**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?

The audience for this project is the New York City TLC company

* 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?

I am trying to help NYC TLC build a model that will help their customers have an estimate of how much the taxi fare will cost before they get in the taxi.

* What questions need to be asked or answered?

What data do we have? Do we have historical data of all taxi fares, what other information do we have like distance, time of ride, date of ride, location/ neighborhood, holidays or special events? Are there trends within the data that can provide insight? What steps can I take to reduce the impact of bias?

* What resources are required to complete this project?

A large dataset with all necessary data, model building software preferably python, server to deploy model on, visualization tools, presentation tools, input from stakeholders

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

A dataset scrubbed for EDA, visualizations, statistical model, regression analysis and/or 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](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?

Once the model is built in the Construct stage, we analyze the results in the Execute stage

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

Why did you select these stages for this task?

To conduct hypothesis testing, we have to iteratively construct the model and evaluate the results in the execute stage

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

Why did you select this stage for this task?

It involves one of the very first steps in the analyze stage where we analyze and explore the data

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

Why did you select these stages for this task?

Data exploration falls under the analyze stage and cleaning the data involves the construct phase

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

Why did you select this stage for this task?

It involves planning the project workflow

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

Why did you select this stage for this task?

Last stage of the project where we’re executing the project by communicating results

1. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

This comes under the construct phase

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

Why did you select these stages for this task?

Building the visualizations in the construct phase and implementing feedback in the execute phase

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

Why did you select this stage for this task?

Initial stage of planning a project

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

Why did you select this stage for this task?

**We analyze the data and build the most appropriate model**

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

Why did you select this stage for this task?

Involves analyzing the data

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

Why did you select this stage for this task?

**Constructing the model**