



Instagram: Improvement Of Quality Engagement & Retention

Nivethon Jayaraman

Table Of Contents:

Introduction & Objective : 3

Methodology : 5-8

Tableau Primary Dashboard : 9

Key Metrics Overview: 10

Engagement Quality vs Total Activity: 11-12

Retention & Dormancy Risk: 13

Monetization Vs User Fatigue: 14-15

Geographic Fatigue Hotspots: 16

Community Health Analysis: 17

Lifestyle Correlation Heatmap: 18

Age Group Vs Churn Risk: 19-20

Secondary Dashboard- Efficient Engager Paradox – Deep Dive: 21

Actionable Steps & Recommendations: 24- 29

Conclusion: 30

Introduction

The Objective:

The goal of this project was to perform a comprehensive "Ecosystem Health Check" on Instagram's user engagement and retention data. We moved beyond tracking simple surface-level metrics like "Time Spent" to calculate **Engagement Quality Efficiency** (EQE)—identifying exactly how much meaningful value users derive from the platform versus the fatigue caused by excessive monetization and passive consumption.

The Methodology:

- To ensure a rigorous behavioral analysis, we utilized a two-step technical approach:
- **Python** (Data Engineering): We engineered complex calculated fields to segment users based on behavioral archetypes and calculated predictive **Churn Risk Scores** by merging interaction density with longitudinal usage patterns.
- **Tableau** (Business Intelligence): We transformed these advanced metrics into an interactive diagnostic suite to isolate "**Fatigue Hotspots**" and visualize the correlation between **ad-load pressure** and **user dormancy**.

The Goal of this Report:

- This report identifies **where** engagement quality is eroding, **why** our highest-value users are currently at the highest risk of churning, and provides **data-driven interventions** to rebalance monetization with long-term platform sustainability.

01

Workflow: STEP 1



Data Engineering And Behavioral Modelling

We utilized **Python** to engineer a high-fidelity dataset that uncovers the hidden relationship between user activity and platform fatigue.

Custom Behavioral Archetyping:

We engineered a multi-variable logic to segment the user base into four distinct behavioral clusters: Power Users, Efficient Engagers, Passive Scrollers, and At-Risk users.

Predictive Churn Risk Scoring: We developed a churn propensity model (**0.0 to 1.0**) that revealed a critical "Quality Flight" trend: Efficient Engagers currently exhibit the highest risk of dormancy at approximately **0.88**, a risk factor invisible to traditional engagement tracking.

EQE (Engagement Quality Efficiency) Metric: Using Python, we defined and calculated **EQE** to quantify the value density of user sessions.



Data Analysis With Python

Ad Load Proxy: A calculated estimate of advertisement density per user session, used to identify segments experiencing over-monetization.

- **Fatigue Analysis via Ad-Load Proxy:**

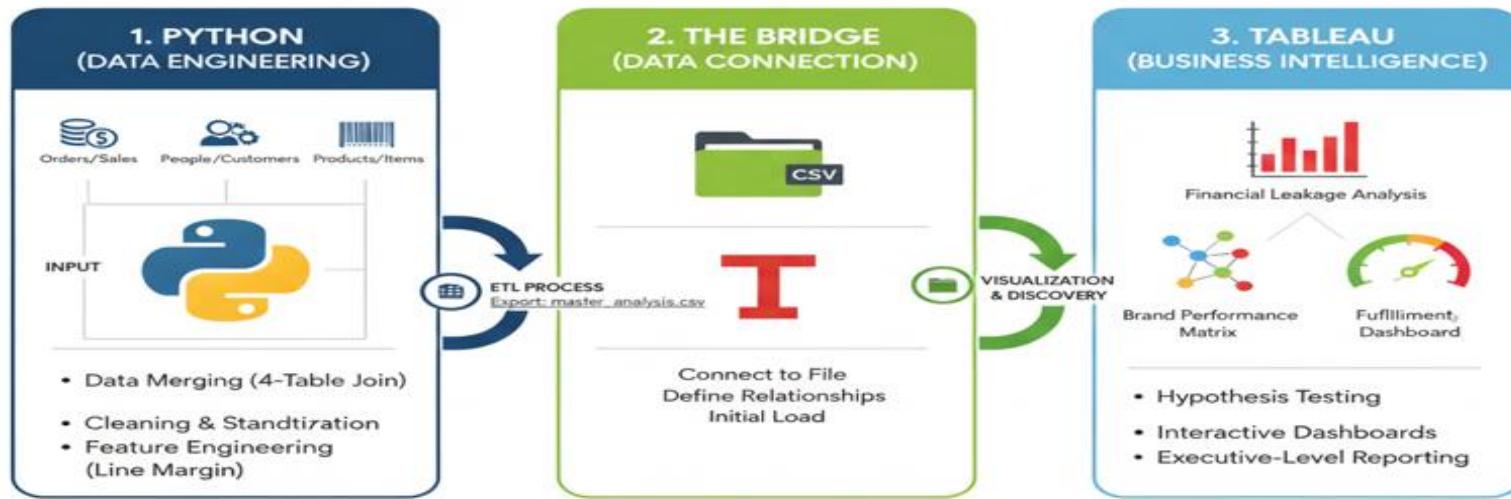
We calculated a normalized ad_load_proxy to measure monetization pressure. Our findings indicate that Efficient Engagers are currently subjected to nearly 3x the monetization load compared to Passive Scrollers, identifying the primary catalyst for high-value user burnout.

- **Forming the Correlation Matrix:**

It is used to find the relationship between lifestyle factors to see if it affects churn .

Analytical Flowchart

DATA ANALYTICS PIPELINE: FROM RAW DATA TO BUSINESS INTELLIGENCE



Turning Data into Business Intelligence (Tableau)

The Core Purpose: This analysis focuses on diagnosing the hidden drivers of user churn by auditing the delicate balance between high-value engagement and monetization-induced fatigue.

Key Objectives of the Analysis:

Segmented Risk Discovery: Identifying "invisible" churn patterns among Efficient Engagers that raw activity numbers fail to capture.

Behavioral Root Cause Analysis: Using interactive drill-downs to pinpoint exactly why our most intentional users—not just passive scrollers—are reaching a breaking point.

Ecosystem Sustainability: Mapping the relationship between Ad Load Proxy and Engagement Quality Efficiency EQE to protect the platform's long-term social health.

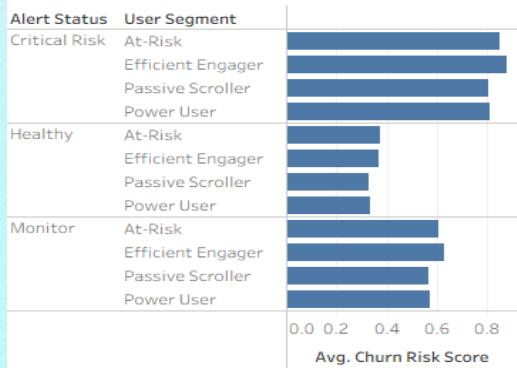
Workflow STEP 2: Tableau Dashboard:

Executive KPI

PLATFORM EQE: 1.707
CHURN RISK: 0.6466
AD FATIGUE: 0.1515

Highest **average churn risk** is in Efficient Engagers

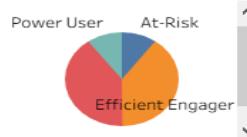
Retention & dormancy risk (2)



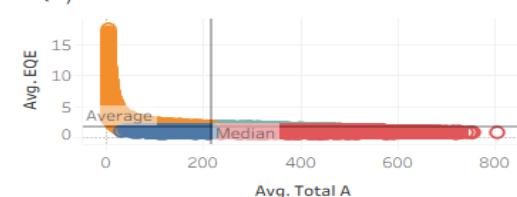
Click the User Segment dropdown on the right :

- Try these to see key insights:
- All → Full platform overview
 - Efficient Engager → Highest average churn risk
 - Passive Scroller → Highest dormancy & passive consumption
 - At-Risk / Power User → Compare fatigue & eng..

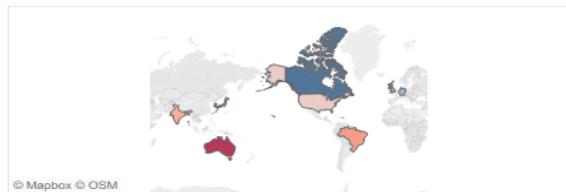
Explore By Segment



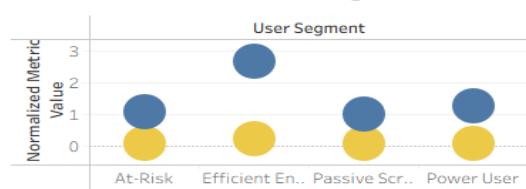
EQE/Total A



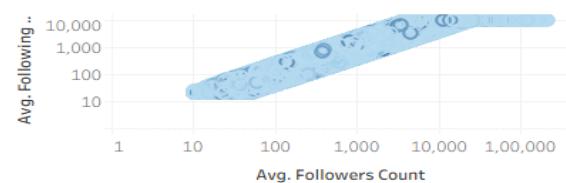
Geographic EQE hotspots



Monetization V User Fatigue



Community Health



Key Metrics Used:

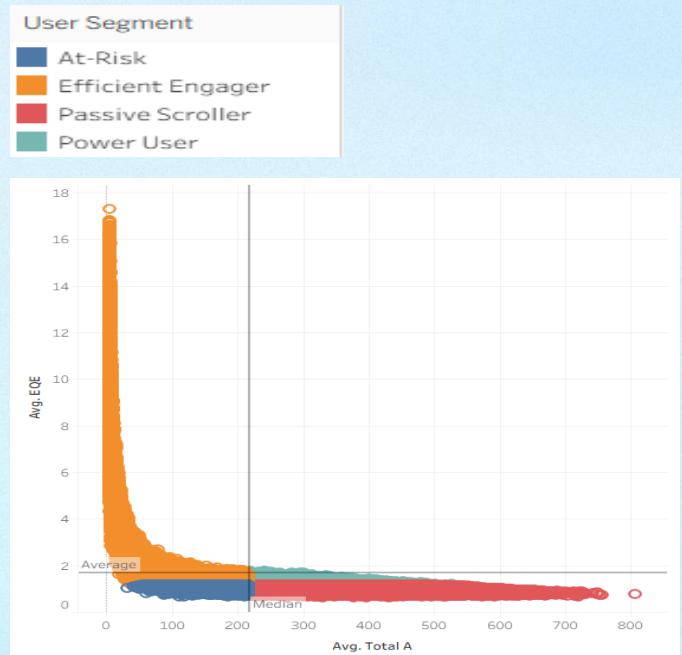
Metric/Field Name	Target Segment	Purpose/Meaning
Engagement Quality Efficiency(EQE)	Passive Scrollers	primary "User Health" metric
AD load Proxy	At Risk Users	Tracks frequency of monetization relative to segment averages
Churn Risk Score	Critical Risk Users (scores>0.7)	Predicts the likelihood of an user becoming dormant
Community Health Ratio	Lopsided Consumers	Follower to Following relationship ratio
User Segment	Category of customers based on EQE levels	Separate and analyze each area like passive scroller, efficient engager

Engagement Quality vs. Total Activity:

Hypothesis:

We hypothesized that while increased activity (Total A) would initially lead to higher engagement, there would be a "point of diminishing returns" where excessive activity shifts from meaningful interaction to passive consumption, lowering the overall quality (EQE) of the session.

Observation: Chart displays a clear L curve , high quality engagement located at lower to moderate levels and passive scrollers are located at the far end of the X axis

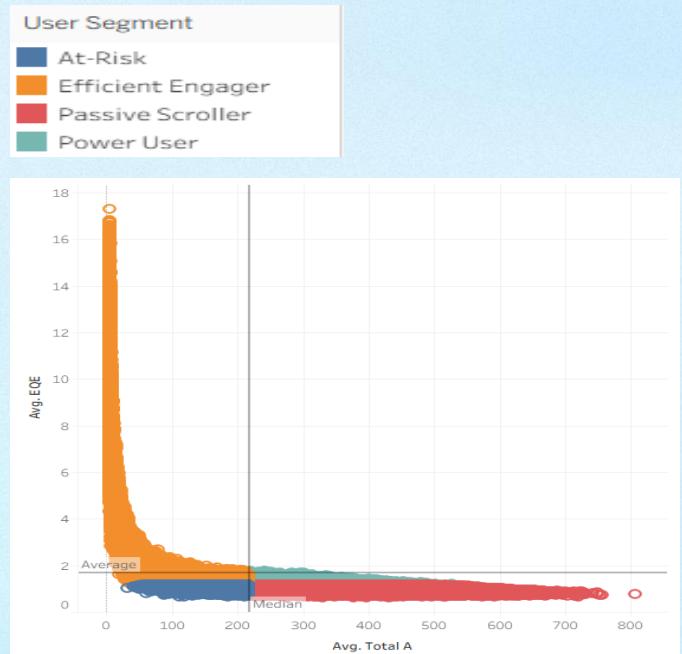


Engagement Quality vs. Total Activity:

Findings:

"Time Spent" is a deceptive metric — high activity volume does not guarantee high engagement quality.

The L-curve shows : **High EQE** at low-to-moderate activity (Efficient Engagers achieve strong results with less effort) Sharp drop at high activity (Passive Scrollers dominate far end with heavy passive consumption and low meaningful interaction).



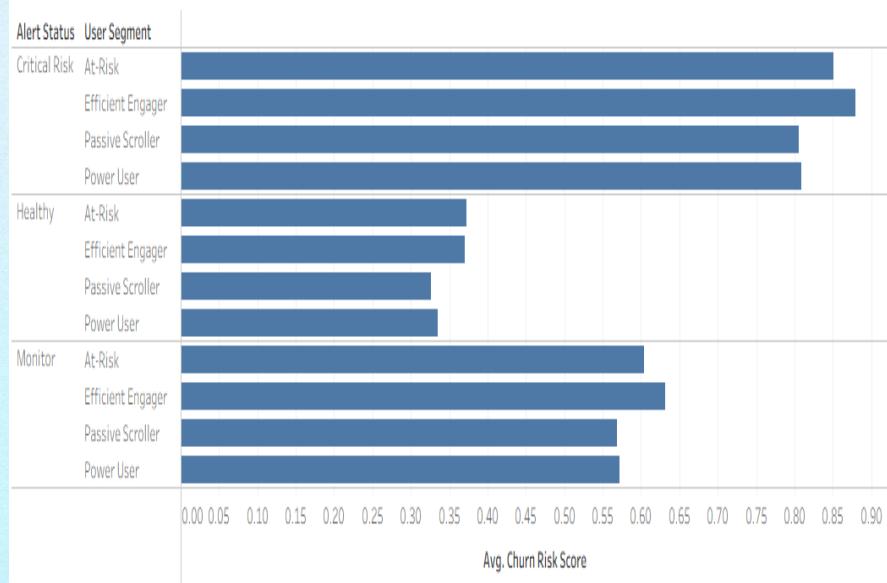
Retention & Dormancy Risk

Hypothesis:

Users with high **EQE** (Efficient Engagers) will be the most loyal and have the lowest churn risk.

Observation: In the "Critical Risk" alert status, **Efficient Engagers** have the highest average churn risk score at approximately **0.88**, surpassing even **At-Risk** users (~0.85) and **Passive Scrollers** (~0.81).

Findings: Hypothesis disproved. High engagement quality does not guarantee retention, in fact quality users are more likely to churn once flagged as at risk , as their use is more intentional than habitual



Monetization Vs User Fatigue

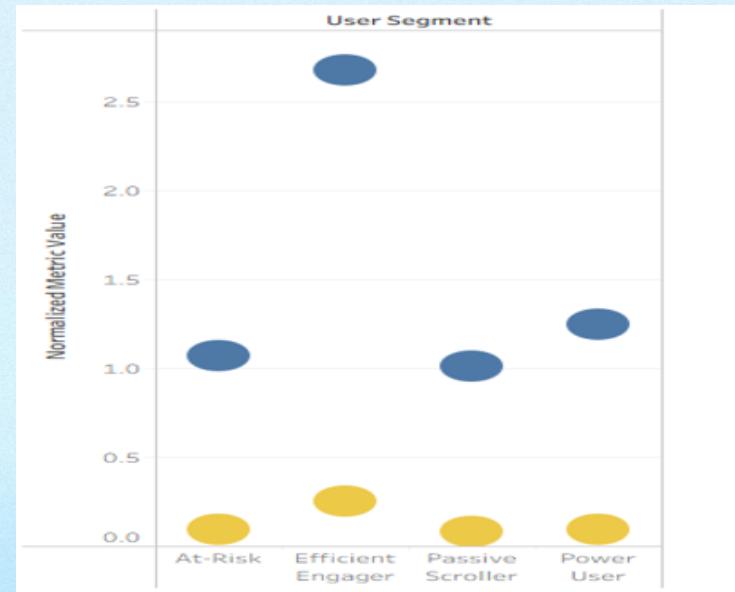
Hypothesis: High value segments- like Efficient engagers are subject to high monetization pressure(ad load proxy) relative to their engagement quality, leading to higher burnout.

Observation: Across all segments, Normalized ad load(blue) consistently exceeds EQE(yellow), Efficient engager shows massive parity with an ad load value of 2.7 while the EQE remains **below 0.3**

Findings:

Proved — The platform is aggressively monetizing its most efficient users. This segment is being served nearly 3X the ad load of At Risk or Passive Users which directly correlates to them having the highest churn risk rate which was also observed on the previous chart.

Metrics Used (In Rows): Avg ad_load_proxy , Avg EQE



Monetization Vs User Fatigue

Segment	Ad Load Level	EQE Level	Strategic Outcome
Efficient Engager	Extreme(2.7)	Low- Moderate	High Burnout Risk-must reduce ad load to save high value segment
Passive Scroller	Baseline(1.0)	Very Low	Stable Habitual Use- high volume, low efficiency
At Risk	Moderate(1.1)	Very Low	Pre Churn-low interaction value making them sensitive to any ad load

Geographic EQE Hotspots

Hypothesis:

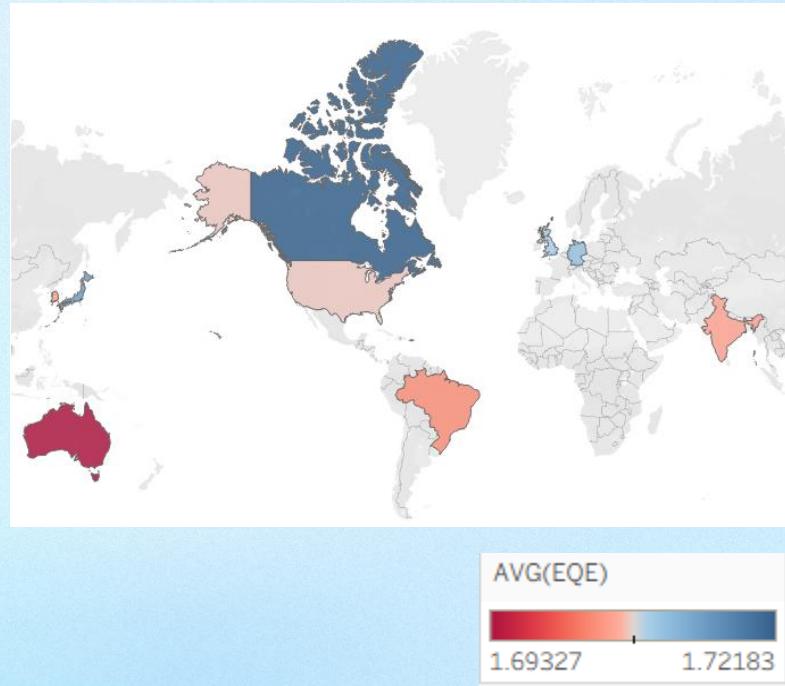
Instagram EQE is globally uniform, with no significant variation in user value density across different geographic regions.

Observation:

The map reveals clear geographic disparities. Regions like Canada and the United Kingdom show high engagement quality (darker blue), while countries such as Australia, India, and Brazil exhibit significantly lower EQE (shades of red/pink).

Findings:

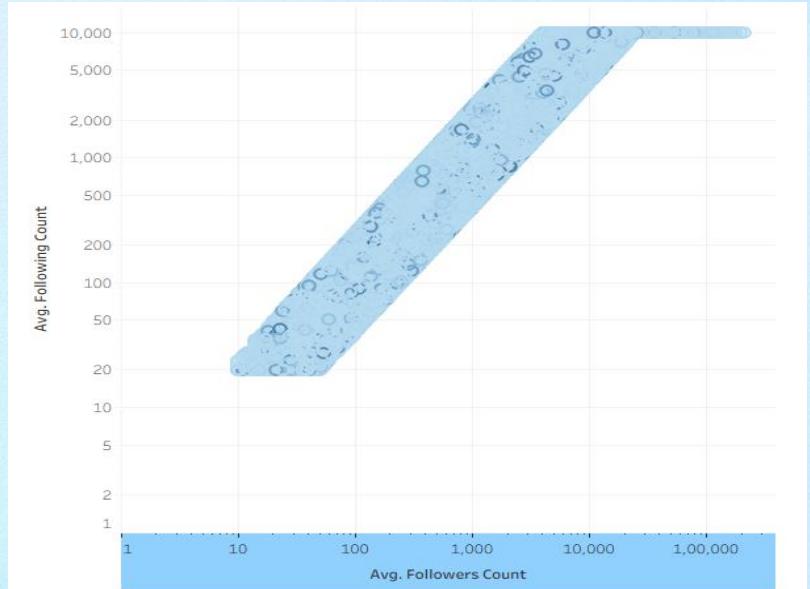
Hypothesis disproved, EQE is highly regional. The low EQE in regions like India & Australia suggests they are Fatigue hotspots where users are engaging in passive consumption.



Community Health

- **Hypothesis:** Healthy community engagement is characterized by a "Circle of Friends" model, where users maintain a relatively balanced **1:1 ratio** between followers and following counts.
- **Observation:** When viewed on a **Logarithmic Scale**, a significant portion of the population deviates from the 1:1 diagonal line. A dense cluster of users occupies the "High Following / Low Follower" zone, representing a **broadcast-consumption** model rather than a reciprocal social community.

Findings: Hypothesis Disproved. The prevailing platform health is currently "lopsided". Most users act as passive consumers of content rather than active nodes in a mutual social network, which correlates with the high volume of **Passive Scrollers** identified in the behavioral segmentation.



Lifestyle Correlation Heatmap

- The chart shows very weak correlations overall, with most off-diagonal values near 0.00 (no meaningful linear relationship).
- The only notable signal is a moderate negative correlation between Stress and Churn (-0.30): higher perceived stress is associated with lower churn risk.
- All other lifestyle factors (Exercise, Sleep, Diet, BMI) have essentially zero correlation with churn risk or each other.

Lifestyle Heatmap

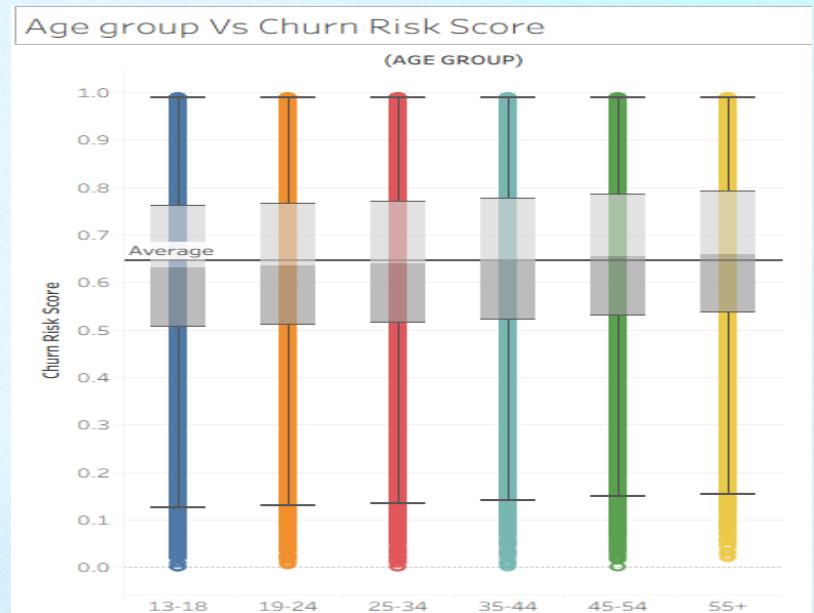
	Var Col					
Var Row	BMI	Churn	Diet	Exercise	Sleep	Stress
Exercise	0.00	0.00	0.00	1.00	0.00	0.00
Sleep	0.00	0.00	0.00	0.00	1.00	0.00
Diet	0.00	0.00	1.00	0.00	0.00	0.00
Stress	0.00	-0.30	0.00	0.00	0.00	1.00
BMI	1.00	0.00	0.00	0.00	0.00	0.00
Churn	0.00	1.00	0.00	0.00	0.00	-0.30

Age Group Vs Churn Risk Score[OUTLIER]

The Hypothesis:

Younger age demographics (13-18 and 19-24) will exhibit significantly higher **Churn Risk Scores** due to more volatile platform usage patterns compared to older, more stable demographics.

Observation: Across all age groups from 13-18 to 55+, the distribution of churn risk remains remarkably consistent, with average scores hovering around the 0.65 threshold. Every demographic shows a wide range of scores, from nearly 0.0 to maximums near 1.0, as indicated by the uniform box plots and whisker lengths.



Age Group Vs Churn Risk Score

Findings: Hypothesis Disproved. Age is not a primary driver of retention risk on the platform. The consistent "Average" line across the X-axis suggests that churn is influenced by behavioral factors (such as **EQE** and **Ad Fatigue**) rather than biological age.

Tight medians across ages: Churn risk medians stay very close (~0.5–0.6) with no clear upward or downward trend by age group.

Outliers are widespread: High and low extremes appear in nearly every age band, suggesting they're driven by individual behavior, not age.

Low predictive power: Age shows no strong or actionable pattern compared to user segment or behavioral metrics like creation and EQE.

Instagram Retention Risk- Efficient Engager Paradox:

KEY INSIGHT

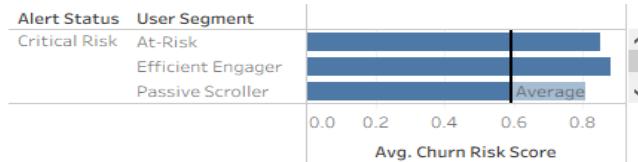
Efficient Engagers get the **highest EQE**

with moderate time, low ads, low creation

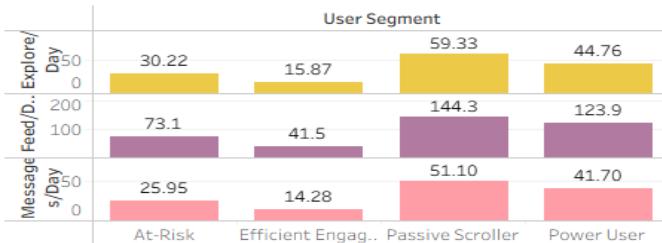
→ but still have the **highest churn risk**

→ Possible reason: Not enough social/growth reward despite efficiency

Churn Risk By User Segment



Average Time Allocation by Segment

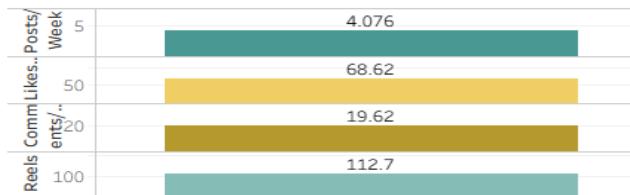


Efficient Engagers = users with highest EQE. Focus: Protect high-value users from subtle dissatisfaction.

Efficient Engagers – Key Metrics



Activity & Creation Levels by Segment



Instagram Retention Risk- Efficient Engager Paradox:

- **Efficient Engagers** show the highest average churn risk despite achieving the **highest EQE** with moderate effort.
- They have low creation (posts/week ≈4, comments/day ≈20) and moderate consumption (Reels time ≈24 min/day, sessions ≈4.8/day) Ad exposure (≈12/day) and notification response (≈50%) are not extreme — rules out classic burnout or overload compared to Passive Scrollers (high passive Reels time, very low creation) and Power Users (high everything), Efficient Engagers appear efficient but unrewarded
- Efficient Engagers are leaving because the platform delivers efficiency without enough social validation, growth, or emotional stickiness. Action focus: Protect high-value users by amplifying creator reward and network effects.

Instagram Retention Risk- Efficient Engager Paradox:

Efficient Engagers — the segment achieving the highest Engagement Quality Efficiency (EQE) with relatively low time investment, low ad exposure, and low content creation — exhibit the highest average churn risk score.

Despite moderate activity levels and no extreme fatigue signals (ads, sessions, notifications), their risk suggests a lack of sufficient social validation, growth feedback, or emotional stickiness.

Retention efforts should prioritize rewarding creation and amplifying network effects for these efficient but potentially unfulfilled users, rather than focusing solely on passive dormancy or heavy-user burnout."

02

Actionable Steps

Monetization Cooling For Efficient Engagers:

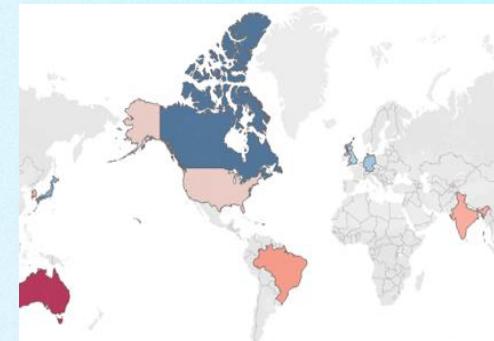
The Problem: We disproved the hypothesis that high-quality users are safe; they actually have the highest Churn Risk Score (0.88) because they are hit with 3x the Ad Load.

The Action: Cap ad frequency specifically for the Efficient Engager segment.

The Goal: Reduce the "Ad Fatigue Gap" to preserve our most intentional and valuable user base before they hit the dormancy threshold.

Fatigue Hotspots:

- **Prioritize ad-load testing** - in high-fatigue regions (e.g., India, Australia) - reduce frequency or improve relevance for users above a certain session threshold.
- **Boost local creator amplification in low-EQE markets** - test region-specific Reels ranking to increase meaningful content exposure.
- **Monitor regional churn delta** - track whether Efficient Engagers in hotspot countries show higher/lower risk than global average.



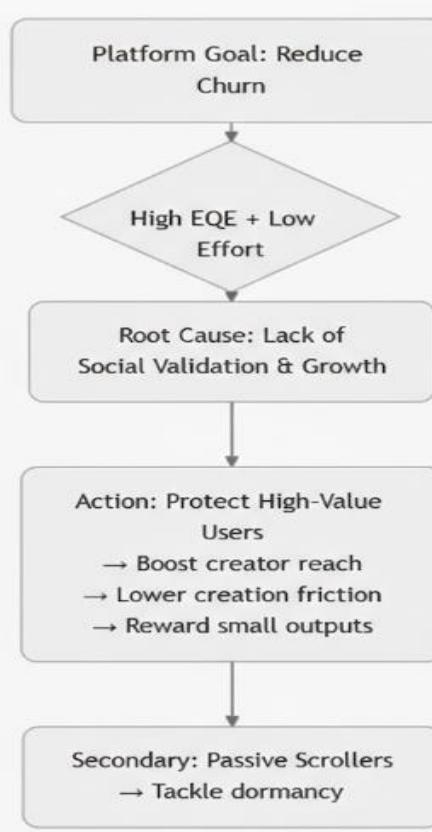
Community Health :

- **The Problem:** We disproved that the platform is a "Circle of Friends"; the Logarithmic Scale showed a massive tilt toward lopsided, non-mutual following.
- **The Action:** Introduce UX features that encourage "Mutual Follows" and "Close Friends" groups specifically for Passive Scrollers.
- **The Goal:** Move users from the "L-Tail" of the scatter plot (high volume, low value) back toward the healthy engagement median.

Shift Retention Focus from Age to Behaviour:

- **The Problem:** We disproved that age drives churn; the Age Group vs. Churn Risk chart showed that risk is uniform across all demographics.
- **The Action:** Discontinue demographic-based retention spending and reallocate the budget toward Behavioral Intervention.
- **The Goal:** Trigger "Save" campaigns based on real-time Churn Risk Scores (>0.7) rather than targeting users by age group.

The Efficient Engager Paradox:



Conclusion:

The core of this analysis reveals that Instagram's current engagement model is at odds with long-term user health. While tracking raw activity volume is necessary, our data-driven deep dive — including the L-Curve and behavioral clustering — highlights a fragile ecosystem where our most valuable users (Efficient Engagers) achieve high EQE with low effort but still face the highest churn risk (0.88). The "L-Curve" proves that high activity eventually erodes meaningful social value, shifting users toward passive scrolling and regional fatigue in hotspots like India and Australia. To secure future growth, pivot from a volume-heavy model to a reciprocal community model: implement Ad Fatigue Caps in high-risk regions, EQE-weighted algorithms to reward efficient users, and creation nudges to boost validation and stickiness.

THANK YOU