**Gun Deaths and Demographics**

Final Project

GEOG 1050 Foundation of GIS

Federico Antolini

Leon Deng

Spring 2020

**Introduction**

In 2012, the Sandy Hook shooting in Newton, Connecticut shocked me as a seventh grader. After five years, another shooting broke out in Las Vegas during a music festival. This incident was followed by the Stoneman Douglas shooting six months later and the El Paso shooting a year later. Seeing the news cover shootings over the last few years, I began to wonder whether the issue with gun violence is common in the western world, if gun violence is getting worse in the United States, and why shootings continue to happen while strong voices call for banning guns. This project provided me with a great opportunity to understand the recent national trend and the current status of gun death rates at the state level and identify the driving forces behind firearm deaths through spatial analysis. My objective of this project is to answer the following questions:

* How does gun violence in the US compare with that of peer nations?
* What is the current trend of gun violence in the US? (Is it getting worse?)
* What does the current pattern of gun death look like across states? Is it consistent, or does it vary widely?
* What are the major driving forces behind gun death rates?

I hope that the fact-based and insightful results of my project will raise the attention of policymakers, law enforcement, the community (including arms manufacturers and dealers, educators, mental health institutions), and individuals.

**Data**

To meet the project objective and conduct spatial analysis, I identified and gathered data on the trend and current status of gun violence from credible data sources such as World Population Review, Gifford’s Law Center, and Statista. I also retrieved a US states shapefile from Data.gov, the home of the US government’s open data – which I saved in my Project folder as a zipped file. The geographic/spatial data I collected for this project are listed below:

Background Data Source:

* Gun Violence Statistics (as of 2018):

<https://lawcenter.giffords.org/facts/gun-violence-statistics/>

Trend Data Source:

* Number of Firearm Deaths in the US from 2007 – 2017 (as of August 9, 2019): <https://www.statista.com/statistics/258913/number-of-firearm-deaths-in-the-united-states/>

Spatial Data Sources:

The coordinate system for all the following data sources is GCS\_North\_American\_1983. The datum for all the data sources is D\_North\_American\_1983.

* Firearm death rate per 100K population by state (as of February 2020): <http://worldpopulationreview.com/states/gun-deaths-by-state/>
* Gun ownership by state (as of February 2020):

<http://worldpopulationreview.com/states/gun-ownership-by-state/>

* Median household income by state (as of February 2020): <http://worldpopulationreview.com/states/median-household-income-by-state/>
* Unemployment rate by state (as of February 2020): <http://worldpopulationreview.com/states/unemployment-rate-by-state/>
* Education level (high school graduation rate) by state (as of February 2020): <http://worldpopulationreview.com/states/high-school-graduation-rates-by-state/>
* Median age by state (as of February 2020):
* <http://worldpopulationreview.com/states/median-age-by-state/>
* Race structure by state (as of February 2020): <http://worldpopulationreview.com/states/states-by-race/>
* Shapefile source:

<https://catalog.data.gov/dataset/tiger-line-shapefile-2017-nation-u-s-current-state-and-equivalent-national>

**Analysis:**

To generate fact-based and insightful project results, I used a comprehensive methodology over the course of this project, including four major steps: data collection, data management and analysis, chart and map development, and spatial analysis.

Data collection

Based on my project objective, I researched, identified, and gathered data from credible sources. After studying and identifying the relevance and credibility of the data, I downloaded the following raw data in the form of CSVs and uploaded them into Excel.

* Gun death rate data of US and peer countries
* US gun death trend data from 2007 to 2017
* Latest state-level data of gun death rates, gun ownership, median household income, unemployment rates, education level, median age, and race structure

Data management and analysis

In Excel, I turned the raw data into structured data using the steps below:

* Created an Excel table for the gun death rate per 100K population of US and major peer countries.
* Created an Excel table for US gun death trend from 2007 to 2017 and calculated the year-by-year change percentages and the change percentage rate of 2017 vs 2007.
* Created an Excel table for current status at the state level by using state name as primary key and the VLOOKUP formula to join the data (gun death rate, ownership, income, unemployment, education, age, and race).
* Used the maximum, minimum, median, average, and standard deviation functions to generate high-level statistical metrics across states.
* Used the CORRELATION function to quantify correlation coefficient between rate of gun deaths and the other demographic factors.

Chart and Map Development

* Chart development in Excel
  + Using the structured data on the gun death rate of US and peer countries table in Excel, I created a bar graph to present a visual view of peer comparison.
  + Using the structured data on the US gun death trend table in Excel, I created a line graph to illustrate the trends over time.
  + Using the structured data on the current status table in Excel, I developed XY scatter charts to display the relationship between gun death rates and each demographic variable and quantified their correlation.
* Map development in ArcMap

Once the trend chart and XY scatter charts were created in Excel, I began developing my bivariate maps in ArcMap for gun death rates and each demographic factor using the following steps:

* Unzipped the US states shapefile to my Bivariate Map folder.
* Opened ArcMap and added the shapefile to the table of contents, changed the layer’s coordinate system to North American Lambert Conical Projection.
* Zoomed in to mainland USA.
* Conducted a table join by selecting the state name column in the shapefile as the field in the layer to join, adding the respective CSV files for gun death rate and gun ownership, and choosing the state name column in the CSV to base the join on.
* Opened the shapefile’s properties and in the symbology tab, selected the quantities section and specified the value as the rate of gun deaths.
* Chose a color ramp using the website Color Brewer for the map with gun death rate only and the color map provided in Lab 4 for gun death rate vs gun ownership.
* Added two more data frames, copied and pasted the shapefile configured with Lambert Conformal Conic Projection, gun death rate CSV, and gun ownership CSV into the two data frames and zoomed in on Alaska and Hawaii respectively. These two frames are added in map layout view so I have all US states.
* Added all required map components (title, legend, scale bar, north arrow, author) in layout view.
* Repeated the steps above for gun death rate and all other demographic factors (race, age, education, unemployment) for different maps.
* Used the Excel XY scatter graphs and ArcMap bivariate maps to analyze the correlation between the gun death rate and gun ownership and other demographic factors.

(The lessons I learned through this project are discussed in the conclusion section.)

Spatial Analysis

After I created the bivariate maps, I conducted spatial autocorrelation analysis using the Global Moran’s I tool in Geoprocessing tools to determine if the spatial distribution of data on gun violence by state is random or not by following the steps listed below:

* Using a query definition to exclude D.C, Alaska, and Hawaii in the map before running Global Moran’s I for spatial autocorrelation, I exported the selected features as a new shapefile. I did not include D.C because there is no value for rate of gun death in D.C. Alaska and Hawaii are also excluded because they are not connected to the continental states, making them outliers for the analysis. Moran’s I does not run with them due to threshold distance issue.
* Using the Project tool, I set the new shapefile’s coordinate system to North American Lambert Conformal Conic Projection, a good choice for an east-west extending mid-latitude country/continent like USA and Canada.
* Opening Global Moran’s I tool, I selected the projected shapefile as the input feature class, the gun death rate as input field, and set the threshold distance as 500,000 meters.
* Specified the threshold distance as 500,000 meters because without it, the tool throws a warning 000853, which states that the minimum distance to ensure every feature has at least one neighbor will be used. I also set the standardization as Row because states are polygons, and the Row option reduces bias for polygon features.
* Selected “Display results”, then ran the test and documented the Moran’s Index, expected value, variance, z-score, and p-value.
* Saved screenshot of results page with bell curve and significant values to the project folder.

**Results**

The peer comparison bar chart shows that gun death rate in the US is much higher than that of peer countries because among these countries, only the US allows individuals to bear arms. The linear 10-year trending graph demonstrates that US gun death rate of 2017 has significantly increased from 2007 by 27.4%, especially since 2015. These findings prove that the US gun violence is in rise and the current gun violence status is horrific, which raises questions on what caused the increase in gun death rates, and whether or not the national gun violence status is consistent across states. The Excel XY scatter graphs and bivariate maps are created to evaluate the gun violence at the state level and analyze multiple contributions of different factors.



Data source: Statisat 2019

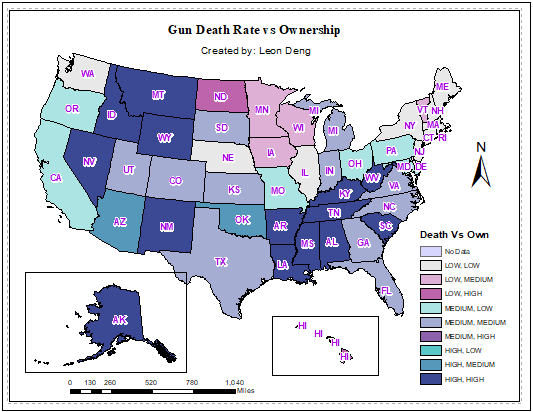
Data source: Giffords Law Center 2018

The scatterplots in Excel and the bivariate maps using spatial attribute and querying in ArcMap visually represent the relationship between gun violence and other demographic factors all at the state level (gun ownership, income, unemployment, education, median age, and race). From the bivariate maps and scatter graphs, I generated the following insightful findings.

The first major finding is that across all states and at the individual state level, there is strong correlation between rate of firearm deaths and percent of gun ownership. In the bottom left, the XY scatterplot shows that the gun death rate increases with gun ownership with a correlation coefficient of 0.7. The bivariate map at the bottom right shows that the consistent correlation between gun death rate vs gun ownership exists at the individual state level.

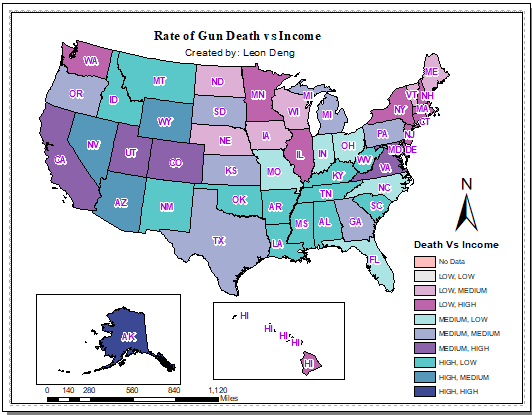
For example, states in the dark violet color, including Alaska, Arkansas, Louisiana, Mississippi, Alabama, and Tennessee have high rates of gun ownership and high rates of gun death. As another example, states that are light purple such as Utah, Colorado, Texas, and Florida have medium rates of gun death rate and ownership. Finally, states that are a whitish pink color like Nebraska, New York, and Illinois have low rates of gun deaths and ownership. The map shows no states that are blue green, which have high rates of gun death and low rates of gun ownership.

This finding indicates that the immediate solution to reduce firearm fatalities is to reduce easy access to guns through stricter gun control laws. However, the overall correlation of gun ownership and gun death rate is not a perfect one. A small number of states such as California, Missouri, Oregon, and Pennsylvania have medium rates of gun death but low rates of gun ownership, and Hawaii has a high rate of gun ownership yet the lowest gun death rate. This indicates that there are other major factors that play a role in gun violence.

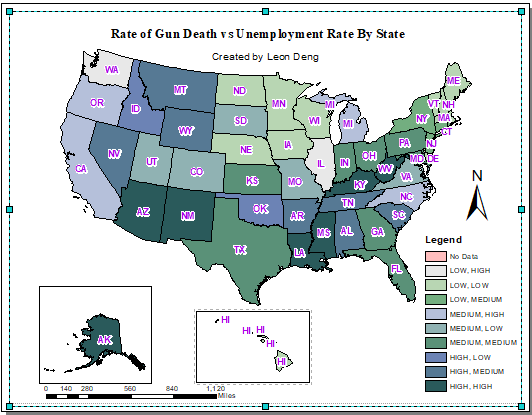
 

In order to find the other contributing factors to gun violence, I analyzed the correlation between gun death rates and other spatial demographic variables such as income, unemployment, education level, median age, and race.

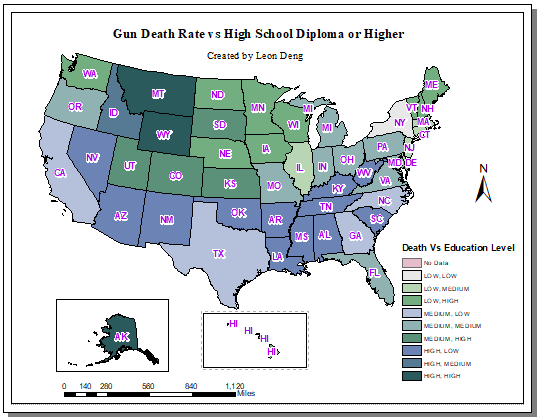
In the XY scatterplot for rate of gun death and average household income, the correlation coefficient is -0.6, and gun death rates decrease as income increases. This indicates that the rate of gun death and average household income has a moderate negative correlation. States with lower income levels generally have higher rates of gun deaths. The negative correlation can be seen in the bivariate map for gun death rate and average household income. States that are light blue, such as New Mexico, Idaho, Arkansas, or Tennessee, have high rates of gun deaths and low income. The states that are lavender, such as Texas, Georgia, Oregon, and Kansas, have medium gun death rates and medium household income. As for states that are light purple like Washington, Minnesota, Illinois, and New York, they have low rates of gun death and high levels of household income. Based on this finding, improving community safety through increasing income and reducing poverty would help lower the gun death rate.

One way to increase average household income is to create more employment opportunities. I analyzed the correlation between unemployment and rate of gun death, which showed an overall positive moderate correlation of 0.5, meaning the gun death rate tends to increase as unemployment rate increases. From the bivariate maps, states like Louisiana, Mississippi, Kentucky, New Mexico, and West Virginia have high unemployment levels and high rates of gun death. For other states such as Minnesota, Wisconsin, Maryland, and Connecticut, they have low unemployment and low rates of gun death. It is clear that increasing employment is an important enabler in reducing gun violence.

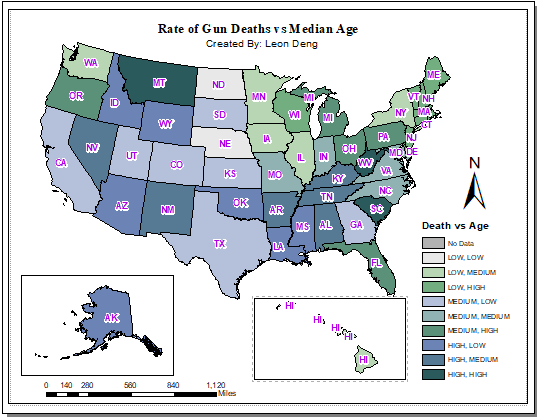
 

Besides analyzing the relationship of unemployment rate and gun death rate, I quantified the correlation of education level and gun death rate which is -0.4. From the scatterplot for rate of gun death and education level, higher levels of education are associated with lower rates of gun death. The map also shows the moderate negative correlation at the state level. States like Kanas, Colorado, and Utah have high levels of education and low levels of gun deaths, while states like Illinois and Nevada, Arizona, Louisiana, and Alabama have low education levels and high rates of gun deaths. Thus, increasing education is key to reducing gun death rates.

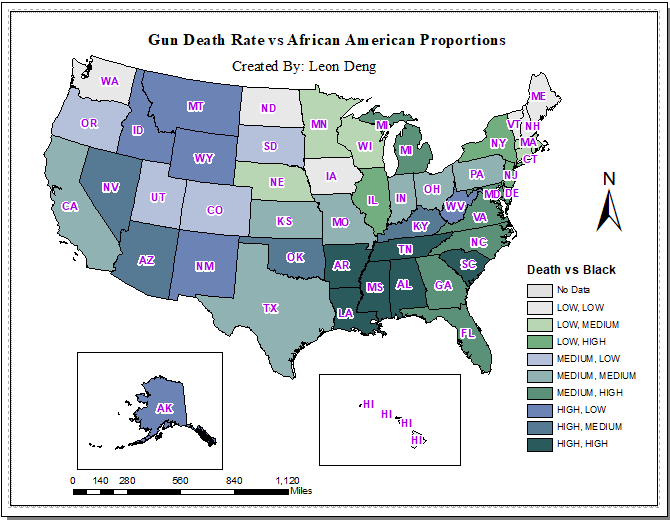
When it comes to education, it is important to focus on youth populations. After analyzing median age and gun death rates, I found that they have a correlation of -0.3, which is weak but still negatively associated. This shows that younger people are more prone to gun violence. In the bivariate map, states with lower median ages such as Idaho, Wyoming, and Louisiana also have high rates of gun death. It is crucial for these states to focus on educating young people on gun safety and creating employment opportunities for them.

There are some exceptions to the negative correlation shown in the scatter plot and bivariate map. Montana, South Carolina, and West Virginia have a midrange median age and high rates of gun deaths as indicated by their dark green color. As for Nebraska and North Dakota, their color is light gray, which indicates low rates of gun death and low median ages. These outliers reiterate that there is not a single force driving gun violence. To really solve the gun violence issue, we need to take a comprehensive approach.

To have a comprehensive analysis, I also analyzed the correlation of gun death rates and races.

From the scatterplot below for rate of gun death and African American proportion, gun death rate increases as the proportion of African American increases with a correlation of 0.3. In the map showing gun death rate vs proportion of African American, the states with the largest proportions of African Americans (Arkansas, Louisiana, Tennessee, Mississippi, Alabama, and South Carolina) also have high rates of gun deaths. In states with a smaller proportion of African Americans, gun death rates appear to be lower.

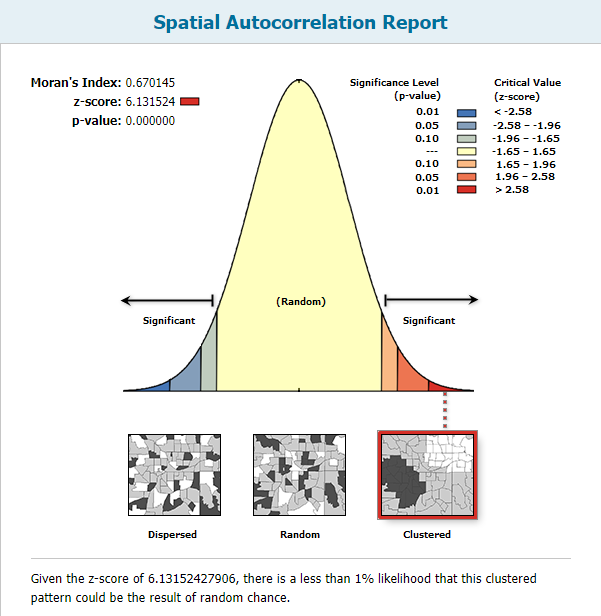
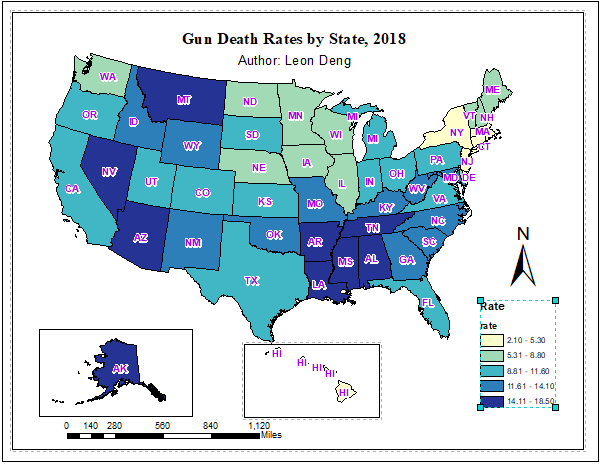


I did the same analysis across races and found that among the races, African American populations and Native American populations are the most at risk, as their correlation with gun death rates of 0.3 are the highest. I also quantified the correlation between each race and rate of gun death, along with other spatial variables (as shown in the table below). Although the proportion of African Americans has a weak negative association with gun ownership, their high gun death rate is associated with lower household income, lower education level, and higher unemployment rate. This finding demonstrates that to reduce gun violence in African American society, the focus should be on improving education and employment. As for Native Americans, the gun violence control efforts will be to reduce easy access to guns and improve education and employment opportunity with young people.



Spatial Autocorrelation Trial:

After running a Global Moran I’s Trial on the map with gun death rates by state, the following results were returned. Spatial autocorrelation always begins with the null hypothesis that the data is randomly distributed. As the p-value for this trial is 0, and the z-score is 6.131524 (which is a high positive value), I rejected the null hypothesis and concluded that gun violence by state is clustered. That is, states with similar rates of gun violence tend to be close to each other spatially. The results of this spatial correlation test indicate that the states in each region should take a similar approach to resolving their gun rate issues.



**Conclusion**

The results of my GIS project showed that gun violence in the US is higher than that of peer countries, and research indicates this is most likely due to the US being the only nation with the right to bear arms among these wealthy countries. Currently, the amount of gun violence in the US has seen an increase over the past few years potentially due to a lack of effective measures to control the issue of gun violence. Also, the rate of gun death does have a pattern by state, with states having similar levels of firearm deaths generally being close together. Factors correlated with gun death, besides gun ownership, include low household income, unemployment, younger populations, low education levels, and areas with a high number of African Americans.

The immediate ideal outcome I hope to see is that stricter policies are imposed to reduce easy access to firearms, as gun ownership is strongly correlated with gun death. State governments should pass stricter laws that regulate firearm use, and business codes must be implemented for arms sellers and weapon industries, such as reviewing a consumer’s background before issuing them a license to own firearms. At the same time, stricter laws alone cannot significantly reduce gun violence. Thus, a comprehensive community effort is needed to address the root of the issue and prevent access to firearms. For example, to ensure that people are employed and away from gun violence, more education programs are needed to teach people the dangers of firearms and to follow the law, especially young people and African Americans. Finally, to effectively implement these comprehensive community strategies, deeper research and analysis is needed at the state level to determine the contributing factors behind each state, since the amount of gun violence is different for each state. Every category of gun violence - suicides, homicides, or law enforcement operations – also needs to be researched to determine the specific causes of each.

If I were given more time on the project, I would have gathered more specific data about the different types of gun violence and analyzed their root causes. For example, to reduce instances of suicide, more mental health support and stress reduction is needed. To control levels of homicide, there needs to be more crime watch in communities. Also, I would have looked the unique causes of gun death in each state because the amount of gun violence is different by state, meaning that every state needs its own policies when addressing gun death. I would have also liked to research the exact causes that led to a spike in the gun death rate starting in 2015, as I focused more on variations across states than across time.

Some parts of the project that went well for me was that I was able to develop proposal and identify relevant data, follow an organized project plan from initial proposal to revision to develop and refine project methods (revision – problem with shapefile), as well as leveraging functions of Excel and Arc Map to determine correlation and trends and spatial relations by state, respectively.

This project was a real eye opener for me, and in doing it I have learned how to better tackle future GIS projects. I ran into a few problems but resolved them and improved my GIS skills as a result. Below are some examples of the lessons I learned through this project:

One problem I had while developing maps was that DC’s data was null because it was unavailable, and that DC is not technically a state. To resolve this issue, I created a category called “No Data”, indicating that DC was not part of the analysis.

Another problem arose while I was creating other bivariate maps for gun death rate vs age, race, income, unemployment, and education. Whenever I added the states shapefile to a new map, the it always contained the joined CSV for the previous map’s data. For example, when making the map for gun death rate vs median age, I found the tables from the gun ownership CSV in the shapefile’s attribute table. The previous map also had an error: when I re-opened the map for gun death rates and gun ownership, it did not have its shapefile. After investigating, I learned that a shapefile can only be joined with one CSV at a time, so I made different copies of the US states shapefile from the zipped folder, joining one CSV to one shapefile. As a result, each shapefile has its own joined CSV and no longer disappears from the map.

Finally, when it came to spatial analysis, I had to exclude Alaska and Hawaii from the layer to analyze because they are not spatially joined with continental states. Alaska is connected Canada, and Hawaii is an island chain.

Over the course of the project, I was able to apply the learnings throughout the semester to this project, such as analyzing the correlation with gun death and demographic factor using bivariate maps. At the same time, I got to explore the new tool of spatial autocorrelation to determine if data is clustered and visually display the correlation in graphs. Most of all, the project raised many interesting questions that were worthy of investigating to better understand the causes and trends of gun violence, and what we can do the prevent gun violence as a society.