Task 3: Customer Segmentation / Clustering:

1. Number of Clusters Formed

The clustering analysis resulted in the formation of **10 distinct clusters**. These clusters represent the natural grouping of the data points based on the applied clustering algorithm. Each cluster indicates a unique subset of the data with similar characteristics.

2. Davies-Bouldin Index

The optimal Davies-Bouldin Index (DBI) value for the clustering result is **1.247137385195408**. The Davies-Bouldin Index is a measure of cluster compactness and separation. A lower DBI value indicates better clustering performance, as it reflects compact clusters that are well-separated from each other.

3. Other Relevant Clustering Metrics

In addition to the DBI, other clustering performance metrics can provide a comprehensive understanding of the clustering results. These include:

- **Silhouette Score**: Assesses how similar an object is to its cluster compared to other clusters. Higher values (closer to 1) indicate better-defined clusters.
- **Intra-cluster Distance**: Measures the average distance between points within a cluster, indicating compactness.
- **Inter-cluster Distance**: Represents the distance between the centroids of different clusters, indicating separation.
- **Cluster Size Distribution**: Ensures no significant imbalance among the clusters, which could lead to skewed interpretations.

4. Interpretation of Results

The DBI value of **1.247** suggests the clustering algorithm has performed well, forming clusters that are reasonably compact and distinct. However, further analysis can be conducted to:

- Evaluate if the number of clusters is optimal using methods such as the Elbow Method or Gap Statistic.
- Verify the consistency of the clustering results across different algorithms.

5. Recommendations

- Visualize the clusters using dimensionality reduction techniques like PCA or t-SNE to gain insights into the structure of the data.
- Validate the clusters with domain-specific knowledge to ensure meaningful patterns.
- Test alternative algorithms or hyper parameter tuning to explore potential improvements in clustering performance.