# DSC 530 Final Project

## Gun Violence

### Jessica Everett

8/11/2019

**Data Set: Gun Violence Data**

Retrieved from: <https://www.kaggle.com/jameslko/gun-violence-data/downloads/gun-violence-data.zip/1>

This dataset includes all gun violence incidences in the USA between 2013 and 2018. Its variables include: date, state, number killed, number injured, and number of guns involved in the incident.

**Hypothetical Question**

Using the descriptions of the elements of this data set, we can see that most gun violence occurs in Illinois. We can also see that most gun violence results in 0 deaths and 0 injured with 1 gun. For my hypothetical question, I would like to determine if the number of guns involved increases the chance of death or injury.

**EDA Outcome**

Using the histograms, it appears that most gun violence involves 1 gun and 0 deaths or injuries. More than one gun is less common and doesn't appear to increase the chance of injury or death. By looking at these histograms, it is obvious that the numerical variables include outliers. This is probably because mass shootings are rare compared with the typical incident. I removed the extraordinary incidents so we can view just what is happening on a typical, day to day basis.

After dropping the outliers, I reran the histograms to get a clearer picture of what is happening in the typical gun violence incident. The number of guns, deaths, and injuries are still mostly 1, 0, and 0 respectively, but after removing the outliers we can focus more on a typical episode of gun violence rather than the occasional extreme events. We can now see more detail in the histograms.

The scatterplot of number of guns and deaths shows that a large majority of gun violence occurs with 1 or 2 guns and between 0 and 1 deaths. The scatterplot of number of guns and injuries shows that a large majority of gun violence occurs with 1 or 2 guns and between 0 and 2 injuries.

The regression analysis shows that as the number of guns involved increases, the number of injured slowly decreases. At this, I will say that the answer to my hypothetical question, “Does more guns increase the number of injuries and deaths?”, is no. In fact, more guns may decrease the number of injuries and deaths, if only slightly.

**Challenges**

I would have liked to examine the states and dates more, but it was such a wide range of data that Jupyter Notebook would freeze every time I tried to do a histogram or any other analysis. I was able to determine the state most involved in gun violence (Illinois), and the date in the last five years with the most gun violence (January 1, 2017).

Some of the other variables in the data set could have been useful, but it was not in a useful format. For example, the genders of the people involved was combined into one variable. If the sex of the perpetrator had been isolated, I could have examined the difference between male and female perpetrators.

**Conclusion**

After removing the outliers of mass shootings, I was able to determine that most everyday gun violence usually only involves one gun. These incidents usually end with 0 or 1 deaths and 0-2 injuries. Injuries are more common than deaths. More guns slightly decreases the chances of injury or death.