The Statistics of Gun Violence within the United States of America

Nghi Huynh, Minh Mai, Joseph Tran

CSE 163

Professor Hunter Schafer

March 15, 2022

Table of Contents

Numbers correspond to research question

1. Finding patterns in instances of gun violence to ultimately determine plausible causes is important towards preventing the occurrence of future deaths. Are there particular states of the United States that are more susceptible to gun violence? If so, why could this be? List the factors that contributed most to the likelihood of the incidents for example, participant age or gender.

   Our Answer: From the data that we have gathered, we have found that the bigger the population, the more gun violence there will be. Illinois, California, Florida, and Texas had significantly more gun violence than most of the country. Chicago, Illinois was not surprising to see on the list as they have many gang-related issues.

2. Displaying how the occurrence of gun violence has increased over time helps us emphasize the need of urgency to combat this issue in America. How did the number of *injuries* and *deaths* by gun violence change as time progressed from 2013 to 2018? What societal events could have influenced the increase, decrease, or stability of injury/death rates?

   Our Answer: Ignoring the outliers of 2013 and 2018, we found a slow, but positive trend in the number of gun violence as the years progress.

3. Although preventing all occurrences of gun violence in the world is pragmatically impossible, we wish to regulate it as much as possible. Analyzing our country in contrast with others is a good way of identifying of how well our community is contributing to preventing gun violence across the world. Making this comparison can also lead to revealing what our country needs to work on regarding gun violence. What is the worldwide relationship between gun ownership and gun violence? Do countries with a higher gun ownership rate also have a higher rate of gun-related deaths?
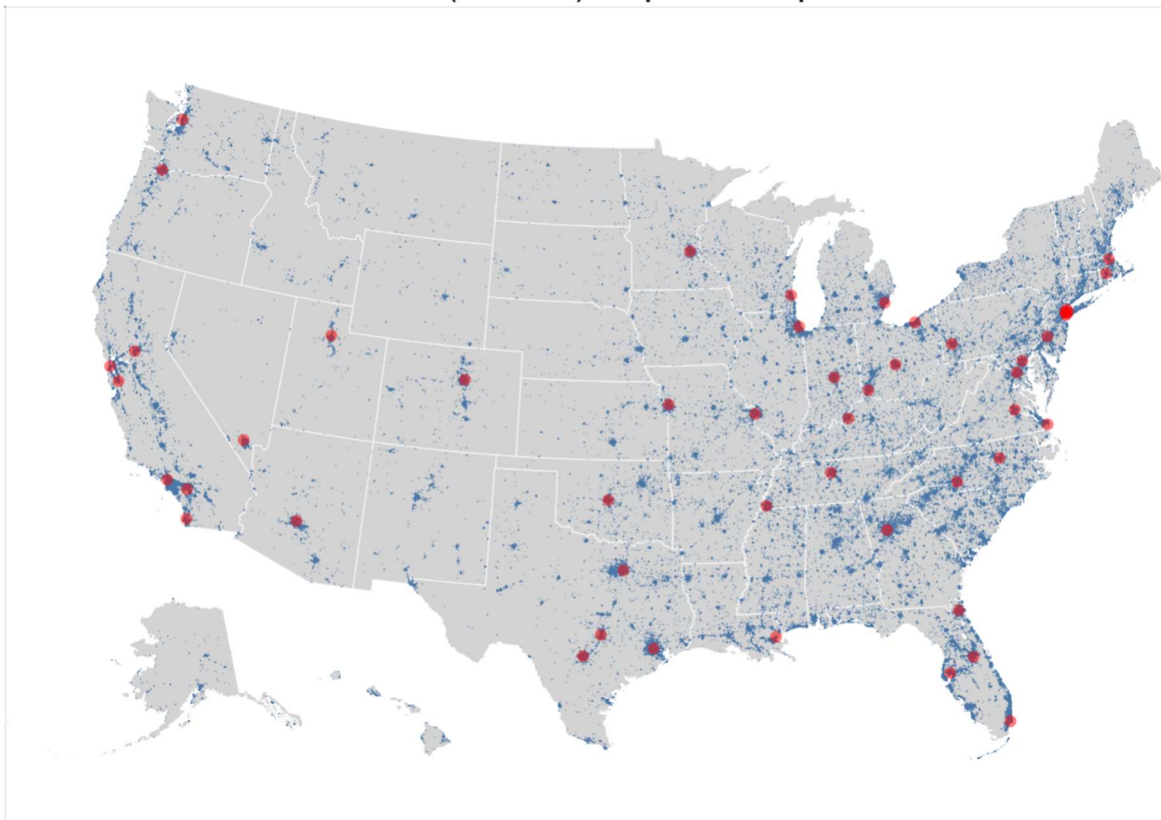
   Our Answer: No correlation between gun ownership and gun violence. United States had the highest gun ownership rate, but a lower mortality rate relative to other countries.

In the United States of America, gun violence is a severe public health issue, being one of the leading causes of premature death. In the last two decades, hundreds of thousands of American lives have been taken as a result of gun violence; while the data varies, many estimates place the number of American fatalities on par with the entire number of American military casualties since World War I started. No other country has even half the amount of gun deaths and comes close to the number of firearms circulating and moving within American borders. Regardless of opinions on the Second Amendment, the United States is a significant outlier in the international world on this topic, and that alone should give every American a pause.

As gun violence happens daily, we therefore want to analyze datasets to identify patterns in the timing and location of these occurrences, in addition to determining whether there is a correlation between gun ownership and gun violence. Preventing gun violence fatalities in our country is one of our greatest concerns, and we wish to utilize this report to emphasize the urgency towards solving this issue in America.

Gun Violence Incidents (2013-2018) & Top 50 Most Populated US Cities

Correlation Between Gun Ownership and Gun Violence

- URL: https://www.openintro.org/book/statdata/index.php?data=gun_violence
- This dataset covers information on gun ownership and gun-related deaths in 75 countries.

Extra Data from Gun Violence Archive

- URL: https://www.gunviolencearchive.org/reports
- Extra data like mass shootings and children's casualties.

Geospatial Data of the World

- URL: https://hub.arcgis.com/datasets/CESJ::world-countries/explore?location=44.064379%2C44.250531%2C2.39
- A comprehensive overlay of nation borders that works best at huge dimensions.

Gun Violence Archive

- URL: https://www.gunviolencearchive.org/reports/number-of-gun-deaths
- This dataset computes the number of gun violence fatalities in 2022.

Gun Violence Data

- URL: https://github.com/jamesqo/gun-violence-data
- These people compiled the Gun Violence Archive's website data on gun violence in the United States from January 1, 2013, to March 31, 2018, into a CSV.

U.S. Geospatial Data

- URL: https://hub.arcgis.com/datasets/CMHS::states-shapefile/explore?location=32.737707%2C57.339835%2C4.00&showTable=true
- The United States' geospatial data to assist plotting.

U.S. Populations

- URL: https://simplemaps.com/data/us-cities
- A simple, accurate, and up-to-date database of cities and towns in the United States.

For each research question, we must start by reading the CSV data along with any necessary geospatial data. For research questions 1 and 2, this will be CSV files regarding gun violence in the United States. For research question 3, this will be a CSV and SHP file regarding gun violence around the world. We then process the data by sorting the Data Frames to only have columns that are useful towards answering our research questions.

Research Question 1:

This question requires the usage of the following columns of the gun violence in the United States dataset: "state", "city_or_county", "n_killed', "n_injured", "latitude", "longitude", and "year". From the United States Population dataset: "city", "state_name", "state_id", "population", "lat", and "lng." And lastly from the United States Geospatial Dataset: "geometry." We used all of these to create advanced interactive visualizations to see which areas have more gun violence than others.

Research Question 2

This question required the usage of the following columns of the United States gun violence dataset: "year", "n_killed", and "n_injured". We did not need to use as many columns as we just needed to compare the difference between the number of people killed and number of people injured from the years 2013-2018 and note any of the changes. We also kept in mind that the year 2013 did not have as many recorded incidents as the later years. Additionally, the year 2018 stops at the end of March.

Research Question 3

This question requires the usage of the following columns of the worldwide dataset: "COUNTRY", "country", "mortality_rate", "ownership_rate", and of course, "geometry." We will use these rows to construct a visualization of mortality rates and gun ownership rates across the world. In addition to this visualization, we will create a bar chart that helps demonstrate the mortality rates versus gun ownership rates of countries with mortality rates greater than 5. We will utilize the plot and chart that we programmed to compare how the United States is contributing to preventing gun violence in contrast with other parts of the world.

During coding, we will use labels, colors, and other useful libraries that will help us to display results from the datasets. After coding, we will analyze and interpret the datasets with the programs that we formulated to write a report summarizing our findings about gun violence in the United States and across the world.
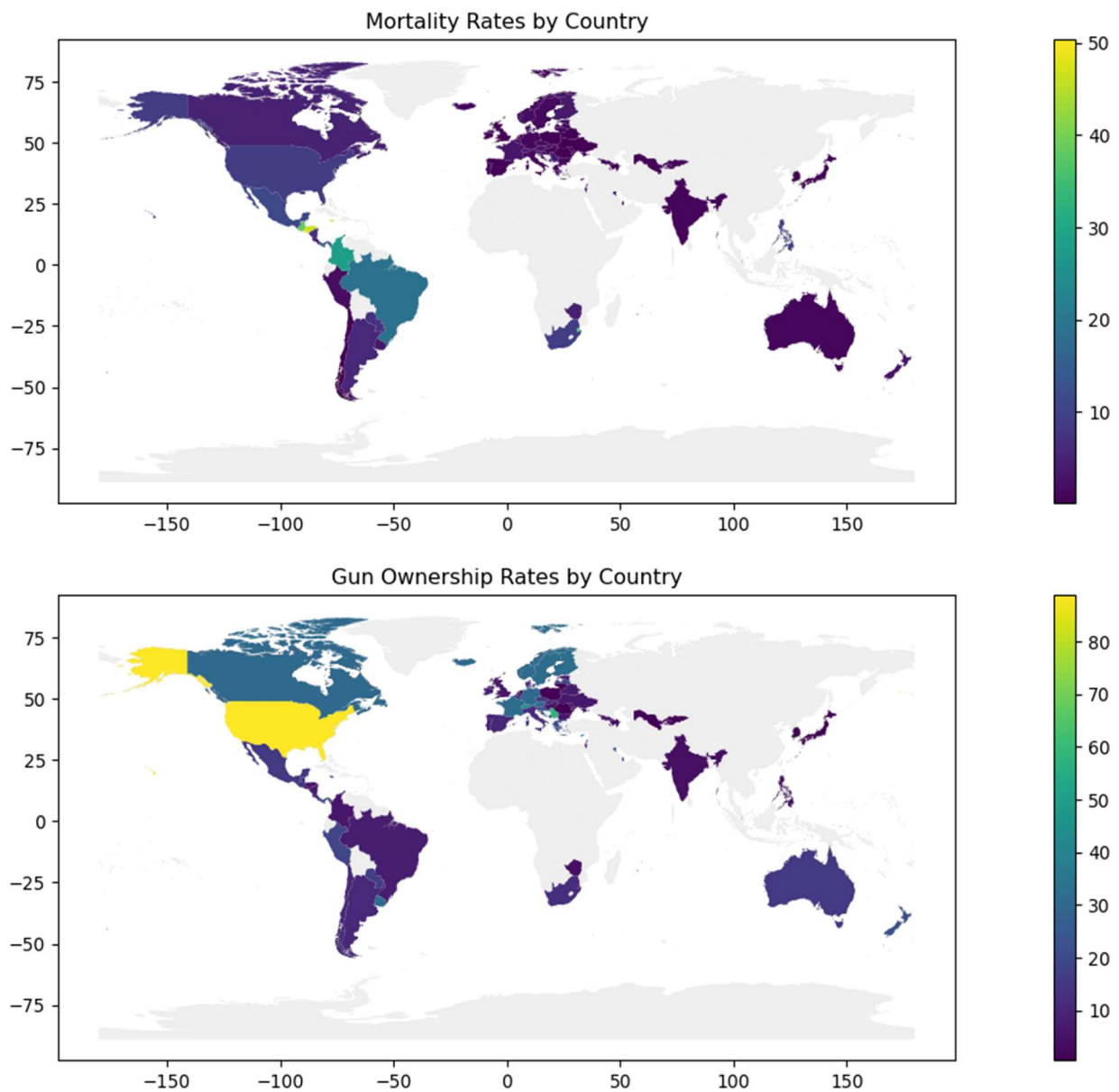

Mortality Rates by Country


Gun Ownership Rates by Country

Are there particular states of the United States that are more susceptible to gun violence? If so, why could this be?

As the map plot indicates, the Midwest and East coast of America have a shockingly greater prevalence of gun violence. Meanwhile, other reports have come from the West Coast, including Washington, Oregon, and California. Although the reasons remain unclear, the East Coast is home to the highest rate of gun violence. More specifically, Florida and all thirteen original colonies hold a tremendous amount of gun violence occurrences throughout the five-year period from 2013 to 2018. By contrast, several states in the West and Midwest, such as Montana, South Dakota, and North Dakota, have a relatively low number of instances. In addition to these bigger states, we have specifically Chicago, Illinois as a city that stood out. This is not surprising as the city has a multitude of gang-related conflicts.
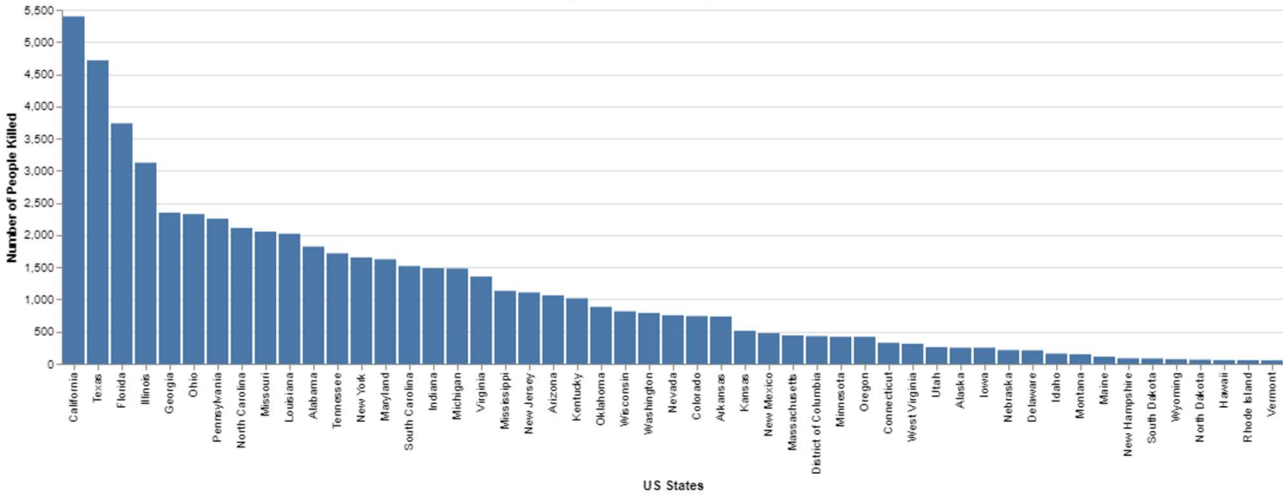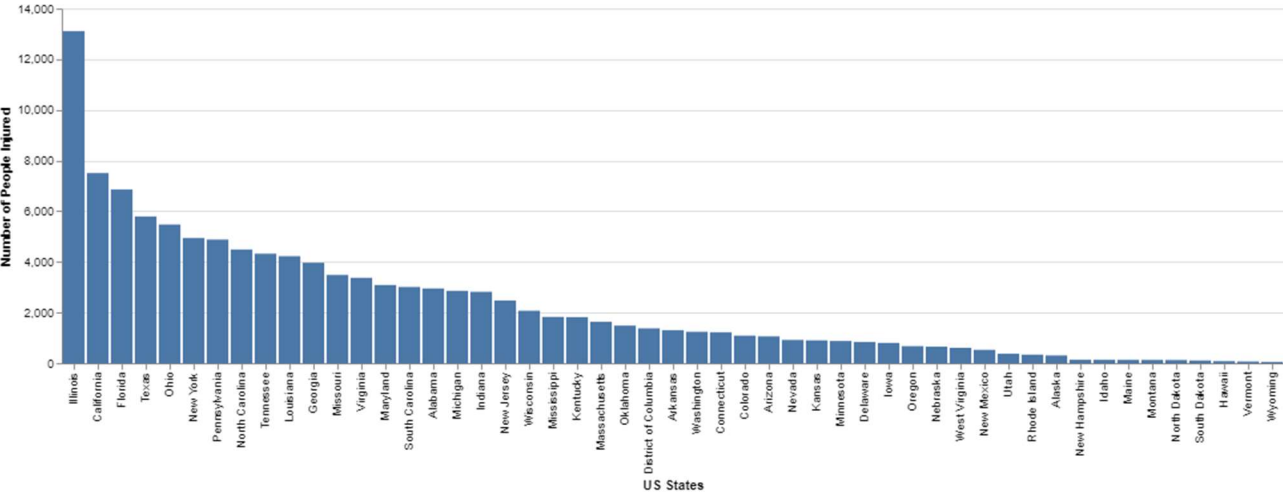

Gun Violence Incidents (2013-2018)
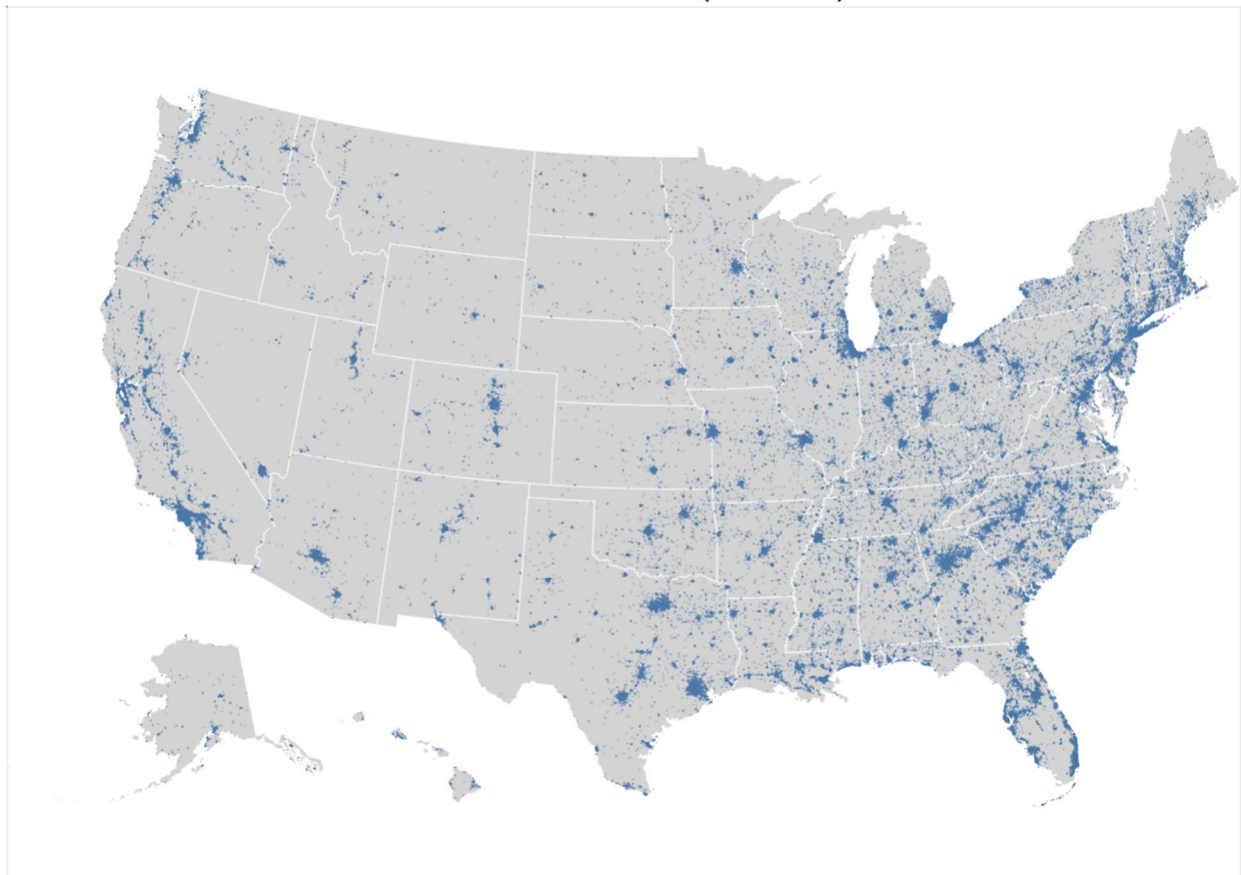
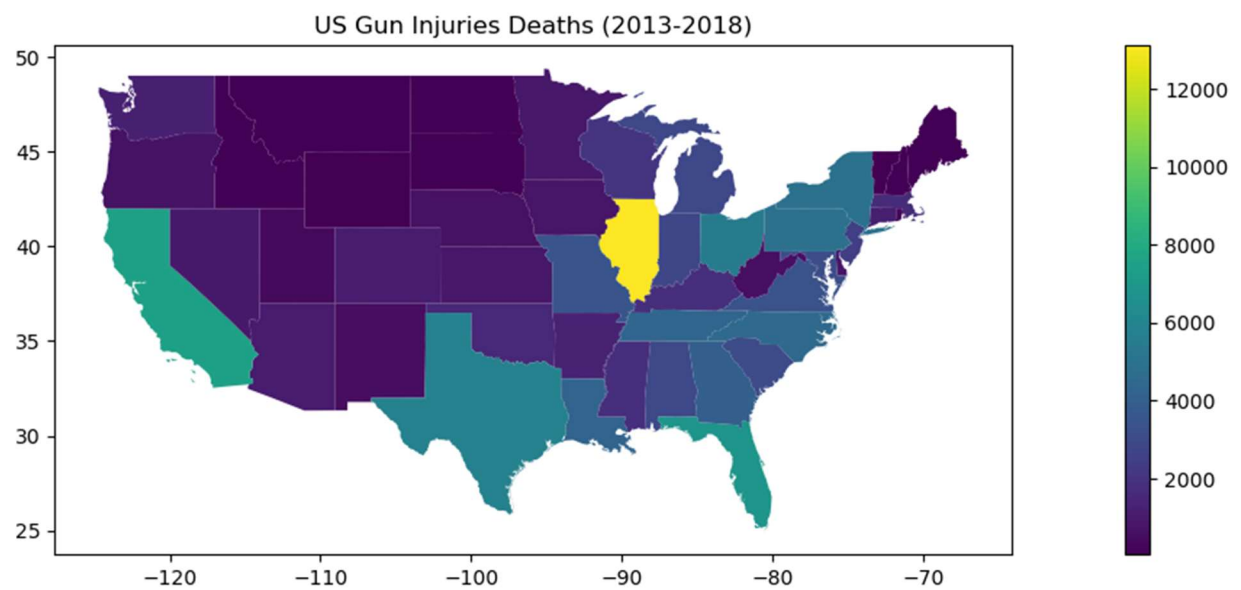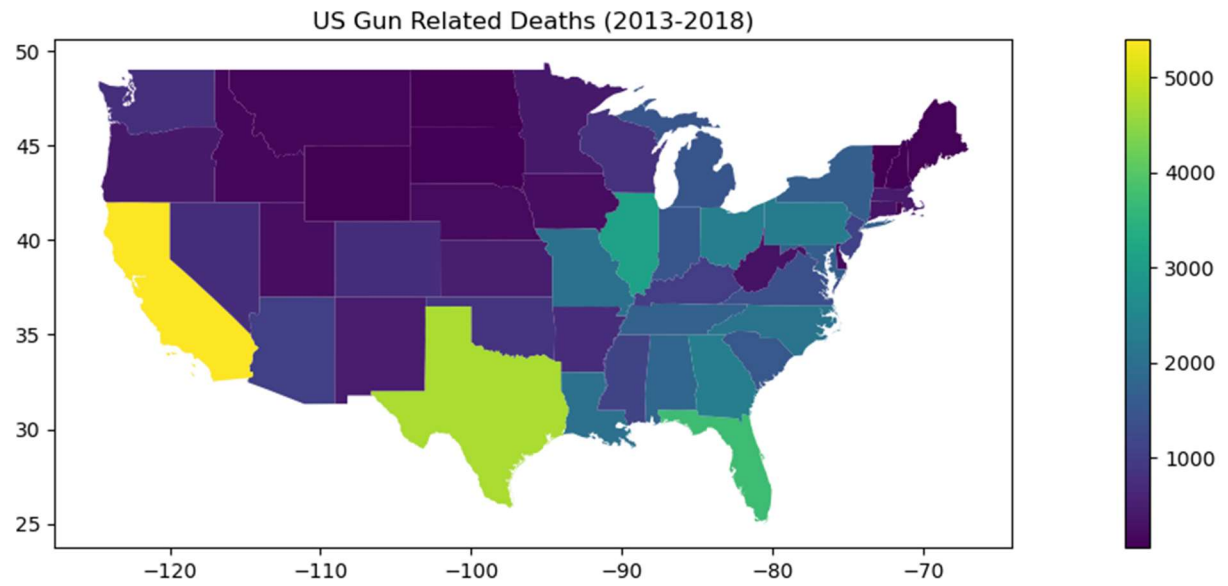## Total Gun Related Incidents per State
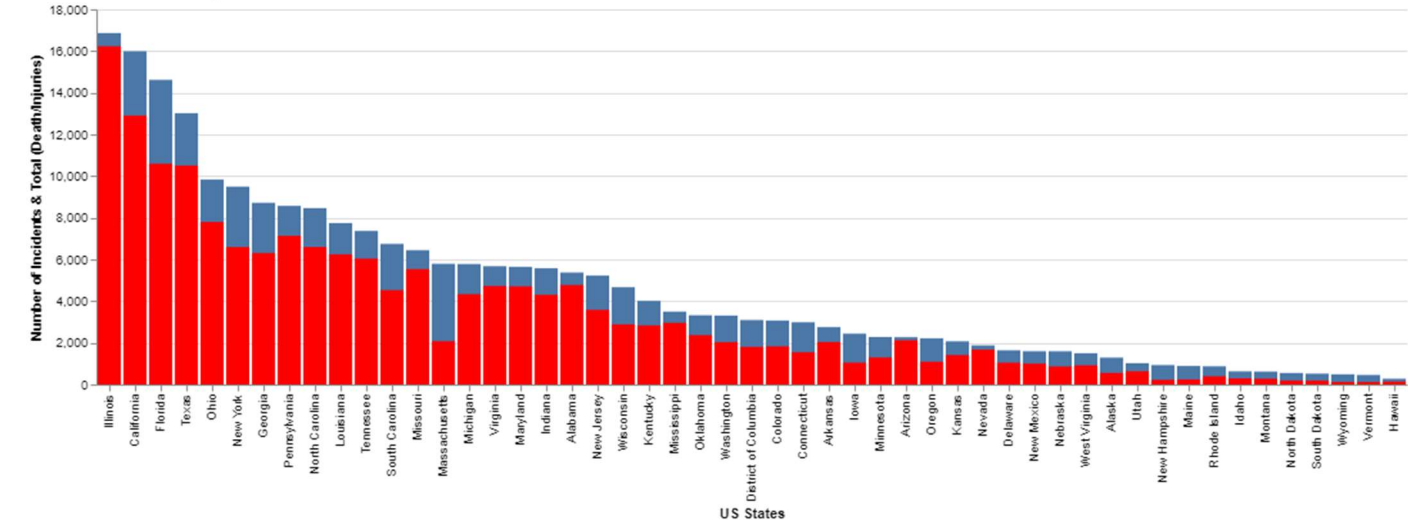
People Killed per State

People Injured per State

US Gun Related Deaths (2013-2018)


US Gun Injuries Deaths (2013-2018)

How did the number of *injuries* and *deaths* by gun violence change as time progressed from 2013 to 2018? What societal events could have influenced the increase, decrease, or stability of injury/death rates?

Ignoring the outliers of our data (2013 and 2018) we can see that there is a positive trend and there is a slow increase in gun violence as the years progress. There were many notable mass shootings in each of those years. For 2015: the San Bernadino Attack, 2016: The Orlando nightclub shooting, 2017: The Las Vegas shooting, and 2018: the Stoneman Douglas high school shooting (Gun Violence Archive).
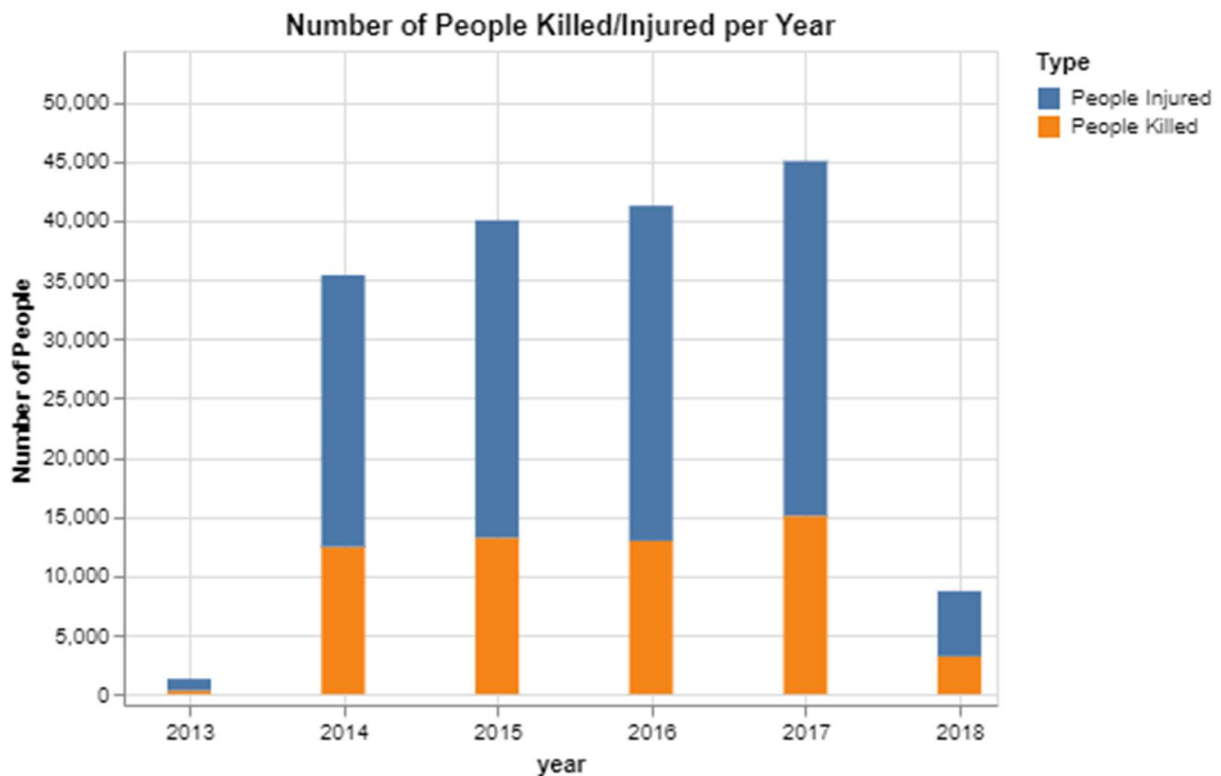
What is the worldwide relationship between gun ownership and gun violence? Do countries with a higher gun ownership rate also have a higher rate of gun-related deaths?

Using a worldwide dataset on mortality rates and gun ownership rates, we surprisingly found that there was no correlation between gun ownership and gun violence. We have also found that the United States had an extreme gun ownership rate in comparison to any other state but had a relatively lower mortality rate than the other countries on our visualization. If there was a relationship between gun ownership and gun violence, the mortality rate and gun ownership rates would appear more similar, however they were not.
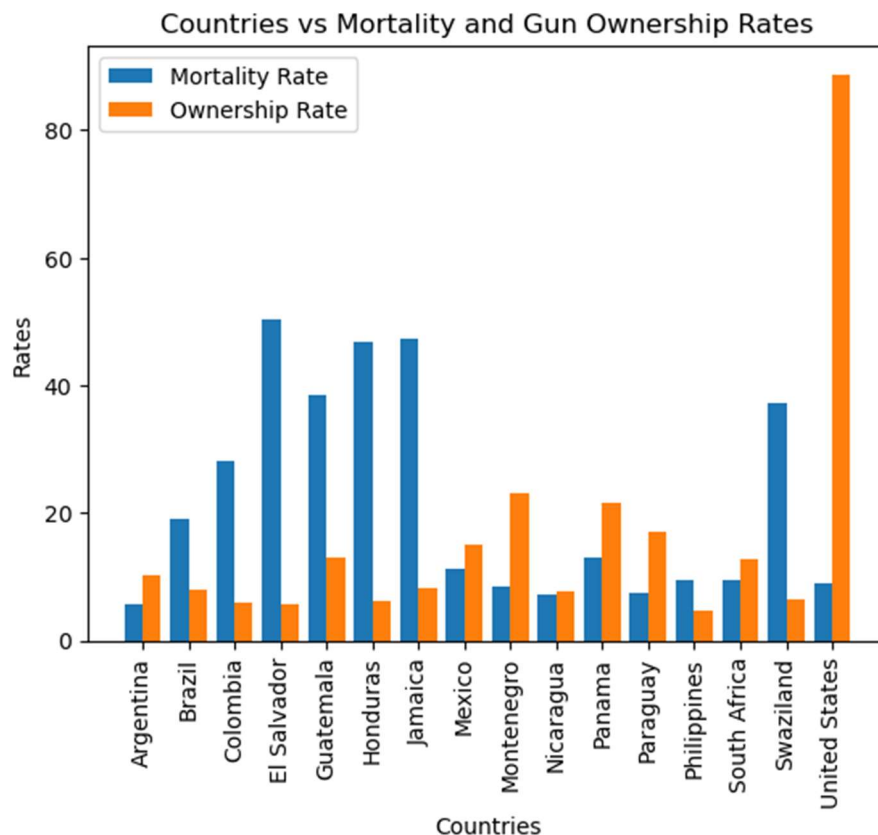
We predict that the data may have resulted in this way due to the United States being a country where it is easy to purchase firearms, but also one of the larger countries with over 300 million people.

Impact and Limitations

Several drawbacks were the inability to locate datasets with sufficient information for analysis. For instance, the statistics on gun violence from 2013 to 2018 are lacking several months from both 2013 and 2018. Without providing acceptable justifications, the individuals that compiled the dataset merely indicated that they did not collect data in 2013 and ceased recording data in March 2018. The outcomes would have been different had they maintained consistency in data collection. Consequently, our conclusions and findings are not completely precise.

Additionally, in plotting maps to evaluate the global association between gun violence and gun ownership, we discovered that the dataset we used omitted Russia and China. Given Russia's geographical mass and China's population, excluding their data will have a major influence on the overall association between gun violence and gun ownership.

To run a report effectively, we need a dataset that has sufficient information. However, we may not be able to get all the information we need due to a variety of variables, such as official cooperation. Therefore, conducting and revising a decent report would take years.

Challenge Goals

New Library

We constructed an advanced visualization of gun violence data across the United States using Altair, a Python Interactive Data Visualization library. This allows us to show the numerical data for each state without crowding the graph with numbers. In addition, it enables us to hover over our gun violence incidents map and immediately identify the names of cities that seem to have more cases.

Multiple Datasets/Result Validity

We utilized a dataset of worldwide gun violence to validate our findings and results about gun violence in the United States. Comparing our nation to others enables us to assess how well our community contributes to global gun violence prevention.

Work Plan Evaluation

1. Prepping (1 hour):
   - Looking for datasets.
   - Diving up tasks.
     - Joseph working on the first research question.
     - Nghi working on the second research question.
     - Minh working on third research question.
2. Coding, interpreting, and testing (60 hours):
   - Coding data and plotting with Geopandas (~10 hours).
   - Learning Altair and creating it (~15 hours).
   - Interpret data and come to conclusions (<5 hours).
   - Testing everything to ensure everything is correct and efficient (~10 hours).
3. Writing and polishing the report (10 hours).
   - Everyone working on the report together while communicating through Discord.

Due to our coordinated effort and communication, everything went according to plan. Everyone was accountable for their own work component and aided one another whenever feasible. We could not have completed this report without the assistance of our team members.

Testing

The way that we tested our code was through ensuring that our programmed visualizations matched the data in the datasets that we used. Many times, throughout and after our coding process, we made sure that the numbers on our plots and charts were lining up accurately.

Collaboration

The report was conducted by Nghi Huynh, Minh Mai, and Joseph Tran. We did not collaborate with any other individuals aside from the course staff of CSE 163.

Interactive visualization, Altair documentation: https://altair-viz.github.io/