Clustering Report

1. Introduction

This report presents the results of a customer segmentation analysis using clustering techniques to group customers based on their profile and transaction data. The objective was to uncover meaningful customer segments to guide marketing strategies and enhance customer engagement.

2. Methodology

Data Preparation

The analysis utilized two datasets:

- **Customers.csv:** Provided demographic and profile information for each customer.
- **Transactions.csv:** Contained transaction histories, including purchase amounts and dates.

Key features engineered for clustering:

- **TotalSpent:** Cumulative spending per customer.
- **TransactionCount:** Number of purchases made by each customer.
- **Recency:** Days since the last purchase.

Clustering Technique

The K-Means clustering algorithm was selected for its efficiency and interpretability. To determine the optimal number of clusters, the **Elbow Method** and **Silhouette Scores** were used.

3. Clustering Results

Number of Clusters Formed

Based on the analysis, the optimal number of clusters was determined to be **4**.

Clustering Metrics

- **Davies-Bouldin Index: 1.0084**, indicating well-separated and compact clusters.
- **Silhouette Score:** Confirmed the clustering quality.

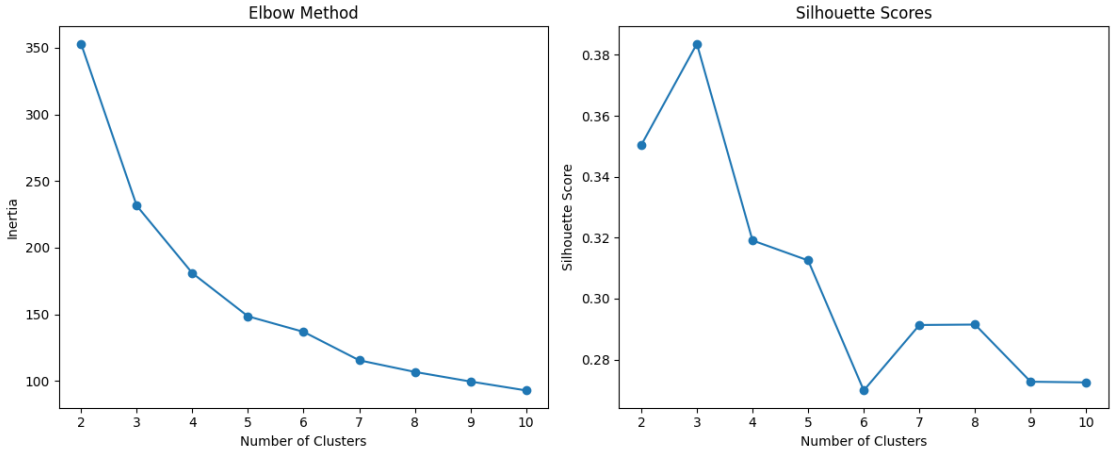
Cluster Characteristics

1. **Cluster 0:** High-value customers with frequent purchases (average spend: $5,967.29).
2. **Cluster 1:** Moderate-value customers with occasional transactions (average spend: $2,187.04).
3. **Cluster 2:** Infrequent low-value customers (average spend: $1,070.56).

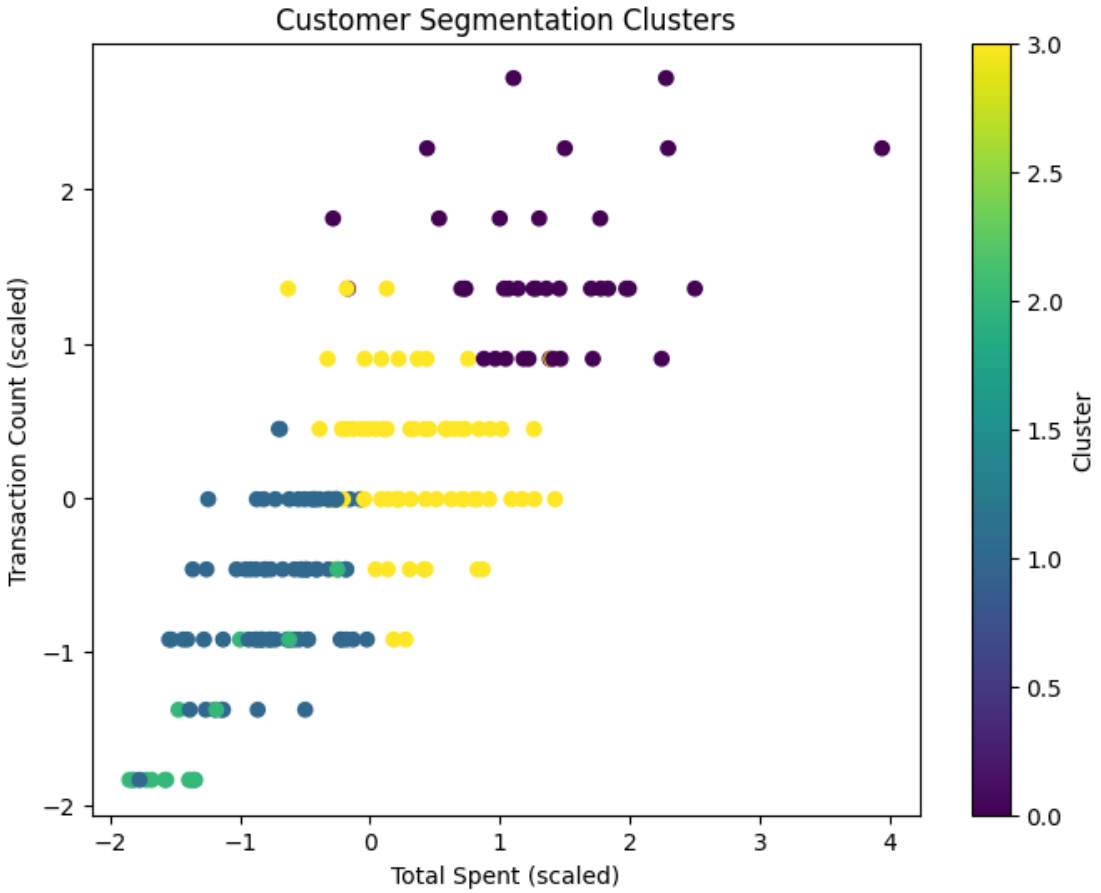
4. **Cluster 3:** Frequent moderate-value customers (average spend: $4,185.33).

4. Visual Representation

- **Elbow Plot:** Showed the point of inflection for the optimal number of clusters.
- **Scatter Plot:** Visualized customer clusters based on total spending and transaction frequency.



- **Silhouette Plot:** Validated cluster compactness and separation.



5. Conclusion

The customer segmentation analysis successfully identified four meaningful clusters. The clusters differ in spending patterns, transaction frequency, and engagement levels, offering actionable insights for personalized marketing strategies.
The DB Index value (1.0084) validates the effectiveness of the clustering approach.

This segmentation provides a foundation for businesses to enhance customer targeting, engagement, and overall marketing performance.

Let me know if you'd like further edits!