

# Visualizing Motor Vehicle Collisions in NYC

Ameya Tayade (N12574690)  
Email: aat404@nyu.edu, Net ID: aat404

Visualizing each incident of motor vehicle collisions on a map will empower the end-user with insights into the street traffic conditions and the factors influencing such incidences. Vehicle drivers will use this tool to safely plan their routes. Governmental civic body planning officials will utilize this tool to locate areas with high number of traffic problems. By locating areas with higher densities of traffic problems, government officials will be able to prioritize repair works to eliminate traffic problems and thus making streets safe for everyone.

This tool will be developed by utilizing NYC open data of motor vehicle collisions available from the url: <https://data.cityofnewyork.us/Public-Safety/NYPD-Motor-Vehicle-Collisions/h9gi-nx95>

Project page (on Github): <https://github.com/NYU-CS6313-Fall16/Motor-Vehicle-Colision-21>

Video: <https://vimeo.com/196913599>

Working demo: <https://ameyatayade90.github.io/>

## What is the problem you want to solve and who has this problem?

The biggest problem on streets is the collision involving motor vehicles as these problems have very high chances of increasing threat to human life and property. To minimize this problem, we should understand where such incidences are reported on a regular basis so that such areas can get higher priorities for repair works. In addition to locating such areas, it is also important to analyze the nature of these incidents because an area may have incidents reported due to faulty road condition and another area may be affected due to unregulated pedestrian traffic. Every distinct problem has its distinct solution. Therefore, in this project I have modeled the location of incidents as well as the nature of incidents. The following abstract taken from nyc.gov website has inspired to attempt to solve this problem:

*“The primary mission of government is to protect the public. New York’s families deserve and expect safe streets. But today in New York, approximately 4,000 New Yorkers are seriously injured and more than 250 are killed each year in traffic crashes. Being struck by a vehicle is the leading cause of injury-related death for children under 14, and the second leading cause for seniors. On average, vehicles seriously injure or kill a New Yorker every two hours.”*  
-- (Taken from: <http://www1.nyc.gov/site/visionzero/index.page>)

By designing a visualization tool to address this issue, I hope to accelerate the process of finding collision incident ‘hot-spot’ areas and also to analyze nature of these incidents.

## **What are the driving analytical questions you want to be able to answer with your visualization?**

The driving analytical questions that we can answer using this visualization tool have been refined. The questions are as follows:

1. Motor vehicle collisions are more prevalent in which parts of NYC?
2. What are the common causes of incidents in an area?
3. What are the factors that influence the collision incidents in an area?
4. Does the distribution of number of incidents in an area change with respect to year and month?

## **What does your data look like? Where does it come from? What real-world phenomena does it capture?**

The visualization tool is developed using JavaScript and NYC Open Data for Motor Vehicle Collisions available at this url: <https://data.cityofnewyork.us/Public-Safety/NYPD-Motor-Vehicle-Collisions/h9gi-nx95>

The attributes of this dataset used in this tool are explained in the table below:

Attribute Name	Attribute Type	Attribute Details
Date	Date in format YYYY-MM-DD	The date when the incident took place
Time	Time in format HH:MM	The 24-hour time when the incident took place
Borough	String	The borough where the incident occurred
Zip Code	String	The zip code of the place where the incident occurred
Latitude	Floating Point Number	The latitude of the location where the incident occurred
Longitude	Floating Point Number	The longitude of the location where the incident occurred
Number of Persons Injured	Integer	The number of people who were injured in this incident
Number of Persons Killed	Integer	The number of people who were killed in this incident
Number of Pedestrians Injured	Integer	The number of pedestrians who were injured in this incident
Number of Pedestrians Killed	Integer	The number of pedestrians who were killed in this incident
Number of Motorists Injured	Integer	The number of motorists who were injured in this incident
Number of Motorists Killed	Integer	The number of motorists who were killed in this incident
Number of Cyclist Injured	Integer	The number of cyclists who were injured in this incident
Number of Cyclist Killed	Integer	The number of people who

Contributing Factor Vehicle 1	String	were killed in this incident This is a categorical string that mentions what was the cause of this incident
-------------------------------	--------	--

### What have others done to solve this or related problems?

The related projects developed to solve this problem are explained below:

#### 1. NYC Vision Zero

Taken from nyc.gov website:

The Vision Zero Crash & Interventions Map is an interactive tool that shows detailed information on traffic injury and fatality crashes in New York City and highlights how the city is responding every day to make our streets safer. The map displays crash types, dates and locations and highlights Vision Zero initiatives such as Arterial Slow Zones, planning workshops and expanded traffic enforcement, major safety projects, as well as other long running safety programs

URL for this tool: <http://www1.nyc.gov/site/visionzero/index.page>

#### 2. NYC Crashmapper

Taken from nyccrashmapper.com website:

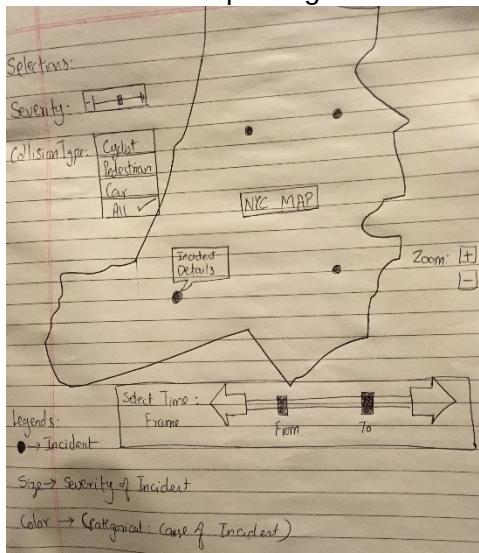
NYC Crashmapper lets you browse and analyze several years of intersection-level collision data on a map of NYC.

URL for this tool: <http://nyc.crashmapper.com/13/6/13/9/transit/collisions/2/12/40.707/-73.891>

Related Article: <https://www.wired.com/2013/08/crashmapper-nyc/>

### Design Iterations

The initial mockup designed to develop this tool is shown in the image below:



#### Modifications Performed:

The mockup developed had filters for selecting incidents based on severity and collision type. These filters were removed from implementation because modeling the causes of incidents was more important than collision types and so the causes have been modeled in the tool.

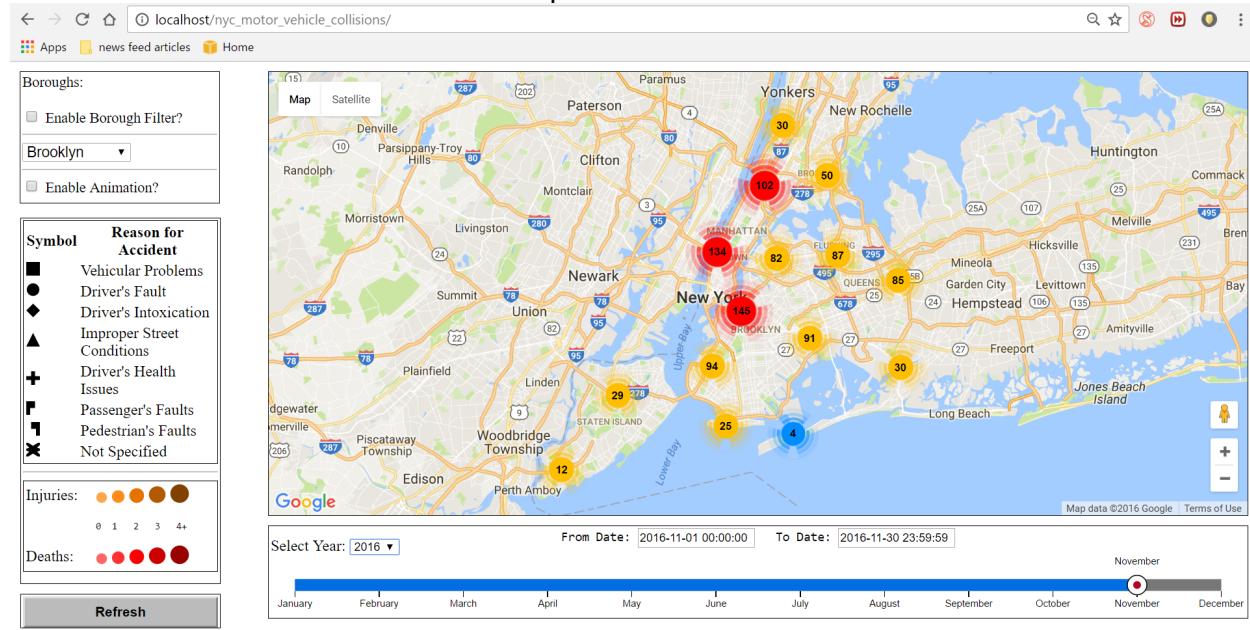
Also, the time slider shown in the mockup is modified to list months and years so that the user can quickly select a year and month and view the incidents that took place in that month, year.

## Modeling Collision Reasons

The dataset contains 48 distinct reasons for collision incidents. But after analyzing the reasons it was seen that multiple distinct reasons could be grouped as a similar incident. For example, there were incidents whose reasons were: Driver Inattention/Distraction, Driver Refused Right-to-Yield, Backing Unsafely. All these reasons denote that the incident occurred due to vehicle driver's fault. Therefore, these incidents are grouped under 'Driver's Fault'. All 48 distinct reasons have been grouped into 8 collision reason types. The mapping is performed in the `collision_reasons.js` file.

## Final Visualization

This tool is developed using JavaScript and D3.js framework. SVG elements were developed to model the 8 distinct collision reasons. Google Map's JavaScript API is used to mark those locations where the collisions have taken place. The screenshot of the tool is shown below:



The image shows the number and distribution of incidents in NYC for the month of November in 2016. The numbers in the markers indicate the number of incidents that have taken place near the area where the marker is located. Zooming in splits the clusters and the individual markers for each incident can be seen.

The elements in this tool are explained below:

### 1. Filters

When the Enable Borough Filter box is checked then only those incidents will be displayed for the borough selected in the drop-down box.

Boroughs:
<input type="checkbox"/> Enable Borough Filter?
Brooklyn ▾
<input type="checkbox"/> Enable Animation?

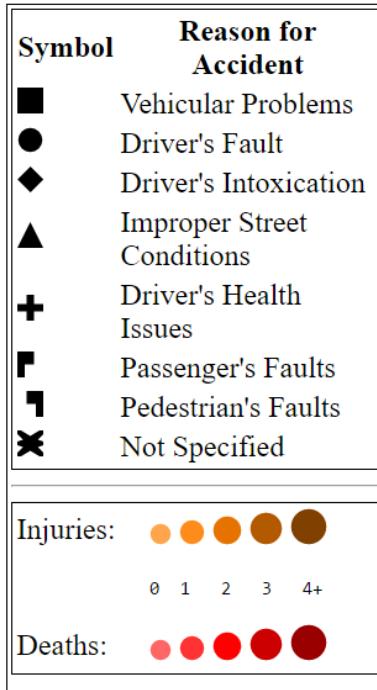
collisions are modeled using as shown in figure on the right

The severity of each collision is intensity and marker size. denote incidents which resulted markers denote the incidents that

Bigger the marker size and more the incident was more severe.

When the Enable Animation box is checked then the animation is enabled where collisions are displayed for each month starting from January 2015 to December 2016.

## 2. Legends



Eight distinct reasons for eight different SVG symbols side.

modeled using both color Orange-colored markers in injuries and red-colored resulted in deaths.

intense colors denote that

## Findings

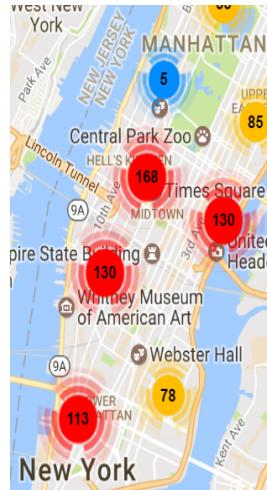
By analyzing the markers displayed in the tool with different combinations of filters we can find answers to analytical questions. The findings in this project are listed below:

### 1. Motor vehicle collisions are more prevalent in which parts of NYC?

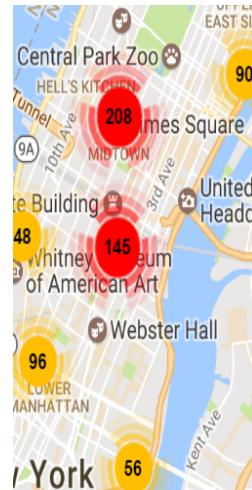
The collision incidents in Manhattan for different month, year combinations are analyzed:



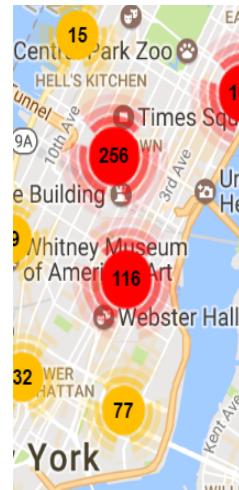
June 2015



September 2015



December 2015



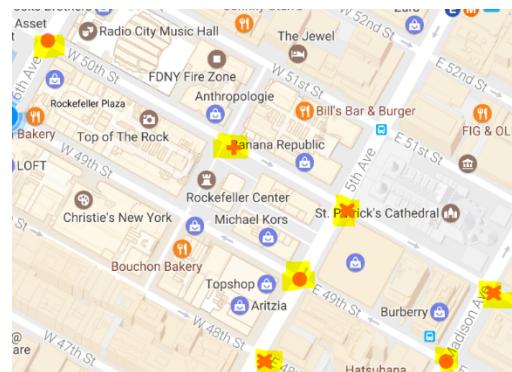
Looking at the screenshots for collisions in Manhattan for four months in 2015 revealed that the area surrounding Time Square is a host-spot for collision incidents.

## 2. What are the common causes of incidents in an area?

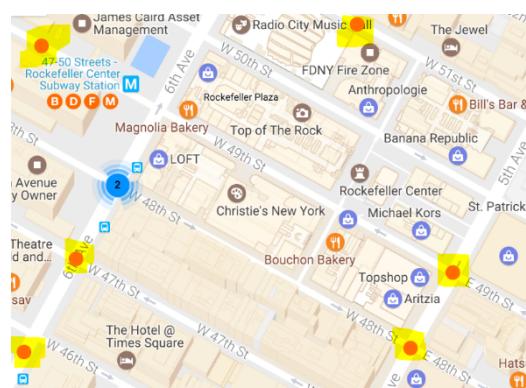
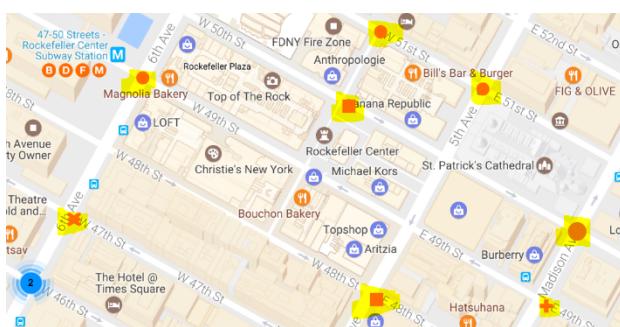
The area surrounding the Rockefeller Center in Manhattan showed the maximum number of incidents. Therefore, analysis of collision markers is performed near this location for different month, year combinations:



April 2015



June 2015



January 2016

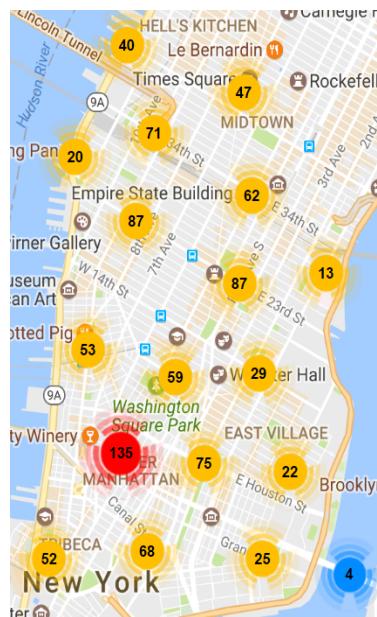
October 2016

Most of the markers are circular and square. This means that most incidents reported over time near Rockefeller center were caused by Driver's Faults and Vehicular Problems.

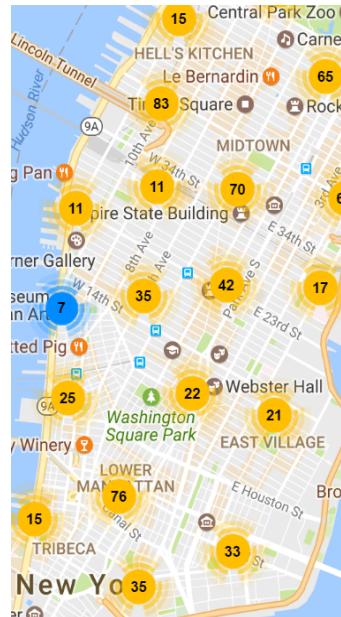
### 3. What are the factors that influence the collision incidents in an area?

What are the factors that influence the collision incidents in an area?  
For this question, the same screenshots for question 2 are considered. Almost every incident reported is located at a **cross-junction**. This is a very valuable observation. Cross-junctions are the areas where the most activity takes place and the chances of errors increase. By keeping this observation in consideration, we can conclude that more work is required to be performed at the cross-junction areas. If proper channels are utilized to ensure proper functioning of cross-junctions, then the number of such incidents can be minimized.

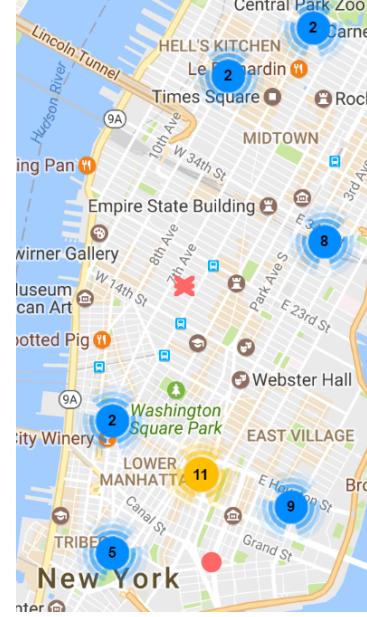
**4. Does the distribution of number of incidents in an area change with respect to year and month?**



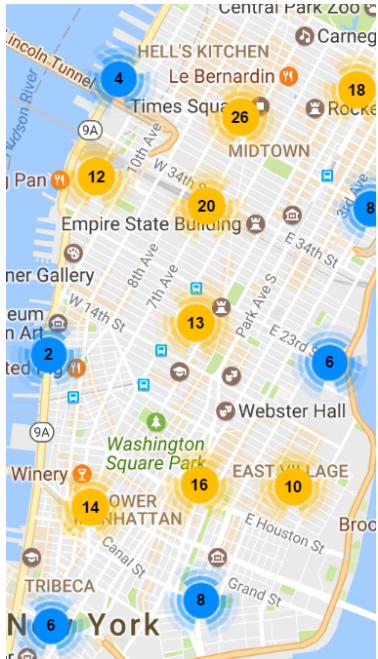
April 2016



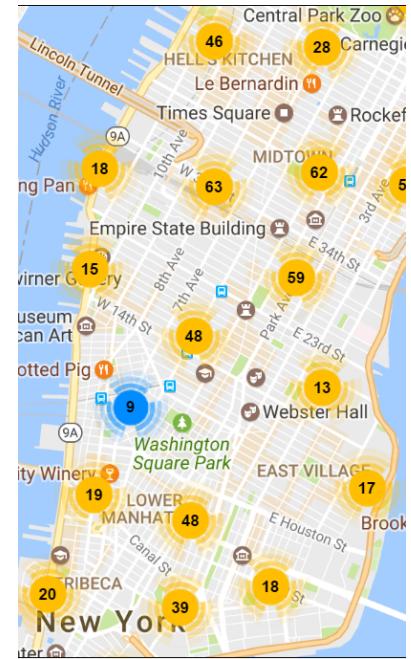
June 2016



August 2016



September 2016



October 2016

The Rockefeller Center has many collision incidents reported throughout the year. But in the year 2016, the number of collisions are high till June. But in the month of August the number of collisions decreased considerably. In September this number increased a little and from October, the numbers reached the high point again. This variation observed is very important. Factors influencing the traffic near Rockefeller center in Manhattan should be considered to understand this sudden drop in the number of collisions. Such observations will further help in determining corrective steps to decrease the number of accidents in NYC.

### **Limitations and Future Works**

The User Interface developed for this tool is at a very nascent stage and improvements are required to improve the users' experience. The time controls given in the tool are not easy to operate. Initially, I also wanted to analyze if there is a particular time in day when the collisions were high. To be able to analyze this, bar charts will be required. These corrections would be performed in the next version of this software tool.

In addition, to corrections, this tool in future will also be integrated with 311 complaints data set to find relationships among accidents occurred due to street problems with the street complaints reported at that time.