

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**



**Facultad de Ingeniería en Electricidad y Computación**

**Maestría En Sistemas De Información Gerencial**

**“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA DE SOPORTE  
A LA DECISIÓN (DSS) PARA OPTIMIZAR EL ANÁLISIS DE LOS  
DATOS DE VENTAS Y PRODUCCIÓN DE UNA EMPRESA  
MANUFACTURERA”**

**TRABAJO DE TITULACIÓN**

Previa a la obtención del título de

**MAGISTER EN SISTEMAS DE INFORMACIÓN GERENCIAL**

**MÓNICA ARACELY ROJAS CUEVA**

**GUAYAQUIL – ECUADOR**

**2017**

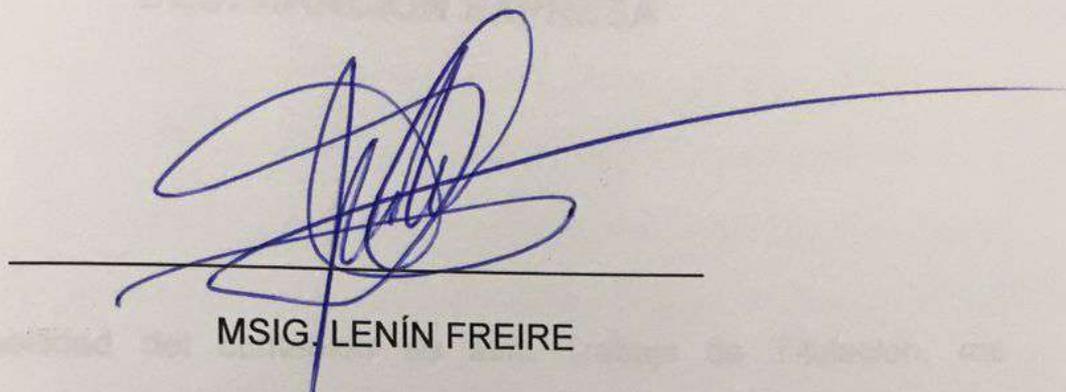
## **AGRADECIMIENTO**

Agradezco a todas las personas que contribuyeron para alcanzar esta meta. A mis profesores, asesor de trabajo de titulación, compañeros, familia y amigos. El apoyo de cada uno de ellos ha sido esencial para culminar con éxito esta experiencia.

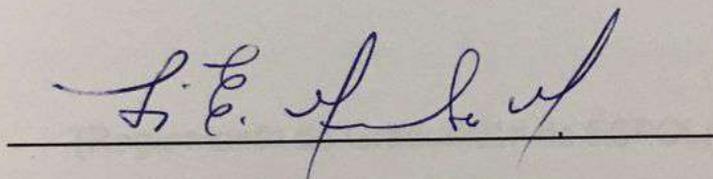
## **DEDICATORIA**

Dedico este trabajo a mi familia, quienes me han impulsado a superarme y a alcanzar las metas que me he propuesto. También a aquellas personas que me han brindado su amistad y apoyo incondicional para cumplir mis objetivos.

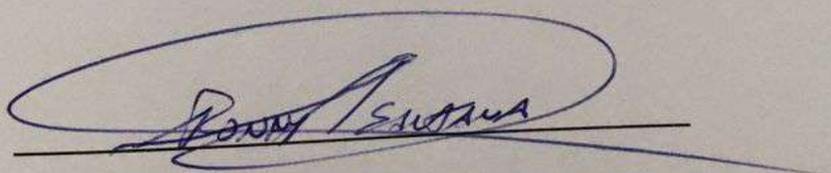
**TRIBUNAL DE SUSTENTACIÓN**

A large, stylized handwritten signature in blue ink, consisting of several overlapping loops and a long horizontal stroke extending to the right. It is positioned above a horizontal line.

**MSIG. LENÍN FREIRE**  
**DIRECTOR MSIG**

A handwritten signature in blue ink, appearing to be 'L. E. Mendoza', written in a cursive style above a horizontal line.

**Ph.D. LUIS MENDOZA**  
**DIRECTOR DEL PROYECTO DE GRADUACIÓN**

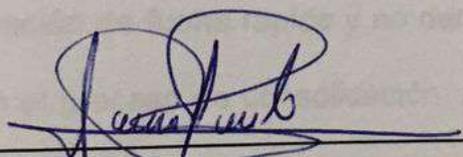
A handwritten signature in blue ink, appearing to be 'Ronny Santana', written in a cursive style above a horizontal line.

**MSIG. RONNY SANTANA**  
**MIEMBRO DEL TRIBUNAL**

## DECLARACIÓN EXPRESA

“La responsabilidad del contenido de este Trabajo de Titulación, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”.

(Reglamento de Graduación de ESPOL)



---

Mónica Aracely Rojas Cueva

## RESUMEN

Toda empresa que desee ser competitiva debe contar con información útil y oportuna que facilite la toma de decisiones. Las empresas de manufactura no son la excepción. Con información histórica y actual podrán establecer mejores estrategias para mejorar sus procesos y por consiguiente lograr mayor satisfacción en sus clientes.

La empresa que será objeto de esta implementación se dedica a la fabricación de soluciones de empaque bajo pedido. Esta empresa no contaba con fácil acceso a la información, en especial en el área comercial. Los informes eran elaborados manualmente, en base a los reportes que ofrece el sistema transaccional. Esto conlleva a no tener la información de forma rápida y no del todo confiable debido a que podían existir errores en el proceso de consolidación. La gerencia necesita de esta información para establecer sus estrategias de ventas.

Para solventar los problemas de información de esta empresa, siguiendo la metodología Scrum, se diseñó e implementó un sistema de soporte a la decisión orientado a datos que abarca el área comercial. Este sistema permite contar con información histórica y actual de las ventas, que puede ser analizada desde varias perspectivas partiendo de lo general a lo específico. Además, se tienen indicadores que facilitan la comprensión de los resultados de la empresa.

La limitante más importante que presenta la solución implementada es que la misma no está integrada con la información de inventarios y cuentas por cobrar, de tal manera que proporcione información integral de la compañía para la toma de decisiones. Sin embargo, este trabajo abre el camino para la creación de almacenes de datos en dichas áreas.

## ÍNDICE GENERAL

AGRADECIMIENTO .....	I
DEDICATORIA .....	II
TRIBUNAL DE SUSTENTACIÓN .....	III
DECLARACIÓN EXPRESA .....	IV
RESUMEN.....	V
ÍNDICE GENERAL .....	VII
ÍNDICE DE FIGURAS.....	XII
ÍNDICE DE TABLAS.....	XVI
INTRODUCCIÓN.....	XVII
CAPÍTULO 1.....	1
GENERALIDADES .....	1
1.1    Antecedentes .....	1
1.2    Descripción del Problema.....	3
1.3    Solución Propuesta .....	5
1.4    Objetivo General.....	8
1.5    Objetivos Específicos .....	8
CAPÍTULO 2.....	9
MARCO TEÓRICO .....	9
2.1    Definición de DSS .....	9

2.2	Inteligencia de Negocios.....	10
2.3	Componentes de un DSS .....	12
2.4	Tipos de DSS .....	14
2.5	Características de los DSS Orientado a Datos .....	15
2.6	Arquitectura de un Sistema BI .....	17
2.7	Modelamiento Dimensional.....	19
2.8	Tipos de Modelamiento .....	20
CAPÍTULO 3.....		22
DEFINICIÓN DE LA SITUACIÓN ACTUAL.....		22
3.1	Situación Actual.....	22
3.2	Levantamiento de Información.....	25
CAPÍTULO 4.....		28
ANÁLISIS Y DISEÑO DE LA SOLUCIÓN.....		28
4.1	Metodología.....	28
4.2	Identificación de Requerimientos .....	31
4.3	Definición del Alcance .....	34
4.4	Diseño de la solución.....	35
4.5	Identificación del modelo relacional .....	37
4.6	Definición de las dimensiones y medidas .....	40
4.7	Diseño del modelo dimensional .....	43

4.7.1	Diseño del Cubo OLAP de Pedidos .....	44
4.7.2	Diseño del Cubo OLAP de Ventas.....	45
4.7.3	Diseño del Cubo OLAP de Entregas.....	46
4.7.4	Diseño del Cubo OLAP de Ingresos .....	47
CAPÍTULO 5.....		49
DESARROLLO DE LA SOLUCIÓN .....		49
5.1	Creación del proceso ETL .....	49
5.1.1	Descripción de los pasos del proceso ETL .....	50
5.1.2	Creación de las transformaciones de las tablas de dimensiones .....	57
5.1.2.1	Transformación ETL dimensión Producto.....	58
5.1.3	Creación de las transformaciones de las tablas de hechos .....	69
5.1.3.1	Transformación ETL tabla hechos Ventas .....	70
5.1.4	Creación de trabajo para las cargas periódicas .....	82
5.2	Creación del cubo OLAP .....	84
5.2.1	Creación del Esquema Comercial.....	85
5.2.2	Creación de las dimensiones compartidas.....	86
5.2.3	Creación de cubo OLAP de Ventas .....	97
5.2.4	Creación de cubo virtual de Ventas e Ingresos.....	105
5.3	Explotación de datos .....	107
5.3.1	Creación de vistas de análisis .....	107

5.3.1.1	Vistas de análisis de las Ventas .....	108
5.3.1.2	Vistas de análisis de comparativos de Ventas y Producción.....	112
5.3.1.3	Vistas de análisis de comparativos de Pedidos por Periodo .....	113
5.3.1.4	Vistas de análisis de Entregas.....	114
5.3.2	Creación de cuadros de mandos .....	117
5.3.2.1	Cuadro de Mando de Ventas.....	117
5.3.2.2	Cuadro de Mando de Toneladas Vendidas y Producidas .....	119
5.3.2.3	Cuadro de Mando de Toneladas Pedidas y Entregadas.....	120
5.3.3	Administración y Seguridad de los Informes .....	121
5.3.3.1	Creación de usuarios.....	123
5.3.3.2	Asignación de accesos a los informes .....	124
5.4	Validación y Verificación de los requerimientos .....	125
5.4.1	Validación y Pruebas en el proceso ETL .....	125
5.4.2	Validación y Pruebas del Cubo OLAP .....	127
5.4.3	Validación y Pruebas de los Informes.....	128
CAPÍTULO 6.....		130
ANÁLISIS DE RESULTADOS.....		130
6.1	Evaluación de los resultados de la implementación.....	130
6.2	Establecer alternativas de mejora.....	133
CONCLUSIONES Y RECOMENDACIONES .....		134

BIBLIOGRAFÍA.....	138
APÉNDICE .....	141

## ÍNDICE DE FIGURAS

Figura 2.1 Arquitectura de un sistema de BI [9] .....	18
Figura 4.1 Partes del proceso de desarrollo del DD-DSS .....	29
Figura 4.2 Modelo Dimensional de Pedidos.....	45
Figura 4.3 Modelo Dimensional de Ventas.....	46
Figura 4.4 Modelo Dimensional de Entregas. ....	47
Figura 4.5 Modelo Dimensional de Ingresos. ....	48
Figura 5.1 Diagrama de Actividad Carga de Datos de Dimensiones. ....	51
Figura 5.2 Diagrama de Actividades Carga de Datos Tabla de Hechos.....	53
Figura 5.3 Esquema ETL Dimensión en PDI.....	55
Figura 5.4 Esquema ETL Tabla de Hecho. ....	57
Figura 5.5 Conexiones a base de datos para ETL de las dimensiones. ....	58
Figura 5.6 Transformación ETL Producto. ....	59
Figura 5.7 Lectura de la tabla Producto .....	60
Figura 5.8 Filtrado de datos de Productos. ....	61
Figura 5.9 Obtención de descripción grupo de artículos. ....	62
Figura 5.10 Obtención de la descripción de la subcategoría.....	63
Figura 5.11 Obtención de la descripción del Mercado. ....	64
Figura 5.12 Reemplazo de Valores Nulos.....	65
Figura 5.13 Obtener Fecha del Sistema. ....	65
Figura 5.14 Escritura de la dimensión Producto.....	66
Figura 5.15 Escritura de datos-Verificación de registro.....	67
Figura 5.16 Escritura de datos – Campos de la dimensión. ....	69

Figura 5.17 Transformación ETL Tabla de Hechos Ventas.....	71
Figura 5.18 Obtención de Fecha de Periodo de Carga.....	72
Figura 5.19 Borrado de Periodos de Carga de Ventas.....	73
Figura 5.20 Lectura Tabla de Facturas y Notas de Crédito.....	74
Figura 5.21 Renombrar Fecha de Borrado.....	74
Figura 5.22 Paso Renombrar Parámetros Carga Hechos.....	75
Figura 5.23 Lectura Tabla de Detalles de Ventas.....	76
Figura 5.24 Obtener Clave Dimensión Producto.....	78
Figura 5.25 Campos para obtener Clave Técnica Dimensión Producto.....	79
Figura 5.26 Ordenamiento de Registros.....	80
Figura 5.27 Eliminar Registros Duplicados.....	81
Figura 5.28 Escritura Tabla de Hechos Ventas.....	82
Figura 5.29 Job para la carga periódica de las transformaciones.....	83
Figura 5.30 Esquema de los Cubos OLAP.....	85
Figura 5.31 Conexión Base de datos del Cubo OLAP.....	86
Figura 5.32 Enlace de dimensión producto con Cubo OLAP.....	87
Figura 5.33 Jerarquía y niveles de Productos.....	88
Figura 5.34 Enlace dimensión Cliente con el cubo OLAP.....	89
Figura 5.35 Jerarquía y Niveles de la dimensión Cliente.....	89
Figura 5.36 Enlace de la dimensión Tiempo con el cubo OLAP.....	90
Figura 5.37 Jerarquía y Niveles de la dimensión Tiempo.....	90
Figura 5.38 Enlace de la dimensión Vendedor con el Cubo OLAP.....	91
Figura 5.39 Jerarquía y Niveles de la dimensión Vendedor.....	91
Figura 5.40 Enlace de la dimensión Localidad con el cubo OLAP.....	92

Figura 5.41 Jerarquía y Niveles de la dimensión Localidad. ....	92
Figura 5.42 Enlace de la dimensión Unidad de Medida con el cubo OLAP. ....	93
Figura 5.43 Jerarquía y Niveles de la dimensión Unidad de Medida. ....	93
Figura 5.44 Enlace de la dimensión Condición Crédito con el cubo OLAP.....	94
Figura 5.45 Jerarquía y Niveles de la dimensión Condición Crédito.....	94
Figura 5.46 Enlace de la dimensión Tipo con el cubo OLAP.....	95
Figura 5.47 Jerarquía y Niveles de la dimensión Tipo.....	95
Figura 5.48 Enlace de la dimensión Estado con el cubo OLAP.....	96
Figura 5.49 Jerarquía y Niveles de la dimensión Estado.....	96
Figura 5.50 Cubo OLAP de Ventas.....	98
Figura 5.51 Cubo OLAP Ventas – Tabla de Hechos. ....	99
Figura 5.52 Enlace de la dimensión Documento con el cubo OLAP.....	100
Figura 5.53 Jerarquía y Niveles de la dimensión Documento.....	100
Figura 5.54 Medida Cantidad Vendida.....	102
Figura 5.55 Medida Monto Bruto.....	102
Figura 5.56 Medida Monto Neto.....	103
Figura 5.57 Medida Valor IVA.....	103
Figura 5.58 Medida Kilos Vendidos.....	104
Figura 5.59 Campo Calculado Toneladas Vendidas. ....	105
Figura 5.60 Cubo Virtual VENTAS-INGRESOS. ....	106
Figura 5.61 Vista de Análisis Ventas por Año. ....	109
Figura 5.62 Vista de Análisis de Ventas por Cliente.....	110
Figura 5.63 Vista de Análisis de Ventas por Localidad.....	111
Figura 5.64 Vista de Análisis de Ventas por Vendedor. ....	112

Figura 5.65 Vista de Análisis de Toneladas Producidas contra Toneladas Vendidas. .....	113
Figura 5.66 Comparativos de Pedidos por Año.....	114
Figura 5.67 Vista de Análisis de Comparativo de Pedidos contra Entregas. ....	115
Figura 5.68 Vista de Análisis de Entregas.....	116
Figura 5.69 Cuadro de Mando de Ventas. ....	118
Figura 5.70 Cuadro de Mando de Toneladas Vendidas y Producidas.....	119
Figura 5.71 Cuadro de Mando de Toneladas Pedidas y Entregadas. ....	120
Figura 5.72 Permisos del rol Administrator .....	121
Figura 5.73 Permisos del rol Business Analyst .....	122
Figura 5.74 Permisos del rol Power User.....	122
Figura 5.75 Permisos del rol Report Author .....	122
Figura 5.76 Creación de usuario.....	123
Figura 5.77 Asignación de rol/usuario a un informe. ....	124

## ÍNDICE DE TABLAS

Tabla 1. Necesidades y requerimientos funcionales. ....	32
Tabla 2. Necesidades y requerimientos no funcionales. ....	33
Tabla 3. Entidades principales de orígenes de datos. ....	37
Tabla 4. Entidades secundarias de orígenes de datos. ....	39
Tabla 5. Dimensiones Compartidas. ....	40
Tabla 6. Dimensiones propias. ....	41
Tabla 7. Tablas de Hecho. ....	42
Tabla 8. Medidas. ....	42
Tabla 9. Esquema ETL Tabla de dimensión. ....	54
Tabla 10. Esquema ETL Tabla de Hecho. ....	55
Tabla 11. Asociación de las dimensiones compartidas con el cubo OLAP. ....	101

## INTRODUCCIÓN

El presente documento está orientado a describir el proceso que se siguió para implementar un sistema de soporte a la decisión orientado a datos en una empresa de manufactura. Se detallan los objetivos generales y específicos del trabajo. Además se define el alcance y las herramientas que se utilizaran para desarrollar la solución.

Este trabajo, se basa en un marco teórico que define las bases para la implementación de la solución. La teoría brinda los conceptos necesarios para diseñar y desarrollar las aplicaciones que permitirán satisfacer el alcance propuesto.

Una vez establecidas las bases teóricas, se determinan las necesidades y requerimientos de los usuarios del área comercial y se plantea la solución para satisfacer dichos requerimientos. También se definen los entregables del proyecto.

Luego, de determinar los requerimientos se procede a analizar y diseñar el modelo dimensional que mejor se ajuste a la realidad y necesidades de la compañía. Se

identifican las entidades del modelo relacional del cual se extraerán los datos para ser almacenados en el datamart.

Posteriormente se procede con el desarrollo de la solución. Se crean los procesos de extracción y carga de datos al datamart. Además, se crean los cubos de datos y las herramientas de visualización de la información como son: vistas de análisis y cuadros de mando.

Finalmente, se evalúan los resultados de la implementación y se sugieren alternativas para mejorar el proyecto realizado.

# **CAPÍTULO 1**

## **GENERALIDADES**

En este capítulo se presentan los antecedentes de la empresa, el problema que enfrenta y la solución que se propone. Asimismo, se definen el objetivo general y los objetivos específicos de este proyecto.

### **1.1 Antecedentes**

La empresa Manufacturera S.A. está dedicada a la producción y comercialización de soluciones de empaque para el mercado bananero y doméstico, a nivel nacional e internacional. Esta empresa forma parte de un grupo de empresas que están alineadas estratégicamente de forma horizontal. Empezando con la empresa de reciclaje de papel que provee la materia prima para la fabricación de bobinas de papel, que a su vez constituyen la principal materia prima para la elaboración de los empaques. Las oficinas centrales de la compañía están ubicadas en la ciudad de Durán,

en donde se localiza la planta de producción. Las operaciones de ventas son realizadas en Durán, Quito, Machala y Cuenca [1].

La visión de la empresa es satisfacer las necesidades de empaques de sus clientes. Para ello posee una moderna planta productora con tecnología de punta que opera en una superficie de 27 mil metros cuadrados. Además, cuenta con personal altamente calificado en cada una de las áreas de la organización. En el área de producción trabajan alrededor de 400 empleados distribuidos en tres turnos.

Manufacturera S.A. dispone de un departamento de diseño estructural para el desarrollo de los planos y fichas técnicas. Este departamento se asegura que el empaque a fabricar se ajuste a las exigencias de resistencia y humedad que demanda el producto que va a ser embalado. Asimismo, tiene un departamento de diseño gráfico que se encarga de realizar el arte que usa el empaque de acuerdo a las especificaciones del cliente. Además, cuenta con los departamentos técnicos para la elaboración de clisés, troqueles y las tintas que se usan en el proceso de producción.

El área de ventas está compuesta por ejecutivos comerciales, servicio al cliente y postventa. Los empleados de esta área, principalmente los ejecutivos comerciales, son los encargados de ofrecer asesoría a los clientes en lo que se refiere a la colocación y seguimiento de pedidos. Los vendedores están distribuidos por sectores comerciales tales como: bananero, floricultor y doméstico (este sector abarca todas los demás

sectores que no incluyen banano ni flores; por ejemplo: sector atunero, frutas, etc.).

Al área de producción llegan los pedidos que se deben fabricar de acuerdo a una planificación previa. El proceso productivo empieza en una máquina conocida como la corrugadora, donde el papel se transforma en láminas que posteriormente pasan a las imprentas; aquí se coloca el diseño que deben tener los empaques con los respectivos colores y troquelado, en el caso que aplique. Una vez terminado este proceso, se almacena el producto terminado en la correspondiente bodega para luego ser distribuido al cliente.

Después de conocer los antecedentes de la empresa manufacturera, las ciudades donde opera y el funcionamiento del negocio, se procederá a describir el problema que actualmente afecta a esta organización en cuanto a necesidades de información para mejorar la gestión comercial.

## **1.2 Descripción del Problema**

La gerencia demanda constantemente información de las ventas y de la producción. Principalmente, para la elaboración de los presupuestos, el análisis del proceso productivo, el análisis de las ventas y la postventa. Esta información debe ser provista de manera rápida para mejorar la gestión de la empresa, pero no existen herramientas que permitan agilizar este proceso.

Actualmente la empresa consolida los datos de las ventas y de la producción de manera manual. Los datos provienen del sistema de administración SAP Business One (SAP BO) y del sistema de producción. Estos sistemas se interconectan a través de interfaces desarrolladas in-house. Los reportes e informes, de los sistemas mencionados, son básicos que proveen información que no es apta para ser analizada directamente, sino que requiere ser consolidada.

Para generar la información que se presenta a la gerencia se exportan a Excel los reportes proporcionados por los sistemas indicados. En esta herramienta los datos son unificados por el personal del área de ventas, de contraloría y por los asistentes de gerencia. Este proceso de consolidación tarda varias horas, e incluso días, para que la información esté lista.

Al ser un proceso manual, la información generada no es completamente confiable ya que pueden existir inconsistencias por la manipulación realizada. Además, la información entregada es estática y no es en tiempo real, lo que dificulta el análisis de la misma y la toma de decisiones en el momento oportuno.

La alta demanda de información por parte de la gerencia y del directorio, sumada a la falta de herramientas que brinden esta información de manera automática y en línea, ha producido que se extiendan las horas de trabajo de los empleados de las áreas involucradas. Aparte de cumplir con las

funciones propias de sus cargos, deben dedicar tiempo y esfuerzo a la unificación de los datos. Pero lo más importante es que la gerencia no dispone de una herramienta que le permita analizar los datos de la compañía de forma dinámica, integrada y a tiempo. Por consiguiente, no se pueden establecer acciones eficaces para aumentar la productividad y rentabilidad del negocio.

Luego de examinar los problemas que afronta la empresa de manufactura, en relación a la gestión de información, se propondrá la solución para solventar los problemas indicados.

### **1.3 Solución Propuesta**

Se plantea la implementación de un Sistema de Soporte a la Decisión Orientado a Datos (Data-Driven DSS, por las siglas en inglés de Decision Support System). Esta herramienta facilitará el análisis de los datos de la compañía permitiendo tener información útil y válida en cualquier momento [2].

El desarrollo de esta solución incluye el proceso de transformación e integración de datos (ETL, por las siglas en inglés de Extract, Transform and Load) provenientes desde las diferentes fuentes que tiene la compañía como son los sistemas SAP BO y de producción. Luego se procederá al diseño del modelo dimensional que incluye la creación del cubo de Procesamiento Analítico en Línea (OLAP, por las siglas en inglés de On-Line

Analytical Processing). Finalmente, se realizará la presentación de los datos en forma de informes dinámicos y gráficos que faciliten la comprensión y el análisis de la información.

La solución propuesta se implementará con tecnología Open Source de la suite BI PENTAHO [3]. Se seleccionó esta plataforma porque cuenta con herramientas potentes para el desarrollo de los diferentes procesos expuestos en el párrafo anterior. Entre las cuales se mencionan las siguientes:

- Data Integration – Kettle: Herramienta para el proceso ETL que permite crear transformaciones y trabajos que extraen los datos dejándolos disponibles para el análisis y presentación en reportes.
- Mondrian: Servidor OLAP.
- Schema-WorkBench: Interfaz de diseño visual que permite la creación y pruebas de esquemas de cubos OLAP Mondrian.
- Pentaho Reporting: Herramienta para la creación de reportes que pueden ser presentados en diferentes formatos: PDF, HTML, CSV.
- Pentaho Dashboard: Herramienta para la integración de los diferentes reportes y vistas de análisis.

Entre los beneficios que brindará la implementación del Data-Driven DSS se mencionan los siguientes:

- Se podrá obtener información de las ventas y de la producción en tiempo real.
  - La información será confiable puesto que se extrae directamente de las fuentes de los datos.
  - Se podrán realizar consultas de información histórica y comparaciones con años anteriores.
  - Se contará con informes dinámicos e interactivos donde los usuarios podrán ajustarlos de acuerdo a sus necesidades de información.
  - Análisis drill down que permitirá ir desde la información general a la particular; como por ejemplo, desglose del informe de ventas totales por mercado e ir al nivel de ventas por producto.
  - Se podrá analizar la información de manera sencilla y práctica puesto que se dispondrán de gráficos que permitirán una mejor comprensión de los datos.
  - Los empleados que intervienen en el proceso de consolidación de información contarán con más tiempo para realizar las tareas de su cargo ya que la herramienta proporcionará la información de manera automática.

Una vez que se ha indicado la solución que se implementará para mejorar el análisis de información en la empresa, se define el objetivo general y los objetivos específicos de este trabajo de titulación.

## **1.4 Objetivo General**

Diseñar e implementar un Data-Driven DSS que permita mejorar el proceso de análisis de datos de ventas y producción, en una empresa de manufactura, con el fin de apoyar la toma de decisiones.

## **1.5 Objetivos Específicos**

- Definir la situación actual con el propósito de identificar y analizar: requerimientos, procesos y procedimientos, sistemas de información existentes, fuentes de datos internas y externas, y necesidades de información.
- Analizar y diseñar la solución Data-Driven DSS que satisfaga las necesidades de información identificadas.
- Desarrollar la solución Data-Driven DSS y establecer el plan de pruebas para asegurar que se cumplan con los requerimientos de información solicitados por la organización.
- Analizar los resultados de la implementación y establecer ajustes y mejoras.

Luego de conocer los antecedentes de la empresa, el problema que afronta, proponer la solución al problema y de establecer los objetivos de este trabajo, en el siguiente capítulo se revisarán los conceptos teóricos sobre los cuales se soporta la solución propuesta.

## **CAPÍTULO 2**

### **MARCO TEÓRICO**

El presente capítulo se enfoca en los conceptos teóricos sobre los que se basa el desarrollo de este trabajo. Se indicarán las definiciones para un DSS y para inteligencia de negocios (BI, por las siglas en inglés de Business Intelligence). Se explicará sobre los componentes y tipos de DSS. Además, se ahondará en los DSS orientado a datos y sus principales características.

#### **2.1 Definición de DSS**

En el año 1971, Morton definió a los DSS como sistemas de computadora que hacen uso de datos y modelos para soportar la toma de decisiones [4].

Otro concepto que es regularmente utilizado, es el propuesto por Keen y Morton en 1978 [4] en el cual indican que los DSS son la unión de las capacidades intelectuales de las personas con las capacidades de los computadores, usadas en conjunto para optimizar la toma de decisiones y resolver problemas semiestructurados [4].

Hay que considerar que un DSS, en términos generales, es cualquier sistema de información con la capacidad de apoyar la toma de decisiones en las empresas [4].

Los DSS y la BI guardan una estrecha relación, por lo que a continuación se aborda la similitud y la diferencia entre estos dos conceptos.

## 2.2 Inteligencia de Negocios

Antes de abordar la relación entre DSS y BI, se mencionan los conceptos básicos relacionados a BI.

**Business Intelligence (BI).** Es el conjunto de metodologías, arquitecturas y tecnologías para facilitar la toma de decisiones en las organizaciones. Con BI se transforman los datos en información y la información en conocimiento [4].

Los sistemas de inteligencia de negocios hacen uso de los datawarehouse; a continuación se define este concepto.

**Datawarehouse.** Es un almacén de datos que contiene información actual e histórica de toda la compañía para facilitar la toma de decisiones en cualquier área de la organización [5]. De acuerdo con Kimball [5], un datawarehouse contiene los datos del sistema transaccional estructurados de tal forma que facilitan la consulta y el análisis. Además, la información

siempre es almacenada en un modelo dimensional. Mientras que W. H. Inmon [5] dice que un datawarehouse es una colección de datos orientado a un tema, integrado, cuya información no es borrada ni modificada y contiene información de diferentes años.

Los datawarehouse hacen referencia a los datamarts, por ésto se los trata a continuación.

**Datamarts.** Es un repositorio de datos que contiene datos específicos de un área de la organización, lo que permite realizar análisis de forma rápida [6]. Los datamarts pueden ser independientes o dependientes de un datawarehouse [6]. Los datamarts independientes se enfocan en un área en particular; por ejemplo, ventas, compras, etc.; no son construidos con un enfoque empresarial. El tipo de datamart independiente puede ser construido en un tiempo relativamente corto. Mientras que los datamarts dependientes toman los datos de un datawarehouse, se los utiliza para mejorar el rendimiento del datawarehouse y mantener la seguridad de los datos debido a que cada datamart maneja sus propios niveles de acceso a los datos [6].

Como se mencionó anteriormente, los sistemas DSS y BI tienen similitudes y diferencias, entre las cuales se mencionan las siguientes [4]:

- BI requiere usar un datawarehouse (que suele ser relativamente costoso) por lo que está orientado a organizaciones grandes. Mientras que los DSS pueden o no usar datawarehouse, lo que permite que sean aplicados a cualquier tipo de organización.
- Los DSS son construidos para dar soporte a la decisión de forma directa y los BI son construidos para brindar información concisa y oportuna y de forma indirecta proporcionan soporte a las decisiones.
- Los sistemas BI están orientados a las áreas de estrategia y ejecutivos, y los DSS se alinean hacia los analistas.
- Los sistemas de BI son construidos con herramientas comerciales que se ajustan a los requerimientos de las empresas. Los sistemas DSS se basan en soluciones para resolver problemas no estructurados y se personalizan de acuerdo a los problemas que se quieran resolver.
- Tanto los sistemas de BI como los DSS usan herramientas como la minería de datos y el análisis predictivo.

En la siguiente sección se profundizará en los conceptos de los DSS, con el fin de tener una mayor comprensión de este tipo de sistemas, objeto de este trabajo de titulación.

### **2.3 Componentes de un DSS**

Los DDS constan de los siguientes componentes [4]:

- **Sistema de gestión de datos:** Este sistema incluye una base de datos que es administrada por un sistema de gestión de base de datos. Esta base de datos puede estar interconectada con el datawarehouse de la compañía.
- **Sistema de gestión de modelos:** Este sistema está compuesto por un paquete de software que incluye modelos financieros, estadísticos, científicos y cuantitativos. Estos sistemas pueden estar conectados a modelos de la compañía o a modelos externos.
- **Sistema de interfaz de usuario:** Se refiere a la interfaz de usuario que comunica a los usuarios con el DSS. Esta comunicación se produce a través del uso de comandos. Los usuarios son considerados parte de este sistema. Los navegadores web proporcionan una interfaz amigable para los usuarios.
- **Sistema de gestión de conocimiento:** Estos sistemas pueden dar apoyo a los otros sistemas o trabajar de forma independiente. Está diseñado para incrementar las capacidades del DSS y mejorar la toma de decisiones. Estos sistemas pueden estar conectados con las bases de conocimiento que tienen ciertas organizaciones.

A continuación se describen los tipos de DSS que son generalmente aceptados.

## 2.4 Tipos de DSS

La Asociación para Sistemas de Información, Grupo de Interés Especial en Sistemas de Soporte a la Decisión (AIS SIGDSS, por sus siglas en inglés de Association for Information Systems Special Interest Group on Decision Support Systems), adoptó la siguiente clasificación de DSS propuesta por Power [4]:

**DSS orientado a comunicaciones y DSS de grupo.** El DSS de grupo es un tipo de DSS híbrido, utiliza las comunicaciones, colaboraciones y tecnologías para el soporte de decisiones. Este tipo de sistema da apoyo a grupos y puede o no incluir toma de decisiones [7].

**DSS orientado a datos.** Este tipo de sistemas incluye datawarehouse y sistemas de análisis. Los sistemas de inteligencia negocio también se incluyen en este tipo. Este sistema consiste en el uso de grandes base datos y archivos que pueden ser internos o externos a la compañía. Con la integración de OLAP se obtiene un alto rendimiento de estos sistemas que permiten analizar grandes cantidades de datos [7].

**DSS orientado a documentos.** Este tipo de DSS ayuda a gestionar los documentos no estructurados que se encuentran principalmente en la web. Se puede acceder a grandes bases de datos de documentos como hipertexto, imágenes, audio y video. Los DSS orientados a documentos

pueden acceder a catálogos de productos, políticas, procedimientos, registros corporativos y correspondencia, entre otros [7].

**DSS orientado a conocimiento.** Los DSS basados en inteligencia artificial se incluyen dentro de este grupo [4]. Estos sistemas se basan en reglas para automatizar el proceso de toma de decisiones. Los sistemas expertos y la minería de datos están incluidos en este grupo. Estos sistemas son costosos debido a los grandes beneficios que ofrecen [4].

**DSS orientado a modelos.** En este tipo de sistemas se incluyen los sistemas que usan modelos financieros, estadísticos y de optimización. En este tipo de sistemas, los usuarios pueden ingresar datos y parámetros para la toma de decisiones pero no requieren ni administran grandes cantidades de datos sino que se centran en acceder y manipular el modelo [7].

En la siguiente sección se presentan las principales características de un DSS orientado a datos, ya que el proyecto se basa en este tipo de DSS.

## 2.5 Características de los DSS Orientado a Datos

De acuerdo con Power, las principales características de los DSS orientados a datos, son [8]:

- **Filtrado y recuperación de datos Ad Hoc.** Estos sistemas permiten a los usuarios filtrar y recuperar datos de forma sistematizada. Los menús

son desplegadas y, por lo general, las vistas son predefinidas. Los usuarios deben tener capacidades de análisis y pueden ir desde lo general a lo específico.

- **Alertas y disparadores.** Algunos sistemas proporcionan reglas para envío de notificaciones de email u otras acciones.
- **Crear visualización de datos.** Los usuarios pueden ser capaces de ver los datos de diferentes formas: histogramas, diagramas de barra y pie. Incluso, pueden reproducir los datos históricos en una secuencia de tiempo.
- **Gestión de datos.** Los usuarios pueden agrupar datos y cambiar el formato de los datos.
- **Datos resumidos.** Los usuarios pueden ver o crear tablas dinámicas, ver los datos resumidos o en detalle. Realizar cálculos de totales y subtotales.
- **Integración con hojas de cálculo.** Los DSS permiten extraer y descargar los datos para análisis posteriores. Incluso hay otros DSS que permiten cargar datos al área de trabajo de los usuarios.
- **Creación y recuperación de metadatos.** Los usuarios pueden agregar metadata al análisis y a los reportes, permitiendo cambiar las etiquetas para presentación de los datos.
- **Diseño, generación y almacenamiento de informes.** Los usuarios pueden crear reportes y gráficos en diferentes formas y pueden guardar estos reportes para su posterior uso. Además, pueden ser distribuidos vía web o exportándolos a formatos como PDF.

- **Análisis estadístico.** Los usuarios pueden calcular estadísticas descriptivas de los datos para resumir los datos y establecer tendencias o encontrar relaciones entre ellos.
- **Vistas de datos predefinidas.** En los DSS, a menudo se publican vistas de datos y reportes predefinidos por los diseñadores del DSS. Estos elementos son agrupados en los dashboards. Algunos DSS muestran también scorecard que son tablas donde se muestran las métricas e indicadores definidos.
- **Vistas de informes de producción.** Los diseñadores de reportes a menudo crean reportes predefinidos para que sean utilizados por los usuarios finales.

Según Power [8], los sistemas de BI son Data-Driven DSS que utilizan gran cantidad de datos históricos y proporcionan reportes periódicos. Los sistemas de BI utilizan minería de datos para automatizar las decisiones. Por lo que se deduce que BI es una evolución de los DSS. A continuación se describe la arquitectura de los sistemas BI.

## 2.6 Arquitectura de un Sistema BI

Los sistemas de inteligencia de negocios tienen su origen en los sistemas transaccionales; de los cuales se extrae, transforma y carga los datos a los almacenes intermedios mediante un proceso conocido como ETL. Los almacenes intermedios ODS (Por las siglas en inglés de Operational Data Store) sirven como intermediarios entre los sistemas fuentes y destinos con

la finalidad de integrar los datos de origen provenientes desde múltiples fuentes [9]. En la Figura 2.1, se muestra la arquitectura de un sistema de BI.

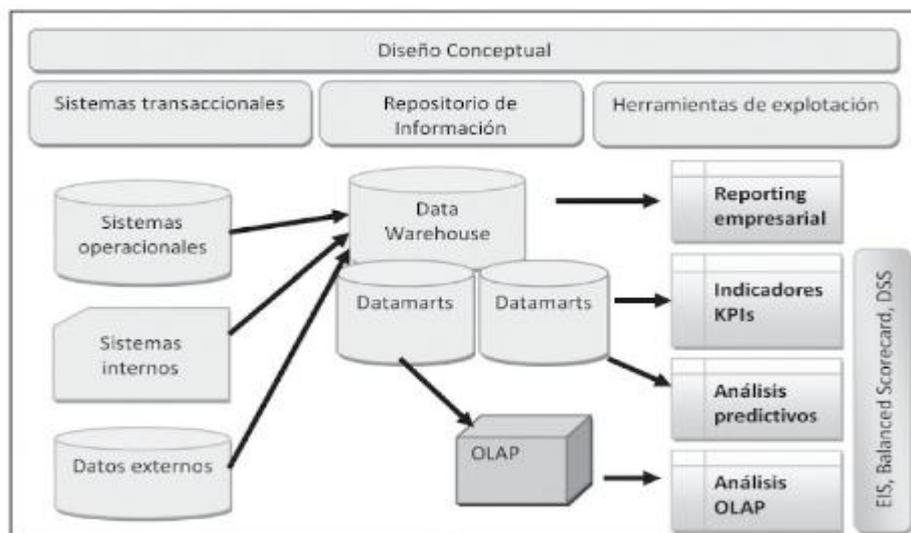


Figura 2.1 Arquitectura de un sistema de BI [9]

Una vez que la información se encuentra depurada y consolidada, se almacena en un datawarehouse corporativo. Desde aquí se pueden construir los diferentes datamarts departamentales que se utilizan para realizar el análisis de los datos a través de base de datos de Procesamiento de Transacciones En Línea (OLTP por sus siglas en inglés de OnLine Transaction Processing) o bases de datos analíticas OLAP [9].

Por último, se realiza el análisis de los datos almacenados en el datawarehouse mediante el uso de las diferentes herramientas existentes para reportería, análisis, cuadros de mandos [9].

Para la creación de un cubo OLAP es necesario realizar el modelamiento de los datos. Por ello, a continuación se trata sobre este tema.

## 2.7 Modelamiento Dimensional

El modelamiento dimensional surge debido a la necesidad de gestionar grandes volúmenes de datos. El modelo dimensional permite mejorar el rendimiento de las consultas en los datawarehouses [10].

El modelado dimensional difiere del modelado relacional principalmente porque el segundo está orientado a las transacciones operativas de una empresa, donde se guarda la consistencia y se evita la redundancia de los datos. Mientras que el primero se orienta al manejo de grandes volúmenes de datos donde prima el rendimiento para proporcionar rapidez en el análisis.

El modelo dimensional agrupa los datos de tal forma que son fácilmente entendibles por los usuarios y se procesan rápidamente en comparación con los modelos relacionales [11].

El modelo dimensional comprende los siguientes elementos principales [10]:

- **Dimensiones:** Son entidades donde se almacenan detalles de las tablas de hechos. Por ejemplo: dimensión de tiempo, productos, clientes.
- **Hechos:** Son entidades que almacenan datos numéricos para conocer el contexto del negocio. Por ejemplo: la tabla de hecho de ventas.

- **Jerarquías:** Son relaciones de muchos niveles entre los atributos de las tablas de dimensión. Por ejemplo; jerarquía de fechas: Año – Mes - Semana
- **Medidas:** Son atributos numéricos asociadas a las tablas de hechos que permiten medir los procesos. Por ejemplo, Cantidad vendida, Total vendido en dólares.

En la siguiente sección se describen los tipos de modelamiento que existen para diseñar base de datos dimensionales.

## 2.8 Tipos de Modelamiento

Existen distintos tipos para el modelamiento dimensional de un datawarehouse; la elección dependerá de las necesidades y requerimientos del negocio. Estos son [10]:

**Esquema Estrella.** El modelo estrella consta de una tabla de hechos y varias tablas de dimensiones unidas a través de sus claves. Las tablas deben estar totalmente desnormalizadas; es decir, deben contener los datos de nombres, descripciones, etc., en las mismas tablas de dimensiones. Esto puede ocasionar redundancia, pero lo primordial es garantizar un buen rendimiento en el acceso a los datos al evitar los saltos a otras tablas para obtener los datos deseados. Además, el espacio adicional utilizado no es muy representativo.

**Esquema Copo De Nieve.** El esquema de copo de nieve se parece a un esquema entidad-relación puesto que las tablas de dimensión se encuentran normalizadas. Al igual que en el esquema de estrella, este esquema consta de una tabla de hechos y varias tablas de dimensiones. Estas a su vez, pueden estar asociadas a otras tablas de dimensiones, que son las denominadas jerarquías. Este modelo es un poco más complejo y puede tornarse lento para el acceso a los datos debido a que pueden crecer significativamente las jerarquías de las dimensiones. Se utiliza este modelo para cuando se quiere segregarse las dimensiones.

**Esquema de Constelación.** El esquema de constelación está compuesto por varios esquemas de estrella; donde existe una tabla principal de hechos y una o varias tablas de hechos secundarias. Las tablas de hechos están unidas a las tablas de dimensiones que pueden o no ser compartidas entre las diferentes tablas de hechos existentes en el modelo.

En este capítulo se revisaron los conceptos teóricos sobre DSS y BI que permiten comprender la base conceptual para el desarrollo de este trabajo. Se pudo conocer los tipos de DSS, las características de los DSS, arquitectura de BI y aprender sobre el modelamiento dimensional muy necesario para este proyecto. En el capítulo tres, se presentará la situación actual de la empresa; así como los requerimientos y necesidades de información para mejorar la toma de decisiones en esta organización.

## **CAPÍTULO 3**

### **DEFINICIÓN DE LA SITUACIÓN ACTUAL**

En este capítulo se establece la situación actual de la problemática de la empresa manufacturera objeto de este trabajo. Se identifican los requerimientos de información necesarios para la mejora en la toma de decisiones de esta organización. Por último, se definirá el alcance del presente proyecto.

#### **3.1 Situación Actual**

El área comercial es el área de mayor criticidad para cualquier empresa. En la Manufacturera S.A., esta área es la responsable que se generen mayores ingresos y rentabilidad para la organización. Para alcanzar ese objetivo es imperativo que se disponga de información útil y a tiempo.

Actualmente, la empresa no cuenta con reportes que proporcionen información para analizar y facilitar la toma de decisiones. Los reportes que se tienen son estáticos que brindan información básica de las ventas,

despachos, toneladas producidas, entre otra información, pero que no permiten profundizar en los datos ni realizar comparaciones con periodos anteriores. La mayor parte de la información que está disponible no permite examinar la situación real de las ventas de la empresa de forma adecuada y en el tiempo oportuno.

En las bases de datos de la empresa se almacena toda la información que se genera de las transacciones diarias que se realizan, tanto del área productiva como del área comercial y administrativa. Para revisar la información almacenada, se tienen consultas en el sistema SAP BO que permiten exportar estos datos a Excel. Cuando se requiere información más al detalle se solicita al área de sistemas que se proporcione reportes específicos en Excel sobre algún tema en particular, siendo esta información útil sólo una vez y para el usuario que la solicita.

Los jefes de las áreas de venta y servicio al cliente descargan a Excel los reportes que tienen disponibles en el sistema SAP BO para generar la información que la gerencia demanda. Desde esta herramienta proceden a elaborar los informes requeridos. Los informes se realizan combinando información de diferentes reportes. Este proceso se efectúa diariamente para entregar a la gerencia la información que necesita para llevar a cabo su gestión. Además, hay información de las ventas que los gerentes solicitan para analizar mensualmente el desempeño de la empresa y compararla con los estados financieros. Los analistas de producción y financieros deben

buscar esta información y prepararla en el formato requerido para la gerencia y el directorio.

El proceso de unificación y consolidación de la información es un trabajo tedioso y complicado para las personas que lo realizan debido a que se tiene que extraer la información desde diferentes fuentes, llevarla a Excel, ponerla en el formato requerido por los jefes y gerentes y validar que los datos sean reales y actualizados. Esta labor consume varias horas de trabajo de los empleados sobretodo del área comercial que, como se mencionó anteriormente, es una de las áreas más críticas de esta organización, lo que provoca estrés laboral e inconformidad en el área.

En resumen, la empresa no dispone de información que le permita gestionar el área comercial de manera más eficiente y rápida. La información con la que cuentan en esta área es generada de forma manual, lo que puede conllevar a errores y, por ende, a que la información no sea totalmente confiable. Además la información no es emitida en línea lo que implica no tener información actualizada. Esto constituye una significativa desventaja competitiva para la organización.

En el siguiente punto se describe el levantamiento de información realizado en el área comercial. Dicha información se obtuvo por medio de entrevistas a los usuarios involucrados en el proceso de ventas.

### 3.2 Levantamiento de Información

Con la finalidad de identificar las necesidades de información requeridas por la empresa se realizaron entrevistas con los usuarios del área comercial y analistas de producción. Estos usuarios son los encargados de preparar los informes que la gerencia comercial requiere y tienen amplio conocimiento en sus respectivas áreas de trabajo. También cuentan con gran experiencia en el sector industrial al cual corresponde esta empresa y conocen el funcionamiento del mercado.

En estas entrevistas, los usuarios detallaron cómo se elaboran los informes para el análisis de las ventas y explicaron el funcionamiento del área comercial desde la puesta de pedidos hasta la entrega del producto al cliente. También manifestaron las mejoras y requerimientos que facilitarían el trabajo en el área comercial y, por consiguiente, beneficiarían a toda la organización. A continuación se detallan los principales procesos del área comercial:

**Proceso de Ingreso de Pedidos de Ventas.** Este proceso involucra la recepción del pedido del cliente, el registro de los datos referentes al pedido en el sistema, tales como: detalle de productos, cantidades, fechas de entrega y otras consideraciones requeridas. Posteriormente, se comunica vía correo electrónico al cliente y demás áreas de la empresa involucradas (crédito y cobranza) que el pedido ha sido registrado.

**Proceso de Aprobación de Pedido.** Una vez que el pedido de ventas es registrado pasa al área de crédito y cobranzas para la respectiva revisión de límites de crédito y facturas vencidas. Dependiendo del análisis realizado por el personal de crédito el pedido será aprobado o rechazado.

**Proceso de Registro de Orden de Producción.** Cuando el pedido ha sido aprobado por el área de crédito de forma automática a través de una interfaz, el pedido se convierte en una orden de producción y pasa a la bandeja del planificador para que dichas órdenes sean programadas para la producción de acuerdo a la fecha de entrega y carga de trabajo de las máquinas.

**Proceso de Ingreso de Producto Terminado.** Cuando se ha fabricado la orden de producción se procede al registro del inventario en el sistema SAP BO. Se ingresan los lotes de productos creados en el área de producción con su correspondiente número de pallet y el personal de bodega de producto terminado procede a sectorizar tanto físicamente como en el sistema los productos fabricados.

**Proceso de Despacho de Producto.** Después que el producto es ingresado en la bodega, el personal de despacho revisa el programa de entregas enviado por servicio al cliente para efectuar la entrega del producto al cliente en las cantidades y fechas de entrega establecidas.

**Proceso de Facturación de Producto.** Luego de registrar la entrega de mercadería en el sistema, el despachador realiza la facturación del producto que se está despachando. Esto lo hace en base a la orden de entrega (guía de remisión).

**Proceso de Devolución de Producto.** En el caso de existir devolución del producto enviado al cliente por motivos de calidad o atrasos en la entrega se procede a emitir la respectiva nota de crédito y el personal de servicio al cliente registra la no conformidad asignándola al área y responsable que corresponda.

Los procesos mencionados forman parte del área comercial o se relacionan con ella durante el ciclo de vida del pedido de ventas. Cada proceso genera información que es importante para la compañía y requiere ser analizada de forma individual y en conjunto, ya sea para la mejora de los procesos o para implementar nuevas estrategias que permitan mejorar la eficiencia de la empresa.

En esta sección se trató sobre la situación actual de la empresa de manufactura, los problemas que enfrenta en cuanto a necesidad de información. La empresa cuenta con los datos en sus sistemas transaccionales pero no tiene una herramienta que le permita explotar dichos datos para beneficio del negocio. En el siguiente capítulo se procederá a realizar el análisis y diseño de la solución propuesta para la empresa.

## **CAPÍTULO 4**

### **ANÁLISIS Y DISEÑO DE LA SOLUCIÓN**

En este capítulo se analizan los requerimientos para diseñar la solución que mejor se adapte a la realidad de la empresa de manufactura. Se describe la metodología que se utilizó para el desarrollo de este trabajo. Se identifica el modelo relacional con el que se trabajará para extraer los datos de los sistemas fuentes, en este caso SAP BO. También se definen las tablas de dimensiones y medidas que serán necesarias para el diseño del cubo OLAP. Por último, se realiza el diseño del modelo dimensional sobre el cual se crea el cubo OLAP y que es la base para la elaboración de los reportes, indicadores, vistas de análisis de la herramienta implementada.

#### **4.1 Metodología**

Luego de identificar la situación actual de la empresa se definió la metodología a seguir para la implementación de la solución propuesta. La

metodología por la que se optó fue Scrum debido a la flexibilidad que esta ofrece; al permitir hacer entregas parciales de forma incremental [12].

El proceso para el desarrollo de este sistema DD-DSS implicó cinco partes, donde las tres primeras partes no son visibles para el usuario final pero implican una gran cantidad de trabajo, tiempo y esfuerzo ya que son la base para la creación de los informes y cuadros de mando (que son la parte visible y funcional para los usuarios). En la Figura 4.1 se muestran las partes que comprendieron este proyecto.

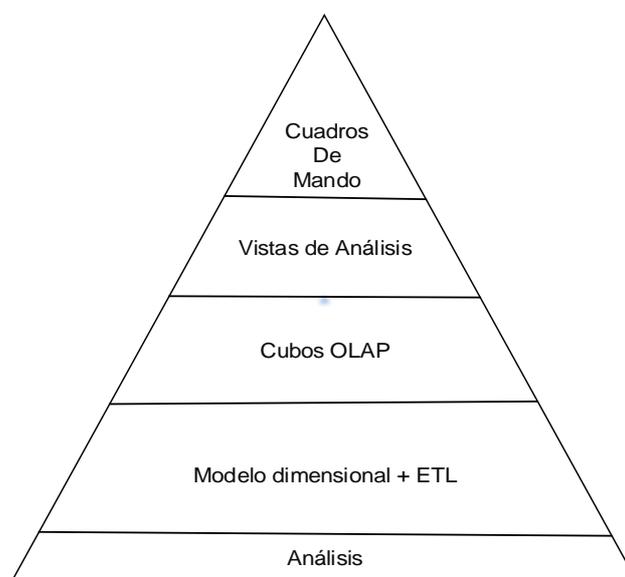


Figura 4.1 Partes del proceso de desarrollo del DD-DSS

Como se aprecia en la Figura de arriba, las partes que comprendieron este trabajo fueron:

- **Análisis:** Incluye el levantamiento de requerimientos y necesidades de los usuarios y el análisis de los mismos.

- **Modelo dimensional y ETL:** Comprende la creación del modelo dimensional basado en el análisis de los requerimientos. Además comprende la creación de los diferentes procesos ETL para alimentar el datamart.
- **Cubos OLAP:** Corresponde a la creación de los cubos OLAP necesarios para cumplir con los requerimientos planteados.
- **Vistas de análisis:** Incluye la creación de los informes en forma de vistas de análisis basados en los cubos OLAP creados.
- **Cuadros de mando:** Concierno a la creación de los cuadros de mando que muestran los resultados de informes y de los indicadores que requieren los usuarios.

Posteriormente en este capítulo y en el capítulo 5, se describirán paso a paso cada una de las partes indicadas.

Las dos primeras partes se realizaron de forma integral, sin dividirla en entregables debido a que este proyecto implica un sólo datamart, para el área comercial, y el análisis y modelo dimensional implica sólo a esta área. Además, las dimensiones y tablas de hecho son compartidas por todos los cubos OLAP que intervienen en este proyecto. Las siguientes partes, desde la creación de los cubos OLAP hasta los cuadros de mando se dividieron en entregables siguiendo la metodología Scrum.

Una vez que se crearon todos los procesos ETL, se procedió a crear los cubos OLAP, vistas de análisis y cuadros de mando de una sola sección del área comercial. De esta forma, luego de terminada la primera sección se pudo validar con los usuarios, realizar las correcciones correspondientes y hacer entregables por partes. Después de asegurarse que la sección seleccionada estaba conforme a lo requerido por los usuarios se dio por terminada esa sección. Luego se continuó con las siguientes secciones, a las cuales se aplicó las correcciones identificadas, en la primera sección, en el análisis y los ETL.

Luego de establecer la metodología de trabajo a seguir se procedió con la siguiente fase de este trabajo, que consiste en la identificación de necesidades y requerimientos de los usuarios del área comercial.

## **4.2 Identificación de Requerimientos**

En base a las entrevistas realizadas, se procedió a identificar los requerimientos a partir de las necesidades de los futuros usuarios de la herramienta a implementar; así mismo, se consideraron las necesidades del área de TI.

Los requerimientos se dividieron en requerimientos funcionales y no funcionales de manera que se pueda identificar claramente qué debe hacer la herramienta (funcional) y cuán bien debe hacerlo (no funcional).

En la Tabla 1 se detallan los requerimientos funcionales:

Tabla 1. Necesidades y requerimientos funcionales.

<b>NECESIDAD</b>	<b>REQUERIMIENTO FUNCIONAL</b>
<p>Contar con información de ventas en toneladas, kilos, dólares y unidades, y poder analizarla de forma resumida y detallada.</p>	<p>La herramienta debe generar reportes y vistas de las ventas realizadas considerando distintos parámetros, tales como: clientes, localidades, productos y en diferentes unidades de tiempo: mensual, semestral, anual. La información se podrá visualizar desde distintos niveles yendo desde lo general a lo específico.</p>
<p>Realizar comparativos entre toneladas producidas y toneladas vendidas. Para poder verificar la eficacia en el despacho puesto que el negocio de la empresa es producir bajo pedido y no mantener inventarios que representan costos.</p>	<p>La herramienta debe generar reportes y vistas que permitan comparar las toneladas que fueron producidas contra las toneladas que fueron vendidas en forma mensual, semestral y anual. Estos reportes deben permitir análisis drill down.</p>
<p>Efectuar comparaciones de los pedidos de ventas de forma mensual de tal forma que se permita determinar el crecimiento o disminución de ventas entre periodos y se pueda determinar si se está perdiendo participación con los clientes.</p>	<p>La herramienta debe generar reportes y vistas para analizar los pedidos de ventas en diferentes periodos: semanal, mensual, semestral. Estos reportes deben permitir análisis drill down.</p>
<p>Determinar la fiabilidad en las entregas. Es importante conocer cuan fiable es el proceso de entrega que tiene la empresa. Para lo cual se debe verificar los cumplimientos de las entregas en términos de tiempo y unidades de producto.</p>	<p>La herramienta debe generar reportes y vistas que muestren las entregas realizadas para poder verificar el cumplimiento en los términos definidos por los usuarios.</p>

En la Tabla 2 se detallan los requerimientos no funcionales:

Tabla 2. Necesidades y requerimientos no funcionales.

<b>NECESIDAD</b>	<b>REQUERIMIENTO NO FUNCIONAL</b>
Seguridad de los datos y la información.	La herramienta debe garantizar la seguridad de los datos y la información que se maneje dentro del datawarehouse. Debe disponer de perfiles para los usuarios que usarán la herramienta. Debe controlar el acceso dependiendo del perfil del usuario.
Confiabilidad de la solución.	La herramienta debe proporcionar la información que los usuarios requieran en cualquier momento con la rapidez y calidad adecuada.
Fiabilidad de la información.	La herramienta debe garantizar que la información que se procese y muestre al usuario final sea útil y válida.
Usabilidad de la herramienta.	La herramienta debe ser fácil de usar por cualquier tipo de usuario. Se deben mostrar gráficos para una mejor comprensión de la información.
Licenciamiento.	La herramienta debe ser implementada con herramientas de código abierto.

En el apéndice A, se tiene el documento de especificación de requerimientos de software (ERS) del proyecto.

Una vez que se identificaron los principales requerimientos del área comercial, se procedió a definir el alcance del proyecto, el cual se describe en el siguiente punto.

### 4.3 Definición del Alcance

Se desarrolló e implementó un Data-Driven DSS que permite el análisis de los datos desde diferentes perspectivas; permitiendo el análisis de las ventas desde un nivel resumido hasta un nivel detallado. Además, se puede comparar los datos actuales con datos históricos para analizar el comportamiento de las ventas en distintos periodos, lo cual permitirá tomar decisiones más acertadas para mejorar la gestión del área comercial. También se dispone de indicadores, tanto en unidades, dólares y toneladas.

A continuación se detalla el alcance definido:

- Almacén de datos donde se consolide la información del área comercial: clientes, pedidos, entregas, facturas, y no conformidades internas y externas.
- Reportes de análisis de las ventas efectuadas que pueden ser filtradas en diferentes periodos: diario, mensual, anual.
- Reportes para comparar toneladas producidas y toneladas vendidas en diferentes periodos: diario, mensual, anual.
- Reportes de análisis de los pedidos ingresados que pueden ser filtradas en diferentes periodos: diario, mensual, anual.
- Reportes de análisis de las entregas para establecer el cumplimiento en términos de unidades y tiempo que pueden ser filtradas en diferentes periodos: diario, mensual, anual.

- Cuadro de mando para integrar los reportes definidos en los puntos anteriores, mostrar comparativos entre diferentes periodos y gráficos para una mejor comprensión de la información.
- Indicadores de: unidades vendidas y unidades producidas, dólares vendidos por unidad de tiempo, toneladas vendidas y producidas por unidad de tiempo.

Los entregables del proyecto son los siguientes:

- Modelos de base de datos dimensionales del área comercial.
- Proceso de extracción ETL (Pentaho Data Integration) del área comercial.
- Cubo del área comercial.
- Cuadro de mando del área comercial.
- Reportes del área comercial (Vistas de Análisis).
- Indicadores.

Después de definir el alcance del proyecto y los entregables del mismo, se procedió a diseñar la solución de acuerdo al análisis realizado.

#### **4.4 Diseño de la solución**

Luego de analizar los requerimientos expuestos en las secciones anteriores, se identificó que para satisfacer las necesidades planteadas es necesario la creación de cuatro cubos OLAP: Gestión de Pedidos, Despacho de Productos, Facturación de Productos e Ingreso de Producto a bodega,

considerando que la información que se debe mostrar proviene de diferentes procesos que involucran al área comercial. Se determinó la necesidad de consolidar los datos de estos procesos, de tal forma que se logre cumplir con los requerimientos de información establecidos en el alcance del proyecto. Se analizaron dos opciones: creación de vistas transaccionales que provean los datos unificados de los distintos procesos o la creación de tablas de hechos por cada proceso. Al utilizar vistas transaccionales, los tiempos de respuesta para los procesos ETL se verían afectados debido a la cantidad de registros que se manejan. Además, se limitaría la cantidad de datos que se podría obtener de las diferentes fuentes y sólo se tendrían los datos para generar los reportes ya definidos. Por lo tanto, se decidió la creación de un cubo de datos por cada proceso del área comercial. Los cubos que se definieron se detallan a continuación:

- Cubo de Ventas: Almacena los datos de las ventas realizadas.
- Cubo de Entregas: Almacena los datos de los despachos realizados.
- Cubo de Pedidos: Almacena los datos de los pedidos de clientes realizados.
- Cubo de Ingresos: Almacena los datos de los ingresos de producto terminado a bodega.

De esta manera, se puede obtener toda la información que se requiera para la generación y presentación de los reportes. También se puede realizar los respectivos análisis de información desde diferentes perspectivas, de

acuerdo a cada proceso o comparando procesos; como, por ejemplo, unidades producidas versus unidades despachadas en determinados periodos.

Una vez definidos los cubos OLAP que se usan para la implementación del proyecto se procedió a identificar las entidades del modelo relacional desde donde se obtienen los datos.

#### 4.5 Identificación del modelo relacional

Como se ha indicado anteriormente, los datos para el cubo OLAP son extraídos desde una fuente de datos principal: SAPBO. Este sistema usa la base datos SQL Server 2012. Desde SAP BO se toman los datos de los pedidos, entregas, ventas y toneladas producidas. A continuación se describen las entidades de las cuales se extraerán los datos para ser almacenados en los datamarts.

Las entidades principales para la creación del datamart comercial se presentan en la Tabla 3.

Tabla 3. Entidades principales de orígenes de datos.

OBJETO	TABLA	DESCRIPCIÓN
Pedido	ORDR RDR1	Entidades que guardan los datos relativos a un pedido de ventas, tales como: Cliente, condición de pago, direcciones de entrega, detalle de productos, cantidades,

		impuestos aplicados, vendedor que realiza la venta.
Entrega	ODLN DLN1	Entidades que guardan los datos referentes a las entregas (despachos) que se realizan a los clientes. Aquí se registran los datos provenientes de los pedidos de ventas. Además, datos importantes sobre cantidades de productos que se entregan, los números de lote (número de pallet) asociados a la entrega, placa del transporte que lleva la carga, chofer, fecha y hora de salida de la planta y horas estimadas de llegada a las bodegas del cliente.
Factura	OINV INV1	Entidades que guardan los datos concernientes a la venta que se hace a los clientes. Se guardan los datos procedentes de la entrega junto con los respectivos precios unitarios, subtotales, tarifa de impuestos, valores de impuestos y totales.
Nota de Crédito	ORIN RIN1	Entidades que guardan los datos relacionados a las notas de créditos emitidas a los clientes. Pueden estar basadas en una factura de venta o ser independientes. Se registran los datos del cliente, las cantidades devueltas y dólares asociadas a la devolución.
Nota de Débito	OINV INV1	Entidades que guardan los datos de las notas de débito emitidas a los clientes. Están basadas en una factura de cliente. Se distingue si es una factura o nota de débito por el campo DOCSUBTYPE. Para facturas se tiene el valor "--" y para notas de débito, "DN".
Ingreso	GSP_vwConsultaInvent arioIngresosPT	Vista que muestra todos los ingresos realizados desde producción a la bodega de producto terminado.

Adicional a las estructuras mencionadas en la Tabla 3, se requiere de estructuras que proporcionen los datos maestros. Las entidades secundarias para crear el datamart del área comercial se presentan en la Tabla 4.

Tabla 4. Entidades secundarias de orígenes de datos.

OBJETO	TABLA	DESCRIPCIÓN
Cliente	OCRD	En esta entidad se guardan los datos del cliente como: Identificación, razón social, tipo de cliente, sector al que corresponde, condición de crédito.
Producto	OITM	En esta entidad se guardan los datos de los productos que son usados tanto para la venta como para la compra. Los datos que se registran son: grupo al que pertenece el producto, categoría y subcategoría, impuestos aplicados, mercado al que corresponde y sector industrial, entre otros.
Vendedor	OSLP	En esta entidad se guardan los datos de los vendedores como nombres, teléfonos, correos electrónicos.
Unidad Medida	EXX_UNIDADES_MEDIDA	En esta entidad se guardan las unidades de medidas que son aplicables a los productos.
Condición Crédito	OCTG	En esta entidad se guardan las condiciones de crédito aplicadas a los clientes.
Dirección	CRD1	En esta entidad se guardan las direcciones asociadas a los clientes. Se clasifican en direcciones de envío y de entrega de factura.

Luego de identificar las entidades del modelo relacional, que son los orígenes de datos para la creación del datamart comercial de la empresa de manufactura, se definieron las dimensiones y medidas que se utilizarán en el modelo dimensional.

#### 4.6 Definición de las dimensiones y medidas

Después de establecer el diseño de la solución e identificar el modelo relacional de la empresa de manufactura, se procedió a definir las tablas de dimensiones y hechos que forman parte del modelo dimensional de la solución propuesta.

Como se definió en la sección anterior, el modelo dimensional consta de cuatro cubos OLAP. Estos cubos tienen dimensiones compartidas y dimensiones propias. Las dimensiones compartidas son: Tiempo, Producto, Cliente, Localidad, Vendedor, Condición de Crédito, Unidad de Medida. En la Tabla 5 se detallan las dimensiones compartidas.

Tabla 5. Dimensiones Compartidas.

DIMENSIÓN	DESCRIPCIÓN
Tiempo	Esta dimensión almacena los distintos periodos que se usarán para el análisis de la información. Se guardarán los días, semanas, meses, semestres, años.
Producto	Esta dimensión guarda los datos de los productos de ventas. Se asociará con el grupo al que pertenece y las correspondientes categorías.
Cliente	Esta dimensión almacena los datos de los clientes, el sector industrial al que pertenecen y permitirá identificar si son locales o del extranjero.
Localidad	En esta dimensión se guardan los datos de las localidades: país, provincia, ciudad.
Vendedor	En esta dimensión se almacenan los datos de los ejecutivos comerciales que forman parte de la empresa y que son parte importante del proceso de ventas. El análisis de ventas por vendedor es parte clave para la organización ya que se pueden encontrar debilidades que necesiten ser fortalecidas u otros aspectos a mejorar.
Condición de crédito	En esta dimensión se almacenan las distintas

	condiciones de crédito que se aplican a los clientes. Estas pueden ser: contado, 30 días, 60 días, 90 días, etc.
Unidad de Medida	En esta dimensión se registran las unidades de medida que son aplicadas a los productos.
Tipo de Pedido	En esta dimensión se almacenan los tipos de pedidos que se manejan en la empresa: Sólo Despacho, Inventario, Inventario y Despacho, Fabricación Externa y Pads.
Estado	En esta dimensión se guardan los diferentes estados de las transacciones: Cancelado, Cerrado, Abierto y Abierto-Impreso.

Las dimensiones propias de cada cubo se detallan en la Tabla 6.

Tabla 6. Dimensiones propias.

CUBO	DIMENSIÓN	DESCRIPCIÓN
Ventas	Documento	En esta dimensión se guardan los datos de las ventas: facturas, notas de crédito y notas de débito. Se guardarán los datos de estos documentos a nivel de cabecera.
Pedido	Pedido	En esta dimensión se guardan los datos de la cabecera de pedidos de clientes.
Entrega	Entrega	En esta dimensión se registran los datos de cabecera de la tabla de entregas del modelo relacional.
Ingreso	Ingreso	En esta dimensión se registran los datos de cabecera de la vista de ingresos a bodega del modelo relacional.

Una vez que se identificaron las dimensiones que son parte del cubo comercial, se establecieron las tablas de hecho necesarias que intervienen en el modelo dimensional de la solución.

En la Tabla 7 se definen las tablas de hecho para el modelo dimensional.

Tabla 7. Tablas de Hecho.

HECHO	DESCRIPCIÓN
VENTAS	En esta tabla se almacenan los datos de las ventas: Clientes, productos, condiciones de crédito y todos los datos de las dimensiones que intervengan con las ventas, así como los valores monetarios y cantidades.
PEDIDO	En esta tabla se guardan todos los datos referentes a los pedidos de ventas de los clientes: cantidades, precios unitarios, totales.
DESPACHO	En esta tabla se almacenan los datos de las entregas que se realizan. Detalle de productos, unidades despachadas, precios unitarios, pesos unitarios, totales en dólares.
INGRESOS	En esta tabla se registran los datos de los ingresos de producto terminado a la bodega: productos, cantidad ingresada, costo de producción.

Asimismo, se establecieron las medidas principales que formarán parte del modelo. En la Tabla 8 se definen las medidas identificadas para el modelo dimensional.

Tabla 8. Medidas.

MEDIDA	DESCRIPCIÓN
Unidades Vendidas	Se utiliza para determinar las cantidades en unidades de producto que fueron facturadas a los clientes.
Peso Vendido	Se emplea para calcular la cantidad de kilos vendidos en base al peso unitario que tiene cada producto.
Total vendido en dólares	Se utiliza para establecer el total vendido en términos monetarios.
Unidades despachadas	Se usa para determinar las cantidades en unidades de producto que fueron despachadas pero que pueden o no haber sido ya facturadas.
Peso Despachado	Se utiliza para calcular la cantidad de kilos despachados en base al peso unitario que tiene cada producto.
Unidades pedidas	Se emplea para definir la cantidad en unidades de productos que fueron pedidas por los clientes

	para producir en determinados periodos. No se considerarán las unidades pedidas para sólo despacho.
Peso Pedido	Se utiliza para calcular la cantidad de kilos pedidos en base al peso unitario que tiene cada producto.
Unidades producidas	Se utiliza para establecer las cantidades en unidades de producto que fueron producidas en determinados periodos.
Peso producido	Se utiliza para calcular la cantidad de kilos producidos en base al peso unitario que tiene cada producto.

Una vez que se definieron las tablas de dimensiones y de hechos que forman parte del modelo dimensional, se procedió a diseñar el modelo que se usó para el desarrollo e implementación de la solución propuesta.

#### **4.7 Diseño del modelo dimensional**

Para el diseño del modelo dimensional se ha optó por el esquema estrella debido a que se tienen todos los datos en las tablas de dimensiones. De esta manera, los datos pueden ser leídos directamente sin necesidad de ir a otras estructuras a extraer descripciones o valores, lo cual mejora el rendimiento de la base de datos.

En el modelo se especifican todas las dimensiones que son parte primordial del cubo comercial y reflejan la realidad de la empresa. Asimismo, se detallan las relaciones existentes entre las tablas de dimensiones y las tablas de hecho.

Es importante indicar que todas las tablas de dimensión tienen tres campos obligatorios, los cuales se detallan a continuación:

- Versión: Indica el número de versión del registro.
- Fecha Inicio: Indica la fecha inicial de validez del registro.
- Fecha Fin: indica la fecha final de validez del registro.

Estos campos se actualizarán cada vez que en la carga de datos se determine que han existido cambios en los datos fuentes.

#### **4.7.1 Diseño del Cubo OLAP de Pedidos**

En la Figura 4.2 se muestra el modelo dimensional de Pedidos. Este modelo permite la explotación de los datos referentes a los pedidos de los clientes y permite analizar el número de pedidos realizados en determinados periodos a nivel de clientes, sectores y mercados.

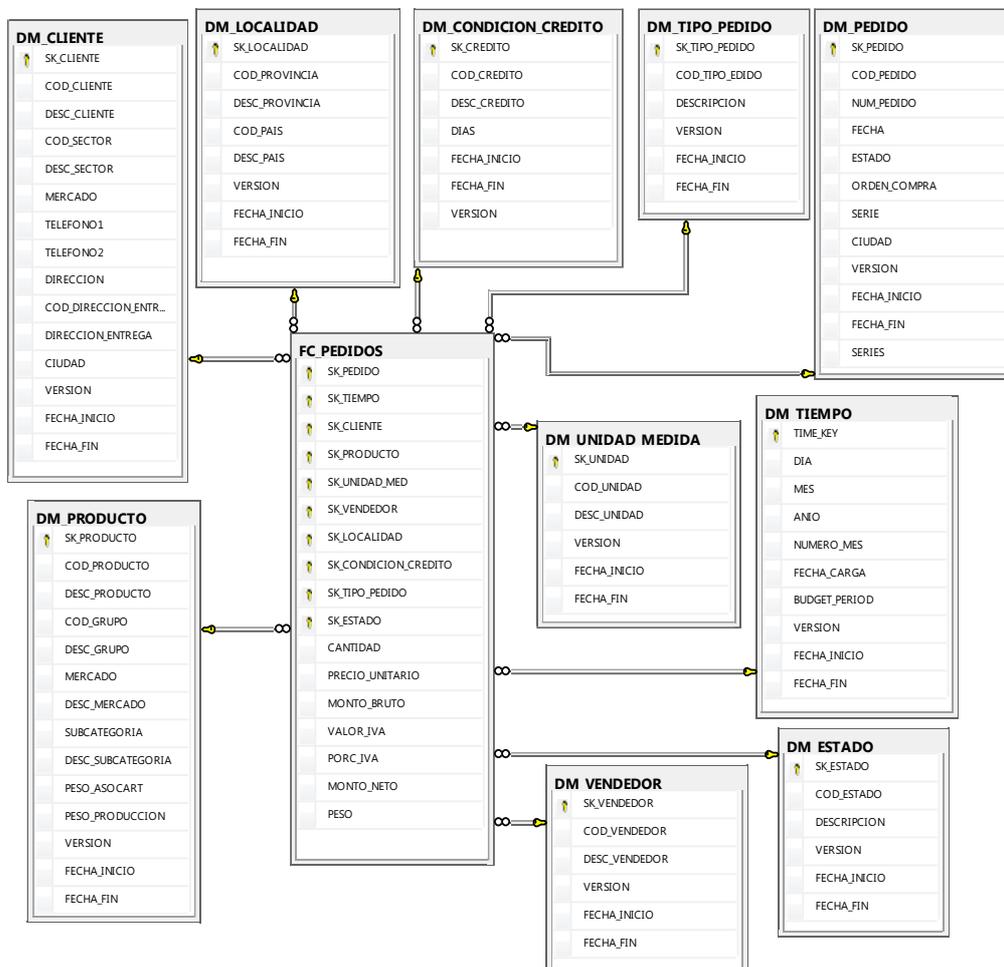


Figura 4.2 Modelo Dimensional de Pedidos.

### 4.7.2 Diseño del Cubo OLAP de Ventas

En la Figura 4.3 se muestra el modelo dimensional de Ventas. Con este modelo se analizan las ventas realizadas por periodos, permitiendo la segregación por clientes, localidades, vendedores y productos.

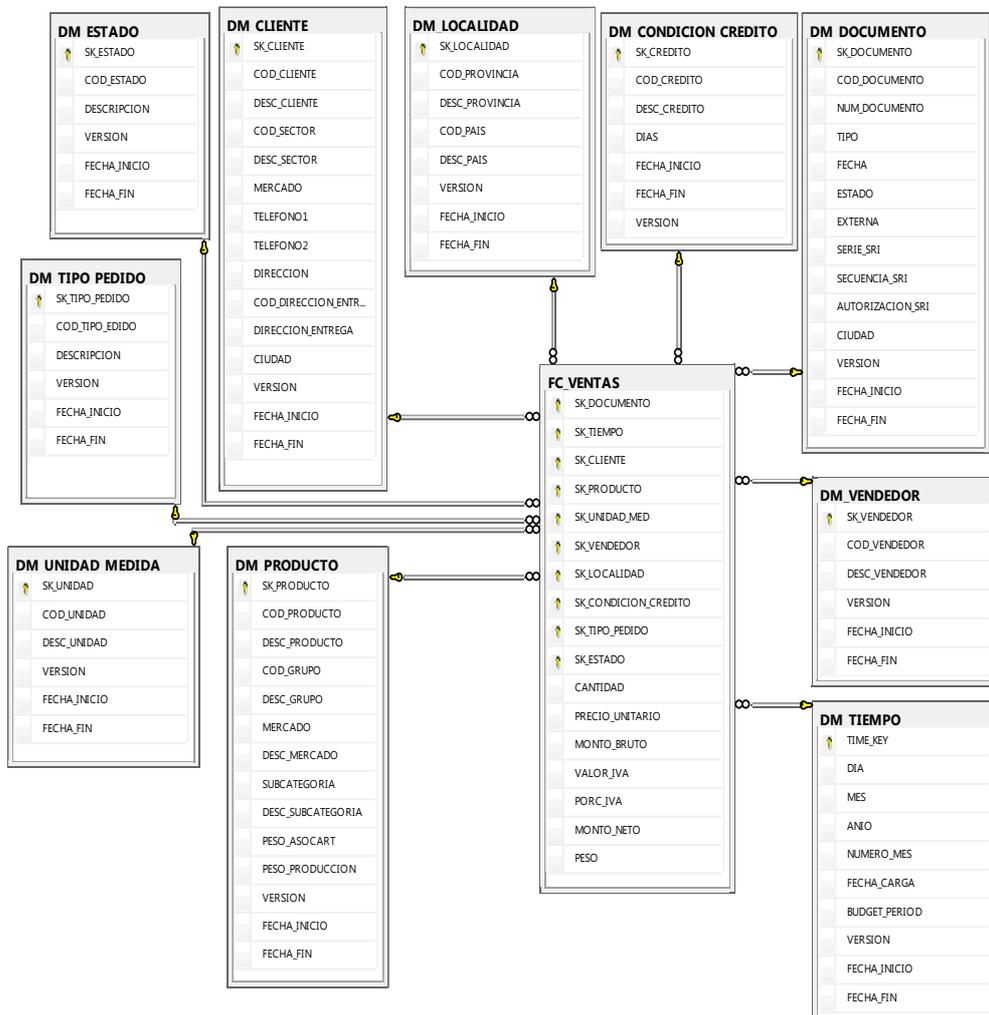


Figura 4.3 Modelo Dimensional de Ventas.

### 4.7.3 Diseño del Cubo OLAP de Entregas

En la Figura 4.4 se muestra el modelo dimensional de Entregas. Con este modelo se examinan las entregas realizadas en determinados periodos para determinar el nivel de cumplimiento en el despacho.

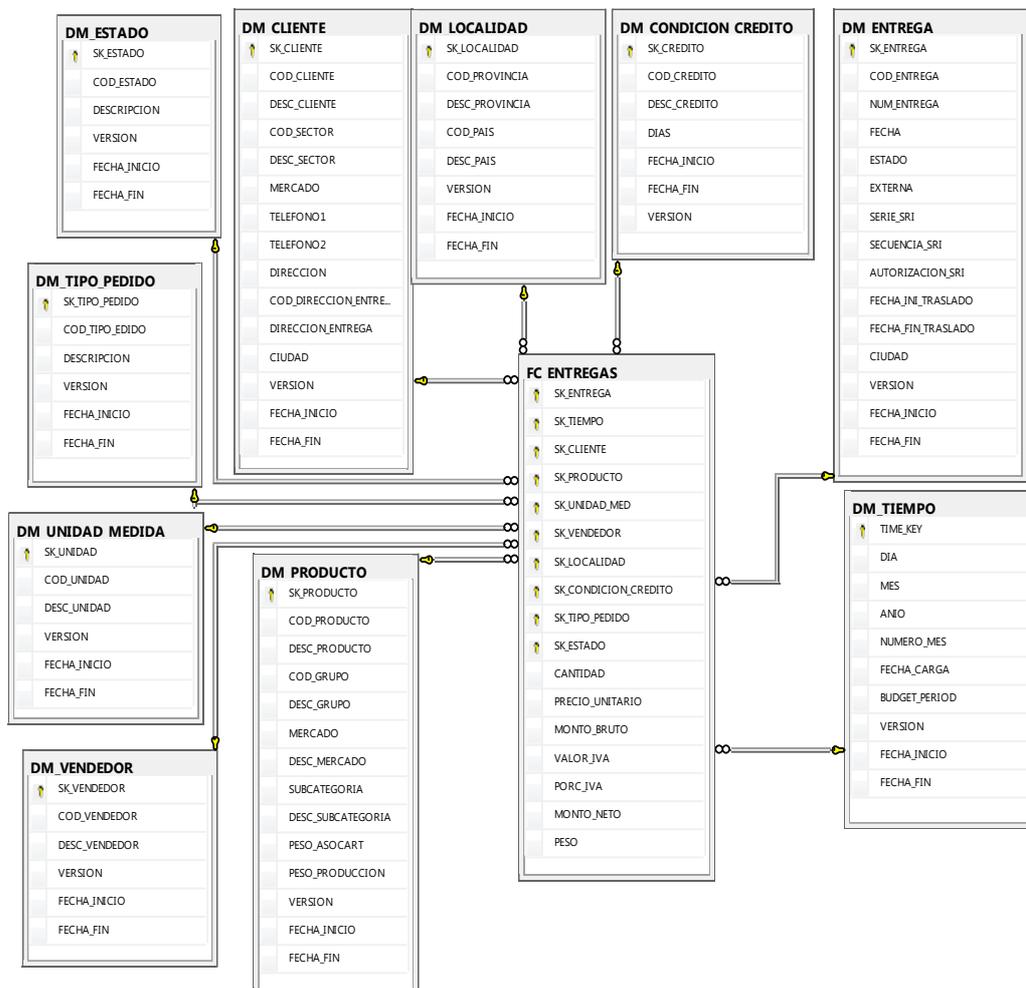


Figura 4.4 Modelo Dimensional de Entregas.

#### 4.7.4 Diseño del Cubo OLAP de Ingresos

En la Figura 4.5 se muestra el modelo dimensional de Ingresos. Con este modelo se determinan los ingresos de producto terminado realizados por periodos a nivel de productos.

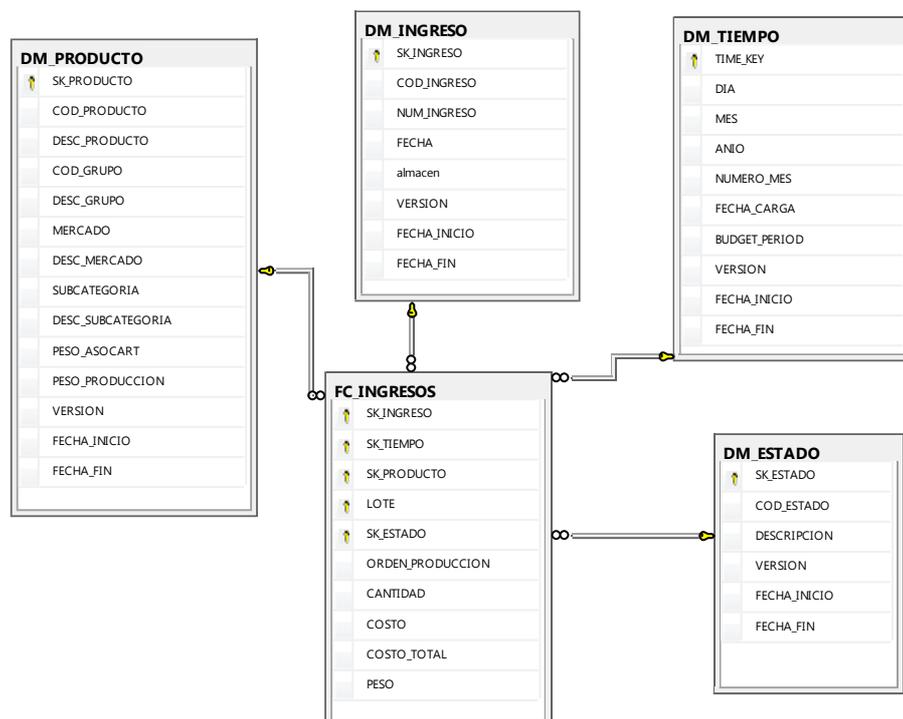


Figura 4.5 Modelo Dimensional de Ingresos.

En el presente capítulo se identificaron las entidades del modelo relacional del área comercial de las cuales se extraen los datos para el proceso ETL. Además, se definieron las dimensiones y tablas de hecho que forman parte del cubo comercial. Finalmente, se diseñó el modelo dimensional sobre el cual se creó el cubo OLAP.

En el siguiente capítulo se describe el desarrollo de la solución propuesta. Se mostrará la creación de los diferentes procesos ETL, del cubo OLAP, de los reportes dinámicos y estáticos, de los indicadores y se realizará la consolidación de los reportes en un cuadro de mando para mejor apreciación y comprensión para los usuarios finales.

## **CAPÍTULO 5**

### **DESARROLLO DE LA SOLUCIÓN**

En este capítulo se describen los procesos que se siguieron para el desarrollo e implementación de la herramienta propuesta. Se detallan los pasos para la creación de los ETL, la creación del cubo OLAP y los diferentes reportes que se integran en el dashboard para el área comercial.

#### **5.1 Creación del proceso ETL**

De acuerdo a lo mencionado en el capítulo anterior, el modelo dimensional consta de dos tipos de tablas: dimensiones y de hecho; cada tipo de tabla requiere ser tratada de forma diferente ya que las primeras almacenan los detalles de la tabla de hechos (datos maestros del modelo relacional) y las segundas almacenan los datos numéricos (datos transaccionales del modelo relacional). Por esta razón, se ha dividido en dos partes el proceso ETL: proceso para las dimensiones y proceso para las tablas de hechos.

A continuación se describen los pasos que se siguieron para la carga de las tablas de dimensiones y las tablas de hechos.

### **5.1.1 Descripción de los pasos del proceso ETL**

Para la carga de las tablas de dimensión se siguen los siguientes pasos:

1. Leer la tabla maestra del modelo relacional de la que se extraerán los datos para la dimensión.
2. Realizar las transformaciones necesarias: aplicación de filtros a los registros, obtener la descripción de los campos que proceden de otras tablas maestras, remplazar valores nulos con valores predeterminados, obtención de la fecha de carga, entre otras. Se agrupó en un solo paso estas acciones debido a que no son obligatorias en todas las dimensiones.
3. Verificar el registro leído. Si el registro es nuevo, se inserta en la tabla de dimensión. De lo contrario, se verifica si hay cambios en los datos. Si el registro cambió, se inserta el registro en la tabla de dimensión con un nuevo código y una nueva versión; también se actualiza el campo de fecha validez del registro anterior a la fecha actual en que se realiza la carga.

En la Figura 5.1 se ilustra con un diagrama de actividad los pasos para la carga de las dimensiones.

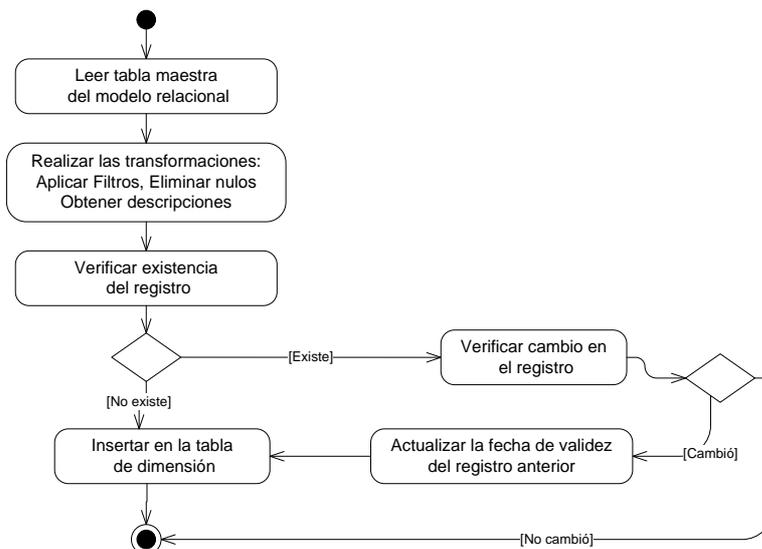


Figura 5.1 Diagrama de Actividad Carga de Datos de Dimensiones.

Para la carga de las tablas de hecho se siguen los siguientes pasos:

1. Obtener la fecha de carga del periodo. El periodo de carga de las tablas de hecho empieza el día uno del mes anterior a la fecha actual.
2. Borrar los registros desde la fecha inicial de carga hasta la fecha actual. Además, para el caso del ETL de Ventas se debe renombrar la fecha de la carga debido a que se extraen los datos desde dos tablas de origen diferente, Facturas y Notas de Crédito.
3. Leer los datos de cabecera de la tabla transaccional relacionada a la tabla de hecho desde la fecha inicial de carga del periodo.
4. Renombrar los parámetros que serán usados en el paso 5 como parámetros de entrada.

5. Leer los datos de detalle de acuerdo al código de la tabla de cabecera.
6. Extraer los códigos de las tablas de dimensiones de acuerdo al código transaccional y a la fecha del registro.
7. Extraer el código de la dimensión de tiempo según la fecha del registro.
8. Ordenar los registros para que pueda aplicarse el siguiente paso.
9. Proceder a eliminar los registros duplicados.
10. Insertar los datos en la tabla de hecho.

En la Figura 5.2 se muestra con un diagrama de actividad los pasos para la carga de las tablas de hecho.

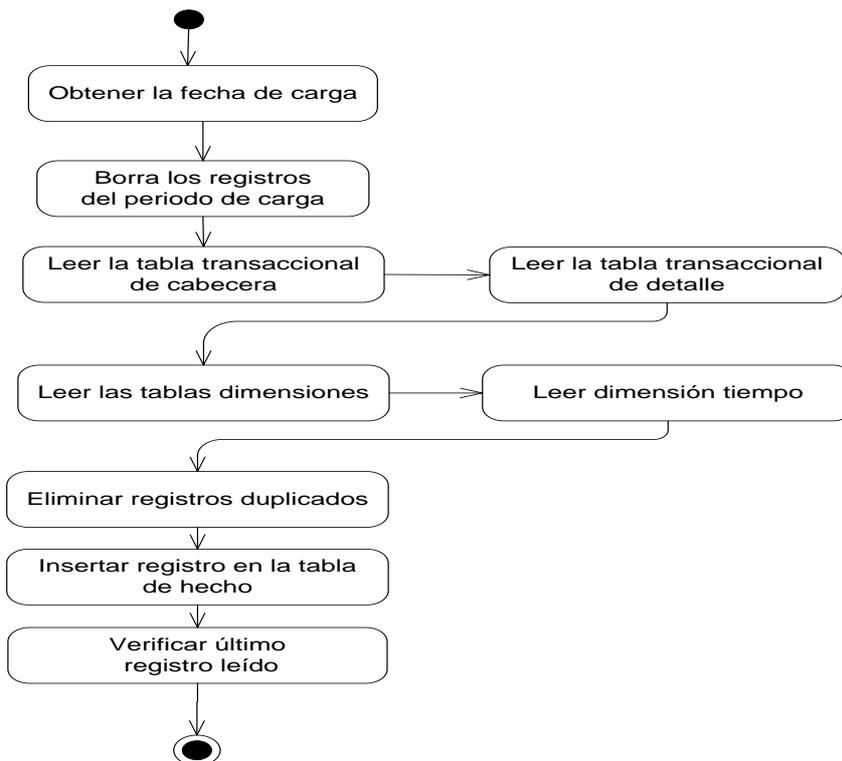


Figura 5.2 Diagrama de Actividades Carga de Datos Tabla de Hechos.

En la Tabla 9 se asocian los pasos de la carga de las dimensiones con los respectivos objetos de la herramienta Pentaho Data Integration (PDI, por sus siglas en inglés) [13]. La columna SEC. PASO identifica el número del paso de acuerdo a lo descrito anteriormente. Este esquema será usado para la creación de las transformaciones de las dimensiones.

Tabla 9. Esquema ETL Tabla de dimensión.

SEC. PASO	PASO/SUBPASO	NOMBRE ESTÁNDAR	OBJETO BD	OBJETO PDI	EXIGIDO
1	Lectura de datos	DBRD_DIMENSION	Tabla maestra principal del modelo relacional	 Table input	S
2	Filtro de datos	FILTRO_DIMENSION	N/A	 Filter rows	N
	Obtener descripción campos	DBDR_DESCAMPO	Otras tablas maestras asociadas a la tabla principal	 Database lookup	N
	Reemplazar valores nulos	ELIMINA_NULLS	N/A	 Null if...	S
	Obtener fecha actual	GET_FECHA	N/A	 Get System Info	S
3	Escritura de datos	DBWR_DIMENSION	Tabla de dimensión	 Dimension lookup/update	S

En la Figura 5.3 se ilustra el proceso ETL en PDI de las tablas de dimensión, donde se implementan los pasos descritos anteriormente.

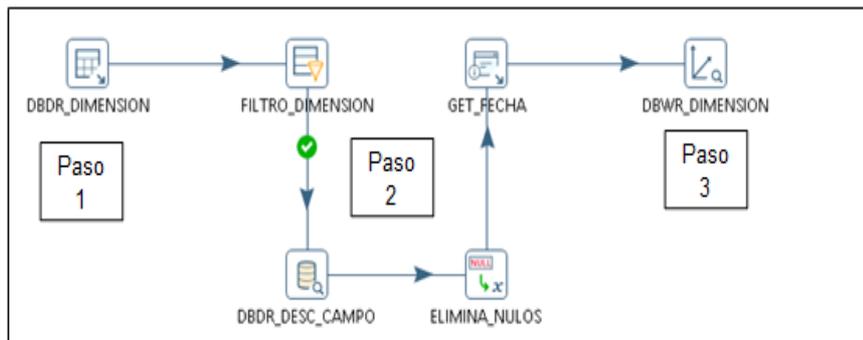


Figura 5.3 Esquema ETL Dimensión en PDI.

En la Tabla 10 se asocian los pasos de la carga de las tablas de hechos con los objetos PDI. Este esquema será usado para la creación de las transformaciones de las tablas de hecho.

Tabla 10. Esquema ETL Tabla de Hecho.

SEC. PASO	PASO/SUBPASO	NOMBRE ESTÁNDAR	OBJETO BD	OBJETO PDI	EXIGIDO
1	Obtener fecha inicial de carga	GET_FECHA_PRO	N/A	 Get System Info	S
2	Borrar los registros del periodo de carga	BORRADO_PERIODO	N/A	 Execute SQL script	S
	Renombrar Fecha	RENOMBRA_R_FECHA	N/A	 Select values	N
3	Lectura datos de cabecera	DBRD_FAC_TABLA	Tabla cabecera transaccional del modelo relacional	 Table input	S

4	Renombrar Parámetros	RENOMBRA R_PARAME TROS	N/A	 Select values	N
5	Lectura de datos de detalle	DBDR_FAC_ DETALLE	Tabla detalle transaccional del modelo relacional	 Table input	N
6	Obtener código de tabla de dimensión	SK_DIMENSI ON	Tabla de dimensión	 Dimension lookup/update	S
7	Obtener código de dimensión tiempo	SK_TIEMPO	Tabla de dimensión Tiempo	 Dimension lookup/update	S
8	Ordenar los registros	ORDENAR	N/A	 Sort rows	S
9	Eliminar registros duplicados	ELIMNAR_D UPLICADO	N/A	 Unique rows	S
10	Escritura de datos	DBWR_FAC _TABLA	Tabla de dimensión	 Dimension lookup/update	S

En la Figura 5.4 se ilustra el proceso ETL en PDI de las tablas de hecho, donde se implementan los pasos explicados arriba.

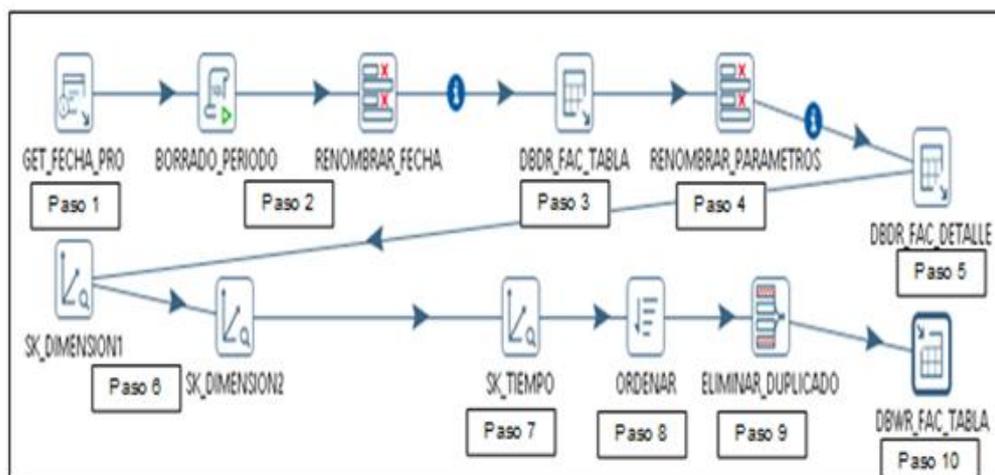


Figura 5.4 Esquema ETL Tabla de Hecho.

Seguidamente se indicará cómo se crearon las transformaciones en la herramienta de PENTAHO. Para la creación de los procesos ETL se usó la aplicación Spoon; un componente de PDI que permite crear las transformaciones que son necesarias para este proyecto.

### 5.1.2 Creación de las transformaciones de las tablas de dimensiones

En esta sección se describe el proceso que se siguió para la creación de las transformaciones de las tablas de dimensión en la herramienta Spoon. Se indica cómo se implementaron los pasos para la creación del ETL descritos anteriormente.

A continuación se detalla paso a paso la creación de la transformación de la dimensión Producto. Se eligió esta dimensión

porque es la más representativa del modelo y la más completa para mostrar la implementación de los pasos en la herramienta. En el Apéndice B se da a conocer a detalle la creación de las transformaciones de las demás dimensiones del modelo.

### 5.1.2.1 Transformación ETL dimensión Producto

Para implementar el proceso ETL, se siguieron los pasos descritos en la sección 5.1.1. Además, se crearon las conexiones a las bases de datos origen y destino que son usados en los pasos de lectura y escritura respectivamente.

En la Figura 5.5 se muestran las conexiones creadas.

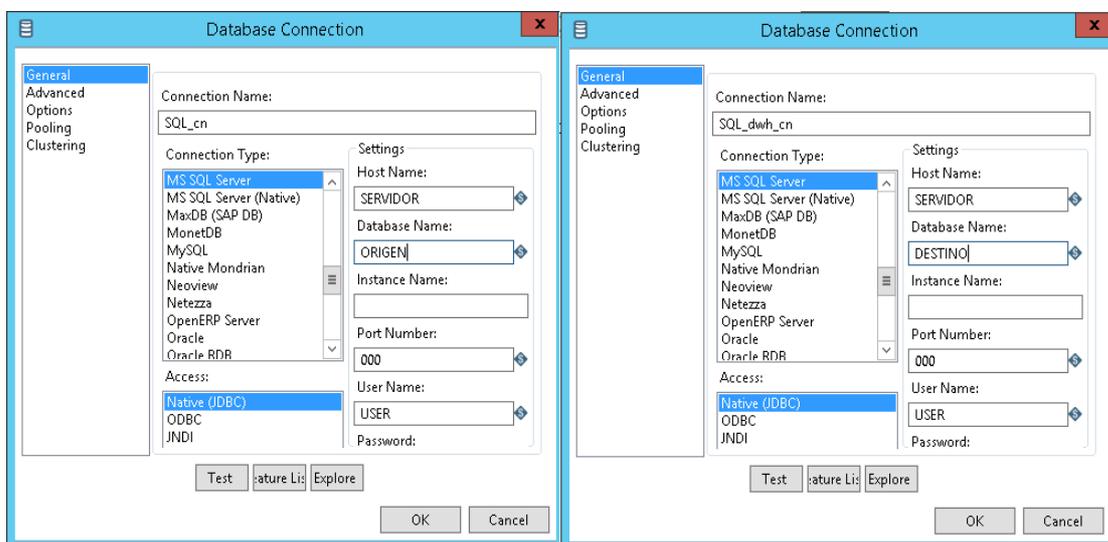


Figura 5.5 Conexiones a base de datos para ETL de las dimensiones.

Para cada transformación se crea su conexión origen (base de datos desde donde se extraerán los datos) y conexión

destino (base de datos en donde se almacenarán los datos transformados). Tanto la base de datos origen (ver Figura 5.5 a la izquierda) como la destino (ver Figura 5.5 a la derecha) son SQL Server 2012. Para el acceso se usó JDBC nativo. Cabe indicar que para realizar la conexión es necesario tener instalado el respectivo driver de SQL Server.

En la Figura 5.6 se muestra la transformación de la dimensión producto. En esta transformación se implementan los pasos indicados en el diagrama de actividades de las dimensiones. Se pueden observar los pasos y objetos PDI de acuerdo al modelo referido en la Figura 5.3.

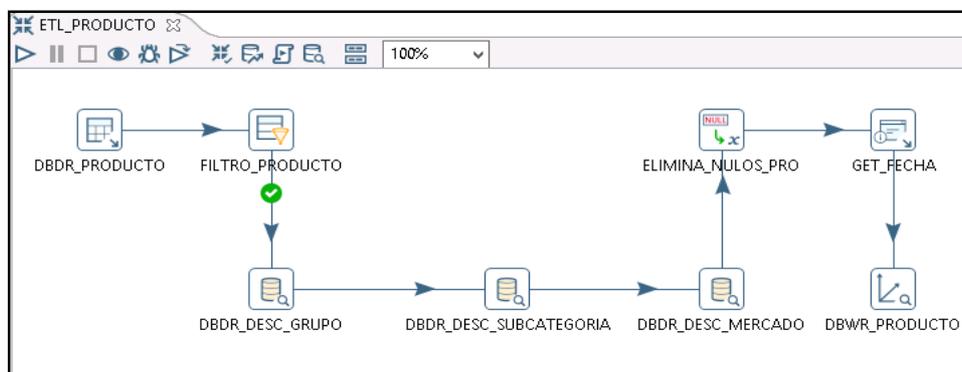


Figura 5.6 Transformación ETL Producto.

A continuación se detallan cada uno de los pasos usados en la transformación de Producto (Figura 5.6), de modo que se pueda comprender el proceso realizado desde la extracción de los datos hasta la carga en el datamart.

**Paso 1: Lectura de datos.** Paso del tipo INPUT; el objeto usado se llama TABLE INPUT. En este paso se recuperan los datos de la tabla maestra del modelo relacional. La tabla que almacena los productos es la OITM. En la Figura 5.7 se muestra la configuración del paso. El objeto tiene el nombre del paso, de acuerdo al nombre estándar definido en la Tabla 9: DBDR\_PRODUCTO. Se asocia a la conexión de datos origen previamente creada. Por último, se ingresa la consulta o procedimiento almacenado con el cual se obtendrán los datos de la base origen. En esta transformación se usó una consulta SQL a la tabla OITM.

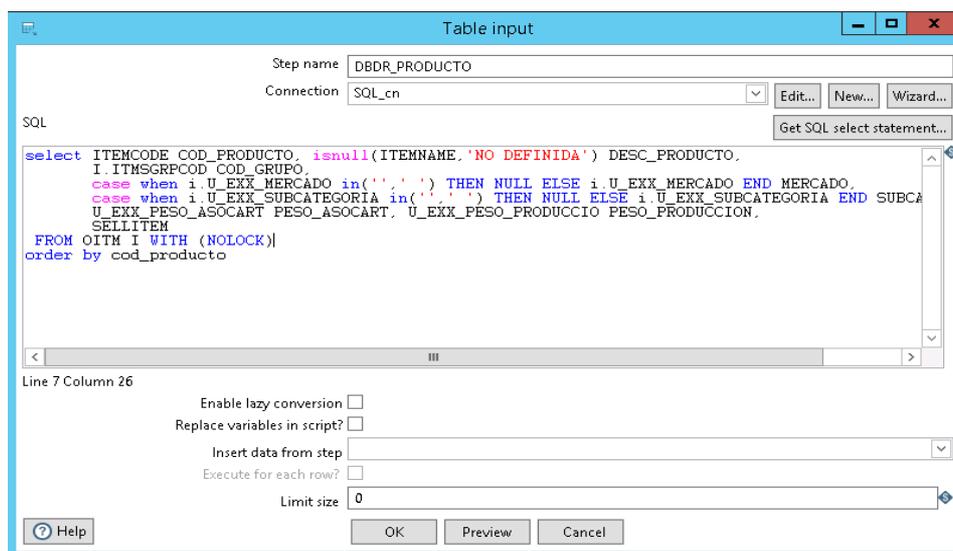


Figura 5.7 Lectura de la tabla Producto

**Paso 2: Filtrado de los datos.** Paso del tipo FLOW; el paso se llama FILTER ROWS. En este paso se filtran sólo los productos señalados para la venta. El nombre del paso es

FILTRO\_PRODUCTO. Cuando el resultado del filtro es verdadero, se continúa con el siguiente paso DBDR\_DESC\_PRODUCTO. En caso contrario, no se realiza ninguna acción. La condición se establece con el campo SELLITEM obtenido del paso 1; se verifica que el valor del campo sea igual a "Y". En la Figura 5.8 se muestra la configuración de este paso.



Figura 5.8 Filtrado de datos de Productos.

**Paso 2: Obtener descripción:** Paso del tipo LOOKUP llamado DATABASE VALUE LOOKUP. En este paso se obtienen las descripciones de los campos que son foráneos a la tabla producto. Se requieren las descripciones del grupo de artículos, de la subcategoría y del mercado al que corresponde el producto. En la Figura 5.9 se muestra la configuración para obtener la descripción del grupo de artículos. El nombre del paso es DBDR\_DESC\_PRODUCTO. Este paso se conecta a la base

de datos origen usando la conexión anteriormente indicada para extraer la descripción de la tabla OITB. Se realiza la búsqueda comparando el campo de la tabla origen llamado ITMSGRPCOD con el campo COD\_GRUPO que proviene del paso 1. El valor devuelto es el campo ITMSGRPNAM que es almacenado en el nombre DESC\_GRUPO en la transformación. No se han definido valores por defecto.

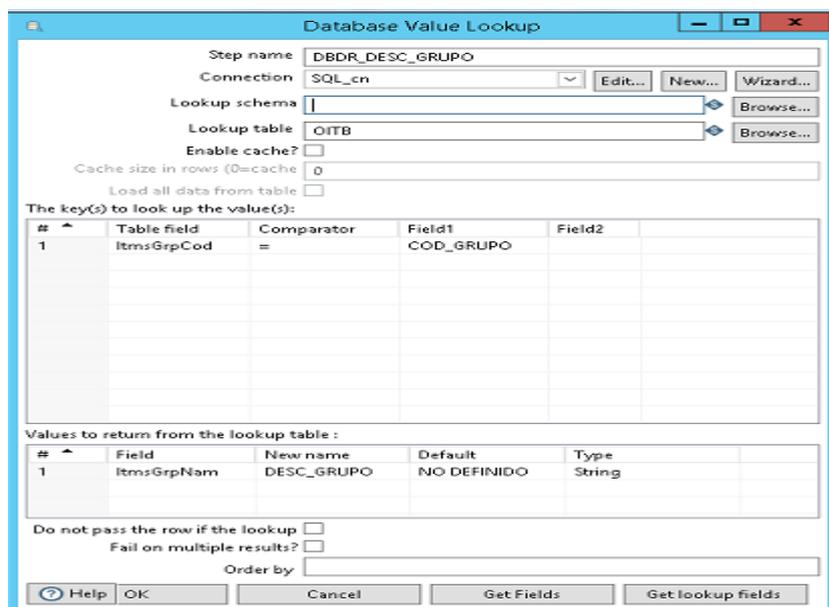


Figura 5.9 Obtención de descripción grupo de artículos.

En la Figura 5.10 se muestra la configuración para la obtención de la descripción de la subcategoría, el nombre del paso es DBDR\_DESC\_SUBCATEGORIA. También se conecta a la base de datos origen y se obtiene el valor de la tabla @EXX\_SUB\_CAT\_ART. En este paso se comparan los campos CODE (Campo de la tabla origen) y

SUBCATEGORIA (Campo de la transformación proveniente del paso 1). El valor se devuelve del campo NAME y se almacena en el campo DESC\_SUBCATEGORIA. No se registran valores por defecto.

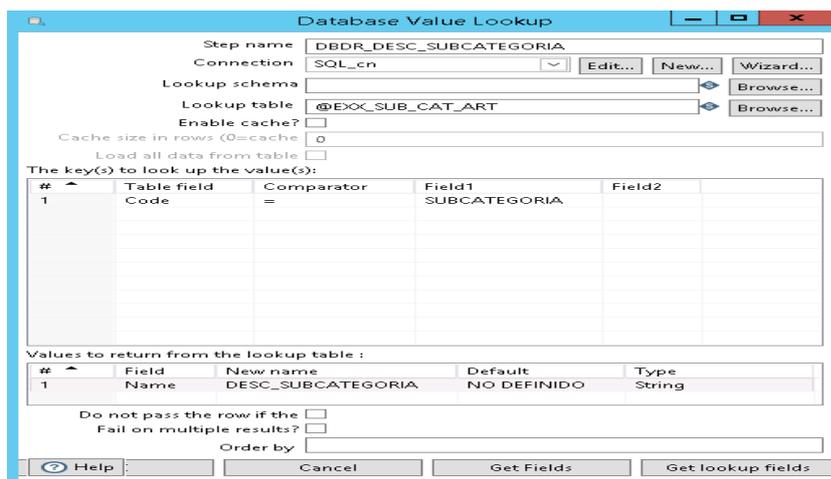


Figura 5.10 Obtención de la descripción de la subcategoría.

En la Figura 5.11 se muestra la configuración para obtener la descripción del mercado; para ello se utiliza la conexión origen que permite el acceso a la tabla OPRC. Los campos que se comparan para la búsqueda son: PRCCODE y MERCADO, de la tabla relacional y de la transformación respectivamente. La descripción se devuelve del campo PRCNAME que es almacenado en el campo DESC\_MERCADO. No se han definido valores por defecto.

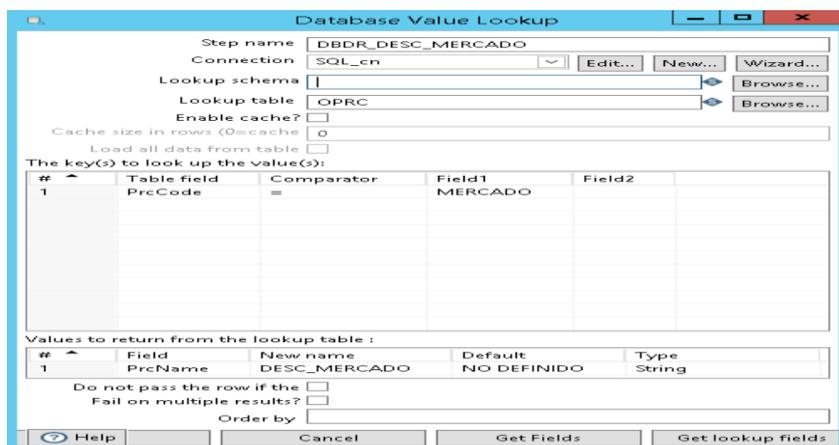


Figura 5.11 Obtención de la descripción del Mercado.

**Paso 2: Remplazar Valores Nulos:** Paso del tipo UTILITY llamado REPLACE NULL VALUE. En este paso se reemplazan los valores nulos, de los campos extraídos en el paso 1, por valores predeterminados. En la Figura 5.12 se muestra la configuración de este paso. El nombre del paso es ELIMINA\_NULOS\_PRO. Se seleccionó la opción para seleccionar los campos de la transformación obtenidos en la lectura de datos. Los campos que se reemplazan son: MERCADO, SUBCATEGORIA, PESO\_ASOCART y PESO\_PRODUCCION, por el valor cero.

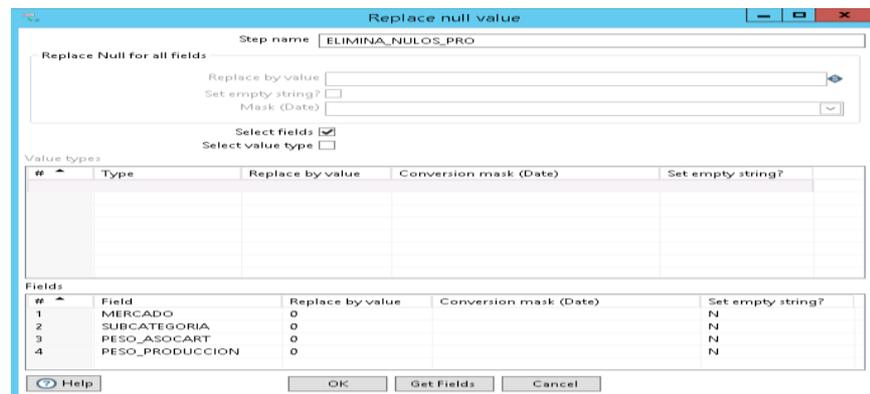


Figura 5.12 Remplazo de Valores Nulos.

**Paso 2: Obtener Fecha Actual:** Paso del tipo INPUT llamado GET SYSTEM DATA. Se utiliza para obtener la fecha actual del sistema que se usará como fecha de carga de la dimensión. En la Figura 5.13 se muestra la configuración del paso. Este paso se llama GET\_FECHA. Se obtiene el valor del sistema del tipo SYSTEM DATE que devuelve la fecha en formato AAAA-MM-DD H24:MI:SS. El nombre del campo que guardará la fecha es FECHA\_EFECTIVA.

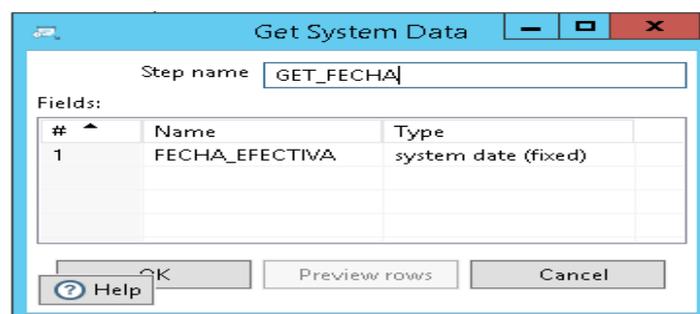


Figura 5.13 Obtener Fecha del Sistema.

**Paso 3: Escritura de los datos:** Paso del tipo DATA WAREHOUSE llamado DIMENSION LOOKUP/UPDATE. En este paso se escriben los datos en la dimensión producto: DM\_PRODUCTO. En la Figura 5.14 se muestra la primera parte de la configuración del paso. El paso se llama DBWR\_PRODUCTO. En este paso se realiza la conexión de la base de datos destino para escribir los datos extraídos y transformados en la dimensión DM\_PRODUCTOS. Debido a la cantidad de registros de la tabla origen se estableció el COMMIT cada mil registros. Es importante indicar que se debe chequear la opción UPDATE THE DIMENSION para que el paso realice las inserciones o actualizaciones necesarias en la dimensión. Si no se chequea el paso sólo realizarán búsquedas de la clave técnica de la dimensión.

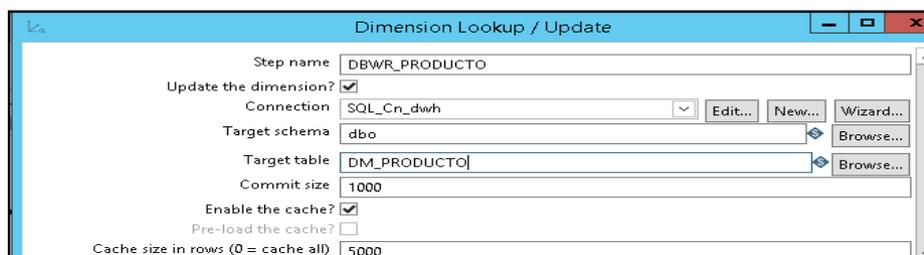


Figura 5.14 Escritura de la dimensión Producto

En la Figura 5.15 se observa cómo en la pestaña KEYS se busca el registro en la dimensión Producto, la búsqueda devolverá la clave técnica del registro en el caso de existir coincidencias. Para ello, se comparan los campos de la

dimensión (COD\_PRODUCTO) y de la transformación (COD\_PRODUCTO). Este campo corresponde a la clave primaria de la tabla relacional de productos. Para la búsqueda es indispensable que se incluya la fecha del registro, puesto que pueden existir diferentes versiones para un mismo código debido a los cambios que se pudieron haber dado en la tabla maestra. El campo de fecha que se utiliza es FECHA\_EFECTIVA que fue configurada en el paso anterior. La fecha se compara entre los campos FECHA\_INICIO y FECHA\_FIN de la tabla dimensión (campos que son obligatorios en todas las dimensiones).

Cache size in rows (0 = cache all) 5000

Keys Fields

Key fields (to look up row in dimension):

#	Dimension field	Field in stream
1	COD_PRODUCTO	COD_PRODUCTO

Technical key field SK\_PRODUCTO New name

Creation of technical key

Use table maximum + 1

Use sequence

Use auto increment field

Version field VERSION

Stream Datefield FECHA\_EFECTIVA

Date range start field FECHA\_INICIO Min. year 1900

Use an alternative start date?  <Select Option>

Table date range end FECHA\_FIN Max. year 2199

OK Cancel Get Fields SQL

Help

Figura 5.15 Escritura de datos-Verificación de registro.

En la Figura 5.16 se muestra la pestaña FIELDS. Donde se comparan los campos de la dimensión y los campos de la transformación para verificar si los datos han cambiado. En la columna TYPE OF DIMENSION UPDATE se seleccionó INSERT. Esto indica que si el valor ha cambiado se insertará un nuevo registro con una nueva clave de dimensión y versión. Además, el sistema procederá a actualizar de forma automática los datos de versión, fecha de inicio y fecha de fin del registro que cambió. Si no han existido cambios no se realiza ninguna acción. Cuando el registro no existe, se procede a realizar la inserción. En la Figura 5.15 se muestra asimismo, la clave técnica de la dimensión que se ha denominado SK\_PRODUCTO. Esta clave se incrementa en uno por cada registro nuevo. También se tiene el campo VERSION que se incrementa en uno por cada cambio en el mismo registro. Por último, se registran las fechas de inicio y de fin. La fecha de inicio comienza en el año 1900 y la fecha de fin culmina en el año 2199. Ese es el rango que se estableció para los registros.

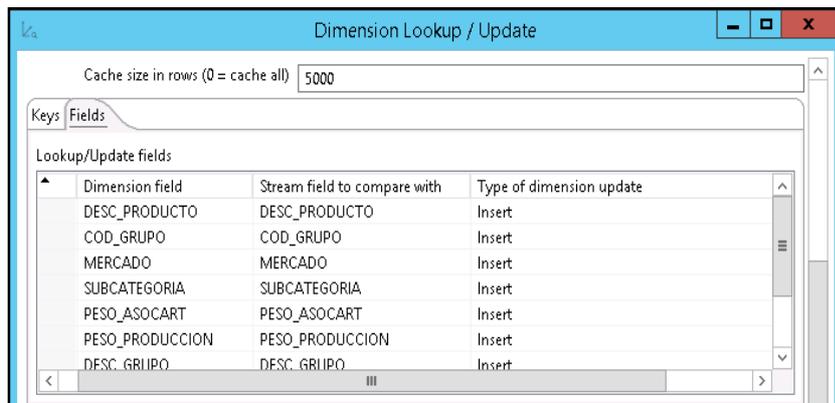


Figura 5.16 Escritura de datos – Campos de la dimensión.

Las configuraciones de los pasos de esta transformación se realizaron conforme a los pasos establecidos en el diagrama de actividades de la carga de las dimensiones.

En el siguiente punto se detallan los pasos que se siguieron para la creación de las transformaciones para las tablas de hechos.

### 5.1.3 Creación de las transformaciones de las tablas de hechos

Una vez que se realizaron las transformaciones para las tablas de dimensión, se procedió a crear las transformaciones de las tablas de hecho. Para ello se siguió el esquema definido en el diagrama de actividades para la carga de tablas de hechos.

Debido a que las ventas representan el mayor interés para el departamento comercial, se escogió la tabla de hechos de Ventas para explicar paso a paso cómo se creó la transformación. En el Apéndice B se detallan las transformaciones para las demás tablas de hecho. A continuación se describe detalladamente la creación de la transformación de Ventas.

### **5.1.3.1 Transformación ETL tabla hechos Ventas**

De la misma manera que en la creación de las transformaciones de las dimensiones, es necesario crear las respectivas conexiones a las bases de datos origen y destino por cada transformación. Para una mejor comprensión revisar la sección 5.1.2.1.

En la Figura 5.17 se muestran los pasos utilizados para la carga de la tabla de hechos de Ventas en la herramienta Spoon, de acuerdo al esquema indicado en la Figura 5.4.

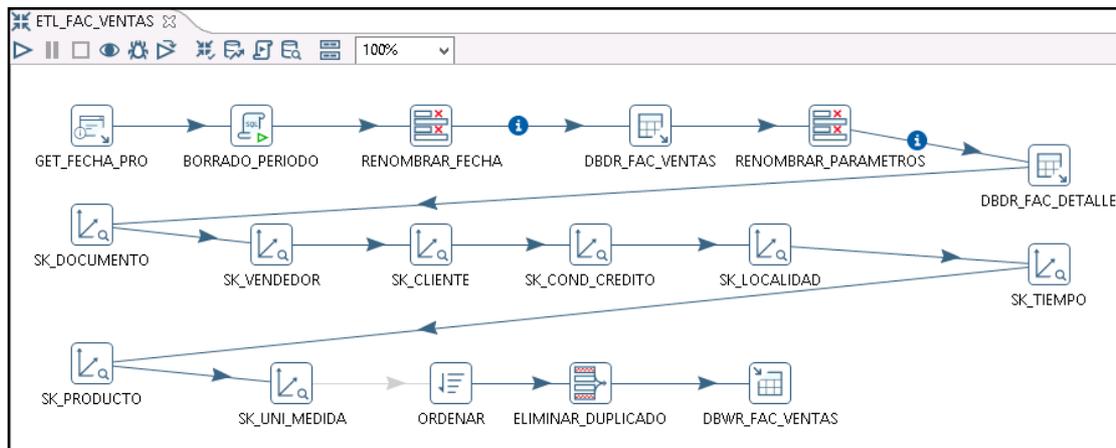


Figura 5.17 Transformación ETL Tabla de Hechos Ventas.

A continuación se detallan los pasos utilizados para la creación de la transformación de la tabla de hechos de Ventas.

**Paso 1: Obtener Fecha de Carga.** Paso del tipo INPUT llamado GET SYSTEM DATA. En este paso se obtiene la fecha de la carga que se utilizará para la eliminación de los registros de la tabla de hechos comprendidos entre esa fecha y la fecha actual y para la respectiva carga de datos en ese mismo periodo. En la Figura 5.18 se puede observar la configuración del paso. El nombre del paso es GET\_FECHA\_PRO. Se almacena el valor de la fecha en la variable FECHA\_BORRADO. El tipo del valor del sistema es el FIRST DAY OF LAST MONTH que devuelve el primer día del mes anterior en formato AAAA-MM-DD.

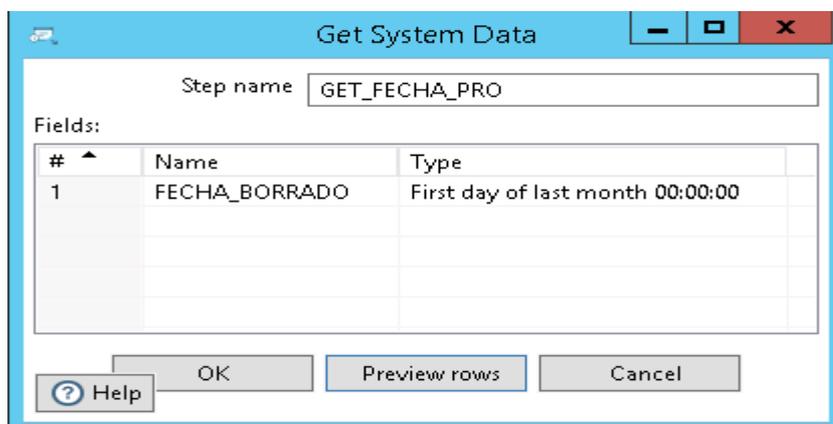


Figura 5.18 Obtención de Fecha de Periodo de Carga.

**Paso 2: Borrado de Registros.** Paso del tipo SCRIPTING llamado EXECUTE SQL STATEMENTS. Se procede a eliminar de la tabla de hechos de Ventas todos los registros desde la fecha de carga del periodo hasta la fecha actual. En la Figura 5.19 se observa la configuración. El paso se llama BORRADO\_PERIODO. Se conecta a la base de datos destino; es decir, la base donde están las tablas del modelo dimensional. En la sección de parámetros se indica el parámetro que recibe, en este caso la fecha de borrado, que es remplazado en la sentencia SQL donde se encuentra el signo “?”.

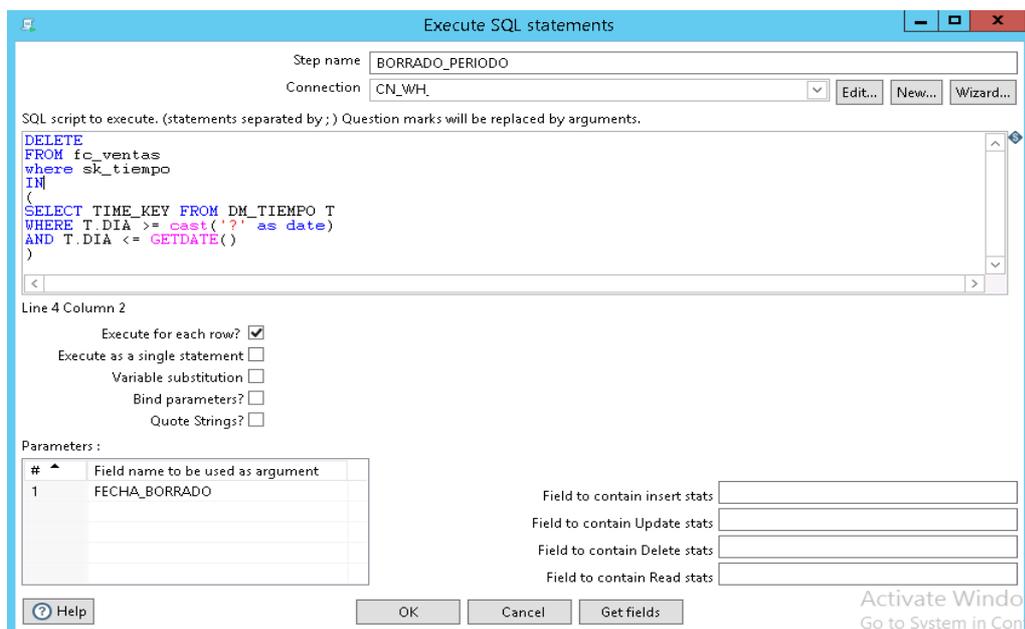


Figura 5.19 Borrado de Periodos de Carga de Ventas.

**Paso 3: Lectura de datos.** Paso del tipo INPUT llamado TABLE INPUT. En este paso se recuperan los datos de la tabla de cabecera transaccional de facturas y notas de crédito: OINV y ORIN. Este paso recibe como parámetro (?) la fecha inicial de carga. En la Figura 5.20 se muestra la configuración de este paso. El paso se llama DBDR\_FAC\_VENTAS. Este paso se conecta a la base de origen para acceder a los datos de las tablas relacionales. Para esta transformación fue necesario crear un paso adicional para renombrar la fecha de carga debido a que los datos se extraen desde dos tablas relacionales y cada una requiere un parámetro de fecha. En la parte inferior se indica

que recibe los valores de los parámetros (?) del paso RENOMBRAR\_FECHA.



Figura 5.20 Lectura Tabla de Facturas y Notas de Crédito.

En la Figura 5.21 se muestra el paso que sólo se utilizó para esta transformación que permite renombrar el parámetro de la fecha de borrado para ser utilizado como parámetros en la lectura de datos. En el siguiente punto se detalla este paso.

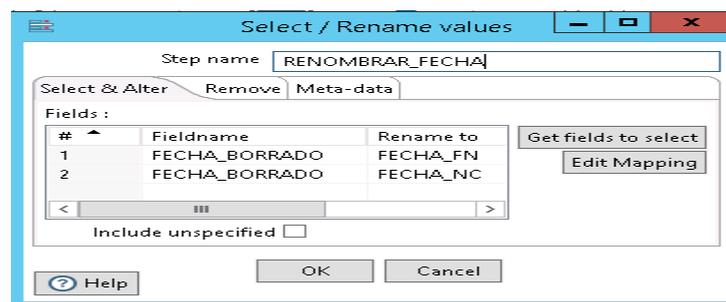


Figura 5.21 Renombrar Fecha de Borrado.

**Paso 4: Renombrar Parámetros.** Paso del tipo TRANSFORM llamado SELECT/RENAME VALUES. Se utiliza este paso para renombrar los parámetros que serán pasados al paso de lectura de detalle debido a que en la cabecera se leen datos de dos tablas: Facturas y Notas de Crédito. En la Figura 5.22 se muestra la configuración. Se selecciona el nombre devuelto en el paso anterior en la lectura de datos que será renombrado y se le asigna el nombre que será usado en el siguiente paso como parámetro.

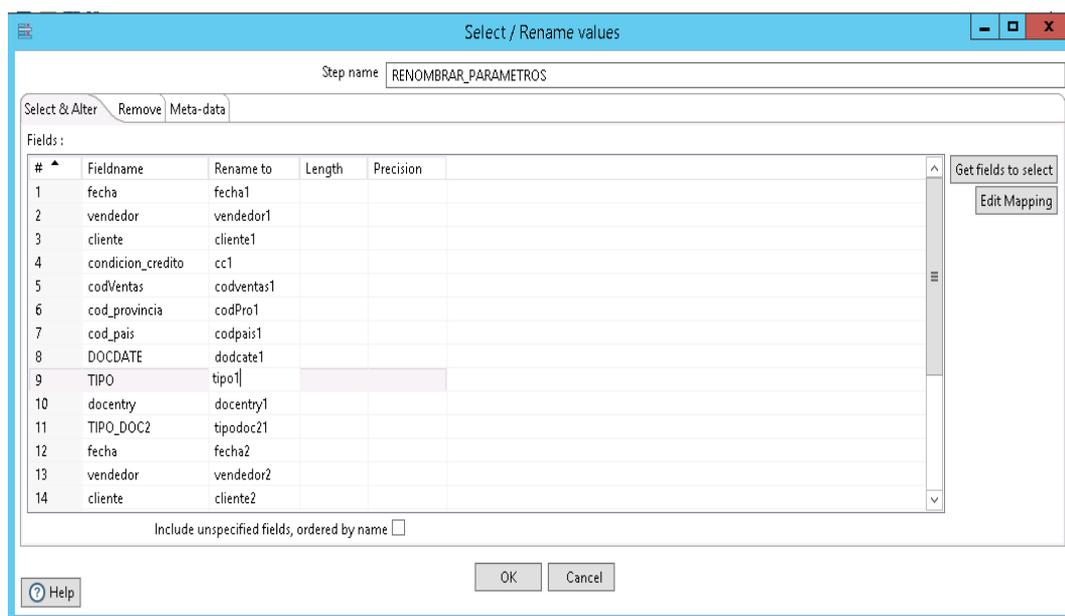


Figura 5.22 Paso Renombrar Parámetros Carga Hechos.

**Paso 5: Lectura de datos de detalle.** Paso del tipo INPUT llamado TABLE INPUT. En este paso se realiza la lectura de los datos de detalle de la tabla transaccional de facturas y

notas de crédito: INV1 y ORIN1. Este paso recibe como parámetro de búsqueda el código de la factura o nota de crédito y el tipo de documento (FACTURA/NOTA CREDITO) (Clave Foránea). Además, recibe como parámetros los campos de la cabecera que serán usados en los pasos posteriores para obtener los códigos de las tablas de dimensión asociadas. En la Figura 5.23 se muestra la configuración del paso. El nombre del paso es DBDR\_FAC\_DETALLE. Se conecta a la base de datos origen para extraer los datos. Adicionalmente, se observa en la parte inferior que recibe los parámetros del paso anterior RENOMBRAR\_PARAMETROS.

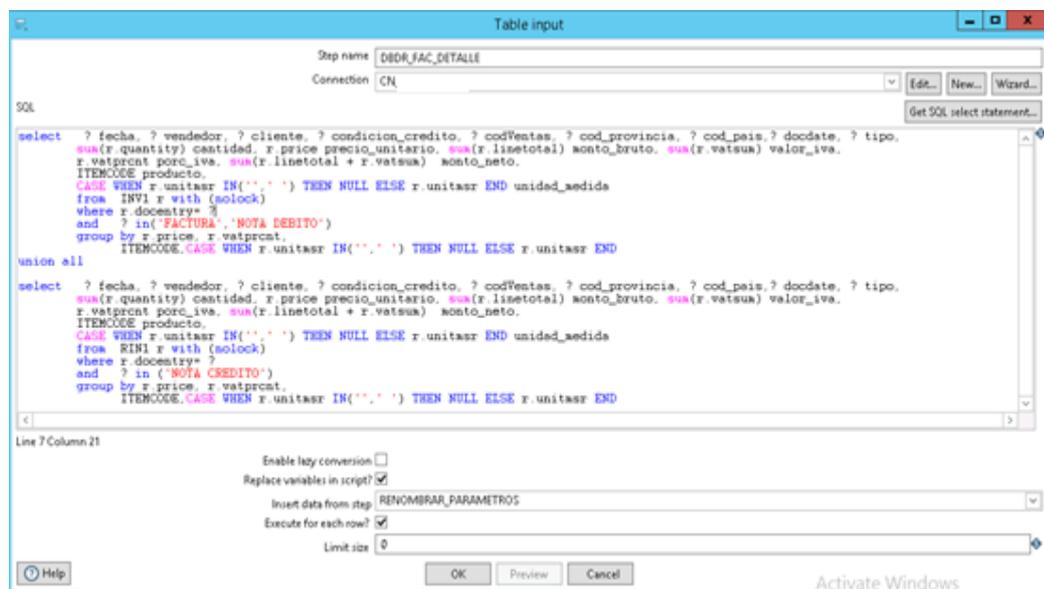


Figura 5.23 Lectura Tabla de Detalles de Ventas.

**Paso 6: Obtener los códigos de las dimensiones.** Paso del tipo DATA WAREHOUSE llamado DIMENSION LOOKUP/UPDATE. En este paso se obtiene el código de las dimensiones asociadas a la tabla de hecho. Reciben como parámetro el código de la tabla transaccional correspondiente y devuelve el código de la dimensión. Es importante indicar que para este caso no se va a realizar ninguna actualización en la dimensión por lo que la opción "Update the dimension?" se encuentra sin chequear. La tabla de hechos de ventas se enlaza con las dimensiones de Producto, Vendedor, Localidad, Unidad de Medida, Cliente y Condición de Crédito, por cada dimensión existe un objeto DIMENSION LOOKUP/UPDATE para extraer las claves técnicas de cada dimensión: SK\_PRODUCTO, SK\_VENDEDOR, SK\_LOCALIDAD, SK\_UNIDAD\_MEDIDA, SK\_CLIENTE y SK\_CONDICION\_CREDITO, respectivamente. En la Figura 5.24 se muestra la configuración del paso para la dimensión de productos. El paso se llama SK\_PRODUCTO. La búsqueda de la clave técnica se la realiza comparando los campos COD\_PRODUCTO de la dimensión Producto con el campo PRODUCTO generado en el paso anterior.

Step name: SK\_PRODUCTO

Update the dimension?

Connection: CN\_WH

Target schema: |

Target table: DM\_PRODUCTO

Commit size: 100

Enable the cache?

Pre-load the cache?

Cache size in rows (0 = cache all): 5000

Keys Fields

Key fields (to look up row in dimension):

#	Dimension field	Field in stream
1	COD_PRODUCTO	producto

Technical key field: SK\_PRODUCTO

New name: |

Help

Figura 5.24 Obtener Clave Dimensión Producto

Además, se compara la fecha del registro que se está cargando, con las fechas de inicio y fin de la dimensión para obtener la clave técnica en esa fecha. En la Figura 5.25 se muestra lo indicado.

Technical key field SK\_PRODUCTO New name

Creation of technical key

Use table maximum + 1

Use sequence

Use auto increment field

Version field version

Stream Datefield fecha

Date range start field FECHA\_INICIO Min. year 1900

Use an alternative start date?  <Select Option>

Table date range end FECHA\_FIN Max. year 2199

OK Cancel Get Fields SQL

Help

Figura 5.25 Campos para obtener Clave Técnica Dimensión Producto

**Paso 7: Obtener código de la dimensión de tiempo.** Paso del tipo DATA WAREHOUSE llamado DIMENSION LOOKUP/UPDATE. En este paso se obtiene el código de la dimensión de tiempo de acuerdo a la fecha de la transacción de venta que se está cargando.

**Paso 8: Ordenar los registros.** Paso del tipo TRANSFORM llamado SORT ROWS. En este paso se realiza el ordenamiento de los registros leídos. En la Figura 5.26 se muestra la configuración de este paso. Se seleccionan los nombres de los campos de la transformación por los cuales se van a ordenar los registros. También se indica si el ordenamiento es ascendente. El nombre del paso es ORDENAR.

Sort rows

Step name: ORDENAR

Sort directory: %%java.io.tmpdir%%

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files?

Only pass unique rows? (verifies)

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Presorted?
1	SK_DOCUMENTO	Y	N	N
2	SK_CLIENTE	Y	N	N
3	SK_LOCALIDAD	Y	N	N
4	SK_PRODUCTO	Y	N	N
5	SK_UNIDAD	Y	N	N
6	SK_CREDITO	Y	N	N

Buttons: Help, OK, Cancel, Get Fields

Figura 5.26 Ordenamiento de Registros.

**Paso 9: Eliminación de valores repetidos.** Paso del tipo OUTPUT llamado UNIQUE ROWS. En este paso se eliminan los registros duplicados para evitar problemas en la escritura de la tabla de hechos. Es importante que previo a este paso se realice el ordenamiento de los registros. En la Figura 5.27 se muestra la configuración del paso. El paso se llama ELIMINAR\_DUPLICADOS. Se seleccionan los campos de la transformación por los cuales se verificarán si existen registros duplicados.

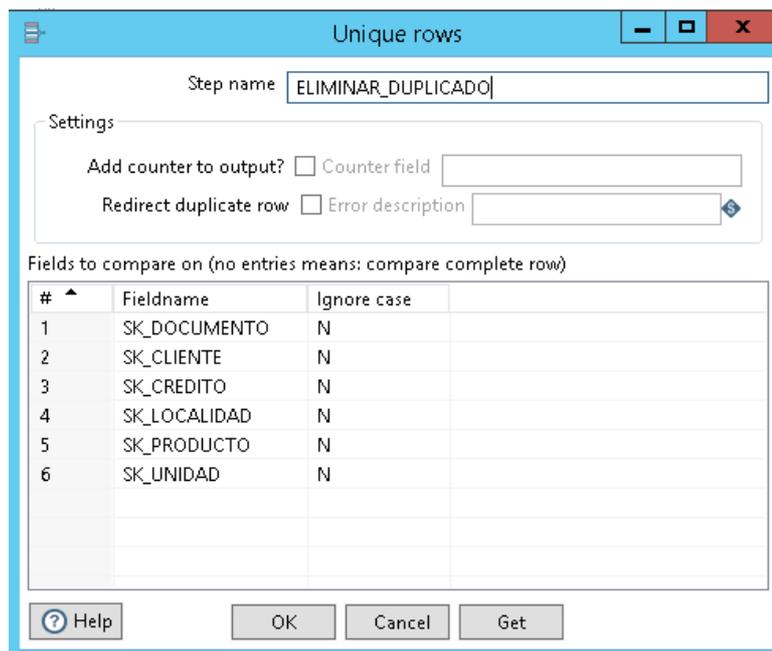


Figura 5.27 Eliminar Registros Duplicados.

**Paso 10: Escritura de datos.** Paso del tipo OUTPUT llamado TABLE OUTPUT. En este paso se escriben los registros leídos y transformados en la tabla de hechos: FAC\_VENTAS. En la Figura 5.28 se muestra la configuración de este paso. El nombre de este paso es DBWR\_FAC\_VENTAS. En este paso se conecta a la base de datos destino para realizar la escritura de los datos. El commit de los datos se hace cada mil registros. Se seleccionan los campos de la tabla con sus respectivos campos de la transformación.

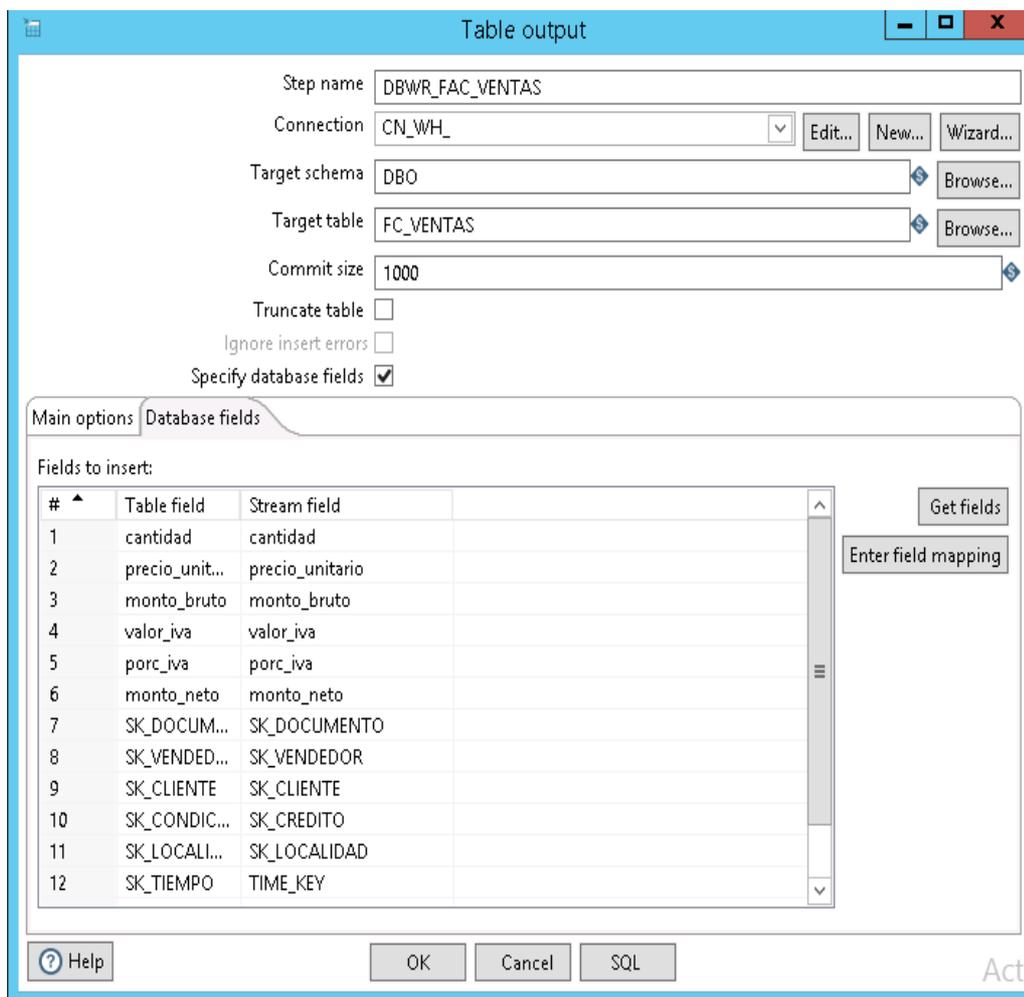


Figura 5.28 Escritura Tabla de Hechos Ventas.

En el siguiente punto se describe el trabajo que se creó para ejecutar los ETL's en forma periódica.

#### 5.1.4 Creación de trabajo para las cargas periódicas

Para que los ETL se ejecuten de manera automática se creó un JOB en la herramienta Spoon, el cual se ejecutará en horario nocturno

para evitar problemas de bloqueos o lentitud en los sistemas transaccionales. Este JOB se encarga de ejecutar cada una de las transformaciones en orden secuencial. En la Figura 5.29 se observa el JOB creado. Las transformaciones de las dimensiones y de las tablas de hechos son llamadas desde este proceso. Las líneas de flujo de color verde indican que la transformación se ejecutó con éxito. En este caso, se ejecuta la siguiente transformación. En caso de ser la última transformación, se envía un correo al personal de sistemas de la empresa indicando que la carga se realizó correctamente. Las líneas de flujo de color rojo indican que hubo error al ejecutar la transformación y se envía un correo al personal de sistemas indicando el error que se presentó; por último, el JOB termina.

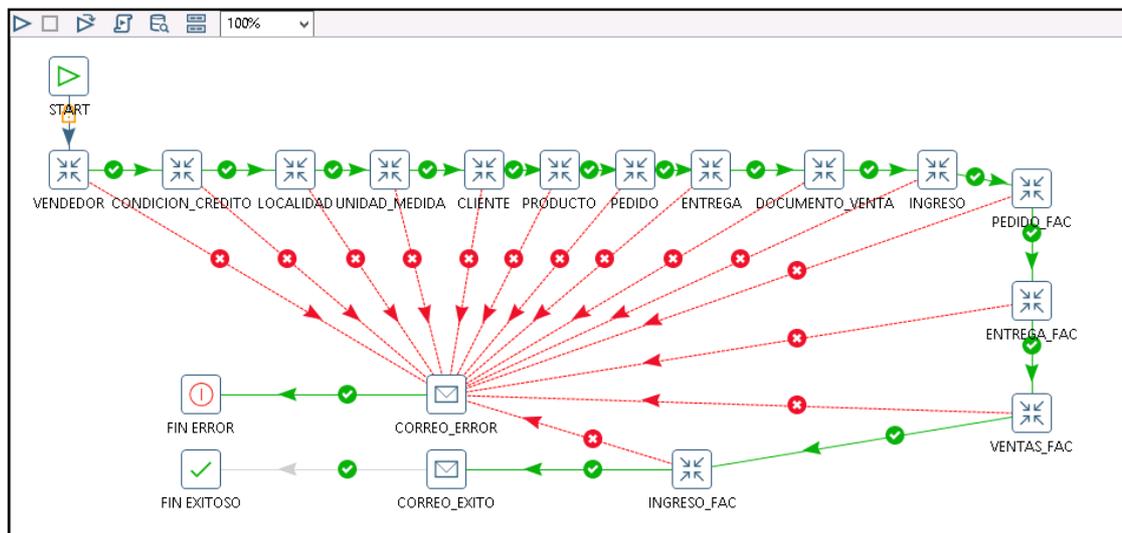


Figura 5.29 Job para la carga periódica de las transformaciones.

Luego de crear las transformaciones para las dimensiones y tablas de hechos que realizarán el proceso ETL necesario para la carga de los datos en el datamart, se procedió a crear en PENTAHO los cubos OLAP definidos en el modelo dimensional en el Capítulo 4.

## **5.2 Creación del cubo OLAP**

En el Capítulo 4 se definió el modelo dimensional que se implementó para satisfacer las necesidades de información de la empresa de manufactura. De acuerdo a ese modelo, se crearon cuatro cubos OLAP: Pedidos, Entregas, Ventas e Ingresos. Además, se crearon dos cubos virtuales para cumplir con los requerimientos definidos. Se creó el cubo virtual VENTAS-INGRESOS que permite comparar los datos del cubo de VENTAS y el cubo de INGRESOS. También, se creó el cubo virtual PEDIDOS-ENTREGAS que permite comparar el cumplimiento de las entregas en relación a los pedidos realizados. Para la creación de los cubos se usó la herramienta Schema Workbench de PENTAHO [14]. Los cubos se encuentran en un solo esquema denominado Esquema Comercial de la Empresa de Manufactura. Dentro de este esquema también se crearon las dimensiones compartidas indicadas en la sección 4.3. En la Figura 5.30 se muestra los cubos OLAP, cubos virtuales y dimensiones compartidas que se crearon según el modelo establecido. Estos cubos serán usados para explotar y presentar la información al usuario final, ya sea en forma de reportes o en vistas de análisis, que están agrupados en un cuadro de mando.

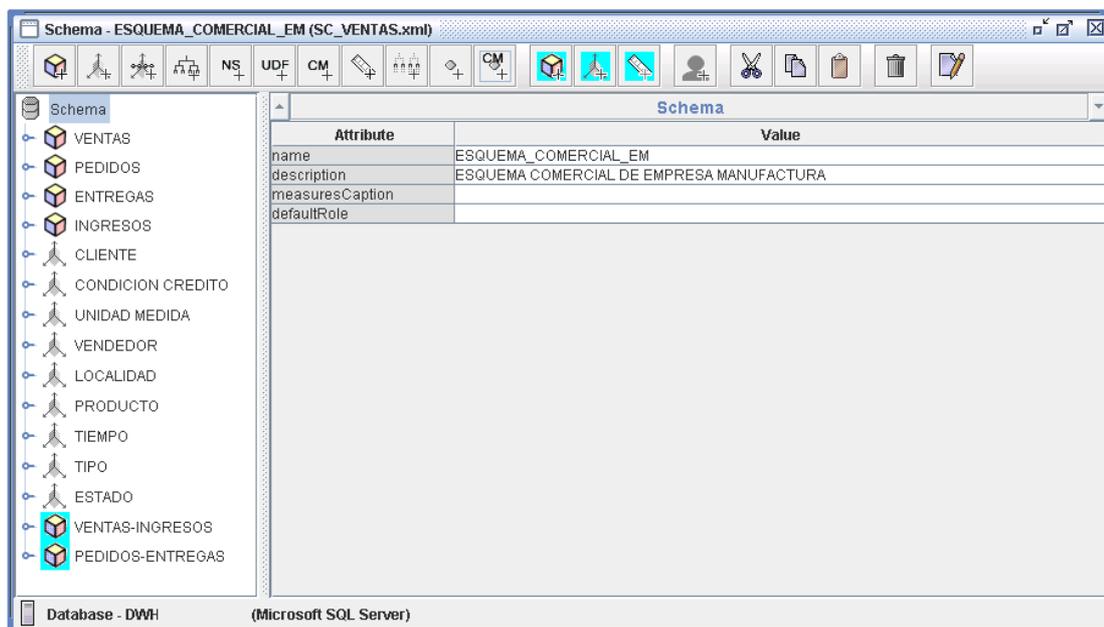


Figura 5.30 Esquema de los Cubos OLAP.

En la siguiente sección se describe la creación del esquema que se usa para agrupar los cubos y dimensiones compartidas.

### 5.2.1 Creación del Esquema Comercial

Como se indicó anteriormente, los cubos OLAP, cubos virtuales y dimensiones compartidas estarán dentro de un solo esquema, al cual se ha denominado ESQUEMA\_COMERCIAL\_EM. El objetivo de este esquema es agrupar los cubos y dimensiones compartidas, para luego ser publicado y poder ser utilizado en el servidor BI de PENTAHO. Dentro de este esquema, se creó la conexión a la base de datos. En este caso sólo se conectará a la base de datos destino debido a que desde este punto, el objetivo será obtener los datos ya

transformados y almacenados en el datamart para su concerniente análisis y presentación en reportes. En la Figura 5.31 se puede observar la conexión realizada.

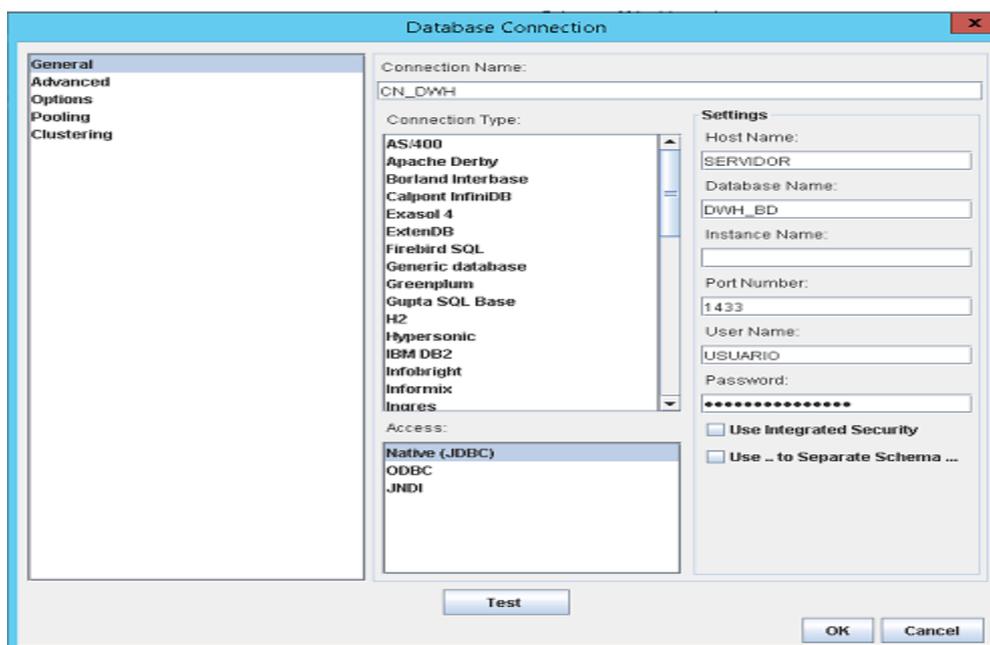


Figura 5.31 Conexión Base de datos del Cubo OLAP.

A continuación se detalla la creación de las dimensiones compartidas que son usadas en los cubos OLAP y cubos virtuales.

## 5.2.2 Creación de las dimensiones compartidas

Una vez que se creó el esquema se procedió a crear las dimensiones compartidas que se definieron en la sección 4.3. A continuación se describe el proceso para la creación de las dimensiones compartidas

del esquema comercial que son usadas en los diferentes cubos OLAP y virtuales.

- **Dimensión Producto:** En la Figura 5.32 se indica cómo se asocia la dimensión con la tabla DM\_PRODUCTO. El tipo de dimensión seleccionado es el STANDARDDIMENSION que indica que no es una dimensión de tiempo. En las dimensiones compartidas no se define el Foreign Key. Esta definición se realiza cuando se asocia la dimensión compartida al cubo OLAP.

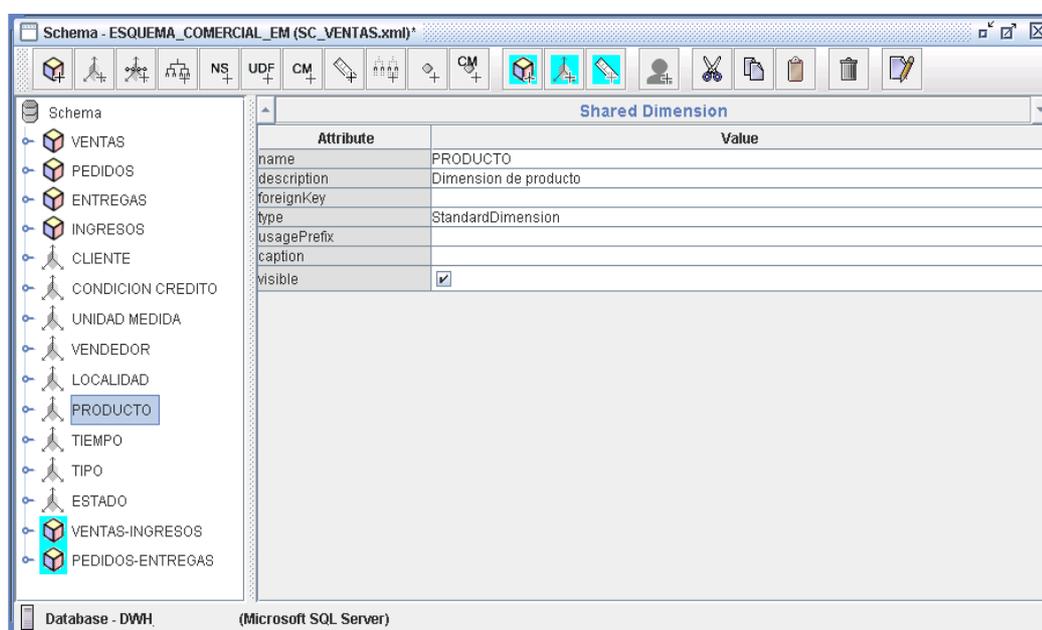


Figura 5.32 Enlace de dimensión producto con Cubo OLAP.

Luego de agregar la dimensión se procede a crear la jerarquía de la dimensión PRODUCTO. Dentro de la jerarquía se definen los niveles de la dimensión. Los niveles de la dimensión de productos son: Grupo, Mercado, Subcategoría y Producto, en el orden dado. En la Figura 5.33 se observan las jerarquías y niveles definidos.

Cada nivel se asocia con su perteneciente campo en la tabla de dimensión.

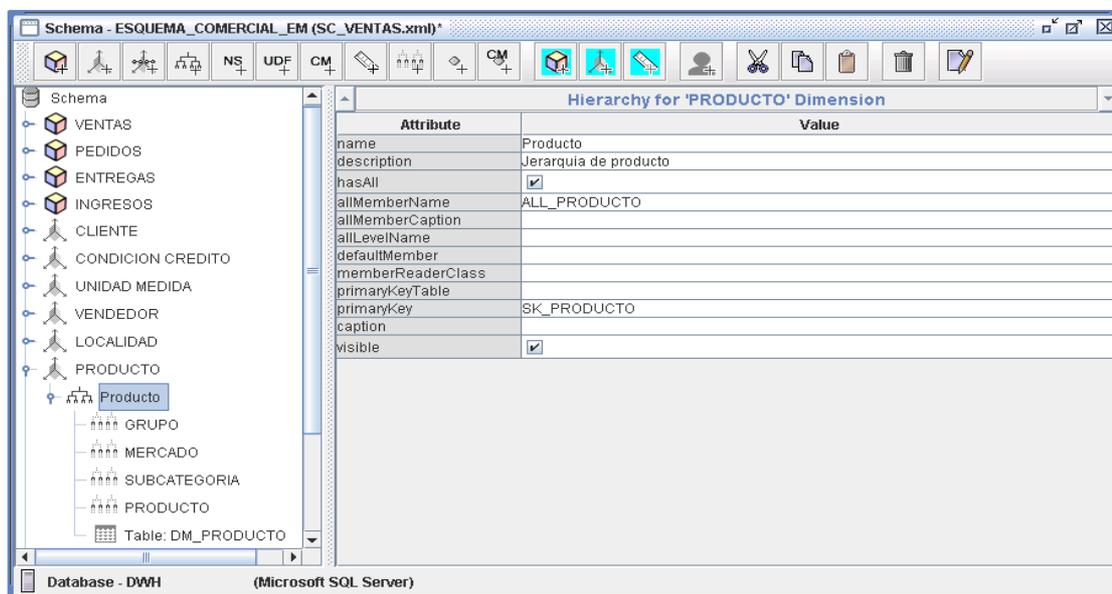


Figura 5.33 Jerarquía y niveles de Productos.

- **Dimensión Cliente:** En la Figura 5.34 se indica cómo se enlaza la dimensión con la tabla DM\_CLIENTE. También esta dimensión es de tipo estándar.

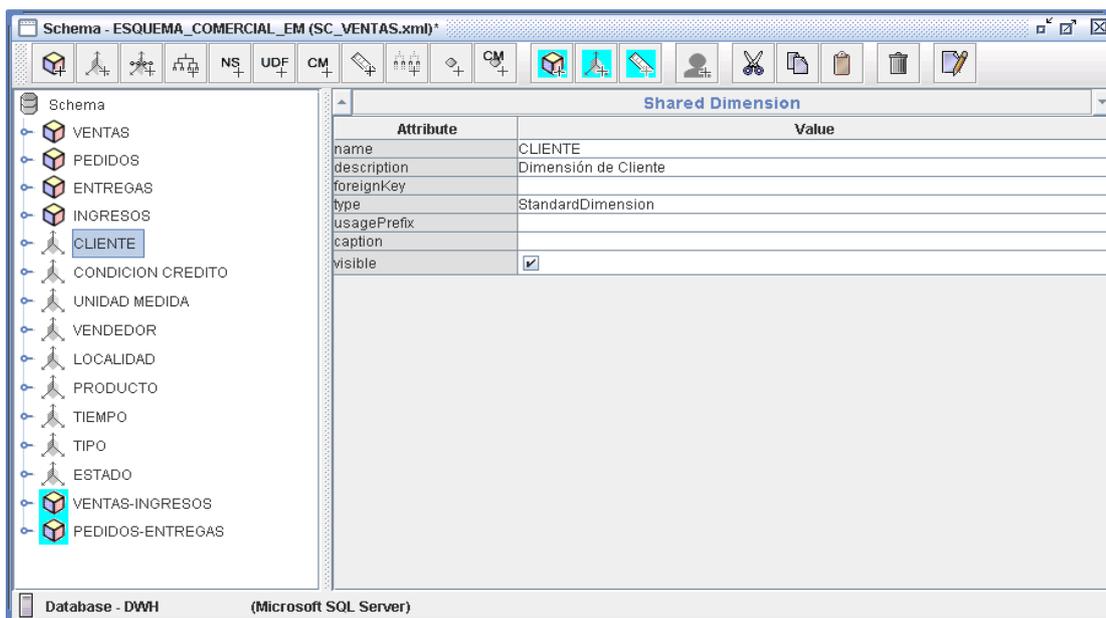


Figura 5.34 Enlace dimensión Cliente con el cubo OLAP.

Se procede a crear la jerarquía para la dimensión CLIENTE. Esta dimensión consta de los siguientes niveles: Mercado, Sector y Cliente, como se ilustra en la Figura 5.35.

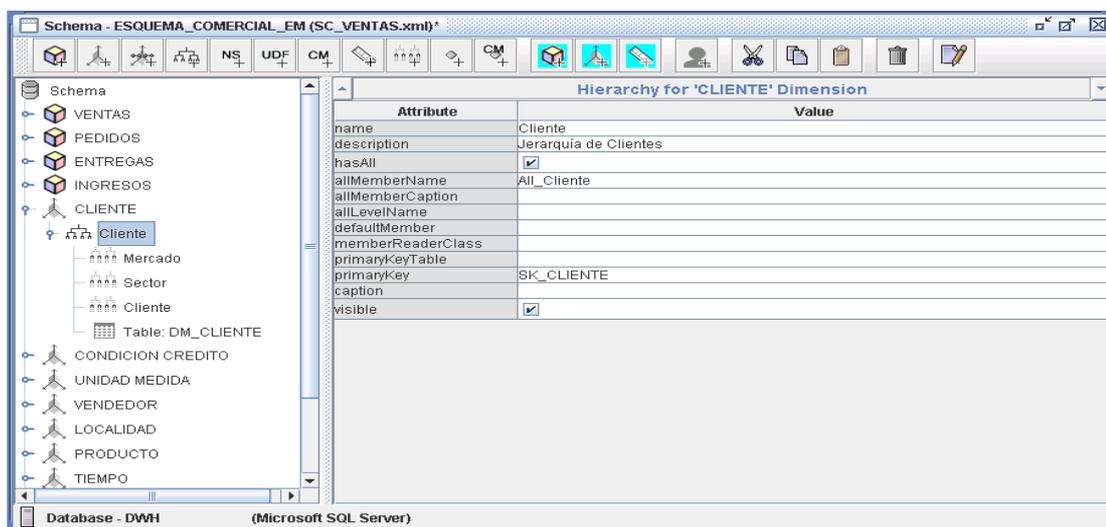


Figura 5.35 Jerarquía y Niveles de la dimensión Cliente.

- **Dimensión Tiempo:** En la Figura 5.36 se puede observar la dimensión Tiempo. La tabla de la dimensión es la DM\_TIEMPO. Este tipo de dimensión se define como TIMEDIMENSION.

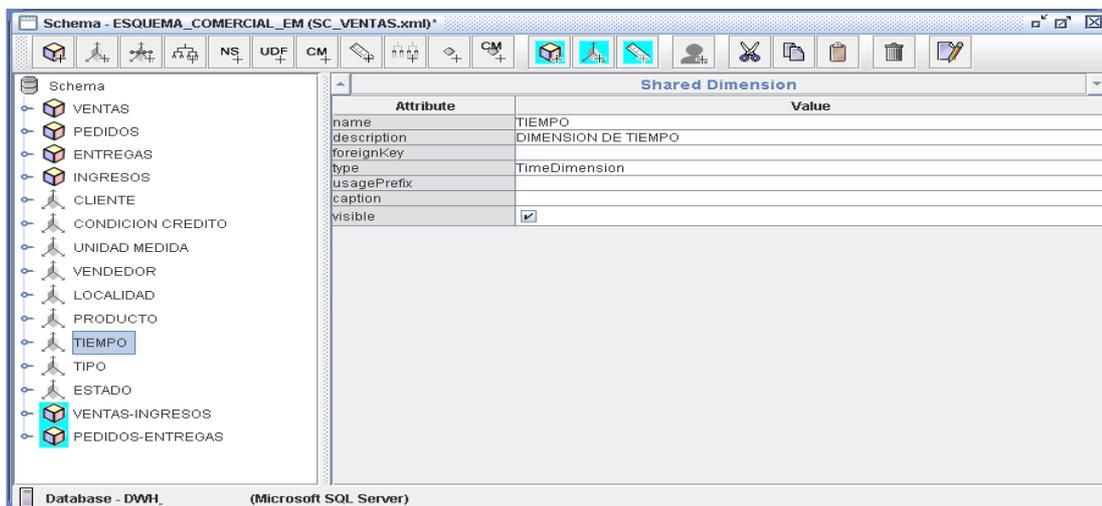


Figura 5.36 Enlace de la dimensión Tiempo con el cubo OLAP.

Se crea la jerarquía para la dimensión tiempo, TIEMPO. Los niveles para esta dimensión son: Año, Mes y Día, de acuerdo a lo mostrado en la Figura 5.37.

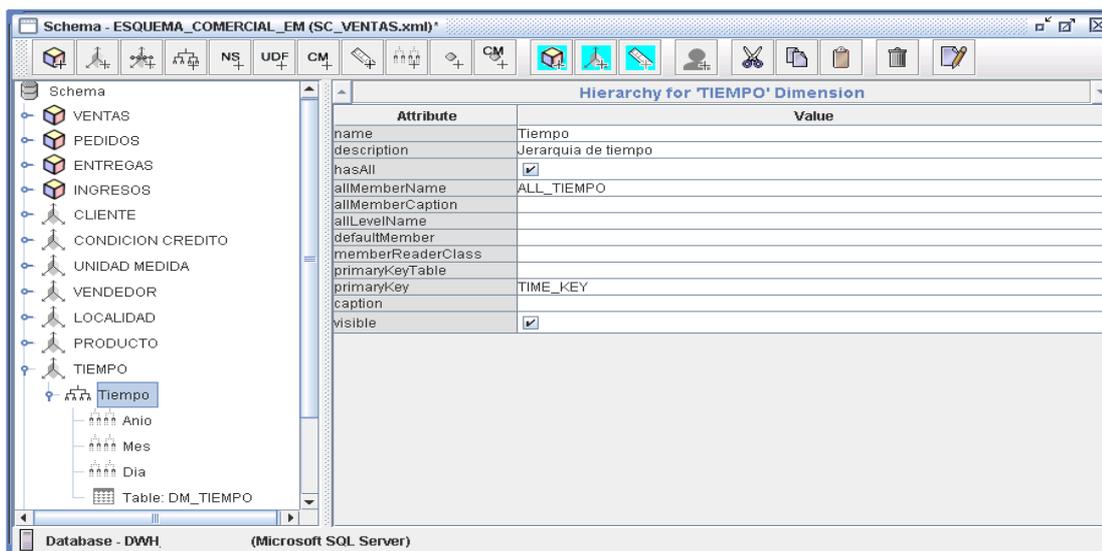


Figura 5.37 Jerarquía y Niveles de la dimensión Tiempo.

- **Dimensión Vendedor:** En la Figura 5.38 se ilustra el enlace entre la dimensión Vendedor con la tabla DM\_VENDEADOR.

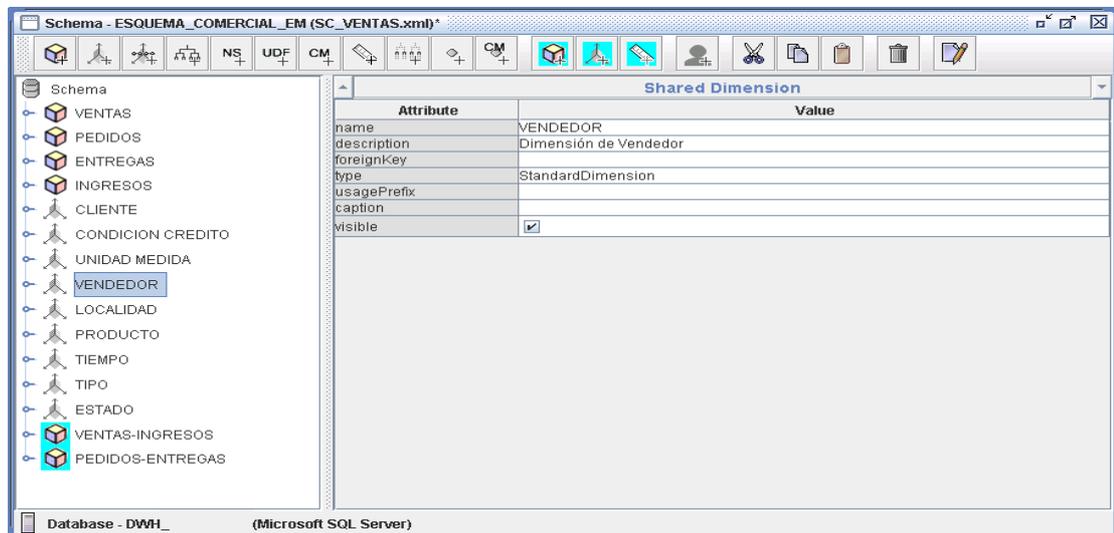


Figura 5.38 Enlace de la dimensión Vendedor con el Cubo OLAP.  
Se procede a crear la jerarquía para la dimensión VENDEADOR.  
Esta dimensión no tiene niveles superiores al vendedor, por lo que únicamente se define ese nivel según lo mostrado en la Figura 5.39.

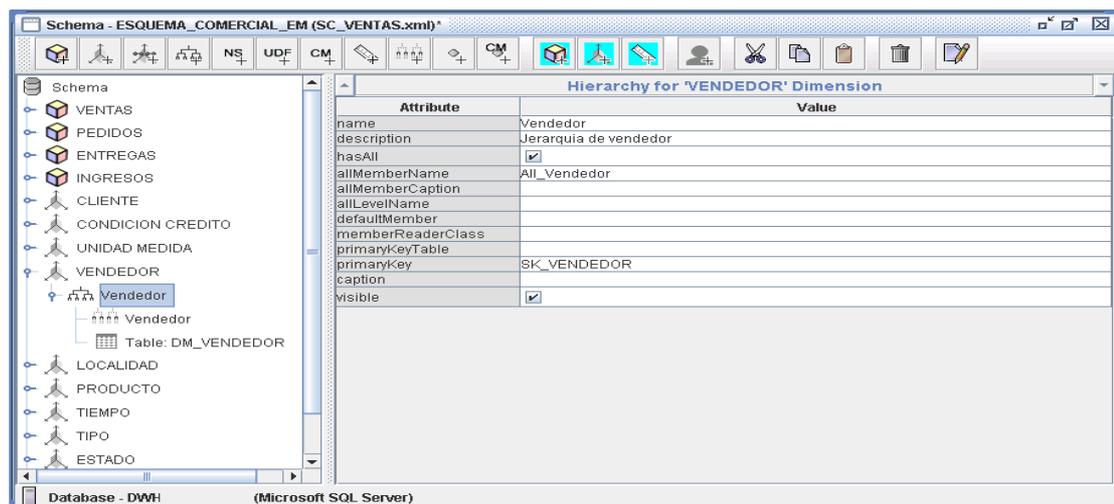


Figura 5.39 Jerarquía y Niveles de la dimensión Vendedor.

- **Dimensión Localidad:** En la Figura 5.40 se ilustra la asociación de la dimensión de localidades con la tabla de dimensión DM\_LOCALIDAD.

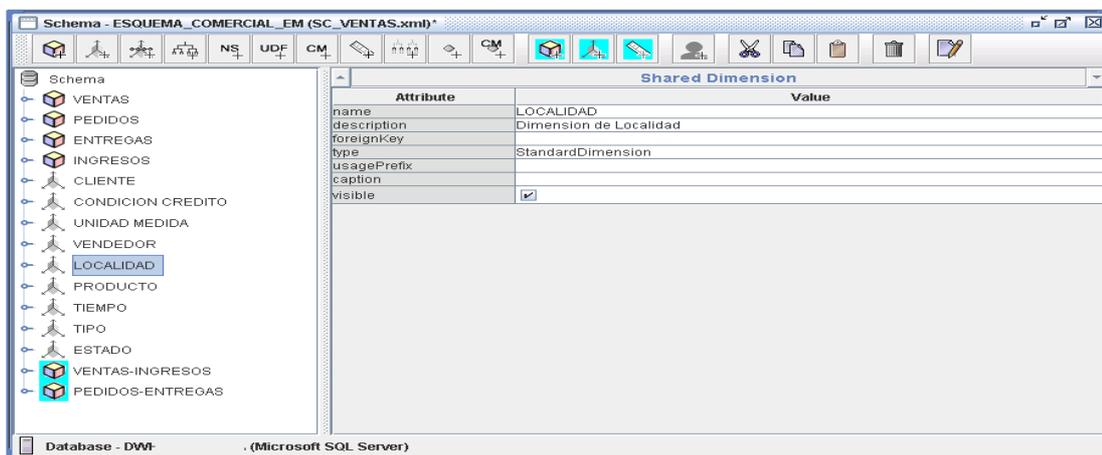


Figura 5.40 Enlace de la dimensión Localidad con el cubo OLAP.

Se continúa con la creación de la jerarquía de la dimensión LOCALIDAD. Los niveles de esta dimensión son País y Provincia, como se puede apreciar en la Figura 5.41.

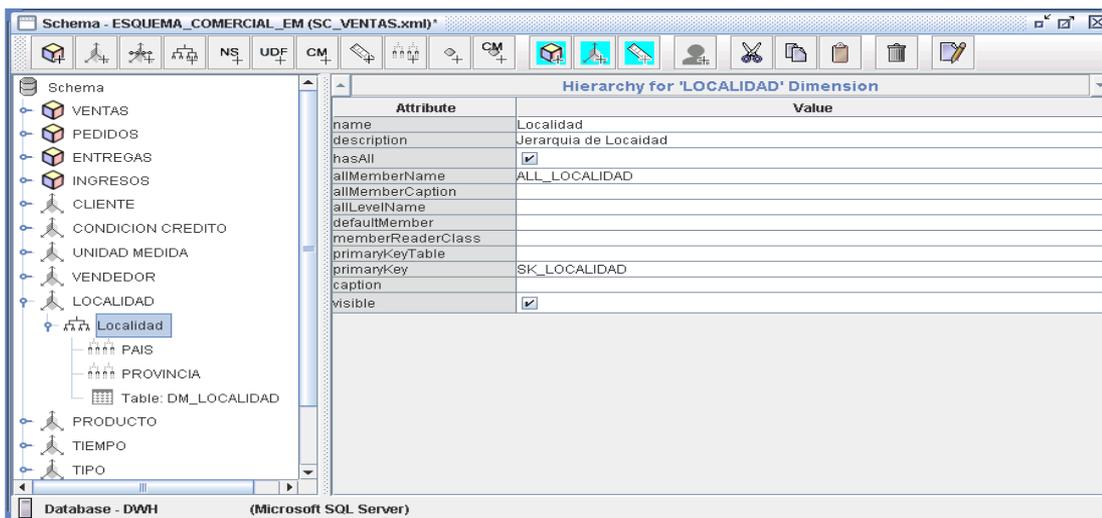


Figura 5.41 Jerarquía y Niveles de la dimensión Localidad.

- **Dimensión Unidad Medida:** En la Figura 5.42 se observa el enlace de la dimensión Unidad de Medida con la tabla DM\_UNIDAD\_MEDIDA.

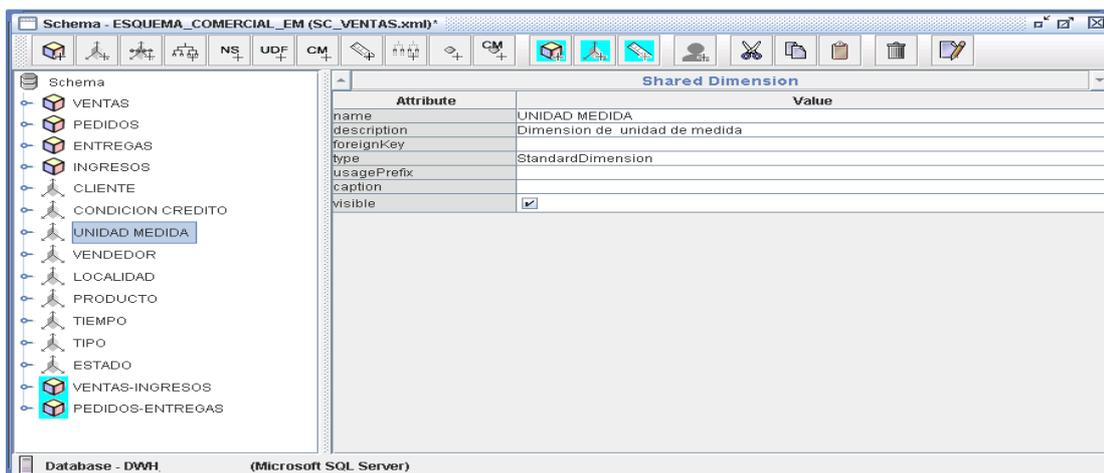


Figura 5.42 Enlace de la dimensión Unidad de Medida con el cubo OLAP.

La jerarquía para la dimensión de unidad de medida es la UNIDAD\_MED. Al igual que la dimensión de vendedores, esta dimensión sólo tiene un nivel que es la propia unidad de medida.

En la Figura 5.43 se observa lo indicado.

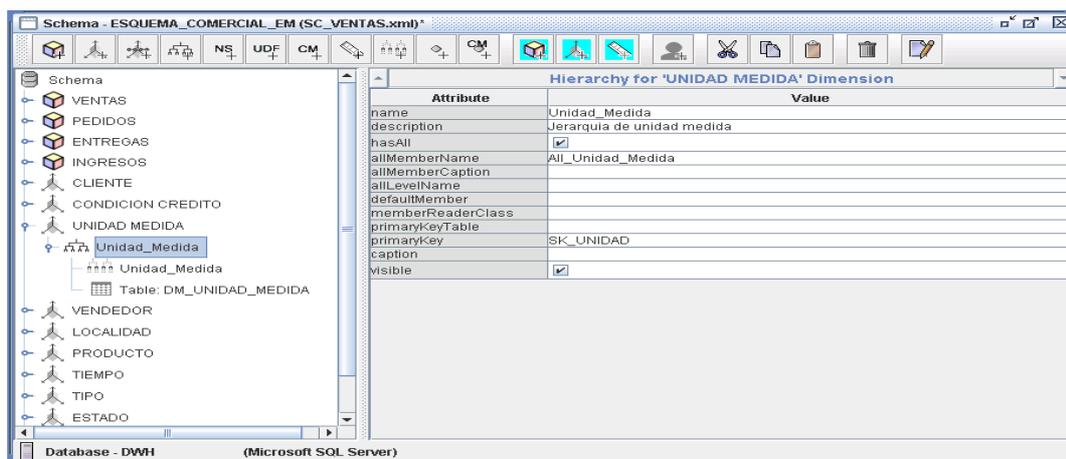


Figura 5.43 Jerarquía y Niveles de la dimensión Unidad de Medida.

- **Dimensión Condición Crédito:** En la Figura 5.44 se muestra la integración de la dimensión condición de crédito con la tabla DM\_CONDICION\_CREDITO.

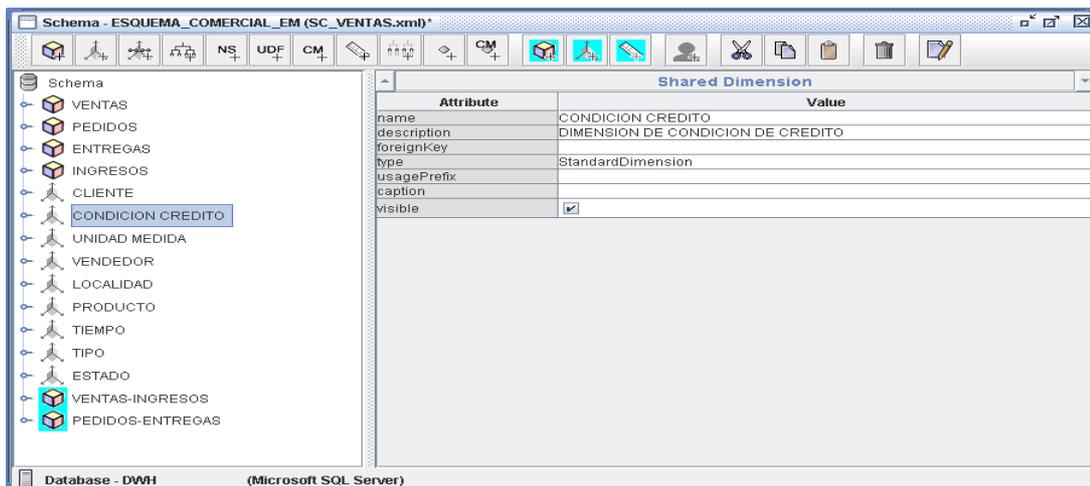


Figura 5.44 Enlace de la dimensión Condición Crédito con el cubo OLAP.

La jerarquía para esta dimensión se llama CONDICION\_CREDITO. El único nivel que utiliza esta jerarquía es la condición de crédito como se aprecia en la Figura 5.45.

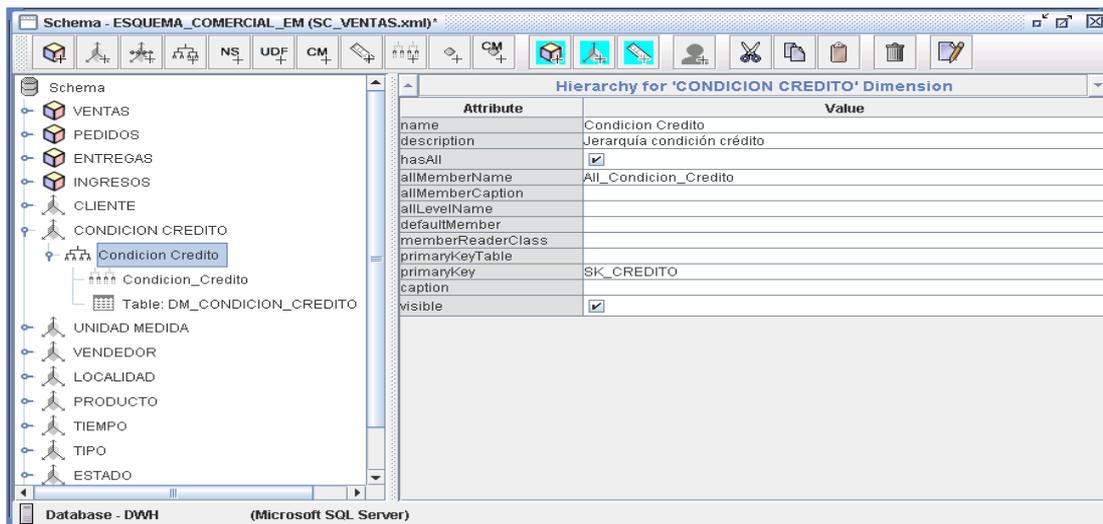


Figura 5.45 Jerarquía y Niveles de la dimensión Condición Crédito.

- **Dimensión Tipo:** En la Figura 5.46 se muestra la integración de la dimensión tipo con la tabla DM\_TIPO.

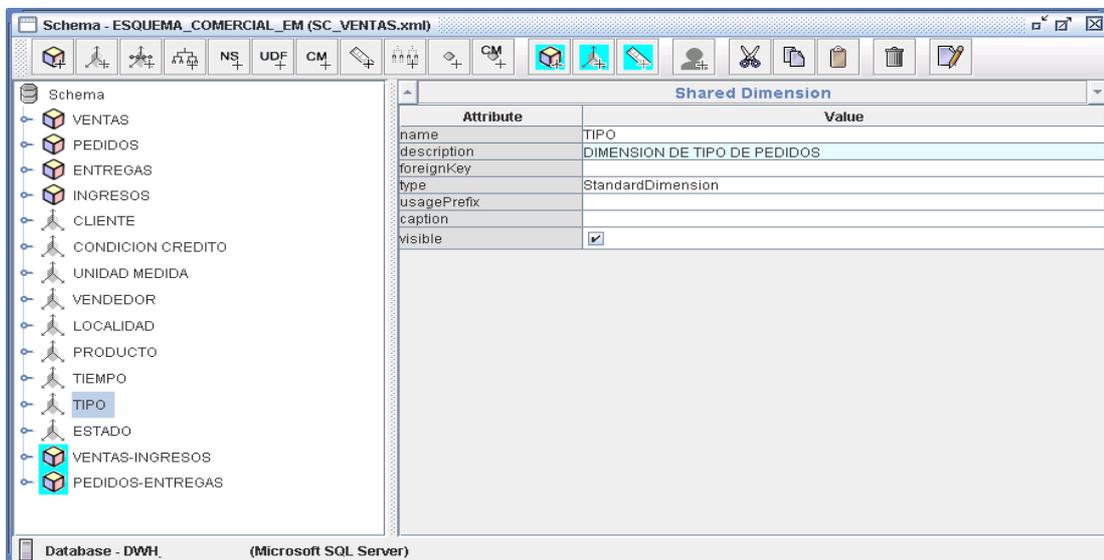


Figura 5.46 Enlace de la dimensión Tipo con el cubo OLAP.

La jerarquía para esta dimensión se llama TIPO. El único nivel que utiliza esta jerarquía es tipo como se muestra en la Figura 5.47.

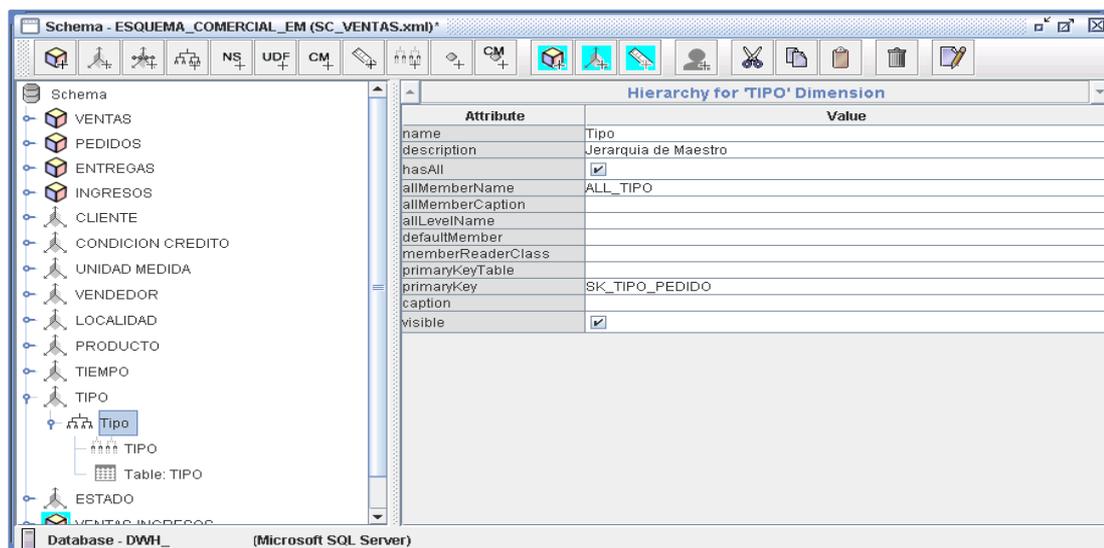


Figura 5.47 Jerarquía y Niveles de la dimensión Tipo.

- **Dimensión Estado:** En la Figura 5.48 se muestra la integración de la dimensión estado con la tabla DM\_ESTADO.

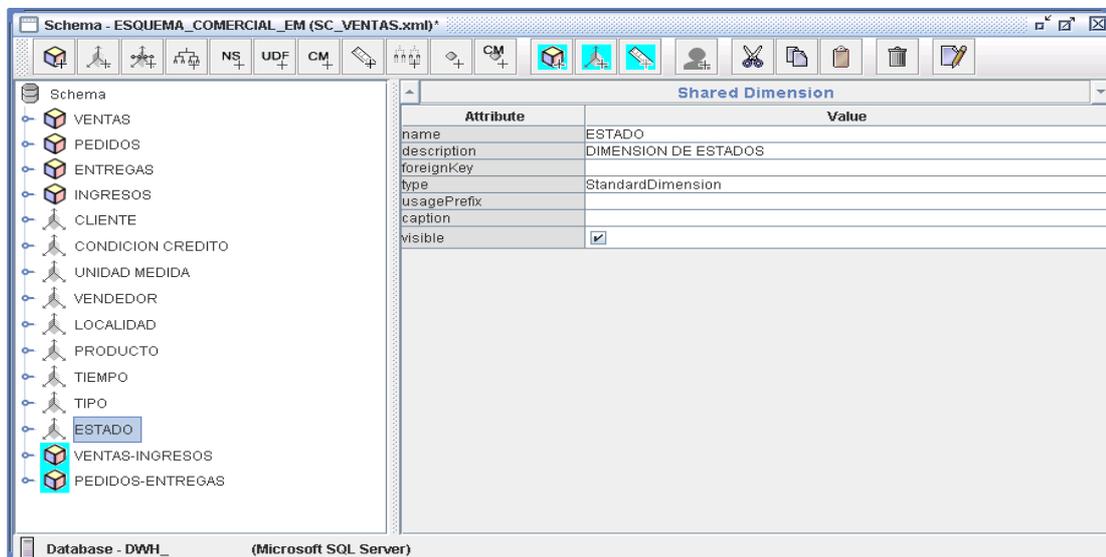


Figura 5.48 Enlace de la dimensión Estado con el cubo OLAP.

La jerarquía para esta dimensión se llama ESTADO. El único nivel que utiliza esta jerarquía es estado como se muestra en la Figura 5.49

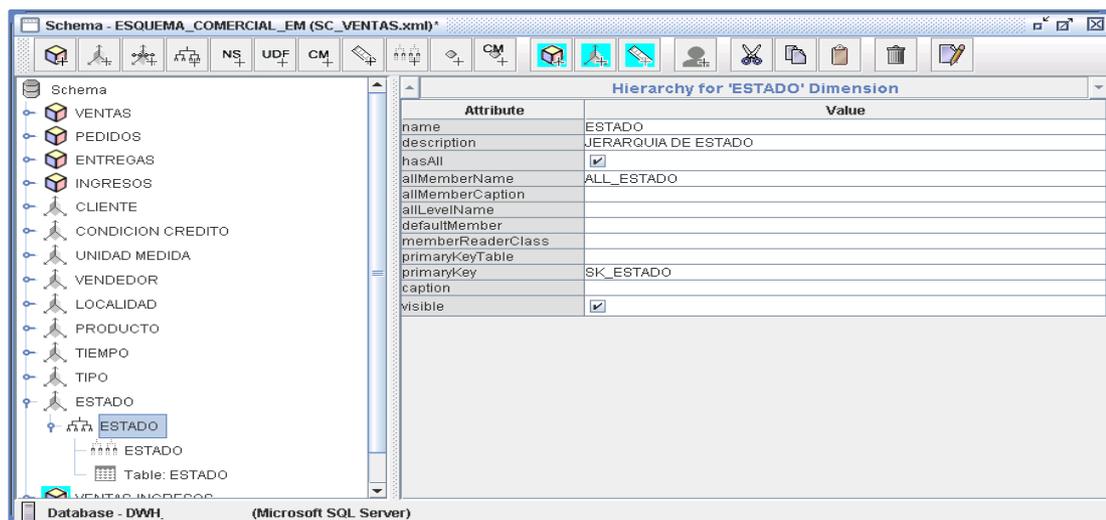


Figura 5.49 Jerarquía y Niveles de la dimensión Estado.

Luego de crear las dimensiones compartidas se procede a crear los cubos OLAP y virtuales. Debido a que el proceso de creación del cubo OLAP es el mismo para todos los cubos definidos, sólo se detallará la creación del cubo de Ventas el cual, como se ha mencionado anteriormente, representa el mayor interés para el área comercial. A continuación se describen los pasos que se realizaron para la creación del cubo OLAP de Ventas.

### **5.2.3 Creación de cubo OLAP de Ventas**

En la creación del cubo OLAP de Ventas se asoció la tabla de hechos con las respectivas dimensiones del modelo: Cliente, Condición de Crédito, Unidad de Medida, Vendedor, Localidad, Producto, Documento, Tipo, Estado y Tiempo. Además, se crearon las medidas que servirán para medir el proceso de ventas: Cantidad, Monto Bruto, Monto Neto y Valor IVA. Asimismo se crearon los campos calculados como son Toneladas Vendidas. En la Figura 5.50 se ilustra la integración de la tabla de hechos con sus dimensiones y medidas.

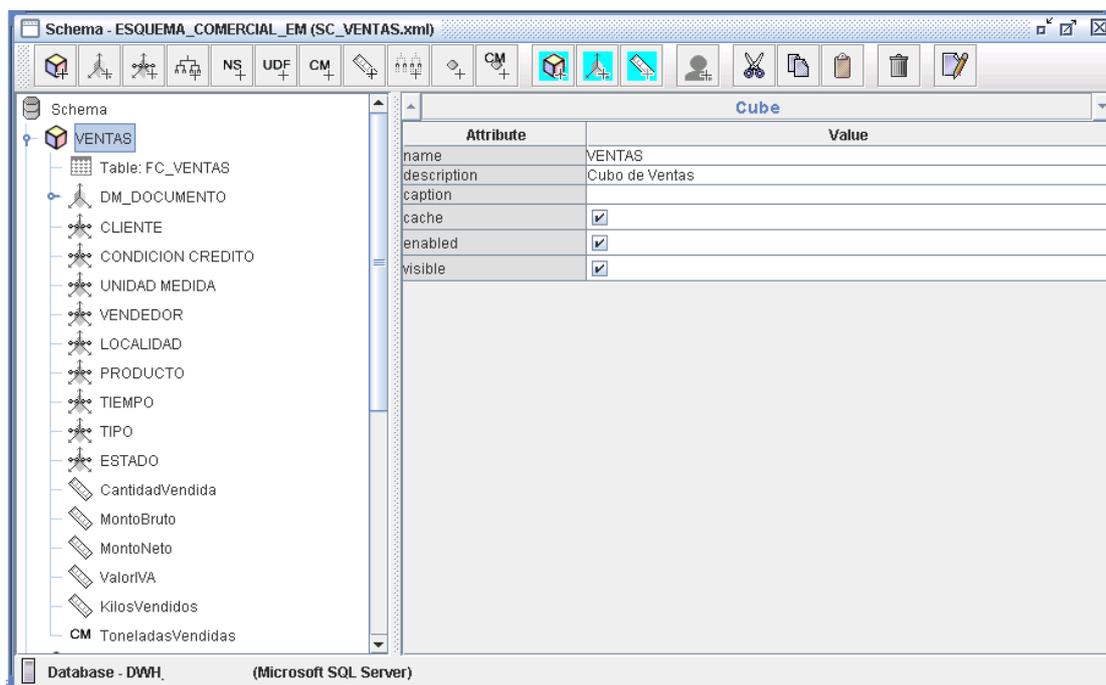


Figura 5.50 Cubo OLAP de Ventas.

A continuación se detallan los pasos que se siguieron para la creación del cubo en Schema Workbench:

**Paso 1: Asociación de la tabla de hechos con el cubo OLAP.** Se asocia el cubo a la tabla de hechos correspondiente, en este caso a la tabla de hechos de Ventas. En la Figura 5.51 se muestra la asociación realizada. Como se observa en la Figura 5.51 se agrega la tabla FC\_VENTAS al cubo. Es importante aclarar que un cubo sólo puede tener una tabla de hechos asociada.

Attribute	Value
schema	dbo
name	FC_VENTAS
alias	FC_VENTAS

Figura 5.51 Cubo OLAP Ventas – Tabla de Hechos.

**Paso 2: Agregar las dimensiones propias del cubo.** El siguiente paso consiste en agregar las dimensiones propias del modelo. El cubo OLAP de Ventas sólo tiene una dimensión propia que se indica a continuación.

- **Dimensión Documento:** La dimensión Documento DM\_DOCUMENTO se enlaza con el cubo OLAP a través del uso de la clave SK\_DOCUMENTO. En la Figura 5.52 se muestra esta integración.

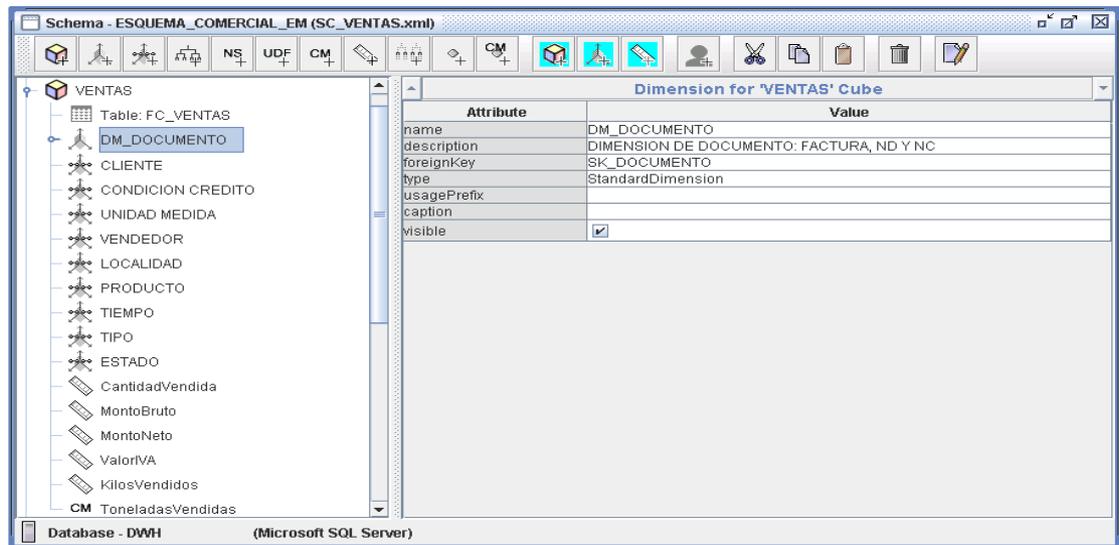


Figura 5.52 Enlace de la dimensión Documento con el cubo OLAP.

La jerarquía para esta dimensión es DOCUMENTO. Esta dimensión que proporciona los datos de la cabecera de las transacciones de venta, cuenta con tres niveles que son: Tipo de Pedido, Estado del Documento y el Número de Documento de la transacción. En la Figura 5.53 se puede observar lo indicado.

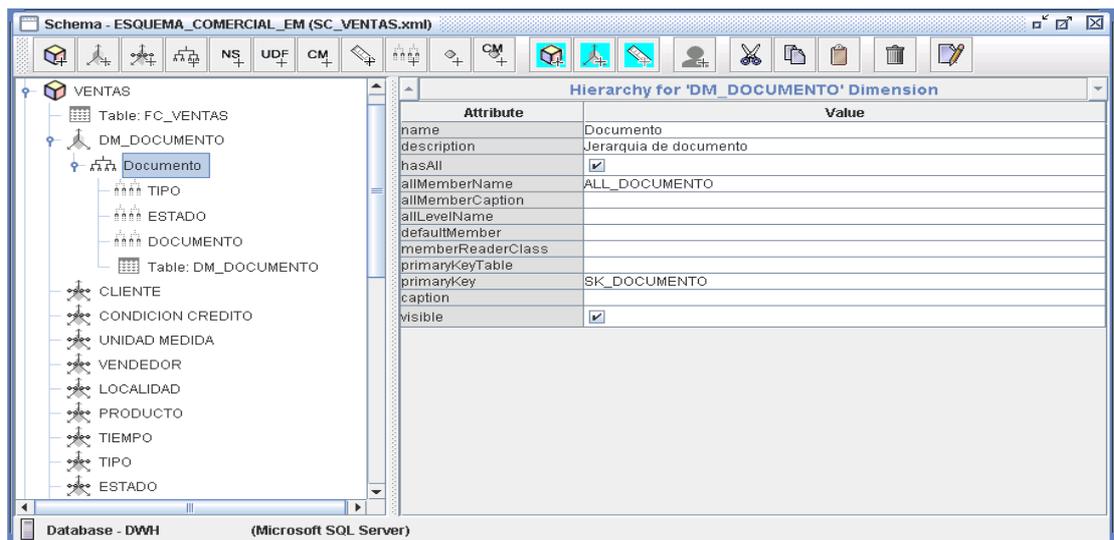


Figura 5.53 Jerarquía y Niveles de la dimensión Documento.

**Paso 3: Agregar las dimensiones compartidas.** El siguiente paso consiste en agregar las dimensiones compartidas. Para ello se usan dimensiones de USO. En la Tabla 11 se muestra el enlace de la dimensión compartida con el cubo OLAP, mediante el uso de la clave foránea de la dimensión con la tabla de hechos definida en el paso 1.

Tabla 11. Asociación de las dimensiones compartidas con el cubo OLAP

DIMENSIÓN COMPARTIDA	CLAVE FORÁNEA PARA ENLACE CON EL CUBO OLAP
Producto	SK_PRODUCTO
Cliente	SK_CLIENTE
Tiempo	SK_TIEMPO
Vendedor	SK_VENDEDOR
Unidad Medida	SK_LOCALIDAD
Condición Crédito	SK_CONDICION_CREDITO
Localidad	SK_LOCALIDAD
Tipo	SK_TIPO
Estado	SK_ESTADO

**Paso 4: Agregar las medidas.** Luego se procede a agregar las medidas que se utilizarán para medir el proceso. Las medidas son: CANTIDAD VENDIDA, MONTO BRUTO, MONTO NETO, VALOR IVA y KILOS VENDIDOS.

En la Figura 5.54 se observa la configuración para la medida CANTIDAD VENDIDA. Esta medida se asocia con el campo Cantidad de la tabla de Hechos. El tipo de función utilizado es la suma.

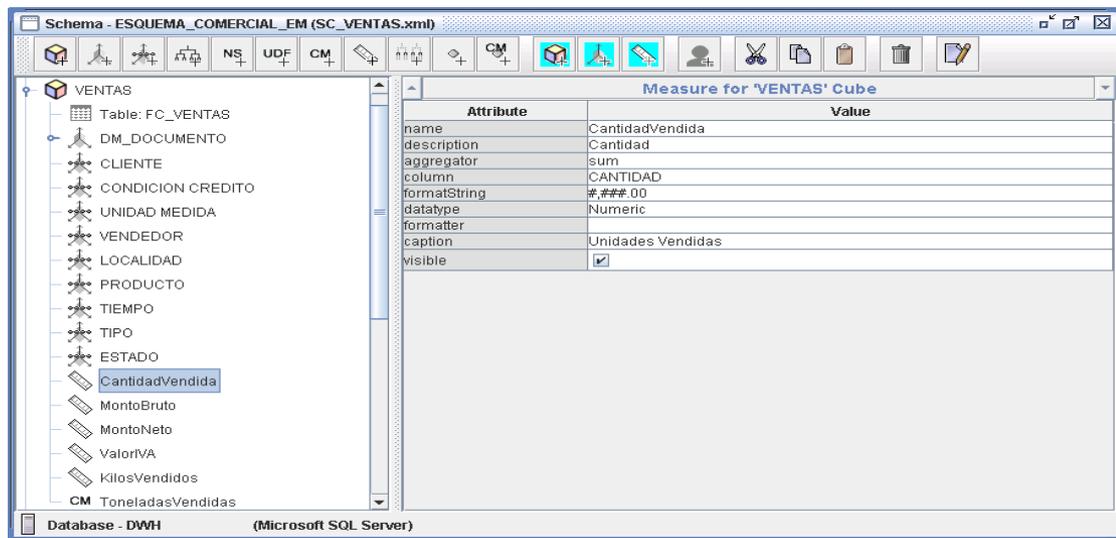


Figura 5.54 Medida Cantidad Vendida.

En la Figura 5.55 se muestra la configuración de la medida MONTO BRUTO. Esta medida se enlaza con el campo Monto Bruto de la tabla de hechos. La función de agregación es la suma.

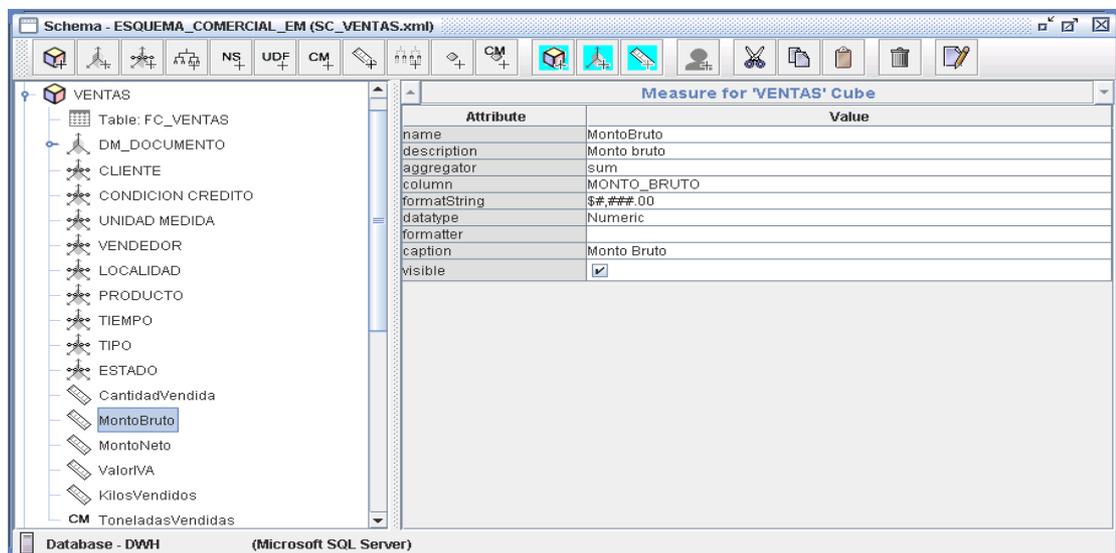


Figura 5.55 Medida Monto Bruto.

En la Figura 5.57 se muestra la configuración de la medida MONTO NETO que se asocia con el campo Monto Neto de la tabla de hechos. La función de agregación es la suma.

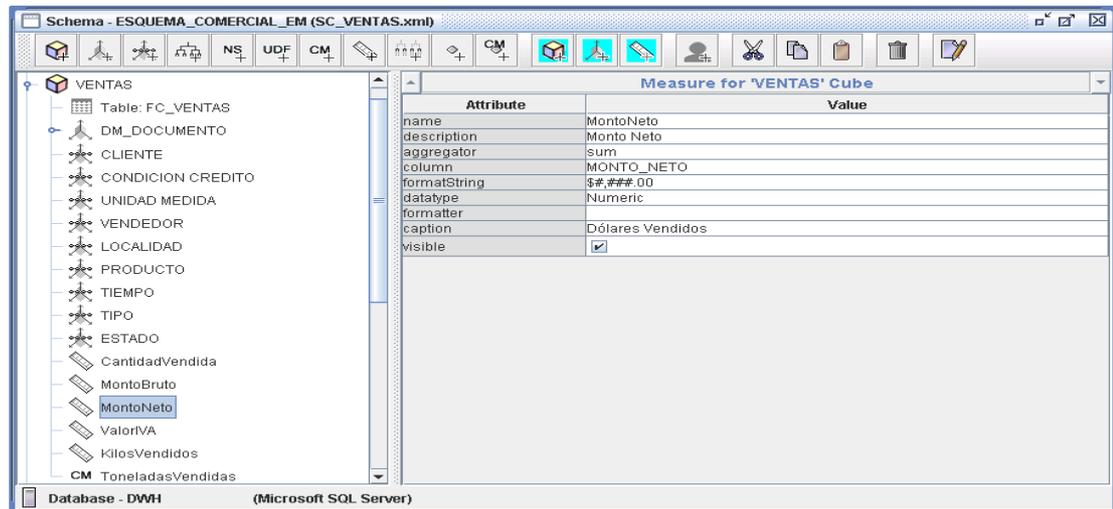


Figura 5.56 Medida Monto Neto.

En la Figura 5.58 se muestra la configuración de la medida VALOR IVA que se asocia con el campo Valor IVA de la tabla de hechos. La función de agregación es la suma.

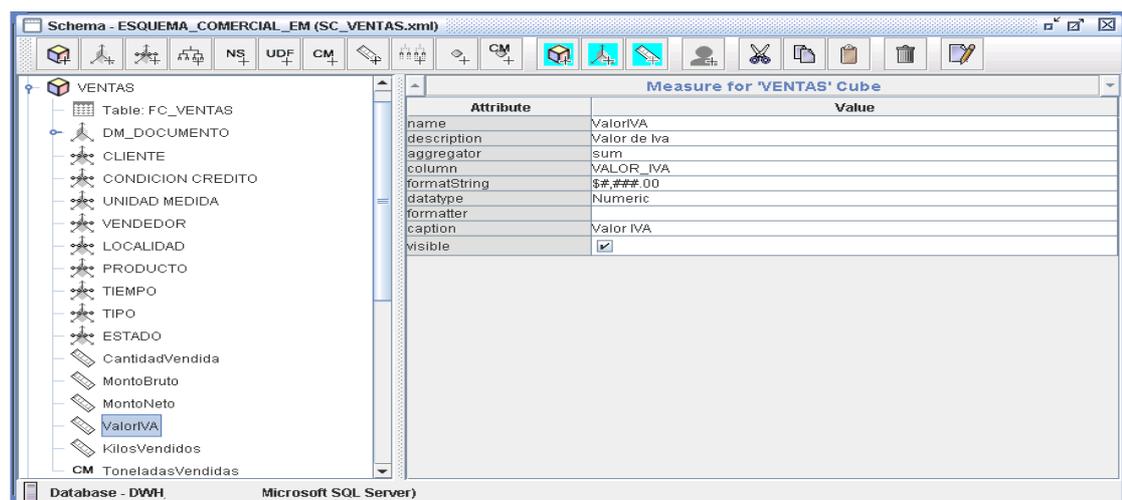


Figura 5.57 Medida Valor IVA.

En la Figura 5.58 se muestra la configuración de la medida KILOS VENDIDOS que se asocia con el campo Peso de la tabla de hechos. La función de agregación es la suma.

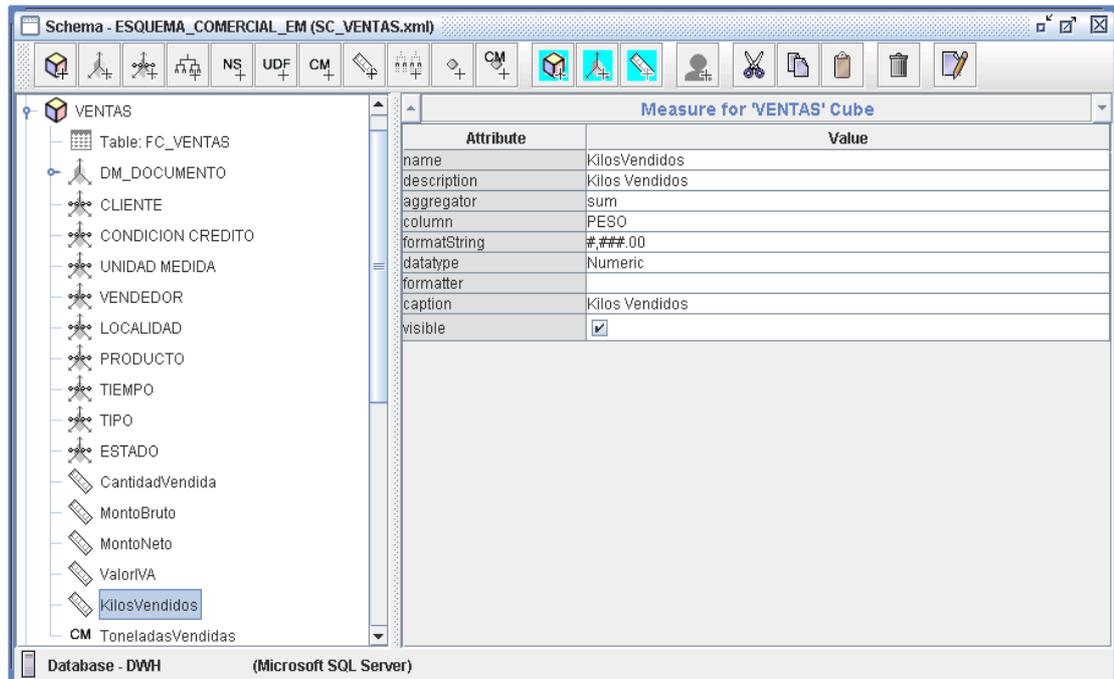


Figura 5.58 Medida Kilos Vendidos.

**Paso 5: Crear campos calculados.** Por último, se procede a crear los campos calculados que se usarán el cubo. Para la empresa de manufactura es de gran importancia conocer las ventas en función de toneladas vendidas por lo que se procedió a crear este campo calculado con la medida PESO aplicando la fórmula PESO/1000. En la figura 5.59 se muestra la creación de este campo calculado.

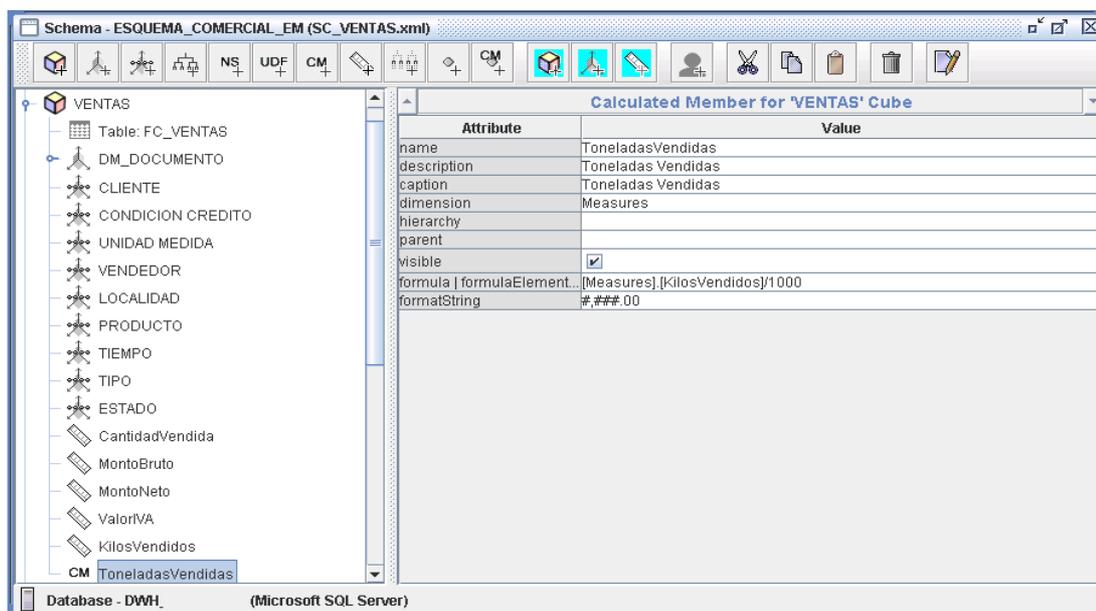


Figura 5.59 Campo Calculado Toneladas Vendidas.

Después de la creación de los cubos OLAP se procedió a crear los cubos virtuales para satisfacer los requerimientos establecidos. De la misma forma que con la creación del cubo OLAP, el proceso de creación de cubos virtuales es el mismo, por lo que sólo se detallará la creación del cubo virtual VENTAS-INGRESOS.

#### 5.2.4 Creación de cubo virtual de Ventas e Ingresos

Este cubo virtual se utilizará para comparar las ventas realizadas contra los ingresos realizados a bodega provenientes de la producción. Los cubos OLAP que intervienen son el cubo de Ventas y el cubo de Ingresos. Para la creación del cubo virtual de VENTAS-INGRESOS se asociaron las tablas de dimensiones compartidas: Producto y Estado. También se agregó la dimensión del cubo de

ventas DOCUMENTO, con el objetivo de filtrar sólo los documentos de tipo FACTURA para este cubo virtual, ya que lo que interesa es comparar las ventas y no las devoluciones (Notas de crédito). Por último, se agregaron las medidas y campos calculados de los dos cubos que intervienen. Las medidas que se utilizaron fueron unidades vendidas y producidas; y kilos vendidos y producidos. Los campos calculados que se usaron fueron toneladas vendidas y producidas. En la Figura 5.60 se muestra el cubo virtual de Ventas e Ingresos.

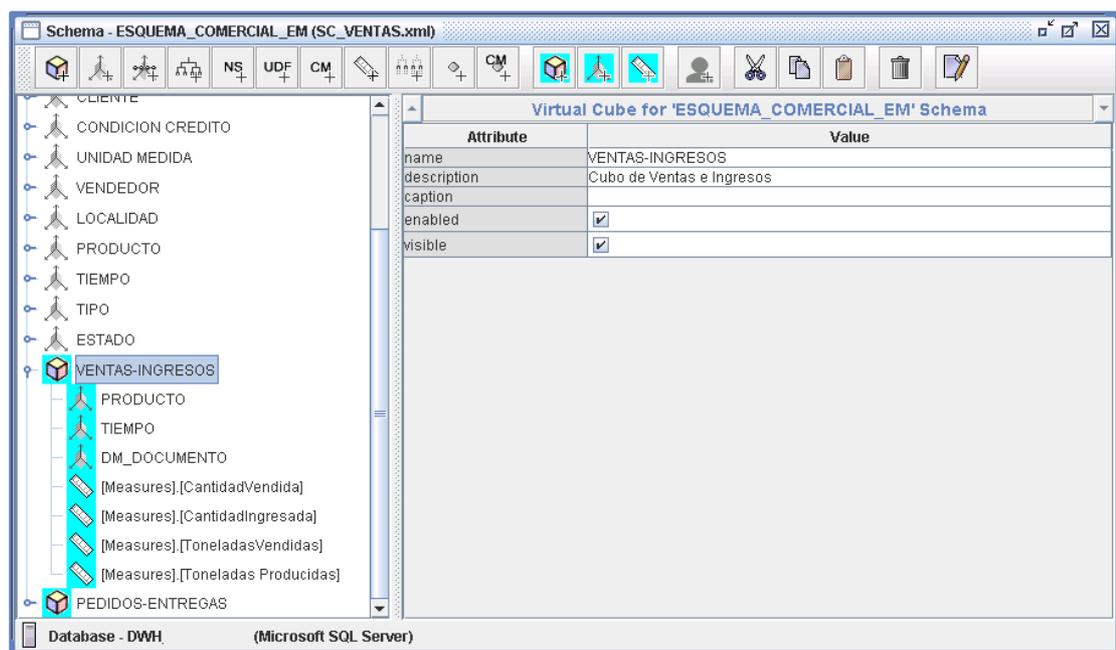


Figura 5.60 Cubo Virtual VENTAS-INGRESOS.

Una vez que se concluyó con la creación de los cubos se procedió a publicar el esquema para que pueda ser utilizado desde el servidor BI de PENTAHO.

En la siguiente sección se procedió a explotar los datos de los cubos que se crearon en el esquema comercial. Para ello se crearon vistas de análisis, reportes estáticos y cuadros de mandos que permitirán satisfacer las necesidades de los usuarios que fueron planteadas en el Capítulo 3.

### **5.3 Explotación de datos**

Para la explotación de los datos se crearon vistas de análisis, cuadros de mandos e indicadores que estarán disponibles para los usuarios de acuerdo al perfil que manejen. Las vistas y cuadros de mando se crean en la Consola de Usuario de Pentaho (PUC, por las siglas en inglés de Pentaho User Console).

A continuación se describen las vistas de análisis que se crearon con la finalidad de cumplir con los requerimientos establecidos en el alcance de este trabajo.

#### **5.3.1 Creación de vistas de análisis**

Las vistas de análisis permiten hacer drill-down de los datos. Esto facilita el análisis de los informes ya que se puede ir desde la información general a la particular. Para la creación de las vistas de análisis se utilizó el componente de PENTAHO llamado JPIVOT

VIEW [15]. Como ya se indicó las vistas de análisis se crean en la PUC. Esta herramienta permite exportar los informes a Excel.

### **5.3.1.1 Vistas de análisis de las Ventas**

Para satisfacer la necesidad de contar con información de las ventas, de acuerdo al requerimiento definido en la Tabla 3.1, se crearon vistas de análisis que están basadas en el cubo OLAP de ventas. A continuación, se describen las vistas que satisfacen este requerimiento:

- **Ventas por Año:** Esta vista permite analizar la información de las ventas por Clientes y Productos en diferentes años. El análisis se puede realizar en términos de dólares y toneladas vendidas. Los datos se muestran de manera resumida por tipo de cliente y grupo de productos. Pero el usuario puede ver los datos de acuerdo al nivel de jerarquía de las dimensiones de cliente, producto y tiempo que se definieron en el cubo de Ventas. En la Figura 5.61 se muestra la vista de ventas por año.

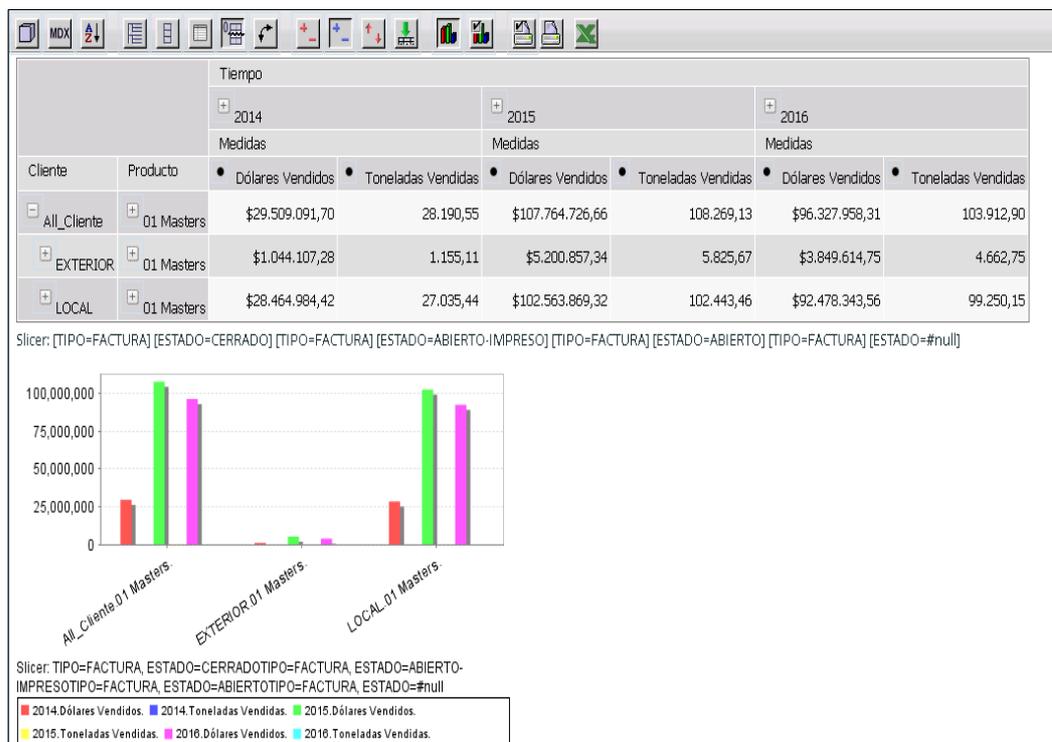


Figura 5.61 Vista de Análisis Ventas por Año.

- Ventas por Cliente:** Esta vista muestra la información de toneladas vendidas por Cliente en un año específico. El usuario puede seleccionar el año que desee ver los datos. En la Figura 5.62 se ilustra las ventas por cliente del año 2016 por grupo y mercado de productos.

Producto								
Cliente	01 Masters	TRANSFERENCIA DE DOMINIO EXPORT	SERVICIO DE CONVERSION	BANANEROS	OTRAS FRUTAS	OTRAS EXPORTACIONES	FLORES OTROS	DOMESTICO
All_Cliente	103.912,90	1.037,75	4,48	67.744,01	1.380,38	7.178,63	2.519,33	19.579,39
EXTERIOR	4.662,75			3.955,53	78,69	25,86		
BANANERO-DOLE	3.971,62			3.955,53		16,09		
EXPORT DIRECTA-VENEZ	227,94							
PESCA-CAMARON	126,98					9,77		
NA CON GIRO DE NEGOC	336,20				78,69			
LOCAL	99.250,15	1.037,75	4,48	63.788,48	1.301,69	7.152,76	2.519,33	19.579,39

Slicer: [TIPO=FACTURA] [ESTADO=CERRADO] [Medida=Toneladas Vendidas] [Año=2016] [TIPO=FACTURA] [ESTADO=ABIERTO-IMPRESO] [Medida=Toneladas Vendidas] [Año=2016] [TIPO=FACTURA] [ESTADO=CERRADO] [Medida=Toneladas Vendidas] [Año=2016]

Figura 5.62 Vista de Análisis de Ventas por Cliente.

- Ventas por Localidad:** Esta vista permite analizar las ventas, en toneladas, de acuerdo a las distintas localidades nacionales y extranjeras donde se distribuye los productos de la empresa. En la Figura 5.63 se muestra la vista de las ventas en toneladas por localidad en el año 2016. El usuario puede filtrar el periodo que necesite analizar.

Localidad	Producto							
	01 Masters	TRANSFERENCIA DE DOMINIO EXPORT	SERVICIO DE CONVERSION	BANANEROS	OTRAS FRUTAS	OTRAS EXPORTACIONES	FLORES OTROS	DOMESTICO
ALL_LOCALIDAD	103.912,90	1.037,75	4,48	67.744,01	1.380,38	7.178,63	2.519,33	19.579,39
Ecuador	18.329,74		4,48	11.900,80	1.061,60	790,42	110,82	4.073,40
GUAYAS	6.934,77			2.038,03	784,63	495,23		3.238,93
MANTA	32,73					,71		32,02
SANTA ELENA	45,65					6,61		39,04
PICHINCHA	614,18		4,48		3,26	5,46	101,86	489,27
MANABI	326,32					247,42		78,90
AZUAY	120,26						2,92	117,34
MACHALA	79,22			73,36				5,86
BABAHOYO	224,02			223,24				,78
LOS RIOS	3.699,87			3.369,85	272,48	27,92		29,62
EL ORO	6.153,77			6.149,88		3,89		
TUNGURAHUA	4,00							4,00
ESMERALDAS	11,20					,72		10,48
SANTO DOMINGO DE LOS TSACHILAS	49,43			33,82	1,23	2,47		11,50
CARCHI	12,61			12,61				
CHIMBORAZO	8,90							8,90
CAÑAR	6,05						6,05	
LOJA	6,75							6,75
#null	85.583,16	1.037,75		55.843,21	318,77	6.388,21	2.408,51	15.505,99

Slicer: [TIPO=FACTURA] [ESTADO=CERRADO] [Medida=Toneladas Vendidas] [Año=2016] [TIPO=FACTURA] [ESTADO=ABIERTO-IMPRESO] [Medida=Toneladas Vendidas] [Año=2016] [TIPO=FACTURA] [ES

Figura 5.63 Vista de Análisis de Ventas por Localidad.

- **Ventas por Vendedor:** Esta vista permite analizar las toneladas vendidas por Vendedor por año. Con estos datos se puede medir el cumplimiento del presupuesto de ventas. Además, se podrán establecer las metas de ventas de los siguientes periodos. En la Figura 5.64 se muestran las toneladas vendidas por vendedor y mercado del año 2016.

Vendedor	Producto							
	01 Masters	TRANSFERENCIA DE DOMINIO EXPORT	SERVICIO DE CONVERSION	BANANEROS	OTRAS FRUTAS	OTRAS EXPORTACIONES	FLORES OTROS	DOMESTICO
All_Vendedor	103.912,90	1.037,75	4,48	67.744,01	1.380,38	7.178,63	2.519,33	19.579,39
#null	108,21				22,10	26,21	1,20	57,15
ADONAI PACHECO	6.056,88			4.698,85	436,05	183,12		726,05
ANGELITA ULLOA	257,95	13,53					7,86	84,07
CHRISTIAN LUNA	3.700,20	1.024,22				319,73	2,04	832,07
CRISTINA MILLAN	3.149,41					89,14	2.415,01	544,44
ERICKA FREIRE	6,01							6,01
GABRIELA ESCALA	32,63					7,28		25,35
NEELTJE ROVELLO	7.238,97			6,98		1.889,64		4.929,60
OFICINA	426,25			8,67	78,69	13,54	34,78	66,25
RICHARD SIONG-TAY	6.105,59					1.808,91	1,58	4.289,31
PATRICIA BRAVO	4.761,56					373,69		4.192,32
ADRIANA MENDOZA	369,96			40,60		13,14		315,91
AURORA TORRES	3.964,12			48,08	711,98	1.970,08		821,78
GEOVANNA ALONSO	3,64					3,64		
SARA MACIAS	835,64			231,40	123,37	125,05		340,70
CARLOS SOLANO	33.929,59			33.929,59				
ELIZABETH SANCHEZ	42,61							6,22
ROBERTO MEJIA	6,56					1,41		5,15

Figura 5.64 Vista de Análisis de Ventas por Vendedor.

### 5.3.1.2 Vistas de análisis de comparativos de Ventas y Producción

Según lo manifestado en el levantamiento de requerimientos, para la empresa de manufactura es importante conocer la eficiencia en el despacho y venta de las unidades producidas. Por lo que se creó una vista de análisis que permite comparar las toneladas producidas contra las toneladas vendidas en distintos periodos. Esta vista está basada en el cubo virtual de ventas-ingresos. En la Figura 5.65 se muestra la vista de análisis que compara los ingresos y las ventas para el año 2016.

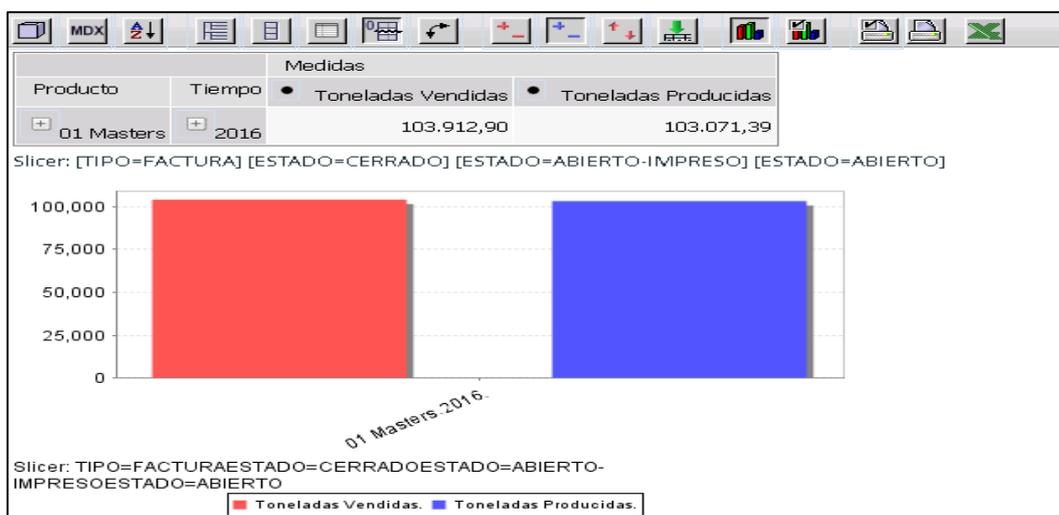


Figura 5.65 Vista de Análisis de Toneladas Producidas contra Toneladas Vendidas.

### 5.3.1.3 Vistas de análisis de comparativos de Pedidos por Periodo

Para poder determinar el crecimiento o disminución de los pedidos de ventas se creó la vista de análisis que permite comparar los pedidos de ventas en toneladas y dólares. Este análisis se puede realizar por tipo de clientes y por tipos de pedidos. En la Figura 5.66 se muestra la vista de la comparación de pedidos por Años. Esta vista de análisis está basada en el cubo OLAP Pedidos.

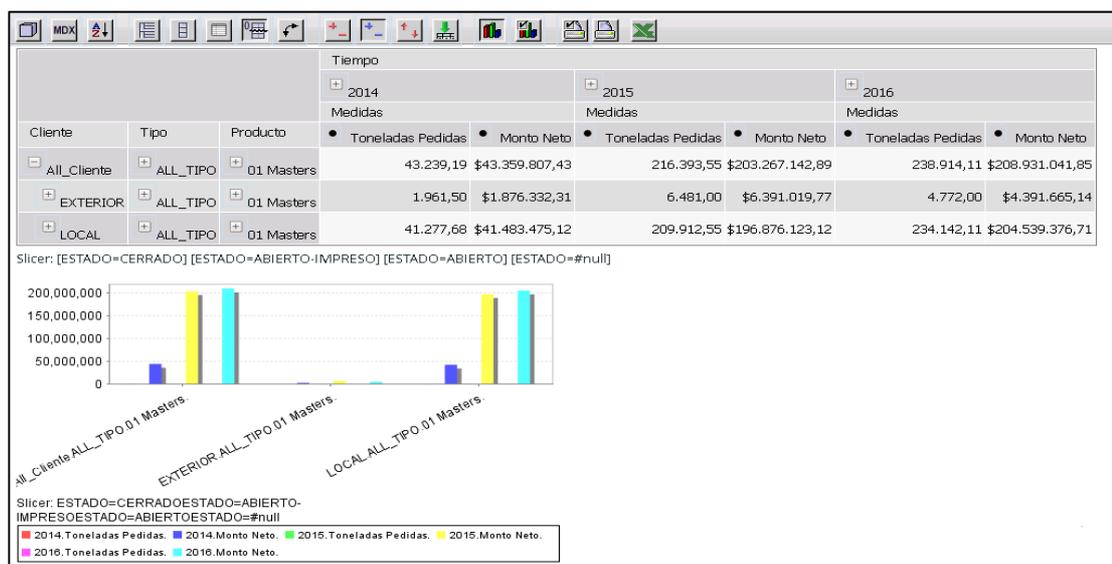


Figura 5.66 Comparativos de Pedidos por Año.

### 5.3.1.4 Vistas de análisis de Entregas

Para poder determinar la fiabilidad de las entregas, se creó la vista de análisis que se ilustra en la Figura 5.67. Esta vista permite comparar las entregas realizadas contra los pedidos de ventas ingresados, en toneladas. De esta forma se puede establecer si se cumplieron las entregas de los pedidos de ventas en las fechas de entregas correspondientes. La vista muestra información por la dimensiones de clientes, tipos de pedido, productos y tiempo. Esta vista está basada en el cubo virtual Pedidos-Entregas.

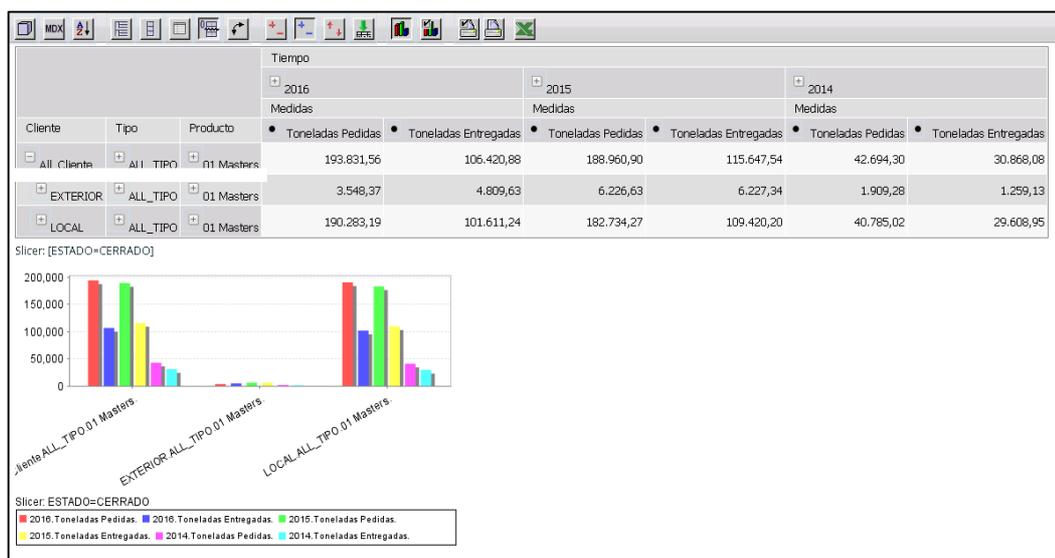


Figura 5.67 Vista de Análisis de Comparativo de Pedidos contra Entregas.

Además, se creó la vista de análisis que se muestra en la Figura 5.68 que permite analizar las entregas realizadas por las dimensiones de clientes, productos y tiempo. Esto permitirá navegar a través de los datos de las entregas, de acuerdo a las jerarquías de las dimensiones establecidas en el cubo OLAP de Entregas.

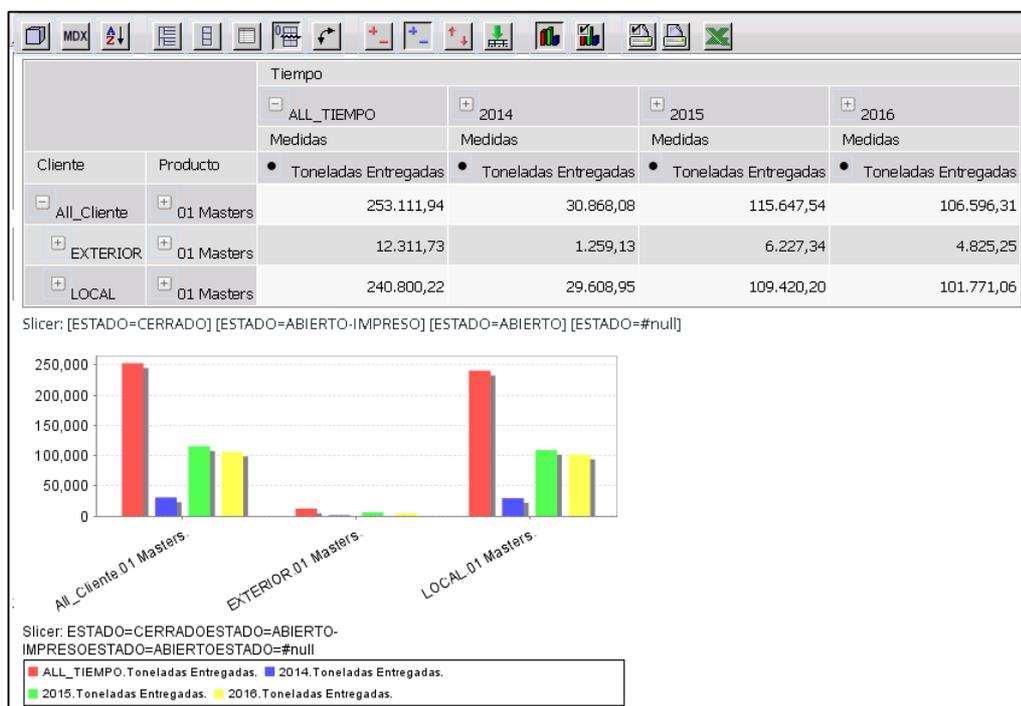


Figura 5.68 Vista de Análisis de Entregas.

Es importante indicar que se han creado las vistas de análisis que cumplen con los requerimientos establecidos en el alcance del proyecto. Pero en base a los cubos OLAP existentes se pueden crear distintas vistas para analizar los datos contenidos en los cubos, de acuerdo a las necesidades de cada usuario. Además, se pueden crear más cubos virtuales para combinar los datos de los cubos OLAP y realizar los análisis que el negocio requiera.

Una vez que se crearon las vistas de análisis, se procedió a crear los cuadros de mando que permiten una mejor visualización y manejo de la información.

### **5.3.2 Creación de cuadros de mandos**

Los cuadros de mando permiten visualizar y analizar la información de una manera gráfica, lo que facilita la comprensión de los resultados. Los cuadros de mando, al igual que las vistas de análisis, se crean en la PUC; para ello se utiliza la herramienta CDE DASHBOARD [16].

#### **5.3.2.1 Cuadro de Mando de Ventas**

Para facilitar el análisis de las ventas se creó el cuadro de mando que se aprecia en la Figura 5.69. Este cuadro de mando muestra las ventas en dólares y toneladas de acuerdo al año seleccionado y segmentadas por los principales mercados que maneja la empresa de manufactura. Los indicadores que se presentan son: Ventas Totales en Dólares, Ventas en dólares por Localidad y Ventas en Toneladas. De acuerdo al indicador seleccionado, se mostrará el gráfico de pie con los datos que correspondan al mercado y a las categorías de producto de ese mercado.

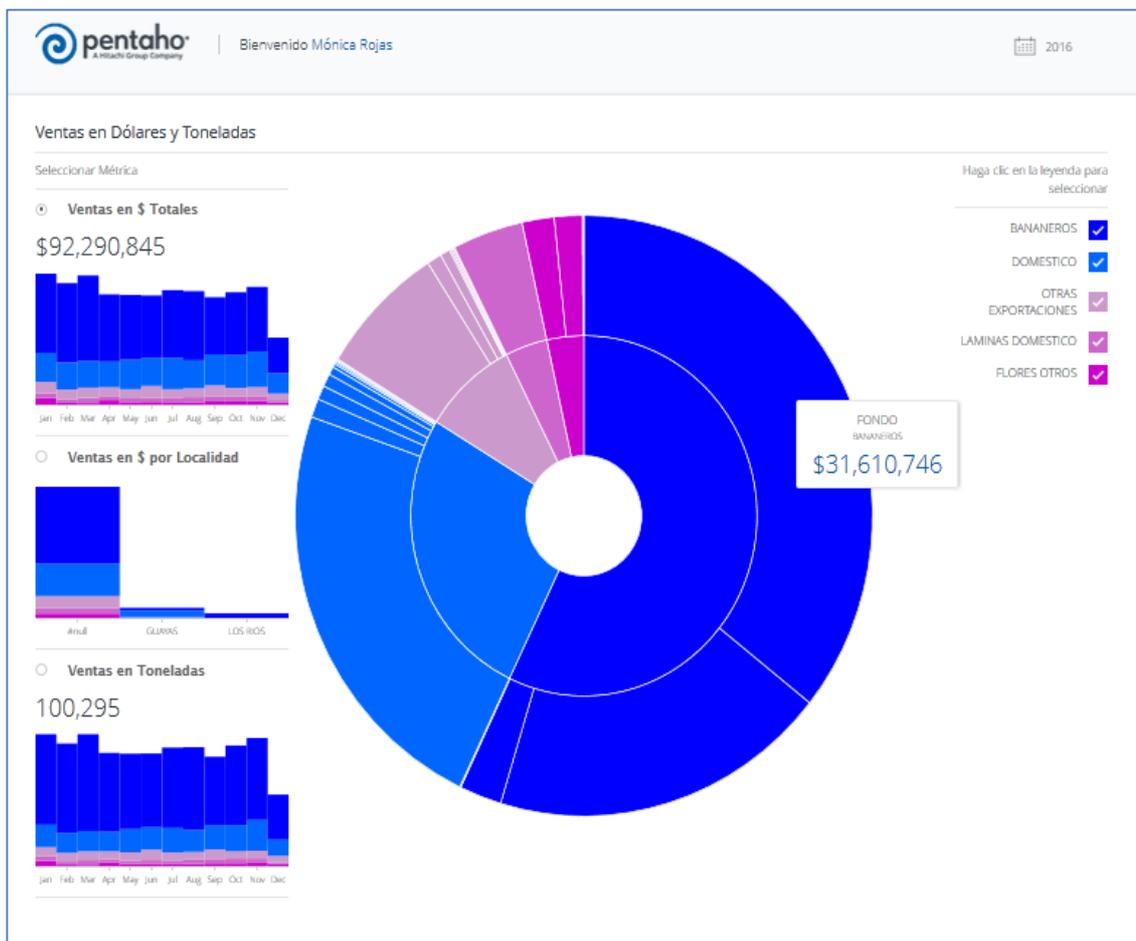


Figura 5.69 Cuadro de Mando de Ventas.

Con este cuadro de mando, el usuario puede visualizar las ventas que se realizaron por año, mercado y categoría de una manera gráfica y dinámica ya que dependiendo los parámetros que seleccione (tiempo y mercado) se presentarán los datos y gráficos.

### 5.3.2.2 Cuadro de Mando de Toneladas Vendidas y Producidas

Para la empresa de manufactura es importante conocer las toneladas que se venden y las toneladas que se producen para evitar la sobreproducción y la existencia de inventarios que aumentan el costo. Por esta razón, se creó el cuadro de mando que se ilustra en la Figura 5.70. Este cuadro presenta los indicadores de Ventas en Toneladas, Producción en Toneladas y Ventas en Toneladas por Localidad. Los parámetros que se utilizan como filtros son el año y el mercado.

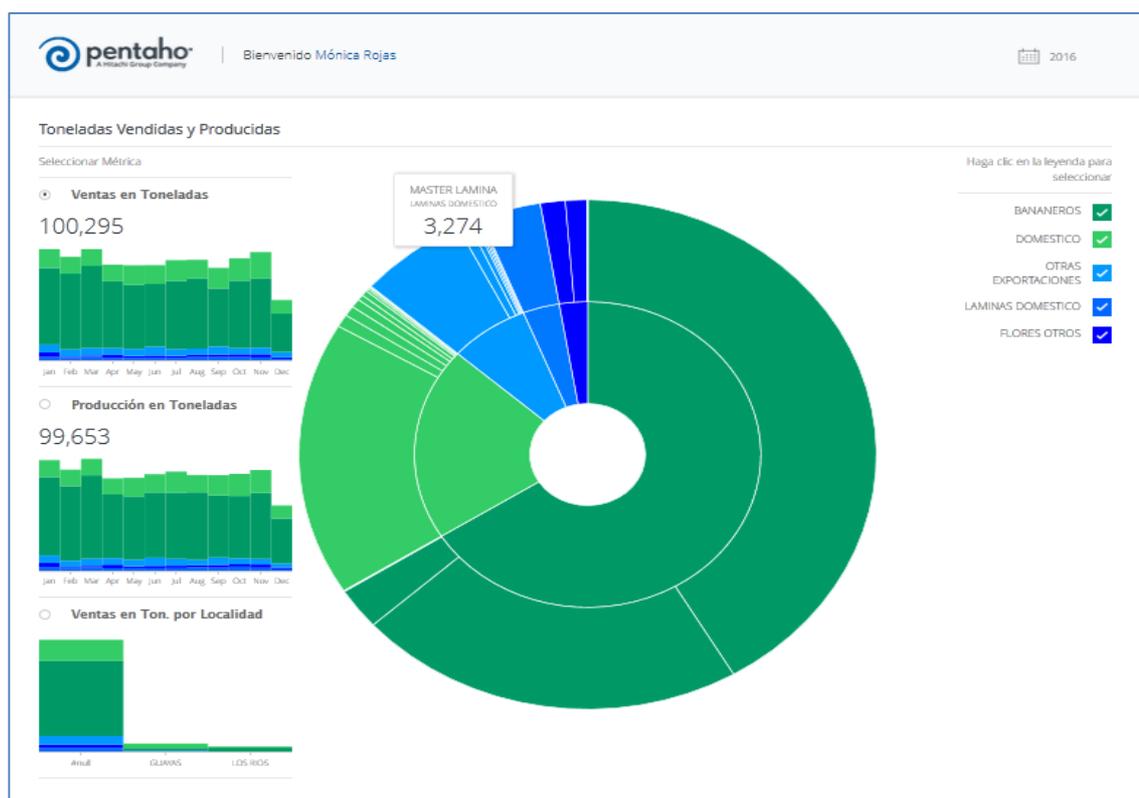


Figura 5.70 Cuadro de Mando de Toneladas Vendidas y Producidas.

### 5.3.2.3 Cuadro de Mando de Toneladas Pedidas y Entregadas

Para el área comercial, es de vital importancia conocer las toneladas que se piden y las que se entregan para poder verificar que los pedidos se despachen de acuerdo a las fechas de entrega establecidas. En el cuadro de mando que se muestra en la Figura 5.71 se puede apreciar los indicadores de toneladas pedidas y entregadas. Los parámetros de este cuadro son el año y el mercado. Dependiendo el indicador seleccionado se podrán visualizar los datos de las toneladas pedidas o entregadas.

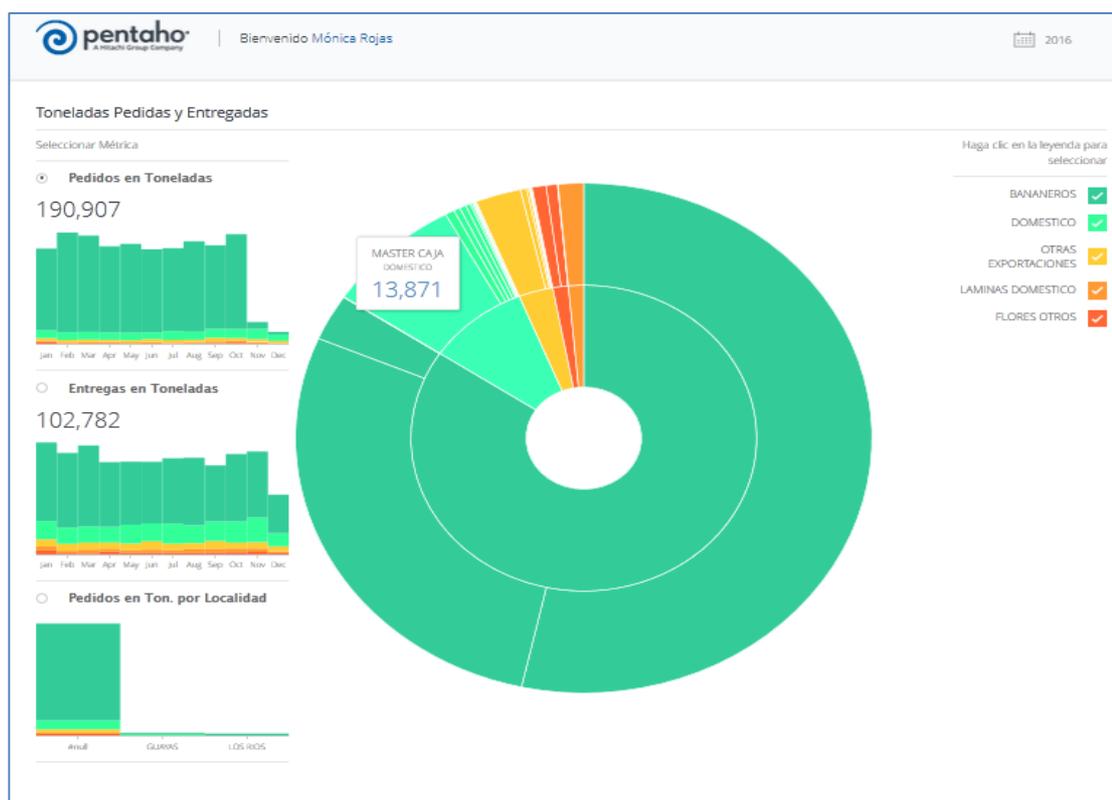


Figura 5.71 Cuadro de Mando de Toneladas Pedidas y Entregadas.

### 5.3.3 Administración y Seguridad de los Informes

Para administrar el acceso a los informes, ya sea de las vistas de análisis o de los cuadros de mandos, PENTAHO proporciona seguridad y permisos a través de roles. Los roles nativos que maneja PENTAHO son los siguientes:

- **Administrator:** En la Figura 5.69 se muestra los permisos que tiene este rol. Por ser el rol administrador los permisos no pueden ser modificados.

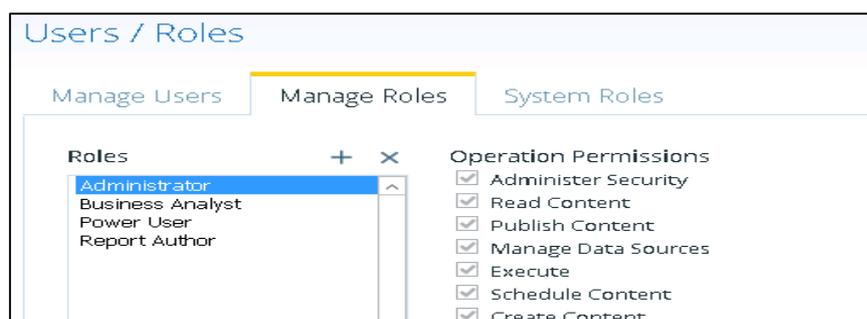


Figura 5.72 Permisos del rol Administrator

- **Business Analyst:** En la Figura 5.70 se muestran los permisos que tiene este rol. Se pueden agregar o quitar los permisos de este rol de acuerdo a las necesidades del negocio.

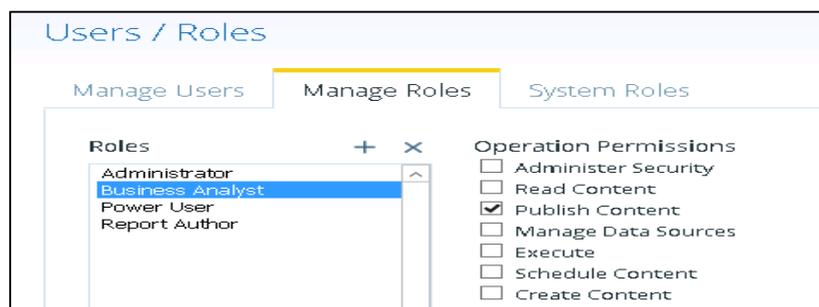


Figura 5.73 Permisos del rol Business Analyst

- **Power User:** En la Figura 5.71 se muestran los permisos nativos de este rol. Este rol permite leer, publicar, ejecutar y crear contenido. Es un rol con mayores permisos.

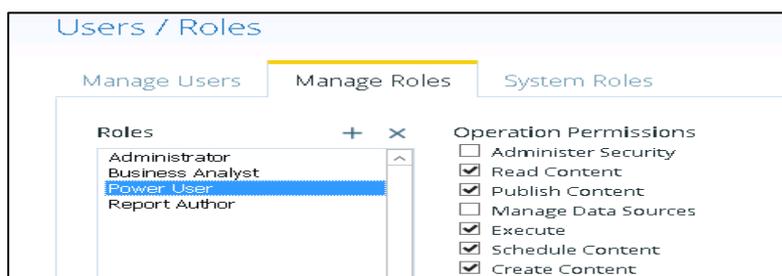


Figura 5.74 Permisos del rol Power User

- **Report Author:** En la Figura 5.72 se muestran los permisos de este rol.

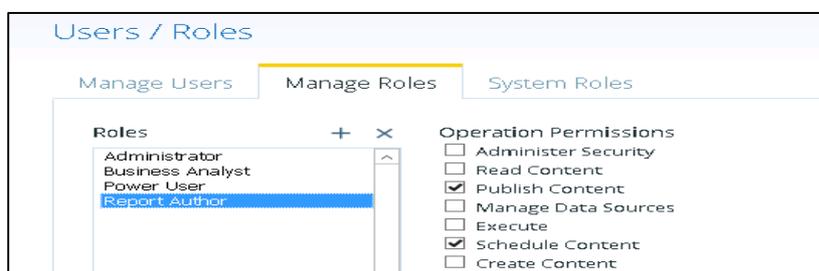


Figura 5.75 Permisos del rol Report Author

Estos roles se asignan a los usuarios y de acuerdo al rol asignado, el usuario tendrá los permisos para crear, publicar, ejecutar o administrar contenido. A continuación se indica cómo se crea un usuario en PENTAHO.

### 5.3.3.1 Creación de usuarios

Para crear un usuario en PENTAHO se debe tener rol de Administrator. Se asigna un nombre y contraseña al usuario y se procede a establecer el rol que tendrá. En la Figura 5.73 se muestra la creación de un usuario.

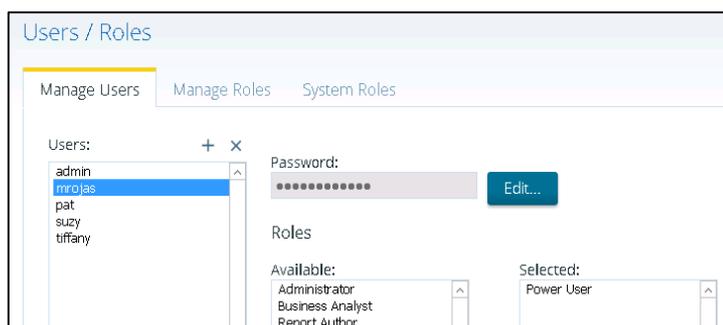


Figura 5.76 Creación de usuario

Siguiendo lo indicado se crearon los usuarios para el área comercial, ejecutivos comerciales, servicio al cliente, jefes de área y gerentes. Por ahora, se ha decidido usar los roles nativos que proporciona PENTAHO con los mismos permisos ya que se acoplan al negocio. Más adelante se revisarán estos roles y, si es necesario, se actualizarán los permisos y/o se crearán nuevos roles.

### 5.3.3.2 Asignación de accesos a los informes

El acceso a un informe se puede dar por rol o por usuario. Para ello se comparte el reporte y se indica a qué usuarios y roles se les permitirá el acceso. Además, se deben seleccionar los permisos que tendrá como son: Control Total, Escritura, Lectura y/o Borrado. En LA Figura 5.74 se muestra la asignación del permiso al informe.

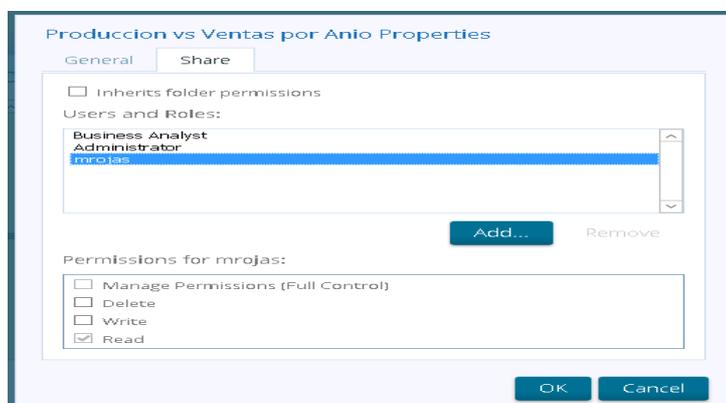


Figura 5.77 Asignación de rol/usuario a un informe.

Los permisos para los informes que se han creado se darán a nivel de usuarios, debido a que inicialmente la gerencia ha decidido probar los informes con un número seleccionado de usuarios. Posteriormente se analizará la creación de perfiles, de acuerdo a las actividades que realicen los usuarios. Lo importante es que PENTAHO facilita la administración y la seguridad en el manejo de los informes, por lo que la gerencia no se debe preocupar por el acceso indebido a los datos.

Después de definir y establecer el manejo de los accesos y seguridad de los informes, se procederá a explicar cómo se validaron y verificaron los datos en los procesos que se han realizado, como son: ETL, Cubos OLAP e informes para los usuarios.

## **5.4 Validación y Verificación de los requerimientos**

En el presente trabajo se han realizado varios procesos hasta llegar a la creación de los informes (vistas de análisis y cuadros de mando) que es el trabajo visible para el usuario final. Durante estos procesos se han realizado pruebas y validaciones de los datos con el objetivo de asegurar que el proceso se realizará de manera correcta y que los datos que se almacenaron en el datamart son confiables. A continuación se describen las pruebas que se llevaron a cabo en los procesos.

### **5.4.1 Validación y Pruebas en el proceso ETL**

Luego de crear el modelo dimensional que se ajustaba a la realidad del negocio, el siguiente punto importante fue la creación de los procesos ETL. Este proceso es el pilar fundamental para que el proyecto sea exitoso, debido a que en este proceso se extraen los datos del modelo relacional para transformarlos y cargarlos al datamart. Si la extracción es hecha de forma errónea, no servirá

tener los mejores y más vistosos cuadros de mando ya que sólo mostrarán información no real y, por ende, no válida para el negocio.

Para la creación del proceso ETL, primero se crearon las transformaciones de las dimensiones y luego de las tablas de hecho. Se empezó creando las transformaciones de las dimensiones con menos cantidad de registros como son: vendedor, unidad de medida, condición de crédito y localidades. Posteriormente se crearon las transformaciones para las dimensiones de cliente y productos que contienen mayor número de registros. Una vez creada la transformación se procedió a ejecutarla en el PDI, en algunos casos se presentaban problemas por tipos de datos o datos nulos en los registros. Para superar estos inconvenientes se añadieron a las transformaciones, objetos propios de PDI que permiten eliminar datos nulos. Luego se superó la carga de datos en las tablas de dimensiones se procedió a validar que los datos cargados sean correctos mediante el uso de consultas SQL, tanto en la base origen como en la destino. La validación básica fue que el número de registros cargados en las dimensiones seas iguales a los registros de las tablas de origen bajo las condiciones establecidas. Después de corroborar que los datos cargados en las tablas de dimensión fueran los correctos, se procedió con la creación de las transformaciones de las tablas de hecho. Estas transformaciones resultaron más complejas puesto que se trataba de extraer los datos de las

transacciones de pedidos, ventas, entregas e ingreso de inventario. El principal problema presentado fue que en algunos casos existían datos duplicados por lo que se debió recurrir a la verificación de los datos en las tablas de origen. Al igual que en la carga de las dimensiones, la verificación de la carga de las tablas de hecho se realizó mediante el uso de consultas SQL, validando que la cantidad de registros sean iguales y que los valores cargados en unidades y dólares sean los correctos.

#### **5.4.2 Validación y Pruebas del Cubo OLAP**

Una vez que se cargó el datamart y se verificó que los datos eran correctos, validados contra los datos de la base de origen. Se procedió a crear los cubos OLAP que se identificaron. Para la validar que los cubos fueron creados correctamente se procedió a publicar el esquema, bajo el cual fueron creados, en el servidor BI de PENTAHO. Luego de publicar el cubo se procedió a crear una vista de análisis de pruebas; si el cubo estaba correctamente creado la vista se creaba sin problemas; de lo contrario, mostraba error y para ello se debía revisar el log generado en el servidor. Por lo general los errores presentados fueron en la creación de medidas y campos calculados. Para superar estos inconvenientes se recurrió a la información técnica que se brinda en la documentación de PENTAHO. También se debió reiniciar varias veces el servidor BI ya

que al ser una herramienta Open Source suele presentar este tipo de errores que se resuelven al reiniciar el servidor. Las validaciones en este proceso fueron más de tipo técnicas, ya que las verificaciones de datos se harán una vez que se creen las vistas de análisis y los cuadros de mando.

### **5.4.3 Validación y Pruebas de los Informes**

Después de verificar que los cubos OLAP y virtuales estuvieran correctamente creados, se procedió a crear las vistas de análisis de acuerdo a los requerimientos definidos en el alcance. La validación de los informes se realizó en dos fases. Primero se validó con los reportes estáticos que usaban los usuarios, se verificó que los datos mostrados en determinado periodo sean los mismos. Esto se lo realizó de manera resumida, es decir a nivel general. La segunda parte se realizó con los usuarios que usan esta información. Se les dio acceso a las vistas para que revisarán los informes y verificarán que se muestre la información que requieren. Se realizaron algunas modificaciones a las vistas para mostrar la información de acuerdo a las necesidades de los usuarios. Una vez que se acordó el formato, se les solicitó que utilizarán los informes en paralelo con los demás reportes que suelen utilizar (cabe aclarar que los informes que utilizan son los preparados por ellos ya que como se indicó al inicio de este documento, el sistema no les brinda reportes dinámicos). Durante este periodo de prueba, en algunas ocasiones los usuarios

mencionaron que no les cuadraban los datos; por lo que se procedió a verificar las fuentes, encontrándose, en la mayoría de los casos, que la falla se encontraba en los informes que ellos solían utilizar. Para los casos que los errores se encontraban en las vistas de análisis de PENTAHO, se procedió a ajustar los ETL ya que se descubrió que algunos datos no estaban en el datamart por el uso incorrecto de los filtros en los ETL.

En el presente capítulo se describió el proceso de desarrollo de la solución propuesta. Se detalló la creación de los ETLs, cubos OLAP, vistas de análisis y cuadros de mando. También se indicó cómo Pentaho administra los accesos a los informes. Por último, se describió cómo se validaron y verificaron los requerimientos.

En el siguiente capítulo se hace un análisis de los resultados de la implementación. Se evalúan los resultados obtenidos con la puesta en producción de la solución y se plantean alternativas de mejora.

## **CAPÍTULO 6**

### **ANÁLISIS DE RESULTADOS**

En este capítulo se describen los resultados obtenidos luego de implementar la solución propuesta. Se analizan los beneficios conseguidos para el negocio. Además, se proponen alternativas para mejorar.

#### **6.1 Evaluación de los resultados de la implementación**

Según disposición de la gerencia, la solución DD-DSS fue probada con un grupo seleccionado de usuarios del área comercial con la finalidad de evaluar la herramienta. Los usuarios a los que se dio acceso a la herramienta fueron:

- Jefe de servicio al cliente: usuario encargado de asegurar y verificar el cumplimiento de las entregas.
- Analista comercial: usuario encargado de analizar las ventas en diferentes periodos para asegurar el cumplimiento del presupuesto de

ventas. Asimismo, se encarga del análisis de los pedidos de los clientes para verificar el crecimiento o pérdida de mercado.

- Analista de producción: usuario encargado de analizar los ingresos de la producción para identificar que la planta produzca de acuerdo a los pedidos ingresados en el tiempo que corresponde.

Estos usuarios son los encargados del análisis diario del área comercial para asegurar el buen funcionamiento y cumplimiento de las metas de esta área. Para ello, utilizaban reportes que les proporciona el ERP, que como se indicó son reportes estáticos que luego son consolidados en Excel.

Una vez que se desarrollaron las vistas de análisis y los cuadros de mando, indicados en el alcance del proyecto, se procedió a dar acceso y capacitar a los usuarios en la herramienta de acuerdo a su perfil. Se hicieron ajustes y correcciones de acuerdo a las necesidades del cargo del usuario.

Se estableció un periodo de prueba y seguimiento de la herramienta de cuatro semanas. Después de transcurrido este tiempo, se realizaron entrevistas con los usuarios para conocer la experiencia que tuvieron trabajando con las vistas de análisis y los cuadros de mando.

De acuerdo a sus opiniones, la herramienta les resultó útil ya que les redujo significativamente el tiempo de obtener la información para analizar y presentar a gerencia. Anteriormente les tomaba 2 horas en tener la

información consolidada y lista. Actualmente con las vistas de análisis, que les proporciona la herramienta, se tardan 5 segundos en obtener la información. Este ahorro de tiempo representa una mejora en el 99% del tiempo de respuesta para la generación de informes. Por consiguiente, los usuarios pueden dedicarse a realizar las tareas propias de su cargo, en lugar de generar informes de forma manual. Además, tienen la seguridad de que los datos proporcionados son fiables debido a que no se ha tenido intervención humana. Los usuarios dijeron también que con las vistas de análisis pueden revisar los datos de una forma dinámica y rápida desde distintos niveles sin tener que ir a otros reportes para obtener los detalles de la información.

Otro punto importante a destacar, es que en las reuniones de gerencia se presentan los resultados de las ventas con los cuadros de mando, que les permiten analizar el comportamiento de las ventas por periodos en términos de dólares y toneladas de manera interactiva. Anteriormente se presentaban los resultados en diapositivas de Power Point. Por lo tanto, ahora la empresa cuenta con mejores herramientas para el análisis y toma de decisiones que son brindadas por la aplicación implementada. Asimismo, cabe indicar que la herramienta es accedida desde un navegador web. Esto facilita el acceso a la misma desde lugares fuera de la oficina que tengan acceso a Internet, por ahora mediante el uso de una red privada virtual (VPN, por sus siglas en inglés Virtual Private Network).

## 6.2 Establecer alternativas de mejora

Considerando la retroalimentación recibida por parte de los usuarios y el análisis realizado, se han definido las siguientes mejoras para la solución implementada:

- Creación de reportes dirigidos al personal operativo de servicio al cliente, para facilitar las revisiones diarias de los movimientos del área comercial. Para esto, se sugiere el uso de Pentaho Reporting que permite crear reportes en base a los cubos OLAP y también de las tablas de dimensiones. Estos reportes deben ser publicados en el servidor para su consumo.
- Creación de un nuevo cubo OLAP para almacenar el presupuesto de ventas. De esta manera se podrá comparar las ventas presupuestadas contra las ventas reales, permitiendo verificar el cumplimiento del presupuesto elaborado al comienzo del año por cada ejecutivo comercial, y corregir posibles desviaciones.
- Creación de un nuevo cubo virtual, entre los cubos OLAP de pedidos y ventas. Este cubo permitirá analizar los pedidos que se convirtieron en ventas reales. Con esta información, el analista comercial podrá realizar el seguimiento de los pedidos con mayor rapidez y facilidad.
- Creación de un nuevo cubo OLAP para registrar las devoluciones de ventas. Este cubo debe permitir identificar las principales causas de devolución y poder establecer las acciones correctivas necesarias.

# CONCLUSIONES Y RECOMENDACIONES

## Conclusiones

1. En este trabajo se ha mostrado el diseño y la implementación de un sistema DD-DSS para una empresa de manufactura, el cual cumple con todos los requerimientos y necesidades de los usuarios y la empresa. El sistema DD-DSS implementado facilita el acceso a los datos relevantes de la compañía de manufactura, brindando información concisa y detallada según lo requieran los usuarios.
2. La fortaleza más importante del DD-DSS desarrollado es que se basa en un modelo dimensional adecuado a las necesidades y la realidad de la empresa. Esto se logró, gracias al apoyo del personal del área comercial que colaboró en todo momento para dar a conocer los problemas que afrontaba esta área con relación a la necesidad de información.

3. Gracias al uso de la metodología Scrum, se pudo satisfacer las necesidades y requerimientos de información de todos los usuarios involucrados en el desarrollo.
4. El desarrollo de ese trabajo permitió corroborar que el tiempo de implementación de la solución es relativamente corto si se diseña un modelo dimensional apropiado. También, hay que considerar que se utilizó la versión open source, por lo que la creación de vistas y cuadros de mando resulta un poco más tediosa y demorada.
5. Cabe indicar que en el presente trabajo se aplicaron varios de los conocimientos adquiridos en la maestría, como son: elección de una correcta metodología ágil que se adapte al entorno de trabajo de la empresa, identificación de los requerimientos funcionales y no funcionales aplicando los conceptos de ingeniería de requerimientos. También fue de gran ayuda lo aprendido en perspectivas de software libre para la selección de una plataforma de código abierto que sea robusta y relativamente fácil de usar. Además, uno de los factores más importantes para el desarrollo de este trabajo fue lo aprendido en el uso estratégico de la información, donde se pudo entender cuán importante es el acceso y disponibilidad de información relevante para cualquier empresa.

## Recomendaciones

1. Establecer políticas y procedimientos para el buen uso de la información. Durante el proceso de implementación se detectó que en el sistema transaccional no se registraban todos los datos maestros de clientes, productos, entre otros, que posteriormente se pueden utilizar para analizar información.
2. Crear datamarts para inventarios y cuentas por cobrar. De modo que se pueda integrar con la información de ventas y proporcionar información integral de la compañía para la toma de decisiones.
3. Capacitar a los usuarios para que puedan crear sus propias vistas de análisis según sus necesidades. Si los cubos están bien diseñados cada usuario estará en capacidad de crear informes en base a los cubos existentes.
4. Implementar minería de datos puesto que mediante la aplicación de modelos matemáticos se podrá tener información que sugiera las posibles estrategias de ventas y de operaciones de la compañía.
5. Para que un proyecto de este tipo sea exitoso, se recomienda que se revisen los datos que alimentarán a los datamarts. Las fuentes desde donde procederán estos datos: sistemas transaccionales, archivos, fuentes externas, etc., de tal forma que se pueda identificar que se estén guardando

los datos correctamente, y de no ser así, la empresa tome las medidas correctivas necesarias. Porque si no se cuenta con datos de calidad, los sistemas DSS no serán de utilidad.

## BIBLIOGRAFÍA

[1] Grupo Surpapel, Quienes somos, <http://www.gruposurpapel.com/>, Consultado Junio 2016.

[2] Stair Ralph, Reynolds George, Principles of Information Systems, Course Technology, Publicado Enero 2013.

[3] Pentaho, a Hitachi Group company, Pentaho Community, <http://community.pentaho.com/>, consultada Junio 2016.

[4] Turban Efraim, Sharda Ramsh, Delen Dursun, Decision Support and Business Intelligence Systems 9th Edition, Prentice Hall, Publicado Febrero 2010.

[5] Curto Diaz Josep, Conesa Caralt Jordi, Introducción al Business Intelligence, Editorial UOC, Publicado Mayo 2012.

[6] Schniederjans Marc J., Schniederjans Dara G., Starkey Christopher M., Business Analytics, principles, concepts and applications, Pearson Education Inc, Publicado Abril 2014.

- [7] Power Daniel, Supporting Decision-Makers: An Expanded Framework, <http://dssresources.com/papers/supportingdm/PowerEBKSupp.pdf>, Publicado Junio 2001.
- [8] Power Daniel, Understanding Data-Driven Decision Support Systems, [https://www.researchgate.net/publication/220630419\\_Understanding\\_Data-Driven\\_Decision\\_Support\\_Systems](https://www.researchgate.net/publication/220630419_Understanding_Data-Driven_Decision_Support_Systems), Publicado Marzo 2008.
- [9] Morales Martínez Maribel, Analítica Web para empresas, Editorial UOC, Publicado Abril 2010.
- [10] Ballard Chuck, Farrel Daniel M., Gupta Amit, Mazuela Carlos, Vohnik Stanislav, Dimensional Modeling; In a Business Intelligence Environment, IBM, Corp., Publicado Marzo 2006.
- [11] Hughes Ralph, Agile Data Warehousing for the Enterprise: A Guide for Solution Architects and Project Leaders, Elsevier, Inc., Publicado Julio 2015.
- [12] Dimes Troy, Conceptos Básicos de Scrum: Desarrollo de software Agile y manejo de proyectos Agile, Babelcube, Inc., Publicado Marzo 2015.
- [13] Casters Matt, Bouman Roland, Dongen Jos Van, Pentaho Kettle Solutions: Building Open Source ETL Solutions with Pentaho Data Integration, Wiley Publishing, Inc, Publicado Enero 2010.

[14] Latino Joel, Ward Harris, Pentaho Analytics for MongoDB Cookbook, Packt Publishing Ltd., Publicado Diciembre 2015.

[15] Patil Manoj R., Rhia Feris, Pentaho for Big Data Analytics, Packt Publishing Ltd., Publicado Enero 2013.

[16] Gaspar Miguel, Learning Pentaho CTools, Packt Publishing Ltd., Publicado Mayo 2016.

# APÉNDICE

## **APÉNDICE A**

Documento de Especificación de Requerimientos de Software

## **APÉNDICE B**

Documento de Creación de ETL's en PDI