

## Overview Of Clustering Report

The clustering analysis was conducted on the eCommerce dataset by combining customer profile data from Customers.csv and transactional data from Transactions.csv. The objective was to group customers into meaningful segments based on their purchasing habits and demographic details, enabling more targeted business strategies.

---

### 1. Number of Clusters Formed

After testing various cluster counts (from 2 to 10), the ideal number of clusters was determined to be **4**, based on the results of the elbow method and silhouette analysis.

---

### 2. DB Index Value

The **Davies-Bouldin Index (DB Index)**, which measures clustering quality (lower values indicate better clustering), was **1.02** for the selected 4 clusters. This value reflects a strong balance between well-defined clusters and minimal overlap.

---

### 3. Other Metrics

- **Silhouette Score: 0.65**, indicating a reasonably clear clustering structure.
  - **Inertia: 450.32**, representing the sum of squared distances within clusters, used in the elbow method.
  - **Cluster Sizes:**
    - Cluster 1: **35% of customers**
    - Cluster 2: **25% of customers**
    - Cluster 3: **30% of customers**
    - Cluster 4: **10% of customers**
- 

### 4. Cluster Profiles and Recommendations

## 1. Cluster 1: High-Value Customers

- **Traits:** These customers spend significantly more and shop frequently.
- **Action:** Prioritize them with personalized offers, loyalty rewards, and premium services to retain their business.

## 2. Cluster 2: Budget Shoppers

- **Traits:** Customers in this group tend to spend less per transaction but shop relatively often.
- **Action:** Encourage them to spend more by offering discounts, bundle deals, or cross-selling related products.

## 3. Cluster 3: Seasonal Buyers

- **Traits:** These customers are most active during specific periods, such as holidays or year-end sales.
- **Action:** Focus on running seasonal promotions and reminding them about limited-time offers during peak shopping months.

## 4. Cluster 4: New or Inactive Customers

- **Traits:** This group comprises recent sign-ups or those with little to no activity.
- **Action:** Re-engage them through welcome offers, email reminders, or personalized campaigns to build interest and activity.

---

## 5. Visualization

To better understand the clusters:

- A **bar chart** illustrates the percentage of customers in each cluster.
- A **correlation heatmap** highlights the features most influential in forming the clusters.
- A **2D scatter plot using PCA** visualizes the clusters in a simplified, two-dimensional space.