**Course One**

# Foundations of Data Science



**PACE Strategy Document**

**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?

The audience for this project is the New York City 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?

The goal is to create a regression model that predicts ride durations based on distance, time of day, season, and additional variables as necessary. This will help clients in optimizing driver schedules, routes and pricing. Targeted marketing and promotions and can also be applied.

* What questions need to be asked or answered?

What data do we have access to?

What assumptions are we making?

What variables are most useful?

Are there any specific questions that need to be answered through analysis of the data?

What methods and tools will be used to analyze the data?

What steps can I take to reduce the impact of bias?

* What resources are required to complete this project?

Access to TLC data on rides.

Expertise to analyze the data.

Visualization tools.

Input from stakeholders.

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

Throughout the project, the stakeholders will be updated on the following:

Formatted database for EDA

Machine learning / regression model

Visualizations

## **THE PACE WORKFLOW**



**[Alt-text: The PACE Workflow with the four stages in a circle: plan, analyze, construct, and execute.]**

### **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?

The model is evaluated in the last stage to check if meets the project’s expectations and goals.

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

Why did you select these stages for this task?

The analysis stage involves any hypothesis tests. During the construction phase, the test is carried out.

1. **Understanding the data: Analyze**

Why did you select this stage for this task?

Analyzing stage. When cleaning and formatting, a deeper understanding of the data can be gained.

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

Why did you select these stages for this task?

Data research and exploration is done in the planning stage and cleaning in analysis.

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

Why did you select this stage for this task?

It should be done in the initial planning stage.

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

Communication is necessary at various points throughout a project. Final insights are shared with stakeholders in the execute phase of the data project workflow.

1. **Compute descriptive statistics: Construct**

Why did you select this stage for this task?

Investigating the statistics within data takes place during the analysis process.

1. **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.

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

Why did you select this stage for this task?

A project proposal is the initial document used to define a project.

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

Why did you select this stage for this task?

The regression model is first constructed. Then, during the analyzing stage the model is examined in detail to be sure it will meet the needs of the task.

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

Why did you select this stage for this task?

Inspecting a data set for missing data would take place while assessing the quality of your data.

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

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