Clustering Report

Number of Clusters Formed:

Using the clustering algorithm K-Means, 4 clusters are formed. The visualizations created showed how the customers were grouped based on their transaction data, where each cluster represents a distinct segment of customers with different spending behaviours.

DB Index Value:

The **Davies-Bouldin (DB) Index** is calculated, and the result is 0.5494920132693868, as reported in the notebook. The DB Index evaluates how well a clustering algorithm is working. It is desirable that the DB Index is lower; it means clusters are better defined with less intra-cluster variance and larger inter-cluster separation.

Other Relevant Clustering Metrics:

Silhouette Score: Although not mentioned as a requirement in the problem statement, this score can be used to measure cluster cohesion. High silhouette scores reflect that clusters are well separated.

Cluster Centroids: The centroids of every cluster describe the average values for the features of customers in that group. It, therefore, signifies regular customer behaviour within each cluster.

Cluster Size: The size of customers within each cluster clarifies if your customer population is distributed. For instance, if one cluster has many more customers, that segment would then represent the majority portion of your users.