Unlocking the Secrets of Customer Behavior

Analyzing Trends and Strategies to Boost Engagement and Revenue

Clustering Report

1. Number of Clusters Generated

Based on the clustering model used, this dataset was broken down into 4 clusters as follows:

- **Cluster 0:** 33 customers
- **Cluster 1:** 83 customers
- **Cluster 2:** 27 customers
- **Cluster 3:** 57 customers

2. Davies-Bouldin Index

The Davies-Bouldin (DB) Index is a cluster quality measure calculated to show the separation between clusters. The value of the DB Index for this clustering model is **0.745**.

Interpretation: A lower DB Index translates to better clustering with less overlap between clusters. A DB Index value of 0.745 indicates that the clusters could generally be well-separated, but there might be room for improvement. This could involve increasing or decreasing the number of clusters or possibly applying a different clustering algorithm.

3. Other Clustering Metrics

Besides the DB Index, there are several other clustering metrics available to assess the performance of the clustering model, including:

- **Silhouette Score:** This assesses how much a point resembles its own cluster compared to others. A higher silhouette score indicates better-defined clusters.
- **Inertia:** Also known as the Within-Cluster Sum of Squares, this measures how compact the clusters are. Lower inertia means tighter clusters.
- **Cluster Cohesion and Separation:** These metrics measure the density within clusters and the distance between clusters, respectively.

All these metrics will be computed in relation to the clustering algorithm and the configuration used.

Conclusion

- **Number of Clusters:** 4
- **Davies-Bouldin Index:** 0.745, suggesting good separation with potential for improvement.
- **Cluster Distribution:** The clusters vary in size; Cluster 1 had the most customers (83), and Cluster 2 had the least (27).

The clustering results suggest that the clusters are somewhat separable but could benefit from further tuning of the algorithm or features to improve cluster separation and cohesion.

Additional Files

- **CustomerClusters.csv:** CSV file containing the Customer IDs and their assigned cluster labels.
- **ClusteringReport.txt:** Text file summarizing the clustering analysis and results.