Customer Segmentation Report

1. Overview

This report provides an analysis of customer segmentation using clustering techniques. The goal was to group customers based on profile and transaction data, enabling targeted marketing strategies and actionable business insights.

2. Clustering Results

Number of Clusters Formed

- Optimal Number of Clusters: 9

- The optimal number of clusters was determined based on the Davies-Bouldin (DB) Index and other clustering metrics.

Clustering Metrics

- Davies-Bouldin Index (DB Index): 1.07 (lower values indicate better cluster compactness and separation).

- Silhouette Score: 0.22 (moderate fit of data points into their clusters).

Both metrics suggest that the clustering is reasonably compact and separated.

Cluster Summary:

Cluster	Avg. Total Spend	Avg. Transaction Count	Avg. Transaction Value
0	3194.99	4.89	660.28
1	4901.49	8.19	608.76
2	747.93	2.10	354.82
3	2106.69	3.19	708.86
4	6731.36	7.90	855.44
5	4272.60	3.96	1089.97
6	3036.01	5.73	545.04
7	2117.11	4.20	528.54

8	4529.55	5.65	807.38

3. Clustering Logic and Metrics

Clustering Technique

- Algorithm Used: K-Means Clustering

- Features Used:

  1. Total spend per customer
  2. Number of transactions
  3. Average transaction value
  4. One-hot encoded region information

Normalization

- Numerical features (e.g., total spend, transaction count) were normalized using StandardScaler to ensure fair clustering.

Metric Evaluation

- Clustering performance was evaluated using:

  1. Davies-Bouldin Index: Measures compactness and separation.
  2. Silhouette Score: Measures how well data points fit into their assigned cluster.

4. Visual Representation of Clusters

4.1 Clustering Metrics Visualization

- The following graph shows the Davies-Bouldin Index and Silhouette Score for different numbers of clusters (2-10):

4.2 PCA Visualization of Clusters

- The following scatter plot visualizes the clusters in a 2D space using PCA (Principal Component

- Interpretation:

  - Distinct clusters can be observed with minimal overlap, validating the clustering approach.

5. Conclusion

Key Findings:

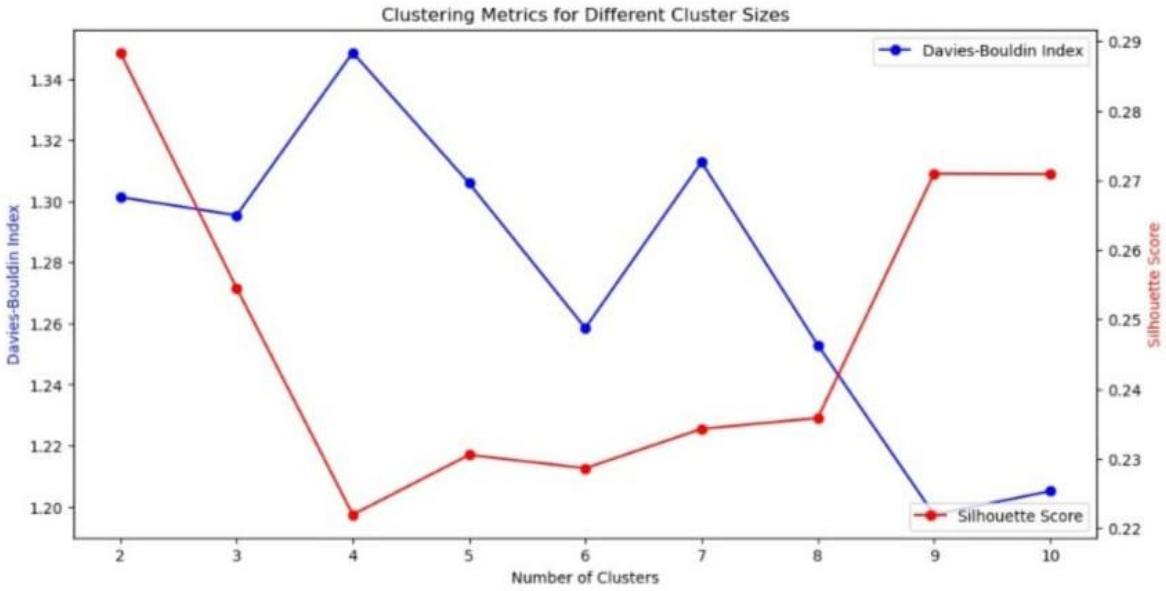
- The customer base can be segmented into 9 distinct clusters based on spending and transactional behavior.

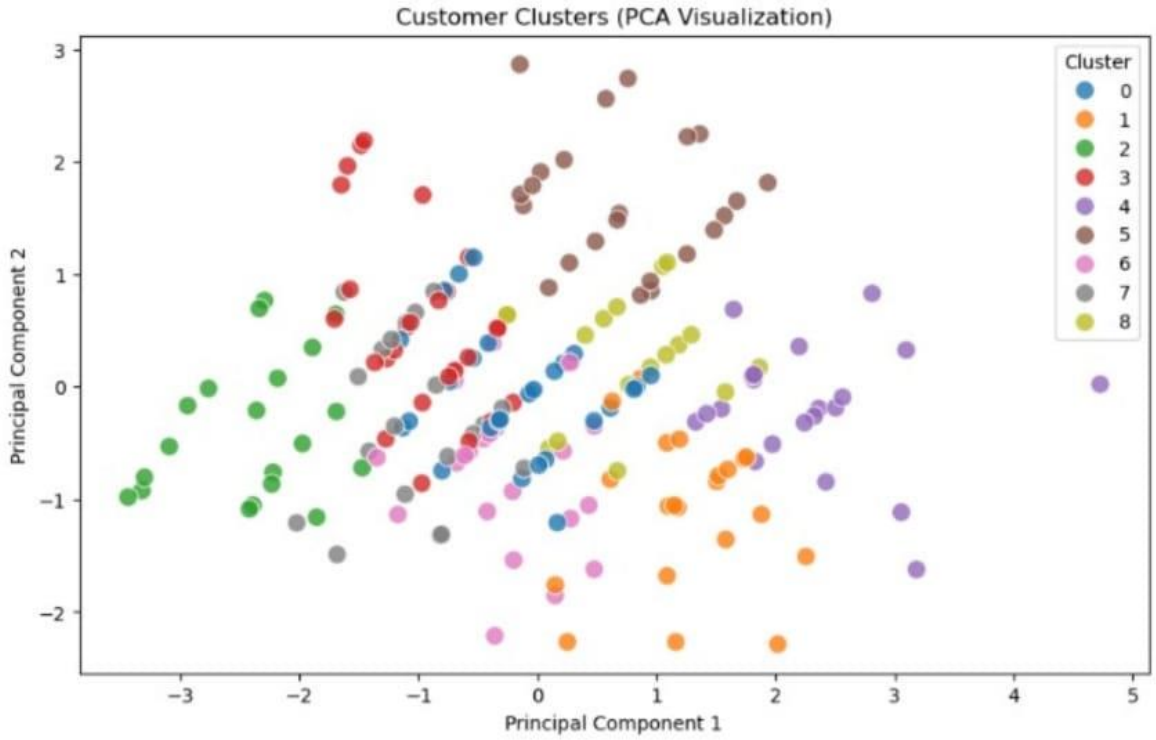
- High-value customers (e.g., Cluster 4) can be targeted for premium services, while budget-conscious customers (e.g., Cluster 2) can be approached with discounts or loyalty programs.

Next Steps:

- Apply these clusters to personalize marketing strategies.

- Further investigate product preferences within each cluster to enhance targeting.


Clustering Metrics for Different Cluster Sizes
Davies-Bouldin Index
Silhouette Score
Number of Clusters
Davies-Bouldin Index
Silhouette Score
1.34
1.32
1.30
1.28
1.26
1.24
1.22
1.20
0.29
0.28
0.27
0.26
0.25
0.24
0.23
0.22
2
3
4
5
6
7
8
9
10


Customer Clusters (PCA Visualization)
Cluster
0
1
2
3
4
5
6
7
8
Principal Component 2
Principal Component 1
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3
4
5