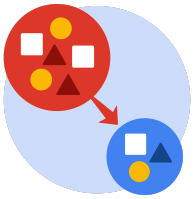
Course Four

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

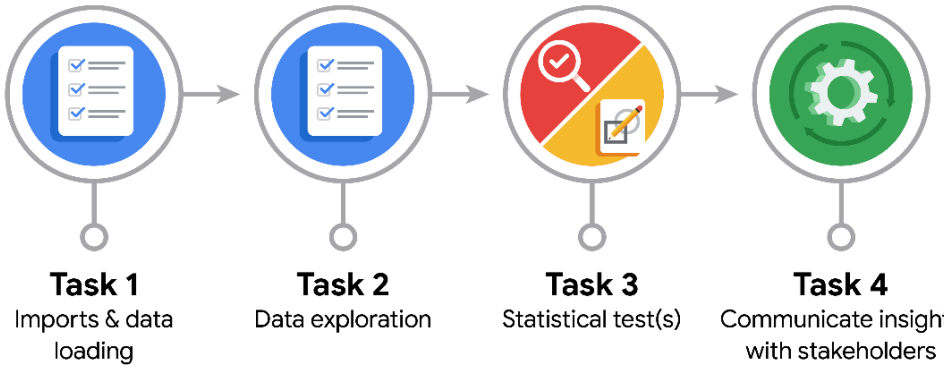
Relevant Interview Questions

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 analyze the relationship between the device type (iPhone or Android) and the mean number of rides taken by Waze users, using statistical hypothesis testing.

- What is your research question for this project?

Is there a notable statistical disparity in the average number of trips taken by users based on whether they are utilizing an iPhone or an Android device?

- What is the importance of random sampling?

Random sampling is crucial because it ensures that the sample data is representative of the entire population. This minimizes bias and allows for valid statistical inferences to be made about the population from the sample.

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

If we only sampled users from a tech-savvy community forum, we might overrepresent iPhone users, as they are often early adopters of technology. This would skew the results and not accurately reflect the overall Waze user base.

PACE: Analyze & Construct Stages

- In general, why are descriptive statistics useful?

Descriptive statistics are useful because they summarize and provide a clear overview of the main features of a dataset, such as central tendency, dispersion, and distribution. They help to identify patterns, outliers, and potential relationships within the data.

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

Descriptive statistics provided a rapid overview of the dataset, allowing for quick comparisons of average trip counts between device categories. This initial assessment helped determine if observed differences warranted further statistical investigation.

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

The null hypothesis (H0) states that there is no significant difference or relationship between variables. The alternative hypothesis (HA) states that there is a significant difference or relationship.

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

Null Hypothesis (H0): The mean number of trips taken by iPhone users is statistically equivalent to the mean number of trips taken by Android users. **Alternative Hypothesis (HA):** There is a statistically significant difference in the mean number of trips between users of iPhones and users of Android devices.

- What conclusion can be drawn from the hypothesis test?

Given the p-value of 0.1433, which exceeds the established significance level of 0.05, we do not have sufficient evidence to reject the null hypothesis. Therefore, we fail to reject the null hypothesis. Meaning there is not a statistically significant difference between the two groups.

PACE: Execute Stage

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

The outcome of the t-test indicates that device type, specifically iPhone versus Android, does not appear to significantly influence the average number of trips taken by users. This finding suggests that Waze's user engagement, as measured by trip frequency, is not significantly driven by the type of mobile operating system used.

- What recommendations do you propose based on your results?

Consequently, efforts to improve user engagement should likely focus on factors other than device type. Further exploration into other variables that could impact user behavior is recommended, as well as the testing of changes to the app itself to observe changes in user behavior.