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.~~

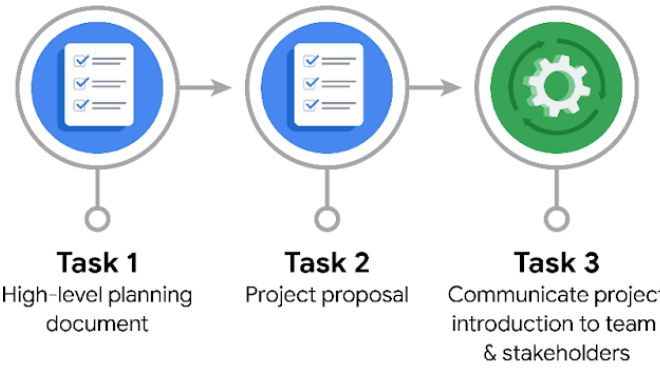
Relevant Interview Questions

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?

  - Internal Stakeholders:
  - Data Analysis Manager (May Santner)
  - Director of Data Analysis (Harriet Hadzic)
  - Senior Data Analyst (Chidi Ga)
  - Senior Project Manager (Sylvester Esperanza)
  - External Stakeholders:
  - Waze Leadership Team
  - Finance and Administration Department Head (Emrick Larson)
  - Operations Manager (Ursula Sayo)

- 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?

  - Objective: Develop a machine learning model to predict user churn on the Waze app.
  - Impact: The model will help Waze prevent churn, improve user retention, and increase growth. By identifying factors contributing to churn, Waze can proactively engage users at risk of churning, enhancing the overall user experience and supporting business decisions for product development and marketing strategies.

- What questions need to be asked or answered?

  - Who are the users most likely to churn?
  - Why do users churn?
  - When do users churn?
  - What factors contribute the most to user churn?
  - How can Waze proactively engage users at risk of churning?

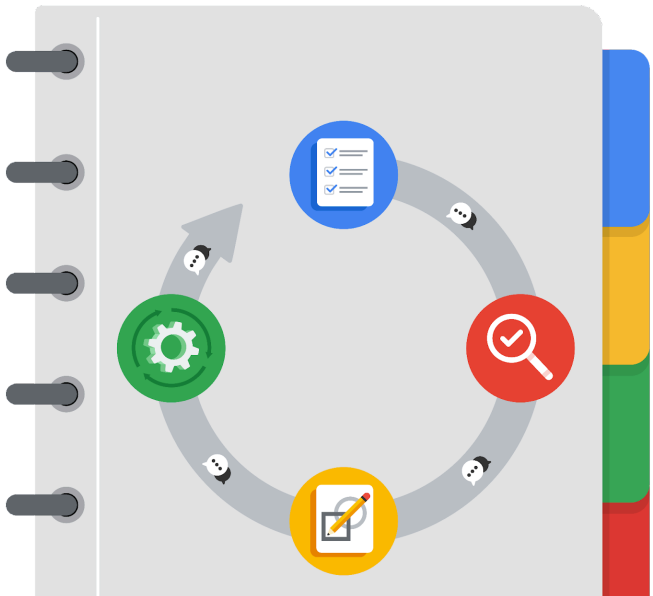
- What resources are required to complete this project?

  - Historical user data from Waze app
  - Data analysis and machine learning tools (e.g., Python, Tableau)
  - Access to Waze's data infrastructure and database
  - Collaboration with Waze's data analysis team and other relevant departments

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

  - Project proposal
  - Workflow structure document
  - Cleaned data files
  - EDA report
  - Tableau dashboard/visualizations
  - Hypothesis testing report
  - Regression model
  - Model evaluation report
  - Final machine learning model
  - Final report to all stakeholders

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?

Evaluation is the final step to assess the performance and effectiveness of the model, ensuring it meets project requirements.

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

   Why did you select these stages for this task?

Analyzing relationships between key variables and constructing tests to validate hypotheses are essential for data-driven decision-making.

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

   Why did you select this stage for this task?

Initial exploration helps identify the available data and any immediate issues before detailed analysis.

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

   Why did you select these stages for this task?

Cleaning and preparing the data ensure quality and reliability, making it ready for analysis and model building.

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

   Why did you select this stage for this task?

Defining the workflow structure ensures a clear path with organized stages and tasks.

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

   Why did you select this stage for this task?

Sharing final insights and model results with stakeholders is crucial for informed decision-making and project closure.

7. **Compute descriptive statistics:** Analyze

   Why did you select this stage for this task?

Generating descriptive statistics provides a summary of the data, highlighting key patterns and trends.

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

   Why did you select these stages for this task?

Creating visualizations helps understand the data better and communicate insights effectively.

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

   Why did you select this stage for this task?

Outlining the project goals, milestones, and deliverables creates a comprehensive plan for the project.

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

    Why did you select this stage for this task?

Developing a regression model helps in predicting user churn by understanding relationships between variables.

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

Why did you select this stage for this task?

Summarizing the data provides a clear overview of the dataset's characteristics and informs further analysis.

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

Why did you select this stage for this task?

Constructing the final machine learning model is the culmination of the analysis aimed at accurately predicting user churn.