

Graphics

(The Good, the Bad, and the Ugly)

Understanding Political Numbers

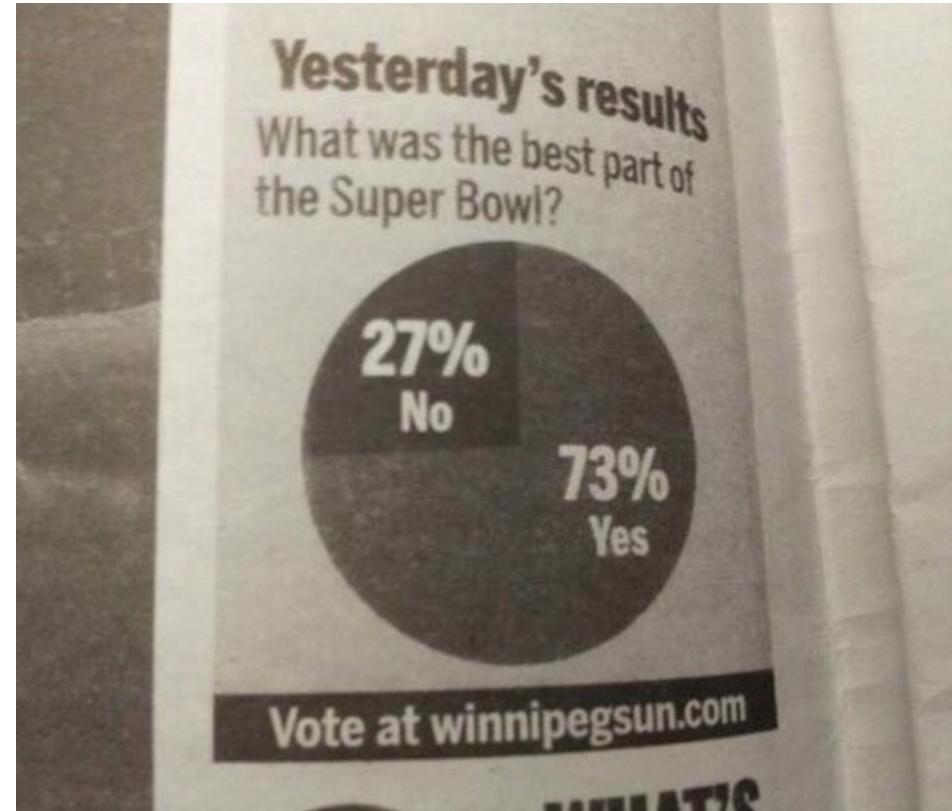
Feb 11, 2019

Agenda

Learning from bad graphics

Principles for making good graphics

Common graphics for common purposes



Graphics. Why?

Graphics. Why?

Is it about lots of info in a *small space*?

Graphics. Why?

Is it about lots of info in a *small space*?

Lots of info presented in an *intuitive, easily interpreted way*

Graphics. Why?

Table 2
McClurg 2006, table 1 (panel A): The political character of social networks

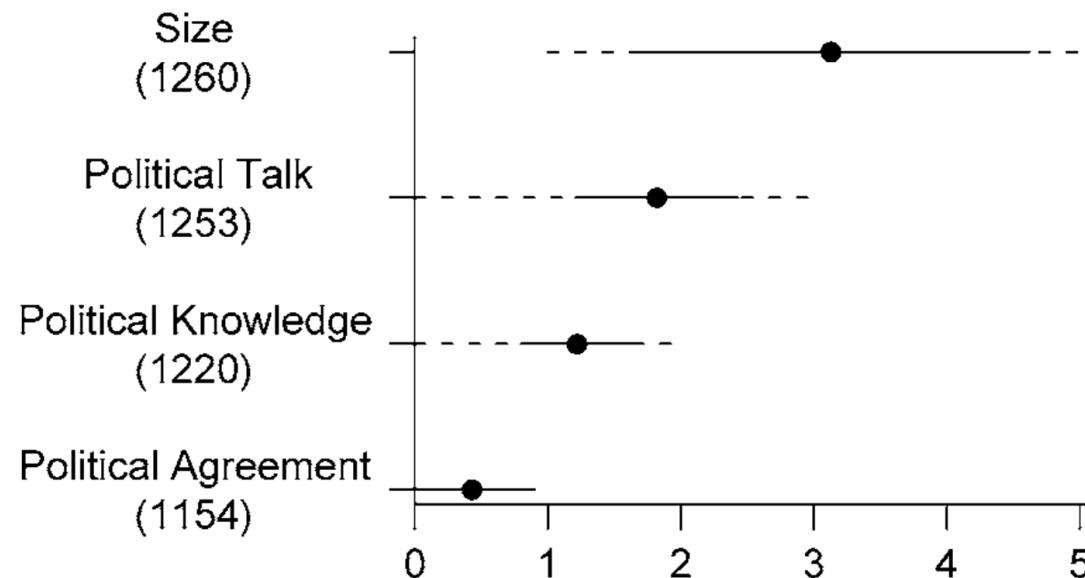
	Mean	Standard Deviation	Min	Max	N
Panel A: Descriptive Statistics					
Size ^a	3.13	1.49	1	5	1260
Political Talk	1.82	0.61	0	3	1253
Political Agreement	0.43	0.41	0	1	1154
Political Knowledge	1.22	0.42	0	2	1220

Notes: This table provides descriptive statistics for the political character of the social networks as perceived by respondents.

^aWhen respondents who report having *no network* are included the mean of this variable drops to 2.57 with a standard deviation 1.81 (n = 1537).

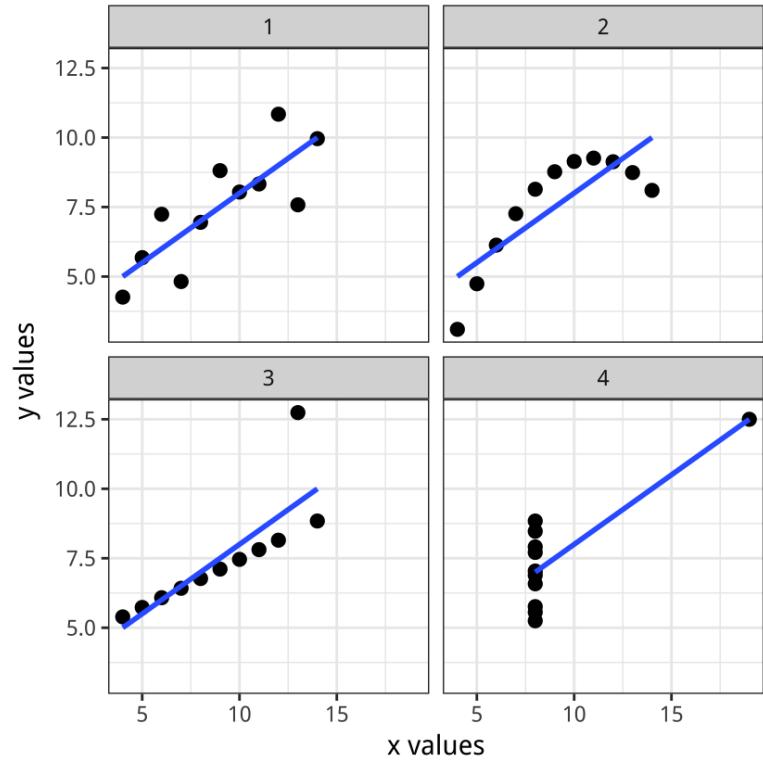
Graphics. Why?

Using a Single Dot Plot to Present Summary Statistics.

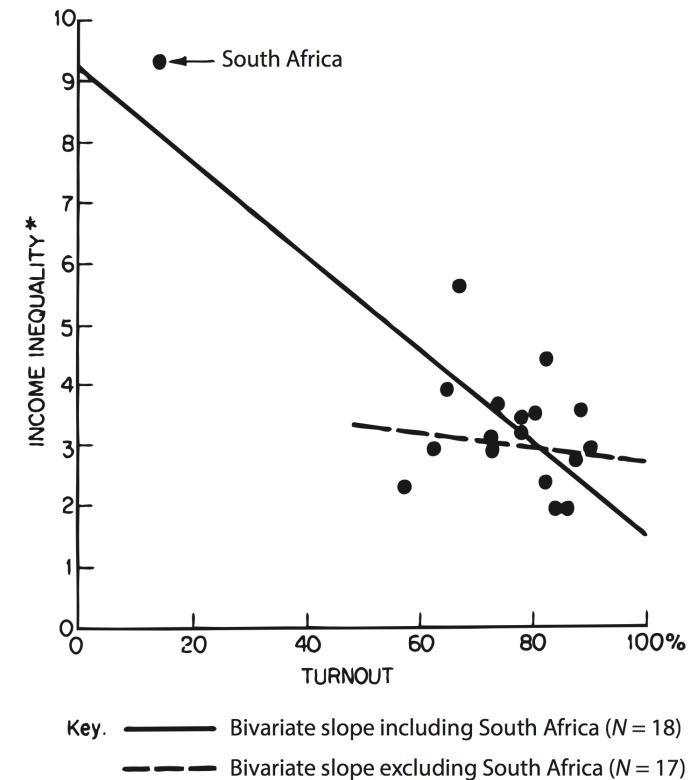
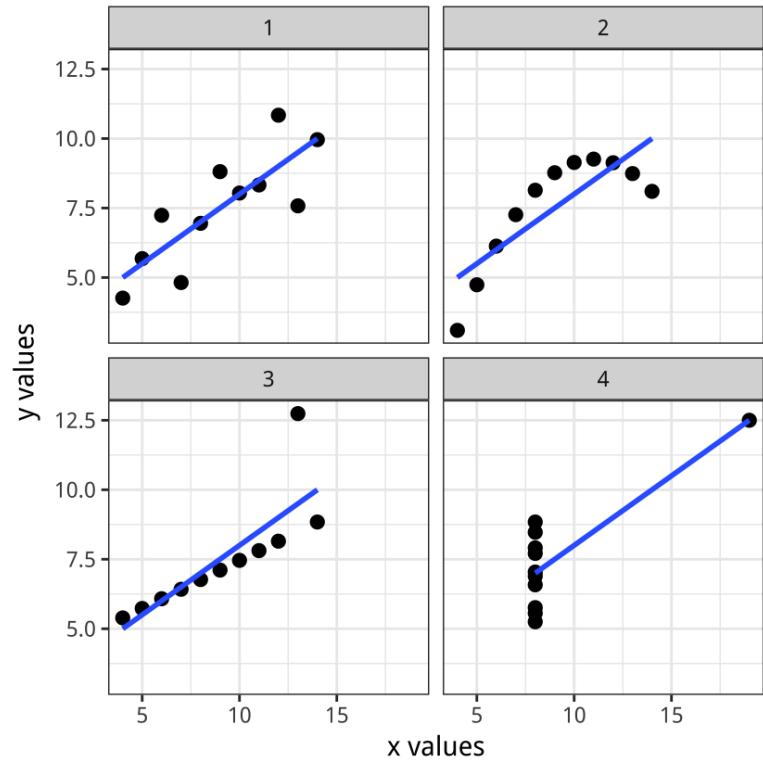


Source: [Kastellec and Leoni \(2007\)](#)

Know thy data



Know thy data



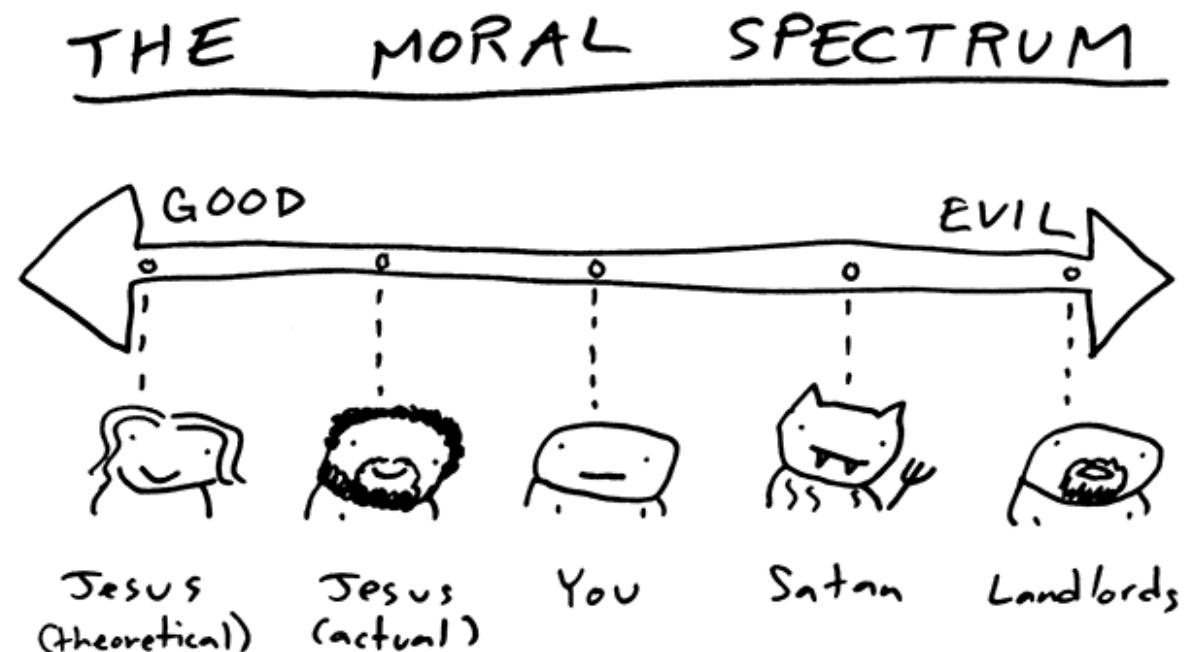
Source: Healy Ch 1

Axes

Graphics are *spatial* (meaning "space") representations of data

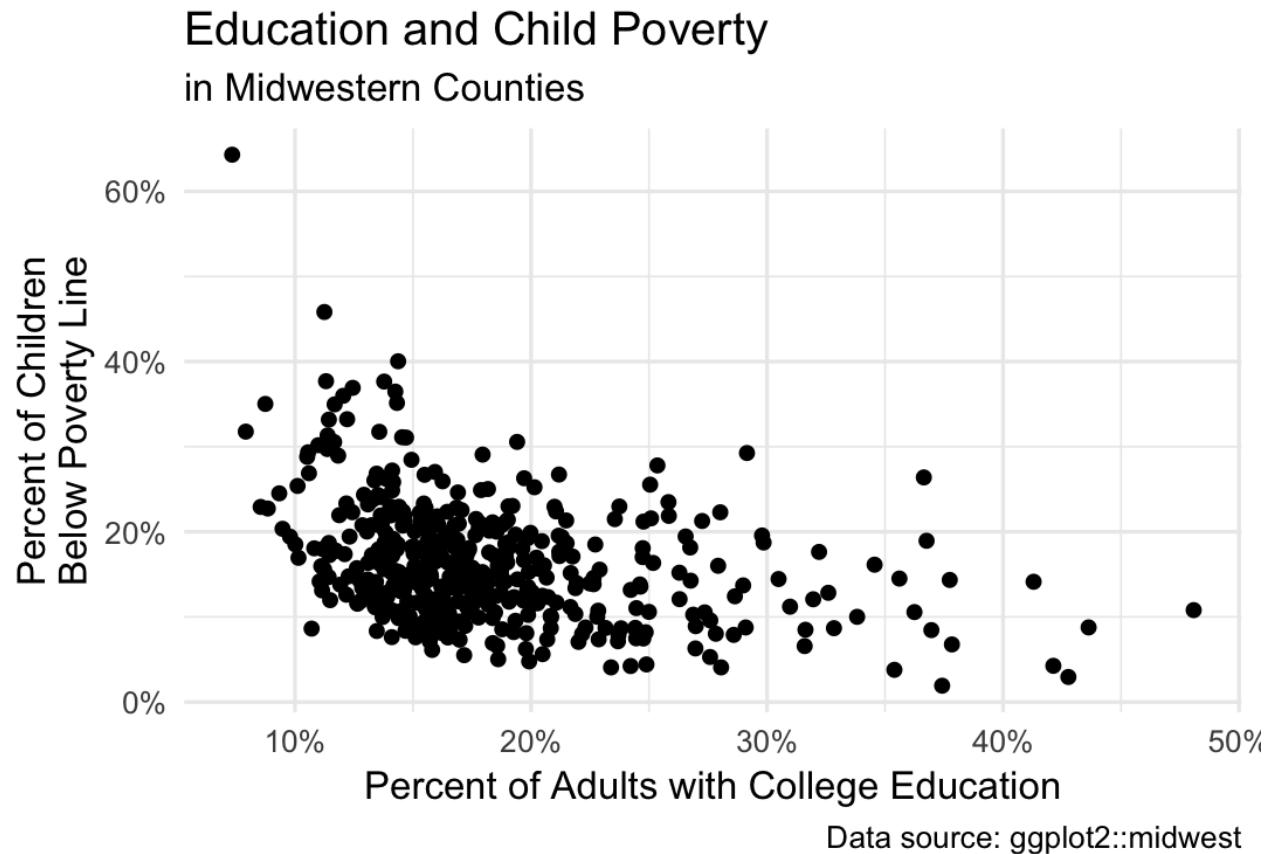
Space needs *dimensions* (axes)

One dimension: a line



Axes

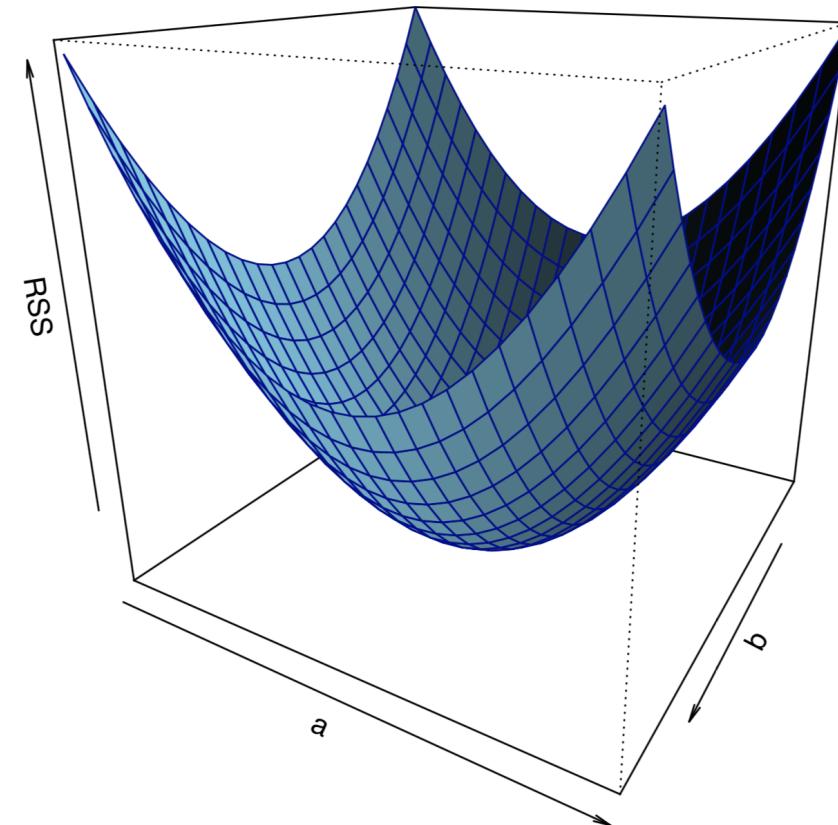
Two dimensions: a coordinate plane



Axes

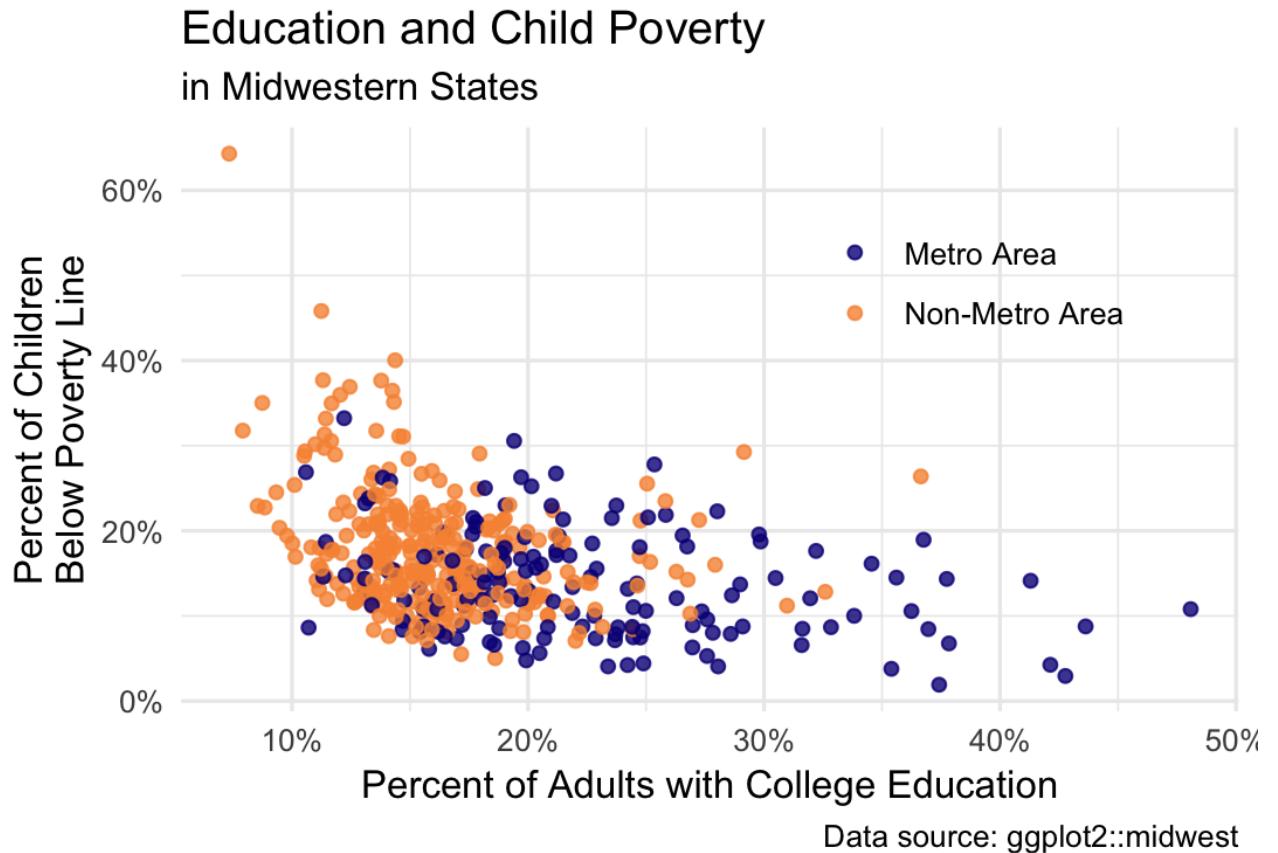
Three dimensions: a surface? (usually overkill)

(Easier to see more dimensions with use of color etc.)



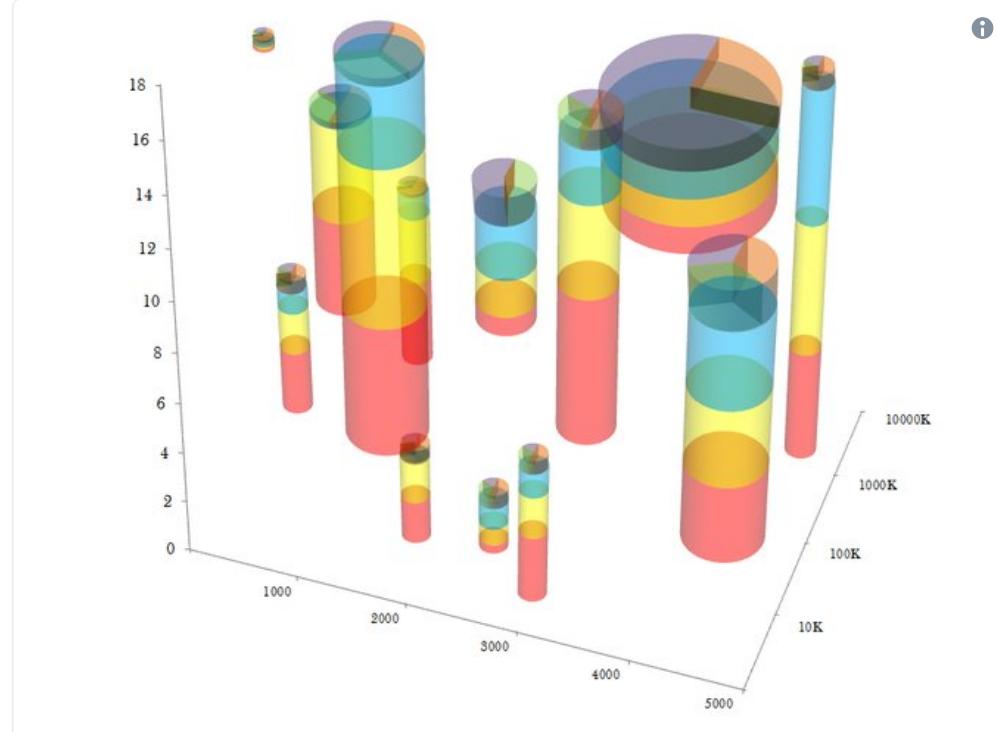
Axes

Three dimensions: a surface? (usually overkill)
(Easier to see more dimensions with use of color etc.)



Graphics are about *comparisons* and *variation*

Less than great examples



Daniel MacArthur
@dgmcarthur



Five years on, it's disappointing that the data visualization community never fully recognized the benefits of 3D stacked scatter pie columns: eagerpies.com/better-than-mi...

♡ 707 8:36 AM - May 28, 2018

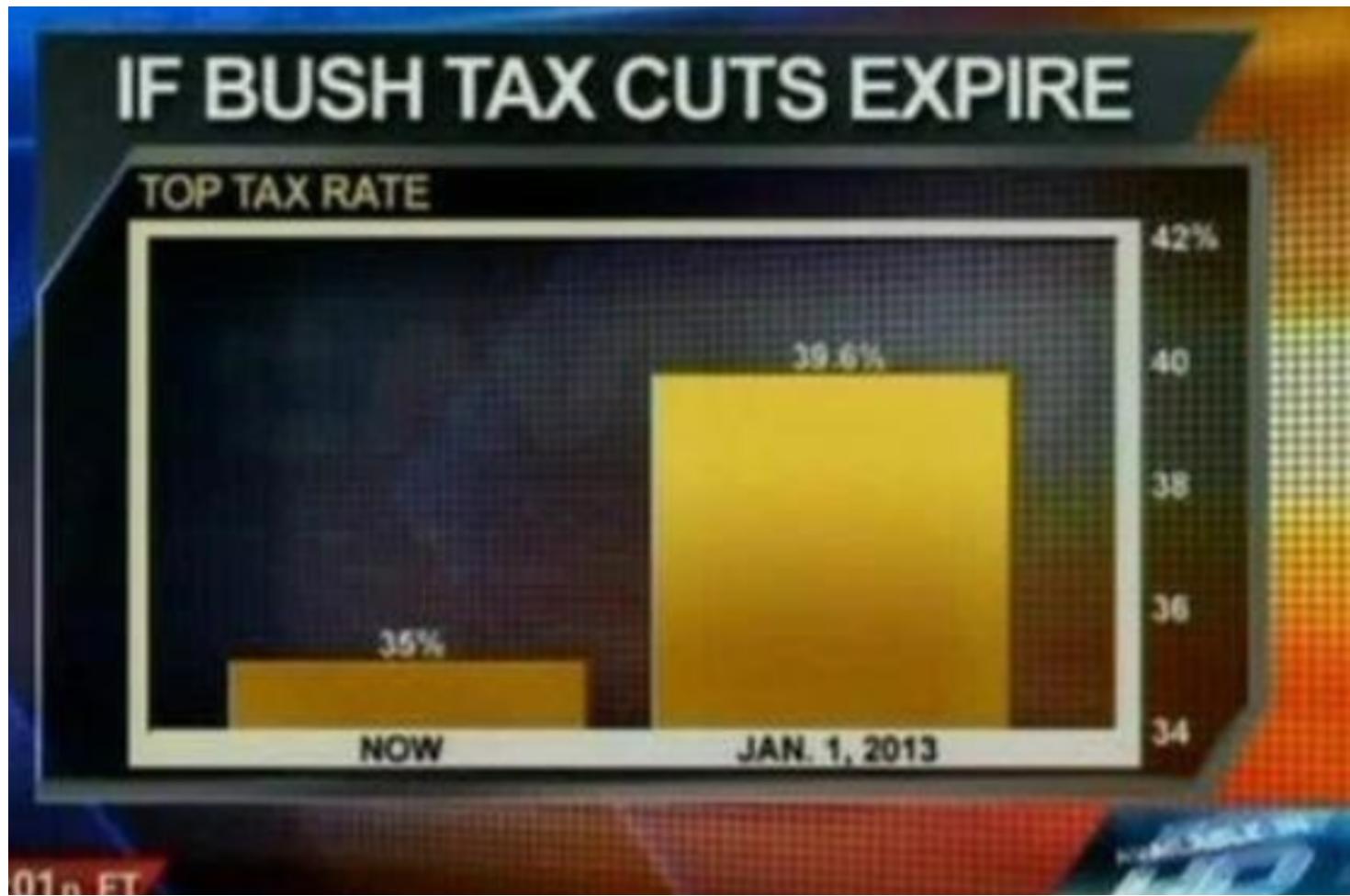
💬 261 people are talking about this

>

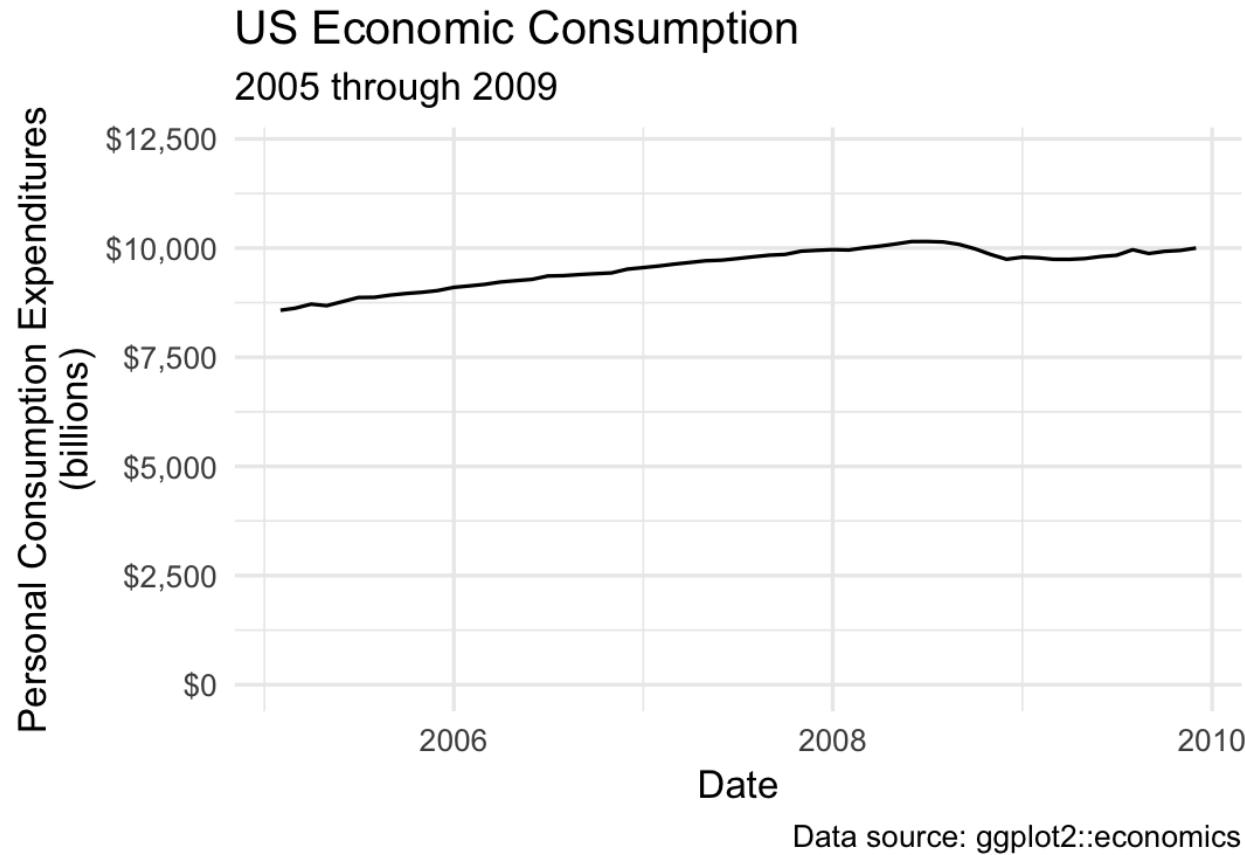
Let's talk about color



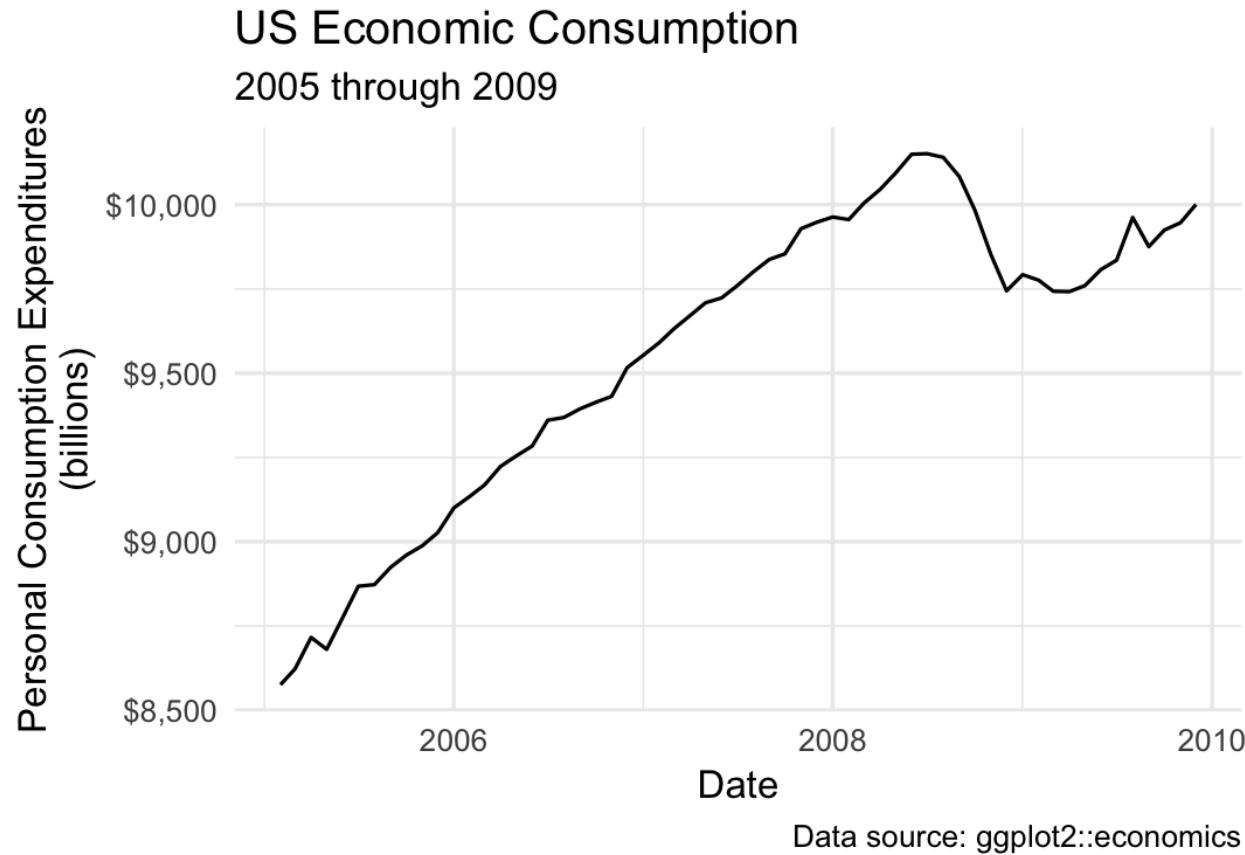
Let's talk about the y-axis



Let's talk about the y-axis



Let's talk about the y-axis



Ed Tufte: "Plot the data, not the zero"

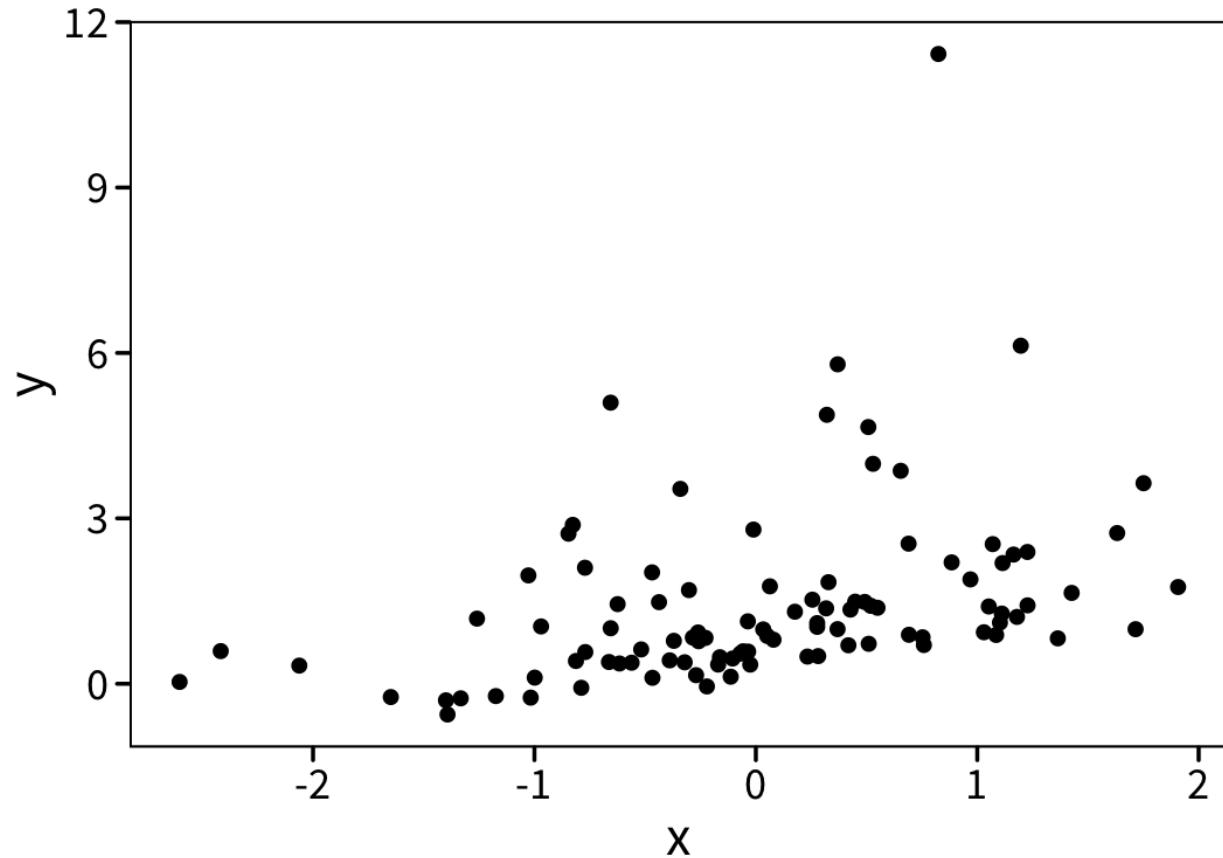
Failed Comparisons



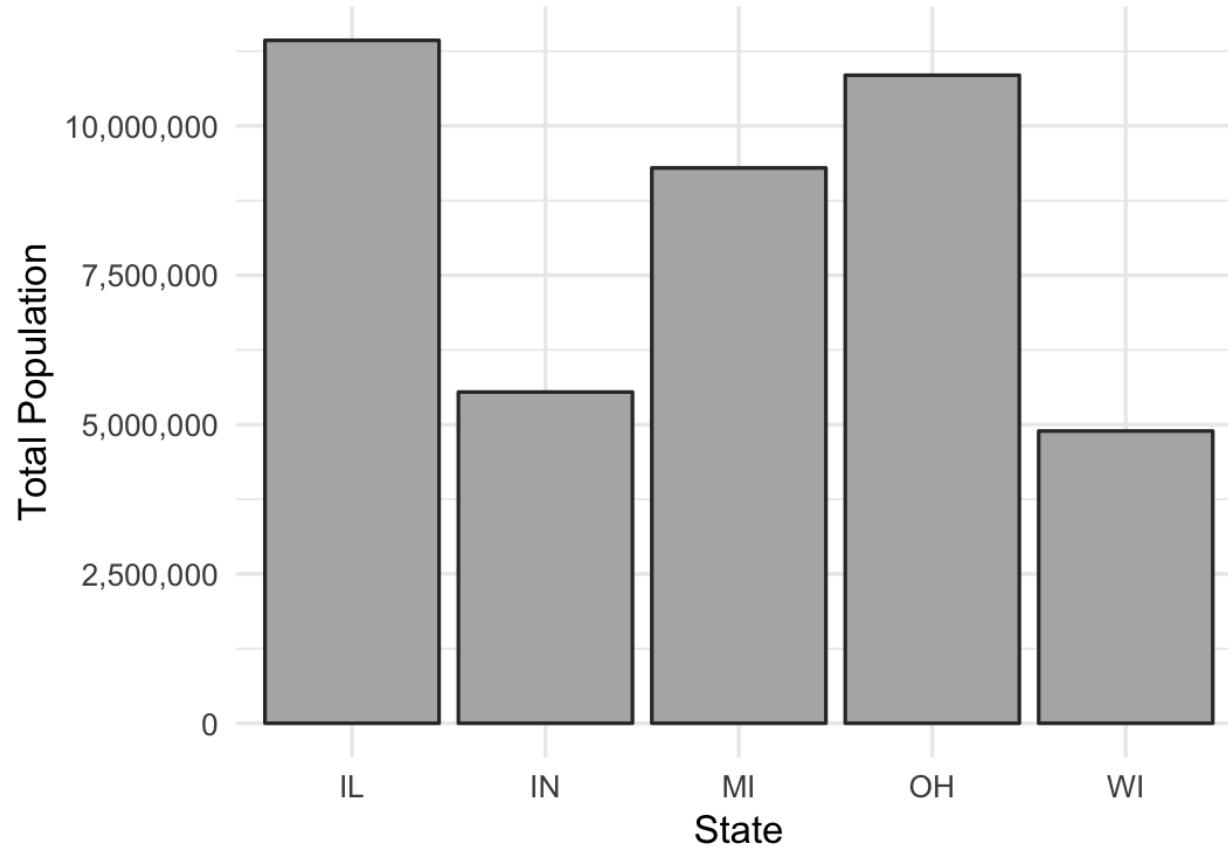
Conceptual dishonesty

Common graphics

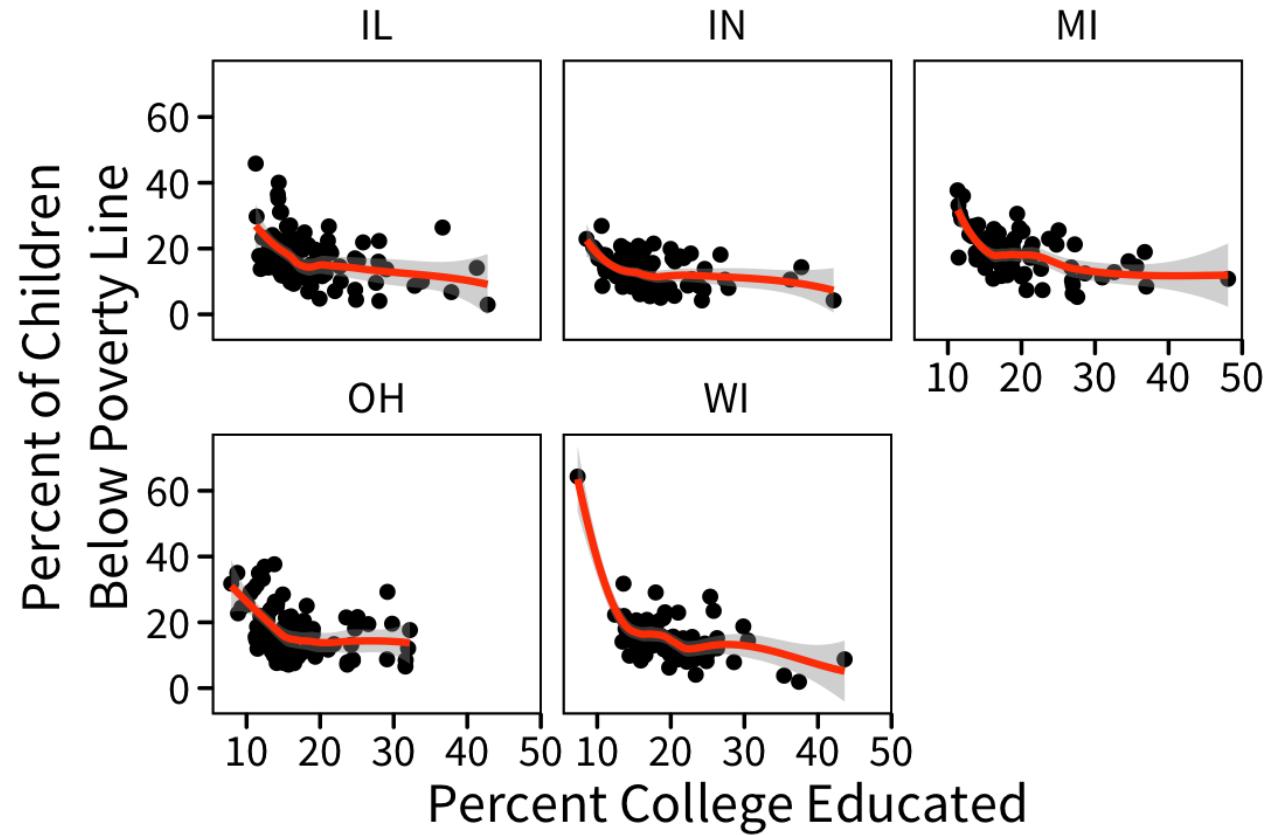
Scatterplots



Bar charts



Small multiples

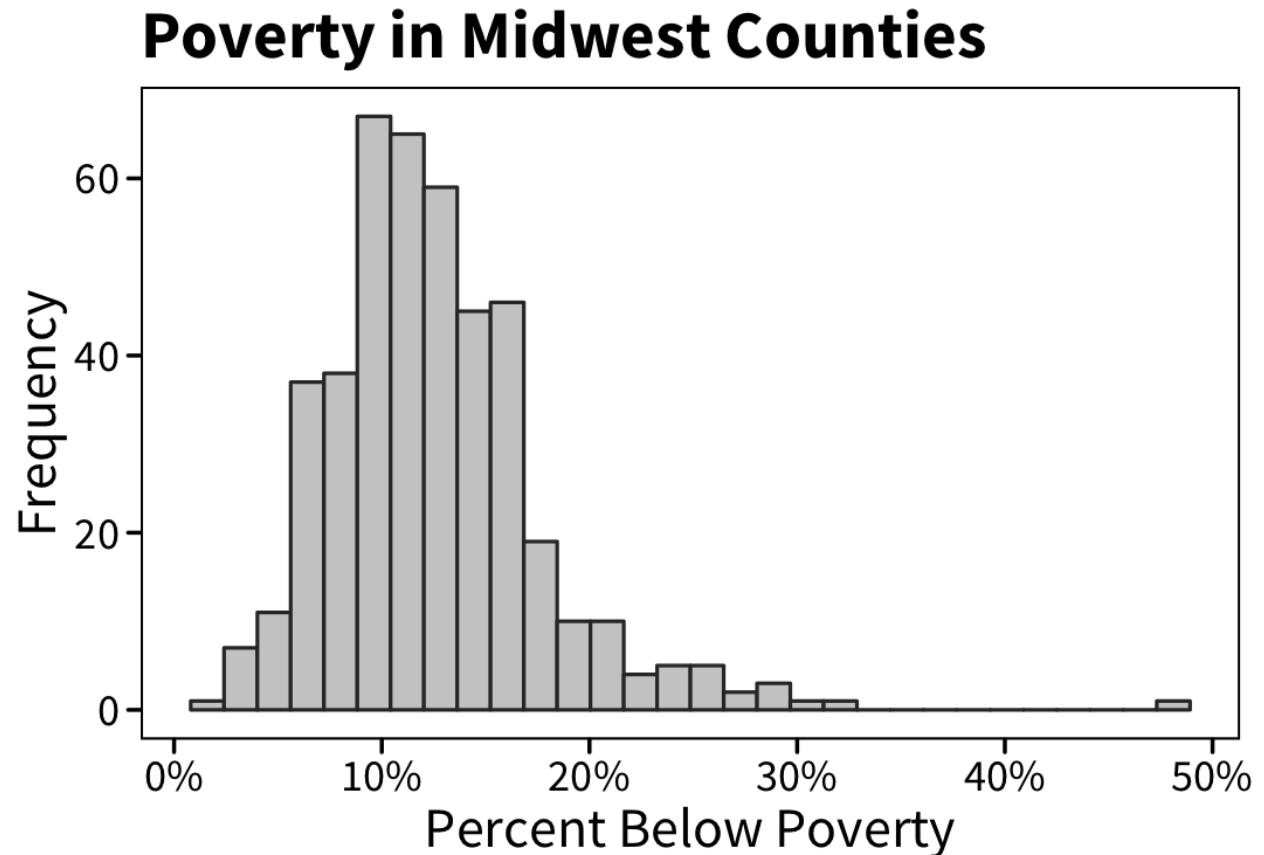


Technical lesson: *distributions*

Look at one variable. Where are the data?

- on the x -axis: What are the possible values?
- on the y -axis: How many of those values are there?

This graphic: a [histogram](#)

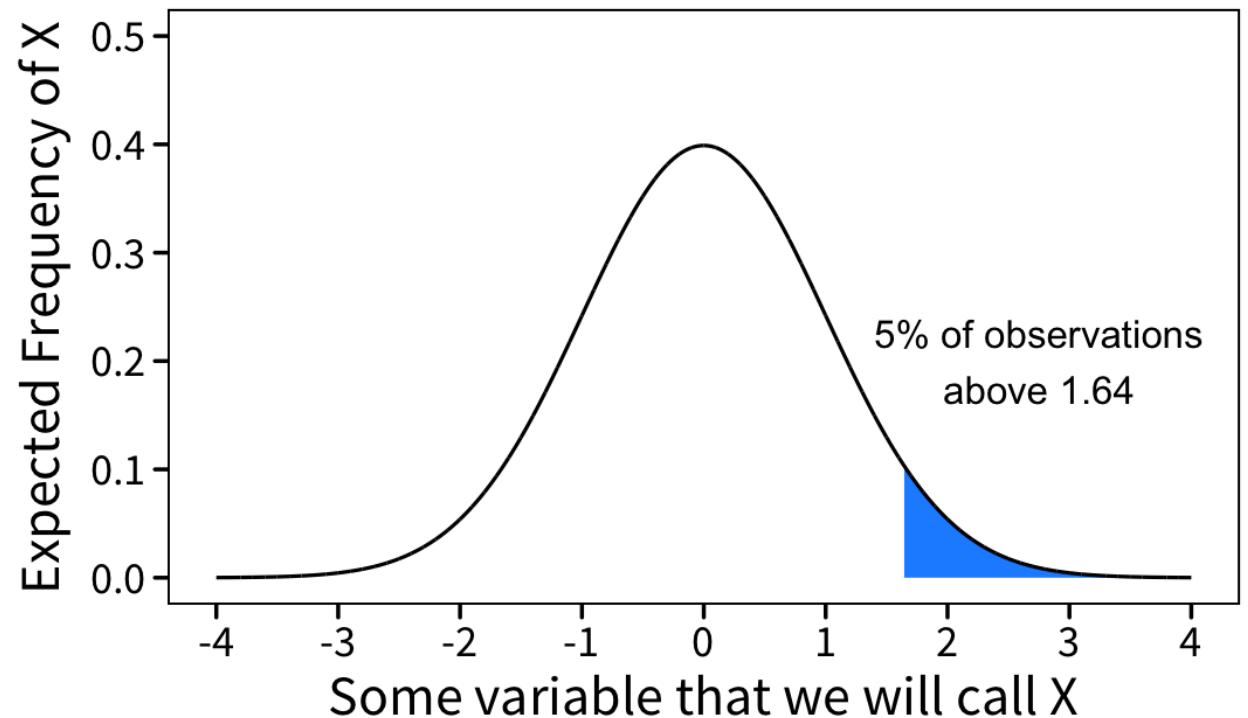


The Normal Distribution

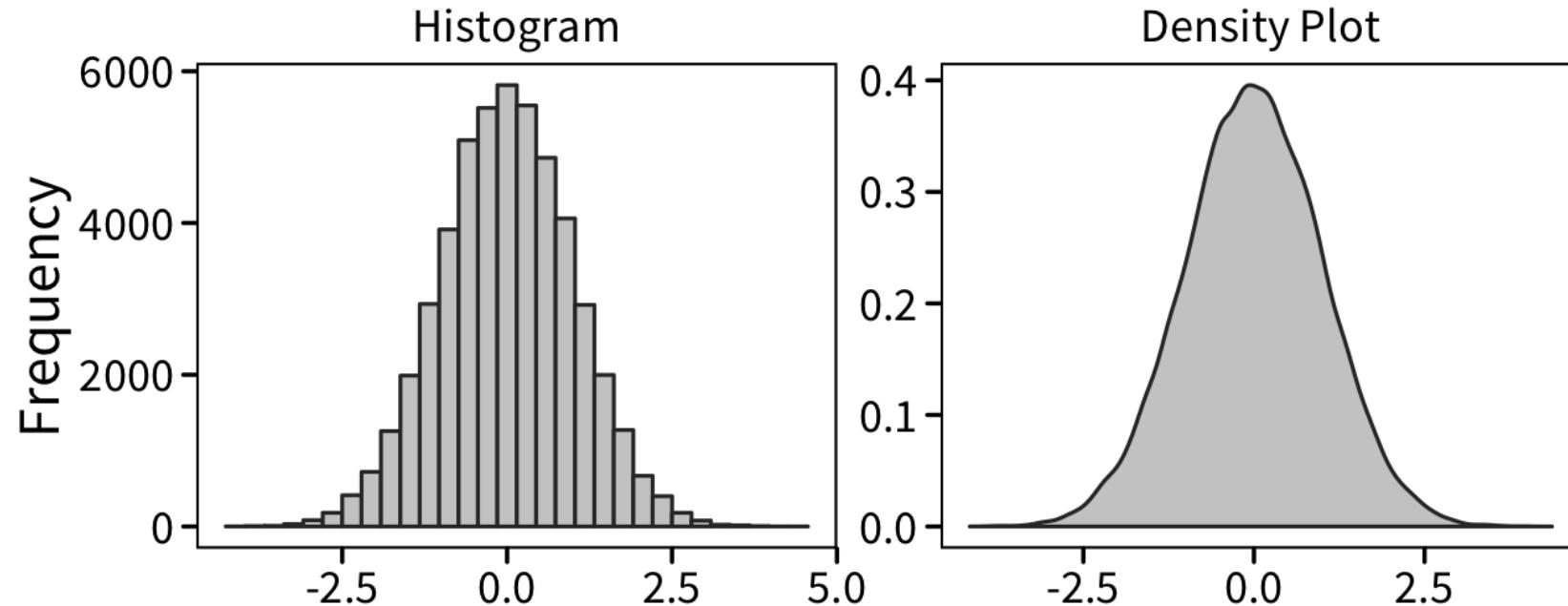
The "standard bell curve"

A.k.a. "Gaussian" distribution

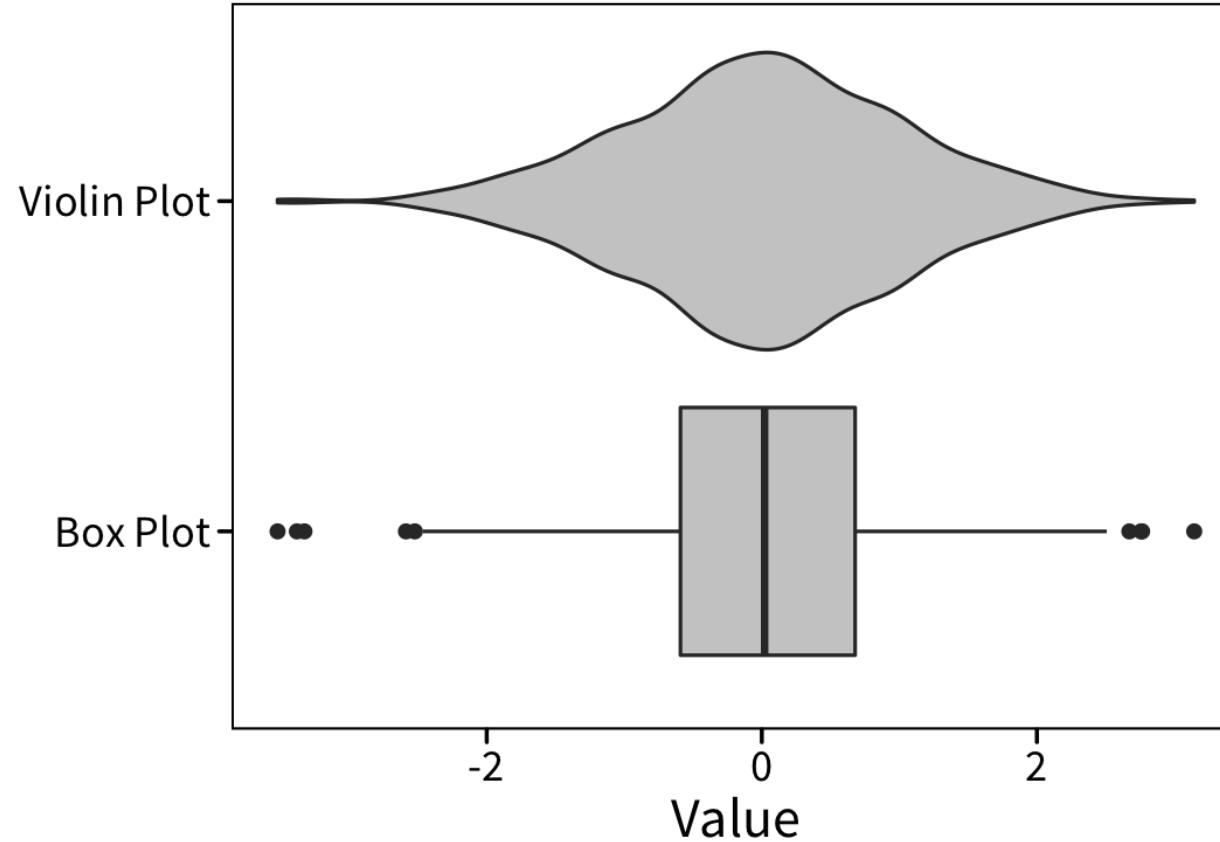
This example: average value of 0 and a
"spread" of 1



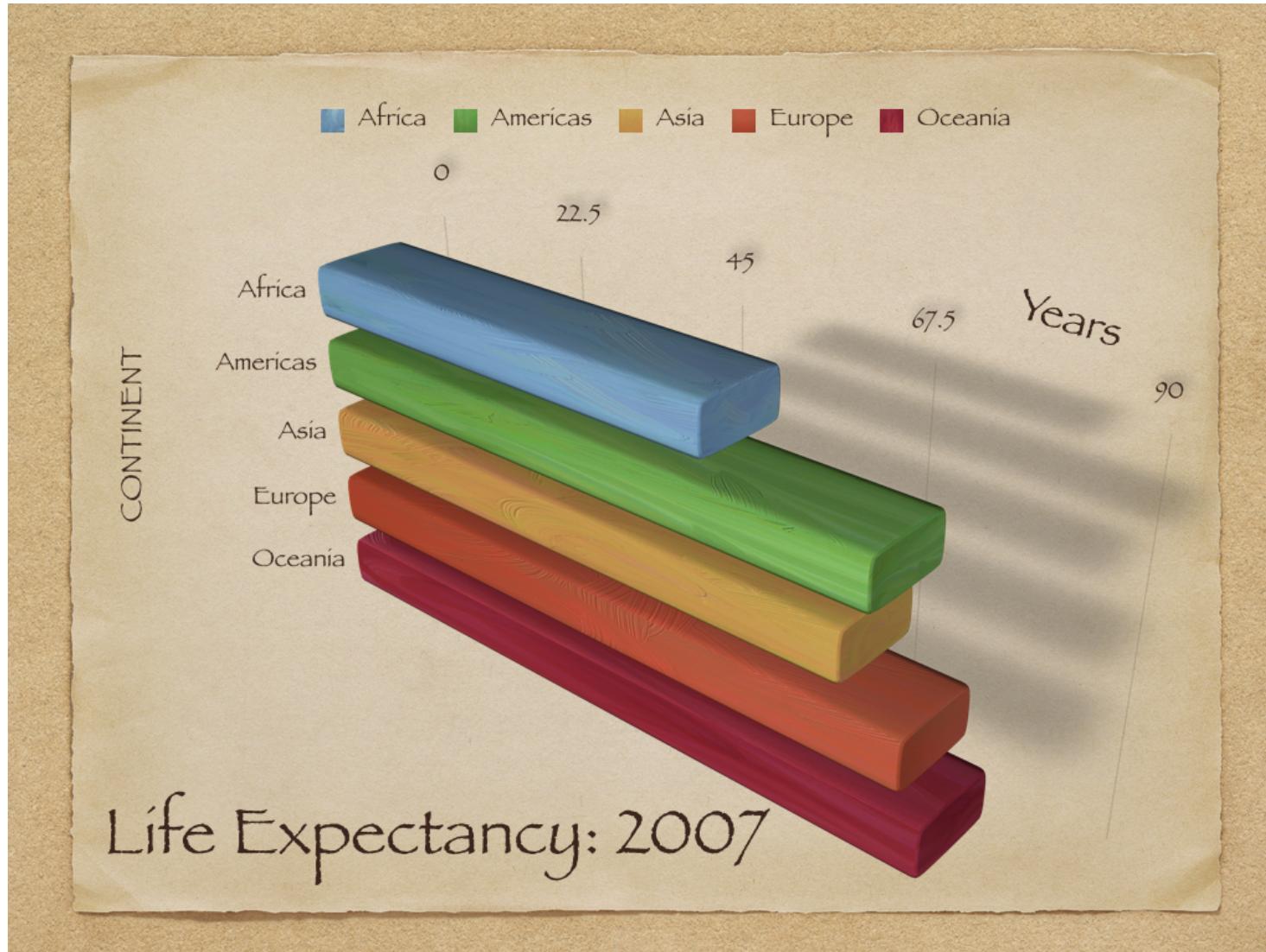
Good ways to plot distributions



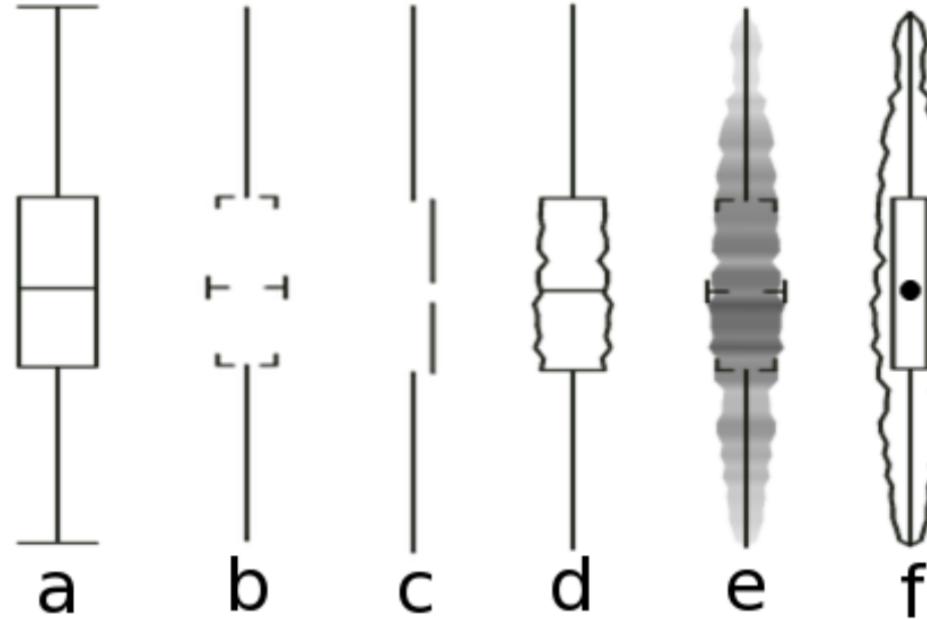
Less good ways to plot distributions



Taste



The "Data-to-Ink Ratio"



Guidelines

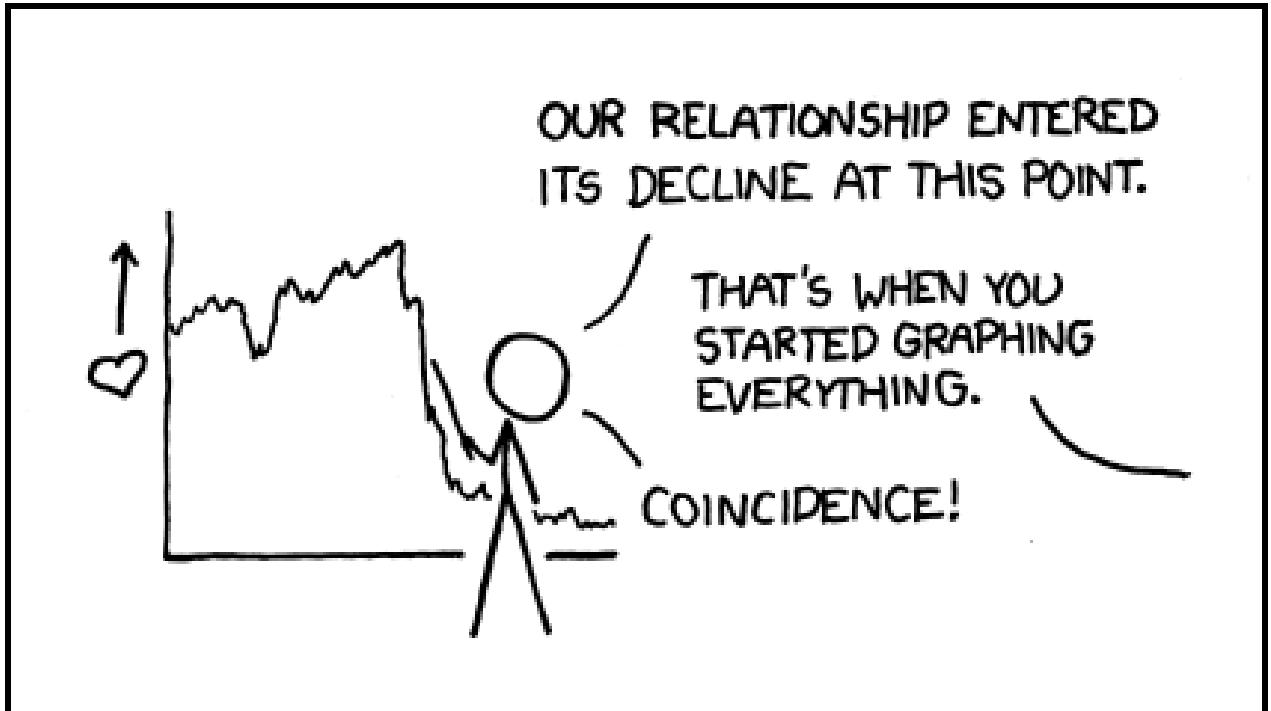
Use graphics to *focus* your point

What to plot:

- distributions
- relationships
- statistical results

Labels go a long way

What goes in the plot vs. in the text?



Looking ahead

On Wednesday: making graphics in R

- the "grammar of graphics" and `ggplot()`
- **Short Essay 1 DUE**
- Assign R Exercise 1

In Section: more `ggplot`