Krysten Tachiyama

Professor Johnson

CMSI 401: Software Engineering Laboratory

Oct 7, 2020

Article Summary

Fatal Force: 1005 People Have Been Shot and Killed by Police in the Past Year

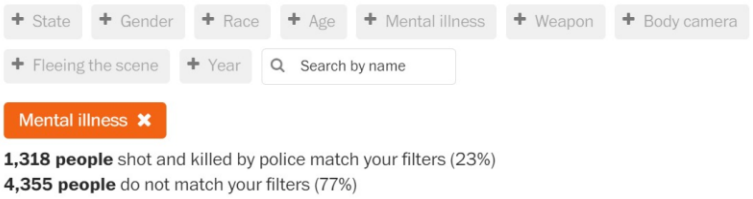
The 2014 fatal shooting of Michael Brown— an eighteen-year-old unarmed black man— rekindled a national debate over law enforcement's use of lethal force and brought attention to the question of how many people are killed by police officers every year. Surprisingly, there was no reliable national database to answer the question. Furthermore, due to the fact that police reporting of shootings is optional, the FBI severely undercounted instances of lethal force leading to fatalities. In response, the Washington Post began to record every U.S. fatal shooting by an on-duty police officer from 2015 to present day.

Some of the data found is exactly what one would expect: African Americans are disproportionately killed by the police compared to other races, and the overwhelming number of victims are young males. Other discoveries are not so obvious, like the fact that the rate of shootings have remained constant throughout the last five years (around 1000/year), or that the highest rates of shootings occur in New Mexico, Alaska, and Oklahoma.

The article does not delve into all of its findings but supplies an interactive element for its readers to find relationships between the data themselves (they even allow you to download the dataset, something very applicable to those of us in CMSI 486 with Dondi). The filtering options include state, gender, race, age, state of mental illness, weapon bore, status of body camera, if the

person was fleeing the scene, and year. Here are some statistics that I found by playing around with the data myself:

- Of those fatally shot, 23% were experiencing a mental health crisis.



- Despite the demand for body cameras rising over the past five years, the number of camera footage provided remains low. The lowest rate of cases that included body camera footage occurred in 2015 (1.3%), and the highest in 2019 (2.4%).

- A black American is more likely to be fatally shot while unarmed by about four to five times the rate of a white American.

$$\frac{\%\ of\ fatal\ shooting, black\ \&\ unarmed}{\%\ of\ black\ Americans} : \frac{\%\ of\ fatal\ shooting, white\ \&\ unarmed}{\%\ of\ white\ Americans}$$
$$= \frac{2.2\%}{13.4\%} : \frac{2.6\%}{72\%}$$
$$= 0.164 : 0.036$$

The importance of representational databases carries a momentous weight in 2020. Covid-19 and gun policy were prevalent issues in the past months and remain top concerns for voters in the upcoming presidential election. Datasets concerning these matters help organize important information for people to internalize and create informed opinions around. For some

reason, there is a pronounced political divide between parties over whether police brutality is systematically racist— It is, and the statistics illustrate it.

The fact that there was no reliable dataset to record fatal police shootings until 2015 is shocking and puts a too much gravity on the single set provided by the Washington Post. If you go to Kaggle and search “police shootings”, the top five results are either exact replicas or more concentrated subsets of data directly pulled from the Post’s fatal shootings database. There’s danger in only having a single source to pull data from. It makes it easier to question the reliability of findings and label them as unsupported. This area of statistics is underrepresented in terms of research and deserves more effort from both reporters and software engineers.

A few other limitations regarding the software system include: 1) The interactive searches provided by the article are limited. There is no way to combine multiple filters in the same category such that readers can see combined subsets of data. For example, you cannot filter the total number of people who had a toy weapon or no weapon at the time they were shot, you must choose one or the other. For some of us who have database experience, this is not a hard query to execute ourselves, but it is safe to say that this is something that is not realistic for most people to do. If you wanted to combine filters via the provided interactive search, you would have to manually add the numbers yourself—a small hindrance for a typical viewer. 2) These searches are also exclusively count queries, with the exception of being able to search for a victim by name. There is no way to list sets of data in ascending/descending order. If a reader wanted to know the top 5 states with the most police shootings, they would have to sift through all the alphabetically ordered states and manually find the ones with the highest counts. 3) There is no regulated system for obtaining data. The Washington Post primarily relies on news accounts, social media, and police reports, which further begs the question of how accurate the

data represents real-world occurrences. For instance, shootings of minority Americans get more media attention than those of white Americans; could this be inaccurately skewing the data? And if so, then how significantly?

The first two problems can be solved by software engineers. For example, a website dedicated to filtering data from the dataset could be created and linked in the article. This way, readers would be able to not only request queries by count, but ordered lists of data, compound filters, and more. Search algorithms could be implemented to provide a more intuitive experience and an advanced search could be included as well. I could see this being a really good project for software engineers interested in the field of data science. Additionally, it would be practical to expand this theoretical website to include data relating to police encounters in general. Instead of just shootings, all fatal encounters by the hands of the police— tasers, suffocation, custody holds, hand combat, etc.— could be added. Data about injuries from encounters, rates of confrontation among neighborhoods/race, and reasons for being pulled over are just a few things that could help create a more aggregated picture of what police encounters look like for the lives of Americans.

With any dataset it is important to consider its limitations and how they may affect the relationships between the data. The database of fatal police shootings provided by the Washington Post is no exception as it contains many limitations of its own that prevent a perfect reporting. Nonetheless, its existence is socially and culturally important, especially in today's political climate and technological age. The Post updates its database consistently, the last being October 5th of this year (2 days ago). As the dataset continues to grow, I hope that more software engineers commit themselves towards informing the general population and promoting social change.

Work Cited

Fatal Force: 1005 People Have Been Shot and Killed by Police in the Past Year. The Washington Post, 5 Oct. 2020, www.washingtonpost.com/graphics/investigations/police-shootings-database/.