**PACE Strategy for NYC TLC Taxi Fare Project**

Purpose: To develop a structured plan and timeline for a regression model that predicts New York City Taxi Fares using data provided by the NYC Taxi and Limousine Commission(TLC). The plan will guide the Automatidata team from project initiation through delivery, ensuring all milestones are clear and aligned with both internal team goals and client expectations.

Audience

* Internal: Data Analysts, Project Manager and Director at Automatidata(technical team)
* External: NYC TLC Finance and Operations Manager(non-technical stakeholders)

Context:

Automatidata has been contracted by the NYC TLC to create a fare prediction tool using existing trip data. This model will help riders estimate fares before the ride. The project is in the initial planning phase, and a structured approach is necessary to manage tasks, milestones and stakeholder communications.

Execution:

The project will involve data inspection, EDA, model development, evaluation and final reporting. Python and related data science libraries will be used. The deliverables include technical insights, visual summaries and presentation materials tailored to each stakeholder group.

**Project Proposal: NYC TLC Taxi Fare Prediction**

Objective

To build a regression model that estimates taxi fares using historical trip data from the NYC Taxi and Limousine Commission.

Key Milestones and Tasks

|  |  |  |  |
| --- | --- | --- | --- |
| Milestone | Task Description | PACE Stage | Duration(Days) |
| M1 | Data Access & Inspection | Plan | 2 |
| M2 | Exploratory Data Analysis | Analyze | 3 |
| M3 | Feature Engineering and Cleaning | Construct | 3 |
| M4 | Model Development(Regression) | Construct | 4 |
| M5 | Model Evaluation and Tuning | Construct | 3 |
| M6 | A/B Testing on variable impact | Construct | 2 |
| M7 | Create Visuals & Reports | Execute | 3 |
| M8 | Final Presentation and Delivery | Execute | 2 |

Stakeholder and Communication Plan

* Internal(Technical): Clear,concise updates with depth; regular progress reports via Slack/Email.
* External(Non Technical-TLC):Visual aids,simplified summaries, periodic presentations

Tools & Techniques

* Python(pandas,numpy,maplotlib,seaborn,scikit-learn)
* Jupyter Notebook
* Google Slides or PowerPoint for presentations

Deliverables

* Cleaned dataset with summary
* EDA report with key findings
* Trained and validated regression model
* Visuals explaining predictions and performance
* Final presentation tailored for TLC stakeholders