



BIBLIOTECA
FAC. DE INFORMÁTICA
U.N.L.P.

Indicadores de negocios para soporte de decisiones en Retail Business.

Trabajo de grado de la Licenciatura en Informática de
Hernán Marrafini
Mauricio Baumler

Director: Dr. Fausto Simonelli.
Co-Director: Mg. Rodolfo Bertone.



Facultad de Informática.
Universidad Nacional de La Plata

<p>TES 06/23 DIF-03105 SALA</p>	<p> UNIVERSIDAD NACIONAL DE LA PLATA FACULTAD DE INFORMÁTICA Biblioteca 50 y 120 La Plata catalogo.info.unlp.edu.ar biblioteca@info.unlp.edu.ar</p> <p> DIF-03105 13</p>
---	--

Índice

1. Introducción.....	3
Motivación.....	3
Desarrollo propuesto.....	4
Resultados.....	5
Estructura de la tesis.....	6
2. Proceso de comercialización.....	8
Definición.....	8
Proceso de gerencia comercial.....	9
Planeamiento de una estrategia comercial.....	10
Mezcla comercial - Variables controlables.....	10
Plan comercial.....	19
Entorno del planeamiento estratégico - Variables incontrolables.....	20
Obtención de información.....	21
3. Comercialización en una empresa de retail.....	24
Definición de retail.....	24
Productividad en el negocio de retail.....	26
Modelo de recursos estratégico.....	28
4. Definiciones técnicas generales.....	32
Sistemas de información.....	32
Toma de decisiones.....	36
El proceso de decisión.....	37
5. Conceptos de Datawarehousing.....	39
Definición de datawarehouse.....	39
Comparación OLTP vs. OLAP.....	41

Modelado de Datos.....	46
Esquema estrella.....	48
Esquema copo de nieve.....	50
Evolución de arquitecturas.....	51
Otras arquitecturas.....	53
6. Descripción de la aplicación.....	54
Consideraciones generales.....	54
Datos disponibles.....	55
Requerimientos del sistema.....	57
Arquitectura.....	58
7. Capa de datos.....	62
Diseño del modelo del datawarehouse.....	62
Data staging area.....	69
Diseño de la Data staging area.....	72
8. Capa de aplicación y presentación.....	78
Capa de aplicación.....	78
Desarrollo de la capa de aplicación.....	79
Capa de presentación.....	80
Desarrollo de la capa de presentación.....	81
9. Conclusiones y Trabajo Futuro.....	96
Conclusiones.....	96
Trabajo Futuro.....	97
Apéndice A: Descripción de reportes.....	98
Referencias.....	113

Capítulo 1

Introducción

En este capítulo se describen la motivación y los objetivos del trabajo de grado; se enumeran brevemente los resultados obtenidos y, por último, se detalla la estructura del informe para ayudar al lector a tener una visión global del mismo.

Motivación

La estrategia de mercado de retail define:

- Mercado destino
- Formas de satisfacer las necesidades del cliente
- Ventajas competitivas sostenibles

La estrategia financiera de retail define:

- Objetivos de venta y beneficios.
- Presupuestos de compra

Estas estrategias gobiernan, las compras que efectuará el gerente comercial de una empresa de retail para satisfacer las necesidades de sus clientes en términos de variedad de productos, abastecimiento y disponibilidad en sus almacenes o puestos de venta.

¿Qué sucede cuando un cliente va a un almacén y no encuentra el producto deseado, o el packaging encontrado no se adapta a sus necesidades particulares? En ambos casos la empresa pierde la venta y la fidelidad del cliente se ve afectada.

Es frecuente encontrar que los puntos de venta de algunas empresas poseen muchos productos que los clientes no desean, y pocos productos que los clientes desean.

En cada categoría de productos, no disponer del stock suficiente, no disponer del tamaño o la marca adecuada, y no promocionar el producto debidamente son las causas principales del descontento del cliente, su pérdida de fidelidad, y por consiguiente la baja en las ventas y en la rentabilidad de la empresa.

Para resolver sus objetivos financieros, los gerentes comerciales se esfuerzan en proporcionar el producto preciso en el lugar correcto y en el tiempo adecuado. También, los gerentes comerciales deben poder anticipar la demanda del mercado y reaccionar rápidamente a sus cambios. Dadas las restricciones de cualquier organización de retail en términos de presupuestos de compra y espacio disponible en los puestos de venta, existen ciertas decisiones clave que los responsables de compras y ventas deben tomar:

- ¿Que comprar?
- ¿Cuánto comprar?
- ¿Cómo, qué y cuándo ubicar estos productos en los estantes de los puestos de ventas?

Planear el abastecimiento y construir el conjunto de productos o servicios correctos para las necesidades del mercado y al mismo tiempo dirigir estas tareas hacia la búsqueda de los objetivos financieros de la empresa convierten este proceso en algo muy complejo.

A pesar de la complejidad y los desafíos implicados, el proceso del planeamiento del abastecimiento puede traducirse en muchas ventajas para los gerentes comerciales:

- Reducción en la variabilidad de precios.
- Reducción de las ventas perdidas
- Reducción del costo de inventario y transporte
- Incrementar la rotación del inventario y GMROI
- Espacio utilizado eficientemente.
- Pocas transferencias de productos entre puestos de venta.

Para alcanzar estos objetivos es imprescindible tener una estrategia comercial eficiente y efectiva, y hacer que las ventajas del proceso sean visibles tanto para el cliente como para la empresa.

La tecnología y la asistencia técnica brindada por sistemas informáticos surgen como el disparador para la realización de cambios fundamentales en los procesos internos de negocios, brindando información detallada de ventas, presentada como indicadores de consumo y rentabilidad, brindando soporte en la toma de decisiones que resulten en ventajas competitivas para el gerente comercial de la empresa.

Desarrollo propuesto

Se desarrollará un modelo de datos n-dimensional destinado a la implementación del modelo de recursos estratégico (SRM), para controlar la

performance de una empresa de retail. También se implementará la carga de datos a partir de un modelo de datos ejemplo, que se destina a la administración operativa de una cadena de venta de productos para la construcción y artículos para el hogar.

Se implementará el motor de análisis que permita extraer los datos del modelo n-dimensional y mostrar los indicadores de performance definidos por el SRM, aplicándolo a un negocio ejemplo consistente en una cadena de venta de productos para la construcción y productos para el hogar.

Dada la información de ventas de un determinado período, el motor de análisis deberá poder generar los siguientes indicadores:

- MN: Margen neto después de impuestos.
- ROI: Ventas sobre inventario.
- ROS: Ventas sobre superficie dedicada.
- ROL: Ventas sobre cantidad de empleados afectados.
- GMROI: Margen sobre inventario.
- GMROF: Margen sobre superficie dedicada.
- GMROL: Margen sobre cantidad de empleados afectados.

También se implementarán reportes complementarios que permiten a un gerente comercial indagar las causas de la alteración de alguno de los indicadores anteriores. Estos reportes también mostrarán la flexibilidad del modelo de warehouse acerca de la información brindada.

El motor de análisis debe permitir la explosión de cada rubro general en sus correspondientes subrubros, de los subrubros en familias, de las familias en productos individuales, calculando los indicadores definidos para cada uno de los niveles cuando el reporte lo permita.

Los reportes generados permitirán múltiples alternativas de ordenamiento y filtrado y el ordenamiento creciente o decreciente para cada uno de los indicadores exhibidos. Todos los reportes tendrán capacidad de visualización gráfica.

Resultados

Como principal resultado de este trabajo se obtuvo el conocimiento del problema de retail desde dos enfoques, uno es el enfoque de negocios y otro es el enfoque desde el punto de vista técnico. El enfoque de negocios permite conocer las características de la actividad y comprender los problemas que se deben resolver en ese contexto. El enfoque técnico permite, a partir del contexto del negocio, realizar un modelo que permita resolver la complejidad de justificar las decisiones tomadas basándose en datos reales.

Otro resultado de este trabajo es la realización de un datawarehouse que implemente el modelo de solución presentado. La construcción de este datawarehouse comprende tanto el modelo multidimensional como los procesos de carga que

permiten mantenerlo actualizado y las consultas que permiten explotar la información que contiene.

Adicionalmente se cuenta con una aplicación de ejemplo que se presenta como herramienta de soporte para la toma de decisiones y a su vez muestra la funcionalidad de la solución de datawarehouse diseñada.

Estructura de la tesis

En el capítulo 2 se define el concepto de comercialización y se describe el proceso realizado por la gerencia comercial de una empresa. También se detallan las variables controlables e incontrolables que componen la estrategia comercial y un proceso de obtención de información que permite planificar la estrategia de acuerdo al conocimiento de dichas variables.

En el capítulo 3 se presenta el concepto de empresa de retail, y se la caracteriza dentro del contexto de variables brindado por el capítulo anterior. Se establece el método para medir la productividad en este tipo de empresas y finalmente se presenta el modelo de recursos estratégicos.

En el capítulo 4 se presentan los aspectos técnicos involucrados en este trabajo. Se definen los conceptos de sistemas de información, bases de datos y sistemas de decisión.

En el capítulo 5 se define el concepto de datawarehouse y se profundiza sobre sus características. También se describen los sistemas que se basan en un datawarehouse, sus distintos tipos, historia y evolución de las arquitecturas.

En el capítulo 6 se realiza una definición de la aplicación desarrollada, presentando el contexto y requerimientos de la misma y definiendo la arquitectura elegida para llevar a cabo dicha aplicación.

En el capítulo 7 se describe la capa de datos. En esta capa se describe el modelado del datawarehouse, especificando la aproximación utilizada. También se describen los procesos de negocios elegidos para la implementación y los procesos de extracción, transformación y carga (ETL) utilizados para alimentar al datawarehouse.

En el capítulo 8 se describen las capas de aplicación y presentación, destinadas a alojar a la aplicación que realiza la interfase con el usuario y la capa que toma los datos y se encarga de los aspectos de visualización de los reportes.

En el capítulo 9 se detallan las conclusiones y los puntos que pueden dar lugar a posibles trabajos futuros.

En el apéndice A se brinda una descripción detallada de cada uno de los reportes desarrollados.

En resumen este trabajo propone una construcción de un modelo multidimensional para solucionar el problema de la toma de decisiones cuando el universo de datos es muy amplio y presenta una aplicación de ejemplo que permite comprobar la funcionalidad del modelo. ↻

Capítulo 2

El proceso de comercialización

En este capítulo se define el concepto de comercialización y se describe el proceso realizado por la gerencia comercial de una empresa. También se detallan las variables controlables e incontrolables que componen la estrategia comercial y un proceso de obtención de información que permite planificar la estrategia de acuerdo al conocimiento de dichas variables.

Definición de comercialización

Según [McCarthy-Perreault94] “La comercialización se define como la ejecución de actividades que tratan de cumplir los objetivos de una organización previendo las necesidades del cliente y estableciendo entre el productor, el intermediario y el cliente una corriente de bienes y servicios que satisfacen las necesidades”.

La comercialización debería comenzar a partir de las necesidades potenciales del cliente, no del proceso de producción o del proceso de venta. Estas necesidades deberían determinar cuales bienes y servicios se producirán y se ofrecerán al cliente.

Como aclara [Braidot96], “la venta esta orientada hacia las necesidades del vendedor, la comercialización o marketing hacia las del comprador. La venta esta preocupada por la necesidad de convertir su producto en dinero líquido; el marketing esta preocupado por la satisfacción de las necesidades del cliente a través del producto y de todo lo que este asociado a su creación, su entrega, su consumo y la satisfacción que el mismo produce, logrando como consecuencia la repetición de compra”.

Las ventas y la comercialización o marketing son términos opuestos mas que sinónimos. En primer lugar, el objetivo del marketing es conocer y entender tan bien al cliente, que el producto o servicio pueda ser definido y ajustado a sus necesidades de manera tal que se venda solo [Braidot96].

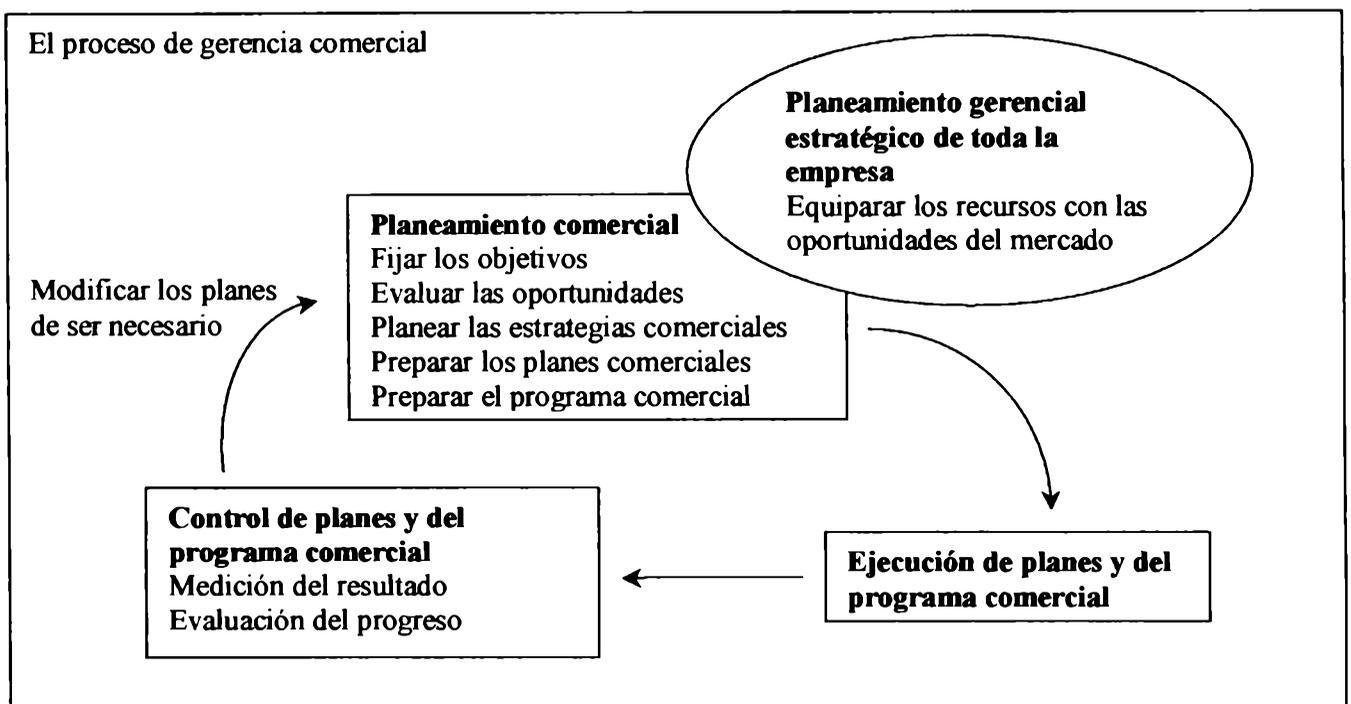
El o los gerentes comerciales de la empresa son los encargados de planificar y controlar el proceso de comercialización de los productos ofrecidos por dicha empresa.

Proceso de gerencia comercial

El proceso de gerencia comercial abarca los siguientes pasos [McCarthy-Perreault94]:

- Planear las actividades comerciales
- Dirigir la ejecución de los planes
- Controlar los planes

En la figura siguiente se ven las relaciones entre estas tres tareas dentro del proceso de gerencia comercial.



Estas tareas conforman un ciclo continuo. En el planeamiento, los gerentes fijan pautas para la tarea de ejecución y especifican los resultados esperados. Luego utilizan estos resultados esperados en la tarea de control, con el propósito de averiguar si todo funcionó de acuerdo a lo previsto. El vínculo de la función de control con la función de planeamiento es especialmente importante ya que la realimentación que de él se logra, a menudo conduce a cambios en los planes, o a planes nuevos.

Por otro lado, dado que los mercados son dinámicos en el sentido de que las necesidades de los clientes, los competidores y el medio ambiente cambian; los gerentes comerciales deben estar atentos a las nuevas oportunidades.

El planeamiento de la estrategia comercial forma parte de un planeamiento más amplio que es el planeamiento estratégico de la empresa, el cual, intenta desarrollar y mantener el equilibrio entre los recursos de la organización y sus oportunidades comerciales.

Planeamiento de una estrategia comercial

Una estrategia comercial especifica un mercado meta y una mezcla comercial afin. Las dos partes que lo componen se definen de la siguiente manera [McCarthy-Perreault94]:

- **Mercado meta:** Es un grupo bastante homogéneo de clientes a los cuales la firma desea atraer.
- **Mezcla comercial:** La constituyen las variables controlables que la empresa define para satisfacer a este mercado meta.

La mezcla comercial típica esta compuesta por un **producto**, ofrecido a un **precio** mediante cierta **promoción** que sirve para informar a los potenciales clientes acerca de las virtudes del artículo, y por la forma de **distribución** hasta llegar al cliente.

La estrategia comercial se dirige a algunos clientes determinados. Este método se llama “comercialización por metas” para distinguirlo de la comercialización masiva. La comercialización por metas afirma que una estrategia comercial se adapta para que satisfaga las necesidades de algún cliente determinado. Por el contrario la comercialización masiva apunta vagamente a todo el mundo con la misma mezcla comercial. Este sistema supone que todo el mundo es igual y considera que todo el mundo es cliente en potencia.

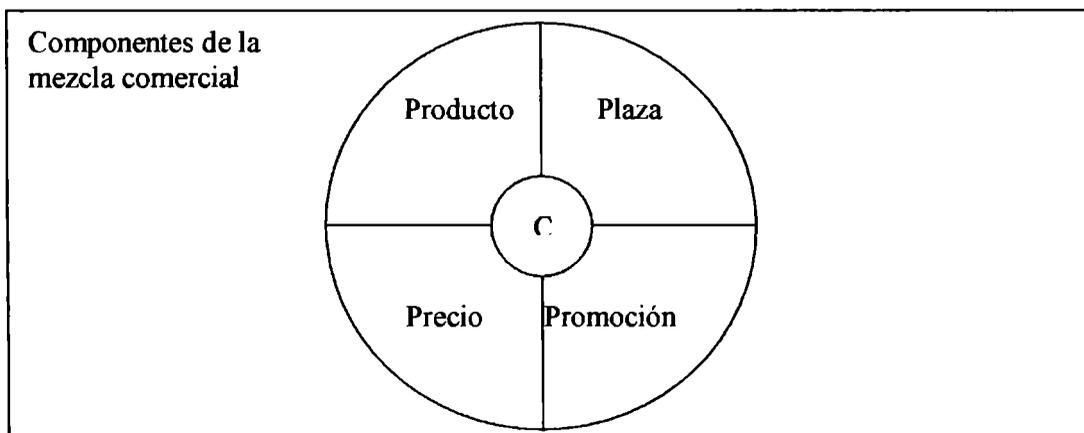
Hay que distinguir entre comercialización masiva y comercializadores masivos. Los comercializadores masivos apuntan a mercados metas claramente definidos. La confusión con comercialización masiva de los mismos suele darse debido a que sus mercados metas suelen ser grandes y dispersos.

Mezcla comercial - Variables controlables

Existen muchas formas posibles de satisfacer las necesidades de los clientes metas. Un producto puede poseer muchos atributos, características y niveles de calidad diferentes. Los niveles de servicios se pueden regular. El envase puede ser de dimensiones, materiales y colores distintos. El nombre de la marca y la garantía son susceptibles de modificarse. Es posible utilizar varios medios de publicidad, como por

ejemplo, diarios revistas, televisión, carteleras. Se pueden marcar precios diferentes o hacer descuentos.

Es útil reducir todas las variables de la mezcla comercial a cuatro grupos básicos [McCarthy-Perreault94]:



Toda mezcla comercial necesita de las cuatro variables, ninguna es más importante que la otra.

Las tareas de elección del mercado meta y la elaboración de una mezcla comercial son tareas interrelacionadas.

Producto

El campo del Producto se ocupa de la creación y definición del concepto que llamaremos “producto”, que será adecuado para el mercado meta. Esta oferta puede implicar un bien físico y simbólico, un servicio o una combinación de ambos. Téngase en cuenta que el producto no se circunscribe a bienes físicos, porque tiene un fuerte significado simbólico que puede impactar en el consumidor e inducirlo a adquirirlo. Un ejemplo de ello es lo que significa para éste, **la marca** de un producto determinado. Para él, **el significado** de un producto de marca equis, puede ser totalmente distinto del que expresa otro similar que compite con él, pero con una marca diferente.

Según [Braidot96], “el producto es (...) el conjunto de satisfactores o beneficios que permitirán al consumidor resolver sus necesidades. Esta integrado por elementos intrínsecos e intangibles. Forman parte del mismo determinadas materias primas como percepciones que el consumidor tiene, sean estas generadas por una publicidad hecha con ese propósito, como por los efectos del precio fijado o el canal utilizado para su distribución”.

Junto con otras decisiones sobre el “producto”, se tiene en cuenta la creación y gestión de nuevos productos y líneas completas de artículos.

Algunos elementos que componen el producto son:

- Bien físico
- Bien simbólico
- Características
- Líneas de productos
- Instalación
- Envase
- Marca

Plaza

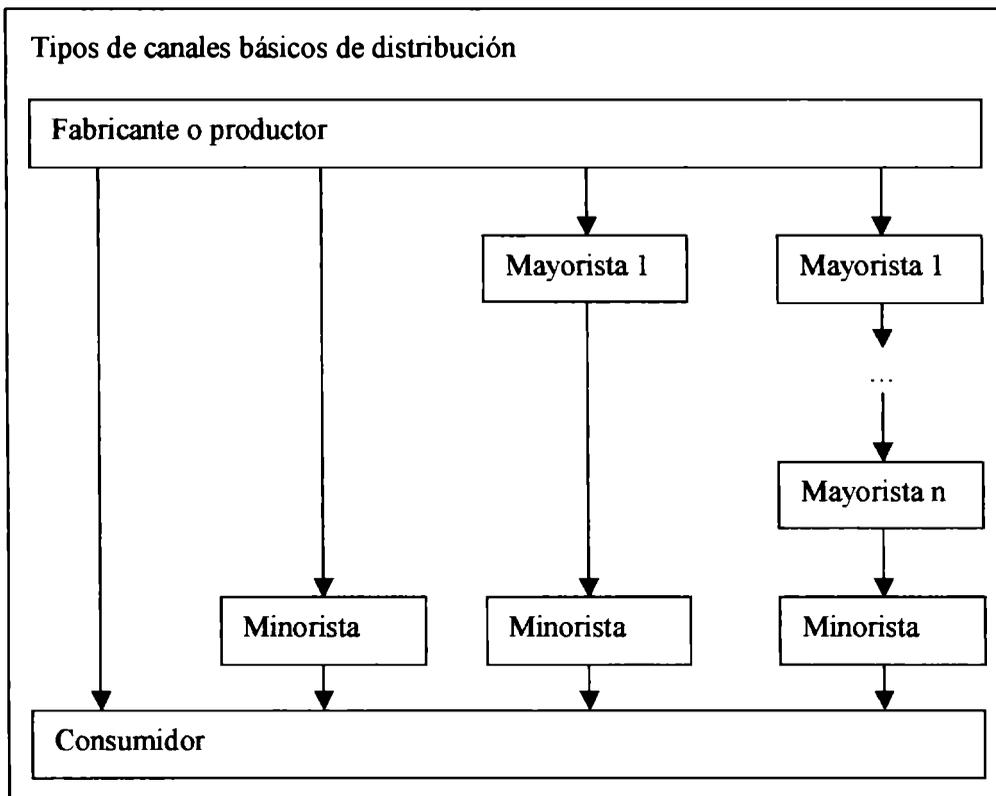
La Plaza consiste en la obtención del producto adecuado respecto del momento y lugar donde el cliente desearía obtenerlo. En otras palabras, consiste en poner a disposición del consumidor el producto adquirido, en el momento, forma y lugar justos.

El producto llega a los clientes por medio de un canal de distribución. Un canal de distribución está formado por empresas o personas que participan del flujo de bienes y servicios desde el productor al consumidor final.

El canal de distribución difiere de acuerdo al mercado meta. Puede ser un canal que va directamente desde el productor al consumidor o puede incluir clases distintas de intermediarios y especialistas. Cuando un gerente comercial tiene varios mercados metas diferentes puede que necesite diferentes canales de distribución.

Cuando el “producto” comprende un bien físico, el planeamiento de la Plaza incluye decisiones sobre distribución física. La distribución física es el transporte y almacenamiento de los bienes, tanto dentro de las empresas mismas como en el canal de distribución, para satisfacer las necesidades de los clientes metas.

[Braidot96] define este proceso como logística de comercialización, el cual consiste en todo el movimiento físico de productos terminados entre las plantas y depósitos, y en general todo el desplazamiento y almacenamiento de productos que se requiere para que lleguen a manos del consumidor, en tiempo, forma y lugar adecuados.



Algunos elementos que componen la Plaza son:

- Objetivos
- Tipo de canal
- Exposición al mercado
- Clases de intermediarios
- Clases y ubicación de los puntos de venta
- Transporte y almacenamiento

El almacenamiento es la función comercial que consiste en la acumulación de mercaderías. Proporciona utilidad de tiempo al tener disponible el producto próximo al sitio de venta. La existencia es la cantidad de mercadería almacenada.

Es preciso almacenar cuando la producción de los bienes no concuerda con el tiempo de consumo de los mismos. El almacenamiento permite a la empresa tener disponible el producto para satisfacer las necesidades del cliente oportunamente.

El almacenamiento genera costos adicionales, ya que se necesita espacio adecuado, seguros sobre la mercadería almacenada y personal disponible para controlar la entrada y salida de los mismos en buen estado. La mayoría de las empresas utilizan depósitos públicos o privados para almacenar los productos.

El costo de manipulación de los productos es el principal gasto de almacenamiento. Se debe tener un proceso acorde que permita el control de la entrada de productos al depósito y de su salida del mismo para ponerlos a la venta.

La opción más utilizada por las empresas es el centro de distribución. Consiste en una clase especial de depósito diseñado para acelerar el flujo de mercancías y evitar gastos de almacenamiento innecesarios.

Ningún intermediario desea quedarse sin los productos necesarios al momento de necesitarlos, pero mantener abundantes existencias es caro. Las empresas prestan atención al costo de las existencias y, junto con la colaboración de los proveedores, deben sincronizarse para recibir las mercaderías en el momento adecuado.

La mayoría de los minoristas trabajan con una gran cantidad de artículos. Para una cadena de supermercados este número puede ser de 20.000 artículos o más. En esta situación es muy difícil o imposible atender a cada artículo en particular originando problemas de almacenamiento o de nula existencia de ciertos productos. Es por ello que con la ayuda de la tecnología se está evolucionando hacia sistemas más perfeccionados de distribución de productos y servicios.

Dado que las empresas compran y acumulan una cantidad tan grande de artículos, deben vigilar las existencias con mucho cuidado. La información de salida de los artículos está dada por las ventas que realiza la empresa. Esta información acumulada le permitiría al gerente comercial conocer bien sus necesidades y plazos al momento de decidir cuáles mercaderías comprar.

Promoción

La Promoción se refiere a la actividad de informar al mercado meta acerca del Producto. Comprende la venta personal, venta masiva y promoción de venta.

La venta personal implica comunicación hablada entre vendedores y clientes potenciales. La venta personal usualmente es cara a cara y también puede realizarse telefónicamente. Este tipo de venta permite al vendedor adaptar la mezcla comercial a cada cliente, pero es cara desde el punto de vista de tiempo y esfuerzo.

La venta masiva consiste en comunicarse con una gran cantidad de clientes al mismo tiempo. La forma principal de venta masiva es la publicidad, toda forma pagada o no de presentación impersonal de bienes o servicios por algún auspiciante determinado.

La promoción de ventas se refiere a todas aquellas estrategias, excepto publicidad, difusión y venta personal, que estimulan el interés o inducen a los clientes finales o a otros, dentro del canal de distribución a probar el producto o comprarlo. Esto suele implicar el uso de cupones, materiales para los lugares de venta, muestras, carteles, catálogos, concursos y circulares.

- Algunos elementos que componen la Promoción son:
- Objetivos
- Vendedores
- Publicidad
- Promoción de ventas
- Difusión

Precio

En una economía de mercado, todo producto o servicio tiene un precio. No obstante, las organizaciones no siempre saben con precisión cual es el valor que el consumidor realmente esta dispuesto a pagar por su producto.

La elección de una estrategia de precios requiere atender a dos aspectos fundamentales; coherencia interna y externa. La coherencia interna contempla los límites fijados por los costos del producto y exigencias de rentabilidad de la empresa. La coherencia externa implica la detección de las expectativas del mercado en relación con el valor del producto, la capacidad de compra del mercado meta y lo que los compradores están dispuestos a pagar por el bien o servicio [Braidot96].

[McCarthy-Perreault94] afirma que además de preparar las tres variables anteriores de manera adecuada, el gerente comercial debe establecer la estrategia de precios correcta. Al fijarlo deben tener en cuenta la clase de competencia que se da en el mercado, el posicionamiento que se le definirá al producto y el costo de toda la mezcla comercial. También les corresponde evaluar aproximadamente la reacción del cliente ante los precios posibles. Además es imprescindible que conozcan las costumbres actuales en cuanto a márgenes, descuentos y otras condiciones de venta.

Algunos elementos que componen el Precio son:

- Objetivos
- Flexibilidad
- Condiciones geográficas
- Posicionamiento del producto
- Descuentos
- Tolerancias

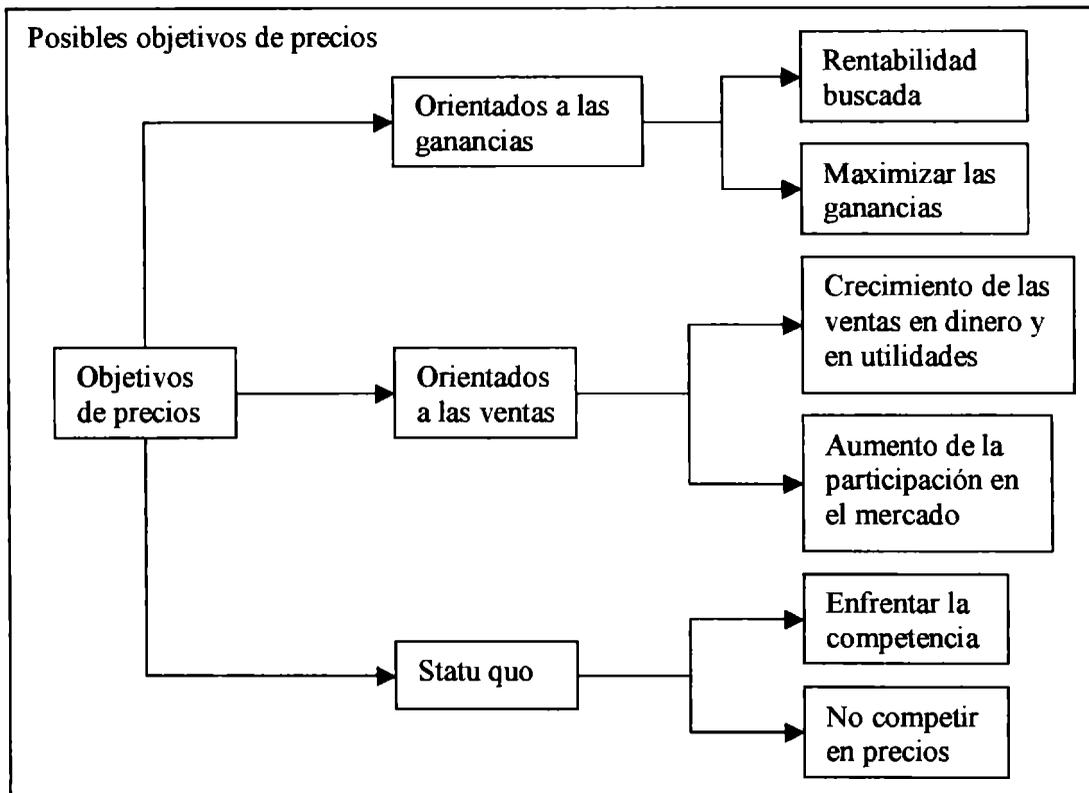
Orientados por los objetivos de la compañía, los gerentes comerciales trazan un conjunto de objetivos y políticas de precios. Así, explican cuales situaciones de precios enfrentará la firma y como las manejará. Estas políticas definen:

- flexibilidad de los precios

- en que nivel se fijarán durante el ciclo de vida del producto
- quienes y cuando serán beneficiarios de descuentos y bonificaciones
- como se liquidarán los costos de transporte

Los objetivos de precios deberían adecuarse a la importancia de la empresa y de los objetivos comerciales. Los objetivos de precios deberían ser expuestos explícitamente, porque influyen directamente en la política de precios y en el método para fijarlos.

Existen distintas clases de objetivos de precios:



Un objetivo de **rentabilidad buscada** establece un nivel concreto de ganancias como objetivo. A menudo este monto se fija como un porcentaje de las ventas o del capital invertido. Este tipo de objetivo tiene ventajas administrativas en una compañía grande, porque los gerentes pueden comparar el rendimiento con la meta buscada.

El objetivo de **maximización de ganancias** busca obtener tanta utilidad como sea posible. Podría expresarse como el deseo de ganar una rápida rentabilidad de la inversión. Fijar precios para lograr el máximo de utilidades no siempre se traduce en precios altos. Los precios bajos pueden ampliar la dimensión del mercado y dar por resultado mayores ventas y ganancias.

Un objetivo orientado al **crecimiento de las ventas** busca algún nivel de ventas en unidades, en dinero o de participación en el mercado, sin hacer referencia a la ganancia.

Se debe tener cuidado con este tipo de objetivos ya que no siempre es cierto que un incremento en las ventas produce un incremento en las ganancias. Cuando se fija este tipo de objetivos hay que dedicar atención adicional a los costos acarreados por el aumento de ventas ya que podrían atentar contra la ganancia de la empresa.

El objetivo de **lograr participación en el mercado** es muy utilizado. Este obliga a los gerentes a prestar atención al comportamiento de la competencia. También suele ser más fácil evaluar la participación del mercado de una empresa, que determinar si las ganancias están siendo llevadas al máximo.

Algunas empresas se esfuerzan por lograr una alta participación en el mercado e incluso controlarlo ya que si lo logran, obtendrán mejores economías de escala que sus competidores.

Los gerentes que están satisfechos con el nivel de ganancias y participación en el mercado a veces adoptan objetivos de **statu quo**. También es normal cuando el mercado no crece. Mantener los precios estables desalentará a la competencia en ese rubro y evitará la toma de decisiones difíciles.

Las políticas de precios suelen conducir a los precios administrados, es decir, a precios fijados conscientemente. En lugar de dejar que las fuerzas diarias del mercado fijen sus precios, la propia empresa es la que los establece.

Dentro de la política de precios debe considerarse la flexibilidad del mismo. Una política de precio único significa ofrecer al mismo precio a todos los clientes que compren productos en las mismas condiciones y cantidades. Por otro lado la política de precios flexibles significa ofrecer el mismo producto y cantidades iguales a distintos clientes pero a precios diferentes. Esta última política de precio tiene la ventaja que permite al vendedor hacer ajustes teniendo en cuenta el precio cobrado por los competidores, la relación con el cliente, y la capacidad negociadora del mismo.

Otro componente de la política de precios es el descuento. Los descuentos son reducciones sobre el precio de lista básico fijado inicialmente para el producto, otorgados por los vendedores a clientes que renuncian a cierta función comercial o la proveen ellos mismos.

Existen varios tipos de descuentos. Los descuentos por volumen son los ofrecidos para estimular a los clientes a comprar en grandes cantidades. Esto permite al vendedor tomar una parte mayor de los negocios del comprador, o trasladar una parte de la función de almacenamiento al comprador, o reducir los costos de envío y venta, o las tres cosas.

Dentro de los descuentos por volumen se identifican los acumulativos y los no acumulativos. Los descuentos acumulativos por cantidad se aplican a compras hechas

durante un período dado y por lo general, el monto del descuento aumenta a medida que lo hace el volumen comprado. El descuento acumulativo tiende a estimular la repetición de la compra por parte del cliente al reducir el costo de sus compras adicionales.

Los descuentos por cantidad no acumulativos se aplican solo a pedidos individuales. Estos descuentos alientan a presentar pedidos más grandes, pero no atan al comprador con el vendedor luego de la compra.

Los descuentos estacionales son los que se ofrecen para estimular a los compradores a acumular existencias antes que lo requiera la demanda actual. Cuando los vendedores lo utilizan este descuento tiende a trasladar la función de almacenamiento a lo largo del canal. También contribuye a distribuir uniformemente las ventas durante el año y por lo tanto permite operar todo el año.

Un precio de liquidación es un descuento temporario sobre un precio de lista básico. Los descuentos por liquidación estimulan la compra inmediata. Este descuento implica que el cliente que quiera aprovecharlo deberá resignar la conveniencia de comprar cuando ellos deseen hacerlo, comprando cuando el vendedor desea. Este tipo de ventas “especiales” proporcionan al gerente comercial de una herramienta de rápida aplicación para responder a los cambios del mercado, sin modificar la estrategia comercial básica. Por ejemplo, se podría utilizar una liquidación para desembarazarse de existencias excesivas o para afrontar la competencia de precios.

Dentro de la política de precios debe considerarse también a las bonificaciones. Al igual que los descuentos, las bonificaciones se ofrecen al cliente a cambio de que este asuma cierta función que normalmente corresponde al vendedor o que resigne algún servicio por parte del vendedor.

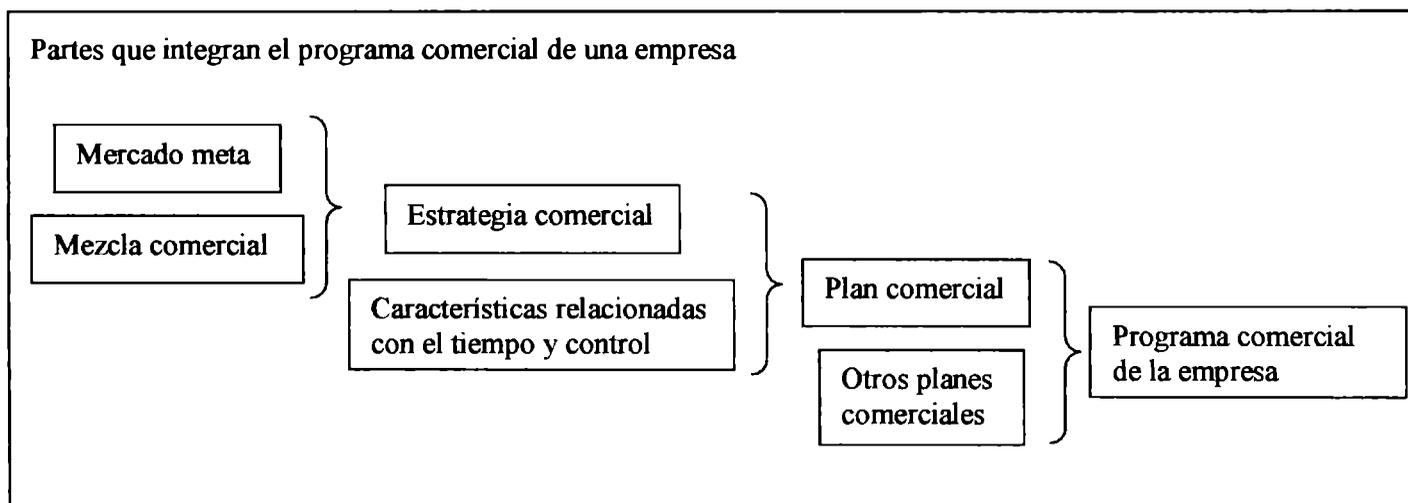
Las bonificaciones por existencias se conceden a un intermediario para conseguir espacio en sus estanterías para un producto. Esto es utilizado en el negocio de retail para estimular la inserción de un producto nuevo dentro de sus puntos de venta. El proveedor generalmente compensa con dinero o productos gratis los costos del comprador que surgen de generar espacio en sus depósitos, manipular y reorganizar sus estanterías para insertar el nuevo producto.

Otro componente del precio que influye es la zona geográfica. Esto implica aplicar un costo de flete promedio al precio del producto para todos los compradores que residan en zonas geográficas específicas. El vendedor paga el costo del flete y factura a cada cliente un monto promedio. Esto también puede ser utilizado como una variable dentro de la política de precios, aplicando el recargo al cliente o ~~no~~ dependiendo de la estrategia adoptada por la empresa. En una estrategia ~~de~~ penetración de la empresa, la misma podría decidir absorber este costo para obtener un precio más competitivo.

Plan comercial

La estrategia comercial determina un mercado meta y una mezcla comercial. El plan comercial es la exposición de la estrategia comercial y el tiempo necesario para llevar a cabo dicha estrategia. Entre otras cosas, el plan comercial deberá contar con una descripción del mercado meta y la mezcla comercial elegida, los resultados esperados (ventas o ganancias mensuales o trimestrales) y procedimientos de control que permitan controlar cualquier desvío en el camino hacia el logro de los objetivos planteados. Estos procedimientos de control pueden ser dados por la comparación de las ventas reales contra las estimadas por ejemplo.

Una vez generado el plan comercial se comenzará con su ejecución y control a lo largo del tiempo. Comúnmente existen varios planes comerciales en la misma empresa destinados a ofrecer una estrategia para todos los productos que una empresa ofrece. De esta manera pueden existir varios gerentes comerciales dedicados a la comercialización de distintos productos o líneas de productos.



El control que proporciona la información permite a los gerentes comerciales modificar sus estrategias. El gerente puede recurrir a numerosos instrumentos como el análisis de ventas por computadora o el análisis contable de costos y ganancias.

Sin la presencia de las tareas de control es imposible determinar si la empresa se conduce por un camino adecuado o debería cambiar su estrategia.

Entorno del Planeamiento Estratégico – Variables

Incontrolables

El planeamiento estratégico de la empresa se desarrolla dentro de un contexto que presenta variables que, a diferencia de las cuatro variables de la mezcla comercial, el gerente no puede controlar. El gerente debe preocuparse por el ambiente competitivo, económico y tecnológico, político y jurídico, cultural y social [Sallenave93]

El ambiente competitivo influye en la cantidad y clase de competidores que el gerente comercial debe enfrentar y en su comportamiento. Si bien los gerentes comerciales no pueden controlar estos factores, podrían estar en condiciones de escoger estrategias que evitaran la competencia frontal o, cuando esta fuera inevitable, podrían planear para ella.

El ambiente económico influye en como las empresas utilizan sus recursos. El ambiente económico sufre las consecuencias de cómo todas las partes de nuestro sistema macroeconómico accionan las unas sobre las otras. Esto a su vez influye sobre el ingreso nacional, el crecimiento económico y la inflación entre otras cosas.

Los cambios en el ambiente económico pueden producirse con gran rapidez y exigen modificar la estrategia comercial. El gerente comercial debe observar el ambiente económico con todo cuidado. En oposición al ambiente cultural y social, las condiciones económicas mudan continuamente y se pueden trasladar con toda rapidez, exigiendo respuestas estratégicas.

El ambiente tecnológico es subyacente al ambiente económico y se compone de los conocimientos técnicos y los equipos que condicionan como los recursos de una economía se transforman en producción. El ambiente tecnológico tiene una gran influencia sobre la actividad de las empresas y sus estrategias de comercialización. Las formas de difundir un producto han evolucionado considerablemente y las computadoras permiten obtener una gran cantidad de información que permite planificar de una manera más eficiente, entre otros aspectos.

El ambiente político comprende las actitudes y reacciones de los consumidores, críticos sociales y gobiernos. Parte de estos elementos son por ejemplo asociaciones de defensa del consumidor o la política que adopte el gobierno actual sobre la comercialización, como por ejemplo el enfoque nacionalista sobre los productos o los trámites necesarios para realizar una actividad comercial.

El ambiente jurídico comprende las normas sobre las cuales se desarrolla la sociedad y por ende cualquier empresa que realiza alguna actividad comercial. Los gerentes comerciales están sujetos, como cualquier otro ciudadano, a las leyes penales y civiles vigentes.

El ambiente cultural y social influye sobre las causas y razones de las personas que determinan como viven y se comportan, lo cual repercute sobre los hábitos de compra del cliente y por consecuencia en los ambientes económicos, políticos y jurídicos. Los cambios en la sociedad y en sus valores se producen con lentitud. La empresa debe descubrir las características del momento y trabajar sobre esas restricciones.

Obtención de información

El planeamiento exitoso de cualquier estrategia comercial exige información relacionada con los mercados meta, y sus posibles respuestas ante las mezclas comerciales, así como sobre la competencia y otras variables incontrolables. También se necesita información para la ejecución y el control de las estrategias. Sin buena información comercial, los gerentes tienen que recurrir a la improvisación o la conjetura, lo cual invita al fracaso en un entorno dinámico.

Los gerentes comerciales de algunas empresas toman decisiones sobre la base casi exclusiva de su criterio, con muy escasos datos confiables. Cuando llega el momento de tomar una decisión es deseable contar con más información, pero dado que la obtención de la misma no es instantánea, no queda más remedio que actuar sin ella [Sallenave93].

Cuando un gerente obtiene información que le facilita la tarea de planificación y decisión, además de utilizarla, tiende a realizar diversos cálculos sobre ella con el objetivo de obtener más información o una mejor calidad de la misma. También les permite vigilar la ejecución de la estrategia, comparar los resultados con lo planificado e introducir modificaciones de ser necesario.

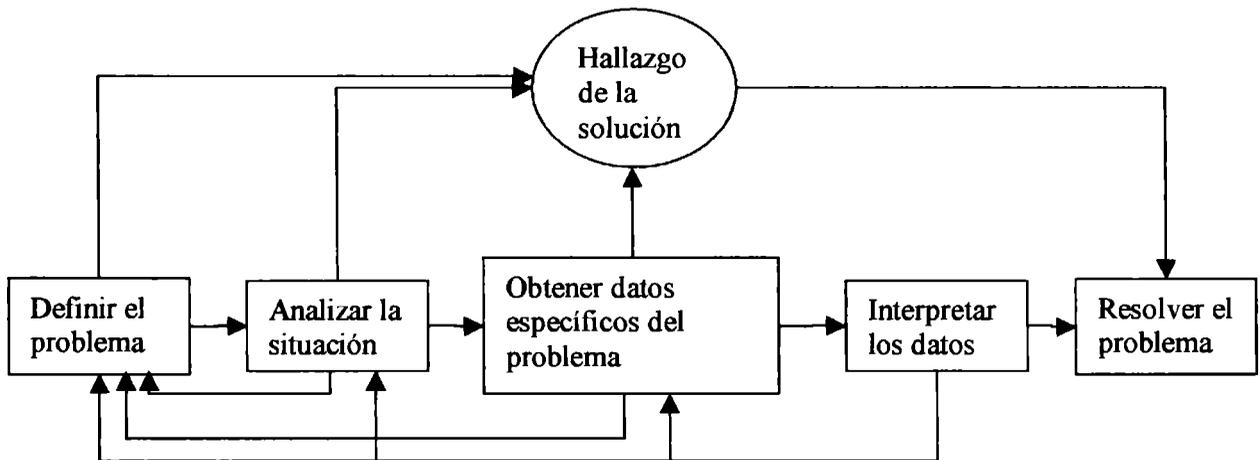
El costo de obtención de esta información no es bajo. Las fuentes de información son muy diversas y la transformación a datos de relevancia para el gerente es un proceso complejo y que consume mucho tiempo.

La investigación de mercado es uno de los procesos que le permiten a una empresa generar información que les permita tomar decisiones justificadas.

La investigación de mercado consta de 5 pasos:

- Definir el problema
- Analizar la situación
- Obtener los datos específicos del problema
- Interpretar los datos
- Resolver el problema

Proceso de investigación de mercado



- **Definir el problema:** Es el paso más importante del proceso y generalmente el que más tiempo consume. Permite obtener una definición clara de los objetivos de la investigación. Se debe tener cuidado en no confundir el problema con los síntomas ya que los objetivos de la investigación se verían desvirtuados.
- **Analizar la situación:** Consiste en un estudio informal de la información que se dispone al momento acerca del problema. Permite caracterizar al problema y determinar que información adicional se necesita.

Un análisis de situación presenta mucha información y es más económico que otro tipo de investigación. Al finalizar el análisis de situación se podrá advertir cuales son los interrogantes de la investigación que permanecen sin respuesta y que medios son necesarios para obtener una respuesta a dichos puntos.

- **Obtener los datos específicos del problema:** Tiene por objeto saber que opinan los clientes sobre algún tema o evaluar sus comportamientos ante determinados eventos. Existen diversos métodos para obtener la información que permita resolver un problema. Entre ellos están la investigación cualitativa y cuantitativa, y la recolección de datos relacionados con los periodos anteriores de actividad de la empresa. La elección de alguno de estos métodos depende del tiempo y el dinero que se dispone para realizar la investigación.
- **Interpretar los datos:** Una vez obtenidos los datos se debe determinar el significado de los mismos. Esto puede involucrar manejos estadísticos de los mismos en el caso de ser datos cuantitativos. Se debe analizar si la muestra tomada es representativa de la realidad y la validez de los mismos. La interpretación de estos datos nunca es exacta sino que se intenta aproximar una situación de la realidad tomando una muestra representativa de la misma. Es responsabilidad del gerente verificar la validez de las conclusiones que surjan a partir del análisis de los datos recolectados.

- Resolver el problema: En este paso se utilizan los resultados de la investigación realizada para tomar decisiones comerciales. Concluido el proceso de investigación el gerente comercial debería poder aplicar los descubrimientos al planeamiento de la estrategia comercial.

La información confiable suele ser costosa y las empresas que pueden afrontar los costos de obtenerla descubren que la información comercial se paga a sí misma. Les permite elaborar una mezcla comercial o elegir un mercado meta correctos o prevenir un problema potencial antes que se convierta en una crisis costosa.

Las herramientas o instrumentos que las empresas utilizan para una mejor obtención de información y gestión de sus negocios deben, como primera prioridad, apuntar a la detección de necesidades del consumidor y tener en cuenta las posibles o muy probables variaciones o evoluciones que en ellas puedan verificarse. En la medida que estos objetivos sean puestos como premisa de acción, la organización estará en condiciones de ofrecer los productos o servicios que realmente se están necesitando o podrá modificar la composición de la línea existente acorde con las variaciones que el mismo mercado aconseja [Braidot96].

Los gerentes comerciales se ven obligados a tomar riesgos en virtud de información incompleta. Eso es parte de su trabajo, pero es preciso que comparen el costo de conseguir más datos, con su probable valor. Si el riesgo no es demasiado grande el valor de obtener mas información puede ser mayor que la perdida potencial causada por una decisión desacertada [Sallenave93].

Según [Braidot96], “dirigir bien una empresa significa saber ‘administrar el futuro’, y para administrar el futuro es necesario contar con la mayor cantidad y mejor calidad de información posible sobre lo que vendrá”.

Una adecuada información de mercado ayuda a los gerentes en el manejo de todas las áreas de manera eficiente, y sobre todo con objetivos bien definidos y fundamentados en la realidad que los rodea.

La utilización y aplicación de la creatividad e inteligencia, apoyadas en recursos técnicos generalmente disponibles, comió los sistemas computarizados y la utilización de servicios especializados de terceros cuando es realmente necesaria una investigación concreta permiten resolver gran parte de las limitaciones propias que suelen encontrarse en las empresas.

El factor de éxito fundamental para una organización es contar con una conducción adecuada y propensa a la utilización de estrategias decisionales fundamentadas técnicamente y compatibilizadas con la realidad de los productos y de los mercados.

No se pretende afirmar que no deban considerarse o tomarse en cuenta las cualidades intuitivas tan comunes en los empresarios emprendedores, Simplemente se debe erradicar la improvisación, manteniendo en alto la intuición como factor de conducción estratégica apoyada en toda la infraestructura técnica disponible [Braidot96].

Capítulo 3

Comercialización en una empresa de Retail

En este capítulo se define el concepto de empresa de retail, y se la caracteriza dentro del contexto de variables brindado por el capítulo anterior. Se establece el método para medir la productividad en este tipo de empresas y finalmente se presenta el modelo de recursos estratégicos.

Definición de Retail

El **retail** se define como el comercio o venta de productos, generalmente en pequeñas cantidades, a consumidores finales que tienen un propósito de consumo y no de reventa de los mismos. Ejemplos de este tipo de empresas son las cadenas de home-centers o supermercados [wikipedia].

Estas empresas generalmente constan de varios puntos de venta que se administran conjuntamente por varios gerentes comerciales, los cuales tienen asignados distintas líneas de producto para todos los puntos de venta.

La magnitud de **productos** ofrecidos por estas empresas es muy grande, promediando más de 60.000 productos distintos.

La **plaza** de estas empresas esta formada por puntos de venta que consisten en grandes instalaciones con depósitos propios y centros de distribución que los abastecen.

Las compras se realizan de manera centralizada para las distintas sucursales de venta que componen la empresa. También se distribuyen los costos de **promoción** y gestión entre las diversas sucursales. Estas estrategias permiten disminuir el costo de la comercialización, pudiendo competir en un mercado de márgenes bajos.

Según [Braidot96], “el fenómeno de la concentración transforma las relaciones entre proveedores y grandes minoristas, provoca la caída de firmas tradicionales y alumbra el surgimiento de nuevos actores en la cadena de distribución y comercialización”.

Debido a la magnitud de estas empresas, no se puede contar con una diversidad de estos puntos de venta por factores económicos y sociales. Los factores económicos que afectan esta decisión son debidos al volumen de mercadería que se ofrece y a la magnitud de la infraestructura invertida para montar un punto de venta. Los factores político-sociales impiden que este tipo de negocios instalen puntos de venta fácilmente dado que la magnitud de las instalaciones genera un gran consumo de espacio y energía, afectando al entorno en el cual se instalan.

Las empresas de retail apelan a la distribución masiva, la cual tiene como objetivo vender a un **precio** mas bajo para lograr una rotación más rápida y mayor volumen de ventas, apelando a mercados más grandes.

Según [Braidto96], estas empresas tienen una estrategia de precios bajos y activos, en donde venden productos indiferenciados en los cuales el precio es el factor más importante en la decisión de compra. Se trata de productos de consumo masivo y de gran rotación. Las cadenas de hipermercados establecen su competencia basados principalmente en esta variable.

Estas empresas generalmente venden productos de conveniencia, pero en gran cantidad. Para que la compra de mucho volumen resulte más fácil, trabajan con gran variedad de rubros de producción. Las instalaciones son muy grandes y están planeadas para que el proceso de venta sea lo más rápido posible. La transacción se realiza utilizando lectores ópticos que registran la operación de salida de productos, y se asigna el espacio de acuerdo a los artículos que mayor movimiento y mayor utilidad poseen. La supervivencia de este tipo de comercios depende de esta eficiencia. La competencia en esta área es intensa y las ganancias netas rondan entre el 1% y el 3% de las ventas dependiendo de la línea de productos observada.

Sumado al bajo margen existe una competencia muy alta ya que los productos ofrecidos son iguales o al menos equivalentes, el acceso por parte de los clientes a los puntos de ventas no difiere entre las distintas empresas y sus costos de abastecimiento y manipulación de los productos no varían demasiado ya que la mayoría de las empresas poseen formatos similares de abastecimiento.

La estrategia comercial de mercado de retail define:

- Mercado destino
- Formas de satisfacer las necesidades del cliente
- Ventajas competitivas sostenibles

La estrategia financiera de retail define:

- Objetivos de venta y beneficios.
- Presupuestos de compra

Estas estrategias gobiernan las compras que efectuará la empresa para satisfacer las necesidades de sus clientes en términos de variedad de productos, abastecimiento y disponibilidad en sus depósitos o puestos de venta.

Dada la magnitud, la competencia y los escasos márgenes en este rubro, todas las decisiones que tome el gerente en su estrategia comercial serán cruciales para la obtención de ganancias. También se debe tener un buen mecanismo de control de la performance de la empresa para poder modificar a tiempo cualquier decisión que genere un bajo rendimiento del negocio.

Productividad en el negocio de Retail

Según [Hernant04] “La definición operacional de la productividad en retailing es muy debatida en el mercado. La mayoría de los investigadores, particularmente los que tienen una orientación económica, definen productividad como una relación entrada-salida expresada en una variedad de formas”. Incluso, se plantea la corrección acerca de las distintas opciones de medir la entrada y la salida.

Por otro lado, existe una definición alternativa en la literatura de performance organizacional. Este enfoque presenta la productividad como un concepto más amplio de cuan eficientemente la salida de una empresa es generada y cuando se alcanzan objetivos específicos.

De hecho, la discusión sobre productividad debería ser considerada en términos de la optimización de los recursos que son utilizados y la eficacia con la que se obtiene la salida. La optimización trata con la utilización de los recursos a través de distintos objetivos, mientras que la eficacia maximiza el retorno de la inversión a largo plazo como objetivo final en la medición de la productividad.

Las medidas de productividad que utilizan medidas financieras son quizá las más antiguas y ampliamente utilizadas para el control financiero. El objetivo es evaluar el uso eficiente de un recurso para crear una salida que la empresa considere de valor.

Un concepto clave en la medida de la productividad es la respuesta a la pregunta ¿Qué es salida? Para la empresa de retail esta respuesta no es tan sencilla de formular, así como lo sería en una empresa dedicada a la producción de algún producto, donde la respuesta es el volumen de unidades producidas (no necesariamente vendidas). Para un gerente comercial de una empresa de retail la respuesta no puede ser conceptualizada de manera tan directa. El gerente comercial agrega valor mediante “servicios relacionados con el producto”, por ejemplo: tiempo, lugar, disponibilidad de mercadería. Estos factores son mucho más difíciles de medir. La variable Producto que compone la estrategia comercial de una empresa de retail es mucho más que el bien físico [Hernant04].

En los establecimientos de retail, que compran mercadería a los proveedores para revenderlas al mercado, esta mercadería es transformada en productos totalmente distintos, aunque su apariencia física no sea modificada. Cuando el producto es

ofrecido por la empresa de retail, el producto original esta embebido dentro de un conjunto de atributos que el gerente comercial ha creado, (lo que llamamos el "concepto" del "producto"), como parte de la estrategia de venta de su compañía.

Cada producto tiene evidentemente una función básica, a la que se añaden una serie de servicios secundarios de naturaleza estética, social y/o cultural. Estas utilidades adicionales pueden ser objetivas o sencillamente preceptuales, es decir, resultado de una imagen de marca o del posicionamiento publicitario logrado [Braidot96].

En el nivel simbólico la relación producto consumidor se hace más compleja, dado que el producto se interpreta como un reflejo del consumidor mismo y de los significantes específicos que este le asigna.

La importancia de detectar estos atributos o servicios derivados radica en que se puede optar por posicionar el producto en relación con alguna de estas características secundarias, diferenciando así la oferta de otros competidores, a pesar de que no tengan relación directa en lo que se refiere a la utilidad básica o satisfactor principal [Braidot96].

La implicancia de esto es que, a excepción de un intermediario, la producción dada por la empresa de retail es un proceso que transforma productos manufacturados junto con varios tipos de trabajo y capital en ofertas más complejas, las cuales se espera que el cliente adquiera. Es esencial distinguir entre la salida de producto por un lado, y las ventas del mismo por otro, para poder medir la productividad de los recursos en la empresa.

Los gerentes comerciales de una empresa de retail pueden ser muy productivos en la creación de un producto extendido, pero la empresa puede encontrarse con que hay muy poco interés en la demanda de ese producto. Este cambio en la demanda implicaría un cambio en las ventas, pero a pesar de eso, no se puede concluir que la productividad de los recursos ha cambiado.

Otro elemento que impide que el volumen de ventas sea una medida aceptable de la salida de una empresa de retail es el hecho que en un punto de venta de la empresa se venden diversos tipos de productos con un nivel de servicio diferente para cada tipo. El volumen de ventas sería adecuado solo en un ambiente de laboratorio o investigativo donde se suponga una empresa de retail dedicada a la venta de un solo producto, pero en la realidad, a pesar de que es una buena medida de la cantidad de unidades vendidas, esta distorsionando la productividad con el peso asignado que tiene el precio de cada producto.

Otra opción es utilizar el margen bruto como medida de salida; suponiendo que uno puede asumir competencia perfecta en el negocio de retail, el valor agregado por una unidad de negocio de retail sería una medida exacta de su salida. En la práctica, la competencia y variedades de productos rara vez cumplen las especificaciones de competencia perfecta.

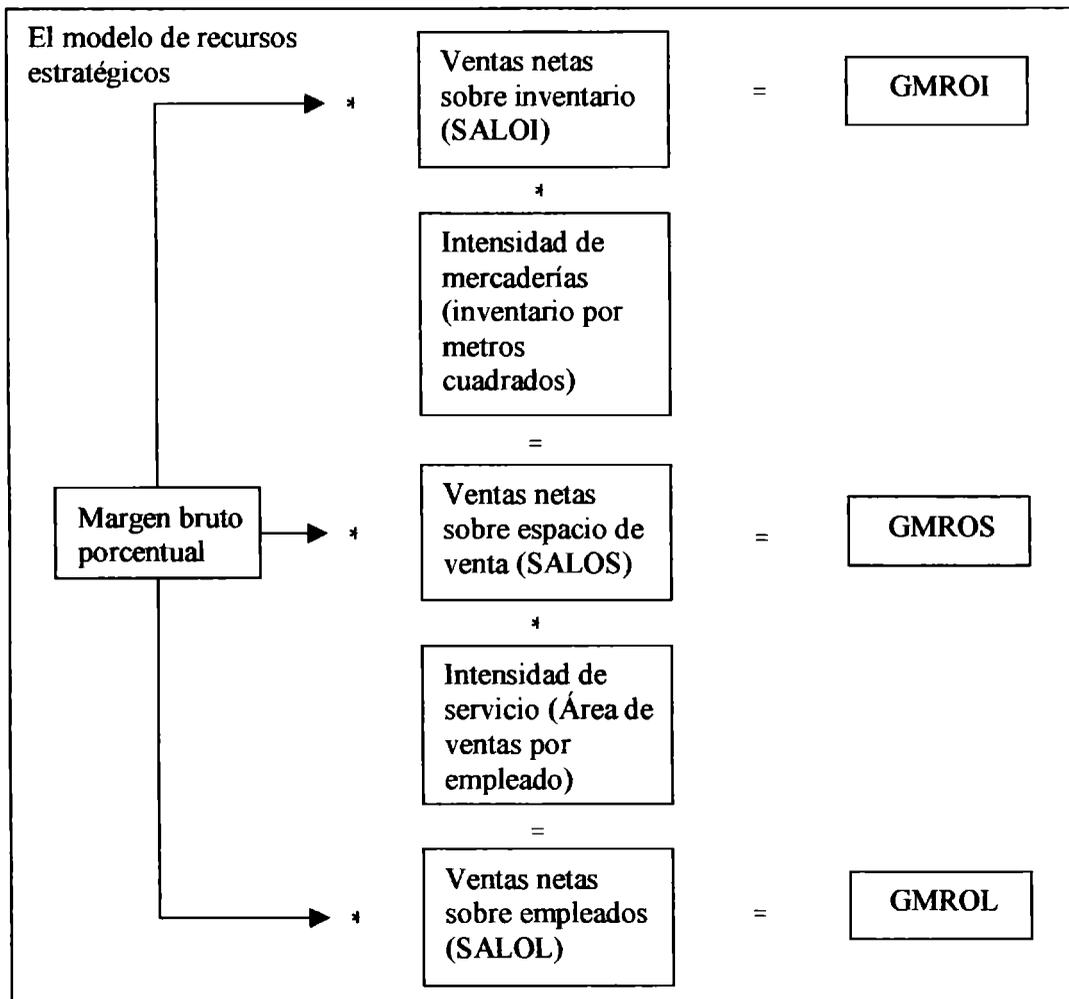
Modelo de recursos estratégico

Para la administración de una empresa de retail, el modelo de recursos estratégicos (Strategic Resource Model, SRM) reconoce a las ventas como la medida de la salida, mientras que la mercadería, el espacio físico y el trabajo son los tres recursos más importantes a ser administrados en la empresa. El modelo SRM integra el manejo del margen bruto y estos tres recursos en un modelo algebraico. Este modelo es similar en orientación al SPM (Strategic Profitability Model, modelo de Dupont), de hecho, se presenta como un complemento al SPM. La finalidad del SPM es proveer información acerca de la performance financiera global, y por ende, no ofrece información de utilidad al administrador de una empresa de retail acerca de cómo mejorar la performance. El SRM es más direccional en el sentido de que se relaciona con la estrategia de precios de la empresa y los recursos inventario, espacio y trabajo [Hernant04].

Para obtener altos niveles de rentabilidad en este negocio, el gerente comercial debe administrar efectiva y eficientemente los tres recursos. Estos tres recursos deben ser manejados colectivamente, es decir, el gerente comercial no puede manejar individualmente la mercadería sin considerar los requerimientos de espacio de la misma, ni el trabajo requerido para vender y manejar la misma.

El SRM provee un framework en el cual se refleja la eficiencia con la cual son manejados los tres recursos. El modelo es algebraico por naturaleza, es decir, mediante la multiplicación de ciertos factores financieros, el modelo proveerá un resultado. Ninguno de los dos modelos provee información acerca de cómo mejorar la performance del negocio, sino que proveen un framework para analizar que sucedería con varias medidas de performance si algún factor o entrada del modelo es alterado.

En el SRM, el precio es integrado al modelo a través del margen bruto. Si la performance global del negocio permanecerá sin afectarse, el modelo establece que una rebaja de precios requiere un incremento en la productividad de la mercadería, espacio y trabajo. El SRM provee herramientas para descubrir que pasaría con varios componentes de la performance, si algún otro componente del modelo es modificado. También, el modelo provee una herramienta de identificación, en el sentido que, a través de la división y subdivisión de las medidas de performance, enfoca áreas dentro del negocio que necesitan esfuerzos de administración.



El SRM provee también una herramienta para generar estrategias dentro del negocio. El modelo sugiere que los administradores pueden desarrollar estrategias que varían en (1) el nivel de margen bruto, (2) el nivel de productividad del inventario, (3) el grado de intensidad de la mercadería y (4) el grado de intensidad de autoservicio. La intensidad de la mercadería está definida como “inventario por metro cuadrado del área de venta”, y su multiplicación con SALOI produce una medida más popular de productividad de mercadería en el retailing: las ventas por metro cuadrado. La intensidad de autoservicio es calculada como área de ventas (metros cuadrados) dividida por el número de empleados, y provee, metafóricamente expresado, una idea de cuantos metros cuadrados tiene que “cubrir” cada empleado con su servicio.

Las medidas evaluativas en el SRM son ventas y margen bruto del retorno del inventario (SALOI, GMROI); ventas y margen bruto retornado del área de ventas (SALOS, GMROS) y ventas y margen bruto del retorno del trabajo (SALOL, GMROL). Todos estos indicadores muestran a que nivel de utilización de cada uno de los factores de entrada han sido convertidos en ventas y rentabilidad bruta para cubrir costos.

El GMROI combina la administración del margen con la administración del inventario y puede ser calculado a distintos niveles: nivel corporativo, nivel de mercado, nivel de negocio o punto de venta, nivel de departamento, nivel de clase de producto, nivel de línea de producto y nivel de SKU (stock keeping unit). De acuerdo al SRM existen dos formas de alcanzar un alto GMROI: alto margen bruto o alta rotación de inventario.

$\frac{\text{Margen bruto}}{\text{Inventario promedio}}$	=	$\frac{\text{Margen bruto}}{\text{Ventas Netas}}$	*	$\frac{\text{Ventas Netas}}{\text{Inventario promedio}}$
GMROI		Margen Bruto Porcentual		Rotación de Inventario

Vale aclarar que el inventario es medido al precio de costo, con el objetivo de proveer una medida válida del retorno del margen bruto.

De la misma manera, existen dos formas de obtener un alto GMROS, el cual es básicamente una función de las ventas por metro cuadrado, intensidad de la mercadería o la rotación del inventario.

$\frac{\text{Margen bruto}}{\text{Área de ventas}}$	=	$\frac{\text{Margen bruto}}{\text{Ventas Netas}}$	*	$\frac{\text{Ventas Netas}}{\text{Inventario promedio}}$	*	$\frac{\text{Inventario promedio}}{\text{Área de ventas}}$
GMROS		Margen Bruto Porcentual		Rotación de Inventario		Intensidad de mercadería

Si el margen bruto es incrementado mientras los otros componentes permanecen constantes, el GMROS se incrementará. Además, un incremento en el GMROS puede estar dado por un incremento en la intensidad de la mercadería, mientras que el margen bruto porcentual y la rotación de inventario permanecen constantes.

El SRM muestra que el GMROL esta determinado por el impacto multiplicativo del margen bruto porcentual, y las ventas netas por empleado. Las ventas netas por empleado es el resultado de multiplicar las ventas netas por metro cuadrado por la intensidad de autoservicio.

$\frac{\text{Margen bruto}}{\text{Empleados}}$	=	$\frac{\text{Margen bruto}}{\text{Ventas Netas}}$	*	$\frac{\text{Ventas Netas}}{\text{Área de ventas}}$	*	$\frac{\text{Área de ventas}}{\text{Empleados}}$
GMROL		Margen Bruto Porcentual		Productividad del espacio		Intensidad de autoservicio

Debe notarse que un incremento en la intensidad de autoservicio resulta en más trabajo delegado al cliente, que resulta en un aumento en la productividad del trabajo. Esto esta basado en la asunción de que el cliente acepta ese trabajo. Si esto no es así, las ventas podrían disminuir como consecuencia del menor nivel de intensidad de autoservicio, induciendo una caída de la productividad del trabajo.

En la práctica, el administrador de una empresa de retail debería esforzarse por maximizar el GMROI y el GMROS, y optimizar el GMROL. La forma más simple de visualizar la diferencia de objetivos para los indicadores es considerar el hecho de agregar un asociado de ventas. Si ese asociado generase ventas adicionales y aumentaría el margen bruto, lo cual la empresa no podría haber generado de otra forma (y el margen bruto cubre la compensación del asociado) entonces la empresa debería contratar a ese asociado. A medida que las ventas por metro cuadrado aumentan, no todos los costos fijos aumentan en proporción. Tener más asociados de ventas podría disminuir las ventas promedio por empleado pero aun así incrementar las ganancias.

~

Capítulo 4

Definiciones técnicas generales

En este capítulo se describen conceptos técnicos generales que sirven para contextualizar técnicamente la solución a construir. Se definen los conceptos de sistemas de información, bases de datos y sistemas de decisión los cuales brindan el marco dentro del cual se plantea el desarrollo.

Sistemas de información

Definición

El término sistema de información tiene los siguientes significados:

- Un sistema automático o manual que comprende gente, máquinas y/o métodos organizados para recolectar, procesar transmitir y diseminar datos que representan información de usuarios.
- Cualquier equipamiento o equipamiento interconectado de sistemas y subsistemas de telecomunicaciones o computación usado en la adquisición, almacenamiento, manejo, movimiento, control, intercambio, transmisión o recepción de voz y datos que incluyen firmware, software y hardware.

Un sistema de información es un conjunto de elementos que interactúan entre sí con el fin de apoyar las actividades de una empresa o negocio.

El equipo computacional o hardware necesario para que el sistema de información pueda operar.

El recurso humano que interactúa con el Sistema de Información, el cual está formado por las personas que utilizan el sistema.

Un sistema de información realiza cuatro actividades básicas: entrada, almacenamiento, procesamiento y salida de información.

Actividades

- **Entrada de información:** Es el proceso mediante el cual el Sistema de Información toma los datos que requiere para procesar la información. Las entradas pueden ser manuales o automáticas. Las manuales son aquellas que se proporcionan en forma directa por el usuario, mientras que las automáticas son datos o información que provienen o son tomados de otros sistemas o módulos. Esto último se denomina interfases automáticas.

Las unidades típicas de entrada de datos a las computadoras son las terminales, las cintas magnéticas, las unidades de disquete, los códigos de barras, los escáner, la voz, los monitores sensibles al tacto, el teclado y el mouse, entre otras.

- **Almacenamiento de información:** El almacenamiento es una de las actividades o capacidades más importantes que tiene una computadora, ya que a través de esta propiedad el sistema puede recordar la información guardada en la sección o proceso anterior. Esta información suele ser almacenada en estructuras de información denominadas archivos. La unidad típica de almacenamiento son los discos magnéticos o discos duros, los discos flexibles o disquetes y los discos compactos (CD-ROM).

- **Procesamiento de información:** Es la capacidad del Sistema de Información para efectuar cálculos de acuerdo con una secuencia de operaciones preestablecida. Estos cálculos pueden efectuarse con datos introducidos recientemente en el sistema o bien con datos que están almacenados. Esta característica de los sistemas permite la transformación de datos fuente en información que puede ser utilizada para la toma de decisiones, lo que hace posible, entre otras cosas, que un tomador de decisiones genere una proyección financiera a partir de los datos que contiene un estado de resultados o un balance general de un año base.

- **Salida de información:** La salida es la capacidad de un Sistema de Información para sacar la información procesada o bien datos de entrada al exterior. Las unidades típicas de salida son las impresoras, terminales, disquetes, cintas magnéticas, la voz, las impresoras y los plotters, entre otros. Es importante aclarar que la salida de un Sistema de Información puede constituir la entrada a otro Sistema de Información o módulo. En este caso, también existe una interfase automática de salida. Por ejemplo, el Sistema de Control de Clientes tiene una interfase automática de salida con el Sistema de Contabilidad, ya que genera las pólizas contables de los movimientos procesales de los clientes.

Tipos y Usos de los Sistemas de Información

Los Sistemas de Información cumplen tres objetivos básicos dentro de las organizaciones:

- Automatización de procesos operativos.
- Proporcionar información que sirva de apoyo al proceso de toma de decisiones.

- Ventajas competitivas a través de su implantación y uso.

Los Sistemas de Información que logran la automatización de procesos operativos dentro de una organización, son llamados frecuentemente Sistemas Transaccionales, ya que su función primordial consiste en procesar transacciones tales como pagos, cobros, pólizas, entradas, salidas, etc. Por otra parte, los Sistemas de Información que apoyan el proceso de toma de decisiones son los Sistemas de Soporte a la Toma de Decisiones, Sistemas para la Toma de Decisión de Grupo, Sistemas Expertos de Soporte a la Toma de Decisiones y Sistema de Información para Ejecutivos. El tercer tipo de sistema, de acuerdo con su uso u objetivos que cumplen, es el de los Sistemas Estratégicos, los cuales se desarrollan en las organizaciones con el fin de lograr ventajas competitivas, a través del uso de la tecnología de información.

Sistemas Transaccionales.

Sus principales características son:

- A través de éstos suelen lograrse ahorros significativos de mano de obra, debido a que automatizan tareas operativas de la organización.
- Con frecuencia son el primer tipo de Sistemas de Información que se implanta en las organizaciones. Se empieza apoyando las tareas a nivel operativo de la organización.
- Son intensivos en entrada y salida de información; sus cálculos y procesos suelen ser simples y poco sofisticados.
- Tienen la propiedad de ser recolectores de información, es decir, a través de estos sistemas se cargan las grandes bases de información para su explotación posterior.
- Son fáciles de justificar ante la dirección general, ya que sus beneficios son visibles y palpables.

Sistemas de Apoyo de las Decisiones.

Las principales características de estos son:

- Suelen introducirse después de haber implantado los Sistemas Transaccionales más relevantes de la empresa, ya que estos últimos constituyen su plataforma de información.
- La información que generan sirve de apoyo a los mandos intermedios y a la alta administración en el proceso de toma de decisiones.
- Suelen ser intensivos en cálculos y escasos en entradas y salidas de información. Así, por ejemplo, un modelo de planeamiento financiero requiere poca información de entrada, genera poca información como resultado, pero puede realizar muchos cálculos durante su proceso.

- No suelen ahorrar mano de obra. Debido a ello, la justificación económica para el desarrollo de estos sistemas es difícil, ya que no se conocen los ingresos del proyecto de inversión.
- Suelen ser Sistemas de Información interactivos y amigables, con altos estándares de diseño gráfico y visual, ya que están dirigidos al usuario final.
- Apoyan la toma de decisiones que, por su misma naturaleza son repetitivos y de decisiones no estructuradas que no suelen repetirse. Por ejemplo, un Sistema de Compra de Materiales que indique cuándo debe hacerse un pedido al proveedor o un Sistema de Simulación de Negocios que apoye la decisión de introducir un nuevo producto al mercado.
- Estos sistemas pueden ser desarrollados directamente por el usuario final sin la participación operativa de los analistas y programadores del área de informática.
- Este tipo de sistemas puede incluir la programación de la producción, compra de materiales, flujo de fondos, proyecciones financieras, modelos de simulación de negocios, modelos de inventarios, etc.

Sistemas Estratégicos

Sus principales características son:

- Su función primordial no es apoyar la automatización de procesos operativos ni proporcionar información para apoyar la toma de decisiones.
- Suelen desarrollarse in house, es decir, dentro de la organización, por lo tanto no pueden adaptarse fácilmente a paquetes disponibles en el mercado.
- Típicamente su forma de desarrollo es a base de incrementos y a través de su evolución dentro de la organización. Se inicia con un proceso o función en particular y a partir de ahí se van agregando nuevas funciones o procesos.
- Su función es lograr ventajas que los competidores no posean, tales como ventajas en costos y servicios diferenciados con clientes y proveedores. En este contexto, los Sistema Estratégicos son creadores de barreras de entrada al negocio. Por ejemplo, el uso de cajeros automáticos en los bancos en un Sistema Estratégico, ya que brinda ventaja sobre un banco que no posee tal servicio. Si un banco nuevo decide abrir sus puertas al público, tendrá que dar este servicio para tener un nivel similar al de sus competidores.
- Apoyan el proceso de innovación de productos y proceso dentro de la empresa debido a que buscan ventajas respecto a los competidores y una forma de hacerlo en innovando o creando productos y procesos.

Toma de Decisiones

Desde su aparición, a mediados de los años '70, las bases de datos (y la teoría sobre bases de datos) han tenido una evolución constante y han impactado en los objetivos y la evolución de los sistemas de información.

Las primeras versiones de las bases de datos se centraron alrededor de un único repositorio sirviendo a todos los propósitos orientados al procesamiento de la información (desde el transaccional, pasando por el procesamiento batch, hasta lo analítico).

En la mayoría de los casos, el principal foco de las primeras bases de datos fueron los sistemas operacionales o transaccionales.

En las últimas décadas, ha surgido una noción más sofisticada de las bases de datos. Por un lado, el objetivo de servir a las necesidades operacionales, y por otro, cubrir las necesidades analíticas de la información.

Una vez satisfecha la necesidad de tener un soporte informático para los procesos básicos de la organización (**sistemas de información para la gestión o transaccionales**).



Las organizaciones exigen nuevas prestaciones de los sistemas de información (**sistemas de información para la toma de decisiones**).

Los Sistemas para la Toma de Decisiones (Decision support Systems) son una clase especializada de sistemas que soportan actividades de toma de decisiones.

“La toma de decisiones se define como el proceso cognitivo de tomar un curso de acción entre muchas alternativas posibles. Ejemplos comunes incluyen decidir que comer y decidir a que o quien votar en unas elecciones.

Una decisión se dice que es una construcción psicológica, esto significa que una decisión no se puede ver, sin embargo, se puede inferir observando el comportamiento resultante que una decisión ha sido tomada.

En ese momento y a partir de la observación, se concluye que ha ocurrido un evento psicológico al que llamamos **Toma de Decisión (Decision Making)**” [wikipedia].

Debido a la gran cantidad de consideraciones implicadas en muchas decisiones, los sistemas de ayuda para la toma de decisiones o **decision support systems** han sido desarrollados para asistir a las personas considerando las consecuencias de diferentes cursos de acción.

Las disciplinas de DSS tratan con el uso de tecnología de información para soportar el proceso de toma de decisiones humanas. Michael Scott Morton quien prácticamente inventó la disciplina en los 70's ofrece esta definición: "los DSS constituyen un soporte basado en computadoras para ayudar a tomar decisiones de management a decisores que se enfrentan con problemas semi-estructurados" [Morton1 in Power1]. Los DSS complementan los recursos intelectuales de las personas con las capacidades de las computadoras para mejorar la calidad de las decisiones.

Para encontrar una definición del término "problemas semi-estructurados" debemos remontarnos hasta 1960 cuando Herbert Simon sentó las bases de lo que fue la Teoría Clásica de Toma de Decisiones en su libro "The New Science of Management Decision". En él, Simon define niveles de decisiones, tipos de decisiones y etapas en las decisiones. Los tipos de decisiones definidos son:

Las **decisiones estructuradas** son aquellas que cuentan con un proceso perfectamente definido para llevarse a cabo, es decir, que se realizan rutinariamente, no cambian; por ejemplo, la selección de personal en una empresa para un puesto determinado se da siempre de la misma manera.

Las **decisiones no estructuradas** no cuentan con un procedimiento predeterminado, quien debe tomar una decisión se basa en criterios, valoración y puntos de vista sobre el problema que se presenta, así como en el entorno en el cual se sitúa la problemática a resolver: por ejemplo, las decisiones sobre las inversiones de una empresa cuando genera ganancias extras.

Las decisiones **semi-estructuradas** consisten en que parte del problema sea estructurado y algunos otros elementos no lo sean. Un ejemplo de lo anterior lo encontramos en los casos de ascensos de personal, ya que algunos factores están predeterminados y otros no [Simon1 in Turban1].

Esta clasificación de tipos de decisiones también se relaciona con diferentes niveles de decisión dentro de la organización. El personal operativo se enfrenta a problemas más o menos estructurados. Quienes hacen la planeación estratégica hacen frente a problemas altamente no estructurados. Sin embargo en cada nivel organizativo existen problemas estructurados y no estructurados.

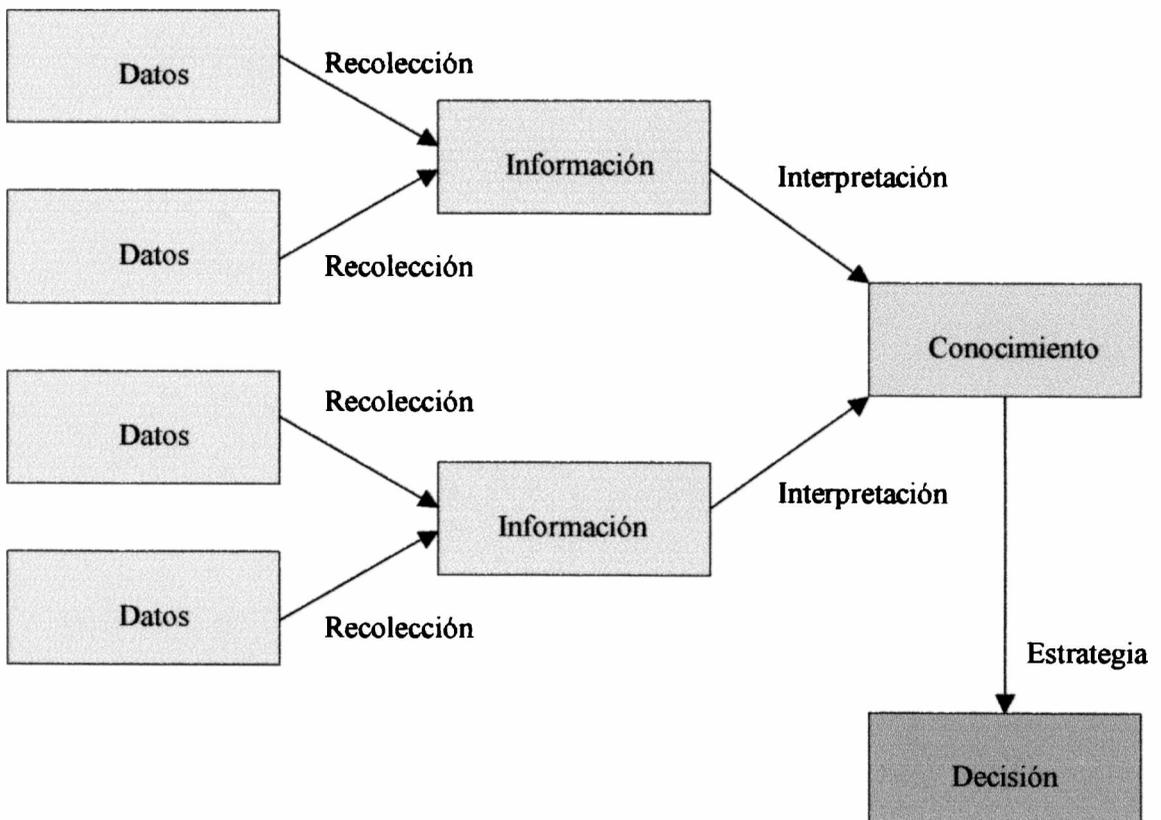
El Proceso de Decisión

En el nivel más alto, el proceso de negocios para la utilización de la información es muy simple y no ha cambiado a pesar de los rápidos avances de la tecnología de la información.

El proceso es en general un flujo que comienza como datos no estructurados que al ser recolectados pasan a ser información, éstos al interpretarse se transforman en conocimiento, y al aplicarse al conocimiento una estrategia se obtiene una decisión.

Entidades de datos discretas (archivos, pedazos de papel, datos técnicos, etc.), son recolectados juntos para conformar conjuntos de información integrada (predicciones de clima, modelos financieros, etc.). Estos conjuntos de información son interpretados dentro de un contexto de conocimientos (análisis económico, casos de negocios, teorías científicas, etc.). La estrategia es aplicada entonces al conocimiento para derivar una decisión (exploit, sell and publish).

El resultado de un proceso de decisión particular forma entonces otra entidad de datos discretos para otro proceso de decisión subsiguiente [Marley1].



Capítulo 5

Conceptos de Datawarehousing

En este capítulo se define el concepto de datawarehouse y se profundiza sobre sus características. También se describen los sistemas que se basan en un datawarehouse, sus distintos tipos, historia y evolución de las arquitecturas.

Definición de Datawarehouse

El término Datawarehouse fue acuñado por Bill Inmon a principios de la década de los '90 y lo definió de la siguiente manera:

“El Datawarehouse es una colección de datos orientados al tema, integrados, no volátiles e historizados, organizados para el apoyo de un proceso de ayuda a la decisión” [Inmon1].

- **Integrado:** Los datos almacenados en el Datawarehouse 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 Datawarehouse. 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 Datawarehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. La información almacenada en el Datawarehouse sirve, entre otras cosas, para realizar análisis de tendencias. El Datawarehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones
- **No Volátil:** El almacén de información de un Datawarehouse existe para ser leído, y no modificado. La información es por tanto permanente, significando la actualización del Datawarehouse la incorporación de los últimos valores que

tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

Obviamente que esta definición, ya clásica, debe tomarse como la definición “pura” sobre datawarehouse.

Después de diez años, sin embargo, algunos términos han sido manejados según las necesidades y capacidades del mercado, dando origen a conceptos como el de Datamart (para referirse a un datawarehouse sobre áreas específicas en lugar del warehouse corporativo) o datawarehouse volátiles, que ante la imposibilidad de almacenar toda la información histórica, almacenan una foto sobre determinado períodos, etc.

Ralph Kimball define datawarehouse de una forma más sencilla y práctica pero igual de importante, un datawarehouse es “una copia de los datos transaccionales específicamente estructurados para consultas y análisis” [Kimball1].

Por nuestro parte, podemos decir que un Datawarehouse es una base de datos orientada al análisis de la información histórica contenida en ella. Dependiendo las necesidades de análisis de la organización puede almacenarse desde unos meses hasta varios años de información.

El modelo que soporta la información que contiene se encuentra diseñado, estructurado e implementado con la finalidad y propósito del análisis y navegación de los datos. Se entiende por navegación o drilling de los datos, la posibilidad de ver información correspondiente a diferentes contextos o entornos, por ejemplo, analizar las ventas anuales y poder “abrirlos” por sucursal, después analizar en más en detalle una sucursal para ver cómo se discriminan las ventas por cada producto, etc.

Según la implementación seleccionada, los datos son almacenados en forma relacional (RDBMS, respetando ciertos estándares a nivel de definición y consulta de datos) o en formato multidimensional (las bases de datos multidimensionales son arquitecturas propietarias definidas por cada proveedor que son frecuentemente actualizadas desde base de datos relacionales).

Típicamente los usuarios tienen sólo permisos de lectura sobre el warehouse (read-only).

Comúnmente se dice que los datawarehouse son fuentes secundarias de información pues no generan información por si mismos, sino que son actualizados desde sistemas fuentes existentes internamente en la organización (sistema de ventas, sistema de presupuestos, etc.) o sistemas externos de información (datos meteorológicos, información de la competencia, cotizaciones de la bolsa, etc.). Por todo esto los datawarehouses son identificados como ambientes OLAP, On-Line Analytical Processing, en contraposición a los ambientes transaccionales clásicos (OLTP, On-Line Transaction Processing) [Bedell1].

Otras Definiciones relacionadas con un datawarehouse son las siguientes:

- **Datawarehouse:** Repositorio completo de datos de la empresa, donde se almacenan datos estratégicos, tácticos y operativos, al objeto de obtener información estratégica y táctica
- **Datamarts:** Repositorio parcial de datos de la empresa, donde se almacenan datos tácticos y operativos, al objeto de obtener información táctica ". Los Datamarts, tienen las mismas características de integración, no volatilidad, orientación temática y no volatilidad que el Datawarehouse. Se basan en una estrategia de "divide y vencerás". Esta estrategia es particularmente apropiada cuando el Datawarehouse central crece muy rápidamente y las distintas áreas de una empresa requieren sólo una pequeña porción de los datos contenidos en él.
- **Data Mining:** Técnicas de análisis de datos encaminadas a obtener información "oculta" en un Datawarehouse. El término Data Mining es una metáfora que surge como analogía de lo anterior, y afirma que con una computadora se puede encontrar de manera automática un "diamante de información" entre toneladas de datos inservibles en una base de datos, teniendo en cuenta que la información es un recurso muy valioso para una compañía tanto como lo es un diamante para un minero. Data Mining se ocupa del descubrimiento de conocimiento oculto, patrones inesperados y nuevas reglas a partir de grandes volúmenes de datos. Actualmente se considera a Data Mining el elemento clave de un proceso mucho más elaborado llamado KDD
- **KDD:** Si bien los términos Minería de Datos (Data Mining) y Descubrimiento de Conocimiento en Bases de Datos (Knowledge Discovery in Databases o KDD) son usados como sinónimos por algunos autores, el término KDD describe el proceso completo de extracción de conocimiento a partir de los datos mientras que Data Mining se refiere exclusivamente a la fase de descubrimiento del proceso completo KDD. Esta confusión se debe a que Data Mining era el término que se usaba anteriormente para describir el proceso entero de KDD.

Comparación OLTP vs. OLAP

Los sistemas operacionales y los Datawarehouses difieren ambos en diseño y funcionalidad. Entre los numerosos puntos de diferencias se encuentran, los objetivos principales de construcción, la orientación o alineación de los datos, la integración, la historicidad, el acceso de datos y manipulación, los patrones de uso, la granularidad de los datos, el perfil de los usuarios, el ciclo de vida, etc.

En ocasiones, un sistema fuente puede actuar también como una base de datos para el soporte de decisión, pero este caso no elimina las diferencias en la naturaleza de una aplicación OLAP y un sistema OLTP, de hecho posiblemente las resalte. Es importante detenerse sobre las diferencias entre ambos ambientes pues, la mejor manera de entender OLAP, es entender las diferencias con los sistemas transaccionales tradicionales [Bedell1].

Objetivos principales

Los OLTP tienen como objetivos principales asistir a aplicaciones específicas, por ejemplo ATM, y mantener integridad de los datos. Mientras que los OLAP apuntan a asistir en el análisis del negocio, identificando tendencias, comparando períodos, gestiones, mercados, índices, etc. mediante el almacenamiento de datos históricos [Bedell1].

Alineación de los datos

Los OLTP están alineados por aplicación. Diferentes sistemas tienen distintos tipos de datos, los cuales son estructurados por aplicación. Se focaliza en el cumplimiento de requerimientos de una aplicación en especial o una tarea específica.

En cambio, los sistemas OLAP están alineados por dimensión. Todos los tipos de datos integrados en un solo sistema. Los datos son organizados definiendo dimensiones del negocio (áreas temáticas o sujetos). Se focaliza en el cumplimiento de requerimientos del análisis del negocio.

Los distintos mercados verticales tienen diferentes perspectivas desde los ambientes transaccionales y desde los ambientes analíticos, por ejemplo, en el mercado bancario existen numerosas aplicaciones de cuentas y préstamos a nivel operacional mientras que en un datawarehouse la información estaría organizada por cliente, tipo de cuenta y tiempo. En los puntos de retail hay aplicaciones sobre registro de ventas, manejo de inventario o stock y presupuesto y en el ambiente de datawarehousing hablaríamos de productos, sucursales, tiempo y las diferentes variables propias de negocio (unidades vendidas, monto neto vendido, monto bruto vendido, precio promedio, unidades presupuestadas, cantidad de piezas en stock, etc.) [Bedell1].

Integración de datos

En los OLTP, los datos se encuentran típicamente no integrados, son calificados como datos primitivos o datos operacionales. Los mismos son estructurados independientemente uno de otros, pudiendo tener diferentes estructuras de claves y convenciones de nombres. Son usualmente almacenados en diferentes formatos de archivos, por ejemplo relacional, VSAM, archivos planos, etc. Incluso, si todos los datos están en formato relacional, los mismos pueden residir en diferentes plataformas de hardware y en distintas RDBMSs. Sin embargo, en los ambientes OLAP, los datos deben estar integrados. Son conocidos como datos derivados o datos DSS dado que provienen de sistemas transaccionales o sistemas de archivos maestros preexistentes en las mismas organizaciones o de sistemas externos de información.

El Datawarehouse, con el objetivo de alinear los datos por áreas temáticas, debe integrar datos operacionales estandarizando estructuras y convenciones de nombres (concepto de diccionario de datos). Este ítem da paso a otro importante tema de interés en la actualidad, Data Cleaning, que se refiere al reformateo de datos para conformar los estándares del datawarehouse. La limpieza debe ser realizada por los programas de extracción para asegurar que los datos estén en un formato consistente antes de ser cargados dentro del Datawarehouse [Bedell1].

Historia

Los OLTP usualmente y dependiendo de la aplicación retienen datos para 60 a 90 días, después son resguardados por los administradores de base de datos en almacenamientos secundarios fuera de línea (por ejemplo cintas o en disco a nivel de backup). También es común que contengan sólo valores corrientes, por ejemplo, el actual balance de cuentas para clientes y no valores históricos. Puede no incluir el tiempo como un componente de la clave. Por ejemplo, sólo el balance corriente de cuentas es almacenado, por lo tanto, no tiene sentido guardar el tiempo como parte de la clave de los datos.

En cambio los OLAP almacenan tanta historia como sea necesario para el análisis del negocio, típicamente dos a cinco años de datos históricos. Retienen valores para cada período (el atributo más atómico de la dimensión tiempo) en la base de datos. Es decir que almacenan una serie de fotos instantáneas de datos operacionales, la frecuencia con la cual define el nivel de detalle es la que se indica en la correspondiente hoja de la dimensión tiempo. Toda esta cantidad y tipo de historia apunta a ayudar a la generación de reportes de comparación de tendencias y períodos de tiempo.

Por otro lado, las bases de datos orientadas al análisis siempre contienen el tiempo como clave dado que una de las principales razones para la construcción del datawarehouse es el almacenamiento de datos históricos y el análisis a lo largo del tiempo [Bedell1].

Acceso y manipulación de los datos

Los sistemas operacionales realizan una manipulación de datos registro por registro con grandes cantidades de inserts, updates y deletes. Además necesitan de rutinas de validación y transacciones a nivel registro (OLTP, on-line transaction processing).

Generalmente poseen pequeñas cantidades de datos involucrados en un solo proceso o transacción y la puesta a punto de la base de datos para el procesamiento de transacciones, se focaliza en mecanismos de bloqueo y asignación de recursos (tuning específico).

En cambio, los datawarehouses tienen una carga y acceso masivo de datos, no se realizan inserts, updates o deletes. La carga y refresco es batch (lo que se conoce como proceso bulk copy). La validación de datos se realiza antes o después de la carga, nunca a nivel registro o transacción. Principalmente se realizan sentencias de select sobre varios registros y tablas (OLAP, on-line analytical processing), teniendo grandes volúmenes de datos involucrados en un único proceso o análisis [Bedell]. Es por ello que generalmente no se respetan las formas normales tan necesarias en los sistemas operacionales clásicos. Las anomalías que tienden a subsanar estas reglas de normalización no se presentan en los sistemas OLAP donde la carga de la información está automatizada y puede permitirse el manejo de redundancia controlada como punto para la mejora de los tiempos de respuesta de las consultas a la base de datos.

Patrones de uso

Los sistemas transaccionales normalmente mantienen un patrón de uso constante requiriendo grandes cantidades de recursos y consumiendo sólo el tiempo referido a la transacción.

En contraposición, los datawarehouses tienen un patrón de uso liviano con picos de usos eventuales en el tiempo (afectados por la disponibilidad de los datos y el flujo de trabajo del negocio). Los picos de uso suceden el mismo día de cada semana y el mismo día de cada mes (cuando los datos están por primera vez disponibles o cuando el negocio necesita por primera vez un reporte) [Bedell].

Granularidad de los datos

En los sistemas operacionales se encuentran los datos a nivel detallado o nivel transaccional. Una transacción incluye a nivel atómico cada uno de los componentes de su estructura (fecha, hora, código de cliente, código de movimiento, importe, etc.).

En un comienzo, los datawarehouses contenían información sumariada hasta cierto nivel (el permitido por los volúmenes de consolidación de los cubos multidimensionales de la década de los '80). Con el avance en el desarrollo de las bases de datos durante la década de los '90 se pudo incorporar la posibilidad de llegar incluso a nivel de detalle (nivel de transacción) en los datawarehouses. (Ver más adelante evolución de los datawarehouses, MOLAP – HOLAP – ROLAP). La diferencia en la granularidad de los datos viene dada por el uso de los mismos. Si bien un datawarehouse puede tener información a nivel transaccional, el objetivo de esta granularidad mínima está asociado con el deseo de realizar ciertos tipos de análisis que requieren que la información esté a ese nivel de detalle (análisis de market basket), pero no significa que veamos la información a nivel transaccional.

Perfil de usuario

Dado que los OLTP tienen como objetivo asistir a aplicaciones específicas y asegurar la integridad de los datos, el perfil de usuario que interactúa con dichos sistemas se encuadra dentro de los empleados operacionales de una organización (comunidad operativa).

Por el contrario, dado el objetivo estratégico y el nivel de información que manejan los DW, el perfil de usuario sobre este tipo de sistemas corresponde a la comunidad gerencial, la cual está a cargo de la toma de decisiones.

Ciclo de vida

Los ambientes operacionales pueden ser desarrollados por el clásico SDLC (Ciclo de vida del desarrollo de sistemas). Según señala Inmon [Inmon1], el datawarehouse operaba bajo un ciclo de vida bastante diferente, a veces denominado CLDS (el inverso de SDLC). El clásico SDLC es guiado por los requerimientos. En una etapa posterior se comienza con el diseño y luego con el desarrollo. El CLDS es casi la inversa, comienza con los datos. Una vez que se identifica a los datos, los mismos son integrados y luego testeados. Más tarde son desarrolladas las aplicaciones de explotación y finalmente son atendidos los requerimientos de consulta de los usuarios. Dado el particular flujo del ciclo de vida, se lo suele llamar data-driven (guiado por los datos) en contraposición al tradicional guiado por los requerimientos (requirement-driven) del SDLC.

Esta visión de Inmon, ha cambiando bastante en los últimos años. Cada vez más se le da importancia a lo que se conoce como query profile, el perfil de consulta de los usuarios, es decir, los futuros análisis que harán y cuáles son los requerimientos para hacer un análisis provechoso que ayuden a cumplir con los objetivos de negocio marcados por las organizaciones. Si bien los datos disponibles tienen un valor fundamental en el modelado y diseño del datawarehouse, en esta nueva visión se lo utiliza más para contrastar con los requerimientos relevados con los usuarios más que como fuente única de modelado. Los requerimientos del negocio son el “centro” del datawarehouse [Kimball2]

Es importante destacar también que el ciclo de vida de un datawarehouse es evolutivo y cíclico ajustándose al ciclo de vida espiral aplicado en otros ambientes de desarrollo. Es por esta y otras razones que comúnmente se dice que Datawarehousing es un proceso.

Como explica Roger Pressman [Pressman1], el paradigma del modelo espiral para la ingeniería del software es actualmente el enfoque más realista para el desarrollo de sistemas de gran escala pues permite al desarrollador y al cliente entender y reaccionar a los riesgos en cada nivel evolutivo utilizando la creación de prototipos, en cualquier etapa de evolución, como un mecanismo de reducción de riesgos.

Modelado de Datos

El modelo dimensional utilizado en el modelado de datawarehousing, organiza y presenta los datos definiendo **dimensiones** (líneas o áreas temáticas del negocio). Por ejemplo, en los modelos de retail es muy común encontrar las mediciones estructuradas por las dimensiones sucursal, producto y tiempo.

De esta forma, permiten analizar la información a distintos niveles de agregación dentro de las diferentes dimensiones. Dentro de cada dimensión se puede definir los niveles de agregación o sumalización para cada análisis, a estos niveles de granularidad se los caracteriza con el nombre de **atributos**.

Dentro de este contexto, un análisis consiste en definir el nivel de agregación que califique el conjunto de datos resultado, independientemente para cada dimensión.

Componentes básicos

A la hora de realizar un modelo dimensional del negocio (BDM) se hace necesaria la presentación de los siguientes componentes

- Dimensiones (Dimensions)
- Atributos (Attributes)
- Elementos (Attribute Elements)
- Relaciones (Attribute Relationships)
- Jerarquías (Hierarchies)
- Variables o Indicadores (Facts o Metrics)

Dimensiones

Son las **áreas temáticas, líneas del negocio o sujetos del negocio**. Las mismas proveen un método general para organizar la información corporativa. Definidas como un grupo de uno o más atributos, separados y distintos uno de otros (es decir, que no se comparten atributos), según el esquema de implementación elegido, puede ser que no se encuentren explícitamente en el datawarehouse, sino que se presenten sólo como un recurso **conceptual o lógico** que ayuda a mostrar múltiples perspectivas de los datos permitiendo realizar análisis por diferentes dimensiones o incluso cruzando información entre distintas dimensiones.

Atributos

Los atributos son una **agrupación de elementos** o ítems dentro de una dimensión. Representan **categorías** o clases de elementos que tienen el mismo nivel lógico dentro de una dimensión donde todos los elementos de un atributo se relacionan con otros atributos de la dimensión de la misma forma. La finalidad de los atributos es ver la información de cada dimensión a **diferentes niveles de detalle** y agrupar los datos para ser analizados.

Permiten definir niveles de agregación y presentar datos agrupados por uno o más atributos. Los atributos son niveles dentro de una dimensión o calificadores de una dimensión. En un modelo entidad-relación, muchos de los atributos serían clasificados como entidades y no como atributos. Por ejemplo, ítem de un producto; los ítems (como entidad del diagrama de entidad-relación) pueden tener algunos atributos característicos como ser el color y el tamaño. En un modelo dimensional, ítem, color y tamaño son todos tratados como atributos de la dimensión producto.

Elementos

Son las **instancias** o valores de los atributos que, como componentes atómicos del modelo, permiten clasificar el rendimiento del negocio. Es importante aclarar que si bien no forman parte del BDM es aconsejable su incorporación para un mayor entendimiento del modelo en etapas tempranas del ciclo de vida (relevamiento)

Relaciones

Son **asociaciones lógicas de atributos** dentro de una jerarquía definida por las instancias de los atributos y transitivas dentro de una jerarquía. Al igual que en el diagrama entidad-relación existen diferentes tipos de relaciones (clasificadas por su cardinalidad):

- Uno-a-uno (1:1)
- Uno-a-muchos (1:M)
- Muchos-a-uno (M:1)
- Muchos-a-muchos (M:N)

Es importante aclarar que los atributos dentro de una dimensión están directamente relacionados uno con otros a través de los diferentes tipos de relaciones antes definidos. En cambio, los atributos en diferentes dimensiones están relacionados uno con otros a través de los indicadores o variables del negocio definidas como intersección de las dimensiones.

Jerarquías

Representadas por un **ordenamiento lógico** dentro de la dimensión, se encuentran formadas por los diferentes tipos de relaciones entre los atributos de una misma dimensión. Pueden existir múltiples jerarquías dentro de una dimensión pero siempre es posible identificar una **jerarquía principal** o columna vertebral de la

dimensión y **jerarquías secundarias** o descriptivas compuestas por atributos característicos definidos desde la jerarquía principal.

Dentro del contexto de navegación del modelo dimensional, se puede decir que las diferentes jerarquías definen el mapa de caminos para el “**drilling**” o la “navegación” de los datos.

Haremos **drill-up** o **roll-up** cuando nos movemos hacia un atributo superior dentro de la jerarquía (navegación ascendente), **drill-down** o **roll-down** cuando analicemos información a mayor nivel de detalle (navegación descendente), **drill-within** cuando nos movamos entre la jerarquía principal y la característica desde un atributo hacia cualquier otro que no sea ni descendiente ni ancestro dentro de la misma dimensión (navegación intradimensional) y **drill-across** para analizar información sobre diferentes dimensiones (navegación interdimensional).

Como convenciones del modelado, la jerarquía principal se dibuja verticalmente desde el atributo más agregado (arriba) hasta el más atómico (abajo) y las jerarquías características se adicionan por los costados.

Indicadores

Son las **variables o métricas** que ayudarán a medir la **performance del negocio**.

Existen dos tipos de indicadores: **básicos y derivados**. Los indicadores básicos, primitivos o crudos existen físicamente en el datawarehouse junto a los atributos que los caracterizan, pueden venir de diferentes sistemas fuentes y tener distintos niveles de granularidad o agregación. Por ejemplo, la variable Venta (\$) es llevada diariamente mientras que el indicador Unidades en Stock es seguido semanalmente.

Por otro lado, los indicadores derivados o métricas calculadas se construyen a partir de los indicadores base y pueden o no estar almacenados físicamente en el datawarehouse (es típicamente una decisión de tuning). Un ejemplo clásico de métrica derivada es Margen de Ganancia, la cual se define como la resta entre Precio y Costo (ambos indicadores básicos).

Esquema Estrella

La técnica de diseño más popular usada para un Datawarehouse es el esquema estrella. El esquema estrella es un paradigma en el cual un único objeto en el centro (conocido como tabla de hechos) está conectado radialmente con otros objetos circundantes llamados tabla de dimensiones formando una estrella.

El esquema estrella puede ser simple o complejo. Un esquema estrella simple consiste de una tabla de hechos y varias tablas de dimensión. Un esquema estrella complejo puede tener más de una tabla de hechos y cientos de tablas de dimensión.

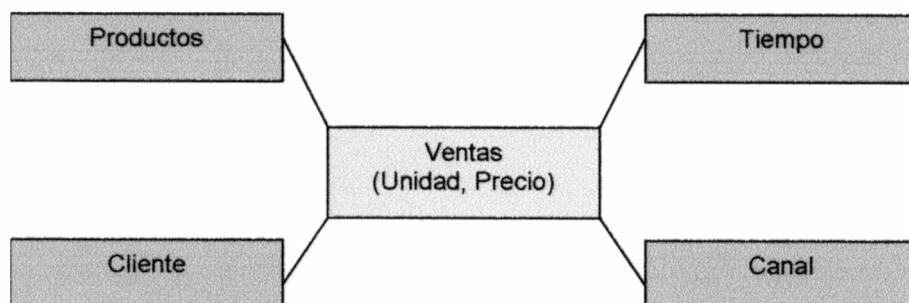
La tabla de hechos contiene las medidas básicas o transacciones del negocio y puede incluir millones de registros.

En las puntas de la estrella se encuentran las tablas de dimensión que contienen los atributos del negocio que se pueden utilizar como los criterios de búsqueda para SQL, y son relativamente pequeñas. Después de que la tabla de hechos es creada, las herramientas OLAP se pueden utilizar para el acceso a la información.

Una consulta se traduce a un ensamble entre una tabla de hechos y múltiples tablas de dimensión. Cada tabla de dimensión se ensambla con la tabla de hechos por la clave primaria de la tabla de dimensión, pero las tablas de dimensión no se ensamblan unas con otras.

Una tabla de hechos normalmente contiene claves y medidas. Por ejemplo, una simple tabla de hechos podría contener la medida Ventas, y la clave de Tiempo, Producto, y Mercado.

Habría tablas de dimensión correspondientes a Tiempo, Producto, y Mercado. La tabla de dimensión Producto, por ejemplo, contendría la información sobre cada número de producto que aparece en la tabla de hechos.



Tablas de Hechos

Tablas de Dimensión

Esquema copo de nieve

El esquema copo de nieve es una extensión del esquema estrella donde cada punta de la estrella se explota en más puntas y su denominación se debe a que el diagrama del esquema se asemeja a un copo de nieve.

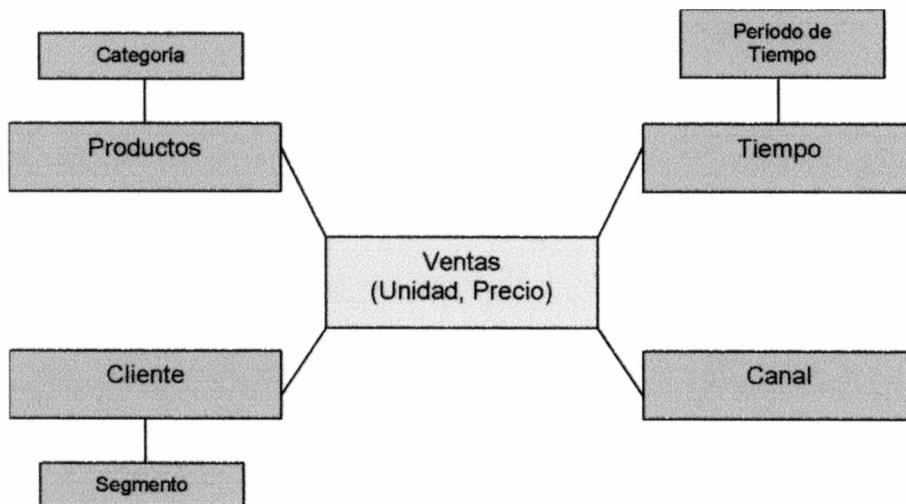
En este esquema, las tablas de dimensión se encuentran más normalizadas para mejorar la performance de las consultas debido a la reducción del almacenamiento en disco para los datos y a la mejora de la performance debido a que los ensambles se realizan sobre tablas de dimensión más pequeñas.

Los esquemas copo de nieve normalizan dimensiones para eliminar redundancia.

Los datos de las dimensiones se agrupan en múltiples tablas en lugar de una tabla grande.

Una tabla de dimensión Producto en un esquema estrella se podría normalizar en una tabla Producto, una tabla Categoría_Producto, y otra de Fabricante_Producto en un esquema copo de nieve.

El resultado es el ahorro de espacio de almacenamiento en disco en perjuicio de un aumento en la cantidad de tablas.



Tablas de Hechos 
Tablas de Dimensión 

Evolución de Arquitecturas

Las arquitecturas de los datawarehouses han evolucionado en base al crecimiento, en funcionalidad y rendimiento, de los DBMS. Es por ello, que la arquitectura reinante de los '80 fue el modelo **MOLAP** (Multidimensional OLAP). Estos Datawarehouses funcionaban con bases de datos propietarias (comúnmente denominados “cubos multidimensionales”) dado que muy pocas bases de datos soportaban el tipo de trabajo al que eran sometidas en ambientes de datawarehousing. A comienzos de los '90, con el gran crecimiento de los RDBMS, comienza a hacerse posible la utilización, en los ambientes de datawarehousing, de las bases de datos relacionales en forma directa y no como fuente secundaria para el armado de cubos multidimensionales. Es así como surge una nueva arquitectura de datawarehouse denominada **ROLAP** (Relational OLAP) la cual ya se ha consolidado como la arquitectura de los '90.

MOLAP

Capacidad de análisis:

Ofrece vistas de objetos multidimensionales

En cuanto a la resolución de consultas tiene un tiempo de respuesta cero, pues tiene todo precalculado.

Si no se precalcula todo (en general todo el precálculo tiene volúmenes inaceptables) la capacidad de análisis se limita a aquellas porciones del cubo que fueron precalculadas. Existen variantes de los MOLAP donde en caso de no poder responder a una consulta, se precalcula el cubo que responde a esa pregunta y a la periferia, pudiendo llevar esta generación varias horas de construcción (generalmente batch).

Sistema de diseño propietario:

Requiere un cubo multidimensional propietario, específico por vendedor, para almacenar los datos de forma multidimensional

Generalmente el cubo se trata de una “caja negra” de datos encriptados que pueden residir de forma local o en un servidor MOLAP.

El cliente interactúa contra la Multidimensional Database (MDDDB) vía un lenguaje de acceso propietario.

La MDDDB es alimentada desde una RDBMS.

Ofrece flexibilidad y escalabilidad limitados.

Los cambios en el modelo dimensional del negocio implican la generación de todos los cubos nuevamente.

Nuevos datos disponibles implican una ventana batch de carga bastante extensa para la generación de todos los cubos multidimensionales.

Unidad de almacenamiento:

Multidimensional data cube, unidad de almacenamiento contra la cual se disparan las consultas.

- Pre-agregación de todas las posibles intersecciones entre las dimensiones.
- Rápida respuesta de consultas obviando la agregación en tiempo de ejecución.
- Ineficiencia a medida que crecen volúmenes de datos y/o dimensiones
- Crecimiento exponencial ante cualquier adición al cubo multidimensional.
- Reducido potencial de almacenamiento de datos (normalmente 5-10 GBs)
- Reducido número potencial de dimensiones (normalmente 8-10)

Ambientes adecuados:

- Modelos dimensionales pequeños y estáticos.
- Instalaciones dónde el tiempo de respuesta sea crítico.
- Pocos volúmenes de datos.
- Análisis de información a nivel agregado.

ROLAP

Capacidad de análisis:

Ofrece vistas de objetos multidimensionales

Tiempos de respuestas que rondan entre los segundos y los minutos. Existen técnicas de tuning, caching, materialización de vistas, indexación y esquema de diseño que mejoran la performance de respuesta de los ROLAP.

La capacidad de análisis abarca todo lo que se encuentra en el datawarehouse o data-mart.

Sistema de diseño abierto:

Permite estrategias con distintos motores de base de datos. El cliente interactúa directamente contra el RDBMS vía SQL. Provee flexibilidad y escalabilidad.

Los cambios en el modelo dimensional del negocio son trasladados al DW e inmediatamente se encuentra disponible para las consultas pertinentes.

La ventana de carga del datawarehouse es menor pues no existe el tiempo de generación de los multi-cubos.

Unidad de almacenamiento:

Datawarehouse, unidad de almacenamiento contra la cual son procesadas las consultas.

Permite almacenar la información en formato relacional.

Iguala la performance de los cubos MOLAP con un apropiado esquema de diseño.

Explota al máximo las capacidades de paralelismo de los RDBMS.

No decae en eficiencia a medida que crecen los volúmenes de datos o números de dimensiones

No tiene límite inherente en cuanto a capacidad de almacenamiento (frecuentemente 100+ GBs). El límite esta dado por el RDBMS. Hay motores que están más preparados que otros para actuar en ambientes de datawarehousing, aunque en la actualidad casi todos los motores ya incorporaron funcionalidad asociada a los datawarehouses.

No tiene límite inherente en cuanto a cantidad potencial de dimensiones (frecuentemente 50+).

Ambientes adecuados:

Modelos dimensionales grandes y dinámicos.

Grandes volúmenes de datos.

Necesidad de análisis a nivel transaccional.

Otras arquitecturas

A medida que fueron consolidándose los ambientes OLAP empezaron a emerger modelos híbridos (**HOLAP, Hybrid OLAP**) los cuales trabajaban con su base de datos propietaria, al igual que los MOLAP, pero disparaban contra la base de datos relacional cuando alguna consulta no podía ser resuelta por los cubos precalculados retornando el resultado desde el RDBMS o generando el cubo correspondiente a una periferia de la respuesta para después retomar los datos desde el cubo generado. El problema de esta arquitectura es que la generación de ese segundo SQL que disparaba directamente contra la base relacional no se encontraba optimizada como para aprovechar las funcionalidades propias del motor y los tiempos de respuesta no eran los esperados. Dada la capacidad de algunos RDBMS de brindar ciertos servicios relacionados con ambientes OLAP, muchos MOLAP se auto definieron como HOLAP montando los “cubos” en bases de datos relacionales en lugar de arquitecturas propietarias.

Capítulo 6

Descripción de la aplicación

En este capítulo se realiza una definición de la aplicación desarrollada, presentando el contexto y requerimientos de la misma y definiendo la arquitectura elegida para llevar a cabo dicha aplicación.

Consideraciones generales

El desarrollo realizado utiliza el dominio dado por una empresa de venta de productos para el hogar con el formato de home center, la cual cuenta con varias sucursales y depósitos en distintas provincias. Cada sucursal dispone de uno o más depósitos para almacenar sus productos.

La estructura de esta empresa se conforma de un gerente general, gerentes por líneas de productos, y gerentes por cada sucursal.

El gerente general es el encargado de fijar los objetivos generales de la empresa junto con los directivos y de delegar objetivos a los gerentes de producto y de sucursales que permitan cumplir los objetivos más generales.

El gerente de línea de producto se encarga de todos los productos que componen la línea. Este gerente tiene a cargo tareas como la compra de productos, logística entre las distintas sucursales, condiciones de venta y promoción de los mismos, estado de existencia en los depósitos y sucursales, rentabilidad de su línea de productos. Mediante estas tareas, el gerente de producto intenta cumplir con los objetivos pautados con el gerente general y presentar información al final del período que avale el cumplimiento de los objetivos e indique aspectos relevantes de la operatoria.

Finalmente el gerente de sucursal se encarga de las tareas administrativas de la sucursal y brinda información relacionada con el sitio en donde se encuentra la sucursal al gerente de producto que le permite tomar decisiones más acertadas.

Esta empresa cuenta con 8 sucursales distribuidas en 5 provincias, las cuales ofrecen más de 60000 productos o SKUs (stock keeping units) agrupados en 11 líneas de producto.

La mayoría de los productos cuentan con un código de barras, también llamado UPC (Universal Product Code) que los identifican. Cada variante de packaging del producto tiene asignada un UPC diferente y por ende constituye un producto diferente.

La minoría de los productos o SKUs que no cuentan con un UPC, generalmente materiales para la construcción o jardín, son identificados por la empresa con un número de SKU y se le asocia un código al producto físico que permita que sea registrada su salida por las líneas de pago de las sucursales.

Las líneas de producto son: aberturas, baños y cocinas, cerámica, construcción, electricidad e iluminación, ferretería, hogar, jardín, maderas, muebles, revestimientos. Un gerente de producto administra una o mas líneas de producto y es el encargado de controlar el ROI, ROS y ROL; junto con otros indicadores complementarios de los productos de la línea con el objeto de poder modificar el precio, la superficie ocupada por el producto en una determinada sucursal, o la cantidad de empleados destinados a brindar servicios sobre esos productos. Estos indicadores, junto con otros indicadores complementarios como las ventas por mes, la rotación de inventario, y rankings de mercaderías vendidas y evolución de estos indicadores le permiten al gerente tener una herramienta que le permite tomar decisiones de manera justificada en lugar de librar las decisiones a su instinto.

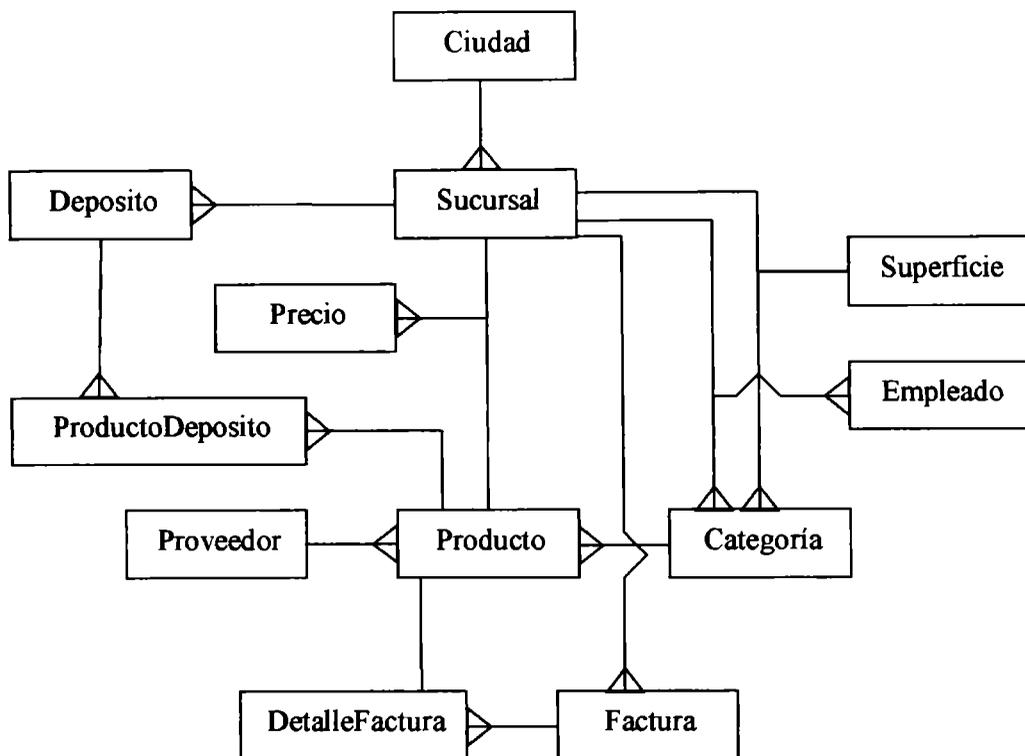
La información necesaria para gerente de producto tiene diversos orígenes como compras de mercadería a proveedores, tickets de venta a los clientes, cantidad de empleados por sucursal, precio y promoción vigente de cada producto. Uno de los puntos más ricos en información es el aportado por las cajas registradoras a medida que los clientes compran productos. El registro de la venta de productos se realiza mediante sistemas de pago o POS, los cuales scanean los códigos de barra de los productos vendidos. Otro punto interesante son los depósitos de la empresa donde los proveedores realizan la entrega de los productos comprados por la empresa, generando información de stock. También se cuenta con un registro de clientes que ingresan a cada sucursal mediante sensores en las puertas del mismo.

Por otro lado, la empresa, coordinando con cada sucursal lleva a cabo una política de descuentos cuyas motivaciones son variadas: estacionalidad, competitividad, etc.

Datos disponibles

La información generada por los procesos descritos anteriormente es concentrada en un sistema transaccional o ERP, el cual es utilizado para la administración de la empresa. Este sistema es un legacy, es decir, no consiste en un producto comprado sino que fue desarrollado a medida para la empresa.

A continuación se muestra un modelo de un subconjunto de las entidades que componen los datos del ERP. Estas entidades son las que interesan desde el punto de vista del sistema de consultas soportado por el warehouse.



De acuerdo al modelo de datos se pueden establecer las siguientes afirmaciones:

En una ciudad existen una o más sucursales, las cuales a su vez poseen uno o más depósitos.

En los depósitos están almacenados los productos que ofrece cada sucursal, aquí se lleva un registro de stock contenido en la tabla ProductoDeposito. La cantidad existente de cada producto se registra bajo demanda, de acuerdo a la llegada y salida de productos. Esta información es histórica.

En la sucursal se exhiben varias “categorías” de productos. El término “categoría” esta usado de manera genérica en este caso dado que la tabla Categoría guarda la jerarquía de categorías que clasifica los productos. Esta jerarquía se compone de: departamento, rubro, familia y producto.

Cada sucursal tiene asociado una determinada cantidad de espacio a cada rubro y una determinada cantidad de personal a cada departamento.

Un producto es provisto por un solo proveedor (que surge de una selección previa) y se actualiza a medida que el producto es provisto por otros proveedores. A este producto se le asigna la jerarquía correspondiente y se envía a los depósitos pertenecientes a las sucursales donde se desea comercializar el producto.

La comercialización de uno o más productos es registrada por un dispositivo POS (Point of Sell) que envía la información de ventas al ERP y se almacena en las tablas Factura y DetalleFactura.

Requerimientos del sistema

El objetivo general del sistema desarrollado es brindar herramientas para el soporte de procesos de toma de decisiones en este tipo de empresas, capaz de integrarse con los sistemas ERP existentes a través de interfaces específicas.

Este objetivo implica varias tareas o subobjetivos.

Se debe establecer un modelo de datos n-dimensional o datawarehouse adecuado para soportar los procesos de investigación y cálculo.

Se debe diseñar un motor de análisis de datos o mecanismo de acceso orientado al problema definido con capacidad para generar reportes de indicadores significativos en varios niveles (departamento, rubro, familia y producto).

Los requerimientos que posee esta aplicación se pueden enumerar de la siguiente manera:

1. Dada la información de ventas de un determinado período, el sistema o motor de análisis deberá poder generar los siguientes indicadores a nivel de rubro general:
 - a. MN (Margen neto después de impuestos)
 - b. ROI (Ventas sobre inventario)
 - c. ROS (Ventas sobre superficie dedicada)
 - d. ROL (Ventas sobre cantidad de empleados afectados)
 - e. GMROI (Margen sobre inventario)
 - f. GMROS (Margen sobre superficie dedicada)
 - g. GMROL (Margen sobre cantidad de empleados afectados)
2. El motor de análisis debe poder la explosión de cada departamento en sus correspondientes rubros, de los rubros en familias, de las familias en productos individuales, calculando los indicadores definidos en el punto anterior para cada uno de los niveles, cuando los datos del reporte lo permitan.
3. Los reportes generados deberán permitir múltiples alternativas de ordenamiento y filtrado incluyendo al menos la capacidad de agrupamiento por sucursal, ciudad y total; y el ordenamiento creciente o decreciente para cada uno de los indicadores exhibidos. Todos los reportes deben tener capacidad de visualización gráfica.
4. El sistema será accedido en distintas ubicaciones geográficas.

A continuación se describirán los diferentes aspectos contemplados en el desarrollo del sistema descrito haciendo énfasis en el área de desarrollo del warehouse. Respecto a la forma de presentación de los datos, no se desarrollará específicamente

sino que se adaptará una herramienta existente ya que los diferentes aspectos y problemas que implica esta área están fuera del alcance de esta tesis.

Arquitectura

La arquitectura utilizada para el desarrollo de la aplicación se basa en el modelo de tres capas. Dicho modelo es la evolución del modelo cliente servidor, en el cual se desacopla la lógica de la aplicación y las reglas del negocio que la utiliza en una nueva entidad llamada comúnmente servidor de aplicaciones. El modelo resultante queda compuesto por la capa de presentación, la capa del servidor de aplicaciones y la capa de datos.

Capa de datos

El almacenamiento persistente es manejado por uno o más DBMSs (gestores de bases de datos). La tecnología de bases de datos es bastante madura, y la mayoría de los productos existentes soportan manejo de transacciones, concurrencia, lenguajes de consulta, especificación de cuestiones de seguridad, resguardo de datos, recuperación de fallas, etc.

Sin embargo, cuando nos referimos a bases de datos, abarcamos cuestiones que van más allá de persistencia pura, como son el soporte de transacciones, manejo de concurrencia, recuperación a fallas y políticas de seguridad.

Estos puntos hacen diferencia entre almacenamiento de datos puro (usar archivos planos) y bases de datos.

Las bases de datos relacionales están basadas en tablas bidimensionales en las cuales cada ítem aparece como una fila en una tabla. Las relaciones entre los datos se encuentran comparando los valores en las distintas tablas.

Los sistemas de bases de datos relacionales son buenos manejando grandes volúmenes de datos, permiten recuperar datos rápidamente pero proveen poco soporte para manipular estos datos (los lenguajes de consulta no son computacionalmente completos). Por otro lado, los lenguajes orientados a objetos permiten expresar relaciones complejas entre los objetos, además de la capacidad de manipularlos.

Las empresas que desarrollan bases de datos han intentado aliviar esta situación haciendo posible correr código en el servidor de bases de datos. El mecanismo utilizado es conocido como "Stored Procedures" (procedimientos almacenados).

Básicamente, un stored procedure es una función que corre en un servidor de bases de datos relacionales. Para escribir stored procedures, generalmente se usa una mezcla de SQL y algún lenguaje propietario de la base de datos en cuestión.

Usando stored procedures no es necesario que los datos crudos, resultado de una sentencia SQL, viajen hasta el cliente para que este los procese. Invocando al procedimiento, los datos se manipulan en la base de datos y se transmite el resultado al cliente (se economiza el uso de las redes y el procesamiento en los clientes) [Cyran].

Los stored procedures son una característica muy poderosa de las bases de datos modernas, pero tienen importantes puntos en contra: la portabilidad es muy limitada debido a que se escriben en un lenguaje propietario, esto no sólo afecta el cambio a una base de datos distinta, sino también a diferentes versiones de la misma base de datos. Los stored procedures aumentan considerablemente el acoplamiento a la estructura de la base de datos ya que estos acceden directamente a las tablas, este acoplamiento reduce la flexibilidad. En consecuencia, algunos autores no recomiendan el uso de stored procedures [Ambler99].

Dentro de esta capa también se encuentra la Data Staging Area, un componente fundamental para el funcionamiento del data warehouse. La Data Staging Area consiste en un área de almacenamiento dentro del motor de base de datos y un conjunto de procesos llamados comúnmente como extracción-transformación-carga (extract-transformation-load, ETL). Comprende todo lo que está entre los sistemas operacionales que sirven de fuente de datos y el área de presentación de datos. En esta área, los datos operacionales crudos son transformados y cargados en el data warehouse en una manera comprensible para el usuario final.

Uno de los requisitos arquitecturales clave en el diseño de la Data Staging Area es que ella está fuera del alcance de los usuarios finales y no provee servicios de consulta y presentación.

Capa de aplicación

Esta capa es llamada "Middle-Tier" (capa del medio o capa intermedia) debido a que se ubica entre la capa de interacción con el usuario y la de almacenamiento de la información.

Aquí se encuentran los artefactos que componen el dominio de la aplicación (por ejemplo, en una aplicación bancaria encontraríamos cuentas de clientes, operaciones financieras, productos financieros, etc.).

Para desarrollar la aplicación usando una arquitectura de tres capas, debemos considerar que el sistema será usado por múltiples usuarios concurrentemente, que el software cliente accederá remotamente a nuestros objetos en el servidor y que la capa de presentación es remota con respecto al Middle-Tier y por lo tanto es costoso actualizarla.

Los problemas que resuelve este servidor de aplicaciones se pueden clasificar en tres tipos:

- La complejidad inherente al dominio de la aplicación, es decir, diseñar un modelo que soporte los requerimientos funcionales de la aplicación (la

funcionalidad esperada del sistema). Estos problemas, no son propios de una arquitectura de tres capas, ocurren en el desarrollo de cualquier sistema de información.

- En una arquitectura de tres capas, el núcleo de la aplicación debe soportar a múltiples clientes interactuando con él de manera concurrente. Esto implica que se tendrá que dar soporte de concurrencia y transacciones.
- División de la funcionalidad del sistema. Una arquitectura de tres capas requiere la especificación de la forma de interacción entre los clientes y los artefactos que forman el núcleo de la aplicación. Esto no sucede en una arquitectura de dos capas donde la presentación y la aplicación corren juntas. En una arquitectura de tres capas, los clientes acceden remotamente a la funcionalidad de la aplicación por lo que se deben definir puntos de acceso a la aplicación y cómo será la interacción en esos puntos. Esto es necesario para evitar una situación caótica en la evolución del sistema.

Capa de presentación

En esta capa se maneja la interacción entre el usuario y el sistema, generalmente se utiliza algún tipo de interfase gráfica.

En una arquitectura de tres capas, resulta indispensable el hecho de que en la interfase gráfica no se manejen cuestiones relacionadas con el dominio de la aplicación, debido a que los distintos componentes correrán remotamente entre ellos.

La estructura de la capa de interacción con el usuario se puede dividir en dos partes: presentación y control del diálogo [Renzel97]. En la presentación sólo se manejan mecanismos para la interacción con el usuario (ventanas, menús, formularios) y en el control del diálogo se define la lógica para acceder a los objetos que componen el dominio de la aplicación y se mantiene el estado de la interacción entre el usuario y el sistema.

Estos dos componentes son la base para distribuir el manejo de la interacción con el usuario entre los clientes y el servidor.

Tenemos dos opciones: presentación distribuida e interfase con el usuario remota. En el caso de una presentación distribuida, el cliente sólo corre la presentación mientras que el control del diálogo se maneja en el servidor. Por otro lado, en una interfase con el usuario remota, la interfase está basada en el cliente, este corre tanto la presentación como el control del diálogo.

Presentación distribuida:

En este caso, los clientes sólo deben manejar aspectos puramente relacionados con la parte gráfica y aspectos de interacción básicos. Esto implica que los requerimientos de procesamiento en el cliente se mantienen relativamente bajos. Por otro lado, el servidor debe dirigir el control del diálogo, esto implica conocer el estado actual de la interacción.

Para manejar la interacción con el usuario se ejecuta el siguiente ciclo: la interfase se genera en el servidor y luego es transmitida al cliente en donde es interpretada y mostrada al usuario. El usuario ejecuta alguna acción y el resultado de la interacción es transmitido al servidor en donde es procesado y, si es necesario, se genera una nueva interfase o se modifica la actual.

Los mensajes que se intercambian entre el cliente y el servidor abarcan solamente cuestiones visuales y de interacción básica (por ejemplo, informar la selección de un elemento en una lista, notificar al servidor que se cambio el texto en un campo, etc.). A causa de esto, crece considerablemente la cantidad de mensajes que intercambian el cliente y el servidor.

Tenemos dos alternativas para implementar la interacción con el usuario utilizando una presentación distribuida: utilizar formularios y web-browsers o transmitir una especificación de la interfase gráfica al cliente de la aplicación.

En la primera alternativa, un browser (o navegador) corre en la máquina cliente y se comunica con el servidor (un "Web Server") usando algún protocolo de comunicación como HTTP.

Cuando el servidor recibe un pedido del cliente genera una nueva "página" (la página es la interfase generada) y es transmitida de vuelta al cliente. El uso de protocolos estandarizados como HTML para describir las páginas y HTTP (sobre TCP/IP) para transmitirlos permite que cualquier navegador compatible pueda ser cliente de nuestra aplicación. Esta opción es muy atractiva debido a que prácticamente no se debe instalar software en los clientes, ya que es muy común que la máquina cliente tenga instalado un "web browser". [Seacord98].

La alternativa al uso de navegadores consiste en transmitir al cliente una interfase gráfica clásica, como la que proveen la mayoría de los sistemas operativos. En esta opción tenemos diferentes niveles para describir la interfase gráfica que se transmitirá al cliente.

Interfase con el usuario remota:

Aquí, el cliente corre la totalidad de la interfase con el usuario. Los requerimientos de capacidad de procesamiento en el cliente son mayores que en el caso de una presentación distribuida y la información intercambiada entre el cliente y el servidor está asociada a la aplicación, es decir, se intercambian mensajes con semántica de la aplicación. A causa de esto, el software que se debe instalar en los clientes es más complejo y de mayor tamaño, lo que es una contra en el momento de la instalación o actualizaciones de la aplicación. Por otro lado, si se lo compara con una interfase generada en el servidor, disminuye el tráfico usado para la comunicación ya que el cliente posee más inteligencia y puede resolver algunos aspectos de la interacción con el usuario sin necesidad de delegarlo al servidor.

En una interfase con el usuario remota, el control del diálogo se maneja en el cliente. Esto implica una mayor complejidad en cuanto al desarrollo de la aplicación, pero también más libertad para hacer optimizaciones apropiadas a la aplicación en particular.

Capítulo 7

Capa de datos

En este capítulo se describe la capa de datos. En esta capa se describe el modelado del datawarehouse, especificando la aproximación utilizada. También se describen los procesos de negocios elegidos para la implementación y los procesos de extracción, transformación y carga (ETL) utilizados para alimentar al datawarehouse.

Diseño del modelo del datawarehouse

El diseño del datawarehouse destinado a soportar la información generada por la empresa se realizó utilizando una aproximación de cuatro pasos [Kimball1]:

1. Selección del proceso de negocio a modelar: Un proceso de negocio es una actividad natural de negocios desarrollada en la empresa, la cual es soportada por un conjunto de datos. Ejemplos de un proceso de negocio son: la compra de productos, manejo de inventario, facturación. Cuando hablamos de proceso de negocios no nos referimos a un departamento o función de la organización. Por ejemplo, se construye un modelo dimensional para manejar órdenes de pedido en lugar de construir modelos separados para el departamento de marketing y de ventas, los cuales ambos acceden a los datos de las órdenes de pedido. El enfoque en el proceso de negocio en lugar de la estructura de la organización permite que se genere información consistente y más económica a lo largo de la organización. Si establecemos límites departamentales en los modelos dimensionales, inevitablemente se duplicaran datos con diferentes etiquetas y terminología. Esto da lugar a la posibilidad de tener inconsistencia en los datos. La mejor forma de prevenir inconsistencia es publicar los datos una sola vez. Esto también reduce el esfuerzo necesario para diseñar y mantener los procesos de ETL, y además el mantenimiento de los datos y almacenamiento.
2. Declaración del nivel de detalle (granularity) de la información: Declarar la granularidad significa especificar exactamente que significa un registro de la fact table. La granularidad conlleva al nivel de detalle con que están asociadas las medidas de la fact table. Ejemplos del nivel de detalle son: un ítem del ticket de venta entregado al cliente o un snapshot diario del nivel de inventario existente en el depósito. Este paso es crítico para el diseño del warehouse, ya que una mala elección del mismo puede invalidar la implementación del warehouse completa. Preferentemente se debe elegir el

nivel de granularidad más pequeño posible, idealmente atómico. Esto implica que al tener más detalle se podrán conocer más facts y en un futuro se evitará la necesidad de subdividir los datos para conocer más información.

3. Elección de las dimensiones que aplican a cada registro de la fact table: Las dimensiones responden a la pregunta “Como describen las personas interesadas los datos que resultan del proceso de negocios?”. La intención es agregar a las fact tables un conjunto de dimensiones representando todas las posibles descripciones que toman valores en el contexto de cada medida. Si la granularidad esta clara, las dimensiones surgirán fácilmente. Con la elección de cada dimensión, listaremos todos los atributos discretos que describen a la fact table. Ejemplos de dimensiones comunes son fecha, producto, cliente.
4. Identificación de las medidas (facts numéricos) que serán incluidos en cada registro de la fact table: Las facts se determinan respondiendo a la pregunta “Qué se esta midiendo?”. Todas las facts candidatas en el diseño deben ser consistentes con la granularidad declarada en el paso 2. Facts que corresponden a distintas granularidades deben estar almacenadas en distintas fact tables. Facts típicas son números aditivos como cantidad ordenada y precio de venta.

Este método de 4 pasos se utiliza desde el paso 1 al paso 4, los pasos 3 y 4 pueden revelar que la granularidad no fue elegida correctamente en el paso 2 con lo que se deberá volver a repetir los pasos 2, 3 y 4.

A continuación se describe el proceso anterior aplicado al diseño de la aplicación para retail:

Paso 1: Procesos de negocios

Describiremos los procesos de negocios a modelar, combinando los requerimientos del negocio junto con los datos disponibles.

De acuerdo a los datos disponibles y requerimientos descriptos anteriormente, los procesos de negocios que tenemos que modelar son los siguientes:

- a. Ventas de productos a clientes mediante líneas de ventas que ingresan un ticket con los productos de venta.
- b. Control de stock de mercaderías en los depósitos de la empresa mediante los datos provistos por los sistemas de monitoreo de stock.
- c. Control de la superficie ocupada por cada sector de productos que ofrece la empresa.
- d. Informe de recursos afectados a cada sector, provisto por las sucursales de la empresa. Estos informes sirven para determinar, por ejemplo, la cantidad de espacio y empleados afectados a un sector determinado.

Paso 2: Definición de granularidad

De acuerdo a la información disponible, se intentó determinar la granularidad más detallada posible. Se describen las granularidades para cada uno de los procesos:

- a. La granularidad para este caso esta dada por el producto correspondiente a una línea de ticket de venta.
- b. En este caso la granularidad es representada por un producto determinado en una de las sucursales de la empresa. En este caso se consideran todos los depósitos afectados a una sucursal. En el caso de productos con más de un envase posible, se distinguen como productos diferentes.
- c. Para el caso de la superficie afectada se define la granularidad como la superficie ocupada por cada categoría de producto en cada sucursal.
- d. Para este proceso la granularidad establecida consiste en los empleados por sector, para cada una de las sucursales que posee la empresa.

Paso 3: Definición de dimensiones

Se definieron las siguientes dimensiones:

- a. Para este caso se definen las dimensiones **FACTURA**, **PRODUCTO**, **FECHA**, **SUCURSAL** y **PROVEEDOR**.
- b. Para este caso se utilizarán las dimensiones de **PRODUCTO**, **FECHA** y **SUCURSAL**.
- c. Al igual que para el caso anterior, se utilizarán las dimensiones de **PRODUCTO**, **FECHA** y **SUCURSAL**.
- d. Para el registro de empleados se utilizarán las dimensiones de **PRODUCTO**, **FECHA** y **SUCURSAL**.

Paso 4: Identificación de las facts:

- a. Las facts identificadas para el proceso a son las siguientes:
CANTIDAD: Cantidad de unidades vendidas de un mismo producto.
COSTO: Precio de costo unitario del producto vendido.
DESIVA: IVA aplicado (unitario) al producto vendido.
FLETE: Gastos de transporte totales.
IIBB: IIBB aplicado (unitario) al producto vendido.
MONTO: Precio de venta unitario del producto vendido.
NRO_FACTURA: Número de factura correspondiente a la venta.

El fact **CANTIDAD** es aditivo a través de las 5 dimensiones que definimos para la fact table. Podemos agrupar y desagrupar por cualquiera de las 5 dimensiones y la suma de esta fact será el correcto.

Los facts **IIBB**, **DESIVA**, **COSTO** y **MONTO** necesitan del fact cantidad para ser aditivos ya que representan montos unitarios.

El fact **FLETE**, no representa realmente el costo de flete del producto unitario sino el costo total del transporte empleado para ese producto, el cual puede ser transportado junto con otros productos correspondientes al mismo ticket. Este

campo se incluye para permitir agregar el costo de flete a los costos de venta de productos cuando se cruza la fact a nivel de agregación de factura completa y no de producto individual. Teóricamente debería separarse este campo y generar una nueva fact table, pero el costo de almacenamiento asociado es demasiado grande.

El fact NRO_FACTURA no es aditivo respecto a ninguna dimensión, solo indica a que factura corresponde el registro que representa un producto individual de la misma.

b. Las facts identificadas para el proceso b son las siguientes:

CANTIDAD: Cantidad en existencia de un producto dado

El fact CANTIDAD es semiaditivo. Esto significa que es aditivo para las dimensiones SUCURSAL y PRODUCTO, pero no para la dimensión FECHA, ya que solo indica la cantidad en un punto en el tiempo. No es posible afirmar si la cantidad en existencia de ayer es la misma que la cantidad de hoy, observando solamente los niveles de inventario.

c. Las facts identificadas para el proceso de negocio c son las siguientes:

M2: Metros cuadrados ocupados por el nivel C2 de productos

El fact M2 es semiaditivo ya que es aditivo para las dimensiones PRODUCTO y SUCURSAL pero no para la dimensión FECHA.

d. Las facts identificadas para el proceso de negocio d son las siguientes:

CANTIDAD: Cantidad de empleados asignados a la atención del nivel C1 de productos

El fact CANTIDAD es semiaditivo ya que es aditivo para las dimensiones SUCURSAL y PRODUCTO pero no lo es para la dimensión FECHA.

Atributos de las dimensiones

Una vez terminado el proceso de cuatro pasos, nos enfocamos en la definición de los atributos de las dimensiones.

Dimensión FECHA:

La dimensión FECHA es una dimensión que se encuentra prácticamente en cualquier datamart ya que todos los datamarts almacenan datos en función del tiempo.

A diferencia del resto de las dimensiones, esta dimensión puede ser construida sin la necesidad de contar con el resto de los datos.

Los atributos definidos son los siguientes:

L_YEAR_ID: Clave interna del año.

L_YEAR_NAME: Nombre del año (ej: 2001)

L_YEAR_END_DATE: Fecha en que termina el año

L_YEAR_TIME_SPAN: Cantidad de días del año
L_YEAR_YR_OF_YEAR: Año
L_YEAR_YR_START_DATE: Fecha en que comienza el año
L_QUARTER_ID: Clave interna del cuatrimestre
L_QUARTER_NAME: Nombre del cuatrimestre
L_QUARTER_END_DATE: Fecha en que termina el cuatrimestre
L_QUARTER_TIME_SPAN: Cantidad de días del cuatrimestre
L_QUARTER_QTR_OF_YEAR: Número de cuatrimestre del año
L_QUARTER_QTR_START_DATE: Fecha en que comienza el cuatrimestre
L_MONTH_ID: Clave interna del mes
L_MONTH_NAME: Nombre del mes
L_MONTH_END_DATE: Fecha en que termina el mes
L_MONTH_TIME_SPAN: Cantidad de días del mes
L_MONTH_MTH_OF_QUARTER: Número de mes del cuatrimestre
L_MONTH_MTH_OF_YEAR: Número de mes del año
L_MONTH_MTH_START_DATE: Fecha en que comienza el mes
L_DAY_DAY: Surrogate Key
L_DAY_NAME: Fecha
L_DAY_END_DATE: Fecha (y horario) en que termina el día
L_DAY_TIME_SPAN: Igual a 1
L_DAY_JULIAN_DATE: Número de día en el calendario Juliano
L_DAY_DAY_OF_WEEK: Número de día de la semana
L_DAY_DAY_OF_MONTH: Número de día del mes
L_DAY_DAY_OF_QUARTER: Número de día del cuarto
L_DAY_DAY_OF_YEAR: Número de día del año
L_DAY_DAY_START_DATE: Fecha (y horario) de comienzo del día

Muchos de estos atributos son incluidos por su capacidad de soportar aritmética de fechas más que por la necesidad de obtener el dato que representan para un registro de la fact table.

Dimensión PRODUCTO:

La dimensión producto describe cada SKU que la empresa ofrece a sus clientes. Tomando en cuenta que la empresa ofrece aproximadamente 60.000 productos, sumado a los diferentes envases y presentaciones que posee cada producto y los productos históricos que ya no se ofrecen; esta tabla puede contener 150.000 registros y llegar hasta el millón de registros fácilmente.

Un grupo importante de atributos dentro de esta dimensión son los que describen la jerarquía de productos. Generalmente los productos individuales se agrupan en marcas, las marcas se agrupan en categorías, y las categorías se agrupan en departamentos.

Los atributos definidos son los siguientes:

TODOS_CODIGO: Código correspondiente a todos los productos
TODOS_DESCRIPCION: Descripción correspondiente a todos los productos
C1_CODIGO: Código correspondiente al nivel C1 de productos.
C1_DESCRIPCION: Descripción correspondiente al nivel C1 de productos.
C2_CODIGO: Código correspondiente al nivel C2 de productos.

C2_DESCRIPCION: Descripción correspondiente al nivel C2 de productos.
C3_CODIGO: Código correspondiente al nivel C3 de productos.
C3_DESCRIPCION: Descripción correspondiente al nivel C3 de productos.
PRODUCTO_SKEY: Surrogate Key.
PRODUCTO_CODIGO: Código correspondiente al producto.
PRODUCTO_NAME: Nombre del producto.

Los atributos TODOS, C1, C2, C3 y PRODUCTO forman la jerarquía completa de los productos de la empresa.

Adicionalmente a la creación de esta dimensión, se crearon las dimensiones PROD_C1 y PROD_C2, las cuales son análogas a la dimensión producto sólo que tienen como último nivel de detalle a C1 y C2 respectivamente. Estas dos dimensiones surgen debido a que los datos de empleados asignados a departamentos y superficie ocupada por los productos tienen una granularidad que está a nivel de C1 y C2 respectivamente.

Dimensión SUCURSAL:

Esta dimensión describe a cada punto de venta de la empresa. Esta es la dimensión geográfica primaria del modelo. Dado que las sucursales están distribuidas en distintas ciudades, la ubicación está definida en cada registro, permitiendo agrupar las sucursales por ciudad.

Los atributos definidos son los siguientes:

TODAS_CODIGO: Código correspondiente a todas las sucursales.
TODAS_DESCRIPCION: Descripción correspondiente a todas las sucursales.
CIUDAD_CODIGO: Código de la ciudad.
CIUDAD_NAME: Nombre de la ciudad.
SUCURSAL_CODIGO: Código de la sucursal.
SUCURSAL_DOMICILIO: Domicilio de la sucursal.
SUCURSAL_NAME: Nombre de la sucursal.
SUCURSAL_SKEY: Surrogate Key.

Dimensión PROVEEDOR:

Describe todos los proveedores que abastecen a la empresa de productos.

Los atributos definidos son los siguientes:

TODOS_CODIGO: Código correspondiente a todos los proveedores
TODOS_DESCRIPCION: Descripción correspondiente a todos los proveedores

PROVEEDOR_CODIGO: Código del proveedor.

PROVEEDOR_NAME: Nombre del proveedor.

PROVEEDOR_SKEY: Surrogate Key.

Dimensión FACTURA:

Describe los tipos de factura posibles que pueden ser entregados al cliente. Esta dimensión podría también contener el número de factura, pero como se habrá notado anteriormente el número de factura es parte de la fact table de ventas. Esto se debe a que

resulta más conveniente almacenar este dato, que forma parte del encabezado de la factura junto con cada ítem de la factura en lugar de crear una dimensión muy grande para almacenar cada número de factura. Si el número de factura fuera el único atributo de interés, la dimensión factura se convertiría en una dimensión degenerada, la cual no existe por si misma sino que esta embebida dentro de la fact table. En nuestro caso, dado que se desea representar el tipo de factura, esta dimensión es creada para especificar ese dato.

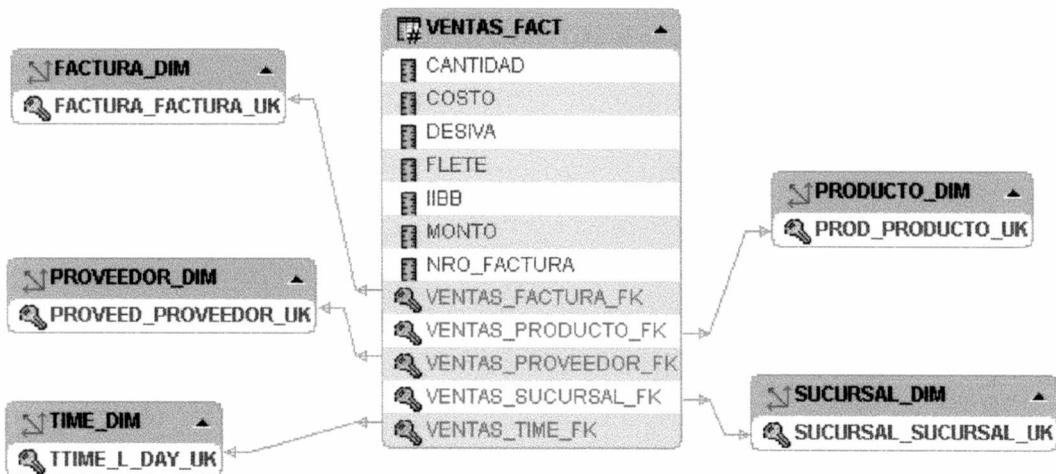
Los atributos definidos son los siguientes:

FACTURA_SKEY: Surrogate Key.

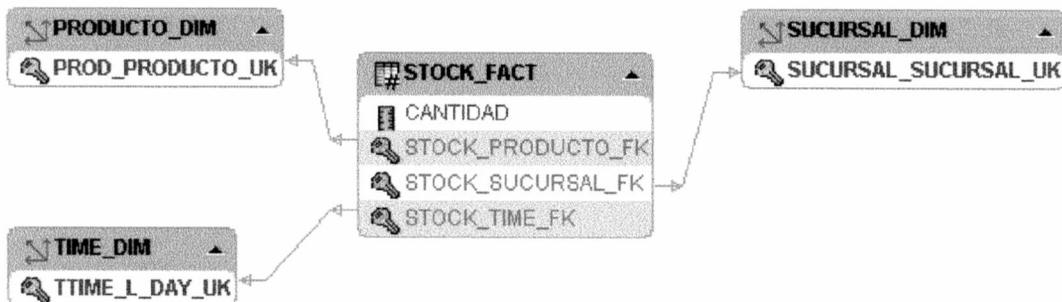
FACTURA_TIPO: Tipo de factura (A, B, D, E, R, T).

El proceso de modelado del datawarehouse se realizó con la herramienta Oracle Warehouse Builder [Alison]. A continuación se presenta una descripción gráfica de cada una de las facts con sus respectivas dimensiones:

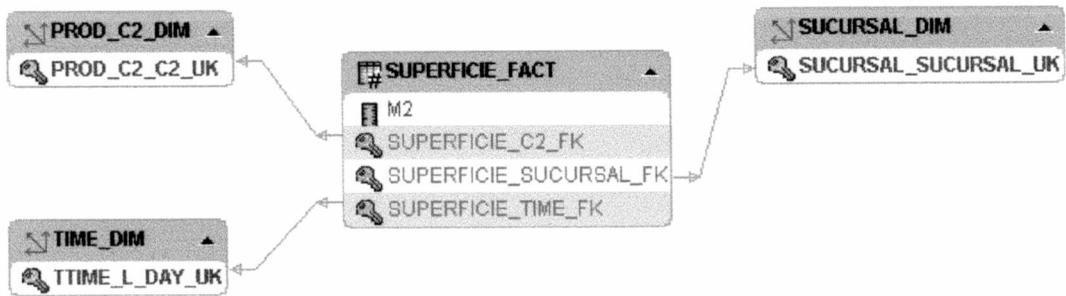
Proceso de negocios A



Proceso de negocios B



Proceso de negocios C



Proceso de negocios D



Si bien se presentan cuatro instancias de las dimensiones **TIME**, **PRODUCTO** y **SUCURSAL** se deba aclarar que estas son dimensiones “consistentes”, es decir, dimensiones que aplican a todas las facts definidas y el significado de sus atributos no varía al cambiar la fact referenciada. Esta es una propiedad importante de las dimensiones ya que implica que su diseño es extensible a las facts con que cuenta la empresa.

Data Staging Area

La data staging area de un data warehouse consiste en un área de almacenamiento y un conjunto de procesos llamados comúnmente como extracción-transformación-carga (extract-transformation-load, ETL). Comprende todo lo que está entre los sistemas operacionales que sirven de fuente de datos y el área de presentación de datos. En esta área, los datos operacionales crudos son transformados y cargados en el data warehouse en una manera comprensible para el usuario final [Scalzo].

Los procesos y datos almacenados en la data staging area no están publicados ni son accesibles por los usuarios finales. Los datos contenidos en esta área son considerados “temporales” en el sentido que están siendo modificados para obtener un dato que tenga un significado apreciable por el usuario final.

La extracción es el primer paso en el proceso de generar los datos necesarios para el data warehouse. Este proceso comprende la lectura, identificación de los datos necesarios para el datawarehouse y transferencia de los datos que interesan a la staging area para su transformación futura.

Una vez que los datos se encuentran en la staging area, existen numerosas transformaciones potenciales a realizar. Estas pueden comprender la “limpieza” de los datos (corregir errores, resolver conflicto de dominios, solucionar la ausencia de elementos), combinación de datos de múltiples orígenes, deduplicación de datos y asignación de claves internas del data warehouse. Todas estas transformaciones preceden a la carga de datos en el data warehouse.

La data staging area consiste básicamente en procesos que realizan actividades de ordenamiento y procesamiento secuencial. En la mayoría de los casos, esta área no está basada en la teoría relacional sino que consiste de un conjunto de archivos planos. Luego de validar los datos y determinar la concordancia de las relaciones uno a uno y uno a muchos con las reglas de negocios, podría no tener sentido realizar la construcción de un modelo de datos físico que cumpla con la tercera forma normal.

Es aceptable crear un modelo normalizado para soportar el proceso de ETL, pero no es el objetivo final. Las estructuras normalizadas no deberían estar al alcance de las consultas de los usuarios finales debido a que atentan contra la performance y la claridad del sistema respecto a los datos que contiene. Por defecto los modelos normalizados son excluidos del área de presentación ya que esta debe ser estrictamente dimensional.

El paso final en el proceso de ETL consiste en la carga de datos al modelo dimensional. La carga en el modelo dimensional generalmente toma la forma de una presentación de tablas dimensionales cuya calidad está aprobada a un proceso de carga masiva. El modelo destino luego deberá indexar los datos por cuestiones de performance. Cuando el modelo completo fue cargado, indexado, recalculado sus agregaciones y validado, los usuarios son notificados acerca de la disponibilidad de nuevos datos. Esta publicación incluye la comunicación de cualquier cambio en las dimensiones o de los nuevos significados o reglas que siguen los datos calculados o las medidas, junto con las causas de estas alteraciones.

A pesar de que muchas personas solo se focalizan en la etapa de extracción y carga, la etapa de transformación es muy importante y no se debe dejar de lado ya que realiza tareas como la combinación de datos, tratar con problemas de calidad de datos, identificar datos actualizados, manejar claves del data warehouse, construir agregaciones y manejar errores.

El diseño de la data staging area puede dividirse en dos partes [Kimball1]:

1. Staging de dimensiones

Dado que las dimensiones necesitan ser reutilizadas por todos los modelos dimensionales del data warehouse, generalmente son diseñadas con una visión centralizada de todo el sistema. Las dimensiones pueden ser procesadas y publicadas concurrentemente y deben ser procesadas en su totalidad antes de iniciar el procesamiento de los datos destinados a las fact tables.

El diseño del proceso de staging de dimensiones comprende los siguientes pasos [Kimball2]:

a. Extraer los datos de las dimensiones del sistema operacional: Los datos destinados a las dimensiones deben ser extraídos del sistema operacional por algún método que puede ser ftp de archivos o transferencia de streams de datos. Se debería auditar y obtener estadísticas de este proceso.

b. “Limpiar” los valores de los atributos: Este paso consiste en realizar acciones que permitan solucionar problemas relacionados con valores inconsistentes, códigos con diferentes significados a través del tiempo, datos inválidos, datos inexistentes; entre otros.

c. Realizar la asignación de claves del datawarehouse o surrogate keys: Debido a que se utilizan surrogate keys en el modelo dimensional se debe llevar un registro o tabla de referencia entre las claves del modelo dimensional y las claves del modelo operativo. Cada vez que se realiza el proceso se debe determinar por cada registro si es un nuevo registro de la dimensión, una actualización de un registro actual o ninguno de los dos.

d. Validar las dimensiones y realizar la publicación: Una vez que las dimensiones reflejan la última actualización se debe validar los datos ingresados y publicar las dimensiones actualizadas.

2. Staging de fact tables

Mientras que las dimensiones están replicadas en cada modelo donde sea necesaria su presencia las fact tables no se duplican. Dado que las fact tables se modelan de acuerdo a los procesos de negocio requeridos, están se encuentran en un lugar único, y permiten adoptar un enfoque con una visión más parcial que el utilizado para las dimensiones.

A continuación se describen los pasos necesarios para el proceso de staging de las fact tables:

a. Extracción de los datos necesarios para la fact table del sistema operacional.

b. Confirmar que las dimensiones ya están actualizadas: Se debe asegurar que se podrán encontrar en las dimensiones los registros correspondientes a los encontrados en las fact tables.

c. Dividir los datos de acuerdo a la granularidad necesaria: se deben separar los datos que se encuentran a diferentes niveles de detalle para poder realizar la carga a la fact table correcta.

d. Transformar los datos de acuerdo a lo que se requiera: algunas transformaciones que se realizan en este paso son cálculos aritméticos, conversiones de tiempo, correspondencia de monedas o unidades de medida y manejo de nulls.

e. Reemplazar las claves operacionales por surrogate keys: Se deben recorrer los datos a ingresar en la fact table, reemplazando cada clave operacional por la más reciente surrogate key definida por las dimensiones. Si la clave operacional de un registro de la fact table no coincide con ninguno de la dimensión, se puede abortar el proceso, rechazar ese registro y guardarlo en un archivo destinado a los registros rechazados, o asignarle una surrogate key que corresponda a un valor dummy o a un valor que indique error en la dimensión.

f. Agregar las claves necesarias de acuerdo al contexto: Algunas veces se deben agregar surrogate keys a los datos de las facts para representar alguna propiedad que no estaba explícitamente almacenada en el modelo operativo. Ejemplos de esto son las claves para las promociones o las claves para determinar la región geográfica.

g. Asegurar la calidad de los datos de la fact table: Se deben realizar comparaciones que aseguren que se cuenta los datos requeridos del operacional destinados al modelo dimensional.

h. Construir o actualizar las agregaciones o fact tables agregadas: Una vez que se cuenta con los datos a cargar se procede a sumarizar de acuerdo a las agregaciones que contenga el modelo.

i. Carga masiva de datos: Durante el proceso de carga, si ocurren errores por colisiones de registros, nuevamente se pueden adoptar varias medidas: abortar el proceso, almacenar los registros conflictivos en un archivo destinado a ese objetivo o actualizar el registro de la fact table con la suma de las medidas de los registros que colisionaron.

j. Informar a los usuarios: Finalmente se puede comunicar a los usuarios que el modelo dimensional cuenta con datos actualizados.

Diseño de la Data Staging Area

A continuación se describe la staging area generada para soportar la aplicación de retail. La staging area desarrollada consta de tres etapas de carga y dos procesos de transformación encargados de realizar la transferencia de datos entre cada etapa.

La primera etapa consiste en un directorio del Sistema Operativo destinado a almacenar los archivos de texto plano que contienen los datos provenientes de la aplicación transaccional, destinados a alimentar el datawarehouse. Cada archivo almacenado en este directorio contiene los datos correspondientes a una tabla de la aplicación. Las tablas de la aplicación fueron descritas en el capítulo 6.

La segunda etapa es un área de staging ubicada dentro de la base de datos, la cual recibe los datos gracias al primer proceso de transformación y se encarga de alojarlos, sirviendo de fuente de origen para el segundo proceso de transformación. Esta área contiene tablas e índices que permiten almacenar y acceder eficientemente a los datos que se encuentran en proceso de carga.

El primer proceso de transformación se encarga de tomar los datos de la primera etapa, realizar algunas comprobaciones y transformaciones sobre los datos y se procede a cargar los datos a la segunda etapa.

Algunas de las transformaciones realizadas en la primera etapa comprenden la eliminación de los productos que no están categorizados, si es que existen. Otros controles realizados son los que establecen que todo detalle de factura debe tener su correspondiente encabezado o que todos los artículos que aparecen en un ticket deben existir en la tabla de stock.

Para todas las comprobaciones realizadas se genera un archivo de log, indicando cuales son los datos que no fueron cargados debido a alguna de las “anomalías” citadas anteriormente.

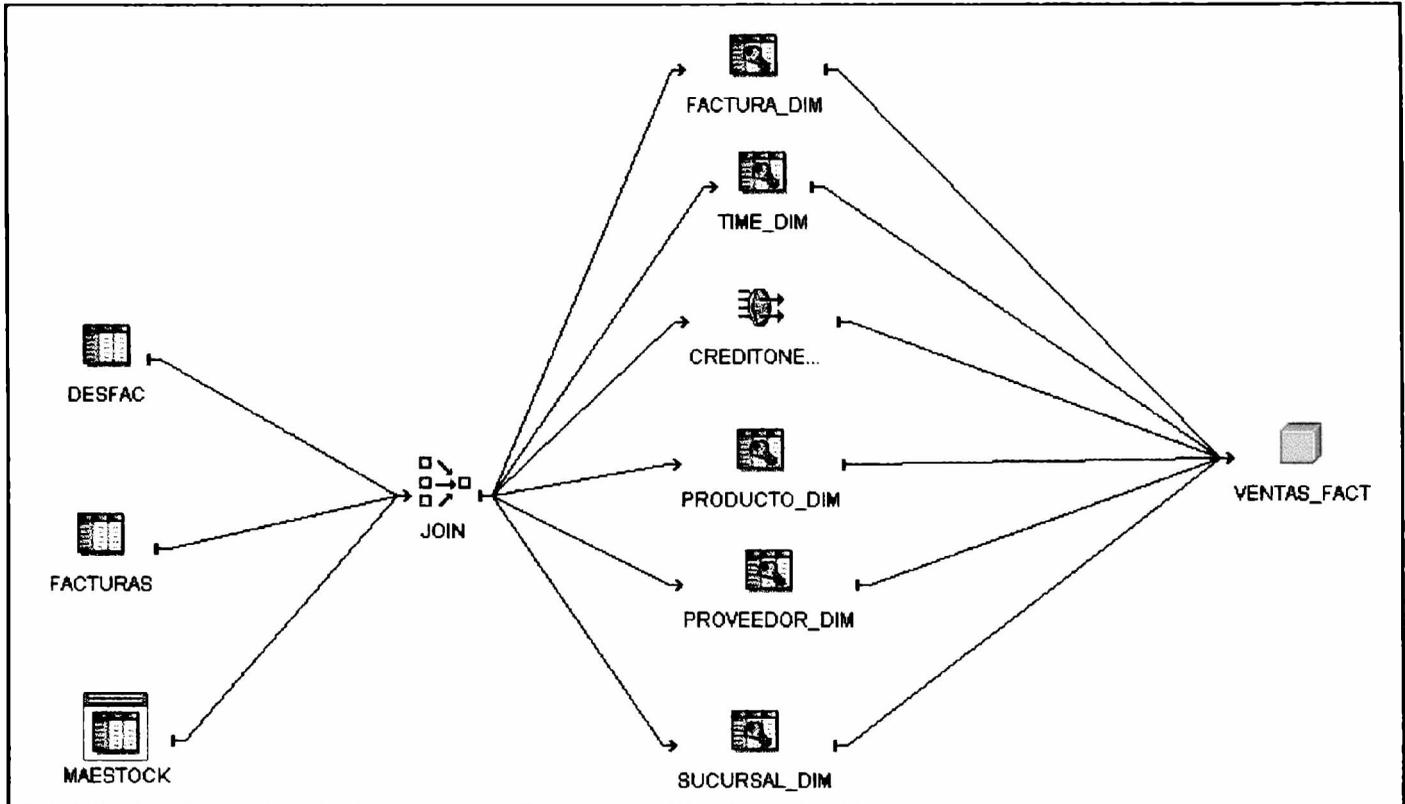
La tercera etapa comprende el modelo dimensional definido en la base de datos y consta de los cubos y dimensiones definidos anteriormente. Este modelo es el que será consultado por el usuario final y se establece como el punto final del proceso de carga y transformación.

El segundo proceso de transformación es el encargado de tomar los datos del area de staging y realizar las transformaciones y comprobaciones necesarias para presentar los datos al data warehouse. Los esquemas del diseño de este proceso se muestran mas abajo en este capitulo.

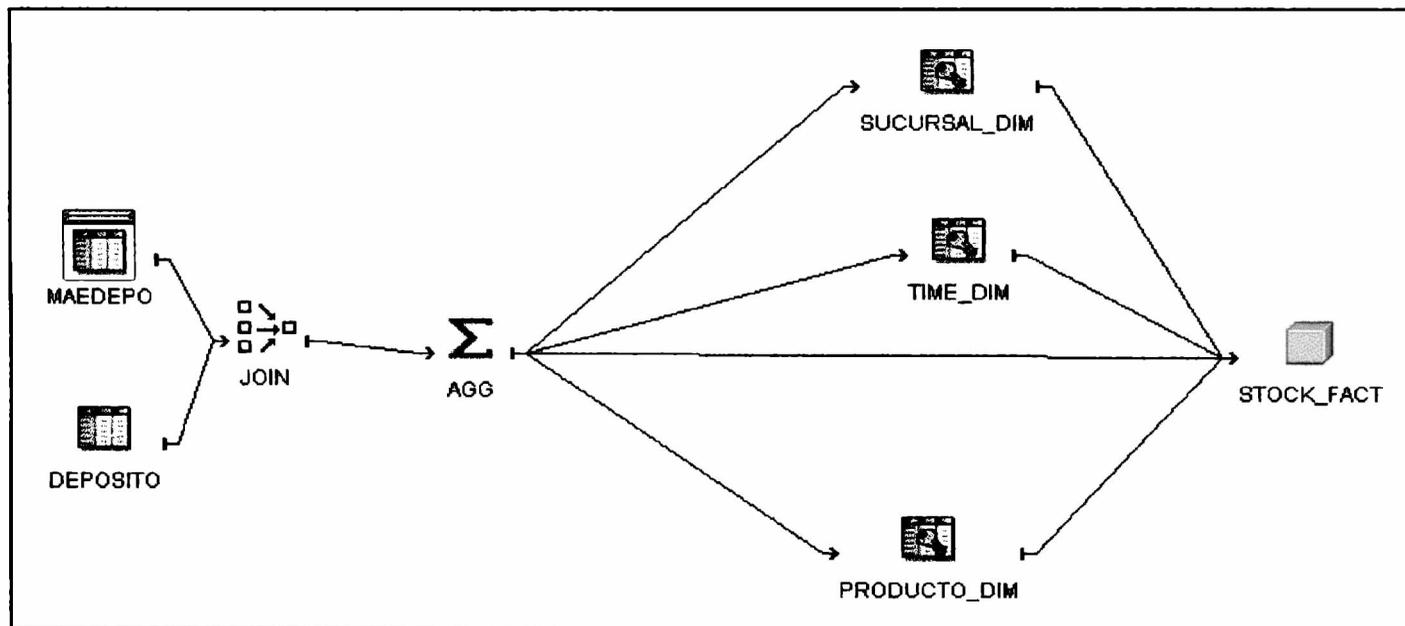
En este proceso las transformaciones que se realizan son joins para generar vistas de datos consistentes con los cubos o dimensiones, agregaciones para cumplir con la granularidad de los datos del modelo dimensional, concatenaciones de datos, eliminación de duplicados y asignación de surrogate keys.

Se muestra el diseño de algunos de los procesos realizados para soportar la alimentación del data warehouse. Vale aclarar que estos procesos detallan las transformaciones realizadas partiendo desde la staging area y llegando al data warehouse.

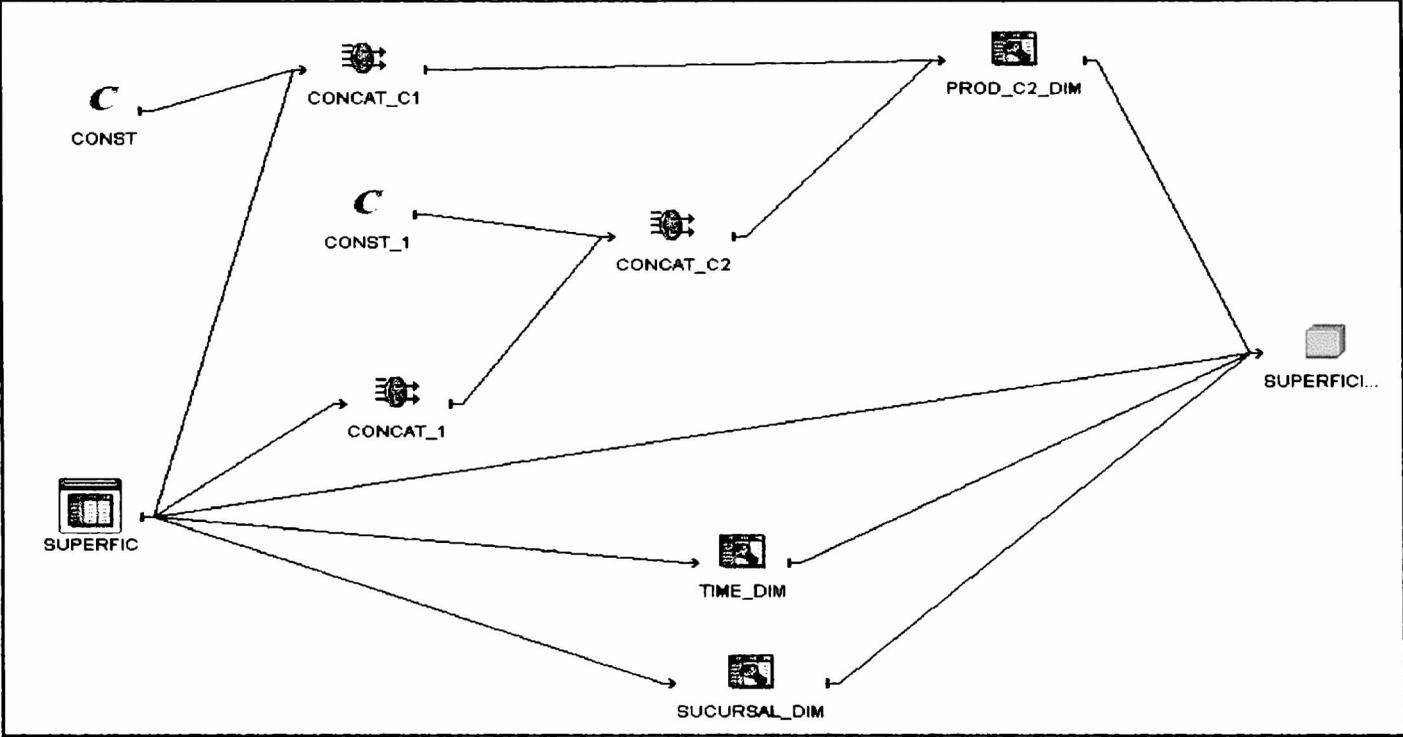
Proceso de carga del cubo de ventas



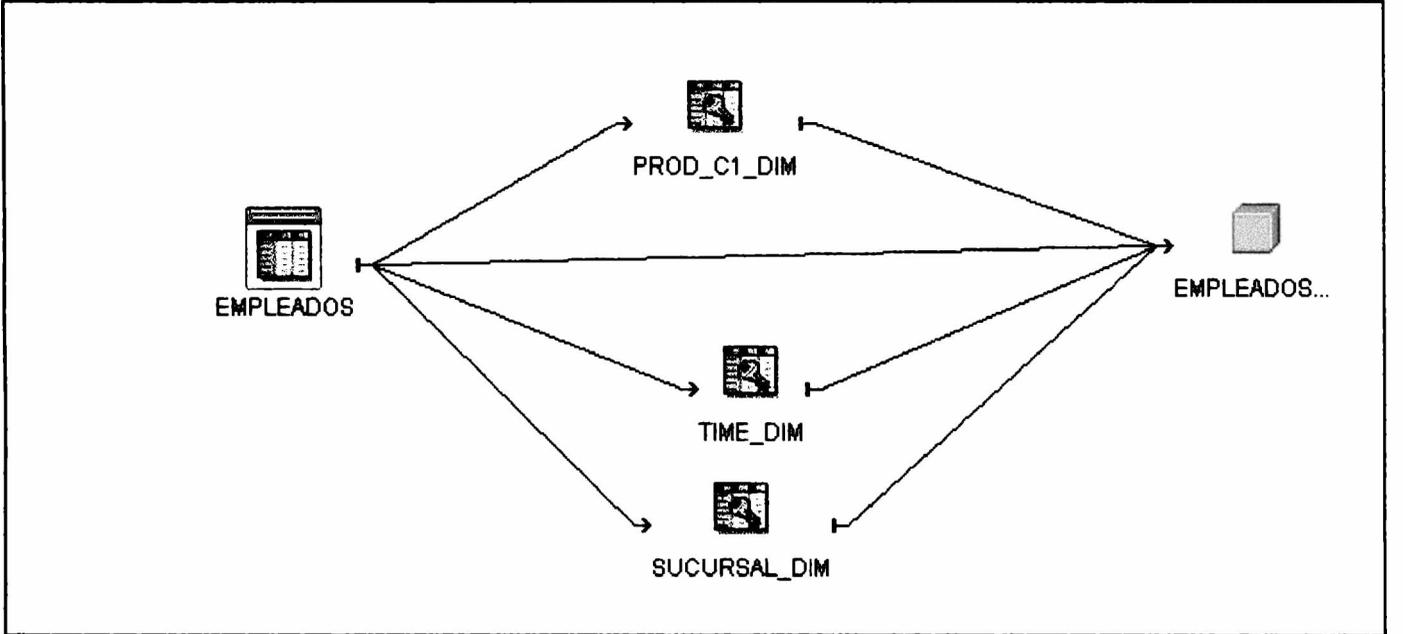
Proceso de carga del cubo de stock



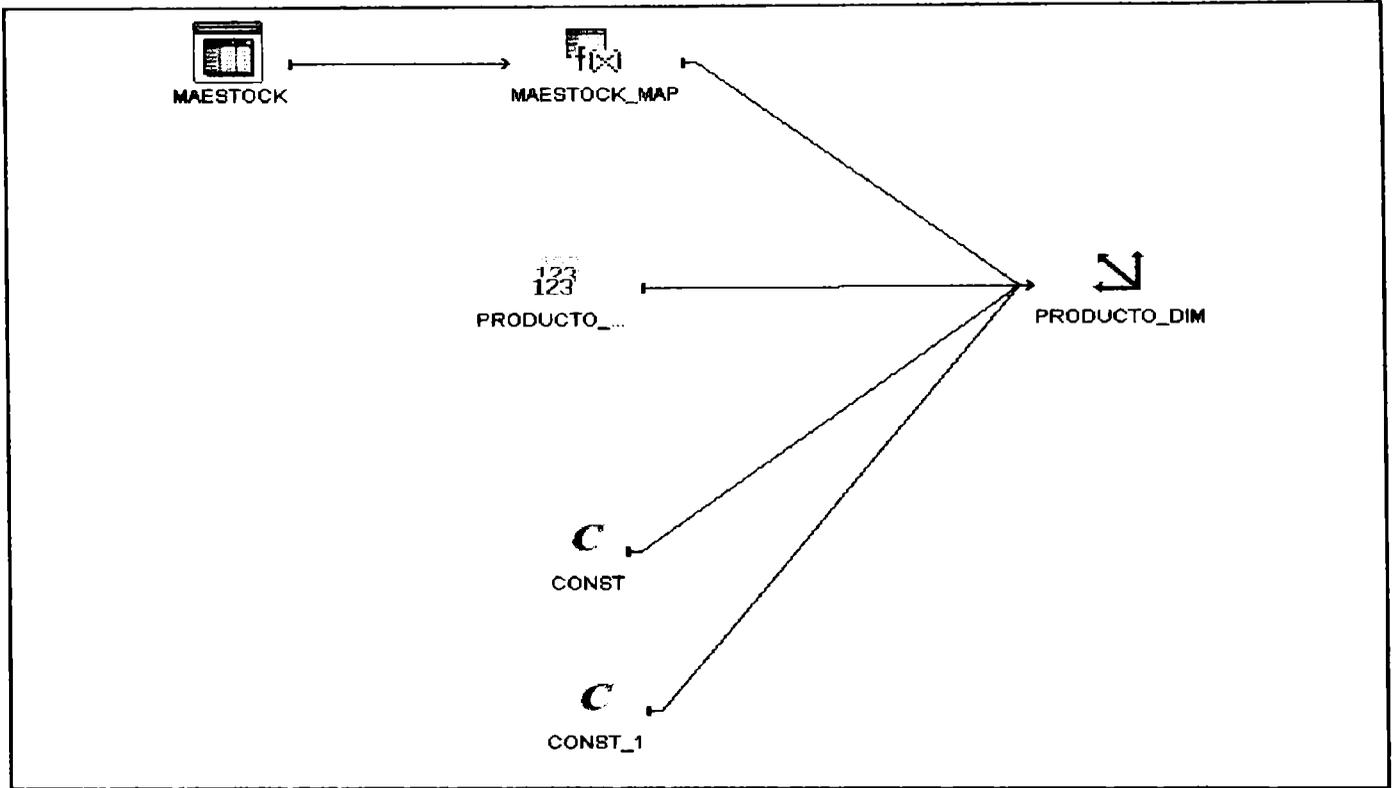
Proceso de carga del cubo de superficie



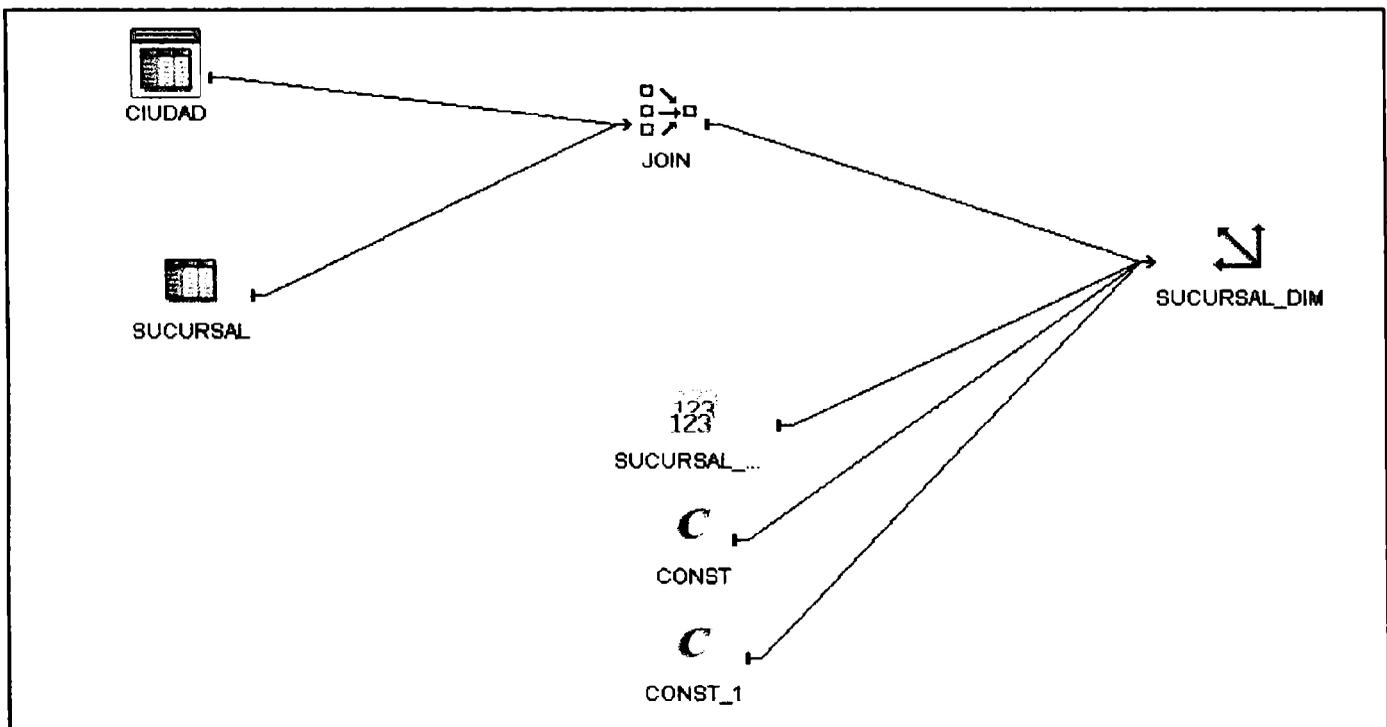
Proceso de carga del cubo de empleados



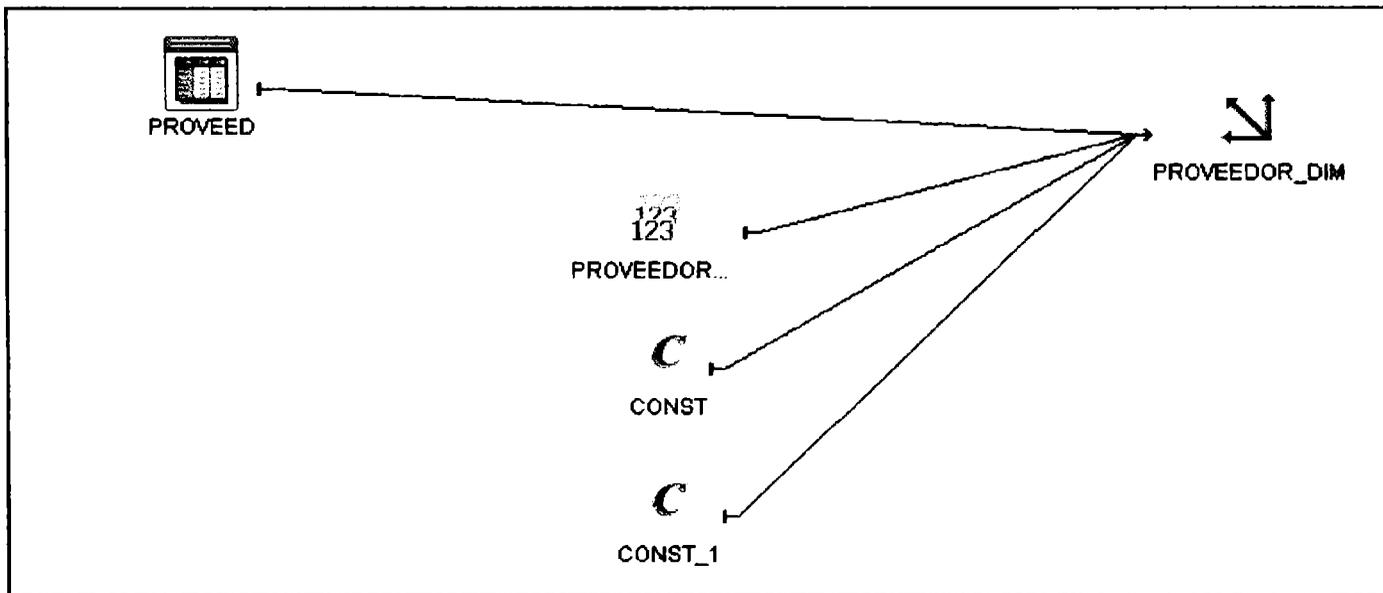
Proceso de carga de la dimensión producto



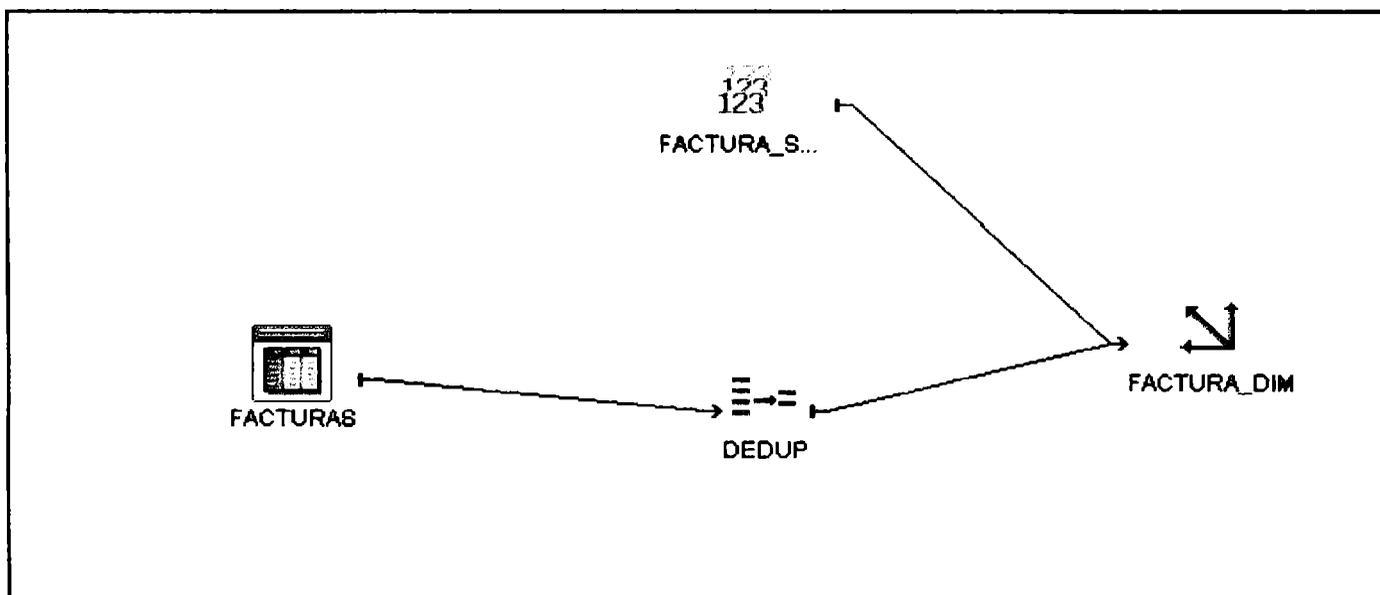
Proceso de carga de la dimensión sucursal



Proceso de carga de la dimensión proveedor



Proceso de carga de la dimensión facturas



Capítulo 8

Capa de aplicación y presentación

En este capítulo se describen las capas de aplicación y presentación, destinadas a alojar a la aplicación que realiza la interfase con el usuario y la capa que toma los datos y se encarga de los aspectos de visualización de los reportes.

Capa de aplicación

Como mencionamos en capítulos anteriores, la capa de aplicación o middle tier se encuentra entre la interfase de usuario y el servidor de base de datos. El middle tier provee la administración de los procesos donde la lógica del negocio y las reglas correspondientes son ejecutadas y provee el manejo de concurrencia entre usuarios proporcionando servicios del tipo de encolamiento y ejecución de aplicaciones.

Para el caso de un datawarehouse la capa de aplicación es donde los datos son integrados, organizados, transformados con el objetivo de facilitar al usuario su comprensión o de presentar los datos a alguna herramienta de reporting o aplicación analítica. Esta capa puede comprender datos de diferentes modelos integrados, que se encuentran contenidos en la capa de datos.

Dado que la capa de datos no es visible al usuario, la capa de aplicación significa el datawarehouse de los usuarios desde el punto de vista de los procesos de negocios.

En esta capa se debe tomar la decisión acerca de cuales entidades del modelo dimensional serán presentadas al usuario y cual será la granularidad permitida para estas entidades. Junto con estas decisiones se deberá decidir si el usuario puede realizar consultas ad hoc al datawarehouse o solo verá la información a través de reportes predefinidos. Dado que consultar los datos del datawarehouse a un nivel de detalle muy pequeño puede plantear problemas cuya solución no es trivial.

Desarrollo de la capa de aplicación

Para la aplicación desarrollada se decidió utilizar el servidor de aplicaciones Oracle 9iAS. Este es un servidor de aplicaciones compatible con los estándares como J2EE y XML. Permite correr sitios web y aplicaciones J2EE [9iAS1].

Las aplicaciones J2EE están compuestas de componentes. Un componente J2EE es una unidad funcional de software auto contenida que esta incluida dentro de una aplicación J2EE junto con sus clases relacionadas y los archivos XML que sirven como descriptores de la comunicación con otros componentes.

Los componentes J2EE están escritos en lenguaje Java y son compilados de la misma forma que cualquier otro programa escrito en Java. Los descriptores de la comunicación utilizan el lenguaje XML para describir las interfaces de los componentes.

Algunos de los elementos que contiene este servidor de aplicaciones y que han sido utilizados en la construcción de la aplicación ejemplo son [9iAS2]:

- HTTP Server: servidor de HTTP standard.
- J2EE Containers: son un conjunto de clases Java que brindan distintas funcionalidades al servidor.
- Discoverer: consiste en la parte de la aplicación destinada a extraer información del datawarehouse, alojada en el servidor de aplicación

El último elemento comprende la aplicación elegida para acceder al datawarehouse y retornar los resultados al usuario. La aplicación Discoverer también se basa en una arquitectura de tres capas. La capa de datos de la aplicación consiste en un repositorio que almacena la metadata necesaria para acceder al datawarehouse. Esta metadata consiste en la definición de una capa que abstrae los elementos del datawarehouse y las consultas que forman los reportes predefinidos. La capa de aplicación se describe a continuación y la capa de presentación se describe más adelante.

Los componentes principales de la aplicación Discoverer son [Russell]:

1. Componentes J2EE: Estos componentes comprenden entre otros un servlet, el cual básicamente consta de clases Java corriendo en la capa de aplicación destinados a recibir requests HTTP del cliente y enviar una respuesta a cada request generada en una página HTML con un procesador de XML/XSL. La máquina virtual de java que corre estos servlets se encuentra en el servidor de aplicaciones Oracle iAS.
2. Componentes CORBA: Los componentes CORBA (Common Object Request Architecture) contenidos son los siguientes:

- a. Un componente de sesión, responsable de activar una sesión cuando el usuario accede con el web browser a la dirección de la aplicación. Realiza operaciones tales como conectarse al servidor de datos y al repositorio de metadata del reporte. Este componente es el que provee el enlace entre el servlet y el repositorio donde se encuentra el datawarehouse. Existe un componente de sesión por cada usuario activo.
- b. Un componente de preferencias el cual almacena las apariencias y comportamiento por defecto de los componentes destinados a la capa de presentación.
- c. Un componente llamado Object Activation Daemon: corre en background y es el encargado de correr concurrentemente y administrar los componentes de sesión.

Capa de presentación

Esta capa comprende las herramientas destinadas a mostrar al usuario la información que este solicita. Se utiliza el término herramienta para referirse a la variedad de aplicaciones y capacidades que pueden ofrecerse a los usuarios de negocios. Su objetivo es interactuar con el usuario y la capa de aplicación para obtener los datos y presentarlos de la manera deseada para facilitar el análisis y toma de decisiones.

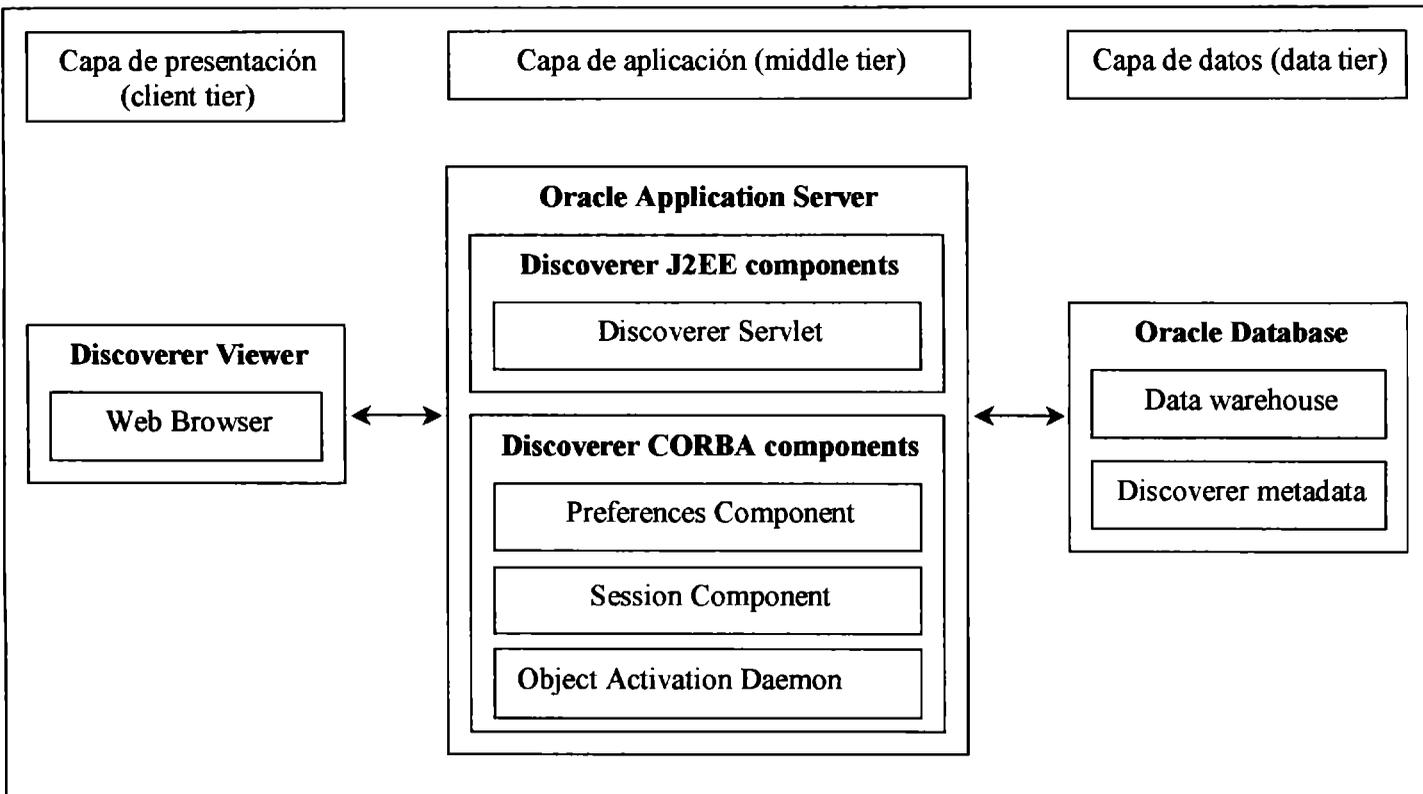
Una herramienta de presentación puede ser tan simple como una interfase para realizar consultas ad hoc o tan compleja como una aplicación de reporting. Las herramientas que consisten en consultas ad hoc, debido al poder que conceden, son utilizadas solo por un porcentaje pequeño de la población potencial capaz de acceder al datawarehouse. La mayoría de los usuarios accederán a través de aplicaciones que ya ofrecen un conjunto de consultas parametrizadas. Aproximadamente el 80% o 90% de los usuarios serán atendidos por estas aplicaciones que esencialmente contienen templates terminados que evitan que el usuario deba construir su propia consulta para explotar el datawarehouse.

En el caso de la aplicación desarrollada para el negocio de retail se provee una presentación utilizando un web browser que se conecta a la capa de aplicación y ofrece distintos reportes parametrizados a los que el usuario puede acceder. También ofrece funcionalidad de drilling y pivoting para que el usuario pueda visualizar los datos de acuerdo a sus preferencias.

Desarrollo de la capa de presentación

Para la capa de presentación se utiliza la opción Discoverer Viewer, la cual es la contraparte de la aplicación que reside en el servidor de aplicaciones y consiste en una herramienta web que permite acceder a reportes predefinidos. La implementación de esta herramienta adopta la opción de presentación distribuida descrita en el capítulo 6 [Discoverer1].

Se brinda un gráfico que detalla la arquitectura de la solución adoptada:



Discoverer Viewer es una herramienta destinada a consultar datawarehouses y brindar la información a usuarios no técnicos de una manera comprensible, permitiéndoles realizar el análisis de los datos. Básicamente consiste en un web browser mediante el cual el usuario accede a la aplicación. Este web browser debe procesar HTML y Javascript.

La herramienta brinda la capacidad de acceder a reportes predefinidos y funcionalidad de drilling y pivoting, pero no permite crear nuevos reportes o alterar las consultas que brindan los datos que forman el reporte [Discoverer2].

La funcionalidad de drilling en un reporte permite al usuario ver la información con diferentes niveles de detalle. Por ejemplo, en un reporte de ventas por rubro en donde inicialmente la información aparece totalizada por rubro, el usuario podría realizar la operación de drill down para obtener la información sumariada por subrubro y si realiza un nuevo drill down accedería a la información sumariada por producto. La operación de drill up es inversa a la operación descrita.

La funcionalidad de pivoting permite al usuario alterar el esquema del reporte

La funcionalidad otorgada por esta herramienta se puede describir de la siguiente manera: Cuando un usuario ejecuta un web browser e ingresa la URL destinada al Discoverer Viewer se produce un acceso al servlet de Discoverer en el servidor de aplicaciones. El servlet retorna la página HTML que solicita los datos de conexión. Una vez que el usuario ingreso los datos de conexión, estos son recibidos por el servlet, el cual utiliza al componente OAD (Object Activation Daemon) para tener acceso a un componente de sesión e iniciar la sesión de usuario.

A partir de aquí el componente de sesión interactúa con el datawarehouse para enviar y recibir datos y el servlet interactúa con el usuario recibiendo requests y enviando respuestas en formato HTML.

A continuación se presenta el listado de los reportes que comprende la aplicación. En el apéndice A se adjunta la descripción detallada de cada reporte. Los reportes principales, que definen las variables necesarias para soportar el modelo SRM presentado en el capítulo 3 son:

1. Margen Neto / ROI (Return of Investment)
2. ROS (Return of Surface)
3. ROL (Return of Labor)
4. Porcentuales
5. GMROI – GMROS – GMROL
6. Evolución GMROI – GMROS – GMROL

También se presentan algunos reportes complementarios que sirven para tener más información acerca del estado de situación y justificar el comportamiento de las variables descriptas anteriormente:

7. Ventas por mes
8. Evolución de ventas por mes
9. Evolución de ventas por sectores por mes
10. Comportamiento de las compras vs. las ventas
11. Informe de ventas general
12. Ranking de ventas por articulo
13. Ranking de ventas por rubro y por articulo
14. Ranking de ventas por proveedor
15. Ranking de ventas por rubro y por proveedor
16. Ranking de ventas totalizado por rubro
17. Informe de ventas por proveedor (Ranking de productos)
18. Informe de ventas por articulo

- 19. Informe de ventas por subrubro
- 20. Mercadería sin movimiento de ventas
- 21. Margen bruto por rubro y por artículo
- 22. Ranking de margen bruto unitario

Ejemplos de reportes

A continuación se muestran ejemplos de algunos de los reportes desarrollados

Margen Neto / ROI (Return of Investment)

MN - ROI										
Sectores	Costos	Venta Bruta	Vta. s/Imp	Part. s/Total	Margen Neto	Margen %	Cont. Marginal	Stock	ROI	
1 ABERTURAS	4853,76	9707,46	3158,27	0,80%	-1771,49	-58,09%	-3,12%	162578,90	-0,01	
2 BAÑOS Y COCINAS	15629,24	31257,71	17474,09	3,30%	1538,85	8,81%	2,71%	789665,15	0,00	
3 CERAMICA	20243,65	40477,94	21777,59	4,12%	1122,92	5,16%	1,98%	407081,80	0,00	
4 CONSTRUCCION	243062,40	485657,34	292131,56	55,22%	45845,16	15,89%	80,75%	1088811,22	0,04	
5 ELECTRICIDAD E ILUMINACION	7704,96	15378,51	8840,85	1,67%	303,89	3,44%	0,54%	618002,41	0,00	
6 FERRETERIA	14693,53	29300,67	17815,95	3,37%	1570,42	8,81%	2,77%	992250,84	0,00	
7 HOGAR	31696,50	63392,24	33232,08	6,28%	1394,58	4,20%	2,46%	527733,94	0,00	
8 INSTALACIONES	62170,81	164262,91	94844,63	17,93%	6632,72	6,99%	11,88%	810299,82	0,01	
9 JARDIN	727,71	1455,10	950,06	0,18%	213,35	22,46%	0,39%	115940,82	0,00	
10 MADERAS	15938,11	31871,51	13092,57	2,47%	-3148,54	-24,05%	-5,55%	320622,45	-0,01	
11 MUEBLES	12061,53	24122,77	14176,45	2,68%	2079,92	14,67%	3,66%	320994,70	0,01	
12 REVESTIMIENTOS	9879,58	19750,39	11509,87	2,18%	995,29	8,65%	1,75%	643347,10	0,00	

ROS (Return of Surface)

generales

11 - Informe de Ventas generales - Tickets

12 - Ranking Vtas. por articulo

13 - Ranking Vtas. por rubro y articulo

14 - Ranking Vtas. por proveedores

15 - Ranking Vtas. por rubro y proveedor

15 - Ranking Cantidad Vtas. por rubro y proveedor mes/año

16 - Informe Vtas. por rubro

17 - Informe Vtas. proveedor

18 - Informe Vtas. por art. Sucursales

19 - Informe de ventas por subrubro Sucursales

20 - Mercaderia sin mov. Vtas.

20 - Mercaderia sin mov. Vtas. s/ultima entrega

20 - Mercaderia sin mov. Vtas. ultima entrega

21 - Margen bruto por rubro y articulo

Data

▼ Hide Data

Ventas x M2

	11/1/05						
	11/30/05						
	HIPERTEHUELCHE - RIO GALLEGOS						
	Venta s/Imp.	Contrib. % Vtas.	M2	% M2 Local	Ventas x M2 ROS	Part. Vtas. X M2	
ABERTURAS	105744.46	3.96%	136	3.42%	777.53	10.07%	
BAÑOS Y COCINAS	141450.95	5.30%	210	5.28%	673.58	8.72%	
CERAMICA	189579.75	7.11%	140	3.52%	1354.14	17.54%	
CONSTRUCCION	1069752.05	40.10%	135	3.42%	7865.82	101.88%	
ELECTRICIDAD E ILUMINACION	82383.42	3.09%	305	7.59%	269.23	3.49%	
EQUIPAMIENTO SUCURSALES	NULL	NULL	5	0.13%	NULL	NULL	
FERRETERIA	234456.66	8.79%	640	21.11%	279.12	3.62%	
FLETES	757.63	0.03%	5	0.13%	151.53	1.98%	
HOGAR	172665.54	6.47%	370	9.30%	466.66	6.04%	
INSTALACIONES	75157.94	2.82%	350	9.05%	208.77	2.70%	
JARDIN	73519.71	2.76%	395	9.95%	185.66	2.40%	

ROL (Return of Labor)

- 5 - Evolucion de ventas por sectores por mes
- 10 - Comportamiento compras vs. ventas
- 11 - Informe de Ventas generales
- 11 - Informe de Ventas generales - Tickets
- 12 - Ranking Vtas. por articulo
- 13 - Ranking Vtas. por rubro y articulo
- 14 - Ranking Vtas. por proveedores
- 15 - Ranking Vtas. por rubro y proveedor
- 15 - Ranking Cantidad Vtas. por rubro y proveedor mes/año
- 16 - Informe Vtas. por rubro
- 17 - Informe Vtas. proveedor
- 18 - Informe Vtas. por art. Sucursales
- 19 - Informe de ventas por subrubro Sucursales
- 20 - Mercaderia sin mov. Vtas.
- 20 - Mercaderia sin mov.

Anio Fin: 2005

Sucursal: 03

Edit Parameters

Data

▼ Hide Data

ROL

11/1/05					
11/30/05					
HIPERTEHUELICHE - RIO GALLEGOS					
	Venta s/Imp.	Part. s/Total	Empleados	\$ x Empleado ROL	Part. \$ por Empleado
ABERTURAS	105744.48	3.96%	NULL	NULL	NULL
BAÑOS Y COCINAS	141450.95	5.30%	27	5238.92	7.85%
CERAMICA	189579.76	7.11%	26	7291.53	10.93%
CONSTRUCCION	1069752.06	40.10%	26	41144.31	61.66%
ELECTRICIDAD E ILUMINACION	62383.42	3.05%	26	3168.58	4.75%
EQUIPAMIENTO SUCURSALES	NULL	NULL	1	NULL	NULL
FERRERIA	234456.68	8.79%	26	9017.56	13.51%

Evolución GMROI - GMROS - GMROL

Evolución GMROI - GMROL - GMROS									
▼ Hide Parameters									
Dia Inicio: 1									
Mes Inicio: 5									
Año Inicio: 2003									
Dia Fin: 31									
Mes Fin: 5									
Año Fin: 2003									
Sucursal: 01									
Edit Parameters									
Data									
▼ Hide Data									
Evolución GMROI - GMROL - GMROS									
	Sectores	GMROI Actual	GMROL Actual	GMROS Actual	GMROI Año ant.	GMROL Año ant.	GMROS Año ant.	GMROI 2 Año ant.	GMROL 2
1	ABERTURAS	-0.01	-177.15	-15.70	0	0	0	0	
2	BAÑOS Y COCINAS	0.00	153.89	3.55	0	0	0	0	
3	CERAMICA	0.00	561.46	2.64	0	0	0	0	
4	CONSTRUCCION	0.04	5093.91	109.94	0	0	0	0	
5	ELECTRICIDAD E ILUMINACION	0.00	303.89	0.31	0	0	0	0	
6	FERRETERIA	0.00	142.77	0.62	0	0	0	0	
7	HOGAR	0.00	232.43	1.33	0	0	0	0	
8	INSTALACIONES	0.01	1658.18	8.42	0	0	0	0	
9	JARDIN	0.00	17.78	0.19	0	0	0	0	
10	MADERAS	-0.01	-629.71	-3.03	0	0	0	0	
11	MUEBLES	0.01	297.13	4.43	0	0	0	0	
12	REVESTIMIENTOS	0.00	331.76	1.51	0	0	0	0	

Ventas por mes

por sectores por mes

10 - Comportamiento
compras vs. ventas

11 - Informe de Ventas
generales

12 - Ranking Vtas. por
artículo

13 - Ranking Vtas. por
rubro y artículo

14 - Ranking Vtas. por
proveedores

15 - Ranking Vtas. por
rubro y proveedor

16 - Informe Vtas. por
rubro

17 - Informe Vtas.
proveedor

17 - Informe Vtas.
proveedor crosstab

18 - Informe Vtas. por
art.

18 - Informe Vtas. por
art. Sucursales

19 - Informe de ventas
por subrubro

19 - Informe de ventas
por subrubro Sucursales

20 - Mercadería sin mov.
Vtas.

21 - Margen bruto por
rubro y artículo

Ventas por mes

	BAHIA BLANCA		COSTA MATERIALES		HIPERTEHUELCHE - COMODORO		HIPERTEHUELCHE - RIO GALLEGOS		Total
	Monto	Contrib. %	Monto	Contrib. %	Monto	Contrib. %	Monto	Contrib. %	
02-MAY-2003	120,20	0,05%	58991,34	24,95%	108458,46	45,88%	68842,64	29,12%	236412,64
03-MAY-2003	0	0	26361,67	10,57%	127986,43	51,31%	95110,41	38,13%	249467,51
04-MAY-2003	0	0	0	0	104106,14	100,00%	0	0	104106,14
05-MAY-2003	0	0	64942,27	32,24%	73991,84	36,74%	62476,50	31,02%	201410,61
06-MAY-2003	0	0	44892,15	21,56%	92618,26	44,48%	70697,70	33,96%	208208,12
07-MAY-2003	0	0	37232,03	20,17%	80524,43	43,63%	66792,92	36,19%	184549,38
08-MAY-2003	0	0	53682,83	26,43%	77904,02	38,35%	71539,89	35,22%	203126,74
09-MAY-2003	0	0	35932,69	16,57%	111189,52	51,26%	69789,29	32,17%	216911,49
10-MAY-2003	0	0	25433,40	9,92%	143304,39	55,89%	87671,11	34,19%	256408,90
11-MAY-2003	0	0	0	0	100049,92	100,00%	0	0	100049,92
12-MAY-2003	0	0	35468,20	19,75%	81391,13	45,34%	62655,48	34,90%	179504,81
13-MAY-2003	0	0	21607,04	13,88%	76673,58	49,37%	57244,12	36,76%	155724,74
14-MAY-2003	0	0	41832,96	23,83%	70201,28	39,65%	65001,70	36,72%	177035,95
15-MAY-2003	0	0	26722,80	17,26%	70002,43	45,20%	58132,23	37,54%	154857,47
16-MAY-2003	0	0	34804,06	19,51%	80066,90	45,38%	62642,56	35,11%	178413,52
17-MAY-2003	0	0	21798,56	10,08%	110317,34	51,03%	84080,32	38,89%	216176,22
18-MAY-2003	0	0	0	0	83327,34	100,00%	0	0	83327,34
19-MAY-2003	0	0	25511,35	14,81%	77997,61	45,29%	68694,43	39,88%	172203,39
20-MAY-2003	0	0	34053,70	15,27%	132396,92	59,37%	56551,46	25,36%	223002,07
21-MAY-2003	0	0	32636,47	23,26%	66967,64	40,80%	50719,41	36,14%	140323,53
22-MAY-2003	0	0	20178,52	17,30%	47170,79	40,45%	49268,23	42,25%	116617,55

Ranking de ventas por artículo

generales

- 12 - Ranking Vtas. por artículo
- 13 - Ranking Vtas. por rubro y artículo
- 14 - Ranking Vtas. por proveedores
- 15 - Ranking Vtas. por rubro y proveedor
- 16 - Informe Vtas. por rubro
- 17 - Informe Vtas. proveedor
- 17 - Informe Vtas. proveedor crosstab
- 18 - Informe Vtas. por art.
- 18 - Informe Vtas. por art. Sucursales
- 19 - Informe de ventas por subrubro
- 19 - Informe de ventas por subrubro Sucursales
- 20 - Mercadería sin mov. Vtas.
- 21 - Margen bruto por rubro y artículo

[Editar Parámetros](#)

Datos

▼ [Ocultar Datos](#)

Ranking de ventas por artículo

	Producto		\$ Período	Part. s/Total
1	CEM01	CEMENTO COMODORO 2000 X 50 KG.	82358,51	8,98%
2	EM6003	CALEFACTOR EMEGE 9050 T.B.PATA6ONIA	37097,25	4,05%
3	TIG022	TUBO CLOAC.RAMAT TIGRE 3.2 110 X 4 M.	18802,48	2,05%
4	DST31	CHAPA NEGRA DD N.18 1,22X2,44 M.-1,27MM.	9559,75	1,04%
5	CAL12	CAL HIDRAT X 30 KG.	6583,61	0,72%
6	CA031	CA#0 REVESTIDO EPOXI 13 MM.(1/2")	4818,75	0,53%
7	EM6002	CALEFACTOR EMEGE 9030 T.B.PATA6ONIA	4470,00	0,49%
8	KLA910	ADHESIVO P/CERAMICO GRANIDUR X30KG	3718,74	0,41%
9	DH3065	CALEFACTOR EMEGE EURO 2105-KC 5400 T.N.	3255,46	0,36%
10	DH053	CALEFACTOR CTZ 9000 C/H T.B. LIN.PESADA	2703,75	0,29%
11	DH052	CALEFACTOR CTZ 9000 C/H T.N. LIN.PESADA	2255,50	0,25%
12	FOC22	TANQUE ROTOPLAS TRICAPA 1100 LT.	2152,55	0,23%
13	TEH706	ANTICONGELANTE FRIOLITE OC 20 KG.	1992,60	0,22%
14	HID800	HIDROMET 1206 BANSHEE LEVER LAV. PER/ORO	1834,92	0,20%
15	IMPO12	EQUIPO DE AUDIO SONY DX-90 (ORTEGA)	1779,94	0,19%
16	FOC21	TANQUE ROTOPLAS TRICAPA 600 LT.	1453,81	0,16%
17	RHEM11	TERMOTANQUE SAIAR 120LTS.MULTIGAS DE PIE	1393,98	0,15%
18	DH1758	TERMOTANQUE ACQUAPIU ESKABE R-800 52 LTS	1380,00	0,15%
19	LI0071	PUERTA LIBERMET 1651 080X200 INYECT.DER.	1347,46	0,15%
20	DH1307	CALEFACTOR EMEGE EURO 2130-SL 3000 T.B.	1196,70	0,13%
21	Total		180145,75	20,74%

Ranking de ventas por rubro y por proveedor

- generales
- 11 - Informe de Ventas generales - Tickets
- 12 - Ranking Vtas. por articulo
- 13 - Ranking Vtas. por rubro y articulo
- 14 - Ranking Vtas. por proveedores
- 15 - Ranking Vtas. por rubro y proveedor
- 16 - Ranking Cantidad Vtas por rubro y proveedor mes/año
- 16 - Informe Vtas. por rubro
- 17 - Informe Vtas. proveedor
- 18 - Informe Vtas. por art. Sucursales
- 19 - Informe de ventas por subrubro Sucursales
- 20 - Mercadería sin mov. Vtas.
- 20 - Mercadería sin mov. Vtas. s/ultima entrega
- 20 - Mercadería sin mov. Vtas. ultima entrega
- 21 - Margen bruto por rubro y articulo

Ranking de cantidad de ventas por rubro y proveedor mes/año

Page 1 of 2 25 rows per page

		NOV
		2005
		20
		Cantidad de Ventas
		CONSTRUCCION
		COSTA MATERIALES
E06	EL TEHUELCHES.A.C.I.O.I.	131742
	Part. s/Total	61.53%
A02	ACINDAR IND. ARG. DE ACEROS S.A.	35988
	Part. s/Total	16.81%
P25	PETROQ.COMODORO RIVADAVIA S.A.	11926
	Part. s/Total	5.57%
C350	CAUCIGH MADERAS DE JEAN CLAUDE CAUCIGH	5400
	Part. s/Total	2.53%

Ranking de ventas totalizado por rubro

generales 11- Informe de Ventas generales - Tickets 12 - Ranking Vtas. por articulo 13 - Ranking Vtas. por rubro y articulo 14 - Ranking Vtas. por proveedores 15 - Ranking Vtas. por rubro y proveedor 15 - Ranking Cantidad Vtas. por rubro y proveedor mes/año 16 - Informe Vtas. por rubro 17 - Informe Vtas. proveedor 18 - Informe Vtas. por art. Sucursales 19 - Informe de ventas por subrubro Sucursales 20 - Mercaderia sin mov. Vtas. 20 - Mercaderia sin mov. Vtas. s/ultima entrega 20 - Mercaderia sin mov. Vtas. ultima entrega 21 - Margen bruto por rubro y articulo	<p>Data</p> <p>▼ Hide Data</p> <p>Informe de ventas totalizado por rubro</p>																																												
	<table border="1"> <tr> <td></td> <td colspan="3">11/1/05</td> </tr> <tr> <td></td> <td colspan="3">11/30/05</td> </tr> <tr> <td></td> <td colspan="3">HIPERTEHUELCHE - RIO GALLEGOS</td> </tr> <tr> <td></td> <td>Venta Bruta</td> <td>Unidades</td> <td>Promedio (\$)</td> </tr> </table>		11/1/05				11/30/05				HIPERTEHUELCHE - RIO GALLEGOS				Venta Bruta	Unidades	Promedio (\$)																												
		11/1/05																																											
		11/30/05																																											
		HIPERTEHUELCHE - RIO GALLEGOS																																											
		Venta Bruta	Unidades	Promedio (\$)																																									
	<table border="1"> <tr> <td>▶ ABERTURAS</td> <td>148540.45</td> <td>858</td> <td>171.52</td> </tr> <tr> <td>▶ BAÑOS Y COCINAS</td> <td>210519.57</td> <td>3685</td> <td>57.13</td> </tr> <tr> <td>▶ CERAMICA</td> <td>281128.58</td> <td>19248</td> <td>14.61</td> </tr> <tr> <td>▶ CONSTRUCCION</td> <td>1474119.88</td> <td>413791</td> <td>3.58</td> </tr> <tr> <td>▶ ELECTRICIDAD E ILUMINACION</td> <td>173223.80</td> <td>43282</td> <td>4.00</td> </tr> <tr> <td>▶ FERRETERIA</td> <td>384723.70</td> <td>40827</td> <td>9.42</td> </tr> <tr> <td>▶ FLETES</td> <td>3163.84</td> <td>172</td> <td>18.39</td> </tr> <tr> <td>▶ HOGAR</td> <td>227472.91</td> <td>2259</td> <td>100.68</td> </tr> <tr> <td>▶ INSTALACIONES</td> <td>219881.35</td> <td>37140</td> <td>5.91</td> </tr> <tr> <td>▶ JARDIN</td> <td>104004.23</td> <td>8764</td> <td>11.87</td> </tr> <tr> <td>▶ MADERAS</td> <td>223115.01</td> <td>15453</td> <td>14.41</td> </tr> </table>	▶ ABERTURAS	148540.45	858	171.52	▶ BAÑOS Y COCINAS	210519.57	3685	57.13	▶ CERAMICA	281128.58	19248	14.61	▶ CONSTRUCCION	1474119.88	413791	3.58	▶ ELECTRICIDAD E ILUMINACION	173223.80	43282	4.00	▶ FERRETERIA	384723.70	40827	9.42	▶ FLETES	3163.84	172	18.39	▶ HOGAR	227472.91	2259	100.68	▶ INSTALACIONES	219881.35	37140	5.91	▶ JARDIN	104004.23	8764	11.87	▶ MADERAS	223115.01	15453	14.41
	▶ ABERTURAS	148540.45	858	171.52																																									
	▶ BAÑOS Y COCINAS	210519.57	3685	57.13																																									
	▶ CERAMICA	281128.58	19248	14.61																																									
▶ CONSTRUCCION	1474119.88	413791	3.58																																										
▶ ELECTRICIDAD E ILUMINACION	173223.80	43282	4.00																																										
▶ FERRETERIA	384723.70	40827	9.42																																										
▶ FLETES	3163.84	172	18.39																																										
▶ HOGAR	227472.91	2259	100.68																																										
▶ INSTALACIONES	219881.35	37140	5.91																																										
▶ JARDIN	104004.23	8764	11.87																																										
▶ MADERAS	223115.01	15453	14.41																																										

Margen bruto por rubro y por artículo

	Producto Codigo Name	\$ Venta	\$ Costo	Margen Bruto	% s/Total	Existencia	Exist. Val. Costo	Cant. Vendida	Cant. Vendida ant.	Importe Fact.	Importe
1	CEM01 - CEMENTO COMODORO 2000 X 50 KG.	8902,08	4461,19	41174,63	16,97%	29812	132698876,29	4292	0	38207727,36	0
2	TEH813 - LADRIBLOCK HIPER 18X18X33CM.	92,68	46,54	30078,24	12,40%	2384	110955,83	59790	0	5641337,20	0
3	TEH814 - LADRIBLOCK HIPER 12X18X33CM.	46,08	22,54	15181,32	6,26%	-4301	-98969,71	40144	0	1808691,52	0
4	SID047 - CHAPA CINCALACAN C-24 1,10 M X 12 PIES	159,56	79,79	11696,57	4,82%	-3364	-269437,50	318	0	60740,08	0
5	AC012 - HIERRO TORSIONADO 12 MM.	402,41	201,29	6020,28	2,48%	14363	2891100,09	551	0	221727,91	0
6	MEX03 - FIELTRO ESPUMA 3 MM. ISOLANT	341,37	170,73	5874,50	2,42%	117	19968,50	4162	0	1420747,80	0
7	AC104 - POSTE DE ACERO P/ALAMBRADO 1,80 M.	46,91	23,46	5365,00	2,21%	-11	-268,85	700	0	32837,00	0
8	ISD002 - MEMBRANA ISOLANT TBA 10 MM. ALUMINIZADA	173,73	86,88	4982,40	2,05%	2708	235312,43	1344	0	233493,12	0
9	OST31 - CHAPA NEGRA DD N.18 1,22X2,44 M. - 1,27MM.	1773,71	886,88	4779,85	1,97%	7700	8829082,43	109	0	193334,39	0
10	AC011 - HIERRO TORSIONADO 10 MM.	776,75	388,41	3961,02	1,63%	41076	15954329,16	646	0	501780,50	0
11	CAL12 - CAL HIDRAT X 30 KG.	1182,46	591,26	3291,51	1,36%	4218	2493904,60	791	0	936326,86	0
12	AC010 - HIERRO TORSIONADO 8 MM.	599,02	299,53	3078,46	1,27%	61154	18317337,81	707	0	423507,14	0
13	AC016 - MALLA SOLDADA P/HORMIGON 15 X 15	1025,83	512,93	2999,48	1,24%	673	345304,48	158	0	162081,14	0
14	AC013 - HIERRO TORSIONADO 16 MM.	223,80	111,90	2890,75	1,19%	-426	-47937,96	155	0	34689,00	0
15	TA0002 - MONTANTE 70 X 35 MM. X 2,8 M.L.	100,29	50,21	2646,51	1,09%	3180	158853,56	747	0	74916,63	0
16	DURL05 - PLACA DURLOCK 9,5MM STD. 120X240 121204	235,78	117,89	2640,42	1,09%	2024	238615,84	265	0	62481,70	0
17	TEH815 - LADRIBLOCK HIPER 08X18X33CM.	6,80	3,43	2525,20	1,04%	1086	3726,63	8310	0	56508,00	0
18	AC009 - HIERRO TORSIONADO 6 MM.	210,82	105,42	2479,11	1,02%	18869	1988114,73	1196	0	261929,90	0
19	AC037 - ALAMBRE GALVANIZADO RECOCIDO LISO N. 8	34,71	17,40	2289,23	0,94%	2568	44506,17	1251	0	43422,21	0
20	TEH335 - LADRIBLOCK HIPER LOSA 12X25X40CM.	11,44	5,73	1978,00	0,82%	588	3368,95	2802	0	32054,88	0
21	KL0910 - ADHESIVO P/CERAMICO GRANIDUR X30KG	1571,21	785,62	1859,33	0,77%	2822	2216909,65	285	0	447794,85	0
22	SAT13 - TUBOS ESTRUCTURALES DD 40 X 40 X 1,60	267,66	133,83	1687,80	0,66%	20	2641,80	89	0	23820,85	0
23	PALL01 - PALLET DE MADERA	225,00	112,50	1575,00	0,65%	-6490	-729000,00	158	0	35550,00	0
24	AC049 - ALAMBRE NEGRO RECOCIDO N 16	329,66	164,88	1664,31	0,64%	9529	1571060,76	1664	0	615703,62	0
25	AC043 - ALAMBRE OVALADO AVR 17/15 X 1000 M.	836,42	418,21	1523,86	0,63%	3064	1281412,17	15	0	12546,30	0

Capítulo 9

Conclusiones y Trabajo Futuro

En este capítulo se describen las conclusiones obtenidas como resultado de esta tesis, y los posibles trabajos futuros que se podrían realizar.

Conclusiones

El proceso de comercialización de un producto o servicio muestra que el establecimiento de una estrategia y un plan comercial, si bien es indispensable, no es una tarea sencilla. De acuerdo a la bibliografía consultada, muchas de las dificultades que afronta la persona encargada de esta tarea están relacionadas con la complejidad de obtener información durante las distintas etapas del planeamiento y durante la etapa de control del plan.

Las empresas dedicadas al negocio de retail, es decir, a la comercialización minorista masiva de muchos productos de diferente clase, afrontan todos los problemas del proceso de comercialización y en particular se enfrentan al problema de determinar que significa la productividad en este negocio. Para esto se plantea el modelo de recursos SRM (Strategic Resource Model), el cual brinda un enfoque que permite calcular la productividad del negocio basándose en los tres recursos principales que posee: mercadería, espacio físico y fuerza de trabajo.

La definición de los conceptos de datawarehousing propuestos en la bibliografía, permiten construir una solución tecnológica que brinde a los responsables de diseñar la estrategia y conducir el plan comercial de una empresa de retail una herramienta capaz de reunir los datos generados por la actividad de la empresa y aplicar el modelo SRM para facilitar y documentar la toma de decisiones realizadas.

La aplicación construida, además de ser una experiencia de desarrollo muy interesante, implementa el modelo SRM y brinda adicionalmente una gran cantidad de información acerca de la operatoria de la empresa. Se adoptó la arquitectura de tres capas para implementar esta solución con el objeto de desacoplar los datos de la aplicación que los accede y presenta al usuario final. De esta manera se logró focalizar el esfuerzo en la construcción del datawarehouse y luego elegir una aplicación de acceso y presentación que permita mostrar los resultados obtenidos.

En resumen, se construyó una solución que permite atacar los problemas de comercialización en empresas de retail, la cual brinda una estructura desacoplada y acceso a toda la información operativa de la empresa además de implementar un modelo que permite medir la productividad. Esta solución es el resultado del estudio y la comprensión de los problemas tecnológicos y de negocio que posee este dominio.

Si bien se presenta una solución terminada con una aplicación de ejemplo, existen varios aspectos que pueden mejorar y complementar la solución presentada como se detalla en la sección siguiente.

Trabajo futuro

A continuación analizaremos los posibles puntos a mejorar en el futuro. El orden de presentación de estos puntos es de temas conceptuales generales a puntos particulares y de implementación.

Uno de las principales tareas a futuro es la que consiste en probar esta aplicación modificando la arquitectura de aplicación propuesta. Como vimos en el capítulo 8 esta solución de datawarehouse se probó con una aplicación de ejemplo que utiliza Oracle Discoverer y el servidor de aplicaciones de Oracle. Sería muy interesante probar la solución con otros servidores de aplicaciones u otras aplicaciones completas.

Los beneficios de esta tarea son interesantes desde el punto de vista de definir una aplicación que respete la arquitectura de tres capas y haga especial énfasis en cumplir con los requerimientos de drilling y pivoting que tienen estas aplicaciones, además de mostrar los datos de una manera amigable para el usuario.

Otra de las tareas que permitirían extender la solución propuesta consiste en tomar el modelo de datawarehouse presentado y aplicar los conceptos de análisis predictivo y data mining.

Esta tarea es muy compleja e importante ya que permitiría extender el modelo para brindar una solución que a partir de la extrapolación de los datos de la actividad de los periodos pasados muestre distintos escenarios de la actividad futura y permita hacer análisis del tipo what if o tomar decisiones mas documentadas al momento de realizar la planificación estratégica.

Otro trabajo futuro que se podría realizar con la aplicación es incorporar la seguridad de acceso a los datos, de manera que no todos los usuarios puedan ver toda la información. Esta validación de acceso podría implementarse a nivel de reporte (ej: los reportes que muestran cifras de margen solo pueden ser vistos por gerentes) y a nivel de datos (ej: los datos correspondientes a las sucursales solo pueden ser vistos solo por el gerente de sucursal responsable y por los gerentes generales).

En conclusión, existen muchos enfoques y áreas bajo los cuales se pueden realizar futuras extensiones a este trabajo planteado.

Apéndice A

Descripción de reportes

En este apéndice se brinda una descripción detallada de cada uno de los reportes desarrollados, con la finalidad de explotar el datawarehouse y mostrar la funcionalidad de la aplicación ejemplo.

Consideraciones generales

Aquí se provee la definición de algunos campos que aparecen en varios reportes. Salvo aclaración explícita, se debe interpretar que la definición válida para el campo es la que se da a continuación:

- **Costos:** El costo del total de productos vendidos. El valor esta acotado por el período determinado por las fechas ingresadas como parámetro, la sucursal dada, la categoría de productos mostrada o el proveedor ingresado. El costo de cada producto esta compuesto por el precio de costo y los descuentos comerciales aplicados. También se aplican los impuestos de IVA e IIBB.
- **Venta bruta:** (Monto * Cantidad) El monto de venta sin descontar impuestos para el total de productos vendidos. El valor esta acotado por el período determinado por las fechas ingresadas como parámetro, la sucursal dada, la categoría de productos mostrada o el proveedor ingresado. El precio de venta bruta de cada producto incluye los descuentos realizados al cliente.
- **Venta s/imp:** (Monto – IVA – IIBB) * Cantidad
- **Margen Neto:** Venta s/imp – Costos – Flete
- **Margen %:** Margen Neto / Venta s/imp
- **Cantidad:** La suma de las cantidades vendidas para los productos del cubo VENTAS_FACT
- **Stock:** Cantidad de productos en existencia, valorizada al costo, al final del período determinado por las fechas ingresadas como parámetro, para cada producto perteneciente a la sucursal ingresada. Se toma el último valor conocido de existencia para el sector/rubro, anterior a la fecha final de facturación dada.
- **M2:** Superficie ocupada por el sector o rubro. Se toma el último valor conocido de m2 ocupados por el sector/rubro, anterior a la fecha final de facturación dada.

- Empleados: Cantidad de empleados afectados al sector. Se toma el último valor conocido de empleados afectados al sector, anterior a la fecha final de facturación dada.
- Sectores: Campo DESCRIPCION (C1) de la dimensión PRODUCTO_DIM
- Costos: COSTO * CANTIDAD del cubo VENTAS_FACT
- ROI: Venta S/Imp / Stock
- ROS: Venta s/imp / M2
- ROL: Venta s/imp / Empleados
- GMROI: Margen Neto / Stock
- GMROL: Margen Neto / Empleados
- GMROS: Margen Neto / M2

Reportes principales

- **Margen Neto / ROI**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Código Sucursal

Descripción de los datos:

1. Costos
2. Venta bruta
3. Venta s/imp
4. $\text{Part. s/total} = \text{Venta s/imp} / \text{SUM}(\text{Venta s/imp})$
5. Margen Neto
6. Margen %
7. $\text{Cont. Marginal} = \text{Margen Neto} / \text{SUM}(\text{Margen Neto})$
8. Stock
9. ROI

Drill:

Este reporte no tiene opciones de drill porque los datos están a nivel de Sector, tampoco es posible hacer drill por fechas dado que el rango de fechas puede ser arbitrario.

- **ROS**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Código Sucursal

Descripción de los datos:

1. Venta bruta
2. Venta s/imp
3. Contrib. % vtas. = $\text{Venta s/imp} / \text{SUM}(\text{Venta s/imp})$
4. M2
5. % M2 local = $\text{M2} / \text{SUM}(\text{M2})$
6. Ventas x M2 (ROS)
7. Part. Vtas. X M2 = $\text{ROS} / \text{SUM}(\text{ROS})$

Drill:

Este reporte no tiene opciones de drill porque los datos están a nivel de Sector, tampoco es posible hacer drill por fechas dado que el rango de fechas puede ser arbitrario.

• ROL**Parámetros:**

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Código Sucursal

Descripción de los datos:

1. Venta s/imp
2. Part. s/total = $\text{Venta s/imp} / \text{SUM}(\text{Venta s/imp})$
3. Empleados
4. ROL = $\text{Venta s/imp} / \text{Empleados}$
5. Part. \$ x Empleado = $\text{ROL} / \text{SUM}(\text{ROL})$

Drill:

Este reporte no tiene opciones de drill porque los datos están a nivel de Sector, tampoco es posible hacer drill por fechas dado que el rango de fechas puede ser arbitrario.

- **Porcentuales**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Código Sucursal

Descripción de los datos:

1. $M2 \% = M2 / \text{SUM}(M2)$
2. $\text{Venta s/imp \%} = \text{Venta s/imp} / \text{SUM}(\text{Venta s/imp})$
3. $\text{Margen Neto \%} = \text{Margen Neto} / \text{SUM}(\text{Margen Neto})$
4. $\text{Stock \%} = \text{Stock} / \text{SUM}(\text{Stock})$
5. $\text{Empleados \%} = \text{Empleados} / \text{SUM}(\text{Empleados})$

Drill:

Este reporte no tiene opciones de drill porque los datos están a nivel de Sector, tampoco es posible hacer drill por fechas dado que el rango de fechas puede ser arbitrario.

- **GMROI – GMROL – GMROS**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Código Sucursal

Descripción de los datos:

1. Margen Neto
2. Margen %
3. $\text{Contrib. \%} = \text{Margen Neto} / \text{SUM}(\text{Margen Neto})$
4. M2
5. Empleados
6. Stock
7. GMROI
8. GMROL
9. GMROS

Drill:

Este reporte no tiene opciones de drill porque los datos están a nivel de Sector, tampoco es posible hacer drill por fechas dado que el rango de fechas puede ser arbitrario.

- **Evolución GMROI – GMROL – GMROS**

Parámetros:

1. Mes de facturación
2. Año de facturación
3. Código Sucursal

Descripción de los datos:

1. GMROI Actual
2. GMROL Actual
3. GMROS Actual
4. GMROI 1 año
5. GMROI Actual - 1 año =
$$(\text{GMROI Actual} - \text{GMROI 1 año}) * 100 / \text{GMROI 1 año}$$
6. GMROL 1 año
7. GMROL Actual - 1 año =
$$(\text{GMROL Actual} - \text{GMROL 1 año}) * 100 / \text{GMROL 1 año}$$
8. GMROS 1 año
9. GMROS Actual - 1 año =
$$(\text{GMROS Actual} - \text{GMROS 1 año}) * 100 / \text{GMROS 1 año}$$
10. GMROI 2 años
11. GMROI Actual - 2 años =
$$(\text{GMROI Actual} - \text{GMROI 2 años}) * 100 / \text{GMROI 2 años}$$
12. GMROL 2 años
13. GMROL Actual - 2 años =
$$(\text{GMROL Actual} - \text{GMROL 2 años}) * 100 / \text{GMROL 2 años}$$
14. GMROS 2 años
15. GMROS Actual - 2 años =
$$(\text{GMROS Actual} - \text{GMROS 2 años}) * 100 / \text{GMROS 2 años}$$

Drill:

Este reporte no tiene opciones de drill porque los datos están a nivel de Sector, tampoco es posible hacer drill por fechas dado que el rango de fechas puede ser arbitrario.

Reportes complementarios

- **Ventas por mes**

Parámetros:

1. Mes de facturación
2. Año de facturación

Descripción de los datos:

1. Sucursal = Venta Bruta
2. Contrib. % = Sucursal / SUM(Sucursal)

Drill:

Este reporte tiene opciones de drill down/up en los campos fecha (año, mes); sucursal (ciudad, sucursal)

- **Evolución de ventas por mes**

Parámetros:

1. Mes de facturación
2. Año de facturación

Descripción de los datos:

1. Sucursales
2. Ventas = Venta Bruta
3. Ventas Anterior = Ídem Ventas pero del año anterior al ingresado
4. Evol. Actual - Anterior = Ventas – Ventas Anterior
5. Evol. Actual - Anterior % = $(Ventas - Ventas Anterior) * 100 / Ventas Anterior$

Drill:

Este reporte no tiene opciones de drill porque los datos están a nivel de Sector, tampoco es posible hacer drill por fechas por las características de comparar datos de diferentes periodos de tiempo.

- **Evolución de ventas por sectores por mes**

Parámetros:

1. Mes de facturación
2. Año de facturación
3. Código Sucursal

Descripción de los datos:

1. \$ = Venta Bruta
2. Unidades = Cantidad del sector en el período ingresado como parámetro
3. Precio promedio = \$ / Unidades
4. Evolución Actual – 1 año = \$ Mes Actual - \$ Mes 1 año
5. Evolución Unidades Actual - 1 año = Unidades Mes Actual – Unidades Mes 1 año
6. Evolución Actual – 2 años = \$ Mes Actual - \$ Mes 2 años
7. Evolución Unidades Actual - 2 años = Unidades Mes Actual – Unidades Mes 2 años

Drill:

Este reporte no tiene drill

- **Comportamiento de las compras vs. ventas**

Parámetros:

1. Año de facturación
2. Mes de facturación

Descripción de los datos:

1. Costos = La suma del valor de costo de la mercadería adquirida para el mes y sucursal correspondientes, obtenido de los remitos de compra
2. Ventas = Venta Bruta.
3. % = Costos / Ventas

Drill:

Este reporte tiene opciones de drill down/up en los fecha (año, mes, día); sucursal (ciudad, sucursal)

• Informe de ventas general

Parámetros:

1. Sucursal
2. Fecha actual de facturación - Inicio
3. Período anterior de facturación: DIARIO/SEMANTAL/MENSUAL. Este parámetro determina el valor de los parámetros del Período actual de facturación – Fin y de los períodos anteriores de facturación.

Descripción de los datos:

1. Rubro/Subrubro = Descripción completa del subrubro.
2. \$ Período = Venta Bruta
3. Variación % Período Ant. =
Si \$ Período = 0:
Si \$ Período Anterior = 0 => Variación % Período Ant. = 0
Si \$ Período Anterior < 0 => Variación % Período Ant = 100
Si \$ Período < 0:
Si \$ Período Anterior = 0 => Variación % Período Ant = 100
Si \$ Período Anterior < 0, Variación % Período Ant = $((\$ \text{ Período} - \$ \text{ Período Anterior}) / \$ \text{ Período Anterior}) * 100$
4. Variación % Año Ant. =
Si \$ Período = 0:
Si \$ Año Anterior = 0 => Variación % Año Ant. = 0
Si \$ Año Anterior < 0 => Variación % Año Ant = 100
Si \$ Período < 0:
Si \$ Año Anterior = 0 => Variación % Año Ant = 100
Si \$ Año Anterior < 0, Variación % Año Ant = $((\$ \text{ Período} - \$ \text{ Año Anterior}) / \$ \text{ Año Anterior}) * 100$
5. Part. S/Total de Vta. =
Si SUM (\$ Período) = 0 => Part. S/Total de Vta. = 0.
Si SUM (\$ Período) < 0 => Part. S/Total de Vta. = $(\$ \text{ Período} / \text{SUM} (\$ \text{ Período})) * 100$
6. Cantidad de tickets o facturas emitidos: Cantidad de comprobantes emitidos para el período actual, para el período anterior y para el año anterior respectivamente

Drill:

Este reporte no tiene drill dado que el rango de fechas puede ser arbitrario

- **Ranking de ventas por artículo**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Cantidad de elementos a mostrar (Top N)

Descripción de los datos:

1. Artículo = Código y nombre del artículo
2. Descripción = Descripción del artículo
3. \$ Período = Venta Bruta
4. Part. s/total = $\$ \text{Período} / \text{SUM}(\$ \text{Período})$. El % mostrado es respecto del total, no de los artículos mostrados en el reporte

Drill:

Este reporte no tiene drill debido a la función necesaria para el cálculo del %

- **Ranking de ventas por rubro y por artículo**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Rubro de productos (código C1)
5. Cantidad de elementos a mostrar (Top N)

Descripción de los datos:

1. Artículo = Código y nombre de artículo
2. Descripción = Descripción de artículo
3. \$ Período = Venta Bruta
4. Part. s/total = $\$ \text{Período} / \text{SUM}(\$ \text{Período})$ El % mostrado es respecto del total, no de los artículos mostrados en el reporte

Drill:

Este reporte no tiene drill debido a la función necesaria para el cálculo del %

• **Ranking de ventas por proveedor**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Cantidad de elementos a mostrar (Top N)

Descripción de los datos:

1. Proveedor código = Código de proveedor
2. Proveedor nombre = Nombre de proveedor
3. \$ Período = Venta Bruta
4. $\text{Part. s/total} = \$ \text{Período} / \text{SUM}(\$ \text{Período})$ El % mostrado es respecto del total, no de los artículos mostrados en el reporte

Drill:

Este reporte no tiene drill debido a la función necesaria para el cálculo del %

• **Ranking de ventas por rubro y por proveedor**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Rubro
5. Cantidad de elementos a mostrar (Top N)

Descripción de los datos:

1. Código Proveedor = Código de proveedor
2. Nombre Proveedor = Nombre del proveedor
3. Cantidad = correspondientes al producto, para el proveedor, sucursal y periodo ingresado
4. $\text{Part. s/total} = \text{Cantidad} / \text{SUM}(\text{Cantidad})$ El % mostrado es respecto del total, no de los artículos mostrados en el reporte

Drill:

Este reporte no tiene drill debido a la función necesaria para el cálculo del %

- **Informe de ventas totalizado por rubro**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal

Descripción de los datos:

1. Rubro = Rubro de productos.
2. Totalizado = Venta Bruta
3. Unidades = Cantidad correspondientes al rubro
4. Promedio (\$) = Totalizado / Unidades

Drill:

Este reporte tiene opciones de drill down/up en los campos sector (sector, rubro, subrubro); no tiene drill por fecha ya que el rango puede ser arbitrario; sucursal (ciudad, sucursal)

- **Informe de ventas por proveedor (ranking de productos)**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Proveedor

Descripción de los datos:

1. Código Rubro = Código del rubro
2. Descripción Rubro = Descripción del rubro
3. Código Artículo = Código del artículo
4. Nombre Artículo = Nombre del artículo
5. Cantidad = Cantidad para el proveedor, sucursal y período ingresado
6. Precio = Venta Bruta

Drill:

Este reporte tiene opciones de drill down/up en los campos sector (sector, rubro, subrubro); no tiene drill por fecha dado que el rango puede ser arbitrario; sucursal (ciudad, sucursal)

- **Informe de ventas por artículo**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Producto

Descripción de los datos:

1. Fecha = Fecha de venta
2. Cantidad = Cantidad para la sucursal y período ingresado
3. Precio = Venta Bruta

Drill:

Este reporte tiene opciones de drill down/up en los campos fecha (año, mes, día); sector (sector, rubro, subrubro); sucursal (ciudad, sucursal)

- **Informe de ventas por subrubro**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Rubro

Descripción de los datos:

1. Rubro/Subrubro = Descripción completa del subrubro
2. Cantidad = Cantidad correspondientes al subrubro, para la sucursal, rubro y período ingresado
3. Part. s/total = Cantidad / SUM(Cantidad)

Drill:

Este reporte tiene opciones de drill down/up en los campos sector (sector, rubro, subrubro); fecha (año, mes, día); sucursal (ciudad, sucursal)

- **Mercadería sin movimiento de ventas**

Parámetros:

1. Fecha inicial de facturación
2. Fecha final de facturación
3. Sucursal
4. Rubro

Descripción de los datos:

1. Código = Código de producto. Se muestran aquellos productos para los cuales no existe registro de ventas durante el período ingresado, para la sucursal ingresada
2. Descripción = Descripción del producto
3. Stock
4. Precio de venta = Precio de venta actual del producto sin impuestos ni descuentos
5. Precio de costo = Precio de costo actual del producto
6. Última fecha de ingreso = La última fecha de compra registrada para el producto correspondiente y la sucursal dada. Se toma la última fecha de ingreso a partir de la fecha final de facturación
7. Código de rubro = Descripción completa del subrubro

Drill:

Este reporte no tiene drill porque los datos son retornados a nivel de producto

- **Margen bruto por rubro y artículo**

Parámetros:

1. Período Actual inicial
2. Período Actual final
3. Período Anterior inicial
4. Período Anterior final
5. Sucursal
6. Rubro

Descripción de los datos:

1. Artículo = Código y descripción del artículo
2. Margen Bruto
3. $\% \text{ s/total} = \text{Margen Bruto} / \text{SUM}(\text{Margen Bruto})$
4. Margen Neto = Margen descontando impuestos correspondientes al producto, para el periodo, rubro y sucursal ingresados. El margen bruto de cada producto incluye los descuentos realizados al cliente
5. $\% \text{ s/total} = \text{Margen Neto} / \text{SUM}(\text{Margen Neto})$
6. Precio de venta = Precio de venta actual del producto sin impuestos ni descuentos
7. Precio de costo = Precio de costo actual
8. Existencia = Cantidad disponible del producto correspondiente en todos los depósitos de la sucursal ingresada. Se muestra la última cantidad registrada a partir de la fecha de fin de facturación ingresada
9. Existencia val. Costo = Cantidad disponible del producto correspondiente valorizada al ultimo precio de costo registrado hasta el final del periodo actual ingresado
10. Cantidad vendida = Cantidad para el rubro, sucursal y periodo ingresados
11. Cantidad vendida ant = Ídem Cantidad Vendida para el periodo anterior
12. Importe fac = La suma de los montos de venta sin descontar impuestos correspondientes al producto, para el rubro, periodo y sucursal ingresados. El precio de venta bruta de cada producto incluye los descuentos realizados al cliente.
13. Importe fac. ant = Ídem Importe fac para el periodo anterior.

Drill:

Este reporte no tiene drill porque los datos son retornados a nivel de producto.

• **Ranking de margen bruto unitario**

Parámetros:

1. Fecha de facturación
2. Sucursal
3. Rubro
4. Ítems:
 - a. Ítems con existencia positiva (si/no)
 - b. Ítems con un % margen mayor a un valor dado
 - c. Top N ítems.

Descripción de los datos:

1. Artículo = Código y descripción del artículo
2. Costo = Costo unitario sin IVA ni flete
3. Venta = Venta bruta
4. % Margen = Margen porcentual sin descontar impuestos correspondientes al producto, para el período, rubro y sucursal ingresados.

Drill:

Este reporte tiene opciones de drill down/up en los campos sector (rubro, subrubro); fecha (año, mes, día); sucursal (ciudad, sucursal)†

Referencias

- [Braidot96] Marketing total
1996
Néstor P. Braidot
- [McCarthy-Perreault94] Fundamentos de comercialización, Principios y
métodos.
1994
E. Jerome McCarthy, William D. Perreault Jr.
- [Sallenave93] Gerencia y Planeación Estratégica
1993
Jean Paul Sallenave
- [Hernant04] Resource Management and Profitability in Retailing
2004
Mikael Hernant
- [Kimball1] The Data Warehouse Toolkit
2002
Ralph Kimball – Margy Ross
- [Kimball2] The Data Warehouse Lifecycle Toolkit
1998
Ralph Kimball, Laura Reeves, Margy Ross, Warren
Thornthwaite
- [Alison] Oracle9i Warehouse Builder User's Guide
2003
Shirinne Alison, Kavita Nayar, Michelle Bird, Julia M.
Stein
- [Scalzo] Oracle DBA Guide to Datawarehousing and Star
Schemas
2003
Bert Scalzo
- [9iAS1] Oracle 9i Application Server Concepts
2002
Oracle
- [9iAS2] Oracle 9i Application Server Administrator's Guide
2002
Oracle

- [Russell] Oracle 9i Application Developer's Guide
2002
John Russell
- [Discoverer1] Oracle 9iAS Discoverer Configuration Guide
2002
Oracle
- [Discoverer2] Oracle 9iAS Discoverer User's Guide
2002
Oracle
- [Bedell1] "Data Warehousing, Data Modeling and Design"
1997
Jeff Bedell
DSSTraining, MicroStrategy.
- [Inmon1] Building the Data Warehouse
1992
William H. Inmon
Wiley-QED John Wiley & Sons
- [Power1] A Brief History of Decision Support Systems
May 31, 2003
D. J. Power.
- [Morton1] Decision Support System: An organizational
perspective
1978
Michael Scott Morton & Peter Keen
- [Pressman1] Ingeniería del Software
1993
Roger Pressman
Mc Graw Hill
- [Marley1] Information Systems Evolution
November 2000
Stephen R. Marley
World Wide Web
<http://geoenabler.geoall.net/SMarley-DIP-SRM>
- [wikipedia] World Wide Web
<http://www.wikipedia.org/>
- [Turban1] Decision Support Systems And Intelligent Systems
2001
Turban, Efraim y Aronson Jay E.
Prentice Hall, New Jersey.

- [Ambler99] Mapping Objects to Relational Databases
1999
Scott W. Ambler
- [Renzel97] Klaus Renzel, Wolfgang Keller. Client/Server
Architectures for Business Information Systems - A
Pattern Language. 1997.
- [Seacord98] Robert Seacord, Scott Hissam. Browsers for Distributed
Systems: Universal Paradigm or Siren's Song? 1998.

Referencias utilizadas para la construcción de la aplicación ejemplo

- [Baylis] Oracle 9i Administrator Guide
2002
Ruth Baylis
- [Cyrán] Oracle 9i Database Concepts
2002
Michele Cyran
- [Lane] Oracle 9i Data Warehousing Guide
2002
Paul Lane
- [Stein] Oracle9i Warehouse Builder Installation and
Configuration Guide
2003
Julia M. Stein