

# Create Automatidata

## Riding the Data Wave: Preparing data for predictive Insights from NYC's Taxi Trails

### Project Overview

The Automatidata project involved consulting for the New York City Taxi and Limousine Commission (TLC) to transform their vast dataset into predictive insights. Automatidata specializes in leveraging client data to optimize business needs, from dashboards to strategic models. For TLC, the goal was to develop a regression model that predicts taxi fares before rides commence. This project marked the initial phase of preparing the data for analysis.

### Key Insights

**Data Structuring:** Initiated with constructing a comprehensive dataframe to house TLC data, ensuring readiness for analysis.

**Type Clarification:** Thorough review and summary of column data types with focus on numeric and categorical definitions to guide transformations.

**Statistics Gathering:** Extracted descriptive statistics facilitated deeper understanding of ride patterns, fare structures, and variable distributions.



### Details

Automatidata, a data consultancy, partners with New York City TLC to utilize taxi and limousine data from nearly one million trips daily, offered by over 200,000 licensees. The project started with foundational tasks crucial for analysis: building a dataframe, examining column data types, and gathering descriptive statistics.

### Problem & Response

#### Problem

How can we translate complex, high-volume taxi operation data into actionable predictive models that accurately estimate fares pre-ride?

#### Response

This project employed Python and pandas to construct and refine a dataframe, systematically verifying data types and producing essential descriptive statistics. These assisted in organizing the data for comprehensive exploratory analysis.

### Next Steps

#### Data Analysis:

Launch detailed exploratory analysis on structured data, steering model development.

#### Model Building:

Develop regression models to predict fares and appraise their performance for real-world usability.

#### Visualization Creation:

Develop strategic dashboards that demonstrate insights visually and aid decision-making.

#### Portfolio Expansion:

Integrate course projects into a cohesive GitHub repository, demonstrating comprehensive expertise gained.