Customer Clustering Analysis Report

Clustering Results Summary

Key Metrics

- **Number of Clusters Formed**: 7
- **Davies-Bouldin Index**: 1.3669
- **Number of Features Used**: 9
- **Total Customers Clustered**: 200

Cluster Characteristics

Based on the clustering metrics, we can identify the following customer segments:

Cluster 0: Premium High-Value Customers

- Highest average transaction value: $6094.58
- Highest transaction frequency: 7.52 transactions
- Highest quantity per purchase: 21.49 items
- Customer base: 39 customers
- Characteristics: Most valuable customer segment with highest engagement across all metrics

Cluster 1: Regular Value Customers

- Average transaction value: $3251.23
- Transaction frequency: 4.40 transactions
- Average quantity: 10.80 items
- Customer base: 35 customers
- Characteristics: Consistent mid-tier customers with moderate purchase behavior

Cluster 2: Moderate Buyers

- Average transaction value: $2387.73
- Transaction frequency: 3.86 transactions
- Average quantity: 9.60 items
- Customer base: 45 customers
- Characteristics: Largest cluster with moderate purchase patterns

Cluster 3: Low-Engagement Customers

- Average transaction value: $568.81
- Transaction frequency: 2.00 transactions
- Average quantity: 3.13 items
- Customer base: 15 customers

- Characteristics: Minimal engagement and lowest value segment

Cluster 4: Mid-High Value Customers

- Average transaction value: $4572.91
- Transaction frequency: 4.71 transactions
- Average quantity: 15.21 items
- Customer base: 24 customers
- Characteristics: Second-highest value segment with good engagement

Cluster 5: Regular Frequency Customers

- Average transaction value: $2987.36
- Transaction frequency: 6.18 transactions
- Average quantity: 12.85 items
- Customer base: 34 customers
- Characteristics: High transaction frequency with moderate value

Cluster 6: Low-Value Customers

- Average transaction value: $1672.66
- Transaction frequency: 2.00 transactions
- Average quantity: 6.27 items
- Customer base: 11 customers
- Characteristics: Low engagement but higher value than Cluster 3

Clustering Performance Analysis

1. **Davies-Bouldin Score Evolution**:
    - Scores ranged from 1.3669 to 1.4934 across different cluster numbers
    - Optimal score achieved at 7 clusters (1.3669)
    - Shows stable clustering performance
2. **Elbow Curve Analysis**:
    - Clear elbow point visible around 4-5 clusters
    - Diminishing returns in inertia reduction after 7 clusters
    - Supports the selection of 7 clusters as optimal
3. **Cluster Distribution**:
    - Good balance in cluster sizes (11-45 customers per cluster)
    - No evidence of cluster collapse or extreme imbalance
    - Natural segmentation of customer base

Methodology

- Algorithm: K-means Clustering
- Optimal clusters determined through Davies-Bouldin Index minimization
- Features scaled using StandardScaler
- Evaluation metrics include transaction frequency, value, and quantity

Visualization Insights

1. **Cluster Sizes**: Relatively balanced distribution with cluster 2 being the largest
2. **Average Total Value**: Clear differentiation between high-value and low-value segments
3. **Transaction Frequency**: Distinct patterns in purchase frequency across segments
4. **Customer Distribution**: Natural segmentation with meaningful size differences

This clustering solution provides a robust segmentation of the customer base, enabling targeted marketing strategies and personalized customer engagement approaches.