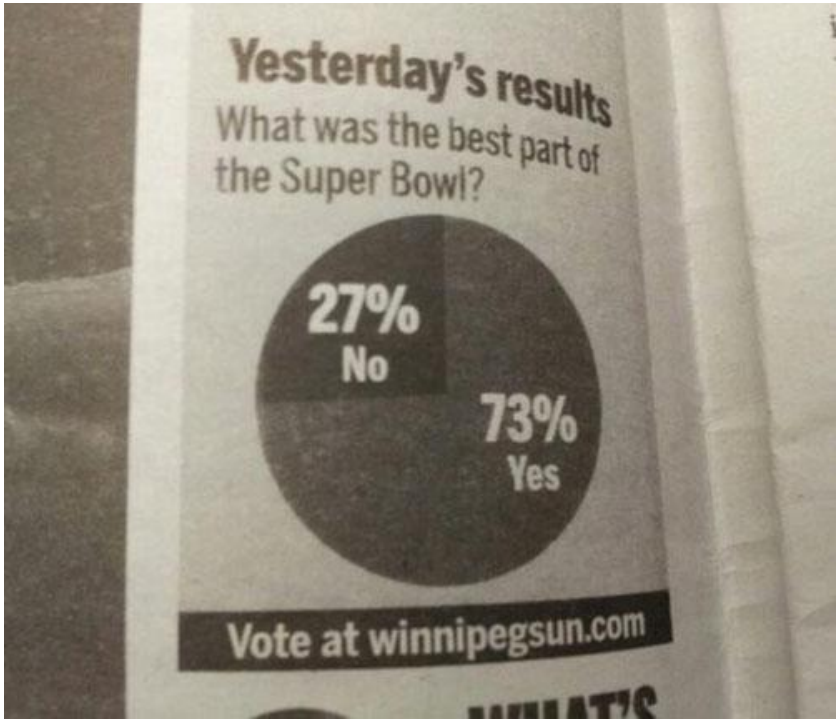
DATA JOURNALISM 101

reporting and visualizing your data like a pro

- Data journalism is journalism.
- Start with a question, not data.
- Can the data really answer the question?
- All data is dirty data.
  - How was the data collected?
  - Data is kept for the agency’s purpose, not your.
  - Read the docs.
  - Don’t believe the docs.



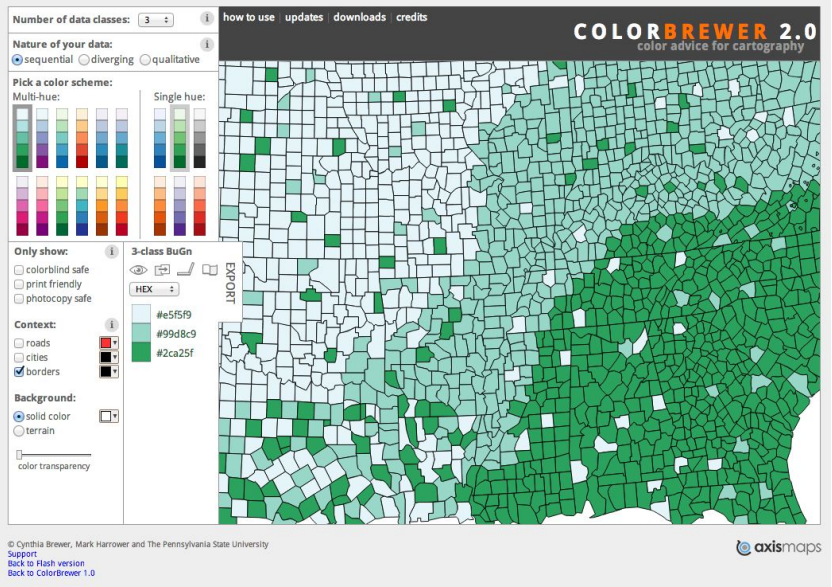
Job *

88% Broadcast reporter
84% Camera Operator
78% Columnist / Commentator
78% Editor
81% Photographer
81% Internet reporter/writer
81% Print reporter/writer
78% Producer
81% Publisher/Owner
78% Technician

VISUALIZING YOUR DATA

COLORS

Try to use no more than six colors. Using color categories that are relatively universal makes it easier to see differences between colors. The more colored categories you try to visualize, the harder it is to do this.



http://colorbrewer2.org/ is great for color scales!

COLORS

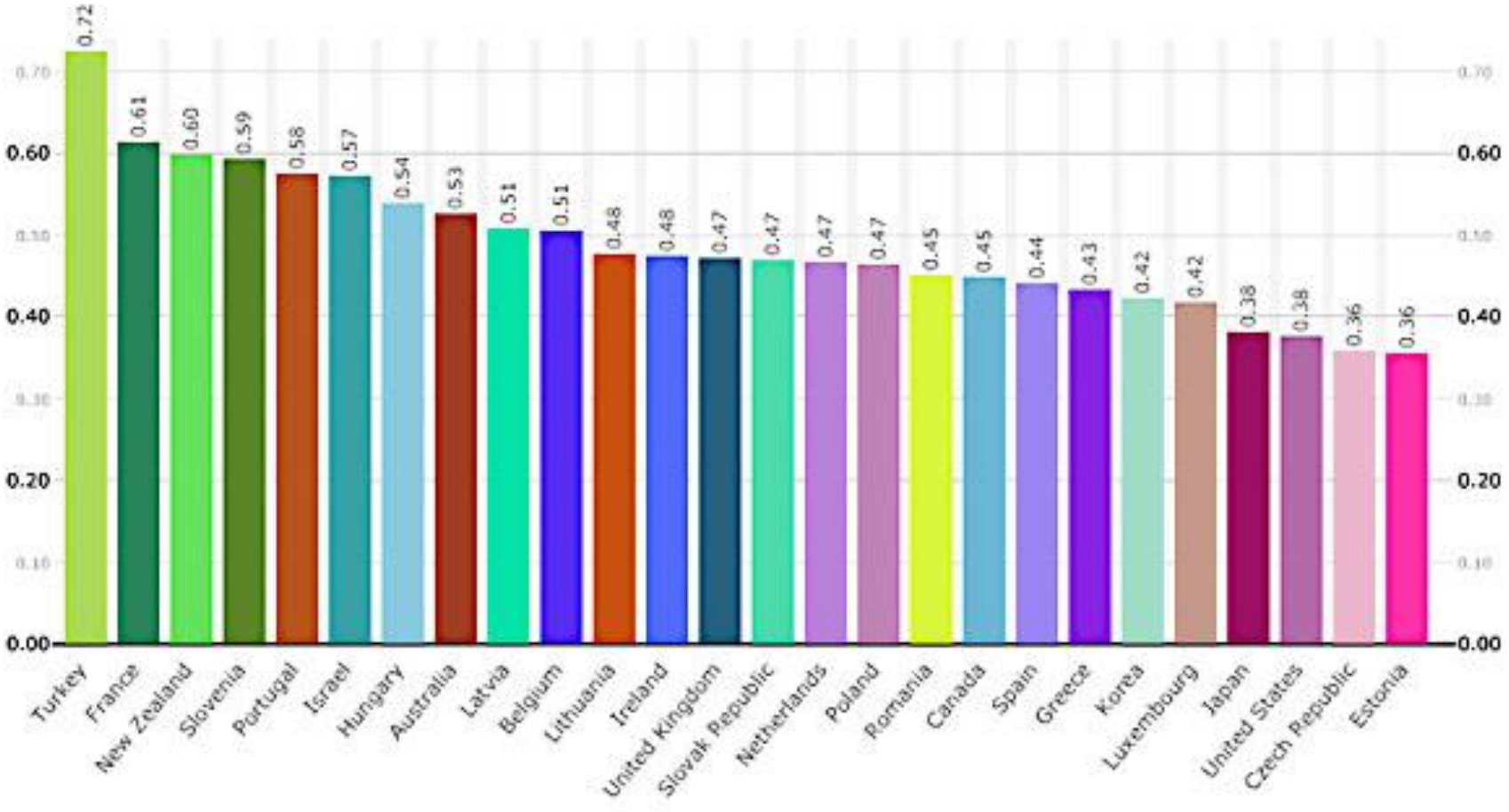
Different colors should be used for different categories/types (e.g., school type, level of government), **not different values in a range** (e.g., age, spending level).

When showing differences in rate of a single variable, stick to a single hue and vary value/lightness.

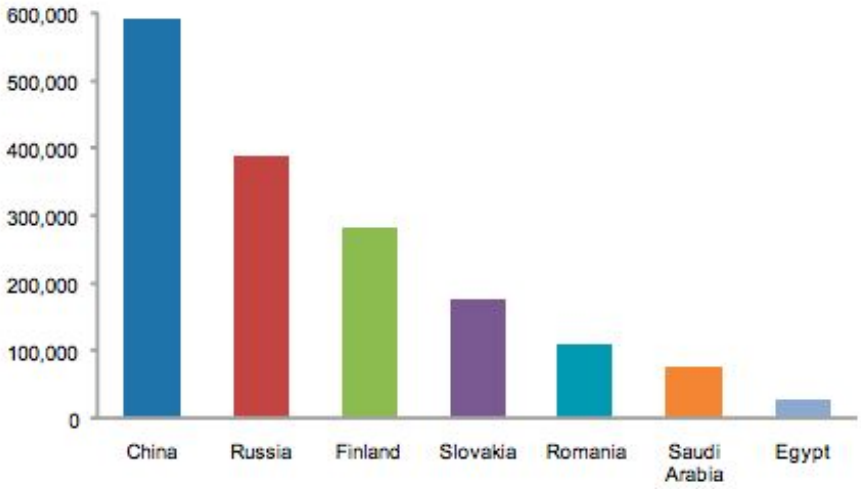
categories

single variable range

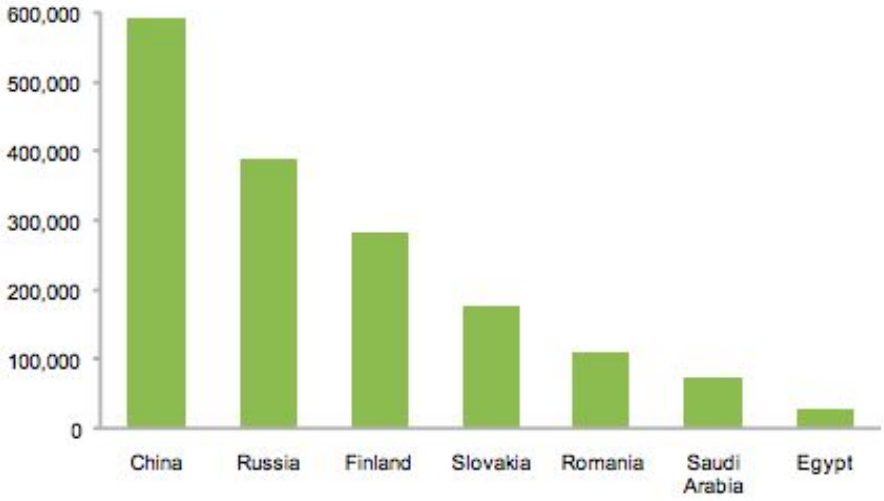
Minimum Wage as % of Median Wage



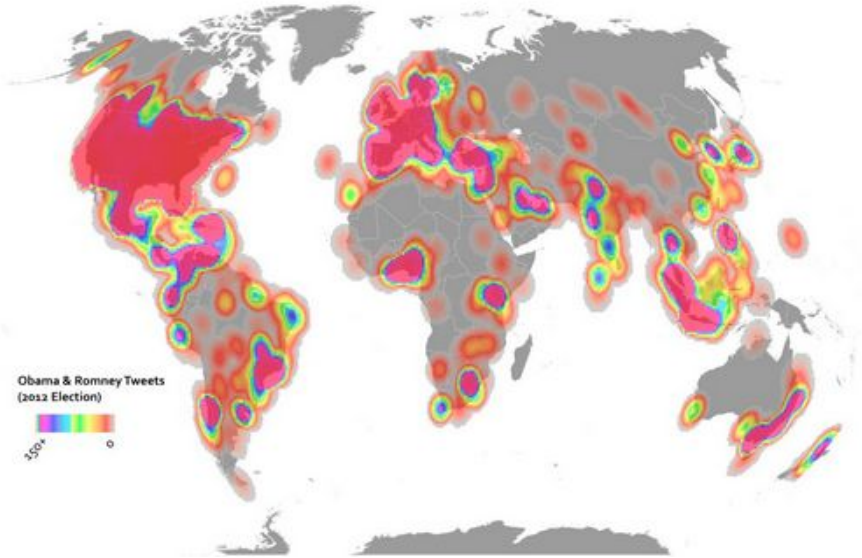
NO


600,000
500,000
400,000
300,000
200,000
100,000
0
China
Russia
Finland
Slovakia
Romania
Saudi Arabia
Egypt

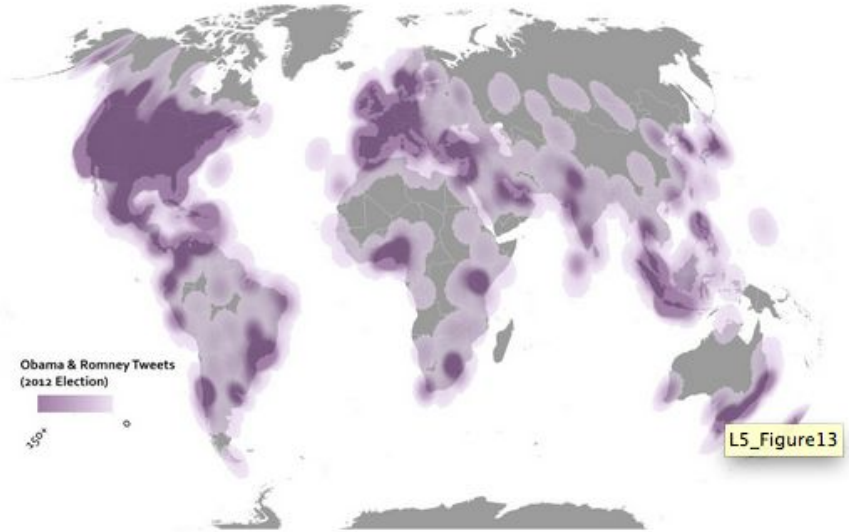
YES


600,000
500,000
400,000
300,000
200,000
100,000
0
China
Russia
Finland
Slovakia
Romania
Saudi Arabia
Egypt

NO


Obama & Romney Tweets
(2012 Election)
150+
0

YES


Obama & Romney Tweets
(2012 Election)
150+
0
L5_Figure13

MAPS

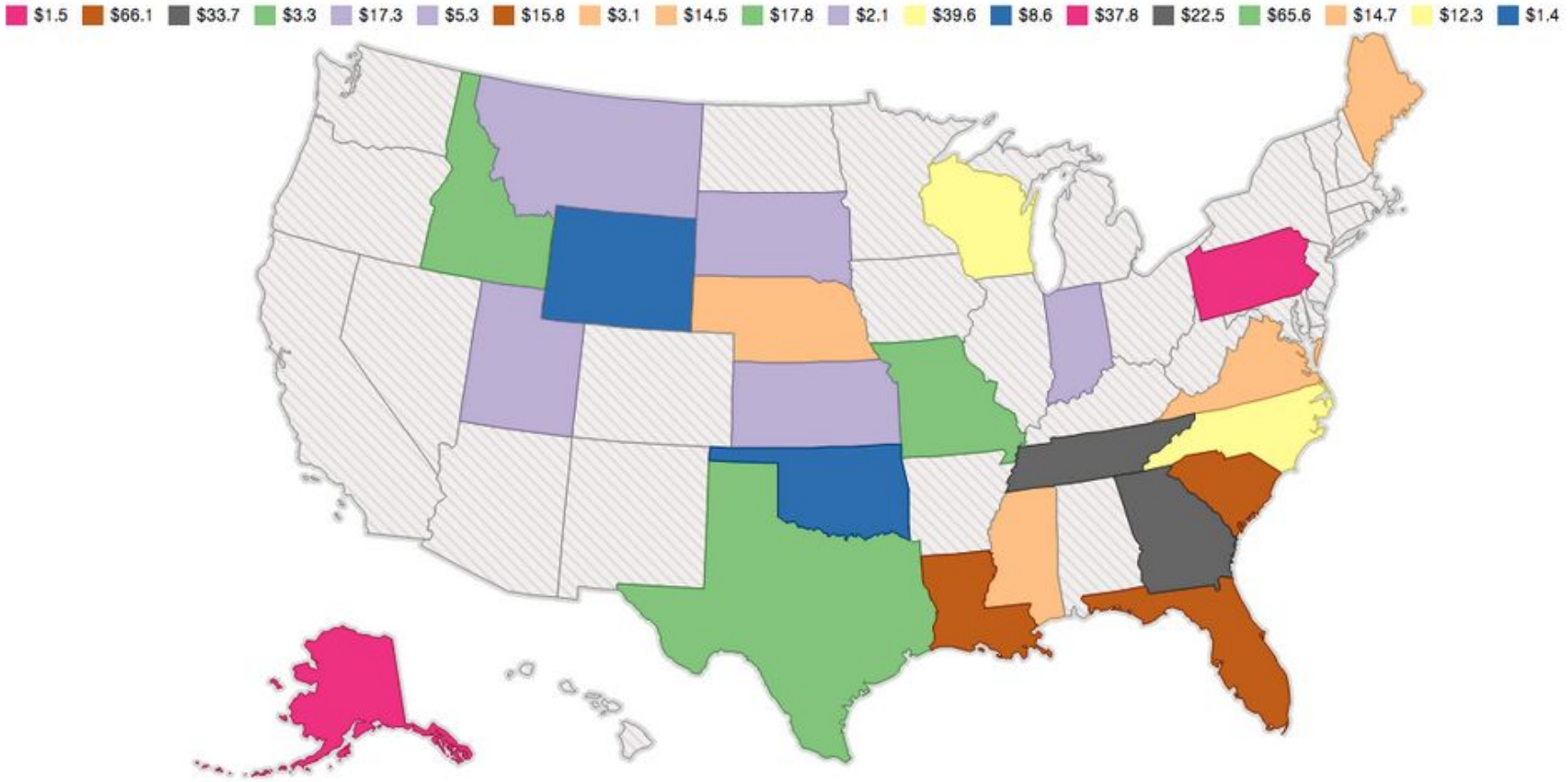
NUMBER ONE MAP RULE:

DON'T make a map just because you have location data. If the data has no geographic correlation, there's no need to show it on a map.

Before mapping your data, always consider if a table or other chart would convey the data as or more effectively.

Leaving Money on the Table

The amount of federal funds, in billions, that 24 states lost by refusing to expand Medicaid

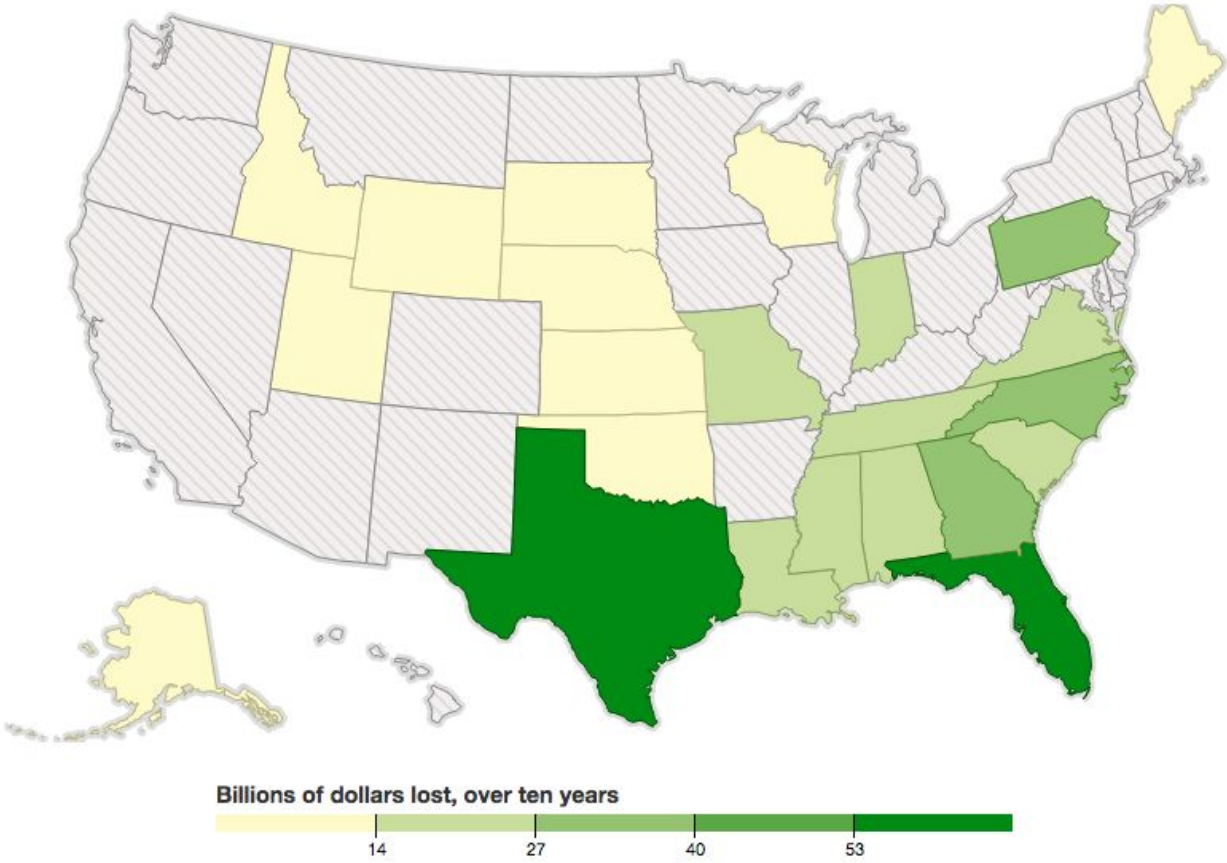


Created with Datawrapper

Source: Urban Institute, Get the data

Leaving Money on the Table

The amount of federal funds, in billions, that 24 states lost by refusing to expand Medicaid



Created with Datawrapper

Source: Urban Institute, Get the data

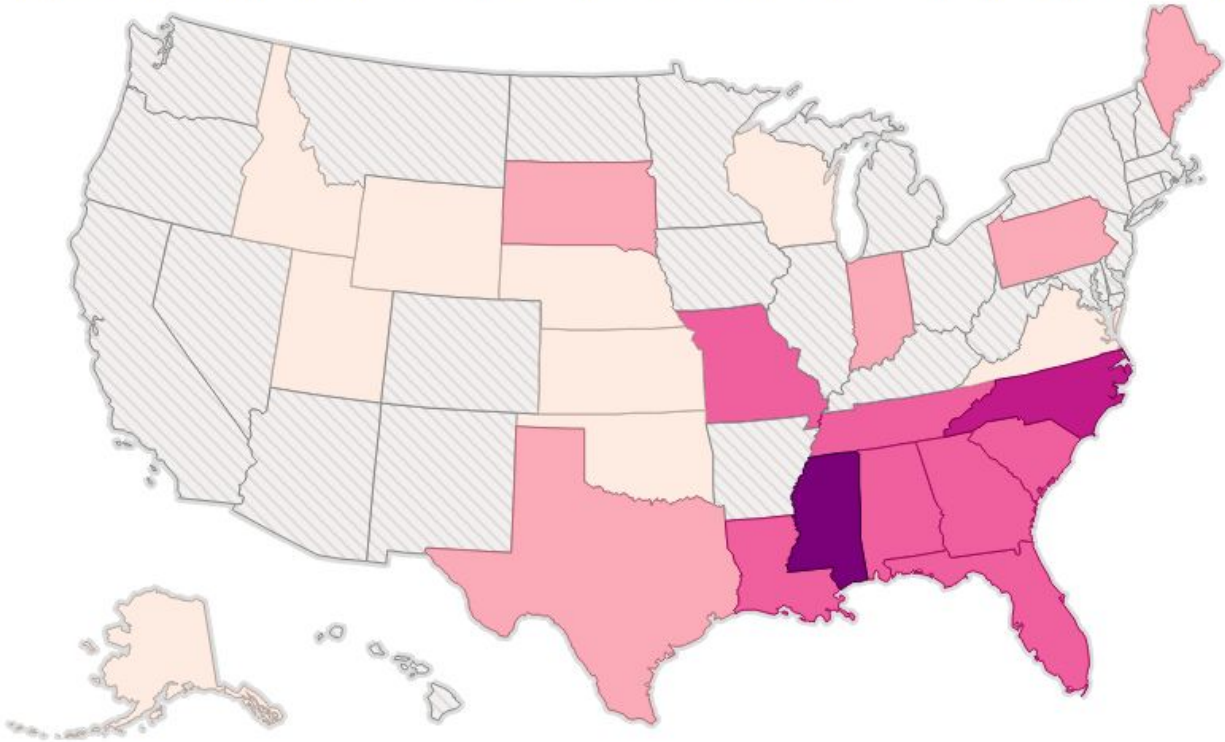
MAPS

All maps must be distinguishable from population maps. To accomplish this, the underlying data should control for population or other denominator.

For example, on a choropleth map of Georgia counties, metro Atlanta counties shouldn't be extremely dark in comparison to the rest of the state. That's a population map.

Leaving Money on the Table

On a per capita basis, here's how much money states are giving up over the next ten years if they continue to refuse Obamacare's Medicaid expansion—and the federal dollar that come with it



Created with Datawrapper

Source: Urban Institute, calculations by QED, Get the data

MAPS

DON'T use a map to make a pie chart, or as decoration for your visualization if it doesn't provide any geographic significance.

GGC'S DIVERSE STUDENT BODY REFLECTS THE DIVERSITY OF GWINNETT COUNTY ATLANTA METRO REGION

WHITE NON-HISPANIC 42%

OTHER 5%

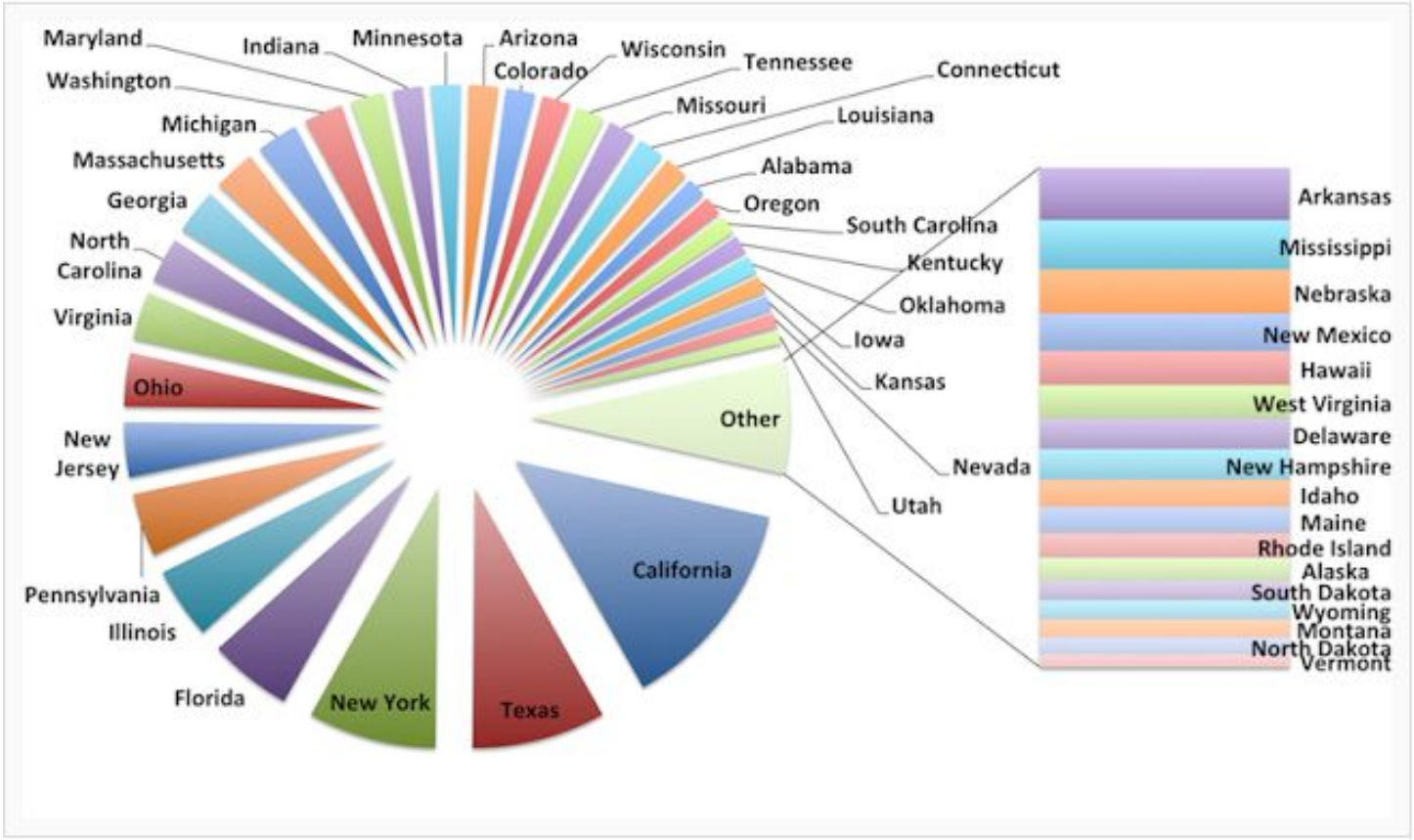
ASIAN 9%

HISPANIC/LATINO 14%

29% BLACK AFRICAN AMERICAN

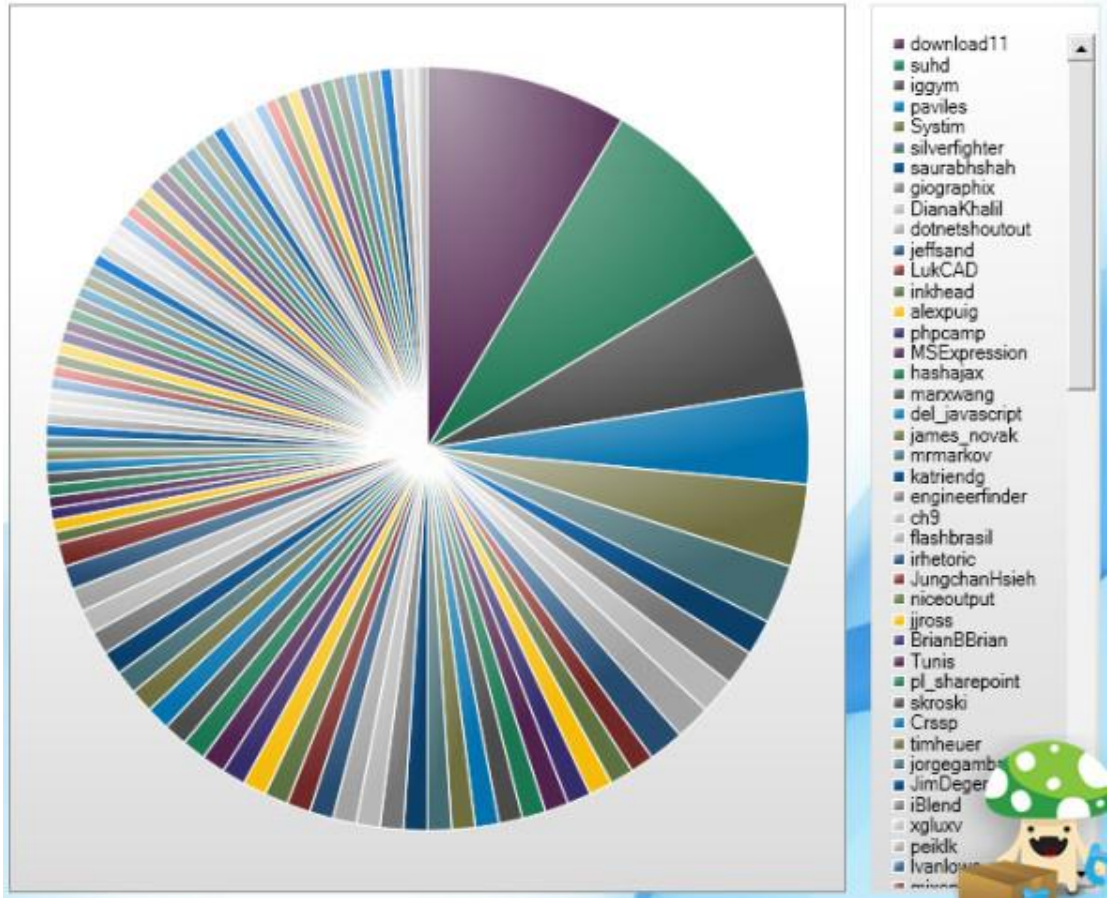
CHART CHOICE

PIE CHARTS: Only use a pie chart to show part of a whole (i.e. no more than two values).

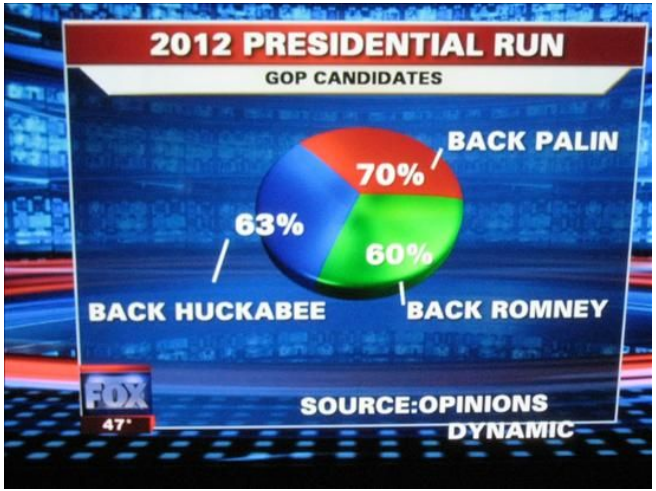


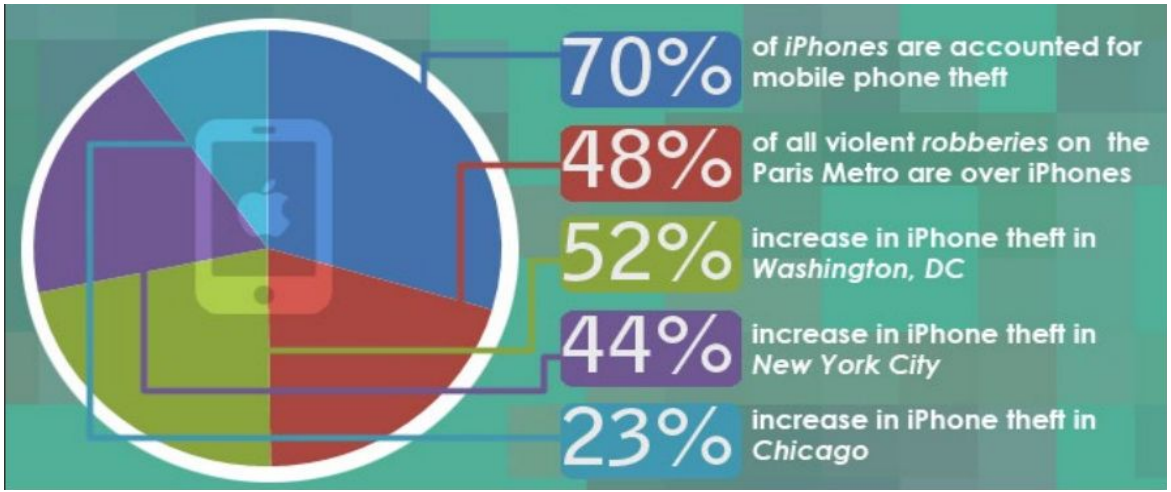
The Unites States GDP for individual states as a contribution to the total US GDP. Fraction of the total US GDP per state was taken from Wikipedia and refers to 2010.

100 Most Active Tweeters



Percentages don't always need to be visualized in a pie chart!





Try another type of chart, like a bar chart!

CHART CHOICE

BAR CHARTS: Use a bar chart to show discrete quantities. Use a stacked bar chart to show the relative values of categories and subcategories of data.

For bar charts, the numerical axis (often the y axis) **must start at zero.**

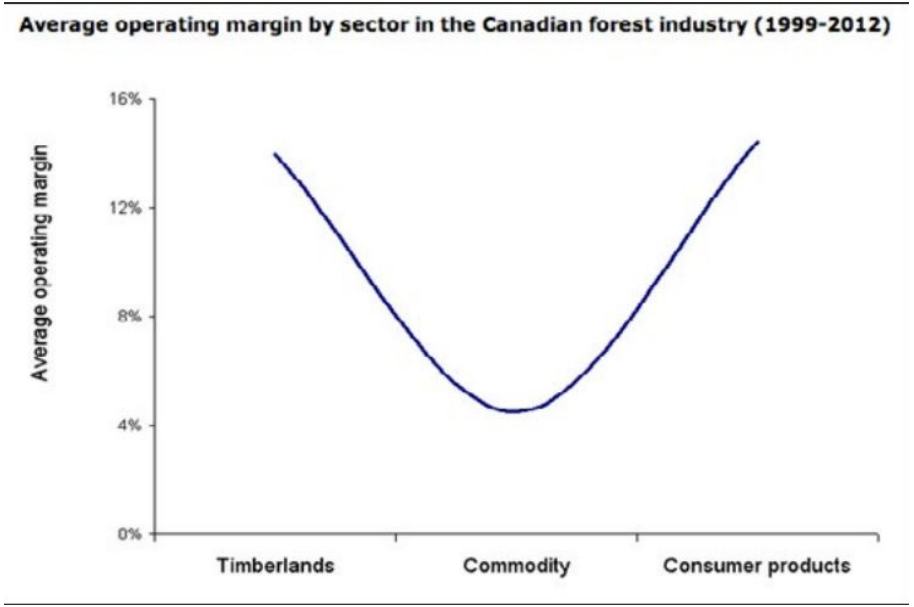
CHART CHOICE

LINE/FEVER CHARTS: Use a line chart to show continuous variables such as time series.

DON'T use a line chart when time isn't a variable.

Continuous data → line chart

Categorical data → another chart



MOBILE

THINK SMALL

http://mobilev.is/

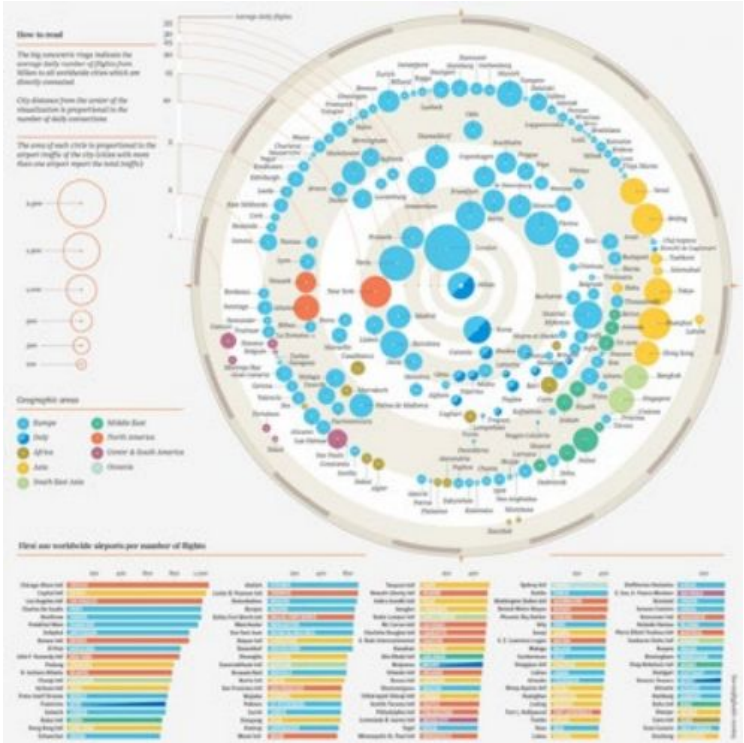
Think vertically, not horizontally.
Especially in tableau. Your charts will not stack, they will only shrink in width.

Pete's story method

BEST PRACTICES

KEEP IT SIMPLE!

- Just because you have a lot of data and a lot of information doesn't mean it all needs to be directly visible in your visualization.
- Data viz is just another vehicle for storytelling – not every piece makes the cut for the story. Use the parts you need to support the hypothesis.
- If you have a lot of data that needs to be shared, think about using a table instead of a visualization



BEST PRACTICES

DON’T MAKE THE USER WORK TOO HARD.

- Your visualization should make the data easier to understand, not just present the data for the user to figure out. If a chart makes it hard to understand an important relationship between variables, do the extra calculation and visualize that as well.
    - This includes using pie charts with wedges that are too similar to each other, or bubble charts with bubbles that are too similar to each other.
- Get feedback from others on the ease of understanding your visualization before publishing. You may think your chart makes perfect sense, but you’ve also been looking at it for a long time. Another set of eyes is always a great idea.

And most importantly...

Just because you have data doesn’t mean you need a data visualization for your story!