

Tema 7

Business Intelligence

SW Corporativo

Curso 2018-19

Objetivos del Tema 7

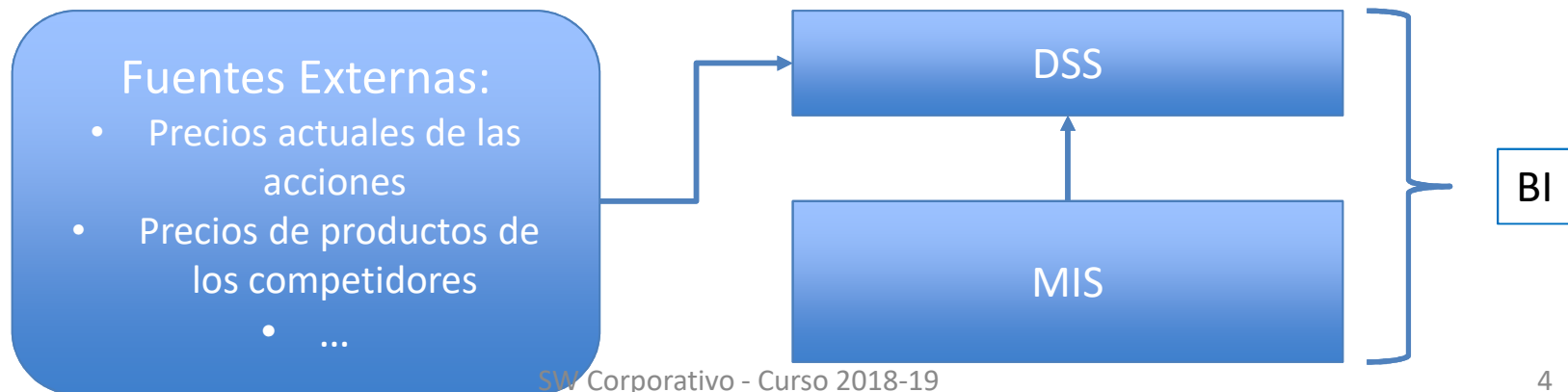
- Conocer los conceptos básicos en torno Business Intelligence (BI) y Business Analytics (BA)
- Entender de importancia de BI dentro de los procesos estratégicos de la Empresa.
- Relación con buena parte de los procesos operativos y de soporte de la empresa e importancia respecto a los objetivos estratégicos.

Recordemos...Objetivos Estratégicos

- Excelencia operacional;
- Nuevos productos, servicios y modelos de negocios;
- Cercanía con clientes y proveedores;
- Toma de decisiones mejorada; ventaja competitiva, y
- Supervivencia continua (competencia).

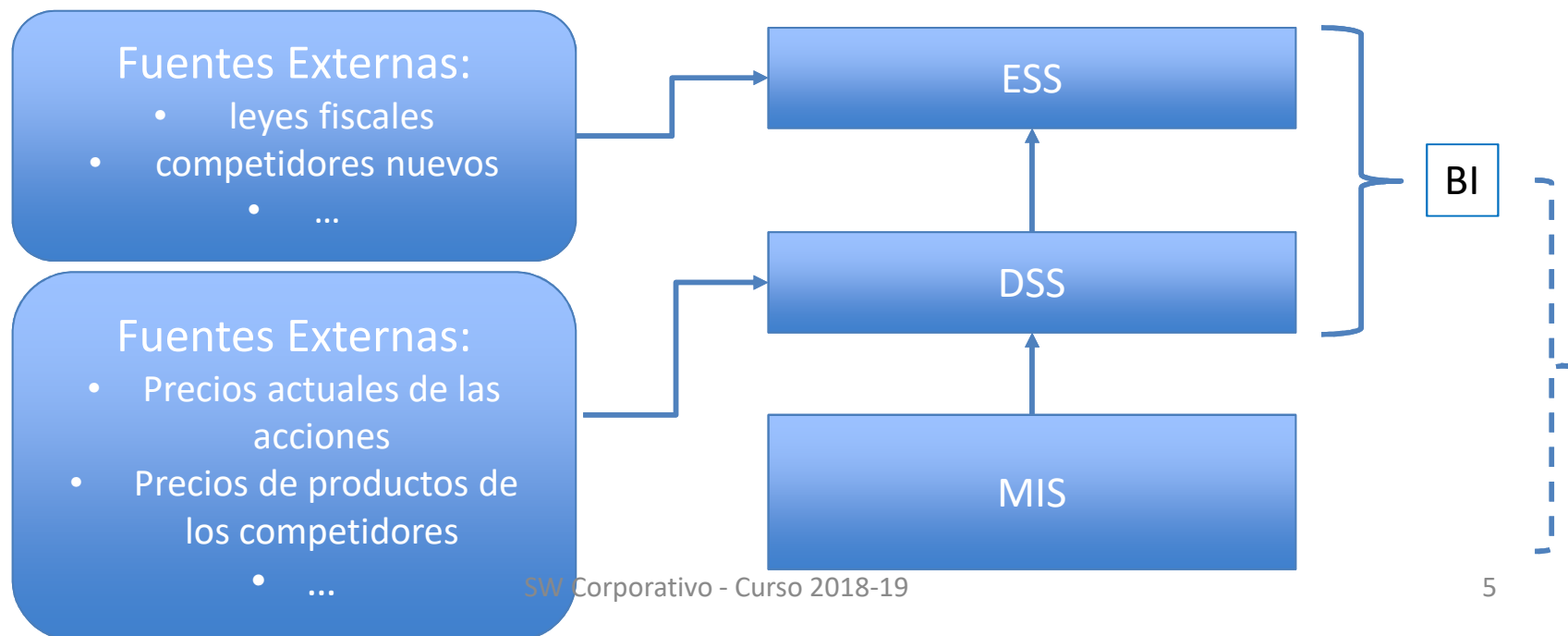
Recordemos.....Sistemas de Soporte a la Decisión (DSS)

- Se puede considerar como un MIS especializado en soporte a las decisiones.
- Dan apoyo a la **toma de decisiones que no es rutinaria**.
- Se enfocan en **problemas que son únicos y cambian con rapidez**, para los cuales el proceso para llegar a una solución tal vez no esté por completo predefinido de antemano.
 - ¿Cuál sería el impacto en los itinerarios de producción si se duplicaran las ventas en el mes de diciembre? ¿Qué ocurriría con nuestro rendimiento sobre la inversión si se retrasara el itinerario de una fábrica por seis meses?



Recordemos...Executive Support System (ESS)

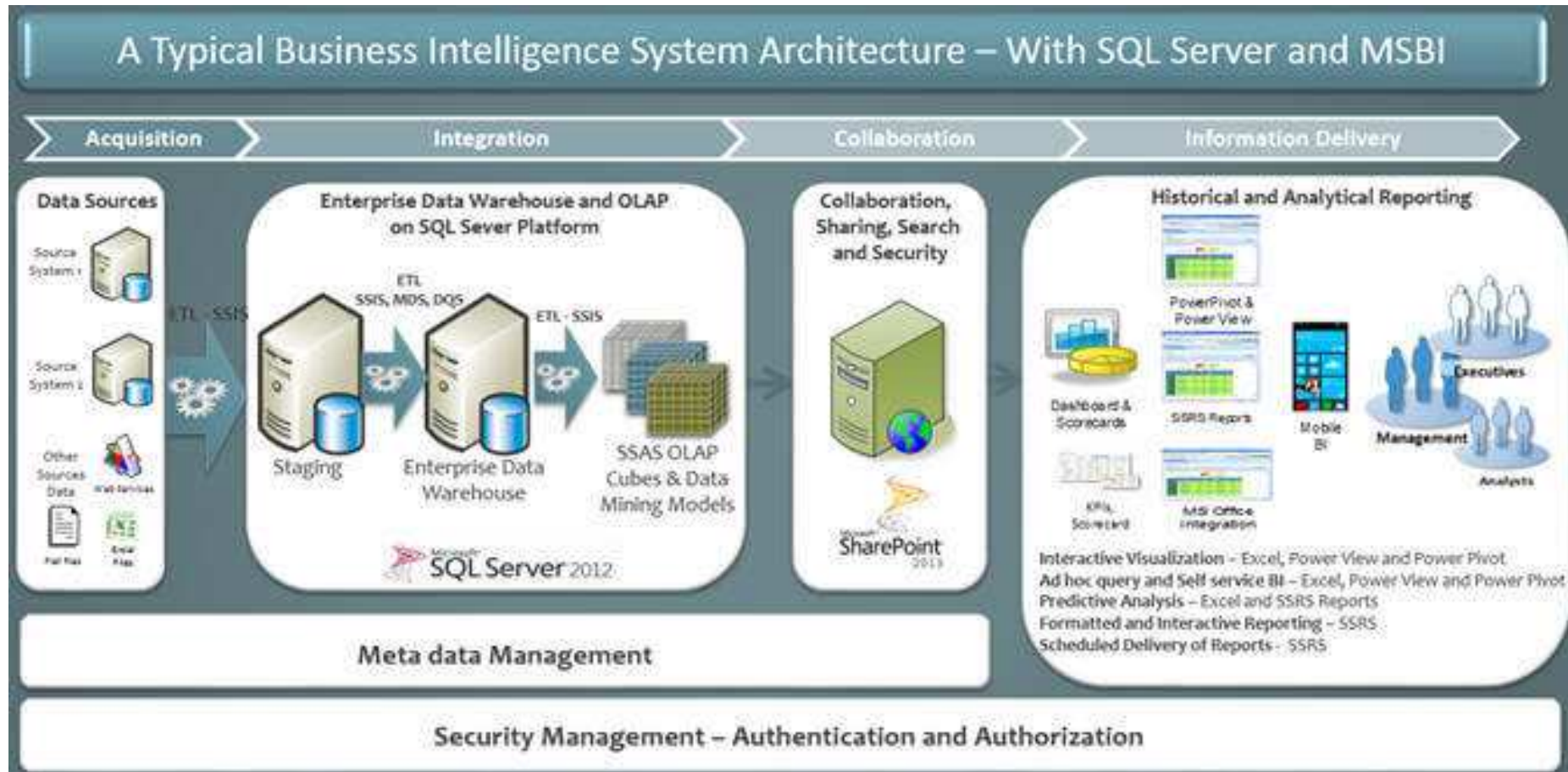
- Los sistemas de apoyo a ejecutivos (ESS) ayudan a los directivos de nivel superior a tomar estas resoluciones.
- Filtran, comprimen y rastrean datos críticos, para mostrar la información de mayor importancia a estos directivos.
- incluyen cada vez en mayor grado los análisis de inteligencia de negocios para analizar tendencias, realizar pronósticos y “desglosar” los datos para obtener mayores niveles de detalle.



Business Intelligence

- ¿Cuáles serán los niveles de empleo en cinco años?
- ¿Cuáles son las tendencias de costos de la industria a largo plazo, y en qué posición se encuentra nuestra empresa?
- ¿Qué productos debemos estar fabricando en cinco años?
- ¿Qué nuevas adquisiciones nos protegerían de las oscilaciones cíclicas de los negocios?

BI



¿Qué más es la inteligencia de negocios?

- Inteligencia de negocios es un **proceso interactivo** para **explotar y analizar información estructurada** sobre un área*, normalmente almacenada en un data warehouse, para descubrir **tendencias o patrones**, a partir de los cuales derivar ideas y extraer conclusiones
- El proceso de inteligencia de negocios incluye la **comunicación de los descubrimientos** y la **realización de los cambios**.
- (*) Las áreas incluyen clientes, proveedores, productos, servicios y competidores

Tendencias...

- En la actualidad, existen tres tendencias que predominan en BI:
 - **BI 2.0**: la inteligencia de negocios mediante el uso de las tecnologías de la web 2.0
 - **BI móvil**: el uso de dispositivos móviles, teléfonos inteligentes y tablets.
 - Adaptado a los tamaños de las pantallas de los dispositivos y a sus características técnicas.
 - **BI social**: el uso e impacto de los medios sociales en las operaciones de negocio.
- Hoy en día se ofrecen los componentes de BI como **servicios en la nube**, de forma que las empresas en lugar de comprar paquetes de software a un proveedor e instalarlos, contratan a un proveedor de BI con servicios en la nube que les permitirá el **acceso a la información en tiempo real**, incluso desde dispositivos móviles en **cualquier lugar y momento**.

BI en Tiempo Real

- Hasta hace poco, BI se centraba en tratamiento de datos históricos.
- El enfoque ha cambiado con la emergencia de tecnología para captura, almacenamiento y empleo de datos en tiempo real.
- Los BI de tiempo real facilitan a los usuarios
 - usar análisis multidimensional,
 - minería de datos, y
 - tomar decisiones para interactuar con clientes con nuevas formas y formatos en tiempo real.

Decisiones

- No obstante, en todo proceso de toma de decisiones, por muy buena información que se posea, existe un ambiente de incertidumbre y que, generalmente, se tiene que asumir un cierto riesgo.
- La toma de decisiones, aunque se basa en informaciones lo más fiables y completas posibles, siempre se realiza asumiendo un **riesgo**.

DECISIÓN = INFORMACIÓN + RIESGO

- Completar la información lo máximo posible, haciendo que sea lo más exacta y completa posible, es un factor de minimización del riesgo.

De los Datos a las Decisiones...

- Los datos se almacenan en un Data warehouse o en un Data mart.
- Pueden provenir de: ERP, CRM, SCM, Sistemas heredados antiguos y/o de la web.
- La comunidad de usuarios analiza los datos utilizando una variedad de aplicaciones (DM, OLAP, DSS...)
- Los resultados de estos análisis se pueden presentar al usuario con otras aplicaciones (dashboards, balanced scorecards, reporting, query y visualización).
- Por ultimo, los directivos usan los resultados en la toma de sus decisiones.

BI vs Business Analytics (Analítica de Negocios)

- Podemos dividir los sistemas de BI en:
 - BI propiamente dicho y
 - Business Analytics
- **BI** está ligado a la captura de los datos correctos por muy complejos que sean para la generar información más relevante para apoyar el proceso de toma de decisiones.
- **La analítica de negocios** se refiere normalmente al uso extensivo de los datos, análisis estáticos, modelos explicativos y productivos para explicar comportamientos y si es posible anticiparse a situaciones futuras.

BI vs Business Analytics

	BI	Analítica de Negocios
Responde a las preguntas	¿Qué sucede? ¿Cuándo? ¿Quién? ¿Cuántos?	¿Por qué sucedió? ¿Ocurrió otra vez? ¿Qué pasaría si cambiáramos por...?
Infraestructura	Reporting o Interfaces de Usuario Sistemas de monitoreo/ alertas Cuadros de mando (dashboard) Balanced scorecards Consultas ad hoc/query OLAP (Cubos)	Análisis estadístico-cuantitativo. Datamining Modelado productivo Análisis multivariables

Caso de éxito: Netflix

- Según las ganancias trimestrales de Netflix, la compañía tiene **actualmente más de 53 millones de clientes** de streaming en todo el mundo.
- Netflix, se ha convertido en uno de los casos de éxito de Business Intelligence, mediante análisis de grandes datos, siendo capaz de rastrear una cantidad interminable de información.
- Aquí se incluye:
 - cuando pasamos,
 - rebobinamos o avanzamos rápido,
 - qué día y la hora hemos visto el contenido,
 - desde donde lo vemos (basado en código postal),
 - qué dispositivo usamos para ver,
 - las calificaciones que da,
 - las cosas que busca y su comportamiento de navegación y desplazamiento.

Caso de éxito: Netflix (cont.)

- Todos estos datos le dan a Netflix una **ventaja incomparable frente a las redes de televisión tradicionales** que tienen que tomar decisiones de radiodifusión basadas en la tradición y la intuición.
- Netflix es capaz de **obtener un conocimiento exacto de sus clientes antes de tomar decisiones.**
- Con esto ha demostrado ser un excelente ejemplo de cómo los datos y la analítica pueden dar una visión vital tanto para administrar un negocio mejor como para ofrecer un producto superior.

Caso de éxito: UPS (empresa de envíos)

- Cuando un conductor va desde un punto A a un punto B puede hacerlo por numerosas rutas.
- ¿Qué es lo que determina el mejor camino a tomar? ¿Es el que tiene menos tráfico? ¿La menor cantidad de semáforos? ¿El límite de velocidad más alto?. ¿La menor distancia física para viajar?
- Para muchos, la ruta más eficiente no es una preocupación importante, pero para compañías como UPS, minimizar el tiempo de conducción es vital.
- Por esta razón, los ingenieros de UPS, segundo de los casos de éxito de Business Intelligence, trataron de diseñar un plan para la optimización de rutas. Utilizaron un sistema llamado **Orion**, abreviatura de **On-Road Integrated Optimization and Navigation**. Este algoritmo utilizó 1.000 páginas de código para analizar 200.000 posibilidades para cada ruta en tiempo real.
- Mediante el uso de estos grandes datos, fueron capaces de entender cuanto podían mejorar sus entregas circulando por diferentes rutas.

Caso de éxito: UPS (cont.)

- Como resultado, entre 2004 y 2012, UPS ahorró 10 millones de bombonas de gas.
- Las emisiones de carbono se redujeron en 100.000 toneladas métricas.
- También ahorró a la compañía 98 millones de minutos ociosos o alrededor de 25 millones de dólares de costo de mano de obra cada año.
- En otras palabras, este simple cambio aumentó los beneficios, satisfizo las demandas de los clientes, mejoró la seguridad y afectó positivamente al medio ambiente.

Componentes de un sistema BI

- Data Warehouse y Data Mart
- Datamining (*extracción inteligente**) y herramientas de reporting, quering y dashboards.
- Herramientas de visualización y de descubrimiento
- Herramientas analíticas (p.e. OLAP)

(*) intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de BBDD.

Data Warehouse

- **Orientado a temas.-** Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden conectados entre sí.
- **Variante en el tiempo.-** Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.
- **No volátil.-** La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de *sólo lectura*, y se mantiene para futuras consultas.
- **Integrado.-** La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.
- **Separación** de los datos usados en operaciones diarias de los datos usados en el almacén de datos para los propósitos de divulgación, de ayuda en la toma de decisiones, para el análisis y para operaciones de control. Ambos tipos de datos no deben coincidir en la misma base de datos, ya que obedecen a objetivos muy distintos y podrían entorpecerse entre sí.

Data Mart

- Un **data mart** es una versión especial de almacén de datos (data warehouse).
- Son subconjuntos de datos con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones.
- Los datos existentes en este contexto pueden ser agrupados, explorados y propagados de múltiples formas para que diversos grupos de usuarios realicen la explotación de los mismos de la forma más conveniente según sus necesidades.

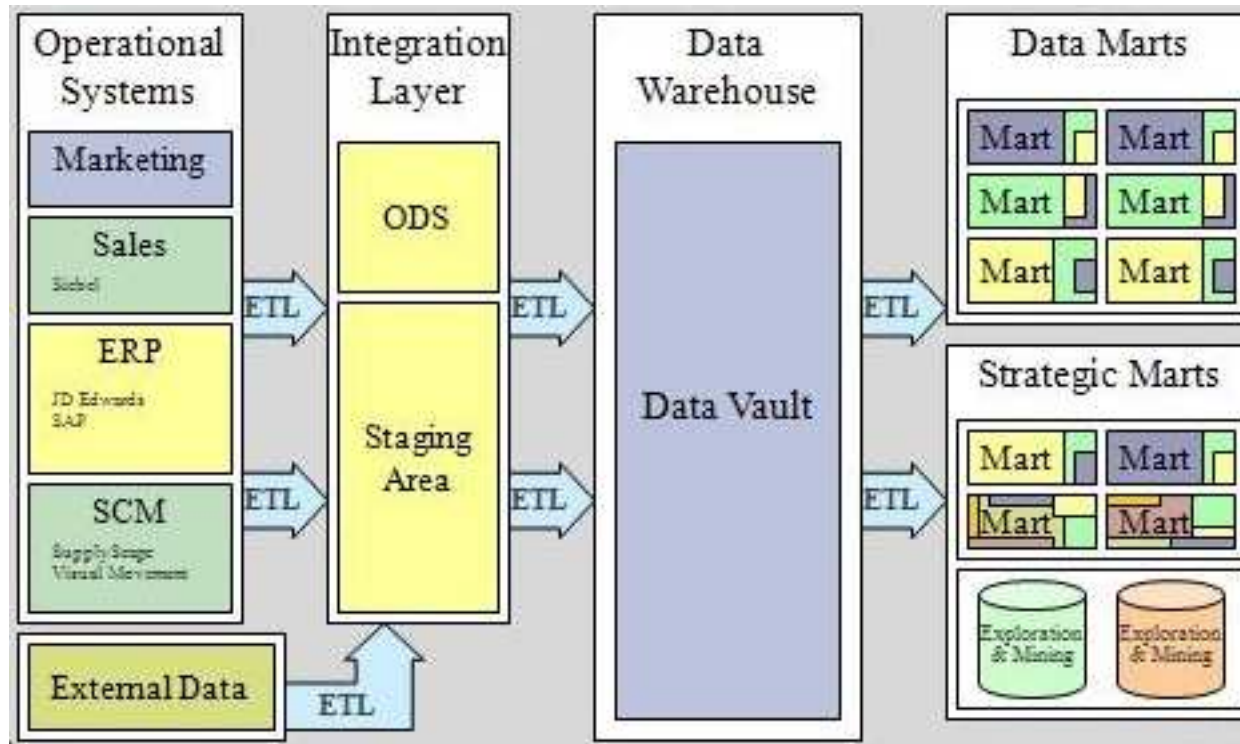
Data Mart

- El Data mart es un sistema orientado a la consulta, en el que se producen procesos batch de carga de datos (altas) con una frecuencia baja y conocida.
- Sobre estas bases de datos se pueden construir **ESS** (Executive Support Systems, Sistemas de Información para Directivos) y **DSS** (Decision Support Systems, Sistemas de Ayuda a la toma de Decisiones).

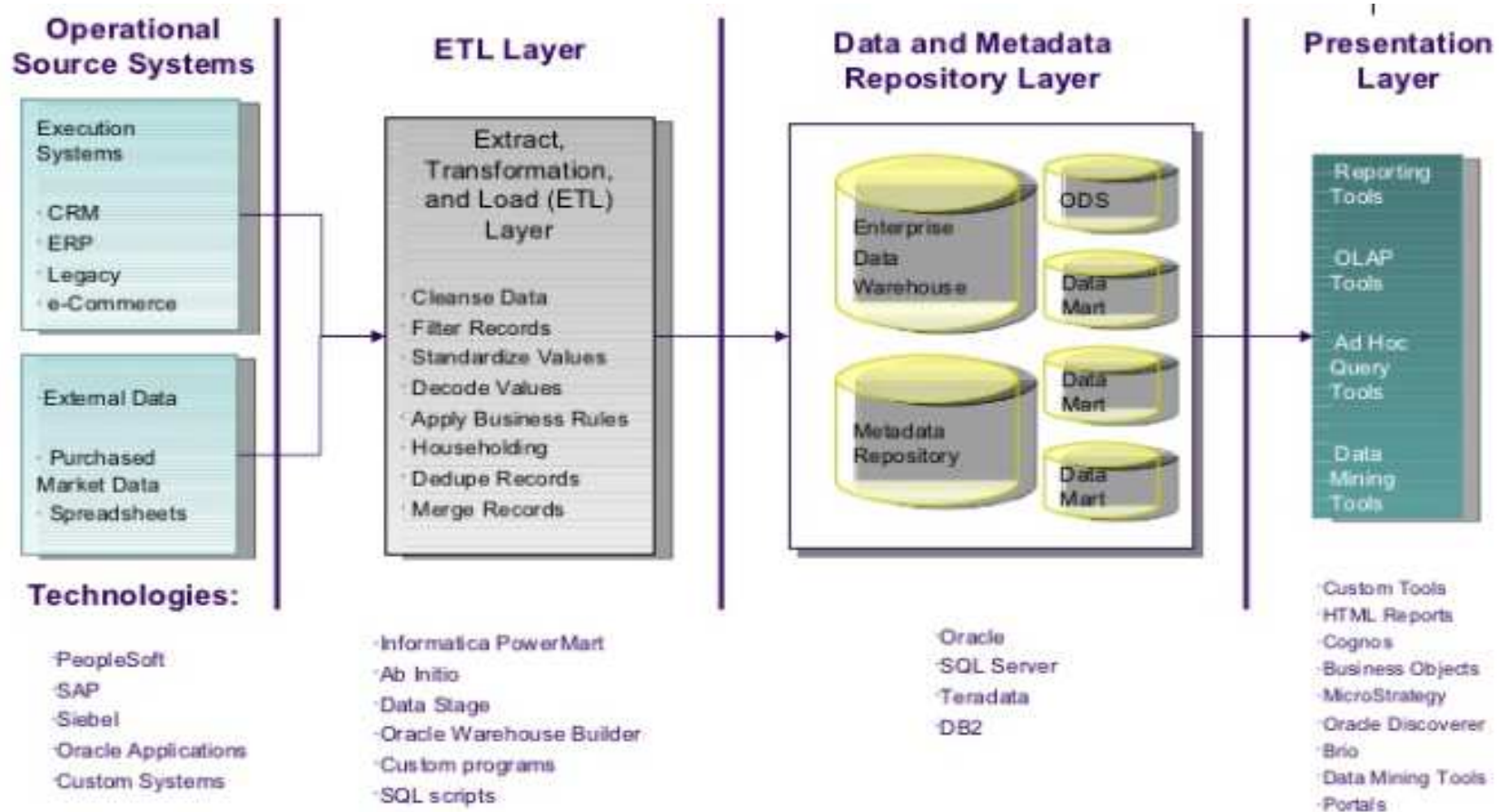
Data Warehouse y Data Mart

- Ambos suelen ser consultados mediante herramientas OLAP (On line Analytical Processing - Procesamiento Analítico en Línea) que ofrecen una visión multidimensional de la información.
- Lo veremos...

Data Warehouse y Data Mart



Arquitetura Data Warehouse



ETL

- Las herramientas de **BI** extraen datos de interés de diferentes fuentes tales como CRM, ERP, SCM, sistemas antiguos, data warehouse, data marts y/o Web. Los datos extraídos no suelen estar en formatos utilizables.
- Los procesos de Extract, transform and load (ETL) son importantes ya que son la forma en que los datos se guardan en un almacén de datos (o en cualquier base de datos).

ETL

- Implican las siguientes operaciones:
 - Extracción. Acción de obtener la información deseada a partir de los datos almacenados en fuentes externas.
 - Transformación. Cualquier operación realizada sobre los datos para que puedan ser cargados en el data warehouse o se puedan migrar de éste a otra base de datos. Estandarización de datos.
 - Carga. Consiste en almacenar los datos en la base de datos final (por ejemplo, en la Base de Datos operacional)
- Los datos estandarizados se pueden analizar y cargar en otros sistemas operacionales o utilizarlos para emitir informes (reporting).

Reporting

- Los **sistemas de reporte** de empresas proporcionan **informes ad hoc (o a medida)** de fuentes de datos.
- Los usuarios acceden a la información e informes que necesitan directamente.
- El sistema/servicio **mejora el control y reduce la latencia de datos.**
- Los **informes rutinarios se generan automáticamente** y se distribuyen periódicamente a listas de suscriptores internos o externos.
- Son muy usuales los informes de ventas semanales, unidades producidas por día, semana o mes...

Herramientas de consulta (query)

- La consulta identifica los datos que el usuario desea examinar.
- Ejemplo: ingresos de venta.
 - Estos datos se pueden visualizar en tres dimensiones: ventas, regiones, productos y tiempo en trimestres.
 - Los resultados de la consulta se pueden configurar como un cubo multidimensional.
- **Cualquier consulta que no esté predefinida es una consulta ad hoc.**
- Las consultas ad hoc permiten a los usuarios solicitar información que no esté disponible en informes periódicos, así como generar nuevas consultas o modificar las antiguas.

Dashboards

- Las **herramientas de informes y consultas** se unen a las **herramientas dashboards** para facilitar interfaces interactivas .
- Los **dashboards** o tableros de control donde **se resume visualmente** la información.
- Los usuarios de negocios utilizan estas herramientas para **monitorización y análisis de información crítica y de métricas**.
- La información se presenta en diagramas, gráficos...
- Los dashboards están diseñados para soportar una función **específica**.
 - Por ejemplo, un tablero de control de marketing informa de métricas tradicionales: ventas, costes, márgenes comerciales...
 - Los tableros de contabilidad de cuentas, facturas pagadas, métricas de rentabilidad...
 - Y otros parámetros más complejos..

Así vende el dashboard...



...la realidad es más parecida a ésta



SAP BusinessObjects BI Platform
SAP Lumira

Mercado

- La mayoría de los proveedores de BI proporcionan aplicaciones para conectar e integrar sistemas ERP y CRM.
- Ejemplos:
 - SAGE Intelligence
 - Business Objects de SAP
 - Oracle Business Intelligence 12^c
 - Cognos de IBM
 -



Selección de features

Deployment

- ☐ Web-Based
- ☐ Installed

Features

- ☐ Ad hoc Analysis
- ☐ Ad hoc Query
- ☐ Ad Hoc Reports
- ☐ Benchmarking
- ☐ Budgeting & Forecasting
- ☐ Dashboard
- ☐ Data Analysis
- ☐ Data Visualization
- ☐ Key Performance Indicators
- ☐ OLAP
- ☐ Performance Metrics
- ☐ Predictive Analytics
- ☐ Profitability Analysis
- ☐ Strategic Planning
- ☐ Trend / Problem Indicators

Herramientas de Análisis

Multidimensional: OLAP

- OLAP (On-Line Analytical Processing) es una técnica conocida también como **análisis multidimensional**.
- Facilita el **análisis de datos a través de dimensiones y jerarquías** (distintos niveles de agrupamiento) utilizando **consultas rápidas predefinidas y subtotales previamente calculados**.
- La razón de utilizar OLAP para consultas es la **rapidez** de respuesta en el caso de múltiples dimensiones.

OLAP

- El modelo multidimensional está constituido por **tres componentes**:
 - **Dimensiones**: Grupos conceptuales que permiten analizar o consolidar datos: productos, clientes, zonas geográficas, tiempo...
 - **Medidas o indicadores**: Valores numéricos que se guardan en la base de datos: unidades vendidas, consumos...
 - **Jerarquías de dimensiones**: Distintos niveles de agregación dentro de una **dimensión**: ventas por día, ventas por meses, ventas por trimestre, ventas por año...
- Se lleva a cabo un proceso de cálculo de subtotales durante la etapa de carga del sistema, con varios niveles de agrupamiento, asignando cada uno de estos subtotales a las distintas celdas que constituyen el cubo multidimensional.

Ejemplo

- Una empresa comercializa 4 productos: TV, frigoríficos, lavadoras, y calefactores en las regiones : norte, sur, este y Oeste.
- Las bbdd tradicionales responden muy bien a consultas del tipo: ¿Cuántas ventas se han realizado en el último trimestre del año en el norte?
- Pero si se desea conocer cuántas lavadoras se vendieron en cada una de las regiones de venta y comparar sus resultados con sus previsiones de ventas, es más rápido con las aplicaciones OLAP.

Características de los sistemas OLAP

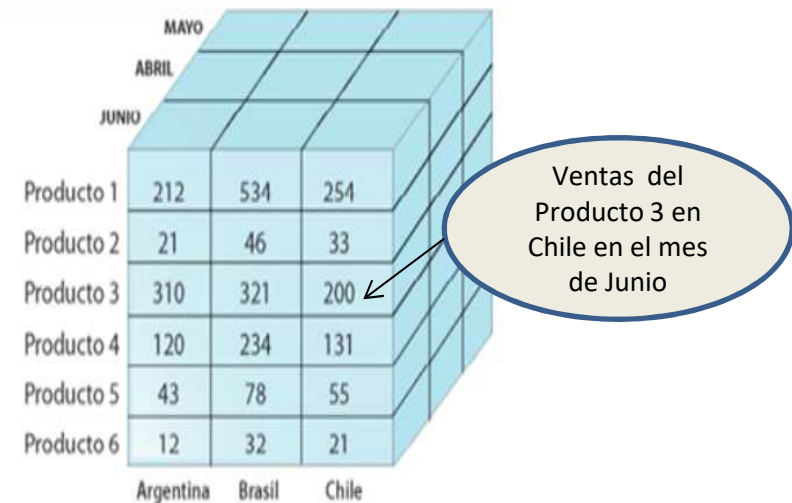
- **Análisis multidimensional:** cada característica de la información de una aplicación de una empresa (productos, precio, artículos vendidos, año, región....) se representa por una dimensión diferente.
- No impone restricciones sobre el número de dimensiones.
- **Respuestas rápidas** a problemas ad hoc.
- **Alto nivel de detalle de cada operación.**
- **Permite definir de forma flexible sobre las dimensiones: restricciones y agregaciones.**
- La **herramienta OLAP genera la correspondiente consulta** y la envía al gestor de consultas del sistema (p.ej. mediante una sentencia **SELECT**).

Operadores y Elementos de Control de los sistemas OLAP

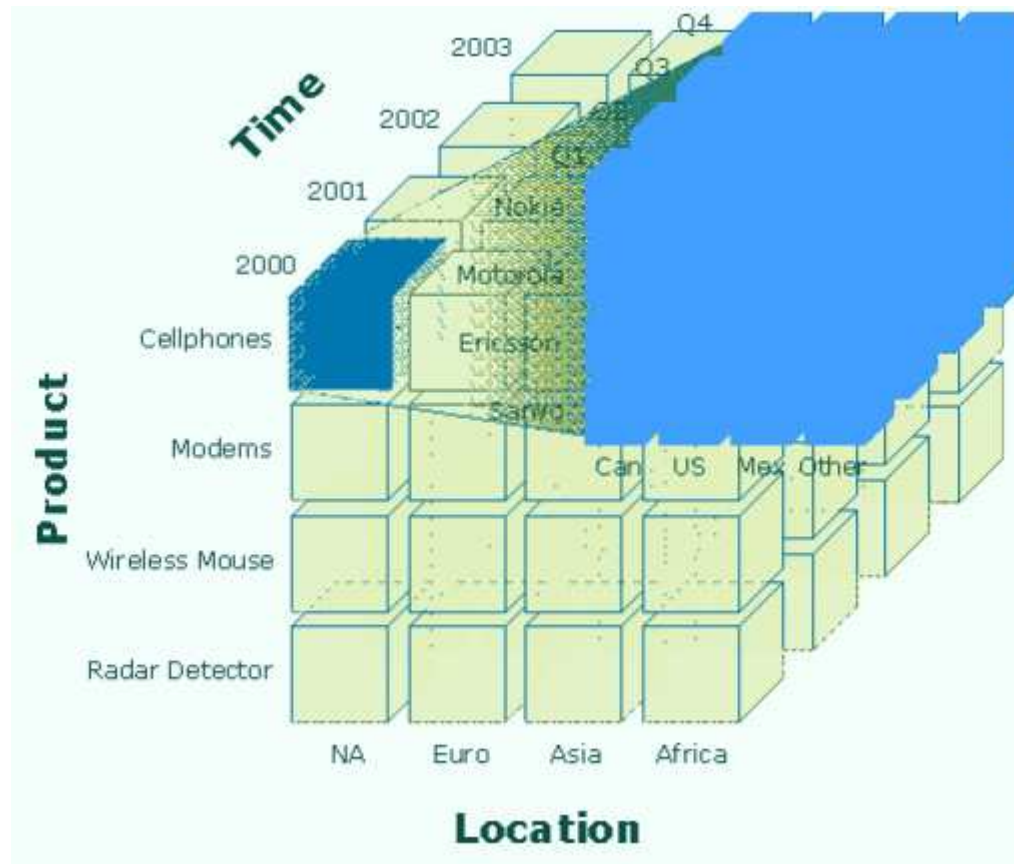
- Lo realmente interesante de las herramientas OLAP son sus operadores de refinamiento o manipulación de consultas: drill , roll , slice & dice, swap.
- Elementos de control:
 - Variables de decisión: representan una medición del negocio.
 - Se basan en el concepto de cubo de datos

Cubos OLAP

- ✓ Para funcionar, las aplicaciones OLAP, utilizan un tipo de base de datos que posee la peculiaridad de ser multidimensional, denominada **cubo OLAP**.
- ✓ Un cubo OLAP es una base de datos que posee varias dimensiones.
- ✓ Se amplían así las posibilidades de las bases de datos relacionales, permitiendo el **procesamiento de importantes volúmenes de información**.
- ✓ Cada una de las dimensiones que posee la base de datos incorpora **un campo determinado para un tipo de dato específico**, que luego podrá ser comparado con la información contenida en el resto de dimensiones.
- ✓ En general los cubos OLAP están compuestos por 3 dimensiones.
- ✓ Una base de datos OLAP puede contener varios cubos.



Cubos OLAP...y esto también



Pasos para el diseño de OLAP

1. Identificar y **seleccionar los procesos** de la empresa que necesitan mediciones.
2. Definir el **nivel de detalle**: ¿Qué nivel de detalle de los datos debe estar en el modelo Dimensional?
 - Hay que definir las **tablas de hechos (fact)** con el nivel de detalle en cada una de ellas.
3. Escoger las **dimensiones** para cada **tabla de hechos**.
 - En los atributos de las dimensiones incluir todas las descripciones posibles con las que luego se quieran referir las métricas de las tablas de hechos.
4. Identificar las **métricas (measures)**: identificar los campos base de mediciones **en cada tabla de hechos**.
 - Los campos que se definan deben corresponder con el nivel de detalle del paso 2
 - Deben ser numéricos y aditivos (que tenga sentido hacerlo)

Ejemplo

Tabla de Dimensión Fechas

Dim_Date	
Id	
Date	
Day	
Day_of_Week	
Month	
Month_Name	
Quarter	
Quarter_Name	
Year	

Tabla de Dimensión Almacenes

Dim_Store	
Id	
Store_Number	
State_Province	
Country	

Tabla de Hecho Ventas

Fact_Sales	
Date_Id	
Store_Id	
Product_Id	
Units_Sold	

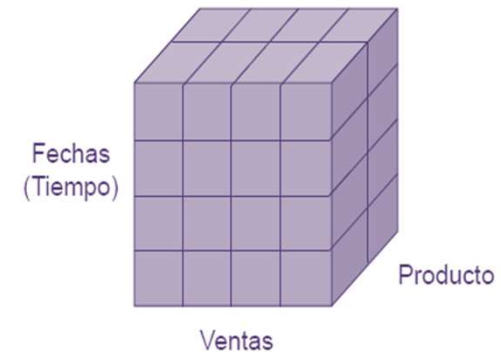
Dim_Product	
Id	
EAN_Code	
Product_Name	
Brand	
Product_Category	

Tabla de Dimensión Productos

*Niveles de Detalle-Levels
y Atributos-attributes*

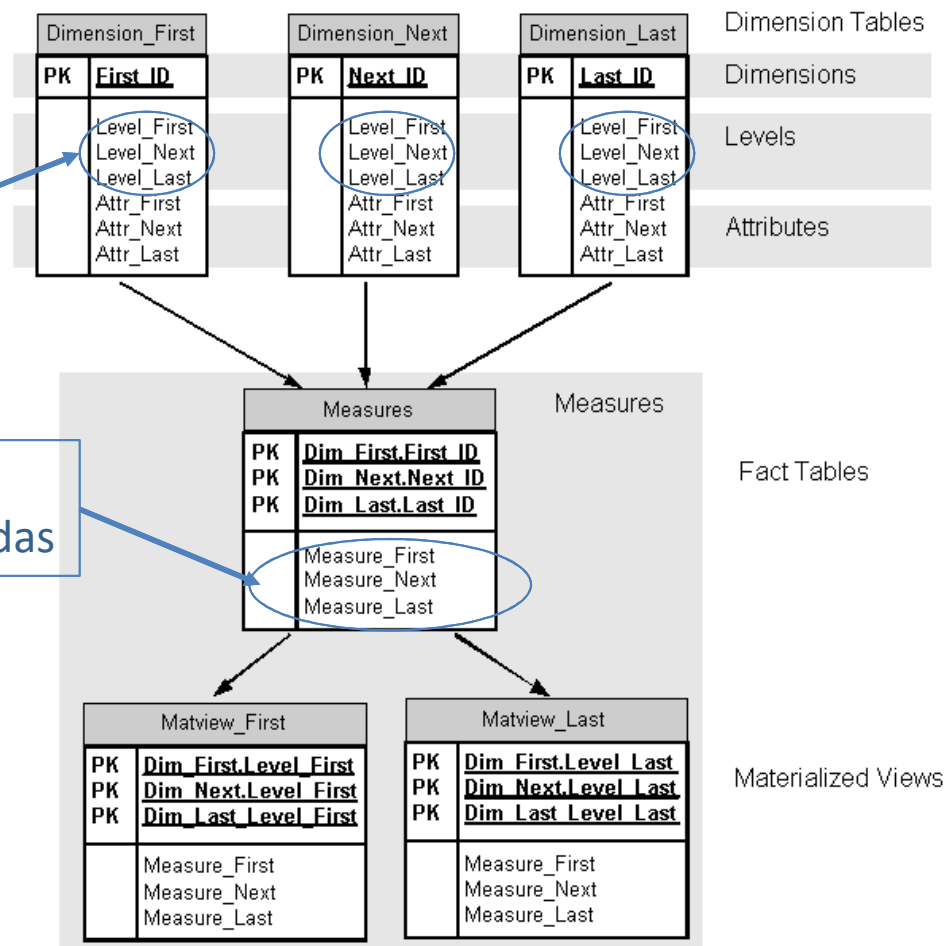
Consulta

- Una consulta a un almacén de datos consiste generalmente en la obtención de medidas (measures) sobre los hechos (facts) y atributos de las dimensiones, restringidas por **condiciones (nivel de detalle)** impuestas sobre las dimensiones.
- Ejemplo: Si los productos son distintos tipos de televisiones, el nivel de detalle de las fechas es trimestral, y las ventas se expresan por ciudades, una consulta podría ser:
 - Importe total de las ventas de este año del producto TV en Chile



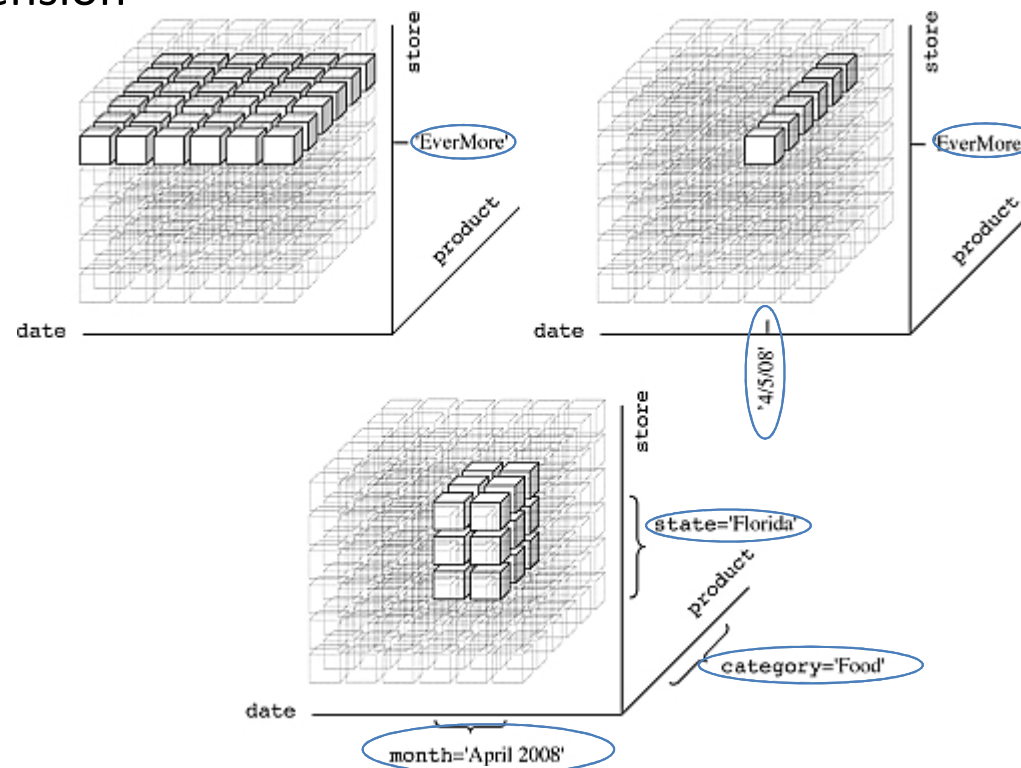
Distintos Niveles
de Detalle,
ya definidos

Distintas Medidas
ya cargadas/calculadas



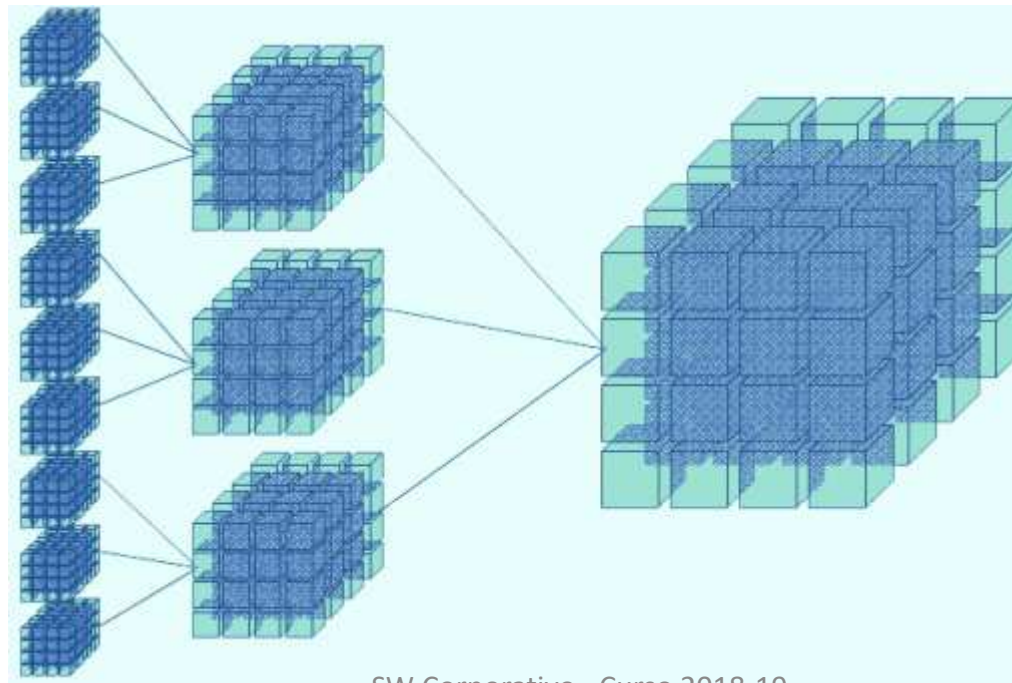
Operaciones básicas en OLAP

- **Drill up/ down:** para presentar los datos con mayor o menor nivel de detalle.
 - Introduce un **criterio más o menos restrictivo** sobre un atributo de una dimensión



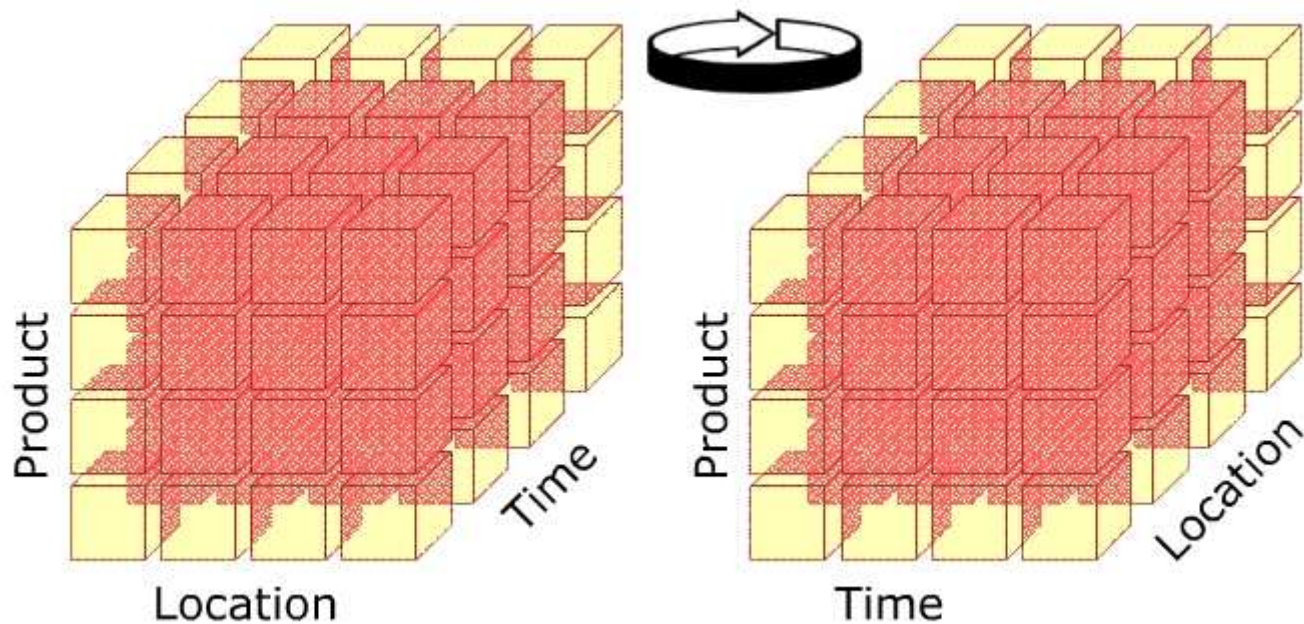
Operaciones básicas en OLAP

- **Roll-up (agregar, consolidar):** se evalúan todas las relaciones para una o más dimensiones, dada una fórmula.



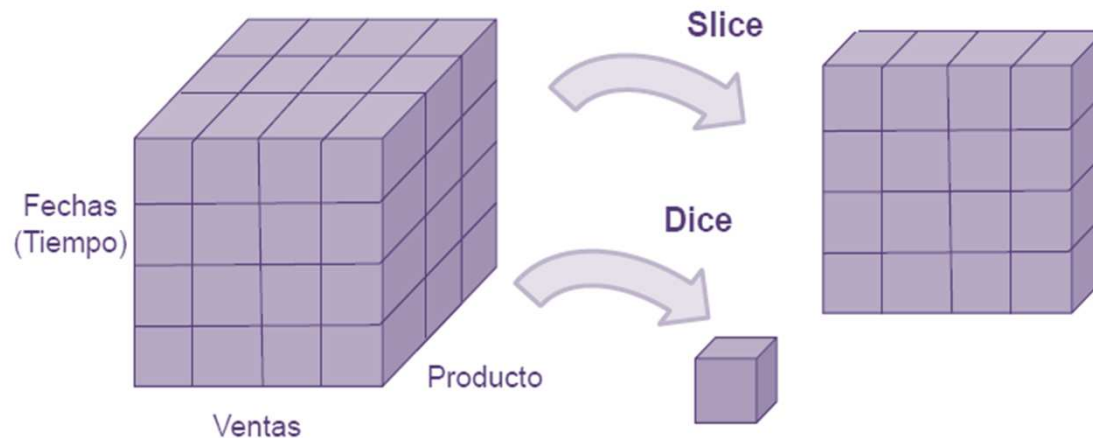
Operaciones básicas en OLAP

- **Pivot (rotar)**: Permite rotar las dimensiones para tener una presentación alternativa de los datos.



Operaciones básicas en OLAP

- **Slice:** produce la selección de una “rebanada” de un cubo fijando un valor para alguna de sus dimensiones.
- **Dice:** produce la selección de un sub-cubo fijando las dimensiones con valores específicos.



Arquitecturas OLAP

- Para realizar el proceso de las consultas o análisis OLAP existen los siguientes tipos de sistemas:
 - ROLAP OLAP Relacional
 - MOLAP OLAP Multidimensional
 - HOLAP OLAP Híbrido (Rolap y Molap)

Sistemas ROLAP

- Implementación OLAP que almacena los datos en una **bbdd relacional** y utiliza un motor ROLAP para proporcionar la función analítica.
- El motor ROLAP:
 - Ejecuta las consultas multidimensionales de los usuarios.
 - Transforma dinámicamente sus consultas a consultas SQL.

Sistemas MOLAP

- Se implementa con una **bbdd multidimensional** donde se almacenan los datos para proporcionar su análisis.
- Los datos se organizan en una estructura tipo cubo que el usuario puede rotar.
- Se realizan con gran velocidad las consultas.
- Para la carga de datos requiere de cálculos intensos de compilación
- Muy adecuado para resúmenes e informes financieros.

Sistemas. HOLAP

- Amacena algunos datos en un motor relacional y otros, en una base de datos multidimensional.
- Trata de combinar las ventajas de MOLAP y ROLAP.
- Los datos que subyacen en los hipercubos son almacenados en una estructura relacional y las agregaciones en una estructura multidimensional.
- Por ejemplo:
 - Si se necesita información tipo resúmenes, HOLAP potencia la tecnología de cubos para el rendimiento más rápido.
 - Si se necesita información muy detallada, un dato concreto, HOLAP se apoya en los datos relacionales.

Proveedores de OLAP

- Microsoft, SQL Server Analysis Services (SSAS)
- SAP, Business Object
- IBM, Cognos OLAP
- Oracle, BI