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?

As a new member of the team, I would ask for the following:
a) Methodology and documentation practices of the company, to learn the rules of the game
b) Any documentation generated so far, related to the project
c) 1-on-1 session with Analytics lead and or project manager to get an overview idea of the project
d) I would ask to participate in work sessions with our customer and the analytics team, and progress update sessions with stakeholders

- How would you plan an analytics project?

So far, I am certified in two methodologies, for simple small projects I would use the Data Workflow Analysis Cycle: **Ask, Prepare Process, Analyze, Share Act**.
For larger dataset and more complex projects I would use the PACE (**Plan, Analyze, Construct and Execute**) workflow methodology because it is a completed circuit that offers freedom and flexibility in a simpler way and also because it is oriented at building models for processing or analysis of the data. This methodology is also more suited to incorporate new information and feedback which can be

incorporated at any part of the process and streamlined communication between all phases that allows me to go back to return to analyzing phase to clarify any aspect.

- What steps would you take to translate a business question to an analytical solution?

  You need to **clearly define the business problem**. This involves understanding the context of the question, identifying the stakeholders involved and determining the objectives of the analysis. Once you have the definition, you can use the information to prepare a proposal and if approved, proceed to the elaboration of the solution under the preferred and agreed methodology

- Why is actively managing data an important part of a data analytics team's responsibilities?

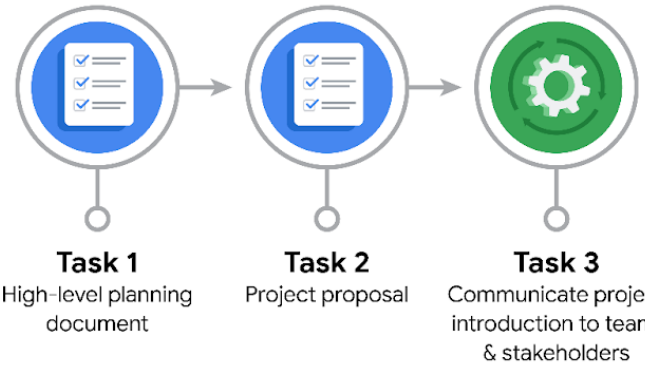
  Because data is a valuable asset that needs protection and compliance. Also, data needs to be generated, processed, combined, aggregated and shared between multiple individuals and applications to be of value. The data analytics team must always be aware of the data the organization owns and its state to be able to work with it to answer different questions and business/model requirements

- What are some considerations you might need to be mindful of when reporting results?

  The audience, their familiarity with analytics terminology, the proper context and detail level, the time available for the presentation or reporting and the following up on questions post presentation
  Make sure you reveal important insight from the data analysis

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?

Waze Leadership

- 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 predicts user churn. The model will be based on a sample of Waze user data. An accurate model will help our company prevent user churn, improve user retention, and increase growth.

- What questions need to be asked or answered?

What is the most useful information, variables or fields in our data?
Who are the users most likely to churn?
Why do users churn?

When do users churn?

- What resources are required to complete this project?

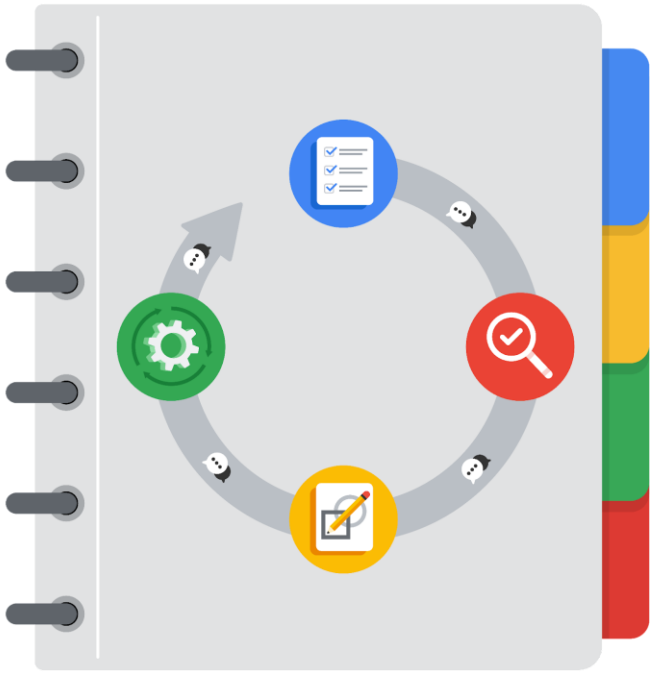
Technical Resources:
-Project dataset
-Python Notebook
-stakeholder's input

People:
Data team:
Harriet Hadzic - Director of Data Analysis
May Santner - Data Analysis Manager
Chidi Ga - Senior Data Analyst
Sylvester Esperanza - Senior Project Manager
Other areas:
Emrick Larson - Finance and Administration Department Head
Ursula Sayo - Operations Manager

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

-Project Proposal
-Project Plan
-Visuals, Dashboard
-Regression Model
-ML Model
-Results

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?

After the model has been constructed, data is run through to evaluate whether it meets the project's expectations and goals.

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

   Why did you select these stages for this task?

   During the analyzing stage, it is determined that a statistical test will be used. During the construction phase, the test is carried out.

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

   Why did you select this stage for this task?

   In the analysis phase we prepare the data for the project, this includes exploratory data analysis to convert data into a usable format. During the analysis phase, you will gain a deeper understanding of the dataset and the information within it.

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

   Why did you select these stages for this task?

   Planning takes place when you first make choices about the methods needed. The cleaning process then takes place in the analyzing stage.

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

   Why did you select this stage for this task?

   In the plan phase we identify needs, goals and strategies. We will take an inventory of the project and the tasks required. Creating an initial project PACE document outlines the workflow and helps to plan how to best approach a project.

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

   Why did you select this stage for this task?

   Execute is the last phase of the project, where we share the result of the analysis and collaboration with stakeholders, as well as value unlocked from the data.

7. **Compute descriptive statistics:** Analyze

   Why did you select this stage for this task?

   Investigating the statistics within data takes place during analysis.

   In the analyze phase is where we compute descriptive statistics to get knowledge and understanding of our data, some of the values we can gather are central tendency, dispersion and data visualization. By understanding the data in the Analyze phase, we provide valuable insights into the data and guiding subsequent modeling efforts.

8. **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.

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

   Why did you select this stage for this task?

   Planning stage. A project proposal is the initial document used to define a project.
   In the planning phase we define the scope of the project, the goals and the strategies, this is where we get the approval and or feedback from our stakeholders to proceed with the project.

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

    Why did you select this stage for this task?

    During the analyzing stage, the model is examined in detail to be sure it will meet the needs of the task.

    Building a regression model takes place in the Construct phase, where we build models to uncover relationships within the data and conduct statistical inference about those relationships.

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

    Why did you select this stage for this task?

    Inspecting a dataset to compile information would take place in the analysis phase.

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

Why did you select this stage for this task?

The construct phase focuses on the actual creation of the model based on the insights gained during the Analyze phase.

The building of a data model would take place in the construct stage.

Disclaimer:

This document is intended for educational and didactic purposes only and should not be considered professional advice. The information contained herein is based on the author's understanding of the subject matter, drawing upon knowledge acquired through coursework and utilizing large language models (LLMs) such as Google Gemini and Microsoft Copilot as research and writing aids.