University at Buffalo
The State University of New York

EAS 503

NYPD Motor Vehicle Collisions

Group 2

1. Karan Nisar
2. Sai Krishna Kanneti
3. Tejas Dhrangadharia

1. Abstract

NYPD motor vehicle collisions data set is a breakdown of every collision in NYC by location and injury. Each record represents a collision in NYC by city, borough, precinct and cross street. The motor vehicle collision database includes the date and time, location (as borough, street names, zip code and latitude and longitude coordinates), injuries and fatalities, vehicle number and types, and related factors for all 1.7 million collisions in New York City during 2011 to 2017.

The objective of the NYPD Motor Vehicle Collisions project is to identify the root cause of vehicle collisions and plotting the various contributing factors versus collisions. Plotting the number of person injured/killed, number of fatal accidents occurred divided on borough and their location mapping in NYC to provide better visualizations.

Collisions are plotted according to borough. This will help to identify the time of the day maximum number of accidents happen, the type of vehicles responsible for frequent accidents and the number of persons (Cyclist, Motorist, and Pedestrian) killed/injured in these accidents.

Matplotlib is used for plotting, Pandas, SQL are used for data querying and numpy is used for analysis. These reports/visualizations can be used by the public to see how dangerous/safe intersections are in NYC. These reports can be used by the local authorities to alert the commuters by including sign boards.

2. Dataset

The data set has 1.17 million rows which represents 1.17 million collisions. There are 29 columns which includes the date and time, location (as borough, street names, zip code and latitude and longitude coordinates), injuries and fatalities, vehicle number and types, and related factors for all the collisions. The data is collected from 2011 to 2017.

This is a breakdown of every collision in NYC by location and injury. Each record includes a collision in NYC by city, borough, cross street, the date and time, location (as borough, street names, zip code and latitude and longitude coordinates), injuries and fatalities, vehicle number and types, and related factors. This data is manually run every month and reviewed by the TrafficStat Unit before being posted on the NYPD website. Each record represents a collision in NYC by city, borough, precinct and cross street

Date	Time	Borough	Zip code	Latitude	Longitude	On Street Name	Number of People Killed	Number of People Injured	Contributing Vehicle 1
2013-01-20	2:00:00	MANHATTAN	10010	40.7407	-73.9818	EAST 26 STREET	1	1	Driver Inattention/Distraction
2013-11-11	12:50:00	MANHATTAN	10029	40.7904	-73.9498	PARK AVENUE	1	3	Failure to Yield Right-of-Way
2013-06-04	8:15:00	MANHATTAN	10025	40.7947	-73.9699	WEST 97 STREET	1	1	Aggressive Driving/Road Rage
2014-01-10	20:48:00	MANHATTAN	10025	40.7959	-73.9727	WEST 97 STREET	1	1	Failure to Yield Right-of-Way
2013-02-06	14:15:00	MANHATTAN	10031	40.8201	-73.9551	WEST 135 STREET	1	6	Failure to Yield Right-of-Way

3. Introduction

Motor vehicle traffic crashes are the leading cause of injury related death for New York residents. During the period from 2012-2014 there was an annual average of 698deaths, 8,093 hospitalizations and 136,913 emergency department visits because of motor vehicle traffic injuries among New York residents.

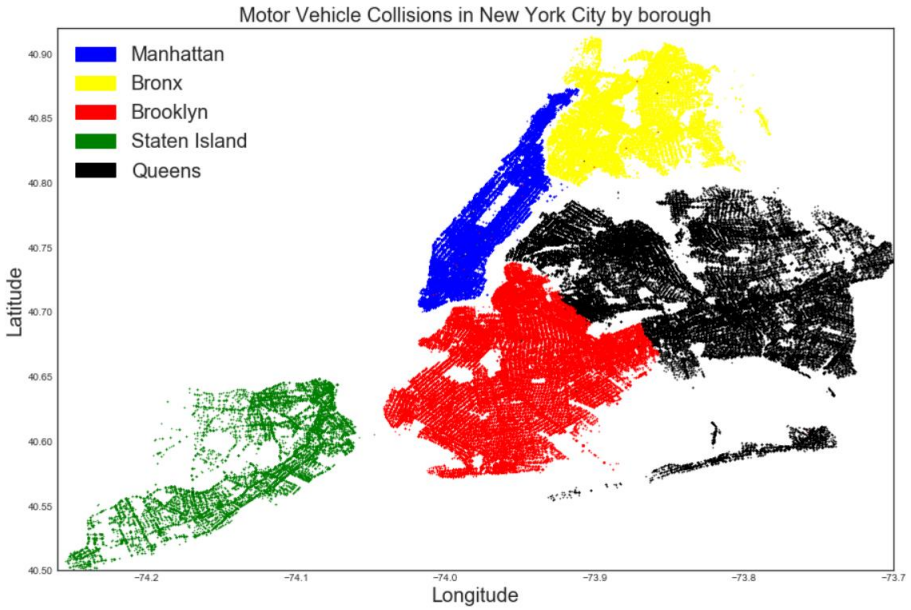
Motor Vehicle Traffic Injuries are a serious public health problem in New York County. They are a leading cause of injury related deaths. Crashes are not only a significant cause of death, pain and suffering, but also an economic burden to New York County. In 2014, the crashes on New York County's roadways resulted in $36.5 million in hospitalization and emergency department (ED) charges. These crashes are not accidents! They are not random, uncontrollable acts of fate, but occur in predictable patterns, with recognizable risk factors and among identifiable populations. A crash-related injury is a predictable and preventable event.

The New York city is divided into 5 boroughs: Manhattan, Bronx, Brooklyn, Staten Island, Queens. The data set is loaded into mysql and pymysql is used to query the database

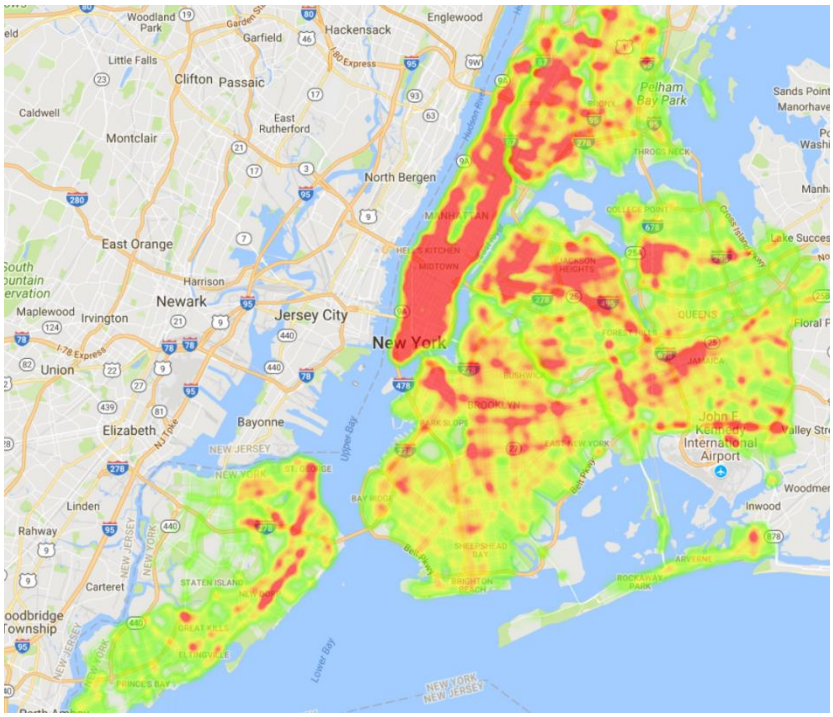
4. Analysis and Results

4.1. Motor Vehicle Collisions in NYC by Borough

- There are 5 boroughs in New York City and during 2016-17 there are total 1048573(As per the data) collisions happened in NYC. Below plot is the scatter plot for number collisions occurred in these 5 boroughs during 2016-17.



- This plot clearly display that collisions happened almost in all 5 boroughs.
- Heat plot will give some more insights about the high accuracy of collisions. Below is the heat map for the number of collisions in NYC,



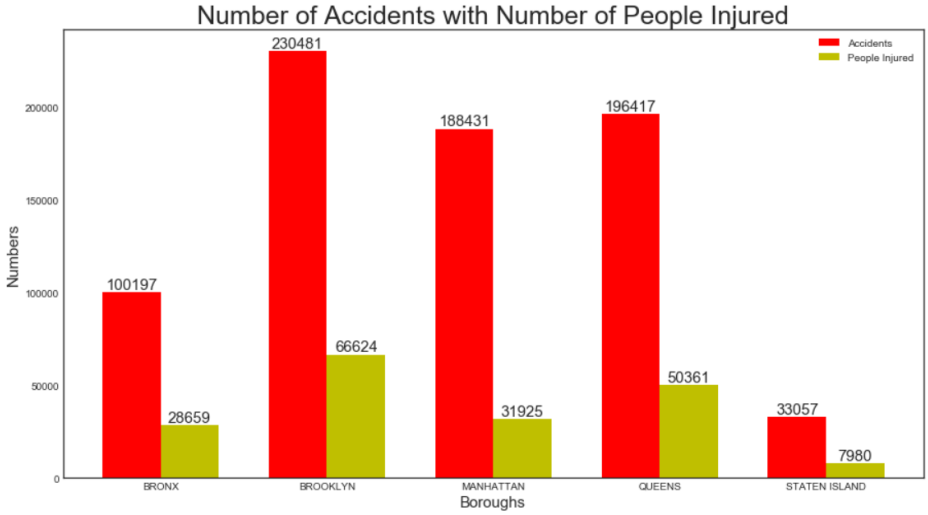
- As Manhattan is **populous** borough in NYC, highest accuracy of collisions is visible in Manhattan.

4.2. Number of People Killed and Injured

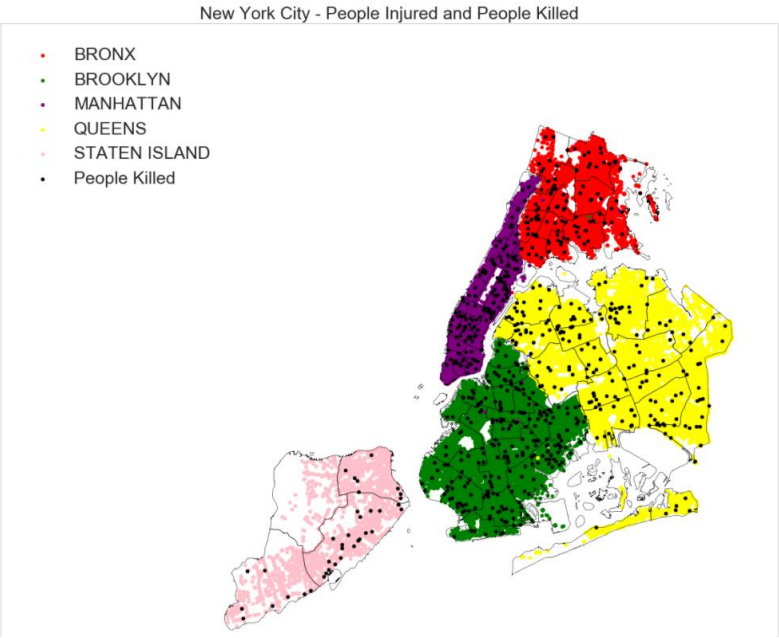
- Below is the comparison of number of accidents with number of people killed and injured,

Borough	Number of Accidents	Number of People Injured	Number of People Killed
Manhattan	1,88,431	31,925	157
Brooklyn	2,30,481	66,624	260
Bronx	1,00,197	28,659	109
Queens	1,96,417	50,361	232
Staten Island	33,057	7,980	42

- And plot for number of number of collisions with number of people injured is given below:



- Displayed number of people injured and number of people killed by borough in a basemap.

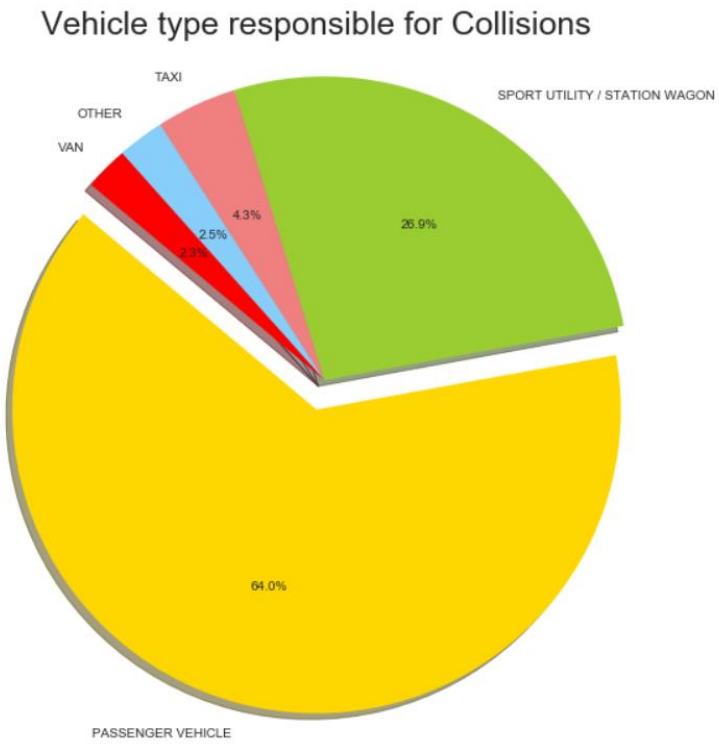


- This plot matches with data shown in the table, in Staten Island number of people killed is less compare to other boroughs.

4.3. Responsible factors for collisions:

4.3.1. Vehicle Type:

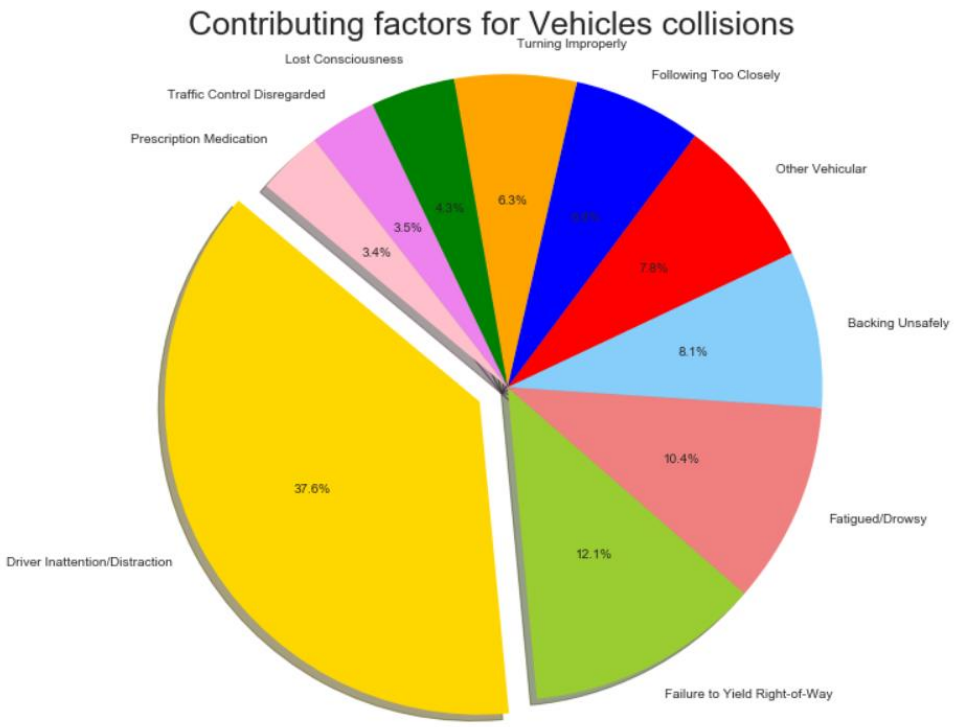
Vehicle type responsible for collisions is displayed in below pie-chart



Most of the collisions(~64%) happened because of the 'Passenger Vehicles' and the second highest vehicle type responsible for collisions is 'Sport Utility / State Wagon'.

4.3.2. Contributing Factors

Below pie chart is the contributing factors for vehicle collisions



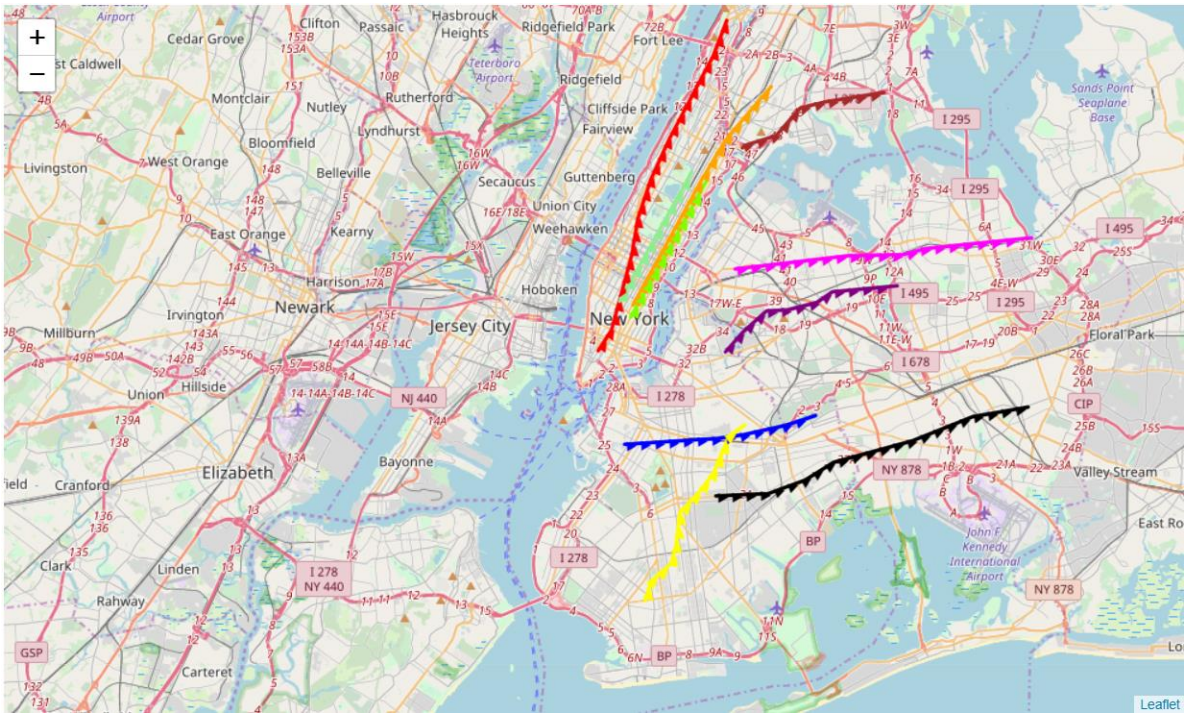
Driver Inattantion/ Distractions is the main factor for vehicle collisions. And, Failure to Yield Right-of-Way is the second highest factor. NYC traffic police can eliminate this by enforcing some strict traffic rules.

4.4. Dangerous Streets/ Intersection in NYC

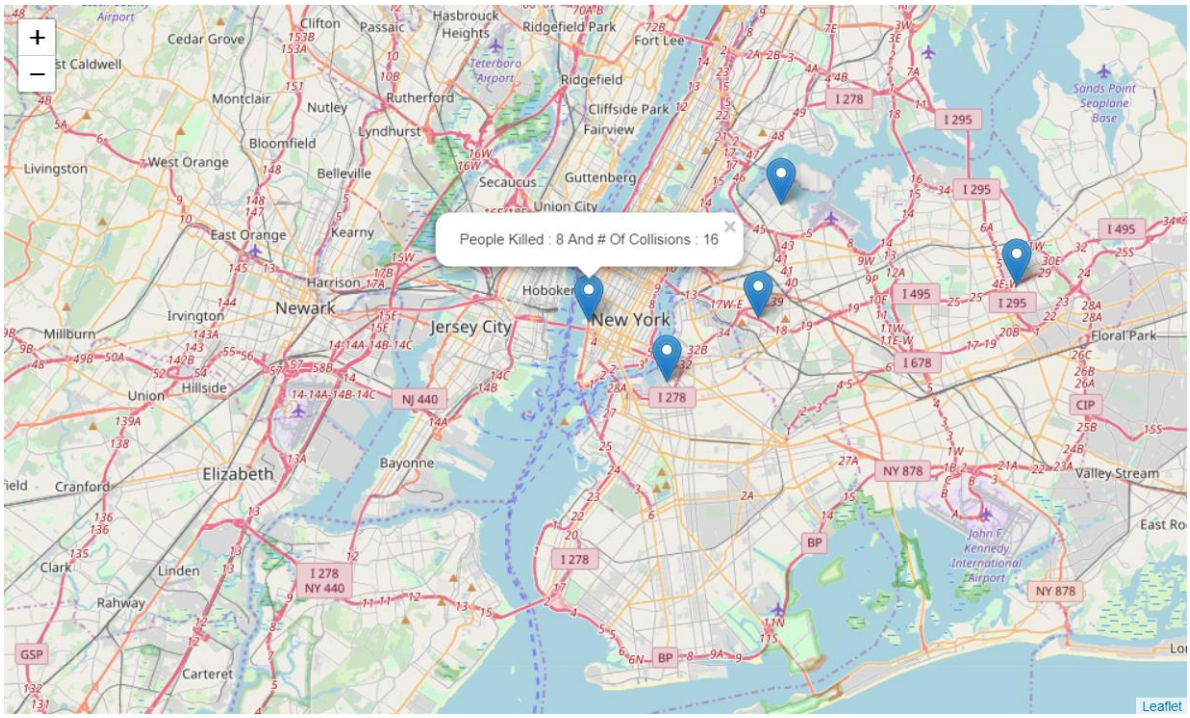
- Data for top 10 dangerous streets in NYC with number of collisions is shown below:

Street	Number of Accidents
BROADWAY	10,854
ATLANTIC AVENUE	9,642
NORTHERN BOULEVARD	7,535
3 AVENUE	7,433
FLATBUSH AVENUE	6,374
QUEENS BOULEVARD	6,154
LINDEN BOULEVARD	5,772
2 AVENUE	5,415
BRUCKNER BOULEVARD	4,624
5 AVENUE	4,330

- Broadway and Atlantic Avenue has the highest number of collisions, those are the top 2 dangerous streets in the NYC.
- Below is the representation of these streets on map



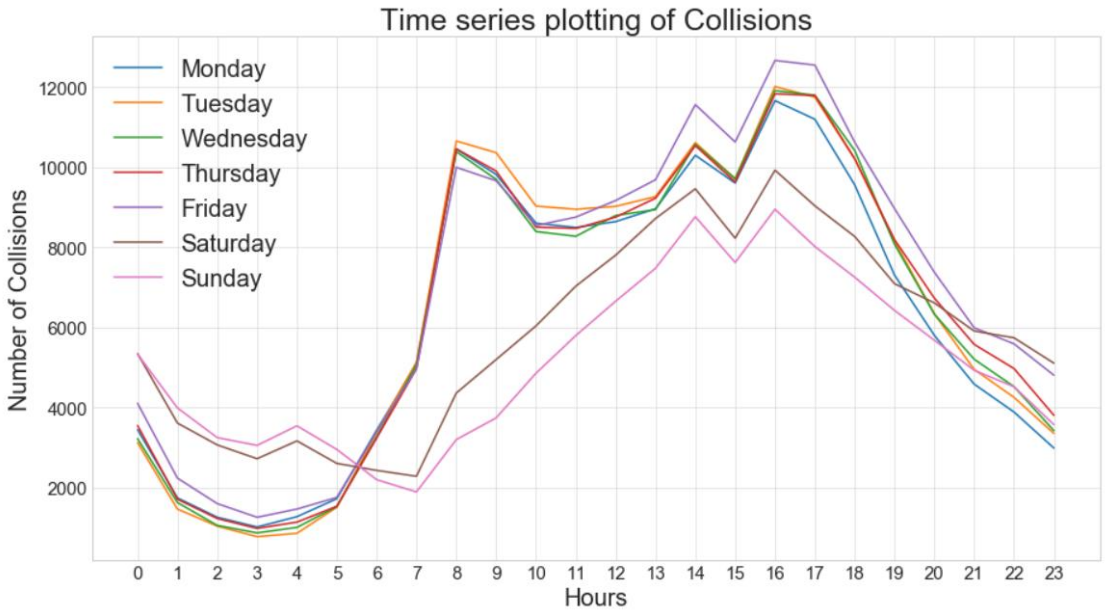
- Details for the dangerous place in NYC where number of people killed is as per below plot



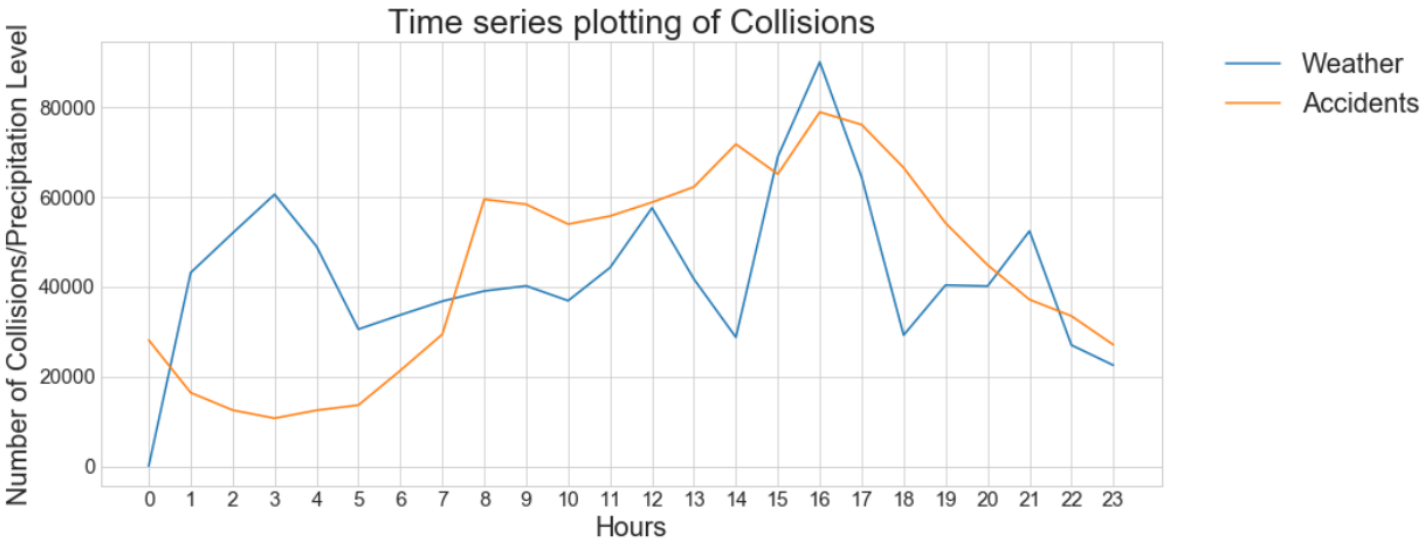
- These are the 5 places where number of people killed is the highest in NYC. So, be aware while someone is passing through this places.

4.5. Time Series Analysis

- Below is the plot for time series plot on daily basis,



- From the plot, travel in night is safe during weekdays and during weekends travel in day time is safe.
- There is some relation between weather and collisions, below is the time series plot and weather precipitation level,



- During the evening, weather precipitation level is high and the number of collisions is also high. So, weather has some relation with collisions.

5. Conclusion

The analysis of Collisions in New York City gave some good insights about the various causes for the accidents. It can be seen that Driver Inattention was one of the major contributing factor of collisions, so such recognizable collisions can be avoided by enforcing strict laws. In addition to it, failure to yield right of way was another major contributing factor for increase in collisions. Therefore, drivers should be vigilant and follow traffic rules. Many passenger vehicles were involved in collisions, so some new laws should be implemented before issuing license to the drivers. Brooklyn has had a large amount of collisions, so one can be attentive while in Brooklyn. The top ten dangerous streets has also been plotted on the map of New York City. These streets have major junctions where collision frequency was highest. Such junctions can be avoided if possible. It can also be seen that traveling during night times on weekends was dangerous while during day time there were more collisions on weekdays. Also the time of around 4pm was found to have high collisions, so such times drivers and pedestrians should be careful. Another factor that was found contributing to increase in number of collisions was the Weather (more precipitation might lead to more slippery roads). The plot of weather vs collision time series confirmed this that high precipitation resulted in increase of accidents.

6. Future Research Directions

By using real time weather data, this information can be used to predict the risk of collisions and alert the drivers. Such implementation would be very useful and can avoid a lot of casualties. The data obtained is taken after every intervals of time. It is not taken at that particular time. Hence the minute to minute analysis cannot be obtained which can help identify the safe time much better. Sensors can be installed at those dangerous junctions that can warn the drivers or pedestrians of any risks. This can help reduce collisions.

7. References

- Data set reference: - https://data.cityofnewyork.us/Public-Safety/NYPD-Motor-Vehicle-Collisions/h9gi-nx95/data
- https://www.health.ny.gov/statistics/prevention/injury_prevention/traffic/county/new_york/2014/new_york_co_crash_fs.pdf