Waze User Churn Project

Milestone 2 - Compile Summary Information

Overview

Waze leadership has asked the data team to build a machine learning model to predict user churn. The model is based on data collected from users of the Waze app.

Problem

- Explore user data to learn important relationships between variables.
- Compile summary information about the data.

Solution

- Built dataframe using Python.
- Analyzed user behavior.
- Collected preliminary statistics

Details

- Our user data contains 14999 rows of user data. There are **82% retanied users and 18% churned users** among them.
- The dataset contains 12 unique variables with data type: object, integer, and float.
- 700 rows contain missing value under variable 'label'. There is no pattern shown in the missing data so there is nothing to suggest a non-random cause of the missing data.
- Churned users averaged ~3 more drives in the last month than retained users.
- Retained users used the app on over twice as many days as churned users in the last month.
- The median churned user drove ~200 more kilometers and 2.5 more hours during the last month than the median retained user.
- Churned users drove around 698 km per day which is around 240% longer distance of retained users.
- Our churned group users seem to have more drive sessions and drive distances in fewer days. We need more exploration on this.

Next Steps

- **Conduct thorough EDA and develop data visualizations** to illustrate the narrative behind the data and guide future project decisions.
- **Recommends gathering more data on the super-drivers.**
  - **Driving so much might be the reason why the Waze app does not meet super-drivers specific set of needs, which may differ from the typical driver.**