

Project Report: Customer Segmentation using RFM and Cohort Analysis

1. Project Overview

The goal of this project is to **analyze customer purchase behavior**, segment customers based on their **value and engagement**, and study **retention trends** using **Cohort Analysis** and **RFM (Recency, Frequency, Monetary) Analysis**.

This helps businesses:

- Identify **high-value customers**
- Understand **customer loyalty and retention patterns**
- Plan **targeted marketing and engagement strategies**

2. Dataset Description

- **Dataset Used:** Online Retail Dataset
- **File:** Online Retail.xlsx
- **Columns:**
 - InvoiceNo
 - StockCode
 - Description
 - Quantity
 - InvoiceDate
 - UnitPrice
 - CustomerID
 - Country

3. Data Preprocessing

1. Removed rows with missing CustomerID.
2. Removed duplicate records.
3. Filtered out transactions with negative or zero Quantity or UnitPrice.
4. Created new columns for analysis:
 - a. InvoiceMonth – month of transaction
 - b. CohortMonth – month of first purchase
 - c. CohortIndex – number of months since first purchase
 - d. TotalPrice – Quantity × UnitPrice for RFM calculation

4. Cohort Analysis

Objective: Understand **customer retention trends** over time.

- Calculated **cohorts** based on first purchase month.
- Created **retention tables** showing % of customers returning in subsequent months.
- Visualized retention using:
 - **Heatmap** – shows retention rate per cohort
 - **Average retention line chart** – overall trend of retention

5. RFM Analysis

Objective: Segment customers based on **purchase behavior and value**.

- Calculated **Recency** (days since last purchase), **Frequency** (number of purchases), and **Monetary** (total spending).
- Combined RFM scores to create customer segments.

Key Insights:

- **High-value customers** can be targeted for loyalty programs.
- **At-risk customers** may need re-engagement campaigns.
- Top 10 customers contributed significantly to revenue.

7. Conclusion

This project demonstrates how **Cohort Analysis** and **RFM segmentation** can help businesses:

- Identify **loyal and high-value customers**
- Detect **drop-off trends**
- Plan **data-driven marketing strategies**

By combining retention analysis and RFM, businesses can improve **customer engagement** and **maximize revenue**.

8. References

1. Online Retail Dataset
2. Python libraries used: pandas, matplotlib, seaborn, openpyxl