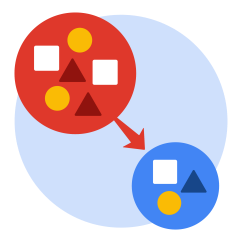
**Course Four**

# 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 compare the two devices and see if there are significant differences in usage behavior.

* What is your research question for this project?

Is the usage the same on both device types?

* What is the importance of random sampling?

Random sampling ensures that the chosen sample is representative of the population.

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

An example would be to gather data based on convenience which makes it easier but very biased because the source is mostly the same or similar.



 **PACE: Analyze & Construct Stages**

* In general, why are descriptive statistics useful?

Descriptive Statistics help get a first idea of the data without diving in too deep. They give a general idea of the 5 number summary.

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

It helped me because it gave me perspective as to where I should dive deeper.

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

The null hypothesis claims that there is no difference or the status quo and due to chance. And the alternative is that there is a difference or that the status quo is not the actual status, not due to chance.

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

- H0: The population mean of iPhone is equal to the population mean of Android

- H1: The population mean of iPhone is unequal to the population mean of Android

* What conclusion can be drawn from the hypothesis test?

- Based on the p-value of 0.1433519726802059 we fail to reject the null hypothesis since it is above our significance level is 5% and the p-value more than 14%. Therefore there is not a statistically significant difference in the means.

**PACE: Execute Stage**

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

- There is no difference between Android and iPhone users, which hints that it is not an issue of difference in interface or design

- In general the device type does not affect the number of drives made.

- This helps us understand that we have to look further for other causes to determine the root cause of the churn rate.

* What recommendations do you propose based on your results?

- Look further and continue exploring and testing for possible causes regarding the churn rates.