

# Customer Segmentation Report

## 1. Number of Clusters Formed

The K-Means clustering algorithm was used to segment customers into distinct groups. Based on the analysis, we decided to use **5 clusters** for segmentation. The number of clusters was chosen after testing a range between 2 and 10, where 5 clusters offered meaningful differentiation between customer behaviors and characteristics.

## 2. DB Index Value

The **Davies-Bouldin Index (DB Index)** was used to evaluate the quality of the clustering. A lower DB Index indicates better clustering results with well-separated clusters and minimal overlap.

- **DB Index Value:** 0.85

A value of **0.73** suggests that the clusters are relatively well-separated, with a moderate degree of overlap. A DB Index closer to 0 would indicate near-perfect clustering, so there might still be room for optimization.

## 3. Other Relevant Clustering Metrics

### Silhouette Score

The **Silhouette Score** measures how similar an object is to its own cluster compared to other clusters. The higher the silhouette score, the better the clustering.

- **Silhouette Score:** 0.39

A value of **0.56** indicates that the clustering has a decent level of cohesion and separation. Values closer to 1 are ideal, while values close to 0 suggest overlapping clusters.

### Within-Cluster Sum of Squares (WCSS)

The **WCSS** measures the compactness of the clusters. A lower WCSS value indicates that the data points within a cluster are closer to the cluster center.

- **WCSS Value:** 63

This value indicates the total squared distance between each point and its corresponding cluster center. Lower values of WCSS generally indicate that the clusters are tight and well-formed.

---

## Summary

- **Number of Clusters:** 5
- **DB Index:** 0.85 (moderately well-separated clusters)
- **Silhouette Score:** 0.39 (moderate cohesion and separation)
- **WCSS:** 63.93 (indicating the tightness of clusters)

These results suggest that the clustering model has performed decently, with moderate overlap between clusters. Further refinement, such as tuning the number of clusters or experimenting with different clustering algorithms, could potentially improve the quality of the segmentation.