PACE strategy document

Introduction

This PACE strategy document will be used to record decisions and reflections as we work through this project. We can use this document as a guide to consider ourresponses and reflections at different stages of the data analytical process.

Data Project Questions & Considerations

 PACE: **Plan Stage**

- Who is your audience for this project?

New York City Taxi and Limousine Commission (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?

Developing a regression model that helps estimate taxi cab fares before the ride based on relevant variables. The clients will be satisfied as they could know taki fares before they take the ride.

- What questions need to be asked or answered?

What is the condition of the provided dataset? What features/variables will be useful? Are there trends or seasonality within the dataset that can provide insights? What steps can be taken to reduce the impact of bias?

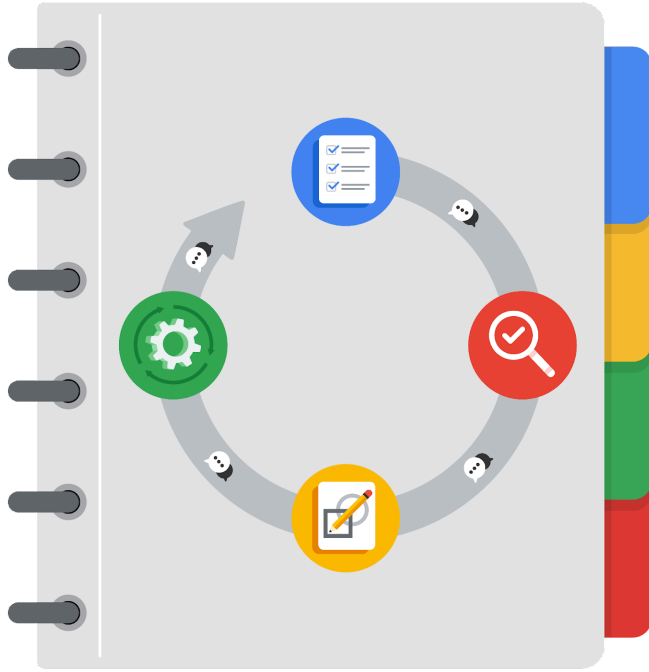- What resources are required to complete this project?

The historical dataset, Jupyter notebook, and stakeholders updates.

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

The deliverables include, a dataset scrubbed for exploratory data analysis, visualizations, statistical model, regression analysis, and machine learning model.

THE PACE WORKFLOW

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

I have been asked to demonstrate for the company's data team how I would use the PACE workflow to organize and classify tasks for the upcoming project.

Project tasks

Following are a group of tasks my company's data team has determined need to be completed within this project. The data analysis manager has asked my to organize these tasks in preparation for the project proposal document. First, identify which stage of the PACE workflow each task would best fit. Next, give an explanation of why I selected the stage for each task.

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

   Why did you select this stage for this task?

   After the model has been constructed, the model needs to be evaluated with new data to see whether the model meets the project's expectations and goals.

2. **Conduct hypothesis testing:** `Analyze ▾` and `Construct ▾`

   Why did you select these stages for this task?

   > During the analyzing stage, we conduct hypothesis testing to determine which statistical test will be used. During the construction phase, those tests are carried out.

3. **Begin exploring the data:** `Analyze ▾`

   Why did you select this stage for this task?

   > During the analysis phase, we explore the data to gain a deeper understanding of the dataset and the insights within it.

4. **Data exploration and cleaning:** `Plan ▾` and `Analyze ▾`

   Why did you select these stages for this task?

   > Planning takes place when we 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 ▾`

   Why did you select this stage for this task?

   > In planning stage, we create 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 ▾`

   Why did you select this stage for this task?

   > 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 ▾

Why did you select this stage for this task?

Investigating and exploring the statistics within data takes place during analysis stage.

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

Why did you select these stages for this task?

Visualization is necessary to begin with data exploring and it is created during the construction stage.

9. **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.

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

Why did you select this stage for this task?

During the analysis stage, the model is examined in detail to ensure it will meet the needs of the task. Building the model takes place in the construction phase.

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

Why did you select this stage for this task?

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

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

Why did you select this stage for this task?

Building the data model takes place in the construct stage.