Jay Grollman (jfg93)
Elizabeth Zelko (efz4)
Jacques Diec (jpd276)

24 April 2015
CS 3300

Project 2 Report

Description of data

Our project has three main components, each which represents an angle of the overarching "story" we are trying to tell. The topic we chose is incarceration. The first part of our visualization is a map which shows the increase in state prison populations over time and the rising national corrections expenditures. In addition to showing the increase in incarcerated persons, we also decided to introduce the variables of race and offense type. In the second part of our visualization, we show the time spent in prison for various offenses over the last two decades, and additionally, highlight the differences between time served by white versus black citizens. For the third part of our visualization, we show the overall demographics of state prison populations in comparison with the demographics of the total U.S. population, for which we used census data.

U.S. Map

Data

We used statistical tools from the Bureau of Justice Statistics to get the prison population data for each state from 1982-2008. We essentially just took this data, edited out Alaska and Hawaii, and for any missing entries, used reasonable intermediate values based on the trends of the other states during the same time and the trend over time for that state. For our data on the national corrections expenditure, we actually got the data by contacting John Schmitt, a senior economist from the Center for Economic and Policy Research, who had done analyses to find the total inflation adjusted costs.

Mapping & Visualization

Through animations using javascript plugins, we used darkening colors to show the increase in state prison populations over time. We accompanied this animation with an animation for the rising corrections expenditures, corresponding to the year shown on the map. We found the extent of this increase pretty incredible, and also the distribution of prisoners per state. Additionally, the figure for corrections expenditures increased by several times in the span of just over two and a half decades.

Bar Graph

Data

Once again, we obtained our data from the Bureau of Justice Statistics website, this time under the National Corrections Reporting Program. The section contains datasets regarding crime, arrest, and incarceration statistics from the previous two decades (1993-2009). The final spreadsheet used for visualization was compiled from 11 spreadsheets (.csv) detailing the sentence length of state prisoners by offense, admission type, sex, and race. Each spreadsheet corresponds to a particular year from 1993 to 2009 (inclusive). In the synthesization of the data, we extracted from each spreadsheet, the mean prison time served by white offenders and the mean prison time served by black offenders for each major offense type/category. Offense types are classified as follows:

1. **All crimes**
2. **Violent crimes**
   Murder, manslaughter, rape, homicide, etc.
3. **Drug crimes**
   Possession, trafficking, distribution, etc.
4. **Property crimes**
   Arson, theft, burglary, robbery, fraud, etc.
5. **Public order offenses**
   Driving while intoxicated, etc.
6. **Other crimes**
   Miscellaneous

Our motivation for creating the bar graph was the measure the effect of race on average time served in prison for different crime types. The variable we are interested in measuring is the mean prison time (months) for black and white offenders. Our control variables are time (years), crime type, and race. As time increases, or as we move the timeline closer to the present, the number of offenders is expected to increase as a result of an increase in the U.S. population. This may explain the overall increase in mean prison time served by both groups (for all offense types) due to an increase in offenders. Other lurking variables may entail the changes in laws, policies, and jurisdictions over time. What we are really interested in measuring is the growth in disparity of prison time for black and white offenders over the last two decades.

Mapping

The height of the bars in the bar graph were scaled linearly by the length of prison time (months) which is a continuous and quantitative variable. We scaled the x-axis with an ordinal scale because crime type is a categorical and discrete variable.

Visualization & Discussion

Visualization of the data suggests a moderate positive correlation between time and the disparity in prison time between black and white offenders. This disparity seems to steadily increase over time, with prison time for black offenders increasing slightly more. Given the status quo, we predict that this gap will continue to widen in the next decade, unless certain measures are taken by policy makers in reforming education and law enforcement institutions over time. However, it is unreasonable to assume that significant improvements will occur in the next few decades, because certain (prejudiced) beliefs, ideologies, and institutions have been deeply ingrained in the fabric of our society for more more than 200 years. Other variables we did not consider were income and highest level of education between different ethnic groups. These variables may nevertheless influence the disparity in prison time between different groups, but our main focus was the effect of race on prison time so we decided to keep our story centralized by not introducing other variables that may detract from it.

Bubble Chart

Data

The data for this piece was by far the most complicated. All of our data was actually created manually from statistics. We found the overall percent of prisoners of each race, and for each race the percent of each gender, age and offense type. Using these percentages, we calculated the number out of 100 that would go in each category and used that to manually create a csv file containing an entry for each hypothetical prisoner. The general population was more complicated because we had to subtract out people under 18 from the data for each race and gender. We then used those numbers to calculate percentages over the total population over 18. Then we used these percentages, as with the prison population statistics, to create a csv file with the corresponding number of entries for each category.

Visualization

The variables we used for our visualization are race, age, gender, and offense type. We show the proportion black, white, hispanic, and other ethnicities/races that belong to each gender and that have been imprisoned for each type of offense, as well as the relative proportions in the general population. We found it especially interesting to see the discrepancies between the number of people in prison versus the percent of the general population for both black and white. It was also interesting and surprising to see how races were distributed over certain age groups, where white prisoners tended to be older and hispanic and black prisoners younger.

Sources

Data

- https://www.census.gov/popest/data/national/asrh/2013/index.html
- http://www.bjs.gov/content/dtdata.cfm#corrections
  (Section: National Corrections Reporting Program)

Map

- http://datamaps.github.io/
- http://eyeseast.github.io/visible-data/2013/08/27/responsive-legends-with-d3/
- http://www.cepr.net/documents/publications/incarceration-2010-06.pdf

Bar Chart

- http://bl.ocks.org/mbostock/3887051
- http://examples.oreilly.com/0636920026938/chapter_09/

Bubble Chart

- https://github.com/thiakx/Educity-ForceBubbleChart
- http://purecss.io/buttons/
- http://www.delimited.io/blog/2013/12/19/force-bubble-charts-in-d3