Task 3 – Report

Clustering Results Report:

1. Number of Clusters Formed:

After applying the clustering algorithm, a total of 6 clusters were formed. These clusters were derived based on the data's intrinsic structure and optimized using appropriate evaluation metrics.

2. Davies-Bouldin Index (DB Index):

The Davies-Bouldin Index for the clustering results is 1.0573276115484942. This metric, which evaluates the average similarity ratio of each cluster to its most similar cluster, indicates that the clustering quality is moderate. A lower DB Index value signifies better clustering performance, indicating well-separated and cohesive clusters.

3. Other Relevant Clustering Metrics:

- **Silhouette Score:** The Silhouette Score is 0.2970203131991246. This metric ranges from -1 to 1, with values closer to 1 indicating that data points are well matched to their clusters and poorly matched to neighboring clusters.

- **Calinski-Harabasz Index:** The Calinski-Harabasz score is 80.82763696828127, reflecting the ratio of the sum of between-cluster dispersion to within-cluster dispersion. Higher values are better, indicating dense and well-separated clusters.

3. Summary of Findings:

The clustering results demonstrate that the dataset exhibits moderate groupings, with 6 clusters providing an optimal balance between separation and cohesion. While the Silhouette Score and Calinski-Harabasz Index highlight the quality of the clusters, the DB Index suggests areas for potential improvement.

This report serves as a foundational analysis of the clustering results and provides a basis for subsequent data-driven decisions.