Elevvo Internship 2025/TASK-02

Submitted By:Aleem Shoukat
Submitted to: Elevvo Internship 2025(Data Analytics)
Date:23/Aug/2025

TASK-02 Customer Sales Analysis Report

1. Introduction

This report presents an end-to-end analysis of customer sales data. The purpose of this task was to explore, clean, and analyze sales records to gain useful insights into customer purchasing behavior. The analysis also helps identify trends, customer value, and sales performance across different countries and months.

2. Dataset Overview

We used an **online retail dataset**, which contains transaction-level sales information.

The dataset includes the following key fields:

- **InvoiceNo** – Transaction number.
- **StockCode** – Product code.
- **Description** – Product description.

- **Quantity** – Number of items purchased.
- **InvoiceDate** – Date of transaction.
- **UnitPrice** – Price per unit of the product.
- **CustomerID** – Unique customer identifier.
- **Country** – Country where the transaction took place.
- **TotalSales** – (calculated) Revenue from each transaction (Quantity × UnitPrice).
- **Date/Month/Year** – Extracted from InvoiceDate for time-based analysis.

This dataset was prepared and cleaned before starting analysis by removing missing values, handling duplicates, and creating new calculated columns.

---

3. Data Preparation and Cleaning

To make the dataset ready for analysis, the following steps were performed:

- Removed transactions with missing or invalid **CustomerID** values.
- Removed duplicate invoices.
- Calculated a new column **TotalSales** for transaction revenue.
- Extracted **Year, Month, Week, and YearMonth** from InvoiceDate for trend analysis.
- Filtered out records with negative quantities or zero sales values.

This ensured the dataset was consistent and ready for exploration.

---

4. Exploratory Data Analysis (EDA)

We performed EDA to understand the sales distribution, trends, and customer patterns.

Key steps included:

- **Top Countries by Sales** → Identified which countries contributed most to overall revenue.
- **Monthly Sales Trend** → Observed how sales changed over time (growth/decline).
- **Product-Level Analysis** → Identified the most sold and most profitable products.
- **Customer Segmentation** → Grouped customers based on their purchasing behavior (using RFM Analysis).

---

5. RFM Analysis (Customer Segmentation)

RFM (Recency, Frequency, Monetary) analysis was performed to measure customer value.

- **Recency (R)** → How recently a customer made a purchase.
- **Frequency (F)** → How often a customer purchased.
- **Monetary (M)** → How much revenue the customer generated.

Summary of RFM results:

- Customers with **low Recency and high Frequency** → Loyal customers.
- Customers with **high Recency and low Frequency** → At-risk or inactive customers.
- Customers with **high Monetary value** → High-value customers driving business revenue.

This analysis helped identify the best customers and those who need re-engagement strategies.

---

6. Data Visualization

To make insights clearer, we used several visualizations:

- **Bar Charts** → Sales by country and top-selling products.
- **Line Charts** → Monthly sales trends.
- **Histograms** → Distribution of Recency, Frequency, and Monetary values.
- **Heatmaps** → Correlation between sales variables.
- **RFM Distribution Graphs** → To visualize customer groups.

These visualizations make it easier for management and non-technical stakeholders to understand the insights.

---

7. Conclusion & Insights

From this analysis, we observed:

- A few countries contribute the majority of sales revenue.
- Monthly sales patterns show clear seasonality and growth opportunities.
- RFM analysis identified a small group of loyal, high-value customers who drive most revenue.
- Many customers purchase only once, meaning there's potential to improve retention.

These findings can guide **marketing campaigns, customer retention programs, and product strategies**.
