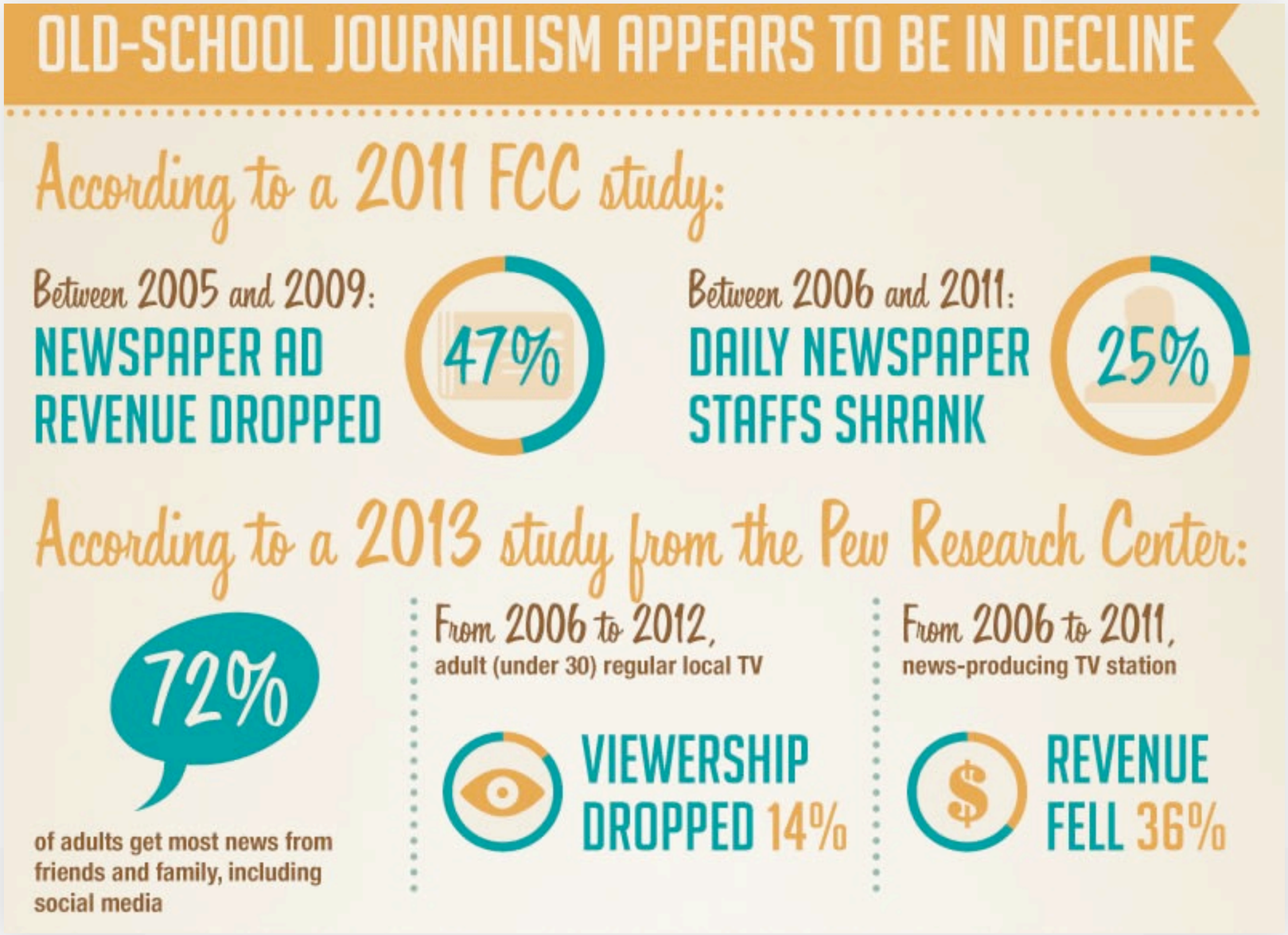
IO Lab: Information & Data Visualization

October 21, 2013

INFO 290TA (Information Organization Lab)
Kate Rushton & Raymon Sutedjo-The

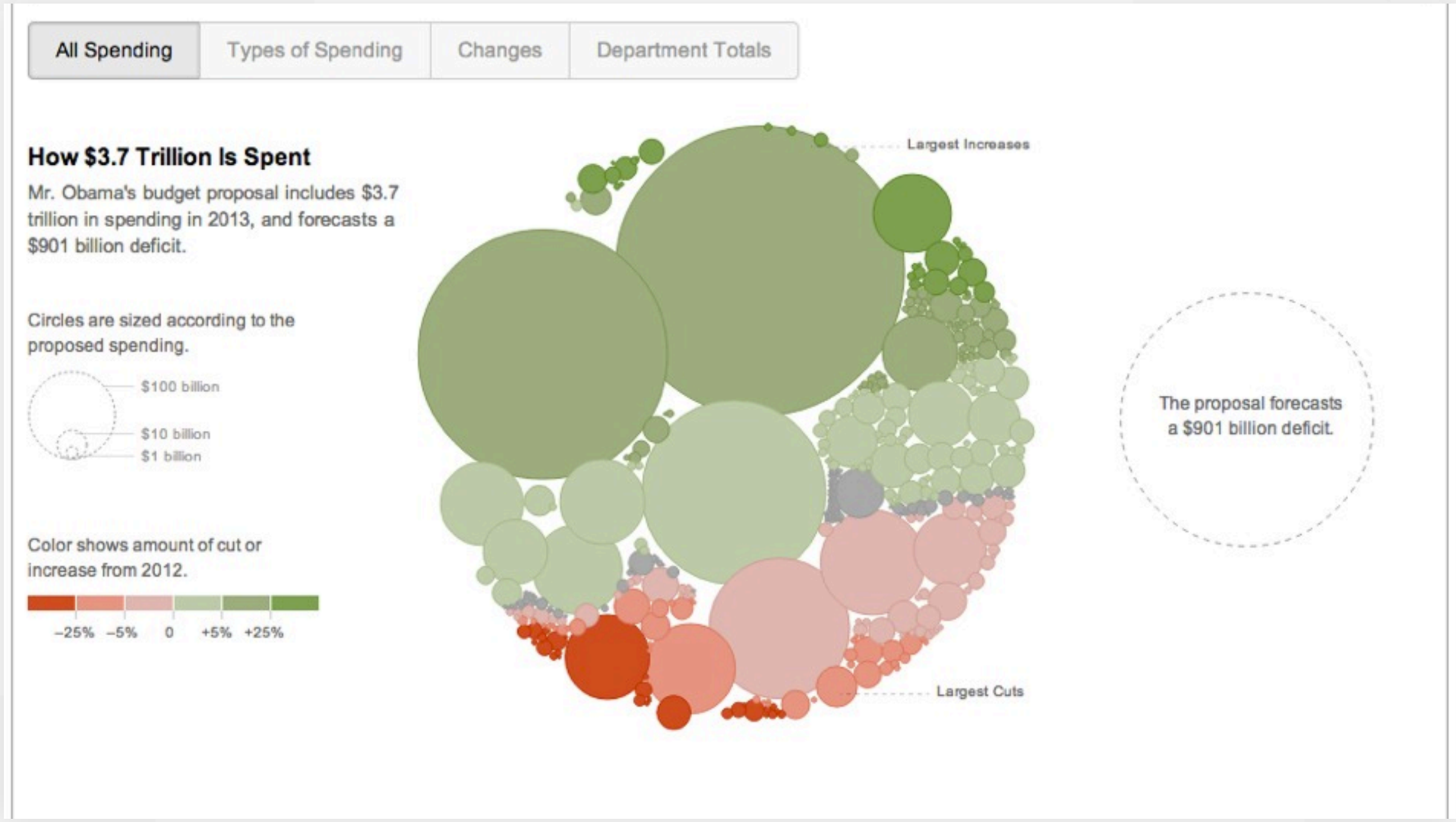
Let’s Start with Some Examples

Traditional Journalism



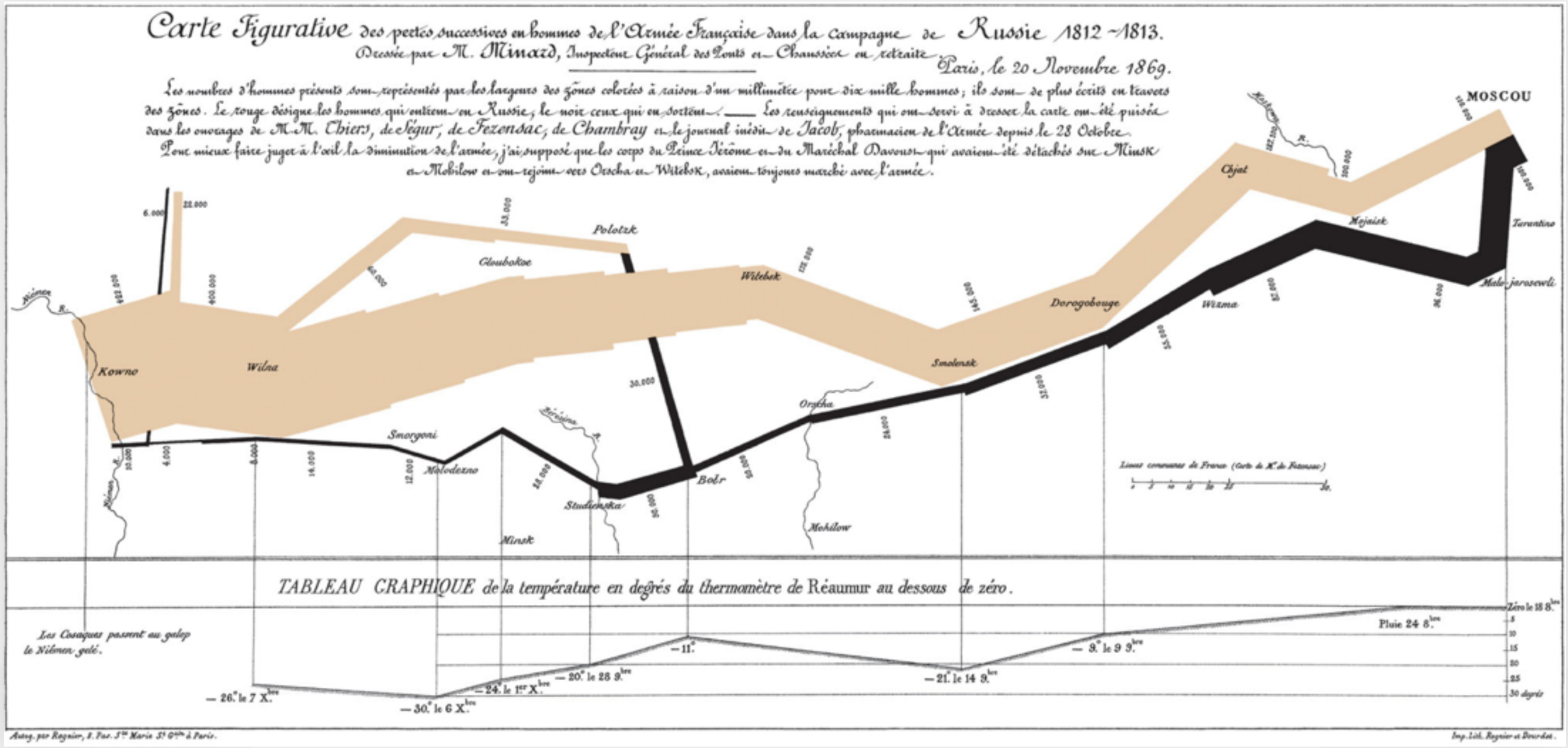
http://visual.ly/traditional-journalism-it-old-news

2013 Budget Proposal



http://www.nytimes.com/interactive/2012/02/13/us/politics/2013-budget-proposal-graphic.html

Napoleon's Invasion of Russia in 1812



http://innovatobase.wordpress.com/2011/12/07/how-to-effectively-express-complex-information/

Why Visualizations?

Why Visualizations?

http://25.media.tumblr.com/tumblr_losdz4Qnjz1qexogfo1_500.jpg

Why Visualizations?

Make data visible

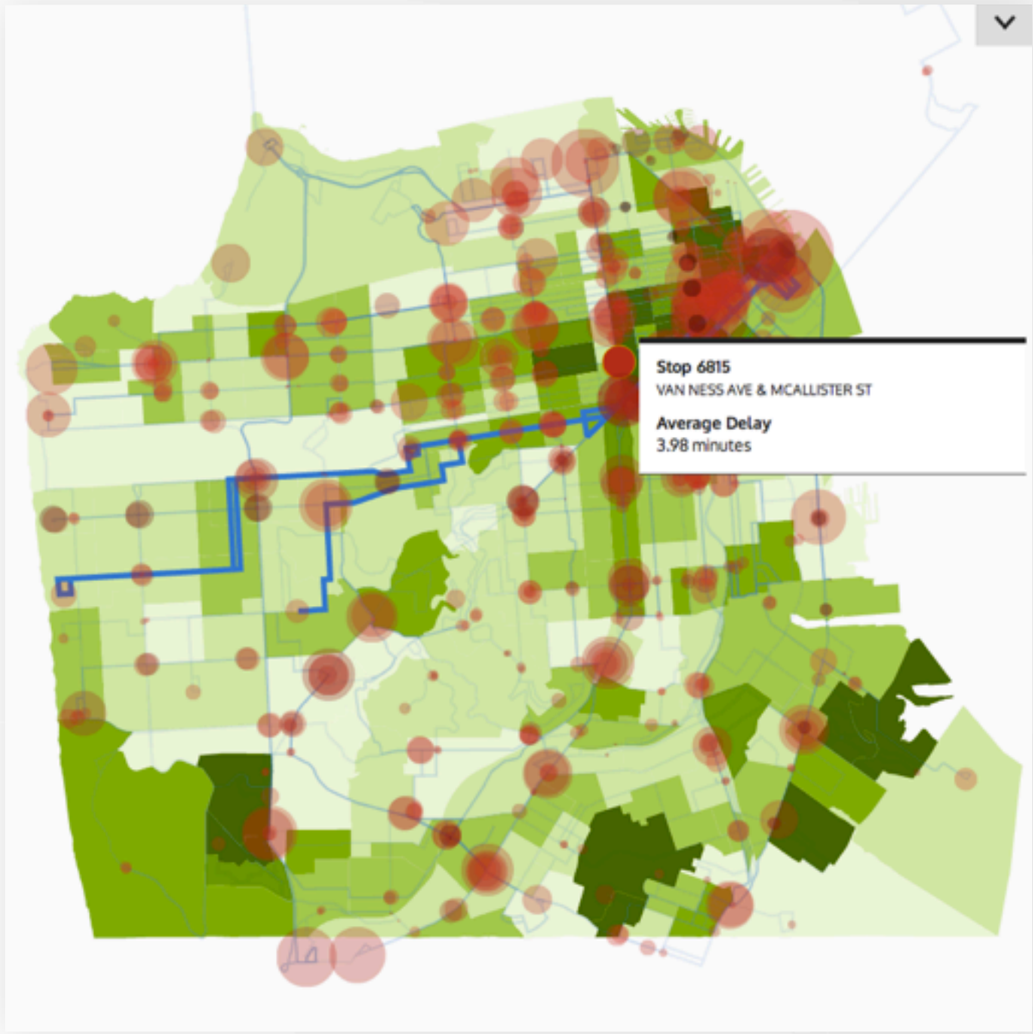
Reveal patterns & insights

Help with “data-driven decisions”

Why Visualizations?

GEO.id	GEO.id2	GEO.display-	HC01_EST_V	HC01_MOE_	HC02_EST_V	HC02_MOE_
1400000US0	6075010100	Census Tract	2204	149	60584	15156
1400000US0	6075010200	Census Tract	2653	197	96370	13748
1400000US0	6075010300	Census Tract	1952	157	99531	30144
1400000US0	6075010400	Census Tract	2477	159	89837	26659
1400000US0	6075010500	Census Tract	1570	143	113281	10420
1400000US0	6075010600	Census Tract	1856	172	26250	6946
1400000US0	6075010700	Census Tract	2716	158	24626	5252
1400000US0	6075010800	Census Tract	2458	173	79932	7879
1400000US0	6075010900	Census Tract	2617	164	104205	15139
1400000US0	6075011000	Census Tract	2576	160	56143	16917
1400000US0	6075011100	Census Tract	2865	180	60420	7718
1400000US0	6075011200	Census Tract	1667	156	73265	11633
1400000US0	6075011300	Census Tract	1595	78	19529	3910
1400000US0	6075011700	Census Tract	918	127	27311	9680
1400000US0	6075011800	Census Tract	700	55	25462	2648
1400000US0	6075011901	Census Tract	1465	219	54059	13267
1400000US0	6075011902	Census Tract	1587	145	47012	7466
1400000US0	6075012000	Census Tract	2397	176	36450	2114
1400000US0	6075012100	Census Tract	2043	175	36328	14265
1400000US0	6075012201	Census Tract	2584	180	32250	7611
1400000US0	6075012202	Census Tract	1546	113	27000	15088
1400000US0	6075012301	Census Tract	1307	147	12591	2894
1400000US0	6075012302	Census Tract	1666	162	25119	7898
1400000US0	6075012401	Census Tract	2206	210	21306	1129
1400000US0	6075012402	Census Tract	1886	165	21707	9802
1400000US0	6075012501	Census Tract	2022	214	11634	704
1400000US0	6075012502	Census Tract	2204	194	12370	1068
1400000US0	6075012601	Census Tract	1419	114	93787	12981
1400000US0	6075012602	Census Tract	1639	168	129716	19704
1400000US0	6075012700	Census Tract	2007	148	89491	22115
1400000US0	6075012800	Census Tract	2230	167	108673	7494
1400000US0	6075012901	Census Tract	1519	118	92820	11958
1400000US0	6075012902	Census Tract	2064	121	103000	10230
1400000US0	6075013000	Census Tract	2340	199	97708	25525
1400000US0	6075013101	Census Tract	2407	161	125066	26758
1400000US0	6075013102	Census Tract	1579	117	99080	6997
1400000US0	6075013200	Census Tract	2392	164	136190	31335

VS



http://ray-mon.com/urbandatachallenge

Two Approaches

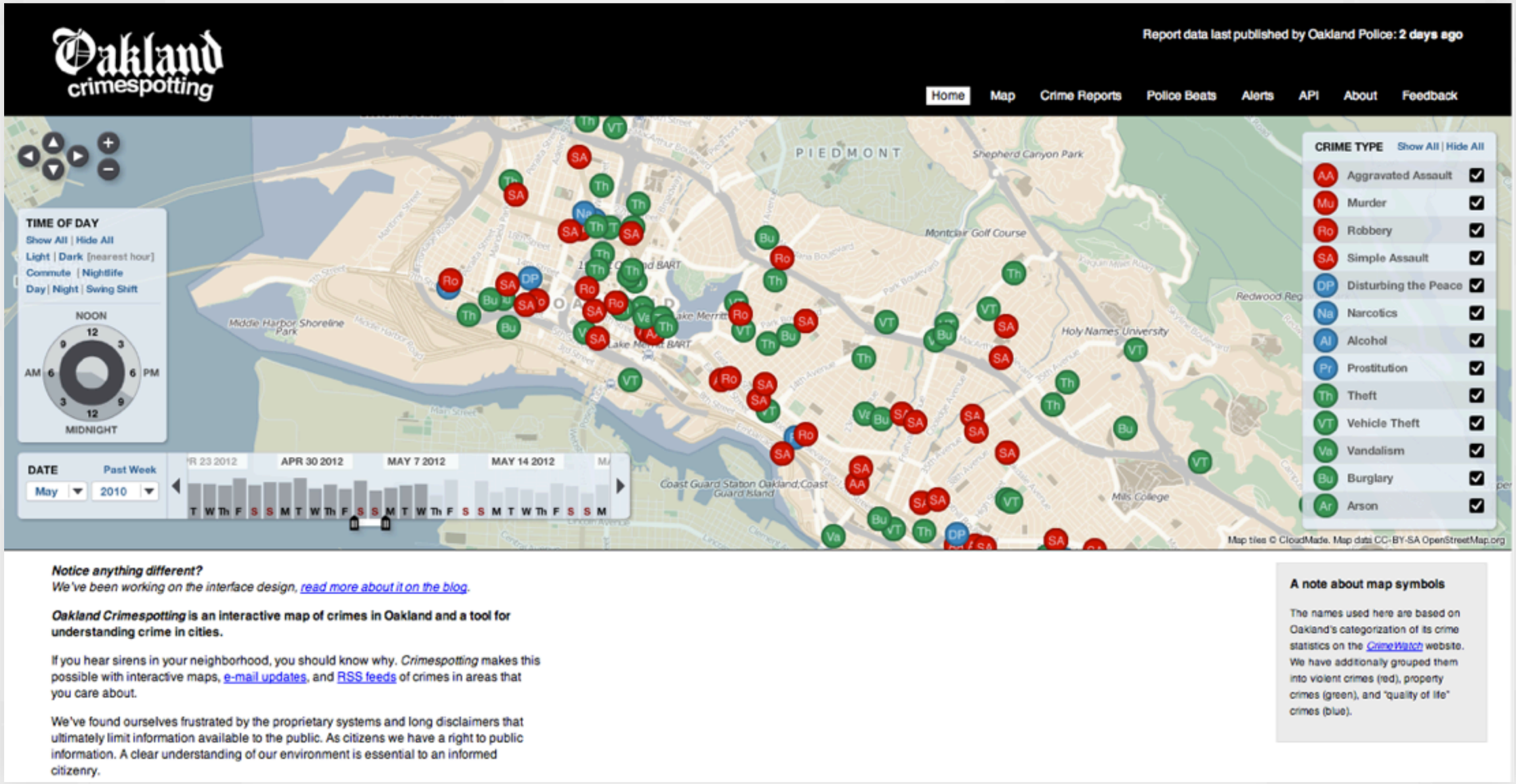
Explore

Analyze

Explain

Communicate

Analyze



Notice anything different?
We've been working on the interface design, read more about it on the blog.

***Oakland Crimespotting* is an interactive map of crimes in Oakland and a tool for understanding crime in cities.**

If you hear sirens in your neighborhood, you should know why. *Crimespotting* makes this possible with interactive maps, e-mail updates, and RSS feeds of crimes in areas that you care about.

We've found ourselves frustrated by the proprietary systems and long disclaimers that ultimately limit information available to the public. As citizens we have a right to public information. A clear understanding of our environment is essential to an informed citizenry.

A note about map symbols

The names used here are based on Oakland's categorization of its crime statistics on the CrimeWatch website. We have additionally grouped them into violent crimes (red), property crimes (green), and "quality of life" crimes (blue).

http://oakland.crimespotting.org/

Communicate

http://visual.ly/then-vs-now-small-businesses

Turning Data into Visualization

Data Types

1D (sets, sequences, text)

2D (maps)

3D (shapes)

Temporal

Multidimensional (relations)

Tree (hierarchies)

Network (graphs)

http://www.mat.ucsb.edu/~g.legrady/academic/courses/11w259/schneiderman.pdf

Data Taxonomy

NOMINAL

Action
Comedy
Drama
Documentary
Horror
Romantic

ORDINAL

Bronze
Silver
Gold

http://www.mpopa.ro/statistica_licenta/Stevens_Measurement.pdf

Data Taxonomy

QUANTITATIVE

Interval (+)
Temperature in °F
(10° F, 20° F, 30° F, ...)

The difference between 10° F & 20° F
is the same as 20° F & 30° F.

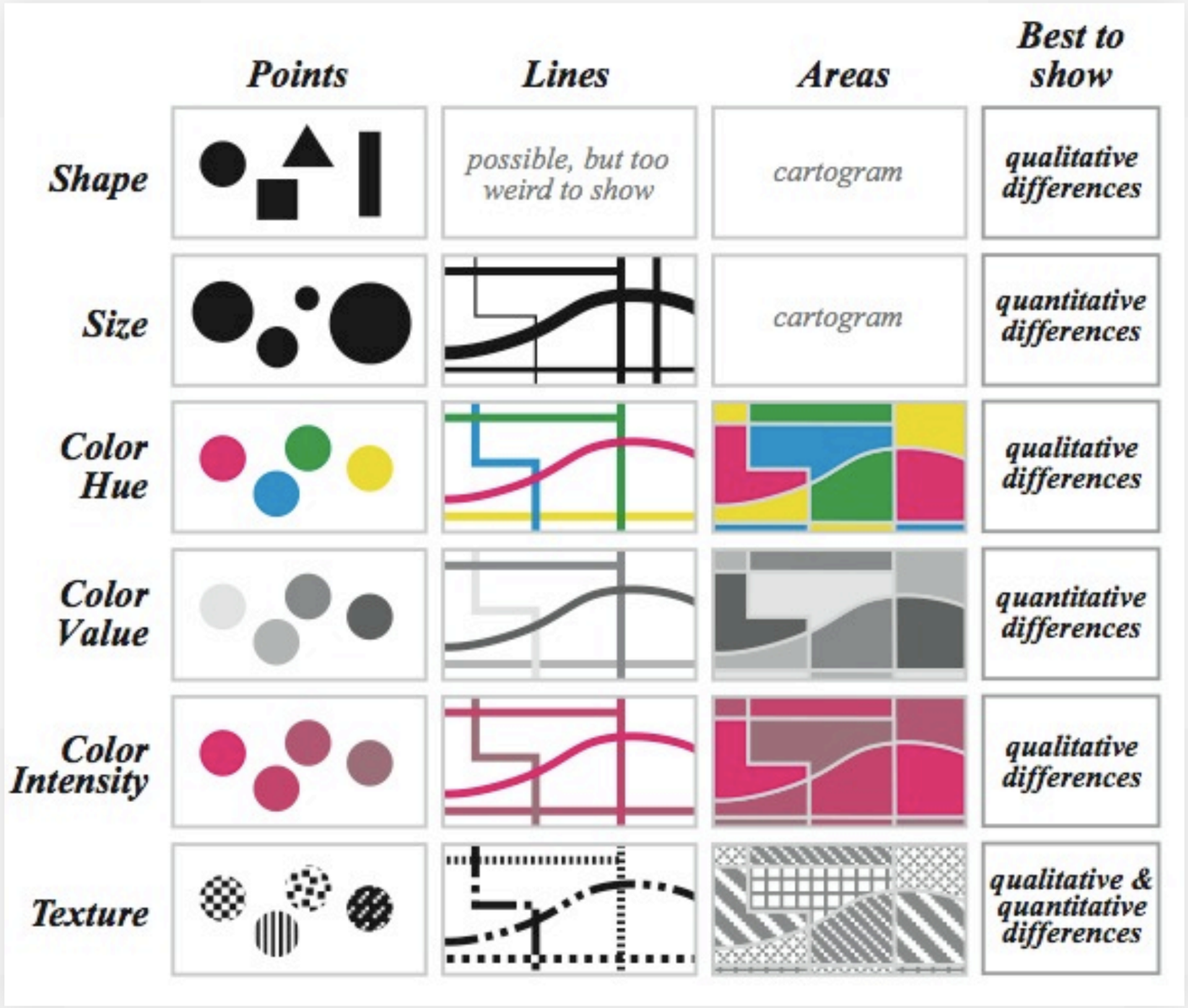
There is no “true zero.”
0° F != no temperature.
20° F != 2x as hot as 10° F.

Ratio (x)
Money
($0, $2000, $4000, ...)

“True zero.”
$0 = absence of money.
$4000 = 2x as much money as $2000.

http://www.mpopa.ro/statistica_licenta/Stevens_Measurement.pdf

Bertin's Semiology of Data



http://www.amazon.com/Semiology-Graphics-Diagrams-Networks-Maps/dp/1589482611

Multivariate Visualizations

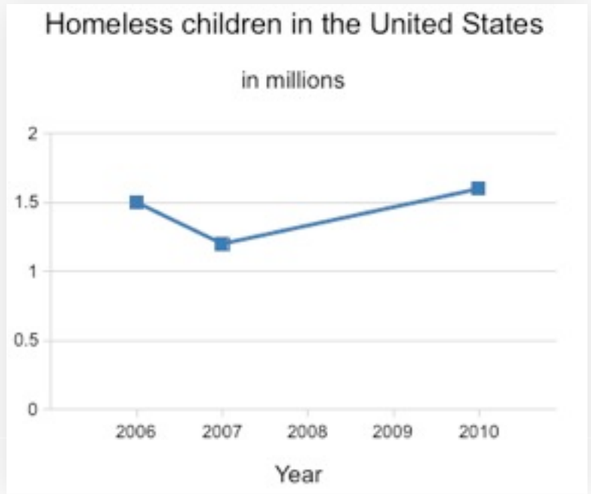
1D, 2D, 3D

1D

15.1%

of the general population lived in poverty (2010)

2D



3D

http://en.wikipedia.org/wiki/Poverty_in_the_United_States
http://www.washingtonpost.com/blogs/wonkblog/wp/2013/05/15/here-are-315-charts-on-poverty-and-inequality-satisfied-yet/

nD

http://michaelporath.com/projects/manifest-destiny/

Human Visual Perception

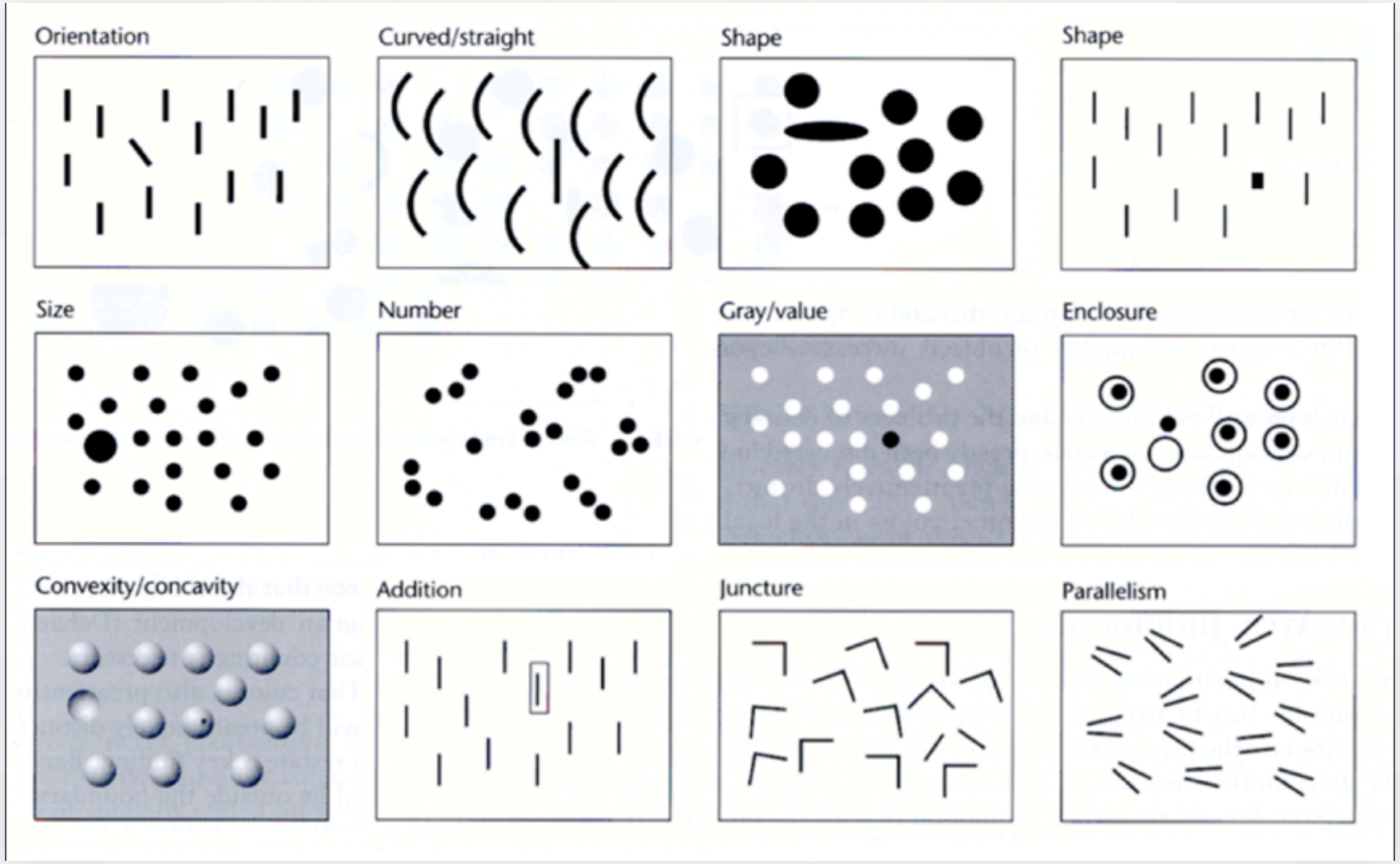
How Many 3s?

1239876245289348924187234784387939487098579485786625356123743456724782128745954194156239906

How Many 3s?

12398762452893489241
87234784387939487098
57948578662535612374
34567247821287459541
9415623906

Pre-Attentive Features



http://www.csc.ncsu.edu/faculty/healey/PP/index.html

Gestalt Principles

A theory of visual perception that describes how the human eyes tend to **organize visual elements into groups or unified wholes** when certain principles are applied.

http://www.scholarpedia.org/article/Gestalt_principles

Gestalt Principles

Proximity

Similarity

Continuity

Closure

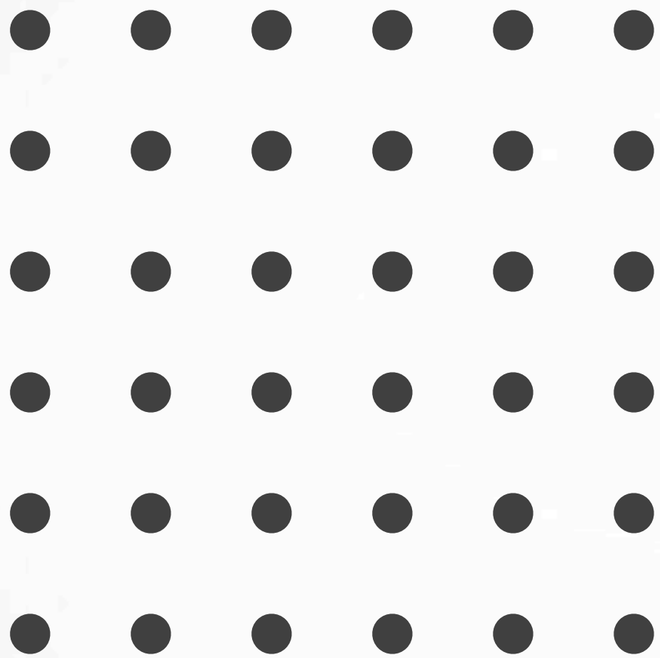
Smallness (figure-ground)

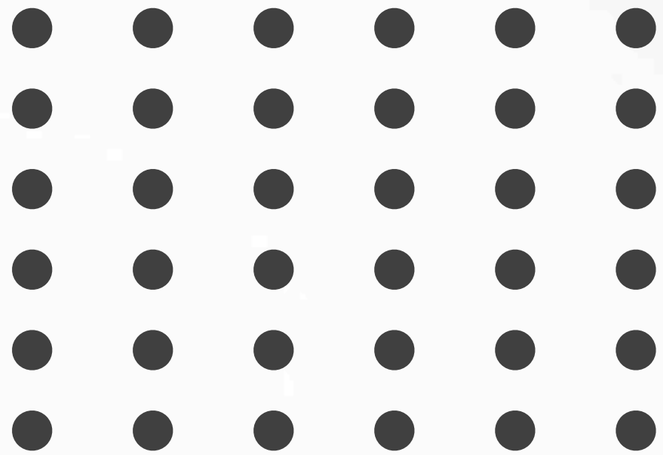
Symmetry

Surroundedness

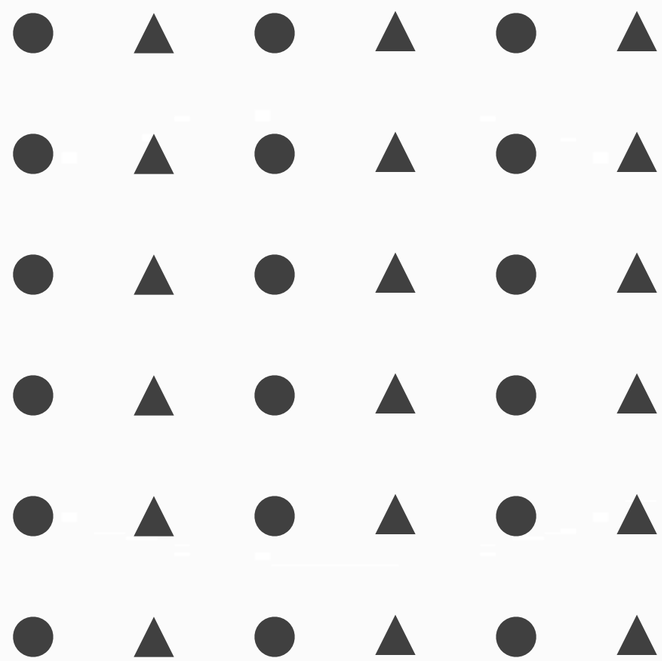
http://www.artinarch.com/vp12.html

Gestalt Principles: Proximity

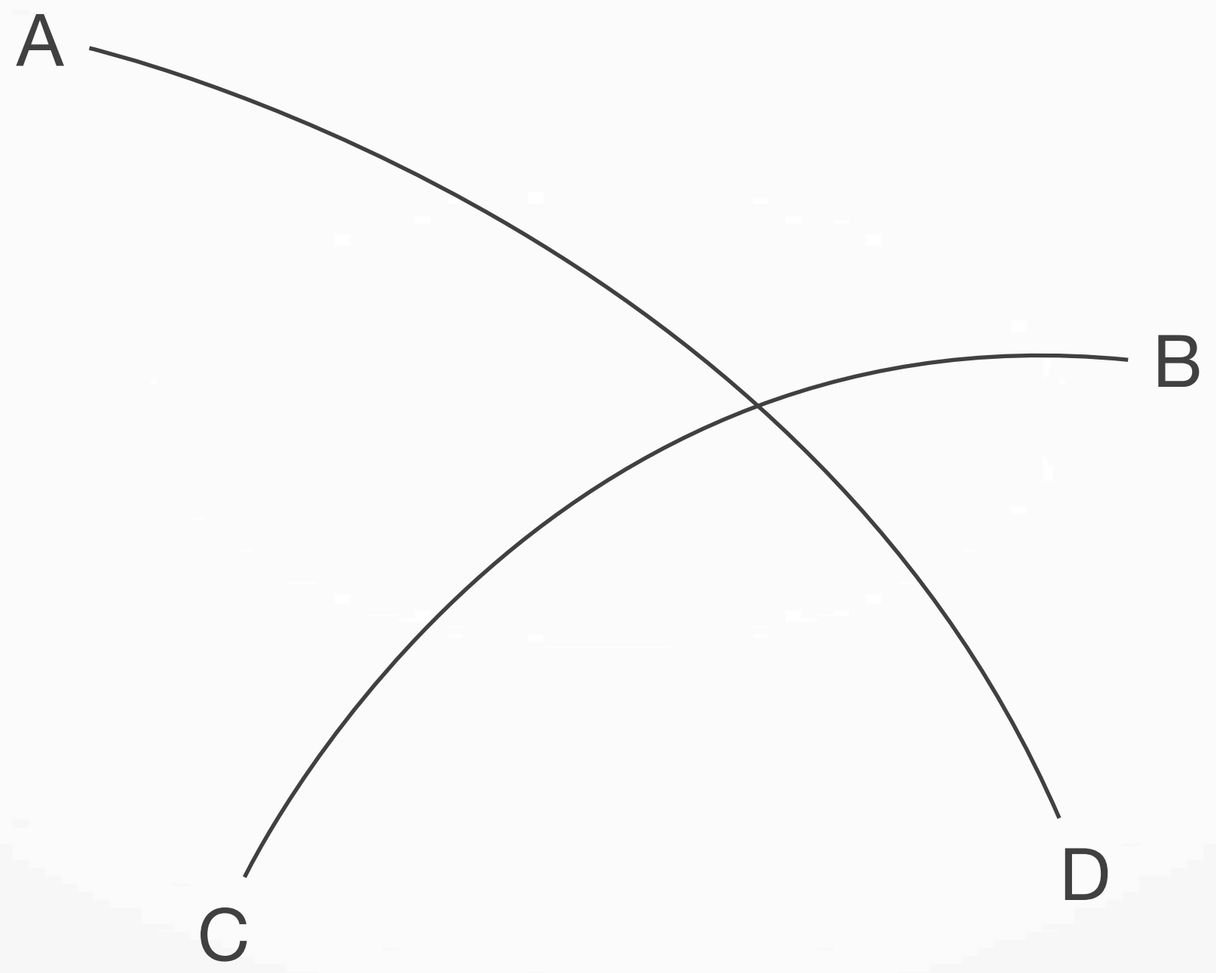




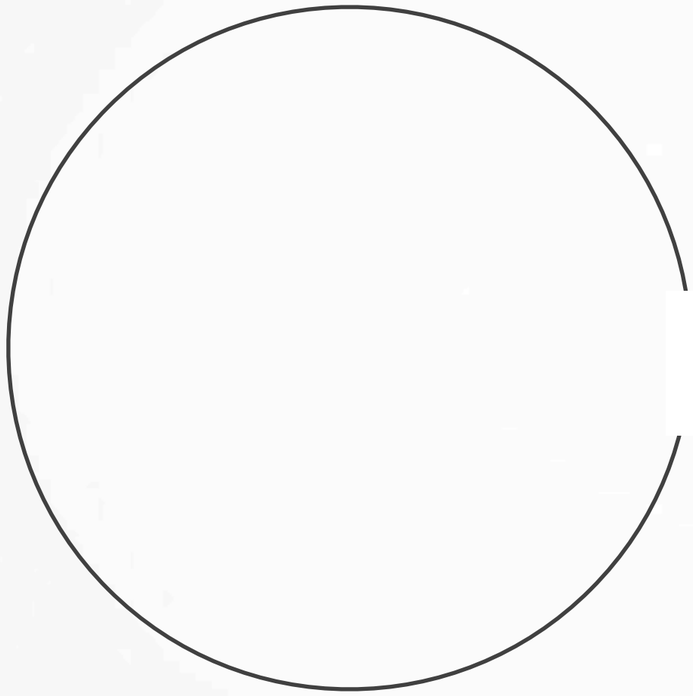
Gestalt Principles: Similarity



Gestalt Principles: Continuity

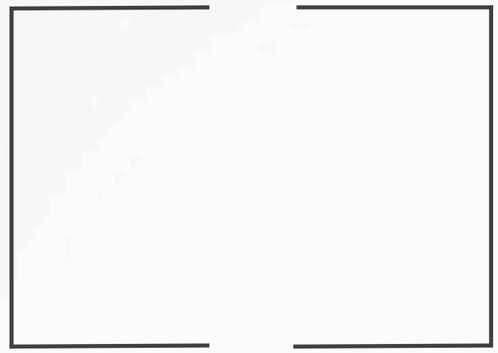


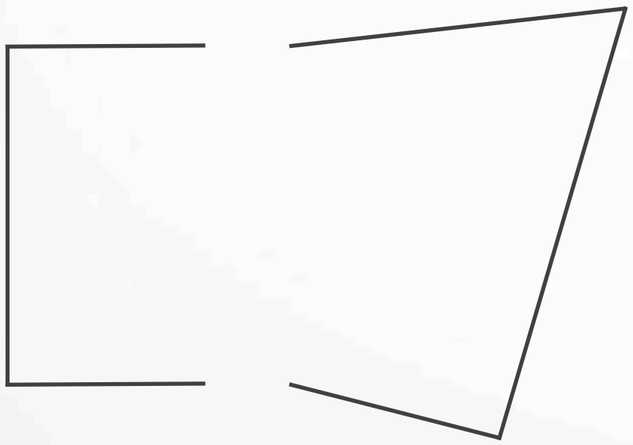
Gestalt Principles: Closure



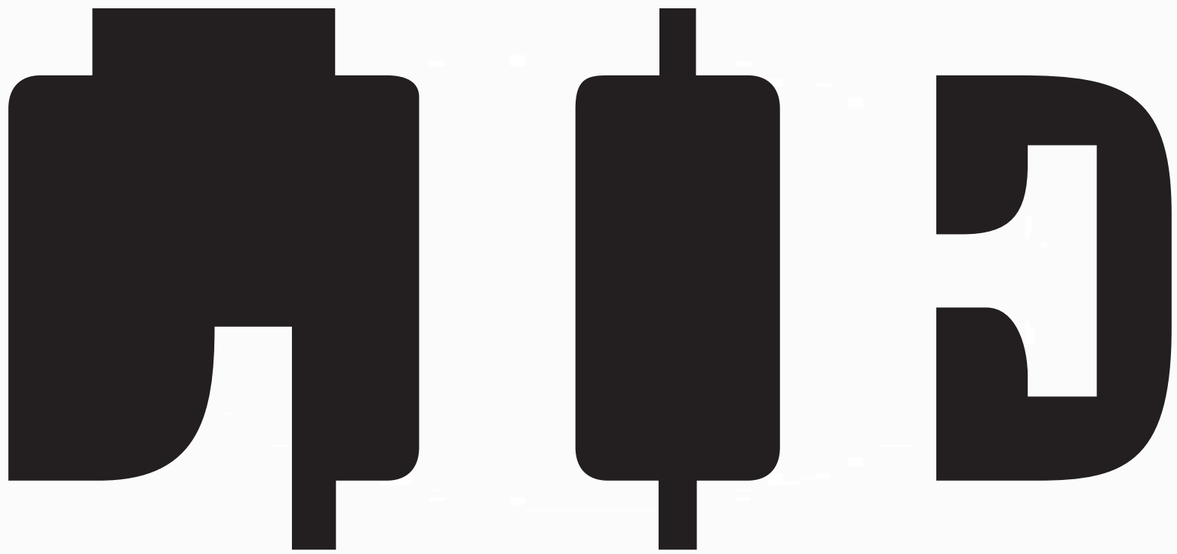
Gestalt Principles: Smallness (Figure-Ground)

Gestalt Principles: Symmetry





Gestalt Principles: Surroundedness



Gestalt Principles: Surroundedness

Storytelling

Why Storytelling?

Structure your data & information

... in an order that makes sense

... to reduce information overload

... and guide the audience

... so the information can be more easily understood

Narrative Structure

PROLOGUE

Act 1

Act 2

...

EPILOGUE

Narrative Structure

Explore

Analyze

Explain

Communicate

reader-driven
no prescribed ordering
no messaging
free interactivity

author-driven
ordering of information
heavy messaging
little interactivity

Elements of Storytelling

VISUAL STRUCTURE

Support
the story

MESSAGING

Tell
the story

INTERACTIVITY

Engage
the story

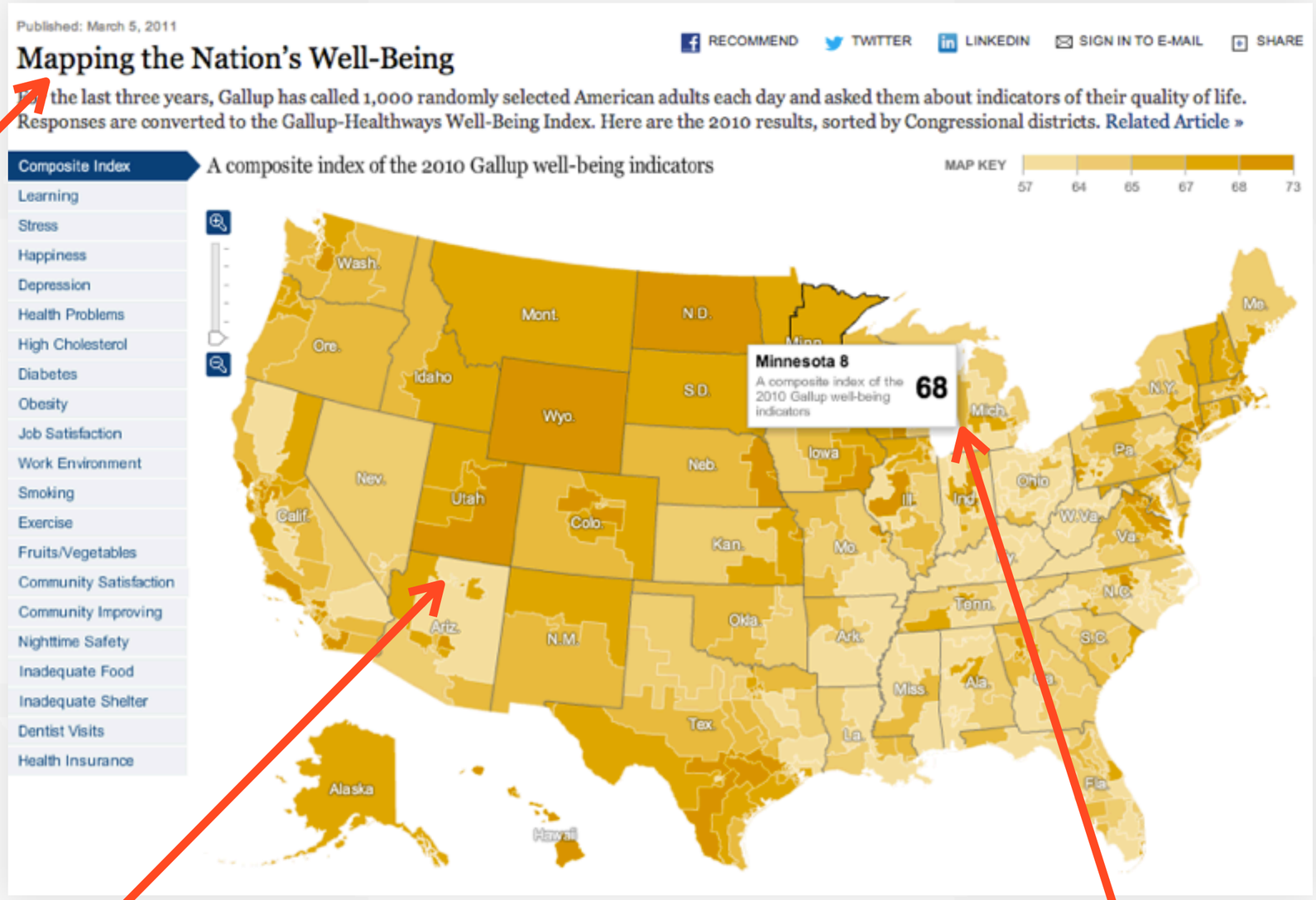
Visual Structure

http://vis.stanford.edu/papers/narrative

Messaging

http://perceptionvsfact.com/1vz9

Interactivity



headline & caption

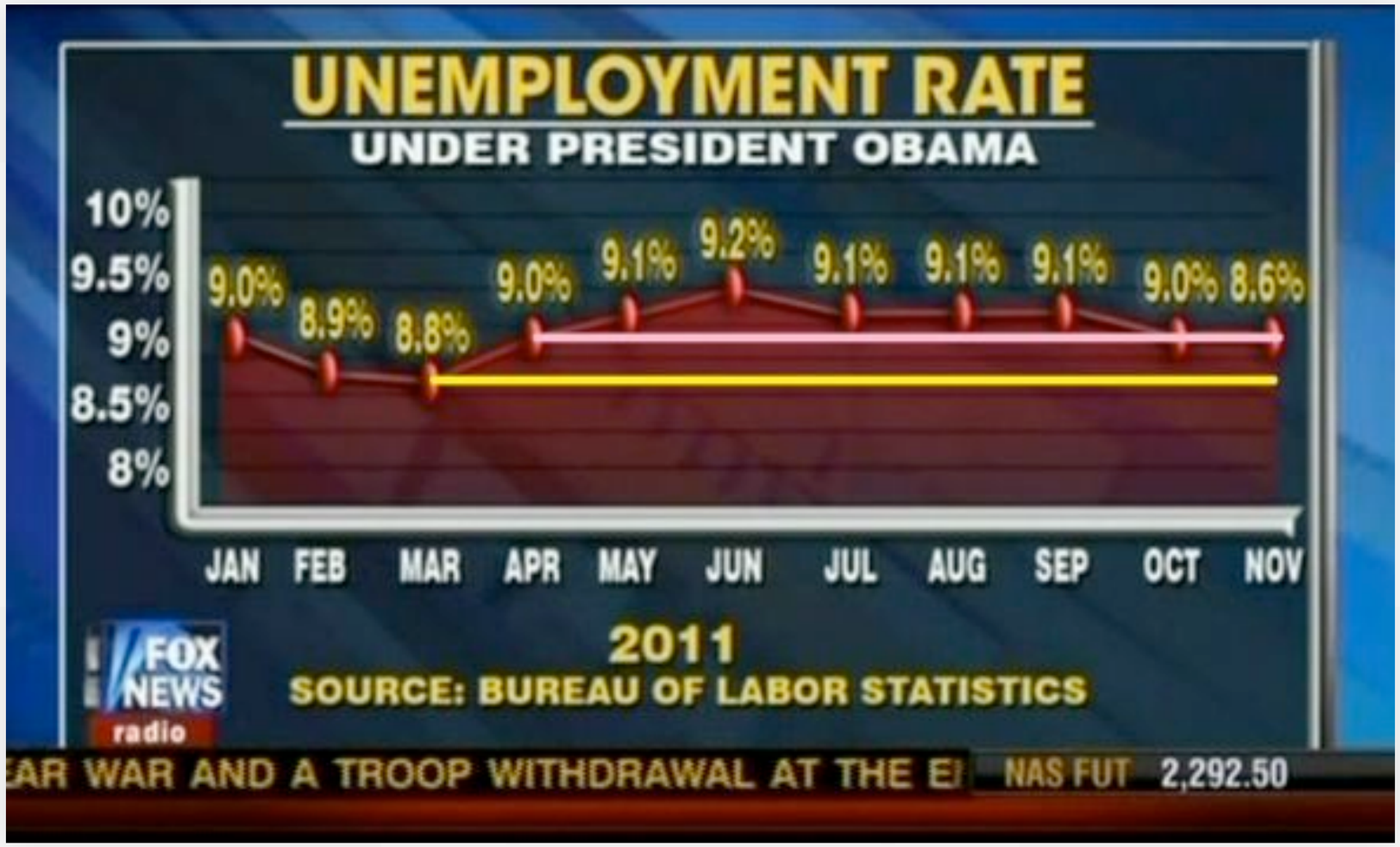
main visualization

annotation

http://www.nytimes.com/interactive/2011/03/06/weekinreview/20110306-happiness.html?_r=0

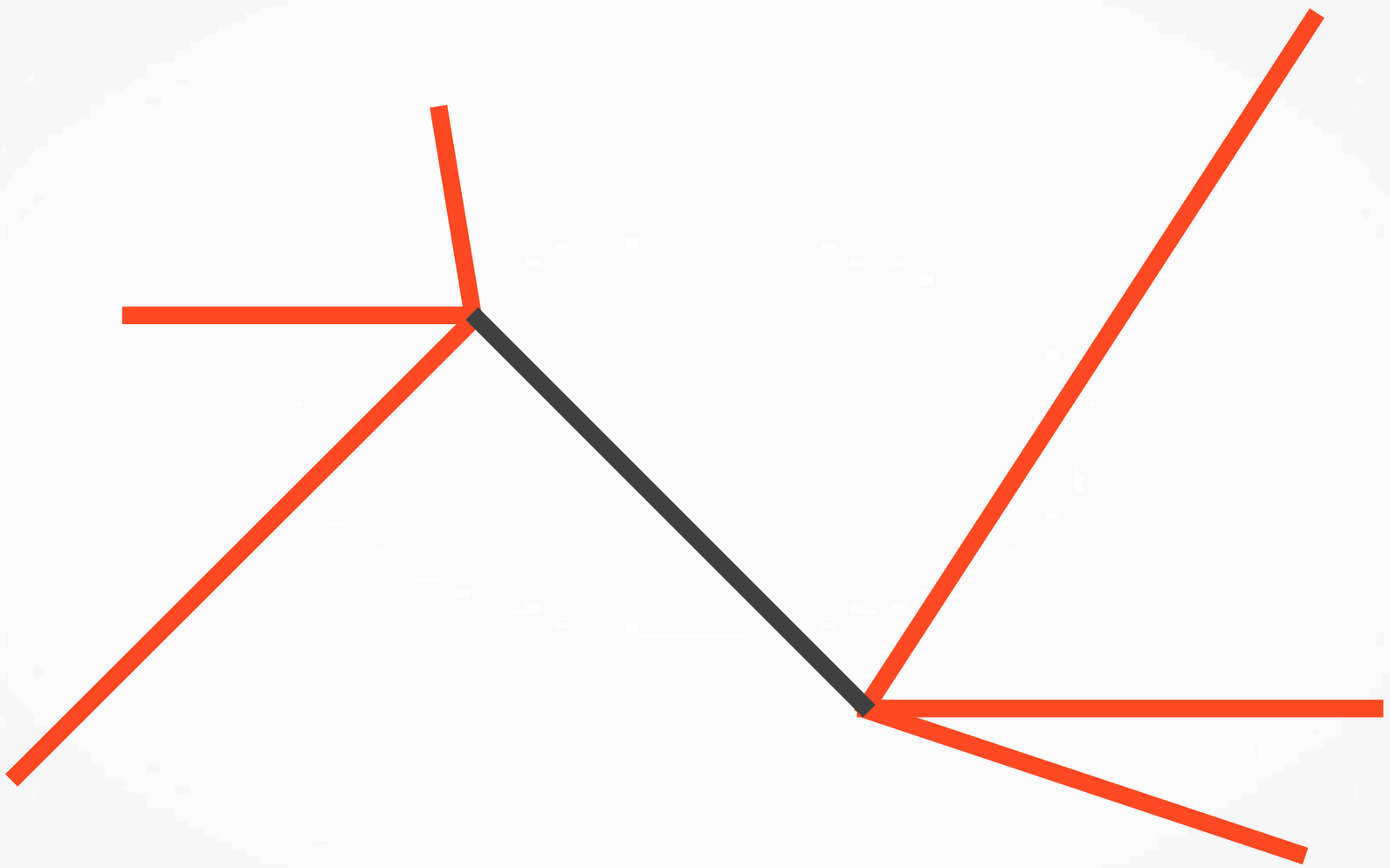
Ethics

Data (Mis)Representation



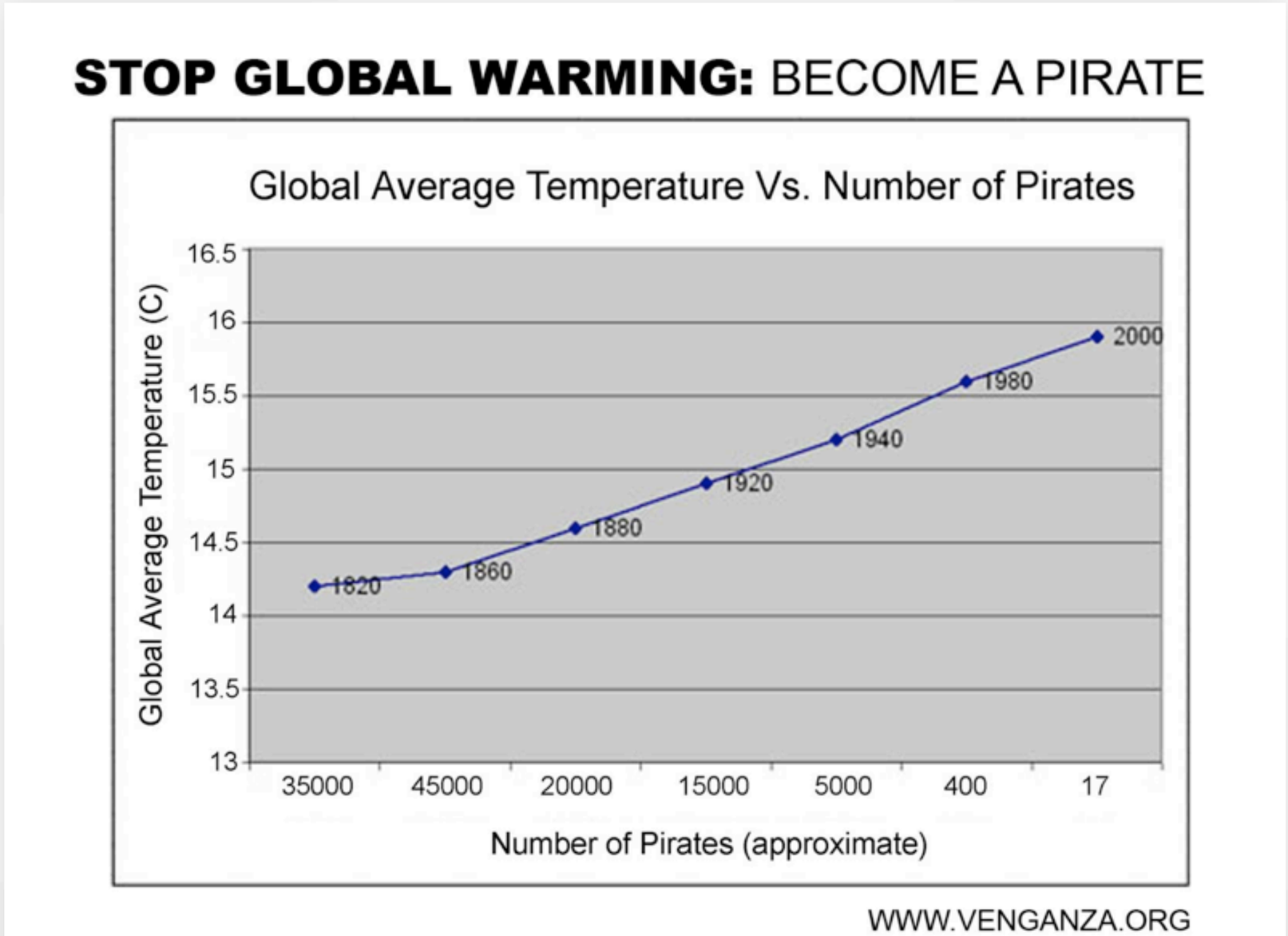
http://mediamatters.org/research/2012/10/01/a-history-of-dishonest-fox-charts/190225

Context Matters



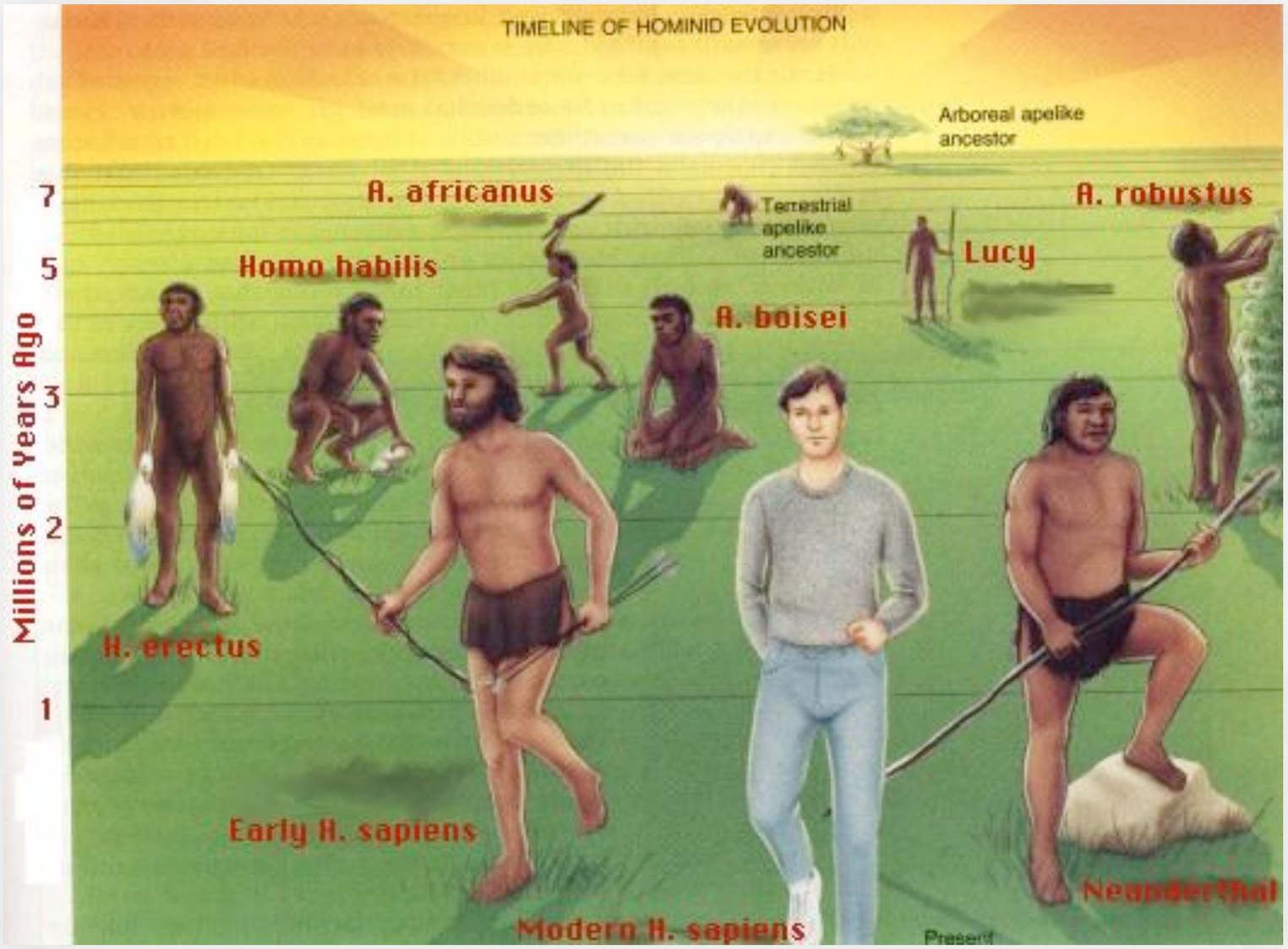
... As Does Scale

Correlation != Causation



http://wp.me/pd52p-3X3

Chart Junk



http://microecos.wordpress.com/2010/03/16/wwtd-what-would-tufte-do/

Interaction & Visualization

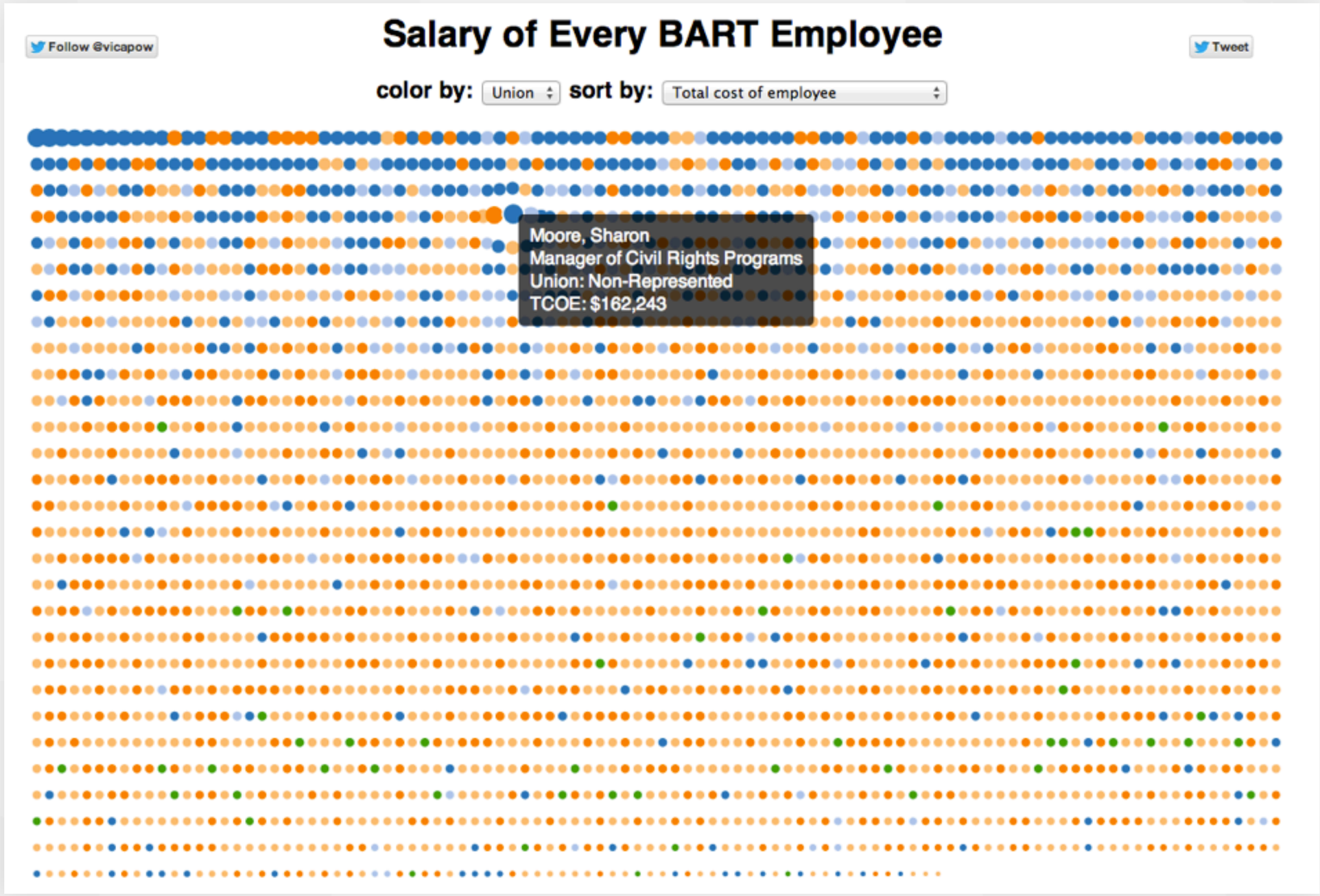
Interaction & Visualization

Interaction can help tremendously with storytelling.

But there are other advantages of interactive visualization:

- User control (to an extent).
- Exploration & filtering data on the fly.
- Presentation of different aspects of the data—“connecting the dots.”

Interaction & Visualization



http://blog.vctr.me/bart/

Be Selective

Just because you can add an interaction to the visualization **does NOT mean you have to.**

“What’s the intent behind interaction X?”

Visualization Tips

Visualization Tips

What are you trying to get across?

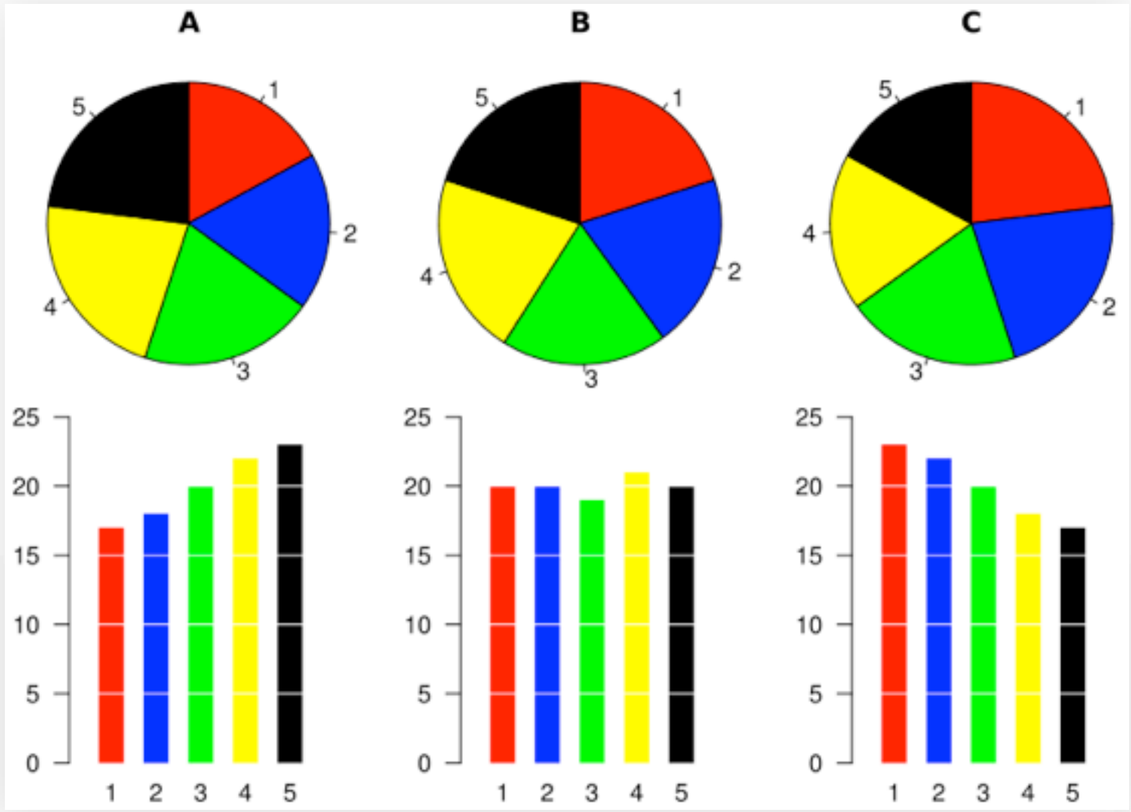
Who is the audience?

How/why is the visualization useful?

Visualization Tips

Consider the data & the best way to represent it.

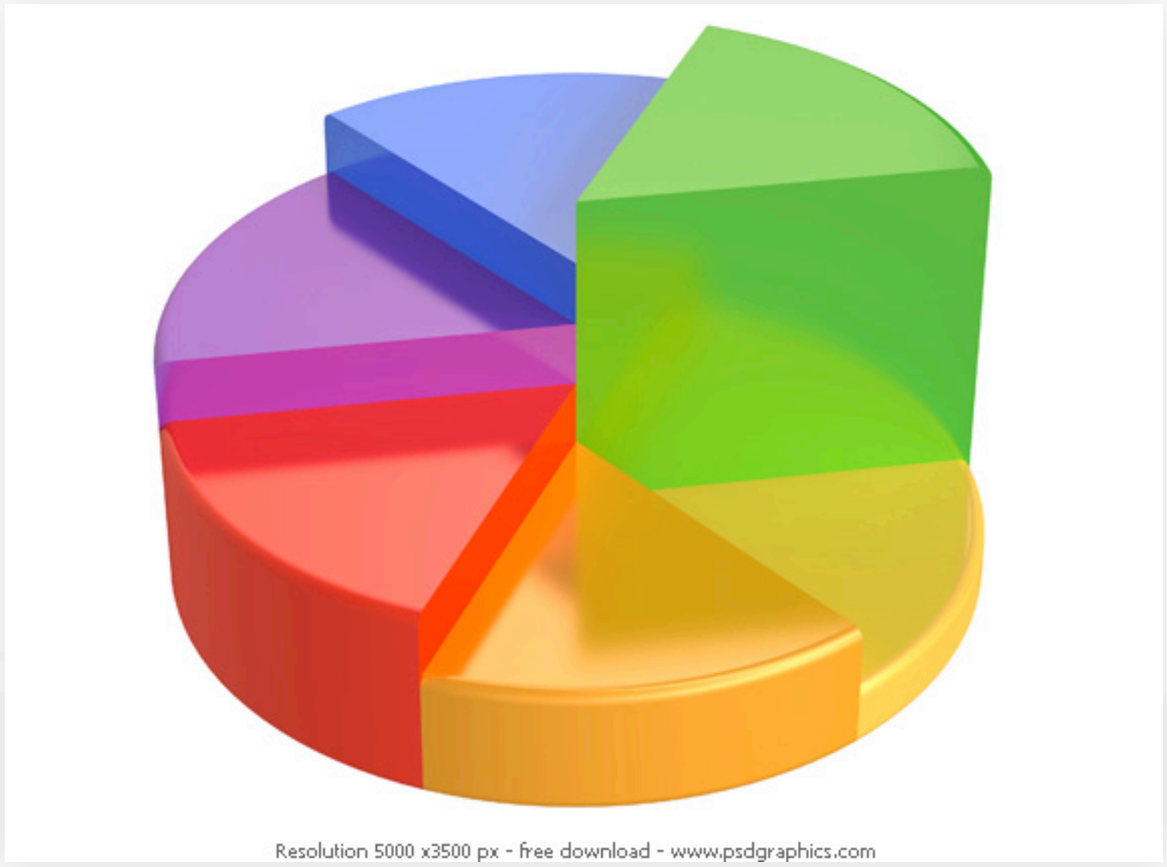
Often, simple is best. (e.g. how is this representation better than a bar chart?)



http://blog.revolutionanalytics.com/2009/08/how-pie-charts-fail.html

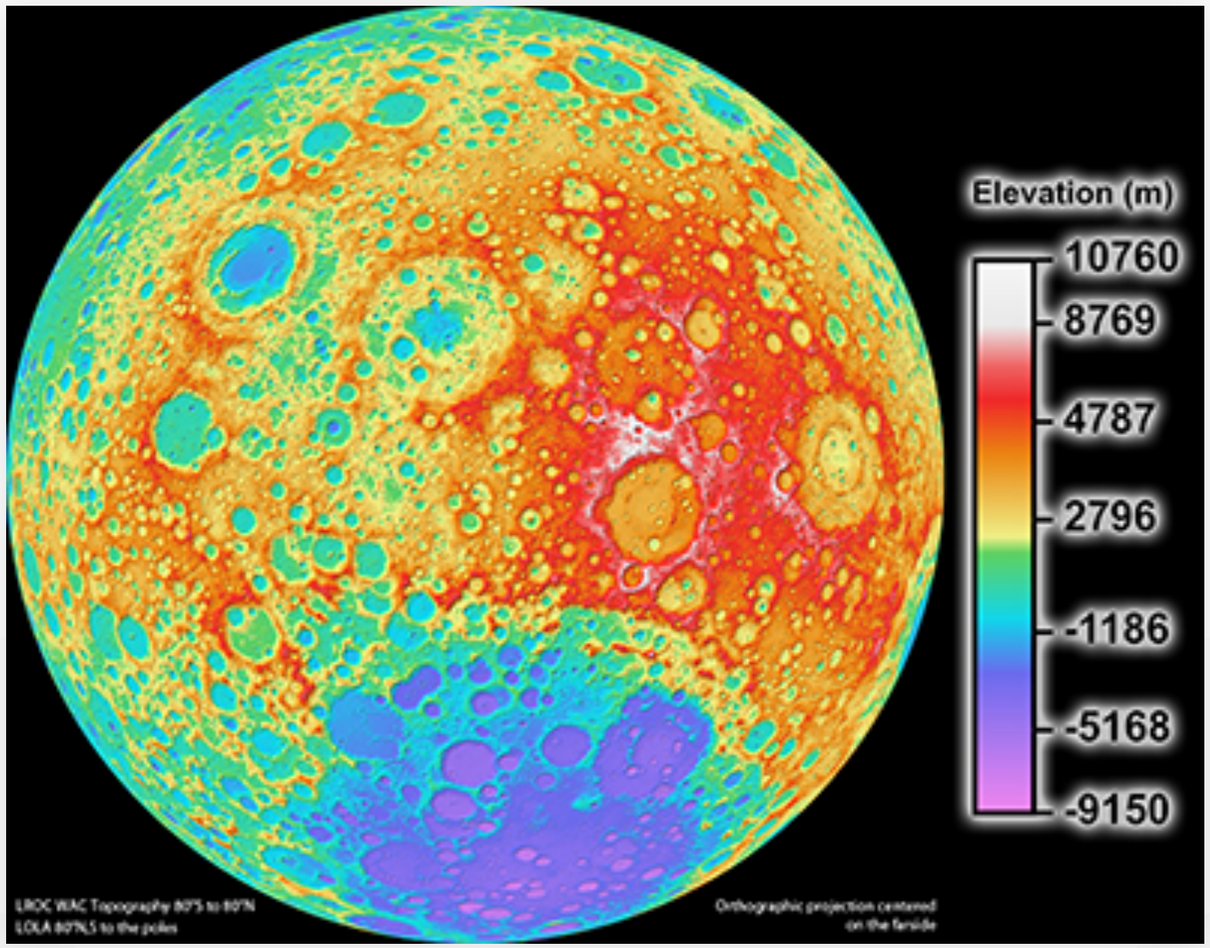
Visualization Tips

Avoid “3D” charts.



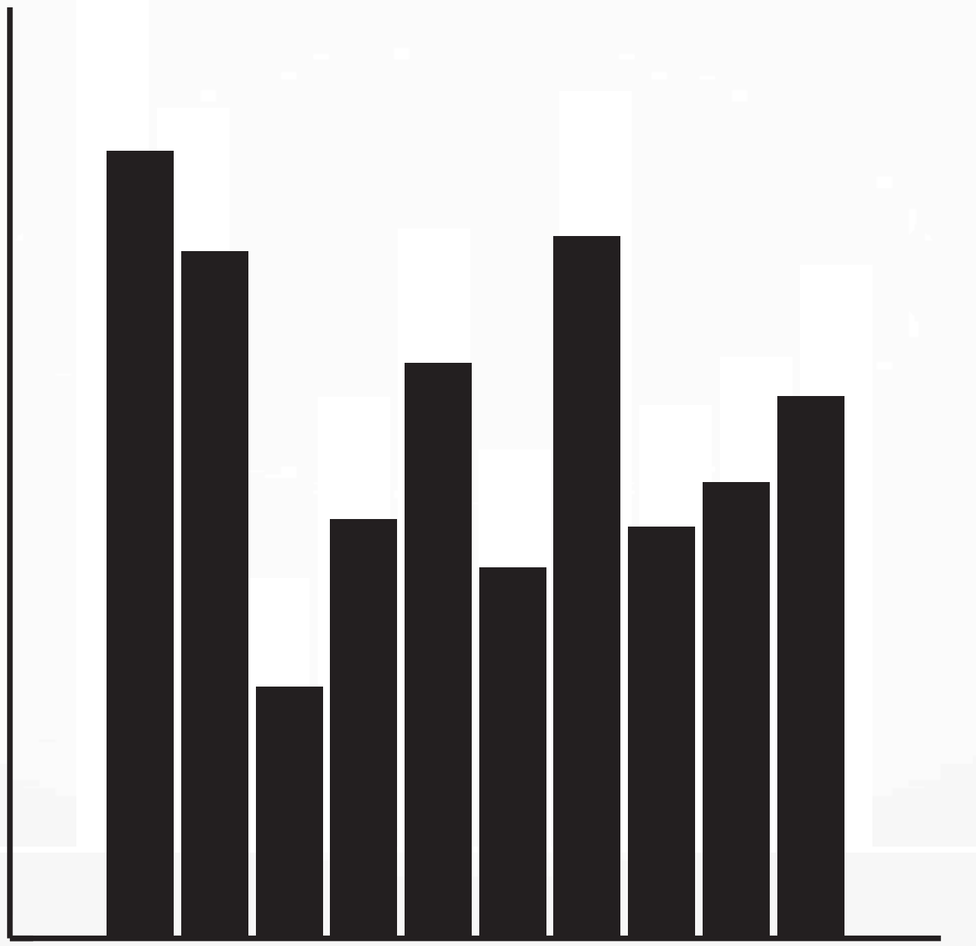
Visualization Tips

Avoid using the rainbow color spectrum to display a continuous range of data.



Visualization Tips

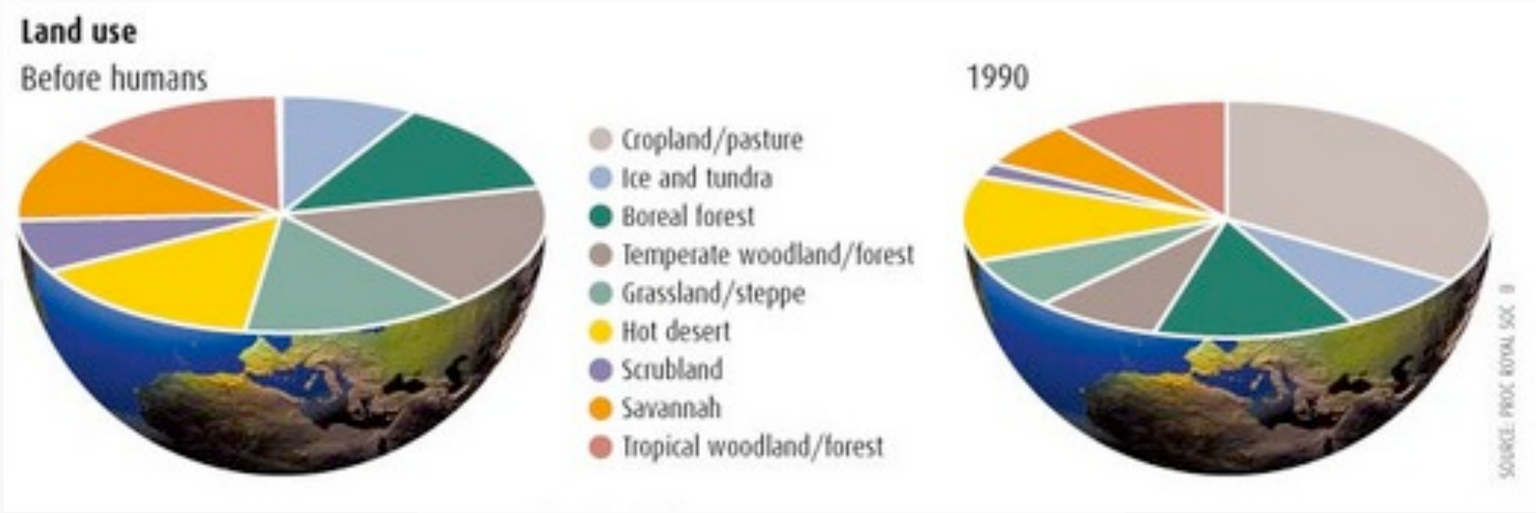
Don’t forget to include titles, axis, & labels.



Visualization Tips

Think about what’s essential for your visualization.

Then, cut out parts that may be redundant (both visual & interaction elements).



Visualization Tips

Look at both good & bad examples.

http://www.informationisbeautiful.net/

http://wtfviz.net/

AND DON'T DO THE BAD ONES.

http://25.media.tumblr.com/tumblr_lxi1ahSqPS1r9ahfuo1_400.gif

Data

Social Data is Ubiquitous

Government & Open Data

US—http://www.data.gov/

California—http://data.ca.gov/

San Francisco—https://data.sfgov.org/

SODA (Socrata Open Data API)—http://dev.socrata.com/

Some Tools

Mr. Data Converter—http://shancarter.github.io/mr-data-converter/

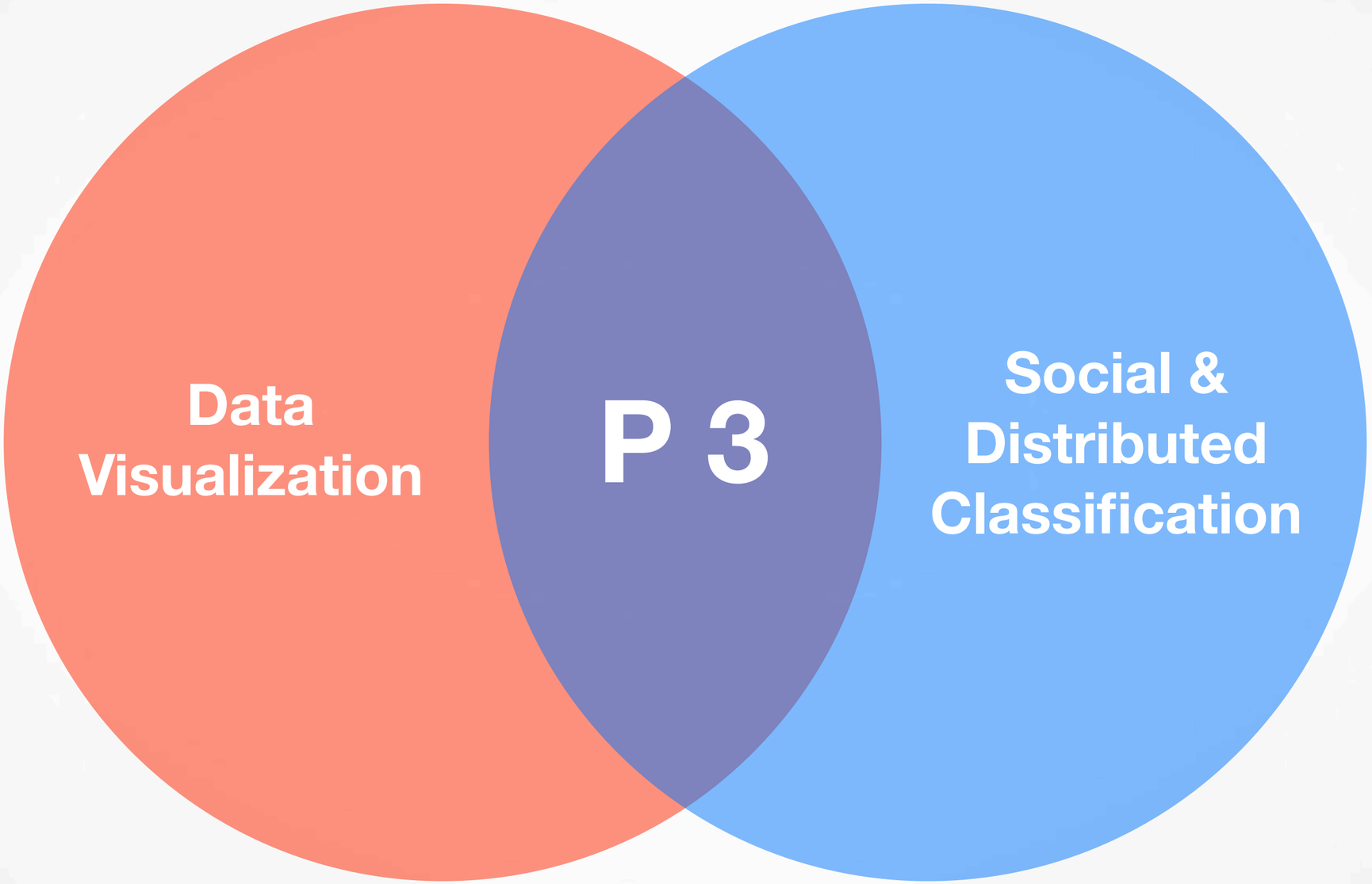
Google Refine—https://code.google.com/p/google-refine/

BeautifulSoup—http://www.crummy.com/software/BeautifulSoup/

Mechanical Turk—https://www.mturk.com/

Project 3

Project 3



PS: Start exploring/gathering data **now**.

Next Class

Next Class

Project 2 Presentations