

Clase 3

Datamart

- Un datamart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica.

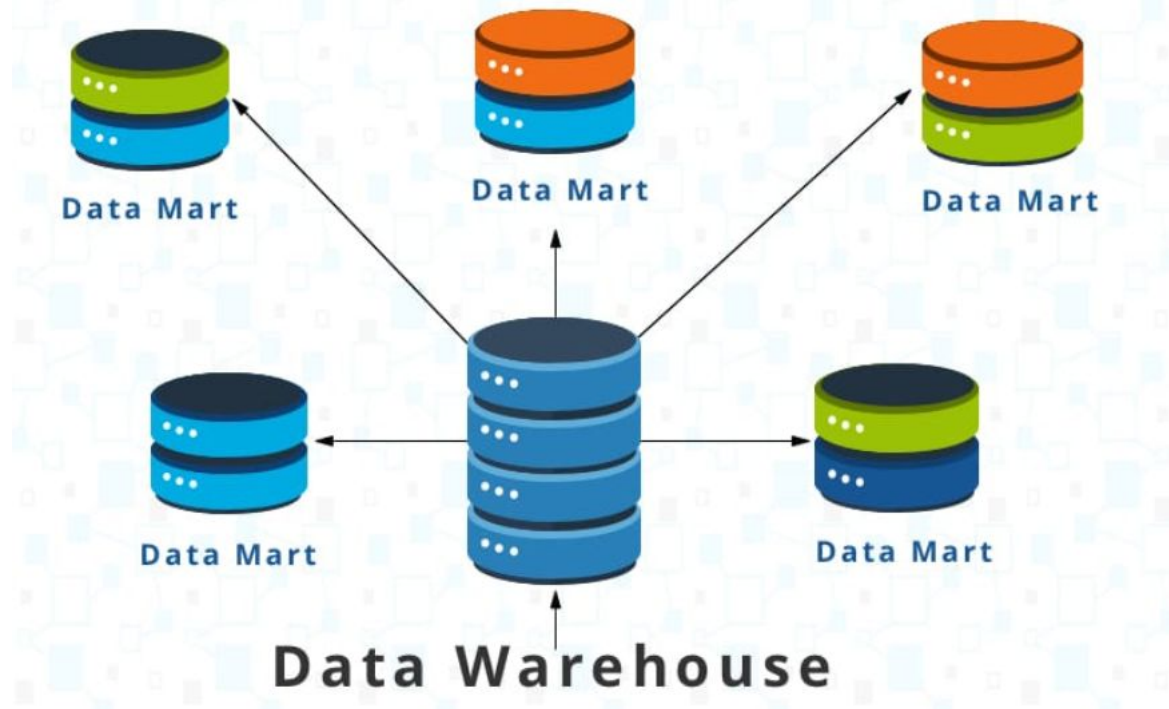
Datamart

- Se caracteriza por disponer la estructura óptima de datos para analizar la información a detalle desde todas las perspectivas que afecten a los procesos de dicho departamento.

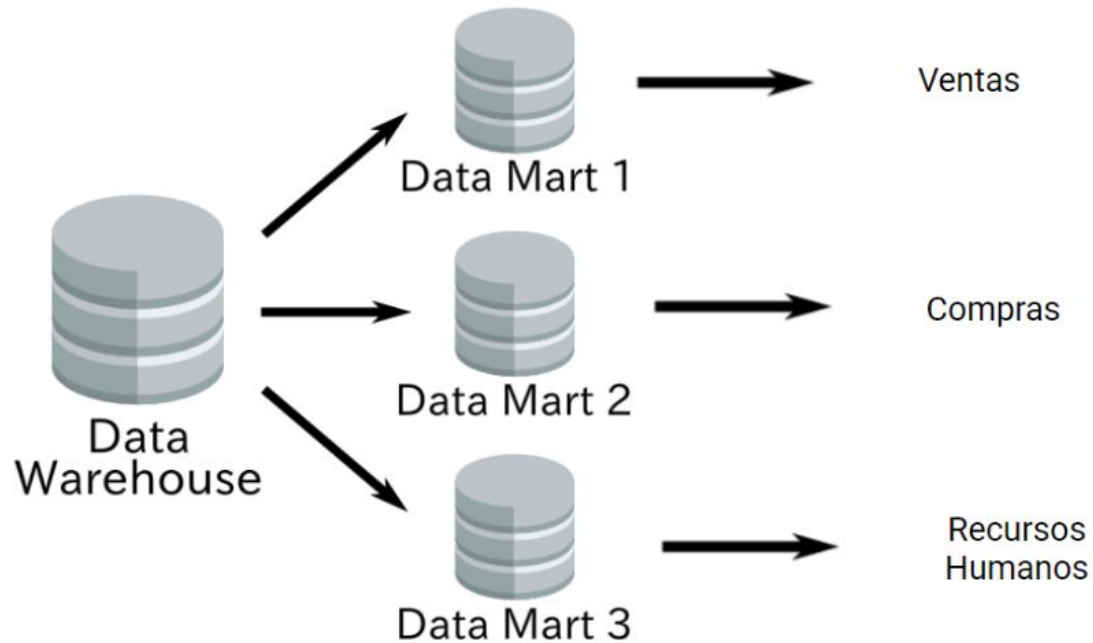
Datamart

- Los datamarts son subconjuntos de los datos del Datawarehouse con el propósito de ayudar a que un área específica dentro del negocio pueda tomar mejores decisiones.

Representación gráfica de un Datamart.



Representación gráfica de un Datamart.



Datawarehouse y Datamart

- Si nos referimos a un Datawarehouse estamos hablando que este contiene **todos** los datos de una organización.
- Mientras que el Datamart solamente obtiene un **subconjunto** de los datos de una organización, lo que hace centrar lo en un área específica dentro de la organización.

Datawarehouse y Datamart

- Un problema que surge es cuando el **datawarehouse** llega a crecer y a tornarse muy complejo. Debido a esto el rendimiento de las consultas decae y el modelo deja de ser óptimo.
- En estos casos la solución es la creación de **datamarts** especializados por áreas como Ventas, Compras, etc.

Ventajas y Desventajas del Datamart

VENTAJAS

- Consultas más rápidas debido al poco volumen de datos a recorrer.
- Fácil acceso a los datos que se utilizan con frecuencia.
- Su costo de construcción es relativamente menor a la de un datawarehouse.

DESVENTAJAS

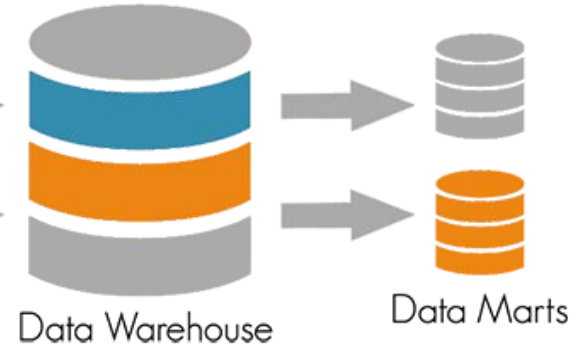
- No maneja grandes volúmenes de información.
- No considera otras fuentes de datos de la empresa.

Tipos de Datamart

- **Dependiente.**
- **Independiente.**
- **Híbridos.**

Tipos de Datamart

- **Dependiente:** se crea a partir de un datawarehouse existente.



Tipos de Datamart

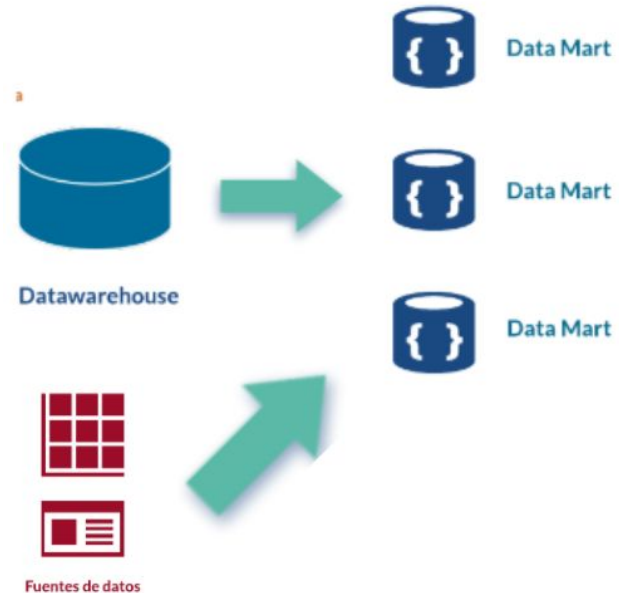
- **Independiente:** es un sistema autónomo que es creado sin utilizar ningun datawarehouse.

Este es conveniente para empresas pequeñas.



Tipos de Datamart

- **Híbridos:** en este tipo combina datos de un datawarehouse con otros sistemas de fuentes de datos.



Datawarehouse vs Datamart

	Datawarehouse	Datamart
Alcance	Almacena información relacionada con todo el sistema.	Se limita a almacenar información de un área de la organización en específico.
Tamaño e integración	Almacena grandes cantidades de datos provenientes de muchas fuentes de datos, por lo que suele ser más grande.	Se concentra en resúmenes de datos totalizados por lo que suele ser más pequeña.

Datawarehouse vs Datamart

	Datawarehouse	Datamart
Creación	La creación es más complicada ya que debe contemplar todos los datos del sistema.	La creación es más simple ya que tiene menos relaciones y están enfocados a sólo un tema.
Costo de manejo	Más costoso, porque requiere más recursos físicos para manejar grandes cantidades de datos.	Es menos costoso ya que requiere menos recurso físicos para manejar los datos requeridos.
Objetivo	Optimizar la obtención de datos, integrando y optimizando los datos fuente.	Es diseñado para entregar de manera óptima la información para el soporte de decisiones de negocio.

Modelos de datos.



Tipos de modelos

- **Modelo Estrella o Star Schema.**
- **Modelo Copo de Nieve o Snowflake Schema.**
- **Modelo Constelación(Copo de Estrellas) o Starflake Schema.**

Modelo Estrella

- Es el más sencillo en su estructura, consta con:
 - Una tabla central de Hechos.
 - Varias tablas de dimensión.
- Lo característico de este modelo es que la única tabla que tiene relación con otras tablas es la de hecho.

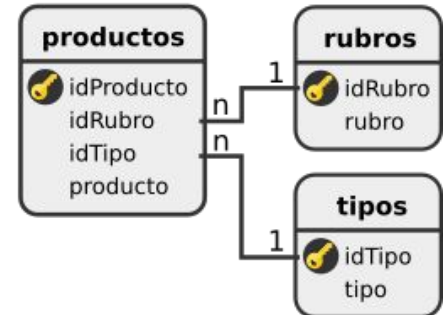
Modelo Estrella

- Las tablas de dimensión sólo están relacionadas con la tabla de hechos.
- Las tablas de dimensión se encuentran desnormalizadas.

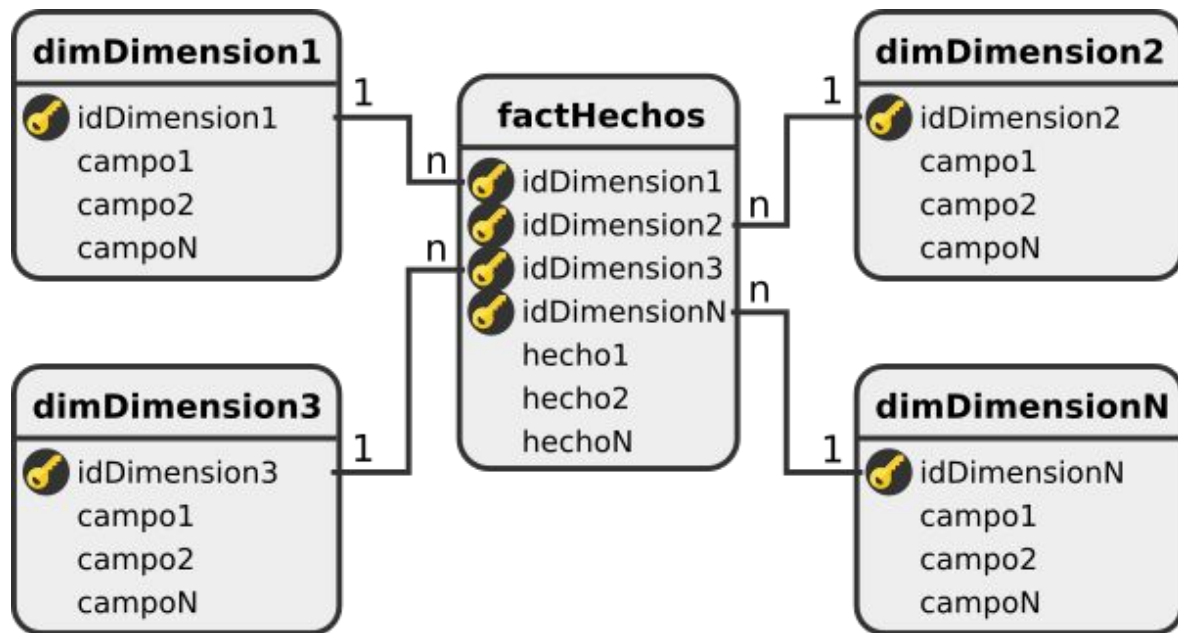
Desnormalizado



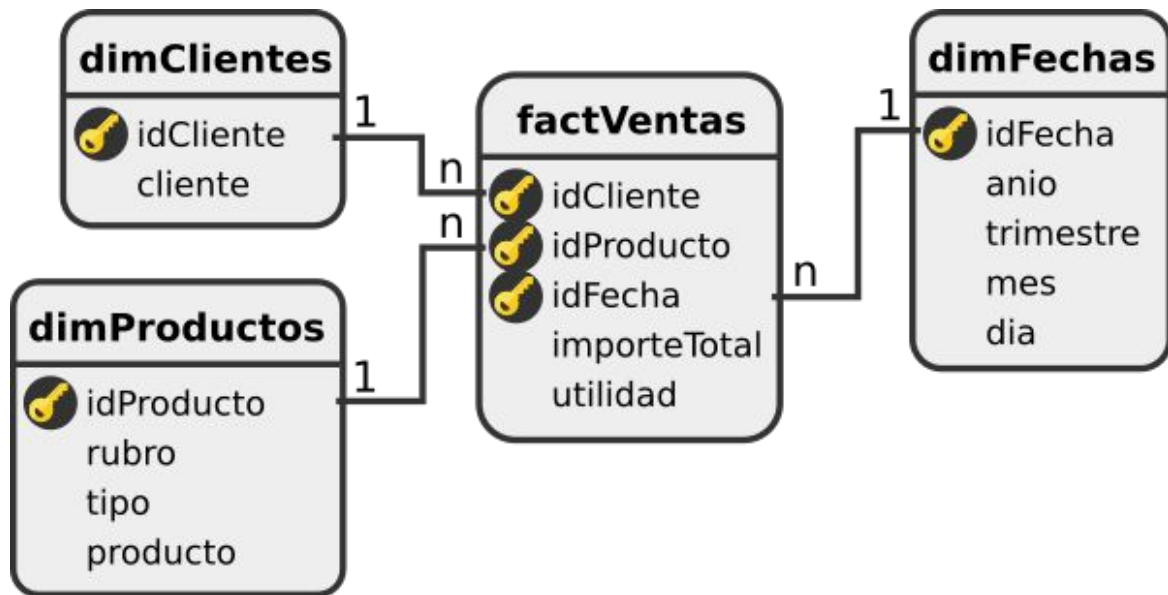
Normalizado



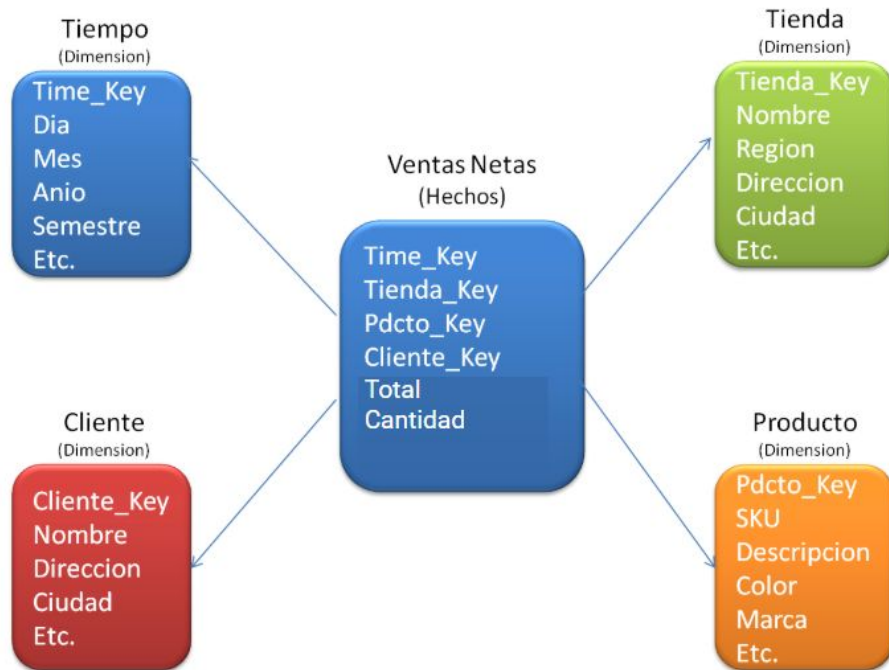
Representación gráfica del modelo Estrella



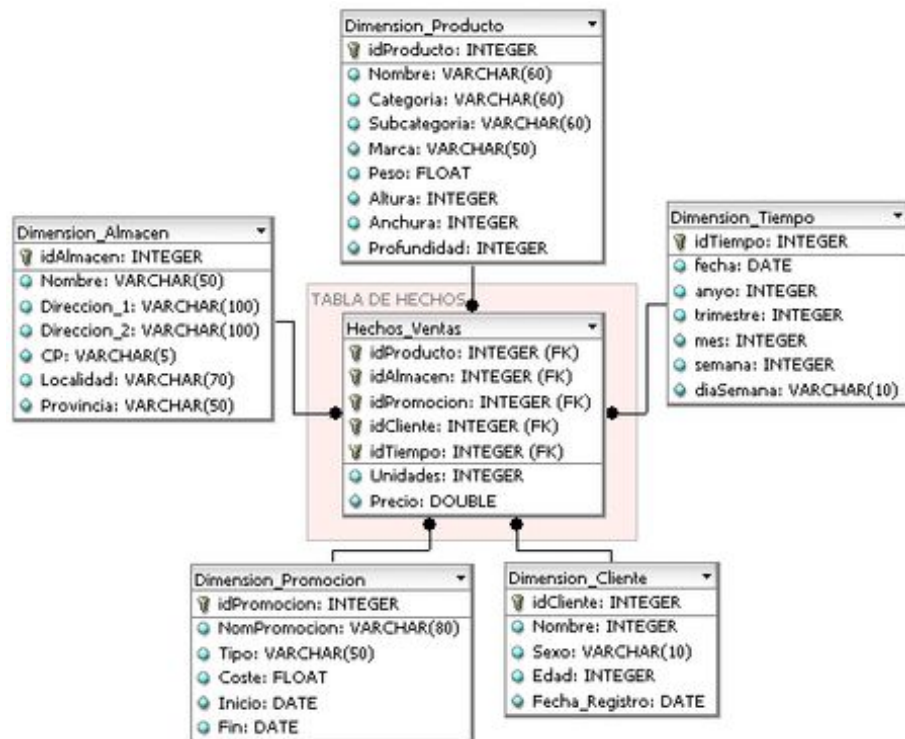
Representación gráfica del modelo Estrella



Representación gráfica del modelo Estrella



Representación gráfica del modelo Estrella



Modelo Copo de Nieve

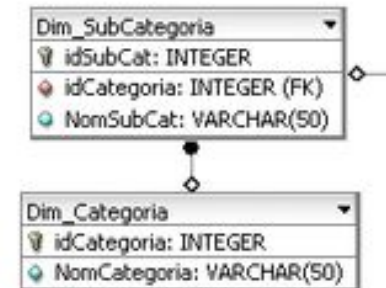
- Es una variación o desviación de un modelo estrella.
- En este modelo la tabla de hechos deja de ser la única relacionada con otras tablas ya que existen otras tablas que se relacionan con las dimensiones.
- Puede implementarse luego de haber desarrollado un Modelo Estrella.

Modelo Copo de Nieve

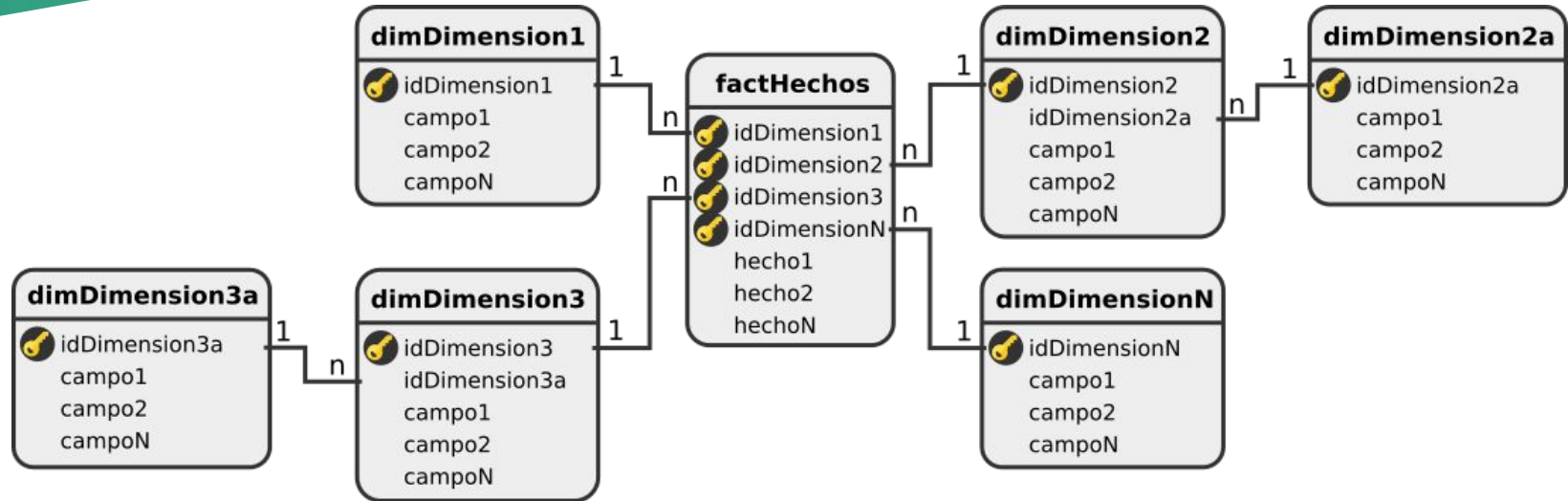
- Existen tablas que no tienen relación directa con la tabla de hechos.
- Este modelo fue creado para facilitar el mantenimiento de las dimensiones.
- La extracción de datos es más difícil y vuelve la tarea de mantener el modelo un poco más compleja.

Modelo Copo de Nieve

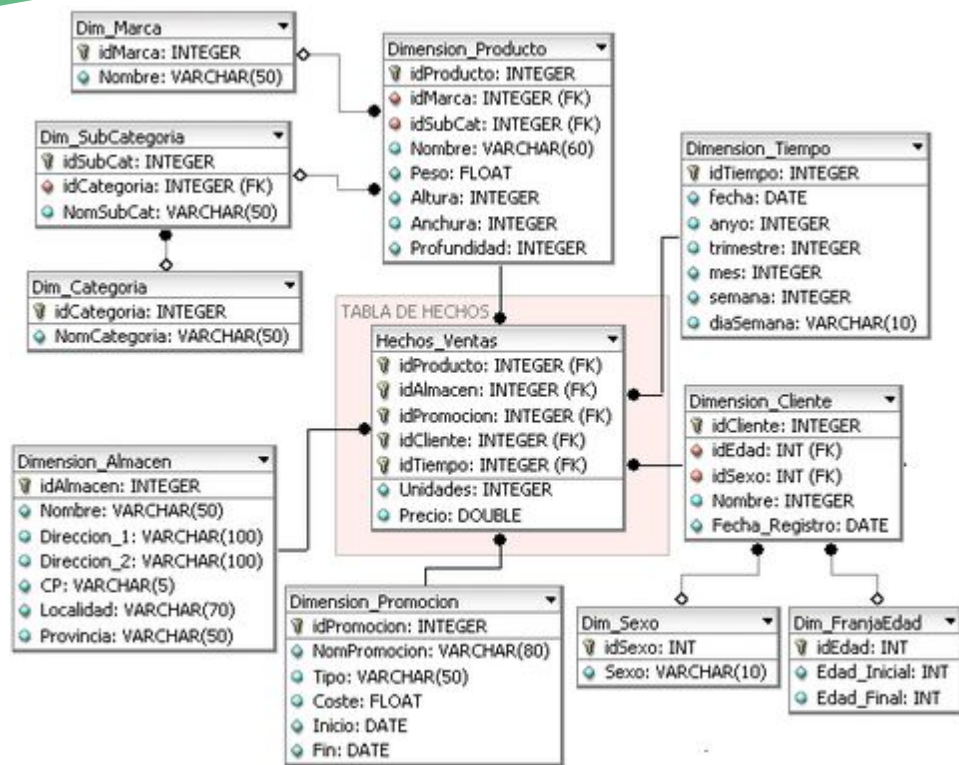
- Su finalidad es normalizar las tablas y así reducir el espacio de almacenamiento al eliminar la redundancia de datos.
- Este modelo puede poseer tablas de dimensiones organizadas en jerarquía.



Representación gráfica del modelo Copo de Nieve



Representación gráfica del modelo Copo de nieve



Modelo Constelación

- Está compuesto por una serie de Esquemas en Estrella.
- Posee lo siguiente:
 - Una tabla de Hechos **principal**.
 - Una o más tabla de Hechos **Auxiliares**, dichas tablas están relacionadas con sus respectivas tablas de Dimensiones.

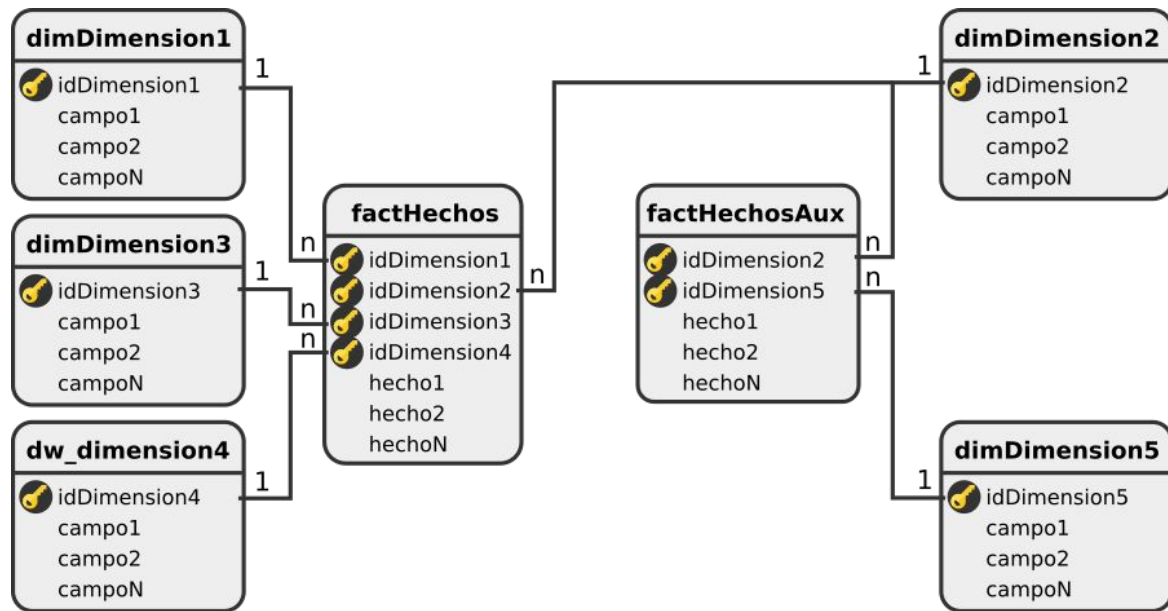
Modelo Constelación

- Las tablas de Hechos Auxiliares pueden vincularse con solo algunas de las tablas de Dimensiones asignadas a la tabla de Hechos Principal, y también pueden hacerlo con nuevas tablas de Dimensiones que se necesiten.

Modelo Constelación

- No es necesario pero se puede dar el caso que las diferentes tablas de Hechos compartan las mismas tablas de Dimensiones.
- Su capacidad analitica es mayor debido a que permite tener más de una tabla de hechos.
- Contribuye a reutilizar tablas de Dimensiones, ya que una misma tabla de Dimensión puede utilizarse para varias tablas de Hechos.

Representación gráfica del modelo Constelación



Modelo Estrella vs Modelo Copo de Nieve

	Estrella	Copo de Nieve
Tablas	Contiene una tabla de hechos rodeada de varias dimensiones.	Contiene una tabla de hechos rodeada de varias dimensiones, que a su vez pueden estar rodeadas de varias dimensiones.
Redundancia	Alta redundancia.	Poca redundancia
Facilidad de uso	Fácil uso.	Difícil de entender, uso mas complicado.

Modelo Estrella vs Modelo Copo de Nieve

	Estrella	Copo de Nieve
Joins	Con un solo join es posible relacionar la tabla de hechos y la de dimensiones.	Requiere múltiples joins para hacer los análisis.
Rendimiento de queries	Ejecuciones más rápidas.	Ejecuciones más complejas, debido a cruces.

Modelo Estrella vs Modelo Copo de Nieve

	Estrella	Copo de Nieve
Estructura	Descentralizada.	Normalizada.
Diseño de bases de datos	Simple.	Complejo.
Cuando usarlo	Cuando las tablas de dimensión tiene pocas filas.	Cuando las tablas de dimensión tienen un tamaño bastante elevado.

¿Dudas o Preguntas?

