***TITLE: ONLINE RETAIL SEGMENT***

## NAME: Zishan Sher

**Roll number : GIL-DSAI-121**

**[xeeshan-dev/Data-Mining-Project (github.com)](https://github.com/xeeshan-dev/Data-Mining-Project)**

DATE: 6/September/2024

***INTRODUCTION***

# OBJECTIVES:

Understanding of data mining

Run queries in sql for data mining

Implementing queries

Find important insights from data to get more reliable data

Implement different techniques in large scale data for good result

***Overview***

It is a data about a online retail segment .in this dataset there are 8 attributes and approximately 13116 rows.

* **InvoiceNo** Unique transactionidentifier.
* **StockCode:** Unique product code.
* **Description:** Product description.
* **Quantity:** Quantity sold per transaction.

** InvoiceDate:** Transaction date and time.

 **Unit Price:** Price per product.

 **CustomerID:** Unique customer identifier.

**Importance of Segmentation:**

Segmentation play a very vital role in any business .it is use to identify the following;

Distinguish between daily, occasional customers help to tailor strategies effectively.

Focus on high-frequency customers with loyalty programs.

Enable personalize promotion to different groups.

Identify trends and patterns in product sales to make informed business decisions and adapt to market changes.

Try better decision making and improve customer statisfication.

***METHODOLOGY***

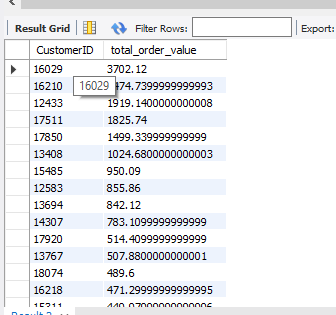
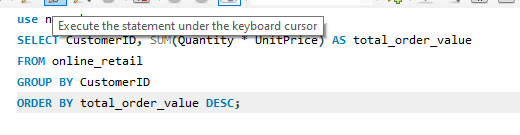
**Data Preparation:**

* **Cleaning:** Address missing values and remove duplicates.
* **Transformation:** Convert date formats and aggregate data where necessary.

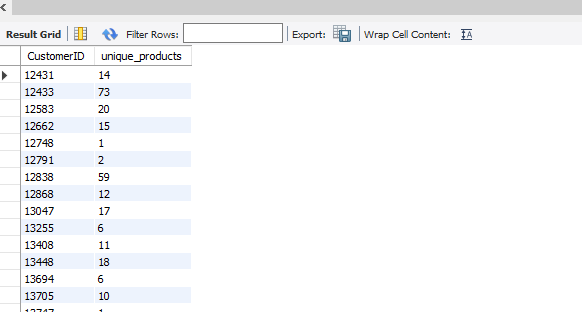
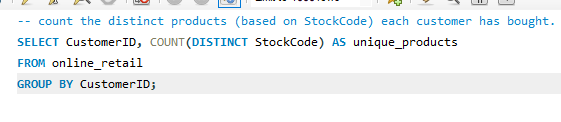
**SQL Queries:**

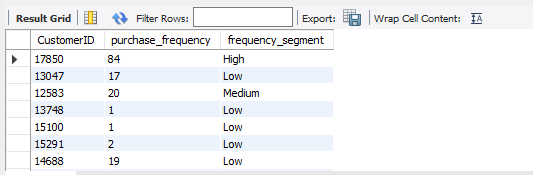
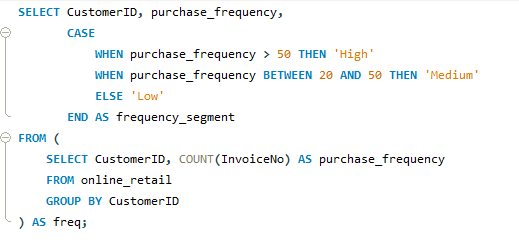
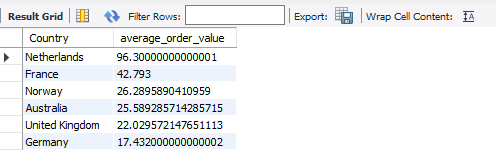
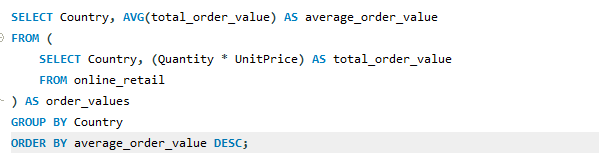
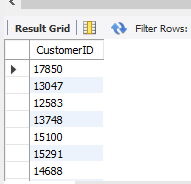
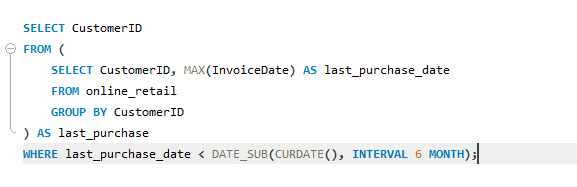
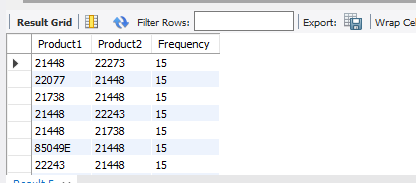
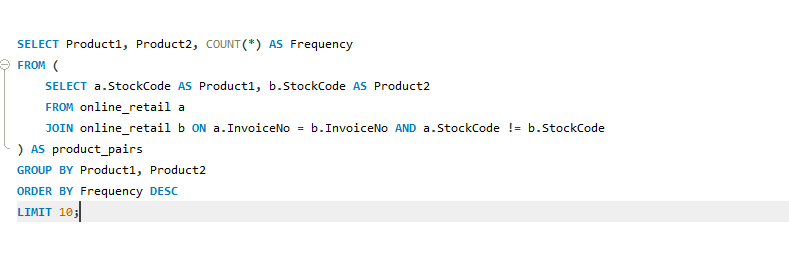
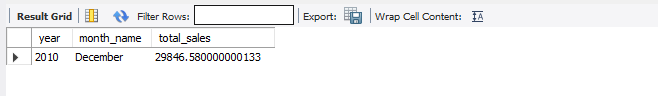
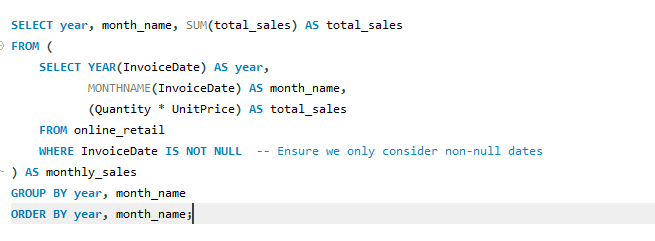
* Beginner Queries:

1



2



* + Products commonly purchased together.
* **Advanced Queries:**
  + **Customer Segmentation by Purchase Frequency:** Classify customers into high, medium, and low frequency segments.
* 
  + **Average Order Value by Country:** Calculate the average order value for each country.
* 
  + **Customer Churn Analysis:** Identify customers who have not made a purchase in the last six months.
  + 
  + **Product Affinity Analysis:** Determine which products are frequently bought together.
* 
  + **Time-based Analysis:** Explore trends in sales patterns over time (monthly or quarterly).
  + 

### ****5. Results****

* **Customer Segmentation:** Identified segments based on purchase frequency revealing insights into customer loyalty and behavior.
* **Average Order Value by Country:** Highlighted countries with higher average order values, indicating where the most valuable customers are located.
* **Customer Churn Analysis:** Found patterns of customer churn allowing for targeted retention strategies.
* **Product Affinity Analysis:** Revealed product associations aiding in product bundling and marketing strategies.
* **Time-based Analysis:** Showed trends and patterns in sales informing inventory management and promotional activities.

### ****6. Conclusion****

Effective segmentation using this dataset has provided valuable insights into customer behavior and purchasing patterns. By understanding customer segments and market trends, the business can tailor its strategies to enhance customer engagement, optimize product offerings, and improve overall performance.

***Recommendations:***

* Implement targeted marketing campaigns based on customer segments.
* Focus on retaining high-value customers and addressing churn.
* Optimize inventory and promotions based on product affinity and sales trends.

**References:**

* Dataset Source: [LMS ]
* SQL Tools Used: [e.g., MySQL Workbench]