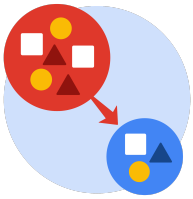
Course Four

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

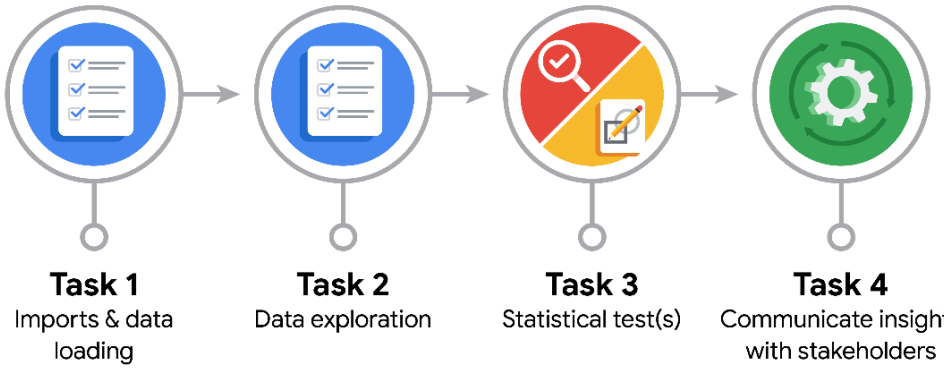
Relevant Interview Questions

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 ride data based on device type (iPhone and Android) to determine if there is a statistically significant difference in the mean number of rides. This analysis will provide insights into user behavior, potentially leading to strategies for improving user experience on specific devices.

- What is your research question for this project?

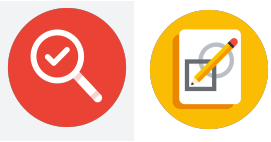
The research question for this project is: "Is there a statistically significant difference in the mean amount of rides between iPhone users and Android users?"

- What is the importance of random sampling?

Random sampling is crucial in ensuring the generalizability of the results to the entire user population. By randomly selecting samples from the population, you reduce the risk of introducing bias and increase the likelihood that the sample is representative of the larger user base. This is essential for drawing valid conclusions.

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

If we didn't use random sampling and, for example, only collected data from users who actively participated in a loyalty program or opted into a survey voluntarily, it could introduce selection bias. This might lead to overrepresenting certain user groups, such as frequent users or those more inclined to participate, while underrepresenting others. This bias could skew the results and not accurately reflect the broader user population.



PACE: Analyze & Construct Stages

- In general, why are descriptive statistics useful?

Descriptive statistics are valuable because they provide a summary of the main features of a dataset. They help in understanding the data by presenting it in a more manageable and interpretable form. Descriptive statistics include measures such as mean, median, standard deviation, and range, which offer insights into the central tendency and variability of the data.

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

Computing descriptive statistics is essential in this project because it allows us to gain a preliminary understanding of the distribution of ride data for iPhone and Android users. By calculating the mean and standard deviation for each group, we can compare their central tendencies and variabilities, which is a crucial step before proceeding to hypothesis testing.

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

In hypothesis testing, the null hypothesis (H0) represents the default or no-effect scenario. It suggests that there is no significant difference or effect in the population. The alternative hypothesis (Ha or H1) is the statement you want to test. It asserts that there is a statistically significant difference or effect in the population.

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

For this project, your null hypothesis (H0) would be: "There is no statistically significant difference in the mean number of rides between iPhone users and Android users."

The alternative hypothesis (Ha) would be: "There is a statistically significant difference in the mean number of rides between iPhone users and Android users."

These hypotheses provide a clear basis for testing whether there's a meaningful distinction between the two user groups.

- What conclusion can be drawn from the hypothesis test?

The conclusion from the hypothesis test will depend on the results obtained. If the test indicates that there is a statistically significant difference between iPhone and Android users in terms of the mean number of rides, this would imply that the choice of device does impact user behavior. Conversely, if the test does not show statistical significance, it suggests that there is no substantial difference in ride numbers between the two groups.

PACE: Execute Stage

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

The key business insight that emerged from the hypothesis test is that there is no statistically significant difference in the mean number of drives between iPhone and Android users on Waze. This means that the choice of device does not significantly influence user behavior in terms of the frequency of using the app for navigation purposes.

- What recommendations do you propose based on your results?

In light of these findings, I propose the following recommendations:

a. Focus on Universal User Experience Improvements: Given that device type doesn't appear to be a significant driver of user behavior, the team should prioritize universal user experience improvements. This entails enhancing the overall usability, features, and performance of the Waze application to benefit all users, regardless of their device.

b. Monitor User Behavior Trends: While the current analysis didn't identify a difference, user behavior can change over time. It's important to maintain ongoing monitoring of user data to detect any emerging trends or shifts. Regular data analysis and user surveys can help uncover evolving user preferences and needs.

c. A/B Testing for Specific Features: Instead of focusing on device-specific changes, consider conducting A/B tests on specific features or user interface elements. By testing changes to individual features, the team can assess their impact on user engagement and make data-driven decisions.

d. User Feedback Integration: Encourage the collection of user feedback and reviews. User comments and suggestions can provide valuable insights into areas that may require improvement or enhancement. This qualitative feedback, combined with quantitative data analysis, can inform product development decisions.

e. Cross-Platform Consistency: Ensure that the user experience remains consistent and high-quality across both iOS and Android platforms. Consistency in features, performance, and usability will contribute to positive user perceptions and satisfaction.