# 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.

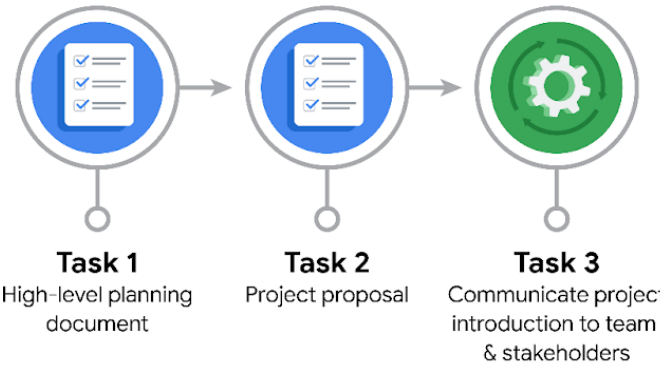
## Relevant Interview Questions

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?

My new team, for whom I have been asked to prepare a planning document, as well as the trust and safety 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?

I am preparing for the development of a machine learning model that will facilitate moderation by classifying claims as opinion or fact.

- What questions need to be asked or answered?

What is the composition and quality of the data? How is it structured? What bias may be present, and how can I ameliorate it?

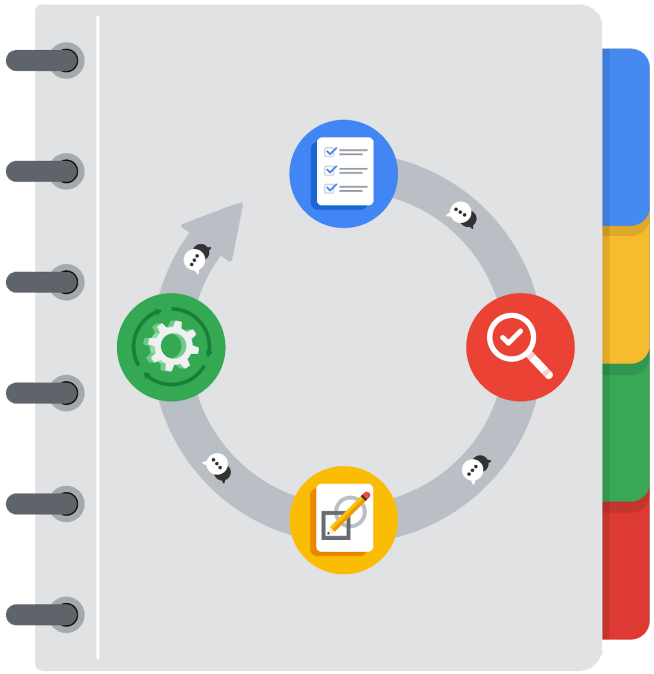
- What resources are required to complete this project?

The data set, computational resources, team input

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

A planning document that specs out the milestones and requirements of this bigger project, and a data set prepared for exploratory data analysis.

## 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?

   | |
   |---|
   | Once the model has been trained, we finalize it during the execute stage. |

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

   Why did you select these stages for this task?

   | |
   |---|
   | During the analyze stage, some basic hypotheses are tested to understand the features of the data and anticipate the kinds of models that will best represent the data.<br><br>During the construction phase, more sophisticated testing is used to begin drawing conclusions related to the underlying problem. |

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

   Why did you select this stage for this task?

   | |
   |---|
   | Exploratory data analysis is used to begin understanding the structure of the data |

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

   Why did you select these stages for this task?

Some basic exploration is necessary to understand how much time and what exact resources will be most appropriate for the given problem. Final cleaning and descriptive statistics begin the Analysis phase.

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

Why did you select this stage for this task?

This is the final step of planning, once the basic task requirements have been specified.

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

Why did you select this stage for this task?

Final insights are the most reliable and objective, and generally will not improve without input from relevant stakeholders.

7. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

This begins the Analyze phase so that we can identify interesting features in the data.

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

Why did you select these stages for this task?

Visualizations are important throughout the process so that transparency is served, and so that decision making is clear.

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

Why did you select this stage for this task?

This is the final stage of planning

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

Why did you select this stage for this task?

We first build the basic model, or models, and then in the construct phase refine their functioning and tune their parameters.

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

Why did you select this stage for this task?

This is the final step of analysis before moving on to predictive methods.

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

Why did you select this stage for this task?

This is the core of the construct phase, where the algorithm builds the model.