Motivation

I chose the *Police Shootings in the US* as it is a highly discussed topic, and it is one of the main causes of many contemporary events in the US (and around the world). I wanted to see how these shootings are distributed over the USA, i.e., whether there are some clear differences between states in the North and South, or between rural and urban states. Moreover, I wanted to know in which states is the situation worst. I was also interested in whether the number of shootings gets lower over the years or not. From the point of view of different attributes, I was interested whether police do use body cameras, or what is the prevalent race of people that are shot.

Data sources and preprocessing

I have used several sources regarding different kinds of data.

- The main dataset regarding the shootings was from Kaggle https://www.kaggle.com/datasets/ramjasmaurya/us-police-shootings-from-20152022
- I have also used SVG and GEOJSON data to plot the maps (whole US containing states, individual states containing cities)
  https://simplemaps.com/resources/svg-us
  https://github.com/ResidentMario/geoplot-data/blob/master/usa-cities.geojson
- I have then used official data regarding the most recent US population census
  https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html

I have done for example following steps during preprocessing:

- I have removed the parts of map & population datasets regarding locations that are not states of the USA (Puerto Rico, various US administrative parts)
- I have pre-processed attributes of the shooting dataset that are to be visualised
  - I have explicitly marked missing values by “N/A”
  - I have merged few rare values of some attributes into new value “Other” (i.e., regarding attribute Race)
  - I have aggregated values of attribute “Age” into intervals, each by 20 years

Design choices

- I chose a map as a main visualization, as I was mostly interested in a distribution over individual states or cities.
- As there are vast differences between state populations, I also include a relative “per 1 million population” view.
- Moreover, as there are lots of cities and towns where shooting happen, I have chosen to include them only in a separate zoomed view. As there was still a lot of cities, I added their names and statistics via mouseover.
- Then I have selected a variant of the pie chart to visualize the distribution of important attributes, where user can choose which attribute to display.
- I decided to include 2 mechanisms for state selection, as sometimes it is more convenient to select something interesting on the map, and sometimes the user may want to select state by its name.
- I added the option to filter by time to observe some changes or just view specific time periods.
- I added a list of names of shot people to signify that these data are not just some numbers and statistics.

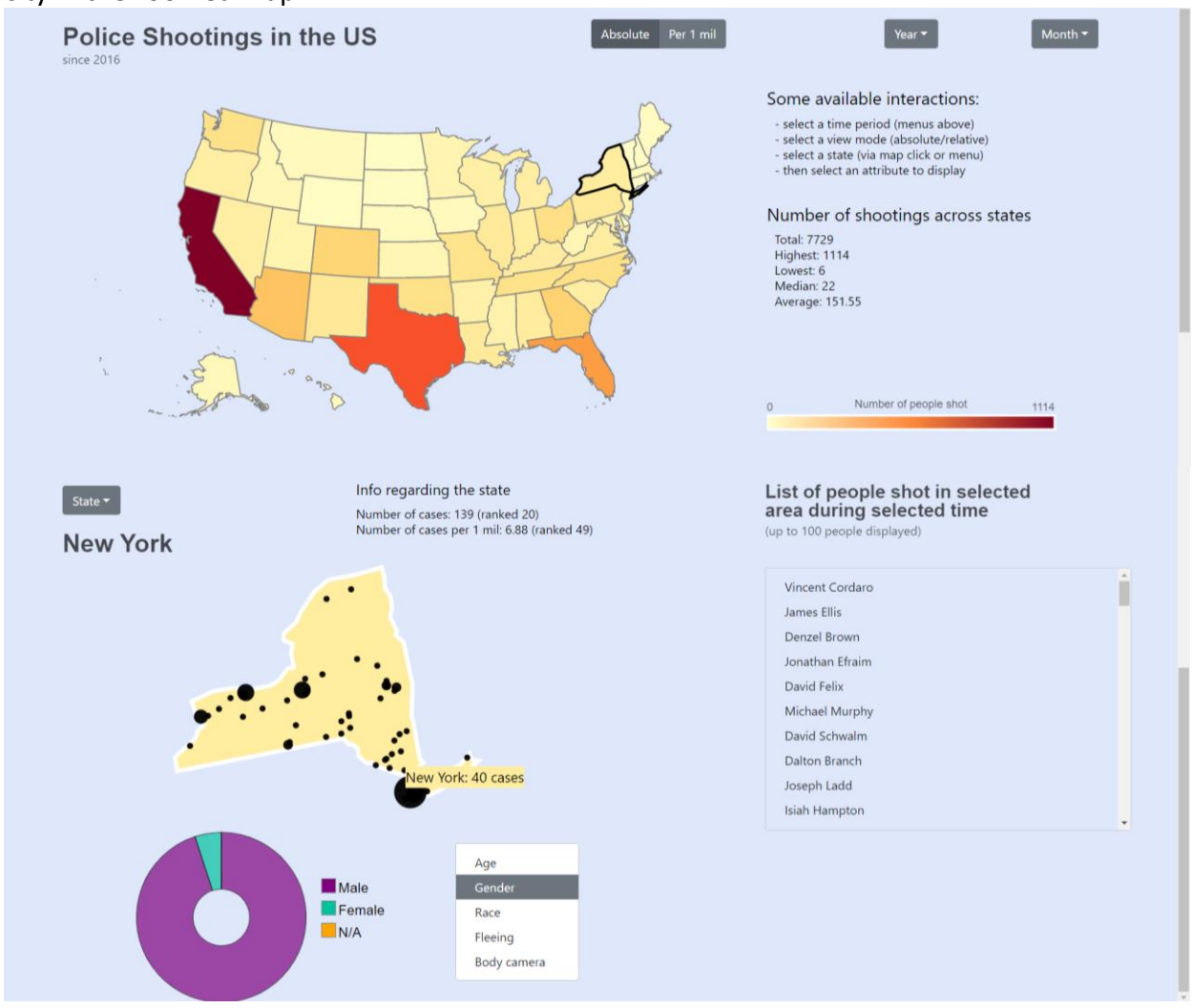
Observations

Some of the things I have observed:

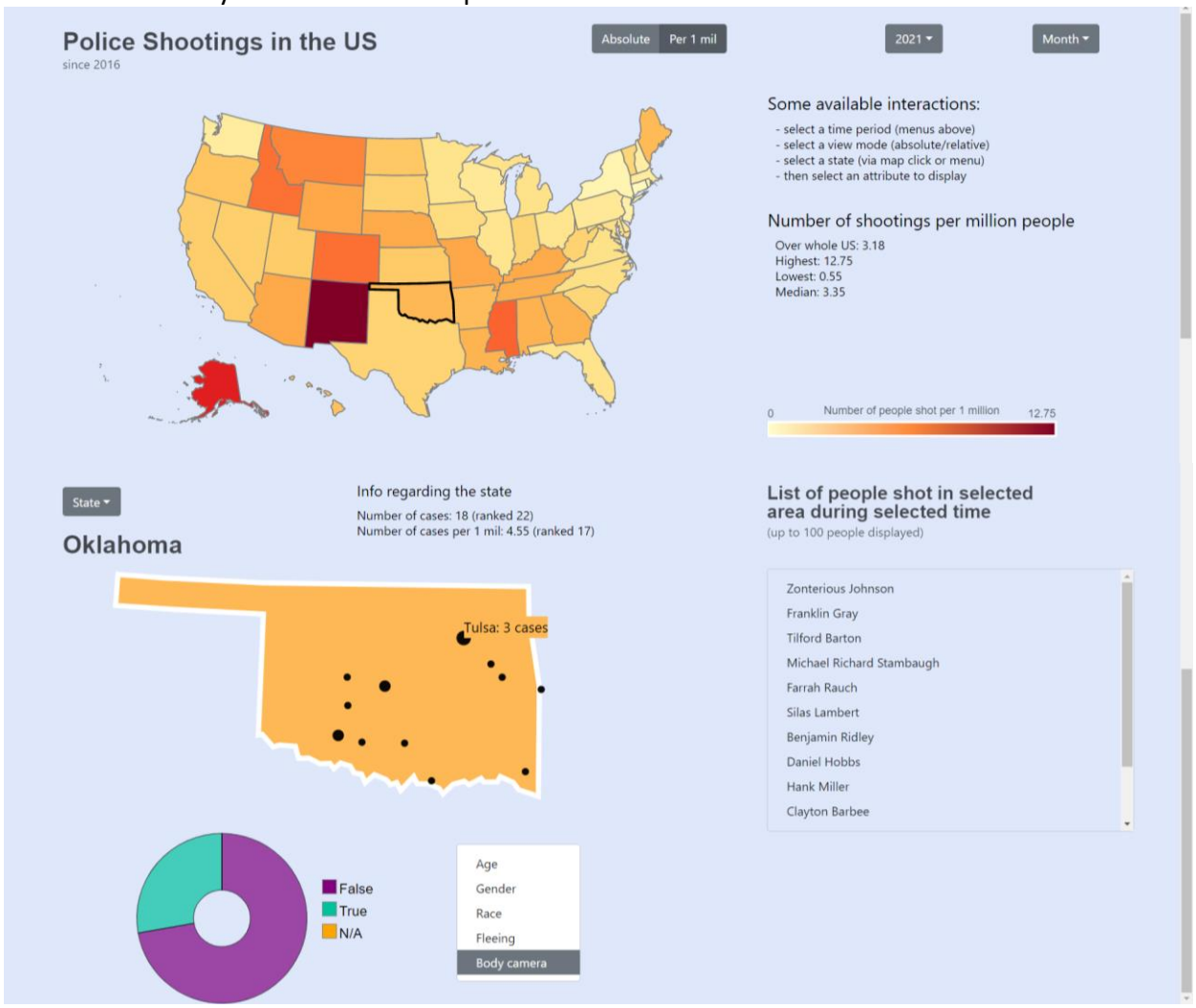
- There are lots of shootings.
- Most of the shootings happen in California, the base of Democrats. This is partially because of its large population, but it is still interesting to me. Los Angeles alone has more cases than 28 of the states (individually).
- States near Mexico border usually have larger number of shootings per capita.
- New York has relatively small number of shootings (both absolute and per capita) – I expected more. On the other hand, Alaska has a quite high number of cases per capita.
- It seems like the situation in North / South is not that different.
- The overall number of shootings goes up every year between 2016-2021 (slightly, but still). There are slightly less shootings in Autumn months.
- Body cameras are not used in most cases.
- Absolute majority of killed people are men.
- In south-west states (such as Cali), often most of the shot people are of Hispanic origin. Overall, throughout the US, the minorities seem to be over-represented, but most of the killed people are white.

Screenshots

Screenshot 1 – Numbers of shootings over the whole period (2016-now), with selected state being NY and displayed attribute Gender. Moreover, the mouse is hovering over the New York city in the zoomed map.



Screenshot 2 – Numbers of shootings per 1 million people over the year 2021, with selected state being Oklahoma and displayed attribute Body camera. Moreover, the mouse is hovering over the Tulsa city in the zoomed map.



Technologies

I have built the visualization as a Javascript application, so I have used combination of HTML, CSS, and JS. For the manipulation with the objects, I have used a bit of Bootstrap and JQuery. For the main visualization, I have used the D3.js library.

Lessons learned + takeaway

As this was my first JS application (and first encounter with JS in general), I have learned several things in this regard. However, it was very painful – I have spent most of the time on partially unrelated tasks (to the visualization task itself).

I believe I learned a bit about the D3, and what it can offer. I have for sure learned that D3 has several incompatible versions, and every time I would search for something, the solutions use different version than I need.

I have also learned how hard it is to propagate some user actions in the app so that everything is updated correctly.

From the visualisation point of view, I now have some experience with coordinate transformations in map views, or some knowledge about what possible colour palettes exist and that it is hard to choose something suitable.