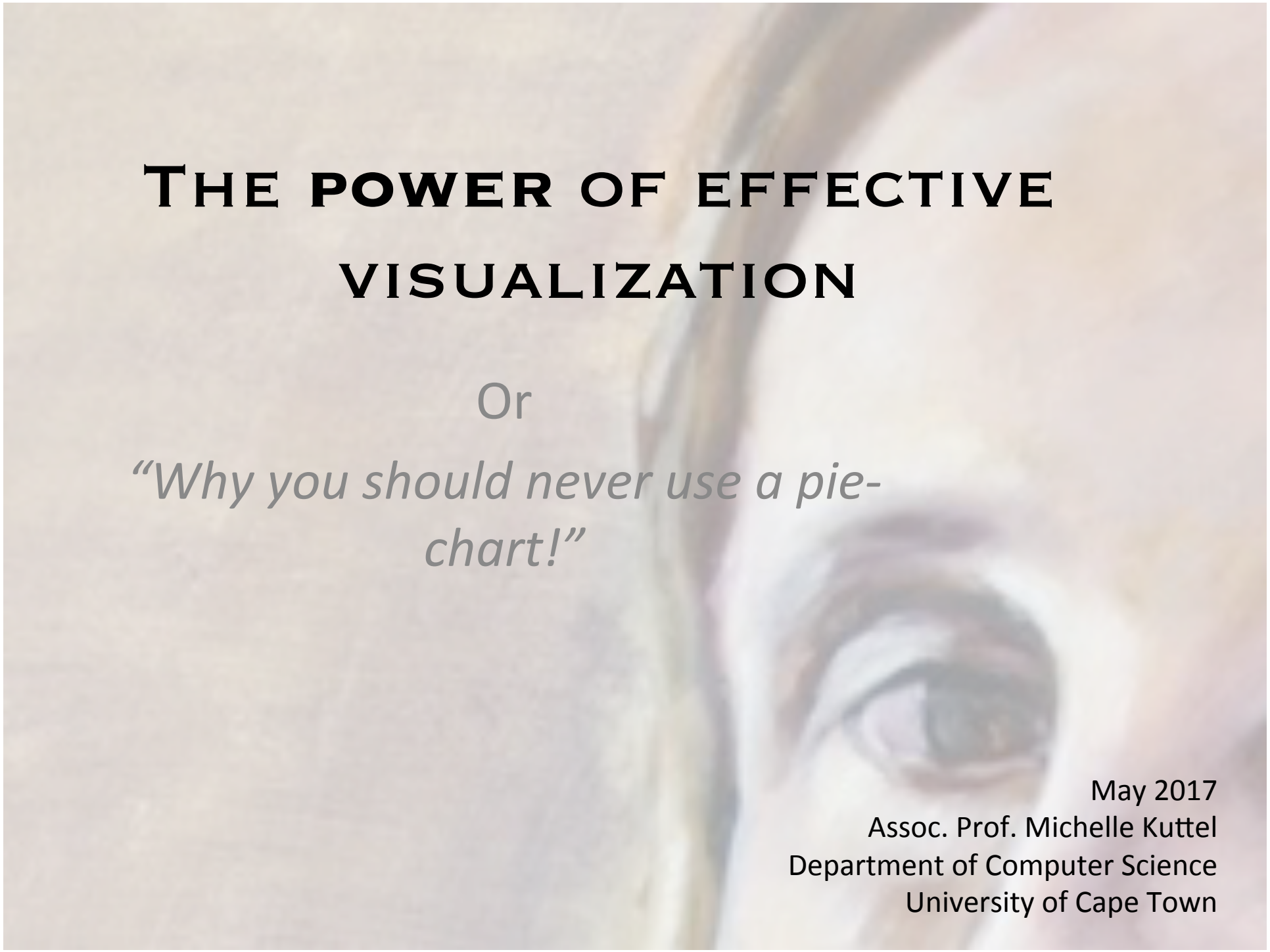


The power of effective visualization

Or

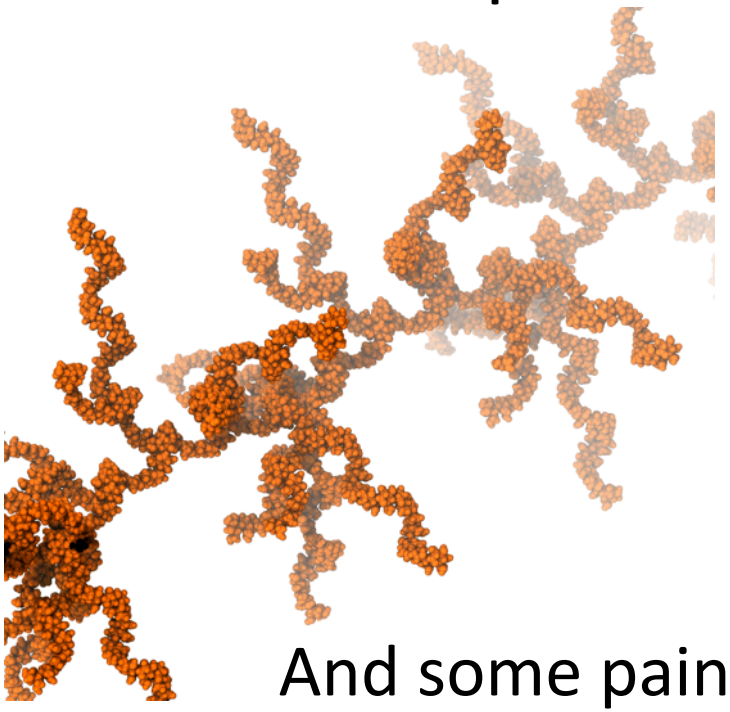
"Why you should never use a pie-chart!"

May 2017

Assoc. Prof. Michelle Kuttel
Department of Computer Science
University of Cape Town

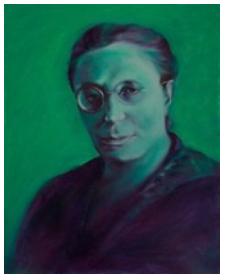
About me

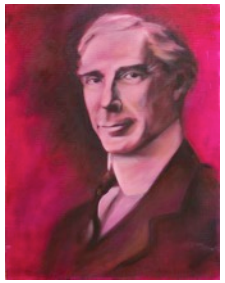
Dept. Computer Science, UCT



- Computational Science
  - Chemistry
  - Astronomy
- High Performance Computing
- **Visualization**
- Software for science

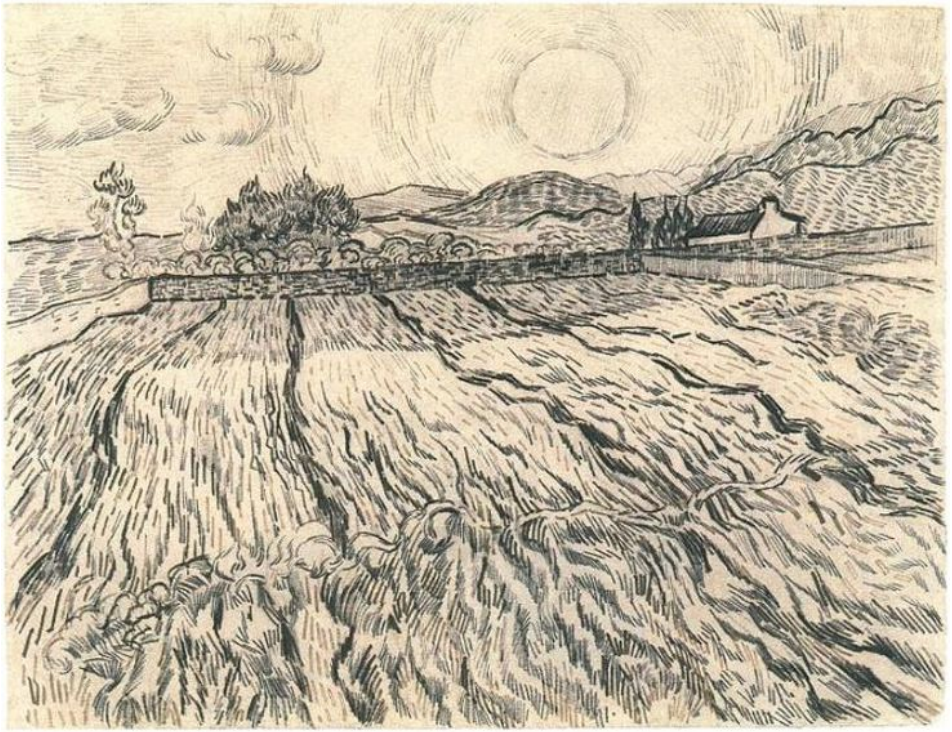
And some painting…

"Science and art have in common *intense seeing,* the wide-eyed observing that generates empirical information."

Edward Tufte in *Beautiful Evidence*



Visualizations for thinking and interrogating

Complex thinking typically requires interaction with **cognitive tools:**

- pencils and paper
- calculators
- **Software**

Visualizations

- graphical representation of data/concepts
- increasingly important for interrogating complex data
  - Interactive visualizations
- **Why are they important and useful?**

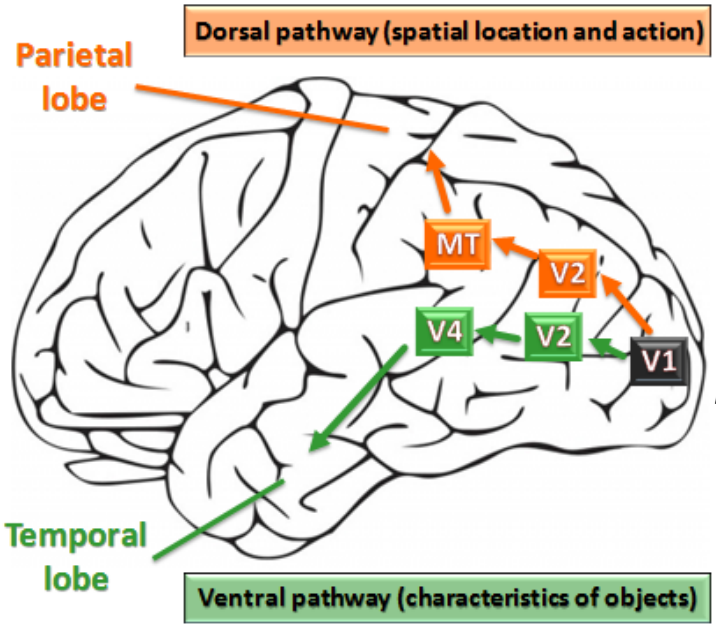
The aim of VIS

- VIS allows people to analyze data when they don't know exactly in advance which questions they need to ask of the data.

- VIS leverages the powerful pattern detection properties of the human visual system.

active vision

- Almost **half** our brain is devoted to vision
- graphic designs are powerful **cognitive tools**



V1 - primary visual cortex

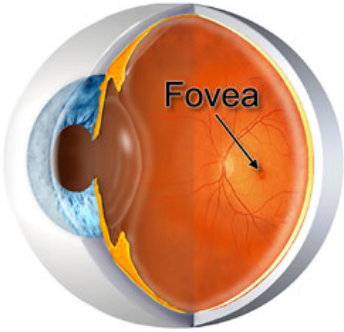
  - diagrams, maps, technical illustrations all help us to understand and solve problems
    - **visual thinking**
  - enhances and amplifies our mental abilities
    - Allows us to surpass the limits of our internal cognition and memory – external memory.

Image from: Lyes Bachatene, Vishal Bharmauria and Stéphane Molotchnikoff (2012). Adaptation and Neuronal Network in Visual Cortex, Visual Cortex - Current Status and Perspectives,.
http://www.intechopen.com/books/visual-cortex-current-status-and-perspectives/adaptation-and-neuronal-network-in-visual-cortex

The process of seeing

Eyes are like digital cameras with a range of light-seeing elements: rods and cones

- brain pixels vary enormously in size over the visual field:
  - tiny ones at the centre (fovea)
  - large ones at the periphery
    - at the end of the visual field, vision is terrible
      - can resolve objects about the size of a human hea



- Non-uniformity means that half our brain power is devoted to viewing less than 5% of our visual world

The advantages of vision

- Our visual system processes significant amounts of visual information in *parallel*
  - Preconscious
    - E.g visual **pop-out**
  - Sound does not work like this – sequential

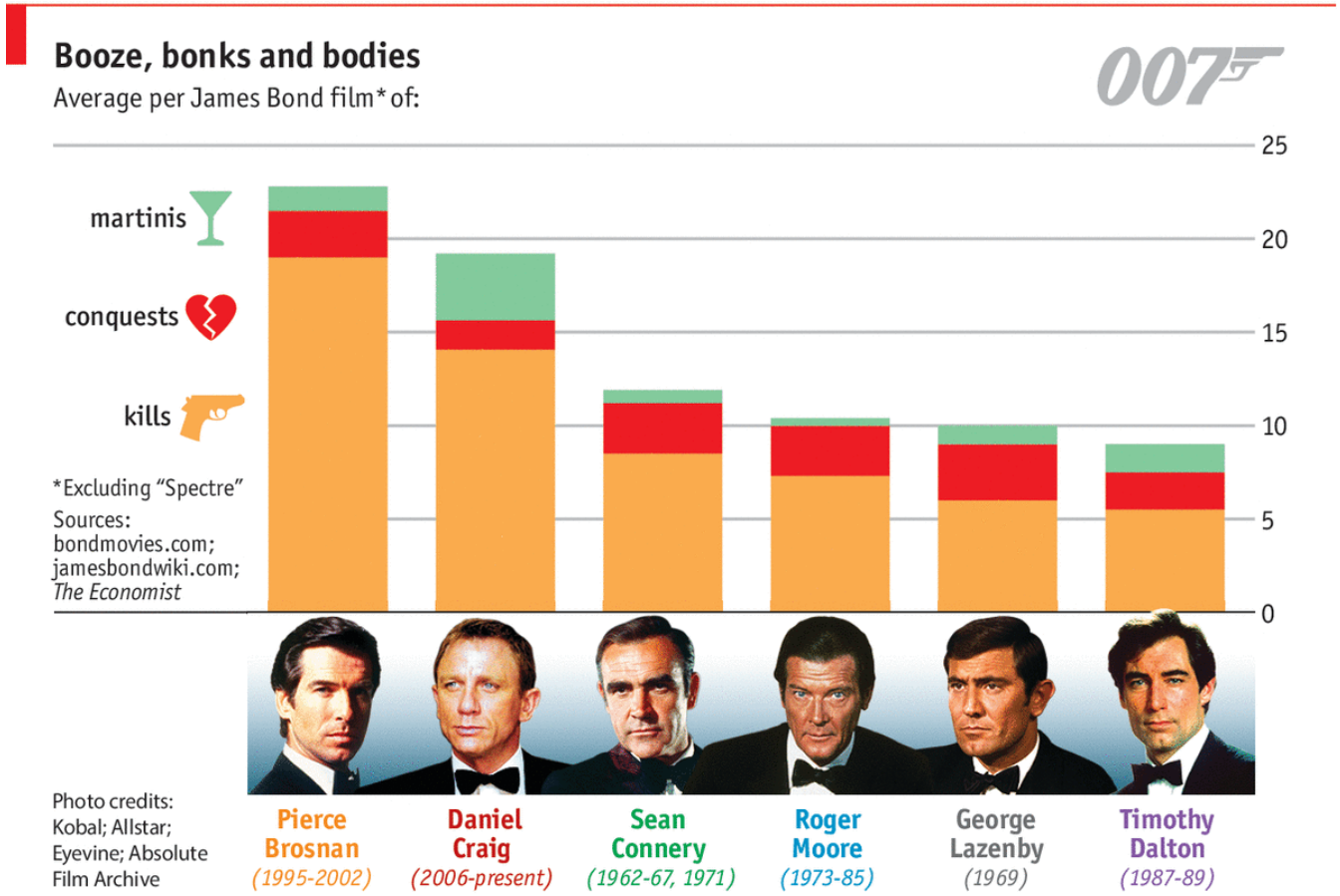
Visualizations

Good data visualizations allow huge quantities of information to be processed rapidly.

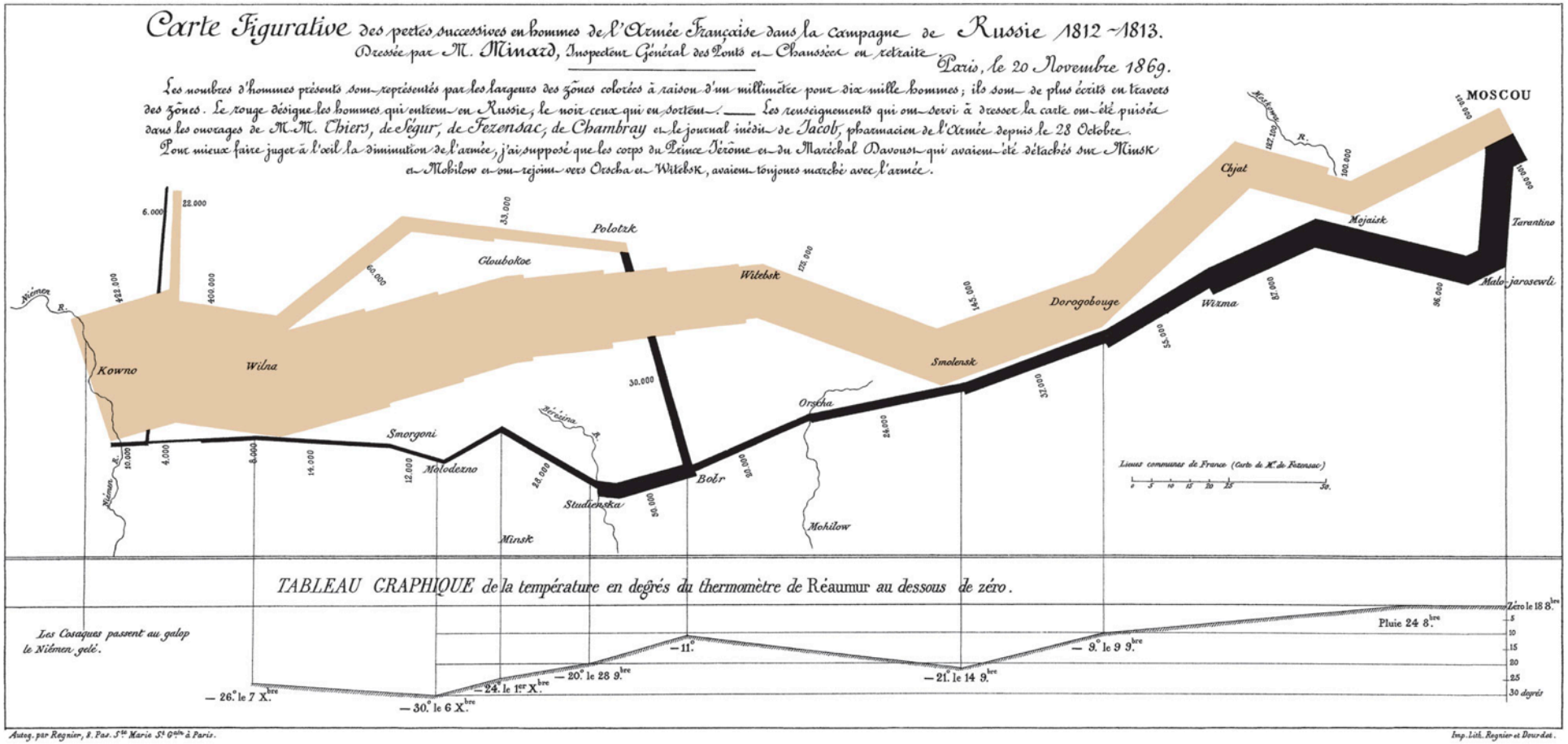
A good visualization allows us to represent, and therefore grasp, an entire process or system – we can focus on a component while still having access to the global view.

What is a “good visualization”?

Visualizations



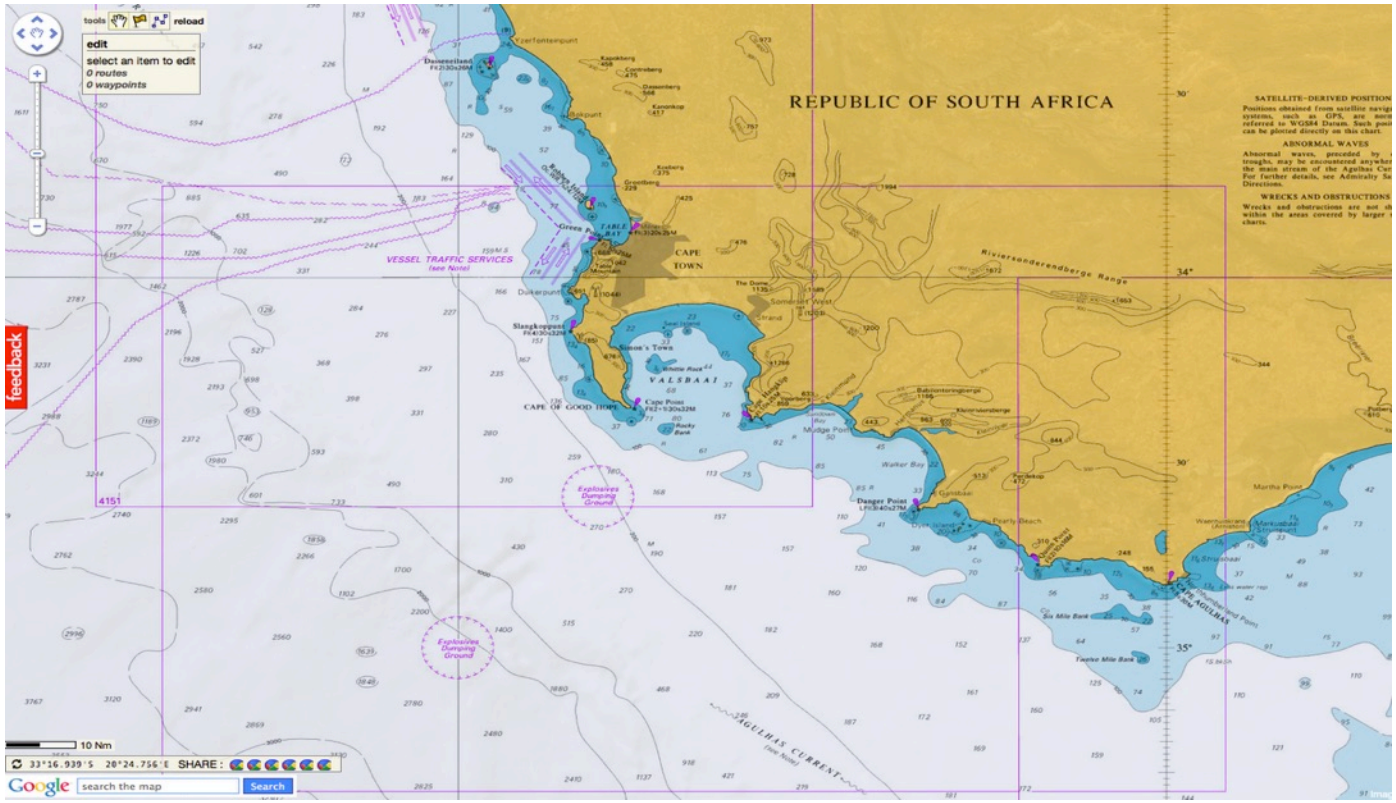
Famous visualization: Minard's Map of Napoleon's march to Moscow



Created in 1861. Described by E. Tufte* as "*Probably the best statistical graphic ever drawn, this map by Charles Joseph Minard portrays the losses suffered by Napoleon's army in the Russian campaign of 1812. Beginning at the Polish-Russian border, the thick band shows the size of the army at each position. The path of Napoleon's retreat from Moscow in the bitterly cold winter is depicted by the dark lower band, which is tied to temperature and time scales.* "

** Image and text from Graphics Press, http://www.edwardtufte.com/tufte/posters*

Maps and Charts



Cartographers have hundreds of years of experience in articulating the difference between the abstraction of a map and the terrain it represents.

How do we design **good** visualizations?

Goal: support user tasks

Focus is on:

- effectiveness
- CORRECTNESS
- *accuracy*
- truth

Most designs are ineffective

- Design goal to SATISFY not OPTIMIZE
- Need to explicitly generate multiple designs in parallel
  - Explore “solution space”

Four levels of design

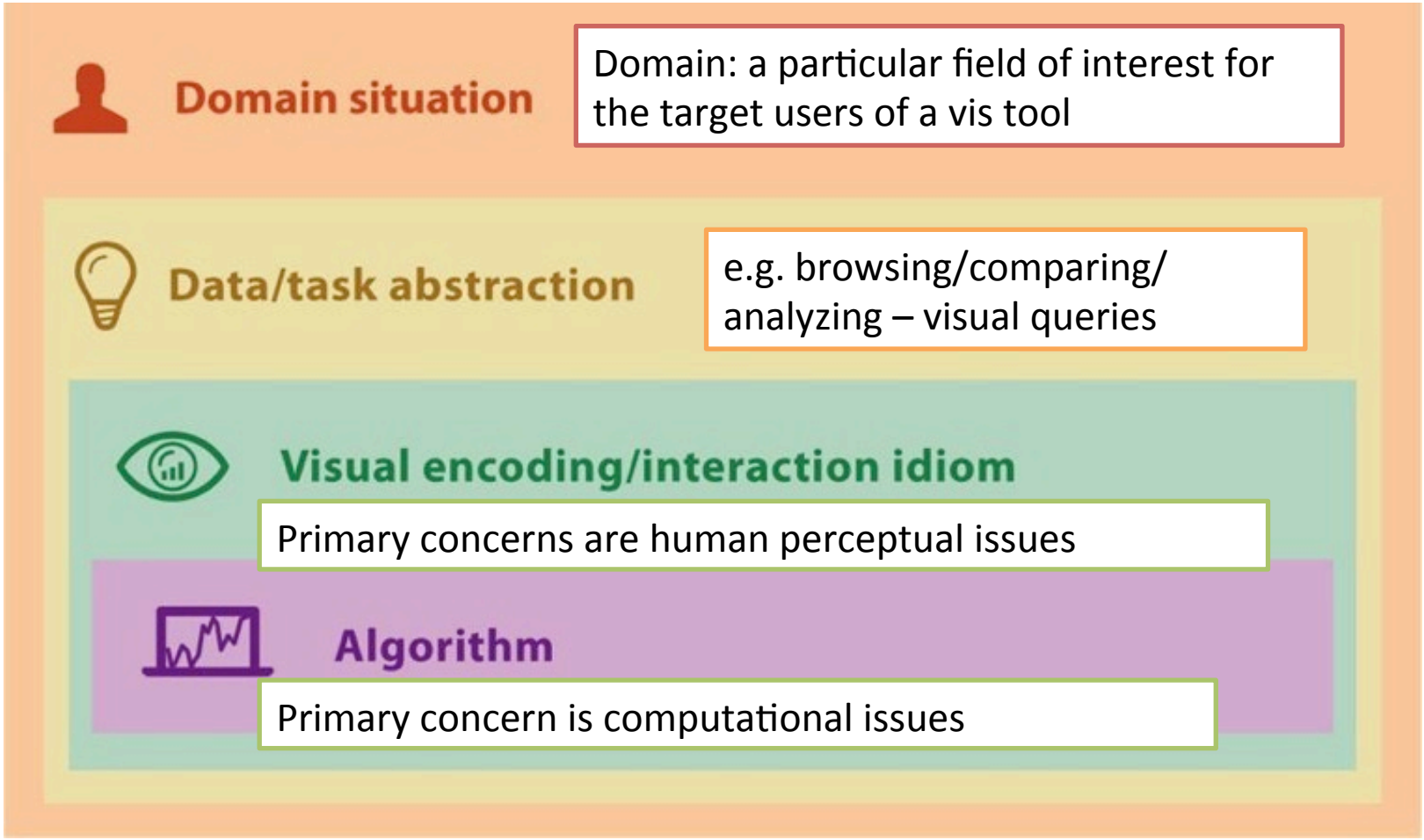


Table from “Visualization Analysis and Design” by Tamara Munzner

Data

Key differential is **spatial** versus **abstract** (non-spatial) data.

Historically this splits the two areas of VIS :

- Scientific visualization
- Information visualization

Characterizing Data

Table from “Visualization Analysis and Design” by Tamara Munzner

Perception and visual attention

Our minds don't work as well as we think they do….

..at any given instant, we apprehend only a tiny fraction of the information stored in our surroundings.

- "The world is its own memory."

We are not immediately conscious of the world:

- conscious of the "field of information" to which we have "rapid access".

The monkey business illusion

by Christopher Chabris and Daniel Simons

The act of perception

- perception is driven by 2 processes:
  - **bottom-up**:
  - driven by external stimulus

  retinal image -> features -> patterns -> objects

  We detect motion, edges of shapes, colours, contours, contrasts through bottom-up processes without conscious awareness

The act of perception

top-down or attention

- driven by need to accomplish some goal, prior knowledge, expectations
  - biased in favour of signals we are looking for
- only get information we need when we need it
  - sequence of rapid eye movements to locate important objects

- Brain functions as a kind of distributed processor

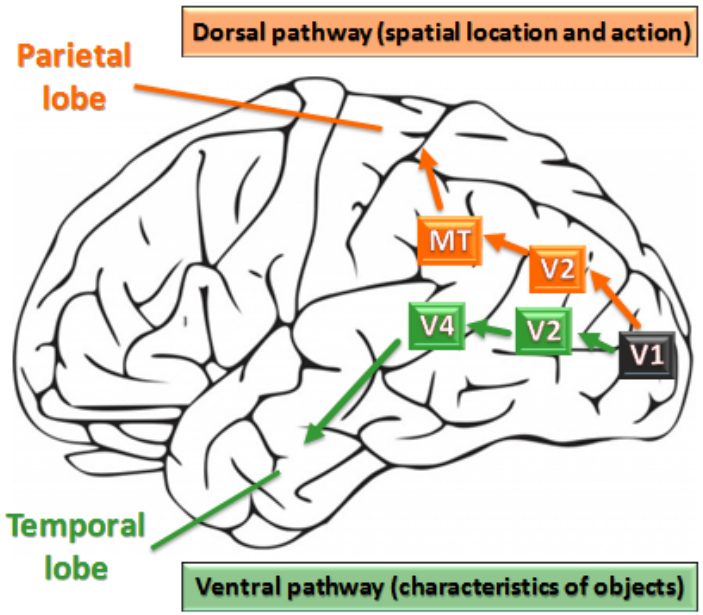
Facilitating visual queries

- Visual attention works like a **spotlight**
- In a graphic or interface, you want to ensure that all **visual queries** can be rapidly and effectively served
- How do we design graphical symbols that can be rapidly located?
- The most important and frequent queries should be supported by the most **visually distinct objects**

What makes a small thing easy to spot?

How humans do low-level feature analysis

- Early stages of visual processing occur in the primary visual cortex:
  - V1 : general scanning
  - V2 : stereo vision
  - V3 : depth and distance
  - V4 : colour
  - V5 (or MT) : motion
  - V6 : objective position of object
- **"Where?"** path: location of objects
  - V1-V2-V3-V5-V6
- **"What?"** path: identification of objects
  - V1-V2-V4



very nice book on the brain ->

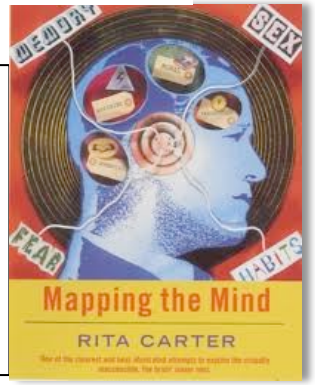


Image from: Lyes Bachatene, Vishal Bharmauria and Stéphane Molotchnikoff (2012). Adaptation and Neuronal Network in Visual Cortex, Visual Cortex - Current Status and Perspectives, Prof. Stephane Molotchnikoff (Ed.),

How to find what you are looking for: biased competition

- If you are looking for a particular colour/ orientation/size of an object, the visual system highlights these
  - e.g. if picking strawberries, all red sensitive cells will "shout louder"

- some things pop out much more easily than others

Pop-out

- Also called **preattentive processing** or **tunable detection.**
- The time to spot an object with **pop-out** does not depend on the number of distractor objects.

What makes objects **pop out**?

- Some kinds of shapes have properties to which our eye-movement programming system is sensitive – they **pop out**
  - can be seen in a single eye movement: at-a-glance
  - processing takes less than a tenth of a second
    - compare with 1-5 seconds for a search
- easiest when single object differs in one feature from all the objects around it
  - to do with degree of contrast to the environment of the object

What makes objects pop out?

- Simplest feature **channels** that lead to pop out:

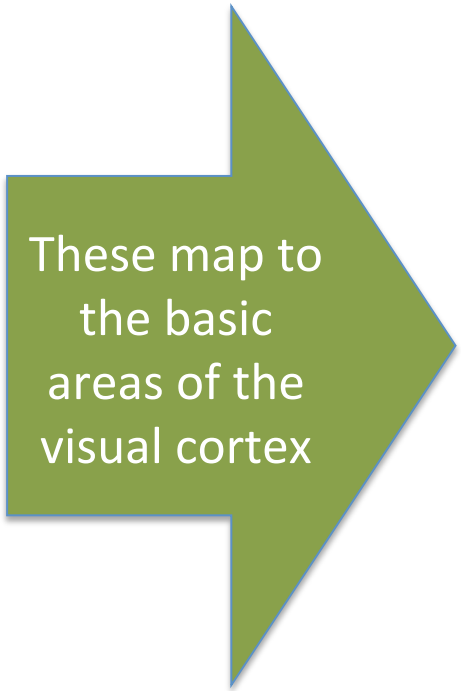
colour

orientation

size

motion

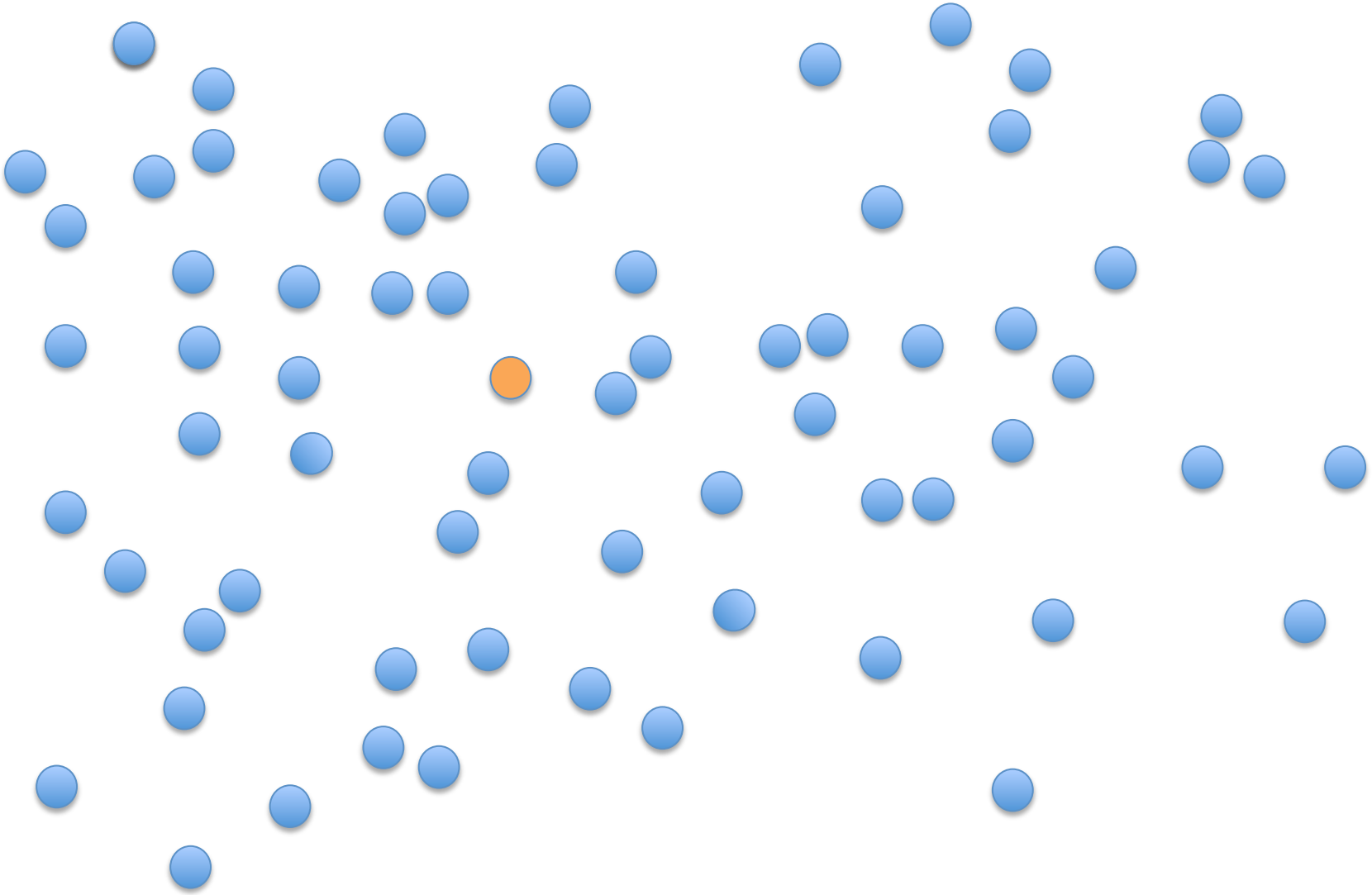
stereoscopic depth



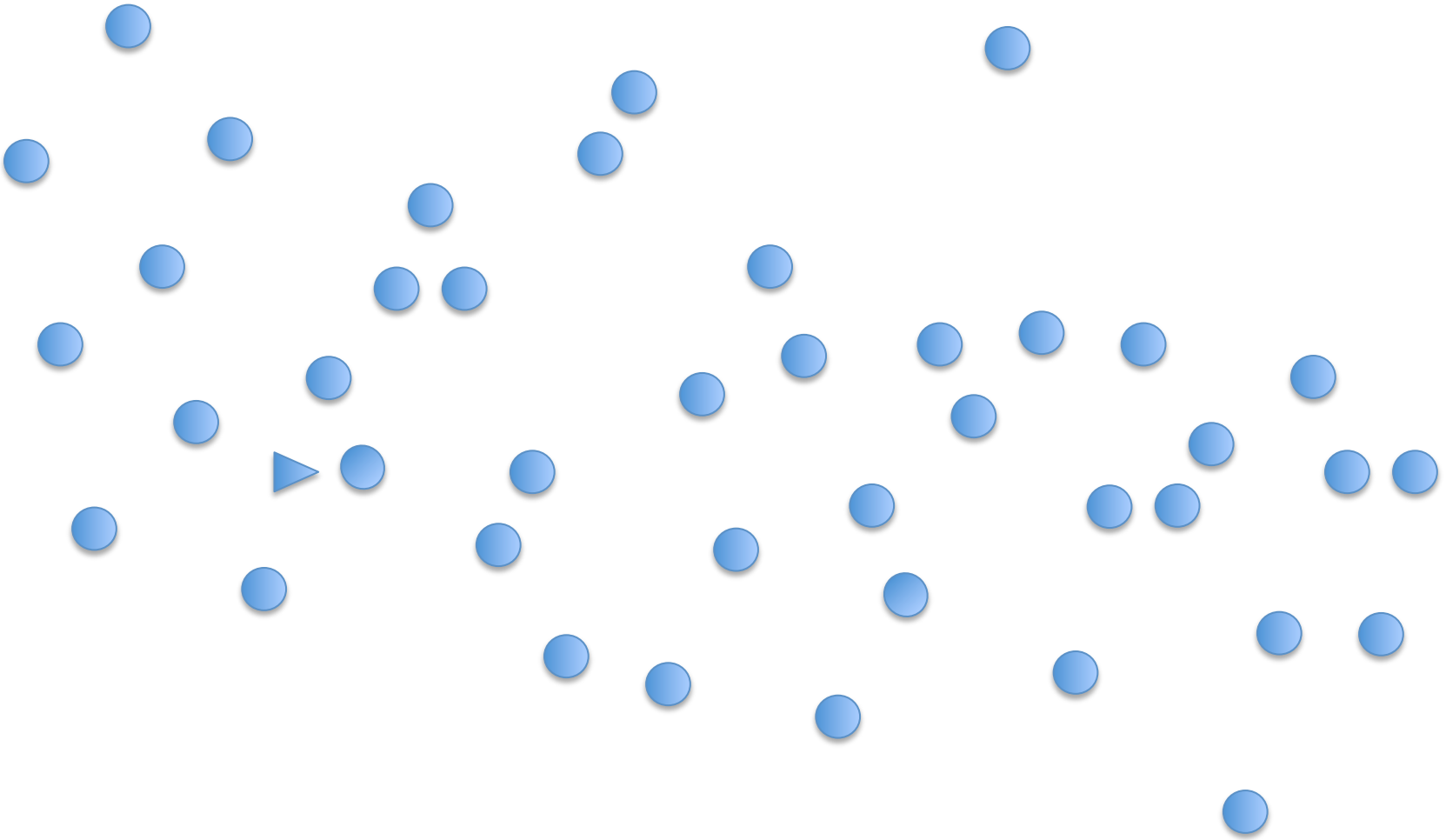
- V1 : general scanning
- V2 : stereo vision
- V3 : depth and distance
- V4 : colour
- V5 (or MT) : motion
- V6 : objective position of object

- Low-level visual system does massive parallel processing on these visual channels

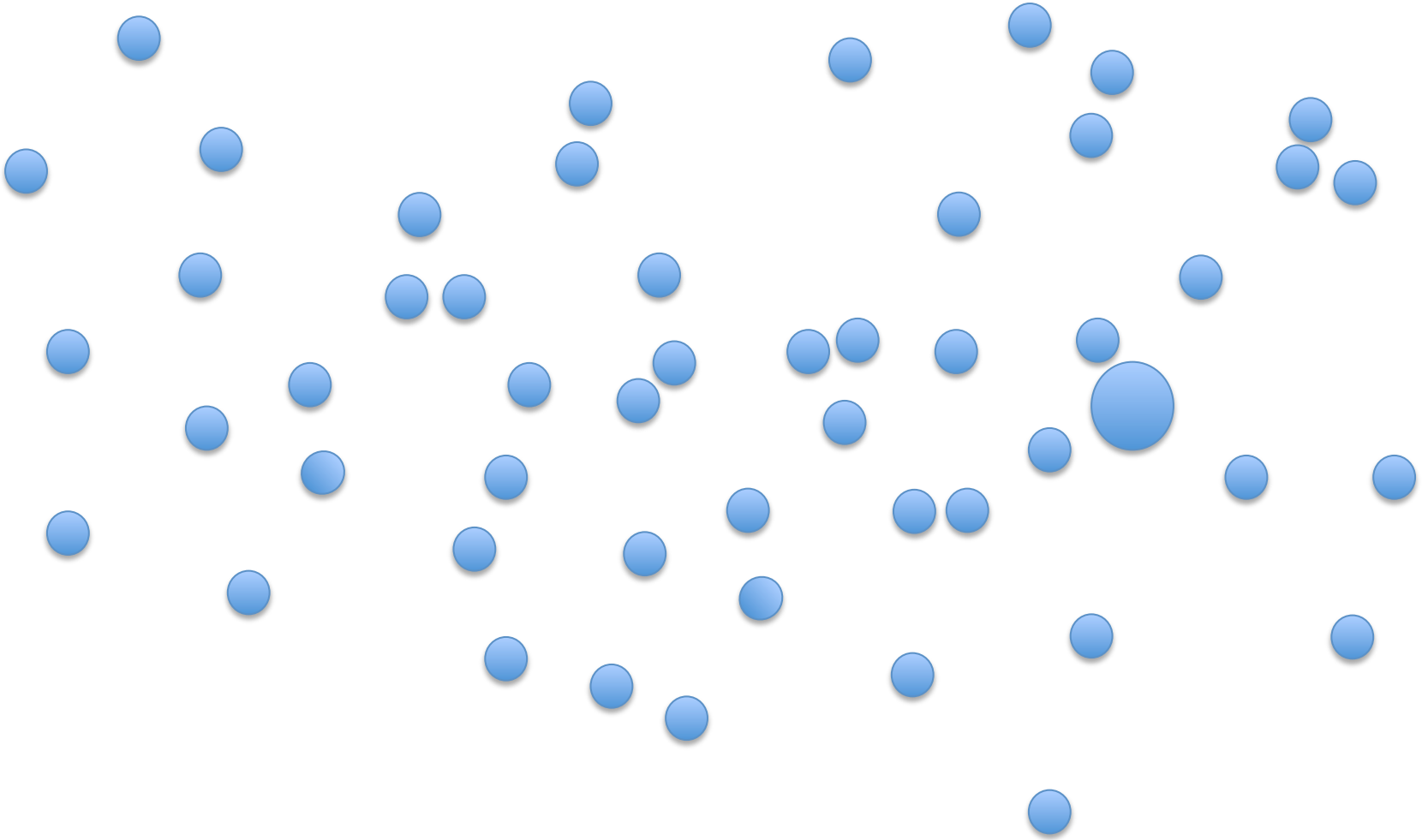
What stands out: colour



What stands out: shape



What stands out: size



Lessons for design

- If you want to make something easy to find, make it ***different*** from its surroundings according to some primary visual channel – size, colour etc.
- A design to support two different kinds of visual query will be most effective if *each query uses a different channel*

(a) Bar charts encode two attributes using a line mark with the vertical spatial position channel for the quantitative attribute, and the horizontal spatial position channel for the categorical attribute.

(b) Scatterplots encode two quantitative attributes using point marks and both vertical and horizontal spatial position.

(c) A third categorical attribute is encoded by adding color to the scatterplot.

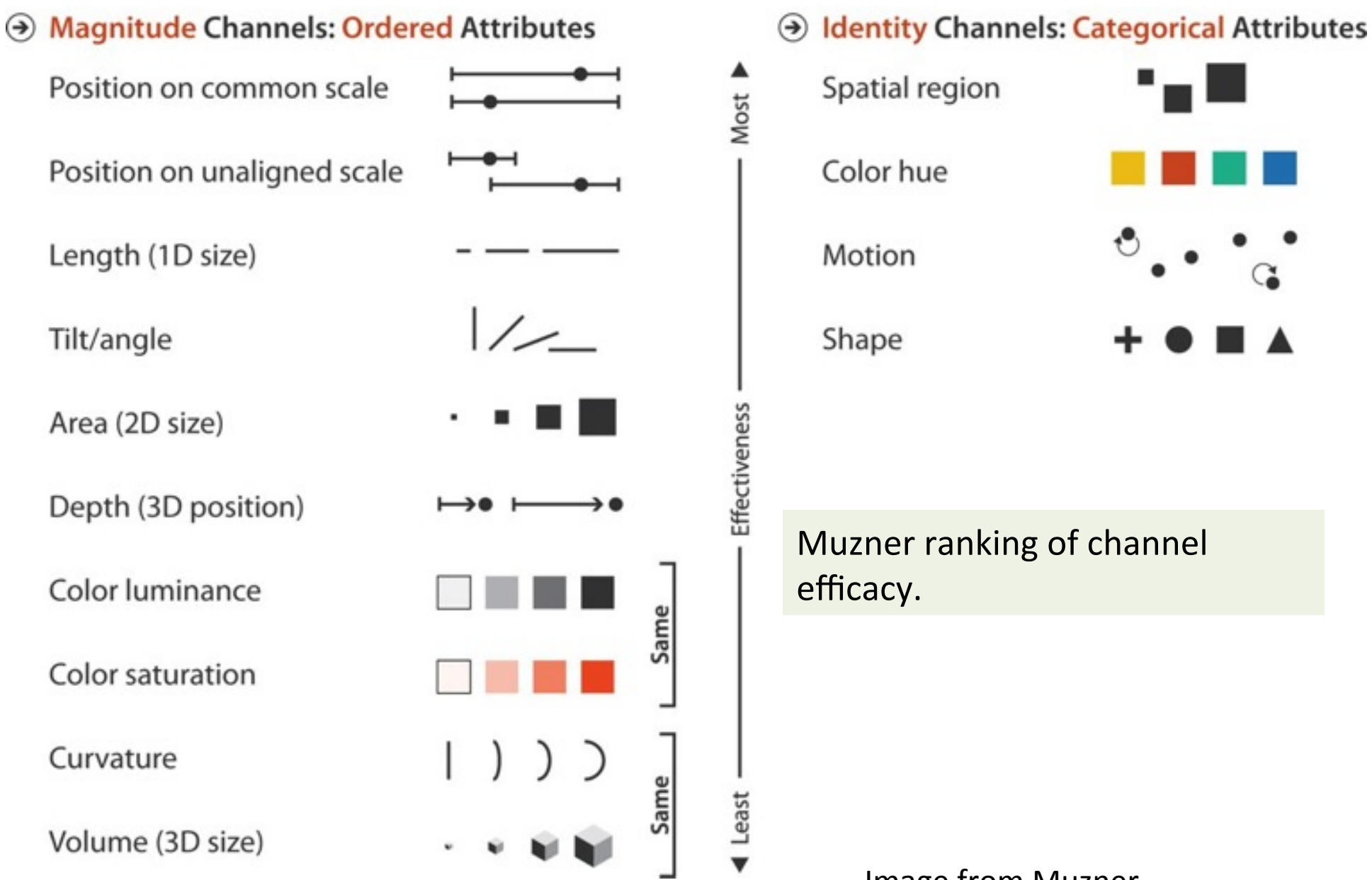
(d) Adding the visual channel of size encodes a fourth quantitative attribute as well.

Image from Muzner.

The Use of channels

- Expressiveness principle
  - The channel should express all of, and only, the information in the data set attributes
    - E.g. don't imply an ordering f it does not exist
- Efficacy principle
  - The most important attributes should be encoded with the most effective channels and vice versa.
  - For example, the choice of which attributes to encode with **position** is central - will dominate user's mental model compared with any other channel.

Channels: Expressiveness Types and Effectiveness Ranks

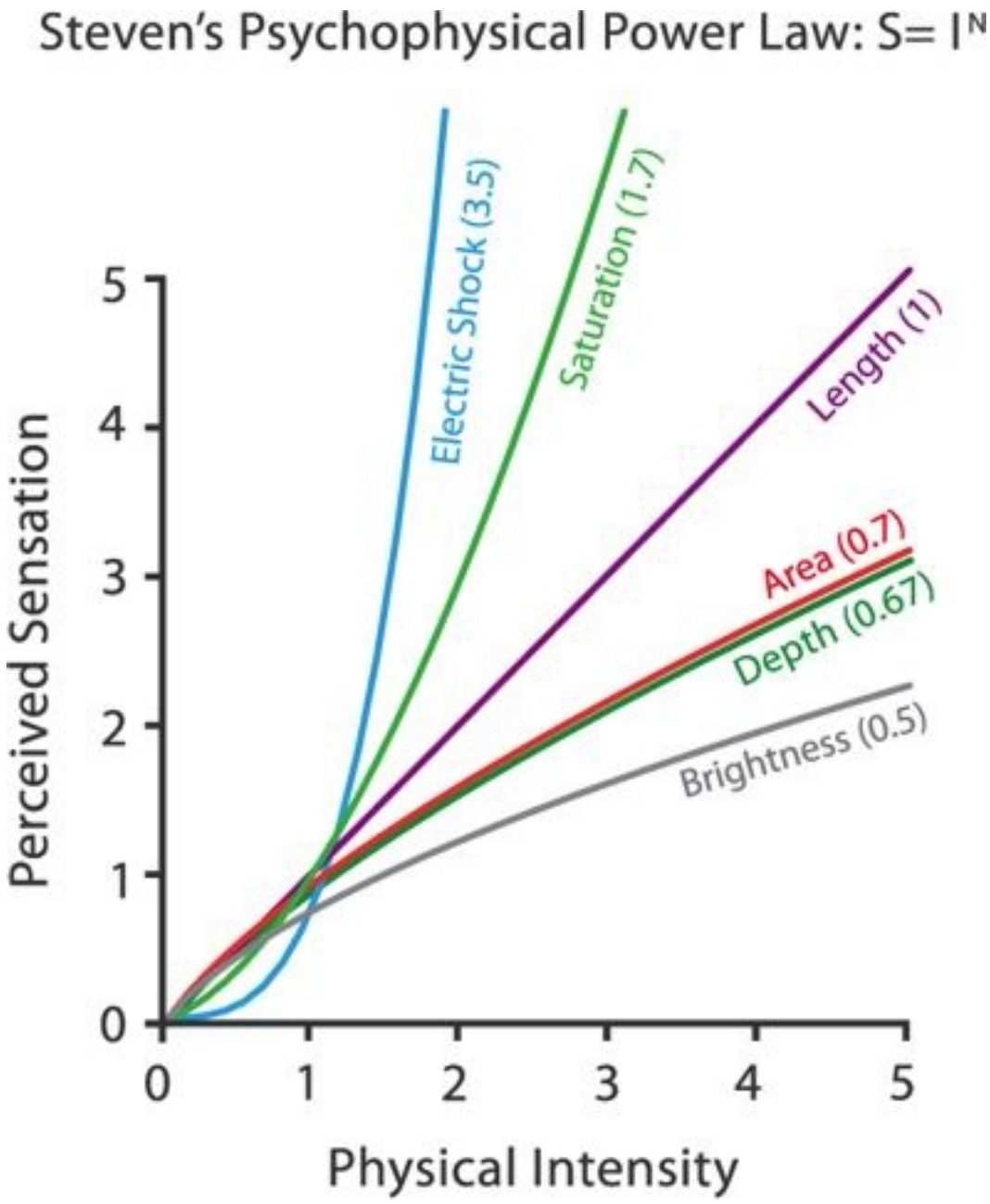


Muzner ranking of channel efficacy.

Image from Muzner.

Channel efficacy: accuracy

- We perceive different visual channel with different levels of accuracy.
- Stevens* showed that some sensations are perceptually magnified compared with their objective intensity (when $n > 1$) and some compressed (when $n < 1$).
- Length perception is completely accurate, whereas area is compressed and saturation is magnified.



*S. S. Stevens. Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects. Wiley, 1975

Channel efficacy: discriminability

- Some channels have a limited number of "bins" (a distinguishable step or level within the visual channel).
  - e.g. colours.

Channel efficacy: separability

Some channels depend on, or interact with, each other.

- Separable channels – e.g. position and hue
- Integral channels – size and colour (not fully separable)

Size interacts with many channels, including shape.

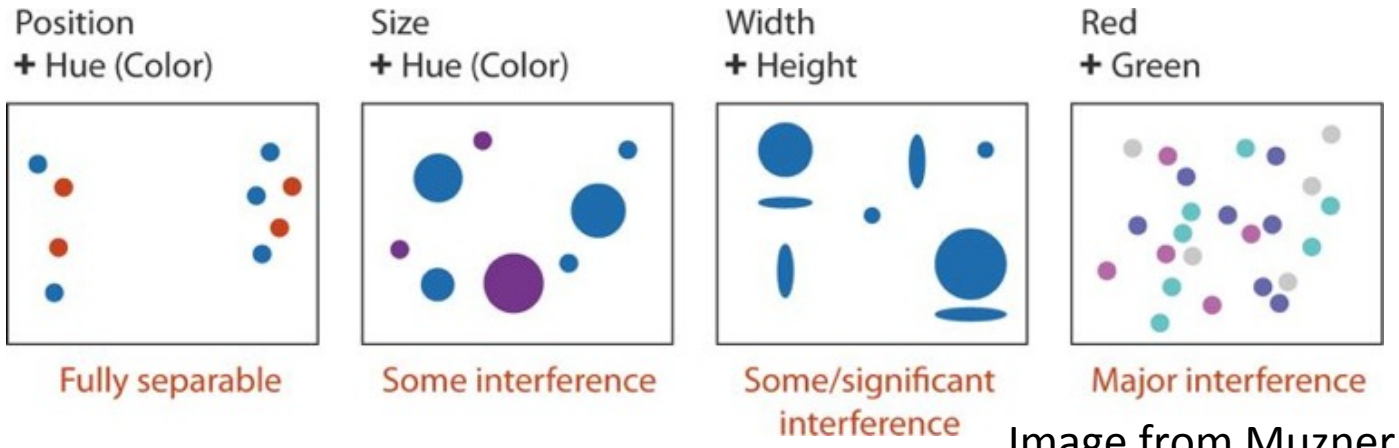


Image from Muzner.

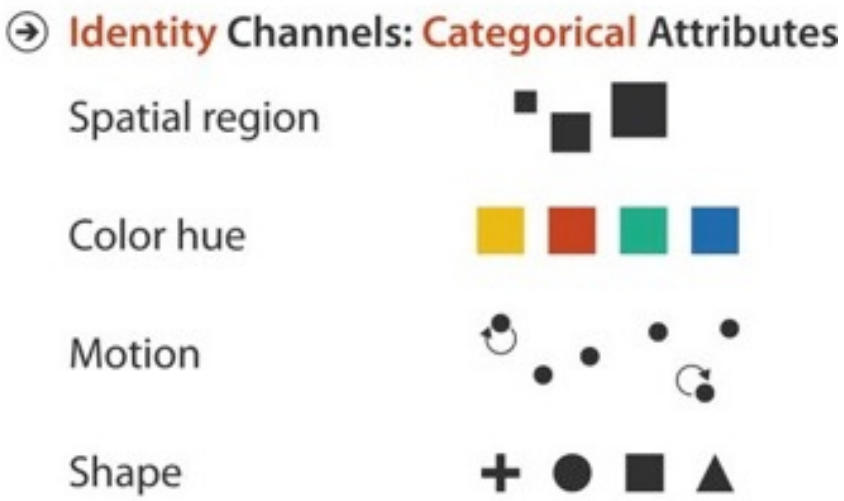
Channel efficacy: pop-out

Pop-out versus serial search –

In general, pop-out is only possible for a single channel at a time.

Channel efficacy: grouping

- Containment is the strongest cue,
- connection the second, then proximity.
- Then similarity - can also show grouping with identity channels.

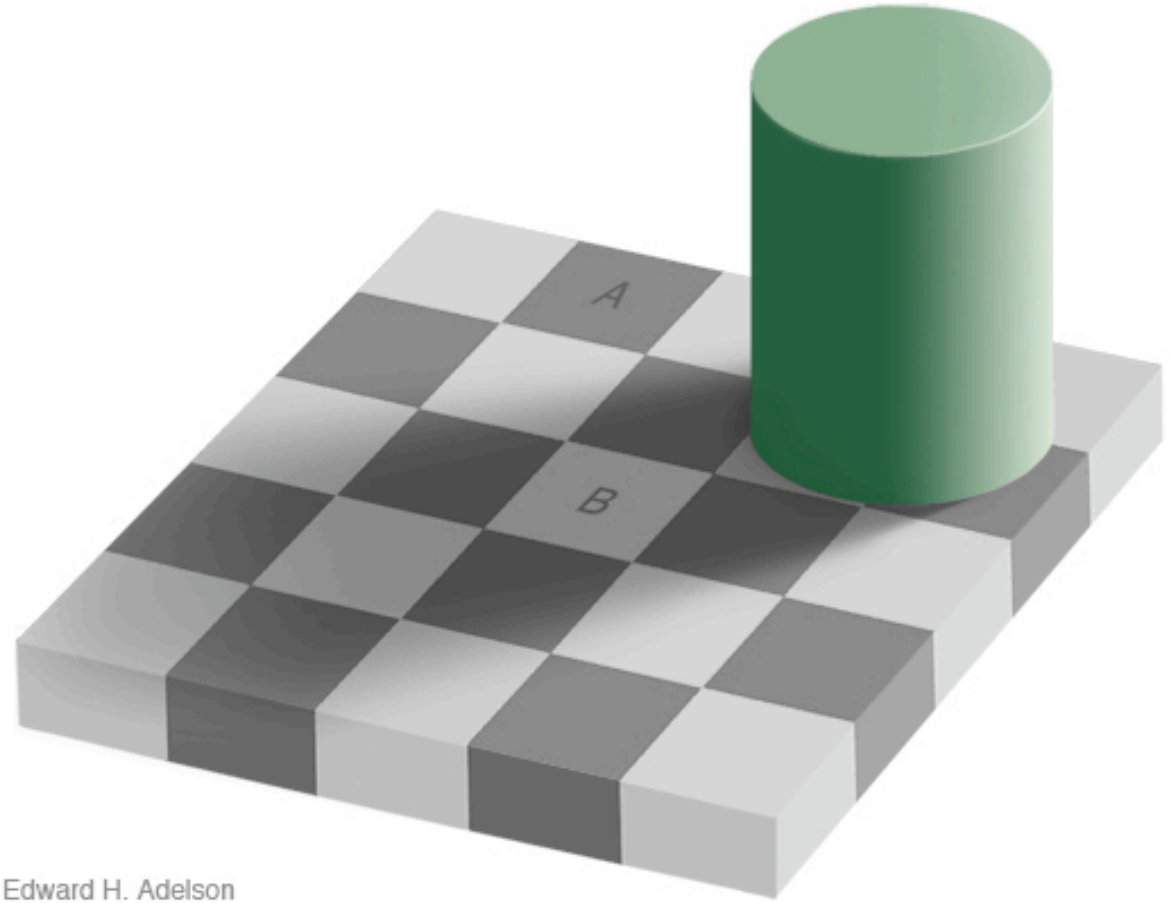


Channel efficacy: Relative versus absolute judgements

- Human perceptual system is based on relative judgements, not absolute ones.
  - E.g. the length difference we can perceive is a percentage of the object's length
  - Colour is affected by the background

Colour appearance

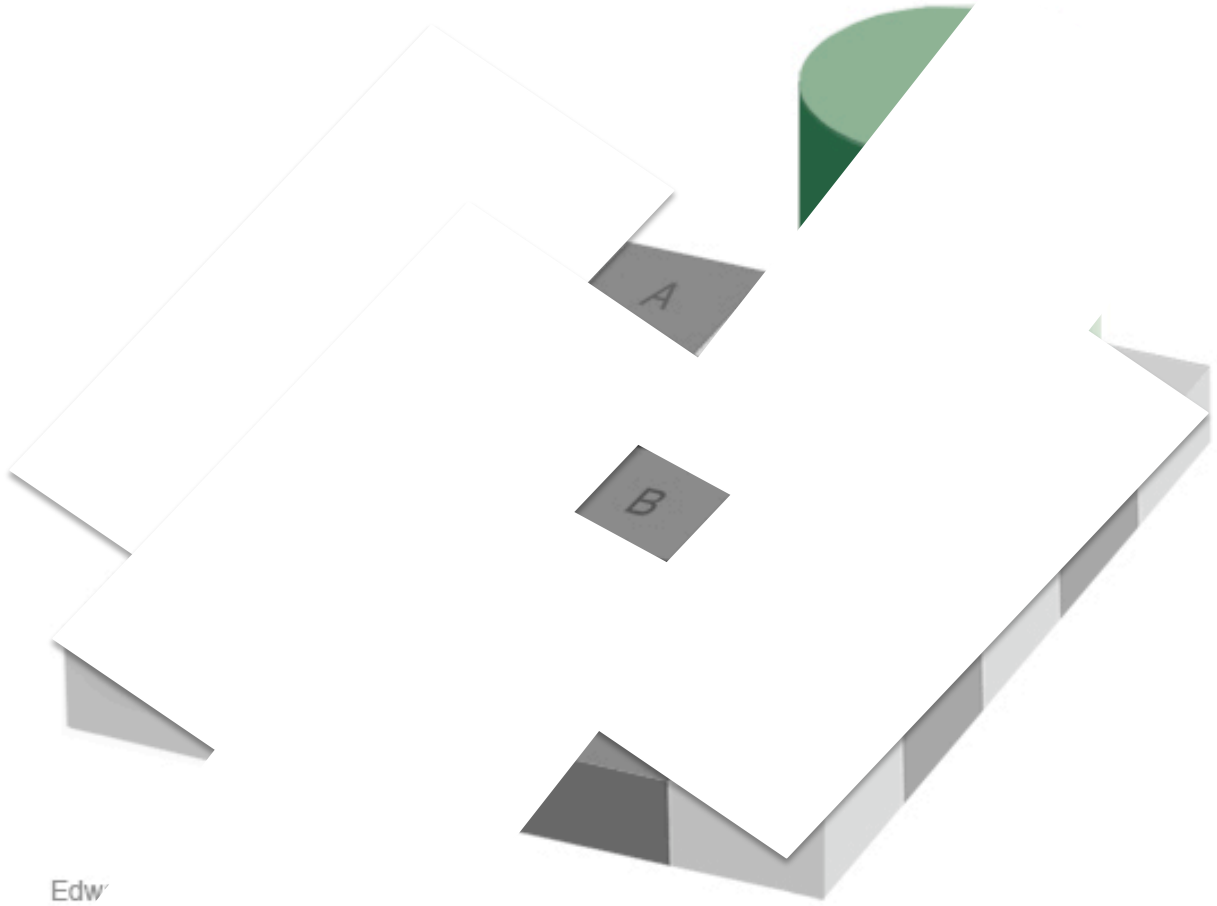
- The checkerboard illusion of Edward Adelson.



The squares marked A and B are the same shade of gray.

Colour appearance

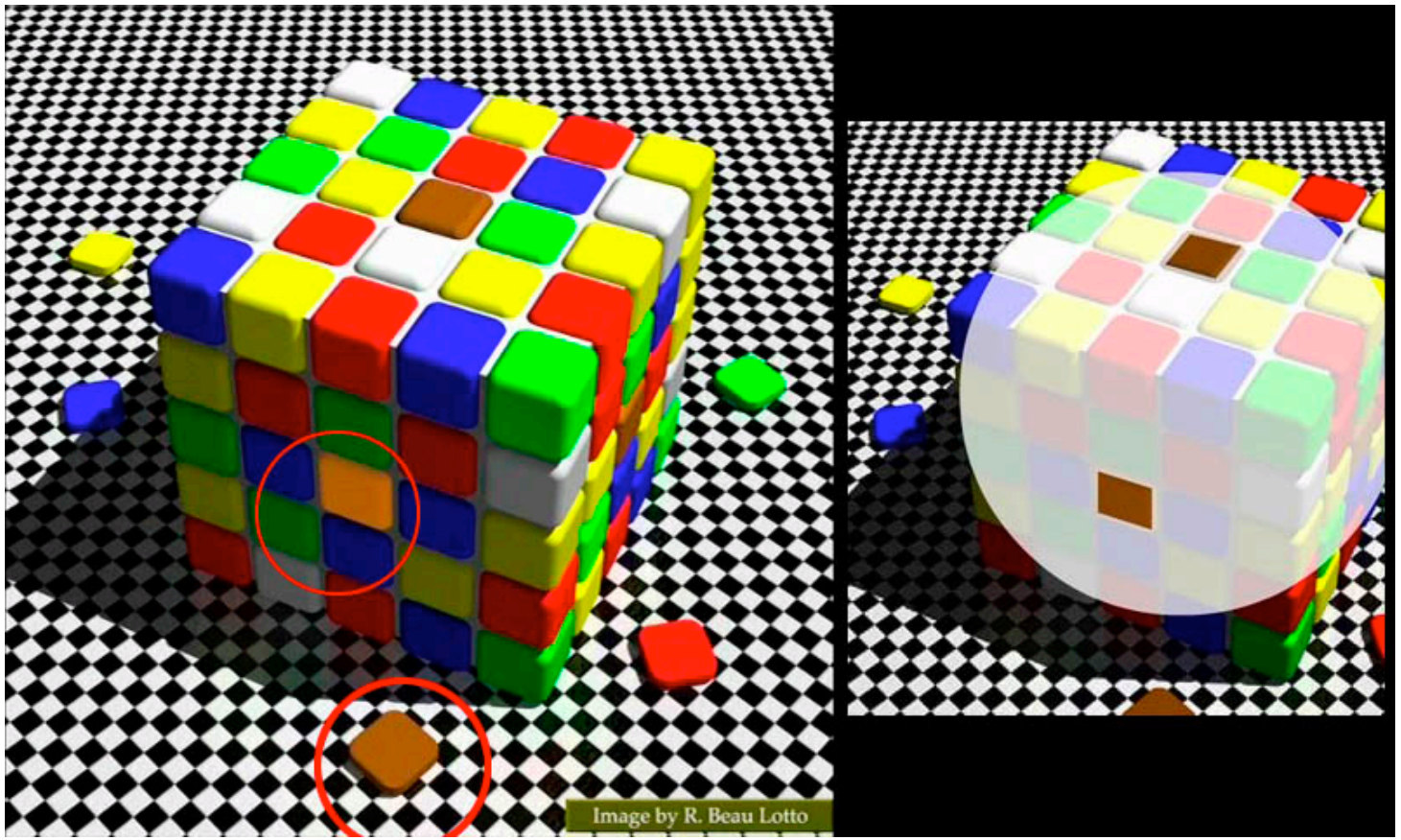
- The checkerboard illusion of Edward Adelson.



The squares marked A and B are the same shade of gray.

Colour appearance

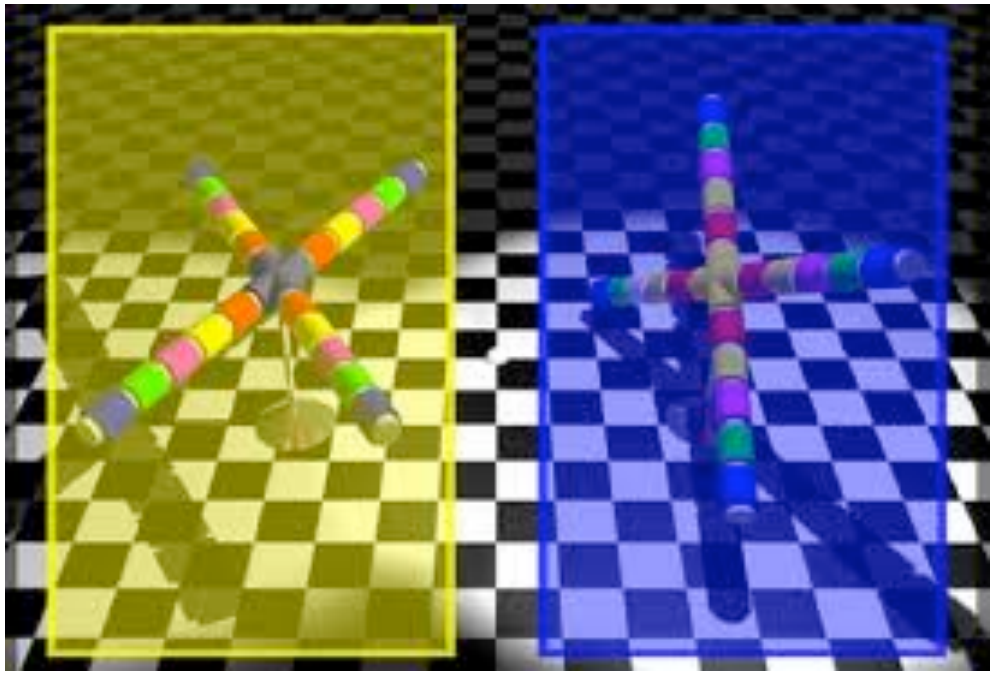
- The cube illusion of R. Beau Lotto. C.



The orange and brown squares are actually the same colour! We **see** the same image colour as being dark brown in the context of strong lighting, and light orange where the same image colour appears in a deeply shaded context.

Colour appearance

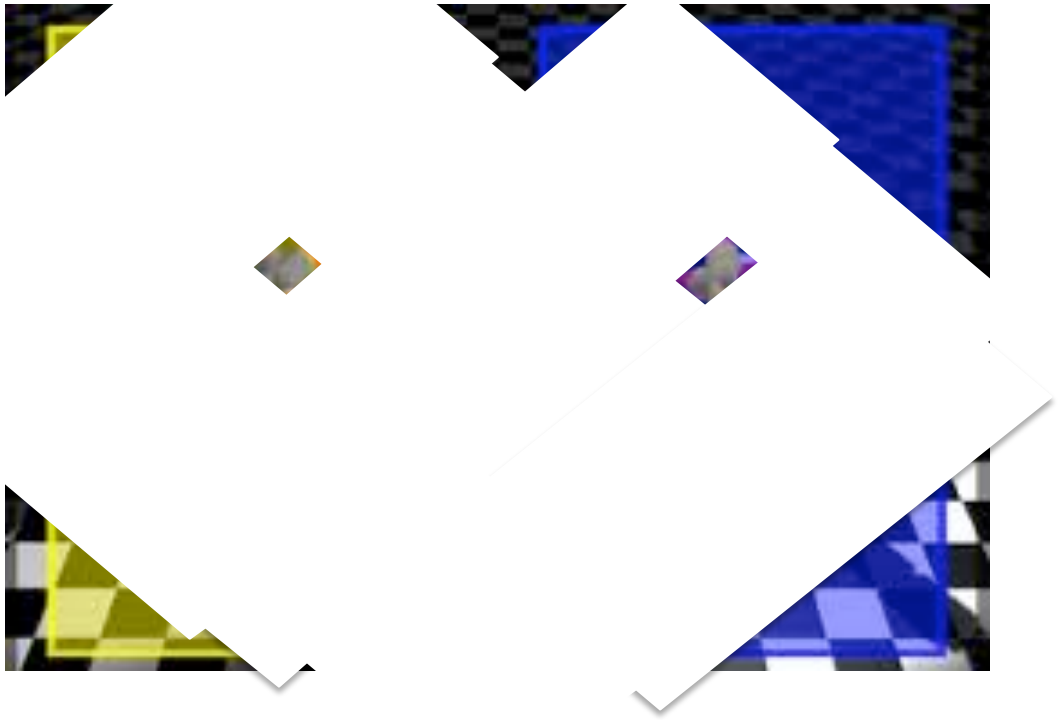
- The cross-piece illusion of R . Beau Lotto



The colour at the intersection of the two rods is actually an identical colour (grey) in both cases, but in the context of apparently yellow illumination on the left and blue illumination on the right, this is judged, **and seen**, to be the reflectance of a blue-grey object and a yellow object respectively.

Colour appearance

- The cross-piece illusion of R . Beau Lotto

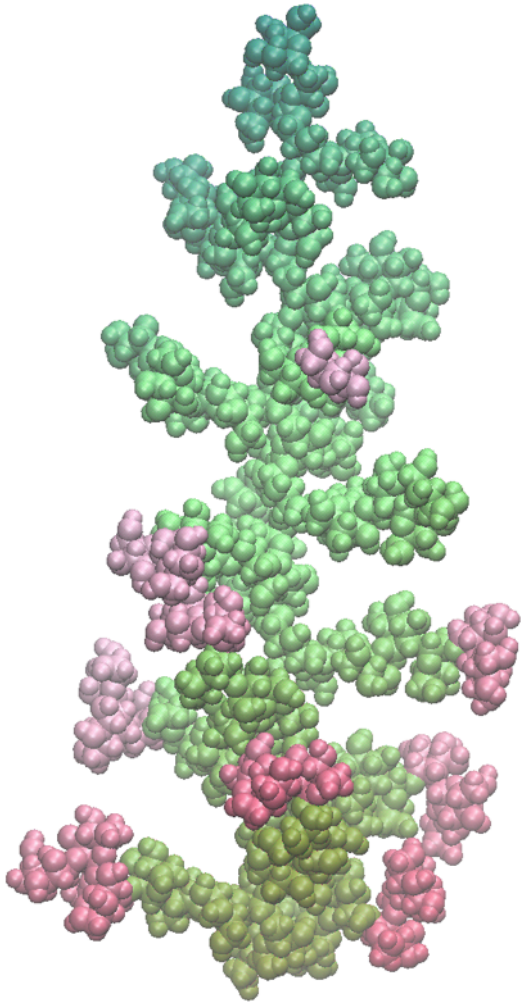


The colour at the intersection of the two rods is actually an identical colour (grey) in both cases, but in the context of apparently yellow illumination on the left and blue illumination on the right, this is judged, **and seen**, to be the reflectance of a blue-grey object and a yellow object respectively.

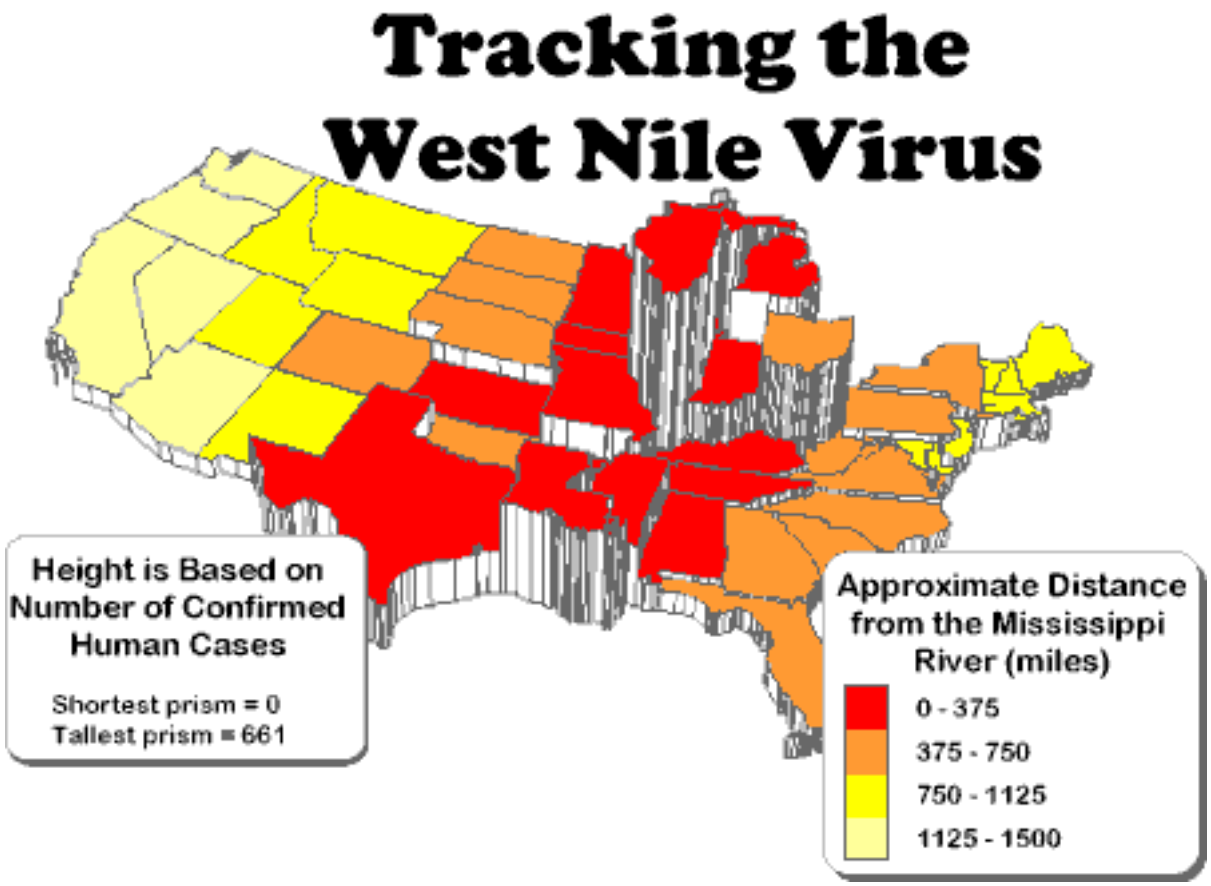
Practical Rules of Thumb (Heuristics)

NO unjustified 3D!

- 3D easy to justify for inherently 3D structures.
- But not necessarily better than 2D otherwise.



3D visualization

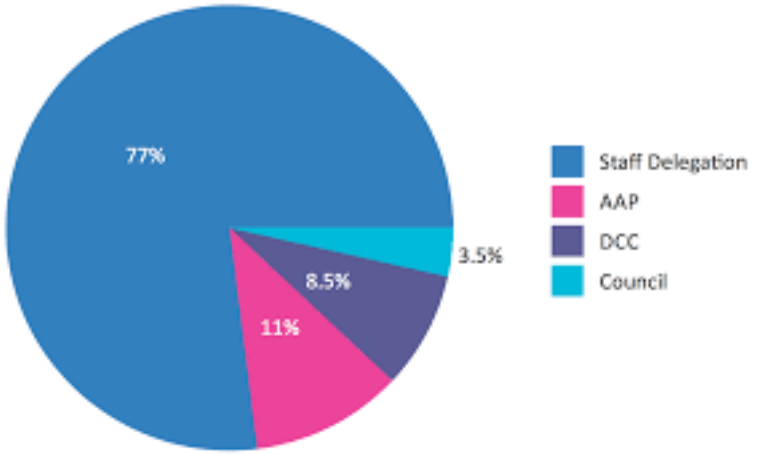


No unjustified 2D!

- 1D lists have several strengths
  - Density of information
  - Excellent for lookup tables

  - 2D good for relationships and topological structure.

Bad chart examples: pie charts

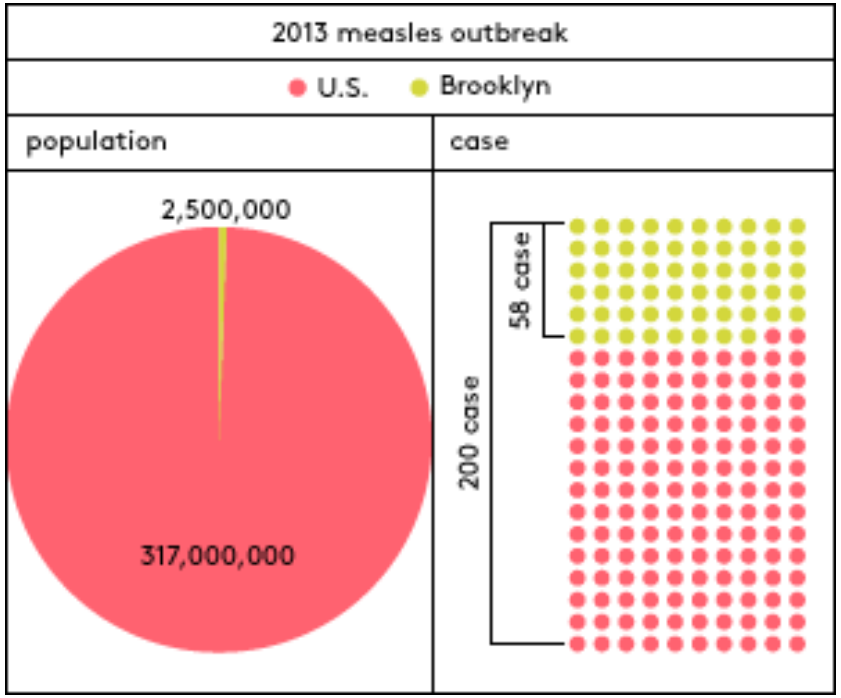


Pie charts require angle and area judgments

- We find area harder to judge than angle.
- Showing part-of-a-whole is equally well done by other means
  - E.g. stacked bar chart.
- Pie charts require more screen area.

Bad chart examples: **pie charts**

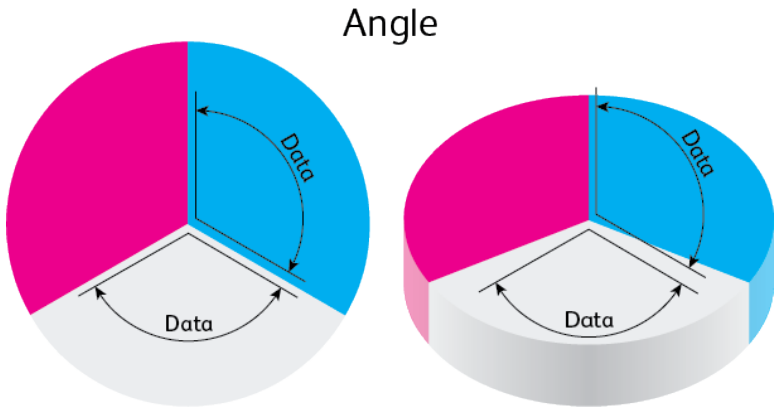
Edward Tufte gives the pie chart a succinct and decisive treatment in "The Visual Display of Quantitative Information":

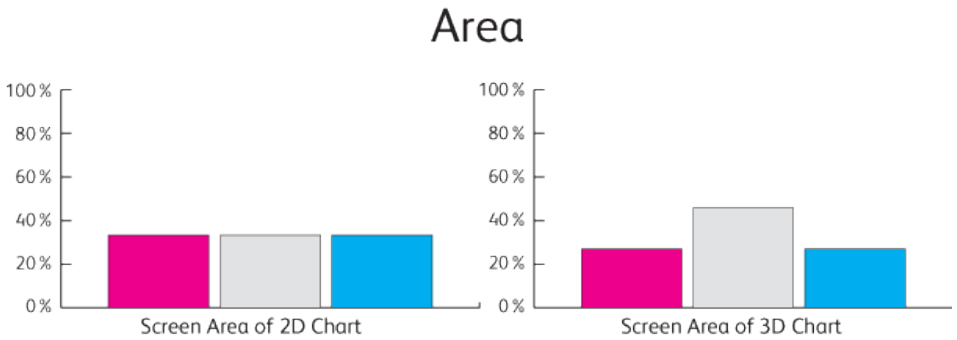


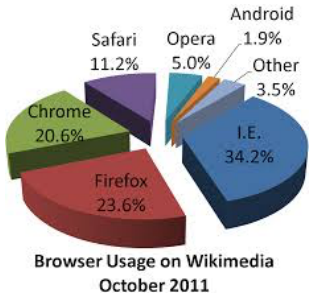
A table is nearly always better than a dumb pie chart; the only worse design than a pie chart is several of them, for then the viewer is asked to compare quantities located in spatial disarray both within and between charts [...] Given their low density and failure to order numbers along a visual dimension, pie charts should never be used.

And worse – 3D pie charts

- 3D pie charts are even worse – distort angles and areas.

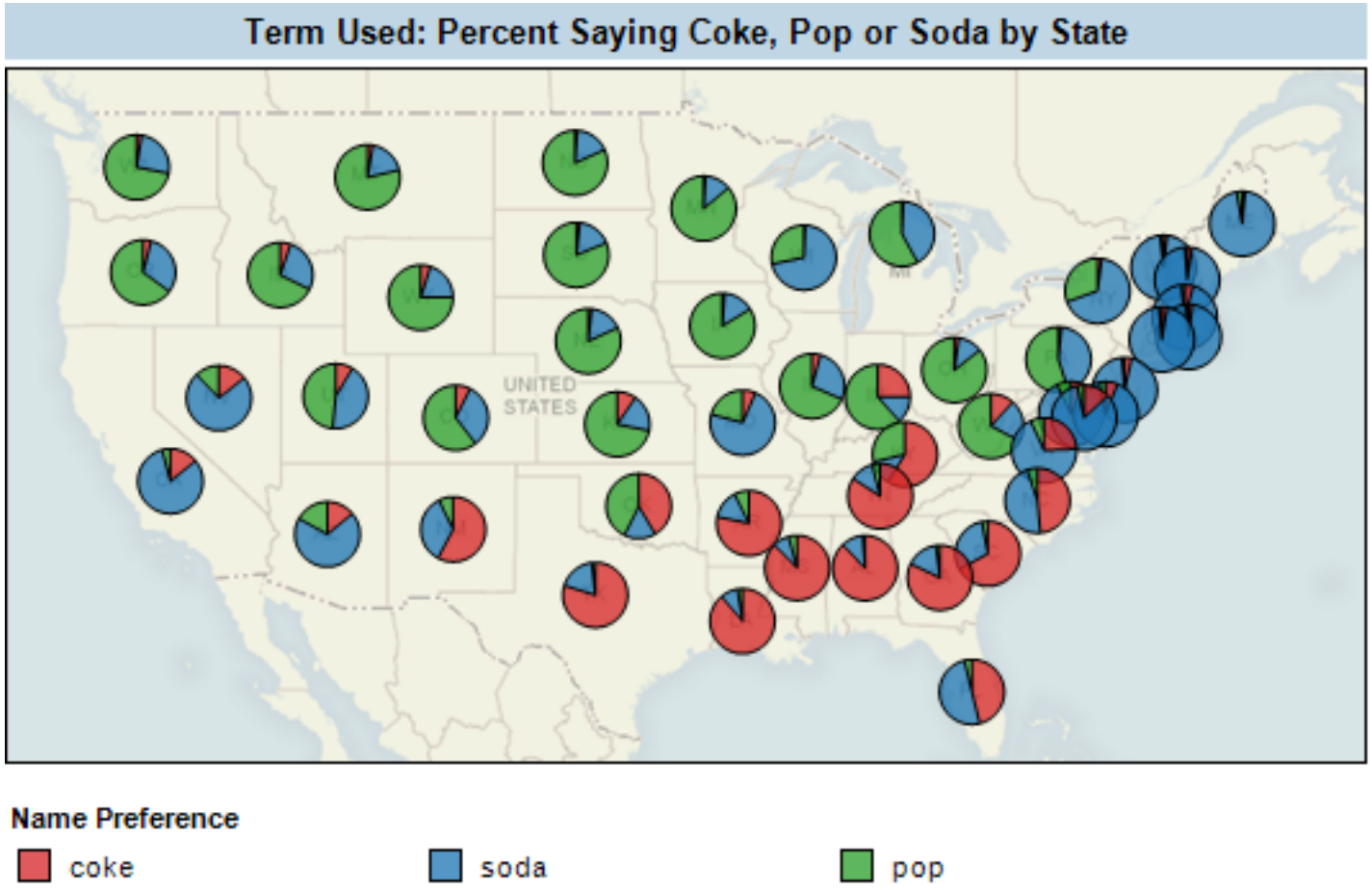






http://peltiertech.com/extra-distortion-in-a-pie-chart/

Do You Say Coke, Soda or Pop?

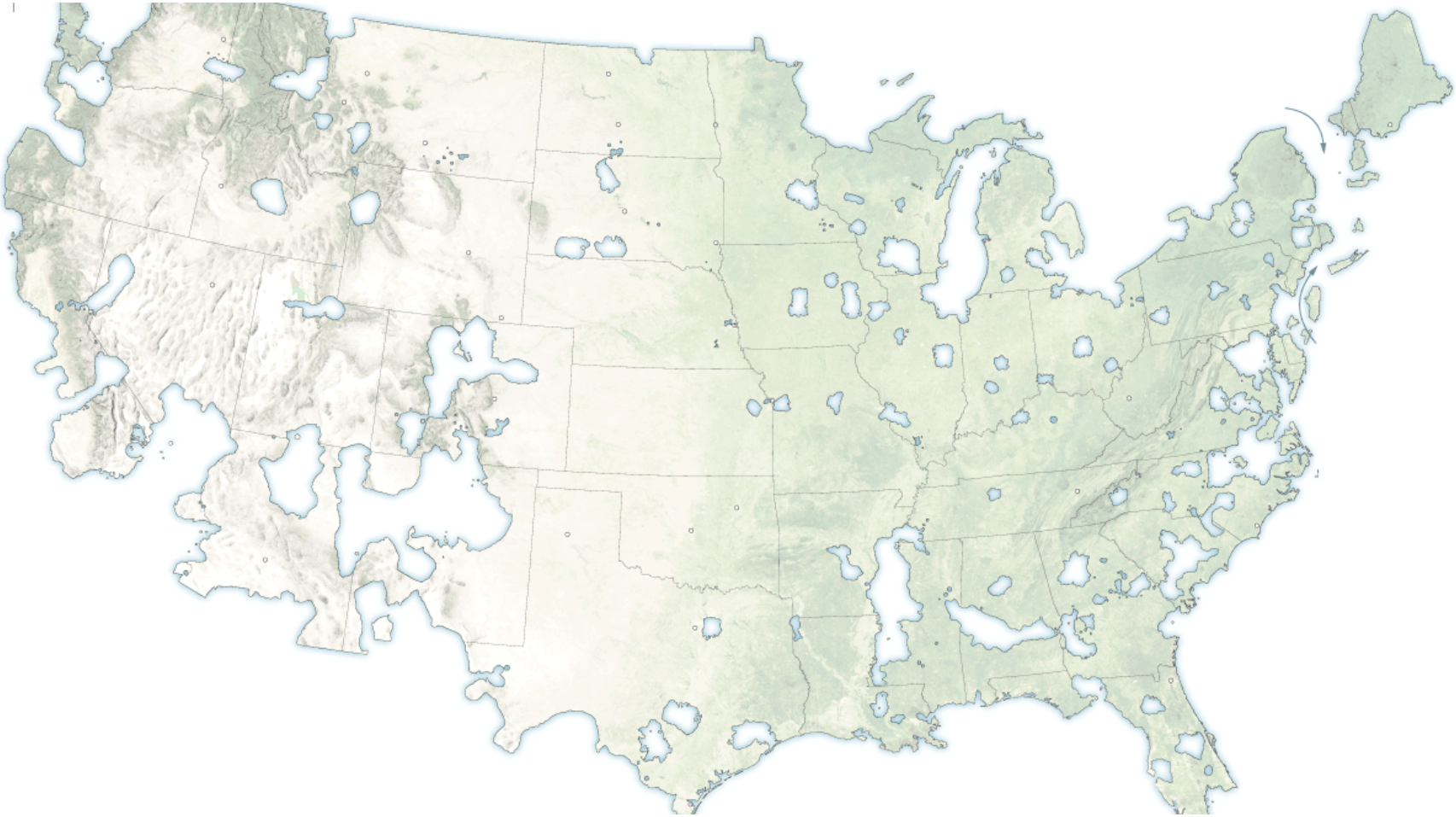


http://www.tableau.com/blog/do-you-say-coke-soda-or-pop-map-visualization-shows-your-likely-answer

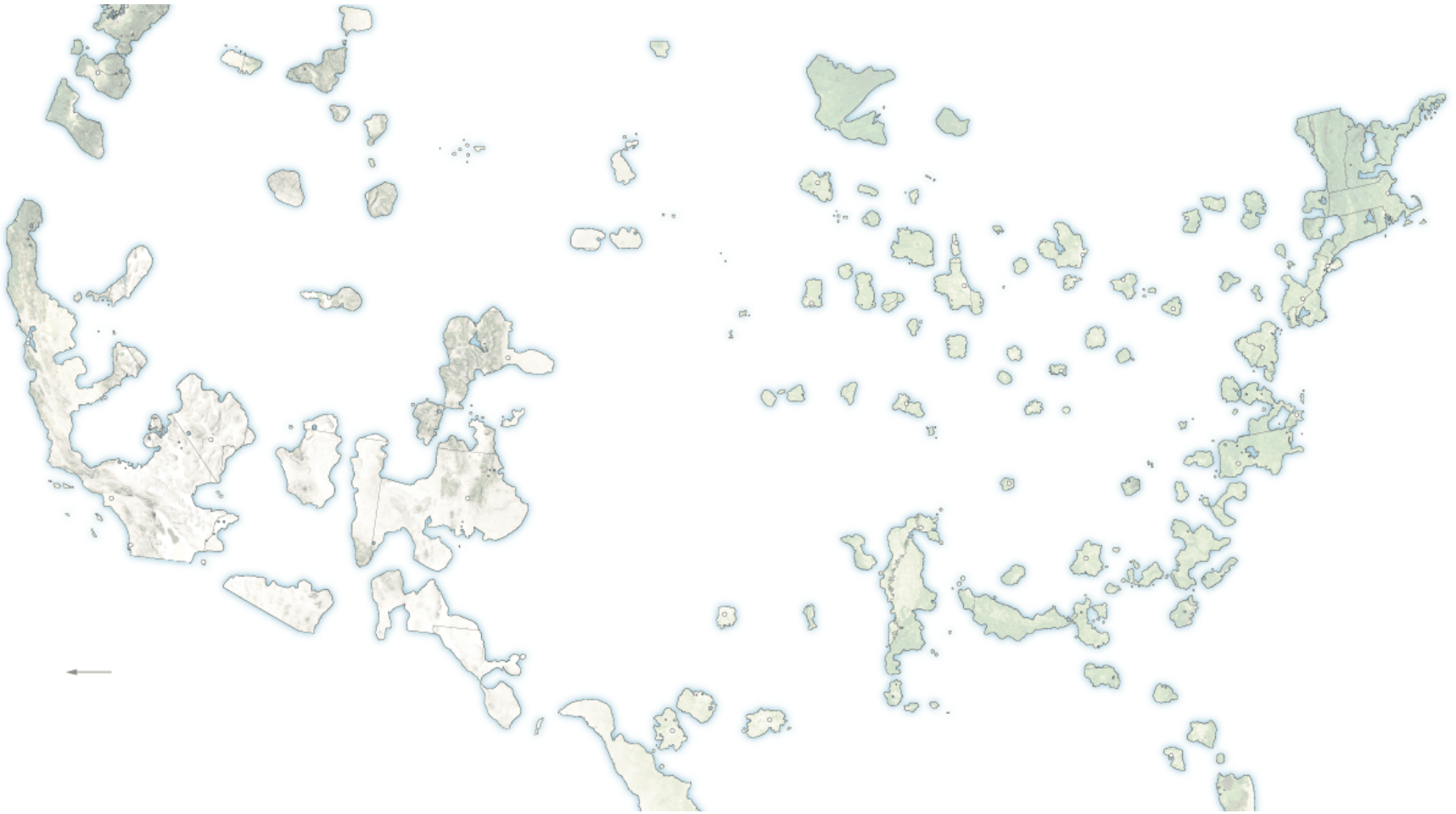
Eyes beat memory

- Switching between/comparing 2 views has a very low cognitive load.
  - Small overview window is very effective.
- We have very limited memory and attention.
- Animation:
  - Jumping between 2 scenes can be very effective – blink comparator idiom
  - However, multiframe animations are difficult to keep track of
    - Having all frames can be more effective than animation (but obviously limit to number).

Trump’s America



Clinton's America



Overview first, zoom and filter, details on demand

- Schneiderman's mantra*
- May need filtering/aggregation for overview

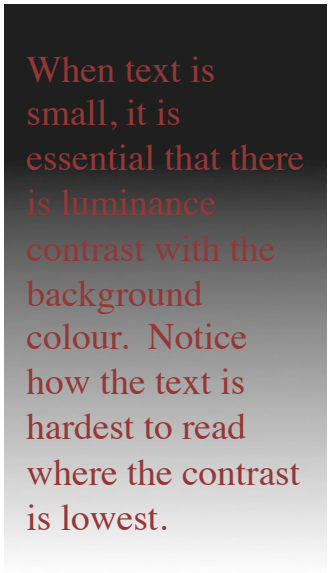
Ben Shneiderman. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." In Proceedings of the IEEE Conference on Visual Languages, pp. 336–343. IEEE Computer Society, 1996.

Get it right in black and white

- the most crucial aspects of visual representation must be legible even if the image is transformed from full colour to black and white.

Colour - Principles for design

1. When small **detail** is important, luminance contrast is necessary:
   - black on white
   - dark blue on white
   - yellow on black



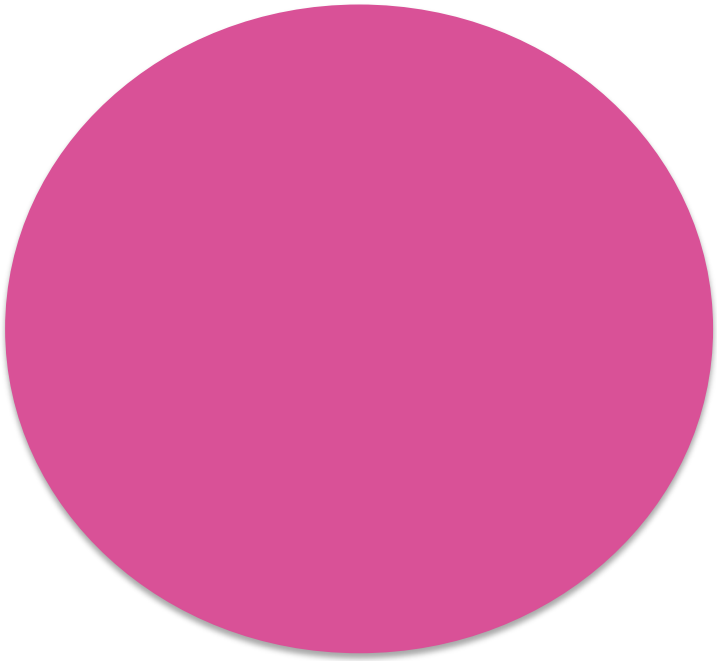
ISO recommends a luminance ratio of at least 3:1 between text and background

2. Colour-coding information

- The most important use of colour is to indicate categories of information
- When designing colour codes need to be concerned with:
  - **visual distinctness**, to support visual search queries
  - **learnability**, so colours come to stand for particular entities
    - for learnability, it is important that unique hues are used first – **red, green, yellow, blue**, - followed by colours that have relatively consistent names: **pink, brown, orange, grey and purple**

Colour names

What colour is this?



Colour names

What colours are these?

- teal
- mauve
- khaki
- puce
- Ochre
- terracotta

2. Colour-coding information

- If a design is complex and symbols are quite small, no more than a dozen codes can be used reliably
  - backgrounds can distort a patch of colour
  - ease of a visual search depends both on the colour and on the background colour
  - small areas should be strongly coloured and have black-white channel differences from large areas to be distinct
  - large areas can have more subdued colours
    - use low saturation colours for large areas

Function, then form

- The best vis designs should be both beautiful and effective.
  - But effective is key,
  - Beautiful is nice.

The Theory of Data Graphics

"There are right ways and wrong ways to show data; there are displays that reveal the truth and displays that do not." [E. R. Tufte]

- Data graphics should draw the viewer's attention to the **sense** and **substance** of the **data**,
  - not to methodology, *graphic design*, something else

  The main purpose of a visualization isn't entertainment — it's effective communication.

  A visualization should represent data in a way that makes information assimilation and comprehension simple.

Graphical Ducks

"It is all right to decorate construction, but never to construct decoration"

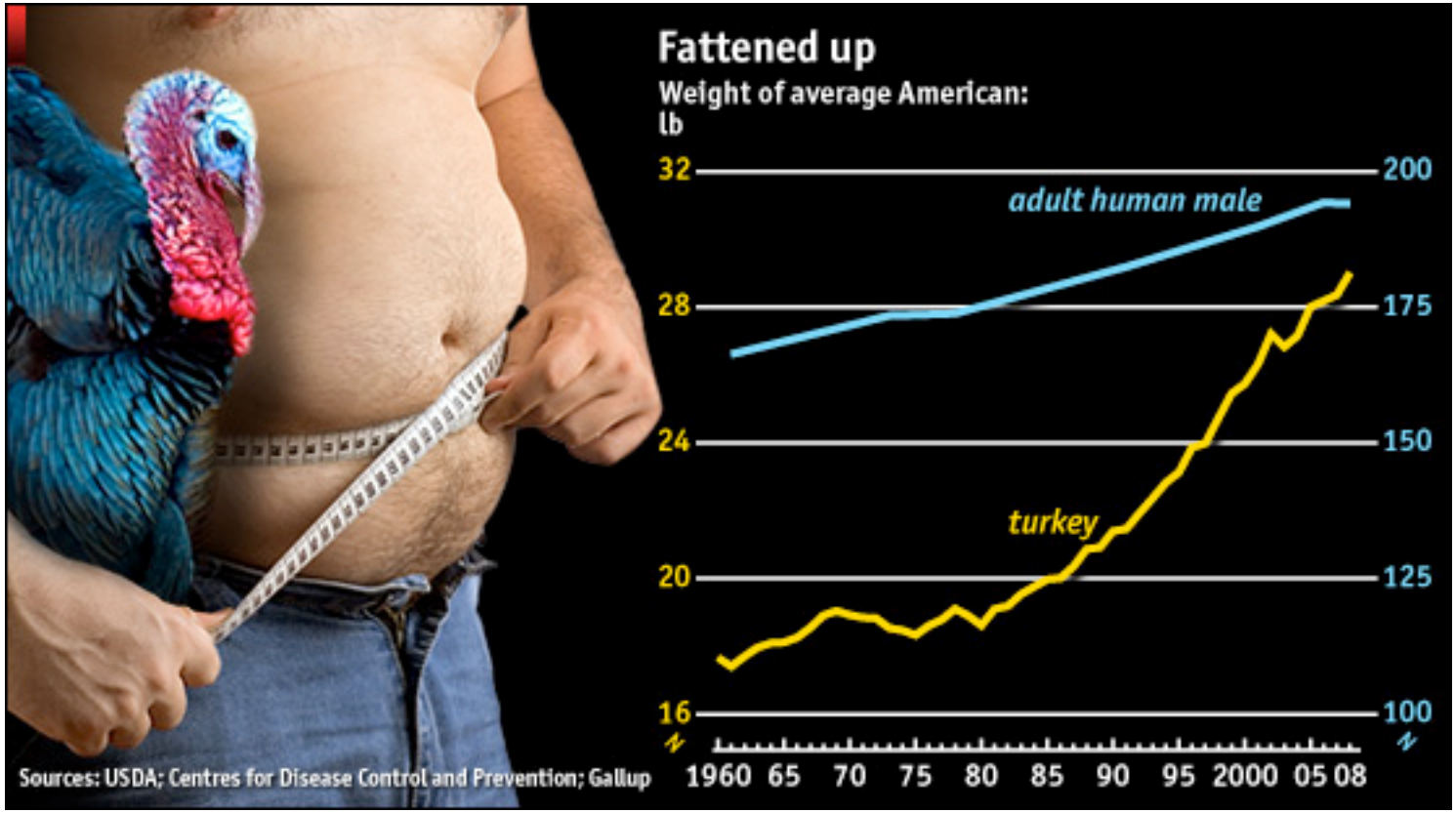
Brown, and Steven Izenour write about the ducks of modern architecture—and their thoughts are relevant to the design of data graphics as well:

> When Modern architects righteously abandoned ornament on buildings, they unconsciously designed buildings that *were* ornament. In promoting Space and Articulation over symbolism and ornament, they distorted the whole building into a duck. They substituted for the innocent and inexpensive practice of applied decoration on a conventional shed the rather cynical and expensive distortion of program and structure to promote a duck. . . . It is now time to reevaluate the once-horrifying statement of John Ruskin that architecture is the decoration of construction, but we should append the warning of Pugin: It is all right to decorate construction but never construct decoration.[2]

[2] Robert Venturi, Denise Scott Brown, and Steven Izenour, *Learning from Las Vegas* (Cambridge, revised edition, 1977), 163. The initial statement of the duck concept is found on 87–103.

Big Duck, Flanders, New York; photograph by Edward Tufte, July 2000.

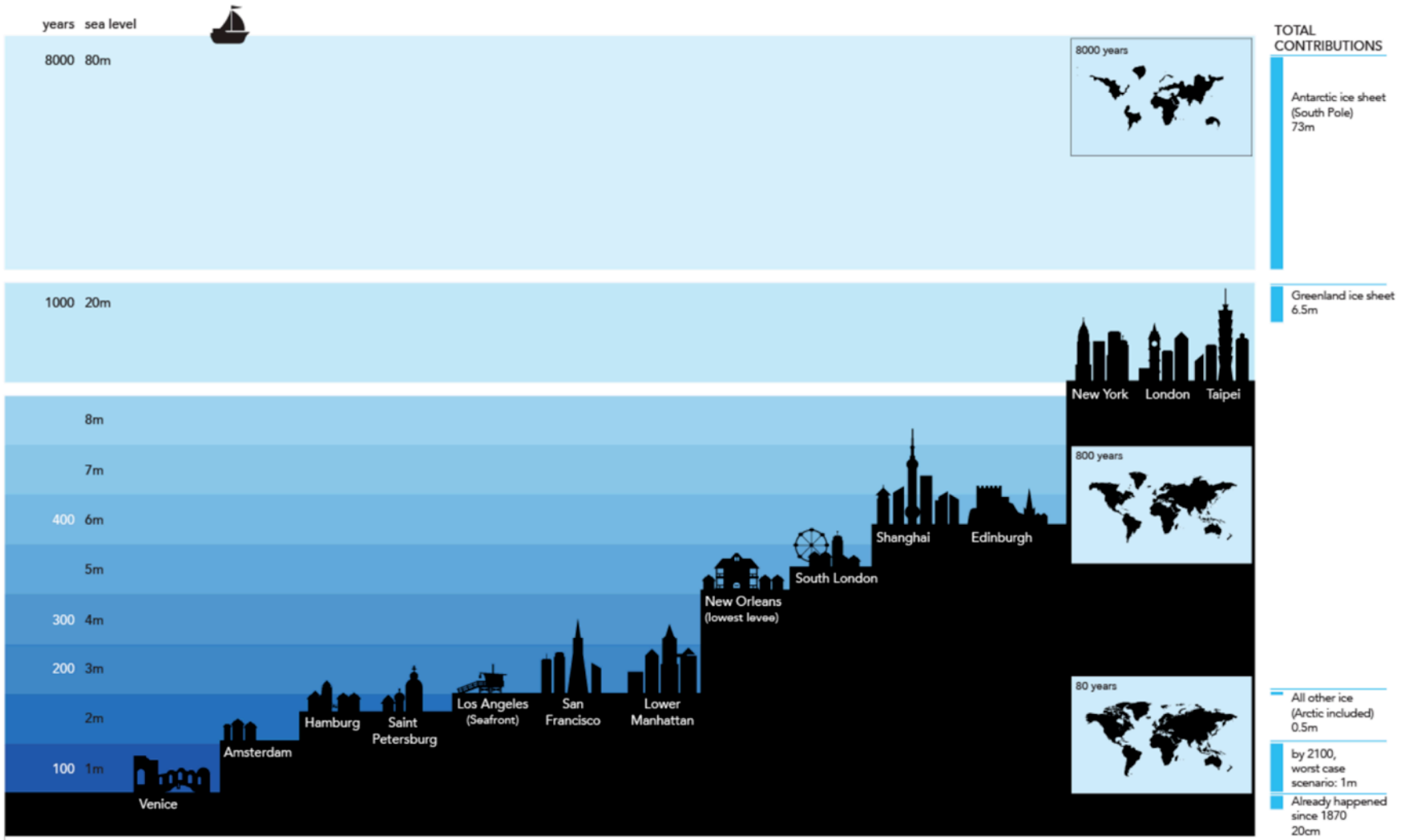
A Turkey or a Duck?



From *The Economist*, Nov 26th 2009 *Web only*

When Sea Levels Attack!

How long have we got?



TOTAL CONTRIBUTIONS

- Antarctic ice sheet (South Pole) 73m
- Greenland ice sheet 6.5m
- All other ice (Arctic included) 0.5m
- by 2100, worst case scenario: 1m
- Already happened since 1870 20cm

David McCandless // V3.0// Feb 2014
Illustration: Laura Sullivan, Joe Swainson, Fabio Bergamaschi

InformationIsBeautiful.net

Note:
Heights above sea level vary across cities. Lowest points used

Source: IPCC, NASA, Realclimate.org, NewScientist.com, Potsdam institute, Sea Level Explorer

Visualization

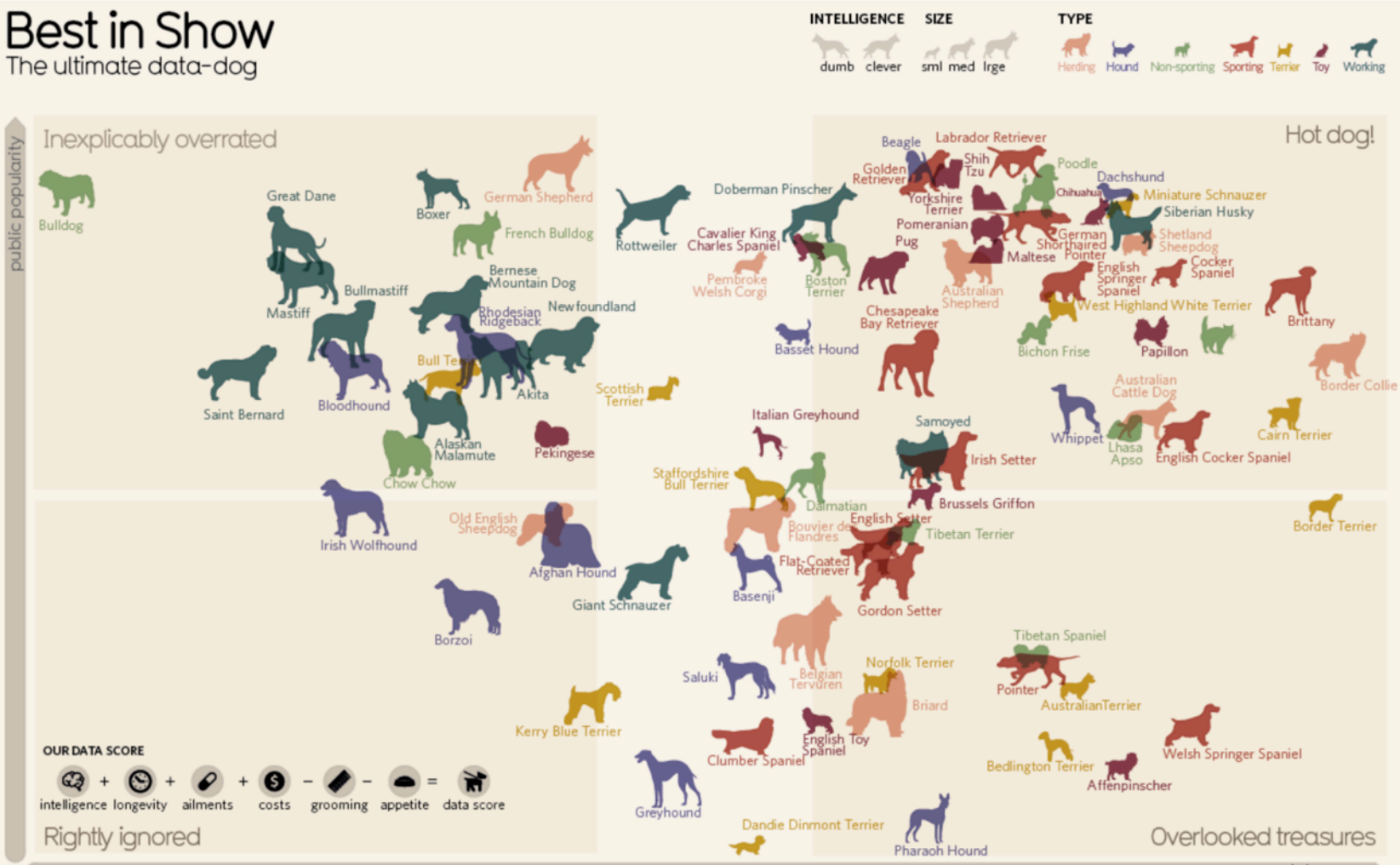
- visualization: graphical representation of some data or concepts
  - tools for visual thinking

"Evidence presentations should be created in accord with the **common analytical tasks** at hand, which usually involve **understanding causality**, making **multivariate comparisons**, examining **relevant evidence**, and assessing the **credibility of evidence and conclusions**."

Edward Tufte in *Beautiful Evidence* [own emphasis]

Best in Show

The ultimate data-dog



by David McCandless

research: @MiriamQuick / dogs: Andrew Park @illustrationkid

informationisbeautiful.net

taken from the new, infographic mega-book

Knowledge is Beautiful

bit.ly/KIB_books

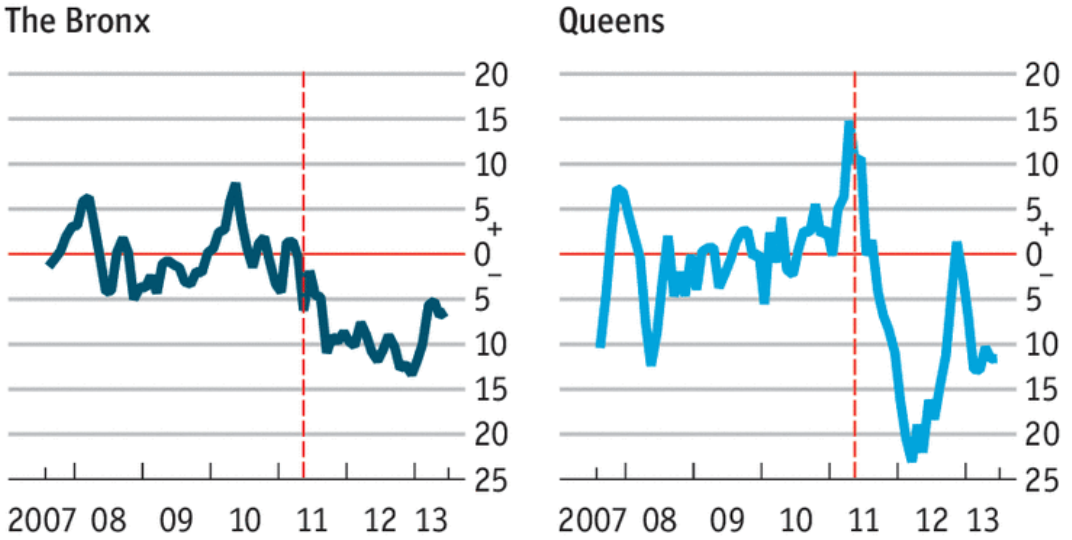
sources: American Kennel Club, Canine Inherited Disorders Database,

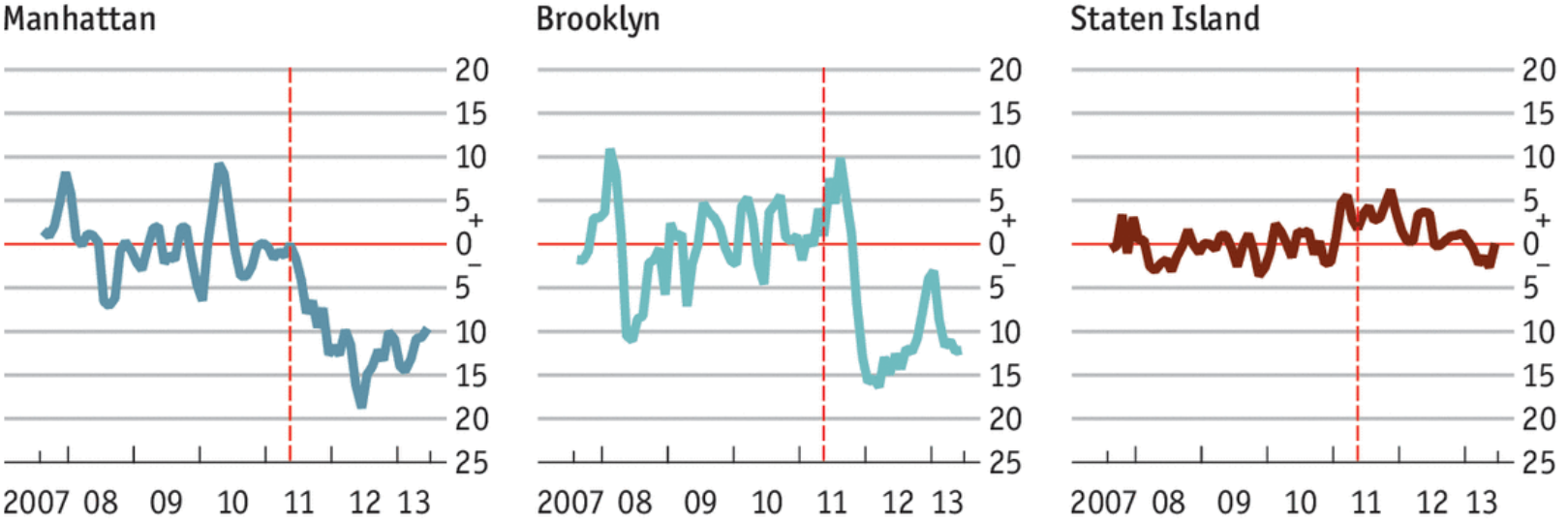
data: bit.ly/KIB_BestDogs

Worth hailing

Alcohol-related crashes in New York City

Difference* in the number of crashes in boroughs when compared with similar counties



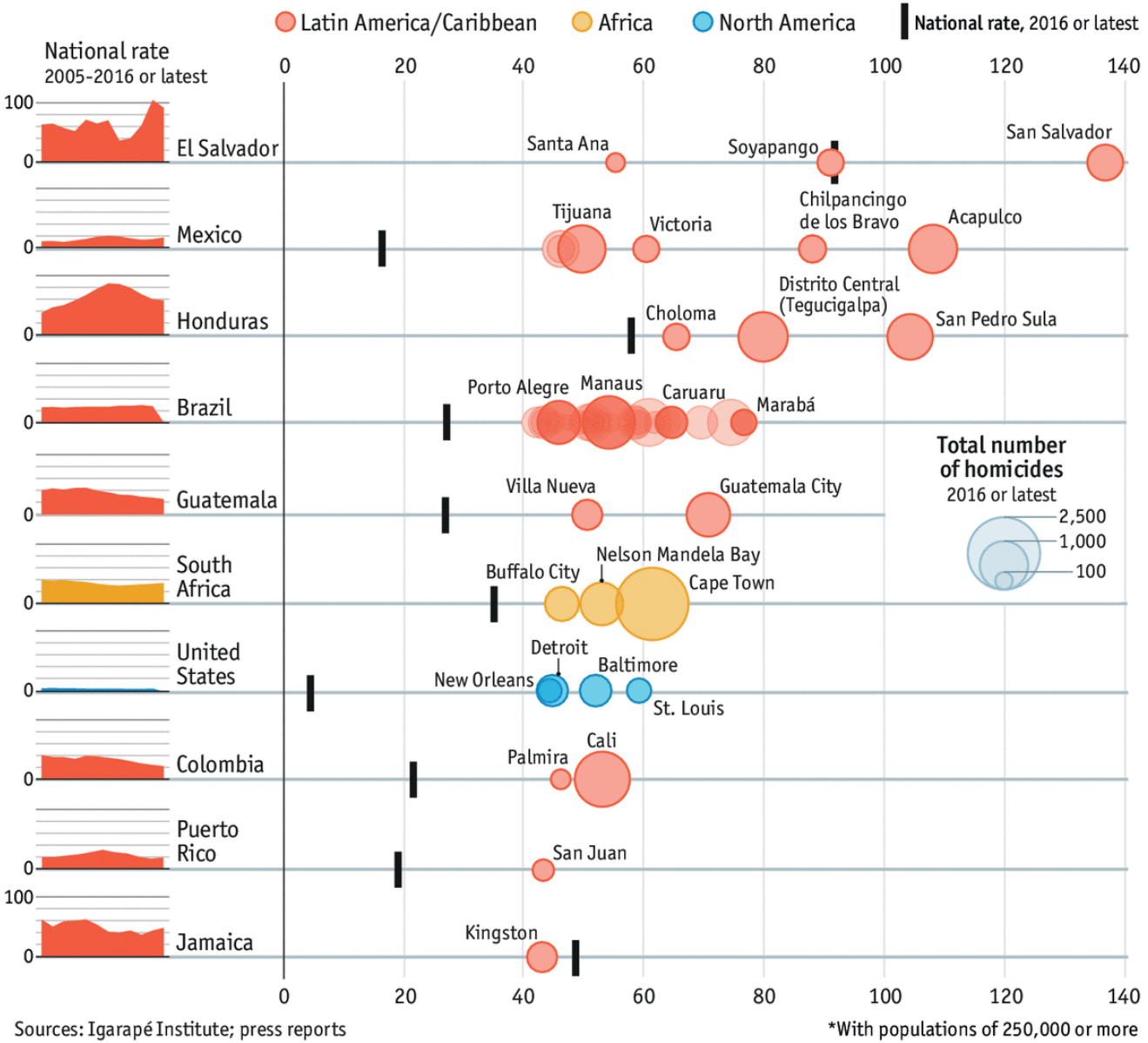


Source: "New York City Drunk Driving After Uber" by J. L. Peck, 2017

*Three-month moving average

Mean streets

Homicides per 100,000 population, 50 highest cities*, 2016 or latest

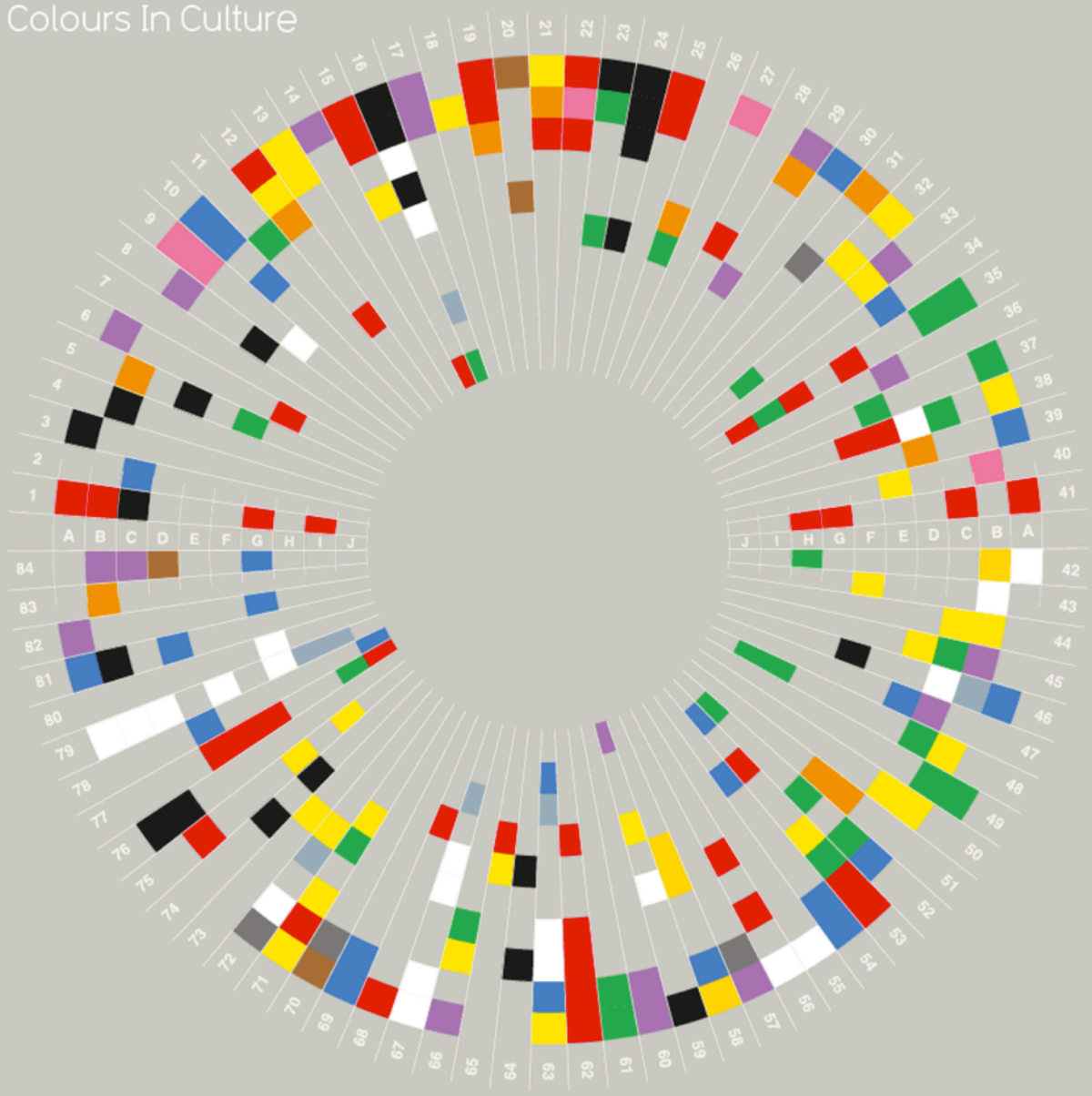


Sources: Igarapé Institute; press reports

*With populations of 250,000 or more

Economist.com

Colours In Culture



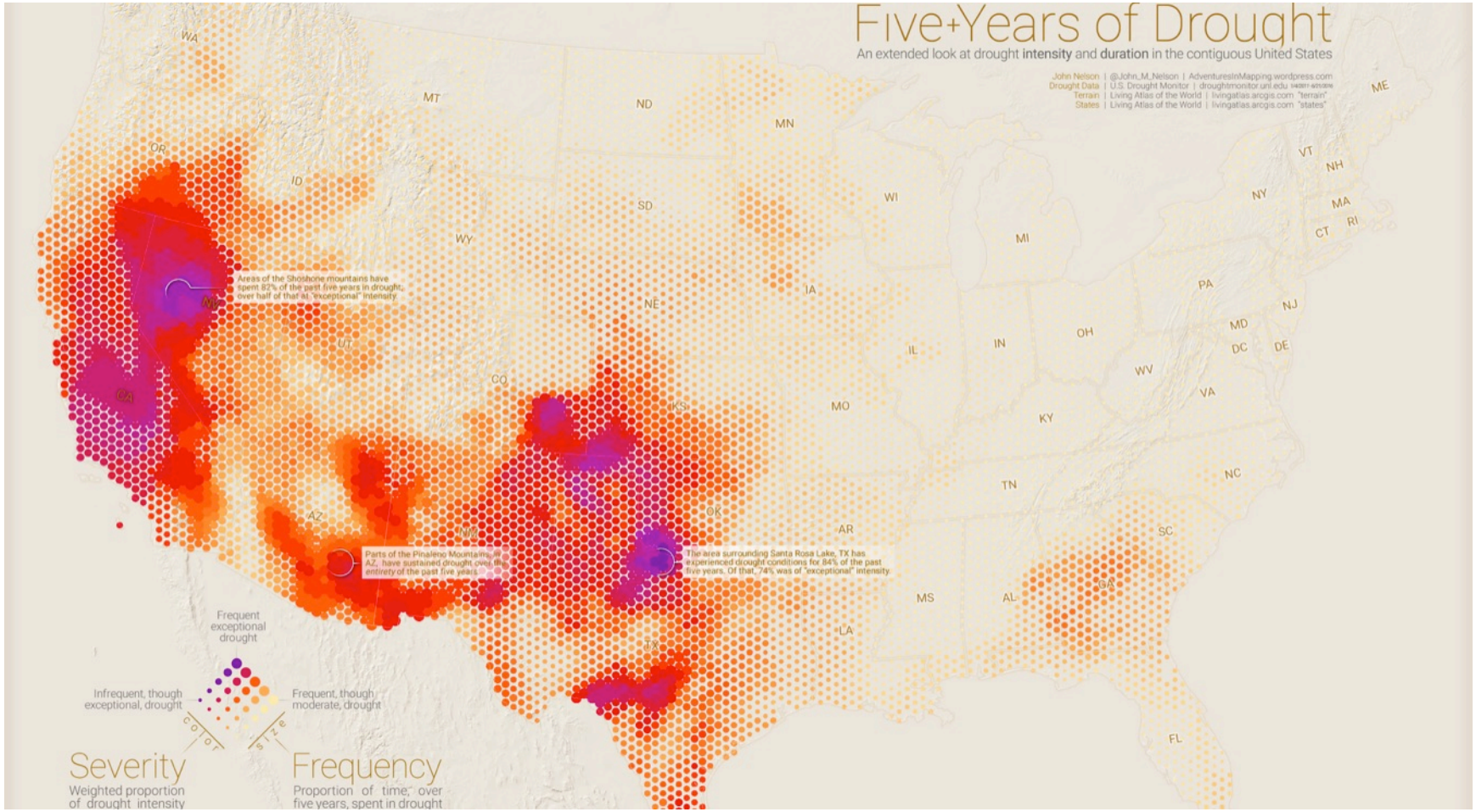
A Western / American	1 Anger	19 Desire
B Japanese	2 Art / Creativity	20 Earthy
C Hindu	3 Authority	21 Energy
D Native American	4 Bad Luck	22 Erotic
E Chinese	5 Balance	23 Eternity
F Asian	6 Beauty	24 Evil
G Eastern European	7 Calm	25 Excitement
H Muslim	8 Celebration	26 Family
I African	9 Children	27 Femininity
J South American	10 Cold	28 Fertility
	11 Compassion	29 Flamboyance
	12 Courage	30 Freedom
	13 Cowardice	31 Friendly
	14 Cruelty	32 Fun
	15 Danger	33 God
	16 Death	34 Gods
	17 Decadence	35 Good Luck
	18 Deceit	36 Gratitude

37 Growth	55 Luxury	73 Royalty
38 Happiness	56 Marriage	74 Self-cultivation
39 Healing	57 Modesty	75 Strength
40 Healthy	58 Money	76 Style
41 Heat	59 Mourning	77 Success
42 Heaven	60 Mystery	78 Trouble
43 Holiness	61 Nature	79 Truce
44 Illness	62 Passion	80 Trust
45 Insight	63 Peace	81 Unhappiness
46 Intelligence	64 Penance	82 Virtue
47 Intuition	65 Power	83 Warmth
48 Religion	66 Personal power	84 Wisdom
49 Jealousy	67 Purity	
50 Joy	68 Radicalism	
51 Learning	69 Rational	
52 Life	70 Reliable	
53 Love	71 Repels Evil	
54 Loyalty	72 Respect	

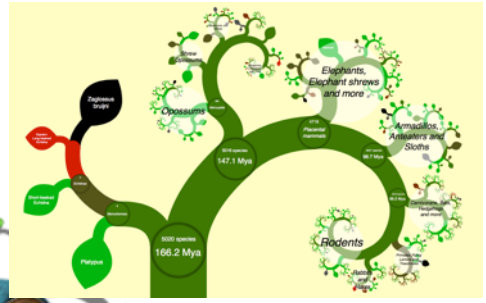
Yellow Grey
Gold Silver

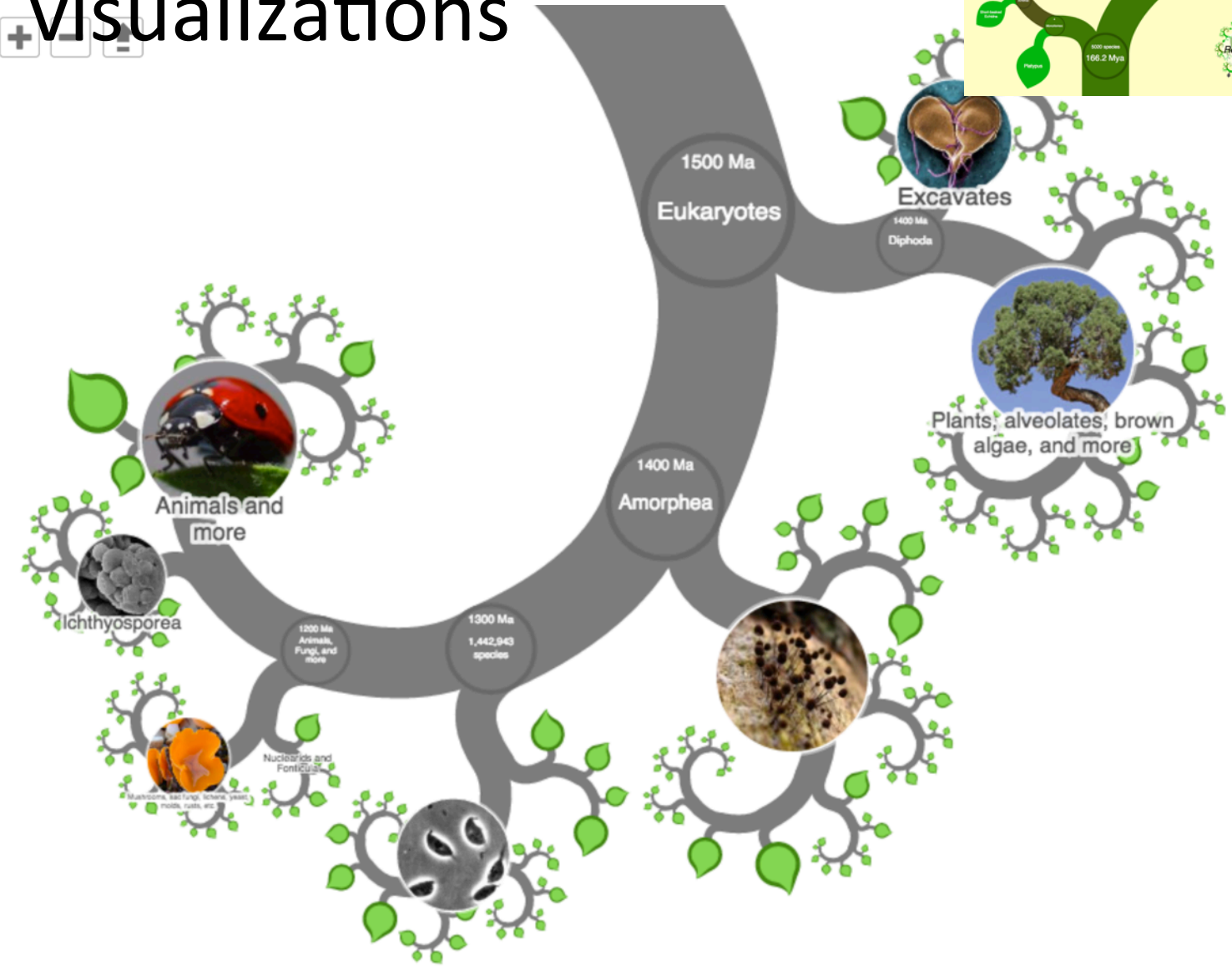
David McCandless & AlwaysWithHonor.com // v1.0 // Apr 09 // InformationIsBeautiful.net

source: Pantone, ColorMatters & web sources

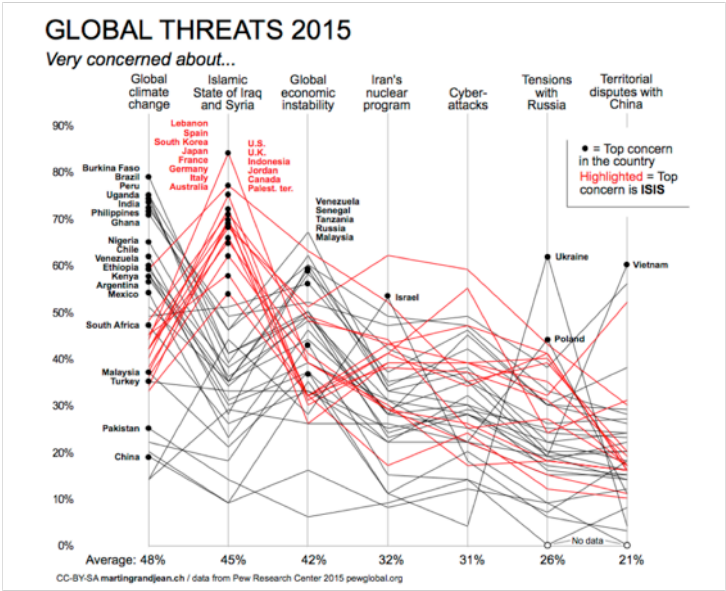

Five+Years of Drought
An extended look at drought intensity and duration in the contiguous United States
John Nelson | @John_M_Nelson | AdventuresInMapping.wordpress.com
Drought Data | U.S. Drought Monitor | droughtmonitor.unl.edu
Terrain | Living Atlas of the World | livingatlas.arcgis.com "terrain"
States | Living Atlas of the World | livingatlas.arcgis.com "states"
Areas of the Shoshone mountains have spent 82% of the past five years in drought, over half of that at "exceptional" intensity.
Parts of the Pinaleno Mountains, AZ, have sustained drought over the entirety of the past five years.
The area surrounding Santa Rosa Lake, TX has experienced drought conditions for 84% of the past five years. Of that, 74% was of "exceptional" intensity.
Frequent exceptional drought
Infrequent, though exceptional, drought
Frequent, though moderate, drought
color
size
Severity
Weighted proportion of drought intensity
Frequency
Proportion of time, over five years, spent in drought
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MI
IN
OH
KY
TN
MS
AL
GA
FL
SC
NC
VA
WV
PA
NY
VT
NH
ME
MA
RI
CT
NJ
MD
DC
DE

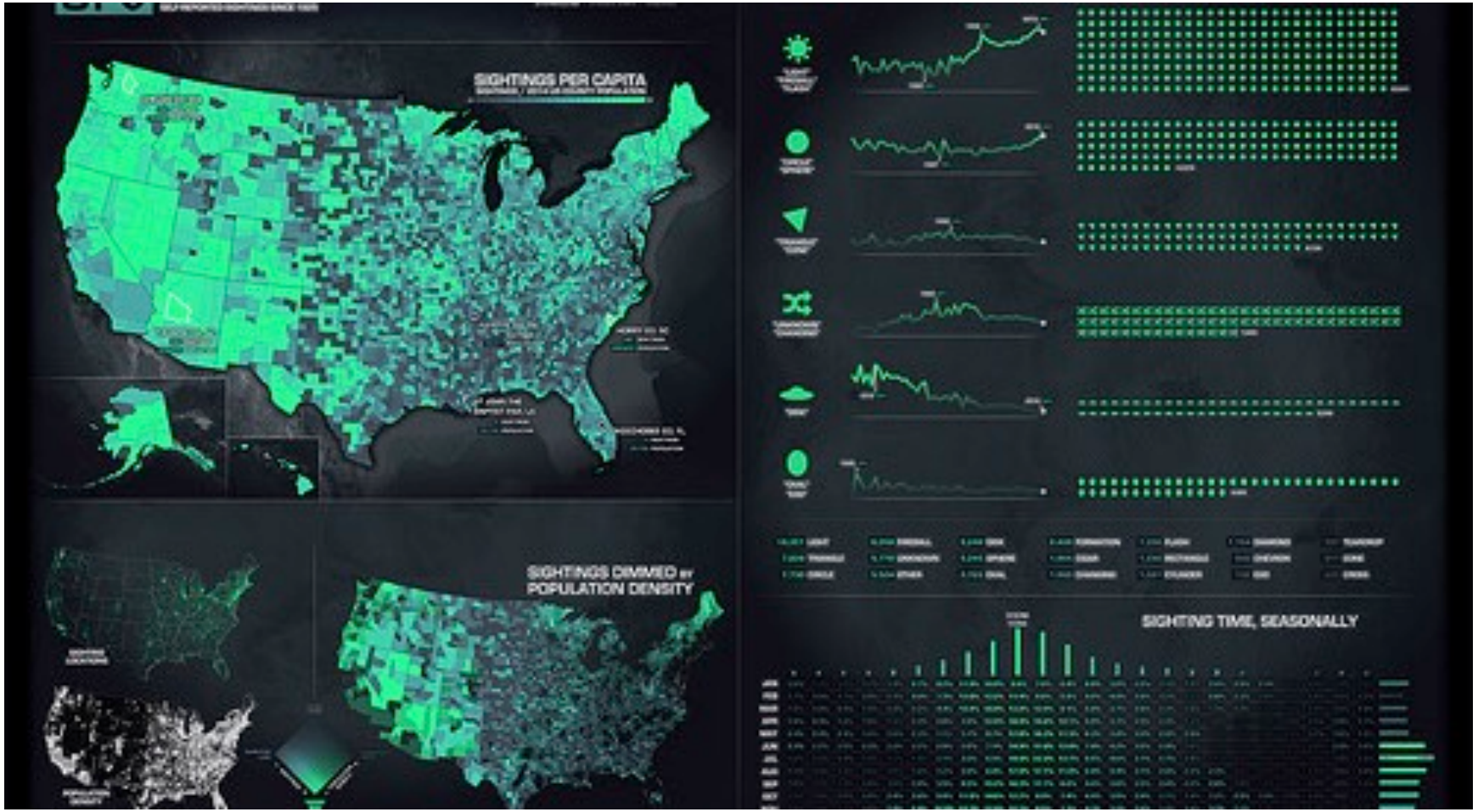
(very) Interactive visualizations

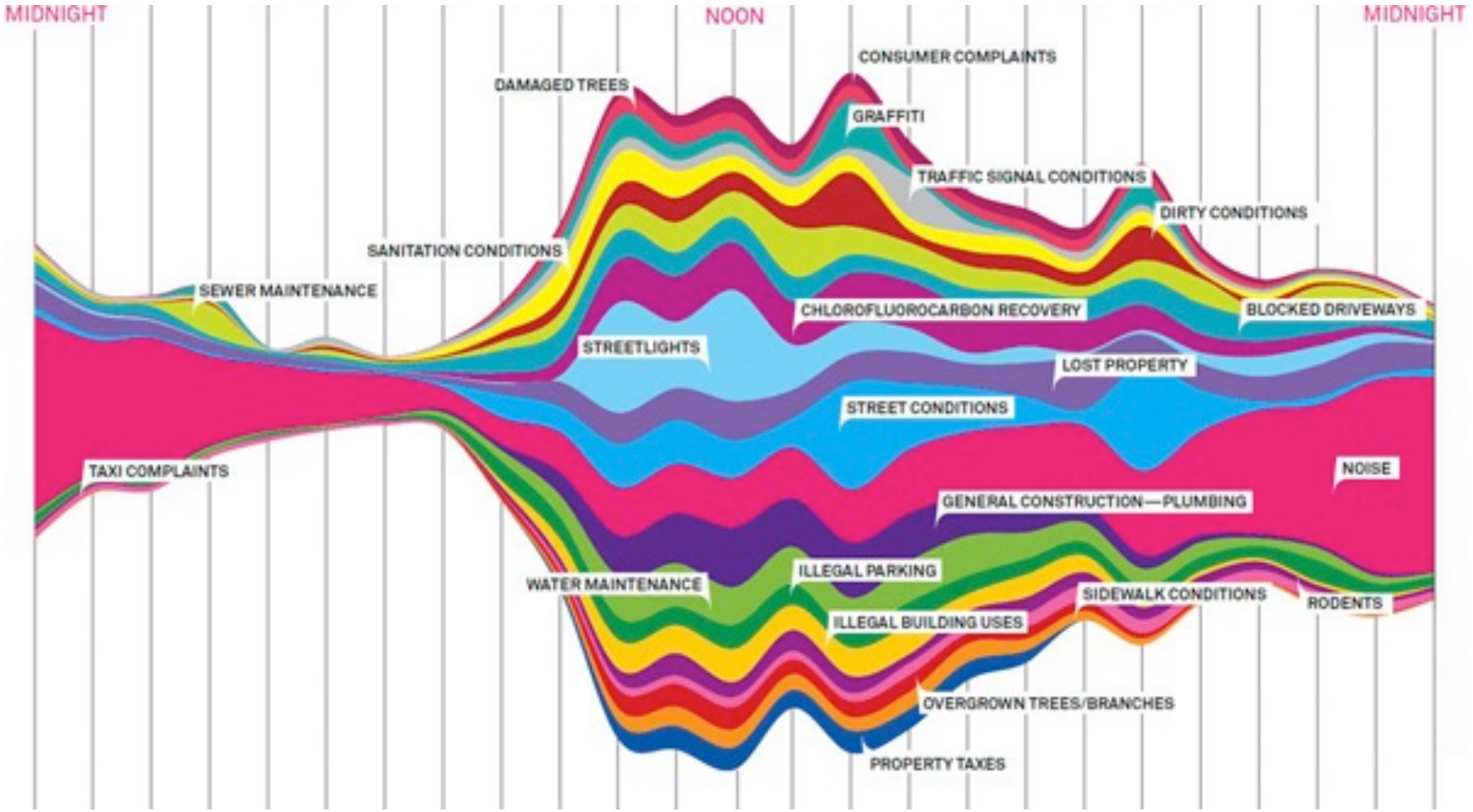




http://www.onezoom.org


GLOBAL THREATS 2015
Very concerned about...
Global climate change
Islamic State of Iraq and Syria
Global economic instability
Iran's nuclear program
Cyber-attacks
Tensions with Russia
Territorial disputes with China
● = Top concern in the country
Highlighted = Top concern is ISIS
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Burkina Faso
Brazil
Peru
Uganda
India
Philippines
Ghana
Nigeria
Chile
Venezuela
Ethiopia
Kenya
Argentina
Mexico
South Africa
Malaysia
Turkey
Pakistan
China
Lebanon
Spain
South Korea
Japan
France
Germany
Italy
Australia
U.S.
U.K.
Indonesia
Jordan
Canada
Palest. ter.
Venezuela
Senegal
Tanzania
Russia
Malaysia
Israel
Ukraine
Poland
Vietnam
No data
Average: 48%
45%
42%
32%
31%
26%
21%
CC-BY-SA martingrandjean.ch / data from Pew Research Center 2015 pewglobal.org

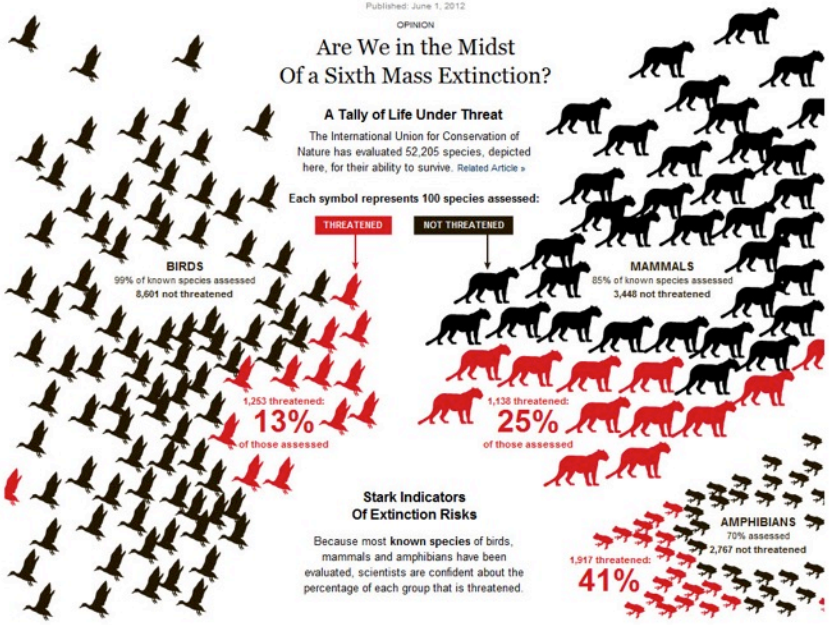

SIGHTINGS PER CAPITA
SIGHTINGS DIMMED BY POPULATION DENSITY
SIGHTING TIME, SEASONALLY

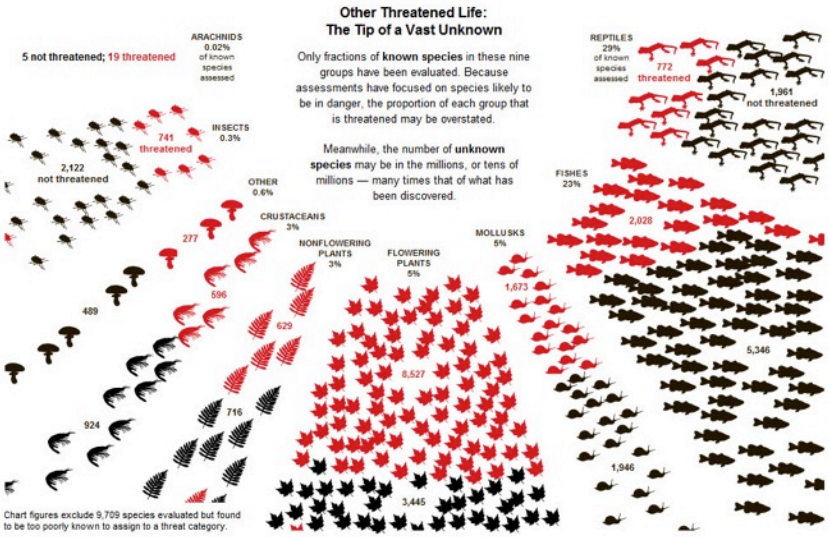

MIDNIGHT
NOON
MIDNIGHT
CONSUMER COMPLAINTS
DAMAGED TREES
GRAFFITI
TRAFFIC SIGNAL CONDITIONS
DIRTY CONDITIONS
SANITATION CONDITIONS
SEWER MAINTENANCE
CHLOROFLUOROCARBON RECOVERY
BLOCKED DRIVEWAYS
STREETLIGHTS
LOST PROPERTY
STREET CONDITIONS
TAXI COMPLAINTS
NOISE
GENERAL CONSTRUCTION—PLUMBING
ILLEGAL PARKING
WATER MAINTENANCE
SIDEWALK CONDITIONS
RODENTS
ILLEGAL BUILDING USES
OVERGROWN TREES/BRANCHES
PROPERTY TAXES

Published: June 1, 2012

OPINION

Are We in the Midst Of a Sixth Mass Extinction?

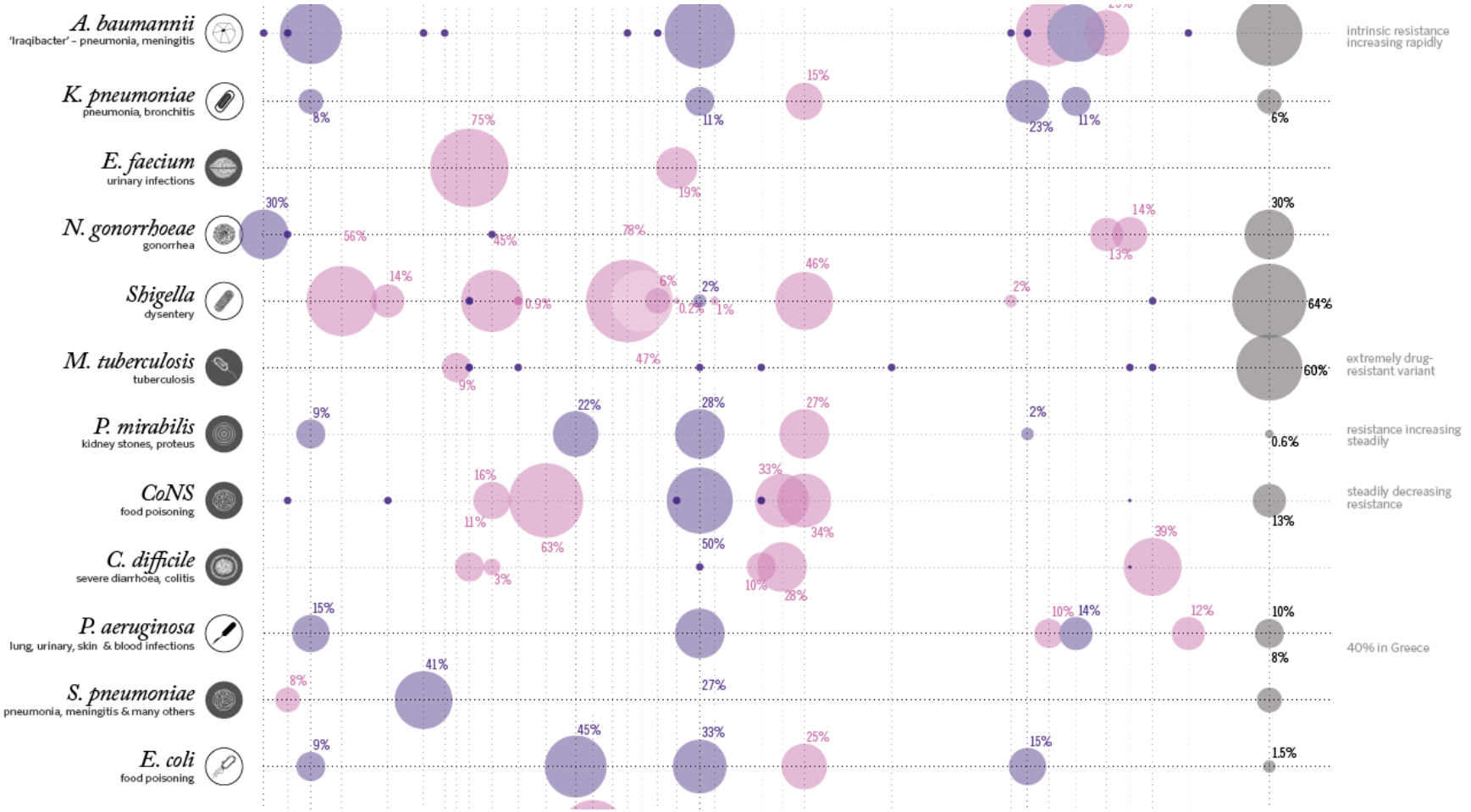




Already Gone	Mollusks	Birds	Flowering plants	Mammals	Fishes	Insects	Amphibians	Reptiles	Crustaceans	Nonflowering plants	Others	No known arachnid extinctions.
Species known to be extinct, or extinct in the wild, since 1500:	327	136	110	79	68	60	39	22	12	10	2	

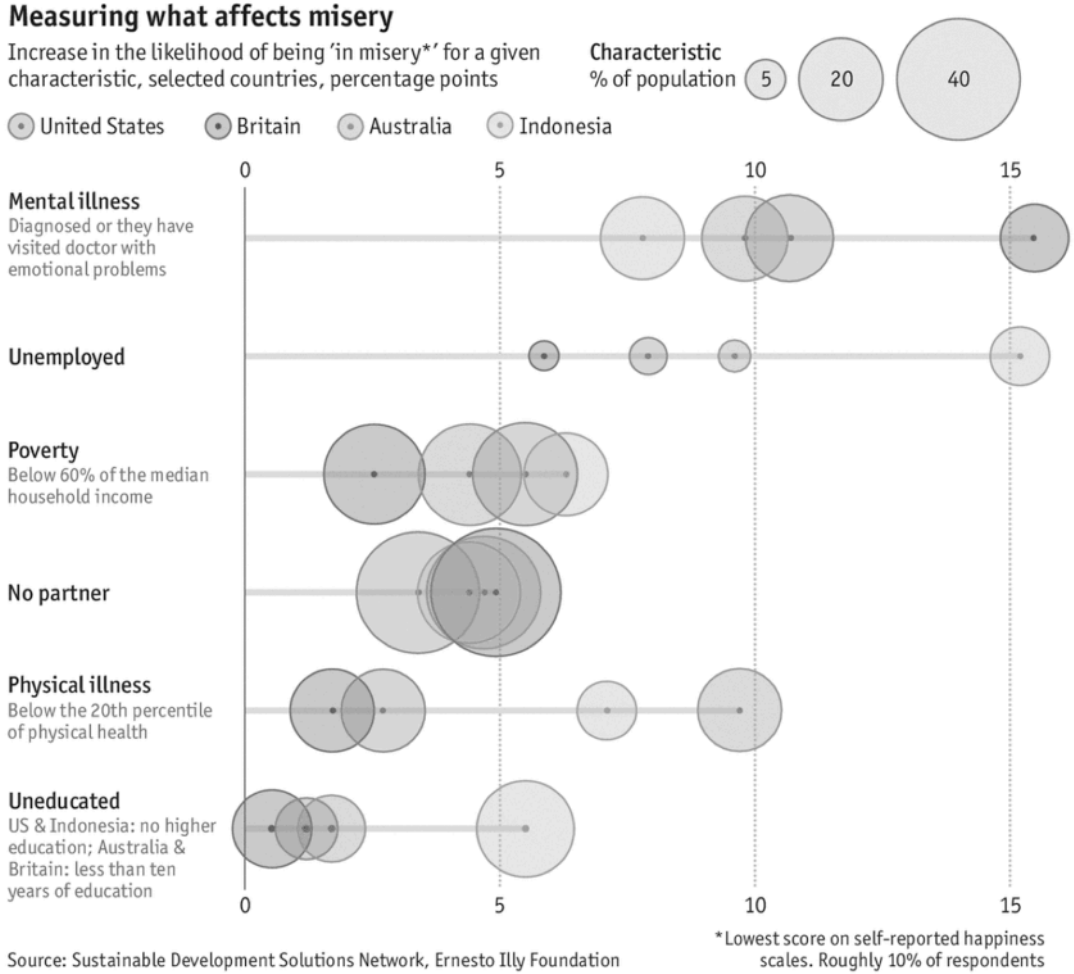
By BILL MARSH | Send Feedback

Source: International Union for Conservation of Nature


A. baumannii
'Iraqibacter' – pneumonia, meningitis
K. pneumoniae
pneumonia, bronchitis
E. faecium
urinary infections
N. gonorrhoeae
gonorrhea
Shigella
dysentery
M. tuberculosis
tuberculosis
P. mirabilis
kidney stones, proteus
CoNS
food poisoning
C. difficile
severe diarrhoea, colitis
P. aeruginosa
lung, urinary, skin & blood infections
S. pneumoniae
pneumonia, meningitis & many others
E. coli
food poisoning
intrinsic resistance increasing rapidly
extremely drug-resistant variant
resistance increasing steadily
steadily decreasing resistance
40% in Greece

Measuring what affects misery

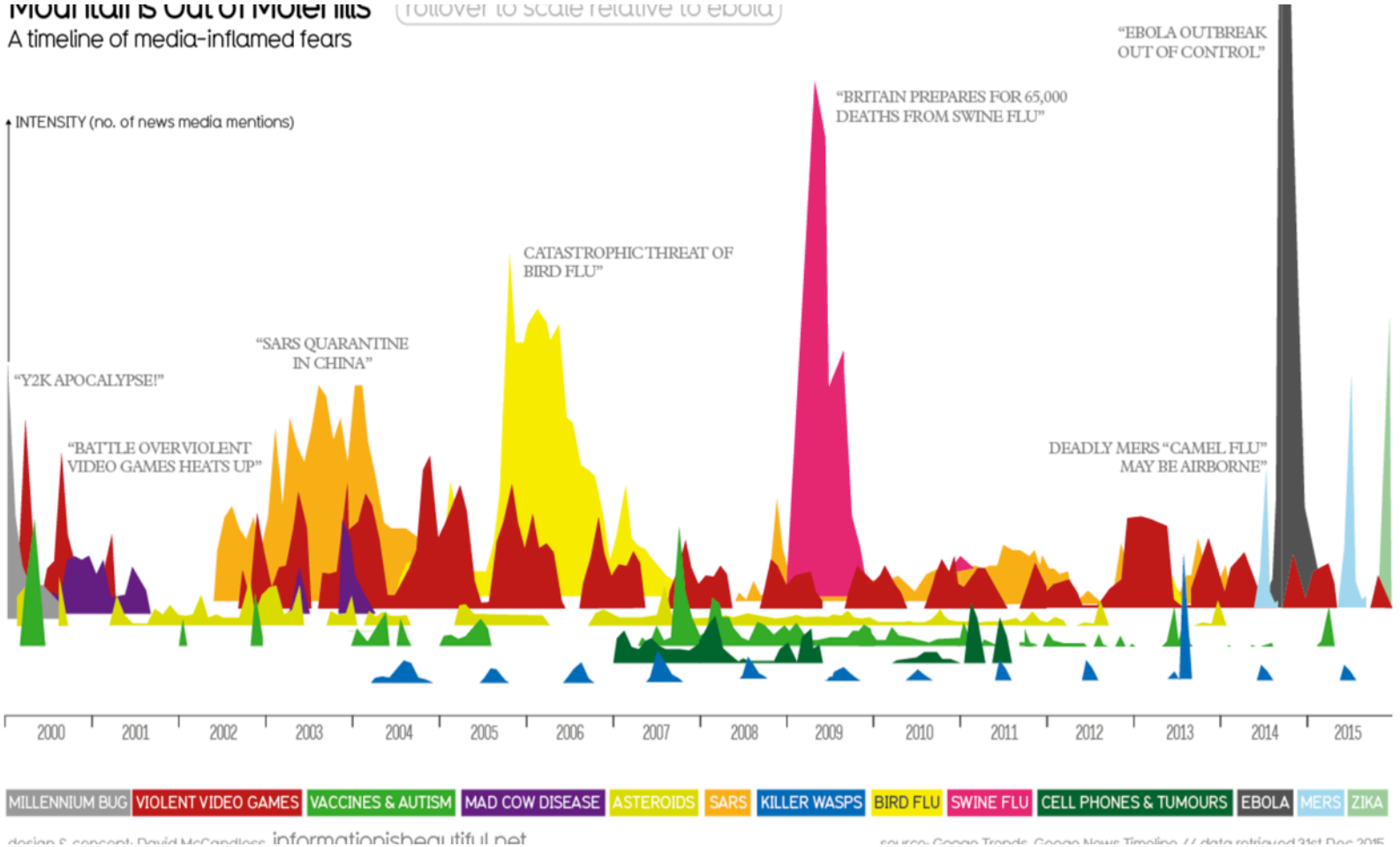
Increase in the likelihood of being 'in misery*' for a given characteristic, selected countries, percentage points



Source: Sustainable Development Solutions Network, Ernesto Illy Foundation

*Lowest score on self-reported happiness scales. Roughly 10% of respondents

Economist.com


Mountain is Out of Molehills
(rollover to scale relative to ebola)
A timeline of media-inflamed fears
INTENSITY (no. of news media mentions)
"Y2K APOCALYPSE!"
"BATTLE OVER VIOLENT VIDEO GAMES HEATS UP"
"SARS QUARANTINE IN CHINA"
CATASTROPHIC THREAT OF BIRD FLU"
"BRITAIN PREPARES FOR 65,000 DEATHS FROM SWINE FLU"
DEADLY MERS "CAMEL FLU" MAY BE AIRBORNE"
"EBOLA OUTBREAK OUT OF CONTROL"
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
MILLENNIUM BUG
VIOLENT VIDEO GAMES
VACCINES & AUTISM
MAD COW DISEASE
ASTEROIDS
SARS
KILLER WASPS
BIRD FLU
SWINE FLU
CELL PHONES & TUMOURS
EBOLA
MERS
ZIKA
design & concept: David McCandless
informationisbeautiful.net
source: Google Trends, Google News Timeline // data retrieved 31st Dec 2015

International Number Ones

Because every country is the best at something

according to data

2016 edition

David McCandless // v 2.1 / Oct 2016
Research: David Mccandless, Stephanie Smith, Esther Kersley
Additional Design: Fabio Bergamaschi

InformationIsBeautiful.net

Sources: CIA, NYTimes, Bloomberg, UN, The Economist, World Bank, Reuters, BBC, Forbes, The Guardian, WSJ & others. // data: http://bit.ly/IIB_IntNoOnes
mostly per capita / % of population data // very small countries excluded

Recommended Reading

- ***Visual Thinking for Design*** by Colin Ware
- ***Visualization Analysis & Design*** *by Tamara Munzner (available for use digitally at UCT from Ebsco books)*
- ***Information visualization: perception for design*** by Colin Ware
- ***The Visual Display of Quantitative Information.*** by Edward R. Tufte (second edition)
- ***Visual Complexity. Mapping Patterns of Information.*** by Manuel Lima (Princeton Architectural Press, New York)
- ***Visual Language for Designers. Principles for creating graphics that people understand.*** by Connie Malamed

all very beautiful books, most are in the UCT library.