

# SQL- Capstone Project: E-Commerce Sales Analysis and Reporting

## Project Objective:

To analyze and generate insights from an e-commerce database that includes information about orders, customers, products, and transactions. The goal is to answer key business questions, generate insights, and create a reporting structure that could be used by the company's analytics or finance team.

---

## Database Structure

Create a database

Create a database with the following tables:

### 1. Customers Table

- **customer\_id** (Primary Key)
- **customer\_name**
- **age**
- **gender**
- **country**
- **email**
- **date\_joined**

### 2. Products Table

- **product\_id** (Primary Key)
- **product\_name**
- **category**
- **sub\_category**
- **unit\_price**
- **unit\_cost**

### 3. Orders Table

- **order\_id** (Primary Key)
- **customer\_id** (Foreign Key referencing Customers.customer\_id)

- **order\_date**
- **order\_status** (e.g., “Completed”, “Pending”, “Canceled”)
- **payment\_method** (e.g., “Credit Card”, “PayPal”, “Bank Transfer”)
- **shipping\_address**

#### 4. Order\_Items Table

- **order\_item\_id** (Primary Key)
- **order\_id** (Foreign Key referencing Orders.order\_id)
- **product\_id** (Foreign Key referencing Products.product\_id)
- **quantity** (quantity ordered)
- **total\_price** (quantity \* Products.unit\_price)

#### 5. Transactions Table

- **transaction\_id** (Primary Key)
- **order\_id** (Foreign Key referencing Orders.order\_id)
- **transaction\_date**
- **amount\_paid**
- **payment\_status** (e.g., “Paid”, “Pending”)

---

### Project Tasks and Key Requirements

#### 1. Data Preparation and Setup

- **Create the Database:** Write SQL scripts to create the tables and relationships.
- **Insert Sample Data:** Populate the tables with data provided in the folder

#### 2. Basic Data Exploration

- Retrieve the total number of customers, products, orders, and transactions.
- Find the earliest and latest order dates in the database.
- Calculate the total revenue, total cost, and profit for all orders.

#### 3. Customer Analysis

- **Demographic Breakdown:** Count of customers by gender and country.
- **Customer Segmentation:** Group customers into age segments (e.g., 18-25, 26-35, etc.) and find the average order count and average spend per segment.

- **Top Customers:** Identify the top 10 customers by total spend and by total orders placed.

#### 4. Product Analysis

- **Top Products:** List the top 10 products by total quantity sold and by revenue.
- **Product Category Analysis:** Calculate total revenue, profit, and quantity sold by each product category and sub-category.
- **Low-Performing Products:** Identify products that have not been sold in the last 6 months.

#### 5. Order Analysis

- **Order Status Summary:** Count of orders by status (Completed, Pending, Canceled).
- **Payment Methods Analysis:** Count of orders by payment method, and percentage of total.
- **Monthly Sales Trend:** Calculate total revenue and number of orders by month to identify sales trends.

#### 6. Revenue and Profit Analysis

- **Total Revenue and Profit:** Calculate overall revenue and profit for completed orders.
- **Profit by Product:** Calculate profit margins by product, product category, and sub-category.
- **High-Profit Orders:** Identify the top 5 orders with the highest profit margins.

#### 7. Transaction Analysis

- **Payment Status:** Count transactions by payment status (Paid, Pending) and calculate the amount of pending revenue.
- **Transaction Completeness:** Identify orders that do not have matching transactions (potential data issue).
- **Transaction Trends:** Calculate total transactions and average transaction amount by month.

#### 8. Advanced Insights

- **Customer Retention:** Calculate the percentage of customers who placed more than one order.
- **Repeat Purchase Rate:** Calculate the repeat purchase rate for customers.
- **Cross-Category Purchases:** Identify customers who have purchased products from more than one category.

#### 9. Reporting and Visualization (Optional if using BI Tools)

- **Revenue and Profit Dashboard:** Use SQL to aggregate data for creating a Power BI or Excel dashboard showing monthly revenue, top products, and key metrics.
- **Customer Dashboard:** Create tables and views that can be used for visualizing customer demographics and spending patterns.

- **Trend Analysis Dashboard:** Set up views and aggregations that support trend analysis over time for orders and revenue.
- 

### Example Queries

- **Top 10 Products by Quantity Sold**

```
SELECT p.product_name, SUM(oi.quantity) AS total_quantity_sold
FROM Order_Items oi
JOIN Products p ON oi.product_id = p.product_id
GROUP BY p.product_name
ORDER BY total_quantity_sold DESC
LIMIT 10;
```

- **Monthly Sales and Profit Trend**

```
SELECT DATE_TRUNC('month', o.order_date) AS month,
       SUM(oi.total_price) AS total_revenue,
       SUM((oi.quantity * p.unit_cost)) AS total_cost,
       (SUM(oi.total_price) - SUM((oi.quantity * p.unit_cost))) AS total_profit
FROM Orders o
JOIN Order_Items oi ON o.order_id = oi.order_id
JOIN Products p ON oi.product_id = p.product_id
WHERE o.order_status = 'Completed'
GROUP BY month
ORDER BY month;
```

- **Repeat Purchase Rate**

```
SELECT COUNT(DISTINCT customer_id) AS total_customers,
       COUNT(DISTINCT CASE WHEN order_count > 1 THEN customer_id END) AS repeat_customers,
       (COUNT(DISTINCT CASE WHEN order_count > 1 THEN customer_id END) * 1.0 / COUNT(DISTINCT
customer_id)) AS repeat_purchase_rate
FROM (
  SELECT customer_id, COUNT(order_id) AS order_count
```

```
FROM Orders

GROUP BY customer_id

) AS customer_orders;
```

---

### Expected Deliverables

1. **SQL Scripts:** Scripts to create and populate the tables, plus all queries you've used for analysis.
  2. **Report:** A summary document detailing insights and answers to key business questions.
  3. **Views or Materialized Views:** Create SQL views for commonly used queries (e.g., monthly sales, top products, repeat customers).
  4. **(Optional) Dashboard:** Create a visual dashboard in a BI tool or Excel to display key metrics and insights.
- 

### Skills Practiced

- **Data Modeling and Database Design:** Setting up tables, keys, and relationships.
- **Data Manipulation:** Using joins, subqueries, and common table expressions (CTEs).
- **Aggregations and Calculations:** Summing, counting, averaging, and using conditional logic.
- **Advanced SQL:** Using window functions, grouping, and filtering data.
- **Insights and Reporting:** Answering business questions with SQL and summarizing findings.

This project will give you hands-on experience with a range of SQL skills, preparing you for data analysis and database management tasks in real-world scenarios.

### Supporting Materials:

- [https://www.w3schools.com/sql/sql\\_exercises.asp](https://www.w3schools.com/sql/sql_exercises.asp)