

Ejercicios Iniciación – Dimensiones

Ejercicio 1 – ¿Cuáles son atributos dimensionales?

A partir de la siguiente descripción de una venta, identifica cuáles serían dimensiones, a qué dimensión podría pertenecer, y cuál sería el hecho. Columnas:

- Fecha de la venta: 10/04/2025
- Producto: Botella de aceite de oliva
- Categoría del producto: Alimentación
- Precio unitario: 4,50 €
- Cantidad vendida: 3
- Importe total: 13,50 €
- Cliente: Juan García
- Edad del cliente: 35
- Ciudad del cliente: Zaragoza
- Tienda: Supermercado Central Zaragoza
- Vendedor: Marta Ruiz

Ejercicio 2 – Creación de entornos

- Sobre postgresql: crea dos nuevos esquemas:
 - OLTP
 - DataWarehouse

Ejercicio 3 - Tabla de dimensiones DIM PRODUCTO

Dado el fichero de ejercicios EJ3_DIM_Producto.txt:

- Ejecuta el código para crear la tabla dim_producto sobre el esquema DataWarehouse
- Inserta los datos en la tabla siguiendo el script

Preguntas a resolver con la tabla de dimensión:

- ¿Qué campos hay en la tabla y cuál es la clave primaria?
- ¿Cuántos productos hay en stock?
- ¿Cuántos productos tiene cada categoría? ¿Cuál es la que mas tiene?
- ¿Qué productos ofrece la marca Logitech?
- ¿Cuántos productos se han añadido en 2024?
- ¿Qué productos no están disponibles (sin stock)?
- Ver todos los productos ordenados por fecha de alta, de más nuevos a más antiguos
- Contar cuántos productos hay por marca

Preguntas de razonamiento

- ¿Por qué id_producto no tiene valor comercial?
- ¿Para qué sirve el campo codigo_producto?
- ¿Qué ventajas tiene tener esta tabla separada?
- ¿Cómo se podría usar junto con una tabla de hechos?

Ejercicio 4 – Creación de DIM_Cliente:

En los sistemas OLTP es común que los datos de una misma entidad estén distribuidos en varias tablas. Por ejemplo, la información de un cliente puede estar dividida en una tabla con sus datos básicos y otra con su información de contacto. Esto se hace para evitar duplicidad y optimizar las operaciones transaccionales.

Sin embargo, en un entorno de análisis (OLAP / Data Warehouse) necesitamos consolidar estos datos en una sola tabla de tipo dimensión, que contenga toda la información necesaria para el análisis.

En este ejercicio vamos a simular ese proceso: partiendo de dos tablas del esquema OLTP, crearemos una vista que nos permita construir la dimensión dim_cliente en el esquema del Data Warehouse.

Objetivo

- Unificar datos de varias tablas OLTP en una dimensión analítica.
- Practicar el uso de JOIN y VIEW.
- Entender por qué en el DWH buscamos estructuras desnormalizadas.

Pasos:

- Crea las dos tablas de cliente en el esquema OLTP con ayuda del script EJ4_Dim_Clientes
- Crea sobre el esquema datawarehouse una vista que una ambas tablas, para crear así la dimensión cliente que se usará para los sistemas de análisis de datos

Preguntas adicionales:

- ¿Por qué es útil separar los datos en el OLTP pero unirlos en el DWH?
- ¿Qué campos serían clave para analizar a los clientes desde un informe?
- ¿Qué pasaría si un cliente no tuviera contacto? ¿Cómo lo gestionamos en el JOIN?.

Ejercicios Iniciación – Hechos

Ejercicio 5 –¿Qué tipo de hecho es cada uno?

Imagina que trabajas con datos relacionados con ventas, actividades bancarias y tráfico web. A continuación, se presentan distintos tipos de hechos que podrías encontrar en una base de datos empresarial.

El objetivo es que determines si cada uno es **aditivo**, **semi-aditivo** o **no aditivo**.

1. **Ventas de productos:** ¿Se pueden sumar las ventas de diferentes días para obtener el total vendido?
2. **Saldo bancario al final del mes:** ¿Se pueden sumar los saldos de varios meses para obtener un saldo total?
3. **Visitas a un sitio web:** ¿Se pueden sumar las visitas de diferentes días para obtener el total de visitas?
4. **Kilómetros recorridos por un vehículo:** ¿Se pueden sumar los kilómetros recorridos durante varios días?
5. **Tiempo dedicado a una tarea:** ¿Se puede sumar el tiempo dedicado durante varios días?
6. **Porcentaje de descuentos aplicados:** ¿Tiene sentido sumar los porcentajes de descuento de diferentes meses?
7. **Inventario de un almacén:** ¿Se puede sumar el inventario de diferentes días?
8. **Número de productos vendidos:** ¿Se pueden sumar los productos vendidos de diferentes días para obtener el total?

Ejercicio 6 - A continuación tienes varios ejemplos de situaciones reales que podrían dar lugar a una tabla de hechos en un modelo dimensional.

Para cada caso, responde a estas tres preguntas:

1. **¿Qué representa exactamente cada fila?** (*Granularidad*)
2. **¿Qué tipo de tabla es?** (*Transaccional, Acumulativa o Snapshot periódico*)
3. **¿El nivel de granularidad es fino o grueso?**

Casos de uso:

- Líneas de productos vendidos en tickets de supermercado.
- Resumen mensual de ventas por vendedor.
- Progreso de cada pedido online (recibido → preparado → enviado → entregado).
- Cada vez que un cliente entra en la web y ve un producto.
- Evolución del estado de un paquete en una empresa de mensajería.
- Stock de productos en almacén cada fin de semana.
- Registro de cada compra que se hace en una app.

Ejercicio 7 – Tabla de hechos

A continuación, en el script EJ7_FCT_Compras.txt se incluyen los datos de la tabla de hechos, con información relevante sobre cuándo se realizaron compras, qué cliente, producto, cantidad e importe total de ventas.

Tareas

- Carga la información en base al script de datos
- Crea un diagrama en el que se representen visualmente las tablas y sus relaciones
- Indica en el diagrama las claves primarias de cada tabla y todos los campos
- Responde a las siguientes preguntas combinando los datos de las tablas de hechos y las tablas DIM_Producto y DIM_Compras ya existentes

Preguntas:

- ¿Cuántos clientes han realizado compras?
- ¿Qué cliente ha gastado más dinero?
- ¿Cuántos productos de la categoría 'Accesorios' se han comprado?
- ¿Qué cliente ha comprado más productos de la marca 'Logitech'?
- ¿Cuántos productos no están disponibles (en_stock = FALSE) pero han sido comprados?
- ¿Cuál fue el importe de compras realizado por los clientes en el mes de marzo de 2024?
- ¿Cuántos productos se han comprado por cada cliente?
- ¿Qué productos de la categoría 'Informática' se compraron más de una vez?
- ¿Cuántos clientes han comprado productos de la marca 'Sony'?
- Productos que ha comprado la gente de Zaragoza ¿Y si tuviésemos que hacerlo sobre el esquema normalizado?
- Productos que ha comprado la gente de Zaragoza y que además el producto esté en stock ¿Y si tuviésemos que hacerlo sobre el esquema normalizado?

Ejercicio 8 – Transformación de un modelo OLTP a un modelo estrella

Una empresa de retail vende productos tecnológicos y del hogar a clientes de varios países europeos. La empresa guarda sus datos operacionales en una base de datos PostgreSQL con un modelo altamente normalizado (OLTP). Tu misión es ayudar al equipo de análisis a construir un modelo estrella (esquema dw) a partir del esquema oltp, generando vistas que permitan un análisis ágil y eficiente con SQL o Power BI.

Objetivos del ejercicio

1. Comprender la diferencia entre modelos OLTP y OLAP.
2. Practicar la creación de vistas para modelar dimensiones.
3. Construir una tabla de hechos basada en datos transaccionales.
4. Ejecutar consultas analíticas que combinen hechos y dimensiones.

Tareas a realizar

1. **Preparación del entorno:**
 - Crear las tablas del esquema oltp (estructura entregada).
 - Insertar los datos de ejemplo.
 - Crear el esquema dw.
2. **Modelado:**
 - Crear vistas de dimensión: dim_producto, dim_cliente, dim_fecha.
 - Crear la vista fact_ventas en el esquema dw.
3. **Esquema visual:**
 - Crea un diagrama en el que se representen visualmente las tablas y sus relaciones
 - Indica en el diagrama las claves primarias de cada tabla y todos los campos
4. **Consultas de análisis (realizar al menos 5):**
 - Total de ventas por país y categoría.
 - Ranking de productos más vendidos.
 - Ventas por ciudad y mes.
 - Clientes que más compran.
 - Tendencia diaria de ventas.