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?

The estimation of the taxi fares(process or booking rates) based on the relevant factors that we will identify.

- What questions need to be asked or answered?

What is the condition of the provided dataset?
Are the trends that are present in the data enough to provide insights?
What factors will be most useful?

- What resources are required to complete this project?

Project Dataset and Python(Jupyter) notebook.

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

Dataset for EDA (exploratory data analysis), visualizations, regression odel and analysis, 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?

Data is run through to evaluate whether it meets the project's expectations and goals.

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

   Why did you select these stages for this task?

In the analyzing stage, it is determined that the statistical test will be used. And, during the construction phase, the test is carried out.

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

   Why did you select this stage for this task?

Analyzing phase helps us gain the deeper understanding of the dataset and the information that is present within it.

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

   Why did you select these stages for this task?

Planning starts when choices are made about the methods and strategies that are needed. The cleaning process then takes place in the analyzing stage.

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

    Why did you select this stage for this task?

    > Creating an initial project PACE document defines the workflow and helps to plan how to best plan the strategy to approach the project.

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

    Why did you select this stage for this task?

    > Final insights are very important to be communicated to the stakeholders.

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

    Why did you select this stage for this task?

    > Investigating the statistics within the data takes place during the analysis phase

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

    Why did you select these stages for this task?

    > We analyze the data thoroughly and then move towards visualization aka construction phase.

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

    Why did you select this stage for this task?

    > A project proposal is aprojets plan and layout.

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

    Why did you select this stage for this task?

    > **The model is examined in detail and then the regression model is built according to its need.**

11. Compile summary information about the data: Analyze ▾

Why did you select this stage for this task?

Inspect a dataset to compile all the necessary information for the summary.

12. Build machine learning model: Construct ▾

Why did you select this stage for this task?

Building of a data model would occur in the construction stage.