**Course One**

# 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?

New York City Taxi and Limousine Commission (NYC TLC)

* 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?

Build a regression model that predict taxi cab fare before each ride based on distance, time of day, and additional variables.

* What questions need to be asked or answered?

1. What is condition of the provided dataset?

2. What variables will be most useful?

3. Are there trends in data that can help in providing insights?

4. What are the steps to ensure minimal impact of bias?

* What resources are required to complete this project?

1. Project dataset

2. Jupyter Notebook

3. Input from Stakeholders

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

1. Dataset processed for EDA (Explanatory Data Analysis)

2. Data Visualization

3. Statistical Model of data

4. Regression Analysis

5. 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 drop down 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](https://www.coursera.org/learn/foundations-of-data-science/supplement/4OtHr/the-pace-stages) and [Communicate objectives with a project proposal](https://www.coursera.org/learn/foundations-of-data-science/supplement/79Ysh/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?

After the model is constructed, data will be ran through to evaluate whether it meets the project’s expectations and goals. The results will then be compiled to decided the efficiency of model, and or if the model needs to be re-trained.

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

Why did you select these stages for this task?

During analyzing stage, the statistical test for the model will be determined. During the construction phase, the test is carried out based upon the model.

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

Why did you select this stage for this task?

During this analysis stage, a deeper understanding of the dataset and the information within it will be gained.

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

Why did you select these stages for this task?

Planning takes place when we make first choices about the methods needed. The cleaning process then takes place during the analyzing phase.

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

Why did you select this stage for this task?

Planning stage includes creation of an initial project and PACE document outlines the workflow and process involved and helps to plan how to best approach a project.

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

Why did you select this stage for this task?

Communication is necessary at various points throughout a project. Final insights are conveyed and shared with the stakeholders at the execute phase of the data 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 data assessment and is created during the construction stage. These visualizations can be used to explain data and the insights easily and effectively to stakeholders.

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

Why did you select this stage for this task?

Planning stage concludes with the drafting of the project proposal which is the initial document used to define a project.

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

Why did you select this stage for this task?

During the analyzing stage, the model is examined properly in detail to be completely sure if it meets the needs of the task. The building of the regression model will take place in the construction phase.

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

Why did you select this stage for this task?

Inspecting the dataset to compile information and remove any discrepancies will take place during the analysis phase.

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

Why did you select this stage for this task?

The building of the data model would take place during the construct phase.