Now you see It!

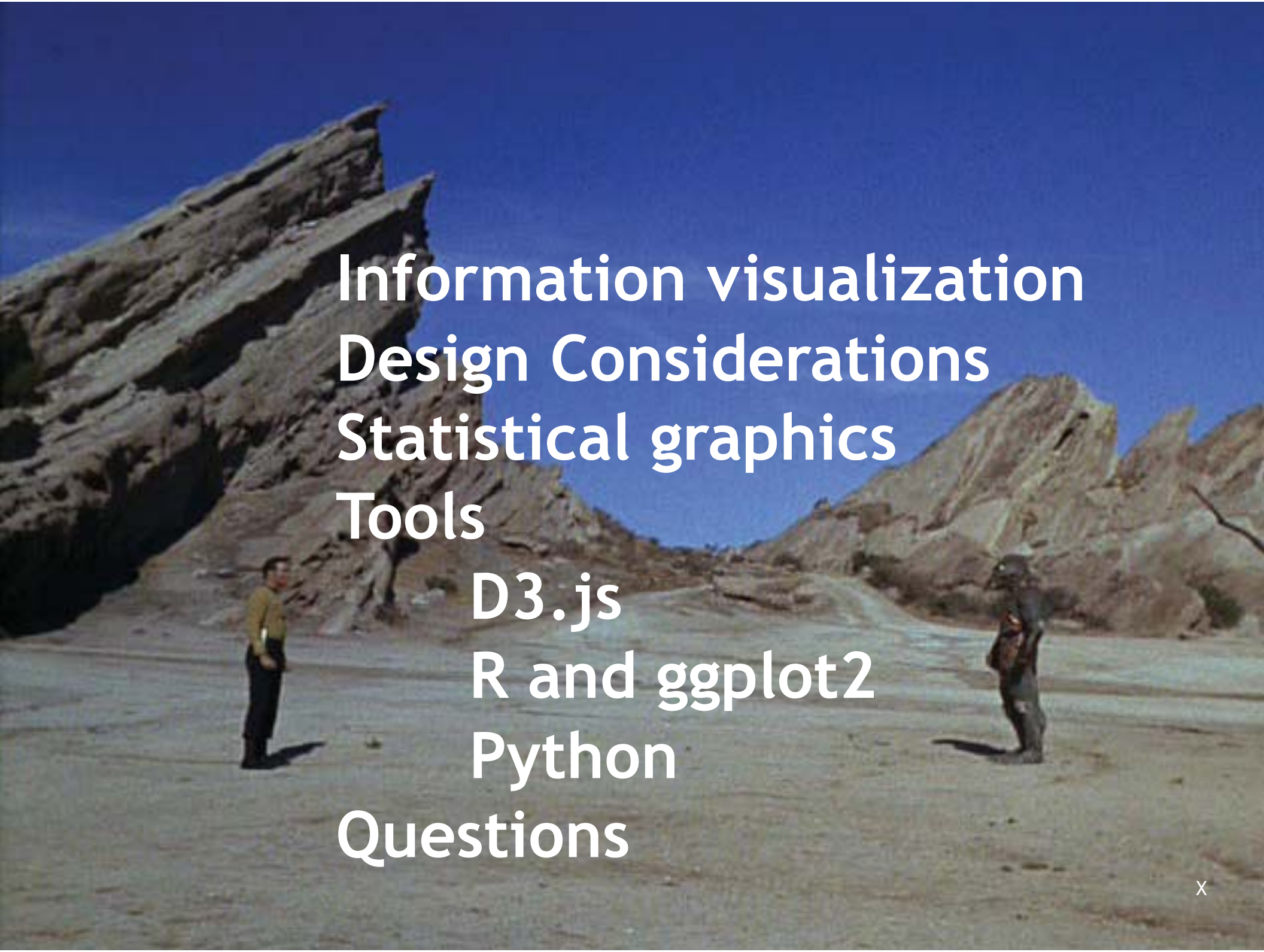
Intro to Data Visualization/Design Elements of Visualization

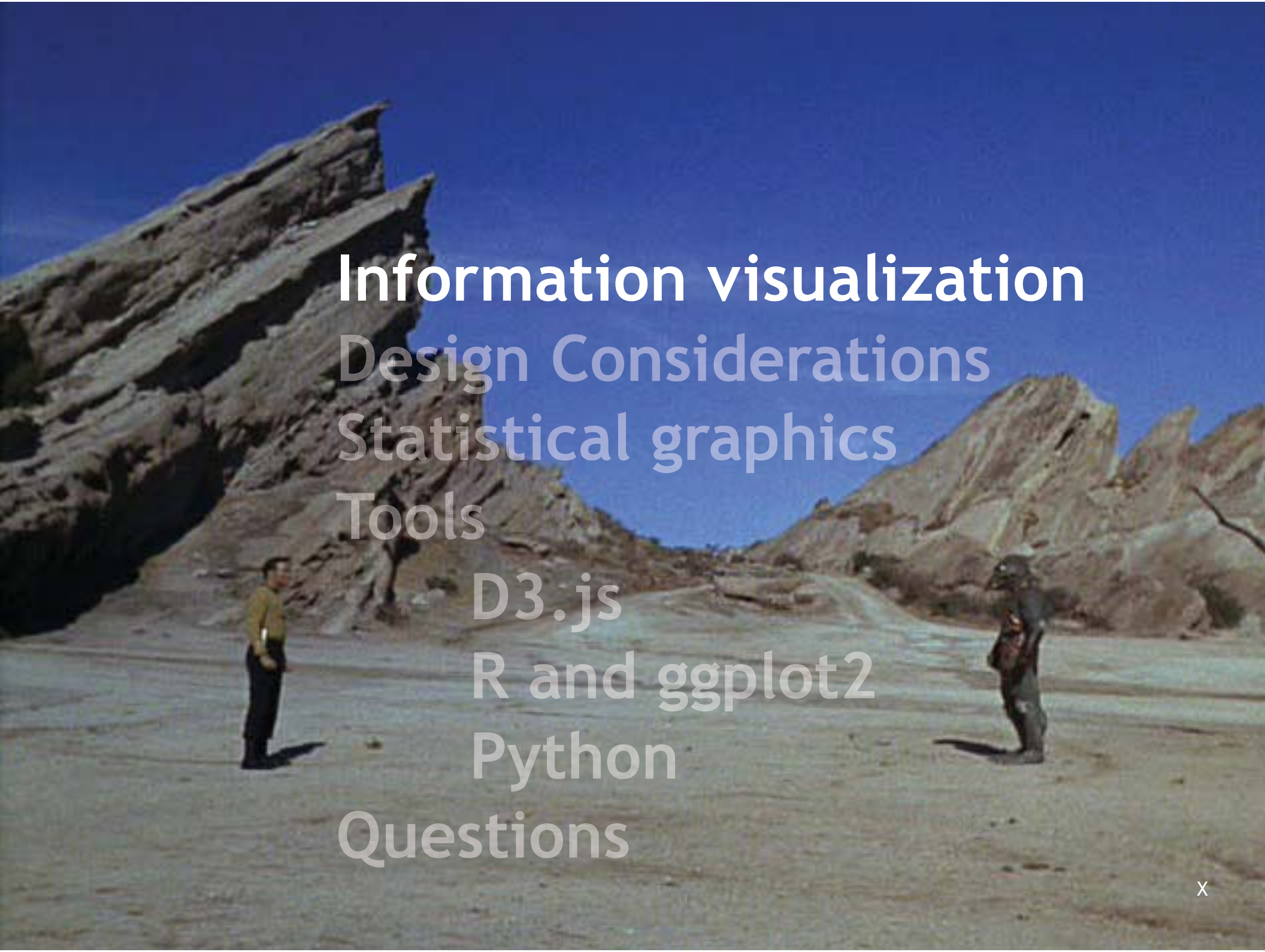
USC Viterbi Data Analytics Boot Camp

Dr. Luciano Nocera
Associate Director Integrated Media Systems Center
Department of Computer Science
University of Southern California

USC Viterbi
Integrated Media
Systems Center

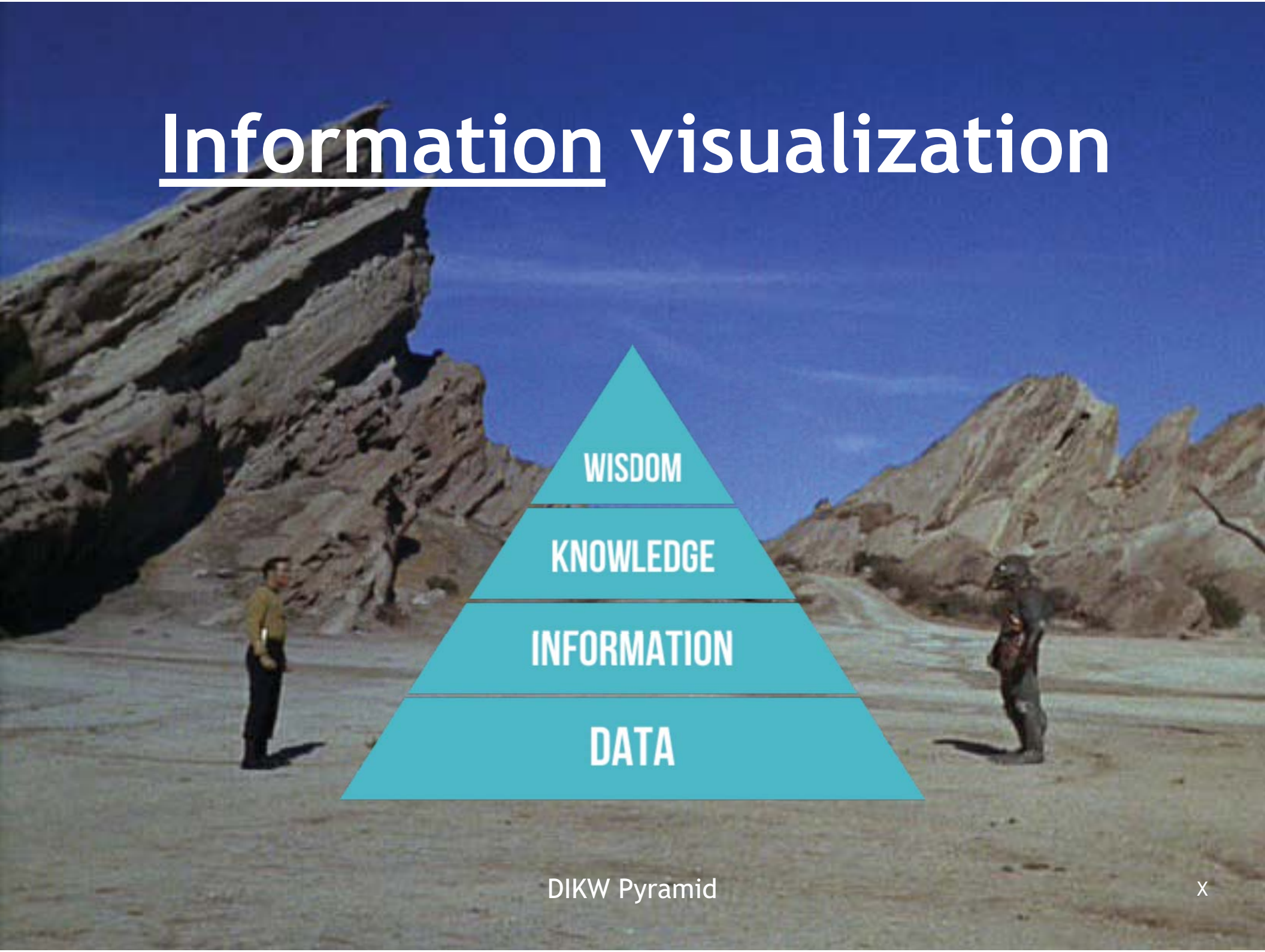
USC University of
Southern California


Information visualization
Design Considerations
Statistical graphics
Tools
D3.js
R and ggplot2
Python
Questions


Information visualization
Design Considerations
Statistical graphics
Tools
D3.js
R and ggplot2
Python
Questions
X

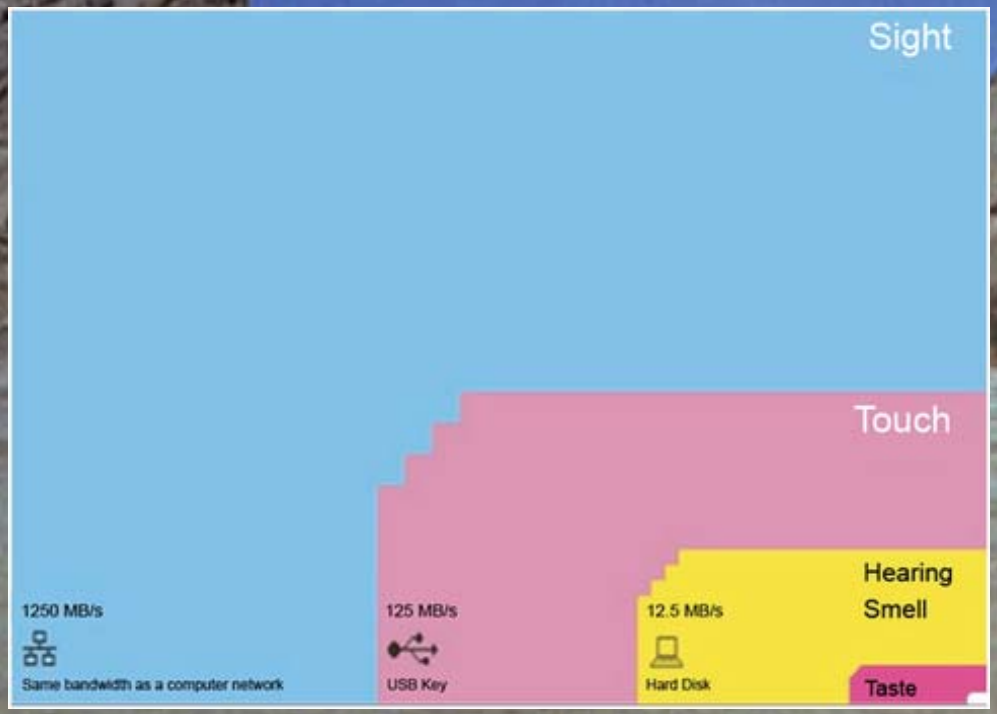
Information visualization

Information visualization



DIKW Pyramid

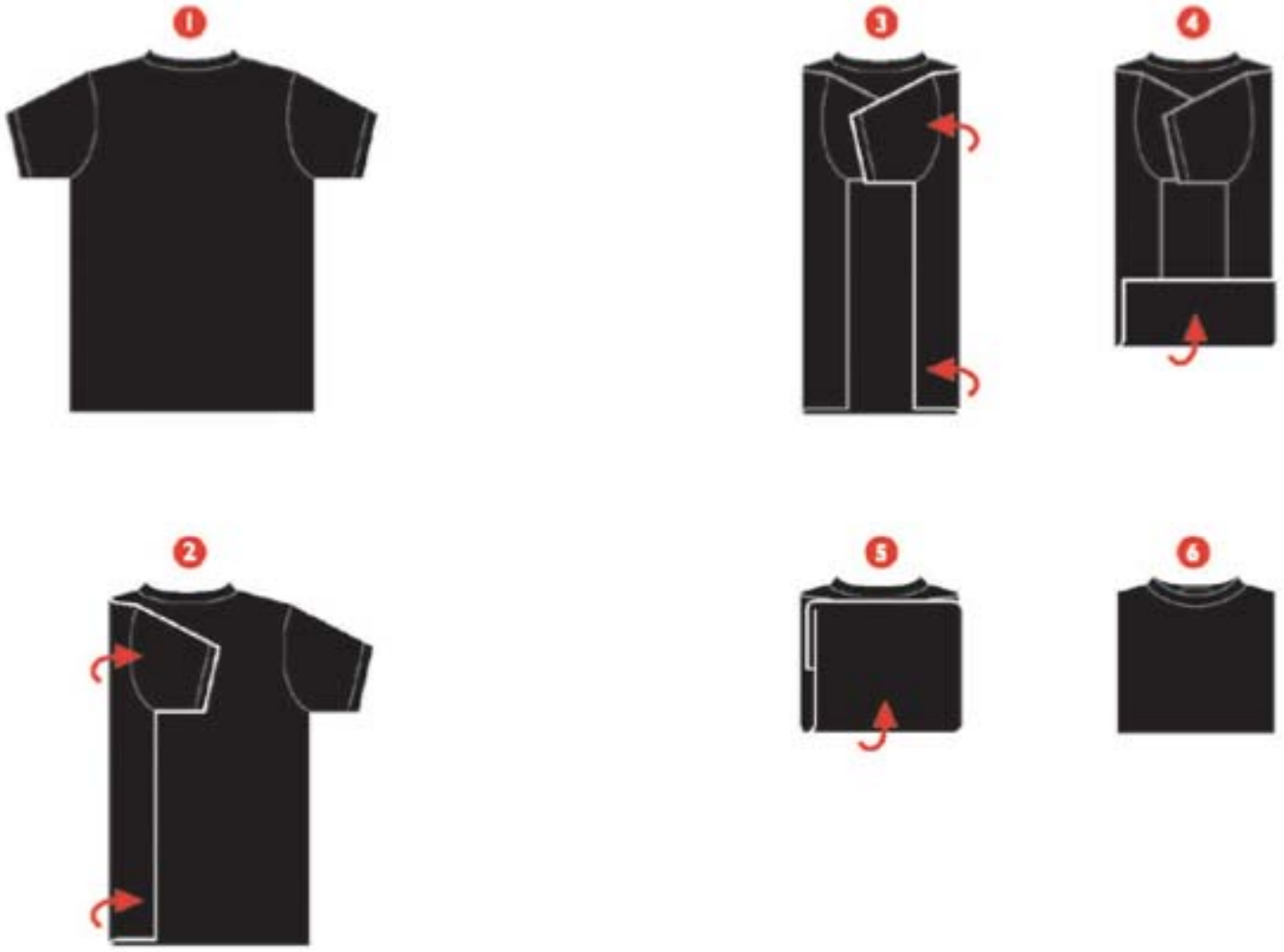
Information visualization



Nørretranders bandwidth of senses
Graphic by David McCandless

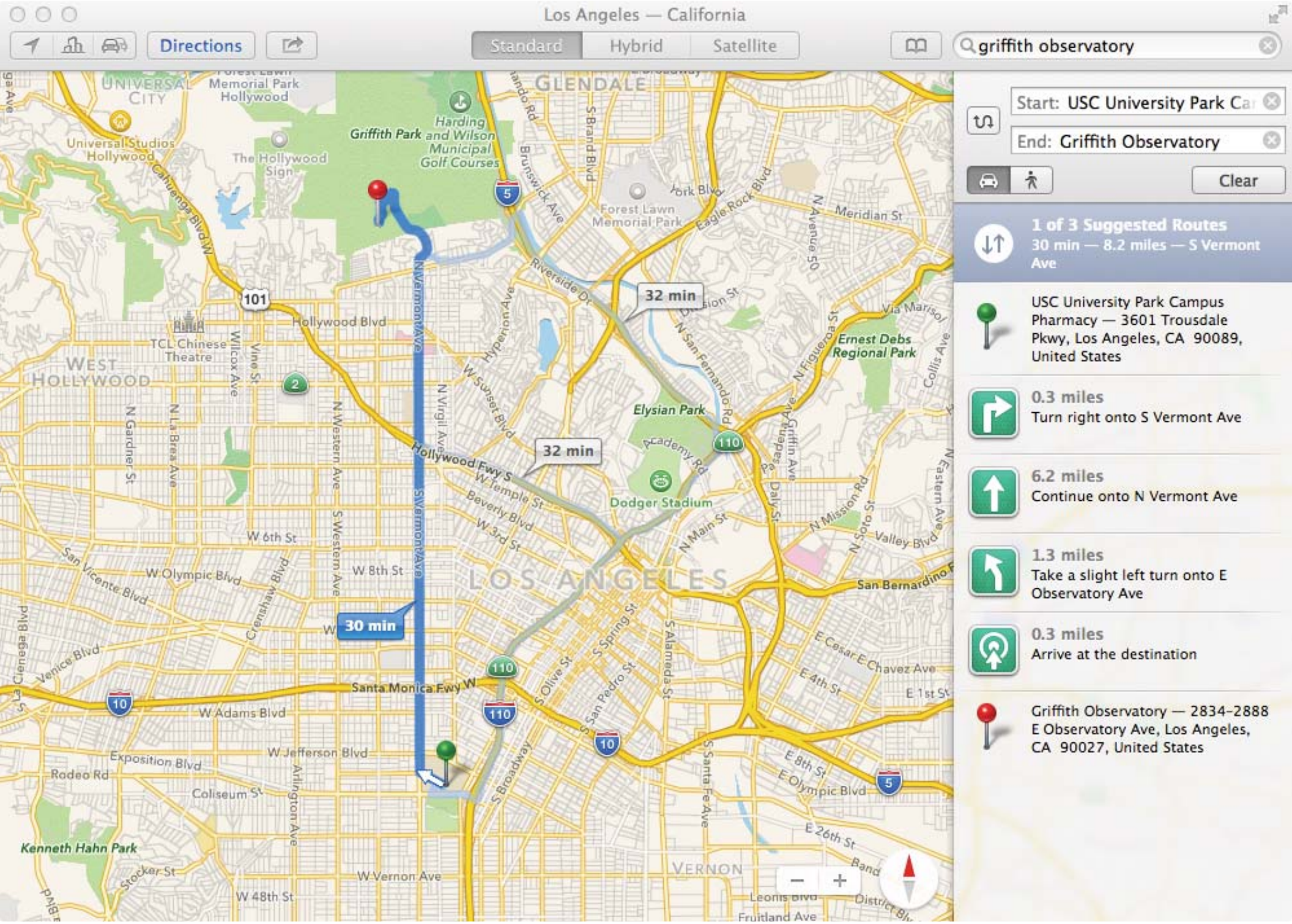
X

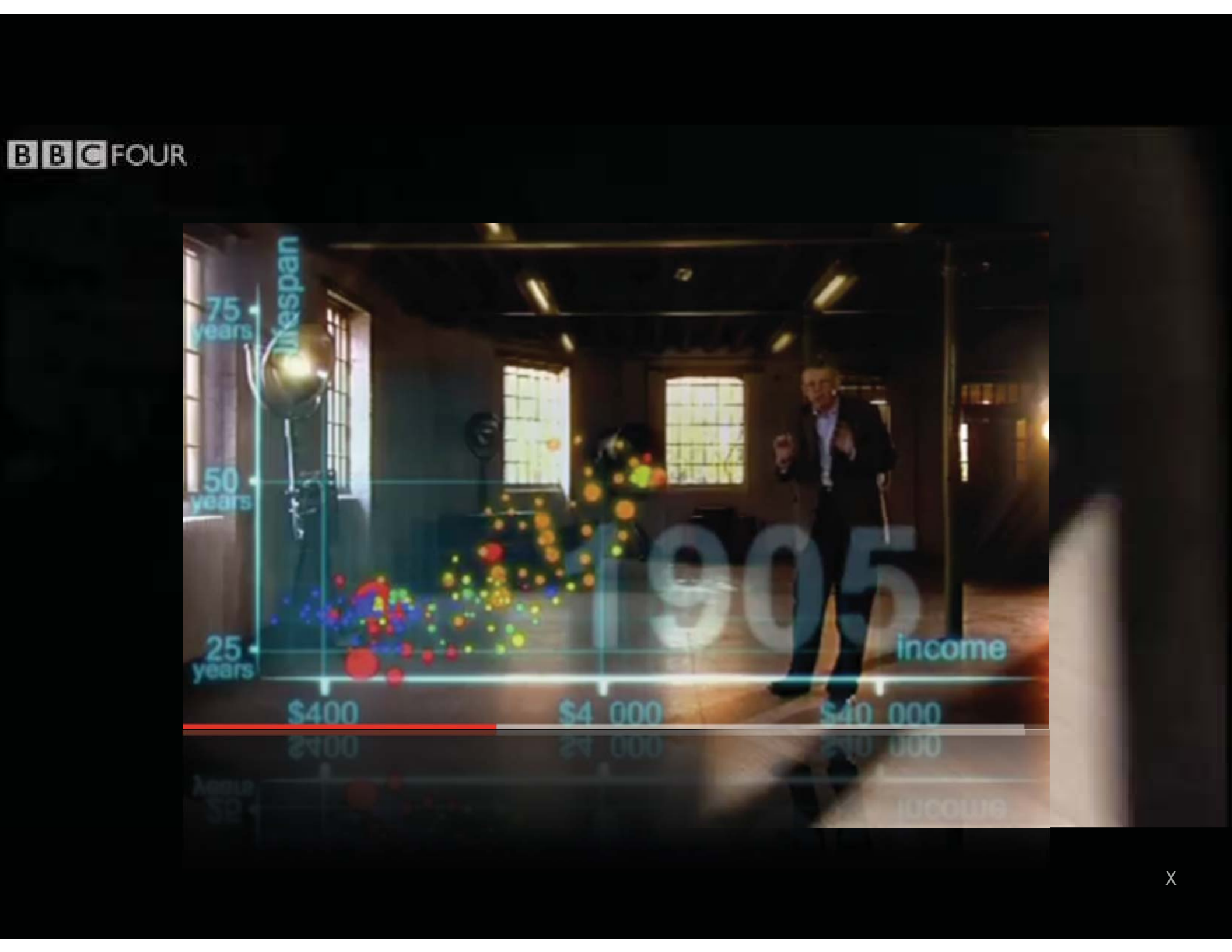

Information visualization
Inform
Describe
Communicate
Explore
Analyze
Decide



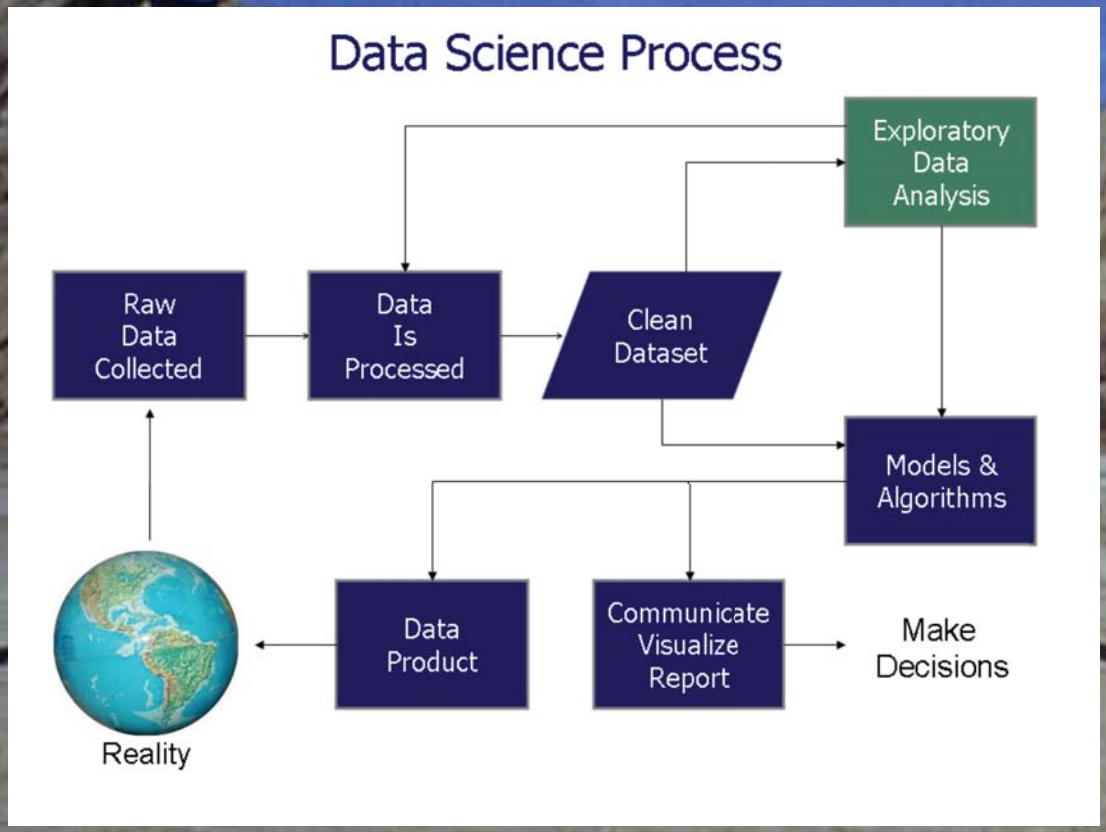
***Wordless Diagrams* (2005) by Nigel Holmes.**

affordances and signifiers

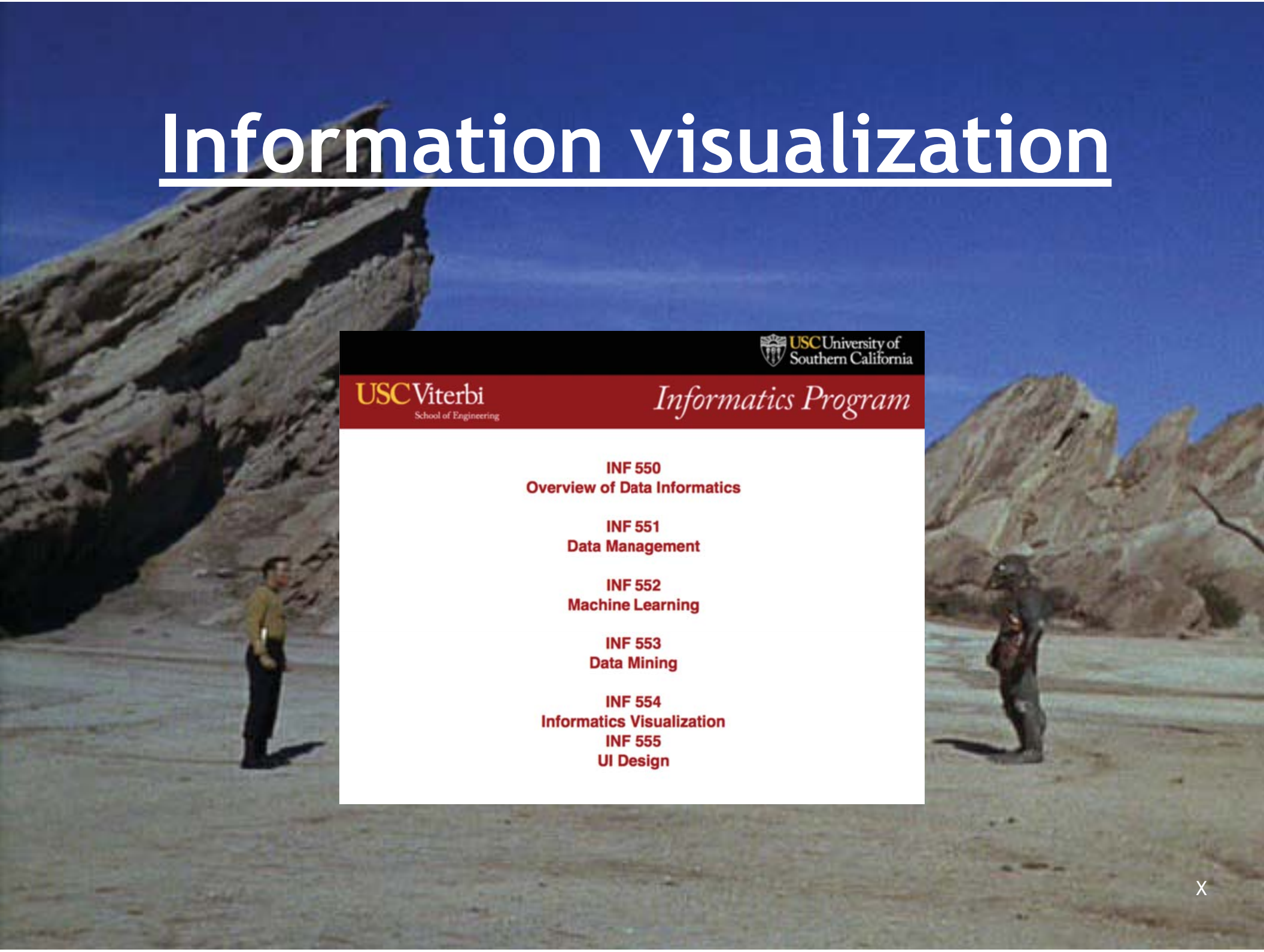

Los Angeles — California
Directions
Standard
Hybrid
Satellite
griffith observatory
Start: USC University Park Ca
End: Griffith Observatory
Clear
1 of 3 Suggested Routes
30 min — 8.2 miles — S Vermont Ave
USC University Park Campus Pharmacy — 3601 Trousdale Pkwy, Los Angeles, CA 90089, United States
0.3 miles
Turn right onto S Vermont Ave
6.2 miles
Continue onto N Vermont Ave
1.3 miles
Take a slight left turn onto E Observatory Ave
0.3 miles
Arrive at the destination
Griffith Observatory — 2834-2888 E Observatory Ave, Los Angeles, CA 90027, United States
32 min
32 min
30 min
Griffith Park
Harding and Wilson Municipal Golf Courses
Forest Lawn Memorial Park
Universal Studios Hollywood
The Hollywood Sign
Hollywood Blvd
Hollywood Fwy S
Santa Monica Fwy W
Elysian Park
Dodger Stadium
Ernest Debs Regional Park
LOS ANGELES
VERNON
GLENDALE
WEST HOLLYWOOD
UNIVERSAL CITY
Kenneth Hahn Park
N Vermont Ave
S Vermont Ave


BBC FOUR
lifespan
75 years
50 years
25 years
1905
income
$400
$4,000
$40,000

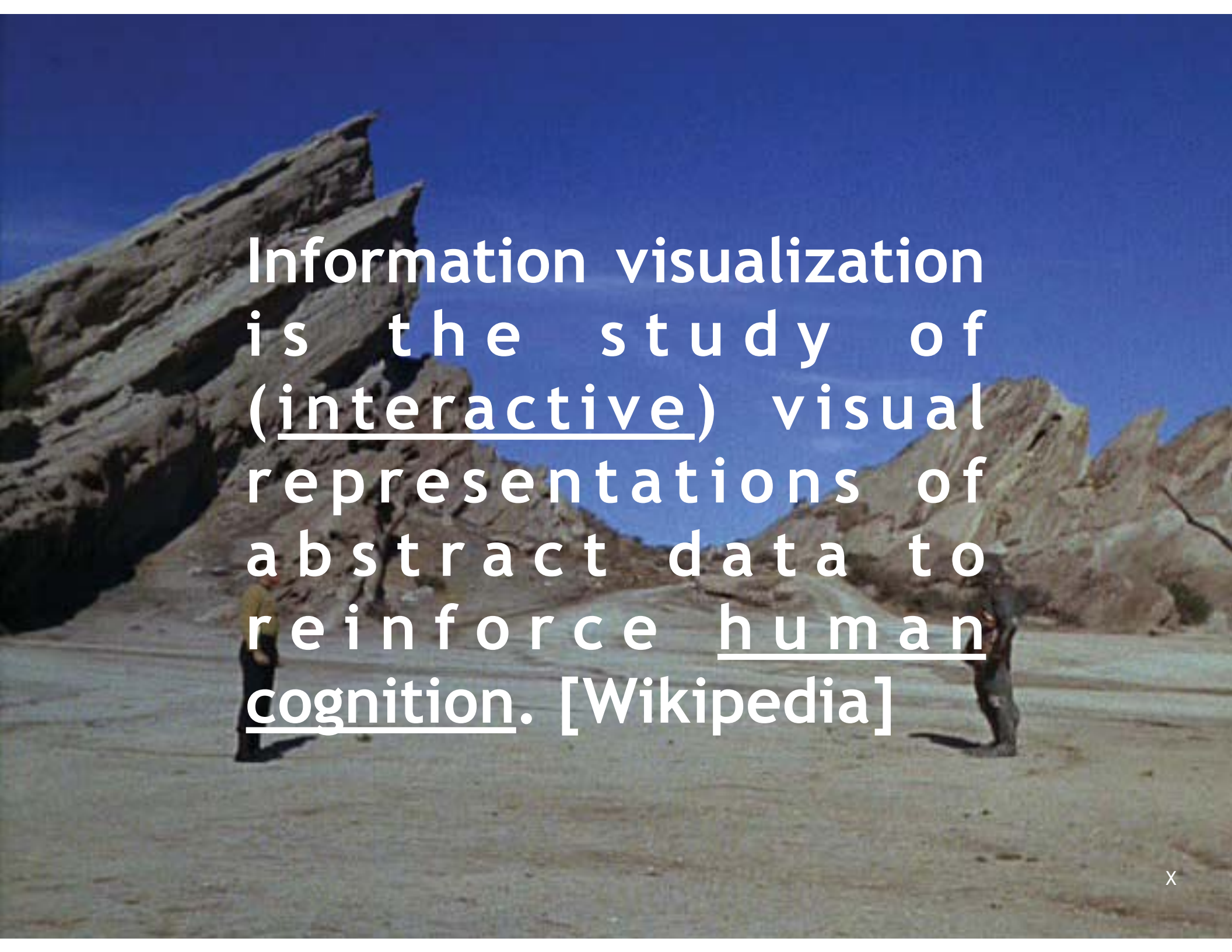
Information visualization



Information visualization



X


Information visualization is the study of (interactive) visual representations of abstract data to reinforce human cognition. [Wikipedia]

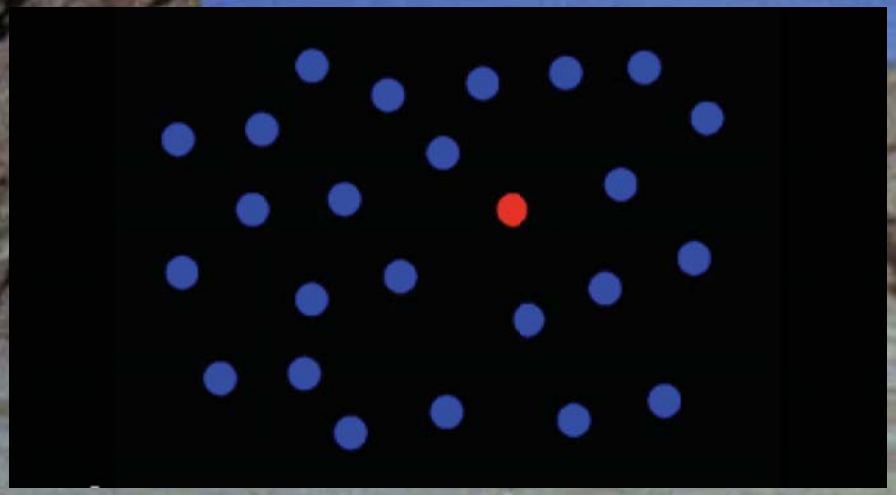
Information visualization is the study of (interactive) visual representations of abstract data to reinforce human cognition. [Wikipedia]

Nørretranders bandwidth of senses
Graphic by David McCandless

You are aware of 0.7%
of what you experience

X

(interactive) visual representations of abstract data to reinforce human cognition. [Wikipedia]

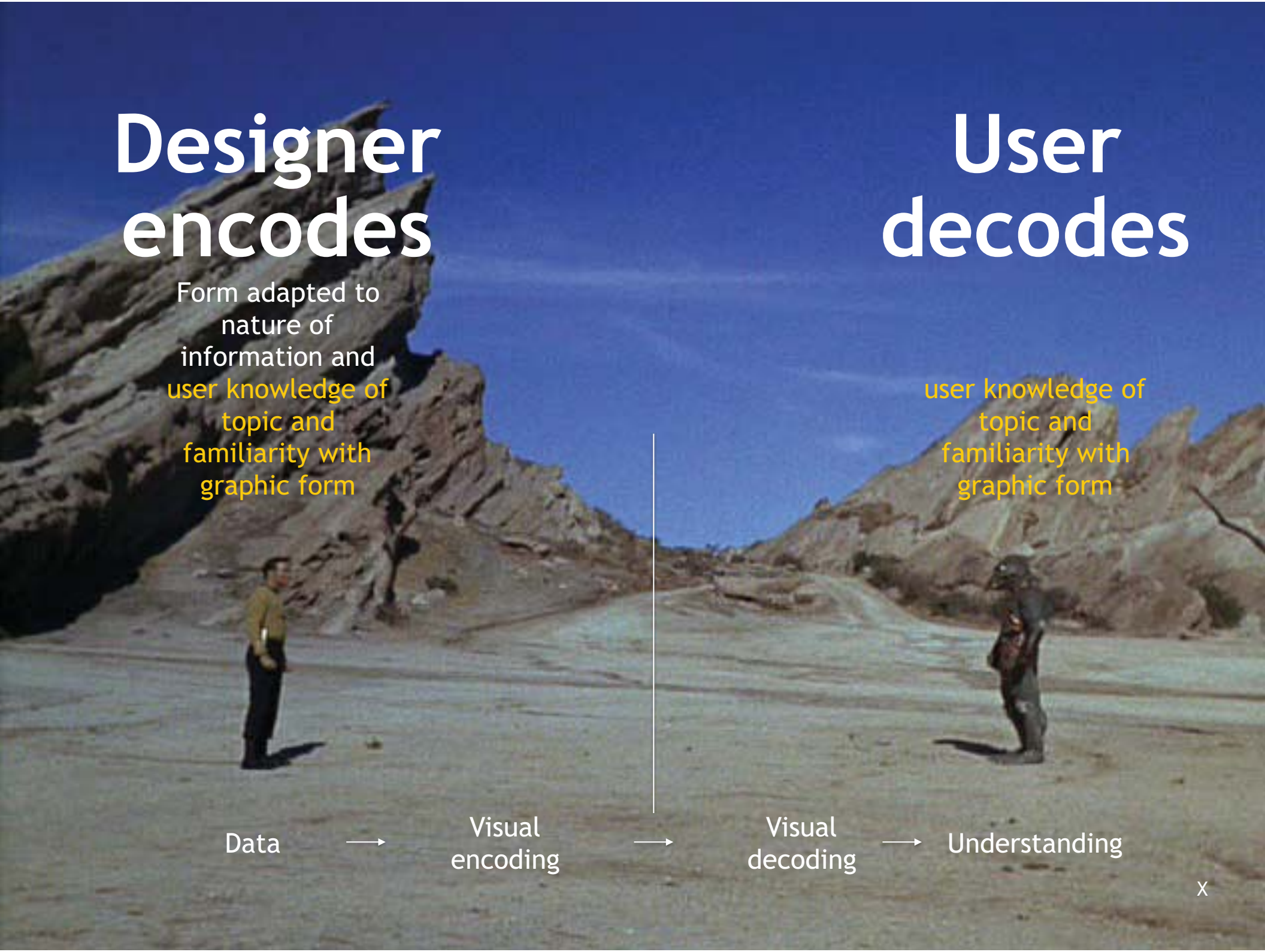


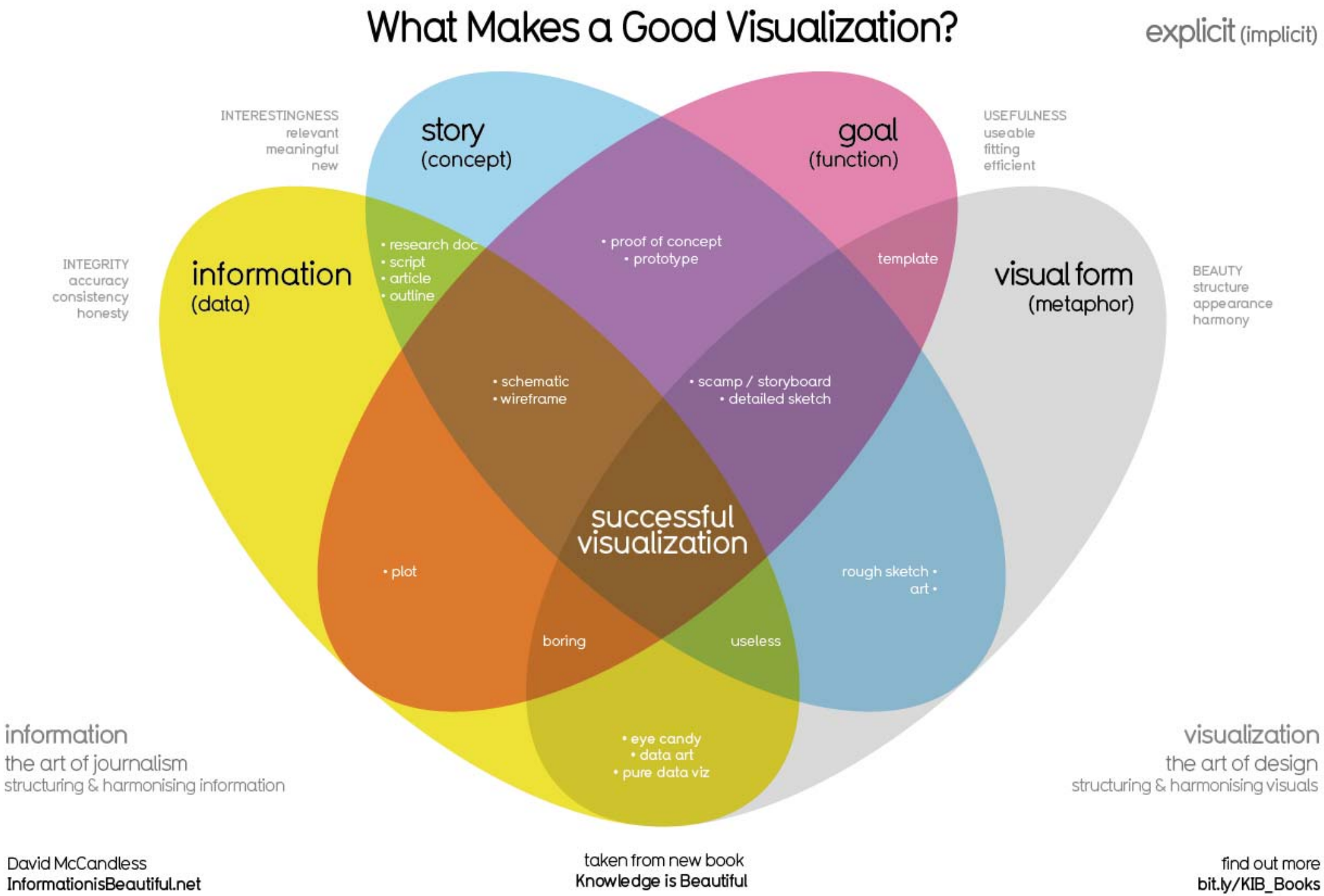
Preattentive features are seen very quickly (typically < 0.1s) Do not require eye movement or focused attention

X

(interactive) visual representations of abstract data to reinforce human cognition. [Wikipedia]


Information visualization
Design Considerations
Statistical graphics
Tools
D3.js
R and ggplot2
Python
Questions
X


Designer encodes
Form adapted to nature of information and user knowledge of topic and familiarity with graphic form
User decodes
user knowledge of topic and familiarity with graphic form
Data
Visual encoding
Visual decoding
Understanding

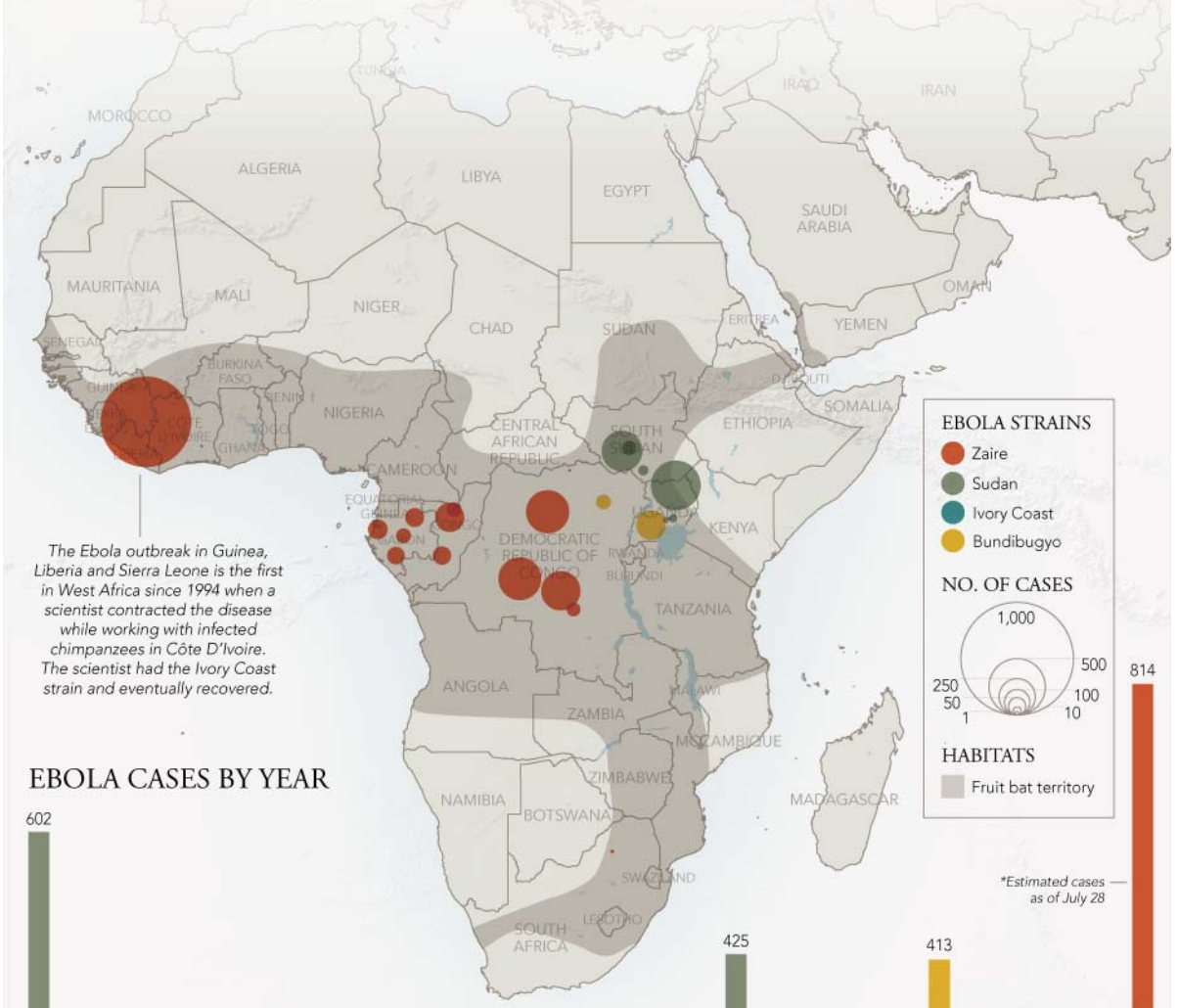


X

David McCandless http://www.informationisbeautiful.net/visualizations/what-makes-a-good-data-visualization/

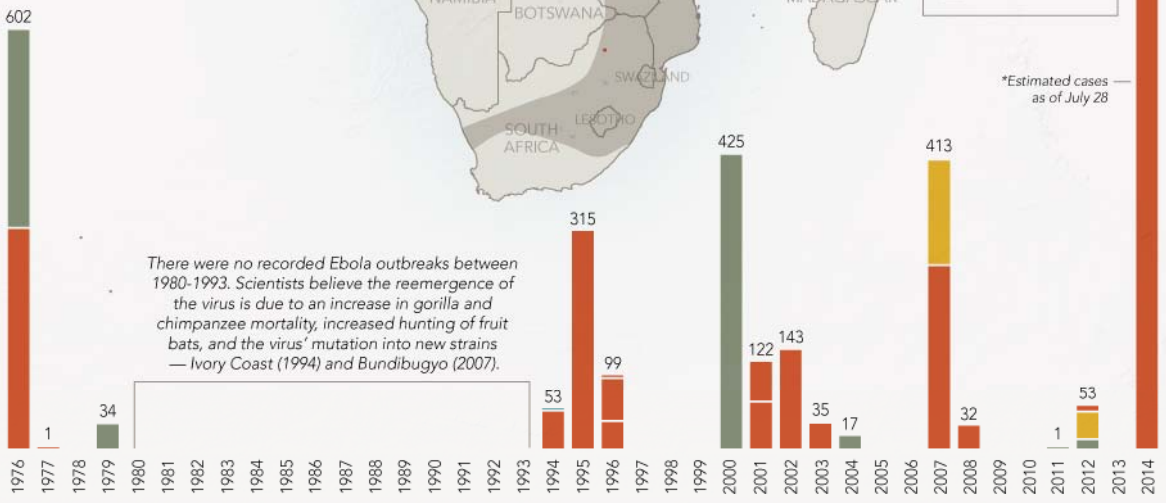
EBOLA'S DEADLY SPREAD

West Africa is experiencing its first Ebola outbreak in two decades, and the deadliest since the first recorded outbreak of the severe viral haemorrhagic fever in 1976 in which 280 deaths were reported. There is no specific treatment or vaccine for Ebola, and the virus can have a fatality rate of up to 90 percent in some outbreaks. Once thought to be carried by primates, new evidence strongly suggests that Ebola is spread by bats.



The Ebola outbreak in Guinea, Liberia and Sierra Leone is the first in West Africa since 1994 when a scientist contracted the disease while working with infected chimpanzees in Côte D'Ivoire. The scientist had the Ivory Coast strain and eventually recovered.

EBOLA CASES BY YEAR



There were no recorded Ebola outbreaks between 1980-1993. Scientists believe the reemergence of the virus is due to an increase in gorilla and chimpanzee mortality, increased hunting of fruit bats, and the virus' mutation into new strains — Ivory Coast (1994) and Bundibugyo (2007).

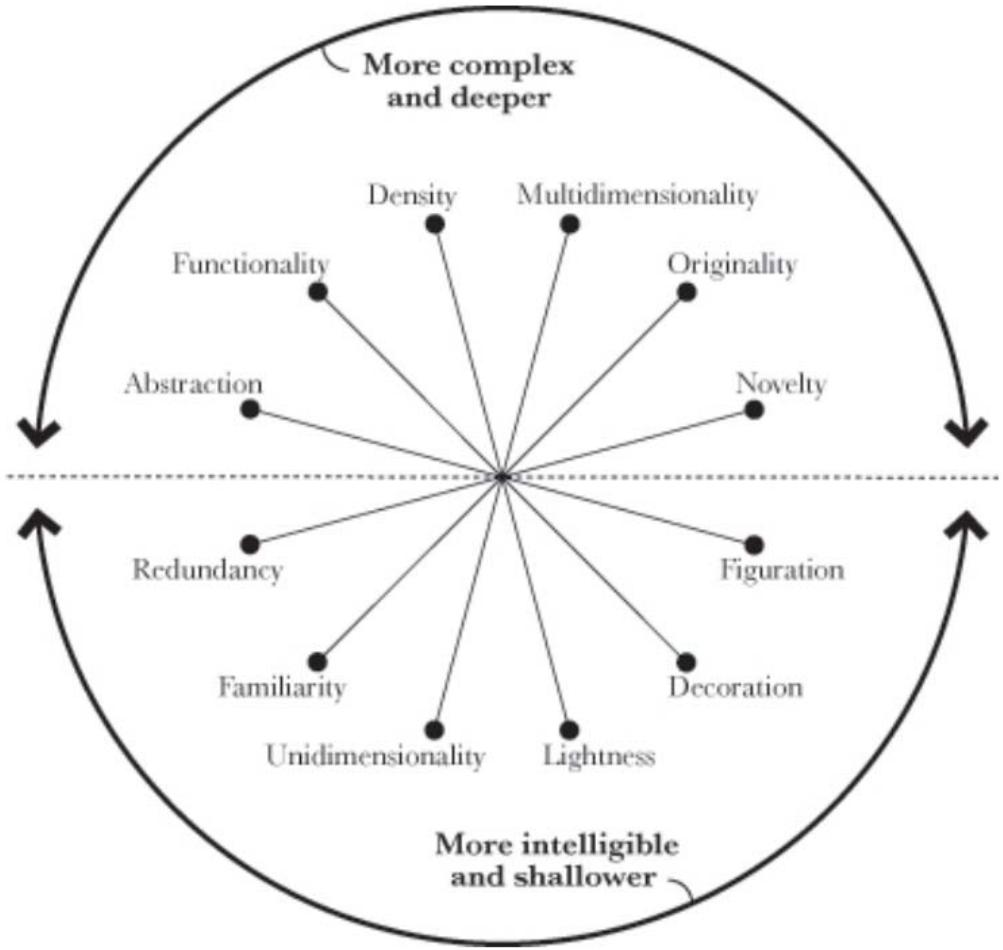
**Estimated cases — as of July 28*

think with Google

What Store Traffic Data Reveals About Black Friday Shopping

This holiday season, mobile will influence more purchases than ever. Consumers are turning to their phones in hundreds of micro-moments throughout the day, helping to inform both online and in-store purchases. In fact, this year 82% of smartphone users will consult their phone while in a store.[1] We took a look at foot traffic patterns[2] over the holiday season and here's what we found:

Cairo's 6D visualization space



The Functional Art: An Introduction to Information Graphics and Visualization, by Alberto Cairo.

More complex and deeper

Abstraction

Distance from referent to representation

Figuration

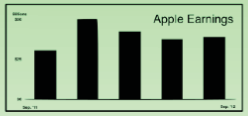
Functionality

Measure of visual elements not directly related to the comprehension of the material

Decoration

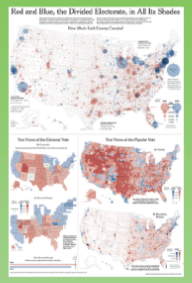
Density

Amount of data displayed in relation to space used

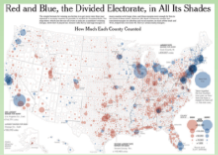


Lightness

Multidimensionality

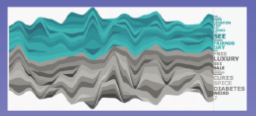


Number of layers of depth, and number of forms used to encode the data

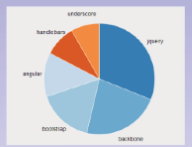


Unidimensionality

Originality



How challenging is the form of visualization used



Familiarity

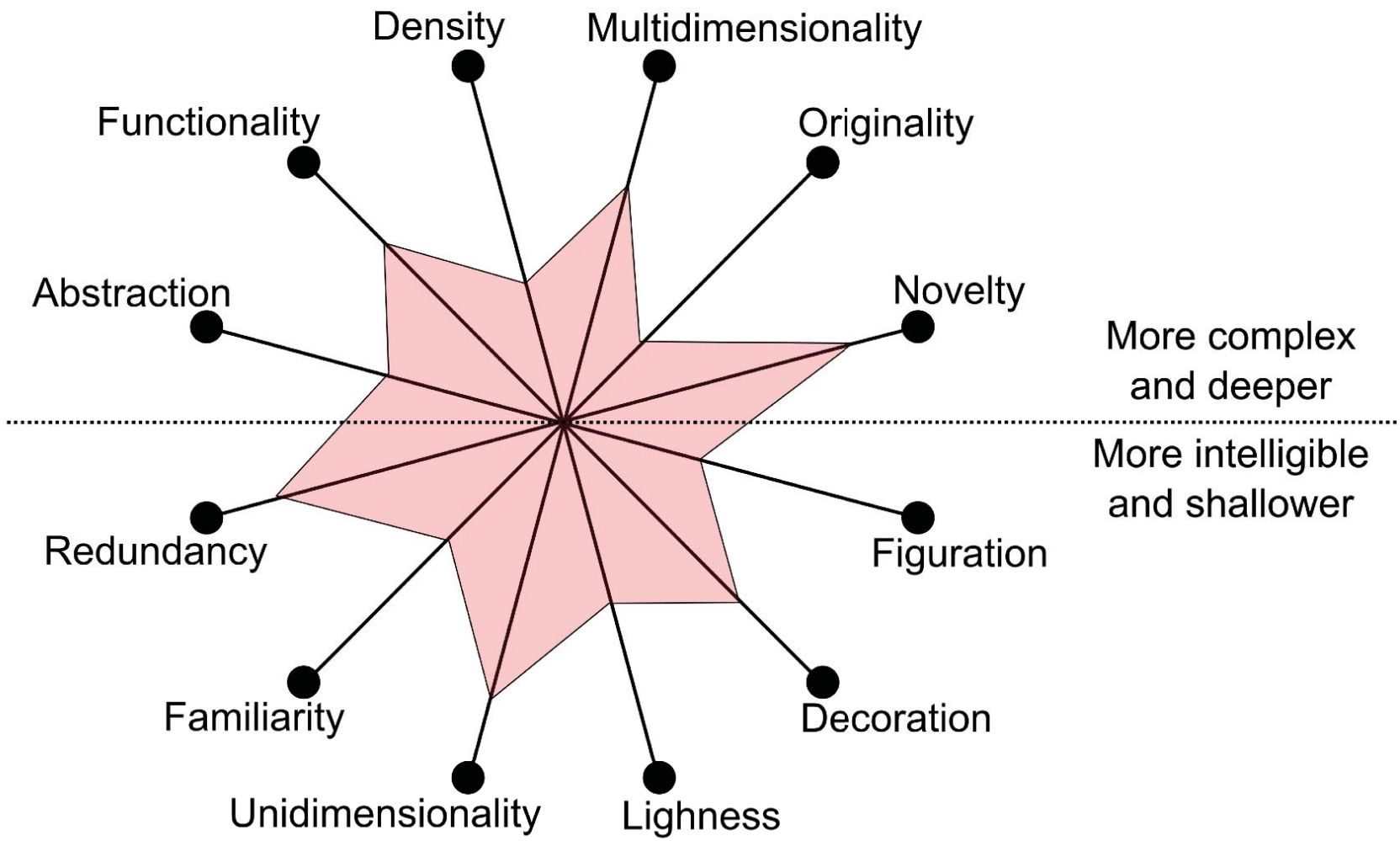
Novelty

Explain multiple things at once or explain the same things multiple times

Redundancy

More intelligeable and shallower

X


Density
Multidimensionality
Functionality
Originality
Abstraction
Novelty
More complex
and deeper
More intelligible
and shallower
Redundancy
Figuration
Familiarity
Decoration
Unidimensionality
Lighness

As for a picture, if it isn't worth a thousand words, the hell with it.
Ad Reinhardt

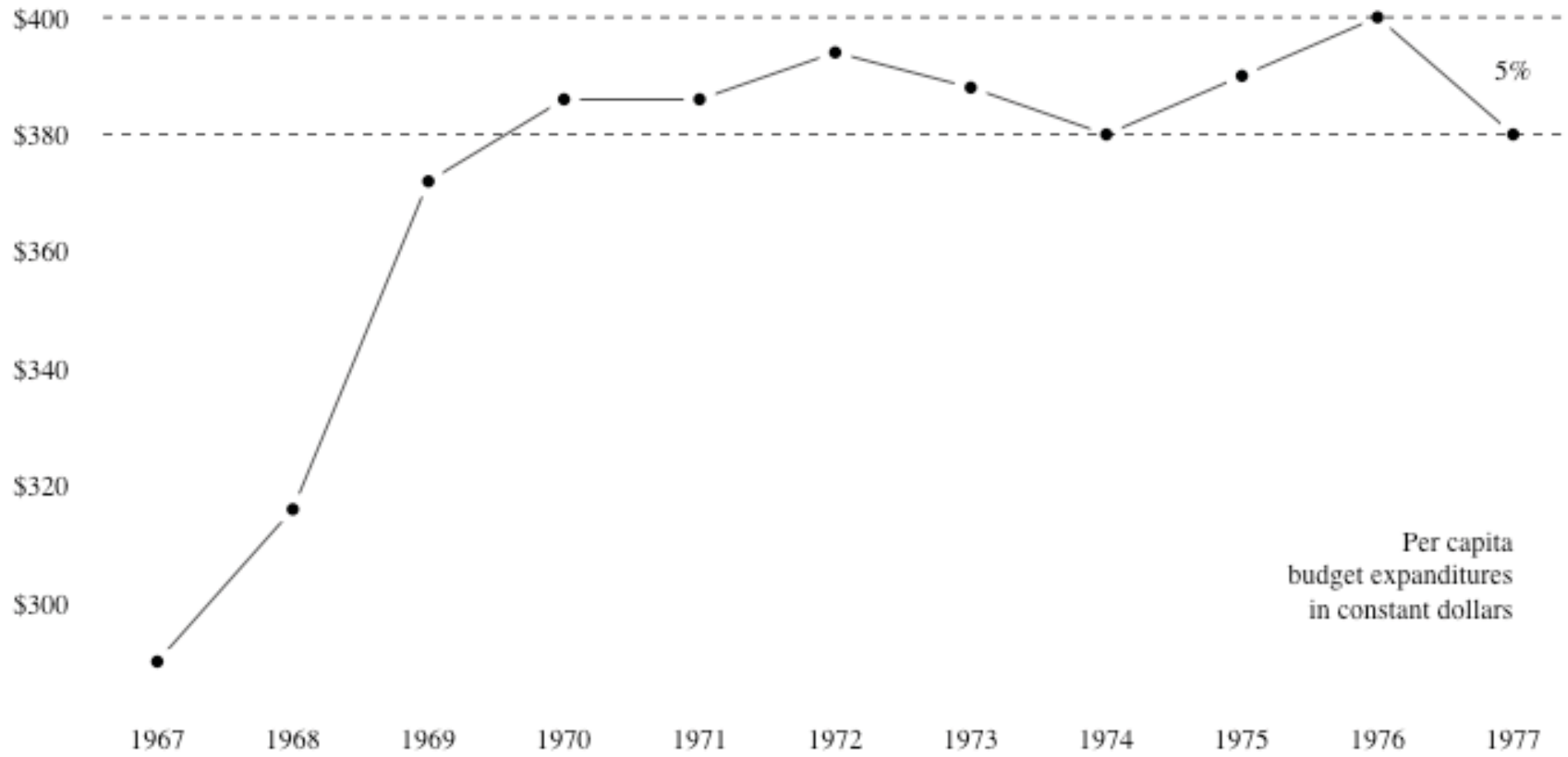
Tufte's principles of design

Elegance in visuals is attained when the complexity of the data matches the simplicity of the design

$$\text{Data-ink ratio} = \frac{\text{Data ink}}{\text{Total ink used}}$$

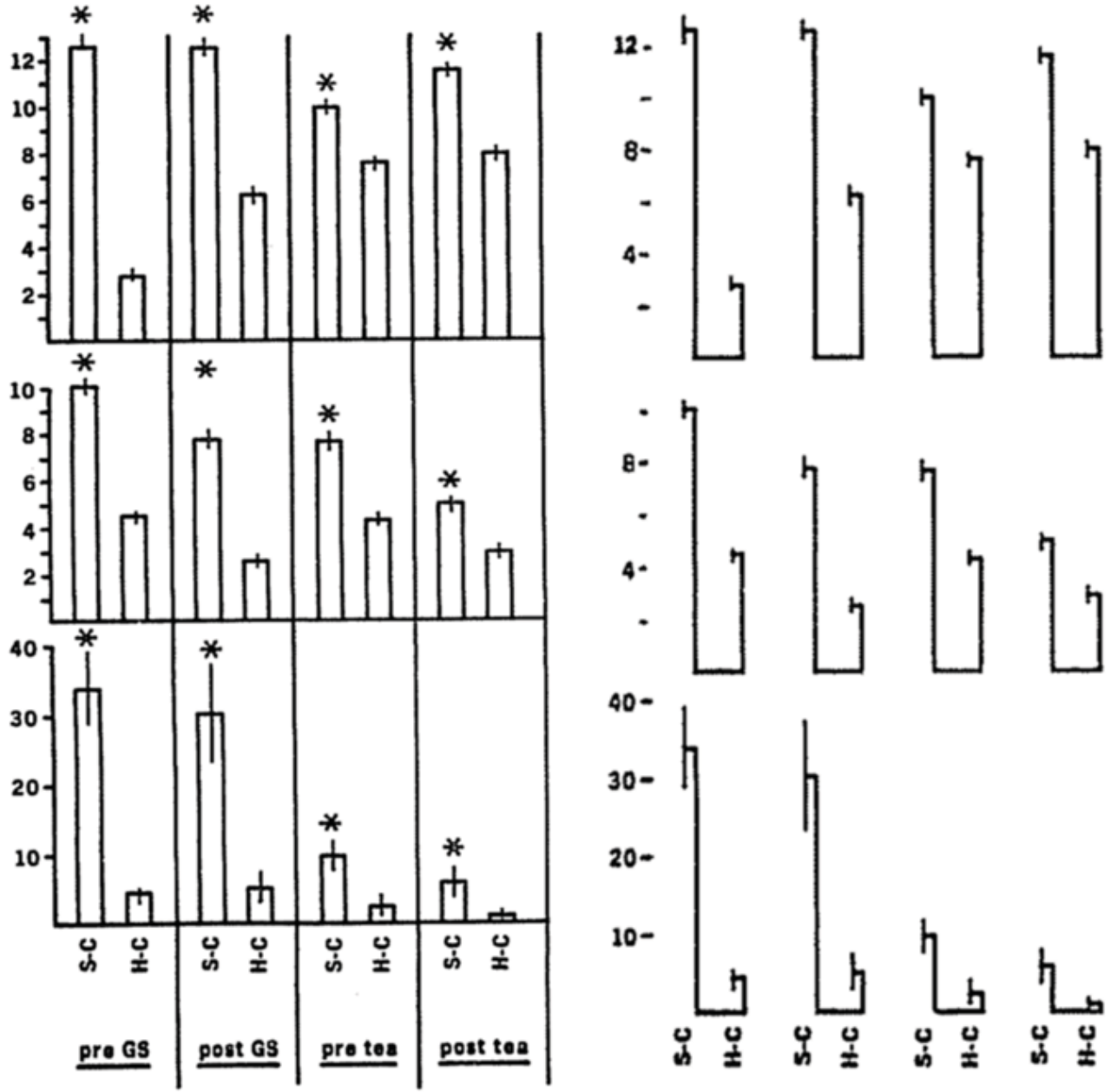
1. Above all else show data.
2. Maximize the data-ink ratio.
3. Erase non-data-ink.
4. Erase redundant data-ink.
5. Revise and edit

Tufte's minimalistic line plot



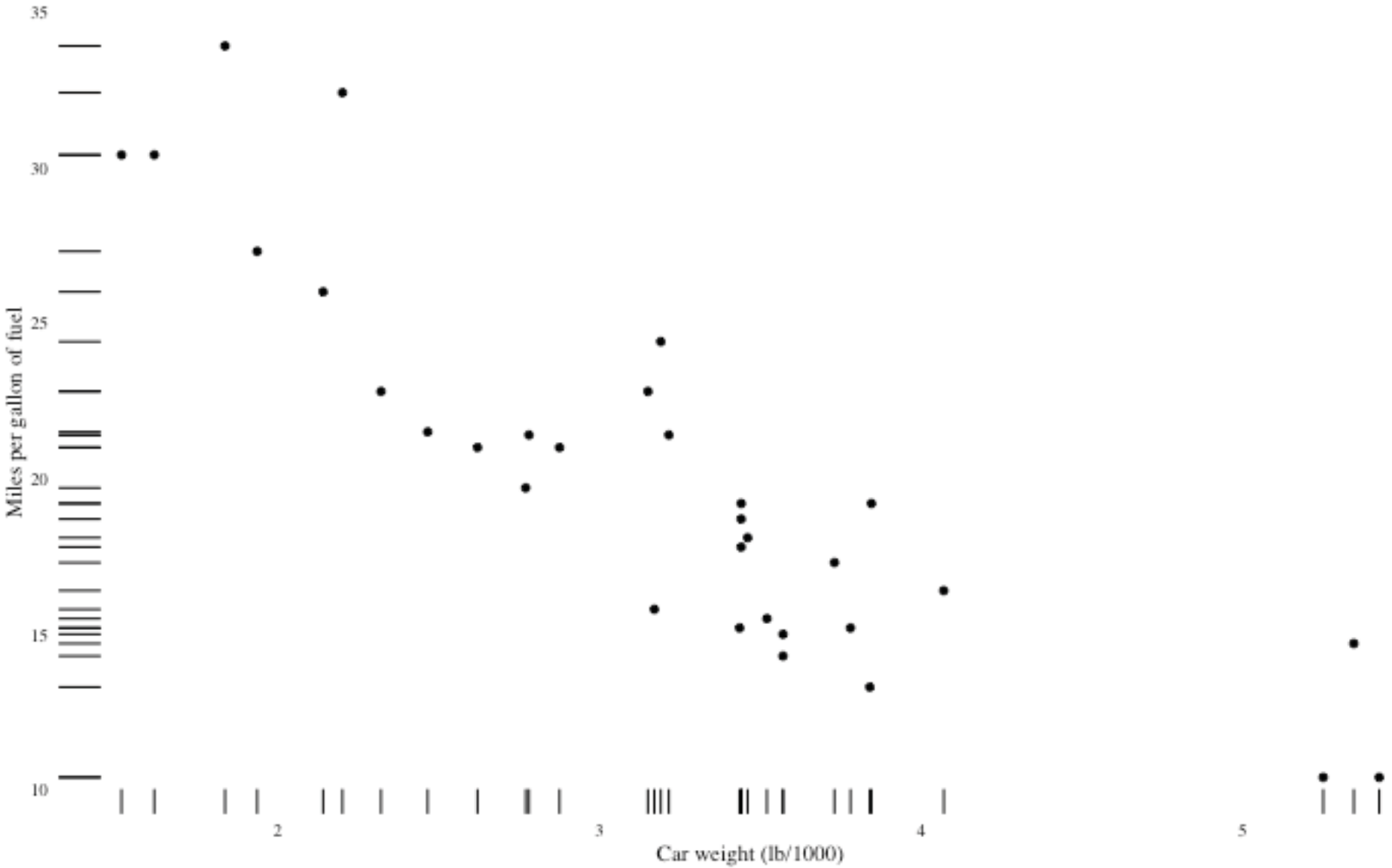
The Visual Display of Quantitative Information, E. Tufte, page 68

Tufte's minimalistic bar chart



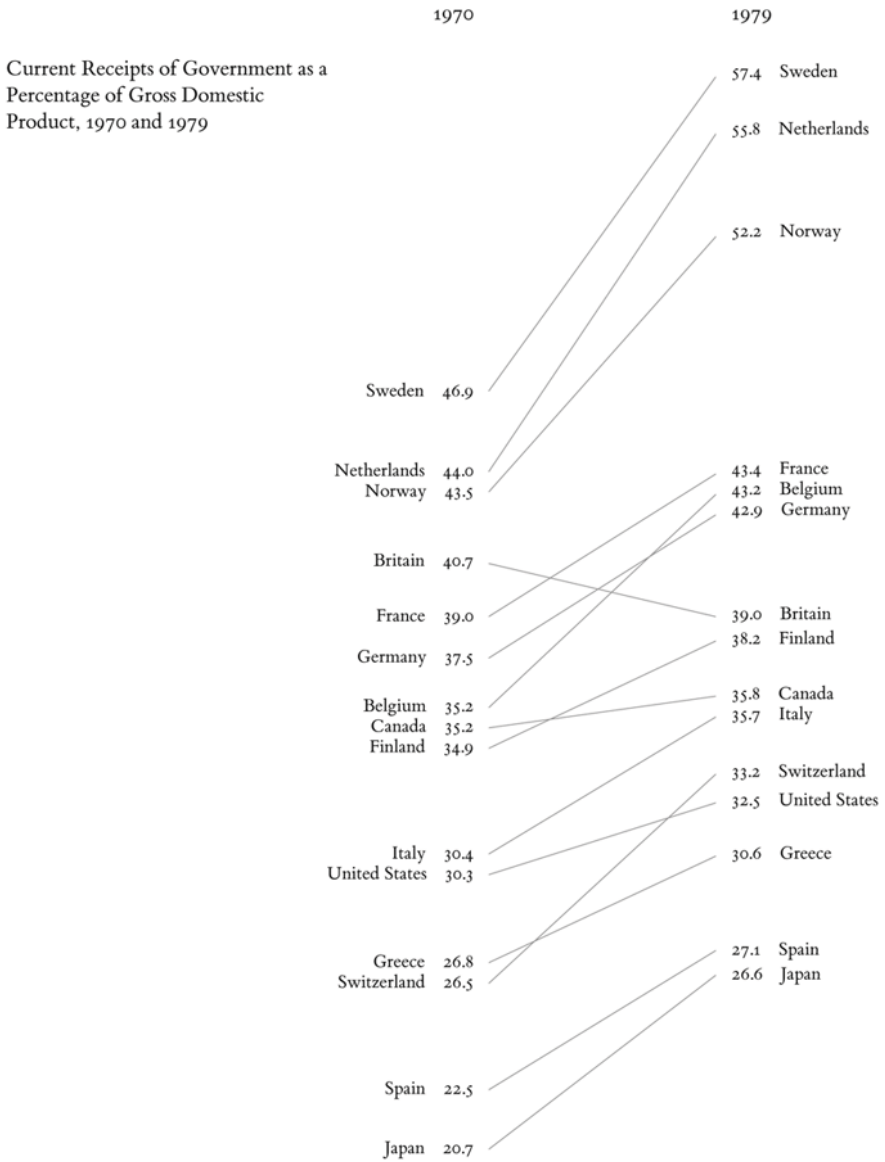
The Visual Display of Quantitative Information, E. Tufte, page 102

Dot-dash plot



The Visual Display of Quantitative Information, E. Tufte, page 133

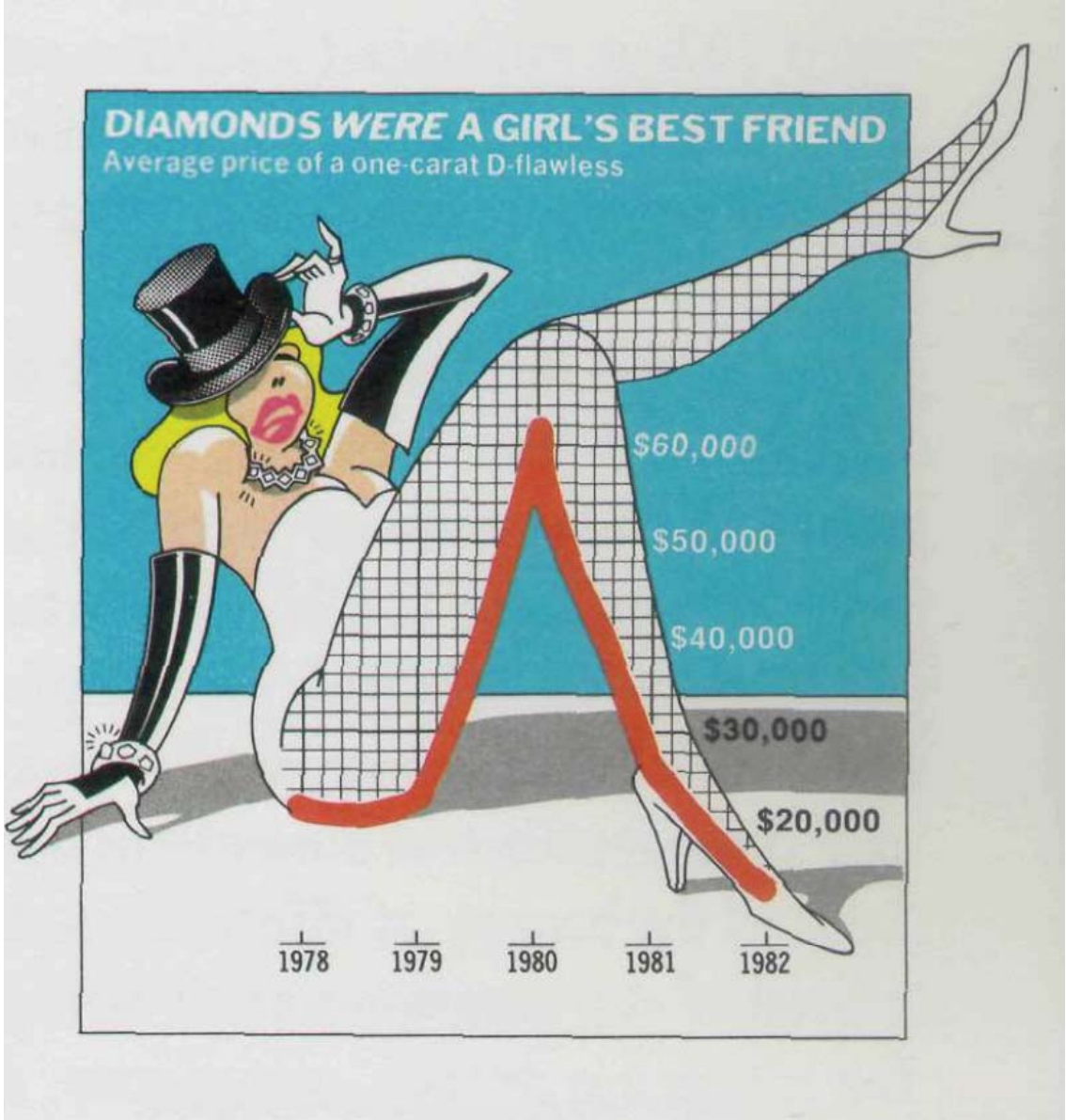
Slopegraph



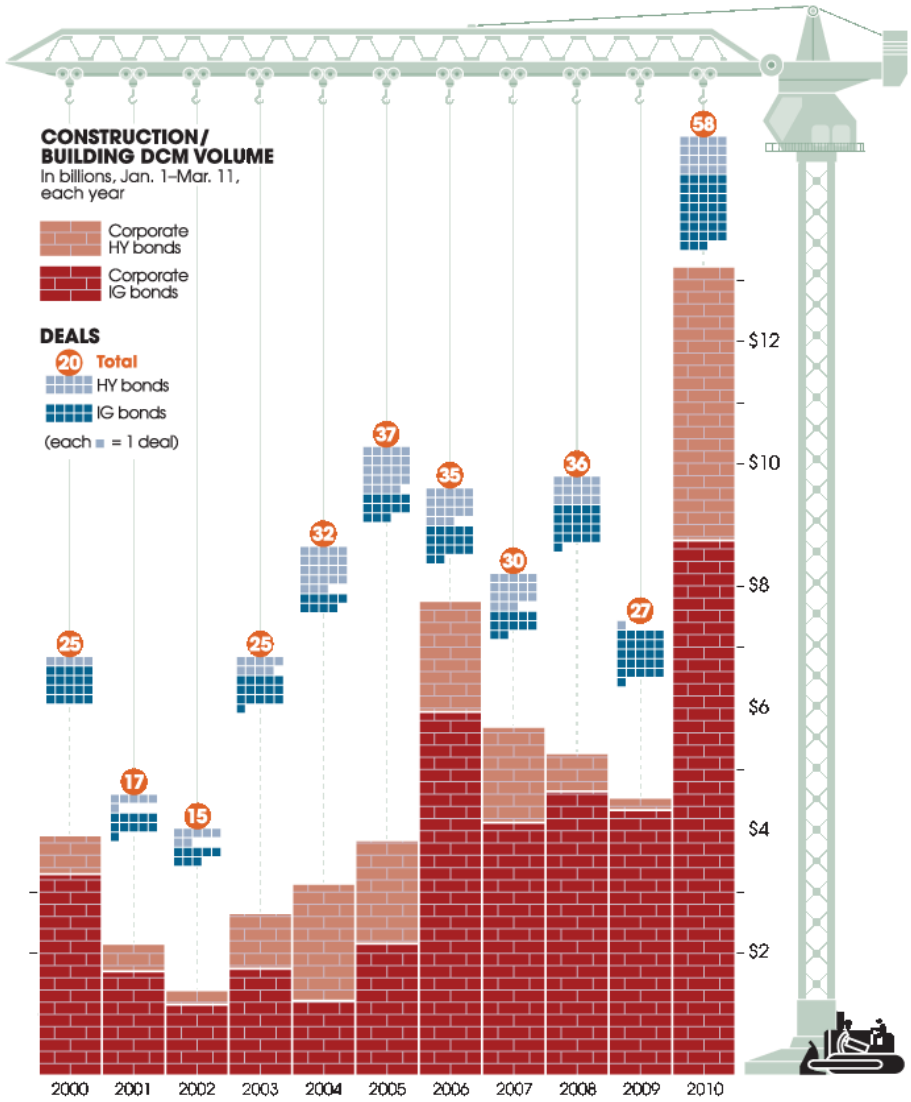
The Visual Display of Quantitative Information" p. 158

Line chart by Nigel Holmes

Use humor to instill affection in readers for numbers and charts



Stacked bar chart by Nigel Holmes



http://nigelholmes.com/graphic/ii-construction/

X

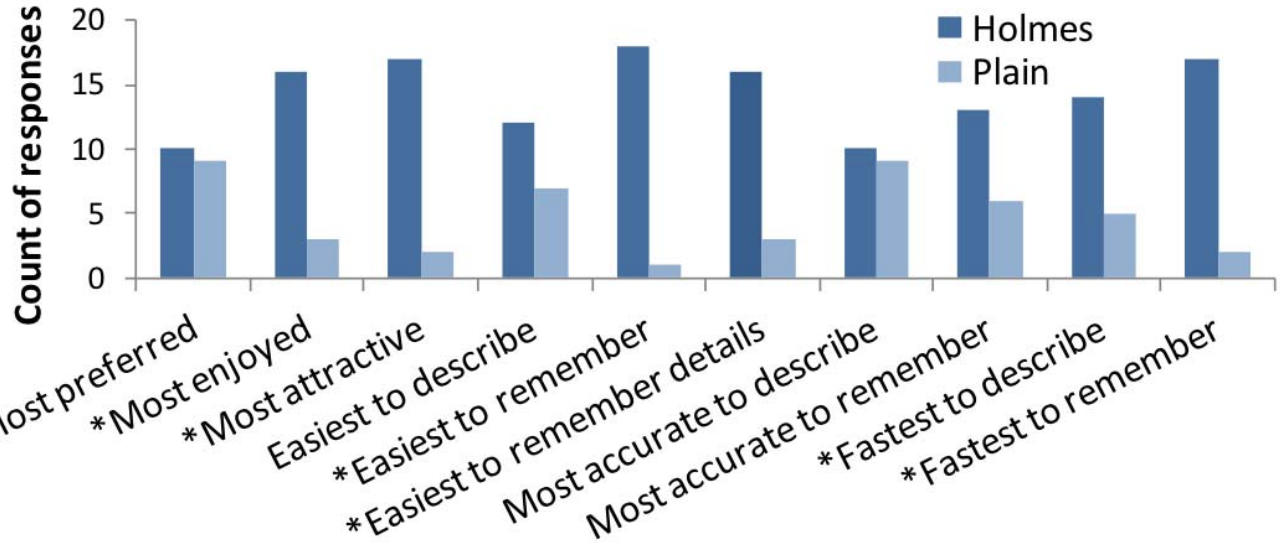
Which is better for comprehension and memorability?

No significant difference:

- between plain & image charts for interactive interpretation & accuracy
- in recall accuracy after a five-minute gap

Significantly better for Holmes:

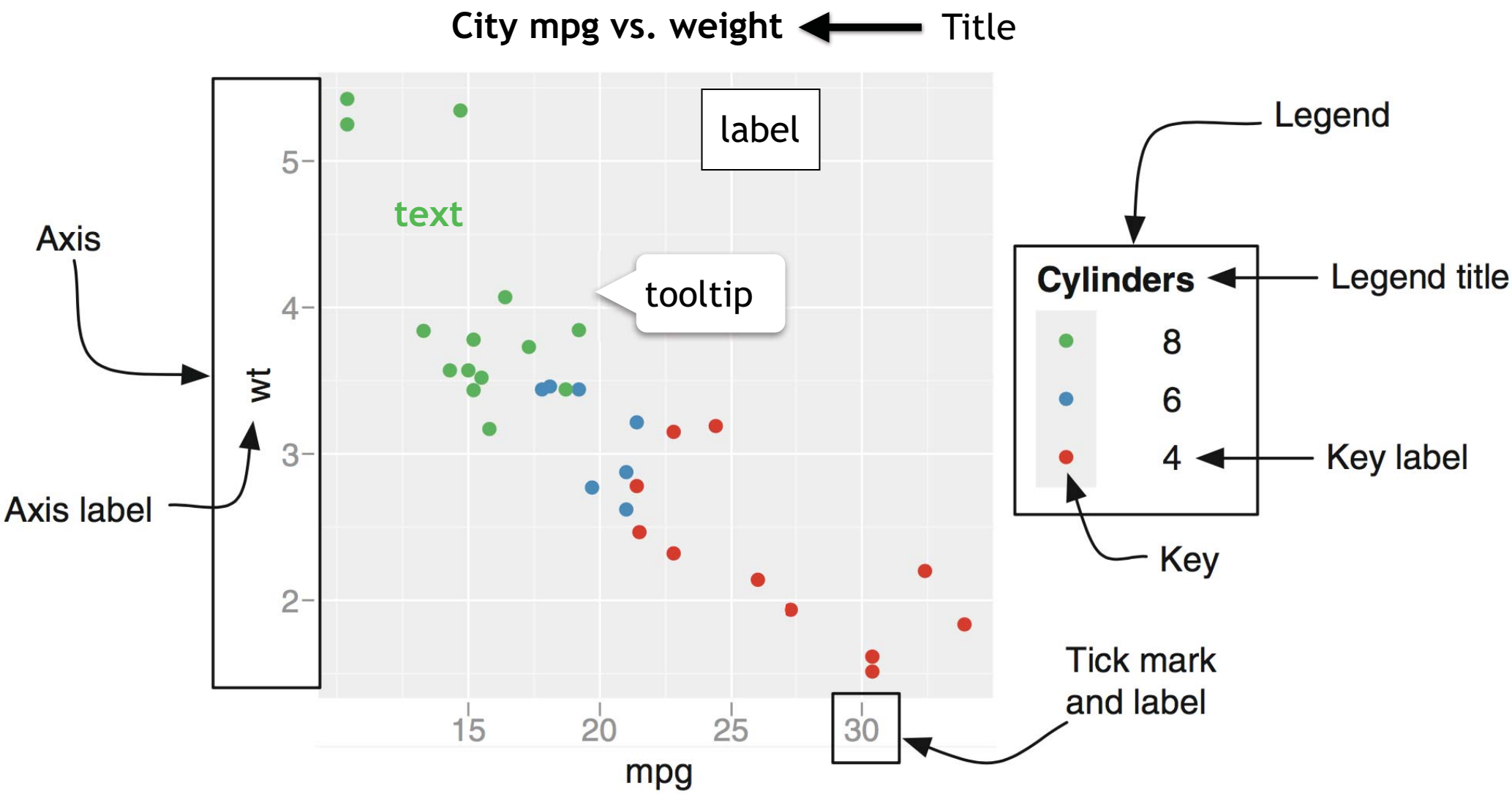
- Holmes better in recall of both chart topic and details after 2-3 weeks
- value messages in the Holmes charts more often than in the plain charts
- Holmes more attractive, most enjoyed
- Holmes were easiest and fastest to remember



Bateman, S., Mandryk, R.L., Gutwin, C., Genest, A.M., McDine, D., Brooks, C. 2010. Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In *ACM Conference on Human Factors in Computing Systems (CHI 2010)*, Atlanta, GA, USA. 2573-2582.

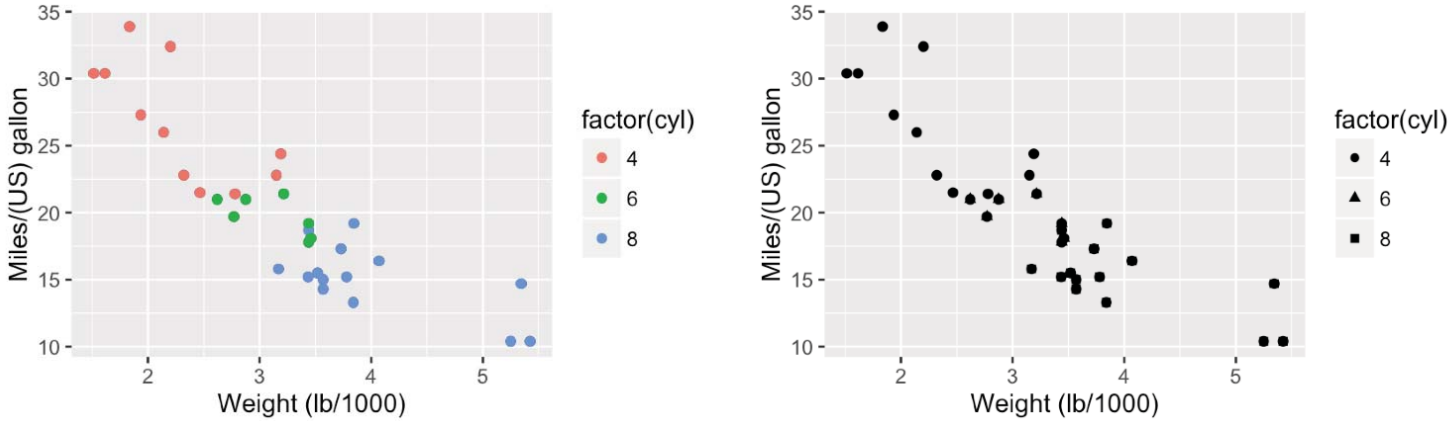
Good Design Practices

Title, axes, and legend

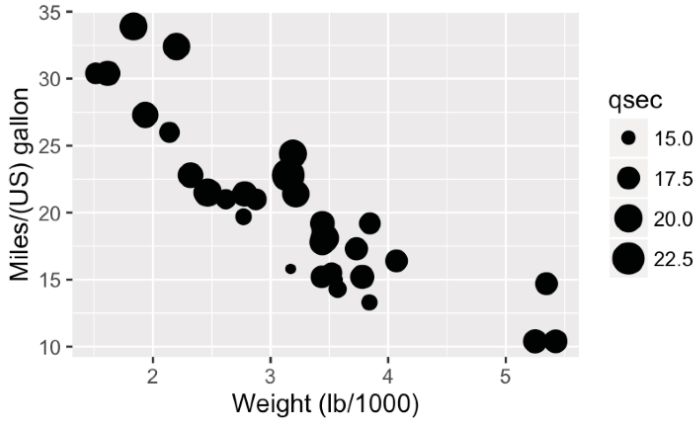


Different features work better with different types of variables

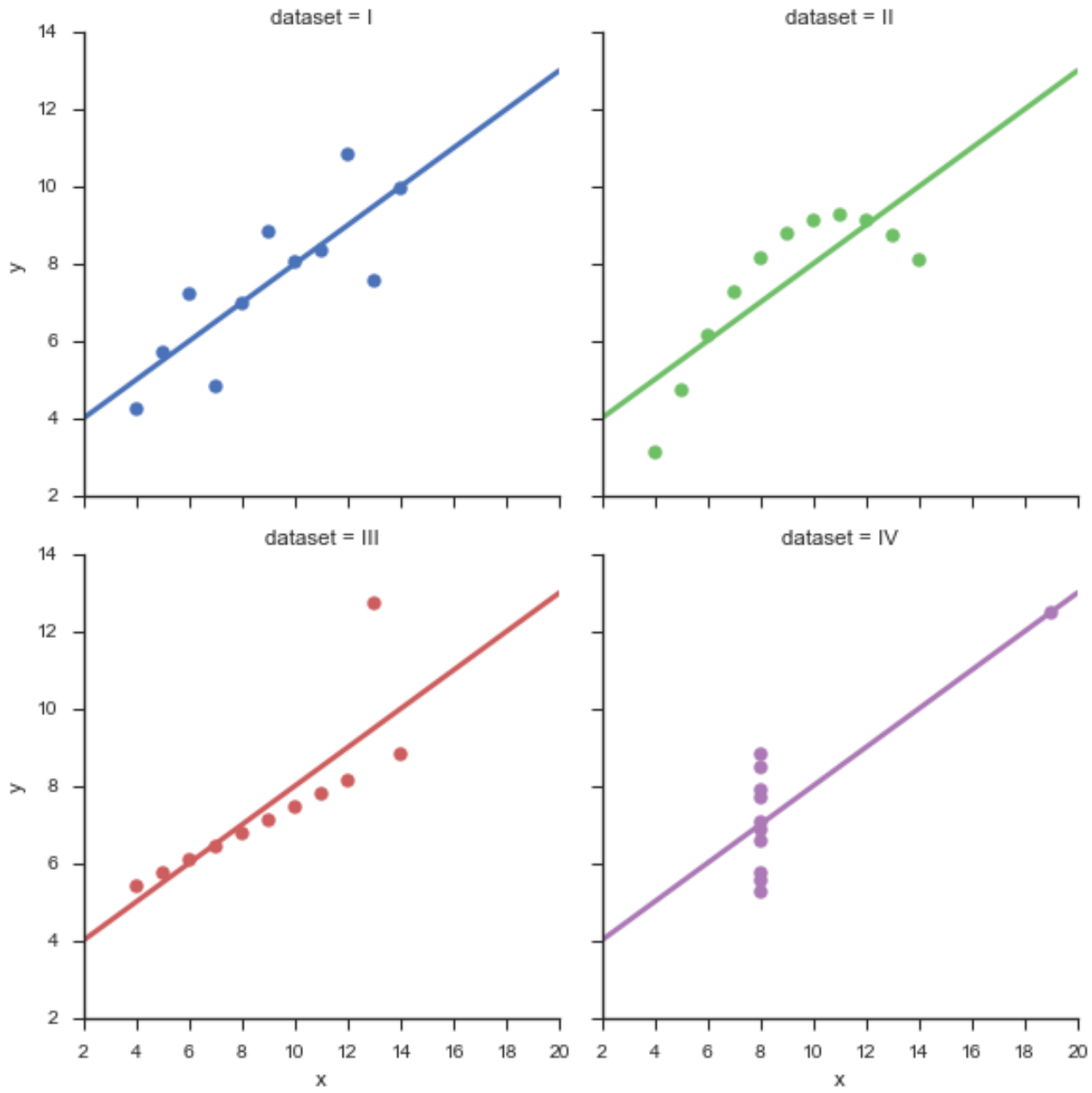
- Color & shape work well with categorical variables



- Size works well with continuous data



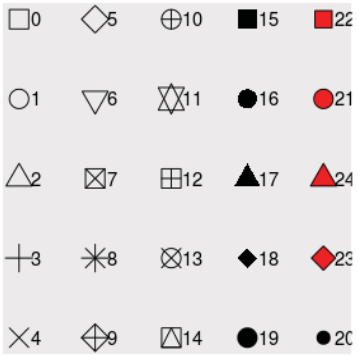
Series of plots usually work better than a single complex plot



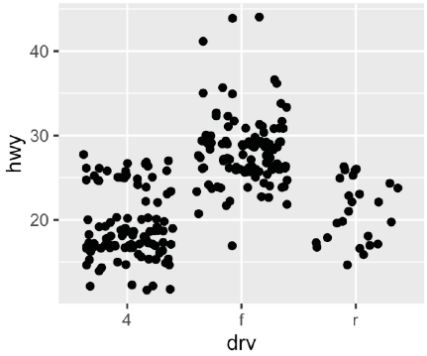
X

Dealing with overplotting

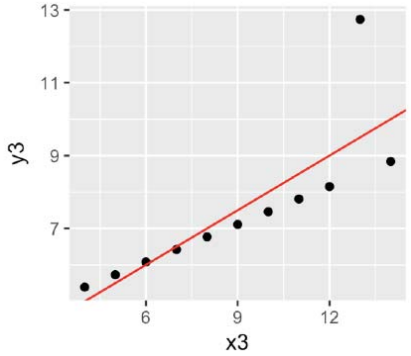
Add transparency or outline shape



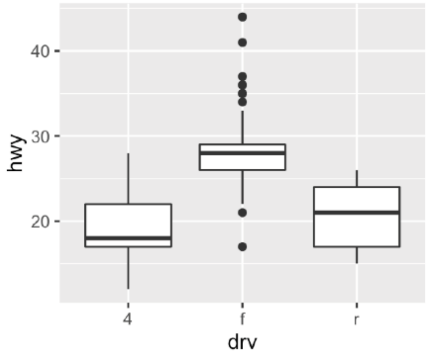
Add Jitter



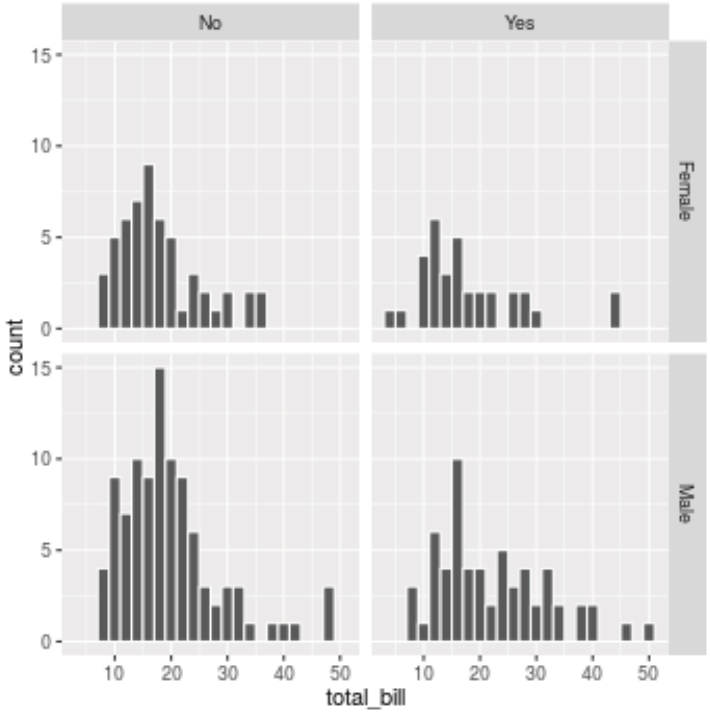
Add additional information



Summarize the data



Split the data & use small multiples

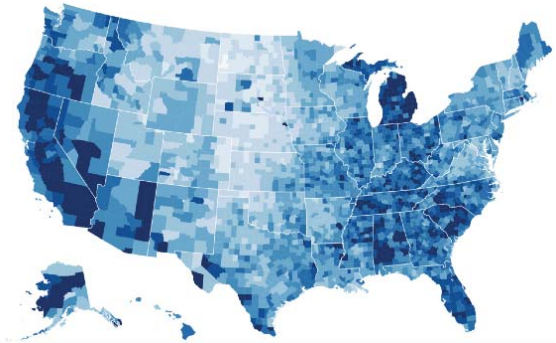


X

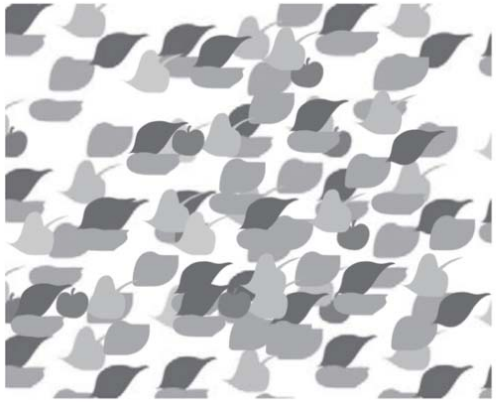
Uses of Color in visualizations

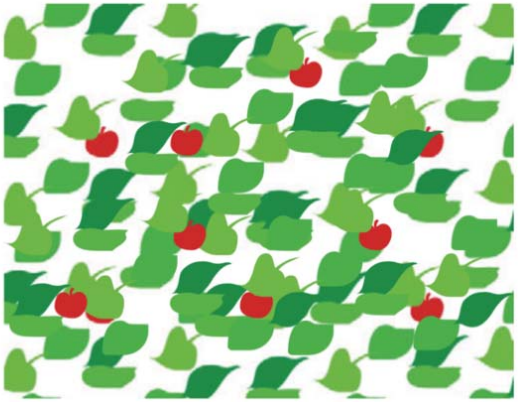
Common uses of colors in visualizations [Tufte]

- Label (identify, highlight, group)
- Measure, represent or imitate reality



- Enliven or decorate

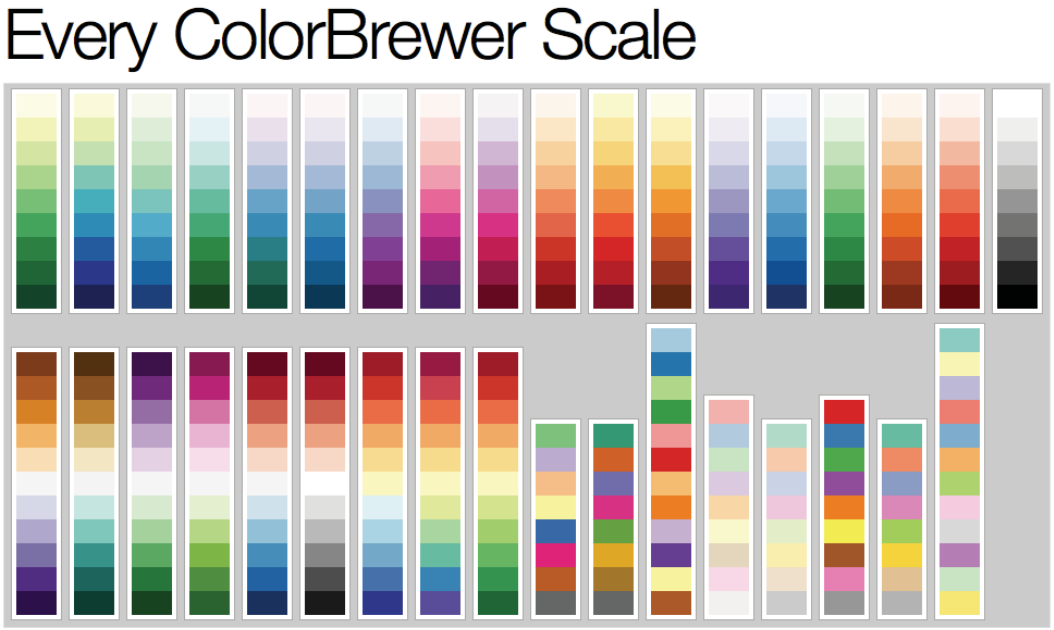




Distinctness & unique hues

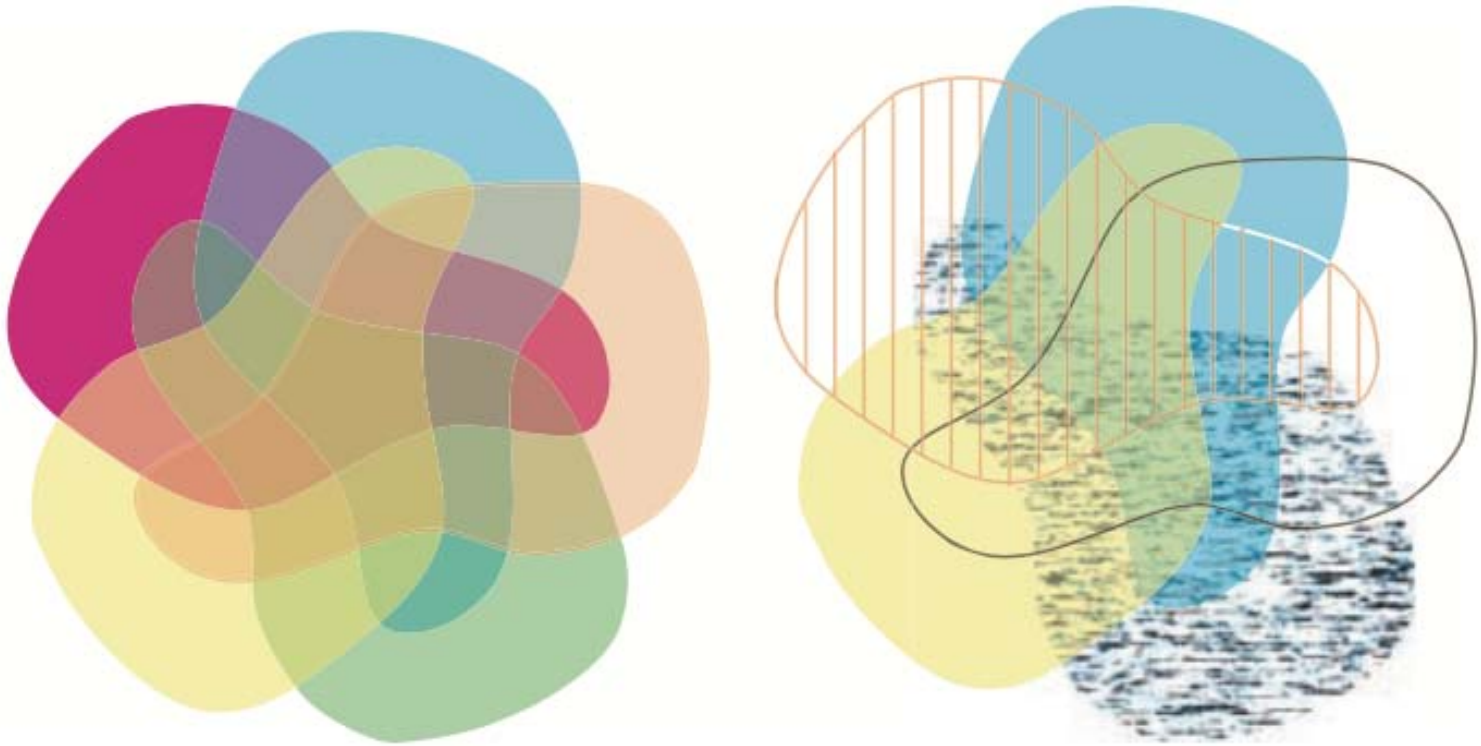
- Limit the number of colors used as keys to 5-7

- Use well established color sequences, e.g., colorbrewer: http://colorbrewer2.org



Maximize selective attentional tuning

- Use **features** as different as possible to superpose information, e.g., in map applications!
- Limit the number of colors and patterns to 4-5



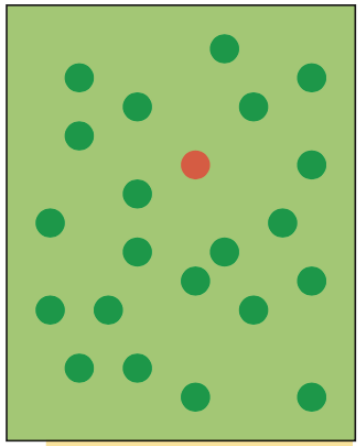
Ware, Colin. "Visual queries: The foundation of visual thinking." *Knowledge and information visualization*. Springer Berlin Heidelberg, 2005.

Distinctness & unique hues

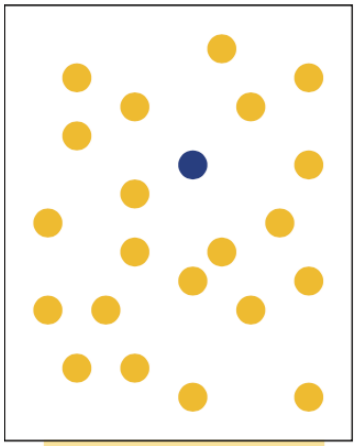
Be mindful of color blind people, e.g., see http://colororacle.org

Color	Color name	RGB (1–255)	CMYK (%)	P	D
	Black	0, 0, 0	0, 0, 0, 100		
	Orange	230, 159, 0	0, 50, 100, 0		
	Sky blue	86, 180, 233	80, 0, 0, 0		
	Bluish green	0, 158, 115	97, 0, 75, 0		
	Yellow	240, 228, 66	10, 5, 90, 0		
	Blue	0, 114, 178	100, 50, 0, 0		
	Vermillion	213, 94, 0	0, 80, 100, 0		
	Reddish purple	204, 121, 167	10, 70, 0, 0		

Use colors in popup effect wisely to support visual searches

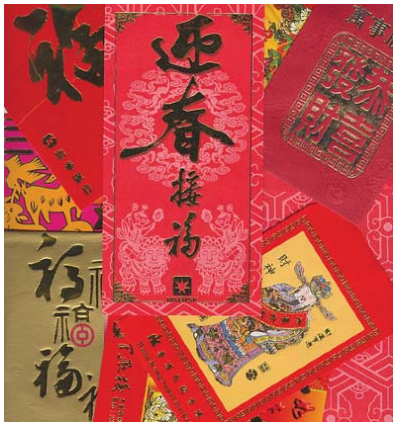


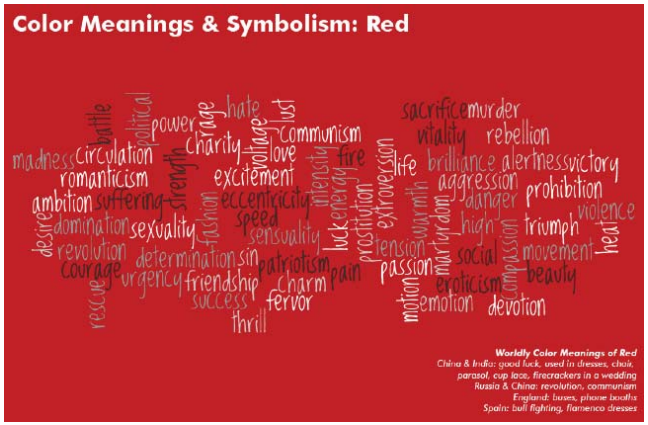
If non-target symbols are similar to the background, they are easy to exclude from the visual search.



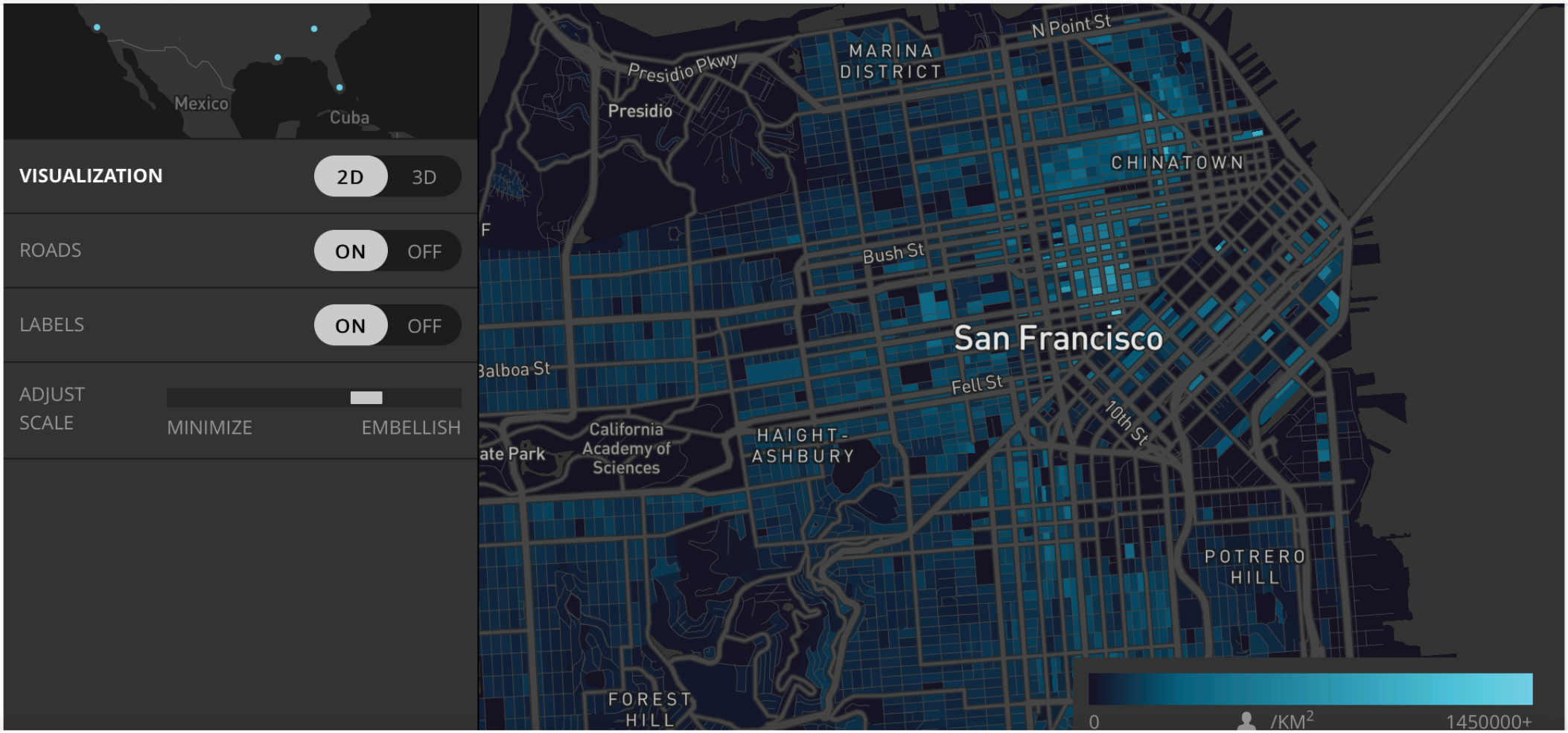
A dark target on a light background with light non-target symbols can be as effective as the reverse.

Observe cultural conventions





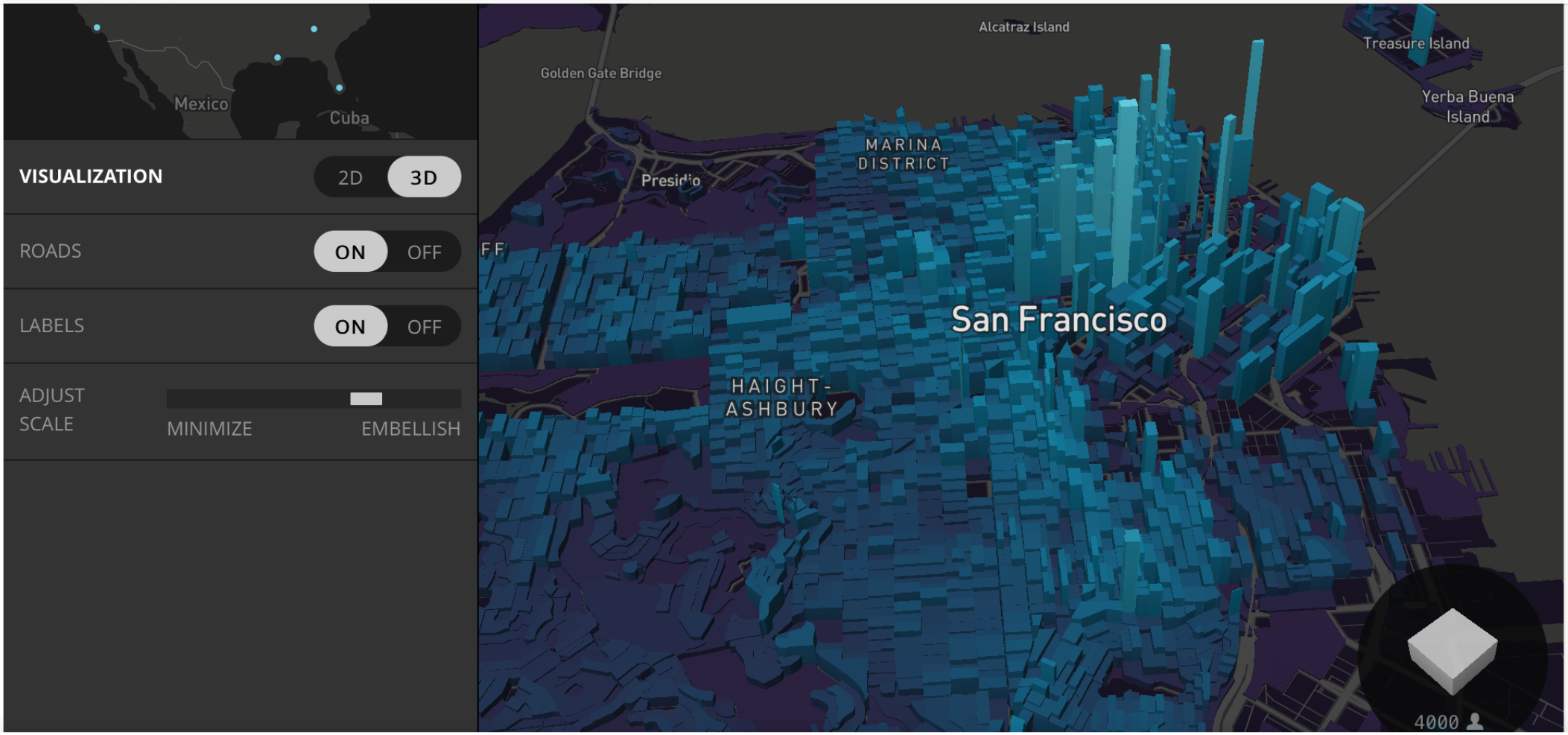
Use of 3D in Maps



https://www.mapbox.com/blog/population-inspector

X

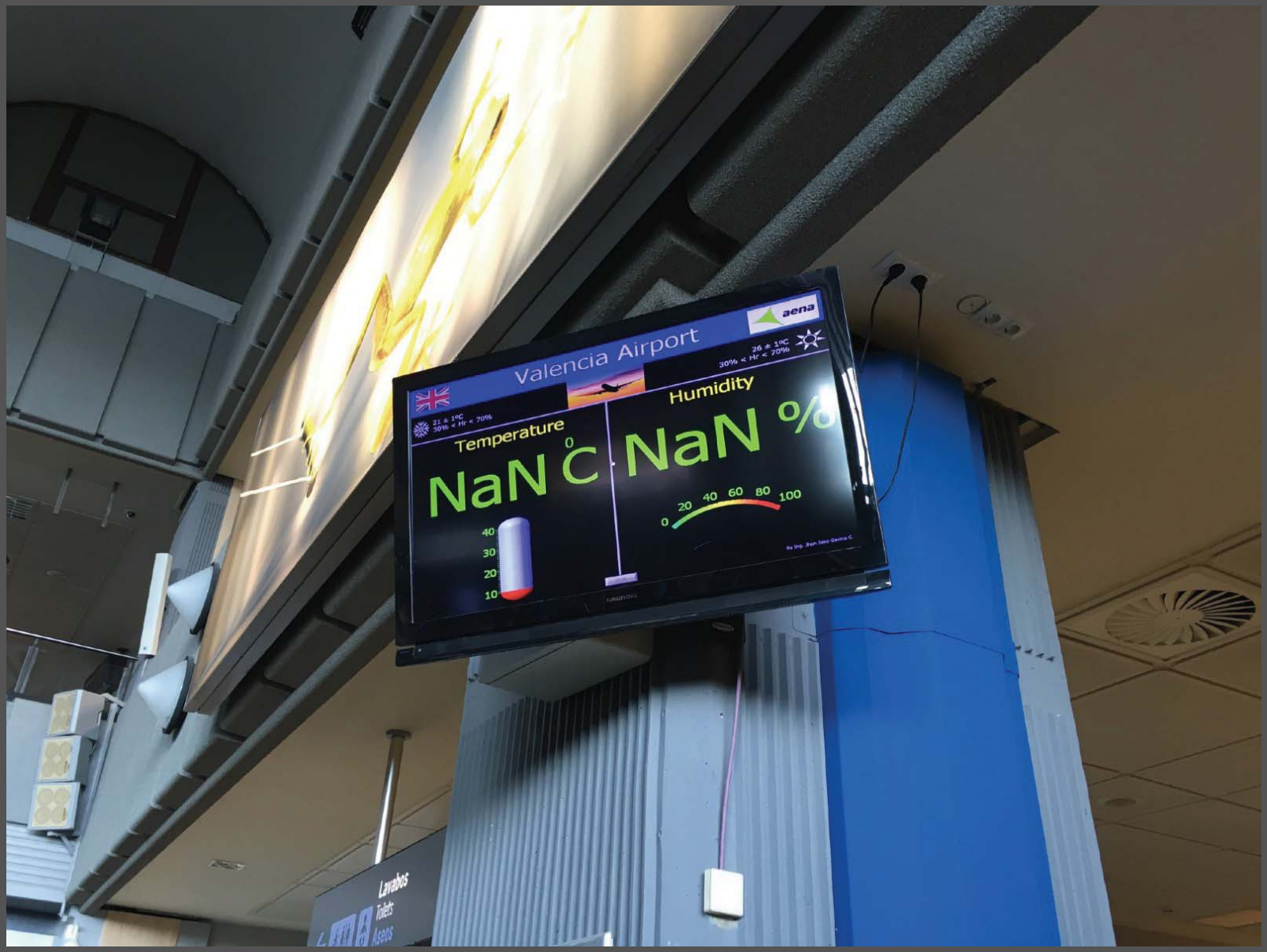
Use of 3D in Maps



https://www.mapbox.com/blog/population-inspector

X

What NOT to do


Valencia Airport
aena
Temperature
NaN C
Humidity
NaN %
Lavabos
Toilets
Aseos

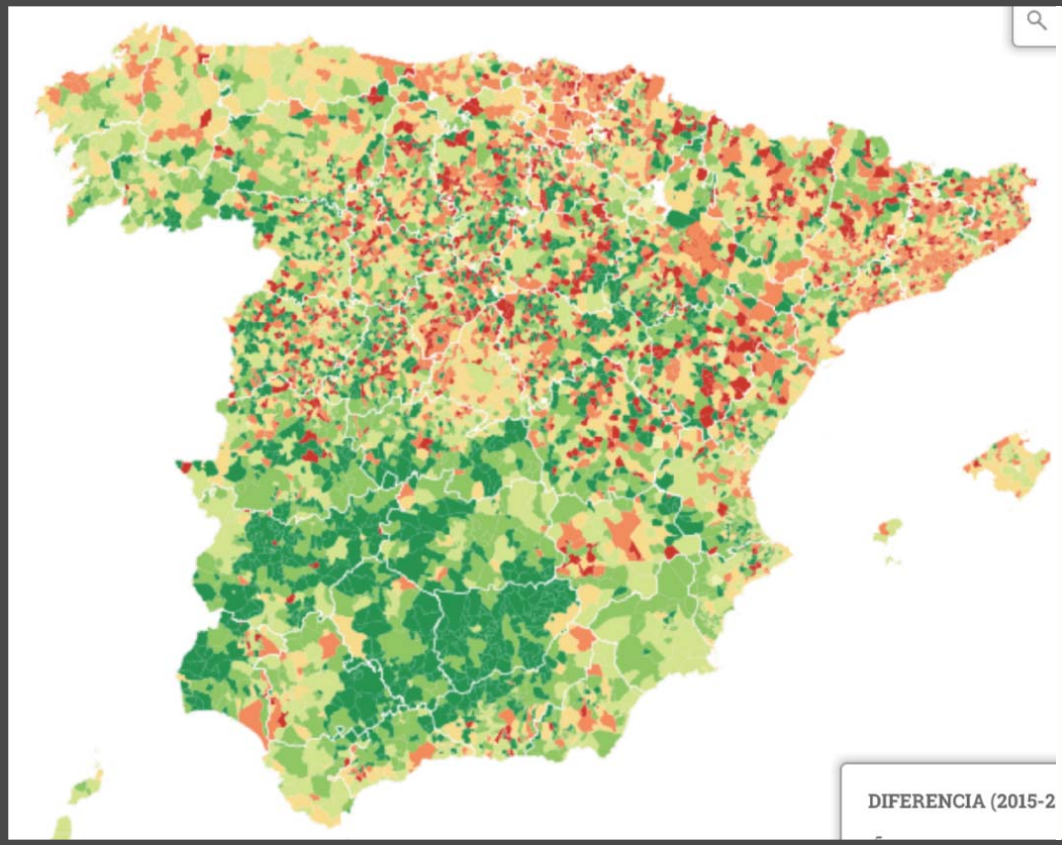
X

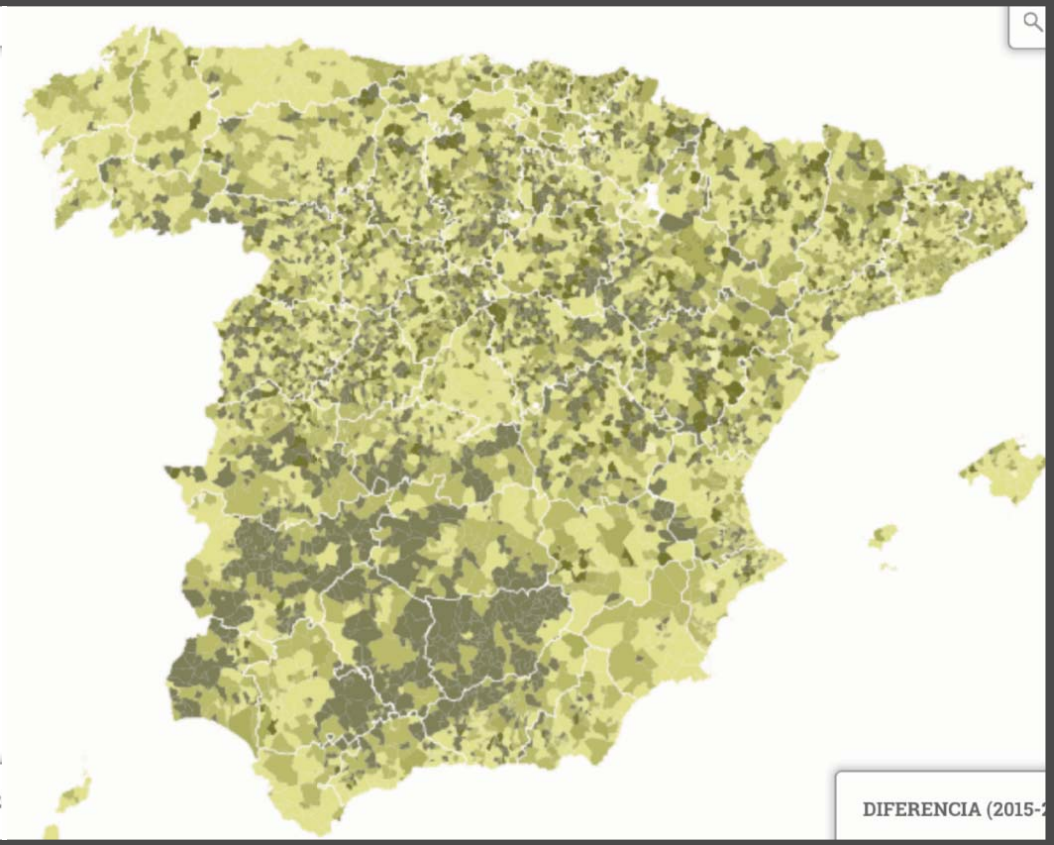

+6%
respecto al
año anterior
Lavabos
Toilets
Aseos

insufficient luminance contrast; small font size not helping

X

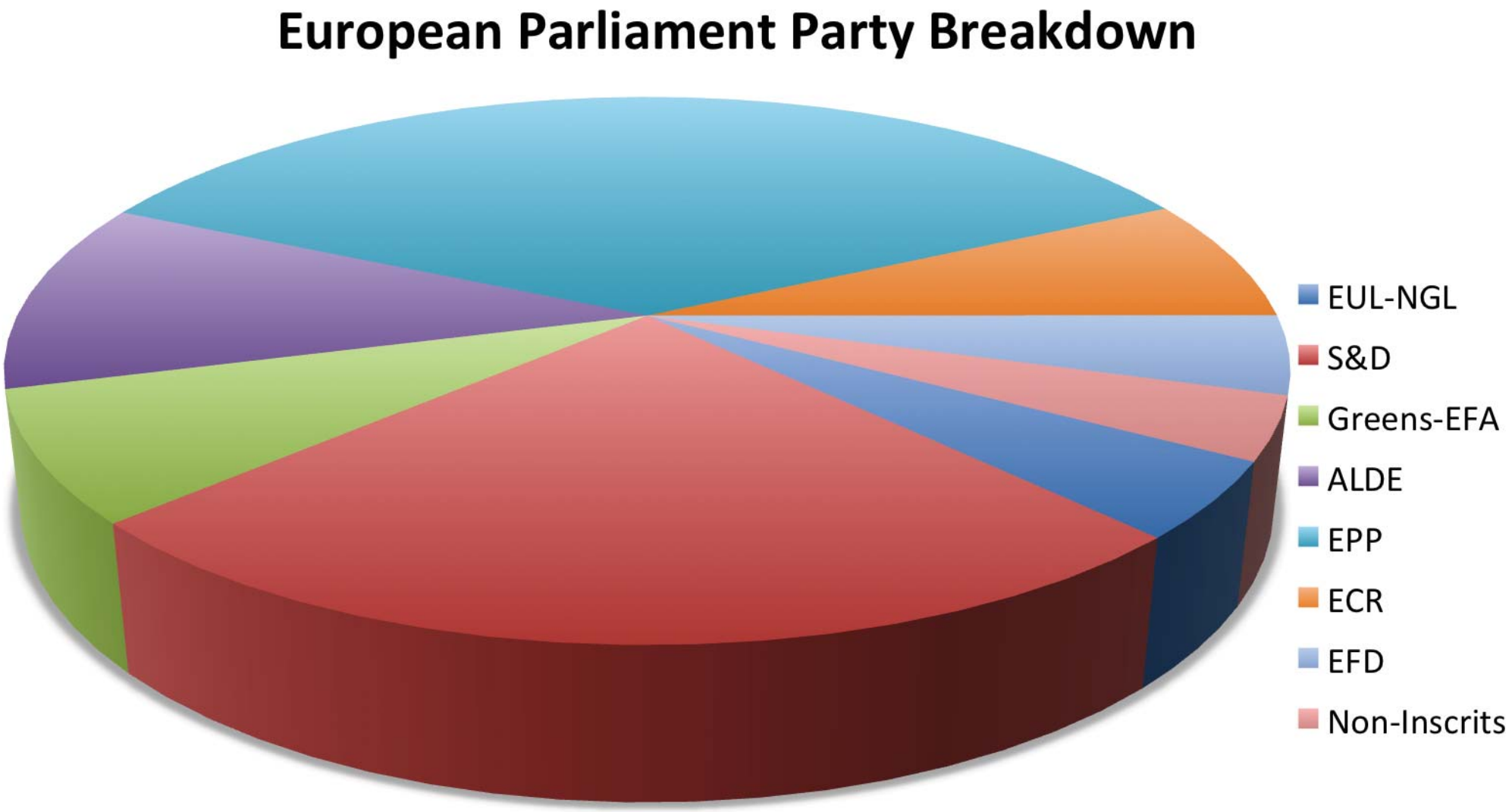
Not colorblind safe





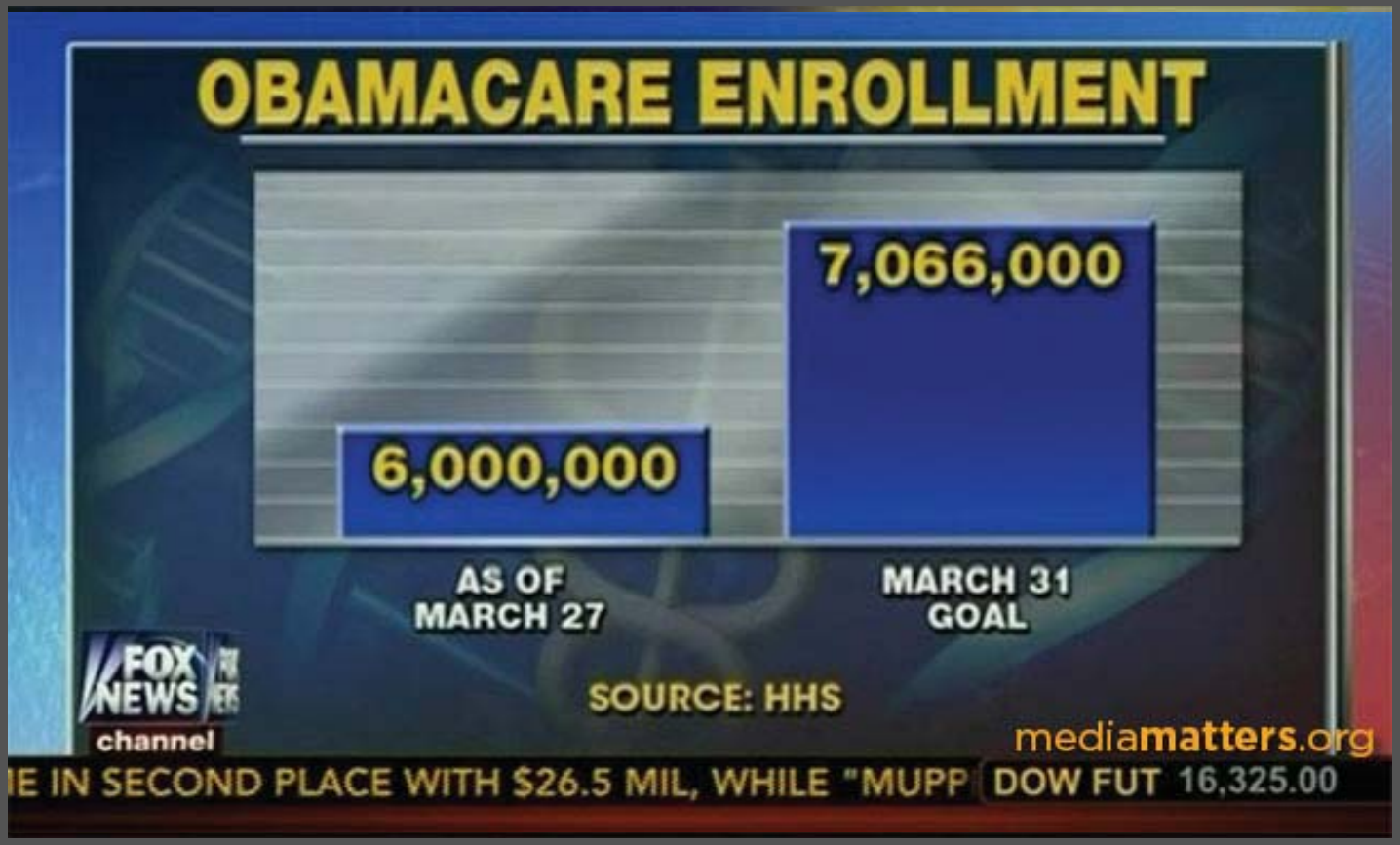
X

Pie chart difficult for comparison, too many colors



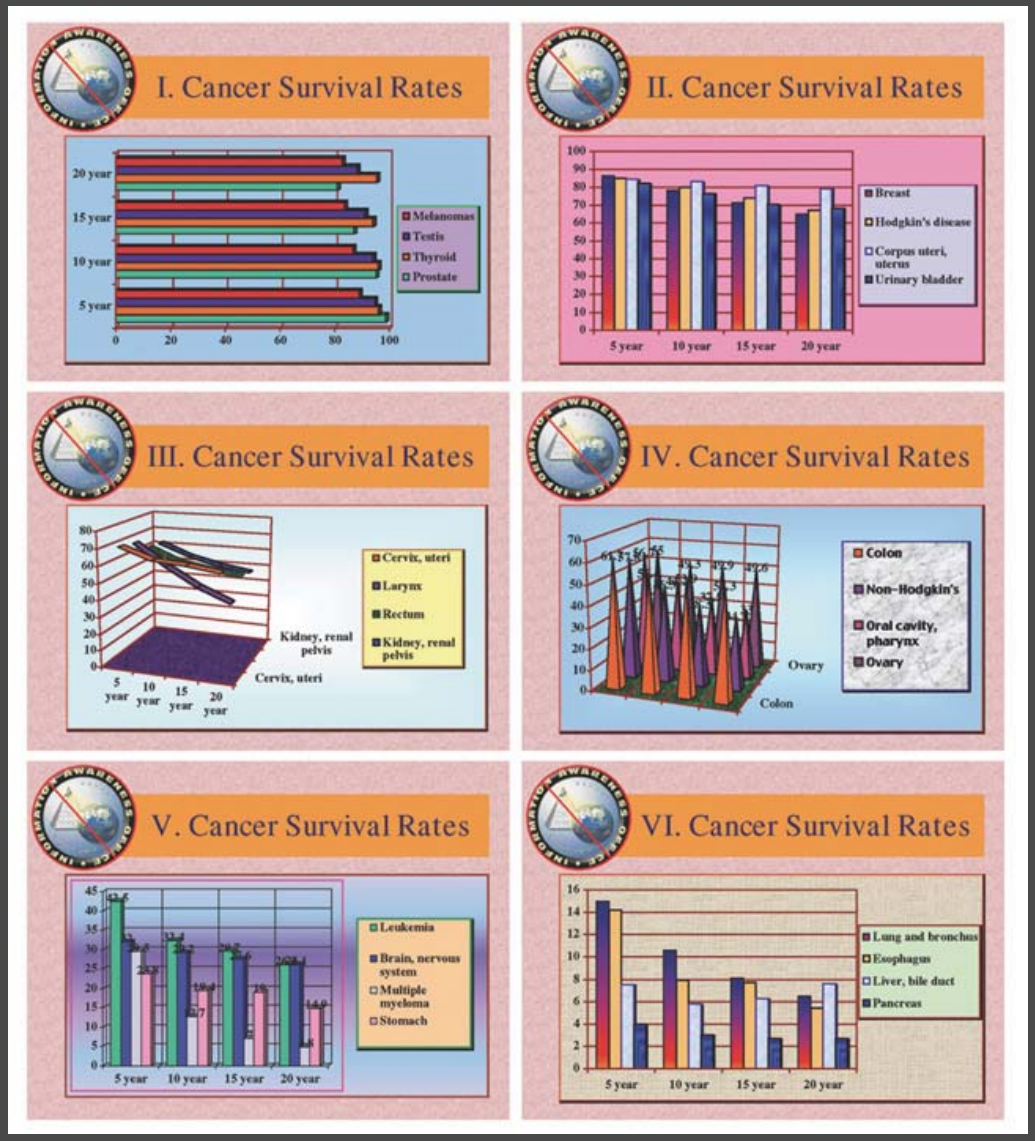
http://www.businessinsider.com/pie-charts-are-the-worst-2013-6

Intentionally misleading comparison



Fox News March 2014

3D & occlusions, distracting colors and effects, bad choices of colors...



http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0000Jr

X