# Foundations of Data Science



# Instructions

Use this PACE strategy document to record decisions and reflections as you work through this end-of-course project. You can use this document as a guide to consider your responses and reflections at different stages of the data analytical process. Additionally, the PACE strategy documents can be used as a resource when working on future projects.

# Course Project Recap

Regardless of which track you have chosen to complete, your goals for this project are:

* Complete the PACE Strategy Document to plan your project while considering your audience members, teammates, key milestones, and overall project goal.
* Create a project proposal for the data team.

# Relevant Interview Questions

Completing this end-of-course project will empower you to respond to the following interview topics:

* As a new member of a data analytics team, what steps could you take to get 'up to speed' with a current project? What steps would you take? Who would you like to meet with?
* How would you plan an analytics project?
* What steps would you take to translate a business question to an analytical solution?
* Why is actively managing data an important part of a data analytics team's responsibilities?
* What are some considerations you might need to be mindful of when reporting results?

**Reference Guide**

This project has three tasks; the following visual identifies how the stages of PACE are incorporated across those tasks.



**Data Project Questions & Considerations**

**PACE: Plan Stage**

* Who is your audience for this project?

The audience will be the main stakeholders of TLC, which include Juliana Soto (Finance and Administrative Department Head) and Titus Nelson (Operations Manager)

* What are you trying to solve or accomplish? And, what do you anticipate the impact of this work will be on the larger needs of the client?

We are attempting to solve the problem of estimating the correct taxi fares in advance of a ride based on relevant variables.

* What questions need to be asked or answered?

What steps can i take to get up to speed with the current project, by meeting with the current team of data analysts. What is the condition of the data set? What are the useful variables? Are their trends in the data? How can I reduce the impact of bias?

* What resources are required to complete this project?

Python notebook and input from stakeholders

* What are the deliverables that will need to be created over the course of this project?

A dataset cleaned for EDA, visualizations, statistical model, regression analysis, and machine learning model.

## 

## **THE PACE WORKFLOW**



**[Alt-text: The PACE Workflow with the four stages in a circle: plan, analyze, construct, and execute.]**

You have been asked to demonstrate for the company's data team how you would use the PACE workflow to organize and classify tasks for the upcoming project. Select a PACE stage from the dropdown buttons. A few tasks involve more than one stage of the PACE workflow. Additionally, not every workplace scenario will require every task. Refer back to the Course 1 end-of-course portfolio project overview reading if you need more information about the tasks within the project.

### 

### **Project tasks**

Following are a group of tasks your company’s data team has determined need to be completed within this project. The data analysis manager has asked you to organize these tasks in preparation for the project proposal document. First, identify which stage of the PACE workflow each task would best fit under using the drop down menu. Next, give an explanation of why you selected the stage for each task. Review the following readings to help guide your selections and explanation: The PACE stages and Communicate objectives with a project proposal. You will later reorder these tasks within a project proposal.

1. **Evaluating the model: Execute**

Why did you select this stage for this task?

The construct stage is where I will be building, interpreting and revising models. Correlations will be found to uncover information from the data, and evaluate if the project meets expectations and goals.

1. **Conduct hypothesis testing: Analyze** **and** **Construct**

Why did you select these stages for this task?

A statistical test will be used in the analyzing stage. These stages are the key parts for examining and interpreting data to test hypotheses, identify patterns, and draw conclusions.

1. **Begin exploring the data: Analyze**

Why did you select this stage for this task?

Exploratory Data Analysis is completed in the analyze stage to gan a deeper understanding of the data set and information in it.

1. **Data exploration and cleaning: Plan** **and Analyze**

Why did you select these stages for this task?

The planning stage is where the first choices about the methods needed. The cleaning process is a part of the analysis stage.

1. **Establish structure for project workflow (PACE): Plan**

Why did you select this stage for this task?

Establishing structure for project workflow would fall under the plan stage because it sets the ground rules for the order of operations for a project.

1. **Communicate final insights with stakeholders: Execute**

Why did you select this stage for this task?

When we “execute” our data analysis to the stakeholders, it includes communicating the final insights to stakeholders. Communication, however, is crucial throughout the PACE project workflow.

1. **Compute descriptive statistics: Analyze**

Why did you select this stage for this task?

Investigating the statistics within data takes place during analysis.

1. **Visualization building: Analyze and Construct**

Why did you select these stages for this task?

Visualization begins with the assessment of data and is performed during the construction stage.

1. **Write a project proposal: Plan**

Why did you select this stage for this task?

Project proposals are one of the first tasks to be done in a workflow

1. **Build a regression model: Analyze and Construct**

Why did you select this stage for this task?

The model is examined in detail to be sure it will meet the requirements of the task. The building of the regression model occurs in the construct stage

1. **Compile summary information about the data: Analyze**

Why did you select this stage for this task?

This is the stage where inspecting a dataset to compile information would take place.

1. **Build machine learning model: Construct**

Why did you select this stage for this task?

This is the stage where building a data model would occur