**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 discover if customers who pay by using a credit card pay a higher fare amount than those who use cash.

* What is your research question for this project?

Is there a connection between a type of payment and a size of payment?

* What is the importance of random sampling?

Random sampling allows to reduce an influence of any extra factors, eliminates bias and consequently increases the reliability of this research.

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

If the population would include only people who are 60 or older the research may return false results, because older people use mostly cash. This would lead to an overrepresentation of one payment type over the other.



 **PACE: Analyze & Construct Stages**

* In general, why are descriptive statistics useful?

Because it allows to summarize large quantities of data by using a rather short summary. While terms like variance, normal distribution may not be known for the majority of people it allows professionals to easily interpret data and present it for their stakeholders/public.

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

Because of large sample size manual analysis is almost impossible. It allows instantaneously to calculate specific variables and visualize data which makes it much easier to interpret.

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

Null hypothesis assumes that an event happened only by chance, without any outside influence. Alternative hypothesis is the opposite - it assumes that an event is connected to something that directly influenced its outcome.

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

Null hypothesis: the amount of fare payment does not influence the selection of a method of payment.

Alternative hypothesis: the amount of fare payment has an influence on the selected method of payment.

* What conclusion can be drawn from the hypothesis test?

The null hypothesis is false - p-value is much lower than the confidence interval.

**PACE: Execute Stage**

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

The main insight is to increase amount of payments made with credit cards.

* What recommendations do you propose based on your results?

First thing should be to determine reason behind those results - why exactly people who pay with credit card spend more? After determining the main reason(s) the next step would be to take appropriate action to increase credit card payments.