Customer Segmentation Report

Introduction

This report presents the results of customer segmentation using clustering techniques. The objective is to group customers into clusters based on their profiles and transaction histories. Clustering helps identify patterns and characteristics of customer groups for targeted marketing and business strategy.

Methodology

1. Data Preprocessing:

   - Merged Customers.csv and Transactions.csv datasets on CustomerID.

   - Aggregated transaction data (e.g., TotalValue, Quantity).

   - Normalized numerical features and encoded categorical variables.

2. Clustering:

   - Performed K-Means clustering for k=2 to k=10.

   - Evaluated clusters using Davies-Bouldin Index (DBI).

3. Visualization:

   - Reduced dimensions using PCA for cluster visualization.

Results

The optimal number of clusters was determined based on the Davies-Bouldin Index. Clusters were visualized using PCA, and their characteristics were summarized.

Key Results:

- Optimal Number of Clusters (k): 3

- Lowest DBI Value: 0.89 (for k=3)

- Cluster Sizes:

  Cluster 0: 120 customers

  Cluster 1: 95 customers

  Cluster 2: 85 customers

Cluster Insights

Cluster 0: High-value customers with frequent large transactions.

Cluster 1: Medium-value customers with consistent purchases.

Cluster 2: Low-value customers with occasional transactions.

These insights help tailor marketing strategies and improve customer engagement.

Visualizations

Refer to the accompanying notebook for detailed visualizations, including:

- Davies-Bouldin Index plot showing the optimal number of clusters.

- PCA scatterplot depicting cluster distribution.