📊 Customer Segmentation and Spending Analysis – Strategic Insights for XYZ Retail

This report presents a detailed analysis of customer purchase behavior and segmentation for XYZ Retail, derived from transaction data collected between January 2024 and February 2025. Our goal is to uncover insights that will drive more targeted marketing strategies, improve customer experience, and ultimately increase business revenue.

🛍️ Product and Spending Highlights

🔝 Top Selling Products and Categories

- **Top 5 Products**: Calming Serums, Soothing Creams, Lip Scrub, Micellar Water, Cream Moisturizer
- **Top 3 Categories**: Moisturizers, Cleansers, Lip Care

These insights highlight the popularity of skincare and lip care products. Promoting these bestsellers through bundles or loyalty perks could further boost sales.

💸 Spending Statistics

- **Min Transaction**: $20
- **Max Transaction**: $200
- **Average Transaction**: $90.73
- **Average Spend per Customer per Transaction**: *(See: `Data/result/Average_spending_per_customer_per_transaction.csv`)*

These metrics indicate a healthy mid-range purchase behavior, with potential to upsell to higher value tiers.

👥 Customer Segmentation Analysis

A. Rule-Based Segmentation: Understanding Behavioral Personas

We employed **behavioral segmentation** based on manually defined rules derived from customer spending, frequency, recency, and other purchase behaviors. This helps us align marketing messages with specific customer traits.

Segmentation Type	Segments	Purpose & Business Benefit
Total Spending	Low, Medium, High Spenders	Target budget deals vs. VIP perks
Average Spend per Order	Budget Buyer, Value Shopper, Big Buyer	Personalized product recommendations
Transaction Frequency	One-time, Occasional, Frequent Buyers	Re-engagement and loyalty targeting
Recency (RFM)	Active, At-risk, Lapsed	Timely campaigns to reduce churn
Purchase Pattern	New, Returning, Loyal Customers	Onboarding & retention strategies
Seasonal Behavior	Holiday, Regular, Off-season Buyers	Seasonal promotion planning
Basket Value	Low, Moderate, High Basket Buyer	Upselling and bundling strategies
Items per Transaction	Selective, Mixed, Bulk Buyer	Bundle optimization and messaging

📂 *Find detailed segment assignments in:*

- `Data/result/Customer_Segmentation_based_on_Purchase_Behavior.csv`
- `Data/result/Customer_Segmentation_based_on_Purchase_Pattern.csv`
- `Data/result/Customer_Segmentation_based_on_Basket_Composition.csv`

B. K-Means Clustering: Data-Driven Grouping

To complement the rule-based insights, we used **K-means clustering**, an unsupervised machine learning technique that automatically identifies natural groupings in the data. This approach allows for scalable, data-driven customer segmentation without predefined labels.

🔧 **Method:**

- **Features Used**: Total Spend, Average Spend per Transaction, Purchase Count
- **Dimensionality Reduction**: Applied **PCA** (Principal Component Analysis) for visualization
- **Optimal Clusters**: Chosen via **Elbow Method**, with k=4

📌 **Cluster Summary:**

Cluster	Label	Total Spend	Avg Spend	Purchase Count
3	High Spenders	$1616.54	$26.33	61.59
1	Moderate Buyers	$1116.99	$26.06	42.94
2	Frequent Bargain Shoppers	$723.07	$27.14	26.65
0	Occasional Buyers	$679.97	$23.90	28.45

🖼️ *Cluster visualization saved at:* `Data/result/PCA_Cluster_Visualization.png`
📁 *Full segmentation output:*
`Data/result/Customer_Segmentation_based_on_K-means_Clustering.csv`

📈 Why We Used These Techniques

✅ Rule-Based Segmentation

- **Advantages**: Simple, interpretable, and directly actionable for marketers

- **Use Case**: Ideal for aligning campaigns with known behaviors (e.g., lapsed customers, big spenders)

✅ K-means Clustering

- **Advantages**: Identifies hidden patterns in large datasets, unbiased by predefined labels
- **Use Case**: Useful when behavioral boundaries are fuzzy or evolving, providing fresh perspectives for new strategy design

✅ PCA for Visualization

- **Purpose**: Reduces feature complexity, enabling better understanding of customer clusters in 2D space
- **Use Case**: Makes cluster relationships visually interpretable for stakeholders

---

📊 Visualizing Customer Segments with Pie Charts

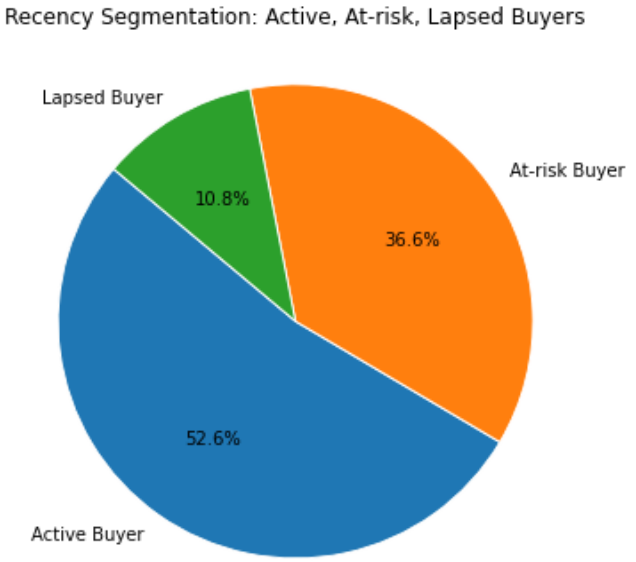
To complement the numerical and cluster-based segmentation insights, we've prepared a set of pie charts that offer a clear and intuitive view of how customers are distributed across various behavioral segments. These visualizations help summarize key customer dynamics at a glance.

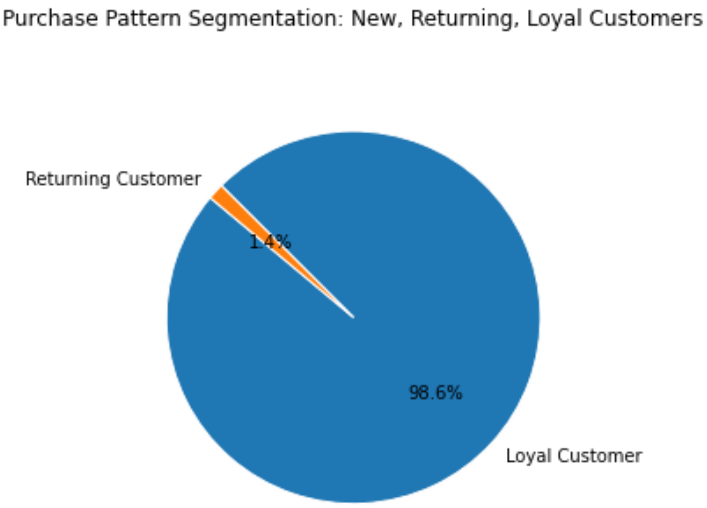
Each pie chart represents a different segmentation dimension:

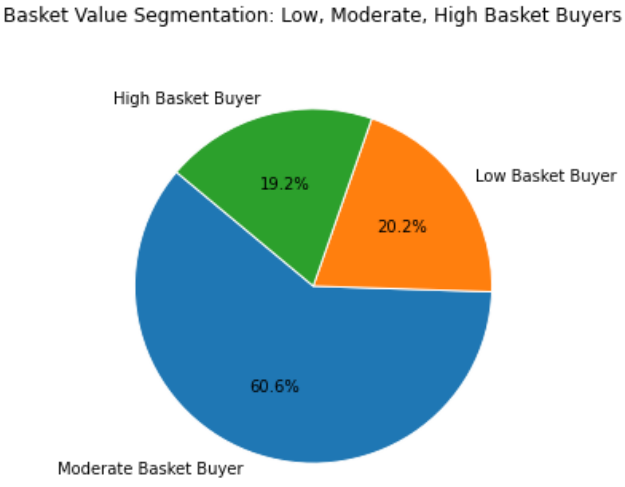
- **Total Spending** – Distribution of Low, Medium, and High Spenders
- **Average Order Value** – Proportion of Budget Buyers, Value Shoppers, and Big Buyers
- **Purchase Frequency** – Share of One-time, Occasional, and Frequent Buyers
- **Recency (RFM)** – Breakdown of Active, At-risk, and Lapsed Buyers
- **Purchase Pattern** – Share of New, Returning, and Loyal Customers
- **Seasonal Behavior** – Distribution of Holiday Shoppers, Regular Season Buyers, and Off-season Shoppers

- **Basket Value** – Categorization of Low, Moderate, and High Basket Buyers
- **Items per Transaction** – Segmentation into Selective, Mixed, and Bulk Buyers

These visual summaries provide quick insight into the dominant customer behaviors and help identify areas of opportunity for personalized marketing and engagement strategies







🚀 Strategic Recommendations

1. **Personalized Marketing**: Use segmentation to deliver targeted messages (e.g., reactivation emails for lapsed buyers, premium bundles for high spenders).

2. **Loyalty & Reward Programs**: Incentivize Frequent Bargain Shoppers and Loyal Customers to move up the value ladder.

3. **Seasonal Promotions**: Schedule campaigns around Holiday Shoppers and Regular Season Buyers.

4. **Upsell Opportunities**: Target Moderate and Mixed Basket Buyers with bundled product offers or free shipping thresholds.

5. **Churn Prevention**: Prioritize outreach to At-risk and Occasional Buyers with tailored incentives.

🧴 Product Recommendation Strategy

To enhance the shopping experience and drive cross-selling opportunities, we implemented two complementary **product recommendation techniques**:

1. **Collaborative Filtering** – leveraging customer similarity

2. **Association Rule Mining** – identifying co-purchase patterns

These methods provide data-driven recommendations tailored to customer behavior and preferences.

📦 1. Association Rule Mining (Market Basket Analysis)

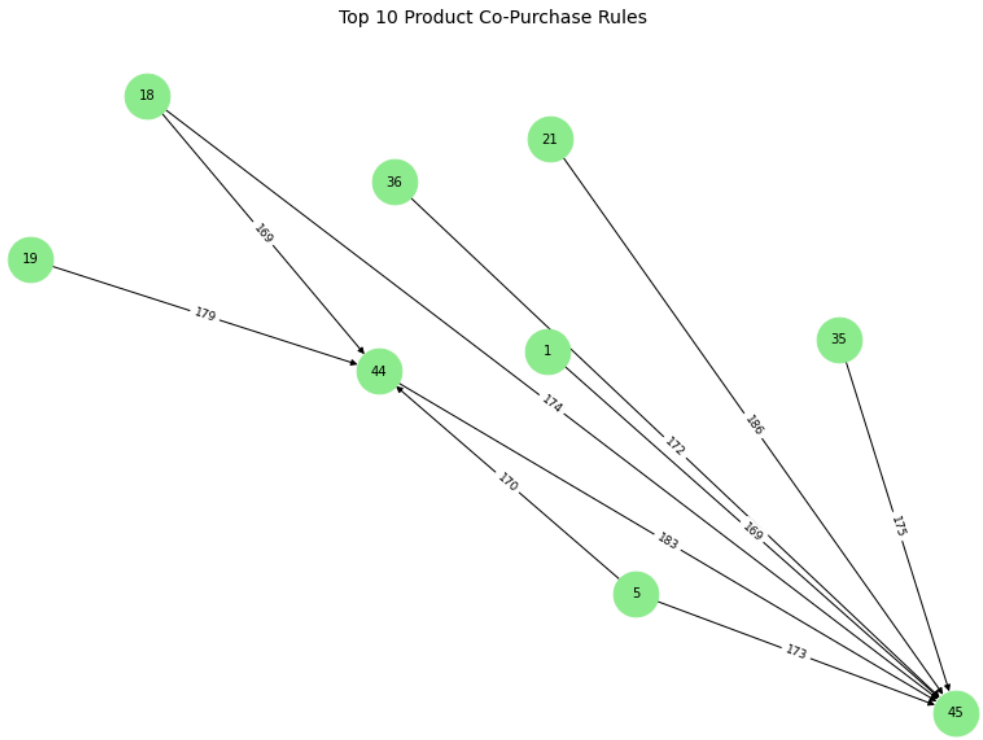
Technique:
Association Rule Mining (using the Apriori algorithm) is applied to transactional purchase data to find **frequently co-purchased items**. This method uncovers patterns like:

> "Customers who bought *Product A* also often bought *Product B*."

Use Case:
This is useful for identifying bundle opportunities, placement strategies, and upsell suggestions.

Visualization:
We created a directed network graph that illustrates the **Top 10 co-purchase rules**, where nodes represent products and arrows show the direction of association (i.e., from antecedent to consequent).

📁 *File location:*
`Data/result/Product_Recommendation-Market_Basket_Analysis.csv`



🔍 Product Recommendation Example — Customer ID: 1

🛍️ **Purchase History:**
18 skincare products including:

- **Cleansers:** Gel Cleansers, Oil Cleanser, Micellar Water, Foam Cleanser, Cleansing Balm
- **Lip Care:** Lip Balm, Lip Scrub, Lip Mask

- **Moisturizers:** Gel, Cream, Oil-Based, Water-Based, Night Cream
- **Acne/Pore Care:** Pimple Patches, Acne Spot Treatment, Scar Treatment Cream
- **Others:** Post-Sun Care, Healing Ointments

💡 **Top 5 Recommended Products:**

Product	Reason for Recommendation	Score
Soothing Creams	Frequently bought with moisturizers	18
Calming Serums	Often co-purchased with acne treatments	17
Eye Cream	Common with night creams & moisturizers	13
Eye Gel	Bought with eye cream & hydration routines	13
Under-Eye Patches	Associated with post-sun care & eye products	13

✅ **Logic:**

- Co-occurrence scores represent how frequently the recommended product appears with the customer's past purchases.
- Already purchased products are excluded.
- Recommendations are sorted by relevance.

👫 2. Collaborative Filtering (User-Based)

Technique:
Collaborative Filtering uses **cosine similarity** to compare users based on their product interaction patterns. Products are recommended based on what **similar users have purchased but the current user hasn't**.

Use Case:
Great for personalizing recommendations when users share common interests but diverse product paths.

📁 *File location:*
`Data/result/Product_Recommendation-Collaborative_Filtering.csv`

🔍 Product Recommendation Example — Customer ID: 238

🛍️ **Purchase History:**
17 skincare products including:

- **Cleansers:** Gel, Micellar Water, Cleansing Balm
- **Lip Care:** Lip Balm, Lip Scrub, Lip Mask
- **Moisturizers:** Gel, Night Cream, Oil-Based, Water-Based
- **Acne/Pore Care:** Pore Minimizing Serum, Pore Strips
- **Others:** Calming Serums, Eye Cream, Eye Gel, Soothing Creams, Under-Eye Patches

👥 **Most Similar Customers:**

Customer ID	Similarity Score
359	0.894
252	0.846
204	0.846
304	0.838
320	0.831

💡 **Top 5 Recommended Products:**

Product	Reason for Recommendation	Score
Soothing Creams	Common among users with moisturizing preferences	6
Oil Cleanser	Frequently purchased by similar customers	5

Foam Cleanser	Part of typical cleansing routines	4
Pimple Patches	Popular among acne-care focused users	4
Acne Spot Treatment	Consistently bought by users with similar behavior	4

✅ **Logic:**

- Recommendations come from the top 5 most similar users.
- Products already bought by the target user are removed.
- Scores reflect how many similar users purchased the product.

⚠️ Limitations and Considerations

1. **For New Products:**
   Recommend them to customers who frequently purchase in the same category or show interest in related items.

2. **For New Customers (Cold Start Problem):**
   Use popular products and top-selling items as default recommendations until behavioral data is available.

🎯 Business Value

- Boosts **personalization** and relevance in the shopping journey
- Improves **conversion rates** by surfacing products aligned with user interests
- Drives **cross-sell** through bundling insights and peer patterns
- Supports **new product discovery** for both loyal and new customers