

# Customer Segmentation Clustering Analysis Report

## Executive Summary

This report presents the results of customer segmentation analysis performed on the eCommerce dataset using K-means clustering. The study identified five distinct customer segments based on transaction patterns and customer behavior.

## Clustering Results

### Key Metrics

- Number of Clusters Formed: 5 (optimal based on Silhouette Score)
- Davies-Bouldin Index: 1.426
- Silhouette Score: 0.257

### Cluster Quality Assessment

The Davies-Bouldin Index of **1.426** and Silhouette Score of **0.257** indicate moderate cluster separation. While the clusters are distinct, some segments overlap, which is common in customer behavior analysis.

## Cluster Characteristics

### Cluster Size Distribution

Total Customers: 200

- Cluster 0: 45 customers (22.5%)
- Cluster 1: 54 customers (27.0%)
- Cluster 2: 43 customers (21.5%)
- Cluster 3: 20 customers (10.0%)
- Cluster 4: 38 customers (19.0%)

## **Cluster Profiles**

### **Cluster 0: Moderate-Value Regular Customers**

- Transaction Count: 4.89 average transactions
- Total Spend: \$3,617.58
- Average Transaction Value: \$747.41
- Recency: 102.31 days
- Account Age: 583.04 days

### **Cluster 1: High-Value Active Customers**

- Transaction Count: 5.46 average transactions
- Total Spend: \$4,011.69 (Highest)
- Average Transaction Value: \$731.39
- Recency: 92.48 days
- Account Age: 635.02 days (Longest)

### **Cluster 2: Regular Customers**

- Transaction Count: 5.60 average transactions (Highest)
- Total Spend: \$3,508.42
- Average Transaction Value: \$650.85
- Recency: 103.58 days
- Account Age: 467.84 days

### **Cluster 3: Low-Value Inactive Customers**

- Transaction Count: 1.75 average transactions (Lowest)
- Total Spend: \$631.20 (Lowest)
- Average Transaction Value: \$416.20 (Lowest)
- Recency: 191.85 days (Highest)
- Account Age: 460.80 days

### **Cluster 4: High-Engagement Recent Customers**

- Transaction Count: 5.50 average transactions
- Total Spend: \$3,870.73
- Average Transaction Value: \$720.77
- Recency: 86.66 days (Most Recent)
- Account Age: 580.74 days

## Model Selection Process

### Clustering Metrics Across Different Numbers of Clusters

Clusters	DB Index	Silhouette Score
2	2.087	0.161
3	2.005	0.142
4	1.871	0.162
5	1.426	0.257
6	1.381	0.233
7	1.382	0.223
8	1.410	0.230
9	1.443	0.216
10	1.401	0.223

Five clusters were selected as optimal based on:

- Highest Silhouette Score (0.257)
- Significant improvement in Davies-Bouldin Index from 4 to 5 clusters
- Meaningful and interpretable customer segments

## Business Implications

### Key Insights

1. Clear distinction between high-value (Cluster 1) and low-value (Cluster 3) customers
2. Significant variation in transaction frequency (1.75 to 5.60 transactions)
3. Large spread in average transaction values (\$416.20 to \$747.41)
4. Recency patterns indicate varying engagement levels (86.66 to 191.85 days)

## **Recommended Actions**

1. Prioritize retention efforts for Cluster 1 (highest total spend)
2. Develop reactivation campaigns for Cluster 3 (inactive customers)
3. Leverage Cluster 4's recent engagement for upselling opportunities
4. Create value enhancement programs for Cluster 0 and 2 to increase transaction values

## **Conclusions**

The five-cluster solution provides a clear segmentation of the customer base, with distinct patterns in spending, engagement, and account history. The moderate Silhouette Score suggests some natural overlap between segments, which is expected in customer behavior analysis. These segments provide a solid foundation for targeted marketing strategies and customer engagement initiatives.