**Analysis of Customer Segmentation, Order Value Distribution, and Product Affinity in an Online Retail Database**

* **Name : Anzal Hussain Anzal**
* **Roll number : DSAI-GIT-208**
* **Email : anzaljussainpoetry@gmail.com**

**1 Introduction**

**Objective:**

* **This project focuses on analyzing various aspects of customer behavior in an online retail environment. The aim is to derive meaningful insights through customer segmentation, order value distribution, product affinity analysis, and monthly sales trends.**

**Tools & Techniques:**

* **The project was executed using SQL queries to manipulate and analyze data within the online\_retail schema in a MySQL environment**

**2 Data Structure and Preparation**

**Schema Overview:**

* **The online\_retail table includes fields like InvoiceNo, StockCode, Quantity, UnitPrice, CustomerID, and InvoiceDate.**

**Data Transformation:**

* + **The InvoiceDate column was transformed into a proper DATETIME format to facilitate time-based analysis.**
  + **Ensured the accuracy of calculations for order values and segmentation.**

**3 Key Analysis and Insights**

**Customer Segmentation by Purchase Frequency**:

* + **Customers were grouped into high, medium, and low purchase frequency segments based on the number of invoices associated with each customer.**
  + **This segmentation helps identify loyal customers (high frequency) and those who may need more attention (low frequency).**

**SELECT CustomerID, COUNT(InvoiceNo) AS purchase\_frequency**

**FROM online\_retail**

**GROUP BY CustomerID;**

**Order Value Distribution:**

* **Total order values were calculated for each customer by multiplying Quantity by UnitPrice. This helped identify high-value customers.**

**SELECT CustomerID, SUM(Quantity \* UnitPrice) AS total\_order\_value**

**FROM online\_retail**

**GROUP BY CustomerID;**

**Product Affinity Analysis:**

* **This query identified products frequently bought together, offering insights into cross-selling and bundling opportunities.**

**SELECT a.StockCode AS Product1, b.StockCode AS Product2, COUNT(\*) AS Frequency**

**FROM online\_retail a**

**JOIN online\_retail b ON a.InvoiceNo = b.InvoiceNo AND a.StockCode != b.StockCode**

**GROUP BY Product1, Product2;**

**Monthly Sales Analysis:**

* **Total sales per month were calculated to identify trends and seasonality**

**SELECT YEAR(InvoiceDate) AS year,**

**MONTHNAME(InvoiceDate) AS month\_name,**

**SUM(Quantity \* UnitPrice) AS total\_sales**

**FROM online\_retail**

**GROUP BY year, month\_name;**

**4 Conclusion and Future Recommendations**

**Key Insights:**

* + **The high-frequency customers contribute the most to overall sales.**
  + **Certain product pairs have a strong correlation and can be bundled for promotional offers.**
  + **Monthly sales trends help in planning inventory and marketing strategies during peak seasons.**

**Future Recommendations:**

* + **Conduct further analysis on customer churn (customers with no purchases in the last 6 months).**
  + **Implement targeted marketing campaigns for high-value customers.**