Early-Stage Retention, Conversion & Persona Profiling analysis for MagicAIBooth

István Péter Járay

MS in Business Analytics, Central European University

In partial fulfilment of the requirements for the degree of Masters Sciences of Business Analytics.

Capstone Project Manager: Eduardo Arino de la Rubia
Project Sponsor: Zoltán C. Tóth

Vienna, 2025.

Table of Contents

Report Summary

KATIKO Kft is an early-stage Hungarian startup with two founders, Zoltán C. Tóth and Viktor Belenyesi. They built a generative AI application called MagicAIBooth that has launched on the 29th of March on the App Store. They recruited a data scientist (the author) to analyze the event logs generated through usage to monitor their app's performance and make downstream decisions from key insights.

A brief overview of the app: the user can upload a selfie and type in a prompt, then receive an ai-edited selfie. The primary use case is for entertainment purposes.

The core user funnel is also intentionally very short: the user opens the app, takes a selfie or uploads one from their gallery, either selects a prompt from a list of templates or types a custom prompt, then they can press the generate button. And finally, if they like the result, they can save or share the image by pressing the share button – through a combined, iOS-specific built-in modal.

Important questions to the client

Upon initiating the project, the team has decided to follow the Lean Analytics framework. According to it's guidelines, and the team's self-assessment the product is positioned between the *Empathy* and *Sticky-engine* stages. *Empathy* implies that an early-stage startup first needs to assess their target group qualitatively, understand their pains and drives, and design the app around a solution to those pains. The *Sticky-engine* is the following step, when the primary goal for a startup is to get their users to keep coming back. As Katiko already had an app, but had no prior understanding on the target group, this in-between position felt validated.

Consequently, our Analytical North Star was 7-day Retention: What ratio of users come back in the seven days following the first open of the application? To also understand conversion more deeply, we decided to assign clicking the share button (as the last stage of the user funnel) as a complementary behaviour we're interested in – explicitly, the ratio of users who press share.

They also requested the analyst to provide some insights on behaviour and app-performance by three key demographic dimensions: age, gender and geographic location.

The delivery process

Data Sources

First, two key services must be mentioned that the founders leveraged as part of the network of services to build the application. PostHog is an analytical platform with the purpose of driving real-time insight generation based on event logs. Due to it's relative ease of use, serverless configuration and free service for early-stage startups, it was thought to be an ideal candidate to rely on for descriptive insights. AWS is a well-known, industry-leading cloud provider providing a wide array of services.

PostHog was designated to be the core analytical platform with real-time insights on customizable dashboards. The founders log all behaviourally relevant events and meaningful metadata on this platform and leverage the automatic GeoIP-enrichment service they have for locational data. Whereas they host both transactional data (e.g., backups, event exports) on a *PostgreSQL* relational database for redundancy, and extract inferred age- and gender-related metadata from the presentation on the selfies through *AWS Rekognition*.

Real-time performance monitoring – interim deliverable

The founders requested upon the project initiation a live dashboard configured to show real-time funnel and retention reports, daily and weekly active users, and new users. I first set up an example dashboard with Markov chain simulated usage which verified that the expectations align. Then a dashboard was integrated into the test-, and finally the live environment.

Salvaging a logging issue

The founders designed a workaround solution to keep the backend safe from malicious activity while aligning with Apple's guidelines by not prompting the users to log in prior to any activity. The implementation wasn't flawless, and resulted in logging each session to a different user. As a result, the live dashboards on the user level weren't reliable, and a custom implementation was required to report the KPIs consistently. In this custom implementation, some user-level connections could be salvaged based on temporal proximity. Consequently, the PostHog dashboard was left out of the Capstone Report.

The final evaluation pipeline

The final evaluation pipeline consists of the following phases: (1) data harmonization; (2) descriptive data analysis; (3) descriptive modelling; (4) predictive modelling.

First, the aforementioned data sources where harmonized and filtered for consistent reports. This step also included content intelligence categorization of the user prompts, including an LLM call. Then, the descriptive analysis section shed light on the 7-day Retention performance and the user-level conversion ratios to each step in the funnel. It also provided insight into the key behavioural differences between demographic cohorts, and temporal dynamics of user activity. The descriptive modelling section provided an overview of the association between key demographics and funnel progression, and the predictive modelling section showcased, that behavioural data combined with demographic metadata can be leveraged to predict conversion to later stages.

Key Insights

Very briefly, I can summarize my findings as the following: : (1) Older users exhibit a higher conversion rate; (2) Gender does not appear to have a significant association; (3) Further exploration of the interaction between gender and age is recommended; (4) Users from Hungary are more likely to convert; (5) Increased usage shows a weak positive correlation with conversion likelihood.

Business Recommendations

The 7-day retention analysis indicates that we achieve approximately 12% retention on the first day of usage, with negligible retention in the following days. This suggests that the application has not yet reached product-market fit, and the founders may need to either pivot or iterate on our current offering. I recommend conducting user interviews and exploring new ideas for further development. Also, I recommend finding a targeted demographic segment, because our results indicate that age, gender and location can significantly influence usage behaviour and patterns.

Learning experience

This project helped me grow as a data scientist by allowing me to work closely with a senior developer team. I learned important aspects of app development, and how important time-to-insight truly is to have a positive impact.

Copyright Notice

Author’s Declaration

I, István Péter Járay, the undersigned candidate for the MA/MSc degree in Business Analytics declare herewith that the present thesis titled *Early-Stage Retention, Conversion & Persona Profiling analysis for MagicAIBooth* is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright.

I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, on 08/06/2025.

Signature:

Full Name: István Péter Járay