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.
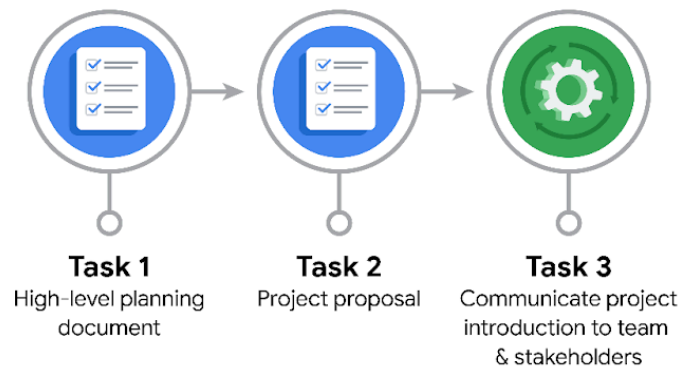
Relevant Interview Questions

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.

Task 1
High-level planning document

Task 2
Project proposal

Task 3
Communicate project introduction to team & stakeholders

Data Project Questions & Considerations

PACE: Plan Stage

- Who is your audience for this project?

New York City Taxi (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?

Create a regression model to estimate the taxi fare before the ride began by analysing previous data and various other conditions which seems to be necessary.

Impacts of this work will be that we can maximise on the revenue by keeping the price dynamic according to the situation. Also it will be convenient for clients to decide whether to take taxi or not.
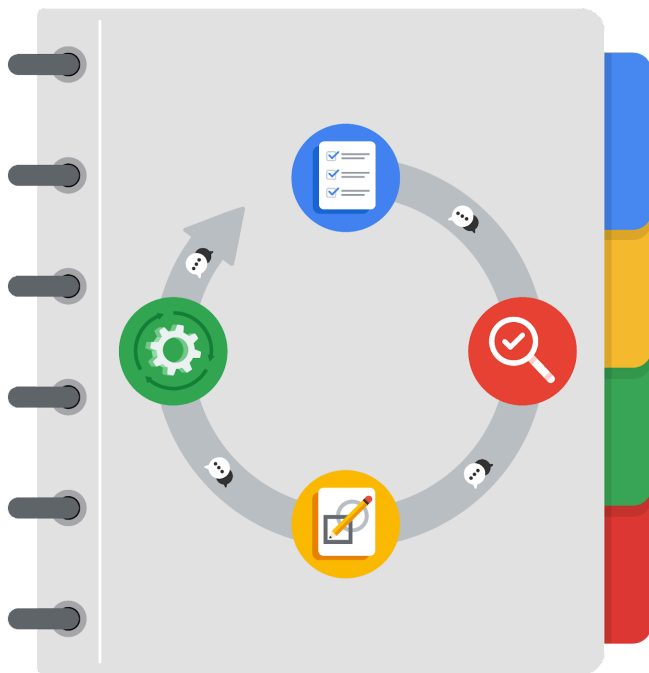
- What resources are required to complete this project?

Python to clean the messy data, sql to keep the data in ordered format and power bi to present the story to client.

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

Model that estimate fares before start of the ride based on the distance, demand and time of the day.

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:** `Construct ▾`

   Why did you select this stage for this task?

   > Creating model and evaluating is a part of making perfect fit model and hence this should be part of construct.

2. **Conduct hypothesis testing:** `Plan ▾` and `Analyze ▾`

   Why did you select these stages for this task?

   > Hypothesis testing should be planned first and then analyse based on the results.

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

   Why did you select this stage for this task?

   > Data Exploration is analysation phase where messy data is cleared to construct ML model on it.

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

   Why did you select these stages for this task?

   > Exploring data comes under planning phase while cleaning data is a analysing phase.

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

   Why did you select this stage for this task?

   > Structure of Project should have to done at the start of project and hence it will come in plan phase.

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

Why did you select this stage for this task?

> Final insights is shared with client afte building and testing the model and hence this will come in execution phase.

7. **Compute descriptive statistics:** `Analyze ▾`

Why did you select this stage for this task?

> Computing descriptive statistics is a analysing phase where data is analyzed before keeping in it ML model.

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

Why did you select these stages for this task?

> Visualization building is the phase where analysis done after we start creating the model.

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

Why did you select this stage for this task?

> Writing Project proposal is the starting point of any project.

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

Why did you select this stage for this task?

> **Building regression model is the middle phase of analyze and construct.**

11. **Compile summary information about the data:** Execute ▾

Why did you select this stage for this task?

Compilation of information is done after running ML model hence it comes under execute phase.

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

Why did you select this stage for this task?

Building ML model is construct phase.