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.

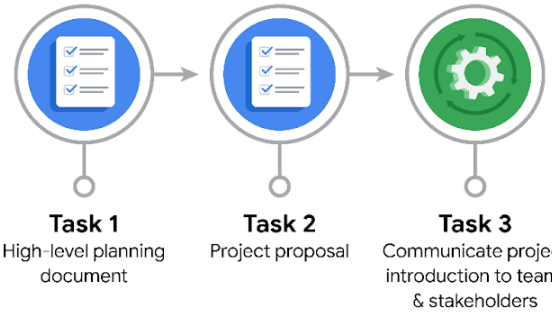
Relevant Interview Questions

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: Planning Stage

- Who is your audience for this project?

New York City Taxi and Limousine Commission

- 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?

We want to create a model that can predict how long taxi and limousine rides will take. By doing this, we can make the taxi and limousine services run smoother and make sure customers have a better experience.

- What questions need to be asked or answered?

Which variables are the most valuable for analysis? Are there any noticeable patterns or trends in the data? How can I minimise bias and ensure fairness in the analysis?

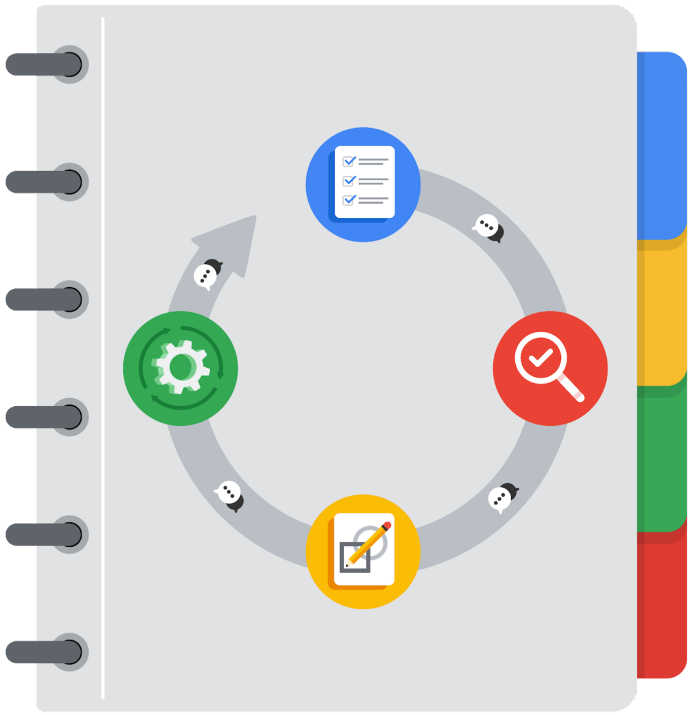
- What resources are required to complete this project?

The TLC data set, Python, Jupyter Notebook

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

Cleaned dataset, visualisations, statistical model, regression analysis, 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 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 assess if the constructed model meets project expectations and goals.

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

Why did you select these stages for this task?

Construct stage carries out the test planned during the analysis stage.

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

Why did you select this stage for this task?

Analyzing stage provides deeper insights through data cleaning and formatting.

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

Why did you select these stages for this task?

Planning stage involves method selection, while cleaning occurs during analysis

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

Why did you select this stage for this task?

Planning stage outlines the workflow and project approach.

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

Why did you select this stage for this task?

Execution phase shares final insights with stakeholders.

7. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

Analysis stage investigates statistics within the data.

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

Why did you select these stages for this task?

Construction and analysis stages involve creating and assessing visualisations.

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

Why did you select this stage for this task?

Planning stage involves defining the project through a proposal.

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

Why did you select this stage for this task?

Construction stage builds the model and the analysis stage examines its suitability.

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

Why did you select this stage for this task?

Analysis stage assesses data quality, including missing data inspection.

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

Why did you select this stage for this task?

Construction stage involves building the machine learning model.