**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?

Waze Data Team and cross-departmental 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?

Develop a machine learning model to predict user churn. Better user retention rates and overall growth of the company is the larger need of the client which is Waze.

* 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 resources are required to complete this project?

Dataset, Python notebook, shareholder’s input.

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

Analyzed dataset, visualizations and a machine learning model.

## 

## **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:** Execute

Why did you select this stage for this task?

After the model has been constructed, the testing data is run on the model to ensure it fits the stakeholder’s needs and expectations.

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

Why did you select these stages for this task?

A Statistical test is selected in the **analyze** stage and carried out in the **construct** stage.

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

Why did you select this stage for this task?

It is during the **analyze** phase we scrub through the data to develop a deeper understanding of the dataset.

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

Why did you select these stages for this task?

In the **planning** phase, we identify and decide upon the methods and process of data cleaning and it is carried out in the **analyze** phase

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

Why did you select this stage for this task?

Creating a PACE document is the first and primary task which outlines the project tasks and activities for an easy and smooth workflow.

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

Why did you select this stage for this task?

Final insights are communicated to the stakeholders after the data has been scrubbed and a model is ready in this case.

1. **Compute descriptive statistics:** Analyze

Why did you select this stage for this task?

Descriptive statistics is computed in the early stages of the **analyze** phase

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

Why did you select these stages for this task?

Visualizations begin in the analyze stage where primary assessment of data is conducted and is carried out extensively in the construct stage.

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

Why did you select this stage for this task?

A Project Proposal is the initial document which specifies the purpose and definition of a project.

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

Why did you select this stage for this task?

Data is prepped and model is examined to meet the needs of the project in the **analyze** phase whereas the building of the model takes place in the **construct** phase.

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

Why did you select this stage for this task?

Summary information of a data is performed in the analyze stage before the EDA process starts.

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

Why did you select this stage for this task?

The building of the machine learning model takes place in the construct phase.