# Analysis Report: Mass Shootings in the United States

Introduction:

Mass shootings have devastating consequences and are a serious issue in the United States. To address this issue effectively, it is important to understand the trends and characteristics of mass shootings. In this report, we analyze a dataset containing information on mass shootings that have occurred in the United States.

Data Cleaning:

Before analyzing the data, we cleaned the dataset to make it easier to work with. We replaced some of the values in the dataset, such as gender and race, and replaced NaN values with appropriate values.

The trend of Fatalities over the Years:

To see the trend of fatalities over the years, we created a bar plot using the year and fatalities columns from the dataset. The plot shows a significant increase in fatalities in recent years. The worst mass shooting in the US occurred on October 2, 2017, and killed 58 people.

Visualization of Mass Shootings on the U.S. Map:

To visualize mass shootings on the U.S. map, we used the Folium library. We added geolocation data for each location, extracted latitude and longitude data from the geolocation data, and created a map centered on the first location. We added markers for each location and displayed the map.

Correlation between Shooter and Race, Gender:

To determine if there is any correlation between the shooter and their race or gender, we used the Seaborn library to create a countplot for gender. The plot shows that there are more male shooters than female shooters. However, we did not perform any statistical tests to determine if there is a correlation between the shooter and their race or gender.

Common Locations and Areas with Most Shootings:

We analyzed a dataset of US mass shootings from 1966 to May 2022. We printed the top 10 cities and states with the most shootings, grouped by city and state, respectively. The top 10 cities with the most shootings were Aurora, Orlando, Orange, Fort Hood, San Francisco, Parkland, Dallas, Seattle, Colorado Springs, and Springfield. The top 10 states with the most shootings were California, Texas, Florida, Colorado, Washington, Pennsylvania, Wisconsin, New York, Illinois, and Ohio. We also grouped the shootings by location type and counted the number of incidents. According to the data, the most common location type was schools, followed by religious institutions, workplaces, and outdoor venues.

Correlation between Fatalities/Injuries and Time:

We created new columns for year and month and calculated the correlation coefficients between these columns and the number of fatalities. The correlation coefficients were negative but very weak, suggesting that there is no significant correlation between fatalities and time.

Conclusion:

In conclusion, our analysis provides insights into the most common locations and areas with the most shootings in the US. We also visualized the trend of fatalities over the years and mapped the shootings on the U.S. map. However, to fully understand the issue of mass shootings, further analysis is needed, including statistical tests to determine the correlation between the shooter and their race or gender.