**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

* 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?

To develop a regression model that helps estimate taxi fares before the ride.

* What questions need to be asked or answered?

Is the dataset cleaned and ready to use? What steps can be taken to reduce the bias? What all insights the data will provide?

* What resources are required to complete this project?

The following are the resources needed for the completion of project: Python, Juptyer Notebook, Data from the stakeholders.

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

Cleaned Data for Analysis, Visualizations for representing the data, Machine Learning Models, Statistical 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](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?

Evaluating the model is done to decide if it has met the stakeholders expectations and is it working properly.

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

Why did you select these stages for this task?

During analysis stage we will decide which test is to be done. And during the construct stage we will actually test it.

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

Why did you select this stage for this task?

During the analyze stage we will get into deeper into the data set which we have received.

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

Why did you select these stages for this task?

During the plan stage we will understand what the data set tells us. During the analysis phase we will get deeper into the data set to identify the trends and clean the data by removing unwanted outliers, data and other things.

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

Why did you select this stage for this task?

During the plan stage we will decide on how to complete the project by creating a project workflow.

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

Why did you select this stage for this task?

During this stage we will communicate our findings with our stakeholders to get their response.

1. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

During the analyze phase the dataset is done various statistical testing to find the trends.

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

Why did you select these stages for this task?

During the construct stage we will visualize the data trends which we got after building the model in the forms of charts, graphs etc.

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

Why did you select this stage for this task?

During the planning stage we will write the project proposal which includes the objectives, milestones, intended stakeholders and the time required for achieving each milestone.

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

Why did you select this stage for this task?

During the analyze stage we will understand the data, perform the tests and in the construct phase we will make regression models.

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

Why did you select this stage for this task?

In this phase we will gather all the data which we have got for the project and group into different categories.

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

Why did you select this stage for this task?

During this stage we will build the machine learning model after doing many statistical tests.