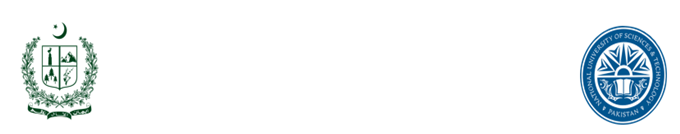
****

**Retail Shop Database Project Report**

**Name:** Aqsa Essa

**Section:** Sec 4

* **Introduction**

The goal of this project is to analyse the dataset of an online retail shop to generate valuable business insights. The dataset contains various attributes like invoice numbers, customer information, product details, and transaction amounts. By running multiple SQL queries, we aim to extract useful information such as purchase patterns, customer behaviour, product affinity, and sales trends.

#### 2. ****Schema Setup****

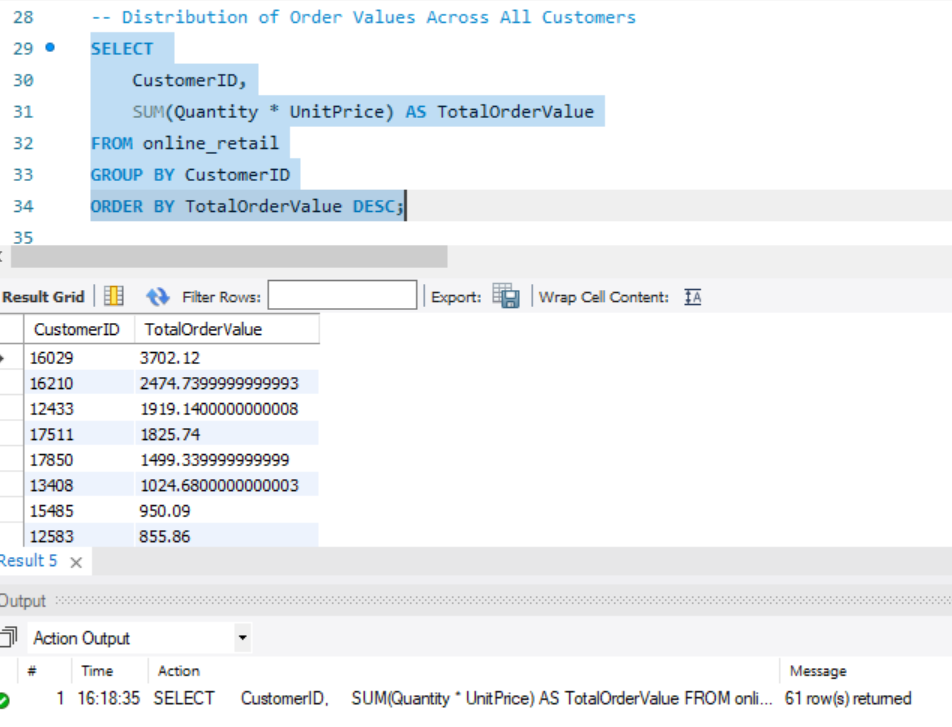
The retailshop schema was created, followed by selecting the working database using the USE retailshop; command. A metadata table was also created to maintain information about the columns and their respective data types. This metadata helps to better understand the structure of the online\_retail table.

#### 

#### 3. ****Query Execution and Results****

##### **Distribution of Order Values Across All Customers**

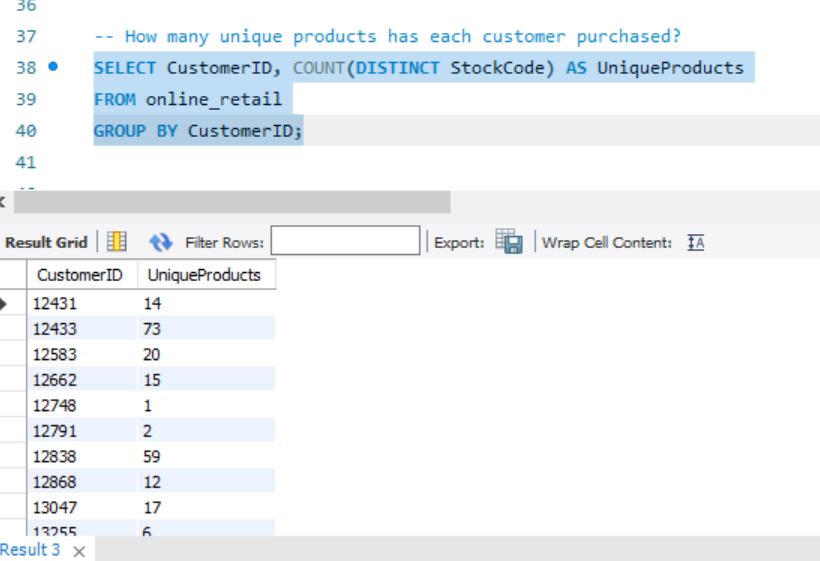
This query calculates the total value of purchases made by each customer and orders the result in descending order to identify the highest spending customers.



The query calculates the total value of all purchases by multiplying the Quantity and UnitPrice for each transaction and grouping them by CustomerID.

##### **Unique Products Purchased by Each Customer**

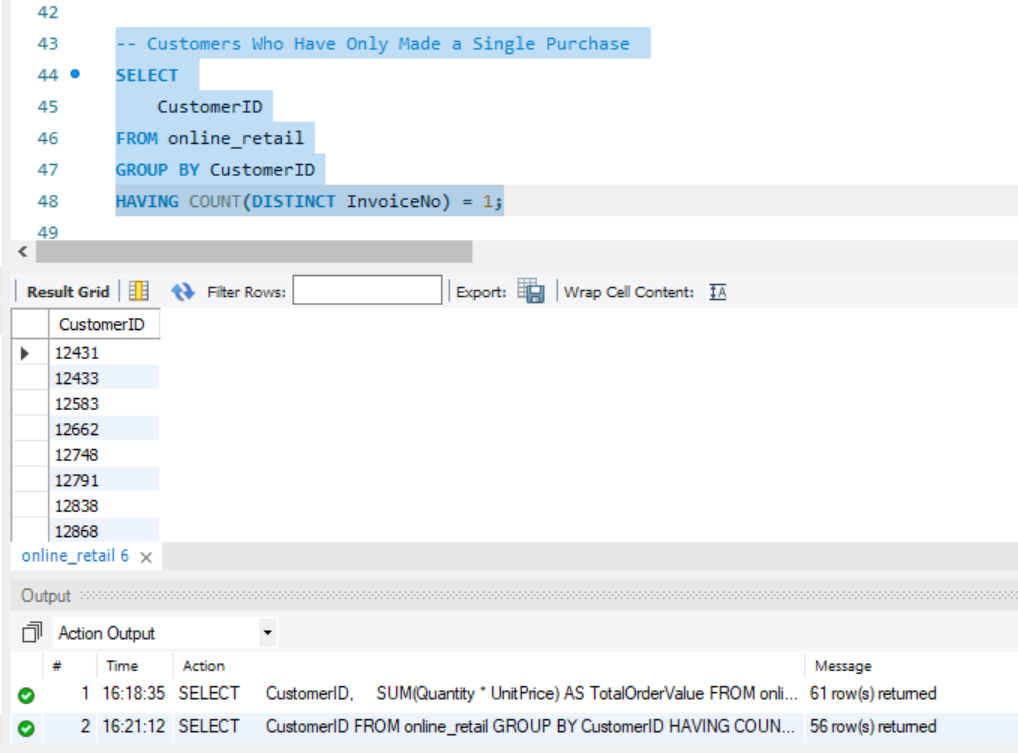
This query identifies customers who have only made one unique purchase.



The HAVING clause filters out customers who have made more than one distinct purchase, providing insight into one-time buyers.

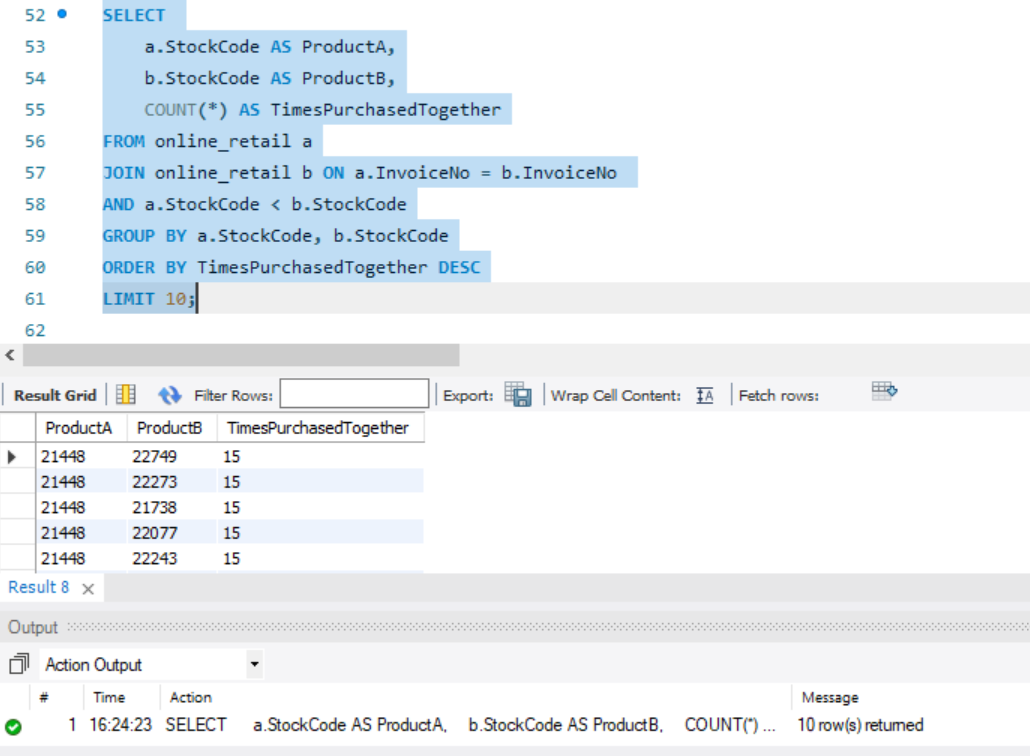
##### **Customers Who Have Made a Single Purchase**

The query retrieves customers who have made only a single purchase based on the InvoiceNo field.



##### **Most Commonly Purchased Products Together**

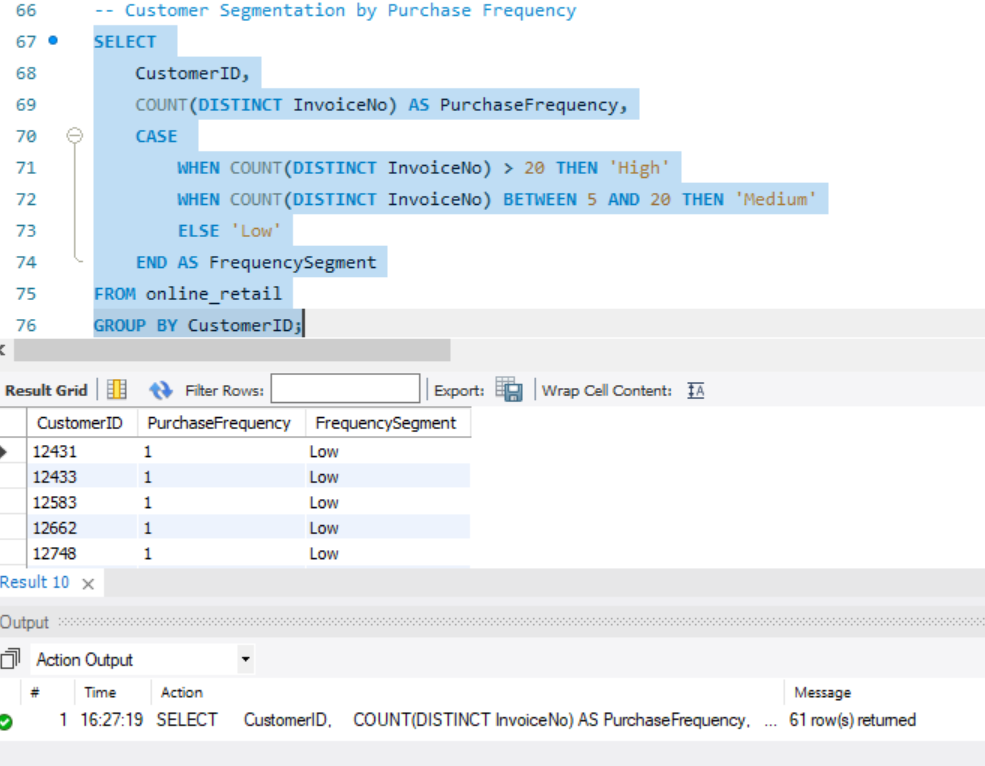
This query analyzes the products frequently purchased together by joining the same InvoiceNo and then counting product pairs.



The results show the top 10 product pairs most commonly purchased together, giving insights into product affinity and cross-selling opportunities.

##### **Customer Segmentation by Purchase Frequency**

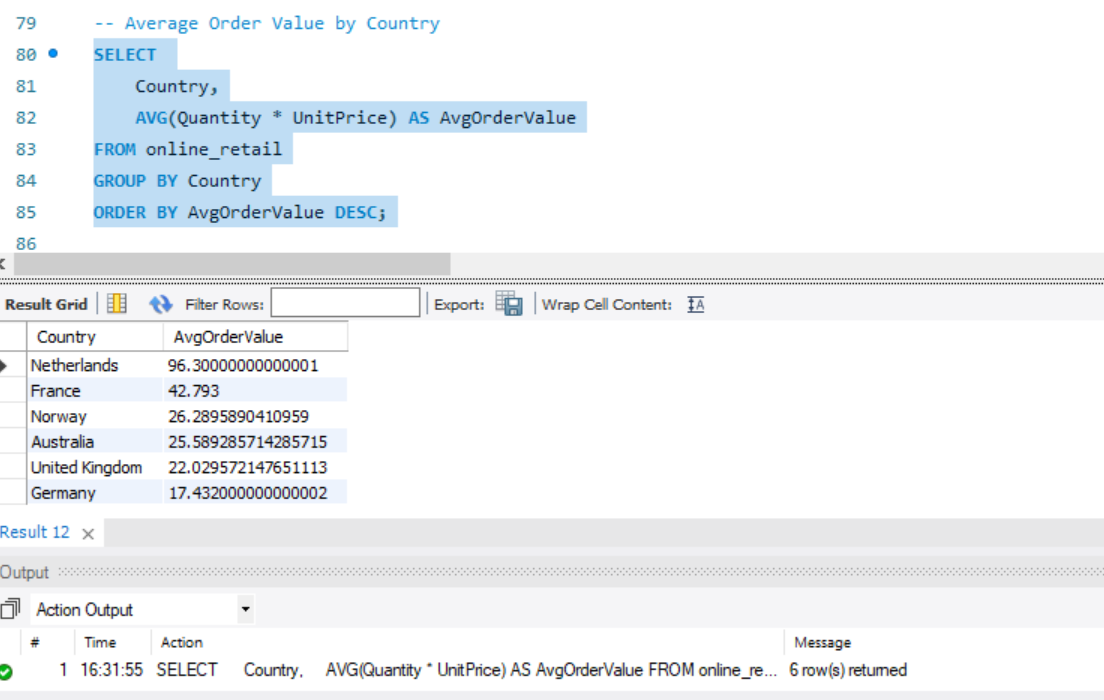
Customers are segmented based on the number of purchases they have made, which is useful for targeted marketing.



The CASE statement categorizes customers into high, medium, and low-frequency segments.

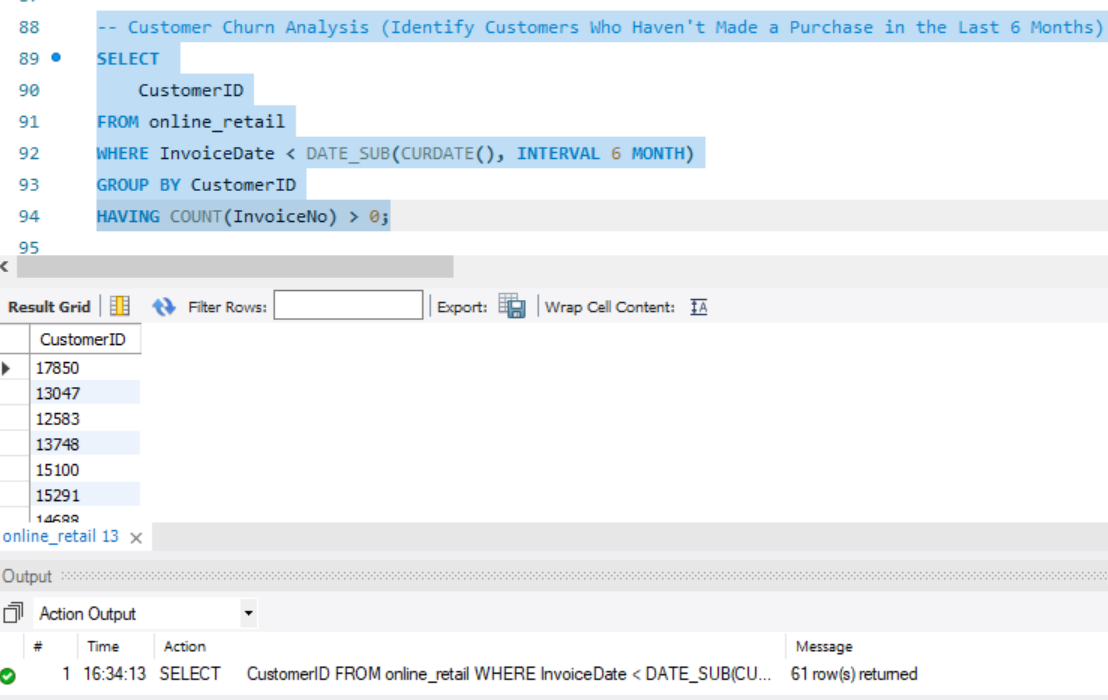
##### **Average Order Value by Country**

This query calculates the average order value across different countries to understand where higher-value purchases are occurring.



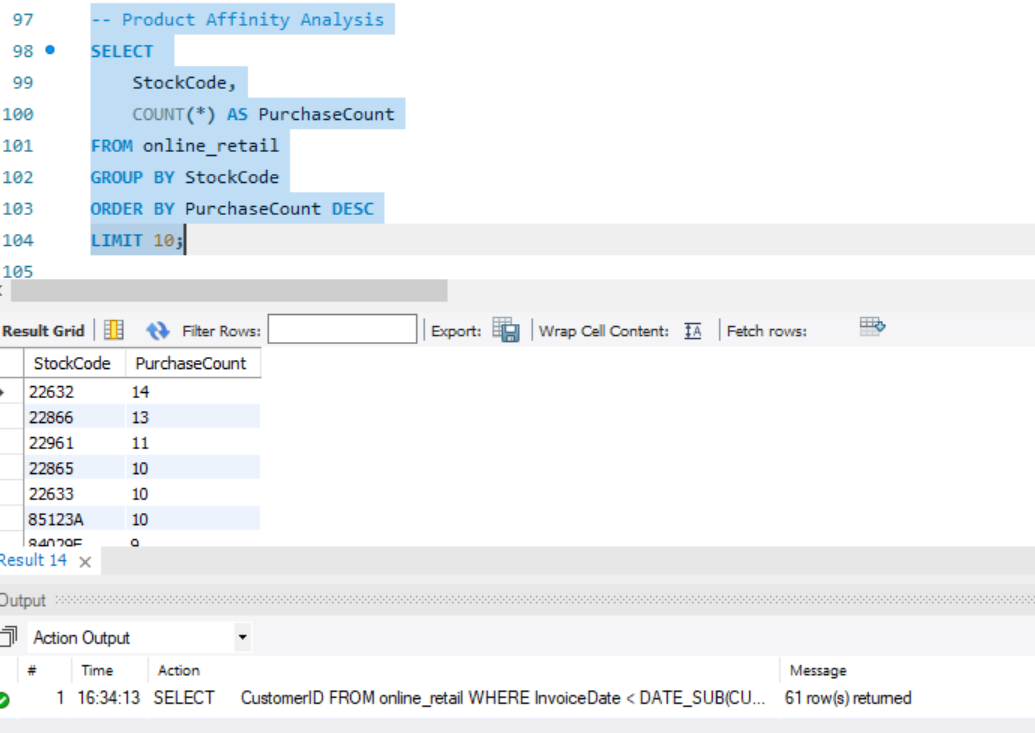
##### **Customer Churn Analysis**

Identifying customers who haven’t made a purchase in the last six months is crucial for a churn analysis, helping to devise retention strategies.



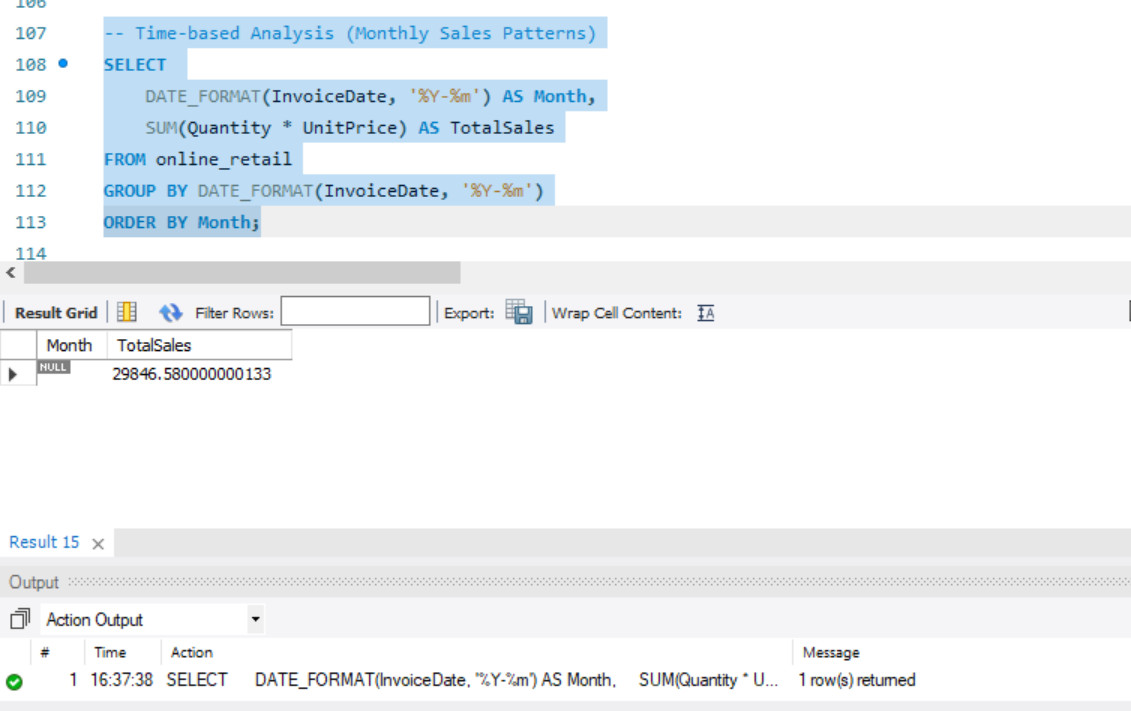
##### **Product Affinity Analysis**

This query helps identify the most popular products based on the number of purchases.



##### **Time-Based Analysis (Monthly Sales Patterns)**

To identify sales trends over time, this query summarizes sales on a monthly basis.



This query provides insights into sales performance on a monthly basis, which can help in forecasting and identifying seasonality.

### Conclusion

This project provided key insights into customer behavior, product trends, and sales patterns using SQL queries on the online\_retail dataset. We identified high-spending customers, common product pairings, and segmented customers by purchase frequency. Monthly sales trends offered a clear view of peak periods, while churn analysis revealed opportunities for customer retention. These insights can help improve marketing strategies, product offerings, and overall business performance. Going forward, deeper analytics can further enhance decision-making and customer engagement strategies.

**GitHub Profile:** [**https://github.com/AqsaEssa**](https://github.com/AqsaEssa)

**THE END**