

- Getting to know your data is super important because you need to know what you can do and what you need to do.
- You get a base understanding of it through summary statistics

- this is a play on what is known as Anscombe's quartet.
- each of these graphs have the sameish summary stats with slight differences on the very tails.
- if you just use summary stats you aren't seeing the patterns in your data. We as humans are good at picking out patterns. Most of us are really bad at looking at a spreadsheet and figuring out patterns in it.

- I am doing a surface level coverage of a huge debate
- I often joke that data viz is art for people with no artistic talent.
- It is a skill that will make you stand out at the work place

- The most important part of sort of "rule" for dataviz is that you are honest.
- the data are accurate as it can be, you are not letting the data tell you things. Data will lie to you if you don't choose the right measures. One example of this was a sort of an alarmist story a few years ago that
- Perceptual issues contains a lot. This can include issue related to colorblindness, low vision. Others are as minor as whether you have the right aspect ratio.
- I usually think of insightful as having two aspects to it. Bang for your buck and whether the audience gets something out of it they otherwise wouldn't. Essentially just the dinograph example.

- The contrast is .
- that is the only thing that is going for it.
- the graph implies that it represents spacial variation which is not true
- we can not understand what is going on in the corners.
- there are better ways to do this.

-
- Minard visualized Napoleon's March to Russia and back from Russia and is often held up as the best data visualization by the father of modern data viz Edward Tufte
- Lets walk through it.
- the brown line is the march to Russia and the Black line represents the . In both black and the brown bars the width represents the number of troops 1mm for every 10,000 troops.
- There are also little tick marks indicating the number of troops.
- You also have a sort of exact route demarked by major cities.
- At the bottom you have how cold it is when winter hits.
- The author also acknowledges some its inaccuracies in the notes.

- Widely used tools like SAS, excel and SPSS give you the option to point and click your way to glory. There is nothing inherently wrong with this.
- Thats how I started and would end up just copying the code into my do file.
- It is super easy however it can come at some costs.

- This is a fairly recent example of when things go wrong
- two notable economists published a paper in a fairly influential journal
- but their results were very wrong because amongst other things they were dragging down to the bottom of their excel
- ultimately it also will just take longer for you to do things to.
It is a slow process

- The internet is your friend.
- For the most part when you are doing stuff in Stata and R that does not require you to build in additional security.
- I learned by copy and pasting stuff then changing it and trying to figure out how I broke the code.
- It is unreasonable to expect that you can memorize every command and to work through fixes yourself.