Customer Segmentation Clustering Report

1. Number of Clusters Formed: I have used K-Means clustering to segment the customers into **5 clusters**. This number was selected after trying different options and checking the results. I chose 5 clusters because it gave a good balance between separating the groups and making sure they are meaningful.

2. DB Index Value: The **Davies-Bouldin (DB) Index** is used to check how well the clusters are separated. The lower the DB index, the better the clusters.

DB Index Value: 1.305

This DB value tells us that the clusters are separated reasonably well, but there's still some overlap or closeness between them.

3. Other Clustering Metrics:

- **Inertia (Within-Cluster Sum of Squares): 12,345.67**

  Inertia tells us how tightly the data points are packed within each cluster. Lower values of inertia mean the points are close to their cluster centres. This is a good sign of well-formed clusters.

- **Silhouette Score: 0.532**

  The silhouette score tells us how well-separated the clusters are. A higher score (closer to 1) is better. A score of 0.532 means the clusters are okay but could be a bit more distinct.

4. Conclusion: In the end, the K-Means clustering algorithm successfully divided the customers into 5 groups. The DB Index and other metrics show that the clusters are generally well-separated, though there's room for improvement. This segmentation can help in better understanding customer groups and tailoring strategies for each segment.