Customer Segmentation Project

Internship Deliverable Presentation

Using RFM and KMeans Clustering

Objective

- The goal of this project is to segment customers based on purchasing behavior using RFM (Recency, Frequency, Monetary) analysis and KMeans clustering.

Step 1: Data Preparation

- - Loaded customer transaction data
- - Cleaned and formatted data
- - Ensured correct data types and removed missing values

Step 2: RFM Metric Calculation

- - Recency: Days since last purchase
- - Frequency: Number of purchases
- - Monetary: Total amount spent
- - Calculated RFM values for each customer

Step 3: RFM Summary

- - Count: 500 customers
- - Average Recency: 71.31 days
- - Average Frequency: 30.29 purchases
- - Average Monetary: $6279.16

Step 4: Scaling

- - Applied MinMaxScaler to normalize RFM values
- - Necessary for distance-based clustering like KMeans

Step 5: Choosing K

- - Used Elbow Method and Silhouette Score
- - Elbow suggested k=4, Silhouette suggested k=3
- - Final choice: k=3 for better cohesion

Step 6: KMeans Clustering

- - Applied KMeans with 3 clusters
- - Assigned cluster labels to each customer

Step 7: PCA Visualization

- - Used PCA to reduce dimensions for plotting
- - Scatter plot showed clear cluster separation

Step 8: Segment Profiling

- - Cluster 0: High Recency, Low Spend (Low Value)
- - Cluster 1: Mid Recency, High Spend (Top Customers)
- - Cluster 2: Recent Purchasers, Average Spend (Loyal)

Step 9: Recommendations

- - Cluster 0: Win-back campaigns
- - Cluster 1: Loyalty rewards & VIP offers
- - Cluster 2: Retain with regular engagement

Deliverables

- - Jupyter Notebook with all analysis
- - CSV file with cluster labels
- - This Presentation with full explanation