

**UNIVERSIDAD PRIVADA ANTENOR ORREGO  
FACULTAD DE INGENIERÍA**

**ESCUELA PROFESIONAL DE INGENIERÍA DE COMPUTACIÓN Y  
SISTEMAS**



---

**“SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA EMPRESAS DE SERVICIO  
DE IMPORTACIÓN Y EXPORTACIÓN DE CALZADO DEL DEPARTAMENTO LA  
LIBERTAD”**

---

**TESIS**

**LINEA DE INVESTIGACIÓN:**

Desarrollo de modelos, aplicaciones y arquitecturas de sistemas para toma de decisiones y  
dirección estratégica (CTI 0403 0703a)

**AUTORES:**

Br. Benites Robles, José Dennis Ivan

Br. López Girón, Stephany Carolina

**ASESOR:**

Ing. Ullón Ramirez Agustin Eduardo

**TRUJILLO**

**2015**

**“SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA EMPRESAS DE SERVICIO  
DE IMPORTACIÓN Y EXPORTACIÓN DE CALZADO DEL DEPARTAMENTO LA  
LIBERTAD”**

**Elaborado por:**

Br. Benites Robles, José Dennis Ivan

Br. López Girón, Stephany Carolina

**Aprobada por:**

---

Mg. Ing. Percy Lucio Carranza Medina  
Presidente  
CIP: 149877

---

Ing. Heber Gerson Abanto Cabrera  
Secretario  
CIP: 106421

---

Ing. Freddy Henry Infantes Quiroz  
Vocal  
CIP: 139578

---

Ing. Agustín Eduardo Ullón Ramírez  
Asesor  
CIP: 137602

# PRESENTACIÓN

Señores Miembros del Jurado:

De conformidad y en cumplimiento de los requisitos estipulados en el reglamento de grados y títulos de la Universidad Privada Antenor Orrego y el Reglamento Interno de la Escuela Profesional de Ingeniería de Computación y Sistemas, ponemos a vuestra disposición el presente trabajo de Tesis: “SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA EMPRESAS DE SERVICIO DE IMPORTACIÓN Y EXPORTACIÓN DE CALZADO DEL DEPARTAMENTO LA LIBERTAD” para obtener el Título Profesional de Ingeniero de Computación y Sistemas.

El contenido de la presente tesis ha sido desarrollado tomando como marco de referencia los lineamientos establecidos por la Facultad de Ingeniería, la Escuela Profesional de Ingeniería de Computación y Sistemas y los conocimientos adquiridos durante nuestra formación profesional, además de consulta de fuentes bibliográficas e información obtenida en la empresa.

Los autores.

---

Br. José Dennis Ivan Benites Robles

---

Br. Stephany Carolina López Girón

## **DEDICATORIA**

***A mi madre Elmira Robles,***

*Gracias por su inmenso apoyo, su paciencia, sus consejos y por la confianza que me sigue dando para lograr todos mis objetivos.*

***A mis hermanos Jhony y Karen Benites Robles,***

*Que más que hermanos son verdaderos amigos, decirles que son mi motivo por el cual sigo y seguiré luchando hasta obtener lo que verdaderamente nos merecemos.*

***Y mi tío Alejandro Robles,***

*Por su gran apoyo diario y de toda la vida.*

***A toda mi familia, profesores, amigos,***

*Ya que gracias a ellos he logrado cumplir muchos objetivos y estoy muy feliz por ello.*

***José Dennis Ivan Benites Robles***

***A Dios.***

*Por haberme permitido llegar hasta este punto y haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.*

***A mi madre Fanny.***

*Por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación constante que me ha permitido ser una persona de bien, pero más que nada, por su amor.*

***A mi padre Miguel Angel.***

*Por los ejemplos de perseverancia y constancia que lo caracterizan y que me ha infundado siempre, por el valor mostrado para salir adelante y por su amor.*

***Stephany Carolina López Girón***

## **AGRADECIMIENTO**

Queremos expresar nuestro agradecimiento al personal de la empresa CARUBI SAC., y en especial al administrador y dueño, quienes nos brindaron todas las facilidades para conocer el quehacer diario y los principales procesos de la empresa.

A nuestro asesor el Ing. Agustín Ullón Ramírez, por su apoyo y asesoramiento en el desarrollo de la presente Tesis.

A todas aquellas personas que nos brindaron su apoyo y su amistad desde los inicios de la carrera.

Muchas Gracias.

**Los Autores.**

## **RESUMEN**

### **“SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA EMPRESAS DE SERVICIO DE IMPORTACIÓN Y EXPORTACIÓN DE CALZADO DEL DEPARTAMENTO LA LIBERTAD”**

Por:

Br. José Dennis Ivan Benites Robles

Br. Stephany Carolina López Girón

La empresa CARUBI SAC., es una empresa dedicada a la producción de calzado. También se dedica al comercio del calzado, dentro y fuera del país. Su objetivo principal es brindar mayor calidad y confort en sus productos, para que sus clientes tengan la plena seguridad de que han realizado la mejor elección.

Una de las principales preocupaciones de la empresa es que no cuentan con un soporte tecnológico que les apoye y les permita tomar mejores decisiones en cuánto a la producción y comercialización de su producto; las decisiones que se han venido tomando a lo largo del tiempo no fueron del todo acertadas, esto ha ocasionado pérdidas considerables en sus ganancias.

Ante este problema la solución que se propone es la Implementación de un Sistema Estratégico, obteniendo como resultado un Datamart, el cual conectaremos con la tecnología Microsoft para poder obtener reportes históricos del área de ventas e inventarios; esta información le permitirá a la gerencia tomar mejores decisiones en cuanto a su producción y comercialización.

Para cumplir con esta necesidad, se utilizó la metodología de Ralph Kimball, SQL Server 2008 R2, SQL Business Intelligent, Tecnología Microsoft para la implementación del Data Mart, y así obtener reportes de gráficos y tablas dinámicas.

En conclusión la implementación del Data Mart y sus reportes brinda un mejor análisis de los datos que maneja la empresa CARUBI SAC. Este análisis dinámico, permitirá que los datos sean accesibles de tal forma que el usuario podrá tener un mejor soporte para la toma de decisiones.

Este Data Mart permitirá a la gerencia desarrollar estrategias para mantenerse mejor posicionados en el mercado.

## **ABSTRACT**

### **BUSINESS INTELLIGENCE SOLUTIONS FOR COMPANIES IMPORT AND EXPORT SERVICE DEPARTMENT LA LIBERTAD SHOES**

**By:**

Br. José Dennis Ivan Benites Robles

Br. Stephany Carolina López Girón

The company CARUBI SAC. Is a company dedicated to the production of footwear. It is also engaged in the trade of shoes, inside and outside the country. Its main purpose is to provide higher quality and comfort in their products, so that their customers have the full assurance that they have made the best choice.

One of the main concerns of the company is that they do not have a technological support to support them and allow them to make better decisions regarding the production and marketing of their product; the decisions that have been taking over time were not entirely successful, this has caused substantial losses in earnings.

Faced with this problem, the proposed solution is the implementation of a strategic system, resulting in a Datamart, which connect with the Microsoft technology to obtain historical reports of the sales and inventories; this information will allow management to make better decisions regarding their production and marketing.

To meet this need, the methodology of Ralph Kimball, SQL Server 2008 R2, SQL Business Intelligence, Microsoft Technology to implement the Data Mart, was used to obtain reports and graphs and pivot tables.

In conclusion, the implementation of the Data Mart and their reports provide a better analysis of the data handled by the company CARUBI SAC. This dynamic analysis allow data accessible so that the user can have better support for decision-making.

Data Mart this will allow management to develop strategies to keep better positioned in the market.

In conclusion, the data mart implemented can provide a better analysis of the data handled by the company the altarpiece. This analysis is dynamic, allows data to be accessible in the way the user intends to analyze.

Data Mart this will allow management to develop strategies to keep better positioned in the market.

# ÍNDICE DE CONTENIDO

PRESENTACIÓN .....	3
DEDICATORIA.....	4
AGRADECIMIENTO.....	5
RESUMEN.....	6
ABSTRACT.....	7
ÍNDICE DE CONTENIDO.....	8
INDICE DE FIGURAS .....	10
INDICE DE TABLAS.....	12
INTRODUCCIÓN.....	13
CAPITULO I.....	42
FUNDAMENTO TEÓRICO Y METODOLOGÍA .....	42
1.1.    Fundamento Teórico .....	42
1.1.1.    Refactoring .....	42
1.1.2.    Proceso de Toma de Decisiones.....	44
1.1.3.    Tecnologías de Información .....	47
1.1.4.    Sistemas de Información .....	47
1.1.5.    Procesamiento Transaccional en Línea (OLTP).....	50
1.1.6.    Data WareHouse .....	51
1.1.7.    Business Intelligence (BI).....	62
1.2.    Metodología .....	62
1.2.1.    Metodología de Ralph Kimball .....	62
CAPITULO II.....	67
DESARROLLO DEL PROYECTO DE TESIS .....	67
2.1.    Planificación del Proyecto .....	67
2.1.1.    Selección de la Estrategia de Implementación.....	67
2.1.2.    Selección de la Metodología de Desarrollo.....	67
2.1.3.    Selección del Ámbito de Implementación.....	67
2.1.4.    Selección del enfoque arquitectónico.....	68
2.1.5.    Desarrollo de un Cronograma y del Presupuesto del Proyecto .....	69
2.1.6.    Desarrollo del escenario del uso empresarial .....	46
2.2.    Determinación de los Requerimientos.....	47



2.2.1.	Requerimientos de los Stakeholders.....	47
2.2.2.	Requerimientos Funcionales del Usuario Final .....	48
2.2.3.	Requerimientos no Funcionales .....	49
2.2.4.	Análisis de Requerimientos .....	50
2.3.	Diseño Técnico de la Arquitectura .....	53
2.3.1.	Nivel de Datos.....	53
2.3.2.	Nivel Técnico.....	54
2.4.	Implementación de la Refactorización .....	55
2.4.1.	Como hacer el Refactoring .....	55
2.4.2.	Diagrama de Base de Datos aplicando Refactoring.....	57
2.5.	Modelado Dimensional .....	58
2.5.1.	Identificación de los Componentes del Modelo.....	58
2.5.2.	Diagrama de la tabla de Hechos .....	61
2.5.3.	Esquema Estrella .....	66
2.6.	Diseño Físico .....	67
2.6.1.	Determinación de las agregaciones.....	71
2.6.2.	Construcción de las Tablas y la Base de Datos en SQL .....	71
2.7.	Proceso de Extracción, Transformación y Carga de Datos.....	74
2.7.1.	Definición de los Pasos de Transformación.....	74
2.7.2.	Definición de los Workflows .....	75
2.7.3.	Creación de los Paquetes de Servicio de Transformación de Datos .....	76
2.8.	Selección de Productos .....	93
2.8.1.	Hardware .....	93
2.8.2.	Software .....	93
2.9.	Especificación de la Aplicación del Usuario Final .....	94
2.9.1.	Estructura de Cubo .....	94
2.10.	Desarrollo de la Aplicación del Usuario Final .....	96
2.10.1.	Estructura de Cubo .....	97
CAPITULO III .....		98
DISCUSIÓN.....		98
3.1.	Formulación del Problema .....	98
3.2.	Hipótesis.....	98
3.2.1.	Manera Presencial.....	98
3.2.2.	Cálculo de los Indicadores de la Hipótesis.....	99

3.2.3. Aplicación del rango de satisfacción a los Indicadores de la Hipótesis.....	103
3.3. Manera no Presencial.....	104
CONCLUSIONES.....	107
RECOMENDACIONES .....	108
REFERENCIAS BIBLIOGRÁFICAS .....	109
ANEXOS.....	110

## INDICE DE FIGURAS

Figura 1: Proceso de toma de decisiones (PTD 2012).....	45
Figura 2: Toma de decisiones (PTD 2012) .....	45
Figura 3: Tipos de Sistemas de Información (Kendall y Kendall, 2012).....	48
Figura 4: Modelo de un ESS (Laudon y Laudon, 2013). .....	50
Figura 5: Estructura básica Datawarehouse (Kimball y Ross., 2013).....	53
Figura 6: Arquitectura básica para OLAP (Kimball y Ross., 2013).....	55
Figura 7: Un esquema de estrella. (Kimball y Ross., 2013).....	56
Figura 8: Un esquema de copo de nieve (Kimball y Ross., 2013). .....	57
Figura 9: Cubos OLAP (Root y Mason, 2012).....	59
Figura 10: Esquemas Relacionales y cubo (Root y Mason, 2012).....	60
Figura 11: Jerarquías de los cubos OLAP (Root y Mason. 2012) .....	60
Figura 12: Jerarquías de los cubos OLAP (Root y Mason. 2012).....	61
Figura 13: Business Intelligence (Kimball y Ross., 2013). .....	62
Figura 14: Metodología de implementación (Kimball, R. Y Ross, M., 2013) .....	63
Figura 15: Cronograma de Actividades .....	45
Figura 16: Requerimiento 1 - Análisis.....	50
Figura 17: Requerimiento 2 - Análisis.....	50
Figura 18 : Requerimiento 3 – Análisis.....	50
Figura 19: Requerimiento 4 - Análisis.....	51
Figura 20: Requerimiento 5 - Análisis.....	51
Figura 21: Requerimiento 6 - Análisis.....	51
Figura 22: Requerimiento 7 - Análisis.....	51
Figura 23: Requerimiento 8 - Análisis.....	52
Figura 24: Requerimiento 9 - Análisis.....	52
Figura 25: Requerimiento 10 - Análisis.....	52
Figura 26: Enfoque Arquitectónico de Base Datos Transaccional.....	53
Figura 27: Diagrama técnico - creación de Data Mart.....	54
Figura 28: Diagrama de BD-Refactoring .....	57
Figura 29: Diagrama de Tabla de Hechos Ventas .....	61
Figura 30: Diagrama de Tabla de Hechos Inventario .....	62
Figura 31: Tablas de Hechos del Data Mart.....	63
Figura 32: Jerarquía de los Atributos de la Dimensión Tiempo.....	63
Figura 33: Jerarquía de los Atributos de la Dimensión Cliente .....	64
Figura 34: Jerarquía de los Atributos de la Dimensión Producto .....	65

Figura 35: Esquema Estrella del Data Mart .....	66
Figura 36: Claves Foráneas de las Tablas de Hechos.....	70
Figura 37: Diseño Físico de la Base de Datos del Data Mart .....	70
Figura 38: Tabla de la Dimensión Tiempo Ventas .....	71
Figura 39: Tabla de la Dimensión Tiempo Compras.....	72
Figura 40: Tabla de la Dimensión Proveedor.....	72
Figura 41: Tabla de la Dimensión Cliente .....	72
Figura 42: Tabla de la Dimensión Vendedor .....	72
Figura 43: Tabla de la Dimensión Transferencia .....	73
Figura 44: Tabla de la Dimensión Producto.....	73
Figura 45: Data Mart para la Empresa Carubi SAC.....	73
Figura 46: Workflow de los pasos de Transformación .....	75
Figura 47: Diagrama Workflow con Restricciones de Precedencia .....	76
Figura 48: ETL del Data Mart .....	77
Figura 49: Ventana de Código para el Paso Limpieza Total.....	78
Figura 50: Ventana de Código para el Paso Limpieza Total.....	79
Figura 51: Conexión con Dim_Tiempo .....	79
Figura 52: Poblamiento de la Dimensión Tiempo Ventas .....	80
Figura 53: Conexión Dim_Tiempo .....	80
Figura 54: Conexión Tiempo.....	81
Figura 55: Poblamiento de la Dimensión Tiempo Compras .....	81
Figura 56: Conexión Dim_Proveedor.....	82
Figura 57: Conexión Dim_Proveedor II.....	82
Figura 58: Poblamiento de la Dimensión Proveedor.....	83
Figura 59: Conexión Dim_Vendedor .....	83
Figura 60: Conexión Dim_Vendedor II .....	83
Figura 61: Poblamiento de la Dimensión Vendedor.....	84
Figura 62: Conexión Dim_Cliente .....	84
Figura 63: Conexión Dim_Cliente II .....	84
Figura 64: Poblamiento de la Dimensión Cliente .....	85
Figura 65: Conexión Dim_Transferencia .....	85
Figura 66: Conexión Dim_Transferencia II .....	86
Figura 67: Poblamiento de la Dimensión Transferencia .....	86
Figura 68: Conexión Dim_Producto.....	87
Figura 69: Conexión Dim_Producto.....	87
Figura 70: Poblamiento de la Dimensión Producto.....	88
Figura 71: Conexión Hechos_Venta .....	89
Figura 72: Conexión Hechos_Venta II.....	90
Figura 73: Poblamiento de la Tabla de Hechos Ventas .....	90
Figura 74: Conexión Hechos_Inventario .....	91
Figura 75: Conexión Hechos_Inventario II.....	92
Figura 76: Poblamiento de la Tabla de Hechos Inventario.....	92
Figura 77: Diseño de Cubo de Ventas.....	94
Figura 78: Medidas y Dimensiones del Cubo Ventas .....	95
Figura 79: Diseño de Cubo de Inventarios.....	95

Figura 80: Medidas y Dimensiones del Cubo Inventario .....	96
Figura 81: Diseño de interfaces .....	96
Figura 82: Reporte de Requerimiento 1 .....	97
Figura 83: Reporte de Requerimiento 8 .....	97
Figura 84: Definición de Variable .....	98

## INDICE DE TABLAS

Tabla 1: Diferencia entre Sistemas Transaccionales y Datawarehouse (Kimball y Ross., 2013) .....	52
Tabla 2: Metodología de aplicación para BI .....	66
Tabla 3: Presupuesto Total .....	69
Tabla 4: Personal involucrado en el Proyecto- Descripción de usuarios del Data Mart .....	46
Tabla 5: Reportes y Unidades de Medición.....	49
Tabla 6: Matriz de Intersecciones.....	59
Tabla 7: Objetivos de las tablas de Hecho.....	60
Tabla 8: Dimensiones de la tabla de Hechos Ventas.....	60
Tabla 9: Dimensiones de la tabla de Hechos Inventario .....	60
Tabla 10: Elección de Hechos .....	61
Tabla 11: Detalle de las claves de las dimensiones .....	62
Tabla 12: Detalle de las Medidas de las Tablas de Hecho .....	63
Tabla 13: Detalle de Dimensión Tiempo Ventas/Compras .....	64
Tabla 14: Detalle de Dimensión Vendedor .....	64
Tabla 15: Detalle de Dimensión Cliente .....	64
Tabla 16: Detalle de Dimensión Producto.....	65
Tabla 17: Detalle de Dimensión Proveedor.....	65
Tabla 18: Detalle de Dimensión Transferencia .....	65
Tabla 19: Nombres estándares para las Tablas Hechos y Dimensiones .....	67
Tabla 20: Nombres estándares para los atributos de las dimensiones .....	67
Tabla 21: Nombres estándares para los atributos de las Tablas de Hechos.....	68
Tabla 22: Tipo de Dato para Dim_Tiempo .....	68
Tabla 23: Tipo de Dato para Dim_Vendedor .....	68
Tabla 24: Tipo de Dato para Dim_Cliente .....	69
Tabla 25: Tipo de Dato para Dim_Proveedor .....	69
Tabla 26: Tipo de Dato para Dim_Producto .....	69
Tabla 27: Tipo de Dato para Dim_Transferencia .....	69
Tabla 28: Determinación de las Agregaciones .....	71
Tabla 29: Software para el proceso de Diseño del Data Mart .....	93
Tabla 30: Descripción de Variable .....	99
Tabla 31: Rango de grado de satisfacción.....	99
Tabla 32: Evaluación de los indicadores de la hipótesis .....	104
Tabla 33: Cuestionario Dirigido al Jefe del Área de Ventas y al Jefe del área de Informática .....	114
Tabla 34: Cuestionario Dirigido al Jefe del Área de Ventas y el Jefe del Área de Informática.....	115

## INTRODUCCIÓN

En estos últimos años, el diseño, desarrollo e implementación de las base de datos estratégicas (Data Warehouse o Data Marts), ha adquirido una especial importancia, lo que permite a la parte gerencial de una organización tener un soporte o apoyo en la toma de decisiones, para lo cual encontramos abundante información sobre su base teórica, metodología de implementación, recomendaciones para su éxito. Todo esto reflejado mediante la emisión de información y reportes requeridos para tomar mejores decisiones.

Esto hace uso de la información de un negocio o sistema para soportar la toma de decisiones a diversos niveles organizacionales. Esta labor ha sido ampliamente usada en diversos sectores. En el sector empresarial, el diseño y desarrollo de Data Warehouse puede apoyar la toma de decisiones administrativa para optimizar los procesos y obtener los mejores resultados financieros.

La Empresa Carubi SAC. es una empresa dedicada a la producción de calzado. Dedicado netamente a las comercialización de su producto. Su objetivo principal es brindar mayor calidad y confort en sus productos, para que sus clientes tengan la plena seguridad de que han realizado la mejor elección.

En la actualidad la empresa Carubi SAC. presenta las siguientes deficiencias en la gestión que estamos analizando, que detallamos a continuación:

- No se tiene información de los gastos de la empresa al adquirir los productos para vender.
- No se tiene información en base a las ganancias o demandas de la empresa.
- No se tiene información de nuevos productos para innovar la empresa.
- No se tiene información del stock de sus productos.
- No tienen un orden de información en base a sus productos.
- Desconocen los productos que son más vendidos, implica pérdida de ventas y clientes.
- Necesitan saber en el transcurso del tiempo como van gestionando el campo de las ventas.
- El tiempo requerido para desplegar la información no es inmediato y con frecuencia no se da acceso a información requerida por ciertas áreas, siendo la recaptura de datos fuente de inconsistencias y errores.

Esta problemática se debe a que dicho sistemas transaccional con el que cuenta no fue desarrollado con el fin de brindar síntesis, análisis, consolidación, búsquedas de datos.

***Estos problemas pueden ser expresados en una sola pregunta:***

¿Con la implementación de un Sistema de Inteligencia de Negocio permitirá dar un mejor soporte a la toma de decisiones para las empresas de servicio de importación y exportación del departamento la Libertad?

***La hipótesis es:***

La implementación de una Solución de Inteligencia de Negocios que mejorará el proceso de Toma de Decisiones para la empresa importadora y exportadora de calzado.

***El objetivo general es:***

Implementar una solución de Inteligencia de Negocio para la empresa Carubi SAC dedicada al servicio de Importación y Exportación de Calzado del departamento La Libertad.

***Los objetivos específicos son los siguientes:***

- ✓ Elaborar un marco ontológico respecto al trabajo evolutivo sobre la base de datos basada en la Refactorización de Base de Datos.
- ✓ Definir los Stakeholders y obtener los requerimientos a través de entrevistas.
- ✓ Realizar el análisis de requerimientos para la elaboración de un Refactoring al Modelado de Datos que se tiene de la organización.
- ✓ Dar solución a los problemas de Integración de datos usando las herramientas ETL.
- ✓ Implementar los cubos y reportes dinámicos de acuerdo a las necesidades del cliente.

El presente proyecto está organizado en tres capítulos que facilitarán el uso y entendimiento del mismo dando a continuación una breve descripción del mismo:

**Capítulo I:** Fundamento teórico y metodología, en este capítulo damos el conocimiento formal de los temas y el modelo de referencia a utilizar para la solución del problema planteado. En esta parte damos los conocimientos básicos de que es un Refactoring, DataMart, Data Warehouse, Soporte de tomas de decisiones y Cubos OLAP.

También en este capítulo referimos la metodología que empleamos para el desarrollo del proyecto, presentando el Enfoque, Esquema y Herramientas que intervendrán en la misma.

**Capítulo II:** Desarrollo del trabajo, en éste capítulo muestra el desarrollo de los pasos enunciados en el Esquema de la Metodología. Mostramos los resultados obtenidos con relación a los objetivos planteados al inicio de este proyecto.

**Capítulo III:** Contrastación de la Hipótesis.

**Conclusiones y Recomendaciones**, en éste último capítulo se encuentra las conclusiones a las que se llegó después de haber culminado el proyecto y las recomendaciones derivadas de la experiencia de la misma.

Además en la parte final del documento se encuentran los anexos con información adicional, referidos a los capítulos enunciados.

Esperamos que este trabajo contribuya a una mejor comprensión de un Data Mart y sirva como guía de consulta para otros trabajos similares que se realicen posteriormente.

# CAPITULO I

## FUNDAMENTO TEÓRICO Y METODOLOGÍA

### 1.1. Fundamento Teórico

#### 1.1.1. Refactoring

Un punto clave en las metodologías ágiles es el refactoring. Refactoring es una forma disciplinada de hacer pequeños cambios en el código fuente para mejorar su diseño, haciendo más fácil la forma de trabajar con el mismo. Refactoring permite evolucionar el código lentamente con el tiempo, para tomar un enfoque evolutivo (iterativo e incremental) de programación.

Un aspecto crítico sobre refactoring es que mantiene el comportamiento semántico del código. No se debe agregar ni sacar nada en un refactoring, sólo se mejora la calidad.

Un ejemplo de refactoring podría ser renombrar una operación o una variable con otro nombre que indique más claramente su propósito. Para implementar un refactoring que cambia el nombre de una operación se debe cambiar el nombre de la misma y luego cambiar cada invocación a la misma a lo largo de todo el código fuente.

Un refactoring no está completo hasta que el código corra igual a como lo hacía antes de aplicar el mismo.

Claramente se necesita una forma sistemática para hacer refactoring de código, incluyendo herramientas y técnicas para hacer eso. Muchos entornos de desarrollo integrados (IDE, Integrated Development Environment) actualmente soportan refactoring de código, lo que es un buen comienzo. Sin embargo, para hacer refactoring en la práctica, se necesita desarrollar un conjunto de tests de integración que validen que el código siga funcionando como antes. No se tendrá la confianza necesaria para hacer refactoring si no se puede asegurar este hecho.

Los desarrolladores que trabajan con metodologías ágiles consideran el refactoring como una técnica primaria de programación. Es tan común hacer un refactoring como introducir una sentencia if o loop en el código. Se debe hacer refactoring en el código porque se trabaja con mayor productividad cuando el código es de mayor calidad.

Más allá de los elementos de cada metodología ágil, en general coinciden que cuando se tiene un nuevo requerimiento que agregar, la primera pregunta que hay que hacer es: “¿Es este el mejor diseño que permite agregar este requerimiento?”

Si la respuesta es sí, se agrega el requerimiento. Si la respuesta es no, primero se hace el refactoring necesario para que el código tenga el mejor diseño posible y luego se agrega el requerimiento.



En principio esto suena como una carga importante de trabajo, en la práctica, sin embargo, si se comienza con un código de alta calidad, y si se aplica refactoring para mantenerla, este enfoque agiliza el desarrollo porque siempre estamos trabajando con el mejor diseño posible.

La adopción de técnicas ágiles y refactoring trae las siguientes ventajas:

- ✓ Se minimiza el trabajo en vano. Un enfoque evolutivo permite evitar pérdidas inherentes a las técnicas en cascada cuando cambia un requerimiento. Una temprana inversión en requerimientos detallados, arquitectura y diseño de artefactos es pérdida cuando un requerimiento es tardíamente encontrado.
- ✓ Se evita rehacer el mismo trabajo. Con las técnicas evolutivas también se necesita hacer algún modelado inicial de los principales requerimientos, los que pueden llevar a rehacer significativa cantidad de trabajo en caso que fueran detectados tardíamente. Sin embargo no se investigan los detalles tempranamente.
- ✓ Siempre se tiene un sistema funcionando. Con un enfoque evolutivo, regularmente se produce software funcionando. Cuando se tiene una nueva versión funcionando del sistema cada una o dos semanas, aunque sea en un ambiente para demo, se reducen drásticamente los riesgos del proyecto.
- ✓ Siempre se sabe que se tiene el mejor diseño posible. Este es el punto clave de lo que se trata el refactoring: mejorar el diseño en pequeños pasos cada vez.
- ✓ Se trabaja en forma compatible con los desarrolladores. Los desarrolladores de aplicaciones trabajan en una forma evolutiva, y si los desarrolladores de bases de datos pretenden formar parte de un equipo de desarrollo moderno, deben trabajar de una forma evolutiva.
- ✓ Se reduce el esfuerzo global. Trabajando en una forma evolutiva, se hace sólo el trabajo que realmente se necesita hoy y nada más. (Fowler, Martín 1999).

### **Como hacer el Refactoring**

Lo ideal es hacer el refactoring sobre la BD Transaccional, nos permite tener un mayor control y organización para redactar el código y así poder corregirlo; lo mejor de poder usar refactoring es que podemos trabajar bajo pruebas unitarias ir implementando poco a poco sin falla y ejecutando lo que se desarrolla.

[La idea la explica perfectamente la metáfora de los dos sombreros. Según esta metáfora, un programador tiene a su disposición dos sombreros. Uno de ellos etiquetado "hacer código nuevo", y el otro con la etiqueta "arreglar código"]. *Metáfora de dos sombreros*

### **VENTAJAS DE HACER REFACTORING**

- Por más claro que se tenga el código original es necesario realizar un análisis unitario para poder detectar fallas futuras

- Al presentarse fallas y desorden en el código aplicar las pruebas unitarias para detectar fallas y aplicar el refactoring etiquetando el código y ordenándolo.
- No es una pérdida de tiempo arreglar el código, al final del mismo se gana dicho tiempo. las modificaciones y añadido tardan menos y se pierde mucho menos tiempo en depurar y entender el código.

## **DESVENTAJAS DE HACER REFACTORING**

- La modificación del código incrementa la complejidad de nuestro diseño, que es justo el efecto contrario del que intentábamos lograr al aplicarlo.
- Convocar y hacer mención que se está aplicando un refactoring; con el equipo de trabajo, para comentar el avance que se está realizando y evitar la sensación de que no se está avanzando. Una sensación que conduce a repercusiones anímicas negativas. A veces, el no comentar con el resto del equipo un refactoring que parece muy necesario, puede afectar a otros compañeros y de los que no éramos conscientes.

### **Refactoring de Base de Datos**

Fowler dice, un refactoring de bases de datos es un simple cambio al esquema de la base de datos que mejora su diseño mientras mantiene su semántica de comportamiento y de datos. En otras palabras, no se puede agregar funcionalidad o cambiar comportamiento existente, ni agregar nuevos datos o cambiar el significado de datos existentes.

Un refactoring de bases de datos es conceptualmente más complicado que un refactoring de código: un refactoring de código sólo necesita mantener el comportamiento semántico, mientras que un refactoring de bases de datos debe también mantener la semántica de la información. (Fowler, Martin 1999).

### **1.1.2. Proceso de Toma de Decisiones**

La Toma de Decisiones es un proceso por el cual se selecciona la mejor opción de entre muchas otras, este es un proceso que no solo se da en las empresas sino también en la vida cotidiana, como por ejemplo la selección de un proveedor, o al inicio de un negocio se toman decisiones que cambian nuestras vidas.

Por lo tanto la toma de decisiones están en todo lugar, no solo en el mundo empresarial sino también en la vida cotidiana, para iniciar, cambiar o concluir algo, siempre tomamos antes una decisión, entonces la vida si es una Toma de Decisiones, pero:

¿Sabemos escoger bien nuestras decisiones?

¿Poseemos algún patrón para tomar nuestras decisiones?

¿El azar es parte de una decisión?

¿Es normal guiarse de la intuición para tomar una decisión?

Leamos entonces la importancia y el proceso de la Toma de Decisiones que suele un tema primordial en la administración de negocios. (Robbins & Coulter 2005)



Figura 1: Proceso de toma de decisiones (PTD 2012)

### ¿Qué es Toma de Decisiones?

La toma de decisiones es un proceso sistemático y racional a través del cual se selecciona una alternativa de entre varias, siendo la seleccionada la optimizadora (la mejor para nuestro propósito).

Tomar una decisión es resolver diferentes situaciones de la vida en diferentes contextos: a nivel laboral, familiar, sentimental, empresarial, etc.

Tomar la correcta decisión en un negocio o empresa es parte fundamental del administrador ya que sus decisiones influirán en el funcionamiento de la organización, generando repercusiones positivas o negativas según su elección. (Hernández Celis, Domingo 2009).



Figura 2: Toma de decisiones (PTD 2012)

### ¿Qué debo saber antes, para tomar una buena decisión?

- ✓ Definir las restricciones y limitaciones.
- ✓ Saber la relación costo beneficio, rendimientos esperados u otros.
- ✓ Saber cuándo se utilizan métodos cuantitativos y cuando los cualitativos.

- ✓ Conocer los factores internos formales (cultura organizacional, políticas internas, estructura, etc.) y los factores internos informales (políticas implícitas, hábitos, experiencia, etc.)
- ✓ Conocer los factores externos (políticos, económicos, sociales, internacionales, culturales)
- ✓ Comprender los cinco puntos anteriores nos ayudara mucho al momento de tomar una buena decisión. (Hernández Celis, Domingo 2009).

### **Etapas del Proceso de Toma de decisiones:**

- Identificar y analizar el problema: Un problema es la diferencia entre los resultados reales y los planeados, lo cual origina una disminución de rendimientos y productividad, impidiendo que se logren los objetivos.
- Investigación u obtención de información: Es la recopilación de toda la información necesaria para la adecuada toma de decisión; sin dicha información, el área de riesgo aumenta, porque la probabilidad de equivocarnos es mucho mayor.
- Determinación de parámetros: Se establecen suposiciones relativas al futuro y presente tales como: restricciones, efectos posibles, costos , variables, objetos por lograr, con el fin de definir las bases cualitativas y cuantitativas en relación con las cuales es posible aplicar un método y determinar diversas alternativas.
- Construcción de una alternativa: La solución de problemas puede lograrse mediante varias alternativas de solución; algunos autores consideran que este paso del proceso es la etapa de formulación de hipótesis; porque una alternativa de solución no es científica si se basa en la incertidumbre.
- Aplicación de la alternativa: De acuerdo con la importancia y el tipo de la decisión, la información y los recursos disponibles se eligen y aplican las técnicas, las herramientas o los métodos, ya sea cualitativo o cuantitativo, más adecuados para plantear alternativas de decisión.
- Especificación y evaluación de las alternativas: Se desarrolla varias opciones o alternativas para resolver el problema, aplicando métodos ya sea cualitativos o cuantitativos. Una vez que se han identificado varias alternativas, se elige la óptima con base en criterios de elección de acuerdo con el costo beneficio que resulte de cada opción. Los resultados de cada alternativa deben ser evaluados en relación con los resultados esperados y los efectos.
- Implantación: Una vez que se ha elegido la alternativa optima, se deberán planificarse todas las actividades para implantarla y efectuar un seguimiento de los resultados, lo cual requiere elaborar un plan con todos los elementos estudiados. (Robbins & Coulter 2005).

Lo presentado anteriormente fueron tan solo las etapas del proceso de toma de decisiones, desarrollarlas dependerá del tipo de problema que se quiera solucionar y del tipo de técnica que deba aplicar para solucionarlo.

En la toma de Decisiones existen también Técnicas Cuantitativas y Cualitativas para la selección de la mejor decisión. (Robbins & Coulter 2005).

- Técnicas Cualitativas: Cuando se basan en criterio de la experiencia, y habilidades
- Técnicas Cuantitativas: Cuando se utilizan métodos matemáticos, estadísticos, etc.

### **1.1.3. Tecnologías de Información**

Las Tecnologías de Información o TICs son aquellas herramientas computacionales e informáticas que procesan, almacenan, sintetizan, recuperan y presentan información representada de la más variada forma. Es un conjunto de herramientas, soportes y canales para el tratamiento y acceso a la información. Constituyen nuevos soportes y canales para dar forma, registrar, almacenar y difundir contenidos informacionales. (Villalobos y Constenia, 2010).

### **1.1.4. Sistemas de Información**

Es un conjunto formal de procesos que, operando sobre una colección de datos estructurada según las necesidades de la empresa, recopilan elaboran y distribuyen la información (o parte de ella) necesaria para las operaciones de dicha empresa y para las actividades de dirección y control correspondientes (decisiones) desempeñar su actividad de acuerdo a su estrategia de negocio.

Otra definición de sistemas de información es "Son procesar entradas, mantener archivos de datos relacionados con la organización y producir información, reportes y otras salidas". Por lo tanto podemos definir un sistema de información como un conjunto de subsistemas que incluyen hardware, software, medios de almacenamiento de datos ya sea primarios, secundarios y bases de datos relacionadas entre sí con el fin de procesar entradas para realizar transformaciones a esas entradas y convertirlas en salidas de información importantes en la toma de decisiones.

El objetivo de un sistema de información es ayudar al desempeño de las actividades que desarrolla la empresa, suministrando la información adecuada, con la calidad requerida, a la persona o departamento que lo solicita, en el momento y lugar especificados con el formato más útil para el receptor. (Kendall y Kendall, 2012).

Tipos de Sistemas de Información:

En la Figura N° 3 muestra los tipos específicos de los sistemas de información que corresponden a cada uno de los niveles de la organización.

- Sistemas de procesamiento de transacciones (TPS)

- Sistemas de información administrativa (MIS)
- Sistemas de apoyo a las decisiones (DSS)
- Sistemas de apoyo a ejecutivos (ESS)

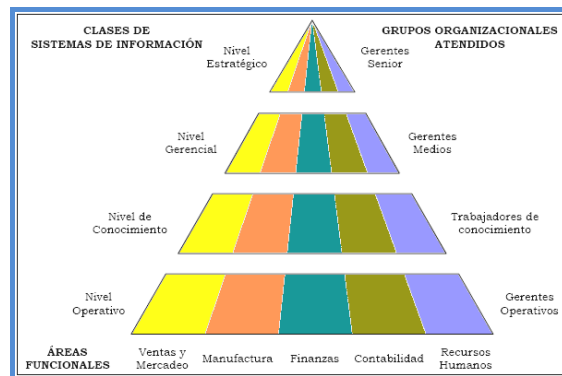


Figura 3: Tipos de Sistemas de Información (Kendall y Kendall, 2012)

### a) Sistemas de Procesamiento de Transacciones:

Los sistemas de procesamiento de transacciones son sistemas de información encargados de procesar gran cantidad de transacciones rutinarias, es decir son todas aquellas que se realizan rutinariamente en la empresa entre estas tenemos el pago de nómina, facturación, entrega de mercancía y depósito de cheques. Estas transacciones varían de acuerdo al tipo de empresa.

Los sistemas de procesamiento de transacción o TPS (Transaccion Procesation System) por sus siglas en inglés, eliminan el trabajo tedioso de las transacciones operacionales y como resultado reducen el tiempo que se empleaba en ejecutarlas actualmente, aunque los usuarios todavía deben alimentar de datos a los TPS.

"Los sistemas de procesamiento de transacciones son sistemas que traspasan sistemas y que permiten que la organización interactúe con ambientes externos. Debido a que los administradores consultan los datos generados por el TPS para información al minuto acerca de lo que está pasando en sus compañías, es esencial para las operaciones diarias que estos sistemas funcionen lentamente y sin interrupción". (Laudon y Laudon, 2013).

### b) Sistemas de información de Gerencial:

Los sistemas de información gerencial (MIS por sus siglas en inglés) no reemplazan a los sistemas de procesamiento de transacciones ni tampoco son los mismos, sino que estos sistemas incluyen procesamiento de transacciones. Los sistemas de información gerencial son sistemas de información computarizada que trabajan con la interacción entre usuarios y computadoras. Requieren que los usuarios, el software (programas de computadora) y el hardware (computadoras, impresoras, etc.) trabajen a un mismo ritmo.

Los sistemas de información gerencial dan soporte a un aspecto más amplio de tareas organizacionales, a comparación de los sistemas de procesamiento de transacciones, los sistemas de información gerencial incluyen el análisis de decisiones y la toma de decisiones.

"Para poder ligar la información, los usuarios de un sistema de información gerencial comparten una base de datos común. La base de datos guarda modelos que ayudan a los usuarios a interpretar y aplicar esos mismos datos.

Los sistemas de información gerencial producen información que es usada en la toma de decisiones. Un sistema de información gerencial también puede llegar a unificar algunas de las funciones de información computarizada, aunque no exista como una estructura singular en ningún lugar del negocio". (Laudon y Laudon, 2013).

### **c) Sistemas de Apoyo a Decisiones:**

Los sistemas de apoyo a decisiones (DSS) ayudan a los gestores a tomar decisiones que son únicas, que cambia rápidamente, y no es fácil definirse de antemano.

Ellos tratan de problemas que el procedimiento para llegar a una solución no puede ser plenamente predefinidos de antemano. Aunque el Departamento de Servicios Sociales uso de información interna de TPS y MIS, que a menudo traen consigo la información de fuentes externas, tales como los precios de las acciones o los precios de los productos de los competidores.

Es evidente que, de diseño, DSS tienen más poder analítico que otros sistemas. Utilizan una gran variedad de modelos para analizar los datos, o se condensan grandes cantidades de datos en un formulario en el que puedan ser analizados por los encargados de adoptar decisiones. DSS están diseñados para que los usuarios puedan trabajar con ellos directamente, estos sistemas incluyen explícitamente el software de uso fácil. DSS son interactivos, el usuario puede cambiar las hipótesis, pedir nuevas preguntas, e incluir nuevos datos. (Laudon y Laudon, 2013)

### **d) Sistemas de Apoyo a Ejecutivos:**

Los Sistemas de Apoyo a Ejecutivos (ESS) son utilizados en el nivel estratégico de la organización. Los ESS no solo están diseñados para incorporar información sobre eventos externos, como las nuevas leyes fiscales o de los competidores, sino que también sacar un resumen de la información interna de los sistemas MIS y DSS.

Estos sistemas pueden filtrar, comprimir, y realizar un seguimiento de datos críticos, mostrando los datos de la mayor importancia para los altos ejecutivos.

Por ejemplo, el CEO de Leiner Health Products, el mayor fabricante de etiqueta privada vitaminas y suplementos en los Estados Unidos, tiene un ESS que ofrece en su escritorio, minuto a minuto la vista de la empresa: estados financieros, medido por capital de trabajo, cuentas por cobrar, cuentas por pagar, flujo de caja, e inventario.

Los Sistemas de Apoyo a Ejecutivos, emplean los más avanzados software de gráficos y puede presentar gráficos y datos de muchas fuentes.

A menudo la información se entrega a los altos ejecutivos a través de un portal, que utiliza una interfaz Web integrada para presentar contenido personalizado de negocio de una variedad de fuentes.

A diferencia de los otros tipos de sistemas de información, los ESS no están diseñados principalmente para resolver problemas específicos. Aunque muchos DSS están diseñados para ser muy analítico, los ESS tienden a hacer menos uso de modelos analíticos.

Los Sistemas de Apoyo a Ejecutivos ayudan a responder a las siguientes preguntas: ¿En qué negocios deberíamos estar? ¿Cuáles son nuestros competidores? ¿Qué nuevas adquisiciones deberíamos realizar? ¿Qué unidades debemos vender más para recaudar más utilidades?

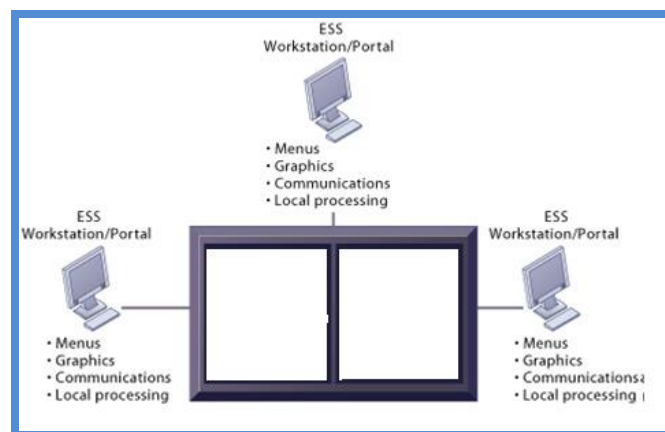


Figura 4: Modelo de un ESS (Laudon y Laudon, 2013).

La ilustración es un modelo de un ESS. Se compone de los puestos de trabajo con menús, gráficos interactivos, y con la capacidad de comunicación que pueden utilizarse para el acceso histórico de los datos corporativos internos y los sistemas de bases de datos externas.

Porque ESS están diseñados para ser utilizados por altos directivos que a menudo tienen poca información, en su caso, el contacto directo con su computadora y la experiencia, basado en los sistemas de información, que sea de fácil uso, con interfaces gráficas. (Laudon y Laudon, 2013)

### 1.1.5. Procesamiento Transaccional en Línea (OLTP)

Procesamiento Transaccional en Línea (Online Transactional Processing, en inglés), tiene como objetivo mantener la integridad de la información (relaciones entre los datos) necesaria para operar un negocio de la manera más eficiente. Sin embargo, este modelo no corresponde a la forma como el usuario percibe la operación de un negocio. El OLTP se basa en un modelo relacional. (Kimball y Ross., 2013).



### 1.1.6. Data Warehouse

El Data Warehouse (DW) convierte entonces los datos operacionales de una organización en una herramienta competitiva, por hacerlos disponibles a los empleados que lo necesiten para el análisis y toma de decisiones.

El objetivo del Datawarehouse es el de satisfacer los requerimientos de información interna de la empresa para una mejor gestión. El contenido de los datos, la organización y estructura son dirigidos a satisfacer las necesidades de información de los analistas y usuarios tomadores de decisiones. El DW es el lugar donde la gente puede acceder a sus datos.

El DW puede verse como una bodega donde están almacenados todos los datos necesarios para realizar las funciones de gestión de la empresa, de manera que puedan utilizarse fácilmente según se necesiten.

Los Datawarehouse (almacenes de datos) generan bases de datos tangibles con una perspectiva histórica, utilizando datos de múltiples fuentes que se fusionan en forma congruente. Estos datos se mantienen actualizados, pero no cambian al ritmo de los sistemas transaccionales.

Muchos datawarehouses se diseñan para contener un nivel de detalle hasta el nivel de transacción, con la intención de hacer disponible todo tipo de datos y características, para reportar y analizar.

Así un datawarehouse resulta ser un recipiente de datos transaccionales para proporcionar consultas operativas, y la información para poder llevar a cabo análisis multidimensional.

De esta forma, dentro de un datawarehouse existen dos tecnologías que se pueden ver como complementarias, una relacional para consultas y una multidimensional para análisis (Kimball y Ross., 2013).

DW está basado en un procesamiento distinto al utilizado por los sistemas operacionales, es decir, este se basa en OLAP -Procesos de Análisis en Línea- (OnLine Analysis Process, en inglés), usado en el análisis de negocios y otras aplicaciones que requieren una visión flexible del negocio.

Para ampliar los conceptos anteriores, en la tabla N°1 se exponen las principales diferencias entre los sistemas Transaccionales (OLTP) y los basados en Datawarehouses.

<b>Transaccionales</b>	<b>Basados en Datawarehouse</b>
Admiten el acceso simultáneo de muchos usuarios -miles- que agregan y modifican datos.	Admiten el acceso simultáneo de muchos usuarios -cientos- que consultan y no modifican datos
Representan el estado, en cambio constante, de una organización, pero no guardan su historial.	Guardan el historial de una organización
Contienen grandes cantidades de datos, incluidos los datos extensivos utilizados para comprobar transacciones.	Contienen grandes cantidades de datos, sumariados, consolidados y transformados. También de detalle pero solo los necesarios para el análisis.
Tienen estructuras de base de datos complejas.	Tienen estructuras de Base de datos simples.
Se ajustan para dar respuesta a la actividad transaccional.	Se ajustan para dar respuesta a la actividad de consultas.
Proporcionan la infraestructura tecnológica necesaria para admitir las operaciones diarias de la empresa.	Proporcionan la infraestructura tecnológica necesaria para admitir análisis de los datos de la empresa.
Los analistas carecen de la experiencia técnica necesaria para crear consultas "ad hoc" contra la compleja estructura de datos.	Pueden combinar datos de orígenes heterogéneos en una única estructura homogénea y simple, facilitando la creación de informes y consultas.
Las consultas analíticas que resumen grandes volúmenes de datos afectan negativamente a la capacidad del sistema para responder a las transacciones en línea.	Organizan los datos en estructuras simplificadas buscando la eficiencia de las consultas analíticas más que del proceso de transacciones.
Los datos que se modifican con frecuencia interfieren en la coherencia de la información analítica.	Proporcionan datos estables que representan el historial de la empresa. Se actualizan periódicamente con datos adicionales, no con transacciones frecuentes.

Tabla 1: Diferencia entre Sistemas Transaccionales y Datawarehouse (Kimball y Ross., 2013)

### 1) Introducción a Datamarts:

El acceso a los datos de toda la empresa a veces no es conveniente (o necesario) para determinados usuarios que solo necesitan un subconjunto de estos datos, en estos casos se utilizan los Datamarts. El concepto Datamart es una especialización del datawarehouse, y está enfocado a un departamento o área específica, como por ejemplo los departamentos de Finanzas o Marketing. Permitiendo así un mejor control de la información que se está abarcando (Kimball y Ross., 2013)

Los principales beneficios de utilizar Datamarts son:

- ✓ Acelerar las consultas reduciendo el volumen de datos a recorrer
- ✓ Estructurar los datos para su adecuado acceso por una herramienta
- ✓ Dividir los datos para imponer estrategias de control de acceso
- ✓ Segmentar los datos en diferentes plataformas hardware
- ✓ Permite el acceso a los datos por medio de un gran número de herramientas del mercado, logrando independencia de estas.

## 2) Arquitectura del Datawarehouse:

La estructura básica de la arquitectura Datawarehouse incluye:

- ✓ Datos operacionales: un origen o fuente de datos para poblar el componente de almacenamiento físico DW. El origen de los datos son los sistemas transaccionales internos de la organización como también datos externos a ésta.
- ✓ Extracción de Datos: selección sistemática de datos operacionales usados para poblar el componente de almacenamiento físico DW.
- ✓ Transformación de datos: procesos para sumarizar y realizar otros cambios en los datos operacionales para reunir los objetivos de orientación a temas e integración principalmente.
- ✓ Carga de Datos: inserción sistemática de datos en el componente de almacenamiento físico DW.
- ✓ Datawarehouse: almacenamiento físico de datos de la arquitectura DW.
- ✓ Herramientas de Acceso al componente de almacenamiento físico DW: herramientas que proveen acceso a los datos. Estas herramientas pueden ser herramientas específicas de mercado para visualización de bases multidimensionales almacenadas en datawarehouses como también aplicaciones desarrolladas dentro de la organización del tipo EIS/DSS (Kimball y Ross., 2013).

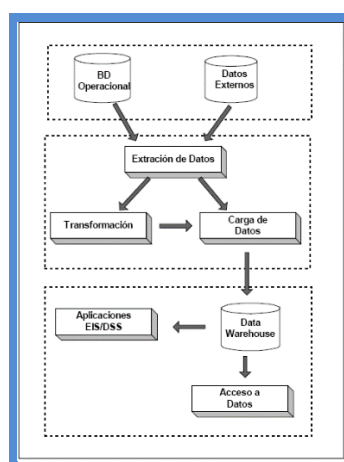


Figura 5: Estructura básica Datawarehouse (Kimball y Ross., 2013).

## a. Introducción al Procesamiento Analítico en Línea

La tecnología de Procesamiento Analítico en Línea –OLAP- (Online Analytical Processing) permite un uso más eficaz de los datawarehouses para el análisis de datos en línea, lo que proporciona respuestas rápidas a consultas analíticas complejas e iterativas utilizada generalmente para sistemas de ayuda para la toma de decisiones. Primero y más importante, el OLAP presenta los datos a los usuarios a través de un modelo de datos intuitivo y natural. Con este estilo de navegación, los usuarios finales pueden ver y entender más efectivamente la información de sus bases de datos, permitiendo así a las organizaciones reconocer mejor el valor de sus datos. (Kimball y Ross., 2013).

En segundo lugar, el OLAP acelera la entrega de información a los usuarios finales que ven estas estructuras de datos como cubos denominadas multidimensionales debido a que la información es vista en varias dimensiones.

Esta entrega es optimizada ya que se prepararan algunos valores calculados en los datos por adelantado, en vez de realizar el cálculo al momento de la solicitud. La combinación de navegación fácil y rápida le permite a los usuarios ver y analizar información más rápida y eficientemente que lo que es posible con tecnologías de bases de datos relacionales solamente. El resultado final: se pasa más tiempo analizando los datos y menos tiempo analizando las bases de datos.

Las aplicaciones OLAP deberían proporcionar análisis rápidos de información multidimensional compartida. Las características principales del OLAP son:

- ✓ Rápido: proporciona la información al usuario a una velocidad constante. La mayoría de las peticiones se deben de responder al usuario en cinco segundos o menos.
- ✓ Análisis: realiza análisis estadísticos y numéricos básicos de los datos, predefinidos por el desarrollador de la aplicación o definido “ad hoc” por el usuario.
- ✓ Compartida: implementa los requerimientos de seguridad necesarios para compartir datos potencialmente confidenciales a través de una gran población de usuarios.
- ✓ Multidimensional: llena la característica esencial del OLAP, que es ver la información en determinadas vistas o dimensiones.
- ✓ Información: acceden a todos los datos y a la información necesaria y relevante para la aplicación, donde sea que ésta resida y no esté limitada por el volumen.
- ✓ El OLAP es un componente clave en el proceso de almacenamiento de datos (data warehousing) y los servicios OLAP proporcionan la funcionalidad esencial para una gran variedad de aplicaciones que van desde reportes corporativos hasta soporte avanzado de decisiones.

Dentro de cada dimensión de un modelo de datos OLAP, los datos se pueden organizar en una jerarquía que represente niveles de detalle de los datos. Por ejemplo, dentro de

la dimensión de tiempo, se puede tener estos niveles: años, meses y días; de manera similar, dentro de la dimensión geografía,

Se puede tener estos niveles: país, región, estado/provincia y ciudad. Una instancia particular del modelo de datos OLAP tendrá valores para cada nivel en la jerarquía. Un usuario que vea datos OLAP se moverá entre estos niveles para ver información con mayor o menor detalle. (Kimball y Ross., 2013).

### ➤ **Arquitectura OLAP**

La tecnología OLAP permite un uso más eficaz de los almacenes de datos para el análisis en línea, lo que proporciona respuestas rápidas a consultas analíticas complejas e iterativas. Los modelos de datos multidimensionales de OLAP y las técnicas de agregados de datos organizan y resumen grandes cantidades de datos para que puedan ser evaluados con rapidez mediante el análisis en línea y las herramientas gráficas. La respuesta a una consulta realizada sobre datos históricos a menudo suele conducir a consultas posteriores en las que el analista busca respuestas más concretas o explora posibilidades. Los sistemas OLAP proporcionan la velocidad y la flexibilidad necesarias para dar apoyo al analista en tiempo real.

A continuación se muestra un ejemplo, la integración del datawarehouse y los procesos OLAP, que generalmente se implementan por medio de una aplicación servidora que accede al datawarehouse y realiza los procesos de análisis. A través de este servicio OLAP, los usuarios acceden a la información residente en las bases de datos (Kimball y Ross., 2013).

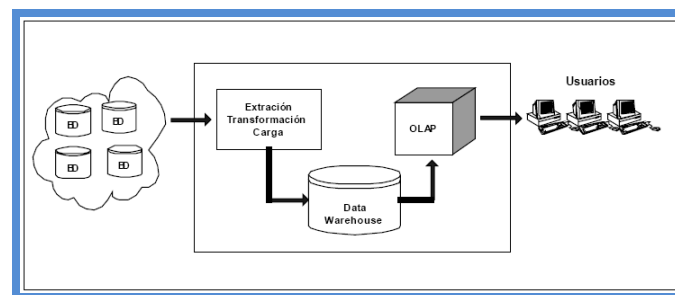


Figura 6: Arquitectura básica para OLAP (Kimball y Ross., 2013).

### ➤ **El modelo de datos OLAP**

En la mayoría de las implementaciones de OLAP, se asume que los datos han sido preparados para el análisis a través del almacenamiento de datos (data warehouse) y que la información se ha extraído de sistemas operacionales, limpiado, validado y resumido antes de incorporarse en una aplicación OLAP.

Este es un paso vital en el proceso, que asegura que los datos que son vistos por el usuario OLAP son correctos, consistentes y que llenan las definiciones organizacionales para los datos.

Cada vez más, la información en un datawarehouse se organiza en esquemas de estrella o de copo de nieve. (Kimball y Ross., 2013).

### a. Esquema Estrella:

El esquema estrella se basa en una tabla de hechos central (las medidas) que se enlaza a las tablas de dimensiones relacionadas (las categorías descriptivas de las medidas), mientras que el esquema copo de nieve, una tabla de hechos central se enlaza a las tablas de dimensiones relacionadas, pero estas a su vez se enlaza a otras tablas dimensionales.

Con este tipo de esquemas simplifica el entendimiento de los datos por parte del usuario, maximiza el desempeño de las peticiones de la base de datos para aplicaciones de soporte de decisiones y requiere menor espacio de almacenamiento para bases de datos grandes.

A continuación se muestra un ejemplo de esquema de estrella. En este tipo de base de datos, una tabla de hechos central se enlaza a las tablas de dimensiones relacionadas.

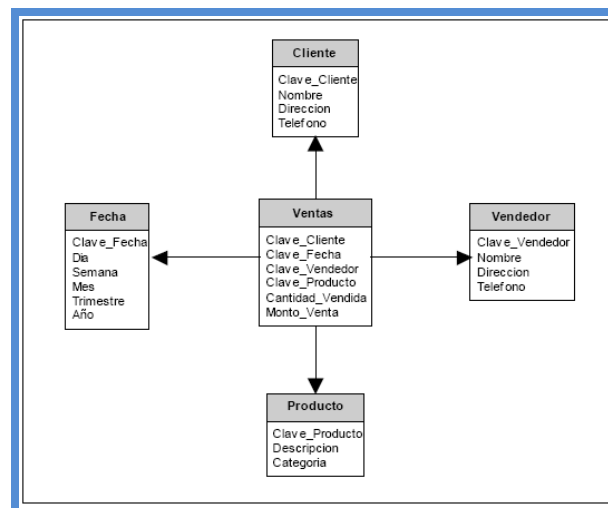


Figura 7: Un esquema de estrella. (Kimball y Ross., 2013).

A continuación se enumeran algunas de las principales ventajas del esquema estrella.

- Crea una base de datos con tiempos de respuesta rápido.
- Diseño fácil de modificar.
- Simula como ven los datos los usuarios finales.
- Simplifica la navegación.
- Facilita la interacción con herramientas.

### b. Esquema Copo de Nieve:

El esquema copo de nieve es una extensión del esquema estrella en donde cada uno de los puntos de la estrella se divide en más puntos. En esta forma de esquema, las tablas de dimensión del esquema estrella contienen más normas. Las ventajas que proporciona el esquema de copo de nieve son mejorar el desempeño de consultas debido a un mínimo almacenamiento en disco para los datos y mejorar el desempeño mediante la unión de tablas más pequeñas con normas.

Así mismo el esquema copo de nieve incrementa la flexibilidad de las aplicaciones debido a la aplicación de normas y por lo tanto disminuye la granularidad de las dimensiones. (Kimball y Ross., 2013)

A continuación se muestra un ejemplo de esquema copo de nieve. Este tipo de esquema se caracteriza por tener tablas dimensionales relacionadas con otras tablas dimensionales además de vincularse a la tabla de hechos.

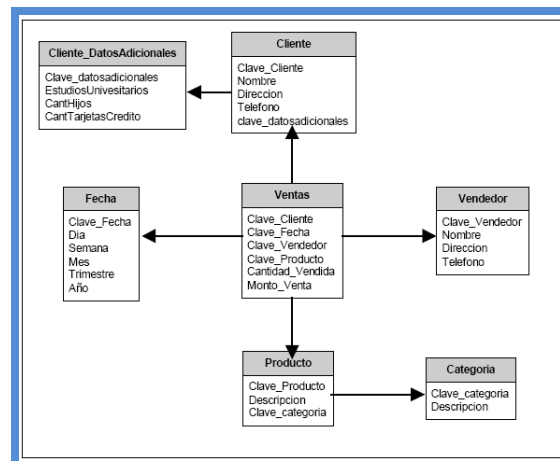


Figura 8: Un esquema de copo de nieve (Kimball y Ross., 2013).

Los esquemas de estrella y copo de nieve son aproximaciones relacionales del modelo de datos OLAP y son un punto de partida excelente para construir definiciones de cubo OLAP. Pocos productos OLAP han tomado ventaja de este hecho.

Generalmente no han provisto herramientas sencillas para mapear un esquema de estrella a un modelo OLAP y como resultado mantienen el costo de construir el modelo OLAP extremadamente alto y el tiempo de desarrollo innecesariamente largo. (Kimball y Ross., 2013).

### ➤ Implementación del OLAP

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

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

En los primeros días de la tecnología OLAP, la mayoría de las compañías asumía que la única solución para una aplicación OLAP era un modelo de almacenamiento no relacional. Después, otras compañías descubrieron que a través del uso de estructuras de base de datos (esquemas de estrella y de copo de nieve), índices y el almacenamiento de agregados, se

podrían utilizar sistemas de administración de bases de datos relacionales (RDBMS) para el OLAP.

Estos vendedores llamaron a esta tecnología OLAP relacional (ROLAP). Las primeras compañías adoptaron entonces el término OLAP multidimensional (MOLAP), estos conceptos, MOLAP y ROLAP,

Un desarrollo reciente ha sido la solución OLAP híbrida (HOLAP), la cual combina las arquitecturas ROLAP y MOLAP para brindar una solución con las mejores características de ambas: desempeño superior y gran escalabilidad. Un tipo de HOLAP mantiene los registros de detalle (los volúmenes más grandes) en la base de datos relacional, mientras que mantiene las agregaciones en un almacén MOLAP separado. (Kimball y Ross., 2013).

#### **a. Sistemas MOLAP:**

MOLAP (OLAP Multidimensional) una copia de los datos de origen del cubo, junto con sus agregaciones, es almacenada en una estructura multidimensional.

Debemos tener en cuenta que mientras los datos de origen cambian directamente con las operaciones, los objetos con almacenamiento MOLAP deben ser procesados para incorporar estos cambios.

El tiempo comprendido entre un procesamiento y el siguiente, crea un periodo de latencia durante el que puede que la información OLAP no coincida con los datos de origen actuales.

Como característica del almacenamiento MOLAP podemos desatacar:

- ✓ Provee excelente rendimiento y compresión de datos.
- ✓ Tiene mejor tiempo de respuesta, dependiendo solo del porcentaje de las agregaciones del cubo.
- ✓ La estructura está muy optimizada para maximizar el rendimiento de las consultas.
- ✓ En general este método, es muy apropiado para cubos con uso frecuente por su rápida respuesta.

#### **b. Sistemas ROLAP:**

En un modelo ROLAP (OLAP Relacional) toda la información del cubo, sus datos, su agregación, sumas, etc., son almacenados en una base de datos relacional.

A diferencia del modo de almacenamiento MOLAP, ROLAP no almacena copia de la base de datos, accede a las tablas originales cuando necesita responder a las consultas, generalmente es mucho más lenta que las otras estrategias de almacenamiento (MOLAP o HOLAP).

ROLAP se utiliza para ahorrar espacio de almacenamiento cuando se trabaja con grandes conjuntos de datos que se consultan con poca frecuencia; por ejemplo, datos exclusivamente históricos. (Kimball y Ross., 2013).



### c. **Sistemas HOLAP:**

HOLAP (OLAP híbrido) combina atributos de MOLAP y ROLAP.

Al igual que MOLAP, HOLAP hace que las agregaciones se almacenen en una estructura multidimensional, y los datos a nivel de detalle, en una base de datos relacional como lo hace el almacenamiento ROLAP.

Para procedimientos de búsqueda que accedan datos sumariados, HOLAP es equivalente a MOLAP. Por el contrario, si los procesos de consultas accedieran a los máximos niveles de detalle, deberían recuperar los datos de la base de datos relacional y esto no sería tan rápido comparado con una estructura MOLAP.

Los cubos almacenados como HOLAP, son más pequeños que los MOLAP y responden más rápidos que los ROLAP.

#### Usos comunes de HOLAP:

- Cubos que requieren rápida respuesta
- Cuando existen sumariaciones basadas en una gran cantidad de datos de origen.
- Solución de compromiso para bajar el espacio ocupado sin perjudicar totalmente el rendimiento de las consultas. (Kimball y Ross., 2013).

### ➤ **Cubos OLAP**

Un cubo es una multidimensional estructura que almacena los datos de tu sistema OLAP. Multidimensional significa que los cubos te permiten mirar tu data en varios caminos o de varias maneras.

El cubo es un elemento clave en el proceso analítico en línea OLAP y consta de Medidas (o datos cuantitativos como ventas o costos) y dimensiones.

La modelación multidimensional de datos es una forma de facilitar el análisis empresarial en línea y de mejorar el rendimiento de las consultas.

El Administrador de OLAP le permitirá convertir los datos almacenados en bases de datos relacionales en información empresarial significativa y fácil de explorar con sólo crear un cubo de datos. Los conceptos y terminología asociados con los cubos se describen en las siguientes imágenes. (Root y Mason, 2012).



Figura 9: Cubos OLAP (Root y Mason, 2012).

## ESQUEMAS RELACIONALES Y CUBOS

La manera más común de administrar datos relacionales para su empleo multidimensional es un esquema de estrella. Un esquema de estrella consiste en una única tabla de hechos que se combina con varias tablas de dimensiones.

La tabla de hechos contiene los datos numéricos que se corresponden con las medidas de un cubo. Las columnas de la tabla de dimensiones, tal como implica su nombre, asignan los niveles jerárquicos de una dimensión. (Root y Mason, 2012).

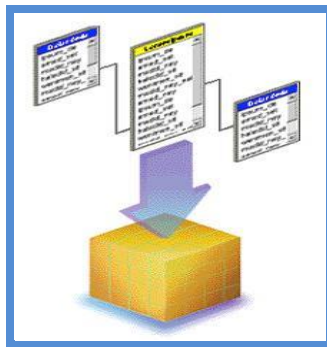


Figura 10: Esquemas Relacionales y cubo (Root y Mason, 2012)

## DIMENSIONES DE UN CUBO

Las dimensiones de un cubo representan distintas categorías para analizar datos empresariales. Categorías tales como fecha, geografía o línea de productos son dimensiones típicas de cubos.

Nota: los cubos no están limitados a tres dimensiones. Pueden contener hasta 64 dimensiones. (Root y Mason, 2012).

## DIMENSIONES Y JERARQUÍAS

Las dimensiones se suelen organizar en jerarquías de información que se asignan a columnas en una base de datos relacional.

Las jerarquías de dimensiones están agrupadas en niveles que constan de los miembros de una dimensión. Podrá unir los niveles de una dimensión para formar los valores de los que constará el siguiente nivel superior.

Por ejemplo, en una dimensión temporal, los días se unen en meses y los meses forman trimestres. (Root y Mason, 2012).

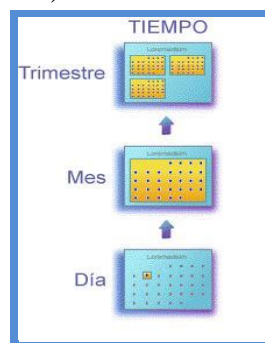


Figura 11: Jerarquías de los cubos OLAP (Root y Mason, 2012)

## MEDIDAS DE UN CUBO

Las medidas son los valores cuantitativos contenidos en la base de datos que desea analizar. Las medidas típicas son ventas, costo y datos presupuestarios.

Las medidas se analizan contra las distintas categorías de dimensiones de un cubo.

Por ejemplo, tal vez desee analizar datos de ventas y de presupuestos (sus medidas) para un determinado producto (una dimensión) correspondientes a varios países (niveles específicos de una dimensión geográfica) durante dos años concretos (niveles de una dimensión temporal). (Root y Mason, 2012).



Figura 12: Jerarquías de los cubos OLAP  
(Root y Mason, 2012)

### ➤ **Agregados**

Los agregados son resúmenes de datos pre calculados que mejoran el tiempo de respuesta a las consultas por el simple hecho de tener preparadas las respuestas antes de que se planteen las preguntas. Por ejemplo, la respuesta a una consulta que solicita el total de ventas semanales de una determinada línea de productos y que se realiza en una tabla de hechos de un almacén de datos que contiene cientos de miles de filas de información, puede llevar mucho tiempo si hay que explorar la tabla de hechos para calcular la respuesta. Por el contrario, la respuesta podría ser casi inmediata si los datos de resumen para la respuesta a esta consulta se han calculado previamente. El cálculo previo de los datos de resumen es la clave para obtener respuestas rápidas en la tecnología OLAP.

Los cubos son la forma en que la tecnología OLAP organiza los datos de resumen en estructuras multidimensionales. Las dimensiones y sus niveles jerárquicos reflejan las consultas que se pueden hacer al cubo. Los agregados se almacenan en la estructura multidimensional en celdas que tienen las coordenadas especificadas por las dimensiones. (Root y Mason, 2012)

### 1.1.7. Business Intelligence (BI)

Inteligencia de Negocios (Business Intelligence – BI) es una disciplina que, junto con sus correspondientes herramientas, hacen centro en el análisis de la información para la correcta toma de decisiones que le permita a la organización cumplir con los objetivos de negocio.

BI es un término "agrupador". El que sea considerado como un conjunto de conceptos le da un poder enorme, pues pueden integrarse funciones que tradicionalmente estaban separadas, tales como el acceso de datos, reportes, explotación, pronóstico y análisis.

De ese modo, al menos en la actualidad en empresas grandes, BI se ha convertido en un apoyo indispensable para la Toma de Decisiones, en cualquier nivel de la organización y mucha gente está explotando el potencial estratégico de los datos operativos.

Bien utilizada, BI puede ser un arma estratégica de la gente de negocios, sustentada en tecnología de sistemas (Kimball y Ross., 2013).

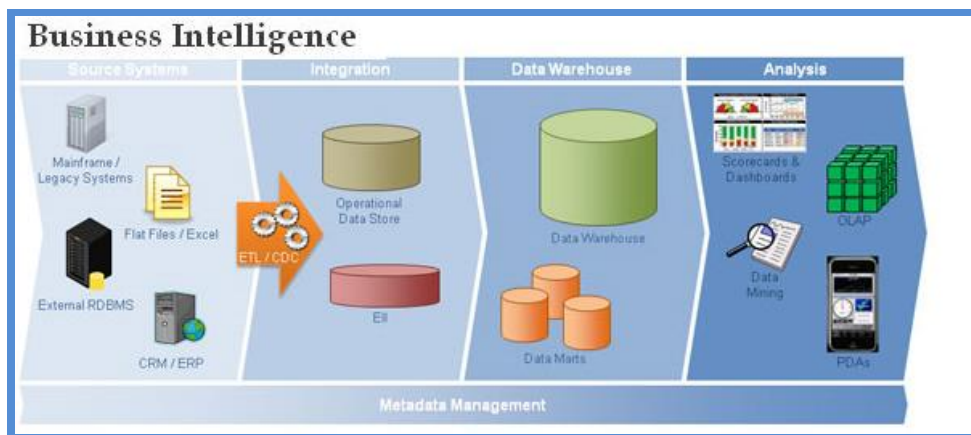


Figura 13: Business Intelligence (Kimball y Ross., 2013).

## 1.2. Metodología

### 1.2.1. Metodología de Ralph Kimball

La metodología de Ralph Kimball se enfoca principalmente en el diseño de la base de datos que almacenará la información para la toma de decisiones.

El diseño se basa en la creación de tablas de hechos, es decir, tablas que contengan la información numérica de los indicadores a analizar, o sea la parte cuantitativa de la información para la toma de decisiones.

Las tablas anteriores se relacionan con tablas de dimensiones, las cuales contienen la información cualitativa, de los indicadores, es decir, toda aquella información que clasifique la información requerida. (Kimball, R. y Ross, M., 2013).

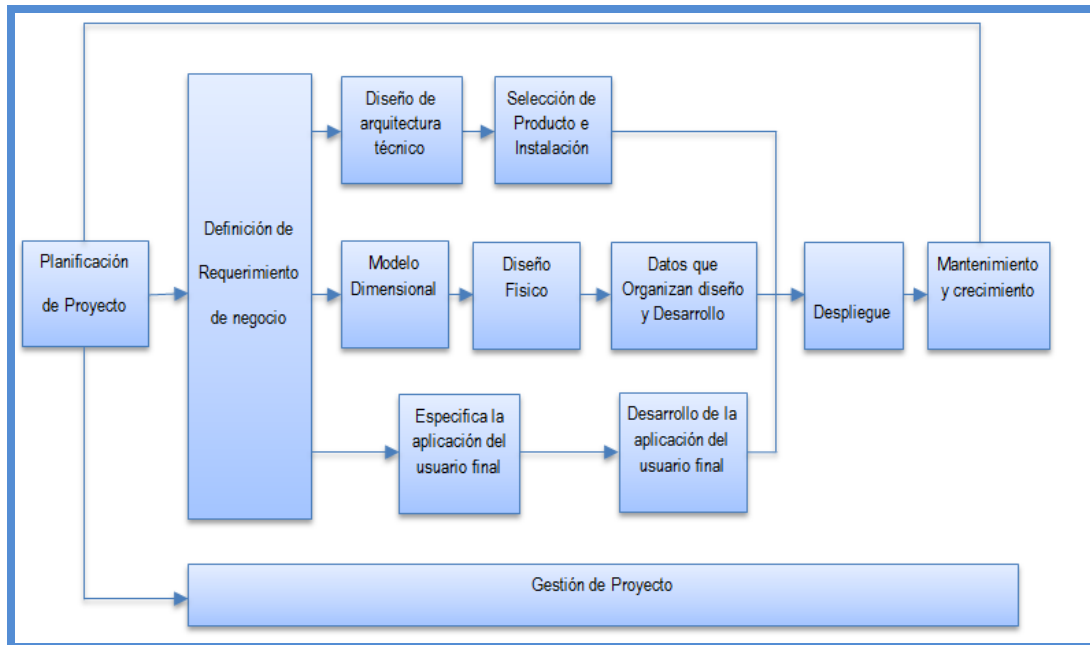


Figura 14: Metodología de implementación (Kimball, R. Y Ross, M., 2013)

### a. Planificación del Proyecto

La planificación busca identificar la definición y el alcance del proyecto de data warehouse, incluyendo justificaciones del negocio y evaluaciones de factibilidad. La planificación del proyecto se focaliza sobre recursos, perfiles, tareas, duraciones y secuencialidad.

El plan de proyecto resultante identifica todas las tareas asociadas con el ciclo de vida del datawarehouse e identifica las partes involucradas. (Kimball, R. y Ross, M., 2013).

### b. Definición de Requerimientos del negocio

Un factor determinante en el éxito de un proceso de Data Warehouse es la interpretación correcta de los diferentes niveles de requerimientos expresados por los diferentes niveles de usuarios. Aquí se identificará la información que requiere el usuario para desempeñar sus tareas. En esta etapa se especifica las funciones específicas que se obtendrán del Data Mart describiendo con claridad los requerimientos tales como:

- Definir los requerimientos del propietario.
- Definir los requerimientos del usuario final.

Estos requerimientos permitirán tener el ambiente operativo en el que se entregará el Data Mart. (Kimball, R. y Ross, M., 2013).

### c. Diseño Técnico de la Arquitectura

Los ambientes de data warehouse requieren la integración de numerosas tecnologías. Se debe tener en cuenta tres factores: los requerimientos del negocio, los actuales ambientes técnicos y las directrices técnicas estratégicas futuras planificadas para de

esta forma poder establecer el diseño de la arquitectura técnica del ambiente de data warehouse. (Kimball y Ross 2013).

#### **d. Modelo Dimensional**

La definición de los requerimientos del negocio determina los datos necesarios para cumplir los requerimientos analíticos de los usuarios. Diseñar los modelos de datos para soportar estos análisis requiere un enfoque diferente al usado en los sistemas operacionales. Básicamente se comienza con una matriz donde se determina la dimensionalidad de cada indicador y luego se especifican los diferentes grados de detalle (atributos) dentro de cada concepto del negocio (dimensión), como así también la granularidad de cada indicador (variable o métrica) y las diferentes jerarquías que dan forma al modelo dimensional del negocio (BDM) o mapa dimensional. (Kimball, R. y Ross, M., 2013).

#### **e. Diseño Físico**

El diseño físico de las base de datos se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Algunos de los elementos principales de este proceso son la definición de convenciones estándares de nombres y seteos específicos del ambiente de la base de datos. La indexación y las estrategias de particionamiento son también determinadas en esta etapa. (Kimball, R. y Ross, M., 2013).

#### **f. Diseño y Desarrollo de Presentación de Datos**

Las principales sub-etapas de esta zona del ciclo de vida son: la extracción, la transformación y la carga (ETL process). Se definen como procesos de extracción a aquellos requeridos para obtener los datos que permitirán efectuar la carga del Modelo Físico acordado. Así mismo, se definen como procesos de transformación los procesos para convertir o recodificar los datos fuente a fin poder efectuar la carga efectiva del Modelo Físico. Por otra parte, los procesos de carga de datos son los procesos requeridos para poblar el Data Warehouse.

Todas estas tareas son altamente críticas pues tienen que ver con la materia prima del data warehouse: los datos. La desconfianza y pérdida de credibilidad del data warehouse serán resultados inmediatos e inevitables si el usuario choca con información inconsistente. Es por ello que la calidad de los datos es un factor determinante en el éxito de un proyecto de data warehouse. Es en esta etapa donde deben sanearse todos los inconvenientes relacionados con la calidad de los datos fuente. (Kimball, R. y Ross, M., 2013).

#### **g. Selección de Productos e Instalación**

Utilizando el diseño de arquitectura técnica como marco, es necesario evaluar y seleccionar componentes específicos de la arquitectura como ser la plataforma de hardware, el motor de base de datos, la herramienta de ETL o el desarrollo pertinente, herramientas de acceso, etc. Una vez evaluados y seleccionados los componentes

determinados se procede con la instalación y prueba de los mismos en un ambiente integrado de data warehouse.

#### **h. Especificación de Aplicaciones para Usuarios Finales**

No todos los usuarios del data warehouse necesitan el mismo nivel de análisis. Es por ello que en esta etapa se identifican los diferentes roles o perfiles de usuarios para determinar los diferentes tipos de aplicaciones necesarias en base al alcance de los diferentes perfiles (gerencial, analista del negocio, vendedor, etc.) (Kimball, R. y Ross, M., 2013).

#### **i. Desarrollo de Aplicaciones para Usuarios Finales**

Siguiendo a la especificación de las aplicaciones para usuarios finales, el desarrollo de las aplicaciones de los usuarios finales involucra configuraciones del metadata y construcción de reportes específicos. Una vez que se ha cumplido con todos los pasos de la especificación y se tiene la posibilidad de trabajar con algunos datos de prueba, comienza el desarrollo de la aplicación. (Kimball, R. y Ross, M., 2013).

#### **j. Implementación**

La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Hay varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación. Todas estas tareas deben ser tenidas en cuenta antes de que cualquier usuario pueda tener acceso al data warehouse. (Kimball y Ross 2013).

#### **k. Mantenimiento y crecimiento**

Data Warehouse es un proceso, de etapas bien definidas con comienzo y fin, pero de naturaleza espiral, pues acompaña a la evolución de la organización durante toda su historia. Se necesita continuar con los relevamientos de forma constante para poder seguir la evolución de las metas por conseguir. Según afirma Kimball, “si se ha utilizado el Ciclo de Vida, el data warehouse está preparado para evolucionar y crecer”. Al contrario de los sistemas tradicionales, los cambios en el desarrollo deben ser vistos como signos de éxito y no de falla. Es importante establecer las prioridades para poder manejar los nuevos requerimientos de los usuarios y de esa forma poder evolucionar y crecer. (Kimball y Ross 2013).

#### **l. Gerenciamiento del Proyecto**

El gerenciamiento del proyecto asegura que las actividades del ciclo del datawarehouse se lleven en forma y sincronizadas. Como lo indica el diagrama, el gerenciamiento acompaña todo el ciclo de vida. Entre sus actividades principales se encuentra el monitoreo del estado del proyecto y la comunicación entre los requerimientos del negocio y las restricciones de información para poder manejar correctamente las expectativas en ambos sentidos. (Kimball, R. y Ross, M., 2013).



<b>FASES</b>	<b>DESCRIPCION</b>	<b>ENTREGABLES</b>
<b>Marco Ontológico</b>	<ul style="list-style-type: none"> <li>▪ Investigación conceptual e histórica de la Metodología a aplicar.</li> <li>▪ Investigación y análisis sobre la Refactorización a una Base de Datos.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Docum. de la Metodología Ralph Kimball.</li> <li>▪ Docum. de mejora de procesos para una Base de Datos utilizando Refactorización.</li> </ul>
<b>Planeación</b>	<ul style="list-style-type: none"> <li>▪ Seleccionar la estrategia de implementación</li> <li>▪ Seleccionar la metodología de desarrollo</li> <li>▪ Seleccionar el ámbito de implementación</li> <li>▪ Seleccionar el enfoque arquitectónico</li> <li>▪ Desarrollar un cronograma y presupuesto del proyecto.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Docum. de la estrategia de implementación</li> <li>▪ Docum. de la metodología de desarrollo.</li> <li>▪ Docum. del enfoque arquitectónico</li> <li>▪ Docum. del programa y presupuesto del proyecto</li> </ul>
<b>Determinación de Requerimientos</b>	<ul style="list-style-type: none"> <li>▪ Definir los requerimientos del usuario final.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Documento requerimiento del usuario</li> </ul>
<b>Análisis de Requerimientos</b>	<ul style="list-style-type: none"> <li>▪ Análisis de cada Requerimiento encontrado mediante el Modelo Starnet.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Modelo Starnet por requerimiento.</li> </ul>
<b>Diseño y Modelización</b>	<ul style="list-style-type: none"> <li>▪ Diseño de la Base de Datos Transaccional (Modelo Estrella).</li> <li>▪ Diseño de la Base de Datos con Refactorización (Modelo Estrella).</li> <li>▪ Diseño del Sistema de Extracción de Datos.</li> <li>▪ Preparación de los datos.</li> <li>▪ Diseño de la administración de datos</li> </ul>	<ul style="list-style-type: none"> <li>▪ Documento Identificación de los Componentes y clases del Modelo.</li> <li>▪ Especificación de las claves primarias y foráneas</li> <li>▪ Esquema estrella de la Base de Datos transaccional y refactoring.</li> <li>▪ Diagrama del Diseño Físico del Data Mart</li> </ul>
<b>Implementación</b>	<ul style="list-style-type: none"> <li>▪ Poblamiento de la Base de Datos Refactoring.</li> <li>▪ Extracción de los datos y transformación.</li> <li>▪ Carga de los datos válidos.</li> <li>▪ Administración del Data Mart.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Construcción de las Tablas y la Base de base refactoring en Ms SQL Server 2008 R2</li> <li>▪ Documento en Ms Excel para la población de Datos.</li> <li>▪ Paquete de Servicio de Transformación de Datos (DTS)</li> <li>▪ Documento Sentencia de SQL para cada paso del DTS</li> </ul>
<b>Prueba</b>	<ul style="list-style-type: none"> <li>▪ Prueba de Performance</li> <li>▪ Prueba de calidad de datos</li> <li>▪ Prueba de los requerimientos</li> </ul>	<ul style="list-style-type: none"> <li>▪ Tablas simples y gráficos dinámicos.</li> <li>▪ Informe final del Sistema Estratégico con los resultados requeridos por el usuario.</li> </ul>

Tabla 2: Metodología de aplicación para BI



## **CAPITULO II**

### **DESARROLLO DEL PROYECTO DE TESIS**

#### **2.1. Planificación del Proyecto**

##### **2.1.1. Selección de la Estrategia de Implementación**

La estrategia de implementación del proyecto contiene las siguientes tareas:

1. Conocer la realidad de la empresa y familiarizarnos con la cultura organizacional.
2. Analizar la documentación y procedimientos de la empresa CARUBI SAC.
3. Analizar los sistemas transaccionales.
4. Recolectar los requerimientos empresariales.
5. Analizar los requerimientos.
6. Realizar el diseño y la arquitectura para los Data Mart.
7. Implementar los Data Mart.

##### **2.1.2. Selección de la Metodología de Desarrollo**

La implementación de un Data Mart es un proceso complejo, es por esto que deben usarse las mejores prácticas existentes. Una de ellas es el método propuesto por Ralph Kimball, el cual se utilizará para desarrollar la metodología de trabajo, adaptándola al caso específico de este proyecto.

##### **2.1.3. Selección del Ámbito de Implementación**

El ámbito de la implementación de los Data Mart fue determinado basándose en los requerimientos de información del personal de nivel medio o Jefatura de la empresa CARUBI SAC.; y el análisis con el apoyo del equipo desarrollador.

Para ello se formuló las siguientes preguntas:

- a) ¿Cuál es el área más recomendable estratégicamente para aplicar los Data Marts?
- b) ¿Cuál es la priorización de desarrollo de los Data Marts en los procesos de las principales áreas de la empresa?
- c) ¿Cuál es el rango de consultas empresariales a los que se debe responder inicialmente los Data Marts?

Después de haber realizado el análisis del funcionamiento de la organización, las respuestas obtenidas fueron las siguientes:

- a) Las áreas de Ventas, es la más recomendable, en el cual debe ser aplicada los Data Marts, por ser la área que brindan mayor información acerca de la situación de la empresa.
- b) De acuerdo a las entrevistas, información y documentación de los sistemas, la priorización de desarrollo será la siguiente
  - 1° Proceso de Ventas de Importación y Exportación
  - 2° Proceso de Inventario
- c) Se cree conveniente considerar un rango no mayor a 10, para poder agilizar el tiempo de respuesta y el usuario tenga una mejor visión de sus resultados.

#### **2.1.4. Selección del enfoque arquitectónico**

El enfoque arquitectónico para los Data Marts estará compuesto de los siguientes niveles

##### ➤ **Back Room**

Está compuesto por:

- 1 Servidor de Base de Datos HP ProLiant G6 Intel Xeon de 2.53 GHz, cuya estructura se encuentra en SQL Server 2008 R2.
- Paquete de acceso para la tecnología Microsoft.

##### ➤ **Front Room**

- PC's clientes, que serán cada una de las computadoras las que se accede a la información que brindaran los Data Marts, a través de la tecnología Microsoft.

## 2.1.5. Desarrollo de un Cronograma y del Presupuesto del Proyecto

### 2.1.5.1. Presupuesto

#### a. Recursos Humanos

Concepto	Cantidad	Costo s/.
Bachiller	2	2000.00
Asesor	1	1000.00

#### b. Bienes: materiales, equipos, software

Concepto	Cantidad	Costo s/.
Agenda de Trabajo	1	20.00
Papel Bond A4(Millar)	2	20.00
Lapicero	2	5.00
Cartucho de tina a color	1	60.00
Cartucho de tinta negro.	1	60.00
DVD's.	3	10.00
Computadora Portátil ASUS, Procesador Intel Core i5.	1	2400.00
Impresora HP	1	200.00
SQL Server 2008 R2 Enterprise	1	450.00
Microsoft Office 2010 Professional	1	400.00
S.O. Windows 8 Professional	1	400.00

Concepto	Cantidad	Costo
Servicio de Internet – Hogar	1	150.00
Fotocopiado	1	30.00
Movilidad	1	50.00
Empastados	3	50.00
Anillado	3	15.00
Telefonía.	1	75.00

#### c. Presupuesto Total

<b>Personal</b>	3000.00
<b>Bienes</b>	3845.00
<b>Servicios</b>	370.00
<b>TOTAL</b>	<b>7215.00</b>

Tabla 3: Presupuesto Total

## 2.1.5.2. Cronograma de actividades

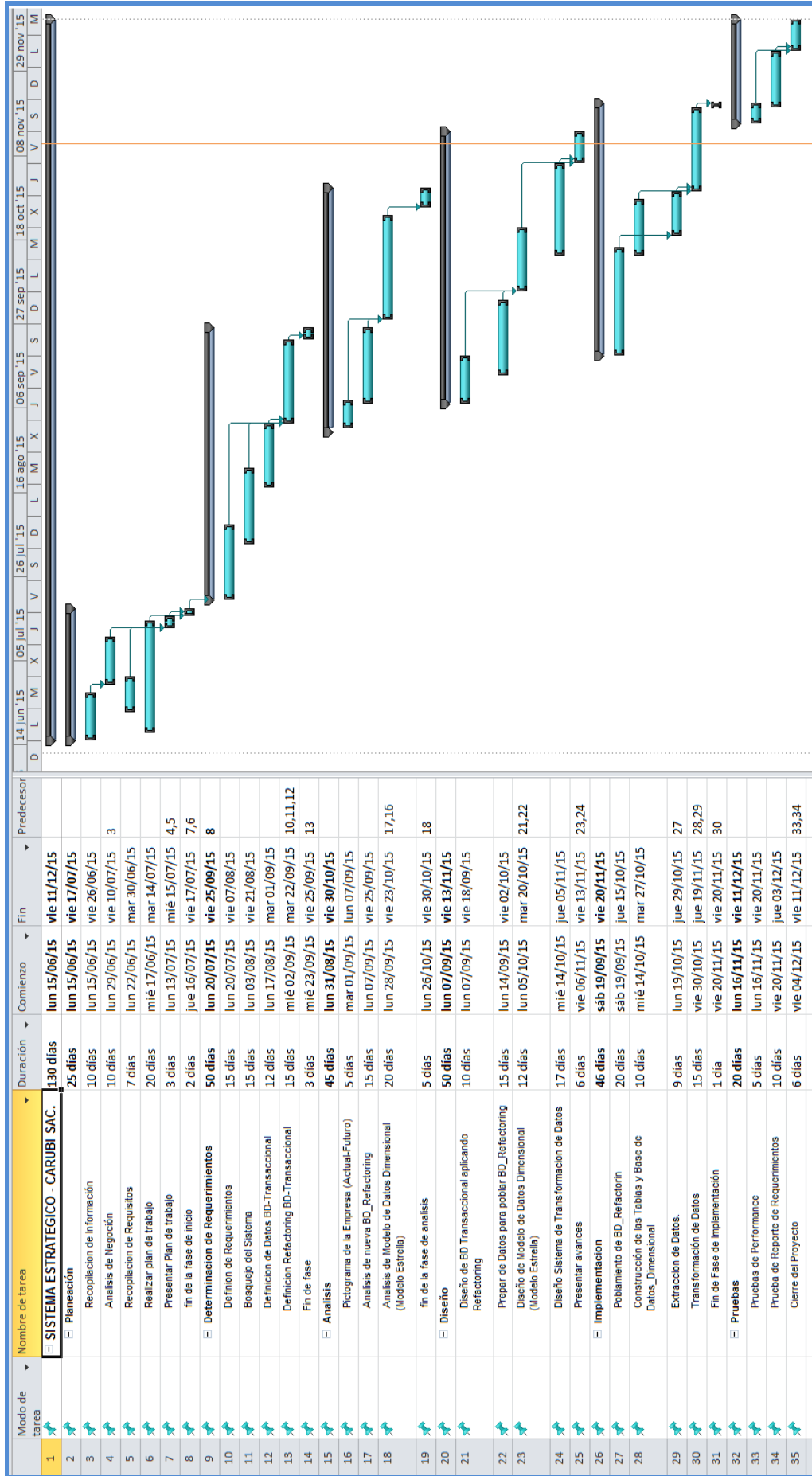


Figura 15: Cronograma de Actividades

## 2.1.6. Desarrollo del escenario del uso empresarial

Para este proyecto de tesis el escenario de uso empresarial es el que se muestra a continuación con sus respectivos componentes:

### a. Descripción de los Stakeholders

- **Personal Involucrado en el Proyecto**

Nombre	Representa	Rol
<b>Gerente General</b>	Personal encargado de Planificar, organizar y dirigir la administración de la empresa, de acuerdo a normas y políticas establecidas.	<ul style="list-style-type: none"> <li>▪ Ejercer la representación legal de la empresa, tomando las decisiones definitivas de acuerdo a las atribuciones conferidas.</li> <li>▪ Controlar el funcionamiento operativo de la empresa, responsabilizándose de la situación económico-financiera de la empresa, tomando decisiones definitivas en provecho del cumplimiento de los objetivos de la empresa.</li> <li>▪ Ejercer el control interno de todas las operaciones y actividades que se realizan en la empresa.</li> </ul>
<b>Jefe de Ventas</b>	Personal encargado de realizar y vigilar el proceso de ventas, recepción y registro de documentación y administración.	<ul style="list-style-type: none"> <li>▪ Plantea las estrategias de ventas.</li> <li>▪ Propone las metas a llegar en ventas</li> <li>▪ Consolidar reportes de ventas por vendedor.</li> <li>▪ Consolidar reportes de ventas totales.</li> </ul>
<b>Jefe de Inventarios</b>	Personal encargado velar por el mantenimiento del kárdex registrando entradas y salidas de productos.	<ul style="list-style-type: none"> <li>▪ Procesar los reportes de entrada de productos con los reportes de salida y que este en equilibrio.</li> <li>▪ Mantener al día el stock de productos.</li> </ul>

Tabla 4: Personal involucrado en el Proyecto- Descripción de usuarios del Data Mart

## 2.2. Determinación de los Requerimientos

### 2.2.1. Requerimientos de los Stakeholders

En la primera entrevista con Stakeholders debemos recopilar todos los requerimientos y aclarar las dudas que tengan; teniendo en cuenta la necesidad principal que es el área de ventas y así poder realizar del Data Mart enfocado a las necesidades de la empresa.

- **¿Por qué construir un Data Mart?**

Para desarrollar estrategias y brindar un soporte a la toma de decisiones, y así lograr una mayor participación en el mercado nacional e internacional; haciendo uso del análisis de sus datos históricos y actuales, beneficiando de esta manera un mejor desempeño de la empresa CARUBI SAC.

- **¿Cuál será el impacto sobre la organización?**

El impacto del Data Mart sobre CARUBI SAC., será beneficioso a mediano y a largo plazo de acuerdo a la rapidez con que se tome las decisiones y a la creatividad que se emplee, permitiendo la reducción de tiempo en la obtención de reportes solicitados por la jefatura, que servirán para tomar decisiones en el momento necesario, lo cual le permitirán tomar ventajas en el mercado y sobre sus competidores.

- **¿Cómo afecta nuestra inversión en Producción?**

La necesidad de adquirir un equipo que cuente con las características necesarias para que pueda soportar un Data Mart y emitir los reportes rápidamente; también se necesita una instalación del Software.

La inversión está rodeando los \$500, esto puede variar si el usuario ya cuenta con el equipo necesario para el Data Mart.

- **¿Cuáles son los riesgos?**

Los riesgos al implementarse el Data Mart serian que la base de datos operacional tenga complicaciones, es decir, genere problemas de consistencia.

### **2.2.2. Requerimientos Funcionales del Usuario Final**

Para determinar los requerimientos empresariales se realizaron entrevistas a los usuarios finales que utilizarán los Data Mart basados en la metodología de Ralph Kimball.

#### **A. Área de Ventas de Importación y Exportación**

R1 ¿Cuál es la cantidad total de ventas en soles de los productos que se importan y exportan en un determinado mes?

R2 ¿Cuál es la cantidad de ventas en soles por vendedor en un mes determinado?

R3 ¿Cuál es la cantidad de productos vendidos que se importan y exportan, por tipo y categoría en un mes determinado?

R4 ¿Cuál es la cantidad de productos que se importan y exportan, comprados por cliente en un determinado mes?

R5 ¿Cuál es la venta total en soles de productos que se importan y exportan comprados por cliente en un determinado mes?

R6 ¿Cuál es el costo en soles de un producto de importación y exportación terminado?

R7 ¿Cuál es la utilidad por mes de un producto terminado?

#### **B. Área de Inventarios**

R8 ¿Cuál es la cantidad de productos que se compraron en un mes determinado?

R9 ¿Cuál es la cantidad de productos que tuvieron alguna transferencia, en un mes determinado y por un motivo?

R10 ¿Cuál es la cantidad de productos comprados por proveedor en un mes determinado?

NRO	CONSULTAS	UNIDAD
01	¿Cuál es la cantidad total de ventas en soles de los productos que se importan y exportan en un determinado mes?	Ventas soles
02	¿Cuál es la cantidad de ventas en soles por vendedor en un mes determinado?	Ventas soles
03	¿Cuál es la cantidad de productos vendidos que se importan y exportan, por tipo y categoría en un mes determinado?	Cantidad de productos
04	¿Cuál es la cantidad de productos que se importan y exportan, comprados por cliente en un determinado mes?	Cantidad de productos
05	¿Cuál es la venta total en soles de productos que se importan y exportan comprados por cliente en un determinado mes?	Ventas soles
06	¿Cuál es el costo en soles de un producto de importación y exportación terminado?	Costo soles
07	¿Cuál es la utilidad por mes de un producto terminado?	Porcentaje Utilidad
08	¿Cuál es la cantidad de productos que se compraron en un mes determinado?	Cantidad productos Comprados
09	¿Cuál es la cantidad de productos que tuvieron alguna transferencia, en un mes determinado y por un motivo?	Cantidad productos Transferidos
10	¿Cuál es la cantidad de productos comprados por proveedor en un mes determinado?	Cantidad de productos comprados

Tabla 5: Reportes y Unidades de Medición

### 2.2.3. Requerimientos no Funcionales

- Brindar mejor información a los usuarios, en el menor tiempo.
- Establecer seguridad adecuada para la administración de los data marts.
- Los Data Marts deben funcionar sobre la plataforma de Microsoft Windows 7, con el manejador de base de datos Microsoft SQL Server 2008 R2 y utilizando el Microsoft SQL Server 2008 Analysis Services como herramienta de diseño y construcción de los cubos.
- Utilizar tecnología Microsoft MS Excel 2010.



- Para el análisis de los requerimientos funcionales y no funcionales utilizaremos los datos que nos proporcionaron la empresa CARUBI SAC., además de los manuales de procedimientos y documentos proporcionados por el personal.

#### 2.2.4. Análisis de Requerimientos

En esta parte se hará un análisis de cada requerimiento del usuario final y definiendo algunas tablas de la Base de Datos Operacional de la empresa, que actuarían como dimensión dentro del Data Mart, también se definirá las posibles medidas dentro del Data Mart.

##### A. Área de Ventas de Importación y Exportación

R1 ¿Cuál es la cantidad total de ventas en soles de los productos que se importan y exportan en un determinado mes?

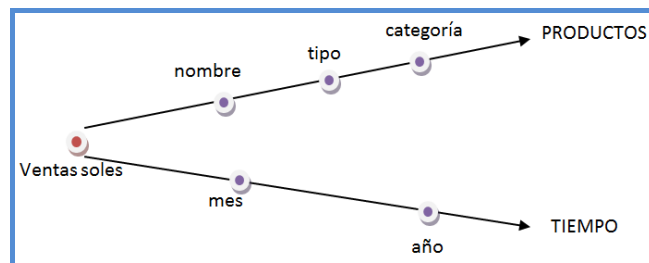


Figura 16: Requerimiento 1 - Análisis

R2 ¿Cuál es la cantidad de ventas en soles por vendedor en un mes determinado?

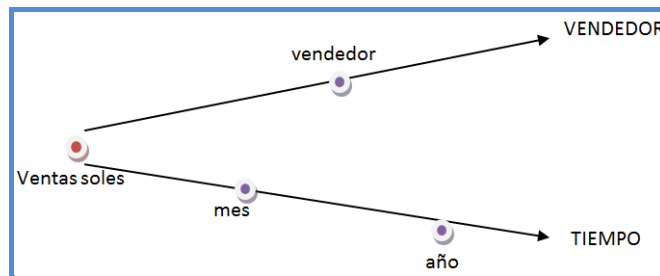


Figura 17: Requerimiento 2 - Análisis

R3 ¿Cuál es la cantidad de productos vendidos que se importan y exportan, por tipo y categoría en un mes determinado?

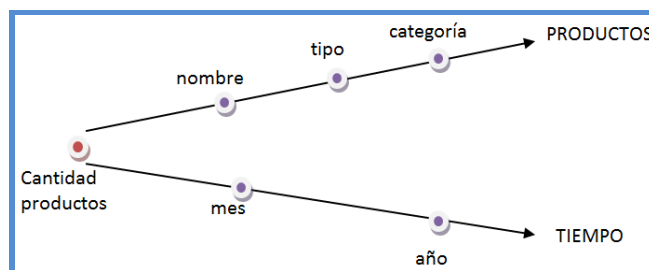


Figura 18 : Requerimiento 3 – Análisis

R4 ¿Cuál es la cantidad de productos que se importan y exportan, comprados por cliente en un determinado mes?

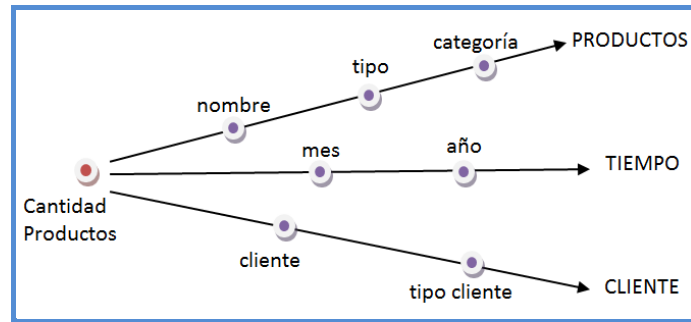


Figura 19: Requerimiento 4 - Análisis

R5 ¿Cuál es la venta total en soles de productos que se importan y exportan comprados por cliente en un determinado mes?

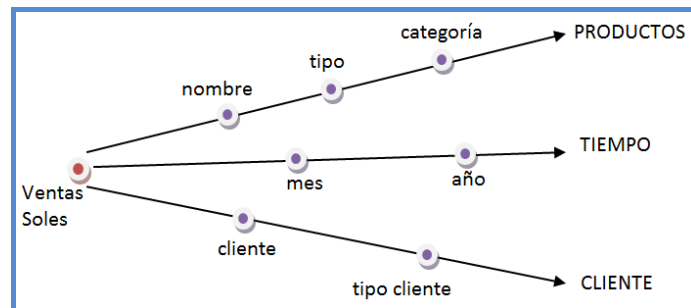


Figura 20: Requerimiento 5 - Análisis

R6 ¿Cuál es el costo en soles de un producto de importación y exportación terminado?

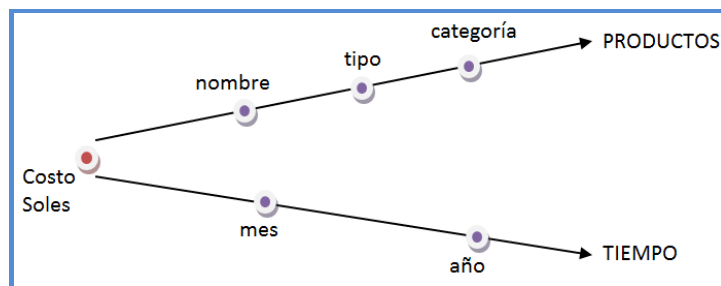


Figura 21: Requerimiento 6 - Análisis

R7 ¿Cuál es la utilidad por mes de un producto terminado?

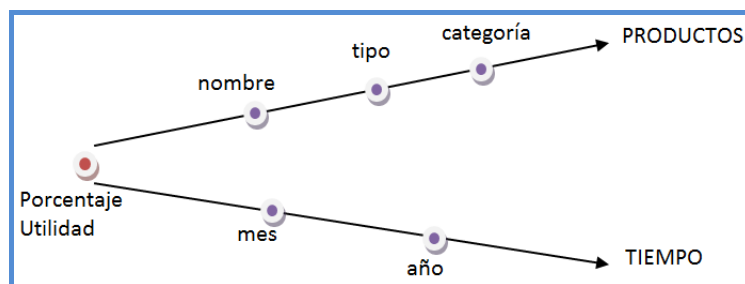


Figura 22: Requerimiento 7 - Análisis

## B. Área de Inventarios

R8 ¿Cuál es la cantidad de productos que se compraron en un mes determinado?

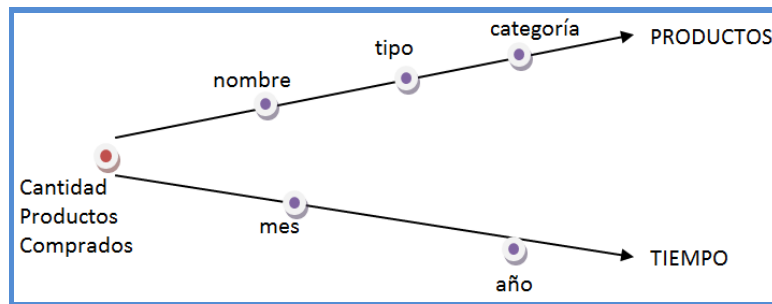


Figura 23: Requerimiento 8 - Análisis

R9 ¿Cuál es la cantidad de productos que tuvieron alguna transferencia, en un mes determinado y por un motivo?

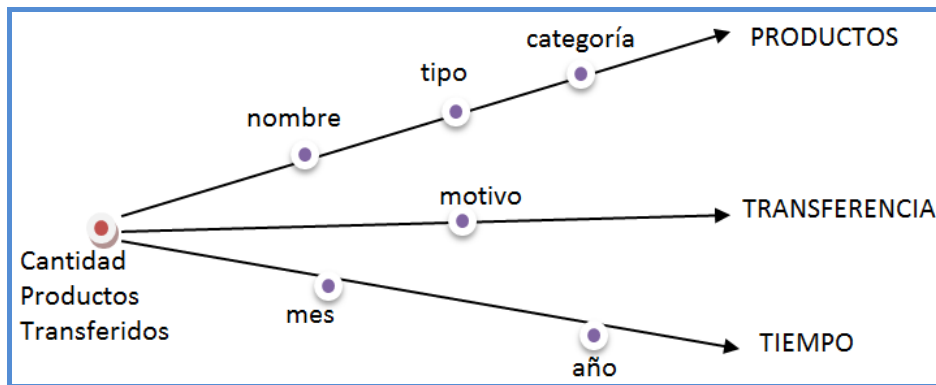


Figura 24: Requerimiento 9 - Análisis

R10 ¿Cuál es la cantidad de productos comprados por proveedor en un mes determinado?

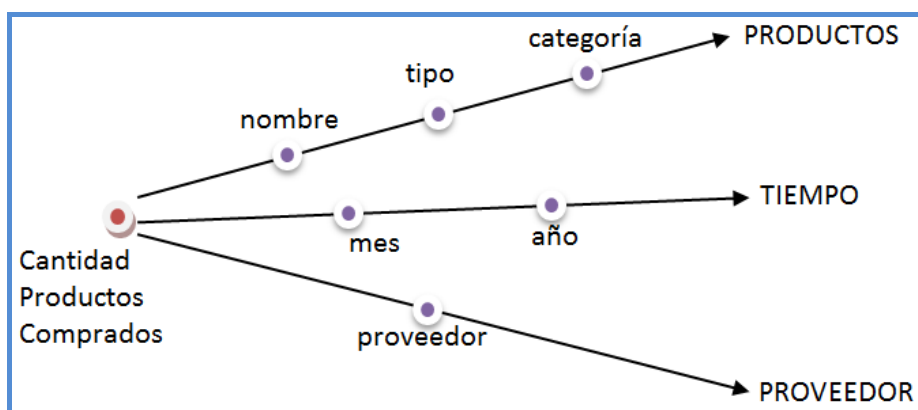


Figura 25: Requerimiento 10 - Análisis

## 2.3. Diseño Técnico de la Arquitectura

### 2.3.1. Nivel de Datos

Para el análisis de los datos, es recomendable empezar analizar los datos fuentes, requerimientos y procesos que maneja la empresa; se creara un tipo de la base de datos y la estructura de las tablas.

**Base de Datos Fuente (Transaccional):** Se utilizó la base de datos Carubi SAC. cuyo modelo esta descrito en la imagen de a continuación.

En este diagrama se muestran las tablas que se relacionan con el área de Inventario así como al área de Ventas, si bien es el modelo de datos es básico; pues, en el transcurso del proyecto se realizará un refactoring lo que permitirá adquirir datos relevantes para la empresa.

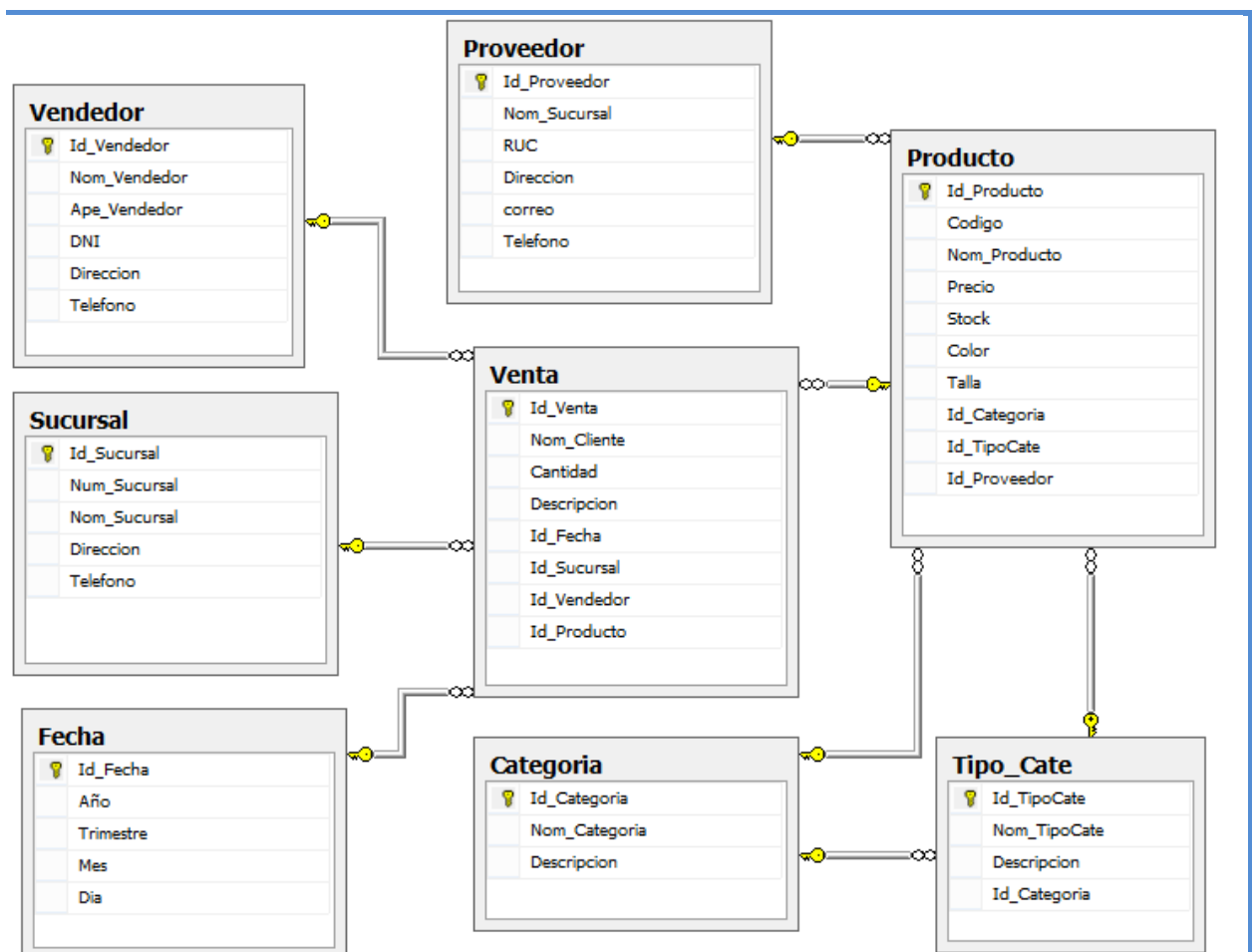


Figura 26: Enfoque Arquitectónico de Base Datos Transaccional

### 2.3.2. Nivel Técnico

A continuación mostramos el diagrama técnico que utilizaremos en la creación de Data Mart y con ellos los reportes solicitados por la empresa y Jefatura.

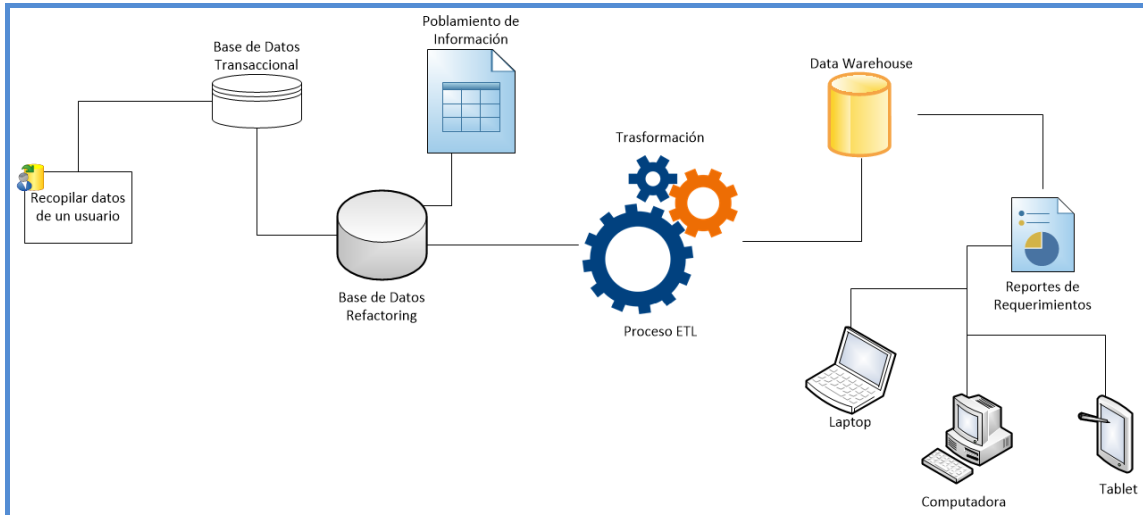


Figura 27: Diagrama técnico - creación de Data Mart

El enfoque arquitectónico Nivel Técnico para el Data Mart estará compuesto de los siguientes niveles

1. Recopilación de Información, esto nos permite saber el manejo de sus procesos; como se está desarrollando actualmente.
2. Base de Datos Transaccional, la BD Transaccional es el enfoque principal de acuerdo a sus necesidades de básicas que tiene el usuario; puede realizar sus ventas y llevar un inventario de sus productos.
3. Base de Datos Refactoring, es el proceso de investigación que se aplicara en el proyecto, ya que nos permitirá consolidar toda la información y realizar una mejora de datos y sin que nos haga falta ningún atributo que posteriormente se necesite en el Data Mart.
4. Poblamiento de Información, el poblamiento nos permitirá obtener datos reales que no se consideraron en el primer modelo de datos, esto permitirá poder obtener un mejor reporte cuando se realice la transformación de datos para el Data Mart.
5. Transformación – Proceso ETL, Tenemos que extraer todos los datos de la BD Refactoring, la cual ha sido creada y poblada para que haya una correcta extracción de datos así poder poblar el Data Mart.
6. Data Mart, Sera nuestra versión especializada de datos, lo cual nos permitirá crear las consultas y reportes necesarios y así satisfacer las expectativas del cliente.

7. Reportes acordes a los requerimientos, Esta parte será el resultado final del proyecto lo cual podrá ser accesible para el uso del cliente se presentara el producto en la plataforma de Microsoft Excel en el cual el cliente podrá interactuar de acuerdo a sus necesidades.

## **2.4. Implementación de la Refactorización**

### **2.4.1. Como hacer el Refactoring**

Lo ideal es hacer el refactoring sobre la BD Transaccional, nos permite tener un mayor control y organización para redactar el código y así poder corregirlo; lo mejor de poder usar refactoring es que podemos trabajar bajo pruebas unitarias ir implementando poco a poco sin falla y ejecutando lo que se desarrolla.

[La idea la explica perfectamente la metáfora de los dos sombreros. Según esta metáfora, un programador tiene a su disposición dos sombreros. Uno de ellos etiquetado "hacer código nuevo", y el otro con la etiqueta "arreglar código"]. *Metáfora de dos sombreros*

#### **VENTAJAS DE HACER REFACTORING**

- Por más claro que se tenga el código original es necesario realizar un análisis unitario para poder detectar fallas futuras
- Al presentarse fallas y desorden en el código aplicar las pruebas unitarias para detectar fallas y aplicar el refactoring etiquetando el código y ordenándolo.
- No es una pérdida de tiempo arreglar el código, al final del mismo se gana dicho tiempo. las modificaciones y añadido tardan menos y se pierde mucho menos tiempo en depurar y entender el código.

#### **DESVENTAJAS DE HACER REFACTORING**

- La modificación del código incrementa la complejidad de nuestro diseño, que es justo el efecto contrario del que intentábamos lograr al aplicarlo.
- Convocar y hacer mención que se está aplicando un refactoring; con el equipo de trabajo, para comentar el avance que se está realizando y evitar la sensación de que no se está avanzando. Una sensación que conduce a repercusiones anímicas negativas. A veces, el no comentar con el resto del equipo un refactoring que parece muy necesario, puede afectar a otros compañeros y de los que no éramos conscientes.

```
BD_Transql - JOSENBRA_VDE_50)
USE [BDtran_Carubi]
GO
/****** Object: Table (dbo).[Fecha] Script Date: 10/23/2015 01:58:23 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[Fecha]([
    Id_Fecha] [int] NOT NULL,
    [Año] [date] NULL,
    [Trimestre] [int] NULL,
    [Mes] [date] NULL,
    [Dia] [date] NULL,
    PRIMARY KEY CLUSTERED
    (
        [Id_Fecha] ASC
    ) WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUE_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]
) ON [PRIMARY]
GO
/****** Object: Table (dbo).[Categoria] Script Date: 10/23/2015 01:58:23 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
SET ANSI_PADDING ON
GO
CREATE TABLE [dbo].[Categoria](

```

Script de la BD Transaccional, este código está implementado en una herramienta tradicional



```
BDRefact_Carubisiql x
1 use BDRefact_Carubi
2 GO
3
4
5 --Creando tabla Vendedor
6 CREATE TABLE Vendedor(
7 Id_Vendedor int not null PRIMARY KEY,
8 Nom_Vendedor VARCHAR(50),
9 Ape_Vendedor VARCHAR(50),
10 DNI int,
11 Direccion VARCHAR(50),
12 Telefono int
13 )
14
15
16 --Creando tabla Cliente
17 CREATE TABLE Cliente(
18 Id_Cliente int not null PRIMARY KEY,
19 Nom_Cliente VARCHAR(50),
20 Ape_Cliente VARCHAR(50),
21 DNI int,
22 Direccion VARCHAR(50),
23 email VARCHAR(50),
24 Telefono int
25 )
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
```

## 2.4.2. Diagrama de Base de Datos aplicando Refactoring

Después de haber realizado el refactoring a la BD Transaccional se genera una mejor estructura y diseño la cual nos permitirá y facilitara un mejor análisis y poblamiento para continuar con el siguiente paso; el Data Mart.

El diagrama generado nos especifica más a detalle las tablas involucradas para poder generar los requerimientos solicitados por el Usuario.

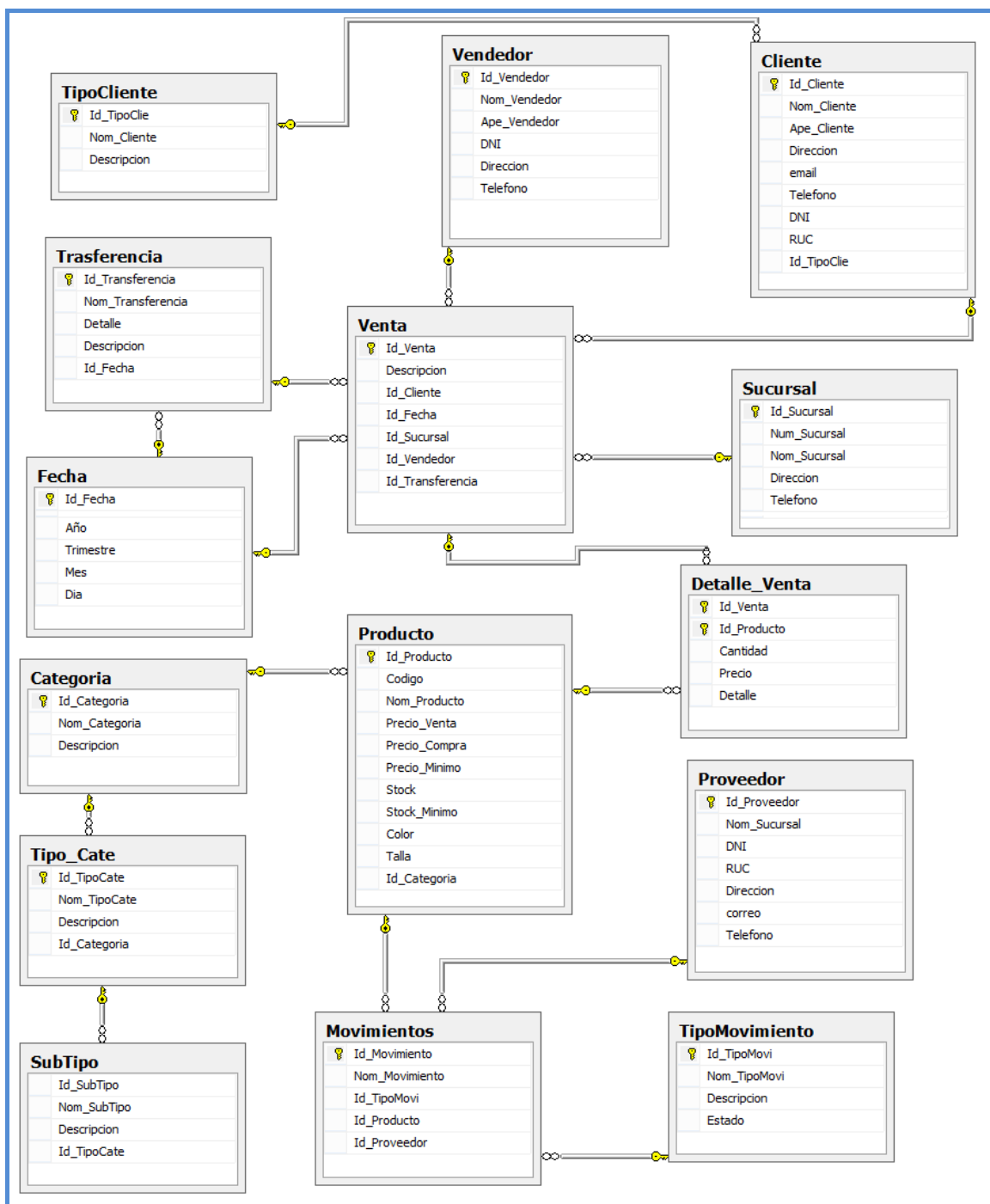


Figura 28: Diagrama de BD-Refactoring



## 2.5. Modelado Dimensional

Esta parte llamada Modelamiento Dimensional, se busca presentar los datos en un marco de trabajo estándar.

Identificaremos los componentes del modelo, desarrollando el método de los cuatro pasos. Luego, se construirá el primer diagrama para el modelo y se detallará cada componente del diagrama.

### 2.5.1. Identificación de los Componentes del Modelo

Para identificar los componentes utilizaremos el método propuesto por Ralph Kimball.

#### El Método de los cuatro pasos

Consiste en cuatro elecciones:

- Elección del Data Mart
- Elección del Objetivo de la Tabla de Hechos
- Elección de las Dimensiones
- Elección de los Hechos

Cada paso está desarrollado a continuación:

#### a. Elección del Data Mart

En este paso identificaremos el posible Data Mart que se puede construir en la empresa Carubi SAC. En nuestro caso identificaremos el área que está en estudio por este proyecto, el cual ya tenemos conocimiento sobre su modelo general de datos.

Entonces, para identificar nuestro Data Mart en estudio recurrimos al siguiente método:

#### Método de la Matriz

Se mencionará los posibles hechos relacionados y que son usados juntos. Listamos las entidades que apoyan a estos hechos.

### **Listado de los Data Marts**

Según la situación general de la empresa, se identificó las posibles fuentes de Data Marts que corresponderán a las filas de la matriz:

Área de Ventas de Importación y Exportación

Área de Inventarios

### **Listado de las Dimensiones**

Para los posibles Data Mart mencionados anteriormente listamos las siguientes posibles dimensiones:

Producto  
Tiempo\_Ventas  
Tiempo\_Compras  
Vendedor  
Cliente  
Proveedor  
Transferencia

### **Marcado de las Intersecciones**

Ordenamos las filas y las columnas en una tabla, dando forma a la matriz, y marcamos las intersecciones donde exista una dimensión relacionada a un Data Mart.

DIMENSIONES	Áreas	
	Área de Inventarios	Área de Ventas
TIEMPO	X	X
VENDEDOR		X
PRODUCTO	X	X
CLIENTE		X
PROVEEDOR	X	
TRANSFERENCIAS	X	

Tabla 6: Matriz de Intersecciones

Para el fácil manejo de las consultas hechas por el usuario, en el Data Mart se consideró dos Tablas de Hechos: Ventas y de Inventario.

## b. Elección del Objetivo de las Tablas de Hechos

Este paso consiste en declarar como es el registro del hecho en las tablas de hechos, es decir, hay que definir claramente y exactamente que registros de cada tabla de hechos figurará en el diseño del modelado del Data Mart.

Tomamos la siguiente definición para cada tabla de hechos:

Tabla de Hechos	Objetivo
Ventas	“Tener un mejor control y gestión de las ventas”.
Inventario	“Controlar las compras y transferencias de productos que maneja la empresa”.

Tabla 7: Objetivos de las tablas de Hecho

Al enunciar el objetivo de cada Data Mart, van a intervenir las posibles dimensiones que se explicarán a continuación.

## c. Elección de las dimensiones

Al establecer el objetivo para cada tabla de hechos, podemos escoger cuales serán nuestras dimensiones para cada una de estas tablas. El objetivo por sí mismo establece cuales serán nuestras dimensiones.

A continuación por cada objetivo escogeremos las dimensiones:

Tabla de Hechos	Objetivos	Dimensiones
Ventas	“Tener un mejor control y gestión de las ventas”	Tiempo_Ventas Producto Cliente Vendedor

Tabla 8: Dimensiones de la tabla de Hechos Ventas

Tabla de Hechos	Objetivos	Dimensiones
Inventario	“Controlar las compras y transferencias de productos de la empresa”.	Tiempo_Compras Proveedor Producto Transferencias

Tabla 9: Dimensiones de la tabla de Hechos Inventario

Como se puede ver la mayoría de las dimensiones elegidas pertenecen a las Entidades del Modelo de Datos General descritos anteriormente. Otras dimensiones son la unión de entidades.

#### d. Elección de los Hechos

El Objetivo de cada tabla de hechos, permite definir los hechos y hace claro el alcance que estos hechos deben tener. En nuestro caso tenemos:

Tabla de Hechos	Hechos
Ventas	Ventas Soles Unidades Vendidas Costo Soles Porcentaje Utilidad
Inventario	Cantidad Productos Comprados Cantidad Productos con Transferencias

Tabla 10: Elección de Hechos

#### 2.5.2. Diagrama de la tabla de Hechos

Para esta etapa se prepara un diagrama lógico para cada Tabla de Hechos completada. Cada Diagrama nombra a la Tabla de Hechos, establece su Objetivo y muestra todas las dimensiones conectadas a la Tabla de Hechos.

##### ▪ Diagrama de la Tabla de Hechos Ventas

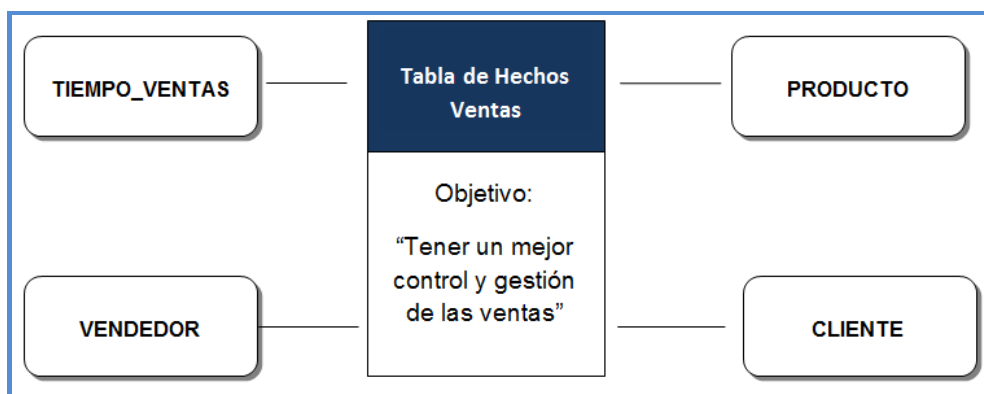


Figura 29: Diagrama de Tabla de Hechos Ventas

▪ **Diagrama de la Tabla de Hechos Inventario**

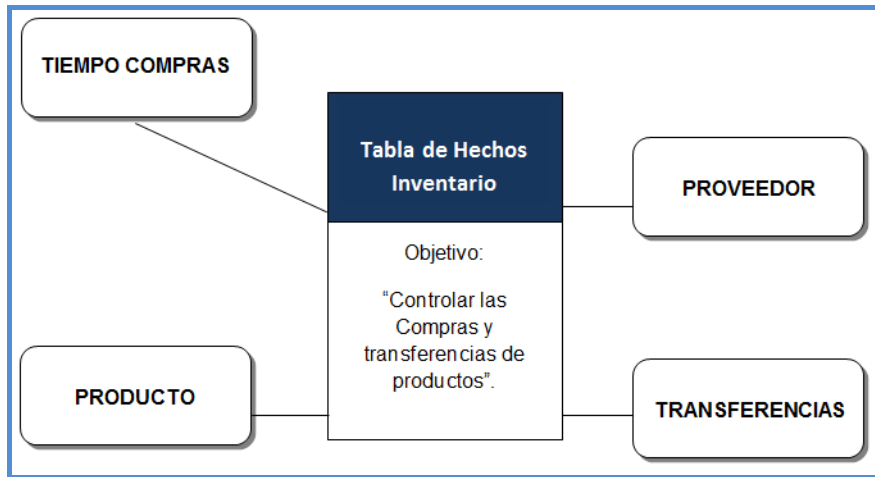


Figura 30: Diagrama de Tabla de Hechos Inventario

▪ **Detalle de las Tablas de Hechos**

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
Dimensión Tiempo Ventas	Tiempo_Key	Llave primaria única para la Dimensión Tiempo Ventas
Dimensión Producto	Producto_Key	Llave primaria única para la Dimensión Producto
Dimensión Vendedor	Vendedor_key	Llave primaria única para la Dimensión Vendedor
Dimensión Cliente	Cliente_Key	Llave primaria única para la Dimensión Cliente
Dimensión Proveedor	Proveedor_Key	Llave primaria única para la Dimensión Proveedor
Dimensión Tiempo Compras	Tiempo_Compras_Key	Llave primaria única para la Dimensión Tiempo Compras
Dimensión Transferencia	Transferencia_Key	Llave primaria única para la Dimensión Transferencia

Tabla 11: Detalle de las claves de las dimensiones

Nombre de la Tabla	Nombre de la Columna	Descripción de la Columna
Hecho Ventas	Cantidad Productos	Cantidad productos vendidos
Hecho Ventas	Ventas Totales	Total de ventas
Hecho Ventas	Porcentaje de Utilidad	Porcentaje de Utilidad
Hecho Ventas	Costo Ventas	Costo Total
Hecho Inventario	Cantidad Productos Transferidos	Cantidad Productos Transferidos
Hecho Inventario	Cantidad Productos Comprados	Cantidad Productos Comprados

Tabla 12: Detalle de las Medidas de las Tablas de Hecho

Por consiguiente, de las dos tablas completadas anteriormente se obtienen las siguientes Tablas de Hechos:

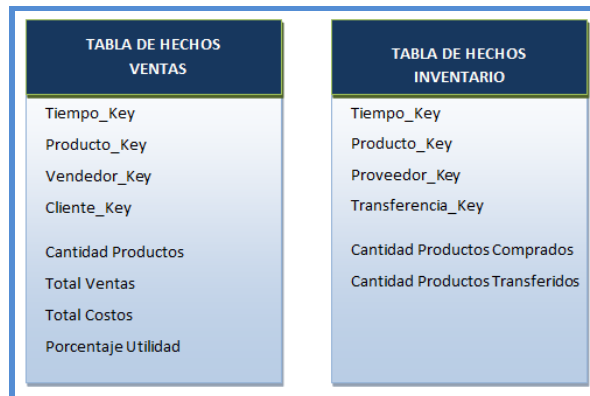


Figura 31: Tablas de Hechos del Data Mart

### ▪ Detalle de las Tablas Dimensión

De acuerdo a nuestras Dimensiones elegidas anteriormente construimos los detalles y sus jerarquías de cada una de ellas.

#### i. Dimensión Tiempo

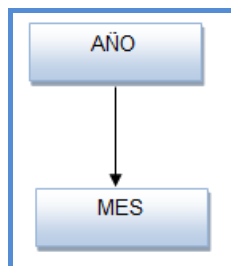


Figura 32: Jerarquía de los Atributos de la Dimensión Tiempo

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Año	Representa el año.	No Actualizar	2012,2013,...
Mes	Representa los meses que posee un año	No Actualizar	Julio, Agosto,...

Tabla 13: Detalle de Dimensión Tiempo Ventas/Compras

## ii. Dimensión Vendedor

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Nombre Vendedor	Representa el nombre del vendedor.	No Actualizar	Roberto Zavaleta

Tabla 14: Detalle de Dimensión Vendedor

## iii. Dimensión Cliente

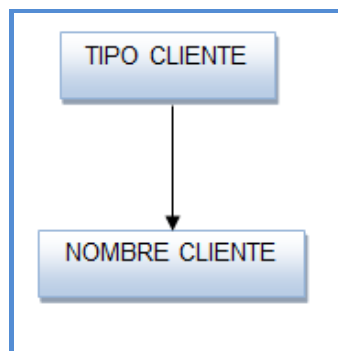


Figura 33: Jerarquía de los Atributos de la Dimensión Cliente

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Tipo Cliente	Representa el tipo de cliente con el que trabaja la empresa.	Actualizar	Natural, Jurídico
Nombre Cliente	Representa el nombre del cliente.	No Actualizar	Liliana S.A.C.

Tabla 15: Detalle de Dimensión Cliente

#### iv. Dimensión Producto

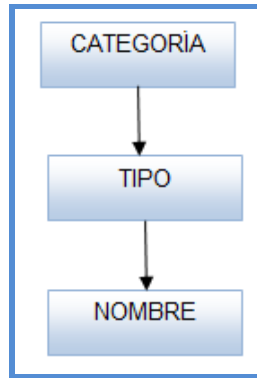


Figura 34: Jerarquía de los Atributos de la Dimensión Producto

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Categoría	Representa a que categoría pertenece un producto	No Actualizar	Ballerinas
Tipo	Representa a qué Tipo de producto pertenece un producto	No Actualizar	Taco x3
Nombre	Representa el nombre de un producto	No Actualizar	Vtx3-0589

Tabla 16: Detalle de Dimensión Producto

#### v. Dimensión Proveedor

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Proveedor	Representa el nombre del proveedor	No Actualizar	Importaciones Mercedes SAC

Tabla 17: Detalle de Dimensión Proveedor

#### vi. Dimensión Transferencia

Nombre del Atributo	Descripción del Atributo	Cambiando Política	Valores de Muestra
Motivo	Representa el motivo de la transferencia.	No Actualizar	Falla de fabrica

Tabla 18: Detalle de Dimensión Transferencia



### 2.5.3. Esquema Estrella

Sabiendo el número de Tablas de Hechos y las dimensiones asociadas a estas, orientamos nuestro Data Mart al Esquema Estrella para una mejor visualización del Modelo para el Diseño Físico.

Para una mejor comprensión, listamos primero los componentes que intervendrán el esquema, luego dividimos el Data Mart en dos gráficos: El esquema estrella de cada Tabla de Hechos y el Diseño Lógico del Data Mart.

#### ▪ Modelo Estrella componente

Componente: TABLA DE HECHOS VENTAS

Componente: TABLA DE HECHOS INVENTARIO

Componente: TABLA DE DIMENSIÓN TIEMPO

Componente: TABLA DE DIMENSIÓN VENDEDOR

Componente: TABLA DE DIMENSIÓN CLIENTE

Componente: TABLA DE DIMENSIÓN PROVEEDOR

Componente: TABLA DE DIMENSIÓN PRODUCTO

Componente: TABLA DE DIMENSIÓN TRANSFERENCIA

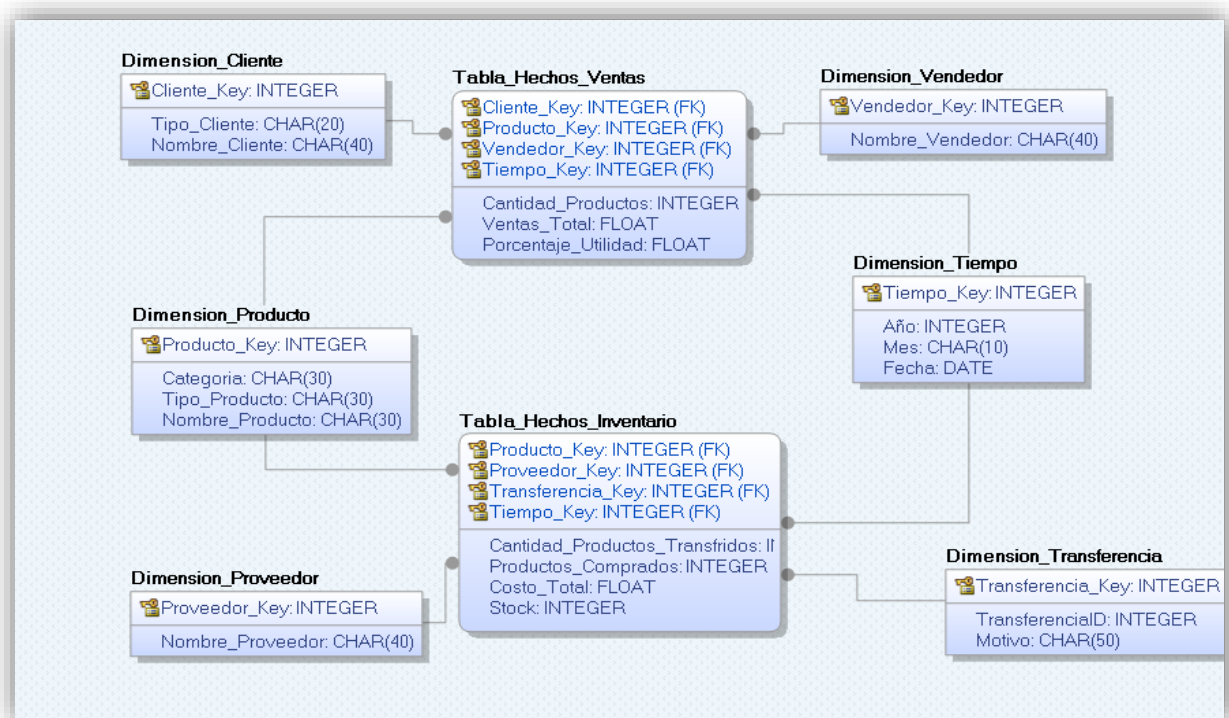


Figura 35: Esquema Estrella del Data Mart

## 2.6. Diseño Físico

Se tiene que tener en cuenta lo siguiente:

Modificar nombres a nombres estándar, si fuera necesario. Para nuestro Data Mart se realizaron los siguientes cambios:

DISEÑO LÓGICO	DISEÑO FÍSICO
Tabla de Hecho Ventas	Tabla_Hecho_Ventas
Tabla de Hecho Inventario	Tabla_Hecho_Inventario
Dimensión_Tiempo	Dim_Tiempo
Dimensión_Cliente	Dim_Cliente
Dimensión_Vendedor	Dim_Vendedor
Dimensión_Producto	Dim_Producto
Dimensión_Proveedor	Dim_Proveedor
Dimensión_Transferencia	Dim_Transferencia

Tabla 19: Nombres estándares para las Tablas Hechos y Dimensiones

NOMBRE DE LA TABLA	DISEÑO LÓGICO	DISEÑO FÍSICO
Dimensión_Tiempo	Año	Año
Dimensión_Tiempo	Mes	Mes
Dimensión_Proveedor	Nombre_Proveedor	Nombre_Proveedor
Dimensión_Vendedor	Nombre_Vendedor	Nombre_Vendedor
Dimensión_Cliente	Nombre_Cliente	Nombre_Cliente
Dimensión_Cliente	Tipo_Cliente	Tipo_Cliente
Dimensión_Producto	Nombre_Producto	Nombre_Producto
Dimensión_Producto	Tipo_Producto	Tipo_Producto
Dimensión_Producto	Categoría	Categoría
Dimensión_Transferencia	Motivo	Motivo

Tabla 20: Nombres estándares para los atributos de las dimensiones

NOMBRE DE LA TABLA	DISEÑO LÓGICO	DISEÑO FÍSICO
Tabla de Hecho Ventas	Cantidad_Productos	Cantidad_Productos
Tabla de Hecho Ventas	Venta_Total	Venta_Total
Tabla de Hecho Ventas	Costo_Total	Costo_Total
Tabla de Hecho Ventas	Porcentaje_Utilidad	Porcentaje_Utilidad
Tabla de Hecho Inventario	Cantidad_Productos_Transferidos	Cantidad_Productos_Transferidos
Tabla de Hecho Inventario	Cantidad_Productos_Comprados	Cantidad_Productos_Comprados

Tabla 21: Nombres estándares para los atributos de las Tablas de Hechos

Determinar el Tipo de Dato para cada Tabla que intervendrá en el Data Mart. Estos cuadros muestran el detalle de cada tabla:

**i. Tabla: Dim\_Tiempo**

NOMBRE DEL CAMPO	TIPO DE DATO
Mes	date
Trimestre	int
Día	date
Año	date

Tabla 22: Tipo de Dato para Dim\_Tiempo

**ii. Tabla: Dim\_Vendedor**

NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD
Nombre_Vendedor	varchar	50

Tabla 23: Tipo de Dato para Dim\_Vendedor

**iii. Tabla: Dim\_Cliente**

NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD
Nombre_Cliente	varchar	50
Tipo_Cliente	varchar	50

Tabla 24: Tipo de Dato para Dim\_Cliente

**iv. Tabla: Dim\_Proveedor**

NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD
Nombre_Proveedor	varchar	50
Identificacion	int	

Tabla 25: Tipo de Dato para Dim\_Proveedor

**v. Tabla: Dim\_Producto**

NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD
Código	int	
Categoría	varchar	50
Tipo_Cate	varchar	50
Nombre_Producto	Varchar	50

Tabla 26: Tipo de Dato para Dim\_Producto

**vi. Tabla: Dim\_Transferencia**

NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD
Motivo	varchar	50

Tabla 27: Tipo de Dato para Dim\_Transferencia

Determinar el tipo de Datos de las claves Primarias. Para nuestro Data Mart se generaran en forma automática y ordenada:

Tiempo\_Key  
 Proveedor\_Key  
 Vendedor\_Key  
 Cliente\_Key  
 Producto\_Key  
 Transferencia\_Key

} Tipo de Datos Entero (4)

Especificar las claves foráneas para cada Tabla de Hechos:

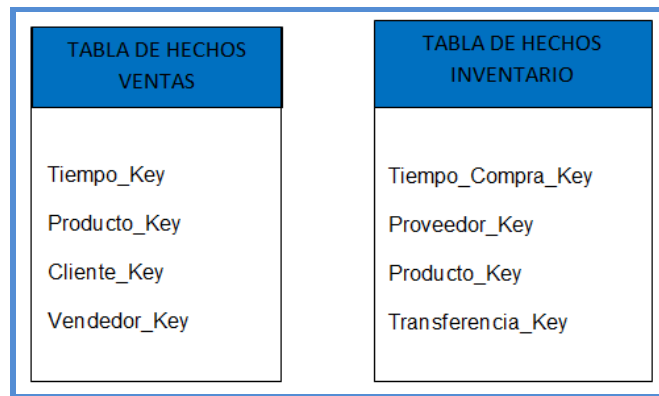


Figura 36: Claves Foráneas de las Tablas de Hechos

Por consiguiente, con los Datos del diseño Lógico y los cambios en la estructura física, se obtiene el siguiente Modelo de Base de Datos Física para el Data Mart.

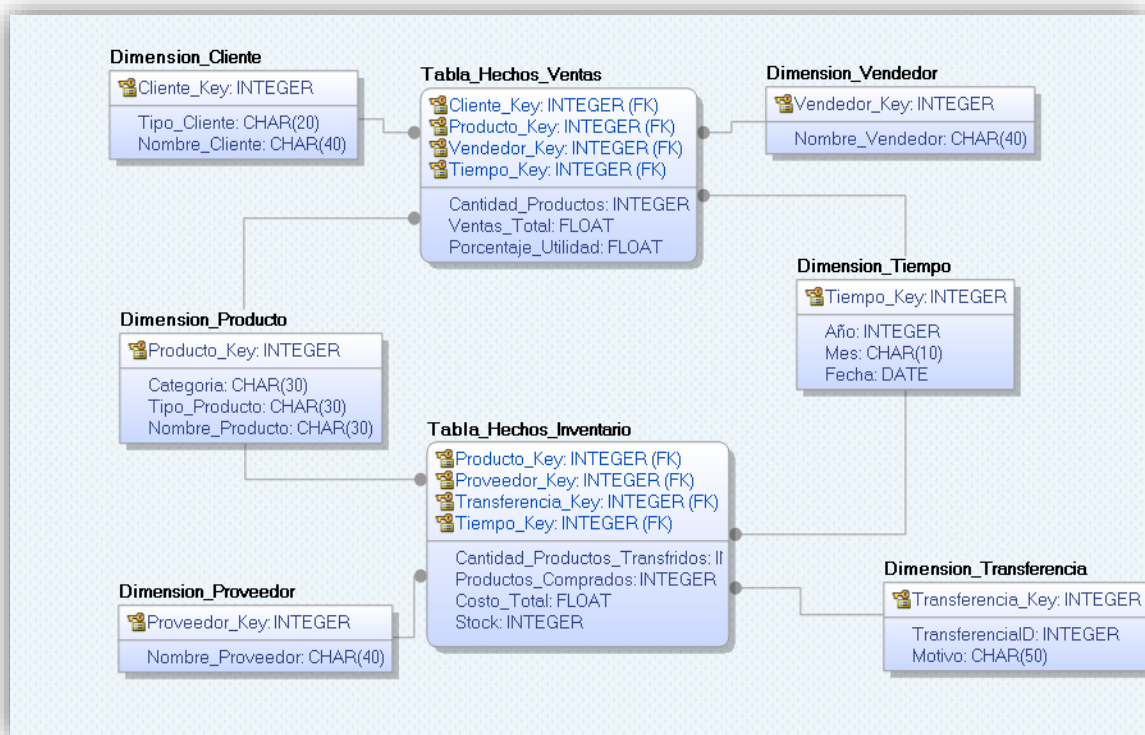


Figura 37: Diseño Físico de la Base de Datos del Data Mart

### 2.6.1. Determinación de las agregaciones

Determinamos las agregaciones por defecto que tendrán cada Hecho o medidas en las Tablas de Hechos. La mayoría de las reglas de agregación son sumas, como veremos a continuación:

Tabla de Hechos	Hecho	Regla de Agregación	Fórmula (SQL Server)
Tabla_Hecho_Ventas	Cantidad_Productos	Sum	Select Cantidad from detalle_comprobante
Tabla_Hecho_Ventas	Venta_Total	Sum	Select cantidad*precio from detalle_comprobante
Tabla_Hecho_Ventas	Costo_Total	Sum	Select cantidad*costo from detalle_comprobante
Tabla_Hecho_Ventas	Porcentaje_Utilidad	Sum	Cantidad*precio – Cantidad*costo
Tabla_Hecho_Inventario	Cantidad_Producto_Comprados	Sum	Select cantidad from detalle_compra
Tabla_Hecho_Inventario	Cantidad_Productos_Transferidos	Sum	Select cantidad from detalle_transf

Tabla 28: Determinación de las Agregaciones

### 2.6.2. Construcción de las Tablas y la Base de Datos en SQL

Después de haber realizado el Diseño Físico de las tablas pasamos a la etapa de construcción de dichas tablas utilizando el SQL Server 2010 R2.

Tabla 1: Dim\_Tiempo\_Ventas

Column Name	Data Type	Allow Nulls
Tiempo_Key	int	<input type="checkbox"/>
Id_Fecha	int	<input type="checkbox"/>
Mes	date	<input type="checkbox"/>
Trimestre	int	<input type="checkbox"/>
Dia	date	<input type="checkbox"/>
Año	date	<input type="checkbox"/>
		<input type="checkbox"/>

Figura 38: Tabla de la Dimensión Tiempo Ventas

**Tabla 2:Dim\_Tiempo\_Compras**

Column Name	Data Type	Allow Nulls
Tiempo_Key	int	<input type="checkbox"/>
Id_Fecha	int	<input type="checkbox"/>
Mes	date	<input type="checkbox"/>
Trimestre	int	<input type="checkbox"/>
Dia	date	<input type="checkbox"/>
Año	date	<input type="checkbox"/>
		<input type="checkbox"/>

Figura 39: Tabla de la Dimensión Tiempo Compras

**Tabla 3:Dim\_Proveedor**

Column Name	Data Type	Allow Nulls
Proveedor_Key	int	<input type="checkbox"/>
Id_Proveedor	int	<input type="checkbox"/>
Nombre_Proveedor	varchar(50)	<input type="checkbox"/>
Identificacion	int	<input type="checkbox"/>

Figura 40: Tabla de la Dimensión Proveedor

**Tabla 4:Dim\_Cliente**

Column Name	Data Type	Allow Nulls
Cliente_Key	int	<input type="checkbox"/>
Id_Cliente	int	<input type="checkbox"/>
Nombre_Cliente	varchar(50)	<input type="checkbox"/>
Tipo_Cliente	varchar(50)	<input checked="" type="checkbox"/>

Figura 41: Tabla de la Dimensión Cliente

**Tabla 5:Dim\_Vendedor**

Column Name	Data Type	Allow Nulls
Vendedor_Key	int	<input type="checkbox"/>
VendedorID	int	<input type="checkbox"/>
Nombre_Vendedor	varchar(50)	<input type="checkbox"/>

Figura 42: Tabla de la Dimensión Vendedor

**Tabla 6:Dim\_Transferencia**

JOSDENBR\SQLXP...m_Transferencia*			
	Column Name	Data Type	Allow Nulls
🔑	Transferencia_Key	int	<input type="checkbox"/>
	Id_Transferencia	int	<input type="checkbox"/>
	Motivo	varchar(50)	<input type="checkbox"/>

Figura 43: Tabla de la Dimensión Transferencia

**Tabla 7:Dim\_Producto**

JOSDENBR\SQLXP...o.Dim_Producto*			
	Column Name	Data Type	Allow Nulls
🔑	Producto_Key	int	<input type="checkbox"/>
	Producto_ID	int	<input type="checkbox"/>
	Codigo	int	<input type="checkbox"/>
	Categoria	varchar(50)	<input type="checkbox"/>
	Tipo_Cate	varchar(50)	<input type="checkbox"/>
	Nombre_Producto	varchar(50)	<input type="checkbox"/>

Figura 44: Tabla de la Dimensión Producto

Una vez construido todas las tablas para el Data Mart, continuamos con la construcción del Diagrama de la Base de Datos:

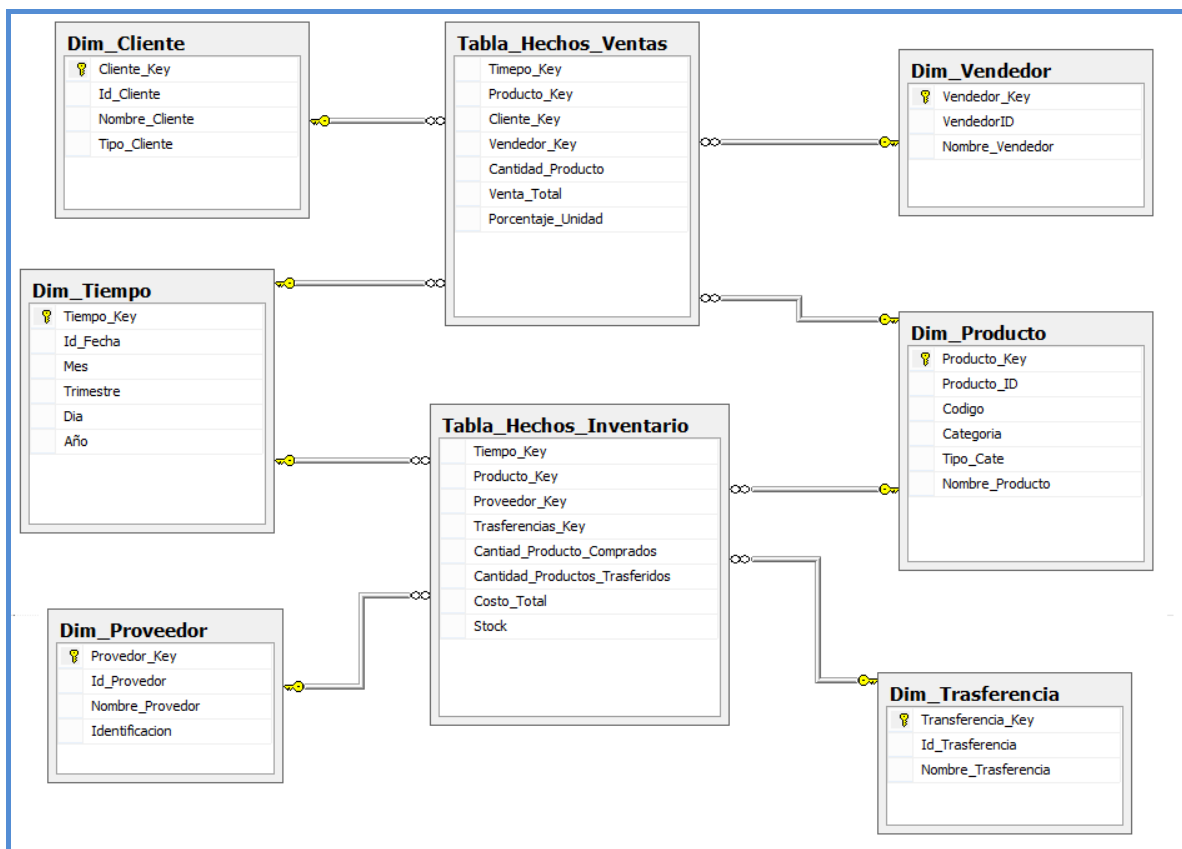


Figura 45: Data Mart para la Empresa Carubi SAC



## 2.7. Proceso de Extracción, Transformación y Carga de Datos

Para completar la construcción del Data Mart debemos poblar cada Tabla de la Base de Datos. Como ya conocemos la estructura de cada tabla, ahora debemos seguir los siguientes pasos para el Poblamiento:

- Definición de los Pasos de Transformación.
- Definición de los Workflows (Flujos de trabajo).
- Creación de los Paquetes de Servicio de Transformación de Datos (DTS).

### 2.7.1. Definición de los Pasos de Transformación

Para un correcto poblamiento de los datos debemos definir los pasos de Transformación, con la unidad de trabajo como parte de un proceso de Transformación. Para poblar el Data Mart se tiene que realizar los siguientes pasos:

**Paso 1:** Limpiando Dimensiones, consiste en eliminar los datos de todas las Dimensiones y Tablas de Hechos, paso que nos permite asegurar de que no pueda existir algún dato que se pudiera duplicar.

**Paso 2:** Poblamiento de la Dimensión Tiempo Ventas, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Tiempo Ventas.

**Paso 3:** Poblamiento de la Dimensión Tiempo Compras, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Tiempo Compras.

**Paso 4:** Poblamiento de la Dimensión Proveedor, consiste en mover los datos de la Tabla Proveedor para poblar la Dimensión Proveedor.

**Paso 5:** Poblamiento de la Dimensión Vendedor, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Vendedor.

**Paso 6:** Poblamiento de la Dimensión Cliente, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Cliente.

**Paso 7:** Poblamiento de la Dimensión Transferencia, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Transferencia.

**Paso 8:** Poblamiento de la Dimensión Producto, consiste en ejecutar una sentencia SQL, para transferir datos a la Dimensión Producto.

**Paso 9:** Poblamiento de las Tabla de Hechos.

### 2.7.2. Definición de los Workflows

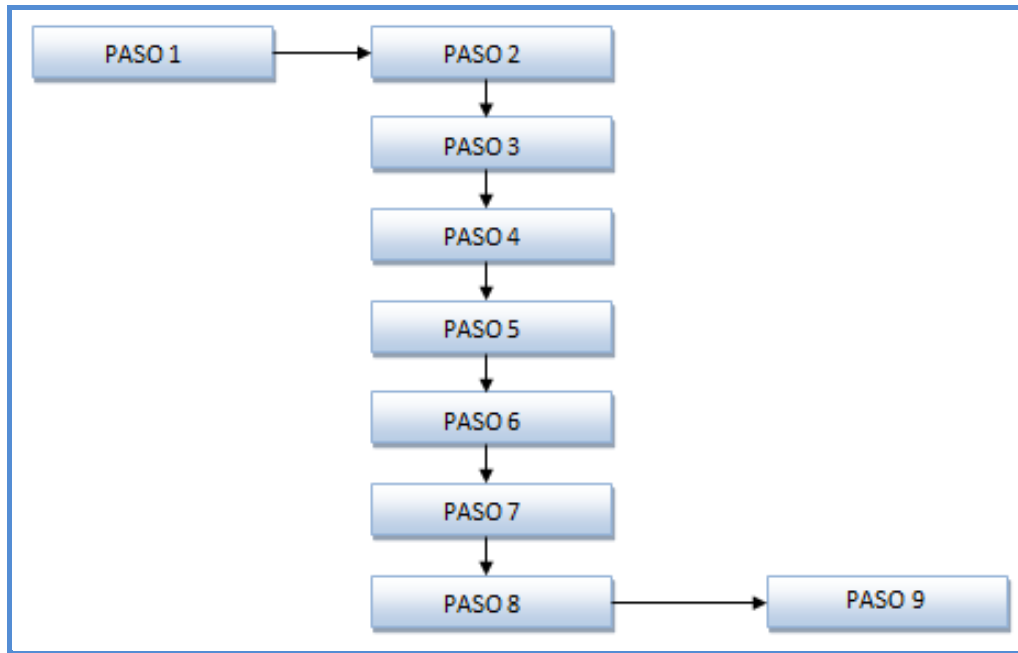


Figura 46: Workflow de los pasos de Transformación

Restricciones de Precedencia:

- La Limpieza de las Dimensiones (Paso 1) debe realizarse al inicio del proceso.
- El Poblamiento de la Dimensión Tiempo Ventas (Paso 2) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- El Poblamiento de la Dimensión Tiempo Compras (Paso 3) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- El Poblamiento de la Dimensión Proveedor (Paso 4) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- El Poblamiento de la Dimensión Vendedor (Paso 5) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- El Poblamiento de la Dimensión Cliente (Paso 6) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- El Poblamiento de la Dimensión Transferencia (Paso 7) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.

- h. El Poblamiento de la Dimensión Producto (Paso 8) debe realizarse sólo cuando se tenga la seguridad de que el paso 1 se ha ejecutado con éxito.
- i. El Poblamiento de las Tablas de Hechos (Ventas e Inventario) (Paso 9) debe realizarse sólo cuando se tenga la seguridad de que el paso 2,3,4,5,6,7,8 se hallan ejecutado con éxito.

Una vez conocido las relaciones de precedencias diagramaremos el Workflows que se necesitará realizar al construir el paquete de poblamiento del Data Mart, para el proceso de Ventas e Inventario de la empresa Carubi SAC.

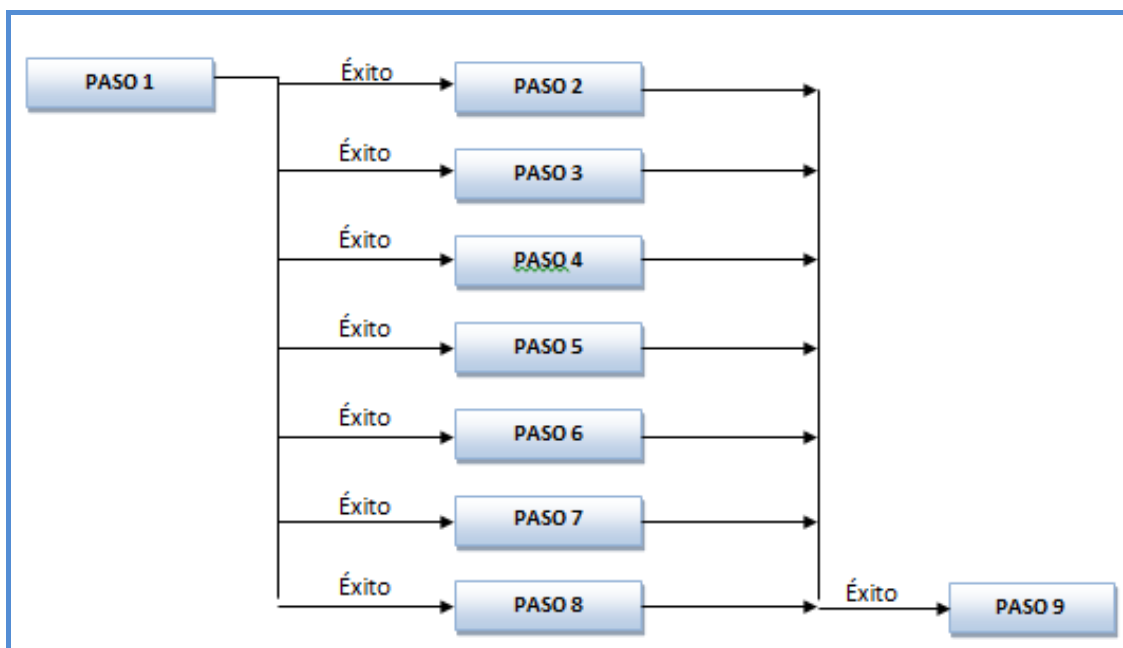


Figura 47: Diagrama Workflow con Restricciones de Precedencia

### 2.7.3. Creación de los Paquetes de Servicio de Transformación de Datos

Definidos los pasos de la transformación de datos y las restricciones de precedencia, podemos crear el paquete Servicio de Transformación de Datos (DTS); recordaremos que un DTS tiene como objetivos importar, exportar y realizar cambios en el formato de datos. Para su construcción utilizaremos el Servicio de Transformación de Datos SQL Server 2008 R2, aquí los datos pueden ser almacenado en varios formatos y en muchos lugares diferentes lo cual no es ningún problema.

Basándonos en Workflow con restricciones de la figura anterior, creamos el paquete DTS al cual llamamos “Poblamiento General Data Mart Carubi SAC.”, con los pasos previos definidos así como todas sus restricciones.

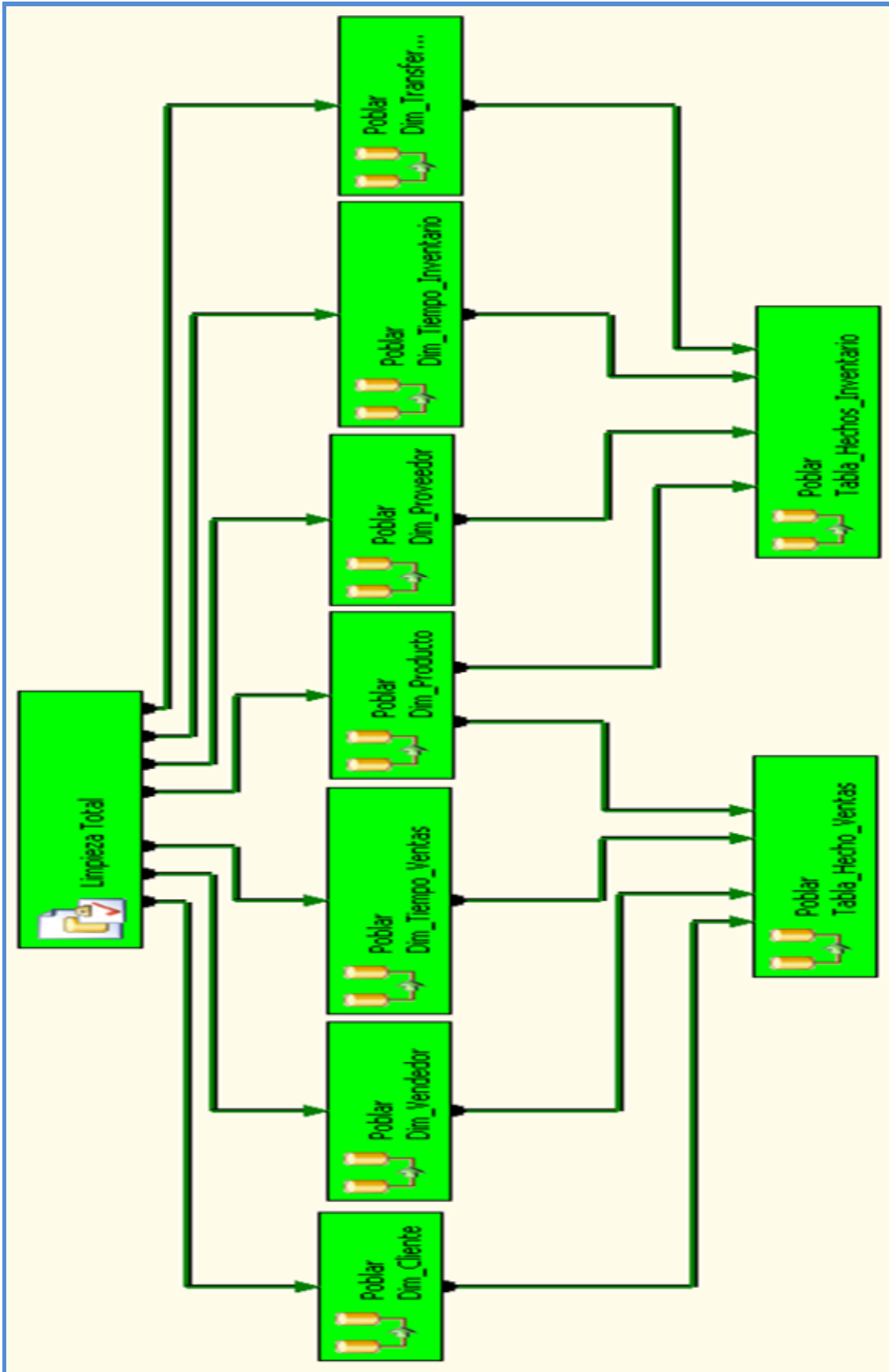


Figura 48: ETL del Data Mart

## SENTENCIAS DE CADA PASO DE WORKFLOW

### Paso 1: Limpiando Dimensiones

Añadimos una tarea al paquete al cual tendrá por nombre “Limpieza Total”, con la siguiente sentencia:

```
DELETE Tabla_Hechos_Ventas
DELETE Tabla_Hechos_Inventario
DELETE Dim_Producto
DBCC CHECKIDENT('Dim_Producto',reseed,0)
DELETE Dim_Tiempo_Ventas
DBCC CHECKIDENT('Dim_Tiempo_Ventas',reseed,0)
DELETE Dim_Tiempo_Inventario
DBCC CHECKIDENT('Dim_Tiempo_Inventario',reseed,0)
DELETE Dim_Proveedor
DBCC CHECKIDENT('Dim_Proveedor',reseed,0)
DELETE Dim_Vendedor
DBCC CHECKIDENT('Dim_Vendedor',reseed,0)
DELETE Dim_Cliente
DBCC CHECKIDENT('Dim_Cliente',reseed,0)
DELETE Dim_Transferencia
DBCC CHECKIDENT('Dim_Transferencia',reseed,0)
```

Esta sentencia nos permite limpiar los datos de todas las tablas de nuestro Data Mart. Esta limpieza asegura que no se dupliquen los datos.

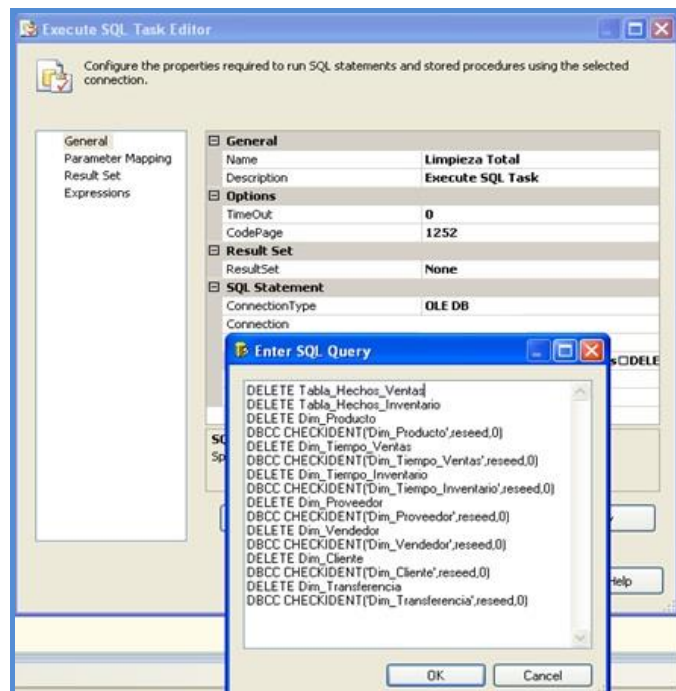


Figura 49: Ventana de Código para el Paso Limpieza Total

## Paso 2: Poblamiento de la Dimensión Tiempo Ventas

Para transferir los datos a la dimensión Tiempo Ventas, se efectúan los siguientes pasos:

- a. Se define una sentencia SQL como origen de datos de la Base de Datos Operacional.

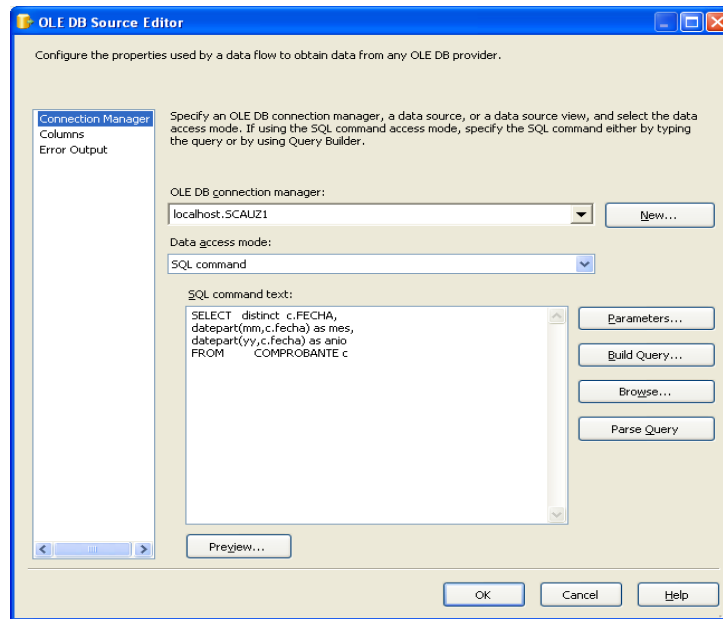


Figura 50: Ventana de Código para el Paso Limpieza Total

- b. Determinamos la Dimensión Tiempo Ventas como destino de Datos.

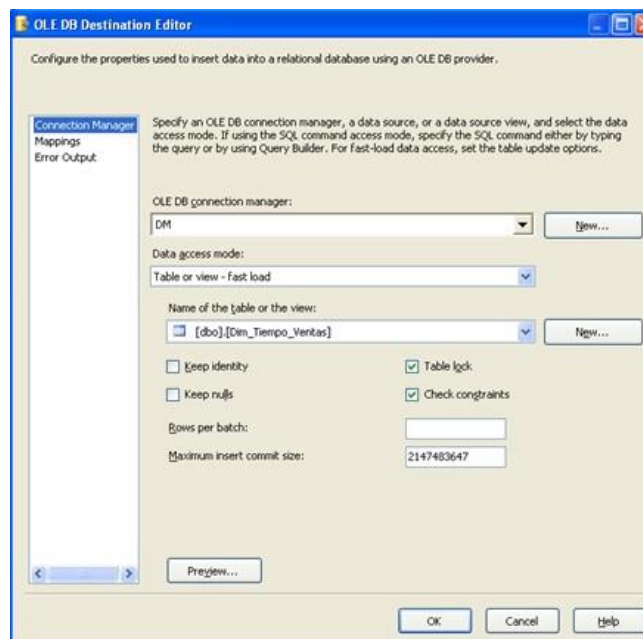


Figura 51: Conexión con Dim\_Tiempo

- c. En las transformaciones se determina que columnas del origen corresponderán con las del destino.

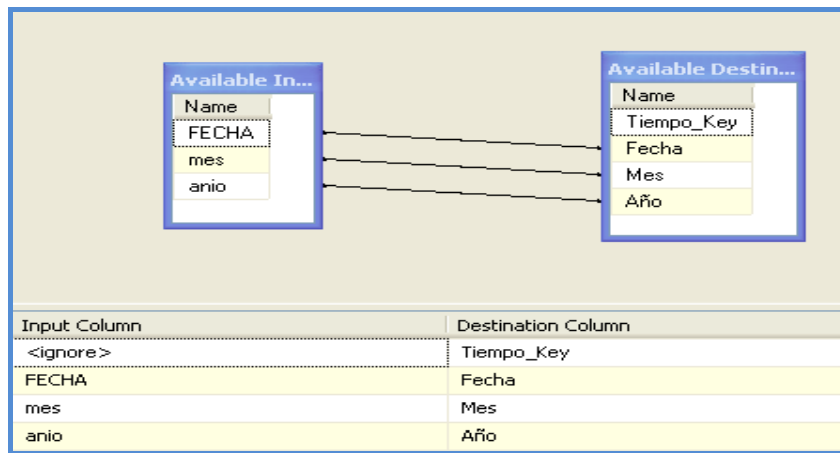


Figura 52: Poblamiento de la Dimensión Tiempo Ventas

### Paso 3: Poblamiento de la Dimensión Tiempo Compras

Para transferir los datos a la dimensión Tiempo Compras, se efectúan los siguientes pasos:

- a. Definimos una sentencia SQL como origen de datos.

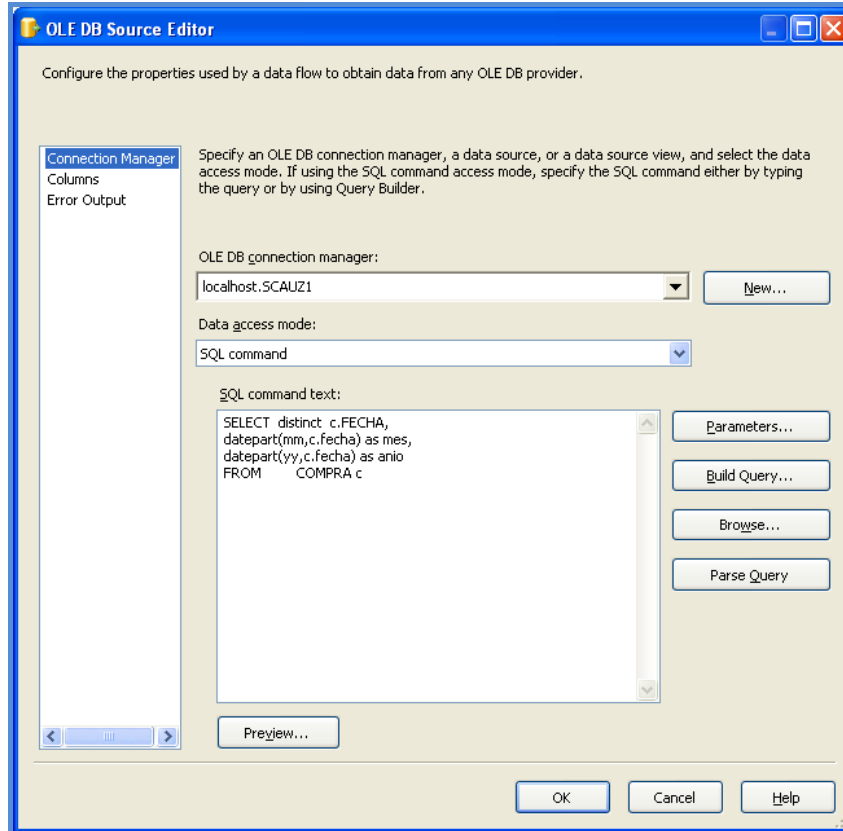


Figura 53: Conexión Dim\_Tiempo

- b. Determinamos la Dimensión Tiempo Compras como destino de datos.

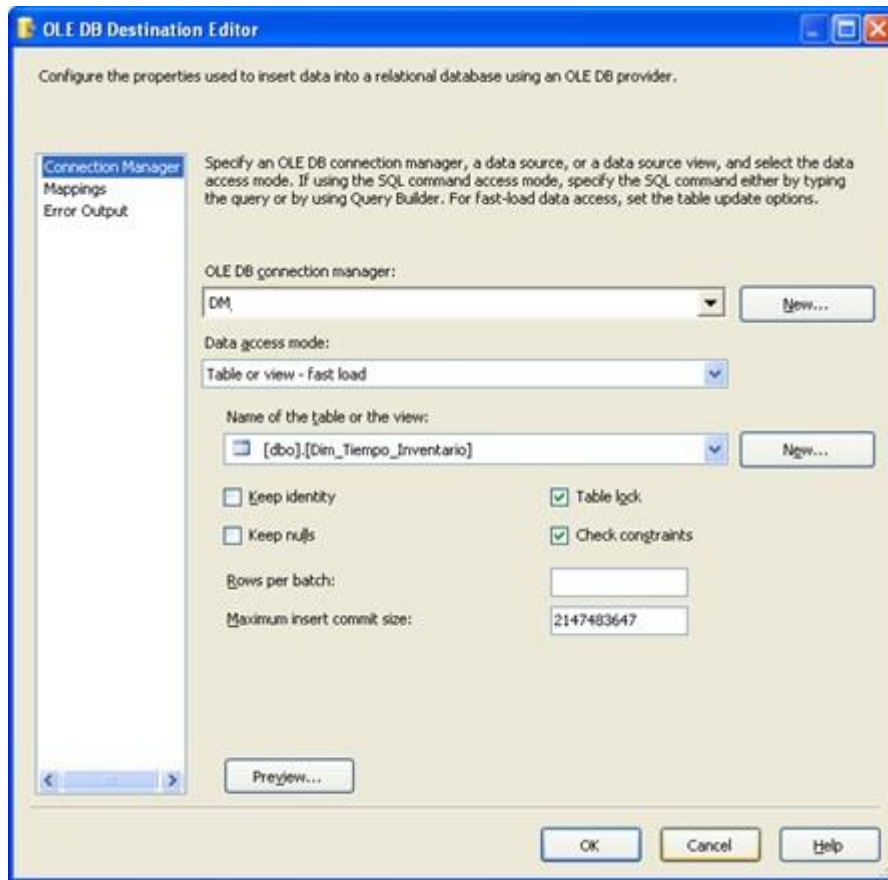


Figura 54: Conexión Tiempo

- c. En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Tiempo Compras

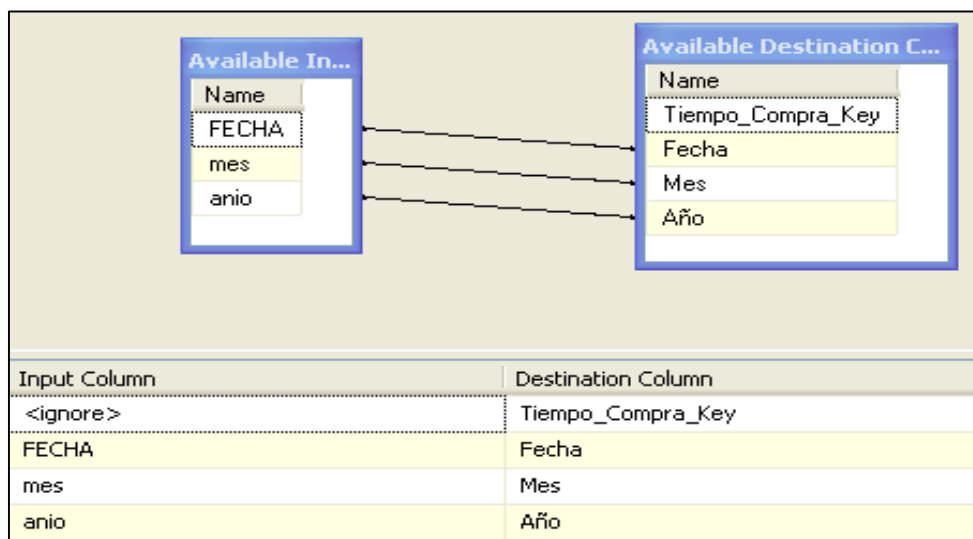


Figura 55: Poblamiento de la Dimensión Tiempo Compras



## Paso 4: Poblamiento de la Dimensión Proveedor

Para transferir los datos a la dimensión Proveedor, se efectúan los siguientes pasos:

- a. Definimos una sentencia SQL como origen de datos.

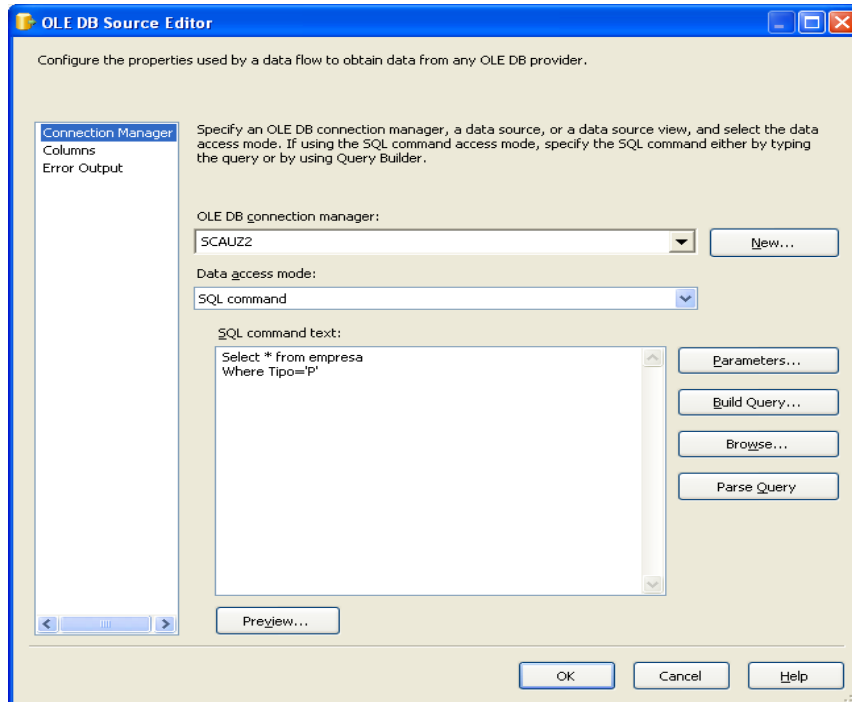


Figura 56: Conexión Dim\_Proveedor

- b. Determinamos la Dimensión Proveedor como destino de datos.

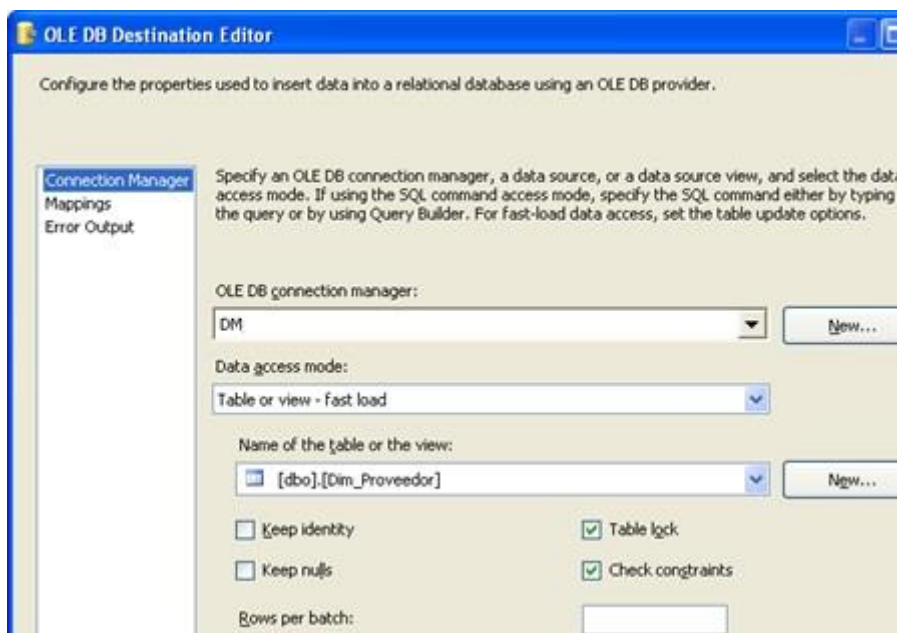


Figura 57: Conexión Dim\_Proveedor II

- c. En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Proveedor.

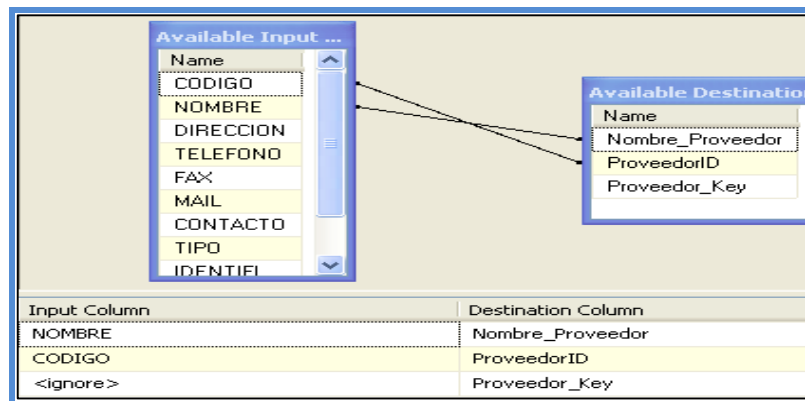


Figura 58: Poblamiento de la Dimensión Proveedor

### Paso 5: Poblamiento de la Dimensión Vendedor

Para transferir los datos a la dimensión Vendedor, se efectúan los siguientes pasos:

- a. Definimos la tabla Operador de la base de datos Operacional como origen de datos.

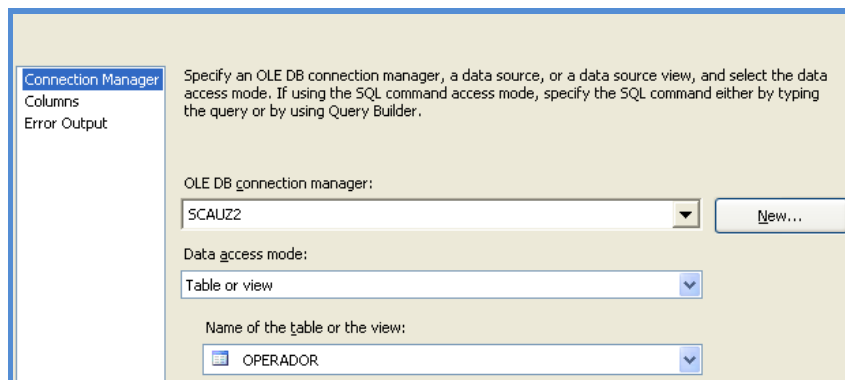


Figura 59: Conexión Dim\_Vendedor

- b. Determinamos la Dimensión Vendedor como destino de datos.

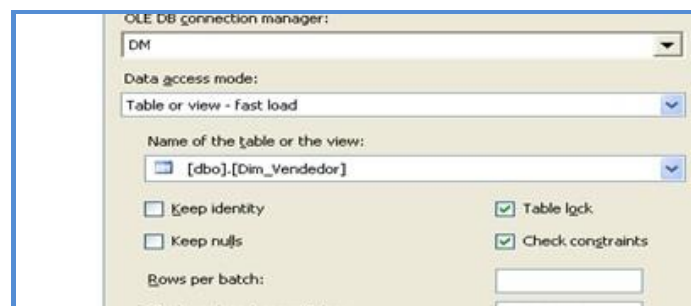


Figura 60: Conexión Dim\_Vendedor II

- c. En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Vendedor.

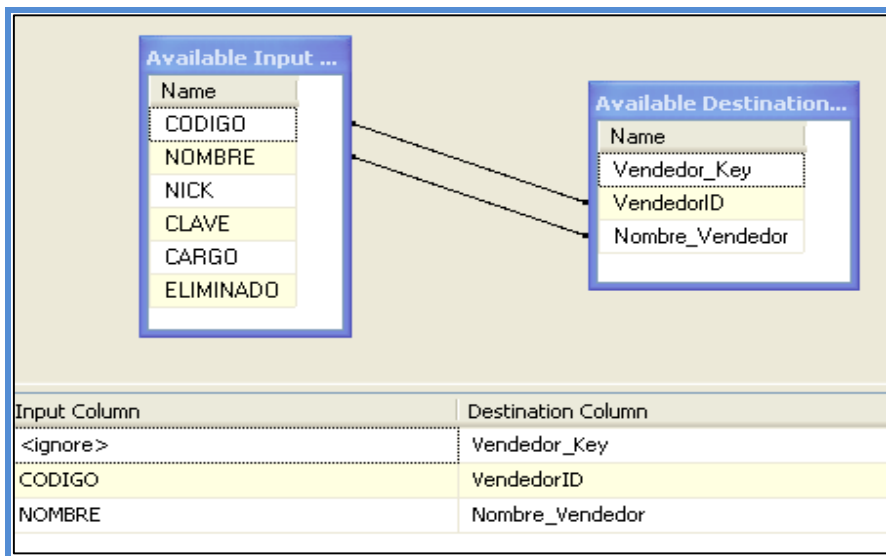


Figura 61: Poblamiento de la Dimensión Vendedor

### Paso 6: Poblamiento de la Dimensión Cliente

Para transferir los datos a la dimensión Cliente, se efectúan los siguientes pasos:

- a. Definimos a la tabla Cliente de la base de datos Operacional como origen de datos.

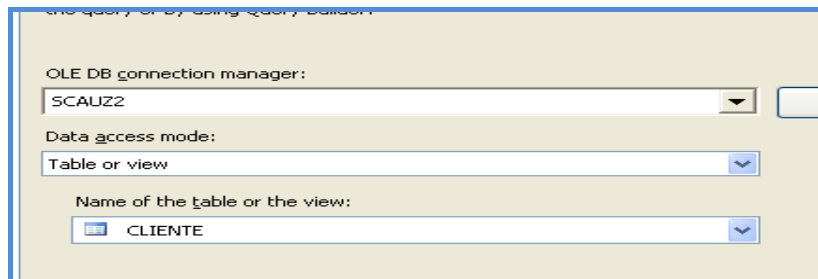


Figura 62: Conexión Dim\_Cliente

- b. Determinamos la Dimensión Cliente como destino de datos.

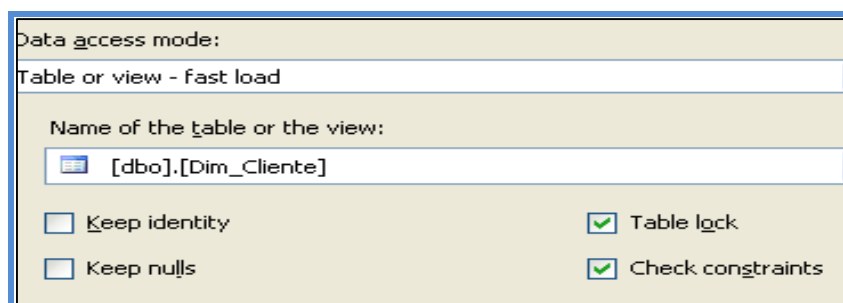


Figura 63: Conexión Dim\_Cliente II

- c. En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Cliente.

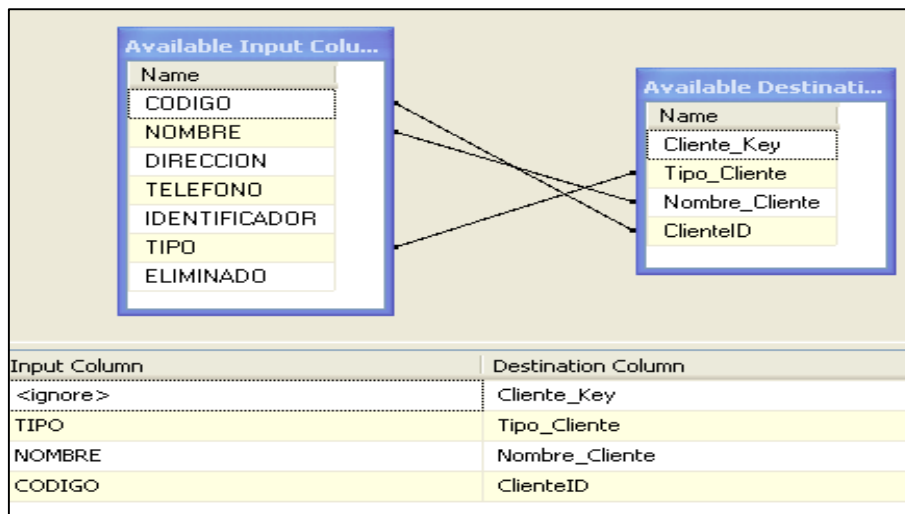


Figura 64: Poblamiento de la Dimensión Cliente

### Paso 7: Poblamiento de la Dimensión Transferencia

Para transferir los datos a la dimensión Transferencia, se efectúan los siguientes pasos:

- a. Definimos una sentencia SQL como origen de datos.

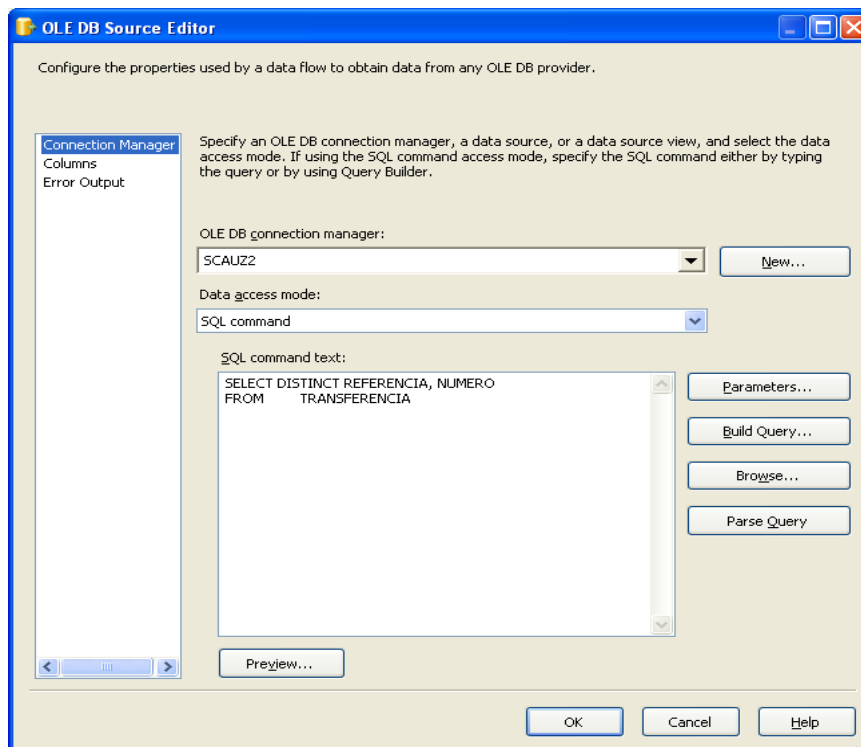
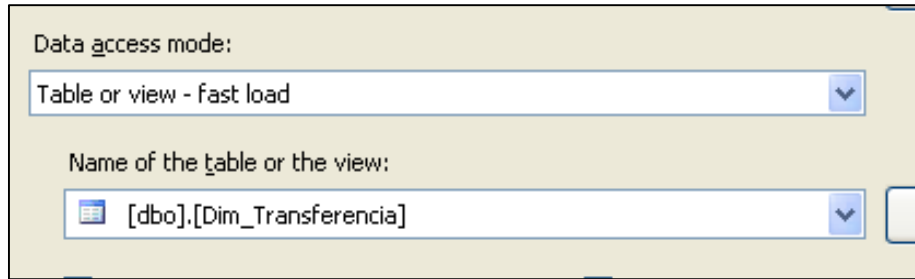


Figura 65: Conexión Dim\_Transferencia

- b. Determinamos la Dimensión Transferencia como destino de datos.

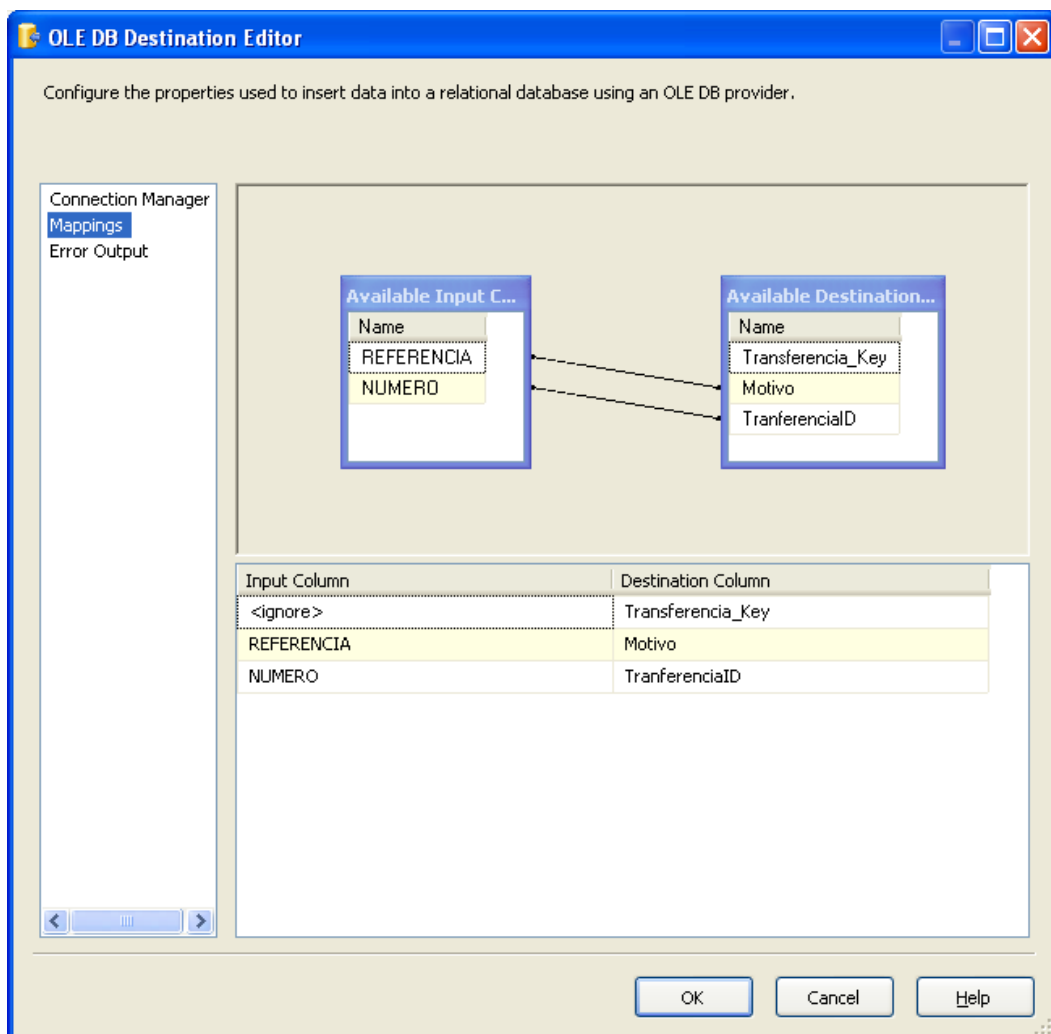


Data access mode:  
Table or view - fast load

Name of the table or the view:  
[dbo].[Dim\_Transferencia]

Figura 66: Conexión Dim\_Transferencia II

- c. En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Transferencia.



OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager  
Mappings  
Error Output

Available Input C...  
Name  
REFERENCIA  
NUMERO

Available Destination...  
Name  
Transferencia\_Key  
Motivo  
TranferenciaID

Input Column	Destination Column
<ignore>	Transferencia_Key
REFERENCIA	Motivo
NUMERO	TranferenciaID

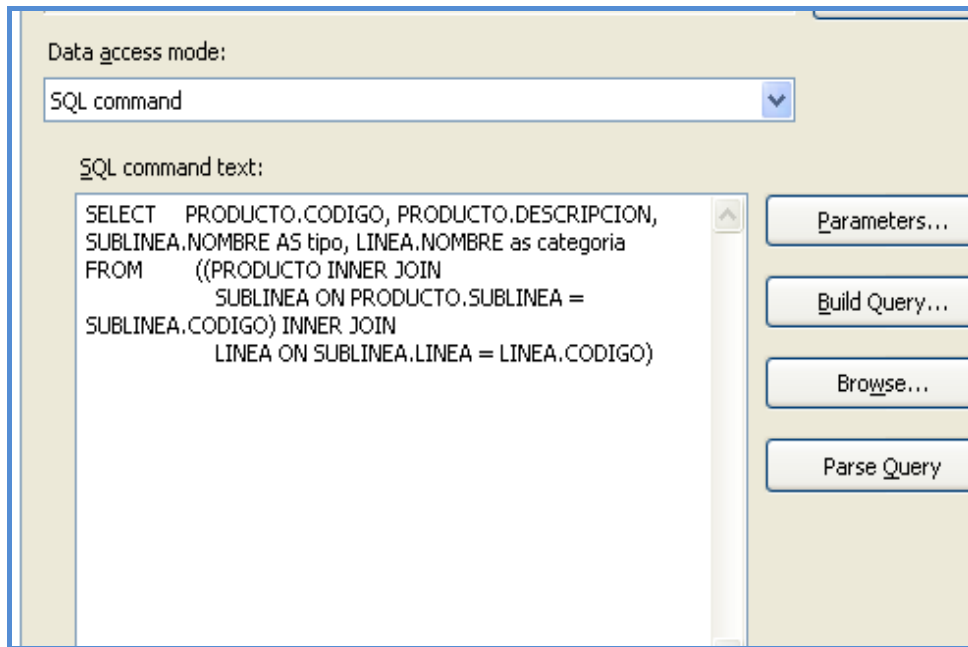
OK Cancel Help

Figura 67: Poblamiento de la Dimensión Transferencia

## Paso 8: Poblamiento de la Dimensión Producto

Para transferir los datos a la dimensión Producto, se efectúan los siguientes pasos:

- a. Definimos una sentencia SQL como origen de datos.



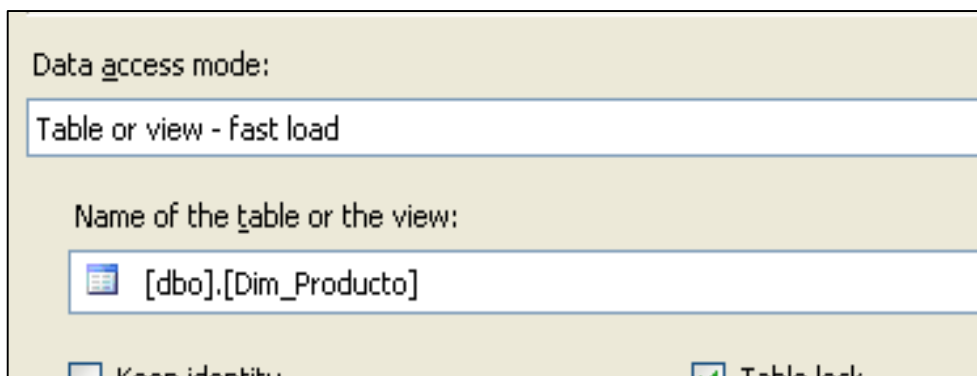
The screenshot shows a window titled "Data access mode:" with a dropdown menu set to "SQL command". Below this is a text area labeled "SQL command text:" containing the following SQL query:

```
SELECT PRODUCTO.CODIGO, PRODUCTO.DESCRIPCION,  
SUBLINEA.NOMBRE AS tipo, LINEA.NOMBRE as categoria  
FROM ((PRODUCTO INNER JOIN  
SUBLINEA ON PRODUCTO.SUBLINEA =  
SUBLINEA.CODIGO) INNER JOIN  
LINEA ON SUBLINEA.LINEA = LINEA.CODIGO)
```

To the right of the text area are four buttons: "Parameters...", "Build Query...", "Browse...", and "Parse Query".

Figura 68: Conexión Dim\_Producto

- b. Determinamos la Dimensión Producto como destino de datos.



The screenshot shows a window titled "Data access mode:" with a dropdown menu set to "Table or view - fast load". Below this is a text area labeled "Name of the table or the view:" containing the text "[dbo].[Dim\_Producto]". At the bottom, there are two checkboxes: "Keep identity" (unchecked) and "Table lock" (checked).

Figura 69: Conexión Dim\_Producto

- c. En las transformaciones se determina que columnas del origen corresponderán con las de la Dimensión Producto

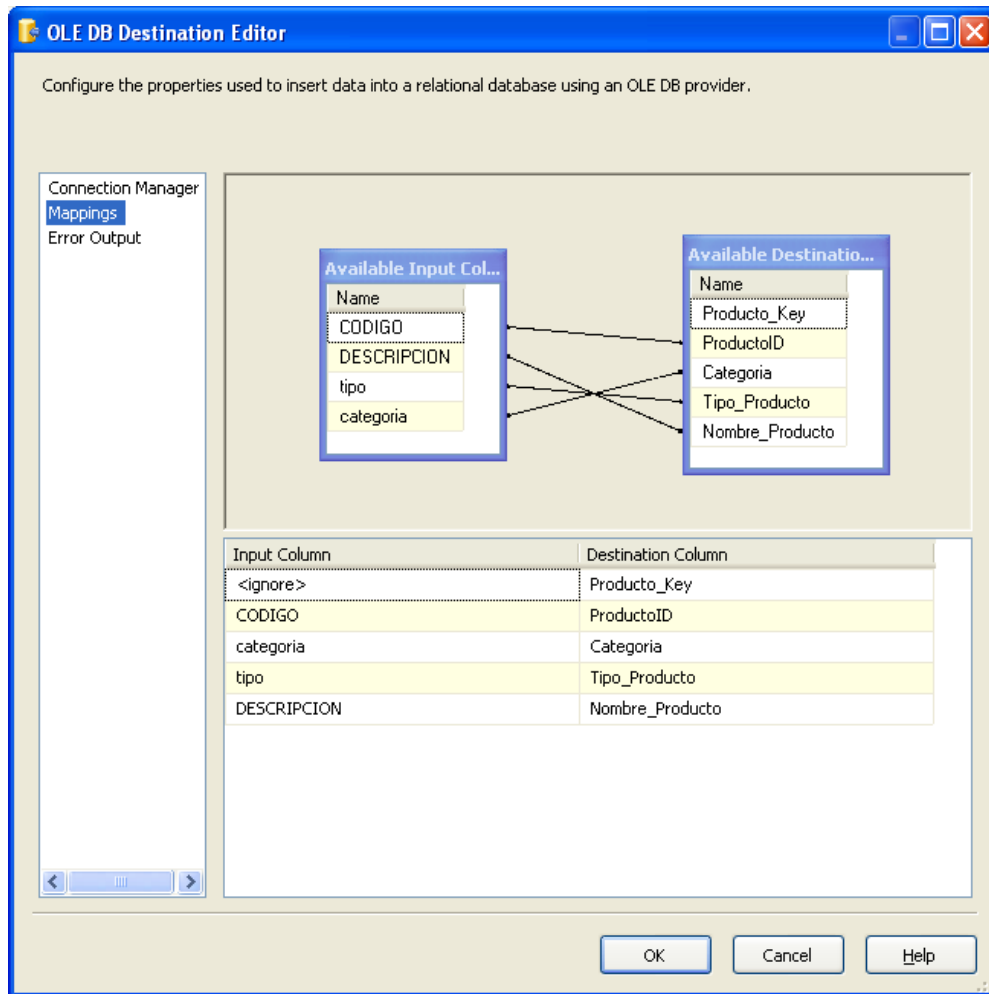


Figura 70: Poblamiento de la Dimensión Producto

## Paso 9: Poblamiento de las Tablas de Hechos

Para poblar la **TABLA DE HECHOS VENTAS**, se efectúan los siguientes pasos:

- a. Definimos una sentencia SQL como origen de datos.

```

SELECT distinct
dt.Tiempo_Key,
dp.Producto_Key,
dcl.Cliente_Key,
dv.Vendedor_Key,
dc.cantidad as Cantidad_Producto,
dc.PRECIO*dc.CANTIDAD as Venta_Total,
p.COSTO*dc.CANTIDAD as Costo_Total,
(dc.precio*dc.cantidad-p.costo*dc.cantidad)/100 as Porcentaje_utilidad

```

```

FROM Comprobante c
INNER JOIN Detalle_Comp dc ON c.Numero=dc.Comprobante
INNER JOIN PRODUCTO p ON dc.PRODUCTO=p.CODIGO
INNER JOIN DM.dbo.Dim_Cliente dcl ON c.cliente= dcl.ClienteID
INNER JOIN DM.dbo.Dim_Producto dp ON dc.Producto=dp.ProductoID
INNER JOIN DM.dbo.Dim_Vendedor dv ON c.Vendedor=dv.VendedorID
INNER JOIN DM.dbo.Dim_Tiempo_Ventas dt ON c.fecha=dt.fecha
WHERE c.fecha IS NOT NULL
group by dt.Tiempo_Key,dp.Producto_Key,
dcl.Cliente_Key,dv.Vendedor_Key,
dc.cantidad,dc.PRECIO*dc.CANTIDAD,
p.COSTO*dc.CANTIDAD,(dc.precio*dc.cantidad-p.cost*dc.cantidad)/100

```

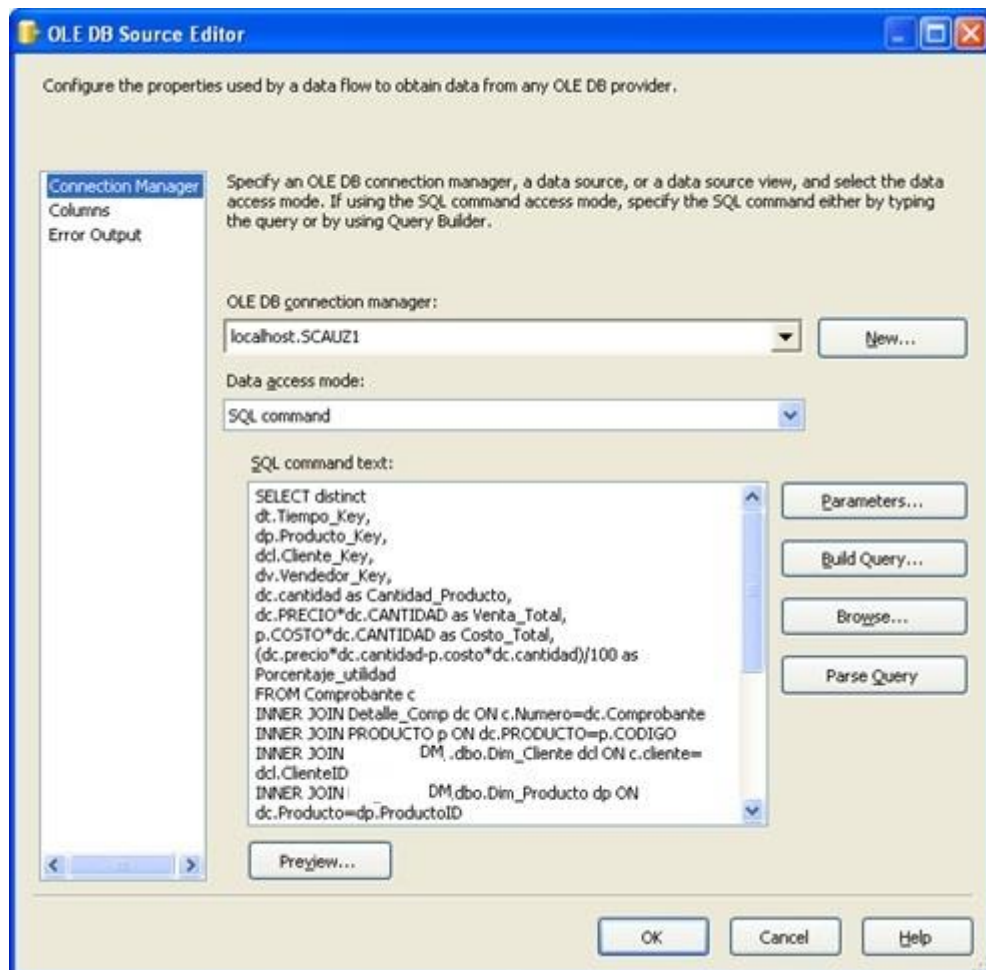
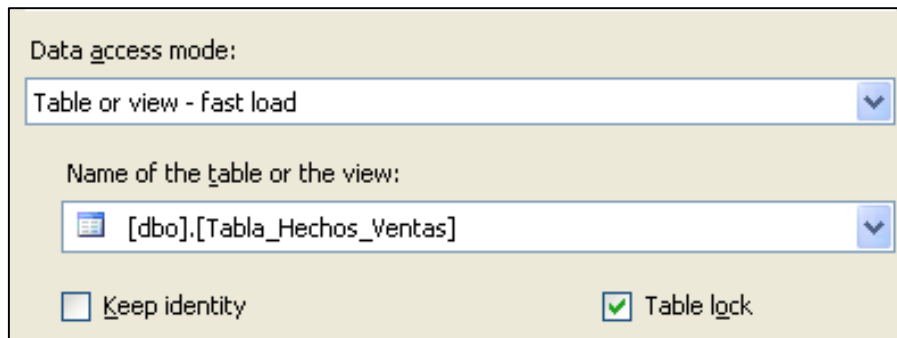


Figura 71: Conexión Hechos\_Venta



- b. Determinamos la Tabla de Hechos Ventas como destino de datos.



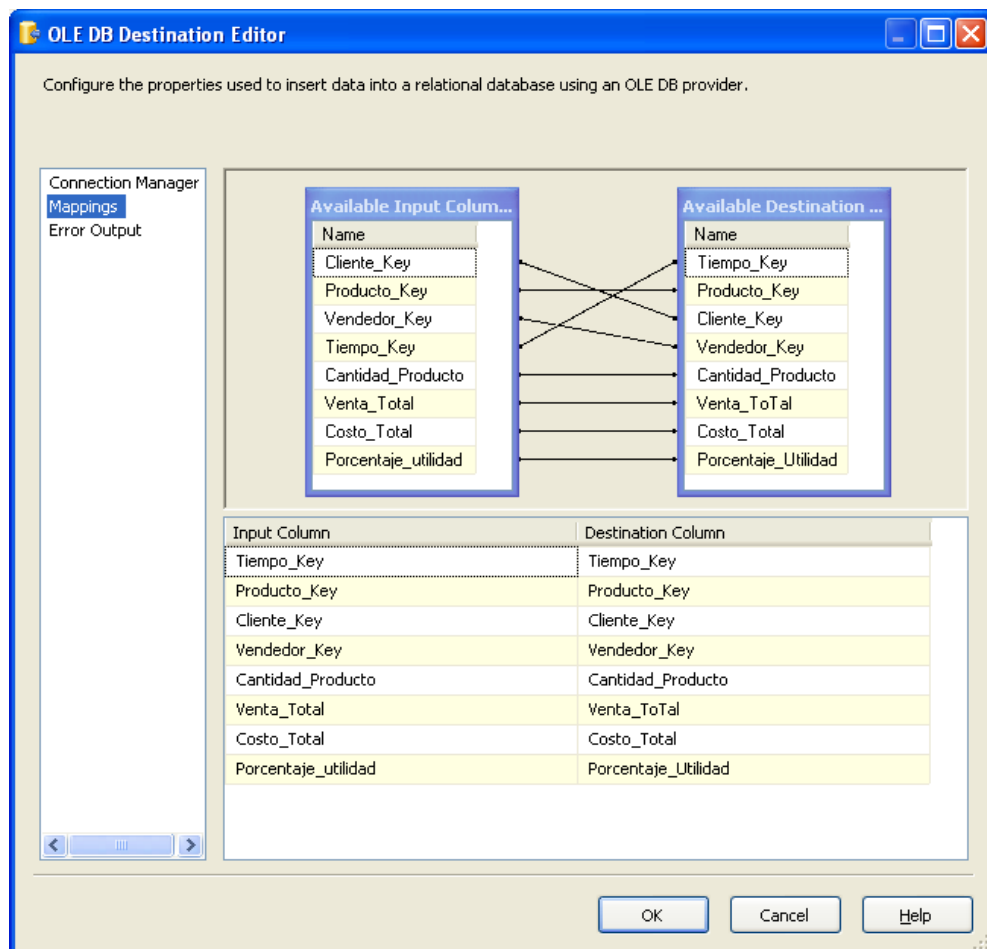
Data access mode:  
Table or view - fast load

Name of the table or the view:  
[dbo].[Tabla\_Hechos\_Ventas]

Keep identity  Table lock

Figura 72: Conexión Hechos\_Venta II

- c. En las transformaciones se determina que columnas del origen corresponderán con las de la Tabla de Hechos Ventas



OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager  
Mappings  
Error Output

Available Input Column...	Available Destination ...
Ciente_Key	Tiempo_Key
Producto_Key	Producto_Key
Vendedor_Key	Ciente_Key
Tiempo_Key	Vendedor_Key
Cantidad_Producto	Cantidad_Producto
Venta_Total	Venta_ToTal
Costo_Total	Costo_Total
Porcentaje_utilidad	Porcentaje_Utilidad

Input Column	Destination Column
Tiempo_Key	Tiempo_Key
Producto_Key	Producto_Key
Ciente_Key	Ciente_Key
Vendedor_Key	Vendedor_Key
Cantidad_Producto	Cantidad_Producto
Venta_Total	Venta_ToTal
Costo_Total	Costo_Total
Porcentaje_utilidad	Porcentaje_Utilidad

OK Cancel Help

Figura 73: Poblamiento de la Tabla de Hechos Ventas

Para poblar la **TABLA DE HECHOS INVENTARIO**, se efectúan los siguientes pasos:

- a. Definimos una sentencia SQL como origen de datos.

```
SELECT distinct(dti.Tiempo_Compra_Key),dp.Producto_Key,  
dpre.proveedor_Key,dtt.Transferencia_Key,  
dc.cantidad as Cantidad_Producto_Comprados,  
dt.ITEM as Cantidad_productos_Transferidos  
FROM Compra c  
INNER JOIN Detalle_Compra dc ON c.CORRELATIVO=dc.COMPRA  
INNER JOIN PRODUCTO p ON dc.PRODUCTO=p.CODIGO  
INNER JOIN DM.dbo.Dim_Proveedordpre ON c.PROVEEDOR=  
dpre.ProveedorID  
INNER JOIN DM.dbo.Dim_Productodp ON dc.Producto=dp.ProductoID  
INNER JOIN DETALLE_TRANS dt ON p.CODIGO=dt.PRODUCTO  
INNER JOIN Transferencia t ON t.NUMERO=dt.TRANSFERENCIA  
INNER JOIN DM.dbo.Dim_Tiempo_Inventariodti ON c.fecha=dti.fecha  
INNER JOIN DM.dbo.Dim_Transferenciadt ON t.NUMERO=dtt.TranferenciaID  
WHERE c.fecha IS NOT NULL  
group by dti.Tiempo_Compra_Key,dp.Producto_Key,  
dpre.Proveedor_Key,dtt.Transferencia_Key,  
dc.cantidad,dt.item
```

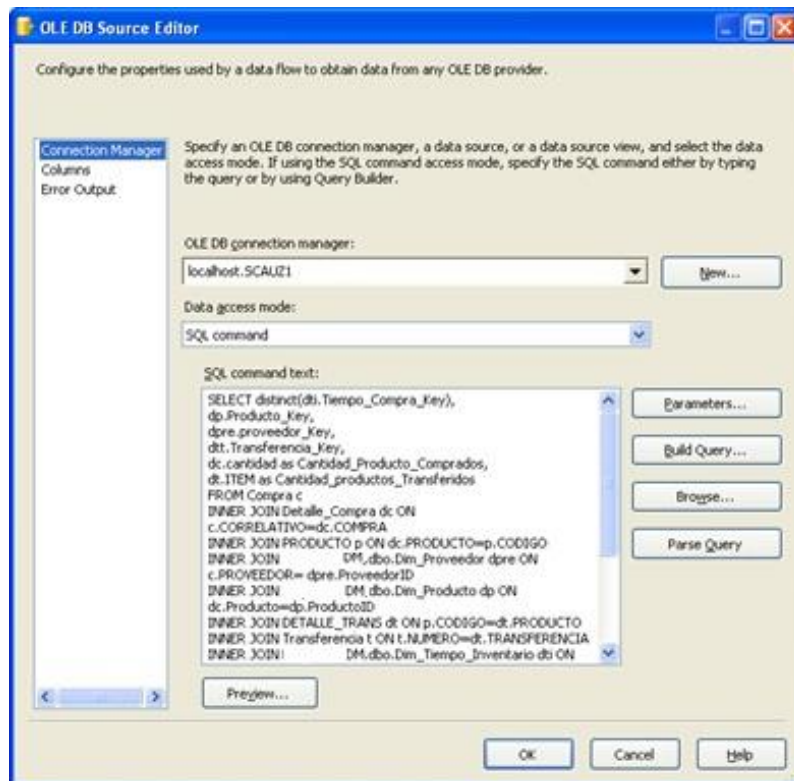


Figura 74: Conexión Hechos\_Inventario

b. Determinamos la Tabla de Hechos Inventario como destino de datos.

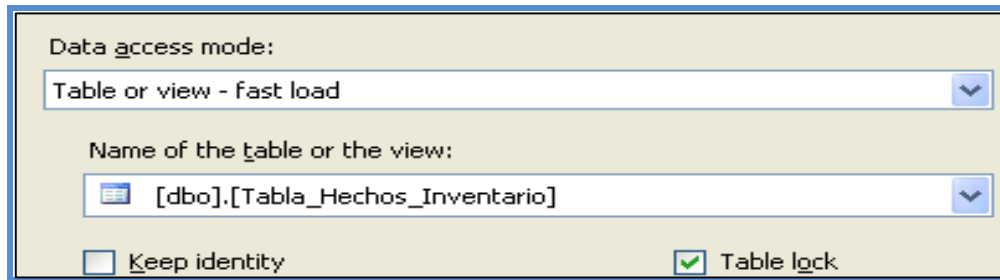


Figura 75: Conexión Hechos\_Inventario II

c. En las transformaciones se determina que columnas del origen corresponderán con las de la Tabla de Hechos Inventario

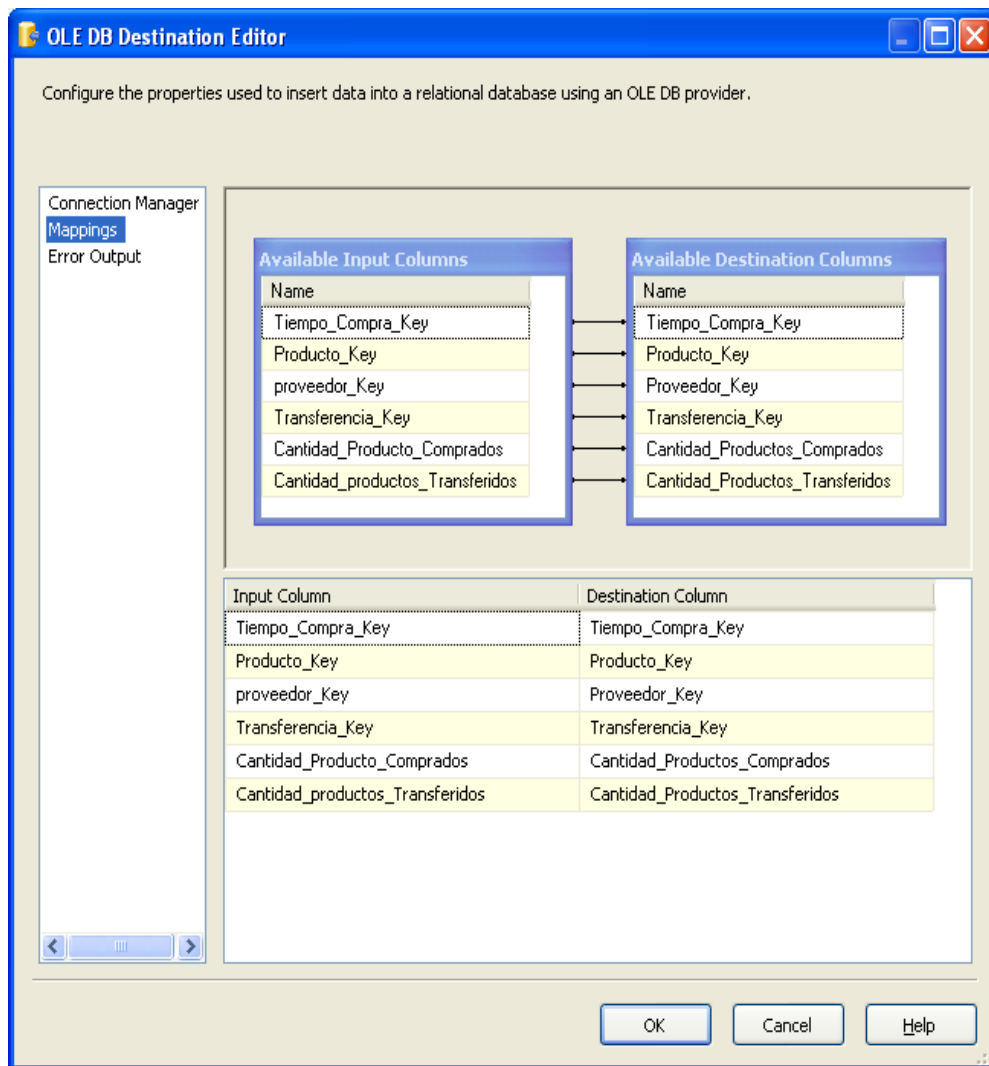


Figura 76: Poblamiento de la Tabla de Hechos Inventario

## 2.8. Selección de Productos

### 2.8.1. Hardware

La base de datos y los servidores OLAP deben de tener la siguiente configuración de hardware:

- RAM: 2GB
- Disco Duro (espacio): 80GB.
- Procesador: Corel 2 Duo Intel Xeon de 2.53 GHz,

### 2.8.2. Software

Para la implementación se trabajará con, MS SQL Server 2008 R2. En las siguientes tablas se detallará la herramienta a utilizar y los procesos a ejecutar.

Componente del Proceso	Herramienta
<b>Construcción de Interfaces</b>	MS Excel 2010
<b>Cubos</b>	AnalysServices
<b>ETL</b>	SQL Server Business Intelligence
<b>DataMart</b>	SQL Server 2008 R2
<b>Modelo Lógico y Físico</b>	MS Visio 2010

Tabla 29: Software para el proceso de Diseño del Data Mart

## 2.9. Especificación de la Aplicación del Usuario Final

### 2.9.1. Estructura de Cubo

La herramienta utilizada para el diseño del cubo es el SSAS (SQL Server Analysis Services) que es el diseñador de cubos de Business Intelligence Development Studio para crear el cubo, incluido las medidas, las dimensiones y sus respectivas relaciones.

- a. El cubo de Ventas se llama “CUBO\_VENTAS”. Las medidas y dimensiones de este cubo son originadas por su equivalente a las tablas del modelo dimensional como se muestra a continuación.

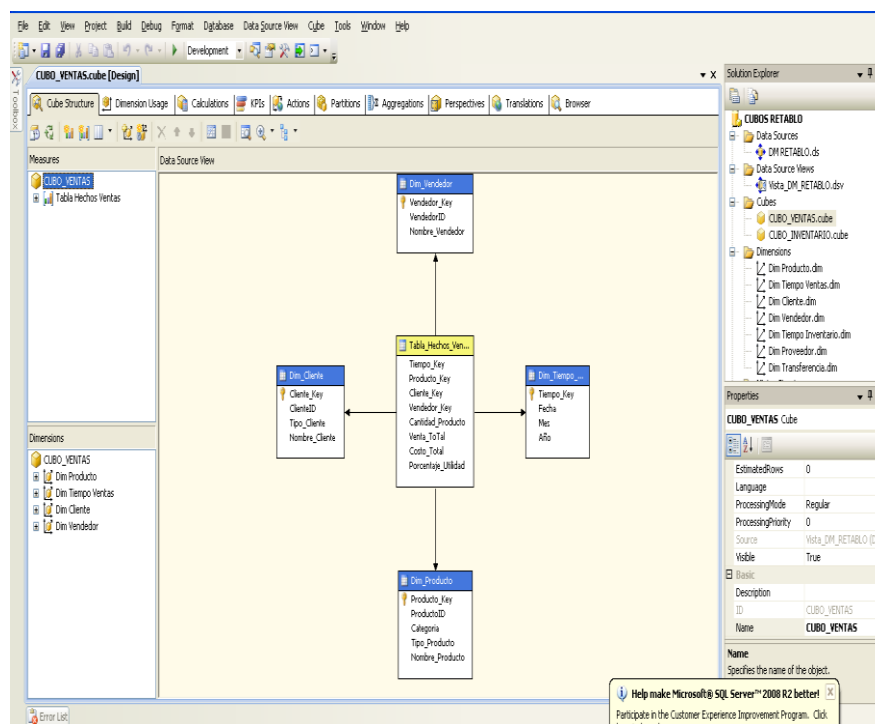


Figura 77: Diseño de Cubo de Ventas

En la Figura N° 78 se puede observar como el diseño del cubo cumple con los requerimientos básicos a descritos en “Análisis de Requerimientos”, donde podemos observar como los datos (las medidas) son mostrados.

		Nombre Producto						
		1: Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantidad
Año	Mes	Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantidad Producto	Cantida
2005	1		11	12	1			
	10		10	20		3		
	11							2
	12						4	
	2		64	14				
	3							
	4		2					
	5	3		21				
	6				4			
7				8		2		
8	6							
9	2							
Total		11	87	75	5	5	4	2
2006		12	10	72	2		2	1
2007		6		24			3	8
2008								9
Grand Total		29	97	171	7	5	9	20

Figura 78: Medidas y Dimensiones del Cubo Ventas

- b. El cubo de Inventario se llama “CUBO\_INVENTARIO”. Las medidas y dimensiones de este cubo son originadas por su equivalente a las tablas del modelo dimensional como se muestra en la Figura N° 79.

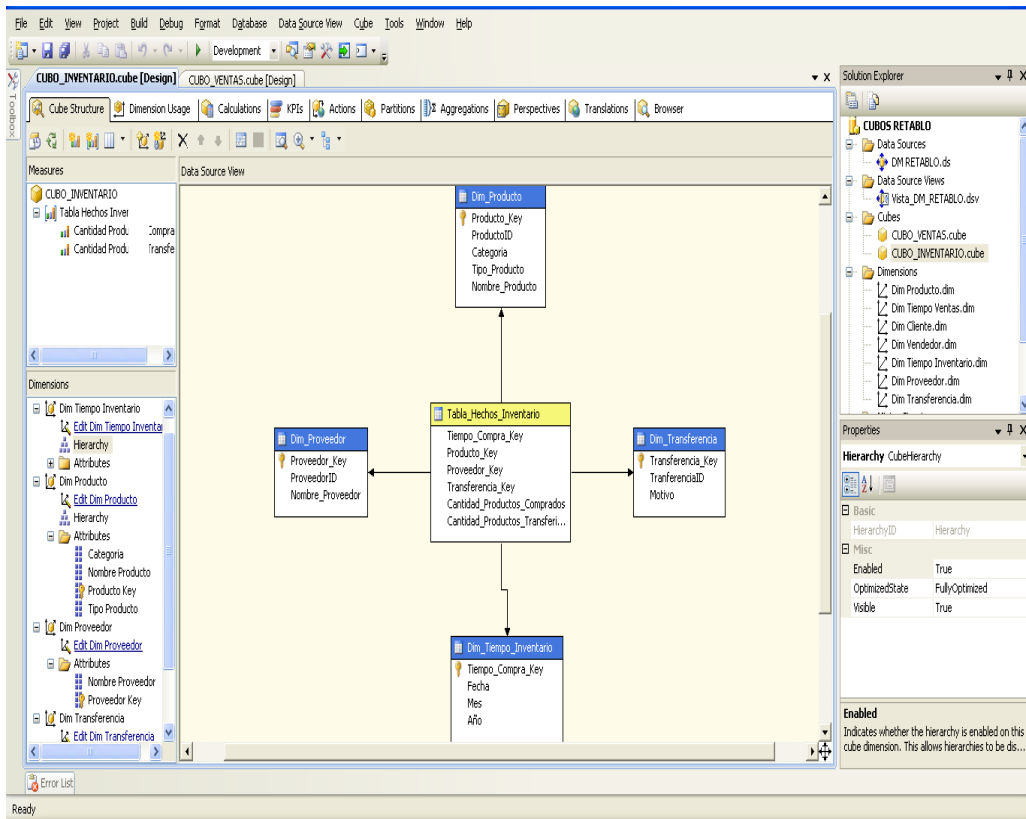


Figura 79: Diseño de Cubo de Inventarios

Y en la Figuras N° 80 se puede observar como el diseño del cubo cumple con los requerimientos básicos a descritos en “Análisis de Requerimientos”, donde podemos observar como los datos (las medidas) son mostrados.

Dimension	Hierarchy	Operator	Filter Expression				
<Select dimension>							
Nombre Proveedor	All						
Nombre Producto	All						
	XLITICO	1000UF 100V (6.90)	1000UF 100V (6.90)				
		1000UF 25V	100PF 50V (101)				
			100UF 20V				
Año	Mes	Comprados	Cantidad Productos Comprados	Cantidad Productos Comprados	Cantidad Productos Comprados	Cantidad Productos Comprados	Cantidad
2004							
2005	1			30			33
	2						132
	3			20			
	4	10	40				
	5						66
	6			40			
	7		48		525		
	Total	10	88	90	525		231
2007							
Grand Total		10	88	90	525		231

Figura 80: Medidas y Dimensiones del Cubo Inventario

## 2.10. Desarrollo de la Aplicación del Usuario Final

El desarrollo de las aplicaciones de los usuarios finales involucra configuraciones de metadata y construcción de reportes específicos. Las Figuras que se van a mostrar a continuación son interfaces personalizadas de cada Requerimiento creadas en Excel con herramientas comunes, utilizando los cubos Ventas e Inventario anteriormente descritos.

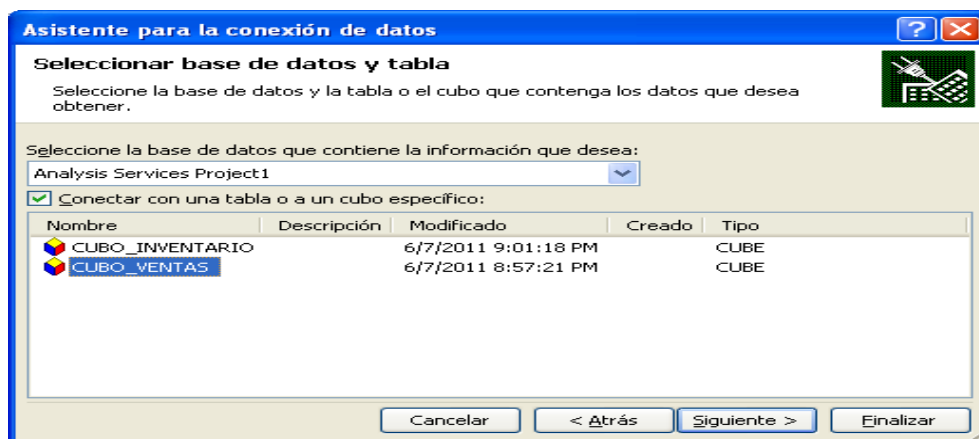


Figura 81: Diseño de interfaces

## 2.10.1. Estructura de Cubo

R1 ¿Cuál es la cantidad total de ventas en soles de los productos que se importan y exportan en un determinado mes?

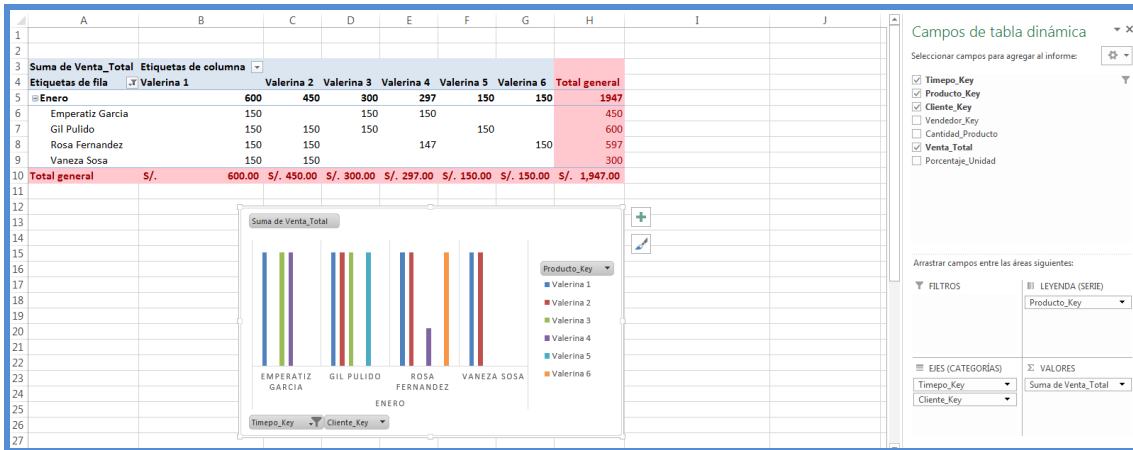


Figura 82: Reporte de Requerimiento 1

R8 ¿Cuál es la cantidad productos que se compraron en un mes determinado?

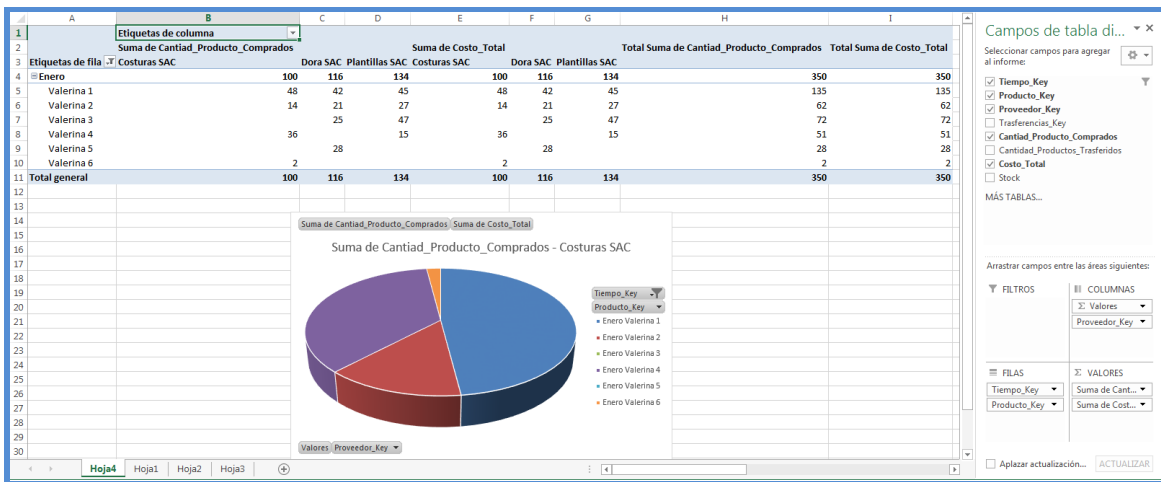


Figura 83: Reporte de Requerimiento 8



## CAPITULO III

### DISCUSIÓN

Para la contrastación de la hipótesis se ha considerado lo siguiente:

#### 3.1. Formulación del Problema

¿Con la implementación de un Sistema de Inteligencia de Negocio permitirá dar un mejor soporte a la toma de decisiones para las empresas de servicio de importación y exportación del departamento la Libertad?

#### 3.2. Hipótesis

La implementación de una Solución de Inteligencia de Negocios que mejorará el proceso de Toma de Decisiones para la empresa importadora y exportadora de calzado. Luego se definen las variables que intervienen en la veracidad o falsedad de la hipótesis:

- Variable Independiente: Implementar una solución de Inteligencia de Negocio
- Variable Dependiente: Mejora en la toma de decisiones para empresas de importación y exportación de calzado

##### 3.2.1. Manera Presencial

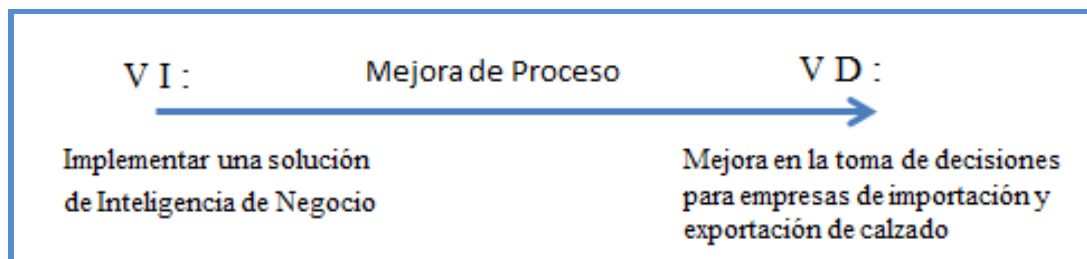


Figura 84: Definición de Variable

Variable	Tipo	Fuente	Dimensión	Escala de Medición	Instrumento
Implementar una solución de Inteligencia de Negocio	Cuantitativa	Investigadores	Estrategias	Nominal	Kimball, R. y Ross
	Cuantitativa	Investigadores	Tecnología unificada	Nominal	Ms SQL Server 2008 R2
Mejora en la toma de decisiones para empresas de importación y exportación de calzado	Cuantitativa	Gerentes	Reportes	Ordinal	Ms Excel
	Cuantitativa	Gerentes	Portabilidad	Ordinal	Internet

Tabla 30: Descripción de Variable

### 3.2.2. Cálculo de los Indicadores de la Hipótesis

Para el cálculo de los indicadores de la hipótesis en el Data Mart Propuesto (DMP) y el Sistema Actual (SA), se realizó un cuestionario (Ver Anexo B) donde se evaluó a los usuarios luego de haber interactuado con el Data Mart.

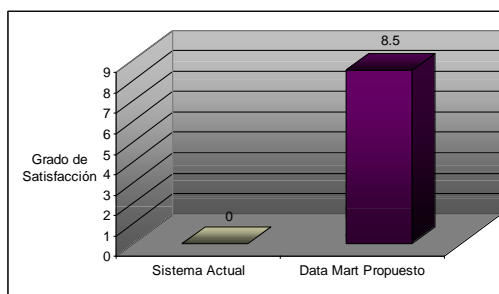
Los valores que los usuarios dieron a las respuestas del cuestionario fueron aplicados según el rango de satisfacción que muestran en la siguiente tabla:

RANGO	GRADO DE SATISFACION
0 – 2.5	Insatisfecho
2.5 – 5.0	Medianamente Satisfecho
5.0 – 7.5	Satisfecho
7.5 – 10.0	Muy Satisfecho

Tabla 31: Rango de grado de satisfacción

#### Pregunta N° 1:

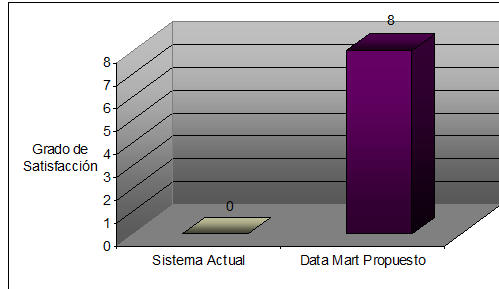
¿Se puede conocer la cantidad total de ventas en soles de varios productos en forma comparativa en un determinado mes?



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre la cantidad total de ventas de varios productos con comparativas de meses mediante tablas y gráficos.

**Pregunta N° 2:**

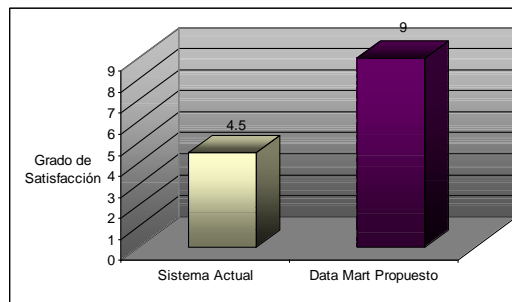
¿Se puede conocer la cantidad de ventas en soles por vendedor en un mes determinado?



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre la cantidad de ventas en soles de los vendedores haciendo una comparación de ventas de meses anteriores.

**Pregunta N° 3:**

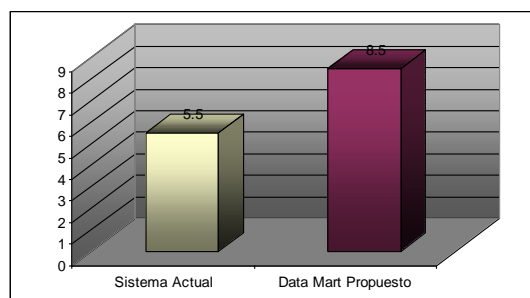
¿Se puede conocer cuál es la cantidad de productos vendidos por tipo y categoría en un mes determinado?



**Resultado:** Los valores obtenidos son: 4.5 (S.A.) y 9 (D.M.P). El Sistema Actual emite un resumen de productos vendidos pero no se puede hacer comparaciones entre periodos. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento.

**Pregunta N° 4:**

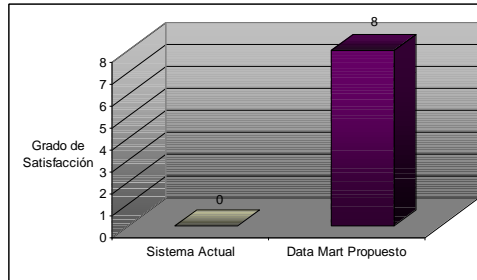
¿Se puede conocer cuál es la cantidad de productos comprados por cliente en un determinado mes?



**Resultado:** Los valores obtenidos son: 5.5 (S.A.) y 8.5 (D.M.P). El Sistema Actual permite conocer las compras de clientes, pero no lo presenta en forma dinámica. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento.

**Pregunta N°5:**

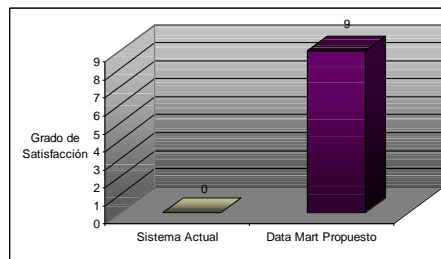
¿Se puede conocer cuál es la venta total en soles de productos comprados por cliente en un determinado mes?



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre las ventas totales que se hizo a un cliente. El Sistema Actual no presenta esta consulta.

**Pregunta N° 6:**

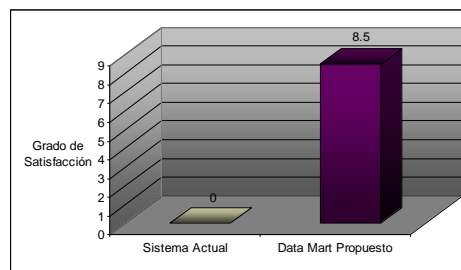
¿Se puede saber cuál es el costo en soles de un producto determinado en diferentes tiempos?



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre los costos de los productos en diferentes periodos de tiempo. El Sistema Actual no presenta esta consulta.

**Pregunta N° 7:**

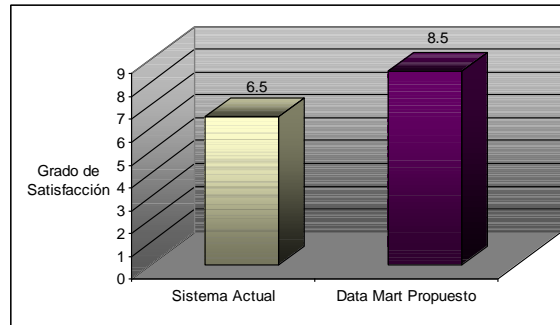
¿Se puede conocer cuál es la utilidad por mes de un producto determinado?



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre las utilidades por mes de un producto en forma dinámica. El Sistema Actual no presenta esta consulta.

**Pregunta N° 8:**

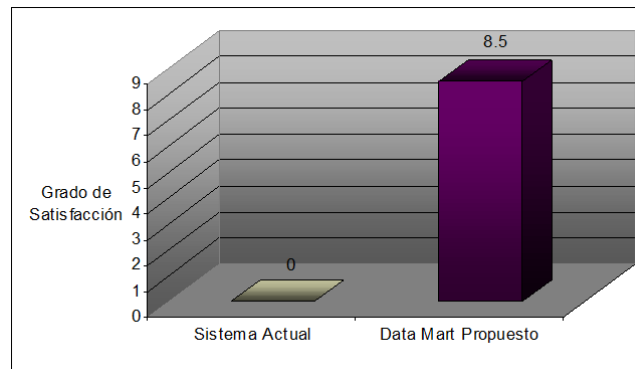
¿Se puede conocer cuál es la cantidad de productos que se compraron en un mes determinado?



**Resultado:** Los valores obtenidos son: 6.5 (S.A.) y 8.5 (D.M.P). El Sistema Actual permite conocer el número de productos que se compraron, pero no lo presenta en forma dinámica. El Data Mart presenta dinámicamente y con gráficos para un mejor entendimiento.

**Pregunta N° 9:**

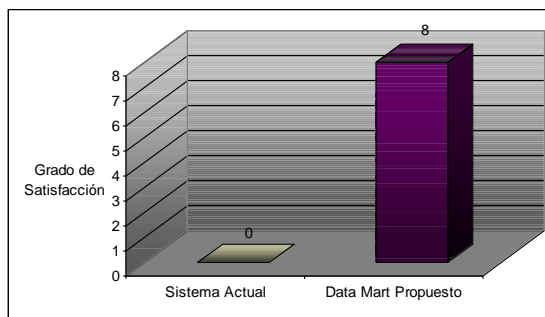
¿Se puede conocer cuál es la cantidad de productos que tuvieron alguna transferencia en un determinado tiempo?



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8.5 (D.M.P). El Data Mart muestra información sobre los productos transferidos y cuál fue su motivo en forma dinámica. El Sistema Actual no presenta esta consulta.

**Pregunta N° 10:**

¿Se puede conocer cuál es la cantidad de productos comprados por proveedor en un mes determinado?



**Resultado:** Los valores obtenidos son: 0 (S.A.) y 8 (D.M.P). El Data Mart muestra información sobre los productos comprados a los proveedores haciendo comparativas de tiempo en forma dinámica. El Sistema Actual no presenta esta consulta.

**3.2.3. Aplicación del rango de satisfacción a los Indicadores de la Hipótesis.**

Los valores aplicados a los indicadores de la hipótesis tanto para el sistema Actual como para el Data Mart Propuesto se muestran en la siguiente tabla:

**Evaluación de los indicadores de la hipótesis:**

<b>INDICADORES</b>	<b>S.A.</b>	<b>D.M.P</b>
¿Se puede conocer la cantidad total de ventas en soles de varios productos en forma comparativa en un determinado mes?	0	8.5
¿Se puede conocer la cantidad de ventas en soles por vendedor en un mes determinado?	0	8
¿Se puede conocer cuál es la cantidad de productos vendidos por tipo y categoría en un mes determinado?	4.5	9
¿Se puede conocer cuál es la cantidad de productos comprados por cliente en un determinado mes?	5.5	8.5
¿Se puede conocer cuál es la venta total en soles de productos comprados por cliente en un determinado mes?	0	8
¿Se puede saber cuál es el costo en soles de un producto determinado en diferentes tiempos?	0	9

¿Se puede conocer cuál es la utilidad por mes de un producto determinado?	0	8.5
¿Se puede conocer cuál es la cantidad de productos que se compraron en un mes determinado?	6.5	8.5
¿Se puede conocer cuál es la cantidad de productos que tuvieron alguna transferencia en un determinado tiempo?	0	8.5
¿Se puede conocer cuál es la cantidad de productos comprados por proveedor en un mes determinado?	0	8
<b>PROMEDIO</b>	<b>1.65</b>	<b>8.45</b>

Tabla 32: Evaluación de los indicadores de la hipótesis

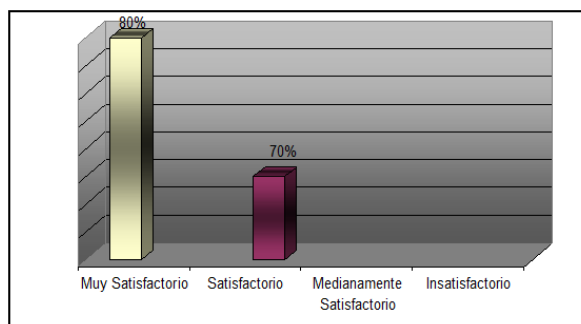
### 3.3. Manera no Presencial

La muestra que se tomó fue la misma a los cuales se les aplicó un cuestionario (Ver Anexo C) el cual comprobará si el Data Mart obtiene información valiosa para los procesos de Ventas e Inventario.

#### Cuestionario dirigido al Jefe de Ventas e Inventario.

##### Pregunta N° 1:

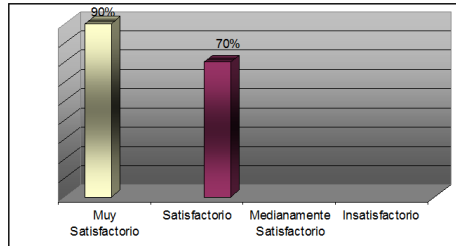
¿De qué manera se realizó el ingreso a los datos que contiene el Data Mart?



**Resultado:** El Jefe del Área de Ventas respondió que el ingreso a los datos que contiene el Data Mart se realizó muy satisfactoriamente (80%). El Jefe del Área de Informática respondió el ingreso a los datos que contiene el Data Mart se realizó satisfactoriamente (70%).

**Pregunta N° 2:**

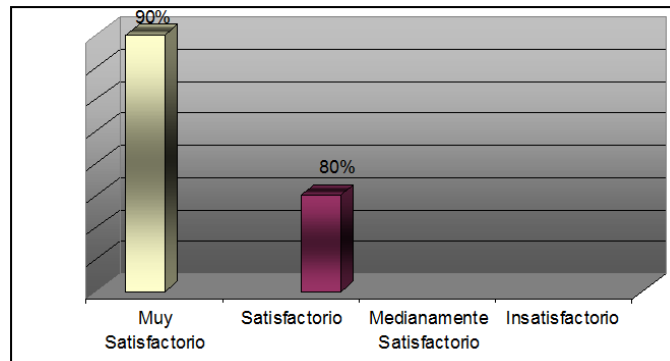
¿El Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas?



**Resultado:** El Jefe del Área de Ventas respondió que el Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas de una manera muy satisfactoriamente (90%). El Jefe del Área de Informática respondió que el Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas de una manera satisfactoriamente (70%).

**Pregunta N° 3:**

¿El Data Mart puede realizar comparaciones de información de diferentes semestres al mismo tiempo?

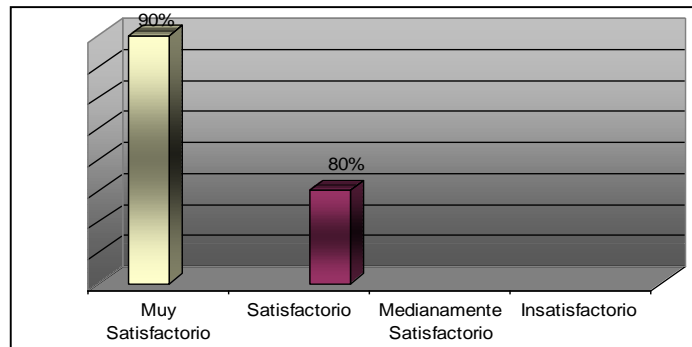


**Resultado:** El Jefe del Área de Ventas respondió que el Data Mart realiza comparaciones de información de diferentes semestres al mismo tiempo de manera muy satisfactoriamente (90%). El Jefe del área de informática respondió que el Data Mart realiza comparaciones de información de diferentes semestres al mismo tiempo de manera muy satisfactoriamente (80%).



**Pregunta N° 4:**

¿El Data Mart permitió un análisis dinámico de los datos?



***Resultado:*** El jefe del Área de Ventas respondió que el Data Mart permitió un análisis dinámico de los datos muy satisfactoriamente (90%). El Jefe del área de Informática respondió que el Data Mart permitió un análisis dinámico de los datos muy satisfactoriamente (80%).

**CONCLUSIÓN:**

Luego de haber evaluado a los usuarios del Data Mart concluimos que, el Data Mart presenta información valiosa y de forma dinámica para un mejor análisis de los datos que pueda dar soporte en la toma de decisiones.

## CONCLUSIONES

- Se determinó que el alcance del proyecto está en el área de Ventas.
- Se determinaron 10 requerimientos que la gerencia necesita, en base a las entrevistas.
- Con el análisis de los requerimientos se logró identificar la jerarquía de los datos en cada consulta hecha por el usuario determinando el modelo de datos Starnet.
- Al realizar el diseño se identificaron 2 tablas hecho y 6 dimensiones.
- La construcción del Datamart se realizó utilizando las herramientas de SQL Server 2008 R2, como son SQL Server Business Intelligence Development y Ms Excel 2010.
- Se realizaron diversas pruebas a fin de garantizar el correcto funcionamiento del Data Mart, estas pruebas se realizaron utilizando el Excel 2010.

## RECOMENDACIONES

- Es necesario realizar coordinaciones con los directivos y el personal para realizar capacitaciones sobre la importancia del manejo de las tecnologías y cultura analítica dentro de la empresa.
- Se debe realizar una evaluación periódica de los procesos de Tecnologías de Información de la empresa, para dirigirlos hacia los lineamientos estratégicos de la institución.
- Realizar un Planeamiento Estratégico de Tecnologías de Información con la finalidad de identificar posibles proyectos de TI, para optimizar los procesos en cuanto al ahorro y/o a la reducción de costos.
- Para la fase de análisis de requisitos se recomienda utilizar el análisis de molde de consulta para una mejor detección de las dimensiones y medidas candidatas a intervenir en el data Mart.
- Para proyectos futuros se recomienda el uso de la metodología de Ralph Kimball y de SQL Server 2010 como herramientas de trabajo para el diseño de DataMart.
- Para próximas versiones se recomienda el uso de dashboard que ofrece la Herramienta de Qlikview para un mejor control de los indicadores.

## REFERENCIAS BIBLIOGRÁFICAS

- CHUNGA, Ana y BOTTON, Jessica. (2013). *DATA MART para los procesos de ventas e inventarios de la empresa Consorcio Electronic SAC utilizando la metodología de Ralph Kimball, SQL Server y Qlikview.*
- FOWLER, MARTÍN. (1999). *Refactoring: mejorar el diseño de código existente.* Edición Addison-Wesley
- GARCÍA-RINCÓN DE CASTRO, César. *9 Metáforas para gestionar conflictos.* Edición Kindle
- HERNÁNDEZ CELIS, Domingo. (2009). *Tesis: Costos, Presupuestos y Toma de Decisiones Empresariales.* Lima.
- JARA, Ernesto y SOBERON, Carlos. (2011). *Desarrollo de un Data Mart para la Gestión de Créditos para la empresa Mi Crédito Perú SAC, basado en la metodología de Ralph Kimball y como herramientas a PENTAHO Y MYSQL.*
- KENDALL, K y KENDALL, J. (2012). *Análisis y Diseño de Sistemas,* 8° Edición. Prentice-Hall. Hispanoamericana.
- KIMBALL y ROSS. (2013). *The Data Warehouse Toolkit.* 3rd Edition.
- LANDAURO, Juan. (2008). *Metodología para crear y/o agregar valor con la incorporación de estrategias de E-Business en empresas tradicionales.*
- LAUDON, K.C. y LAUDON, J.P.(2013), *Management Information Systems: Managing the Digital Firm.* 13° Edition. Prentice-Hall PTR. Estados Unidos.
- ROBBINS y COULTER. (2005). *Management.* 8° Edition Prentice Hall
- RODRIGUEZ, Jesús. (2011). *Toma de decisiones. Más allá de la intuición,* Primera edición. CreateSpace Independent Publishing Platform.
- ROMERO, Liliam y VINCES, Erika. (2012). *Diseño de un Datamart para el Soporte a la Toma de Decisiones en el área de Ahorro de la Caja Municipal de Ahorro y Crédito de Trujillo usando la Metodología Ralph Kimball y SQL Server 2008 BUSINESS INTELLIGENCE.*
- ROOT, R. y MASON C. (2012). *Pro SQL Server 2012 BI Solutions.* 1° Edition. Apress. Estados Unidos.
- VILLALOBOS y CONSTENIA. (2010). *Impacto del Uso de Tecnologías de Información y Comunicación (Tics) en la Docencia Universitaria.* Editorial Chile.

# ANEXOS

## ANEXO A

### ENTREVISTAS Y CUESTIONARIOS

#### CUESTIONARIO PARA ENTREVISTA N° 1: Gerente

##### 1. LAS RESPONSABILIDADES

- Describe su área y su relación con el resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

##### 2. LOS OBJETIVOS COMERCIALES Y PROBLEMAS

- ¿Cuáles son los objetivos de su área?
- ¿Qué usted está tratando de lograr con estos objetivos?
- ¿Cuáles de estos objetivos son su prioridad para alcanzar sus metas dentro de su organización?
- ¿Cuáles son sus factores críticos de éxito?
- ¿Cómo usted sabe que usted está haciendo bien?
- ¿Qué tan menudo usted mide los factores de éxito importantes?
- ¿De los departamentos que funcionan? ¿cuáles son cruciales para asegurar que los factores de éxito importantes se logren?
- ¿Qué roles cumplen estos departamentos?
- ¿Cómo ellos trabajan para asegurar el éxito juntos?
- ¿Cuáles son los importantes problemas que usted enfrenta hoy dentro de su función? Y ¿Cuál es el impacto en la organización?
- ¿Cómo usted identifica sus problemas en su Área o sabe que usted se dirige hacia el problema?

##### 3. ANALISIS DE LOS REQUISITOS

- En el análisis de los datos ¿Qué papel juega las decisiones que usted y otros gerentes toman en la ejecución del negocio?

- ¿Qué información importante se exige a hacer o a apoyar las decisiones que usted hace en el proceso de lograr sus metas y superar los obstáculos? ¿Cómo usted consigue esta información hoy?
- ¿Está allí otra información que no está disponible a usted hoy y que usted cree tendría el impacto significativo en ayudar a encontrar sus metas?
- ¿Están allí los cuellos de botella específicos a llegar a la información?
- ¿Qué informes usted usa actualmente?
- ¿Qué datos en el informe son importantes?
- ¿Cómo usted usa la información?
- ¿Qué problemas encuentra en estos informes?
- ¿Cuánta información histórica se requiere?
- ¿Si el informe fuera dinámico, en que lo haría diferentemente?
- ¿Qué capacidades analíticas le gustaría tener?
- ¿Qué oportunidades existen para mejorar dramáticamente su negocio basándose en el acceso mejorado de la información?
- ¿Cuál es el impacto financiero usted piensa que tendría?

## **CUESTIONARIO PARA ENTREVISTA N° 2: Jefe área de Ventas**

### **1. LAS RESPONSABILIDADES**

- Describe su organización y su relación con el resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

### **2. LOS OBJETIVOS COMERCIALES Y PROBLEMAS**

- ¿Cuáles son los objetivos de en el desempeño de su función?
- ¿Qué usted está tratando lograr con estos objetivos?
- ¿Cuáles de estos objetivos son su prioridad para alcanzar sus metas dentro de su función?
- ¿Cuáles son sus factores críticos de éxito?
- ¿Qué tan menudo usted mide los factores de éxito importantes?
- ¿Cuáles son los importantes problemas que usted enfrenta hoy?
- ¿Qué le impide cubrir sus objetivos?

- ¿Cuál es el impacto de estos problemas en la organización?
- ¿Cómo usted identifica sus problemas en su organización o sabe que usted se dirige hacia el problema?
- ¿Cómo es la relación que tiene con el Gerente?

### **3. ANALISIS LOS REQUISITOS**

- ¿Qué tipo de análisis rutinario usted realiza actualmente? ¿Qué datos se usa?
- ¿Cómo usted consigue los datos actualmente?
- ¿Qué usted hace una vez con la información que usted obtiene?
- ¿Qué informes usted usa actualmente?
- ¿Qué datos en el informe es importante?
- ¿Cómo usted usa la información?
- ¿Qué oportunidades existen para mejorar dramáticamente su negocio basándose en el acceso mejorado de la información?

## **CUESTIONARIO PARA ENTREVISTA N° 3: Área de Sistemas**

### **1. LAS RESPONSABILIDADES**

- Describe su organización y su relación al resto de la compañía.
- ¿Cuáles son sus responsabilidades primarias?

### **2. EL APOYO AL USUARIO/ LOS ANÁLISIS Y REQUISITOS DE LOS DATOS**

- ¿Cuál es el proceso actual para hacer llegar (obtener) la información?
- ¿Qué herramientas son usadas para acceder y analizar la información hoy? ¿Quién las usa?
- ¿Le piden que realice los análisis rutinarios?
- ¿Usted crea los informes estandarizados?
- Describa las demandas de información ad hoc típicas. ¿Cuánto tiempo toma para cumplir estas demandas?
- ¿Quién son los demandantes más frecuentes de análisis y / o datos?
- ¿Cómo es el mecanismo de apoyo?

- ¿Cuál es el cuello de botella más grande / los problemas con los datos actuales que encuentran en el proceso?
- ¿Hay unos atrasos en enviar información a los demandantes?

### **3. DATOS DISPONIBILIDAD Y CALIDAD**

- ¿Qué sistemas de la fuente se usan para la información frecuentemente-pedida?
- ¿Cuál es la granularidad?
- ¿Qué tan menudo son los datos puesto al día?
- ¿Cuánta información histórica está disponible? Y ¿Cuánta necesita para realizar los informes?
- ¿Cuál es un tamaño estimado de estos datos (preliminar #de filas)?
- ¿Cuáles son los archivos principales que usted tiene? Describa el mantenimiento de éstos archivos.
- ¿Usted tiene los archivos de la fuente comunes actualmente?
- ¿Quién mantiene los archivos de la fuente?
- ¿Cómo las llaves se mantienen? ¿Son las llaves reasignadas?
- ¿Cuál es el cardinalidad (#los valores distintos)?



**ANEXO B**

**CUESTIONARIO DIRIGIDO: JEFE DEL AREA DE VENTAS Y AL JEFE DEL AREA DE INFORMATICA**

PREGUNTAS	VALORES										
	0	1	2	3	4	5	6	7	8	9	10
1. ¿Se puede conocer la cantidad total de ventas en soles de productos en forma comparativa en un determinado mes?											
2. ¿Se puede conocer la cantidad de ventas en soles por vendedor en un mes determinado?											
3. ¿Se puede conocer cuál es la cantidad de productos vendidos por tipo y categoría en un mes determinado?											
4. ¿Se puede conocer cuál es la cantidad de productos comprados por cliente en un determinado mes?											
5. ¿Se puede conocer cuál es la venta total en soles de productos comprados por cliente en un determinado mes?											
6. ¿Se puede saber cuál es el costo en soles de un producto determinado en diferentes tiempos?											
7. ¿Se puede conocer cuál es la utilidad por mes de un producto?											
8. ¿Se puede conocer cuál es la cantidad de productos que se compraron en un mes determinado?											
9. ¿Se puede conocer cuál es la cantidad de productos que tuvieron alguna transferencia en un determinado tiempo?											
10. ¿Se puede conocer cuál es la cantidad de productos que tuvieron alguna transferencia en un determinado tiempo?											

Tabla 33: Cuestionario Dirigido al Jefe del Área de Ventas y al Jefe del área de Informática

## ANEXO C

### CUESTIONARIO DIRIGIDO AL JEFE DEL AREA DE VENTAS Y EL JEFE DEL AREA DE INFORMATICA

PREGUNTAS	VALORES										
	0	1	2	3	4	5	6	7	8	9	10
1. ¿De qué manera se realizó el ingreso a los datos que contiene el Data Mart?											
2. ¿El Data Mart permitió realizar consultas que se consideran valiosas para el Área de Ventas?											
3. ¿El Data Mart puede realizar comparaciones de información de diferentes semestres al mismo tiempo?											
4. ¿El Data Mart permitió un análisis dinámico de los datos?											

Tabla 34: Cuestionario Dirigido al Jefe del Área de Ventas y el Jefe del Área de Informática