**Homework 1**: Digital and Analog Visualization

**Objective:** Take a data table and use it to create a data visualization that communicates a message.

**Task 1**: **Examine the data.**

Distribution of Prison Sentences for Black and White Men Convicted for Similar Crimes. Data source: *Starr, Sonja B, M. M. Rehavi (2014) "Racial Disparity in Federal Criminal Sentences." J. Pol. Econ. 122, no. 6: 1320-54.*

|  |  |  |
| --- | --- | --- |
| Prison time (X axis) | Black | White |
| No Time | 20% | 32% |
| Up to 1 year | 6% | 8% |
| 1-5 years | 39% | 41% |
| 5-10 years | 20% | 12% |
| Over 10 years | 15% | 7% |

**Task 2:** **Create visuals on the computer.** (40%)

Make two different visualizations of this data using a computer. You can make it in Excel or any other program.

Get some inspiration: <https://datavizproject.com/> <https://datavizcatalogue.com/> <https://depictdatastudio.com/charts/> <https://raw.githubusercontent.com/Financial-Times/chart-doctor/main/visual-vocabulary/poster.png> <https://www.datylon.com/blog/types-of-charts-graphs-examples-data-visualization>[https://www.visual-literacy.org/periodic\_table/periodic\_table.html#](https://www.visual-literacy.org/periodic_table/periodic_table.html)

**2a.** Paste your charts here. Use as much space as you’d like. (30%)

**2b.** Which one of your two visualizations more effectively communicates the message you want to send and why? (10%)

In this case, the line graph more effectively communicates the data because it clearly illustrates the trends in prison time distribution between Black and White individuals across the given time interval. By connecting each of the data points, we can easily compare how the percentages change as prison time increases among our two groups. This data visualization can also show the patterns between the two groups across the various sentencing lengths, in this case highlighting how White individuals don’t get as many large sentences compared to Black individuals even though they committed similar crimes. While the bar graph conveys this message, it lacks the ability to show progression and trends as clearly as the line graph does, requiring a deeper analysis.

**Task 3:** **Create visuals using an ‘analog’ (non-digital, not computer generated) method.** Now, make a chart using stuff around the house (and take a photo). Some students have used food, pillows, tea light candles, scraps of paper, paper plates, etc. The chart doesn’t have to be different than one of the ones you chose above. You can make two if you’d like (as you did in task 2a) (optional). (30%)

**Paste your photo(s) here.**

**A graph of lipsticks on a desk

Description automatically generated**

**Task 4: Reflect on data. Answer the following questions:** (30%)

**4a.** Look at your dataset from Task 1. In the data, the columns each sum to 100%, meaning that the data are already normalized to show a meaningful split (i.e., you don’t have to divide each cell by the column total—it’s done already). (a) Would it make sense to normalize the rows by the row totals? (b) What question would you be answering (instead) if you did it that way? (c) What is gained or lost? (20%)

1. No, it would not make sense to normalize the rows by the row totals. The data is currently presented in a way that shows the comparison of how prison time is distributed between both racial groups. However, normalizing the rows would change it to a comparison of the distribution of racial groups across different prison time categories, which is a completely different statistic (not the goal of the original data).
2. What percentage of people in each prison time category are black or white?
3. You would gain information into the racial composition within each prison time category which can be helpful in understanding how racial demographics are distributed among different sentence lengths. But you would lose the ability to directly compare how sentencing time is distributed between black and white individuals. This shift would make it harder to see the trends in racial disparities in sentencing, which is crucial for understanding issues in the justice system—the primary focus of the original data.

**4b.** When given percentages, you are missing out on raw values (i.e., number of people)—reflect on this decision and why it could be helpful or harmful for the message you want to send. (10%)

Percentages are useful because they allow for simple comparisons between groups of different sizes, making it easier to spot proportional differences and inequalities. In this case, showing how prison sentences are distributed between Black and White individuals in percentage terms highlights disparities without being influenced by the size of each group.

But focusing solely on percentages can be problematic because it overlooks the actual number of people involved. For example, a small percentage difference might represent many individuals if the group sizes are very different. Without raw data, the true extent of the problem can be underestimated, potentially minimizing the real-world impact and significance of the disparities observed.

For this data set, I believe it would’ve been beneficial to include raw values, to show the extremity of the situation and more heavily emphasize the difference between each ethnic group

**Turn in your answers as a word document (preferred) or PDF on Canvas.**

My checklist:

* 1. ~~I have two different computer-generated charts.~~
  2. ~~My computer-generated charts have axis labels and titles.~~
  3. ~~I have made a chart (or two) using an analog method.~~
  4. ~~I have answered the 3 questions (2b, 4a, 4c) (~500 words~~ **~~total for all 3 questions~~** ~~max).~~