

Waze User Churn Mitigation: Data-Driven Insights & Strategic Roadmap

Milestone 2 - Leveraging Preliminary Analysis to Enhance User Retention

Overview

This report presents the findings of our initial data inspection and analysis, focusing on user churn within the Waze application. The objective is to identify key behavioral patterns and data anomalies that will inform our strategic efforts to reduce churn and enhance user retention. This preliminary analysis, conducted as part of Milestone 2, lays the groundwork for subsequent exploratory data analysis (EDA) and the development of targeted retention strategies.

Problem

Waze is experiencing a monthly user churn rate of 18%, representing a significant loss of potential revenue and user engagement. This churn, defined as the discontinuation of Waze app usage (including uninstallation or prolonged inactivity), necessitates a data-driven approach to identify and mitigate underlying causes.

Solution

By leveraging preliminary data inspection and analysis, we have identified key behavioral patterns and data anomalies that will guide the development of targeted user retention strategies. This involves segmenting users based on their driving behavior, understanding the needs of high-usage "super-drivers," and ultimately improving the overall user experience.

Details

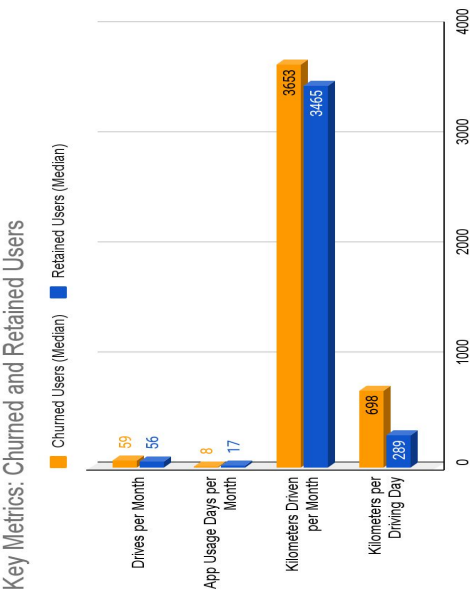
Data Characteristics: The dataset comprises 14,999 observations with 12 variables (objects, floats, integers). The 'label' column, indicating churn, has 700 missing values, which appear to be randomly distributed.

Behavioral Patterns:

- Churned users exhibit higher drive frequency but lower app usage frequency compared to retained users.
- Churned users drive longer distances and durations, particularly within a "super-driver" segment.
- The median churned user drove ~698 kilometers per driving day, 240% more than retained users.

Data Anomalies: High maximum values in 'driven_km_drives' and 'duration_minutes_drives' suggest potential outliers and a non-typical driver population.

Median vs. Mean: Using the median to analyze driving habits, removed the effect of outliers, and gave a more accurate representation of the data.



Next Steps

- Address Missing Data:** Investigate and resolve the missing values in the 'label' column to ensure data integrity.
- Conduct In-Depth EDA:** Perform comprehensive EDA to uncover deeper patterns and correlations related to user churn.
- Refine User Segmentation:** Develop detailed user segments based on driving behavior and other relevant variables.
- Investigate "Super-Drivers":** Conduct targeted research to understand the needs and pain points of high-usage "super-drivers."
- Data Augmentation:** Gather supplementary data (demographics, user feedback) to enhance user profiling and segmentation.
- Develop Data Visualizations:** Create compelling visualizations to communicate key insights to stakeholders.
- Validate User Base:** Verify if the dataset is a random sample of users, or a specific subset.