

Clustering Analysis Report

1. Number of Clusters Formed:

I tested out different values of clusters ranging from 2 to 10 to find the most optimal one. Each of these clusters was evaluated on its corresponding DB score to arrive at the most optimal number of clusters.

2. DB Index Value:

The Davies-Bouldin Index (DB Index) is a metric used to evaluate the quality of the clustering solution. A lower DB Index value indicates better clustering, as it reflects a smaller average similarity ratio between clusters, meaning the clusters are well-separated and compact.

3. Other Relevant Clustering Metrics:

- Silhouette Score: The silhouette score is used to assess the quality of the clusters. It measures how similar each point is to its own cluster compared to other clusters. The silhouette score ranges from -1 to +1, where a value closer to +1 indicates well-separated and dense clusters.

- Inertia: The inertia is a metric used in clustering to measure the sum of squared distances between each data point and its cluster centroid. A lower inertia value indicates that the points within the clusters are closer to the centroids, suggesting that the clustering is more compact.

Best number of clusters based on DB score: 7

Clustering Metrics Scores:

Clusters: 2,	DB Index: 1.972,	Silhouette Score: 0.163,	Inertia: 300.601
Clusters: 3,	DB Index: 1.791,	Silhouette Score: 0.225,	Inertia: 256.898
Clusters: 4,	DB Index: 1.420,	Silhouette Score: 0.303,	Inertia: 205.996
Clusters: 5,	DB Index: 1.696,	Silhouette Score: 0.251,	Inertia: 195.338
Clusters: 6,	DB Index: 1.608,	Silhouette Score: 0.256,	Inertia: 180.965
Clusters: 7,	DB Index: 1.419,	Silhouette Score: 0.259,	Inertia: 169.805
Clusters: 8,	DB Index: 1.521,	Silhouette Score: 0.244,	Inertia: 160.629
Clusters: 9,	DB Index: 1.476,	Silhouette Score: 0.247,	Inertia: 152.569

Metrics for 7 Number of Clusters:

Davies-Bouldin Index: 1.42
Silhouette Score: 0.26