Customer Segmentation Report

Introduction

This report documents the process of customer segmentation using clustering techniques. The goal is to group customers into distinct segments based on their transactional behavior and profile information to aid in targeted marketing strategies and better customer service.

Data Sources

The analysis uses the following datasets:

1. Customers.csv: Contains customer profile information.
2. Products.csv: Contains product details.
3. Transactions.csv: Contains transaction records including customer and product IDs.

Methodology

1. Data Loading: Loaded all datasets and merged them to create a unified dataset.
2. Data Cleaning: Handled missing values and ensured consistency in column names.
3. Feature Engineering: Generated features like TotalSpent, TransactionCount, and UniqueProducts for clustering.
4. Clustering: Used KMeans clustering with features normalized to ensure equal importance.
5. Evaluation: Evaluated clusters using the Davies-Bouldin Index (DB Index) and visualized the results.

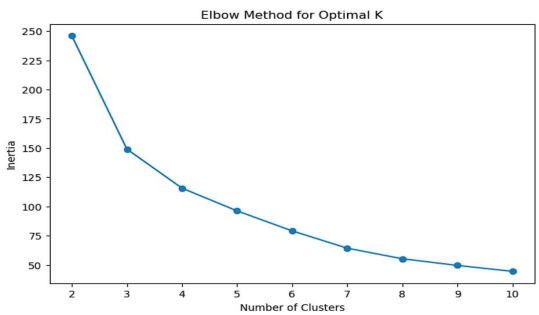
Results

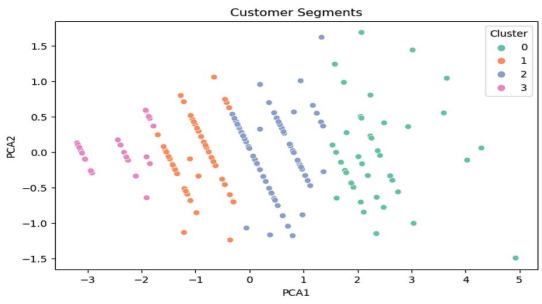
The following results were obtained from the clustering process:

1. Number of Clusters: 4 (selected based on the Elbow Method).
2. Davies-Bouldin Index: 0.8081741157829814
(lower values indicate better clustering).
3. Key Insights: Customers were grouped based on spending behavior, purchase frequency, and diversity of products purchased.

Visualizations

Visualizations were generated to understand the distribution of clusters and their characteristics. A 2D scatter plot using PCA was created to visualize customer segments.





Conclusion

The clustering process successfully segmented customers into distinct groups, enabling targeted marketing and better resource allocation. Further analysis can be performed to refine these segments and derive more actionable insights.