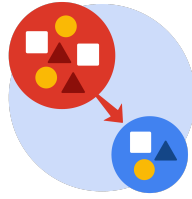# Course Four
## From Data to Insight: The Power of Statistics



## 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.

**Task 1**
Imports & data loading

**Task 2**
Data exploration

**Task 3**
Statistical test(s)

**Task 4**
Communicate insights with stakeholders

## Data Project Questions & Considerations

### PACE: Plan Stage

- What is the main purpose of this project?

  To determine if different payment methods have an impact on fare amounts.

- What is your research question for this project?
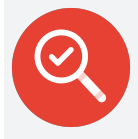
  Is there a relationship between the total fare amount and payment type?

- What is the importance of random sampling?

  Each member of the population gets an equal chance to be in the sample.

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

  > Undercoverage bias. The sample is no longer representative of the target population.

**PACE: Analyze & Construct Stages**

- In general, why are descriptive statistics useful?

  > Descriptive statistics are useful because they let us quickly explore and understand large amounts of data.

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

  > Computing descriptive statistics helps us quickly compare the average total fare amount among different payment types.

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

  > Null hypothesis is assumed to be true and to have happened by chance until proven otherwise. Alternative hypothesis is opposite of null hypothesis. It is true when the null hypothesis is proven wrong.

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

  > Analyzed what the stakeholder's needs were very thoroughly.

- What conclusion can be drawn from the hypothesis test?

> There is a statistically significant difference in the average fare amount between customers who use credit cards and customers who use cash.

## PACE: Execute Stage

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

> The key business insight is that encouraging customers to pay with credit cards can generate more revenue for taxi cab drivers.

- What recommendations do you propose based on your results?

> This project requires an assumption that passengers were forced to pay one way or the other, and that once informed of this requirement, they always complied with it. The data was not collected this way; so, an assumption had to be made to randomly group data entries to perform an A/B test. This dataset does not account for other likely explanations. For example, riders might not carry lots of cash, so it's easier to pay for longer/farther trips with a credit card. In other words, it's far more likely that fare amount determines payment type, rather than vice versa.