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.

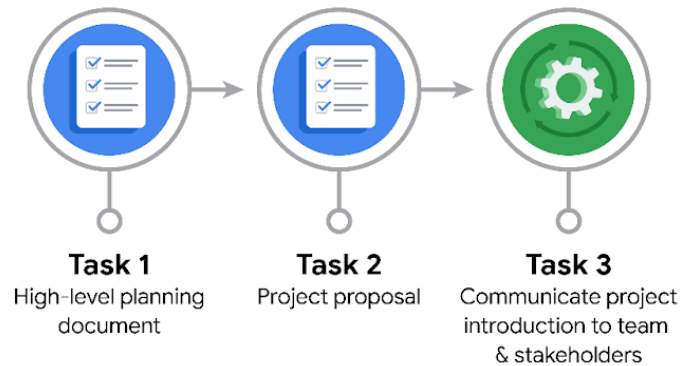☐ Create a project proposal for the data team.

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?

Automidata scenario: New York City Taxi and Limousine Commission.

TikTok scenario: TikTok data team and cross-functional team members.

Waze scenario: Waze data team and cross-functional team members.

- 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?

Automatidata Scenario: Estimate taxi fares before the ride based on different variables.

Tiktok Scenario: Using a predictive model to determine whether a video contains a claim or it offers an opinion.

Waze Scenario: Working to develop a machine learning model to predict user churn.

- What questions need to be asked or answered?

1) How is the provided dataset?
2) What variables will we need?
3) Are there any trends that can provide us insight?

4) How will I reduce bias?

- What resources are required to complete this project?

Data set
Python
Subject matter expert

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

- Project document - Familiarizing with the data - EDA - Visualizations - Statistical Model - CDA - Regression Analysis - Machine learning model - A/B testing exploratory data analysis, visualizations, statistical model, regression analysis and/or 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?

   > After running the data through the model we can check if it meets project requirements hence evaluate the model.

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

   Why did you select these stages for this task?

   > Analyze stage: at this stage, the statistical test to be executed is determined
   >
   > Construct stage: Execution of the test is a part of construct stage

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

   Why did you select this stage for this task?

   > We Understand data and information within it.

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

    Why did you select these stages for this task?

    > Data exploration is a part of the planning stage.
    >
    > Cleaning is done in the analyze phase.

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

    Why did you select this stage for this task?

    > Planning is when you create the PACE strategy document and establish structure for project workflow

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

    Why did you select this stage for this task?

    > Communicating final insights with the stakeholders is essential in the execution of the project.

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

    Why did you select this stage for this task?

    > Descriptive statistics is used to understand the statistics of data during analysis

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

    Why did you select these stages for this task?

We understand data better using visualizations and hence analyze it, making these visualizations is a part of construct phase.

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

Why did you select this stage for this task?

Project Proposal is a part of the planning stage at it is the initial document that explains the project.

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

Why did you select this stage for this task?

A regression model is made in the construction stage and understanding this model is a part of analyze stage.

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

Why did you select this stage for this task?

Summarising data means to better understand it and we do it in analyze stage where we analyse its key features.

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

Why did you select this stage for this task?

Building of the machine learning model will be a part of construct phase.