CLUSTER ANALYSIS OF CONSUMER SPENDING PATTERNS FOR A MIDWESTERN RETAIL CHAIN USING SPSS

1. Background and Problem Statement

A mid-sized retail chain operating across four U.S. Midwestern states wanted to optimize its promotional strategies by segmenting its customer base based on spending behavior. The marketing team lacked actionable customer profiles despite collecting transactional data over two years. Traditional demographic-based segmentation was no longer driving effective campaigns. The company commissioned a data mining project to use SPSS for unsupervised learning—specifically cluster analysis—to identify distinct customer groups based on purchasing frequency, average transaction value, category preferences, and seasonal variation in spending.

2. Objectives

- To perform k-means clustering in SPSS to segment customers based on spending patterns
- To analyze characteristics of each customer cluster for actionable marketing strategy design
- To visualize cluster profiles for easy interpretation by non-technical stakeholders
- To recommend specific campaign strategies based on the derived segments

3. Methodology

3.1 Data Source and Preparation

- Dataset: 4,132 customer records (Jan 2022–Dec 2023)
- Variables used for clustering:
    - Total annual spend
    - Number of transactions per year
    - Average basket size
    - Percentage of spend on discount items
    - Spend by season (Spring, Summer, Fall, Winter)

Preprocessing in SPSS:

- Log transformation for right-skewed monetary values
- Standardization of variables using Z-scores
- Outlier removal based on interquartile range analysis
- Final dataset checked for multicollinearity and suitability for clustering using hierarchical method

3.2 Cluster Analysis in SPSS:

- Hierarchical clustering using Ward's method to determine the optimal number of clusters (found k = 4)
- Followed by k-means clustering with k = 4
- Cluster membership saved for analysis
- ANOVA and post hoc tests conducted to validate group differences
- Cluster profiles created using descriptive statistics and visualizations

4. Results and Cluster Profiles

Cluster 1: Bargain Shoppers (28%)

- High percentage spend on discount items (avg. 64%)
- Low average basket size but frequent transactions
- Most active in Winter season
- Campaign strategy: Weekly coupons and loyalty point accelerators

Cluster 2: High-Value Loyalists (16%)

- Highest total annual spend (avg. \$3,800)
- Balanced seasonal spending, low discount dependency
- High average basket size, consistent across categories
- Campaign strategy: Exclusive early access offers and VIP reward tiers

Cluster 3: Seasonal Big Spenders (22%)

- 60%+ of total spend concentrated in Summer and Fall
- High transaction value, fewer annual visits
- Strong preference for seasonal and limited-edition products

- Campaign strategy: Targeted seasonal catalogues and flash sale notifications

Cluster 4: Casual Buyers (34%)

- Low transaction frequency and total spend
- Impulse purchases, often under \$30
- Higher churn risk
- Campaign strategy: Win-back emails and personalized product recommendations

5. Recommendations

- Redesign email marketing flows based on cluster-specific triggers
- Allocate advertising budget based on cluster lifetime value (e.g., Cluster 2 gets high-value campaigns)
- Enable store managers to identify cluster membership of loyalty card users at POS terminals
- Automate cluster assignment using real-time SPSS syntax for new customers monthly

6. Deliverables

- SPSS .sav file with cluster membership and variables
- Syntax file for reproducible k-means clustering
- Visualizations: Cluster summary dashboard and comparative bar charts
- Written report outlining segment profiles and suggested campaign tactics

7. Stakeholder Relevance

Academic:

- Demonstrates practical use of unsupervised learning in a retail setting
- Suitable case for marketing analytics, retail operations, and SPSS coursework

Corporate:

- Framework adaptable to other regional or national retailers
- Supports customer relationship management (CRM) and data-driven promotion planning