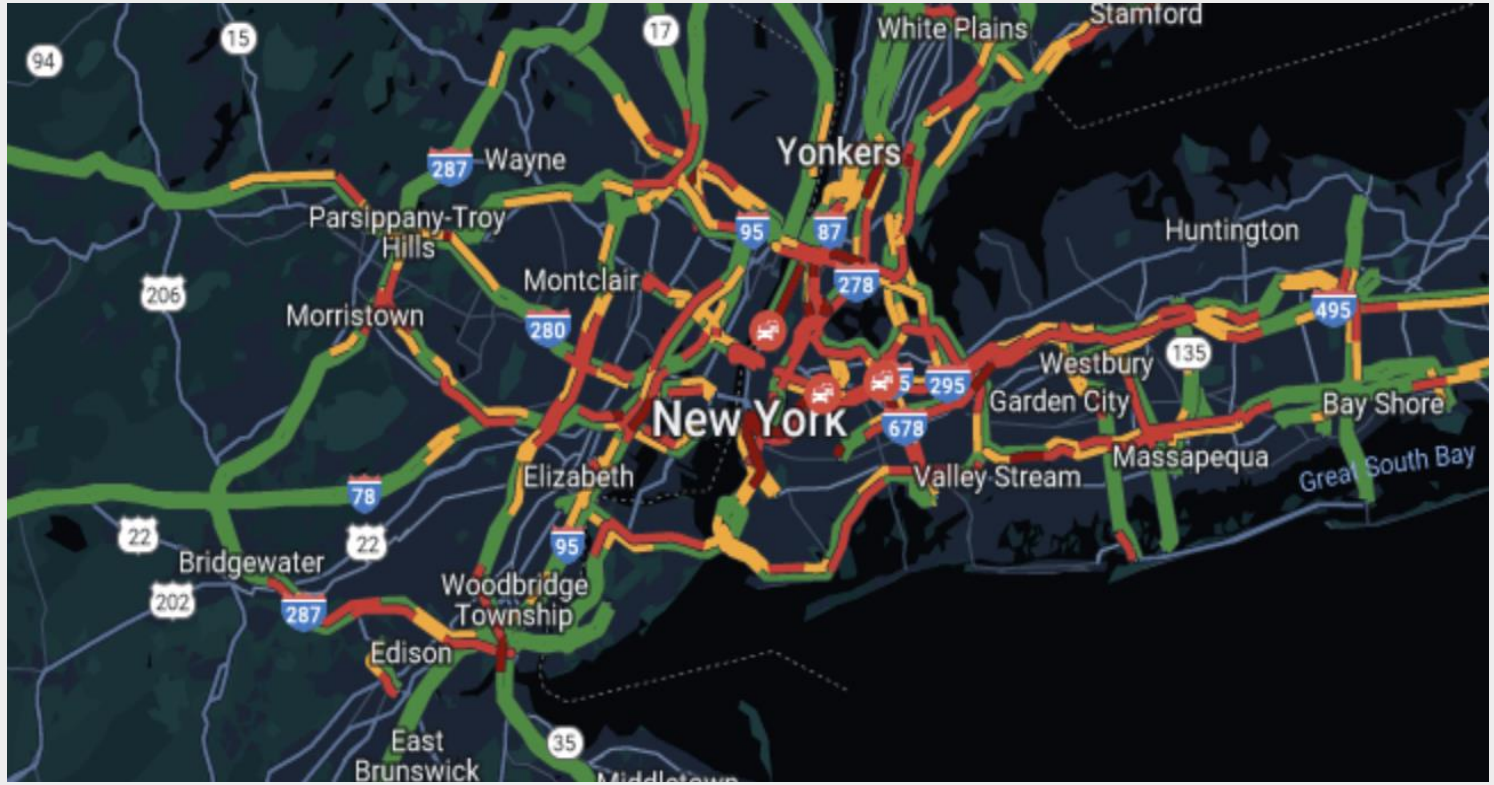
Analysis of Traffic Accidents In New York City

Modelling data for all collisions to understand any patterns or consistencies



Authors: **Joseph Olaniyan**
Prachi Shah
Levan Getia
Angelo Hovarth

Introduction

General NYC characteristics

- New York city is home to 8.5 million people
- In 2022, there were around 100,50*8* total traffic accidents in NYC
- This breaks down to about roughly 625 accidents per day

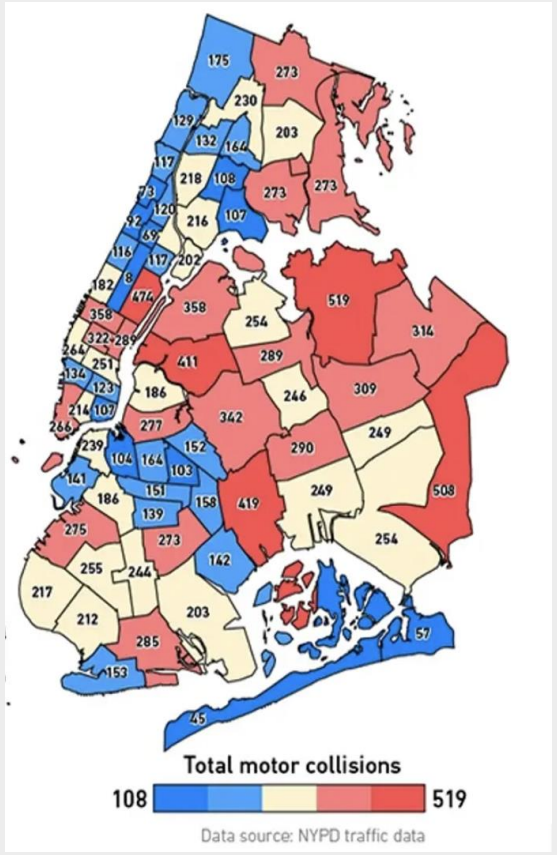
Reasons for choosing NYC as a model city

- NYC is a very traffic accident prone city
- It has vast data of accidents that can be analyzed
- 38% of all traffic accidents in NYC end with fatality or injury

Project Outline

The Data

- Firstly, data is obtained from CSV file, which contains accident entries
- Afterwards it is cleaned using jupyter notebook and loaded to mongoDB
- Afterwards API is set up to access data through server



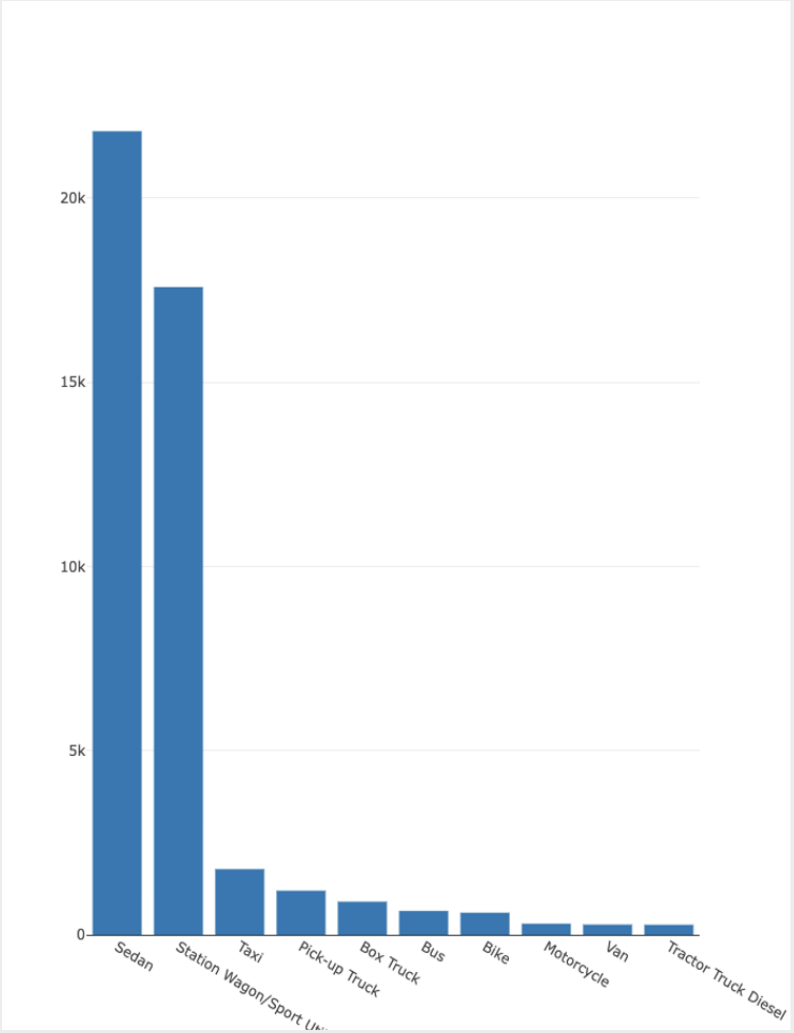
Modelling and Analysis of data

- Analysis is mainly performed through javascript usage
- We developed webpage which will serve as a main project host
- Our project performs analysis of whole data and models patters
- Mainly, our goal was to capture vehicles that are most susceptible to accidents and what regions of NYC, or also what time of day is the riskiest, in terms of getting in an accident

Project attributes

Visualization

- Project offers map with several layers, which gives out clustered data of all the accidents according to their location, as well as top 3 and bottom 3 cities where accidents are most or least likely to happen.
- We also developed a pie chart, which tries to capture the part of day when most accidents happen statistically.
- Lastly, this project offers bar chart, modelling the top 10 vehicle types that are prone to getting into the accidents



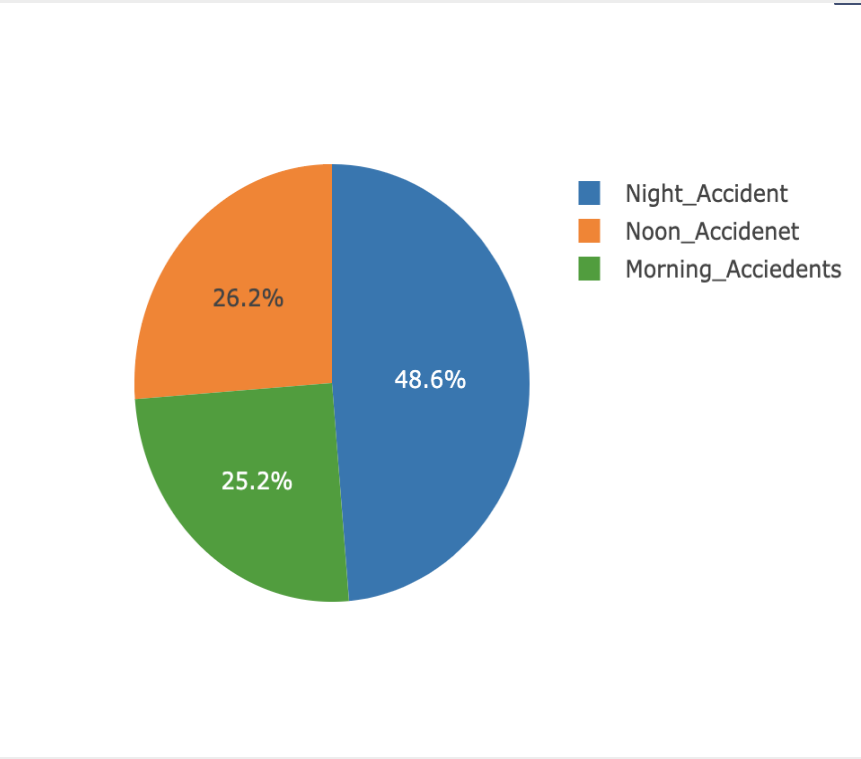
Hard Data

- Project provides cleaend csv file with no null entries
- There are 3 different APP routes that can be accessed after running app.py
- Data is collection of all motor collisions reported to NYPD from January 2020 to August 2020

Key Findings

Accident Statistics

- Nearly half of the accidetns took place at night time, specifically from 12 am to 5pm
- Morning and Afternoon crashes both shared quarter of total data with afternoon crashes higher with just 1 %
- Sedan is number one in terms of most accidents, followed by SUV, both have accounted for more than 15 thousand records each, compared to just 301 or 277 for motorcycle and van respectively, in the same timeframe.



NYC Regional acccident data

- Queens and Brooklyn have recorded most of tha accidents for the given period.
- Staten Island was safest region with just under 1400 accidents
- Manhattan recorded second to last in terms of accidents with Riverside Drive, happened to be one of the least accidental streets inside Manhattan.

References

Data

Initial data is pipelined through csv from Maven Analytics : https://mavenanalytics.io/data-playground?page=6&pageSize=5

Add-ons Used: Leaflet plugins : https://leafletjs.com/plugins.html
OpenStreetMap: https://www.openstreetmap.org/#map=3/71.34/-96.82

Software and Languages used: Jupyter Notebook, Python, JavaScript, Google Chrome.

GitHub Link: https://github.com/Prachi-Shah1002/Project_3

Thank you ! Lets expxlore the webpage