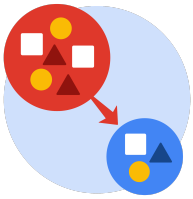
Course Four

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~~

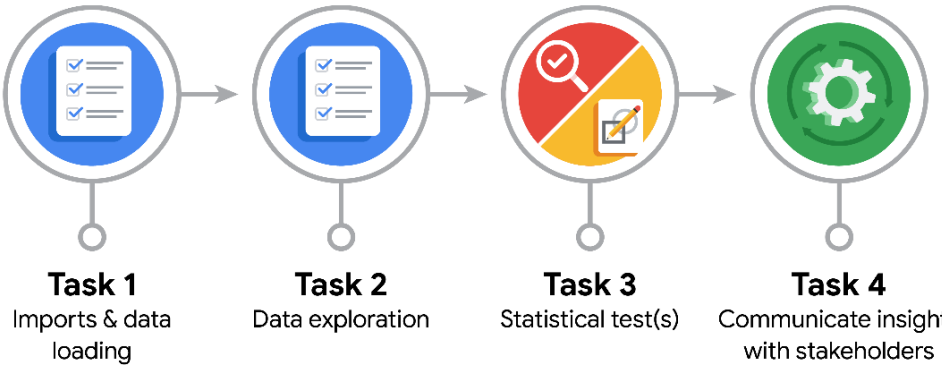
Relevant Interview Questions

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 of this project is to implement a 2 sample hypothesis test in order to arrive to a data driven conclusion.

- What is your research question for this project?

Is there a statistically significant difference between the number of views on videos published by unverified artists to the number of views on videos published by verified artists?

- What is the importance of random sampling?

It avoids introducing bias into the work being done.

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

It could skew the results towards a sample that is not representative of the population of interest.

PACE: Analyze & Construct Stages

- In general, why are descriptive statistics useful?

They provide a comprehensive summary of the location and scale of the population.

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

It help me gauge the order of magnitude difference between these two populations.

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

The null hypothesis states the observed relation is due to chance while the alternative hypothesis states the opposite. Rejecting the null hypothesis in this case would suggest that there is some statistical significance to this relationship.

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

I formulated the null hypothesis such that it stated that the difference between the two populations was non-present.

- What conclusion can be drawn from the hypothesis test?

We concluded that the is statistically significant difference between the number of views given the author categories.

PACE: Execute Stage

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

Unverified authors get more views than verified authors, and this difference is not due to chance.

- What recommendations do you propose based on your results?

The is a concerning discrepancy between the number of unverified artists to verified artists. Further, investigation about the guidelines of verifying an author should be conducted.