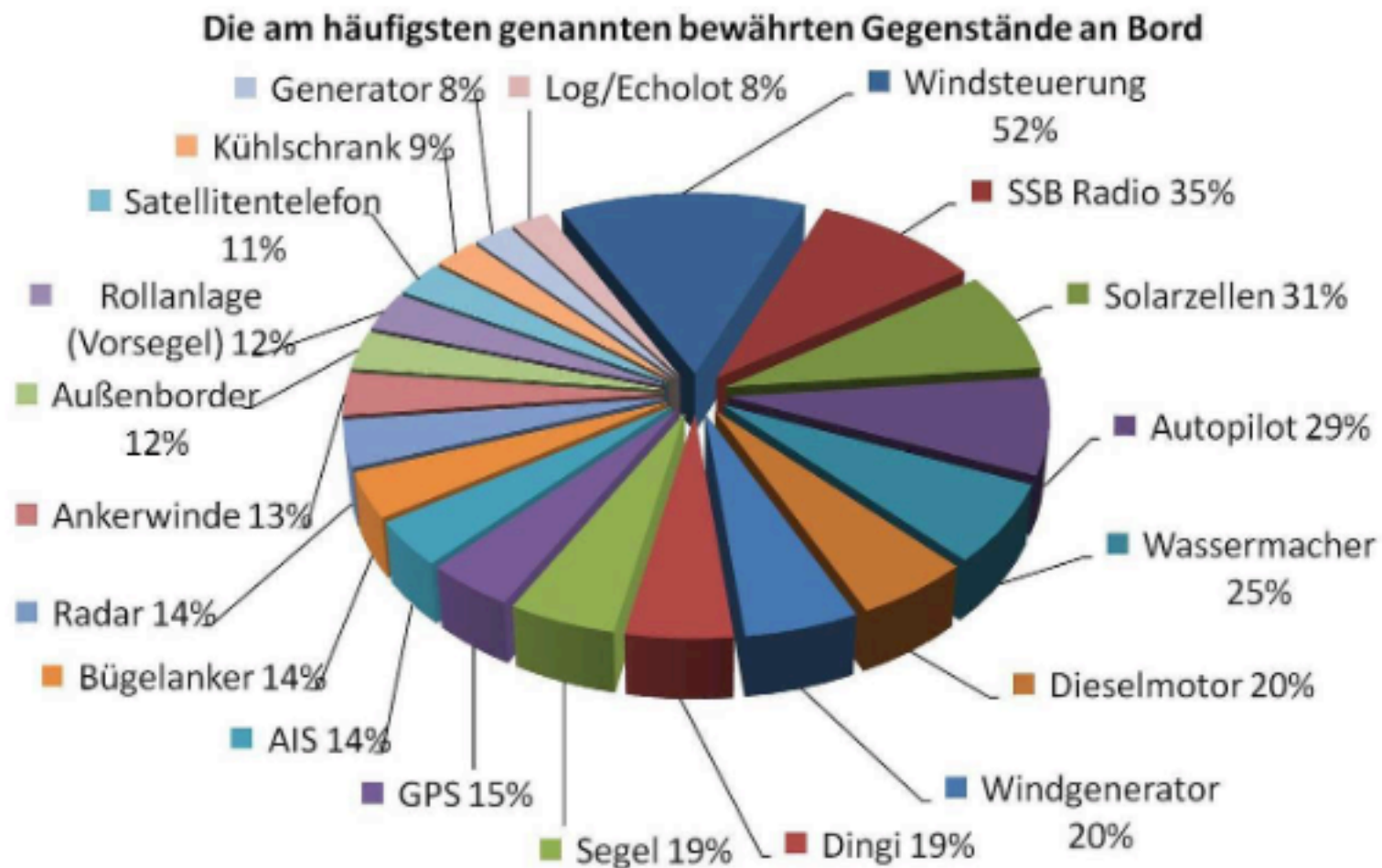


An exploration of dataviz bad practices

An overview of dataviz caveats

[Yan-holtz.com/teaching](https://yan-holtz.com/teaching)

What's wrong with this chart?



Source: [WTF Visualizations](#)

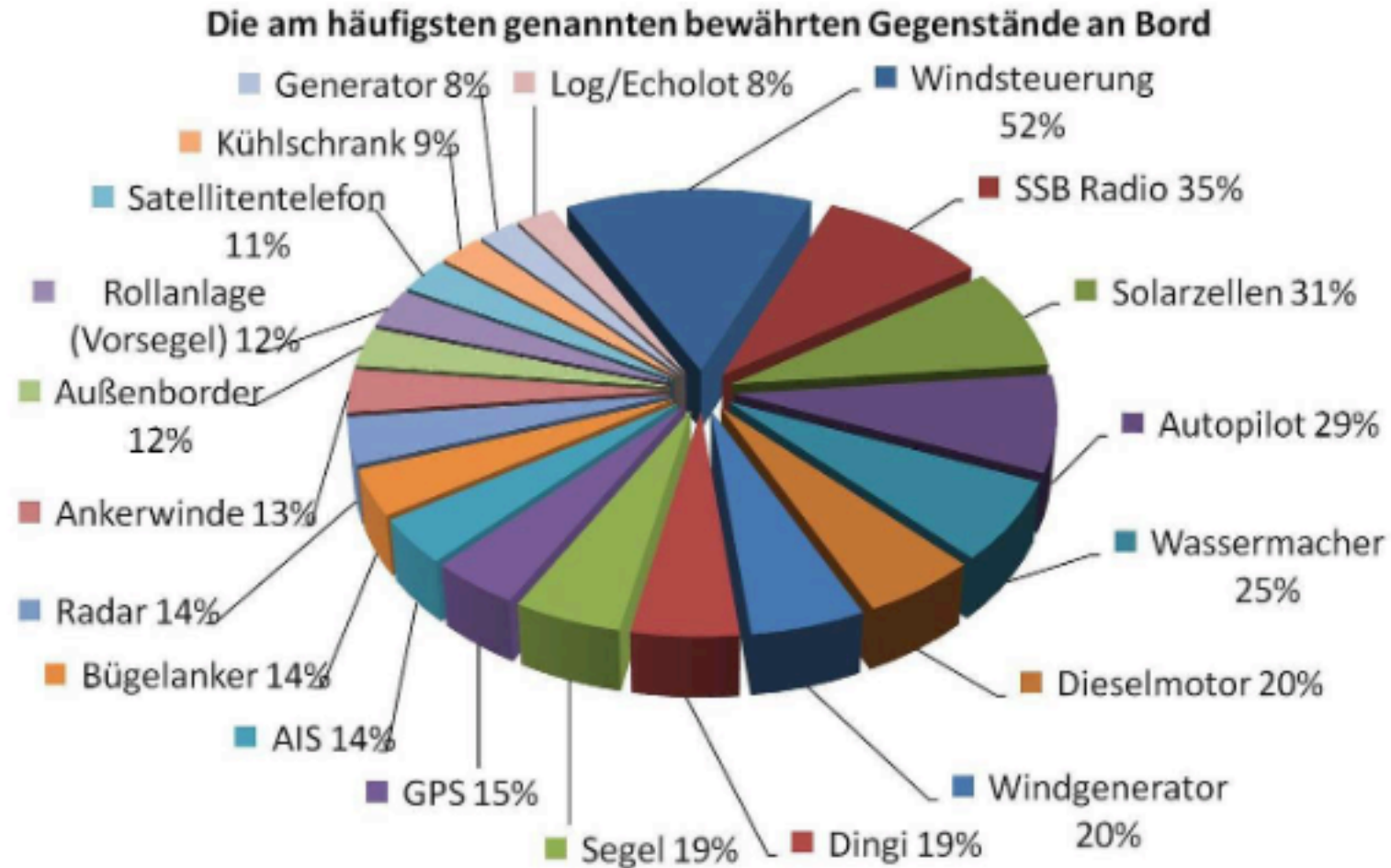
Calculation

+

3D

+

Pie chart



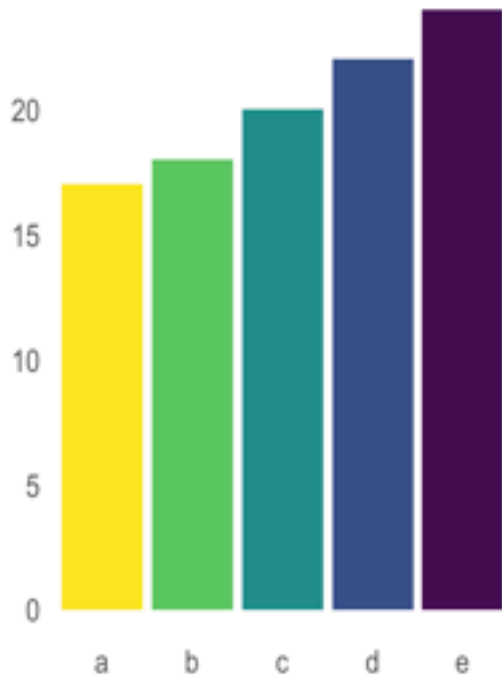
Source: [WTF Visualizations](#)

What's wrong with pie chart?

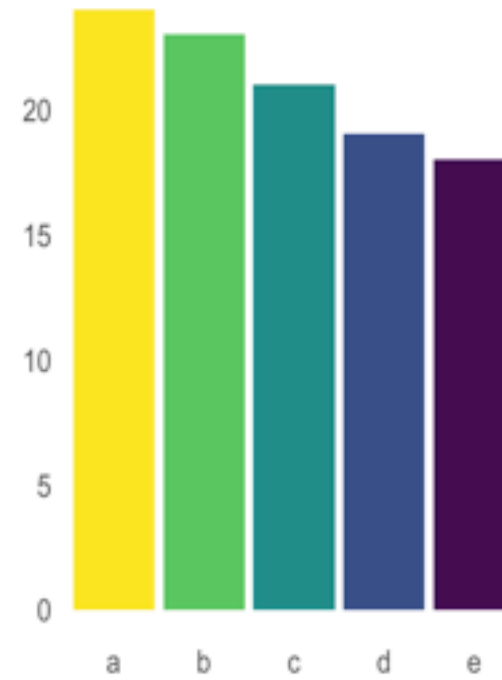
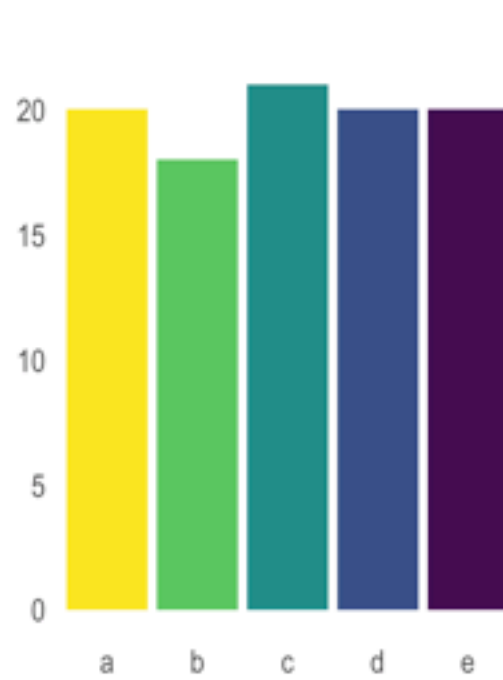


What can you see?

What's wrong with
pie chart?

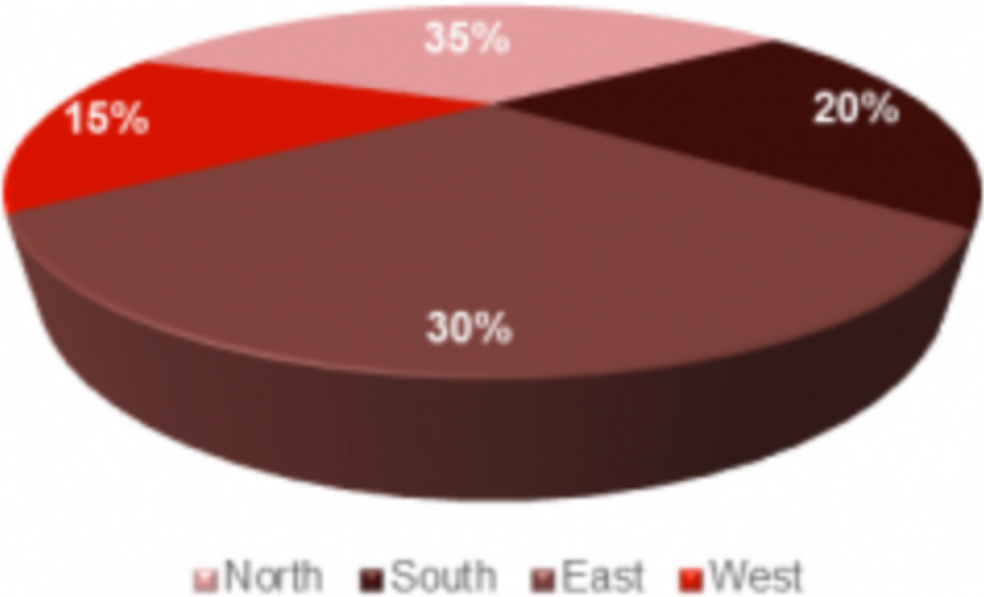


It is hard to
distinguish angles

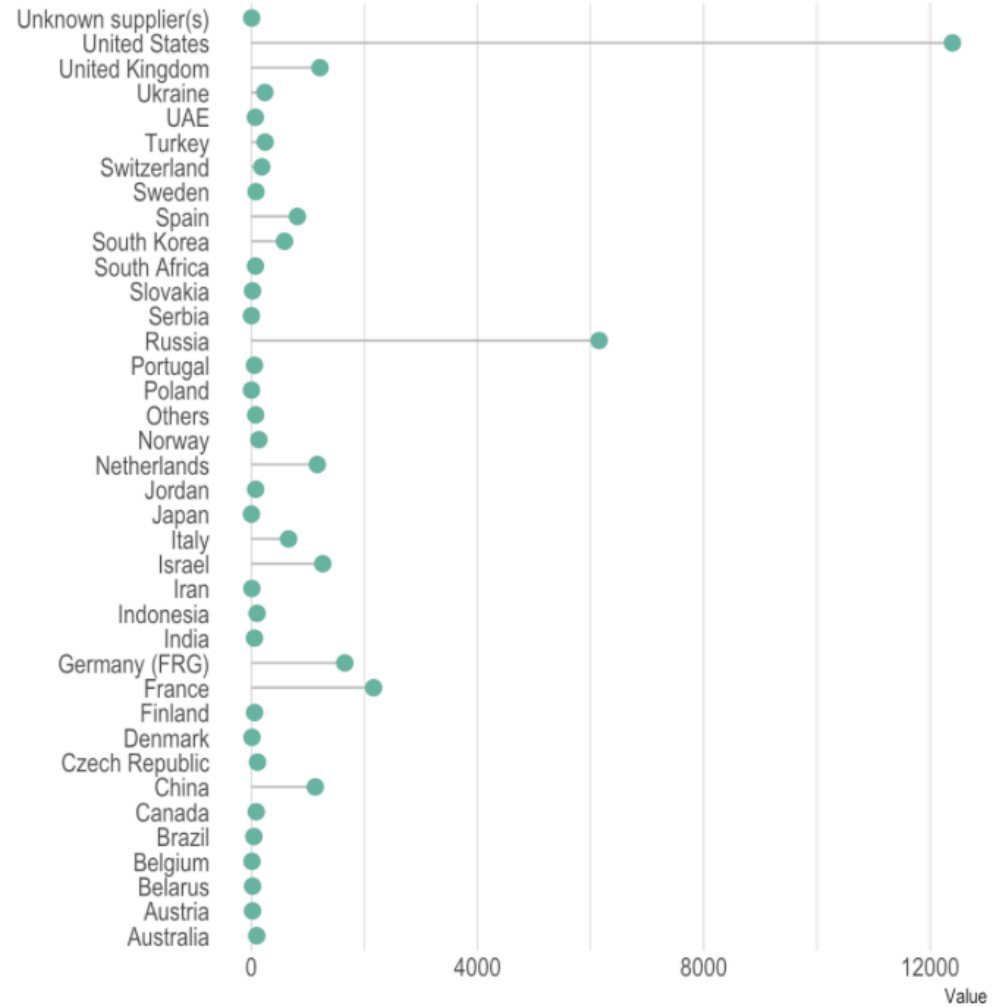


What's wrong with 3D

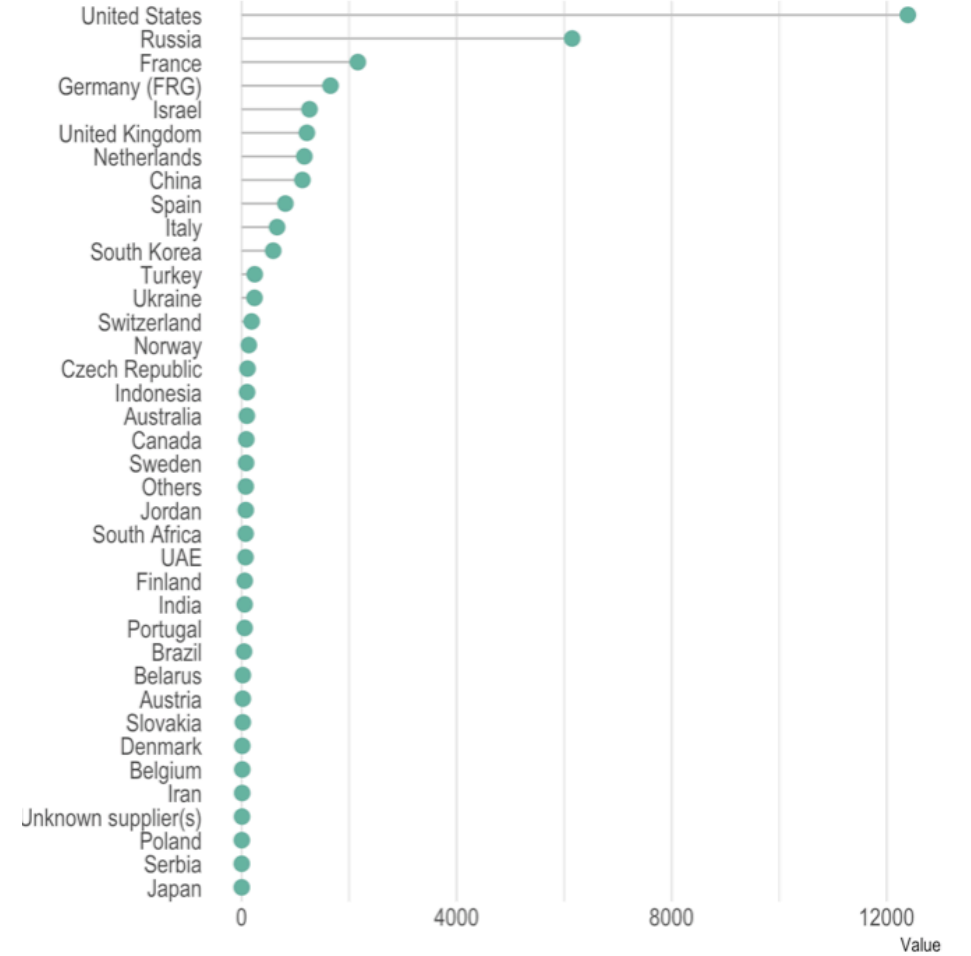
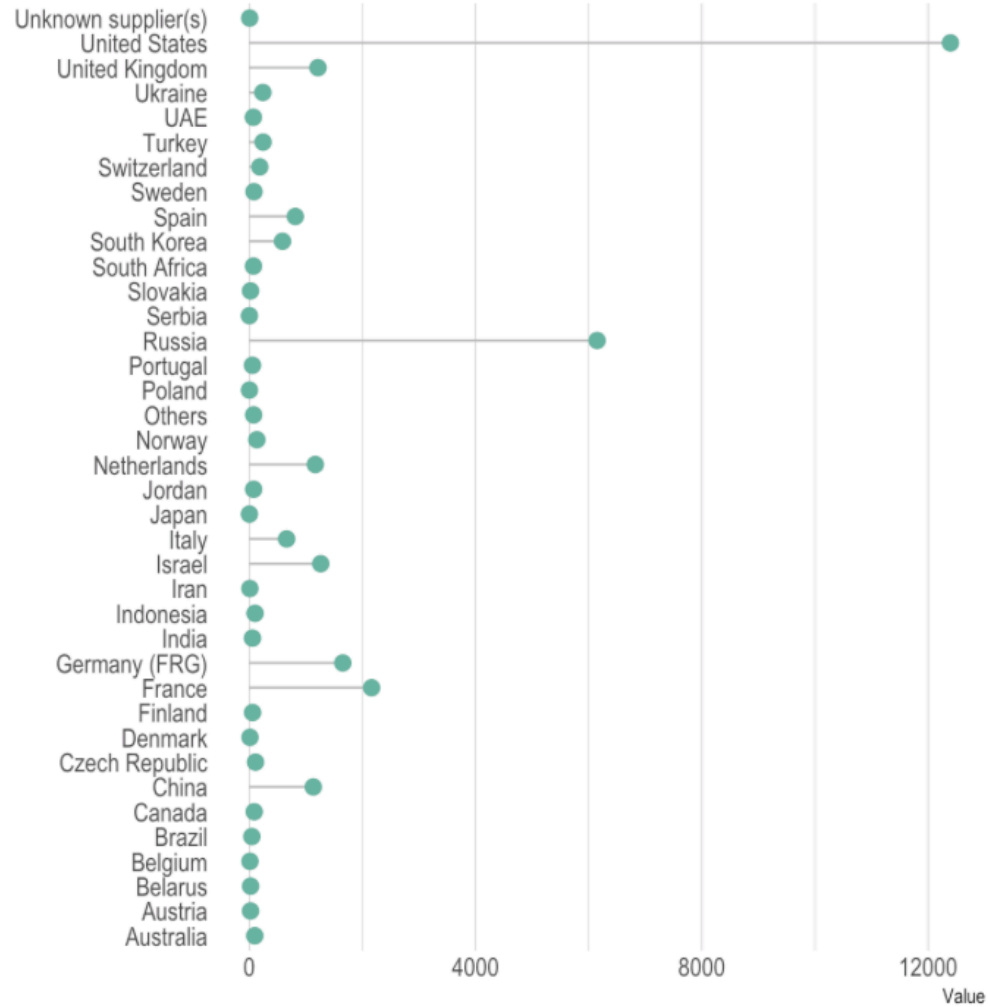
Sales Proportions By Region



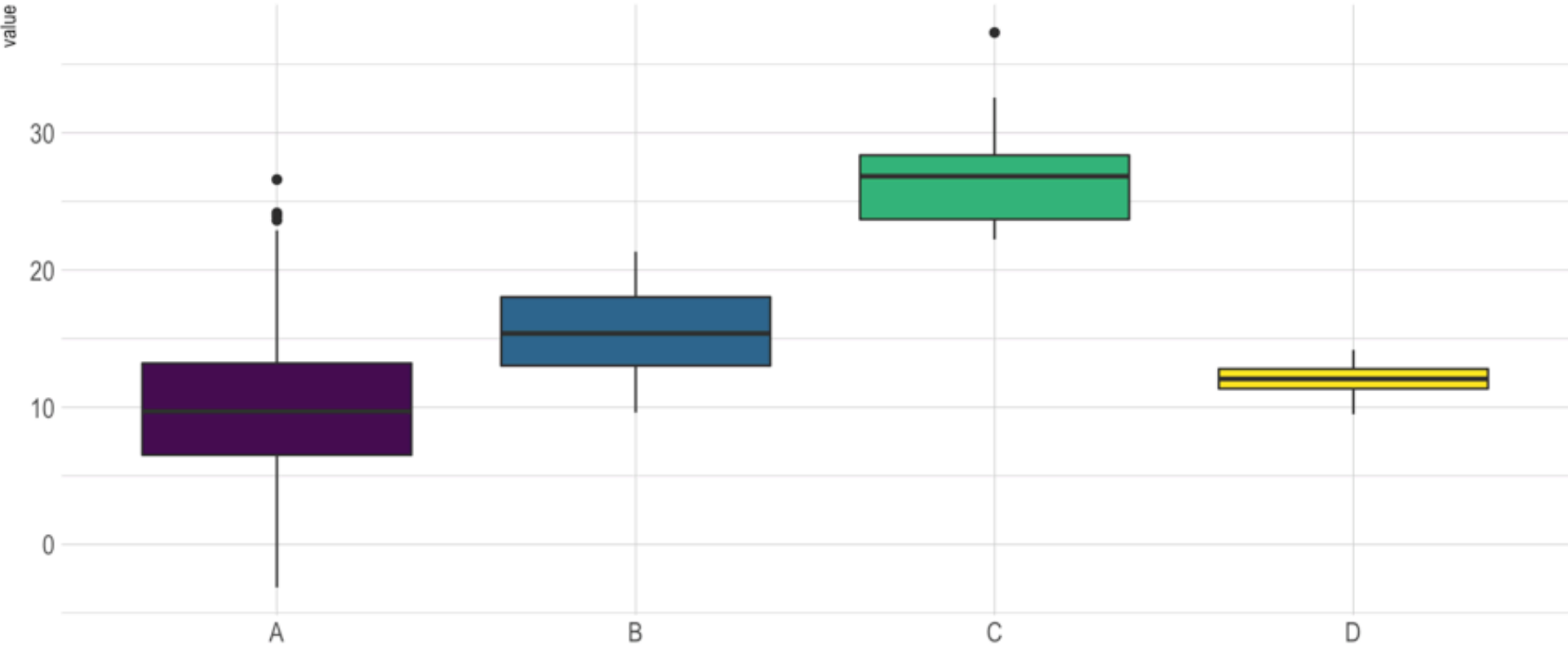
What could be better here?



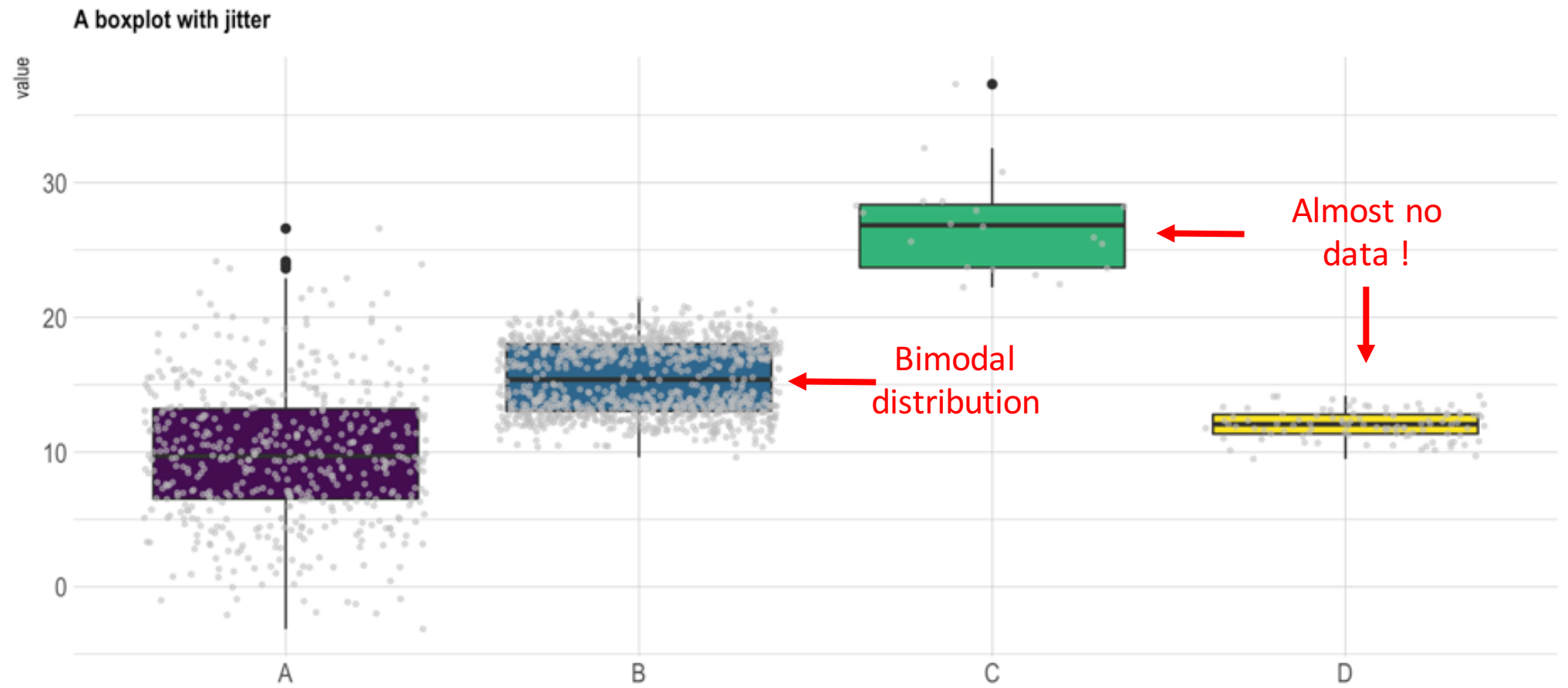
Order your data



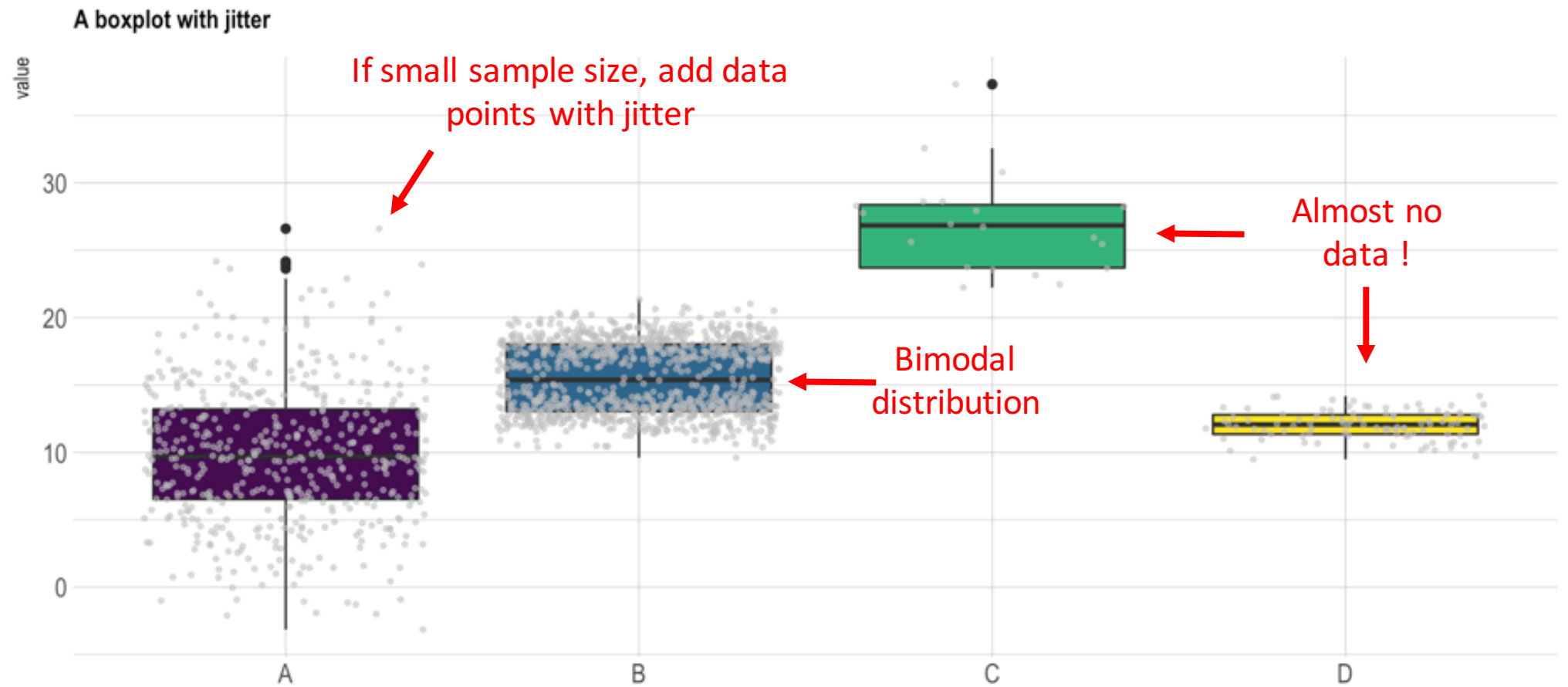
Anything wrong here?



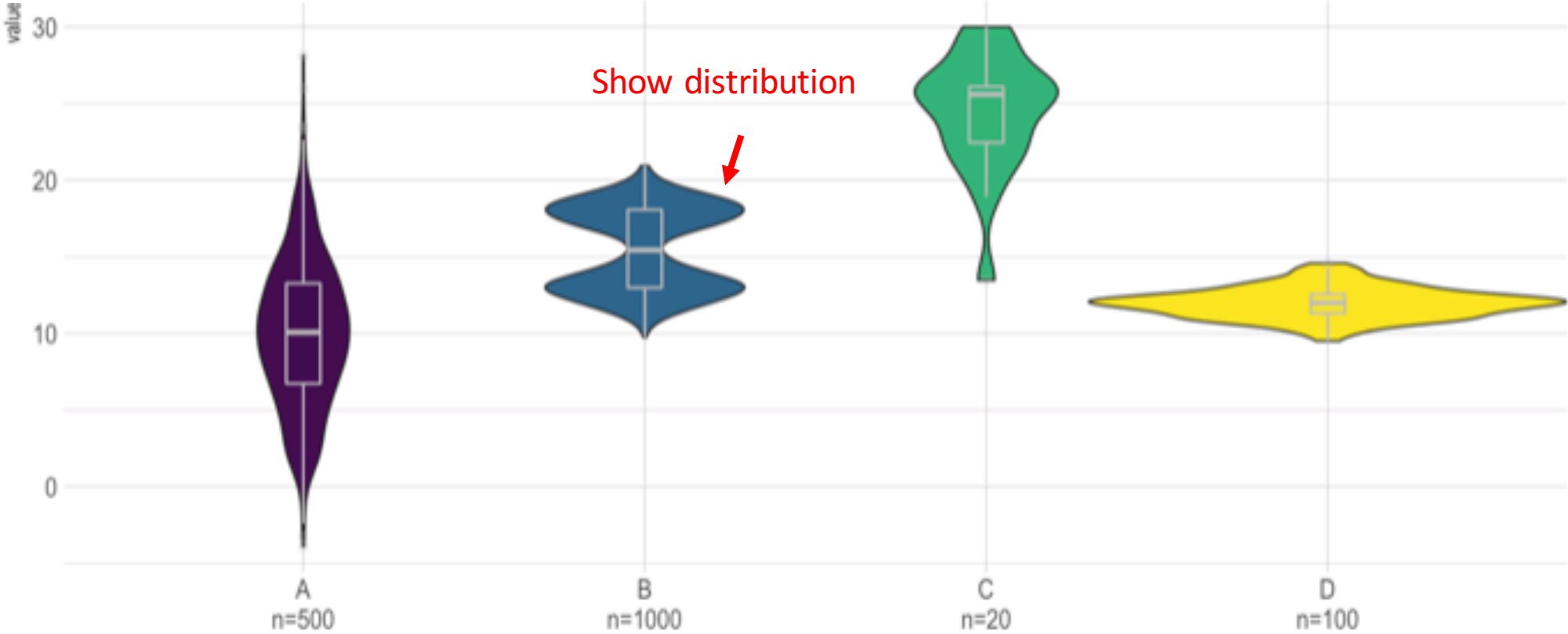
Boxplot = hide information



Boxplot = hide information



If big sample size, use violin plot



Show distribution



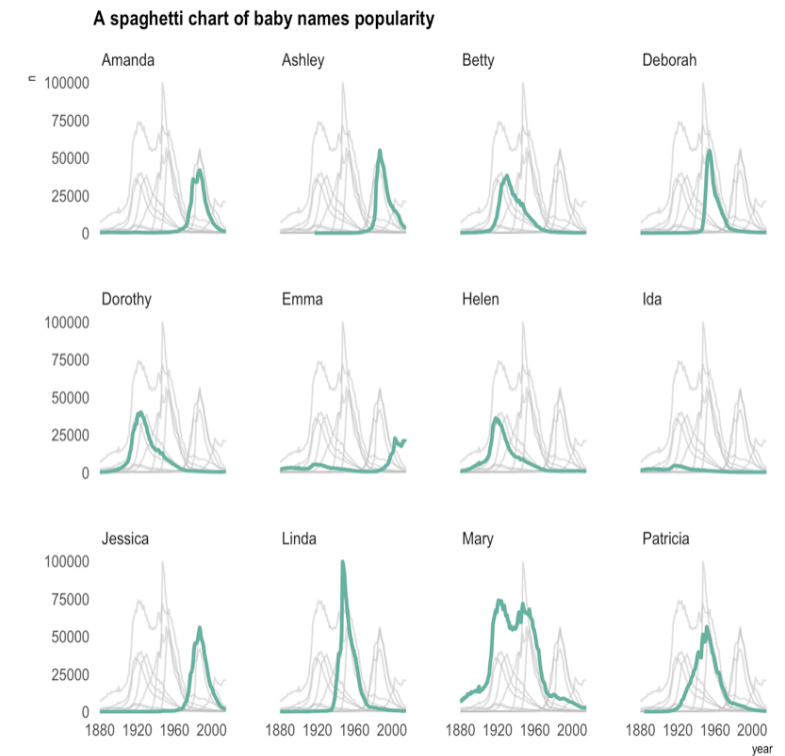
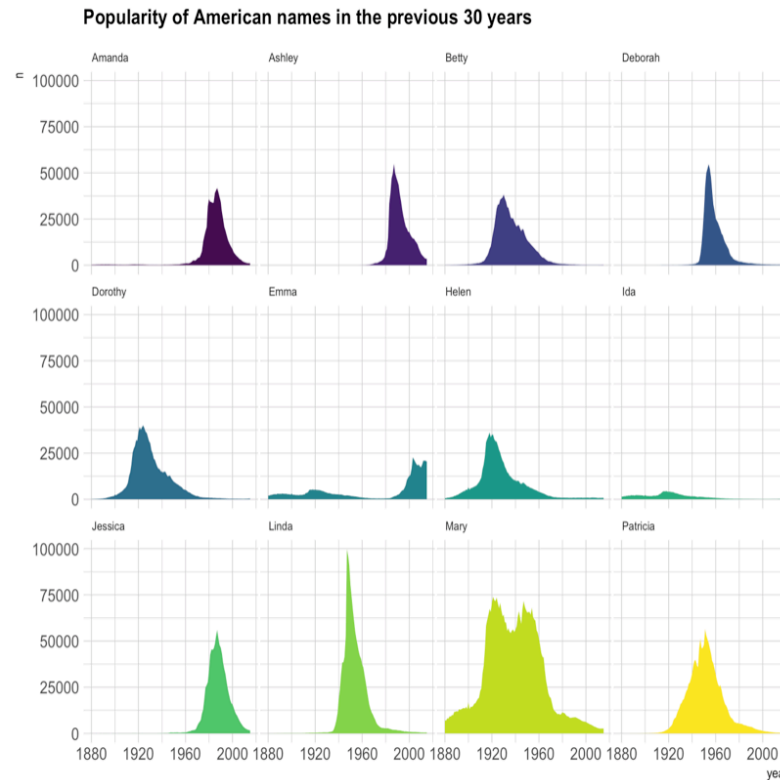
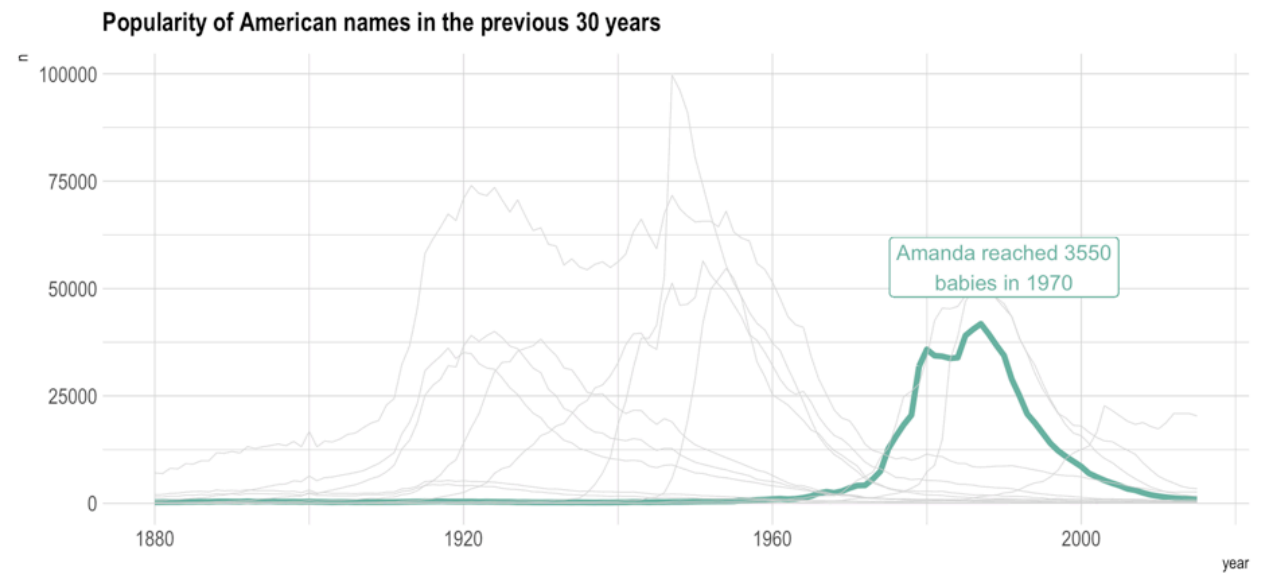
Show sample size



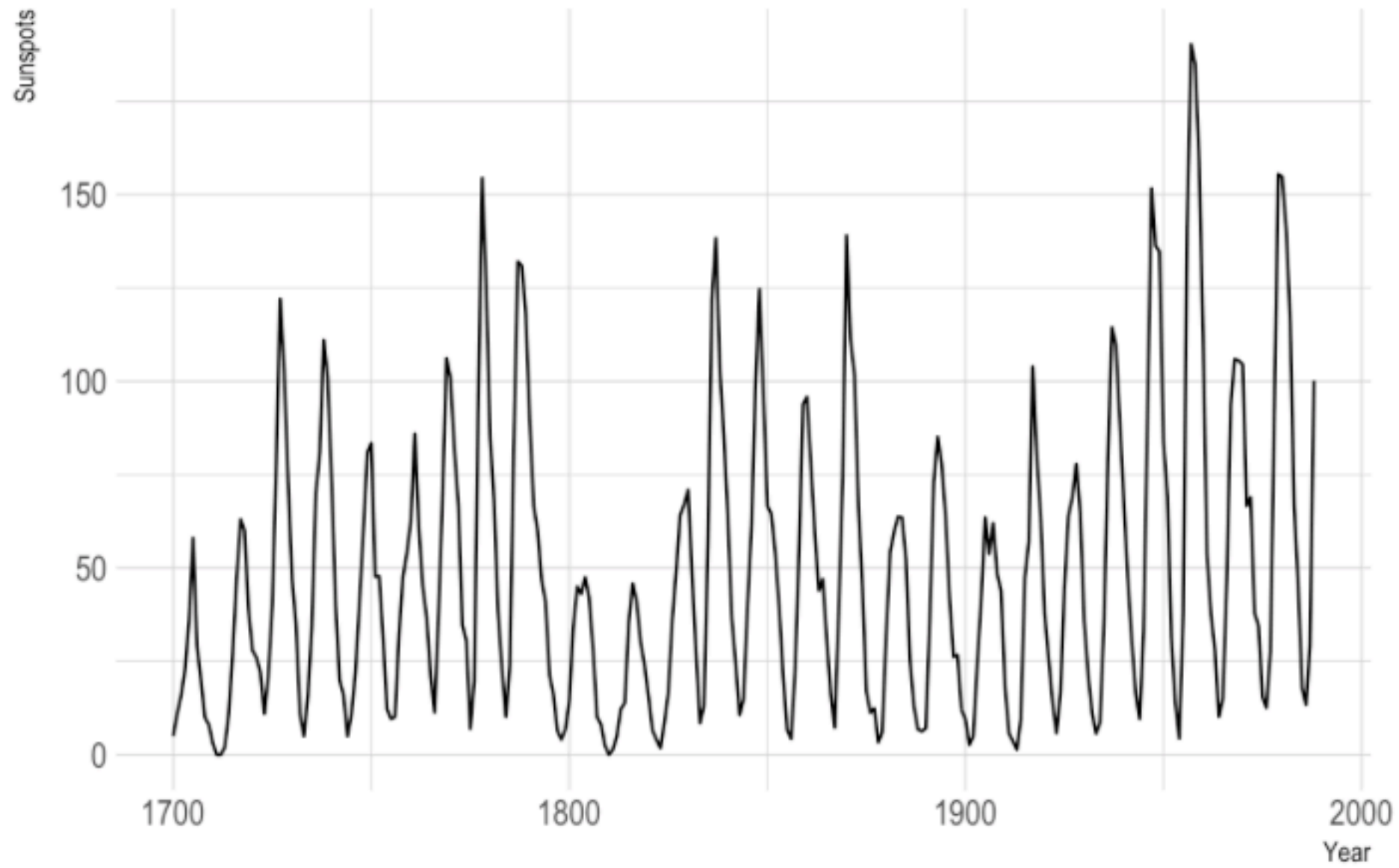
What do you remember?

The Spaghetti chart

[Read more](#)

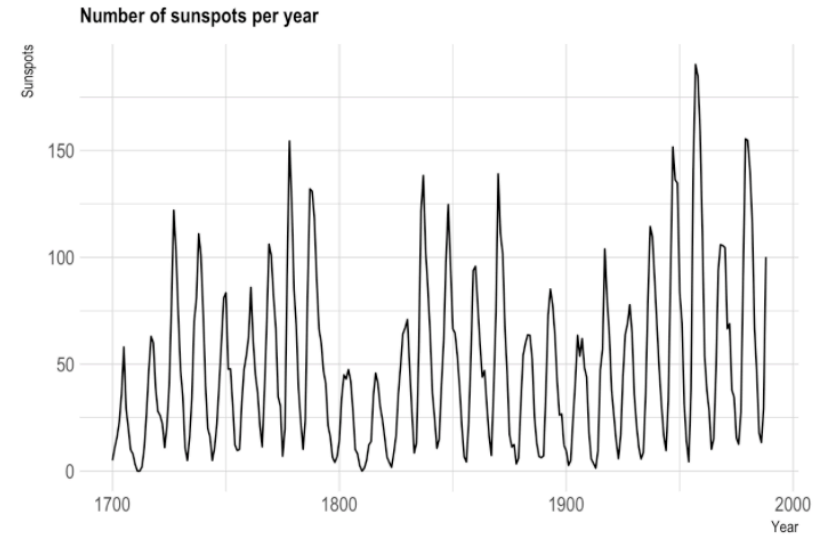


Number of sunspots per year

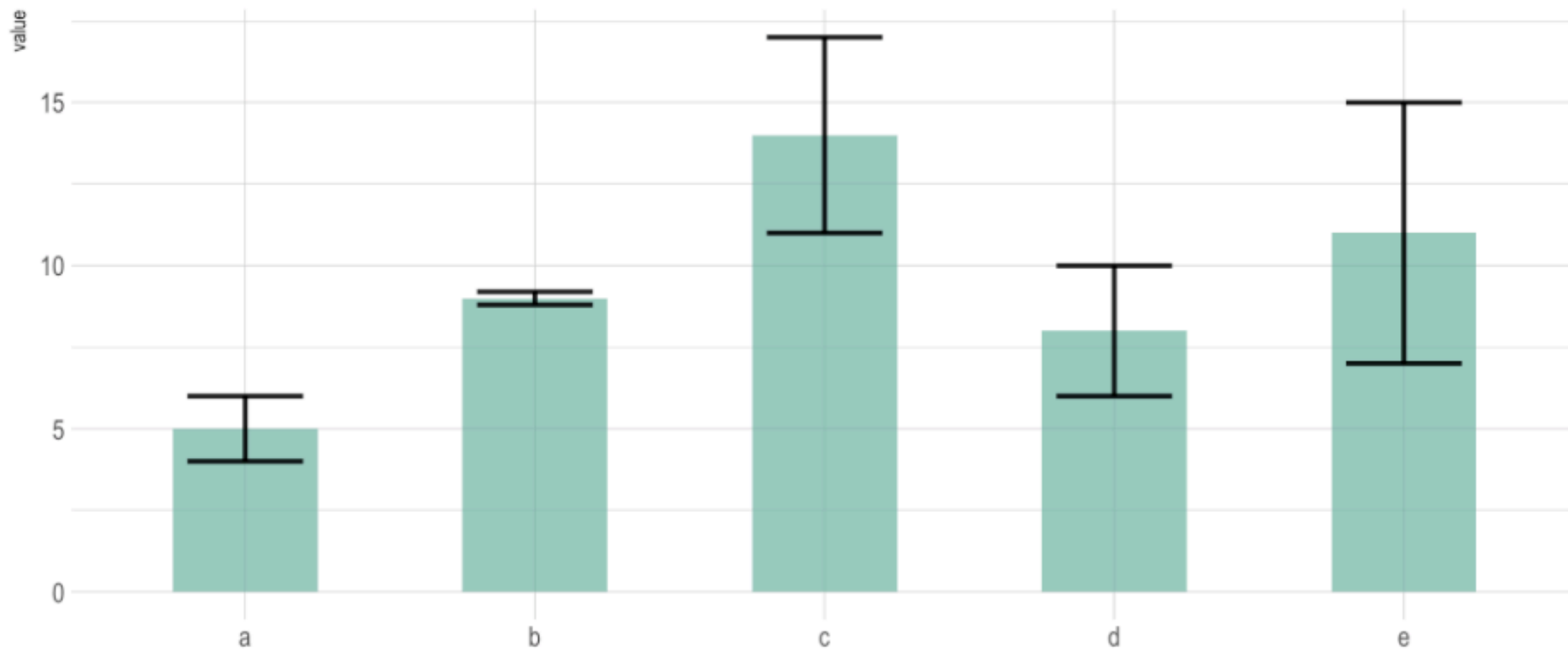


Mind the aspect ratio

[Read more](#)

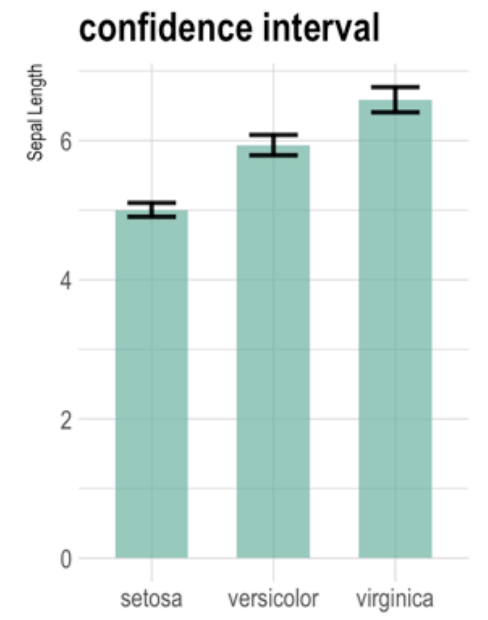
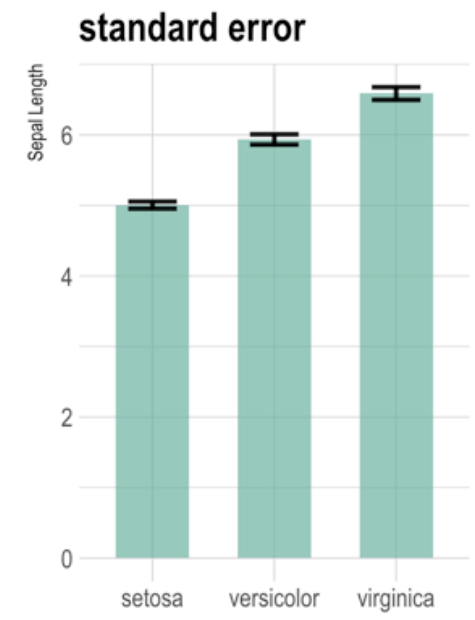
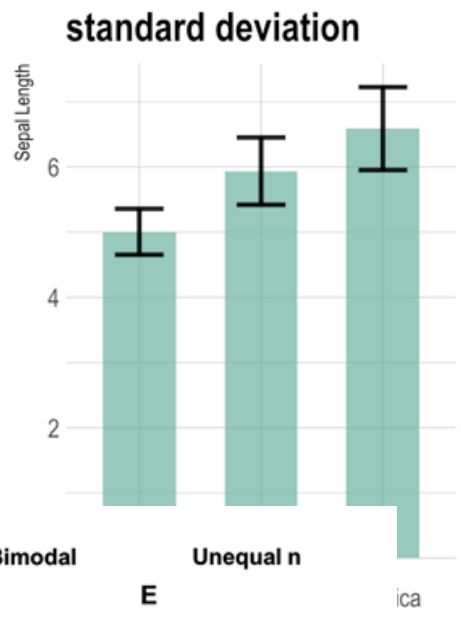
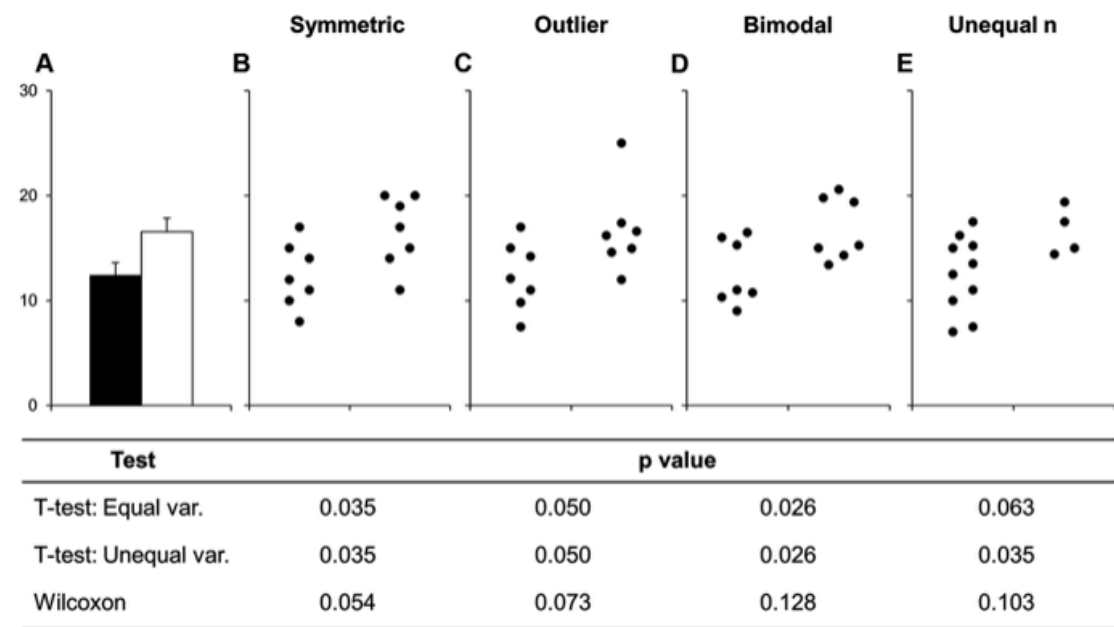


A barplot with error bar



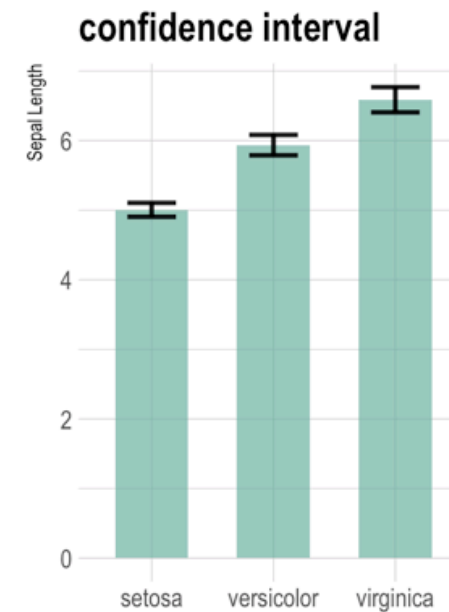
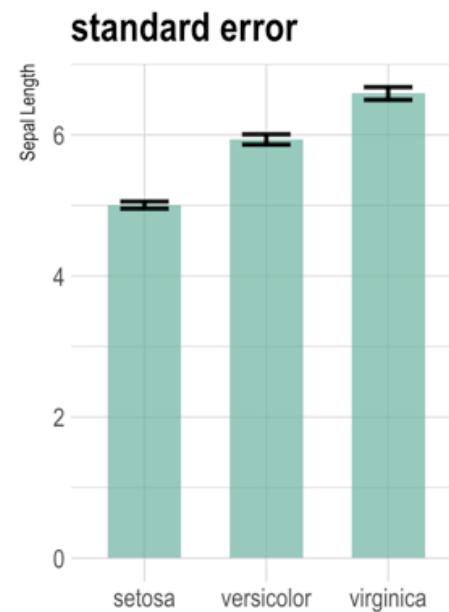
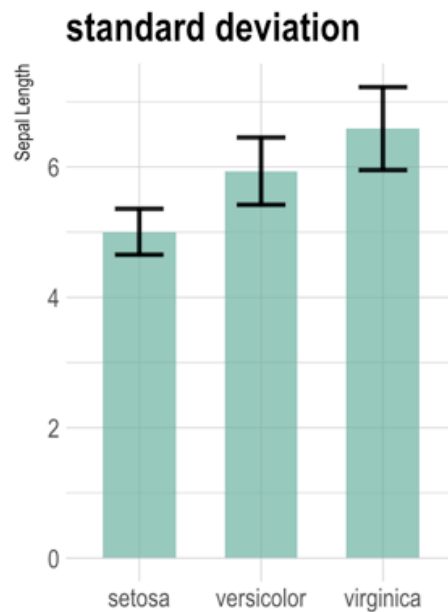
The issue with error bars

[Read more](#)



The issue with error bars

[Read more](#)





AMERICA'S ECONOMY

2010 GROSS DOMESTIC PRODUCT



United States
\$14.6 TRILLION



China \$5.7 TRILLION



Japan \$5.3 TRILLION

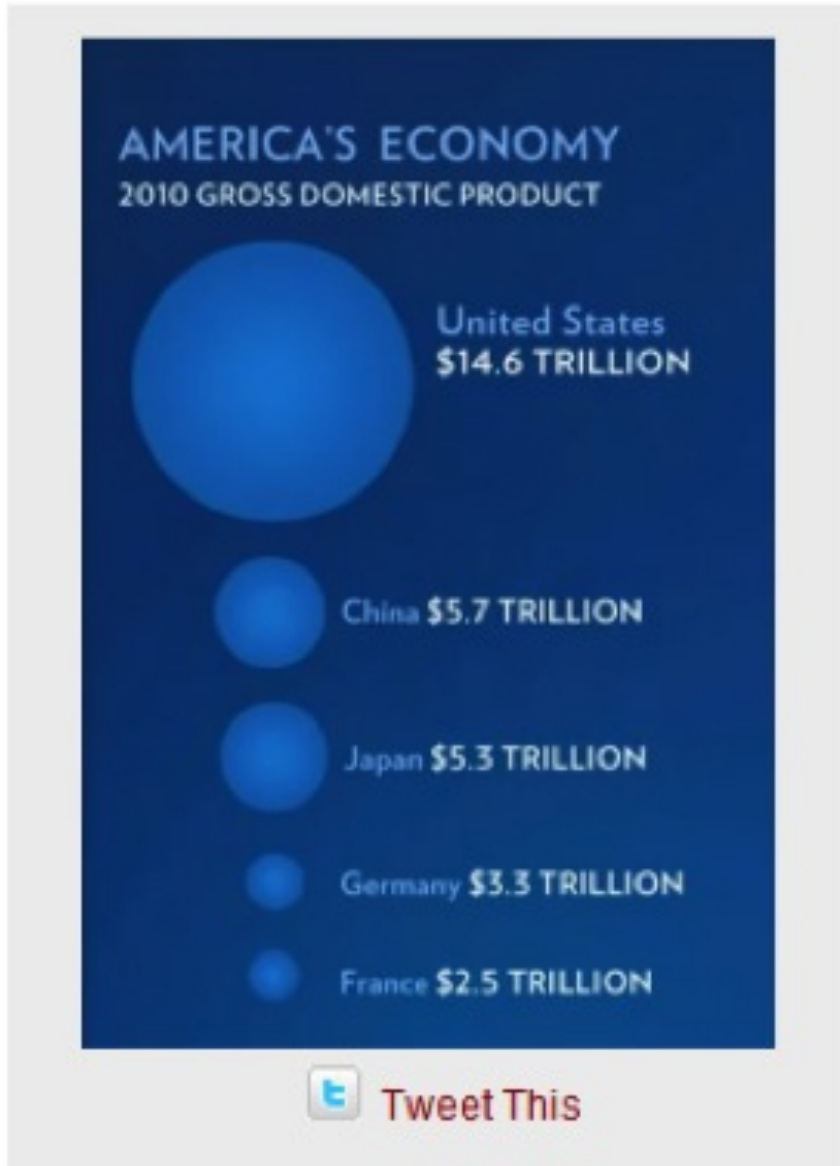


Germany \$3.3 TRILLION

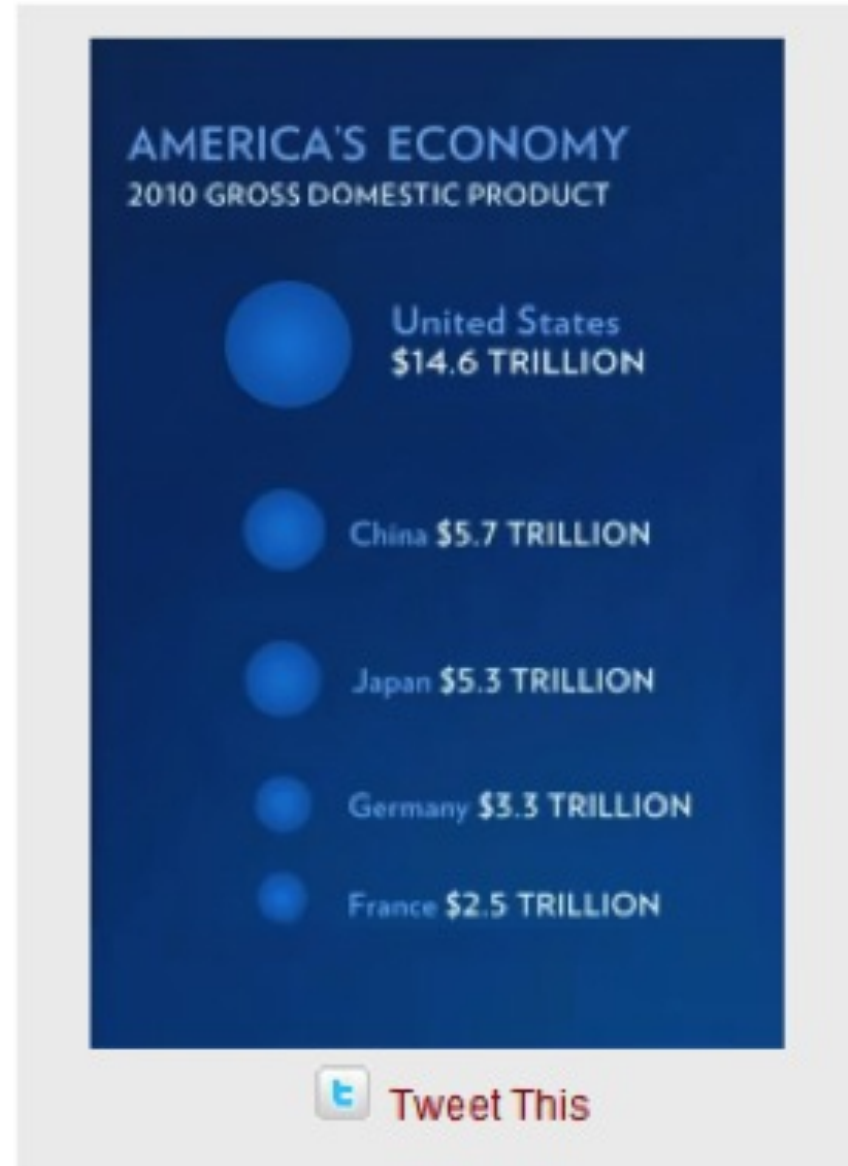


France \$2.5 TRILLION

Size = radius



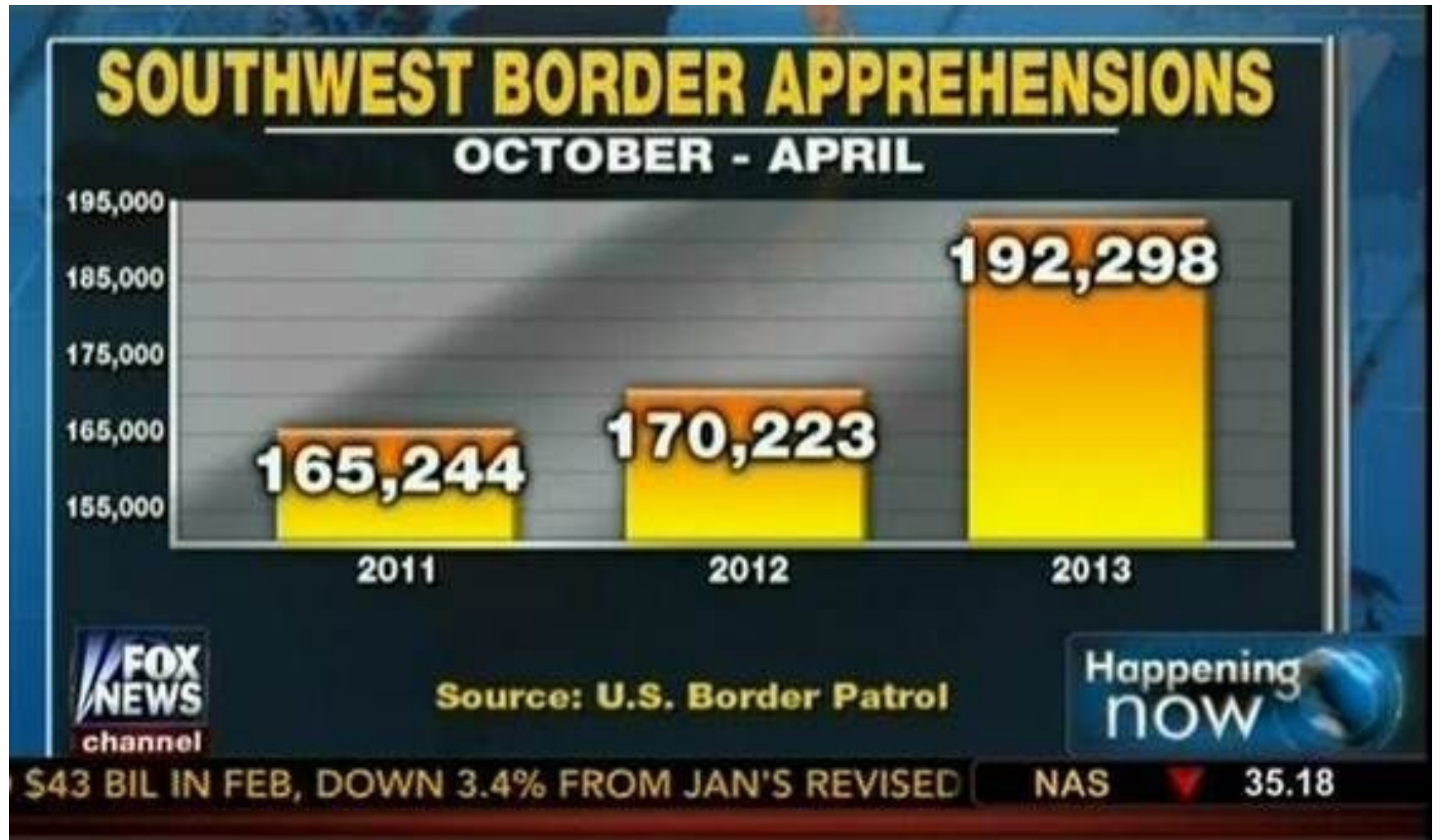
Size = area



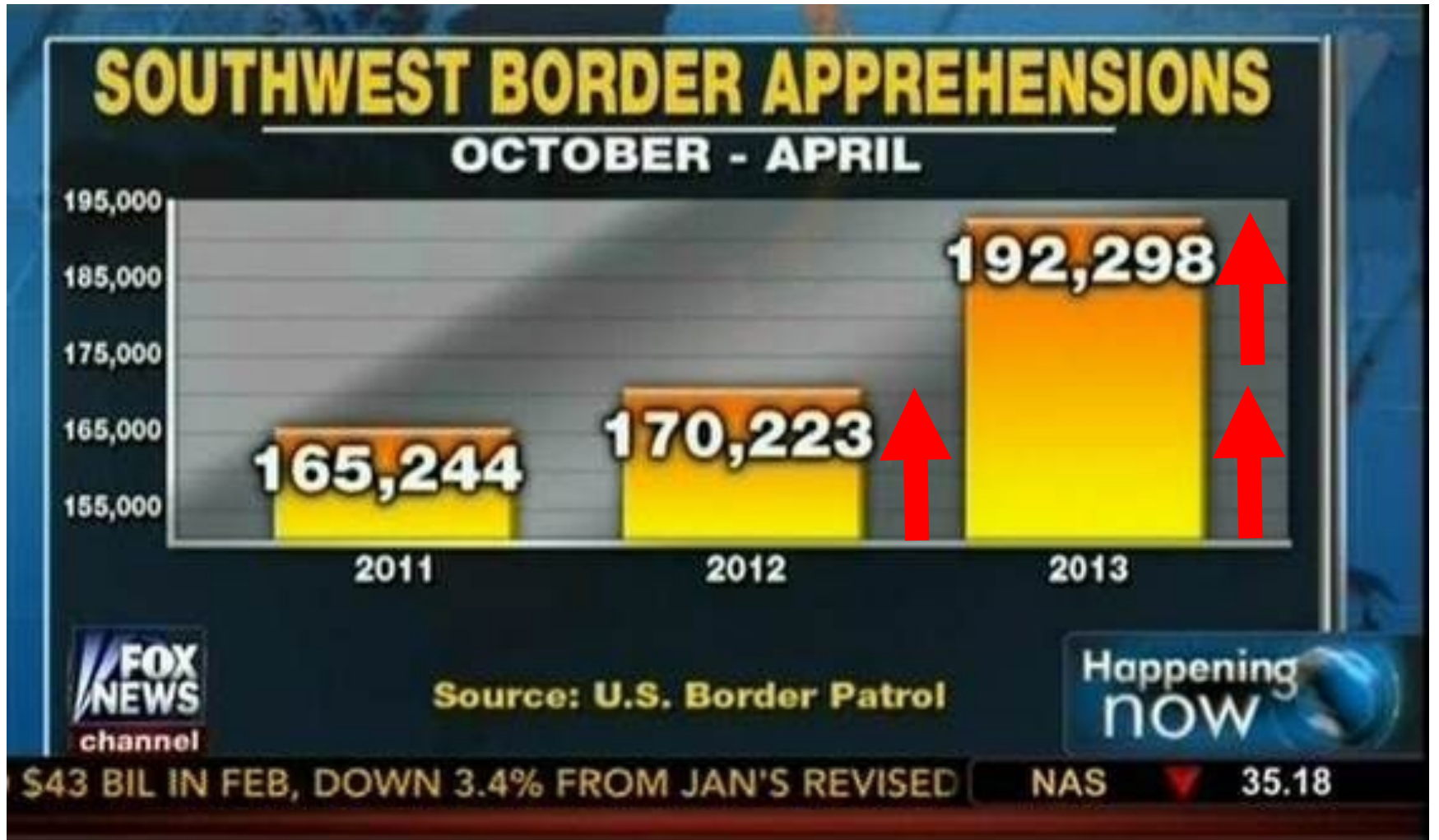
Size = radius

Size = area





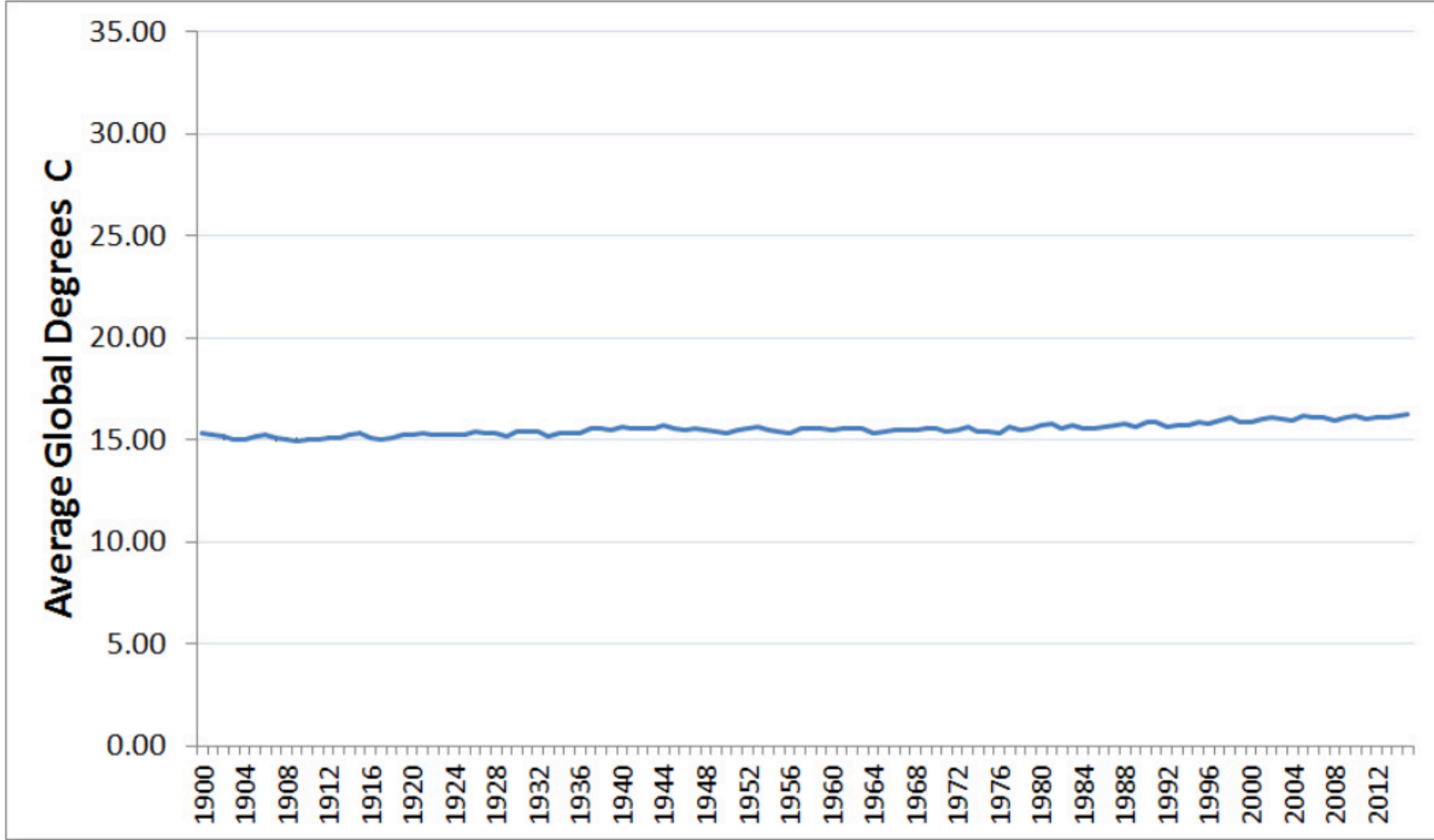
Source: Fox News, via [Media Matters for America](#)

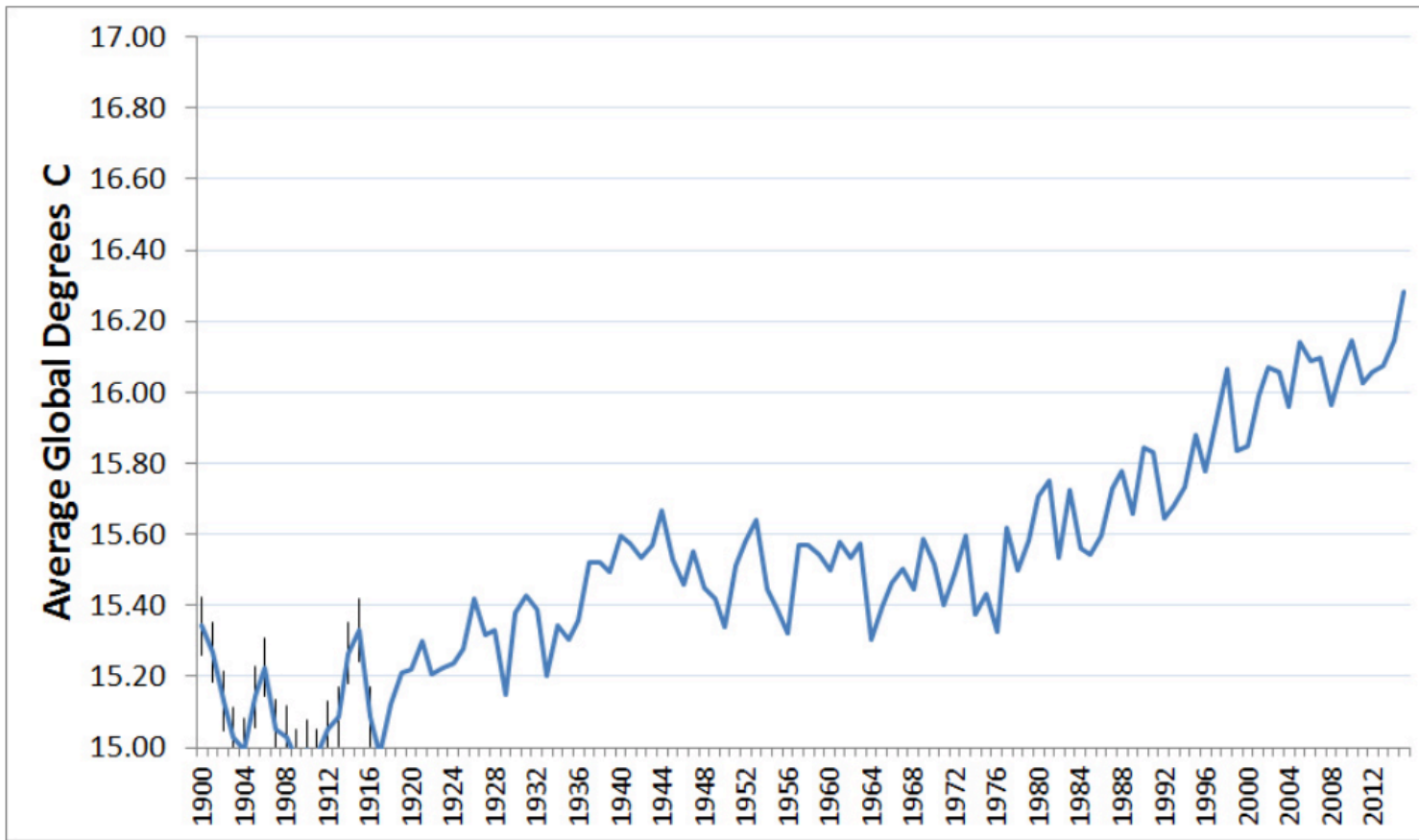


Obama
administration

=
2x more
apprehensions?

Source: Fox News, via [Media Matters for America](#)

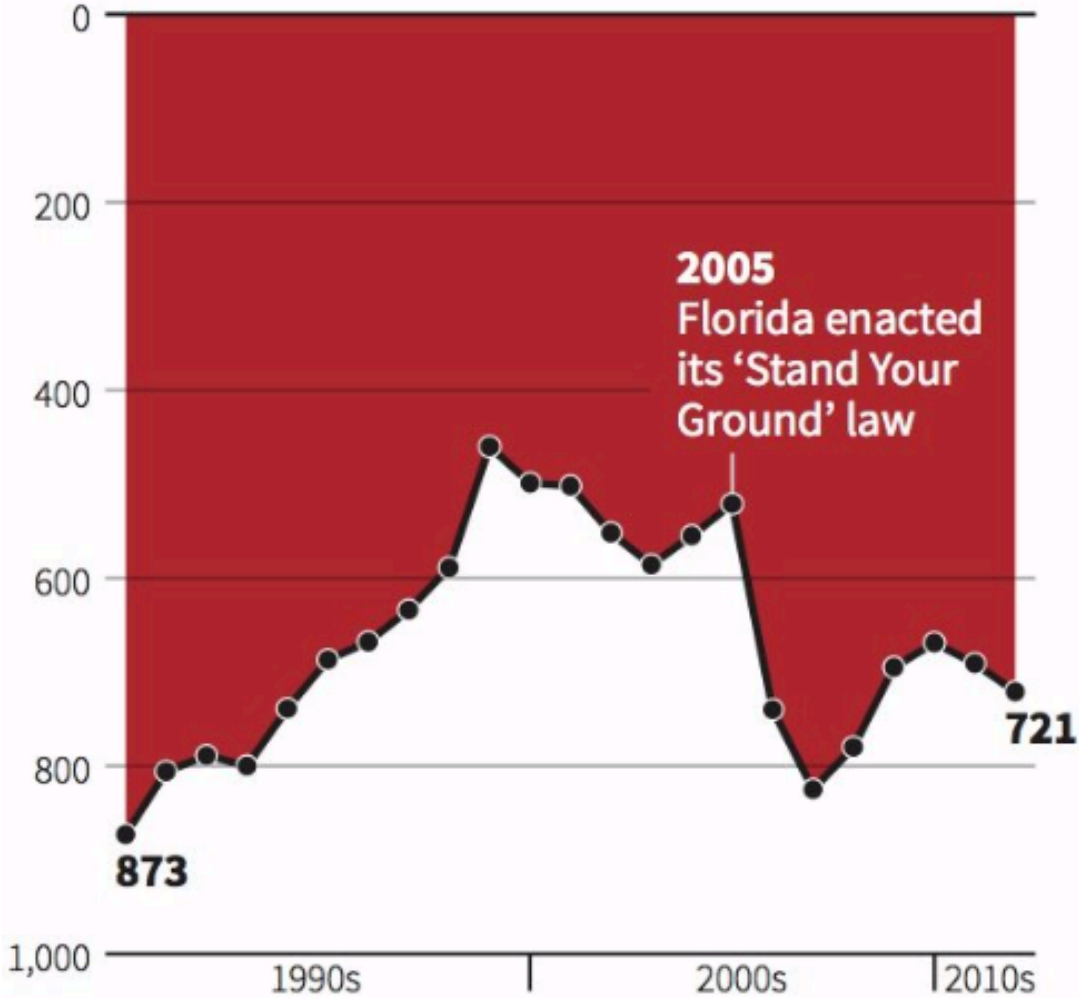




“In general, in a time-series, use a baseline that shows the data not the zero point” - [Edward Tufte](#)

Gun deaths in Florida

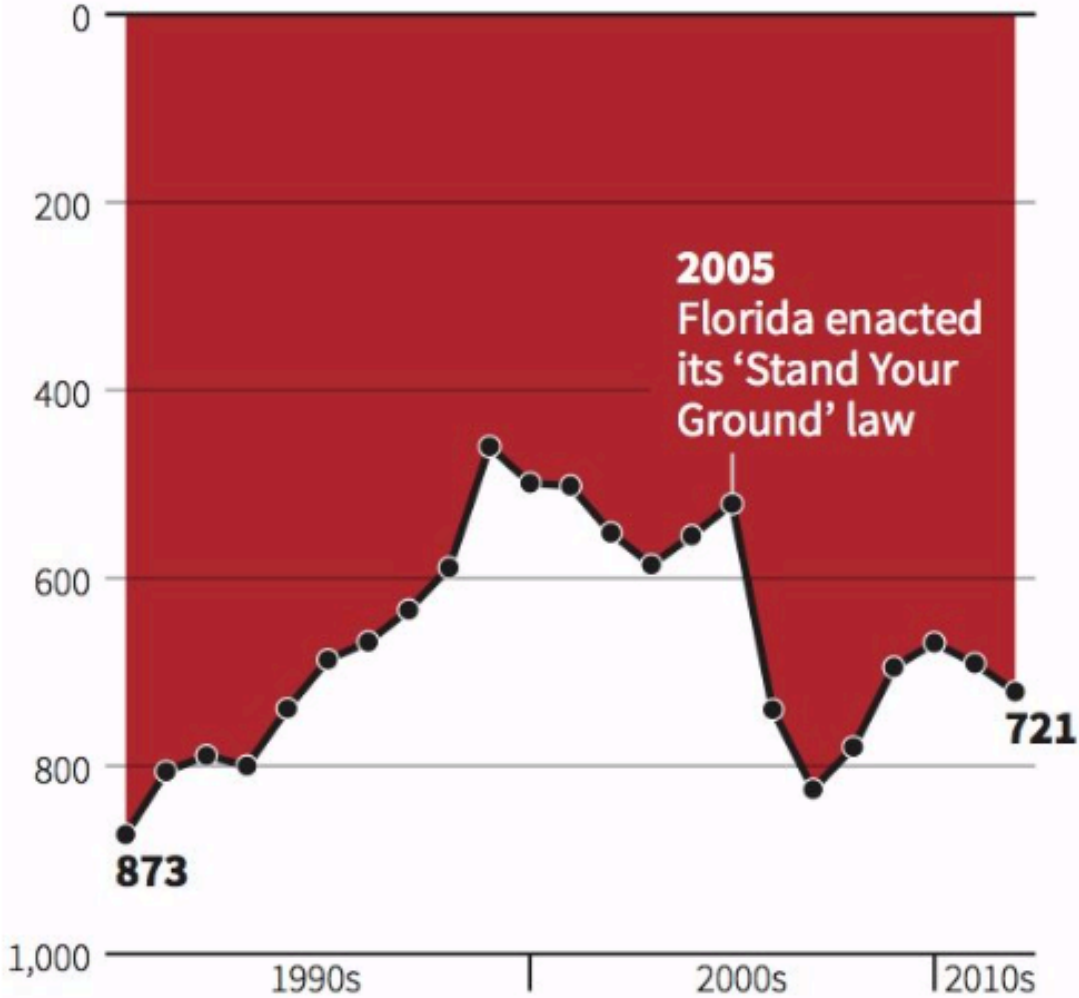
Number of murders committed using firearms



Source: Florida Department of Law Enforcement

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

Do not be counter intuitive

Data to Viz

A collection of dataviz caveats



Order your data

When displaying the value of several entities, ordering them makes the graph much more insightful.



To cut or not to cut?

Cutting the Y-axis is one of the most controversial practice in data viz. See why.



The spaghetti chart

A line graph with too many lines becomes unreadable: it is called a spaghetti graph.



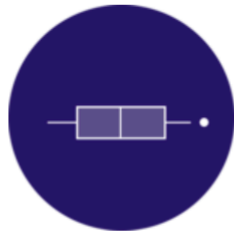
Pie chart

The human eye is bad at reading angles. See how to replace the most criticized chart ever.



Play with histogram bin size

Always try different bin sizes when you build a histogram, it can lead to different insights.



Do boxplots hide information?

Boxplots are a great way to summarize a distribution but hide the sample size and their distribution.



The problem with error bars

Barplots with error bars must be used with great care. See why and how to replace them.



Too many distributions.

If you need to compare the distributions of many variables, don't clutter your graphic.

Thanks

Slides:

[Yan-holtz/teaching](https://github.com/Yan-holtz/teaching)

Data to viz:

Data-to-viz.com



Yan.holtz.data@gmail.com



www.yan-holtz.com



[@R_Graph_Gallery](https://twitter.com/R_Graph_Gallery)



github.com/holtzy