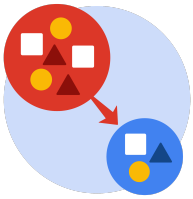
Course Four

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

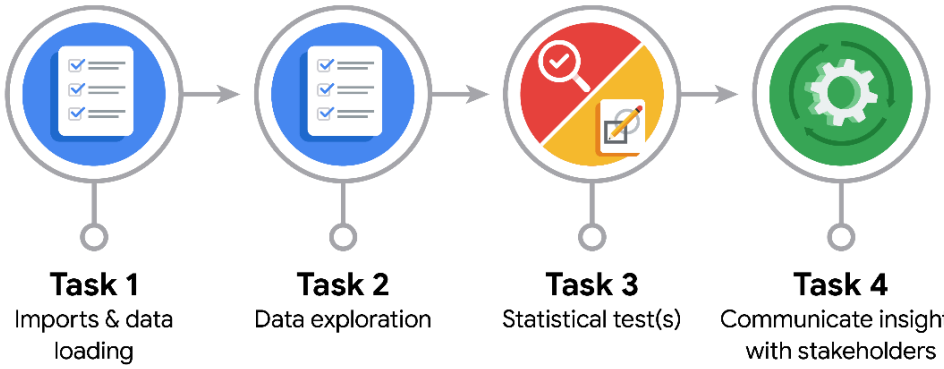
Relevant Interview Questions

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 New York City Taxi and Limousine Commision (TLC) wants to determine the fare ride before the clients request a trip.

- What is your research question for this project?

  What possible variables affect the taxi cab fare? Is the fare amount, trip distance, and type of amount the ones that are important? Is there a relationship between total fare amount and payment type?

- What is the importance of random sampling?

  Random sampling helps to select a portion of the population that is representative

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

  In the scenario that a company wants to determine how users interact in an old website and in the new website. If you don't select random sampling and you opted by choosing the users who enter in the website during weekdays with a computer. In this example, you exclude other users that enters in the weekends, and other factors.

PACE: Analyze & Construct Stages

- In general, why are descriptive statistics useful?

It gives you an insight of your data by knowing the measure of central tendency, the measure of variability, and distribution.

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

You can determine values or factors that appear frequently in your data. It also helps you determine some outliers or data that may not be relevant to your data. Additionally, it can determine if there is a relation between variables or not.

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

The null hypothesis is the one that occurred by any chance and that there is no relationship between variables. And the alternative hypothesis is the opposite.

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

Basically, if you want to see that there is no relationship in any of your variables, is the null hypothesis. Otherwise, it became the alternative hypothesis.

- What conclusion can be drawn from the hypothesis test?

If there is or not a relation between the variables that you establish.

PACE: Execute Stage

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

- What recommendations do you propose based on your results?