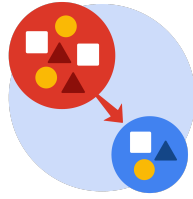# Course Four
From Data to Insight: The Power of Statistics

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.

Task 1
Imports & data loading

Task 2
Data exploration

Task 3
Statistical test(s)

Task 4
Communicate insights with stakeholders

Data Project Questions & Considerations

PACE: Plan Stage

- What is the main purpose of this project?

> T0 concludes if there is a relationship between fares and credit card payment

- What is your research question for this project?

> - The relationship between fare amount and payment type.
> - 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 the importance of random sampling?

> There is more chance to get unbiased data.
> Most likely to collect items from every type of data

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

> For example, we need to survey if employees of private companies are happier than government employees. And we take Voluntary response samples, not making sure if employees from both types of companies participated.
>
> Or we take samples from private employees only to fulfill the purpose of the study.
>
> Or conveniently takes samples within own research company only.
>
> All these are biased samples and are most likely to conclude wrong results.
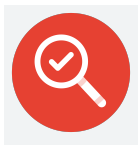
PACE: Analyze & Construct Stages

- In general, why are descriptive statistics useful?

> Descriptive statistics summarize the useful features of data and are helpful in analyzing large data sets.

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

> For this project i used descriptive statistics to analyze the total average fare amount for each payment type.

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

-The null hypothesis says there is no effect on the population. alternative hypothesis states that there is an effect on the population.

-A null hypothesis is always assumed to be true unless there is convincing evidence to the contrary. Alternative hypothesis is only assumed to be true if there is any convincing evidence.

- symbols used for null hypothesis are =,<=,>= . and for alternative hypothesis are !=,< ,>

- typically keywords used for null hypothesis are "No difference,No effect, No relationship, no change"

And for alternative hypothesis are "an effect, a difference , a relationship, a change"

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

Generally null hypothesis states that there is no effect and alternative hypothesis states there is a difference or effect in the population.

For this project
H sub o : There is no difference in the average fare amount between customers who pay with credit card and customer who pay cash.

H sub a: there is a difference in the average fare amount between customers who pay with credit card and customers who pay cash.

- What conclusion can be drawn from the hypothesis test?

There is a significant difference between the customers who pay from credit card and customers who pay cash. Customers who pay with credit card have higher average.

PACE: Execute Stage

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

Key business insight is that encouraging customers to pay with credit card will generate more revenue.

- What recommendations do you propose based on your results?

Should encourage customers to pay from credit card.