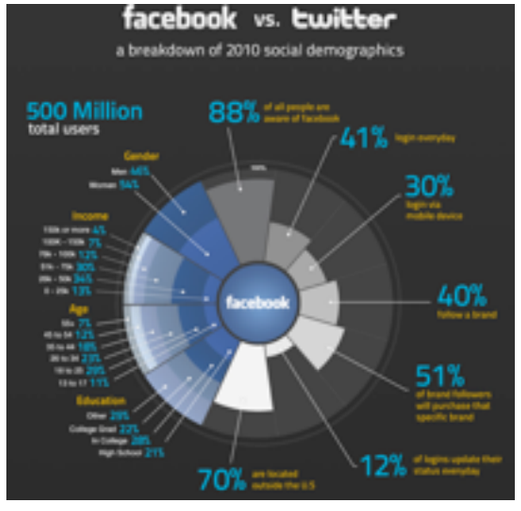
A few points of clarification to add to my lecture, as there were questions afterwards, and I wasn't sufficiently clear on a few points.

We know from Cleveland and McGill (https://secure.cs.uvic.ca/twiki/pub/Research/Chisel/ComputationalAestheticsProject/cleveland.pdf p536, among others) that we're better at comparing length (with a common or unshared baseline) than any other visual quantity (area, angle, etc.).

They state their premise very well:
A graphical form that involves elementary perceptual tasks that lead to more accurate judgements than another graphical form (with the same information) will result in better organization and increase the chances of a correct perception of patterns and behavior.

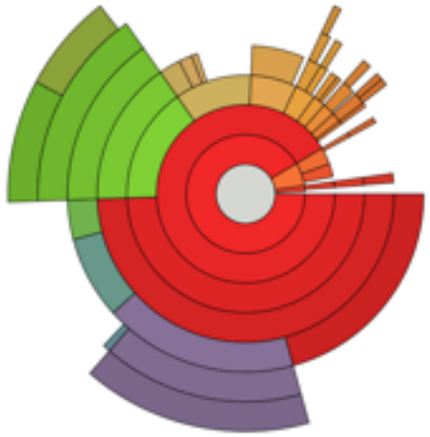
The practical result is that rectangular bars and stacked bars are easier to accurately compare (because they're a linear comparison) than graph types that involve areas of varying widths (rectangles with similar widths can be compared linearly), and therefore should be the preferred format when accurate interpretation of the data matters.

Now, the Nightingale roses, and any similar radius-based stacked bar graphs, are ostensibly fine, because the proposition is that the radial distance matters, not area. This format includes the Facebook vs Twitter graph I discussed.



However, we perceive the area, not just height, and give disproportionate wight to the outer portions of the slice. (Examples of the original roses and less distorted variations are here: http://mbostock.github.com/protovis/ex/crimea-rose.html)

Similarly, the hierarchical example shown later attempts to use angular fractions of the whole to show relative volume, but we end up seeing children and grandchildren with larger areas than the parent, because the outer rings comprise so much more area then the inner rings. This too is a huge distortion of the data.



All that said, circular layouts of data are totally valid in situations where there is no requirement to compare relative area at different distances from the center. Examples include formats such as: http://mbostock.github.com/d3/ex/tree.html and http://circos.ca/, both commonly used in the life sciences, as well as representations of cyclical patterns.

I hope that helps to clarify the ambiguity around the circular formats I claimed were inappropriate.

Thanks for the invitation to speak, I had an excellent time.

-Noah

Effective Data Visualizations

Noah Iliinsky • @noahi

UW Genome Sciences Combi • December 7, 2011

Why Visualization?

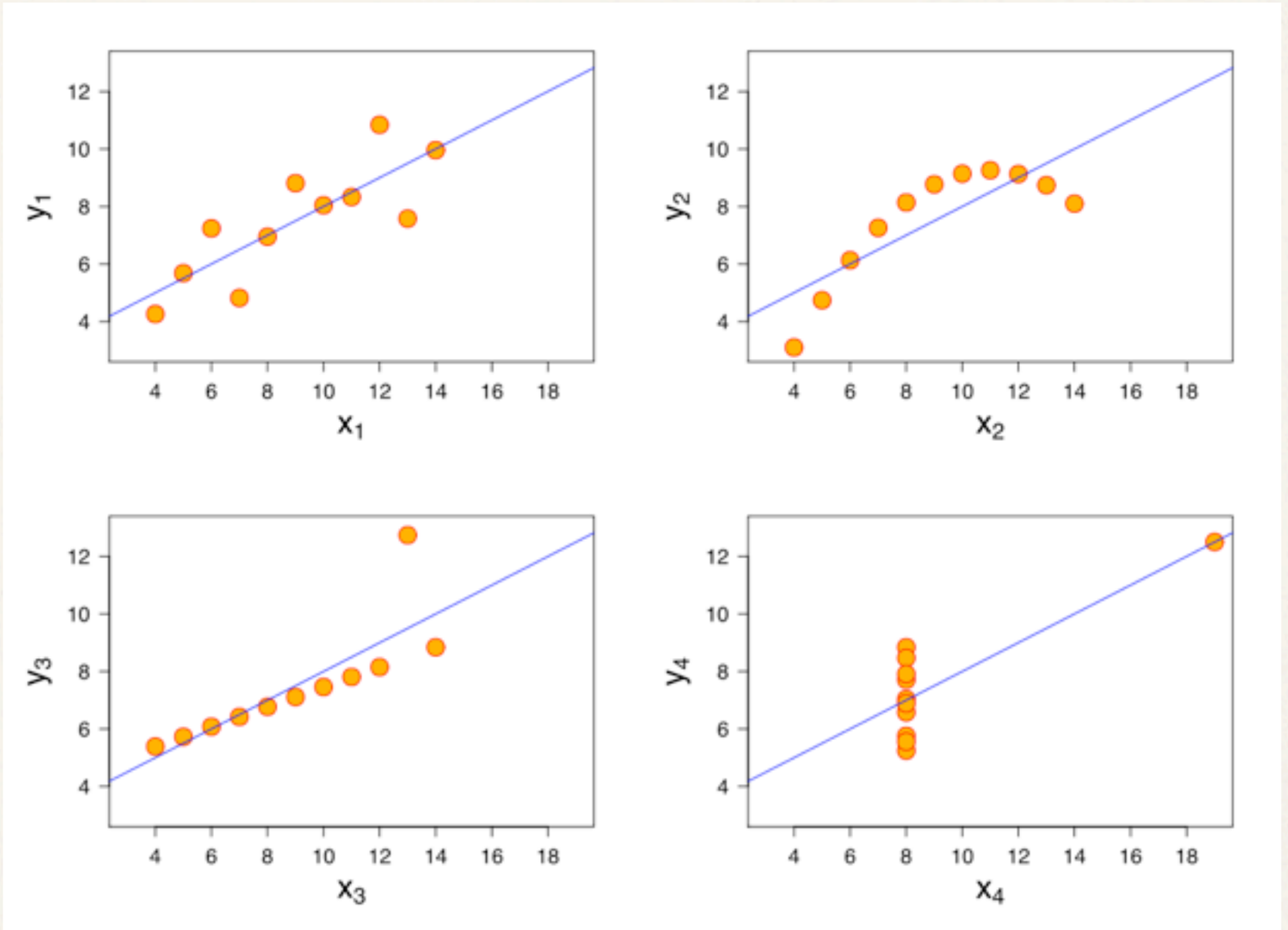
Why visualization?

Anscombe's Quartet

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

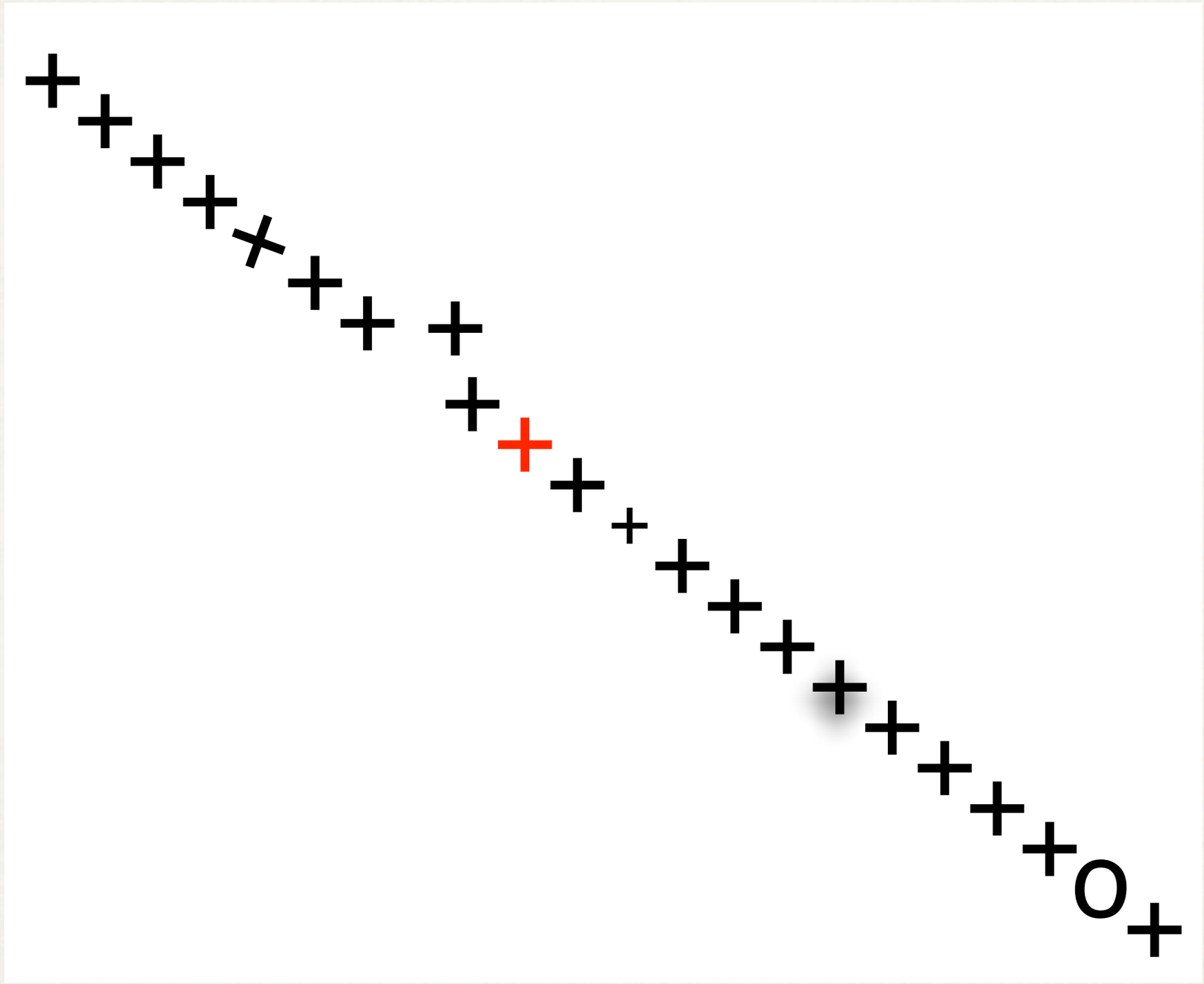
http://en.wikipedia.org/wiki/Anscombe%27s_quartet

Visualization makes data *accessible*.

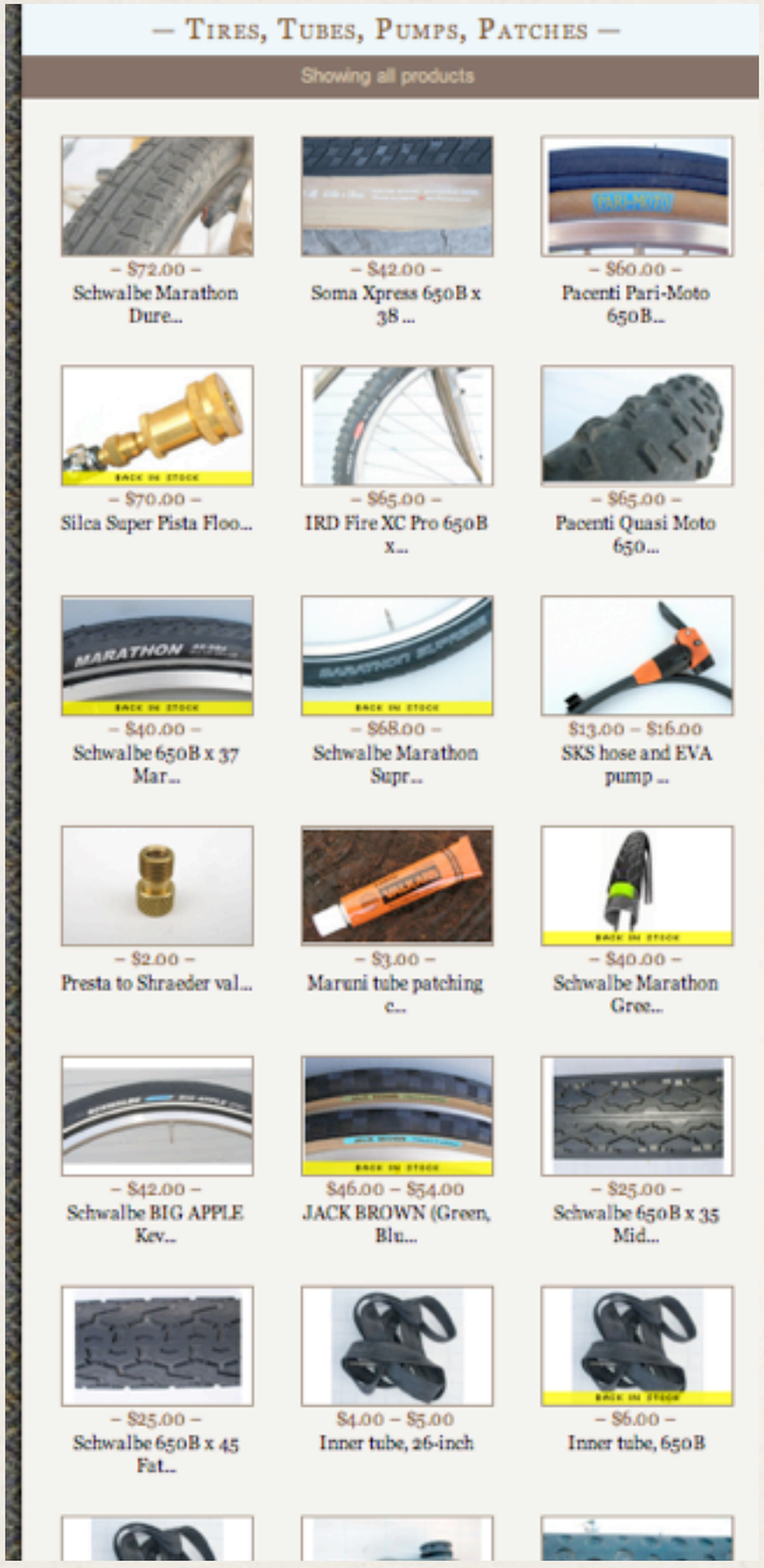


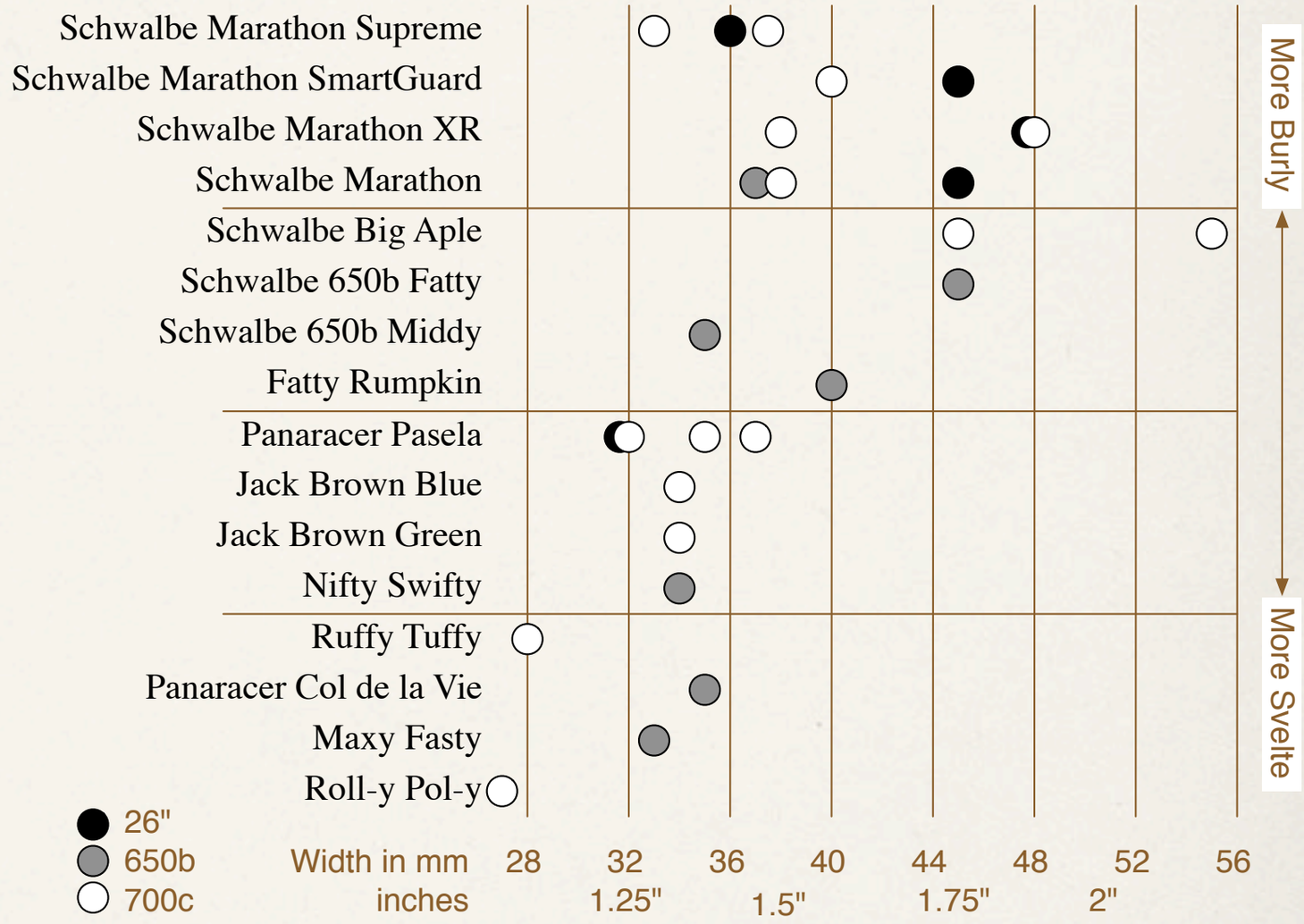
http://en.wikipedia.org/wiki/Anscombe%27s_quartet

Visualization leverages the amazing abilities of our eyes and brains



Visualization gives faster access to actionable insights

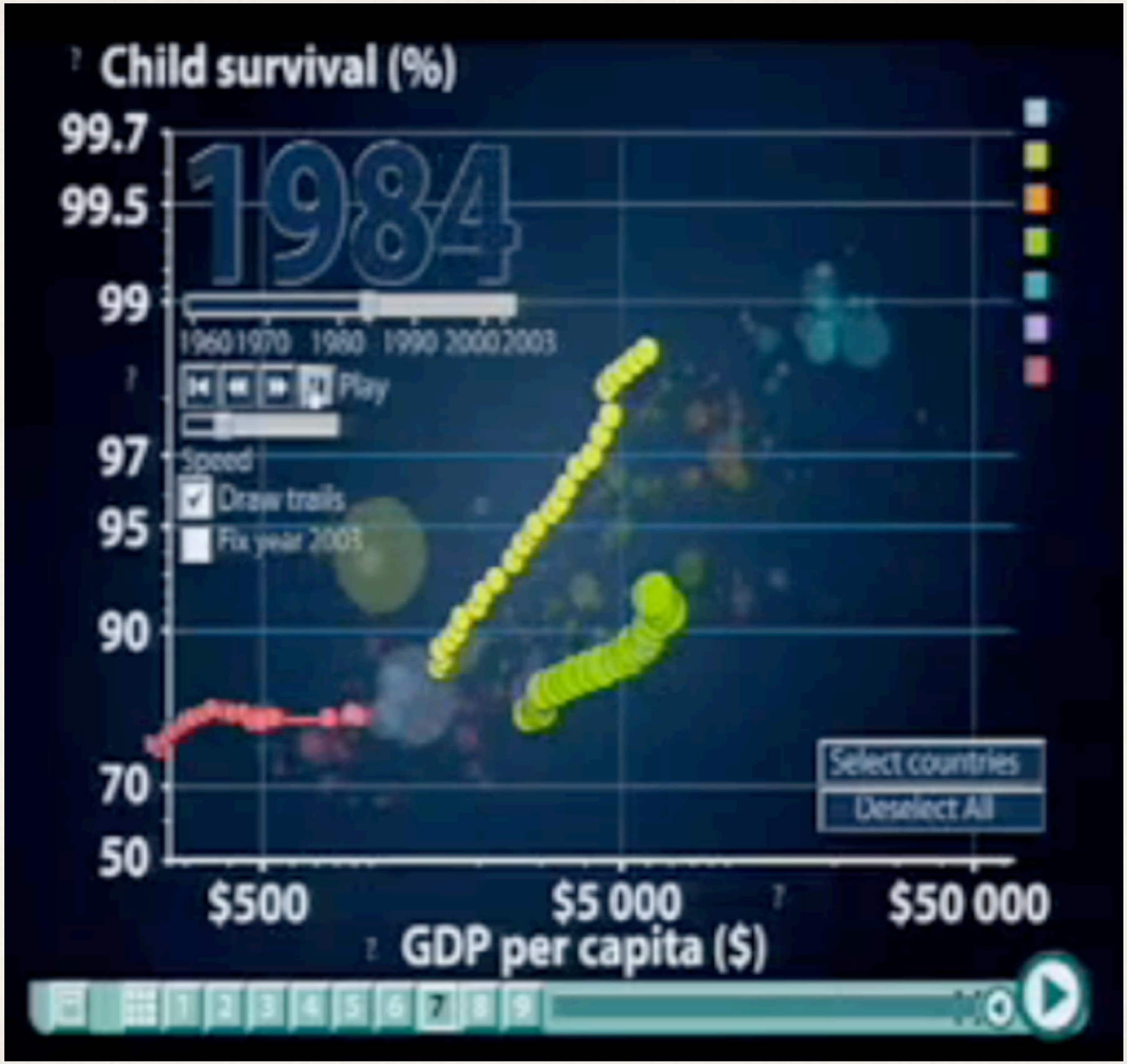




http://www.rivbike.com/products/list/tires_tubes_pumps_patches?a=1&page=all

http://complexdiagrams.com/2009/03/12/tire-chart/

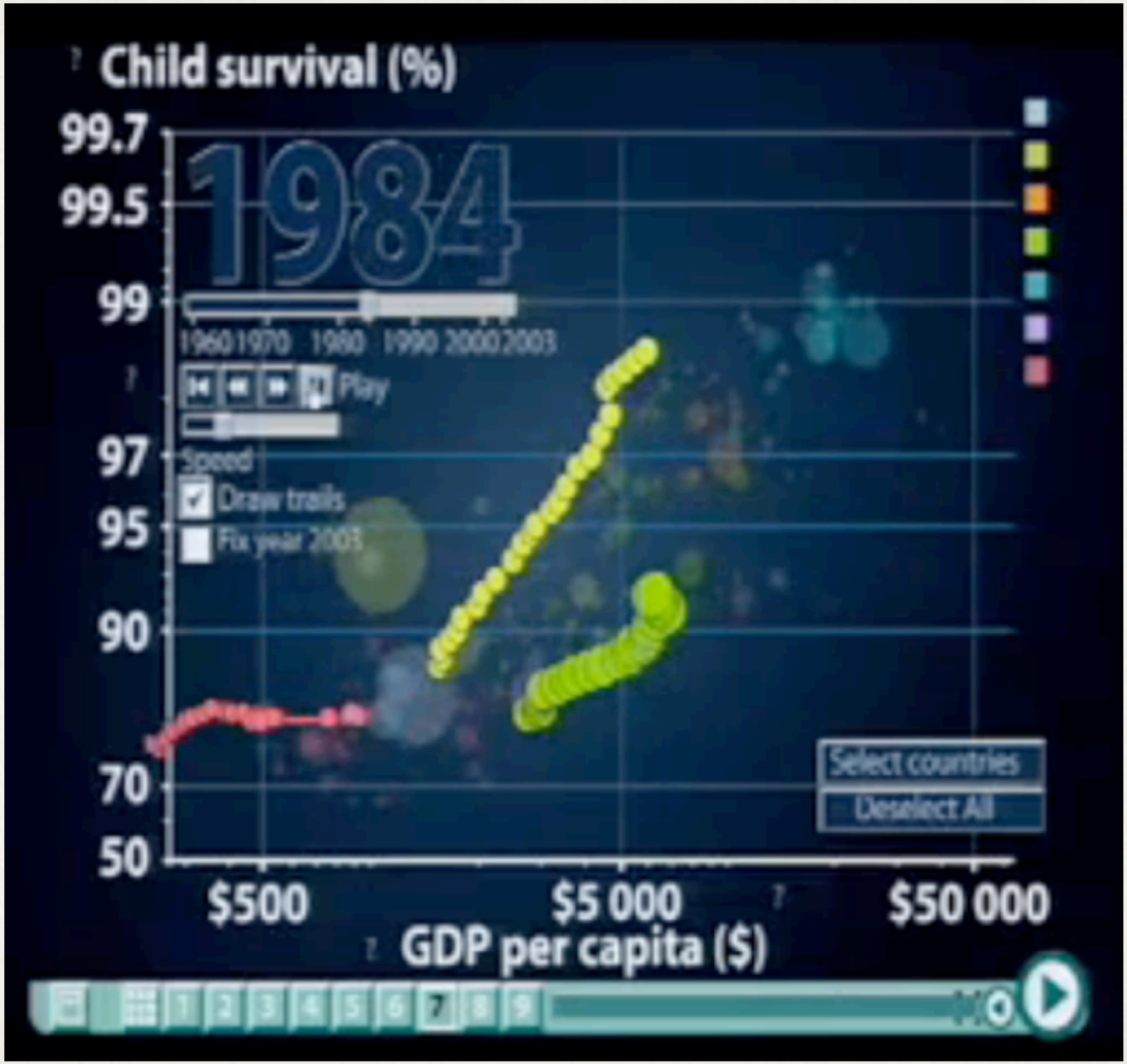
Visualization allows access to huge amounts of data



http://www.youtube.com/watch?v=hVimVzgtD6w

Why Stories?

Stories make data *relevant*.

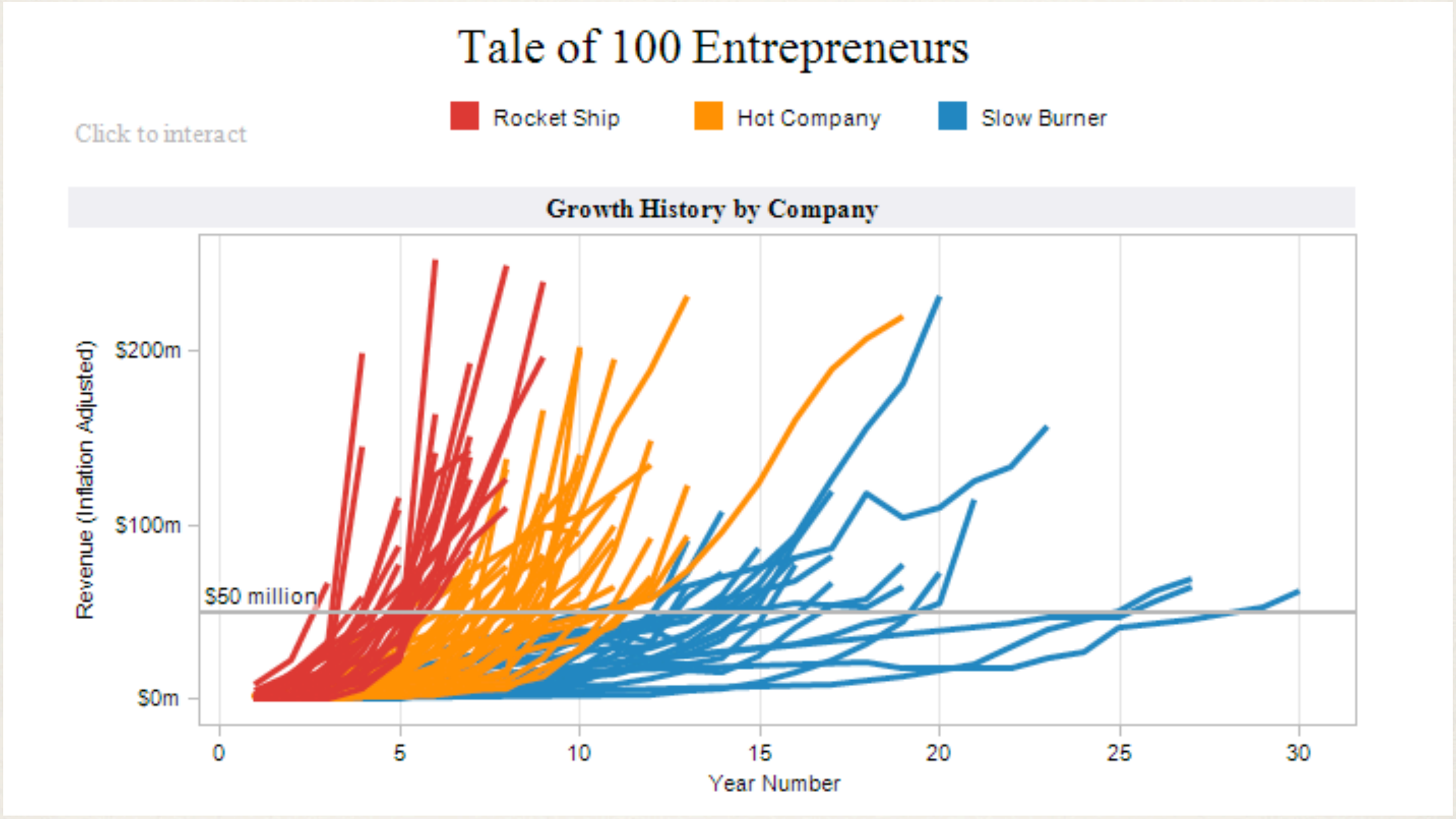


http://www.youtube.com/watch?v=hVimVzgtD6w

Part One: Concepts & Definitions

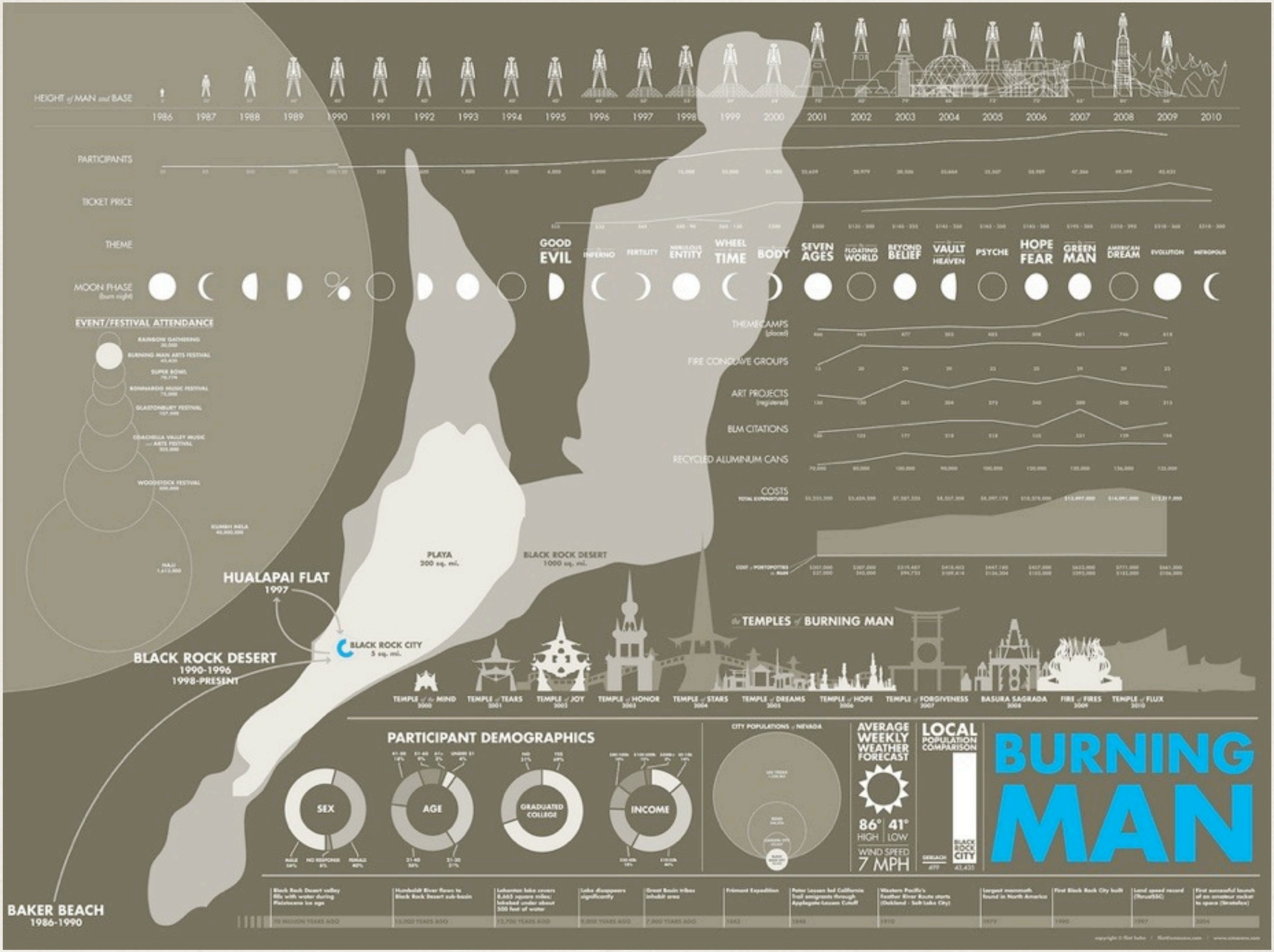
Data Visualization vs Infographics

Data Visualizations are generated by software.



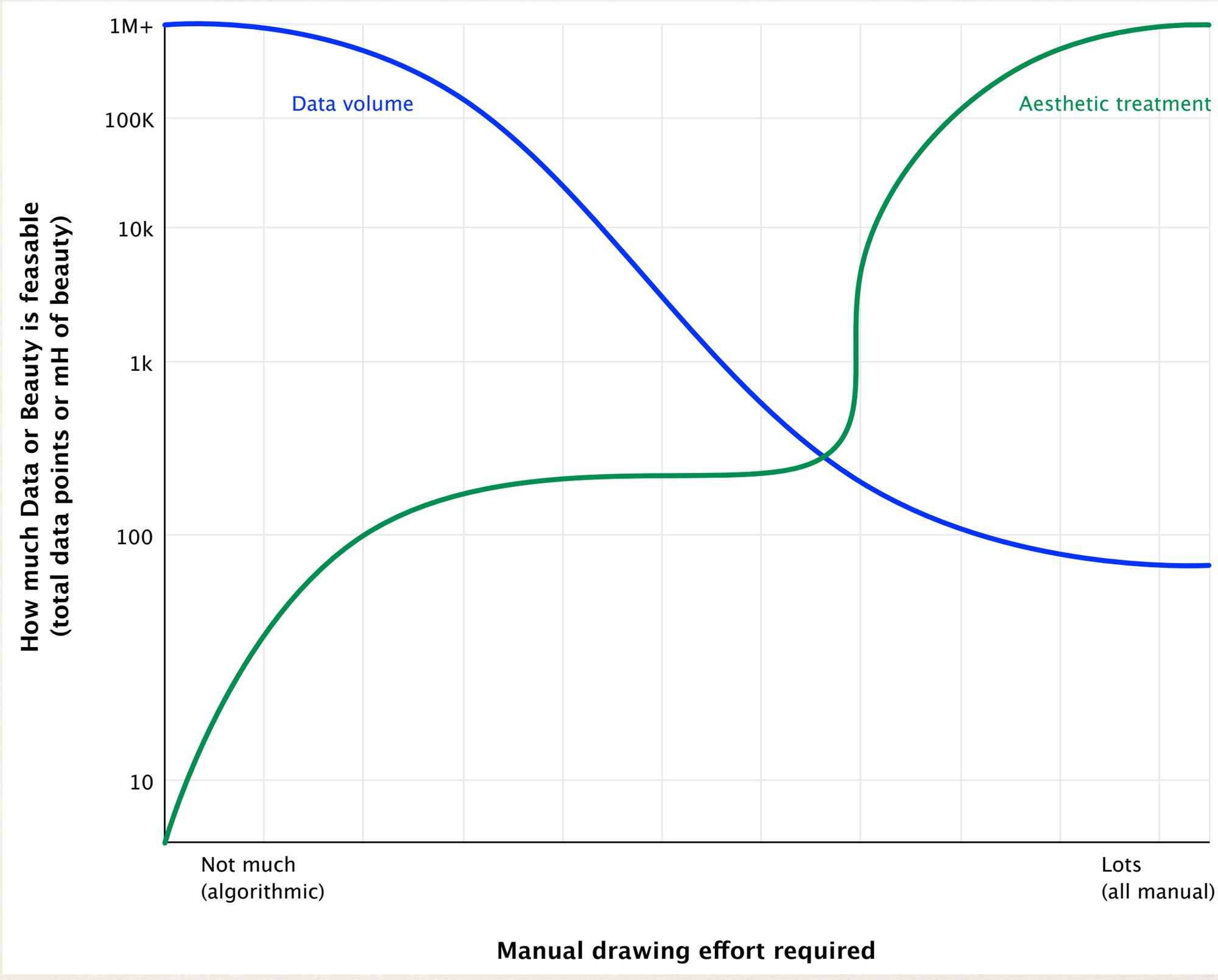
http://www.tableausoftware.com/public/gallery/taleof100

Infographics are manually drawn.



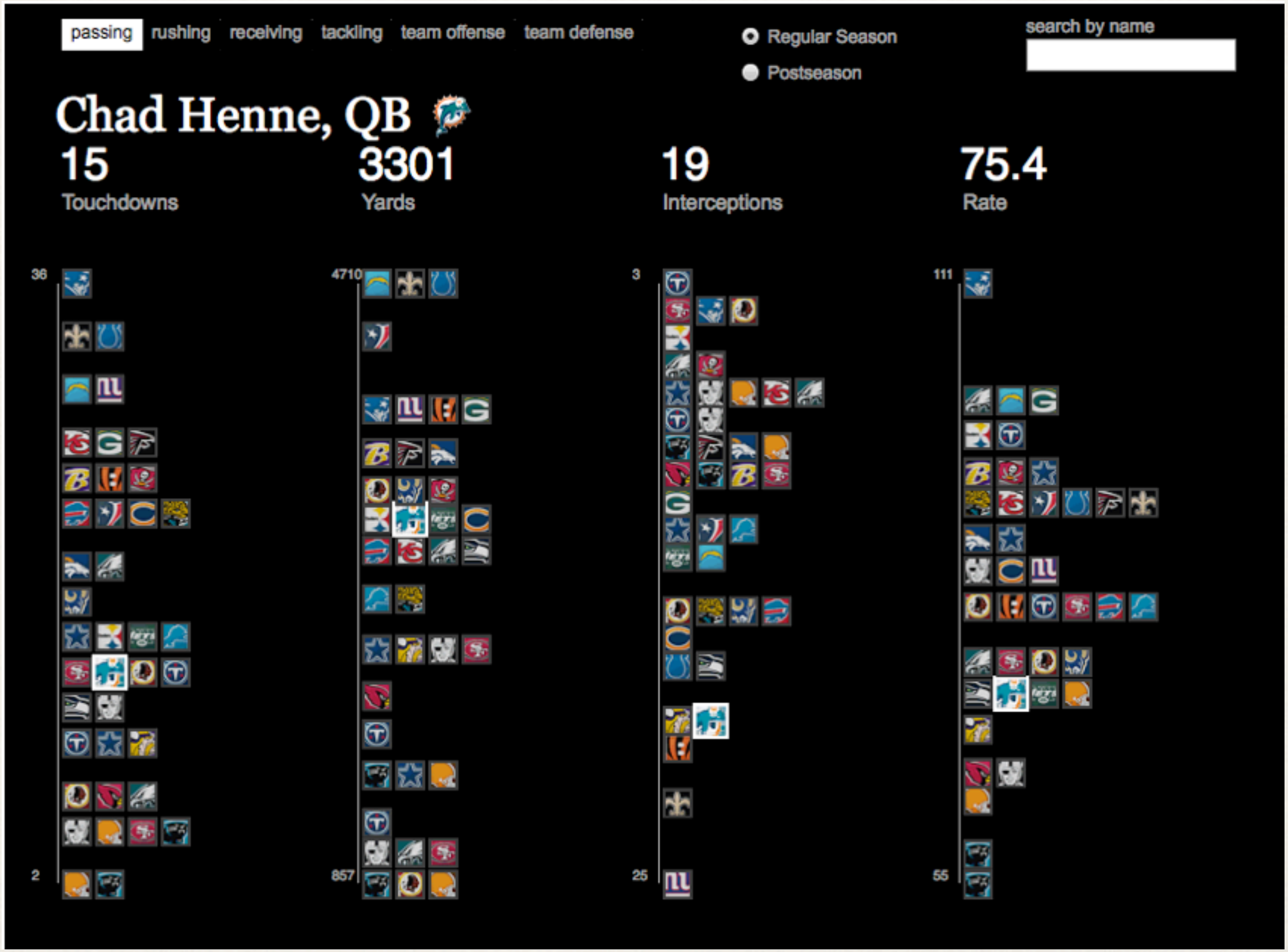
http://www.flickr.com/photos/xmasons/4841339241/sizes/l/in/photostream/

Data visualizations vs Infographics



Exploration vs Explanation

Visualization for *exploration*, when you don't (yet) have a story to tell.



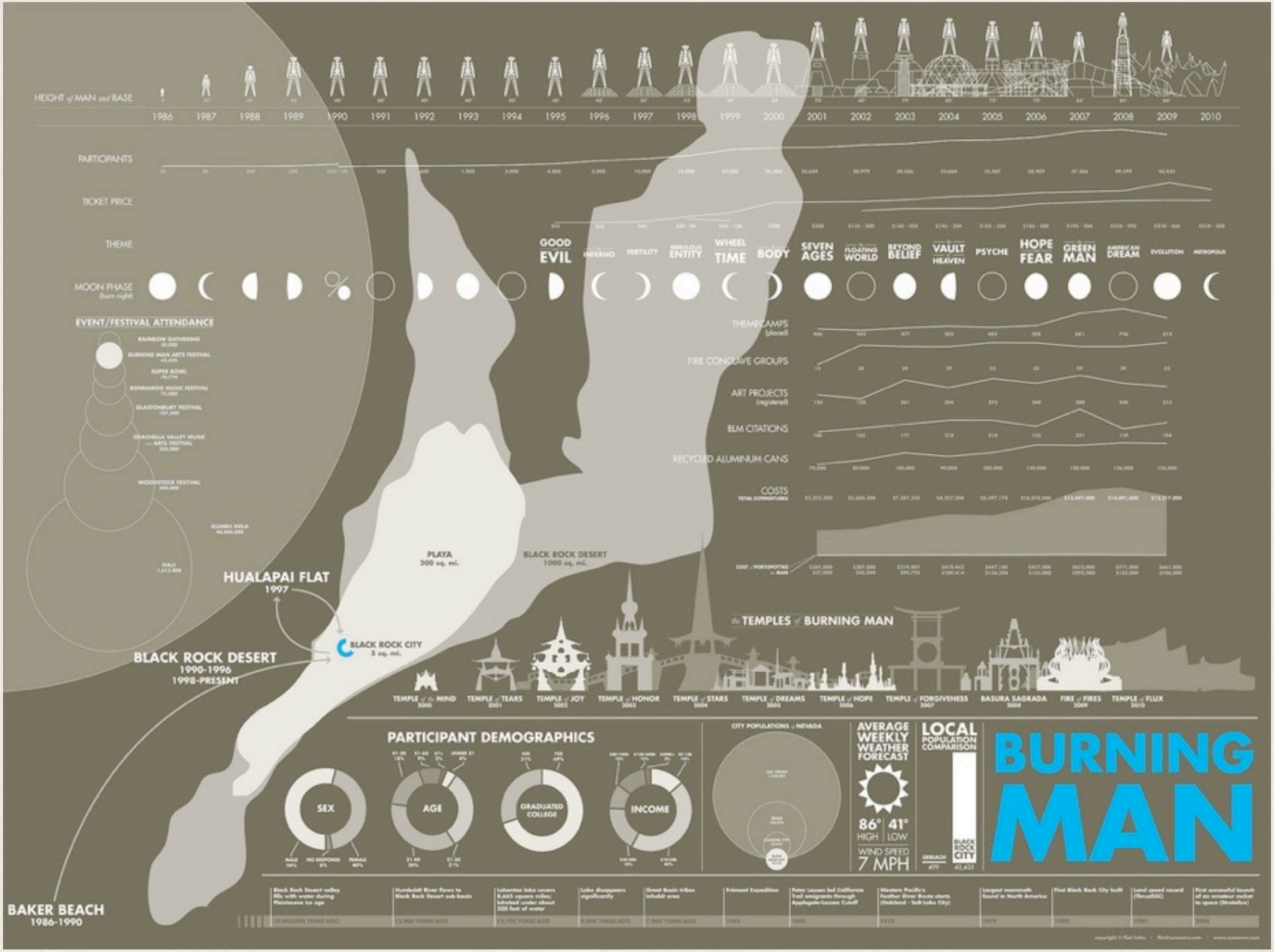
http://www.juiceanalytics.com/nfl-visualization/

Visualization for *explanation*, when you do have a story to tell.

https://www.msu.edu/~howardp/softdrinks.pdf

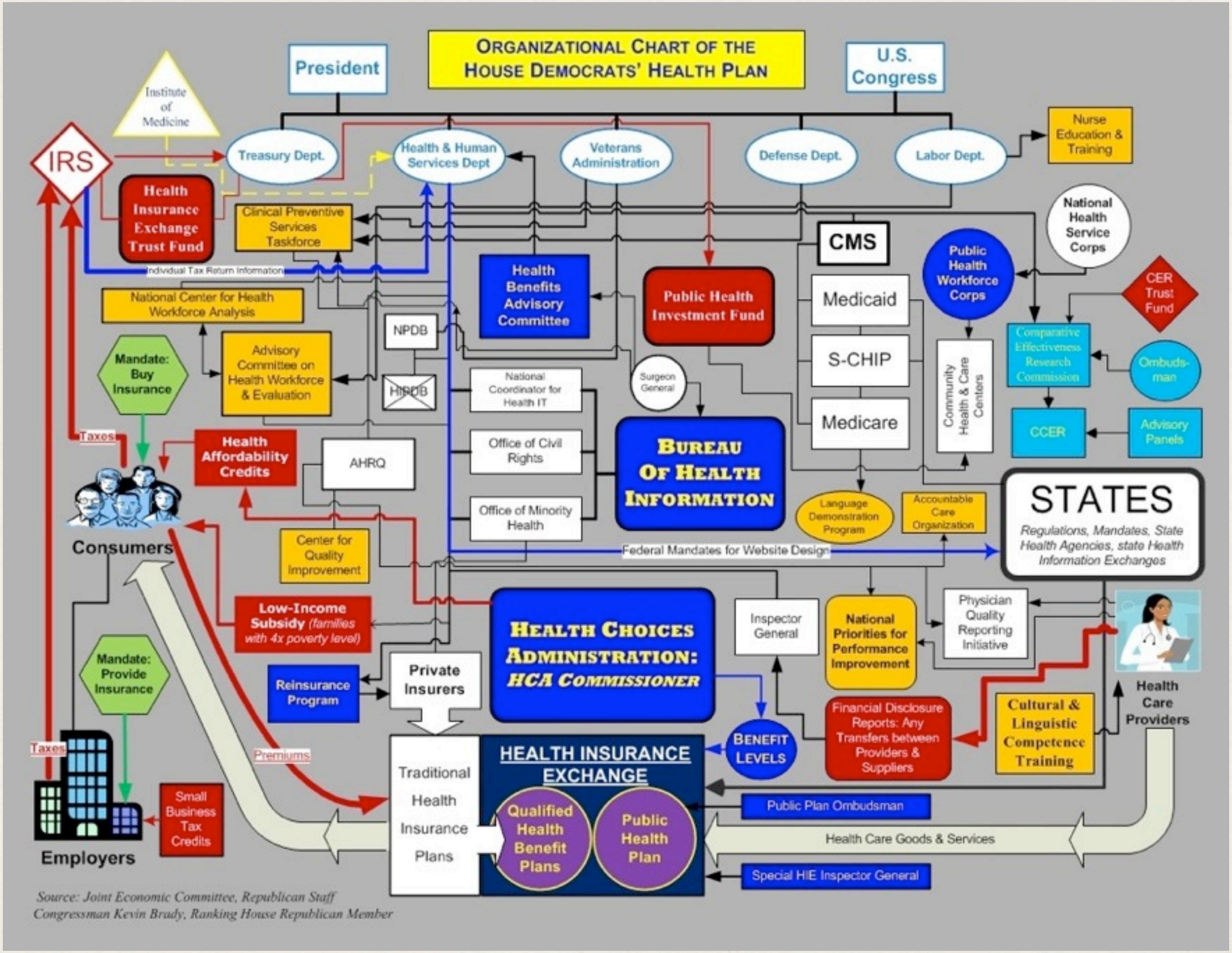
Education vs Persuasion

Visualization for *education*



http://www.flickr.com/photos/xmasons/4841339241/sizes/l/in/photostream/

Visualization for *persuasion* (or *propaganda*)



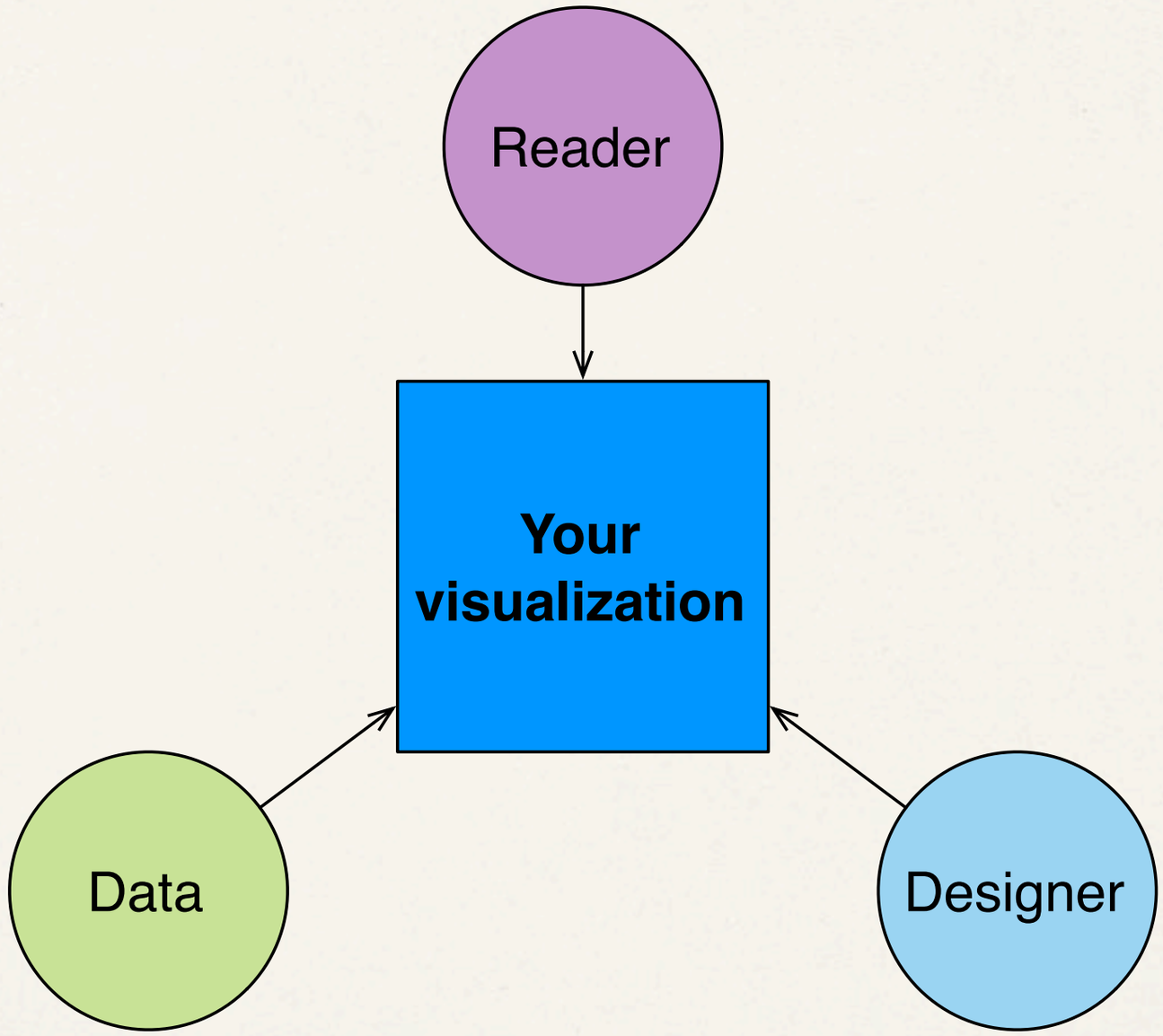
http://www.house.gov/apps/list/press/tx08_brady/71509_hc_chart.html

Part Two: How To Do It

Make good choices.

Good Choices are Guided by Three Inputs

Three inputs.



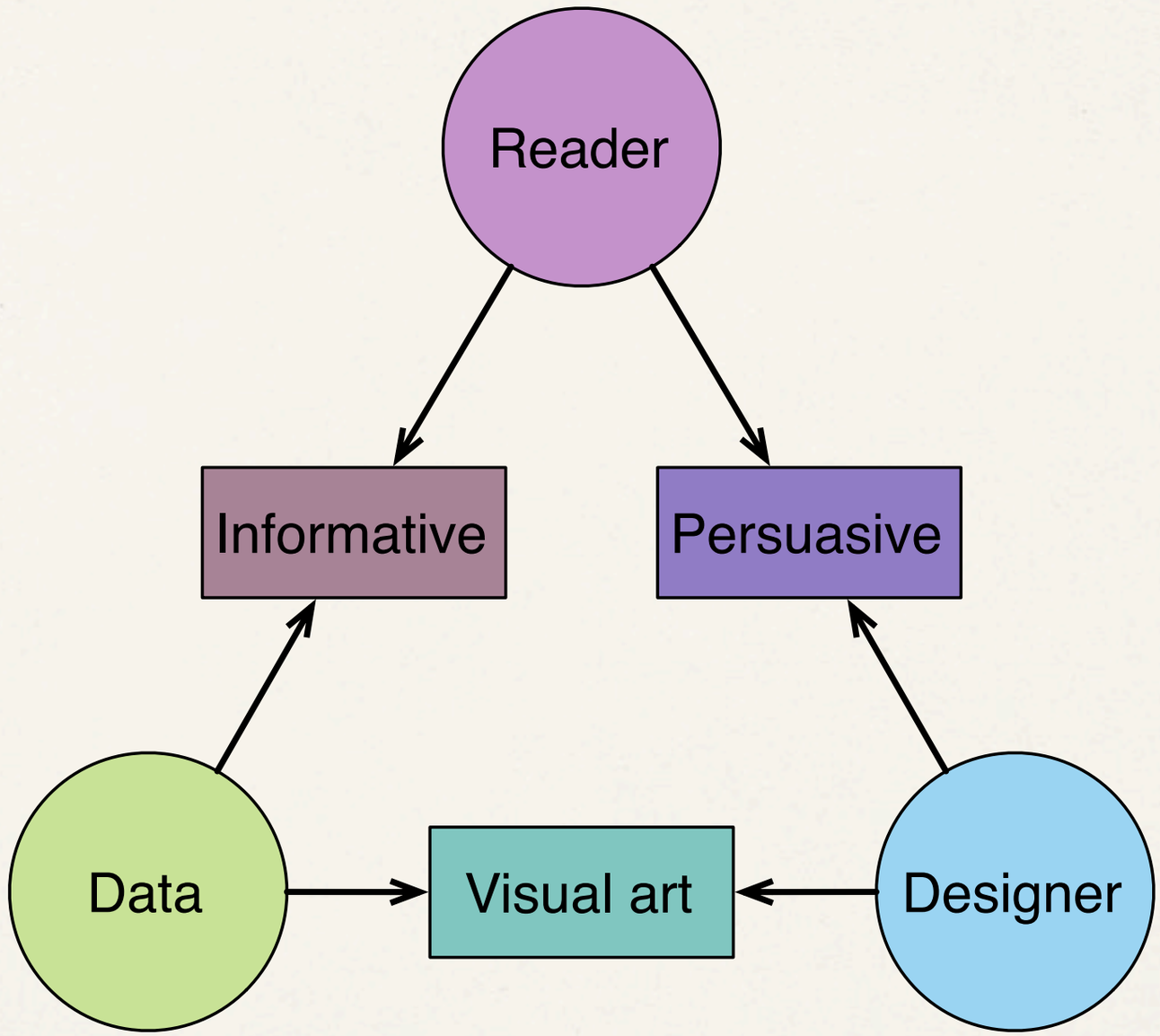
You have goals.

You have goals.

Why are you here?

If you can't concisely articulate your goal, you're doing it wrong.

Three types of information products.



Your Reader has Needs.

Your reader has needs.

Your success is defined by
your readers' success.

Your reader has needs.

Your success is defined by your customers' success.

If you can't satisfy **their** needs, **you** have failed.

Identify your audience!

Understand their
hopes, dreams, and
favorite flavors!

Understand their jargon,
identity, and contexts of
use!

Consider the contexts and needs of:

- a lab mate
- a geneticist
- a scientist
- a member of the general public

Data has Properties

Data has properties.

* Wheel size: numeric (actually categorical)
* Tire width: continuous
* Price: continuous
* Anti-puncture: binary
* Foldable: binary

Now we start designing.

Statement of Goals

Statement of goals.

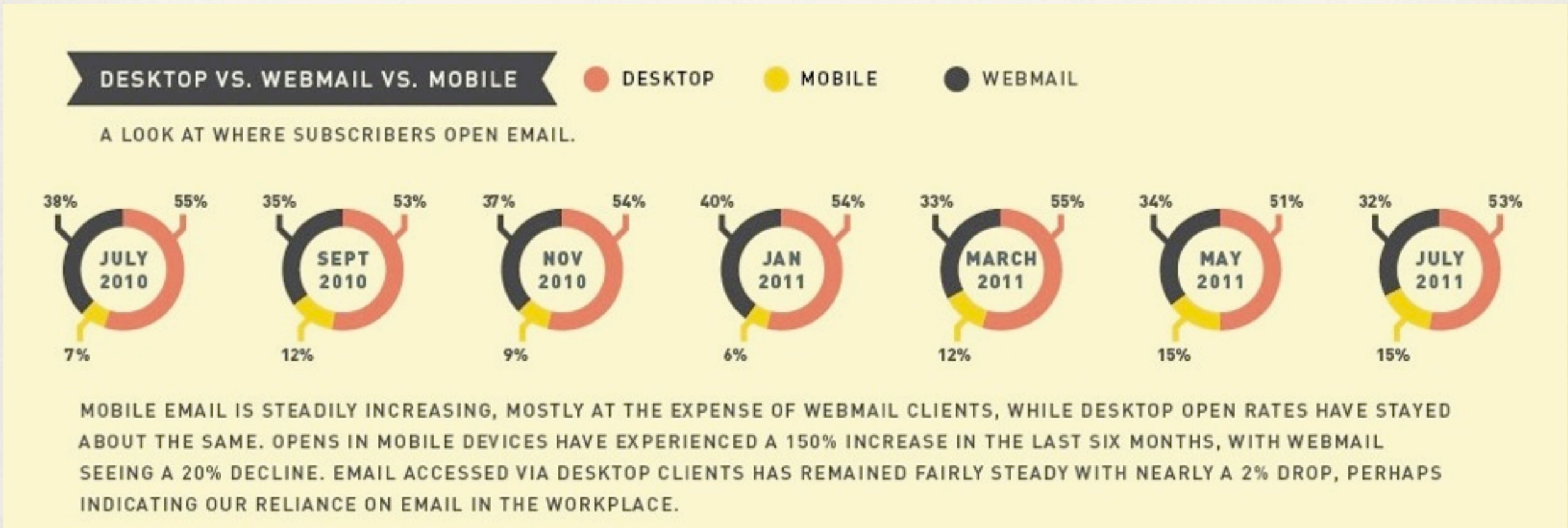
WRONG

"Show the results."

"Show the specific effects of three different drugs across the selected mutations and a control."

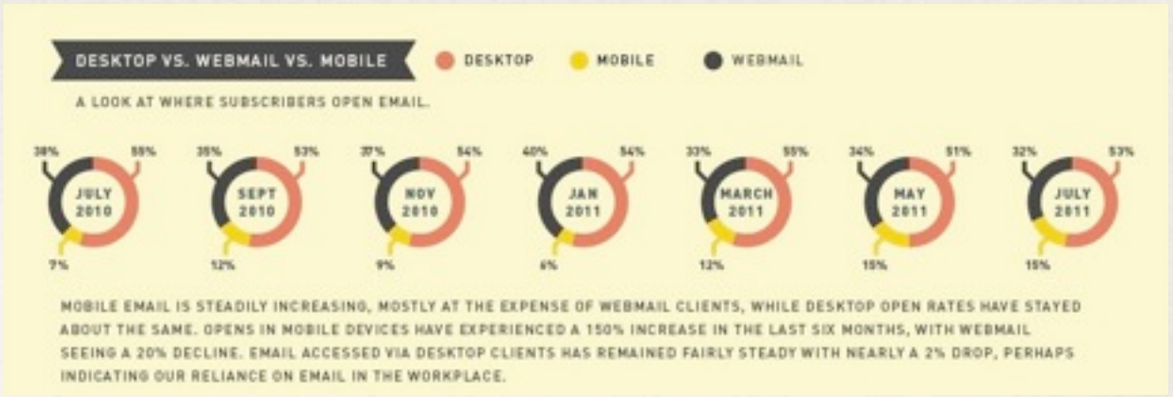
Define Desired Knowledge *Before* Structure

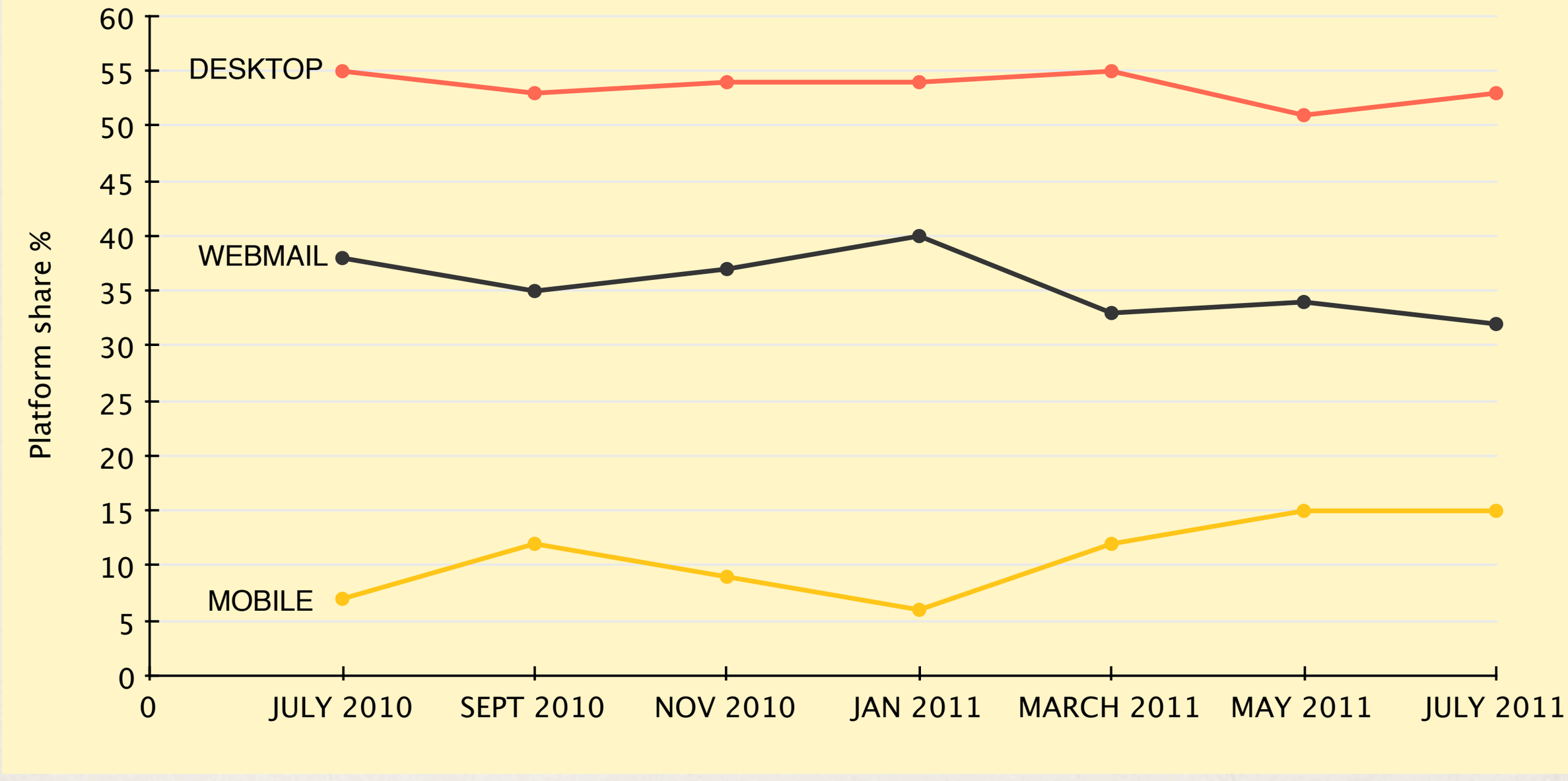
Knowledge *before* structure.



http://litmus.com/blog/email-client-market-share-infograph

Knowledge *before* structure.





Appropriate Encodings

Data has properties.

* Wheel size: numeric (actually categorical)
* Tire width: continuous
* Price: continuous
* Anti-puncture: binary
* Foldable: binary

Encoding well:

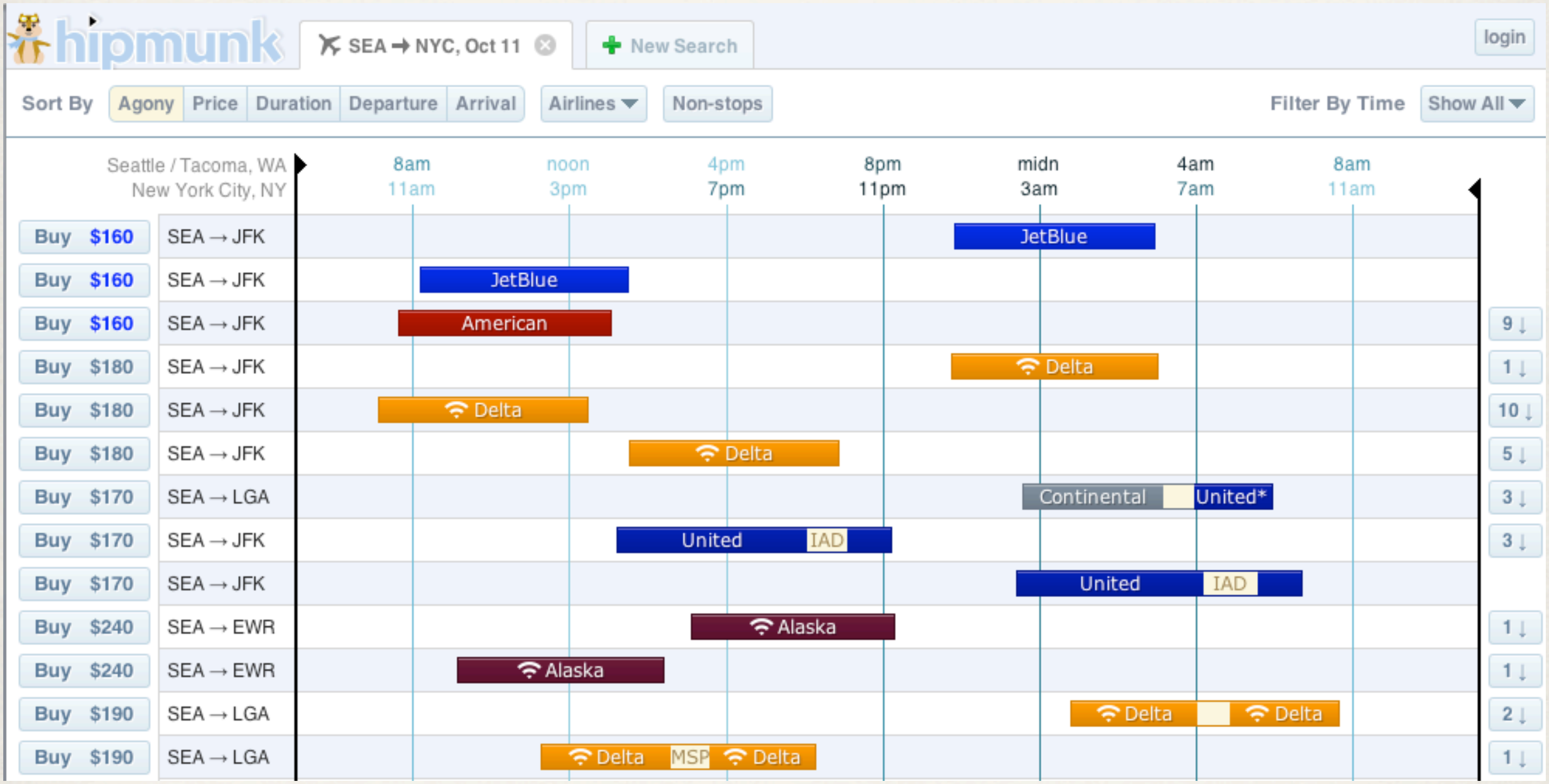
1. Position is everything.

2. Color is difficult.

- @moritz_stefaner

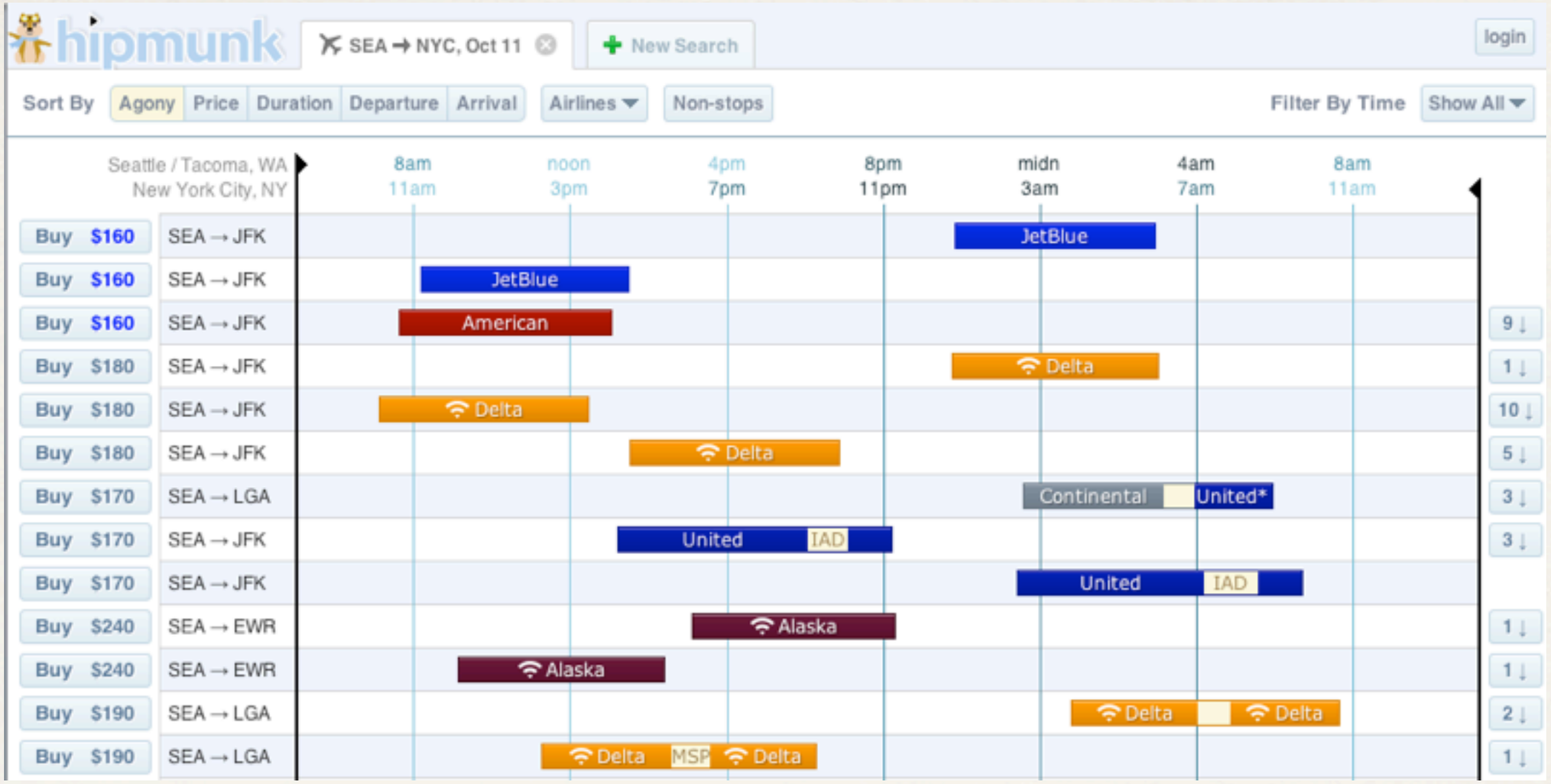
Position is Everything.

Position is everything.



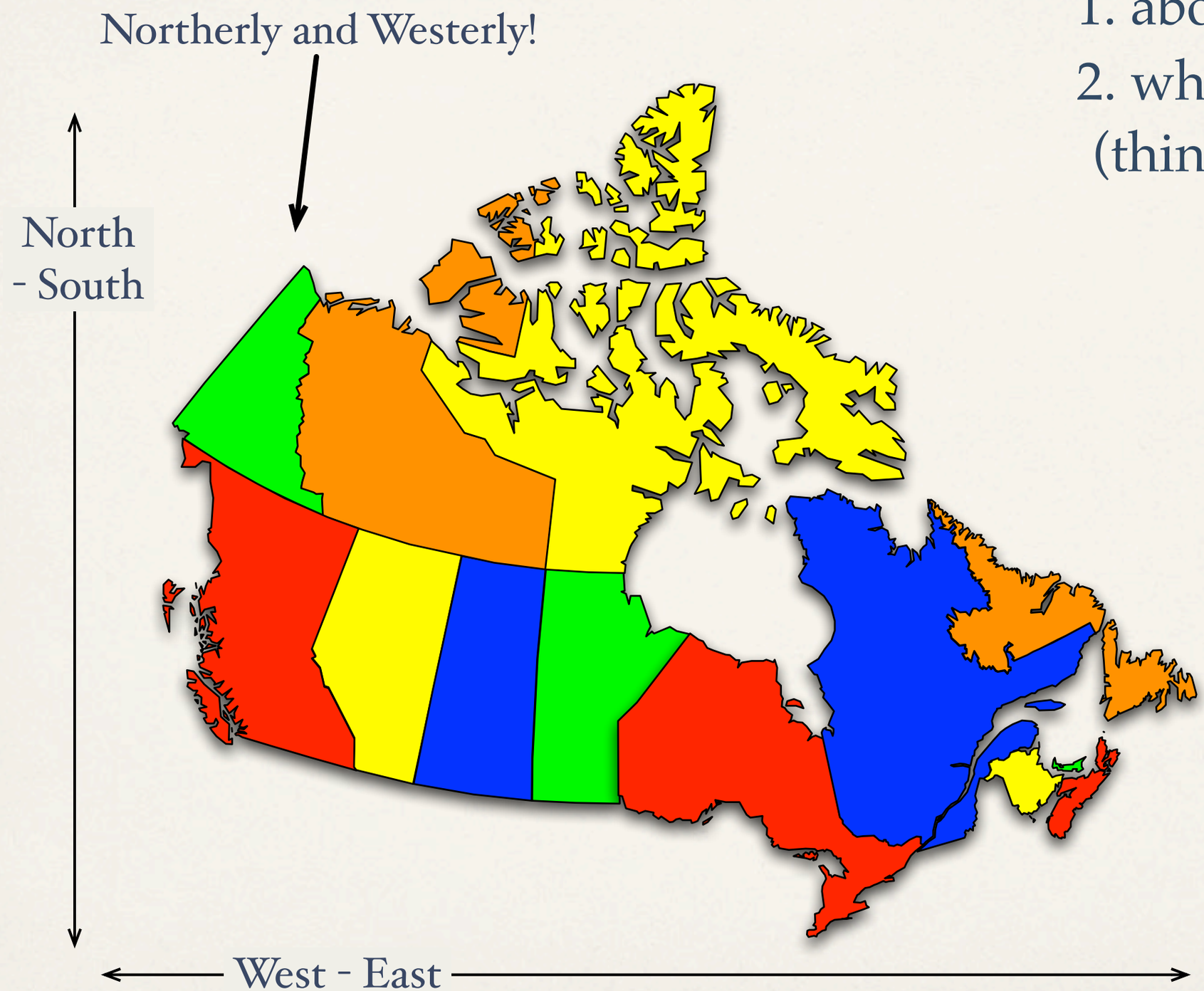
http://hipmunk.com

Position is everything.



- absolute & relative departure time (continuous)
- absolute & relative arrival time (continuous)
- absolute & relative length of trip (continuous)
- stopovers (binary)
- absolute & relative stopover duration (continuous)
- absolute & relative stopover start & stop time (continuous)
- sort order (ranked)

Axes give you information for free!



1. about targets
2. when searching
   (think grouping)

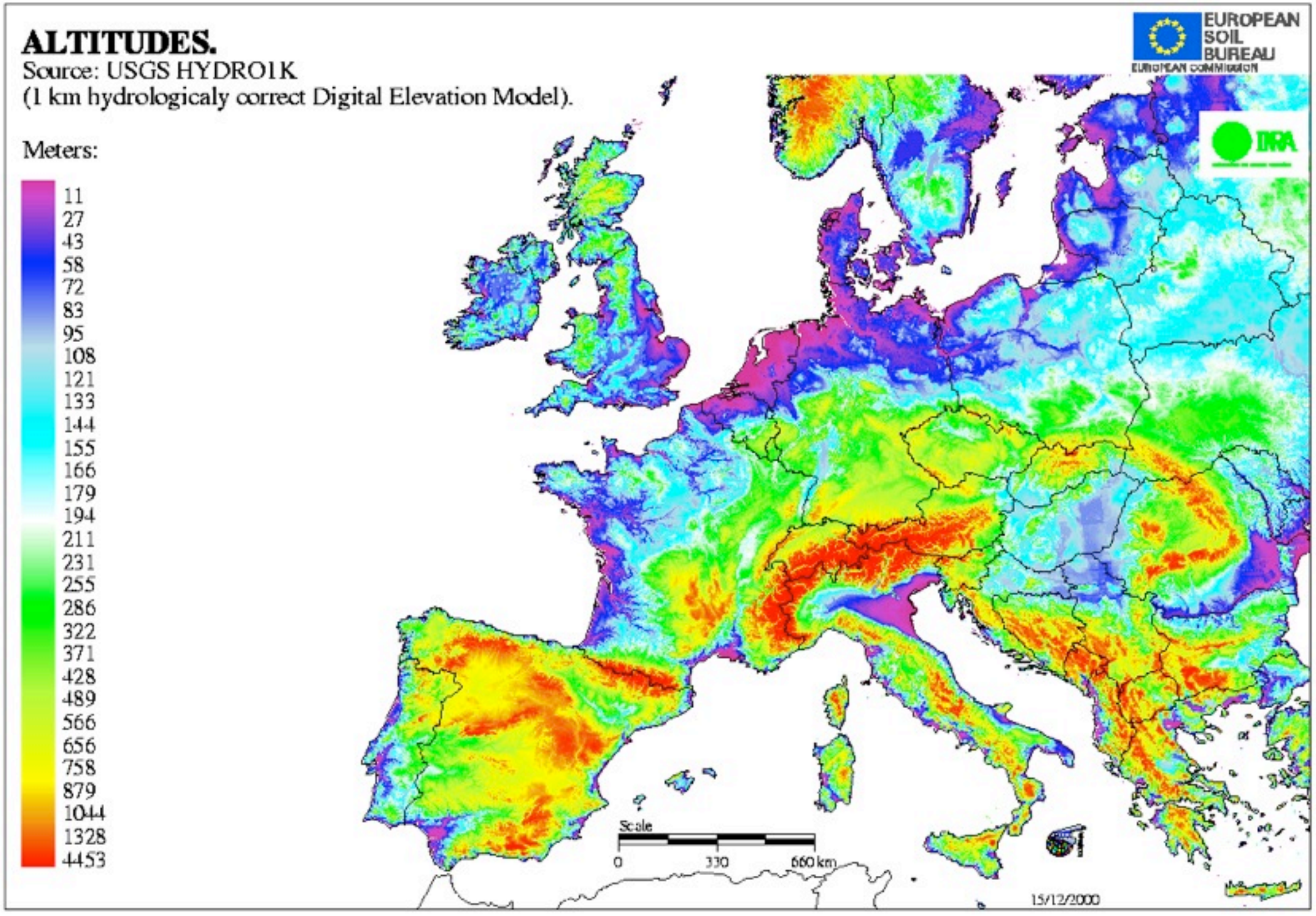
Lack of axes gives you spaghetti!

http://commons.wikimedia.org/wiki/File:Spaghetti_alle_vongole.jpg

CC SOME RIGHTS RESERVED

Color is Difficult.

Color is difficult.

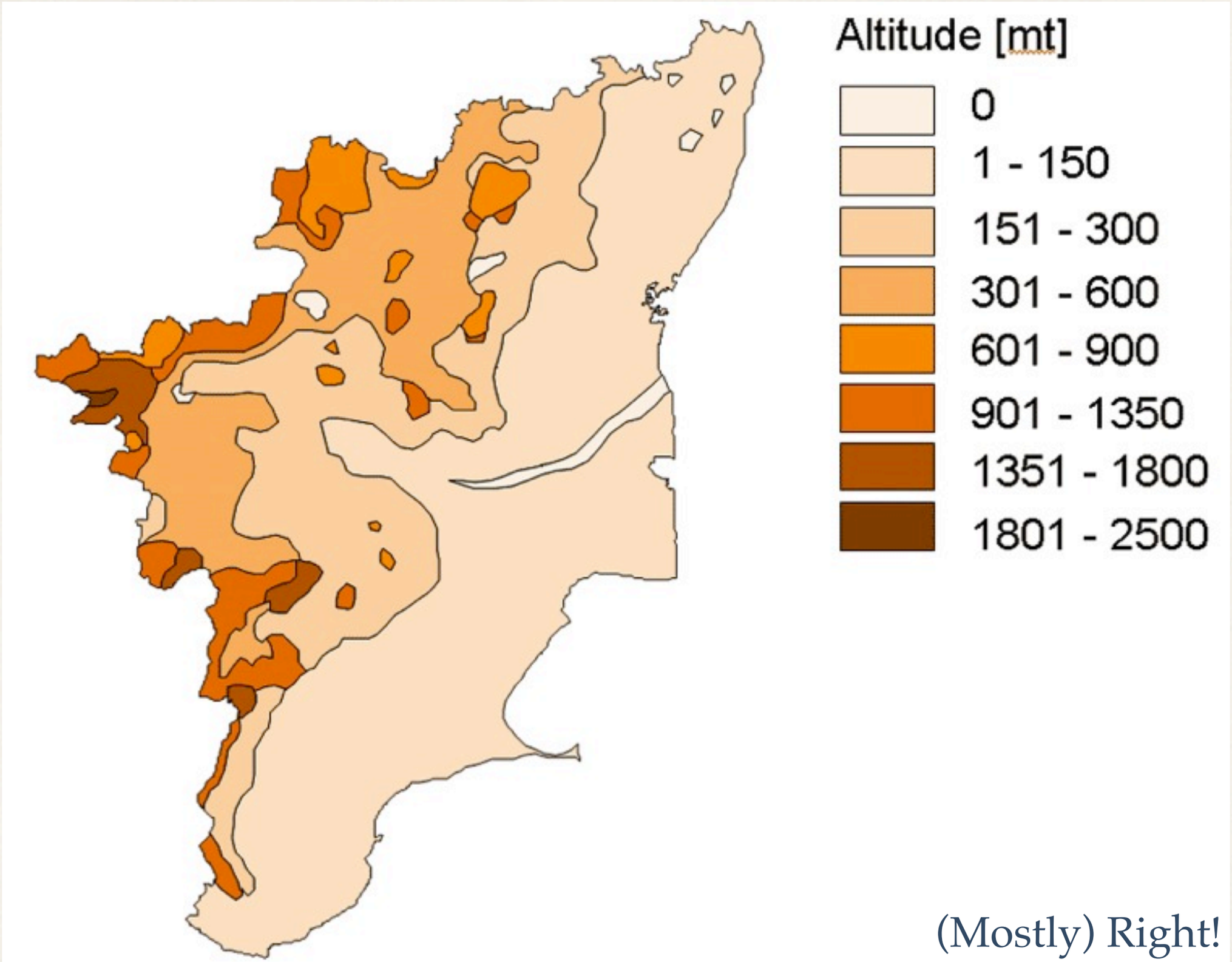


Wrong!

http://eusoils.jrc.ec.europa.eu/esdb_archive/serae/GRIMM/erosion/inra/europe/analysis/maps_and_listings/web_erosion/maps_and_listings/altitude_a3.gif

Color is not ordered.

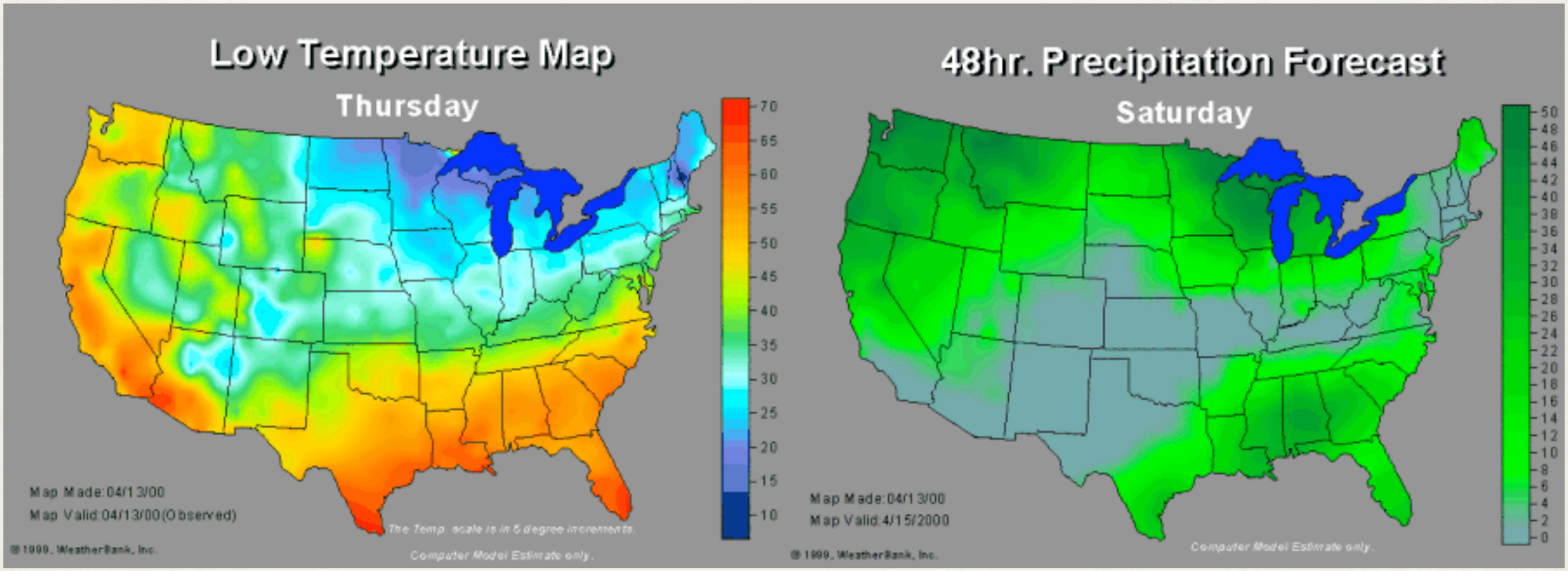
Color is difficult.



(Mostly) Right!

http://mapsof.net/uploads/static-maps/topographic_(altitude)_map_tamil_nadu.png

Color is difficult.



Not bad...

http://www.goldensoftware.com/gallery/gallery-11.shtml

Color is meaningful.

Gender

Nationality

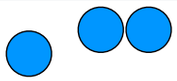
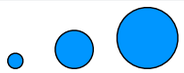
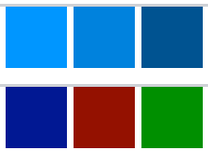
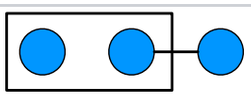
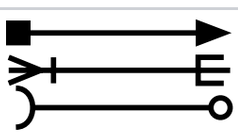
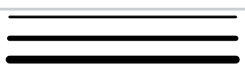
Politics

Religion

Morality

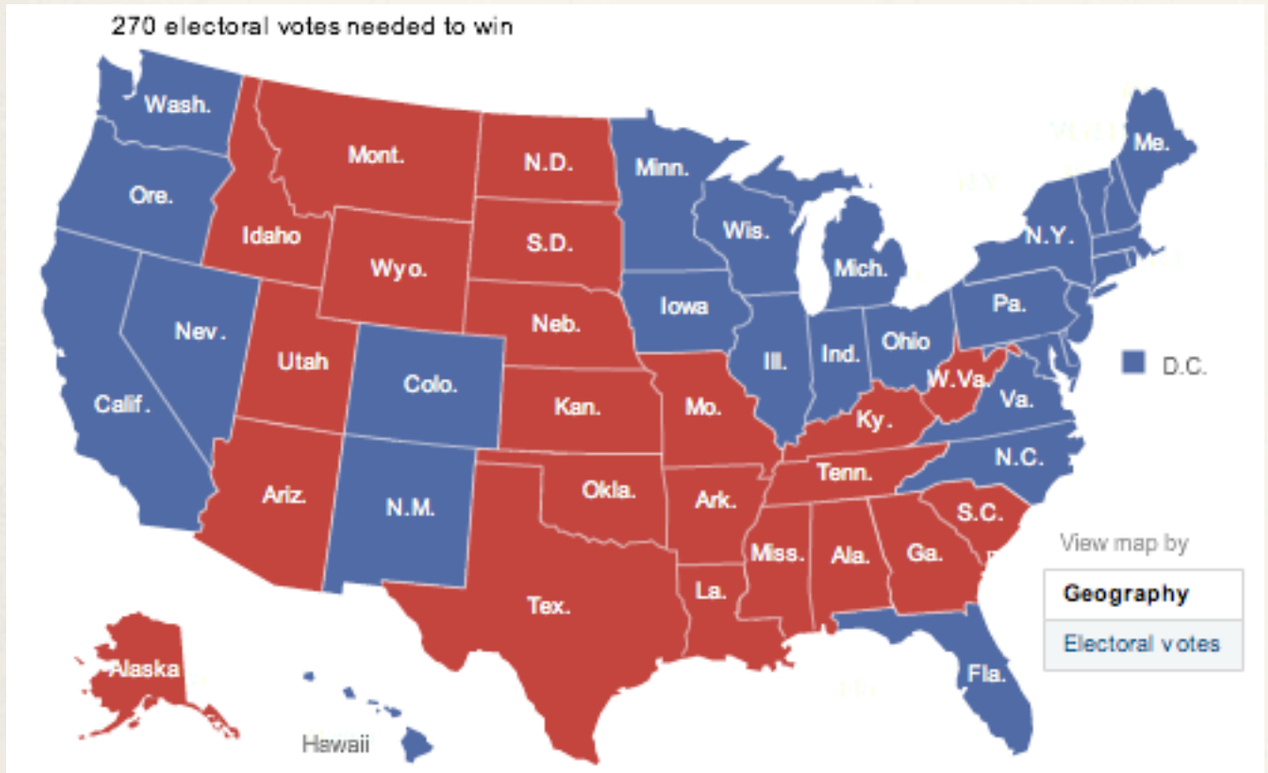
Nature

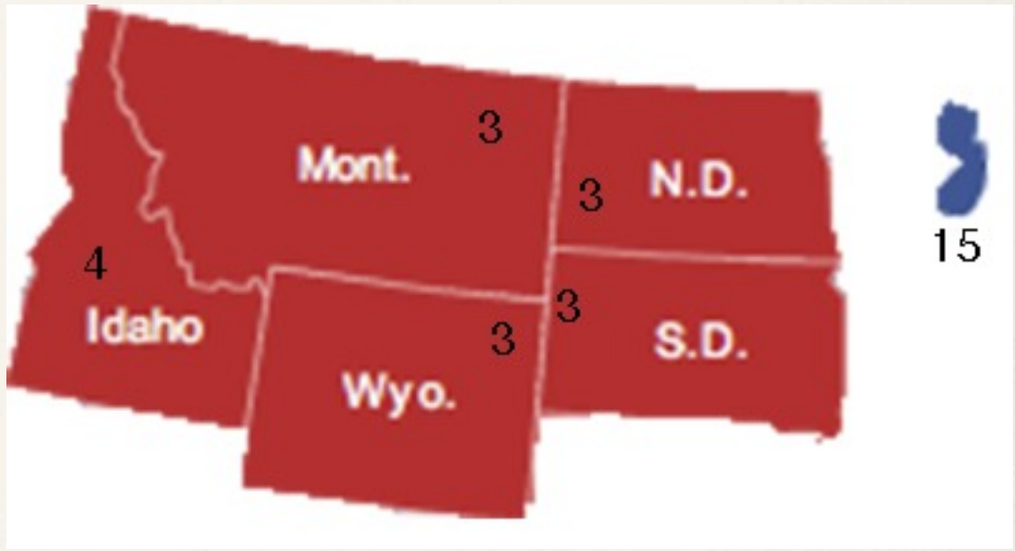
Appropriate encodings

Example	Encoding	Ordered	Useful values	Quantitative	Ordinal	Categorical	Relational
	position, placement	yes	infinite	Good	Good	Good	Good
1, 2, 3; A, B, C	text labels	optional (alphabetical or numbered)	infinite	Good	Good	Good	Good
	length	yes	many	Good	Good		
	size, area	yes	many	Good	Good		
	angle	yes	medium	Good	Good		
	pattern density	yes	few	Good	Good		
	weight, boldness	yes	few		Good		
	saturation, brightness	yes	few		Good		
	color	no	few (< 20)			Good	
	shape, icon	no	medium			Good	
	pattern texture	no	medium			Good	
	enclosure, connection	no	infinite			Good	Good
	line pattern	no	few				Good
	line endings	no	few				Good
	line weight	yes	few		Good		

Use defaults.

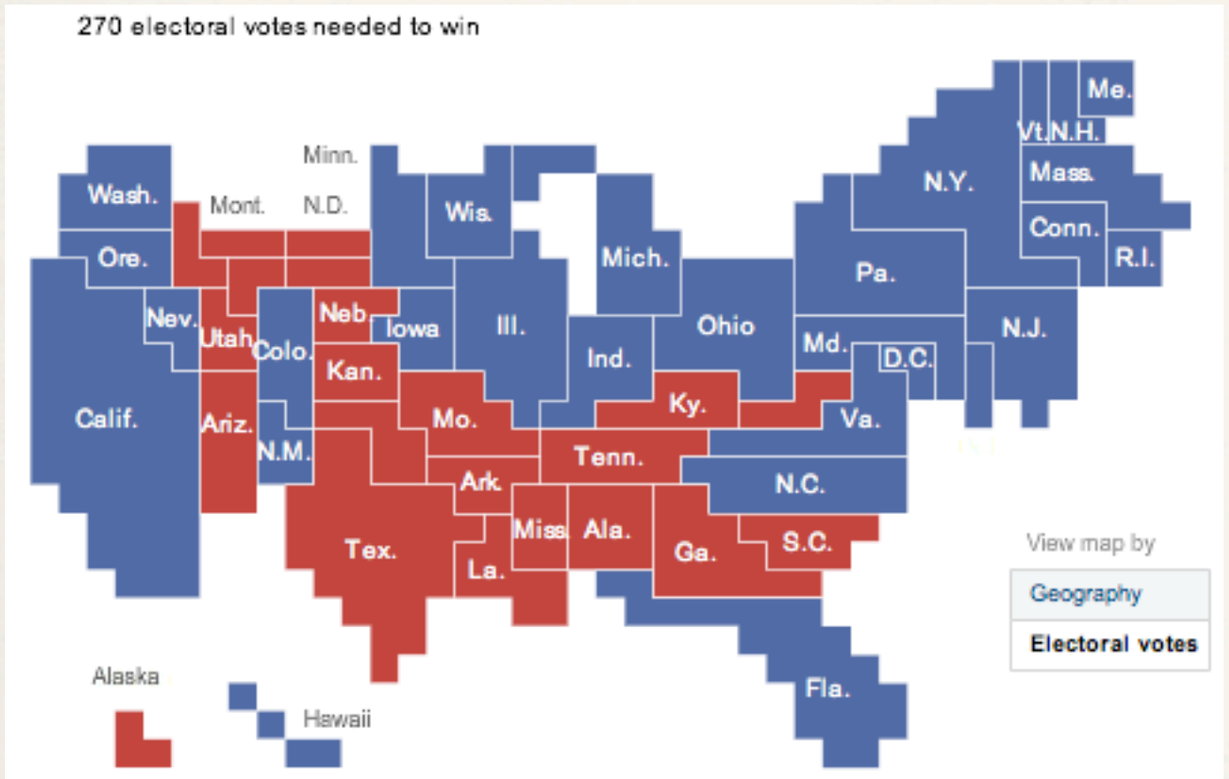
Use defaults.

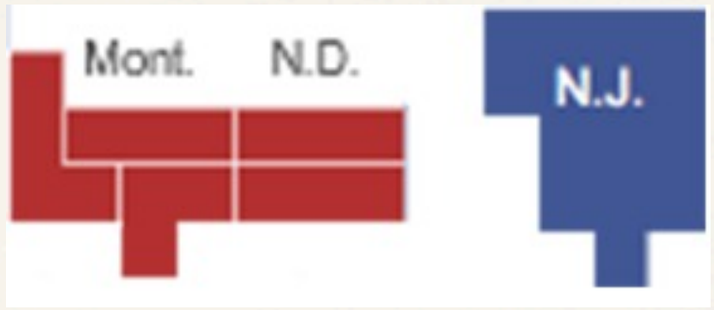




http://elections.nytimes.com/2008/president/whos-ahead/key-states/map.html

Unless...





http://elections.nytimes.com/2008/president/whos-ahead/key-states/map.html

Unless you've got something better.

Design strategies

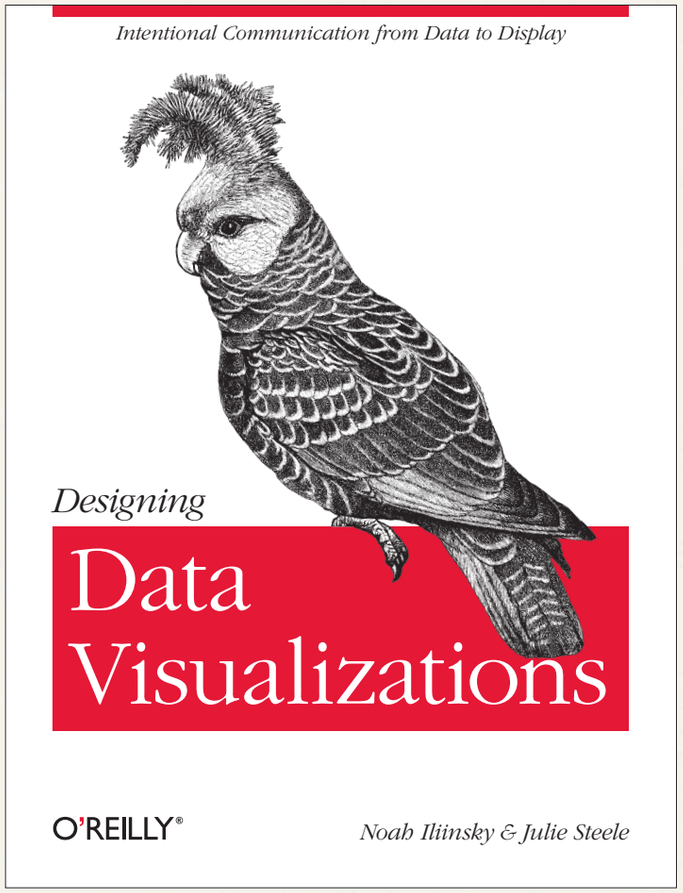
* Limit the data you include
* Use position for your most important relationship(s)
* Try different axes
* Consider default formats
* Use color for categories, not rank
* Encode other data and relationships with appropriate properties

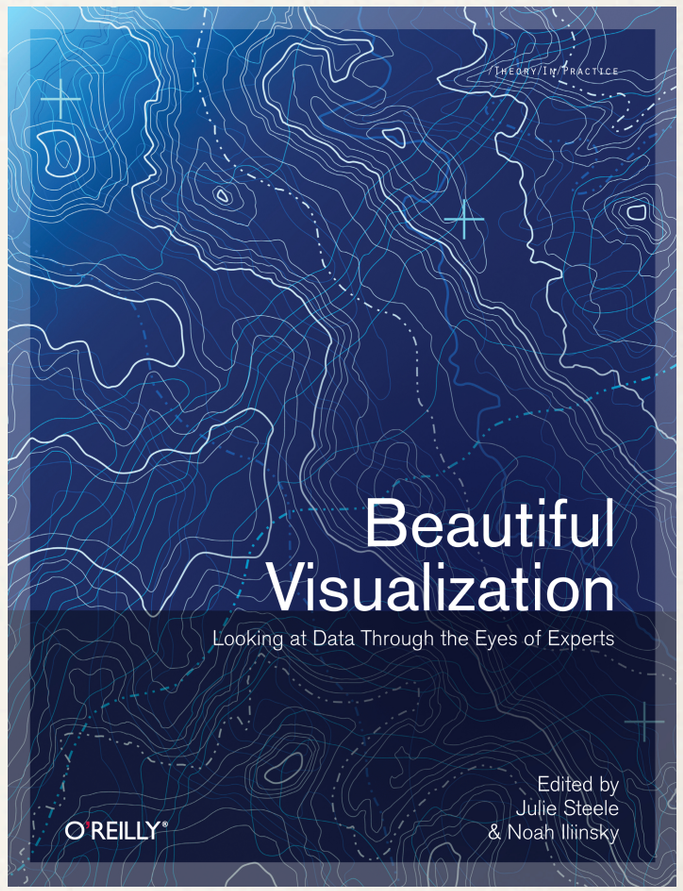
Tools

- D3.js / protovis: structured frameworks
- processing: flexible, great for data art
- R + ggplot2: stats & analytics
- Tableau: visual exploration & analytics

Thank you!

@noahi
gmail: iliinsky
ComplexDiagrams.com

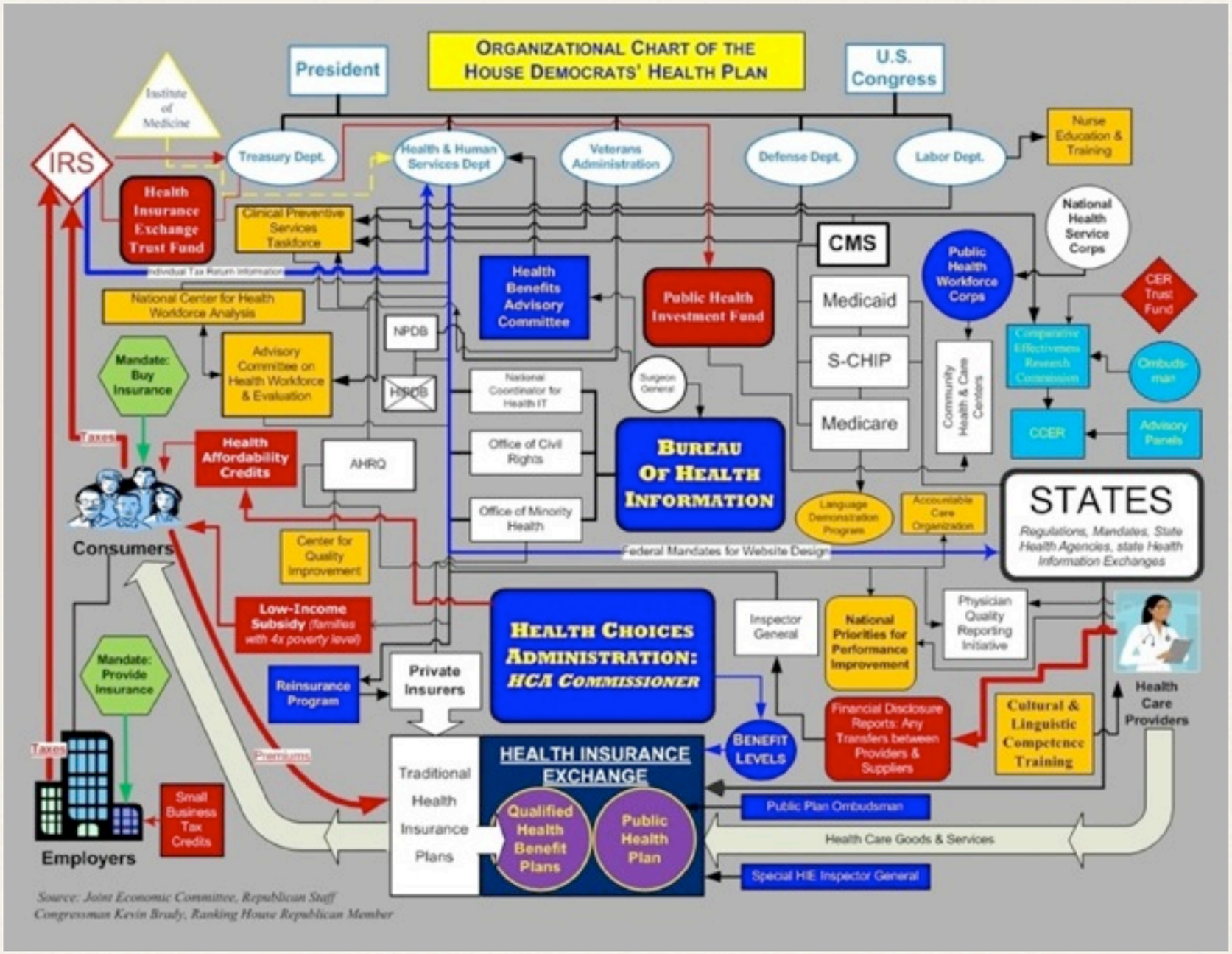




How to do it
(this talk)

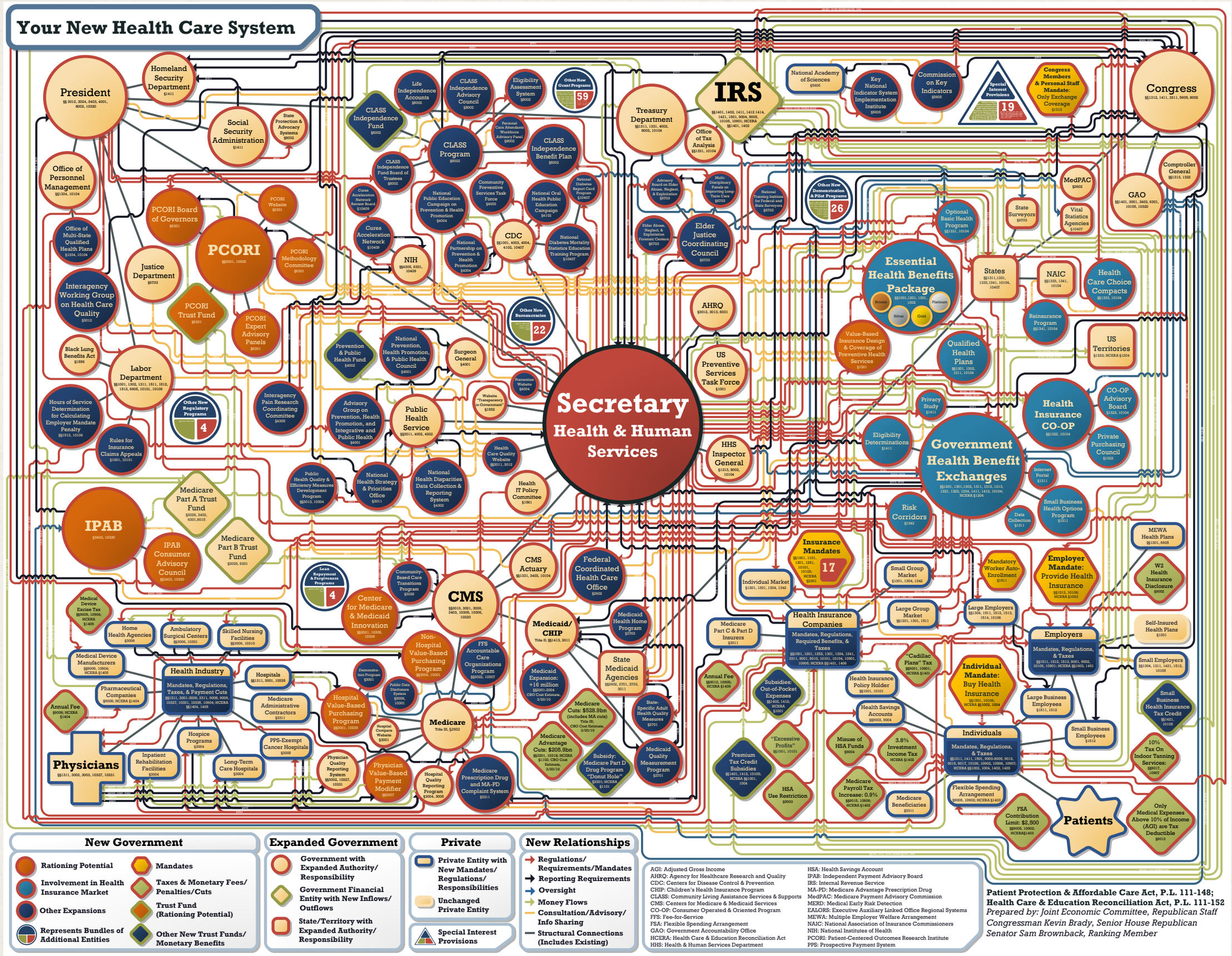
How they did it
(20 case studies)

Visualization for *persuasion* (or *propaganda*)



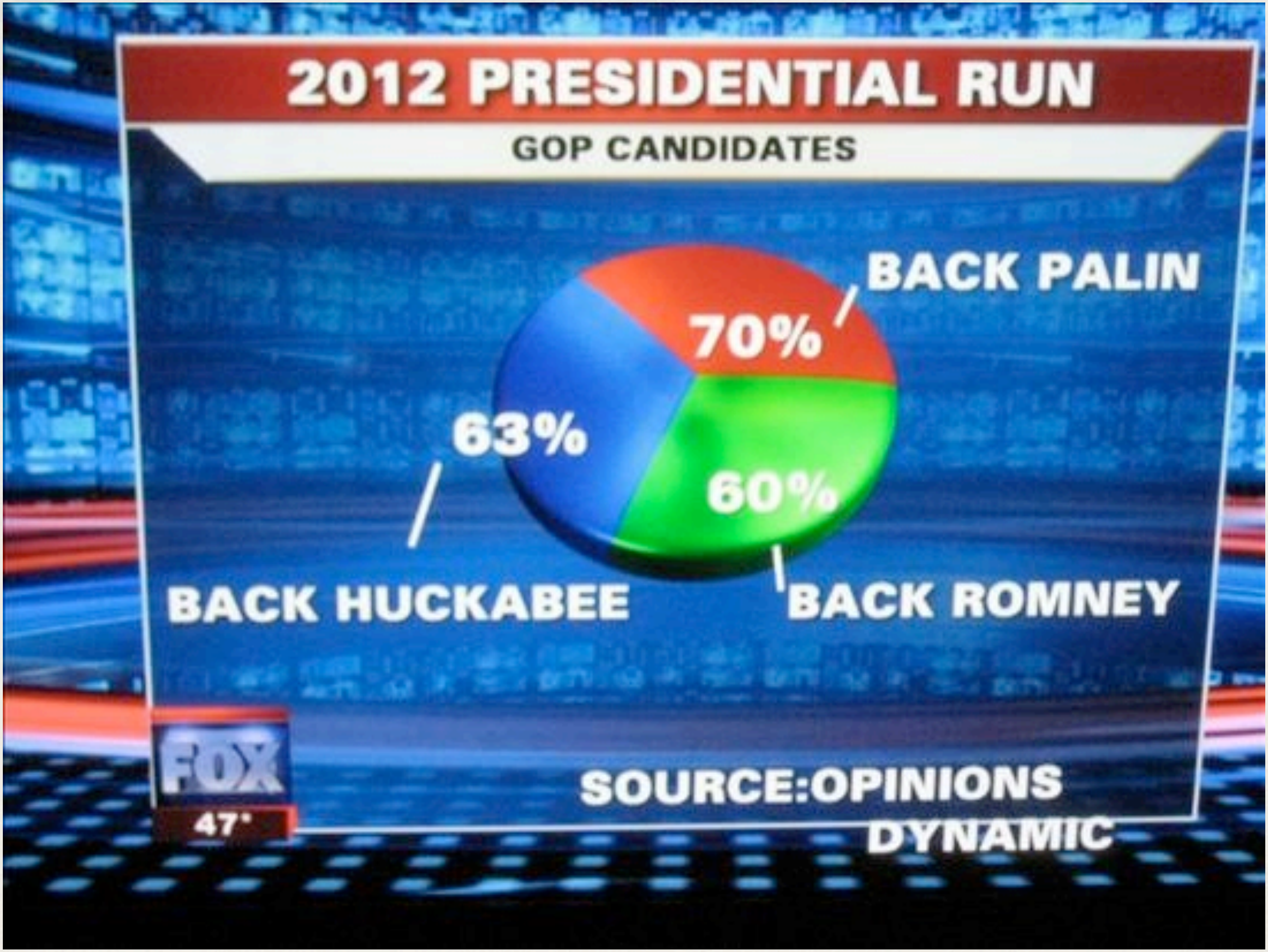
http://www.house.gov/apps/list/press/tx08_brady/71509_hc_chart.html

Visualization for *persuasion* (or *propaganda*)



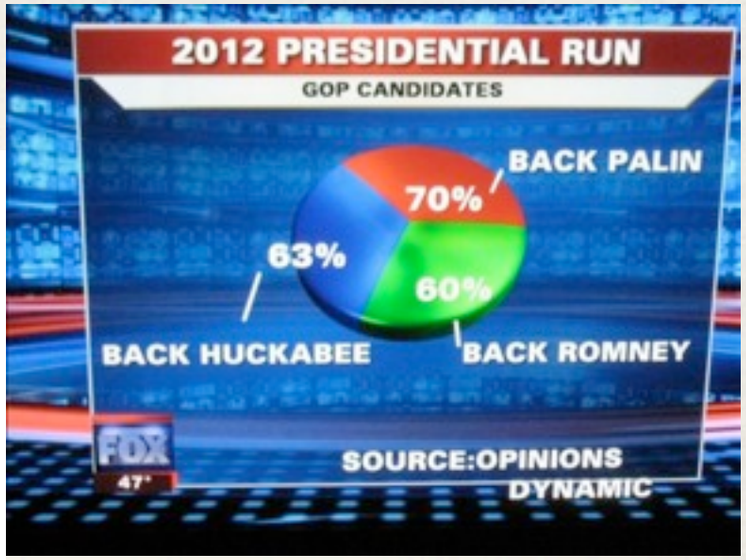
http://www.house.gov/jec/news/2010/obamacare_chart.pdf

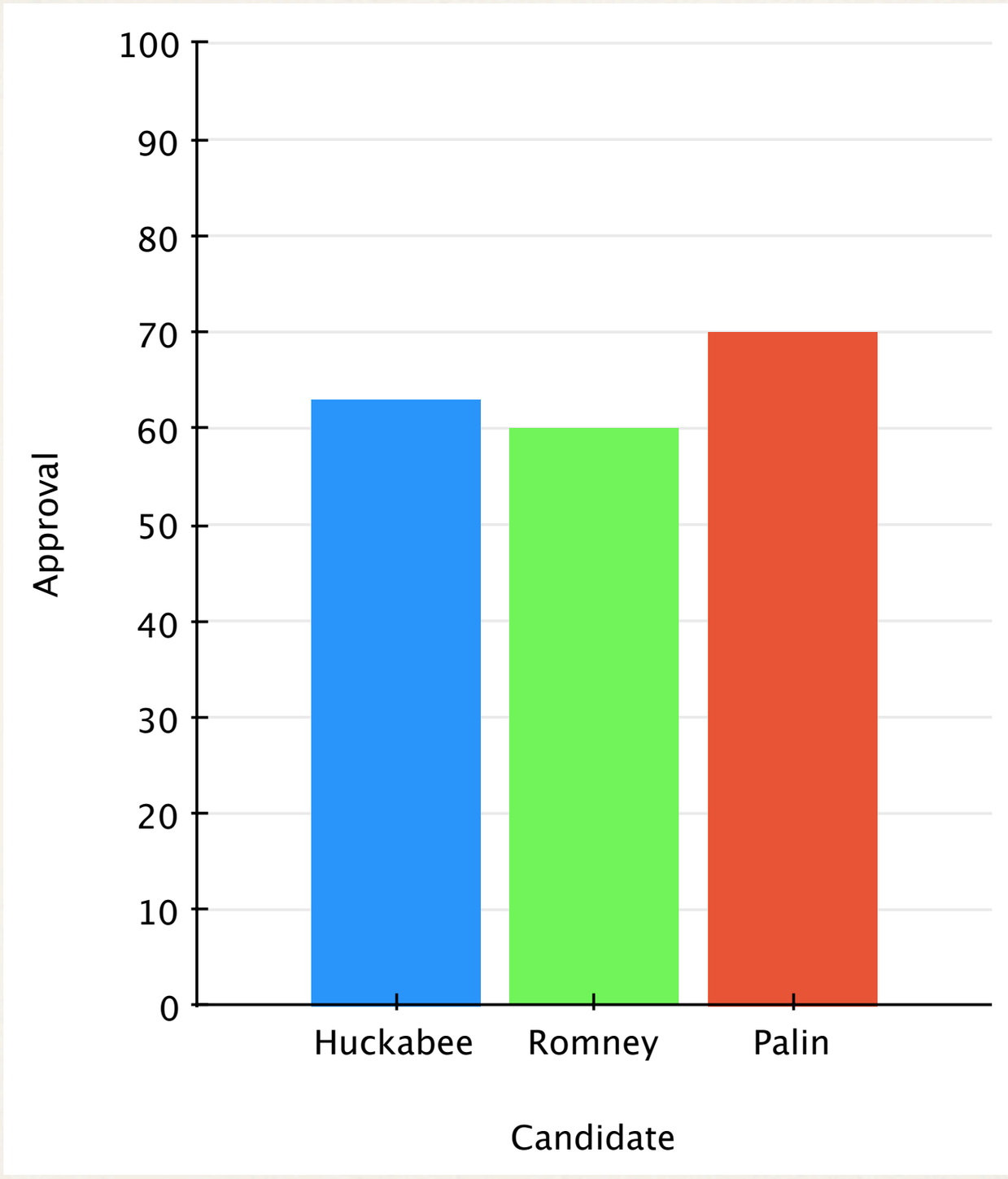
Thursday, December 8, 2011



Wrong!

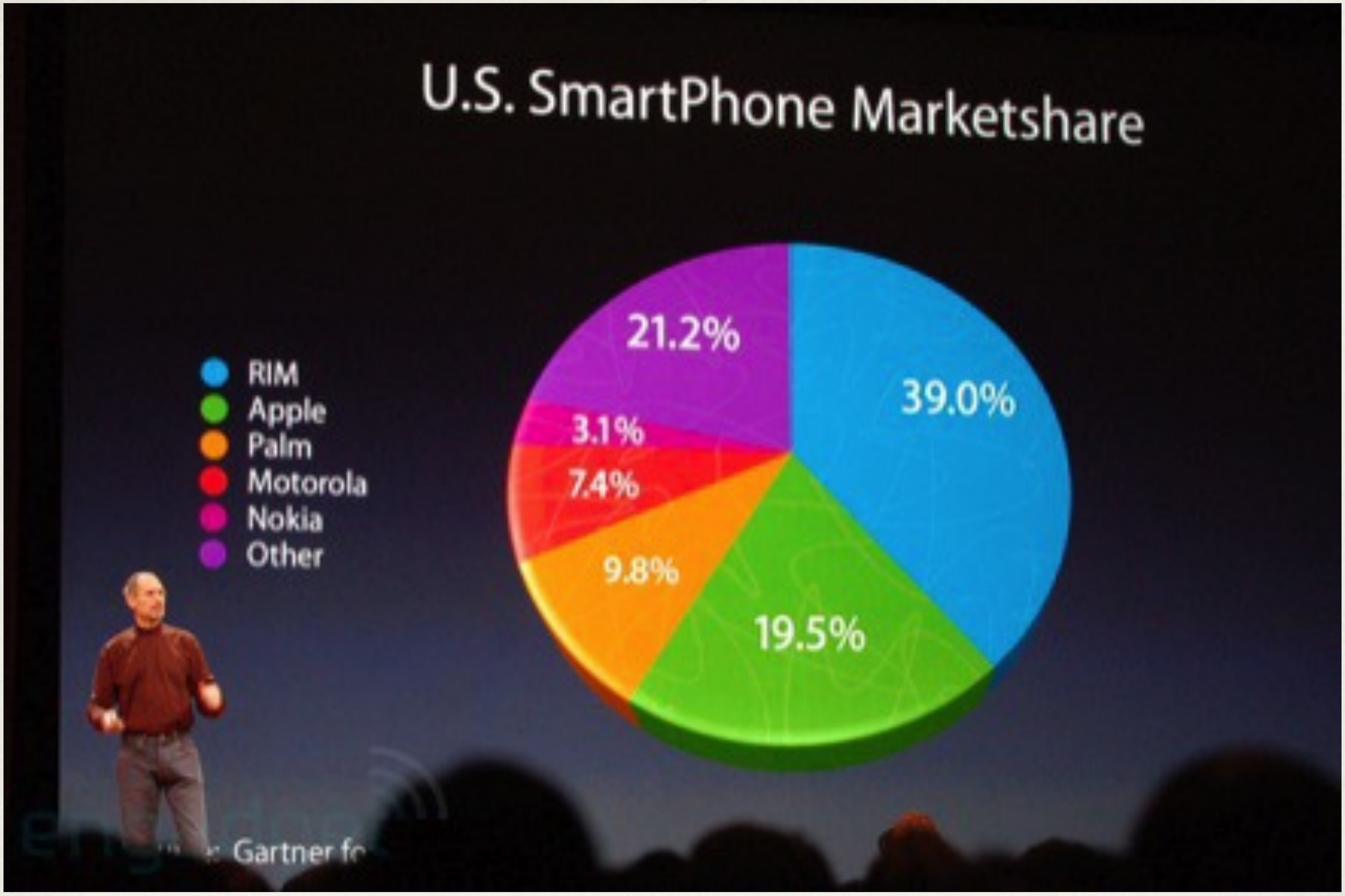
http://thinkprogress.org/2009/11/23/fox-pie-chart/



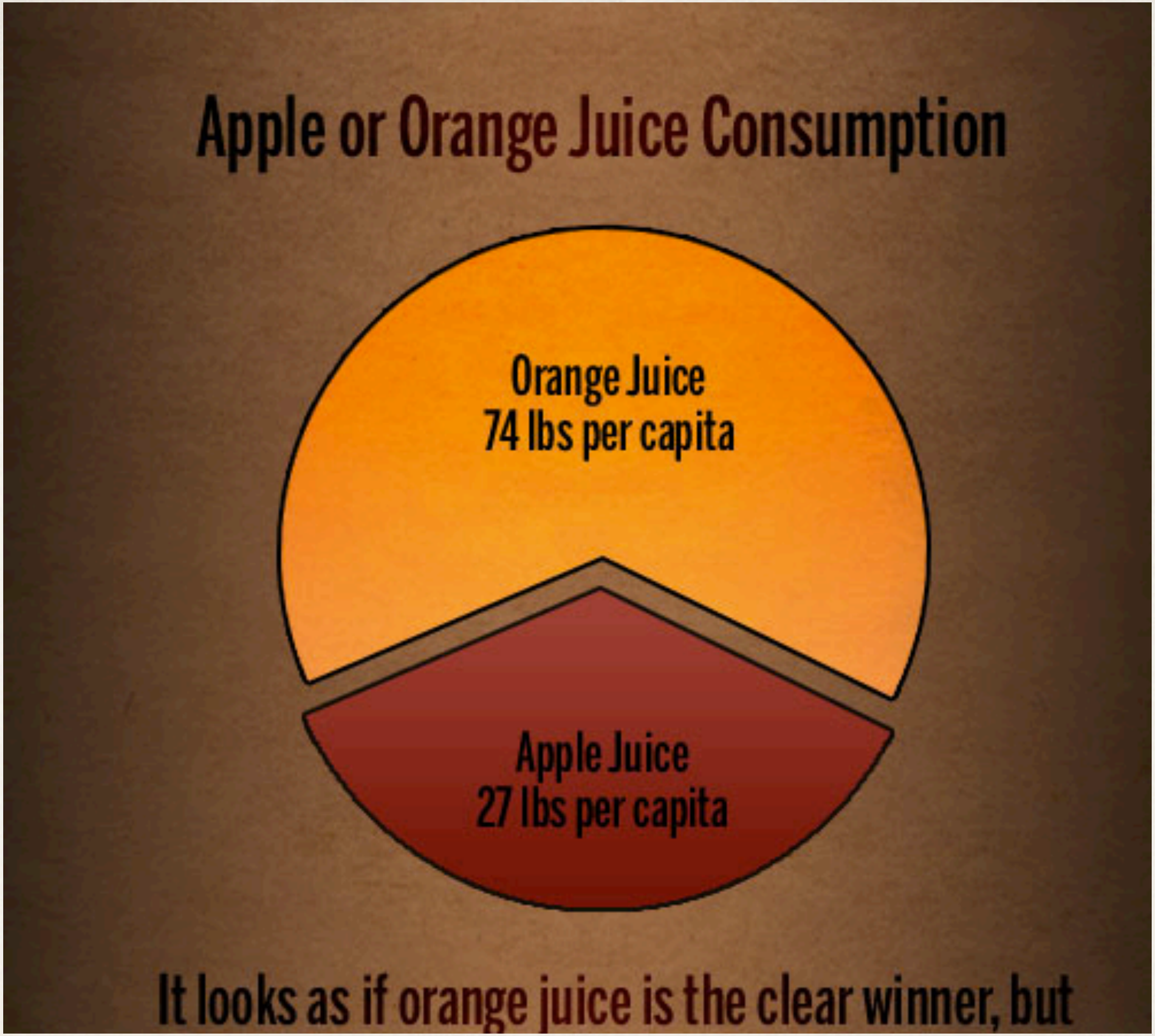


Right!

http://thinkprogress.org/2009/11/23/fox-pie-chart/

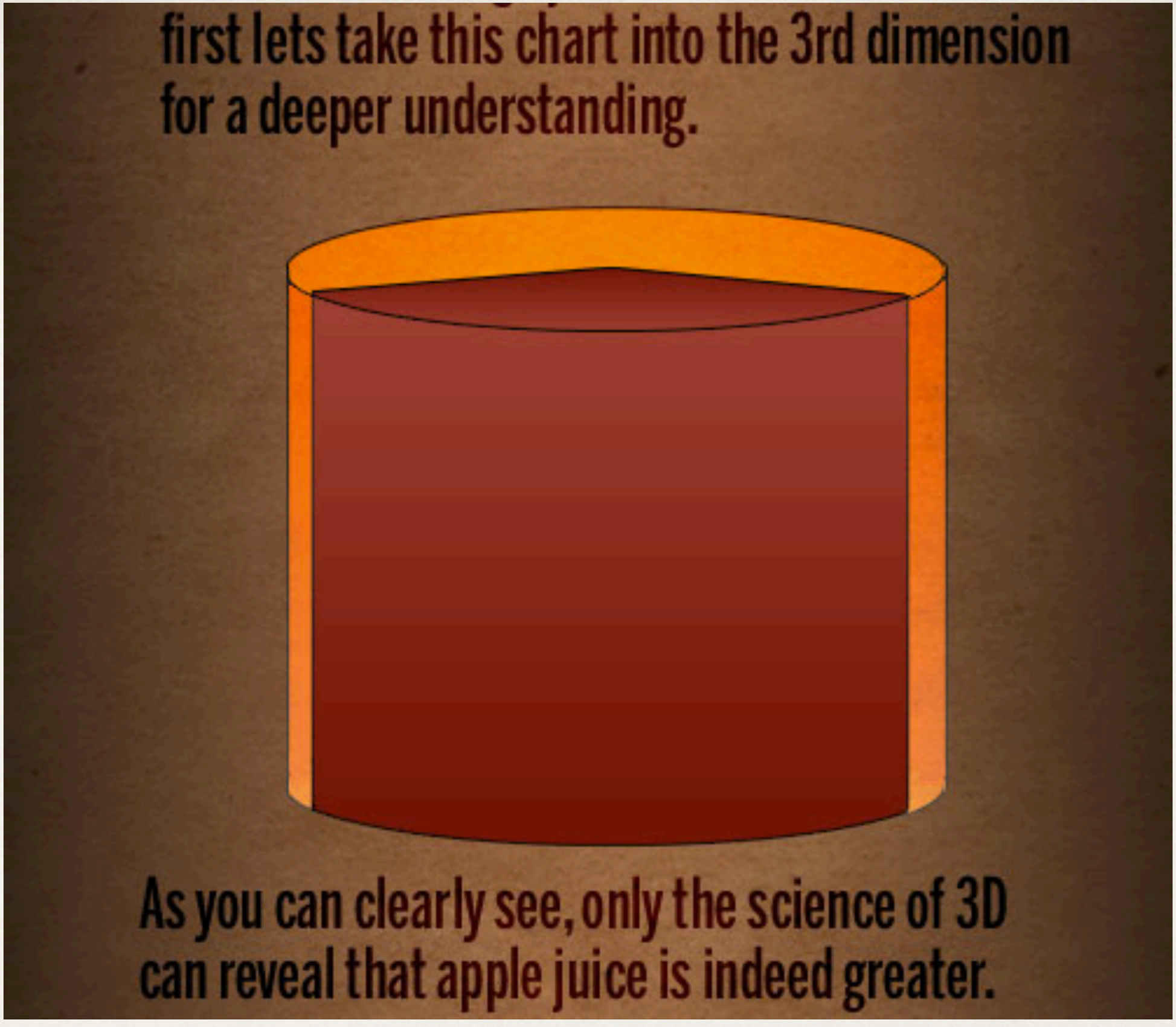

U.S. SmartPhone Marketshare
RIM
Apple
Palm
Motorola
Nokia
Other
21.2%
39.0%
3.1%
7.4%
9.8%
19.5%
Gartner fo

Wrong!



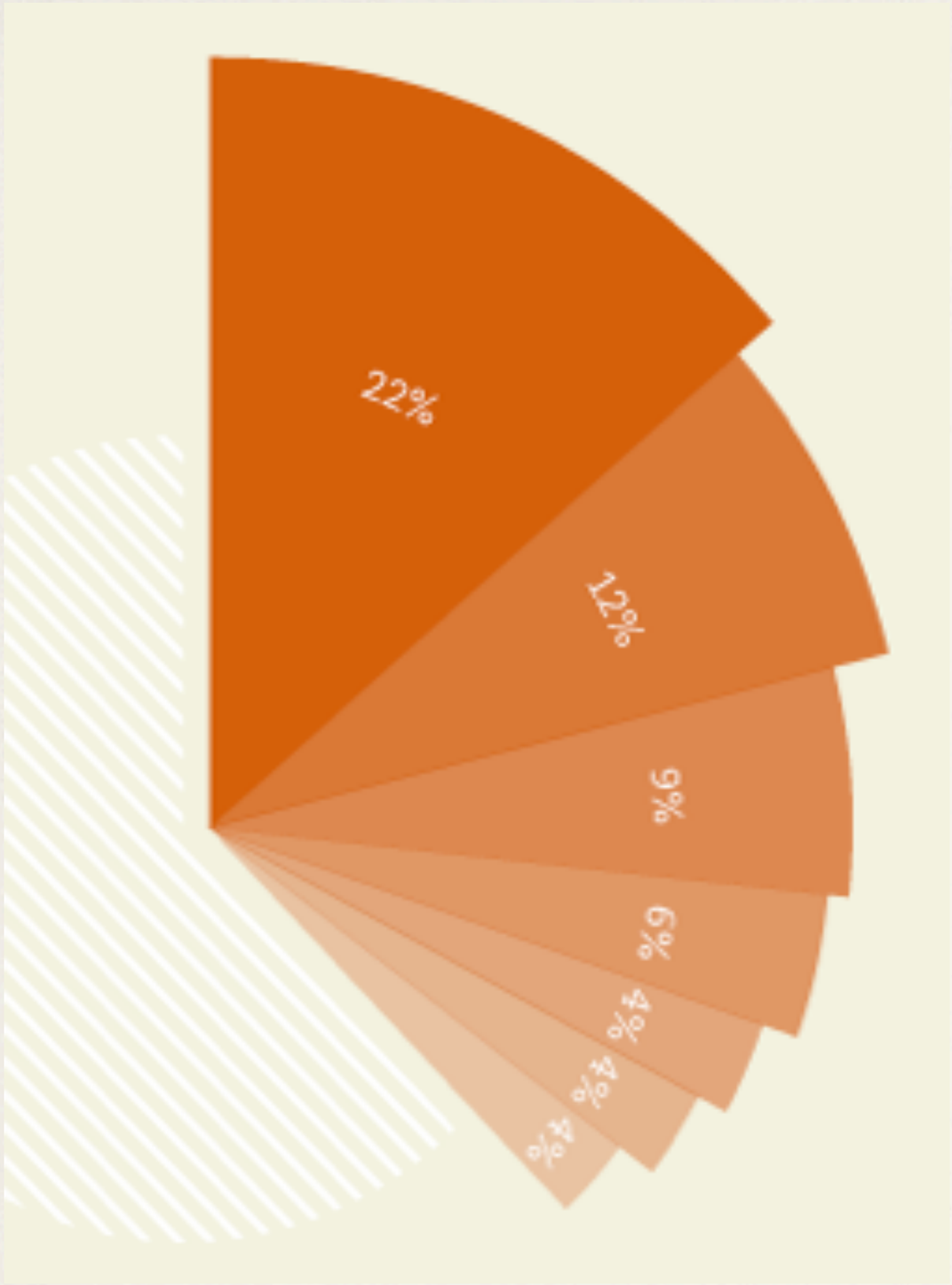
Right...

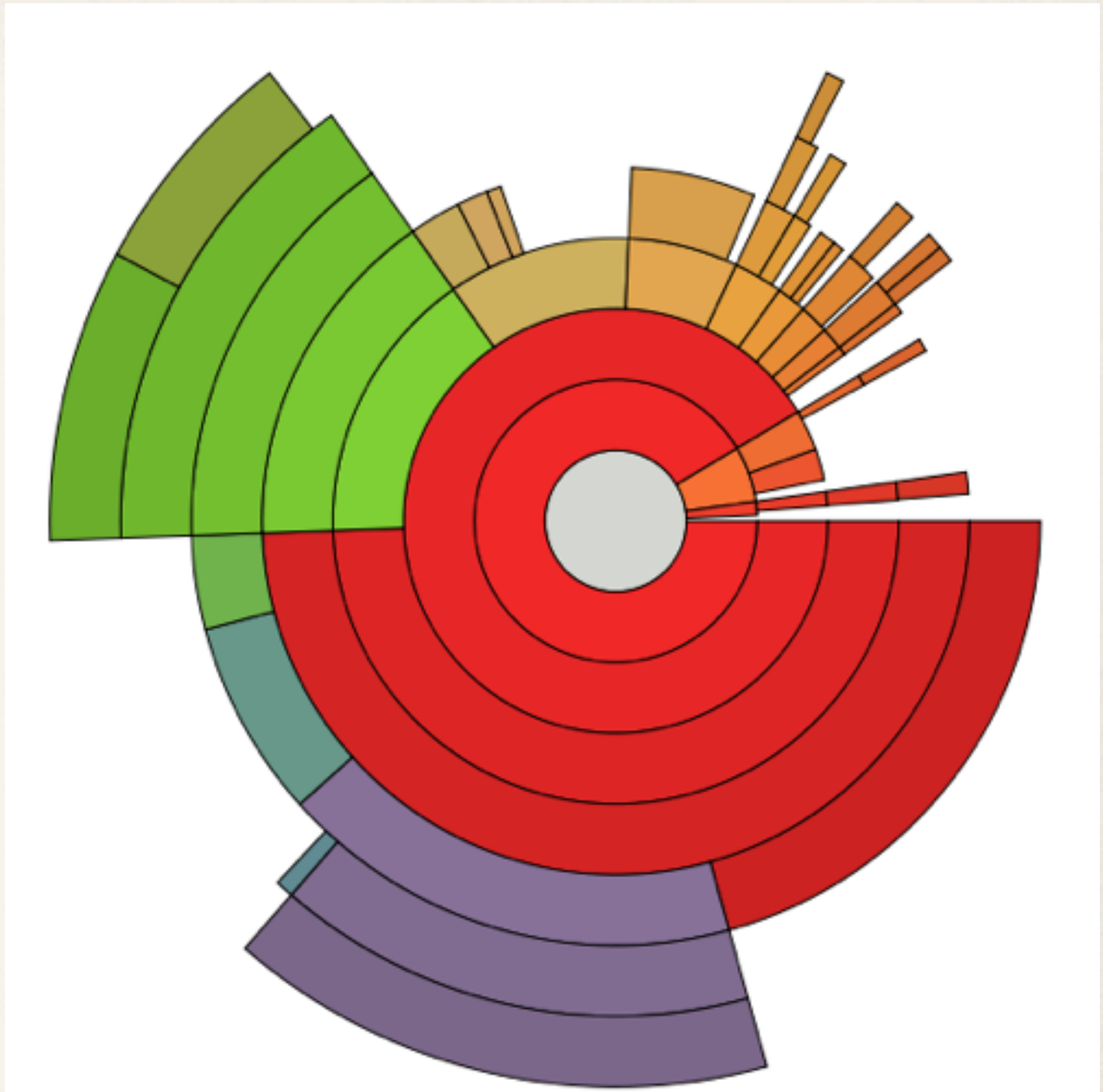
http://www.smarter.org/research/apples-to-oranges/



Wrong!

http://www.smarter.org/research/apples-to-oranges/

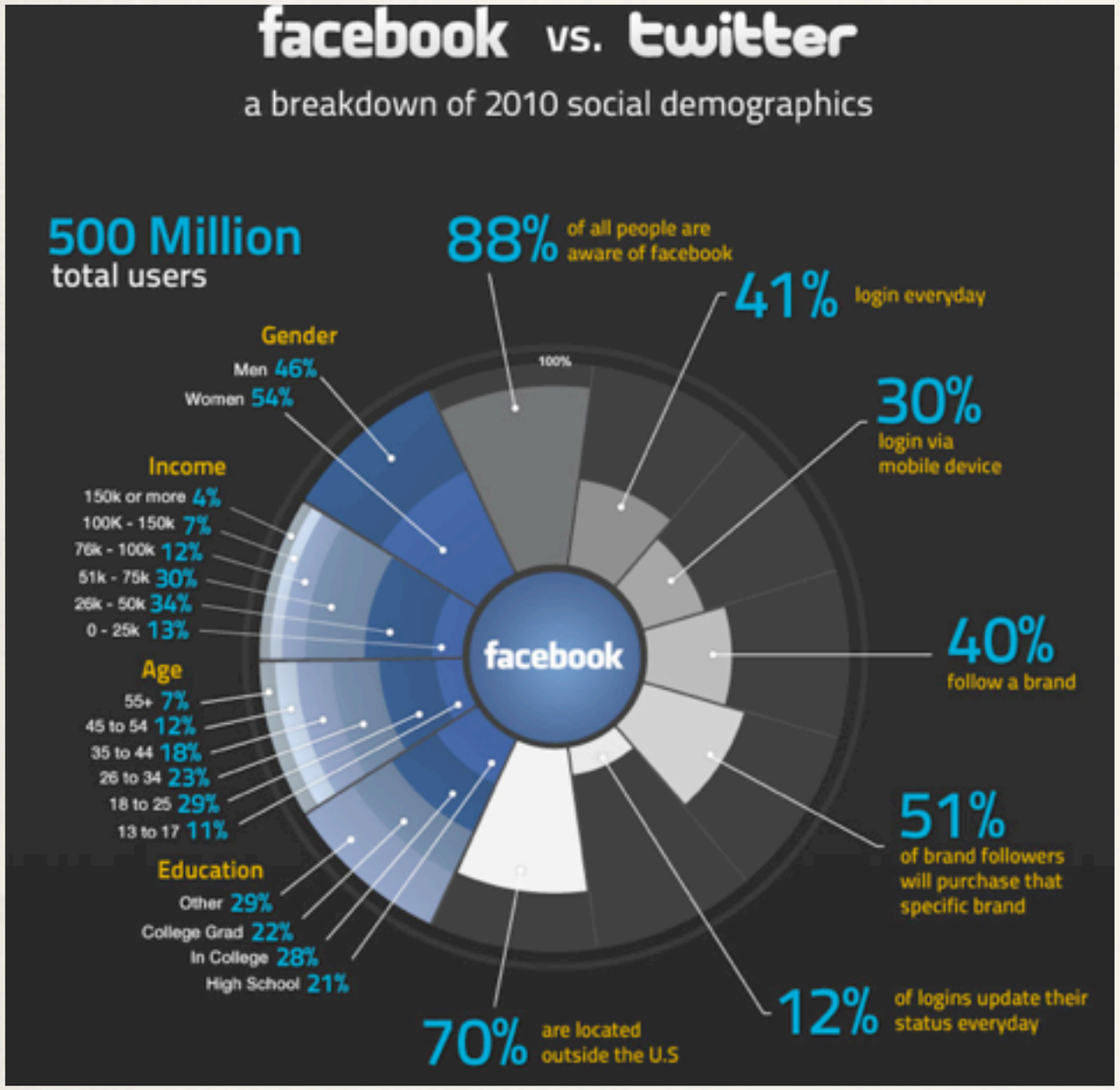


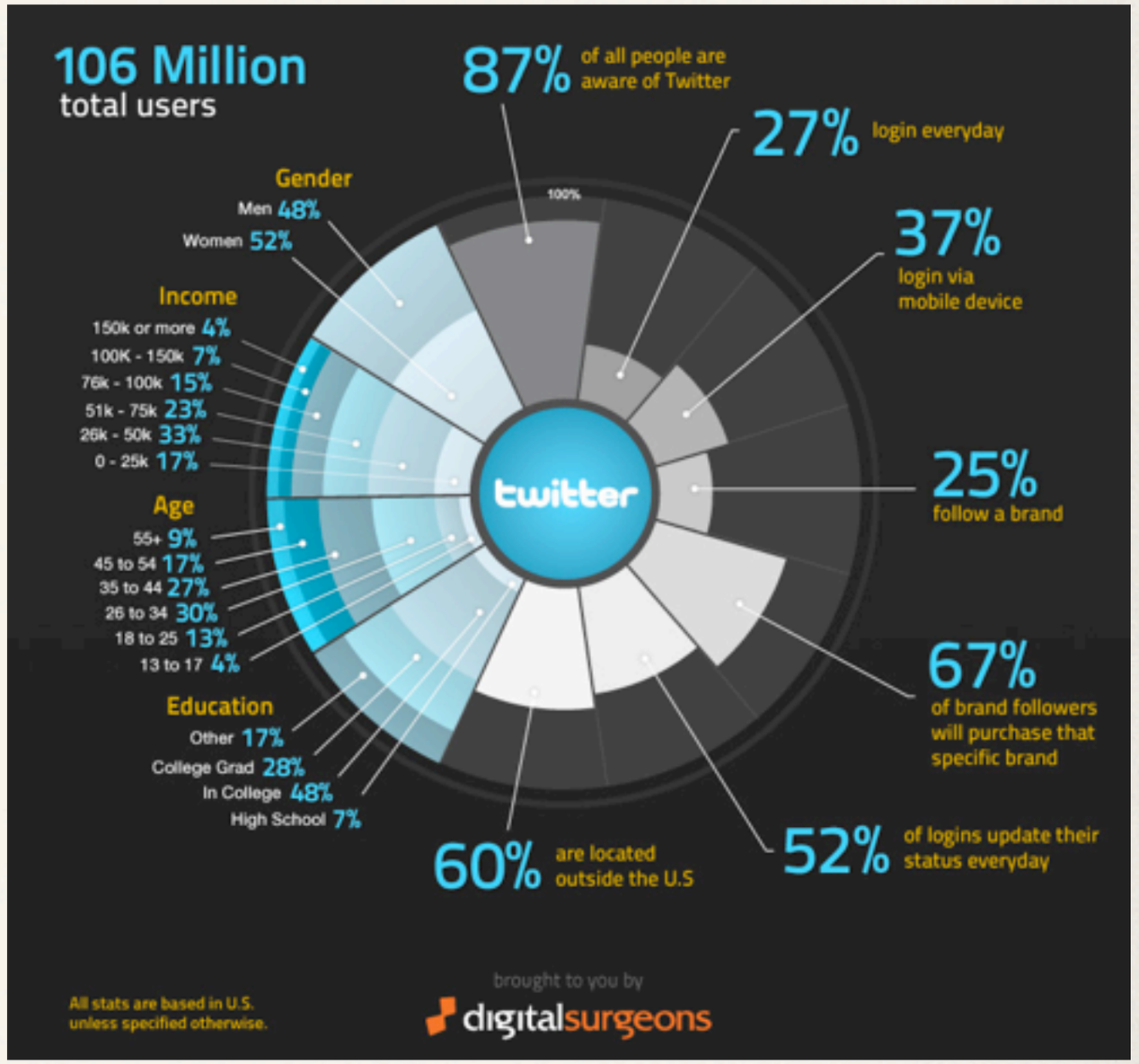


Wrong!

http://www.coursehero.com/blog/2011/08/25/infographic-the-pulse-of-education-around-the-world/

http://en.wikipedia.org/wiki/File:Disk_usage_%28Boabab%29.png



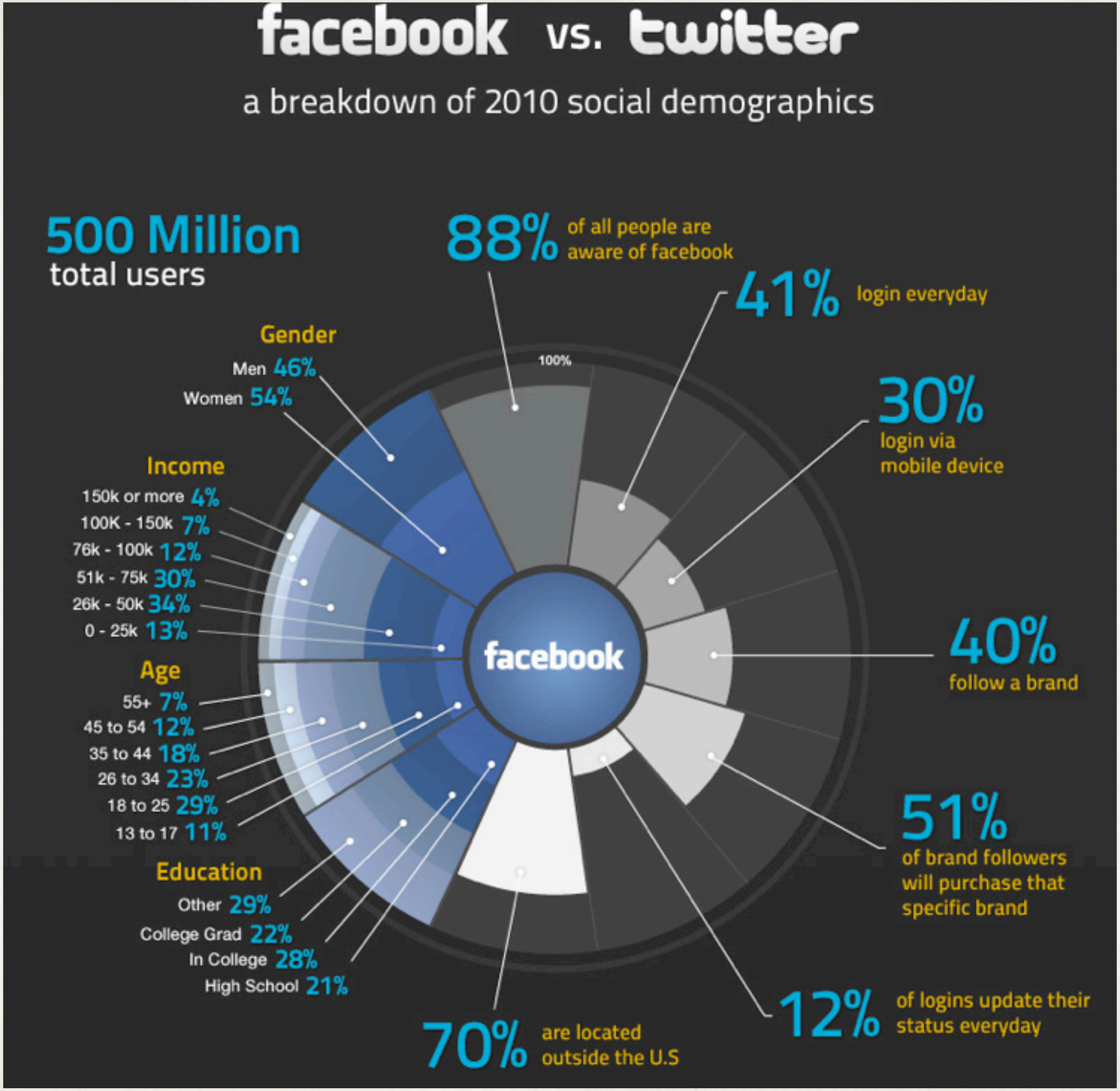


Wrong!

http://www.digitalsurgeons.com/facebook-vs-twitter-infographic/

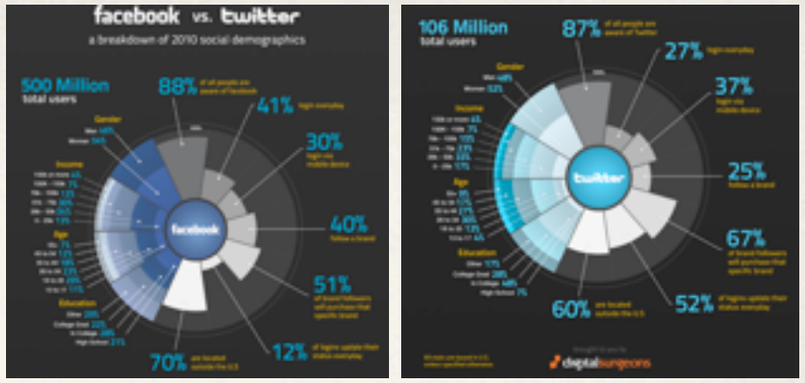
facebook vs. twitter

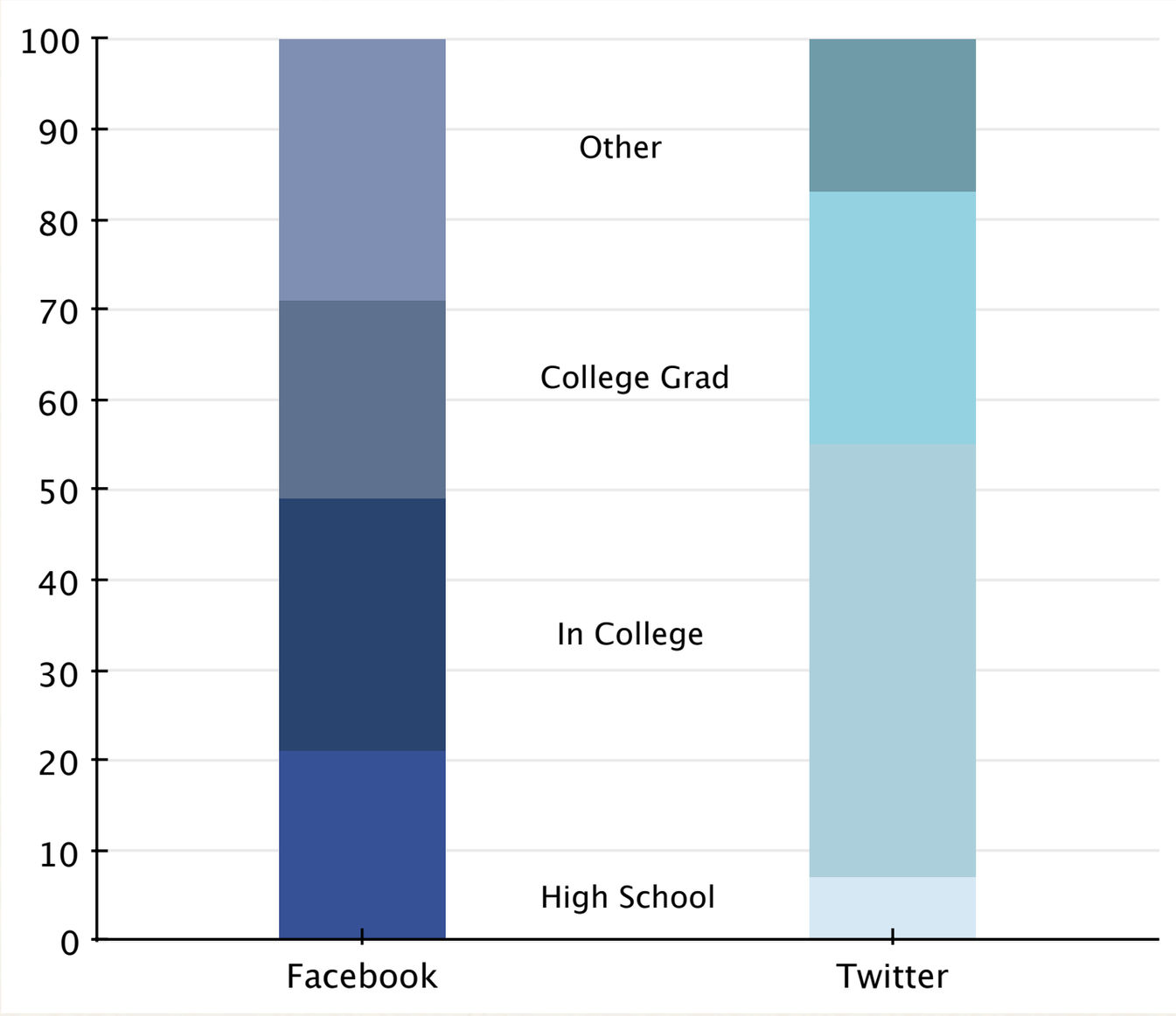
a breakdown of 2010 social demographics



http://www.digitalsurgeons.com/facebook-vs-twitter-infographic/

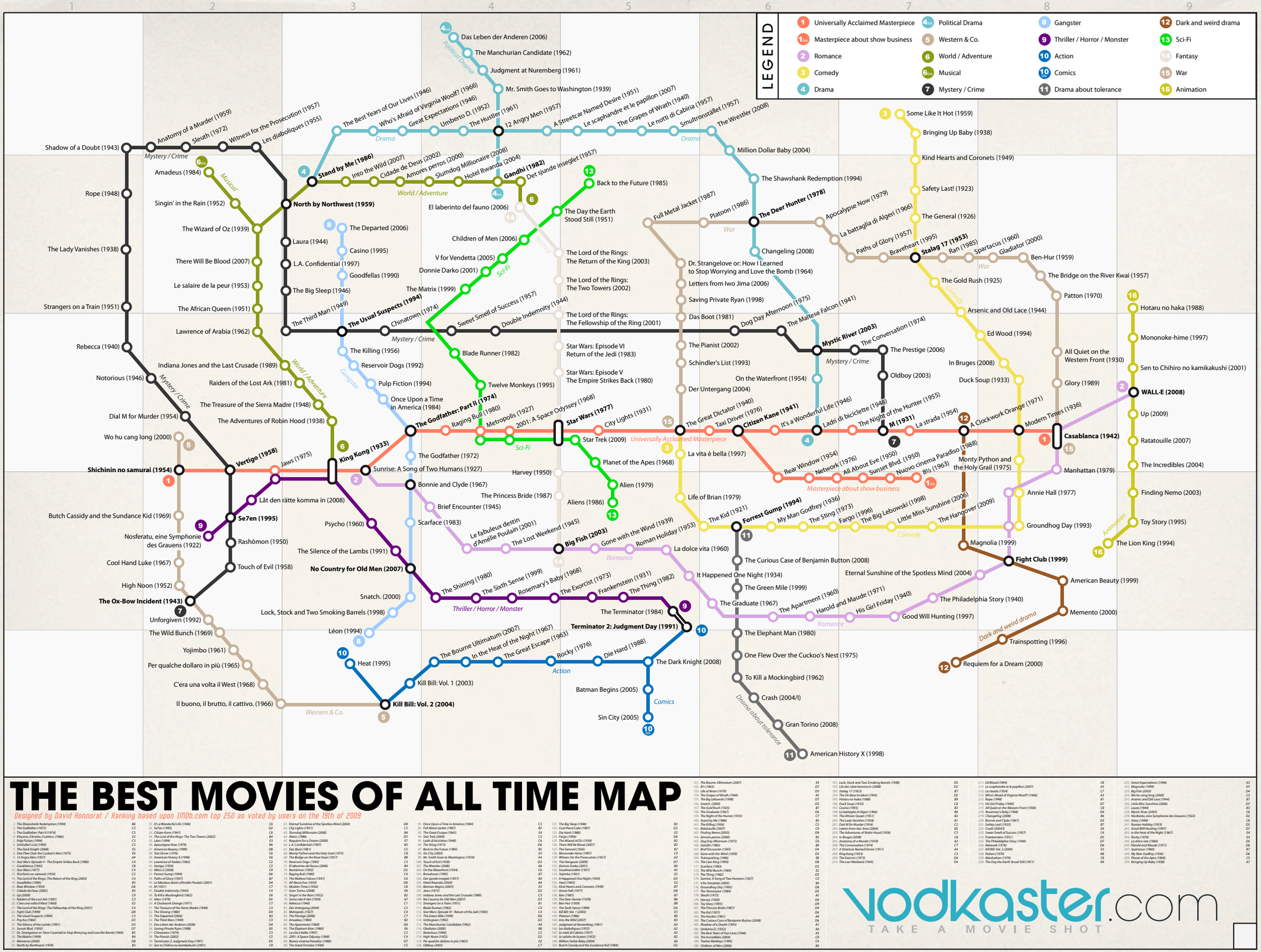
Use a format that fits the data





Better.

http://www.digitalsurgeons.com/facebook-vs-twitter-infographic/



THE BEST MOVIES OF ALL TIME MAP

Designed by David Honnorat / Ranking based upon IMDb.com top 250 as voted by users on the 19th of 2009

1. The Shawshank Redemption (1994) B6
2. The Godfather (1972) C3
3. The Godfather: Part II (1974) C3
4. 12 Angry Men (1957) D2
5. Schindler's List (1993) C3
6. The Dark Knight (2008) E5
7. One Flew Over the Cuckoo's Nest (1975) E6
8. Star Wars: Episode V - The Empire Strikes Back (1980) C4
9. Pulp Fiction (1994) C6
10. The Lord of the Rings: The Return of the King (2003) C4
11. Goodfellas (1990) B3
12. Fight Club (1999) D4
13. Star Wars (1977) C4
14. Casablanca (1942) C8
15. Shichinin no samurai (1954) C2
16. The Lord of the Rings: The Fellowship of the Ring (2001) C4
17. Goodfellas (1990) B3
18. Raiders of the Lost Ark (1981) C3
19. Rear Window (1954) D4
20. Cidade de Deus (2002) B3
21. Up (2009) C9
22. Raiders of the Lost Ark (1981) C3
23. Cinema Paradiso (1988) E2
24. The Lord of the Rings: The Fellowship of the Ring (2001) C4
25. Fight Club (1999) D8
26. The Usual Suspects (1994) C3
27. Psycho (1960) D3
28. The Silence of the Lambs (1991) D3
29. Sunset Blvd. (1950) D7
30. Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb (1964) B5
31. The Matrix (1999) B4
32. Memento (2000) D8
33. North by Northwest (1959) B3
34. It's a Wonderful Life (1946) C6
35. City Lights (1931) D2
36. Citizen Kane (1941) C6
37. The Lord of the Rings: The Two Towers (2002) B4
38. Léon (1994) E7
39. American History X (1998) E6
40. Lawrence of Arabia (1962) C2
41. Vertigo (1958) C2
42. WALL-E (2008) D2
43. Forrest Gump (1994) C4
44. Paths of Glory (1957) B7
45. Le fabuleux destin d'Amélie Poulain (2001) D4
46. M (1931) C7
47. Double Indemnity (1944) C4
48. To Kill a Mockingbird (1962) E6
49. Alien (1979) A7
50. A Clockwork Orange (1971) C1
51. The Treasure of the Sierra Madre (1948) C3
52. The Shining (1980) D4
53. The Departed (2006) C3
54. The Third Man (1949) C3
55. Das Leben der Anderen (2006) D4
56. Saving Private Ryan (1998) B5
57. Chinatown (1974) C3
58. The Pianist (2002) C5
59. Terminator 2: Judgment Day (1991) D5
60. Sen to Chihiro no kamikakushi (2001) C9
61. Eternal Sunshine of the Spotless Mind (2004) D8
62. City Lights (1931) D2
63. Slumdog Millionaire (2008) D3
64. Aliens (1986) D5
65. Requiem for a Dream (2000) C6
66. L.A. Confidential (1997) B3
67. Das Boot (1981) B5
68. Monty Python and the Holy Grail (1975) C8
69. The Bridge on the River Kwai (1957) B8
70. Reservoir Dogs (1992) C3
71. El laberinto del fauno (2006) B4
72. Rashômon (1950) D2
73. Raging Bull (1980) C4
74. The Maltese Falcon (1941) C6
75. All About Eve (1950) B4
76. Modern Times (1936) C8
77. Gran Torino (2008) E6
78. Singin' in the Rain (1952) B2
79. Some Like It Hot (1959) D5
80. Rebecca (1940) C1
81. Der Untergang (2004) C5
82. Metropolis (1927) D2
83. The Prestige (2006) B3
84. Amadeus (1984) D2
85. The Apartment (1960) D6
86. The Elephant Man (1980) E6
87. La vita è bella (1997) C5
88. 2001: A Space Odyssey (1968) D2
89. Nuovo cinema Paradiso (1988) D7
90. The Great Dictator (1940) C5
91. Once Upon a Time in America (1984) C3
92. Full Metal Jacket (1987) B5
93. The Great Escape (1963) E4
94. Star Trek (2009) C6
95. Ladri di biciclette (1948) B2
96. The Sting (1973) B3
97. Back to the Future (1985) B5
98. Sin City (2005) E5
99. Mr. Smith Goes to Washington (1939) A4
100. Touch of Evil (1958) C3
101. The Wrestler (2008) B4
102. On the Waterfront (1954) B5
103. Braveheart (1995) C7
104. Det sjunde inseglet (1957) D5
105. Hotel Rwanda (2004) B4
106. Batman Begins (2005) B7
107. Jaws (1975) D2
108. Indiana Jones and the Last Crusade (1989) C3
109. No Country for Old Men (2007) C3
110. Strangers on a Train (1951) B1
111. Blade Runner (1982) C4
112. Star Wars: Episode VI - Return of the Jedi (1983) C4
113. The Green Mile (1999) D6
114. Unforgiven (1992) C2
115. The Manchurian Candidate (1962) A4
116. Gladiator (2000) B8
117. Notorious (1946) C2
118. High Noon (1952) D2
119. Per qualche dollaro in più (1965) C6
120. Oldboy (2003) D2
121. The Big Sleep (1946) B3
122. Cool Hand Luke (1967) D3
123. Die Hard (1988) E5
124. Fargo (1996) D6
125. The Wizard of Oz (1939) B2
126. There Will Be Blood (2007) B2
127. The General (1926) B7
128. Monnoké-hime (1997) C3
129. Witness for the Prosecution (1957) A2
130. The Hangover (2009) D7
131. Donnie Darko (2001) B4
132. Smultronstället (1957) A5
133. Yojimbo (1961) E2
134. It Happened One Night (1934) D5
135. Heat (1995) E3
136. Wild Hearts and Coronets (1949) B7
137. Annie Hall (1977) D8
138. Rain (1983) B7
139. The Deer Hunter (1978) B6
140. Ben-Hur (1959) D4
141. The Sixth Sense (1999) D4
142. Kill Bill: Vol. 1 (2003) E3
143. Platoon (1986) B6
144. Into the Wild (2007) B3
145. Judgment at Nuremberg (1961) A4
146. Les diaboliques (1955) A2
147. Le notti di Cabiria (1957) A5
148. Le salaire de la peur (1953) B2
149. Million Dollar Baby (2004) A6
150. Butch Cassidy and the Sundance Kid (1969) D2
151. The Bourne Ultimatum (2007) E4
152. 8½ (1963) D7
153. Life of Brian (1979) D5
154. The Grapes of Wrath (1940) A5
155. The Big Lebowski (1998) D7
156. Snatch. (2000) D3
157. The Gold Rush (1925) B7
158. The Graduate (1967) D6
159. The Night of the Hunter (1955) C7
160. Stand by Me (1986) B3
161. The Killing (1956) C2
162. Ratatouille (2007) B5
163. Finding Nemo (2003) D3
164. Amores perros (2000) B3
165. Dog Day Afternoon (1975) B2
166. Gandhi (1982) B4
167. Brief Encounter (1945) B4
168. Gone with the Wind (1939) C3
169. Trainspotting (1996) E8
170. The Lion King (1994) D8
171. Scarface (1983) D3
172. The Wild Bunch (1969) E2
173. The Thing (1982) D5
174. Sunrise: A Song of Two Humans (1927) C3
175. V for Vendetta (2005) E3
176. Groundhog Day (1993) B7
177. The Terminator (1984) D5
178. Sleuth (1972) A2
179. Harvey (1950) D4
180. Toy Story (1995) D4
181. The Princess Bride (1987) D4
182. The Kid (1921) D6
183. The Hustler (1961) A4
184. The Curious Case of Benjamin Button (2008) D6
185. Shadow of a Doubt (1943) A1
186. Umberto D. (1952) A4
187. The Best Years of Our Lives (1946) A3
188. The Incredibles (2004) C9
189. Twelve Monkeys (1995) A6
190. Children of Men (2006) B4
191. Lock, Stock and Two Smoking Barrels (1998) D3
192. Låt den rätte komma in (2008) D2
193. Stalag 17 (1953) B7
194. The Ox-Bow Incident (1943) D2
195. Hotaru no haka (1988) B9
196. Duck Soup (1933) C8
197. Casino (1995) B3
198. La battaglia di Algeri (1966) B6
199. The African Queen (1951) B2
200. The Lady Vanishes (1938) B1
201. Dial M for Murder (1954) B7
202. Letters from Iwo Jima (2006) E6
203. The Adventures of Robin Hood (1938) C3
204. In Bruges (2008) C8
205. Anatomy of a Murder (1959) A2
206. The Conversation (1974) C7
207. A Streetcar Named Desire (1951) A4
208. King Kong (1933) C3
209. Trainspotting (1996) E8
210. The Last Weekend (1945) D4
211. Ed Wood (1994) C8
212. Le scaphandre et le papillon (2007) A5
213. La strada (1954) C7
214. Who's Afraid of Virginia Woolf? (1966) A3
215. Rope (1948) B9
216. His Girl Friday (1940) D7
217. All Quiet on the Western Front (1930) C8
218. Rosemary's Baby (1968) D4
219. Changeling (2008) B6
220. Bonnie and Clyde (1967) D3
221. Safety Last! (1923) B7
222. Crash (2004/I) E6
223. Sweet Smell of Success (1957) D5
224. Frankenstein (1931) C4
225. The Philadelphia Story (1940) A2
226. Network (1976) D6
227. Kill Bill: Vol. 2 (2004) E3
228. Patton (1970) B8
229. Manhattan (1979) C8
230. The Day the Earth Stood Still (1951) B4
231. Great Expectations (1946) A3
232. Magnolia (1999) D7
233. Big Fish (2003) D4
234. Wo hu cang long (2000) C2
235. Arsenic and Old Lace (1944) B7
236. Little Miss Sunshine (2006) D7
237. Laura (1944) B3
238. Mystic River (2003) C6
239. Nosferatu, eine Symphonie des Grauens (1922) D2
240. Glory (1989) C8
241. Roman Holiday (1953) D5
242. Good Will Hunting (1997) D7
243. In the Heat of the Night (1967) E4
244. Rocky (1976) E4
245. La dolce vita (1960) D5
246. Harold and Maude (1971) D6
247. Spartacus (1960) E3
248. My Man Godfrey (1936) D6
249. Planet of the Apes (1968) C5
250. Bringing Up Baby (1938) A7

vodkaster.com
TAKE A MOVIE SHOT

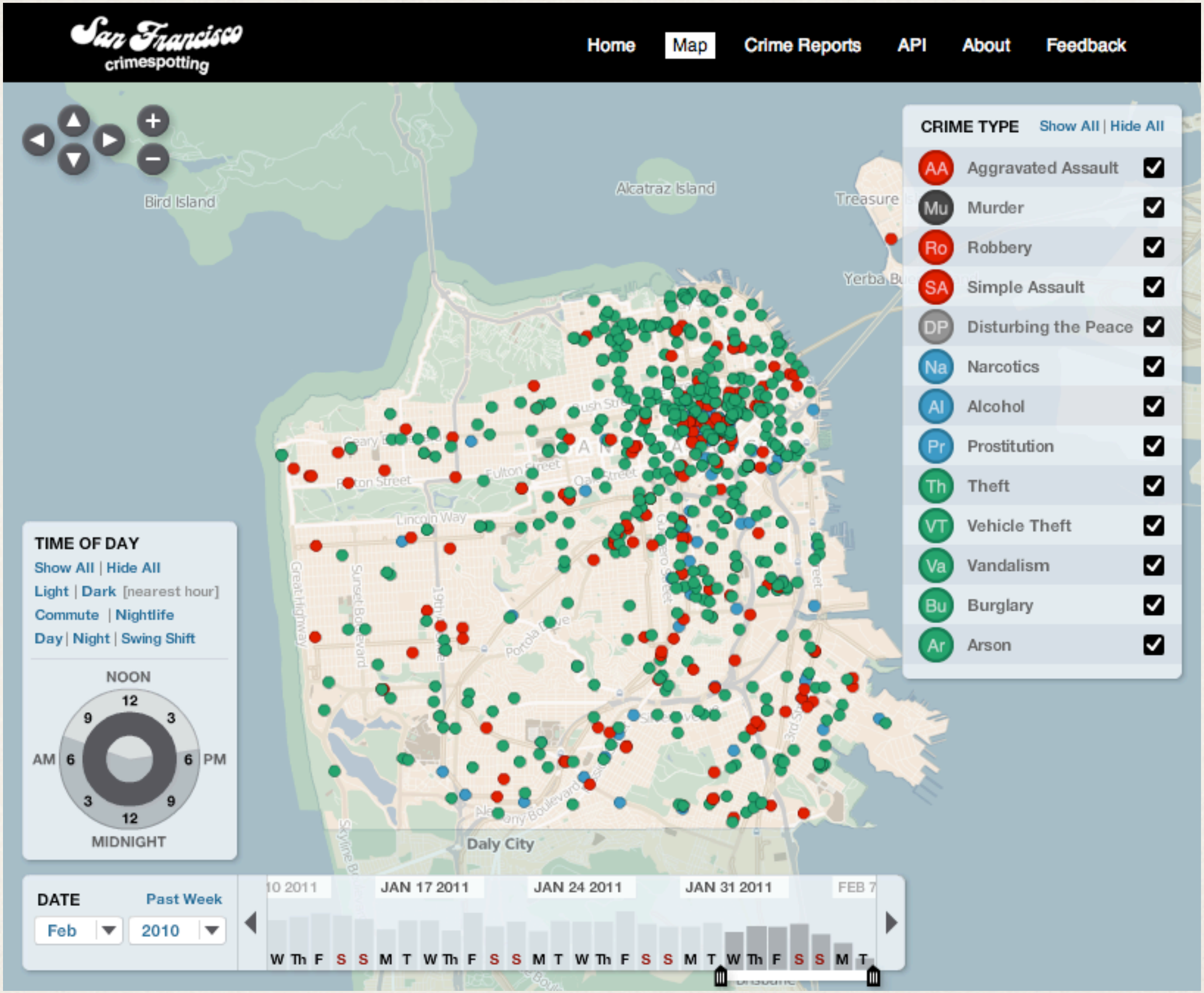
Wrong!

http://blog.vodkaster.com/2009/06/25/the-top-250-best-movies-of-all-time-map/

Thursday, December 8, 2011

Right!

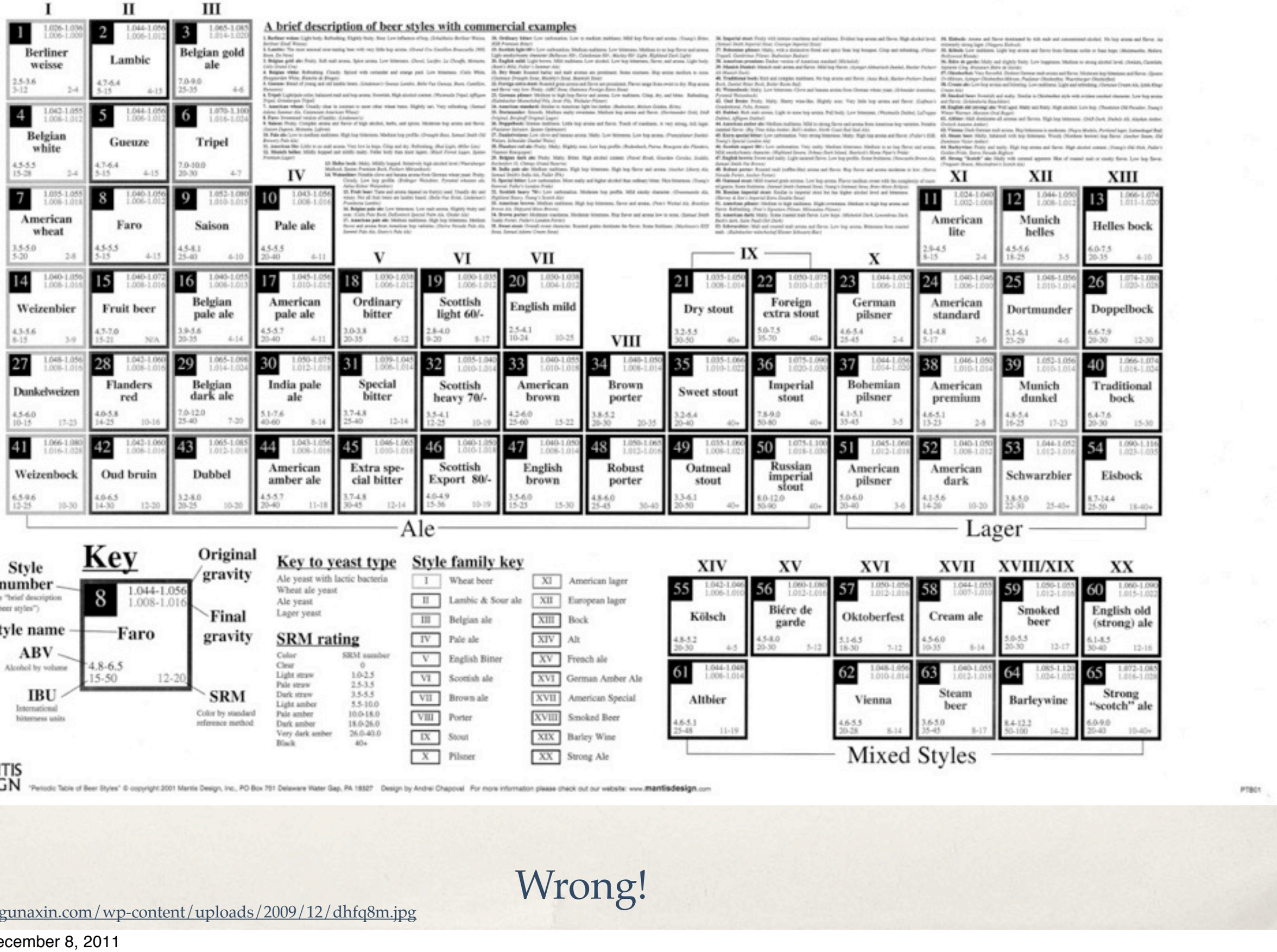
http://www.futurechanges.org/2009/11/12/eisenhower-interstate-system-london-underground-style-map/



Right!

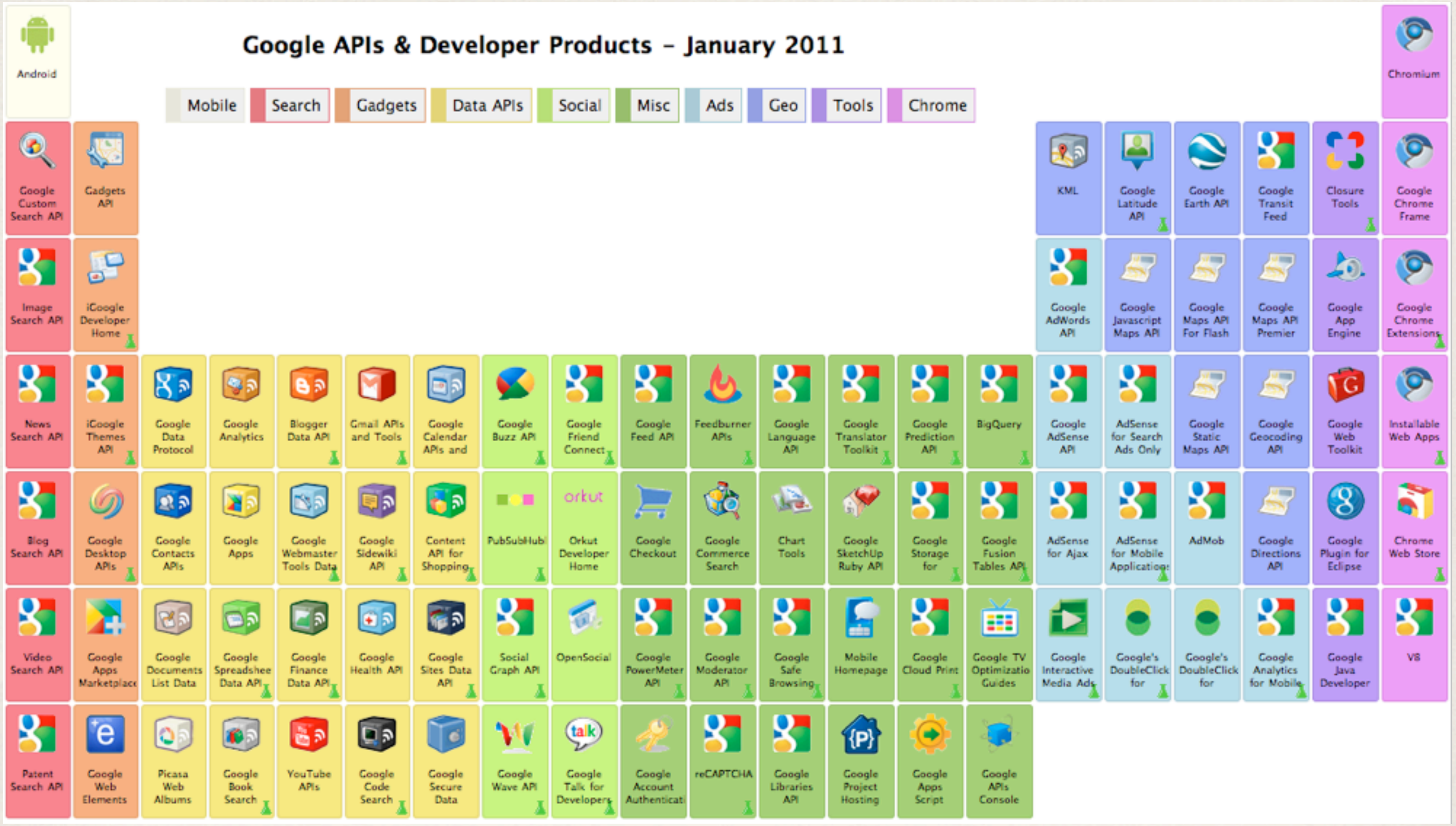
http://sanfrancisco.crimespotting.org/

Periodic Table of Beer Styles



Wrong!

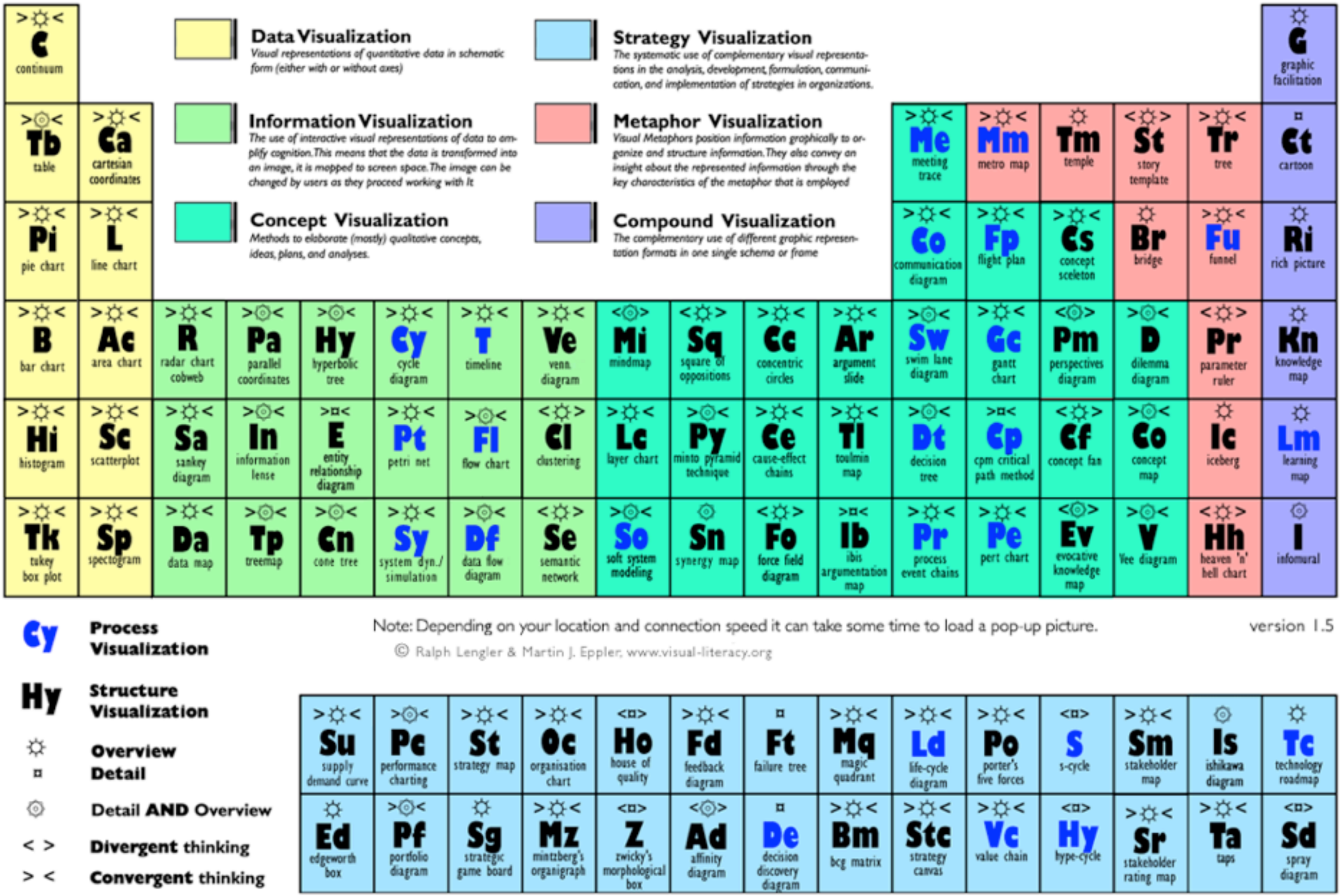
http://www.gunaxin.com/wp-content/uploads/2009/12/dhfq8m.jpg



Wrong!

http://code.google.com/more/table/

A PERIODIC TABLE OF VISUALIZATION METHODS



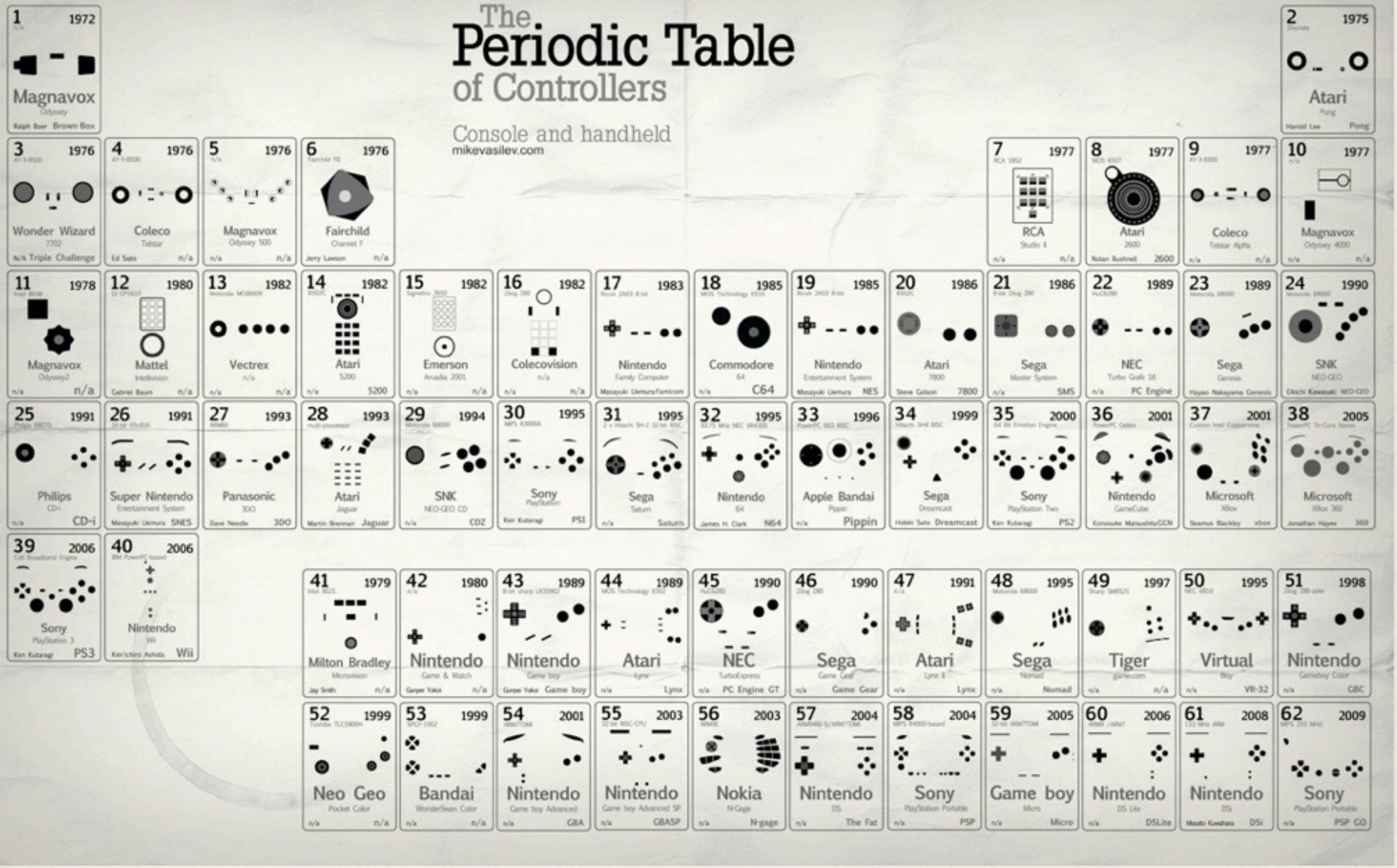
Wrong!

http://www.visual-literacy.org/periodic_table/periodic_table.html

The Periodic Table of Controllers

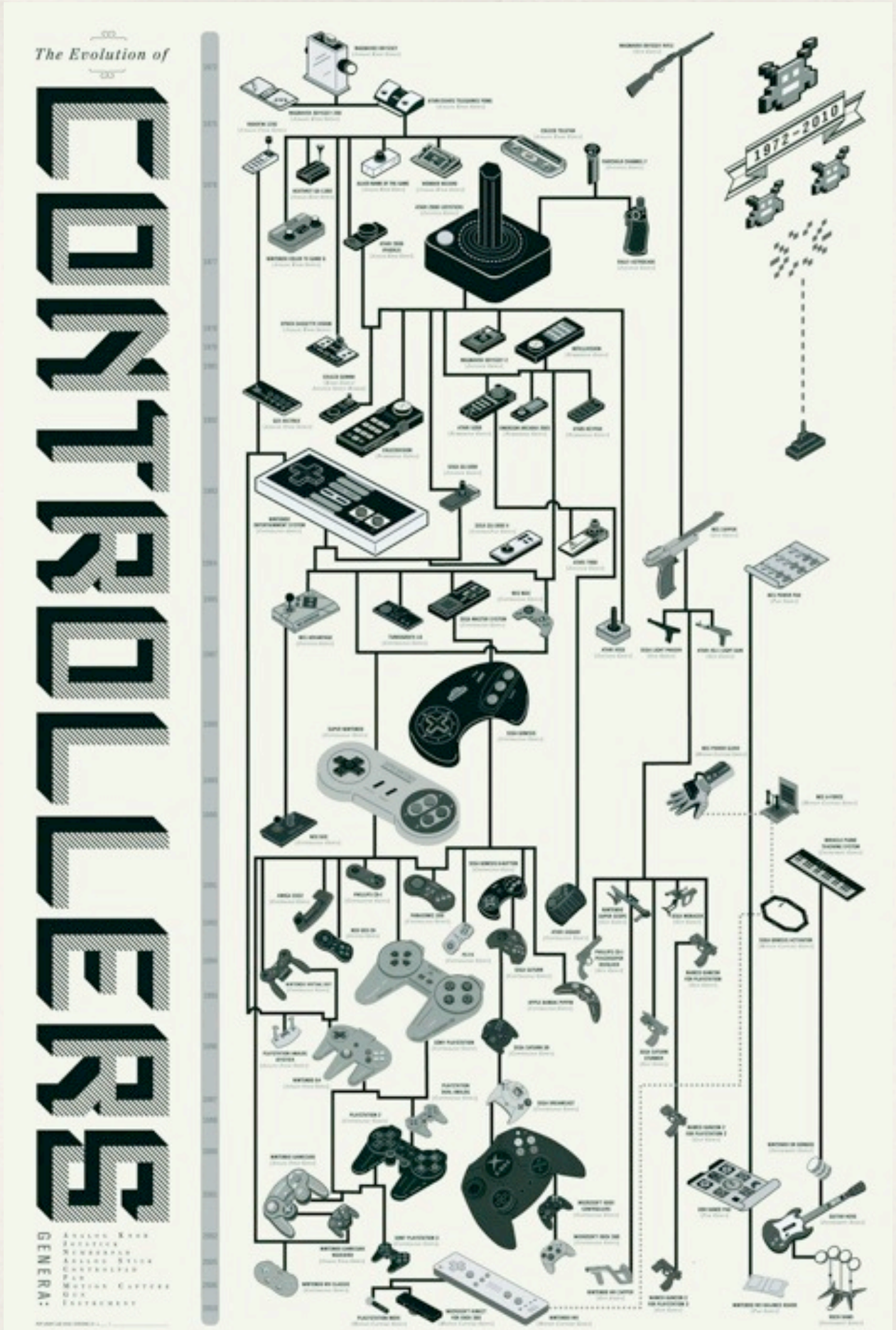
Console and handheld

mikevasilev.com



Wrong!

http://web.me.com/mvasilev/mikevasilev.com/PTOC.html



Right!

http://popchartlab.com/index.php/poster_detail/the_evolution_of_video_game_controllers/