

## Seminario:

*Construyendo una solución de BI paso a paso con SQL Server 2005*

Ing. José Mariano Álvarez  
[jose.mariano.alvarez@SqlTotalConsulting.com](mailto:jose.mariano.alvarez@SqlTotalConsulting.com)



## Modelamiento de Datos Modelos Multidimensionales



### Retos del Diseño de un DW / Data Mart

- Diseñar esquemas físicos legibles para el usuario de negocio que provean respuestas rápidas y predecibles que soporten el análisis impredecible de largos volúmenes de datos históricos y su contexto.



### Los esquemas deben:

- Facilitar la formulación de preguntas de negocio.
- Proveer respuestas rápidas y predecibles para múltiples vías de acceso.
- Proveer facilidad para carga y mantenimiento.
- Ser extensibles.
- Ser sensibles al cambio.
- Proveer una vista exacta de cambios históricos.



### Diseño Dimensional

- Centrado en el análisis de procesos de negocios y eventos.
- Modelamiento en términos de dimensiones y hechos.
- Denormalización de las dimensiones jerárquicas (denormalización administrable).



### Opciones de Modelo de Datos

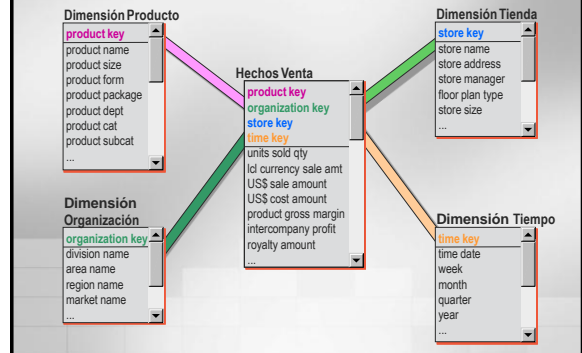


## ¿Qué es un Star Schema?

- Es un Esquema de base de datos relacional para representar datos de forma multidimensional.
- Es la forma más simple de un esquema de Data Warehouse que contiene una o más dimensiones y tablas de hechos.
- El Centro de la estrella es una tabla de hechos grande que está conectada a las tablas dimensionales.
- Las ventajas del Star Schema son:
  - Fácil entendimiento de los datos
  - Alto rendimiento (Desempeño)



## Ejemplo de un Star Schema



## Tabla de Hechos

- Contiene las medidas relacionadas a los procesos o eventos que son analizados por las dimensiones contenidas en las tablas de dimensiones.
- Las filas registran que un evento ha ocurrido.
- Define la relación entre 2 o más dimensiones.



## Características de las Tablas de Hechos

- Son usualmente las tablas más grandes.
- Son altamente normalizadas.
- Normalmente solo se añaden registros.
- Pueden crecer rápidamente.
- Pueden contener detalles o datos resumidos.
- Las medidas no necesariamente son agregadas.
- Son unidas a tablas de dimensiones mediante llaves foráneas que referencia a las llaves primarias de las tablas de dimensiones.



## Tipos de Tabla de Hechos

- Las tablas de hechos contienen medidas relacionadas con:
  - Transacciones.
  - Eventos.
  - Valores de puntos en el tiempo (instantáneas).
  - Otros.



## Ejemplo de Tabla de Hechos

- Ventas.
- Ordenes.
- Envíos.
- Cuentas.
- Reservaciones
- Rentas.
- Nivel de Inventario.
- Balance Contable.
- Chequeo de TxS.
- TxS ATM.
- Suscripciones.
- Llamadas (CDRs).



## Granularidad

- Todas las medidas deben estar al mismo nivel de detalle.
- Detalle de Ordenes
  - Id de Fecha
  - Id de Cliente
  - Id de Producto
  - Unidades
  - Monto
  - Monto total de la orden (no pertenece)



## Contenido de las tablas de Hechos

ProductSales Fact

product key
organization key
store key
time key
units sold qty
lcl currency sale amt
US\$ sale amount
US\$ cost amount
product gross margin
intercompany profit
royalty amount
...

Usualmente solo llaves de dimensiones y medidas.

## Ejemplo

- Tabla de Hechos

Product Key	Order Date Key	Due Date Key	Ship Date Key	Customer Key	Territory Key	Sales Amount	Tax Amount	Sales Order Number	Freight
1	2	3	4	12	1	100	21.57	SO43701	79.20
1	2	2	2	11	2	150	30.20	SO43702	75.30
2	2	2	3	11	2	200	14.56	SO43700	70.20



## Tablas de Dimensiones

- Contienen perspectivas del negocio.
- Utilizan terminología de negocio.
- Columnas extensas.
- Contienen dato textual y discreto.
- Son usualmente más pequeñas que las tablas de hechos.
- Pueden contener jerarquías denormalizadas (colapsadas).



## Tablas de Dimensiones

- Están unidas a una tabla de hechos mediante una llave foránea que referencia a su llave primaria.



## Las Tablas de Dimensiones contienen los parámetros analíticos

- La cantidad vendida es analizada por día, mes, trimestre o año.
- Cantidades vendidas en los días soleados vs. los días lluviosos.
- Cantidad de inventario analizado por almacén, por producto.
- Ganancia analizada por producto, categoría, departamento, tienda, distrito, región.



## Ejemplo de Tabla de Dimensiones

- Clientes.
- Almacén.
- Productos.
- Servicios.
- Empleados.
- Partes.
- Contrato.
- Vendedor.
- Cuentas.
- Ubicación.
- Tiempo.

SQL TOTAL CONSULTING

## Ejemplos

### • Dimensión Tiempo

TimeKey	FullDate	DayName	Month	Quarter	Semester	Year
1	2001-07-01	Sunday	July	3	2	2001
2	2001-07-02	Monday	July	3	2	2001
3	2001-07-03	Tuesday	July	3	2	2001
4	2001-07-04	Wednesday	July	3	2	2001

### • Dimensión Producto

Product Key	Product Code	Product Name	Product Subcategory	Color	Status
218	HL-U509	Sport-100 Helmet, Black	31	Black	Current
219	SO-B909-M	Mountain Bike Socks, M	23	White	NULL
220	SO-B909-L	Mountain Bike Socks, L	23	White	NULL
221	HL-U509-B	Sport-100 Helmet, Blue	31	Blue	NULL

## Ejemplos

### • Dimensión Empleado

Employee Key	Parent Employee Key	Sales Territory Key	First Name	Last Name	Marital Status	Email Address
3	14	11	Roberto	Brown	2	rbrown@adw.com
4	3	11	Rob	Ford	2	rford@adw.com
5	3	11	Kevin	Walters	2	kwalters@adw.com
6	267	11	Mary	Higa	2	mhiga@adw.com

### • Dimensión Territorio

Territory Key	Territory Code	Sales Territory Region	Sales Territory Country	Sales Territory Group
1	1	NorthWest	United States	North America
2	2	NorthEast	United States	North America
3	3	Central	United States	North America

SQL TOTAL CONSULTING

## Pasos para Diseñar un Star Schema

- Identificar un proceso de negocio para el análisis (ejemplo: Ventas)
- Identificar las medidas o hechos (Monto en Pesos).
- Identificar las dimensiones para los hechos (Producto, Ubicación, Tiempo, Organización).
- Listar los campos que describen cada dimensión (nombre de la región, nombre de la sucursal, categoría de producto, etc.)
- Determinar el nivel más bajo (granularidad) de agregación en la tabla de hechos




SQL TOTAL CONSULTING

## Aspectos Importantes de los Star y Snowflake Schemas

- En un Star Schema cada dimensión deberá tener una Llave Primaria
- En un Star Schema hay una tabla por dimensión
- En un Snowflake schema una tabla dimensional tendrá una o más tablas padre
- Las jerarquías están en una sola tabla en el star schema
- En un Snowflake las jerarquías están distribuidas en diferentes tablas

SQL TOTAL CONSULTING

## Consideraciones para el diseño de la Dimensión Tiempo

- 
 Se debe usar una tabla dimensional o un server time dimension?
- 
 Qué granularidad debo usar?
- 
 Qué calendarios y jerarquías se requieren?

SQL TOTAL CONSULTING

# Qué son Dimensiones Lentamente Cambiantes (Slowly Changing Dimensions) ?

- Type 1 slowly changing dimension**  
Cambio de un atributo sobrescribe filas.
- Type 2 slowly changing dimension**  
Un cambio crea nuevos registros. Marca de registro actual o expirado.
- Type 3 slowly changing dimension**  
Se mantiene el dato original, el nuevo, y la fecha efectiva de cambio.



# SCD Ejemplos

Customer Key	Name	State
1001	Christina	Illinois

→ California

## SCD Tipo 1

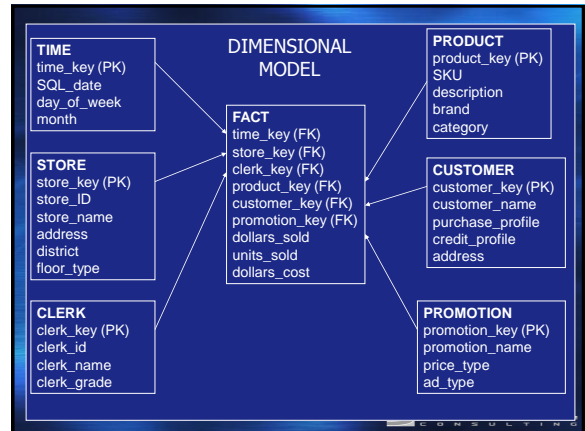
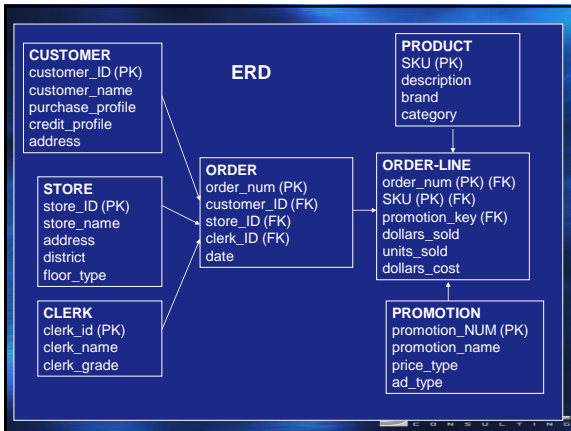
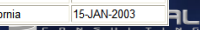
Customer Key	Name	State
1001	Christina	California

## SCD Tipo 2

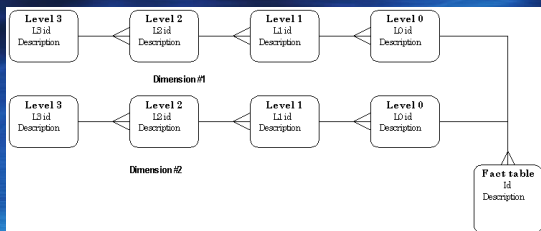
Customer Key	Name	State
1001	Christina	Illinois
1005	Christina	California

## SCD Tipo 3

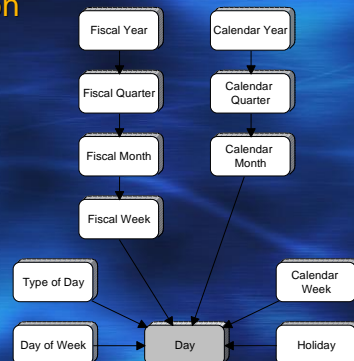
Customer Key	Name	Original State	Current State	Effective Date
1001	Christina	Illinois	California	15-JAN-2003



# Jerarquías



# Dimensión Tiempo



Attribute Name	Attribute Description	Sample Values
Day	The specific day that an activity took place.	06/04/1998; 06/05/1998
Day of Week	The specific name of the day.	Monday; Tuesday
Holiday	Identifies that this day is a holiday.	Easter; Thanksgiving
Type of Day	Indicates whether or not this day is a weekday or a weekend day.	Weekend; Weekday
Calendar Week	The week ending date, always a Saturday. Note that WE denotes	WE 06/06/1998; WE 06/13/1998
Calendar Month	The calendar month.	January, 1998; February, 1998
Calendar Quarter	The calendar quarter.	1998Q1; 1998Q4
Calendar Year	The calendar year.	1998
Fiscal Week	The week that represents the corporate calendar. Note that the F	F Week 1 1998; F Week 46 1998
Fiscal Month	The fiscal period comprised of 4 or 5 weeks. Note that the F in the data	F January, 1998; F February, 1998
Fiscal Quarter	The grouping of 3 fiscal months.	F 1998Q1; F1998Q2
Fiscal Year	The grouping of 52 fiscal weeks / 12 fiscal months that comprise the financial year.	F 1998; F 1999

## Tipos de Dimensiones

- Regular
- Role-playing
- Fact
- Many-many
- Referenced



## Regular Dimension

- Un foreign key existe entre la tabla de hechos y la tabla de dimensión




## Role-playing Dimension

- Una dimensión puede ser usada múltiples veces sobre el mismo cubo/measure group




- Se ven dos dimensiones distintas

e.g.  
OrderDate  
ShippedDate




## Fact Dimension

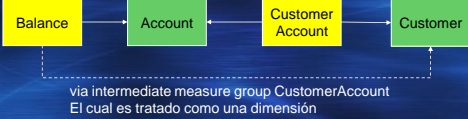
- Una dimensión basada sobre una fact table




- Un Multidimensional (MOLAP) measure group puede tener una dimensión relacional (ROLAP)



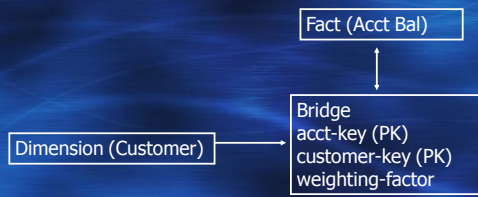
## Many-many Dimension



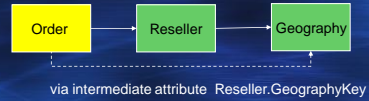
via intermediate measure group CustomerAccount  
El cual es tratado como una dimensión



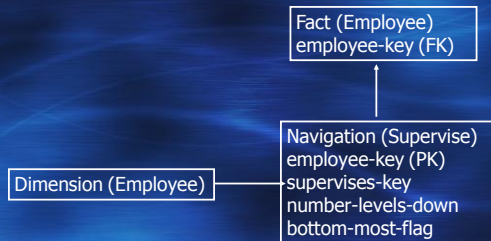
## Many to many



## Referenced Dimension



## Recursiva



## Modelo Analítico Multidimensional

## Elementos de un modelo multidimensional - OLAP

- Métricas
- Dimensiones
- Jerarquías
- Atributos

## Métricas

- Lo que queremos analizar
- Indicadores de gestión
- Valores numéricos del análisis
- Ejemplos:
  - Ventas en Monto
  - Ventas en unidades
  - Cantidad de Clientes
  - Utilidad
  - Margen de Contribución
  - Costo
  - Stock

## Dimensiones

- Son las perspectivas de análisis
  - Analizamos las ventas
    - Por Tienda
    - Por Vendedor
    - Por Producto
    - Por Semana
    - Por Departamento
    - Por Marca
    - Por Mes
    - Por Territorio de Ventas
    - Por País



## Jerarquías

- Son los niveles de agrupación dentro de una dimensión.
- También se conocen como los caminos de “drill-down” o “roll-up”
- Una dimensión puede tener más de una jerarquía.
- Ejemplo: Dimensión Tiempo
  - Jerarquías:
    - Año-Trimestre-Mes-Día
    - Año-Semana-Día



## Atributos

- Son características de los miembros de las dimensiones
- Pueden ser usados para clasificar, filtrar, o buscar información dentro de la base de datos
- Ejemplo:
  - Dimensión = Producto, Miembro = Producto
  - Atributos: Unidad de medida, Empaque, Color, Categoría, Sub Categoría, Marca
- Algunos atributos se pueden convertir en miembros de una dimensión en una de sus jerarquías.
- Todos los miembros de una dimensión son atributos de la misma, no todos los atributos de una dimensión pertenecen a una jerarquía.



**Microsoft**  
Your potential. Our passion.™

<http://www.microsoft.com/sql/bi>  
<http://www.microsoft.com/sql/>

© 2003-2005 Microsoft Corporation. All rights reserved. This presentation is for informational purposes only. Microsoft makes no warranties, express or implied, in this document.

