

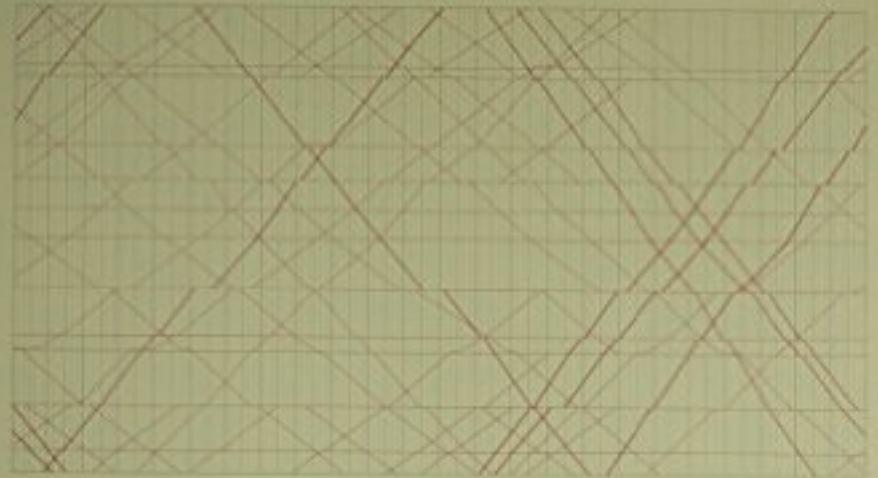
Improving Figure Clarity

Steven Braun

Data Analytics and Visualization Specialist

May 9, 2017

Simplify charts and graphs
to aid communication



SECOND EDITION

The Visual Display of Quantitative Information

EDWARD R. TUFTE

Edward R. Tufte

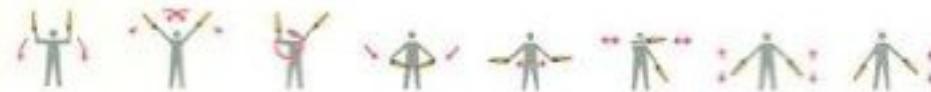
Envisioning Information



ESCAPING FLATLAND



MICRO / MACRO READINGS



LAYERING AND SEPARATION



SMALL MULTIPLES



COLOR AND INFORMATION



NARRATIVES OF SPACE AND TIME

GRAPHICAL INTEGRITY

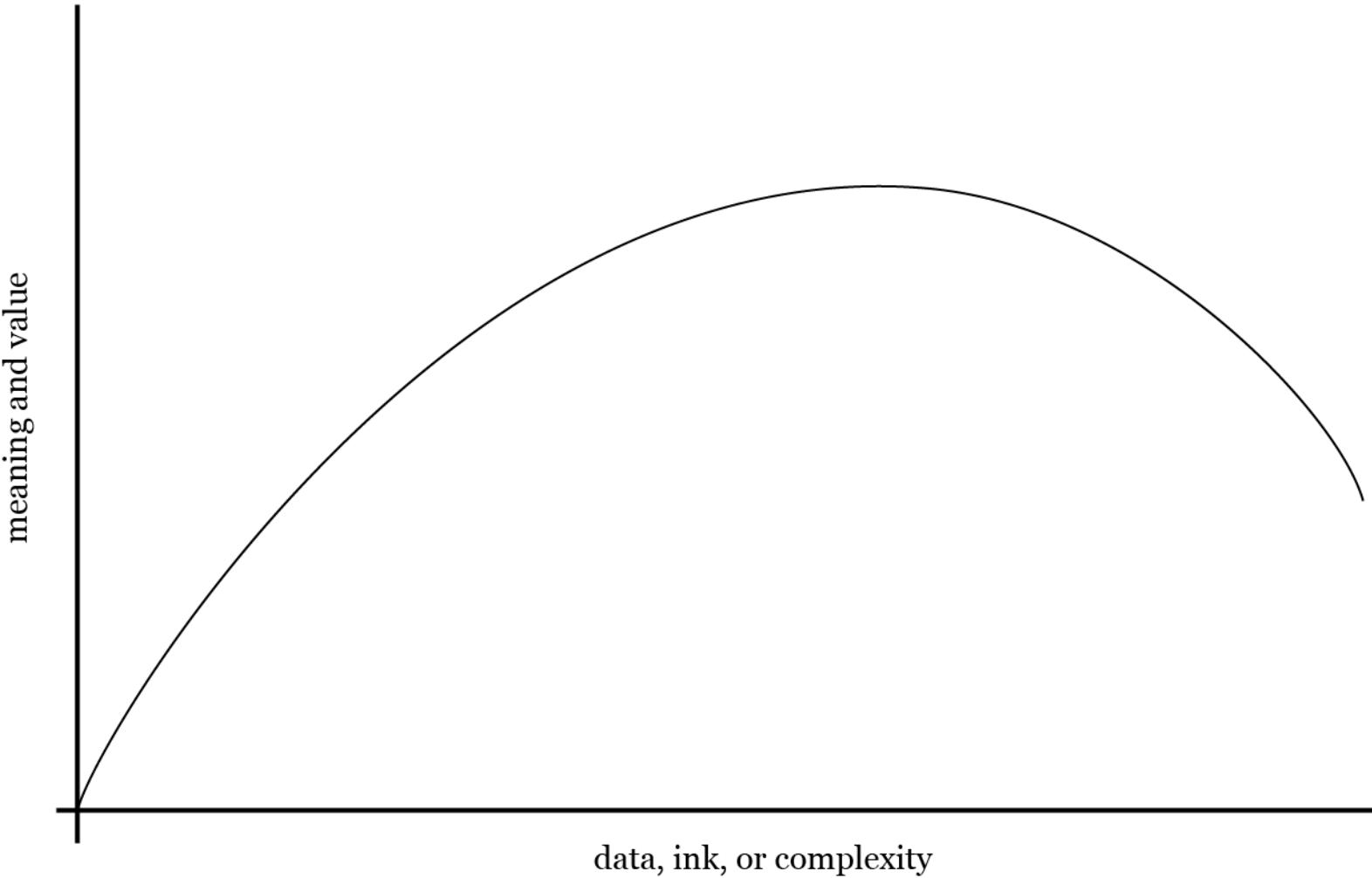
“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

Edward Tufte, “Visual Display of Quantitative Information”

TUFTE'S ROLES OF GRAPHICAL DISPLAYS

1. Show the data
2. Avoid distorting what the data have to say
3. Encourage comparisons
4. Reveal the data at several levels of detail
5. Serve a reasonably clear purpose
6. Be closely integrated with the statistical
and verbal descriptions

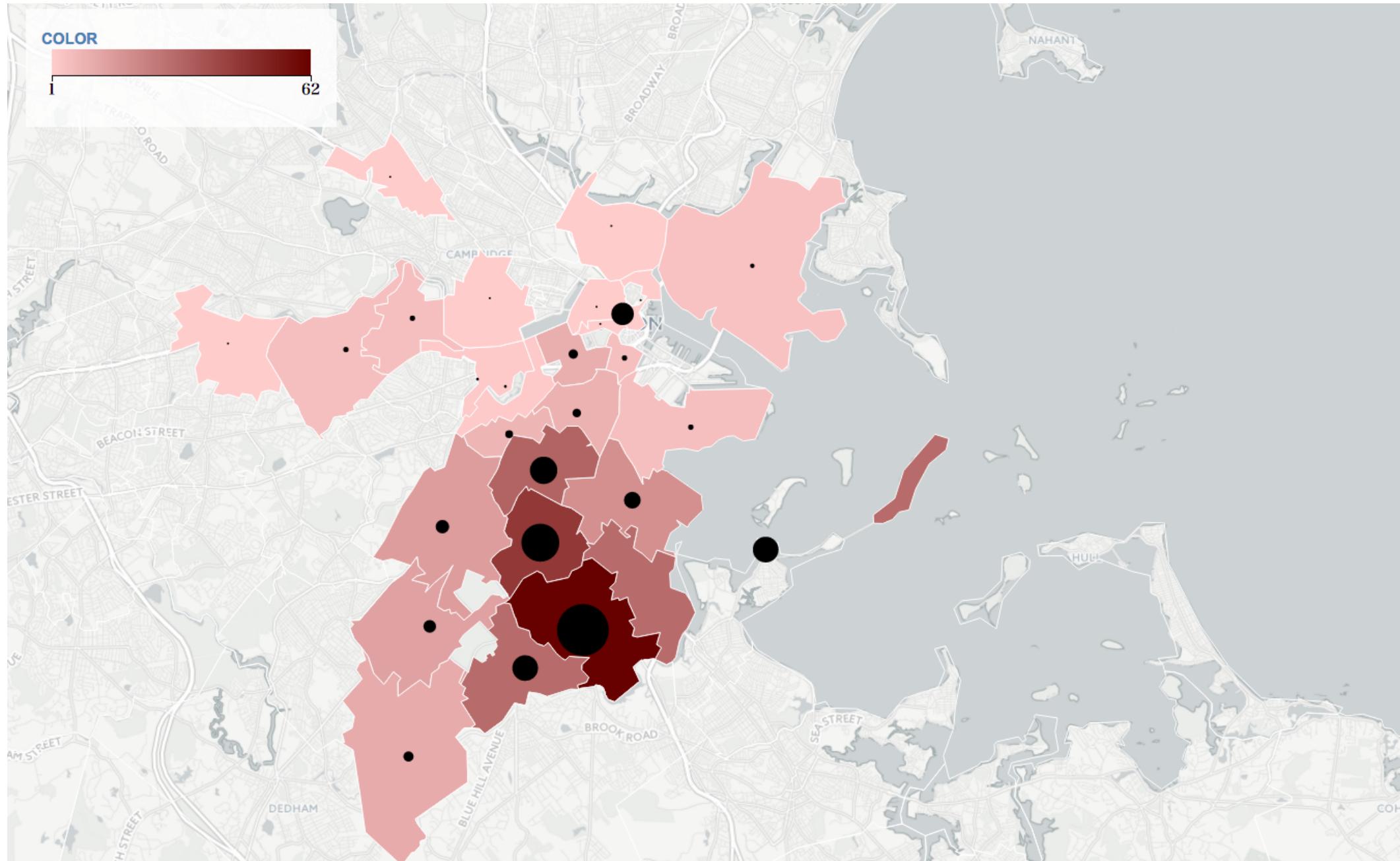
“data-ink ratio” = $\frac{\text{data ink}}{\text{total ink used to print the graphic}}$



56.8 Million Fantasy Players

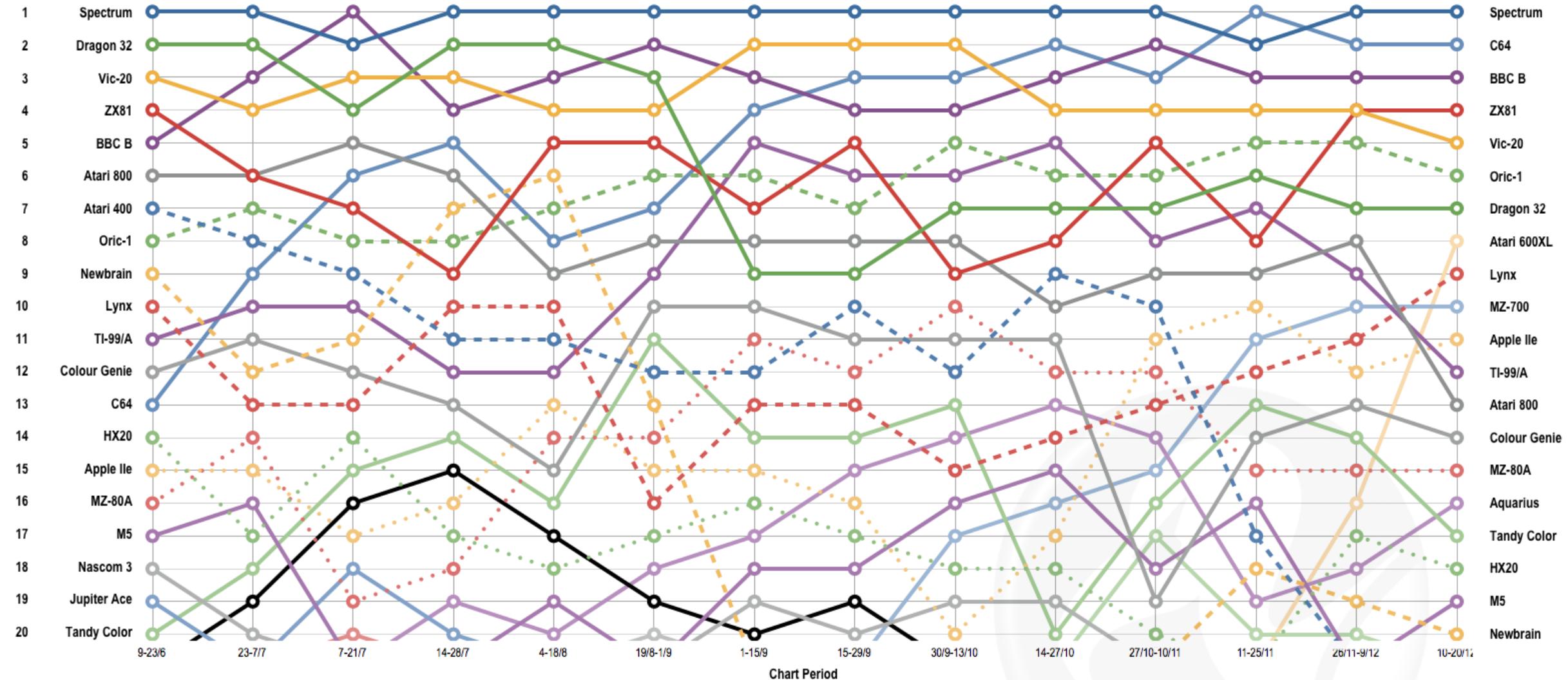
Aged 12+ in North America



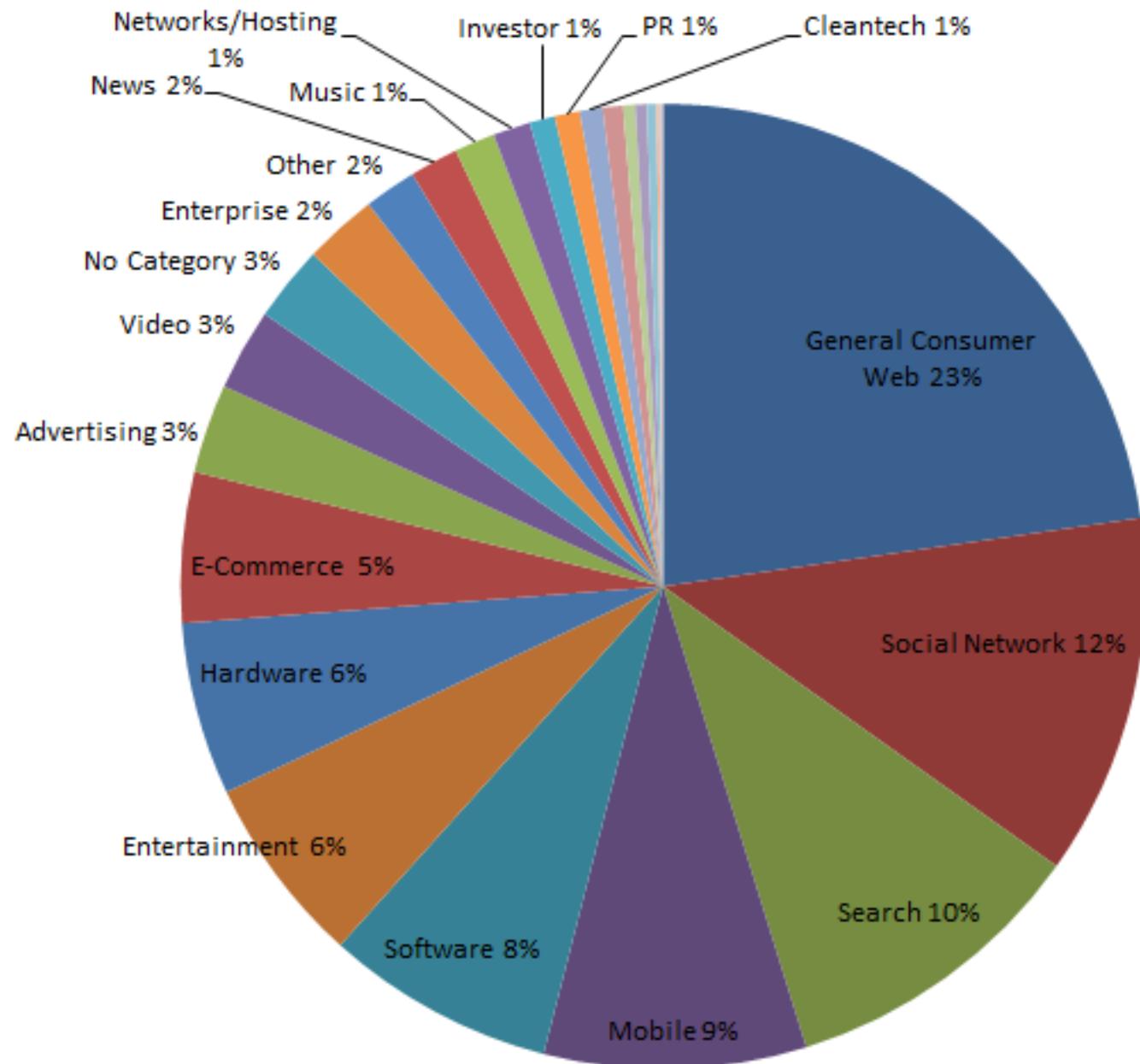


Is data-ink ratio a good measure of effectiveness?

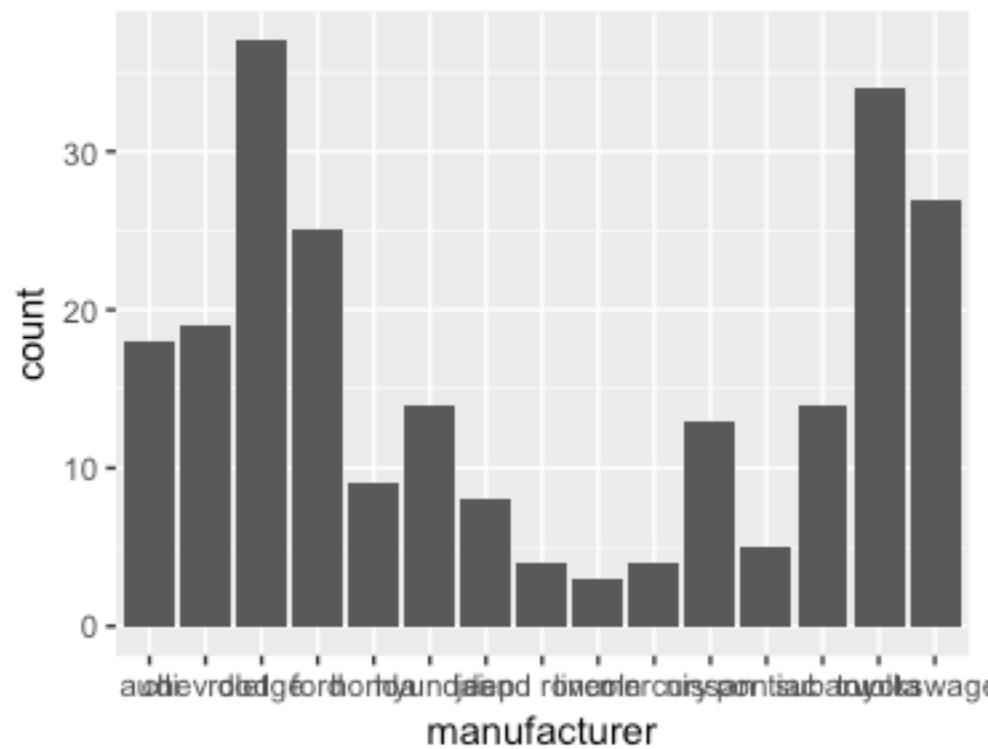
Sometimes.

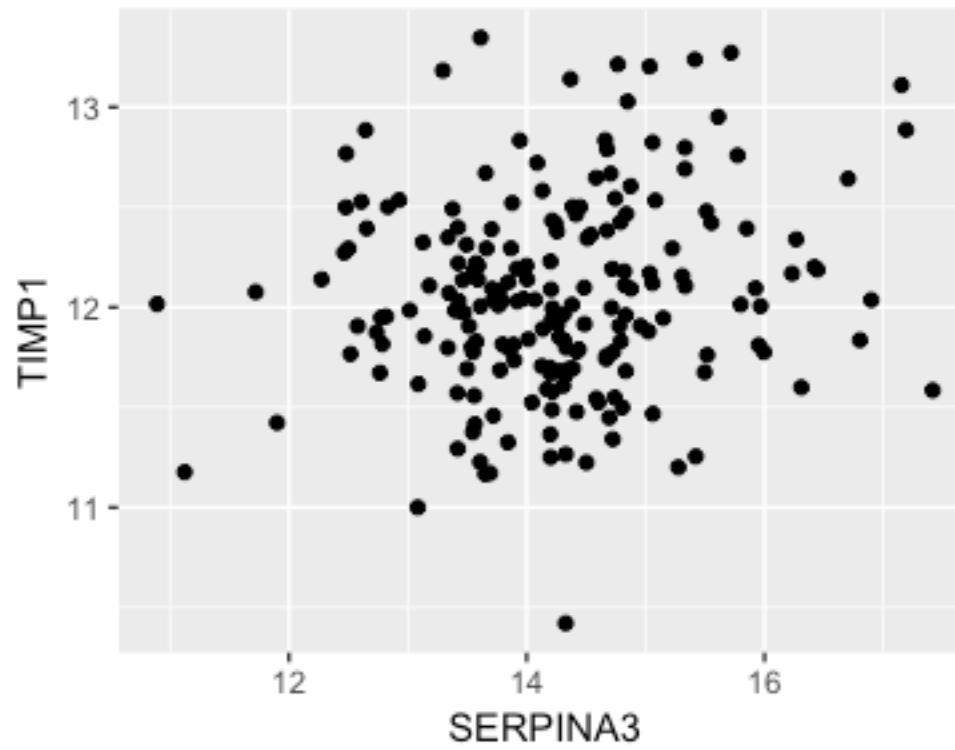


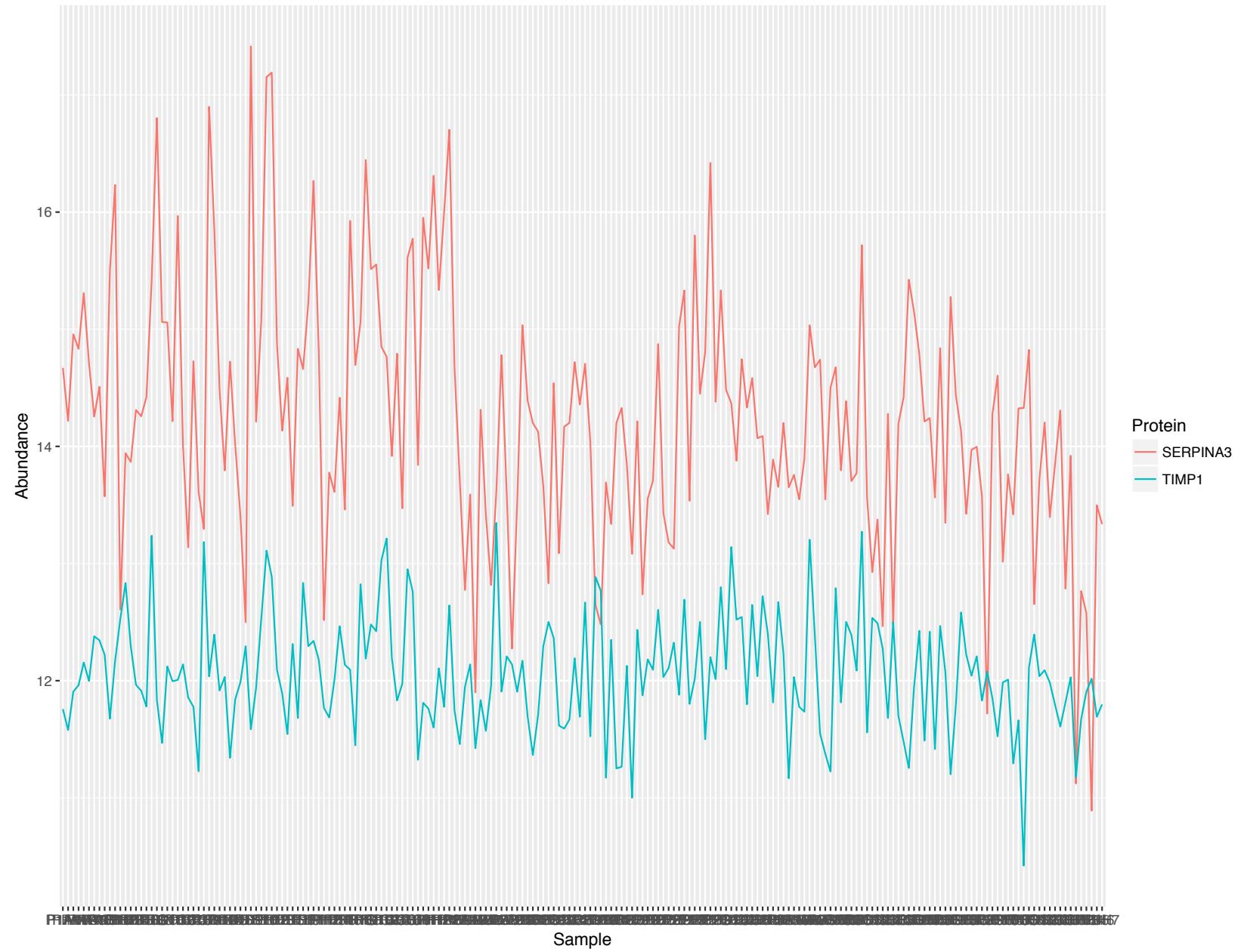
- Spectrum
- Dragon 32
- CBM Vic 20
- Sinclair ZX81
- Acorn BBC B
- Atari 800
- Atari 400
- Oric 1
- Newbrain A
- Lynx 48
- Texas TI-99/A
- Colour Genie
- CBM 64
- Epson HX20
- Apple IIe
- Sharp MZ-80A
- Sord M5
- Nascom 3
- Jupiter Ace
- Tandy Color
- Sharp PC-1500
- Acorn Atom
- Mattel Aquarius
- Micropresso
- Sharp MZ-700
- Acorn Electron
- Atari 600XL



Often, we need to work within and
against **software defaults** that may
impede chart simplification







SIMPLIFYING CHARTS AND GRAPHS

Reduce number of elements

Fewer elements means fewer bits to process

Use Gestalt principles

Alignment, spacing, grouping, and hierarchy

Use strong visual encodings

Position, length, proportion, hue, and more

Examples of charts and processes of
simplification/enhancing clarity

Use contrasting visual encodings to
enhance the salience
of particular data elements

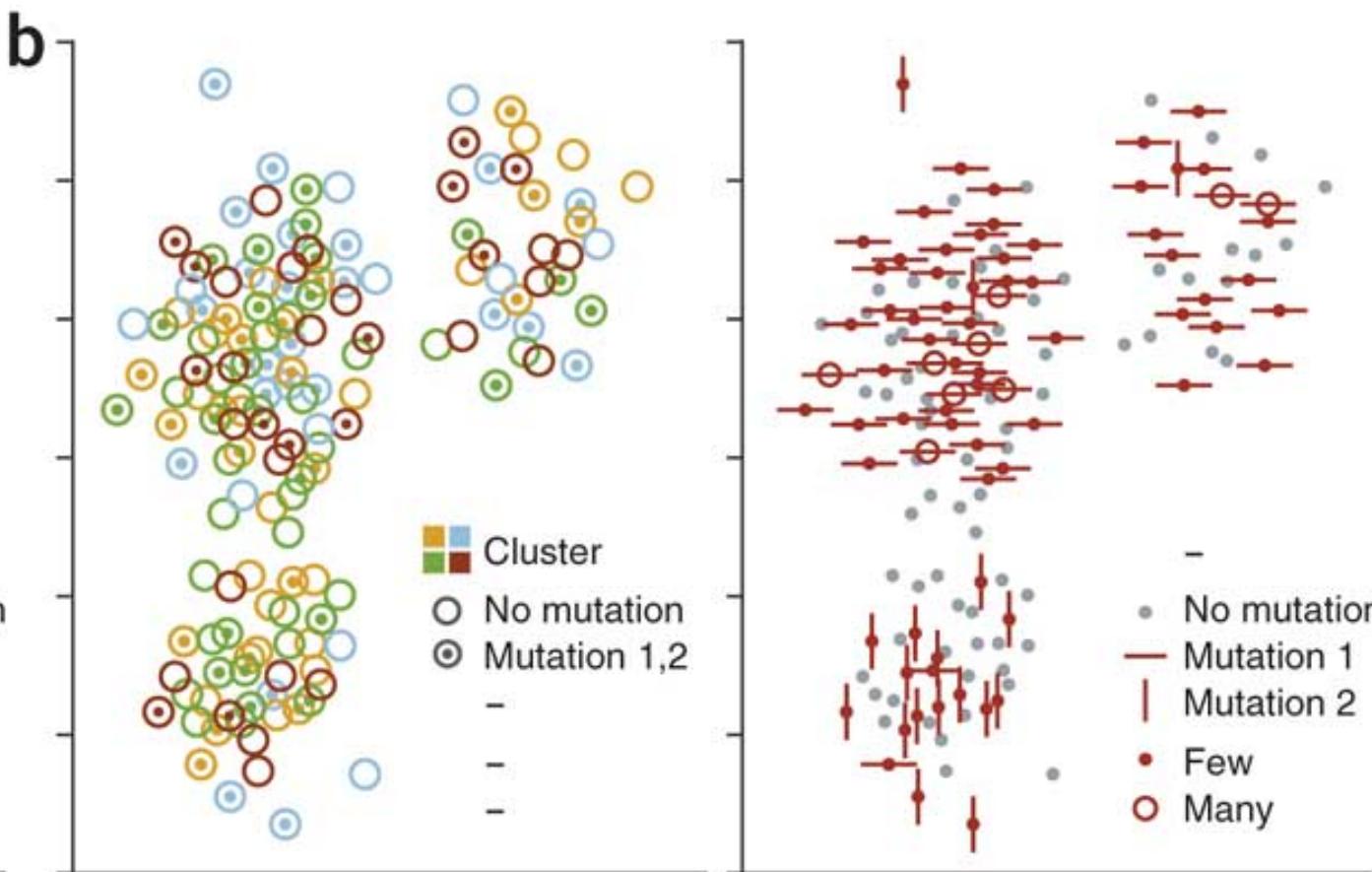
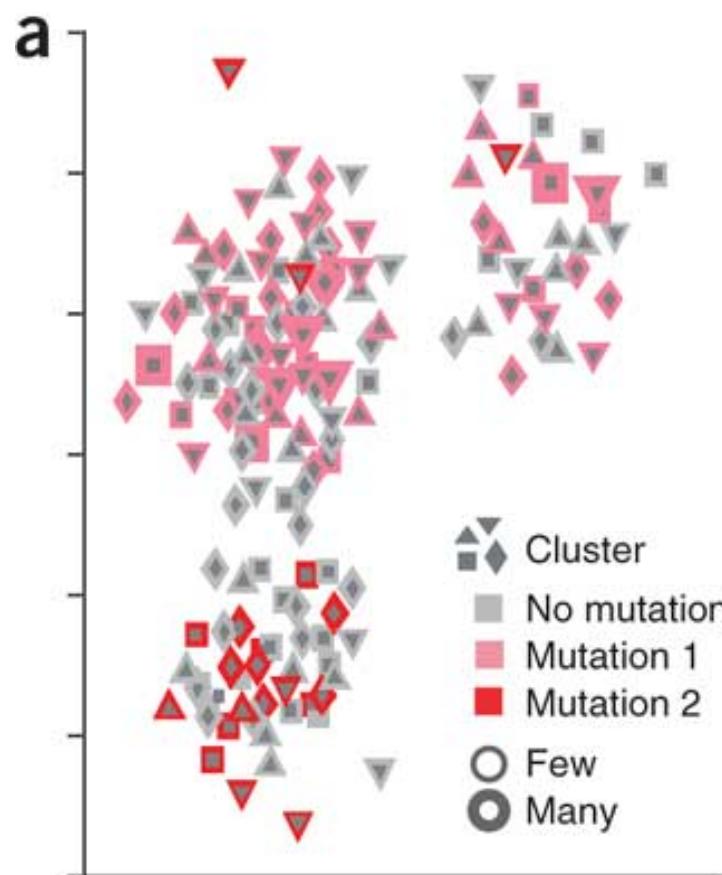
VISUAL ENCODINGS

Color		Sequence		Length	
Value/Gradation		Size + Scale		Area	
Texture		Orientation		Proportion	
Symbol		Proximity/Density		Count	

Instances of **A** are easy to find,
but instances of **P** are not

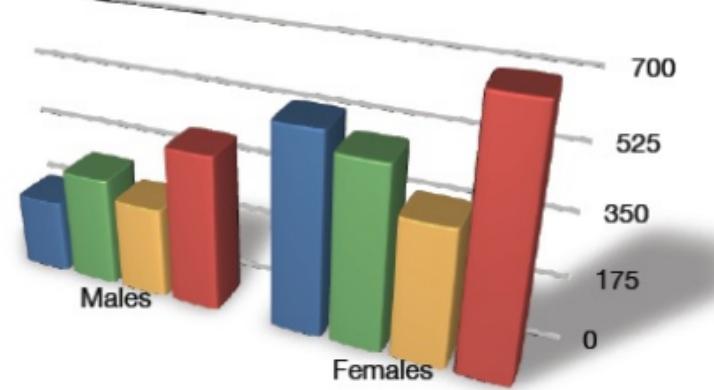
MSVTLHTVFCERTPKTC
EMESRCVPQEGVQWRDL
GSALQPGFGGFQVFCL
SLPRTGRGGNSIWWGKK
FEDEYSEYSEYLKH**AVR**
GVVSMSNNGPNTNGSQF
FITYGKQPHLDMKYTVF
GKVIDGLEKAPVNEKTY
RPLNDVHIKDITIHNPF

Redundant encodings can aid perception, but visual conjunctive searches come at a cost

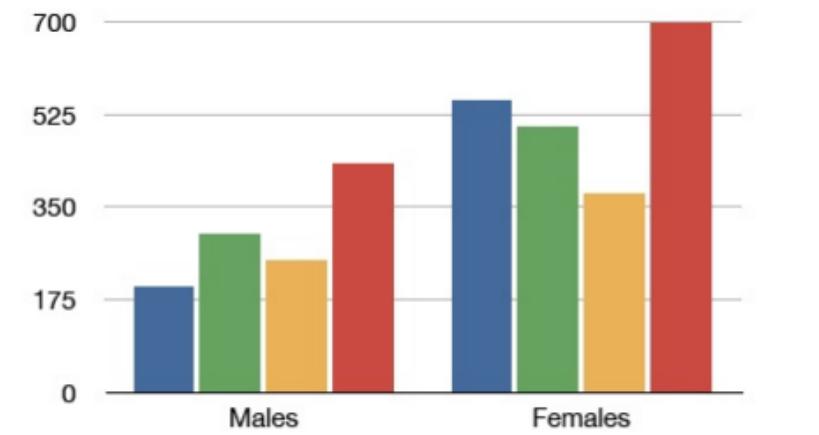


Simple Tips for Improving Basic Charts

Avoid any unjustified use of
3D representation



■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

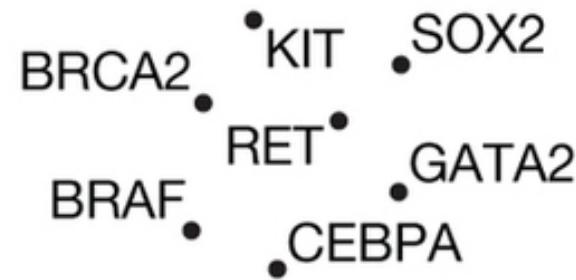


■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

Use data labels judiciously
and position them without obscuring
important parts of the visualization

a

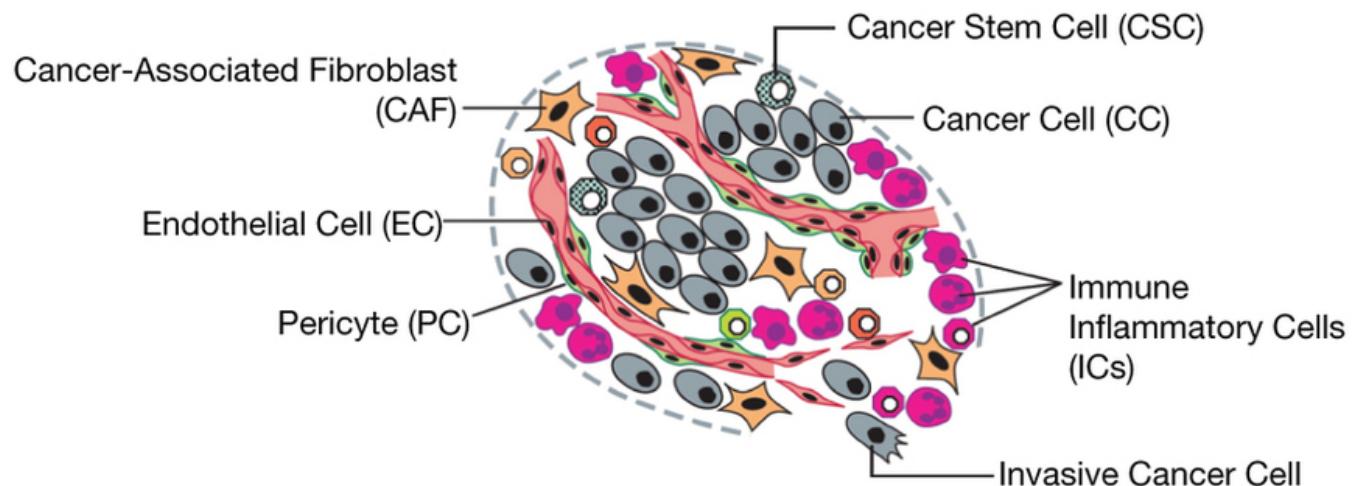
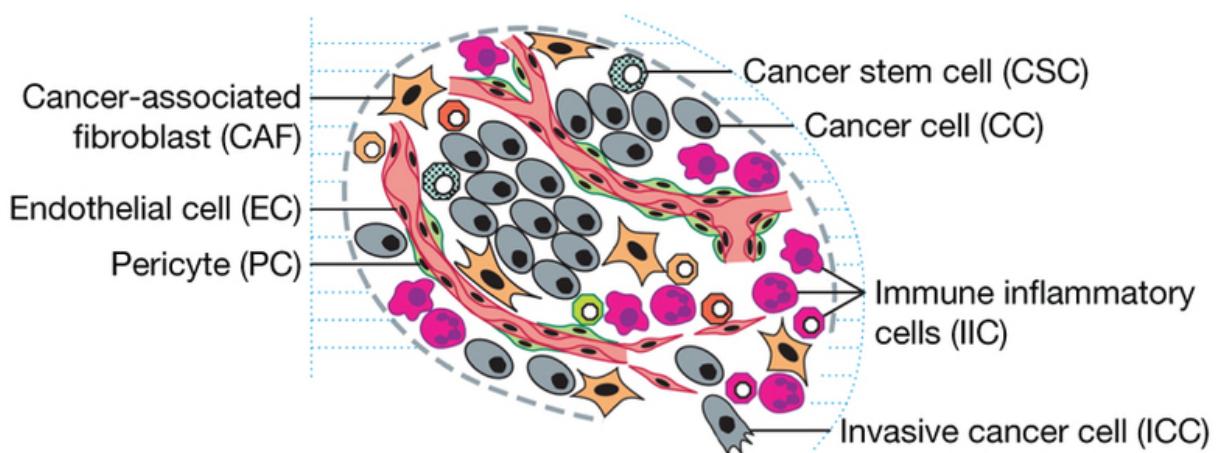
Poor label placement

b

Good label placement

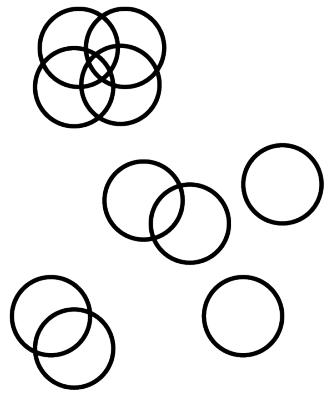
c

Label placement priority

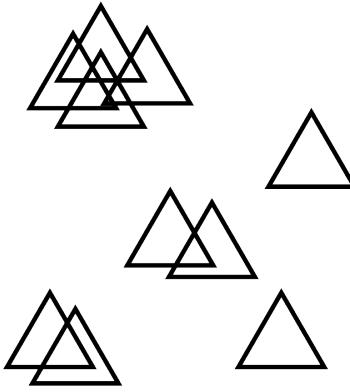
a**b**

For dense plots, use open markers
instead of closed

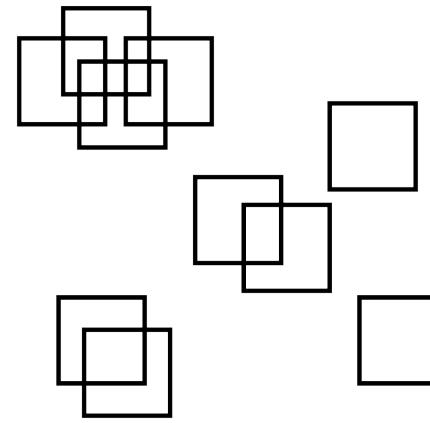




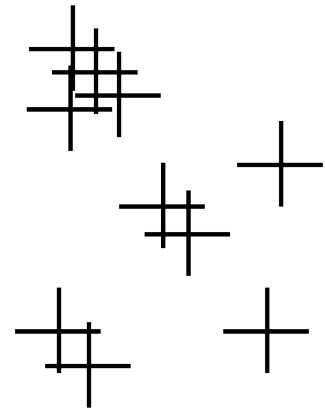
Circle



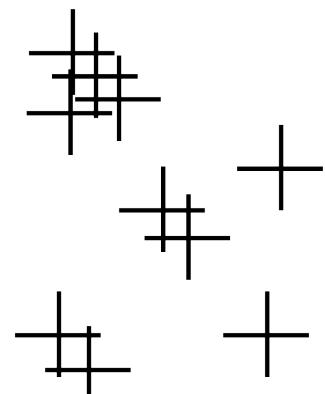
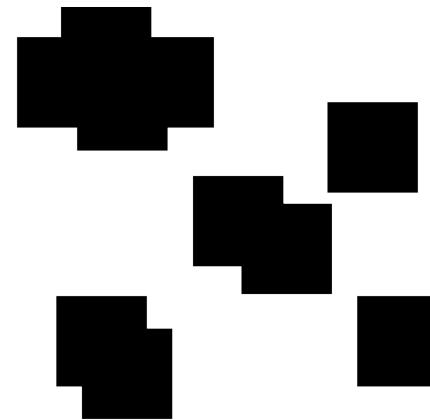
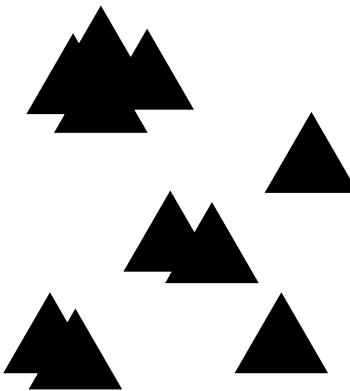
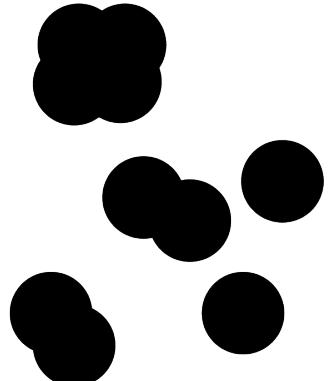
Triangle



Square

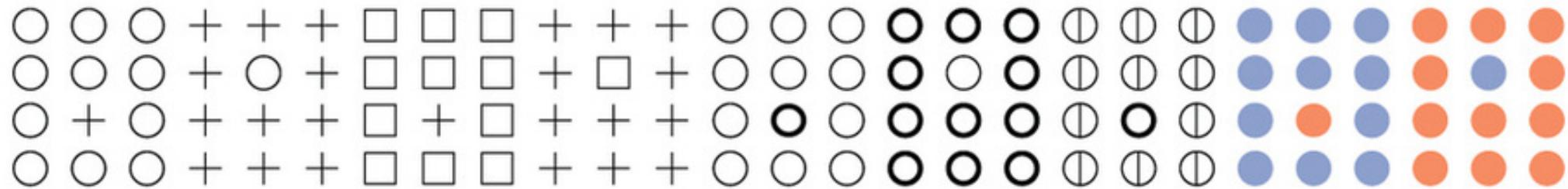


Crosshair

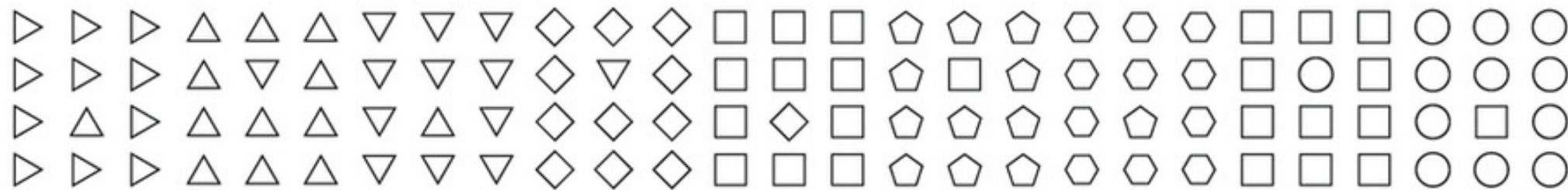


a

Strong visual boundaries



Weak visual boundaries

**b**

A N T A

c

○ □ △ ○

d

A T G T

○ △ ★ △

e

A C N N

○ ▽ □ □

f

A A A A

○ ○ ○ ○

e

○ ○ ○ ○

○ ○ ○ ○

○ ● ○ ○

○ ○ ○ ○

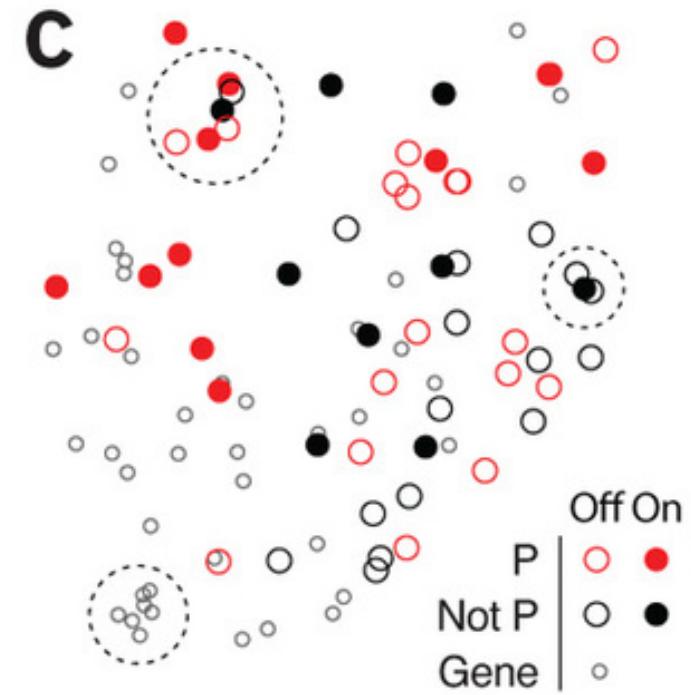
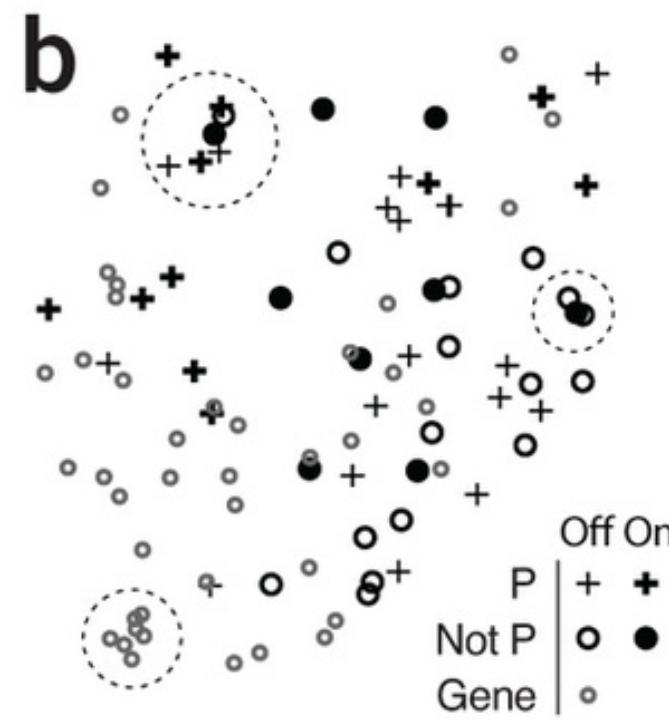
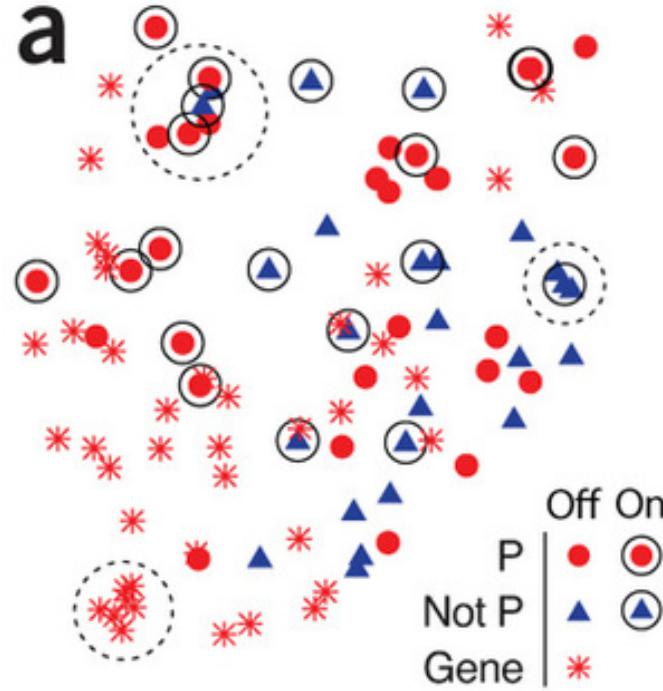
f

○ w < ○

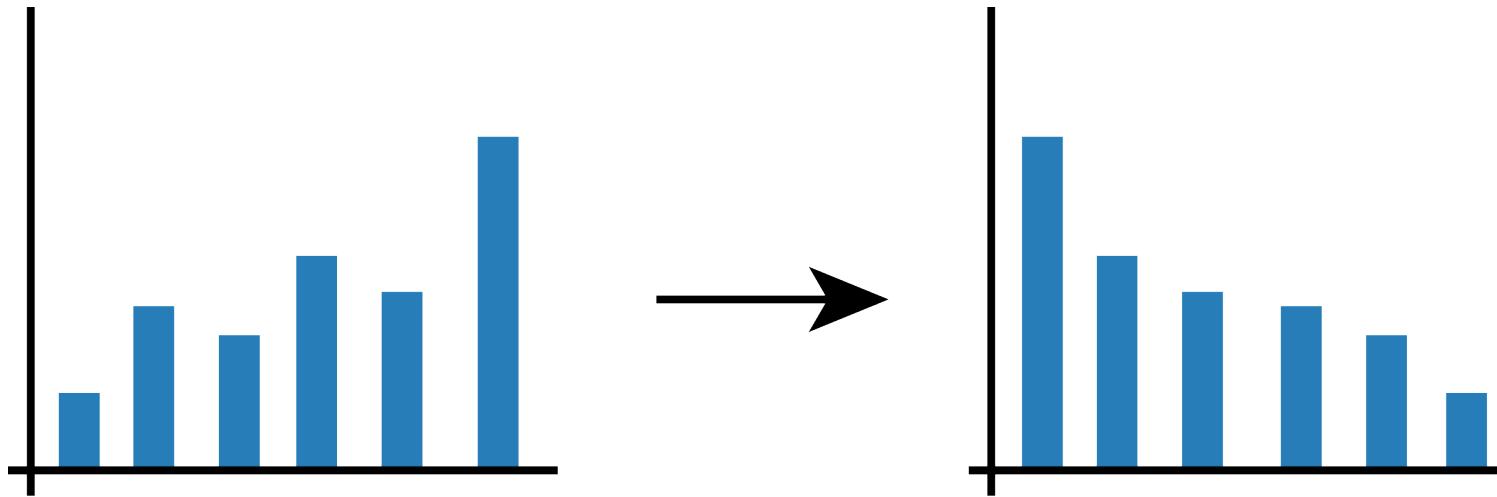
○ < S <

○ + w w

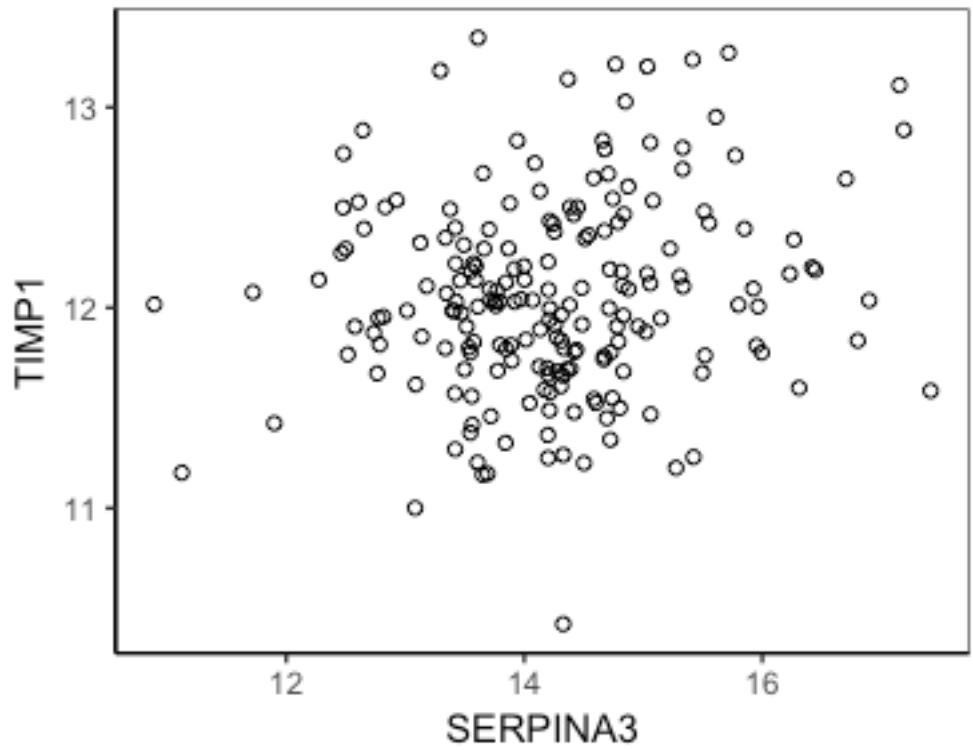
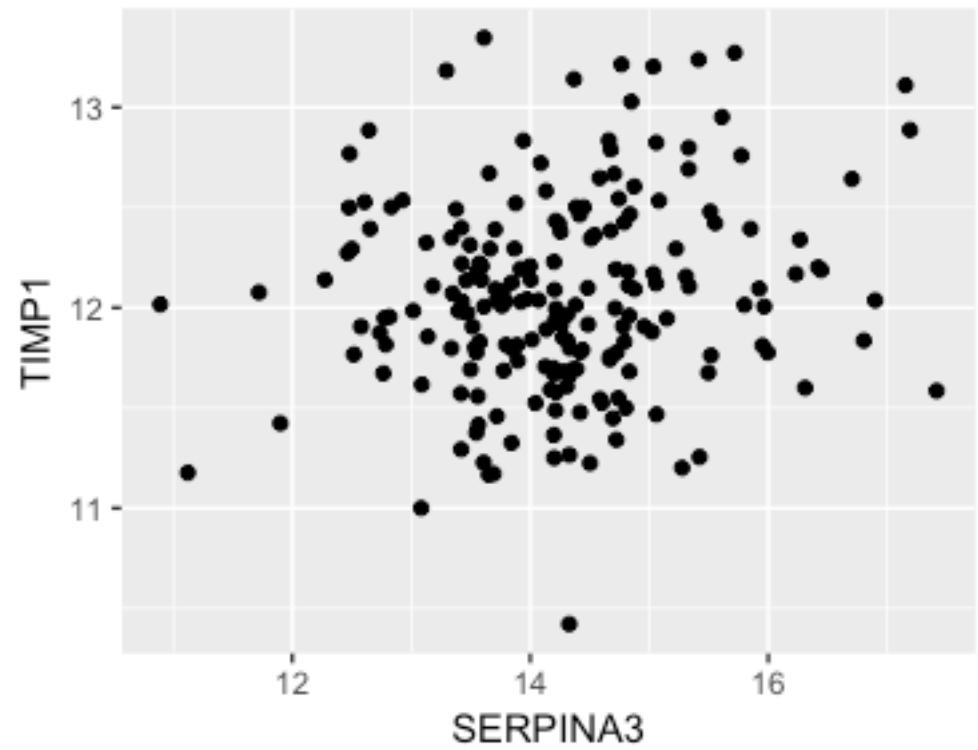
○ ○ ○ ○



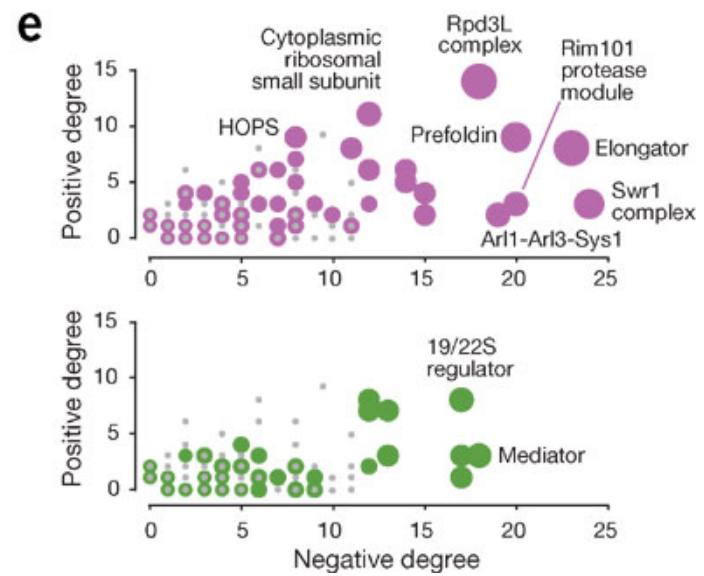
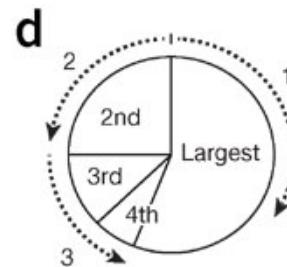
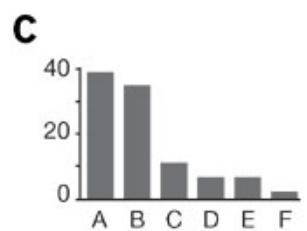
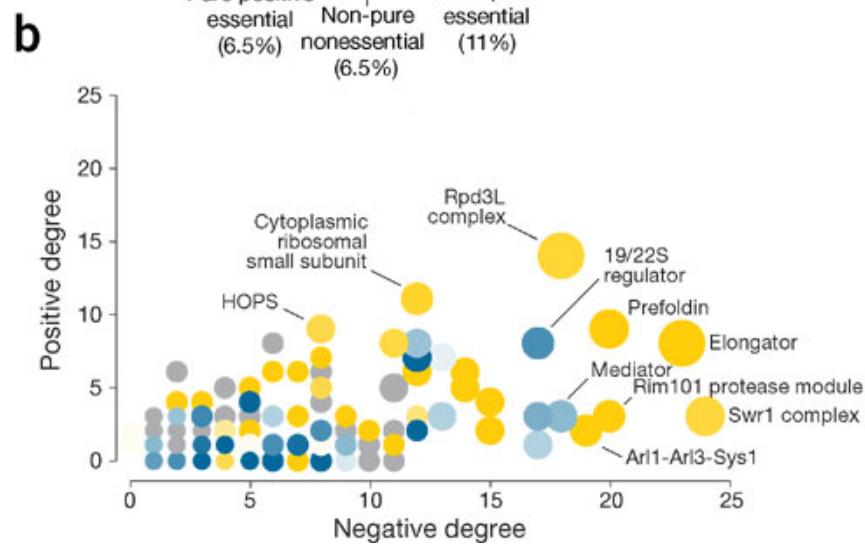
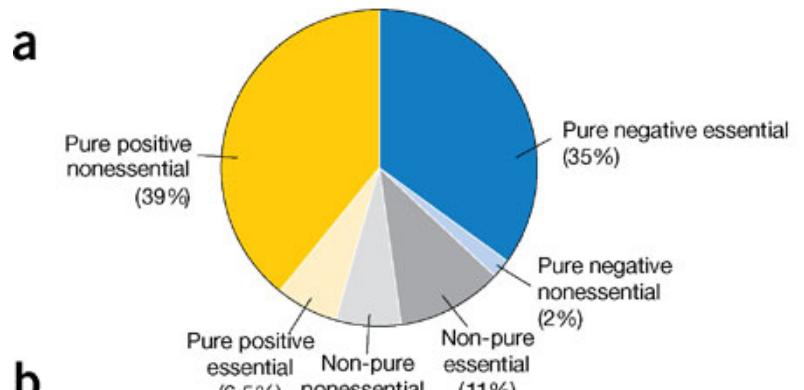
Order bar charts by
descending magnitude



Eliminate or reduce unnecessary
default elements, such as gridlines
and excessive tick marks



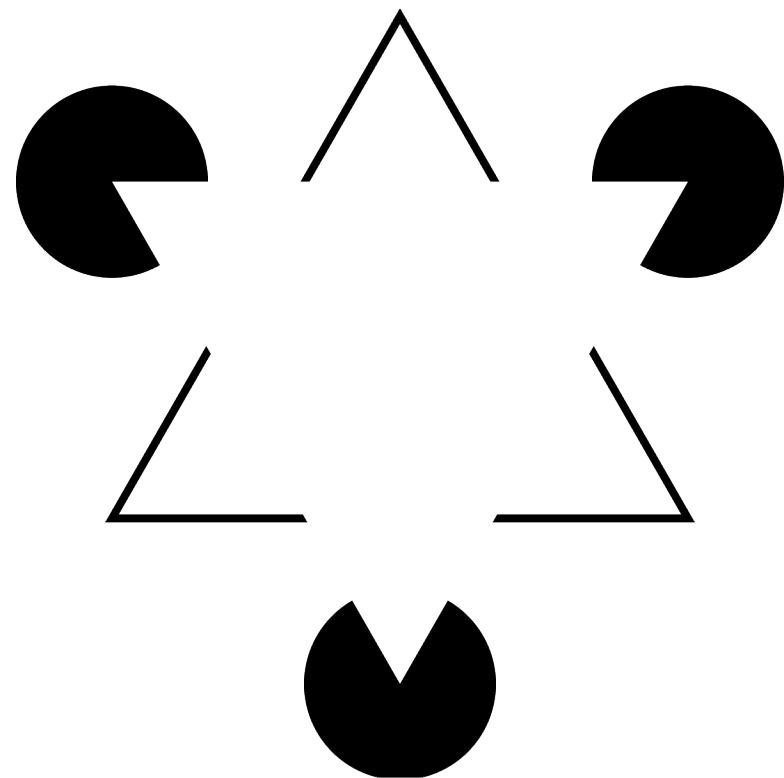
Arrange the two largest wedges of a pie chart at the 12:00 position and order the remaining wedges counterclockwise in descending size



Avoid Gestalt illusions that can
engender incorrect readings of a
visualization

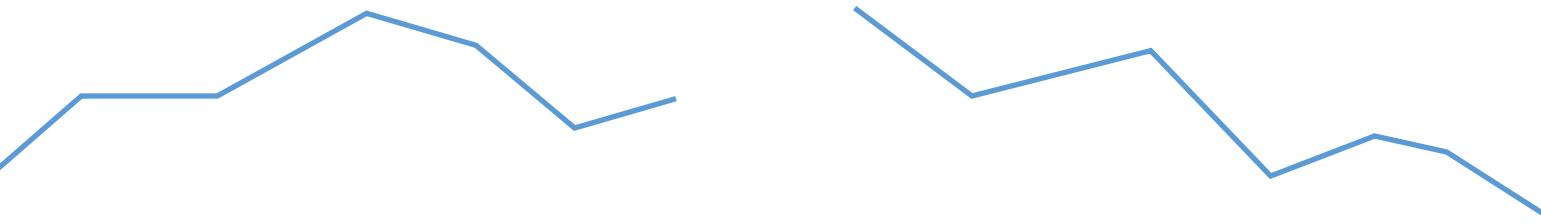
GESTALT ILLUSIONS

Closure



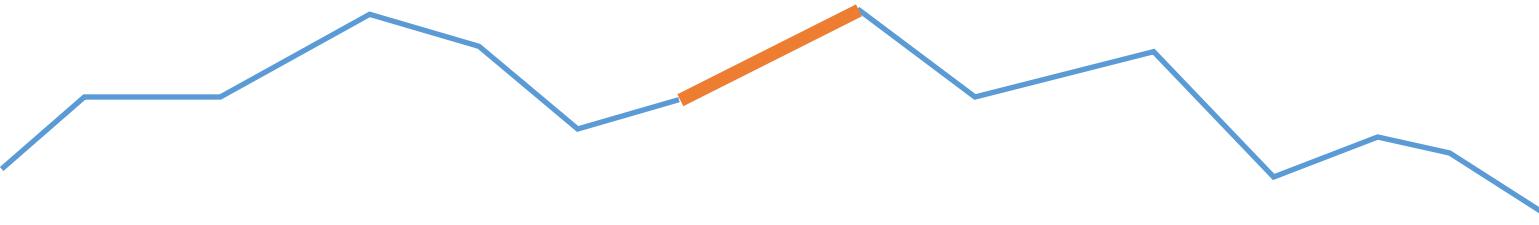
GESTALT ILLUSIONS

Closure



GESTALT ILLUSIONS

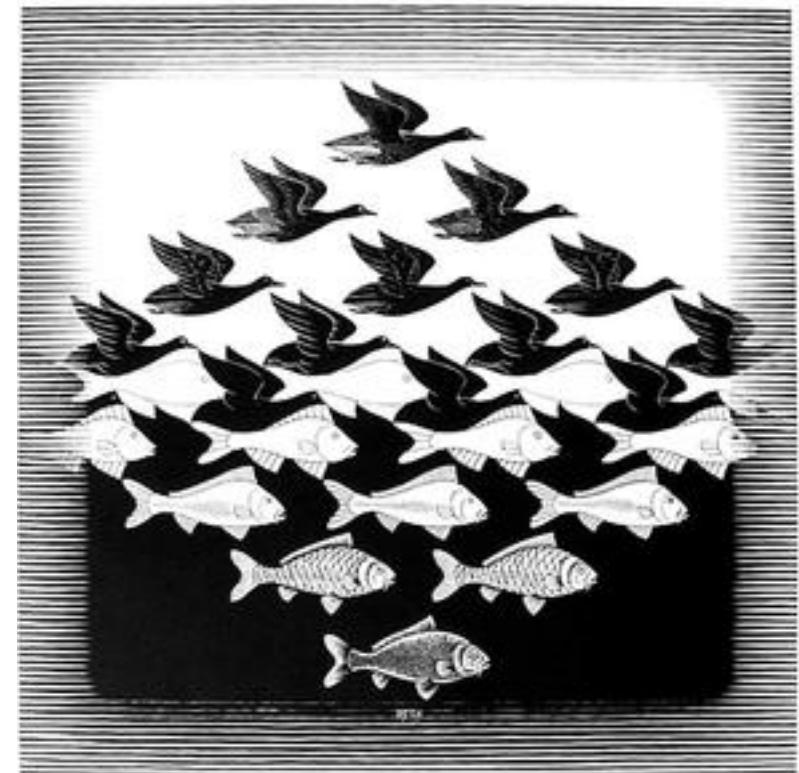
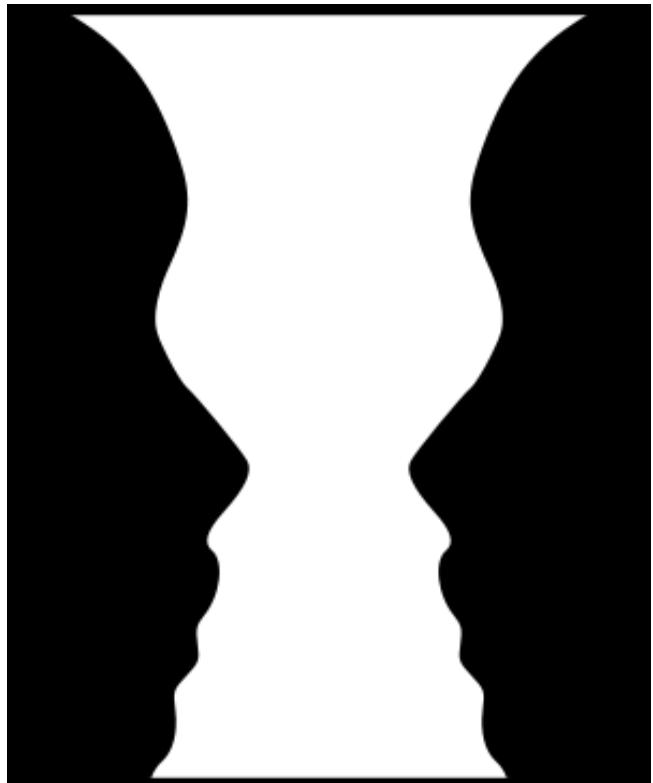
Closure



It is generally better to explicitly encode null values than leave them implied in your representation

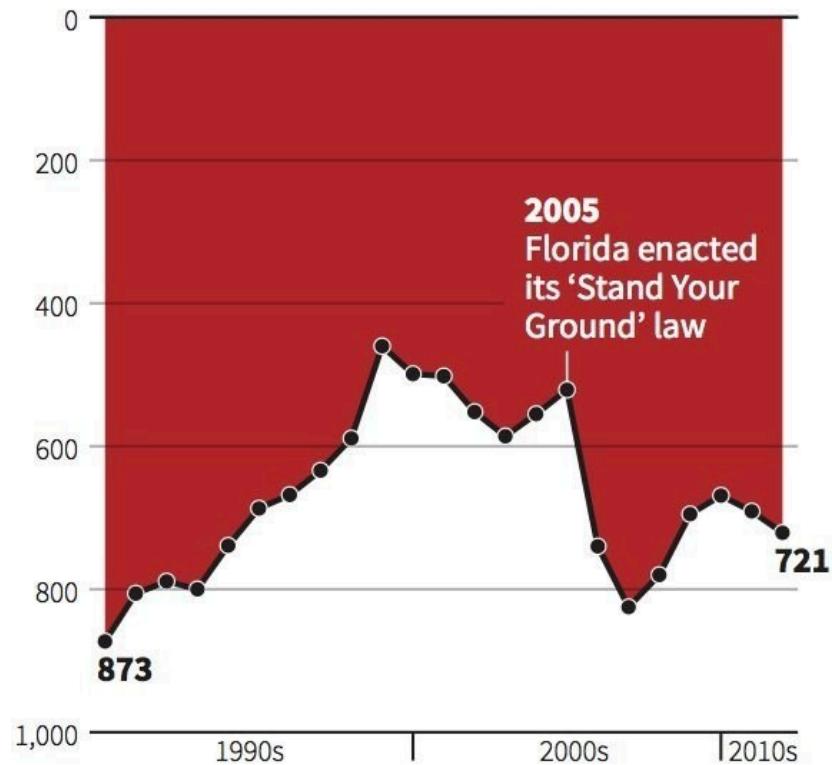
GESTALT ILLUSIONS

Figure/ground



Gun deaths in Florida

Number of murders committed using firearms



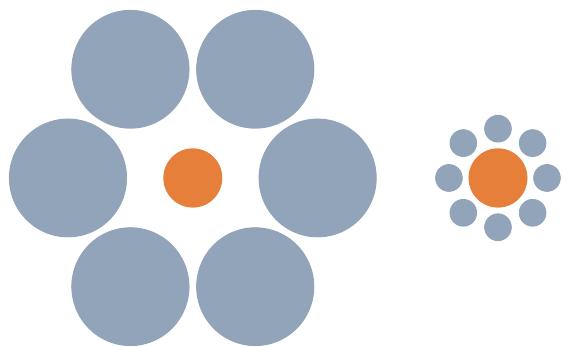
Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

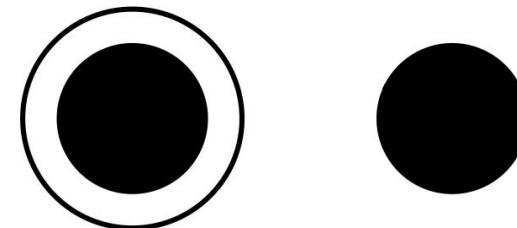
REUTERS

ILLUSIONS OF SCALE

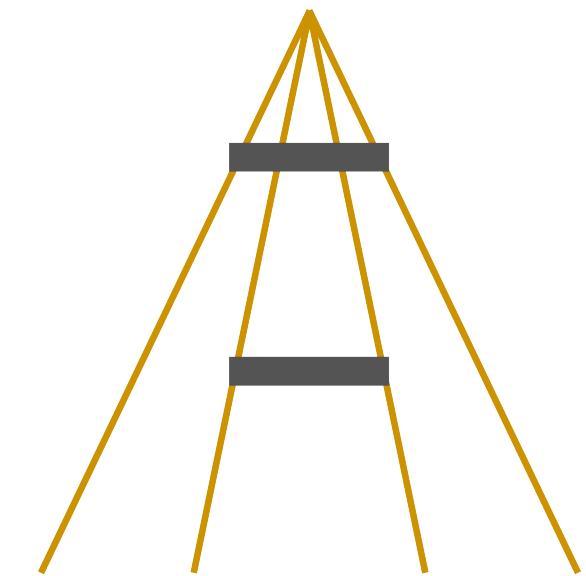
Ebbinghaus



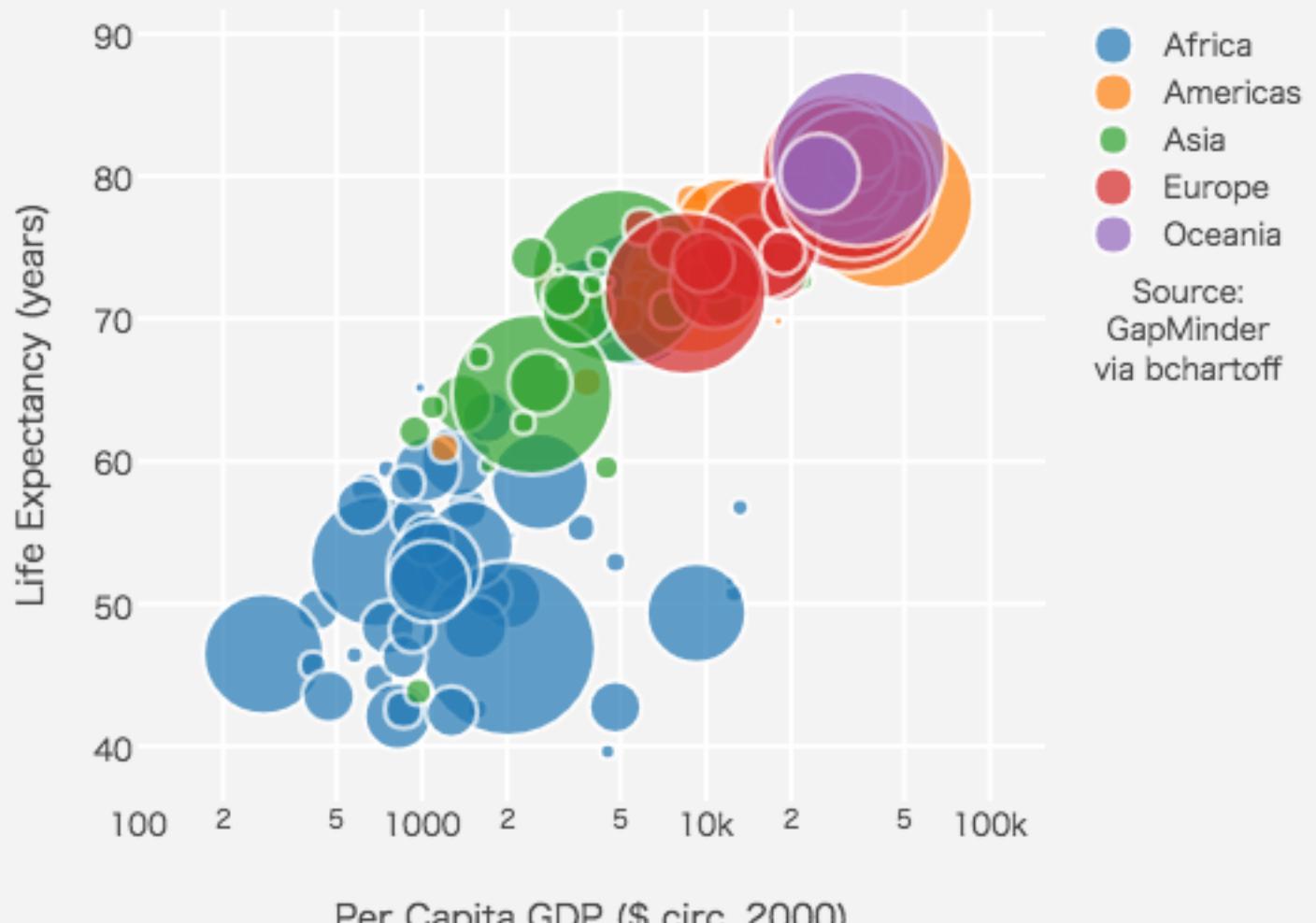
Delboeuf



Ponzo



Life Expectancy v. Per Capita GDP, 2007



Encode null values appropriately to aid understanding

Null Values

Nulls and zeroes are important placeholders that tell us something significant about our data

TYPES OF NULL VALUES

Missing value

No answer, “N/A” (survey)

NaN

Type error

No signal

Nulls resulting from data that do not fit schema

Improving Figure Clarity

<http://www.random.org/integers/>

1. In small groups, determine an interpretation for the data set provided
2. Work together to produce a visualization of the data that clearly communicates that interpretation
3. Pass your visualization on to another table
4. In each group, use the visualization received from another table to produce another visualization that communicates an *alternative* interpretation of the data, as provided
5. Repeat this process 3 times

Improving Figure Clarity

<http://www.random.org/integers/>

24	54	71	75	34
9	76	74	34	43
71	17	4	81	60
9	89	67	80	37

1. In small groups, determine an interpretation for the data set provided
2. Work together to produce a visualization of the data that clearly communicates that interpretation
3. Pass your visualization on to another table
4. In each group, use the visualization received from another table to produce another visualization that communicates an *alternative* interpretation of the data, as provided
5. Repeat this process 3 times

Discussion

1. How accurate were initial interpretations?

2. Did you find interpretations of your data that were surprising?

3. Are there ways in which you could have communicated your data more effectively?