

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: Plan Stage

- Who is your audience for this project?

Customers who are using New York City Taxi and Limousine Commission (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?

We want to estimate the fare before the customer requests a taxi or limousine.

- What questions need to be asked or answered?

What are the variables that affect the cab fare the most?

What is the condition of the dataset?

How can we reduce bias from the data?

- What resources are required to complete this project?

Data provided from TLC.

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

PACE Strategy Document and the Project Proposal

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?

Once you finish the data processing, you need to evaluate if your model fulfill with the goals.

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

Why did you select these stages for this task?

In the analysis phase, you determine the statistical model that you will use. In the construct phase, you perform several tests to see if the results are sufficient.

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

Why did you select this stage for this task?

In the analyze phase, you begin to collect the data and see if data formatting and cleaning is necessary.

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

Why did you select these stages for this task?

In the plan phase, you need to be familiar with your data to make the questions of your project. In the analyze phase, you need to format and clean your data to work with your model.



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

Why did you select this stage for this task?

It's the stage where you establish milestones and deliverables for each stage.

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

Why did you select this stage for this task?

Once you analyze the results on how the variable affect the cab fare, it's time to present your results to the stakeholders

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

Why did you select this stage for this task?

In this phase you select the statistical model that you are going to work with.

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

Why did you select these stages for this task?

In the analyze phase, you create quick visualizations to gain quick insights about your data. In the construct phase, you make visualizations of the final results or test.

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

Why did you select this stage for this task?

It's where you establish the goals, the milestones, the time for each stage and so on.



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

Why did you select this stage for this task?

In the analyze phase, you perform a general data analysis to know the model that works for your goals. In the construct phase, you create now with the model to obtain the results you want.

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

Why did you select this stage for this task?

You make a summary to have a raw idea of the data you work with and select which model works best.

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

Why did you select this stage for this task?

It's the phase where you work with Python.