Course 1: Foundations of Data Science

New York City TLC Project Proposal

Overview:

The objective of this project is to utilize the data collected by New York City Transit Limousine Commission and develop a regression model for them to predict the fare price before the rides.

Prepared by: Nawaz Quraishi

Milestone	Tasks	Deliverables/Reports	Relevant Stakeholder
1	Establish structure for project workflow (PACE) Plan	• Global-level project document	Deshwan Washington - Data Analysis Manager
1a	Write a project proposal Plan	• Project Proposal	Uli King - Senior Project Manager
1b	Select software/languages Plan	• Establishing software/hardware needs, languages to be used	Luana Rodriguez - Senior Data Analyst
2	Data exploration and cleaning Analyze	• Data scrubbed, converted and formatted	Luana Rodriguez - Senior Data Analyst

2b	Begin exploring the data Analyze	• Data files ready for EDA	Luana Rodriguez - Senior Data Analyst
2c	Compute descriptive statistics Analyze	• EDA report	Luana Rodriguez - Senior Data Analyst
3	Build a regression model Analyze and Construct	• Analysis of testing results between two important variables that would be best suited for constructing the prediction model	Luana Rodriguez - Senior Data Analyst
4	Build a machine learning model Construct and Execute	• Machine learning model is constructed	Deshwan Washington - Data Analysis Manager
5	Evaluate the model Execute	• Determine the success of the model, incorporate feedback • Final model	Udo Bankole - Director of Data Analytics Deshwan Washington - Data Analysis Manager
5a	Communicate final insights with stakeholders Execute	• Tableau dashboard/visualizations	Udo Bankole - Director of Data Analytics Deshwan Washington - Data Analysis Manager Luana Rodriguez - Senior Data Analyst

The estimated timelines for the Milestones are as follows:
Milestone 1: 1 to 2 days

Course 1: Foundations of Data Science

Milestone 2: 4 to 5 days
Milestone 3: 1 week
Milestone 4: 1 week
Milestone 5: 1 to 2 weeks