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

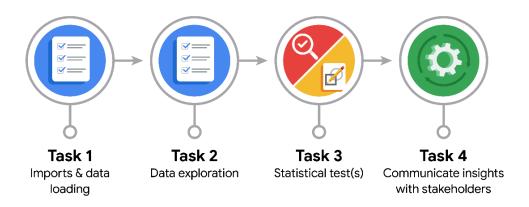
Relevant Interview Questions

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



What is the main purpose of this project?

To determine whether different payment methods affect average fares paid.

• What is your research question for this project?

Do the customers who use credit card pay higher amounts than the customers who pay using cash?

What is the importance of random sampling?

The importance of random sampling is to get a sample that accurately represents the population.

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

Non-response bias.





PACE: Analyze & Construct Stages

• In general, why are descriptive statistics useful?

Descriptive statistics are useful because they would provide an initial understanding of the averages and the distribution of our target variables.

How did computing descriptive statistics help you analyze your data?

Computing descriptive statistics help me to get an initial understanding of the average fare paid amount between the two payment methods.

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

The null hypothesis is which we assume to be true unless there is evidence proving the contrary. Alternative hypothesis contradicts the null hypothesis and it's only true if and only if there is clear and convincing evidence in favor of it.

How did you formulate your null hypothesis and alternative hypothesis?

Null hypotheses follow an equation sign. In this scenario, the null hypothesis is that there is no difference between the average amount paid between the two payment methods.

Alternative hypothesis contradicts the null hypothesis. If the null hypothesis assumes no difference between the average, the alternative hypothesis got to be that there is difference between the average amount paid.

• What conclusion can be drawn from the hypothesis test?

There is statistically significant evidence that customers who pay using credit card tends to pay more than the customers who pay using cash.



PACE: Execute Stage

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

The key business insight would be that there is statistically significant evidence that people who paid using credit cards could generate more revenue for drivers.

What recommendations do you propose based on your results?

Installing signs that say "Credit Card Payments Preferred" and telling drivers to encourage (not force) customers to pay using credit card.