**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?

Data Science Lead and Manager, Project management officer and Operations Lead

* 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?

Correctly classify comments between complaints or opinions. The impact of this would be on the policies of the content that is permitted to be published in the platform and general recommendations for content creators to improve engagement.

* What questions need to be asked or answered?

Is there a specific region or category of content that needs to be prioritized for the project?

What is the definition of complaint and opinion?

Do we have limitations to considered in regards to the data access, like privacy restrictions?

* What resources are required to complete this project?

Access to the data sources, and access to meetings with stakeholders to validate progress and collect feedback.

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

The plan itself, the initial EDA of the data to validate which data to use in the model, the ML 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?

Run data through the model to see if it works.

1. **Conduct hypothesis testing: Analyze** **and** **Construct**

Why did you select these stages for this task?

In analyze decide the statistical test. During construction the test is carried out

1. **Begin exploring the data: Analyze**

Why did you select this stage for this task?

EDA

1. **Data exploration and cleaning: Plan** **and Analyze**

Why did you select these stages for this task?

Decide about EDA and carry it out

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

Why did you select this stage for this task?

Defining tasks before doing

1. **Communicate final insights with stakeholders: Execute**

Why did you select this stage for this task?

Final step is to communicate

1. **Compute descriptive statistics: Analyze**

Why did you select this stage for this task?

EDA

1. **Visualization building: Analyze and Construct**

Why did you select these stages for this task?

Visualization begins with data assessment and is created during the construction stage.

1. **Write a project proposal: Plan**

Why did you select this stage for this task?

Is the very beginning of the project

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

Why did you select this stage for this task?

Analyze the model to make sure it fits the scenario, then build it.

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

Why did you select this stage for this task?

EDA

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

Why did you select this stage for this task?

After deciding on the model, build it in construction phase