CREDIT CARD CUSTOMER SEGMENTATION

Project Description

Credit Card Customer Segmentation using KMeans

Performed customer segmentation on financial behaviour data using KMeans clustering. Conducted EDA, analyzed skewness ,handled missing values with median imputation, applied feature scaling, identified optimal clusters using Elbow and Silhouette methods, visualized clusters using PCA, and derived actionable customer segments for marketing and risk analysis.

Project Implementation

1.**Data collection**- kaggle dataset

2.**Data preprocessing**

- ❖ Handle missing values
- ❖ Scale data using StandardScaler

Clustering Models

KMeans Clustering

1.Choose optimal K using:

- Elbow method
- Silhouette score

2.Cluster Interpretation & Naming

3. Cluster Profiling

Segment Interpretation

Cluster	Cluster Name	Business Insight
Cluster 0	High-Risk Revolving Credit Users	High balance, high cash usage

Cluster	Cluster Name	Business Insight
Cluster 1	Low-Usage / Inactive Customers	Inactive, low spend
Cluster 2	High-Value Active Customer	Best customers
Cluster 3	Cash-Dependent Financially Stressed Customers	Risky, low repayment

Cluster 0 – High-Balance, Cash-Advance Heavy Customers (RISKY)

BALANCE: 5010 (highest)
CASH_ADVANCE: 5065 (very high)
CASH_ADVANCE_FREQUENCY: 0.51
PRC_FULL_PAYMENT: 0.04 (very low)
CREDIT_LIMIT: 8209 (highest)

Cluster 1 – Low Usage, Average Payers (INACTIVE / BASIC)

BALANCE: 108 (very low)
PURCHASES: 350
PURCHASES_TRX: 4.8
PRC_FULL_PAYMENT: 0.25
CASH_ADVANCE: very low

Cluster 2 – High value Active Customer

PURCHASES: 2182 (highest)
PURCHASES_TRX: 32.7 (very high)
PURCHASES_FREQUENCY: 0.93
INSTALLMENTS_PURCHASES: high
PRC_FULL_PAYMENT: 0.28

Cluster 3 – Cash-Dependent but Low Spend

CASH_ADVANCE: 782

PURCHASES: 283
PAYMENTS: 947
PRC_FULL_PAYMENT: 0.02 (lowest)
CREDIT_LIMIT: low

- **Cluster 0** represents high-balance customers with frequent cash advances and poor full-payment behaviour, indicating higher credit risk.
- **Cluster 1** includes low-usage customers with minimal balances and moderate payment discipline, representing inactive or casual cardholders.
- **Cluster 2** represents highly active, high-spending customers with frequent transactions and strong engagement, making them ideal for premium offerings.
- **Cluster 3** contains cash-dependent customers with low spending and poor payment behaviour, indicating potential financial distress

Business Action:

This project supports data-driven decision making.

Cluster 0

- Risk monitoring
- Payment reminders
- Limit control strategies

Cluster 1

- Activation campaigns
- Fee-free usage incentives
- Cashback offers

Cluster 2

- Premium rewards

❖ Loyalty programs
❖ Credit limit increases

Cluster 3

❖ Financial counselling
❖ Spending controls

Streamlit Application

Using Streamlit created Components:

- Dataset preview
- Cluster visualization using PCA
- Customer segment summary