Customer Segmentation Analysis Report

Executive Summary

This report presents the results of a customer segmentation analysis performed using clustering techniques on customer profile and transaction data. The analysis identified five distinct customer segments, achieving an optimal Davies-Bouldin Index of 0.8819, indicating well-separated and cohesive clusters.

Methodology

The analysis employed clustering algorithms on combined customer profile and transaction data, focusing on key metrics:

- Total Spend
- Average Spend
- Transaction patterns
- Customer profile characteristics

Clustering Results

Optimal Number of Clusters

After evaluating cluster configurations from 2 to 10 clusters, the analysis determined that **5 clusters** provided the optimal segmentation, based on the following metrics:

- Davies-Bouldin Index: 0.8819 (lowest among all configurations)
- Visual cluster separation (as evident in the scatter plots)
- Business interpretability of the resulting segments

Davies-Bouldin Index Analysis

The DB Index values across different cluster numbers showed:

- Best performance (lowest DB Index): 5 clusters (0.8819)
- Worst performance: 3 clusters (1.1139)
- Notable patterns:
  - Performance degraded with 6 clusters (1.0436)
  - Relatively stable performance between 7-10 clusters (~0.90-0.99)

Cluster Distribution
Based on the Customer_Clusters.csv data, the distribution of customers across the five clusters shows:

- Cluster 0: 43 customers (21.5%)
- Cluster 1: 45 customers (22.5%)
- Cluster 2: 24 customers (12%)
- Cluster 3: 35 customers (17.5%)
- Cluster 4: 53 customers (26.5%)

Visual Analysis

Cluster Visualization Interpretation
The provided scatter plots reveal several key patterns:

1. Total Spend vs. Average Spend Plot
   - Clear linear relationship between total and average spend
   - Distinct separation between clusters along both axes
   - Evidence of spending-based segmentation

2. Density Distributions
   - Distinct peaks for each cluster
   - Minimal overlap between cluster distributions
   - Clear separation in spending patterns

Cluster Characteristics

Based on the visualization:

Cluster 0 (Blue)
- Lower total and average spend
- Highest density in the lower-left quadrant
- Likely represents budget-conscious customers

Cluster 1 (Orange)
- Moderate total spend with varying average spend
- Middle-range position in both metrics
- Represents typical regular customers

Cluster 2 (Purple)
- Scattered distribution
- Mixed spending patterns
- Potentially represents occasional high-value purchasers

Cluster 3 (Pink)
- Higher average spend
- Moderate to high total spend
- Represents premium customers

Cluster 4 (Green)
- Highest total spend values
- Consistent average spend
- Represents the most valuable customer segment

Recommendations

1. Targeting Strategies
- Develop specific marketing campaigns for each segment
- Focus on upgrading Cluster 0 customers to higher-value segments
- Maintain engagement with Cluster 4 customers

2. Product Development
- Align product offerings with cluster spending patterns
- Create premium options for Clusters 3 and 4
- Develop value propositions for Clusters 0 and 1

3. Customer Experience
- Customize service levels based on cluster characteristics
- Implement targeted retention strategies for high-value clusters
- Develop upgrade paths for lower-value clusters

Conclusion

The five-cluster solution provides a robust and actionable segmentation of the customer base. The low Davies-Bouldin Index (0.8819) confirms the statistical validity of this segmentation, while the clear visual separation and interpretable cluster characteristics support its practical utility for business decision-making.