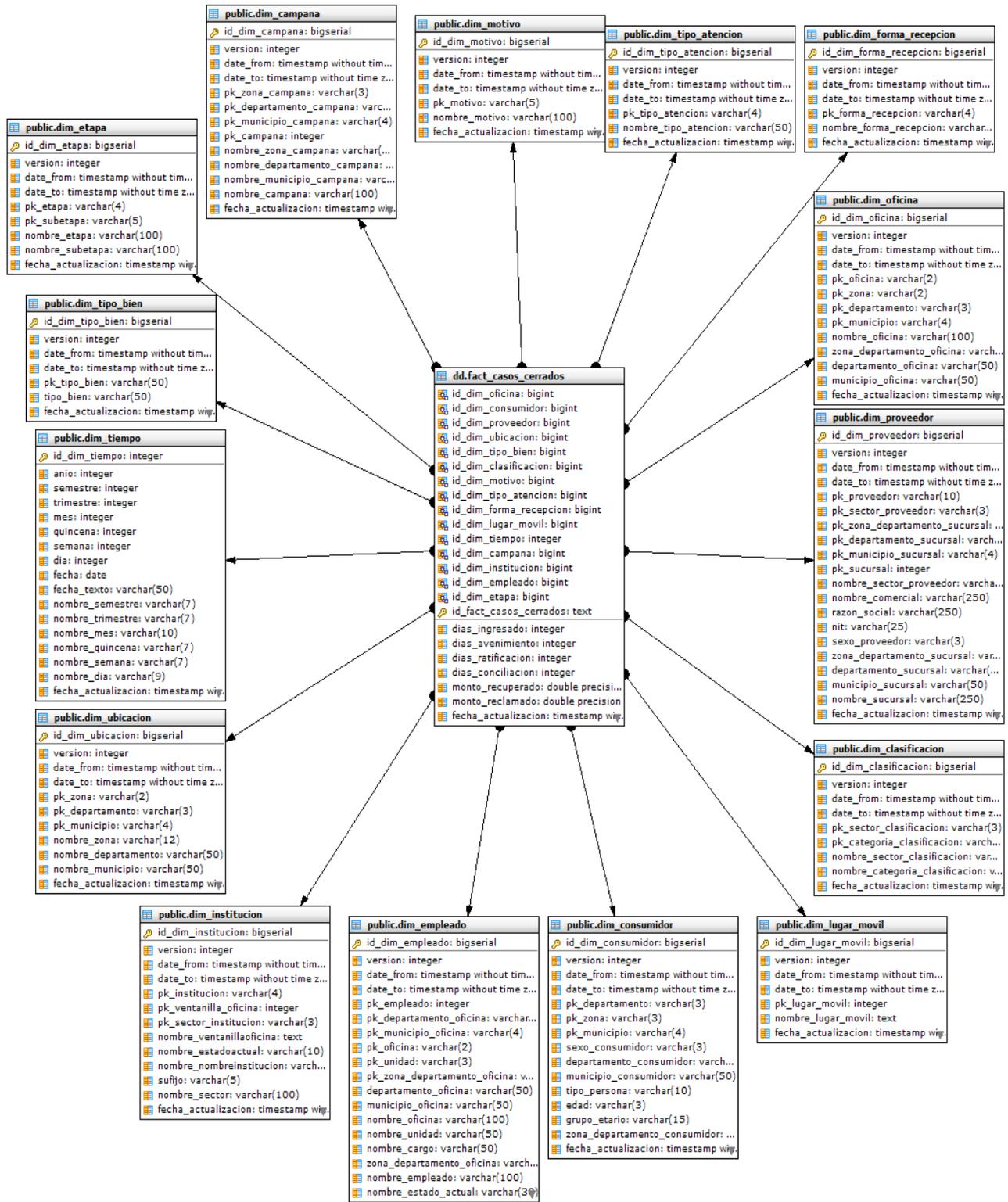


Figura 16.12: Diseño físico de base de datos, fact table casos cerrados, Dirección de Descentralización.

16.5.3 Diseño Físico Base de Datos Staging Area

A continuación, se muestra el diseño físico de base de datos del staging área para el data mart de Dirección de Descentralización. Las tablas se muestran en un solo diagrama.

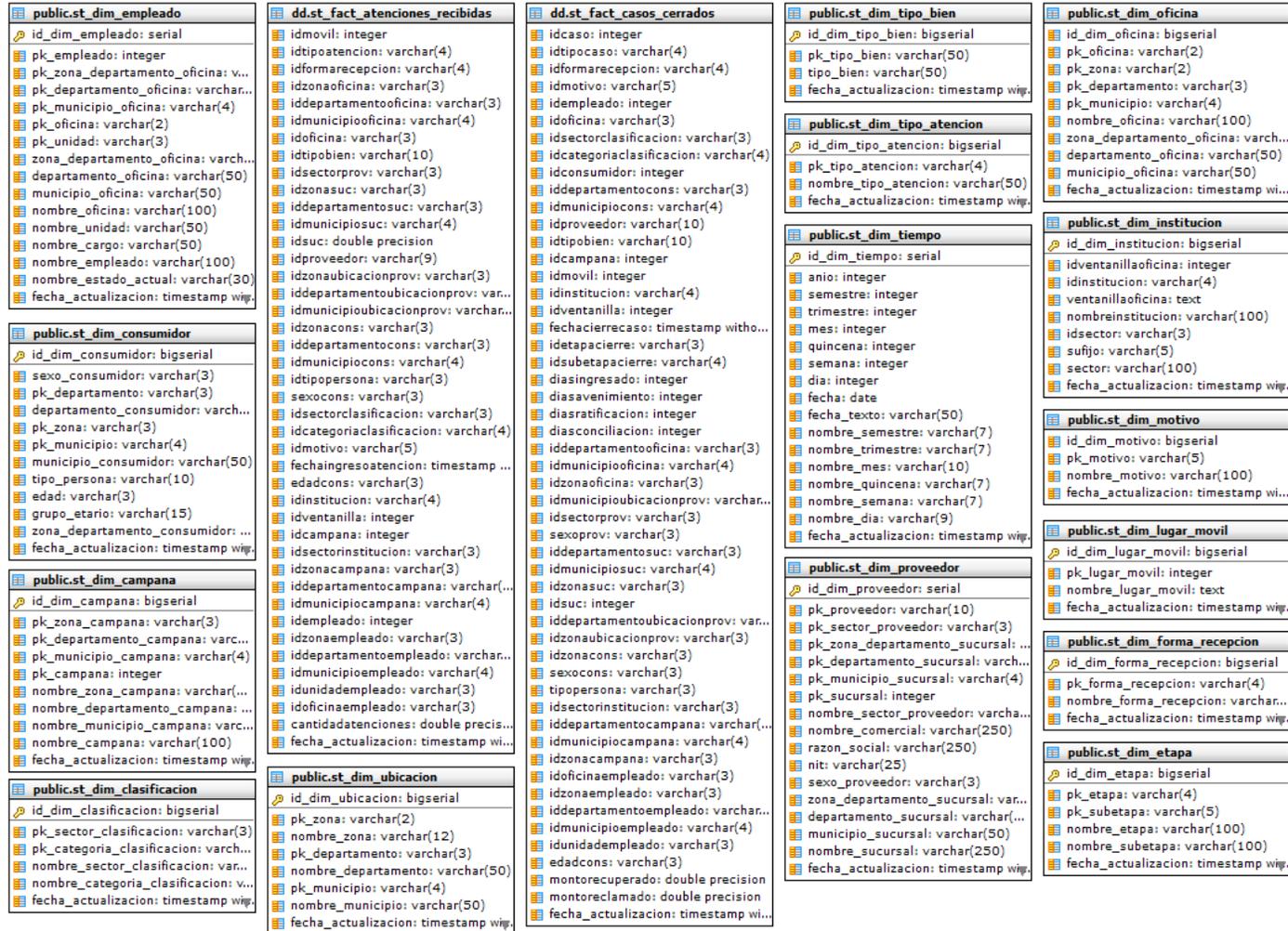


Figura 16.13: Diseño físico de base de datos para data mart de Dirección de Descentralización

16.6 Diseño de Procesos ETL

16.6.1 Procesos ETL Staging

En los siguientes apartados se listan los procesos ETL de Staging, seguido de un diagrama representativo para esta unidad. Los diseños de los procesos restantes se colocarán en el CD que acompaña al documento en la carpeta [CD-EISI\Diseños de Procesos ETL\dd \(Direccion de Descentralizacion\)](#)

Diagramas desarrollados:

- St_fact_atenciones recibidas
- St_fact_casos_cerrados

16.6.2 Procesos ETL Data Mart

A continuación, se listan los procesos ETL que mueven datos del Staging Area hacia el Data Mart (Procesos ETL Data Mart) desarrollados para esta unidad, seguido de un ejemplo (diagrama) para ilustrar uno de los procesos ETL.

Diagramas desarrollados:

- Dwt_fact_atenciones recibidas
- Dwt_fact_casos_cerrados

A continuación, se muestra uno de los diseños de proceso. Se muestra el diagrama Dw_fact_atenciones_recibidas

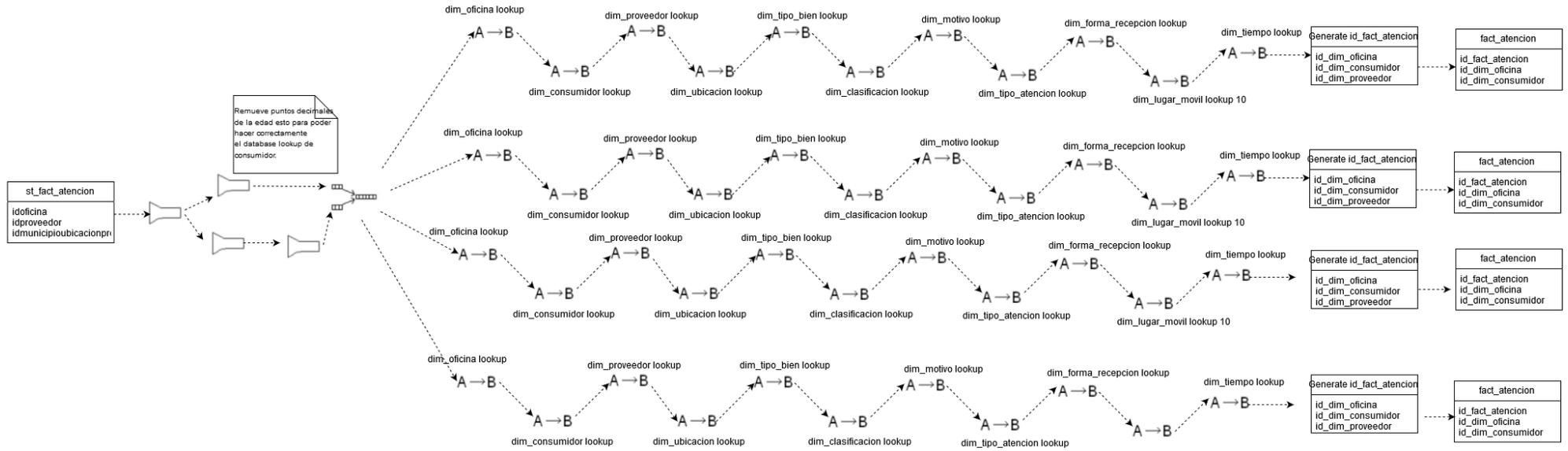


Figura 16.14: Proceso ETL Dw_fact_atenciones_recibidas.

16.7 Pruebas

16.7.1 Pruebas de Procesos ETL

La realización de las pruebas consistió en ejecutar los Jobs almacenados en el servidor de Staging Área. Las transformaciones finales para el Data Mart de Dirección de Descentralización se muestran en la Tabla 16.2. Se puede visualizar que no hay procesos ETL para llenado de dimensiones esto es porque se han reutilizado las dimensiones existentes.

No	Transformación	Descripción
1	dw_incremental_fact_atenciones_recibidas.ktr	Transformación para cargas incrementales desde tabla st_fact_atenciones_recibidas (staging area) hacia tabla fact_atenciones_recibidas (data warehouse).
2	dw_incremental_fact_casos_cerrados.ktr	Transformación para cargas incrementales desde tabla st_fact_casos_cerrados (staging area) hacia tabla fact_casos_cerrados (data warehouse).
3	dw_inicial_fact_atenciones_recibidas.ktr	Transformación para carga inicial desde tabla st_fact_atenciones_recibidas (staging area) hacia tabla fact_atenciones_recibidas (data warehouse).
4	dw_inicial_fact_casos_cerrados.ktr	Transformación para carga inicial del data warehouse desde tabla st_fact_casos_cerrados (staging area) hacia tabla fact_casos_cerrados (data warehouse).
5	st_incremental_fact_atenciones_recibidas.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia tabla st_fact_atenciones_recibidas (staging area).
6	st_incremental_fact_casos_cerrados.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia tabla st_fact_casos_cerrados (staging area).
7	st_inicial_fact_atenciones_recibidas.ktr	Transformación para carga inicial del data warehouse desde tablas fuentes hacia tabla st_fact_atenciones_recibidas (staging area).
8	st_inicial_fact_casos_cerrados.ktr	Transformación para carga inicial del data warehouse desde tablas fuentes hacia tabla st_fact_casos_cerrados (staging area).

Tabla 16.2: Transformaciones para Data Mart Dirección de Descentralización

Los Jobs finales de Primer Nivel para Data Mart de Dirección de Descentralización se muestran en Tabla 16.3.

No	Job Primer Nivel	Transformación	
		No	Nombre Transformación
1	job_incremental_fact_atenciones_recibidas.kjb	1	st_incremental_fact_atenciones_recibidas.ktr
		2	dw_incremental_fact_atenciones_recibidas.ktr
2	job_incremental_fact_casos_cerrados.kjb	1	st_incremental_fact_casos_cerrados.ktr
		2	dw_incremental_fact_casos_cerrados.ktr
3	job_inicial_fact_atenciones_recibidas.kjb	1	st_inicial_fact_atenciones_recibidas.ktr
		2	dw_inicial_fact_atenciones_recibidas.ktr
4	job_inicial_fact_casos_cerrados.kjb	1	st_inicial_fact_casos_cerrados.ktr
		2	dw_inicial_fact_casos_cerrados.ktr

Tabla 16.3: Jobs de Nivel 1 para Dirección de Descentralización

Los Jobs finales de Segundo Nivel para Data Mart de Dirección de Descentralización se muestran la Tabla 16.4.

No	Jobs Segundo Nivel	Jobs primer nivel ejecutados	
		No	Nombre Job
1	job_inicial_dd.kjb	1	job_inicial_fact_casos_cerrados.kjb
		2	job_inicial_fact_atenciones_recibidas.kjb
2	job_incremental_dd.kjb	1	job_incremental_fact_casos_cerrados.kjb
		2	job_incremental_fact_atenciones_recibidas.kjb

Tabla 16.4: Jobs de Nivel 2 Dirección de Descentralización

La finalización del proceso ETL se notifica mediante correo electrónico como se muestra en la Figura 16.15. La figura muestra solo parte del contenido del mensaje email.

```

job_inicial_dd
job_inicial_dd : : start : Start of job execution (2019/03/17 23:27:30.844)
job_inicial_dd : : START : start : Start of job execution (2019/03/17 23:27:30.850)
job_inicial_dd : : START : [nr=0, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:27:30.856)
job_inicial_dd : : job_inicial_fact_casos_cerrados.kjb : Followed unconditional link : Start of job execution (2019/03/17 23:27:30.862)
job_inicial_fact_casos_cerrados
job_inicial_fact_casos_cerrados : : START : Start of job entry : Start of job execution (2019/03/17 23:27:30.938)
job_inicial_fact_casos_cerrados : : START : [nr=1, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:27:30.939)
job_inicial_fact_casos_cerrados : : st_inicial_fact_casos_cerrados.ktr : Followed unconditional link : Start of job execution (2019/03/17 23:27:30.950)
job_inicial_fact_casos_cerrados : : st_inicial_fact_casos_cerrados.ktr : [nr=3, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:30:11.586)
job_inicial_fact_casos_cerrados : : dw_inicial_fact_casos_cerrados.ktr : Followed link after success : Start of job execution (2019/03/17 23:30:11.588)
job_inicial_fact_casos_cerrados : : dw_inicial_fact_casos_cerrados.ktr : [nr=4, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:31:35.918)
job_inicial_fact_casos_cerrados : : Mail : Followed link after success : Start of job execution (2019/03/17 23:31:35.919)
job_inicial_fact_casos_cerrados : : Mail : [nr=4, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:31:37.813)
job_inicial_fact_casos_cerrados : : Success : Followed link after success : Start of job execution (2019/03/17 23:31:37.814)
job_inicial_fact_casos_cerrados : : Success : [nr=4, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:31:37.815)
job_inicial_dd : : job_inicial_fact_casos_cerrados.kjb : [nr=1, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:31:37.872)
job_inicial_dd : : job_inicial_fact_atenciones_recibidas.kjb : Followed link after success : Start of job execution (2019/03/17 23:31:37.872)
job_inicial_fact_atenciones_recibidas
job_inicial_fact_atenciones_recibidas : : START : Start of job entry : Start of job execution (2019/03/17 23:31:37.885)
job_inicial_fact_atenciones_recibidas : : START : [nr=2, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:31:37.887)
job_inicial_fact_atenciones_recibidas : : st_inicial_fact_atenciones_recibidas.ktr : Followed unconditional link : Start of job execution (2019/03/17 23:31:37.888)
job_inicial_fact_atenciones_recibidas : : st_inicial_fact_atenciones_recibidas.ktr : [nr=4, errors=0, exit_status=0, result=true] : Job execution finished (2019/03/17 23:33:46.188)
    
```

Figura 16.15: Contenido correo email de notificación de finalización de los procesos ETL (Carga Inicial)

Las siguientes figuras muestran los Jobs de segundo nivel, los cuales son los Jobs principales que son ejecutados en el servidor. La Figura 16.16 muestra la secuencia del job que realiza la carga inicial, mientras que la Figura 16.17 muestra la secuencia del job que servirá para realizar las cargas incrementales.



Figura 16.16: Job ETL para carga inicial del Data Mart Dirección de Descentralización

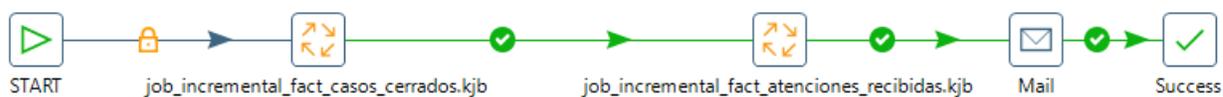


Figura 16.17: Job ETL para carga incremental del Data Mart Dirección de Descentralización

16.7.1.1 Ejecución de Procesos ETL iniciales

El comando para ejecutar el proceso ETL se muestra a continuación:

```
#/root/data-integration/kitchen.sh -file=/root/jobs/dd/job_inicial_dd.kjb -level=Debug
```

A cotinuacion se muestra en resumen la ejecución de las transformaciones y Jobs desarrollados para para la Dirección de Descentralización.

Catidad de Transformaciones/ job	Tiempo de ejecución	Cantidad de registros movidos
6	00:12:10.684	381,710

Para más detalle de las pruebas realizadas en el Sprint 5 acceder al documento Release 2.pdf, sección 9.9.1 incluido en el CD que acompaña este documento en la carpeta: CD-EISI\Documentación\

La finalización del proceso ETL se notifica mediante correo electrónico como se muestra en la Figura 16.18



Figura 16.18: Notificación por correo finalización de proceso ETL (Carga Incremental)

16.7.2 Pruebas Cubos OLAP

Dado que para Dirección de Descentralización se han identificado dos hechos a medir, se realizaron dos cubos para dar respuesta a lo solicitado por esta unidad. A continuación, se muestran los cubos realizados para el Data Mart de Dirección de Descentralización.

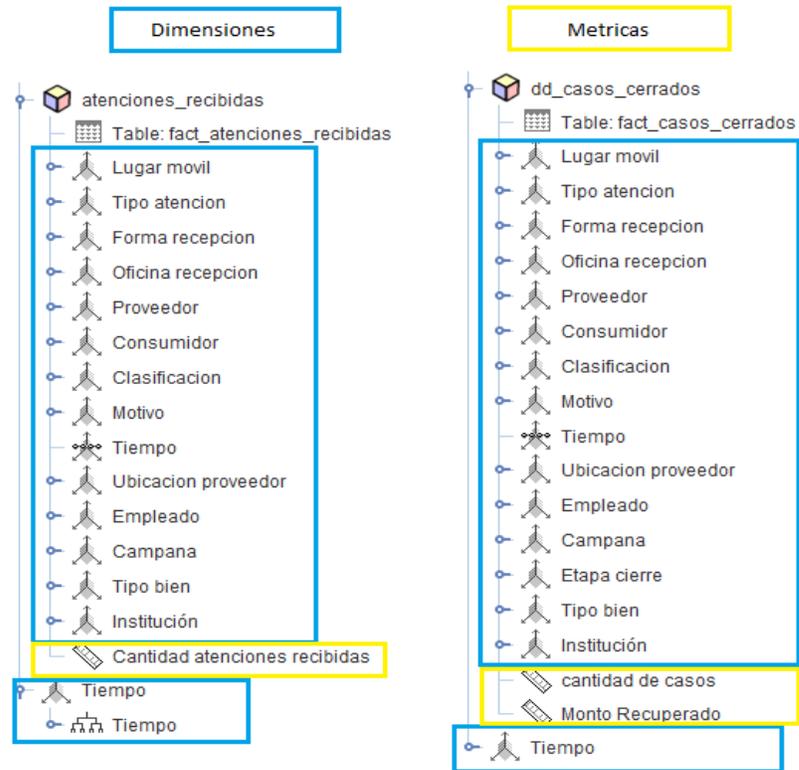


Figura 16.19: Cubos para Dirección de Descentralización.

Para estos cubos se han generado y publicado su correspondiente archivo XML en el Servidor de Pentaho Server BI para su posterior visualización.

Como se ha dicho anteriormente, con el uso del Plugin JPivot para Pentaho Server se logra visualizar las metricas de estos cubos desde las diferentes dimensiones asociadas. Las siguientes figuras muestran los dos cubos realizados para esta unidad organizacional.

Oficina recepcion	Institucion	Medidas						
		Cantidad atenciones recibidas						
		Tiempo						
		Todos los años	2008	2009	2010	2011	2012	2013
Todas las oficinas	Instituciones	133.665	8.544	23.744	21.540	10.564	3.604	20.608
	Administración Nacional de Acueductos y Alcantarillados (ANDA)	1						
	CENADE	2						
	Ciudad Mujer	36						
	Consejo Superior de Salud Pública	9		4	1			2
	Corte Suprema de Justicia	2						
	DOCUSAL	6	1	3	1	1		
	Dirección General de Migración y Extranjería	1						
	Dirección General del Instituto Salvadoreño del Seguro Social	1						1
	Dirección Nacional de Medicamentos	6						2
	Gobernaciones Departamentales	1.433		4	38	100	20	367
	INSAFOCOOP	19				2		2
	ISS	1					1	

Figura 16.20: Visualización del Cubo atenciones recibidas en la Defensoría del Consumidor.

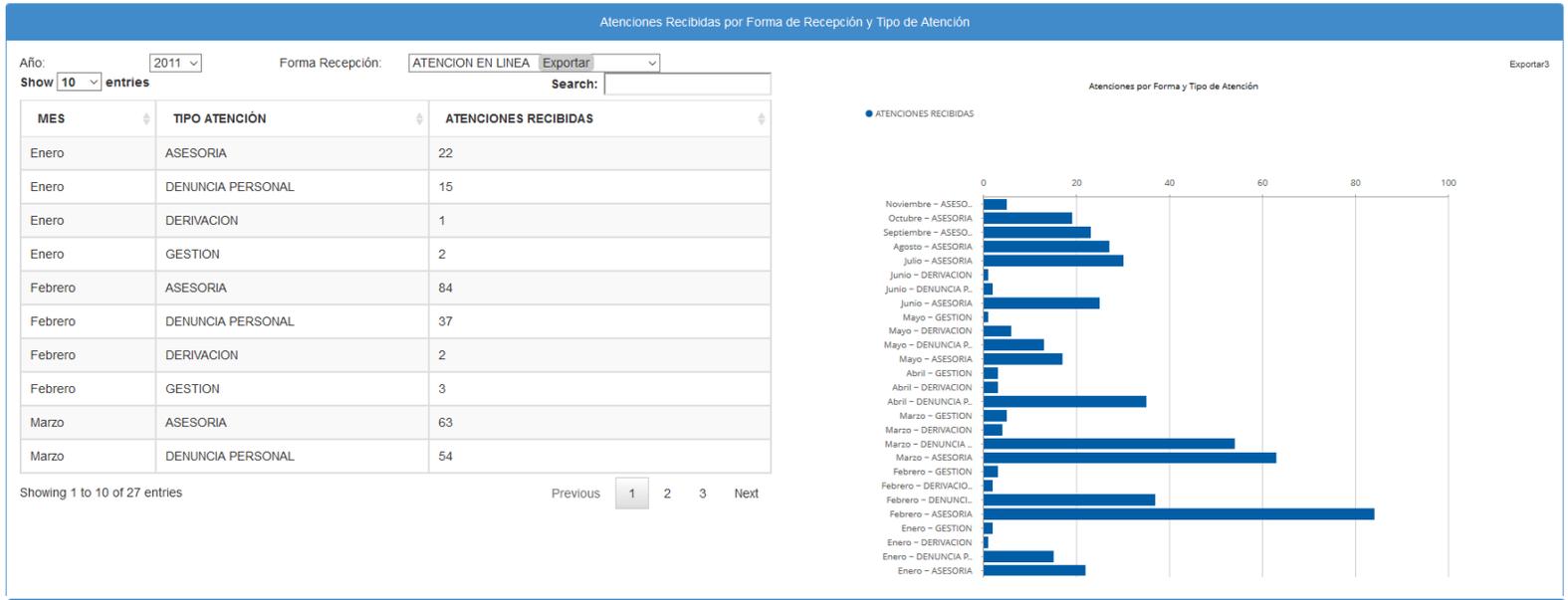
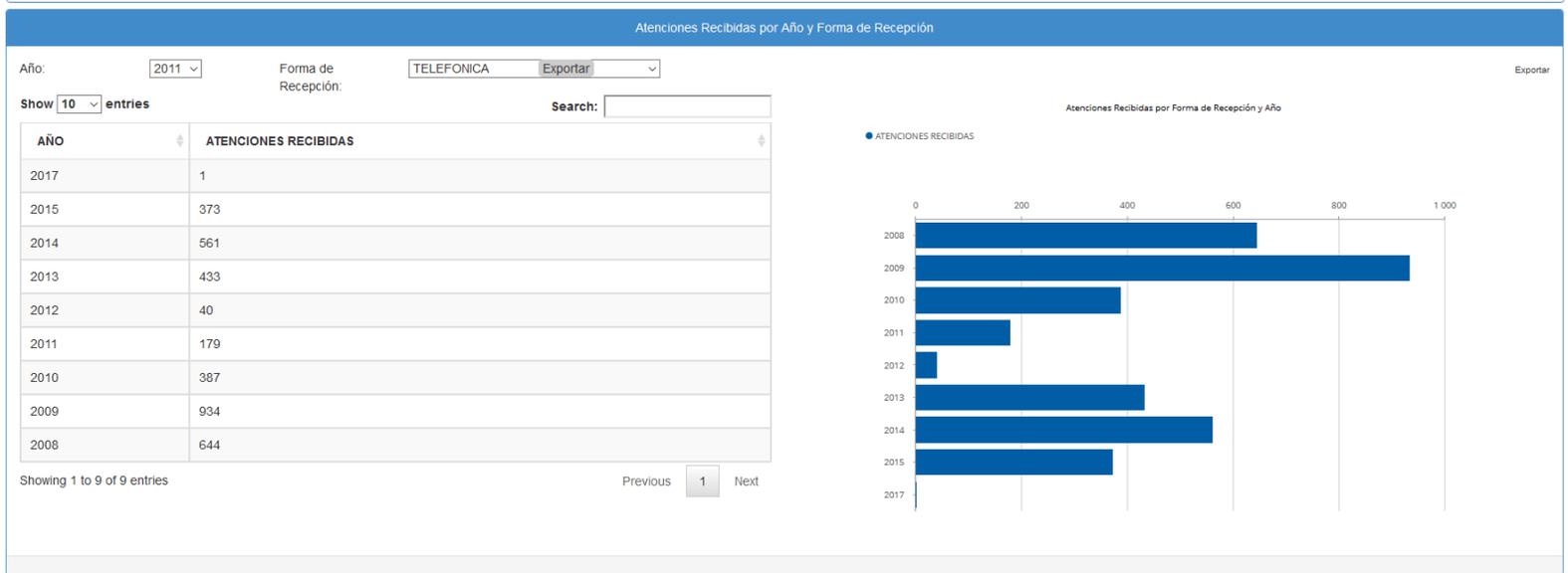
Motivo	Tiempo	Institucion	Medidas	
			Monto Recuperado	cantidad de casos
Todos los motivos	Todos los años	Instituciones	12.150.836,03	58.686
	2008	Instituciones	889.794,29	4.084
		Ministerio de Economía (MINEC)	0	2
		N/A	889.794,29	4.079
		Superintendencia de General de Electricidad y Telecomunicaciones (SIGET)	0	3
	Primer semestre	Instituciones	89.388,92	586
	Segundo semestre	Instituciones	800.405,37	3.498
	2009	Instituciones	2.618.453,85	11.089
	2010	Instituciones	2.431.573,93	16.547
	2011	Instituciones	1.419.136,55	6.420
	2012	Instituciones	278.461,12	1.843
	2013	Instituciones	1.309.582,81	5.801
	2014	Instituciones	1.995.672,59	7.852
	2015	Instituciones	1.208.149,89	5.040

Figura 16.21: Visualización del Cubo casos cerrados en la Defensoría del Consumidor.

16.7.3 Pruebas Dashboard

El Dashboard realizado para Dirección de Descentralización se muestra en la siguiente página.

16.7.3.1 Prueba de Dashboard Casos Cerrados



17 SPRINT 6

17.1 Sprint Backlog

A continuación, se presenta el Sprint Backlog, que refleja las historias de usuarios, su estimación y asignación para este Sprint.

#	Historia	Punto Estimado	Asignado
1.	Redacción de la plantilla para el refinamiento de requerimientos	20	Todos
2.	Poblar la base de datos de desarrollo SARA	13	Jorge
3.	Diseño del modelo multidimensional	8	Samuel – Ivis
4.	Diseño BD del Staging Area	3	Jorge – Samuel
5.	Diseño del proceso ETL	8	Todos
6.	Diseño de vistas	3	Ivis – Esau -- Samuel
7.	Creación de Script SQL del Staging Area	13	Jorge – Ivis – Samuel
8.	Creación de Script SQL modelo multidimensional	13	Jorge – Ivis – Samuel
9.	Desarrollo de los procesos ETL	100	Jorge – Ivis – Samuel
10.	Desarrollo de cubos OLAP	40	Jorge – Ivis – Samuel
11.	Desarrollo de vistas	13	Ivis – Esau -- Samuel
12.	Implementación e integración de componentes	40	Todo
13.	Pruebas	20	Jorge – Ivis – Samuel
14.	Documentación	40	Todos

Tabla 17.1: Sprint Backlog, Sprint 6

17.2 Descripción de Historias de Usuario

A continuación, se presenta la descripción de las historias de usuario detallando: el responsable, al sprint que pertenece, una breve descripción, sus criterios de validación y los puntos de estimación. Previamente a las historias de usuario se presenta la historia épica del Sprint 6.

<i>Historia Épica #7</i>	Data Mart Unidad de análisis y consumo de mercados.
Responsable	Jorge Iraheta/Esau Serpas/Ivis Navarro/Samuel Núñez Sprint: 6
Descripción	Como Unidad de Análisis y Consumo de Mercados (jefatura) Necesito: Un mecanismo para analizar la información (Cantidad de atenciones ingresadas, Cantidad de casos finalizados junto a los montos recuperados y reclamados) registrados en la defensoría del consumidor. Para: visibilizar el trabajo que se desempeña como unidad, optimizando la carga de trabajo de la unidad.
Criterios de aceptación	- La cantidad de atenciones recibidas deben poder ser visualizadas por las siguientes perspectivas: <ul style="list-style-type: none"> • Por fecha • Por lugar móvil • Por tipo atención • Por forma de recepción • Por oficina de recepción • Por proveedor • Por consumidor • Por clasificación (Sector y Categoría) • Por motivo

	<ul style="list-style-type: none"> • Por campaña • Por ubicación del proveedor • Por tipo bien • Por institución • Por empleado • Por etapa <p>- La cantidad de casos cerrados junto a los montos recuperados y montos reclamados, los cuales se deben poder analizarse desde las siguientes perspectivas:</p> <ul style="list-style-type: none"> • Por fecha • Por proveedor • Por ubicación del proveedor • Por consumidor • Por empleado • Por oficina de recepción • Por clasificación (Sector y Categoría) • Por motivo • Por forma de recepción • Por lugar móvil • Por campaña • Por tipo atención • Por etapa Cierre • Por tipo bien • Por institución que toma la denuncia.
Puntos estimados	334

Historia de usuario #1	Refinamiento del requerimiento de información
Responsable	Jorge Iraheta Sprint: 6
Descripción	<p>Como equipo Scrum necesito el refinamiento del requerimiento de información de la unidad de Análisis y Consumo de Mercados: Realizar reuniones con los interesados y especificar los elementos de información requeridos para la unidad de Análisis y Consumo de Mercados.</p> <ul style="list-style-type: none"> • Llenar la herramienta de recolección. • Revisar y desglosar los requerimientos solicitados • Firmar documento por el usuario como aceptación de los de los requerimientos. <p>Para tener claridad de la información que desean almacenar, visualizar y analizar mediante el Data Mart o modelo Multidimensional para ésta unidad.</p>
Criterios de aceptación	<ul style="list-style-type: none"> - Tener llena la plantilla de refinamiento de requerimientos de información. - Incluir únicamente requerimientos de información con naturaleza histórica y no información transaccional. - Tener un documento firmado de aceptación de requerimientos. - Identificar (mediante un listado) las tablas fuentes que servirán para cumplir con los requerimientos de información solicitados.
Puntos estimados	20

<i>Historia de usuario #2</i>	Poblar base de datos de desarrollo SARA
<i>Responsable</i>	Samuel Núñez Sprint: 6
<i>Descripción</i>	Como Product Owner necesito: migrar los datos necesarios según las tablas fuentes de datos que se identifiquen durante el proceso de toma de requerimientos. Para poseer datos de prueba para el desarrollo del Data Mart.
<i>Criterios de aceptación</i>	- Se deben poseer Scripts de INSERT para las tablas Fuente. - Todas las tablas fuentes identificadas en el refinamiento de requerimientos deben poseer datos. - Debe existir al menos una prueba de conexión y acceso remoto exitoso a los datos. - Datos cargados en la base de datos
<i>Puntos estimados</i>	13

<i>Historia de usuario #3</i>	Diseño del modelo multidimensional
<i>Responsable</i>	Ivis Navarro/Samuel Núñez Sprint: 6
<i>Descripción</i>	Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de la unidad de Análisis y Consumo de Mercados. Para tener la base (diagrama) para el desarrollo de Scripts de Base de Datos SQL para Data Mart.
<i>Criterios de aceptación</i>	-Esquema del modelo multidimensional de la unidad de Análisis y Consumo de Mercados. -El esquema debe cumplir las reglas de un modelo de tipo Estrella. -El esquema debe poseer las relaciones necesarias para la búsqueda según los criterios de búsqueda definidos por la Unidad de Análisis y Consumo de Mercados. -El esquema debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
<i>Puntos estimados</i>	8

<i>Historia de usuario #4</i>	Diseño BD del Staging Area.
<i>Responsable</i>	Jorge Iraheta/Samuel Núñez Sprint: 6
<i>Descripción</i>	Como Equipo Scrum necesito: el diseño del modelo multidimensional para el Data Mart de la unidad de Análisis y Consumo de Mercados. Para tener el insumo para la creación de Scripts de Base de Datos SQL para Staging Area.
<i>Criterios de aceptación</i>	-Deberá existir una o más tablas por cada dimensión. -Deberá existir una o más tablas de hecho. -El diseño debe especificar los tipos de datos y tamaños si aplica. -Deberá existir diccionario de Datos para dicho diseño.
<i>Puntos estimados</i>	3

<i>Historia de usuario #5</i>	Diseño del proceso ETL
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Núñez/Esau Serpas Sprint: 6
<i>Descripción</i>	Como Equipo Scrum necesito: Diseñar los Proceso ETL <ul style="list-style-type: none"> • Extracción • Transformación • Carga Para cumplir con el Diseño del proceso ETL y tenerlo como insumo a la hora de desarrollar dichos procesos ETL.
<i>Criterios de aceptación</i>	-Se deberá utilizar nomenclatura UML para los pasos en cada proceso ETL. -Deberá incluir todas las tablas dimensiones y tablas de hechos. -Los nombres de las tablas o etiquetas deben poderse visualizar fácilmente.
<i>Puntos estimados</i>	8

<i>Historia de usuario #6</i>	Diseño de vistas
Responsable	Ivis Navarro/Samuel Núñez/Esau Serpas Sprint: 6
Descripción	Como Administrador de Data Warehouse necesito: Realizar el diseño de vistas necesarias para visualizar la información del Data Mart de la unidad de Análisis y Consumo de Mercados. Para Diseñar y digitalizar la vista de presentación de datos.
Criterios de aceptación	-Deberá existir un diseño de vista o dashboard para cada hecho a medir. -Deberá explicarse cada sección del dashboard o vista. -Etiquetas y textos deben ser legibles a simple vista.
Puntos estimados	3

<i>Historia de usuario #8</i>	Creación de Script SQL del Staging Area
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 6
Descripción	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el Esquema Estrella en el SGBD. Para la implementación del esquema donde será procesada, calculada y/o limpiada la información.
Criterios de aceptación	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Staging Area. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
Puntos estimados	13

<i>Historia de usuario #7</i>	Creación de Script SQL Modelo Multidimensional
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 6
Descripción	Como Equipo Scrum necesito: Desarrollar el script SQL para implementar el esquema estrella en el SGBD. Para la implementación del esquema donde será almacenada la información.
Criterios de aceptación	-El Script debe poderse ejecutar en la base de datos sin ningún error. -Los nombres, etiquetas y tipos de datos deben coincidir con los plasmados en los diagramas del Modelo Multidimensional. -El Script deberá estar disponible para todos los miembros del equipo mediante la nube.
Puntos estimados	13

<i>Historia de usuario #9</i>	Desarrollo de los procesos ETL.
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 6
Descripción	Como Equipo Scrum necesito: Desarrollar los procesos ETL correspondientes Para obtener, trasladar y transformar los datos que serán almacenados en el Data Mart de la unidad de Análisis y Consumo de Mercados.
Criterios de aceptación	-Los nombres de los procesos ETL deben poseer la nomenclatura definida por el equipo. -Los procesos ETL deben poseer la documentación necesaria en forma de comentarios dentro de la herramienta PDI. -Las Transformaciones ETL deben haber sido probados en desarrollo (Equipo Local) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de pruebas . -Las Transformaciones ETL deben haber sido probados en producción (Servidor) usando la base de datos pública de final . -Las transformaciones correspondientes a una dimensión o hecho deben ser incorporados un Job de PDI. -Deberá crearse un Job que incorpore los otros Jobs correspondientes a cada dimensión y hecho.

	-Los Jobs de PDI deben haber sido probados con las bases de datos de prueba y base de datos final. -Las bitácoras de los procesos ETL deben almacenarse en una base de datos.
Puntos estimados	100

Historia de usuario #10	Desarrollo de Cubos OLAP
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 6
Descripción	Como Equipo Scrum necesito: una investigación sobre la herramienta para el desarrollo de cubos OLAP e implementar dicha herramienta para la creación de cubos OLAP. Para implementar el cubo correspondiente a la unidad de Análisis y Consumo de Mercados.
Criterios de aceptación	-El cubo deberá poderse visualizar correctamente y sin errores. -Deberá existir como mínimo un cubo OLAP por Hecho. -Deberán poderse realizar las operaciones como mínimo de: Roll Up y Drill Down.
Puntos estimados	40

Historia de usuario #11	Desarrollo de vistas
Responsable	Ivis Navarro/Samuel Núñez/Esau Serpas Sprint: 6
Descripción	Como Equipo Scrum necesito: Desarrollar vistas o dashboards según el diseño previamente elaborado. Para presentar información relevante hacia la unidad para la toma de decisiones.
Criterios de aceptación	-Vista o dashboard funcionando sin errores técnicos. -Debe coincidir con el diseño realizado previamente. -Debe existir al menos un dashboar por tabla de hechos.
Puntos estimados	13

Historia de usuario #12	Implementación e integración de componentes
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez/Esau Serpas Sprint: 6
Descripción	Como Equipo Scrum necesito: La integración de todos los elementos del modelo multidimensional <ul style="list-style-type: none"> - ETL - Fuentes - Stagin Área - BD multidimensional - Pentaho Server (Vistas) Para el desarrollo del modelo multidimensional de forma integral y probar el funcionamiento correctamente de la totalidad del Data Mart.
Criterios de aceptación	- Se deberán modificar los procesos ETL para que utilicen las bases de datos finales (no las de pruebas). - Se deberán realizar pruebas en la visualización de los dashboards. - Todo el entorno (desde los procesos ETLs, hasta los dashboards o vistas) deberán tener el visto bueno de todo el equipo.
Puntos estimados	40

Historia de usuario #13	Pruebas
Responsable	Jorge Iraheta/Ivis Navarro/Samuel Núñez Sprint: 3
Descripción	Como Equipo Scrum necesito: Realizar pruebas al modelo multidimensional terminado. Para verificar que funciona de la manera que se espera.
Criterios de aceptación	Pruebas exitosas en pruebas de aceptación del usuario
Puntos estimados	20

<i>Historia de usuario #14</i>	Documentación
<i>Responsable</i>	Jorge Iraheta/Ivis Navarro/Samuel Nuñez/Esau Serpas Sprint: 3
<i>Descripción</i>	Como Equipo Scrum necesito: Documentación completa con historias, nomenclatura y diagramas empleados. Para verificar el óptimo desarrollo de los cubos (la unidad de Análisis y Consumo de Mercados).
<i>Criterios de aceptación</i>	Documentos completos de análisis y diseño para modelo multidimensionales
<i>Puntos estimados</i>	40

17.3 Refinamiento del Requerimiento de Información

Para la realización del refinamiento de los requerimientos de información se partió con la investigación del proceso general que sigue la unidad de Análisis y Consumo de Mercados dentro la Defensoría del Consumidor. Además de establecer el proceso general de la unidad se hizo uso de los paquetes de información para refinar el requerimiento de información.

17.3.1 Proceso BPMN de la Unidad

A continuación, presentamos el diagrama BPNM que engloba las actividades de la unidad de Análisis y Consumo de Mercados.

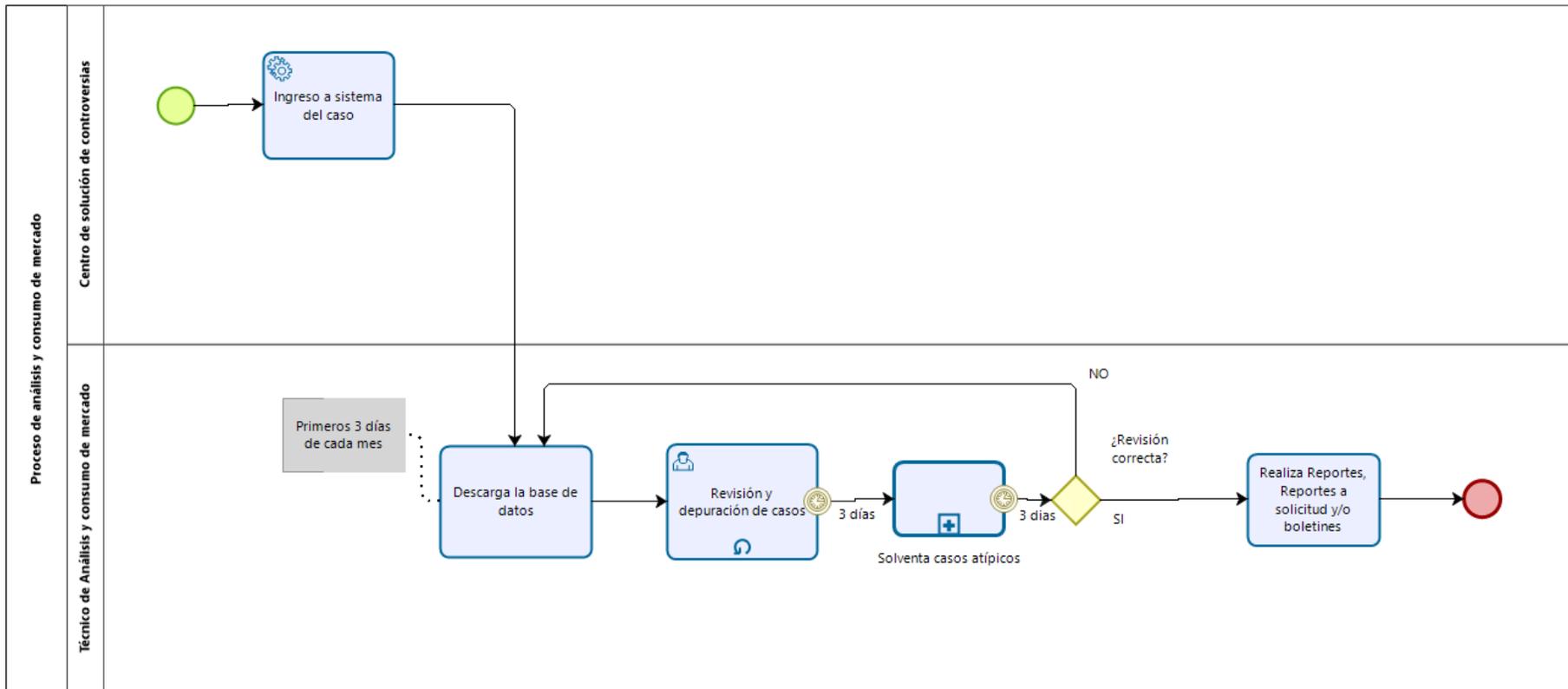


Figure 17.1: BPNM de Unidad de Análisis y Consumo de Mercados

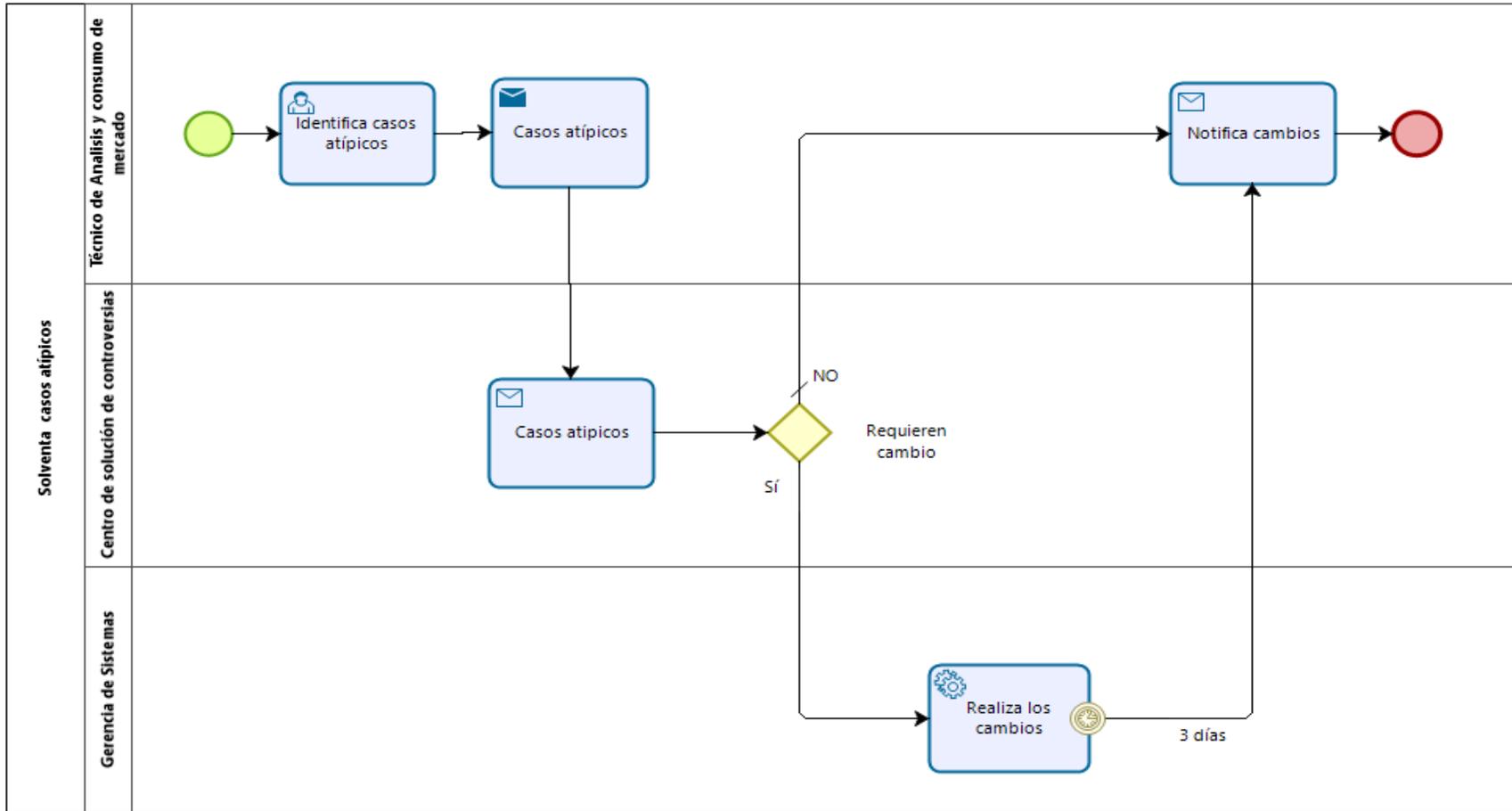


Figure 17.2: BPMN de Sub-Proceso: Solventa Casos atípicos

17.3.2 Paquetes de información

A continuación, se presenta los paquetes de información para la unidad de Análisis y Consumo de Mercados.

Para la unidad de Análisis y Consumo de Mercados de la Defensoría del Consumidor se identificaron dos paquetes de información para los siguientes temas:

- Figura 17.1 Paquete de información para cantidad de atenciones recibidas en la DC
- Figura 17.2 Paquete de información para cantidad y montos de los casos cerrados en la DC

Tema:		Cantidad de atenciones recibidas en la Defensoría del consumidor.												
J E R A R Q U I A S	Tiempo	Lugar Movil	Tipo Atencion	Forma de Recepción	Oficina Recepción	Proveedor	Consumidor	Clasificación	Motivo	Campaña	Ubicacion proveedor	Tipo Bien	Institucion que toma la denuncia	Empleado
	Año	Nombre Lugar	Nombre tipo atencion	Nombre forma recepcion	Zona	NombreSector	Sexo	Nombre sector	Nombre Motivo	Zona Campaña	Zona	Nombre Tipo Bien	Nombre de la institucion	Zona Departamento
	Semestre				Departamento	NombreComercial	Zona	Nombre categoría		Departamento	Departamento		Sector de la ventanilla	Departamento de la Oficina
	Trimestre				Municipio	RazonSocial	Departamento			Municipio	Municipio		Nombre ventanilla oficina	Municipio de la Oficina
	Mes				Oficina	NIT	Municipio			Nombre campaña				Oficina del Empleado
	Quincena					Departamento Sucursal								Cargo del Empleado
	Semana					MunicipioSucursal								Nombre Empleado
	Día					NombreSucursal								
	Hechos Medidos:	Cantidad de atenciones recibidas												

Figura 17.1 Paquete de información para cantidad de atenciones recibidas en la DC

Tema:		Casos Cerrados por Dirección del Centro de Solución de Controversias													
J E R A R Q U I A S	Tiempo	Proveedor	Ubicación del Proveedor	Consumidor	Empleado	Oficina Recepción	Clasificación	Motivo	Forma Recepción	Lugar Movil	Campaña	Tipo Atención	Etapa Cierre	Tipo Bien	Institucion que toma la denuncia
	Año	Nombre Sector	Zona	Sexo	Zona Departamento	Zona	Nombre sector clasificación	Nombre Motivo	Nombre forma recepcion	Nombre Lugar	Zona	Nombre Tipo Atención	Nombre Etapa	Nombre Tipo Bien	Nombre de la institucion
	Semestre	Nombre Comercial	Departamento	Zona	Departamento de la Oficina	Departamento	categoria clasificación				Departamento		Nombre Subetapa		
	Trimestre	Razon Social	Municipio	Departamento	Municipio de la Oficina	Municipio					Municipio				
	Mes	NIT		Municipio	Oficina del Empleado	Oficina					Nombre Campaña				
	Quincena	Departamento Sucursal			Cargo del Empleado										
	Semana	Municipio Sucursal			Nombre Empleado										
	Día	Nombre Sucursal													
	Hechos Medidos:		Monto Reclamado, Monto Recuperado, Cantidad de Casos (Medida calculada)												

Figura 17.2 Paquete de información para cantidad y montos de los casos cerrados en la DC

17.4 Especificación de Requerimientos

17.4.1 Caso de uso cantidad de atenciones recibidas

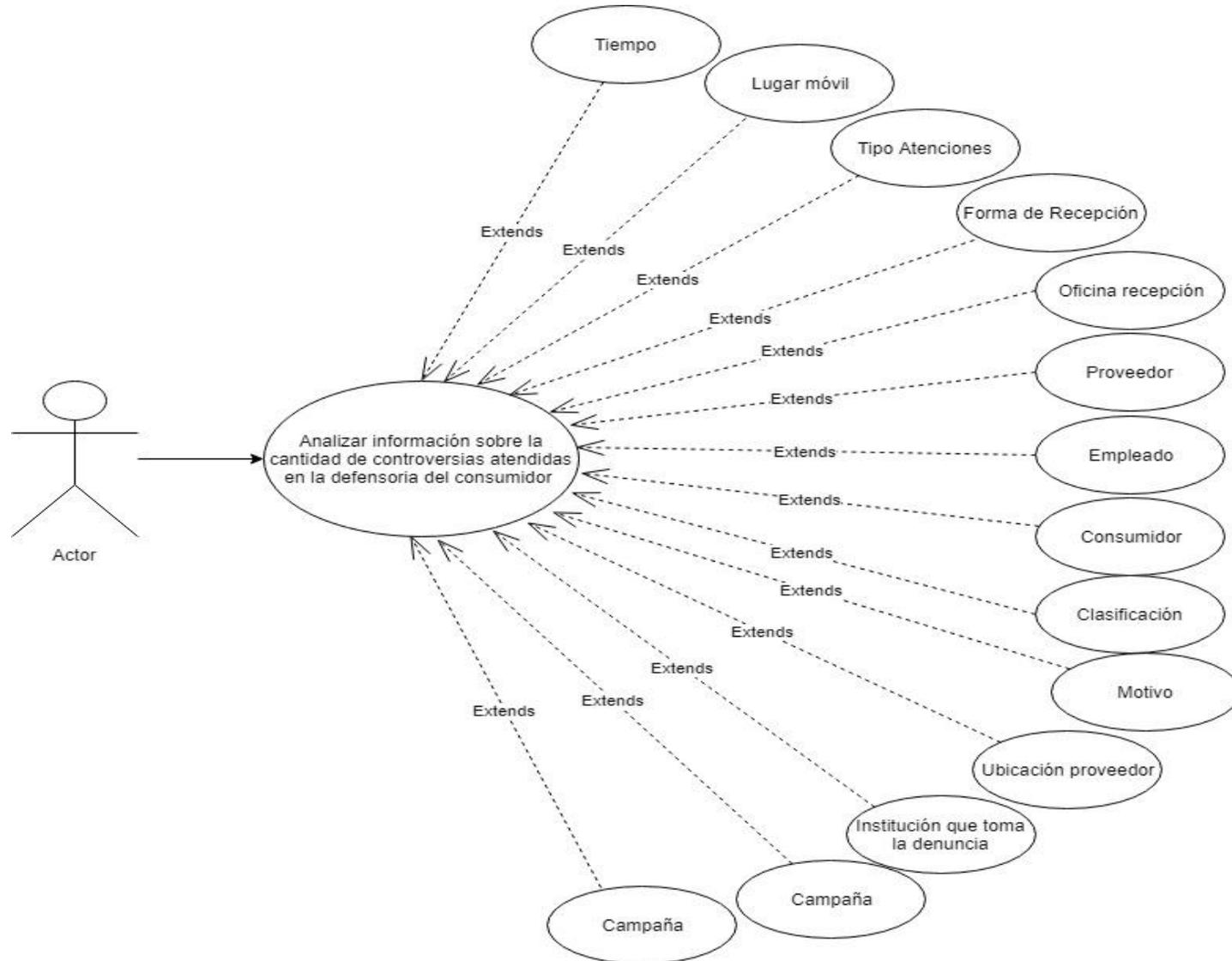


Figura 17.3: Caso de uso cantidad de atenciones recibidas en la defensoría del consumidor

17.4.2 Caso de uso análisis de cantidad y montos de los casos

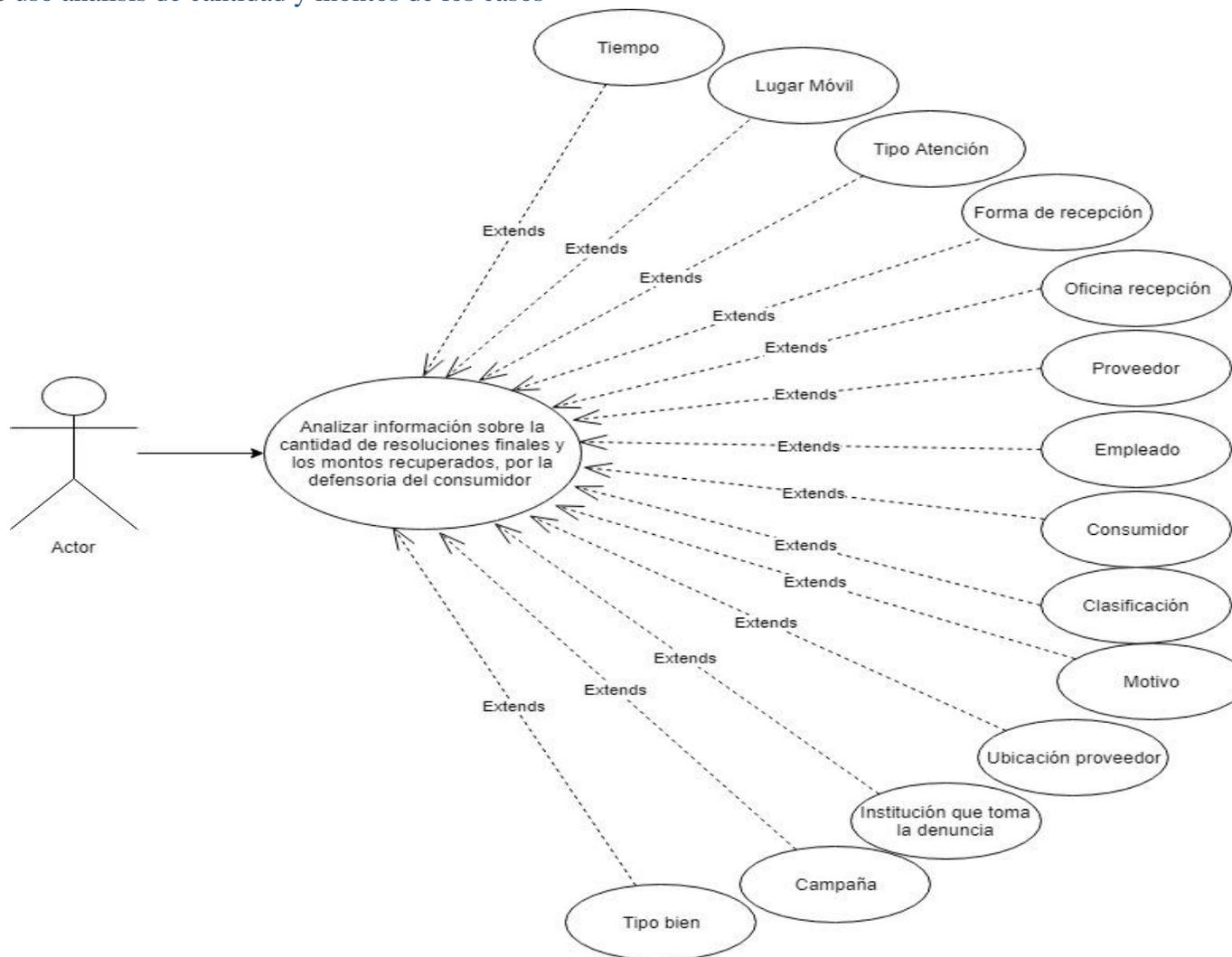


Tabla 17.2: Caso de uso análisis de cantidad y montos de los casos cerrados en la Defensoría del Consumidor

17.5 Diseño Multidimensional

17.5.1 Diseño Conceptual del Data Mart (UML)

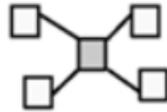
17.5.1.1 Nivel 1: Definición del modelo

Nivel 1:

Unidad: la unidad de Analisis y Consumo de Mercados

Especialización: Analizar las atenciones atendidas.

Hechos principales: Cantidad de atenciones atendidas, Cantidad de casos cerrados



Unidad de análisis y consumo de mercados

Figura 17.4: Nivel 1 Diagrama conceptual Data Mart Unidad de Análisis y Consumo de Mercados

17.5.1.2 Nivel 2: Definición del esquema estrella

Atenciones Recibidas:

Paquete de Hecho: Atenciones recibidas

Paquetes Dimensionales

- | | | |
|-----------------|---------------------|-------------------|
| 1. Oficina | 2. Consumidor | 3. Proveedor |
| 4. Ubicación | 5. Tipo bien | 6. Empleado |
| 7. Institución | 8. Campana | 9. Tiempo |
| 10. Lugar móvil | 11. Forma recepción | 12. Tipo Atención |
| 13. Motivo | 14. Clasificación | |

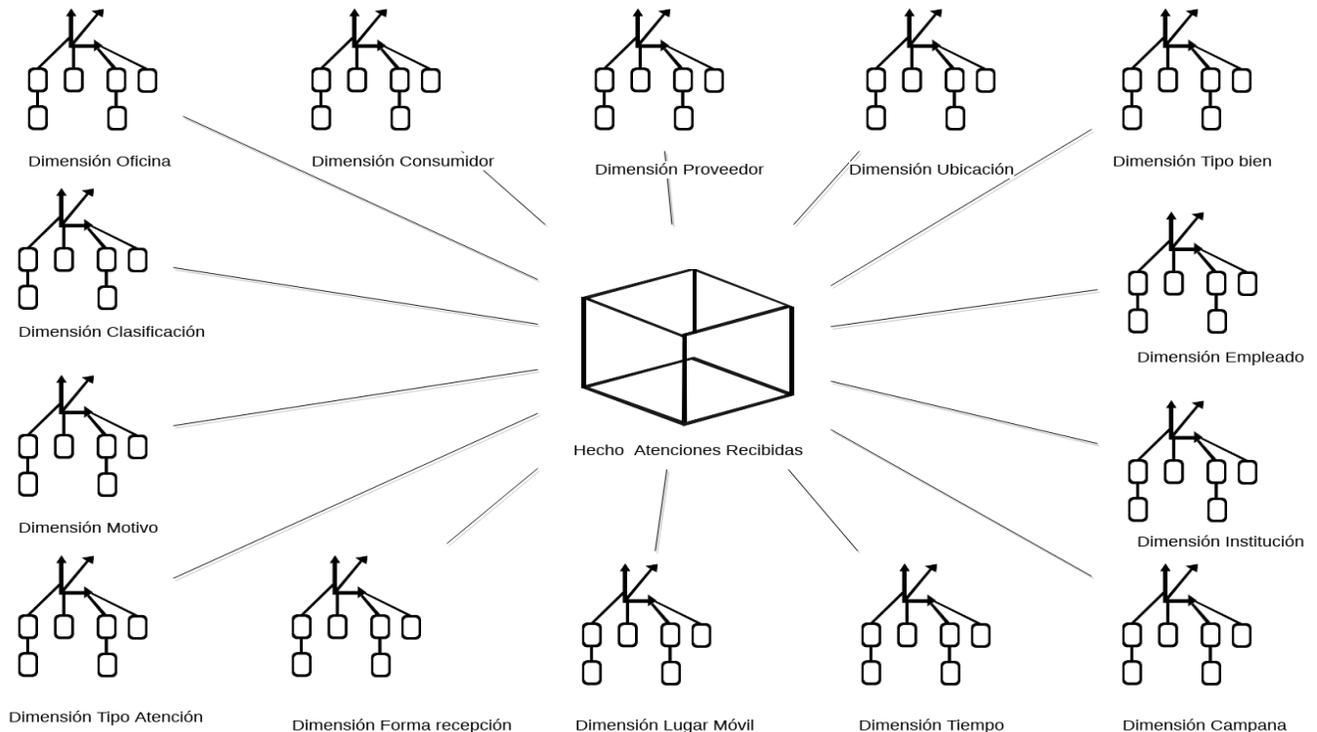


Figura 17.5: Diagrama Conceptual Data Mart (Atenciones recibidas)

Casos cerrados:

Paquete de Hecho: Cantidad de casos cerrados

Paquetes Dimensionales:

- | | | |
|-------------------|-----------------|---------------------|
| 1. Oficina | 2. Consumidor | 3. Proveedor |
| 4. Ubicación | 5. Tipo bien | 6. Campaña |
| 7. Institución | 8. Empleado | 9. Etapa |
| 10. Tiempo | 11. Lugar Móvil | 12. Forma Recepción |
| 13. Tipo Atención | 14. Motivo | 15. Clasificación |

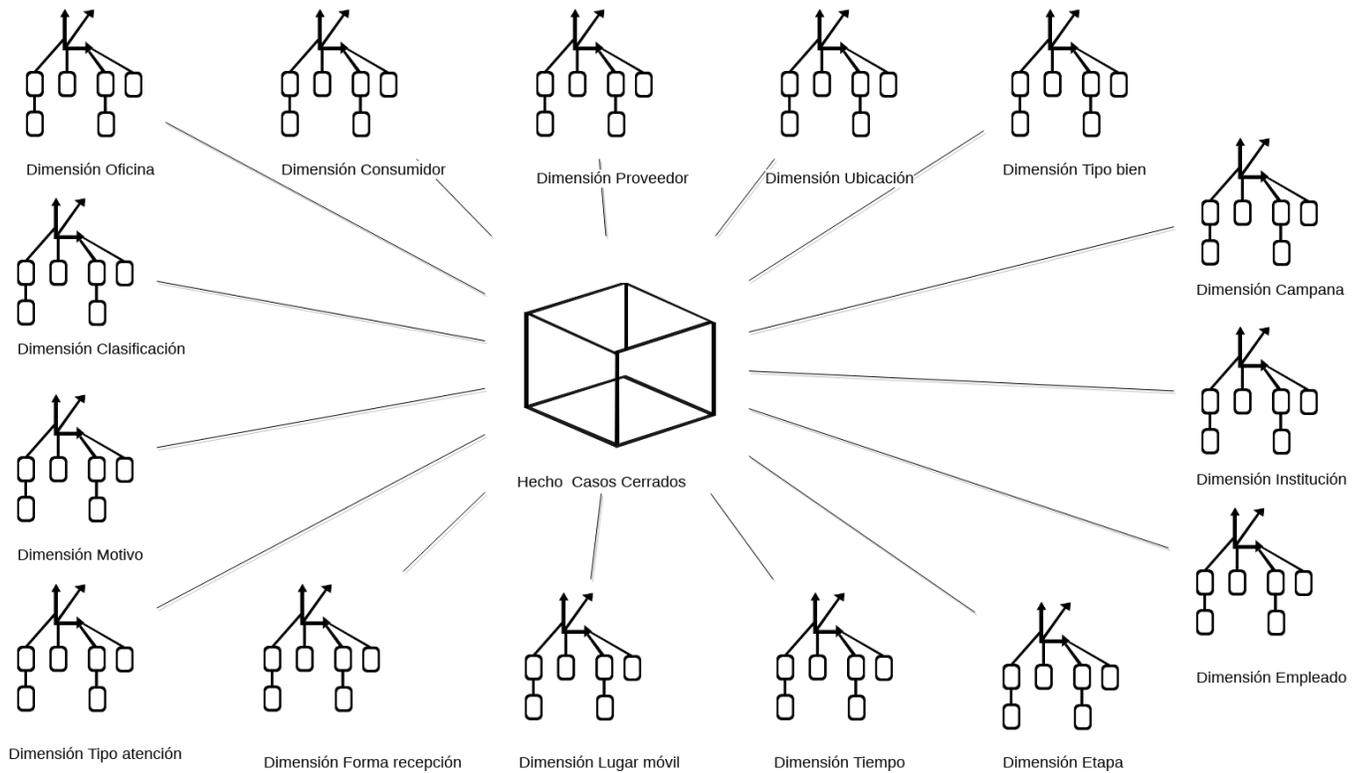


Figura 17.6: Nivel 2 Diagrama Conceptual Data Mart (Casos cerrados)

17.5.1.3 Nivel 3: Dimensiones/Hechos

17.5.1.3.1 Fact Atenciones recibidas

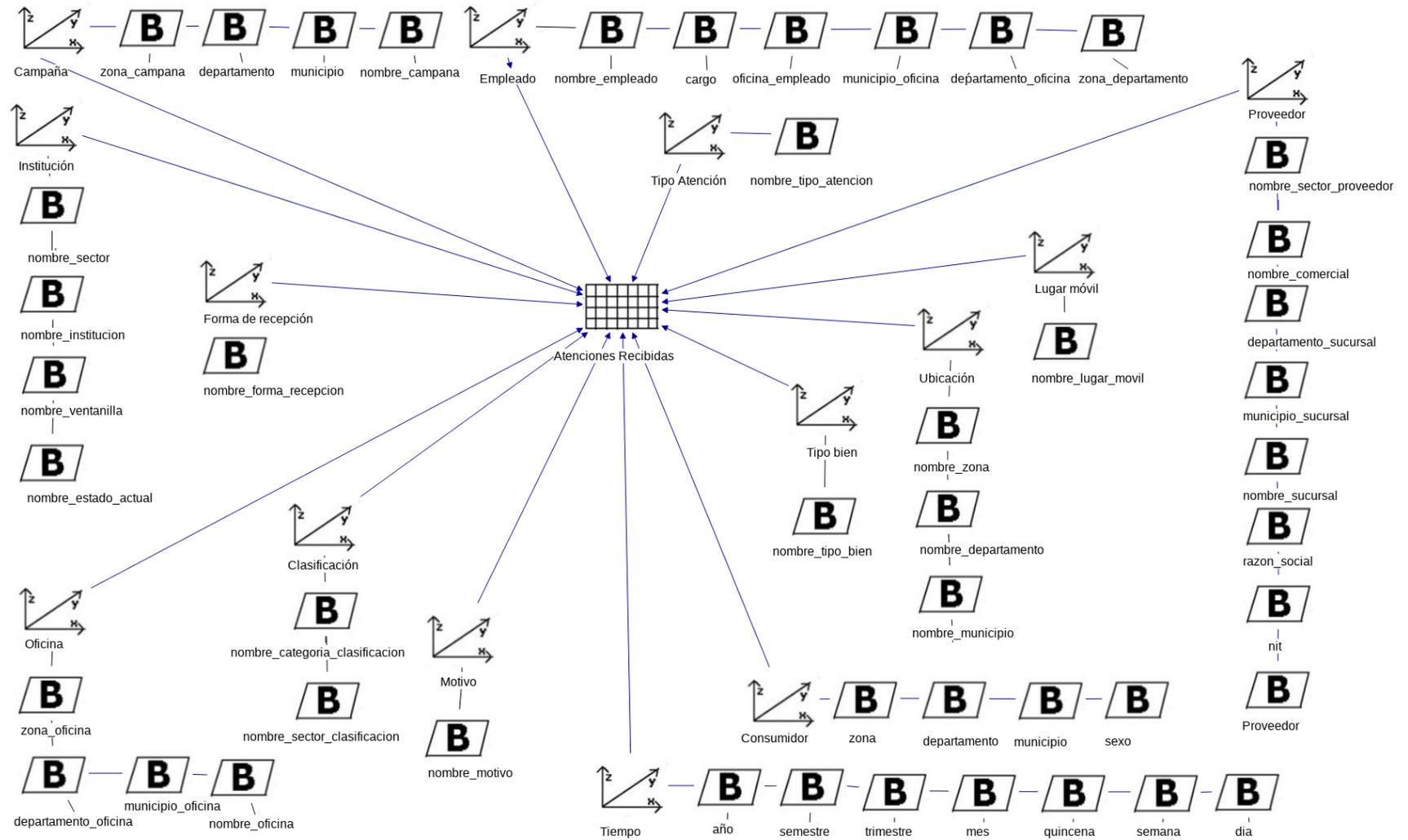


Figura 17.7: Nivel 3 Diagrama conceptual Data Mart (Hechos: Atenciones cerradas)

17.5.1.3.2 Casos cerradas

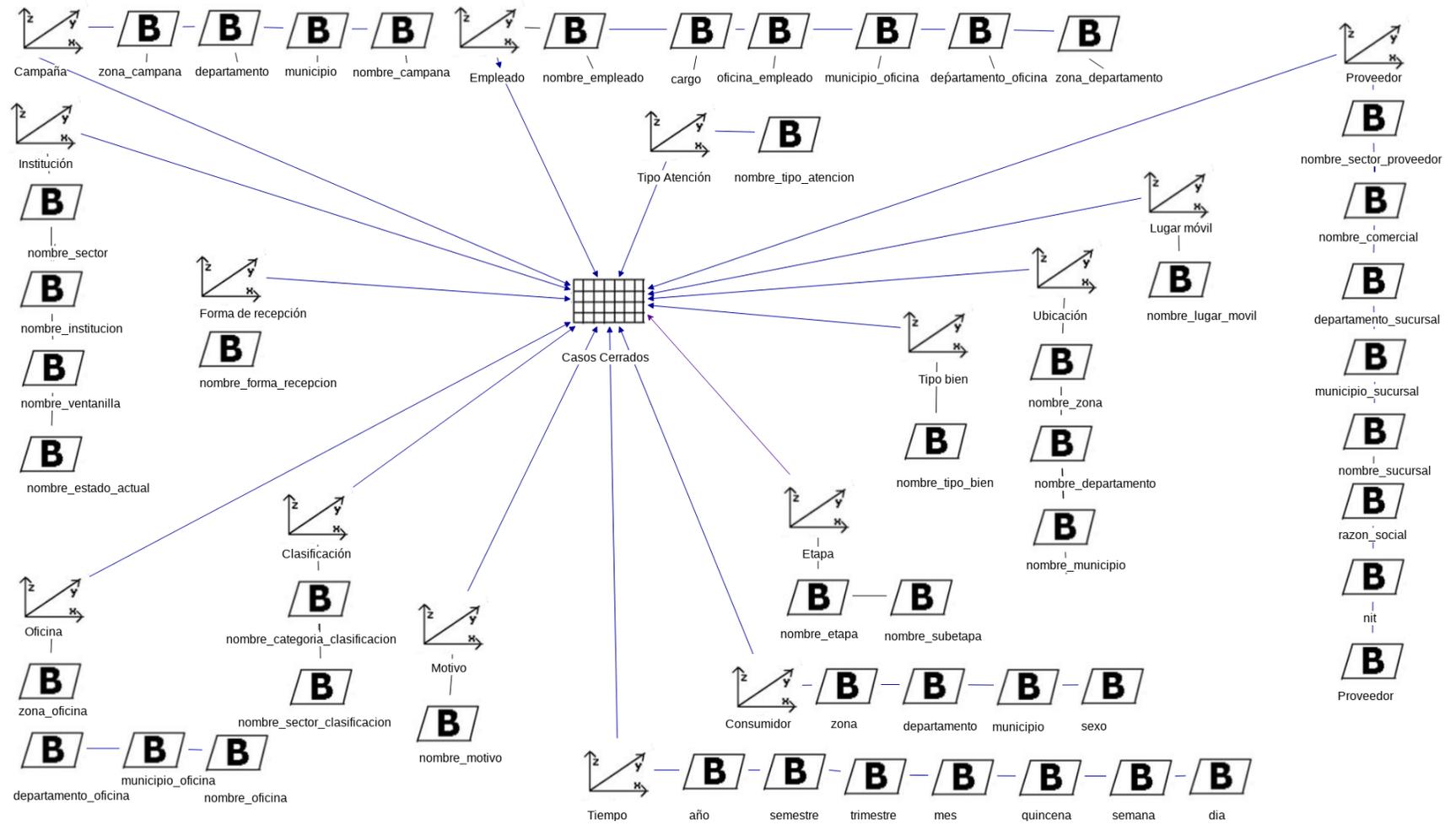


Figura 17.8: Nivel 3 Diagrama conceptual Data Mart (Hechos: casos cerrados)

17.5.2 Diseño Físico Base de Datos Data mart

A continuación, se muestra el diseño físico de base de datos para el data mart de Unidad de Analisis y Consumo de Mercados. Dado que se tienen 2 fact al igual que Dirección de Descentralización también se han separado los diagramas (un diagrama por cada fact table). La Figura 17.9 muestra el diagrama para la fact table atenciones recibidas, mientras que la Figura 17.10 muestra el diagrama para la fact table de casos cerrados.

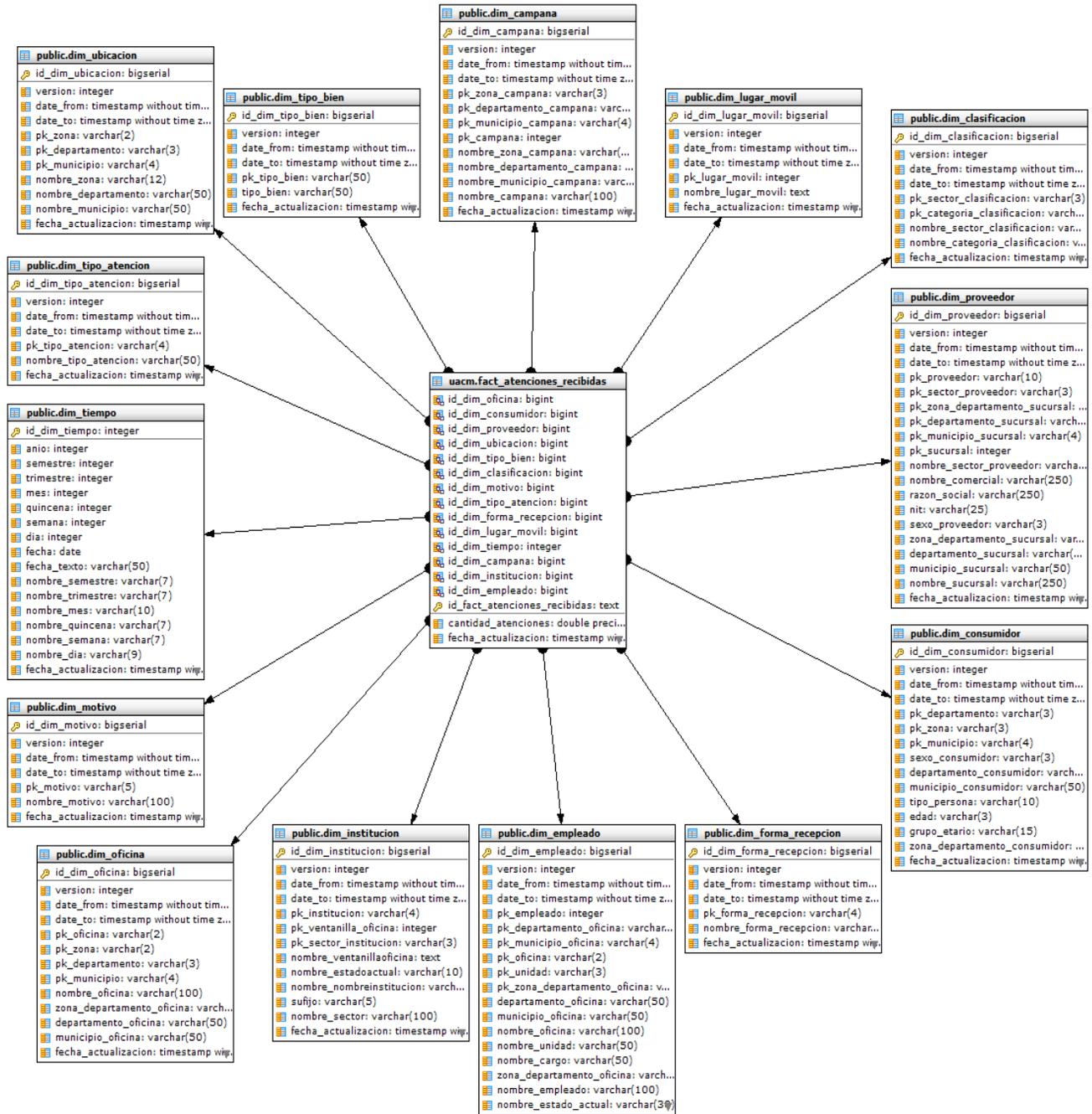


Figura 17.9: Diagrama físico de base de datos, fact table atenciones recibidas, Unidad de Analisis y Consumo de Mercados.

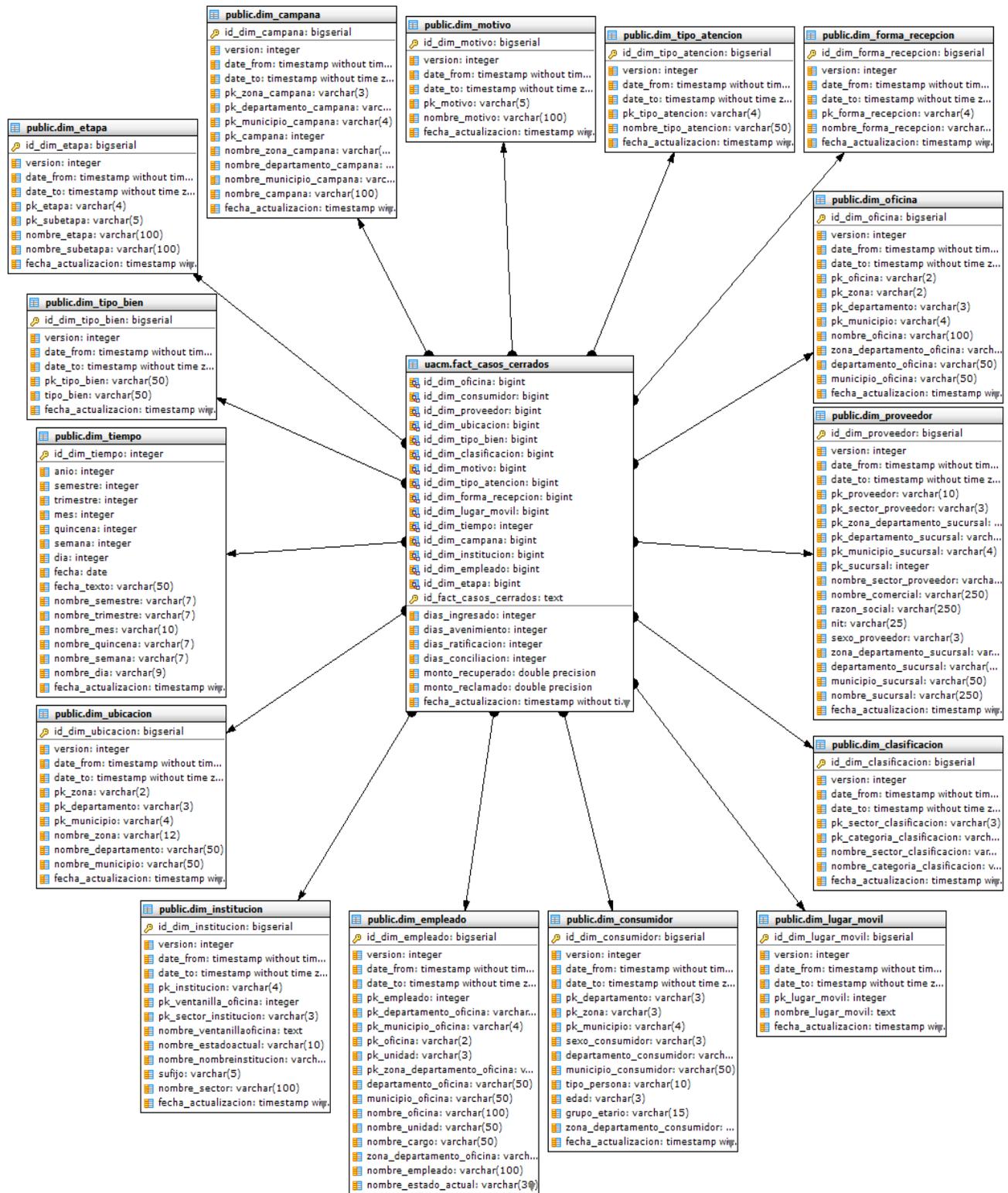


Figura 17.10: Diseño físico de base de datos, fact table casos cerrados, Unidad de Analisis y Consumo de Mercados

17.5.3 Diseño Físico Base de Datos Staging Area

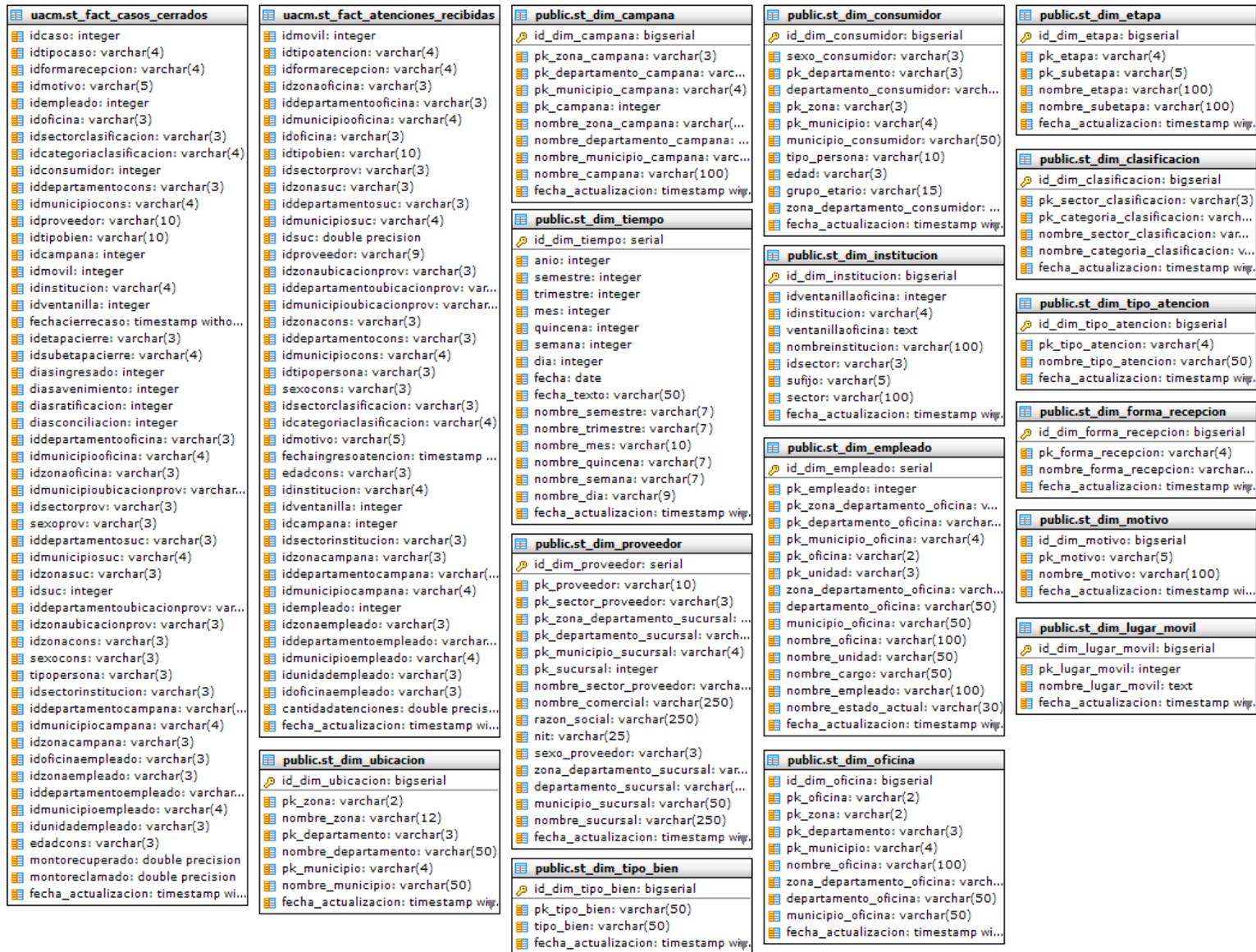


Figura 17.11: Diseño físico de base de datos para staging area de data mart Unidad de Analisis y Consumo de Mercados.

17.6 Diseño de Procesos ETL

17.6.1 Procesos ETL Staging

En los siguientes apartados se listan los procesos ETL de ambos grupos de diagramas, seguido de un diagrama representativo para esta unidad y los diseños de los procesos restantes se colocarán en el CD que acompaña al documento en la carpeta [CD-EISI\Diseños de Procesos ETL\uacm \(Unidad de Analisis y Consumo de Mercados\)](#)

Diagramas desarrollados:

- St_fact_atenciones recibidas
- St_fact_casos_cerrados

17.6.2 Procesos ETL Data Mart

17.6.2.1 Dw_fact_atenciones recibidas

A continuación, se listan los procesos ETL que mueven datos del Staging Area hacia el Data Mart (Procesos ETL Data Mart) desarrollados para esta unidad, seguido de un ejemplo (diagrama) para ilustrar uno de los procesos ETL.

Diagramas desarrollados:

- Dwt_fact_atenciones recibidas
- Dwt_fact_casos_cerrados

A continuación, se muestra uno de los diseños de proceso. Se muestra el diagrama Dw_fact_casos_cerrados

Proceso ETL DW casos cerrados

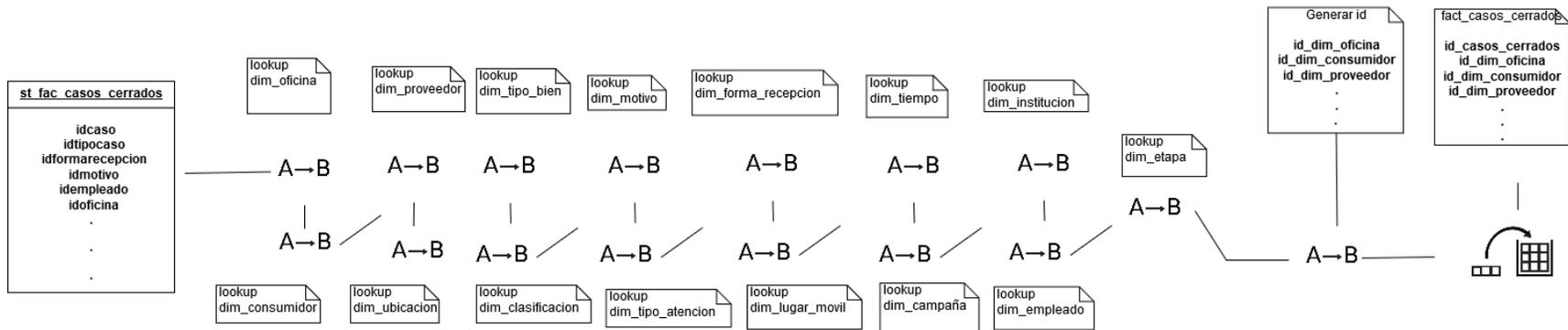


Figura 17.12: Proceso ETL Dw_fact_casos_cerrados

17.7 Pruebas

17.7.1 Pruebas de Procesos ETL

El proceso de pruebas de proceso ETL consistió en ejecutar los Jobs almacenados en el servidor de Staging Área. Las transformaciones finales para el Data Mart de Unidad de Analisis y Consumo de Mercados se muestran en la Tabla 17.3. Se puede visualizar que no hay procesos ETL para llenado de dimensiones esto es porque se han reutilizado las dimensiones existentes en sprints anteriores.

No	Transformación	Descripción
1	dw_incremental_fact_atenciones_recibidas.ktr	Transformación para cargas incrementales desde tabla st_fact_atenciones_recibidas (staging area) hacia tabla fact_atenciones_recibidas (data warehouse).
2	dw_incremental_fact_casos_cerrados.ktr	Transformación para cargas incrementales desde tabla st_fact_casos_cerrados (staging area) hacia tabla fact_casos_cerrados (data warehouse).
3	dw_inicial_fact_atenciones_recibidas.ktr	Transformación para carga inicial desde tabla st_fact_atenciones_recibidas (staging area) hacia tabla fact_atenciones_recibidas (data warehouse).
4	dw_inicial_fact_casos_cerrados.ktr	Transformación para carga inicial del data warehouse desde tabla st_fact_casos_cerrados (staging area) hacia tabla fact_casos_cerrados (data warehouse).
5	st_incremental_fact_atenciones_recibidas.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia tabla st_fact_atenciones_recibidas (staging area).
6	st_incremental_fact_casos_cerrados.ktr	Transformación para cargas incrementales del data warehouse desde tablas fuentes hacia tabla st_fact_casos_cerrados (staging area).
7	st_inicial_fact_atenciones_recibidas.ktr	Transformación para carga inicial del data warehouse desde tablas fuentes hacia tabla st_fact_atenciones_recibidas (staging area).
8	st_inicial_fact_casos_cerrados.ktr	Transformación para carga inicial del data warehouse desde tablas fuentes hacia tabla st_fact_casos_cerrados (staging area).

Tabla 17.3: Transformaciones finales para Data Mart Unidad de Analisis y Consumo de Mercados

Los Jobs finales de Primer Nivel para Data Mart de Unidad de Analisis y Consumo de Mercados se muestran la Tabla 17.4. También muestra el orden en que ejecutan las transformaciones de cada uno de los Jobs, el orden en que aparecen los Jobs se muestra en Tabla 17.5.

No	Job Primer Nivel	Transformación	
		No	Nombre Transformación
1	job_incremental_fact_atenciones_recibidas.kjb	1	st_incremental_fact_atenciones_recibidas.ktr
		2	dw_incremental_fact_atenciones_recibidas.ktr
2	job_incremental_fact_casos_cerrados.kjb	1	st_incremental_fact_casos_cerrados.ktr
		2	dw_incremental_fact_casos_cerrados.ktr
3	job_inicial_fact_atenciones_recibidas.kjb	1	st_inicial_fact_atenciones_recibidas.ktr
		2	dw_inicial_fact_atenciones_recibidas.ktr
4	job_inicial_fact_casos_cerrados.kjb	1	st_inicial_fact_casos_cerrados.ktr
		2	dw_inicial_fact_casos_cerrados.ktr

Tabla 17.4: Jobs de Nivel 1 para Data Mart Unidad de Analisis y Consumo de Mercados

Los Jobs finales de Segundo Nivel para Data Mart de Unidad de Analisis y Consumo de Mercados se muestran la Tabla 17.5. La tabla también muestra el orden en que ejecutan las transformaciones cada uno de los Jobs, los Jobs presentados a continuación reflejan el orden en que deberán ser ejecutados.

No	Jobs Segundo Nivel	Jobs primer nivel ejecutados	
		No	Nombre Job
1	job_inicial_dd.kjb	1	job_inicial_fact_casos_cerrados.kjb
		2	job_inicial_fact_atenciones_recibidas.kjb
2	job_incremental_dd.kjb	1	job_incremental_fact_casos_cerrados.kjb
		2	job_incremental_fact_atenciones_recibidas.kjb

Tabla 17.5: Jobs de Nivel 2 para Data Mart Unidad de Analisis y Consumo de Mercados

La Figura 17.13 y Figura 17.14 muestra los Jobs de segundo nivel los cuales son los Jobs principales que son ejecutados en el servidor. La figura Figura 17.13 muestra la secuencia del job que realiza la carga inicial, mientras que la Figura 17.14. muestra la secuencia del job que servirá para realizar las cargas incrementales.



Figura 17.13: Job ETL para carga inicial del Data Mart Unidad de Analisis y Consumo de Mercados



Figura 17.14: Job ETL para carga incremental del Data Mart Unidad de Analisis y Consumo de Mercados

17.7.1.1 Ejecución de Procesos ETL iniciales

A cotinuacion se muestra en resumen la ejecución de las transformaciones y Jobs desarrollados para para la Unidad de Anaisis y Consumo de Mercados.

Catidad de Transformaciones	Tiempo de ejecución	Cantidad de registros movidos
6	00:11:48.749	381,689

Para más detalle de las pruebas realizadas en el Sprint 6 acceder al documento Release 2.pdf, sección 10.9.1 incluido en el CD que acompaña este documento en la carpeta: CD-EISI\Documentación\

La finalización del proceso ETL se notifica mediante correo electrónico como se muestra en la Figura 17.15.

☐ ☆ ▶ DATAWAREHOUSE	INCREMENTAL ETL FINISHED FOR sprint6 (Unidad de Analisis y Consumo de Mercados) - EL ...	15:10
☐ ☆ ▶ DATAWAREHOUSE 3	INCREMENTAL ETL FINISHED FOR fact_atenciones_recibidas - EL PROCESO ETL INCREMEN...	15:10
☐ ☆ ▶ DATAWAREHOUSE 3	INCREMENTAL ETL FINISHED FOR fact_casos_cerrados - EL PROCESO ETL INCREMENTAL P...	15:09

Figura 17.15: Notificación de Finalización de procesos ETL (Carga Incremental)

17.7.2 Pruebas Cubos OLAP

Dado que para la Unidad de Analisis y Consumo de Mercados se identificaron dos hechos de los cuales se desea analizar información, se han realizado dos cubos para dicha unidad. A continuación, se muestran los cubos realizados para el Data Mart de Unidad de Analisis y Consumo de Mercados.

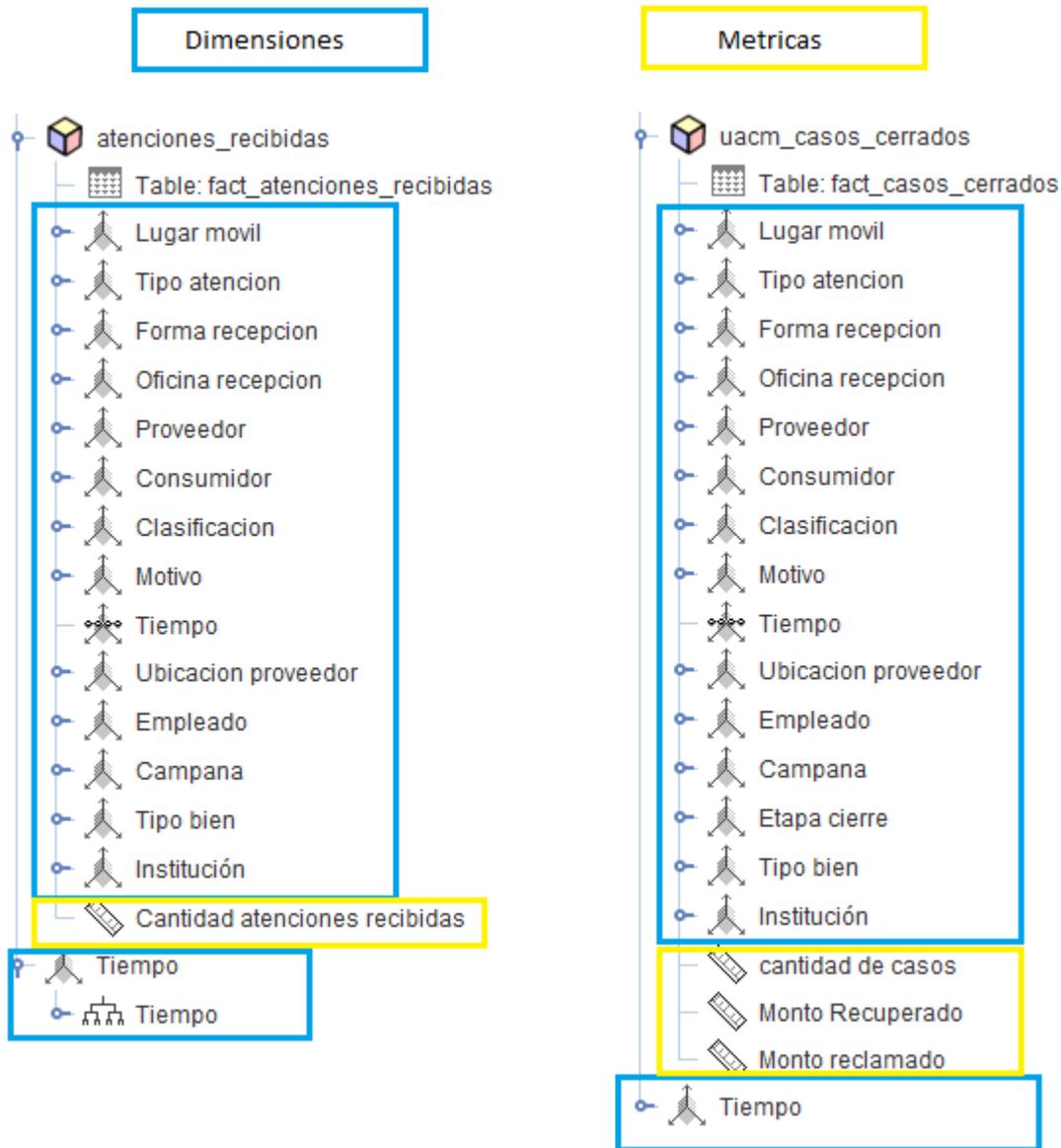


Figura 17.16: Cubos para Unidad de Analisis y Consumo de Mercados.

Por cada uno de los cubos de realizados para esta unidad, se ha publicado un archivo XML en el Servidor de Pentaho Server BI. Con el uso del Plugin JPivot para Pentaho Server se logra visualizar la información de los cubos tal como se muestra en las siguientes imágenes:

					Medidas				
					Cantidad atenciones recibidas				
					Oficina recepcion				
	forma recepcion	Nombre sector	Clasificacion	Consumidor	Tiempo	Todas las oficinas	Central	Occidental	Oriental
iones	Recepcion	Todos los sectores	Todas las clasificaciones	Todos los consumidores	Todos los años	133.665	85.362	27.516	20.787
				F	Todos los años	63.893	40.005	13.557	10.331
					2008	3.968	2.516	876	576
					2009	11.247	6.883	2.437	1.927
					2010	10.879	6.353	2.435	2.091
					2011	4.910	3.057	956	897
					2012	1.636	894	378	364
					2013	9.704	6.231	1.973	1.500
					2014	13.571	8.694	3.117	1.760
					2015	7.958	5.360	1.383	1.215
					2016	4	3		1
					2017	16	14	2	
				M	Todos los años	68.732	44.448	13.884	10.400

Figura 17.17: Visualización del Cubo Atenciones recibidas en la Defensoria del Consumidor.

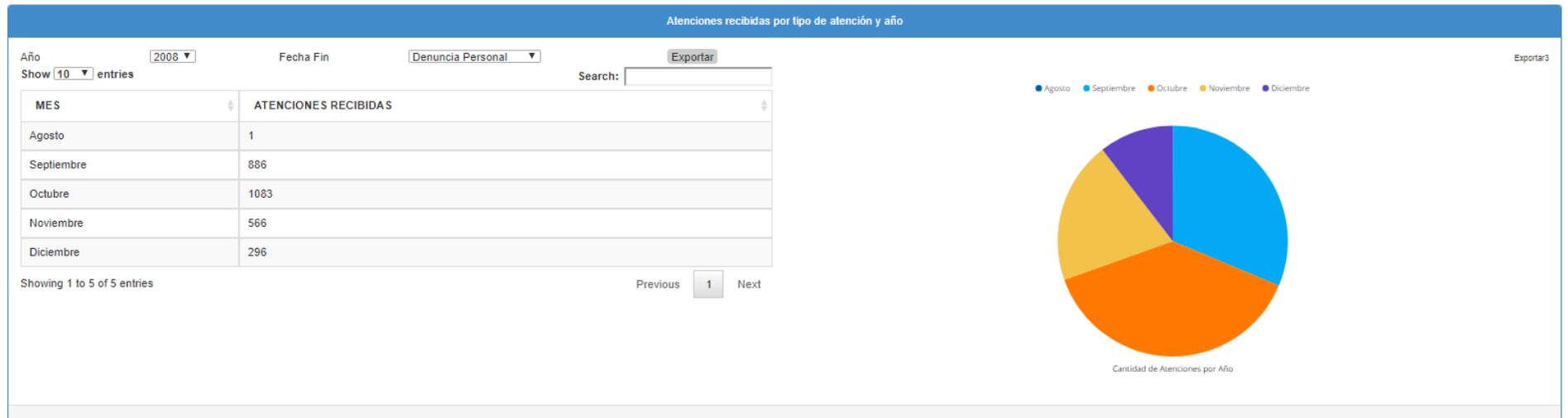
				Ubicacion proveedor		Ubicacion proveedor		Ubicacion provee	
				Todas las ubicaciones		Todas las ubicaciones		Todas las ubicaciones	
				Campana		Campana		Campana	
				Todas las campanas		Todas las campanas		Todas las camp	
				Tipo bien		Tipo bien		Tipo bien	
				Tipos de bien		Tipos de bien		Tipos de bien	
				Etapa cierre		Etapa cierre		Etapa cierre	
				Todas las etapas		Todas las etapas		Todas las etap	
as atenciones	Empleado	Institucion		58.686	5	12.367	46.314	1	
as atenciones	Todos los empleados	Instituciones		2.302		16	2.286	1	
as atenciones	Todos los empleados	Instituciones		27.526	2	159	27.365		
as atenciones	Todos los empleados	Instituciones		65		38	27		
as atenciones	Todos los empleados	Instituciones		4.897		3.776	1.121		
as atenciones	Todos los empleados	Instituciones		1.728		1.374	354		
as atenciones	Todos los empleados	Instituciones		108		6	102		
as atenciones	Todos los empleados	Instituciones		15		2	13		

Figura 17.18: Visualización del Cubo Casos Cerrados por Defensoria del Consumidor.

17.7.3 Pruebas Dashboard

A continuación, se presenta uno de los dashboards desarrollados para la Unidad de Análisis y Consumo de Mercado. Se muestra el dashboard de atenciones recibidas

17.7.3.1 Prueba de Dashboard Atenciones Recibidas



18 ENTREGABLES

La lista de entregables del proyecto se describe en la

Tabla 18.1.

ENTREGABLE	DESCRIPCIÓN
Esquema de base de datos Staging Area	Imágenes de los esquemas de base de datos correspondientes a cada una de las unidades.
Esquema de base de datos.Data Warehouse	Imágenes de los esquemas de base de datos correspondientes a cada una de las unidades.
Archivos fuentes de Pentaho BI Server	Archivos de instalación de Pentaho BI server personalizados (logo, colores, etc).
Script de base de datos Staging Area	Script SQL para creación de esquema de base de datos de staging área.
Script de base de datos Data Warehouse	Script SQL para creación de esquema de base de datos de data warehouse.
Manual de despliegue	Manual donde se encuentran las indicaciones para realizar el despliegue de los servidores, y que hacer con los archivos fuentes y entregables.
Manual de usuario	Manual donde se explica el uso de Pentaho BI Server.
Manual técnico	Manual donde se explican los aspectos técnicos de el data warehouse o modelo multidimensional. En el se encuentran diferentes diagramas técnicos, historias de usuario, etc.
Manual de instalación	Manual donde se explica únicamente la instalación de las diferentes herramientas como SQL server, Pentaho BI Server, PostgreSQL, etc.
Dashboard de las unidades	Archivos de dashboards para cada unidad.

Tabla 18.1: Entregables del proyecto

19 CONCLUSIONES

- En definitiva, el principal uso del Modelo Multidimensional dentro de la Defensoría del Consumidor será el auto servicio de la información por parte de las Unidades de la Institución mejorando de esta manera el proceso de obtención, procesamiento y análisis de la información que la institución genera y facilitando de esta forma la toma de decisiones de las diferentes Gerencias, Jefaturas y Direcciones involucradas.
- El Modelo Multidimensional para el Sistema de Atención a Reclamos y Asesorías para la Defensoría del Consumidor finaliza con la construcción y documentación de 6 data marts correspondientes a las unidades Dirección Jurídica (DJ), Dirección de Ciudadanía y Consumo (DCC), Tribunal Sancionador (TS), Dirección del Centro de Solución de Controversias (DCSC) Dirección de Descentralización (DD) y Unidad de Análisis y Consumo de Mercados (UACM).
- En cada uno de los data marts se dio respuesta a diferente cantidad de requerimientos principales de cada unidad, siendo 2 requerimientos para DJ, 1 requerimiento para DCC, 3 requerimientos para TS, 3 requerimientos DCSC, 2 requerimientos para DD y 2 requerimientos para UACM, dando solución en total a 13 grandes requerimientos de información para la Defensoría del Consumidor, mejorando de esta manera el proceso de obtención, procesamiento y análisis de la información que la institución genera y facilitando de esta forma la toma de decisiones de las diferentes Gerencias, Jefaturas y Direcciones involucradas.
- El Modelo Multidimensional desarrollado, también marca las bases para nuevas formas novedosas de explotación de información como lo es la minería de datos. A medida que las tablas de hechos de la base de datos del Modelo Multidimensional crezcan con el tiempo, la explotación de manera automática o semi automática mediante el uso de algoritmos estadísticos de búsqueda próximos a la Inteligencia Artificial y a las redes neuronales brindará información importante que puede ser utilizada para saber por ejemplo, donde enfocar los esfuerzos de la Defensoría del Consumidor dentro del país, cuando doblar recurso humano debido a mayor afluencia de consumidores, que parte de la población debería ser capacitado en ciertas áreas de consumo, que sector de comercio debería ser educado en materia de consumo, entre otros.
- Se espera que la Defensoría del Consumidor pueda servir como ejemplo para las demás instituciones del estado y que estas logren ver los grandes beneficios que ofrecen las soluciones de Inteligencia de Negocios y como pueden no solo aplicarse al sector privado, si no también como las instituciones publicas pueden beneficiarse de ellas independientemente del rubro al cual pertenecen y las funciones para las cuales han sido creadas.

20 RECOMENDACIONES

- Se recomienda superar los requerimientos mínimos de hardware de los servidores a la hora de implementar la arquitectura de servidores en un ambiente de producción.
- Hacer uso de dispositivos que agilicen la lectura de las bases de datos, tales como Discos Duros de Estado Solido.
- Agregar validaciones a los sistemas fuentes para impedir que se guarde información no útil o con inconsistencias en los bases de datos fuentes.
- Antes de implementar en producción, se recomienda realizar pruebas de los procesos ETL con datos reales para realizar las respectivas modificaciones en procesos ETL ante cualquier posible situación que no se haya considerado o que no haya surgido con los datos de prueba otorgados en la elaboración del proyecto.
- Se recomienda mantener la lógica de un data mart diferente por cada unidad organizacional y separar dicho data mart en esquemas de bases de datos aún cuando los requerimientos de información sean iguales. Esto con el objetivo de lograr escalabilidad cuando una unidad organizacional solicite cambios.
- Asegurarse de forma periodica mediante la lectura de los correos de confirmación sobre la actualización del data warehouse mediante las cargas incrementales programadas periódicamente.
- Programar la ejecución de cargas incrementales en horas no hábiles, para que no afecte las transacciones del Sistema SARA.
- Capacitar y proveer los respectivos manuales a los usuarios (de las unidades organizacionales) en cuanto al uso de la herramienta Pentaho BI Server para aprovechar las bondades de la herramienta.

21 BIBLIOGRAFÍA

- Álvarez, T. V. (23 de Marzo de 2014). Obtenido de Comunidad IEBS: <https://comunidad.iebschool.com/metodologiasparaelcambio/2014/03/23/priorizando-historias-de-usuario/>
- Análisis de Brecha. (2017). En S. Study, *Una guía para el cuerpo de conocimiento de Scrum* (pág. 150). Estados Unidos: VMEdU, Inc.
- Análisis FODA. (2017). En S. Study, *Una guía para el Cuerpo de Conocimiento de Scrum* (pág. 150). Estados Unidos: VMEdU, Inc.
- Dertiano, V. (02 de Mayo de 2018). *BIGeek*. Obtenido de Comparativa entre Inmon y Kimball: <http://blog.bi-geek.com/arquitectura-comparativa-inmon-y-kimball/>
- Garzas, J. (2 de Octubre de 2013). Obtenido de Javiergarzas.com: <http://www.javiergarzas.com/2013/10/wsjf-safe.html>
- Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit*. Canada: John Wiley & Sons, Inc.
- Learn data modeling*. (15 de Abril de 2015). Obtenido de <https://learndatamodeling.com/blog/tag/what-is-dolap/>
- Mantilla, J. (2011). *docshared*. Obtenido de <http://docshare02.docshare.tips/files/23619/236195121.pdf>
- Marco Institucional Defensoría del Consumidor*. (s.f.). Obtenido de Defensoría del Consumidor: <https://www.defensoria.gob.sv/institucion/marco-institucional/>
- Pichler, R. (09 de January de 2009). *Scrum Alliance*. Obtenido de The Product Vision: <https://www.scrumalliance.org/community/articles/2009/january/the-product-vision>
- Ponniah, P. (2010). *Data Warehousing Fundamentals For IT Professionals*. New Jersey: John Wiley & Sons, Inc.
- PONNIAH, P. (2010). *DATA WAREHOUSING FUNDAMENTALS FOR IT PROFESSIONALS*. New Jersey: JHON WILEY & SONS, INC.
- Power Data. (6 de Agosto de 2013). *La Limpieza de Datos: la etapa previa a los Procesos ETL*. Obtenido de <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/312597/la-limpieza-de-datos-la-etapa-previa-a-los-procesos-etl>
- Power Data. (16 de Julio de 2013). *Procesos ETL: Extracción*. Obtenido de <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/312587/Procesos-ETL-Extracci-n-En-qu-consiste>
- Power Data. (s.f.). *Data Warehouse: todo lo que necesitas saber sobre almacenamiento de datos*. Obtenido de Power Data: <https://www.powerdata.es/data-warehouse>
- Priya, G., & Sultana, A. R. (s.f.). *ETL Process in Data Warehouse*.

- Rivadera, G. R. (s.f.). *UCASAL*. Obtenido de La metodología de Kimball para el diseño de almacenes de datos: <http://www1.ucasal.edu.ar/htm/ingenieria/cuadernos/archivos/5-p56-rivadera-formateado.pdf>
- Satpathy, T. (2017). *A Guide to the Scrum Body Of Knowledge (SBOOK Guide)*. Arizona: VMEdU, Inc.
- Satpathy, T. (2017). *Una Guía para el cononocimiento de SCRUM (Guía SBOK) 3ra Edición*. Arizona: SCRUMStudy.
- SBook. (2017). *A guide to the Scrum Body of Knowledge*. Avondale, Arizona: SCRUMstudy.
- Schwartz, F. (s.f.). Obtenido de scrumcolombia.org: <https://scrumcolombia.org/refinamiento-del-product-backlog/>
- Sinergia e Inteligencia de Negocio S.L.* (2016). Obtenido de Sinnexus: http://www.sinnexus.com/business_intelligence/olap_vs_oltp.aspx
- Sinergia e Inteligencia de Negocio S.L.* (2016). Obtenido de Sinnexus: http://www.sinnexus.com/business_intelligence/olap_avanzado.aspx
- Trujillo, S. L. (2002). En S. L. Trujillo, *Extending UML for multidimensional Model*. Alicante, España: Alicante.
- WikiEQI. (s.f.). *Wiki EQI*. Obtenido de GESTIÓN DE RIESGOS en Gestión de proyecto: http://www.eoi.es/wiki/index.php/GESTI%C3%93N_DE_RIESGOS_en_Gesti%C3%B3n_de_proyectos
- Wikipedia. (18 de Noviembre de 2017). *Caja negra (sistemas)*. Obtenido de [https://es.wikipedia.org/wiki/Caja_negra_\(sistemas\)](https://es.wikipedia.org/wiki/Caja_negra_(sistemas))
- Wikipedia. (26 de Octubre de 2017). *Data mart*. Obtenido de https://es.wikipedia.org/wiki/Data_mart
- Wikipedia. (4 de Marzo de 2018). *Operational data store*. Obtenido de https://en.wikipedia.org/wiki/Operational_data_store

22 GLOSARIO

A

Asesoría
Circunstancias en la que una persona busca apoyo en otra para aclarar algo, 8

Atención
Entiendase como el servicio que brinda una institución a la ciudadanía, 8

C

Contingenciales
Lo que puede suceder o no, 10

Cubo de información
Base de datos multidimensional, 7

CyC
Ciudadanía y Consumo, 40

CSC
Centro de Solución de Controversias, 40

D

Denuncias
Notificación que se hace a la autoridad de que se ha cometido un delito o de que alguien es el autor de un delito, 9

Diagramas
Representación gráfica de las variaciones de un fenómeno o de las relaciones que tienen los elementos o las partes de un conjunto, 8

Diagramas
Base de datos corporativa que integra información de una o más fuentes, 10

Drill Down
Entiendase como especificar un concepto, 41

DD
Dirección de Descentralización, 40

DJ
Dirección Jurídica, 40

H

Hallazgo
Cosa material o inmaterial que se halla o descubre, 9

I

Incumplimientos
Falta de cumplimiento, 9

Inspección
Acto de examinar o reconocer una cosa para buscar si cumple con parámetros, 9

J

Jobs
Entiendase como una tarea programada, 75

P

Proceso ETL
Proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base, 48

R

Resoluciones
Determinación definitiva que se toma en lo referente a una cuestión problemática, 9

Roll Up
Entiendase como Generalizar un concepto, 58

S

Staging Area
Área para limpieza de datos, 56

T

TS
Tribunal Sancionador, 40

U

UACM
Unidad de Análisis y Consumo de Mercado, 40

23 ANEXOS

23.1 Anexo 1: Visión del producto.

THE PRODUCT VISION BOARD

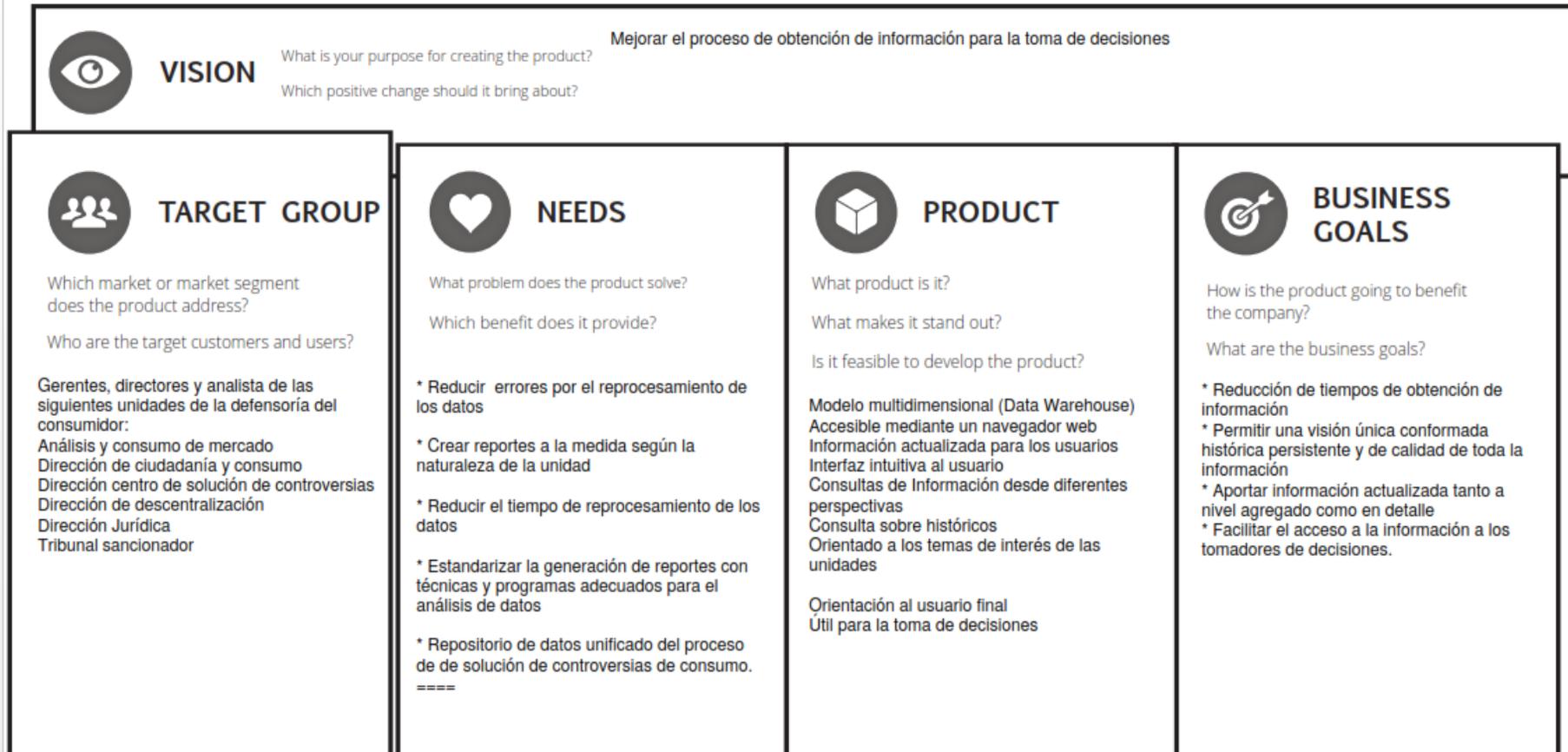


Figura 23.1: Visión del producto.

23.2 Anexo 2: Documentación de Excepciones.

N°	Campo	Tipo	Valor	Valor por defecto	Transformación	Descripción
1	IdSucursal	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
2	IdDepartamentoSuc	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
3	IdMunicipioSuc	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
4	IdZonaSuc	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
5	IdMovil	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
6	IdMunicipioUbicaciónProv	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
7	IdDepartamentoUbicacionProv	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
8	EdadCons	String	NULL / 0	N/A	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con N/A
9	IdZonaUbicacionProv	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
10	IdMunicipioCons	String	NULL / 0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
11	SexoCons	String	NULL / 0	N/A	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con N/A

12	IdSectorClasificacion	String	NULL /0	S00	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con S00
13	IdCategoriaClasificacion	String	NULL /0	C000	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con C000
14	SexoProv	String	NULL /0	N/A	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
15	IdFormaRecepcion	String	NULL /0	FR00	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con FR00
16	TipoBien	String	NULL /0	-1	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
17	IdSectorProv	String	NULL /0	S00	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con -1
18	IdMotivo	String	NULL /0	MO000	St_inicial_fact_atencion	Si el valor del atributo es NULL o 0 será reemplazado con MO000
19	Sufijo	String	NULL	A o B	St_inicial_fact_resolucion	Se agrega un sufijo para diferenciar de qué fuente proviene el proveedor. A: Proveniente de Vigilancia y Mercado. B: Proveniente de Centro de Solución de Controversias. Posteriormente se concatena ese sufijo al IdProveedor.
20	IdTipoResolución	String	NULL /0	-1	St_inicial_fact_resolucion	Si el valor del atributo es NULL o 0 será reemplazado con -1
21	IdTipo Establecimiento	String	NULL /0	-1	St_inicial_fact_resolucion	Si el valor del atributo es NULL o 0 será reemplazado con -1

22	IdZonaCampana	String	NULL /0	-1	St_inicial_fact_resolucion	Si el valor del atributo es NULL o 0 será reemplazado con -1
----	---------------	--------	---------	----	----------------------------	--

