**1. Sensory overload**

There are two basic measures of how good your data visualization is.

**2. Measures of a good visualization**

The first measure is how many insights can the reader get from your plot. For far too many plots, the answer is zero. In this video we'll look at some ways to avoid that situation. The second measure is how quickly the reader will get those insights. If your plot is being presented in a meeting, you might reasonably expect your audience to concentrate on your plot for twenty seconds. Often, you need to get your message across quickly.

**3. Chartjunk**

Chartjunk refers to anything in the plot that makes it harder for the reader to get insight into the data. We'll look at some common problems in the following slides. One term that's less well known is skeuomorphism. That means adding things that happen in the real world to virtual objects. For example you could add shadows to bars in a bar plot.

**4. Pictures**

Here's an example from the BBC TV network, found on the Junk Charts blog. This probably ought to be a bar plot, but since the y-axis doesn't begin at zero, it suggests that maybe the BBC was aiming for a dot plot. In which case, is the height represented by the top of the head of the center of the head? Using a picture of a man instead of a bar makes things harder to understand, not easier.

1. 1 https://junkcharts.typepad.com/junk\_charts/2013/09/the-incredibly-expanding-male.html

**5. Skeuomorphism**

This example was created by the Fox News TV channel, and found by Media Matters for America. You can see some familiar problems with the bar plot, like the y-axis not including zero, distorting the relative heights of the bars. Also notice that the time periods have been arbitrarily limited to October to April, omitting what happened between May and September each year. Let's focus on the skeuomorphic elements of the plot. There's a shadow across the panel of the plot, which adds no value, but since it travels upwards and to the right, it gives a subconscious signal that numbers are increasing. The bars have also been given a slight depth to them. The 3D aspect makes it harder to accurately judge where they lie against the y-axis, so the bar heights have to be written in text to compensate. By stripping away all the junk, the data is easier to see.

1. 1 https://www.mediamatters.org/fox-news/fox-news-newest-dishonest-chart-immigration-enforcement

**6. Extra dimensions**

This plot was found on the dataisugly subreddit, and has many problems. Try pausing the video and counting the issues you can find. Let's concentrate on the unnecessary use of 3D perspective. Bar plots are inherently two dimensional. You have categories on one axis, and a continuous variable on the other axis. By making bars three-dimensional, you don't provide additional information, but simply make it harder to judge the lengths of bars. A 2D bar plot is easier to read, although one category label is missing, and the percentages don't add up to one hundred, so it's unclear what the y-axis represents.

1. 1 https://www.reddit.com/r/dataisugly/comments/dh6yra/an\_actual\_graph\_presented\_in\_my\_psychology

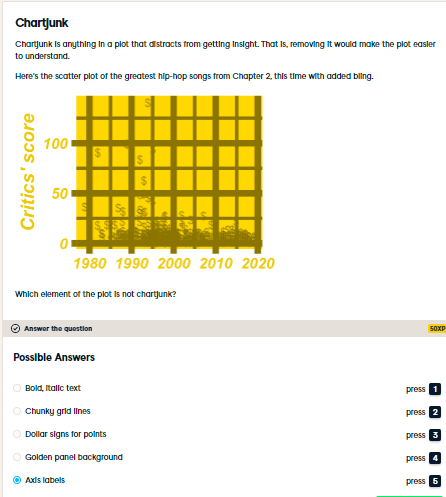
**7. Ostentatious colors and lines**

Here's another plot from dataisugly, showing seats in the German Federal Council. There are sixteen states and seven political parties. Parties typically have to enter into a power-sharing coalition to get seats in each state. Each political party has a color associated with their brand: CDU is black, SPD is red, Grüne is green, and so on. Each state on the map or slice of pie plot has the color of the party which won the most votes and stripes colored by the other parties sharing power. The problem is that the plot is just too ugly to look at, so no-one can get any insights from it. In general, stripes or other forms of hatching should be avoided in plots because they're just plain difficult to stare at. One reddit user tried to improve the plot, as shown on the right. By removing the hatching, it's much nicer to look at the plot.

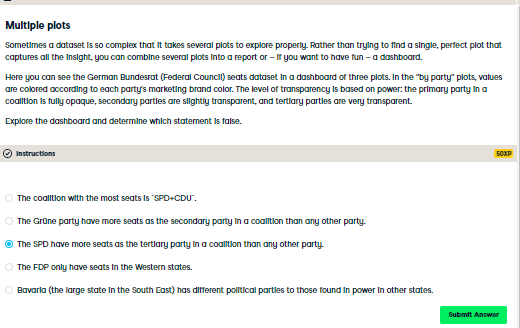
1. 1 https://www.reddit.com/r/dataisugly/comments/cyhle4/wikipedias\_image\_of\_the\_distribution\_of\_seats\_in
2. 2 https://imgur.com/y27mRLj

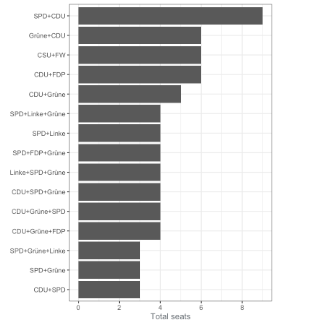
**8. Let's practice!**

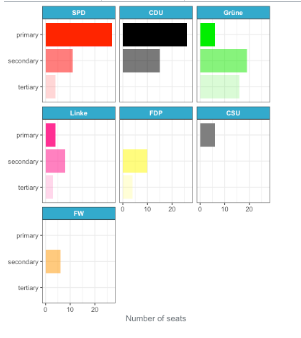
Let's go to war on chartjunk!

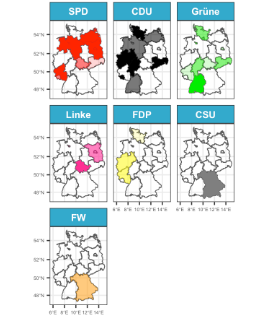


Joyous junk detection! The font face may have been terrible, but having axis labels helps readers interpret the plot, so they aren't chart junk. In general, anything that makes it harder to interpret a plot should be removed.









Perfect political analysis! The Grune party had more seats as the tertiary party in a coalition. For complex datasets, it is often best to draw lots of simpler plots that each answer a couple of questions, rather than trying to draw a single plot that answers everything.