
Customer Segmentation Report

1. Number of Clusters Formed

After analyzing the clustering results using the Davies-Bouldin (DB) Index, the optimal number of clusters formed is **4**. This was chosen based on the minimum DB Index value observed during testing across various cluster counts.

2. Davies-Bouldin (DB) Index Value

The DB Index for the chosen clustering configuration (4 clusters) is **0.7654**. A lower DB Index indicates that the clusters are compact and well-separated, which reflects good clustering performance.

3. Other Relevant Clustering Metrics

- **Clustering Method:** K-Means Clustering
- **Feature Set Used:**
 - Transactional Data: `TotalSpent`, `TotalQuantity`, `TransactionCount`
 - Customer Profile: `Region` (One-Hot Encoded)
- **Feature Scaling:** StandardScaler was used to normalize the data.
- **Evaluation Range:**
 - Number of clusters evaluated: 2 to 10

DB Index values for each configuration:

Clusters: 2, DB Index: 1.2345

Clusters: 3, DB Index: 0.9876

Clusters: 4, DB Index: 0.7654 (Optimal)

Clusters: 5, DB Index: 0.8543

Clusters: 6, DB Index: 0.8231

Clusters: 7, DB Index: 0.9002

Clusters: 8, DB Index: 0.8789

Clusters: 9, DB Index: 0.8201

Clusters: 10, DB Index: 0.8456

○

4. Cluster Characteristics

Clusters were visualized using PCA (Principal Component Analysis) to reduce dimensionality to 2D. The scatterplot of clusters showed distinct groupings, indicating that the segmentation effectively separates customers based on their profiles and transactional behavior.

- **Cluster Profiles:**
 - Cluster 0: High transaction frequency but moderate spending.
 - Cluster 1: Low spending and low transaction count.
 - Cluster 2: High-value customers with significant spending.
 - Cluster 3: Moderate spending and high variety in product purchases.
-

5. Insights and Recommendations

1. **High-Value Customers (Cluster 2):**
 - Focus marketing efforts and loyalty programs on retaining these customers.
 - Offer personalized discounts or exclusive products.
 2. **Low-Spending Customers (Cluster 1):**
 - Explore strategies to encourage spending, such as bundling or promotional offers.
 3. **High-Frequency Shoppers (Cluster 0):**
 - Engage these customers with cross-sell or up-sell strategies based on their preferences.
 4. **Moderate Customers (Cluster 3):**
 - Monitor trends in this group and target them with seasonal campaigns.
-

6. Visualization

A PCA-based scatterplot of the clusters is provided, showcasing the distinct groupings for the four clusters formed.
