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?

Mary Joanna Rogers is the project manager and will be the one who will deliver the project upon completion of the project. The secondary audience is the Tik Tok executives who will be receiving the visualizations of our project.

- 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 predict when a content creator makes a claim that needs to be reviewed by the moderators. Our goal is to keep our community safe and inclusive and the use of moderators time is limited due to the inflow of content.

- What questions need to be asked or answered?

What constitutes a claim?

What claims create the most disruption to the safety and inclusivity of our community?
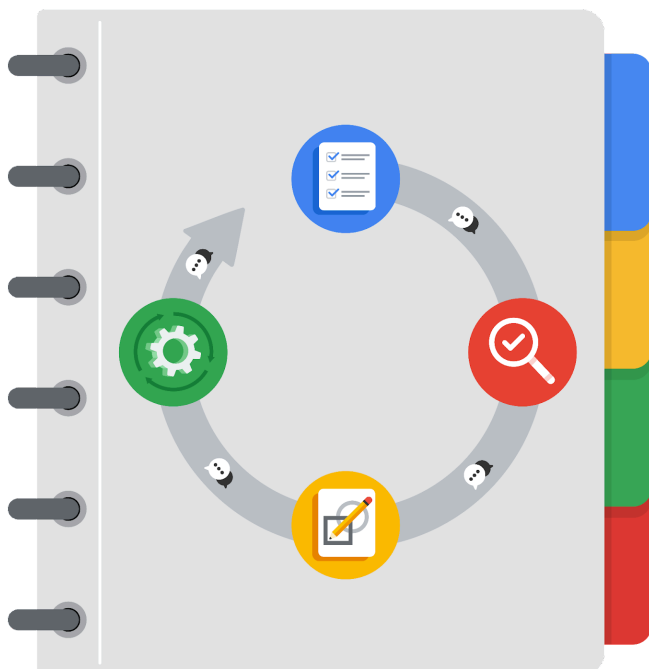
What is an opinion?

- What resources are required to complete this project?

I am given willow jaffey, Rosie Mae Bradshaw, and Orion Rainier, Mary, Margery Adebowale, and Maika Abad

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

A machine learning model to filter the data before it arrives at the moderators.

A report to the stakeholders of Tic Tok that contains visualization for simplification of complex data science.

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?

    The model has already been assembled, and it needs to be tested or iterated though to produce the most reliable tool yet.

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

    Why did you select these stages for this task?

    We have built most of our ideas now its time to build the code necessary for disproving our hypothesis

    After that we use data to select the best set of ideas to start working on developing the machine learning model.

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

    Why did you select this stage for this task?

After preparing, It time to see the data in action to determine with visualizations

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

Why did you select these stages for this task?

Every data project has a phase where we need to make the data set accessible for use in the project to prevent code errors from improperly formatted data.

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

Why did you select this stage for this task?

The planning phase is about constructing a work flow plan to efficiently solve the business task.

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

Why did you select this stage for this task?

The team has completed the project and is ready to bring it to stakeholders.

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

Why did you select this stage for this task?

These statistics help us build our understanding of the data to generate insights.

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

Why did you select these stages for this task?

> After analyzing the data we need to build the visualizations to increase our analytical prowess and discover the insights.

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

Why did you select this stage for this task?

> Before any work begins it's often best to consider a set of valid plans to improve the decision making process. There might be many valid project proposals to be chosen from.

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

Why did you select this stage for this task?

> After analyzing, constructing a mathematical model needs to be constructed based on the data.

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

Why did you select this stage for this task?

> The information helps us inform the processes needed to analyze the data to most efficiently approach the task.

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

Why did you select this stage for this task?

> With all the data analyzed, the model constructed, it needs to be built by the process of iteration.