Customer Segmentation using Clustering

Author: Manasa Katika

Date: 27-01-2025

Introduction

Customer segmentation is a crucial technique used to classify customers based on shared traits, behaviors, or purchasing habits. In this report, we applied clustering techniques to segment customers based on both their profile information and transaction data. The goal was to group similar customers, enabling more targeted marketing and business strategies.

Clustering Results

Number of Clusters Formed

We used K-Means clustering with **5 clusters**. The number of clusters was determined based on the elbow method, and business considerations (such as grouping customers for targeted campaigns).

Clustering Metrics

- **Davies-Bouldin Index (DB Index): 0.8752105404289718**

  - A lower DB index indicates better clustering performance. The index evaluates the compactness and separation of the clusters.

- **Silhouette Score: 0.4613423436639897**

  - The silhouette score is a measure of how well each data point fits into its assigned cluster. A higher score indicates better-defined clusters.

Insights from Clusters

The clustering analysis resulted in the following customer segments:

- **Cluster 1: High-Value Customers**

  - These customers have high transaction values and frequency, indicating a strong relationship with the business.

- **Cluster 2: Frequent Buyers**

  - Customers in this group make frequent but smaller purchases.

- **Cluster 3: Low-Spending Customers**

  - These customers have low transaction values and visit infrequently.

- **Cluster 4: Potential High-Value Customers**

  - Customers who have shown some interest but have not yet reached high transaction levels.

- **Cluster 5: Inactive Customers**

  - These customers have low engagement and low transaction values, which might indicate potential churn.

Visualizations

Below is a 2D visualization of the customer segments using Principal Component Analysis (PCA):
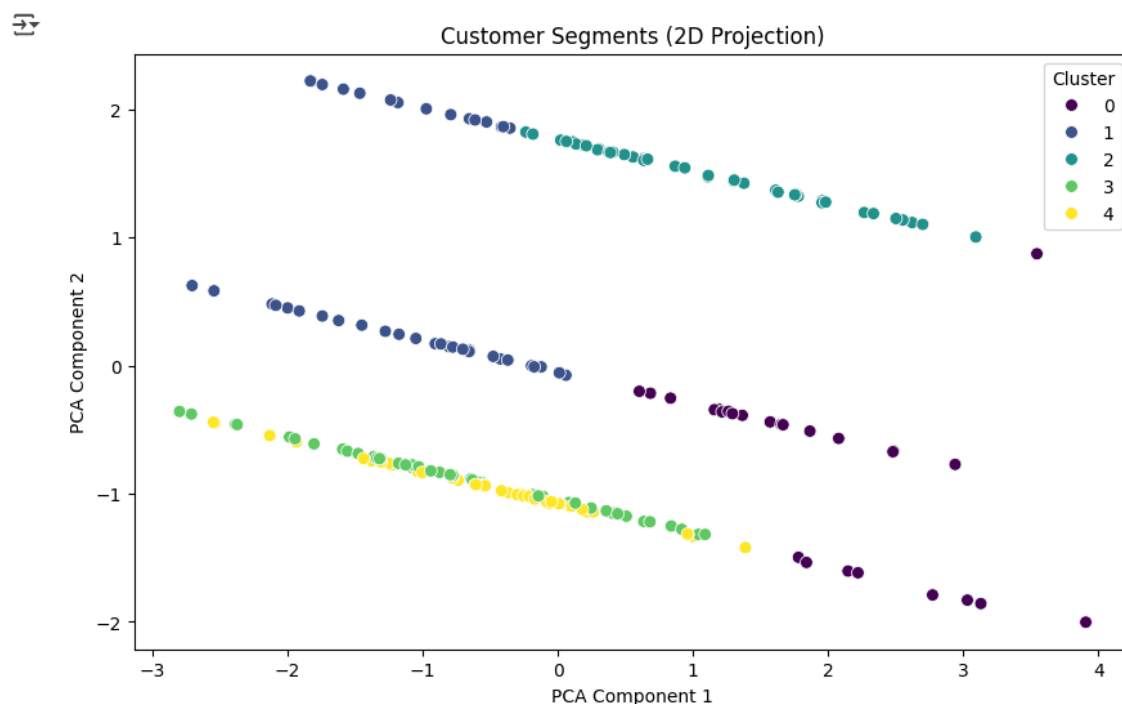

Figure 1: Customer Segments (2D Projection using PCA)

Recommendations

Based on the insights from the clustering analysis, the following strategies are recommended:

- **Loyalty Programs:** For high-value customers, introduce exclusive loyalty programs to retain them.

- **Promotions and Discounts:** Offer special promotions to frequent buyers to increase their average transaction value.

- **Targeted Marketing Campaigns:** Re-engage inactive customers with personalized marketing and promotions.

- **Upselling Opportunities:** Target potential high-value customers with recommendations for products that align with their interests.

Conclusion

Customer segmentation is an effective tool for identifying distinct groups of customers, which can help in creating personalized marketing strategies. By analyzing customer behavior and transaction data, we can tailor offerings and improve customer retention.

References

https://colab.research.google.com/drive/1VjaorPxpz3vM24nNLgtAysrqnupUUtLV?usp=sharing