CUSTOMER SEGMENTATION WITH PYTHON CLUSTERING FOR A U.S. FASHION E-COMMERCE BRAND

1. Background

A rapidly scaling fashion e-commerce company in the U.S. wanted to move beyond generic campaign targeting. Their marketing team handled hundreds of thousands of customers across the country, but lacked structured customer segments based on real behavior. Personalization strategies were underperforming due to reliance on basic filters like gender and geography.

We were brought in to use Python for advanced customer segmentation using behavioral data—purchases, recency, and browsing engagement—to power a targeted marketing strategy.

2. Objective

- To segment active customers into distinct behavioral groups using clustering in Python
- To apply RFM (Recency, Frequency, Monetary) modeling as a base for K-Means clustering
- To visualize and interpret customer profiles and use them to guide email, offer, and retargeting strategies
- To generate re-usable Python code and dashboard-ready outputs

3. Data Used

Source: Internal CRM and transaction log (Jan 2022 – Mar 2023)

Fields Included:

- Customer_ID, Last_Purchase_Date, Total_Orders, Total_Spend, Average_Order_Value, Product_Category_Preference, Email_Clicks, Site_Visits

Data Size: 97,000 unique customers after filtering for activity in last 12 months

4. Methodology

4.1 RFM Feature Engineering

- **Recency** = Days since last purchase

- **Frequency** = Total number of transactions
- **Monetary** = Total spend in USD
- Applied log transformation and min-max scaling

4.2 Clustering Approach

- Chose **K-Means Clustering** for interpretability and scalability
- Optimal number of clusters determined using:
    - **Elbow Method**
    - **Silhouette Score**
- Final model: **5 clusters**
- Libraries used: pandas, scikit-learn, matplotlib, seaborn

4.3 Visualization

- Used 2D PCA to visualize cluster separability
- Generated segment profiles based on statistical summaries

5. Segment Results

Segment Name	Size	Description
VIP Loyalists	8%	High frequency + high spend + recent activity
Bargain Shoppers	21%	Moderate spend, high frequency, strong coupon engagement
One-Time Buyers	34%	Single or two purchases, long recency
Window Shoppers	19%	High visits, low conversion, high click-through rates
Churn Risk Customers	18%	Inactive for 6+ months, historically above-average spend

6. Interpretation and Strategy

- **VIP Loyalists** were offered early access to product drops and “thank you” rewards
- **Bargain Shoppers** targeted with flash sale and coupon-heavy campaigns
- **Churn Risk Customers** received automated win-back sequences
- **Window Shoppers** were shown retargeting ads for the most browsed product categories

- **One-Time Buyers** were nudged with cross-sell product bundles based on initial purchase

7. Reporting Output

- **Python Jupyter Notebook**:
    - Clean RFM pipeline
    - Cluster generation and export
    - PCA + elbow/silhouette visualizations
- **PDF Report (15 pages)**:
    - Executive summary of segments
    - RFM distribution by segment
    - Visual profiles and suggested marketing actions
- **Excel Sheet**:
    - Customer ID with assigned segment
    - Summary tables for email campaign mapping
    - Filters by state, gender, and product category

8. Business Impact

- Within 6 weeks:
    - **Email conversion rate** rose by **28%**
    - **Cart abandonment recovery rate** improved by **19%** among Bargain Shoppers
    - Marketing cost per acquisition dropped by **11%** through better targeting
    - Segment-based strategy now used for **quarterly re-activation campaigns**

Statssy