Project: E-commerce SQL Analysis

1. Introduction

This project analyses e-commerce sales data using Google BigQuery and SQL. The goal is to generate business insights, including top customers, revenue trends, repeat customer behaviour, product category performance, and average order value.

2. Dataset Overview

- Customers Table: customer_id, name, region
- Products Table: product_id, product_name, category, price
- Orders Table: order_id, customer_id, product_id, quantity, order_date

3. SQL Queries

The following SQL queries were used to analyse the data:

1. Top 5 Customers by Total Spend

```
SELECT Customers.customer_id, Customers.name, SUM(orders.quantity * products.price) AS
total_spent
FROM model-ruler-427606-r7.Ecommerce.Orders
JOIN model-ruler-427606-r7.Ecommerce.Customers
ON orders.customer_id = customers.customer_id
JOIN model-ruler-427606-r7.Ecommerce.Products ON orders.product_id =
products.product_id
GROUP BY customers.customer_id, customers.name
ORDER BY total_spent DESC
LIMIT 5;
```

2. Monthly Revenue Trend

```
SELECT FORMAT_DATE('%Y-%m', orders.order_date) AS month,
       SUM(orders.quantity * products.price) AS revenue
FROM model-ruler-427606-r7.Ecommerce.orders

JOIN model-ruler-427606-r7.Ecommerce.products ON orders.product_id =
products.product_id
GROUP BY month
ORDER BY month;
```

3. Repeat vs New Customers

```
WITH first_purchase AS (
    SELECT customer_id, MIN(order_date) AS first_order
    FROM model-ruler-427606-r7.Ecommerce.orders
    GROUP BY customer_id
)
SELECT COUNT(DISTINCT CASE WHEN orders.order_date = first_purchase.first_order THEN
orders.customer_id END) AS new_customers,
       COUNT(DISTINCT CASE WHEN orders.order_date > first_purchase.first_order THEN
orders.customer_id END) AS repeat_customers
FROM model-ruler-427606-r7.Ecommerce.orders
JOIN first_purchase ON orders.customer_id = first_purchase.customer_id;
```

4. Best-Selling Product Categories

```
SELECT products.category, SUM(orders.quantity) AS total_sold
FROM model-ruler-427606-r7.Ecommerce.orders
JOIN model-ruler-427606-r7.Ecommerce.products ON orders.product_id =
products.product_id
GROUP BY products.category
ORDER BY total_sold DESC;
```

5. Average Order Value

```
SELECT AVG(order_value) AS avg_order_value
FROM (
    SELECT orders.order_id, SUM(orders.quantity * products.price) AS order_value
    FROM model-ruler-427606-r7.Ecommerce.orders
    JOIN model-ruler-427606-r7.Ecommerce.products ON orders.product_id =
products.product_id
    GROUP BY orders.order_id
);
```

4. Insights

- Top 5 customers contributed ~78% of total revenue.
- Monthly trend shows peak sales in the festive season (Nov-Dec).
- Repeat customers are ~90% of the base but generate ~99% of revenue.
- Best-selling categories: Home > Fashion > Groceries.
- Average order value is around ₹2,697.

5. Conclusion

The analysis provided valuable business insights for customer retention, sales strategy, and inventory planning. This project demonstrates the practical use of SQL in BigQuery for real-world business analytics.