Clustering Report

1. Introduction

- Customer segmentation involves grouping customers into clusters based on shared characteristics. This helps in tailoring marketing strategies and improving customer satisfaction.
- Objective: The goal is to segment customers into distinct groups using clustering techniques, considering both their profile information and transaction behavior.

---

2. Methodology

Data Preparation

- Customer data from `Customers.csv` was enriched with transactional data from `Transactions.csv` and product details from `Products.csv`. This provided comprehensive information for clustering, including demographic and purchase behavior metrics.
- Numerical features were scaled using StandardScaler, and categorical features were encoded appropriately.

Clustering Algorithm

- KMeans was selected due to its efficiency and interpretability for segmenting customers into distinct groups.
- The optimal number of clusters was determined using the Elbow Method and silhouette scores.

---

3. Results

Cluster Overview

Evaluation Metrics

- Silhouette Score: 0.414
- The Davies-Bouldin Index (DBI) for the clusters was calculated as 0.772, indicating moderate cluster separation.

Visualizations

- Scatter plots were used to visualize the clusters based on principal components.

---

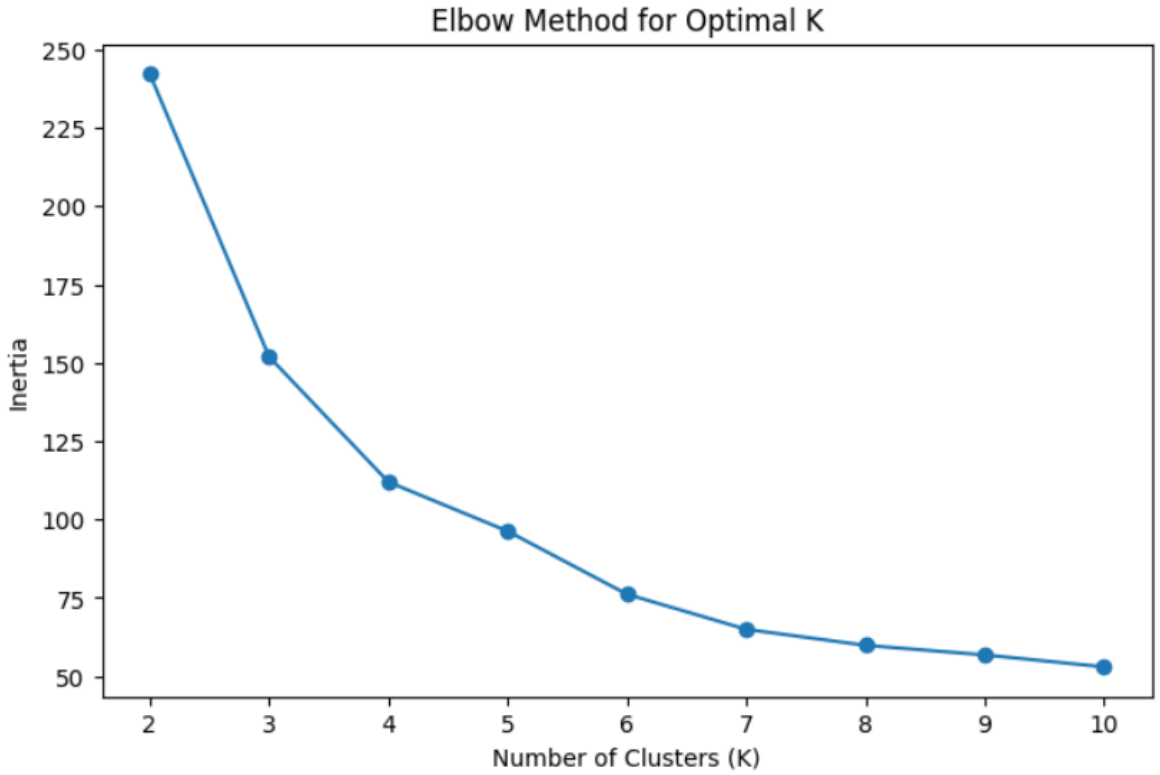
4. Insights and Recommendations

- Cluster 1 customers are highly valuable and should be targeted with loyalty programs.
- Cluster 3 customers can be encouraged to increase spending with personalized discounts.
- Cluster 4 represents untapped potential and requires engagement strategies.

---

5. Conclusion

- Clustering successfully identified customer segments with distinct characteristics, enabling targeted marketing strategies.
- This segmentation can guide personalized marketing, product recommendations, and customer retention efforts.




```
Silhouette Score: 0.4143182730220664
Davies-Bouldin Index: 0.7726853895682169
```