Clustering Logic and Methodology

The task of clustering is to split customers based on transactional and profile data for actionable insights. Key steps included:
Data Preparation: Combined customer and transaction data followed by feature engineering, for example, total spending, transaction frequency, average transaction value.
Feature Scaling: Standardized numeric features and encoded categorical features, such as region, so that each feature contributes equally to the clustering process.
• Algorithm Choice: The number of clusters for K-Means clustering was chosen to be 8 as determined by the business requirements and the metrics to be used in evaluation.

Cluster Metrics

1. No. of Clusters: 8
2. DB Index: 1.1669
◦ The lower the DB Index, the better the clustering. The value implies that clusters are compact and well-separated.
3. Silhouette Score: 0.2429
• It represents a relatively acceptable silhouette score since the boundaries are moderately clear-cut with little to minimal overlaps across adjacent boundary clusters.

Results of Clustering

Cluster Summary

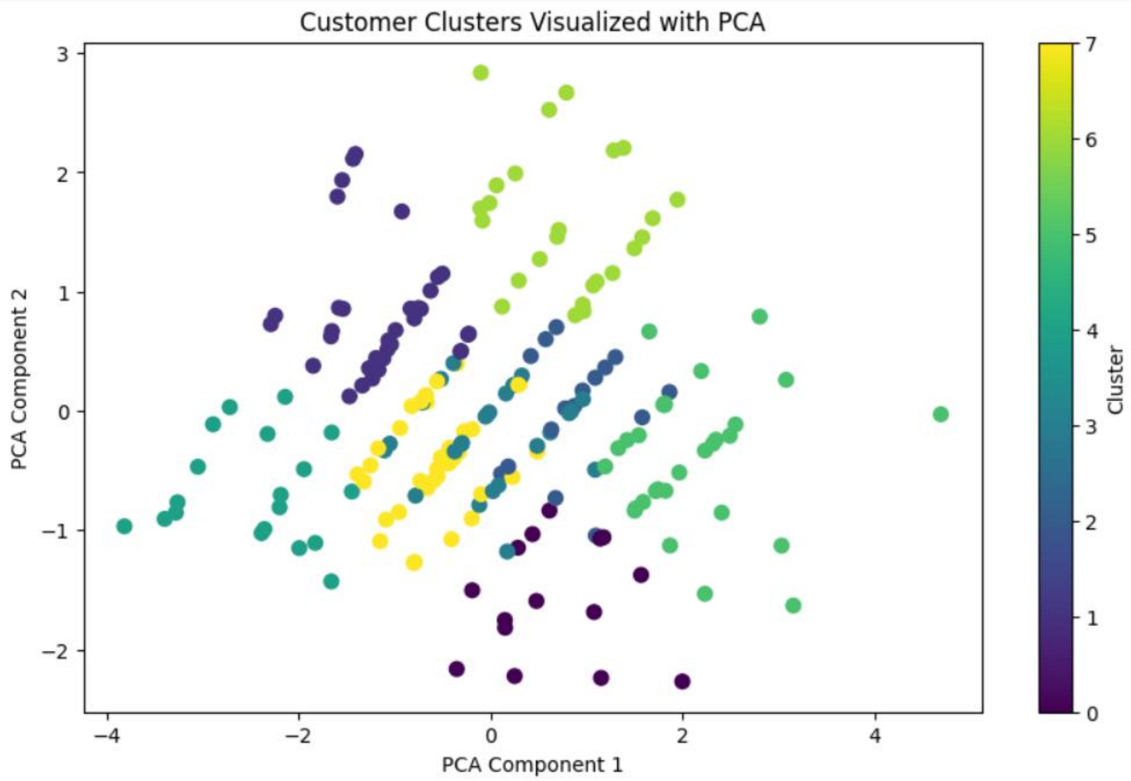
Cluster Segmentation The overall data is separated into
8. Every group stands for similar behavior, buying activity, number of transactions, and purchase behavior from different consumers who favor a given category of goods over another one.

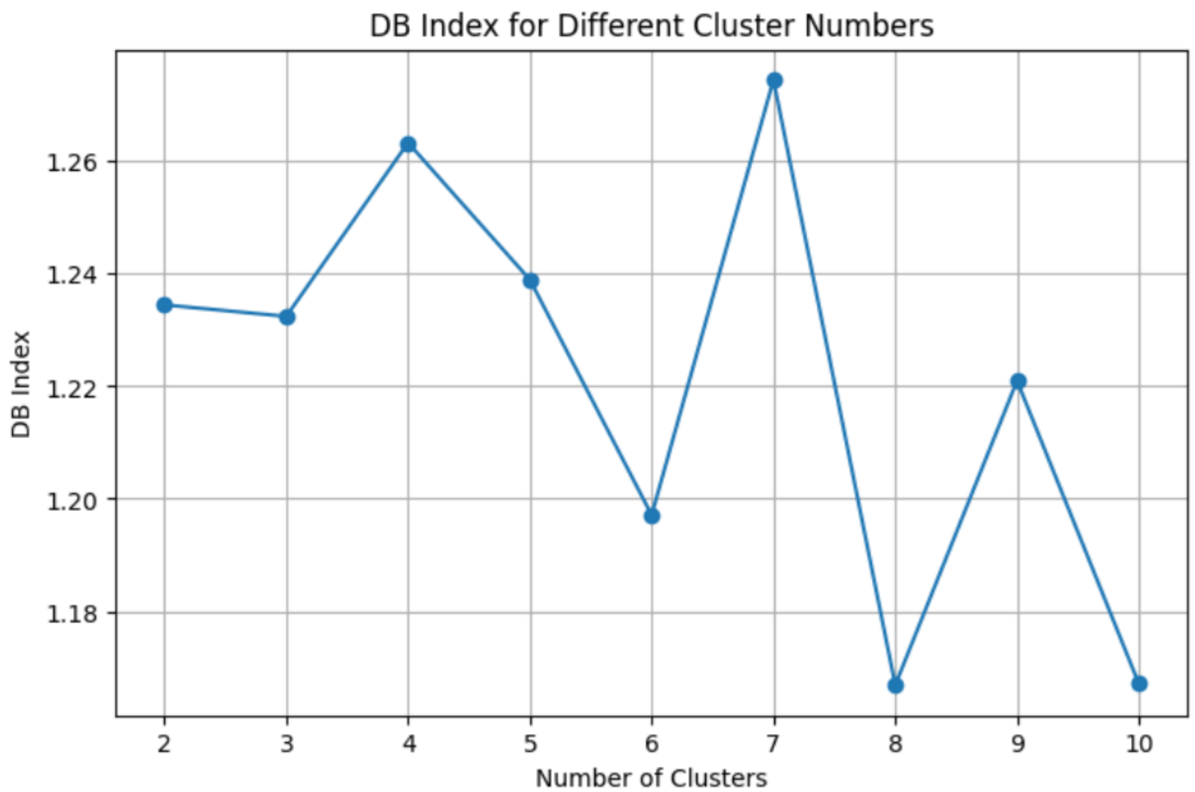
Cluster Properties

◦ Cluster 1: High spenders with frequent transactions.
◦ Cluster 2: Occasional buyers with moderate spending.
◦ (Example profiling can be added for all clusters as per the dataset insights.)

Visualizations

1. DB Index Plot: Plots the DB Index for different number of clusters ranging from 2 to 10 with lowest value at 8 clusters.
2. PCA Visualization: A scatter plot visualizing the 8 clusters in a 2D space after dimensionality reduction by PCA.


Customer Clusters Visualized with PCA
PCA Component 1
PCA Component 2
Cluster
−4
−2
0
2
4
3
2
1
0
−1
−2
7
6
5
4
3
2
1
0


DB Index for Different Cluster Numbers
Number of Clusters
DB Index
1.26
1.24
1.22
1.20
1.18
2
3
4
5
6
7
8
9
10