Clustering Results Report

1. Number of Clusters Formed:

 The clustering algorithm (KMeans) grouped the dataset into 4 distinct clusters.

 Cluster centers represent the average position of each cluster in the feature space.

2. Evaluation Metrics:

a. Davies-Bouldin Index (DBI):

 Value: 0.9666

 The DBI measures the compactness and separation of clusters.

 Lower values indicate better clustering, with minimal intra-cluster distance.

b. Silhouette Score:

 Value: 0.3103

 The silhouette score measures how similar data points are within their own cluster compared to others.

 The value of 0.3103 suggests moderate clustering quality. Ideal scores are closer to 1.

3. Other Observations:

a. Cluster Centers:

The centroids of the clusters (scaled values) are:

[

 [1.43, 1.46, 1.51, 0.15],

[-0.93, -0.52, -0.91, -1.07],

[-0.76, -0.92, -0.72, 1.03],

[0.30, 0.20, 0.22, -0.03]

]

Each centroid represents the mean characteristics of the respective cluster.

b. Cluster Visualization:

- A scatter plot was generated to visualize clusters based on:

  - X-Axis: Number of Transactions (NumTransactions)

  - Y-Axis: Total Spend (TotalSpend)

- Different clusters are color-coded, showing distinct groupings.

4. **Insights and Recommendations:**

Moderate Clustering Quality:

The DBI and Silhouette scores indicate reasonable but not perfect clustering quality.

Consider optimizing the number of clusters or preprocessing the data further.

Cluster Utility:

The clusters can be used for customer segmentation, enabling personalized marketing strategies or customer retention programs.