In this video, we're going to look at scales.

This first plot shows the average height of a 21-year-old male in centimeters.

The x-axis is time, starting in 1871, and ending in 1975.

Each person represents the height, at a different point in time, and the points are evenly spaced in time, so the x-axis is OK.

The y-axis ranges from just under 160 to 180 centimeters, which isn't inherently bad, but does overstate the change.

The real problem is the bars.

If it was accurate, we would only really see the heads of the men, but instead we see their whole bodies, making it seem as if people have not only doubled in height, but they've also double in width.

This next plot also has issues with scale.

The total range of the plot is 8% to 10%, although all the numbers fall in the range of 8.6% to 9.2%.

If we plotted the y-axis on a 0% to 10% scale, the conclusion would be that nothing is really changing at all.

The last point in the chart is at the wrong height, and the size of the markers makes the relative locations hard to distinguish.

Also notice that the gap between 9.0% and the 8.9% markers on the far left side, and the 8.9% and 8.8% markers, have a different gap.

This plot shows the relative breakdown of teachers by race in a certain teaching program.

The Caucasian bar is truncated, which is a risky choice, but could be appropriate in some situations.

A much bigger problem is that the scale of each blue bar is entirely different.

For example, the Native American bar is about a third of the length of the African American bar, but there are more than 10 times as many African Americans in this program as Native Americans.

In fact, visually, this plot is completely meaningless.

The only useful thing about it is the numbers.

But even there, there is a bit of confusion, as Native Americans are given to one decimal place, but the others are rounded.

Which when combined with the confusing scales, casts doubt on the correctness of the numbers.

Here is a before and after of the same data.

On the left, we see the US military expense in the right axis, and troop count on the left axis.

Both the line and bar plots are individually OK, but the combination is misleading.

Because you have mixed two units, dollars and people, there is a false impression of some sort of crossover point in 1995 that does not exist.

On the right is the same data presented in a different way.

We now have troops on the x-axis, and dollars on the y-axis.

The line moves through time now, allowing us to see when moments of change occurred, such as decreases in troop count, through the 90s, at the end of the Cold War, the increase in spending of the 2000s, and the recent decreases in military spending.

The final visualization I want to show you today is all about the different types of household.

The US Census Bureau periodically determines how many households are comprised, for example, of married couples with and without children, people living alone, and so on.

First of all, I'm not saying this is a bad visualization.

In fact, if we are interested in the relative share of each type of household in a particular year, it's actually pretty good.

However, if what we're interested in is the rates of change across the years, this is next to useless.

The key problem is that the x-axis is completely off.

The gap between the first two columns is 10 years, but the gap between the last two columns is only 2 years, meaning that the rates are hard to read from this.

If we're not interested in the rates of changes, but just want to compare two years at a time, it's not bad, but it's not easy either.

Try comparing 1970 married without children to 2010 married without children, without looking at the numbers.

Can you tell if it has grown or shrunk?

Finally, and more generally, this chart shows relative numbers.

If you look at absolute numbers, we might find the total number of couples married with children is actually constant, but the number of other households has increased.

We are now going to change into R to try plotting this data as a line chart.