

Universidad Central “Martha Abreu Estévez” de Las Villas

Facultad de Matemática, Física y Computación



DISEÑO DE UN MERCADO DE DATOS DEL ÁREA COMERCIAL PARA LA EMPRESA VÁLVULAS URREA S.A. DE C.V.

**Tesis presentada en opción al Título Académico de
Master en Computación Aplicada**

**Autor: Ing. Felipe Ramirez Salcido
Tutor: Dr. RAMIRO PEREZ VÁZQUEZ**

**Guadalajara, Jalisco, México
2003**

Índice

INTRODUCCIÓN.....	1
ANTECEDENTES.....	1
OBJETIVOS.....	4
CAPÍTULO 1. REVISIÓN BIBLIOGRÁFICA	5
1.1 CONCEPTOS FUNDAMENTALES	5
ERP	5
MODELO DIMENSIONAL	6
RELACIÓN DEL MODELO ENTIDAD RELACIÓN VS EL MODELO DIMENSIONAL	10
PROBLEMAS DEL SQL.....	12
OLAP (Online Analytic Processing)	13
ROLAP (Relational OLAP)	14
MOLAP (multidimensional OLAP)	15
COMPARACIÓN ENTRE ROLAP Y MOLAP.....	17
ALMACENES DE DATOS.....	18
MERCADOS DE DATOS	20
CAPÍTULO 2. PLANTEAMIENTO DEL PROBLEMA.....	22
CAPITULO 3. SOLUCIÓN AL PROBLEMA COMPUTACIONAL.....	25
3.1 DIMENSIONES DEL MERCADO DE DATOS PARA VENTAS	29
Descripción de las Dimensiones	30
Detalles de la Dimensión Fecha.....	31
Descripción de los Atributos de la Dimensión Fecha.....	32
Detalles de la Dimensión Cliente.....	33
Descripción de los Atributos de la Dimensión Cliente	33
Detalles de la Dimensión Producto	35
Descripción de los Atributos de la Dimensión Producto	36
Detalles de la Dimensión Promociones.....	38
Descripción de los Atributos de la Dimensión Promociones.....	38
Detalles de la Dimensión Condiciones de Pago	40
Descripción de los Atributos de la Dimensión Condiciones de Pago	40
Detalles de la Dimensión Canal de Distribución	42
Descripción de los Atributos de la Dimensión Canal de Distribución	42
Detalles de la Dimensión Sector	43
Descripción de los Atributos de la Dimensión Sector	43
Detalles de la Dimensión Línea	44
Descripción de los Atributos de la Dimensión Línea.....	44
Detalles de la Dimensión Familia	44
Descripción de los Atributos de la Dimensión Familia	45
Detalles de la Dimensión Oficina de Ventas	45
Descripción de los Atributos de la Dimensión Oficina de Ventas	46
Detalles de la Dimensión Destinatario de Mercancías.....	46
Descripción de los Atributos de la Dimensión Destinatario de Mercancías.....	46
3.2 DETALLES DE HECHOS.....	48
Tabla de Hechos de Pedidos	48
Descripción de los hechos de Pedidos	49
Tabla de Hechos de Facturación.....	50
Descripción de los hechos de Facturación	50
3.3 DEMOSTRACIÓN DEL MODELO EN POWER PLAY.....	51

CONCLUSIONES.....	56
BIBLIOGRAFÍA.....	57
REFERENCIAS BIBLIOGRÁFICAS.....	59

Resumen

El presente trabajo está orientado al diseño de un mercado de datos para el área comercial de la empresa Válvulas Urrea, cuyo origen de datos será el sistema ERP R/3 de SAP, para lo cual se estudia aquí la teoría de los almacenes de datos y del análisis multidimensional.

En la empresa ha sido implementado desde el año 2000 el sistema ERP R/3 de SAP para el control de las operaciones de la organización. Sobre los datos que en dicho sistema se genera se extraerá información específica para crear el mercado de datos (*Data Mart*) para cubrir las necesidades de análisis del área comercial.

El objetivo primordial es facilitar la toma de decisiones de los miembros del comité directivo, al apoyarlas con herramientas que les permitan hacer análisis de los datos que son difíciles de obtener solamente con el sistema ERP.

Summary

The present work it's oriented to the design of one data mart for the comercial area of the Válvulas Urrea enterprise, the data will be extracted from an ERP system SAP R/3. To develop this work it's necessary to study the data warehousing theory and the multidimensional analysis.

The Válvulas Urrea Corp. had implemented in the 2000 year the ERP SAP R/3 system for the control of the whole operations in the company. From the data generated in that system It will be extracted specific information for making the comercial data mart, and to satisfy the analysis needs of that department of the company.

The primary object of this *work* is to make it easy to the managers to make fast and right decisions, supported by easy tools for that data analysis.

INTRODUCCIÓN

ANTECEDENTES

Hasta antes del año 2000 la empresa Válvulas Urrea basaba sus operaciones en sistemas hechos a la medida por sus empleados, generados en la plataforma AS/400 de IBM. Cada sistema de cada departamento de la compañía estaba aislado de los sistemas de las otras áreas, formando “islas” separadas e independientes que ocasionaban el problema de que la información generada por cada departamento no podía integrarse rápidamente para los directores de la empresa y para lograrlo debían dedicarse muchas horas de trabajo de personal asignado a esas actividades de “concentrar información”, además que en muchas ocasiones no coincidían los datos aún cuando se hablaba del mismo tema.

Desde mayo del año 2000 se implementó en la compañía el sistema ERP (*Enterprise Resource Planning*) llamado R/3 de la compañía SAP con la intención de poder controlar todas las operaciones de la empresa en un solo sistema global e integral que permitiera consolidar la información en tiempo real, con el objetivo de disminuir costos y dar ventajas competitivas a la compañía al tener la información en tiempo real para todas las áreas de la compañía.

El problema de concentrar la información en un solo sistema ha quedado resuelto mediante la adquisición de un conjunto de “módulos” del sistema R/3 que están ligados entre sí, sin embargo, el sistema tiene solamente funcionalidad transaccional (OLTP), es decir que está diseñado para guardar la información de la operación pero no cuenta con herramientas poderosas suficientes para hacer consultas y reportes, sino que todavía deben invertirse muchas horas de trabajo para generar algunos reportes estadísticos que apoyen de diversas maneras a la toma de decisiones.

La compañía SAP ha desarrollado la herramienta de tipo OLAP (*Online Analytical Processing*) llamada BIW (*Business Information Warehouse*) la cual es sumamente poderosa y eficiente para realizar análisis multidimensionales. Uno de los proyectos de la empresa a mediano-largo plazo es la implementación de esta herramienta, pero dado que sus prioridades de inversión actualmente son otras, dicho proyecto se ha pospuesto debido a que esta solución es extremadamente costosa. Actualmente no se tiene la certeza de que dicho proyecto pueda ser reactivado en el futuro próximo.

Mientras tanto, el sistema actual obliga a los directivos a realizar exhaustivos y pesados procedimientos de análisis de información para la planeación estratégica y desarrollo de la empresa, con el riesgo de que se pueden cometer errores graves derivados de factores humanos como cansancio, salud, anímicos, etc.

Uno de los motivos de la labor de planeación estratégica mencionada, es el de buscar permanentemente una ventaja competitiva en el mercado, razón por la cual se plantea el presente proyecto de diseño de un mercado de datos para el área comercial, debido a que este departamento es el que más requiere de una herramienta de análisis expedito para sustentar sobre bases más sólidas y confiables el proceso de toma de decisiones y su inmediata aplicación.

A pesar de que la empresa ha hecho grandes inversiones en la implementación del sistema R/3 de SAP y en el hardware necesario para la integración de datos, prevalece actualmente la carencia de “informes” que apoyen la toma de decisiones referentes a la mercadotecnia-producción, debido a que el sistema SAP es meramente transaccional y que su diccionario de datos consta de miles de tablas, esto complica enormemente el desarrollo de consultas por las dificultades que tiene el modelo relacional para unir muchas tablas, además debido a la complicada operación de la empresa, la base de datos crece muy rápido lo que complica cada vez más el proceso de datos.

La necesidad antes mencionada obliga al personal de la dirección a dedicar enormes esfuerzos de compilación y análisis de datos que además de ser agotadores y lentos, están supeditados a factores humanos que pueden conducir a graves errores deductivos y en algunos casos, que finalmente se traducen en pérdidas económicas.

Esta necesidad no se ha atendido debido a la fuerte inversión que representaría adquirir la solución de SAP *ad-hoc* a la empresa y que además se llevan a cabo otros proyectos más prioritarios, sin embargo, se cuenta con la licencia (no aprovechada) para uso de una herramienta OLAP (Power Play, de Cognos) el que podría ser utilizado de manera sustancial en el presente proyecto.

La viabilidad de la implementación del presente proyecto dependerá de la confirmación de resultados de acuerdo a las hipótesis y objetivos que se plantean, sin embargo, se cuenta con el interés y el apoyo necesario por parte de la empresa para llevar a cabo la presente investigación en el presente año.

Por tanto en base a lo descrito anteriormente los objetivos son:

OBJETIVOS

1. Diseñar un Mercado de Datos (*Data Mart*) en el que se almacenarán los datos relevantes a las ventas orientado al análisis para efectos de la planeación estratégica de la compañía.

Objetivos Específicos

1. Minimizar el tiempo invertido por los empleados en preparar los informes que solicitan los ejecutivos.
2. Reducir el impacto “negativo” en el desempeño global del sistema SAP ocasionado por las consultas que ejecutan los usuarios para generar los informes.
3. Ayudar a agilizar la toma de decisiones de los ejecutivos relacionados con el área comercial mediante el uso de herramientas disponibles que aprovechen la información del sistema ERP (*Enterprise Resource Planning*).
4. Minimizar la inversión económica para el proyecto mediante la utilización de la infraestructura de software que existe en la empresa (Power Play) para realizar el análisis OLAP a partir de los datos extraídos del sistema R/3 de SAP.

CAPÍTULO 1. REVISIÓN BIBLIOGRÁFICA

1.1 CONCEPTOS FUNDAMENTALES

ERP

Planeación de recursos empresariales o ERP (*Enterprise resource planning*) por sus siglas en inglés, son los sistemas transaccionales (OLTP, *online analytical processing*) encargados de ayudar a controlar importantes partes de una empresa incluyendo planeación de la producción, compras, manejo de inventarios, ventas, etc. Típicamente un sistema ERP utiliza o está integrado con un sistema de bases de datos relacional. Así que cuando una transacción se genera en uno de los módulos, los datos asociados y resultados son automáticamente actualizados para las demás aplicaciones. De esta forma los procesos del negocio son integrados completamente.

Los más grandes proveedores de sistemas ERP son Baan, J.D. Edwards, Oracle, PeopleSoft, y SAP.

Los objetivos fundamentales de los sistemas ERP siempre han sido rehacer y reemplazar los sistemas primarios operacionales de las organizaciones con un complejo conjunto de aplicaciones bien integradas que implementan un amplio conjunto de procesos de negocios. Las instalaciones más exitosas de sistemas ERP han sido aquellas donde el cliente ha construido sus procesos de negocio en torno al software ERP, en lugar de hacerlo al revés [KIM00].

Los sistemas ERP están basados en el Modelo Relacional y sus implementaciones están creadas en gestores de bases de datos relacionales como Informix, Oracle, SQL Server, etc.

Debido a la naturaleza de estos sistemas, basada en el modelo relacional, los usuarios de dichas aplicaciones tienen problemas para obtener lo que los ejecutivos necesitan: reportes y consultas de su información para hacer consolidaciones y análisis dimensionales, que sirvan para apoyar la toma de decisiones, pues deben acoplar varias tablas para obtener la información necesaria de un reporte.

Algunas de las dificultades más importantes de los sistemas ERP para explotar la información son principalmente debido a los siguientes puntos:

- Los sistemas ERP han estado más interesados con el procesamiento transaccional que en el soporte a las decisiones.
- Los esquemas primarios de Base de Datos de los sistemas ERP son extremadamente complejos, incluyendo miles de tablas de base de datos. Aunque es posible extraer datos de un sistema ERP hacia un Almacén de datos, se requieren conocimientos especializados y un poderoso software de extracción de datos.
- Aunque los proveedores de ERP han introducido entornos de almacenes de datos en sus ofertas de productos, hasta ahora no han demostrado convincentemente su deseo de facilitar las actividades de importar y exportar datos, aun cuando eso es lo que el cliente desea [KIM00]

MODELO DIMENSIONAL

El modelado dimensional es también la técnica de diseño lógico usada por los almacenes de datos.

El modelado dimensional parte del principio de que el objetivo principal de un sistema de decisión es el análisis del rendimiento. Este rendimiento puede materializarse a través de un conjunto de indicadores.

Estos indicadores se analizarán a través de **dimensiones**. El tiempo, dimensión que se encuentra casi sistemáticamente en toda aplicación de decisión, permitirá seguir la evolución de una cifra de negocio, por trimestre, por año, etc. Así mismo será interesante obtener esta cifra de negocio por producto, por región, por cliente, que son otras tantas dimensiones de análisis pertinentes para este indicador.

El modelado dimensional DM (*Dimensional Modeling*) es una técnica de diseño lógico que busca presentar los datos en un marco estándar e intuitivo que permita acceso con alto desempeño. En este modelo, los indicadores de base se agrupan en una tabla central, llamada tabla de hechos. Una tabla de hechos agrupa todos los indicadores que comparten el mismo conjunto de dimensiones y que no pueden ser deducidos de otros indicadores. En la figura 2 se muestra un ejemplo.

A dichos diagramas se les conoce como modelo de estrella (*Star Join*). Dicho modelo Estrella fue popularizado por el Dr. Ralph Kimball para describir el proceso de desnormalización que simula la estructura de una base de datos multidimensional. Kimball describe la desnormalización como la pre-conjunción (pre-join) de tablas, para que la aplicación en tiempo de ejecución no realice los joins de las mismas. [BUR96]

Una tabla de hechos siempre expresa una relación de muchos a muchos, debido a que contiene una llave primaria multiparte creada por dos o más llaves extranjeras. Las tablas de hechos más útiles también contienen una o más mediciones numéricas, o “hechos” que ocurren por la combinación de llaves que definen cada registro.

Alrededor de esta tabla de hechos figuran todos los elementos que caracterizan las dimensiones del análisis. Las tablas de dimensiones contienen información textual descriptiva. Los atributos de las dimensiones se utilizan como la fuente de las consultas que se hacen al almacén de datos.

Alrededor del principio de consultas en el modelo de estrella está la introducción de datos altamente redundantes por razones de desempeño. Esencialmente, una tabla de hechos es una representación en la Primera Forma Normal (1FN) de la base de datos con un alto grado de redundancia añadido a sus tablas.[BUR96]

Para comprender el proceso de desnormalización veamos el siguiente ejemplo. En la figura 1 se muestra una base de datos estándar en la Tercera Forma Normal (3FN) para representar la venta de artículos.

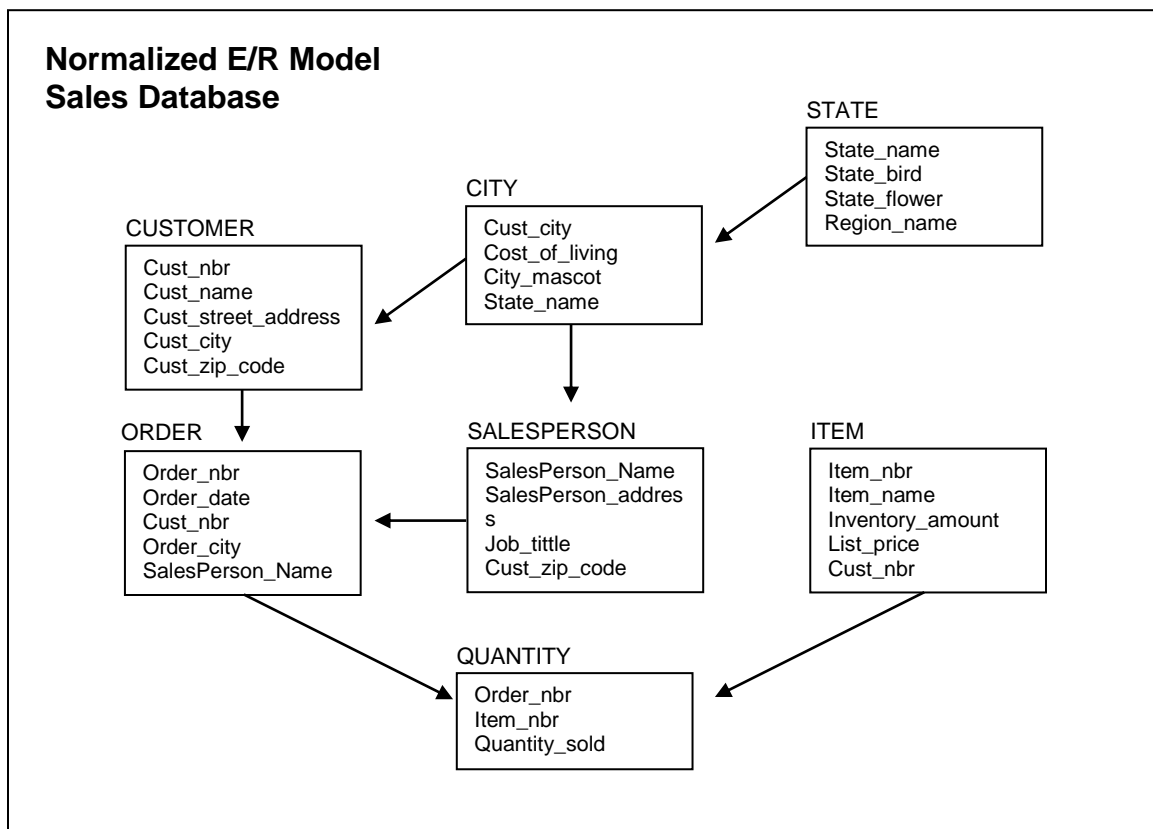


Figura 1. Ejemplo de modelo E/R completamente Normalizado.

En la Figura 2 se muestra el diseño desnormalizado, el cual aumenta considerablemente la simplicidad pero a costa de la redundancia de la información.

La nueva Tabla de Hechos contendrá una fila para cada partida de cada pedido, resultando en una tremenda cantidad de información redundante.

Nótese también las tablas de dimensiones que rodean la tabla de hechos. Algunas de las tablas de dimensiones contienen datos que pueden ser añadidas a las consultas con joins, mientras que otras dimensiones como la dimensión REGION no contiene ningún dato y solamente sirve como un índice para los datos.

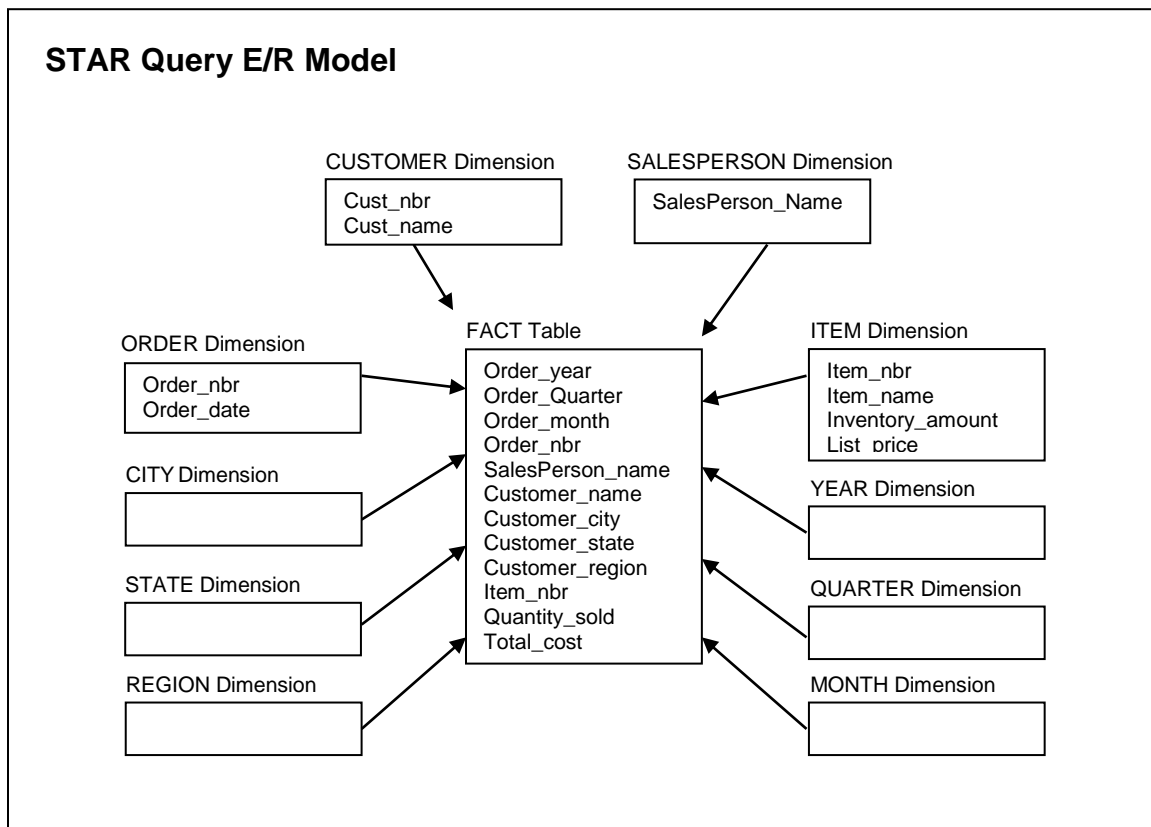


Figura 2. Modelo dimensional detallado de venta de artículos.

El modelo en estrella, y más generalmente todos los modelos multidimensionales que se derivan de él, parten del principio de que son principalmente los análisis de los indicadores (dimensiones) los que interesarán al usuario [FRA97].

Éste es un punto fundamental porque caracteriza las dos mayores fuerzas de este tipo de modelo: la legibilidad y el rendimiento. La legibilidad ante todo: este modelo es muy elocuente para el usuario y presenta de manera clara su finalidad. Está orientado al tema y define claramente los indicadores de análisis. Luego el rendimiento: los caminos de acceso a la base de datos son previsibles. En este modelo la tabla de hechos puede contener varios millones de filas, pero las tablas de dimensiones son más reducidas. Es más fácil controlar y optimizar los accesos a las tablas de hechos, porque se accede a ella después de haber efectuado selecciones sobre las tablas de dimensiones. La tabla de hechos contienen una gran cantidad de filas, pero el tamaño de cada fila es reducido.

RELACIÓN DEL MODELO ENTIDAD RELACIÓN VS EL MODELO DIMENSIONAL

El modelo Dimensional es como hemos dicho un modelo lógico, también el modelo Entidad Relación cae en esta categoría.

La clave para comprender la relación entre modelo dimensional (DM) y modelo entidad relación (ER) es que un solo diagrama ER se puede convertir en múltiples diagramas DM.

Así que el primer paso para convertir un diagrama ER en un conjunto de diagramas DM es separar el diagrama entidad relación en procesos discretos de negocio y modelar a cada uno de ellos por separado.

El segundo paso es separar esas relaciones muchos a muchos en el modelo ER que contienen información numérica y también los datos que se pueden sumar que no forman parte de una llave y designarlos como tablas de hechos.

Para materializar el modelo ER se utiliza habitualmente el modelo relacional, el cual ha contribuido enormemente al fenomenal éxito de poner grandes cantidades de datos en bases de datos relacionales. Sin embargo, el modelo relacional no contribuye a las habilidades de los usuarios para consultar datos.

Los diagramas ER son útiles, pero fueron concebidos para ser vistos en secciones pequeñas, no todo al mismo tiempo.

En general, las siguientes definiciones se aplican a los almacenes de datos que se basan en el modelo dimensional, a diferencia de los sistemas OLTP que se basan en el modelo relacional:

- *Datos orientados a Temas:* A diferencia de una aplicación orientada a transacción, la cual está enfocada en una transacción específica del negocio, un almacén de datos intenta reunir todo lo que se conoce sobre un tema desde todas las fuentes de información dentro de la organización.
- *Solo lectura durante consultas:* Los almacenes de datos son actualizados durante las horas de la noche y se utilizan para solicitudes de lectura durante el día.
- *Estructuras altamente desnormalizadas:* A diferencia de un sistema OLTP con muchas tablas “estrechas” los almacenes de datos efectúan pre-joins de tablas, creando tablas “robustas” con gran redundancia de columnas.
- *Los datos son precalculados:* Al contrario de los sistemas OLTP, los almacenes de datos precalculan totales para mejorar el desempeño del sistema. Nótese que el pre-sumarizar totales es antirelacional, esto significa que el modelo relacional aboga por calcular totales en tiempo de ejecución, solamente permitiendo que se almacenen componentes de datos atómicos.

Así, cuando contrastamos a un almacén de datos con un sistema orientado a transacciones, las diferencias son evidentes. Dichas diferencias se muestran en la tabla 1. [BUR96]

Diferencias entre Sistemas OLTP y un Almacén de Datos		
	Sistemas OLTP	ALMACEN DE DATOS
Normalización	Alta (3FN)	Baja (1FN)
Tamaño de Tablas	Pequeño	Grande
Número de filas/tablas	Pequeño	Grande
Tamaño/duración de transacción	Pequeño	Grande
Número de usuarios en línea	Alto (hasta miles)	Pequeño (<100)
Actualizaciones	Frecuentes	Por las noches
Lecturas completas de tablas	Raramente	Frecuentemente
Datos históricos	<90 días	Años

Tabla 1. Diferencias entre un sistema OLTP y un Almacén de Datos.

PROBLEMAS DEL SQL

El área más difícil de la creación de los almacenes de datos es la traducción de análisis de negocios simples a SQL. SQL no fue diseñado con reportes de negocios en mente. SQL fue un lenguaje diseñado para permitir que la semántica de las tablas relacionales fuera expresada en una forma conveniente y accesible.[KIM96]

Algunas de las limitantes antes mencionadas de los sistemas ERP tienen su origen en las dificultades que existen en el lenguaje de consultas estructurado SQL (*Structured Query Language*) para hacer consultas complejas, por ejemplo no hay una forma directa en SQL de comparar este año al año anterior, o comparar un grupo de clientes con otro.

Desarrolladores de aplicaciones y usuarios finales de todo el mundo están desperdiciando enormes cantidades de su valioso tiempo implementando simples comparaciones y simples cálculos secuenciales en sus aplicaciones clientes.

OLAP (Online Analytic Processing)

OLAP ("Procesamiento Analítico en Línea") describe una clase de tecnologías que están designadas para el acceso a datos "en vivo" y análisis de ellos. Mientras que el proceso de transacción generalmente se basa en bases de datos relacionales, OLAP se ha convertido en el sinónimo de vistas multidimensionales de datos de negocios. Estas vistas multidimensionales proveen la base técnica para los cálculos y el análisis requerido por las aplicaciones de Inteligencia para Negocios.

En general, estos sistemas OLAP deben:

- Soportar requerimientos complejos de análisis
- Analizar datos desde diferentes perspectivas
- Soportar análisis complejos contra un volumen ingente de datos

La funcionalidad de los sistemas OLAP se caracteriza por ser un análisis multidimensional de datos corporativos, que soportan los análisis del usuario y unas posibilidades de navegación, seleccionando la información a obtener.

Normalmente este tipo de selecciones se ve reflejada en la visualización de la estructura multidimensional, en unos campos de selección que nos permitan elegir el nivel de agregación (jerarquía) de la dimensión, y/o la elección de un dato en concreto, la visualización de los atributos del sujeto, frente a una(s) dimensiones en modo tabla, pudiendo con ello realizar, entre otras las siguientes acciones:

Rotar (Swap): alterar las filas por columnas (permutar dos dimensiones de análisis)

Bajar (Down): bajar el nivel de visualización en las filas a una jerarquía inferior

Detallar (*Drilldown*): informar para una fila en concreto, de datos a un nivel inferior

Expandir (Expand): id. anterior sin perder la información a nivel superior para éste y el resto de los valores

Colapsar (Collapse): operación inversa de la anterior.

Existen dos arquitecturas diferentes para los sistemas OLAP: OLAP multidimensional (MOLAP) y OLAP relacionales (ROLAP).

ROLAP (Relational OLAP)

Un conjunto de interfaces de usuario y aplicaciones que dan a las bases de datos relacionales el sabor dimensional.

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

El sistema ROLAP utiliza una arquitectura de tres niveles. La base de datos relacional maneja los requerimientos de almacenamiento de datos, y el motor ROLAP proporciona la funcionalidad analítica.

- El nivel de base de datos usa bases de datos relacionales para el manejo, acceso y obtención del dato.
- El nivel de aplicación es el motor que ejecuta las consultas multidimensionales de los usuarios.
- El motor ROLAP se integra con niveles de presentación, a través de los cuales los usuarios realizan los análisis OLAP.

Después de que el modelo de datos para el Data Warehouse se ha definido, los datos se cargan desde el sistema operacional. Se ejecutan rutinas de bases de datos para agregar el dato, si así es requerido por el modelos de datos.

Se crean entonces los índices para optimizar los tiempos de acceso a las consultas.

Los usuarios finales ejecutan sus análisis multidimensionales, a través del motor ROLAP, que transforma dinámicamente sus consultas a consultas SQL. Se ejecutan estas consultas SQL en las bases de datos relacionales, y sus resultados se relacionan mediante tablas cruzadas y conjuntos multidimensionales para devolver los resultados a los usuarios.

La arquitectura ROLAP es capaz de usar datos precalculados si estos están disponibles, o de generar dinámicamente los resultados desde los datos elementales si es preciso. Esta arquitectura accede directamente a los datos del Data Warehouse, y soporta técnicas de optimización de accesos para acelerar las consultas. Estas optimizaciones son, entre otras, particionado de los datos a nivel de aplicación, soporte a la desnormalización y joins múltiples.

MOLAP (multidimensional OLAP)

Conjunto de interfaces de usuario, aplicaciones, y tecnologías propietarias de bases de datos que tienen un fuerte sabor dimensional.

La arquitectura MOLAP usa unas bases de datos multidimensionales para proporcionar el análisis, su principal premisa es que el OLAP está mejor implantado almacenando los datos multidimensionalmente a diferencia del ROLAP.

El sistema MOLAP utiliza una arquitectura de dos niveles: La bases de datos multidimensionales y el motor analítico.

- La base de datos multidimensional es la encargada del manejo, acceso y obtención del dato.
- El nivel de aplicación es el responsable de la ejecución de los requerimientos OLAP. El nivel de presentación se integra con el de aplicación y proporciona un interfaz a través del cual los usuarios finales visualizan los análisis OLAP. Una arquitectura cliente/servidor permite a varios usuarios acceder a la misma base de datos multidimensional.

La información procedente de los sistemas operacionales, se carga en el sistema MOLAP, mediante una serie de rutinas batch. Una vez cargado el dato elemental en la Base de Datos multidimensional (MDDDB), se realizan una serie de cálculos en batch, para calcular los datos agregados, a través de las dimensiones de negocio, rellenando la estructura MDDDB.

Tras rellenar esta estructura, se generan unos índices y algoritmos de tablas hash para mejorar los tiempos de accesos a las consultas.

Una vez que el proceso de compilación se ha acabado, la MDDDB está lista para su uso. Los usuarios solicitan informes a través del interfase, y la lógica de aplicación de la MDDDB obtiene los datos.

La arquitectura MOLAP requiere de cálculos intensivos de compilación. Lee de datos precompilados, y tiene capacidades limitadas de crear agregaciones dinámicamente o de hallar ratios que no se hayan precalculados y almacenados previamente.

COMPARACIÓN ENTRE ROLAP Y MOLAP

Cuando se comparan las dos arquitecturas, se pueden hacer las siguientes observaciones:

- El ROLAP delega la negociación entre tiempo de respuesta y el proceso batch al diseño del sistema. Mientras, el MOLAP, suele requerir que sus bases de datos se precompilen para conseguir un rendimiento aceptable en las consultas, incrementando, por tanto los requerimientos batch.
- Los sistemas con alta volatilidad de los datos (aquellos en los que cambian las reglas de agregación y consolidación), requieren una arquitectura que pueda realizar esta consolidación ad-hoc. Los sistemas ROLAP soportan bien esta consolidación dinámica, mientras que los MOLAP están más orientados hacia consolidaciones batch.
- Los ROLAP pueden crecer hasta un gran número de dimensiones, mientras que los MOLAP generalmente son adecuados para diez o menos dimensiones.
- Los ROLAP soportan análisis OLAP contra grandes volúmenes de datos elementales, mientras que los MOLAP se comportan razonablemente en volúmenes más reducidos (menos de 5 Gb)

ALMACENES DE DATOS

Tras las dificultades de los sistemas tradicionales en satisfacer las necesidades informacionales, surge el concepto de Data Warehouse, como solución a las necesidades informacionales globales de la empresa. Este término acuñado por Bill Inmon, se traduce literalmente como Almacén de Datos. No obstante si el Data Warehouse fuese exclusivamente un almacén de datos, los problemas seguirían siendo los mismos que en los Centros de Información [INM92].

La ventaja principal de este tipo de sistemas se basa en su concepto fundamental, la estructura de la información. Este concepto significa el almacenamiento de información homogénea y fiable, en una estructura basada en la consulta y el tratamiento jerarquizado de la misma, y en un entorno diferenciado de los sistemas operacionales. Según definió Bill Inmon, el Data Warehouse se caracteriza por ser:

Integrado: los datos almacenados en el Data Warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.

Temático: sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del Data Warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.

Histórico: el tiempo es parte implícita de la información contenida en un Data Warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente.

Por el contrario, la información almacenada en el Data Warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el Data Warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.

No volátil: el almacén de información de un Data Warehouse existe para ser leído, y no modificado. La información es por tanto permanente, significando la actualización del Data Warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

Otra característica del Data Warehouse es que contiene datos relativos a los datos, concepto que se ha venido asociando al término de metadatos. Los metadatos permiten mantener información de la procedencia de la información, la periodicidad de refresco, su fiabilidad, forma de cálculo, etc., relativa a los datos de nuestro almacén.

Estos metadatos serán los que permitan simplificar y automatizar la obtención de la información desde los sistemas operacionales a los sistemas informacionales.

Los objetivos que deben cumplir los metadatos, según el colectivo al que va dirigido, serían:

- Soportar al usuario final, ayudándole a acceder al Data Warehouse con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Ayudar a construir consultas, informes y análisis, mediante herramientas de navegación.
- Soportar a los responsables técnicos del Data Warehouse en aspectos de auditoria, gestión de la información histórica, administración del Data Warehouse, elaboración de programas de extracción de la información, especificación de las interfaces para la realimentación a los sistemas operacionales de los resultados obtenidos, etc.

Para comprender el concepto de Data Warehouse, es importante considerar los procesos que lo conforman. A continuación se describen dichos procesos clave en la gestión de un Data Warehouse:

Extracción: obtención de información de las distintas fuentes tanto internas como externas.

Elaboración: filtrado, limpieza, depuración, homogeneización y agrupación de la información.

Carga: organización y actualización de los datos y los metadatos en la base de datos.

Explotación: extracción y análisis de la información en los distintos niveles de agrupación.

MERCADOS DE DATOS

Mercado de Datos (*Data Mart*) es un repositorio de datos que se obtiene de bases de datos operacionales y está diseñado para servir a un área particular de la empresa (ej.: área comercial).

El término Mercado de Datos (*Data Mart*) significa evitar la imposibilidad de atacar el trabajo de planear el almacén de datos completo de la empresa [KIM00] .

El almacén de datos no es más que la unión de los mercados de datos que lo componen.

Según define Meta Group, "un Data Mart es una aplicación de Data Warehouse, construida rápidamente para soportar una línea de negocio simple". Los Data Marts, tienen las mismas características de integración, no volatilidad,

orientación temática que el Data Warehouse. Representan una estrategia de "divide y vencerás" para ámbitos muy genéricos de un Data Warehouse.

Esta estrategia es particularmente apropiada cuando el Data Warehouse central crece muy rápidamente y los distintos departamentos requieren sólo una pequeña porción de los datos contenidos en él. La creación de estos Data Marts requiere algo más que una simple réplica de los datos: se necesitarán tanto la segmentación como algunos métodos adicionales de consolidación.

CAPÍTULO 2. PLANTEAMIENTO DEL PROBLEMA

En la empresa se ha identificado un problema que radica básicamente en la falta de una completa integración entre el procesamiento de información que hace el sistema ERP llamado R/3 de SAP y la base de datos que éste actualiza con la función de la planeación estratégica y la toma de decisiones de la alta dirección. Este problema consiste en los siguientes aspectos.

Actualmente toda la información de la empresa se actualiza en línea sin embargo, todos estos datos no son directamente utilizables por los ejecutivos de la dirección porque aunque el sistema cuenta con reportes predefinidos, el sistema no cuenta con las herramientas de consolidación de información y análisis a la medida de los usuarios, sino que deben extraerse manualmente los datos del sistema R/3 de SAP para crear reportes en Excel de modo que los ejecutivos puedan llevar a cabo su función de planeación estratégica.

Otro problema derivado del anterior es la gran demanda de recursos de procesamiento (Memoria RAM y CPU del servidor de SAP) que ocasiona la labor de extracción manual de los datos del sistema ERP para la elaboración de los reportes mencionados anteriormente, afectando el desempeño global del sistema. Esto causa que los procesos críticos de la compañía se vean afectados, tales como la generación e impresión de facturas, el reporte de la producción, la captura de pedidos de clientes, etc. Este impacto negativo en el desempeño del sistema se debe a que constantemente los usuarios del área comercial deben correr consultas que abarcan muchos registros de la base de datos de SAP.

Por toda la problemática descrita anteriormente, es muy conveniente implementar un mercado de datos que permita se realicen análisis multidimensionales directamente por los ejecutivos utilizando herramientas apropiadas que permitan hacer análisis multidimensionales, de tal manera que no solamente se eviten los problemas mencionados con anterioridad sino que además se optimice la eficiencia en el desempeño del software R/3 de SAP.

La solución ideal a la problemática planteada sería la implementación de un almacén de datos para abarcar toda la información de la empresa, sin embargo no es accesible por su alto costo y porque hay otras prioridades en la empresa, es por eso que el presente proyecto consiste en un mercado de datos cuya implementación sería de un costo bastante razonable y que puede proporcionar una gran experiencia antes de buscar una solución ideal.

Sin embargo, la exitosa implementación del mercado de datos depende también de que se puedan resolver los siguiente problemas con la extracción de información de la base de datos de SAP:

- Los datos están normalizados en miles de tablas dentro de SAP, lo que complicaría enormemente hacer joins suficientes para reunir información.
- No hay documentación suficiente, o la que existe en muchas de las ocasiones están en mnemónicos o abreviaciones del Alemán.
- Muchos de los datos dentro de la base de datos de SAP, están en un formato interno propietario, lo que complica la extracción directa de los mismos mediante herramientas de terceros.
- En muchos casos aunque se trate de un mismo valor de una tabla a otra a veces el tipo de datos no lo es.

Por las razones descritas anteriormente el presente proyecto será enfocado al diseño del mercado de datos, dejando la implementación para un proyecto posterior una vez resueltos los problemas de extracción de información.

Para poder llevar a cabo un diseño adecuado del mercado de datos es necesario aclarar las actividades del área comercial así como la estructura organizacional de empresa con respecto a las ventas.

El área comercial es la encargada de generar los pedidos de los clientes, es decir hacer labor de ventas con ellos para que se generen nuevos pedidos, también dar de alta a las escalas de descuentos para los diferentes productos, participan también en conjunto con otras áreas como mercadotecnia para definir los nuevos precios de los productos y promociones, y con el área de producción para definir los planes de producción, entre otras cosas.

También hay personal del área de ventas encargado de hacer los informes que las áreas directivas de la empresa necesitan para tomar decisiones y para ello, el personal hace extracciones manuales del sistema SAP y los resume con el formato necesario.

CAPITULO 3. SOLUCIÓN AL PROBLEMA COMPUTACIONAL

Para llevar a cabo el diseño del mercado de datos de ventas nos hemos basado en las técnicas y la metodología del Dr. Ralph Kimball [KIM00]. Para seguir tal metodología y llegar al modelo dimensional, a continuación se describe la información del área de ventas, en este caso la estructura organizacional de ventas de la compañía y la jerarquía de productos, para con esto llegar al modelo dimensional y a el o los diagramas de estrella.

En el proyecto de implantación del sistema SAP en la empresa se definió la estructura de ventas que se muestra a continuación, ver figura 3.

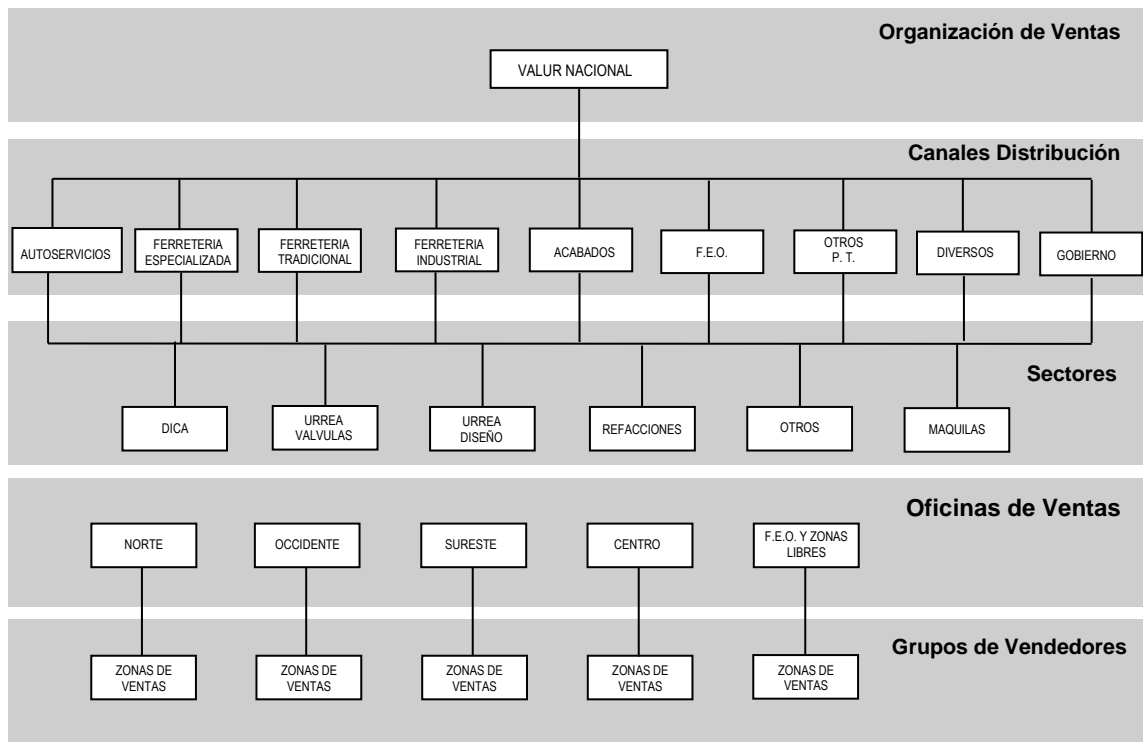


Figura 3. Estructura de Ventas del Grupo Ureca, División Grifería y Válvulas.

Como se muestra en la figura 3, la estructura de ventas de la compañía se estableció en función de los siguientes criterios: clasificación de los clientes, familias de productos, y por la división geográfica. El significado de cada nivel de la estructura de Ventas es el siguiente:

- **Canales de Distribución:** Se clasifican a los clientes por los tipos de productos que consumen o comercializan ya que algunos son distribuidores propios.
- **Sectores:** Es la clasificación interna que existe para diferenciar nuestros productos, ya que van orientados a diferentes mercados, como son por ejemplo los productos para la construcción o para la industria.
- **Oficinas de Ventas:** Para las ventas nacionales se instaló oficinas de ventas en puntos estratégicos según la situación geográfica y de mercado del país para atender de manera eficiente a todos nuestros clientes nacionales.
- **Grupos de Vendedores:** en cada oficina de ventas existe además otro nivel de subdivisión denominado Grupos de Vendedores o también Zonas de Ventas, que varían en número según la situación geográfica del país y las necesidades de los clientes de las diferentes oficinas de ventas.

Además de la estructura organizacional definida anteriormente del área de ventas, también se describe a continuación la jerarquía de productos que se manejan en la empresa (ver figura 4); que consiste en la clasificación de nuestros productos en familias según la estructura organizacional de ventas y los materiales utilizados para fabricarlas así como el mercado al que van dirigidos.

En la figura 4, la jerarquía se expresa de izquierda a derecha y corresponden a los Sectores, Líneas y Familias. Todas las consultas que los usuarios analistas requieren tienen dependen de la estructura organizacional de ventas y la jerarquía de productos.

Por esta razón, a fin de que la información que resulte del mercado de datos sea congruente y satisfaga las necesidades de los analistas.

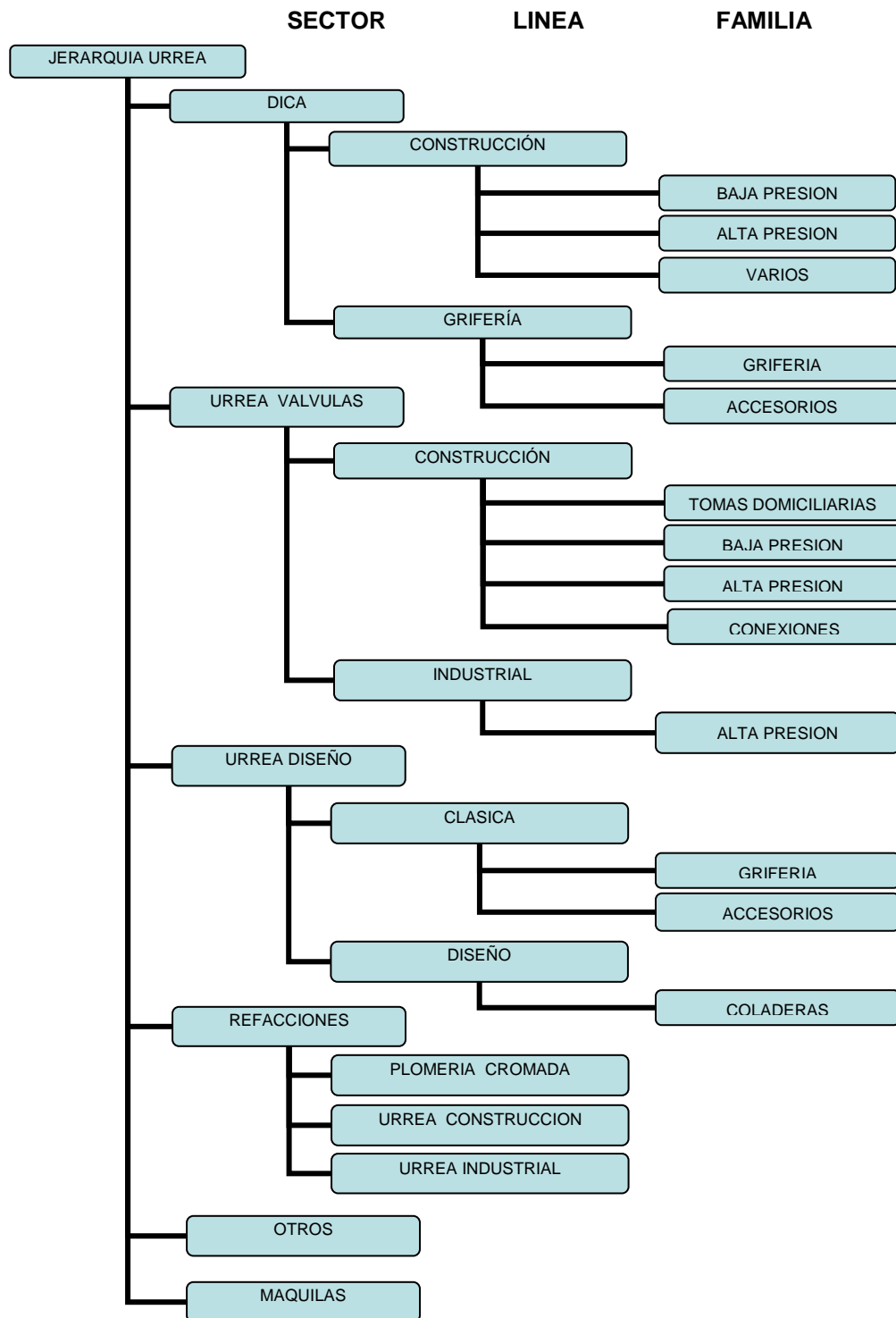


Figura 4. Jerarquía de productos del Grupo Urrea, División Grifería y Válvulas

De acuerdo a la metodología de Kimball, los pasos que deben seguirse para el diseño del mercado de datos son los siguientes:

- 1) Definir los mercados de datos que se consideren factibles de acuerdo a los alcances del proyecto, en este caso el área comercial.
- 2) Definir las dimensiones del mercado de datos e identificar el nivel más básico de información con el que se va a trabajar.
- 3) Deducir las dimensiones.
- 4) Especificar los hechos, correspondientes a los datos cuantificables al más profundo nivel posible, como por ejemplo el importe total de facturación, la cantidad de productos vendidos , etc.
- 5) Obtener los diagramas de dimensiones y las tablas de hechos que servirán como fundamento para el diseño lógico y físico de la base de datos donde se implementará el mercado de datos.

Los diagramas correspondientes a la metodología descrita anteriormente se presentan a continuación, iniciando con el esquema sobre las dimensiones que serán consideradas para este proyecto.

3.1 DIMENSIONES DEL MERCADO DE DATOS PARA VENTAS

Las dimensiones identificadas para el área comercial de la empresa se muestran a continuación en la figura 5.

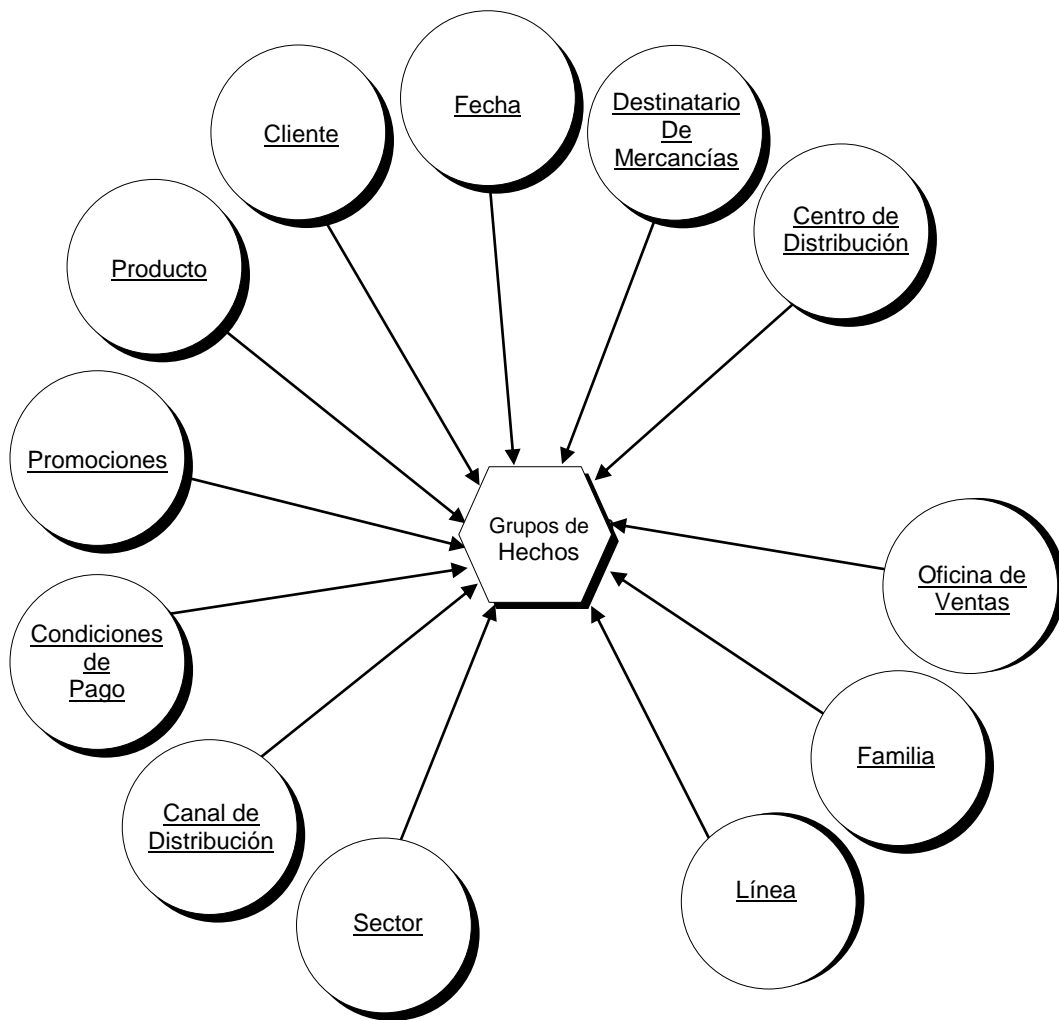


Figura 5. Diagrama de Dimensiones

En la siguiente tabla se describen con más detalle las dimensiones que se identificaron en la figura 5.

Descripción de las Dimensiones

Nombre	Descripción
Fecha	Contiene todos los atributos asociados con la fecha en que se dan las actividades.
Cliente	Contiene todos los atributos asociados con los clientes.
Producto (material)	Contiene información acerca de los productos que se venden.
Promociones	Describe las promociones aplicadas a los clientes cuando hacen sus pedidos.
Condiciones de Pago	Condiciones de pago en forma de tipos de descuento y plazos de pago.
Canal de Distribución	Canal de distribución que permite clasificar la entrega de la mercancía.
Sector	Agrupación de materiales o productos. Por mediación del sector, se determina el área de ventas y la división a las que se ha asignado un material o producto.
Línea	Clasificación de productos según el mercado al que van dirigidos.
Familia	Subclasificación de las líneas según el tipo de producto.
Oficina de Ventas	Delegación (p. ej. una sucursal) responsable de la comercialización de determinados productos y servicios en una determinada zona geográfica.
Centro de Distribución	Contiene los atributos referentes al centro de distribución de donde se embarca el producto.

Nombre	Descripción
Destinatario de Mercancías	Datos de ubicación del cliente quien recibirá los productos.

A continuación se muestran los diagramas referentes a cada una de las dimensiones mencionadas anteriormente desglosando los atributos que las componen.

Detalles de la Dimensión Fecha

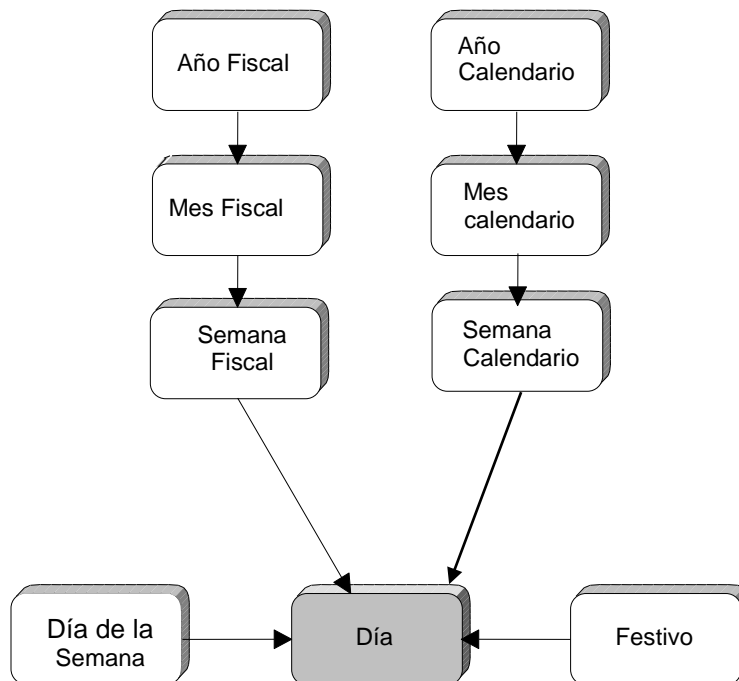


Figura 6. Diagrama detallado de la Dimensión Fecha

Descripción de los Atributos de la Dimensión Fecha

Nombre Atributo	Descripción	Valores de Ejemplo
Día	El día específico cuando tuvo lugar la actividad.	09/02/2000
Día de la Semana	El día específico de la semana	Lunes, Martes
Festivo	Identifica que ese día es un día festivo	Navidad, Día de la Raza, etc.
Semana Calendario	La fecha de fin de semana, siempre sábado. Note que FS denota Fin de Semana	FS 06/06/1998; FS 06/13/1998
Mes Calendario	El mes calendario	Febrero, 2000, Marzo, 2003
Año Calendario	El año calendario	2003
Semana Fiscal	Semana fiscal	
Mes Fiscal	Periodo fiscal	
Año Fiscal	La agrupación de 52 semanas fiscales o 12 meses fiscales que completan el año fiscal	2002, 2003

Detalles de la Dimensión Cliente

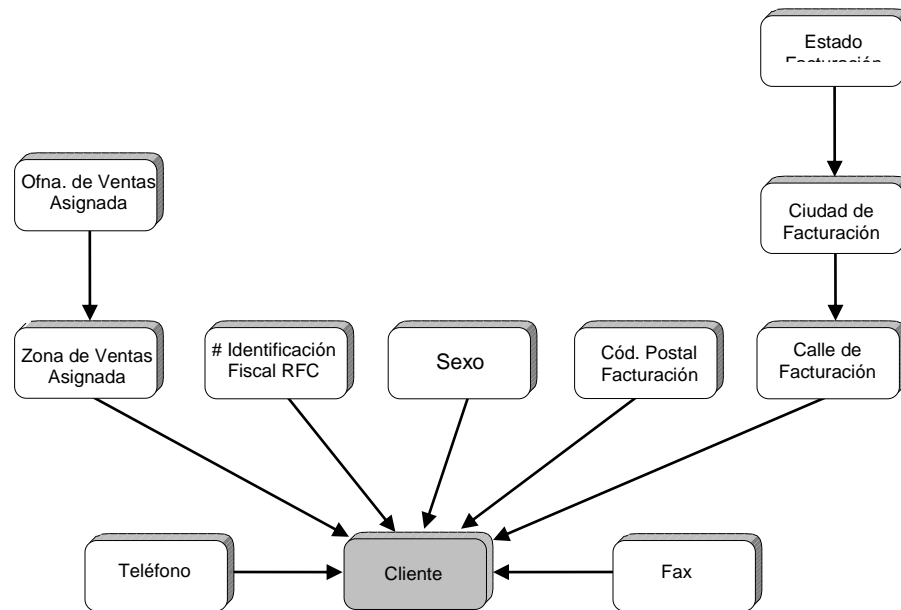


Figura 7. Diagrama detallado de la Dimensión Cliente

Descripción de los Atributos de la Dimensión Cliente

Nombre Atributo	Descripción	Valores de Ejemplo
Punto de entrega para el cliente	Especifica el destino de las mercancías.	
Prefijo	Clave que se antepone a un número telefónico según la localidad	
Clave Lada	Clave que identifica a una localidad para cuando se hacen llamadas de Larga distancia.	Guadalajara =33, DF 55, etc.
Localidad	Nombre del poblado o municipio	Zapopan, Tlaquepaque, etc.

Nombre Atributo	Descripción	Valores de Ejemplo
Ciudad	Nombre de la ciudad del cliente	Guadalajara, Zacatecas, etc.
Zona de Ventas Asignada	Clasificación del cliente según la región donde se encuentra.	Norte, Pacífico, Centro, etc.
Código Postal		
Calle	Nombre de la calle del domicilio del cliente.	
Estado	Entidad federativa donde se encuentra el cliente	Chihuahua, Monterrey, Jalisco, etc.
Cliente	Nombre del cliente	
Sexo	Identifica el género	Masculino, Femenino
Oficina de Ventas Asignada	Clasificación según la regiones donde radica el cliente.	
Fecha de Nacimiento	Fecha de nacimiento del cliente	
Nivel de Escolaridad	Nivel académico del cliente	
Calle Facturación	Nombre de la calle donde se factura	
Ciudad Facturación	Nombre de la ciudad donde se factura al cliente.	
Código Postal Facturación	Código postal del domicilio de facturación	
Estado Facturación	Estado o entidad del domicilio de Facturación del cliente	
Nº Identificación Fiscal	Registro Federal de	MASE750523PV4

Nombre Atributo	Descripción	Valores de Ejemplo
RFC	Contribuyente	

Detalles de la Dimensión Producto

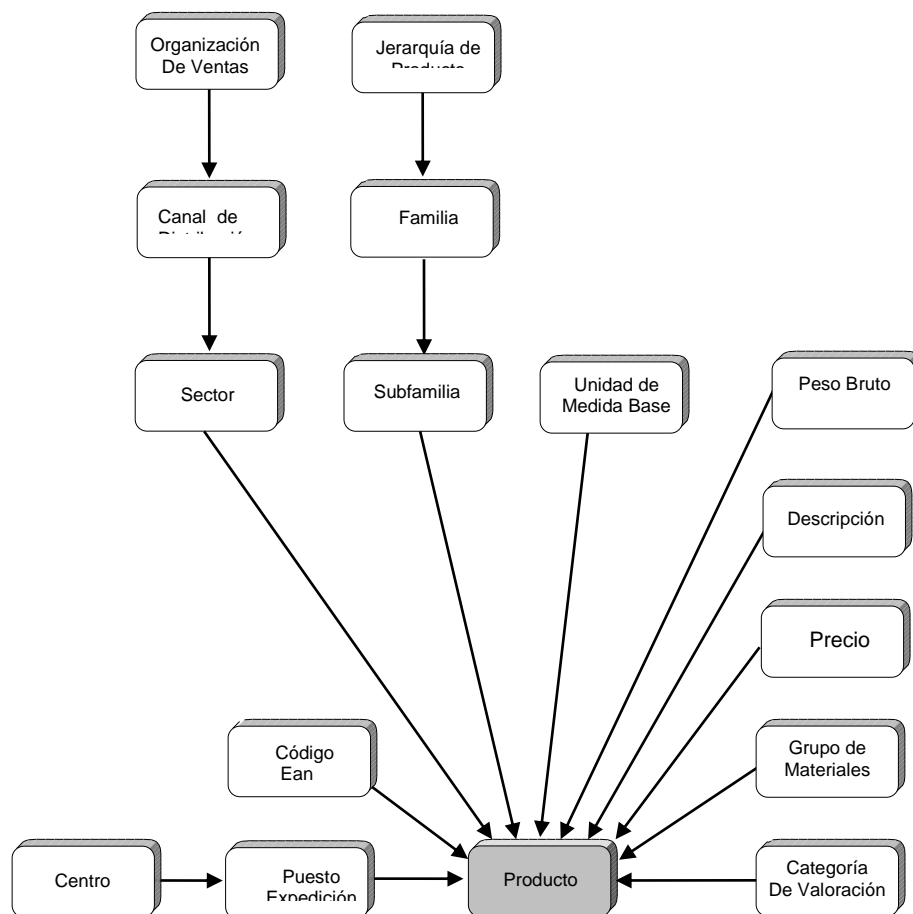


Figura 8. Diagrama detallado de la Dimensión Producto

Descripción de los Atributos de la Dimensión Producto

Nombre Atributo	Descripción	Valores de Ejemplo
Producto	Código que identifica a cada producto.	46, 9471Z.1,320LB
Descripción	Descripción del material	PORTA CEPILLERO
Grupo de Materiales	Clasificación para agrupar materiales según su embalaje.	001 – Cajas 002 - Granel
Código EAN	Código que se imprime en etiquetas de código de Barras y que es único para cada producto.	7501973704474
Categoría de Valoración	Sirve para clasificar los valores de stock del mismo tipo en determinadas cuentas contables.	Z006 – Producto Terminado
Precio	Precio del producto.	
Centro	Identifica de cual centro se expide el producto.	1120 - Guadalajara
Puesto de Expedición	Lugar donde se procesan las entregas al hacer la expedición.	1114 – Pto. Exp. León
Unidad de Medida Base	Unidad de medida en que se gestiona el stock de material	PZA – Pieza, KG Kilos
Peso Bruto	Peso Bruto del Producto	0.723 Kg
Sector	Agrupación de materiales o productos. Por mediación	01 – DICA, 02 – URREA

Nombre Atributo	Descripción	Valores de Ejemplo
	del sector, se determina el área de ventas y la división a las que se ha asignado un material o producto.	VALVULAS
Línea	Clasificación de productos según al mercado al que van dirigidos.	URREA INDUSTRIAL
Familia	Subclasificación de las líneas según el tipo de producto.	COLADERAS, REFACCIONES P. CONSTRUCCION
Canal de Distribución	Canal de distribución que permite clasificar la entrega de la mercancía.	02 – FERRETERIA ESPECIALIZADA 03 – FERRETERIA TRADICIONAL
Oficina de Ventas	Agrupación central de las ventas de todas las regiones y zonas de venta.	1101 – VALUR NACIONAL
Jerarquía de Producto	Secuencia alfanumérica de caracteres que sirve para agrupar materiales mediante la combinación de diferentes características y se utiliza para evaluaciones y determinaciones de precios.	12010401101040 2000

Detalles de la Dimensión Promociones

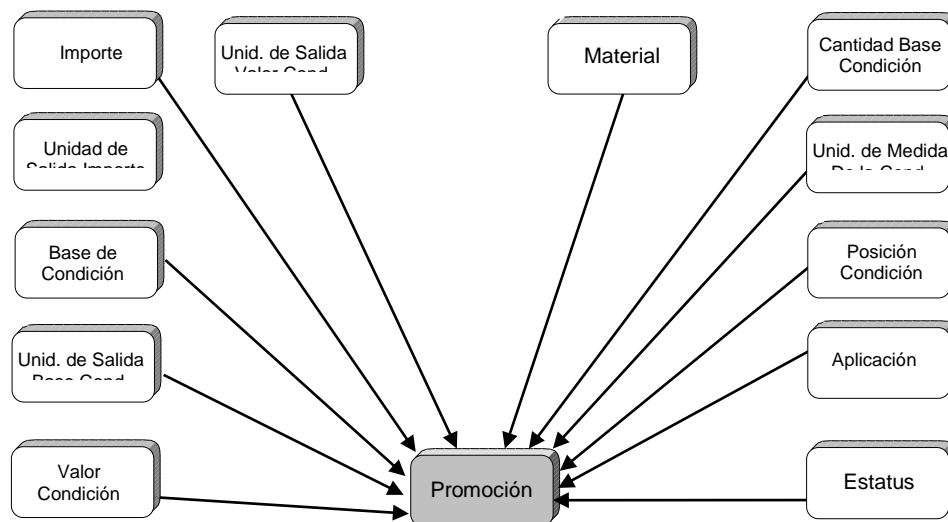


Figura 9. Diagrama detallado de la Dimensión Promociones

Descripción de los Atributos de la Dimensión Promociones

Nombre Atributo	Descripción	Valores de Ejemplo
Código de Promoción	Código que identifica a cada promoción.	PR00, K030
Estatus	Indica si la promoción está activa o no para poder usarla	
Aplicación	Indica si la condición aplica para Ventas o para Compras	Sí,No
Posición Condición	Indica el No de Posición en el pedido donde se agregó la promoción	
Unidad de Medida de la Condición	Unidad de medida de referencia para el importe de la condición.	PZA, %

Cantidad Base Cond.	Cantidad base con la cual se determina el importe o porcentaje.	PZA
Código de Material	Código que identifica al producto asociado a la promoción.	46, 9471Z.1,320LB
Unidad de Salida Valor Cond.	Unidad con la cual se imprime el importe de una condición	
Importe	Importe o porcentaje de la promoción	
Unidad de Salida Importe	Unidad con la cual se imprime el importe de una condición	
Base de Condición	Valor base que el sistema utiliza para calcular esta condición.	
Unidad de salida Base Cond.	Unidad con la cual se imprime el importe de una condición	
Valor Condición	Valor resultante de la condición para una determinada condición, suma o subtotal en un documento.	

Detalles de la Dimensión Condiciones de Pago

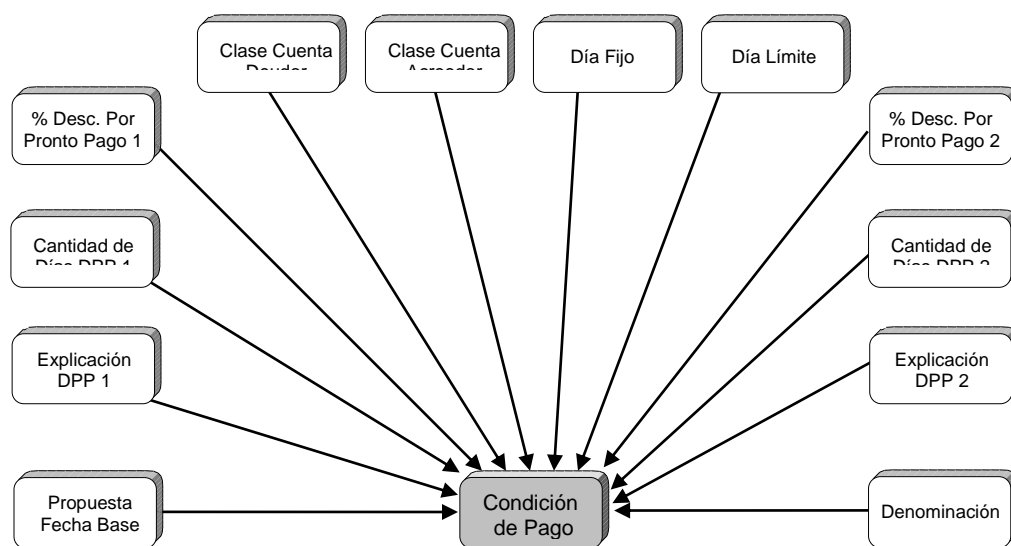


Figura 10. Diagrama detallado de la Dimensión Condiciones de Pago

Descripción de los Atributos de la Dimensión Condiciones de Pago

Nombre Atributo	Descripción	Valores de Ejemplo
Clave Condición de Pago	Código que identifica a cada condición de pago.	Z009,ZCC1
Denominación	Descripción de las condiciones de pago	Pagos a plazos 30, 60 ,90 días
Día límite	Día del mes hasta el que son válidas las correspondientes condiciones.	
Clase de Cuenta Deudor	Indicador que determina si la condición aplica a Deudores	Sí, No
Clase de Cuenta Acreedor	Indicador que determina si la condición aplica a Acreedores	Sí, No
Día Fijo	Día natural para fecha base de plazo de pago.	

Nombre Atributo	Descripción	Valores de Ejemplo
% Desc. Por Pronto Pago 1	Porcentaje de descuento por pronto pago que se concede al pagar dentro del plazo previsto.	
Cantidad de Días DPP 1	Cantidad de días que transcurren después de la fecha base para plazo de pago.	
Explicación DPP 1	Descripción del descuento por pronto pago 1.	
% Desc. Por Pronto Pago 2	Porcentaje de descuento por pronto pago que se concede al pagar dentro del plazo previsto.	
Cantidad de Días DPP 2	Cantidad de días que transcurren después de la fecha base para plazo de pago.	
Propuesta Fecha Base	Especifica como será la propuesta del sistema en el campo de la fecha base para plazo de pago al entrar un documento.	Sin Propuesta, Fecha de Contabilización, Fecha de documento o Fecha de entrada

Detalles de la Dimensión Canal de Distribución

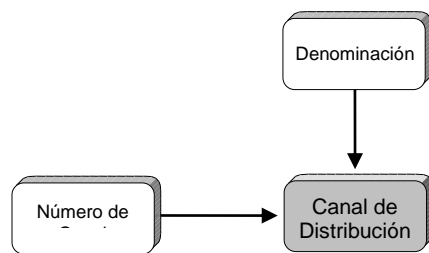


Figura 11. Diagrama detallado de la Dimensión Canal de Distribución

Descripción de los Atributos de la Dimensión Canal de Distribución

Nombre Atributo	Descripción	Valores de Ejemplo
Número de Canal	Número que identifica al Canal de distribución	01, 02
Denominación	Descripción del canal de distribución.	Autoservicio, Ferretería especializada.

Detalles de la Dimensión Sector

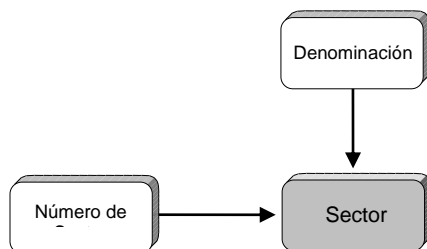


Figura 12. Diagrama detallado de la Dimensión Sector

Descripción de los Atributos de la Dimensión Sector

Nombre Atributo	Descripción	Valores de Ejemplo
Número de Sector	Número que identifica al Sector	01, 02
Denominación	Descripción del Sector.	DICA, URREA VÁLVULAS.

Detalles de la Dimensión Línea

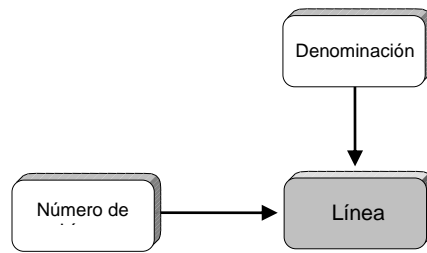


Figura 13. Diagrama detallado de la Dimensión Línea

Descripción de los Atributos de la Dimensión Línea

Nombre Atributo	Descripción	Valores de Ejemplo
Número de Línea	Número que identifica a la Línea	01, 02
Denominación	Descripción de la Línea.	PLOMERÍA CROMADA

Detalles de la Dimensión Familia

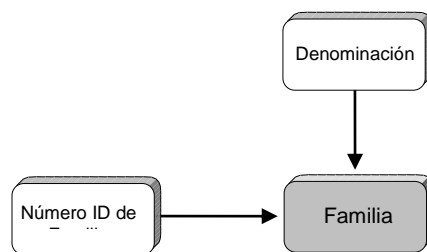


Figura 14. Diagrama detallado de la Dimensión Familia

Descripción de los Atributos de la Dimensión Familia

Nombre Atributo	Descripción	Valores de Ejemplo
Número ID de Familia	Número que identifica a la Familia	01, 02
Denominación	Descripción de la familia.	BAJA PRESION, ALTA PRESION

Detalles de la Dimensión Oficina de Ventas

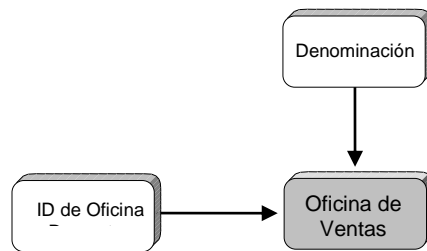


Figura 15. Diagrama detallado de la Dimensión Oficina de Ventas

Descripción de los Atributos de la Dimensión Oficina de Ventas

Nombre Atributo	Descripción	Valores de Ejemplo
ID de la Oficina de Ventas	Clave que identifica a la Oficina de Ventas	0300
Denominación	Clasificación según la regiones donde radica el cliente.	NORTE, OCCIDENTE

Detalles de la Dimensión Destinatario de Mercancías

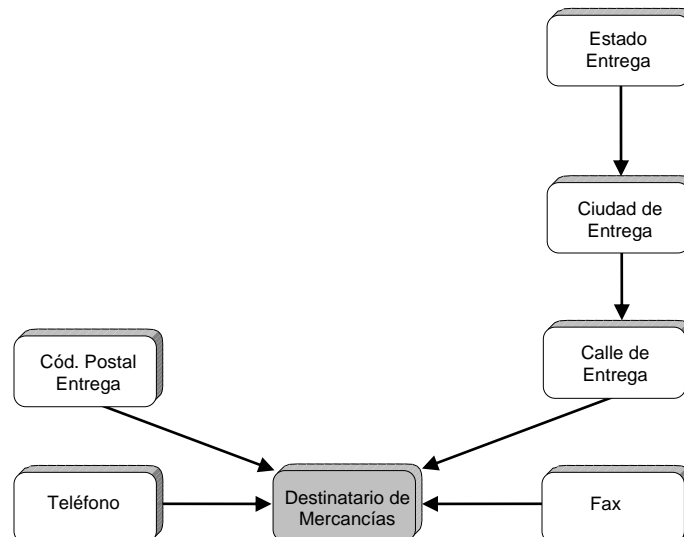


Figura 16. Diagrama detallado de la Dimensión Destinatario de Mercancías

Descripción de los Atributos de la Dimensión Destinatario de Mercancías

Nombre Atributo	Descripción	Valores de Ejemplo
Teléfono	Numero de teléfono del cliente quien es el destinatario de mercancías	

Nombre Atributo	Descripción	Valores de Ejemplo
Fax	Número de Fax del destinatario de mercancías.	
Código Postal Entrega	Código Postal del destinatario de mercancías.	45160
Calle entrega	Nombre de la calle a donde se entregan las mercancías.	
Ciudad	Nombre de la ciudad del destinatario de mercancías.	Guadalajara, Zacatecas, etc.
Estado entrega	Nombre del estado en donde reside el destinatario de mercancías	Jalisco, Michoacán, etc.

3.2 DETALLES DE HECHOS

Para el presente proyecto se han considerado los dos hechos más importantes para el área comercial, estos son Pedidos y Facturación porque para los analistas del área comercial, estos son los más importantes .

La suma de los pedidos recibidos expresados en pesos se evalúa contra las metas mensuales de ventas que es uno de los indicadores más importantes de toda la empresa. Otro análisis consiste en medir cuantos de los pedidos recibidos son procesados hasta llegar a la facturación.

A continuación se detalla el diseño de cada uno de los hechos.

Tabla de Hechos de Pedidos

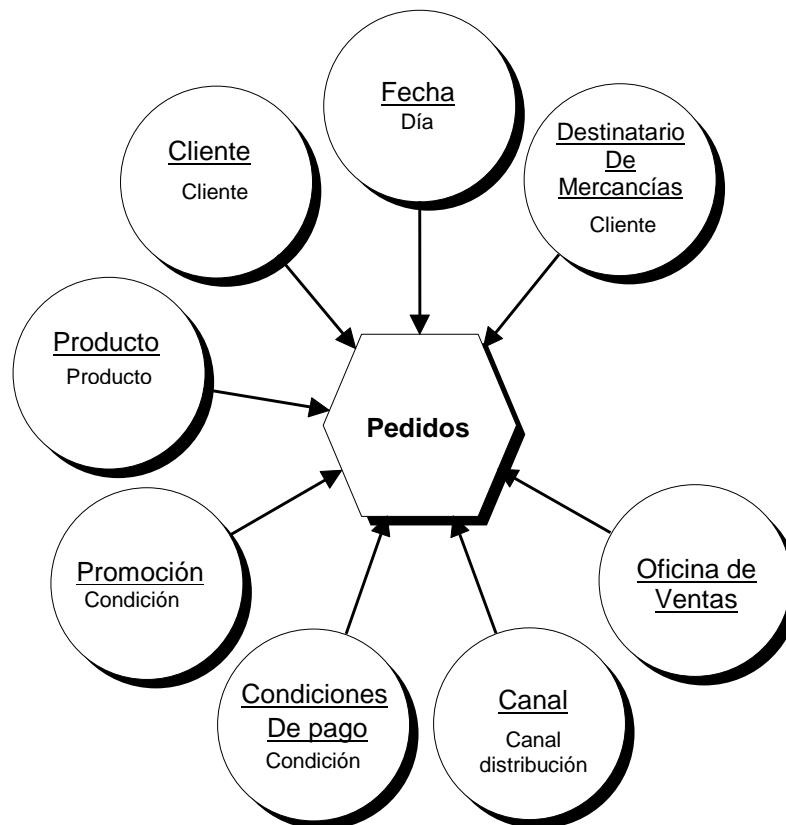


Figura 17. Diagrama de Tabla de Hechos de Pedidos

Descripción de los hechos de Pedidos

Nombre del Hecho	Descripción	Regla de cálculo por defecto
Pedido	Número de Pedido	N/A
Cantidad	Cantidad de materiales en la partida	Suma
Total Posición	Importe total de la partida	Suma
Precio	Precio base del material	N/A
Descuento	Descuento aplicado a la partida	Suma
Peso Neto	Peso Bruto – Peso del embalaje	
Peso Bruto	Peso neto + Peso del embalaje	
IVA	Impuesto al valor agregado	Suma

Tabla de Hechos de Facturación

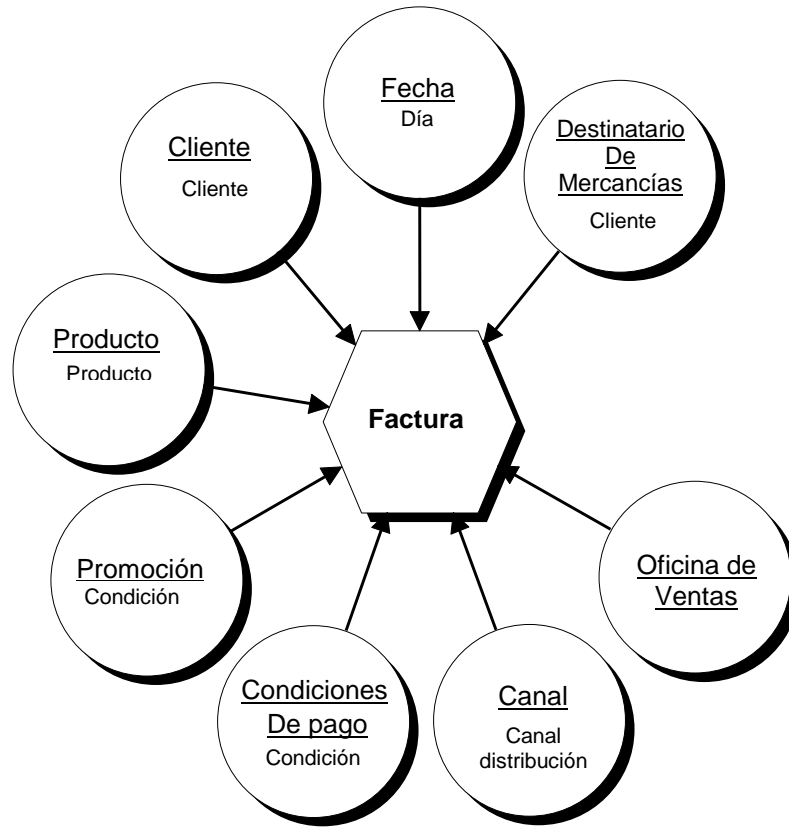


Figura 19. Diagrama de Tabla de Hechos de Facturación

Descripción de los hechos de Facturación

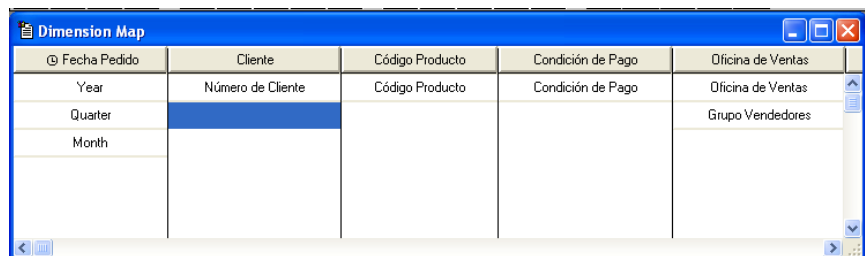
Nombre del Hecho	Descripción	Regla de cálculo por defecto
Factura	Número de Pedido	N/A
Cantidad	Cantidad de materiales en la partida	Suma
Total Posición	Importe total de la partida	Suma
Precio	Precio base del material	N/A

3.3 DEMOSTRACIÓN DEL MODELO EN POWER PLAY

Para efectos de probar el modelo que hemos diseñado en el presente trabajo, hemos creado una representación de la tabla de hechos de pedidos en formato de Microsoft Access y con esta base de datos implementamos un hipercubo.

En las siguientes impresiones de pantalla de la salida del software Power Play podrán apreciarse los puntos importantes en la implementación del modelo y también podrán apreciarse algunos detalles de la funcionalidad olap del software Power Play.

Para poder crear hipercubos en Power Play, es necesario que primero se construya el modelo con el que se va a trabajar, esto se hace en la herramienta denominada PowerPlay Transformer, ver figura 20.



Fecha Pedido	Ciente	Código Producto	Condición de Pago	Oficina de Ventas
Year	Número de Cliente	Código Producto	Condición de Pago	Oficina de Ventas
Quarter				Grupo Vendedores
Month				

Figura 20. Pantalla del Mapa de Dimensiones

Después que se ha creado lo que se denomina Mapa de Dimensiones, enseguida se puede construir el diagrama para las dimensiones. Ver figura 21.

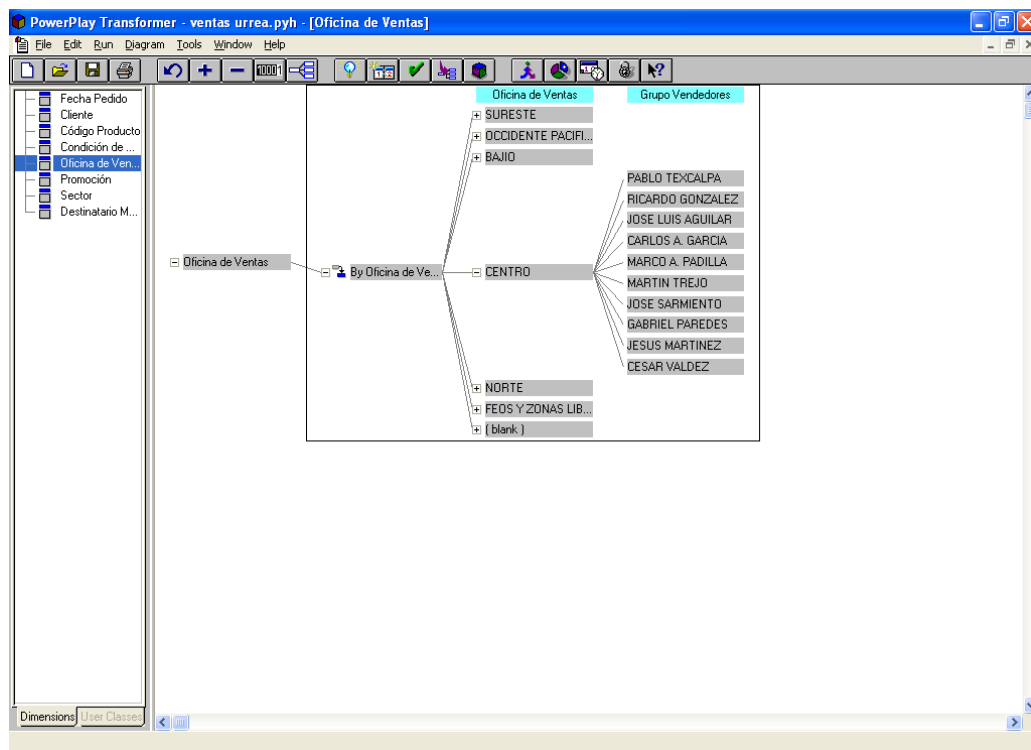


Figura 21. Diagrama de la dimensión Oficina de Ventas

En la figura 22 se muestra el diagrama de la dimensión Fecha, en el cual podemos apreciar algunas de las facilidades que tiene Power Play para utilizar dimensiones de fechas.

Una vez que el modelo ha sido representado en PowerPlay Transformer es necesario crear los hipercubos necesarios, para poder ahora sí hacer análisis OLAP con el PowerPlay. En la figura 22 se muestra una gráfica de código de barras, para ejemplificar algunas de las características del software de Powerplay.

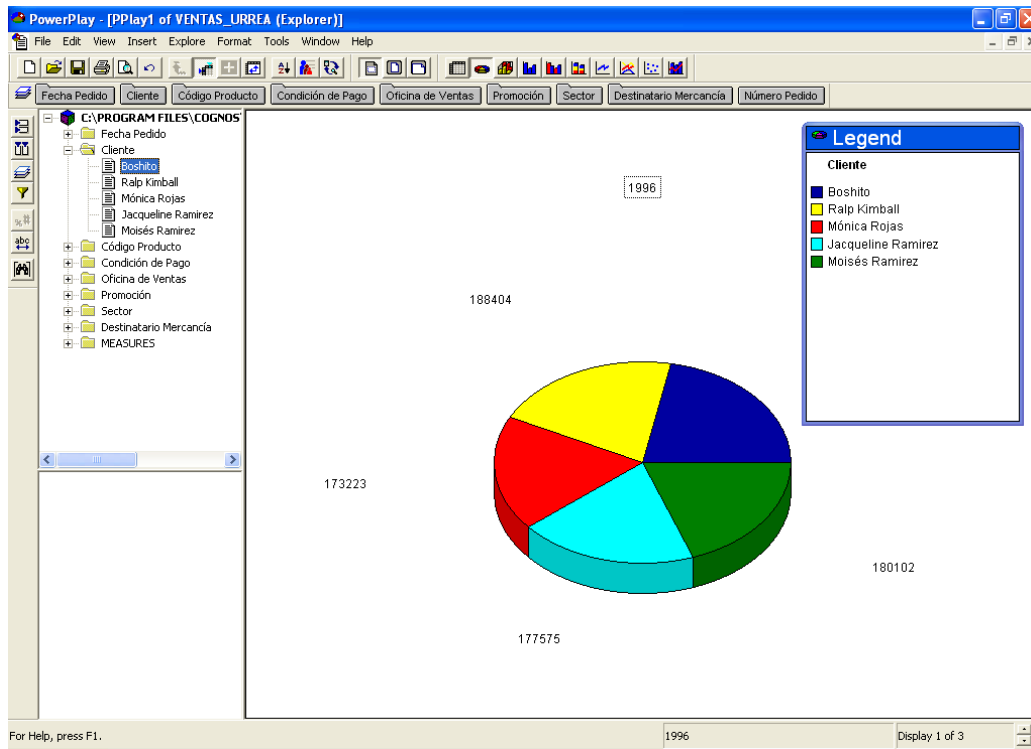


Figura 22. Diagrama deHipercubo para Pedidos

Además de las gráficas de pay es posible utilizar barras, líneas, etc. Independientemente del esquema seleccionado para los gráficos, en cualquiera de ellos es posible hacer drill down , siempre que se haya especificado esta característica para la dimensión específica en el modelo.

En la figura 23 se muestra una gráfica de barras en la cual se ha especificado cierto nivel de filtrado para el año 1996 y el cliente Boshito.

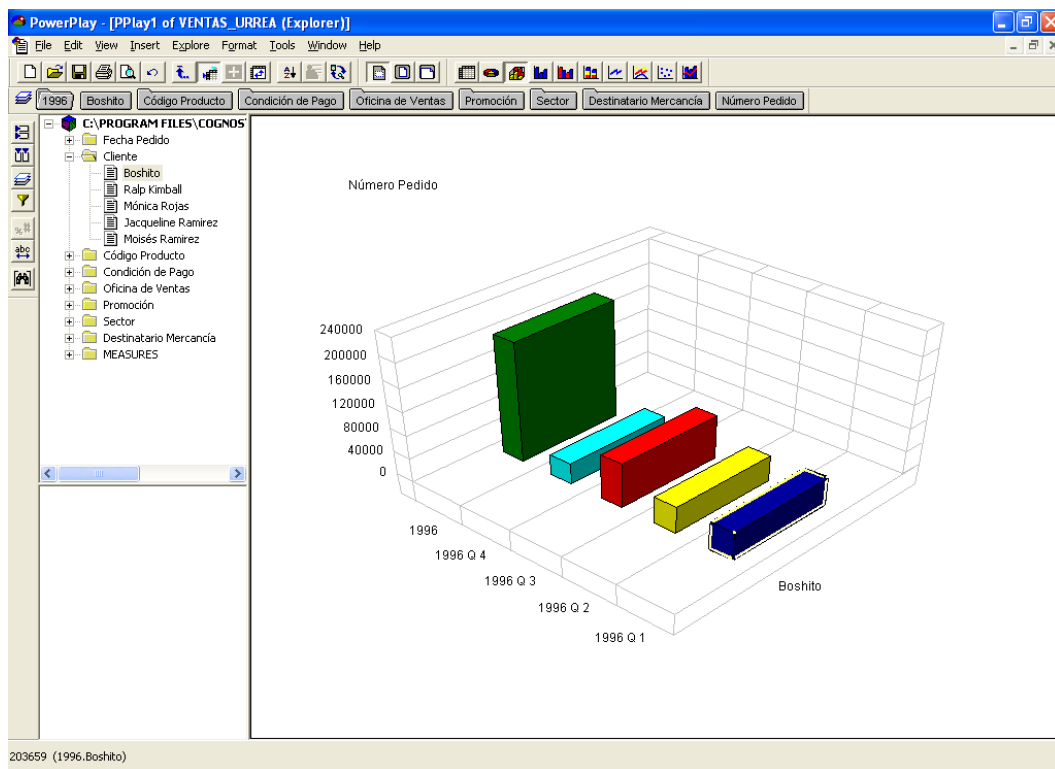


Figura 23. Diagrama deHipercono para Pedidos filtrado por cliente y año

También es posible dar doble click sobre alguna de las barras de alguna gráfica, y si en el modelo fue definido un camino de drill down para la dimensión correspondiente. Ver figura 24

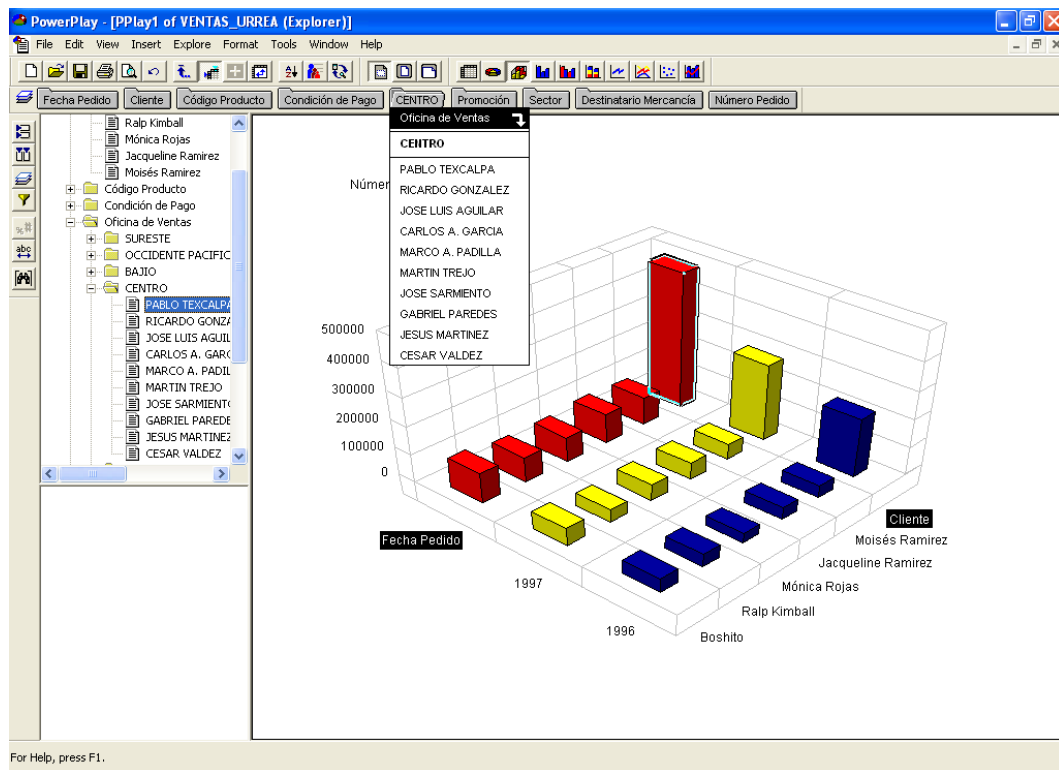


Figura 24. Diagrama deHipercono mostrando el submenú para filtrar por Oficina de Ventas

De esta forma se ha probado el modelo diseñado mediante el hipercono creado en PowerPlay,

CONCLUSIONES

- Se diseñó el Modelo del mercado de datos para el área comercial de la empresa siguiendo la metodología propuesta por Ralph Kimball.
- Mediante la creación de Cubos y reportes en Power Play, se demostró el diseño propuesto.

RECOMENDACIONES

El cumplimiento de los objetivos específicos propuestos en este trabajo están condicionados a que se puedan resolver los problemas de extracción de información de la base de datos de SAP que hemos encontrado. Por lo tanto es necesario evaluar las posibilidades de extracción de información del sistema SAP ya sea con herramientas de terceros o con desarrollos propios.

Aún si no pueden resolverse los problemas mencionados anteriormente, el diseño propuesto definitivamente servirá para futuros proyectos en los que se consideraría la propuesta de SAP para OLAP denominada “*Business Information Warehouse*”.

BIBLIOGRAFÍA

Kimball, R.; Reeves, L.; Ross, M.; Thornthwaite, W. "The Data Warehouse Lifecycle Toolkit"

Data Webhouse , John Willey & Sons. 1998

http://searchsap.techtarget.com/sDefinition/0,,sid21_gci213946,00.html

Definición de ERP , por SAP.

Kimball, Ralph "ERP Vendors: Bring Down Those Walls"

Data Webhouse , Junio 26, 2000, Volumen 3 - Número 10.

Kimball, Ralph "A Dimensional Modeling Manifesto"

DBMS Online, Julio 10, 1997

Kimball, Ralph "The problem with Comparisons"

DBMS Online, - Enero 1996

Kimball, R.; Reeves, L.; Ross, M.; Thornthwaite, W. "The Data Warehouse Lifecycle Toolkit"

Data Webhouse , John Willey & Sons. 1998, Pág. 154

Inmon ,Bill , "Building the data warehouse", QED Press/John Wiley, 1992 *

Burleson, Donald K., "High-Performance Oracle Database Applications"

Consejo Superior de Informática, Ministerio de Administraciones públicas

<http://www.map.es/csi/csi.htm>

<http://www.map.es/csi/silice/DW225.html>

Universidad de Chile, Departamento de Ciencias de la Computación
<http://www.dcc.uchile.cl/~ruribe/cc60k/pres1/pag3.htm>

Inmon ,Bill “SAP and Data Warehousing”, 1999

REFERENCIAS BIBLIOGRÁFICAS

[KIM00] Kimball, Ralph “ERP Vendors: Bring Down Those Walls” Data Webhouse ,Junio 26, 2000, Volumen 3 - Número 10.

[BUR96] Burleson, Donald K., “Hig-Performance Oracle Database Applications”

[FRA97] Franco, Jean-Michel “El data warehouse, el data mining”, Ediciones Gestión 2000, Barcelona 1997.

[KIM96] Kimball, Ralph “The problem with Comparisons”
DBMS Online, - Enero 1996

[INM92] Inmon ,Bill , “Building the data warehouse”, QED Press/John Wiley, 1992 *