Lecture 2: Introduction to Data - Graphics Principles - Critiquing Statistical Graphics

Michael Lopez, Skidmore College

What is data?

How do we describe data?

Two measurements to describe datasets

Matrix form:

Types of Data

Categorical:

Continuous:

Graphics and Their Goals (via Edward Tufte, Alberto Cairo)

Graphics: visually display measured quantities by combining points, lines, coordinate system, numbers, symbols, words, shading, color

Goals:

- make large, complicated datasets more coherent
- encourage comparisons
- reveal levels of detail
- describe, explore, tabulate, or decorate
- be closely integrated with statistical/verbal descriptions of dataset

Warning signs

1. Data-ink ratio
2. Distortion
3. Lie factor
4. Truth-continuum

Data-ink ratio

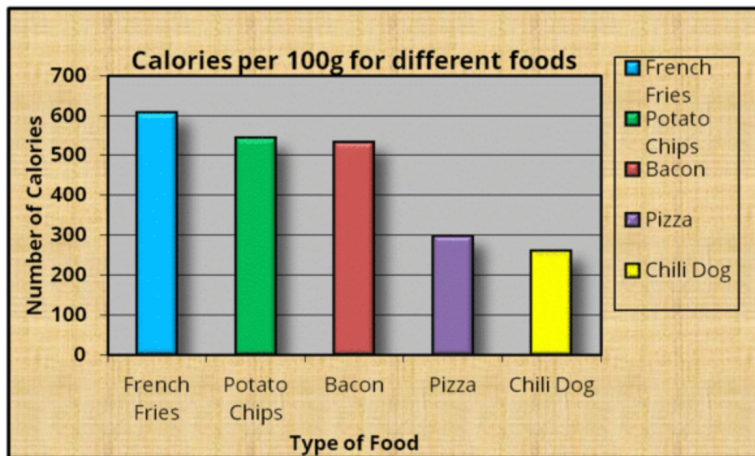Extras - anything that `gets in the` way of viewing data

Note: Clear and concise labels

Note: Data-ink should primarily present information from the data: the non-erasable, non-reduntant core part of the graphic

Data-ink ratio:

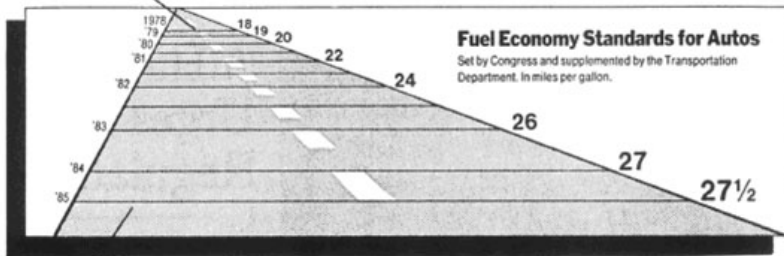# Data-ink ratio
Remove non-data-ink, remove redundant data-ink (click for link)

Figure 1:Data-ink

Distortion

Visual representation of data is inconsistent with numerical representation:
The graph doesn't match the data

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.

Fuel Economy Standards for Autos
Set by Congress and supplemented by the Transportation Department. In miles per gallon.

1978 '79 '80 '81 '82 '83 '84 '85

18 19 20 22 24 26 27 27½

This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

Figure 2:Fuel standards

Lie Factor

Tufte suggests optimizing the Lie Factor:

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.
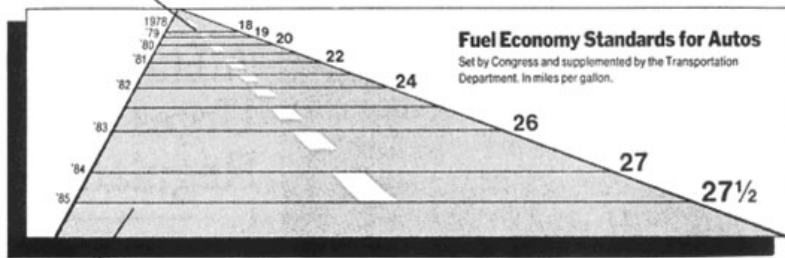
Fuel Economy Standards for Autos
Set by Congress and supplemented by the Transportation Department. In miles per gallon.

1978
'79
'80
'81
'82
'83
'84
'85

18 19 20
22
24
26
27
27½

This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

Figure 3:Fuel standards

Truth Continuum

"Study finds more than a quarter of journalist grads wish they'd chosen another career" (link)

Critiques of statistical graphics

http://www.statisticshowto.com/misleading-graphs/

https://arxiv.org/pdf/1503.00781.pdf

http://gator.gatewayk12.org/~smcgrail/myweb/powerpoint/
misleading_graphs/here_are_some_examples_of_mislea.htm

http:
//qz.com/580859/the-most-misleading-charts-of-2015-fixed/

http://themonkeycage.org/2013/01/
how-2012-stacks-up-the-worst-graph-on-record/

https://www.buzzfeed.com/katienotopoulos/
graphs-that-lied-to-us?utm_term=.grvny7OrGr#.hyWQEdzapa

Example
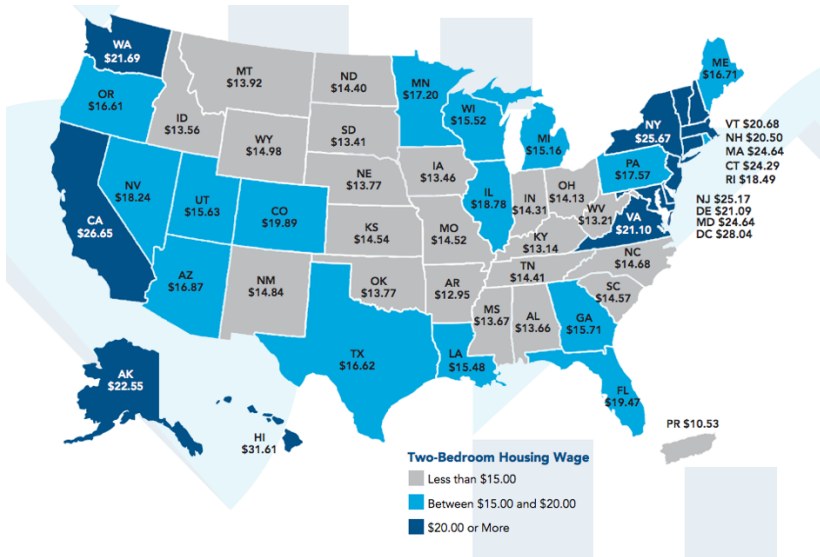
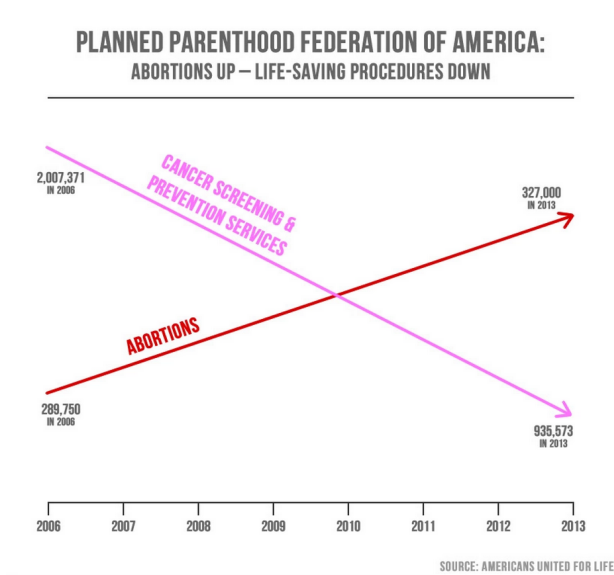Figure 4:Hourly wages
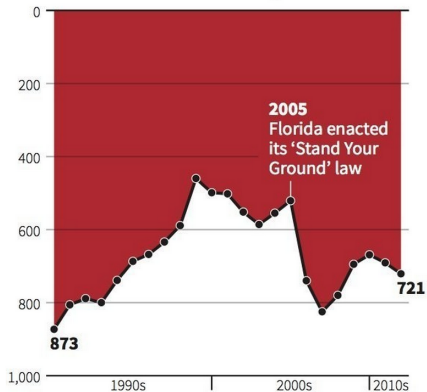
Example
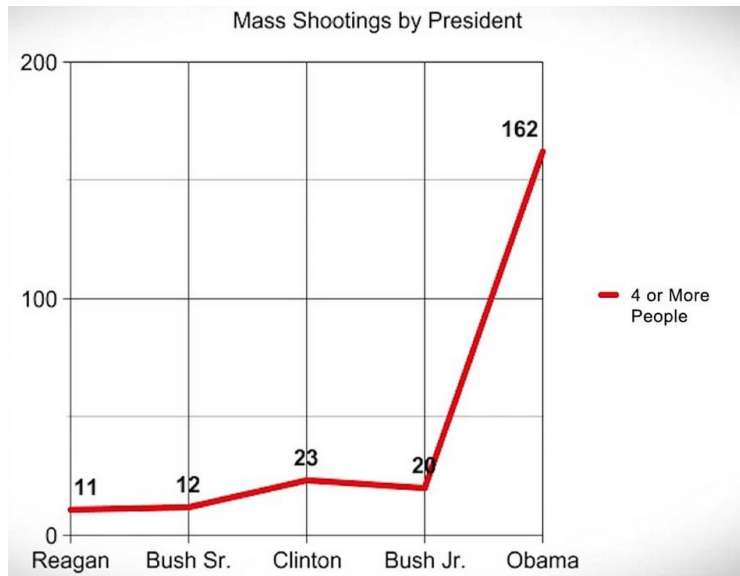
Figure 5:Hourly wages

Example

Figure 6:Hourly wages

Example

Figure 7:Hourly wages

Example

Figure 8:Hourly wages