IBM – Coursera
Car Accident Severity Project

BY

K UDAY KARTHIK

INTRODUCTION:

- Every Life on this planet is very much valuable for someone or the other.

- This problem is about numerous car accidents that occur in our daily lives. People should be more careful while driving the car. Every second, many people across the world will be losing their lives on the road.

INTRODUCTION

- Their family members will be waiting for them in their home. In this project We will be analyzing the car accident data, how the accidents are caused generally. What to do to mitigate accidents. Weather and temperature also might affect the environment.

- Driving during night times is too risky. In countries where there will be snowfall, It will be very difficult for the driver to drive the car.

BUSINESS UNDERSTANDING

- **Requirements:**
- What is the severity rate?
- On what dates, accidents happened more ?
- On what dates, accidents happened less?

DATA PREPARATION

- I have chosen the open-source dataset from
- **https://www.kaggle.com/sobhanmoosavi/us-accidents**
- It is free to use.

DATA CLEANING

- The dataset is very huge. So to simplify it. I have chosen a particular state called
- “New York” (NY) from all the states in the USA.
- The records show the accident details of “New York”.
- In the dataset, I have used the columns like Accident Severity, Date, Time, Latitude, Longitude, Temperature, WInd Speed, Day / Night and many more. Since accidents are dependent on all these factors.

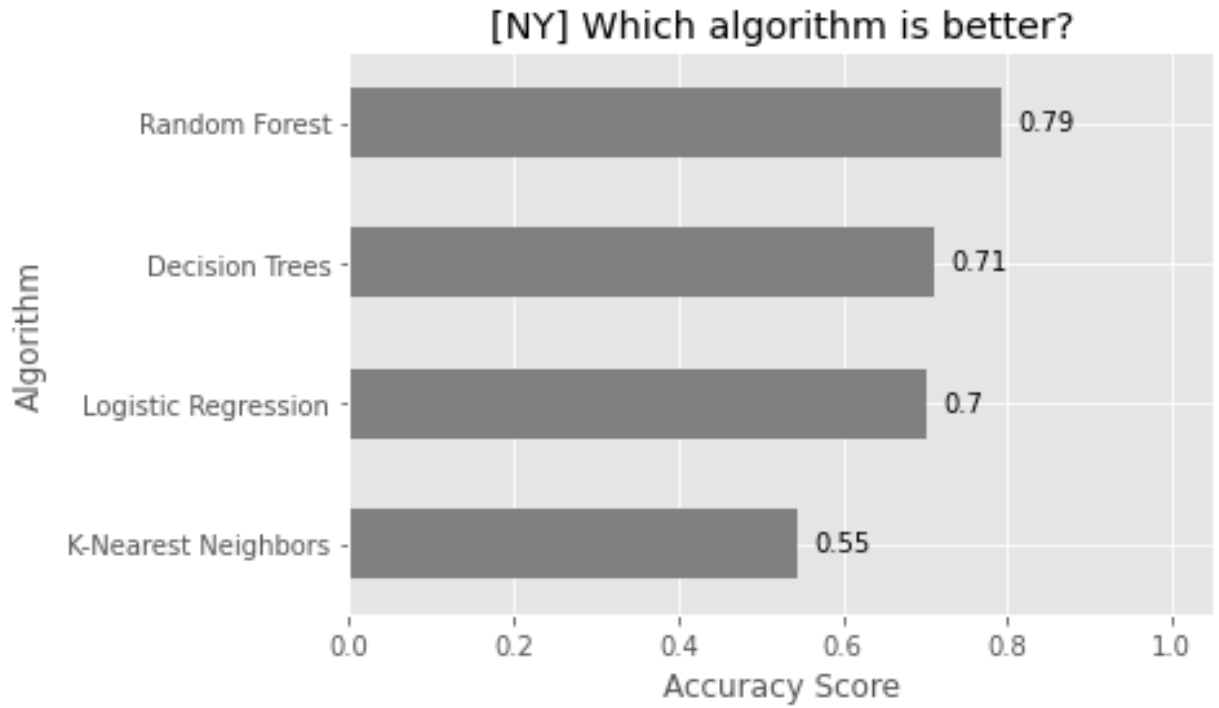
DATA CLEANING

- Actually there were numerous attributes. I have cleaned the data by considering only relevant attributes that are necessary for my project.
- Missing Values all have been checked.

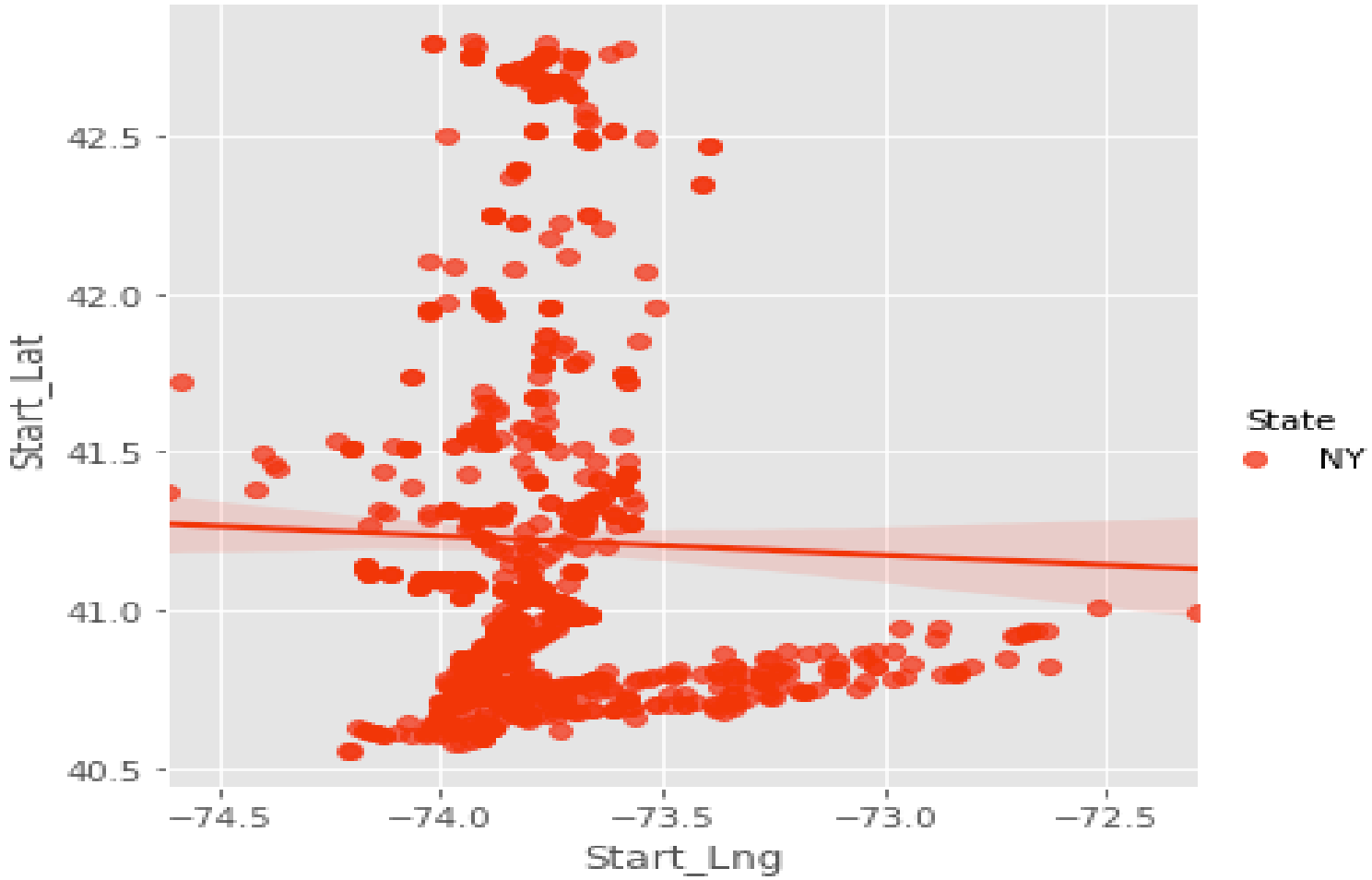
MODEL EVALUATION

- I considered mainly 4 models for this project.
- Logistic Regression, Decision Tree Regression, KNN and Random Forest Regression.
- I have fit my data to all these models and predicted accident severity.
- To my knowledge, I have observed with given dataset, Random Forest was a better one.

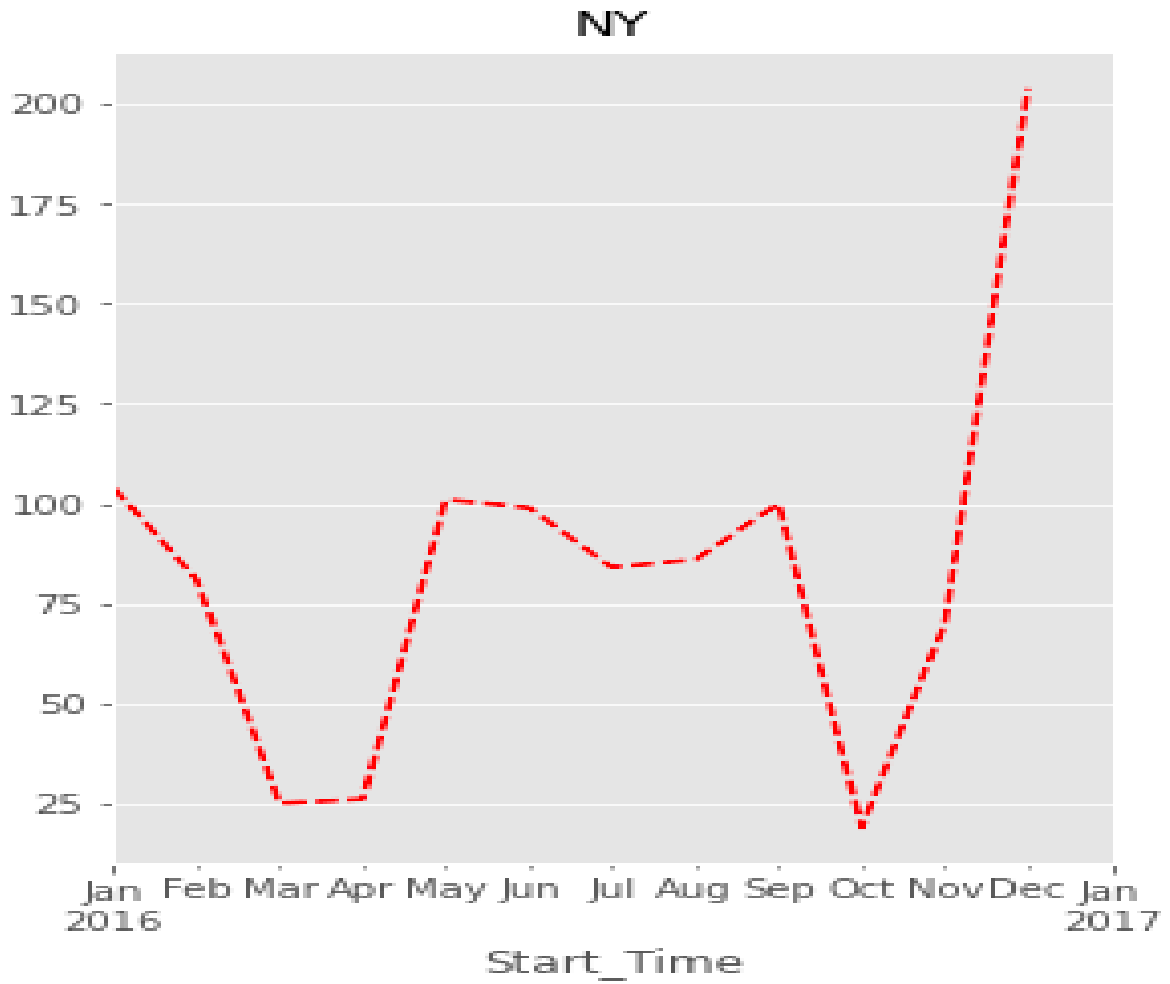
VISUALIZATIONS



VISUALIZATIONS



VISUALIZATIONS



VISUALIZATIONS

For New York in this dataset:
There are 999 total accidents.
There are 999 total days.
There are 14 unique days.
On average, there are 71 accidents per day.

CONCLUSION

- From Visualizations, we can interpret that # of accidents was high enough during year-end. This might of because of Christmas and New Year effect.
- Most of the people will be outdoors during night time. As Driving during night time is too risky. Some may feel drowsy too. Most of them drink and drive too.
- So during this season, strict police checking must be done for drunk and drive cases too. Hotspot regions should be under police control.

THANK YOU!