**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?

All people using tiktok

* 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?

 Build a machine learning model that classifies whether a TikTok video contains a **claim** or an **opinion**.

 Anticipated impact: Reduce backlog of user reports, prioritize moderation, improve safety and trust on the platform.

* What questions need to be asked or answered?

 What features in the data help distinguish claims vs opinions?

 How reliable and accurate is the model in classifying reports?

 What assumptions must be tested in regression modeling?

 How will success be measured (accuracy, recall, F1-score)?

* What resources are required to complete this project?

 Historical TikTok user report dataset.

 Python environment for analysis and modeling.

 Statistical and ML libraries (pandas, scikit-learn, etc.).

 Visualization tools (Tableau, matplotlib).

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

 Project plan with milestones.

 Cleaned dataset ready for EDA.

 EDA summary report and descriptive statistics.

 Hypothesis testing results.

 Regression + machine learning model outputs.

 Visualizations for executives.

 Final report and presentation

## 

## **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](https://www.coursera.org/learn/foundations-of-data-science/supplement/4OtHr/the-pace-stages) and [Communicate objectives with a project proposal](https://www.coursera.org/learn/foundations-of-data-science/supplement/79Ysh/communicate-objectives-with-a-project-proposal). You will later reorder these tasks within a project proposal.

1. **Evaluating the model:** Analyze

Why did you select this stage for this task?

Requires testing performance (Construct) and verifying success for deployment (Execute).

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

Why did you select these stages for this task?

Data testing begins with analysis of variables (Analyze) and informs regression modeling (Construct).

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

Why did you select this stage for this task?

Early exploration identifies patterns, anomalies, and possible features.

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

Why did you select these stages for this task?

Data cleaning = analysis, while preparing features = construct for modeling.

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

Why did you select this stage for this task?

Workflow setup belongs to the planning stage.

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

Why did you select this stage for this task?

Final deliverables and presentations to TikTok executives.

1. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

Descriptive stats summarize dataset and guide modeling.

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

Why did you select these stages for this task?

Supports EDA internally (Analyze) and executive communication (Execute).

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

Why did you select this stage for this task?

Proposal = project planning deliverable.

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

Why did you select this stage for this task?

Construct = model building; Analyze = validate results.

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

Why did you select this stage for this task?

Part of EDA and data understanding.

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

Why did you select this stage for this task?

Model development is part of construct phase.