Mini Project - PDSML

Domain : Logistics.

Content : This data set is a masked data set which is similar to what data analysts at Uber handle.

Features Information

Request id : A unique identifier of the request.

Pickup point : The point from which the request was made.

Driver id : The unique identification number of the driver.

Status : The final status of the trip, that can be either completed, canceled by the driver or no cars available.

Request timestamp: The date and time at which the customer made the trip request.

Drop timestamp : The drop-off date and time, in case the trip was completed.

Deliverables

Data Analysis : In this stage extracts useful information statistically. Such as check outliers, skewness, compare the features by graph and many more. Do all the required steps as well.

This process involves performing

1. **Univariate Analysis :** In this part, first check target features and start univariate analysis.
2. **Bivariate analysis :** This analysis involves studying two variables and their relationship, recalling some of the hypotheses that we generated earlier.
3. **Removing Missing values if any :** After exploring all the variables in our data, we can now impute the missing values and treat the outliers because missing data and outliers can have an adverse effect on the model performance and accuracy.

Dataset link : **Uber Data**

Deadline to submit the project : One Week

ALL THE BEST