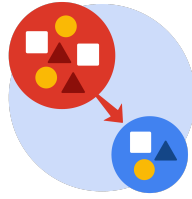


Course Four

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?

To build a prediction model to classify claims from videos.

- What is your research question for this project?

To test whether there is a significant difference between the video views for verified versus unverified accounts.

- What is the importance of random sampling?

To have a representative sample of the population by minimizing bias and other unnecessary influences.



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

In this case, a sampling bias can be including only active authors for verified.



PACE: **A**nalyze & **C**onstruct Stages

- In general, why are descriptive statistics useful?

They help in describing what is already shown by the data, help in better understanding of the data itself before trying to draw conclusions from it.

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

Help identifying missing values, outliers, mean and/or dispersion.

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

Null hypothesis states there is no change in the status quo while alternative states there is or has been a change in the status quo.

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

By looking at the research project's given objective, i.e., to test whether there is a significant difference between the video views for verified versus unverified accounts.



- What conclusion can be drawn from the hypothesis test?

There is a statistically significant difference in the video views between verified and unverified accounts.



PACE: Execute Stage

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

The analysis shows that there is a difference in the number of views between TikTok videos posted by verified accounts and TikTok videos posted by unverified accounts. As a result, these findings suggest there might be fundamental behavioral differences between these two groups of accounts: verified and unverified.

- What recommendations do you propose based on your results?

I suggest moving forward towards building a regression model to help analyze the behaviour in these verified and unverified group of accounts.