Daniel Sanchez

MIS 505

Dr. Christopher Sibona

Module 7 – Final Project Write-Up

August 16th, 2020

For my project, I chose to analyze a report of the use of fatal force by police between January 2015 and August 2020.  
  
This dataset can be found here: <https://github.com/washingtonpost/data-police-shootings>

I reviewed overall trends in the number of killings by police, including the cumulative total, and annual trends. I also examined trends relating to age, race, geographic location and circumstances surrounding the killings.

I have included five dashboards that illustrate my key findings.  
  
[Dashboard 1](#One): Overall, Annual Trends, and Distribution by Age

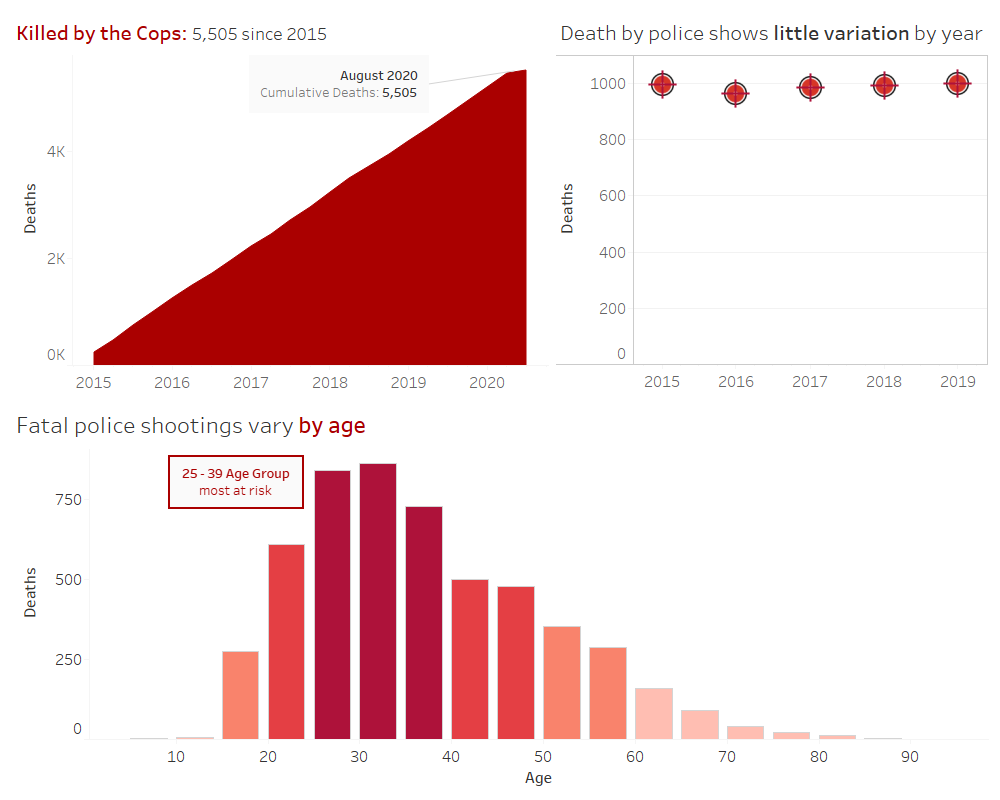
[Dashboard 2](#Two): Race Distribution – By Age, Overall, and by Year

[Dashboard 3](#Three): State Rankings

[Dashboard 4:](#Four) State and City Top 10 with Race Distribution

[Dashboard 5](#Five): Circumstances of Killings – Unarmed, Fleeing, and Mental Illness by Race

**Dashboard 1: Overall, Annual Trends, and Distribution by Age**

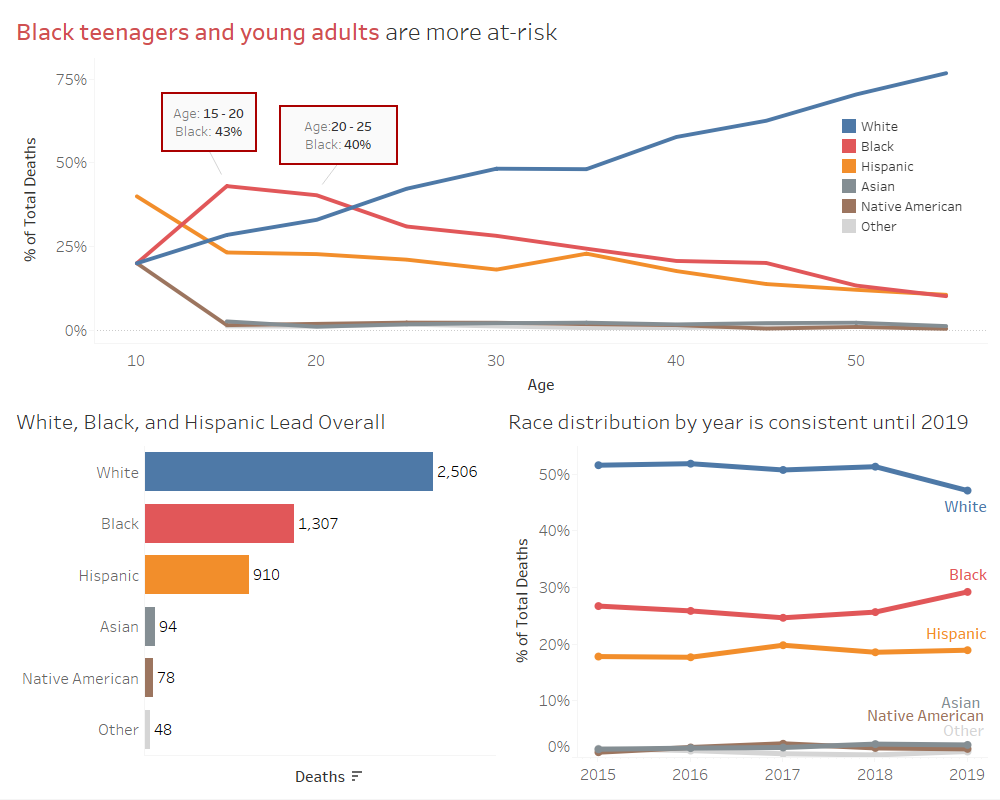


This dashboard illustrates some overall elements of the dataset. On the top left, I used a running total area chart to show the size and scope of the data (January 2015 - August 2020). The title and annotation both show the total number of deaths reported.

On the top right, I created cross-hair markers to show the relative consistency by year. I had first tried using a line chart, but I went with the cross-hair markers to provide more intrigue. I excluded 2020 from this chart, since that year did not provide an entirely complete set of data (only counting up to August). I kept the gridlines here, which makes the chart itself actually look like a firing range.

On the bottom, I placed an age histogram, with 5-year bins. While somewhat duplicative, I chose to emphasize the distribution with a color scale. I was a bit shocked by the maximum and minimum ages represented here, but I instead chose to focus on the most prevalent age grouping. The use of crimson red in this first dashboard is meant to elicit a visceral reaction to the data.

**Dashboard 2: Race Distribution – By Age, Overall, and by Year**

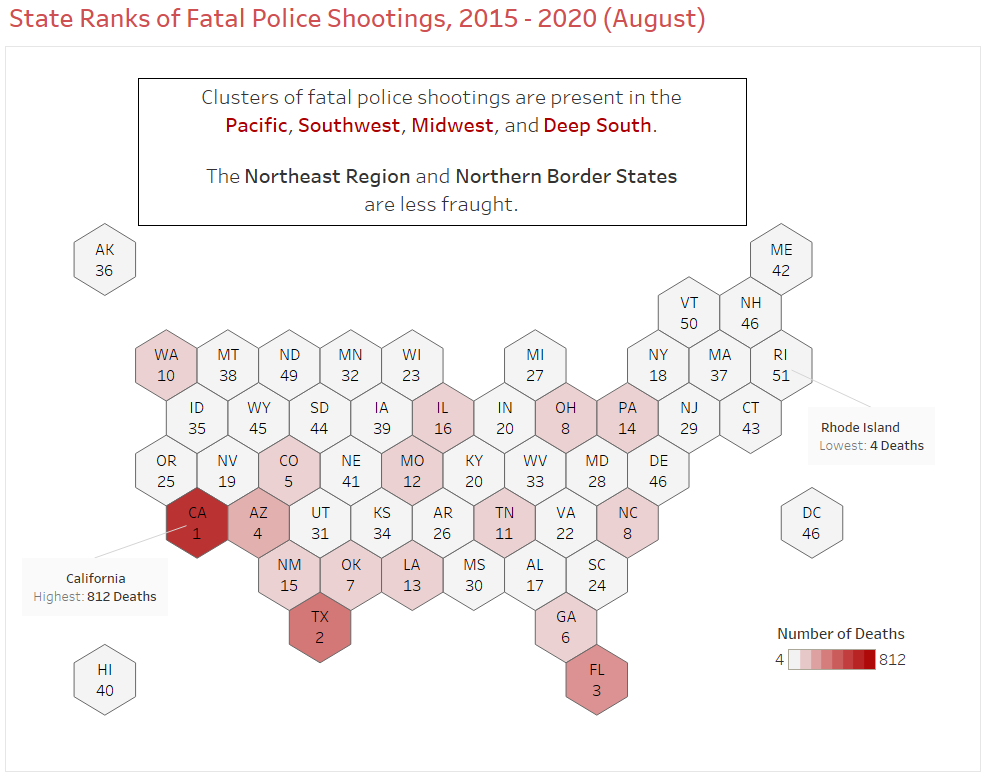


This dashboard drills-down into racial trends in the data. I placed the most insightful chart at the top, which reflects a disturbing finding for black individuals between the ages of 15 and 25. Without looking at any other data surrounding these killings, it would seem as though this demographic is disproportionately more at-risk. A colored line chart represents the age series for each race. I have the legend floating in the chart’s white space, to help the reader investigate further. The key findings, though, are represented in the title and boxed annotations for the target demographic.

On the bottom-left is a sorted bar chart with color-coding by race. This distribution is not particularly insightful on its own, considering the make-up of the United States overall. However, I included this with a simple title to show the reader an example of ordered totals (and not just percentages). This chart also serves as an additional legend for the reader to orient themselves to subsequent dashboards.

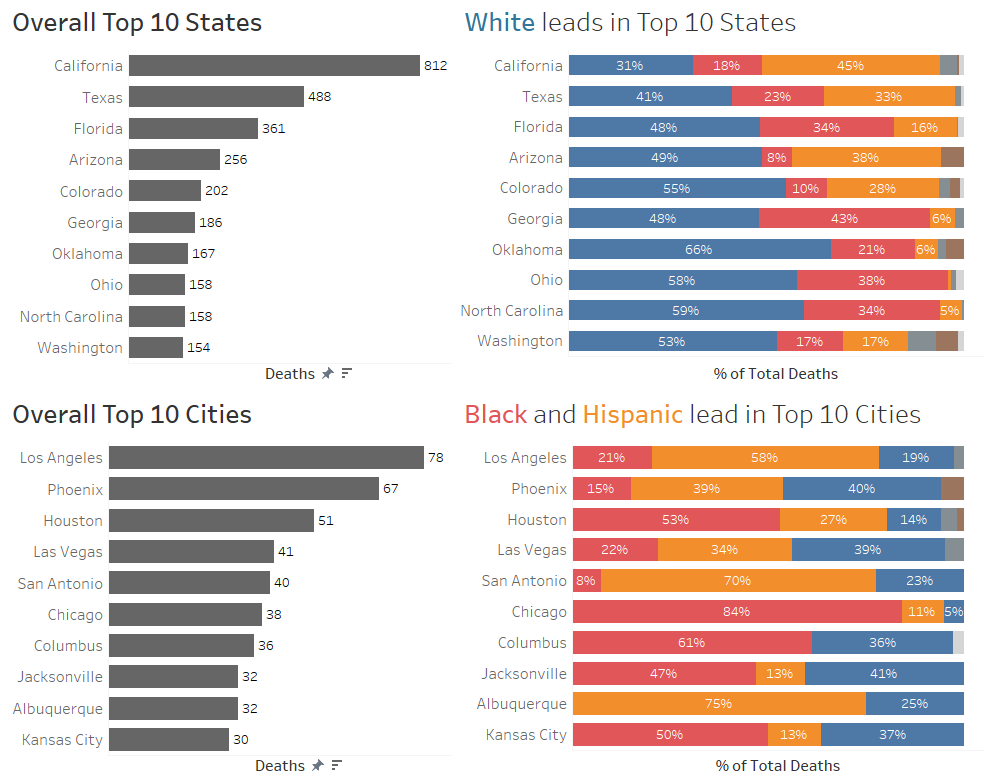
On the bottom right, this line chart with markers shows relative overall consistency among the races across the years. This is meant to contrast with the line chart above, since the reader might expect the ages or overall racial trends to mirror each other.

**Dashboard 3: State Rankings**



I chose to use a hex map to represent state rankings in this chart. While the title provides some idea of what the map represents, I’ve also labeled the state ranks here, along with a crimson color scale. The emphasis aligns with the total death count in the top states. I’ve annotated the highest and lowest ranked states (California and Rhode Island, respectively). I also included a boxed annotation to illustrate overall findings regarding the regional distribution of fatal force. Without any other data in this visualization, it is difficult to determine what is driving the state rankings. Some of that may have to deal with population density and socioeconomic status, neither of which is represented in this data set.

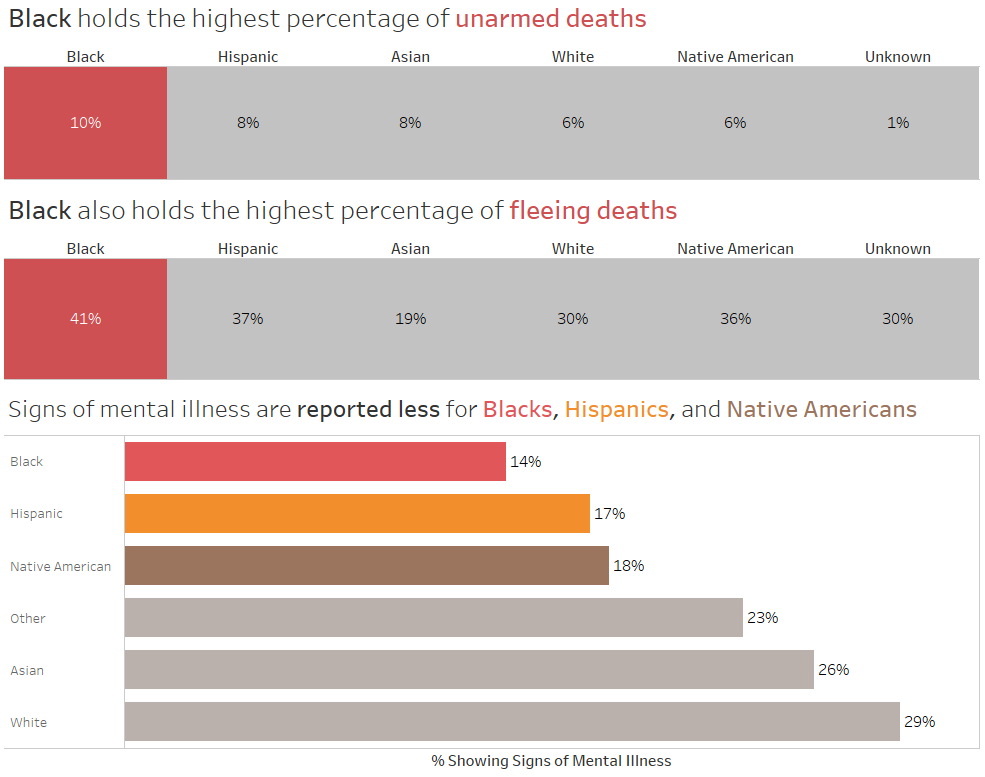
**Dashboard 4: State and City Top 10 with Race Distribution**



To delve more deeply into geographic trends, I’ve listed the top ten states and cities in this dashboard. The top ten bar charts on the left are sorted by death count. I’ve removed the axis ticks and instead marked the labels directly. I stuck with a dark gray color for the bars, as to not overload the dashboard with color; the sorting and labels should speak for themselves.

More interesting insights lie in the race distribution on the right side of this dashboard. I used the title as an opportunity to explain the key findings in each case. The stacked bar charts show the percentage of total deaths in each location, delineated by race. I sorted these charts by the races I meant to feature. I’ve also only marked the labels for the top three races represented.

**Dashboard 5: Circumstances of Killings – Unarmed, Fleeing, and Mental Illness by Race**



In this dashboard, I focused on characteristics under which it would be difficult to justify the use of fatal force: the individual *being* *unarmed, fleeing*, or *showing signs of mental illness*.

While dealing with relatively small percentages, I found it striking that the black population in this data set was the only race with an unarmed percentage in the double digits. I emphasized this with a red color coding in the title and a two-tone heat map that highlights the black percentage. I eliminated the row label, since the chart title fills that role.

For fleeing deaths, I used a similar strategy. While the unarmed deaths chart was sorted in decreasing order, I sorted the fleeing deaths chart manually so that the races are vertically aligned between both charts.

In the mental illness chart, I chose to feature the bottom three races showing signs of mental illness. This characteristic seems like it would be easily misread, ignored, or misunderstood in a fatal force scenario. The implications of this could be providing cultural or sensitivity training to a police force, if it is found that the disparity is otherwise unwarranted.

**Excluded Charts**

The charts that made their way into the final analysis all had a number of iterations. I did not include those step-by-step improvements here. I also used a handful of visualizations for exploratory analysis, in order to determine which variables and relationships to feature. Some of these include the following:

Gender Count: I found that it was somewhat interesting to have such a wide disparity between male and female. However, I was hesitant to use such a basic analysis in the final report. I instead chose to focus on variables with more than two categories.

Average Age by Race: This was also interesting to a certain extent. However, instead of focusing on averages, I chose to examine distributions across the span of age, having race occupy a percent of the total observations.

Body Camera Usage: I was not able to discern any meaningful patterns or disparities to feature with visualization.

Mental Illness by State: I briefly considered featuring regional trends in mental illness signs. While there may be something there, I felt as though there were too many factors for providing a concise set of key findings.

Threat Level by Race: I debated including this chart in my final analysis. While there are differences between the threat levels across the races, they did not appear to be as stark as with other characteristics.