Final project report

**Janhavi Anantprakash Kharmale**

**University of South Florida**

**ISM** **ISM6419.901S24 -Data Visualization Prof. Han Reichgelt**

**May. 02, 2024**

Table of Contents

1. [Introduction 3](#_bookmark0)
2. Methodology [4](#_bookmark1)
3. Analysis [5](#_bookmark2)
4. Conclusion [14](#_bookmark3)

# Introduction

In recent years, mass shootings have tragically become a recurring and concerning phenomenon in the United States, prompting widespread public concern and calls for effective intervention strategies. Understanding the complex interplay of factors contributing to mass shootings at the local level is crucial for developing targeted prevention efforts and policy responses.

This project focuses on investigating county and state level factors associated with mass shootings across the entire United States. By analyzing demographic, socioeconomic, and geographic data, we aim to uncover patterns and insights that shed light on the contextual dynamics of mass shootings.

This study aims to investigate critical research questions related to mass shootings in the United States at the state and county level, focusing on the intersection of demographic, socioeconomic, and geographic factors. The first question examines how the incidence rate of mass shootings varies across counties within each state and explores the relationship between this variation and population. Understanding these dynamics can provide insights into the spatial distribution of mass shootings and the potential impact of population density on such incidents.

The second research question delves into demographic factors such as race, immigration status and income levels to identify the frequency of mass shootings. By analyzing these demographic characteristics, we aim to uncover patterns and disparities that may contribute to the prevalence of mass shootings in specific communities.

Furthermore, this study seeks to explore whether socioeconomic indicators, including poverty rates, unemployment rates, and income per capita, can predict the likelihood of mass shootings occurring in certain counties. Identifying these socioeconomic predictors can inform targeted interventions aimed at addressing underlying social and economic factors associated with mass violence.

Lastly, the study investigates the prevalence and impact of mass shootings occurring within educational institutes, including schools, colleges, and universities. This analysis aims to understand the frequency and severity of mass violence incidents specifically targeting educational settings. By examining the data on mass shootings in educational institutes, we seek to identify patterns, trends, and underlying factors associated with these tragic events.

By addressing these questions, we aim to contribute valuable insights into the multifaceted nature of mass shootings and inform evidence-based strategies for prevention and intervention.

This study will leverage comprehensive datasets on mass shootings and demographic characteristics, enabling a nuanced analysis that explores the intersection of social, economic, and geographic factors in relation to mass violence incidents. Through this analysis, we seek to support informed decision-making and policy development aimed at reducing the occurrence and impact of mass shootings in our communities.

1. **Methodology**

**Data Sources:**

**Mass Shooting Dataset:**

The Mass Shooting Dataset is sourced from The Violence Project's Mass Shooter Database, which provides comprehensive information on recorded mass shootings in the United States. This dataset includes details on the location, date, number of victims, offender demographics, and contextual factors associated with each incident.

Dataset Source: The Violence Project - Mass Shooter Database

**Demographic Dataset:**

The Demographic Dataset contains population statistics by state and county, encompassing a wide range of demographic and socioeconomic indicators such as employment rates, gender distribution, racial demographics, poverty rates, and income levels.

Dataset Source: U.S. Census Bureau - American Community Survey (ACS)

**Data Integration and Analysis:**

The datasets will be integrated using common geographic identifiers (state and county) to enable county-level analysis across the United States.

# Analysis

**Description of Data Use and Visualization for Research Question 1:**

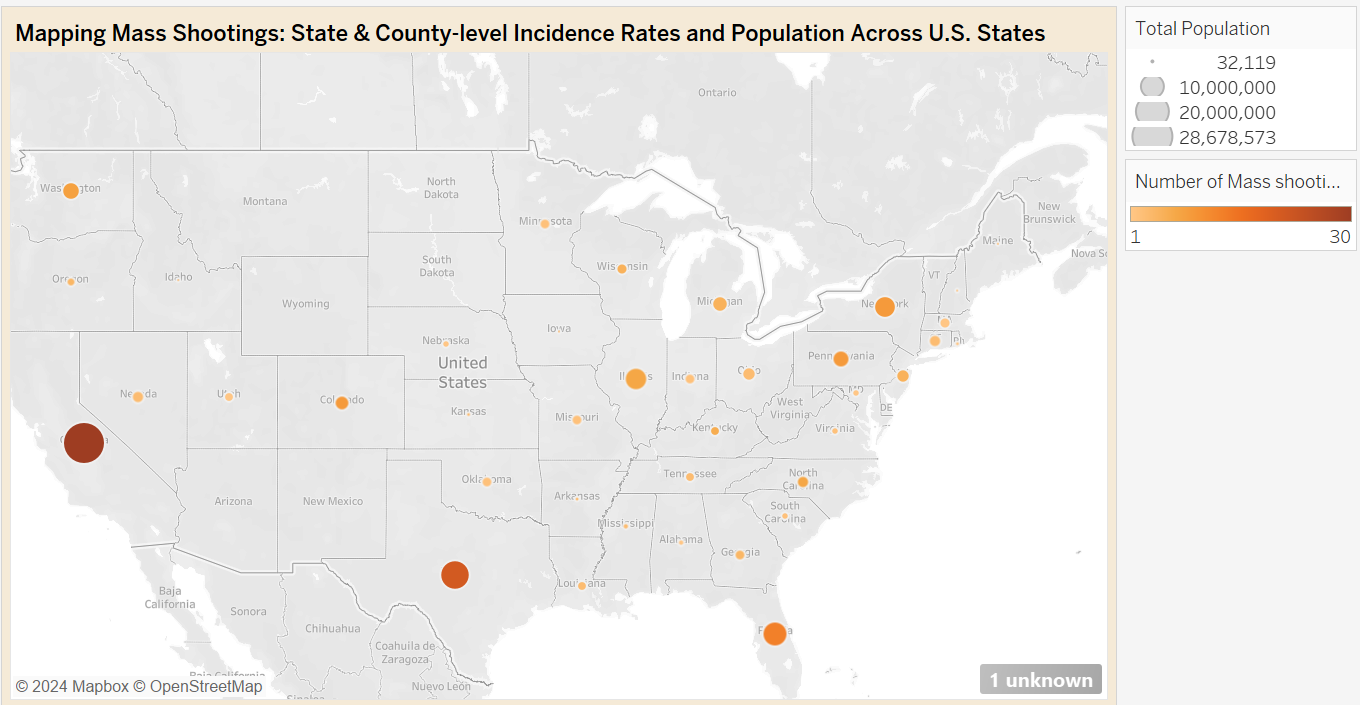
To address the research question regarding the variation in mass shooting incidence rate across counties within each state and its relationship to population density, I utilized a combination of the Mass Shooter Database and the Demographic Dataset, joined by state and county identifiers.

**Visualization 1:**

* I created a map visualization using Tableau to represent the number of mass shootings and total population for each state.
* The size of the circle on the map correlates with the population size of each state, providing a visual representation of population density.
* The color of the circle indicates the number of mass shooting incidents within each state, allowing for a comparison of mass shooting rates relative to population size.

**Key Insights and Interpretation:**

* The visualization reveals insights into states with higher concentrations of mass shootings relative to their population size, indicating potential hotspots or areas of concern.
* States like California, Texas, and Florida stood out with both large populations and higher numbers of mass shootings, suggesting a potential link between population density and mass shooting occurrence.
* This visual analysis provided a clear understanding that areas with denser populations tended to experience more mass shootings, aiding in the exploration of geographical patterns and contributing factors associated with this tragic phenomenon.

****

**Description of Data Use and Visualization for Research Question 2:**

To address the research question concerning demographic factors such as race, immigration status and income levels associated with the frequency of mass shootings, I employed data from the integrated Mass Shooter Database and Demographic Dataset. This analysis focused on uncovering patterns and disparities related to mass shootings within specific communities based on demographic characteristics.

**Visualization 2.1:**

* I created a scatter plot visualization using Tableau to represent the relationship between different racial demographics (e.g., Asian, Black, White) and the count of mass shootings.
* Each data point (circle) on the scatter plot represents a racial demographic category and its corresponding mass shooting count.
* The color of each circle reflects the employment status associated with the demographic category, providing additional insights into the socioeconomic context.

**Key Insights and Interpretation:**

* The scatter plot visualization facilitates the identification of racial demographics that may be disproportionately affected by mass shootings.
* Through this visualization, I observed that the highest number of mass shootings were attributed to the white demographic category, particularly among those who were unemployed.
* This insight sheds light on potential correlations between race, employment status, and mass shooting incidents.

**A screenshot of a graph

Description automatically generated**

**Visualization 2.2:**

* I created a bar chart, by plotting immigration status categories (e.g., immigrants vs. non-immigrants) against the number of mass shooting incidents.
* The bar graph visualization I created was useful in examining the relationship between immigration status and the frequency of mass shootings.

**Key Insights and Interpretation:**

* The visualization revealed that the majority of mass shooting cases were committed by individuals categorized as non-immigrants.
* It highlighted the importance of considering immigration status as a demographic factor when studying mass shooting occurrences and their correlates.

**A graph of immigration status

Description automatically generated**

**Description of Data Use and Visualization for Research Question 3:**

How does the incidence of mass shootings per capita vary across U.S. states, whether socioeconomic indicators, including poverty rates and unemployment rates, can predict the likelihood of mass shootings occurring in certain states and counties.

**Visualization 3.1:**

* For this analysis, I employed a calculated field to determine the mass shootings per capita for each state. This calculated field was derived from the number of mass shootings and the total population of each state, allowing for a standardized comparison of mass shooting incidence relative to population size across different states.
* I then created a visualization using Tableau, where the mass shootings per capita were plotted on the x-axis and states were represented along the y-axis.
* The color encoding in the visualization represented the state-level unemployment rate, providing insights into the relationship between mass shooting rates and unemployment across various states.

**Key Insights and Interpretation:**

* **California's Low Mass Shooting Rate and High Unemployment:**

Conversely, California showed a lower mass shooting rate per capita despite experiencing a higher unemployment rate compared to New Hampshire. This contrast underscores the complexity of factors influencing mass shooting incidents, indicating that unemployment alone may not be a sole determinant.

* **Interpretation of Relationship:**

The visualization suggests that while unemployment rates may influence certain socio-economic contexts, they do not exclusively dictate the incidence of mass shootings. Other socio-economic, cultural, and environmental factors likely play significant roles in shaping regional variations in mass shooting rates.

**A graph with a bar chart

Description automatically generated with medium confidence**

**Visualization 3.2:**

* For this analysis, I created a bar chart using Tableau to visualize the poverty rate across the different counties.
* To address this question, I utilized data from the Demographic Dataset to visualize and analyze poverty rates across all U.S. counties.

**Key Insights and Interpretation:**

* **Variation in Poverty Rates:**

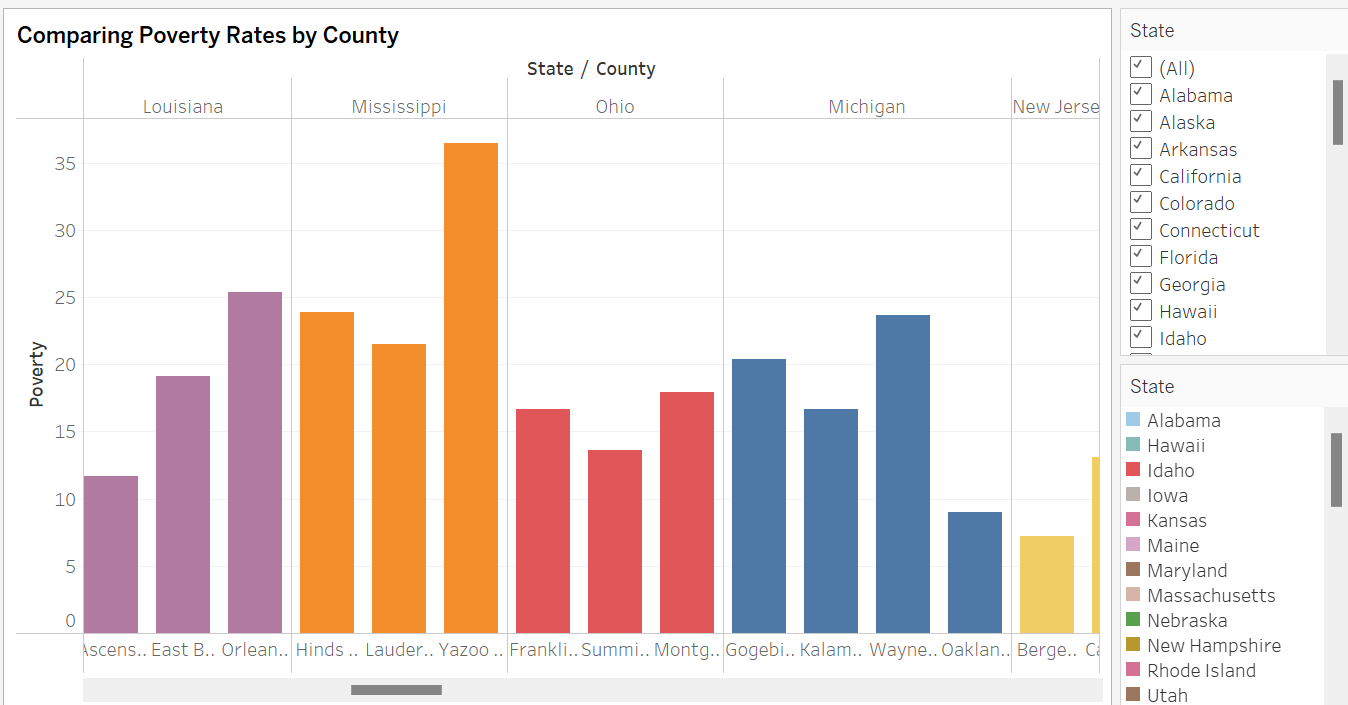
The visualization reveals significant variation in poverty rates across U.S. counties, highlighting disparities in economic conditions at the local level. Some counties exhibit higher poverty rates, indicating greater economic challenges, while others have lower poverty rates, reflecting better socio-economic conditions.

* **Regional Patterns and Disparities:**

By observing poverty rates within states, we can identify regional patterns and disparities in economic well-being. Certain states may have clusters of counties with similar poverty profiles, suggesting shared socio-economic factors influencing poverty levels.

* **Factors Influencing Poverty:**

The analysis prompts consideration of factors influencing poverty rates, such as educational attainment, employment opportunities, healthcare access, and social policies. Understanding these determinants is crucial for developing targeted interventions to address poverty and promote economic growth at the county level.

****

**Description of Data Use and Visualization for Research Question 4:**

What are the variations in the number of victims killed in mass shootings across different U.S. states, and what factors may contribute to the differences observed?

**Visualization 4:**

* For this analysis, I created a bar chart visualization using Tableau to represent the number of victims killed in mass shootings per state. The x-axis displays the states, and the y-axis represents the count of victims killed.
* The data used for this visualization was sourced from the Mass Shooting Database, providing comprehensive information on mass shooting incidents across U.S. states.

**Key Insights and Interpretation:**

* **California's High Number of Victims Killed:**

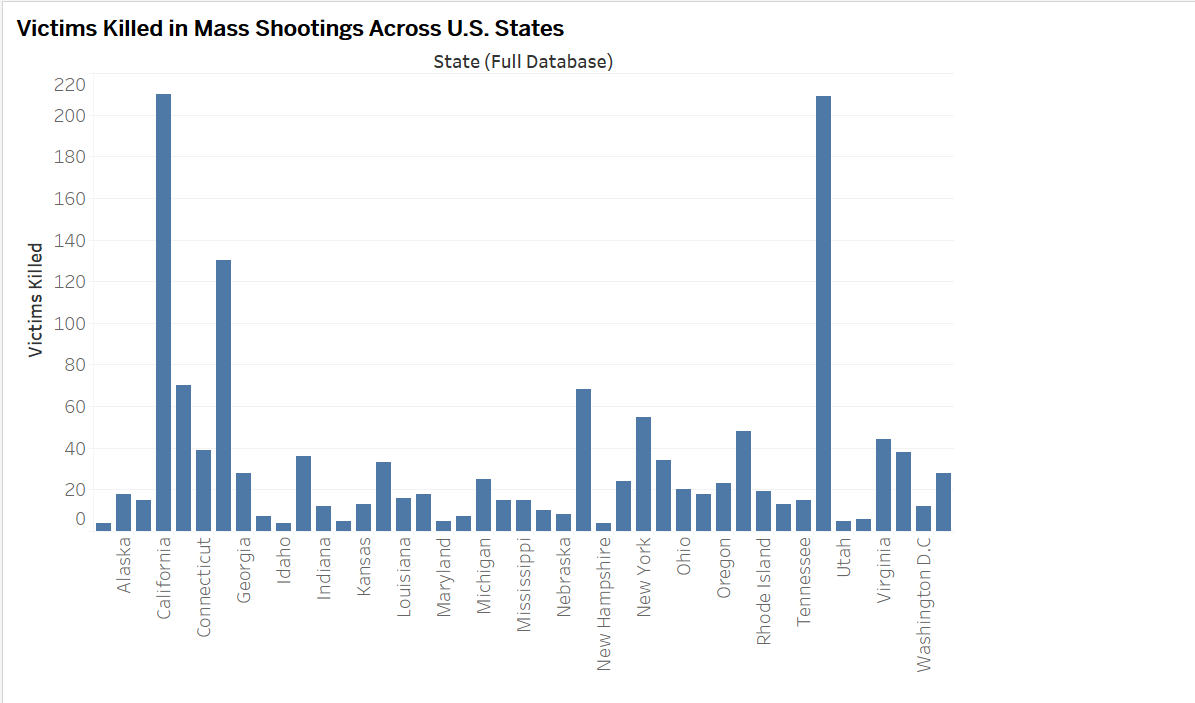
The visualization reveals that California has the largest number of victims killed in mass shootings among all analyzed states. This finding suggests a concerning trend of severe mass shooting incidents in California compared to other states.

* **Texas and Florida as Follow-Up States:**

Texas and Florida emerge as the states with the next highest number of victims killed in mass shootings after California. This ranking underscores the prevalence of mass shooting impacts in these populous states.

* **Factors Contributing to Differences:**

Several factors may contribute to variations in mass shooting casualties across states, including population density, gun laws and access, socio-economic disparities, and access to mental health services. These factors influence the frequency and severity of mass shootings within each state.

****

**Description of Data Use and Visualization for Research Question 5:**

What are the characteristics and impacts of mass shootings at educational institutions in terms of incident frequency, fatalities, and injuries?

**Visualization 5:**

* For this analysis, I created a bar chart using Tableau to visualize the characteristics and impacts of mass shootings at educational institutions.
* The chart displays three key metrics: the number of mass shooting incidents, the number of victims (fatalities), and the number of injuries specifically occurring at educational institutions.

**Key Insights and Interpretation:**

* **Frequency of Incidents at Educational Institutions:**

The visualization reveals the frequency of mass shooting incidents occurring at educational institutions, highlighting the scope and prevalence of such events in educational settings.

This visualization helped uncover that over the years, there were 25 mass shooting incidents at educational institutions, resulting in a significant number of fatalities and injuries, typically ranging between 200 to 300 individuals.

* **Implications for Safety Measures and Prevention:**

Analyzing incident characteristics and impacts can inform the development of targeted safety measures and prevention strategies tailored to educational settings. Insights gained from this analysis are valuable for enhancing security protocols and reducing risks within schools and universities.

**A screenshot of a computer screen

Description automatically generated**

**Description of Data Use and Visualization for Research Question 6:**

How does the incidence of mass shootings vary across different community type in the United States.

**Visualization 6:**

* For this analysis, I created a bar chart using Tableau to visualize the number of mass shootings in rural, urban, and suburban areas.

**Key Insights and Interpretation:**

* By visually representing the number of mass shooting cases in each community category (urban, suburban, rural), the visualization clearly revealed that majority of mass shooting incidents occurred in urban areas.
* This visualization provided a straightforward and impactful illustration of the distribution of mass shootings across different community types, offering valuable insights into the geographical patterns.

**A screenshot of a graph

Description automatically generated**

**Dashboard:**

**A screenshot of a computer screen

Description automatically generated**

**A screenshot of a computer screen

Description automatically generated**

# Conclusion

In conclusion, this project has provided valuable insights into the complex dynamics surrounding mass shootings in the United States, focusing on county and state-level factors that influence the incidence, characteristics, and impacts of these tragic events. By leveraging comprehensive datasets on mass shootings and demographic characteristics, we have explored critical research questions related to mass violence and its correlates across diverse communities.

**Key Findings and Contributions:**

**Geographic Patterns and Population Density:**

Our analysis revealed that areas with higher population density, such as urban centers in states like California, Texas, and Florida, tend to experience more mass shootings. This suggests a potential link between population density and mass shooting occurrence, highlighting the importance of considering geographical factors in understanding the spatial distribution of mass violence.

**Demographic Disparities**:

We identified demographic factors such as race and immigration status as potential correlates of mass shooting frequency. Notably, the analysis showed that the majority of mass shooting incidents were committed by individuals categorized as non-immigrants, emphasizing the relevance of demographic characteristics in studying mass violence.

**Socioeconomic Predictors:**

Our examination of socioeconomic indicators, including poverty rates and unemployment rates, offered insights into their relationship with mass shooting incidence at the state and county levels. While unemployment rates were associated with certain socio-economic contexts, they did not exclusively dictate the incidence of mass shootings, suggesting the influence of multifaceted factors.

**Educational Settings:**

Mass shootings at educational institutions emerged as a significant area of concern, with a notable frequency of incidents resulting in substantial fatalities and injuries. This underscores the critical need for targeted safety measures and prevention strategies within educational settings.

**Community Comparisons:**

By comparing mass shooting rates across urban, suburban, and rural areas, we highlighted the concentration of incidents in urban centers. This visualization provided valuable insights into the distribution of mass shootings across different community types, aiding in understanding geographical patterns and influencing factors.

**Implications and Policy Recommendations:**

The findings from this study have important implications for policy development and intervention strategies aimed at preventing and responding to mass shootings. Targeted interventions addressing demographic disparities, socioeconomic inequalities, and geographical variations can contribute to mitigating the risk of mass violence and enhancing community safety. Additionally, enhancing security measures within educational institutions and implementing evidence-based prevention programs tailored to specific community types are crucial steps in reducing the impact of mass shootings.

**Further exploration:**

Moving forward, additional research should explore how mass shooting rates and patterns in the United States compare with those in other countries with varying firearm regulations and socio-cultural contexts, aiming to derive lessons for policy and prevention efforts. This comparative analysis could shed light on the impact of firearm accessibility and prevalence within communities on variations in mass shooting incidence rates across different geographic areas. Understanding these relationships is crucial for developing evidence-based policies and interventions aimed at reducing the occurrence and impact of mass shootings globally.