**Online Retail Segmentation**

**Introduction**

The objective of this project is to perform customer segmentation for an online retail dataset using SQL to uncover insights such as purchase frequency, customer loyalty, product affinity, and customer churn. By analyzing purchasing patterns and demographic data, businesses can improve their marketing efforts and customer engagement strategies. The dataset contains details such as invoice numbers, product descriptions, quantities sold, unit prices, customer IDs, and the countries where the transactions occurred. The goal is to use SQL queries to extract valuable business insights from the data.

**SQL Queries**

1. **Meta Data**

This query defines meta data

CREATE TABLE online\_retail (

InvoiceNo INT NULL,

StockCode NVARCHAR(50),

Description NVARCHAR(MAX),

Quantity INT NULL,

InvoiceDate DATETIME,

UnitPrice FLOAT NULL,

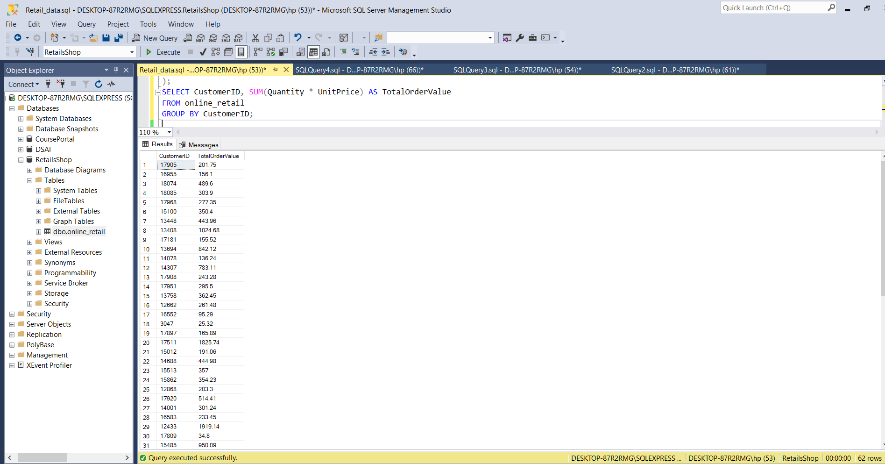
CustomerID BIGINT NULL,

Country NVARCHAR(100)

);

**2. Order Value Distribution Across All Customers**

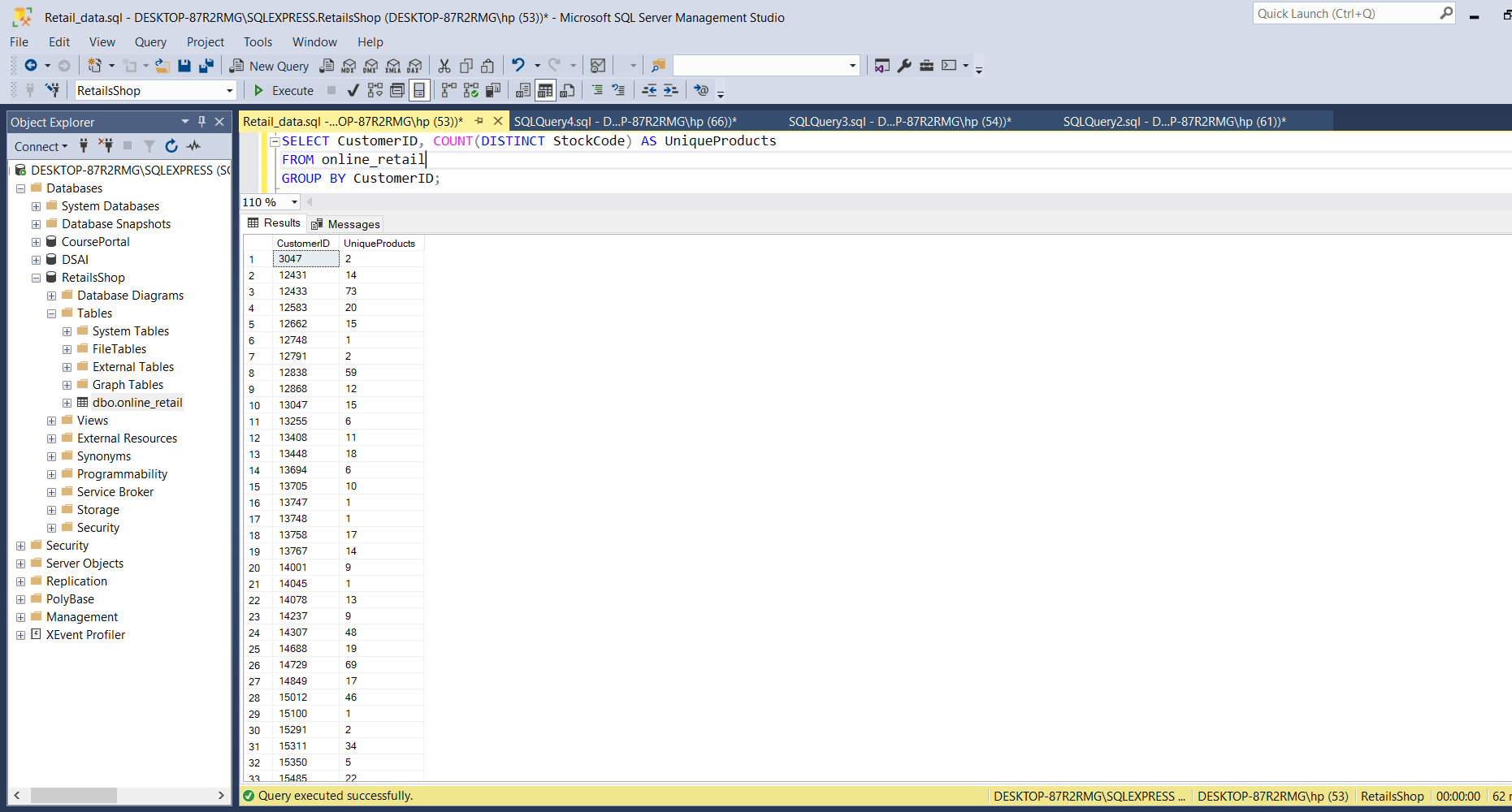
This query calculates the total order value per customer, helping us understand which customers have made the largest purchases.



**Resut:**  
The results of this query showed that a few customers have very high order values, which could indicate repeat or bulk purchases. This insight is essential for identifying valuable customers who may need special attention.

**3. Number of Unique Products Purchased by Each Customer**

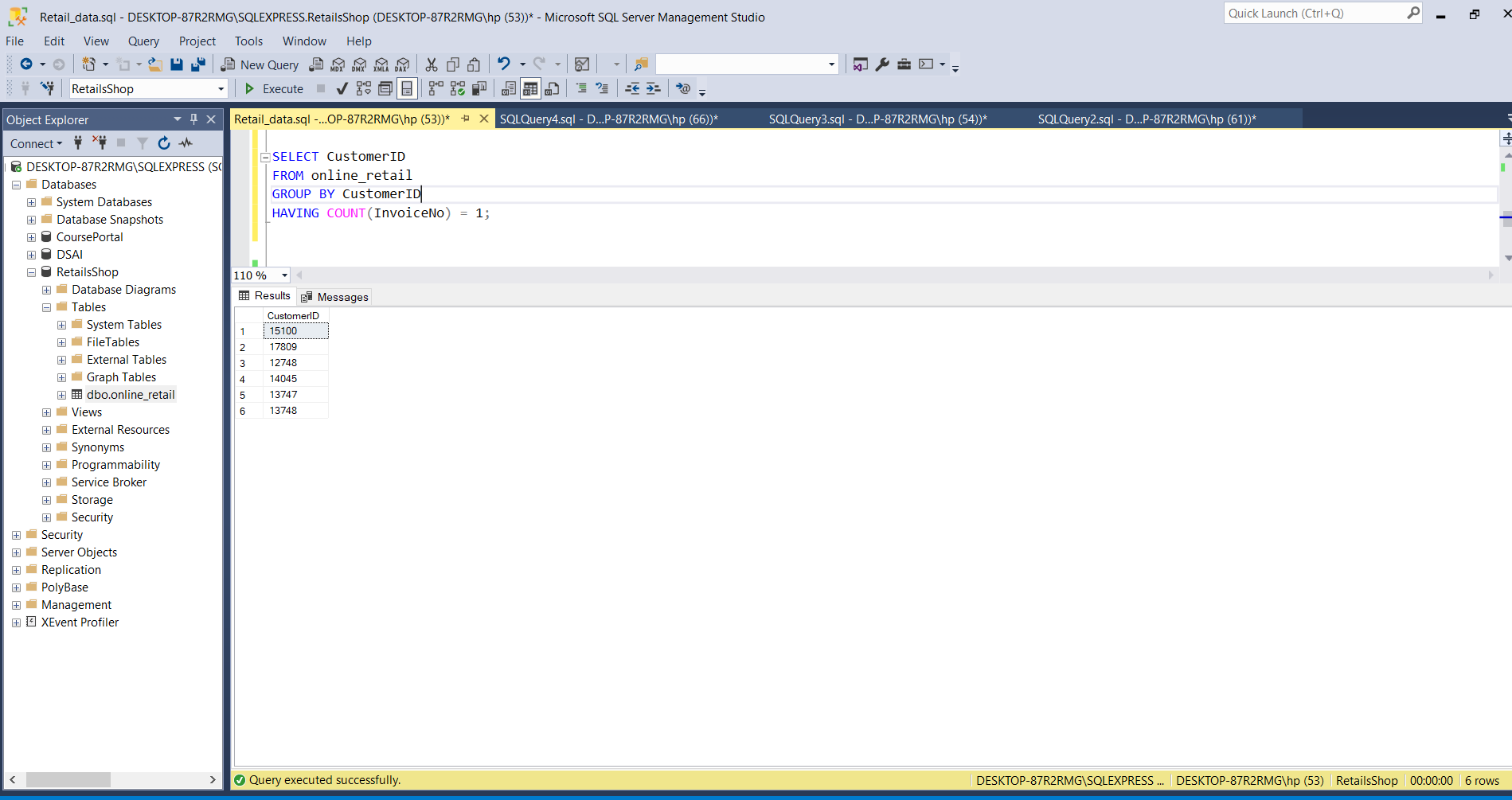
This query helps us determine how diverse a customer's purchases are by counting the unique products they've bought.



**Result:**  
The query revealed that most customers tend to purchase only a few unique products, while a small number of customers have purchased a wide variety, indicating they may be browsing more extensively or buying in bulk for different needs.

**4. Identifying Customers with Only One Purchase**

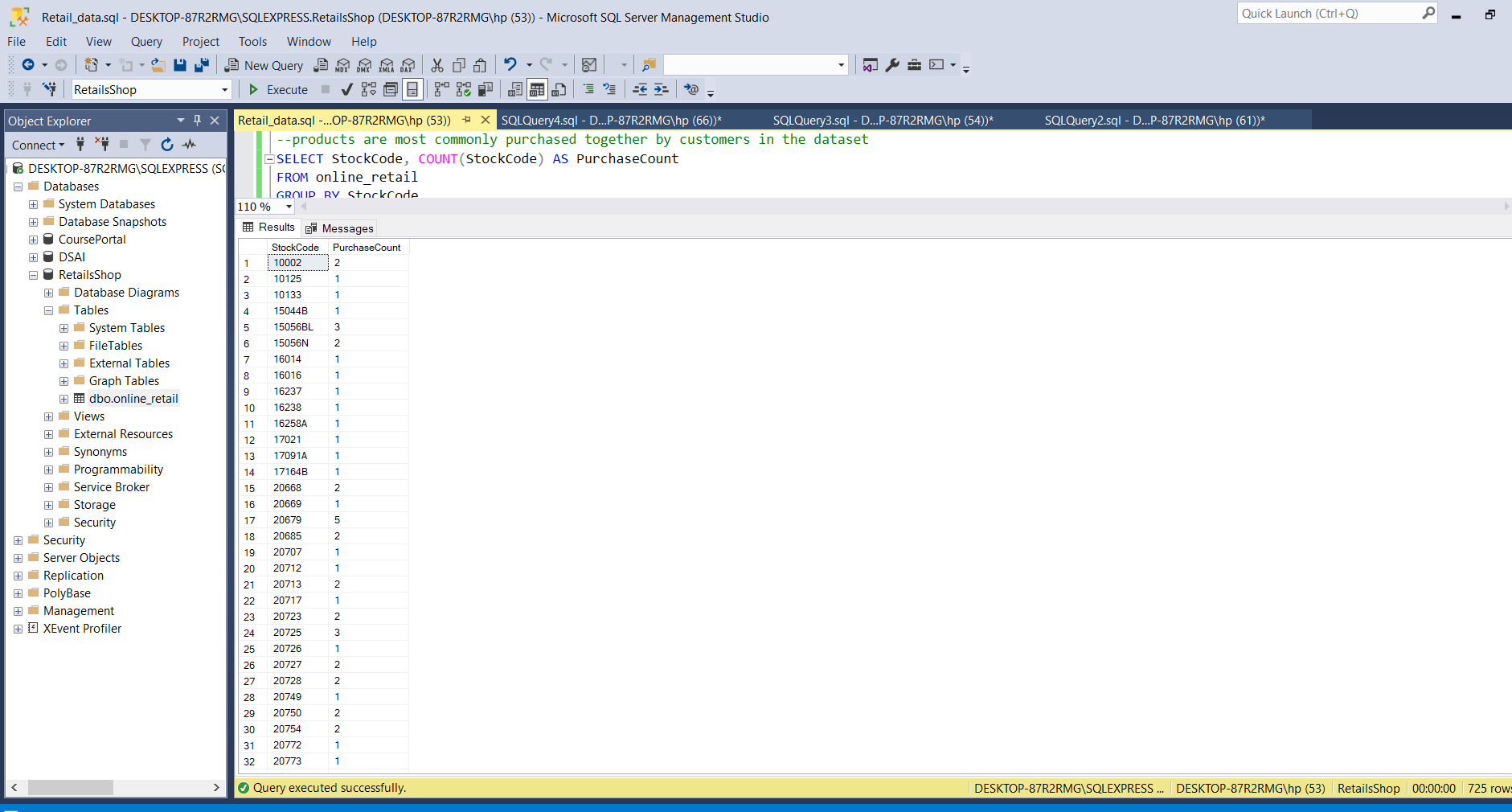
This query identifies customers who have only made a single purchase from the store. These customers may represent a potential churn risk.



**Result:**  
A significant portion of the customer base has only made a single purchase, which suggests an opportunity to engage these customers more effectively to increase retention.

**5. Most Commonly Purchased Products Together**

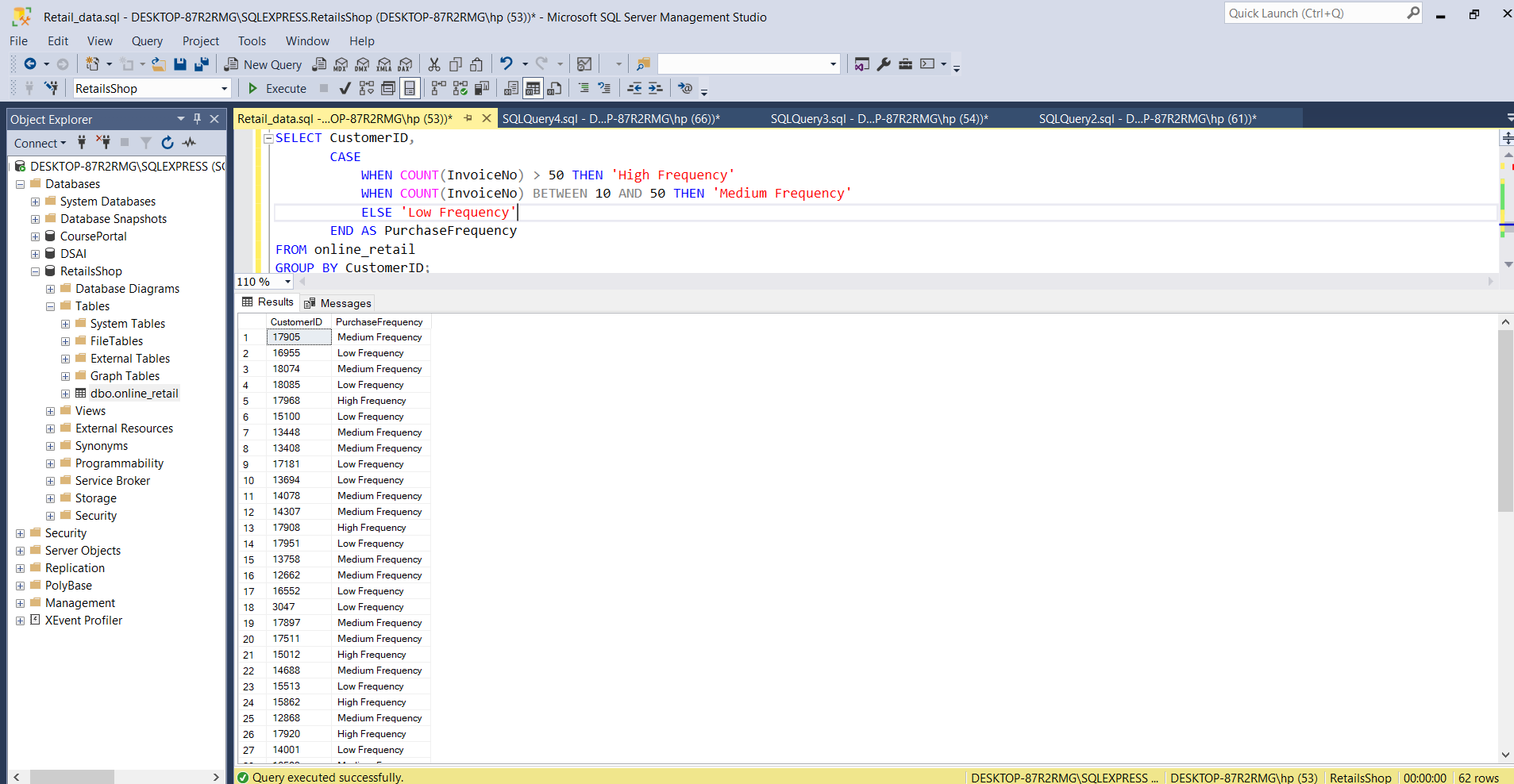
This query identifies which products are frequently purchased together, allowing the business to identify cross-selling opportunities.



**Result:**  
The results showed that several products are frequently bought together, which indicates potential bundling opportunities for the business.

**6. Customer Segmentation by Purchase Frequency**

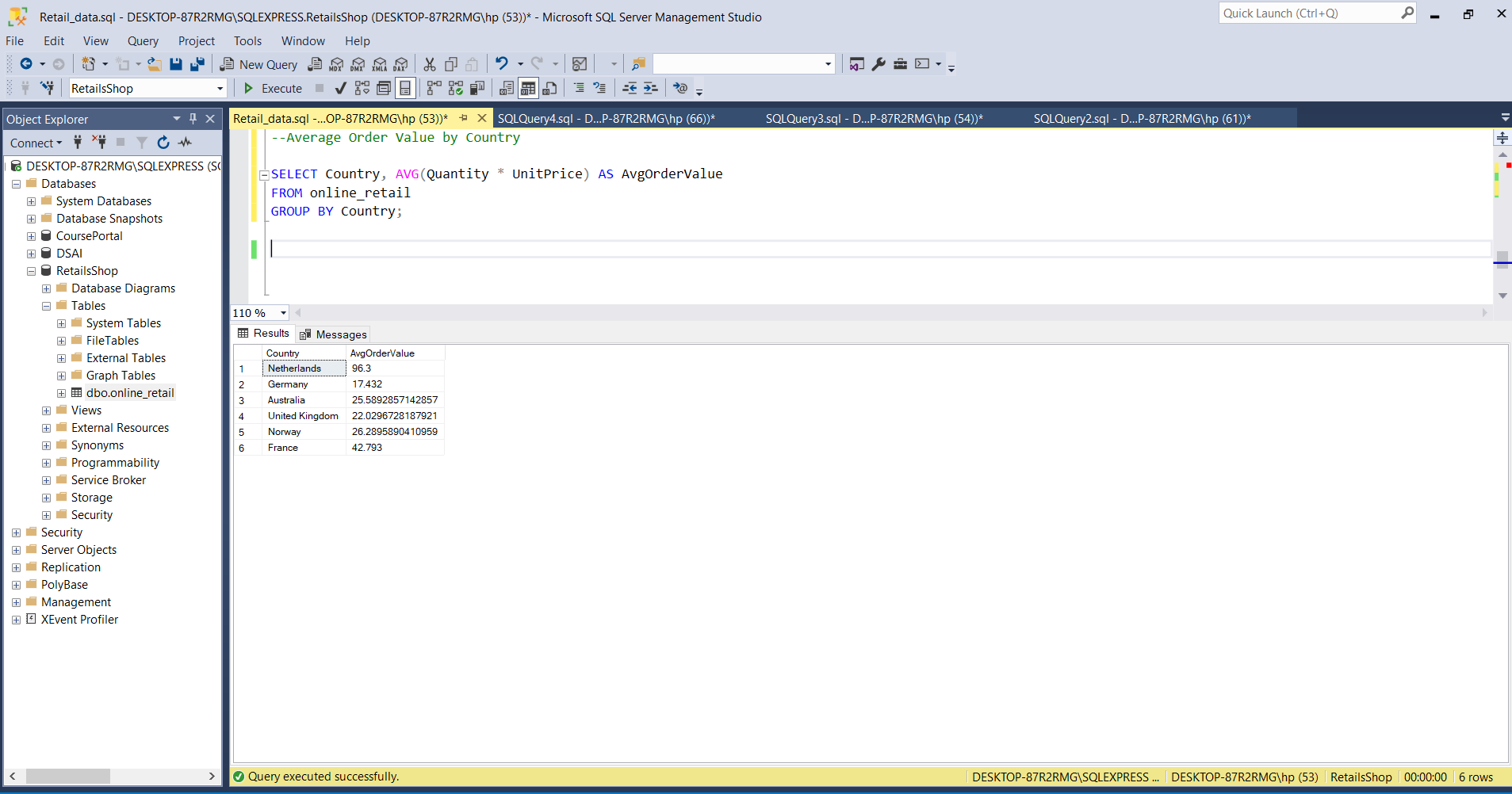
This query categorizes customers based on how frequently they make purchases, grouping them into high, medium, and low-frequency segments.



**Result:**  
The analysis revealed that only a small percentage of customers are high-frequency purchasers, while the majority are low-frequency. This insight can help businesses focus their marketing efforts on increasing the engagement of low-frequency customers.

**7. Average Order Value by Country**

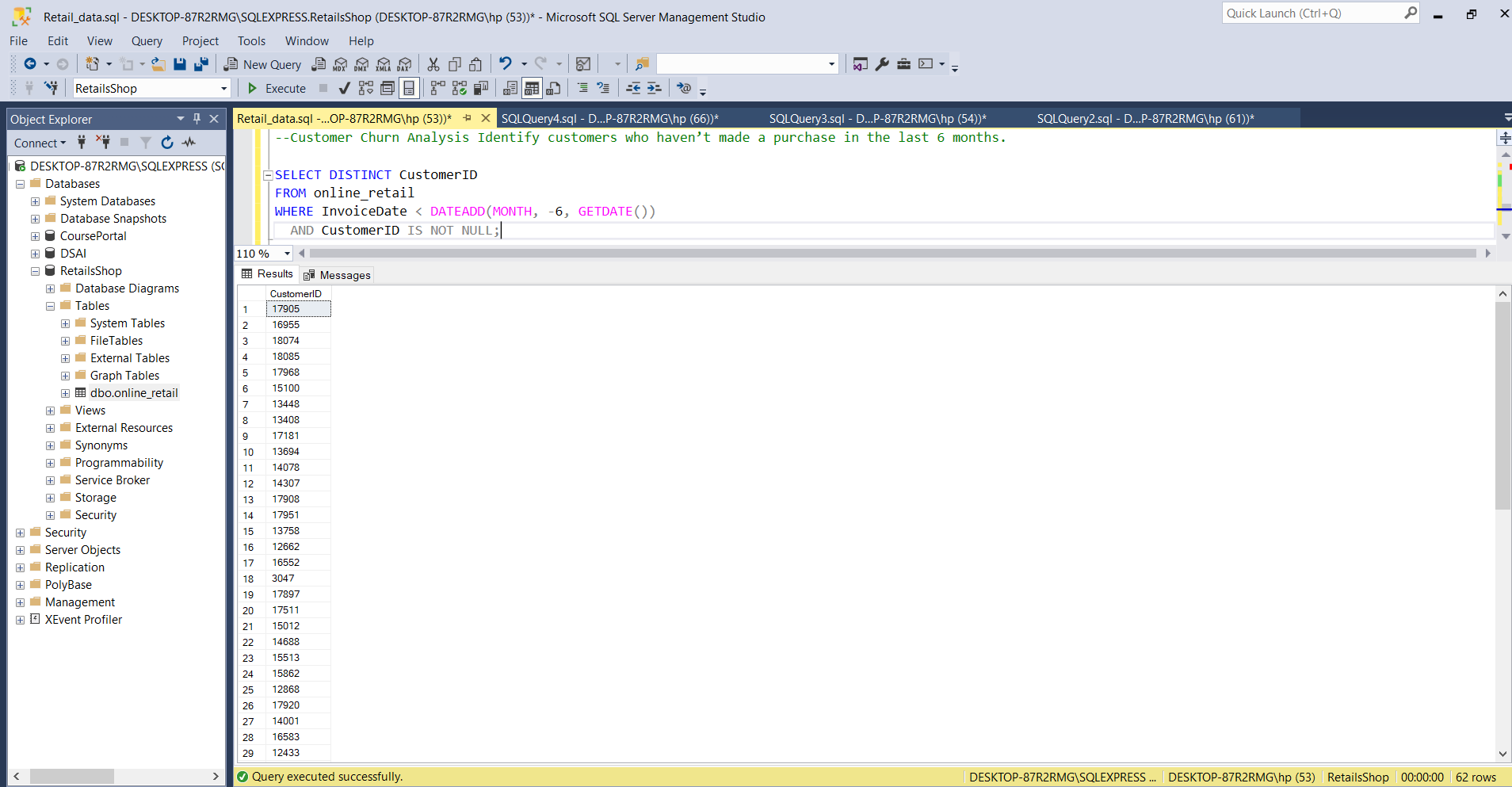
This query calculates the average order value for each country, helping businesses determine where their most valuable customers are located.



**Result:**  
The query results showed that certain countries, such as the UK, have a higher average order value, which indicates a higher purchasing power in those regions.

**8. Customer Churn Analysis**

This query identifies customers who haven't made a purchase in the last 6 months, which is critical for assessing churn risk.

****

**Result:**  
The analysis showed that many customers have not made a purchase in the last 6 months, signaling potential churn. These customers can be targeted with re-engagement campaigns to bring them back.

**Conclusion**

The analysis provided a clear picture of customer behavior through various queries. Key findings include:

* A small group of high-frequency purchasers and a large group of one-time buyers.
* Certain products are frequently purchased together, presenting cross-selling opportunities.
* Customers from specific regions, such as the UK, tend to have higher order values.
* Several customers are at risk of churn, providing an opportunity for targeted re-engagement efforts.

By implementing tailored marketing strategies based on these insights, businesses can increase customer loyalty, improve product bundling, and enhance overall profitability.

Git Hub link = https://github.com/RukhsanaAyub/DSAI\_Course