**Course One**

# Foundations of Data Science



# Instructions

Use this PACE strategy document to record your decisions and reflections as you work through this end-of-course project. As a reminder, this document is a resource that you can reference in the future and a guide to help consider responses and reflections posed at various points throughout projects.

# Course Project Recap

Regardless of which track you have chosen to complete, your goals for this project are:

* Understand and assess the proposed scenario
* Demonstrate foundational knowledge of the data science workflow - PACE
* Articulate a data project proposal in the planning stage for cross-functional team members

# 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: Planning Stage**

* Who is your audience for this project?

New York City Taxi and Limousine Commission (TLC)

* 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 regression model that predicts ride durations based on many variables such as distance, time of day, season, and so on. This model will help the New York City Taxi and Limousine Commission to update its policies, and maybe hire more taxi drivers for a specific time depends on ride need.

* What questions need to be asked or answered?

What’s the outline of the project?

Do we need a model? If yes, What’s the input and output of that model?

How the model prediction will help the client?

Witch variables we need in that model? There’s any relationship between them?

* What resources are required to complete this project?

In this project we need the ride Data, Computer with high performance, Programming Language such as R or Python, SQL, visualization software like Tableau, PowerBI.

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

-The summary of model of prediction

-The steps of choosing the model

-The visualization explaining the results

-The executive Part or Prototype

## **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: Automatidata](https://www.coursera.org/learn/foundations-of-data-science/supplement/XxgHa/course-1-end-of-course-portfolio-project-overview-automatidata) 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:** Construct

Why did you select this stage for this task?

After we construct the model, we need to test it and evaluate it and execute then we get the good results in evaluating.

1. **Conduct a hypothesis test:** Construct **and** Analyze

Why did you select these stages for this task?

With hypothesis test, we can make sure that the model is ready for execute stage, and in analyzing data, make hypothesis test of the data quality before start using it in analyzing stage.

1. **Understanding the data:** Plan

Why did you select this stage for this task?

Understanding the data is the one of the important steps to do in the beginning of any data science project, because that will help us to make a plan of the project by defining the problem and propose a solution.

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

Why did you select these stages for this task?

In the stage of analyze we interact with data, especially organize and clean it and in the stage of construct we make sure that the data is very explored if not we do that again, like correct any error in data.

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

Why did you select this stage for this task?

Before start any project, we have to establish a structure for project workflow.

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

**Communicate final insights with stakeholders is the last in the PACE process then it available in the execute stage.**

1. **Compute descriptive statistics:** Plan

Why did you select this stage for this task?

Compute descriptive statistics must be in the plan stage.

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

Why did you select these stages for this task?

Visualization building can be shown in the Construct phase to make sure of the result or in the Execute stage.

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

Why did you select this stage for this task?

Plan stage is stage where we have to make a project proposal for working with it along the project.

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

Why did you select this stage for this task?

**Build a regression model must be done in the Construct phase and it can be updated in the execute phase.**

1. **Inspect the data set for missing data:** Analyze

Why did you select this stage for this task?

**All inspection of the quality of the data is task in the analyze stage.**

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

**The build of any model is done the phase of modeling or construct stage.**