

Step 1: Scenario

You are building a **mini Data Warehouse** for an **E-Commerce company**.

Business wants to analyze:

- How much revenue is generated by category, product, and time.
- How customer information changes over time (name, email).

We will use **Star Schema**:

- **Fact table:** `fact_orders`
 - **Dimension tables:** `dim_customers`, `dim_products`, `dim_date`
-

Step 2: Tasks

Part A: Design Schema

1. Create the following tables in PostgreSQL:

`dim_customers`

- `customer_sk` (PK, surrogate key, auto-increment)
- `customer_id` (business key)
- `name`
- `email`
- `effective_date` (date when record became active)
- `end_date` (date when record expired; NULL = active)
- `is_active` (Y/N)

`dim_products`

- `product_sk` (PK, surrogate key)
- `product_id` (business key)
- `name`
- `category`

`dim_date`

- `date_sk` (PK)
- `full_date` (date)
- `year`, `month`, `day_of_week`

`fact_orders`

- order_id (PK)
- order_date_sk (FK → dim_date.date_sk)
- customer_sk (FK → dim_customers.customer_sk)
- product_sk (FK → dim_products.product_sk)
- quantity
- price
- total_amount (quantity × price)

☞ **Learning here:** surrogate keys, star schema, date dimension.

Part B: Populate Data

1. Insert at least:
 - 5 customers
 - 5 products
 - 10 orders
 2. Implement **SCD Type-2** in dim_customers:
 - Example: customer changes email → old record end_date updated, new record inserted with same customer_id but new surrogate key.
-

Part C: Queries

Write SQL queries for:

1. Show **total revenue per product category**.
 2. Show **monthly revenue trend** (use dim_date).
 3. Find customers whose information changed over time (SCD-2 history).
 4. Using a **CTE**, find the top 2 customers by total spend in each month.
 5. Using a **window function**, show each customer's order **rank by date**.
-

Step 3: Deliverables

1. A .sql file containing:
 - CREATE TABLE statements (schema).
 - INSERT statements (sample data).
 - Queries (with comments).
2. A short write-up (half page):
 - Difference between OLTP and OLAP modeling.
 - Why surrogate keys are important in dimensions.

- What you learned about Slowly Changing Dimensions.