User Churn Project | Exploratory Data Analysis

Prepared for: Waze Leadership Team

Project Overview

The Waze data team is currently developing a data analytics project aimed at increasing overall growth by preventing monthly user churn on the Waze app. Thorough exploratory data analysis (EDA) enables Waze to make better decisions about how to proactively target users likely to churn, thereby improving retention and overall customer satisfaction. This report offers details and key insights from Milestone 3, which impact the future development of the overall project.

Key Insights

Nearly all variables were either heavily right-skewed or uniformly distributed. For right-skewed distributions, most users had values at the lower end of the range. For uniform distributions, users were generally equally likely to have values across the entire range.

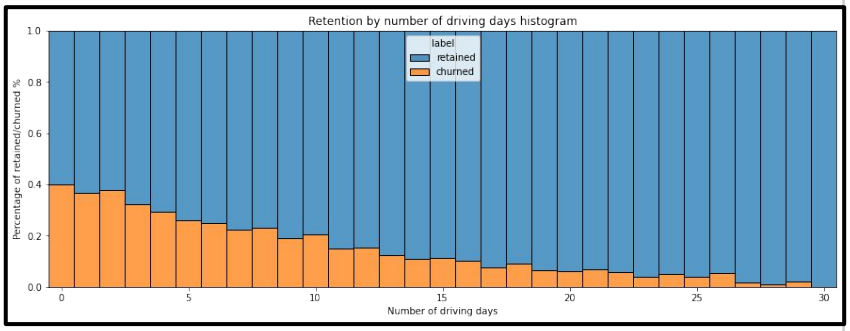
Most of the data was reliable, with no single variable appearing entirely inaccurate. However, several variables exhibited highly improbable or potentially impossible outliers, such as driven_km_drives. Additionally, some monthly variables, like activity_days and driving_days, might be problematic, as one has a maximum value of 31 while the other has 30, suggesting data collection may not have occurred within the same month for both variables.

Less than 18% of users churned, with approximately 82% being retained.

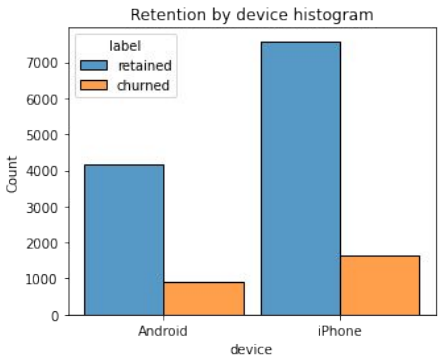
Distance driven per driving day positively correlated with user churn—the farther a user drove each day, the more likely they were to churn. Conversely, the number of driving days had a negative correlation with churn; users who drove more frequently in the last month were less likely to churn.

Users of all tenures, ranging from brand new to nearly 10 years, were relatively evenly represented in the data. This is supported by the histogram for n_days_after_onboarding, which shows a uniform distribution for this variable.

Details



- Churn decreases with increased usage of Waze over time.
- The proportion of churned users to retained users is consistent between device types.



Next Steps

- Investigate the erroneous or problematic discrepancies between number of sessions, driving_days, and activity_days.
- Continue to explore user profiles with the greater Waze team; this may glean insights on the reason for the long distance drivers' churn rate.
- Plan to run deeper statistical analyses on the variables in the data to determine their impact on user churn.