

# Project Documentation: Customer Segmentation Using RFM Analysis

## Introduction

This project aims to segment customers of a business using **RFM (Recency, Frequency, Monetary)** analysis to identify high-value groups, optimize marketing strategies, and improve customer retention. By analysing transactional data, I categorized customers into distinct clusters to understand their behaviour and contribution to revenue.

---

## Objective

1. **Identify customer segments** based on purchasing behaviour.
  2. **Prioritize high-value customers** for retention and **reactivate at-risk groups**.
  3. **Optimize marketing budgets** by tailoring strategies to each segment.
- 

## Methodology

1. **Data Preparation:** Cleaned raw transactional data in Excel, removing duplicates and outliers.
  2. **RFM Scoring:** Calculated Recency (days since last purchase), Frequency (total transactions), and Monetary (total spend) for each customer.
  3. **Clustering:** Applied the Elbow Method in R to determine the optimal number of clusters (4 clusters).
  4. **Analysis & Visualization:** Analysed clusters in Power BI to assign labels and quantify their impact on revenue.
- 

## Key Findings

1. **Low-Value Customers (Cluster 1):**
  - **25% of customers**, contributing **2% (\$0.37M)** of total sales.
  - Low RFM scores: Infrequent, low-spending, and inactive.
  - **Action:** Deprioritize or target with low-cost reactivation campaigns.
2. **At-Risk Customers (Cluster 2):**
  - **21.7% of customers**, contributing **1.97% (\$0.36M)** of sales.
  - High recency (recently purchased) but low frequency and spend.
  - **Action:** Engage with personalized offers to prevent churn.
3. **Loyal/High-Value Customers (Cluster 3):**

- **33.9% of customers**, driving **63.95% of revenue**.
  - High RFM scores: Frequent, high-spending, and recent purchasers.
  - **Action**: Reward with loyalty programs; focus on retention.
4. **Big Spenders (Cluster 4)**:
- **18.9% of customers**, contributing **32% of sales**.
  - High monetary spend but low recency and frequency.
  - **Action**: Re-engage with win-back campaigns (e.g., limited-time offers).
- 

## Recommendations

1. **Reallocate Resources**: Shift focus to **Cluster 3 (Loyal/High-Value)** and **Cluster 4 (Big Spenders)** to maximize ROI.
  2. **Reactivate At-Risk Customers (Cluster 2)**: Use email/SMS campaigns with discounts on repeat purchases.
  3. **Cost-Effective Retention**: For **Cluster 1 (Low-Value)**, phase out high-cost outreach unless reactivation tests prove viable.
  4. **Monitor Big Spenders**: Target Cluster 4 with exclusive offers to incentivize repeat purchases.
- 

## Tools Used

- **Excel**: Data cleaning and RFM scoring.
- **R Programming**: K-means clustering with the Elbow Method.
- **Power BI**: Cluster analysis, visualization, and reporting.