**Course Four**

# From Data to Insight: The Power of Statistics



# Instructions

Use this PACE strategy document to record decisions and reflections as you work through this end-of-course project. As a reminder, this document is a resource that you can reference in the future, and a guide to help you consider responses and reflections posed at various points throughout projects.

# Course Project Recap

Regardless of which track you have chosen to complete, your goals for this project are:

* Complete the questions in the Course 4 PACE strategy document
* Answer the questions in the Jupyter notebook project file
* Compute descriptive statistics
* Conduct a hypothesis test
* Create an executive summary for external stakeholders

# Relevant Interview Questions

Completing this end-of-course project will empower you to respond to the following interview topics:

* How would you explain an A/B test to stakeholders who may not be familiar with analytics?
* If you had access to company performance data, what statistical tests might be useful to help understand performance?
* What considerations would you think about when presenting results to make sure they have an impact or have achieved the desired results?
* What are some effective ways to communicate statistical concepts/methods to a non-technical audience?
* In your own words, explain the factors that go into an experimental design for designs such as A/B tests.

**Reference Guide**

This project has four tasks; the visual below identifies how the stages of PACE are incorporated across those tasks.



**Data Project Questions & Considerations**

**PACE: Plan Stage**

* What is the main purpose of this project?

The main purpose is to analyze user data to develop a predictive model that identifies patterns and factors contributing to user churn on the Waze app.

* What is your research question for this project?

Is there a statistically significant difference in the mean number of drives between iPhone and Android users?

* What is the importance of random sampling?

Random sampling ensures that every individual has an equal chance of being selected, which reduces bias and increases the generalizability of the results.

* Give an example of sampling bias that might occur if you didn’t use random sampling.

If the sample only included users from a specific region or users who frequently use the app, it could result in biased conclusions that do not accurately represent the overall user base.



 **PACE: Analyze & Construct Stages**

* In general, why are descriptive statistics useful?

Descriptive statistics summarize and provide insights into the main features of a dataset, helping to identify patterns, trends, and outliers.

* How did computing descriptive statistics help you analyze your data?

Computing descriptive statistics provided an overview of the user engagement and driving patterns, highlighting differences between iPhone and Android users.

* In hypothesis testing, what is the difference between the null hypothesis and the alternative hypothesis?

The null hypothesis (H0) states that there is no effect or difference, while the alternative hypothesis (H1) states that there is a significant effect or difference.

* How did you formulate your null hypothesis and alternative hypothesis?

Null Hypothesis (H0): There is no difference in the average number of drives between iPhone users and Android users.

Alternative Hypothesis (H1): There is a difference in the average number of drives between iPhone users and Android users.

* What conclusion can be drawn from the hypothesis test?

Based on the p-value obtained from the two-sample t-test, if the p-value is less than 0.05, we reject the null hypothesis, indicating a significant difference in the average number of drives between the two groups. If the p-value is greater than 0.05, we fail to reject the null hypothesis, indicating no significant difference.

**PACE: Execute Stage**

* What key business or organizational insight(s) emerged from your A/B test?

The hypothesis test results indicate that there is no significant difference in the average number of drives between iPhone and Android users. This suggests that device type does not play a crucial role in influencing user engagement in terms of the number of drives.

* What recommendations do you propose based on your results?

Since there is no significant difference in user engagement between iPhone and Android users, focus should be shifted to other factors that might influence user churn. Further analysis should be conducted on variables such as session frequency, driving distance, and user demographics. Additionally, strategies to improve overall user experience, regardless of device type, should be considered, such as personalized incentives and enhanced app features.