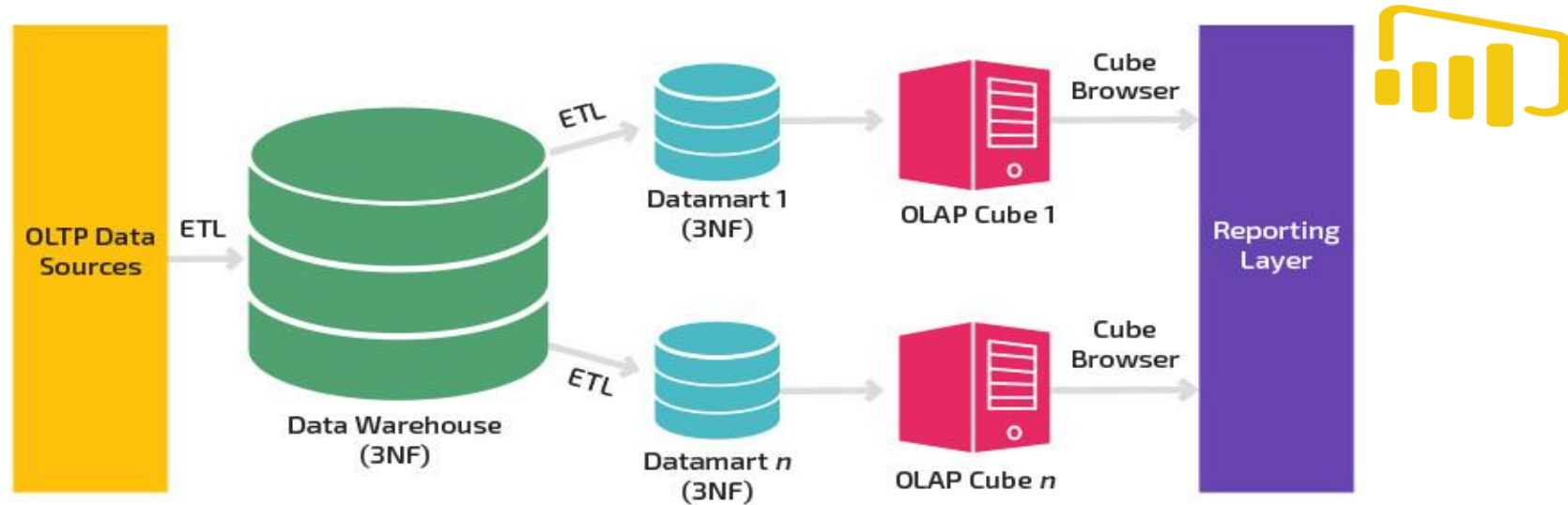


MINERÍA DE DATOS Y DATA WAREHOUSE

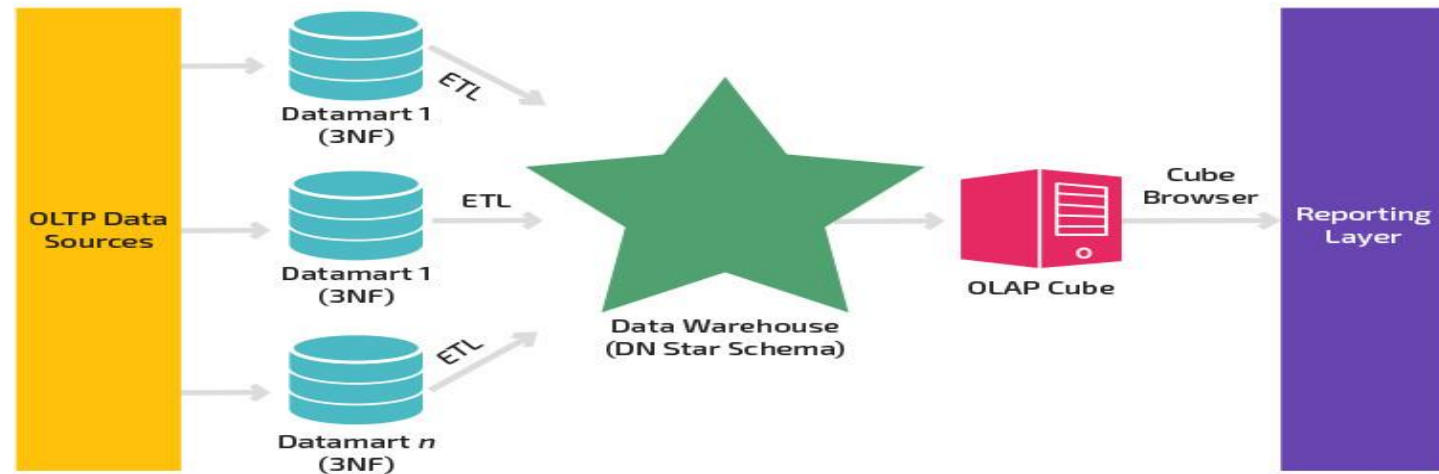
Profesor: Pablo Argeñal

1. Arquitectura de Data Warehouse

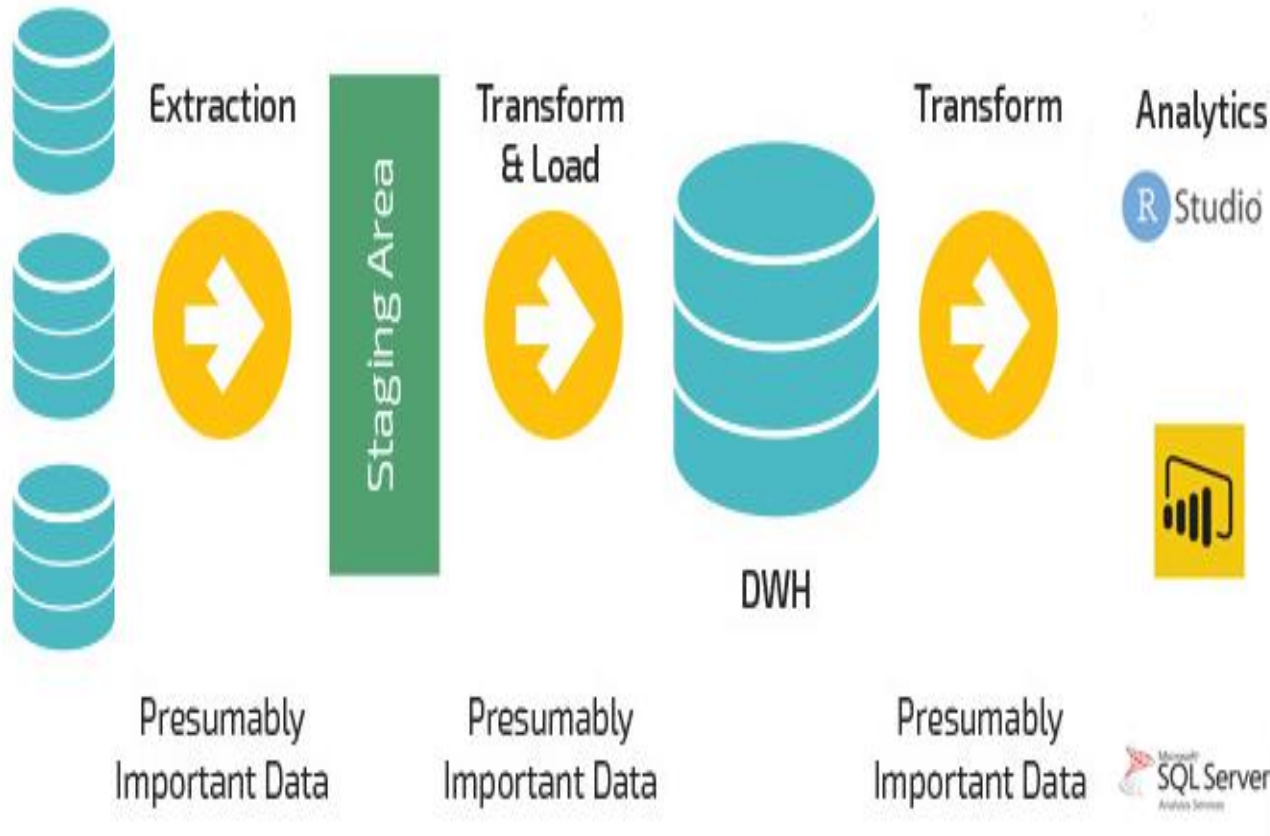
Inmon Model



Kimball Model



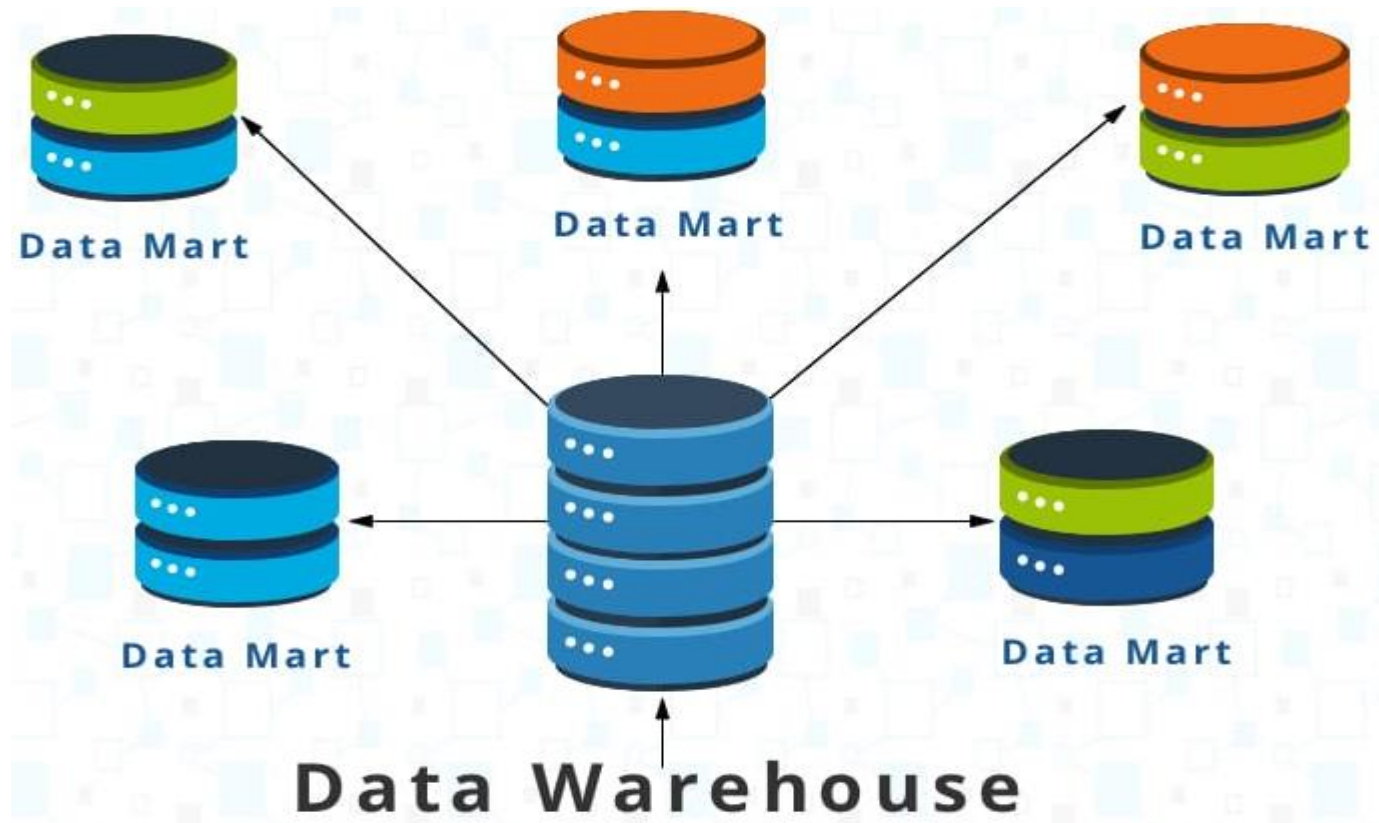
1.2.Principales componentes de datos



- **Datamart**

- **Data Warehouse**

1.2.Principales componentes de datos



1.3.Modelado

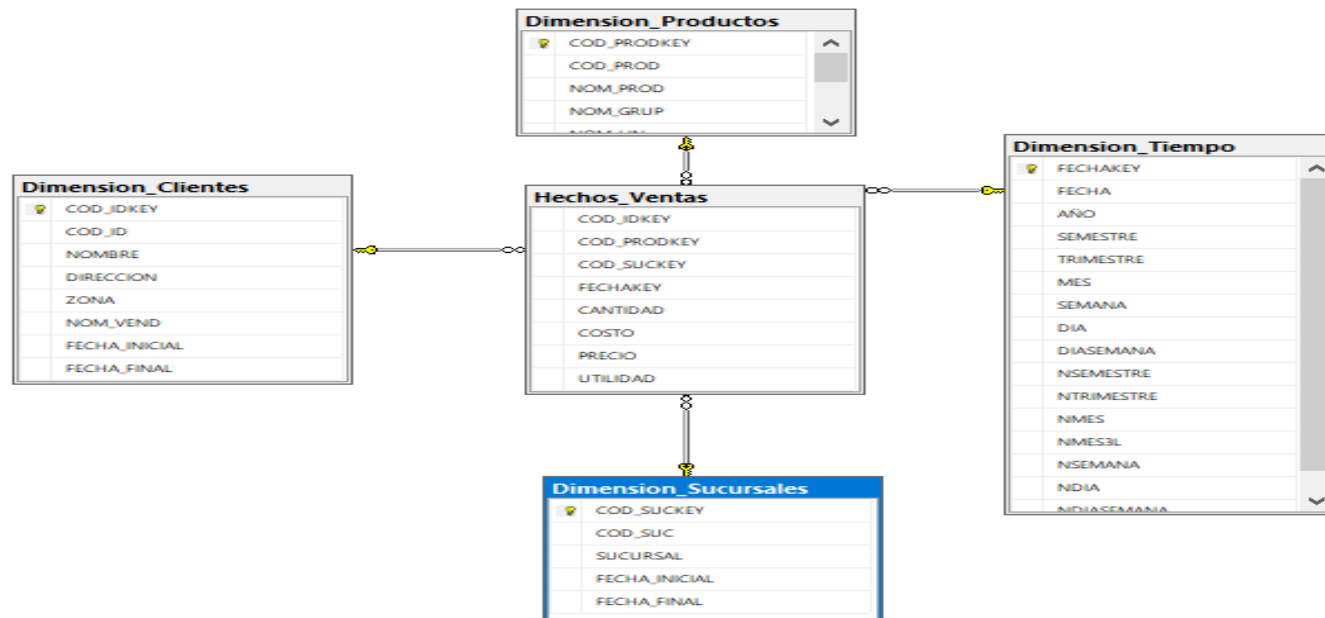
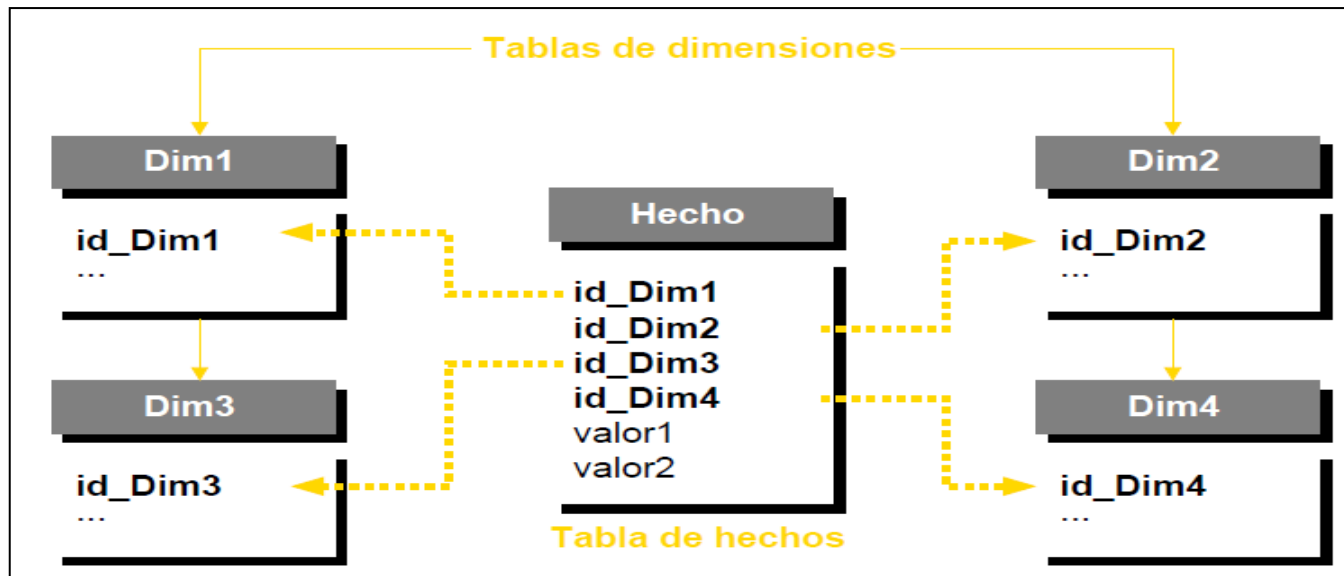


Esquema en estrella: formado por una tabla de hechos con una única tabla para cada dimensión.

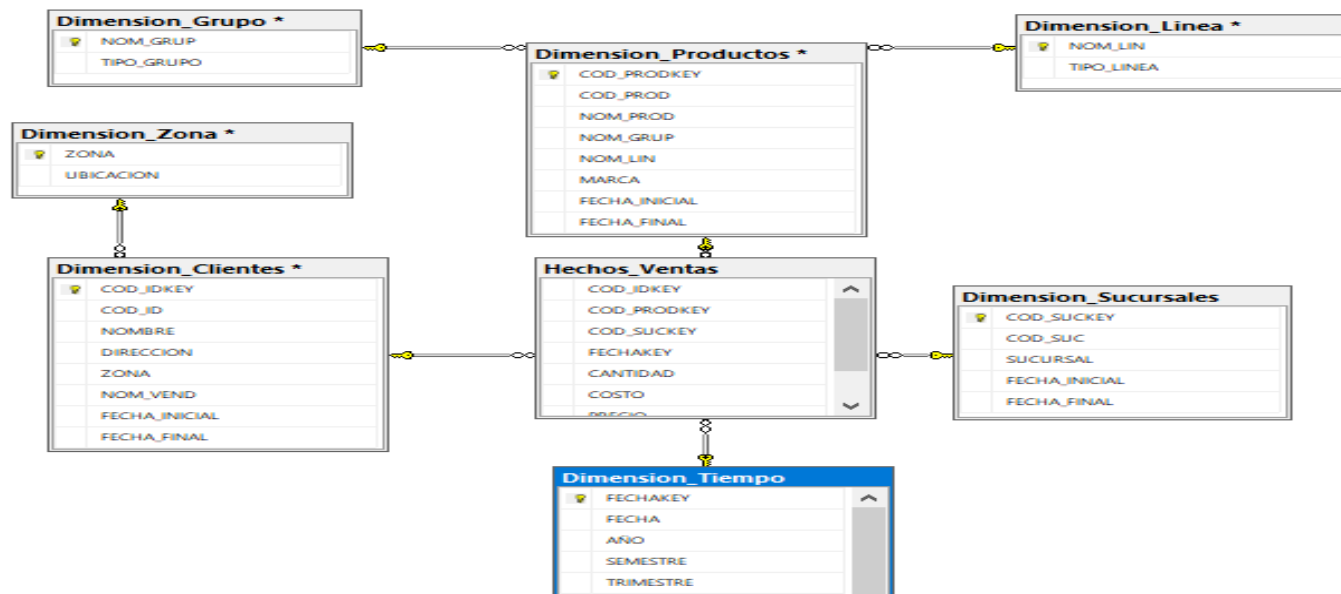
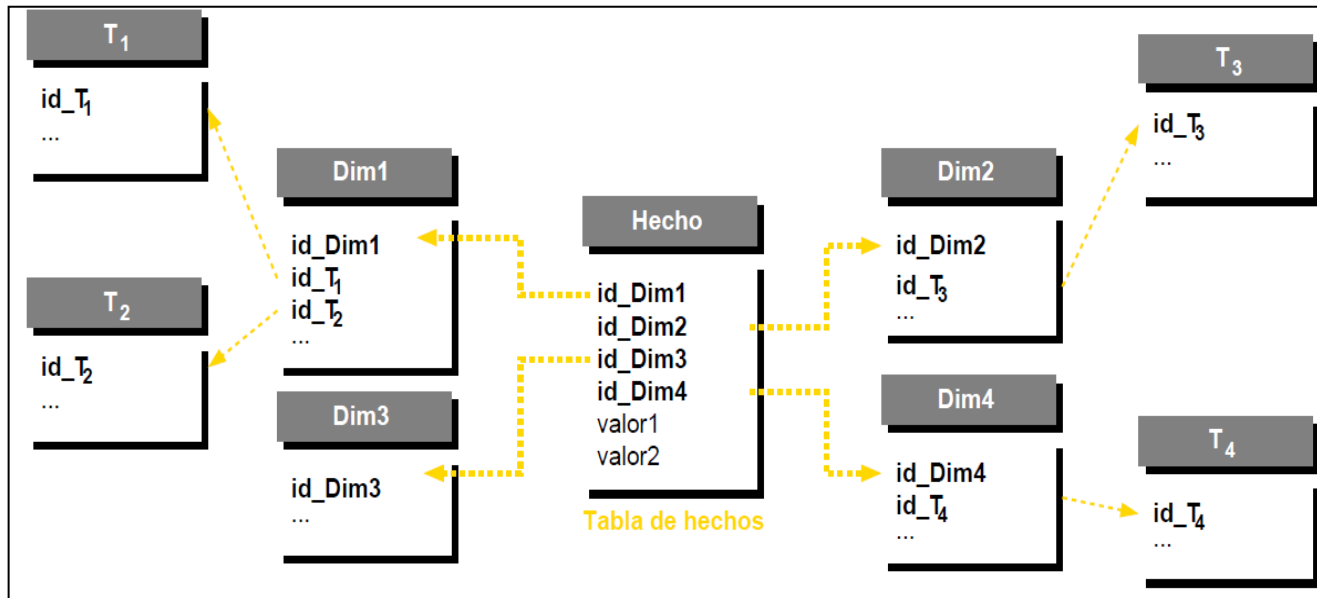
Esquema en copos: es una variante del esquema de estrella en el que las tablas dimensionales de este último se organizan jerárquicamente mediante su normalización.

Constelación de hechos: es un conjunto de tablas de hechos que comparten algunas tablas de dimensiones.

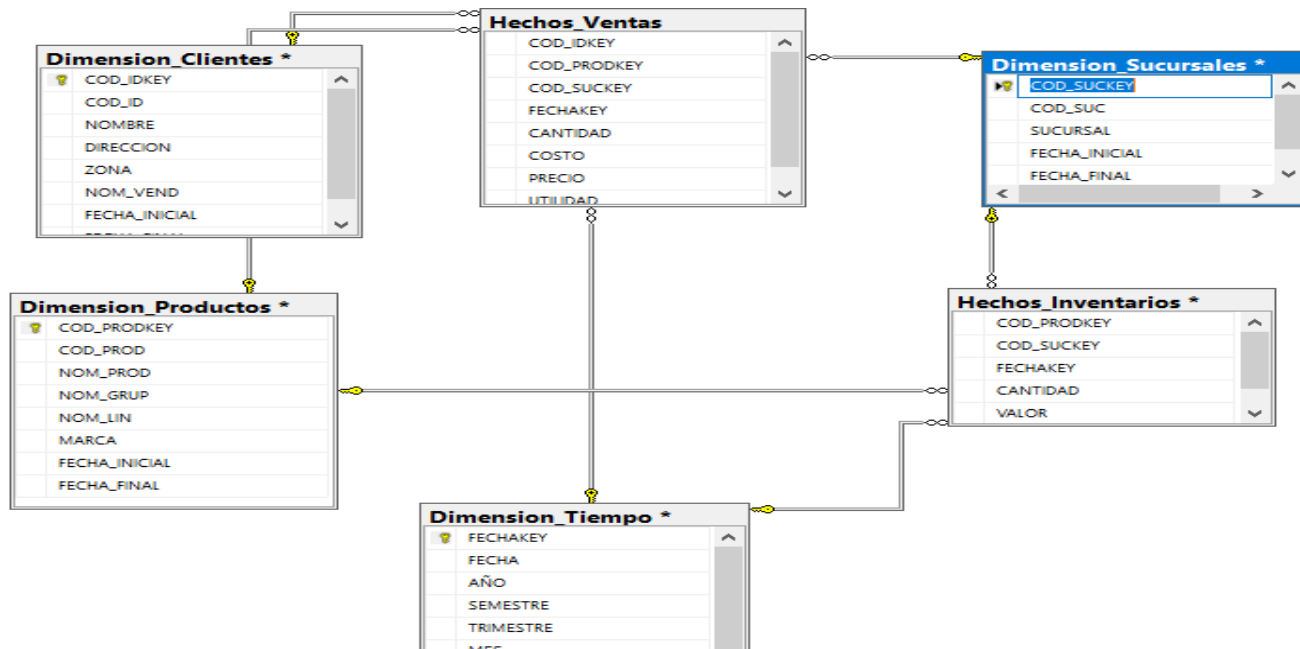
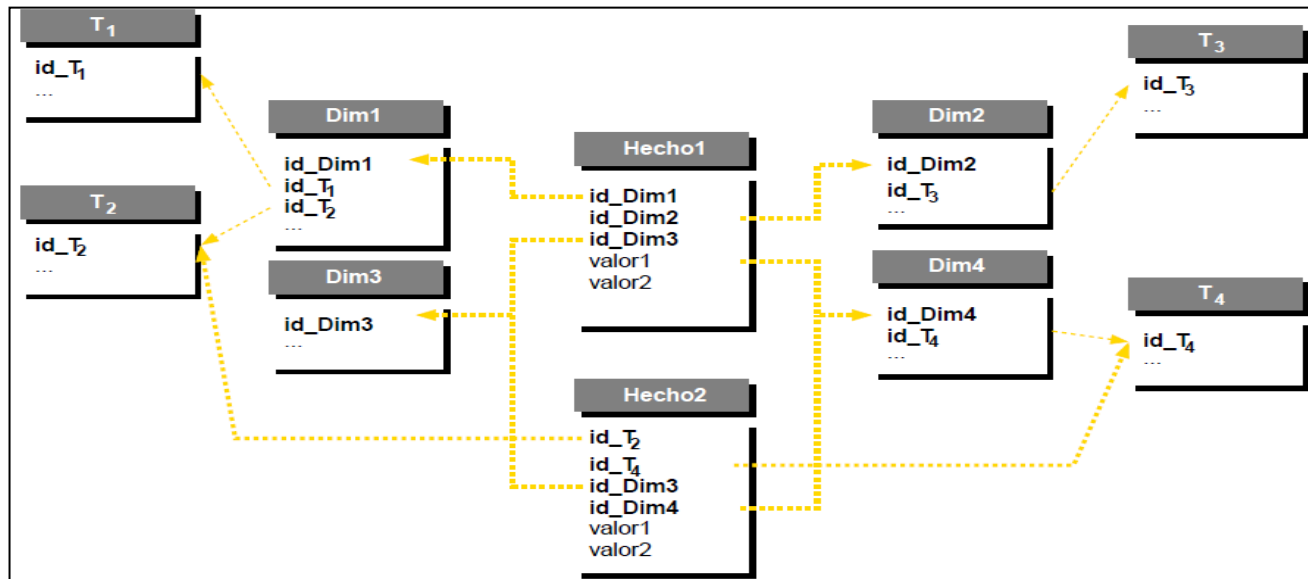
1.4.ESQUEMA DE ESTRELLA



1.5. Esquema de copos



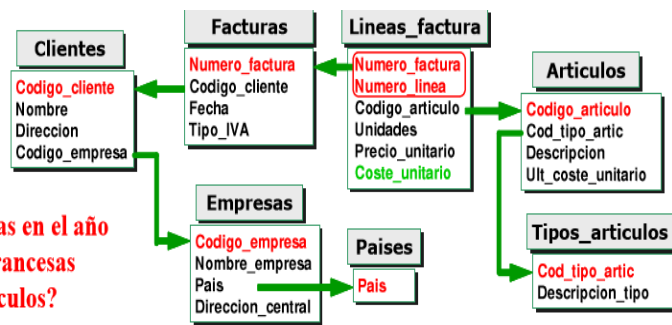
1.6.Constelación de hechos



1.7.ESQUEMA DE ESTRELLA A FONDO



Sistema Transaccional



¿Beneficio de ventas en el año
2004 a empresas francesas
según tipos de artículos?

```

SELECT SUM(Unidades*(Precio_unitario - Coste_unitario)), Cod_tipo_artic
FROM Empresas INNER JOIN
  (Clientes INNER JOIN
    (Facturas INNER JOIN
      (Lineas_factura INNER JOIN Articulos
        ON Lineas_factura.Codigo_articulo=Articulos.Codigo_articulo)
      ON Facturas.Numero_factura=Lineas_factura.Numero_factura)
    ON Clientes.Codigo_cliente=Factura.Codigo_cliente)
  ON Empresas.Codigo_empresa=Clientes.Código_empresa
WHERE Fecha BETWEEN '1/1/2004' AND '31/12/2004' AND Pais='Francia'
GROUP BY Cod_tipo_artic
    
```

Modelo de Estrella

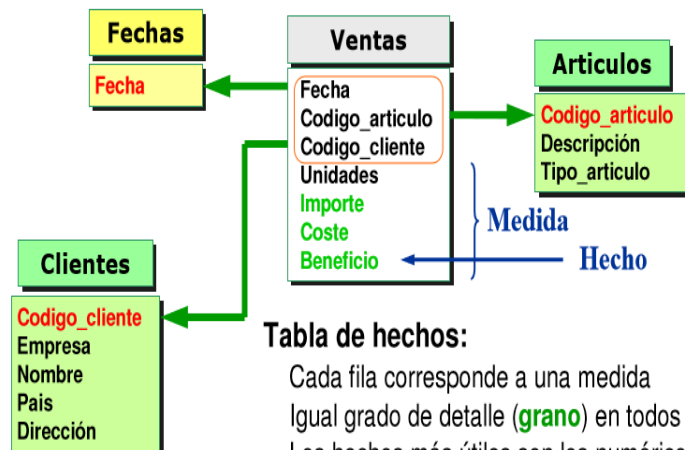


Tabla de hechos:

Cada fila corresponde a una medida
Igual grado de detalle (**grano**) en todos los hechos
Los hechos más útiles son los numéricos y aditivos

Tablas de dimensión:

Contienen descriptores textuales
Son los puntos de entrada en la tabla de hechos

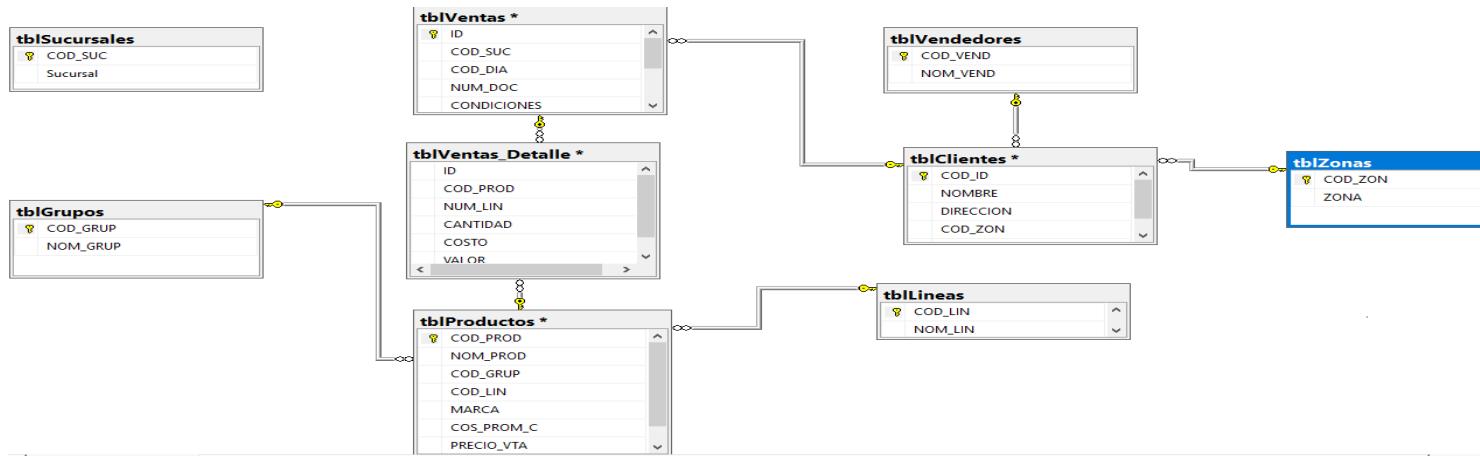
Transformación de datos

Desnormalización

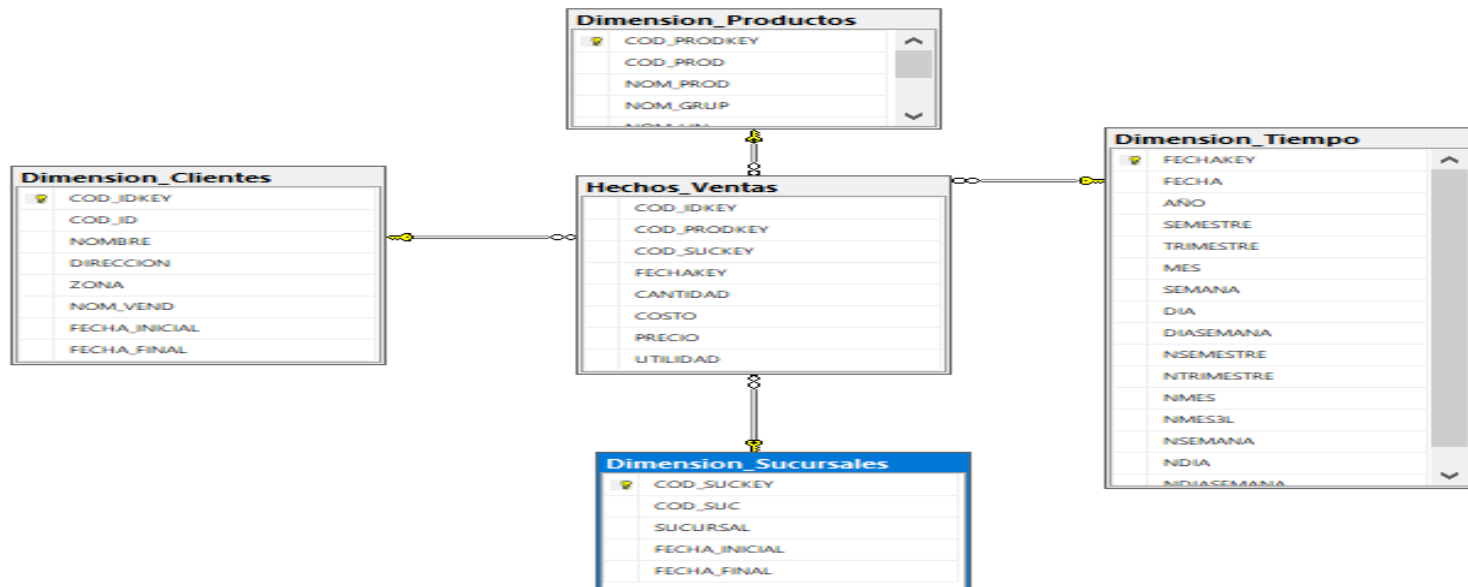
Aún no es el modelo dimensional

1.7.ESQUEMA DE ESTRELLA A FONDO

Sistema Transaccional



Modelo de Estrella



1.8.PREGUNTAS AL MODELO



Realizar las preguntas apropiadas al modelo según los campos de las dimensiones

Dimensiones	Hechos_Ventas (Unidades,Costos,Ventas,Utilidad)
Cientes	
Zona	Cuántas unidades se vendieron por zona?
Vendedor	Cuánta utilidad se generó por vendedor?
Productos	
Linea	Cuántas fueron mis ventas por línea de producto?
	Cuántas fueron mis ventas por sublínea de producto?
Tiempo	
Año	Comparar mis ventas vrs ventas año pasado?

1.9.DIMENSIONES LENTAMENTE CAMBIANTES (SCD 1 Y 2)



Una SCD Dimension son dimensiones en las que sus atributos cambian a lo largo del tiempo.

Dimensiones Tipo 1: Se sobrescriben los valores

Es el tipo de carga más básico y el más habitual. En este caso, se sobrescriben los valores existentes en el maestro con los nuevos valores.

COD_SUCKEY	COD_SUC	SUCURSAL
1	01	SUCURSAL 01
COD_SUCKEY	COD_SUC	SUCURSAL
1	01	CENTRO DE DISTRIBUCION

1.10.DIMENSIONES LENTAMENTE CAMBIANTES (SCD 1 Y 2)



Dimensiones Tipo 2: Se guarda la historia de cambios

En los casos que sí que nos interesa guardar la historia se utiliza la estrategia conocida como "SCD Tipo 2". En este caso, la tabla de dimensión incluye los campos de *fecha inicial* y *fecha final*. Estas fechas nos permiten determinar en que estado estaba la dimensión en cualquier fecha del calendario, el registro actual se representa por la fecha final con el valor de 31/12/999,

COD_SUCKEY	COD_SUC	SUCURSAL	FECHA_INICIAL	FECHA_FINAL
1	01	SUCURSAL 01	22/07/2020	31/12/9999
1	01	SUCURSAL DISTRIBUCION	22/07/2020	25/07/2020
2	01	CENTRO DE DISTRIBUCION	25/07/2020	31/12/9999

1.11.ETL & BI



El Business Intelligence trata de extraer los datos de la empresa a través de distintas fuentes mediante herramientas o técnicas llamadas ETL (extraer, transformar y cargar), luego de los datos ser pasados por este proceso se continúa con una carga de almacén de datos.

Finalmente, todo este análisis deberá permitir incrementar el nivel administrativo, financiero y la toma decisiones en cuanto a las acciones de la empresa.

1.12. una solución de BI efectiva permite



Observar: *¿Qué está pasando?*

Comprender: *¿Por qué ocurre?*

Predecir: *¿Qué ocurriría?*

Colaborar: *¿Qué debe hacer el equipo?*

Decidir: *¿Qué camino debe seguir la empresa?*