

Exploratory Data Analysis for NYC TLC Data

Executive Summary Report Commission Prepared by Automatidata

Project Overview

The NYC Taxi and Limousine Commission (TLC) has collected detailed taxi and rideshare trip data. Our objective is to create a regression model to predict taxi fares.

Details

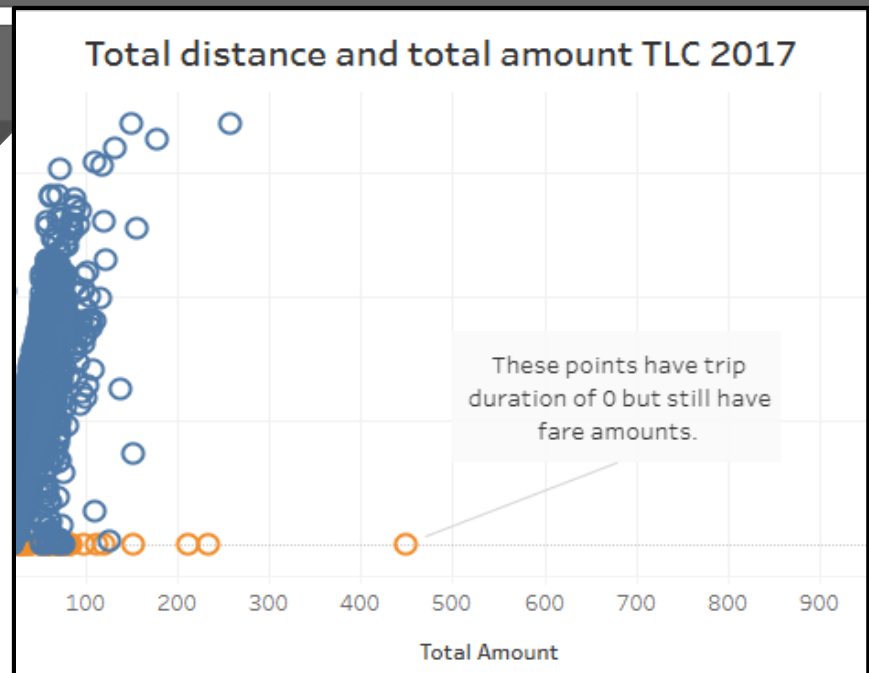
Key Insights

Problem:

After conducting initial EDA it is observed that some data points will be hurdle for predicting accurate ride fair prediction.

Proposed solution:

We recommend removing outliers with total distance of 0



Scatter Plot Between 'total_distance' and 'total_amount'.

Next Steps

- Clean data and remove outliers.
- Determine the variables that have the largest impact on trip fares.
- Conduct statistical analysis and parameter tuning, and run regression to filter relevant columns.