

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS E INGENIERÍA



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS PARA EL ÁREA DE IMPORTACIONES EN UNA EMPRESA COMERCIALIZADORA/IMPORTADORA

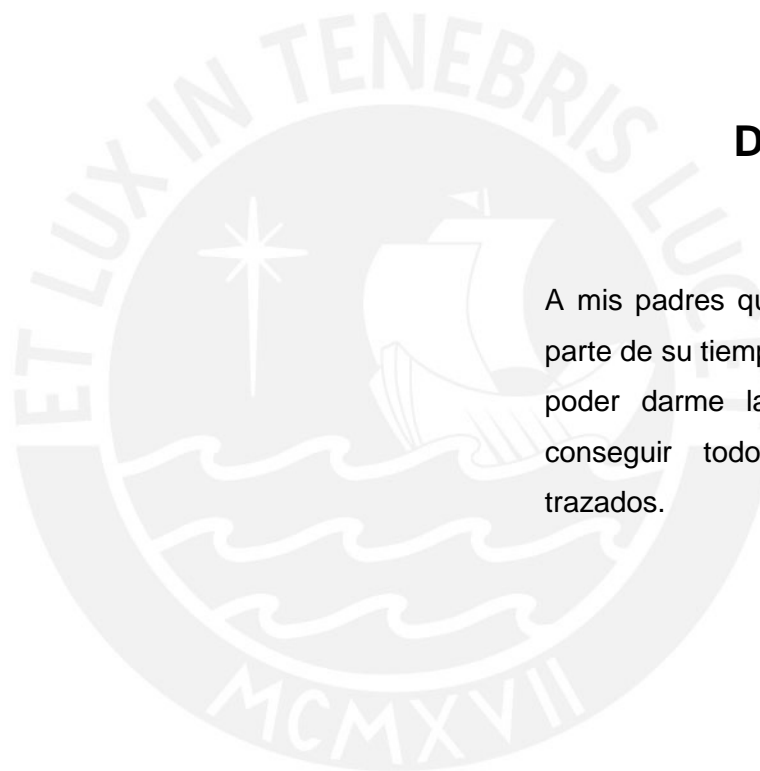
Tesis para optar por el Título de Ingeniero Informático, que presenta el alumno:

José Eduardo Córdova Yupanqui

ASESOR:

Ing. Raúl Valencia Ibarra

Lima, abril de 2013



Dedicatoria

A mis padres que dedicaron gran parte de su tiempo y esfuerzo para poder darme la oportunidad de conseguir todos mis objetivos trazados.

Resumen

El presente proyecto de tesis consiste en analizar, diseñar e implementar una solución de inteligencia de negocios, que permita realizar el análisis a nivel gerencial del área de Importaciones y Logística dentro de una empresa que comercializa e importa bienes y que tenga implantado un ERP para las operaciones dentro de sus diversos procesos de negocio. La solución consiste en una serie de elementos gráficos, flexibles y de acceso eficiente a los datos ofrecidos desde distintos orígenes; permitiendo con ello, lograr un análisis adecuado de los datos por volumen y distribuirlos por distintos filtros como fechas, ubicación geográfica, proveedores internacionales, entre otros y dar la facilidad a los usuarios para que interpreten mejor dicha información. De igual modo, la solución permitirá a los usuarios de la alta dirección de las organizaciones de este tipo de empresas tomar mejores decisiones a nivel de gestión en relación a las compras de importación.

Tabla de Contenido

Dedicatoria.....	2
Resumen.....	3
Tabla de Contenido de Figuras	6
Tabla de Contenido de Cuadros	8
Introducción	9
1. CAPÍTULO 1: GENERALIDADES	11
1.1. Definición del Problema	11
1.2. Marco Conceptual del Problema	14
1.2.1. ERP (<i>Enterprise Resource Planning</i>):	14
1.2.2. Sistemas OLTP (<i>On-Line Transaction Processing</i>):	14
1.2.3. Datamart:.....	15
1.2.4. Data Warehouse:	15
1.2.5. Data Mining (Minería de datos):.....	17
1.2.6. Sistemas OLAP (<i>On-line Analytical Processing</i>)	17
1.2.7. Arquitectura ROLAP (<i>Relational On-line Analytical Processing</i>)	18
1.2.8. Arquitectura MOLAP (<i>Multimentional On-line Analytical Processing</i>) ..	18
1.2.9. Arquitectura HOLAP (<i>Hybrid On-line Analytical Processing</i>).....	19
1.2.10. Metadata.....	19
1.2.11. Métrica	20
1.2.12. Jerarquía.....	20
1.2.13. Cluster	20
1.2.14. <i>Slow Moving</i> (Movimiento Lento)	20
1.2.15. Incoterms	20
1.2.16. Exportación a Título Definitivo.....	21
1.2.17. Depósitos Aduaneros	21
1.2.18. Tránsito Aduanero	21
1.2.19. Traslado	21
1.3. Plan de Proyecto.....	21
1.4. Estado del Arte	24
1.4.1. Soluciones Alternativas Actuales.....	24
1.4.2. Importancia de las soluciones de Inteligencia de Negocios en el mundo 26	
1.4.3. Soluciones de Inteligencia de Negocios Actuales	26
1.4.3.1. IBM Data Stage	27
1.4.3.2. Java – Octopus.....	27
1.4.3.3. Qlik View.....	28
1.4.3.4. MicroStrategy	28
1.4.3.5. Business Objects.....	28
1.4.4. Comparación de Herramientas	29
1.5. Descripción y sustentación de la solución	30
1.5.1. Objetivo General	31
1.5.2. Objetivos Específicos.....	31
1.5.3. Alcances y Limitaciones.....	32
2. CAPÍTULO 2: ANÁLISIS.....	34
2.1. Metodología aplicada para el desarrollo de la solución.....	34
2.1.1. Planificación	35
2.1.2. Análisis	35
2.1.3. Diseño	36
2.1.4. Construcción	36
2.1.5. Pruebas	37

2.1.6.	Implantación.....	37
2.2.	Análisis de la situación actual.....	38
2.3.	Identificación de requerimientos.....	42
2.4.	Análisis de la solución.....	47
2.4.1.	Requerimiento de información de los usuarios.....	47
2.4.2.	Análisis Técnico y económico del proyecto.....	49
3.	CAPÍTULO 3: DISEÑO.....	52
3.1.	Arquitectura de la solución.....	52
3.1.1.	Modelado de la Arquitectura-Diseño de la Base de Datos.....	52
3.1.2.	Fuentes de Información.....	53
3.1.3.	Procesos ETL.....	55
3.1.4.	Datamart.....	55
3.1.5.	Servidor de Presentación.....	56
3.2.	Diseño de la solución.....	57
3.2.1.	Justificación de la arquitectura OLAP.....	57
3.2.2.	Elementos para el diseño del modelamiento.....	58
3.2.2.1.	Hecho.....	59
3.2.2.2.	Medida.....	59
3.2.2.3.	Tabla de Hechos (Fact Table).....	59
3.2.3.	Lineamientos para el diseño lógico dentro del modelamiento.....	59
3.2.3.1.	Números Enteros.....	59
3.2.3.2.	Números decimales.....	60
3.2.3.3.	Fecha.....	60
3.2.3.4.	Cadena de Caracteres.....	60
3.2.4.	Diseño de la extracción.....	60
	Medidas.....	62
	Medidas.....	64
3.3.	Diseño de la explotación de datos.....	65
4.	CAPÍTULO 4: CONSTRUCCIÓN Y PRUEBAS.....	71
4.1.	Construcción.....	71
4.1.1.	Construcción de la ETL.....	71
4.1.2.	Construcción del Servidor de Explotación.....	77
4.2.	Pruebas.....	80
4.2.1.	Planificación del plan de Pruebas.....	81
4.2.2.	Diseño del plan de pruebas.....	82
4.2.3.	Determinación de los casos de prueba.....	82
4.2.4.	Ejecución y análisis de los casos de prueba.....	83
5.	CAPÍTULO 5: OBSERVACIONES, CONCLUSIONES Y RECOMENDACIONES.....	93
5.1.	Observaciones.....	93
5.2.	Conclusiones.....	94
5.3.	Recomendaciones.....	95
6.	ANEXOS.....	¡Error! Marcador no definido.
6.1.	ANEXO A - Plan de Proyecto.....	¡Error! Marcador no definido.
6.2.	ANEXO B - Modelo de Base de Datos del Data Warehouse.....	¡Error! Marcador no definido.
	Marcador no definido.	
6.3.	ANEXO C – Documento de Extracción.....	¡Error! Marcador no definido.
6.4.	ANEXO D – Documento de Explotación.....	¡Error! Marcador no definido.
6.5.	ANEXO E - Glosario de Términos.....	¡Error! Marcador no definido.

Tabla de Contenido de Figuras

FIGURA 1.1-	INTEGRACIÓN DE MÓDULOS EN ERP	14
FIGURA 1.2-	DATAMARTS EN UNA EMPRESA	16
FIGURA 1.3-	INTEGRACIÓN DE BASES DE DATOS PARA UN DATA WAREHOUSE	16
FIGURA 1.4-	CUBO OLAP (ON LINE ANALYTICAL PROCESSING)	18
FIGURA 1.5-	ARQUITECTURAS OLAP	19
FIGURA 1.6-	FASES DE EJECUCIÓN ENFOQUE PMBOK	22
FIGURA 1.7-	CICLO DE VIDA PARA IMPLEMENTACIÓN DEL PROYECTO.....	23
FIGURA 1.8-	EDT DEL PROYECTO	25
FIGURA 1.9-	DIAGRAMA GANTT DEL PROYECTO	26
FIGURA 2.1-	DIAGRAMAS DE PROCESOS DE IMPORTACIÓN	40
FIGURA 2.2-	CREACIÓN DE PROVEEDOR INTERNACIONAL.....	41
FIGURA 2.3-	PROCESO DE REVALORIZACIÓN DE ARTÍCULOS POR IMPORTACIÓN.....	41
FIGURA 3.1-	ESQUEMA DE ARQUITECTURA.....	53
FIGURA 3.2-	DIAGRAMA RELACIONAL DE BASE DE DATOS SAP FUENTE	54
FIGURA 3.3-	ESQUEMA LÓGICO DEL MODELO ESTRELLA DE LA TABLA FACT_IMPORTACION	61
FIGURA 3.4-	ESQUEMA LÓGICO DEL MODELO ESTRELLA DE LA TABLA FACT_OPERADOR	63
FIGURA 3.5-	MAPEO HACIA LAS TABLAS FÍSICAS DENTRO DE LA FASE DE EXPLOTACIÓN	66
FIGURA 3.6-	JERARQUÍA DE LA TABLA PRODUCTO	67
FIGURA 3.7-	DISEÑO DE EXPLOTACIÓN DE HERRAMIENTA ORACLE WAREHOUSE BUILDER.....	67
FIGURA 3.8-	REPORTE DE PROVEEDORES INTERNACIONALES	69
FIGURA 3.9-	REPORTE DE GASTOS REFERENTE A OPERADORES	69
FIGURA 3.10-	ANÁLISIS DE CONDICIONES INTERNACIONALES	70
FIGURA 4.1-	PAQUETES DE EXTRACCIÓN DE HECHOS Y DIMENSIONES.....	72
FIGURA 4.2-	EXTRACCIÓN Y CARGA DE LA DIMENSIÓN PRODUCTO	73
FIGURA 4.3-	EXTRACCIÓN Y CARGA DE LAS DIMENSIONES PROVEEDOR Y OPERADOR	74
FIGURA 4.4-	EXTRACCIÓN Y CARGA DE LAS DIMENSIONES ANEXAS	74
FIGURA 4.5-	CONSULTA DE EXTRACCIÓN DE LA TABLA PRODUCTO	75
FIGURA 4.6-	PROCESO DE EXTRACCIÓN DE LA TABLA FT_IMPORTACION	76
FIGURA 4.7-	PROCESO DE EXTRACCIÓN DE LA TABLA FT_OPERADOR	76
FIGURA 4.8-	DISEÑO FÍSICO, LÓGICO Y DE PRESENTACIÓN	77
FIGURA 4.9-	DIMENSIONES Y TABLAS DE HECHOS EN EL CUBO	78
FIGURA 4.10-	CONFIGURACIÓN DE REPORTE EN EL CUBO	79
FIGURA 4.11-	RESULTADOS DE CONSULTA	79
FIGURA 4.12-	CARPETAS DE CONSULTAS AL CUBO	80
FIGURA 4.13-	PROCESO ETL PARA TABLA DE HECHOS FT_IMPORTACION	87
FIGURA 4.14-	PROCESO ETL PARA TABLA DE HECHOS FT_OPERADOR.....	90
FIGURA 4.15-	CONFIRMACIÓN DE LA CONSULTA EXTRAÍDA DESDE EL CUBO	91



Tabla de Contenido de Cuadros

CUADRO 1.1-	CARACTERÍSTICAS OLTP VS OLAP	15
CUADRO 1.2-	COMPARATIVO DE HERRAMIENTAS EN LA FASE DE ETL	29
CUADRO 1.3-	COMPARATIVO DE HERRAMIENTAS EN LA FASE DE PRESENTACIÓN	30
CUADRO 2.1-	REQUERIMIENTOS FUNCIONALES PARA EL PROYECTO	45
CUADRO 2.2-	REQUERIMIENTOS NO FUNCIONALES PARA EL PROYECTO.....	46
CUADRO 2.3-	INDICADORES PARA EL APROVISIONAMIENTO DE COMPRAS	48
CUADRO 2.4-	INDICADORES PARA EL ANÁLISIS DE TRANSPORTE	48
CUADRO 2.5-	INDICADORES PARA EL ANÁLISIS DEL OPERADOR LOGÍSTICO	49
CUADRO 2.6-	PRESUPUESTO DE COSTOS PARA EL PROYECTO	51
CUADRO 3.1-	DIFERENCIAS TÉCNICAS ENTRE SISTEMAS OLAP Y OLTP.....	57
CUADRO 3.2-	CAMPOS DE TABLA DE HECHOS IMPORTACIÓN	62
CUADRO 3.3-	CAMPOS DE TABLA DE HECHOS FT_OPERADORES	64
CUADRO 3.4-	MAPEO DE DIMENSIONES Y TABLAS DE HECHOS	65
CUADRO 3.5-	REPORTES PARA EL ÁREA DE IMPORTACIONES.....	68
CUADRO 4.1-	OBJETIVOS DE PRUEBA Y SUS CRITERIOS DE ACEPTACIÓN	82
CUADRO 4.2-	CASO DE PRUEBA PARA EL DATAMART FT_IMPORTACION	83
CUADRO 4.3-	TABLA FUENTE DE COMPRAS (TABLA OPCH SAP).....	86
CUADRO 4.5-	CASO DE PRUEBA PARA EL CUBO DE MANTENIMIENTO.....	90

Introducción

Las empresas comercializadoras que se enfocan principalmente en la gestión de requerimientos de los clientes tienen una primera decisión importante dentro de sus procesos base, la cual consiste en determinar cómo escoger la mejor propuesta de compra y almacenaje para los productos que requieren comercializar.

Atender el punto anterior implica también considerar un entorno de cambios y un mundo globalizado. Las decisiones a nivel de gestión que influyan directamente en la cadena de suministro deben ser elegidas eficaz y eficientemente en base a información realmente útil para el negocio, que permita el cumplimiento de los objetivos específicos a mediano y largo plazo.

En el Perú, la mayoría de las empresas del rubro se encarga de cubrir parte del proceso de importación (almacenaje, transporte local y gestión logística en general) apoyada principalmente por los operadores logísticos para el almacenamiento temporal de sus productos, por las empresas que ofrecen servicios de envío de documentos y por las empresas que gestionan la atención en aduanas. Debido a esta dependencia la integración y al análisis de la información se vuelve crítica para la logística de importaciones de las empresas.

La productividad del área de importaciones de las empresas hace que el flujo de inventario de stocks y de costos sea correctamente analizado por las demás áreas operativas de las mismas. Por esta razón, dicha información es vital para el área de ventas (para analizar la proyección de ventas de acuerdo a disponibilidad), el área de almacén (para analizar la redistribución de entregas de clientes) y el área de contabilidad y costos (para determinar el flujo de costos de los productos en el tiempo). En este marco, el objetivo principal del área de importaciones es el análisis de la rentabilidad de las ventas tomando en consideración los costos de la mercadería involucrada en cada importación en las empresas de este rubro.

En el presente proyecto de tesis se explicaran los obstáculos que impiden que se realice un correcto análisis de la información de este tipo de empresas y el

planteamiento para la solución de los mismos (según plataformas de hardware y software justificadas).

En este contexto se propone implementar una solución de inteligencia de negocios como una herramienta que se ajuste a los requerimientos reales de una empresa del rubro en el área de importaciones para el aprovechamiento de la información de gestión relativa para la toma de decisiones en dicha área.

El proyecto de tesis se divide en los siguientes capítulos: generalidades, análisis, diseño, construcción y pruebas, y finalmente el capítulo de observaciones, conclusiones y recomendaciones, los cuales son tratados a detalle en el presente documento.

Dentro de las generalidades se describen el marco conceptual, la descripción del problema encontrado y la solución propuesta. Asimismo se examina brevemente el estado del arte existente para el proyecto.

Dentro del análisis se detallan los requerimientos funcionales y no funcionales para el problema planteado, la metodología o forma de trabajo usada para el presente proyecto y el análisis de solución en detalle.

En el diseño se detallan los componentes del modelo dimensional, la explicación de la arquitectura seleccionada y descripción del proceso de extracción y explotación orientado al problema planteado.

Dentro de la fase de construcción y pruebas se ilustran las fases de desarrollo del proceso de extracción de los componentes y los resultados de la ejecución como también el esquema de presentación de los reportes solicitados.

CAPÍTULO 1: GENERALIDADES

En este capítulo se describen el problema de los usuarios finales respecto al análisis de datos y la obtención de información, cuya solución es el objetivo principal del desarrollo del presente proyecto de tesis. Asimismo se describe el entorno en el que se presenta dicho problema y las diversas alternativas que existen actualmente para dar solución a todo o parte del problema. Finalmente se incluye una breve descripción de las características de la solución planteada para el proyecto.

1.1. Definición del Problema

Según el levantamiento de información realizado, se han identificado las siguientes consideraciones:

- La mayoría de los pedidos de importación del área de compras se rigen por las exigencias del área de ventas y, en general, el proceso se inicia con el pedido del cliente confirmado de una manera reactiva. Estas exigencias, si no han sido resueltas a tiempo por la gestión del área de importaciones, incurren en exigencias y movilización de parte de la logística interna para satisfacer las órdenes de ventas de clientes potenciales disconformes.

- Gracias a la dinámica interna del consumo, en todo el país en los últimos años, las importaciones han ido aumentando a un ritmo de 30% anual llegando hasta los 40 mil millones de dólares, según Instituto Nacional de Estadísticas e Informática (INEI).
- Según el INEI, en una importación el 65% de las empresas en el país utilizan el servicio de depósito temporal gracias a los operadores logísticos por lo cual que realizan nacionalizaciones parciales según la necesidad.
- Debido a esta coyuntura, las empresas establecen, en su gran mayoría, almacenes diversificados para sus productos como: almacenes para productos que se encuentren temporalmente en sucursales de los operadores logísticos, almacenes aduaneros (temporal sin costo), almacenes de productos dentro del inventario de seguridad de la empresa (almacén propio) y almacenes virtuales de la mercadería solicitada de importación.
- Los factores que afectan a los costos de los productos dentro de una importación como: costo base, flete, seguro, derechos arancelarios (derechos de aduana), gastos de internamiento, impuesto selectivo son determinados según la condición internacional de negociación (abreviado en inglés como incoterm) acordado entre las partes.
- La política de almacenaje de las empresas es variable usándose técnicas como FIFO (*First In-First Out*, Primeras Entradas – Primeras Salidas) y LIFO (*Last In-First Out*, Últimas Entradas – Primeras Salidas) y no se considera dentro del análisis de presupuesto para alguna compra de importación.

Dentro de una empresa importadora comercializadora el problema principal es la falta de un correcto análisis de los datos globalizados requeridos por cada línea de negocio. Esto parte de la diversificación de sistemas y la falta de integración entre ellas; además de la ausencia de herramientas de almacenamiento central que unifique la data dispersa.

Estas limitantes ocasionan que se organice la información por separado. Estos análisis pierden valor por el tiempo que toma en organizarlos manualmente, por la manipulación de los datos y posibles errores humanos que ello ocasiona.

Esto genera, dentro de los análisis internos del área de importaciones, información desactualizada (información de productos solicitados, ranking de operadores y proveedores, ranking de eficiencia de distribución). Además se corre el peligro de realizar análisis imprecisos de los costos de los productos y por consecuencia, posibles diferencias en la utilidad esperada (inclusive, convertirse en pérdida).

Del mismo modo, los datos que son explotados para la toma de decisiones en las demás áreas operativas, debido a los mismos riesgos expuestos, pueden ser considerados como información inoportuna e insuficiente. El área comercial, al no tener información del stock por venir, no puede identificar o hacer el seguimiento de la atención de los pedidos de los clientes potenciales. El área de almacén se ve afectado ya que al no tener información confiable de la llegada de los productos, comienza a redistribuir las prioridades de algunas o todas las entregas confirmadas para poder atender las ventas incluidas o relacionadas a los pedidos de importación de compra pendientes por llegar. El área contable o de costeos de las empresas también necesita de dicha información confiable para el correcto análisis de la evolución de los costos en el tiempo y evaluar el presupuesto de costos eficientemente.

A todo esto se adiciona el esfuerzo realizado muchas veces por el personal que se encarga de preparar de forma manual los archivos que servirán para el análisis de las importaciones actuales.

Por lo expuesto, queda claro que se tienen muchos inconvenientes en la obtención de información para la toma de decisiones relacionadas al área de importaciones, que no son cubiertas de forma suficiente por las áreas de compras, ventas y almacén a través de los procesos involucrados.

En este contexto el presente proyecto de tesis propone implementar una solución de inteligencia de negocios como una herramienta que se ajuste a los requerimientos reales de la empresa en el área de importaciones, añadiendo el análisis dimensional y los indicadores necesarios para el aprovechamiento de la información de las personas involucradas en las decisiones de empresas de este rubro.

1.2. Marco Conceptual del Problema

Para comprender el problema dentro del contexto planteado es necesario revisar los siguientes conceptos técnicos:

1.2.1. ERP (*Enterprise Resource Planning*):

Son sistemas estructurados que brindan soporte a la parte operativa de la organización. Permiten monitorear a nivel de registro el estado de los procesos al interior de la empresa, los avances de cada área, entre otros [DAT 001]. El reto es unir los elementos de las áreas y proporcionar a los usuarios una manera universal de utilizar la información almacenada en diferentes sistemas. La Figura 1.1 [WWW005] muestra la interacción de las necesidades de todas las áreas de la organización frente a la existencia de un ERP organizacional.

1.2.2. Sistemas OLTP (*On-Line Transaction Processing*)

Sirven para la captura de las transacciones cotidianas (ventas, compras, control de almacén, cuenta corriente, generación de notas de crédito, control de la producción, contabilidad, etc.) y es la fuente principal de las soluciones analíticas. Se encarga de dar soporte a los procesos diarios de ingreso y mantenimiento de datos en tiempo real. Proporciona soporte muy limitado y poco eficiente para la elaboración de reportes y, en consecuencia, para la toma de decisiones de la empresa. Es la estructura de base de datos de un sistema transaccional y/o sistemas ERP. El Cuadro 1.1 [JCT001] hace referencia a las diferencias que existen entre un sistema basado en arquitectura OLTP y una arquitectura OLAP.



Figura 1.1 - Integración de módulos en ERP. Fuente [WWW005]

Reportes Analíticos	OLTP
Información histórica a analizar	Para soportar la información del día a día
Data necesita de ser integrada	Data coleccionada a nivel de transacción
Diseño de base de datos desnormalizada	Diseño de base de datos normalizados

Cuadro 1.1- Características OLTP vs OLAP

1.2.3. Datamart

Representa un repositorio de datos que tienen la información de un área en específico; como lo muestra la Figura 1.2, está orientado a departamentos dentro de una organización. Puede ser implementado como una solución para problemas inmediatos. Los datamarts pueden ser de diversas bases de datos dimensionales dependiendo del tipo de análisis que se quiere desarrollar [BIT 2002].

1.2.4. Data Warehouse

Consiste en una colección de datamarts en la cual se encuentra integrada la información de una empresa y que es usada como soporte para el proceso de tomas de decisiones gerenciales. Las aplicaciones para soporte de decisiones basadas en un Data warehouse pueden hacer más práctica y fácil la explotación de los datos [JCT001].

Es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas de datos para luego procesarlas permitiendo el análisis requerido, además del tratamiento jerarquizado de la información. En la Figura 1.3 [WWW002] se ilustra la integración las bases de datos para un data warehouse.

Las decisiones estratégicas son aquellas que afectan a toda la empresa (o a una buena parte de la misma) durante un largo periodo de tiempo. Influyen, por lo tanto, en los objetivos generales de la empresa y en su modelo de negocio. Estas decisiones son tomadas por los máximos responsables de las compañías (ejecutivos, presidentes, directores generales y comités de dirección).



Figura 1.2 - Datamarts en una empresa. Fuente [WWW005]

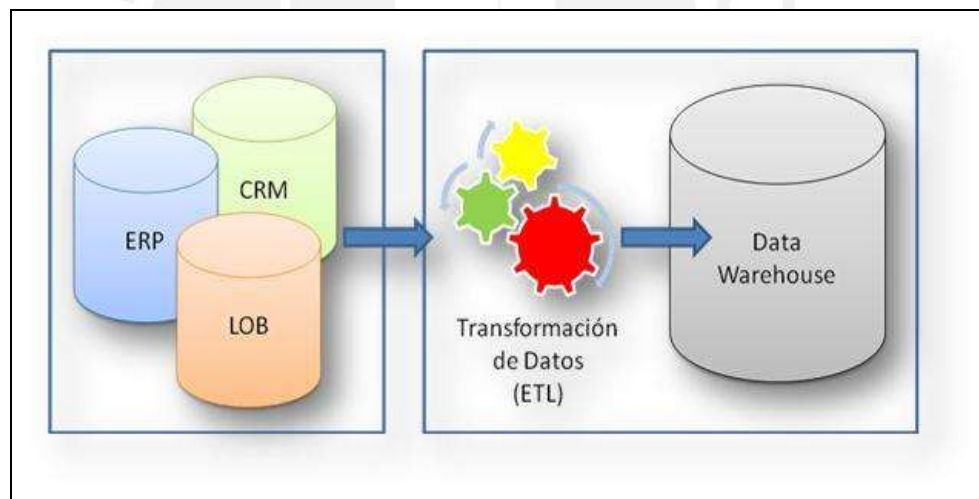


Figura 1.3 - Integración de Bases de Datos para un Data Warehouse. Fuente [WWW002]

Características Generales de un Data warehouse:

- Consolidada:
La data se centraliza desde diferentes orígenes o base de datos en un repositorio central único a la cual tienen accesos los usuarios de toda la organización.
- Consistente:

Los usuarios deben obtener una única versión de los datos no importando de qué área provengan las consultas ni el momento en que ellas se realicen.

- Orientada al objetivo:

Sólo contiene información relevante para la toma de decisiones para cada área dentro de las líneas de negocio pertinentes, de esta manera la solución se personaliza y se estructura en base a los requerimientos apropiados.

- Histórica:

Se almacena toda la información histórica de la empresa, permitiendo de esta manera realizar comparaciones entre periodos actuales e históricos. La granularidad en una solución de inteligencia de negocios hace que puedan analizarse periodos, sub-periodos y comparaciones entre los mismos.

- No volátil:

Estas soluciones únicamente se diseñan y optimizan solo para realizar consultas, las operaciones de actualización y de borrado de datos son exclusivas de los sistemas transaccionales.

- No es atómica:

Un data warehouse contiene datos sumariados que permiten eficiencia en las operaciones de consulta.

1.2.5. Data Mining (Minería de datos)

Consiste en el proceso de correr datos en algoritmos completamente sofisticados, relevando patrones y correlaciones. Esto puede ser usado para el mejor entendimiento del negocio y explotar el rendimiento de éste y mejorar el análisis [PEN008].

Es la extracción de información oculta y predecible de grandes bases de datos. Esta es una poderosa tecnología con gran potencial que ayuda a las compañías a concentrarse en la información más importante de sus bases de datos. El objetivo principal de un proceso de data mining es la búsqueda de patrones de comportamiento.

1.2.6. Sistemas OLAP (*On-line Analytical Processing*)

Se denomina al proceso en el que se emplean herramientas sofisticadas que permiten agilizar el proceso de análisis de grandes cantidades de datos desde diferentes fuentes, organizada en perspectivas (dimensiones) y métricas, permitiendo ejecutar análisis complejos de datos, en base a los cuales se

tomarán las decisiones del negocio. OLAP, permite a los usuarios una fácil y amigable navegación por la información, obteniendo el nivel de granularidad (ver Figura 1.4 [WWW007]) que requiera para la toma de decisiones. Asimismo, puede generar cálculos adicionales en base de datos existentes.

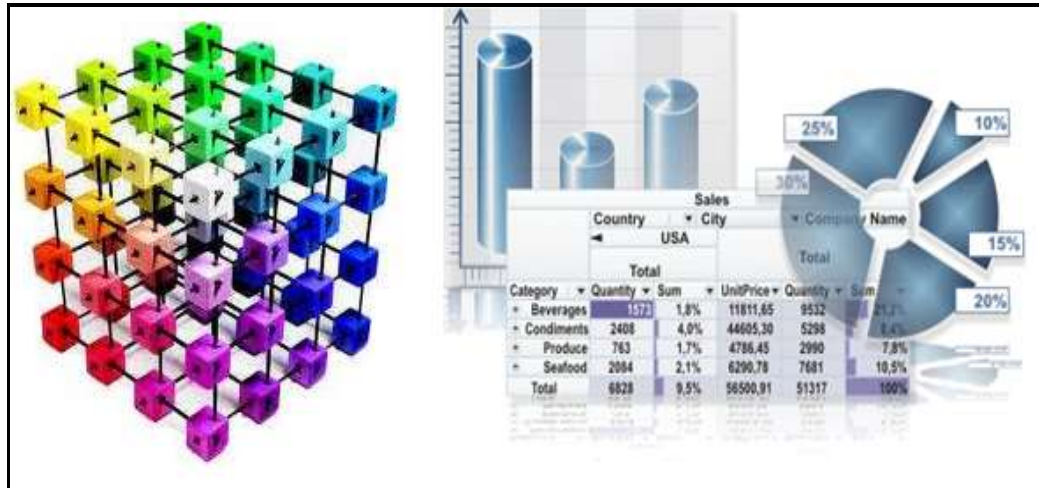


Figura 1.4 - Cubo OLAP (On Line Analytical Processing). Fuente [WWW007]

1.2.7. Arquitectura ROLAP (Relational On-line Analytical Processing)

Se denomina a la arquitectura analítica para la explotación de datos basados en una base de datos relacional en una Tercera Forma Normal. Se encuentra compuesta por un servidor de base de datos relacional donde se encuentra la data detallada de la organización y el servidor de inteligencia de negocios que se encuentra en un servidor especializado donde se trabajará la data para la presentación solicitada por el usuario final. Esta arquitectura se basa en tener como fuente de datos para la fase de presentación una base de datos relacional; la(s) fuente(s) original(es) de datos realizan un proceso ETL (Extract, Transform and Load - Extracción, Transformación y carga) trabajando sobre partes necesarias de cada una de las estructuras para dar como resultado del proceso una base de datos que mantiene una estructura de tablas que se encuentran normalizadas según la tercera Forma Normal.

1.2.8. Arquitectura MOLAP (Multidimensional On-line Analytical Processing)

Se denomina a la arquitectura analítica de la cual se explota una base de datos multidimensional en su integridad. El usuario no tiene acceso a la base de datos relacional sino que manipula directamente los datos del cubo ya

procesado. Mejora el rendimiento y tiempo de respuesta de las consultas pero existe una limitación con respecto al tamaño de la base de datos cuando tiene mayor volumen.

1.2.9. Arquitectura HOLAP (*Hybrid On-line Analytical Processing*)

Un sistema OLAP híbrido mantiene los registros detallados en la base de datos relacional, mientras que los datos resumidos o agregados se almacenan en una base de datos multidimensional separada.

Este método de almacenamiento es una combinación de los dos anteriores e intenta rescatar lo mejor de cada uno. En la Figura 1.5 [WWW007] se aprecia la diferencia de la interacción que tiene el usuario con dichas tecnologías.

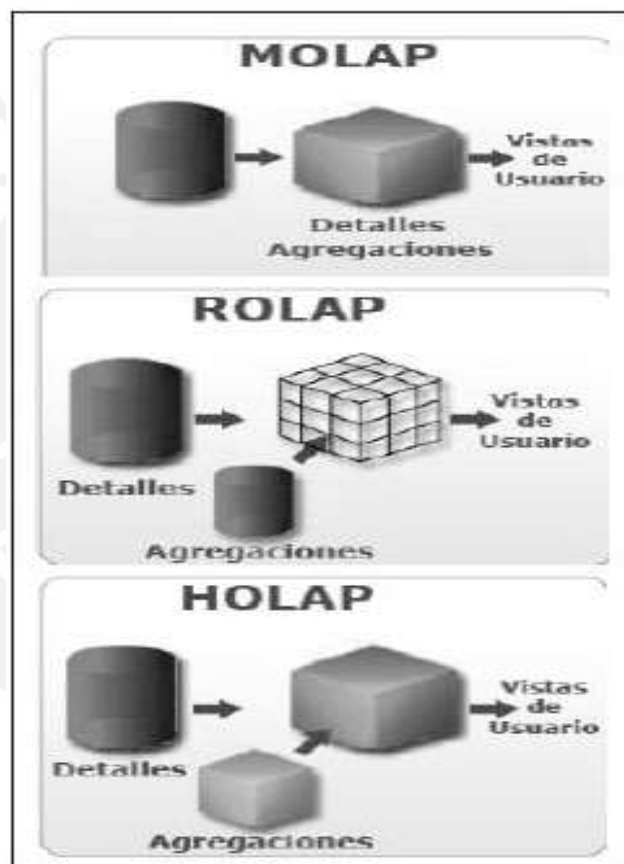


Figura 1.5 - Arquitecturas OLAP. Fuente [WWW007]

1.2.10. Metadata

Término usado para revisar el detalle acerca de los datos; contiene información sumariada de otras fuentes de datos para tener una semántica más relevante, es considerado como el diccionario de datos de un Data warehouse [WWW003].

1.2.11. Métrica

Representa un conjunto de datos de negocios en relación a una o más dimensiones. [Pen 2008].

La información numérica resumida contribuye a entender el comportamiento del negocio, esta información está contenida dentro de las tablas de hechos, esta información se denomina métrica.

1.2.12. Jerarquía

Las jerarquías pueden existir en una dimensión en la cual sirven como navegaciones predefinidas y están compuestas de uno o más niveles. Una dimensión puede tener una o más jerarquías. [Pen 2008]

Para el mejor entendimiento del proyecto también es necesaria la definición de conceptos de negocio relacionados a este, por eso se presentan los siguientes conceptos.

1.2.13. Cluster

Derivado de la palabra en inglés clúster, la cual significa conjunto o grupo. En el presente proyecto se aborda el concepto para la realización de un analítico conglomerado, donde se divide un conjunto de objetivos de compras en grupos, de forma que los perfiles de los objetivos en un mismo grupo tengan características muy similares entre si y los de los objetivos de clúster diferentes sean distintos [WWW005].

1.2.14. Slow Moving (Movimiento Lento)

El slow moving es un indicador usado en el área de compras que muestra los días que un producto no tiene movimiento. Lo óptimo es minimizar el tiempo transcurrido entre la fecha de compra y la fecha de venta de un producto.

1.2.15. Incoterms

Se refiere al conjunto de términos o condiciones usados en el rubro de importaciones y facilitan un conjunto de reglas que determinan los principales términos dentro de un contrato de compra-venta internacional [WWW004].

Estas reglas definen las obligaciones sea para el comprador o al vendedor internacionalmente cuando se realiza importaciones, se definen los siguientes aspectos:

- Lugar y forma de entrega.
- Transmisión de los riesgos y responsabilidades.

- Distribución de los gastos de operación.
- Tramites, gestiones y diligencias formales oficiales que se relacionan con la compra-venta.

1.2.16. Exportación a Título Definitivo

Este término se utiliza para las mercancías de libre circulación que abandonen el territorio aduanero y que estén destinadas a permanecer fuera de este.

1.2.17. Depósitos Aduaneros

Este concepto es utilizado para las mercancías importadas son almacenadas bajo el control de la aduana por un tiempo acordado en un lugar habilitado para esta finalidad sin pagar derechos e impuestos a la importación.

1.2.18. Tránsito Aduanero

Es la terminología usada para describir el estado de las mercancías transportadas dentro de aduanas a otra oficina bajo el control aduanero.

1.2.19. Traslado:

Es la acción por la cual se transfiere las mercancías que son retiradas del medio de transporte utilizado y cargada a otro medio de transporte para el internamiento de la mercancía en territorio del importador para la entrega definitiva esto bajo la jurisdicción de una oficina aduanera.

1.3. Plan de Proyecto

Para gestionar la implementación de la solución de inteligencia de negocios en una empresa comercializadora se han determinado las actividades a realizar de acuerdo a las buenas prácticas de gestión de proyectos propuestas por el PMBOK (*Project Management Body of Knowledge*) del PMI (*Project Management Institute*) organización sin fines de lucro con sede y bases en la Universidad de Cambridge. Esta metodología se entiende como la administración central de diversos factores como la gestión de recursos, tiempo, costo y comunicaciones.

Considerando que el presente proyecto de tesis no es de gran complejidad, en donde sólo participa el tesista, bajo la supervisión de su asesor, se vio conveniente abordar solo las siguientes áreas de conocimiento: la gestión del tiempo, del alcance y de la calidad del proyecto. En la Figura 1.6 se

esquematiza las fases de ejecución de los procesos establecidos por la metodología en el tiempo.

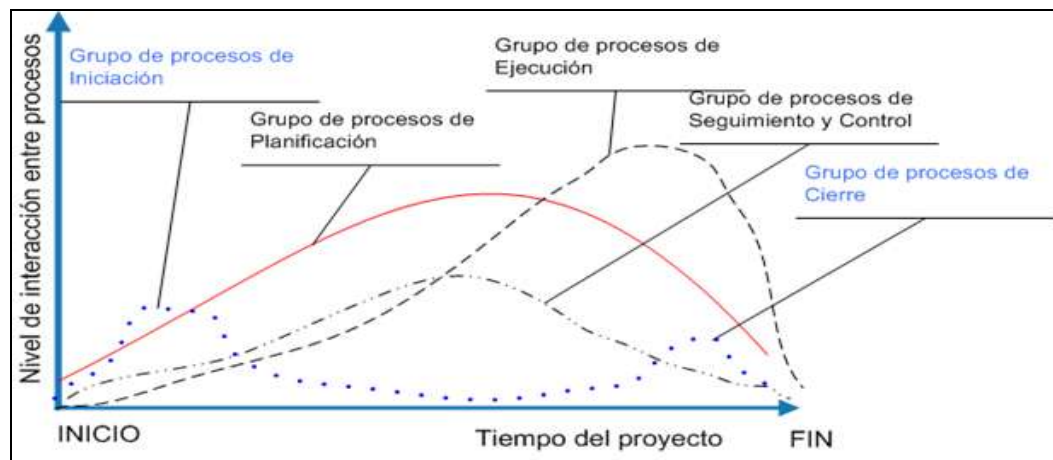


Figura 1.6 - Fases de Ejecución enfoque PMBOK. Fuente [WWW007]

Dentro del ámbito de la inteligencia de negocios, se considera a Bill Inmon y Ralph Kimball como los mayores referentes para el modelamiento de un data warehouse, ambos proponen una arquitectura de inteligencia de negocios similar, pero presentan variantes en la implantación del proyecto en una empresa.

Por un lado está la arquitectura de inteligencia de negocios llamada CIF (Corporate Information Factory) o fábrica de información corporativa, propuesta por Inmon y por el otro, arquitectura de inteligencia de negocios que se conoce como MD (Multidimensional), propuesta por Kimball. Si bien es cierto, que ambas arquitecturas tienen el mismo propósito, su principal diferencia radica en la forma en la que se modelan los datos.

Inmon sugiere utilizar el modelo entidad-relación (E-R) para el diseño de un data warehouse y el modelo de datos dimensional para el diseño de los datamarts; Kimball sugiere utilizar el modelo de datos dimensional, tanto para el diseño del data warehouse, como para el de los datamart. La arquitectura CIF tiene un enfoque descendente, siendo su primer objetivo proveer una vista empresarial de los datos a través de un data warehouse y luego ir hacia abajo construyendo los diferentes datamarts de las unidades de negocio que se requieran.

Al contrario, la arquitectura MD tiene un enfoque ascendente, ya que primero se parte desde los requerimientos funcionales de cada unidad de negocio o

departamento de una empresa construyendo un datamart. El conjunto de todos estos constituye la arquitectura del data warehouse.

Para el desarrollo o modelamiento de la solución propuesta en el presente proyecto de tesis se utilizará la metodología creada por Kimball. Se requerirá el enfoque ascendente para inteligencia de negocios ((de abajo hacia arriba)) dado que se va a requerir de la creación inicial de datamarts separadamente para consolidar todo el data warehouse. Cada entregable se encuentra detallado en el capítulo 2 del presente documento tomando en cuenta los hitos y entregables dentro de cada fase del proyecto en particular. Para una mejor comprensión se diseñó la estructura mostrada en la Figura 1.7 [PEN007] donde se aprecian las fases o hitos para la entrega de cada artefacto dentro del proyecto de desarrollo y configuración.

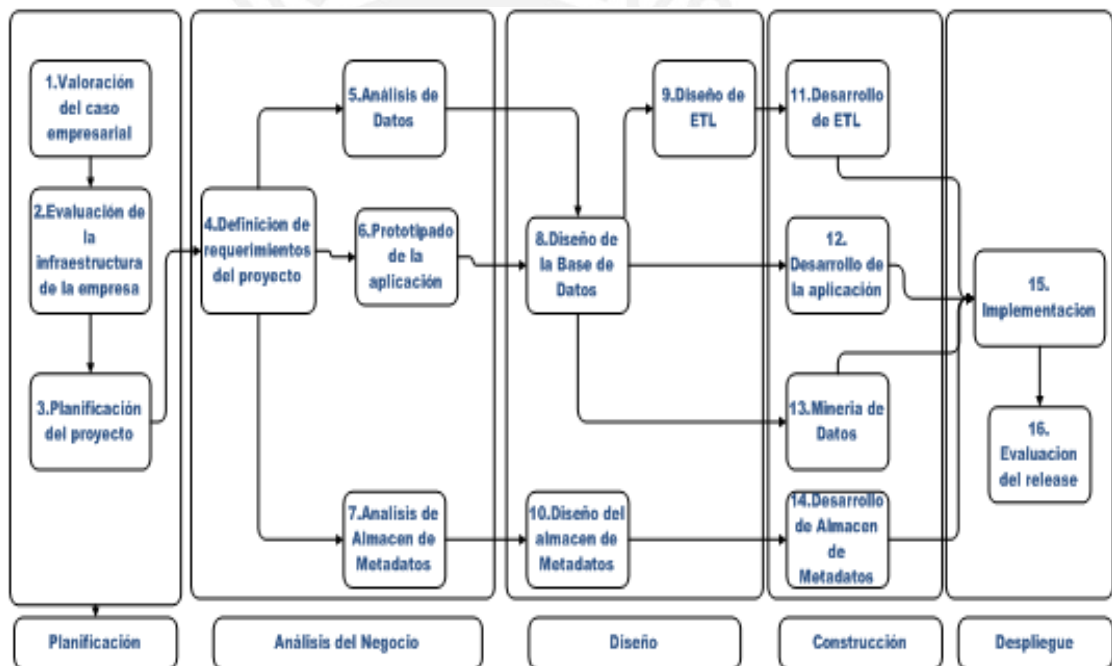


Figura 1.7 - Ciclo de vida para implementación del Proyecto. Fuente [PEN007]

El trabajo realizado en el presente proyecto de tesis se enmarca dentro de los cursos de Proyecto de Tesis 1 y Proyecto de Tesis 2.

Para poder identificar adecuadamente las tareas en el proceso de implementación de la solución de inteligencia de negocios para el área de importaciones de una empresa comercializadora/importadora de bienes se ha diseñado el EDT (Estructura de Descomposición del Trabajo) correspondiente

mostrado en la Figura 1.8. Aquí se identifican las actividades principales en cada fase del proyecto; así como los entregables y artefactos que se generen.

Habiendo completado el diseño de la descomposición de tareas para el proyecto, se elabora el cronograma fijando la duración aproximada para la realización de cada una de las tareas. Éstas se agruparon con las fases principales del proyecto para llevar el control idóneo del mismo. En la Figura 1.9 se presenta el Diagrama Gantt del Proyecto.

1.4. Estado del Arte

Muchas organizaciones a nivel global vienen implementando soluciones de inteligencia de negocios, según el INEI que desde el 2005 la tasa de crecimiento del uso de este tipo de herramientas en nuestro país va incrementándose sostenidamente con una tasa acumulativa del 13.27%.

La demanda de las empresas para el análisis eficiente de la información base de todas sus áreas motiva el crecimiento de la existencia de más herramientas para este tipo de soluciones.

1.4.1. Soluciones Alternativas Actuales

Actualmente la mayoría de empresas que necesitan este tipo de análisis buscan soluciones alternativas económicas que aplaquen temporalmente ciertos requerimientos. Las empresas normalmente recurren a este tipo de soluciones:

- La utilización de librerías dinámicas basadas en Excel de Microsoft Office para explotar la información de los ERP's existentes.
- La creación de reportes utilizando herramientas reporteadoras como Crystal Report de SAP, Jreport de Jinfo, Ireport de JasperSoft, entre otras.
- La ejecución directa de consultas hacia las bases de datos en todas sus formas: Archivos planos, bases en Access y bases de datos relacionales como Oracle o SQL.

Cada una de ellas encuentra diversas limitaciones y consideraciones de hardware y software que hacen que no sea la manera óptima de solucionar este tipo de requerimientos.

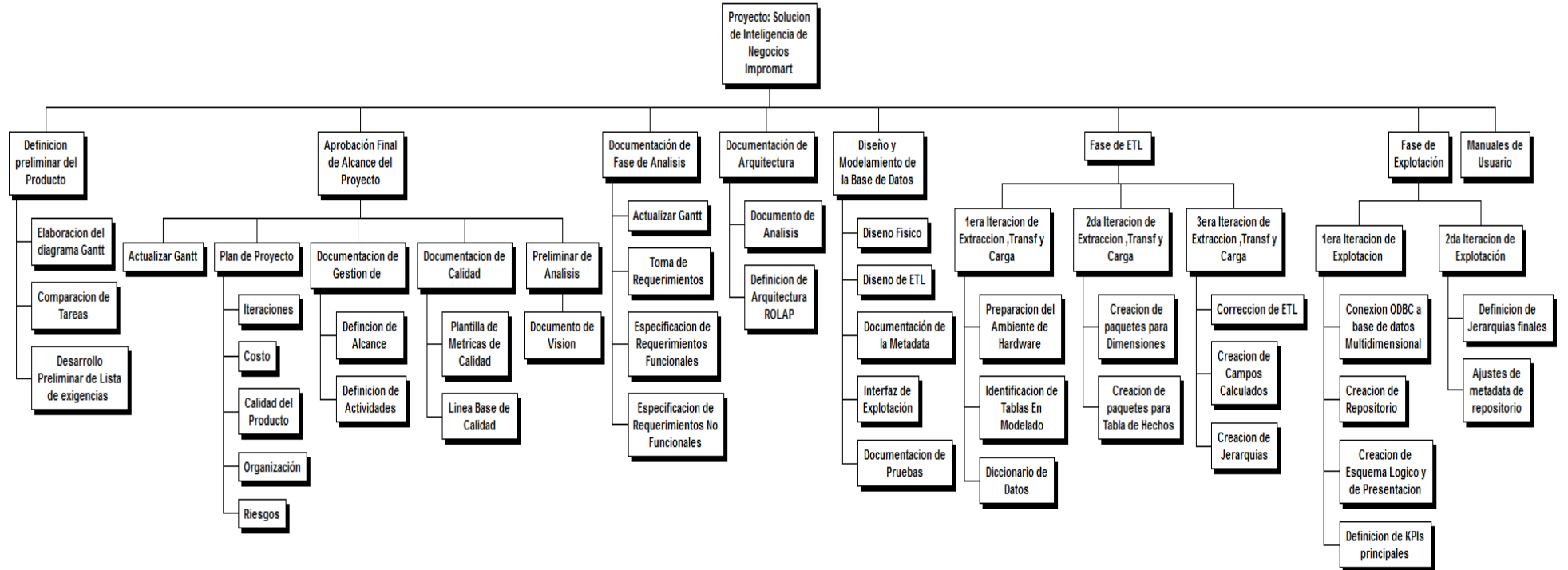


Figura 1.8- EDT del Proyecto

Id	Nombre de tarea	Duración	Comienzo	Fin	Predecesoras
1	Proyecto de Tesis	92 días	jue 02/08/12	vie 07/12/12	
2	Plan de Tesis	28 días	jue 02/08/12	lun 10/09/12	
3	Definición de Alcance	4 días	jue 02/08/12	mar 07/08/12	
4	Identificación del Problema	6 días	mié 08/08/12	mié 15/08/12	3
5	Definición del Marco Conceptual	5 días	jue 16/08/12	mié 22/08/12	4
6	Definición del Estado del Arte	5 días	jue 16/08/12	mié 22/08/12	4
7	Definición de los Procedimientos e Investigación	5 días	jue 23/08/12	mié 29/08/12	6
8	Primera Revisión del Plan de Tesis	2 días	jue 30/08/12	vie 31/08/12	7
9	Cronograma del Proyecto de Tesis	2 días	lun 03/09/12	mar 04/09/12	8
10	Segunda Revisión del Plan de Tesis	2 días	lun 03/09/12	mar 04/09/12	8
11	Corrección del Primer Entregable del Plan de Tesis	4 días	mié 05/09/12	lun 10/09/12	10
12	Revisión de Segundo Entregable	2 días	mié 05/09/12	jue 06/09/12	10
13	Marco Conceptual	9 días	mar 11/09/12	vie 21/09/12	2
14	Definición del Problema	2 días	mar 11/09/12	mié 12/09/12	12
15	Conceptos Relacionados	2 días	mar 11/09/12	mié 12/09/12	12
16	Plan de Proyecto	4 días	jue 13/09/12	mar 18/09/12	15
17	Estado del Arte	3 días	mié 19/09/12	vie 21/09/12	16
18	Revisión General de Marco Conceptual	2 días	jue 13/09/12	vie 14/09/12	15
19	Análisis	21 días	mié 19/09/12	mié 17/10/12	18
20	Metodología	4 días	mié 19/09/12	lun 24/09/12	16,18
21	Requerimientos funcionales para el area de Importación	8 días	mar 25/09/12	jue 04/10/12	20
22	Requerimientos no funcionales para el area de Importación	3 días	vie 05/10/12	mar 09/10/12	21
23	Software a utilizar	2 días	mié 10/10/12	jue 11/10/12	22
24	Revisión general del análisis	4 días	vie 12/10/12	mié 17/10/12	23
25	Diseño	27 días	jue 18/10/12	vie 23/11/12	
26	Diseño de ETL	15 días	jue 18/10/12	mié 07/11/12	24
27	Diseño de Explotación	12 días	jue 08/11/12	vie 23/11/12	26
28	Construcción y Pruebas	7 días	lun 26/11/12	mar 04/12/12	27
29	Instalación y Configuración del Software	7 días	lun 26/11/12	mar 04/12/12	27
30	Observaciones, Conclusiones y Recomendaciones	3 días	mié 05/12/12	vie 07/12/12	
31	Observaciones	3 días	mié 05/12/12	vie 07/12/12	29

Figura 1.9- Diagrama Gantt del Proyecto

1.4.2. Importancia de las soluciones de Inteligencia de Negocios en el mundo

Se sabe que en la actualidad es mayor el número de empresas que consideran la implementación y adquisición de herramientas de inteligencia de negocios como soluciones estratégicas. Estas soluciones facilitan el seguimiento y control de los objetivos de negocio mediante la generación de indicadores y métricas que permiten medir la performance de sus procesos de negocio y poder tomar decisiones oportunas.

Actualmente, las soluciones inteligencia de negocios se encuentran dentro del ranking de las diez mayores prioridades de tecnología para las empresas a nivel mundial, según encuesta ejecutada por Gartner Institute en el 2011.

1.4.3. Soluciones de Inteligencia de Negocios Actuales

Hoy en día existen distintos tipos de soluciones inteligencia de negocios con un rol importante en la ejecución y éxito de las estrategias de las empresas a

través de la información brindada y los indicadores que permiten explicar comportamientos de la eficiencia de los procesos del negocio. Las tendencias actuales e innovaciones permiten contar con grandes repositorios de datos, incluyendo datos no estructurados por sistemas transaccionales, y poder generar información al personal de dirección de las empresas para una rápida y mejor toma de decisiones, ayudando a que las organizaciones sean más competitivas en el mercado gracias a la capacidad de tomar mejores decisiones en y en corto tiempo.

Actualmente la Inteligencia de Negocios se está moviendo dentro del contexto de los procesos del negocio, no solo para permitir una experiencia de los usuarios más efectiva, sino también para permitir la optimización y retroalimentación de los procesos del negocio.

A continuación se describen las características de los productos del mercado involucrados con soluciones de inteligencia de negocios.

1.4.3.1. IBM Data Stage

Herramienta de IBM para crear y mantener rápidamente los datamarts y los data warehouses. Soporta la extracción, la integración y la transformación de volúmenes altos de datos complejos. [IBM 2009].

Las soluciones con este software especializado requieren de un hardware de múltiples procesadores, esto permite que DataStage pueda brindar altos niveles de escalabilidad y trabajar eficientemente con grandes volúmenes de datos. Numerosos "conectores" soportan una amplia variedad de formatos de datos de origen y destino, que incluye a los famosos sistemas de gestión de bases de datos de IBM y OEM (licenciamiento por volumen), las fuentes de datos ODBC (Origen de Datos), aplicaciones de terceros, mensajes en tiempo real generados por programas de colocación de mensajes en colas, XML, servicios Web, y los formatos de archivos más utilizados.

1.4.3.2. Java – Octopus

Es una herramienta de extracción y carga dentro del framework de Java que permite realizar transformaciones definidas en fichero XML independiente de la base de datos a utilizar, se utiliza la conexión JDBC. [JAV 2009].

Diferentes tipos de bases pueden ser fuente de datos (MS SQL, Oracle, DB2, QED, JDBC-ODBC con Excel y Access, Misal, CSV, y XML).

1.4.3.3. Qlik View

Muestra la forma de manejo de los datos que se cargan a la memoria RAM para hacer más rápido el acceso a los datos, con respuestas inmediatas. La visualización de dashboards y reportes dinámicos.

Visualiza datos en tablas, gráficos en 2-D o 3-D, puede hacer un Zoom in Zoom out de los gráficos y además revisar el detalle de los mismos. Se puede crear nuevas visualizaciones para una comprensión más profunda.

1.4.3.4. MicroStrategy

Herramienta de Inteligencia de negocios que permite crear informes y análisis de datos almacenados en una base de datos relacional y de otras fuentes. Esta herramienta ofrece una flexibilidad y eficiencia en la explotación y minería de datos [WWW007].

Permite el soporte para los tres niveles BI (del inglés Business Intelligence – Inteligencia de Negocios) que son el estratégico, analítico y operacional y, a la vez, ofrece un proceso sencillo de migración para fusionar aplicaciones de BI departamentales y de equipos de trabajo dentro de un marco que complementa la información de toda la organización. La seguridad implementada permite mantener autonomía e independencia en cada área y al mismo tiempo la compañía en general puede contar con información de negocios altamente eficiente, integral y consistente. Gradualmente permite fusionar aplicaciones de BI muy diversas para transformarlas en un activo homogéneo y confiable para la compañía.

Presenta un diseño en web amigable e integridad con aplicaciones SAP BW metadatos, y la mejora de la presentación de informes para incorporar a las bases de datos heterogéneas.

1.4.3.5. Business Objects

Herramienta que elabora cubos para la extracción de data de sistemas transaccionales propietario. Su lenguaje de programación llamado ABAP, es el motor de una serie de herramientas para la dirección estratégica de empresas propietario de SAP [WWW008]. Las soluciones SAP para el almacenamiento de datos proporcionan una base sólida para los datos que le permitirá recopilar, almacenar, transformar y gestionar los datos en un almacén de datos empresariales escalable o en un centro de datos flexible y ágil para la eficiente explotación de datos. Ofrece amplia gama de servicios y soluciones para las distintas versiones de SAP existente desde SAP R/3 hasta SAP Business One.

1.4.4. Comparación de Herramientas

La comparación con dichos productos fue realizada teniendo en cuenta factores como complejidad en implementación, costo, integración con otros sistemas, tecnología de reportes, complejidad en el uso, eficiencia o tiempo de respuesta. La calificación de estas herramientas se ha realizado en base a la experiencia del tesista con cada una de ellas y además de experiencia tomada por consultores especializados en diversas soluciones de este tipo.

Los criterios a consideración, fueron evaluados y valorados en un rango de puntajes, éstos dependiendo del resultado en las pruebas de concepto realizadas en cada herramienta, la toma de constantes fue realizada bajo el mismo escenario de software y hardware para todas las herramientas.

La plataforma tecnológica como unidades de almacenamiento, cantidad de funciones de inteligencia de negocios presentadas sirve de base de estudio para tal evaluación. En los Cuadros 1.2 y 1.3 se muestran los cuadros comparativos de las características que presentan las diferentes soluciones mencionadas anteriormente; en cuanto a las capacidades de dichas herramientas en las fases de carga de ETL y la fase de explotación de datos. La medición se define en escala del 1 al 5 (siendo la calificación 1 como básica hasta la 5 que es considerada muy buena).

Funcionalidad /Producto	SQL	ORACLE	Data Stage	DB2	Microstrategy	Business Object
Extracción de datos	4	5	5	4	2	2
Carga de datos	4	5	5	4	2	2
Herramientas de transformación	4	5	5	4	2	2
Integración con hojas de calculo	4	5	5	4	3	4
Complejidad de la herramienta	4	4	2	2	3	3
Capacidad de storage	4	5	4	3	2	2
Integración	4	4	4	4	5	5
Licenciamiento	3	5	4	4	3	5
Costo	5	4	4	4	4	4
Arquitectura	3	5	3	3	5	5
Soporte	4	4	4	4	4	4

Cuadro 1.2 - Comparativo de Herramientas en la Fase de ETL

Funcionalidad /Producto	SQL Reporting Services	Oracle Warehouse Builder	Microstrategy	Business Object
Cuadros de mandos analíticos	2	4	5	4
Cuadro de mando integral o estratégico(Balance Scorecard)	2	4	5	4
Query & Reporting	5	5	3	5
Preparación de Querys para entorno OLAP	4	5	3	5
Complejidad de uso	4	4	4	4
Herramientas de Datamining	3	4	3	5

Cuadro 1.3 - Comparativo de Herramientas en la Fase de Presentación

1.5. Descripción y sustentación de la solución

Se ha tomado como modelo una empresa comercializadora de bienes, dicha empresa realiza importaciones totales y parciales, como también utiliza depósitos aduaneros para almacenar grandes grupos de bienes a distribuir.

En base al levantamiento de información obtenido de las diversas entrevistas y reuniones con personal del área de importaciones, ventas y almacén se realizó el análisis, diseño e implementación de este proyecto. Se sigue la metodología de Kimball para el desarrollo del proyecto a través de un enfoque descendente (de abajo hacia arriba), escalando todos los requerimientos funcionales de cada unidad de negocio o departamento de una empresa para consolidar finalmente el data warehouse.

La solución abarca el modelo del negocio, fuentes de información, modelamiento de datos y necesidades de información a nivel de gestión de la empresa comercializadora.

La solución no abarca la creación de módulos dentro del proceso operacional de la empresa o ERP existente, en base a estas fuentes de información se obtendrá mejores decisiones para retroalimentar procesos y realizar transacciones diversas.

El control de las decisiones tomadas en la importación se basa en monitorear los factores a utilizar como tiempo, costo, garantía del proveedor, competitividad del sector, competitividad del producto, integración con el

proveedor (acuerdo, pacto o trato a nivel de integración vertical), estos factores servirán de índices de rendimiento para la elaboración de la solución. Dentro la ponderación previa adicional para la ejecución del proceso de importación se indica los siguientes aspectos:

- Rentabilidad sobre cada operación de importación.
- Volumen de compra y capacidad de almacenes propios o tercerizados.
- Frecuencia de Compra a determinados proveedores y alternativas de solución.
- Facturación anual sobre los productos y definición de prioridades para el control de stocks en almacén.

Este levantamiento de información permite aplicar esta solución a otras empresas del mismo rubro, como salvedad de que se tengan que realizar adecuaciones propias de cada empresa.

1.5.1. Objetivo General

El presente proyecto tiene como objetivo general realizar el análisis, diseño e implementación de una solución de inteligencia de negocios para el área de importaciones en una empresa comercializadora/importadora.

1.5.2. Objetivos Específicos

En base a lo que se tiene como objetivo general, se plantean los siguientes objetivos específicos para el proyecto:

- A. Definir los lineamientos para el levantamiento de información con las áreas de importaciones y almacén. Estas reuniones establecerán en cada iteración las reglas del negocio, establecer y reestructurar prioridades e indicadores claves que se fundamentarán en los reportes a elaborarse dentro de la solución.
- B. Analizar y desarrollar los requerimientos funcionales y no funcionales planteados en el levantamiento de información para el sistema.
- C. Definir una arquitectura para el proceso ETL para la correcta obtención de los datos, a fin de que se conviertan en información útil para la toma de decisiones. Administrar y ejecutar los componentes que capturan los datos desde su origen hasta llevarlos hasta el repositorio del datamart.

- D. Definir una arquitectura robusta y eficiente para la explotación de los datos por parte del usuario final.
- E. Diseñar la estructura de datos idónea para la explotación de datos. Configurar el motor de inteligencia de negocios después del proceso de ETL para la presentación de manera eficiente de los datos.
- F. Elaborar un modelo de presentación amigable que permita la preparación de sentencias lógicas simples para el usuario.
- G. Preparar la documentación del sistema a implementar.

1.5.3. Alcances y Limitaciones

El alcance del proyecto abarca desde la gestión del mismo, pasando por la implementación del ETL y concluyendo con la elaboración y presentación de los reportes que ayuden a la toma de decisiones en el negocio.

La solución permitirá filtrar el análisis de la data por múltiples formas de negociación dentro de una importación, como también por si es una importación total o parcial.

La solución permitirá tener un seguimiento de importaciones realizadas por proveedor, según la línea de producto. Se permitirá conocer a que proveedores alternativos contactar para agilizar el proceso de compra y la consecución de la mercadería de compra.

La solución permitirá filtrar por indicadores o parámetros respectivos que serán utilizados por los analistas de negocios para la toma de decisiones dentro de cada importación. Los factores a utilizar como tiempo, costo, garantía del proveedor, competitividad del sector, competitividad del producto, integración con el proveedor (acuerdo, pacto o trato a nivel de integración vertical) están definidos dentro del filtro para el análisis personalizado del área de compras y de almacén.

La solución permitirá al personal de logística mostrar los días en que el producto no es vendido, clasificado por línea de producto. Los días transcurridos entre la última venta del producto y la fecha actual, esta rotación ayudará a un mejor análisis de las compras regulares a realizarse. Este reporte permitirá decidir cuales productos comprar nuevamente y que productos impulsar para incrementar las ventas.

La solución permitirá visualizar el índice de disponibilidad de los almacenes frente a las órdenes y estadísticas de movimientos de los productos en las áreas de ventas.

La solución permitirá tras el análisis de información de la solución crear nuevos acuerdos de niveles de servicio con la empresa encargada de las operaciones de importaciones, esto para optimizar los tiempos de reacción frente a las necesidades de producto y poder contar con la data de movimientos, por ejemplo, en menos días.

La solución se permitirá la reestructuración flexible de esta solución, tanto en la fase de extracción como en la fase de explotación de los datos, según cambios futuros en los requerimientos de los usuarios para la solución.

La solución permitirá explotar la información de diversas fuentes de origen de datos, en este caso se explotará desde una base de datos que está destinada para la fuente de datos del ERP SAP Business One en su versión 8.8.

Fundada en 1972 como Systems Applications and Production Data Processing (Sistemas, Aplicaciones y Productos para Procesamiento de Datos), SAP es reconocida como líder mundial en soluciones de colaboración de negocios para todos los sectores de la economía y mercados verticales. Atendiendo a más de 41.200 clientes en todo el mundo, SAP es la mayor empresa de software de gestión empresarial y la tercera mayor proveedora independiente de software en la clasificación mundial. SAP tiene una notable historia de innovación y crecimiento que la convirtió en una verdadera líder de mercado. Actualmente, SAP emplea a más de 41.900 empleados en más de 50 países. Sus profesionales están determinados a ofrecer el más alto nivel de servicio y soporte para sus clientes.

SAP ha aprovechado su vasta experiencia para ofrecer un amplio conjunto de soluciones para reforzar cada aspecto de las operaciones corporativas. Al usar las soluciones SAP, las organizaciones de todos los tamaños pueden reducir costos, mejorar el desempeño y ganar agilidad para reaccionar ante las necesidades de negocios en transformación.

Teniendo estas consideraciones se sustenta la solución propuesta para una empresa comercializadora importadora independientemente de la ubicación geográfica en donde se encuentre.

CAPÍTULO 2: ANÁLISIS

En este capítulo se describe inicialmente la metodología a utilizar para poner en práctica el proyecto de tesis. También detalla los requerimientos y las necesidades de los usuarios de las diferentes áreas de la organización. Finalmente muestra el análisis de la solución para orientar el desarrollo en la fase de construcción.

2.1. Metodología aplicada para el desarrollo de la solución

La metodología a utilizar para el proyecto tiene como base el ciclo de vida de los data warehouses elaborado por Ralph Kimball. El esquema presentado por Kimball con el nombre de BDL (*Business Dimensional Lifecycle* o ciclo dimensional del negocio) muestra las diferentes etapas por la que debe pasar todo proyecto de data warehousing. Se ilustra la secuencialidad de tareas de alto nivel requeridas para el efecto del diseño, desarrollo e implementación de un proyecto de inteligencia de negocios [PEN 007]

Los recursos y el tiempo dentro de los proyectos de data warehousing son diferentes para cada etapa. El BDL se enfoca en la secuencialidad y en la concurrencia de los términos de los procesos y no en los plazos.

A continuación se expone un resumen de lo que comprende cada una de las etapas en el presente proyecto de tesis, incluyendo los artefactos correspondientes:

2.1.1. Planificación

La planificación del proyecto implica todos los procesos involucrados para tener una visión previa de la empresa y el desarrollo del anteproyecto para saber cómo afrontar un proyecto de data warehouse: dentro de este modelo se identifican las siguientes fases:

- Diagnóstico de la empresa.
- Plan de proyecto.

Después de dichas fases se tiene como entregables finales los siguientes:

- Documento de Misión y Visión.
- Estudio de Factibilidad.
- Documento de toma de requerimientos.
- Plan de proyecto.
- Cronograma de actividades y seguimiento.

Esta fase se encuentra documentada en el anexo A - Plan de Proyecto.

2.1.2. Análisis

Dentro del análisis se definen los requerimientos involucrados y la definición de la arquitectura adecuada para la solución. La arquitectura en donde se enmarcará la configuración de toda la solución debe plantearse y justificarse según las prioridades evaluadas en el análisis de factibilidad y los requerimientos funcionales y no funcionales planteados.

Se divide en estas fases y entregables:

- Definición de Requerimientos.
- Definición de la arquitectura.

Después de dicha fase se tiene como entregables finales los siguientes:

- Documento de Análisis
- Documento de Arquitectura

2.1.3. Diseño

En esta fase se tiene que definir la estructura a alto nivel para la preparación de la construcción de los diversos módulos y paquetes a realizarse en el proyecto.

Se divide en las siguientes fases:

- Diseño físico.
- Diseño dimensional.
- Diseño de ETL y reportes.

Entre el diseño físico y el diseño lógico o dimensional hay una retroalimentación que se realizó conjuntamente con las áreas de importaciones y el área de sistemas para la determinación de la arquitectura necesaria y factible, además de las medidas y métricas necesarias, ya que algunas de las decisiones que se tomen durante el diseño físico pueden afectar a la estructura del esquema lógico y viceversa.

En esta fase también se diseñan los reportes analíticos según información tomada en la etapa de la definición de requerimientos. Estos reportes se caracterizan por facilitar la obtención de información en forma rápida y oportuna llegando incluso si se quisiera a un mayor detalle (esquema *drill down*).

Después de dicha fase se tiene como entregables finales los siguientes:

- Documento de la metadata.
- Documento de modelamiento dimensional de base de datos.

Esta fase se encuentra documentada en el anexo B - Modelo de Base de Datos del Data Warehouse.

2.1.4. Construcción

La calidad de los datos es un factor determinante en el éxito del proyecto de data warehousing. Es en esta etapa donde deben sanearse todos los inconvenientes relacionados con la calidad de los datos fuente.

Según Ralph Kimball dentro de la construcción se tienen que presentar los siguientes artefactos, se presenta el mismo orientado al trabajo del presente proyecto:

- Configuración e Instalación del software: En esta etapa se configura e instala el motor de base de datos Microsoft SQL Server 2008 y así como también la plataforma de Integration Services 2008.

- La etapa de construcción de ETL consta de tres sub-etapas principales: extracción, transformación y carga de datos (procesos ETL).
- Extracción: Durante esta sub etapa se siguen los procesos necesarios para obtener los datos que permiten efectuar la carga del modelo físico.
- Transformación: En esta sub etapa se siguen los procesos para convertir los datos fuente a fin de calcular las métricas y mantener un formato estándar de los datos.
- Carga: Durante la carga de datos se siguen los procesos necesarios para poblar el datamart.

Después de dicha fase se tiene como entregables finales los siguientes:

- Documento de Diseño (Extracción y Explotación).
- Documento de Interfaz Gráfica.

Esta fase se encuentra documentada en el anexo C – Documento de Extracción y en el anexo D – Documento de Explotación.

2.1.5. Pruebas

Luego de la construcción de la ETL y los reportes se pone en práctica el plan de pruebas que permite identificar los errores originados en las etapas anteriores. Las pruebas son realizadas utilizando diferentes parámetros de entrada. Las pruebas toman como factores los tiempos de respuesta para la ejecución de cada reporte o gráfico, también se comprueba la consistencia de la data y los indicadores claves de desempeño (*KPI - Key Performance Indicators*) de acuerdo a los requerimientos funcionales presentado para el presente proyecto de tesis.

Después de dicha fase se tiene como entregables finales los siguientes:

- Documento de Pruebas.

2.1.6. Implantación

La implantación representa el contar con el producto final instalado en las sucursales de la empresa. Existen varios factores extras que aseguran el correcto funcionamiento del producto, entre ellos se encuentran la capacitación a los usuarios finales, el soporte técnico, la comunicación y las

estrategias de feedback. Todas estas tareas deben ser consideradas antes de que se de uso a la solución.

Según lo indicado anteriormente para el presente proyecto sólo se está cubriendo las siguientes áreas de conocimiento del PMBOK con sus respectivos artefactos o entregables:

Gestión del tiempo del proyecto

- Plan de gestión del tiempo
- EDT del Proyecto
- Cronograma de Hitos
- Actas de Reunión
- Reportes de Avance
- Checklist de Cierre.

Gestión del alcance del proyecto

- Plan de gestión de alcance.
- Definición del alcance

Gestión de calidad del proyecto

- Project Charter
- Plan de gestión de calidad
- Plantilla de Métrica de calidad
- Línea base de Calidad

2.2. Análisis de la situación actual

Para comprender mejor los procesos que se realizan en el área de importaciones de una empresa comercializadora e importadora de bienes en la parte operativa y en la parte de consecución de información para el análisis para la toma de decisiones se realizaron diversas entrevistas con el área de importaciones de varias empresas de este tipo.

Las empresas importadoras en su mayoría realizan las siguientes acciones a nivel operativo:

- Registro de proveedores internacionales y los tipos de condiciones de negociación acordadas regularmente.
- Registro de operadores logísticos clasificados por el tipo de servicio brindado dentro de toda la logística de entrada.

- Registro de almacenes intermedios para el almacenamiento virtual de los productos que tienen almacenamientos temporales en el operador logístico o en la misma agencia de aduanas.
- Emisión de Órdenes de Compra con los respectivos proveedores elegidos.
- Emisión de Facturas anticipadas al exterior exentas del IGV para el costeo inicial de los productos.
- Según la negociación realizada se pueden registrar diversos gastos de importación como son el flete, el registro de los impuestos aduaneros (DUA-Documento Único Administrativo) y los gastos por estibamiento de la carga.
- De acuerdo a las condiciones de la mercadería por recibir nacionalización al 100% o una nacionalización parcial; para lo cual se registra los gastos en los depósitos aduaneros.
- Cálculo de indicadores o inductores de los gastos adicionales incurridos en los productos según historial realizado en los últimos años. Los gastos adicionales se analizan por facturas de importación gestionadas desde la gestión realizada para el acuerdo con el proveedor extranjero hasta la liberación del producto de aduanas y envío a nuestros almacenes.
- Costeo de los productos para el ingreso a la mercadería nacionalizada y lista para poder comercializarla.
- Adicionar costos a los productos a causa de entregas de documentos de flete, seguro, o algún adicional que se encuentre añadido en el resto de la importación, estos gastos adicionales pueden aumentar o disminuir el costo de los productos.

En la Figura 2.1 (elaborada por el tesista) se muestran los procesos ejecutados por el área de importaciones en conjunto con el área de almacén y contabilidad en las tareas de costeo.

Por otro lado las Figuras 2.2 y 2.3 ilustran a manera de ejemplo un grupo de las tareas ejecutadas dentro del trabajo elaborado en las importaciones de una empresa dentro del sistema SAP Business One desde la definición del proveedor extranjero hasta el costeo de los productos según el volumen.

Adicionalmente se tiene que para cubrir los objetivos y alcances del área de importaciones referente al análisis macro de los datos que se usualmente tiene la Gerencia de Compras de Importación se realizan una serie de tareas como requisito para la retroalimentación en la toma de sus decisiones. Estas son las siguientes:

- Captar información de licitaciones, ventas específicas y obtener información de pedidos de importación por recibir para abastecer almacenes para un periodo.

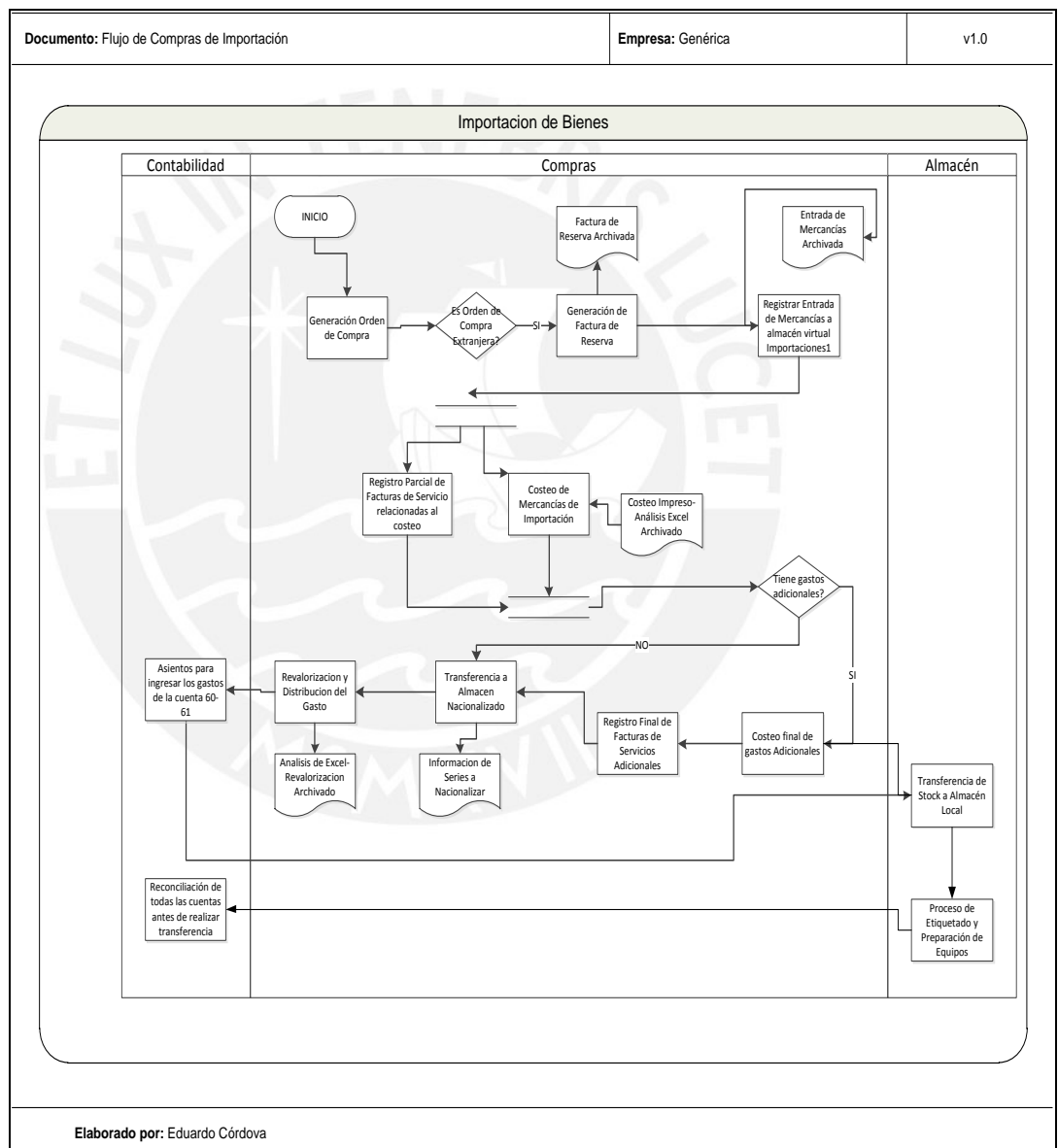


Figura 2.1 - Diagramas de Procesos de Importación. Fuente: [Elaboración Propia]

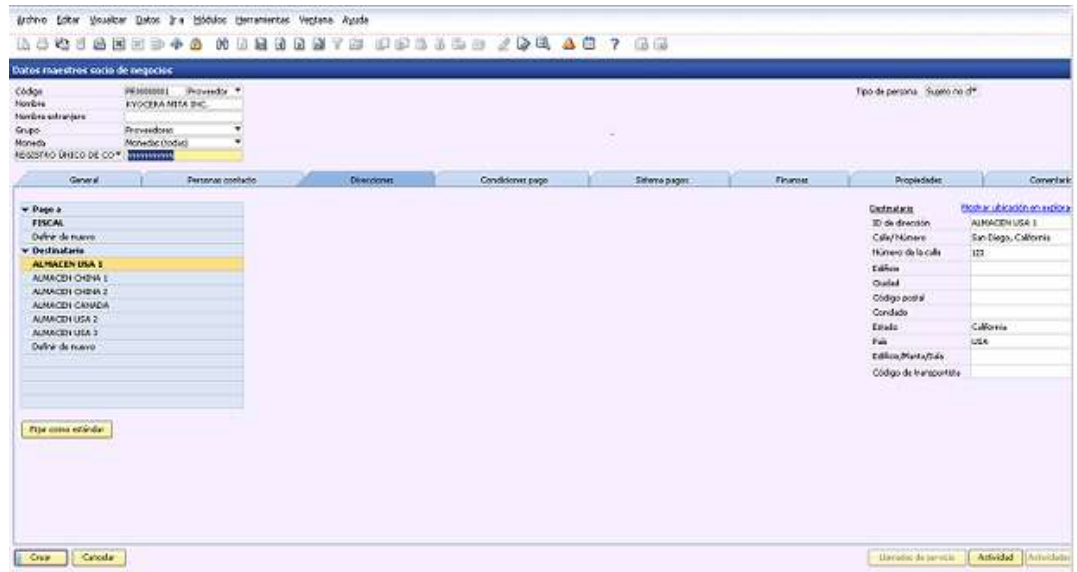
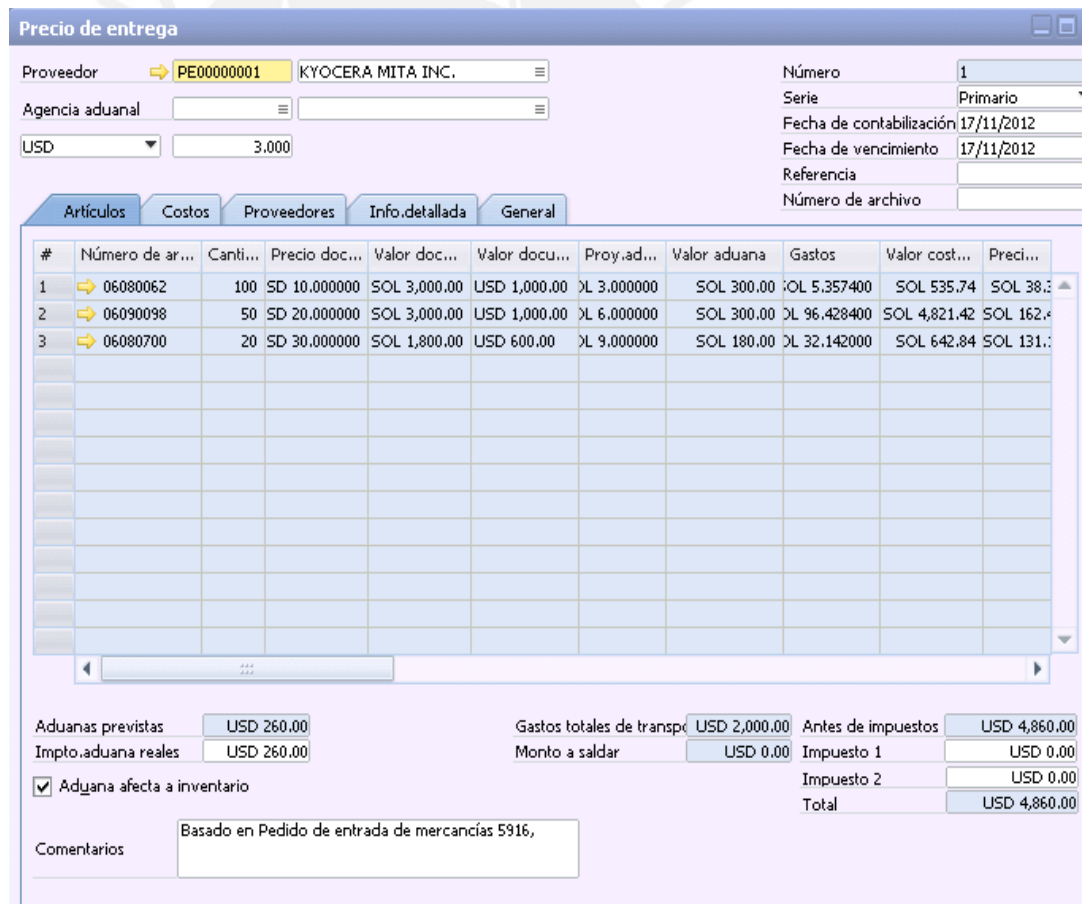


Figura 2.2 - Creación de Proveedor Internacional. Fuente: [WWW008]



#	Número de ar...	Canti...	Precio doc...	Valor doc...	Valor docu...	Proy.ad...	Valor aduana	Gastos	Valor cost...	Preci...
1	06080062	100	SD 10.000000	SOL 3,000.00	USD 1,000.00	DL 3.000000	SOL 300.00	DL 5.357400	SOL 535.74	SOL 38.2
2	06090098	50	SD 20.000000	SOL 3,000.00	USD 1,000.00	DL 6.000000	SOL 300.00	DL 96.428400	SOL 4,821.42	SOL 162.4
3	06080700	20	SD 30.000000	SOL 1,800.00	USD 600.00	DL 9.000000	SOL 180.00	DL 32.142000	SOL 642.84	SOL 131.:

Aduanas previstas: USD 260.00
 Impto. aduana reales: USD 260.00
 Aduana afecta a inventario
 Comentarios: Basado en Pedido de entrada de mercancías 5916,

Gastos totales de transporte: USD 2,000.00
 Monto a saldar: USD 0.00
 Antes de impuestos: USD 4,860.00
 Impuesto 1: USD 0.00
 Impuesto 2: USD 0.00
 Total: USD 4,860.00

Figura 2.3 - Proceso de Revalorización de artículos por Importación. Fuente: [WWW008]

- Recepcionar las diversas ofertas de los distintos proveedores o diferentes sucursales del mismo proveedor con los diversos niveles de tratamiento en el transporte de la mercadería.
- Aproximar los gastos adicionales involucrados de acuerdo a la gestión acordada, esta aproximación la plasman con inductores que se adicionan a los costos de importación proporcionalmente.
- Simular las diversas penalidades o posibles multas a realizarse en el proceso y analizar si la operación es rentable
- Generar el estimado de costos para los productos según el historial de gastos adicionales de importación involucrados.
- Certificar la capacidad instalada de los almacenes aduaneros para el volumen de los productos de importación; esto se conoce ya que previamente se tiene el conocimiento de estos datos dentro de la negociación.
- Certificar la calidad y el tiempo de entrega de los productos e informar el estado de los mismos para gestionar la devolución o la no consideración de la factura de los mismos.

2.3. Identificación de requerimientos

La definición de los requerimientos, tanto los funcionales como no funcionales fueron obtenidos a partir de la problemática planteada y reuniones con los usuarios finales en una empresa dedicada a la importación y comercialización de bienes.

Los requerimientos funcionales que se plantearon se presentan en el Cuadro 2.1.

Nro.	Descripción	Dif.	Prio	Ex
1	<p>Reporte de compras realizadas: El reporte permite elegir según la línea de producto y el periodo de tiempo a analizar. Este reporte muestra el proveedor y la forma de pago de una compra, así como la cantidad de productos comprados y el costo de estos. El reporte permite tomar decisiones sobre que proveedores elegir para la compra de productos y la mejor forma de pago. Además otorga un margen de negociación con el proveedor según las cantidades compradas en un periodo de tiempo para obtener tarifas rebajadas y formas de pago que beneficien a la empresa.</p>	2	1	E
2	<p>Reporte de Slow Moving: El reporte muestra los días en que el producto no es vendido. Se elige una línea de producto específico. El reporte muestra un indicador denominado Slow moving cuyo valor son los días transcurridos entre la última venta del producto y la fecha actual. El indicador es de color verde cuando la rotación del producto es favorable, de color amarillo cuando la rotación es aceptable y de color rojo cuando ha sobrepasado el tiempo máximo permitido. El reporte refleja una alarma de los productos que tienen bajas ventas y decidir si impulsarlos con campanas o dejar de incluirlos en compras futuras. Este reporte permite decidir qué productos comprar nuevamente y que productos impulsar para incrementar las ventas.</p>	2	1	E

Nro.	Descripción	Dif.	Prio	Ex
3	Reporte de stock por sucursal. Este reporte muestra la cantidad total de productos en una sucursal y cuantos existen en stock físico, cuantos están comprometidos, cuántos se ha solicitado y cuanto está disponible para ventas. Este tipo de reportes Permite la toma de decisiones referente a la mejor distribución por sucursal, así como conocer que tiendas tienen mayor demanda de productos. Permitir visualizar la disponibilidad de los almacenes frente a las órdenes y estadísticas de movimientos de los productos.	2	1	E
4	Reporte de compras a realizar de un producto: Este reporte muestra la cantidad de productos que se necesita para cubrir las ventas destinadas a 30 días. El reporte permite elegir la línea de producto y este muestra los productos, el stock, la cantidad de productos en exhibición y la venta a 30 días. Este reporte ayuda a decidir qué productos se debe comprar y en qué cantidad para cubrir la demanda de los mismos. Permitir la creación de escenarios tras la decisión de una forma de importación simple o múltiple para cubrir backorder(s) o cubrir el inventario de seguridad de la empresa.	2	1	D
5	Reporte de proveedores Se elige un periodo de tiempo para el análisis. El reporte permite conocer los proveedores con los que trabaja la empresa en el periodo de tiempo y encontrar con quienes se tienen las mayores compras. Este reporte permite una mejor negociación con el proveedor buscando los mejores precios de compra y el mejor margen de ganancia	2	1	D
6	Reporte de estados de compra. Este reporte muestra el estado de las compras (pedidas, atendidas y rechazadas). Permite tener un seguimiento de las compras realizadas por proveedor, según la línea de producto elegido. El reporte permite conocer a que proveedores contactar para pedir agilizar el proceso de compra. Este reporte también ayuda a tener control sobre el proceso de compra.	2	1	D

Nro.	Descripción	Dif.	Prio	Ex
7	Reporte de multas o retrasos. Este reporte muestra el monto de las multas pagadas por importación en un periodo determinado dentro de los proceso de compra realizados. Permite tener un seguimiento de la eficiencia del área y del proveedor.	2	1	E
8	Permitir la elección de múltiples formas de negociación (incoterms) dentro de una importación total o parcial dentro de la ejecución de reportes.	2	1	E
9	Permitir ingresar los indicadores o parámetros respectivos que serán utilizados por los analistas de negocios para la toma de decisiones dentro de cada importación.	2	1	E
10	El sistema muestra en todos los reportes como marca de agua el logo de la empresa de bienes	2	3	D
11	Los usuarios tienen la opción de imprimir los reportes	3	2	E
12	Se generan gráficos tipo columnas en base a los reportes	1	2	E
13	El sistema presenta dashboards para visualizar el estado de los indicadores , como los semáforos dentro del reporte de Slow moving	2	2	D
14	.Todos los reportes pueden ser exportados a otros archivos con formato: PDF, XLS, HTML y texto	2	2	E
15	El usuario define los campos y filtros que muestra el reporte	2	1	E

Cuadro 2.1 - Requerimientos Funcionales para el Proyecto

Los requerimientos no funcionales que se plantearon pueden verse en el Cuadro 2.2.

Para la explicación se presenta estos aspectos:

Dificultad	
Valores	Descripción
1	Alta
2	Media
3	Baja

Importancia/Prioridad	
Valores	Descripción
1	Alta
2	Media
3	Baja

Exigibilidad	
Valores	Descripción
E	Exigible
D	Deseable

Nro.	Descripción	Dif.	Prio	Ex
1	La solución es elaborada con licencia de Integration Services	3	1	E
2	La solución tiene como motor de base de datos Microsoft SQL Server 2008	2	1	E
3	La solución tiene para la elaboración de reportes la herramienta de Reportan Services de Microsoft	2	1	E
4	Es de fácil manejo y tiene una interfaz amigable, provee un sistema intuitivo y fácil de usar en la elaboración de reportes y análisis	1	2	D
5	El sistema tiene el soporte de 24x7	3	2	D

Cuadro 2.2- Requerimientos No Funcionales para el Proyecto

Los requerimientos funcionales y no funcionales indicados en el catálogo de requerimientos anterior fueron obtenidos principalmente por dos fuentes. La primera de ellas proviene la teoría ofrecidas por una consultora de soluciones informáticas (principalmente en soluciones de ERP y de inteligencia de negocios), en dicha investigación se deseaba obtener las necesidades y requerimientos para la explotación de la tablas origen del ERP SAP y explotar la información captada por los procesos de importación. La segunda fuente se encuentra basada en entrevistas realizadas personalmente a potenciales usuarios dentro de una empresa comercializadora e importadora de impresoras, fotocopiadoras y repuestos de las mismas. Se les preguntó acerca de las necesidades que

tienen y las limitaciones con respecto a las funcionalidades del ERP existente enfocándose en los análisis y reportes que presenta el sistema utilizaban actualmente. En esta investigación personal se entrevistó a tres tipos de usuario: a un usuario de análisis de costos de importación, a un usuario de almacén y a un representante de la gerencia de ventas.

2.4. Análisis de la solución

En esta sección se identifican las necesidades principales de los usuarios, las cuales serán cubiertas en su totalidad por la solución a desarrollar en el presente proyecto. Luego se presenta el análisis de la viabilidad del proyecto, un análisis técnico y económico del mismo, y finalmente se detallan diagramas de reportes preliminares para el análisis de la información.

2.4.1. Requerimiento de información de los usuarios

El principal beneficiario de la solución es el personal del área de importaciones, en particular se encuentra el analista de costos y de compras de importaciones ya que realiza el análisis principal de la información de mercadería importada y de los pedidos de importación.

Se debe tener en cuenta que los usuarios de esta solución deberán tener un conocimiento intermedio-avanzado de computación ya que la forma de trabajo actual para poder resolver dichos análisis se realiza por medio del uso de librerías dinámicas de herramientas de ofimática. Se debe considerar también que dichos usuarios desempeñarán sus labores sólo dentro de este tipo de empresas, independientemente de la zona geográfica en donde se encuentre. Dentro de los principales objetivos de los usuarios para mejorar sus procesos de análisis de información se tienen las siguientes:

- Evaluar la eficiencia de los proveedores internacionales y operadores logísticos
- Evaluar los tiempos de respuesta de la logística interna
- Mejorar el costeo y realizar un eficaz presupuesto o proyección de costos.
- Aminorar los gastos adicionales involucrados en las importaciones
- Identificar las falencias administrativas para tener una mercadería inmovilizada dentro de las operaciones internacionales.

La solución que se plantea consiste en desarrollar un datamart de compras y almacén que integre dichas áreas para que sirva como apoyo para la toma de decisiones dentro del área de importaciones. Este proyecto nace para que pueda ser implantado en cualquier empresa que realicen importaciones de bienes.

Los indicadores que se plantean para dar encuentro a los requerimientos planteados se presentan en los Cuadros 2.3, 2.4 y 2.5:

Indicador	Cálculo de Indicador
Plazo de aprovisionamiento	Fecha de emisión del pedido - Fecha de envío de mercadería
Ciclo de vida de Pedidos de Compra	Fecha de registro de OC - Fecha de entrega de pedido en aduanas
Error de previsión de demanda	Previsión de demanda-Demanda Real/Demanda Real*100
Cumplimiento de plazos-Entregas a tiempo	Numero de pedidos en el plazo previsto/Número total de pedidos realizados * 100
Contador de entrega a tiempo	Valor de pedidos entregados en el plazo previsto/Valor total de pedidos entregados
Suministro de inventario mensual	Inventario disponible/Promedio de uso mensual de almacén intermedio *100

Cuadro 2.3 - Indicadores para el Aprovisionamiento de Compras

Indicador	Indicador
Costo de transporte medio unitario	Costo total de transporte/Número total de mercadería comprada * 100
Mix de carga	Volumen por modo de transporte(avión-tren-barco)/Volumen total expedido * 100
Utilización del transporte	Costo total de transporte por modo (avión-tren-barco) /Kilogramos totales movidos por modo (avión-tren-barco) *100
Costo de transporte por tipo de operación	Costo de transporte internacional/Costo total de la operación*100
Rotación de inventarios	Costo de Ventas/Costo total de inventario * 100

Cuadro 2.4 - Indicadores para el análisis de Transporte

Indicador	Cálculo de Indicador
Participación de Operadores externos	Gastos por operador logístico/Compras dentro de un periodo*100
Tiempo de mercadería inmovilizada	Nro. de días en el almacén aduanero / Ciclo de vía de Pedidos de Compra * 100
Gastos adicionales según incoterm	Gasto Flete, Seguro, Documento Aduana anual/Costo de compras internacionales * 100
Recepción de productos defectuosos	Nro. de Productos defectuosos/Total de productos recibidos*100
Tasa de manejo de inventario	(Gastos Adicionales y Administrativos (Almacenamiento, Manejo, Obsolescencia, Defectuosos, G. Administrativos, Pérdidas ocasionadas)/Costo de Inventario Promedio *100)+ Costo de Oportunidad (%)+Seguro(%)+ Flete(%)

Cuadro 2.5 - Indicadores para el análisis del Operador Logístico

La solución de inteligencia de negocios propuesta será realizada bajo una arquitectura heterogénea basándonos en la realidad de las grandes compañías que presentan múltiples bases de datos.

2.4.2. Análisis Técnico y económico del proyecto

La solución de inteligencia de negocios que se implementará considera los siguientes procesos principales:

- Analizar los datos a limpiar, revisar y depurar para lograr la calidad de los datos y el análisis respectivo del almacén de los metadatos.
- Revisar la infraestructura de red existente y plantear soluciones para tener una arquitectura robusta sin pérdidas de conexión hacia las distintas bases de datos
- Verificar y mapear de la base de datos transaccional y la base de datos OLAP.
- Construir, transformar y la cargar todo el datamart representativo hacia la base de datos de explotación. En estos procesos se hará el cálculo de las medidas, incluidas en las tablas principales, y que posteriormente nos servirán en los datamarts.

- Crear y modificar las dimensiones. Asimismo, crear las relaciones, jerarquías y niveles correspondientes.
- Identificar las consultas lógicas y físicas hacia la base de datos requeridas para la creación de reportes y gráficos que solucionen los requisitos.
- Realizar el diseño para la adecuada visualización de los datos, de acuerdo a las especificaciones de los usuarios, a través de los reportes que contendrán los principales indicadores de gestión y que ayudarán en la toma de decisiones

La percepción del costo de una solución BI es comúnmente denominado con el TCO (*Total Cost of Ownership*, Costo Total de Propiedad) al cual se le atribuyen ciertos factores que participan en el cálculo del mismo como:

- Software:
Que incluyen componentes como el sistema operativo, la plataforma de inteligencia de negocios y las herramientas integradas necesarias.
- Servicios:
En el que se encuentra la instalación y el mantenimiento del hardware, la implementación y el desarrollo del software como las licencias correspondientes, la capacitación a los usuarios y la consultoría en procesos de negocios.
- Hardware:
Dentro de este aspecto se encuentran los servidores utilizados, la conectividad subyacente, respaldos de equipos cliente y servidores; como también el mantenimiento correspondiente.

Cada uno de estos rubros pueden tener muchos más elementos que lo componen y se puede apreciar una aproximación en el Cuadro 2.6, asimismo existen rubros de costo escondidos como la cultura organizacional, el rechazo al cambio, el aprovechamiento correcto de soluciones tecnológicas que son difíciles de medir y que pueden dejar fuera de la realidad dicha interpretación.

Teniendo analizado los requerimientos reales de la empresa y los procesos dentro de cada área (especialmente en el área de importaciones) se podrá aprovechar al máximo la información de gestión relativa para la toma de decisiones en todas las áreas involucradas en un proceso de importación.

Rol	Costo x hora USD
Analista	20
Responsable de Calidad	40
Arquitecto	60
DBA	50
Programador	50
Jefe de Proyecto	80

Actividades	Esfuerzo (horas)	Responsable	Horas asignada	Costo (USD)
Planificación y Captura de Datos	40	Analista	40	800
Modelamiento del Proyecto	15	Analista	15	300
Análisis de la Arquitectura	130	Analista	20	400
		Arquitecto	50	3000
		DBA	60	3000
Implementación Data Warehouse	150	Programador	120	6000
Extracción		Analista	20	400
		Responsable de Calidad	10	400
Implementación Data Warehouse	180	Programador	150	6000
Transformación		Analista	20	1000
		Responsable de Calidad	10	500
Pruebas	40	Programador	20	1000
		Responsable de Calidad	20	800
Implantación	40	Responsable de Calidad	40	1600
Gestión y Coordinación	50	Jefe de Proyecto	50	4000
TOTAL HORAS	645		TOTAL USD	\$ 29,200.00

Motivo	Costo (USD)
Licencias de Motor BD y Desarrollo	1000
Servidores para el motor y el DW	3500
Material Informático	400
Servicio de Luz	120
Servicio Internet	200
Servicio Teléfono y Celular	250
Gastos diversos	2600
Materiales de estudio	200
Total USD	8270

Cuadro 2.6 - Presupuesto de Costos para el proyecto

CAPÍTULO 3: DISEÑO

Este capítulo describe la arquitectura que conformará la solución, se revisarán tanto el diseño de los datamarts como los procesos de extracción y de explotación de la solución de inteligencia de negocios para el área de importaciones. Además se ilustrará el diseño de explotación de datos en base a reportes configurados según lo establecido en la toma de requerimientos.

3.1. Arquitectura de la solución

La arquitectura de la presente solución de inteligencia de negocios presenta los componentes necesarios para implementar una solución de este tipo y las conexiones entre los diferentes puntos de los mismos.

3.1.1. Modelado de la Arquitectura-Diseño de la Base de Datos

Aunque las soluciones basadas en inteligencia de negocios tienen más de una década dentro del mercado; recién ahora se está dando más énfasis e importancia el uso de este tipo de aplicaciones por lo que la mayoría de empresas procura manejar una infraestructura y arquitectura acorde a la dimensión creciente de la solución. Dada la necesidad de tener la mayoría de información automatizada y almacenada por medios informáticos hacen que las arquitecturas sean más robustas y de acuerdo a estándares para que los datos y análisis respectivos sean efectuados de una manera eficiente. Es por

ello que se realiza la estandarización para adaptar una solución de este tipo hacia base de datos OLAP brindando consultas multidimensionales.

La arquitectura básica comprende los componentes de extracción y explotación de datos necesario para la construcción de toda la solución de inteligencia de negocios para el área de importaciones de una empresa comercializadora.

En la Figura 3.1 (elaborada por el alumno) se presentan los componentes en concreto de la arquitectura lógica para la correcta explotación de los datos.

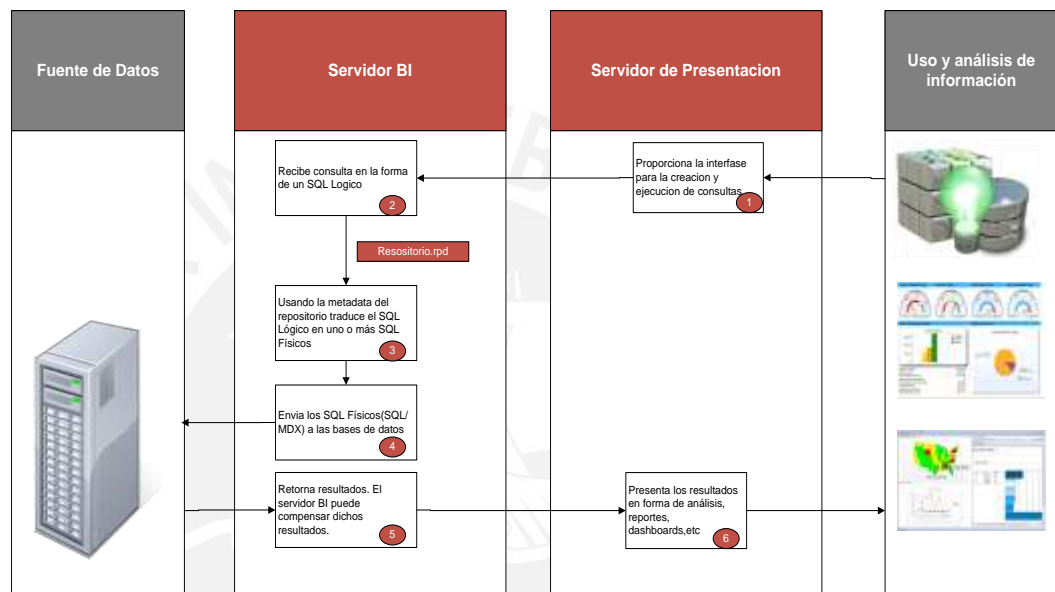


Figura 3.1 - Esquema de Arquitectura. Fuente [Elaboración Propia]

La presente sección describe cada componente necesario para la elaboración del esquema de arquitectura.

3.1.2. Fuentes de Información

Los datos son extraídos desde base de datos relacionales OLTP ubicados en el motor de base de datos SQL Server 2008. Los sistemas transaccionales o ERP's, en este caso SAP Business One 8.8, servirán de fuente original de los requerimientos del usuario en sus procesos diarios. Este sistema principal de la organización contiene las bases de datos operacionales fuente para la construcción de la base de datos OLAP y la información a los datamarts. El diagrama de clases de alguna de las tablas de la base

de datos fuente en el módulo de compras se muestra en la Figura 3.2.

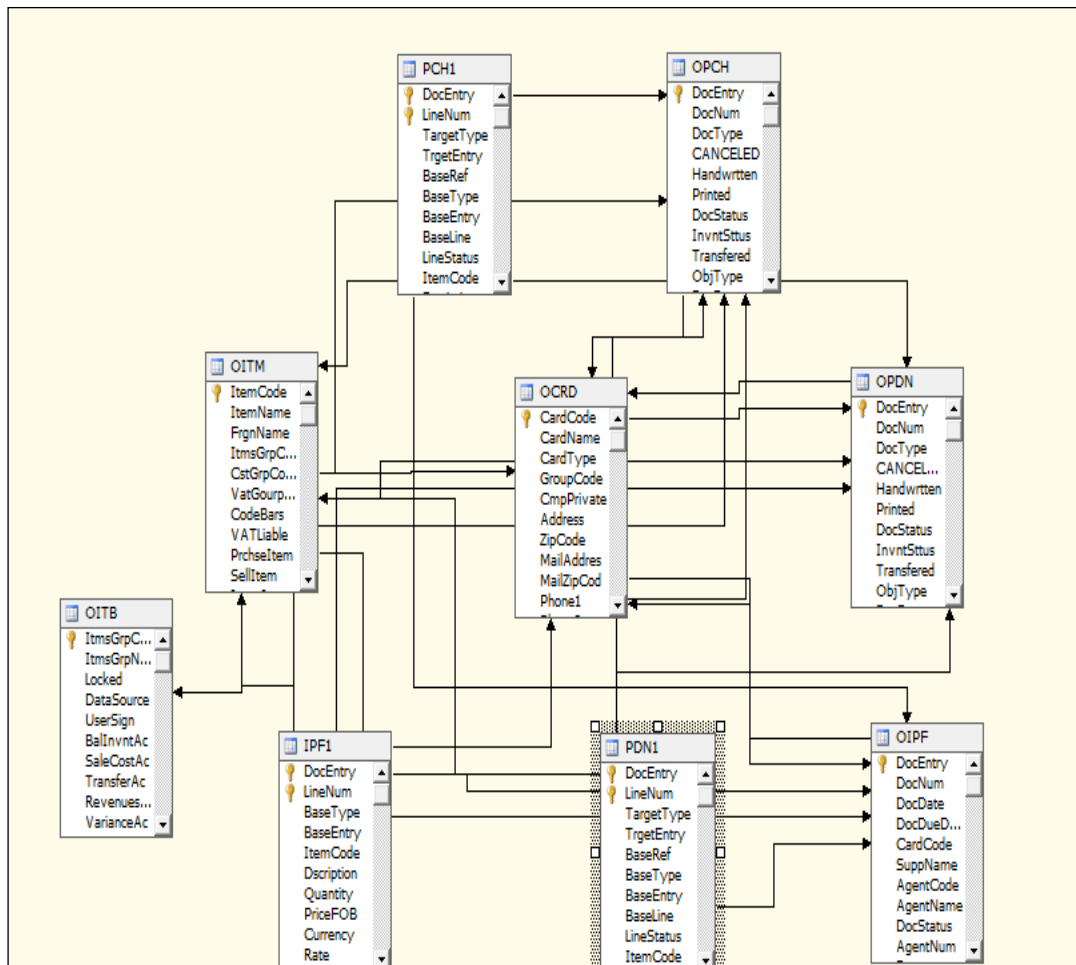


Figura 3.2 - Diagrama Relacional de Base de Datos SAP

Para ello, se ha establecido cargar el modelo fuente con el sistema SAP Business One en modo prueba con una licencia temporal de 30 días; allí se realizó el trabajo de ingreso de datos simulando la operatividad diaria del área de importaciones según como se viene trabajando, utilizando los distintos módulos del ERP SAP Business One en su versión 8.8. La razón de ser de este proceso es tener información confiable y que los resultados sean verificados y aceptados por los usuarios, en resumen, que reflejen con veracidad la información de la operatividad del área.

3.1.3. Procesos ETL

El proceso ETL se encarga básicamente del paso de información de un modelo de base de datos relacional a un modelo multidimensional. El proceso extrae la información que está almacenada en el modelo relacional, adicionando información de diferentes orígenes (archivos de texto, backup de sistemas heredados, etc), realiza la transformación de los tipos de datos de los campos correspondientes y carga de datos de las tablas de dimensiones y hechos hacia las bases de datos multidimensionales.

- Extracción: Hace referencia a la captura de información desde los sistemas operacionales, se necesita de la integración.
- Transformación: Hace referencia a la adaptación de los datos fuente al formato destino en el datamart como agregar datos numéricos, transponer información, componer textos a partir de otros, etc.
- Carga: Hace referencia al proceso en que los nuevos datos son finalmente almacenados en el datamart en su formato definitivo.

El primer paso para el diseño de la ETL es realizar el mapeo de la base de datos fuente con la base de datos del Datamart. De esta manera se establece la relación de campos migrados base en ambas bases de datos.

Luego de realizar el mapeo respectivo de todas las tablas fuentes se realizan las transformaciones adecuadas que sufrirá cada campo de las tablas para el eficiente ingreso de datos.

3.1.4. Datamart

Este componente está formado por la entidad de base de datos multidimensional. A nivel físico, estas tablas se agrupan en tablas de dimensiones y tablas de hechos. Aparte de permitir una mejor organización, permite agregar de manera más simple las futuras tablas a medida que crezca la base de datos.

Los servicios OLAP por medio de los datamarts permiten consolidar los datos de los cubos multidimensionales para su mejor análisis. Se empleará Utilizando la herramienta Microsoft Integration Services (MS SQL Server 2008) se logra la extracción, la transformación y la carga correspondiente de los datos fuente iniciales.

En esta base de datos se procesan las consultas MDX ligeras o consultas lógicas que retornan resultados multidimensionales adicionando las medidas necesarias para la correcta evaluación dentro de la toma de decisiones. Se puede dentro de estas bases optimizar el manejo de cache para las consultas más solicitadas por los usuarios de la solución de inteligencia de negocios y apoyar a la construcción de agregados para un mejorar rendimiento en la carga.

3.1.5. Servidor de Presentación

Es importante reconocer que conforman la parte final de la explotación de información el Servidor OLAP, el Servidor de Presentación y la interfaz en donde se encontrarán los terminales conectados del usuario.

Desde donde el usuario puede explotar la información completa del datamart. Estos entornos se componen de reglas de negocio, consultas asistidas, clusters de informes y perfiles por usuario.

Además dentro del diseño se tiene que considerar los siguiente elementos físicos de la arquitectura: Planeación de la capacidad, diagramación de la red para el uso compartido, automatización de tareas y aislamiento, elementos configurables, administración y recomendaciones de respaldo.

La conexión a la base de datos necesarias para la solución de inteligencia de negocios para el área de importaciones necesita la conexión hacia la base de datos origen (ERP SAP) y las diversas fuentes de tipo texto existentes almacenadas en carpetas compartidas tienen una interconexión con las bases de datos intermedias.

La manipulación de los datos, al mismo tiempo, se interrelacionan con la base de datos donde están alojados los datamart.

Se debe monitorear además la conexión eficiente del Servidor de Presentación con las máquinas cliente para que se realice la explotación de los mismos.

Todo esto para cumplir con los requerimientos no funcionales del área de importaciones las cuales se basa en los criterios de eficiencia y de confianza en la data manipulada.

3.2. Diseño de la solución

3.2.1. Justificación de la arquitectura OLAP

La fundamentación para la elección de explotar la información sobre esta arquitectura radica en los beneficios que ofrece como el de soportar análisis complejos de datos, permitir analizar datos desde diferentes perspectivas y soportar análisis contra un volumen ingente de datos los cuales irán creciente al paso de tiempo según la operatividad de la empresa.

Las consultas hacia la base de datos se realiza accediendo a un menor número de tablas lo cual optimiza el tiempo de respuesta. La clave primaria de una tabla de hechos está formada por las llaves primarias de las dimensiones y sólo existirá una tabla de dimensiones por cada dimensión. Se decide utilizar este modelo ya que está dirigido a bases con una gran cantidad de datos independiente del tamaño de la empresa, esto optimizará el tiempo de respuesta de los reportes.

El diseño de bases de datos *On-Line Analytical Processing* (OLAP) presenta diferencias fundamentales respecto de los principios de diseño de las bases de datos *On-Line Transaction Processing* (OLTP). El Cuadro 3.1 (elaborado por el alumno) muestra las principales características de ambos tipos de almacenamiento de datos.

OLTP (Transaccional)	OLAP (Análisis)
Se definen entidades normalizadas	Se definen entidades desnormalizadas
Se define una normalización hasta la tercera forma normal	
Diseños complejos de bases de datos	Define diseños sencillos de base de datos fáciles de entender por los usuarios
Almacena información al menor nivel de detalle transaccional	Almacena datos a nivel transaccional y resumizados
Incrementa el numero de joins para la realización de consultas	Reduce el número de joins (consulta a nivel lógico) hacia las consultas a la base de datos OLAP
Normalmente mantiene información estática	Normalmente dinámico

Cuadro 3.1 - Diferencias Técnicas entre sistemas OLAP y OLTP

El asistente para el diseño de cubos basado en la herramienta SQL Integration Services ayudará a la creación de los datamarts facilitando la definición de las dimensiones, los niveles de jerarquía, la tabla de hechos y la conexión a la base de datos relacional en un fichero XML.

La información proporcionada por un sistema OLAP debe cumplir las siguientes características:

- Presentación en un formato intuitivo y fácil de usar para el usuario.
- Alta performance para acceder a búsquedas complejas que involucren grandes cantidades de información.
- Modelo multidimensional

Las diferencias técnicas de diseño entre los sistemas OLTP y los sistemas OLAP hacen que se concluya la elección por la segunda opción como alternativa para la arquitectura.

Se concluye en esta sección por estas justificaciones también la utilización de la arquitectura ROLAP (modelo estrella) dentro de solución de inteligencia de negocios para el área de importaciones en una empresa comercializadora/importadora y para ello se utiliza como herramienta el motor ROLAP de la plataforma SQL Analysis Services 2008. Las expresiones recompiladas permiten optimizar los reportes y consultas, tratando de evaluar la expresión en memoria.

Asimismo, se ha optado por esta arquitectura, pues está compuesta por un servidor relacional que almacena los datos detallados y agregados además de un motor OLAP en otro servidor especializado que almacena la estructura de los cubos de información.

3.2.2. Elementos para el diseño del modelamiento

Dimensión (Lookup table): Se define dimensión o tabla de dimensión a los criterios por el cual se va a cruzar la información con la tabla de hechos. Las dimensiones contienen las descripciones de las entidades principales del negocio, respecto de las cuales se calcularán las medidas o métricas dentro de la tabla de hechos. Por ejemplo el tiempo, el cual está compuesto por días, meses, años semestres, etc. dependiendo el nivel de granularidad o detalle a donde se quiere llegar.

Las características en una tabla de dimensión son:

- Tienen una relación uno a muchos con la tabla de hechos (Fact Tables).
- Incluyen una clave primaria, de preferencia numérica.

3.2.2.1. Hecho

Es un elemento de la fact table. Representa un evento específico, por ejemplo se puede mencionar las compras realizadas por un proveedor extranjero con una modalidad de importación en cierta fecha.

3.2.2.2. Medida

Este elemento es quizá el más importante, pues es propio de la realización de un datamart. Son los valores cuantitativos que almacenan métricas del negocio. Están representados por columnas numéricas en la fact table representadas por columnas derivadas de la carga o datos que pueden derivar de campos calculados de un conjunto de columnas.

3.2.2.3. Tabla de Hechos (Fact Table)

Es la tabla central que representa datos numéricos en el contexto de las entidades del negocio.

Esta tabla de hechos tiene las siguientes características:

- Posee una clave primaria compuesta por los campos que representan sus relaciones con las tablas de dimensión.
- Poseen columnas numéricas para las medidas o métricas.

3.2.3. Lineamientos para el diseño lógico dentro del modelamiento

Para diseñar un modelo lógico primero se deben definir los campos de las tablas con sus tipos de datos idóneos para el correcto funcionamiento. Todo campo a emplear en el modelo del datamart se encuentra en grupos de tipos de datos: como: números enteros, números decimales, tipo fecha y cadena de caracteres.

3.2.3.1. Números Enteros

Se empleo el tipo de dato INTEGER para los identificadores de las tablas tanto de dimensiones como de tablas de hechos exceptuando el tipo fecha que se ejecuto con el tipo DATE. Esto debido a que las búsquedas y consultas son más eficientes y rápidas cuando la llave primaria es un número entero.

3.2.3.2. Números decimales

Se empleó el tipo de dato NUMERIC (10,2) para los valores de los montos de compras, gastos de importación y flete incurridos en las importaciones del negocio.

3.2.3.3. Fecha

Se empleo el tipo de dato DATE para los valores de las fechas. No se empleo el tipo de dato TIMESTAMP por no ser compatible con la versión de Oracle a usar en la migración del datamart hacia el servidor de presentación.

3.2.3.4. Cadena de Caracteres

Se empleó el tipo de datos VARCHAR para los valores de texto de longitud variable que son de uso descriptivo para los códigos de las tablas.

3.2.4. Diseño de la extracción

El diseño de extracción de datos comprende la carga de los datamarts desde las tablas transaccionales fuente, el mapeo respectivo y la descripción de las tablas destino del datamart.

El esquema de extracción permite conocer el orden en que realiza la carga de las dimensiones y las tablas de hechos.

Proceso de Extracción, Transformación y Carga:

El diseño del proceso ETL dentro de un proyecto de inteligencia de negocios sirve básicamente para mapear correctamente los datos a poblar en el datamart.

El proceso que se realiza para la extracción, transformación y carga de tablas tomando tomará las siguientes consideraciones:

- La herramienta SQL Integration Services que se incluye dentro del software de Microsoft SQL Server 2008 será el que se usará para los procesos ETL.
- Existirá un paquete por cada grupo de tablas a migrar; por ello existirán paquete para tablas de Hechos y paquete para tabla de Dimensiones.
- Dentro de cada paquete, a través de las diversas tareas se podrá acceder a la base de datos haciendo uso de queries o stored procedures.

- Se realiza la transformación de datos fuente como cambios de tipo de dato, truncamiento de datos, cambios de formato para la presentación; para estandarizar, permitir el mapeo adecuado y realizar la limpieza de los datos finales hacia los datamarts.

El esquema lógico de diseño de la tabla importación con sus respectivas dimensiones es mostrado en la Figura 3.3 (elaborada por el alumno).

Además se presentará en el Cuadro 3.2 el mapeo de cómo se cargará el modelo dimensional destino resultante de la carga ETL para la tabla de hechos de Importación.

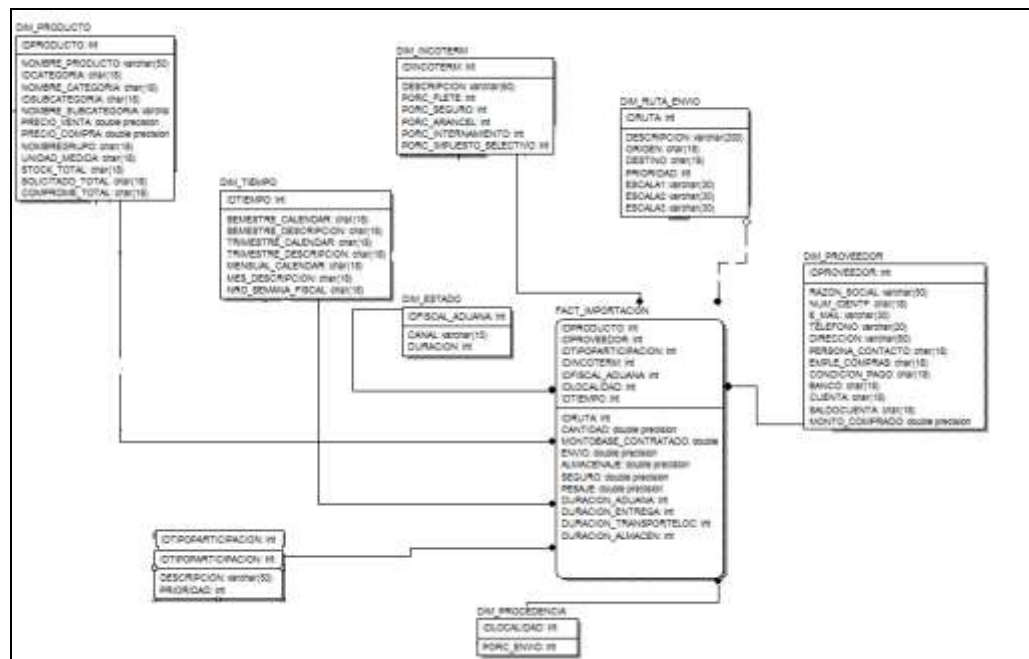


Figura 3.3 - Esquema Lógico del Modelo Estrella de la Tabla FACT_IMPORTACION

N°	Nombre de la Dimensión	Descripción	Llave Primaria
1.	Dim_Tiempo	Representa el momento en el tiempo en que se dio la compra de importación	IdTiempo

2.	Dim_Proveedor	Representa al proveedor que recibe la solicitud de mercancías.	IdProveedor
3.	Dim_TIPOPARTICIPACION	Representa el tipo de participación dentro del proceso	IdTipoParticipacion
4.	Dim_Condicion	Representa la especialidad o condición (incoterm) a la que corresponde el servicio.	IdCondicion
5.	Dim_ATENCIONES	Representa a la atención registrada.	IdAtención
6.	Dim_Procedencia	Representa la zonificación del origen de la entrega de productos	IdLocalidad

Medidas

N°	Nombre	Descripción
1.	Duracion_Aduana	Duración en un determinado canal
2.	Costo de envío	Costo de envío total de mercadería en soles
3.	Costo_Almacenaje	Costo de almacenaje total de mercadería en soles
4.	Compra_total	Compra total en soles

Cuadro 3.2- Campos de Tabla de Hechos Importación

La tabla de hechos de operador logístico analiza el comportamiento de los servicios de distribución logística de terceros. El modelo dimensional destino resultante de la carga ETL para esta tabla de hechos es mostrado en la Figura 3.4 (elaborada por el alumno).

La descripción de los campos de dichas tablas multidimensionales y sus respectivas medidas se muestran en el Cuadro 3.3.

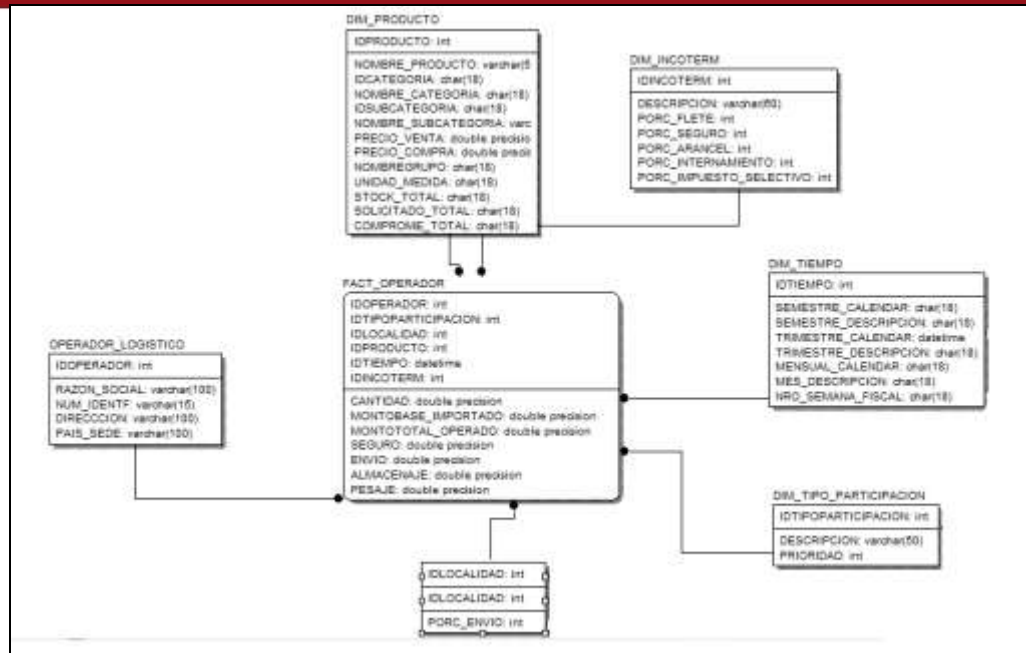


Figura 3.4 - Esquema Lógico del Modelo Estrella de la Tabla FACT_OPERATOR

Nº	Nombre de la Dimensión	Descripción	Llave Primaria
1.	Dim_Tiempo	Representa el momento en el tiempo en que se dio la atención.	IdTiempo
2.	Dim_PROGRAMACION	Representa la programación de la atención de importación.	IdProgramacion
3.	Dim_TIPOPARTICIPACION	Representa el Tipo de participación dentro de la operación	idTipoParticipacion
4.	Dim_OPERADORLOGISTICO	Representa al operador logístico que ejecutó parte del servicio.	IdOperador
5.	Dim_ESPECIALIDAD	Representa la especialidad a la que corresponde el servicio del operador logístico.	IdEspecialidad
6.	Dim_ATENCIONES	Representa a la atención registrada	IDAtencion

7.	Dim_ALMACEN	Representa la zonificación del origen de la entrega de productos	IDAlmacen
8.	Dim_PRODUCTO	Representa a la mercadería involucrada en la entrega de productos	IDProducto

Medidas

N°	Nombre	Descripción
1.	Duracion_Aduana	Duración en un determinado canal
2.	Costo de envio	Costo de envío total de mercadería en soles
3.	Costo_Almacenaje	Costo de almacenaje total de mercadería en soles
4.	Compra_total	Compra total en soles
5.	Numero_Atenciones	Volumen de Atenciones
6.	Tiempo_Atencion	Duración de atención del operador de acuerdo al tipo de participacion

Cuadro 3.3 - Campos de Tabla de Hechos FT_Operadores

Como parte del proyecto en el Anexo B – Modelo de la Base de Datos del Data Warehouse se detallará todas las dimensiones, medidas y tablas de hechos referenciados para el proyecto.

Dimensiones	Facts		
	FACT_IMPORTACION	FACT_OPERADO	FACT_ALMACEN
Dim_TIEMPO	✓	✓	✓
Dim_PROVEEDOR	✓		
Dim_ESTADOADUANA	✓		
Dim_OPERADORLOGISTICO	✓	✓	✓
Dim_TIPOPARTICIPACION	✓	✓	✓
Dim_PRODUCTO	✓	✓	✓
Dim_INCOTERM	✓		
Dim_PROCEDENCIA	✓	✓	✓
Dim_CONDICION	✓	✓	✓
Dim_ALMACEN		✓	✓

Cuadro 3.4 - Mapeo de Dimensiones y Tablas de Hechos

Para la ejecución de diagrama se desarrollaron los mapeos de data para las dimensiones como para la tabla de hechos; dicho mapeo de dichas tablas multidimensionales se muestra en el Cuadro 3.4.

3.3. Diseño de la explotación de datos

La explotación de datos presenta fases donde se recupera la información de los datamarts físicos en un primer lugar, y en una etapa posterior se configura el mapeo de las tablas para realizar consultas lógicas simples hacia el servidor de presentación y aligerar la ejecución de reportes, estas consultas lógicas configuradas hacia los datamarts se almacenan en carpetas para ser accedidos por los usuarios finales

El mapeo de tablas físicas se refleja también en el modelo lógico de la solución para que el motor entienda la existencia de un modelo multidimensional como se muestra en la Figura 3.5.

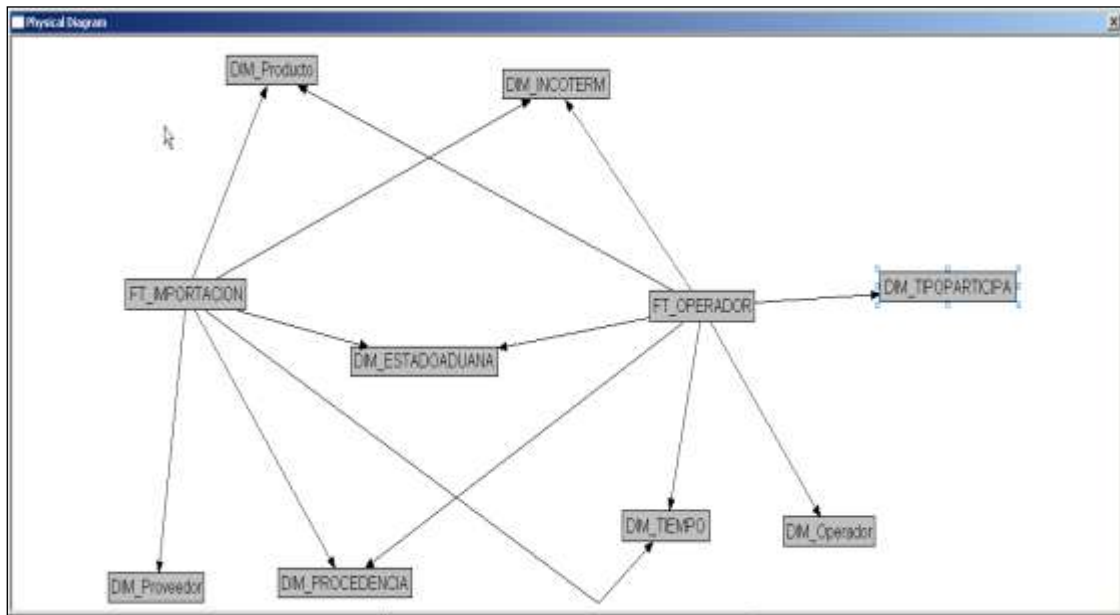


Figura 3.5 - Mapeo hacia las tablas físicas dentro de la fase de explotación

Para que estos análisis realmente tengan un rol protagónico en la toma de decisiones se añadirán métricas adicionales como la duración en aduanas, el costo de envío, el costo de almacenaje, la compra total y el costo total actual para el análisis comparativo o tasas.

Dentro de la parte de mapeos de las tablas se añade el diseño de la granularidad de los filtros de búsqueda de tablas importantes como las tablas de producto, tiempo. Estas tendrán una jerarquía de datos tal que ayude al análisis respectivo.

Como ejemplo se muestra en la Figura 3.6 se presenta el diseño de la jerarquía de la dimensión que almacena la información relacionada a los productos o mercadería importada que mantiene la empresa.

Nivel	Atributos
Nivel 1	IdCategoría
Nivel 2	IdSubcategoría
Nivel 3	IdProducto

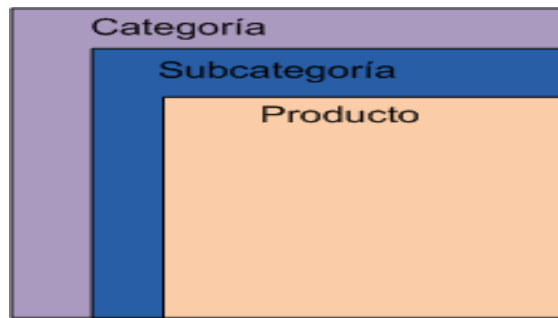


Figura 3.6- Jerarquía de la Tabla Producto

Para la construcción de reportes se define el uso extensivo de las dimensiones y realizando los cruces correspondientes para cada requerimiento solicitado. Para ello se toma la información contenida en el datamart creado tras el proceso ETL; se construye el repositorio necesario para la explotación de los datos. Luego de este desarrollo se ejecuta las pruebas y la publicación de los reportes.

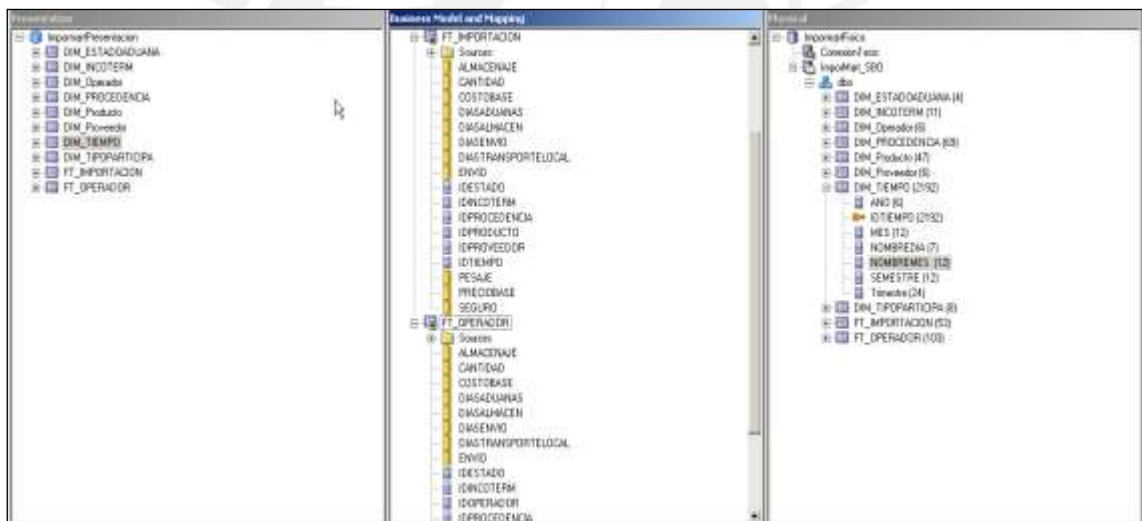


Figura 3.7 - Diseño de Explotación de Herramienta Oracle Warehouse Builder

Los análisis previamente documentados en la toma de requerimientos hacen posible que se tenga una idea de los inputs y los posibles outputs de cada reporte; pero principalmente es la manera como presentarlo en el diseño para que en verdad dicho reportes tengan un valor agregado para las tomas de decisión. Para ello se realizó el esquema ejemplo para ver los resultados esperados en lo que respecta al diseño de los reportes a presentar.

La definición de los primeros reportes se da en base a la mejora realizada sobre los que ya existen y a otros que, con la ayuda de retroalimentación con el área se va obtener y poder generar.

El diseño se basa en las propiedades gráficas de la aplicación, la data representada en forma numérica y su correspondiente objeto gráfico representativo. Adicionalmente; los campos como nombre del reporte, nombre de la empresa, se añadirán a los requerimientos dentro del layout.

A continuación, el Cuadro 3.3 muestra la definición de reportes ejemplo que se presentan de aquí en adelante.

Para ello se muestran ejemplos de los diseños requeridos para el análisis; en la Figura 3.8 se presenta el diseño para analizar el primer requerimiento del Cuadro 3.5.; este se logra explicar diseñando un pye o gráfica circular 2D el cual diferencia el porcentaje de compras de importación de cada proveedor internacional. Las Figuras 3.9 y 3.10 se presenta la ejecución del segundo y tercer reporte y en ellas se utiliza gráfico de barras; cada una de ella presenta filtro de tiempo (en este caso de forma anual). Se utiliza para ello la herramienta de explotación OWB (*Oracle Warehouse Builder*)

Nro	Nombre del Reporte	Descripción del Contenido	Información Contenido	Filtros Especiales
1	Reporte de proveedores internacionales	El reporte permite conocer los proveedores internacionales con los que trabaja la empresa en el periodo de tiempo y encontrar con quienes se tienen las mayores compras y analizar su relación con el costo. Este reporte permite una mejor negociación con el proveedor.	*Nombre de Proveedor *Monto Total Facturado *Monto Total Costo Base de los Productos	*Tipo de condicion internacional (Incoterm) *Períodos mensual, Trimestral, anual
2	Reporte de gastos referente a operadores	Este reporte muestra el monto de las gastos adicionales a la importación en un periodo determinado dentro de los proceso de compra realizados y la mercadería inmovilizada	*Condicion Asociada a los transportes *% de Multas sobre el Costo *Tiempo de mercadería inmovilizada promedio en aduana	* Todos Los Operadores Logísticos *Tipo de condicion internacional (Incoterm) *Tipos de estado o canal en aduana *Períodos mensual, Trimestral, anual
3	Análisis de condiciones internacionales	Este reporte muestra el monto de las gastos, compras y comparativo con el costo base por condicion internacional o incoterm acordado por periodos anuales	*Año *Incoterm *Costo Base *Monto Total *Gastos de envío *Gastos por almacenaje *Gastos por seguro	*Período anual *Incoterm acordado

Cuadro 3.5 - Reportes para el área de importaciones

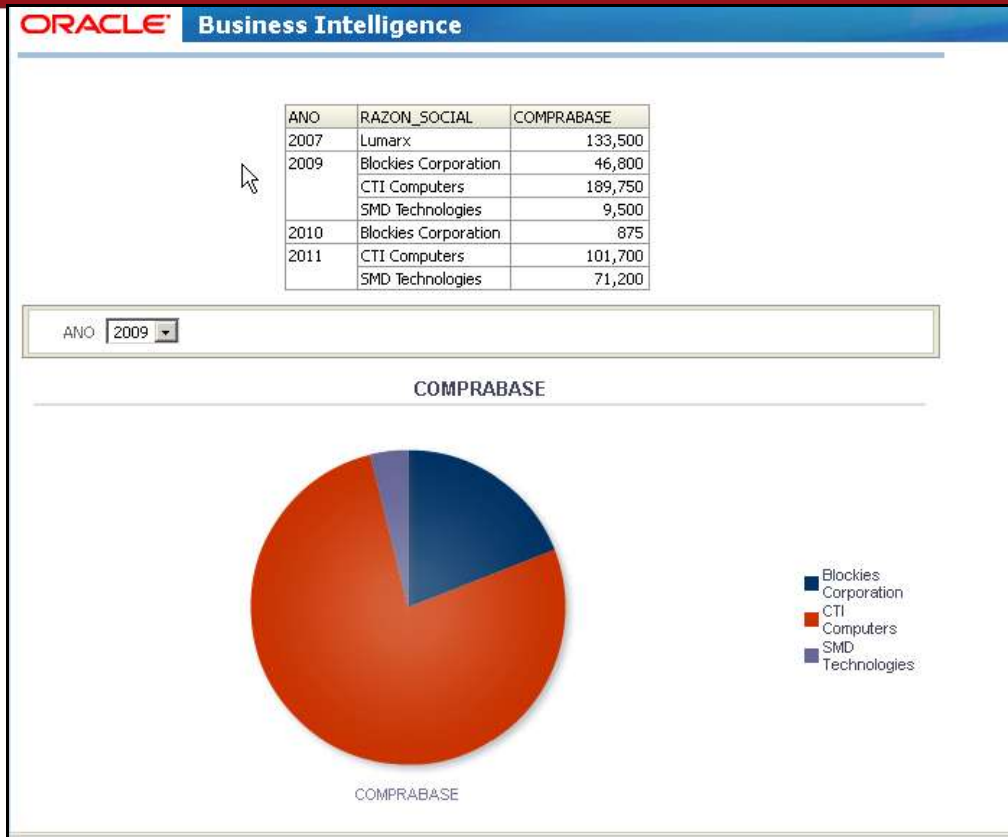


Figura 3.8 - Reporte de proveedores internacionales

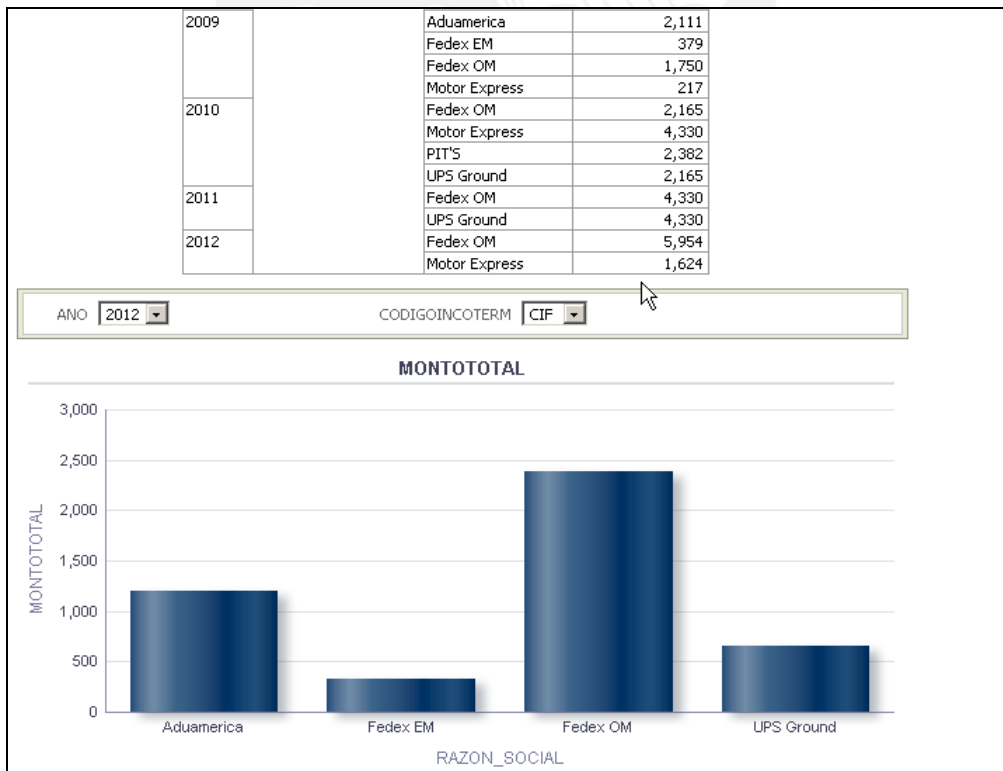


Figura 3.9 - Reporte de gastos referente a operadores

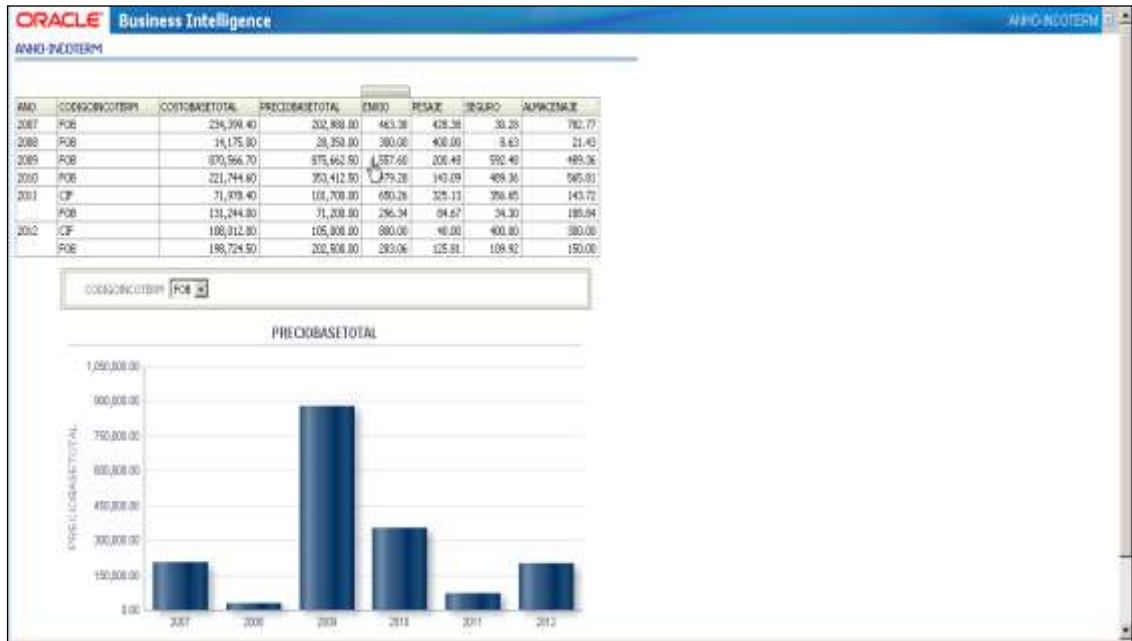
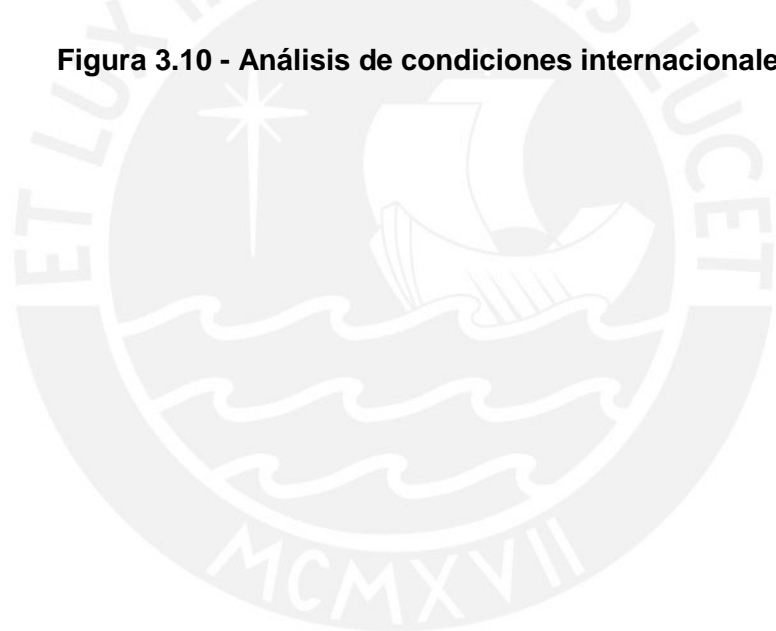


Figura 3.10 - Análisis de condiciones internacionales



CAPÍTULO 4: CONSTRUCCIÓN Y PRUEBAS

El presente capítulo describe la fase de construcción y pruebas de cada proceso, tanto la extracción y carga de los datamarts como la explotación de los datos con la visualización y tendencia respectivas. Se revisará la configuración de todo el software utilizado, construcción de procesos de carga y la ejecución de pruebas de proyecto.

4.1. Construcción

4.1.1. Construcción de la ETL

La fase de construcción del proyecto está definida principalmente por la implementación de los procesos ETL; se asume que la carga de datos fuente se realiza en una etapa inicial y que las tablas de entrada de los datamarts se encuentran creadas según el modelado previo en el capítulo anterior para la recepción de dichos datos; esto para que se ejecuten los procesos ETL de forma correcta. Para el desarrollo de la solución de inteligencia de negocios se realizaron paquetes de extracción que se dividen en extracción de tablas de dimensiones y en la extracción de las tablas de hechos.

Dado que se divide la carga de datos en tablas de hechos y de dimensiones se muestran los diagramas de cada uno de estos paquetes en las Figura 4.1.

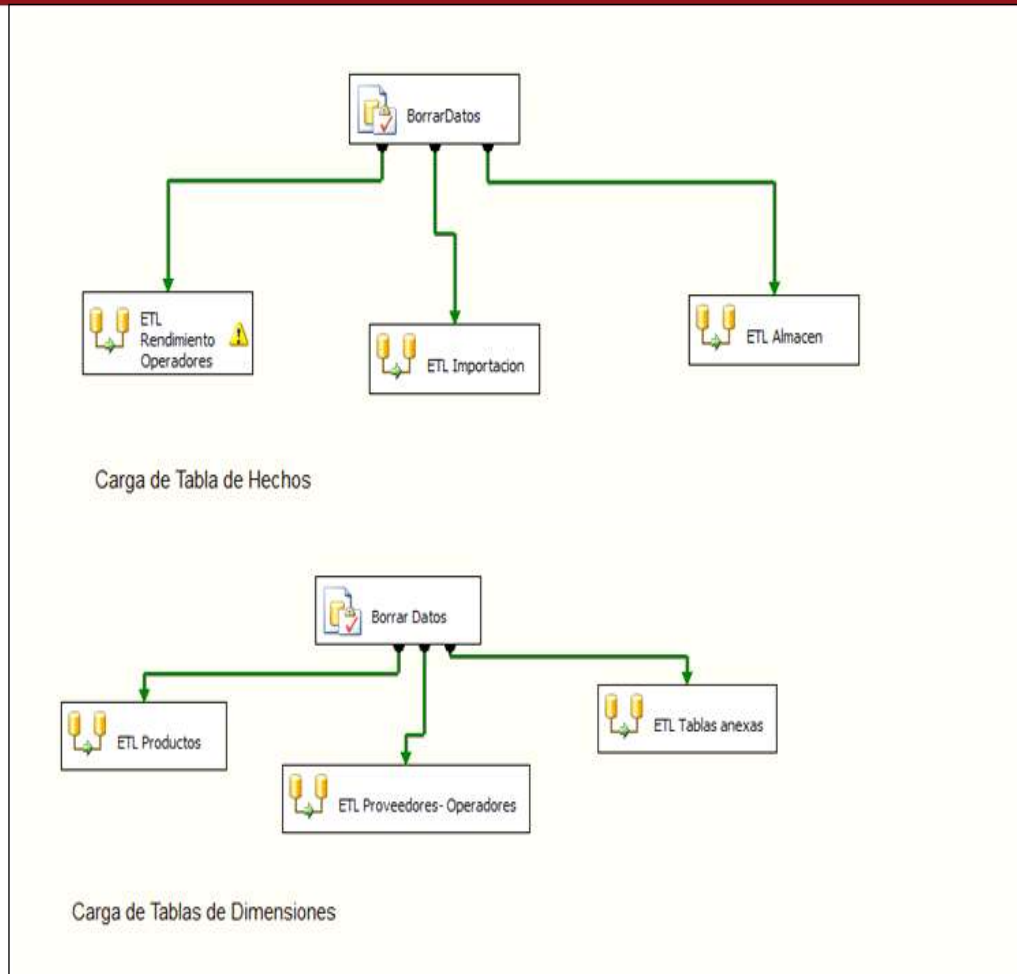


Figura 4.1 - Paquetes de extracción de Hechos y Dimensiones

Las transformaciones generadas se realizaron para estandarizar las consultas hacia el cubo. Dentro de estas transformaciones se encuentran: Estandarización de nombres, formato de campos, lenguaje o idioma, tamaño de campos permitido y campos pre-calculados. Estas operaciones se realizan en las distintas secuencias que pasa la tabla fuente, este flujo por el que pasa la información se diseña por medio de características ofrecidas por la herramienta de SQL Integration Services, el cual permite analizar el flujo de los datos y las distintas agregaciones y filtros por la que recorre la información antes de residir en las tablas destino.

El mapeo de los datos se valida en cada fase de los procesos gracias a las distintas características que ofrece la herramienta.

En el caso de cargar de la tabla Productos se desarrolla una limpieza de los productos que ya no se encuentran activos en la empresa en el momento de la

carga aunque tuvieron compras en años anteriores; esto se muestra en la Figura 4.2.

En caso de los proveedores dividimos la carga para la extracción de proveedores internacionales y los proveedores que otorgan servicios de operación logística. Para incorporar las tablas anexas se elaboraron sentencias de transformación directa sin emplear diseño desde el modelador; dicho proceso se muestra en la Figura 4.3.

El proceso de la Figura 4.4 corresponde al ETL que realiza la extracción, transformación y carga de todas las dimensiones anexas del datamart. En algunos casos la fuente de información es una tabla de base de datos y en uno es la carga de archivo de texto como en el caso de la tabla Tiempo.

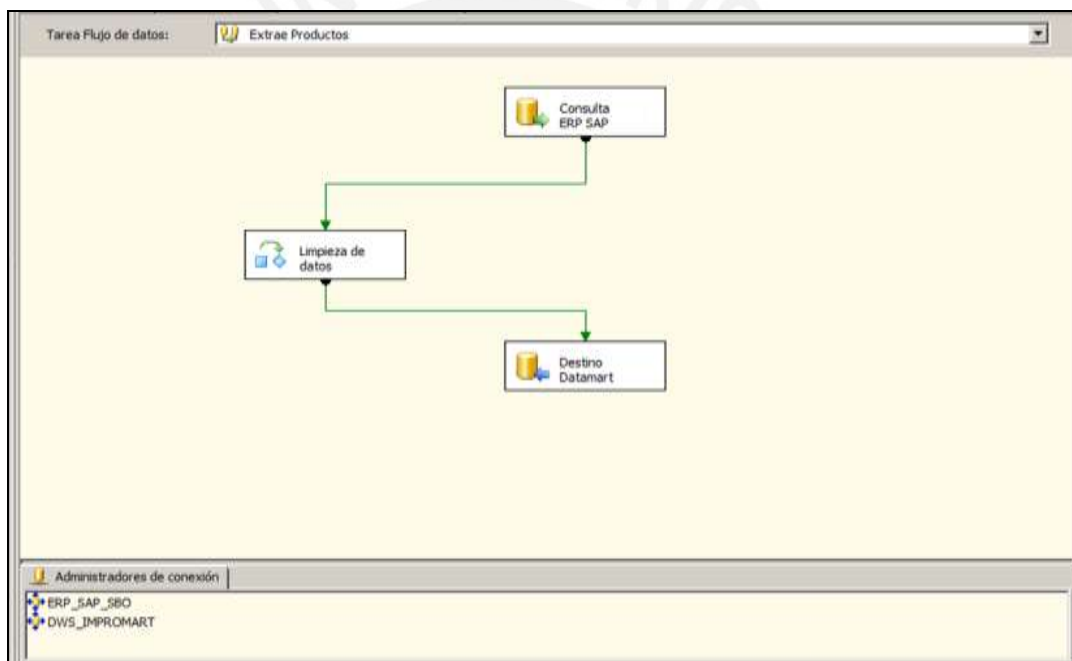


Figura 4.2 - Extracción y Carga de la dimensión Producto

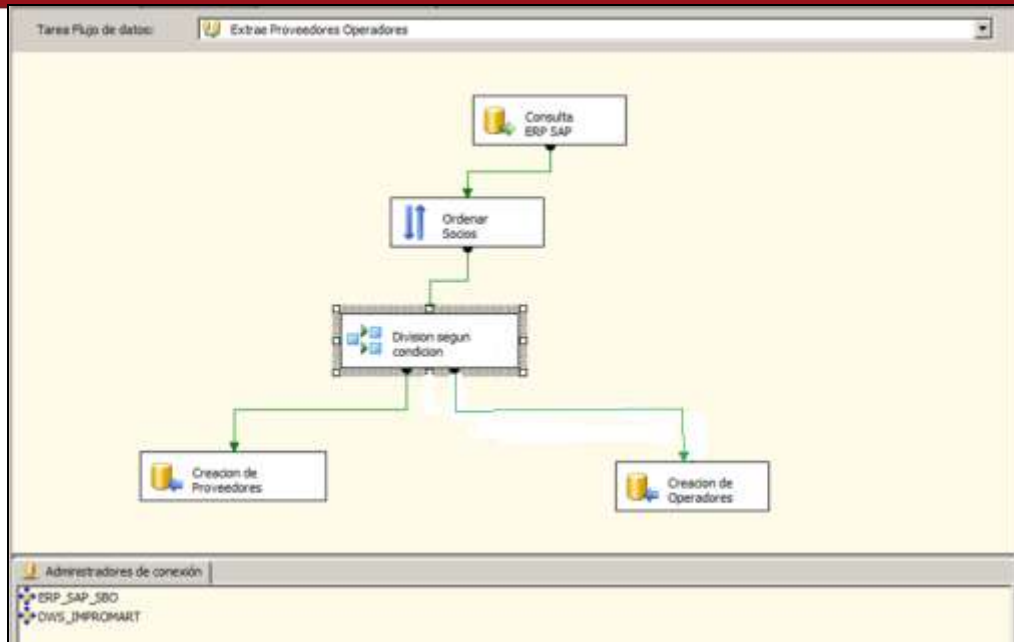


Figura 4.3 - Extracción y Carga de las dimensiones Proveedor y Operador

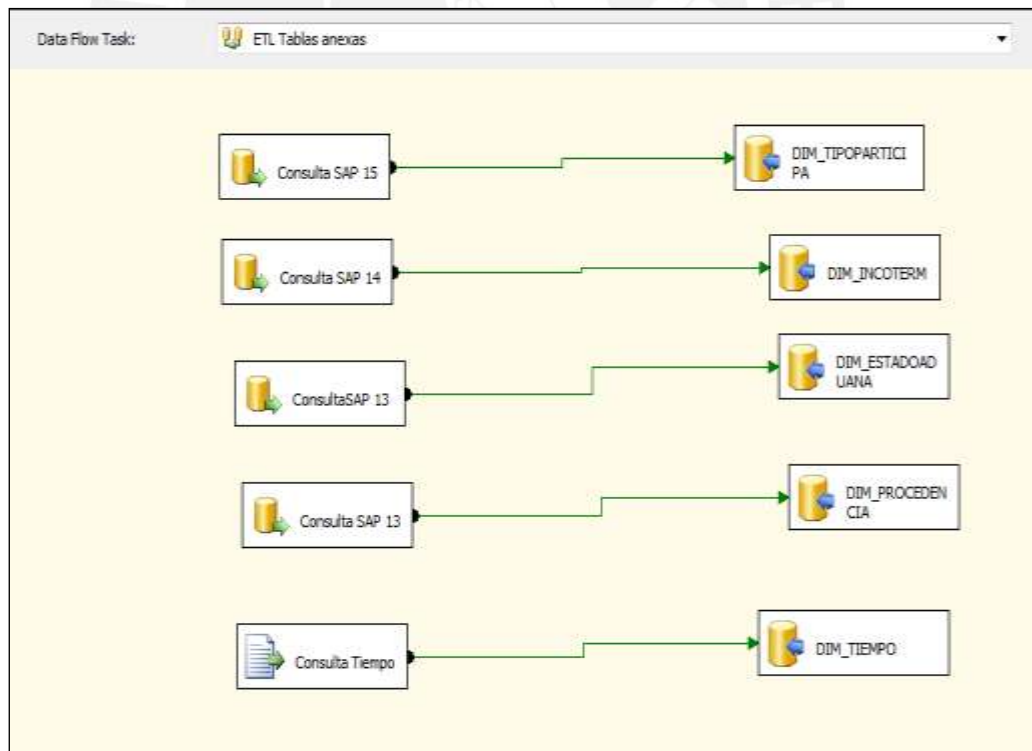


Figura 4.4 - Extracción y Carga de las dimensiones Anexas



Figura 4.5 - Consulta de Extracción de la tabla Producto

En la Figura 4.5 se muestra las consultas internas para las extracciones y transformaciones de la tabla de productos antes de la limpieza dentro del proceso ETL.

El mapeo de las tablas principales del proyecto son las tablas de hechos de Importaciones y Rendimiento de Operadores. El proceso de la Figura 4.6 corresponde al ETL que realiza la extracción, transformación y carga de la información de documentos de importación; cuando se realizan las extracciones iniciales y las transformaciones de datos se comienza a realizar búsquedas hacia otras tablas para conseguir información sumariada de gastos de importación y clasificar la información según la necesidad de las tablas destino. Extrae la información de la tabla de detalle de Importaciones y el detalle de se presenta los mapeos a dichas tablas de hechos. A la par se muestra el proceso ETL para cargar el datamart de Operadores en la Figura 4.7.

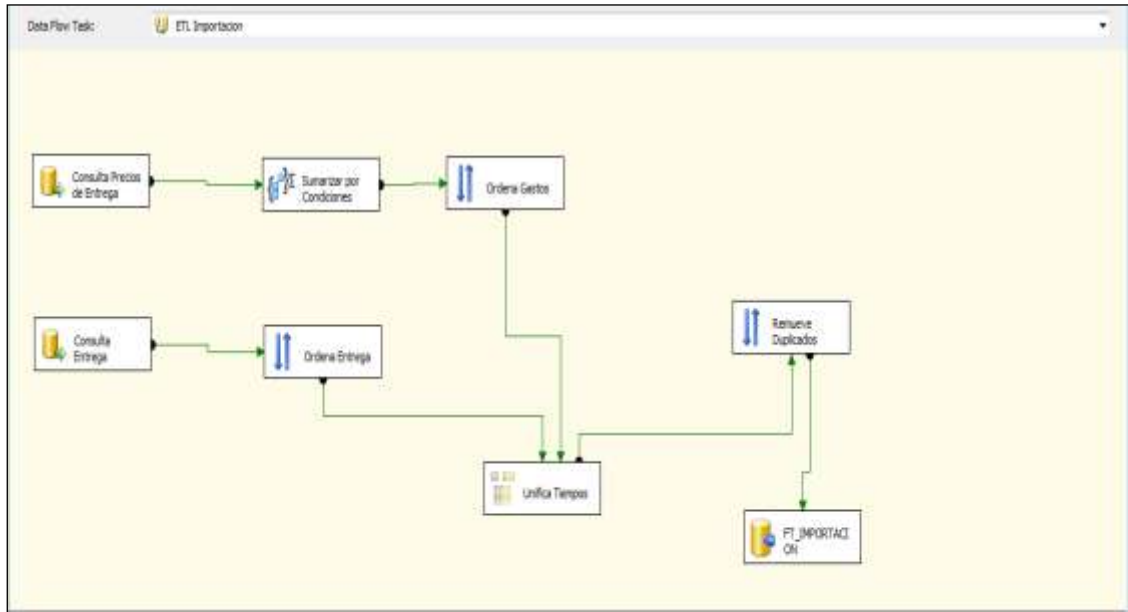


Figura 4.6 - Proceso de Extracción de la tabla FT_IMPORTACION

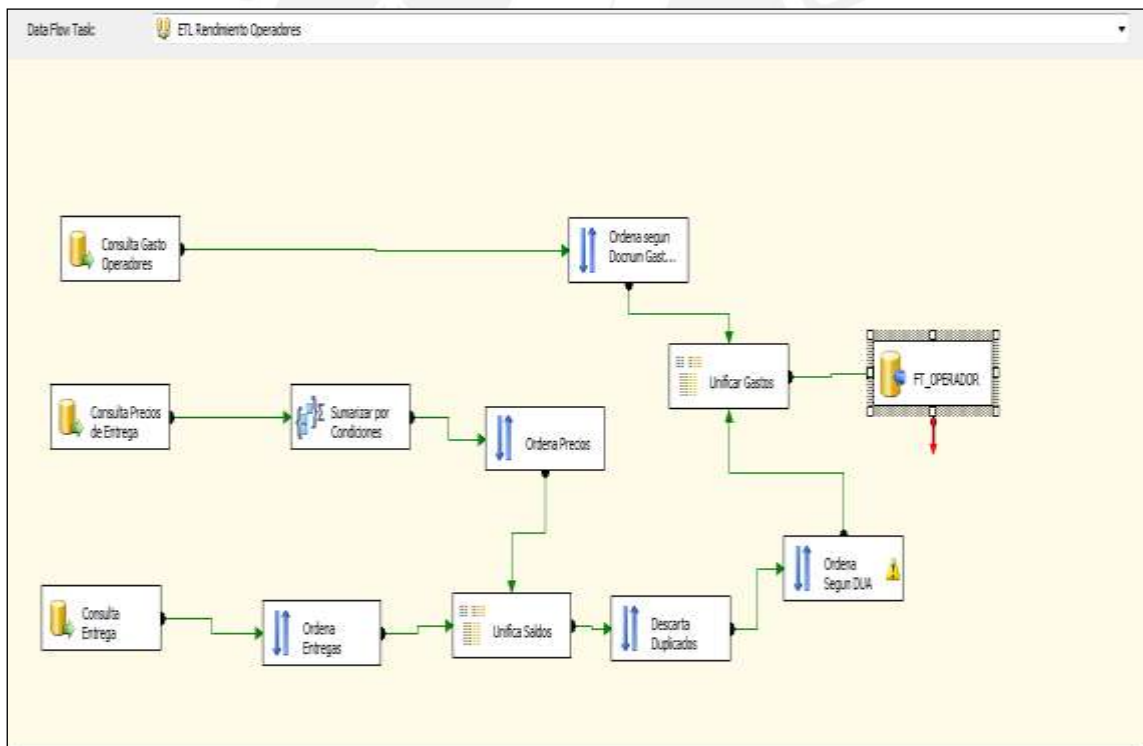


Figura 4.7 - Proceso de Extracción de la tabla FT_OPERADOR

Para la extracción, transformación y carga de datos se utiliza la tecnología ofrecida por el motor de base de datos SQL (SQL Analysis Services e Integration Services) y

para la fase de explotación de datos se utilizará la tecnología ofrecida por Oracle (Oracle Warehouse Builder). La experiencia bajo estas tecnologías representa una garantía para la viabilidad de la ejecución del proyecto.

Dentro de la fase de carga y transformación se envía la información a la base de datos de explotación de las tablas de dimensiones y de hechos que se hayan modificado, especificando el trabajo del agente de SQL SERVER para que se realice dicho trabajo automáticamente en horas programadas según la carga de trabajo del servidor origen y destino. En tal sentido, como parte del proyecto, en el Anexo C – Diseño de Extracción se detallará todo lo relacionado a la extracción.

4.1.2. Construcción del Servidor de Explotación

Ahora para la fase de explotación se utiliza la herramienta de ODBC para generar la conexión hacia el datamart fuente y poder crear el archivo de repositorio con extensión .rpd para generar las métricas y mapeo lógico entre las tablas. Dentro de la fase de presentación se divide dentro de la herramienta de Oracle 3 fases importantes: la fase física; desde donde se exportarán todas las tablas; la fase lógica donde se añadirán las métricas y jerarquías necesarias para la explotación. Estas fases se muestran en la Figura 4.8.

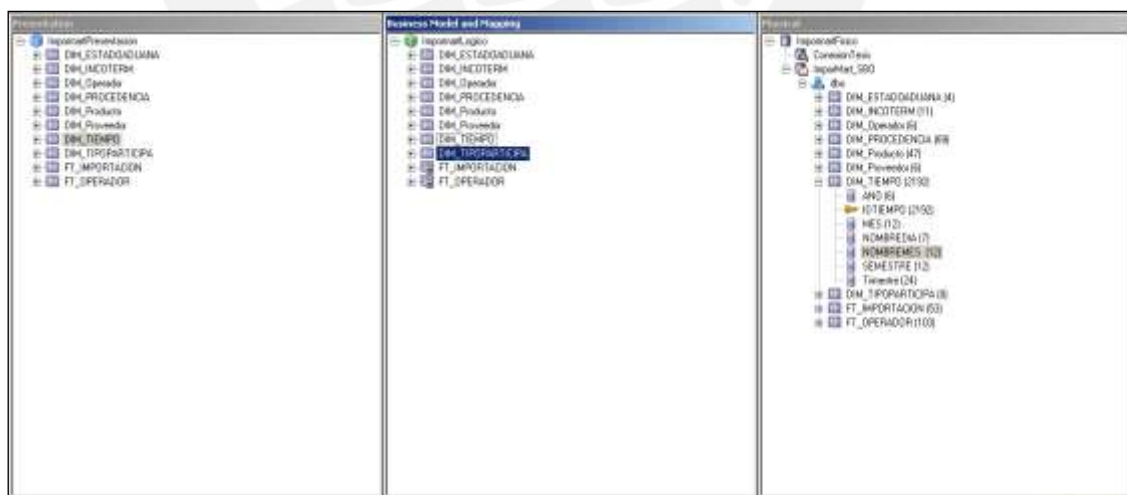


Figura 4.8 - Diseño Físico, Lógico y de Presentación

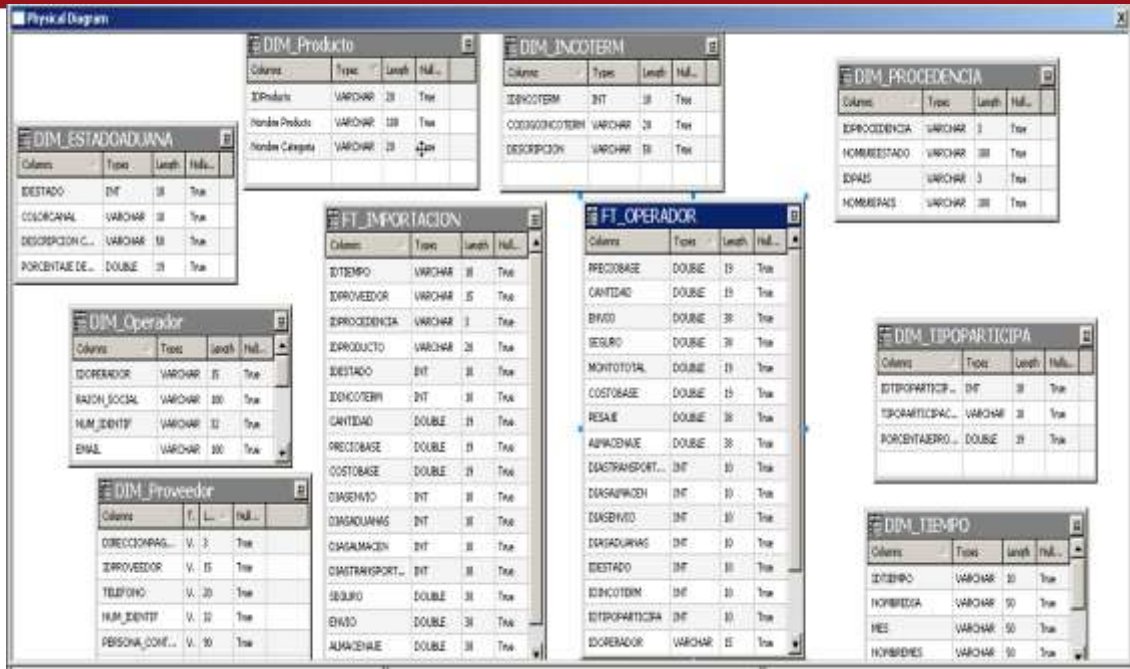


Figura 4.9 - Dimensiones y Tablas de Hechos en el Cubo

En este punto, se realizan las siguientes acciones:

- Configuración de las diferentes sentencias join que unen los códigos identificatorios de las tablas de dimensiones y los campos de identidad de las tablas de hechos. La configuración de las tablas físicas del datamart exportadas por la herramienta de explotación se muestran en la Figura 4.9.
- Resolver los campos calculados o kpis que se han definido para el proyecto (como campos sumariados, promedio, tipo contador, máximo, mínimo). Los campos sumariados que se definieron para la tabla de hechos FT_Importación fueron los siguientes:
 Días Envío: Duración en días de transporte internacional
 Días Aduanas: Duración en días en la agencia de aduanas.
 Días Trans. Local: Duración de envío en territorio nacional.
 Días Almacén: Duración en días almacenado dentro de operador Logístico.
 Monto Seguro: Monto en Dólares de Seguro Contratado para la importación.
 Monto Envío: Monto en Dólares de Transporte o Flete
 Monto Almacenaje: Monto en Dólares de almacenaje intermedio.
 Precio Base o FOB: Monto en Dólares contratado inicialmente fuera de gastos extras.
- Configuración de las jerarquías de Tiempo como fecha, mes, trimestre, semestre, año y jerarquías de Producto como producto, Grupo, Subgrupo. En las Figura 4.10 se muestra la configuración de reportes o análisis del cubo dentro de la herramienta Oracle Fusion Middleware de Oracle. Aquí se

diseña las consultas al cubo y se revisa la optimización del mismo. En la Figura 4.11 se detallan los resultados del mismo.

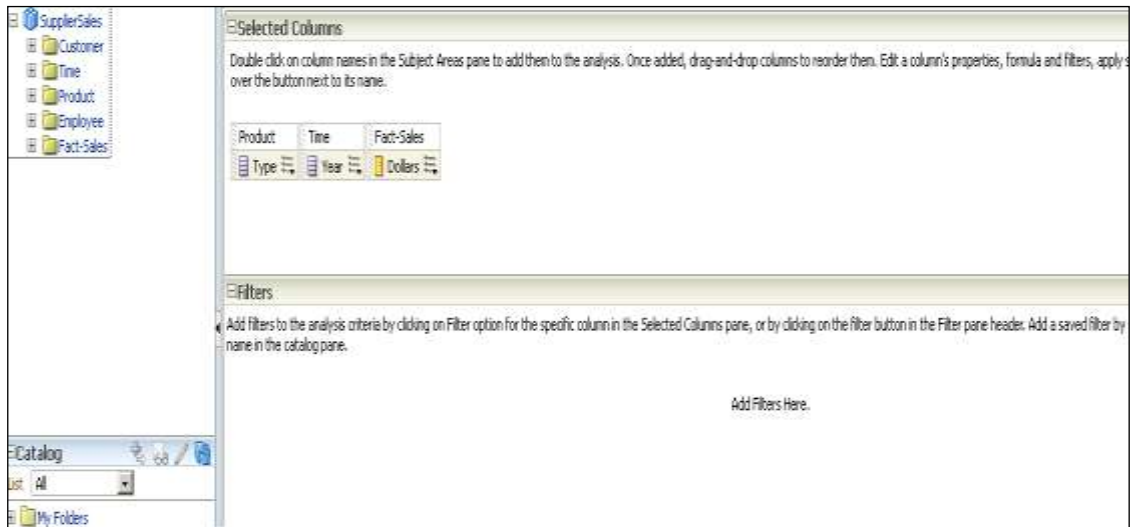
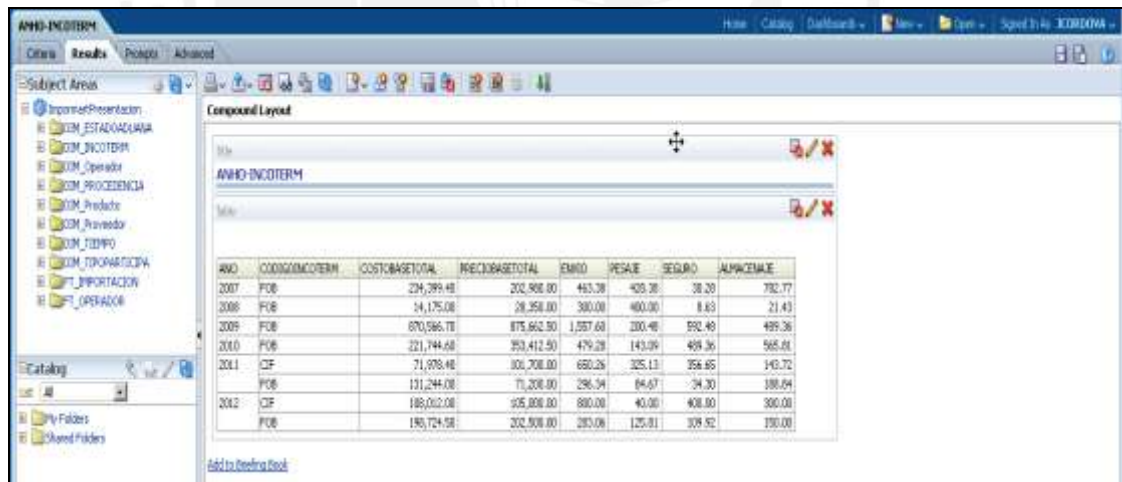


Figura 4.10 - Configuración de reporte en el Cubo



AÑO	CODIGOINCOTERM	COSTOBASSETOTAL	INECIOBASSETOTAL	EMBO	PESAJE	SEGURO	ALMACENAJE
2007	FOB	234,399.48	202,508.00	463.38	429.38	38.20	732.77
2008	FOB	14,175.08	28,350.00	300.00	400.00	8.63	21.43
2009	FOB	870,586.70	875,662.90	1,557.60	200.48	592.48	499.36
2010	FOB	221,744.68	353,412.90	479.28	143.09	499.36	565.61
2011	CFR	71,078.48	301,208.00	680.28	325.13	358.85	143.72
	FOB	131,244.08	71,208.00	296.34	84.67	34.30	188.64
2012	CFR	185,032.08	935,808.00	800.00	40.00	408.00	800.00
	FOB	195,724.58	202,508.00	283.08	125.81	309.92	330.00

Figura 4.11 - Resultados de Consulta

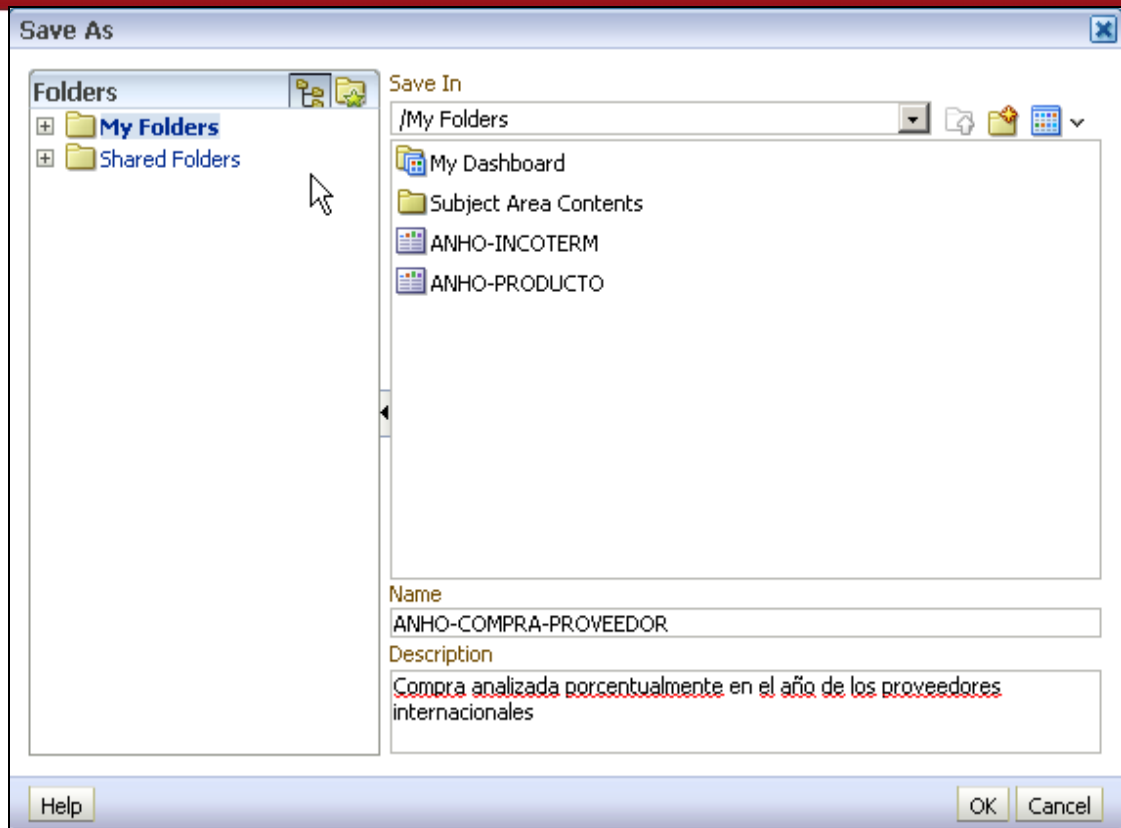


Figura 4.12 - Carpetas de consultas al cubo

Las consultas configuradas se guardan en carpetas personalizadas dentro del servidor de presentación como se muestra en la Figura 4.12, haciendo de fácil acceso al usuario.

4.2. Pruebas

Dentro del proceso de la construcción de la solución, las pruebas constituyen una de las actividades indispensables debido a que a través de ella se asegura la calidad del producto. Es imperativo por ello definir de forma clara la metodología del plan de pruebas que permita verificar los componentes de la solución y así asegurar que se ha construido un producto que cumpla con los requerimientos planteados. Dado que se sigue un esquema top-down (de abajo hacia arriba) las pruebas se realizan en cada iteración del proyecto. A continuación se determinan los tipos de técnicas de eficiencia dentro de una solución de inteligencia de negocios; esto para comprobar el correcto funcionamiento de la solución.

Los casos de prueba propuestos se desarrollaron, en primer lugar, fijando los datos en la persistencia y luego, realizando la operación con persistencia, verificando los valores finales con los valores esperados y, finalmente, borrando la operación.

Los datos de prueba fueron ingresados en la base de datos operacional según los procesos propuestos por la empresa importadora ejemplo.

Los casos de prueba que corresponden con la persistencia se diseñan de tal manera que sea posible volver a ejecutarlos y permitan hacer la comparación de manera automática.

4.2.1. Planificación del plan de Pruebas

Las etapas de extracción, transformación, carga y la explotación de los datos son verificadas por separado.

Se define primeramente el objetivo del plan de pruebas la cual radica en la calidad de datos proporcionados por el Datamart cubriendo específicamente los procesos ETL y asegurando la veracidad de la información mostrada en las maquinas usuario. Esto significa el poder verificar que la información que se encuentra en los sistemas fuente sea igual a la información mostrada a través de la herramienta de explotación. El objetivo fundamental de la realización de pruebas unitarias es asegurar el correcto funcionamiento de las interfases, o flujo de datos entre componentes de extracción y de explotación. También se realizaron pruebas de integridad para el correcto enlace entre los componentes mencionados.

Estas pruebas consiguen una gran cobertura ya que favorecen la granularidad, disminuyen la necesidad de depuración, ayudan a mejorar el diseño y sirven como documentación.

Desde el servidor de base de datos en la solución de inteligencia de negocios se realizaron las siguientes demostraciones como prueba de carga:

- Configuración y monitoreo del sistema en el Servidor (Uso de CPU).
- Administrar el ciclo de vida de los procesos (ETL y Explotación de datos).
- Convertir el mapeo de Dimensiones y tabla de hechos en información de Diseño gracias a la herramienta IDT.

Desde el equipo del usuario final se realizará las siguientes acciones:

- Pre-construir un dashboard según lo establecido en los requerimientos.

- Pruebas de caja negra para comprobar que se cumplen los requisitos del software.
- Pruebas con gran cantidad de datos: stress y volumen de datos.
- Comprobar la visualización en dispositivos avanzados según los requerimientos

Este plan de pruebas deberá ser ejecutado según el plan de proyecto, dentro de este entregable se puede observar esta actividad programada.

4.2.2. Diseño del plan de pruebas

Se definirán los casos de prueba. Se irá registrando y se comparará dicho resultado con la salida proyectada para dicha entrada. De esta manera se determinará si cada objeto de prueba cumple con los criterios de calidad.

Los criterios de aceptación para los casos de prueba se muestran en el Cuadro 4.1.

Objeto de Prueba	Proceso Involucrado	Criterio
Procesos ETL y Base de Datos	Extracción	Los datos extraídos de la base de datos fuente deben ser el mismo que los datos cargados en la base del datamart.
Servicios OLAP	Explotación	La información del análisis final del usuario debe ser igual a los datos del datamart.

Cuadro 4.1 - Objetivos de prueba y sus criterios de aceptación

4.2.3. Determinación de los casos de prueba

Se ha diseñado una plantilla en el Cuadro 4.2 que servirá para especificar los casos de prueba que conformarán el plan de pruebas del proyecto. El diseño tendrá los siguientes campos:

- Nombre de casos de prueba
- Objeto de análisis en el caso de prueba
- Objetivo del caso de prueba
- Datos de entrada del caso de prueba
- Resultado final del caso de prueba

Se revisa el ejemplo del caso de Prueba

Caso de Prueba Nro 1	
Objeto de la prueba:	Revisar ETL y carga de Datamart
Objetivo de la prueba:	Verificar que la tabla FT_OPERADOR teniendo como tablas fuente las tablas de Entradas de Mercancías Importadas y tablas de Importaciones
Precondición:	Disponibilidad de la base de datos fuente
Parámetros requeridos	Año de análisis
Datos de Entrada	OPCH,PCH1,OIPF,IPF1, OPDN,PDN1,OITM,OITB
Descripción de la prueba:	Probar la cantidad de registros or las tablas fuente frente a la cantidad de registro del datamart
Resultados Esperados:	La tabla FT_OPERADOR muestra la misma cantidad de documentos que la tabla de detalle de importaciones expresa en la tabla PDN1
Resultado Obtenido:	Se cargaron los datos correctamente.

Cuadro 4.2 - Caso de Prueba para el Datamart FT_IMPORTACION

4.2.4. Ejecución y análisis de los casos de prueba

Una vez definidos los casos de prueba, se ejecutarán cada una de ellas. Se registrarán por las especificaciones del caso de prueba y se registrará los resultados positivos o negativos de la prueba.

Luego se determinarán los criterios de aceptación o rechazo del producto final con la observación de los requerimientos solicitados inicialmente. Cualquier error podría ser sujeto a correcciones hasta lograr el resultado esperado.

De manera de ejemplo, se presenta dos casos de prueba; el primero sobre el ETL de la tabla de hecho FT_IMPORTACION (proceso de extracción) y el segundo sobre el cubo de Rendimiento de Operadores (proceso de explotación). Se muestra en el Cuadro 4.3 la tabla fuente de la consulta de la base de datos operacional SAP, el resultado indica que existen 53 filas de detalle para las compras de importación y 69 filas para las compras nacionales en total 122 filas de resultados

docentry	FECHA	IDPROVEEDOR	ITEMCODE	CANTIDAD	PRECIOBASE
268	39319	V50000	I00010	50	450
268	39319	V50000	I00011	50	9
268	39319	V50000	I00012	50	33.75
268	39319	V50000	LM4029D	50	26.25
268	39319	V50000	LM4029PH	50	15
269	39370	V10000	C00010	30	20
269	39370	V10000	C00011	30	40
269	39370	V10000	I00001	30	3
270	39479	V10000	LM4029APCD	40	5
270	39479	V10000	A00001	30	400
271	40183	V10000	LM4029MC	40	50
271	40183	V10000	LM4029SB	40	100
271	40183	V10000	LM4029ACA	40	10
271	40183	V10000	LM4029APCD	40	5
271	40183	V10000	A00001	30	400
272	39889	V70000	C00009	40	10
272	39889	V70000	C00010	40	10
272	39889	V70000	C00011	40	20
272	39889	V70000	I00001	40	1.5
272	39889	V70000	I00002	40	6
273	39908	V60000	A00005	40	300
273	39908	V60000	C00001	40	300
273	39908	V60000	C00002	40	225
273	39908	V60000	C00003	40	97.5
273	39908	V60000	C00004	40	26.25
274	40063	V1010	A00002	50	150
275	40070	V30000	C00007	30	250
275	40070	V30000	C00008	30	100
276	40161	V30000	I00012	30	22.5
276	40161	V30000	LM4029D	30	17.5
279	40268	V30000	C00004	50	17.5
284	40619	V60000	LM4029MC	50	37.5
284	40619	V60000	LM4029SB	50	75
284	40619	V60000	LM4029ACA	50	7.5
284	40619	V60000	LM4029APCD	50	3.75
284	40619	V60000	A00001	40	300
287	40484	V1010	LM4029PH	50	15
287	40484	V1010	LM4029MC	50	37.5
287	40484	V1010	LM4029SB	50	75
287	40484	V1010	LM4029ACA	50	7.5

287	40484	V1010	LM4029APCD	50	3.75
288	40909	V1010	A00002	30	150
288	40909	V1010	A00003	30	225
288	40909	V1010	A00004	30	375
288	40909	V1010	A00005	30	300
288	40909	V1010	C00001	30	300
289	40575	V70000	I00010	40	300
289	40575	V70000	I00011	40	6
289	40575	V70000	I00012	40	22.5
289	40575	V70000	LM4029D	40	17.5
289	40575	V70000	LM4029PH	40	10
292	41000	V1010	A00003	100	225
293	40996	V1010	A00006	100	300
277	40179	V30000	A00003	40	150
277	40179	V30000	A00004	40	250
277	40179	V30000	A00005	40	200
277	40179	V30000	C00001	40	200
277	40179	V30000	C00002	40	150
278	40237	V10000	A00003	50	300
278	40237	V10000	A00004	50	500
278	40237	V10000	A00005	50	400
278	40237	V10000	C00001	50	400
278	40237	V10000	C00002	50	300
280	40270	V23000	C00001	40	400
280	40270	V23000	C00002	40	300
280	40270	V23000	C00003	40	130
280	40270	V23000	C00004	40	35
280	40270	V23000	C00005	40	60
281	40492	V23000	C00007	30	500
281	40492	V23000	C00008	30	200
281	40492	V23000	C00009	30	20
281	40492	V23000	C00010	30	20
281	40492	V23000	C00011	30	40
282	40390	V30000	C00006	40	7.5
282	40390	V30000	C00007	40	250
282	40390	V30000	C00008	40	100
282	40390	V30000	C00009	40	10
282	40390	V30000	C00010	40	10
283	40544	V30000	I00001	30	1.5
283	40544	V30000	I00002	30	6
283	40544	V30000	I00003	30	10

283	40544	V30000	I00004	30	15
283	40544	V30000	I00005	30	45
286	40633	V60000	I00005	40	67.5
286	40633	V60000	I00006	40	60
286	40633	V60000	I00007	40	21
286	40633	V60000	I00008	40	29.25
286	40633	V60000	I00009	40	112.5
290	40544	V70000	C00005	30	30
290	40544	V70000	C00006	30	7.5
290	40544	V70000	C00007	30	250
291	40650	V20000	A00001	40	200
291	40650	V20000	A00002	40	100
291	40650	V20000	A00003	40	150
291	40650	V20000	A00004	40	250
291	40650	V20000	A00005	40	200
294	40996	V20000	A00006	80	320
295	41060	V50000	C00011	40	30
295	41060	V50000	I00001	40	2.25
295	41060	V50000	I00002	40	9
295	41060	V50000	I00003	40	15
295	41060	V50000	I00004	40	22.5
296	40999	V50000	I00006	30	60
296	40999	V50000	I00007	30	21
296	40999	V50000	I00008	30	29.25
296	40999	V50000	I00009	30	112.5
296	40999	V50000	I00010	30	450
297	41077	V20000	C00002	30	150
297	41077	V20000	C00003	30	65
297	41077	V20000	C00004	30	17.5
297	41077	V20000	C00005	30	30
297	41077	V20000	C00006	30	7.5
298	41160	V70000	LM4029SB	30	50
298	41160	V70000	LM4029ACA	30	5
298	41160	V70000	LM4029APCD	30	2.5
298	41160	V70000	A00001	50	200
298	41160	V70000	A00002	50	100

Cuadro 4.3 - Tabla Fuente de Compras (Tabla OPCH SAP)

Ahora se realiza la ejecución de la carga y transformación de datos; ejecutamos el paquete de tabla de hechos con el mapeo mostrado en la Figura 4.13. Aquí podremos ver la ejecución y filtración de la data. Inicialmente se ve que la extracción llamada Consulta Entrega indica la obtención de 122 filas en su ejecución; tras los filtros de compras internacionales, la determinación de los precios y los gastos de cada importación; termina insertando 53 filas de datos en el datamart de importación. El resultado final del datamart se visualiza en el Cuadro 4.4. Con dichos resultados finalmente se comprueba que los datos son los esperados.

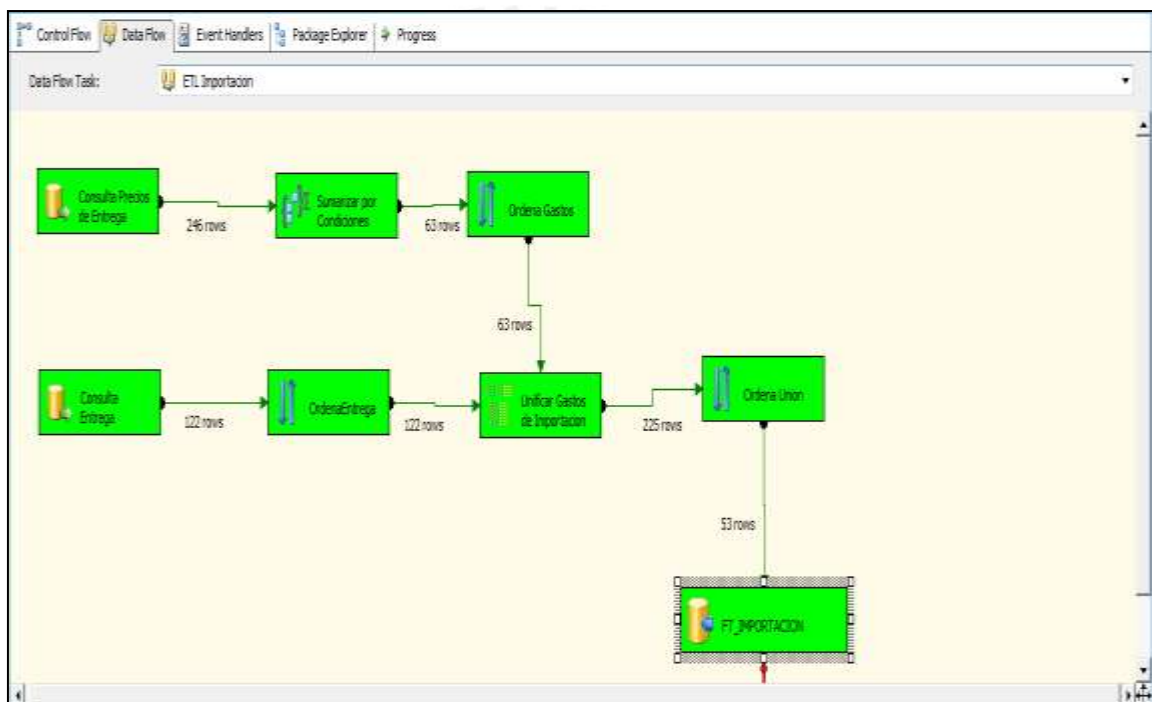


Figura 4.13 - Proceso ETL para Tabla de Hechos FT_IMPORTACION

IDTIE MPO	IDPROV EEDOR	IDPROCE DENCIA	IDPRO DUCTO	IDEST ADO	IDIN COT ERM	CANTI DAD	PRECI OBAS E	COSTOB ASE	DIA S EN VIO	DIAS ADU ANAS	DIAS ALM ACEN	DIA S TRA NSP OR TE LOC AL	SEGU RO	ENVIO
25/08/2007	V50000	BC	I00010	1	3	50	450	527.1	45	10	50	5	4.3	207.61
25/08/2007	V50000	BC	I00011	1	3	50	9	11.79	45	10	50	5	5.11	3.46
25/08/2007	V50000	BC	I00012	1	3	50	33.75	40.92	45	10	50	5	1.47	1.25
25/08/2007	V50000	BC	LM4029D	1	3	50	26.25	41.19	45	10	50	5	4.3	207.61
25/08/2007	V50000	BC	LM4029PH	1	3	50	15	35.22	45	10	50	5	5.11	3.46
15/10/2007	V10000	PA	C00010	1	3	30	20	10.54	45	10	50	5	6.29	16.67
15/10/2007	V10000	PA	C00011	1	3	30	40	20.83	45	10	50	5	3.18	13.33
15/10/2007	V10000	PA	I00001	1	3	30	3	1.82	45	10	50	5	0.54	10
01/02/2008	V10000	PA	A00001	1	3	30	400	200	45	10	50	5	4.32	150
01/02/2008	V10000	PA	LM4029APCD	1	3	40	5	2.5	45	10	50	5	4.32	150
17/03/2009	V70000	PA	C00009	1	3	40	10	10.45	45	10	50	5	1.27	3.29
17/03/2009	V70000	PA	C00010	1	3	40	10	10.54	45	10	50	5	1.27	3.29
17/03/2009	V70000	PA	C00011	1	3	40	20	20.83	45	10	50	5	7.33	4.38
17/03/2009	V70000	PA	I00001	1	3	40	1.5	1.82	45	10	50	5	4.65	7.12
17/03/2009	V70000	PA	I00002	1	3	40	6	6.68	45	10	50	5	14.49	5.48
05/04/2009	V60000	PA	A00005	1	3	40	300	309.09	45	10	50	5	12.78	217.39
05/04/2009	V60000	PA	C00001	1	3	40	300	303.29	45	10	50	5	37.24	130.43
05/04/2009	V60000	PA	C00002	1	3	40	225	150	45	10	50	5	38.3	434.78
05/04/2009	V60000	PA	C00003	1	3	40	97.5	65	45	10	50	5	12.78	217.39
05/04/2009	V60000	PA	C00004	1	3	40	26.25	17.5	45	10	50	5	3.63	86.96
07/09/2009	V1010	CA	A00002	1	3	50	150	100	45	10	50	5	100	100
14/09/2009	V30000	NY	C00007	1	3	30	250	250	45	10	50	5	143.79	169.21
14/09/2009	V30000	NY	C00008	1	3	30	100	200	45	10	50	5	143.79	169.21
14/12/2009	V30000	NY	I00012	1	3	30	22.5	40.92	45	10	50	5	35.59	4.33
14/12/2009	V30000	NY	LM4029D	1	3	30	17.5	41.19	45	10	50	5	35.59	4.33
05/01/2010	V10000	PA	A00001	1	3	30	400	200	45	10	50	5	4.51	41.44
05/01/2010	V10000	PA	LM4029ACA	1	3	40	10	5	45	10	50	5	189.53	41.44
05/01/2010	V10000	PA	LM4029APCD	1	3	40	5	2.5	45	10	50	5	9.26	48.34
05/01/2010	V10000	PA	LM4029MC	1	3	40	50	42.41	45	10	50	5	63.27	63.54
05/01/2010	V10000	PA	LM4029SB	1	3	40	100	50	45	10	50	5	4.51	41.44

31/03/2010	V30000	NY	C00004	1	3	50	17.5	17.5	45	10	50	5	138.66	0
02/11/2010	V1010	CA	LM4029ACA	1	3	50	7.5	5	30	40	50	10	6.41	48.34
02/11/2010	V1010	CA	LM4029APCD	1	3	50	3.75	2.5	30	40	50	10	11.19	41.44
02/11/2010	V1010	CA	LM4029MC	1	3	50	37.5	42.41	30	40	50	10	3.37	41.44
02/11/2010	V1010	CA	LM4029PH	1	3	50	15	35.22	30	40	50	10	52.22	63.54
02/11/2010	V1010	CA	LM4029SB	1	3	50	75	50	30	40	50	10	6.41	48.34
01/02/2011	V70000	PA	I00010	1	3	40	300	527.1	30	40	50	10	6.37	123.82
01/02/2011	V70000	PA	I00011	1	3	40	6	11.79	30	40	50	10	2.63	7.43
01/02/2011	V70000	PA	I00012	1	3	40	22.5	40.92	30	40	50	10	9.46	20.64
01/02/2011	V70000	PA	LM4029D	1	3	40	17.5	41.19	30	40	50	10	9.46	20.64
01/02/2011	V70000	PA	LM4029PH	1	3	40	10	35.22	30	40	50	10	6.37	123.82
17/03/2011	V60000	PA	A00001	1	6	40	300	200	45	10	50	5	191.27	125.65
17/03/2011	V60000	PA	LM4029ACA	1	6	50	7.5	5	45	10	50	5	32.56	125.65
17/03/2011	V60000	PA	LM4029APCD	1	6	50	3.75	2.5	45	10	50	5	61.91	144.5
17/03/2011	V60000	PA	LM4029MC	1	6	50	37.5	42.41	45	10	50	5	61.91	144.5
17/03/2011	V60000	PA	LM4029SB	1	6	50	75	50	45	10	50	5	8.99	109.95
01/01/2012	V1010	CA	A00002	1	3	30	150	100	30	40	50	10	21.9	43.55
01/01/2012	V1010	CA	A00003	1	3	30	225	230.06	30	40	50	10	21.9	43.55
01/01/2012	V1010	CA	A00004	1	3	30	375	382.39	30	40	50	10	16.61	65.32
01/01/2012	V1010	CA	A00005	1	3	30	300	309.09	30	40	50	10	27.61	87.1
01/01/2012	V1010	CA	C00001	1	3	30	300	303.29	30	40	50	10	21.9	43.55
28/03/2012	V1010	CA	A00006	1	6	100	300	310	45	10	50	5	300	500
01/04/2012	V1010	CA	A00003	1	6	100	225	230.06	45	10	50	5	100	300

Cuadro 4.4 - Resultados de Datamart FT_Importacion

Para las pruebas del cubo de Rendimiento de Operadores del proceso de explotación genera el caso de prueba descrito en el Cuadro 4.5

Caso de Prueba Nro 1	
Objeto de la prueba:	Servicios OLAP
Objetivo de la prueba:	Verificar que el cubo de Rendimiento de Operadores muestre la misma información que
Precondición:	Disponibilidad de la base de datos fuente el datamart FT-Operadores. Disponibilidad del servidor de Presentación o Cubo.
Parámetros requeridos	Año de análisis
Datos de Entrada	Valores del datamart frente a los valores del cubo
Descripción de la prueba:	Realizar la consulta del datamart y del cubo
Resultados Esperados:	El cubo de Rendimiento de Operadores muestre la misma cantidad de registros que la tabla FT-Operadores dentro del datamart
Resultado Obtenido:	Muestra la misma información

Cuadro 4.5 - Caso de Prueba para el Cubo de Mantenimiento

Se compara con los resultados del datamart para el rendimiento del operador mostrado en la Figura 4.14 y que evidencia la presencia de 102 filas dentro del datamart. La Figura 4.15 muestra la confirmación del número de filas expresado con los resultados del datamart.

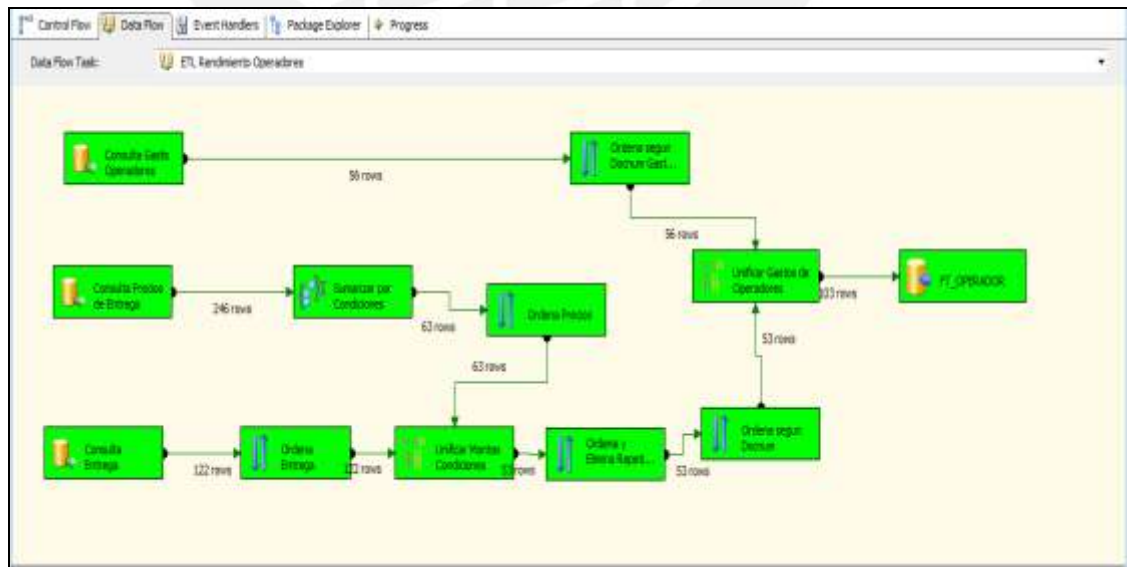
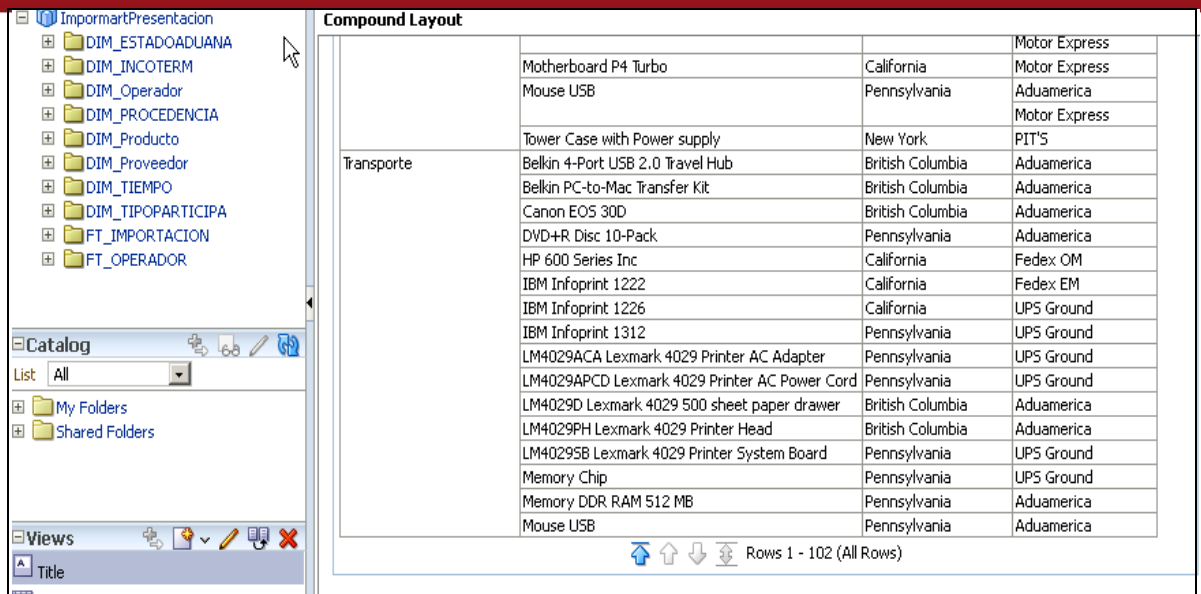


Figura 4.14 - Proceso ETL para tabla de hechos FT_OPERADOR



	Motherboard P4 Turbo	California	Motor Express
	Mouse USB	Pennsylvania	Aduamerica
			Motor Express
	Tower Case with Power supply	New York	PIT'S
Transporte	Belkin 4-Port USB 2.0 Travel Hub	British Columbia	Aduamerica
	Belkin PC-to-Mac Transfer Kit	British Columbia	Aduamerica
	Canon EOS 30D	British Columbia	Aduamerica
	DVD+R Disc 10-Pack	Pennsylvania	Aduamerica
	HP 600 Series Inc	California	Fedex OM
	IBM Infoprint 1222	California	Fedex EM
	IBM Infoprint 1226	California	UPS Ground
	IBM Infoprint 1312	Pennsylvania	UPS Ground
	LM4029ACA Lexmark 4029 Printer AC Adapter	Pennsylvania	UPS Ground
	LM4029APCD Lexmark 4029 Printer AC Power Cord	Pennsylvania	UPS Ground
	LM4029D Lexmark 4029 500 sheet paper drawer	British Columbia	Aduamerica
	LM4029PH Lexmark 4029 Printer Head	British Columbia	Aduamerica
	LM4029SB Lexmark 4029 Printer System Board	Pennsylvania	UPS Ground
	Memory Chip	Pennsylvania	UPS Ground
	Memory DDR RAM 512 MB	Pennsylvania	Aduamerica
	Mouse USB	Pennsylvania	Aduamerica

Figura 4.15 - Confirmación de la consulta extraída desde el Cubo

Herramientas de SQL como Performance Monitor para medir el rendimiento, Tunning Advisor para mejorar las consultas preparadas por el usuario y determinar el tamaño de las tablas son herramientas de control de procesos realizados dentro de la fase de extracción. Se presenta en la Figura 4.16 el trabajo de tablas con Tunning Advisor.

Herramientas de Oracle como Oracle Database 11g Performance Tuning y Analytic SQL for Data Warehousing logran detectar las consultas de explotación directa de los usuarios y la transformación a las consultas físicas a la base de datos operacional. Existen también otras herramientas externas como SpotLight o TBufeo que realizan pruebas exhaustivas a las base de datos dimensionales para ver el performance respectivo.

Aparte de ello ORACLE ofrece el monitoreo de todas las consultas lógicas y físicas a las bases de datos operacionales, detectar las que ralentizan los procesos sobre diversas tablas criticas o no, esto para dar soluciones alternativas como el particionamiento o indexamiento de tablas.

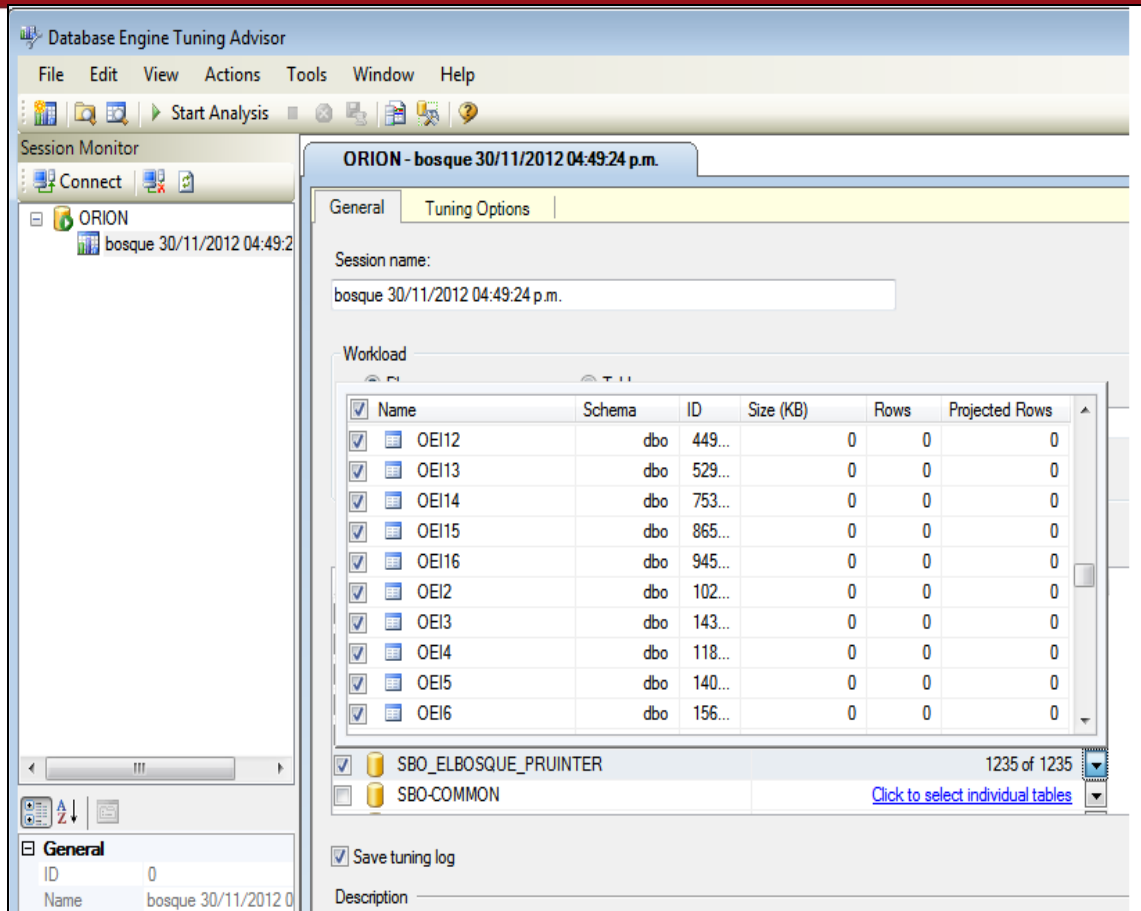


Figura 4.16 - Determinación de tamaño de tablas Dimensionales

CAPÍTULO 5: OBSERVACIONES, CONCLUSIONES Y RECOMENDACIONES

El presente capítulo muestra las observaciones sobre el proyecto, así como las conclusiones y las recomendaciones para proyectos futuros relacionados al tema.

5.1. Observaciones

Se decidieron usar herramientas de software licenciado para la base de datos y para la herramienta de BI para los procesos de explotación y carga dada la posibilidad de auspicio para el mismo.

La herramienta de software elegida para el uso de la interfaz de reportes permitió un manejo intuitivo y sencillo a los usuarios finales para generar sus propios reportes y realizar el análisis correspondiente.

Se toma en cuenta que los reportes satisfacen las necesidades de los usuarios para la adecuada toma de decisiones. Aparte de ello, les ayuda a reducir los tiempos de respuesta en el procesamiento y análisis de la información, lo que indica que pueden ser empresas sostenibles en el tiempo bajo un entorno competitivo.

Se concluye que el manejo de las herramientas de inteligencias de negocios, licenciados o no, son considerados como soporte de la solución; pero el buen manejo de la teoría de estas soluciones hacen las diferencia entre la eficiencia o no de las mismas.

Las herramientas de inteligencias de negocios utilizados como el de SQL o el de la marca Oracle son sólo componentes de la solución pero no las más importantes. Un buen modelado de datos acorde a las necesidades del negocio es el eje principal de todo proyecto de inteligencia de negocios.

La solución está orientada al uso de todo el potencial que brinda Microsoft SQL y Oracle, desde su versión 10g, y sus respectivas herramientas de inteligencia de negocios que son más amigables para la programación y diseño.

La retroalimentación de información con los usuarios finales hacen que esta solución pueda expandirse poco a poco para tener una poderosa fuente de información del trabajo y las operaciones del día a día.

5.2. Conclusiones

Las conclusiones obtenidas a la investigación y desarrollo de la presente tesis son las siguientes:

- Se cumple con el objetivo de realizar las coordinaciones necesarias con las áreas involucradas para el levantamiento respectivo y el correspondiente análisis del problema en la empresa de estudio. Se definieron las reglas del negocio, y se estructuraron las prioridades e indicadores claves que se fundamentarán en los reportes a elaborarse dentro de la solución. Se analizaron y desarrollaron los requerimientos funcionales y no funcionales de cada área involucrada.
- La selección de una arquitectura adecuada permite el desarrollo de un sistema de manera rápida y organizada. Para el caso, se logra que el acceso a datos, la lógica de negocio y la interfaz gráfica sean componentes independientes, lo cual permite un sistema flexible y de fácil mantenimiento. De esta forma si se desea por ejemplo cambiar base de datos por algún motivo, este cambio ocasionaría un impacto mínimo en los componentes. La correcta administración y ejecución de los componentes dentro de la fase de extracción como en la explotación garantizan la eficiencia para la captura los datos desde su origen hasta llevarlos hasta el repositorio del datamart.
- Se desarrolló el mapeo respectivo para la extracción, transformación y carga de todos los elementos correspondiente a los datamarts que incluyen la

solución de inteligencia de negocios implementada. Se definieron los lineamientos de comparación de arquitecturas y de motores de búsqueda para que se obtenga el óptimo desempeño.

El tiempo invertido en las etapas de análisis y diseño minimiza riesgos en la implementación y reduce pérdidas de tiempo, debido a que una acertada definición de funcionalidades permite que los cambios posteriores sean mínimos. Por ejemplo si se hubiese realizado un análisis rápido y no concienzudo del diagrama de clases, se habrían presentado muchos problemas en la fase de construcción de la jerarquía de bloques, debido a la complejidad presentada.

- Cada fase del proyecto fue ejecutado sobre una base metodológica enfocando la gestión de alcance, calidad y tiempo de la metodología PMBOK. El uso de esta metodología para el manejo del proyecto de tesis y los conocimientos en la metodología de programación del ciclo de vida de Ralph Kimball hicieron que se enfoque correctamente las necesidades y cambios dentro del entorno físico y técnico de la solución durante el proceso del mismo.
- El desarrollo de los procesos de extracción transformación y carga son los apropiados según la información requerida por el área.
- Se diseñó un modelo dimensional adecuado según la cantidad y la profundidad de datos que requiere la organización y posee el datamart.
- La ejecución de las consultas de los usuarios fueron debidamente empaquetados para su correcta configuración personalizándose según las necesidades de cada área. La interfaz de consulta es flexible para cada ejecución de los usuarios finales pero existe una configuración inicial base para las consultas básicas la cual se mantiene almacenado como también la configuración de seguridad para el acceso pertinente de cada tipo de usuario.

5.3. Recomendaciones

A manera de recomendación se sugiere implementar nuevas funcionalidades que no solo sea una herramienta de inteligencia de negocios para una empresa de

bienes sino que también pueda emplearse en las distintas empresas peruanas que ofrecen servicios en el país y en el extranjero.

Entre dichas funcionalidades se podrían mencionar: permitir el ingreso de servicios y las asignaciones de cuentas contables de importación, permitir la inclusión del costo de un servicio o garantía en la compra de bienes.

Se necesitara quizás crear módulos adicionales o reorganizar los ya existentes pero lo ya implementado es en realidad lo básico para manejar la información relevante dentro de los procesos de importación.



Bibliografía

Textos de consulta

- [BIT 2002] Vitt, Elizabeth; Luckevich, Michael; Misher, Stacia 2002
Business Intelligence Técnicas de análisis para la toma de decisiones
estratégicas. 1era Edición, Madrid: McGraw-Hill/Interamericana de
España S.A.U.
- [POR 2001] Poirier, Charles 2001
Administración de cadenas de aprovechamiento. 1era Edición,
Madrid: Oxford University Press México S.A.
- [PEN 008] Joy Mundy, Warren Thornthwaite with Ralph Kimball 2011
The Microsoft Data Warehouse Toolkit with SQL Server 2008 R2
Wiley Segunda Edición
- [PEN 007] Ralph Kimball, Margy Ross, Joy Mundy, Warren Thornthwaite, Bob
Becker 2008
The Data Warehouse Lifecycle Toolkit. Practical Techniques for
Building Data Warehouse and Business Intelligence Systems
Wiley Segunda Edición
- [LOG 001] Sandoval Aguilar, Luis Alberto 2002
Logística internacional de importaciones
Superintendencia Nacional de Aduanas
- [LOG 002] Guardia Yamamoto, Gustavo 2001
Una función de importaciones para el Perú
PUCP CISEPA
- [LOG 003] Los incoterms y el comercio internacional
Trade Books 2000
- [COD 001] E.F.Codd The Relational Model for Database Management.
1990
Adisson Wesley
Segunda Edición
- [DAT 001] C.J. Date Introducción a los sistemas de bases de datos.
2001
Prentice Hall
Primera Edición
- [JCT001] Norberto Mazon Lopez, Lopez Pardillo Vela, Juan Carlos Trujillo
Mondéjar. Diseño y Explotación de almacenes de datos: Conceptos
básicos de modelado multidimensional
2011
Editorial Club Universitario
- [CAU001] Peter Rob, Carlos Coronel
Sistemas de base de datos: diseño, implementación y administración.
Thompson
Quinta Edición

Sitios de consulta

- [WWW001] Red de información BI La Rioja-Proyecto Ibermática 2007
http://www.larioja.org/upload/documents/458269_gestion_proyectos.pdf
- [WWW002] Modelamiento multidimensional
<http://www.inf.udec.cl/~revista/ediciones/edicion4/modmulti.PDF>
- [WWW003] Notas sobre BI, manual practico
<http://science.kennesaw.edu/~bsetzer/3310fa10/build/notes/notes1129.html>
- [WWW004] Nueva estrategia de gestión aduanera-Superintendencia Nacional de Aduanas
http://www.aduanet.gob.pe/aduanas/informfis/Nueva_Estrategia.htm
- [WWW005] ERP Flap System
<http://www.flapsystem.com/Servicios/ERPFlapSystem/Funcionamiento/tabid/82/Default.aspx>
- [WWW006] Historia de un Viejo Informático. El Data Warehouse entró en nuestras vidas...para quedarse.
<http://eltamiz.com/elcedazo/2009/06/08/el-data-warehouse-entro-en-nuestras-vidas-para-quequedarse/>
- [WWW007] N. Pendse The OLAP Report
<http://www.olapreport.com>
- [WWW008] Sap Support Portal
<https://websmp108.sap-ag.de/support>

