

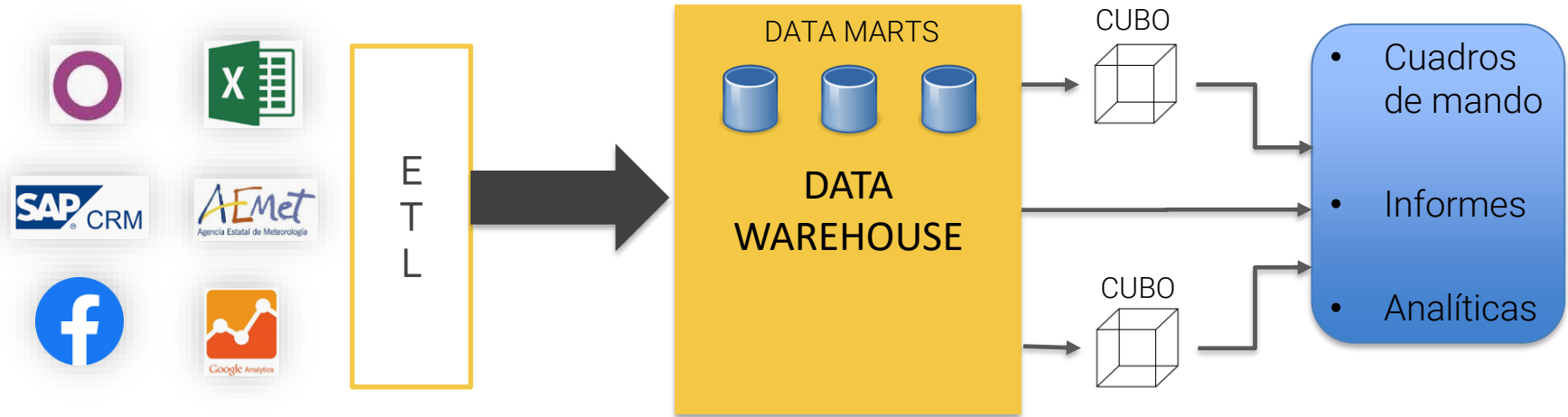


1. Arquitectura Data Warehouse
2. ETL
3. BD Data Warehouse
4. Cubos

Arquitectura Data Warehouse

Arquitectura Data Warehouse

1. Arquitectura Data Warehouse



ETL



- Extraer datos de fuentes heterogéneas y distintos formatos
- Aplicar transformaciones para facilitar análisis
- Cargar en base de datos Data Warehouse

ETL (Extract-Transform-Load)

- Extract
- Transform
- Load

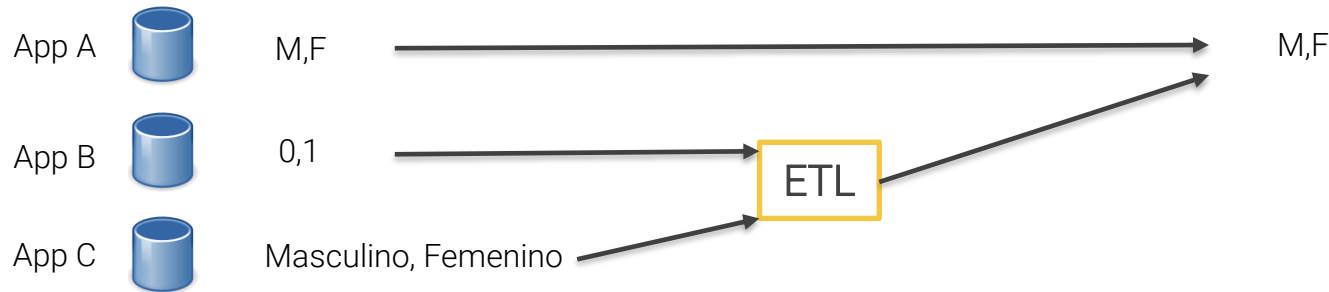
ETL (Extract-Transform-Load)

- **Extract:** obtención de información de las distintas fuentes de datos.
- Transform
- Load



ETL (Extract-Transform-Load)

- **Extract:** obtención de información de las distintas fuentes de datos.
- **Transform:** filtrado, limpieza, depuración y homogeneización y agrupación de la información.



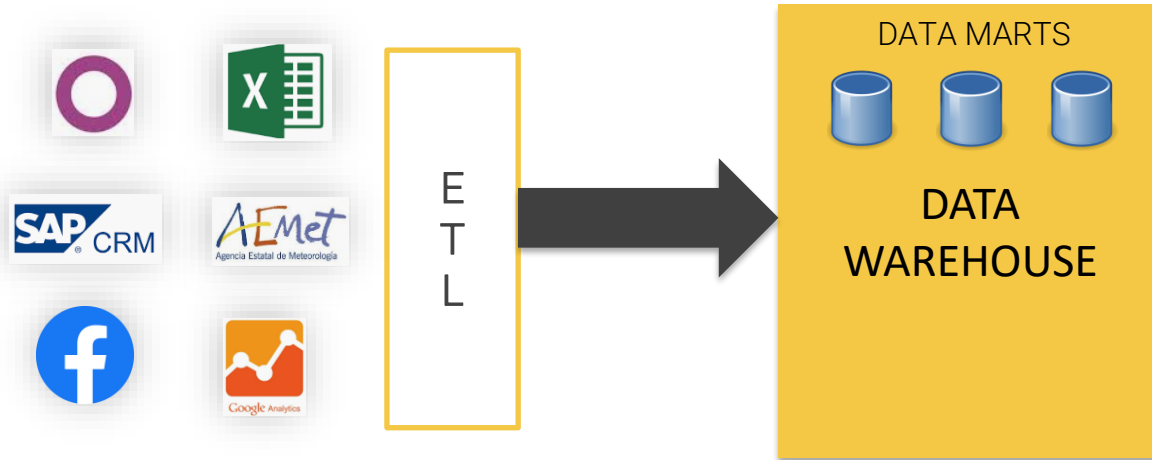
ETL (Extract-Transform-Load)

- **Extract:** obtención de información de las distintas fuentes de datos.
- **Transform:** filtrado, limpieza, depuración y homogeneización y agrupación de la información.
- **Load:** organización y actualización de los datos y metadatos en la BD Data Warehouse.



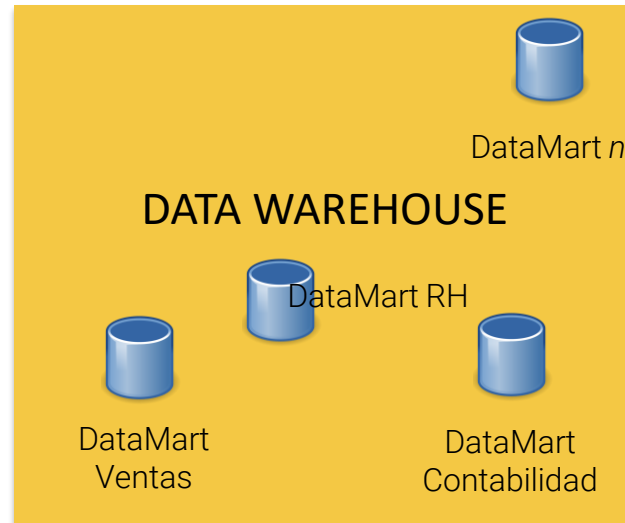
BD Data Warehouse

3. BD Data Warehouse



Data Marts

- Un DataMart es un subconjunto de datos centrados en un tema o una área de negocio específica.



Elementos que componen una BD DataWarehouse y DataMart

- Hechos
- Dimensiones
- Métricas

Elementos que componen una BD DataWarehouse y DataMart

- Hechos
- Dimensiones
- Métricas



Elementos que componen una BD DataWarehouse y DataMart

- **Hechos:** proceso de negocio de la organización. Ejemplo: Ventas (número ventas e importe)
- Dimensiones
- Métricas

Elementos que componen una BD DataWarehouse y DataMart

- Hechos: proceso de negocio de la organización. Ejemplo: Ventas (número ventas e importe)
- **Dimensiones:** distintas vistas para un proceso de negocio. Ejemplo: cliente, producto, tiempo

Elementos que componen una BD DataWarehouse y DataMart

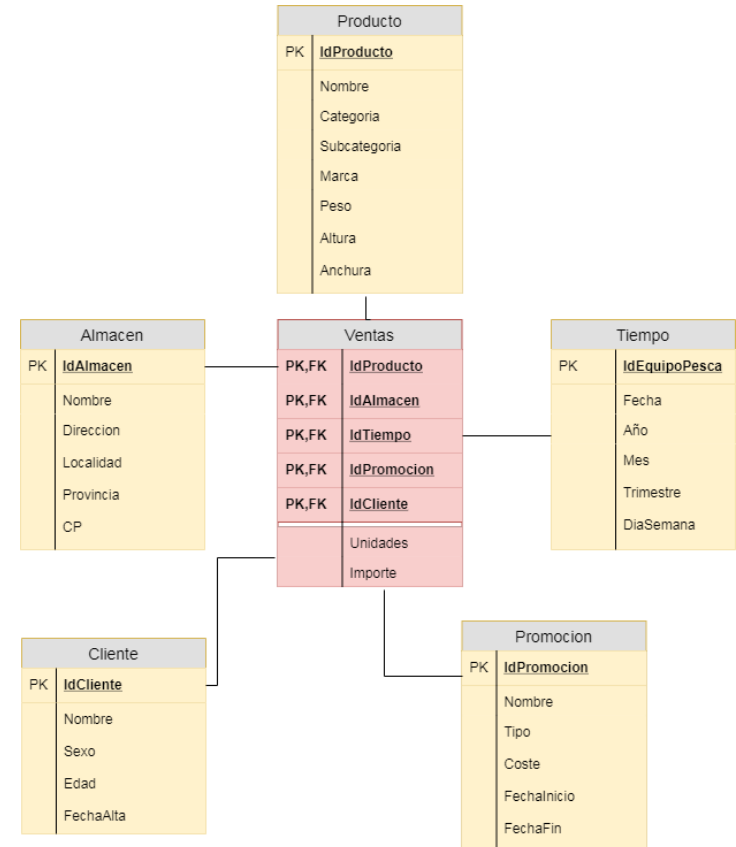
- Hechos: proceso de negocio de la organización. Ejemplo: Ventas (número de ventas e importe)
- Dimensiones: distintas vistas para un proceso de negocio. Ejemplo: cliente, producto, tiempo
- **Métricas:** indicadores cuantificables para un hecho. Ejemplo: suma, media, máx., min...

Modelado de una BD DataWarehouse y DataMart

- Modelo de datos en estructura de **estrella**
- Modelo de datos en estructura de **copo de nieve**

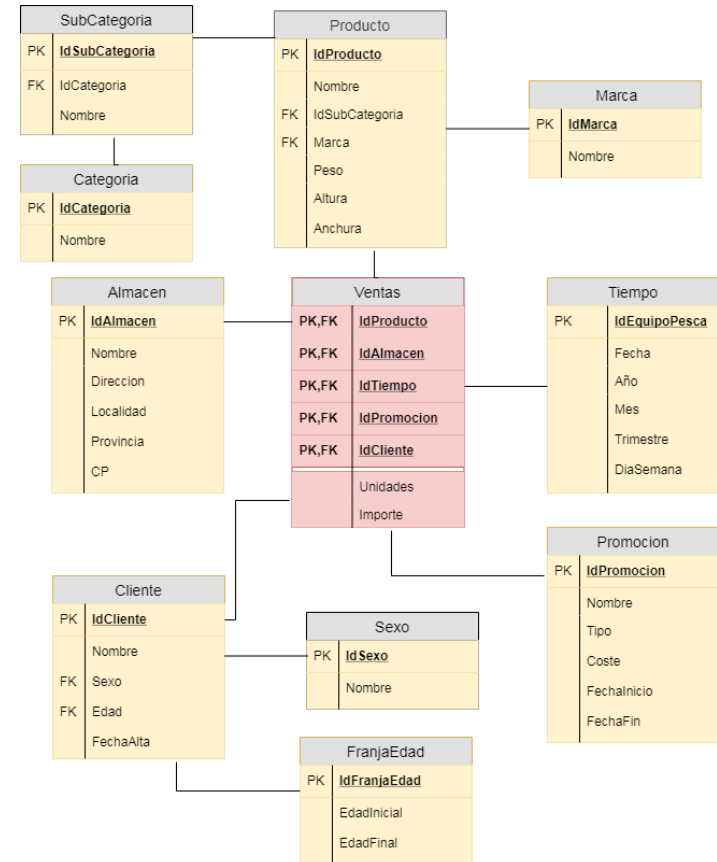
Modelo en estrella

- Una tabla de Hechos (Ventas).
- N tablas de dimensiones (Producto, Cliente, etc).
- Una dimensión se debe de implementar en una única tabla.
- Relaciones entre hechos con las dimensiones.



Modelo copo de nieve

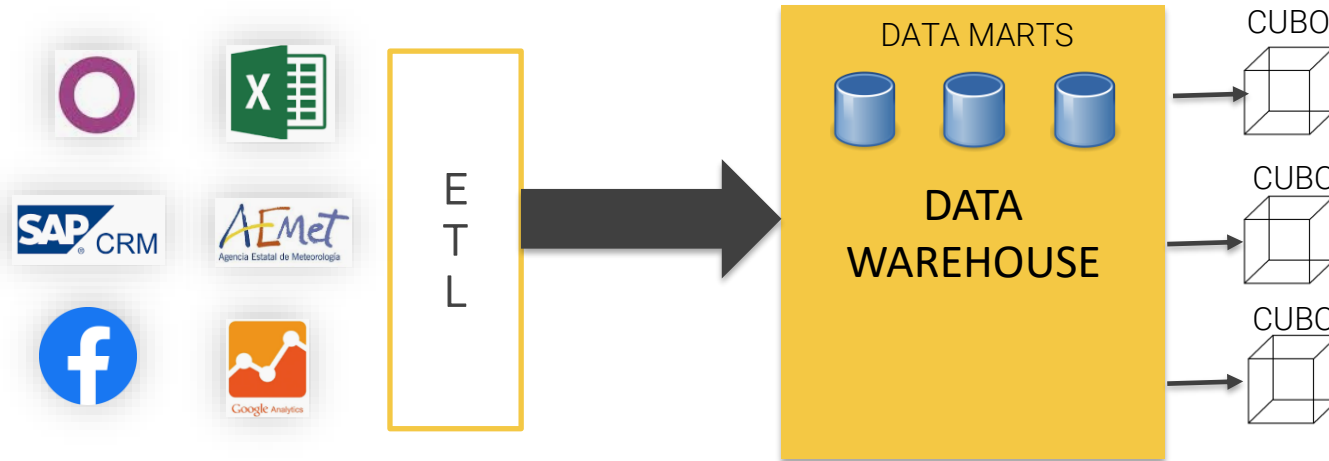
- Una tabla de Hechos (Ventas).
- N tablas de dimensiones (Producto, Cliente, etc).
- Una dimensión se puede implementar con más de una tabla.



Comparativa entre modelo estrella y copo de nieve

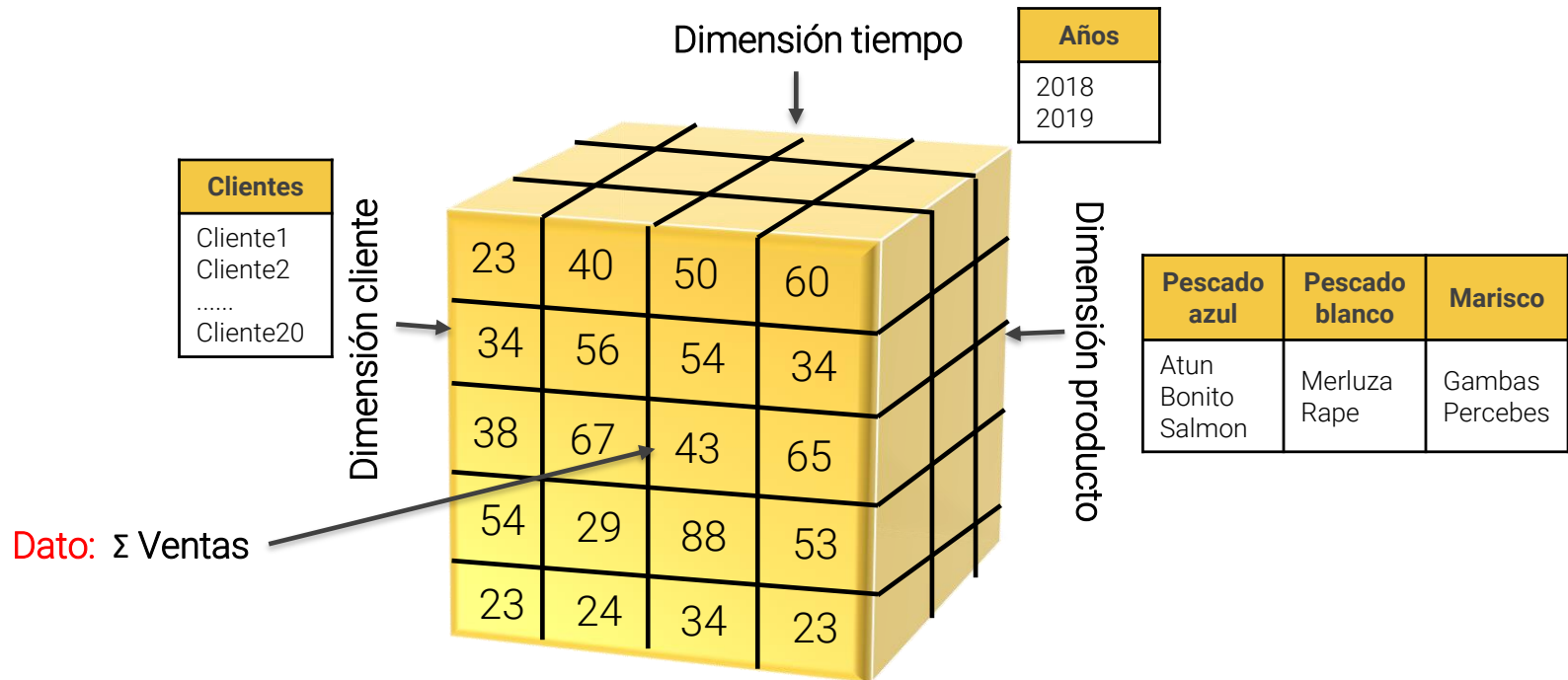
Modelo estrella	Modelo copo de nieve
Diseño más sencillo	Diseño más estructurado
Existe redundancia de datos	NO existe redundancia de datos
La extracción de los datos es más rápida	La extracción de los datos no es tan rápida
Diseño de base de datos muy desnormalizado con tablas horizontales (<i>muchas columnas en las tablas de hechos</i>)	Diseño de base de datos menos desnormalizado y más próximo a un diseño relacional clásico

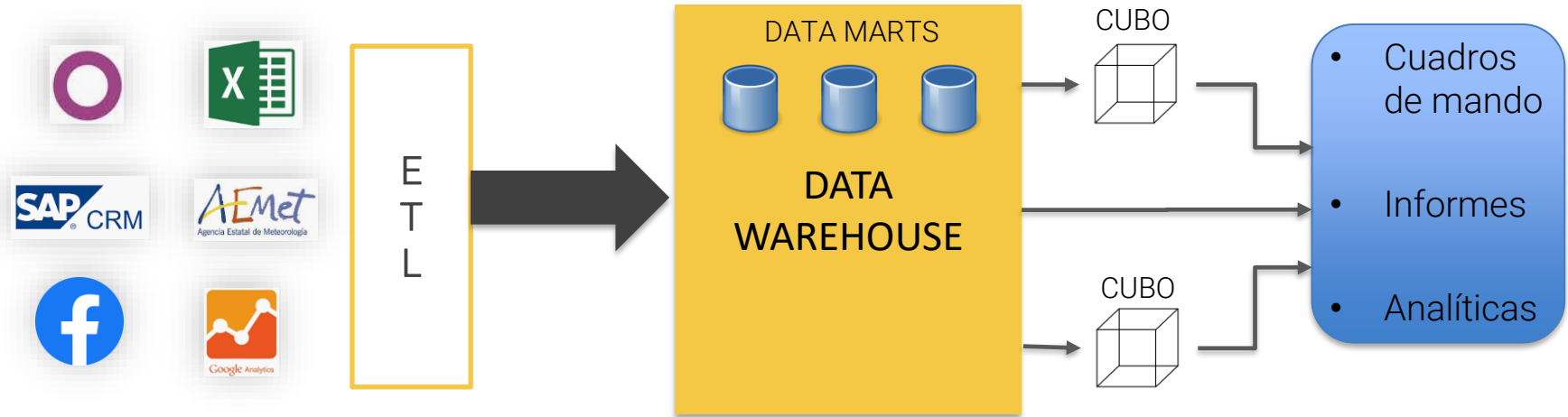
Cubos



4. Cubos

Subconjuntos de datos, organizados y sumados dentro de una estructura multidimensional.





Si puedes controlar la información, puedes
controlar a la gente.

Tom Clancy

