# 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.~~

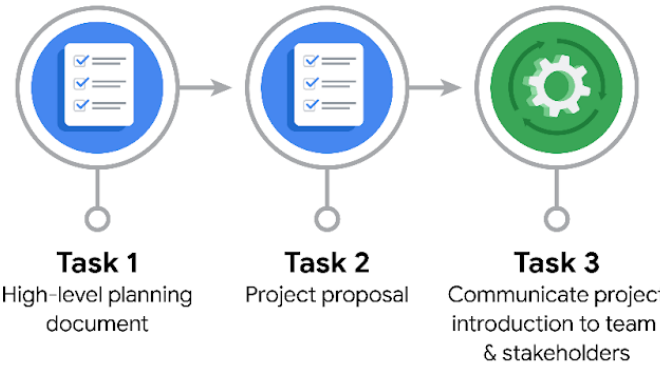
## Relevant Interview Questions

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 Taxi and Limousine Comission** (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?

The problem we are trying to solve is to estimate the taxi fares for TLC riders based on the relevant factors we can identify.

- What questions need to be asked or answered?

What is the condition of the provided dataset? What variables will be most useful? Are there trends in the data that can provide insight? What steps can I take to reduce the impact of bias?

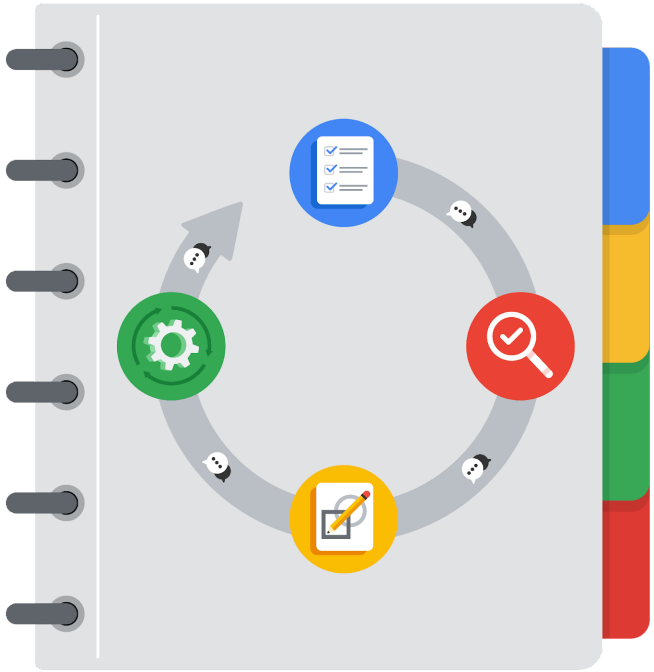
- What resources are required to complete this project?

| |
|---|
| The dataset for the project, Python notebook, and input from the stakeholders. |

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

| |
|---|
| Clean dataset for Exploratory Data Analysis (**EDA**), Develop regression and/or machine learning model, visualizations, Statistical Inference |

## 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 and 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 creating the model is created, it is tested using data to check if it meets the goals and objectives of the project. |

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

   Why did you select these stages for this task?

   | |
   |---|
   | During the analyze stage, it is determined a statistical model will be used. During the construct phase, the testing will occur. |

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

   Why did you select this stage for this task?

   | |
   |---|
   | In the analyze phase, you will gain a better understanding of the data by doing an Exploratory Data Analysis (**EDA**). |

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

   Why did you select these stages for this task?

In the Planning phase, you determine what kind of data you need. In the analyze phase, the data is cleaned and explored.

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

Why did you select this stage for this task?

In the planning phase, creating a PACE strategy document will help outline the workflow of the project and the best way to approach it.

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

Why did you select this stage for this task?

During the Execute phase, you share your final insights with stakeholders. Communication with the stakeholders occurs throughout the entire project.

7. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

Computing descriptive statistics takes place during the analysis phase.

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

Why did you select these stages for this task?

The data is assessed during the analyze phase, then visualizations are created in the construct phase.

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

Why did you select this stage for this task?

| |
|---|
| In the planning phase, a project proposal is used to define the scope of the project. |

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

Why did you select this stage for this task?

| |
|---|
| In the analyze phase, the data is explored to ensure it meets the needs of the task. In the construct phase, the regression model is created. |

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

Why did you select this stage for this task?

| |
|---|
| In the analyze phase, summary information about the data is found through Exploratory Data Analysis (**EDA**) |

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

Why did you select this stage for this task?

| |
|---|
| In the Construct phase, a machine learning model is created. |