

Project Report: E-commerce Customer Segmentation & Strategic Recommendations (RFM Customer Segmentation Analysis)

- **Project Type:** Business Analysis, Data Analytics, & Marketing Strategy
 - **Stakeholders:** Head of Marketing, Head of Sales, CRM Manager, CFO
-

1. Executive Summary

This analysis was conducted to address critical business challenges of inefficient marketing spend and low customer retention (15% repeat-purchase rate). By applying the **RFM** (**R**ecency, **F**requency, **M**onetary) framework to 541,909 transactions, we segmented 4,316 unique customers, revealing a significant Pareto Principle: the top **21% of customers** ("Champions" & "Loyalists") generate **67% (over \$10.7M)** of total revenue.

Conversely, the largest segment by count, "Lost Customers" (30% of the base), contributes only 5% of revenue. This report confirms that the current "one-size-fits-all" marketing strategy is highly inefficient.

Key findings identify a high-value, **\$2.24M "At Risk" segment** requiring immediate intervention and a high-potential "New Customer" segment (420 customers) critical for future growth. This report proposes a data-driven, 4-point strategic action plan to optimise marketing ROI, nurture high-potential customers, and mitigate churn.

2. Business Problem & Project Objectives

2.1. Business Problem Statement

The company's current undifferentiated marketing strategy is inefficient. It results in low customer retention (15% repeat-purchase rate) and a high Customer Acquisition Cost (CAC), as significant resources are spent on low-value or inactive customers. This failure to maximize Customer Lifetime Value (CLV) directly impacts profitability.

2.2. Project Objectives

1. **Analyze & Segment:** To analyze the transactional dataset (UCI Online Retail) and segment the entire customer base into distinct, behavior-based groups using the RFM model.
 2. **Identify Key Segments:** To isolate and profile high-value, high-potential, and high-risk customer segments.
 3. **Develop Strategy:** To recommend specific, data-driven marketing and CRM strategies for each key segment to optimize resource allocation, improve retention, and maximize profitability.
-

3. Methodology & Toolkit

This project employed a structured analytical process, translating raw data into strategic recommendations.

1. **Framework:** The **RFM (Recency, Frequency, Monetary)** model was selected as the industry-standard framework to quantify and categorize customer behavior as a proxy for CLV.
 2. **Data Ingestion & Cleaning (ETL):** **Power Query (in Excel)** was used for the initial Extract, Transform, and Load (ETL) process on the 541,909-row dataset. This tool was chosen for its robust capabilities in handling missing data, filtering, and transforming data at scale before analysis.
 3. **Data Aggregation & Scoring:** **Excel (PivotTable with Data Model & Formulas)** was used to aggregate the cleaned transaction-level data to a customer-centric view.
 - PivotTable (Data Model): Enabled Distinct Count of InvoiceNo (Frequency) and Sum of TotalPurchaseValue (Monetary) to be calculated per CustomerID.
 - Excel Formulas: DATE functions were used to calculate Recency (Days). PERCENTILE.INC was used to establish the four cut-off points (quintiles) for scoring. Nested IF formulas were then applied to assign R, F, & M scores (1-5) to all 4,316 customers.
 4. **Visualization & Analysis:** **Power BI** was used to create an interactive, 1-page dashboard. This tool was chosen for its strong data visualization capabilities (Treemap, Heatmap Matrix) and its ability to clearly communicate insights to non-technical stakeholders via interactive slicers.
-

4. Data Integrity & Cleaning Process

Data integrity was paramount. The raw 541,909-row dataset was subjected to a rigorous cleaning process using Power Query.

- **Missing CustomerIDs:** 135,080 rows (24.9%) were removed as they were non-attributable and thus unusable for RFM segmentation.
 - **Cancellations:** 8,905 rows (where InvoiceNo started with 'C') were removed to prevent negative quantities from skewing monetary analysis.
 - **Invalid Transactions:** 40 rows with a UnitPrice of 0 were removed as they do not represent valid monetary sales.
 - **Feature Engineering:** A TotalPurchaseValue (Quantity * UnitPrice) column was created to calculate the value of each transaction line.
 - **Result:** A clean, analysis-ready dataset of **397,884 valid transactions** across **4,316 unique customers** was established.
-

5. Key Findings & Analysis of Segments

The dashboard analysis of the 4,316 customer segments revealed significant concentration and actionable insights.

Segment	Customer Count	% of Customers	Total Revenue (Approx.)	% of Revenue	Key Insight
Champions	463	11%	\$7.37M	46%	(High Value) Your most valuable asset.
Loyal Customers	433	10%	\$3.37M	21%	(High Value) Your consistent, core buyers.
At Risk	478	11%	\$2.24M	14%	(High Risk) High-value customers you are about to lose.
Average Customers	1,100	25%	~\$1.60M	10%	(Low Value) The "middle," low-revenue group.
New Customers	420	10%	~\$0.80M	5%	(High Opportunity) Your next generation of Champions.
Lost Customers	1,300	30%	~\$0.80M	5%	(High Waste) Your largest, least valuable segment.

- **Insight 1: Revenue is Heavily Concentrated.** The Pareto Principle is strongly evident. The top 21% of customers ("Champions" & "Loyal") generate 67% (over \$10.7M) of total revenue.
 - **Insight 2: A High-Value Segment is at High Risk.** The "At Risk" segment, representing **\$2.24M** in historical revenue, consists of previously frequent buyers who are now inactive (Low Recency, High Frequency) and require immediate intervention.
 - **Insight 3: Significant Marketing Inefficiency Identified.** The "Lost Customers" segment is the largest by count (30% of the database) but contributes only 5% of revenue. Marketing spend directed at this group is likely yielding a negative ROI.
 - **Insight 4: A Clear Conversion Opportunity Exists.** The "New Customers" segment (420 customers) is the foundation of future growth. They are recently active (High Recency) but have only purchased once (Low Frequency).
-

6. Strategic Recommendations & Business Case

Based on these findings, a "one-size-fits-all" strategy is detrimental. We recommend adopting the following 4-point segmented action plan to optimize resource allocation.

1. Segment: "Champions" (11% of Customers, 46% of Revenue)

- **Recommendation: PROTECT & REWARD.**
- **Action Plan:** Implement a VIP program (e.g., "UrbanThread Gold") offering non-monetary perks such as early access to new products, free shipping, and dedicated customer service. Avoid discounts, as this group is not price-sensitive and it would erode margins.
- **Business Case:** Focuses on loyalty and recognition, which is a low-cost, high-impact method to protect our largest revenue stream (\$7.37M) and increase their lifetime value.

2. Segment: "At Risk" (11% of Customers, 14% of Revenue)

- **Recommendation: WIN-BACK (URGENT).**
- **Action Plan:** Launch an aggressive, personalized "We Miss You" campaign. Offer a significant, one-time, high-value discount (e.g., 20% off) to incentivize re-engagement.
- **Business Case:** This is a high-ROI "rescue" mission. Rescuing even 10% of this segment could recover **\$224,000** in revenue that is otherwise on the verge of being lost.

3. Segment: "New Customers" (10% of Customers, 5% of Revenue)

- **Recommendation: NURTURE & CONVERT.**

- **Action Plan:** Implement an automated 3-part email onboarding series focused on securing the *second* purchase. (Email 1: Welcome & brand story. Email 2: Product recommendations. Email 3: Small incentive for their next purchase).
- **Business Case:** This strategy increases the probability of converting a high-cost, single-purchase customer into a profitable, long-term "Loyal Customer," directly improving retention rates.

4. Segment: "Lost Customers" (30% of Customers, 5% of Revenue)

- **Recommendation: DE-PRIORITIZE & REDUCE SPEND.**
- **Action Plan:** Immediately remove this 1,300-customer segment from all expensive marketing campaigns (e.g., print catalogs). Reduce email contact to a minimal, low-cost re-engagement effort (e.g., 2-3 times per year).
- **Business Case:** This action immediately frees up a significant portion of the marketing budget, which can be re-allocated to the high-ROI "Win-Back" and "Nurture" campaigns, drastically improving overall marketing efficiency.