**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 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?

Trying to estimate Taxi fares across the city based on relevent factors that need to be identified.

* What questions need to be asked or answered?

What variables are going to be most useful in determining the fares of taxi rides?

* What resources are required to complete this project?

Input from Stakeholders, Python, and data

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

Dataset that contains only the useful information, visualizations, and models (stats, ML, or regression)

## 

## **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?

The model has to be created to evaluate it, therefor it goes after the Construct phase.

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

Why did you select these stages for this task?

Analyze, we will choose a statistical method to use, and build it in the Construct phase. So we can test at the end of the Construct phase.

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

Why did you select this stage for this task?

Planning is for making the plan before the data is available, in the Analyze stage is for looking at the data.

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

Why did you select these stages for this task?

Planning is when we take the first steps in choosing method requirements. Cleaning takes place in Analyze.

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

Why did you select this stage for this task?

Establishing Structure should be completed or at least a resonable framework before you look at the data. This can be changed as time goes on.

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

Why did you select this stage for this task?

While you need to communicate throughout the process, the final presentation of insights is given in the Execute stage.

1. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

Descriptive Statistics is used to see what is in the data, therefore it is in the Analyze stage.

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

Why did you select these stages for this task?

Initial Visualizations are built as part of EDA, and more polished Visuals are created in Construct

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

Why did you select this stage for this task?

A project proposal should be given at the end of the Planning Stage, it will give your manager a way of knowing what is going on.

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

Why did you select this stage for this task?

**During Analyze, you will ensure that the details are present in the model. Then it will be fleshed out in the construct phase.**

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

Why did you select this stage for this task?

General information will be there at the end of the Analyze step, and you should know what your data looks like at that time.

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

Why did you select this stage for this task?

**Machine Learning models are part of the construct phase, and you will need to have completed EDA beforehand.**