# Waze User Churn Regression Testing Summary

### Overview

The data team is developing a data analytics project aimed at increasing overall growth by reducing monthly user churn on the Waze app. Churn is defined as the number of users who have uninstalled the Waze app or stopped using the app. **This report offers details and key insights from Milestone 5, which impact the future development of the overall project**.

### Project Status

**Target Goal:** Apply user data to build and analyze a binomial logistic regression model.

Methods:

- Created features of interest to the stakeholders and business scenario.
- Assessed features for multicollinearity.
- Built the regression model.
- Evaluated model performance

**Impact:** With enough data, binomial logistic regression model results can reveal important variable relationships and predict binary outcomes, which can inform decisions for marketing and product development, for example.

### Next Steps

Due to the model results, we recommend using the key insights from this project milestone to guide further exploration.

This model should not be used to make significant business decisions; however, it has valuable insights insofar as it demonstrated a great need for additional data (features) that correlates with user churn, and also a possible need to better define the user profile Waze seeks to target in their aim to increase overall growth by preventing monthly user churn on the app.

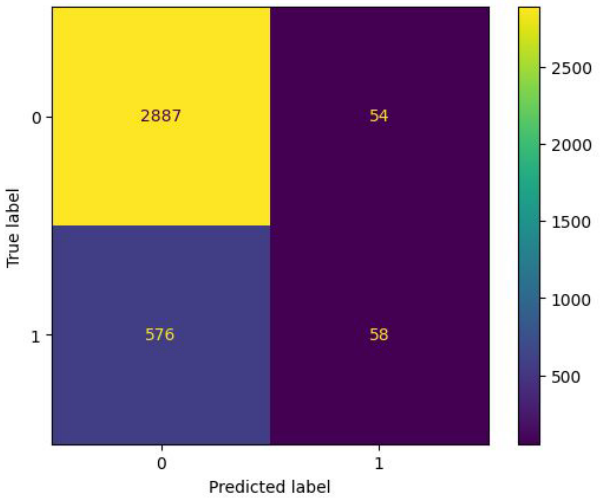
### Key Insights

The efficacy of a binomial logistic regression model is determined by accuracy, precision, and recall scores; in particular, recall is essential to this model as it shows the number of churned users.

The model has mediocre precision (53% of its positive predictions are correct) but very low recall, with only 9% of churned users identified. This means the model makes a lot of false negative predictions and fails to capture users who will churn.

Activity_days was by far the most important feature in the model. It had a negative correlation with user churn.

In previous EDA, user churn rate increased as the values in km_per_driving_day increased. In the model, distance driven per day was the second-least-important variable.




Note: 1 = churned and 0 = retained