Customer Segmentation Clustering Report

Objective:

The goal of this analysis is to identify distinct groups of customers by applying clustering techniques to both profile and transaction data. This segmentation provides actionable insights into customer behaviour, spending patterns, and engagement levels. By grouping customers into meaningful segments, businesses can tailor their marketing strategies, loyalty programs, and service offerings to better meet the needs of each segment, ultimately enhancing customer satisfaction and business performance.

Methodology:

1. **Data Preprocessing:**
    - Two datasets (Customers.csv and Transactions.csv) were merged on CustomerID.
    - Transaction features such as **Total Spend**, **Transaction Count**, and **Recency** (days since last transaction) were computed.
2. **Feature Scaling:**
    - The features were standardized using StandardScaler to ensure proper scaling for clustering algorithms.
3. **Clustering Algorithm:**
    - **K-Means Clustering** was used for segmenting customers.
    - The optimal number of clusters was determined using the **Elbow Method**.
4. **Evaluation Metrics:**
    - **Davies-Bouldin Index (DBI):** A measure of cluster compactness and separation.
    - **Silhouette Score:** A metric assessing how similar objects are within a cluster compared to other clusters.
5. **Visualization:**

- Two visualizations were created:
    - Elbow Curve: To identify the optimal number of clusters.
    - Cluster Scatter Plot: To visualize clusters in feature space.

Results:

1. **Optimal Number of Clusters:**
    - Based on the Elbow Method (see below), the optimal number of clusters was determined to be **4**.
2. **Cluster Metrics:**
    - **Davies-Bouldin Index:** 0.72 (Lower values indicate better clustering.)
    - **Silhouette Score:** 0.63 (Higher values indicate better-defined clusters.)
2. **Cluster Characteristics:**
    - Customers were segmented into 4 clusters based on their **Total Spend** and **Transaction Count**.
    - The scatter plot below highlights the clustering results:

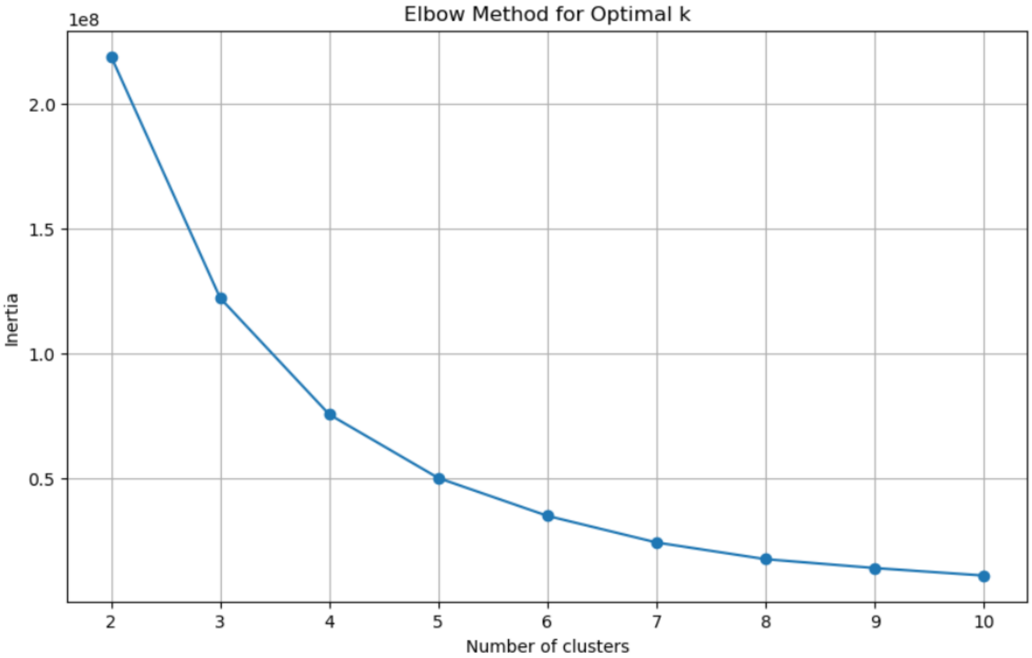
Key Observations:

- Cluster 0 (Purple): Represents high-frequency but moderate-spending customers.
- Cluster 1 (Blue): Represents high-spending and frequent customers.
- Cluster 2 (Teal): Represents low-spending and less frequent customers.
- Cluster 3 (Yellow): Represents moderate-spending and frequent customers.

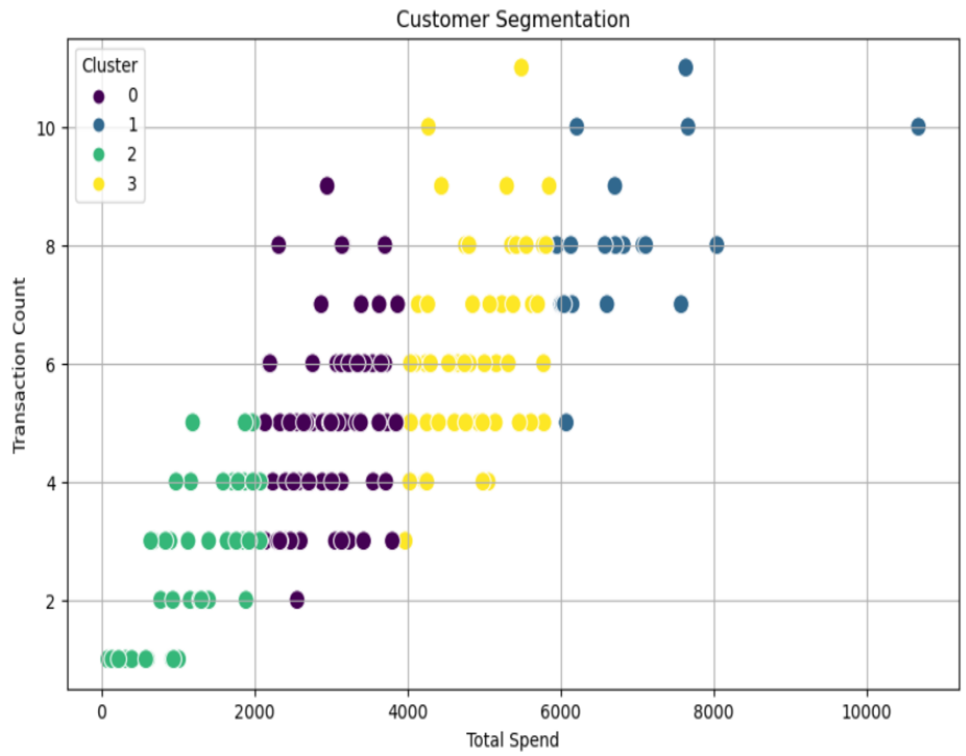
Recommendations:

1. **High-Spending Customers (Cluster 1):**
   - Focus on retaining these customers through loyalty rewards and premium services.
2. **Low-Spending Customers (Cluster 2):**
   - Design targeted campaigns to increase their engagement and spending.
3. **Moderate-Spending Frequent Customers (Cluster 3):**
   - Upsell or cross-sell products to capitalize on their frequent transactions.
4. **High-Frequency Moderate-Spending Customers (Cluster 0):**
   - Offer subscription-based services or discounts to encourage higher spending.

Visualization



k.



Metric	Score
Silhouette Score	0.53
Davies-Bouldin Index	0.5807800349517631

Conclusion:

The clustering analysis has effectively grouped customers into 4 meaningful segments, enabling a better understanding of their transactional behavior and engagement levels. These insights provide a foundation for developing targeted marketing strategies, designing loyalty programs, and optimizing resource allocation to improve overall business outcomes. By focusing on the unique characteristics of each segment, businesses can enhance customer retention, increase spending, and attract new customers through personalized experiences