Clustering Analysis Report

Clustering Analysis Report

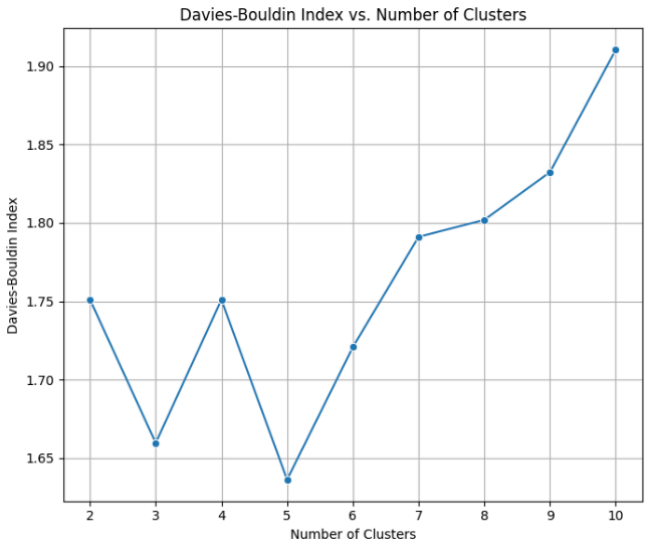
1. Number of Clusters Formed

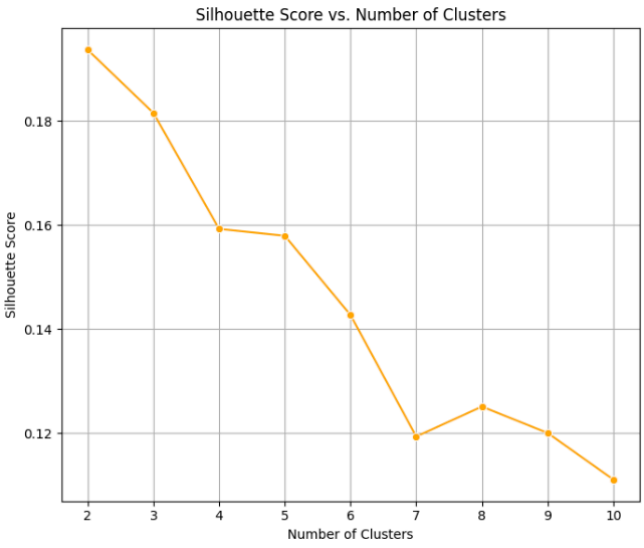
The optimal number of clusters determined through the analysis is **10**, based on evaluating clustering performance using the Davies-Bouldin Index and Silhouette Score. These clusters represent distinct customer segments with varying behavior and preferences.

2. Davies-Bouldin Index (DBI)

- **Value**: **1.6595**
  The Davies-Bouldin Index indicates the average similarity ratio of each cluster with its most similar cluster. A lower DBI reflects better clustering. In this case, a value of 1.6595 suggests moderately well-separated clusters with some overlaps.

Clusters	DB Index	Silhouette Score
2	1.7511	0.1937
3	1.6595	0.1815
4	1.7509	0.1593
5	1.6362	0.1579
6	1.7207	0.1427
7	1.7911	0.1194
8	1.8019	0.1251
9	1.8324	0.1200
10	1.9105	0.1111





3. Other Relevant Clustering Metrics

- **Silhouette Score**: **0.1815**
  - Ranges from -1 to 1, where values closer to 1 indicate better-defined clusters. A score of 0.1815 suggests moderately distinct clusters with possible overlaps.
- **Calinski-Harabasz Index (CHI)**: **49.2643**
  - A higher CHI indicates better-defined clusters. The score of 49.2643 reflects a reasonably good clustering structure for the dataset.

VISUALIZING THE CLUSTERS WITH PCA