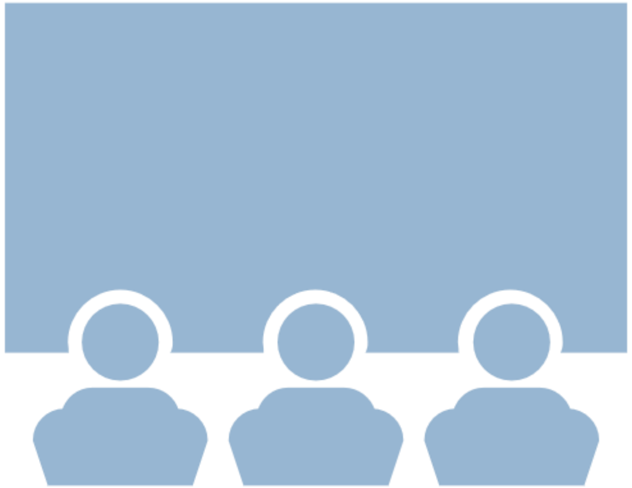
SaaS Customer Churn Analysis & Retention Experiment Design

Behavioral Analysis, Product Friction, and Risk-Based Intervention Design

Created by - Deeksha Gosain

OUTLINE



- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

EXECUTIVE SUMMARY

- This project was conducted to diagnose **rising customer churn** in a subscription-based SaaS product, with the objective of identifying **early churn drivers** and enabling **targeted retention interventions**.
- Customer churn increased sharply starting **early 2024**, signaling a **structural retention issue** rather than random attrition.
- Cohort analysis revealed a **steep retention drop within the first 1–3 months after signup**, leading to the definition of **early churn as churn within the first 90 days**.
- Behavioral analysis uncovered a **counterintuitive pattern**: **highly engaged trial users churned more than low-engagement users**.
- Further investigation showed churn was driven by **product friction (early-life errors)** rather than lack of activation or usage.
- A churn risk model identified - a **high-risk segment with ~70% early churn**, compared to **~21% early churn** among lower-risk users.
- Overall, results indicate that **early churn is primarily a product experience and reliability problem**, making **error reduction during early usage** the highest-impact retention lever.

INTRODUCTION

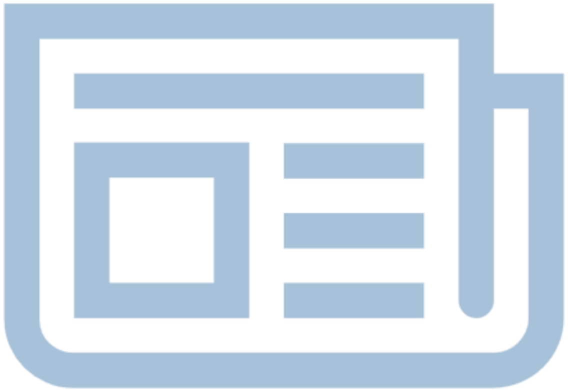
- This project analyzes **SaaS Subscription & Churn Analytics Dataset (Kaggle)**,a synthetic dataset representing a subscription-based software company with customer accounts, subscriptions, churn and product usage data.

- The project is guided by three core analytical questions:
  - **Is customer churn worsening over time, and is the increase structural rather than random?**
  - **When in the customer lifecycle does churn primarily occur, based on signup cohorts?**
  - **Which early behavioral signals (engagement and product friction) are associated with early churn?**

- The goal of the analysis is to move from high-level churn metrics to **behavior-driven insights** that explain *why* customers churn and *who* should be prioritized for retention efforts.

- The findings demonstrate how combining **cohort analysis, behavioral feature engineering, and churn risk modeling** can translate raw subscription and usage data into **actionable retention strategy and experiment design insights** for a SaaS business.

METHODOLOGY



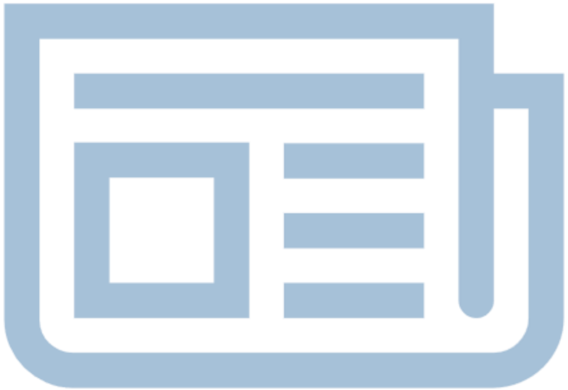
1. **Collection & Data Preparation (SQL & Python)**
   - Used the SaaS Subscription & Churn Analytics Dataset (Kaggle) and loaded it into SQLite database
   - Created SQL views to serve as a single source of truth for churn KPIs

2. **Exploratory Data Analysis**
   - Analyzed monthly churn trends and active account growth
   - Examined churn distribution by trial status, industry, and geography

3. **Cohort Analysis**
   - Built signup-month cohorts to analyze retention over time
   - Used retention heatmaps to identify early lifecycle drop-offs

4. **Behavioral Feature Engineering**
   - Engineered early-life behavioral metrics (first 30 days):
     - Engagement (active days, usage events, features used)
     - Friction (error counts, errors per active day)
   - Segmented users into activation-based behavioral cohorts

METHODOLOGY



5. Churn Risk Modeling

- Trained an interpretable logistic regression model using early behavioral features
- Generated churn risk scores to rank users by early churn likelihood
- Evaluated separation between high-risk and low-risk user segments

6. Experiment & Dashboard Design

- Designed a targeted A/B testing framework focused on high-risk trial users
- Defined success metrics and statistical evaluation approach
- Built a BI dashboard to monitor churn KPIs, cohort retention, behavioral insights, and risk segmentation

DASHBOARD TAB 1

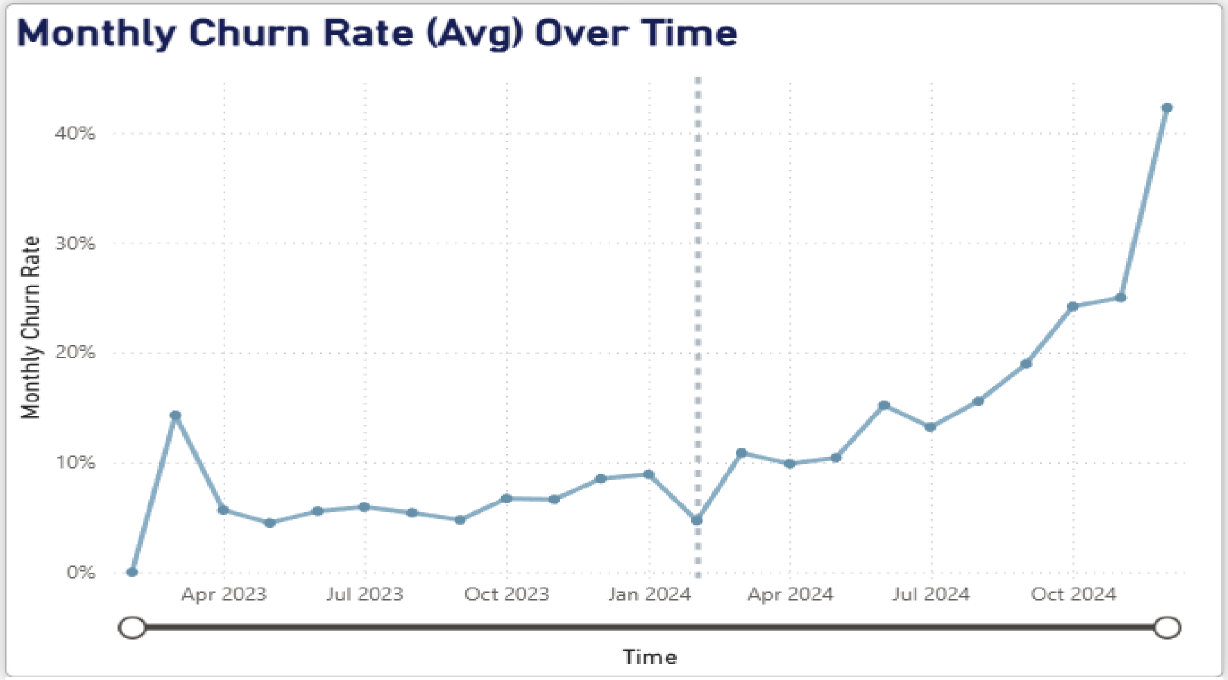
Churn Trend & Structural Shift

42.29%
Monthly Churn Rate (Max)

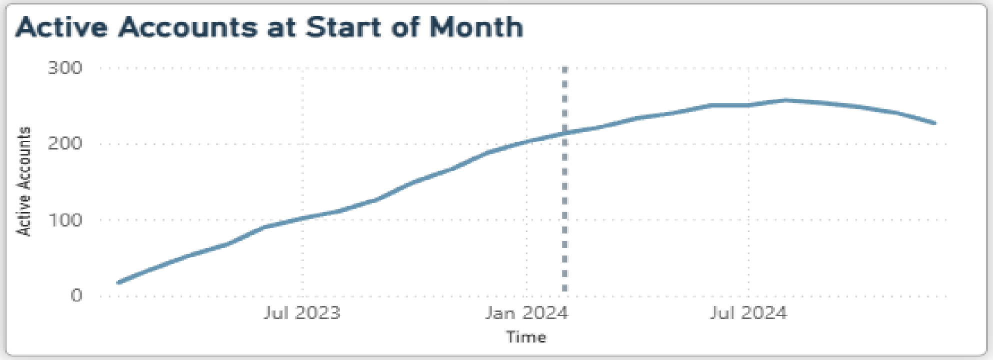
546
Total Churn Events

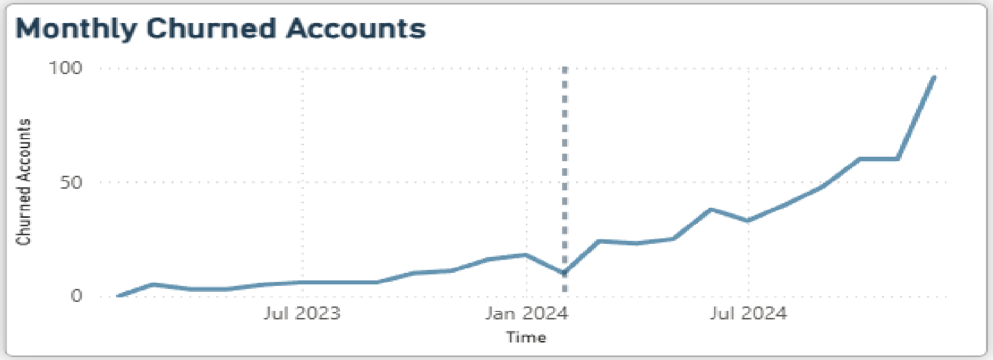
11.61%
Average Monthly Churn Rate

500
Distinct Accounts (Obser...



The sharp increase in monthly churn rate from early 2024 is attributed to decreasing growth of active accounts as well as increasing churned accounts.





DASHBOARD TAB 2

Cohort Retention Breakdown

37.60%
Early Churn Rate

81.80%
Month-1 Retention

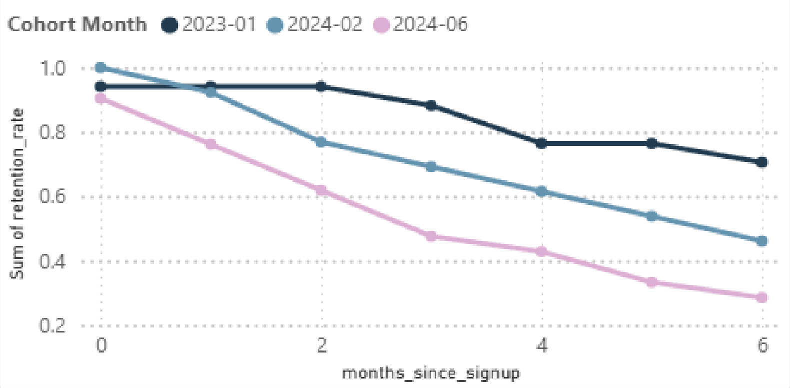
62.40%
Month-3 Retention

Retention Heatmap

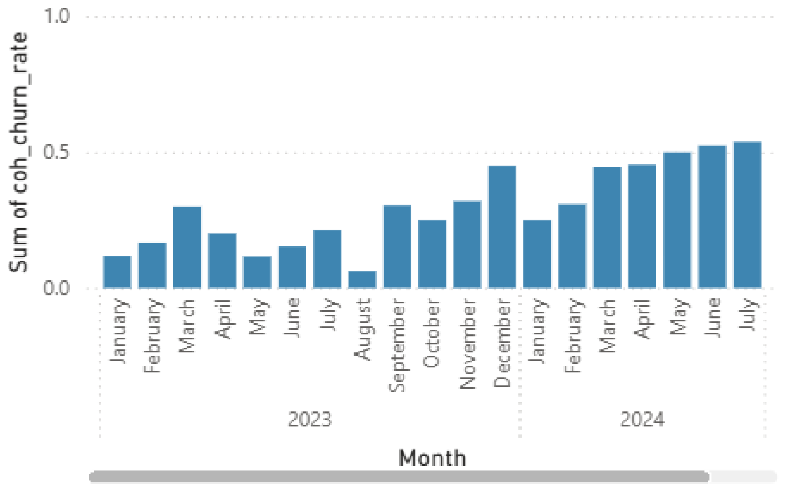
Year	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
⊞ **2023**	**11.50**	**10.73**	**10.40**	**9.35**	**8.77**	**8.37**	**7.76**	**7.29**	**6.89**	**6.36**	**5.89**	**5.36**	**5.05**	**4.53**	**4.08**	**3.50**	**3.14**
⊟ **2024**																	
January	0.94	0.88	0.81	0.75	0.56	0.44	0.38	0.31	0.31	0.25	0.25	0.25					
February	1.00	0.92	0.77	0.69	0.62	0.54	0.46	0.46	0.38	0.38	0.38						
March	0.89	0.78	0.63	0.56	0.52	0.41	0.41	0.33	0.33	0.30							
April	0.95	0.73	0.68	0.55	0.41	0.32	0.32	0.27	0.23								
May	0.95	0.68	0.50	0.50	0.45	0.41	0.32	0.32									
June	0.90	0.76	0.62	0.48	0.43	0.33	0.29										
July	0.96	0.81	0.65	0.46	0.38	0.38											
August	0.86	0.67	0.57	0.48	0.43												
September	0.80	0.36	0.28	0.16													
October	0.97	0.61	0.39														
November	0.75	0.25															
December	0.41																

Churn increase is driven by worsening early-life retention in newer acquisition cohorts, not by a change in customer mix.

Retention Curves (Pre vs Post 2024)



Early Churn Rate By Signup Month



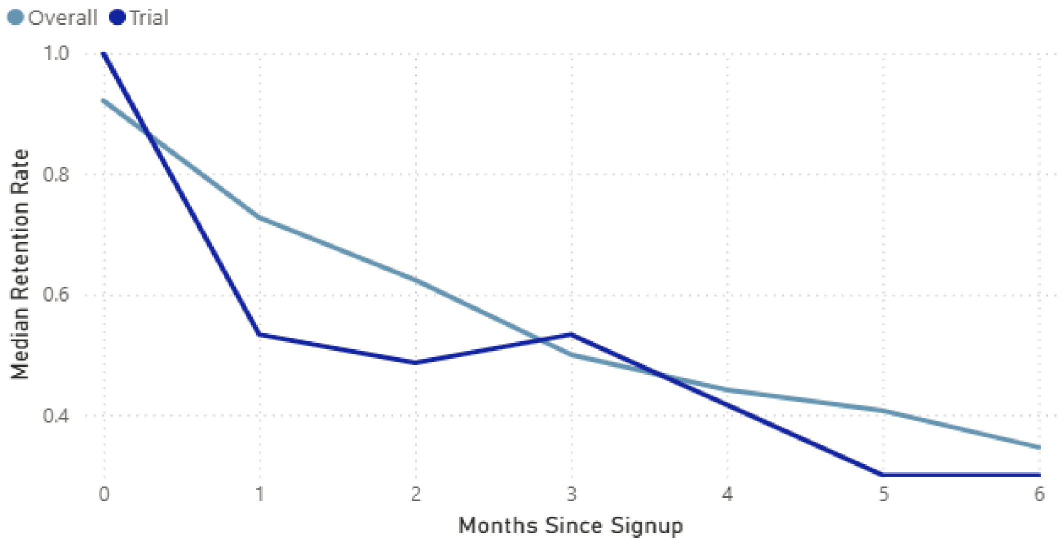
DASHBOARD TAB 3

Diagnostic Population Rationale

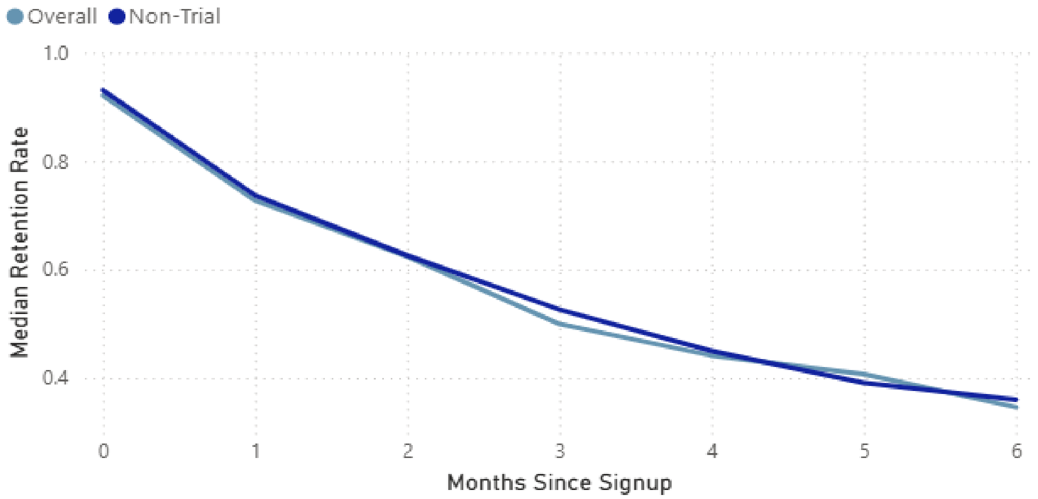
33%
Median Cohort Early Churn Rate – Trial

29%
Median Cohort Early Churn Rate -Non Trial

Median Cohort Retention — Trial vs Overall (2024 Cohorts)
Typical cohort behavior, first 6 months since signup



Median Cohort Retention — Non-Trial vs Overall (2024 Cohorts)
Typical cohort behavior, first 6 months since signup



While both trial and non-trial users exhibit early churn in post-2024 cohorts, trial users show faster retention deterioration in the first 1–3 months. Trial users are therefore used as a diagnostic population because they surface product issues earlier and are less confounded by renewal timing, brand loyalty, or refund friction.

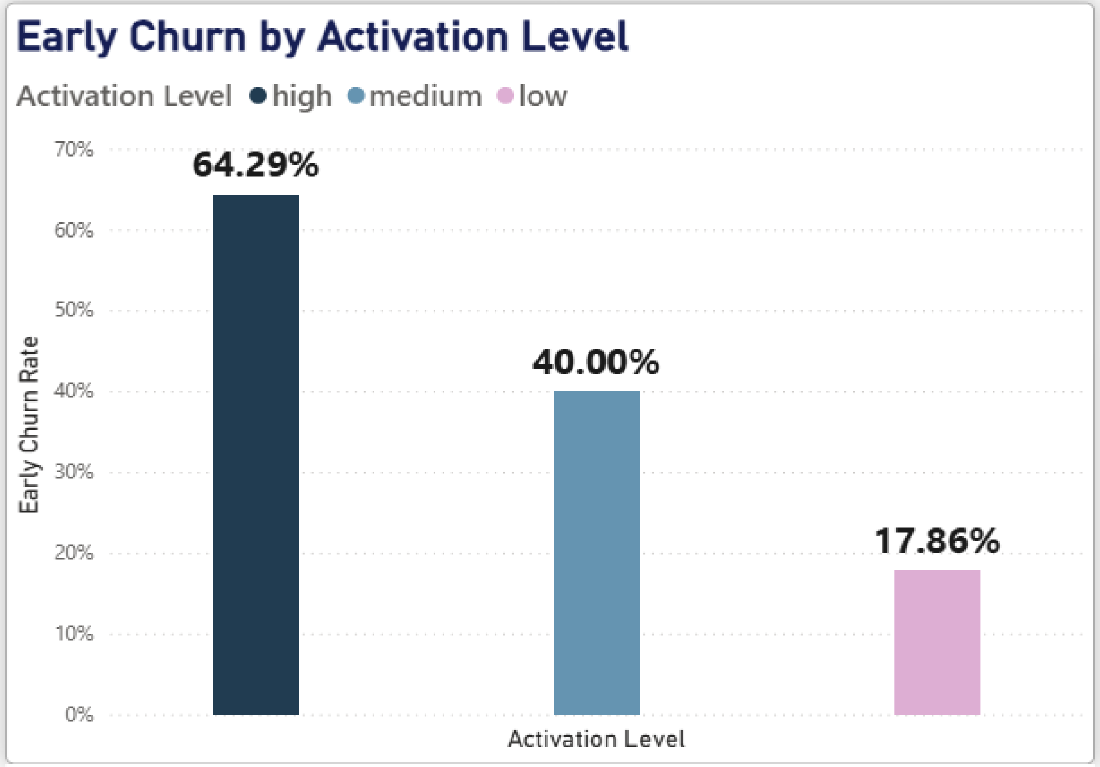
DASHBOARD TAB 4

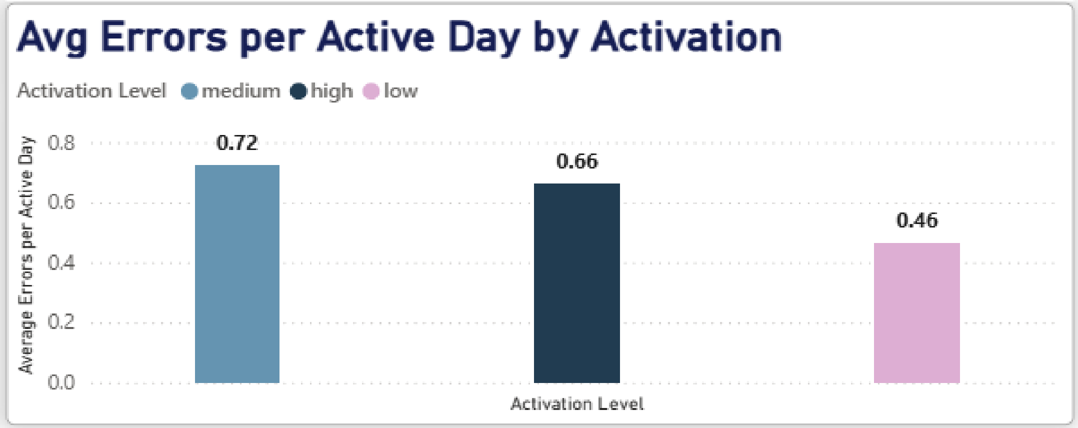
Product Friction: Usage, Errors & Features

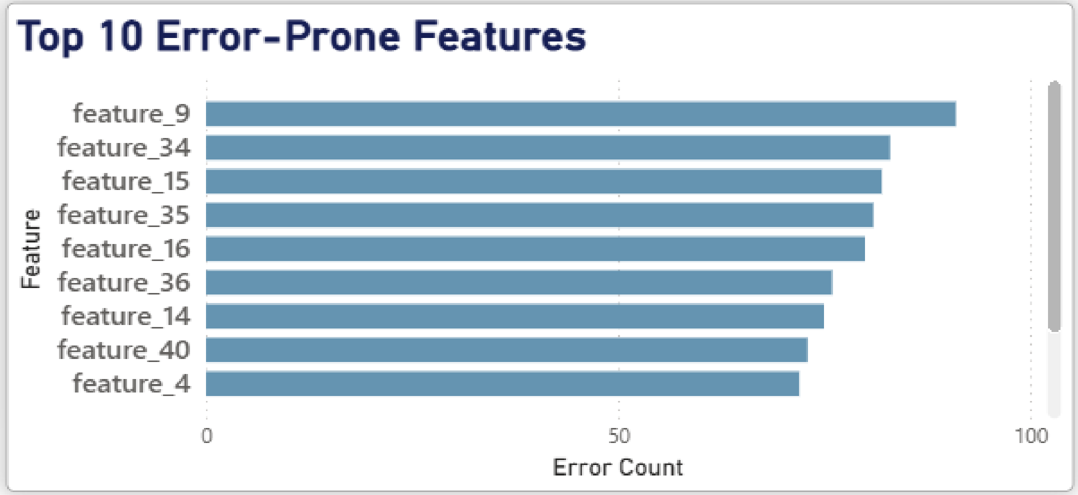
0.60
Mean Errors per Active Day

58.2%
% Users with ≥1 Error (First 30 Days)

Note: This dashboard focuses on trial users and is intended for diagnostic analysis only. Insights are directional and used to identify potential churn drivers.







This dashboard shows that higher early engagement leads to greater error exposure, and those errors — concentrated in specific features — are the primary driver of early churn.

DASHBOARD TAB 5

Risk-Based Action

70.00%
Early Churn Rate (High Risk Users)

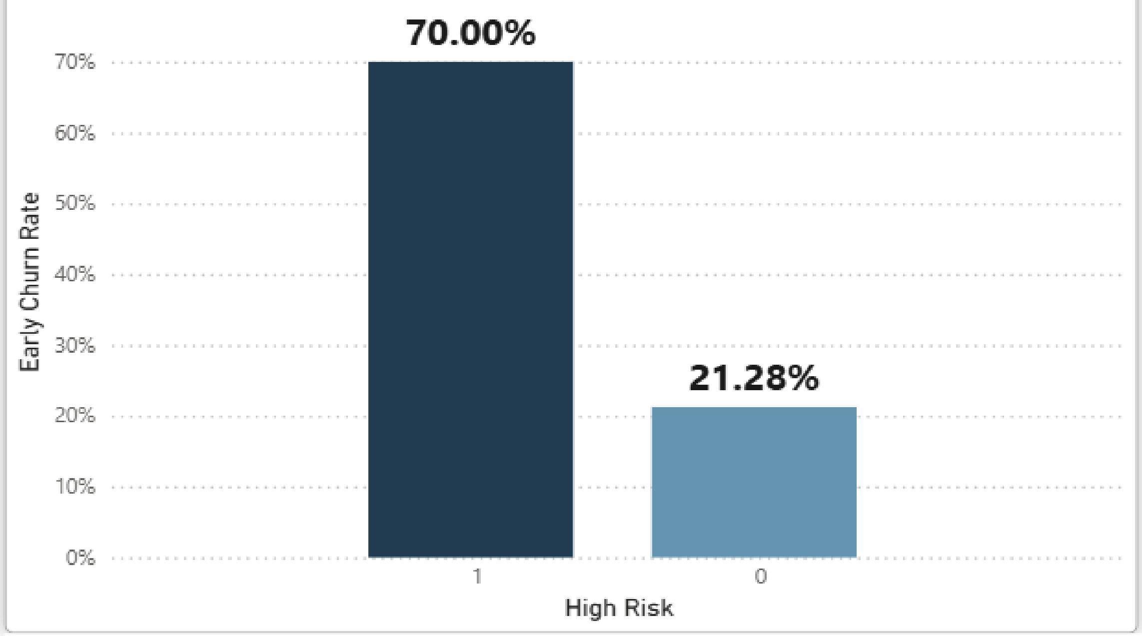
5
Median Active Days (30d) - High Risk

0.46
Median Errors / Active Day (30d) — High R...

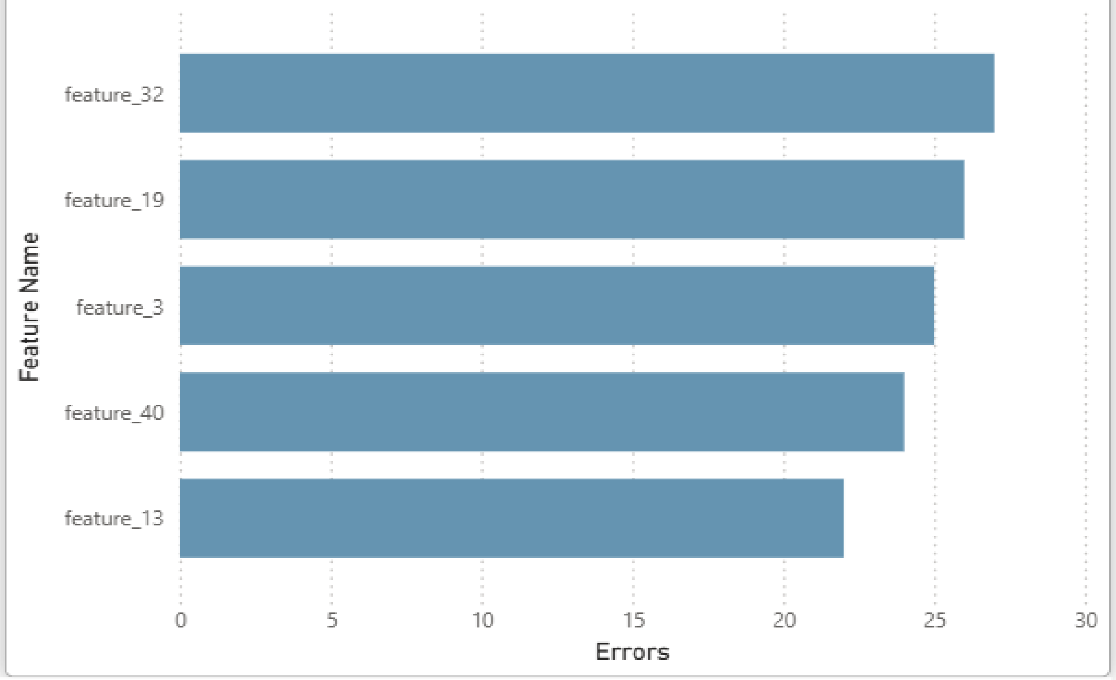
Unit- seconds
19K
Median Usage Time (30d) - High Risk

High Risk vs Low Risk Early Churn

1 = High Risk Users ; 0 = Low Risk Users



Top Error-Driving Features Among High-Risk Users



Focus A/B testing on high-risk trial users and prioritize fixes in features with the highest error exposure. Additionally, target high-risk trial users with proactive product guidance and error-specific nudges during days 1–14. Measure success via reductions in early churn and error intensity.

OVERALL FINDINGS & IMPLICATIONS

Findings

- **Customer churn increased structurally over time**, with a sharp rise observed from early 2024 rather than random month-to-month variation.
- **Early churn is concentrated within the first 90 days after signup**, indicating that customer outcomes are largely decided early in the lifecycle.
- **Product friction (errors) increases with engagement intensity**, exposing active users to more failure points.
- **Behavioral cohorts based on early activation show clear churn separation**, outperforming static attributes such as plan tier.
- **Churn risk modeling identifies a high-risk segment where ~70% of users churn early,** compared to ~21% among lower-risk users.

Implications

- **Retention efforts should prioritize early lifecycle intervention like improving onboarding quality and early lifecycle intervention**, rather than focusing on long-tenure customers.
- **Reducing early product friction is a key retention lever**, especially for highly engaged users.
- **Product experience improvements are likely to yield higher retention impact than market-level segmentation alone**
- **Risk-based prioritization allows retention teams to focus on users with the highest churn likelihood**, maximizing ROI.

CONCLUSION

- The analysis demonstrates that customer churn in this SaaS environment is driven primarily by **early lifecycle behavior**, with a significant portion of churn occurring within the first 90 days after signup rather than later in the customer journey.
- Cohort and behavioral analysis reveal that churn is **not caused by lack of engagement**, but by **friction experienced during early product usage**, particularly among highly engaged trial users who encounter elevated error rates.
- Risk-based segmentation shows that early churn is **highly concentrated within a small subset of power users**, enabling targeted retention efforts to be significantly more effective than broad, population-wide interventions.
- Overall, the findings highlight the value of **combining cohort analysis, behavioral feature engineering, and churn risk modeling** to move from descriptive churn metrics to **actionable retention strategy and experiment-driven decision-making** in a SaaS context.