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.
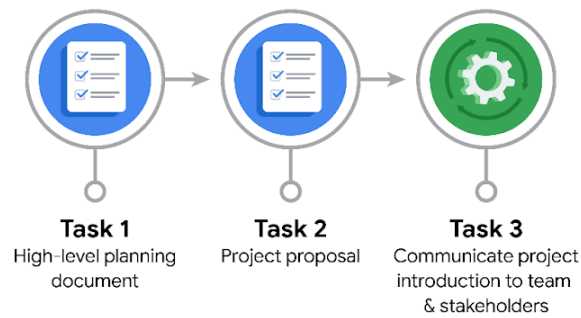
Relevant Interview Questions

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 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?

  Estimation of taxi fares before ride for NYCTLC to better regulate the industry

- What questions need to be asked or answered?

  How is the dataset formulated? Are there any trends
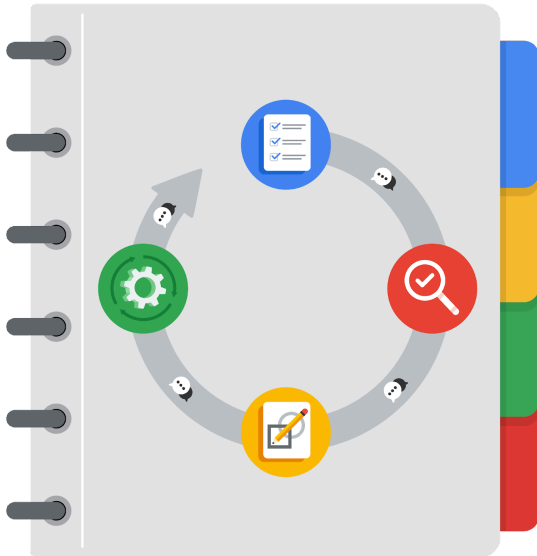
- What resources are required to complete this project?

  Project dataset, Python notebook, and input from stakeholders.

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

Dataset scrubbed for exploratory data analysis, visualizations, statistical model, regression 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?

   It has to meet the client expectations and goals after the model is constructed

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

Why did you select these stages for this task?

Analyze stage allows for hypothesis testing while construct stage allows us conduct testing of the model

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

Why did you select this stage for this task?

This stage allows us to get a better understanding of the data.

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

Why did you select these stages for this task?

Planning allows us to understand the goals of the project while analysis allows us to clean and explore the data according to our planning

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

Why did you select this stage for this task?

It allows us to understand the goals of the project and improve workflow

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

Why did you select this stage for this task?

It is done in the final stage of PACE.

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

Why did you select this stage for this task?

To better understand the dataset

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

   Why did you select these stages for this task?

   > Visualisation begins in analyse and is used to construct

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

   Why did you select this stage for this task?

   > Planning as it is the first stage

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

    Why did you select this stage for this task?

    > Analysis to ensure it follows our project goals and model is build in the construction page

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

    Why did you select this stage for this task?

    > It allows for an easy understanding of data for further analysis or construction of the data.

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

    Why did you select this stage for this task?

    > The building of a data model would take place in the construct stage.