A brief on the approach used to solve the problem.

Taxi demand prediction is Regression problem. So, we have to predict how many number of taxi demand on particular date and hour. In given dataset there is 2 dependent variables

1. Date
2. Hour

Task is predict taxi demand from this 2 variables. from Explorartry data analysis we get below information. to create feature.

1. in day there is rush time when taxi demands are high
2. on weekend there is less taxi demand compare to normal days
3. on holiday there is less taxi demands.

from this key points will create features. and feed it to different regression algorithm. and fine tune that algorithm.

Which Data-preprocessing / Feature Engineering ideas really worked? How did you discover them?

1. Holiday or not

2. weekday or weekends

3. rush hours

from EDA we get insights about features so we have created that.

What does your final model look like? How did you reach it?

After checking few model i get to know XGBoost is working good on this problem compare to other models. I have try different models like ridge, Random Forest and KNN. after that I fine tune the model.