Guidelines for making accurate and informative graphs have long existed. These include choosing the right type of plot, selecting a good color palette, adding informative axes labels, and so on. These guidelines are based on our perception of length, area, shape, and color, and on the kinds of data being plotted. More recently, additional guidelines are being created that have an equity lens that aims to avoid reinforcing stereotypes and biases conveyed through choices of graphing elements. A core consideration is to bring empathy into data visualization by considering how the communities whose data are being plotted and the target audience will perceive the graphs. Today, more people are interacting with visualizations in the media. Unlike in the past, the community of people on which the data were collected are reading and interpreting plots. These changes have led to a more in depth study of how bias can enter into a plot through the symbols and language used in its presentation.

1. Giorgia Lupi carried out an unusual project where she and another information designer mailed postcards to one another every week for a year. Each week, they agreed on a topic about themselves, collected data, drew a graphic, and mailed it to each other. This project led Lupi to identify what she calls "data humanism" in visualization.

   Watch the short TED talk by Giorgia Lupi. What does Lupi mean by *data humanism*?

   > **Solution:** We should think about the data behind a visualization more holistically as coming from people who are more complex than a single (or a few) quantitative dimensions.

2. Read the 4-page paper "Applying Racial Equity Awareness in Data Visualization" by Schwabish and Feng of The Urban Institute. (Also available as a blog post). Note that the authors are not focusing on trickery in plots, where the visualization is purposefully distorting or hiding the truth. They are discussing design choices that have a more inclusive approach to the people represented by the visualization.

   Summarize each of sections 2 through 9 of the paper into one to three main points.

   > **Solution:**
   >
   >   1. Language
   >
   >   - Label your data using people-first terminology

- Terms and phrases continue to change and up-to-date version should be used

2. Ordering

   - If your study focuses on a specific community, what other communities, if any, should be included?
   - Is there a quantitative relationship that guides the ordering of groups? Would ordering it in increasing vs decreasing order reduce the emphasize on a group?
   - Does the ordering reflect your argument?

3. Missing Groups

   - Can the visualization be disaggregated further to be more informative?
   - Is an "Other" category necessary, and, if so, is there a more inclusive label for it?

4. Color

   - Is color needed?
   - Be aware of the connotation of color and avoid stereotypic colors.
   - Select color palettes for visually impared viewers and to represent categories and scales fairly (see Cindy Brewer's palettes).

5. Shapes

   - Use gender- and race- sensitive imagery, if needed
   - Be careful that the size and area of images does not misrepresent the magnitude of the measurements.

6. Empathy

   - Consider what type of plot will resonate with the reader
   - Try to express the presence of people in the plot

7. Default approaches

   - Make the effort to tailor the plot to the audience and to best present the findings, rather than take the default settings for plots.

8. Reflect lived experience

   - Ask others to review and make suggestions for your visualization, especially those that are represented in it.

Consider the following plots from a data humanism and equity perspective. Identify shortcomings of each plot and suggest how it might be fixed.

3. For background, according to Kaba et al, "the care of persons with mental illness in the United States is inextricably linked to the criminal justice system. .... Approximately one third of these persons have an identified mental illness diagnosed before or during incarceration. Treatment and discharge planning for this population represent considerable challenges." Do you see any problems with color, labels, ordering?
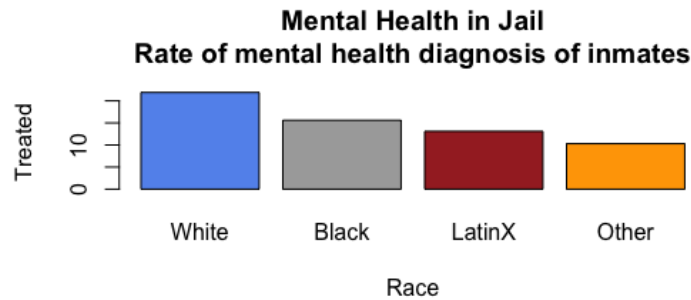
Solution:

Problems:

- Title is not people-first because it calls people inmates and has a generic main title that doesn't convey the aim of the study.

- The label LatinX is not a preferred label by the group referred to (at this point in time)

- The colors are similar to skin tones, and are not necessary to convey the information.

- The label on the y-axis is inaccurate.

Possible Solution:

- Update the titles to be people-first and to point out the disparity in access, e.g., Access to Mental Health Diagnosis in Jail: People of Color less likely to receive diagnosis

- Relabel LatinX to Hispanic

- Remove the color from the bars. Alternatively, use colors that are easy to differentiate, have the same level of perception, and are not connected to skin color. See Cindy Brewers' color palettes for examples.

- Correct the label on the y-axes to "percent receiving diagnosis"
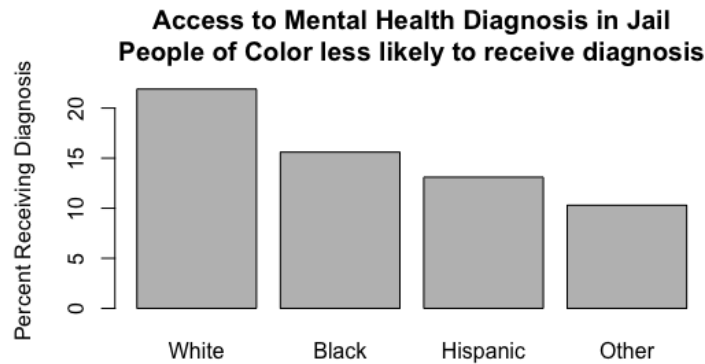
4. The figure below breaks several of the traditional guidelines for graphics, as well as the new standards. Compare the icon for Lithuania and India. How much bigger is the Lithuanian? That is, how many icons for India would fit inside the one from Lithuania. Does that make sense? What is causing this misrepresentation?
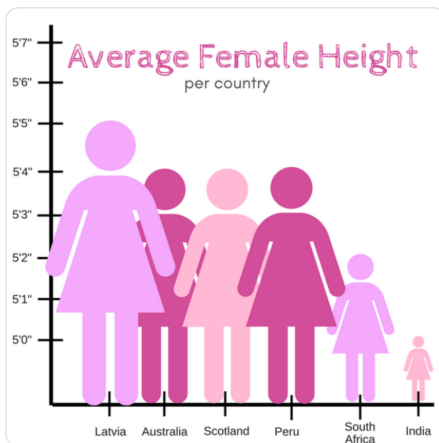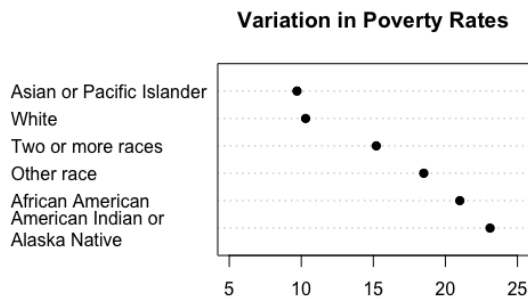

Solution: Problems:

- Use of traditional pink for female.

- Y-axis doesn't start at 0, which makes the shorter groups appear even smaller than they are proportionally.

- Icons use height to represent the average, but the taller icons are also larger in terms of area (and girth) which exaggerates the difference between groups.

> • The selection of only 6 countries makes it seem like the shortest group is in the minority. In addition, the population size of these countries varies widely, and the representation of women in the world is lost in this plot.
>
> Possible solution:
>
> • Use dots or same sized icons to represent the average height, or a dot chart might be even better.
>
> • Include 0 on the axis denoting average height.
>
> • Include more countries, and arrange countries by geographic region.
>
> • Eliminate color, or if countries are organized by geographic region, use color to represent regions.

5. Below, the percentage of individuals living below the poverty line are plotted according to race. Bring what you know about the composition of these categories to suggest a possibly more informative plot.

Variation in Poverty Rates

> **Solution:**
>
> Problems:
>
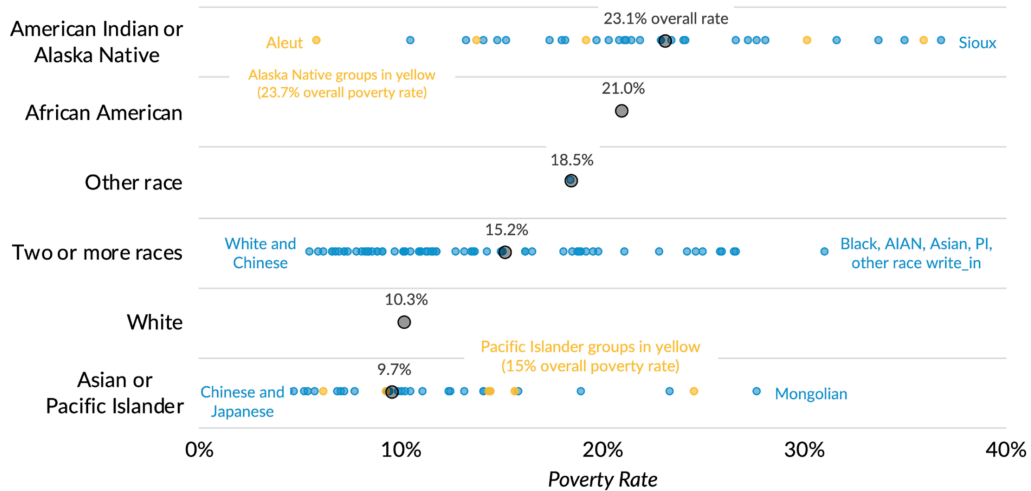> • The axis for the dots begins at 5, which makes comparisons difficult.
>
> • There is no label on the x-axis so we don't know if 5 is a percent or a count or what.
>
> • The group labeled Asian or Pacific Islander contains many subgroups which might have quite different poverty rates.
>
> Possible solution:

- Begin the numeric axis at 0.

- Label the numbers as percentages.

- Split the Asian group into subgroups to get a more nuanced picture.

Variation in Poverty Rates

American Indian or Alaska Native — Aleut ... 23.1% overall rate ... Sioux
Alaska Native groups in yellow (23.7% overall poverty rate)

African American — 21.0%

Other race — 18.5%

Two or more races — White and Chinese ... 15.2% ... Black, AIAN, Asian, PI, other race write_in

White — 10.3%

Asian or Pacific Islander — Chinese and Japanese ... 9.7% ... Mongolian
Pacific Islander groups in yellow (15% overall poverty rate)

0%   10%   20%   30%   40%
Poverty Rate

Source: Schwabish and Feng (2021).

6. Use your summary of Schwabish & Feng and the experiences that you just gained from critiquing and fixing plots to create a list of five equity-focused guidelines for visualization. Your list should include the topics: language, order, color, and comparisons. They may combine ideas from more than one of the eight sections in the article.

Solution: The following is one of many possible solutions. However, it is based in good principles of visualizations.

1. Do the title, axis label, and tick mark labels bring a human-centered perspective to the visualization that is sensitive to today's descriptors?

2. Does the ordering of bars, dots, facets, etc. reflect the goals of the comparison, make comparisons easy while avoiding the traditional power hierarchies, and focus on the important findings?

3. Is color needed or is it a distraction? If color conveys useful information, has the color palette been selected to match the type of data, give equal visibility to all groups, and stay away from negative connotations?

4. Does the breakdown into subgroups align with the goals of your analysis? Are all of the comparisons needed? Should additional subgroups be created? Would an overlaid plot or facets better convey the information?

5. Imagine/ask how a person in the study might see themselves in the visualization, and revise accordingly.