Course 5 workplace scenarios

Waze

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:

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 more than halfway through their user churn project. Earlier you completed a project proposal, used Python to analyze and visualize Waze's user data, and conducted a hypothesis test. As a next step, leadership asks your team to build a regression model to predict user churn based on a variety of variables.

Course 5 tasks:

- Check model assumptions
- Build a binomial logistic regression model
- Evaluate the model
- Share an executive summary with the Waze leadership 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.*

Key Takeaways

In Course 5, Regression Analysis: Simplify Complex Data Relationships, you practiced modeling variable relationships, and investigated linear and logistic regression to better understand data modeling. Additionally, you reviewed model assumptions and evaluation techniques that will help you interpret and articulate relationships in datasets.

Course 5 skills:

- Conduct statistical analysis
- Conduct regression modeling

- Create predictive models
- Expand Python coding
- Share Insights and Ideas with stakeholders

Course 5 end-of-course project:

- Regression model within a Python notebook
- Executive summary with results of model and insights

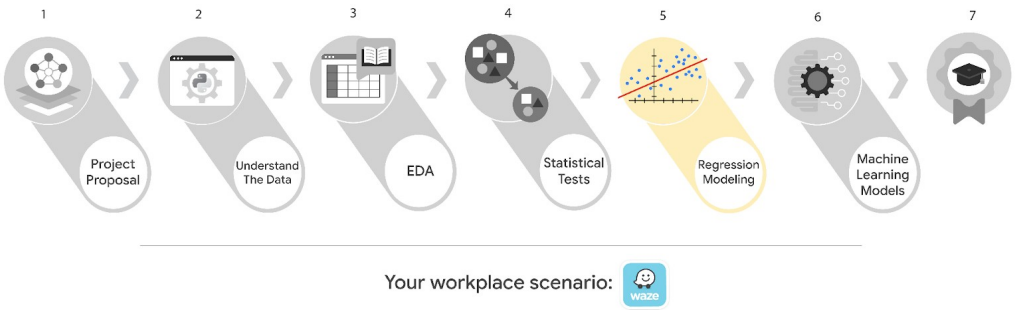
The end-of-course portfolio projects are designed for you to apply your data analytical skills within a workplace scenario. No matter which scenario you work with, you will practice your ability to discuss data analytic topics with coworkers, internal team members, and external clients.

As a reminder, you are required to complete one project for each course. To gain additional practice, or to add more samples to your portfolio, you may complete as many of the scenarios as you wish.

Course 5 end-of-course portfolio project overview: Waze

Learn about the Course 5 Waze workplace scenario!

The end-of-course project in Course 5 focuses on your ability to build regression models using Python. As a reminder, in Course 1 you developed a project proposal that outlined milestones, which progress with each of the end-of-course projects. A visual representation is provided in the graphic shown here:



Learn more about the project, your role, and expectations in this reading.

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 working on the churn project. The following tasks are needed at this stage of the project:

- Determine the correct modeling approach
- Build a regression model
- Finish checking model assumptions
- Evaluate the model
- Interpret model results and summarize findings for cross-departmental stakeholders within Waze

Your assignment

You will create a regression model for the churn project. You'll determine the type of regression model that is needed and develop one using Waze's churn project data.

Team members at Waze

Data team roles

- Harriet Hadzic - Director of Data Analysis
- May Santner - Data Analysis Manager

- Chidi Ga - Senior Data Analyst
- Sylvester Esperanza - Senior Project Manager

Data team members have technical experience with data analysis and data science. However, you should always be sure to keep summaries and messages to these team members concise and to the point.

Cross-functional team members

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

Your Waze team includes several managers overseeing operations. It is important to adapt your communication to their roles since their responsibilities are less technical.

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.*

Specific project deliverables

With this end-of-course project, you will gain valuable practice and apply your new skills as you complete the following:

- Complete the questions in the Course 5 PACE strategy document
- Answer the questions in the Jupyter notebook project file
- Build a binomial logistic regression model
- Create an executive summary to share your results

Good luck with this project! Waze looks forward to seeing how you communicate your creative work and approach problem-solving!

Key takeaways

The Google Advanced Data Analytics Certificate end-of-course project is designed for you to practice and apply course skills in a fictional workplace scenario. By completing each course's end-of-course project, you will have work examples that will enhance your portfolio and showcase your skills for future employers.

Activity Overview

In this activity, you will showcase your ability to use Python to build a multiple linear regression (MLR) model. You will also update team members and stakeholders through an executive summary, demonstrating your ability to organize and communicate key information.

For additional information on how to complete this activity, review the previous readings: *End-of-course project introduction* and *Course 5 end-of-course portfolio project overview: Waze*.

Be sure to complete this activity before moving on. The next course item will provide you with completed exemplars to compare to your own work. You will not be able to access the exemplars until you have completed this activity.

Scenario

The data team at Waze is more than halfway through their project for developing a machine learning model to predict user churn. Earlier, you completed a project proposal, used Python to explore and analyze Waze's user data, created data visualizations, and conducted a hypothesis test. Now, leadership wants your team to build a regression model to predict user churn based on a variety of variables.

You check your inbox and discover a new email from Ursula Sayo, the Operations Manager at Waze. Ursula asks your team about the details of the regression model. You also notice two follow-up emails from your supervisor, May Santner. The first email is a response to Ursula, and says that the team will build a binomial logistic regression model. In her second email, May asks you to help build the model and prepare an executive summary to share your results.

Note: *Team member names used in this workplace scenario are fictional and are not representative of Waze.*

Email from Ursula Sayo, Operations Manager

Subject: Details on Regression Model

From: "Ursula Sayo," Ursula@waze

Cc: "Harriet Hadzic," Harriet@waze**;** "Chidi Ga," Chidi@waze**;** "Sylvester Esperanza," Sylvester@Waze; "May Santner," May@waze

Hello data team,

I really appreciate your work, and thanks for the explanation of the next phase of the algorithm creation.

I was hoping to get a bit more detail on regression. Will you be applying a linear or logistic regression model? It wasn't clear in the meeting, and I want to align on expectations.

Thank you,

Ursula Sayo

Operations Manager

Waze

- - - - - - - - - – - – - - - -

Email from May Santner, Data Analysis Manager

Subject: RE: Details on Regression Model

From: “May Santner,” May@waze

Cc: “Harriet Hadzic,” Harriet@waze**;** “Chidi Ga,” Chidi@waze**;** “Sylvester Esperanza,” Sylvester@Waze; “Ursula Sayo,” Ursula@waze

Thank you for your email.

Apologies that the details were not made clear in our meeting.

To answer your question, we will build a binomial logistic regression model. Because we want to predict user churn, the binomial logistic regression model will be our confirmation for how best to proceed with the ML algorithm in the final phase of the project.

Our team will be working on getting you the results of our analysis this week.

Feel free to reach out with additional questions.

Many thanks,

May Santner

Data Analysis Manager

Waze

- - - - - – – -

Email from May Santner, Data Analysis Manager

Subject: RE: Details on Regression Model

From: “May Santner,” May@waze

Cc: “Chidi Ga,” Chidi@waze

Hello team!

Would you two mind completing the following?

- Build a binomial logistic regression model in a code notebook
- Write an executive summary of your results

I’d appreciate a chance to review your work before you send it over to Ursula, but write the summary as if you’re addressing the leadership team.

Best regards,

May Santner

Data Analysis Manager

Waze

Step-By-Step Instructions

Follow the instructions to complete the activity. Then, go to the next course item to compare your work to a completed exemplar.

Step 1: Access the templates

To use the templates for this course item, click each link below and select *Use Template*.

Link to templates:

- Course 5 PACE strategy document
- Course 5 executive summary

OR

If you do not have a Google account, you can download the templates directly from the attachments below:

Step 2: Access the end-of-course project lab

***Note**: The following lab is also the next course item. Once you complete and submit your end-of-course project activity, return to the lab instructions' page and click **Next** to continue on to the exemplar reading.*

To access the end-of-course project lab, click the following link and select *Open Lab*.

- Course 5 Waze project lab

Your Python notebook for this project includes a guided framework that will assist you with the required coding. Input the code and answer the questions in your Python notebook to build a regression model. You'll find helpful reminders for tasks like:

- Model building and evaluation
- Checking model assumptions
- Interpreting model results

You will also discover questions in this Python notebook designed to help you gather the relevant information you'll need to write an executive summary for your team.

Use your completed PACE strategy document and Python notebook to help you prepare your executive summary in the next step.

Data Dictionary

This project uses a dataset called waze_dataset.csv. It contains synthetic data created for this project in partnership with Waze. Examine each data variable gathered.

The dataset contains:

14,999 rows – each row represents one unique user

12 columns

Column name	Type	Description
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

Step 3: Complete your PACE strategy document

The **Course 5 PACE strategy document** includes questions that will help guide you through the Course 5 Waze project. Answer the questions in your PACE strategy document to prepare for using Python to organize your data and build a regression model.

As a reminder, the PACE strategy document is designed to help you complete the contents for each of the templates provided. You may navigate back and forth between the PACE strategy document and the Python notebook. Make sure your PACE strategy document is complete before preparing your executive summary.

Step 4: Prepare an executive summary

Your executive summary will keep your Waze teammates and stakeholders informed of your progress. The one-page format is designed to respect teammates and stakeholders who may not have time to read and understand an entire report.

First, select one of the executive summary design layouts from the provided template. Then, add the relevant information. Your executive summary should include the following:

- A summary of the variables analyzed in your regression model

- The results of your analysis
- Recommendations or insights based on your results

Complete your executive summary to effectively communicate your results to the Waze leadership team.

Pro Tip: Save the templates

Finally, be sure to save a blank copy of the templates you used to complete this activity. You can use them for further practice or in your professional projects. These templates will help you work through your thought processes and demonstrate your experience to potential employers.

What to Include in Your Response

Later, you will have the opportunity to self assess your performance using the criteria listed below. Be sure to address the following elements in your completed activity.

Course 5 PACE strategy document:

- Answer the questions in the PACE strategy document

Course 5 Python notebook:

- Build a regression model

Course 5 executive summary:

- Identify the outcome and impact of your work for this data project