Nastassia Pukelik
Pro. Jairam
CIS 9660
August 13, 2025

Business Report – Customer Segmentation & Market Basket Analysis

Summary:
We identified 2 customer segments across 25,325 customers. Segment sizes: C0: 5938, C1: 19387. The cleaned dataset contains ~40,417 transaction rows (date range: 2020-01-01 – 2025-09-05) ~ Product catalog analyzed: ~11 unique items.

Key Findings: Distinct segments differ by recency, purchase frequency, spending (Monetary), and average order value.

Average feature levels by segment:
Cluster 0: Recency=983.36, Frequency=1.0, Monetary=2717.83, AvgOrderValue=2717.83, UniqueProducts=1.0
Cluster 1: Recency=1055.05, Frequency=1.0, Monetary=726.76, AvgOrderValue=726.76, UniqueProducts=1.0

Top product associations with highest lift:
If buy {'Wall Clock'}, recommend{'Backpack'} (support=0.005, confidence=0.05, lift=0.44)
If buy {'Blue Pen', recommend {'Backpack'} (support=0.005, confidence=0.05, lift=0.43)
If buy {'Office Chair'}, recommend {'Desk Lamp'} (support=0.005, confidence=0.05, lift=0.42)

Actionable Recommendations

- Champions/Loyal: VIP rewards, early access, and exclusive bundles to increase retention and AOV.
- At-Risk: Win-back emails with targeted discounts on previously purchased categories.
- New/Low-Value: Onboarding sequences, first-purchase incentives, and recommendations based on popular items.
- Use high-lift rules to drive cross-sell modules on PDP/cart; A/B test recommendation placements.

ROI Projection (Illustrative)
Expected incremental revenue ≈ N × p × AOV, where N = customers targeted, p = conversion rate, AOV = average order value. Example: 1,000 targets × 5% × $80 ≈ $4,000.

Methodology & Validation
RFM feature engineering, outlier handling, and standardization preceded clustering. K-Means guided by elbow/silhouette determined K; hierarchical clustering provided corroboration. FP-Growth mined frequent itemsets; rules evaluated by support, confidence, and lift.

Limitations & Next Steps

- Expand segmentation analysis by experimenting with additional clustering methods and incorporating new behavioral features beyond RFM.
- Enhance the Market Basket module with advanced filtering, seasonal trend detection, and integration of real-time transactional data to generate up-to-date, personalized product recommendations.