Activity Overview

In this activity, you will showcase your ability to use statistical methods to analyze and interpret data. In particular, you will compute descriptive statistics and conduct a hypothesis test. 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 4 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

Your Waze team is nearing the midpoint of their project to develop a machine learning model to predict user churn. So far, you've completed a project proposal, and used Python to explore and analyze Waze's user data. You've also used Python to create data visualizations. The next step is to use statistical methods to analyze and interpret your data.

You receive a new email from Sylvester Esperanza, your project manager. Sylvester tells your team about a new request from leadership: to analyze the relationship between mean amount of rides and device type. You also discover follow-up emails from three other team members: May Santner, Chidi Ga, and Harriet Hadzic. These emails discuss the details of the analysis. A final email from Chidi includes your specific assignment: to conduct a two-sample hypothesis test.

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

Email from Sylvester Esperanza, Senior Project Manager

Subject: New Request - Analyze rides based on device type

From: "Sylvester Esperanza," Sylvester@Waze

Cc: "May Santner," May@waze; "Harriet Hadzic," Harriet@waze; "Chidi Ga," Chidi@waze

Hello, data team!

Excellent work so far. The leadership team is impressed with the results, especially the analysis on the last report!

On that note, they have requested a new deliverable be added to the initial project scope. They would like a statistical analysis of ride data based on device type. In particular, leadership wants to know if there is a statistically significant difference in mean amount of rides between iPhone® users and Android™ users. Should you conclude that the difference in mean amount of rides between iPhone users and Android users is statistically significant, please discuss next steps: what are your thoughts on strategies our team could implement to address these differences, such as improving user experience on a specific device?

Many thanks!

Sylvester Esperanza

Senior Project Manager

Waze

- -

Email from May Santner, Data Analysis Manager

Subject: RE: New Request - Analyze rides based on device type

From: "May Santner," May@waze

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

Thanks, Sylvester.

It's great to hear leadership is happy. I'm reminded again what a great data team we have!

Please tell leadership we will provide them with this analysis in two weeks time.

@Chidi, my initial thought is for us to conduct a two-sample t-test to analyze the difference in the mean amount of rides between iPhone users and Android users. What do you think?

Thanks,

May Santner

Data Analysis Manager

Waze

- -

Email from Chidi Ga, Senior Data Analyst

Subject: RE: New Request - Analyze rides based on device type

From: “Chidi Ga,” Chidi@waze

Cc: “May Santner,” May@waze;“Harriet Hadzic,” Harriet@waze; “Sylvester Esperanza,” Sylvester@Waze

Hi all,

@May, I agree with you on hypothesis testing. We’ll share a summary of the results before we present to leadership.
We’ll get started right away.

Thank you,

Chidi Ga

Senior Data Analyst

Waze

- -

Email from Harriet Hadzic, Director of Data Analysis

Subject: New Request - Analyze rides based on device type

From: Harriet Hadzic,” Harriet@waze

Cc: “May Santner,” May@waze; “Chidi Ga,” Chidi@waze; “Sylvester Esperanza,” Sylvester@Waze

I support this plan of action. Thank you all.

Harriet Hadzic

Director of Data Analysis

Waze

- -

Email from Chidi Ga, Senior Data Analyst

Subject: New Request - Analyze rides based on device type

From: “Chidi Ga,” Chidi@waze

Hi there, fellow data guru!

You’ve been handling all of this work really well, by the way. Excellent job.

I was wondering if you’d like to try the hypothesis test on the user data yourself? Based on what you’ve shared with me, I’m confident you have all the skills and experience needed for this task.
What do you think?

Also, as I said in my email to May, you’ll need to draft an executive summary of the results to share with Harriet and the rest of the leadership team.
Thanks so much!

Chidi Ga

Senior Data Analyst

Waze

–

“You can have data without information, but you cannot have information without data.”

—-Daniel Keys Moran

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 the following links and select *Use Template*.

Links to templates:

Course 4 PACE strategy document

Course 4 Executive summary

OR

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

Activity Template_ Course 4 PACE strategy document
DOCX File

Activity Templates_ Executive summaries
PPTX File

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 4 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 conduct a hypothesis test. You'll find helpful reminders for tasks like:

Computing descriptive statistics

Conducting a two-sample hypothesis test

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

Column name	Type	Description
n_days_after_onboarding	int	The number of days since a user signed up for the app
otal_navigations_fav1	int	Total navigations since onboarding to the user's favorite place 1
otal_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 4 PACE strategy document includes questions that will help guide you through the Course 4 Waze workplace scenario project. Answer the questions in your PACE strategy document to prepare for using Python to inspect and organize your data.
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 teammates at Waze 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 statistical methods involved in your analysis

The results of your hypothesis test

Recommendations or insights based on your results

Complete your executive summary to effectively communicate your results to your teammates.

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 4 PACE strategy document:

Answer the questions in the PACE strategy document

Course 4 Waze project lab:

Compute descriptive statistics

Conduct a hypothesis test

Course 4 executive summary:

State the hypothesis test results clearly

List recommended next steps for the data project