Clustering Analysis Report

1. Number of Clusters Formed:

The K-Means algorithm was applied to the dataset, and based on the analysis, the optimal number of clusters selected was 4. This number was chosen by evaluating the Davies-Bouldin Index and Silhouette Score over a range of cluster values.

2. Davies-Bouldin Index (DB Index):

The calculated Davies-Bouldin Index value for the selected clustering configuration is 1.060424039992303. A lower DB Index indicates better clustering with well-separated and compact clusters.

3. Other Relevant Clustering Metrics:

- **Silhouette Score:** The silhouette score for the chosen number of clusters is 0.3135106549790538. This metric measures the cohesion and separation of clusters, with values closer to 1 indicating better-defined clusters.
- **Inertia Values:** Inertia values were analyzed for different cluster sizes, and the optimal number of clusters was selected based on the elbow method. The inertia for 4 clusters is 79.46684322018922, indicating the within-cluster sum of squares.

4. Visualizations and Insights:

- Cluster distributions were visualized using scatter plots to observe the separation.
- Feature standardization was performed before clustering to ensure consistency.

5. Conclusion:

The clustering process successfully grouped customers based on their transaction behavior, identifying key patterns in spending, transaction frequency, and average transaction value. These insights can be used for targeted marketing and customer segmentation strategies.