Project Brief

Project Title: Customer Purchase Behavior and Sales Trend Analysis for Online Retail Optimization

Group: 3

Project Overview:

Analyze the **Online Retail Dataset** from the UCI Machine Learning Repository(https://archive.ics.uci.edu/dataset/352/online+retail) , which contains transactional data for a UK-based online retail company that sells gifts and other products. The objective of this project is to gain insights into customer purchase behavior and provide actionable recommendations for improving sales strategies, identifying key customers, and understanding product performance.

Dataset Description:

The dataset contains transactional data from December 2010 to December 2011, and includes the following key columns:

- **InvoiceNo:** Unique identifier for each transaction.
- **StockCode:** Unique identifier for each product.
- **Description:** Product description.
- **Quantity:** Number of units sold.
- **InvoiceDate:** Date and time of the transaction.
- **UnitPrice:** Price per unit.
- **CustomerID:** Unique identifier for each customer.
- **Country:** The country where the customer is located.

This dataset offers a rich source of information for customer segmentation, sales performance analysis, and understanding product popularity.

Project Goals:

The aim of this project is to analyze customer purchasing patterns, identify top-performing products, and explore sales trends across different markets. Your team will provide recommendations on how the company can enhance its customer engagement and optimize sales.

Project Tasks:

Your tasks will involve querying the data using SQL to provide insights into key business metrics. The specific tasks include:

1. **Customer Segmentation by Purchase Frequency:**
   o Group customers based on their purchase frequency (e.g., one-time, repeat, and high-frequency buyers).

o Provide insights into how much revenue is generated from each segment and suggest potential engagement strategies.
2. **Top 10 Most Purchased Products:**
   o Identify the top 10 most purchased products based on total quantity sold.
   o Provide a breakdown of which countries purchase these products the most.
3. **Revenue Analysis by Country:**
   o Calculate total sales revenue for each country, identifying the top 5 countries that contribute the most to overall sales.
   o Provide insights into customer behavior in these countries and suggest opportunities for market expansion.
4. **Monthly Sales Performance:**
   o Analyze monthly sales data to identify seasonal trends and key periods of high or low sales.
   o Provide recommendations on how to capitalize on peak seasons or improve sales during slower months.
5. **Customer Lifetime Value (CLV) Analysis:**
   o Calculate the Customer Lifetime Value (CLV) by analyzing the total revenue generated by repeat customers over the dataset's time span.
   o Identify the top 5 customers based on total sales value and provide an analysis of their purchasing behavior (e.g., preferred products, purchasing frequency).
6. **Product Performance Analysis by Category (if available):**
   o Analyze which product categories (if categories can be derived from descriptions or stock codes) generate the most sales revenue.
   o Provide insights into product trends and suggest which categories should be prioritized for sales and marketing efforts.

Expected Deliverables:

1. **SQL Queries:**
   o Submit all SQL queries used to perform the analysis for each task.
   o Ensure the queries are properly commented, with clear explanations of the logic behind each step.
2. **Data Insights Report:**
   o Write a report summarizing the key findings from the analysis. The report should include insights on customer behavior, product performance, and sales trends.
   o Use visualizations (charts, graphs) to support your findings, especially for time-based trends and product comparisons.
3. **Presentation:**
   o Prepare a 5-10 slide presentation summarizing your findings and recommendations.
   o Be prepared to present your approach, analysis, and business insights derived from the data.

Skills and Concepts to Apply:

- **SQL Queries:** Use SQL to retrieve and manipulate data, using joins, aggregate functions (SUM, COUNT), GROUP BY, ORDER BY, window functions e.t.c

- **Customer Segmentation:** Segment customers based on their behavior (e.g., purchase frequency, total spend).
- **Sales Metrics:** Calculate total sales, average order value, customer lifetime value, and product performance metrics.
- **Trend Analysis:** Analyze sales trends over time (e.g., monthly sales) and across geographies (e.g., country-level sales performance).
- **Reporting:** Summarize insights clearly and provide actionable business recommendations based on the data.

Tools:

- SQL Server
- Data visualization tools like Excel, Power BI, or Tableau for enhancing the report and presentation.

Evaluation Criteria:

- Correctness and complexity of the SQL queries used in analysis.
- Quality of insights derived from customer, product, and sales data.
- Clarity and effectiveness of the report and presentation.
- Application of advanced SQL techniques (e.g., window functions, subqueries).
- Team collaboration and communication during the presentation.