



**UNIVERSIDAD DEL BÍO-BÍO**

FACULTAD DE CIENCIAS EMPRESARIALES

DEPARTAMENTO CIENCIAS DE LA COMPUTACIÓN Y TECNOLOGÍAS DE LA INFORMACIÓN

# **IMPLEMENTACIÓN DE BUSINESS INTELLIGENCE PARA UNA PYME LOCAL DEL RUBRO ELÉCTRICO**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL EN INFORMÁTICA

**AUTORES**

**: PEÑA GUTIÉRREZ GONZALO ANDRÉS  
PINCHEIRA HENRÍQUEZ IVÁN MAURICIO**

Profesor Guía

: Dr. Gutiérrez Retamal Gilberto

CHILLÁN, 2014

## Resumen

Este proyecto se presenta para dar conformidad a los requisitos exigidos por la Universidad del Bío-Bío en el proceso de titulación para la carrera de Ingeniería Civil en Informática. El proyecto titulado Implementación de Business Intelligence para una Pyme del Rubro Eléctrico, muestra el diseño e implementación de un DataMart para el área de ventas clientes de la empresa VALFI, la cual es una comercializadora del rubro eléctrico e insumos computacionales de la ciudad de Chillán. A partir de una serie de entrevistas con los dueños y empleados de VALFI se consiguió recolectar los requisitos con los que se logró diseñar el DataMart.

Previo a la implementación del DataMart, se realizó un análisis de las herramientas disponibles en el mercado para el proceso de extracción, transformación y análisis de los datos.

En una etapa previa a los procesos ETL y luego de realizado un análisis de las herramientas, se realizó un proceso de modelado dimensional, usando como referencia la metodología Kimball, para la definición de las dimensiones y hechos que compondrán el DataMart.

Luego se realizó el proceso de extracción, transformación y carga de los datos a partir de la base de datos transaccional de VALFI hacia un DataMart.

En una etapa siguiente se muestra el análisis de datos a partir del DataMart previamente creado, con el objetivo de crear un análisis OLAP y un cuadro de mando integral o Dashboard.

Las posibilidades que entrega el análisis de datos permiten concluir que nuestra aplicación representa una herramienta que puede impactar positivamente en VALFI. Por ejemplo, ver el comportamiento de las ventas durante los doce meses del año, ver la procedencia de los clientes con más ventas o ver cuáles son los productos más vendidos durante un año.

## **Abstract**

In this report, is shows the design and implementation of a Sell Client's DataMart for the Company VALFI, which is a marketing Electric Sector and computer supplies of Chillan City. Starting from a series of interviews with the owners and employees of VALFI was obtained collect the requirements with which it was possible the DataMart Design.

Previous of the implementation of Datamart, an analysis of tools commercially available is performed for the Extract, Transform and Analysis of Data.

Then was performed the Extract, Transform and Load Process of data starting of transactional Data Base of VALFI to DataMart.

In a following step show the Data Analysis starting of DataMart previously created, with the objective of create OLAP Analysis and Scorecard or Dashboard.

The possibilities delivered by the analysis be concluded our application represents a tool that can positively impact VALFI. For example view the sales performance during the twelve months of the year, see the origin of customers with more sales or see what the best sellers are for a year.

# Índice General

<b>1</b>	<b>INTRODUCCIÓN.....</b>	<b>10</b>
1.1	OBJETIVO GENERAL.....	12
1.2	OBJETIVOS ESPECÍFICOS .....	12
1.3	ALCANCE Y LÍMITES DEL PROYECTO .....	12
1.4	PROBLEMÁTICA DE LA EMPRESA VALFI .....	12
1.5	ORGANIZACIÓN DEL INFORME .....	13
<b>2</b>	<b>DESCRIPCIÓN DE LA EMPRESA .....</b>	<b>14</b>
2.1	IDENTIFICACIÓN DE LA EMPRESA.....	14
2.2	LOGO DE LA EMPRESA.....	14
2.3	POLÍTICAS DE LA EMPRESA .....	15
2.3.1	VISIÓN DE LA EMPRESA VALFI .....	15
2.3.2	MISIÓN DE LA EMPRESA VALFI.....	15
2.4	VALFI ELECTRICIDAD.....	15
<b>3</b>	<b>CONCEPTOS .....</b>	<b>16</b>
3.1	BUSINESS INTELLIGENCE (BI) .....	16
3.2	DATAWAREHOUSE (DW) Y DATAMART (DM).....	16
3.3	METODOLOGÍA KIMBALL.....	17
3.4	CONCEPTOS UTILIZADOS EN EL DISEÑO DE LA SOLUCIÓN .....	18
3.4.1	DIMENSIÓN .....	18
3.4.2	MEDIDAS.....	19
3.4.3	JERARQUÍA .....	20
3.4.4	TABLA DE HECHOS.....	20
3.4.5	ESQUEMA ESTRELLA (STARSCHEMA) .....	21
3.4.6	EXTRACCIÓN, TRANSFORMACIÓN Y CARGA (ETL) .....	22
3.4.7	HERRAMIENTAS ETL.....	22
3.4.8	PENTAHO DATA INTEGRATION (KETTLE) .....	23
3.4.9	CUADRO COMPARATIVO DE HERRAMIENTAS ETL .....	23
3.4.10	CUBO OLAP .....	24
3.4.11	ANÁLISIS OLAP.....	25
3.4.12	CUADRO DE MANDO (DASHBOARD/ BALANCEDSCORECARD). .....	25
3.4.13	CUADRO COMPARATIVO DE HERRAMIENTAS DE ANÁLISIS.....	26
<b>4</b>	<b>SOLUCIÓN .....</b>	<b>27</b>
4.1	DEFINICIÓN DE REQUERIMIENTOS.....	28
4.2	MODELO DE DIMENSIONES, JERARQUÍAS Y HECHOS.....	28
4.2.1	DIMENSIONES.....	28
4.2.2	TABLA DIMENSIÓN “FECHA” .....	29
4.2.3	TABLA DIMENSIÓN “PRODUCTO” .....	31
4.2.4	TABLA DIMENSIÓN “CLIENTE” .....	33
4.2.5	TABLA DE HECHOS.....	37
4.2.6	TABLA DE HECHO FINAL .....	39
4.3	PROCESO EXTRACCIÓN, TRANSFORMACIÓN Y CARGA (ETL).....	40

4.3.1	JOB PRINCIPAL .....	41
4.3.2	EXTRACCIÓN .....	42
4.3.3	TRANSFORMACIONES DIMENSIONES.....	43
4.3.4	TRANSFORMACIÓN DIMENSIÓN “PRODUCTO” .....	44
4.3.5	TRANSFORMACIÓN DIMENSIÓN “CLIENTE” .....	45
4.3.6	TRANSFORMACIÓN DIMENSIÓN “FECHA” .....	46
4.3.7	TRANSFORMACIÓN HECHO “VENTA” .....	47
4.3.8	CARGA DE DIMENSIONES Y HECHO .....	48
<b>4.4</b>	<b>ANÁLISIS OLAP.....</b>	<b>51</b>
<b>4.5</b>	<b>CUADRO DE MANDO INTEGRAL (DASHBOARD/BALANCEDSCORECARD).....</b>	<b>56</b>
<b>5</b>	<b><u>PRUEBAS.....</u></b>	<b>58</b>
<b>6</b>	<b><u>CONCLUSIONES Y FUTURAS EXTENSIONES.....</u></b>	<b>59</b>
6.1	CONCLUSIONES .....	59
6.2	TRABAJO FUTURO.....	61
<b>7</b>	<b><u>BIBLIOGRAFÍA .....</u></b>	<b>62</b>
<b>8</b>	<b><u>ANEXO: DETALLE JOB PRINCIPAL.....</u></b>	<b>63</b>
<b>9</b>	<b><u>ANEXO: DETALLE EXTRACCIÓN.....</u></b>	<b>64</b>
<b>10</b>	<b><u>ANEXO: DETALLE JOB TRANSFORMACIÓN .....</u></b>	<b>65</b>
<b>11</b>	<b><u>ANEXO: DETALLE TRANSFORMACIÓN “PRODUCTO” .....</u></b>	<b>66</b>
<b>12</b>	<b><u>ANEXO: DETALLE TRANSFORMACIÓN “CLIENTE” .....</u></b>	<b>67</b>
<b>13</b>	<b><u>ANEXO: DETALLE TRANSFORMACIÓN “FECHA” .....</u></b>	<b>68</b>
<b>14</b>	<b><u>ANEXO: DETALLE TRANSFORMACIÓN HECHO “VENTA” .....</u></b>	<b>69</b>
<b>15</b>	<b><u>ANEXO: DETALLE JOB CARGA .....</u></b>	<b>70</b>
<b>16</b>	<b><u>ANEXO: DETALLE CARGA “CLIENTE” .....</u></b>	<b>71</b>
<b>17</b>	<b><u>ANEXO: DETALLE CARGA “PRODUCTO” .....</u></b>	<b>72</b>
<b>18</b>	<b><u>ANEXO: DETALLE CARGA “FECHA” .....</u></b>	<b>73</b>
<b>19</b>	<b><u>ANEXO: DETALLE CARGA “VENTA” .....</u></b>	<b>74</b>
<b>20</b>	<b><u>ANEXO: PLAN DE PRUEBAS .....</u></b>	<b>75</b>
<b>20.1</b>	<b>CASO DE PRUEBA JOB EXTRACCIÓN .....</b>	<b>75</b>
20.1.1	CASO DE PRUEBA 1 .....	75
20.1.2	EJECUCIÓN CASO DE PRUEBA 1 .....	76
<b>20.2</b>	<b>CASO DE PRUEBA JOB TRANSFORMACIÓN .....</b>	<b>89</b>
20.2.1	CASO DE PRUEBA 2 .....	89
20.2.2	EJECUCIÓN CASO DE PRUEBA 2 .....	89
<b>20.3</b>	<b>CASO DE PRUEBA JOB CARGA .....</b>	<b>97</b>
20.3.1	CASO DE PRUEBA 3 .....	97
20.3.2	EJECUCIÓN CASO DE PRUEBA .....	97
<b>20.4</b>	<b>CASO DE PRUEBA CUBO OLAP .....</b>	<b>103</b>
20.4.1	CASO DE PRUEBA 4 .....	103
20.4.2	EJECUCIÓN DEL CASO DE PRUEBA.....	103

<b>21</b>	<b><u>ANEXO: MANUAL DE CONFIGURACION.....</u></b>	<b><u>106</u></b>
<b>21.1</b>	<b>CUBO OLAP.....</b>	<b>106</b>
<b>21.2</b>	<b>CONFIGURACIÓN PENTAHO BUSINESS ANALYTICS.....</b>	<b>108</b>
21.2.1	PLUGIN PENTAHO PARA CREAR DASHBOARD .....	112
21.2.2	PLUGIN PENTAHO ANÁLISIS OLAP .....	112

## Índice Tablas

Tabla 1 COMPARACIÓN HERRAMIENTAS ETL .....	24
Tabla 2 Comparativo Herramientas de Análisis .....	27
Tabla 3 Atributos Dimensión "Fecha" .....	31
Tabla 4 Atributos Dimensión "Producto" .....	33
Tabla 5 Atributos Dimensión "Cliente" .....	36
Tabla 6 Atributos Hecho Principal "Venta" .....	38
Tabla 7 Detalle Job Principal .....	63
Tabla 8 Detalle Extracción.....	64
Tabla 9 Detalle Job Transformación.....	65
Tabla 10 Detalle Transformación "Producto" .....	66
Tabla 11 Detalle Transformación "Fecha" .....	68
Tabla 12 Detalle Hecho "Venta" .....	69
Tabla 13 Detalle Job Carga.....	70
Tabla 14 Detalle Carga "Cliente" .....	71
Tabla 15 Detalle Carga "Producto" .....	72
Tabla 16 Detalle Carga "Fecha".....	73
Tabla 17 Detalle Carga "Venta" .....	74
Tabla 18 Caso de Prueba 1 .....	75
Tabla 19 Caso de Prueba 2 .....	89
Tabla 20 Caso de Prueba 3 .....	97
Tabla 21 Caso de Prueba 4 .....	103

## Índice Figuras

Figura 1 Imagen Corporativa VALFI .....	14
Figura 2 Pasos Metodología Kimball.....	18
Figura 3 Ejemplo Dimensión.....	19
Figura 4 Ejemplo Medidas.....	19
Figura 5 Ejemplo Jerarquía .....	20
Figura 6 Ejemplo Esquema DM .....	21
Figura 7 Ejemplo Esquema Estrella (Star Schema).....	22
Figura 8 Jerarquía Dimensión "Fecha" .....	29
Figura 9 Tabla Dimensión "Fecha" .....	30
Figura 10 Jerarquía Dimensión "Producto" .....	31
Figura 11 Tabla Dimensión "Producto" .....	32
Figura 12 Jerarquía dimensión "Cliente" .....	34
Figura 13 Tabla Dimensión "Cliente" .....	35
Figura 14 Tabla Dimensión "Venta" .....	37
Figura 15 Tabla de Hechos Final.....	39
Figura 16 Job Principal.....	41
Figura 17 Extracción.....	42
Figura 18 Job Transformación .....	43
Figura 19 Transformación "Producto" .....	44
Figura 20 Transformación Dimensión "Cliente" .....	45
Figura 21 Transformación Dimensión "Fecha" .....	46
Figura 22 Transformación Hecho "Venta" .....	47
Figura 23 Job Carga.....	48
Figura 24 Carga "Cliente" .....	49
Figura 25 Carga "Producto" .....	49
Figura 26 Carga "Fecha" .....	49
Figura 27 Carga "Venta" .....	50
Figura 28 Cubo OLAP Mondrian .....	51
Figura 29 Análisis OLAP Ventas por fechas y procedencia clientes.....	52
Figura 30 Venta por procedencia de los clientes.....	53
Figura 31 Productos con mayores ganancias.....	54
Figura 32 Comportamiento de las ventas durante un año.....	55
Figura 33 Cuadro de Mando Integral.....	57
Figura 39 Tabla Venta_Detalle.....	76
Figura 40 Tabla Rubro.....	77
Figura 41 Tabla Cliente .....	78
Figura 42 Tabla Familia .....	79
Figura 43 Tabla Marca.....	80
Figura 44 Tabla Producto .....	81
Figura 45 Logging Extracción.....	82
Figura 46 Step Metrics Extracción .....	82
Figura 47 Tabla Venta_Detalle.....	83
Figura 48 Tabla Rubro.....	84
Figura 49 Tabla Cliente .....	85
Figura 50 Tabla Familia .....	86

Figura 51 Tabla Marca.....	87
Figura 52 Tabla Producto .....	88
Figura 53 Archivo DPA_acotado.....	90
Figura 54 Logging Dimensión Cliente .....	90
Figura 55 Step Metrics Dimensión Cliente.....	91
Figura 56 Step Metrics Dimensión Producto.....	91
Figura 57 Logging Dimensión Fecha.....	92
Figura 58 Step Metrics Dimensión Fecha.....	92
Figura 59 Logging Hechos.....	92
Figura 60 Step Metrics Hechos.....	93
Figura 61 Dimensión Cliente.....	94
Figura 62 Dimensión Producto.....	95
Figura 63 Dimensión Fecha.....	95
Figura 64 Hechos Ventas.....	96
Figura 66 Logging Dimensión Cliente .....	98
Figura 67 Step Metrics Dimensión Cliente.....	98
Figura 68 Step Metrics Dimensión Producto.....	98
Figura 69 Logging Dimensión Producto .....	99
Figura 70 Logging Dimensión Fecha.....	99
Figura 71 Step Metrics Dimensión Fecha.....	99
Figura 72 Logging Hechos Ventas .....	99
Figura 73 Steps Metrics Hechos Ventas .....	100
Figura 74 Dimensión Cliente.....	100
Figura 75 Dimension Producto.....	101
Figura 76 Dimensión Fecha.....	101
Figura 77 Hechos Ventas.....	102
Figura 78 consulta SQL DM cantidad vendidas 2014 .....	104
Figura 79 Consulta cubo OLAP cantidad vendidas 2014 .....	104
Figura 80 Consulta SQL cantidad ventas por región .....	105
Figura 81 Consulta cubo OLAP cantidad de ventas por región.....	105
Figura 82 Directorio Pentaho .....	108
Figura 83 Directorio Biserver-CE ubicación archivo "start-pentaho.bat".....	108
Figura 84 Plataforma Pentaho BA Server de acceso y Credenciales de Acceso.....	109
Figura 85 Pantalla Principal Plataforma Pentaho BA Server.....	110
Figura 86 Pentaho Marketplace .....	111
Figura 87 Detalle Plugin CDF.....	111
Figura 88 CDTools para Cuadros de Mando – Creados por la Empresa Webdetails.....	112
Figura 89 Herramienta de Análisis Saiku .....	112
Figura 90 Pantalla Data Sources.....	113
Figura 91 Configuración DataSource JDBC.....	114
Figura 92 Import Analysis.....	115

---

## 1 INTRODUCCIÓN

---

La informática en los últimos años ha sido un área que ha ayudado en el surgimiento de las grandes organizaciones, donde le ha facilitado agilizar sus procesos en la toma de decisiones para el cumplimiento de sus objetivos. Pero ha traído un problema, que cada vez más los clientes de estas organizaciones demandan con mayor rapidez y eficiencia la entrega de productos o servicios, donde muchas veces estas organizaciones no han sido capaz de responder dentro de los tiempos acordados con estos requerimientos, por la gran cantidad de datos que almacenan sus sistemas. Lo que ha traído como consecuencia la emigración de estos clientes a la competencia. Para esto las organizaciones han tenido que implementar soluciones como Business Intelligence (BI) para poder responder de forma rápida y eficiente a sus obligaciones, además de facilitar la toma de decisiones de los altos mandos.

BI es la habilidad de transformar los datos en información y la información en conocimiento con el gran objetivo de facilitar la toma de decisiones de las organizaciones.

Pero BI en la actualidad es una exclusividad de las grandes organizaciones ya que para su implementación es necesario adquirir licencias de herramientas sofisticadas y profesionales con formación en esta área, que para una pequeña o mediana empresa (PYME) es prácticamente muy difícil que puedan implementarlo por su alto costo económico.

Nosotros en este proyecto de tesis implementaremos BI en una PYME, ya que conforman el mayor universo de empresas en el país donde dan empleo alrededor del 80% de los trabajadores. La PYME, donde realizamos nuestro proyecto se llama VALFI y es empresa creada en el año 2006 que se dedica a la venta de productos eléctricos e insumos computaciones, abasteciendo al público de la provincia de Ñuble, especialmente de la ciudad de Chillán.

El área de análisis que realizamos nuestro trabajo fue Ventas Clientes de la empresa VALFI. Para la implementación de nuestra solución se utilizaron herramientas OpenSource, ya que la empresa no tenía presupuesto para adquirir una licencia de pago. Con el gran objetivo de generar cuadros de mandos integrales o

dashboard, para poder ver en forma de gráficos el nivel de cumplimientos de sus objetivos. La implementación de BI se realizó a través de la creación de un DataMart (DM) usando como fuente su base de datos transaccional, además aplicamos la metodología Kimball para el proceso de extracción, transformación y carga de los datos hasta su análisis.

## 1.1 Objetivo General

Implementar Business Intelligence para el área de ventas clientes, creando un DataMart (DM) en una PYME local de rubro eléctrico.

## 1.2 Objetivos Específicos

- Explotación de datos transaccionales del área de ventas clientes de la organización, transformándolos en información para la toma de decisiones estratégicas.
- Determinar las necesidades de información (Requerimientos de la Empresa).
- Aplicar una metodología de BI.
- Utilización de un DataMart (DM) usando herramientas OpenSource para el proceso ETL, la creación de cubos OLAP para la generación de informes y BalancedScoreCard.

## 1.3 Alcance y Límites del Proyecto

Se realizará un proyecto de BI solamente en el área de ventas clientes de VALFI debido al tiempo limitado que contamos para desarrollar este proyecto, para esto se aplicará la metodología Kimball para esta solución que tiene como objetivo la construcción de un DataMart, utilizando como fuente la base de datos transaccional de VALFI. Además se utilizarán herramientas OpenSource desde el proceso de extracción, transformación, carga de los datos, el diseño del cubo OLAP, finalmente la construcción de los informes y el BalancedScoreCard para facilitar la toma de decisiones.

## 1.4 Problemática de la Empresa VALFI

La problemática de la empresa VALFI, es que la actualidad cuenta con base de datos que es consumida por una aplicación de escritorio en visual basic el cual comenzó con pocos módulos, pero a medida que las necesidades de la empresa fueron creciendo también lo hizo el sistema, que trajo como consecuencia la lentitud en el sistema la cuál afectaba a todos los usuarios dentro del sistema. Además en el ámbito de los reportes no cuentan con nivel de detalle de sus productos y sus clientes.

En la primera entrevista con los gerentes de la empresa, nos comentaron que deseaban migrar su sistema a un sistema web, pero tenían la incertidumbre si aun con la

migración, el nuevo sistema al momento de generar reportes, seguiría afectando el rendimiento del sistema a nivel global para todos los usuarios, (problemática que se venía suscitando), a lo cual se les comento, que para motivos de análisis y reportes que no afectaran el rendimiento de la base de datos transaccional, una opción era realizar un DataWarehouse que en un comienzo se constituiría de un DataMart del área de Ventas, la cual al ser una fuente independiente de la base de datos transaccional no afectaría su rendimiento.

## **1.5 Organización del Informe**

EL resto del informe está organizado de la siguiente manera. En el capítulo 2 hacemos una descripción de la empresa VALFI, detallando su misión, visión y al rubro que pertenece. El capítulo 3 está destinado a definir conceptos aplicados en el diseño, análisis e implementación de nuestra solución. En el capítulo 4 está destinado a explicar cómo se realizó toda la implementación de nuestra solución, desde el proceso de extracción, transformación y carga de los datos hasta el análisis OLAP. En el capítulo 5 se explicará los tipos de pruebas que se realizaron en este proyecto para asegurar su calidad. Finalmente, en el capítulo 6 se darán a conocer todas las conclusiones y experiencias rescatadas en este proyecto y el trabajo futuro que se puede realizar.

---

## 2 DESCRIPCIÓN DE LA EMPRESA

---

En este capítulo se dará a conocer datos de la empresa VALFI donde realizamos nuestro trabajo de implementación de Business Intelligence.

### 2.1 Identificación de la empresa

VALFI nace en el año 2006 publicada en el diario Oficial, como **Sociedad Comercial Valdés y Figueroa Limitada**, formada por dos emprendedores de reconocida experiencia en el área de Ventas de Productos Eléctricos y Asesoría Técnica[1].

### 2.2 Logo de la Empresa



**Figura 1 Imagen Corporativa VALFI**

El logo de VALFI es un guepardo o más conocido como chita, que es sinónimo de agilidad y rapidez en la entrega de soluciones integrales a todos sus clientes. Esta rapidez ha permitido a VALFI establecer un fuerte lazo comercial entre la empresa y sus clientes.

## **2.3 Políticas de la Empresa**

A continuación se presentarán las políticas de la empresa VALFI, la Visión y Misión han sido obtenidas de la página web de la empresa.

### **2.3.1 Visión de la Empresa VALFI**

“Potenciar y fortalecer nuestra empresa comercializadora, para así obtener reconocimiento por la venta y asesoría técnica de materiales eléctricos y productos computacionales.

Capacitar de forma constante a nuestro personal técnico y administrativo, desarrollando eficiencia para cumplir con los requerimientos exigidos por nuestros clientes y la empresa.

Posicionar la marca VALFI ELECTRICIDAD en el mercado como:

La Tienda Comercial que además de vender materiales eléctricos y productos computacionales es capaz de entregar el Mix de soluciones que su empresa necesita”.

### **2.3.2 Misión de la Empresa VALFI**

“Atender a nuestros clientes como personas y solucionar en el más corto plazo posible sus requerimientos.”

## **2.4 VALFI Electricidad**

Es una empresa que comercializa productos eléctricos y computacionales de las más reconocidas Marcas: Lexo, Lovato, Vkb, Delixi, legrand, Bticino, Chint, Fujitel, Gewiss, Visual Light, Hyundai, Lg, Castillo, Tigre, Saime, Covisa, Osram, Duracell, Drl, Philips, Sacco, Roker, Tp-Link, Black Box etc. y así abastece a contratistas eléctricos, establecimientos educacionales, hospitales, constructoras, ingenieros, técnicos y público en general.

Además VALFI cuenta con una bodega donde se maneja un completo stock de tuberías, conduit, cañerías galvanizadas, luminarias, reflectores, cámaras, cables, cordones y una variada gama de conductores eléctricos.

---

## 3 CONCEPTOS

---

En este capítulo mencionaremos y definiremos todos los conceptos que utilizaremos en la implementación de Business Intelligence para el área de Ventas Clientes de la empresa VALFI.

Estos conceptos se aplicarán tanto en el diseño de la solución, como el proceso de extracción transformación y carga de los datos, análisis y finalmente en el proceso de pruebas.

### 3.1 Business Intelligence (BI)

BI es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en las organizaciones [2]. Las condiciones ideales para que una organización necesite aplicar BI, es cuando no se tiene claro si alcanzaron o cumplieron sus objetivos, o si dedican mucho tiempo en el análisis de datos dentro de sus departamentos. Para aplicar BI es necesario conocer dos conceptos que son muy importantes y que sirven de partida para su implementación, como son DataWarehouse (DW) y DataMart (DM).

### 3.2 DataWarehouse (DW) y DataMart (DM)

Estos dos conceptos se traducen como almacén o repositorio de datos, en el cual residen datos provenientes de una o varias fuentes y tienen como principal característica que los datos existen para ser leídos y no modificados manteniendo un histórico de los datos. Además se organizan en temas para facilitar su acceso o su entendimiento para adecuarse a las necesidades de los usuarios y tienen como gran objetivo el análisis de datos a través de variadas herramientas de software.

La gran diferencia que podemos encontrar en estos dos conceptos es que el DW involucra datos de todos los departamentos o áreas de la organización (general), mientras que el DM es solamente de un área o departamento, lo que sería un DW departamental (específico).

### 3.3 Metodología Kimball

Antes de realizar cualquier estudio o implementación de BI es necesario aplicar una metodología. En este tema existe una gran cantidad de literatura, pero para nuestro trabajo aplicamos solamente una metodología que es muy conocida en el área de BI, esta metodología se conoce como Kimball, que fue creada por Ralph Kimball.

Esta metodología la escogimos para realizar nuestra solución de implementar BI en VALFI por una de sus principales características, que es la creación de un DM en su primera fase, pues el sistema que actualmente poseen, sólo registra información de ventas y genera reportes estáticos, lo que se adecua a nuestro principal requerimiento que es la aplicación de BI a solamente un área, Ventas Clientes.

Para crear el DM, es necesario identificar bien cuáles serán las fuentes de los datos y la principal estructura que esta tendrá, su tabla de hechos que a su vez esta tiene los siguientes componentes: hecho principal, las dimensiones y jerarquías.

Una gran ventaja que tiene esta metodología es que sirve para realizar trabajos futuros dentro del mismo proyecto, como por ejemplo realizar BI en otra área de la organización. Para esto será necesario construir otro DM que puede ser el área de producción identificando todos sus componentes, en donde si es necesario se puede utilizar por ejemplo dimensiones y jerarquías de otros DM ya implementados. De esta forma la unión de todos los DM de las áreas de la organización va ser igual al DW.

Esta metodología tiene los siguientes pasos, ver Figura 2:



**Figura 2 Pasos Metodología Kimball**

### **3.4 Conceptos Utilizados en el Diseño de la Solución**

Para comenzar con el diseño de un DW o DM que utilizaremos en nuestro trabajo, es necesario conocer algunos conceptos que nos permitirán construir nuestra solución como son dimensiones, jerarquías y la tabla de hechos.

#### **3.4.1 Dimensión**

Una dimensión corresponde a una o más categorías que se desea incluir en el DW o DM, por ejemplo, “Fecha”, ver Figura 3.

Fecha
año mes_nombre semestre trimestre semana dia_nombre dia fecha

Figura 3 Ejemplo Dimensión

### 3.4.2 Medidas

Son valores concretos asociados a una tabla de hechos, estos valores dependerán de las reglas de negocio. Ejemplo “**monto\_venta**”, “**costo\_venta**”, “**ganacia\_venta**” y “**cantidad**”, ver Figura 3.

Venta	
PK	id_venta
PK,FK1	id_fecha
PK,FK2	id_producto
PK,FK3	id_lugar_cliente
	monto_venta costo_venta ganacia_venta cantidad

Figura 4 Ejemplo Medidas

### 3.4.3 Jerarquía

Representa la especificación de los niveles que relacionan los diferentes atributos dentro de una dimensión, ver Figura 5.



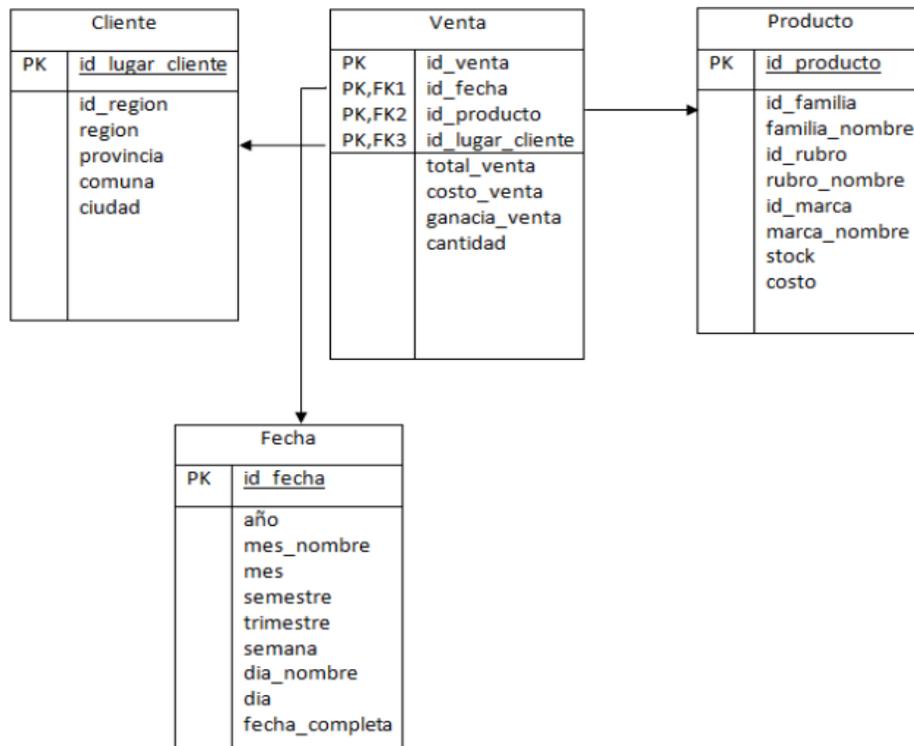
Figura 5 Ejemplo Jerarquía

### 3.4.4 Tabla de Hechos

Es un esquema que contiene los valores de las medidas del negocio. En la Figura 6 se puede observar la tabla de hecho principal con el nombre “Venta”, la cual

tiene en su clave primaria llamada “**id\_venta**”. Además, tiene tres claves foráneas “**id\_fecha**”, “**id\_producto**”, “**id\_lugar\_cliente**” que sirven para unir a la tabla principal con las dimensiones “Cliente”, “Fecha” y “Producto” las que categorizan las ventas con las medidas “**total\_venta**”, “**costo\_venta**”, “**ganacia\_venta**” y “**cantidad**”.

Finalmente con este esquema se puede obtener para una venta, la fecha que se realizó, el producto que se vendió, la procedencia del cliente y el monto total de la venta.



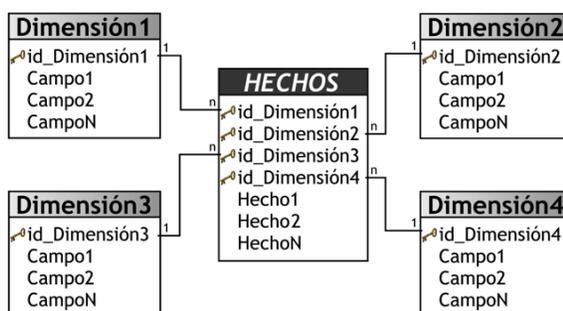
**Figura 6 Ejemplo Esquema DM**

La tabla de hechos, según su diseño además se puede clasificar según su esquema, por ejemplo esquema estrella (StarSchema), o copo de nieve (SnowflakeSchema).

### 3.4.5 Esquema Estrella (StarSchema)

El esquema estrella está compuesto por una tabla de hecho central y un conjunto de dimensiones alrededor asemejando su composición a la de una estrella.

Las dimensiones sólo tienen conexión a la tabla de hechos y a ninguna más; al contrario de copo de nieve que tiene dimensiones enlazadas con otras dimensiones.



**Figura 7 Ejemplo Esquema Estrella (Star Schema)**

### 3.4.6 Extracción, Transformación y Carga (ETL)

Al tener ya diseñado nuestra tabla de hechos, tenemos que pasar a nuestro siguiente paso que será realizar el proceso de Extracción, Transformación y Carga (ETL) de los datos, desde una o más fuentes de origen hasta un DM o DW. Comúnmente las fuentes que se utilizan son bases de datos transaccionales (OLTP) donde una de sus características es que estas son orientadas al procesamiento, esto quiere decir que involucran operaciones de inserción, modificación y eliminación de los datos, almacenan datos actuales y tienen un elevado número de transacciones. A diferencia del DM o DW que son bases de datos orientados al análisis (OLAP), en donde tienen un bajo número de transacciones y almacenan datos históricos que solamente son para la lectura y análisis.

Con respecto al proceso ETL, se tienen las siguientes etapas:

- Extracción: Obtención de información de las distintas fuentes tanto internas como externas [3].
- Transformación: Filtrado, limpieza, depuración y agrupación de la información [3].
- Carga: Organización y actualización de los datos y los metadatos en la base de datos DW o DM [3].

### 3.4.7 Herramientas ETL

Para realizar este proceso es necesario ocupar una o más herramientas ETL. En el área de BI existen una gran variedad de estas herramientas de las cuales podemos encontrar aplicaciones con licencia como también OpenSource, como por

ejemplo Pentaho Data Integration (Kettle) [4], Talend [5], Apatar [6], SQLServerIntegrationServices (SSIS) [7], etc.

La herramienta escogida para nuestra implementación de BI es Pentaho Data Integration (Kettle), ya que esta aplicación cumple otro principal requerimiento de nuestra solución y de la empresa, pues la gerencia no puede costear una solución de pago, y nuestra herramienta a utilizar tiene que ser OpenSource. Además es compatible con las bases de datos que utilizaremos como fuente y destino de datos SQL Server. Otro elemento que nos ayudó a elegir Pentaho data integration fue nuestra experiencia previa en la herramienta.

#### **3.4.8 Pentaho Data Integration (Kettle)**

Es una herramienta de Extracción, Transformación y Carga, (ETL), la cual permite, a través de una interfaz gráfica y a un conjunto de herramientas, realizar la implementación de un DW o DM de una manera sencilla.

Pentaho está conformado por un conjunto de herramientas, de las cuales son las que proveen las funcionalidades de Job y transformaciones que nos darán las facultades de realizar procesos de ETL.

#### **3.4.9 Cuadro comparativo de herramientas ETL**

En la Tabla 1, se muestra un cuadro comparativo de las distintas herramientas ETL que existen en el mercado, en donde se compararán con distintos atributos como

tipo (pagada o gratuita), SO cuales son los Sistemas Operativos que la soporta y BD las bases de datos que son compatibles con cada herramienta.

Herramienta	Tipo	SO	BD
<b>Pentaho Data Integration (Kettle)</b>	<b>Pagado, Open Source</b>	<ul style="list-style-type: none"> <li>• <b>Windows</b></li> <li>• <b>Macintosh</b></li> <li>• <b>Linux</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>MySQL</b></li> <li>• <b>Microsoft SQL Server</b></li> <li>• <b>Oracle</b></li> <li>• <b>Otras 32 Bases de datos.</b></li> </ul>
Talend	Pagado, Open Source	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Unix</li> <li>• Linux</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> </ul>
Apatar	Open Source	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Linux</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> </ul>
SQL Server Integration Services (SSIS).	Pagado	<ul style="list-style-type: none"> <li>• Windows</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> </ul>
PALO	Open Source	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Linux</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> </ul>

**Tabla 1 COMPARACIÓN HERRAMIENTAS ETL**

### 3.4.10 Cubo OLAP

Luego que se tiene implementada nuestro DM, tenemos que seguir con nuestro siguiente paso, que es el análisis de los datos. Para esto se tiene implementar el Cubo OLAP, que es una base de datos Multidimensional, que tiene los mismos componentes de nuestra tabla de hechos, dimensiones y jerarquías.

Para realizar el análisis de datos, se tiene que utilizar otro tipo de herramientas para la creación de un Cubo OLAP. Para nuestra solución utilizaremos Mondrian, el cual es un motor para la creación de Cubos OLAP. A través de una herramienta de interfaz denominado Pentaho Schema Workbench.

Luego de nuestro proceso ETL, creación del Cubo OLAP se debe pasar a una etapa de análisis y presentación del DM, que se denominan Análisis OLAP (Business Analytics) y Cuadro de Mando (Dashboard/ BalancedScoreCard) los cuales se definirán a continuación:

#### **3.4.11 Análisis OLAP**

Como su nombre lo indica, es el análisis de los datos que contiene un Cubo OLAP en el cual se puede ver desde varios puntos de vistas (dimensiones) un conjunto de valores (medidas), dando la posibilidad de ver estos datos en distintos niveles de detalles (jerarquías).

Es la visualización del cubo dimensional del DM, presentando los datos de una manera mucho más rápida, pues solo se realizan consultas sobre los datos que tenga el cubo OLAP, estos análisis son muy utilizados en áreas de marketing y ventas e informes.

#### **3.4.12 Cuadro de Mando (Dashboard/ BalancedScoreCard).**

Es una herramienta de control empresarial que permite establecer y monitorizar los objetivos de una organización y de sus diferentes áreas o departamentos.

También se puede considerar como una aplicación que ayuda a una organización a expresar los objetivos e iniciativas necesarias para cumplir con su

estrategia, mostrando de forma continuada cuándo la empresa y los empleados alcanzan los resultados definidos en su plan estratégico [8].

Para nuestra solución utilizaremos las siguientes herramientas. Para análisis utilizaremos la herramienta Opensource Saiku, un potente visor OLAP de análisis. Luego para Dashboard utilizaremos la herramienta Opensource CDF (Community Dashboard Framework), un marco de trabajo para la realización de Dashboard.

Estas herramientas escogidas estarán contenidas en la herramienta Pentaho Business Analytics, el cual es un servidor de aplicaciones donde estas herramientas tienen que ser ejecutadas (se explicará proceso en solución).

### 3.4.13 Cuadro comparativo de herramientas de análisis.

En la Tabla 2, muestra un cuadro comparativo de las distintas herramientas de análisis que existen en el mercado, en donde se compararán con distintos atributos como Tipo en donde mencionara sí la herramienta es de paga o gratuita, SO cuales son los Sistemas Operativos que la soporta y BD las bases de datos que son compatibles con cada herramienta.

Herramienta	Tipo	SO	BD
<b>Mondrian</b>	<b>Pagado, Open Source</b>	<ul style="list-style-type: none"> <li>• <b>Windows</b></li> <li>• <b>Linux</b></li> <li>• <b>Unix</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Bases de Datos Relacionales (MySQL, Oracle, PostgreSQL, Microsoft SQL Server, etc)</b></li> </ul>
Jpivot	Pagado, Open Source	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Linux</li> <li>• Mac OS</li> <li>• BSD</li> <li>• Solaris</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> <li>• Otras 32 Bases de datos.</li> </ul>
<b>Saiku</b>	<b>Pagado, Open Source</b>	<ul style="list-style-type: none"> <li>• <b>Windows, Linux y Unix</b></li> </ul>	<ul style="list-style-type: none"> <li>• <b>MySQL</b></li> <li>• <b>Microsoft SQL Server</b></li> <li>• <b>Oracle</b></li> <li>• <b>Otras 32 Bases de datos.</b></li> </ul>

PALO	Pagado, Open Source	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Linux</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> <li>• Otras 32 Bases de datos.</li> </ul>
Spago BI	Pagado, Open Source	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Linux</li> <li>• Mac OS</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> <li>• Otras 32 Bases de datos.</li> </ul>
Pentaho	Pagado	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Linux</li> <li>• Mac OS</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> <li>• Otras 32 Bases de datos.</li> </ul>
JarperSorft	Pagado	<ul style="list-style-type: none"> <li>• Windows</li> <li>• Linux</li> <li>• Mac OS</li> <li>• Unix</li> </ul>	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• Microsoft SQL Server</li> <li>• Oracle</li> <li>• Otras 32 Bases de datos.</li> </ul>

**Tabla 2 Comparativo Herramientas de Análisis**

---

## 4 SOLUCIÓN

---

En este capítulo se realizará la implementación de nuestra solución BI para la empresa VALFI en el área de ventas cliente aplicando la metodología Kimball, con el

objetivo de implementar un DM para la generación de un Dashboard. Se comenzará definiendo los componentes del DM que son: dimensiones, jerarquías, hechos y medidas.

Con la solución ya diseñada, se realizará el proceso de Extracción, Transformación y Carga (ETL) desde la base de datos transaccional de origen, hacia un StageArea, y finalmente a un DM, para luego implementar nuestro cubo OLAP con las dimensiones y medidas.

Finalmente, se presentará, como a través del cubo OLAP creado, se realizará el Cuadro de Mando Integral (Dashboard/BalancedScoreCard), y el análisis dimensional (Análisis OLAP), que es lo que observará el usuario final de la empresa VALFI para facilitar la toma de decisiones.

#### **4.1 Definición de Requerimientos**

- Información de las ventas (cantidad, costo, monto y la ganancia) por la procedencia de sus clientes.
- Información sobre las cantidades de productos vendidos por regiones y periodo de tiempo anual.
- Información de la relación de los productos categorizados (Familia), con respecto a costos, montos y ganancias.
- Tener una vista resumen (ventas por región, ventas por provincias, ventas por comunas, ventas por periodo mensual) en un período anual determinado.

#### **4.2 Modelo de Dimensiones, Jerarquías y Hechos**

##### **4.2.1 Dimensiones**

Para nuestra solución BI, se han diseñado dimensiones las cuales son necesarias en el DM, para este trabajo tras analizar el negocio, la empresa y cómo se

abordará la solución, se han escogido los siguientes aspectos: “Cuándo” se realizó la venta, “Qué” fue vendido, “Procedencia” del cliente que realizó la compra.

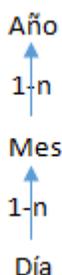
Para responder esas preguntas, se han diseñado 3 dimensiones, los cuales a través de jerarquías, entregan una solución a estas dudas, que debe responder el DM.

- “Producto”.
- “Cliente”.
- “Fecha”.

La dimensión “Fecha”, nos entregará información de “Cuándo” se realizó una venta, la dimensión “Producto”, nos entregará información de “Qué” producto fue vendido, y la dimensión “Cliente”, nos entregará información de la “Procedencia” del cliente efectuó la compra.

#### 4.2.2 Tabla Dimensión “Fecha”

Para esta dimensión “Fecha” tendrá la siguiente jerarquía, ver Figura 8.



**Figura 8 Jerarquía Dimensión "Fecha"**

En la Figura 8 se define la jerarquía de la dimensión “Fecha”, donde se definen las siguientes agregaciones:

- Un día de la semana pertenece solo a una semana del mes. Una semana del mes tiene uno o más días de la semana.
- Una semana del año pertenece solo a un mes. Un mes tiene una o más semanas.
- Un mes del año pertenece solo a un año. Un año tiene uno o más meses del año.

La Tabla Dimensión “Fecha” quedará de la siguiente forma, ver Figura 8.

Fecha	
PK	<u>id_fecha</u>
	anio mes_nombre mes dia_nombre dia

**Figura 9 Tabla Dimensión "Fecha"**

- La nueva tabla de dimensión tendrá el nombre “Fecha”.
- Se agregará una clave primaria con el nombre “id\_fecha”, que se utilizará para unirla con el hecho principal “Venta”.
- Utilizando como Base la metodología de Kimball, aplicándolo a nuestro proyecto, se agregarán los siguientes atributos para lograr la jerarquía presentada en la Figura 8, los atributos tendrán los siguientes nombres:
  - **“anio”**: Es el año que pertenece a la fecha.
  - **“mes\_nombre”**: Es el nombre del mes que pertenece a la fecha.
  - **“mes”**: Es el identificador del mes que pertenece a la fecha.
  - **“dia\_nombre”**: Es el nombre del día de la semana que pertenece a la fecha.
  - **“dia”**: El día del mes que pertenece a la fecha.

La dimensión “Fecha” está definida por los siguientes atributos, ver Tabla 3.

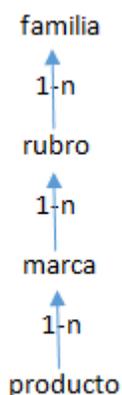
Nombre Atributo	Tipo	Ejemplo
id_fecha	Datetime	“1995-07-05” (año-mes-día)
Anio	Int	1995
mes_nombre	text	“Julio”

Mes	Int	7
dia_nombre	text	"Martes"
Dia	Int	4

**Tabla 3 Atributos Dimensión "Fecha"**

#### 4.2.3 Tabla Dimensión "Producto"

Para esta dimensión "Producto" tendrá la siguiente jerarquía, ver Figura 10.



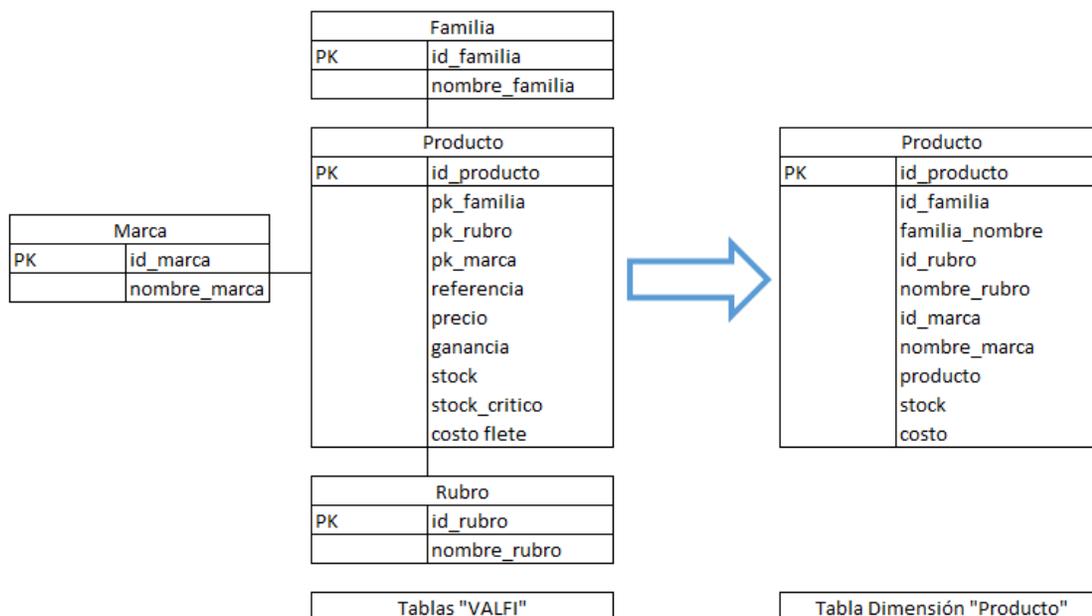
**Figura 10 Jerarquía Dimensión "Producto"**

En la Figura 10 se define la jerarquía de la dimensión "Producto", en donde se definen las siguientes agregaciones:

- Un producto en especial pertenece solo a una marca. Una marca puede tener uno o más productos.
- Una marca en especial pertenece solo a un rubro. Un rubro puede tener una o más marcas.
- Un rubro en especial pertenece solo a una familia. Y una familia puede tener una o más categorías.

La Tabla Dimensión "Producto" quedará de la siguiente forma ver Figura 11, como se observa en la figura, se presenta la tabla producto con sus respectivas relaciones que posee VALFI en su base de datos (izquierda) y la tabla Dimensión

Producto que tendrá el DM (derecha).



**Figura 11 Tabla Dimensión "Producto"**

- La nueva tabla de dimensión tendrá el nombre "Producto".
- Se mantendrá la clave primaria con el nombre "id\_producto", que se utilizará para unirla con el hecho principal "Venta".
- Utilizando como Base la metodología de Kimball, aplicándolo a nuestro proyecto, para la realización de la dimensión producto en base a la realidad de la empresa VALFI, Se eliminarán los siguientes atributos "ganancia", "stock\_critico", "costo\_flete", el atributo "referencia" se renombrará por "producto" y ganancia por "costo".
- Se tendrá los siguientes atributos:
  - "id\_familia": Es el identificador de la familia que pertenece el producto.
  - "familia\_nombre": Es el nombre de una serie de productos de una misma necesidad, pero son diferentes.
  - "id\_rubro": Es el identificador del rubro que pertenece el producto.
  - "rubro\_nombre": Es el nombre genérico de un producto.
  - "id\_marca": Es el identificador de la marca que pertenece el producto.
  - "marca\_nombre": Es el nombre de la marca que pertenece un producto.
  - "producto": Es el nombre del producto.
  - "stock": Es la cantidad de productos disponibles para una venta.

- “costo”: Es el precio del producto.  $\text{costo} = \text{precio} - \text{costo\_flete}$

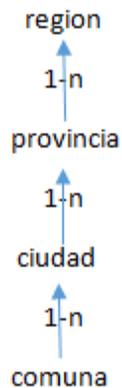
La dimensión “Producto” está definida por los siguientes atributos, ver Tabla 4.

Nombre Atributo	Tipo	Ejemplo
id_producto	int	1
id_familia	int	1
familia_nombre	text	“Productos Domésticos”
id_rubro	int	1
rubro_nombre	text	“CAJAS (TABLEROS)”
id_marca	int	1
marca_nombre	text	“Legrand”
Producto	text	“01512 23 EMB. 26 CTOS. PTA BLANCA LEGRAND”
Stock	int	200
Costo	int	10000

**Tabla 4 Atributos Dimensión “Producto”**

#### 4.2.4 Tabla dimensión “Cliente”.

Para esta dimensión “Cliente” tendrá la siguiente jerarquía, ver Figura 12.

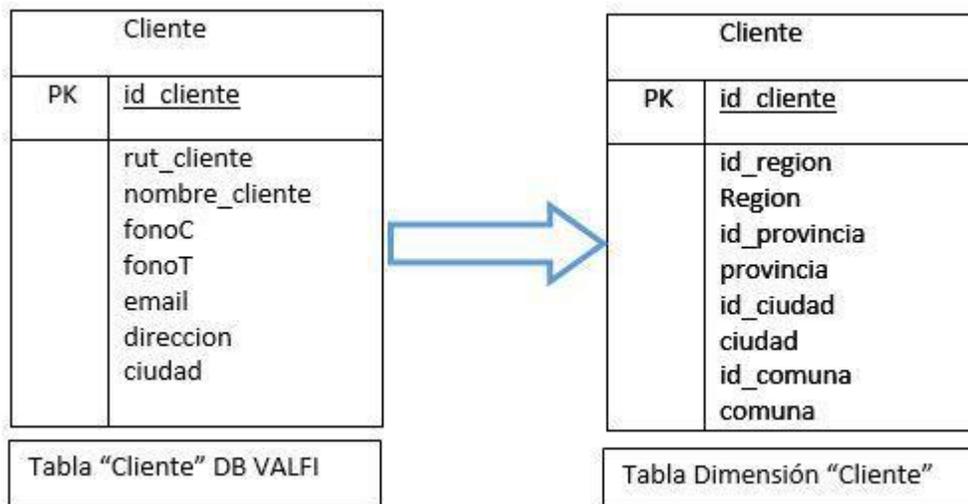


**Figura 12 Jerarquía dimensión "Cliente"**

En la Figura 12 se define la jerarquía de la dimensión "Cliente", en donde se definen las siguientes agregaciones:

- Una comuna en especial pertenece solo a una ciudad. Una ciudad puede tener una o más comunas.
- Una ciudad en especial pertenece solo a una provincia. Una provincia puede tener una o más ciudades.
- Una provincia en especial pertenece solo a una región. Y una región puede tener una o más provincias.

La Tabla Dimensión “Cliente” quedará de la siguiente forma ver Figura 13, como se observa en la figura, se presenta la tabla cliente que posee VALFI en su base de datos (izquierda) y la tabla Dimensión Cliente que tendrá el DM (derecha).



**Figura 13 Tabla Dimensión "Cliente"**

- La nueva tabla de dimensión tendrá el nombre “Cliente”.
- Se agregará una clave principal con el nombre “id\_cliente”.
- Utilizando como Base la metodología de Kimball, aplicándolo a nuestro proyecto, para la realización de la dimensión Cliente en base a la realidad de la empresa VALFI, para cumplir con el esquema definido en la figura 12 de la dimensión “Cliente” y con la jerarquía presentada en la figura 11, se eliminarán los siguientes atributos que no influyen en el desarrollo del proyecto para esta dimensión, los atributos son “rut\_cliente”, “nombre\_cliente”, “fonoT”, “fonoC”, “email”, “direccion”, que no son considerados, la dimensión representará la procedencia de los clientes de VALFI.
- Se mantendrá el atributo “ciudad”.
- Tendrá los siguientes atributos
  - **“id\_region”**: Es el identificador de la región en la que pertenece el cliente.
  - **“region”**: Es el nombre de la región a la que pertenece el cliente.
  - **“provincia”**: Es el nombre de la provincia a la que pertenece el cliente.
  - **“comuna”**: Es el nombre de la comuna a la que pertenece el cliente.
  - **“ciudad”**: Es el nombre de la ciudad a la que pertenece el cliente.

La dimensión “Cliente” está definida por los siguientes atributos, ver Tabla 5.

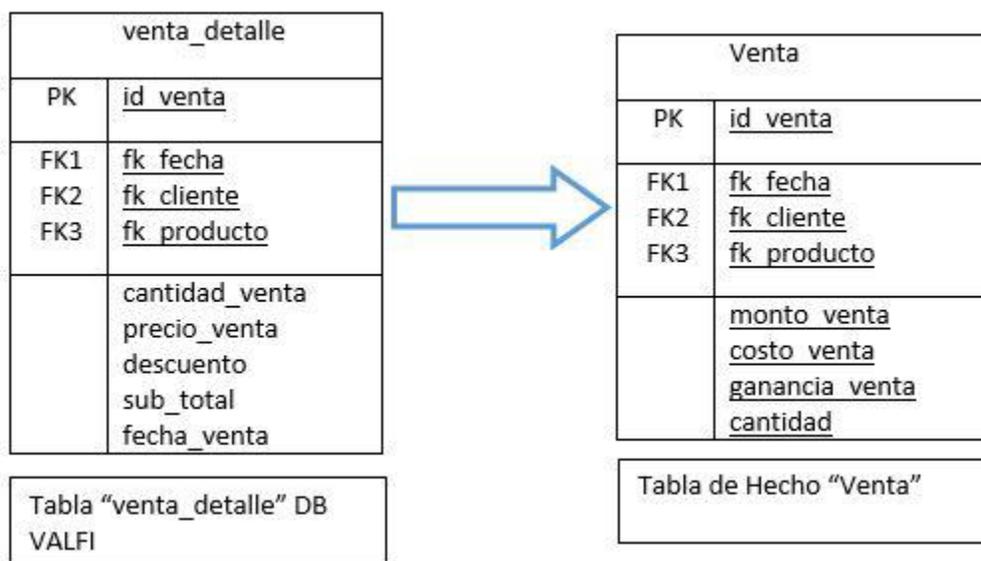
Nombre Atributo	Tipo	Ejemplo
id_cliente	int	1
Region	text	8
provincia	text	"Ñuble"
Comuna	text	"Chillán"
Ciudad	text	"Chillán"

**Tabla 5 Atributos Dimensión "Cliente"**

#### 4.2.5 Tabla de Hechos

Para la solución de este problema se diseñó un hecho principal "Venta". Además esta tabla está diseñada para poder dar respuestas a preguntas que la empresa VALFI necesita para tomar mejores decisiones.

El hecho principal "Venta" quedará de la siguiente forma ver Figura 14, como se observa en la figura, se presenta la tabla "venta\_detalle" que posee VALFI en su base de datos (izquierda) y la tabla de Hechos que tendrá el DM (derecha).



**Figura 14 Tabla Dimensión "Venta"**

- El nuevo Hecho principal tendrá el nombre "Venta".
- Se agregará una clave principal con el nombre "id\_venta".
- Basándonos en la metodología de Kimball, y para lograr lo definido en la Figura 14, se agregará las siguientes claves foráneas "id\_fecha", "id\_producto" y "id\_cliente", para unirlas con las dimensiones "Producto", "Cliente" y "Fecha".
- El hecho "Venta" tendrá las siguientes medidas:
  - **"monto\_venta"**: Es el valor de la venta.
  - **"costo\_venta"**: Es el costo del producto que se vende.
  - **"ganancia\_venta"**: Es la ganancia total de la venta.  
**"ganancia\_venta"="monto\_venta" - "costo\_venta"**.
  - **"cantidad"**: Es la cantidad de producto(s) que se vendió.

El hecho principal "Venta" está definida por los siguientes atributos, ver Tabla 6:

Nombre Atributo	Tipo	Ejemplo
id_venta	int	1
id_fecha	Datetime	"2014-05-05"
id_producto	int	1
id_cliente	int	1
monto_venta	decimal(19,4)	10000
costo_venta	decimal(19,4)	5000
ganancia_venta	decimal(19,4)	5000
cantidad	int	1

**Tabla 6 Atributos Hecho Principal "Venta"**

#### 4.2.6 Tabla de Hecho Final

En la Figura 15, se muestra el diseño de la solución final, donde se aprecia en el centro el hecho principal llamado “Venta” y a sus lados las dimensiones “Cliente”, “Producto” y “Fecha”. Estas tablas están unidas a través de las claves primarias de las respectivas de las dimensiones (PK) con las claves foráneas del hecho principal (“FK1”, “FK2” y “FK3”).

Para esta solución se aplicó el esquema estrella para facilitar la complejidad y la velocidad de las consultas SQL.

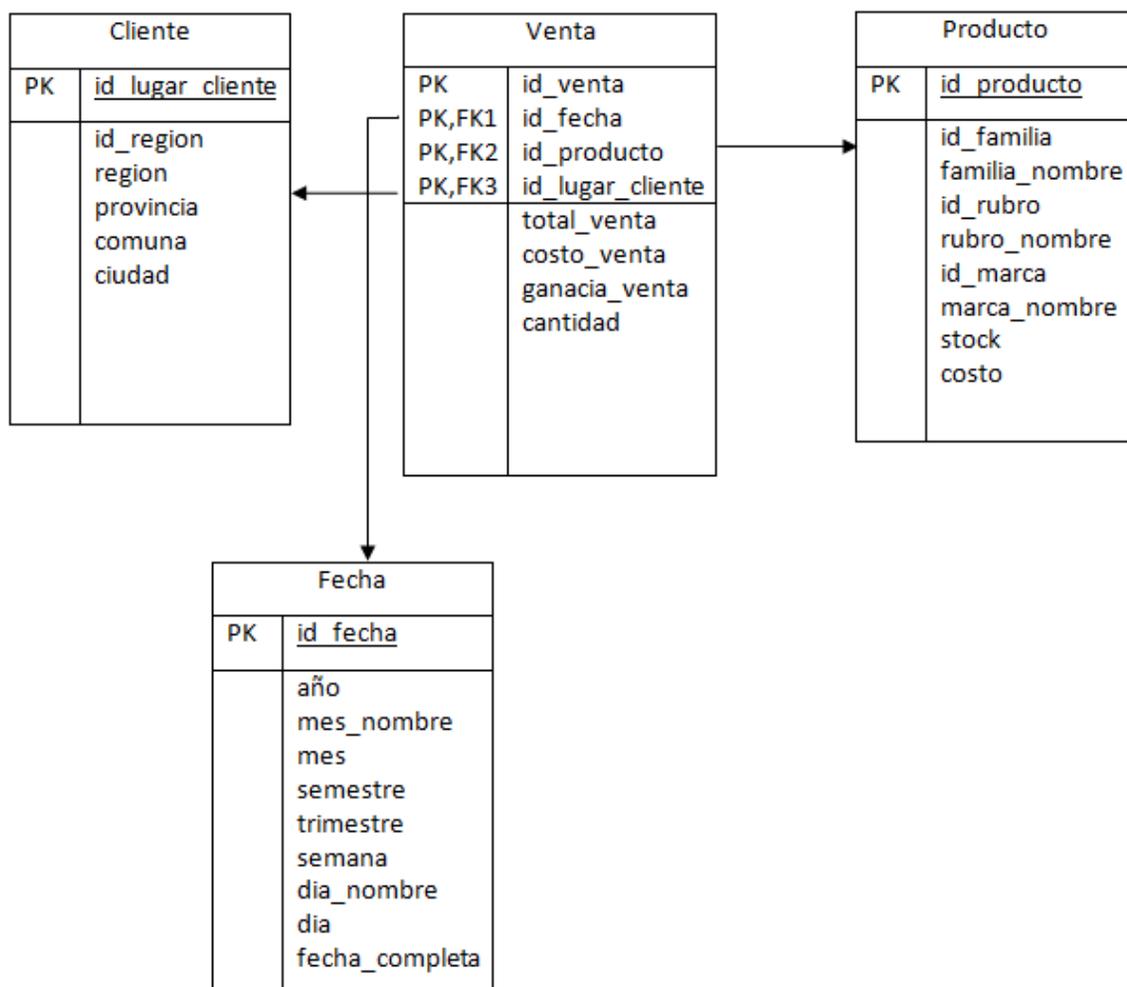


Figura 15 Tabla de Hechos Final

### **4.3 Proceso Extracción, Transformación y Carga (ETL)**

Después de haber diseñado la solución (Tabla de Hechos) de la implementación de BI para el área de venta clientes de VALFI, es necesario seguir con el siguiente paso, el proceso ETL.

En la implementación de esta solución, se utilizaron datos de muestra los cuales son lo más cercano a la realidad de la empresa. Estos datos fueron tabulados en un archivo Excel, los que fueron traspasados a una base de datos (SQL Server) para simular la fuente de datos OLTP y realizar el proceso ETL para cargarlo a un DM.

Para este proceso se utilizará la Suite Pentaho, que consta de las herramientas para realizar el proceso ETL, desde nuestra fuente de datos origen hasta nuestro DM siguiendo el esquema definido en 4.1. y 4.2.

### 4.3.1 Job Principal

Para realizar el proceso ETL es necesario crear un Job Principal, el que se encargará de ejecutar este proceso.

Pentaho Data Integration, se basa en “Steps” o “Pasos”, que comienza con un paso “Start” y finaliza con un “Success” al ejecutarse correctamente, en el caso contrario se ejecutan los pasos “Write to Log” que escriben en el log de la herramienta, y finalizando en el paso “Abort Job”. Además los pasos son conectados a través de flechas con dirección, que indican el flujo que realiza. En la Figura 16 se visualizan tres tipos de flujos, el flujo verde que se ejecuta si el paso tiene éxito, el rojo en caso que el paso falle y los flujos con el símbolo de candado son pasos obligatorios que se realizarán siempre. Los pasos “Extracción”, “Transformación”, “Carga” son Jobs que ejecutan las transformaciones específicas. Para ver más detalle del Job Principal, ver 8 Anexo: Detalle Job Principal.

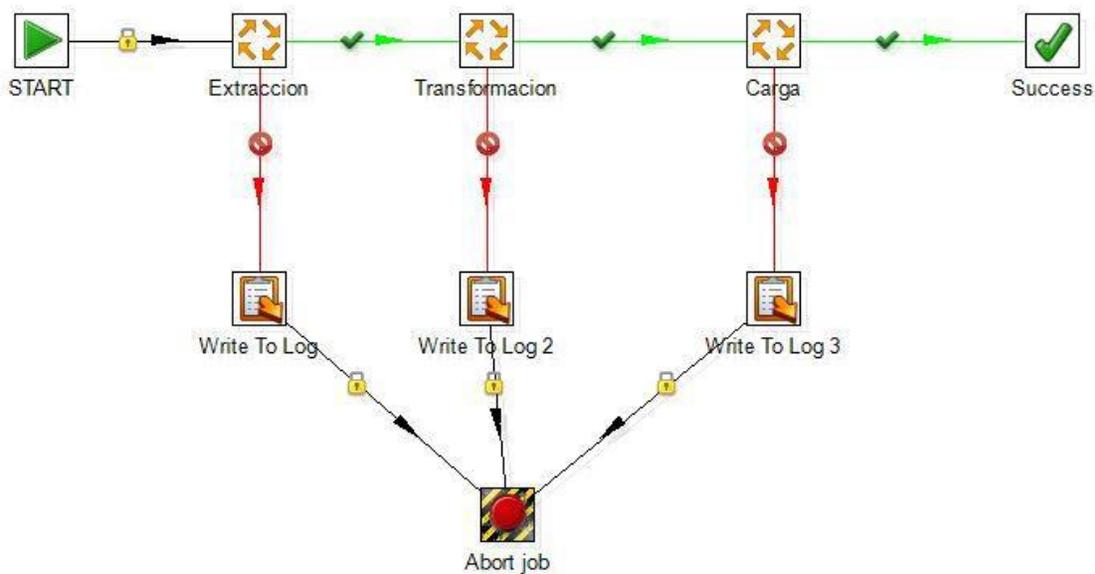


Figura 16 Job Principal

### 4.3.2 Extracción

Para que el DM propuesto en la solución cumpla con esquema definido, es necesario realizar una Extracción desde la fuente de datos de VALFI hacia un StagingArea, que es una base de datos temporal, que sirve para poder trabajar con los datos para no saturar el sistema origen, lo cual se visualiza en la Figura 17.

Para ver con más detalle la Extracción, ver 9 Anexo: Detalle Extracción.

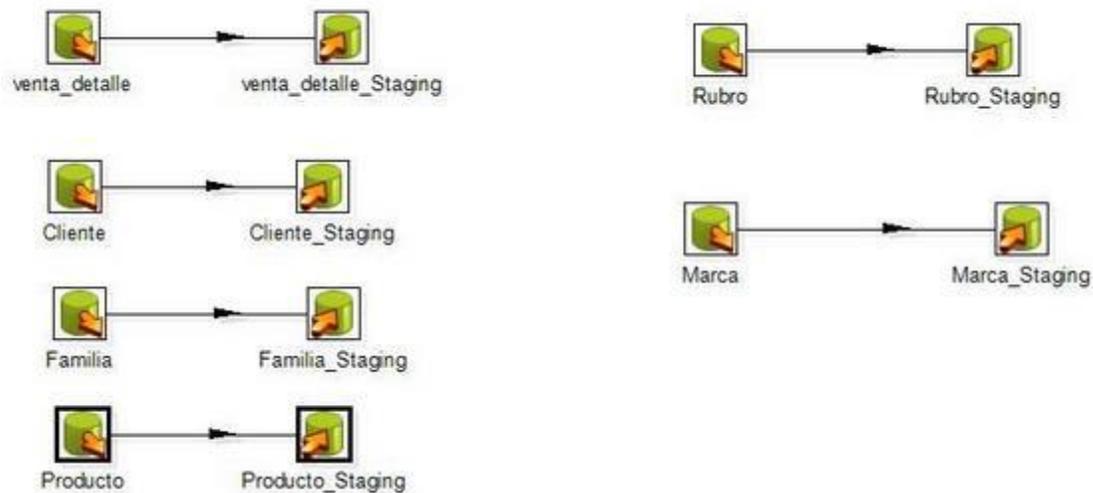


Figura 17 Extracción

### 4.3.3 Transformaciones Dimensiones

Para el proceso de Transformación, se ejecutará un job, el que se encargará del flujo de las transformaciones y en caso que haya algún error en cada paso, creará un archivo log para ver el detalle del error, lo que se observa en la Figura 18.

Para ver con más detalle Job Transformación, Ver 10 Anexo: Detalle Job Transformación.

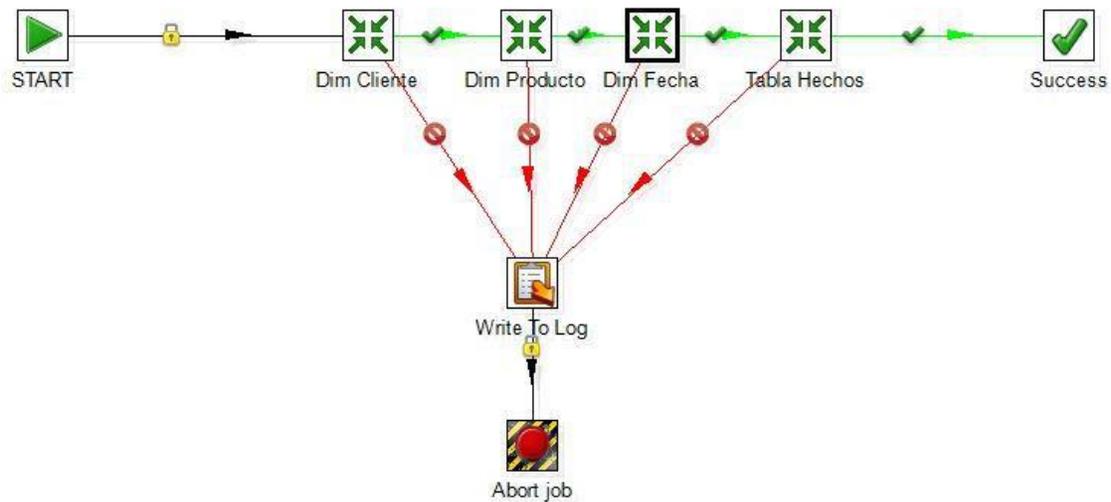


Figura 18 Job Transformación

#### 4.3.4 Transformación Dimensión “Producto”

Para la realización de la transformación “Producto” es necesario realizar sentencias “join” entre las tablas de origen “Producto”, “Familia”, “Rubro” y “Marca”, para poder crear la dimensión que se necesita “Dimensión Producto DV”. Como se puede observar en la Figura 18, se utilizan las tablas necesarias para las jerarquías a través de uniones para luego escoger los campos de la dimensión que serán almacenados en una zona intermedia al DM para su validación.

Para ver con más detalle la Transformación “Producto”, ver 11 Anexo: Detalle Transformación “Producto”.

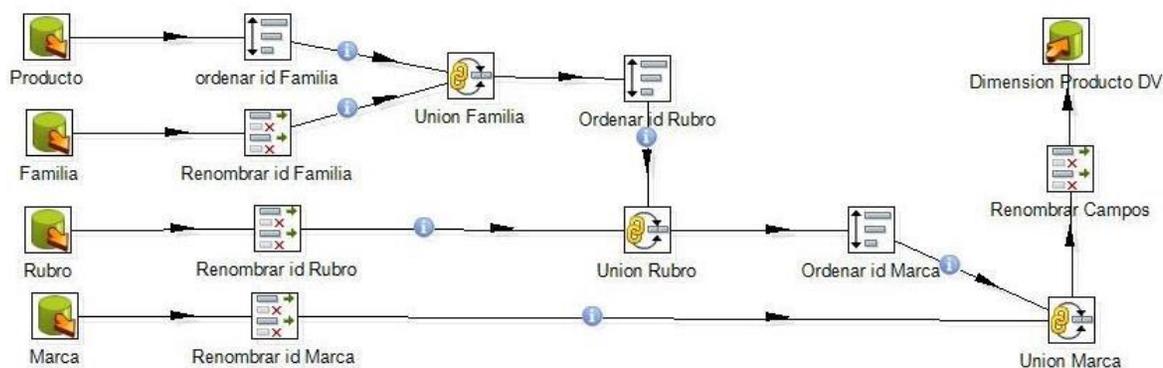


Figura 19 Transformación "Producto"

#### 4.3.5 Transformación Dimensión “Cliente”

Para la transformación de la Dimensión “Cliente” se utiliza la tabla de origen “Cliente” junto con un archivo externo en formato Excel “Comuna-Provincia-Ciudad-Region.xls” que contiene los datos para la jerarquía de localización (Región – Provincia – Comuna – Ciudad). Luego de unir la tabla de origen con el archivo en Excel, se selecciona los datos necesarios para la dimensión “Cliente” que serán almacenados en una zona intermedia al DM “DV Dimensión Cliente”, para su validación.

Para ver con más detalle la Transformación “Cliente”, ver 12 Anexo: Detalle Transformación “Cliente”.

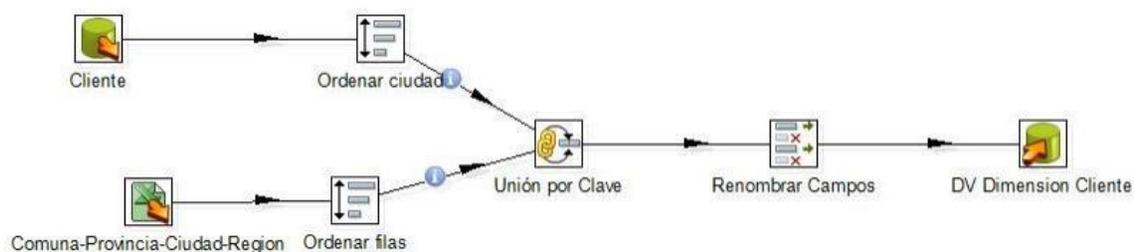
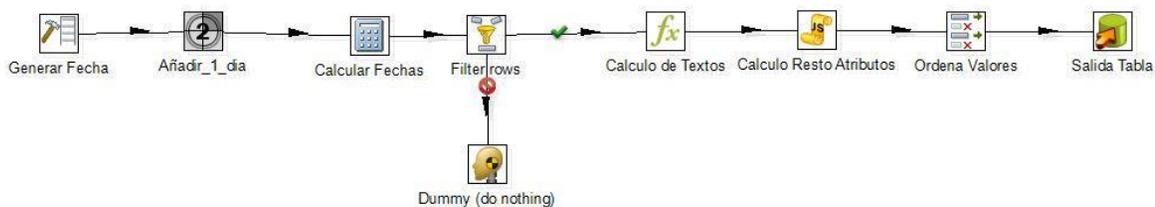


Figura 20 Transformación Dimensión "Cliente"

#### 4.3.6 Transformación Dimensión “Fecha”

Para la Dimensión “Fecha” se ha realizado una transformación la cual nos crea la dimensión completa con los campos y datos necesarios para su jerarquización y almacenándolo en una zona intermedia para su validación antes de la carga, lo cual se observa en la Figura 21.

Para ver con más detalle la Transformación “Fecha”, ver 13 Anexo: Detalle Transformación “Fecha”.

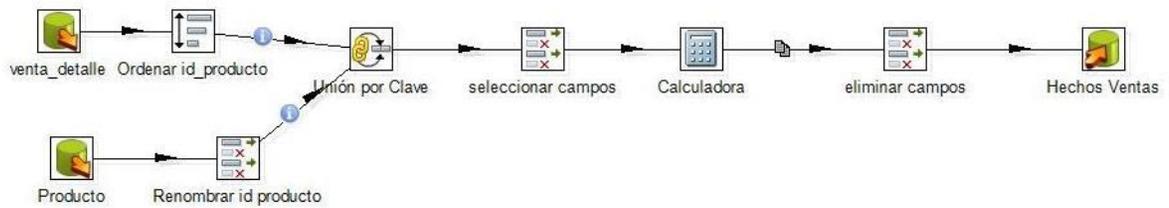


**Figura 21 Transformación Dimensión "Fecha"**

#### 4.3.7 Transformación Hecho “Venta”

En la Figura 22, se puede observar la transformación realizada para la tabla de hechos Venta, esta transformación recibe como fuentes de datos, las tablas “venta\_detalle” (que contiene la venta realizada) y la tabla producto (que contiene la información sobre los costos del producto, flete y ganancia) necesarios para las medidas definidas, luego de obtener los campos necesarios, se realiza los cálculos para determinar las medidas “monto\_venta”, “costo\_venta” y “ganancia”, luego se continúa escogiendo los campos definidos para los hechos y traspasarlos a una zona intermedia (DM) para su validación.

Para ver con más detalle el Hecho “Venta”, ver 14 Anexo: Detalle Transformación Hecho “Venta”.

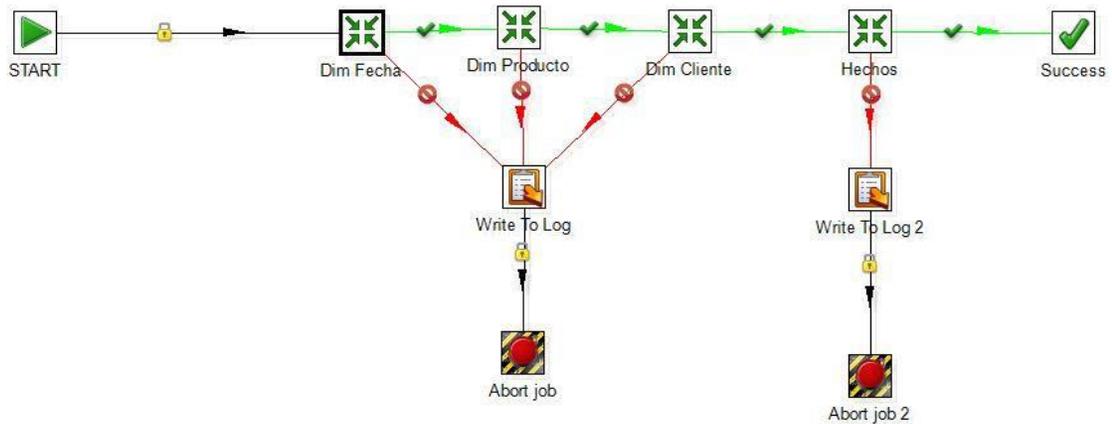


**Figura 22 Transformación Hecho "Venta"**

#### 4.3.8 Carga de Dimensiones y Hecho

Para el proceso de Carga, se ejecutará un job, el que se encargará del flujo de las transformaciones y en caso que haya algún error en cada paso, creará un archivo log para ver el detalle del error, lo que se observa en la Figura 23.

Para ver con más detalle Job Carga, Ver 15 Anexo: Detalle Job Carga”.



**Figura 23 Job Carga**

En el proceso de la carga se efectúa un proceso similar de las dimensiones “Cliente”, “Fecha” y “Producto” y el Hecho “Venta”, se extraen los datos desde la fuente intermedia en la que fueron validados y se trasladan a la fuente destino DM. Donde es muy importante para nuestra solución ya que esta carga enviará los datos y campos, hacia nuestro DM, el cual será usado para los análisis para la creación del Dashboard.

#### 4.3.8.1 Carga Dimensión “Cliente”

Para ver con más detalle el Carga “Cliente”, ver 16 Anexo: Detalle Carga “Cliente”.



Figura 24 Carga "Cliente"

#### 4.3.8.2 Carga Dimensión “Producto”

Para ver con más detalle el Carga “Producto”, Ver 17 Anexo: Detalle Carga “Producto”.



Figura 25 Carga "Producto"

#### 4.3.8.3 Carga Dimensión “Fecha”

Para ver con más detalle el Carga “Fecha”, ver 18 Anexo: Detalle Carga “Fecha”.



Figura 26 Carga "Fecha"

#### 4.3.8.4 Carga Hecho "Venta"

Para ver con más detalle el Carga "Venta", ver 19 Anexo: Detalle Carga "Venta".

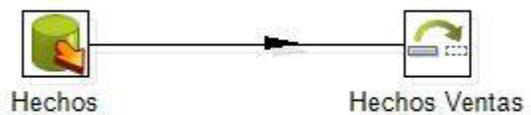


Figura 27 Carga "Venta"

#### 4.4 Análisis OLAP

Luego de haber realizado el proceso ETL, debemos crear nuestro cubo OLAP, para esto se utiliza Mondrian que es un motor de cubos dimensionales, que es necesario para poder realizar un análisis de cubo OLAP (análisis OLAP), ver Figura 28.

Para ver con más detalle el cubo OLAP, ver Anexo 21.1.



Figura 28 Cubo OLAP Mondrian

Este tipo de análisis ayudará a la empresa VALFI, a ver sus datos de una manera distinta, a través de las dimensiones establecidas (“Cliente”, “Fecha” y “Producto”). Donde se podrá obtener información general o a un alto nivel de detalle, esto dado por las jerarquías de cada dimensión. A continuación se presentará el análisis OLAP realizados a nuestro cubo OLAP creado en el DM, mostrando los beneficios que puede traer a VALFI.

Por ejemplo sí, la empresa quisiera saber cuáles son los productos que más vende y la procedencia de los clientes, el análisis OLAP nos permite obtener esta respuesta de una manera sencilla y rápida. El cubo OLAP permite responder a esta pregunta, puesto que en su diseño contamos con la medida "cantidad" y con las dimensiones "Cliente", "Producto" y "Fecha", lo que proporciona mostrar esta información de manera general o con un alto nivel de detalle. En la Figura 29 se puede observar como a través de una consulta en MDX se puede obtener la información, de las cantidades vendidas durante el año 2014 (Columna), por Región, Familia y Rubro (Fila). Además se observa cuadros vacíos pues dichos productos no han tenido ventas durante esa fecha en esa región pero sí en otras. Sí se necesita un mayor nivel de detalle (Provincia - Comuna) de la dimensión “Cliente” o (Marca y Producto) de la Dimensión “Producto”, solo deben seleccionarse y se realizará la consulta. Cabe señalar que una ventaja de este análisis OLAP es que las consultas se realizan al DM, para no saturar la base de datos OLTP de VALFI.

			Todos								
			2014								
			abril	enero	febrero	julio	junio			marzo	mayo
			6	6	6	6	4	5	6	6	6
Region	Familia	Rubro	Cantidad								
De La Araucanía	PRODUCTOS DE CONTROL	SECCIONADOR							2		
		REGLETAS									
Del BíoBío	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)								2	
	PRODUCTOS DE CONTROL	OBTURADOR								2	
		REGLETAS	2	2							2
	PRODUCTOS DE FUERZA	INTERRUPTOR AUTOMATICO	8	8	6		6	12	4	2	8
Del Libertador B. O'Higgins		FUSIBLES	5	5		4	2			5	1
	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	3	3	3						3
	PRODUCTOS DE CONTROL	OBTURADOR								2	
Del Maule	PRODUCTOS DE CONTROL	SECCIONADOR					2				
Metropolitana de Santiago	PRODUCTOS DE CONTROL	SECCIONADOR	1	1			2			1	1
		REGLETAS							2		
	PRODUCTOS DE FUERZA	INTERRUPTOR AUTOMATICO	3	3			2			3	3

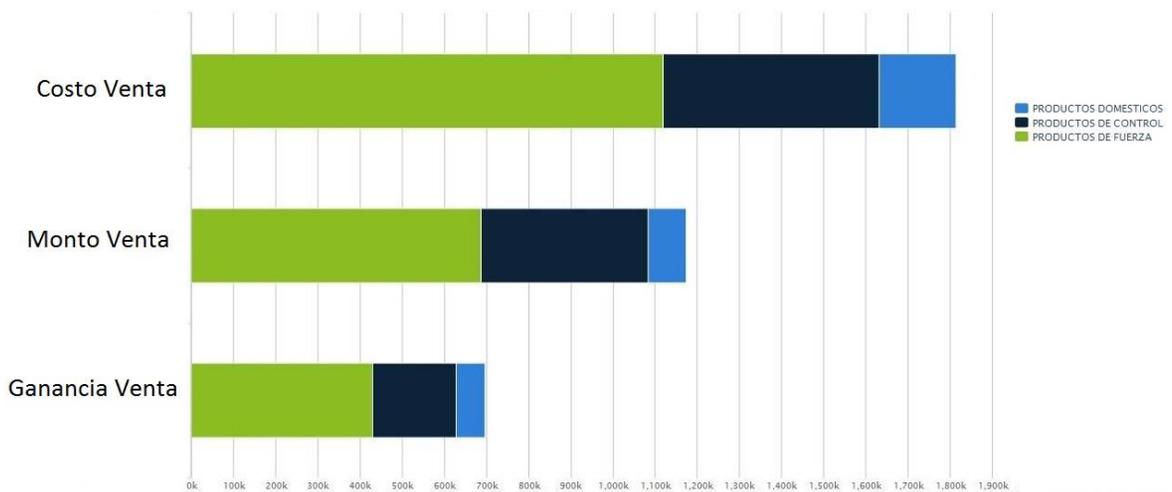
Figura 29 Análisis OLAP Ventas por fechas y procedencia clientes

De la misma manera se puede realizar el análisis OLAP sobre las ventas (cantidad- ganancia - costo - monto). En la Figura 30, se presenta un análisis utilizando las medidas "ganancia", "costo venta" y "monto venta" de la tabla de hechos "Venta" y la jerarquía completa de la dimensión "Cliente". Esta jerarquía nos permite no solo ver los montos totales, también nos entrega una segunda visión o perspectiva dada por la dimensión "Cliente".

(All)	Region	Provincia	Comuna	Ciudad	Cantidad	Costo Venta	Monto Venta	Ganancia Venta	
Todos	De La Araucanía	Cautín	Temuco	Temuco	4	81.900	57.700	31.500	
	Del Biobío	Bío- Bío	Los Angeles	Los Angeles	12	107.900	74.700	41.500	
			Concepción	Concepcion	Concepción	10	130.000	90.000	50.000
		Ñuble	Bulnes	Bulnes	Bulnes	2	26.000	18.000	10.000
			Chillan	Chillan	Chillan	44	561.600	316.800	216.000
			Pinto	Pinto	Pinto	10	234.000	162.000	90.000
			Quillon	Quillon	Quillon	2	26.000	18.000	10.000
			San Carlos	San Carlos	San Carlos	6	61.100	63.900	23.500
		San Nicolas	San Nicolas	San Nicolas	2	26.000	18.000	10.000	
	Del Libertador B. O'Higgins	Cachapoal	Rancagua	Rancagua	14	182.000	90.000	70.000	
	Del Maule	Talca	Talca	Talca	2	29.900	20.700	11.500	
	Metropolitana de Santiago	Santiago	Santiago	Santiago	22	347.100	244.300	133.500	

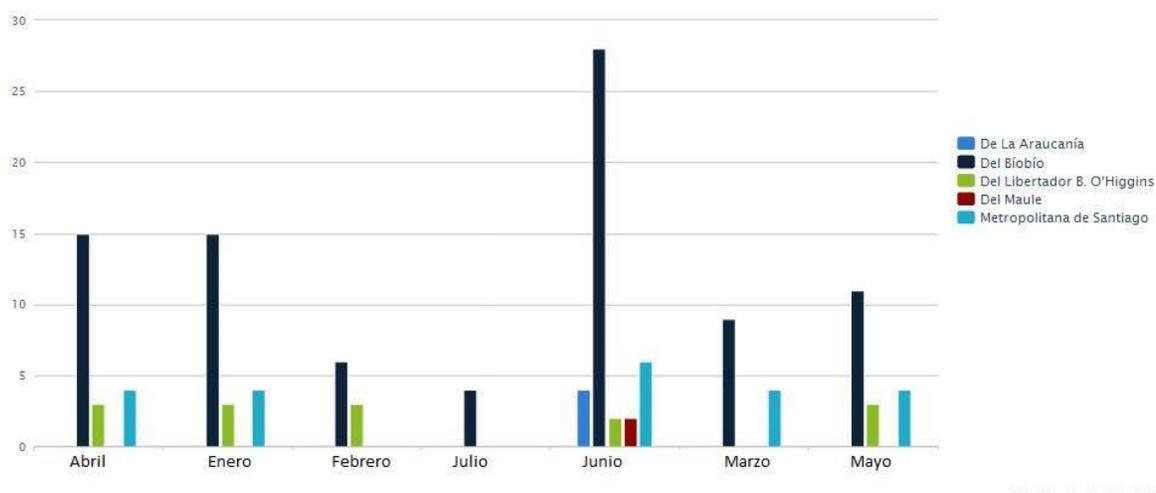
**Figura 30 Venta por procedencia de los clientes**

Estos análisis OLAP no solo pueden ser mostrados como tablas sino que también como gráficos. Donde se puede realizar una consulta al cubo OLAP y tener como resultado un gráfico, como se puede observar en la Figura 31. El análisis presentado es un gráfico de barras apilado en la que se observa el “costo”, “monto” y “ganancia” de las ventas pero bajo la perspectiva de las familias de productos dimensión “Producto”. Esto puede servir para observar que productos (en este caso familias de productos) presentan mayores ganancias con respecto a sus costos y precios de venta.



**Figura 31 Productos con mayores ganancias**

Otra forma en la que el análisis OLAP puede beneficiar a la empresa VALFI, es mostrando el comportamiento de las ventas que realizan. En la Figura 32 se puede observar un gráfico de barras, que nos presenta el comportamiento de las ventas (cantidades vendidas), durante el año (dimensión “Fecha”), para las diferentes regiones en las que se han realizado ventas (Dimension “Cliente”). Esto puede ayudar a VALFI a visualizar como sus ventas se ven afectadas durante una época particular del año, que meses las ventas bajan y en que regiones (clientes) sucede. Esto ayudaría a planificar y presupuestar un stock para esos casos.



**Figura 32 Comportamiento de las ventas durante un año**

Se han presentado varios ejemplos en los cuales el análisis OLAP puede ser útil para la empresa VALFI, no solo para consultar datos, sino también para ver comportamientos en sus ventas, no solo como tablas y números (Nivel Operativo), sino también como gráficos (Gerencia). Donde una gran ventaja de esta solución es que estas consultas pueden ser guardadas, actualizadas y cargadas al DM, lo que llevará a cumplir que la información sea actualizada en el cubo OLAP.

#### 4.5 Cuadro de Mando Integral (Dashboard/BalancedScoreCard)

El cuadro de mando Integral es una manera de presentar información de manera general y lo más simple (respecto a los datos), con indicadores clave (KPI) del negocio, todo esto para que sea cómodo ver la información para el nivel gerencial.

Para nuestra solución en el marco de la empresa VALFI, un cuadro de mando Integral presenta de manera resumida y general indicadores para la empresa.

En este cuadro de mando se mostrarán la cantidad de ventas realizadas por la empresa durante un año. Donde la importancia de este cuadro de mando Integral para la empresa, lo ayudará a obtener información como la siguiente:

- Saber que Regiones / Provincias / Comunas la empresa vende sus productos.
- Ver el comportamiento de esas ventas durante los doce meses del año.

Además el cuadro de mando integral es dinámico, donde al seleccionar un año se actualizan todos los gráficos y al seleccionar una región los demás gráficos mostrarán información con respecto a la información escogida.

Los datos del cuadro de mando integral provienen del DM, donde al tener actualizado el DM se tendrá actualizado este cuadro de mando, proporcionando esta información resumida.

# VALFI

*Electricidad*

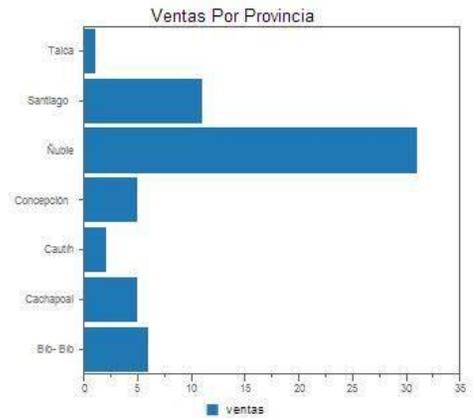


*"Mix de Soluciones"*

Ventas por año

Año

2014 ▾



Ciudad	Cantidad Ventas
Bulnes	1
Chillan	21
Concepcion	10
Los Angeles	10
Pinto	10
Quillon	1
Rancagua	10
San Carlos	10
San Nicolas	1
Santiago	11

Showing 1 to 10 of 12 entries

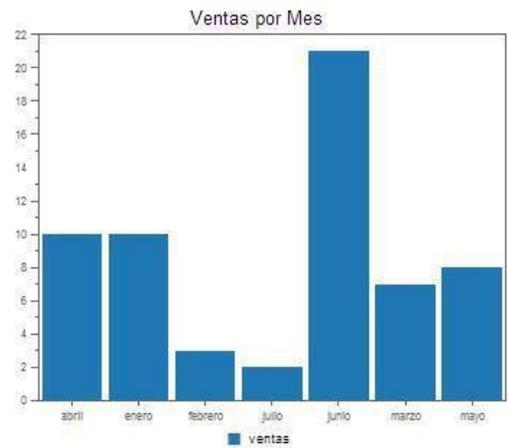


Figura 33 Cuadro de Mando Integral

---

## 5 PRUEBAS

---

En este capítulo se hablarán sobre los casos de prueba que se aplicaron en nuestro proyecto, para comprobar la calidad del proceso ETL.

Para comprobar la calidad de los procesos ETL, se ha tenido que realizar un plan de pruebas en las que se verifican, los datos de entrada (origen), el proceso ETL asociado, y los datos de Salida, (destino), se establecen las precondiciones para los casos de prueba en caso de ser necesarios, además de la ejecución exitosa del Proceso (verificar si se producen fallos durante la ejecución que detengan el proceso), esto también se verifica revisando los archivos Log de los procesos.

Además, se realiza un caso de prueba para el Cubo de análisis OLAP, el cual se realiza, para comprobar tras la ejecución de los procesos ETL, y luego de creado el Cubo Dimensional, los datos que se muestran corresponden a los datos que presenta el DM creado, esto se realiza a través de consultas SQL al DM, que luego serán realizadas en el Cubo Dimensional, los cuales deben presentar los mismos datos para cada consulta realizada.

Para observar los casos de prueba realizados, ver 20 Anexo: Plan de Pruebas.

---

## 6 CONCLUSIONES Y FUTURAS EXTENSIONES

---

Luego de haber realizado el proyecto de implementación de BI en el área de ventas clientes de VALFI, se puede sacar varias conclusiones debido a lo amplio de su tema.

### 6.1 Conclusiones

De acuerdo a todos los objetivos planteados en nuestro proyecto se obtuvieron las siguientes conclusiones:

- **Explotación de datos transaccionales del área de ventas clientes de la organización, transformándolos en información para la toma de decisiones estratégica.**

El poder explotar los datos transaccionales del área de ventas, según los requerimientos que necesitaba la organización, conocer el modelo de datos, los campos necesarios, el tipo de datos para verificar si era necesario alguna conversión en los procesos ETL, fue muy importante para la explotación de los datos y su incorporación en un DM para su paso a ser información para toma de decisiones.

- **Determinar las necesidades de información (Requerimientos de la Empresa).**

Determinar los requerimientos es muy importante, ya que con esto se puede ver que lo realmente necesita la organización, en este caso fue la implementación de BI en el área de ventas clientes, en donde además se definió las jerarquías de como tenían que ir en las dimensiones para adecuar las características de la organización en el diseño de los informes o el BalancedScoreCard.

- **Aplicar una metodología de BI.**

Se analizaron varias metodologías de BI en donde se compararon todos sus atributos y se llegó a la conclusión de aplicar la metodología Kimball en nuestro estudio, ya que se adecuaba más a nuestro requerimiento de implementar BI en el área de ventas clientes, ya que con esta metodología ya no va a ser necesario crear un DW sino que un DM, un almacén de datos de una sola área, lo que facilita además un trabajo futuro para esta misma empresa VALFI, si se

quiere realizar otro estudio de otra área de esta misma empresa se tendrá que implementar otro DM y si es necesario utilizar componentes de DM de ventas clientes como por ejemplo sus dimensiones.

- **Utilización de un DataMart (DM) usando herramientas OpenSource para el proceso ETL, la creación de cubos OLAP para la generación de informes y BalancedScoreCard.**

Para nuestro requerimiento de implementar BI en una pyme utilizando herramientas OpenSource es muy importante, ya que en la actualidad el BI es prácticamente una exclusividad de las grandes organizaciones públicas como privadas y una causa es la elevada suma que cuesta adquirir una licencia de una herramienta que puede costar ciento o miles de dólares, por lo cual las herramientas OpenSource es una gran solución para PYMEs de cualquier rubro en donde se puede obtener muy buenos resultados desde el proceso ETL hasta la construcción de los informes o de un BalancedScoreCard.

Todo esto representado en el capítulo de solución de nuestro proyecto, en la cual con un caso práctico se logra demostrar que un proyecto BI puede ser realizado con herramientas OpenSource y como pueden resolver problemas concretos de una empresa para su toma de decisiones a través de un DM que luego es utilizado para presentar información detallada utilizando Análisis OLAP en base a consultas e información resumida utilizando Cuadros de Mando.

## 6.2 Trabajo Futuro

Nuestra solución se puede extender para implementar BI en el área de ventas clientes en cualquier empresa, especialmente para una PYME, ya que para la implementación se utilizará herramientas OpenSource, porque es muy difícil que este tipo de empresas pueda adquirir una licencia de este tipo de herramientas. Además sí la implementación de BI es para una empresa diferente al rubro eléctrico va ser necesario realizar cambios en algunos de los componentes del DM como las dimensiones, jerarquías, porque están creado especialmente para la lógica de negocio de VALFI.

Otro trabajo futuro que se puede realizar es para este mismo proyecto, que es la implementación de BI para otra área de la empresa, esto se debe a la metodología aplicada (Kimball). Para esto va ser necesario la implementación de otro DM, donde sí es necesario se puede reutilizar los componentes del DM de ventas clientes ya implementado, como por ejemplo sus dimensiones.

---

## 7 BIBLIOGRAFÍA

---

[1] VALFI. [en línea]

< [http://www.valfi.cl/es\\_valfi.php](http://www.valfi.cl/es_valfi.php) > [consulta: 10 Octubre 2014]

[2] Sinnexus. ¿Qué es Business Intelligence?. [en línea]

< [http://www.sinnexus.com/business\\_intelligence/html](http://www.sinnexus.com/business_intelligence/html) > [consulta: 01 Julio 2014]

[3] Sinnexus. Datawarehouse. [en línea]

<[http://www.sinnexus.com/business\\_intelligence/datawarehouse.aspx](http://www.sinnexus.com/business_intelligence/datawarehouse.aspx)> [consulta: 01 Julio 2014].

[4] Pentaho. Data Integration – Kettle. [en línea]

<<http://community.pentaho.com/projects/data-integration/>> [consulta: 03 Julio 2014].

[5] Talend. talend\*. [en línea]

< <http://www.talend.com/> > [consulta: 03 Julio 2014].

[6] Apatar. Apatar connecting data. [en línea]

<<http://www.apatar.com/>> [consulta: 03 Julio 2014].

[7] Microsoft. SQL Server Integration Services. [en línea]

<<http://msdn.microsoft.com/es-cl/library/ms141026.aspx>> [consulta: 03 Julio 2014].

[8] Sinnexus. Cuadro de Mando Integral. [en línea]

<[http://www.sinnexus.com/business\\_intelligence/cuadro\\_mando\\_integral.aspx](http://www.sinnexus.com/business_intelligence/cuadro_mando_integral.aspx)> [consulta: 10 Julio 2014].

## 8 ANEXO: DETALLE JOB PRINCIPAL

Símbolo	Nombre	Descripción
	“START”	Comienzo del Job Principal.
	“Extraccion”	Job que ejecuta la Extracción.
	“Transformacion”	Job que ejecuta la Transformación.
	“Carga”	Job que ejecuta la Carga.
	“Write To Log”	En caso de error en el job de Extracción se escribirá un registro en el Log.
	“Write To Log 2”	En caso de error en el job de Transformación se escribirá un registro en el Log.
	“Write To Log 3”	En caso de error en el job de Carga se escribirá un registro en el Log.
	“Abort”	Se aborta el job en caso de cualquier error.
	“Success”	En caso de éxito se culmina el job.

**Tabla 7 Detalle Job Principal**

## 9 ANEXO: DETALLE EXTRACCIÓN

Símbolo	Nombre	Descripción
	“venta_detalle”	Extracción tabla “venta_detalle” del origen de datos.
	“Cliente”	Extracción tabla “Cliente” del origen de datos.
	“Familia”	Extracción tabla “Familia” del origen de datos.
	“Producto”	Extracción tabla “Producto” del origen de datos.
	“Rubro”	Extracción tabla “Rubro” del origen de datos.
	“Marca”	Extracción tabla “Marca” del origen de datos.
	“venta_detalle_Staging”	Carga la tabla “venta_detalle” a StagingArea.
	“Cliente_Staging”	Carga la tabla “Cliente” a StagingArea.
	“Familia_Staging”	Carga la tabla “Familia” a StagingArea.
	“Producto_Staging”	Carga la tabla “Producto” a StagingArea.
	“Rubro_Staging”	Carga la tabla “Rubro” a StagingArea.
	“Marca_Staging”	Carga la tabla “Marca” a StagingArea.

**Tabla 8 Detalle Extracción**

## 10 ANEXO: DETALLE JOB TRANSFORMACIÓN

Símbolo	Nombre	Descripción
	“START”	Comienzo del Job Transformación.
	“Dim Cliente”	Ejecución Transformación Dimensión Cliente.
	“Dim Producto”	Ejecución Transformación Dimensión Producto.
	“Dim Fecha”	Ejecución Transformación Dimensión Fecha.
	“Tabla Hechos”	Ejecución Transformación Tabla de Hechos.
	“Write To Log”	En caso de error en la Transformación de Transformación se escribirá un registro en el Log.
	“Abort Job”	Se aborta el job en caso de cualquier error.
	“Success”	En caso de éxito se culmina el job.

**Tabla 9 Detalle Job Transformación**

## 11 ANEXO: DETALLE TRANSFORMACIÓN “PRODUCTO”

Símbolo	Nombre	Descripción
	“Cliente”	Entrada para los datos de la tabla Cliente desde el StagingArea.
	“Comuna- Provincia-Ciudad- Region”	Entrada para los datos de la planilla Excel que contiene la información de la jerarquía.
	“Ordenar Ciudad”	Ordena la tabla Cliente por el campo ciudad.
	“Ordenar filas”	Ordena los datos de la planilla Excel por el campo “Nombre Comuna”.
	“UnionPorClave”	Paso en la que se realiza un Join entre lospasos Ordenar Ciudad y Ordenar filas.
	“Renombrar Campos”	Paso que Renombra los campos por los establecidos para la dimensión y elimina los campos no necesarios.
	“DV Dimension Cliente”	Salida para una tabla de una base de datos, en la que los datos se guardan para el proceso de carga.

**Tabla 10 Detalle Transformación "Producto"**

## 12 ANEXO: DETALLE TRANSFORMACIÓN “CLIENTE”

Símbolo	Nombre	Descripción
	“Cliente”	Entrada para los datos de la tabla Cliente desde el StagingArea.
	“Comuna- Provincia-Ciudad- Region”	Entrada para los datos de la planilla Excel que contiene la información de la jerarquía.
	“Ordenar Ciudad”	Ordena la tabla Cliente por el campo ciudad.
	“Ordenar filas”	Ordena los datos de la planilla Excel por el campo “Nombre Comuna”.
	“UnionPorClave”	Paso en la que se realiza un Join entre lospasos Ordenar Ciudad y Ordenar filas.
	“Renombrar Campos”	Paso que Renombra los campos por los establecidos para la dimensión y elimina los campos no necesarios.
	“DV Dimension Cliente”	Salida para una tabla de una base de datos, en la que los datos se guardan para el proceso de carga.

### 13 ANEXO: DETALLE TRANSFORMACIÓN “FECHA”

Símbolo	Nombre	Descripción
	“GenerarFecha”	Genera 10000 tuplas con el campo en formato fecha desde 1995-07-04.
	“Añadir1Dia”	Se añade un día a las tuplas para las fechas.
	“CalcularFechas”	Se generar varios campos para la dimensión.
	“FilterRows”	Se filtran los campos inferiores al año 1990.
	“Dummy”	Los campos que sean inferiores al año 1900, se van a este paso que no realiza nada.
	“CalculoDeTextos”	Se calculan los nombres de descripción de cada componente de la dimensión.
	“CalculoRestoAtributos”	Se calculan a través de JavaScript los atributos restantes.
	“OrdenaValores”	Se renombran los nombres de los campos.
	“SalidaTabla”	Salida para una tabla de una base de datos, en la que los datos se guardan para el proceso de carga.

Tabla 11 Detalle Transformación "Fecha"

## 14 ANEXO: DETALLE TRANSFORMACIÓN HECHO “VENTA”

Símbolo	Nombre	Descripción
	“Venta detalle”	Entrada para los datos de la tabla venta_detalle desde el StagingArea.
	“Producto”	Entrada para los datos de la tabla producto desde el StagingArea.
	“Unión por Clave”	Join entre las tablas venta_detalle y producto.
	“Calculadora”	Se calcula ganancia de una venta y el costo de la venta.
	“Seleccionar campos”	Se seleccionan los campos necesarios luego del join.
	“Eliminar campos”	Se seleccionan los campos y renombran para la tabla de hechos, según el modelado.
	“Hechos Ventas”	Salida para una tabla de una base de datos, en la que los datos se guardan para el proceso de carga.

Tabla 12 Detalle Hecho "Venta"

## 15 ANEXO: DETALLE JOB CARGA

Símbolo	Nombre	Descripción
	“START”	Comienzo del Job Carga.
	“Dim Fecha”	Ejecución Transformación Carga Dimensión Fecha.
	“Dim Producto”	Ejecución Transformación Carga Dimensión Producto.
	“Dim Cliente”	Ejecución Transformación Carga Dimensión Cliente.
	“Hechos”	Ejecución Transformación Carga Tabla de Hechos.
	“Write To Log”	En caso de error en la Transformación de Carga de las Dimensiones se escribirá un registro en el Log.
	“Write To Log 2”	En caso de error en la Transformación de Carga de la Tabla de Hechos se escribirá un registro en el Log.
	“Abort job”	Se aborta el job en caso de cualquier error en la carga de las dimensiones.
	“Abort job 2”	Se aborta el job en caso de cualquier error en la carga de la tabla de hechos.
	“Success”	En caso de éxito se culmina el job.

**Tabla 13 Detalle Job Carga**

---

## 16 ANEXO: DETALLE CARGA “CLIENTE”

---

Símbolo	Nombre	Descripción
	“Cliente”	Entrada en la cual se obtienen los datos de la fase de datos preparados para la carga.
	“DimensionCliente”	Se carga la dimensión en el DM, como dimensión lentamente cambiante.

**Tabla 14 Detalle Carga "Cliente"**

---

**17 ANEXO: DETALLE CARGA “PRODUCTO”**

---

Símbolo	Nombre	Descripción
	“Producto”	Entrada en la cual se obtienen los datos de la fase de datos preparados para la carga.
	“DimensionProducto”	Se carga la dimensión en el DM, como dimensión lentamente cambiante.

**Tabla 15 Detalle Carga "Producto"**

---

**18 ANEXO: DETALLE CARGA “FECHA”**

---

Símbolo	Nombre	Descripción
	“Fecha”	Entrada en la cual se obtienen los datos de la fase de datos preparados para la carga.
	“DimensionFecha”	Se carga la dimensión en el DM, como dimensión lentamente cambiante.

**Tabla 16 Detalle Carga "Fecha"**

---

**19 ANEXO: DETALLE CARGA “VENTA”**

---

Símbolo	Nombre	Descripción
	“Hechos”	Entrada en la cual se obtienen los datos de la fase de datos preparados para la carga.
	“HechoVentas”	Se carga los hechos en el DM, como dimensión lentamente cambiante.

**Tabla 17 Detalle Carga "Venta"**

## 20 ANEXO: PLAN DE PRUEBAS

A continuación se presenta el plan de pruebas realizado para este proyecto de Inteligencia de Negocios, el objetivo de este anexo es especificar los casos de prueba, secuencia de pruebas y el cómo se ejecutaron las diversas pruebas.

Para poder planificar el plan de pruebas, se debe definir a que componentes del proyecto se les realizaran las pruebas, se han definido realizar un plan de pruebas para los procesos ETL realizados en el proyecto para comprobar su funcionamiento y ejecución y tiempos. Además se realizaran pruebas para el cubo dimensional muestre la misma información que debería mostrar el DataMart.

A continuación se presentaran casos de prueba y ejecución para los procesos ETL, se realizara un caso de prueba para cada Transformación desarrollada en la creación del DataMart con la herramienta Pentaho Data Integration.

### 20.1 Caso de Prueba Job Extracción

#### 20.1.1 Caso de Prueba 1

Caso de Prueba N°1				
Objeto a analizar del caso de prueba:				
Proceso de Extracción y Staging Area.				
Objetivo del caso de prueba:				
Verificar que las tablas involucradas en la creación del DataMart sean correctamente cargadas al Staging Area (Base de Datos Intermedia).				
Pre-requisitos del caso de prueba:				
<ul style="list-style-type: none"> <li>- Disponibilidad de la BD fuente.</li> <li>- BD Intermedia creada.</li> <li>- El Job Extracción debe estar creado y listo para usarse.</li> </ul>				
Datos de Entrada:				
<ul style="list-style-type: none"> <li>- Tabla venta_detalle</li> <li>- Tabla Cliente</li> <li>- Tabla Familia</li> <li>- Tabla Producto</li> <li>- Tabla Rubro</li> <li>- Tabla Marca</li> </ul>				
N°	Función a probar	Acción	Resultado Esperado	Resultado Obtenido
1	Ejecución Transformación, y comparación cantidad de registros en BD fuente frente a cantidad de registros en Staging Area.	Ejecutar el Job Extracción	Las tablas homologas en la Staging Area, muestren los mismos registros que BD fuente.	Job ejecutado correctamente, tablas generadas en el Staging Area

**Tabla 18 Caso de Prueba 1**

## 20.1.2 Ejecución Caso de Prueba 1

### 20.1.2.1 Datos de Entrada

A continuación se presentaran las tablas fuentes, visualizadas a través del motor de base de datos utilizada como fuente (SQL Server).

The screenshot shows a SQL Server query window with the following SQL code:

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT [id_venta_detalle]
      ,[pk_producto]
      ,[pk_cliente]
      ,[cantidad_venta]
      ,[precio_venta]
      ,[descuento]
      ,[sub_total]
      ,[fecha_venta]
FROM [Valfi].[dbo].[Venta_detalle]
    
```

The results pane displays the following data:

	id_venta_detalle	pk_producto	pk_cliente	cantidad_venta	precio_venta	descuento	sub_total	fecha_venta
1	1	1	1	2	20000	2000	18000	2014-06-06 00:00:00.000
2	2	2	55	2	20000	2000	18000	2014-06-06 00:00:00.000
3	3	2	59	2	20000	2000	18000	2014-06-06 00:00:00.000
4	4	5	60	2	20000	2000	18000	2014-06-06 00:00:00.000
5	5	5	61	2	20000	2000	18000	2014-06-06 00:00:00.000
6	6	10	36	2	20000	2000	18000	2014-06-05 00:00:00.000
7	7	10	6	2	20000	2000	18000	2014-06-05 00:00:00.000
8	8	11	6	2	20700	0	20700	2014-06-05 00:00:00.000
9	9	12	1	2	20700	0	20700	2014-06-05 00:00:00.000
10	10	12	1	2	20700	0	20700	2014-06-05 00:00:00.000
11	11	31	43	2	39000	3000	36000	2014-06-05 00:00:00.000

Query executed successfully. IVAN-NOTE\SQLSERVER (11.0 ... IVAN-NOTE\ivan (54) master 00:00:00 61 rows

Figura 34 Tabla Venta\_Detalle

SQLQuery5.sql - IV...AN-NOTE\Ivan (57)\* X

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT [id_rubro]
, [nombre_rubro]
FROM [Valfi].[dbo].[Rubro]
    
```

100 %

Results Messages

	id_rubro	nombre_rubro
1	1	CAJAS (TABLEROS)
2	2	OBTURADOR
3	3	INTERRUPTOR AUTOMATICO
4	4	SECCIONADOR
5	5	RELOJ
6	6	P
7	7	BARRA
8	8	REGLETAS
9	9	REPARTIDOR
10	10	PORTAFUSIBLE
11	11	DIFERENCIAL

Query executed successfully. IVAN-NOTE\SQLSERVER (11.0 ... | IVAN-NOTE\Ivan (57) | master | 00:00:00 | 44 rows

**Figura 35** Tabla Rubro

SQLQuery1.sql - IV...AN-NOTE\Ivan (58)\* X

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT [id_cliente]
      ,[rut_cliente]
      ,[nombre_cliente]
      ,[fonoC]
      ,[fonoT]
      ,[email]
      ,[direccion]
      ,[ciudad]
FROM [Valfi].[dbo].[Cliente]
    
```

100 %

Results Messages

	id_cliente	rut_cliente	nombre_cliente	fonoC	fonoT	email	direccion	ciudad
1	1	79.938.086	AGRICOLA LA HIGUERA LTDA.	222222	99887766	aaa@aaa.cl	calle aaa 121	Chillar
2	2	13.128.907	ARREGUI LAGOS JORGE	222222	99887766	aaa@aaa.cl	calle aaa 122	Chillar
3	3	14.169.681	BARRA CANDIA GUSTAVO	222222	99887766	aaa@aaa.cl	calle aaa 123	Chillar
4	4	06.165.711	BETANCOURT MARTINEZ MARIO	222222	99887766	aaa@aaa.cl	calle aaa 124	Chillar
5	5	99.558.800	CABLE CENTRAL S.A.	222222	99887766	aaa@aaa.cl	calle aaa 125	Chillar
6	6	81.826.800	CAJA DE COMPENSACIÓN DE LOS ANDES	222222	99887766	aaa@aaa.cl	calle aaa 126	Chillar
7	7	08.138.992	CARTES FUENTEALBA JUAN	222222	99887766	aaa@aaa.cl	calle aaa 127	Chillar
8	8	85.783.800	CEMET S.A.C.I.	222222	99887766	aaa@aaa.cl	calle aaa 128	Chillar
9	9	11.088.091	CERDA RIQUELME VICTOR	222222	99887766	aaa@aaa.cl	calle aaa 129	Chillar
10	10	78.784.950	COMERCIAL D Y D LTDA.	222222	99887766	aaa@aaa.cl	calle aaa 130	Chillar
11	11	96.994.840	COMERCIAL WOOD MARKET S.A.	222222	99887766	aaa@aaa.cl	calle aaa 131	Chillar
12	12	76.227.030	CONSTRUCTORA CARLOS MUÑOZ E HIJOS LTDA	222222	99887766	aaa@aaa.cl	calle aaa 132	Chillar
13	13	82.642.300	CONSTRUCTORA NAZARETH LIMITADA	222222	99887766	aaa@aaa.cl	calle aaa 133	Chillar
14	14	78.117.710	CONSTRUCTORA PICHILEMU LTDA.	222222	99887766	aaa@aaa.cl	calle aaa 134	Chillar
15	15	78.000.500	CONSTRUCTORA RICARDO FUENTES CALMUCHO S.R.L.	222222	99887766	aaa@aaa.cl	calle aaa 135	Chillar

Query executed successfully. | IVAN-NOTE\SQLEXPRESS (11.0 ... | IVAN-NOTE\Ivan (58) | master | 00:00:00 | 65 rows

Figura 36 Tabla Cliente

SQLQuery2.sql - IV...AN-NOTE\Ivan (51)\* x

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT [id_familia]
, [nombre_familia]
FROM [Valfi].[dbo].[Familia]
    
```

100 %

Results Messages

	id_familia	nombre_familia
1	1	PRODUCTOS DOMESTICOS
2	2	PRODUCTOS DE CONTROL
3	3	PRODUCTOS DE FUERZA
4	4	PRODUCTOS DE CONTROL
5	5	VARIOS
6	6	ARTICULOS ELECTRONICA
7	7	APARATOS (Placas, Modulos, Int, Enchufes)
8	8	ILUMINACIÓN

Query executed successfully. | IVAN-NOTE\SQLEXPRESS (11.0 ... | IVAN-NOTE\Ivan (51) | master | 00:00:00 | 8 rows

**Figura 37** Tabla Familia

The screenshot shows a SQL Server Enterprise Manager window with a query executed. The query is as follows:

```
/****** Script for SelectTopNRows command from SSMS *****/  
SELECT [id_marca]  
      , [nombre_marca]  
FROM [Valfi].[dbo].[Marca]
```

The results pane shows the following data:

	id_marca	nombre_marca
1	1	LEGRAND
2	2	bticino

The status bar at the bottom indicates: Query executed successfully. | IVAN-NOTE\SQLEXPRESS (11.0 ... | IVAN-NOTE\Ivan (51) | master | 00:00:00 | 2 rows

**Figura 38** Tabla Marca

SQLQuery4.sql - IV...AN-NOTE\Ivan (54)\* X

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT [id_producto]
,[pk_familia]
,[pk_rubro]
,[pk_marca]
,[referencia]
,[precio]
,[ganancia]
,[stock]
,[stock_critico]
,[costo_flete]
FROM [Valfi].[dbo].[Producto]
    
```

100 %

Results Messages

	id_producto	pk_familia	pk_rubro	pk_marca	referencia	precio	ganancia	stock	s
1	1	1	1	1	01512 23 EMB. 26 CTOS. PTA BLANCA LEGRAND	10000	5000	200	:
2	2	2	2	1	01654 2 LEGRAND	10000	5000	200	:
3	3	3	3	1	03382 INTERRUPTOR AUTOMATICO 1 X 6 A. 6KA LEGRAND	10000	5000	200	:
4	4	3	3	1	03384 INTERRUPTOR AUTOMATICO 1 X 10 A. 6KA LEGRAND	10000	5000	200	:
5	5	3	3	1	03386 INTERRUPTOR AUTOMATICO 1 X 16 A. 6KA LEGRAND	10000	5000	200	:
6	6	3	3	1	03387 INTERRUPTOR AUTOMATICO 1 X 20 A. 6KA LEGRAND	10000	5000	200	:
7	7	3	3	1	03388 INTERRUPTOR AUTOMATICO 1 X 25 A. 6KA LEGRAND	10000	5000	200	:
8	8	3	3	1	03389 INTERRUPTOR AUTOMATICO 1 X 32 A. 6KA LEGRAND	10000	5000	200	:
9	9	3	3	1	03449 INT. AUT. CURVA C 3 X 10 A. 6KA LEGRAND	10000	5000	200	:
10	10	3	3	1	03451 INT. AUT. CURVA C 3 X 16 A. 6KA LEGRAND	10000	5000	200	:
11	11	3	3	1	03452 INT. AUT. CURVA C 3 X 20 A. 6KA LEGRAND	11500	5750	200	:
12	12	3	3	1	03453 INT. AUT. CURVA C 3 X 25 A. 6KA LEGRAND	11500	5750	200	:
13	13	2	4	1	03730 5 CONTROL (15 MIN.) S/RES.	11500	5750	200	:
14	14	3	5	1	03740 5 CONTROL (15 MIN.) S/RES.	11500	5750	200	:

Query executed successfully. | IVAN-NOTE\SQLEXPRESS (11.0 ... | IVAN-NOTE\Ivan (54) | master | 00:00:00 | 736 rows

Figura 39 Tabla Producto

### 20.1.2.2 ETL

A continuación se presenta imágenes durante la ejecución del proceso ETL, la herramienta Kettle presenta una visualización de los pasos con datos de entrada y salida en cada paso y una vista de log, indicando cada suceso en el proceso ETL.

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

```

2014/09/30 02:01:18 - Version checker - OK
2014/09/30 02:01:26 - Spoon - Spoon
2014/09/30 02:01:53 - Spoon - Transformación abierta.
2014/09/30 02:01:53 - Spoon - Ejecutando transformación [Extraccion]...
2014/09/30 02:01:53 - Spoon - Se ha iniciado la ejecución de la transformación.
2014/09/30 02:01:53 - Extraccion - Iniciado despacho de la transformación [Extraccion]
2014/09/30 02:01:54 - Familia_Staging.0 - Connected to database [stagingarea] (commit=1000)
2014/09/30 02:01:54 - Cliente_Staging.0 - Connected to database [stagingarea] (commit=1000)
2014/09/30 02:01:54 - Rubro_Staging.0 - Connected to database [stagingarea] (commit=1000)
2014/09/30 02:01:54 - venta_detalle_Staging.0 - Connected to database [stagingarea] (commit=1000)
2014/09/30 02:01:57 - Marca_Staging.0 - Connected to database [stagingarea] (commit=1000)
2014/09/30 02:01:57 - Producto_Staging.0 - Connected to database [stagingarea] (commit=1000)
2014/09/30 02:02:17 - Marca.0 - Finished reading query, closing connection.
2014/09/30 02:02:17 - Familia.0 - Finished reading query, closing connection.
2014/09/30 02:02:17 - Rubro.0 - Finished reading query, closing connection.
2014/09/30 02:02:17 - venta_detalle.0 - Finished reading query, closing connection.
2014/09/30 02:02:17 - Marca.0 - Procesamiento finalizado (I=2, O=0, R=0, W=2, U=0, E=0)
2014/09/30 02:02:17 - venta_detalle.0 - Procesamiento finalizado (I=61, O=0, R=0, W=61, U=0, E=0)
2014/09/30 02:02:17 - Rubro.0 - Procesamiento finalizado (I=44, O=0, R=0, W=44, U=0, E=0)
2014/09/30 02:02:17 - Familia.0 - Procesamiento finalizado (I=8, O=0, R=0, W=8, U=0, E=0)
2014/09/30 02:02:17 - Cliente.0 - Finished reading query, closing connection.
2014/09/30 02:02:17 - Cliente.0 - Procesamiento finalizado (I=65, O=0, R=0, W=65, U=0, E=0)
2014/09/30 02:02:17 - Producto.0 - Finished reading query, closing connection.
2014/09/30 02:02:17 - Producto.0 - Procesamiento finalizado (I=736, O=0, R=0, W=736, U=0, E=0)
2014/09/30 02:02:17 - Marca_Staging.0 - Procesamiento finalizado (I=0, O=2, R=2, W=2, U=0, E=0)
2014/09/30 02:02:17 - Familia_Staging.0 - Procesamiento finalizado (I=0, O=8, R=8, W=8, U=0, E=0)
2014/09/30 02:02:17 - Cliente_Staging.0 - Procesamiento finalizado (I=0, O=65, R=65, W=65, U=0, E=0)
2014/09/30 02:02:17 - Rubro_Staging.0 - Procesamiento finalizado (I=0, O=44, R=44, W=44, U=0, E=0)
2014/09/30 02:02:17 - venta_detalle_Staging.0 - Procesamiento finalizado (I=0, O=61, R=61, W=61, U=0, E=0)
2014/09/30 02:02:18 - Producto_Staging.0 - Procesamiento finalizado (I=0, O=736, R=736, W=736, U=0, E=0)
2014/09/30 02:02:18 - Spoon - La transformación ha finalizado!!
    
```

Figura 40 Logging Extracción

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo	Tiempo	Velocidad (r/s)
1	Cliente	0	0	65	65	0	0	0	0	Finalizado	0.0s	1.354
2	venta_detalle	0	0	61	61	0	0	0	0	Finalizado	0.0s	1.419
3	Producto	0	0	736	736	0	0	0	0	Finalizado	0.1s	10.082
4	Rubro	0	0	44	44	0	0	0	0	Finalizado	0.0s	1.023
5	Familia	0	0	8	8	0	0	0	0	Finalizado	0.0s	186
6	Marca	0	0	2	2	0	0	0	0	Finalizado	0.0s	48
7	Rubro_Staging	0	44	44	0	44	0	0	0	Finalizado	0.5s	90
8	Cliente_Staging	0	65	65	0	65	0	0	0	Finalizado	0.5s	133
9	Familia_Staging	0	8	8	0	8	0	0	0	Finalizado	0.5s	18
10	Producto_Staging	0	736	736	0	736	0	0	0	Finalizado	1.1s	673
11	Marca_Staging	0	2	2	0	2	0	0	0	Finalizado	0.3s	7
12	venta_detalle_Staging	0	61	61	0	61	0	0	0	Finalizado	0.5s	124

Figura 41 Step Metrics Extracción

Las figuras anteriores muestran la ejecución sin problemas de la ETL Extracción, además de proporcionar información de los datos de entradas y salida de cada paso, y datos escritos en la Staging Area, ahora solo queda comprobar que los registros fueron pasados al Staging Area.

### 20.1.2.3 Tablas Finales

Tras la ejecución del Proceso ETL de Extracción, se deben verificar que las tablas han sido replicadas en la Staging Area con todos sus registros, esto lo podemos comprobar observando en el Motor de la Staging Area (MySQL).

id_venta_detalle	pk_producto	pk_cliente	cantidad_venta	precio_venta	descuento	sub_total	fecha_venta
1	1	1	2	20000	2000	18000	2014-06-06
2	2	55	2	20000	2000	18000	2014-06-06
3	2	59	2	20000	2000	18000	2014-06-06
4	5	60	2	20000	2000	18000	2014-06-06
5	5	61	2	20000	2000	18000	2014-06-06
6	10	36	2	20000	2000	18000	2014-06-05
7	10	6	2	20000	2000	18000	2014-06-05
8	11	6	2	20700	0	20700	2014-06-05
9	12	1	2	20700	0	20700	2014-06-05
10	12	1	2	20700	0	20700	2014-06-05
11	31	43	2	39000	3000	36000	2014-06-05
12	11	1	2	20700	0	20700	2014-06-05
13	15	49	2	21000	300	21700	2014-06-05

Figura 42 Tabla Venta\_Detalle

SQL File 1 rubro x

1 • `SELECT * FROM tesis_stagingarea.rubro;`

Result Set Filter:  Export: Wrap Cell Content:

	id_rubro	nombre_rubro
▶	1	CAJAS (TABLEROS)
	2	OBTURADOR
	3	INTERRUPOTOR AUTOMATICO
	4	SECCIONADOR
	5	RELOJ
	6	P
	7	BARRA
	8	REGLETAS
	9	REPARTIDOR
	10	PORTAFUSIBLE
	11	DIFERENCIAL
	12	FUSIBLES
	13	CANAI FTA

rubro 1 x

Output

Action Output

	Time	Action	Message
✓	1 00:04:48	SELECT * FROM tesis_stagingarea.rubro	44 row(s) returned

Figura 43 Tabla Rubro

SQL File 1 cliente x

1 • `SELECT * FROM tesis_stagingarea.cliente;`

Result Set Filter: [ ] Export: [ ] Wrap Cell Content: [ ]

	id_cliente	rut_cliente	nombre_cliente	fonoC	fonoT	email	direccion	ciudad
▶	1	79.938.086	AGRICOLA LA HIGUERA LTDA.	222222	99887766	aaa@aaa.cl	calle aaa 121	Chillan
	2	13.128.907	ARREGUI LAGOS JORGE	222222	99887766	aaa@aaa.cl	calle aaa 122	Chillan
	3	14.169.681	BARRA CANDIA GUSTAVO	222222	99887766	aaa@aaa.cl	calle aaa 123	Chillan
	4	06.165.711	BETANCOURT MARTINEZ MARIO	222222	99887766	aaa@aaa.cl	calle aaa 124	Chillan
	5	99.558.800	CABLE CENTRAL S.A.	222222	99887766	aaa@aaa.cl	calle aaa 125	Chillan
	6	81.826.800	CAJA DE COMPENSACIÓN DE LOS ANDES	222222	99887766	aaa@aaa.cl	calle aaa 126	Chillan
	7	08.138.992	CARTES FUENTEALBA JUAN	222222	99887766	aaa@aaa.cl	calle aaa 127	Chillan
	8	85.783.800	CEMET S.A.C.I.	222222	99887766	aaa@aaa.cl	calle aaa 128	Chillan
	9	11.088.091	CERDA RIQUELME VICTOR	222222	99887766	aaa@aaa.cl	calle aaa 129	Chillan
	10	78.784.950	COMERCIAL D Y D LTDA.	222222	99887766	aaa@aaa.cl	calle aaa 130	Chillan
	11	96.994.840	COMERCIAL WOOD MARKET S.A.	222222	99887766	aaa@aaa.cl	calle aaa 131	Chillan
	12	76.227.030	CONSTRUCTORA CARLOS MUÑOZ E HIJOS LTDA	222222	99887766	aaa@aaa.cl	calle aaa 132	Chillan
	13	82.642.300	CONSTRUCTORA NAZARETH LIMITADA	222222	99887766	aaa@aaa.cl	calle aaa 133	Chillan

cliente 2 x

Output

Action Output

	Time	Action	Message
✓	1 23:59:54	SELECT * FROM tesis_stagingarea.cliente	65 row(s) returned

**Figura 44** Tabla Cliente

SQL File 1 familia x

1 • `SELECT * FROM tesis_stagingarea.familia;`

Result Set Filter:  Export: Wrap Cell Content:

	id_familia	nombre_familia
▶	1	PRODUCTOS DOMESTICOS
	2	PRODUCTOS DE CONTROL
	3	PRODUCTOS DE FUERZA
	4	PRODUCTOS DE CONTROL
	5	VARIOS
	6	ARTICULOS ELECTRONICA
	7	APARATOS (Placas, Modulos, Int, Enchufes)
	8	ILUMINACIÓN

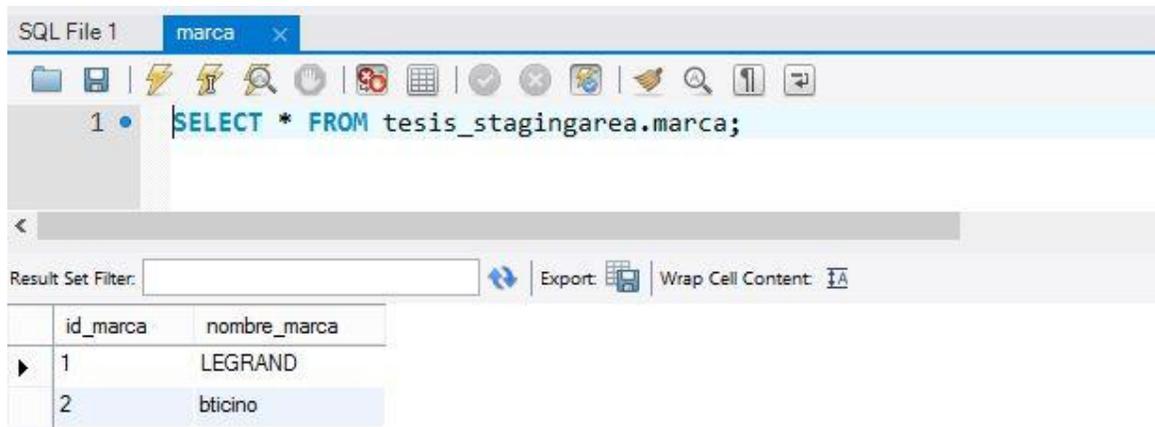
familia 1 x

Output

Action Output

	Time	Action	Message
✓	1 00:00:25	SELECT * FROM tesis_stagingarea.familia	8 row(s) returned

Figura 45 Tabla Familia



**Figura 46** Tabla Marca

SQL File 1 producto x

```
1 • SELECT * FROM tesis_stagingarea.producto;
```

Result Set Filter: Export Wrap Cell Content

id_producto	pk_familia	pk_rubro	pk_marca	referencia	precio	ganancia	stock	stock_critico	costo_flete
1	1	1	1	01512 23 EMB. 26 CTOS. PTA BLANCA LEGRAND	10000	5000	200	20	3000
2	2	2	1	01654 2 LEGRAND	10000	5000	200	20	3000
3	3	3	1	03382 INTERRUPTOR AUTOMATICO 1 X 6 A. 6KA LEGRAND	10000	5000	200	20	3000
4	3	3	1	03384 INTERRUPTOR AUTOMATICO 1 X 10 A. 6KA LEGRAND	10000	5000	200	20	3000
5	3	3	1	03386 INTERRUPTOR AUTOMATICO 1 X 16 A. 6KA LEGRAND	10000	5000	200	20	3000
6	3	3	1	03387 INTERRUPTOR AUTOMATICO 1 X 20 A. 6KA LEGRAND	10000	5000	200	20	3000
7	3	3	1	03388 INTERRUPTOR AUTOMATICO 1 X 25 A. 6KA LEGRAND	10000	5000	200	20	3000
8	3	3	1	03389 INTERRUPTOR AUTOMATICO 1 X 32 A. 6KA LEGRAND	10000	5000	200	20	3000
9	3	3	1	03449 INT. AUT. CURVA C 3 X 10 A. 6KA LEGRAND	10000	5000	200	20	3000
10	3	3	1	03451 INT. AUT. CURVA C 3 X 16 A. 6KA LEGRAND	10000	5000	200	20	3000
11	3	3	1	03452 INT. AUT. CURVA C 3 X 20 A. 6KA LEGRAND	11500	5750	200	20	3450
12	3	3	1	03453 INT. AUT. CURVA C 3 X 25 A. 6KA LEGRAND	11500	5750	200	20	3450
13	2	4	1	03730 5 CONTROL (15 MIN 1 S./RFS.	11500	5750	200	20	3450

producto 1 x Read Only

Output

Action Output

Time	Action	Message	Duration / Fetch
1 00:04:11	SELECT * FROM tesis_stagingarea.producto	736 row(s) returned	0.000 sec / 0.000 se

**Figura 47** Tabla Producto

Tras Revisar las tablas en el Staging Area, se puede comprobar la ejecución exitosa del proceso ETL, y las tablas replicadas son consistentes con las tablas fuentes.

## 20.2 Caso de Prueba Job Transformación

### 20.2.1 Caso de Prueba 2

Caso de Prueba N°2				
Objeto a analizar del caso de prueba:				
Procesos de Transformación.				
Objetivo del caso de prueba:				
Verificar que los procesos de transformación para la dimensiones se efectúen de manera correcta.				
Pre-requisitos del caso de prueba:				
<ul style="list-style-type: none"> <li>- Disponibilidad de la Staging Area.</li> <li>- BD Intermedia creada (base de datos con los datos ya limpiados y procesados).</li> <li>- El Job Transformación debe estar creado y listo para usarse.</li> </ul>				
Datos de Entrada:				
<ul style="list-style-type: none"> <li>- Tabla Cliente (Staging Area)</li> <li>- Tabla Familia (Staging Area)</li> <li>- Tabla Marca (Staging Area)</li> <li>- Tabla Producto (Staging Area)</li> <li>- Tabla Rubro (Staging Area)</li> <li>- Tabla venta_detalle (Staging Area)</li> <li>- Archivo DPA_acotado.xls (Dato Externo)</li> </ul>				
N°	Función a probar	Acción	Resultado Esperado	Resultado Obtenido
1	Ejecución Transformación, y verificación que los datos hayan sido traspasados luego del proceso.	Ejecutar el Job Transformación	La Transformación se ejecuta exitosamente y las tablas generadas con éxito.	Ejecución Sin Problemas del Job.

**Tabla 19 Caso de Prueba 2**

### 20.2.2 Ejecución Caso de Prueba 2

#### 20.2.2.1 Datos de Entrada

A Continuación se Presenta una imagen con la información del Dato Externo utilizado durante la transformación de la Dimensión Cliente (Necesario para obtener su Localidad) el archivo Externo es un archivo en formato Excel, obtenido desde datos.gob.cl (datos abiertos), con la información a nivel nacional de regiones, provincias y comunas, los otros datos de entrada están presentados en el caso de prueba anterior (verificación Staging Area)

	A	B	C	D	E	F	G	H
	Nombre Comuna	Código Comuna desde 2010	Provincia desde 2010	Código Provincia desde 2010	Nombre Región desde 2008	Código Región desde 2008	Nombre Servicio de Salud desde 2008	Código Servicio de Salud desde 2008
1								
2	Arica	15101	Arica	151	De Arica y Parinacota	15	Arica	01
3	Camarones	15102	Arica	151	De Arica y Parinacota	15	Arica	01
4	Putre	15201	Parinacota	152	De Arica y Parinacota	15	Arica	01
5	General Lagos	15202	Parinacota	152	De Arica y Parinacota	15	Arica	01
6	Iquique	01101	Iquique	011	De Tarapacá	01	Iquique	02
7	Camiña	01402	Tamarugal	014	De Tarapacá	01	Iquique	02
8	Colchane	01403	Tamarugal	014	De Tarapacá	01	Iquique	02
9	Huara	01404	Tamarugal	014	De Tarapacá	01	Iquique	02
10	Pica	01405	Tamarugal	014	De Tarapacá	01	Iquique	02
11	Pozo Almonte	01401	Tamarugal	014	De Tarapacá	01	Iquique	02
12	Alto Hospicio	01107	Iquique	011	De Tarapacá	01	Iquique	02
13	Antofagasta	02101	Antofagasta	021	De Antofagasta	02	Antofagasta	03
14	Mejillones	02102	Antofagasta	021	De Antofagasta	02	Antofagasta	03
15	Sierra Gorda	02103	Antofagasta	021	De Antofagasta	02	Antofagasta	03
16	Taltal	02104	Antofagasta	021	De Antofagasta	02	Antofagasta	03
17	Calama	02201	El Loa	022	De Antofagasta	02	Antofagasta	03
18	Ollagüe	02202	El Loa	022	De Antofagasta	02	Antofagasta	03
19	San Pedro de Atacama	02203	El Loa	022	De Antofagasta	02	Antofagasta	03
20	Tocopilla	02301	Tocopilla	023	De Antofagasta	02	Antofagasta	03
21	María Elena	02302	Tocopilla	023	De Antofagasta	02	Antofagasta	03

Figura 48 Archivo DPA\_acotado

20.2.2.2 ETL

A Continuación se Presentan las vistas de Logging y Step Metrics para cada Transformación que conforma el Job Transformación, para verificar el correcto funcionamiento de cada transformación que conforma el Job.

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

- 2014/10/14 00:45:37 - Version checker - OK
- 2014/10/14 00:46:27 - Spoon - Transformación abierta.
- 2014/10/14 00:46:27 - Spoon - Ejecutando transformación [Transformacion Dimension Cliente]...
- 2014/10/14 00:46:27 - Spoon - Se ha iniciado la ejecución de la transformación.
- 2014/10/14 00:46:28 - Transformacion Dimension Cliente - Iniciado despacho de la transformación [Transformacion Dimension Cliente]
- 2014/10/14 00:46:32 - DV Dimension Cliente.0 - Connected to database [tesisdv] (commit=1000)
- 2014/10/14 00:46:51 - Cliente.0 - Finished reading query, closing connection.
- 2014/10/14 00:46:51 - Cliente.0 - Procesamiento finalizado (I=65, O=0, R=0, W=65, U=0, E=0)
- 2014/10/14 00:46:51 - Ordenar ciudad.0 - Procesamiento finalizado (I=0, O=0, R=65, W=65, U=0, E=0)
- 2014/10/14 00:46:51 - Comuna-Provincia-Ciudad-Region.0 - Procesamiento finalizado (I=347, O=0, R=0, W=347, U=0, E=0)
- 2014/10/14 00:46:51 - Ordenar filas.0 - Procesamiento finalizado (I=0, O=0, R=347, W=347, U=0, E=0)
- 2014/10/14 00:46:52 - Unión por Clave.0 - Procesamiento finalizado (I=0, O=0, R=412, W=65, U=0, E=0)
- 2014/10/14 00:46:52 - Renombrar Campos.0 - Procesamiento finalizado (I=0, O=0, R=65, W=65, U=0, E=0)
- 2014/10/14 00:46:52 - DV Dimension Cliente.0 - Procesamiento finalizado (I=0, O=0, R=65, W=65, U=0, E=0)
- 2014/10/14 00:46:52 - Spoon - La transformación ha finalizado!!

Figura 49 Logging Dimensión Cliente

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	Comuna-Provincia-Ciudad-Region	0	0	347	347	0	0	0	0	Finalizado
2	Cliente	0	0	65	65	0	0	0	0	Finalizado
3	Ordenar filas	0	347	347	0	0	0	0	0	Finalizado
4	Ordenar ciudad	0	65	65	0	0	0	0	0	Finalizado
5	Unión por Clave	0	412	65	0	0	0	0	0	Finalizado
6	Renombrar Campos	0	65	65	0	0	0	0	0	Finalizado
7	DV Dimension Cliente	0	65	65	0	65	0	0	0	Finalizado

**Figura 50 Step Metrics Dimensión Cliente**

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 00:48:41 - Spoon - Transformación abierta.  
 2014/10/14 00:48:41 - Spoon - Ejecutando transformación [Transformacion Dimension Producto]...  
 2014/10/14 00:48:41 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 00:48:41 - Transformacion Dimension Producto - Iniciado despacho de la transformación [Transformacion Dimension Producto]  
 2014/10/14 00:48:41 - Dimension Producto DV.0 - Connected to database [tesisdv] (commit=1000)  
 2014/10/14 00:48:42 - Familia.0 - Finished reading query, closing connection.  
 2014/10/14 00:48:42 - Familia.0 - Procesamiento finalizado (I=8, O=0, R=0, W=8, U=0, E=0  
 2014/10/14 00:48:42 - Rubro.0 - Finished reading query, closing connection.  
 2014/10/14 00:48:42 - Rubro.0 - Procesamiento finalizado (I=44, O=0, R=0, W=44, U=0, E=0  
 2014/10/14 00:48:42 - Renombrar id Familia.0 - Procesamiento finalizado (I=0, O=0, R=8, W=8, U=0, E=0  
 2014/10/14 00:48:42 - Renombrar id Rubro.0 - Procesamiento finalizado (I=0, O=0, R=44, W=44, U=0, E=0  
 2014/10/14 00:48:42 - Marca.0 - Finished reading query, closing connection.  
 2014/10/14 00:48:42 - Producto.0 - Finished reading query, closing connection.  
 2014/10/14 00:48:42 - Marca.0 - Procesamiento finalizado (I=2, O=0, R=0, W=2, U=0, E=0  
 2014/10/14 00:48:42 - Producto.0 - Procesamiento finalizado (I=736, O=0, R=0, W=736, U=0, E=0  
 2014/10/14 00:48:42 - Renombrar id Marca.0 - Procesamiento finalizado (I=0, O=0, R=2, W=2, U=0, E=0  
 2014/10/14 00:48:42 - ordenar id Familia.0 - Procesamiento finalizado (I=0, O=0, R=736, W=736, U=0, E=0  
 2014/10/14 00:48:42 - Union Familia.0 - Procesamiento finalizado (I=0, O=0, R=744, W=736, U=0, E=0  
 2014/10/14 00:48:42 - Ordenar id Rubro.0 - Procesamiento finalizado (I=0, O=0, R=736, W=736, U=0, E=0  
 2014/10/14 00:48:43 - Union Rubro.0 - Procesamiento finalizado (I=0, O=0, R=780, W=736, U=0, E=0  
 2014/10/14 00:48:43 - Ordenar id Marca.0 - Procesamiento finalizado (I=0, O=0, R=736, W=736, U=0, E=0  
 2014/10/14 00:48:43 - Union Marca.0 - Procesamiento finalizado (I=0, O=0, R=738, W=736, U=0, E=0  
 2014/10/14 00:48:43 - Renombrar Campos.0 - Procesamiento finalizado (I=0, O=0, R=736, W=736, U=0, E=0  
 2014/10/14 00:48:44 - Dimension Producto DV.0 - Procesamiento finalizado (I=0, O=736, R=736, W=736, U=0, E=0  
 2014/10/14 00:48:44 - Spoon - La transformación ha finalizado!!

**Figura 51 Step Metrics Dimensión Producto**

### Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 00:49:48 - Spoon - Transformación abierta.  
 2014/10/14 00:49:48 - Spoon - Ejecutando transformación [TransformacionDimensionTiempo]...  
 2014/10/14 00:49:48 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 00:49:48 - TransformacionDimensionTiempo - Iniciado despacho de la transformación [TransformacionDimensionTiempo]  
 2014/10/14 00:49:48 - Salida Tabla.0 - Connected to database [tesisdv] (commit=1000)  
 2014/10/14 00:49:49 - Generar Fecha.0 - Procesamiento finalizado (I=0, O=0, R=0, W=10000, U=0, E=0)  
 2014/10/14 00:49:49 - Añadir\_1\_dia.0 - Procesamiento finalizado (I=0, O=0, R=10000, W=10000, U=0, E=0)  
 2014/10/14 00:49:49 - Calculo Resto Atributos.0 - Optimization level set to 9.  
 2014/10/14 00:49:50 - Calcular Fechas.0 - Procesamiento finalizado (I=0, O=0, R=10000, W=10000, U=0, E=0)  
 2014/10/14 00:49:50 - Filter rows.0 - Procesamiento finalizado (I=0, O=0, R=10000, W=10000, U=0, E=0)  
 2014/10/14 00:49:50 - Calculo de Textos.0 - Procesamiento finalizado (I=0, O=0, R=10000, W=10000, U=0, E=0)  
 2014/10/14 00:49:51 - Calculo Resto Atributos.0 - Procesamiento finalizado (I=0, O=0, R=10000, W=10000, U=0, E=0)  
 2014/10/14 00:49:51 - Ordena Valores.0 - Procesamiento finalizado (I=0, O=0, R=10000, W=10000, U=0, E=0)  
 2014/10/14 00:49:54 - Salida Tabla.0 - Procesamiento finalizado (I=0, O=10000, R=10000, W=10000, U=0, E=0)  
 2014/10/14 00:49:54 - Spoon - La transformación ha finalizado!!

Figura 52 Logging Dimensión Fecha

### Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	Generar Fecha	0	0	10000	0	0	0	0	0	Finalizado
2	Añadir_1_dia	0	10000	10000	0	0	0	0	0	Finalizado
3	Calcular Fechas	0	10000	10000	0	0	0	0	0	Finalizado
4	Filter rows	0	10000	10000	0	0	0	0	0	Finalizado
5	Calculo de Textos	0	10000	10000	0	0	0	0	0	Finalizado
6	Calculo Resto Atributos	0	10000	10000	0	0	0	0	0	Finalizado
7	Ordena Valores	0	10000	10000	0	0	0	0	0	Finalizado
8	Dummy (do nothing)	0	0	0	0	0	0	0	0	Finalizado
9	Salida Tabla	0	10000	10000	0	10000	0	0	0	Finalizado

Figura 53 Step Metrics Dimensión Fecha

### Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 00:51:00 - Spoon - Transformación abierta.  
 2014/10/14 00:51:00 - Spoon - Ejecutando transformación [Transformacion Tabla Hechos]...  
 2014/10/14 00:51:00 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 00:51:00 - Transformacion Tabla Hechos - Iniciado despacho de la transformación [Transformacion Tabla Hechos]  
 2014/10/14 00:51:00 - Hechos Ventas.0 - Connected to database [tesisdv] (commit=1000)  
 2014/10/14 00:51:01 - Producto.0 - Finished reading query, closing connection.  
 2014/10/14 00:51:01 - Producto.0 - Procesamiento finalizado (I=736, O=0, R=0, W=736, U=0, E=0)  
 2014/10/14 00:51:01 - Renombrar id producto.0 - Procesamiento finalizado (I=0, O=0, R=736, W=736, U=0, E=0)  
 2014/10/14 00:51:01 - venta\_detalle.0 - Finished reading query, closing connection.  
 2014/10/14 00:51:01 - venta\_detalle.0 - Procesamiento finalizado (I=61, O=0, R=0, W=61, U=0, E=0)  
 2014/10/14 00:51:01 - Ordenar id\_producto.0 - Procesamiento finalizado (I=0, O=0, R=61, W=61, U=0, E=0)  
 2014/10/14 00:51:01 - Unión por Clave.0 - Procesamiento finalizado (I=0, O=0, R=797, W=61, U=0, E=0)  
 2014/10/14 00:51:01 - seleccionar campos.0 - Procesamiento finalizado (I=0, O=0, R=61, W=61, U=0, E=0)  
 2014/10/14 00:51:01 - Calculadora.0 - Procesamiento finalizado (I=0, O=0, R=61, W=61, U=0, E=0)  
 2014/10/14 00:51:01 - eliminar campos.0 - Procesamiento finalizado (I=0, O=0, R=61, W=61, U=0, E=0)  
 2014/10/14 00:51:01 - Hechos Ventas.0 - Procesamiento finalizado (I=0, O=61, R=61, W=61, U=0, E=0)  
 2014/10/14 00:51:01 - Spoon - La transformación ha finalizado!!

Figura 54 Logging Hechos

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	venta_detalle	0	0	61	61	0	0	0	0	Finalizado
2	Producto	0	0	736	736	0	0	0	0	Finalizado
3	Ordenar id_producto	0	61	61	0	0	0	0	0	Finalizado
4	Renombrar id_producto	0	736	736	0	0	0	0	0	Finalizado
5	Unión por Clave	0	797	61	0	0	0	0	0	Finalizado
6	seleccionar campos	0	61	61	0	0	0	0	0	Finalizado
7	Calculadora	0	61	61	0	0	0	0	0	Finalizado
8	eliminar campos	0	61	61	0	0	0	0	0	Finalizado
9	Hechos Ventas	0	61	61	0	61	0	0	0	Finalizado

**Figura 55 Step Metrics Hechos**

### 20.2.2.3 Tablas Finales

Tras la Ejecución del Proceso ETL de Transformación se debe verificar que los datos limpiados, y transformados para ser procesados como las dimensiones y hechos del DataMart, hayan sido cargados exitosamente a la base de datos Intermedia “tesis\_dv”.

dimension\_cliente x

1 • `SELECT * FROM tesis_dv.dimension_cliente;`

Result Set Filter:  Export: Wrap Cell Content:

	id_cliente	ciudad	comuna	provincia	region	id_region
▶	60	Bulnes	Bulnes	Ñuble	Del Bío Bío	08
	1	Chillan	Chillan	Ñuble	Del Bío Bío	08
	2	Chillan	Chillan	Ñuble	Del Bío Bío	08
	3	Chillan	Chillan	Ñuble	Del Bío Bío	08
	4	Chillan	Chillan	Ñuble	Del Bío Bío	08
	5	Chillan	Chillan	Ñuble	Del Bío Bío	08
	6	Chillan	Chillan	Ñuble	Del Bío Bío	08
	7	Chillan	Chillan	Ñuble	Del Bío Bío	08
	8	Chillan	Chillan	Ñuble	Del Bío Bío	08
	9	Chillan	Chillan	Ñuble	Del Bío Bío	08
	10	Chillan	Chillan	Ñuble	Del Bío Bío	08
	11	Chillan	Chillan	Ñuble	Del Bío Bío	08
	12	Chillan	Chillan	Ñuble	Del Bío Bío	08

dimension\_cliente 2 x

Output

Action Output

	Time	Action	Message
✓	1 01:19:25	SELECT * FROM tesis_dv.dimension_cliente	65 row(s) returned

**Figura 56 Dimensión Cliente**

dimension\_producto x

```
1 • SELECT * FROM tesis_dv.dimension_producto;
```

Result Set Filter: [ ] Export: [ ] Wrap Cell Content: [ ]

id_producto	id_familia	id_rubro	id_marca	stock	costo	familia_nombre	rubro_nombre	marca_nombre	nombre_producto
1	1	1	1	200	3000	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	01512 23 EMB. 26 CTOS. PTA BLANCA LEGRAND
266	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	680613 CAJA HIDROBOX 3 MOD. LEGRAND
267	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	689030 CAJA UNIVERSAL CHUQUI LEGRAND
287	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	16205 CAJA DERIVACION164X140X70 BTICINO
291	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25401 CAJA 1 PUESTO S/P IP40 BTICINO
292	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25402 CAJA 2 PUESTO S/P IP40 BTICINO
293	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25403 CAJA 3 PUESTO S/P IP40 BTICINO
294	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25404 CAJA 4 PUESTO S/P IP40 BTICINO
295	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25404/2V CAJA 4 PUESTO S/P IP40 BTICINO
296	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25406 CAJA 6 PUESTO S/P IP40 BTICINO
297	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25406V CAJA 6 PUESTO S/P IP40 BTICINO
298	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (TABLEROS)	LEGRAND	25408 CAJA 8 PUESTO S/P IP40 BTICINO

dimension\_producto 1 x Read Only

Output

Time	Action	Message	Duration / Fetch
1 01:19:59	SELECT * FROM tesis_dv.dimension_producto	736 row(s) returned	0.000 sec / 0.000 sec

Figura 57 Dimensión Producto

dimension\_tiempo x

```
1 • SELECT * FROM tesis_dv.dimension_tiempo;
```

Result Set Filter: [ ] Export: [ ] Wrap Cell Content: [ ] Fetch rows: [ ]

id_fecha	dia_nombre	dia	semana	semana_nombre	mes	mes_nombre	trimestre	trimestre_nombre	anio
1995-07-04 00:00:00	martes	4	27	Semana 27	7	julio	3	3 Trimestre	1995
1995-07-05 00:00:00	miércoles	5	27	Semana 27	7	julio	3	3 Trimestre	1995
1995-07-06 00:00:00	jueves	6	27	Semana 27	7	julio	3	3 Trimestre	1995
1995-07-07 00:00:00	viernes	7	27	Semana 27	7	julio	3	3 Trimestre	1995
1995-07-08 00:00:00	sábado	8	27	Semana 27	7	julio	3	3 Trimestre	1995
1995-07-09 00:00:00	domingo	9	27	Semana 27	7	julio	3	3 Trimestre	1995
1995-07-10 00:00:00	lunes	10	28	Semana 28	7	julio	3	3 Trimestre	1995
1995-07-11 00:00:00	martes	11	28	Semana 28	7	julio	3	3 Trimestre	1995
1995-07-12 00:00:00	miércoles	12	28	Semana 28	7	julio	3	3 Trimestre	1995
1995-07-13 00:00:00	jueves	13	28	Semana 28	7	julio	3	3 Trimestre	1995
1995-07-14 00:00:00	viernes	14	28	Semana 28	7	julio	3	3 Trimestre	1995
1995-07-15 00:00:00	sábado	15	28	Semana 28	7	julio	3	3 Trimestre	1995
1995-07-16 00:00:00	domingo	16	28	Semana 28	7	julio	3	3 Trimestre	1995

dimension\_tiempo 1 x

Output

Time	Action	Message
1 01:22:00	SELECT * FROM tesis_dv.dimension_tiempo	10000 row(s) returned

Figura 58 Dimensión Fecha

The screenshot shows a SQL query execution interface. The query executed is `SELECT * FROM tesis_dv.hechos_ventas;`. The result set contains 61 rows of sales data. Below the table, the 'Action Output' section shows a successful execution at 01:22:29, returning 61 rows.

	id_venta	id_fecha	id_producto	id_cliente	monto_venta	costo_venta	ganancia_venta	cantidad
▶	1	2014-06-06 00:00:00	1	1	18000	26000	10000	2
	22	2014-04-06 00:00:00	1	55	18000	39000	15000	3
	32	2014-01-06 00:00:00	1	55	18000	39000	15000	3
	42	2014-02-06 00:00:00	1	55	18000	39000	15000	3
	52	2014-05-06 00:00:00	1	55	18000	39000	15000	3
	2	2014-06-06 00:00:00	2	55	18000	26000	10000	2
	3	2014-06-06 00:00:00	2	59	18000	26000	10000	2
	4	2014-06-06 00:00:00	5	60	18000	26000	10000	2
	5	2014-06-06 00:00:00	5	61	18000	26000	10000	2
	23	2014-04-06 00:00:00	5	36	18000	26000	10000	2
	33	2014-01-06 00:00:00	5	36	18000	26000	10000	2
	43	2014-02-06 00:00:00	5	36	18000	26000	10000	2
	53	2014-05-06 00:00:00	5	36	18000	26000	10000	2
	-	-	-	-	-	-	-	-

Output

	Time	Action	Message
✓	1 01:22:29	SELECT * FROM tesis_dv.hechos_ventas	61 row(s) returned

**Figura 59 Hechos Ventas**

Tras Revisar las tablas en la base de datos tras la limpieza y formación de las dimensiones y hechos (“tesis\_dv”), se puede comprobar la ejecución exitosa del proceso ETL.

## 20.3 Caso de Prueba Job Carga

### 20.3.1 Caso de Prueba 3

Caso de Prueba N°3				
Objeto a analizar del caso de prueba:				
Procesos de Transformación.				
Objetivo del caso de prueba:				
Verificar que el proceso de Carga de las dimensiones y hechos se realice de manera exitosa.				
Pre-requisitos del caso de prueba:				
<ul style="list-style-type: none"> <li>- Disponibilidad de la base de datos "tesis_dv".</li> <li>- DataMart Creado</li> <li>- El Job carga debe estar creado y listo para usarse.</li> </ul>				
Datos de Entrada:				
<ul style="list-style-type: none"> <li>- Tabla dimensión_cliente</li> <li>- Tabla dimensión_producto</li> <li>- Tabla dimensión_fecha</li> <li>- Tabla hechos_ventas</li> </ul>				
N°	Función a probar	Acción	Resultado Esperado	Resultado Obtenido
1	Ejecución Transformación, y verificación que los datos hayan sido traspasados luego del proceso al DM.	Ejecutar el Job Carga	La Transformación se ejecuta exitosamente y las tablas generadas con éxito.	Ejecución Sin Problemas del Job.

**Tabla 20 Caso de Prueba 3**

### 20.3.2 Ejecución Caso de Prueba

#### 20.3.2.1 Datos de Entrada

Los datos de entrada, son los mostrados anteriormente que fueron procesados y limpiados y almacenados en la base de datos "tesis\_dv", para ser cargadas al DM definitivo.

#### 20.3.2.2 ETL

A Continuación se Presentan las vistas de Logging y Step Metrics para cada Transformación que conforma el Job Carga, para verificar el correcto funcionamiento de cada Transformación que conforma el Job.

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 01:45:19 - Spoon - Transformación abierta.  
 2014/10/14 01:45:19 - Spoon - Ejecutando transformación [Carga Dim Cliente]...  
 2014/10/14 01:45:19 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 01:45:19 - Carga Dim Cliente - Iniciado despacho de la transformación [Carga Dim Cliente]  
 2014/10/14 01:45:19 - Cliente.0 - Finished reading query, closing connection.  
 2014/10/14 01:45:19 - Cliente.0 - Procesamiento finalizado (I=65, O=0, R=0, W=65, U=0, E=0  
 2014/10/14 01:45:19 - Dimension Cliente.0 - Procesamiento finalizado (I=65, O=0, R=65, W=65, U=0, E=0  
 2014/10/14 01:45:19 - Spoon - La transformación ha finalizado!!

Figura 60 Logging Dimensión Cliente

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	Cliente	0	0	65	65	0	0	0	0	Finalizado
2	Dimension Cliente	0	65	65	65	0	0	0	0	Finalizado

Figura 61 Step Metrics Dimensión Cliente

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 01:35:43 - Spoon - Transformación abierta.  
 2014/10/14 01:35:43 - Spoon - Ejecutando transformación [Carga Dim Producto]...  
 2014/10/14 01:35:43 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 01:35:43 - Carga Dim Producto - Iniciado despacho de la transformación [Carga Dim Producto]  
 2014/10/14 01:35:43 - Producto.0 - Finished reading query, closing connection.  
 2014/10/14 01:35:43 - Producto.0 - Procesamiento finalizado (I=736, O=0, R=0, W=736, U=0, E=0  
 2014/10/14 01:35:44 - Dimension Producto.0 - Procesamiento finalizado (I=736, O=0, R=736, W=736, U=0, E=0  
 2014/10/14 01:35:44 - Spoon - La transformación ha finalizado!!

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	Producto	0	0	736	736	0	0	0	0	Finalizado
2	Dimension Producto	0	736	736	736	0	0	0	0	Finalizado

Figura 62 Step Metrics Dimensión Producto

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 01:30:33 - Spoon - Transformación abierta.  
 2014/10/14 01:30:33 - Spoon - Ejecutando transformación [Carga Dim Tiempo]...  
 2014/10/14 01:30:33 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 01:30:33 - Carga Dim Tiempo - Iniciado despacho de la transformación [Carga Dim Tiempo]  
 2014/10/14 01:30:33 - fecha.0 - Finished reading query, closing connection.  
 2014/10/14 01:30:33 - fecha.0 - Procesamiento finalizado (I=10000, O=0, R=0, W=10000, U=0, E=0  
 2014/10/14 01:30:38 - Dimension Fecha.0 - Procesamiento finalizado (I=10000, O=0, R=10000, W=10000, U=0, E=0  
 2014/10/14 01:30:38 - Spoon - La transformación ha finalizado!!

Figura 63 Logging Dimensión Producto

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 01:30:33 - Spoon - Transformación abierta.  
 2014/10/14 01:30:33 - Spoon - Ejecutando transformación [Carga Dim Tiempo]...  
 2014/10/14 01:30:33 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 01:30:33 - Carga Dim Tiempo - Iniciado despacho de la transformación [Carga Dim Tiempo]  
 2014/10/14 01:30:33 - fecha.0 - Finished reading query, closing connection.  
 2014/10/14 01:30:33 - fecha.0 - Procesamiento finalizado (I=10000, O=0, R=0, W=10000, U=0, E=0  
 2014/10/14 01:30:38 - Dimension Fecha.0 - Procesamiento finalizado (I=10000, O=0, R=10000, W=10000, U=0, E=0  
 2014/10/14 01:30:38 - Spoon - La transformación ha finalizado!!

Figura 64 Logging Dimensión Fecha

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	fecha	0	0	10000	10000	0	0	0	0	Finalizado
2	Dimension Fecha	0	10000	10000	10000	0	0	0	0	Finalizado

Figura 65 Step Metrics Dimensión Fecha

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

2014/10/14 01:46:20 - Spoon - Transformación abierta.  
 2014/10/14 01:46:20 - Spoon - Ejecutando transformación [Carga Hechos Ventas]...  
 2014/10/14 01:46:20 - Spoon - Se ha iniciado la ejecución de la transformación.  
 2014/10/14 01:46:20 - Carga Hechos Ventas - Iniciado despacho de la transformación [Carga Hechos Ventas]  
 2014/10/14 01:46:20 - Hechos.0 - Finished reading query, closing connection.  
 2014/10/14 01:46:20 - Hechos.0 - Procesamiento finalizado (I=61, O=0, R=0, W=61, U=0, E=0  
 2014/10/14 01:46:20 - Hechos Ventas.0 - Procesamiento finalizado (I=61, O=0, R=61, W=61, U=0, E=0  
 2014/10/14 01:46:20 - Spoon - La transformación ha finalizado!!

Figura 66 Logging Hechos Ventas

**Execution Results**

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	Hechos	0	0	61	61	0	0	0	0	Finalizado
2	Hechos Ventas	0	61	61	61	0	0	0	0	Finalizado

**Figura 67 Steps Metrics Hechos Ventas**

Las Imágenes anteriores nos verifican el correcto funcionamiento de los procesos ETL asociados a la carga de las dimensiones y la tabla de hechos hacia el DataMart, ahora solo queda verificar los datos en las tablas.

### 20.3.2.3 Tablas Finales

Tras la ejecución del Proceso ETL de Carga, se deben verificar que las tablas han sido replicadas en el DM, esto lo podemos comprobar observando en el Motor del DM (MySQL).

dimension\_cliente x

```
SELECT * FROM tesis_datamart_venta.dimension_cliente;
```

Result Set Filter: [ ] Edit: [ ] Export/Import: [ ] Wrap Cell Content: [ ]

	id_cliente_sk	version	date_from	date_to	id_cliente	ciudad	comuna	provincia	region	id_region
▶	0	1	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
	1	1	2014-07-09 21:14:19	2200-01-01 00:00:00	60	Bulnes	Bulnes	Ñuble	Del BíoBío	08
	2	1	2014-07-09 21:14:19	2200-01-01 00:00:00	1	Chillan	Chillan	Ñuble	Del BíoBío	08
	3	1	2014-07-09 21:14:19	2200-01-01 00:00:00	2	Chillan	Chillan	Ñuble	Del BíoBío	08
	4	1	2014-07-09 21:14:19	2200-01-01 00:00:00	3	Chillan	Chillan	Ñuble	Del BíoBío	08
	5	1	2014-07-09 21:14:19	2200-01-01 00:00:00	4	Chillan	Chillan	Ñuble	Del BíoBío	08
	6	1	2014-07-09 21:14:19	2200-01-01 00:00:00	5	Chillan	Chillan	Ñuble	Del BíoBío	08
	7	1	2014-07-09 21:14:19	2200-01-01 00:00:00	6	Chillan	Chillan	Ñuble	Del BíoBío	08
	8	1	2014-07-09 21:14:19	2200-01-01 00:00:00	7	Chillan	Chillan	Ñuble	Del BíoBío	08
	9	1	2014-07-09 21:14:19	2200-01-01 00:00:00	8	Chillan	Chillan	Ñuble	Del BíoBío	08
	10	1	2014-07-09 21:14:19	2200-01-01 00:00:00	9	Chillan	Chillan	Ñuble	Del BíoBío	08
	11	1	2014-07-09 21:14:19	2200-01-01 00:00:00	10	Chillan	Chillan	Ñuble	Del BíoBío	08
	12	1	2014-07-09 21:14:19	2200-01-01 00:00:00	11	Chillan	Chillan	Ñuble	Del BíoBío	08

dimension\_cliente 2 x

Output

Action Output

	Time	Action	Message
✓	1 01:53:19	SELECT * FROM tesis_datamart_venta.dimension_cliente	66 row(s) returned

**Figura 68 Dimensión Cliente**

dimension\_producto

1 • SELECT \* FROM tesis\_datamart\_venta.dimension\_producto;

Result Set Filter: [ ] Edit: [ ] Export/Import: [ ] Wrap Cell Content: [ ]

id_producto_sk	version	date_from	date_to	id_producto	id_familia	id_rubro	id_marca	stock	costo	familia_nombre	rubro_nombre
0	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1	1	1	1	200	3000	PRODUCTOS DOMESTICOS	CAJAS (T/
2	1	1900-01-01 00:00:00	2200-01-01 00:00:00	266	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
3	1	1900-01-01 00:00:00	2200-01-01 00:00:00	267	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
4	1	1900-01-01 00:00:00	2200-01-01 00:00:00	287	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
5	1	1900-01-01 00:00:00	2200-01-01 00:00:00	291	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
6	1	1900-01-01 00:00:00	2200-01-01 00:00:00	292	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
7	1	1900-01-01 00:00:00	2200-01-01 00:00:00	293	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
8	1	1900-01-01 00:00:00	2200-01-01 00:00:00	294	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
9	1	1900-01-01 00:00:00	2200-01-01 00:00:00	295	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
10	1	1900-01-01 00:00:00	2200-01-01 00:00:00	296	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
11	1	1900-01-01 00:00:00	2200-01-01 00:00:00	297	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/
12	1	1900-01-01 00:00:00	2200-01-01 00:00:00	298	1	1	1	200	7200	PRODUCTOS DOMESTICOS	CAJAS (T/

dimension\_producto 2 x [Apply] [Cancel]

Output

Action Output

Time	Action	Message	Duration / Fetch
1 01:53:50	SELECT * FROM tesis_datamart_venta.dimension_producto	737 row(s) returned	0.000 sec / 0.000 sec

Figura 69 Dimension Producto

dimension\_tiempo

1 • SELECT \* FROM tesis\_datamart\_venta.dimension\_tiempo;

Result Set Filter: [ ] Edit: [ ] Export/Import: [ ] Wrap Cell Content: [ ] Fetch rows: [ ]

id_fecha_sk	version	date_from	date_to	id_fecha	dia_nombre	dia	semana	semana_nombre	mes	mes_nombre	trimestre	t
0	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-04 00:00:00	martes	4	27	Semana 27	7	julio	3	3
1	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-05 00:00:00	miércoles	5	27	Semana 27	7	julio	3	3
2	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-06 00:00:00	jueves	6	27	Semana 27	7	julio	3	3
3	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-07 00:00:00	viernes	7	27	Semana 27	7	julio	3	3
4	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-08 00:00:00	sábado	8	27	Semana 27	7	julio	3	3
5	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-09 00:00:00	domingo	9	27	Semana 27	7	julio	3	3
6	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-10 00:00:00	lunes	10	28	Semana 28	7	julio	3	3
7	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-11 00:00:00	martes	11	28	Semana 28	7	julio	3	3
8	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-12 00:00:00	miércoles	12	28	Semana 28	7	julio	3	3
9	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-13 00:00:00	jueves	13	28	Semana 28	7	julio	3	3
10	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-14 00:00:00	viernes	14	28	Semana 28	7	julio	3	3
11	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-15 00:00:00	sábado	15	28	Semana 28	7	julio	3	3
12	1	1900-01-01 00:00:00	2200-01-01 00:00:00	1995-07-16 00:00:00	domingo	16	28	Semana 28	7	julio	3	3

dimension\_tiempo 2 x [Apply] [Cancel]

Output

Action Output

Time	Action	Message	Duration / Fetch
1 01:54:20	SELECT * FROM tesis_datamart_venta.dimension_tiempo	10001 row(s) returned	0.000 sec / 0.015 sec

Figura 70 Dimensión Fecha

The screenshot shows a SQL query execution interface. The query executed is `SELECT * FROM tesis_datamart_venta.hechos_ventas;`. The result set is a table with 8 columns: `id_venta`, `id_fecha`, `id_producto`, `id_cliente`, `monto_venta`, `costo_venta`, `ganancia_venta`, and `cantidad`. The table contains 15 rows of data. Below the table, the 'Output' section shows a log entry: '1 01:54:45 SELECT \* FROM tesis\_datamart\_venta.hechos\_ventas' with the message '61 row(s) returned'.

	id_venta	id_fecha	id_producto	id_cliente	monto_venta	costo_venta	ganancia_venta	cantidad
▶	1	2014-06-06 00:00:00	1	1	18000	26000	10000	2
	22	2014-04-06 00:00:00	1	55	18000	39000	15000	3
	32	2014-01-06 00:00:00	1	55	18000	39000	15000	3
	42	2014-02-06 00:00:00	1	55	18000	39000	15000	3
	52	2014-05-06 00:00:00	1	55	18000	39000	15000	3
	2	2014-06-06 00:00:00	2	55	18000	26000	10000	2
	3	2014-06-06 00:00:00	2	59	18000	26000	10000	2
	4	2014-06-06 00:00:00	5	60	18000	26000	10000	2
	5	2014-06-06 00:00:00	5	61	18000	26000	10000	2
	23	2014-04-06 00:00:00	5	36	18000	26000	10000	2
	33	2014-01-06 00:00:00	5	36	18000	26000	10000	2
	43	2014-02-06 00:00:00	5	36	18000	26000	10000	2
	53	2014-05-06 00:00:00	5	36	18000	26000	10000	2

**Figura 71 Hechos Ventas**

Tras Revisar las tablas en el DataMart, se puede comprobar la ejecución exitosa del proceso ETL, y las tablas replicadas son consistentes con las tablas fuentes.

## 20.4 Caso de Prueba Cubo OLAP

### 20.4.1 Caso de Prueba 4

Caso de Prueba N°4				
Objeto a analizar del caso de prueba:				
Servicio Cubo OLAP.				
Objetivo del caso de prueba:				
Verificar que el cubo de Venta muestre la misma información que la tabla "hechos_venta" del DataMart.				
Pre-requisitos del caso de prueba:				
<ul style="list-style-type: none"> <li>- Disponibilidad de la BD DataMart.</li> <li>- Disponibilidad del Cubo de Ventas.</li> </ul>				
Datos de Entrada:				
<ul style="list-style-type: none"> <li>- Realizar la Misma Consulta.</li> </ul>				
N°	Función a probar	Acción	Resultado Esperado	Resultado Obtenido
1	Valores del DataMart Frente a Valores del Cubo OLAP.	Realizar consulta al DataMart y al Cubo de Ventas.	El Cubo de ventas debe presentar la misma información que muestre el DataMart consultando la tabla de Hechos.	Muestra la Misma Información.

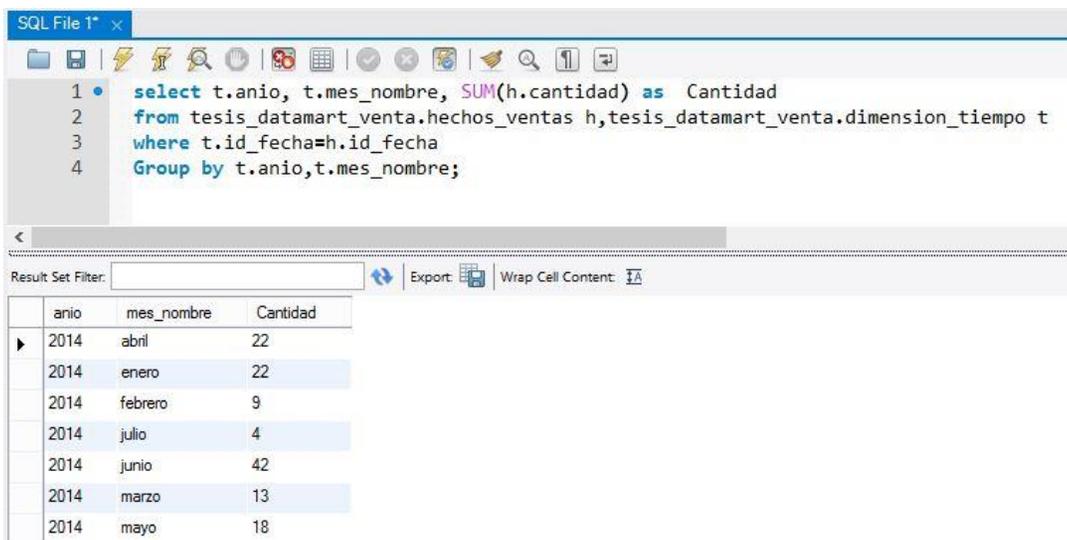
Tabla 21 Caso de Prueba 4

### 20.4.2 Ejecución del Caso de Prueba

#### 20.4.2.1 Primera Consulta a Realizar

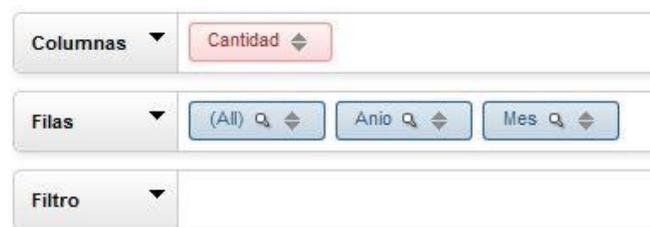
Mostrar la cantidad de ventas realizadas para cada periodo del 2014.

A Continuación se presenta una imagen donde aparece la consulta SQL y el Resultado del DataMart.



**Figura 72 consulta SQL DM cantidad vendidas 2014**

La Misma Consulta se efectúa al Cubo Ventas a través de la Herramienta Saiku, se presenta el resultado de la consulta.



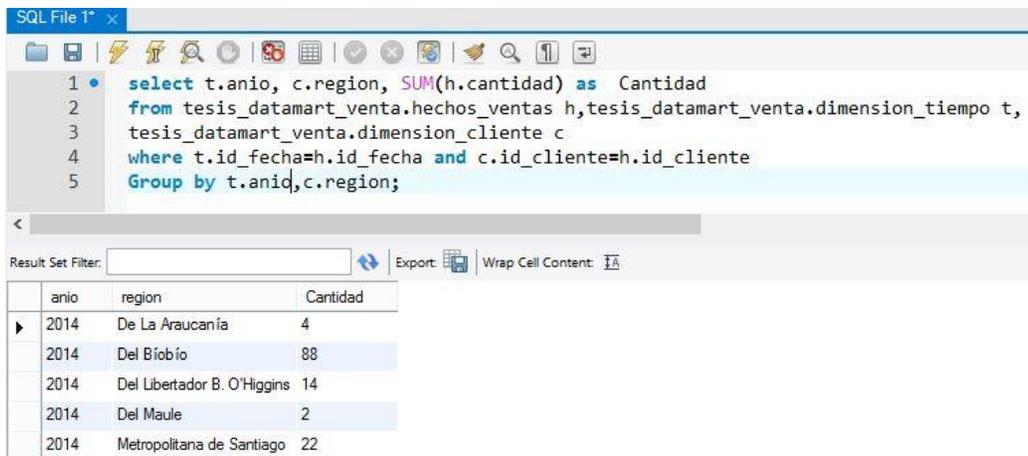
(All)	Anio	Mes	Cantidad
Todos	2014	abril	22
		enero	22
		febrero	9
		julio	4
		junio	42
		marzo	13
		mayo	18

**Figura 73 Consulta cubo OLAP cantidad vendidas 2014**

#### 20.4.2.2 Segunda Consulta a Realizar

Mostrar la Cantidad de Ventas por Región.

A continuación se presenta una imagen donde se observa la consulta SQL y el resultado del DataMart



**Figura 74 Consulta SQL cantidad ventas por región**

Tras observar los resultados del DataMart, se mostraran los resultados del Cubo OLAP, realizado la herramienta Saiku.



(All)	Anio	Region	Cantidad
Todos	2014	De La Araucanía	4
		Del Biobío	88
		Del Libertador B. O'Higgins	14
		Del Maule	2
		Metropolitana de Santiago	22

**Figura 75 Consulta cubo OLAP cantidad de ventas por región**

Tras Observar los resultados de las pruebas tanto en el DataMart como en el Cubo OLAP, se da por conclusión que los resultados son los esperados.

---

## 21 ANEXO: MANUAL DE CONFIGURACION

---

### 21.1 Cubo OLAP

Una vez ejecutados los procesos ETL, debe ser creado el cubo Dimensional, para los análisis OLAP, esto se realiza con la herramienta "Pentaho Schema Workbench", el cual genera un archivo XML, que el motor Mondrian puede interpretar. Luego el archivo XML debe ser cargado en el servidor de análisis a continuación se presenta la definición del cubo dimensional en XML.

```
<?xml version="1.0" encoding="UTF-8"?>
<Schema name="TesisDataMart">
<Dimension type="StandardDimension" highCardinality="true" name="Cliente"
caption="Cliente">
  <Hierarchy hasAll="true" allMemberName="Todos" allMemberCaption="Todos"
primaryKey="id_cliente">
    <Table name="dimension_cliente"/>
    <Level name="Region" column="region" ordinalColumn="region" uniqueMembers="true"
levelType="Regular" type="String"/>
    <Level name="Provincia" column="provincia" ordinalColumn="provincia"
uniqueMembers="true" levelType="Regular" type="String"/>
    <Level name="Comuna" column="comuna" ordinalColumn="comuna"
uniqueMembers="true" levelType="Regular" type="String"/>
    <Level name="Ciudad" column="ciudad" ordinalColumn="ciudad" nameColumn="ciudad"
uniqueMembers="true" levelType="Regular" type="String"/>
  </Hierarchy>
</Dimension>

<Dimension type="StandardDimension" highCardinality="true" name="Producto"
caption="Producto">
  <Hierarchy hasAll="true" allMemberName="Todos" allMemberCaption="Todos"
primaryKey="id_producto">
    <Table name="dimension_producto"/>
    <Level name="Familia" column="id_familia" ordinalColumn="id_familia"
nameColumn="familia_nombre" uniqueMembers="true" levelType="Regular"
type="Integer"/>
    <Level name="Rubro" column="id_rubro" ordinalColumn="id_rubro"
nameColumn="rubro_nombre" uniqueMembers="true" levelType="Regular"
type="Integer"/>
    <Level name="Marca" column="id_marca" ordinalColumn="id_marca"
nameColumn="marca_nombre" uniqueMembers="true" levelType="Regular"
type="Integer"/>
    <Level name="Producto" column="id_producto" ordinalColumn="id_producto"
nameColumn="nombre_producto" uniqueMembers="true" levelType="Regular"
type="Integer">
      <Property name="Stock" column="stock" type="Numeric"/>
    </Level>
  </Hierarchy>
</Dimension>
```

```

    <Property name="Costo" column="costo" type="Numeric" />
  </Level>
</Hierarchy>
</Dimension>

<Dimension type="TimeDimension" highCardinality="false" name="Fecha">
  <Hierarchy hasAll="true" allMemberName="Todos" allMemberCaption="Todos"
primaryKey="id_fecha">
    <Table name="dimension_tiempo" />
    <Level name="Anio" column="anio" uniqueMembers="true" levelType="TimeYears"
type="Integer" />
    <Level name="Mes" column="mes_nombre" uniqueMembers="false"
levelType="TimeMonths" type="String" />
    <Level name="dia" column="dia" uniqueMembers="false" levelType="TimeDays"
type="String" />
  </Hierarchy>
</Dimension>
<Cube name="Ventas" cache="false" enabled="true">
  <Table name="hechos_ventas" />
  <DimensionUsage source="Cliente" name="Cliente" foreignKey="id_cliente" />
  <DimensionUsage source="Fecha" name="Fecha" foreignKey="id_fecha" />
  <DimensionUsage source="Producto" name="Producto" foreignKey="id_producto" />

  <Measure name="Cantidad" column="cantidad" aggregator="sum" visible="true" />
  <Measure name="Costo Venta" column="costo_venta" aggregator="sum" visible="true" />
  <Measure name="Monto Venta" column="monto_venta" aggregator="sum"
visible="true" />
  <Measure name="Ganancia Venta" column="ganancia_venta" aggregator="sum"
visible="true" />
</Cube>

</Schema>

```

## 21.2 Configuración Pentaho Business Analytics

Pentaho Business Analytics Platform, es una herramienta que ofrece la suite de pentaho, para realizar los trabajos de análisis OLAP y cuadros de mando (Dashboard) utilizando el cubo dimensional creado anteriormente, esta aplicación corre en un servidor Tomcat (Java), esta herramienta está actualmente en la versión 5.0.1, la cual ha implementado considerable cambios con respecto a su anterior versión (4.8), esta versión 5.0.1 CE (Community Edition), será la que se utilizara a continuación.

Como se mencionaba anteriormente esta aplicación corre en un servidor Tomcat, ya que esta hecho en Java, para poder ejecutar esta herramienta, simplemente se descargar el archivo .rar desde la página de descargas, y se puede descomprimir el contenido en cualquier directorio, y luego ingresar a la carpeta “biserver-ce”, en la imagen se presenta el directorio en el cual tenemos las distintas herramientas pentaho para BI, en la imagen se aprecia la plataforma que utilizaremos.



**Figura 76 Directorio Pentaho**

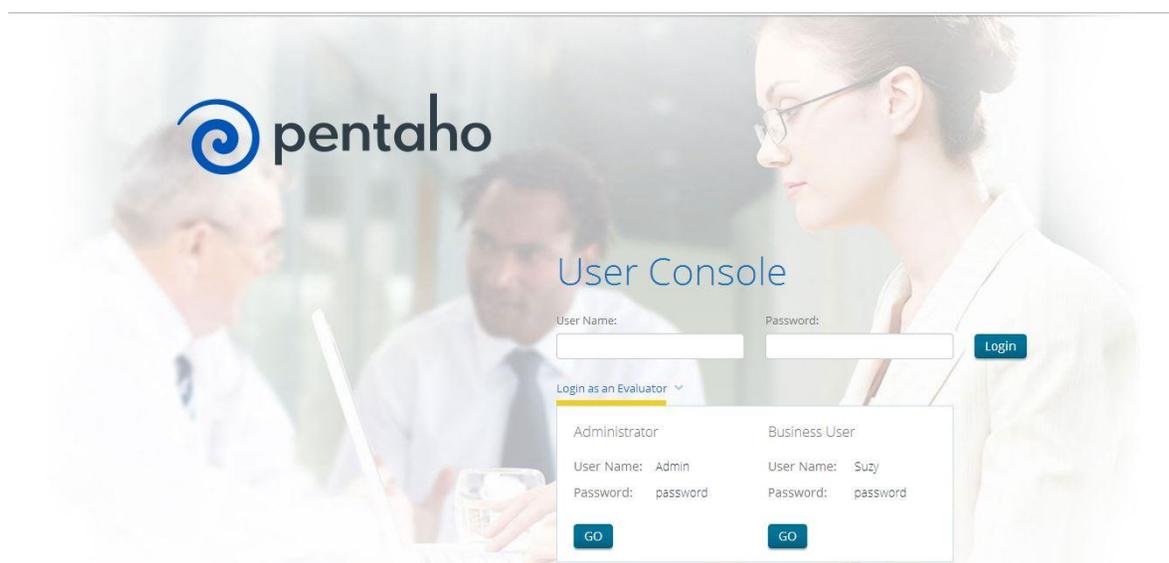
Luego de acceder al directorio mencionado antes se debe ejecutar el archivo “start-pentaho.bat”, este archivo es el que ejecuta las sentencias y los procesos para que el servidor pueda comenzar a correr en el equipo. Una vez que logra cargar todos los archivos necesarios, se puede ingresar al servidor a través de la siguiente url “http://localhost:8080/pentaho/Login “, de haber cargado correctamente todos los archivos, la página de inicio del servidor se mostrara a continuación, con su ventana de acceso de usuario.



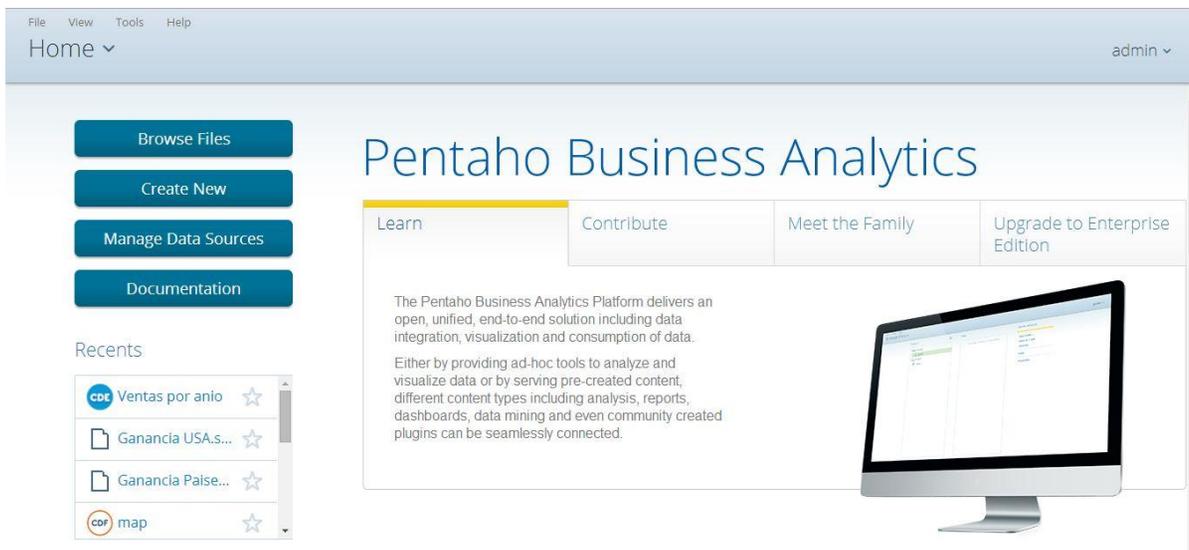
**Figura 77 Directorio Biserver-CE ubicación archivo "start-pentaho.bat"**

Se pueden ingresar utilizando los usuarios por defecto que trae pentaho, para luego ingresar al servidor y poder crear nuevos usuarios y permisos, aquí entraremos con el usuario administrador que posee todos los permisos, el cual se puede ingresar inmediatamente.

Una vez realizado el acceso de usuario, se presenta la interfaz del servidor. Lo primero que se debe realizar para poder utilizar las bondades de esta herramienta y realizar los cuadros de mando y análisis OLAP, ya que para la instalación recién descargada, necesita las herramientas para dicho trabajo, las cuales pueden instalarse directamente desde el servidor, por lo cual debemos ingresar al “Marketplace” para poder instalar los plugins necesarios.

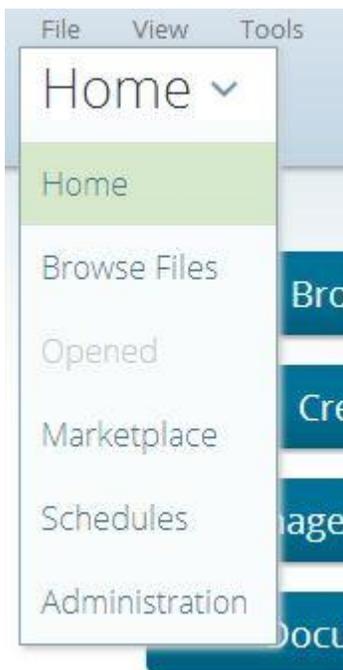


**Figura 78 Plataforma Pentaho BA Server de acceso y Credenciales de Acceso**



**Figura 79** Pantalla Principal Plataforma Pentaho BA Server

El Marketplace es una herramienta, la cual facilita la incorporación de plugins al servidor pentaho los que agregan funcionalidades al servidor, estos plugins son administrados tanto por la comunidad de pentaho que desarrollan dichos plugins, los que son necesarios para realizar este proyecto fueron los siguientes:



En esta figura de la izquierda se puede observar donde está la opción para poder ir al Marketplace para poder visualizar los plugins, mientras en la imagen de abajo, se puede observar el Marketplace, en la imagen se muestran los plugins instalados necesarios para la realización del proyecto, con información adicional en línea, indicando si están actualizados, hay una nueva versión.

<b>Pentaho Marketplace</b> Pentaho <b>Installed Version:</b> TRUNK-SNAPSHOT (TRUNK)	View details	Upgrade
<b>Community Dashboards Framework</b> WebDetails <b>Installed Version:</b> 14.03.07 (STABLE)	View details	Upgrade
<b>Community Data Access</b> WebDetails <b>Installed Version:</b> 14.03.07 (STABLE)	View details	Upgrade
<b>Community Dashboard Editor</b> WebDetails <b>Installed Version:</b> 14.03.07 (STABLE)	View details	Upgrade
<b>Community Graphics Generator</b> WebDetails <b>Installed Version:</b> 14.03.07 (STABLE)	View details	Upgrade
<b>Saiku Analytics</b> Analytical Labs <b>Installed Version:</b> 2.6-PENTAHO5-SNAPSHOT (TRUNK)	View details	Up to Date
<b>Saiku Chart Plus</b> IT4biz <b>Installed Version:</b> RC4 (TRUNK)	View details	Up to Date

**Figura 80 Pentaho Marketplace**

Además al seleccionar un Plugin se entrega una descripción del mismo, del autor, la función que cumple, la versión que se encuentra en línea, capturas de pantalla del plugin, la licencia y las dependencias que necesita el plugin. En la siguiente imagen se muestra un detalle del plugin CDF.

**Community Dashboards Framework**  
14.03.07 (STABLE)

Close

**INFO**

**Version**  
14.03.07

**Created by**  
WebDetails

**License**  
MPL 2.0

**Dependencies**

**DESCRIPTION**

Community Dashboard Framework (CDF) is the engine that allows you to create friendly, powerful and fully featured Dashboards on top of the Pentaho Business Intelligence server.

**SCREENSHOTS**

14.05.06 (Stable)

Install

Upgrade

Uninstall

**Figura 81 Detalle Plugin CDF**

### 21.2.1 Plugin Pentaho para crear Dashboard

Community Dashboard Editor, Community Dashboard Framework, Community Chart Components, Community Graphics Generator.

Estas herramientas son necesarias para la creación de Dashboards, entregan al desarrollador un marco de trabajo para creación y edición, una herramienta de visualización, y generación de fuentes de datos (consultas SQL, consultas MDX), gráficos, todo esto para poder realizar cuadros de mando sencillos con información global de la empresa.



**Figura 82 CDTTools para Cuadros de Mando – Creados por la Empresa Webdetails**

### 21.2.2 Plugin Pentaho Análisis OLAP

Saiku Analytics, Saiku Chart Plus.

Estas herramientas son necesarias para la visualización y creación de Análisis Olap, las cuales son consultas al cubo dimensional generado anteriormente, este plugin Saiku, se selecciona el Cubo que servirá como fuente, y se puede pasar a realizar consultas en el cubo a través de la opción Arrastrar y soltar, (medidas y dimensiones), el plugin inmediatamente realizara las consultas al cubo, mostrando los resultados, estas consultas realizas luego pueden guardarse para volverlas a ver en cualquier momento con los datos del cubo dimensional actualizados.

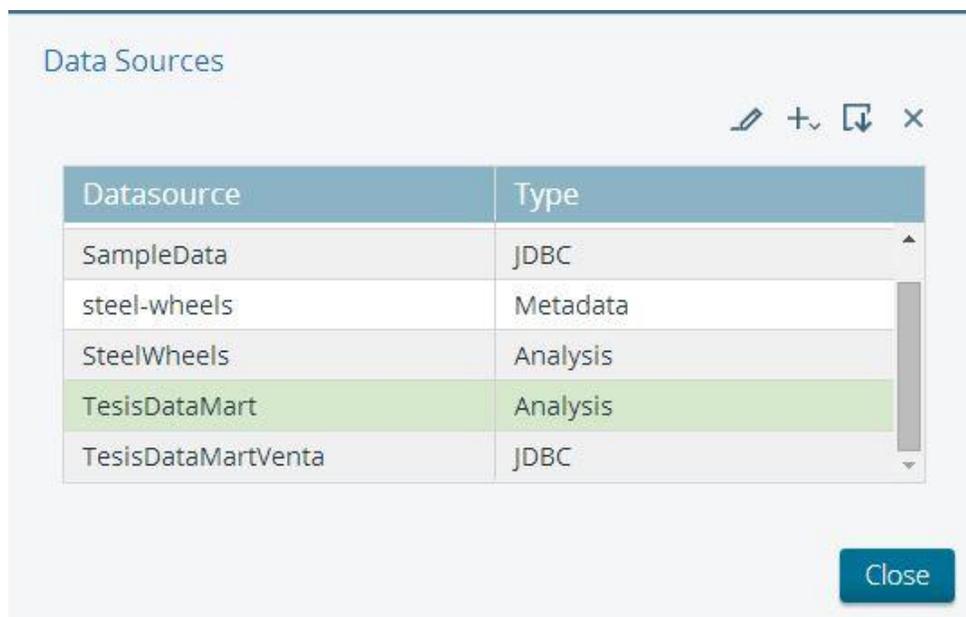
Saiku Chart Plus es un plugin para saiku el cual agrega nuevas opciones y mejoras para los gráficos de saiku, no es obligación usarlo, pero da la utilidad de usar mapas.



**Figura 83 Herramienta de Análisis Saiku**

Luego de tener los plugins necesarios, se necesita cargar nuestro cubo recién creado en el repositorio de Pentaho, ya que de no hacerlo no se podría tener acceso al cubo creado pues no sería visible. Entonces se debe importar nuestro cubo con extensión xml y seleccionar que es para análisis.

Para importar nuestro cubo debemos ir a la pantalla principal del servidor y seleccionar la opción “Manage Data Sources”, el servidor nos mostrara las data sources existentes, para este proyecto se crearan dos tipos de data sources que se muestran en la siguiente imagen: “JDBC” y “Analysis”



**Figura 84** Pantalla Data Sources

Se mostrará a continuación la configuración de las fuentes de datos:

TesisDataMartVenta

Este DataSource es del tipo JDBC, el cual nos proporcionara la conexión con la base de datos en la cual se encuentra nuestro DM de Ventas, esto nos proporcionara la facilidad de poder realizar todo tipo de consultas a nuestro DM, tanto desde Análisis OLAP, como desde los Cuadros de Mando y Reportes.

La configuración es la misma que se realiza con los DataSource en Pentaho Data Integration, se coloca un nombre a la conexión, host, nombre de la base de datos, puerto, nombre de usuario de la base de datos y la contraseña (en caso de tenerla), en este DataSource se configura una conexión nativa JDBC hacia el DM que se encuentra en una base de datos MySQL.

Database Connection

General  
Advanced  
Options  
Pooling

Connection Name:  
TesisDataMartVenta

Database Type:  
MySQL  
Hadoop Hive  
Hadoop Hive 2  
Hypersonic  
Impala  
MS SQL Server  
PostgreSQL  
H2  
MonetDB

Settings  
Host Name:  
localhost  
Database Name:  
tesis\_datamart\_venta  
Port Number:  
3306  
User Name:  
root  
Password:  
.....

Access:  
Native (JDBC)  
ODBC  
JNDI

[Adding Databases](#)

Test

OK Cancel

**Figura 85 Configuración DataSource JDBC**

### Tesis DataMart

Este DataSource es del tipo Analysis (OLAP), por lo cual en este paso lo que se realiza es la importación de nuestro Cubo Dimensional realizado anteriormente en Pentaho Schema Workbench, lo primero se selecciona el Cubo Mondrian con extensión XML, para indicar que es nuestro cubo dimensional, el segundo paso es indicarle cual es nuestro datasource del cual nuestro cubo consultara los datos, en el cual escogemos el DataSource JDBC creado anteriormente (TesisDataMartVenta), esto le dirá a nuestro servidor Pentaho, que nuestro cubo de ventas realizara las consultas a la base de datos en MySQL donde se encuentra el DM (Cabe recordar que Mondrian es un motor ROLAP, que interpreta nuestras consultas MDX un lenguaje dimensional en lenguaje SQL para consultar una base de datos relacional o columnar).

Al hacer clic en el botón “Save” estaremos guardando esta configuración y nuestro cubo estará importado, listo para comenzar el análisis dimensional.

The image shows a software dialog box titled "Import Analysis". It contains the following elements:

- A label "Mondrian File:" above a text input field containing "TesisDataMart.mondrian.xml".
- Two radio button options:
  - Select from available data sources.
  - Manually enter data source parameter values.
- A label "Data Source:" above a dropdown menu showing "TesisDataMartVenta".
- Two buttons at the bottom right: "Save" and "Close".

**Figura 86 Import Analysis**

Luego de haber importado el cubo, podemos comenzar con nuestro análisis del cubo dimensional con Saiku, para lo cual, nos iremos a “Create New” y escogemos “New Saiku Analytics”, inmediatamente se cargara el plugin saiku en una pestaña en nuestro servidor.