**Automatidata Project Proposal**

## **Overview:**

Automatidata will perform regression analysis to create a predictive model that would optimize New York City TLC’s transport operation. To best optimize New York City TLC’s operation, regression analysis is needed to create a predictive model that will produce trip duration based on distance, time of day, season, and other variables.

|  |  |  |  |
| --- | --- | --- | --- |
| **Milestone** | **Tasks** | **Deliverables/Reports** | **Relevant Stakeholder (Optional Activity)** |
| **1** | Establish structure for project workflow (PACE)  Plan | * Global-level project document |  |
| **1a** | Write a project proposal  Plan |  |  |
| **2** | Data exploration and cleaning  Analyze | * Data files ready for EDA (Exploratory Data Analysis) |  |
| **2a** | Inspect data files for missing data  Analyze |  |  |
| **3** | Understand the data  Analyze **and** Construct | * EDA results |  |
| **3a** | Visualization building  Analyze **and** Construct | * Tableau dashboard/   visualizations |  |
| **4** | Build a regression model  Analyze **and** Construct | * Analysis of testing results between two important variables * Share results of testing |  |
| **4a** | Build machine learning  Construct |  |  |
| **5** | Conduct hypothesis test  Execute | * Review testing results * Determine the success of the model |  |
| **5a** | Evaluate the model  Execute |  |  |
| **6** | Communicate final insights with stakeholders  Execute |  |  |
| **6a** | Not necessary for this project  **Select PACE stage** |  |  |