**Automatidata Project Proposal**

## **Overview**

The New York City Taxi and Limousine Commission (TLC) seeks a way to utilize the data collected from the New York City area to predict the fare amount for taxi cab rides.

|  |  |  |  |
| --- | --- | --- | --- |
| **Milestones** | **Tasks** | **Deliverables/Reports** | **Relevant Stakeholder (Optional)** |
| **1** | **Establish structure for project workflow (PACE)Plan** | * Global-level project document | Deshawn Washington — Data Analysis Manager |
| **1a** | **Write a project proposal**  **Plan** |  | Uli King — Senior Project Manager |
| **2** | **Compile summary information about the dataAnalyze** | * Data files ready for EDA | Luana Rodriquez — Senior Data Analyst |
| **2a** | **Begin exploring the data**  **Analyze** |  | Deshawn Washington — Data Analysis Manager |
| **3** | **Data exploration and cleaning**  **Plan and Analyze** | * EDA report * Tableau dashboard/visualizations | Luana Rodriquez — Senior Data Analyst |
| **3a** | **Visualization building**  **Construct and Analyze** |  | Uli King — Senior Project Manager |
| **4** | **Compute descriptive statisticsAnalyze** | * Analysis of testing results between two important variables * Share results of testing | Deshawn Washington — Data Analysis Manager |
| **4a** | **Conduct hypothesis testing**  **Analyze and Construct** |  | Udo Bankole — Director of Data Analysis |
| **5** | **Build a regression model**  **Analyze and Construct** | * Review testing results * Determine the success of the model | Luana Rodriquez — Senior Data Analyst |
| **5a** | **Evaluate the model**  **Execute** |  | Udo Bankole — Director of Data Analysis |
| **6** | **Communicate final insights with stakeholders**  **Execute** |  |  |
| **6a** | **Not necessary for this project**  **Select PACE stage** |  |  |

*Note: The estimated times for the milestones in the example equate to the length of the courses where you will learn the necessary skills. Realistic timelines when working with actual clients and data scientists as a data scientist would most likely have tight deadlines, for example:*

Milestone 1: 1-2 days

Milestone 2: 2-3 weeks

Milestone 3: 1 week

Milestone 4: 1 week

Milestone 5: 1-2 weeks