

# Customer Segmentation and Clustering Report

## 1. Introduction

This report presents the results of customer segmentation performed using clustering techniques. The segmentation utilizes customer profiles and transaction data to group similar customers together. The goal is to identify patterns and similarities that can help businesses understand their customers better.

## 2. Methodology

- Data Used: Customers.csv and Transactions.csv
- Preprocessing: Data cleaning, handling missing values, and feature scaling were performed.
- Clustering Algorithm: K-Means clustering was chosen based on initial analysis.
- Number of Clusters: The optimal number of clusters was determined using the Elbow Method.
- Evaluation Metrics: Davies-Bouldin Index (DB Index) was used as the primary metric.

## 3. Clustering Results

- Number of Clusters Formed: X clusters
- DB Index Value: Y (lower values indicate better cluster separation)
- Additional Metrics:
  - Silhouette Score: Measures similarity within clusters.
  - Cluster Distribution: Proportion of customers in each cluster.
  - Transaction Trends Per Cluster: Purchasing behavior per group.

## 4. Cluster Insights

1. High-Value Customers Cluster: Frequent, high-value transactions.
2. Moderate Buyers Cluster: Occasional, mid-range transaction values.
3. Low-Engagement Customers Cluster: Minimal transactions.
4. Newly Acquired Customers Cluster: Recent sign-ups with growth potential.
5. Discount-Driven Shoppers Cluster: Customers engaging in discounted purchases.

## 5. Visual Representation

- Elbow Method Plot: Showing optimal K selection.

- Scatter Plot of Clusters: Customer group visualization.
- Histogram of Transaction Volumes per Cluster: Purchasing frequency.

## **6. Conclusion & Recommendations**

- The segmentation identifies distinct customer groups for better marketing strategies.
- Retention strategies for high-value customers are recommended.
- Discount shoppers may benefit from targeted offers.
- Exploring alternative clustering algorithms and feature engineering for better segmentation.