**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 TLC

Stakeholders

Data Team

* 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?

Using regression model, predict NYC TLC ride duration based on distance, time of day, season, and other additional variables.

First impact for the client: being able to optimize routes for each vehicle / driver.

Second impact: being able to provide estimated wait time for customers, allowing both parties to plan and schedule accordingly.

* What questions need to be asked or answered?

What is the condition of the provided dataset?

What variables will be the most useful?

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?

The stored data

Team of analyst

Software / Programs

Expertise

Input from stakeholders

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

Results of the data analyzed, scrubbed for exploratory data analysis

Visualization

Regression model

Prediction / statistical model

Implementation of the prediction model / ML

## **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?

Execute: if we are evaluating the model, that means the analysis is complete and the model was created in the construction stage

The model has been constructed; data is tested to determine if the results meet the project’s goal

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

Why did you select these stages for this task?

Analyze and construct: After analysis is complete, construct a regression model and test the hypotheses.

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

Why did you select this stage for this task?

Analyze: in the analyze stage, the data is explored

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

Why did you select these stages for this task?

Plan and Analyze: in this stage the data is explored, clean, and analyze for a deeper understanding

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

Why did you select this stage for this task?

Planning: in this stage, the project workflow is created before the project can begin

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

Why did you select this stage for this task?

Execute: the analysis is complete and the results should be presented visually to clients

1. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

Analyze: part of the analysis

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

Why did you select these stages for this task?

Analyze and Construct: Analysis needed to be completed before visualization can be built

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

Why did you select this stage for this task?

Planning:

A project should not start without a proposal

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

Why did you select this stage for this task?

**Construct and Execute:**

**In these stages, the model is built and executed.**

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

Why did you select this stage for this task?

Analyze:

After the analysis, a summary / visualization report needs to be generated for stakeholders

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

Why did you select this stage for this task?

**Construct:**

**Building a machine learning model**