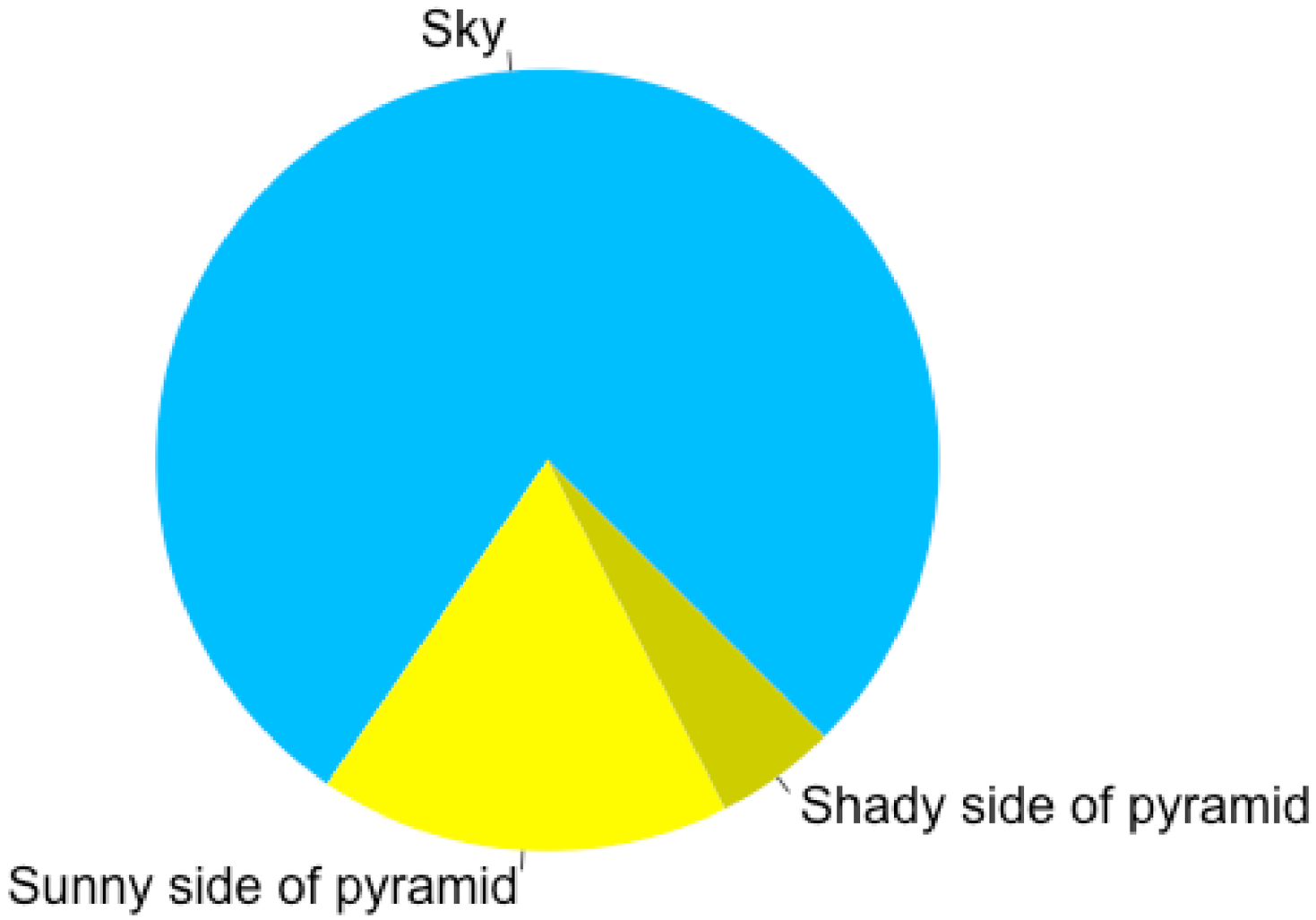
Polar coordinates

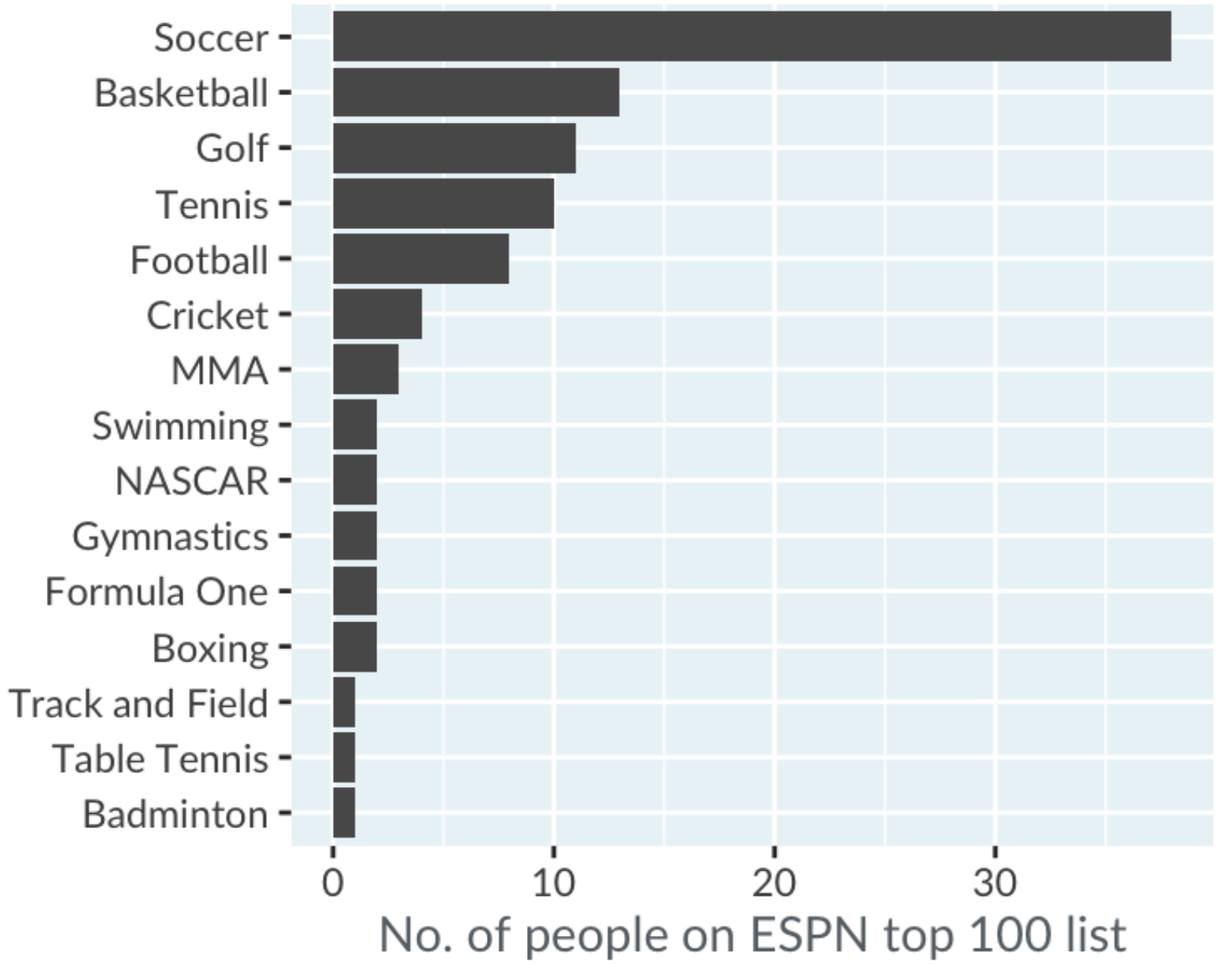
UNDERSTANDING DATA VISUALIZATION

Richie Cotton
Data Evangelist at DataCamp

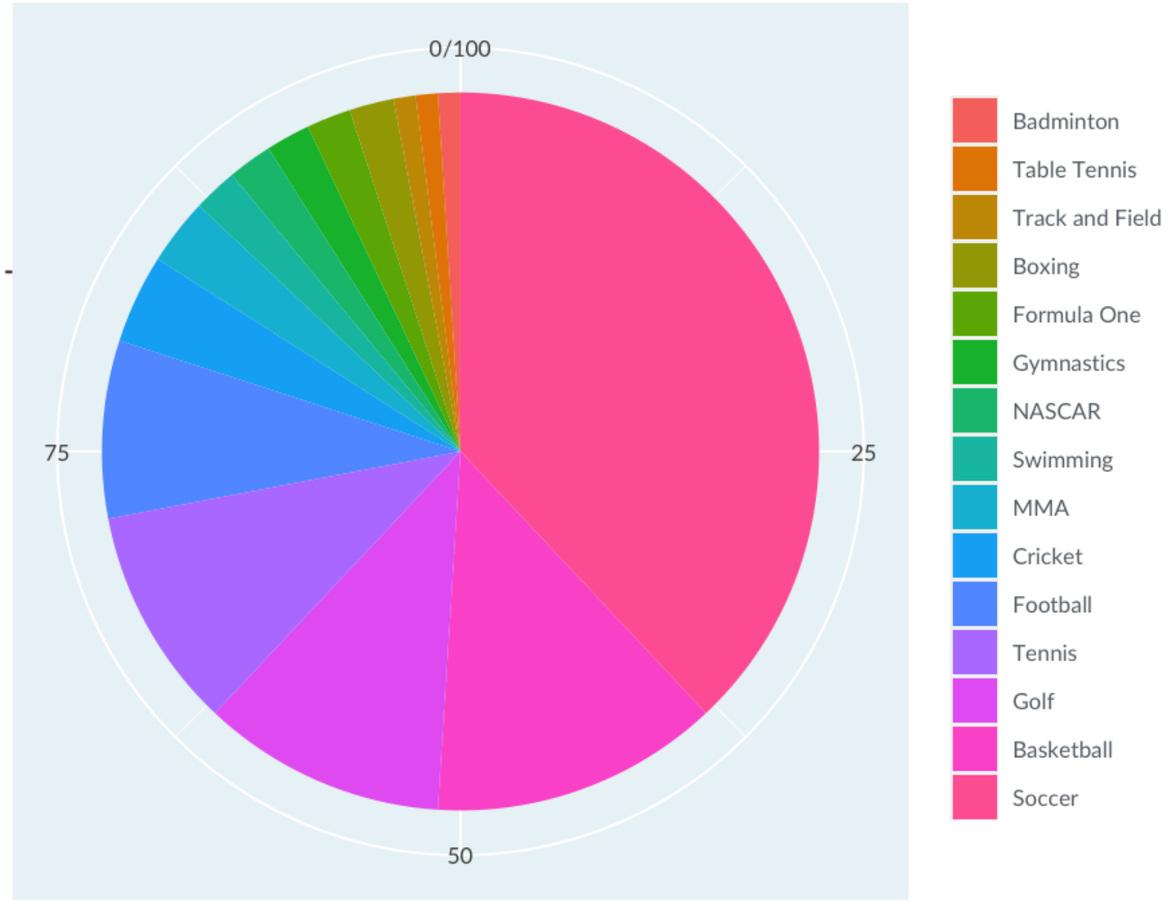
Pie plots



ESPN famous athletes, by sport



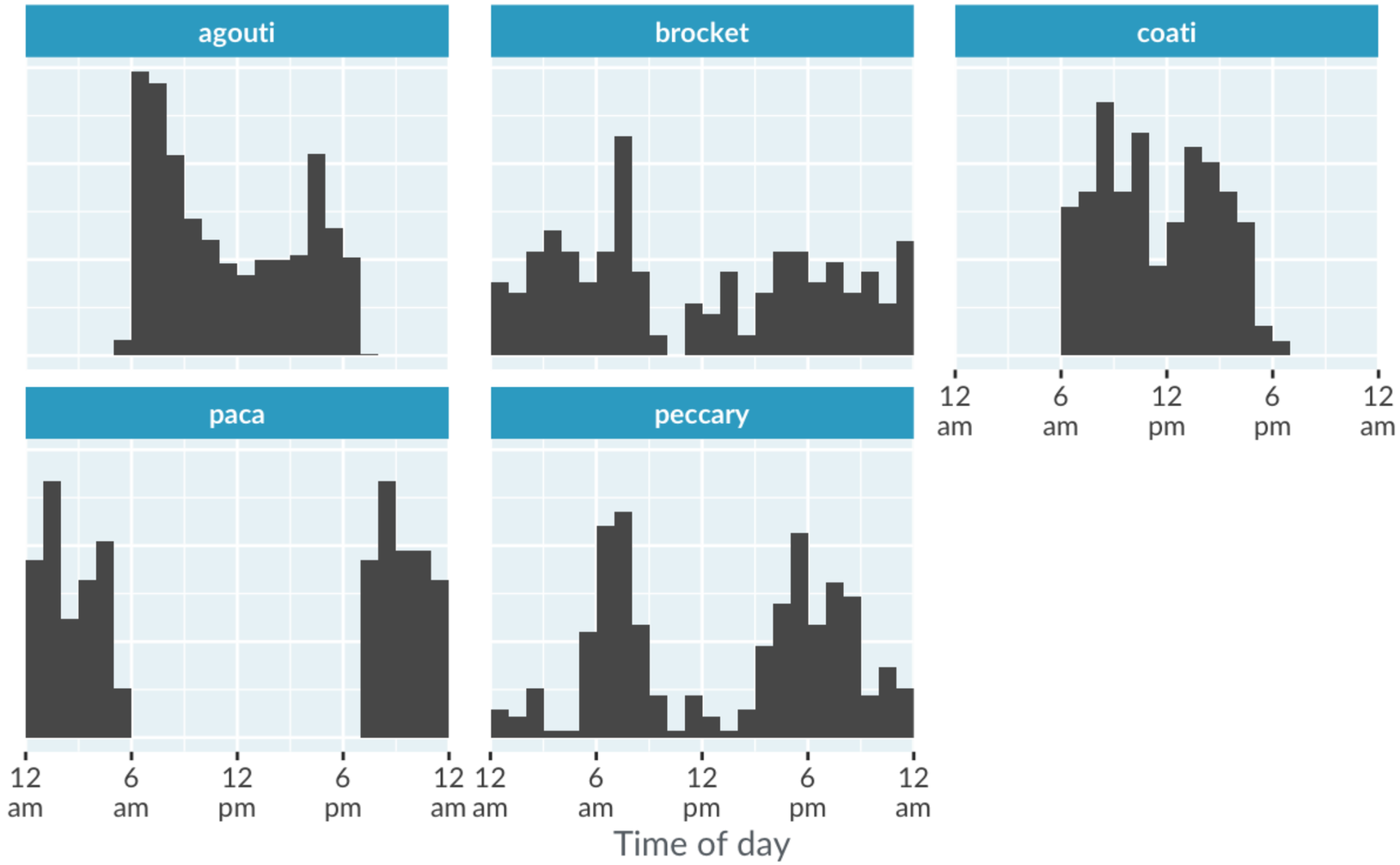
Bar plot + polar coords = pie plot



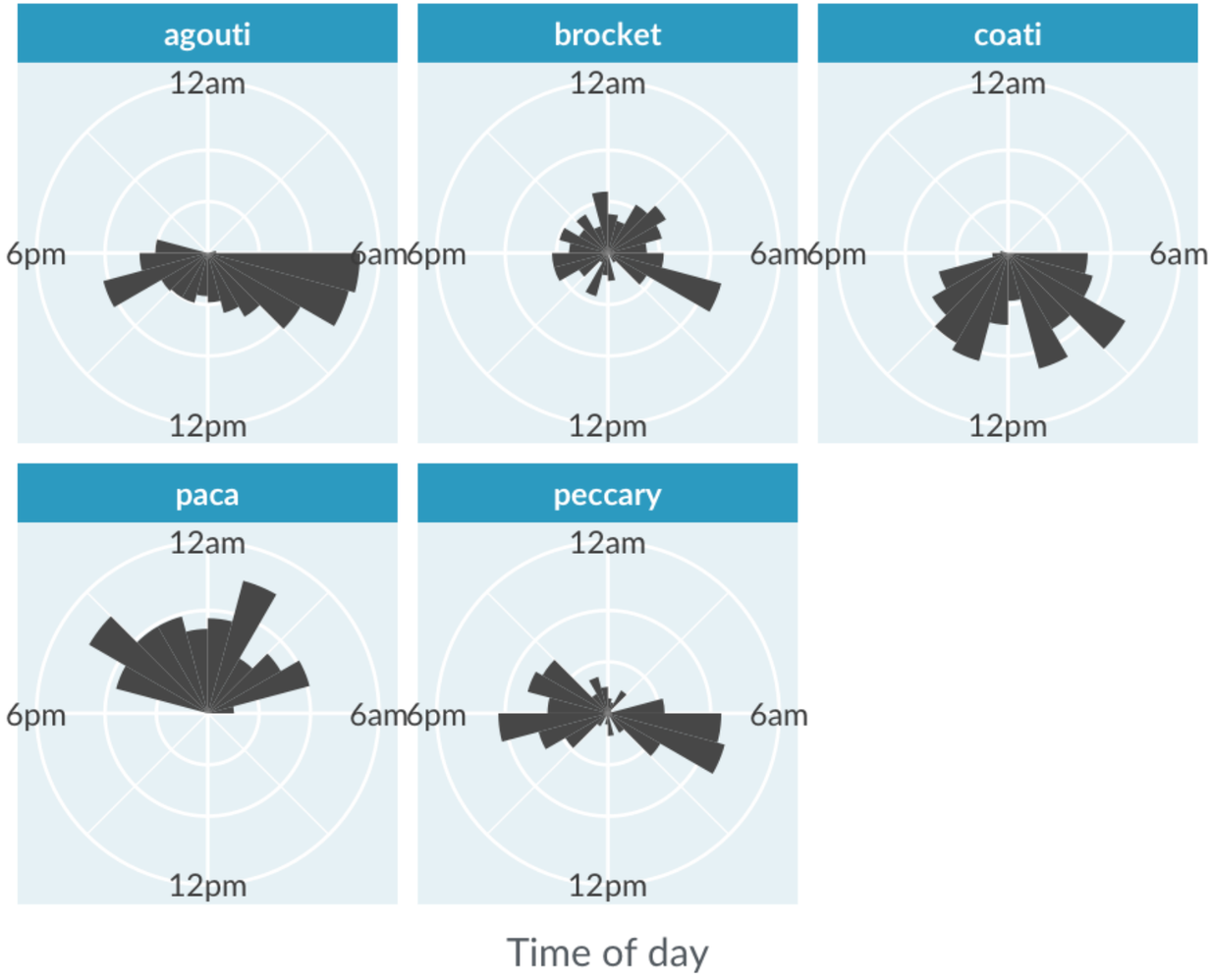
When should you use polar coordinates?

- Almost never.
- If you have a variable that is naturally circular (time of day, compass direction).

Histogram of animal activity



Histogram + polar coords = rose plot



Let's practice!

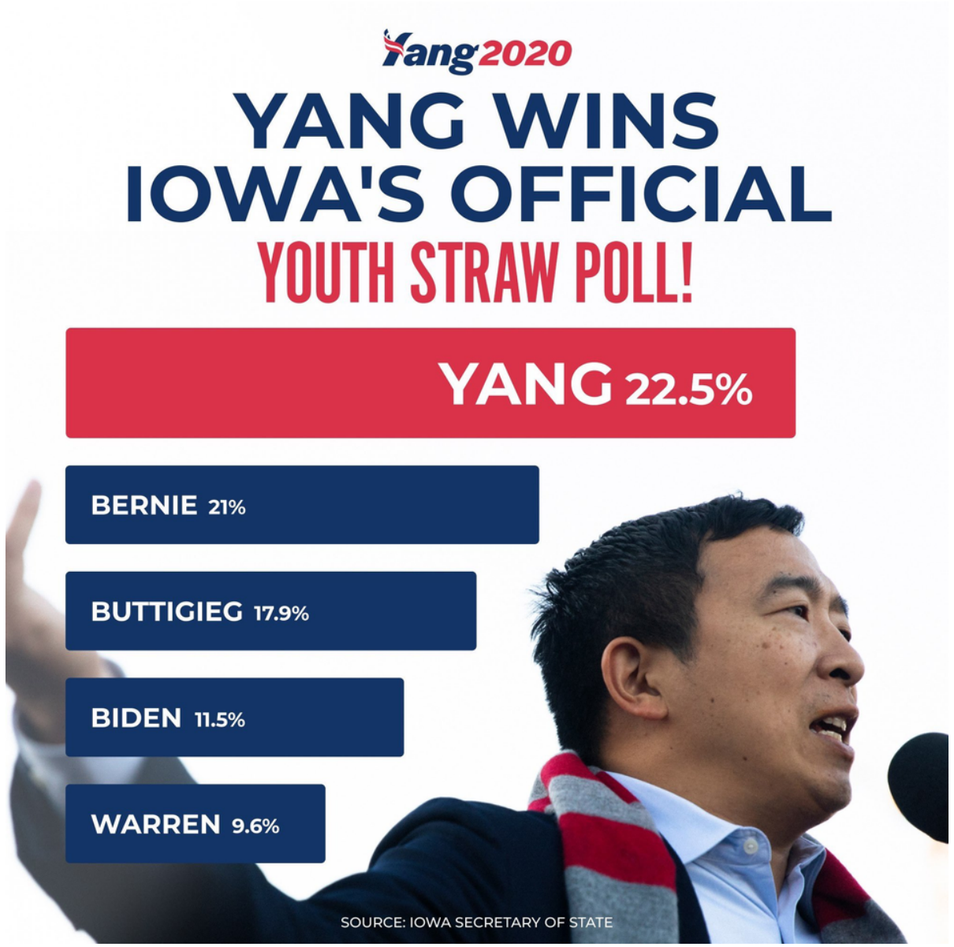
UNDERSTANDING DATA VISUALIZATION

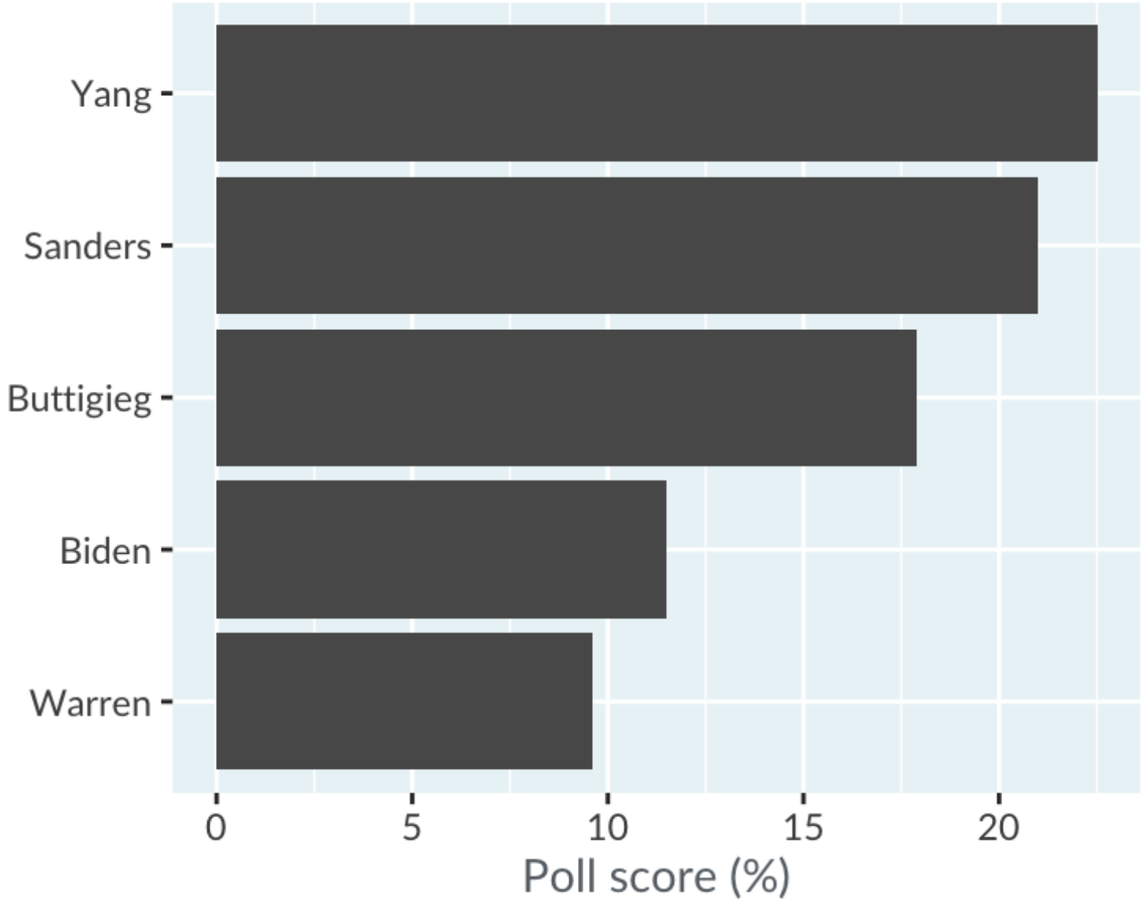
Axes of evil

UNDERSTANDING DATA VISUALIZATION

Richie Cotton
Data Evangelist at DataCamp

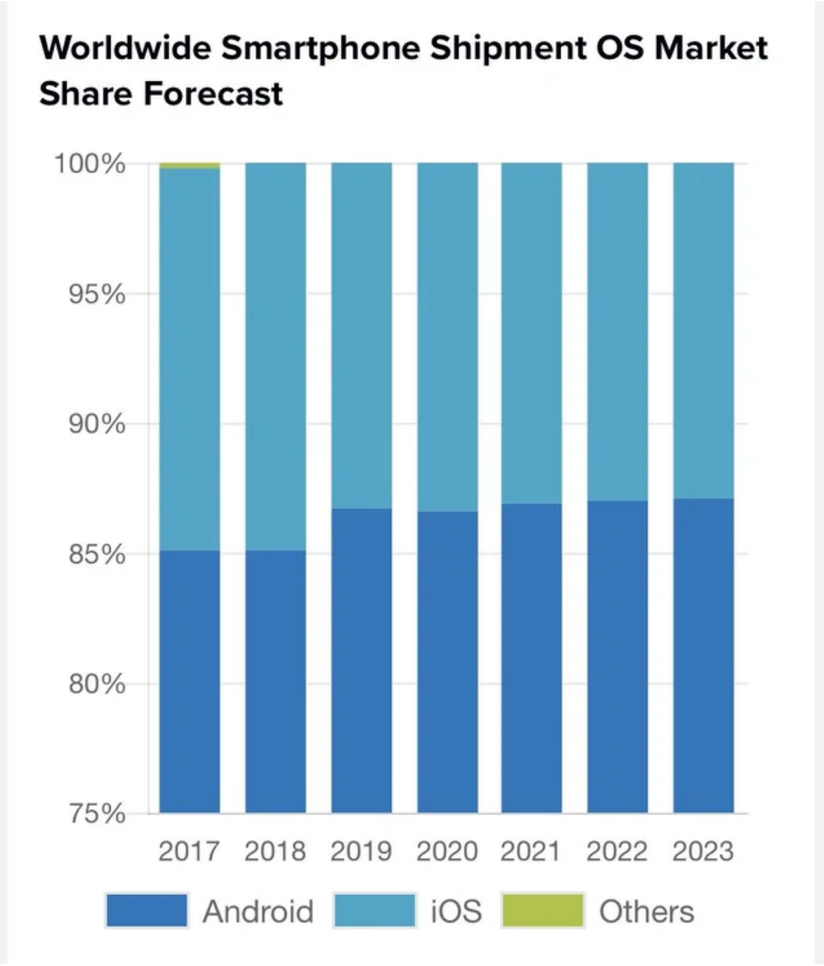
Nonsense bar lengths

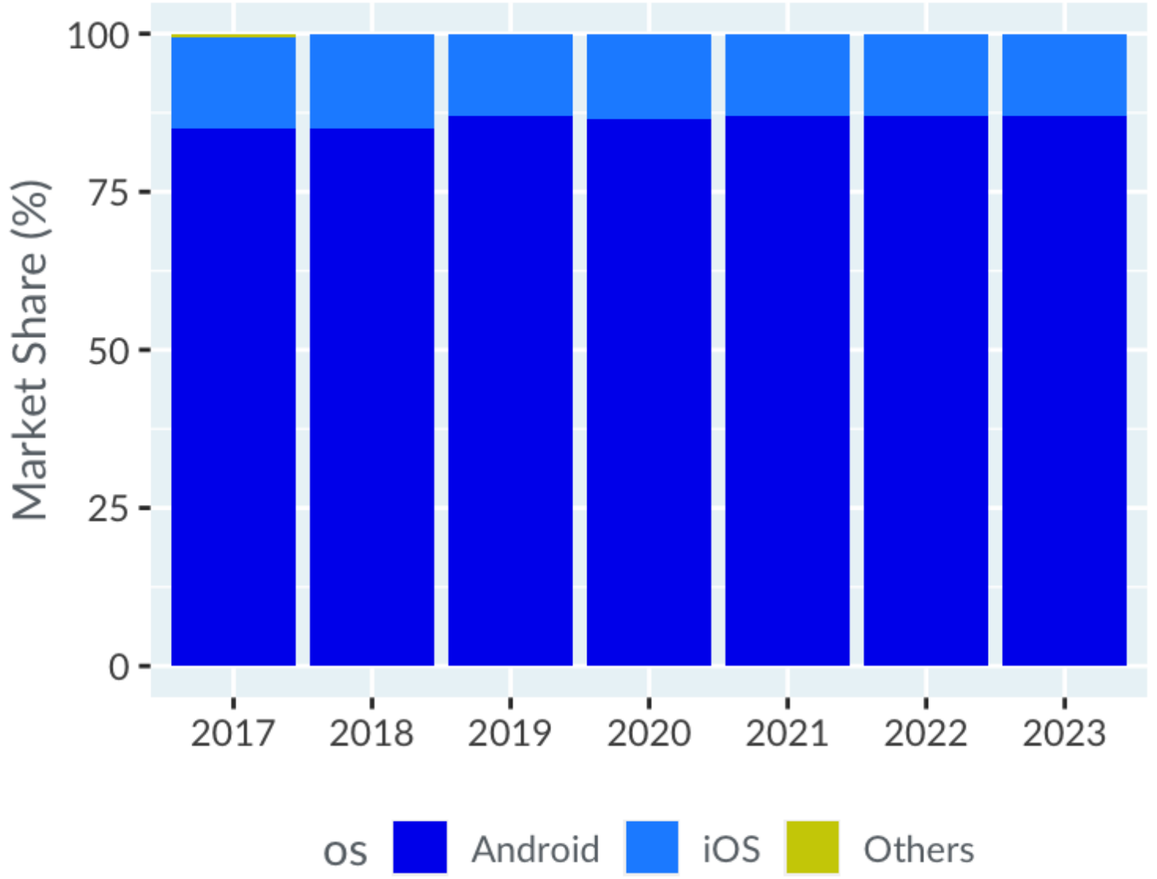




[1] https://www.reddit.com/r/dataisugly/comments/exewcc/thats_quite_a_large_15

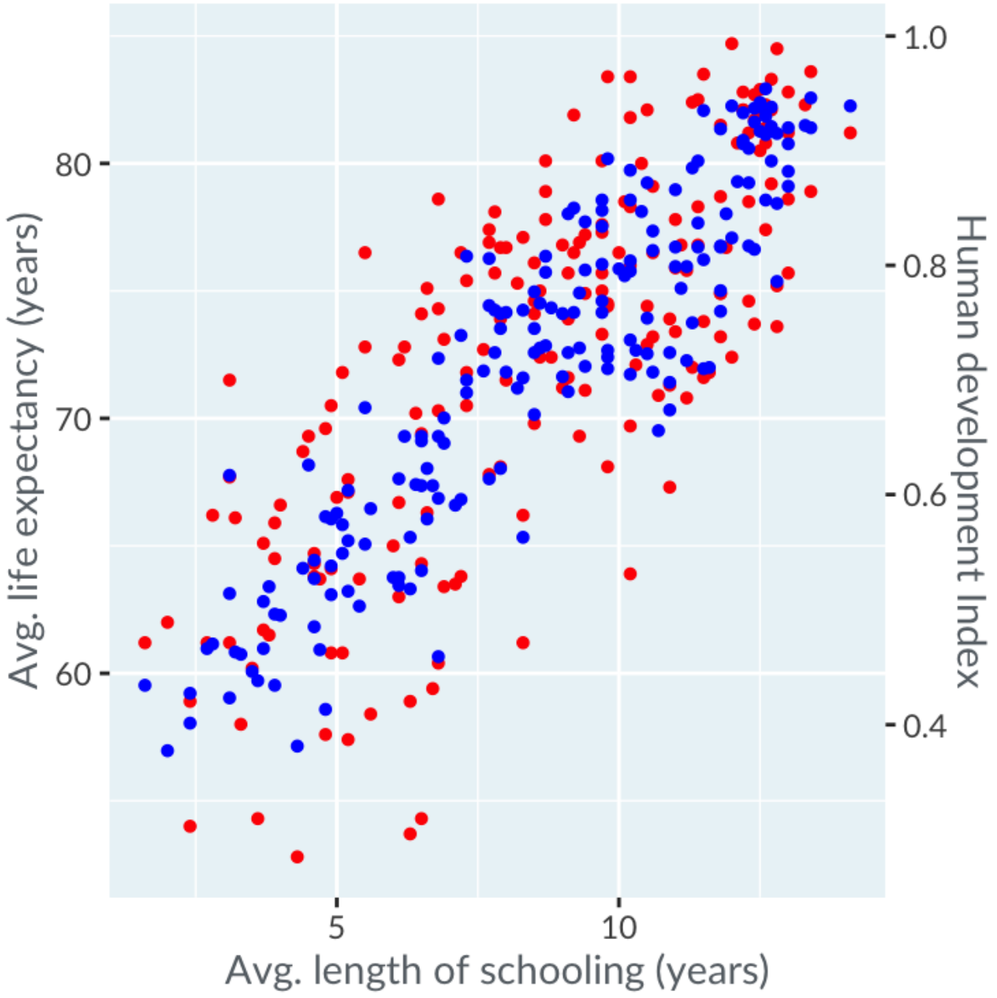
The same applies to stacked bar plots

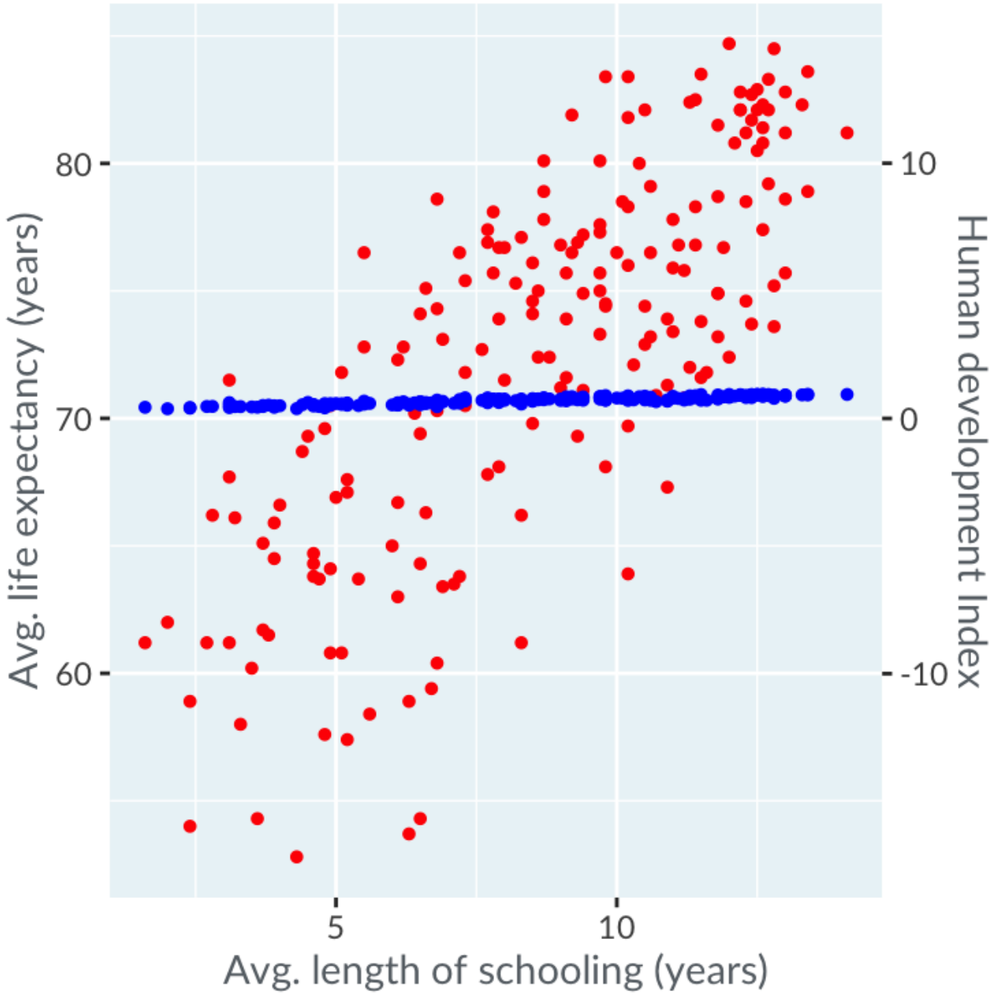




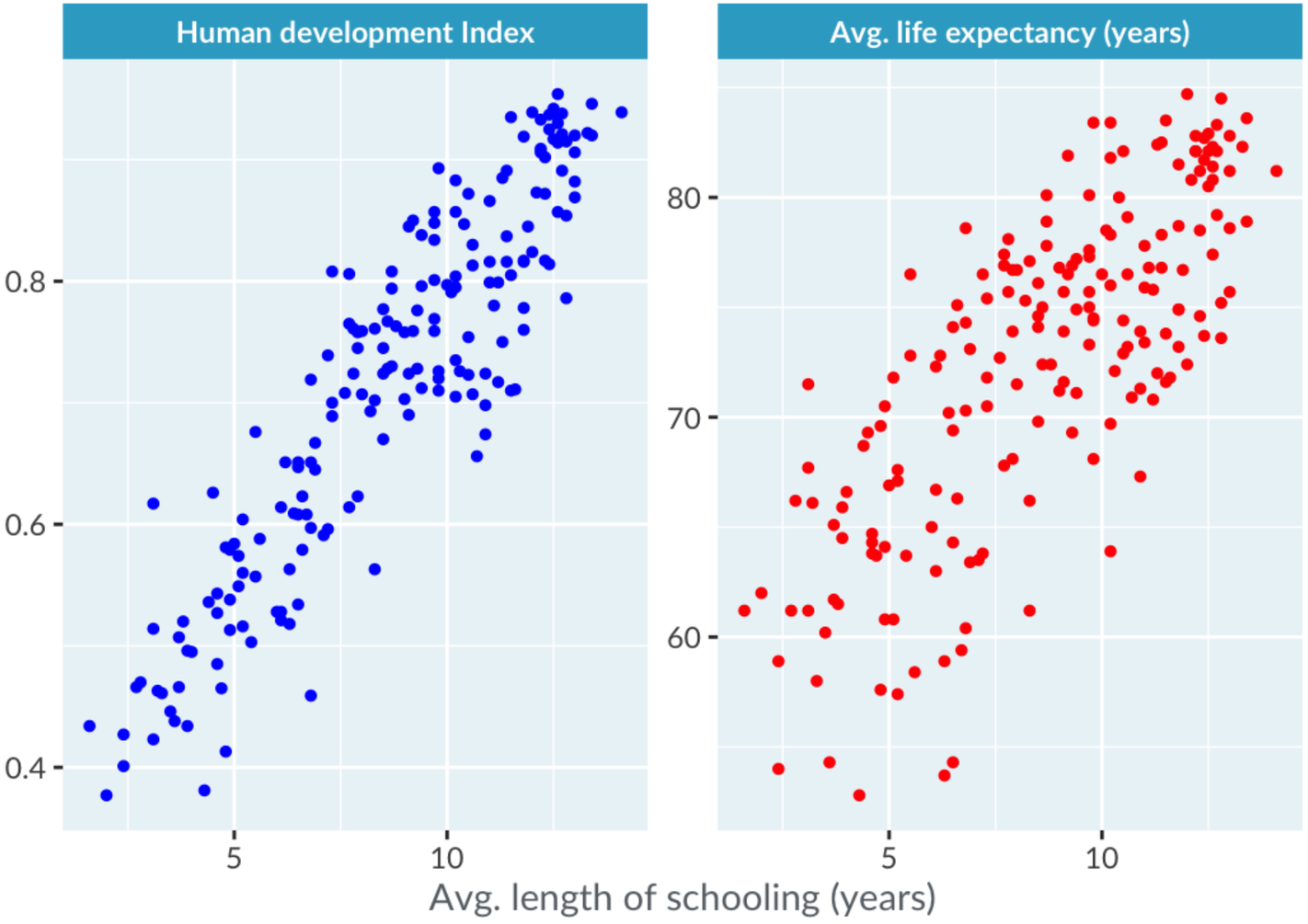
[1] https://www.reddit.com/r/dataisugly/comments/d76ixt/lets_make_13_vs_87_market_share_look_like_5050

Dual axes are misleading





Better to use multiple panels



Let's practice!

UNDERSTANDING DATA VISUALIZATION

Sensory overload

UNDERSTANDING DATA VISUALIZATION

Richie Cotton
Data Evangelist at DataCamp

Measures of a good visualization

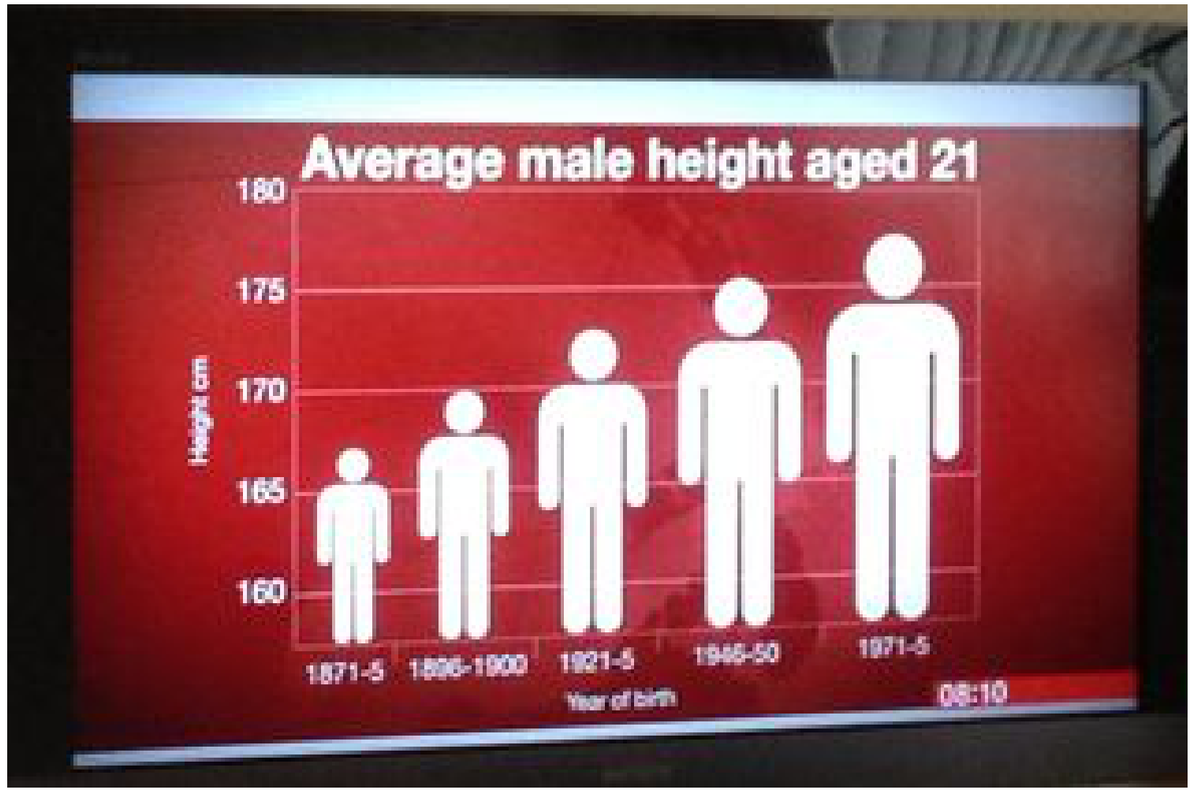
- How many interesting insights can your reader get from the plot?
- How quickly can they get those insights?

Chartjunk

> Any element of the plot that distracts from the reader getting insight

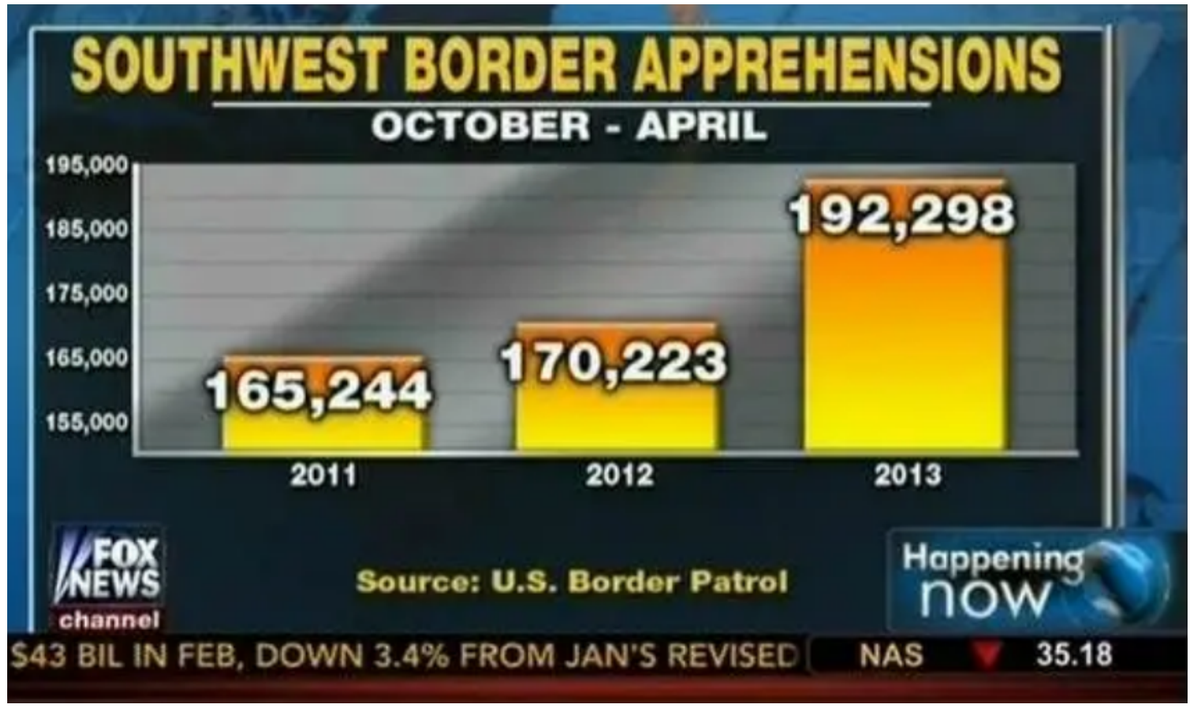
- Pictures
- Skueomorphism: reflections, shadows, etc.
- Extra dimensions
- Ostentatious colors or lines

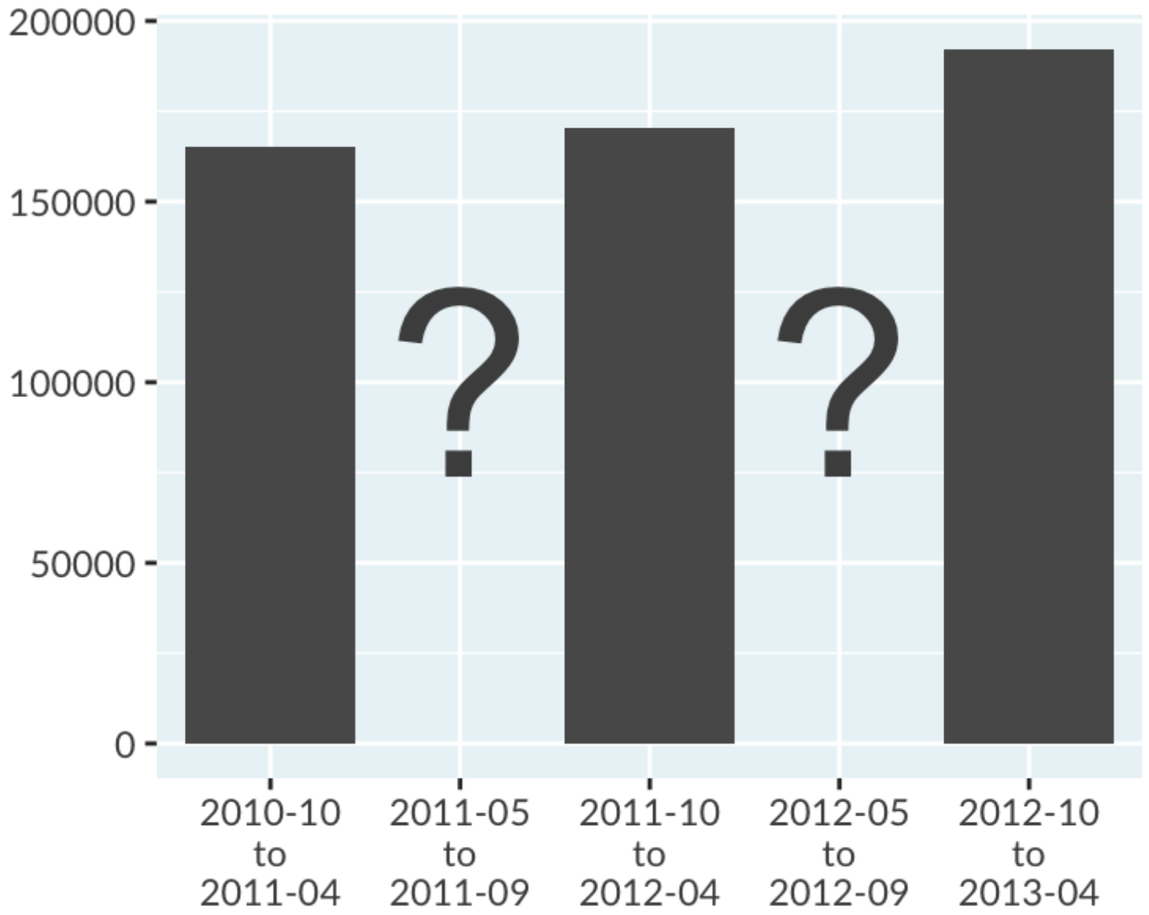
Pictures



[1] https://junkcharts.typepad.com/junk_charts/2013/09/the-incredibly-expanding-male.html

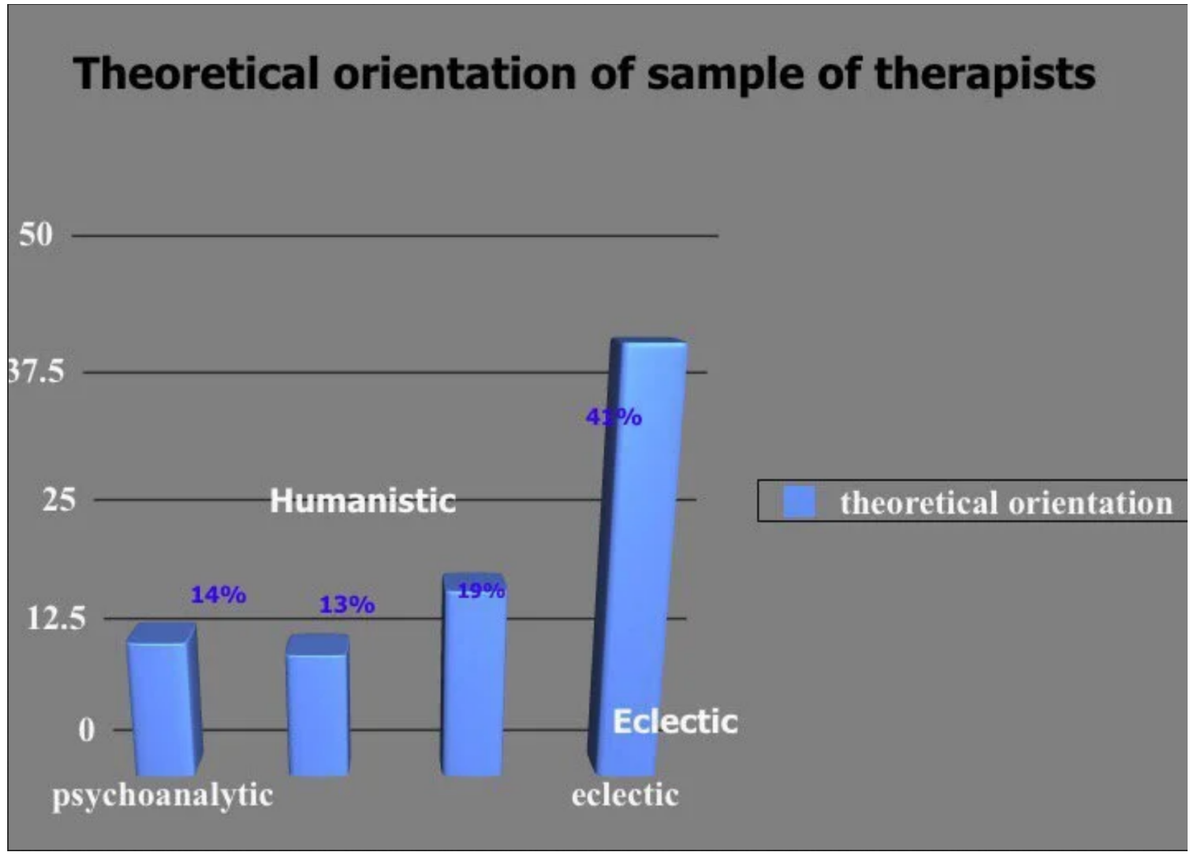
Skeuomorphism

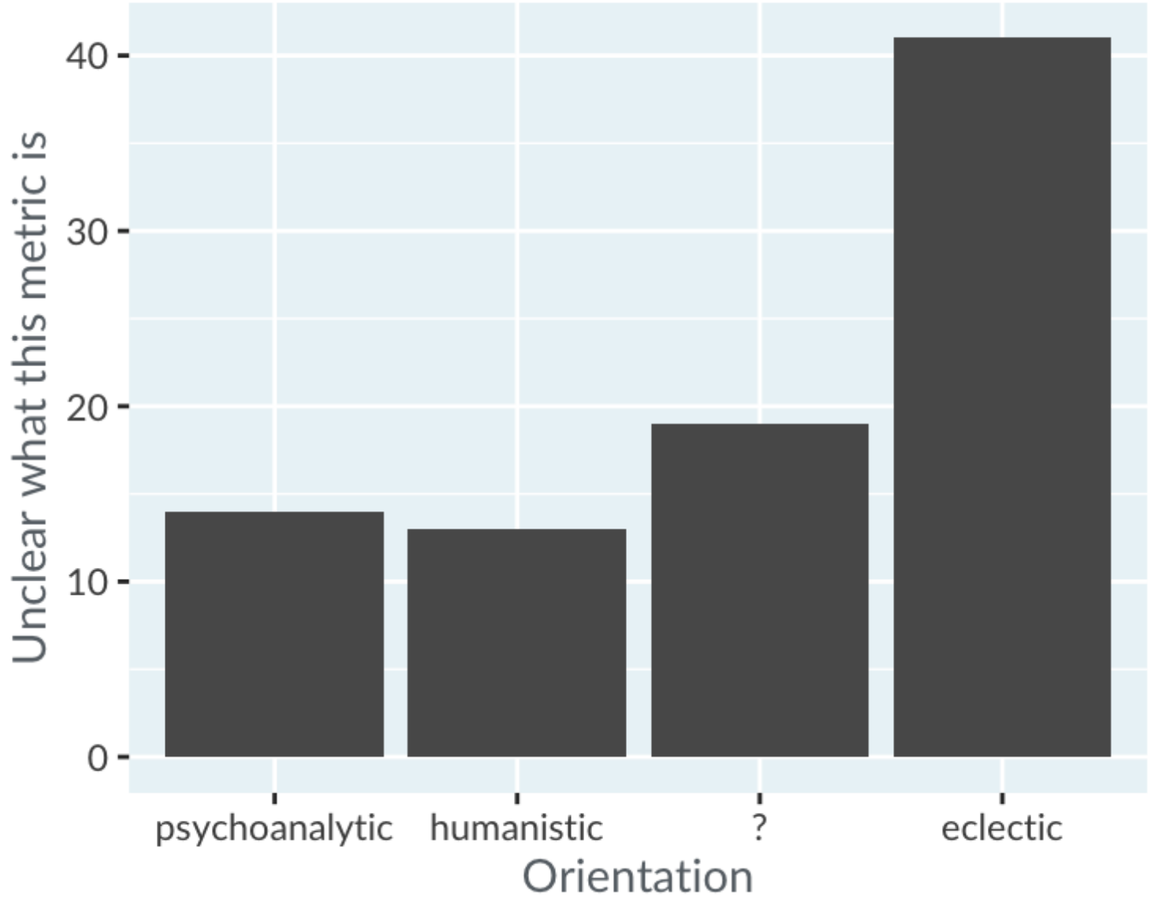




[1] https://www.mediamatters.org/fox-news/fox-news-newest-dishonest-chart-immigration-enforcement

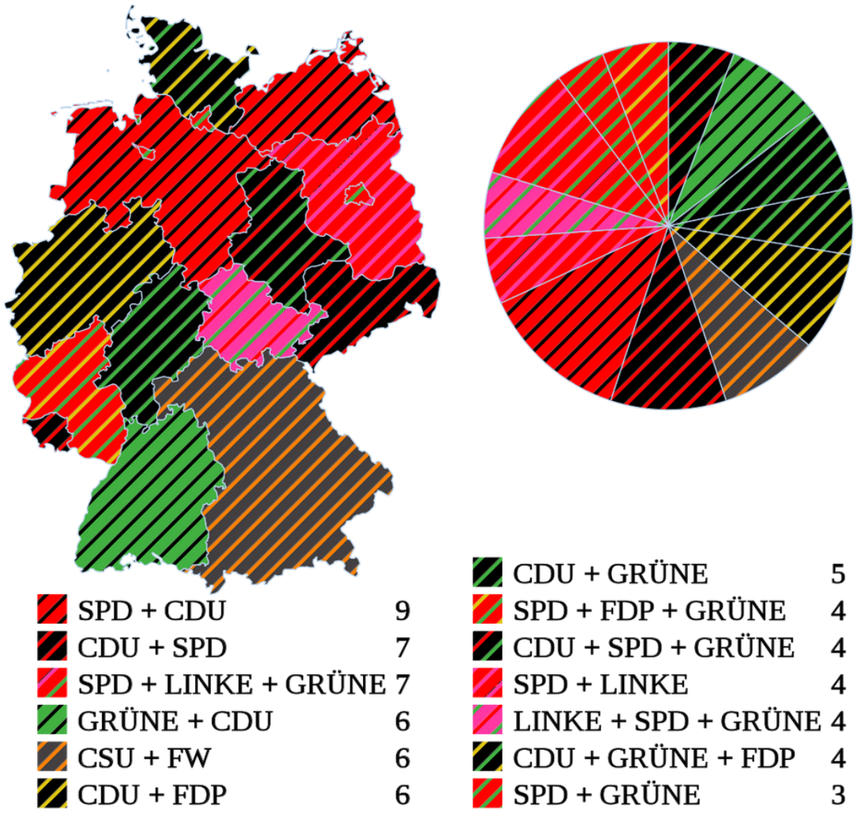
Extra dimensions





[1] https://www.reddit.com/r/dataisugly/comments/dh6yra/an_actual_graph_presented_in_my_psychology

Ostentatious colors and lines



[1] https://www.reddit.com/r/dataisugly/comments/cyhle4/wikipedias_image_of_the_distribution_of_seats_in [2] https://imgur.com/y27mRLj

Let's practice!

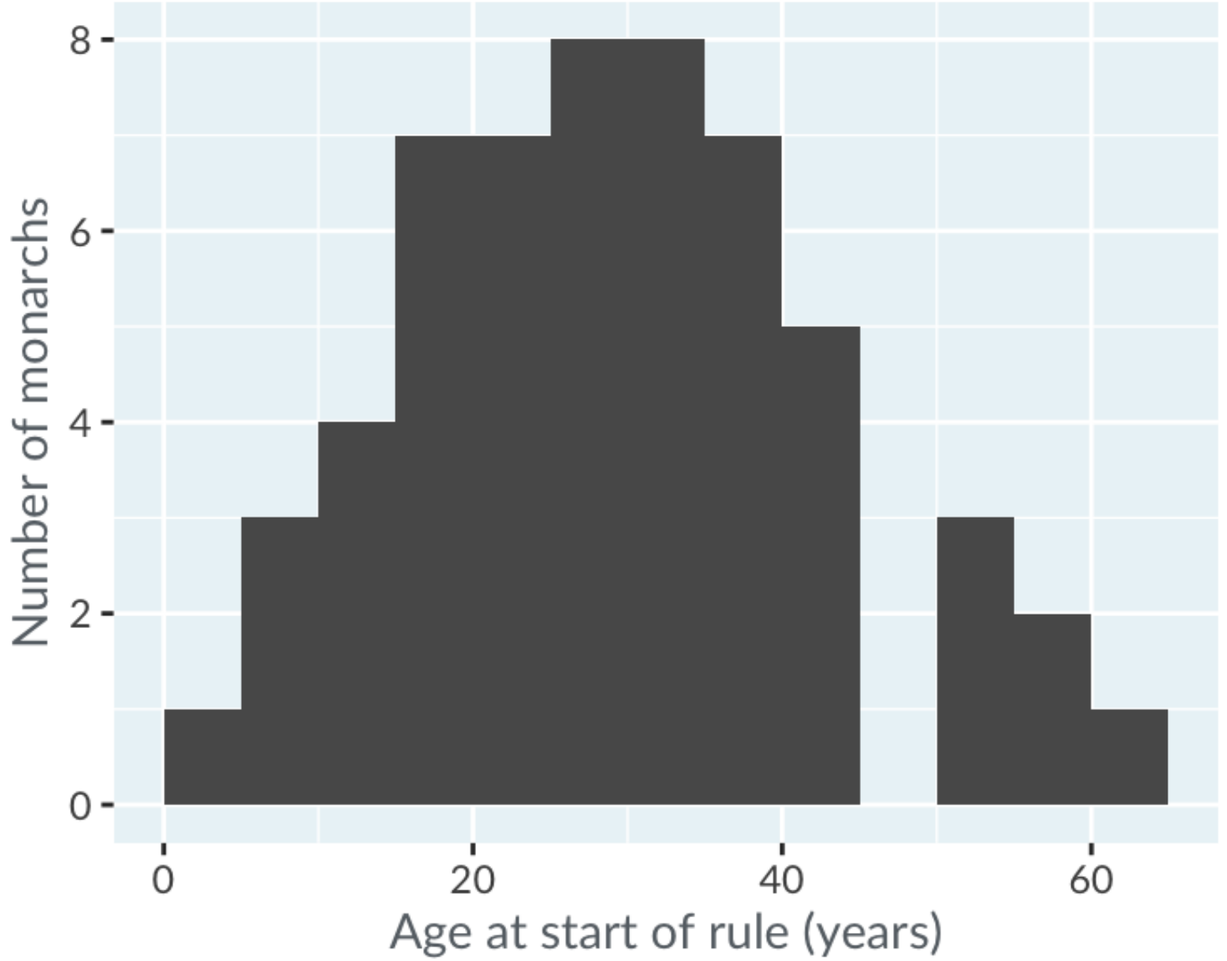
UNDERSTANDING DATA VISUALIZATION

Congratulations

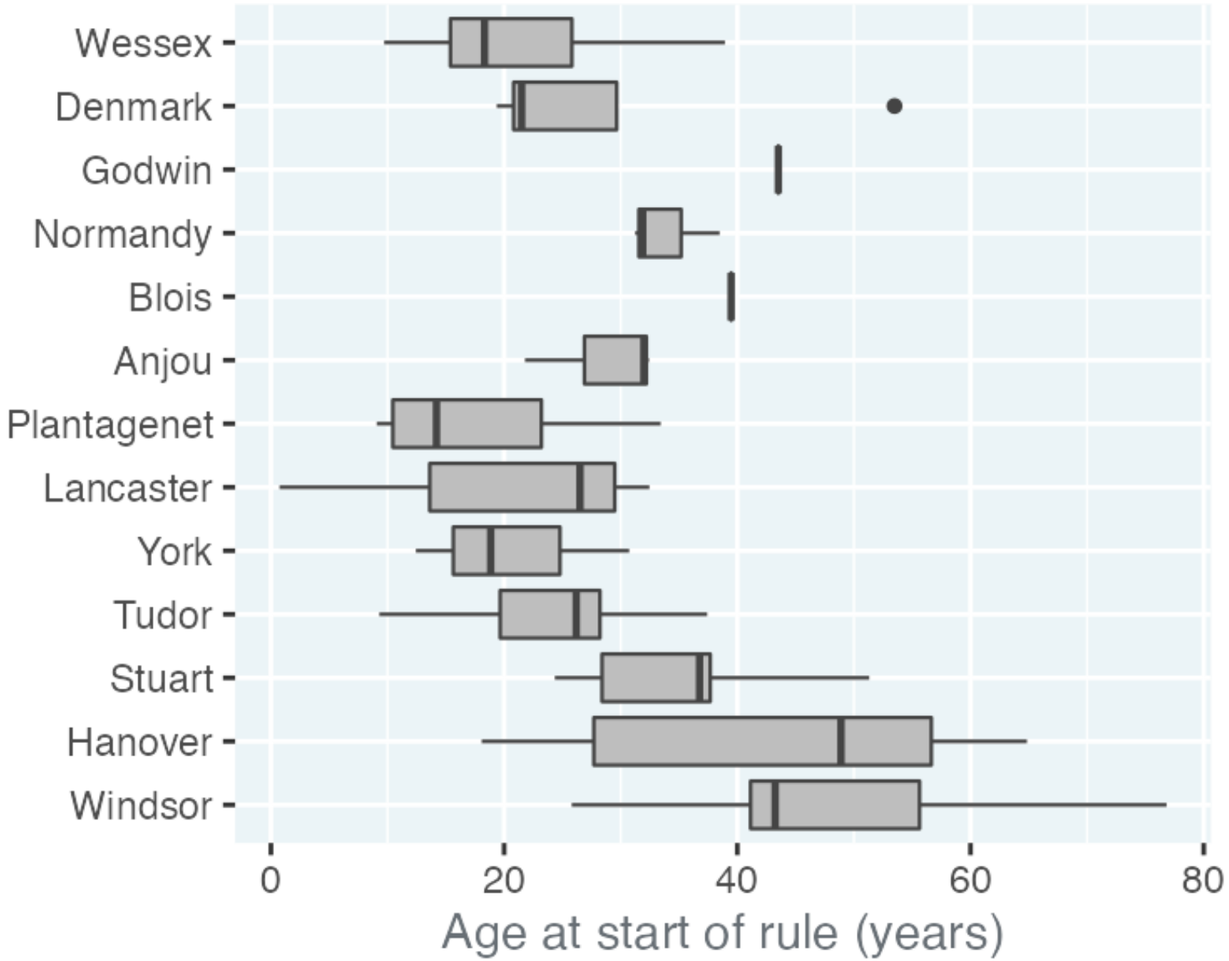
UNDERSTANDING DATA VISUALIZATION

Richie Cotton
Data Evangelist at DataCamp

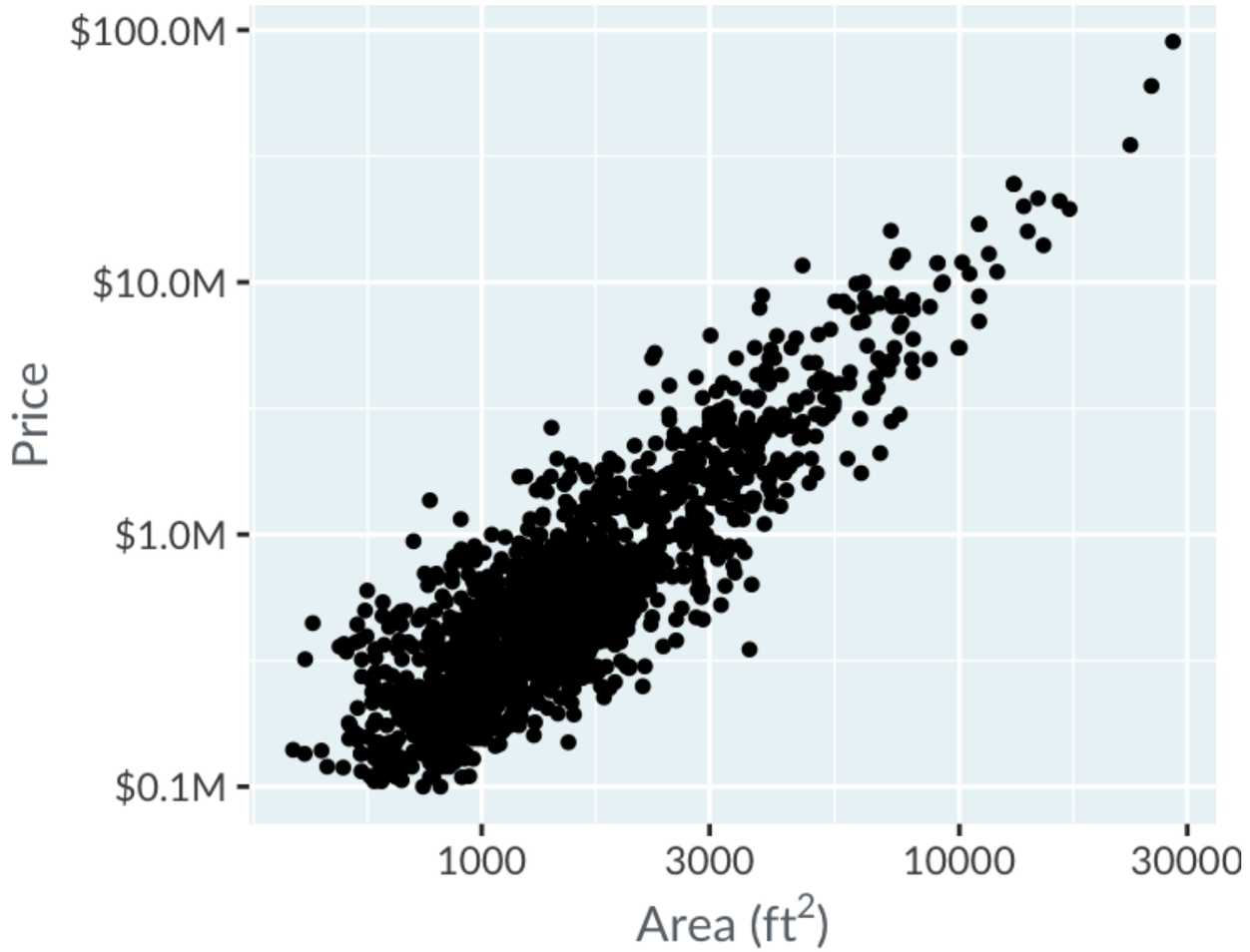
Histograms: show a distribution



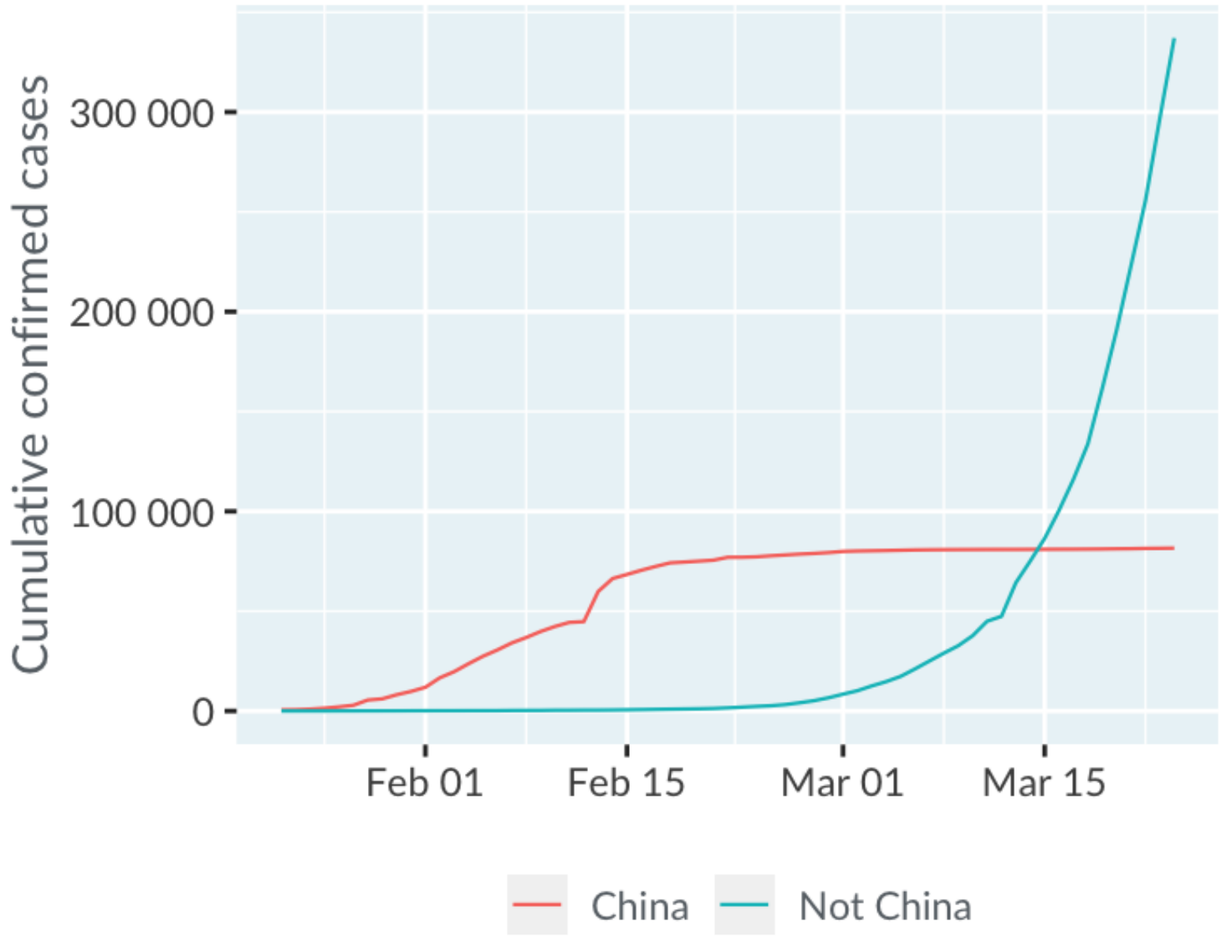
Box plots: show lots of distributions



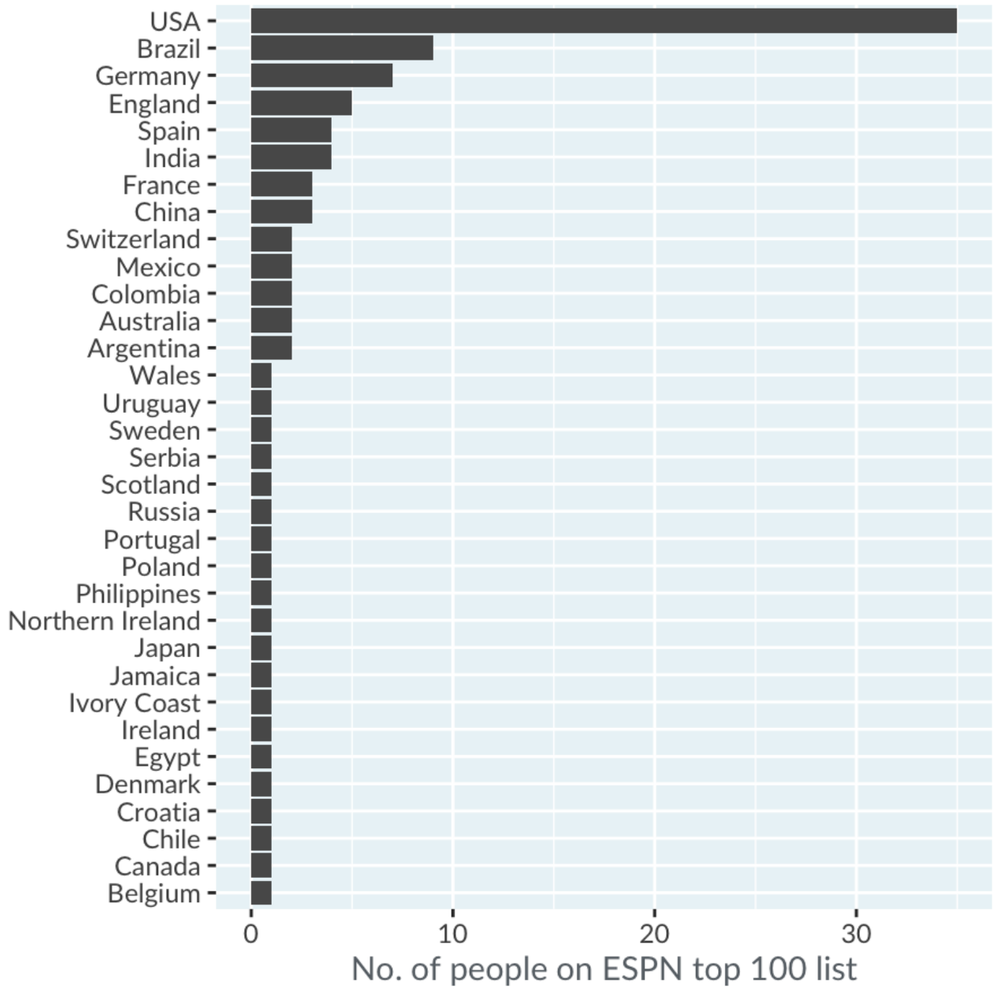
Scatter plots: compare two numeric variables



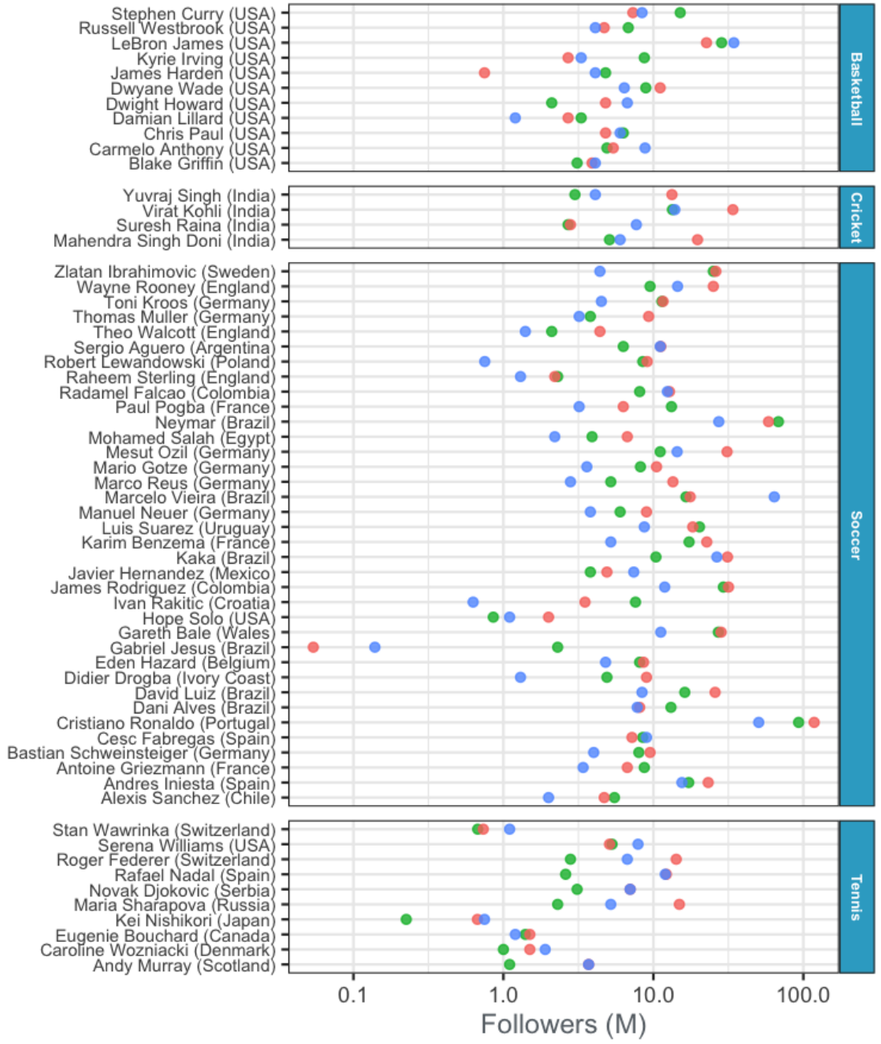
Line plots: show trends over time



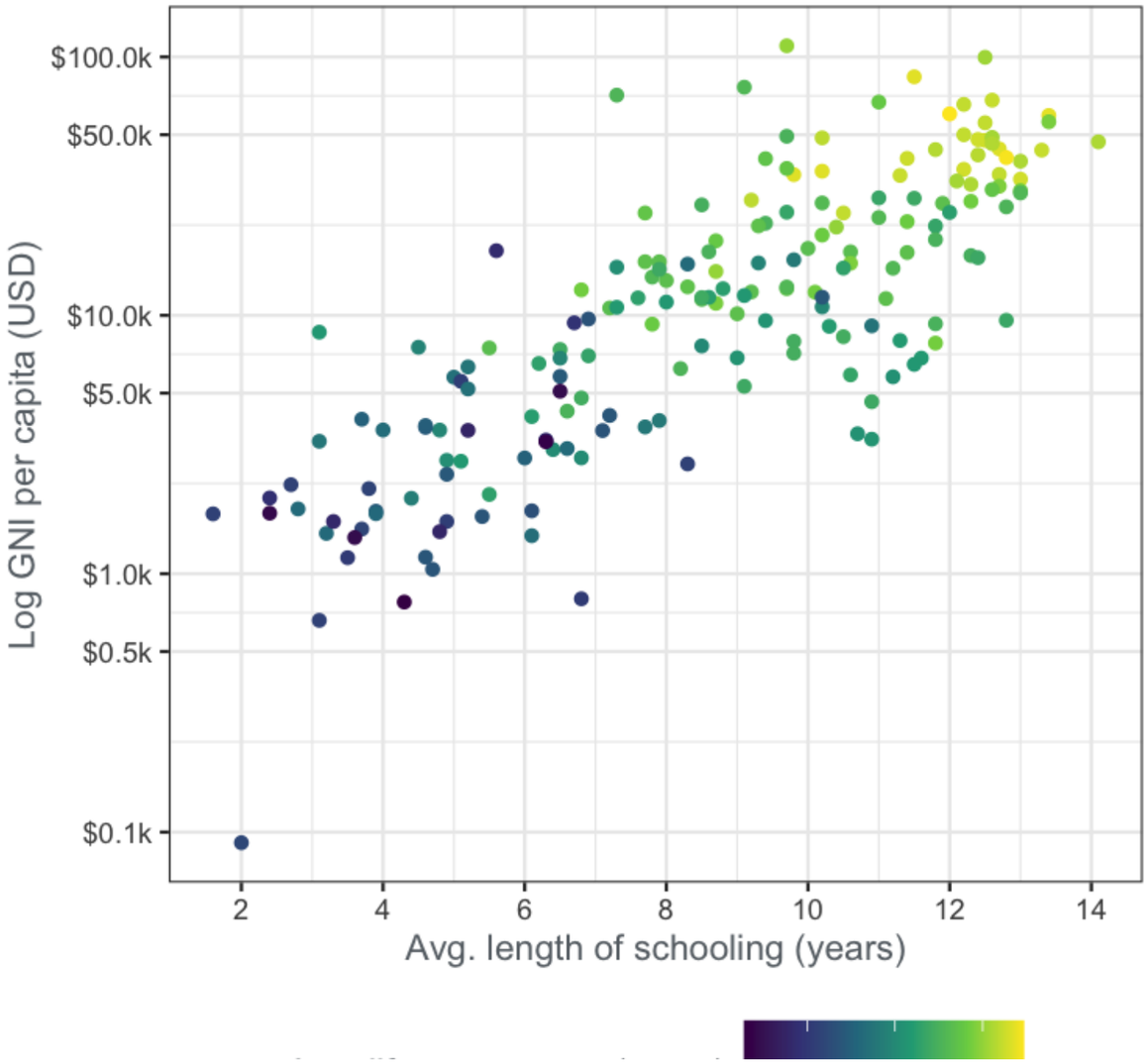
Bar plots: show counts by category

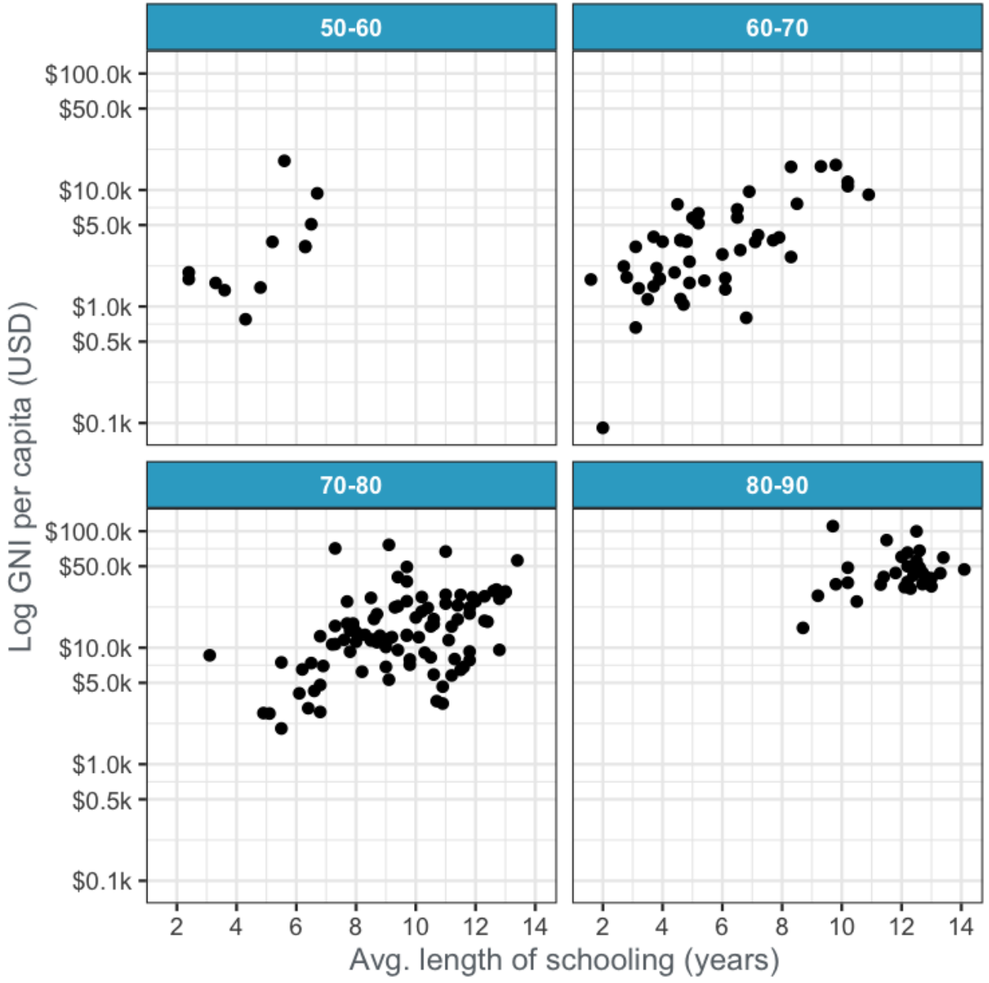


Dot plots: show log scale metrics by category

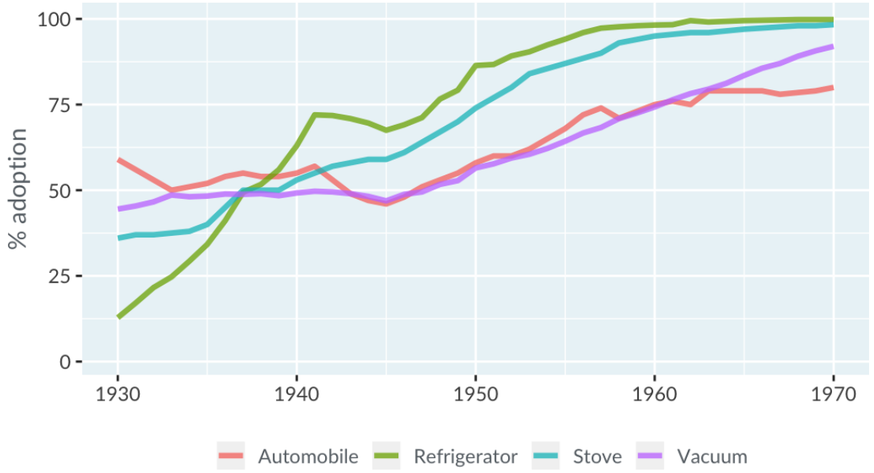


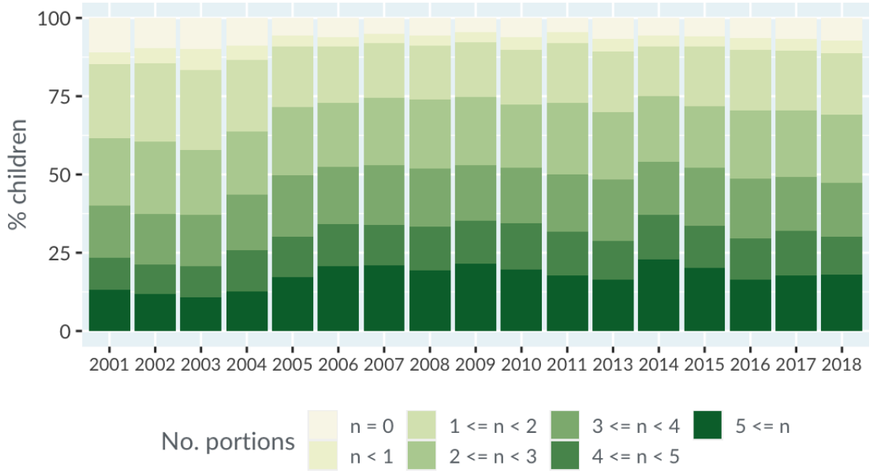
Extra dimensions

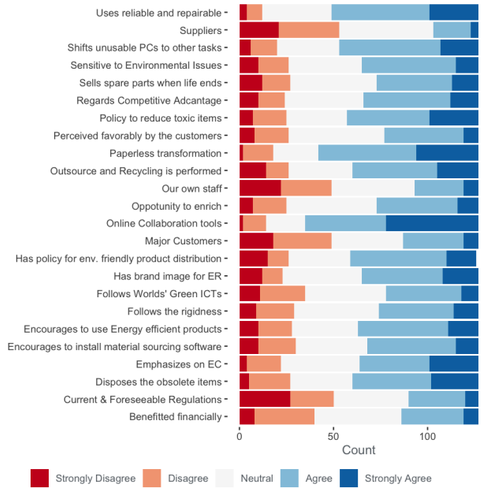




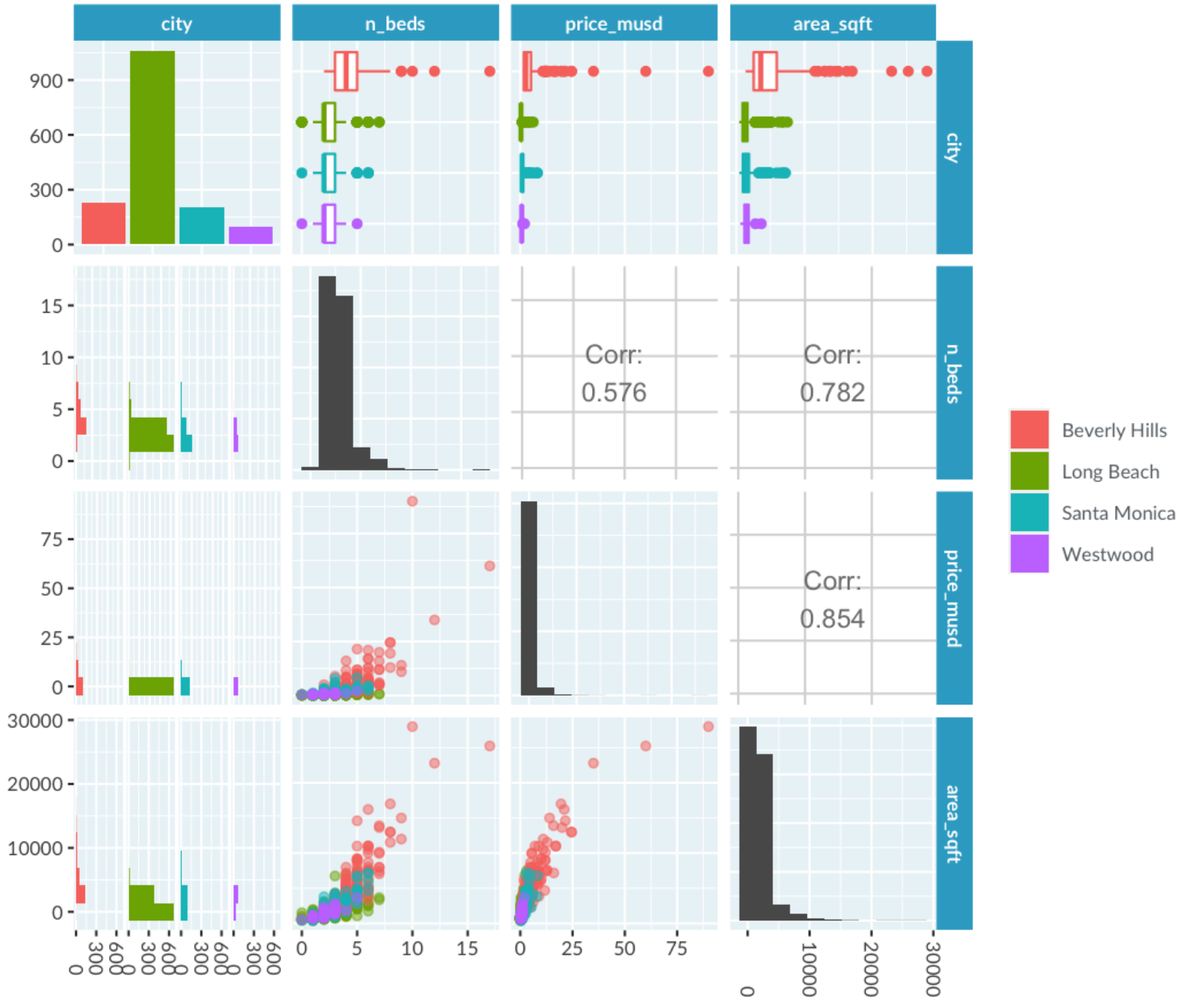
3 types of color scale



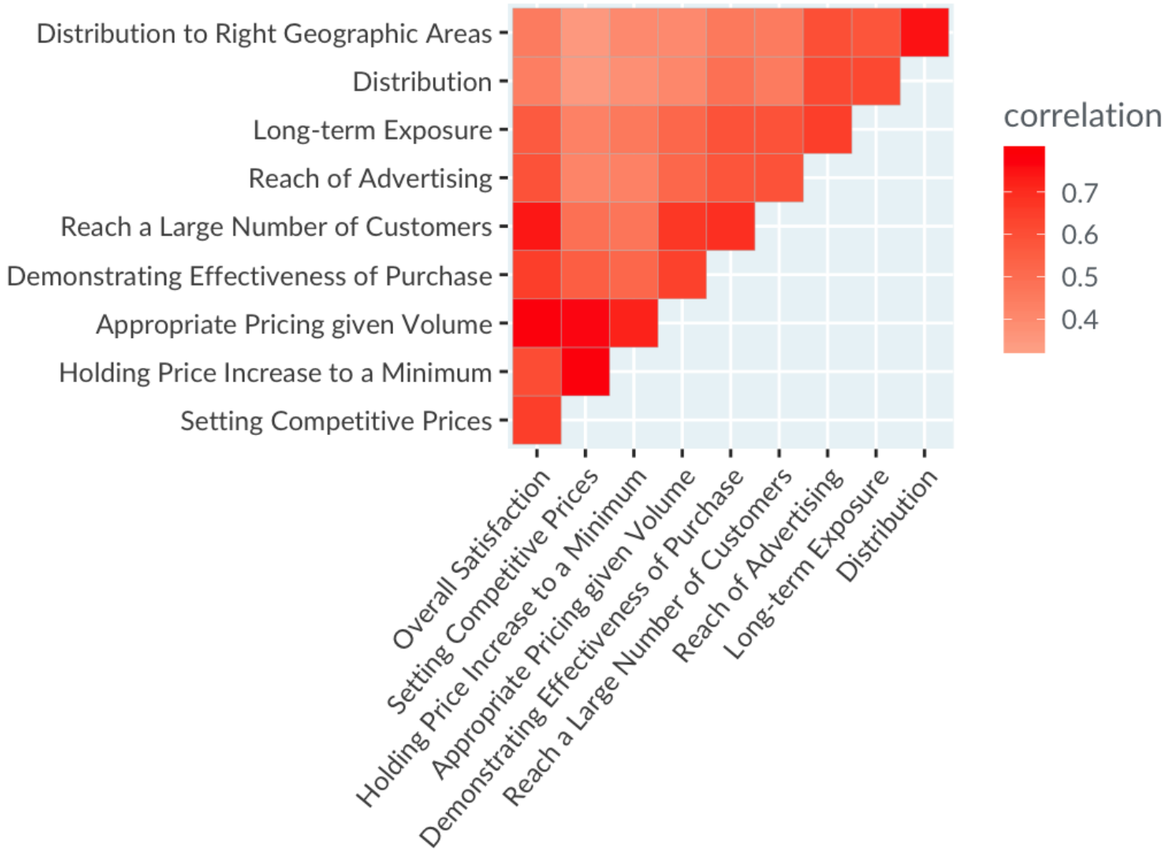




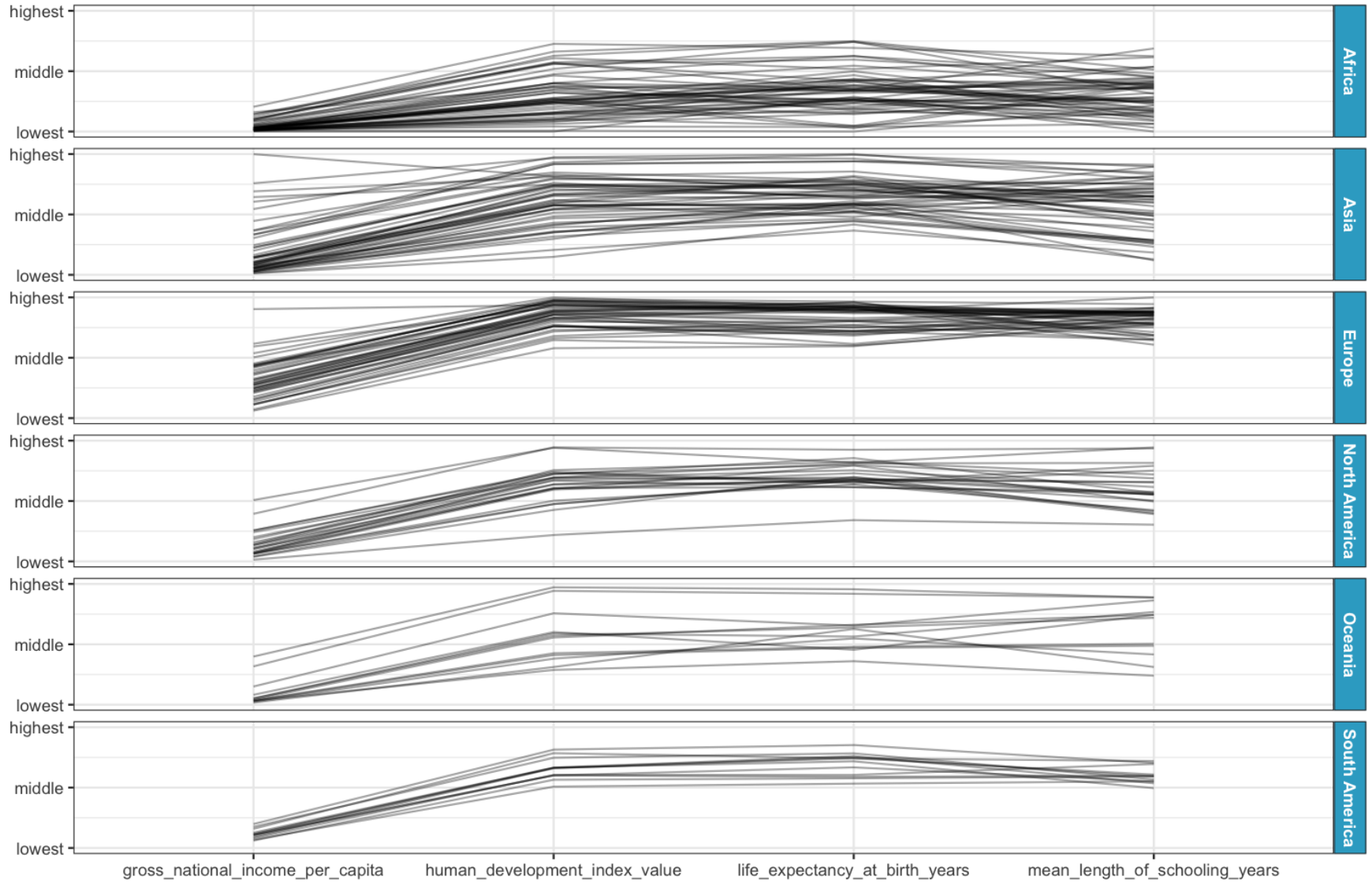
Pair plot: compare many variables



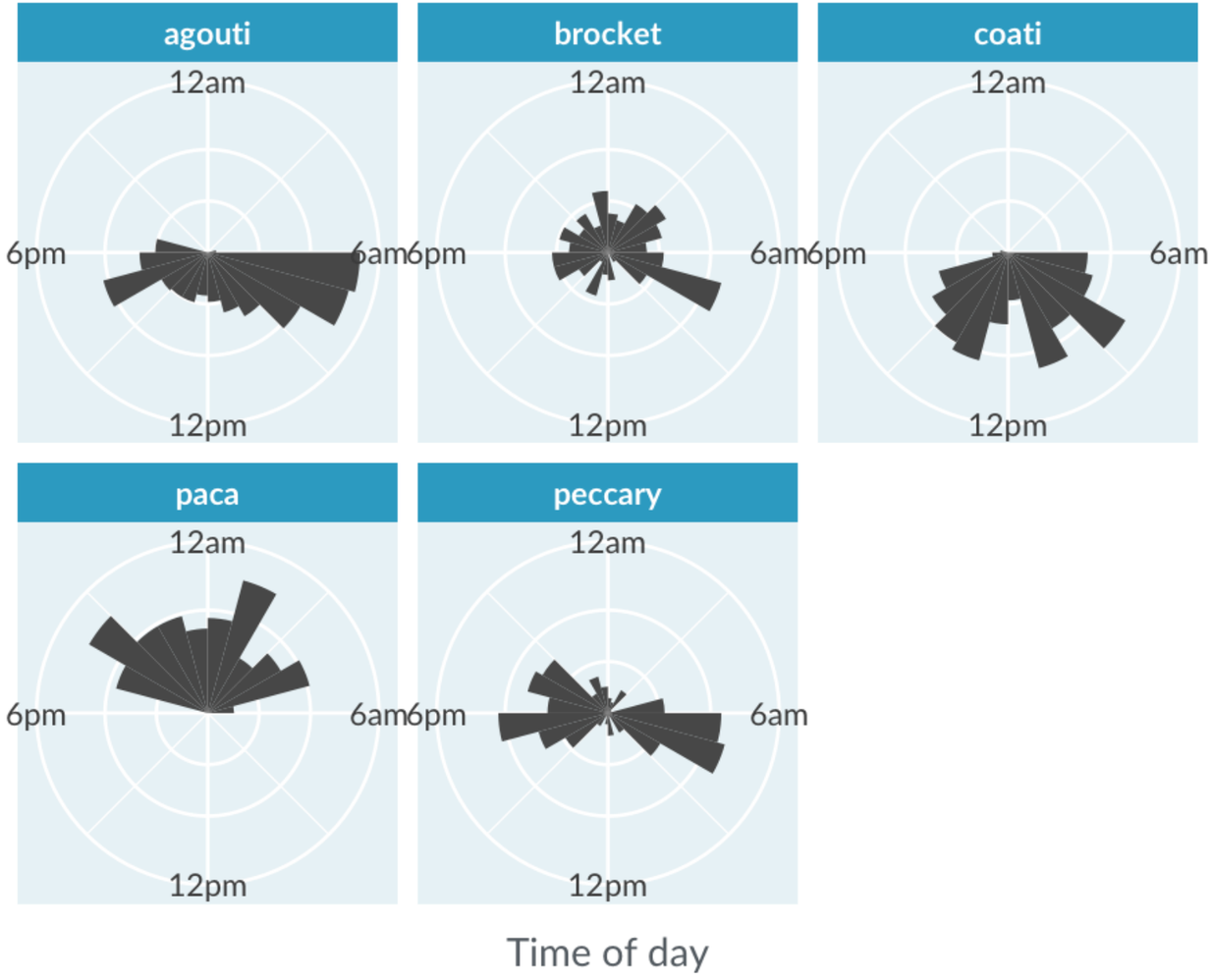
Correlation heatmap: show related variables



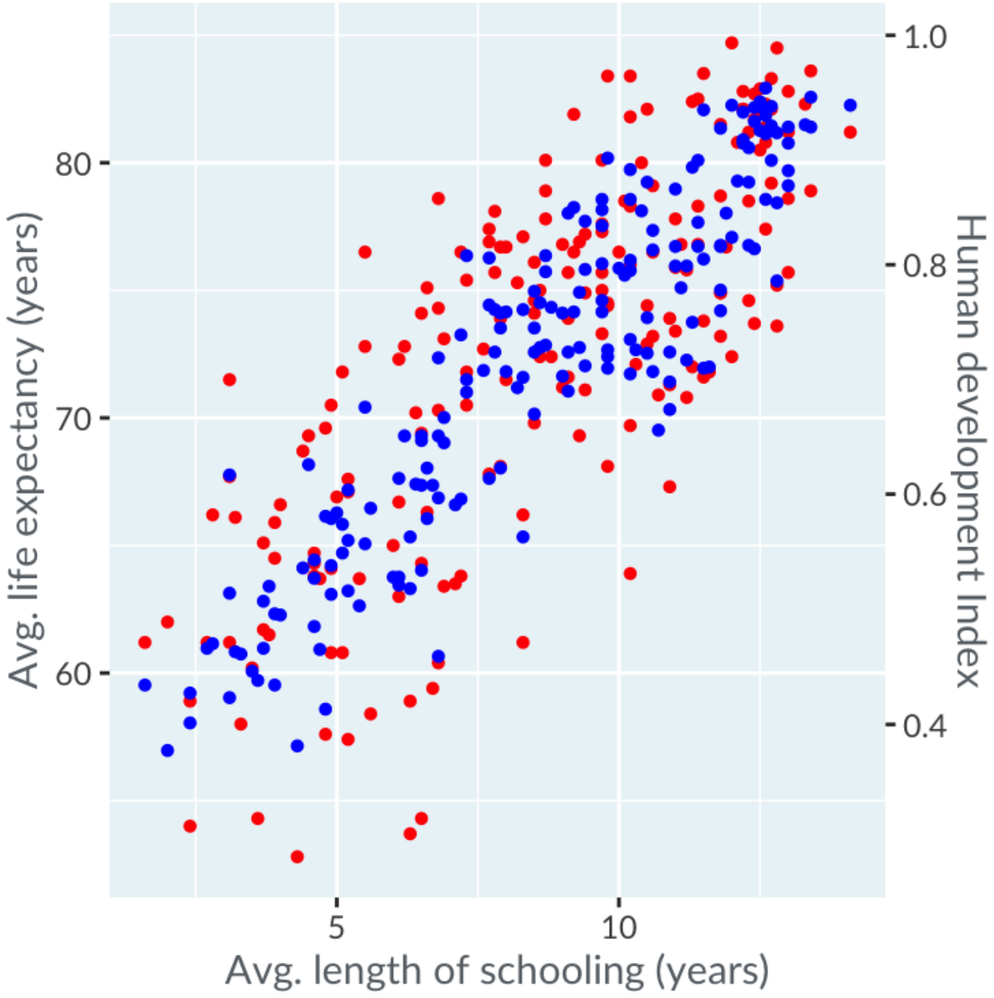
Parallel coordinates plot: find patterns across variables

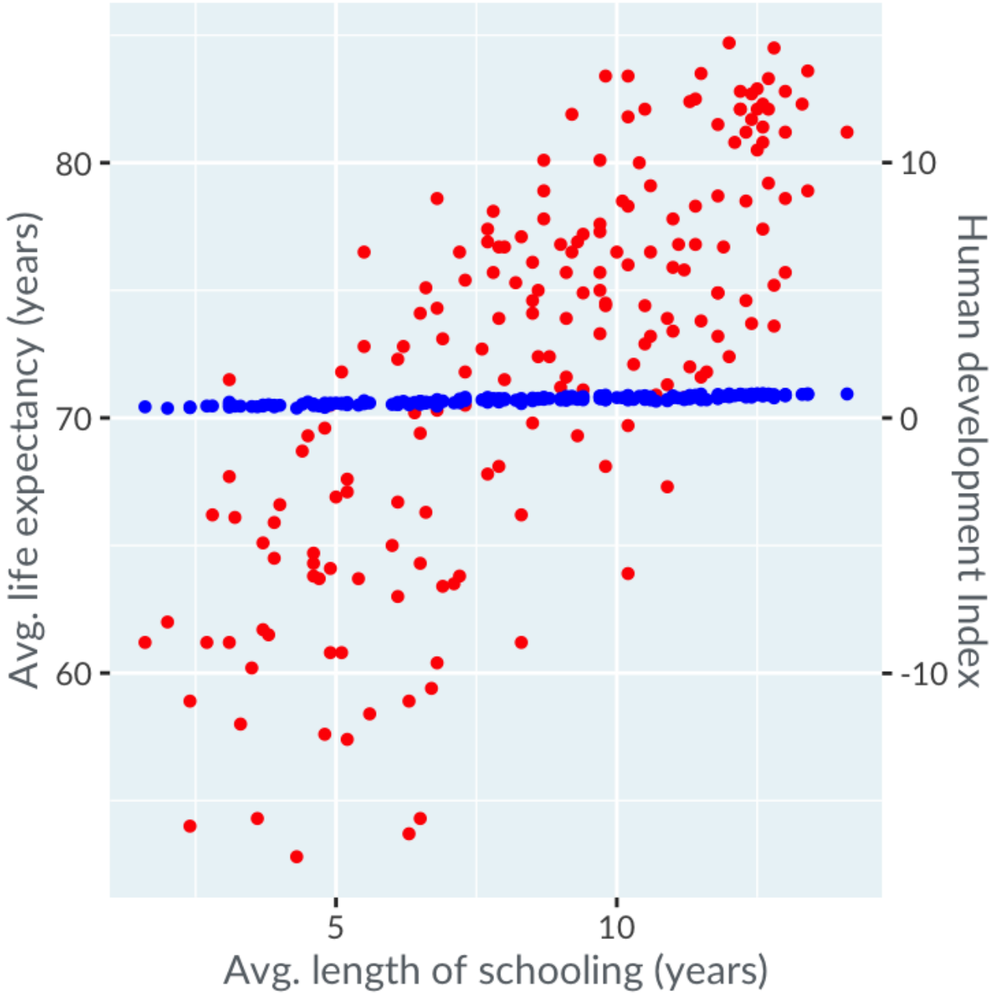


Rose plot: show a cyclical distribution

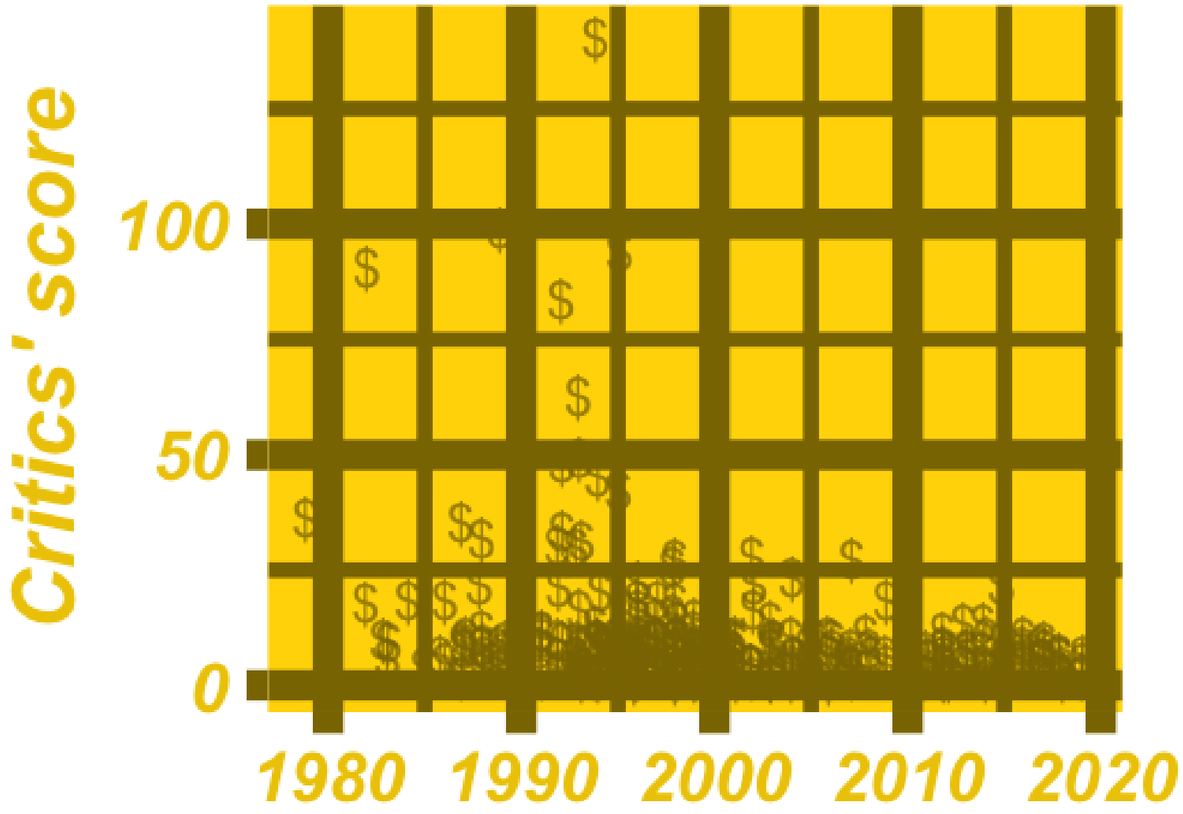


Dual axes are bad





Eliminate chartjunk



Next steps

- Introduction to Data Visualization with ggplot2
- Introduction to Tableau
- Introduction to Data Visualization with Matplotlib
- Introduction to Data Visualization with Seaborn

You made it!

UNDERSTANDING DATA VISUALIZATION