Final Project Writeup 6.859

Benton Wilson* Elizabeth Weeks†

Video: https://youtu.be/cyY8JSKUUFo
Website: https://6859-sp21.github.io/final-project-marijuana/src/index.html
Github: https://github.com/6859-sp21/final-project-marijuana

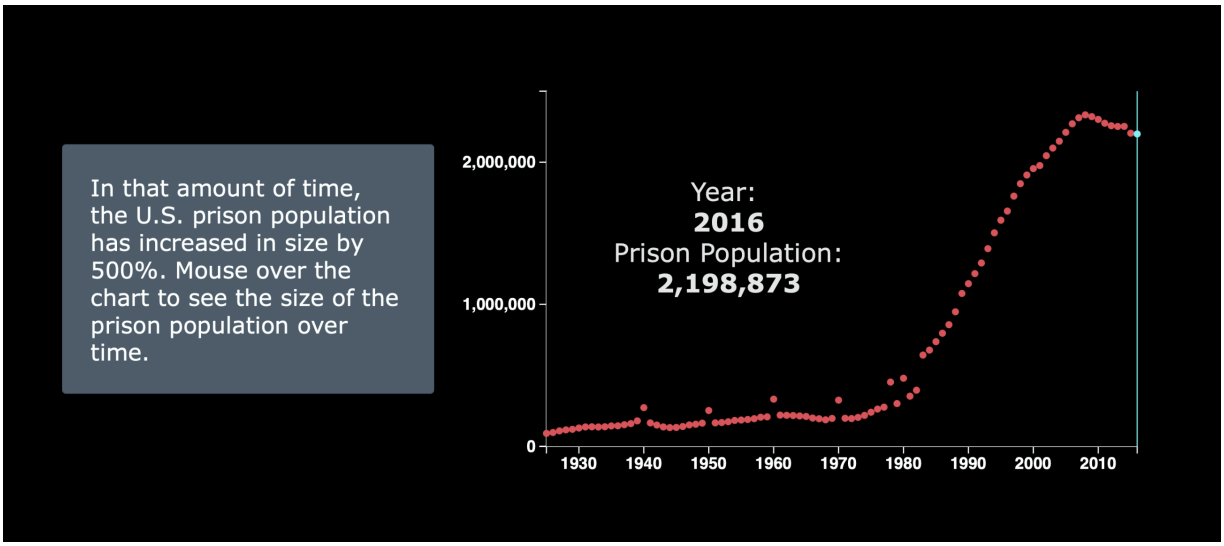


Figure 1: This is the first visual that users see when scrolling through the website. It is meant to help give context to later visuals.

Abstract

Mass incarceration is a large, but often hidden issue in the United States. The number of prisoners in the U.S. has increased drastically in the last few decades, and it is an important issue to understand so that we can start to solve the problem as we try to overcome systemic racism. Using the D3 library as well as techniques learned during 6.859: Iteractive Data Visualization, we constructed a scrollytelly website in order to visualize the makeup of the U.S. prison population, relying mostly on a dynamically updating matrix of humanoid icons. At the end of the narrative walk through, users are invited to both explore the data further as well as visit websites of various organizations trying to take action against mass incarceration.

1 Introduction

During the past year, many events unfolded which helped to bring systemic racism into the forefront of the public mind. While police brutality tends to make the biggest headlines, mass incarceration also plays a huge role in perpetuating biases in society.

Initially, we planned on examining the issue of mass incarceration by looking specifically at marijuana and imprisonment rates. This is important because there are lots of prisoners who we believe are wrongfully in prison for a drug that is in many cases, now legal in the state where they were originally convicted. We were curious about some of the mandatory minimums and how they might relate to mass incarceration. However, as we began to do some research, we realized that there was a lot that we didn't know about mass incarceration in general.

As a result, we decided to explore more generally, looking at information about mass incarceration in the United States as a whole, since there are many sides to the issue. Mass incarceration often ends up more hidden from public view (at least for us, it doesn't seem to be something that we see as often online).

In the end, we tried to do two things. First, answer the question of "What does the prison population look like?", by allowing the user to look at various breakdowns of the population, but also put that prison population in context in a humanizing way. As a result, we chose icons to represent most of our charts, and we have a scrollytelling style website that enables us to control how a user walks through the data (and then at the end a user can do their own exploration).

2 Related Work

Most of the research that we used involved techniques taken from 6.859: Interactive Data Visualization [3], as well as the library commonly known as d3.js [1]. We also pulled starter code for our icons from an example on observable. [4] We feel that these resources were unbelievably helpful in designing our website (we both have zero web development experience before this class).

As far as our data goes, the majority of the data was pulled from an extensive survey done by the Bureau of Justice Statistics. They surveyed around 24,000 prisoners in various state and federal facilities, and the end result was a dataset with around 2,000 columns [2]. Since there were so many fields, they were all coded, and we spent a good chunk of time at the start of our project simply

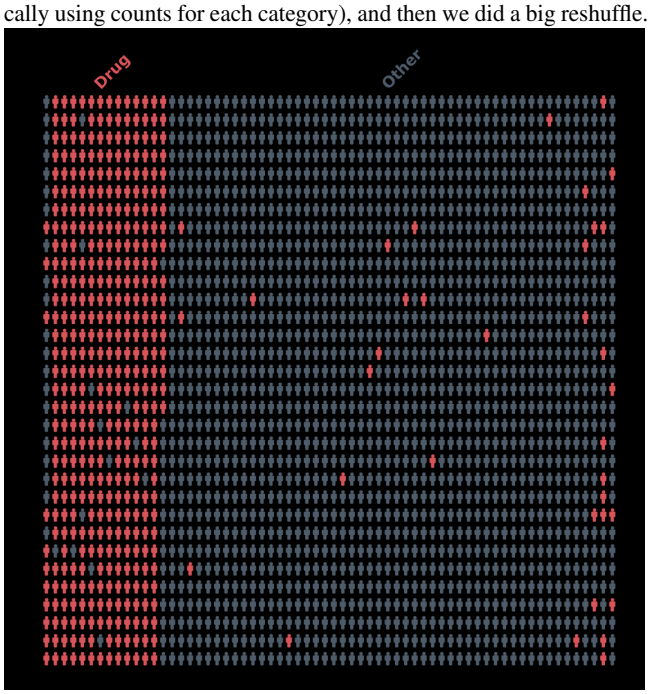
*e-mail: benton@mit.edu
†e-mail: eweeks@mit.edu

relabeling and filtering out certain columns with less complete data.

3 METHODS

The crux of our visualization revolved around the grid of icons that changed color and then moved, so we spent an extensive amount of time trying to improve and optimize how the people moved, since we wanted it to look like a crowd with a group of people lining up in their corresponding location.

Our initial implementation was pulled from an Observable notebook [4]. Since we wanted the people to color and then move (rather than just making certain chunks appear colored in groups), we attached the notion of a *graphID* with each person. This was a monotonically increasing value such that each icon had a unique spot in the graph. Then, we could update that icon's location accordingly based on their new *graphID*. Initially, we just scanned through all the data, updated the values of *graphID*s in the order that we saw them (basically using counts for each category), and then we did a big reshuffle.



This presented a few problems. The first problem is that oftentimes (especially when there are fewer categories), many of the icons moved when they did not need to. We wanted to reduce the amount of motion (which also helps with the amount of lag and processor overhead), by not actually moving any of the icons if they were already in the correct section (for instance, if an icon was colored blue, and after moving the icons to be sorted, that location should have a blue icon, then the original icon just stays in place). To do this we wanted to come up with an algorithm that did not make too many passes over the data.

We made one pass over the icons, and for each icon that is already in an assigned location, we put it's ID in a set of used IDs. Then, on the second pass, we monotonically assign the remaining icons a new *graphID*, from the set of *graphID*s that have not been used yet for each category. This technique greatly improved the frame rate in the browser, and also helped make our visualization have less noisy movement.

Another problem that we encountered was that we wanted the icons to move at a constant rate (like a crowd). However, d3 does not allow us to set rates for transitions, so we first calculate the distance that an icon needs to move, and then scale our transition duration accordingly (icons moving further have longer transition times).

Finally, we found that the frame rate of the visualization dramatically improved when we randomly delayed all of the transitions in the range of 0-1000ms. We believe that this occurs because the number of icons that start moving at once is reduced, which reduces the peak amount of work for the browser (there is less of a spike in computation).

4 RESULTS

In the end, we were happy with the results of our visualization. The following is a case study in how a user might interact with the visual (thanks to Lucy, our friend for being the test subject).

First, Lucy opened the web page and read the striking text at the top about increasing incarceration rates. Then, Lucy scrolled a bit further, and saw the chart (shown in figure 1), that presented the change in prison population over time.

After exploring that data a bit, Lucy continued down the page, and the grid of icons showed up, reflecting the percent of prisoners in the U.S. and then Louisiana. Lucy tried changing the state to her home state of Maryland, and then moved on, continuing to scroll down the page.

When she hit the panel about the Center for Prison Reform, she opened the link in a new tab and looked at a few of the sites, before returning to our visual. Finally, she was able to scroll to the bottom and explore the data a bit further, clicking some buttons before being done with the visual.

When we asked what she had learned, Lucy said that before the visual, she didn't realize just how many people were in prison in the United States. She found it interesting to see the prison population broken down by crime, and overall she liked that she was able to see how incarceration rates vary by state.

5 DISCUSSION

One thing that we found extremely helpful in evaluating our visualization was the midpoint milestone. We tried to make, as much as possible, the minimal version of our website clean in order to get feedback on how we might improve the user experience.

The following is a list of the main points of feedback that we were able to respond to in our final product:

- We added more useful information to our tooltip; originally, it only contained the same information as the label for a section of the people, but we altered it to also show the percentage for each category.
- We also got feedback that it would be nice if the tooltip color matched the color of the icons.
- We got some feedback that it was annoying to have to scroll slowly to the bottom of the page. This was partly due to an unresolved bug in the milestone, where the visualization would freeze if the user scrolled too quickly. For the final, we fixed the visualization so that new transitions no longer freeze previous ones.
- We also got feedback that having a "category selection" button and a "move" button was confusing for the user (originally, the user would color the icons according to a category, and then move them by clicking the move button). In response to feedback, we changed the behavior to simply color the icons, and after a small delay, automatically begin to move them.
- In our midway milestone visual, we highlighted the rate of incarceration in Louisiana, but we received feedback that users really wanted to be able to look at data for other states as well, so we added the state selector into our final.

Our intention was to enable the user to explore some of the data surrounding mass incarceration in the United States, and to contextualize it with a humanistic representation of the data. We hope that a user who sees the visual is able to gain an appreciation

for just how large of a problem mass incarceration is. Then, we point the user towards ways to get involved.

6 Future Work

While we are happy with the visual that we created, there are many ways that it could be extended or even further improved.

In our visual, we mostly focus on data from the United States, but it would be interesting to make a visual that includes data from other countries as well. For instance, our visual leaves certain questions unanswered which may help us understand why incarceration rates in the U.S. are so high. For instance, why do certain countries have extremely low incarceration rates? Is there anything that those countries do that the U.S. could do as well?

Another thing that we really wanted to do was to try and humanize the data as much as possible. This was one of the large motivations for using human like icons, and having them move like a crowd, but it would be even better if in the future, we had each icon link to a story about someone in prison. This way, we would allow the user to immediately see the human impact of imprisonment, which may be more powerful in creating empathy for prisoners. This would also potentially make the user more likely to get involved through some of the organizations that we link to.

Finally, it might be interesting to incorporate multivariate filters into our visualization. For instance, we could look at prisoners arrested for a drug related offense by race. Currently, since we are downscaling our data, we didn't have a clean way of representing this, but in the future this could be a really cool way to dig deeper into the data.

Acknowledgments

We wish to thank the 6.859 staff for their helpful critiques and teaching throughout the semester.

References

[1] M. Bostock, V. Ogievetsky, and J. Heer. D^3 data-driven documents. *IEEE transactions on visualization and computer graphics*, 17(12):2301–2309, 2011.

[2] U. S. B. of Justice Statistics, Mar 2021.

[3] A. Satyanarayan. 6.859 : Interactive data visualization.

[4] A. Smith. Pictogram grid in d3js.