**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?

Members of the data team

Members of the TLC team

* 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 build a regression model to estimate the taxi fare based on the distance, time of the day, and other variables.

Having known the taxi fare beforehand, the raiders of the TLC can get into the cab confidently, knowing that they're paying a fair price, eventually improving the trust the customers have in TLC

* What questions need to be asked or answered?

What’s the accuracy of the prediction model be.

How complete and clean is the data

* What resources are required to complete this project?

What software/programming languages to be used.

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

Project proposal

EDA report

Model, and Insights

Final reports

## 

## **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?

Once the model is constructed, it will be evaluated with data at the Execute state

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

Why did you select these stages for this task?

During the Analysis phase the type of hypothesis testing will be identified and in the Construct stage the hypothesis testing will be carried out.

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

Why did you select this stage for this task?

During this phase, EAD will be carried out to get an understanding of the data.

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

Why did you select these stages for this task?

During the planning stage, the methods needed for the data exploration are discussed and determined, and the data cleaning is done at the Analyze state.

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

Why did you select this stage for this task?

Establishing the project workflow is done at the Planning stage

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

Why did you select this stage for this task?

Communicating the final insights with the stakeholders is the last done at the Executing stage

1. **Compute descriptive statistics: Analyze**

Why did you select this stage for this task?

EDA and Calculating statistics need to be performed at the Analyze stage to get data insights

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

Why did you select these stages for this task?

The analysis performed at the Analyze stage is visualized at the Construct state via dashboards. Thus, the visualization-building begins at the Analyze stage and extends to the Construct stage

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

Why did you select this stage for this task?

During this stage, the document that defines the project the project proposal is made to identify the milestones, the deliverables, and appropriate stakeholders for each stage of the project.

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

Why did you select this stage for this task?

During the analysis phase, the initial groundwork for a model building, like EDA, getting insight into the relationship between data is laid out,, and during the construct phase,, the regression model is built.

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

Why did you select this stage for this task?

Summary information about the data is the final document you get after analyzing the data

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

Why did you select this stage for this task?

Building the model will be done at the Construct stage.