

3300 Project 2 Writeup
Anna Shats, Yvonne Chan, Nathan Bala

The data sets used for this project were taken from a couple of places. The main data set was taken from the Washington Post, containing a list of civilians who have died due to the United State police since 2015. This information included data such as each individual's ethnicity, if they were armed/what they were armed with, as well as if they were fleeing/by what manner they were fleeing from the police. Also included in the data set was the city and state where the murder took place. Our second data set was taken from Kaggle and contained information on United State counties such as average income levels, racial demographics, and poverty percentages. The dataset that contained the 2015-2019 shootings had the cities where the shootings happened, but unfortunately did not include the latitude and longitude data. To solve this problem, another dataset with all U.S. cities and their respective latitude and longitude data was found at [this website](#). Then, a python script was written using Pandas in order to parse and match all "city, state" combinations in both the files and then add the latitude and longitude of the specific cities to the U.S. Shootings data set. We also used the Python library Geopy in order to reverse address lookup the longitude and latitude of each individual to get the corresponding county where their location occurred, and used another python script in order to add more information for each individual in the main data set. Another way the shooting dataset had to be handled was to split up the date column into year, month, and day to properly work with the slider that maps the datapoints based on year. The longitude and latitude data was also used to accurately plot the locations of each individual's death on the map in the second visualization. We used the US.json to create the map, which details the shape and arcs of each state.

For this project, the overall rationale behind our design was to display a stark and very matter of fact tone. For the first visualization, more than 3,400 small rectangles were created, each representing one of the individuals who died since 2015 due to actions by the American police. While making larger rectangles that represented multiple people at a time may have led to a more clean/less cluttered view, it would take away from the impact of having a representation that clearly illustrated the massive number of those who had been killed. Users are allowed to apply multiple filters concerning a multitude of categories at a time to view the wide range of situations people found themselves in during their deaths. When a filter is applied, the corresponding amount of rectangles moves down the screen, crosses a line, and turns to a bright red color. This transition was done to create a definite divide between the amount of people who fit a given filter versus those who did not. This allows users to easily see trends in the data and how to relate the size over certain subpopulations of the overall data.

Another design idea we were initially considering for the circle locations on the map was to have the various years accumulate on top of each other with the sliding of the slider, so as to show the severity of the past few years all together. The reason we decided against that and to

show each separately, is to more clearly distinguish the severity of shootings in specific places through an assigned opacity value for the circles of .5. So that when multiple shootings happened in the same city in that year, it is very apparent on the graph (check out LA!). If we were to do this for all the data (2015-2019) it would have not had the same effect. On the other hand, we noticed an issue with hovering for victim information. Due to the large concentration of victims in single cities, it became difficult to view the demographics of overlapping circles. This was a tradeoff with showcasing the variance in opacity of the circles, so we could not display the background information of all victims located in the exact same city.

The story we wanted to tell with our visualizations was the long-term impact police killings have on the population in the U.S. By showing data from 2015 to 2019, audiences can see the kinds of people who died at the hands of police. By allowing users to filter through characteristics like ethnicity, armed status, and majority demographic at the death's location, viewers can see the percentage of people who fit the filter compared to those who do not. The first visualization offered many surprising observations about the data in an interesting way. For instance, one can see that the majority of African Americans who died were located in a white majority neighborhood. Another interesting observation is that the percentage of white people who died is almost twice the amount of African American people who died. While the first visualization told the story of the quantity of different types of victims in police killings, the second focused on mapping the location of deaths in over the entire country. The map allows users to view where the killings are most concentrated and dispersed in each year from 2015 to 2019, as well as see the basic background information of each of the killings when hovered over. We felt it was important to view the background information of victims because it gives viewers a better understanding of the kind of people killed in what areas, in addition to adding a more humane aspect to the visualization.

Two interesting findings were revealed when looking at the locations of the shootings on the map. First, the places with the biggest clusters of shootings stay pretty consistent across the years (you can tell LA has a lot!). Second, the amount of shootings per year vary only slightly per year such that from 2015 to 2018, the total amount of individuals shot by police U.S. is staying pretty consistent. This is very interesting considering the attention given recently to police violence and large cases in the past few years on police violence especially among the African American population. Despite the emphasis on police violence in recent years, our data does not display any drastic changes in total shootings per year and leads us to assume there has not been a lot of changes in regards to policy on these issues.

Team Contributions:

Nathan Bala:

- Created python scripts to reverse search longitude/latitude to gain access to specific addresses, as well as scripts to combine the data set concerning the police killings and

the data set concerning United States Counties and their respective demographics and economic situations.

- Created the First Visualization (The filtering population visualization), spending the most time working on the Javascript logic to filter and display correctly the rectangles as well as basic styling.

Anna Shats:

- Created python script for finding latitude and longitude of the U.S. cities in the shooting dataset as well as adding the latitudes and longitudes into to the original shooting dataset.
- Created the Second Visualization (map) including the slider and zooming functionality. Also worked on some of the styling of the map, the slider, and the legend related to the second visualization.

Yvonne Chan:

- Created alternative visualization for the filtering population one, which had different styling and formatting for filtering, but ultimately did not end up being used
- Added styling and dividers to make main page cleaner
- Added hovering to show states on the visualization second visualization (map) and also showed details of specific victims when hovering over specific circles