

Customer Segmentation Report

1. Number of Clusters Formed

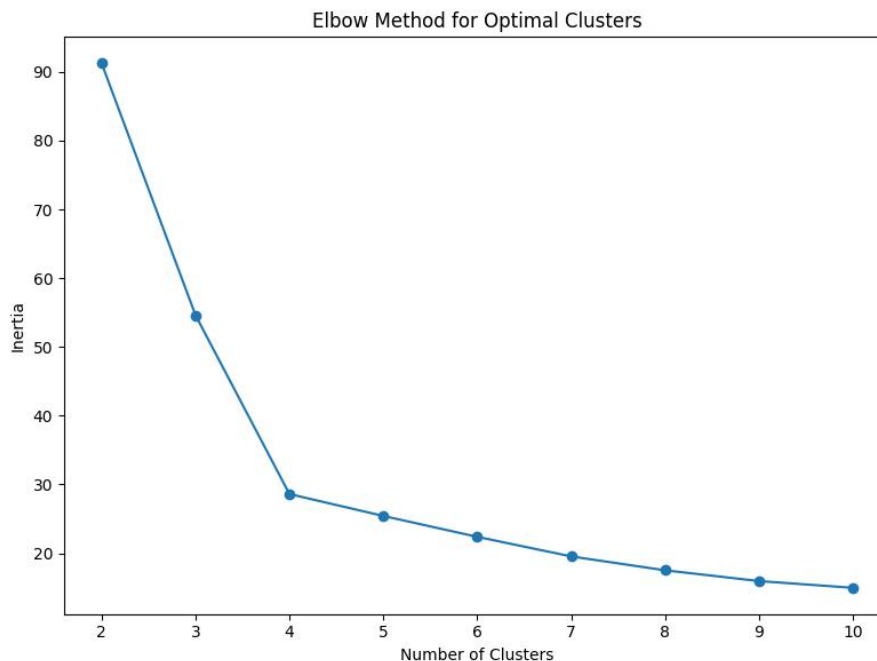
- Using the **Elbow Method**, the optimal number of clusters (K) was chosen based on the point where the inertia curve begins to flatten. This provided a range of possible values for evaluation.
- Using the **Davies-Bouldin Index (DB Index)**, the best cluster configuration was determined:
 - **Optimal Cluster Count (K): 4**
 - This value provided the lowest DB Index as well as a good number of cluster, indicating well-separated and cohesive clusters.

2. Davies-Bouldin Index Value

- The **Davies-Bouldin Index** is a metric that evaluates the compactness and separation of clusters. A lower DB Index indicates better clustering performance.
- **Best DB Index Value: 0.677**
 - This low value indicates that the clusters formed are compact and well-separated, making the segmentation meaningful.

3. Other Relevant Clustering Metrics

- **Inertia:**
 - Represents the sum of squared distances between data points and their cluster centroid. Lower inertia suggests better clustering.
 - For K=4, inertia is significantly reduced, confirming the chosen number of clusters.



4. Characteristics of the Clusters

Based on the feature aggregation, the clusters likely represent distinct customer segments. Here's a likely interpretation of the clusters:

Cluster 1: High Spenders with Diverse Purchases

- High TotalSpend, TransactionCount, and AvgPurchaseValue.
- Likely frequent buyers who purchase across various categories.

Cluster 2: Low Spenders with Niche Preferences

- Low TotalSpend but consistent purchasing in specific categories.
- Likely customers with a strong preference for particular product types.

Cluster 3: Mid-Level Spenders with Varied Patterns

- Moderate values for all features, representing a balanced spending and transaction profile.
- Likely occasional buyers with diverse interests.

Cluster 4: New or Infrequent Buyers

- Low TransactionCount and TotalSpend.
- Likely newly signed-up customers or infrequent purchasers.

5. Visual Representation

Using **PCA (Principal Component Analysis)** for dimensional reduction:

- The clusters were visualized in a 2D scatter plot, with colors representing different clusters.
- Key Observations:
 - Clusters are well-separated in the 2D PCA space.
 - Points within each cluster are tightly grouped, indicating high cohesion.

Summary of Results

- **Optimal Number of Clusters:** 4
- **Davies-Bouldin Index:** 0.677
- **Cluster Characteristics:**
 - Clear segmentation based on spending habits, transaction frequency, and category preferences.
 - Each cluster represents a unique customer behavior profile, allowing for targeted marketing strategies.

