Group: Kate Lautenbach, Stephanie Jacinto, Emily Kelly
DS 2001
Professor Lange
21 October 2024

Police Fatality Analysis

Introduction:

For our final project, we analyze police brutality rates in the United States and explore how different factors have contributed to these trends in the last 20 years. This has been a highly debated and ongoing social issue for many decades in the United States, due to disagreement of the presence of systemic bias and critique of law enforcement practices. With the rise in media coverage regarding these incidents, we find it important to keep track of these rates and develop methods to keep them in a steady decrease. We analyzed factors such as the race, gender, age, and location of the civilians, and tracked the trend of the total fatalities by police encounters throughout the years. From this information, we can identify biases, unethical practices, and gaps in police training that contribute to harm in communities.

This data comes from Fatal Encounters, which is a website that has collected information on police-related deaths from 1999 to 2021 to gauge what people affected look like, where they live, and how they died. Not only does this data set encompass interactions from police brutality, but also all instances of self-defense, suicide, and medical emergencies.

Methodology:

Using skills we have used within this course we have created multiple graphs identifying, age, race, gender, location, and the years that were most affected by fatal police encounters. We cleaned the data by removing the lengthy descriptions, removing the names of the victims, and the links to the news articles. We removed these elements so we were able to process the CSV faster with more ease and because those sections were not relevant to our research.The majority of the cleaning we performed was to convert numerical data into float or integer data types, and separate composite attributes to find information we need. The computations we programmed included: counting the number of equal values within a column and determining the minimum, maximum, and mean values of a column. We comprised and analyzed 5 different graphs. Three bar graphs displaying information regarding the demographics (age, race, and gender) of people who died during police encounters, one scatter plot displaying the locations of the fatalities, and one bar graph displaying the number of fatalities during police encounters by year.

Findings:

The graph titled “Number of victims by 10 year intervals” shows a wide age range of people who died during a police encounter, the youngest being 0 and the eldest being 107. The graph shows that the highest rate of deaths occurred between the ages of 21-30.

The graph presenting the racial makeup of the deceased civilians is labeled: “Number Of Victims Per Race Group”. The two most notable groups with disproportionate fatalities to census data were European-American/white people whose fatalities involving police made up 32% of occurrences tracked by the data set even though white people make up just over 60 percent of America’s population, and African-American/Black people whose fatalities made up 22% of the total fatalities tracked while African Americans make up only 14% of the total population. These numbers suggest a racial bias within police brutality and the care police take to prevent a fatal situation through medical attention or de-escalation techniques.

The graph labeled: “Number of Victims per Gender Group” reveals major disparity in genders experiencing fatalities during police encounters. Over 90 percent of the fatalities are men, which shows an inclination by the police to cause more harm on the basis of gender.

The graph “Plot by Latitude and Longitude” shows each place where the fatality took place. The location graph is similar to what was expected, with high density of data points in major cities. However, there are some cities, not clearly seen on the graph, that are disproportionate to the overall population. For example, Houston has almost double the deaths of New York City despite having a quarter of the population.

“Deaths by Year” tracks the fatalities that occur each year included in the data set. The annual fatalities graph is skewed by lack of data in the earlier years however it improves after 2013 where it looks similar to other graphs of annual fatalities caused by police. Even though the data is lacking, the fact that there is data on this topic previous to 2013, when many other datasets started, provides a wealth of information not easily found elsewhere. The numbers represented in this graph are also high because it shows any encounter resulting in a fatality not just fatalities directly caused by police. The graph shows a decline after 2014, likely due to the social backlash from the death of Michael Brown. Then there was an upward trend in the graph again in 2017, with slightly lower numbers in 2021 due to the data being incomplete. There was a large discrepancy between deaths directly caused by police and fatalities that happen during police encounters when comparing our graph to one that studies fatal counts of police brutality exclusively. This difference highlights the need for police to improve in medical training and de-escalation techniques so they can better protect communities.

Limitations:

Despite the set missing instances of fatal police encounters this data is one of the most comprehensive records available. It can help to find overlooked murders by police and also help identify areas where police need better training to serve communities. However, there are limitations to only using one source. It is missing data from earlier years of its range which is not an accurate representation of whether fatal police encounters have been increasing or decreasing from the beginning of the data set. Also without cross-referencing with population census data, there is a limit on the information that can be extracted. For example, by comparing the data to census population data we would be able to provide a better understanding of cities or zip codes disproportionately affected by fatal police encounters and therefore which areas need the most change in policy. A similar approach can be applied for analyzing the identities of the victims. If this project were to be expanded, graphs could be made to analyze the fatalities based on the intersectionality of the identity of the victims, and there could be clear visualizations of the disproportionate fatalities based on factors such as race, gender, and age.