INFO 5100 – P2
Ryan Enderby (rae83)
Di Jiang (dj327)
Xing Wei (xw395)

Hate Crimes Across the United States

The most dangerous cities in the US for hate crimes

Data

Sources

Our data comes from two sources: the FBI and the US Census Bureau. The bulk of our data came from the 2013 FBI report on hate crimes. A downloadable CSV file was found on GitHub[1], while the information that generated this file came from the official FBI release.[2] We supplemented this data set with population data for the cities in our map so that we would be able to scale data by population to show per capita statistics. Some of this population data came from the FBI as well, although their population dataset was incomplete, and did not contain all of the cities on our map. As such, we used US Census Bureau data on population by city in 2013 to complete our dataset.[3]

Since the FBI dataset reported hate crimes by city, we needed a way to place city names on a physical map. As such, we used another dataset that contained latitude and longitude coordinates by city.

Variables

Following the addition of population data, our dataset contained the following variables:

- State, Agency Name
  - State and Agency Name represent state and city, respectively.
- Race, Religion, Sexual Orientation, Disability, Gender, Ethnicity
  - The numbers of hate crimes reported for each of these bias motivations.
- 1st, 2nd, 3rd, 4th quarter
  - Total number of reported hate crimes in the given city for each quarter / season. These numbers are not broken down by hate crime type.

---

[1] https://github.com/emorisse/FBI-Hate-Crime-Statistics/tree/master/2013
[2] https://www.fbi.gov/about-us/cjis/ucr/hate-crime/2013
[3] http://www.census.gov/popest/data/index.html

- Population
- State Code
    - State Code is just the two-letter abbreviation for each state, and we used this in text areas presenting cities in the format of "city, state." Given that cities in different states sometimes have the same name (e.g. the abundance of Springfields and Rochesters), we felt it necessary to always display cities with their state as well. We use the us-states-names.tsv file to map from state in the FBI database to the two-letter state code.

Reformatting

A number of edits and reformatting changes had to be made for our purposes. We had to manually edit several Agency Names to match the city names on our latitude-longitude plotting dataset, allowing us to map each city to the JSON map.

Originally, the dataset was over 1,800 lines. However, most of these cities had two or fewer crimes reported for the whole year. As such, we chose to look at only the worst offending cities. For each different hate crime motivation, we chose the 25 cities with the highest number of crimes reported in the year. We selected top 25 as the boundary after looking through the data and seeing where a noticeable drop-off of reported hate crimes occurred. Some cities were highest offenders in more than one hate crime motivation, so we removed duplicate cities from the dataset.

Integrating datasets

To integrate these different datasets, we combined the population data into the same file as the FBI hate crime statistics and city names. However, when it came time to plot cities on the map, we used name matching to coordinate a city to the latitude-longitude values specified in the separate CSV file. We used the same US map JSON file that was presented in class (us.json).

Mapping

Overview donut chart

For the interactive donut chart, hate crime motivation is mapped to color, such that each different color represents a different hate crime motivation. Like most pie and donut charts, the area of each slice represents the relative proportion of that crime out of all hate crimes for a given city. This visualization defaults to showing the entire US, but by selecting a city from the dropdown menu, you can see the hate crime statistics for a given city. Hovering over a slice of the chart reveals a tooltip that shows the motivation, number of cases, and the percentage that this number of cases represents out of all hate crimes in the given city.

Map

The map shows you for a given hate crime motivation, which cities have the highest incidence of crime. Buttons along the top of the map allow you to choose motivation, and once a motivation is selected, circles appear on the cities that have the highest numbers of reported hate crimes. The color gradient of each city is scaled according to the number of crimes reported (e.g. dark red has more crimes than lighter red, etc.).

A toggle switch below the row of buttons allows you to scale the reported information to instead show the cities that have the highest number of reported crimes *per capita*. This way, you can toggle on and off seeing cities with the highest overall number of reported crimes, and the highest number of crimes per capita.

Hovering over a dot opens a tooltip with more information about the city. Information presented includes city names, the number of crimes reported for the selected motivation, the population of the city, where the city ranks in the US in overall incidence of crime for the given motivation, and a pie chart showing the proportional relations of hate crime motivations in that city. Much like the overview pie chart, this maps motivation to color and relative proportion of crimes to area. Due to the size of the pie chart, any slice that represents less than 6% of cases does not have a value label directly over it.

Seasonal chart

The final visualization section illustrates seasonal variation in hate crime incidence. A dropdown menu along the top allows you to either see the whole US or to select individual cities. Buttons next to the menu allow the user to select an individual season to see how it compares to other seasons in terms of number of crimes.

This data is visualized by human figures that become colored to represent the number of hate crimes that happen in that season relative to the most dangerous season in the chosen area. Color of the figures corresponds to season, and the number illuminated is proportional to the number of crimes in that season.

A line chart adjacent to the figures shows the trend in crimes over the course of the year. When a specific city is chosen, the corresponding line on the chart becomes illuminated. Since most of the cities have very few hate crimes per season, and a few have considerably more, a power scale is used to spread the lines out such that they can actually be distinguished.

Story

Through this project, we sought look at the landscape of hate crimes in the US and see the most dangerous cities in America in terms of incidence of hate crimes. It became clear that most hate crimes in the US are still racially motivated. More so, hate crimes tend to be concentrated in a few cities: they are not uniformly distributed across the US. For example, Columbus, OH has one of the highest rates of hate crimes motivated by race, ethnicity, sexual orientation and disability. More so, most of the crimes here happen in only a few seasons, with 44.1% occurring in the fall and 33.8% occurring in the winter.

Through our visualizations, we are able to identify New York in the winter, Columbus in the fall, and Boston and Los Angeles in the summer as the places to avoid if you are concerned about hate crime incidents.

One interesting finding was that for the most part, the cities that had the highest number of hate crimes were also the same cities that had the highest number per capita. As such, it wasn't just that some major metropolitan areas had a greater number of hate crimes just because they had a greater population; it was that these cities also had a disproportionately high number of hate crimes compared to population.