**GUN DEATHS IN AMERICA**

AIT 580 - Introduction to Big Data and Analytics

Group 1

Final Project Report

Siddharth Sajjansingh Sandu (G 01414077)

Vedant S. Limbhare (G 01413578)

Vijayalakshmi Padmanabhan (G 01337140)

**ABSTRACT**

The gun violence dataset provides an extensive collection of instances involving the usage of guns in the United States. The dataset contains details about each incident's location, date, gun type, number of victims, offenders, state, city, number of victims killed and injured, number of guns involved, and so on. It also includes information on the victims' and perpetrators' ages, genders, races, and other characteristics. The dataset spans several years, from 2013 to 2018 and is consistently updated with new data as it becomes available. Researchers, decision-makers, and the general public can use this information to understand better the scope and makeup of gun violence in the US.

**INTRODUCTION**

Gun violence is a complicated and diverse issue with serious ramifications for public health and safety. Gun-related occurrences have garnered more media attention in recent years and generated discussions about gun regulation, mental health, and other related topics. Numerous organizations have gathered data on gun-related incidents to understand better the scope and character of gun violence in the United States. The analysis of one such dataset, the gunshot dataset, which offers comprehensive data on occurrences involving guns across the nation, will be the main subject of this report. We can learn more about the patterns and trends in gun violence by looking at this dataset, and we may even be able to develop some prevention or mitigation tactics. We will also come up with the analysis patterns of the gunshot occurrences, who have been affected the most, how many have been injured, how they have been attacked, whether they were attacked based on their gender, race, ethnicity, what is the motive behind the attack.

This paper will concentrate on raising awareness of gun violence by analyzing a dataset of gunshots that contains comprehensive data on instances involving guns. By examining this dataset, we can learn more about the trends and patterns of gun violence and the conditions under which these episodes occur. In addition to outlining the dataset's methodology, breadth, and limits, this report will highlight significant discoveries about gun violence in the US. Additionally, it will evaluate the effects of gun violence on society, the economy, and public health and consider possible remedies. In the end, this report aims to increase awareness of gun violence and motivate communities and people to take action to avoid and lessen the harm that firearms do. It focuses on the below research questions:

# **RESEARCH QUESTIONS**

1. What are the temporal patterns of gun violence incidents in the United States? Are there certain months, days of the week, or times of day when incidents are more likely to occur?

2. What are the geographic patterns of gun violence incidents in the United States? Are certain states, cities, or neighborhood’s more likely to experience incidents?

3. What are the factors associated with the severity of gun violence incidents (e.g., number of victims and fatalities)? Are certain types of incidents more likely to be deadly than others?

4. How have patterns of gun violence changed over time (i.e., from 2013 to 2018)? Have there been any notable trends or changes in the types or severity of incidents or the demographic characteristics of victims and perpetrators?

**MATERIALS AND METHODS**

With this initiative, we hope to change that by making a database of over 260,000 instances of gun violence, along with precise information about each incidence, available in CSV format. We anticipate that this will make it simpler for statisticians and data scientists to investigate gun violence and make accurate projections about future patterns.

Between January 2013 and March 2018, inclusive, data for all incidences of gun violence that were reported in the US are included in the CSV file.

**DATA CLEANING**

Data cleaning is locating and eliminating database mistakes, discrepancies, and inaccuracies. The data quality can significantly impact the accuracy and validity of any insights or conclusions derived from the data, making this a crucial phase in the data analysis process.

There are often numerous phases involved in data cleansing, including:

●     **Data auditing:** Analyzing the dataset for problems such as duplicates, outliers, and missing numbers.

●     **Data Cleaning:** Correcting or eliminating errors, inconsistencies, and inaccuracies from the dataset. This may entail adding missing values, eliminating duplicates, and fixing formatting errors in the data.

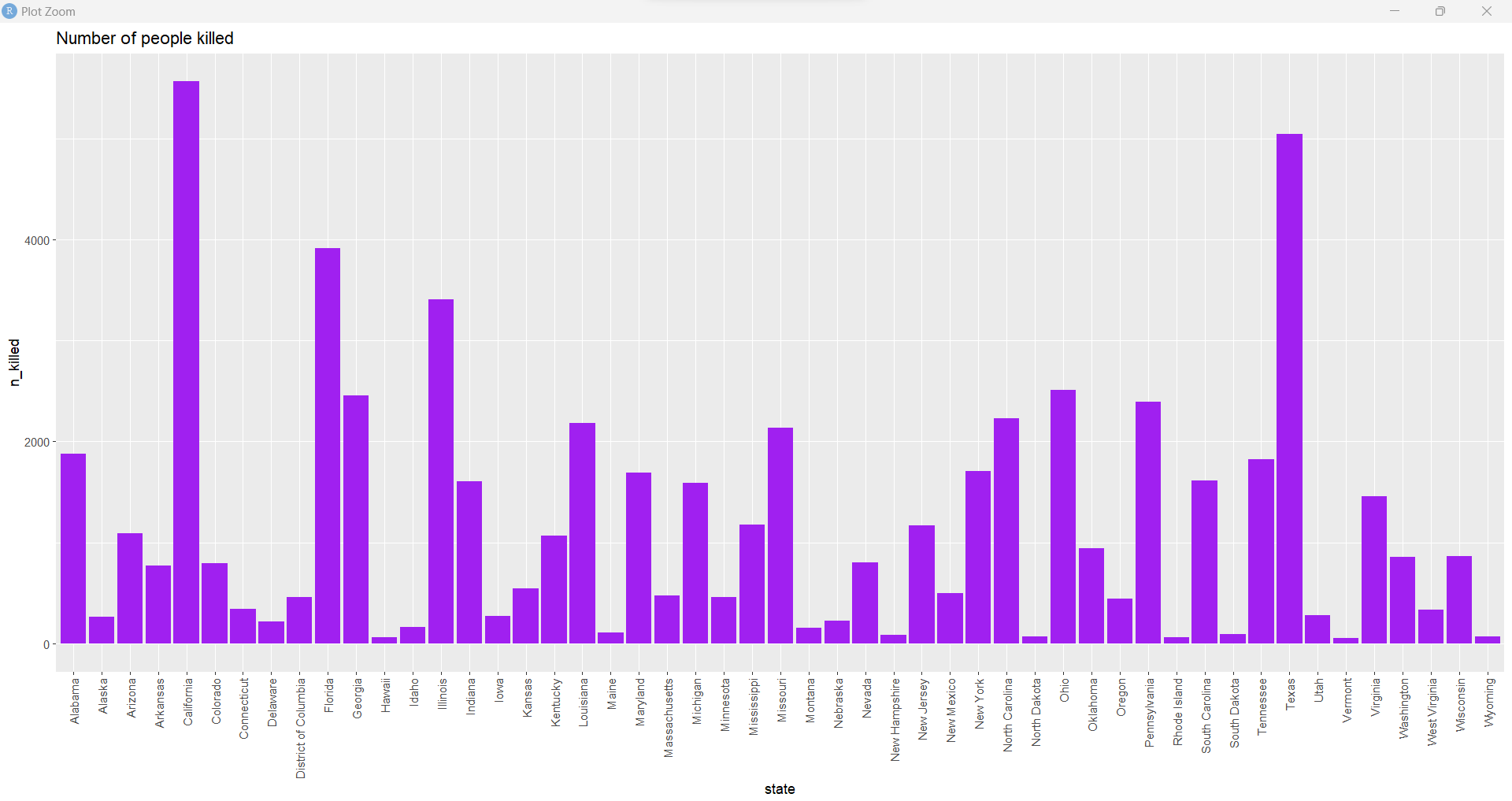
●     **Data validation:** Ensuring that the cleaned dataset complies with the necessary criteria and is appropriate for the analysis that will be performed.

The quality and dependability of data analysis can be increased with effective data cleansing, resulting in more precise insights and better decision-making.

Due to the sensitive and arguably contentious nature of the data involved, cleaning a gun shoot dataset can be difficult. In our dataset, we have performed the data cleaning for the null columns. Also, we have eliminated the unwanted columns delaying the data visualization process during the project flow. The following steps can be used to clean a gunshot dataset.

To be more precise, we have changed the null values with NA in the categorial column. The empty columns in the numerical column are changed with a “unk”, to avoid issues during the data visualization.

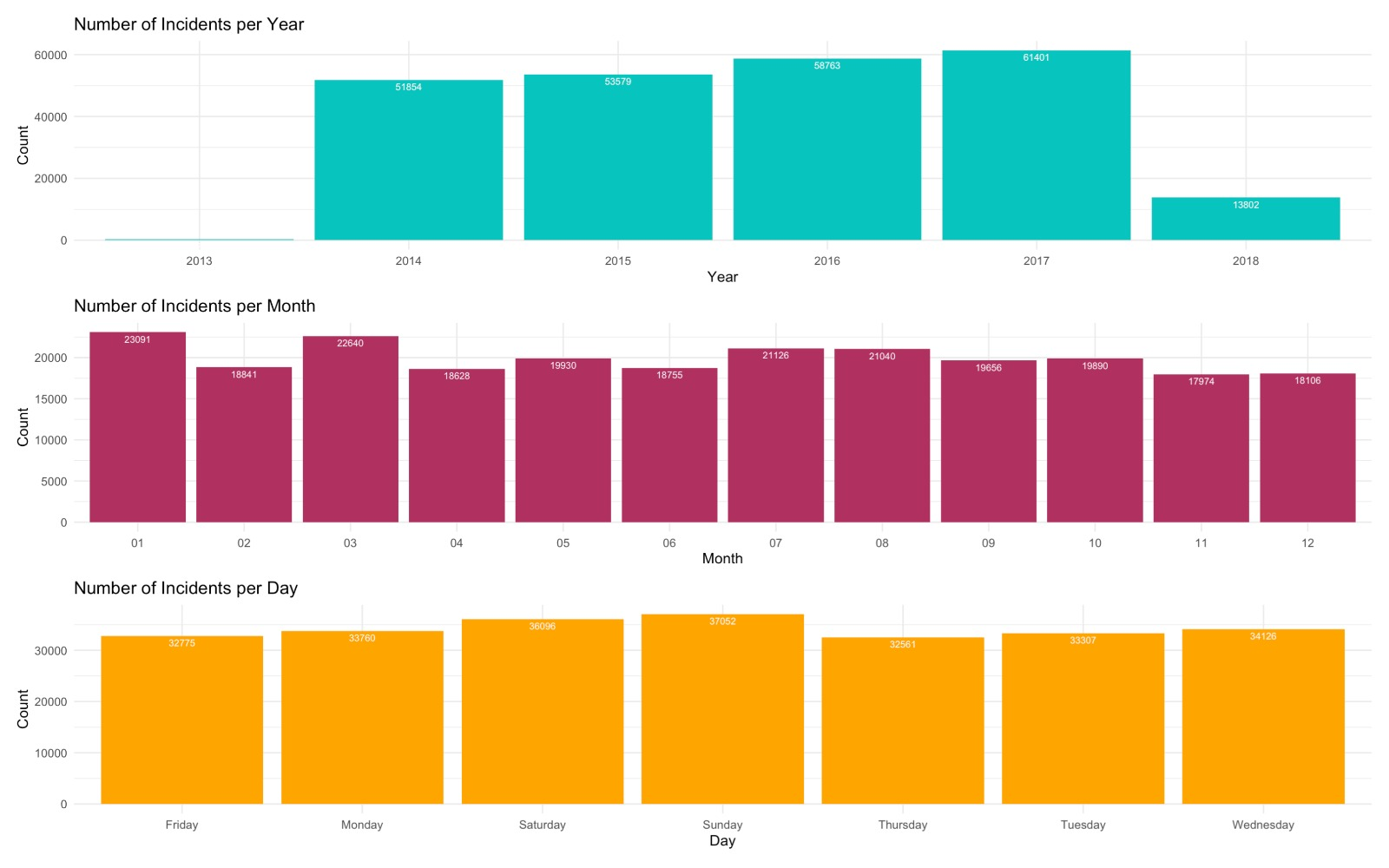
**DATA VISUALIZATION:**



The most gun shootings in recent years have occurred in California has the highest occurrences of 16306-gun shoots that have been reported. Statistics show that the state has experienced a concerning number of gun-related occurrences, with Chicago being among the most badly hit cities. The high frequency of shootings has been attributed to elements including gang violence, easy access to firearms, and ineffective gun control regulations. Gun violence has increased in the state in recent years, with young people making up most of the victims. Policymakers, law enforcement organizations, and community leaders have all taken efforts to address the situation after it sparked concerns. California still struggles with the issue of gun violence despite efforts to reduce it. The high rate of shootings in these states is a result of a multitude of factors, including gang violence, drug trafficking, and insufficient gun control regulations. Policymakers, law enforcement organizations, and community leaders have all taken efforts to address the situation after it sparked concerns. Stricter gun restrictions, neighborhood-based initiatives, and gun buyback programs have all been put into place to reduce gun violence. To protect the safety and security of citizens in these states, more extensive steps are required because the issue still exists. To address this issue and lower the frequency of events involving guns, parties must work together.

**RESEARCH QUESTIONS IMPLEMENTATION:**

1. *What are the temporal patterns of gun violence incidents in the United States? Are there certain months, days of the week, or times of day when incidents are more likely to occur?*

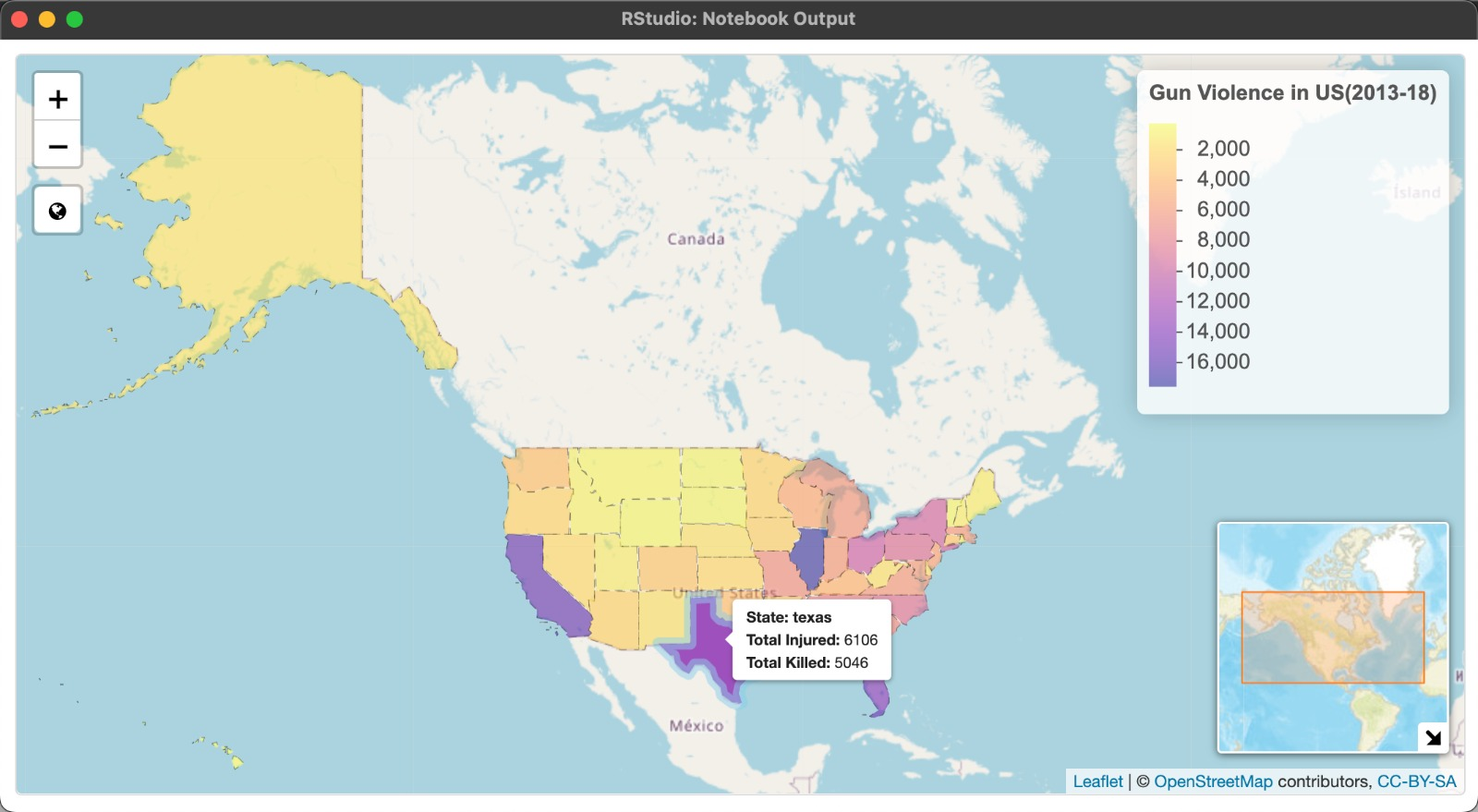


Incidents per Year: 1. The gun violence incidents per year show an upward trend from 51000 in 2014 to 61000 in 2017. 2. Although the data for 2018 is insufficient (only for the first 3 months of 2018), there are already 13000 incidents have taken place.

Average Number of Incidents per Month Over the Years: 1. From the graph we could see wavered readings. As there is a change in readings every month. It started off with 23091 in the first month, which recorded the highest count of that year and since then the graph has had an upward and downward movement. In the month of November which marks 11th in the above graph, this recorded the lowest count of 17974.

Average Number of Incidents per Day over Years: 1. From the graph during the phase of the week. We can clearly see that during the weekends the numbers are skyrocketing which clearly depicts the victims being targeted in mass numbers in a public place gathering.

1. *What are the geographic patterns of gun violence incidents in the United States? Are certain states, cities, or neighborhoods more likely to experience incidents?*

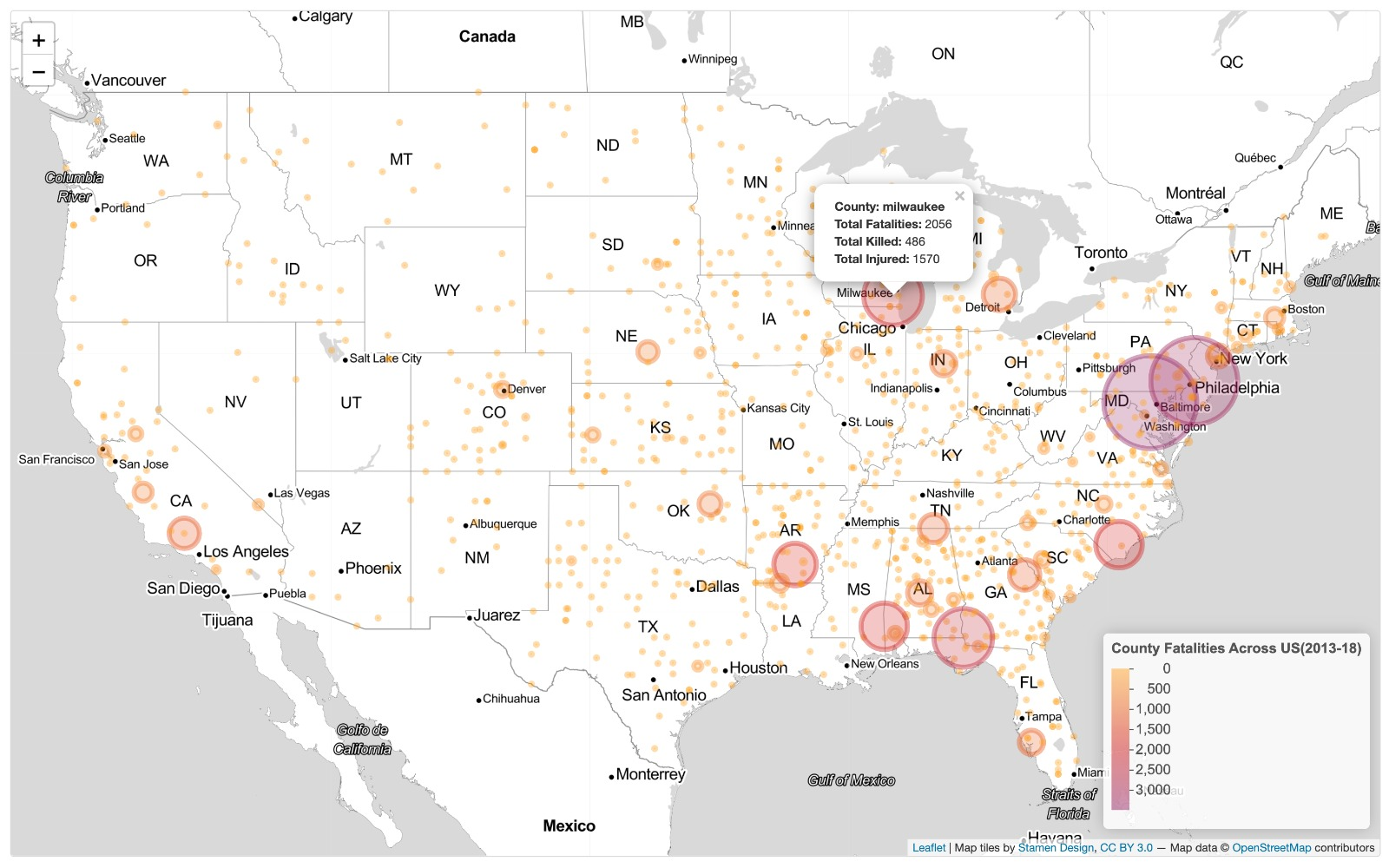


From the above choropleth map, there are high chances of crime happening in states such as Illinois which has the highest record of Incidents followed by California, Texas, and Florida. From the above-depicted map displayed, we can easily control and take measures to get rid of the gun license provided to the public. Also, we can read other places and get sufficient data on how the crime rate is, whether the rate is kept under control, and how the rate of crime is increasing. Is it a drastic increase or a gradual upward movement?

This provides a proper overview of how measures can be taken and what type of rules can be imposed.

For the above choropleth map, we have made use of “leaflet” library in R and made use of shapefile for US state boundaries. This has helped us in getting the correct representation of each US state map and the count of Gun Violence taken place in that specific state. Further we have used color coding with the help of addition number of incidents and the number of killed which ranges from < 2000 to >16000. We have added additional functionalities such as Mini Map and easy zoom toggle button for effective visualization.

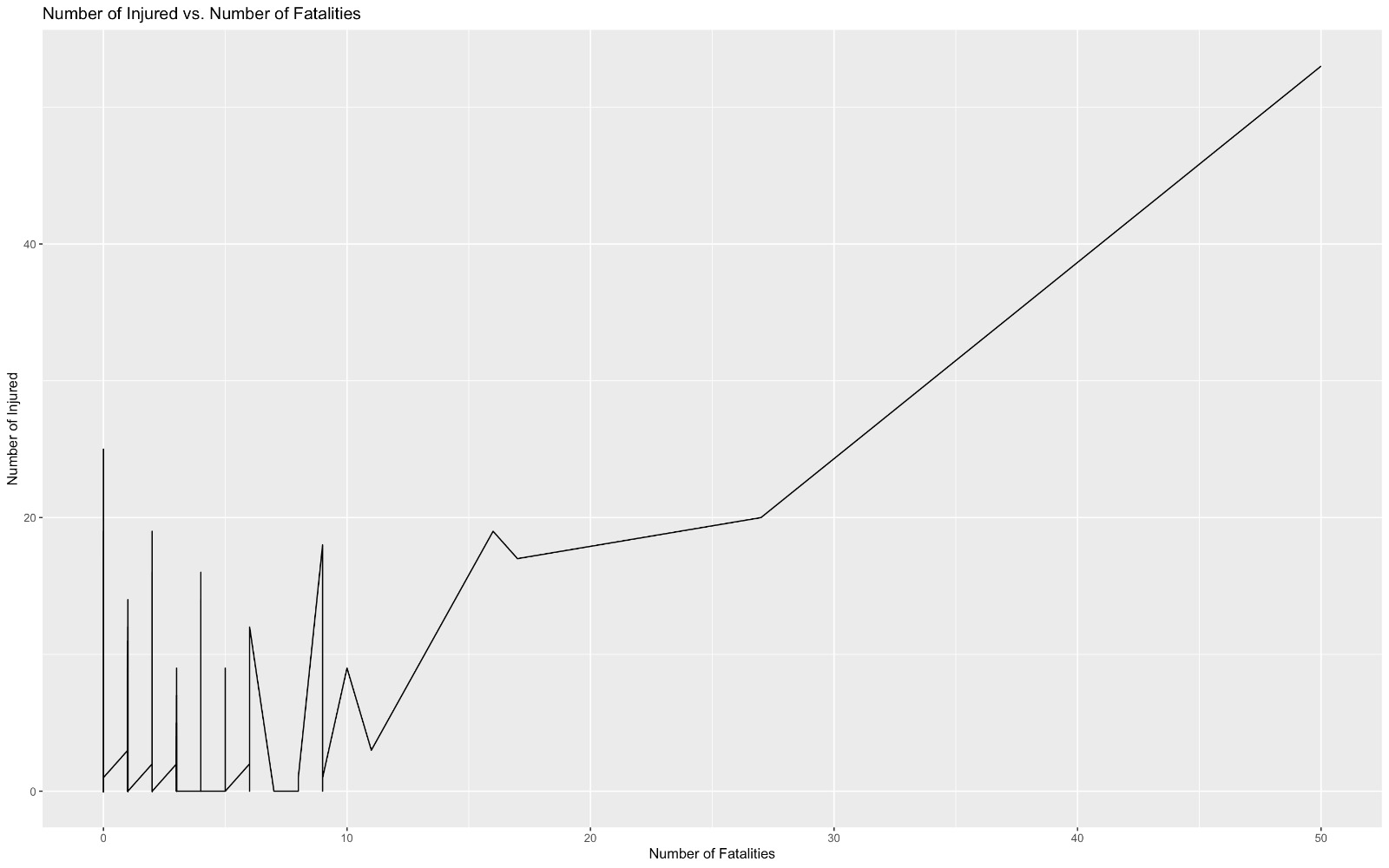
By using the county\_city column, we tried to map a similar choropleth map for each county and tried to represent the areas with high volume of incidents occurred during the span of 2013-2018.



From the above leaflet, we can see a different scenario where even though states such as Illinois and California had the most incidents across the state, when we compare across county, the highest count of incidents occurred in Baltimore followed by Philadelphia and Houston. With the help of both these graphs, we can see the various aspects of Geographical patterns of Gun Violence Incidents across the US.

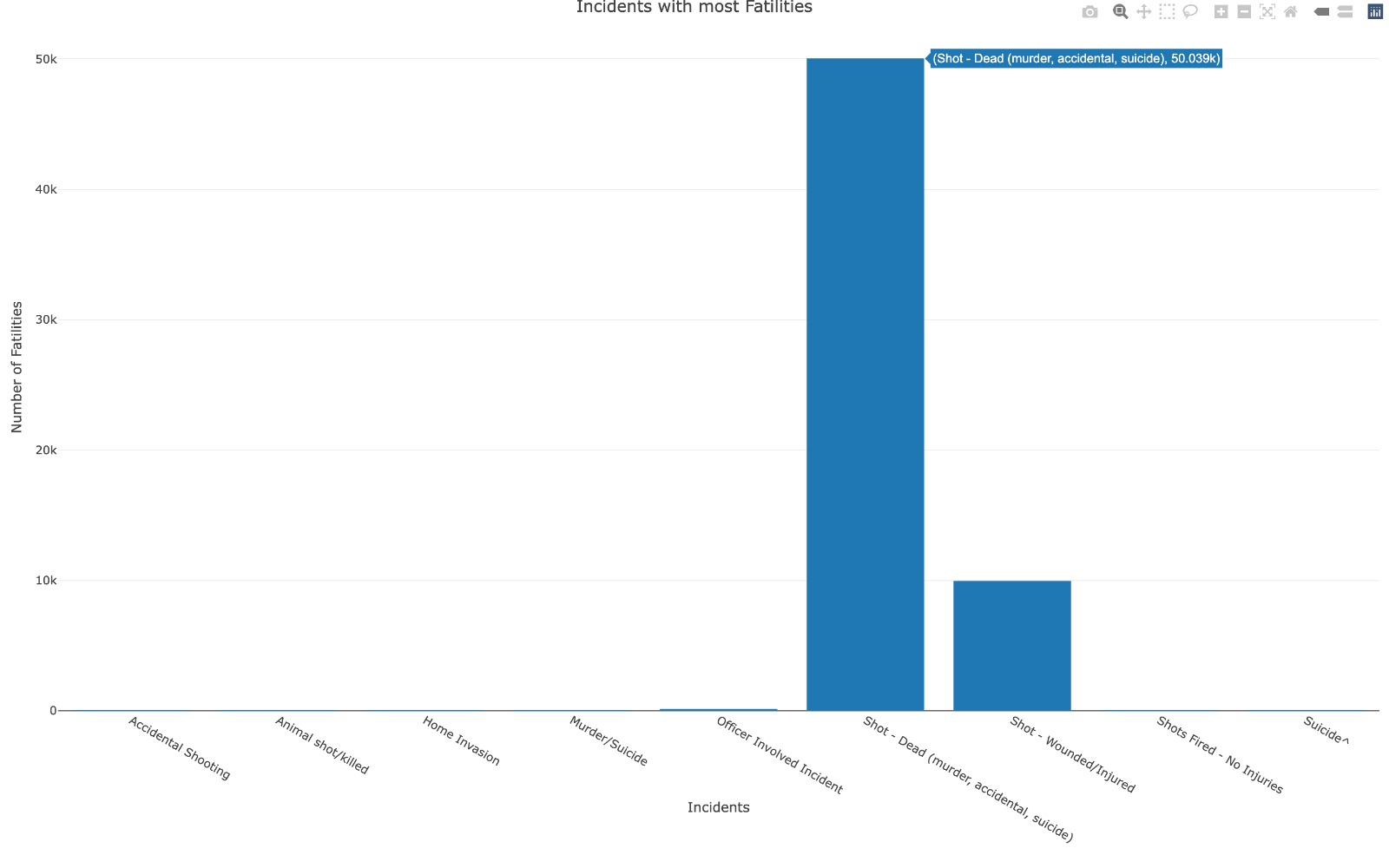
Both above choropleth maps help us in getting a clear representation of states and county’s which are at high risk of a Gun Violence incident happening. Effective measures such as strict gun license’s, secretly notifying devices in schools, mental health camps can be organized in the above hotspot areas.

1. *What are the factors associated with the severity of gun violence incidents (e.g., number of victims, number of fatalities)? Are certain types of incidents more likely to be deadly than others?*

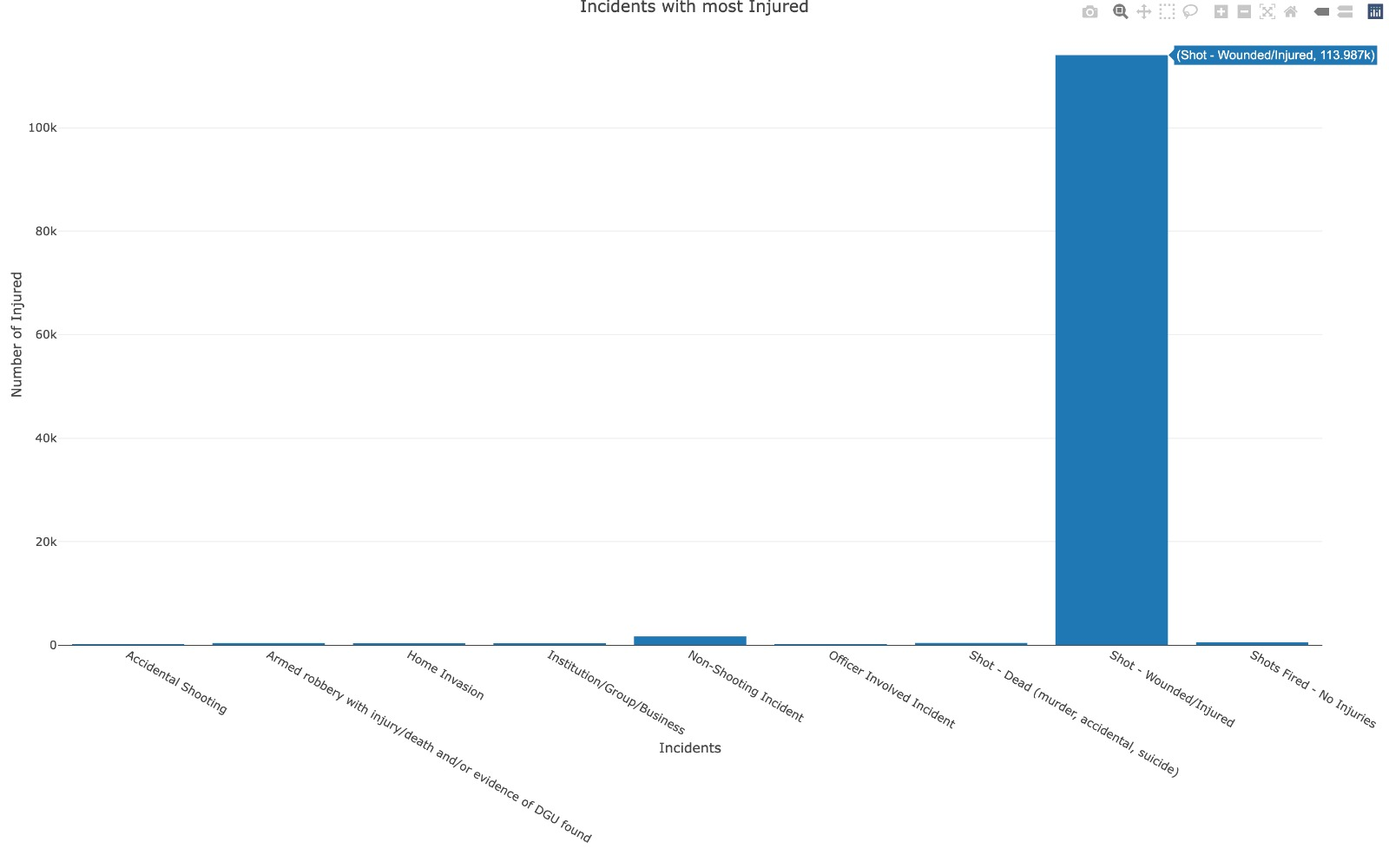


The positive trend, as shown in the figure, depicts that the number of fatalities and injured people are directly proportional. The fatalities that occurred due to the gun shoot were tremendously increasing. These situations take place in different scenarios such as discrimination in race, ethnicity, or even illegal gun ownership. As you see in the graph, there is a steep inclination in the occurrences of fatalities and injuries.

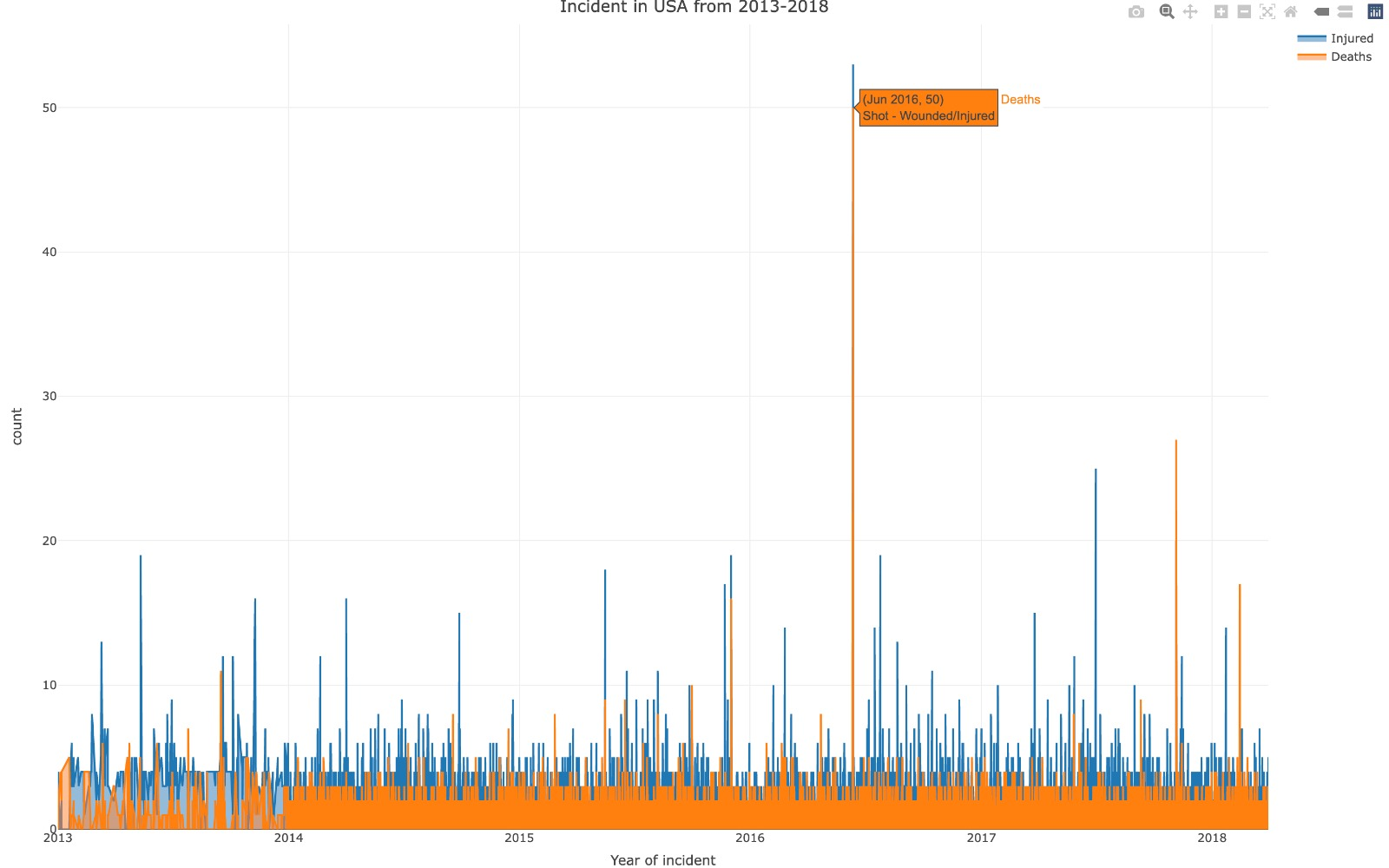
These alarming figures emphasize the need for ongoing initiatives to lessen gun violence and encourage responsible gun ownership. It is crucial to have thorough background checks and other safeguards in place to stop guns from ending up in the wrong hands, as well as to give those who might be at danger of killing themselves or others access to mental health services and other forms of help.



* This graph shows the number of incidents that took place between 2013-2018. According to the above graph, it is depicted that almost more than 90% of the gunshots were shot dead.
* The number of fatalities caused by the gunshot and those who are dead because of the incident is terribly high with a count of 50,039 deaths. Despite all the other instances of gunshot incidents, this 50K of death is caused mainly by accidental gunshot, murder, or gun fired during suicide.
* Giving gun owners instruction on gun safety is one method of reducing the number of gunshots. This can include guidance on safe gun handling, storage, and knowledge of the dangers that might come with gun ownership. By limiting access to guns for people who are highly likely to conduct violent actions, such as those with a history of domestic abuse or those who have specific mental health disorders, gun control legislation can also play a role in reducing gun violence.

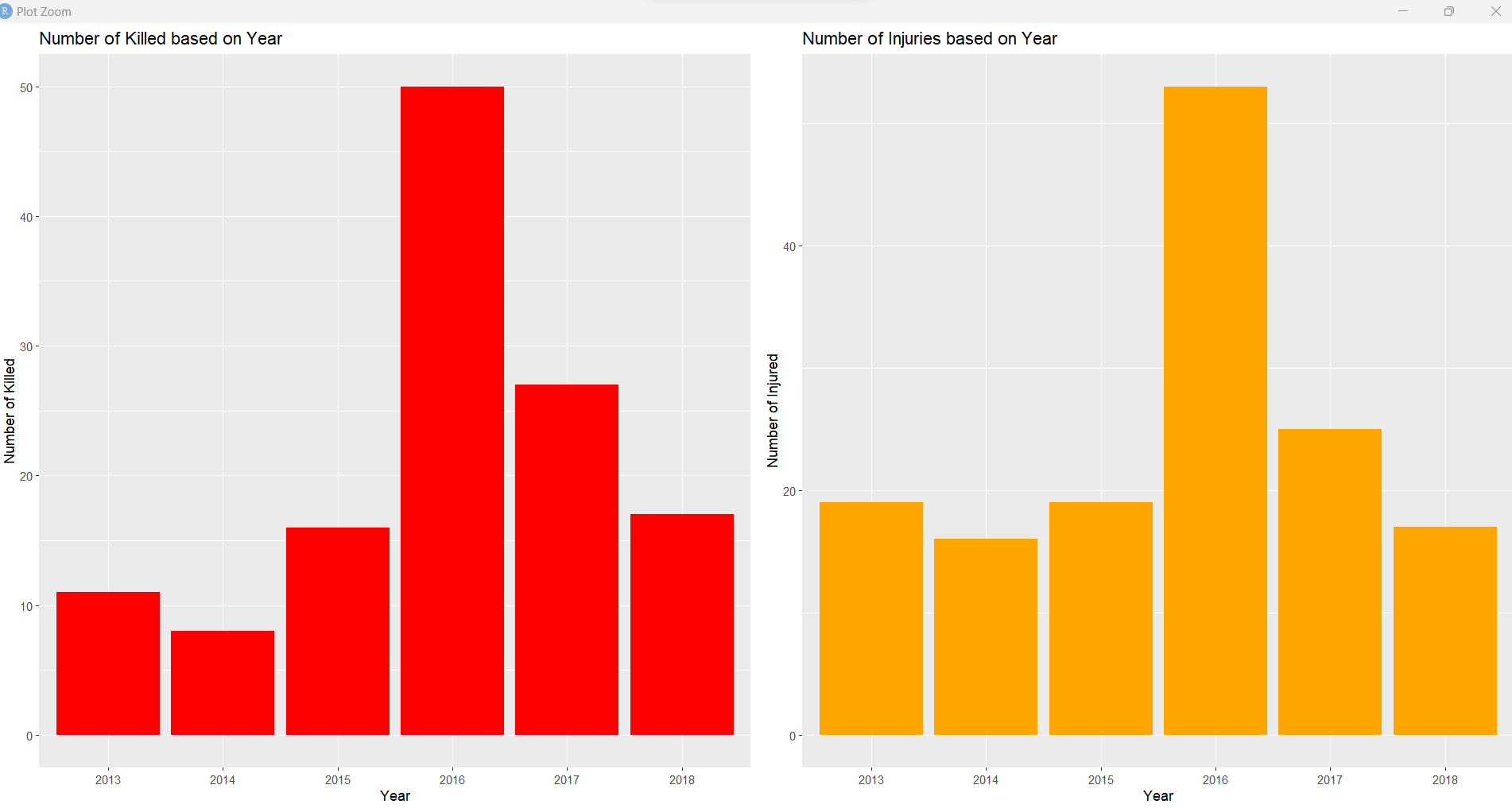


The graph depicts the number of people that were injured during a gun violence incident. From this graph, we can provide one comparison the number of people who are shot dead during a gunshot is way lesser than the number of people who are injured because of a gunshot incident. By looking at the graph we can see that around 113,987 people were left injured due to Gun Violence, whereas in the previous graph, we can see that the people who were shot dead because of a gun shoot were about 50,039 people. Both situations are heartbreaking, but if provided with proper education or awareness of the misuse of the usage of guns.



The graph illustrates the trend of incidents related to gun violence that occurred during a six-year period, from 2013 to 2018, and the corresponding number of people who were either killed or injured as a result of these incidents. The data shows that there has been a gradual increase in the frequency of gun violence incidents over the years, which is a concerning trend. The escalating number of fatalities and injuries is indicative of the detrimental impact of gun violence on the lives of people, and how it can cause harm and trauma to individuals, families, and communities. It underscores the urgency of addressing the issue of gun violence and implementing measures to prevent and reduce such incidents.

1. *How have patterns of gun violence changed over time (i.e., from 2013 to 2018)? Have there been any notable trends or changes in the types or severity of incidents, or in the demographic characteristics of victims and perpetrators?*



The above graphs depict the number of killed and injured during the year 2013 - 2018. It is depicted that according to the dataset, there was a greater number of kills in the year 2016 to an approximately 50 kills. Following this, the next year also has quite a lot of guns shoot occurrences to approximately 25+ kills.

The graph that depicts the number of injuries during the year 2013 - 2018 indicates that there were a high number of injuries in the year 2016 during a gun shoot to approximately 50+ kills. In addition to this, it is also depicted that the year 2017 also has the second highest gun shoot incidents. It is also found that the year 2013 and 2015 has the same number of incidents with the same pattern/trend.

On the positive side, it is found that the gun shots after the year 2016 gradually reduced in the next years, and since our dataset has data about gunshots in America from the year 2013 till 2018. According to the dataset, the last year 2018 had fewer gun shoots when compared to the previous years.

**CONCLUSION**

To conclude with our report regarding the gunshots in America, we have made all possible analyses through the dataset which we found and have understood that there are multiple reasons for a gunshot to take place and there weren’t any personal reasons involved behind every gunshot instance. AndSS we were able to analyze which state had the greatest number of attacks and how many were injured. To be more precise, we even plotted the month when there was the highest incident happened and how it happened. Gun violence has been a significant issue in the United States for decades, with an estimated 36,000 deaths annually attributed to firearms. With the help of the dataset, we found that from 2013 to 2018, there were 50k deaths and 110k injuries because of gunshots.

However, compared to other industrialized nations, the United States still has a high prevalence of gun violence despite several attempts to address the problem. This is caused by a variety of complex causes, such as the accessibility of guns, mental health problems, and a culture that frequently elevates violence.

Policymakers must adopt a multipronged strategy to solve this problem, including reasonable gun control laws, better access to mental health care, and a change in cultural attitudes about gun violence. Only then the United States will be able to start making strides toward reducing the senseless deaths brought on by guns.

**REFERENCES**

J. K. (2018, January 1). Gun Violence Data. Kaggle. Retrieved April 17, 2023

<https://www.kaggle.com/datasets/jameslko/gun-violence-data>

(2016, July 13). Gun Deaths in America. Data.World.

<https://data.world/azel/gun-deaths-in-america>

(2023, January 6). Tiger Shapefile. Census.gov