Course 1: Foundations of Data Science

Background on the Waze scenario

Welcome to your new role at Waze! We're thrilled to have you on the data team!

Waze's free navigation app makes it easier for drivers around the world to get to where they want to go. Waze's community of map editors, beta testers, translators, partners, and users helps make each drive better and safer. Waze partners with cities, transportation authorities, broadcasters, businesses, and first responders to help as many people as possible travel more efficiently and safely.

You'll collaborate with your Waze teammates to analyze and interpret data, generate valuable insights, and help leadership make informed business decisions. Your team is about to start a new project to help prevent user churn on the Waze app. Churn quantifies the number of users who have uninstalled the Waze app or stopped using the app. This project focuses on monthly user churn. In your role, you will analyze user data and develop a machine learning model that predicts user churn.

This project is part of a larger effort at Waze to increase growth. Typically, high retention rates indicate satisfied users who repeatedly use the Waze app over time. Developing a churn prediction model will help prevent churn, improve user retention, and grow Waze's business. An accurate model can also help identify specific factors that contribute to churn and answer questions such as:

- Who are the users most likely to churn?
- Why do users churn?
- When do users churn?
-

For example, if Waze can identify a segment of users who are at high risk of churning, Waze can proactively engage these users with special offers to try and retain them. Otherwise, Waze may simply lose these users without knowing why.

Your insights will help Waze leadership optimize the company's retention strategy, enhance user experience, and make data-driven decisions about product development.

Project background

Waze's data team is in the earliest stages of the churn project. The following tasks are needed before the team can begin the data analysis process:

- A project proposal identifying the following:
  - Organize project tasks into milestones
  - Classify tasks using the PACE workflow
  - Identify relevant stakeholders

Your assignment

For your first assignment, Waze is asking for a project proposal that will create milestones for the tasks within this project. Remember to take into account your audience, team, project goal, and PACE stages of each task in planning your project deliverable.

Note: *This project's dataset was created for pedagogical purposes and does not represent Waze's actual data.*

Team members at Waze

As a new data analyst, you'll work closely with a talented team of experienced data professionals. You'll also collaborate with Waze co-workers outside the data team as your project requires.

Data team roles

- Harriet Hadzic - Director of Data Analysis
- May Santner - Data Analysis Manager
- Chidi Ga - Senior Data Analyst
- Sylvester Esperanza - Senior Project Manager

Other roles in the scenario

- Emrick Larson - Finance and Administration Department Head
- Ursula Sayo - Operations Manager

Note: *The story, all names, characters, and incidents portrayed in this project are fictitious. No identification with actual persons (living or deceased) is intended or should be inferred. The data shared in this project has been altered for pedagogical purposes.*

The members of the data team have technical expertise in data analysis, and are familiar with terms and concepts from statistics and machine learning. To communicate effectively with data team members, make sure your messages are clear, concise, and focused on key points.

Waze co-workers outside the data team may not have a background in data analysis, or be familiar with technical terms and concepts. When communicating with these co-workers, make sure to adjust your language to fit your audience, and clearly explain technical terms if necessary.

Meeting notes

After onboarding, you are given access to the company network and set up with a company email account (your first initial and last name, followed by @Waze).

Opening your inbox, you notice an email from your supervisor, May Santner.

From: "May Santner," May@waze.com

Subject: Your first task: Review meeting notes

If you are reading this message, then your company accounts have been created! Now is the perfect time to get started. Last week, I attended an internal meeting with our leadership team about a new project. You'll receive more information in the next few days, but I'd like you to be aware of some key points made by leadership. Here is an excerpt from my meeting notes. I've organized the points by the person who made them.

Sylvester Esperanza, Senior Project Manager

- The data team will need a global-level project document to outline project goals and milestones.
- We will need to generate some visuals to share with the Waze executives.

Chidi Ga, Senior Data Analyst

- The dataset has to be inspected before any analysis can begin.

- The data team needs to learn more about the data through exploratory data analysis (EDA).
- Eventually, our team will need to test the model to find out if it's delivering consistent results.

Harriet Hadzic, Director of Data Analysis

- Before we share any insights, we'll need to determine whether or not our model meets the project requirements.
- Once we finalize the model, I'll need to know the main talking points for our presentation to the leadership team.

My thoughts and concerns…

- I think it's best to use Python for this project. I'll have someone on my team set that up as soon as we have the project proposal in place.
    - It will be important to establish the relationship between key variables in the dataset. I'd suggest the data team consider hypothesis testing.

Review the meeting notes and proposed tasks from each team member to become familiar with the project's context. I'll ask you to identify project tasks and come up with a structure to guide the data team through this project. After our discussion about your experience in the Google certificate program, I know that your efficient communication style and strong problem-solving skills will help the data team achieve its goals.
There will be more details sent to you soon.
Welcome to the team,
May Santner
Data Analysis Manager
Waze
(P.S. Don't forget about our Data Superstars discussion group. Our next session is this Friday at 3pm. Each month, the team gets together to discuss the work of an innovative and inspiring historical figure. This month, it's Ada Lovelace, often considered the world's first computer programmer. This event is also a great opportunity to get to know your new teammates - and, if you want, try the outstanding coffee in the Waze cafe! Please join us!)

Waze project proposal

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.

Milestones	Tasks	Deliverables/Reports	Relevant Stakeholder (Optional)
1	Establish structure for project wo... Plan	• Global-level project document	May Santner — Data Analysis Manager
1a	Write a project proposal Plan		Sylvester Esperanza — Senior Project Manager
2	Compile summary information a... Analyze	• Data files ready for EDA	Chidi Ga — Senior Data Analyst
2a	Begin exploring the data		

	Analyze		
3	Data exploration and cleaning Plan and Analyze	• EDA report	Chidi Ga — Senior Data Analyst
3a	Visualization building Analyze and Construct	• Tableau dashboard/visualizations	Sylvester Esperanza — Senior Project Manager
4	Compute descriptive statistics Analyze	• Analysis of testing results between two important variables	
4a	Conduct hypothesis testing Analyze and Construct		May Santner — Data Analysis Manager
5	Build a regression model Analyze and Construct		

5a	Evaluate the model Execute	• Determine the success of the model	Harriet Hadzic — Director of Data Analysis
6	Build a machine learning model Construct	• Final model	
6a	Communicate final insights with ... Execute	• Report to all stakeholders	Harriet Hadzic — Director of Data Analysis

Note: The estimated times for the milestones in the example equate to the length of the courses where you will learn the necessary skills. The project timeline would most likely have tighter deadlines in a professional context involving actual clients and data scientists. For example:

Milestone 1: 1-2 days
Milestone 2: 2-3 weeks
Milestone 3: 1 week
Milestone 4: 1 week
Milestone 5: 1-2 weeks

Project goal:
Waze leadership has asked your data team to develop a machine learning model to predict user churn. Churn quantifies the number of users who have uninstalled the Waze app or stopped using the app. This project focuses on monthly user churn. An accurate model will help prevent churn, improve user retention, and grow Waze's business.
Background:

Course 1: Foundations of Data Science

Waze's free navigation app makes it easier for drivers around the world to get to where they want to go. Waze's community of map editors, beta testers, translators, partners, and users helps make each drive better and safer.

Scenario:

Your team is in the early stages of their user churn project. Your project proposal has been approved and your team has been given access to Waze's user data. To get clear insights, the data must first be inspected, organized, and prepared for analysis.

Course 2 tasks:

- Import data
- Create a dataframe
- Inspect data
- Identify outliers
- Create a data visualization
- Share an executive summary with the Waze data team

Note: *The story, all names, characters, and incidents portrayed in this project are fictitious. No identification with actual persons (living or deceased) is intended or should be inferred. And, the data shared in this project has been created for pedagogical purposes.*

Background on the Waze scenario

Waze's free navigation app makes it easier for drivers around the world to get to where they want to go. Waze's community of map editors, beta testers, translators, partners, and users helps make each drive better and safer. Waze partners with cities, transportation authorities, broadcasters, businesses, and first responders to help as many people as possible travel more efficiently and safely.

You'll collaborate with your Waze teammates to analyze and interpret data, generate valuable insights, and help leadership make informed business decisions. Your team is about to start a new project to help prevent user churn on the Waze app. Churn quantifies the number of users who have uninstalled the Waze app or stopped using the app. This project focuses on monthly user churn.

This project is part of a larger effort at Waze to increase growth. Typically, high retention rates indicate satisfied users who repeatedly use the Waze app over time. Developing a churn prediction model will help prevent churn, improve user retention, and grow Waze's business. An accurate model can also help identify specific factors that contribute to churn and answer questions such as:

- Who are the users most likely to churn?
- Why do users churn?
- When do users churn?

For example, if Waze can identify a segment of users who are at high risk of churning, Waze can proactively engage these users with special offers to try and retain them. Otherwise, Waze may lose these users without knowing why.

Your insights will help Waze leadership optimize the company's retention strategy, enhance user experience, and make data-driven decisions about product development.

Project background

Waze's data team is in the earliest stages of the churn project. The following tasks are needed before the team can begin the data analysis process:

- Build a dataframe for the churn dataset
- Examine data type of each column
- Gather descriptive statistics
- Your assignment

14,999 rows – each row represents one unique user
13 columns

Column name	Type	Description
ID	int	A sequential numbered index
label	obj	Binary target variable ("retained" vs "churned") for if a user has churned anytime during the course of the month
sessions	int	The number of occurrence of a user opening the app during the month
drives	int	An occurrence of driving at least 1 km during the month
device	obj	The type of device a user starts a session with
total_sessions	float	A model estimate of the total number of sessions since a user has onboarded
n_days_after_onboarding	int	The number of days since a user signed up for the app
total_navigations_fav1	int	Total navigations since onboarding to the user's favorite place 1

total_navigations_fav2	int	Total navigations since onboarding to the user's favorite place 2
driven_km_drives	float	Total kilometers driven during the month
duration_minutes_drives	float	Total duration driven in minutes during the month
activity_days	int	Number of days the user opens the app during the month
driving_days	int	Number of days the user drives (at least 1 km) during the month