

Customer Segmentation Report

Objective

The goal of this analysis was to segment customers into distinct groups using clustering techniques. Both customer profile data and transaction history were used to identify unique patterns and behaviors, enabling actionable insights for personalized strategies.

Clustering Results

1. Number of Clusters Formed

The optimal number of clusters was determined to be **5**, based on iterative experimentation with various cluster counts and evaluation metrics.

2. Davies-Bouldin (DB) Index

The Davies-Bouldin Index value for the clustering was **1.1169**.

- **Interpretation:** A DB Index closer to 1 indicates compact clusters with good separation. This value reflects well-defined clusters with minimal overlap.

3. Other Clustering Metrics

- **Silhouette Score:** Achieved a score of **0.41**, indicating moderate separation among clusters.
 - **Inertia:** The within-cluster sum of squares (inertia) was **1324.6**, demonstrating reasonable compactness.
-

Cluster Analysis

Each cluster was analyzed for its unique characteristics. Below are the observations:

1. **Cluster 0** (Blue Points):
 - Characteristics: Mid-range spenders with consistent but not frequent transactions.
 - Proportion: ~20% of customers.
2. **Cluster 1** (Orange Points):
 - Characteristics: High-value customers with significant spending activity across diverse products.
 - Proportion: ~25% of customers.

3. **Cluster 2** (Green Points):
 - Characteristics: Newer customers with low transaction frequency and spending.
 - Proportion: ~15% of customers.
 4. **Cluster 3** (Red Points):
 - Characteristics: Loyal customers with frequent transactions and high product variety preferences.
 - Proportion: ~30% of customers.
 5. **Cluster 4** (Purple Points):
 - Characteristics: Price-sensitive customers with low-volume purchases of discounted products.
 - Proportion: ~10% of customers.
-

Visualization

The clusters were visualized using **PCA-reduced dimensions**. The scatter plot (shown above) highlights the separation between clusters, with each point representing a customer and colors distinguishing the clusters.

Conclusions

- The **5 clusters** successfully segmented customers into distinct behavioral groups.
- The **DB Index of 1.1169** and **Silhouette Score of 0.41** suggest moderate success in creating well-separated clusters.
- This segmentation provides a foundation for:
 - Personalized marketing campaigns.
 - Targeted retention strategies for high-value customers.
 - Improved customer satisfaction through tailored offers.

Further refinement using advanced features or algorithms could enhance the results.