

MODELO DIMENSIONAL

Eric Gustavo Coronel Castillo

youtube.com/c/DesarrollaSoftware

www.desarrollasoftware.com

gcoronel@uni.edu.com

Logro Esperado

Al finalizar esta presentación, se espera que el participante aprenda a realizar un modelo dimensional de una solución de BI.



Presentación



Fue propuesto en el año 1996 por RALPH KIMBALL como propuesta para el diseño de un dataware house partiendo de la **visión multidimensional** que lo usuarios tienen de los datos empresariales cuando se enfrentan a ellos con propósitos de análisis.

El análisis multidimensional consiste en analizar los datos que hacen referencia a **hechos**, ya sean económicos o de otros tipos, utilizando algún tipo de **métrica o medida de negocio**.



```

graph TD
    Sales((Sales)) --- Region((Region))
    Sales --- Product((Product))
    Sales --- SalesChannel((Sales channel))
    Sales --- Date((Date))
    Sales --- Setter((Setter))
    Sales --- Customer((Customer))

```

Region

- Country
- City
- Post code

Product

- Product line
- Name
- Price

Sales channel

- Medium
- Name
- Contact

Date

- Year
- Month
- Day

Setter

- Name
- Contact
- City

Customer

- Customer group
- Name
- Post code
- City
- Age

Sales

- Region_ID
- Product_ID
- Sales_Channel_ID
- Date_ID
- Setter_ID
- Customer_ID
- Quantity

Fact table

Dimension table

Dimensional Modeling (DM) es una técnica de diseño lógico que tiene como objetivo presentar los datos dentro de un marco de trabajo estándar e intuitivo, para permitir su acceso con un alto rendimiento.

Definición



Cada Modelo Dimensional está compuesto por una tabla con una llave combinada, llamada tabla de hechos, y un conjunto de tablas más pequeñas llamadas tablas de dimensiones.



Figura 1. Modelo Estrella



Hechos y Medidas:

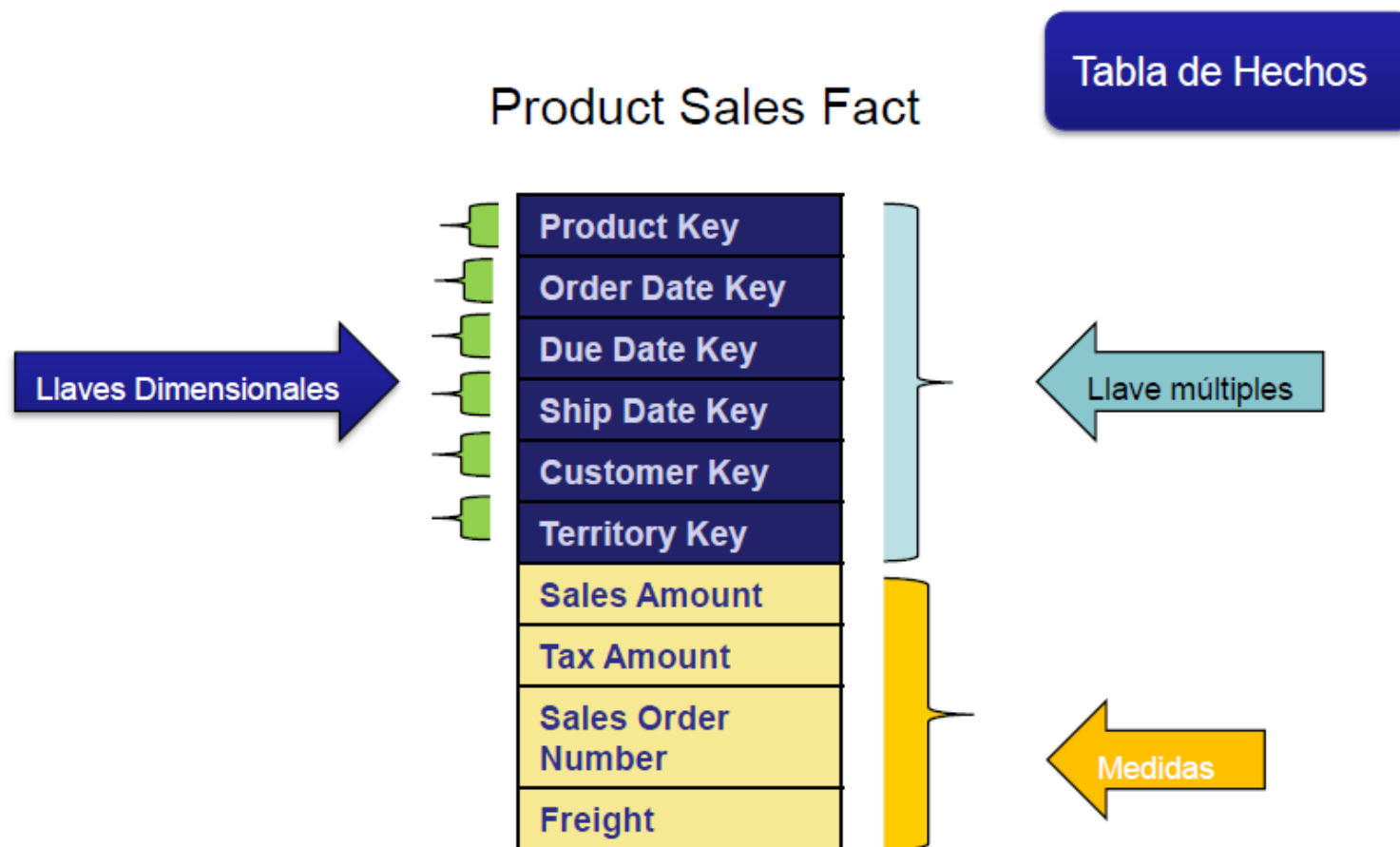
- Son las variables y métricas que ayudarán a medir la performance del negocio.
- Las medidas mas usadas son numéricas y aditivas.
- Una fila en una tabla de hechos corresponde a una medida.
- Todas las medidas en la tabla de hechos deben estar al mismo nivel de granularidad.

Hechos y Medidas

Ejemplos:

- Ventas en Monto
- Ventas en unidades
- Cantidad de Clientes
- Utilidad
- Margen de Contribución
- Costo
- Stock

Contenido de la Tabla de Hechos



Adicionalmente puede tener atributos.


```

graph TD
    Sales((Sales)) --- Region((Region))
    Sales --- Product((Product))
    Sales --- SalesChannel((Sales channel))
    Sales --- Date((Date))
    Sales --- Setter((Setter))
    Sales --- Customer((Customer))
  
```

Region

- Country
- City
- Post code

Product

- Product line
- Name
- Price

Sales channel

- Medium
- Name
- Contact

Date

- Year
- Month
- Day

Setter

- Name
- Contact
- City

Customer

- Customer group
- Name
- Post code
- City
- Age

Sales

- Region_ID
- Product_ID
- Sales_Channel_ID
- Date_ID
- Setter_ID
- Customer_ID
- Fact

Facty table

Dimension table

- 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).
- Dentro de cada dimensión se puede definir los niveles de agregación o sumariaización para cada análisis, a estos niveles de granularidad se los caracteriza con el nombre de atributos.

```

graph TD
    Sales((Sales)) --- Region((Region))
    Sales --- Product((Product))
    Sales --- SalesChannel((Sales channel))
    Sales --- Date((Date))
    Sales --- Setter((Setter))
    Sales --- Customer((Customer))

```

Region

- Country
- City
- Post code

Product

- Product line
- Name
- Price

Sales channel

- Medium
- Name
- Contact

Date

- Year
- Month
- Day

Setter

- Name
- Contact
- City

Customer

- Customer group
- Name
- Post code
- City
- Age

Sales

- Region_ID
- Product_ID
- Sales_Channel_ID
- Date_ID
- Setter_ID
- Customer_ID
- Fact

Facts table

Ejemplo: Analizamos las ventas

- Por Tienda
- Por Vendedor
- Por Producto
- Por Semana
- Por Departamento
- Por Marca
- Por Mes
- Por Territorio de Ventas
- Por País

Tabla de Dimensiones



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

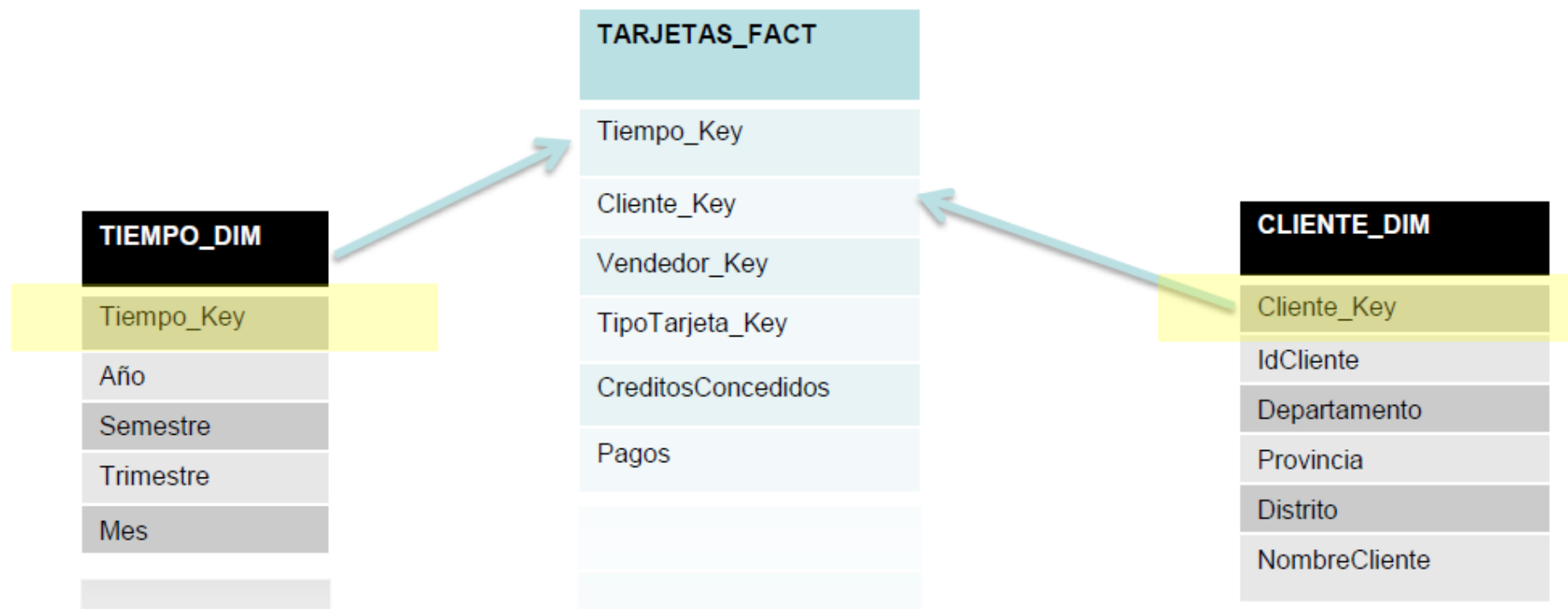
Product Key	Product Code	Product Name	Product Subcategory	Color	Status
218	HL-U509	Sport-100 Helmet, Black	31	Black	Current
219	SO-B909-M	Mountain Bike Socks, M	23	White	NULL
220	SO-B909-L	Mountain Bike Socks, L	23	White	NULL
221	HL-U509-B	Sport-100 Helmet, Blue	31	Blue	NULL

Componentes



Dimensiones: *Regular Dimensión*

Un foreign key existe entre la tabla de Hecho y la tabla dimensional



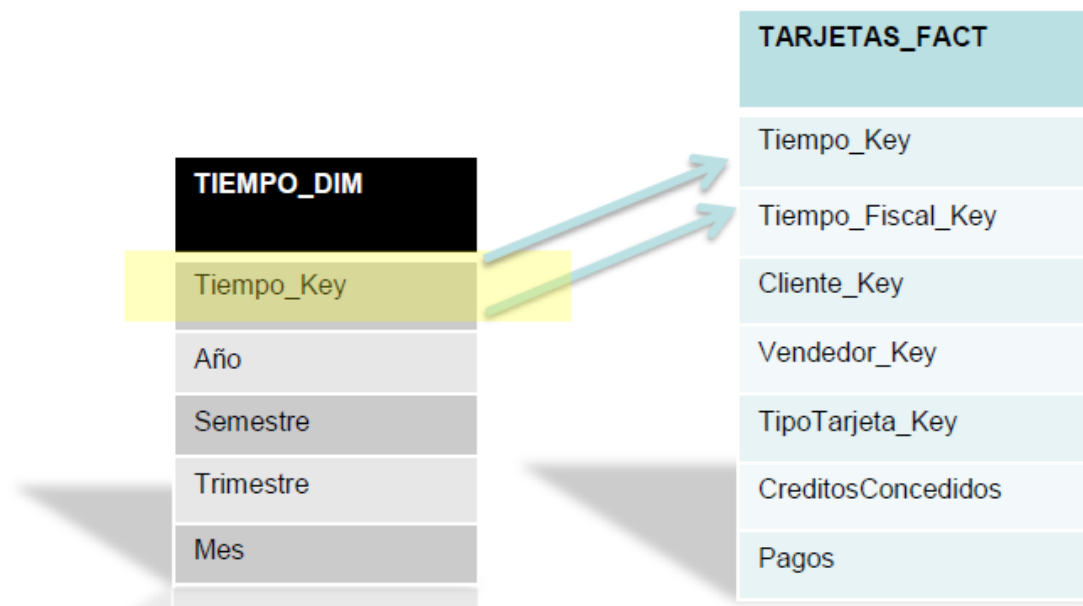
Componentes



Dimensiones:

Role-Playing Dimensión

- Una dimension puede ser usada multiples veces sobre el mismo cubo/measure group.
- Se ve dos dimensiones distintas :
 - Tiempo
 - TiempoFiscal

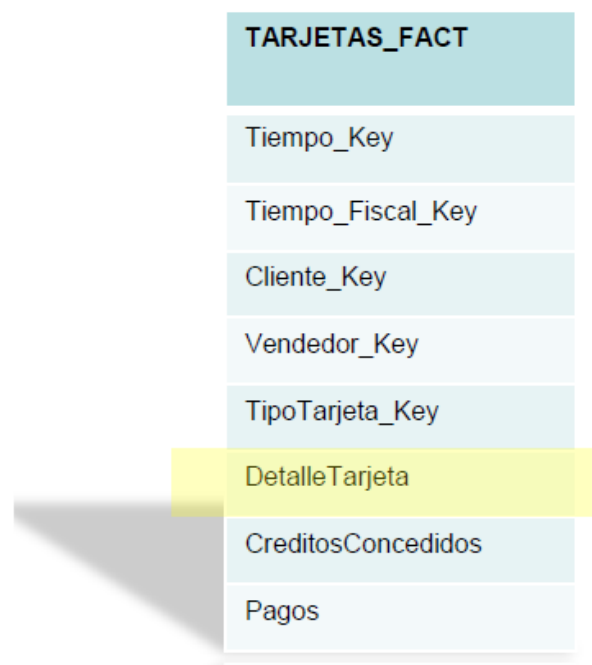


Componentes



Dimensiones: *Fact-Dimensión*

- Una dimensión basada sobre un atributo de una tabla de hechos.



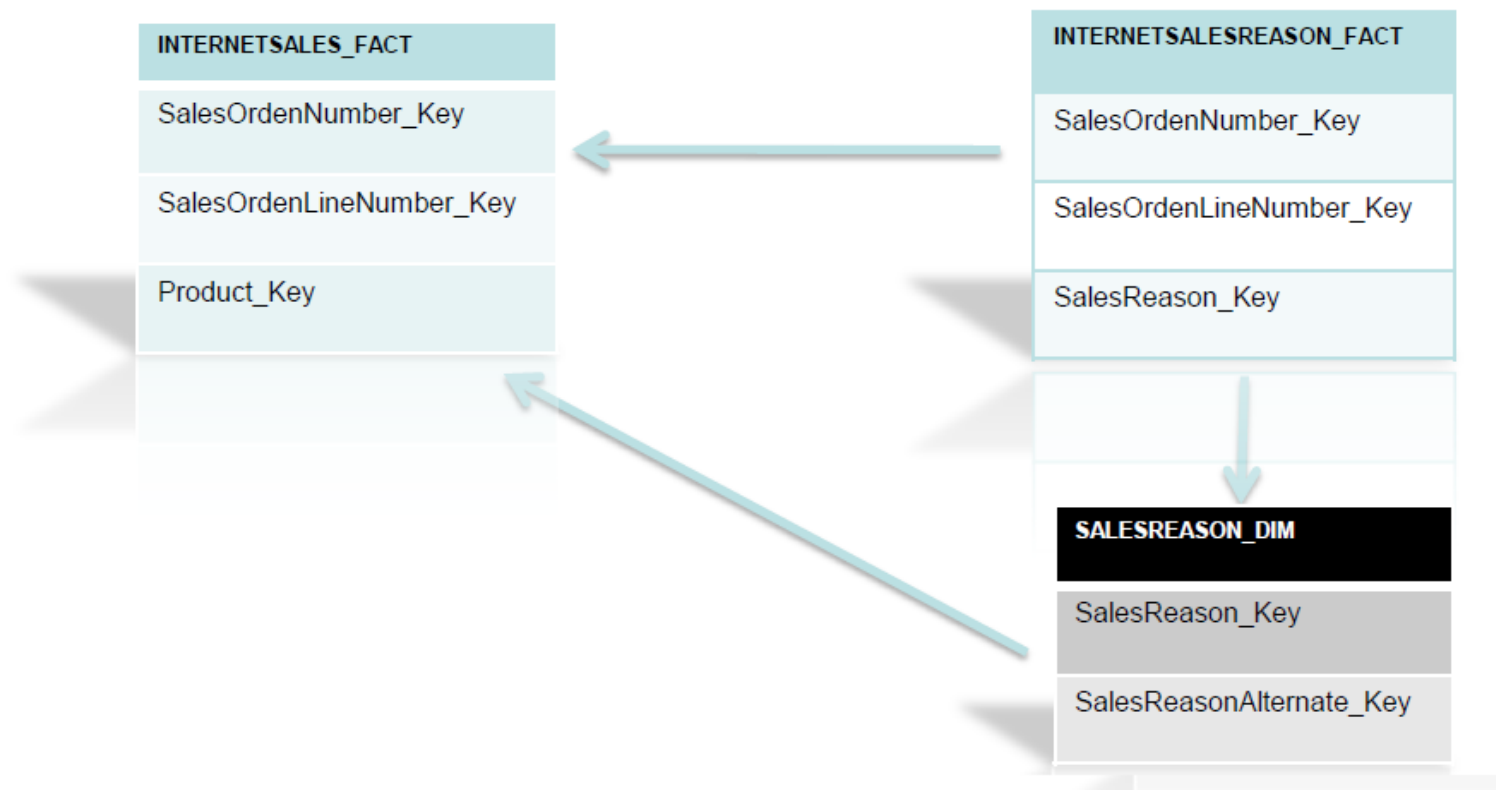
- Tambien es conocida como dimensión degenerada

Componentes



Dimensiones:
Many to many

- Utiliza un grupo de medidas intermedia

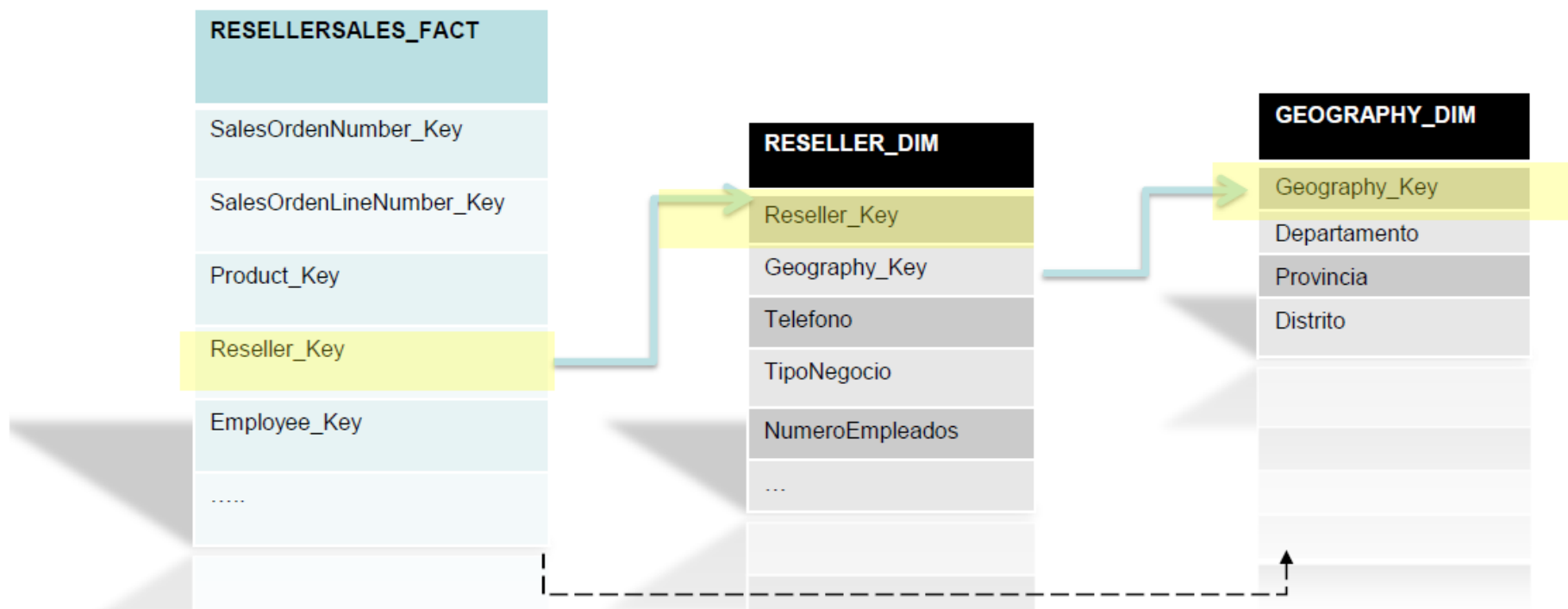


Componentes



Dimensiones: *Referenciada*

- Es una dimensión que no esta directamente relacionada con la tabla de hechos





Atributos:

- Los atributos son una agrupación de elementos o items 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.
- La finalidad de los atributos es ver la información de cada dimensión a diferentes niveles de detalle y agrupar los datos para ser analizados.



Relaciones:

- Los atributos dentro de una dimensión están directamente relacionados uno con otros a través de los diferentes tipos de relaciones antes definidos.
- 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.
- Son asociaciones lógicas de atributos dentro de una jerarquía definida por las instancias de los atributos y transitivas dentro de una jerarquía:
 - Uno-a-uno (1:1)
 - Uno-a-muchos (1:M)
 - Muchos-a-uno (M:1)
 - Muchos-a-muchos (M:N)



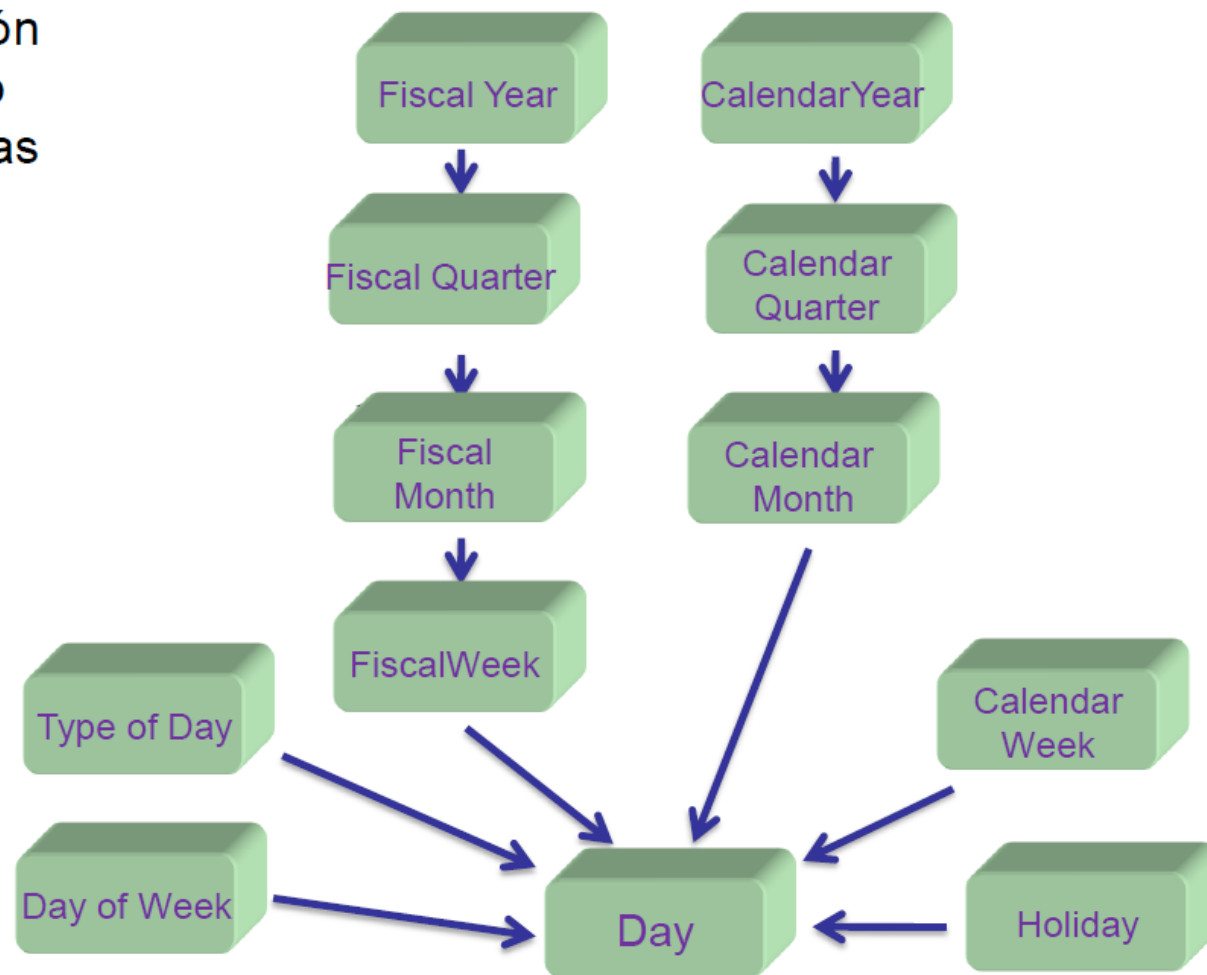
Jerarquías:

- 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.
- 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.
- También se conocen como los caminos de “drill-down” o “roll-up”

Componentes



Dimensión
Tiempo
Jerarquías



Componentes



Ejemplo:

Empresa

Cadena de supermercados

Actividad objetivo de análisis

Venta de productos

Información registrada sobre una venta

“Se vendió **5 unidades** del **producto A** en el **almacén** numero 1 el **día** 20/01/2010 por un **total** de 250.000 solares”

Para el análisis, interesa las ventas por productos en los distintos almacenes de cadena



Dimensiones

- Producto
- Almacén
- Tiempo

Medidas

- Importe
- Unidades

Solución



Dimensiones: desde los que se pueden analizar la actividad

TIEMPO_DIM
Tiempo_Key
Año
Semestre
Trimestre
Mes
Semana
Día

Ventas
Tiempo_Key
Producto_Key
Almacen_Key
Importe
Unidades

PRODUCTO_DIM
Producto_Key
NroProducto
Tipo
Departamento
Categoria
Descripcion
Marca

ALMACEN_DIM
Almacen_Key
Ciudad
Region
Tipo
Almacen

Actividad que es objeto de análisis con indicadores que intenta analizar


```

graph TD
    Sales((Sales)) ---|Region_ID| Region[Region]
    Sales ---|Product_ID| Product[Product]
    Sales ---|Sales channel_ID| SalesChannel[Sales channel]
    Sales ---|Date_ID| Date[Date]
    Sales ---|Seller_ID| Seller[Seller]
    Sales ---|Customer_ID| Customer[Customer]

```



Caso de Estudio 1:

Identificar Medidas y Dimensiones

Cadena de Supermercados

Una cadena de supermercados, desea analizar las ventas de sus productos. Cada Producto pertenece a una Sub-Familia y cada Sub-Familia pertenece a una Familia.

El Administrador desea entender las compras de los clientes desde el Punto de venta También se desea analizar las ventas por almacén.

Cada Cliente pertenece a una Zona, y cada zona pertenece a un distrito, cada distrito a una provincia y cada provincia a un departamento.

Se desea hacer comparativo de ventas por día de semana (Lunes, Martes, Miércoles, Jueves, Viernes, Sábado, Domingo).

También se desea visualizarse como totalizados anuales, semestrales, trimestrales y mensuales.