

Principles of Visualization

123 GO – what's your word of the day?

Announcements:

- Assignments will be more open ended, and you can choose how deep to go.

Outcomes:

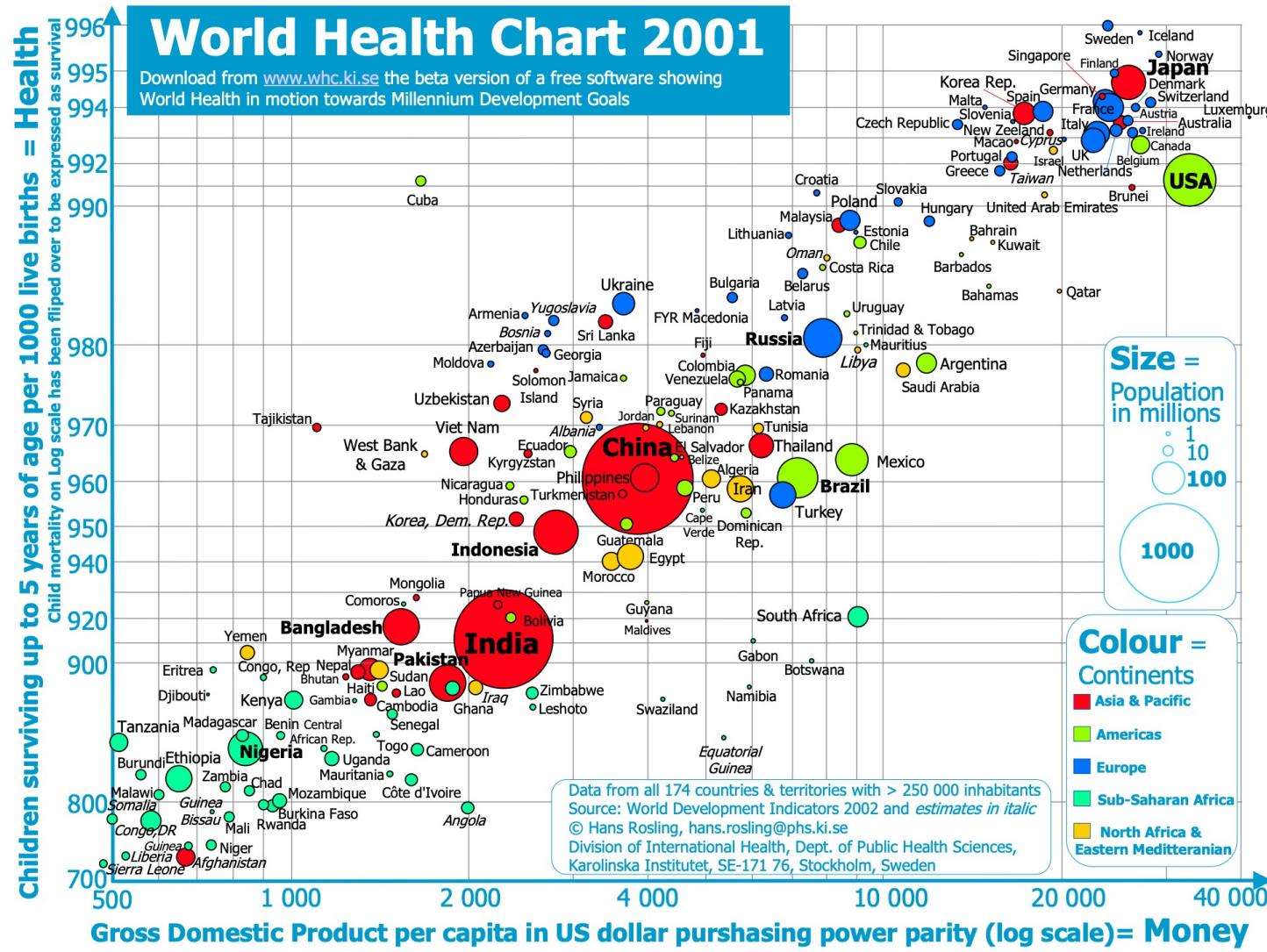
- Identify elements of plots
- Create clear, meaningful statistical graphics
- Contrast EDA with presentation quality graphics

Source: https://www.textbook.ds100.org/ch/06/viz_intro.html

Statistical graphics have two main purposes.

Audience	Time to produce	Purpose
Exploratory Data Analysis (EDA)		
Publication Quality		

Vocabulary - Parts of a plot

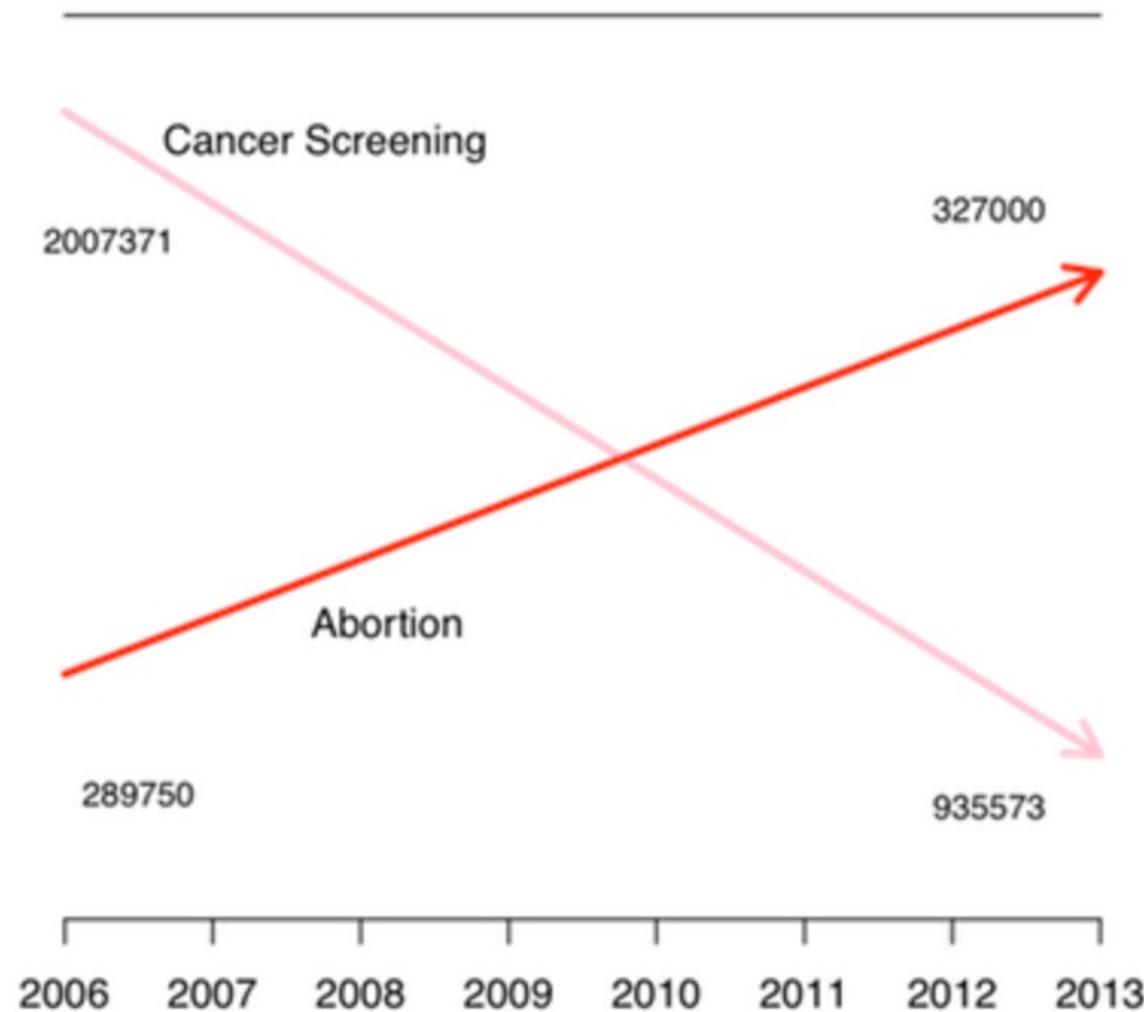


Source: Michael Friendly
<http://www.datavis.ca/gallery/>
Hans Rosling

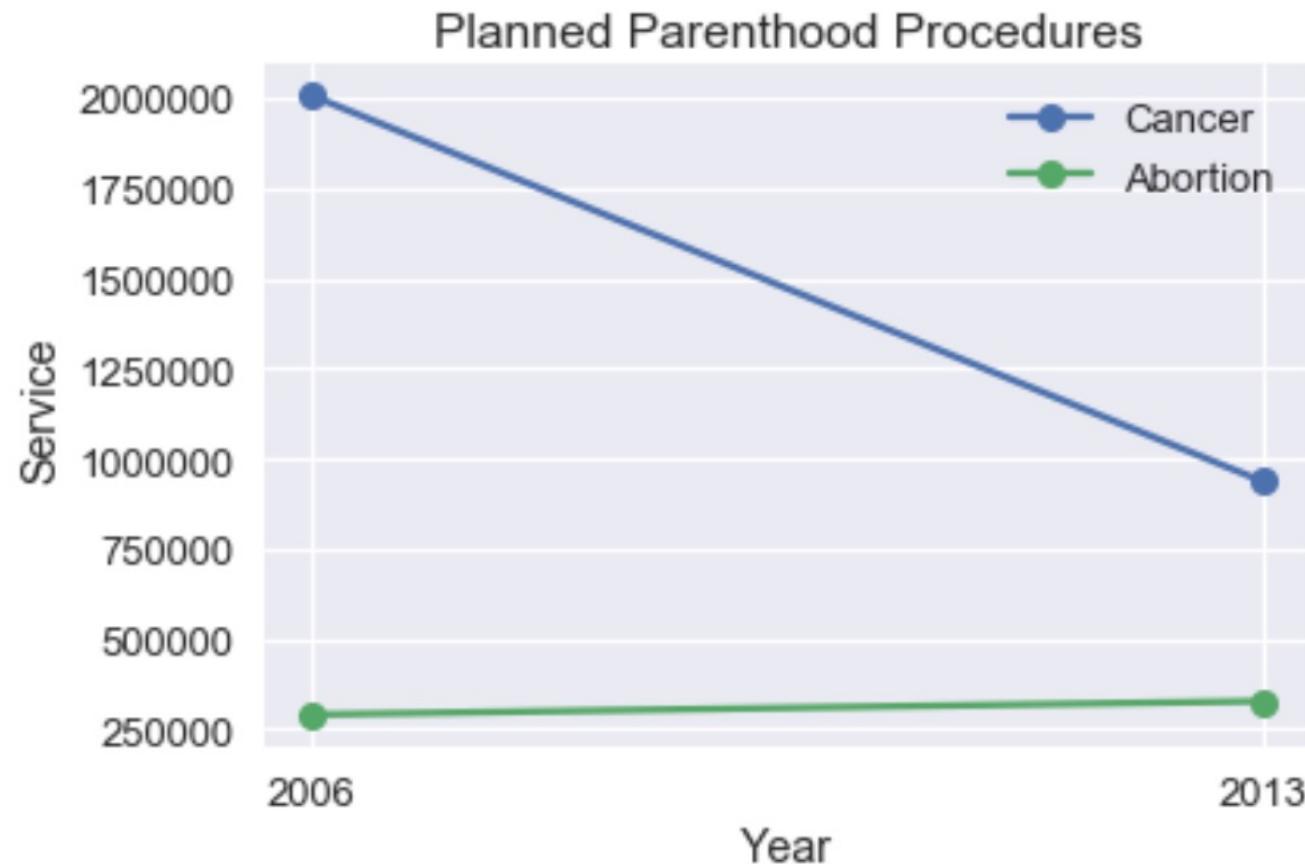
In a 2015 US Congressional hearing, representative Chaffetz discussed an investigation of Planned Parenthood programs. He presented the following plot that originally appeared in a report by Americans United for Life. It compares the number of abortion and cancer screening procedures, both of which are offered by Planned Parenthood. (The full report is available at <https://oversight.house.gov/interactivepage/plannedparenthood>.)

What is suspicious about this plot? How many data points are plotted?

Color



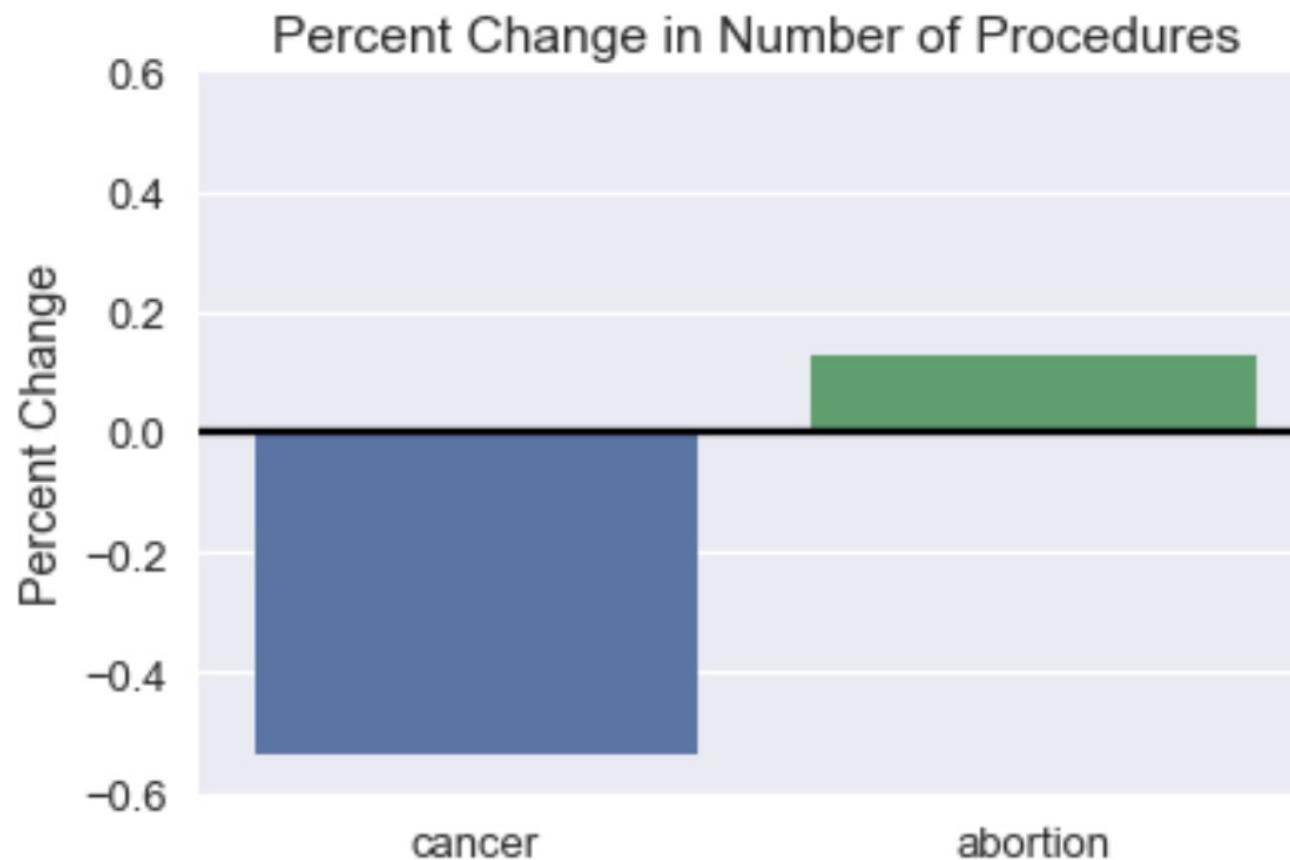
Does this plot show the pattern in the data?



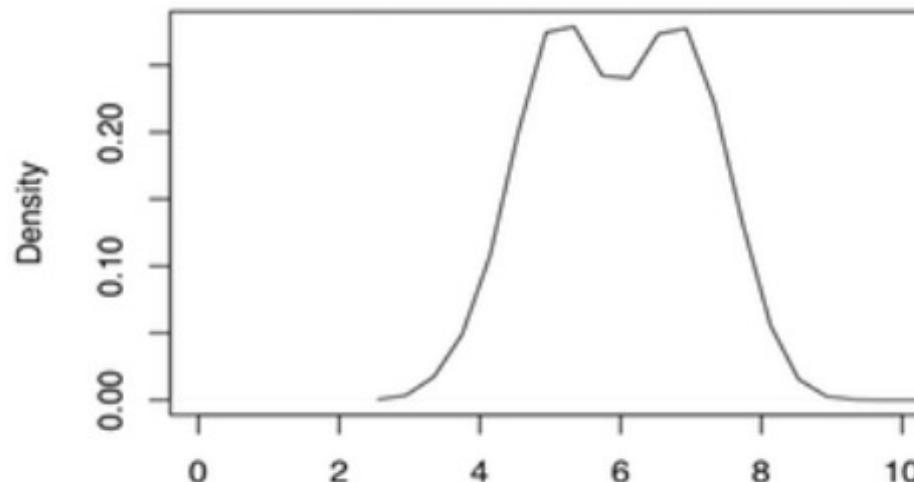
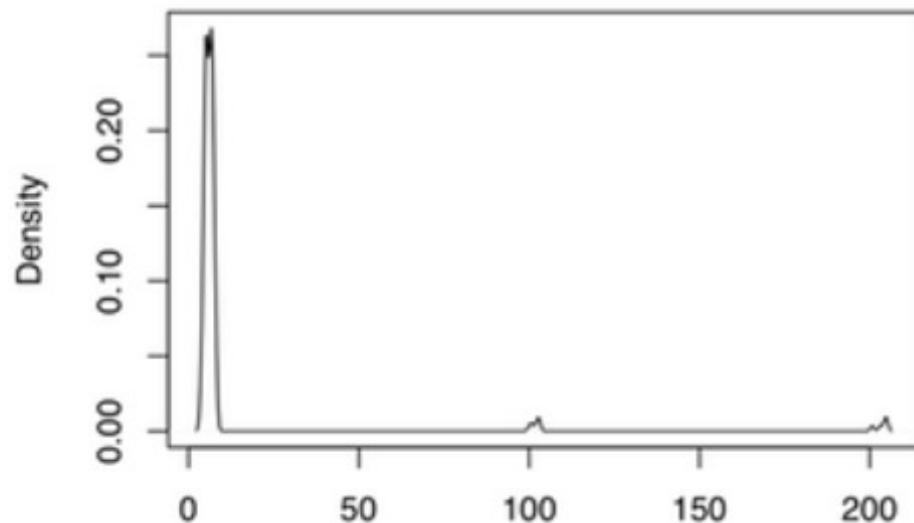
Choose the correct scale.

Relative scales allow comparison of numbers on different orders of magnitude.

Percent?

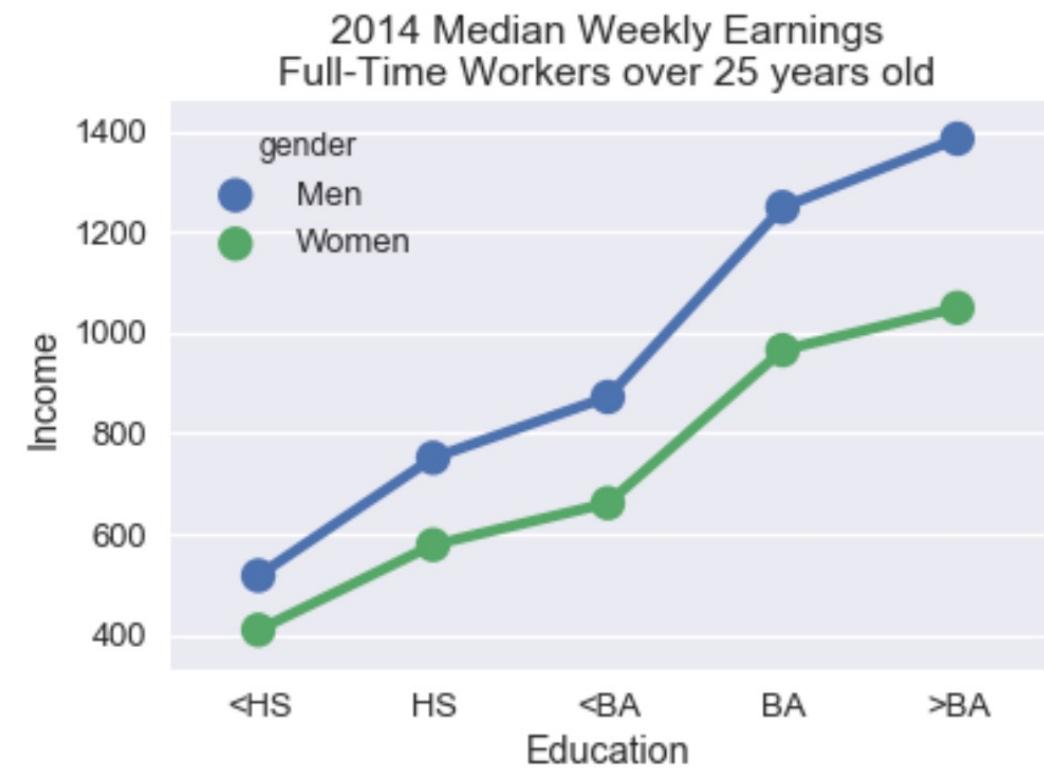
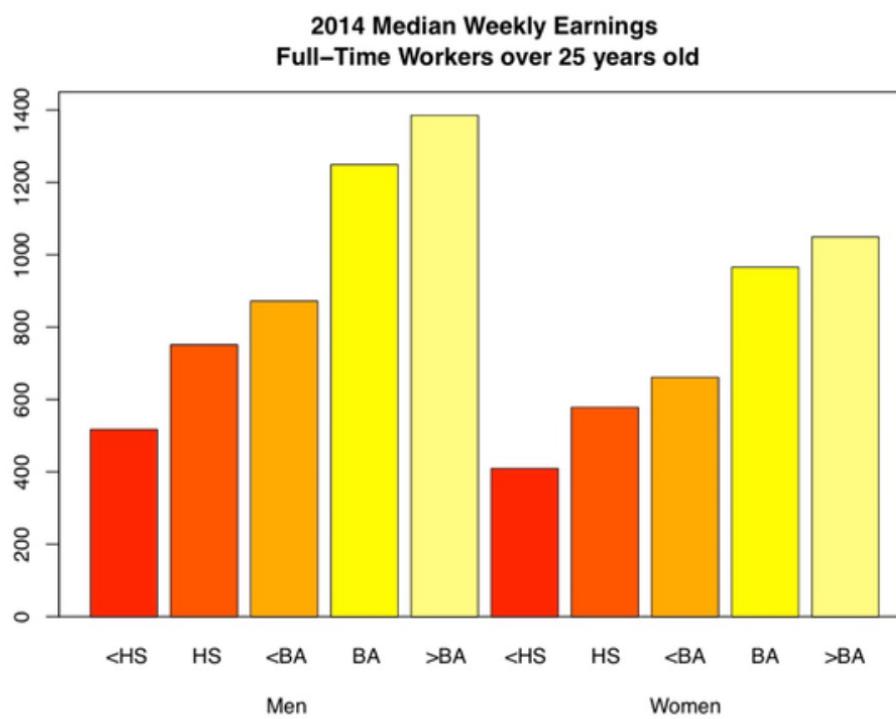


Focus on the area of interest.



Group subsets of the data to compare them.

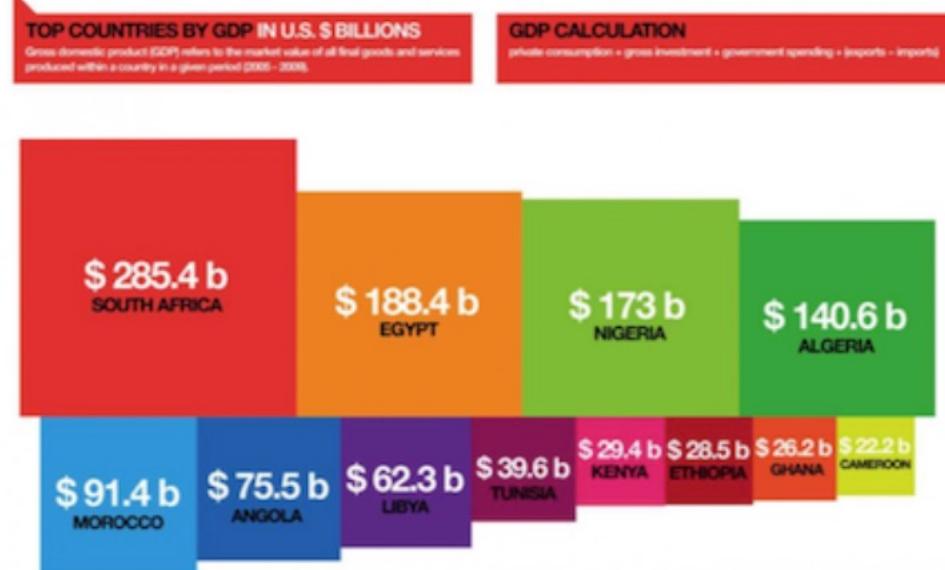
Semantically meaningful color



Human eyes are better at comparing lengths than angles or areas.

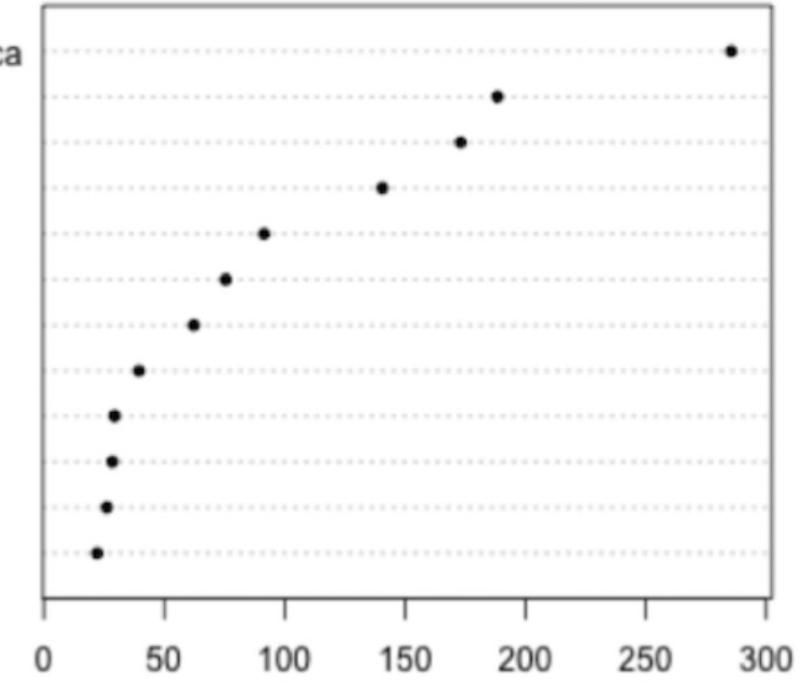
Compare Algeria and South Africa

African Countries by GDP

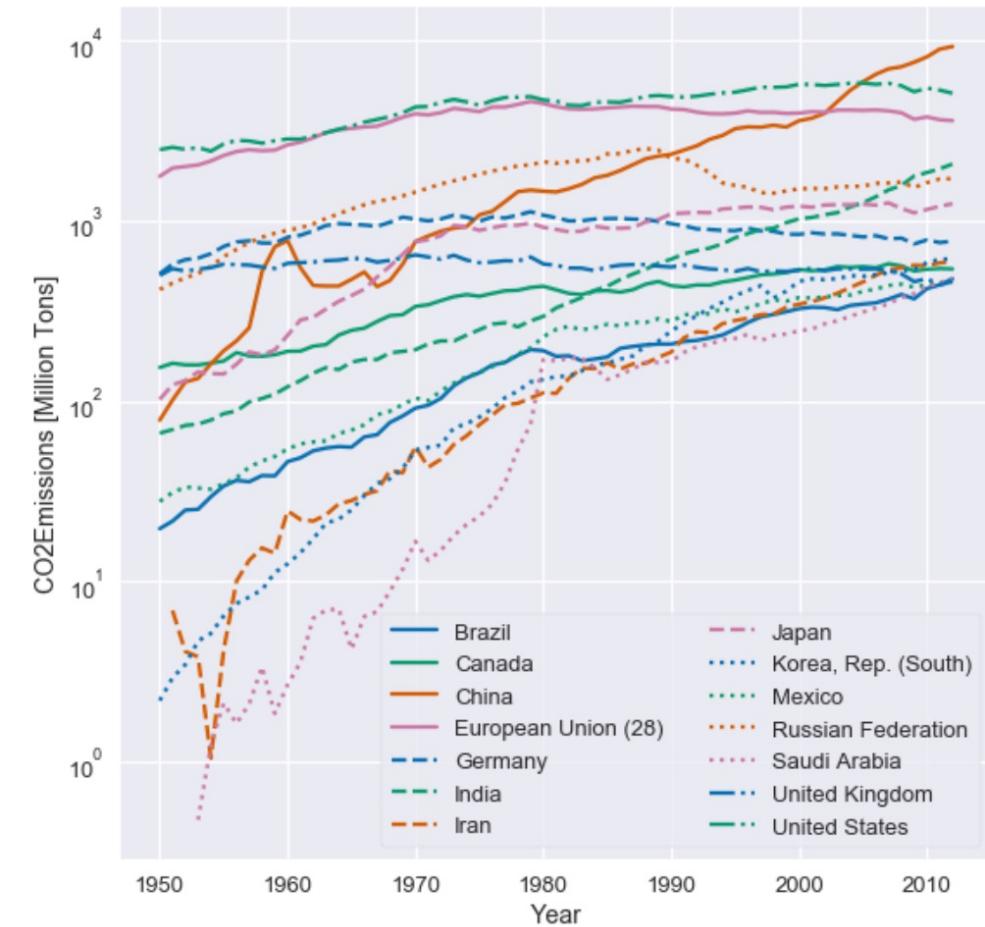
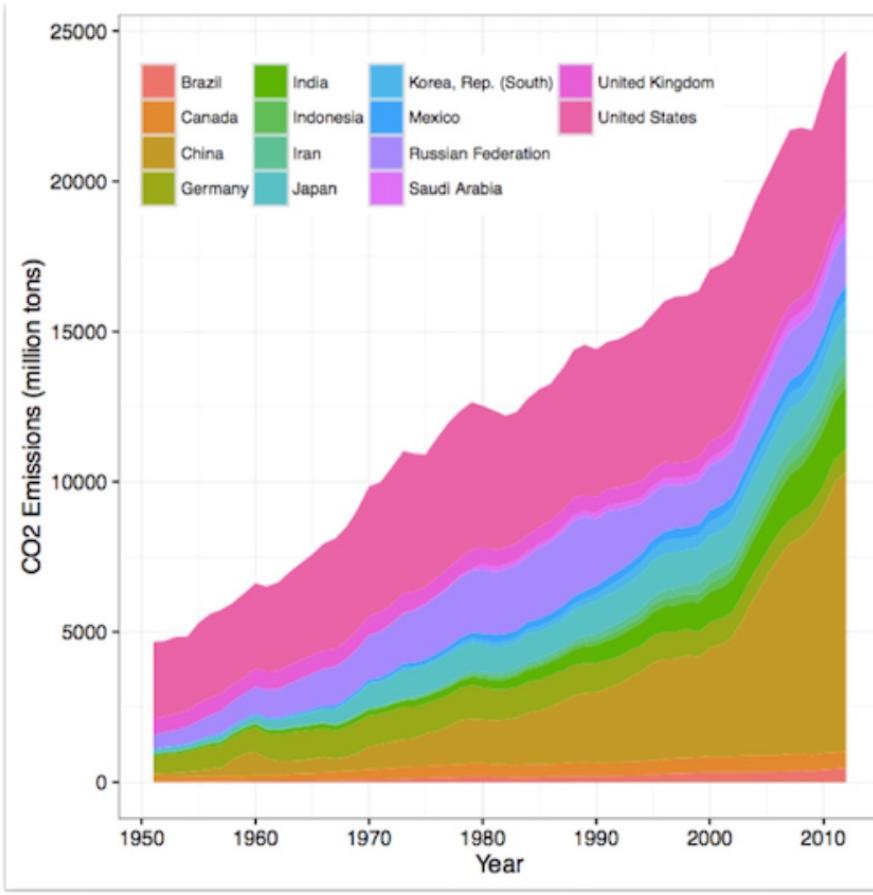


African Countries by GDP

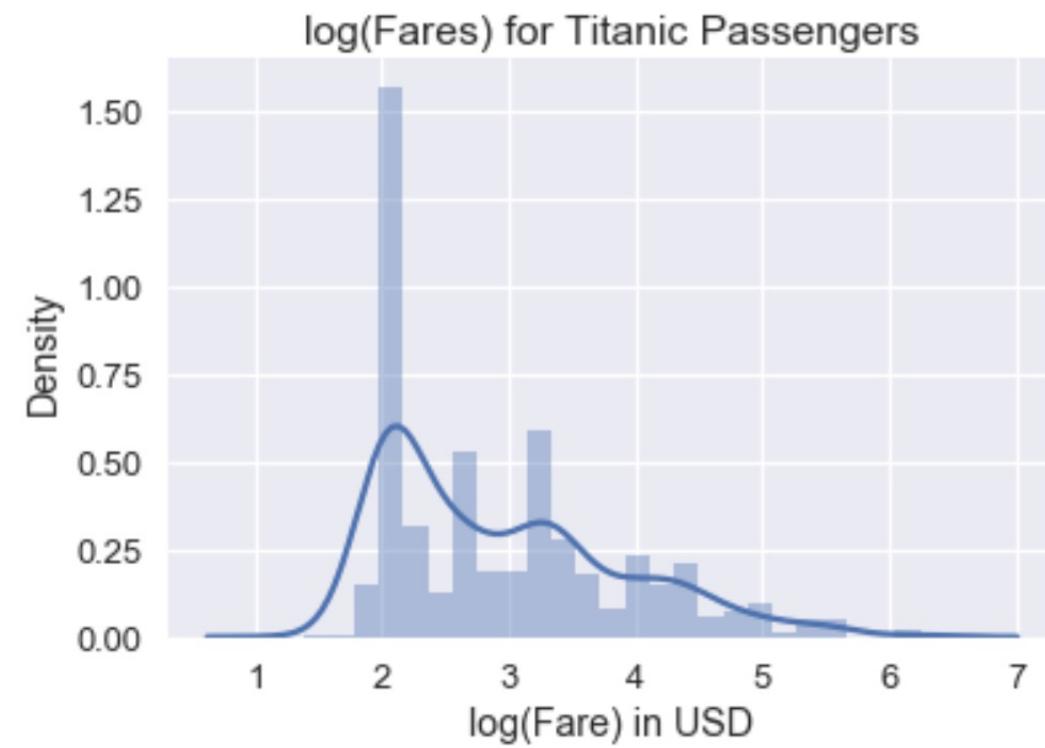
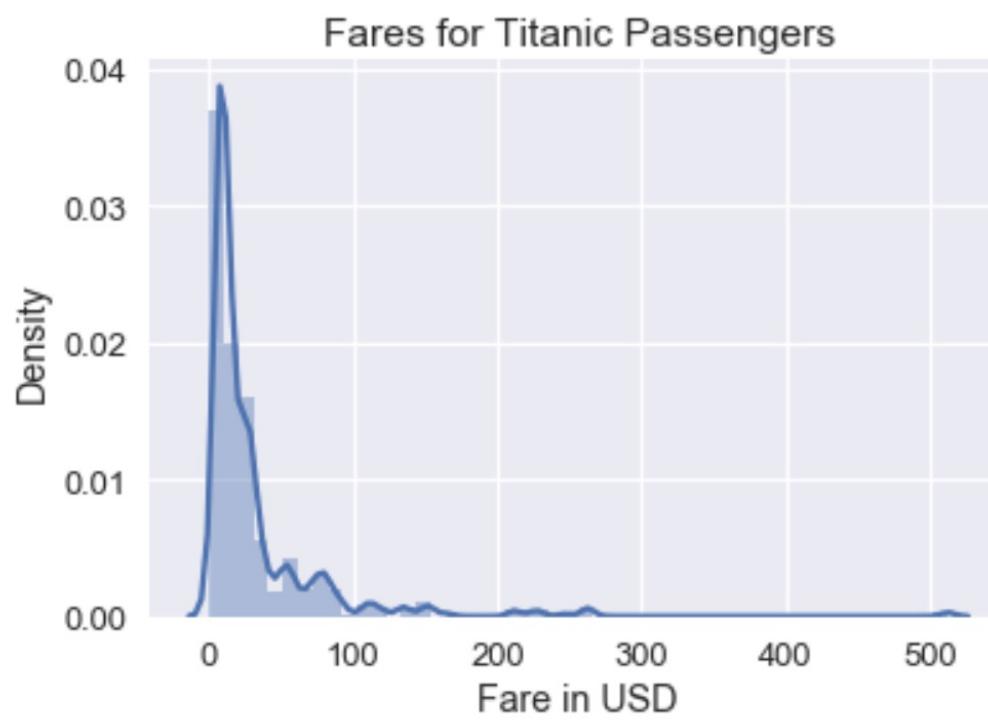
South Africa
Egypt
Nigeria
Algeria
Morocco
Angola
Libya
Tunisia
Kenya
Ethiopia
Ghana
Cameroon



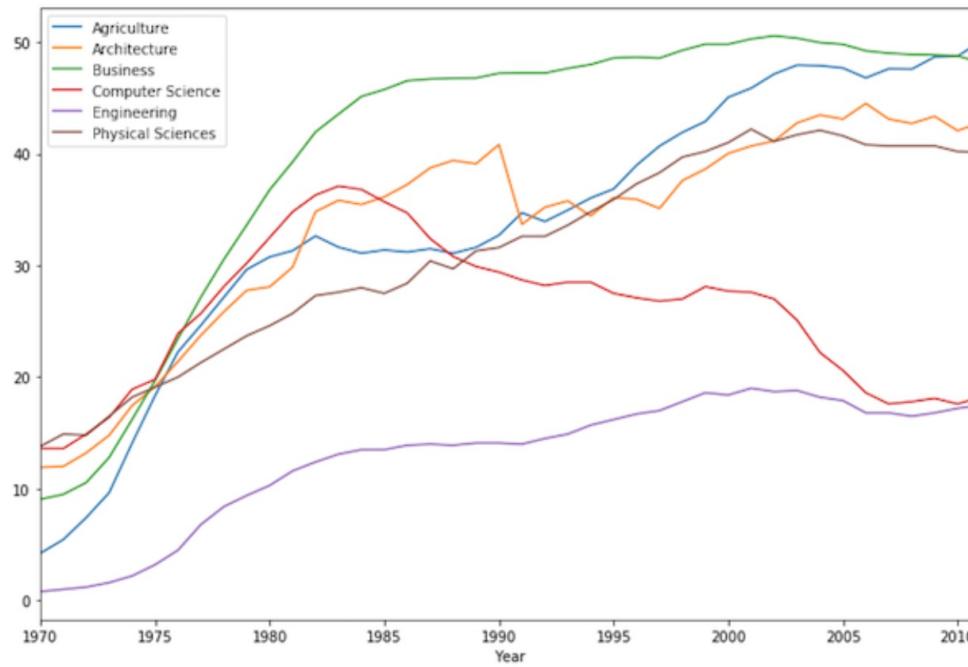
Keep the same baseline.



Consider log transformations for skewed data.

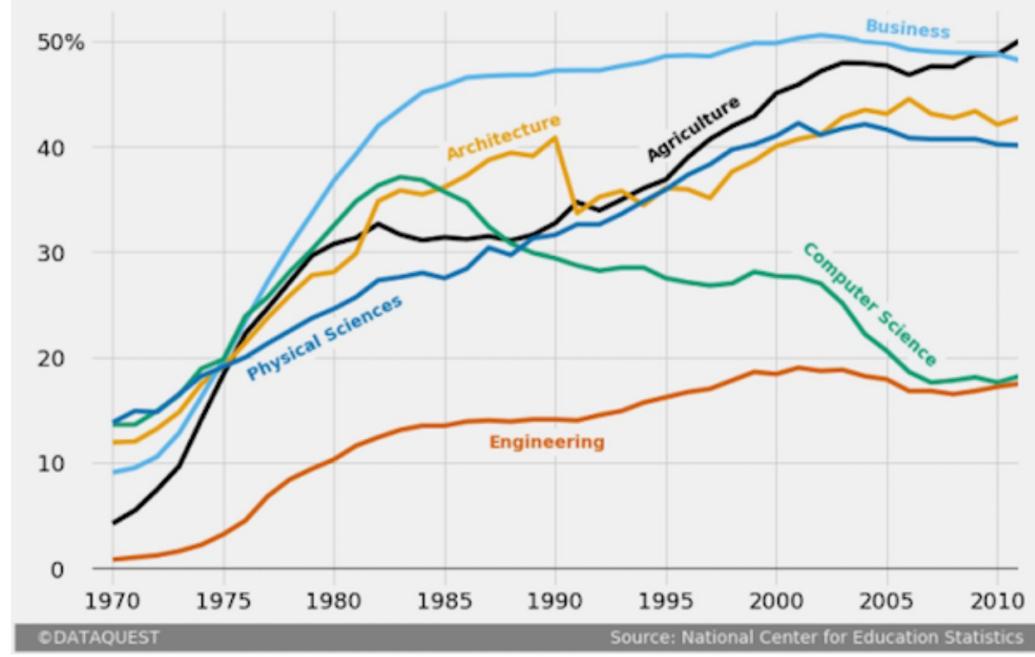


Add context through titles, labels, callouts, and captions.

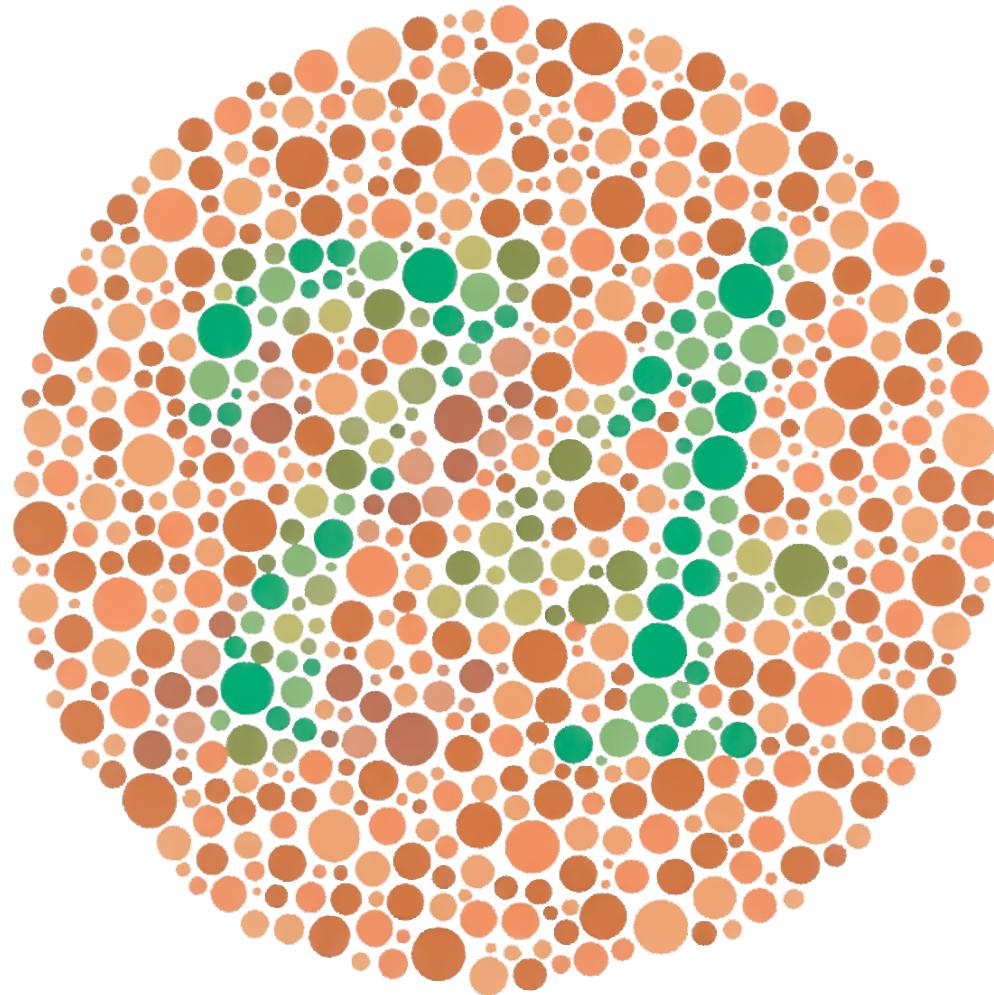


The gender gap is transitory - even for extreme cases

Percentage of Bachelors conferred to women from 1970 to 2011 in the US for extreme cases where the percentage was less than 20% in 1970

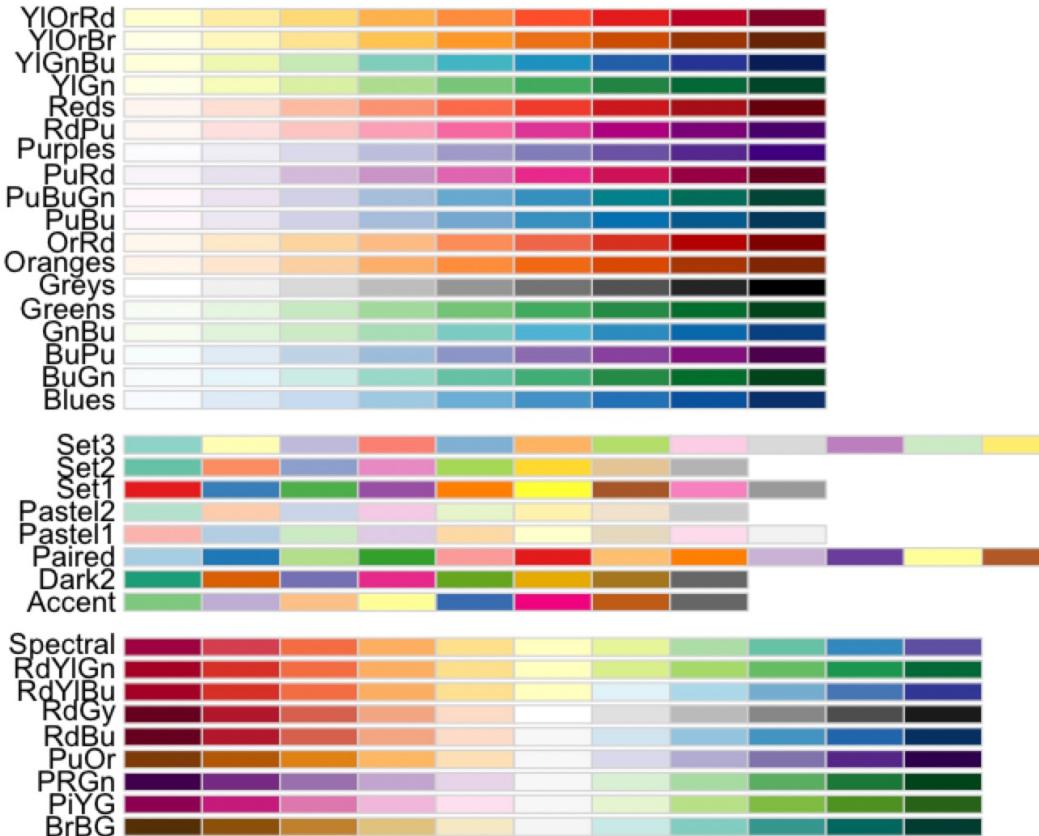


Choose colors easy to perceive.

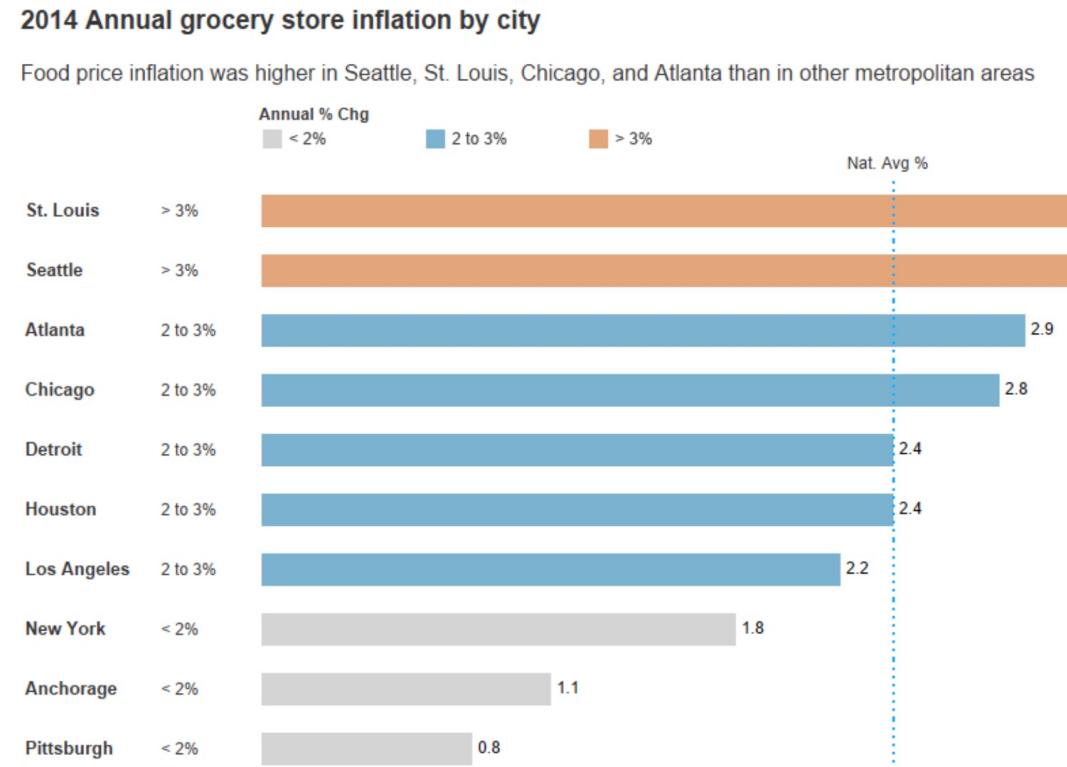
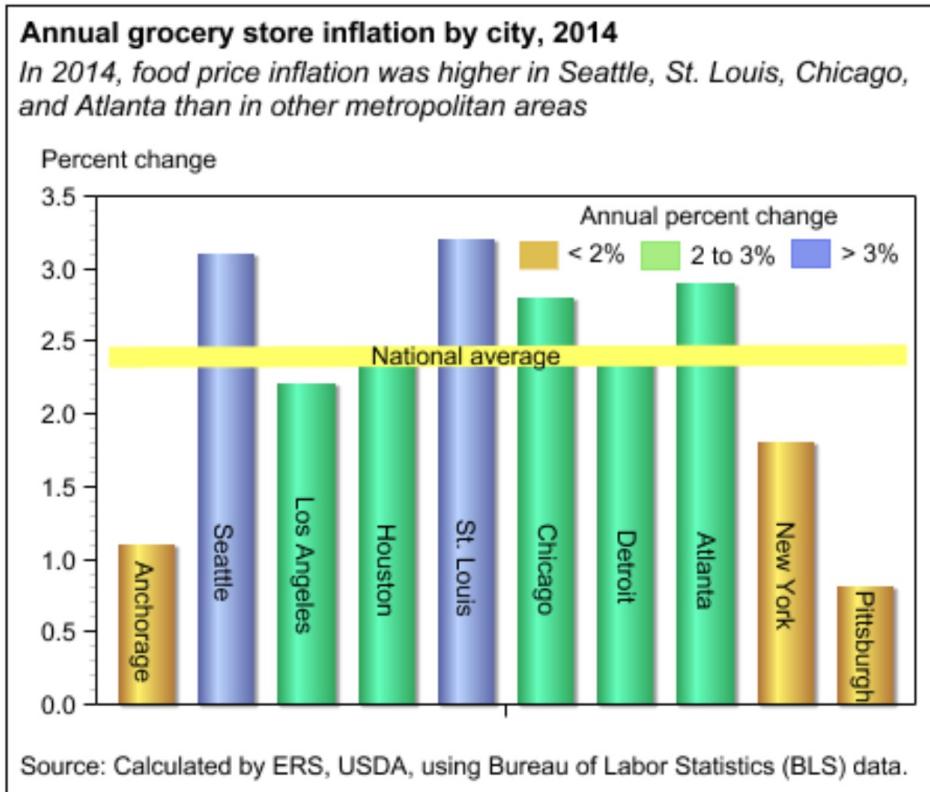


The Brewer color palettes work well.

```
library(RColorBrewer)  
par(mar=c(3,4,2,2))  
display.brewer.all()
```



Every part of visualization should contribute to understanding.

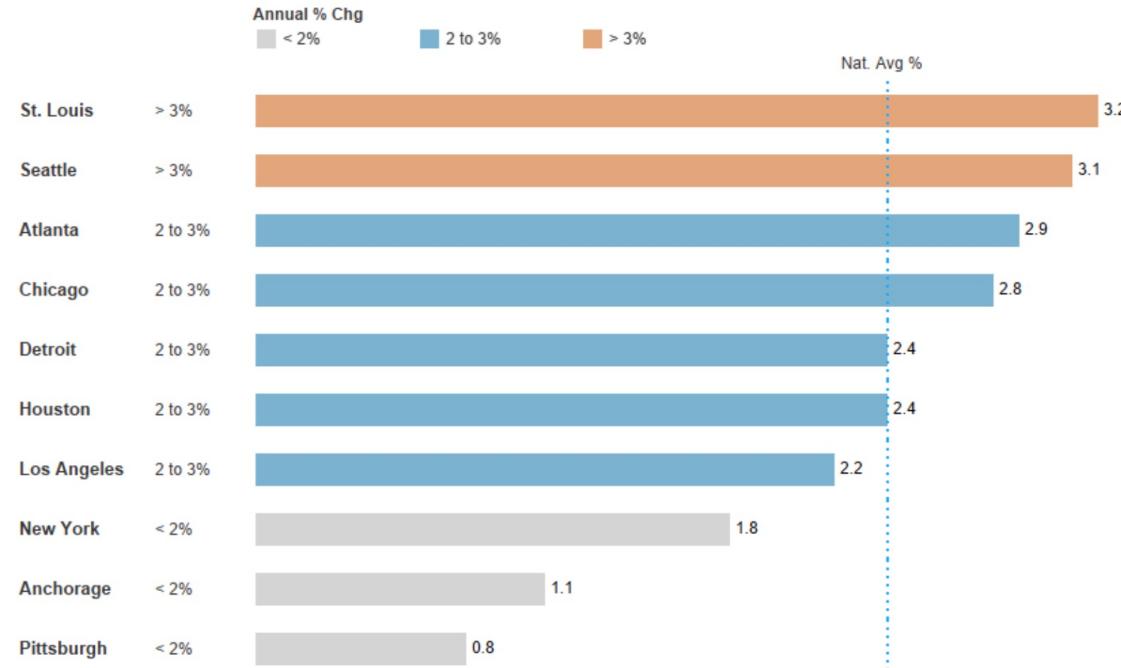


Source: <https://www.bluegranite.com/blog/data-visualization-remove-chart-clutter-and-focus-on-the-insights>
I would go further...

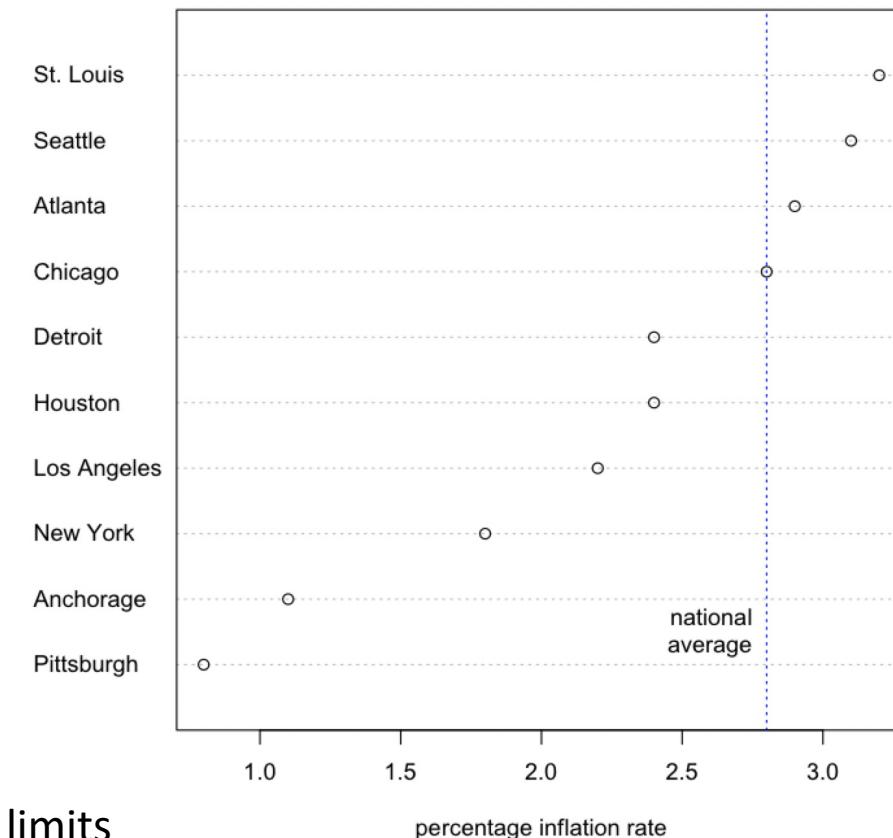
Keep your graphics “as simple as possible, but no simpler.”

2014 Annual grocery store inflation by city

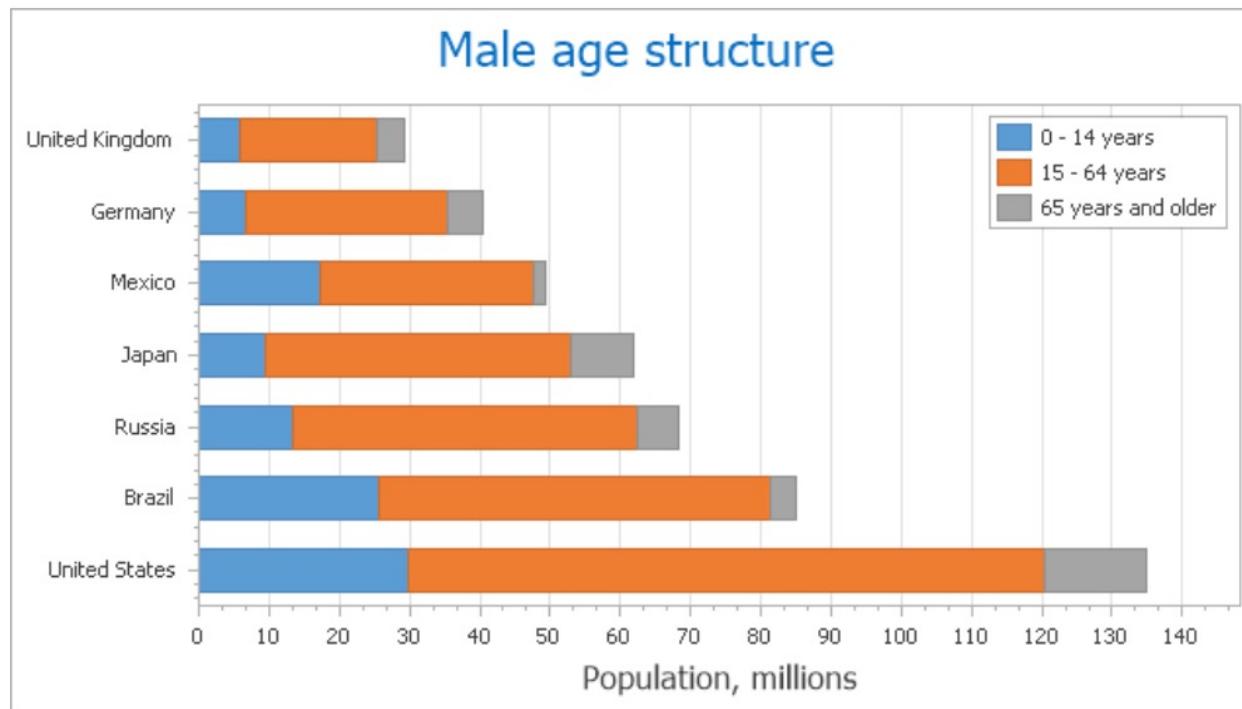
Food price inflation was higher in Seattle, St. Louis, Chicago, and Atlanta than in other metropolitan areas



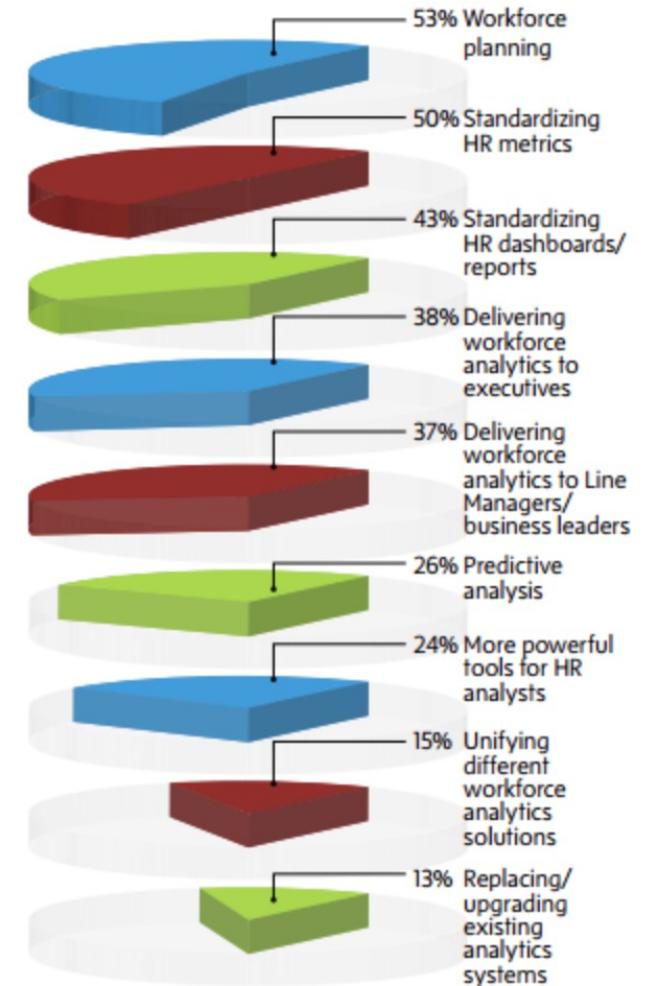
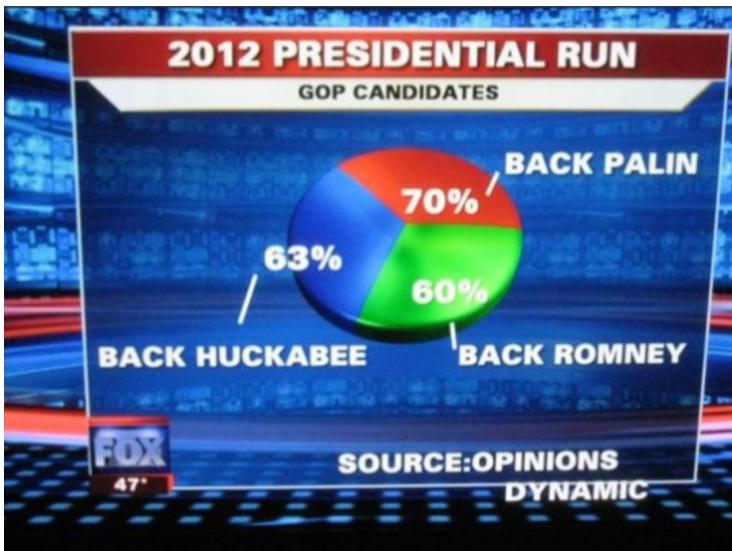
2014 annual grocery store inflation by city



How could we improve this graphic?



Just for fun 😊



Source: <https://flowingdata.com/2009/11/26/fox-news-makes-the-best-pie-chart-ever/>