**Predictive Modeling for Ride Durations in NYC Taxi and Limousine Trips Proposal**

## **Overview:**

The project aims to develop a regression model for predicting ride durations based on various factors using data from the New York City Taxi and Limousine Commission (TLC). The project will involve tasks such as data exploration and cleaning, computing descriptive statistics, building a regression model, evaluating the model, and communicating insights with stakeholders.

| **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** | **Begin exploring the data**  **Analyze** | * Data files ready for EDA |  |
| **2a** | **Data exploration and cleaning**  **Analyze** |  |  |
| **3** | **Compute descriptive statistics**  **Analyze and Construct** | * EDA report |  |
| **3a** | **Visualization building**  **Analyze and Construct** | * Tableau dashboard/visualizations |  |
| **4** | **Conduct hypothesis testing**  **Analyze** | * Analysis of testing results between two important variables |  |
| **4a** | **Select software/languages**  **Plan and Construct** |  |  |
| **5** | **Build a regression model**  **Analyze and Construct** |  |  |
| **5a** | **Evaluate the model**  **Execute** | * Determine the success of the model |  |
| **6** | **Compile summary information about the data**  **Analyze** | * Final model |  |
| **6a** | **Communicate final insights with stakeholders**  **Execute** | * Report to all stakeholders |  |