High Impact Skills

Development Program

AI & Data Science

Shakeel Ahmad

Roll No # GIL-DSAI-061

GitHub: <https://github.com/ShakeelOfficial99>

Data Mining Module

**Online Retail Environment of Customer Behavior Analysis**

**Introduction**

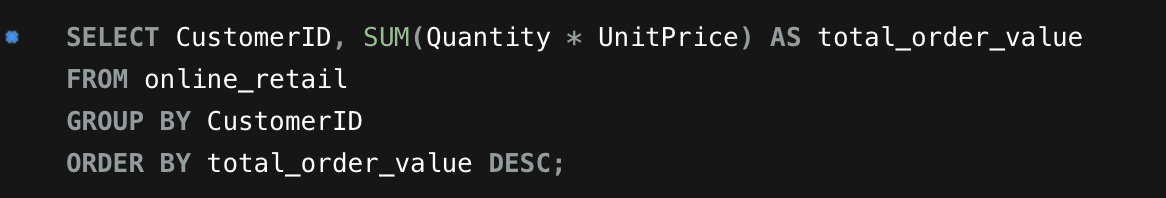
* **Objective:**  
  This project focuses on analyzing various aspects of customer behavior in an online retail environment. Our primary goals are to derive actionable insights through customer segmentation, assess order value distribution, understand product affinity, and identify monthly sales trends. By achieving these objectives, we aim to enhance customer targeting, optimize sales strategies, and improve overall business performance.
* **Tools & Techniques:**  
  The analysis was conducted using SQL queries within a MySQL environment, utilizing the online\_retail schema. This schema provides comprehensive data on customer transactions, enabling detailed and insightful analysis.

**Data Structure and Preparation**

* **Schema Overview:**  
  The online\_retail table includes several key fields:
* **InvoiceNo:** Unique transaction identifier
* **StockCode:** Product identifier
* **Quantity:** Number of units purchased
* **UnitPrice:** Price per unit
* **CustomerID:** Unique customer identifier
* **InvoiceDate:** Date of transaction.
* **Data Transformation:**  
  To ensure analysis accuracy:
* The InvoiceDate column was converted to a proper DATETIME format to facilitate effective time-based analysis, enabling insights into sales trends and seasonal variations.Calculations for order values and customer segmentation were verified for accuracy, ensuring the reliability of subsequent analyses.
* Calculations for order values and customer segmentation were verified for accuracy to ensure the reliability of the analysis.

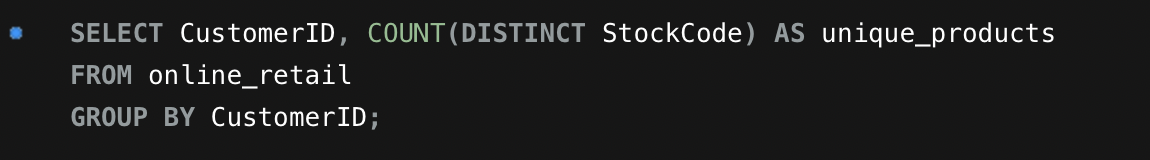
**Key Analysis and Insights**

* **Customer Segmentation by Purchase Frequency:**  
  Customers were segmented based on their purchase frequency into high, medium, and low frequency groups. This segmentation was determined by counting the number of invoices associated with each customer.



* **This segmentation helps in identifying:**
* High-Frequency Customers: Typically loyal customers who frequently engage with the brand.
* Low-Frequency Customers: Individuals who may need targeted strategies to increase their purchase frequency.

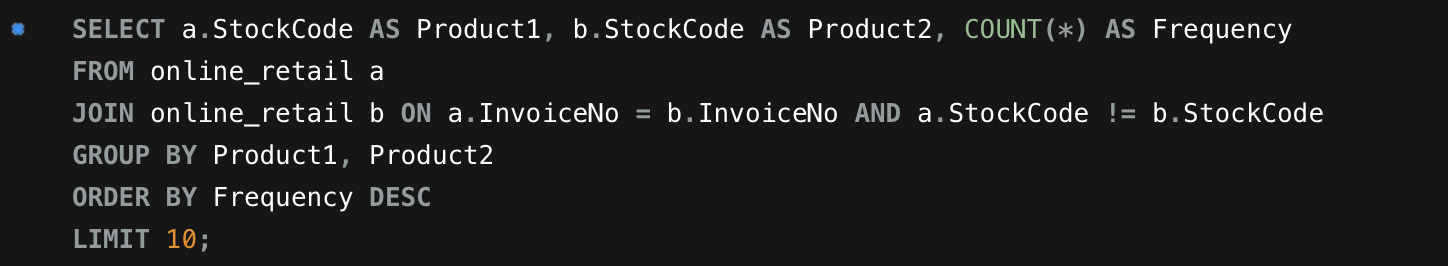
**Order Value Distribution:**  
To understand order value distribution, the total order value for each customer was calculated by multiplying the quantity of items purchased by their unit price.



**This analysis helps in:**

* Identifying high-value customers who contribute significantly to overall revenue.
* Targeting these customers for loyalty programs or special offers.

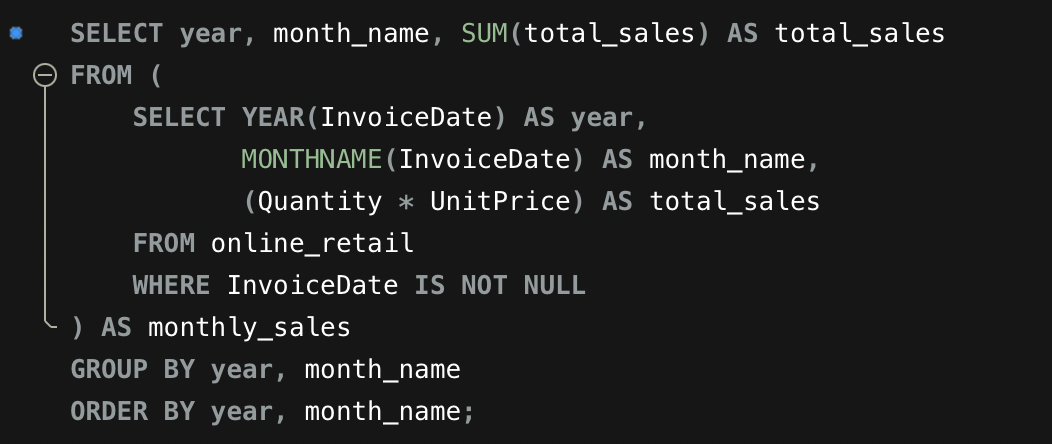
**Product Affinity Analysis:**  
Product affinity was analyzed to identify products frequently purchased together. This provides insights into potential cross-selling and bundling opportunities.



**This analysis reveals:**

* Common product pairs that could be bundled for promotions.
* Opportunities to enhance sales through targeted marketing efforts.

**Monthly Sales Analysis:**  
Total sales were computed on a monthly basis to identify trends and seasonal patterns.



**This analysis provides insights into:**

* Seasonal trends that influence sales.
* Peak periods where marketing efforts and inventory management should be intensified.

**Conclusion and Future Recommendations**

**Key Insights:**

* High-Frequency Customers: They are the primary contributors to overall sales, indicating the importance of nurturing these relationships.
* Product Affinity: Certain products have a high propensity to be purchased together, suggesting opportunities for effective bundling and promotional strategies.
* Monthly Sales Trends: Understanding these trends assists in optimizing inventory levels and marketing strategies to align with peak sales periods.

**Future Recommendations:**

* Customer Churn Analysis: Conduct a detailed analysis of customer churn, focusing on those who have not made purchases in the last six months. This will help in understanding the reasons behind customer disengagement and formulating re-engagement strategies.
* Targeted Marketing Campaigns: Implement targeted marketing campaigns aimed at high-value customers. Personalizing offers and promotions for these individuals can drive further sales and enhance customer loyalty.

By following these recommendations, the online retail business can better serve its customers, improve sales performance, and develop more effective marketing and inventory strategies.