

Retail Fashion Company



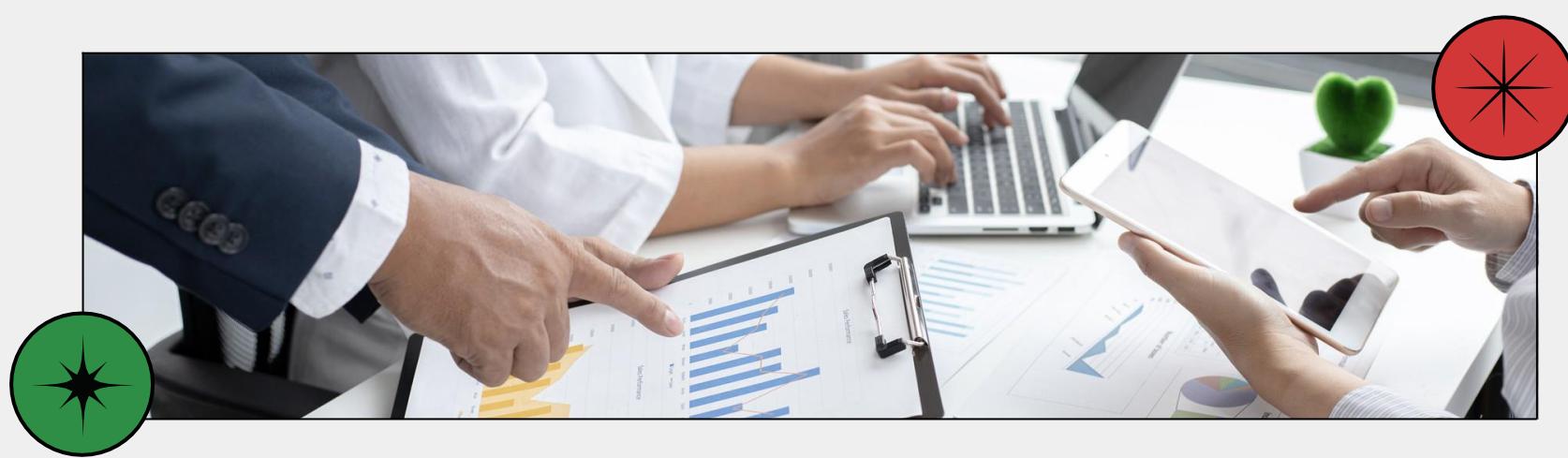
Cohort & Retention Analysis

with RFM Segmentation



Faraj Hafidh |





LIST OF CONTENT

- ◆ Project Overview
- ◆ Business Problem
- ◆ Problem Statement & Objective
- ◆ Methodology
- ◆ Analysis & Insight
- ◆ Recommendations
- ◆ Appendix

Project Overview

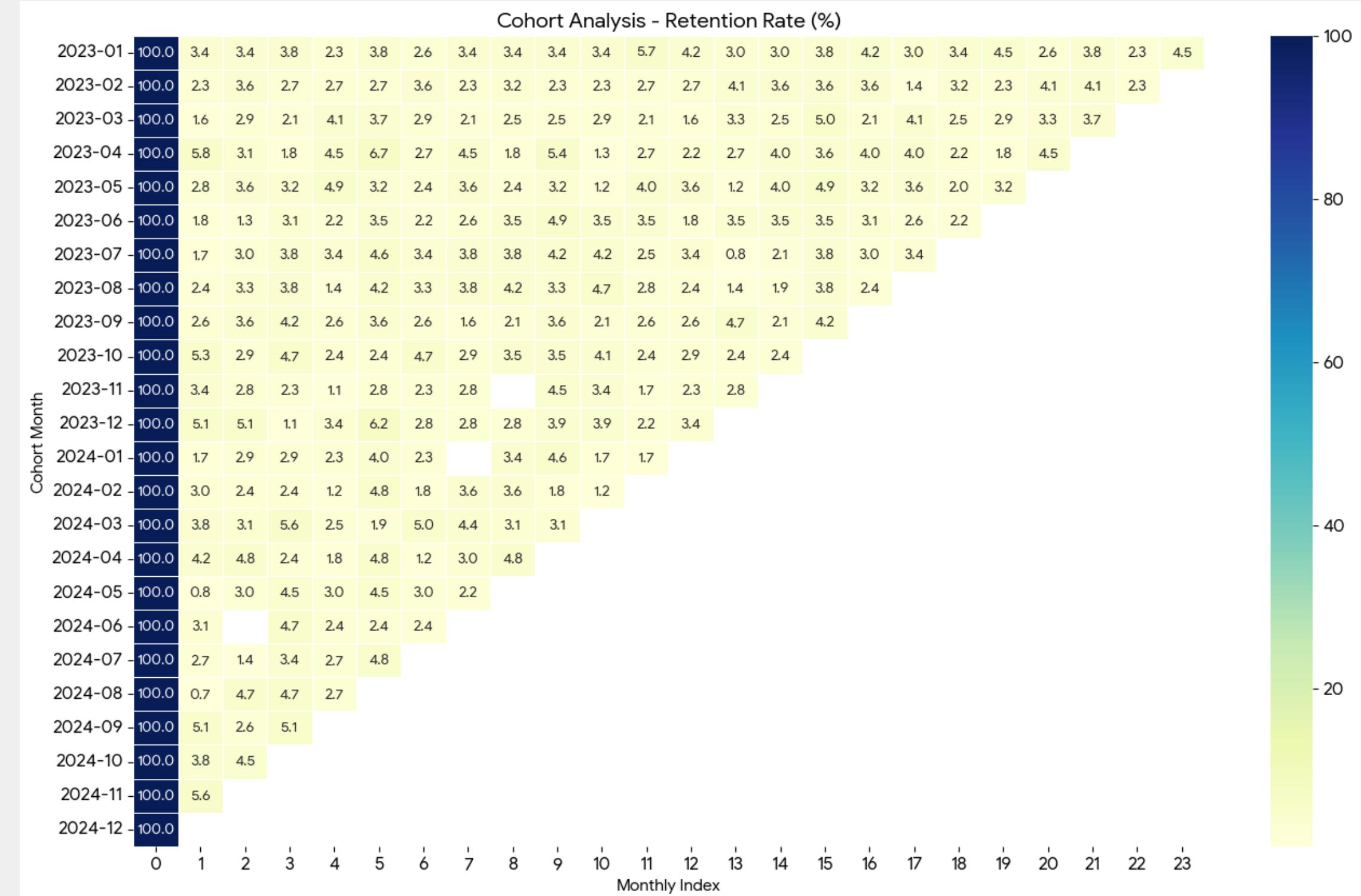
This project provides actionable business insights by combining Cohort Retention Analysis and RFM Customer Segmentation to understand both when customers churn and which customers are worth retaining.

This dataset covers retail fashion transactions from 2021-2024, but this analysis focuses on 2023-2024 to reflect current business conditions and provide relevant strategic recommendations



BUSINESS PROBLEM

A retail fashion company faces a critical retention crisis where 95% of customers never return after their first purchase, leading to a 99.6% collapse in new customer acquisition (265 customer \rightarrow 1 customer over 2 years). The business is projected to lose 15-20% of revenue without immediate intervention to shift acquisition-dependent to retention-focused strategy.

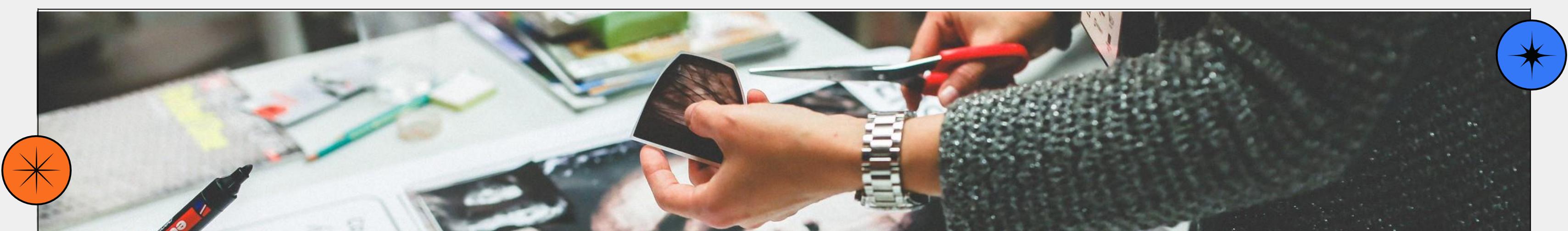


Problem Statement

Analyzing important factors of customer retention performing, and which customer segments contribute most revenue during the 2023-2024.

Objective

- To understand when customer churn and which customer are worth to retaining.
- Giving business recommendation to create targeted strategies.



Methodology

Define Objective

Determine the objective and scope

Data Cleaning

Cleaning Dataset using SQL.

Insight & Recommendations

Provide Insights & Recommendations

Data Gathering

Conduct data gathering with SQL

Data Analysis

Conduct data analysis with EDA using SQL, Excel, PowerBI



Appendix : Data Gathering, Cleaning and Clustering

Data Gathering : with SQL (start from here)

```
1  SELECT
2      s.transaction_id,
3      s.date,
4      s.customer_id,
5      -- Data dari Customer
6      c.age_group,
7      c.gender_cleaned,
8      c.city AS customer_city,
9      -- Data dari Product
10     p.category_cleaned,
11     p.list_price,
12     p.cost_price,
```

```
1  -- Data dari Store
2      sd.store_name,
3      sd.region AS store_region,
4      -- Data Transaksi
5      s.quantity,
6      s.returned,
7      s.discount
8  FROM
9      v_sales_cleaned AS s
10     INNER JOIN v_customer_cleaned AS c ON
11         s.customer_id = c.customer_id
12     INNER JOIN v_product_cleaned AS p ON
13         s.product_id = p.product_id
14     INNER JOIN store_data AS sd ON
15         s.store_id = sd.store_id
16     WHERE
17         s.date >= '2023-01-01'
```

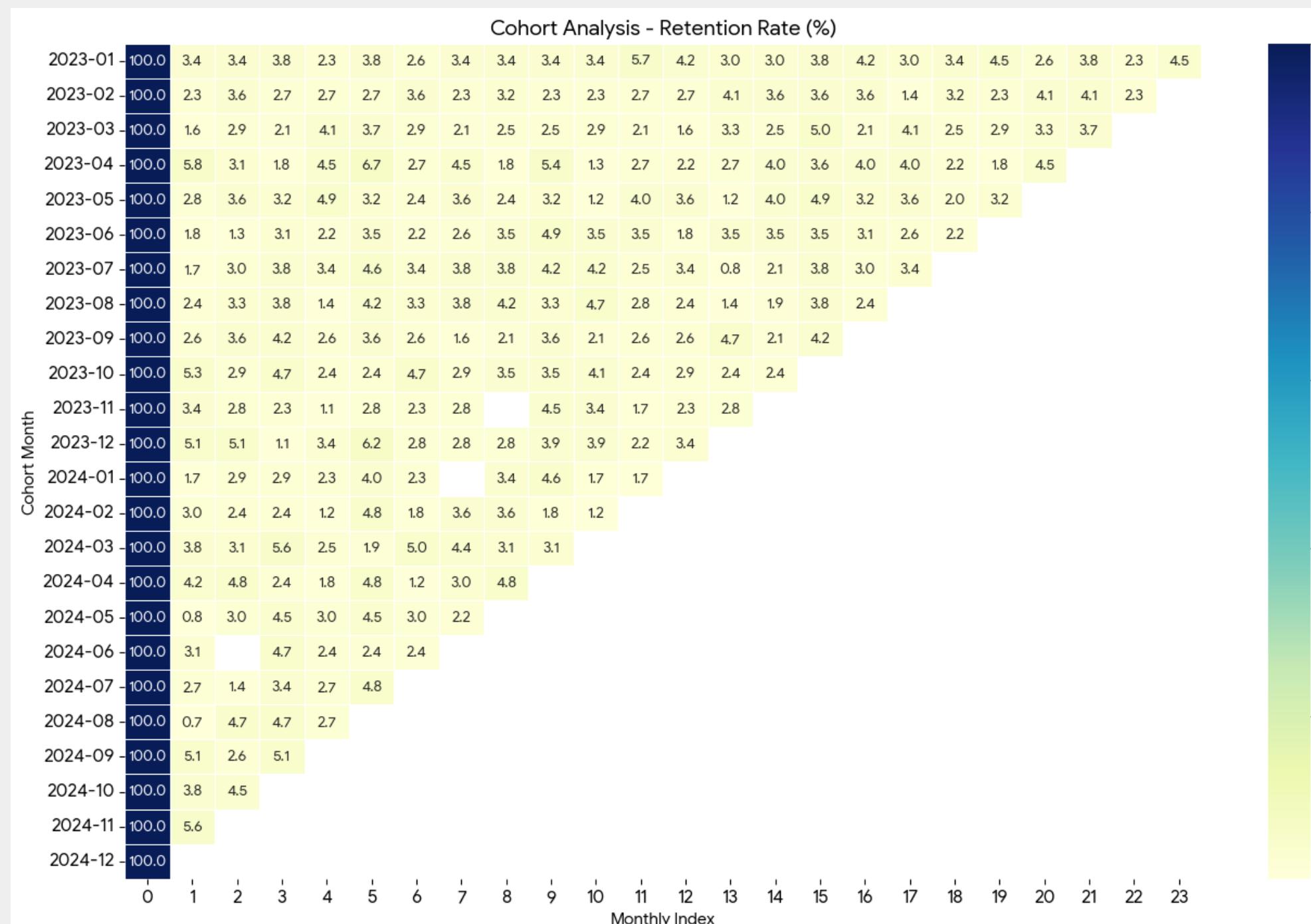
Data Cleaning



RFM and Cohort Query



Cohort Retention Analysis – to understanding WHEN Customer Churn



The business faces a critical retention crisis with 95% of customer churning immediately after their first purchase across all cohort from 2023-2024. This systematic problem is evident in the January 2023 cohort where 265 customers were acquired but only 3.4% returned in month 2, losing 256 customers instantly. The Situation has deteriorated significantly as new customer acquisition collapsed 99.6% from 265 customers in January 2023 to just 1 customer in December 2024, indicating severely weakening acquisiton channels and highlighting an unsustainable business model dependent on constant new customer influx.

While Q4 Periods (Oct-Dec) consistently show relatively stronger retention rate Of ~5%, with peak performance in April 2023 (5.8%) and November 2024(5.6%), the business experienced failure during Q2-Q3 2024.May and august 2024 record with the lowes retention with (0.75% and 0.67%), suggesting potential operational issues, product quality problems, or market saturation during mid-year.

Customer Segmentation Using RFM Analysis - Understanding WHO our valuable Customers Are

Why RFM Analysis Matters ?

Cohort analysis shows that 95% of customers never return after their first purchase. However, it cannot answer:

- Which customers in that remaining 5% are actually driving revenue ?
- Are we losing high-value customers or low-value ones ?
- Which segments should receive priority in retention campaigns ?

RFM (recency, frequency, monetary) analysis segmenting customers based on behavioral pattern to identify :

- **High-value Customers** worth retaining (*Champions, Loyal*)
- **At-risk Customers** who used to be valuable (*Need Attention*)
- **Lost Customers** Who may be reactivable (*Hibernating, Lost*)
- **Low Value Customers** not worth *retention investment*

RFM Methodology Summary:

Step	Description
Data Integration	Sales, customer, product, and store tables are joined after data cleaning. Analysis period is limit to 2023-2024.
Aggregation	Transactions are aggregated at customer level to calculate recency, frequency, monetary and average order value
Scoring	RFM Values are scored using quantile-based segmentation (NTILE 1-5), where higher scores indicate stronger performance
RFM Cell	RFM Cell is created by concatenating R,F and M values (e.g., 555). Total RFM scores is calculated by summations.
Segmentation	Customers are grouped into business segments such as Champion, Loyal, Potential Loyal, Hibernating Lost and At Risk.

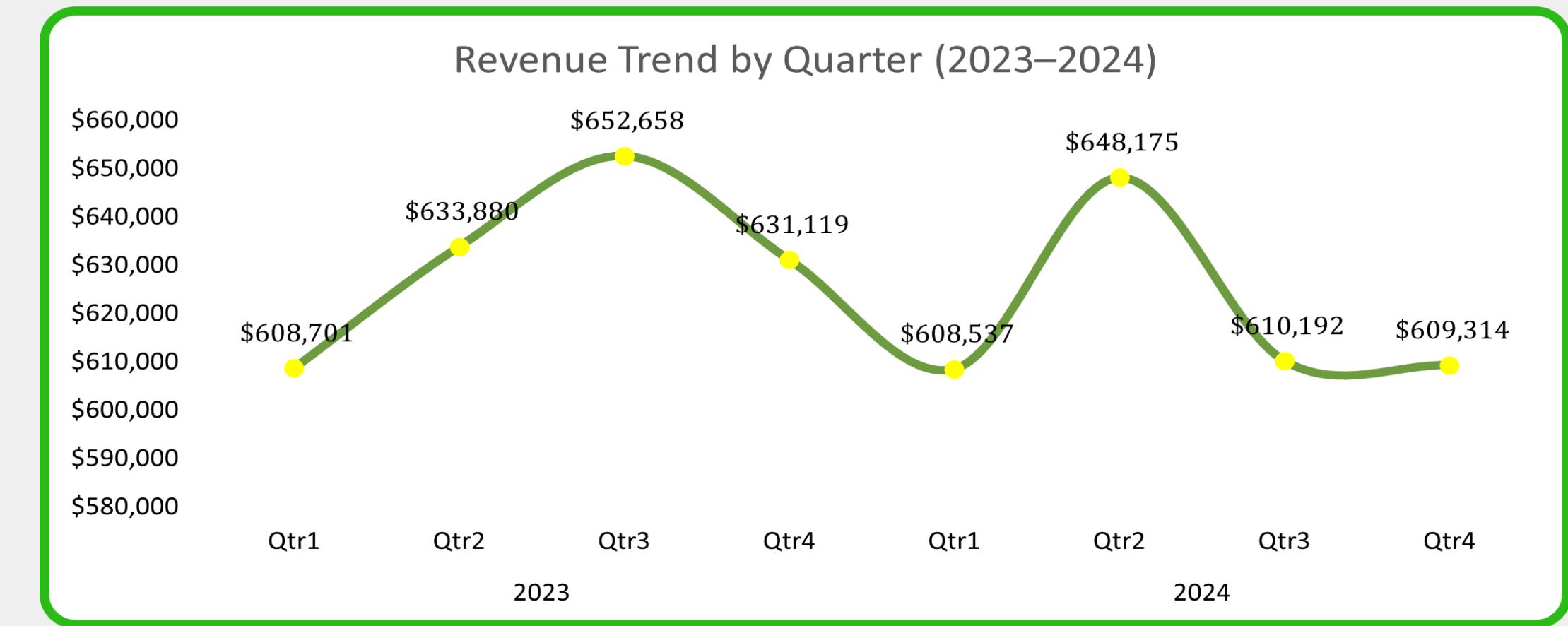
Combined Insights - Connecting Retention Patterns to customer Value

Quarterly revenue trends for 2023 to 2024

In Q1-Q3 2023, revenues grew from **\$608K** to **\$652K**. In other words, revenues grew by **+7.2%** in the first half of 2023. However, in Q4 there was a significant **decline to \$631K (-3.2%)**. Overall, this year the retail sector generated a total revenue of **\$2.526M**.

In **Q1-Q2 2024**, revenue showed a strong recovery to **\$608K** and, followed by a peak in Q2 at **\$648K (+2.7% compared to Q4 2023)**. However, Q3-Q4 experienced a continuous **decline from \$610K to \$609K (-6% from the Q1-Q2 peak)**. Overall, in 2024, retail generated a total revenue of **\$2.476M**, representing a **-1.98% decrease** compared to the **previous year**.

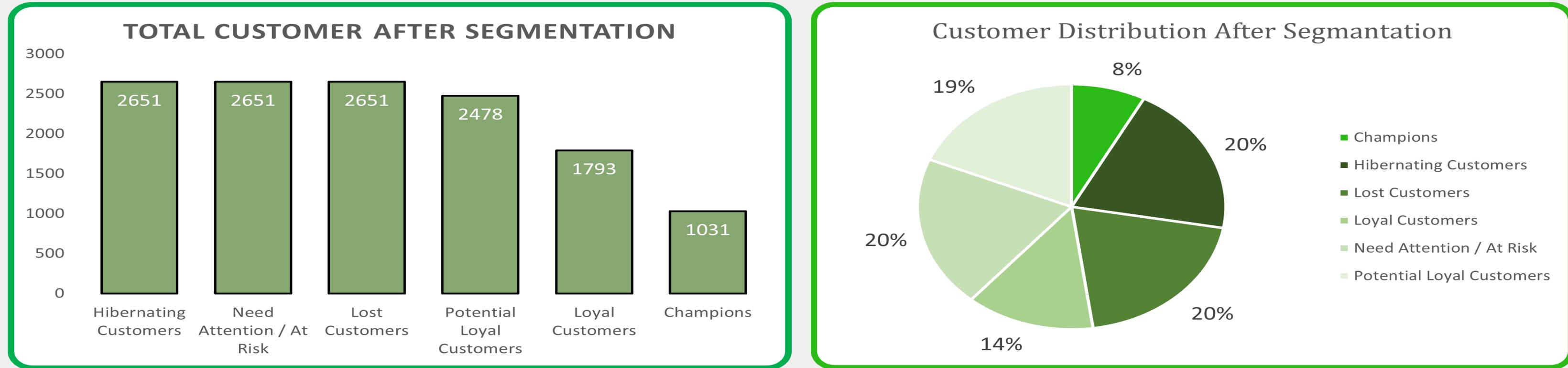
* **Key Insight:** The business is **trappe** in a "**Leaky Bucket**" model - constantly needing new customers because existing ones don't return, making revenue entirely dependent on acquisition volume.



Connecting to Cohort Insights: When we overlay cohort data with revenue trends:

1. Q4 revenue drops align with cohort size decline
 - > Fewer new customers = lower revenue potential.
2. Q1 up to Q2 revenue spikes coincide with larger cohorts
 - > Acquisition volume still drives short-term revenue.
3. Mid-year revenue decline (Q3 2024) matches retention collapse
 - > May-Aug 2024 cohorts had **0.67-0.75%** retention and contributed minimal revenue.

Distribution of Customer after we do segmentation using RFM



Connecting to Cohort Churn

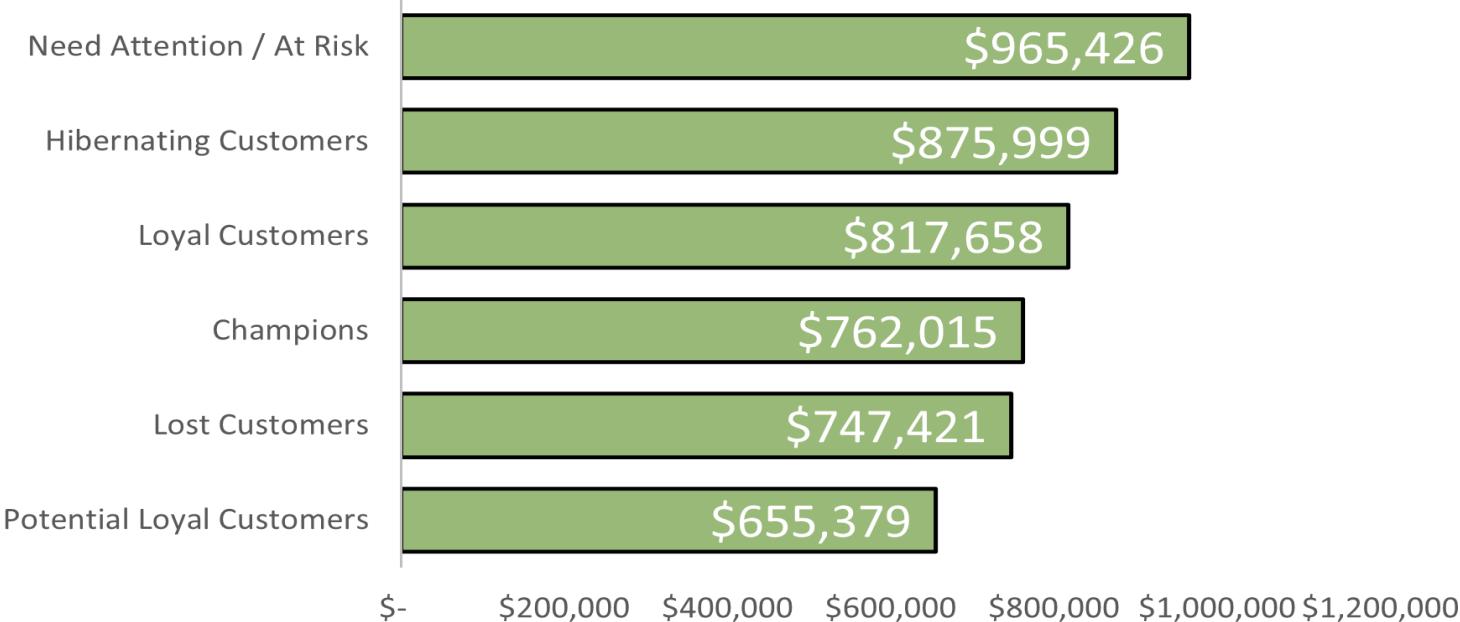
The 60% concentration in Hibernate/At-Risk/Lost segments directly validates the cohort finding that 95% of customers churn immediately:

- Most customers make 1-2 purchases then disappear
 - > classified as "Hibernating" or "Lost".
- Very few into "Loyal" or "Champions"
 - > only 22% combined.
- The small Champion segment (~8%) represent the 3-5% who actually return in cohort analysis

This Confirms: The cohort churn pattern is not just about timing - it's about fundamental customer behavior where the majority are one-time buyers

Analyzing Revenue Distribution Across RFM Segments

TOTAL REVENUE BY CUSTOMER SEGMENT



The Retention Paradox:

1. Champions (8% Customers) generate 16% of revenue

- High customers lifetime value, but we're failing to create more Champions (cohort analysis shows 95% churn).

2. At-Risk + Hibernating + Lost (60% of customers) generate 53% of revenue

- The business is heavily dependent on customers who already churning or inactive
- This explains why revenue is declining in 2024 - we're losing the customers who drive revenue

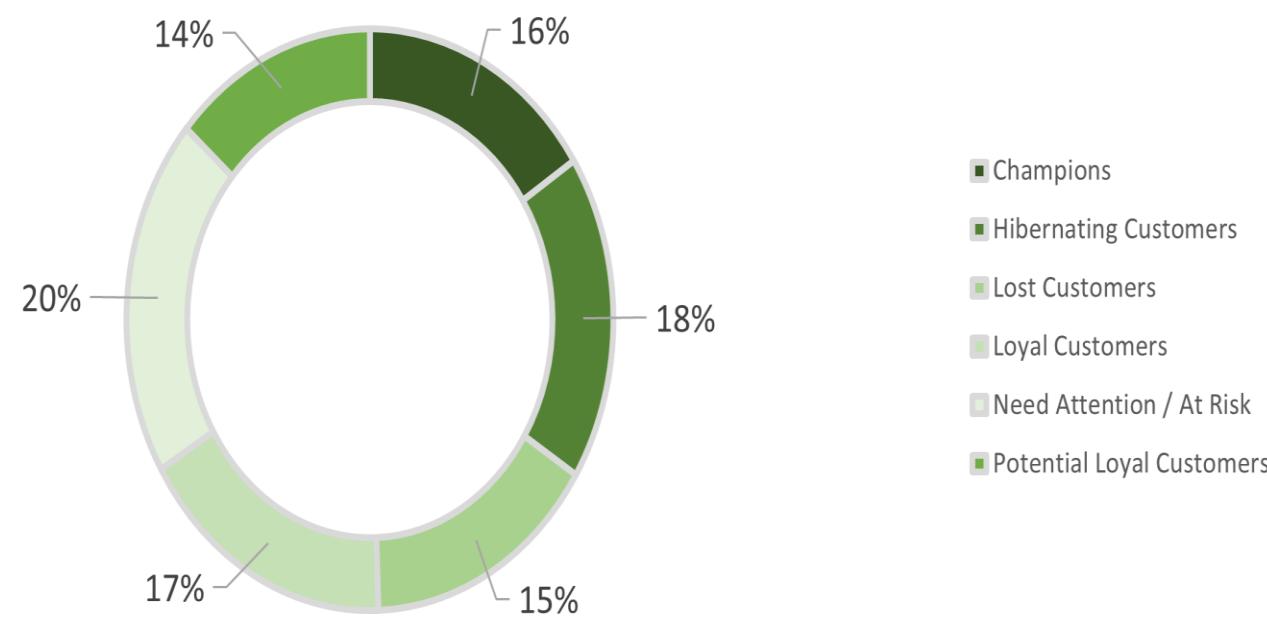
3. Revenue concentration in declining segments = unsustainable model

- As cohort size shrink (265 -> 1 customers) and retention stays at 3-5%, the pool of "at risk" and "hibernating" customers will eventually run dry.

4. Game-Changing Insight: Loyal Customers are Undervalued

- Loyal customers (16.95%/17% of revenue) match Champions (16%) in total contribution.
- With 1,793 Loyal vs 1031 Champions, Loyal segment has 1.7x more growth potential.
- Increasing Loyal customer spending by just 25% would make them the #1 revenue contributor.
- Strategy shift: Focus on Loyal → Champion conversion, not just creating Champions from scratch

Revenue Distributin by Customer Segment



After all of this, do you think will answers our cohort question: "Are the 3-5% who return valuable enough to retain ?"

Answer: Yes

Those who become Champions deliver 2.6x more revenue per customer than the average. The problem is we create too few of them.

RECOMMENDATIONS-1

Fix the first-month experience

Cohort analysis shows reveals a dramatic drop-off after the initial purchase.

RFM demonstrate that 60% go into hibernating/at risk/lost, that show there is no engagement after purchase

Action:

- Implement a 30-to-90 day onboarding program for new customers.
- Send personalized follow-ups within 7 days of the first transaction.
- Offers a 'second purchase incentive' within 14 days (data reveals most churn occurs in the first month)
- Survey first-time buyers to identify friction issues

Protect and Grow Champions

Too few customers develop into champions despite their high value.

Customers cannot become high-value champions due to a lack of differentiated experiences and clear growth pathways

Actions:

- Create VIP loyalty program for champions and loyal customers.
- Provide early access to new products, exceptional savings, or customized style.
- Assign dedicated account management to the top 1031 champions.
- Analyze the buying patterns of champions vs. one-time buyers to identify 'Champions indications'

RECOMMENDATIONS -2

Reactivate At-Risk,Lost and Hibernating Customers (53% of revenue)

More than half of revenue comes from customers who are already disengaged.

Without proactive retention strategies, valuable customers can disengage unnoticed, resulting in income loss.

Actions:

- Create bespoke 'we miss you' offers to re-engage hibernating consumers.
- Re-engagement program for at-risk clients includes exclusive discounts and new product releases
- Segment by prior purchase category and deliver tailored product recommendations.
- Reactivate consumers with a previous AOV of \$100 or more (greater ROI potential)

Optimize Acquisition Strategy (Address Declining Cohort Size)

New customers acquisition dropped from 265 in January 2023 to 1 in December 2024

The current acquisition technique draws low-quality leads who do not want to become repeat customers.

Strategic Shift: Prioritize Quality over Quantity

- Avoid mass acquisition initiatives as they often result in dissatisfied customers rather than loyal supporters.
- Target customers with champions profiles:
 - Higher income populations.
 - Fashion fanatic (social media engagement)
 - Previous brand references indicators
- Analyze Q4 Cohorts (Oct-Dec) which show 5% retention vs 3% average - duplicate what works

RECOMMENDATIONS - 3

Address Q2-Q3 Seasonal Weakness

Problem:

- Cohort analysis: May-Aug 2024 had the lowest retention (0.67-0.75%)
- Revenue analysis: Q2-Q3 2024 experienced saw steady continuous decline.

Actions:

- Investigate mid-year operational difficulties such as delivery delays, stocks-outs, and quality problems.
- Launch seasonal initiatives in Q2-Q3 to address prior weakness.
- Replicate Q2 2023's success (which saw high sales growth). Analyzed that works
- Consider hosting mid-year loyalty events (summer bargains, special previews) to increase engagement.

Leverage Q4 Acquisition Strength

Q4 indicated improved retention (~5%) but revenue still declines.

Insufficient activation and engagements strategies fail to convert higher-quality Q4 customers into repeat purchases despite their great retention potential.

Actions:

- Double focus on Q4 on marketing and devote 40% of the annual acquisition budget here.
- Analyze what differentiate Q4 customers (Christmas shoppers versus brand loyalist)
- Increase Q4 engagement into Q1 with 'New year refresh' advertisement
- Track Q4 cohorts through 2025 to evaluate if 5% retention is sustained.

Retail Fashion Company

**THANK
YOU!**

