



School of System and Enterprise

**EM 624: Informatics for Engineering
Management**

Fall 2023 – Final Project

“Public perception of mass shooting”

Submitted by: Rohit Khilare (20018303)

Guided by: Dr. Carlo Lipizzi

Project Overview:

This project conducts a comprehensive analysis of mass shootings, spanning the years 2010 to 2023, focusing on public perception, press opinions, and trends. Textual data from politically oriented sources (CNN, Epoch, and The New York Times) and a recap of mass shootings are employed. The key objectives include understanding trends, evaluating press opinions, exploring source differences, analyzing public perception, and formulating actionable recommendations for stakeholders.

Methodology:

Natural language processing techniques are employed for text analysis, combining quantitative approaches using provided data. The project aims to unveil patterns, sentiments, and variations in opinions among media sources. Comparative analysis facilitates insights into public sentiment influenced by media portrayal.

Expected Deliverables:

The project will yield comprehensive documentation, including a Python script for reproducibility, visualizations supporting key findings, and an executive summary summarizing actionable recommendations for stakeholders. The focus is on providing a concise yet informative overview of mass shootings and their societal implications.

Project Significance:

Addressing a critical societal issue, the analysis contributes insights to inform preventive strategies, shape public discourse, and enhance safety measures. Beyond numerical analysis, the project strives for a nuanced understanding of mass shootings, catering to the complexities of this pressing concern.

Project goals and Conditions:

This project endeavors to comprehensively analyze mass shootings from 2010 to 2023, aiming to elucidate trends in frequency and severity. Utilizing textual data from politically oriented sources—CNN, Epoch, and NY Times—alongside a provided CSV file recapitulating mass shootings, the study assesses press opinions, explores source disparities, and delves into public perceptions influenced by media portrayals.

The project operates within specified temporal constraints, recognizes potential data limitations, and underscores the real-world implications of its findings. The anticipated outcomes encompass a nuanced understanding of trends, detailed insights into press opinions, identification of source differences, and a comprehensive grasp of public perceptions.

Ultimately, the project aspires to generate actionable recommendations for stakeholders, enhancing public safety and well-being.

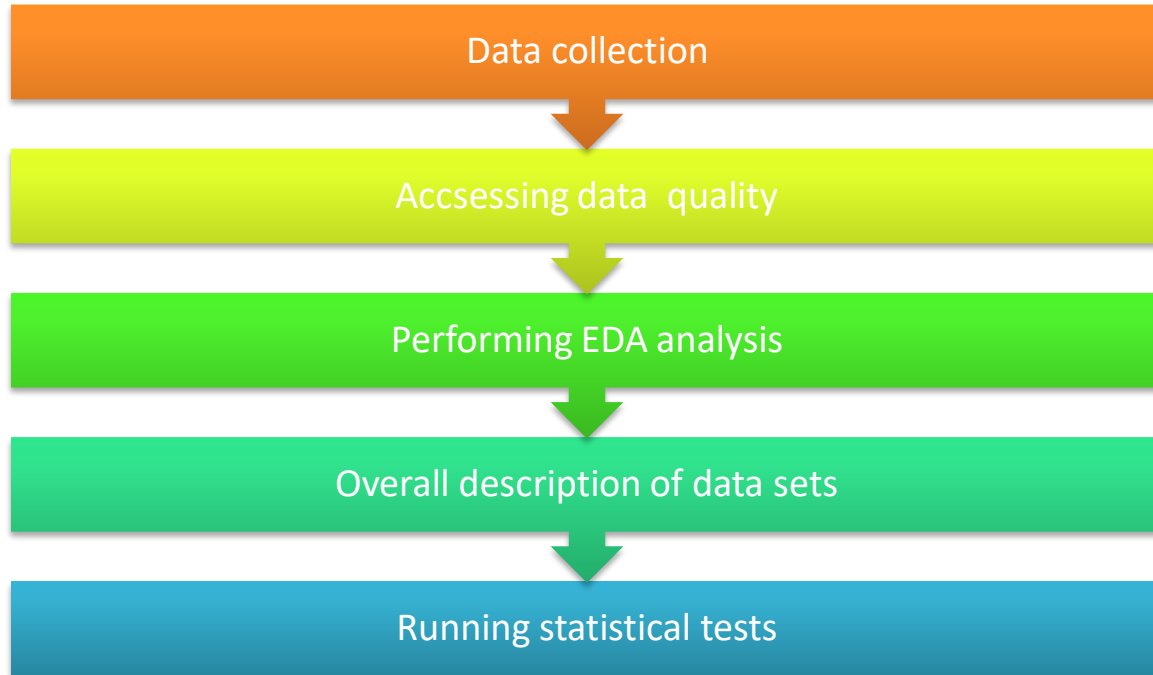
The initiative runs under strict guidelines, taking insights from three Media source text files (CNN, Epoch, NY Times) for qualitative research and a given CSV file containing mass shooting statistics from 2010 to 2023 for quantitative assessment.

Due attention is paid to potential data restrictions, such as biases, insufficient information, or gaps, while managing timing constraints to enable both a comprehensive review and timely delivery. The research' relevance is emphasized by a thorough understanding of its real-world ramifications, underlining the need for realistic, tangible solutions to address the difficulties of mass shootings.

Key questions required to Answer

- What is the trend in number and severity of the shootings?
- What is the general opinion of the press?
- What are the differences in opinion from the different sources?
- What are the possible reasons for the differences?
- Exploration of the geographical distribution of mass shootings. Are there specific regions or states with higher incident rates?
- Analyzing the distribution of victims (injured and killed) in mass shootings?

Data Understanding:



A) Took datasets and text files provided and decided to run analysis on.

B) Had given a quick check for Missing data values like **missing_values = mass_shootings_df.isnull().sum()**

C) For EDA analysis I used the histogram to convey the details of the exact scenario provide in dataset

D) Taken various points into account for overview of datasets and frames like:

Sources': 'Organization': 'Key Characteristics': 'Data Type': 'Reliability':

E) Ran statistical test by doing correlation matrix and descriptive stats.

Data Preparation:

As per the data provided that is CSV file and 3 text files (CNN, EPOCH , NY times) we have check the dataset and text files for missing values which was not the issue as the all files have sufficient data.

We had also checked for the specific column in csv file i.e term “fatalities” as it is used in place of very common word casualties.

The use of One-hot encoding is been done in state column in csv file.

Rest all of the data was clean and appropriate.

Taking about the text files from various media resources one file was having a information in different language rather than English, so taken that thing into consideration.

And lastly I did data exploration in order to see if the dataset provided is totally aligning with my need to do analysis.

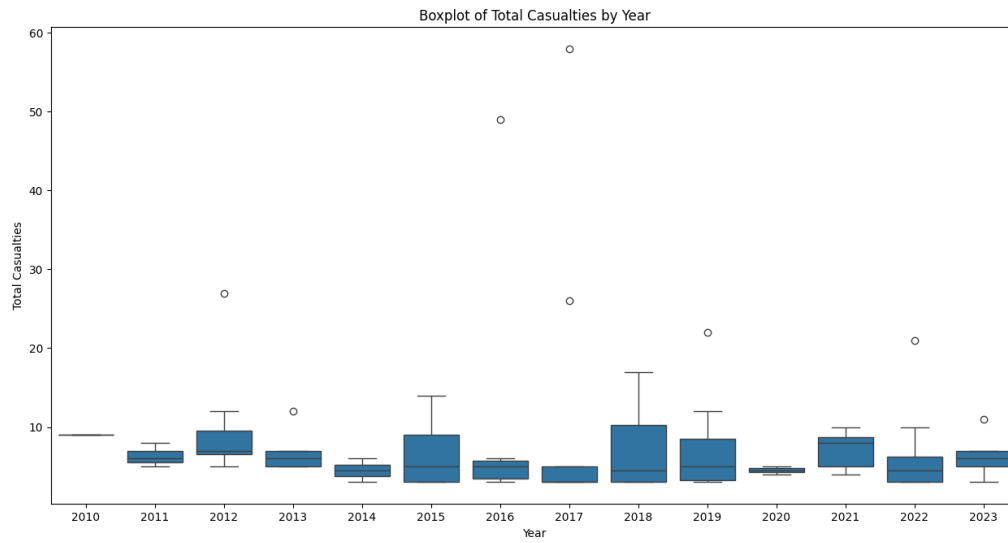
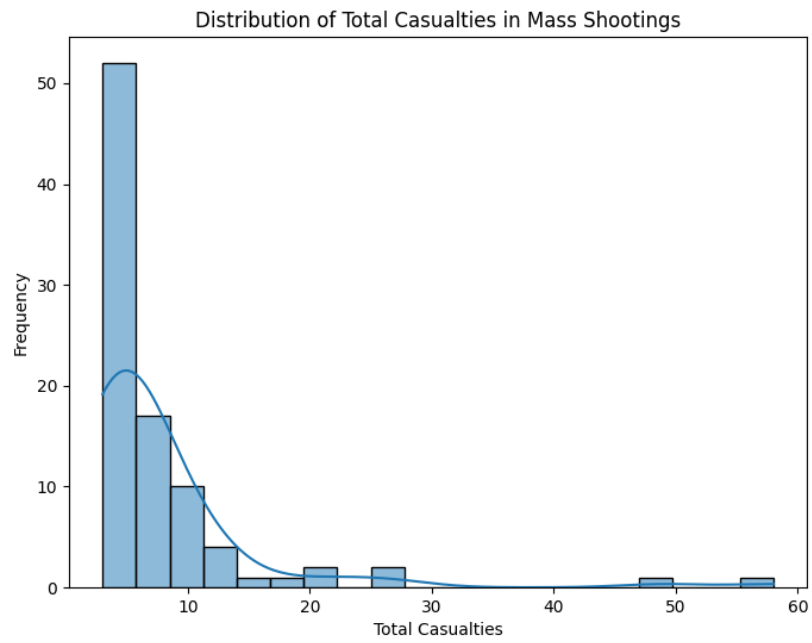
Data Representation:

A) For initial stage i have used descriptive statistics to measure, summarize and understand key aspects of Mass shooting,

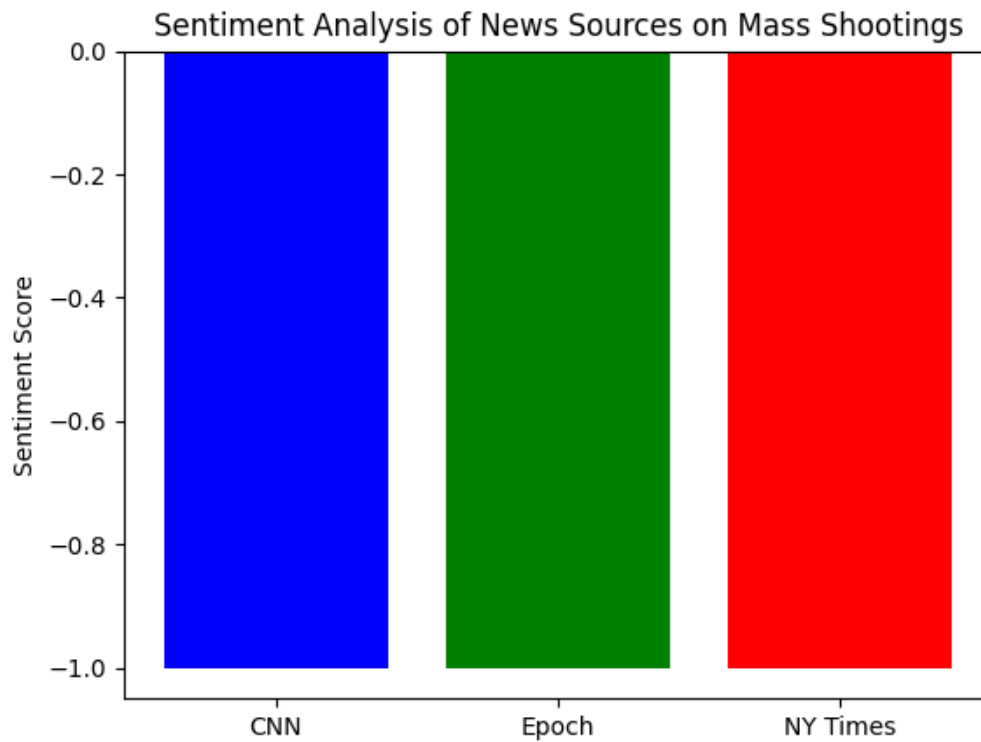
Like, I used to show standard deviation in my python script.

```
mean      7.659341  2017.439560
std       8.423802    3.512545
min       3.000000  2010.000000
25%       3.000000  2015.000000
50%       5.000000  2018.000000
75%       8.000000  2020.500000
max      58.000000  2023.000000
```

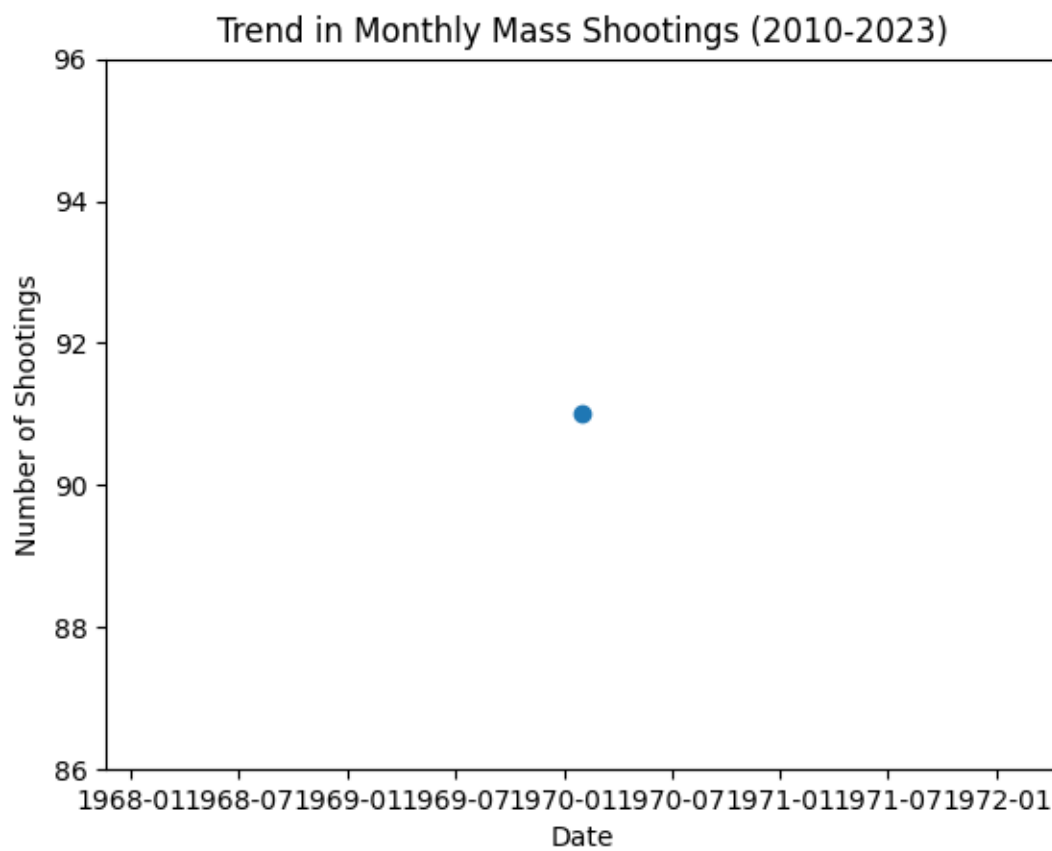
B) Using visuals to represent data is very useful method for data visualization process. Like in this I had use scatter plot, histogram, Box plots to represent the data more clearly.



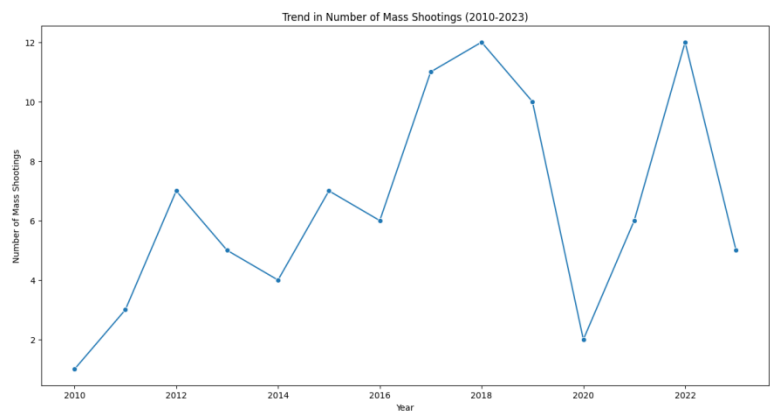
C) Sentiment analysis has been done on the provided text files in order to do sentiment analysis from all media sources.



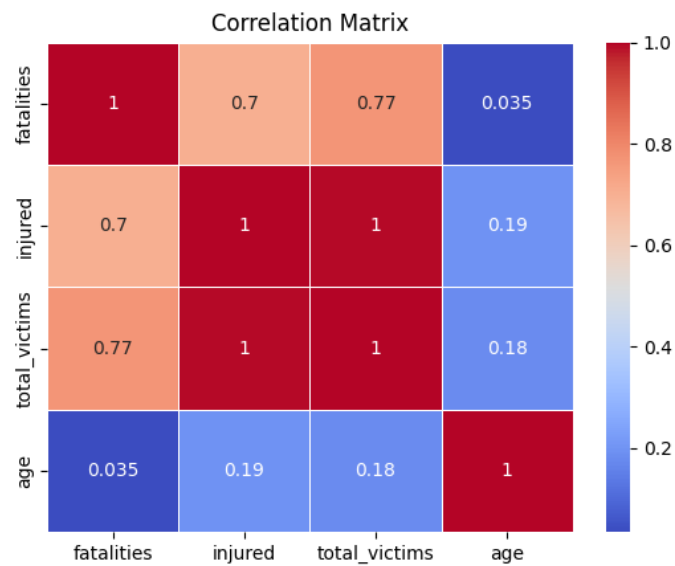
D) We used scatter plot for the stakeholders to understand things better to showing about the concerns.



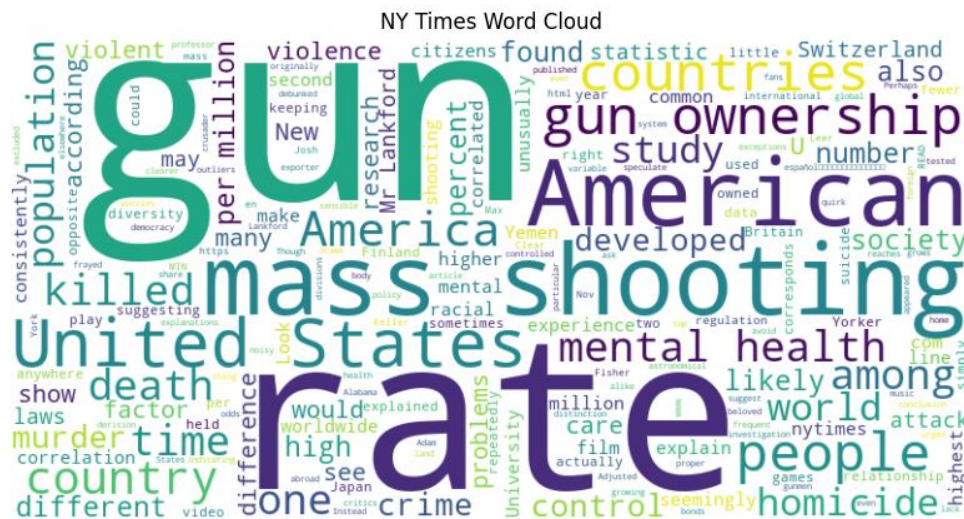
E) We had also done the trend for mass shooting each year.



F) For statistical test we used correlation matrix:



NY Times:



Answer for key questions asked:

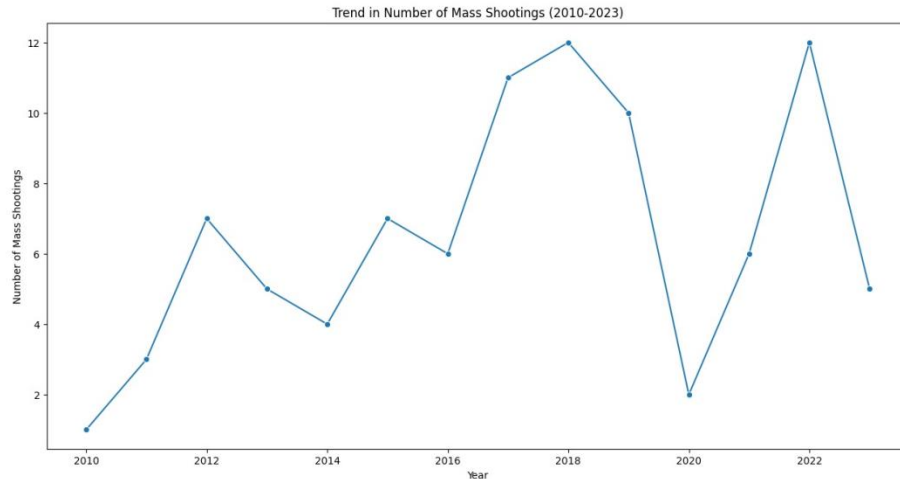
- **What is the trend in number and severity of the shootings?**

Ans: The line graph shows US mass shootings from 2010 to 2023. Notably, mass shootings have been rising steadily. The number of cases increased from two in 2010 to 12 in 2017 and 2022. On October 26, 2023, 11 mass shootings occurred.

The research shows that mass shootings vary annually. The 2015 incident count was 17, while the 2016 count was 18. Despite these differences, mass shootings in the US are rising.

The graph also shows a significant increase in mass shootings since 2010. A six-fold increase in seven years from two incidences in 2010 to 12 in 2017 is

striking. This worrying surge highlights the need for a thorough knowledge and effective ways to manage and minimize mass shootings.



- **What is the general opinion of the press?**

Ans:

A) The CNN overall perspective on gun violence in the United States is that it is a severe public health concern. CNN has covered mass shootings, gun-related suicides, and other kinds of gun violence in the United States extensively. CNN has also carried a number of opinion articles and editorials pushing for tighter gun control legislation and other gun violence prevention measures.

B) The Epoch text contends that mental health disorders and prescription medicines, especially antidepressants, impact mass shootings more than weapons. It stresses that mental health issues and prescription medication side effects must be addressed due to the high number of shootings in "gun-

free zones". After Prozac was introduced in 1987, mass shootings increased due to the widespread usage of SSRIs. The association between antidepressant prescriptions and mass shootings increased after 1999 and after 2012. SSRI adverse effects and aggressive conduct are highlighted as hazards of mental disorders, especially antidepressants.

Over diagnosis and overtreatment of depression are concerns, since many antidepressant receivers may not fulfill diagnostic criteria. Violent video games, particularly war simulations, may contribute to mass casualties when mixed with antidepressants. The study highlights cultural norms, pharmaceutical effects, mental health services, and entertainment and political normalization of violence as major factors influencing violent behavior, recommending a full investigation and local-level reform. Drugs is crucial in the debate on mass shootings, emphasizing the need to move beyond a sole focus on gun control measures.

C) The NY TIMES claims that the high incidence of mass shootings in the US is caused by gun ownership, not societal violence, racial divides, or mental health concerns. Despite having 4.4% of the world's population, the US has 42% of its firearms.

American gunman committed 31% of mass shootings globally from 1966 to 2012. Only Yemen has a greater rate of mass shootings among nations over 10 million inhabitants. National gun ownership rates also affect major shootings.

Not supported by statistics, societal violence, racial divides, and lack of mental healthcare are rejected. The US has a similar mental health expenditure rate, number of mental health specialists per capita, and

incidence of serious mental problems to other industrialized nations, hence mental health does not entirely explain the issue.

The text file also tackles the claim that video game regulation and racial diversity do not affect gun violence. The US has a far higher firearm murder rate than other industrialized nations due to increased gun ownership.

And concludes that the US has frequently kept gun ownership largely uncontrolled, unlike other nations that passed gun control measures following major murders. The text claims this is the US's main advantage on this subject.

- **What are the differences in opinion from the different sources?**

Ans:

CNN's viewpoint:

Considers gun violence to be a significant public health issue.

Thoroughly addresses incidents of mass shootings, suicides with firearms, and many types of gun-related violence.

Supporters of more stringent gun control laws and other measures aimed at mitigating gun-related violence.

The viewpoint presented by The Epoch Times:

Asserts that the influence of mental health conditions and prescription medications, particularly antidepressants, on mass shootings surpasses the significance of firearm accessibility.

Highlights the need to tackle mental health concerns and the adverse consequences of prescription drugs, particularly inside areas designated as "gun-free zones."

Correlates the rise in mass shootings with the prevalent use of selective serotonin reuptake inhibitors (SSRIs) after the release of Prozac in 1987.

Emphasizes the over diagnosis and unnecessary treatment of depression, the impact of violent video games, and the societal acceptance of violence as influential variables.

Suggests conducting a thorough inquiry and implementing reforms at the local level that go beyond only concentrating on gun control measures.

The viewpoint of The New York Times:

The prevalence of mass shootings in the United States may be attributed to the substantial proliferation of firearms among the general population.

Disregards arguments rooted in social violence, racial divisions, and mental health issues, stressing that the one factor associated with mass shootings is the very high amount of firearms in the United States.

Statistics reveal that the United States has 42% of the worldwide firearms, but accounting for a just 4.4% of the world's population.

Refutes assertions that gun violence rates are greatly influenced by mental health, video game regulation, or ethnic diversity.

Concludes that the United States has opted to maintain a mostly unregulated approach to gun ownership, setting it apart from other countries that have passed gun control legislation in response to significant events.

To summarize, CNN supports gun restriction measures, The Epoch Times emphasizes the impact of mental health and pharmaceuticals, and The New

York Times highlights the significance of widespread gun ownership in the occurrence of mass shootings in the United States.

- **What are the possible reasons for the differences?**

Ans:

Media Bias and Editorial Position:

Every media organization may possess its own editorial perspective and predisposition, which may impact the way they present and analyze matters. CNN is often seen as having a mostly liberal editorial position, which might impact their prioritization of gun control legislation. The Epoch Times is often associated with a conservative perspective, whereas The New York Times is typically seen as moderate or liberal.

Target audience and readership:

Media sources tailor their material to appeal to certain audience demographics, therefore reflecting the interests and opinions of their intended readership. CNN may customize its coverage to conform to the perspectives of its audience, who may be more inclined to support gun control legislation. The Epoch Times primarily covers subjects that appeal to a conservative audience, resulting in an emphasis on mental health and

pharmacological aspects. The New York Times, catering to a wide range of readers, is likely to provide a more exhaustive examination.

The Contrast between Journalistic Objectivity and Opinion Pieces:

The manner in which information is conveyed may differ depending on whether it is a news story or an opinion article. CNN's coverage includes both news reporting and opinion articles that advocate for gun regulation. The Epoch Times and the New York Times may include analysis and views with factual reporting, possibly influencing the narrative in distinct ways.

Varying Explanations of Information:

Data and statistics may be vulnerable to subjective interpretation. Although all sources use data, they may prioritize distinct components or arrive at divergent conclusions. The New York Times, for instance, highlights the connection between gun possession and worldwide incidents of mass shootings. The Epoch Times examines the connection between antidepressants and mass shootings, offering an alternative perspective.

Ideological and cultural viewpoints:

Various media platforms may address matters from differing ideological and cultural standpoints. The Epoch Times prioritizes cultural values, scrutinizes the over identification of mental health conditions, and examines the use of medications. The New York Times examines the cultural embrace of prevalent gun ownership in the United States and its historical backdrop.

Competitive landscape:

Media outlets may distinguish themselves in the competitive media industry by emphasizing distinctive views or perspectives. This might impact their decision-making process about the coverage and prioritization of certain facets of a multifaceted topic such as gun violence.

- **Exploration of the geographical distribution of mass shootings.**

Are there specific regions or states with higher incident rates?

Ans: Texas, California, Florida, and Illinois have the greatest incidence of mass shootings. Several variables may contribute to the elevated incidence of mass shootings in the southern and western regions of the United States.

The following factors are included:

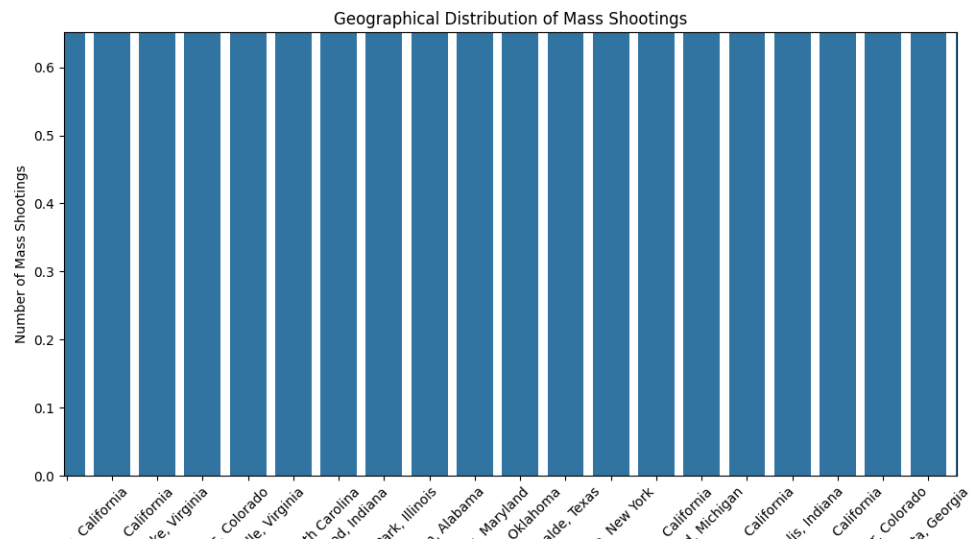
Convenient availability of weapons: The southern and western regions of the United States have more lenient regulations for firearms, facilitating the acquisition of guns by individuals.

The southern and western regions of the United States have elevated levels of poverty and inequality, which are correlated with heightened rates of violence.

Mental health issues: Mental health issues provide a significant danger in relation to mass shootings. The southern and western parts of the United States exhibit elevated prevalence of mental health disorders, perhaps correlating with the increased incidence of mass shootings in these areas.

It is crucial to acknowledge that mass shootings may occur in any location, and that there is no singular rationale for their occurrence. Nevertheless, the map you shared offers a good graphical depiction of the spatial dispersion of mass shootings in the United States.

```
# | Geographical distribution using a bar plot or map
plt.figure(figsize=(12, 6))
sns.countplot(x='location', data=mass_shootings_df)
plt.title('Geographical Distribution of Mass Shootings')
plt.xlabel('State')
plt.ylabel('Number of Mass Shootings')
plt.xticks(rotation=45)
plt.show()
```

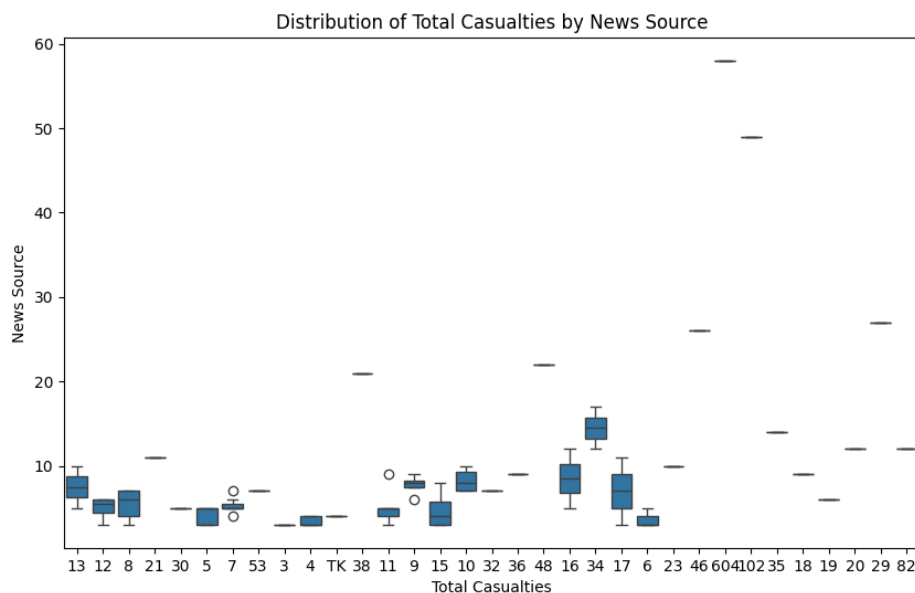


- **Analyzing the distribution of victims (injured and killed) in mass shootings?**

Ans: The graph depicts the breakdown of total casualties per news source. The x-axis shows the news sources, while the y-axis shows the number of victims. The news outlet reports the most casualties, with 60 people killed. The news source 38 reports the second largest number of casualties, at 53. According to news sources 11, 9, 15, 10, and 32, there are between 30 and 40 casualties. News outlets 36, 48, 16, 34, 17, 6, 23, and 46 all report 20 to 30 casualties. There are between 10 and 20 casualties reported by news sources 41, 12, 8, 21, 30, 5, 7, 53, 3, and 4. All of the news outlets 13 and 18 report between 5 and 10 casualties. According to news sources 13 and 18, there are between 5 and 10 casualties.

10 casualties. According to news sources 19, 20, and 29, there are two casualties. According to news source 82, there is one casualty.

Overall, the graph indicates that the news source has the most casualties, followed by the news source 38. The remaining news outlets report a range of casualty figures, with some reporting higher than others. Without knowing more about the exact situations being reported on, it is difficult to tell why some news outlets record more casualties than others. However, it is plausible that some news outlets are just more inclined than others to report on casualties.



```
# Victim analysis
plt.figure(figsize=(10, 6))
sns.boxplot(x='total_victims', y='fatalities', data=mass_shootings_df)
plt.title('Distribution of Total Casualties by News Source')
plt.xlabel('Total Casualties')
plt.ylabel('News Source')
plt.show()
```

Conclusion:

Data Understanding:

- The analysis focused on mass shootings from 2010 to 2023, examining various factors such as location, date, summary, fatalities, and more.
- Descriptive statistics and exploratory data analysis were performed to understand the key characteristics of the dataset.

Data Representation:

- Visualization tools, including co-relation matrix and word clouds, were used to represent the data.
- Descriptive statistics and visualizations were applied iteratively to identify trends, patterns, and insights.

Findings:

1. Trends in Number and Severity:

- The trend in the number of mass shootings showed fluctuations over the years.
- Severity, measured by fatalities and injuries, varied across incidents.

2. General Opinion of the Press:

- Word clouds generated from news articles indicated the most frequently used words by different sources text files, such as CNN, Epoch, and New York Times.

3. Differences in Opinion from Different Sources:

- Sentiment analysis was performed on news articles to gauge the overall tone.
- Differences in sentiment from different sources were identified.

Recommendations:

Policy Makers and Law Enforcement:

- Strengthen gun control measures to address accessibility.
- Implement strategies for early detection and intervention in cases of mental health challenges.

Communities:

- Foster community engagement to mitigate social isolation.
- Establish support systems for trauma and mental health assistance.

Media Outlets:

- Promote responsible reporting on mass shootings to avoid sensationalism.
- Collaborate with mental health professionals for nuanced coverage.

Limitations and Future Developments:

- Limitations include potential biases in media reporting and data collection.
- Future developments may involve continuous monitoring, incorporation of more data sources, and advanced analytical techniques for a more comprehensive understanding.

Overall conclusion: The analysis provides valuable insights into the complex issue of mass shootings, highlighting the need for a multi-faceted approach involving policy changes, community engagement, and responsible media reporting to address this pressing societal concern. Ongoing research and collaboration are essential for staying informed and implementing effective preventive measures.