Customer Retention Risk & Revenue Stability Analysis

This project analyzes customer behavioral and transactional data to identify churn risks and quantify their impact on long-term revenue stability. Our objective is to proactively detect high-value customers with declining retention and pinpoint product categories contributing to churn-risk revenue.

Project Overview

1

Identify At-Risk Customers

Detect high-value customers showing signs of churn.

2

Quantify Revenue Impact

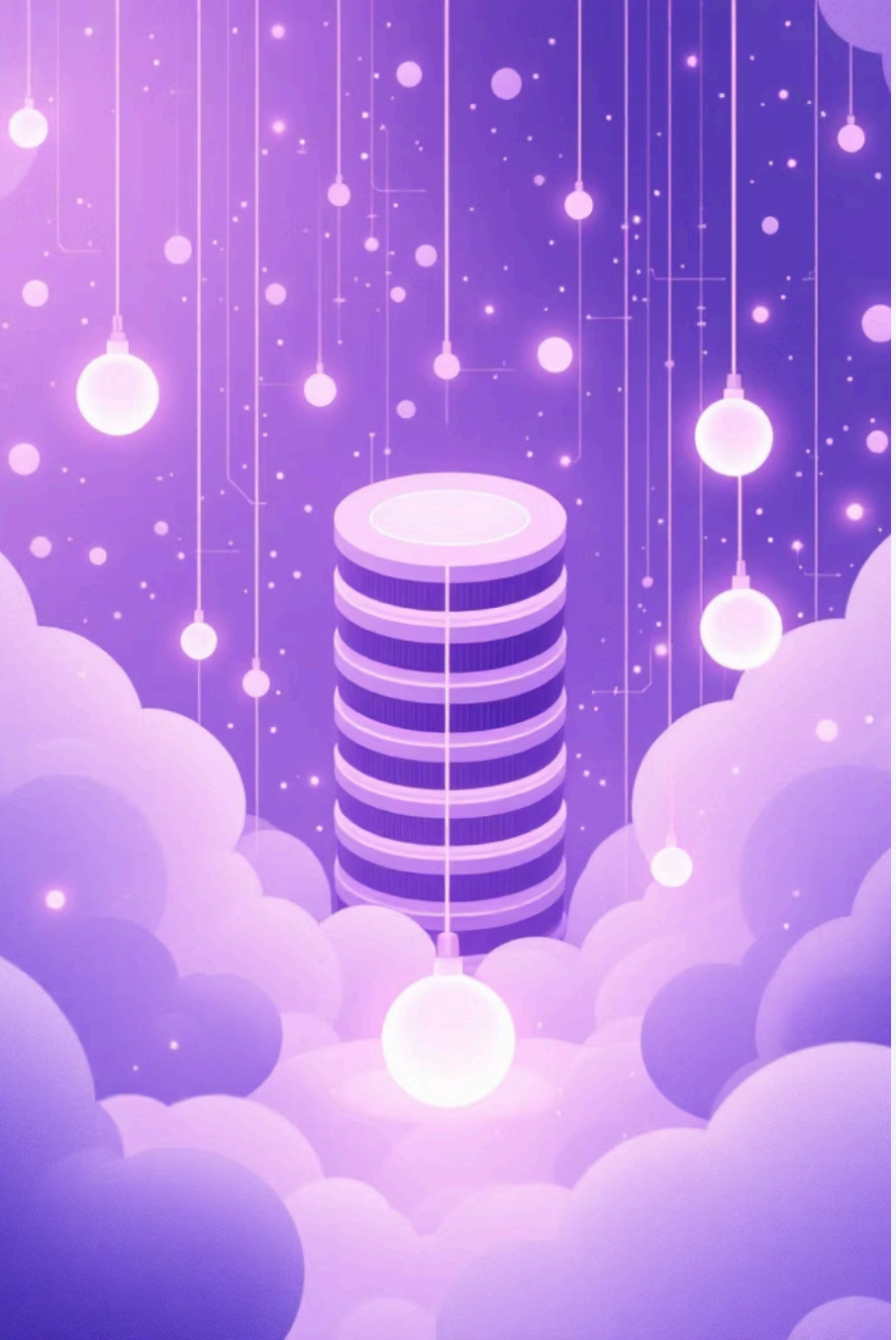
Assess potential loss from declining customer retention.

3

Pinpoint Product Categories

Identify categories driving churn-risk revenue.

Made with GAMMA



Dataset Summary

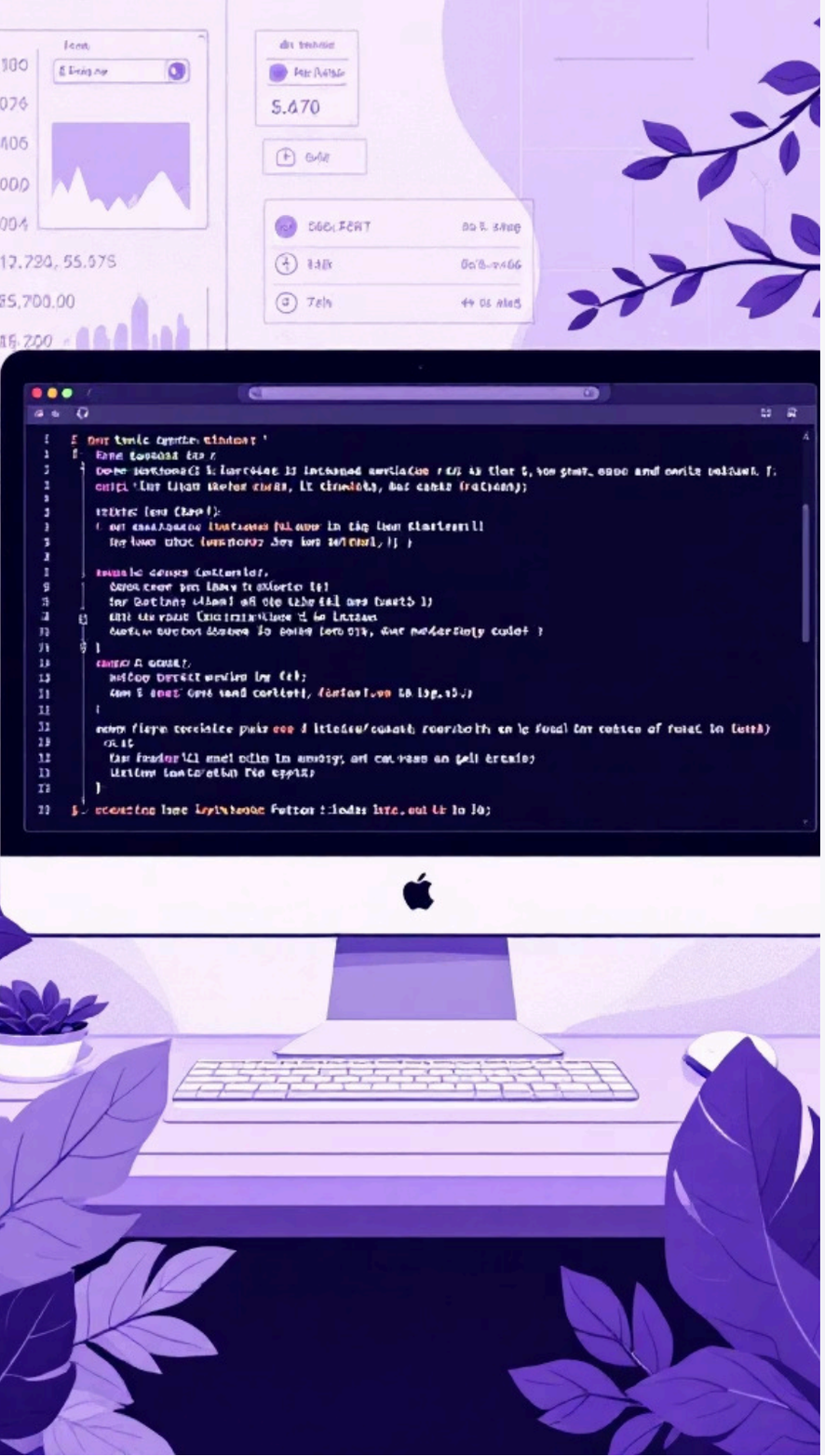
Our dataset comprises 3,900 rows and 18 columns, offering a comprehensive view of customer behavior.

- **Rows:** 3,900
- **Columns:** 18
- **Missing Data:** 37 values in Review Rating

Key Features

- Customer demographics (Age, Gender, Location, Subscription Status)
- Purchase details (Item, Category, Amount, Season, Size, Color)
- Shopping behavior (Discount, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type)

Made with GAMMA



Exploratory Data Analysis (Python)

We initiated our analysis with data preparation and cleaning using Python, leveraging pandas for efficient data handling.

Data Loading

Imported dataset using pandas.

Initial Exploration

Used `df.info()` and `.describe()` for structure and summary statistics.

Missing Data Handling

Imputed missing 'Review Rating' values using median per product category.

Column Standardization

Renamed columns to snake case for readability.

Made with GAMMA

Feature Engineering & Database Integration

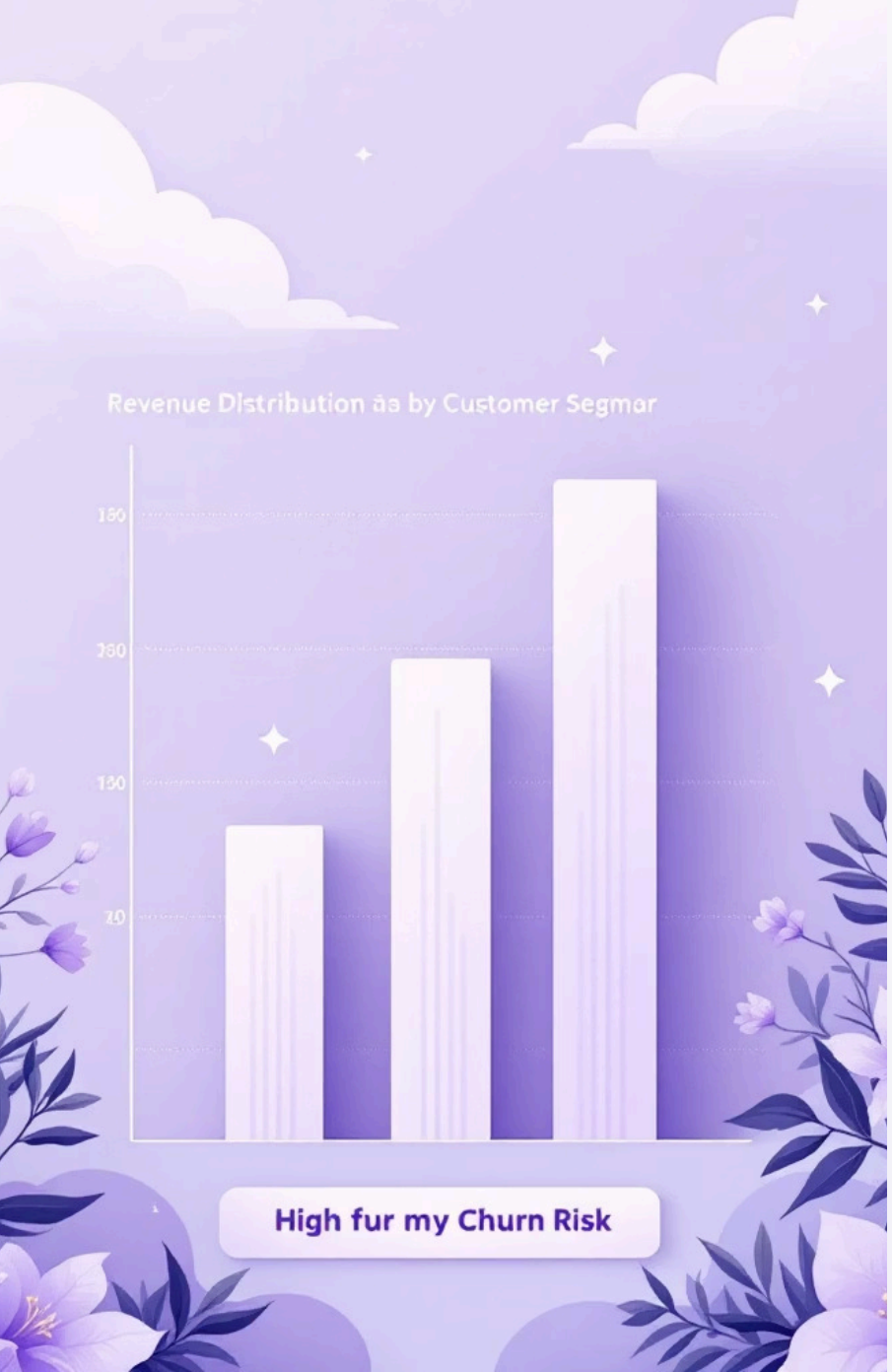
Key Engineered Features

- **Age Group:** Segmented customers into demographic cohorts.
- **Purchase Frequency:** Mapped for retention analysis.
- **Retention Score:** Measured customer engagement.
- **Churn Risk Groups:** Segmented into Highly Retained, Moderate Risk, High Churn Risk.
- **Revenue At Risk:** Flagged Moderate and High Churn Risk customers.
- **Customer Lifetime Value (CLV):** Estimated for high-value identification.

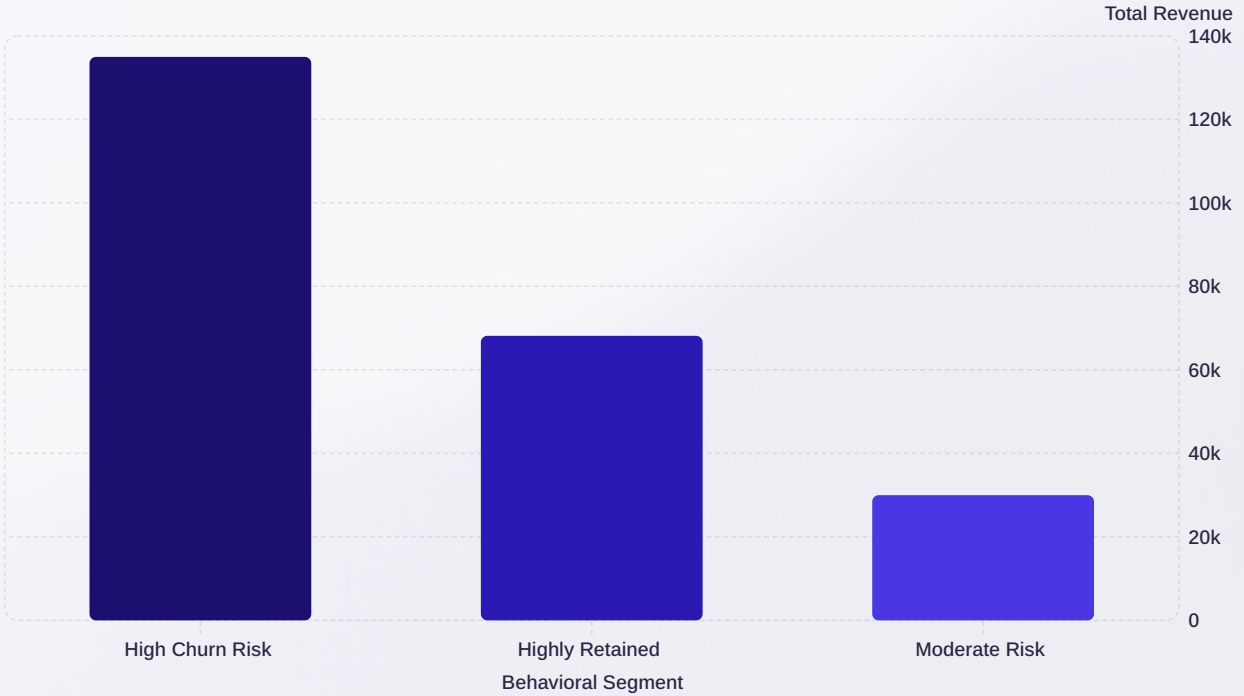
Database Integration

Connected Python script to PostgreSQL and loaded the cleaned DataFrame for SQL analysis.

Made with GAMMA



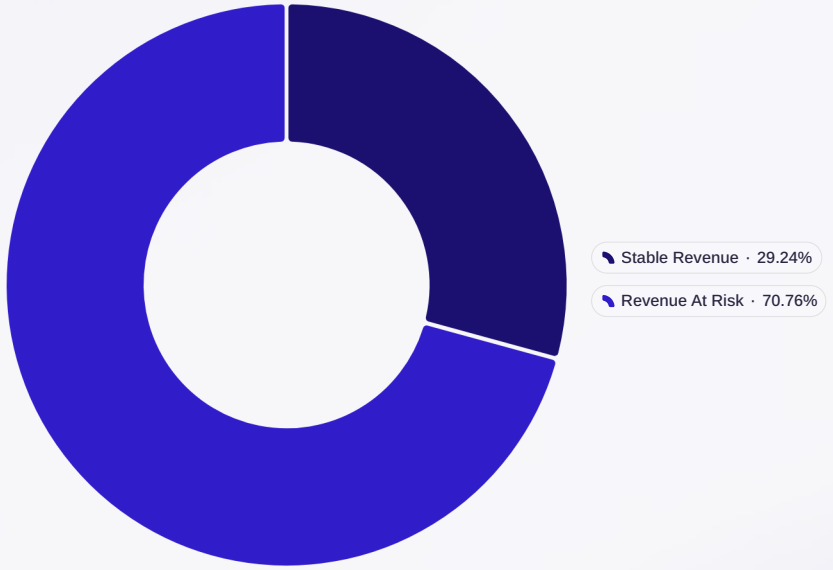
Revenue Distribution Across Segments



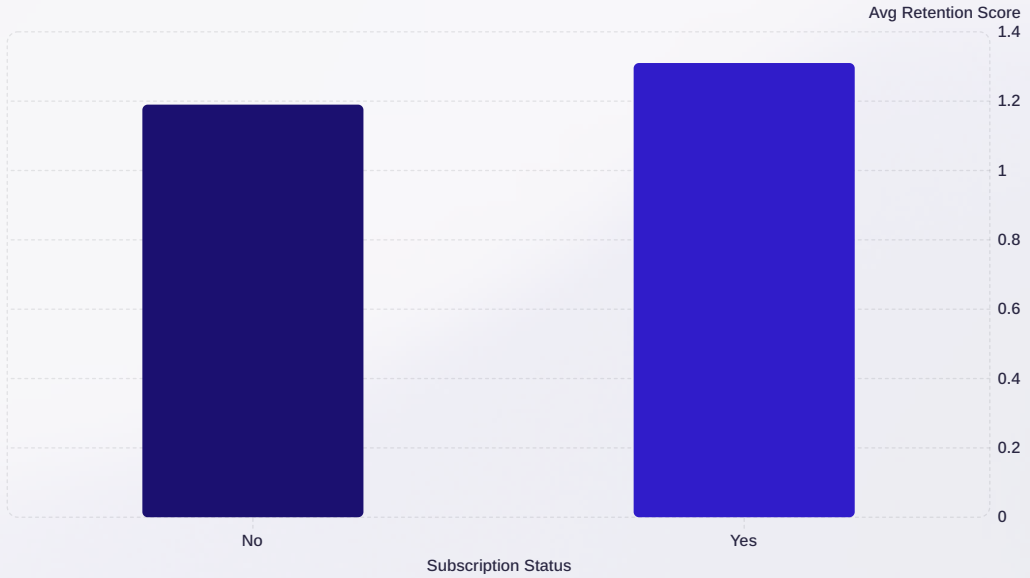
Analysis reveals a significant portion of revenue is tied to customers in the "High Churn Risk" segment, highlighting potential vulnerability. This distribution underscores the importance of targeted retention strategies.

Made with GAMMA

Revenue Vulnerability & Subscription Impact



A substantial 70.76% of our total revenue is currently "At Risk," originating from customers identified as Moderate or High Churn Risk. This indicates a critical need for proactive retention efforts.



Subscribed customers exhibit a higher average retention score (1.31) compared to non-subscribers (1.19), emphasizing the positive impact of subscription-based engagement on customer loyalty.

Made with GAMMA

Product Categories & Churn Risk

Category	Behavioral Segment	Total Revenue
Clothing	High Churn Risk	60723
Accessories	High Churn Risk	43439
Footwear	High Churn Risk	20204
Clothing	Moderate Risk	13539
Outerwear	High Churn Risk	10580

Clothing and Accessories are the top categories generating significant revenue from high churn-risk customers. This indicates these segments require immediate and focused retention interventions to mitigate potential losses.

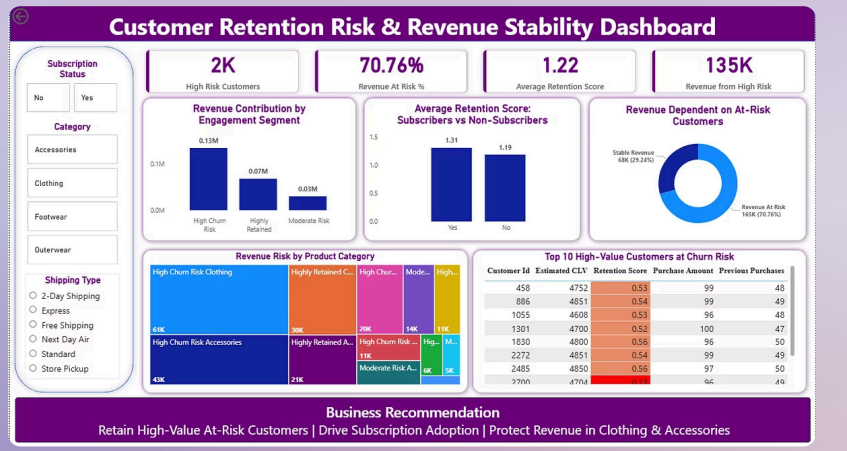
Made with GAMMA

Top 10 High-Value Customers at High Churn Risk

Customer ID	Purchase Amount	Previous Purchases	Estimated CLV	Retention Score
2272	99	49	4851	0.54
886	99	49	4851	0.54
2485	97	50	4850	0.56
2843	100	48	4800	0.53
1830	96	50	4800	0.56
458	99	48	4752	0.53
2700	96	49	4704	0.13
1301	100	47	4700	0.52
2806	97	48	4656	0.53
1055	96	48	4608	0.53

These high lifetime-value customers are currently exhibiting low retention behavior. Prioritizing targeted retention campaigns for these individuals is crucial to prevent significant revenue loss.

Made with GAMMA



Customer Retention Risk & Revenue Stability Dashboard

Our interactive Power BI dashboard provides a visual summary of customer retention risks and revenue stability. It highlights key metrics such as revenue contribution by segment, subscription impact, and product categories with high churn risk.

Made with GAMMA