**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?

Practice and showcase the ability to conduct hypothesis testing and A/B testing

* What is your research question for this project?

Sample and analyze whether there is a relationship between payment type and total fare amount, to be more specific: discover if users who use credit cards pay higher fare amounts than users using cash

* What is the importance of random sampling?

First, it helps to ensure that the results of the study are generalizable to the population as a whole. If the sample is not representative of the population, then the results of the study may not be accurate.

Second, random sampling helps to reduce bias in the study. If the researcher is not able to select the sample, then there is less chance that they will select members of the population who are more likely to agree with their hypothesis.

Third, random sampling helps to ensure that the results of the study are reliable. If the study is repeated using a different sample, then the results should be similar.

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

**Convenience sampling:** This is a type of non-random sampling where the researcher selects the sample based on their own convenience. For example, a researcher might survey people who are walking by their office or who are willing to participate in their study. This type of sampling can lead to bias because the researcher is more likely to select people who are similar to them or who are interested in the topic of the study.

**Volunteer bias:** This is a type of non-random sampling where the researcher selects the sample from people who volunteer to participate in the study. This type of sampling can also lead to bias because volunteers are more likely to be interested in the topic of the study or to have strong opinions about the topic.

**Quota sampling:** This is a type of non-random sampling where the researcher sets quotas for different groups of people in the population. For example, a researcher might set a quota for 50% of the sample to be women and 50% to be men. This type of sampling can lead to bias if the researcher does not carefully select the groups that they are sampling from.



 **PACE: Analyze & Construct Stages**

* In general, why are descriptive statistics useful?

They summarize data in a way that is easy to understand. This can be helpful for researchers, who can use descriptive statistics to quickly get a sense of the data they are working with. It can also be helpful for non-technical audiences, who may not be familiar with statistical terms.

They can help to identify patterns in the data. By looking at descriptive statistics, researchers can identify trends, outliers, and other patterns in the data. This can help them to better understand the data and to ask more informed questions.

They can be used to make inferences about the population. If the sample is representative of the population, then descriptive statistics can be used to make inferences about the population as a whole. This can be helpful for making decisions about the population, such as whether to launch a new product or service.

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

Computing descriptive statistics helped me to analyze my data by summarizing the data in a way that was easy to understand. This allowed me to identify patterns in the data and to ask more informed questions.

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

The null hypothesis is a statement of no difference between the two groups being compared. The alternative hypothesis is a statement of a difference between the two groups being compared.

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

I formulated my null hypothesis and alternative hypothesis based on the descriptive statistics that I computed.

* What conclusion can be drawn from the hypothesis test?

Based on the hypothesis test results, a conclusion can be drawn that there is a statistically significant difference in the average total fare amount between customers who use credit cards and customers who use cash. The calculated t-statistic value of 20.346 and the extremely low p-value of 4.53e-91 provide strong evidence to reject the null hypothesis.

Therefore, the conclusion is that there is a significant difference in the average total fare amount between credit card users and cash users. This implies that the payment method used by customers has an impact on the fare amount they incur during their taxi rides.

**PACE: Execute Stage**

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

**Payment Method Impact:** The analysis revealed that there is a significant difference in the average total fare amount between credit card users and cash users. This suggests that the choice of payment method influences the fare amount incurred by customers during their taxi rides.

**Payment Promotion:** Based on the results, it can be recommended that the taxi company encourages and promotes the use of credit card payments. This could be done through incentives, discounts, or loyalty programs to attract more customers to opt for credit card payments.

**Pricing Strategy:** The findings indicate that there may be underlying factors contributing to the difference in fare amounts between credit card and cash payments. It would be valuable to further investigate these factors and consider how pricing strategies can be optimized to maximize revenue and customer satisfaction.

**Customer Education:** The results emphasize the importance of educating customers about the potential benefits of using credit card payments, such as convenience, security, and potential cost savings. This can be done through marketing campaigns, customer communications, and providing clear information about payment options.

* What recommendations do you propose based on your results?

Encourage Credit Card Payments: Implement strategies to incentivize and promote credit card payments, such as offering exclusive discounts or rewards for customers using credit cards.

Pricing Analysis: Conduct a thorough analysis of fare structures and pricing strategies to ensure they align with the preferences and behaviors of different payment methods, aiming to optimize revenue generation.

Customer Communication: Develop targeted communication campaigns to educate customers about the benefits of credit card payments and address any concerns or misconceptions they may have.