Clustering Report

1. Number of Clusters Formed

The optimal number of clusters was determined to be 4 based on the Davies-Bouldin Index.

2. Davies-Bouldin(DB) Index

DB Index is a measure of clustering quality, and a lower DB index indicates better-defined clusters. For k = 4 , the DB Index was calculated as 0.383335553763368 and this indicates reasonably well-defined clusters.

3. Inertia and Silhouette Score

a) Inertia: Measures the sum of squared distances of samples to their nearest cluster center. A lower inertia means the clusters are more compact.
Here, **Inertia (inertia) = 22.527815985312436**

b) Silhouette Score: Measures how similar each point is to its own cluster compared to other clusters. A higher silhouette score indicates better clustering.
Here, **Silhouette Score(silhouette_avg) = 0.6205322658864421**

4. Visualizations

a) DB Index vs. Number of Clusters

This plot shows the DB Index for different numbers of clusters. The optimal number of clusters corresponds to the minimum DB index

b) Pair-plot of Customer Features with Clusters

This pair plot visualises how customers are grouped based on their key features: Total Spending, Average Transaction Value, and Purchase Frequency. Each cluster is coloured differently.