**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?

TikTok data team

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?

To generate a ML model to automate the classification of user reports as claims or opinion. The model should be able to fasten the process and reduce backlogs.

* What questions need to be asked or answered?

What all data should be consider for the model creation and execution?

What is the story that can be made from the data?

Which testing and analysis strategy will be the best choice?

* What resources are required to complete this project?

The data for the model.

Human resources like Data scientists.

Tools/ software for the project.

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

Stakeholder updations.

Dataset.

Documentations & Visualizations

ML mode / Regression modell.

## 

## **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 the model is constructed and executed it needs to be evaluated by stakeholders to check if the performance is up to expectations.

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

Why did you select these stages for this task?

During analysis its prepared for testing and as well as during construction to see if all the model is working on the sample data provided at each stage of its construction.

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

Why did you select this stage for this task?

Data needed for the project is gathered from multiple sources for analysis and other testing and statistical types are planed.

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

Why did you select these stages for this task?

Data is gathered and cleansed for the model construction.

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

Why did you select this stage for this task?

It's the most crucial step for the seamless flow of the project.  
Details about the tools that should be used and the testing methods etc are discussed during the planning stage.

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

Why did you select this stage for this task?

Communication of final insights is the culmination of the project's work and is where the results are delivered and put into action. It's the step where you "execute" the delivery of the project's outcomes to the people who will use them to make decisions.

1. **Compute descriptive statistics: Analyze**

Why did you select this stage for this task?

Descriptive statistics are used to summarize and describe the main features of a dataset. They provide a concise overview of the data without making inferences about a larger population.

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

Why did you select these stages for this task?

When initially exploring the data to understand its patterns, distributions, and relationships, visualization is a key part of the Analyze phase.  
When creating visualizations to communicate final insights and results to stakeholders, this is part of the Execute phase.

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

Why did you select this stage for this task?

A project proposal gives a brief idea about the tasks and milestones that need to be achieved.

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

Why did you select this stage for this task?

The analyzed data is taken and the solution model is created and tested.

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

Why did you select this stage for this task?

This stage is all about understanding the data. It helps to get a grasp of the data's characteristics, identify any potential issues, and inform subsequent analysis.

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

Why did you select this stage for this task?

This is the solution to the problem or the project goal. The model is constructed to perform the task of classification.