## Reference Guide
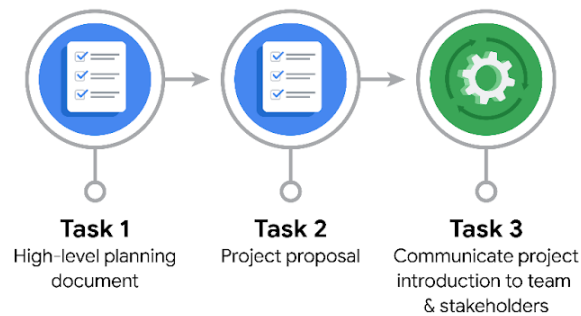
This project has three tasks; the following visual identifies how the stages of PACE are incorporated across those tasks.



**Task 1**
High-level planning document

**Task 2**
Project proposal

**Task 3**
Communicate project introduction to team & stakeholders

## Data Project Questions & Considerations



**P**ACE: **Plan Stage**

- Who is your audience for this project?

  *Y*our audience will be the New York City Taxi and Limousine Commission.

- What are we trying to solve or accomplish? And, what do you anticipate the impact of this work will be on the larger needs of the client?

  *T*he problem you are trying to solve is the estimation of taxi fares based on relevant variables that you identify.

- What questions need to be asked or answered?

  What is the condition of the provided dataset? What variables will be the most useful? Are there trends within the data that can provide insight? What steps can I take to reduce the impact of bias?

- What resources are required to complete this project?

  *We* will need the project dataset, Python notebook, and input from stakeholders.

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

> *T*he deliverables include *a* dataset scrubbed for exploratory data analysis, visualizations, statistical model, regression analysis and/or machine learning model.

## Project tasks

1. **Evaluating the model:** `Execute ▾`

> 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 ▾`

> 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 ▾`

> 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 ▾`

> 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 ▾`

> Planning stage. 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 ▾`

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

7. **Compute descriptive statistics:** `Analyze ▾`

Investigating the statistics within data takes place during analysis.

8. **Visualization building:** `Analyze ▾` **and** `Construct ▾`

Visualization begins with data assessment and is created during the construction stage.

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

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

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

During the analyzing stage, the model is examined in detail to be sure it will meet the needs of the task. The building of the regression model will take place in the construction phase.

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

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

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

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