

Lecture 2: Introduction to Data - Graphics Principles

Michael Lopez, Skidmore College

What is data? How do we describe data?

Data

- ▶ Two characteristics:
- ▶ 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

Grammar of graphics

Plots are composed of...

Data

Aesthetic mappings

Layers

Grammar of graphics

... as well as

Scales

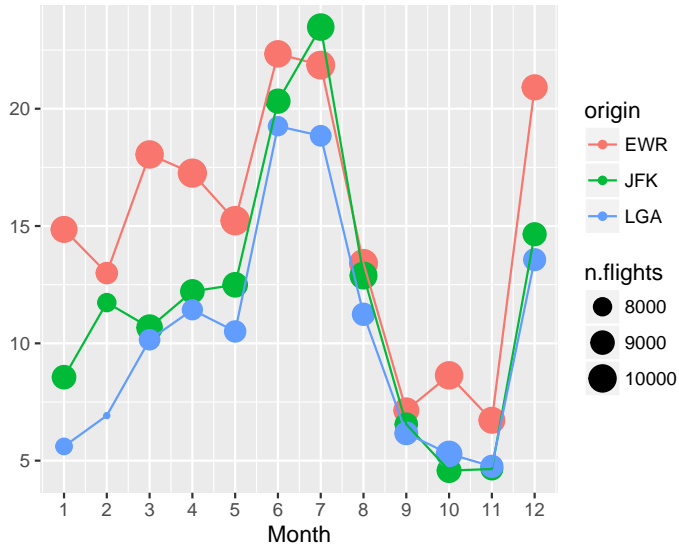
Coordinates

Faceting

Theme

Example

Average flight delays at New York airports



Example

Data

Aesthetic mappings

Layers

Scales

Coordinates

Faceting

Theme

Terminology

1. Data-ink ratio
2. Distortion
3. Lie factor

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-redundant core part of the graphic

Data-ink ratio:

Data-ink ratio

Ways to improve the data-ink ratio

remove non-data-ink

remove redundant data-ink

Data-ink ratio

Click for link to gif**

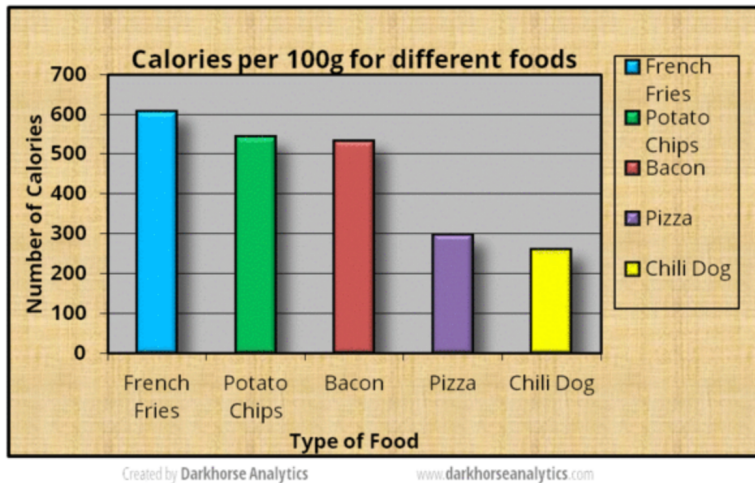


Figure 1: Data-ink

Distortion

Visual representation of data is inconsistent with numerical representation

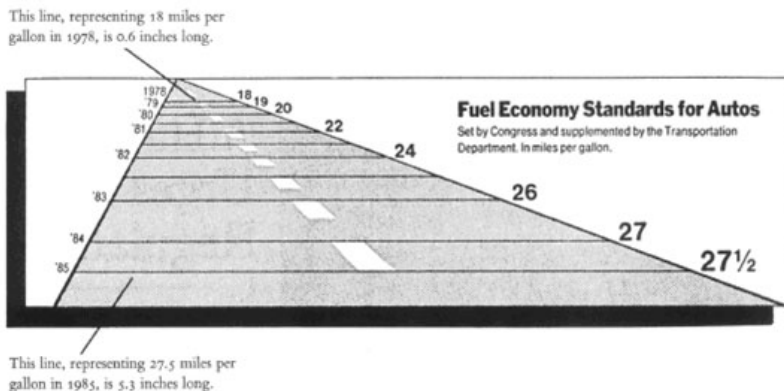
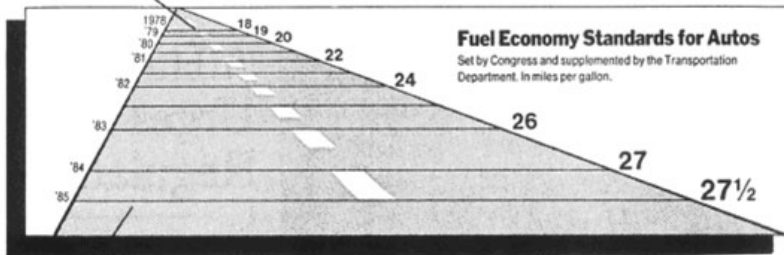


Figure 2: Fuel standards

Lie Factor

Calculating the lie factor

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



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

Example

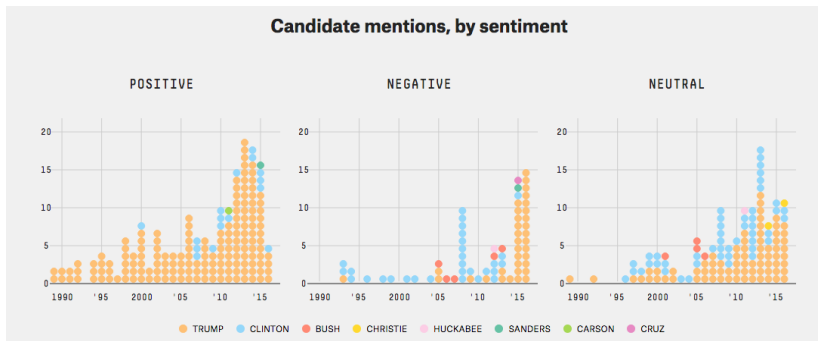


Figure 3: Trump

Example

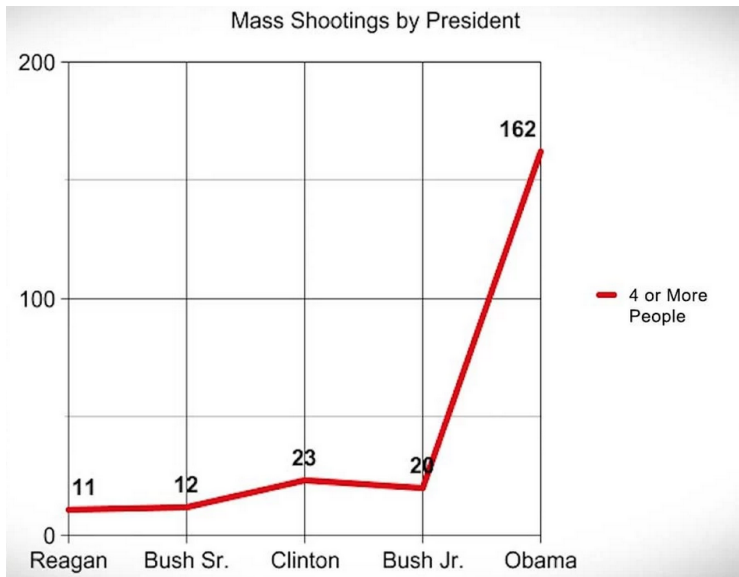


Figure 4: Mass shootings

Example

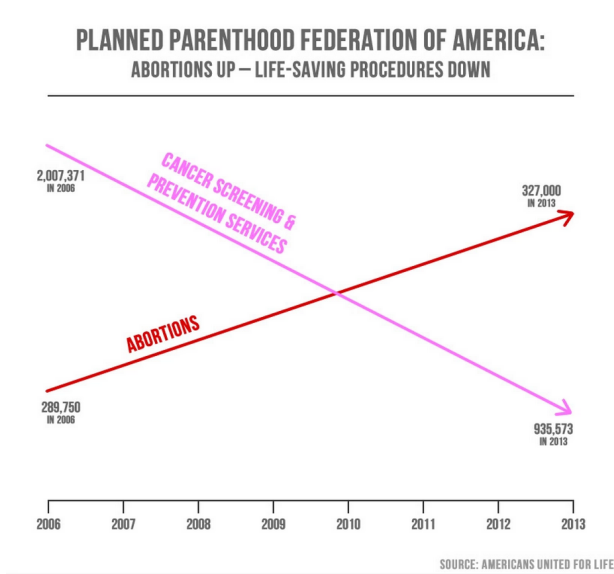


Figure 5: Life saving procedures

Example



Figure 6: Hourly wages

Example

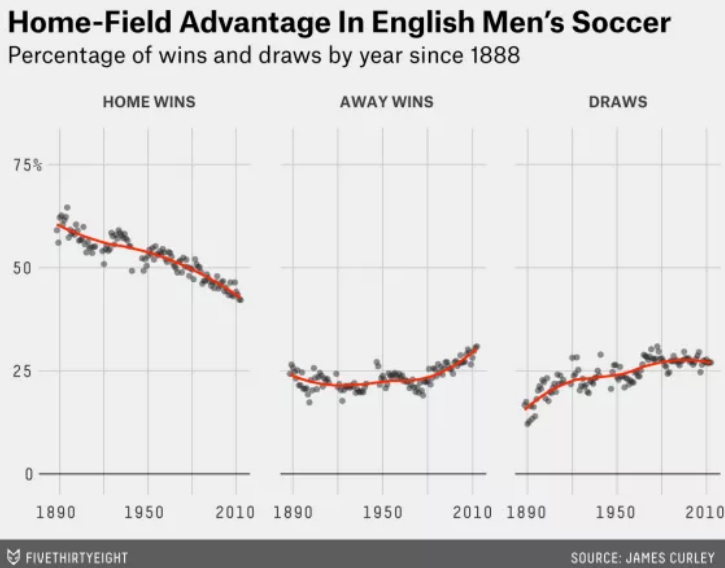


Figure 7: Soccer outcomes

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

[//qz.com/580859/the-most-misleading-charts-of-2015-fixed/](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/](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](https://www.buzzfeed.com/katienotopoulos/graphs-that-lied-to-us?utm_term=.grvny7OrGr#.hyWQEdzapa)