**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 payment type (cash vs credit card) and fare amount. For this an statistical analysis of payment type and A/B hypothesis testing are required.

Test the hypothesis that customers who use a credit card pay higher fare amounts.

Should you conclude that there is a statistically significant relationship between credit card payment and fare amount, discuss what the next steps should be: what are your thoughts on strategies our team could implement to encourage customers to pay with credit card?

* What is your research question for this project?

The research questions for this project is whether there is a statistically signficant relationship between payment type and fare amount.

* What is the importance of random sampling?

Random sampling is key to avoid any bias in the data used in the analysis and associated conclusions.

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

We could select a higher % of longer distance fares in the group of trips paid with credit card. This would lead to conclude that trips paid by credit card result in higher payment. However, the reason for higher fares would be the longer trip.



 **PACE: Analyze & Construct Stages**

* In general, why are descriptive statistics useful?

Descriptive statistics are useful to evaluate, explore and understand the data.

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

Descriptive statistics provide a count of the data as well as a view of outliers.

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

The null hypothesis is the statement or claim being made (which we are trying to disprove) and the alternative hypothesis is the hypothesis that we are trying to prove and which is accepted if we have sufficient evidence to reject the null hypothesis.

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

The null hypothesis represents the status quo and generally indicates no difference is expected between two samples. The alternative hypothesis states one sample (or population) is different than the one compared with.

* What conclusion can be drawn from the hypothesis test?

Either we reject the null hypothesis or fail to reject it, based on the p-level being above or below the significance level.

**PACE: Execute Stage**

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

Using the available dataset, we conclude that trips which were paid with credit card do result in total fares that are statistically higher than those paid with cash.

* What recommendations do you propose based on your results?

Incentive use of credit card payments as a means to increase overall revenue for TLC drivers.