

Incarceration in the US Visualization

Final Report

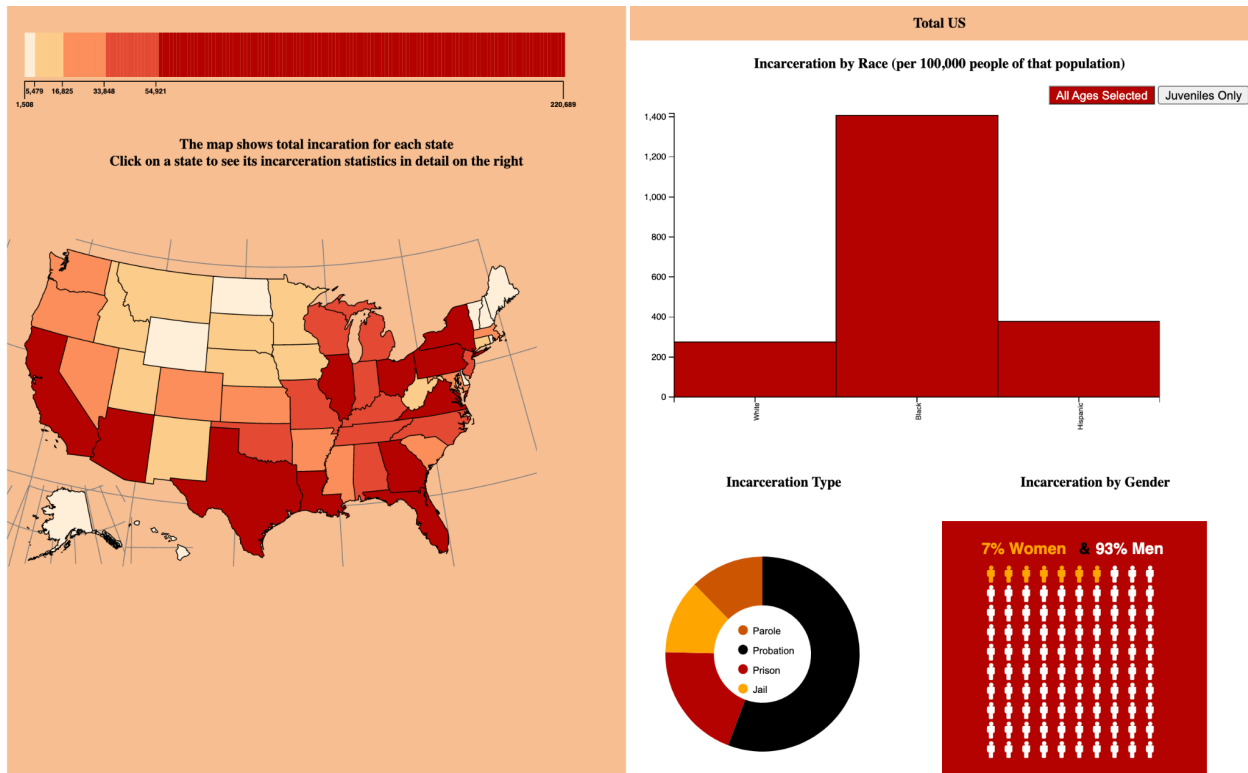
Dataset - <https://www.sentencingproject.org/criminal-justice-facts/>
- <https://bjs.ojp.gov/content/pub/pdf/p19.pdf>

Contents

- 1) Final Data Visualization
- 2) Information about our Data
- 3) Design Rationale
- 4) Data Visualization Description
- 5) Team Contributions

Final Data Visualization

This Data Visualization demonstrates the incarceration statistics in the US visualized by race, gender, and type. The map on the left gives the total number of incarcerated people in that state. Clicking on a state will give you a pie chart representing the amount of people incarcerated per 100,000 of that states population for each type of incarceration, there is also a bar chart representing the races with two view (juvenile and all ages view), and also a pictograph representing the gender breakdown of incarceration in that state.



Information About the Data

Data (1) - Bar Graph and Pictogram

Data Description

The data is a combination of state-level criminal justice data from a variety of sources. To achieve our goal of visualizing state imprisonment and custody by gender and race, we subtracted the information data from this website and pasted it in CSV format. Data citation can be found here: <https://www.sentencingproject.org/criminal-justice-facts/>

Data Rows

1. Total incarcerated, prison and jail
2. Men in prison
3. Women in prison
4. White Adult
5. Black Adult
6. Hispanic Adult
7. Juvenile custody rate
8. White
9. Black
10. Hispanic custody rate
11. American Indian
12. Asian

Data Columns

States names in the U.S. (US_Total, Alabama, Alaska...)

Data Cleaning

The main cleanings we did are

1. **Extract only necessary rows**
 - a. For Juvenile data = [White, Black, Hispanic, American India, Asian]
 - b. For all ages data = [White Adult, Black Adult, Hispanic Adult]
 - c. For pictogram = [Men in Prison, Women in Prison]
2. **Clean rows with empty column values**
3. **Replace the value of “-” or “~” with the value of 0**
4. **Converted pictogram data to percentages**

Data Limitations

- 1) 0 value for certain races and “-” or “~”.

_____ There are certain cases where we get 0 for that race even though it's most likely not the case. And there are also cases where we simply don't have the data. This mostly happens for Asian and American Indian Juveniles. It would take a deep dive to analyse where this data is coming from and how we can get the missing values. We decided to leave it as it is as

incarceration data can often be incomplete due to difficulties in gathering and we are only trying to show a general trend. This did not happen too often as well so we foresee not much change even with the complete data.

Data (2) - Pie Chart

Data Description

The data is a combination of state-level criminal justice data from a variety of sources. To achieve our goal of visualizing the type of incarceration by state we copied the data from the website: <https://nicic.gov/projects/state-statistics-information>

Data Columns

- 1) ID - The ID number of the state in my data
- 2) State - the state
- 3) Prison - people in prison per 100,000 of that population
- 4) Jail - people in jail per 100,000 of that population
- 5) Parole - people on parole per 100,000 of that population
- 6) Probation - people in probation per 100,000 of that population
- 7) Total - The total of the previous four columns to make the pie chart calculation easy

Data Cleaning

For numerical data in Prison, Jail, Parole, Probation, and Total columns they initially show up in string format. I convert them to Number.

I also filter by State when a particular state is selected giving me a single row with the above data columns for that state. I then create a dictionary of the form:

Data Limitations

For some states, the data appears to say there are 0 people in Jail. This seems incorrect but I assume it is because that state has a unified jail with another state. There was not much I could do in preprocessing to fix this problem so I have left it as it is.

Data (3) - Map

Data Description

The data is a combination of state-level criminal justice data from a variety of sources. To achieve our goal of visualizing state imprisonment and custody by gender and race, we subtracted the information data from this website and pasted it in CSV format. Data citation can be found here: <https://www.sentencingproject.org/criminal-justice-facts/>. Gathered all the data from this page to Incarceration_data.json. Then used us_copy.json for topological ordering.

Data Cleaning

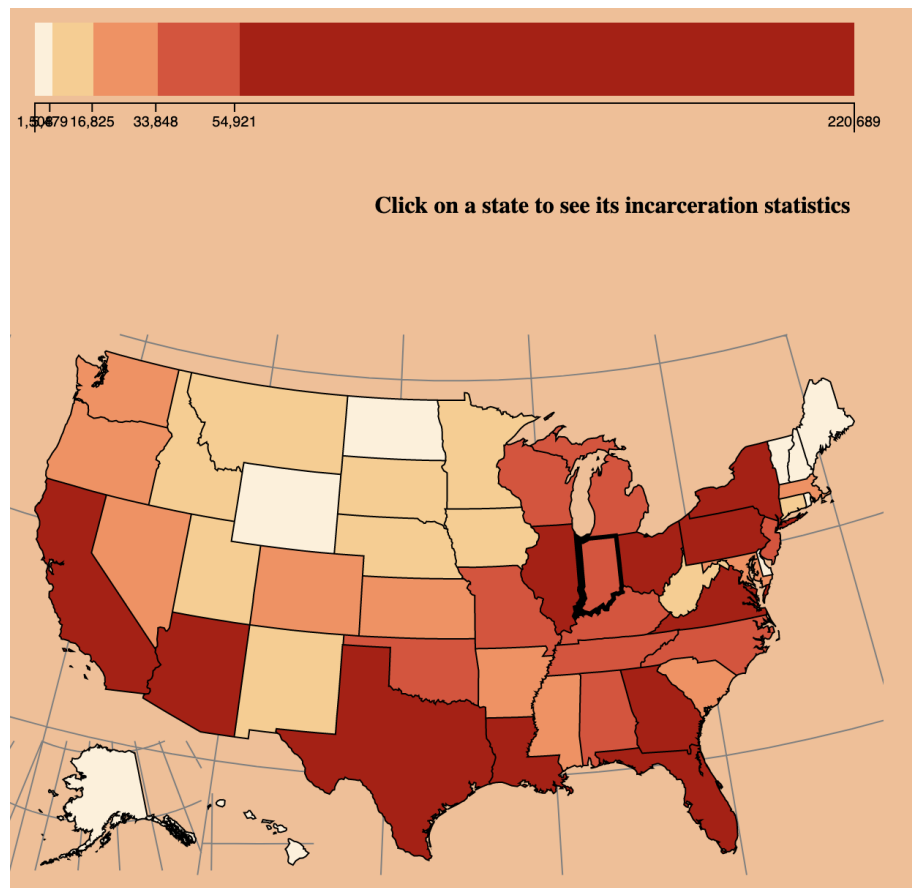
Not much cleaning was needed on this data. The main things are we pretty much only used the "Total_Incarcerated" field and additionally we create mappings from the state name to its ID and vice versa. We also create a mapping from the state ID to the total incarcerated value.

Data Limitations

All data needed was present. Would have maybe liked a per 100, 000 breakdown instead of the complete total.

Design Rationale

Visualization One - Map



Final design explanation and rationalization:

- 1) Used the d3.geoAlbersUsa projection to emphasize the continental US. The actual size and shape of the states do not matter as much as their shade so there was no need for a map such as the mercator.
- 2) Even though the size and shape of the states were not as important, we still used a graticule for the users' curiosity.
- 3) The five colors used for the categories are distinct in color so there are no clashing visual channels - #fef0d9 (dark red), #fdcc8a (salmon), #fc8d59 (light orange), #e34a33 (beige), #b30000 (white).

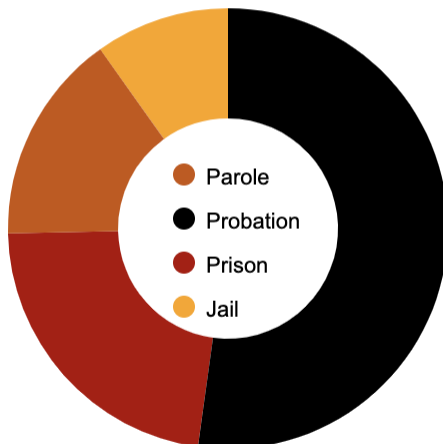
- 4) For the actual shading of the states, we used quintiles of different sizes because there was a huge disparity between the incarceration rates themselves, as seen in the legend (16,825 - 33,848 compared to 54,921 - 220,689).
- 5) Emphasized state lines so that we could clearly distinguish the states
- 6) On click state border becomes thicker to indicate state is selected. Makes it easy for user to tell which state is selected.
- 7) On hover we have tooltip that tells us what state it is and the actual total incarcerated data. Makes it easy for user to pick the state they want and also provides the info in a readily accessible manner.

Design Tradeoffs:

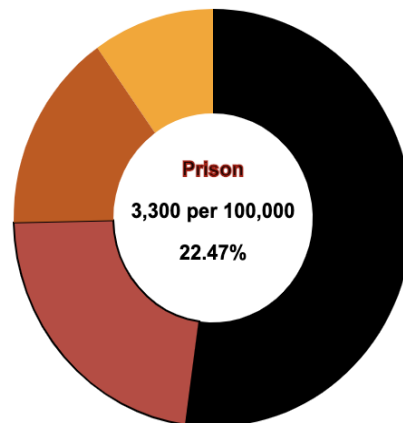
- 1) We used the d3.geoAlbersUsa projection instead of more geographically accurate projections because the shade of each state is what mattered most.
- 2) We don't change anything about the state as you hover over it except showing you the tooltip. This is because we already use a thicker border for select and we don't want to confuse users by using that for hover as well. The other thing we could have tried was opacity or fill change but both of these mess with color channel too much (tried opacity and it starts to make one color on the map look like a different color that is also on the map and changing fill will no longer make it correspond with the initial legend).

Visualization Two - Pie Chart

Incarceration Type



Incarceration Type



Final design explanation and rationalization:

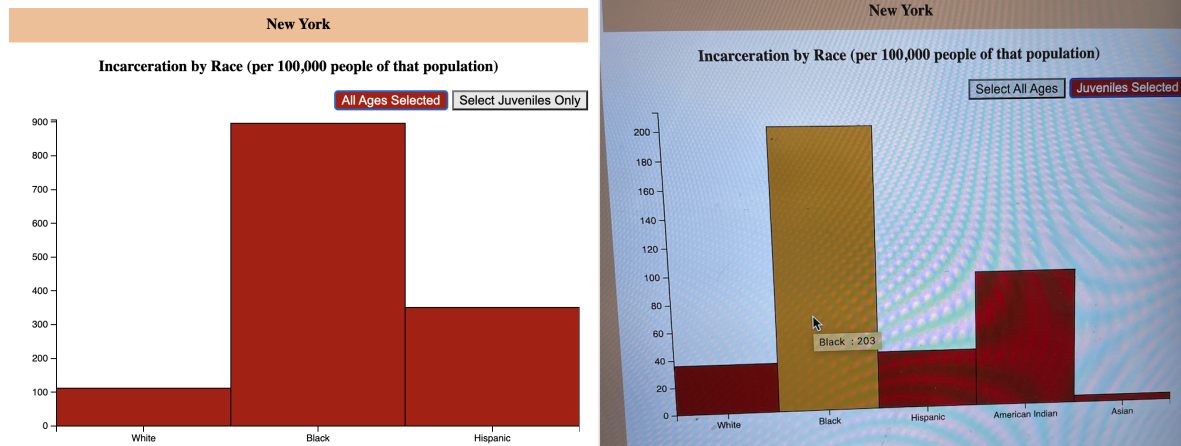
- 1) There are four incarceration types in my data and the pie chart has 4 categories - prison, jail, probation and parole - to represent this.
- 2) The four colors used for the categories are distinct in color so there are no clashing visual channels - black, red, orange, yellow (shades of)
- 3) I keep the legend inside the pie chart instead of outside so that space is conserved and it looks compact. It also invites the user to look at it immediately
- 4) The slices are sorted from largest to smallest going in a clockwise direction to help users distinguish size differences.
- 5) When the user clicks on a slice the percentage of that slice, the slice category name, and the actual number value associated with the category are displayed in the center. Doing this allows users to look at the information in detail without having to puzzle it out themselves. I choose to use click instead of hover here because I want users to be able to keep looking at a slice's info even as they explore other charts if they'd like. The slice category name is outlined in that category color so it is easy to associate the info with the slice.
- 6) Clicking on the same slice brings back the legend. We have this deselect action as clicking again to deselect is intuitive.
- 7) When a slice is selected the stroke becomes black and the opacity becomes 0.8. This makes it clear a slice is selected.

A few design tradeoff examples:

- 1) Version 1: I initially used the same color scheme as the map graph
Version 2: Changed to more distinct colors - both from graph and each other
Reason: The map graph uses an ordinal color scheme to display total incarceration data. The pie chart is categorical and displays the number of people incarcerated by type of incarceration. Using the exact same color scheme as the graph could cause scale and data confusion. Additionally, it's not a good idea to use an ordinal scale where a categorical scale is required. The final colors I use to create a categorical scale while also not disrupting the general color scheme of the graph.
- 2) Version 1: Legend Outside
Version 2: Legend Inside
Reason: Conserves space and immediately draws the user's attention
- 3) Version 1: Only 0.8 opacity for select
Version 2: 0.8 opacity and black stroke for select
Reason: In some cases, the 0.8 opacity might not be very obvious. I included the black stroke as it was more immediately noticeable.
- 4) Version 1: Legend always in the center
Version 2: Legend replaced by details when you click on a slice
Reason: While it might be a little inconvenient to not always have the legend, you can easily get it back by deselecting and there are only four colors to remember. I thought it

was more important to be able to see the details of a slice when you want to than to have the legend always in the center.

Visualization Three - Bar Chart



- 1) There was originally a dropdown menu to see the interaction of incarceration rate by races between states, including Hispanic, White, Black, Latino, American Indian and Asian. Each bar represents the incarceration rate per 100,000 people. When scrolling down the dropdown menu, we can choose among states.
- 2) If we hover over each bar that is presented and pause for one second, the color of the bar will change and it will give us a text box showing the incarceration rate in that race. I added the transition to the interactive function so that the bar change between each state becomes smoother.
- 3) We also have two views. In the first one you see the races for all ages (Black, White, Hispanic) and the second only for juveniles (Black, White, Hispanic, Asian, Native American). This allows you to take a deeper dive into the data. You can toggle between the two views using buttons that change background color, text color and text content to indicate they've been clicked. We made it explicit to make sure there is no confusion about which view you're on.
- 4) When the bars transition between modes the x axis rescales to make sure there are no awkward white spaces.
- 5) When the bars transition between modes and also between states the y axis rescales so that no bars overshoot the axis (we have allowance of 10 extra after the max value of that group to make sure no bar gets cut off)

Design tradeoffs:

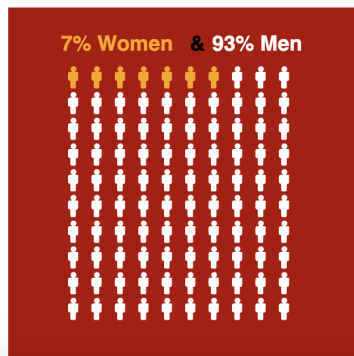
- 1) When combining the map and the bar chart together, we chose not to have the dropdown box on the top. Instead, we made an interaction between the map and the bar chart.
- 2) We considered just putting the max possible value across the data as the y axis limit and letting the y axis stay static instead of constantly rescaling it. A static y axis might let users compare more easily against different states and juveniles vs all ages. In the end we decided to make it rescale because some states really just had much more

incarcerated per 100,000 than others and sometimes there could be large differences between juvenile and all ages race breakdown numbers for the same state. Using the same scale would either cause some data to overshoot the graph bounds or minimize other data too much.

- 3) We change text after you select a mode so it's very clear you have selected that mode. We were worried that the red background might actually look like we are inviting a user to click instead of it already being selected.

Visualization Four - Pictogram

Incarceration by Gender



Final design explanation and rationalization:

- 1) We made a pictogram because it spreads out the info in a really clear manner and we made a 100 icons for data out of 100 per cent so that it is easy to read.
- 2) The pictogram shows orange people icons to represent women and white people icons to represent men.
- 3) There is an orange text label which is women incarcerated percentage and a white text label which is men incarcerated percentage. Having this makes it easy for the user to immediately pick up the information. The "&" that bridges the two text labels is black - a color we don't use elsewhere in this graph. This makes sure that the "&" appears neutral and not part of any other category.
- 4) There are 100 people icons and each represents 1%. I considered making a legend that explicitly states this but decided it is not necessary as this is a fairly intuitive conversion and there are text labels above that explicitly state the percentages.

A few design tradeoff examples:

- 1) Version 1: Background same white as other detail charts
Version 2: Background red
Reason: As the pictogram chart is a bunch of icons rather than a compact chart having the same background as every other chart close by made it seem like there was no separation. Giving it a red background makes it feel like a compact chart of its own and provides clear separation from the other charts

- 2) Version 1: Men comes first and Women icons are orange
Version 2: Women come first and Women icons are orange
Reason: Orange is a striking color and it draws attention compared to white. For this reason, I thought it was better to have the typically lower percentage group use orange so it wouldn't strain user's eyes, and would still maintain its novelty.
 - 3) Version 1: All text was black
Version 2: Text corresponds with category color
Reason: Linking the text color with category color allows users to more immediately make an association between the data the pictogram icons are and the data the pictogram text tells us.
-

The Story

There are several things that are clear to us when presented with these visualizations. The first trend we see is that the highest incarceration rates seem to cluster around both the southern part and the northeastern part of the United States. Some factors that can contribute to this trend is both the higher population density within these areas and the demography of the population. Another trend that we see is that the distribution of the type of incarceration for each state seems to have the same pattern throughout. Probation is the largest incarceration type, while parole is the lowest. Likewise, with each state, we can clearly see that men are more likely to get incarcerated, as they have a 9:1 ratio with women for incarceration rates. Lastly, when comparing the race/ethnicity incarceration rates, there is an apparent trend that historically marginalized groups such as Black, Hispanic/Latinos, and Native Americans have higher incarceration rates compared to that of other groups such as Whites and Asians across juveniles and all ages.

Team Contribution

Through this project, we successfully followed the timeline and the milestone instruction. For each step of the project, we assigned work to each individual, clarified their tasks, and met once or twice weekly to check each other's progress and discuss further changes/improvements to make. Active participation of all group members ensured a good pace for finishing our project.

Contribution outline (tasks individual done)

- Divya
 - Came up with the project idea
 - Research on data for further idea development
 - Collected and cleaned data 2
 - Made the Pie Chart & Pictograph (static and interactions)
 - Made the Map Graph interactions
 - Combined all the graphs together (unified color scheme, made layout, added graph titles and descriptions, etc.)

- Made the interactions between graphs
 - Made two different modes for bar graph
 - Helped write final report
- Lang
 - Designed the bar chart
 - Research on some project ideas
 - Collected and cleaned the data 1
 - Made the bar chart (static and interaction)
 - Helped organize the meeting and outlined the final report
- Brian
 - Made Map Graph
 - Made Map Graph Legend
 - Helped Develop the Final Report (the story)

Code references

<http://bl.ocks.org/jonahwilliams/2f16643b999ada7b1909>

<https://jeffrz.github.io/info3300-fa2021/notes/21.10.13.notes.htm>

<http://bl.ocks.org/henryjameslau/3299be02d548f999c8ae01699ab1ead2>

<https://jonsadka.com/blog/how-to-create-adaptive-pie-charts-with-transitions-in-d3>