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.

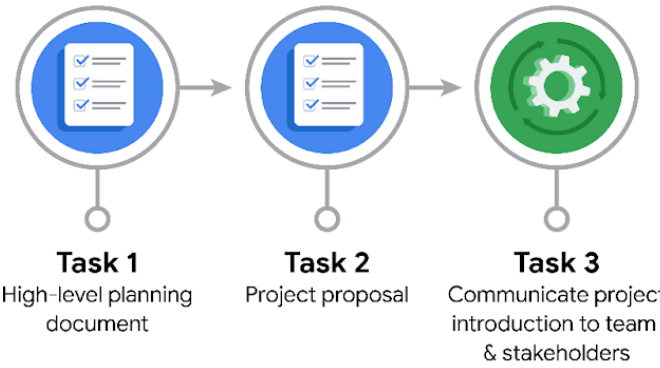
Relevant Interview Questions

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?

Audience is mainly the leadership, data, and cross-functional teams as they're the ones who identified the need and are asking for this project.

- 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?

Solving the challenge of moderators trying to classify every single piece of content (videos) on whether it offers a claim or opinion. The predictive model will do it instead. The impact will be big because will save time for moderators and they can do other tasks.

- What questions need to be asked or answered?

What format is the current data in? Are there data privacy points I need to keep in mind? Condition of dataset? Steps to reduce bias?

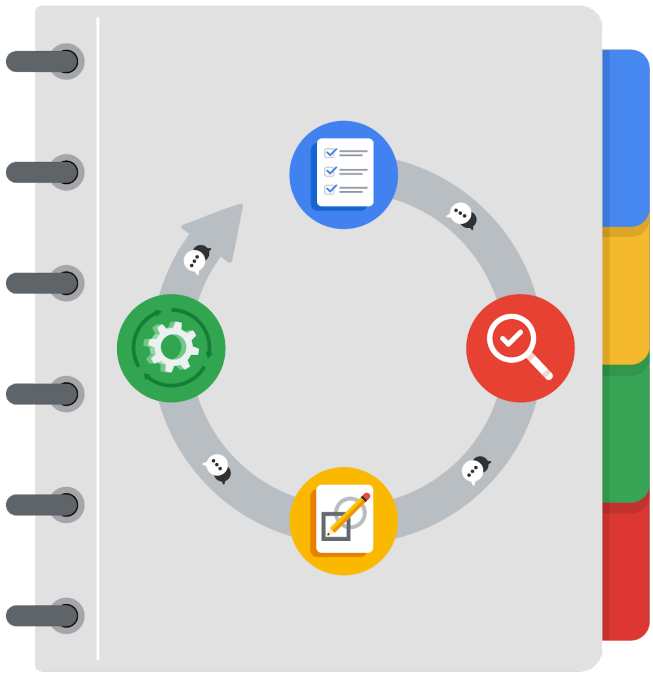
- What resources are required to complete this project?

Python, data viz software (e.g. Power BI), dataset, stakeholder input

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

Machine learning model, dashboard

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?

   To see how well the model worked and if it meets the requirements/goals, it can only be done after the model is built, which is in the earlier stage of construct.

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

   Why did you select these stages for this task?

   Analyze stage - a statistical test is determined, construct stage - test is carried out.

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

   Why did you select this stage for this task?

   This is to see what is in the data and understand it better.

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

   Why did you select these stages for this task?

   You plan what methods are needed when exploring the data and then clean it as part of analyze.

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

   Why did you select this stage for this task?

   How the workflow of the project will go needs to be planned out before the project is tackled. PACE document outlines workflow and helps to determine how to best approach a project.

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

   Why did you select this stage for this task?

   Final insights to share results with stakeholders.

7. **Compute descriptive statistics:** Analyze

   Why did you select this stage for this task?

   Because this is where after you clean the data it's ready to be broken down and its statistics analyzed.

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

   Why did you select these stages for this task?

   To make sure the data is ready for visualizations, the data needs to be prepared and assessed which happens during the analyze stage. Visualizations are created during construct.

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

   Why did you select this stage for this task?

   This is for the beginning to plan out and define the project.

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

Why did you select this stage for this task?

To build the model you need to have data cleaned and then construct the model.

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

Why did you select this stage for this task?

This is the initial gathering of information about the data by inspecting the data.

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

Why did you select this stage for this task?

This is where you build the machine learning model.