

AN INSIGHT INTO CRIME IN BOSTON

The background of the entire page is a grainy, high-angle photograph of the Boston skyline at night. The city lights create a dense, glowing texture against a dark sky. Notable landmarks like the Prudential Center and various skyscrapers are visible.

BOSTON CRIME REPORT

Cem Yilmaz x18191681

THE INTERSECTION OF CRIME AND DATA

The coming together of criminology and data science has created a largely under investigated space called criminal analysis. Criminal analysis or as Interpol names it, criminal intelligence analysis, is the practice of using data to understand crime trends and behaviour. According to interpol, "Timely and accurate intelligence analysis is key to understanding the inner-workings and driving factors of crime phenomena and criminal enterprises " (amorq.com, 2019). Crime analysis, is becoming a more and more accepted area of study. In fact, some police departments are now even using data marketplaces like kaggle to invite data analyst and scientists alike to dig deep into the data to understand patterns. The underlying data used in this report is also taken from Kaggle.

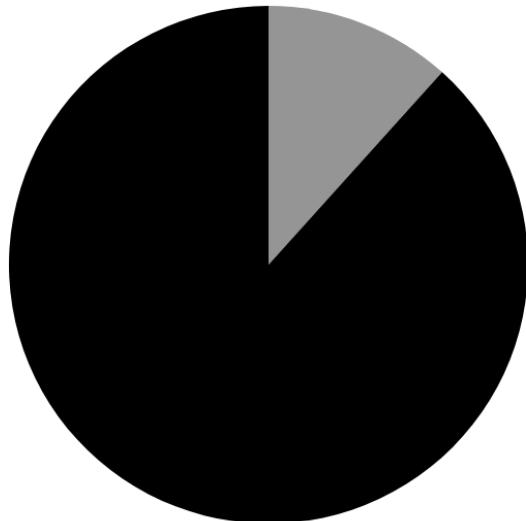
The specific crime report data studied for this report is Boston's crime report. Boston, known for its bloody history, is a city well known for its crime ridden past. Boston has been the home and breeding ground for many large gangs such as the Dorchester gang (Interpol.int, 2019). Even the Irish mobs have their share piece of the pie of committed crimes in Boston (Bostonherald.com, 2019). According to a more recent study, rate of homicides have increased by 50% in just one year from 2017 to 2018 (Universal Hub, 2019)

Data can be only non-mendacious in nature. This report will look at real data provided by the boston police force on reported crime and will attempt to reveal interesting insights and trends packed into a nicely fitted and well presentable report. This report will also host some interesting animated charts and graphs and some more static yet intriguing graphs.



SHOOTINGS

■ Fatal ■ Non-Fatal



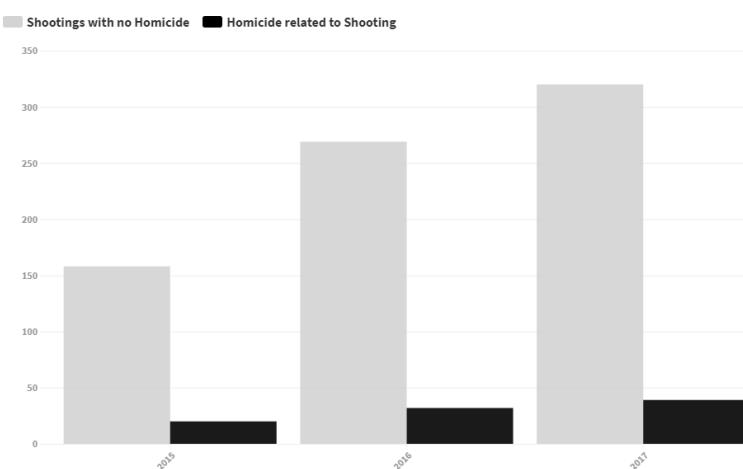
Boston is home to many shootings. In fact, there were 978 shootings in total, 114 of which caused human death. A pie chart and a barchart are the selected visual techniques to accompany these figures. In the pie chart, we can see that fatal shootings constitute a small greyish slice and the black represents the total shootings minus the number of fatal shootings. Pie charts are simple yet effective visual representations that tell a story quickly and effectively. Pie charts, in my opinion, work best when there are a limited number of variables. When there are more than 3 or 4 variables, the pie chart starts becoming hard to follow. As there were only two variables in this particular example, a pie chart seemed a good fit for the purpose of presenting this data.

In order to understand trends, one must analyse the frequency of occurrence of events over a specific period of time. In order to understand what direction the frequency of events are taking, events can be divided among the years and analysed for its percentage increase or decrease. In this case, there is a 51.12% increase in the number of crimes related to shootings from 2015 to 2016. Whereas, homicides or fatal shootings increased by 60% between 2015 and 2016. In this bar chart, one can see instantly how fatal shootings pair up to their non-fatal shooting related crimes counterpart instantly. By having each bar sit side by side arranged by their year of occurrence, it makes the task of understanding the data conceptually simple and quick. In 2017 there were 39 homicides related to shooting which is a multiplier of 1.95 fold more homicides related to shooting occurring in 2015. In more common language, this almost means that there were almost 2 deaths for every 1 person being killed in 2017 compared to 2015.

51%
INCREASE IN SHOOTING CRIMES
IN 1 YEAR

1.95X
TIMES INCREASE IN HOMICIDE

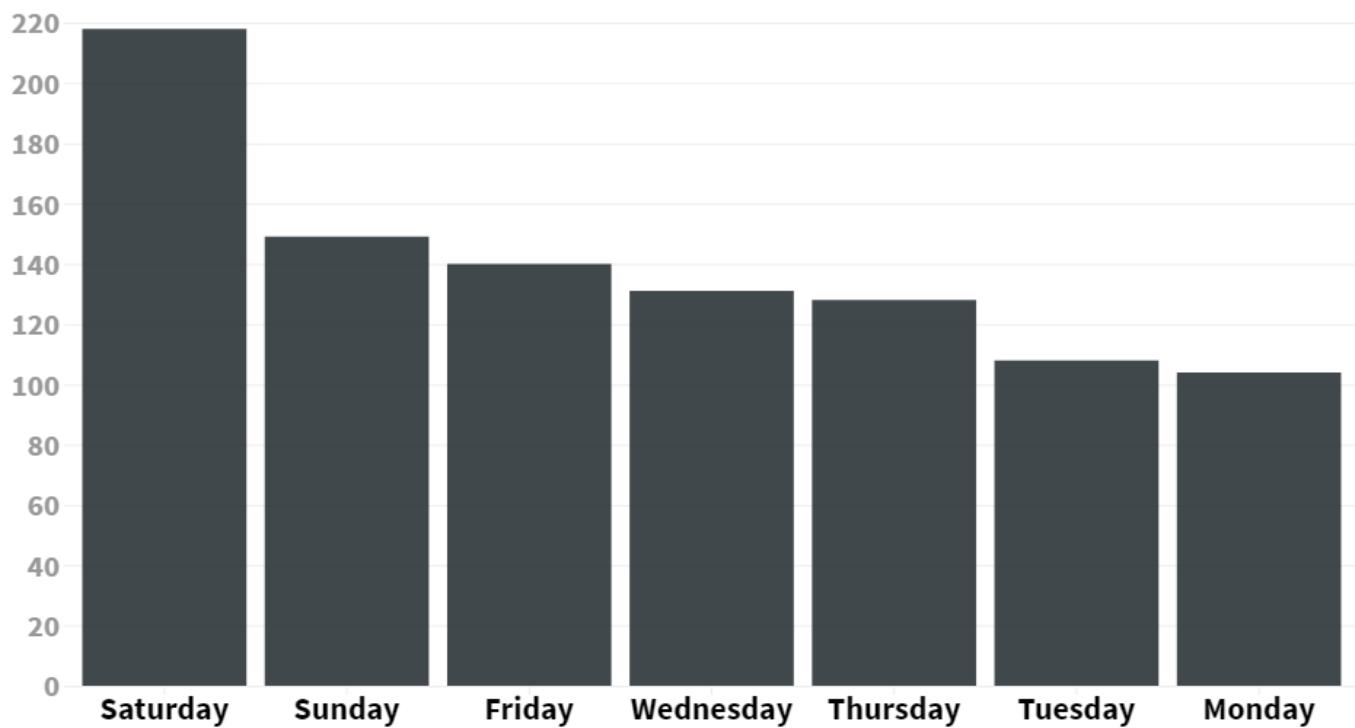
2017 witnessed 1.95 times as many more shootings than 2015.



SHOOTINGS - DAY OF WEEK

An interesting phenomena to understand when one has access to the date an event took place in is to see what on what specific day of the week is the event more likely and less likely to unfold. In terms of shootings taking place, the data has revealed to us that Saturday is the leading day which hosts to most number of deaths out of all the days of the week. In fact, there is 0.22 probability or a 22% chance that a shooting will take place on a Saturday, the biggest probability granted to any single day of the week. One could make many assumptions as to why shootings occur most frequently on a saturday. However, it would be very premature to do so given that we have only utilized one dataset and ultimately resource. Similar to the previous chart used to represent data, the bar chart used to represent this data is also a barchart. Visually, we can see that 'Saturday', is hovering over the other days of the week.

Shootings Grouped by Day of Week



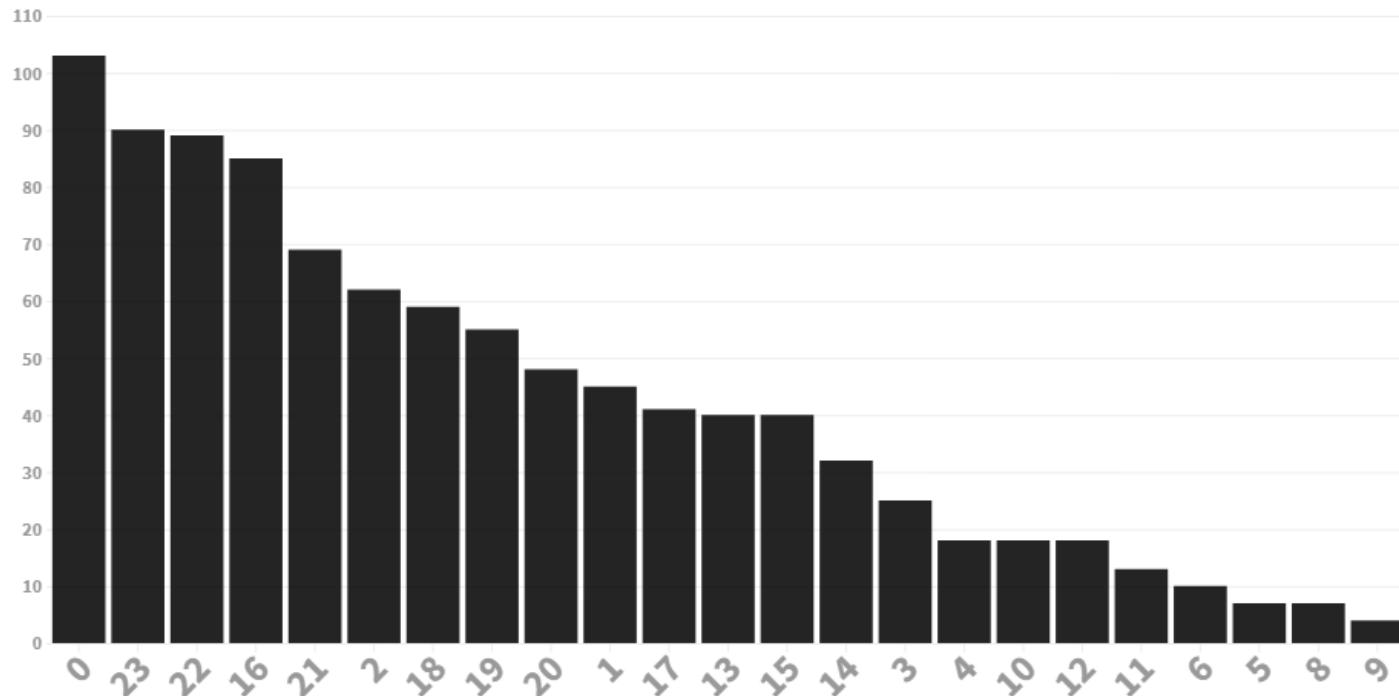
SHOOTINGS - HOUR OF THE DAY

Understanding what time of the day a shooting is more likely to occur may be beneficial to many different stakeholders involved. For example, knowing that there is a higher chance of shootings occurring at “x” O clock, may prompt city or district councils to have more medical staff in hospitals or even more police on patrol around that time. In this dataset, the data has shown us that there is the greatest chance of shootings to occur at midnight 12:00 am. In fact, there is a 10% chance that the shooting will occur at midnight. Similarly to the previously used chart, a bar chart will also be used to picture this. To make it more precise, only fatal shootings have been taken into consideration for this one.

Fatal Shootings per the Hour

24 hour clock format

■ Freq



00:00

TIME SHOOTINGS OCCUR THE MOST FREQUENTLY

1 in 10 shootings occur at midnight

09:00

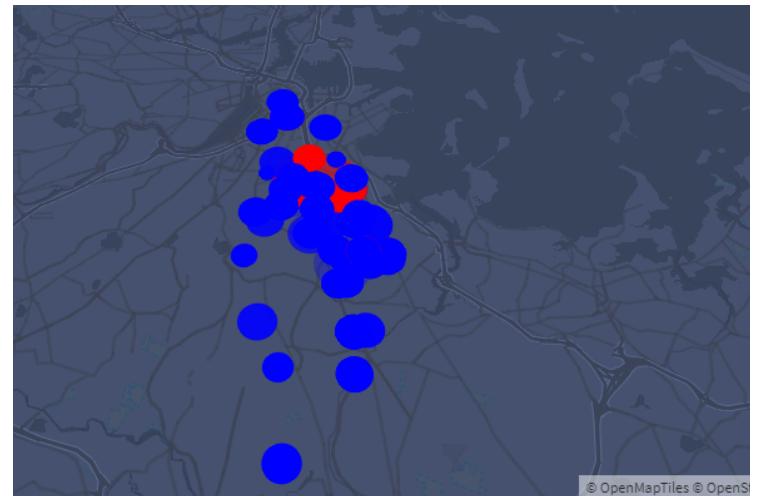
TIME SHOOTINGS OCCUR THE LEAST FREQUENTLY

There are the least number of shootings early in the morning.

INTERACTIVE CHARTS

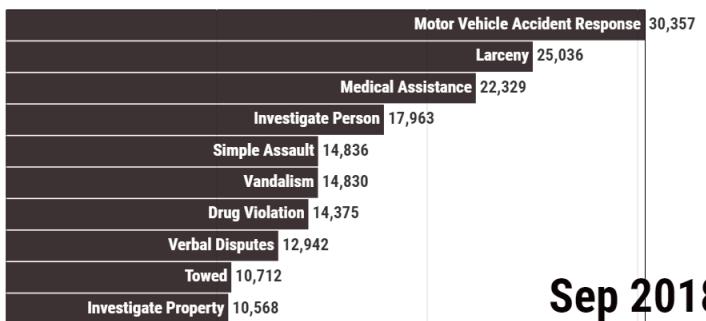
Shootings Mapping

When one has access to the latitude and longitude of data, fantastic visual representations can be produced. Using flourish.app, I've been able to present all shootings taking place in a linear orderly accordingly to their specific location. I have colored fatal shootings in red to further elaborate its intensity of seriousness. Here is the link to the published version. Find screenshot of it below. <https://public.flourish.studio/visualisation/584201/>



Crime Type Race - Boston
Crime Race

[Replay](#)



Bar Race

Another interesting way of depicting the frequency of occurrence of crimes when one has time variable as well is to use form it into a bar race. Flourish.App provide aesthetically appealing designs just for this. Here at link, you may find the public version of this interactive graph in action. <https://public.flourish.studio/visualisation/584293/>

SHOOTINGS - GROUPED BY DISTRICTS IN BOSTON

Understanding where shootings most take place can have a lot of impact. One may make the assumption that districts with high number of shootings will have relatively low value properties, reduced amount of amenities such as schools possibly, entertainment facilities, businesses and so on. It would definitely be interesting to see if these assumptions are correct by studying other data corresponding to these districts and to understand if they suffer from any form of inferiority relative to other districts. In this dataset, it turns out that the b3 districts hosts the highest number of shootings. In fact, over a third of all shootings occur in district b3 which is quite a large amount. Knowing this, would help police forces locate a shooting incident quicker, its root causes, who's behind the shooting and so on. A bar chart is used to deliver this data in a visually aesthetic fashion.

35%

OF ALL SHOOTINGS TAKE PLACE IN ZONE B2

b3 is the most common district for shootings

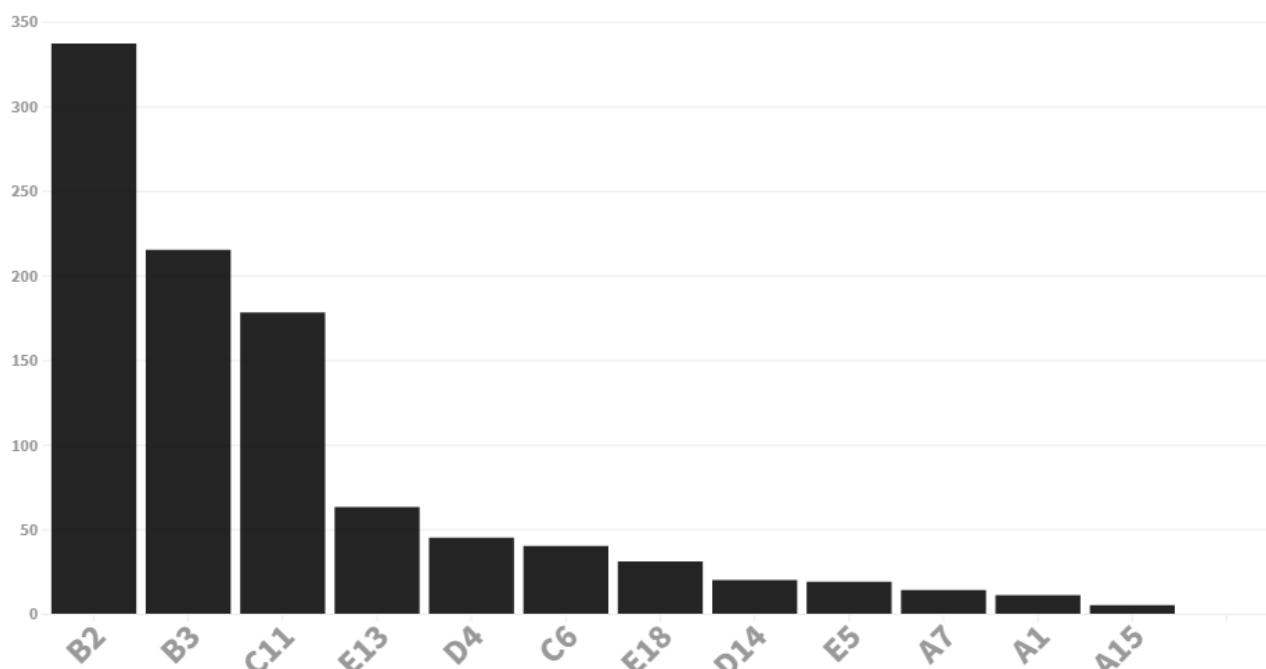
1 IN 3

FATAL SHOOTINGS TAKE PLACE IN ZONE B2

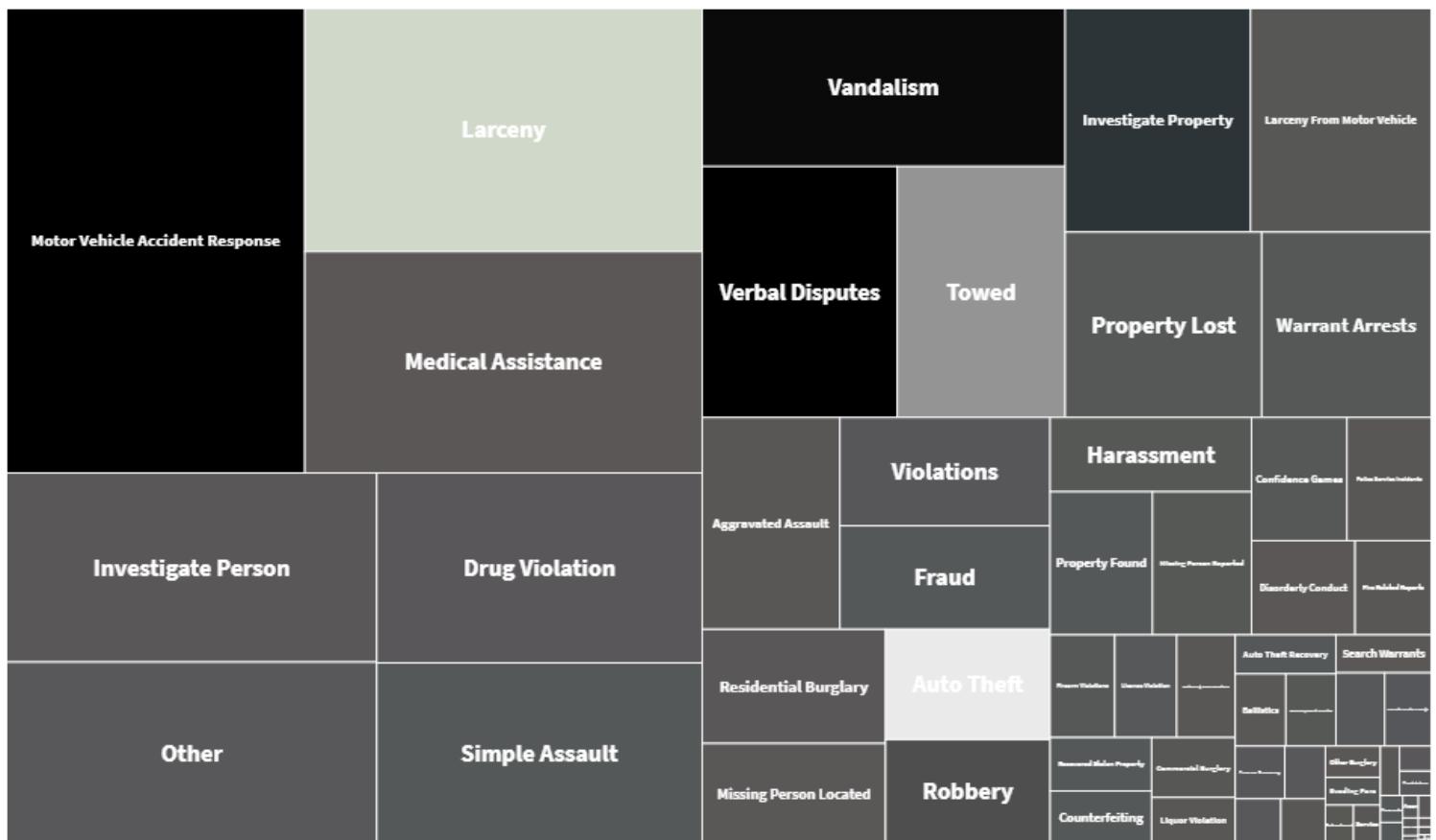
b2 overall hosts the largest portion

Shootings per District

■ Freq



MOST OCCURRING CRIMES



Understanding the frequency of occurring crime can help one identify what type of crime is most likely to occur using machine learning algorithms such as decision trees. Understanding frequency of occurrence will enable many data practitioners alike to join forces and come up with interesting insights and predictions. Here below is a treemap of crimes scaled to their frequency of occurrence. Treemaps are straight-forward, simple and easy to comprehend. In this treemap, we can see that the most occurring report is motor vehicle accident reports. Interestingly enough, the correlation rate between vehicle accident reports and automobile thefts is .94 which signify strong correlation. There may or may not be a relationship between the two.

PREDICTIONS

Predictions backed by historic data can be very accurate.

Machine learning techniques involving categorical variable driven techniques such as naive bayes or even aprioris algorithm otherwise knowns as association rules may be used to understand and predict what crime type is most likely to occur based on other factors. Decision trees may also prove fruitful in identifying patterns and creating predictions. However, for the purpose of this project, soft statistically assumptive predictions only are presented. More harder techniques such as exponential smoothing or other time-series analysis could be made to create predictions. Assumptive predictions are as follows. For example, prostitution increased between 2017 and 2018 as it rose up by 37%. This may suggest that it will keep increasing. It is however, open ended in nature. Bomb hoax are down 72% between 2018 and 2017 so is residential burglary decreasing by 14.33% in comparison to 2017. A line chart is used to depict their trend. There is a line chart given to each crime type. The Y axis is constant for all 4 variables. This places the frequency of occurrence of each crime type relative to each other.

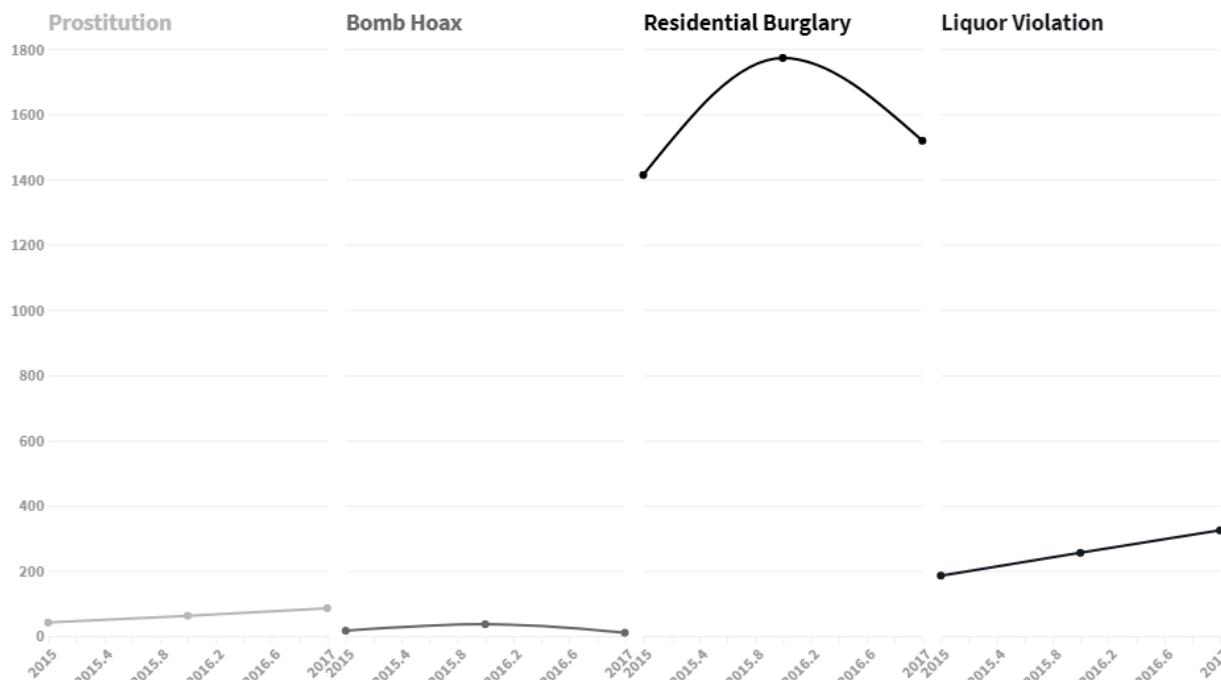
37%

PROSTITUTION ROSE BY 37% BETWEEN 2017 AND 2018.

72%

BOMB HOAX REPORTS FELL BY 72% BETWEEN 2017 AND 2018.

Crime Trends



A REPORT TO BE CONTINUED..

Representing this dataset by means of static and animated graphs have been a lot of fun. The dataset being so rich in data has helped immensely. Had I more time, I would like to experiment with presenting visuals to represent machine learning backed insights. Displaying statistical predictions such as simple moving averages or other time series analysis based predictions would be definitely interesting to play around with. Given the latitude and longitude of data, I can only imagine there are many creative ways in which one can display data. I would definitely be interested in using other. I would have liked to replace the bottom graph on the infographic with something a little more suitable.

The colors and textures seem to not fit well with the rest of the items on the infographic. Overall, however, I am of the opinion that this project turned out quite nice and most important of all visually satisfying.

