INFM 603: Information Technology and Organizational Context

Session 9: Visualization

Jimmy Lin
The iSchool
University of Maryland

Thursday, November 6, 2014

Material borrowed from slides by Ben Shneiderman for his visualization course

Visualization

- Why visualizations?
  - Human perceptual skills are remarkable
  - Human image storage storage is fast and vast
- What are visualizations for?
  - A good visualization tells a story
  - A great visualization delivers insight
- Science as well as an art…

Raw Ingredients

- Types of data
  - Categorical data
  - Ordinal data
  - Numerical data
- Ways of representing data
  - Height
  - Width
  - Thickness
  - Area
  - Angles
  - Size
  - Shape
  - Color

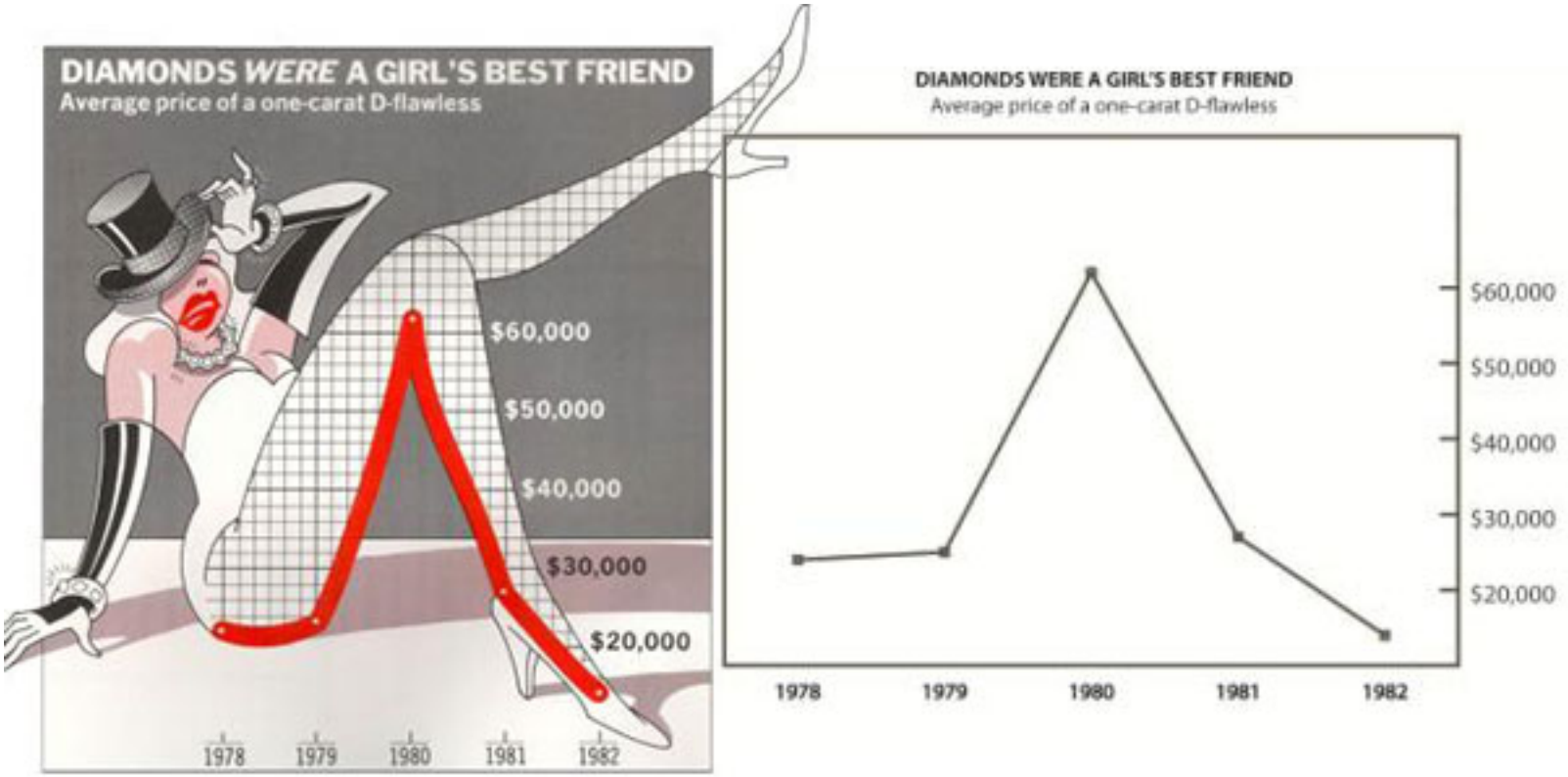
SAN DIEGO
ZOO
Panda Viewing
9am - 5:30pm

Static vs. Interactive

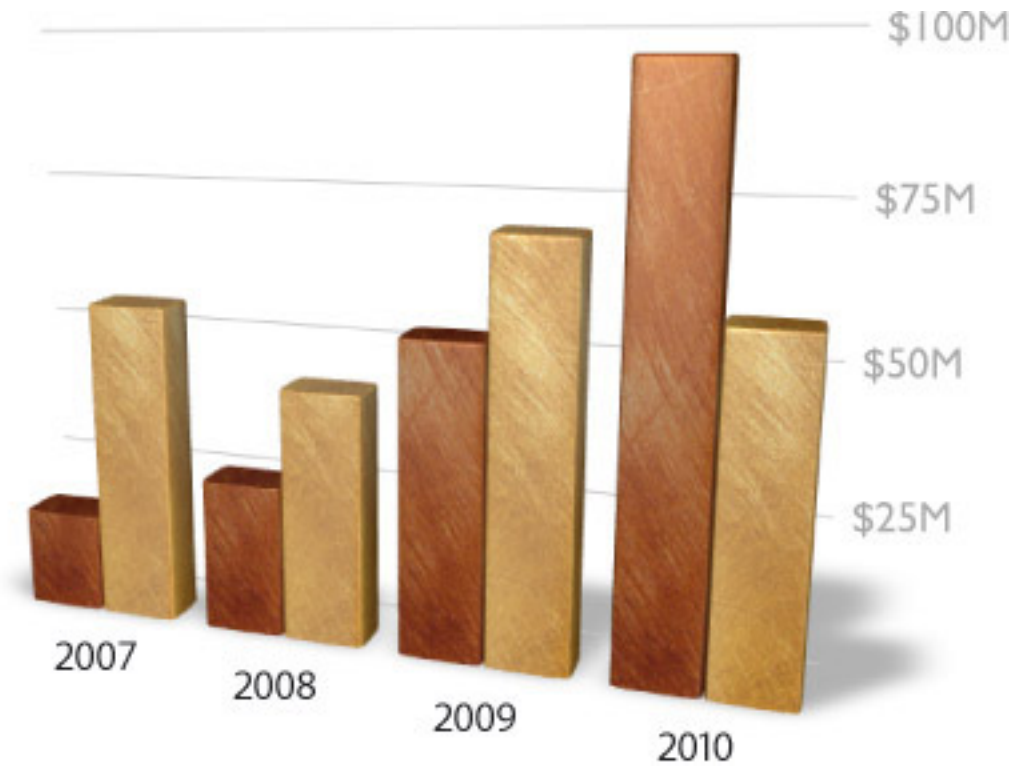
- Interactive visualizations allow users to “explore” the data
  - Overview
  - Zoom and filter
  - Details-on-demand
- Brushing and linking is a powerful technique

Info Porn

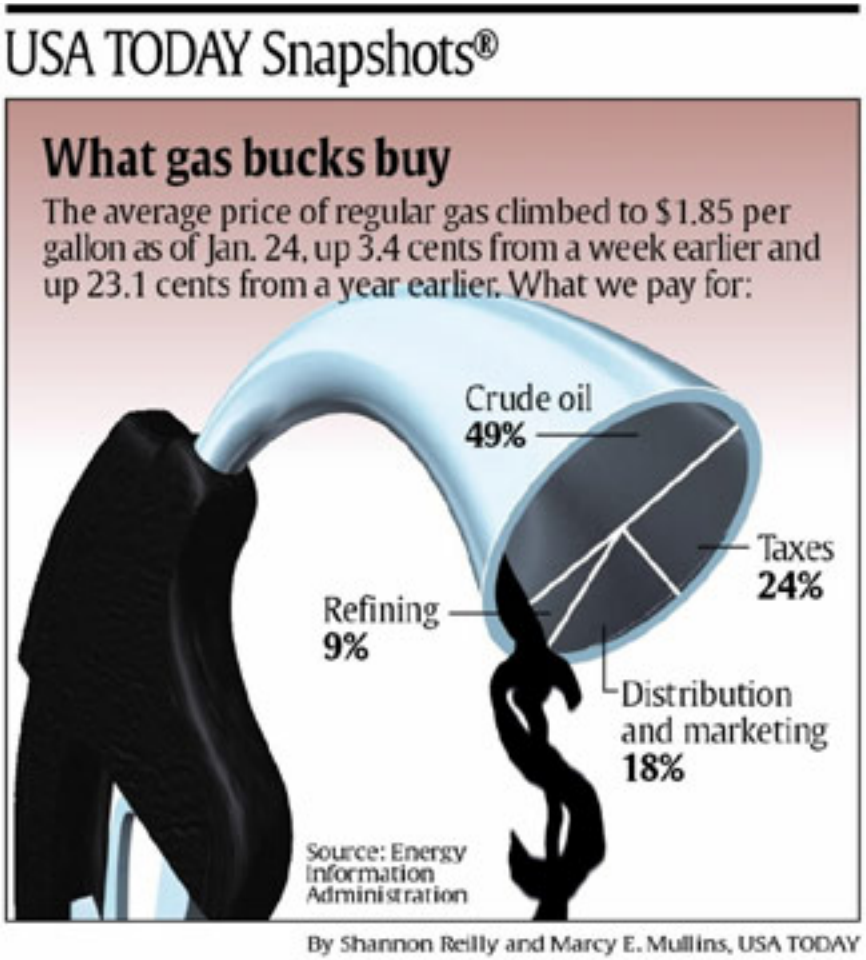
- Informative visualizations vs. info porn
- Beware of infographics and chart junk



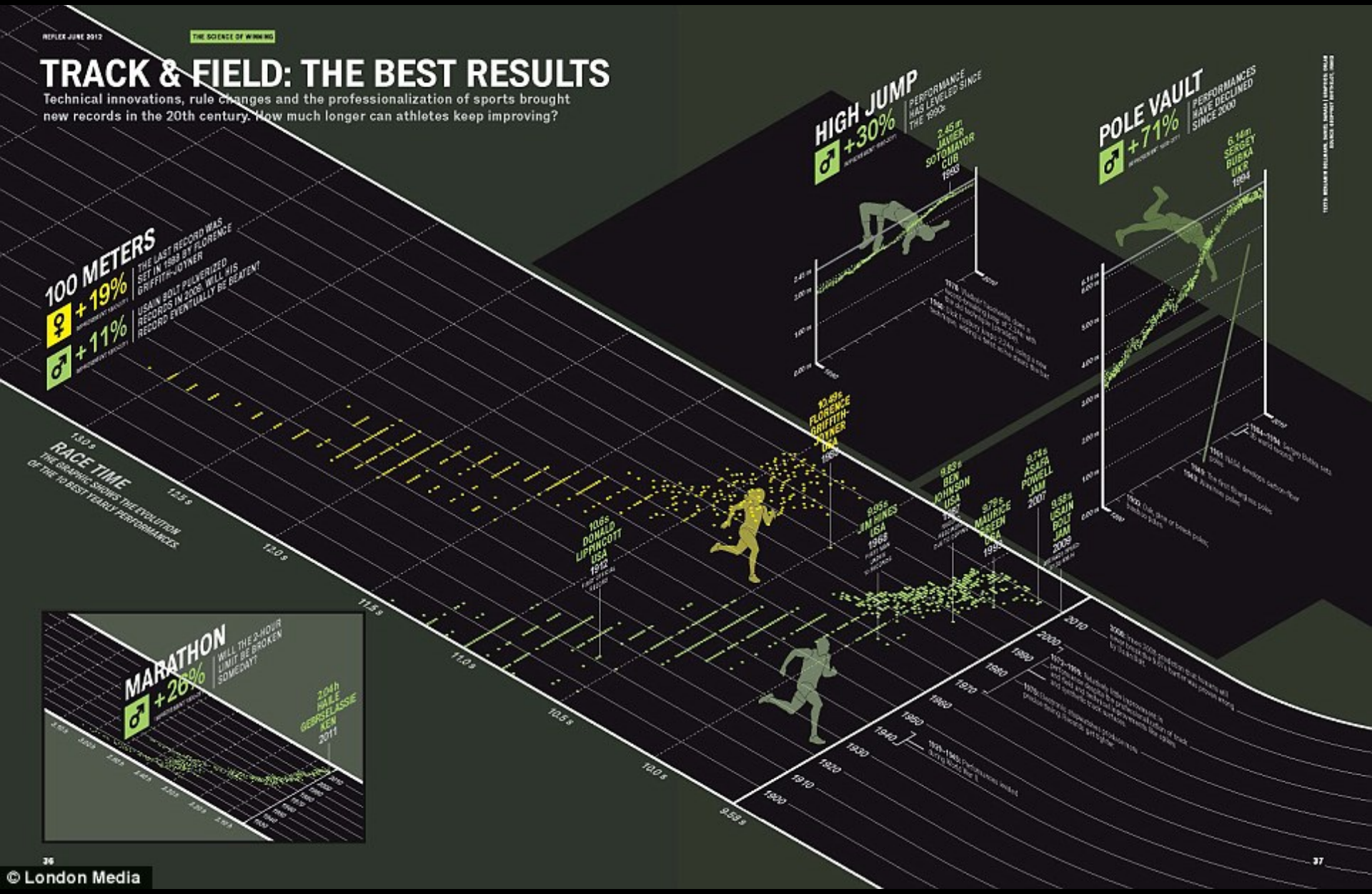
http://infosthetics.com/archives/chart_junk_effective.jpg



http://thedoublethink.com/wp-content/uploads/2009/07/keynote-chartjunk.jpg

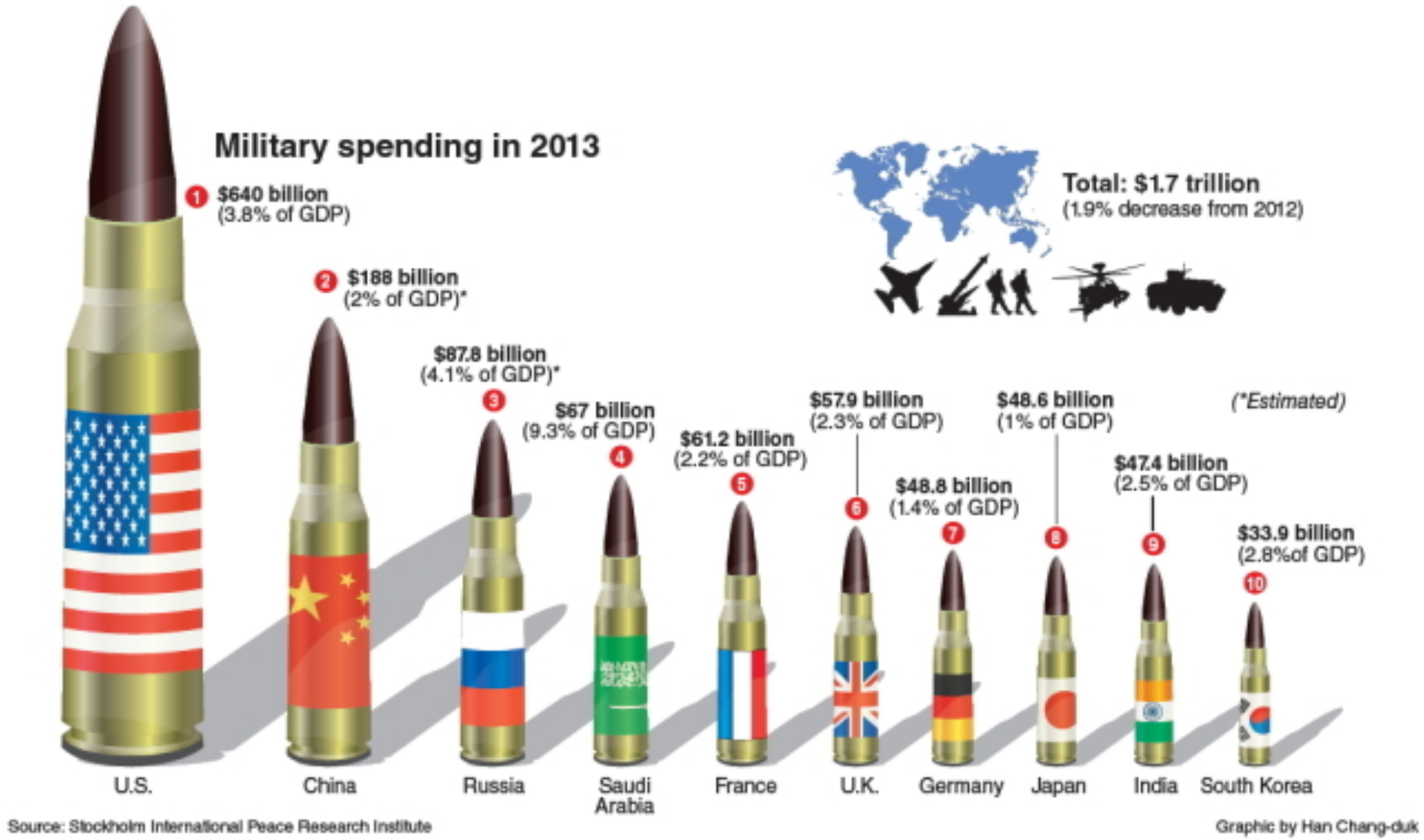


http://www.captech.org/vcp/mathmania/February/images/GasGraph.jpg



http://junkcharts.typepad.com/junk_charts/music/

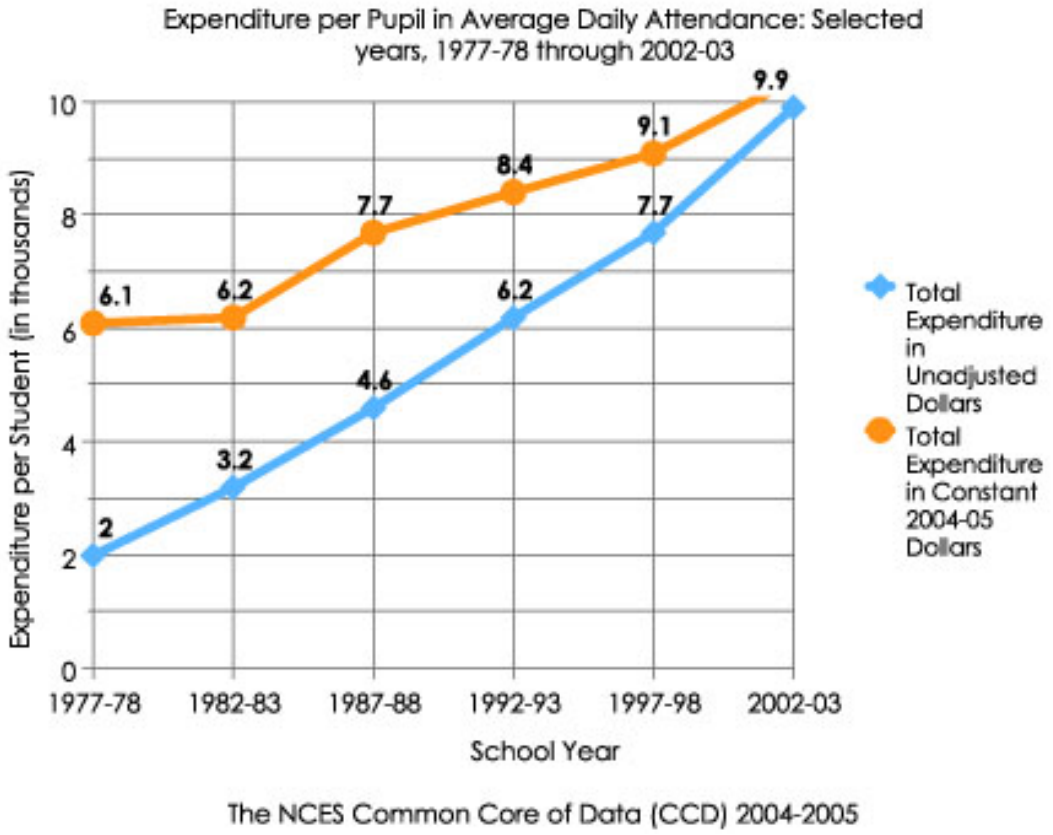
Compare: New York Times interactive visualization



http://junkcharts.typepad.com/junk_charts/2014/10/misguided-warheads-in-the-classroom.html

Line Graphs (+ Variants)

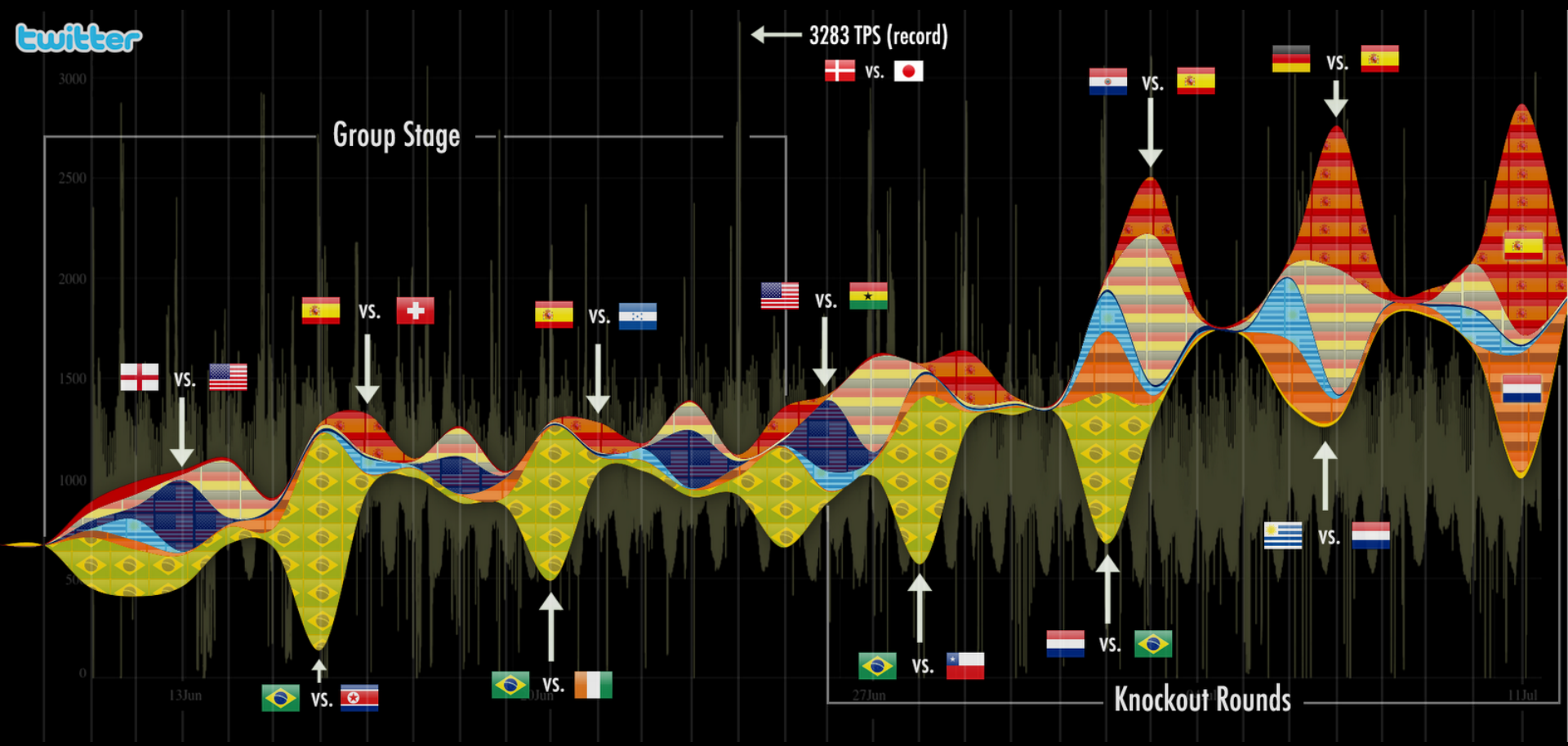
- Suitable for what types of data?
- What makes a good line graph?

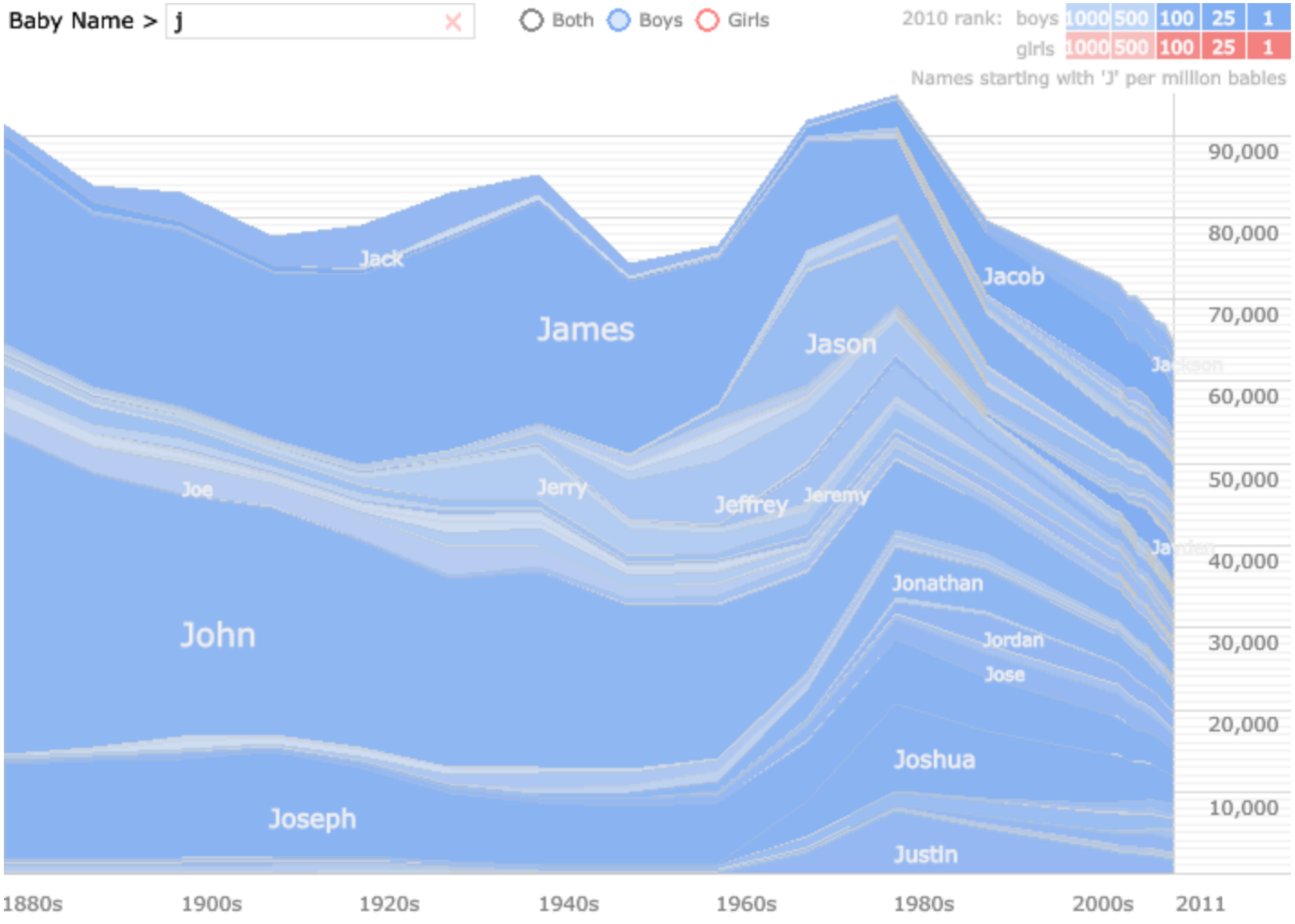


http://nces.ed.gov/nceskids/help/user_guide/graph/images/line.jpg

Stacked Graphs

- Multiple time-series data "stacked" on top of each other
  - Useful for showing aggregates
- Sometimes called "stream graphs"

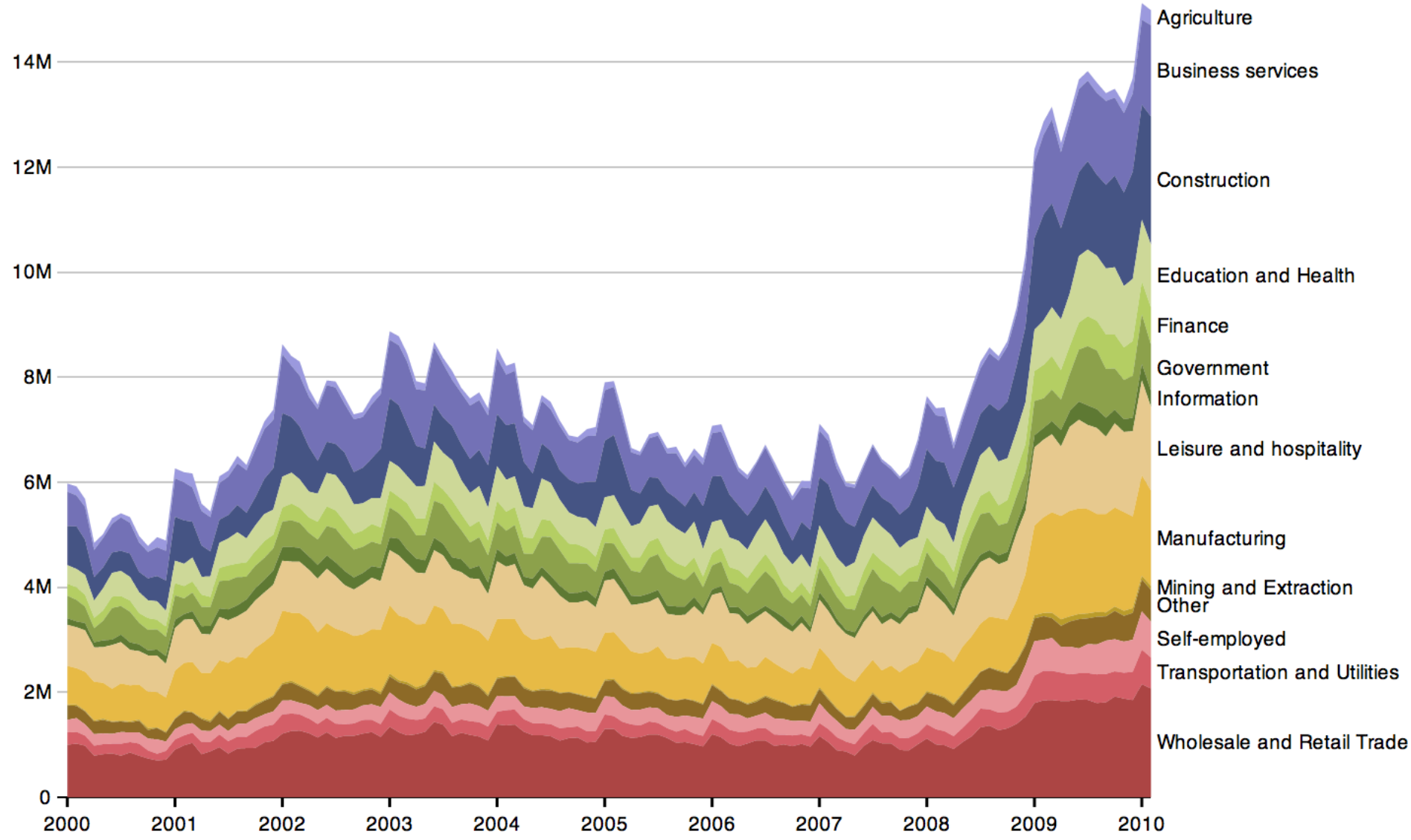

twitter
3283 TPS (record)
vs.
Group Stage
Knockout Rounds
3000
2500
2000
1500
1000
0
13Jun
27Jun
11Jul



http://www.babynamewizard.com/voyager#

Stacked Graph of Unemployed U.S. Workers by Industry

View: Counts (Zero-Baseline)

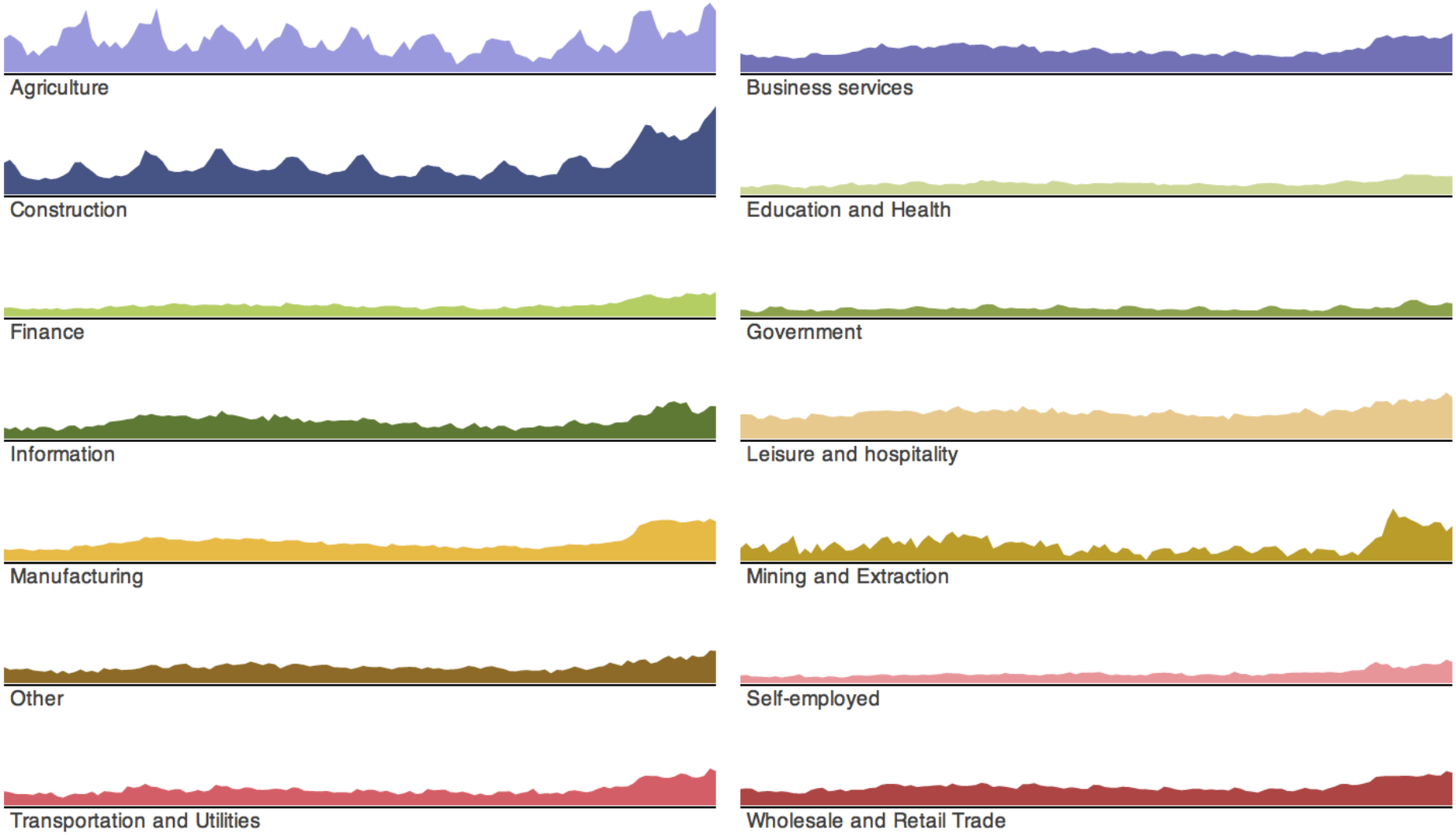


Total counts of unemployed persons per industry, 2000-2010.

Source: U.S. Bureau of Labor Statistics

http://hci.stanford.edu/jheer/files/zoo/

Unemployment Rate of U.S. Workers by Industry, 2000-2010



Unemployment rates of U.S. workers per industry.

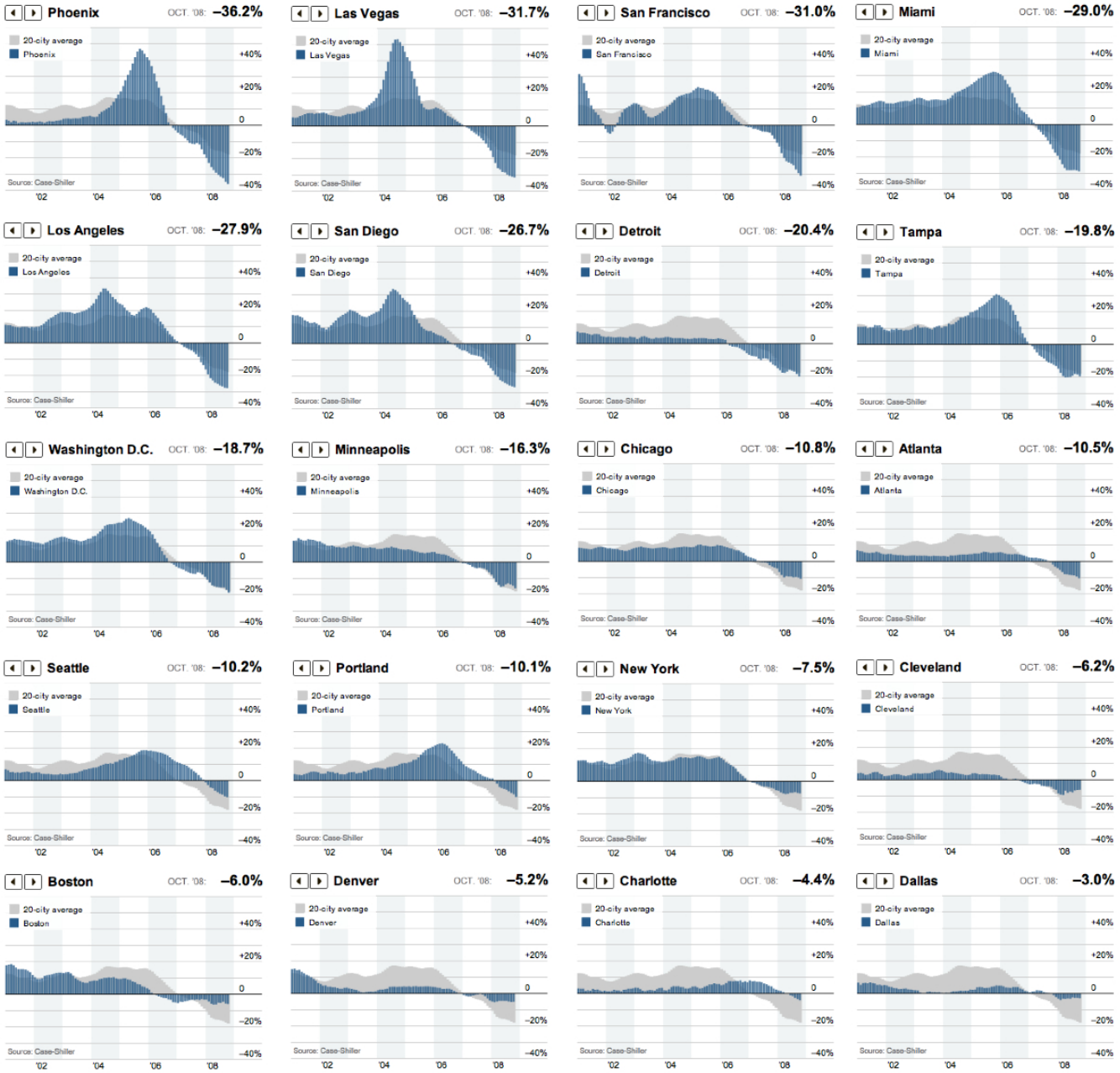
Source: U.S. Bureau of Labor Statistics

Small Multiples

- Multiple small graphs arranged in a grid
- Advantages and disadvantages?

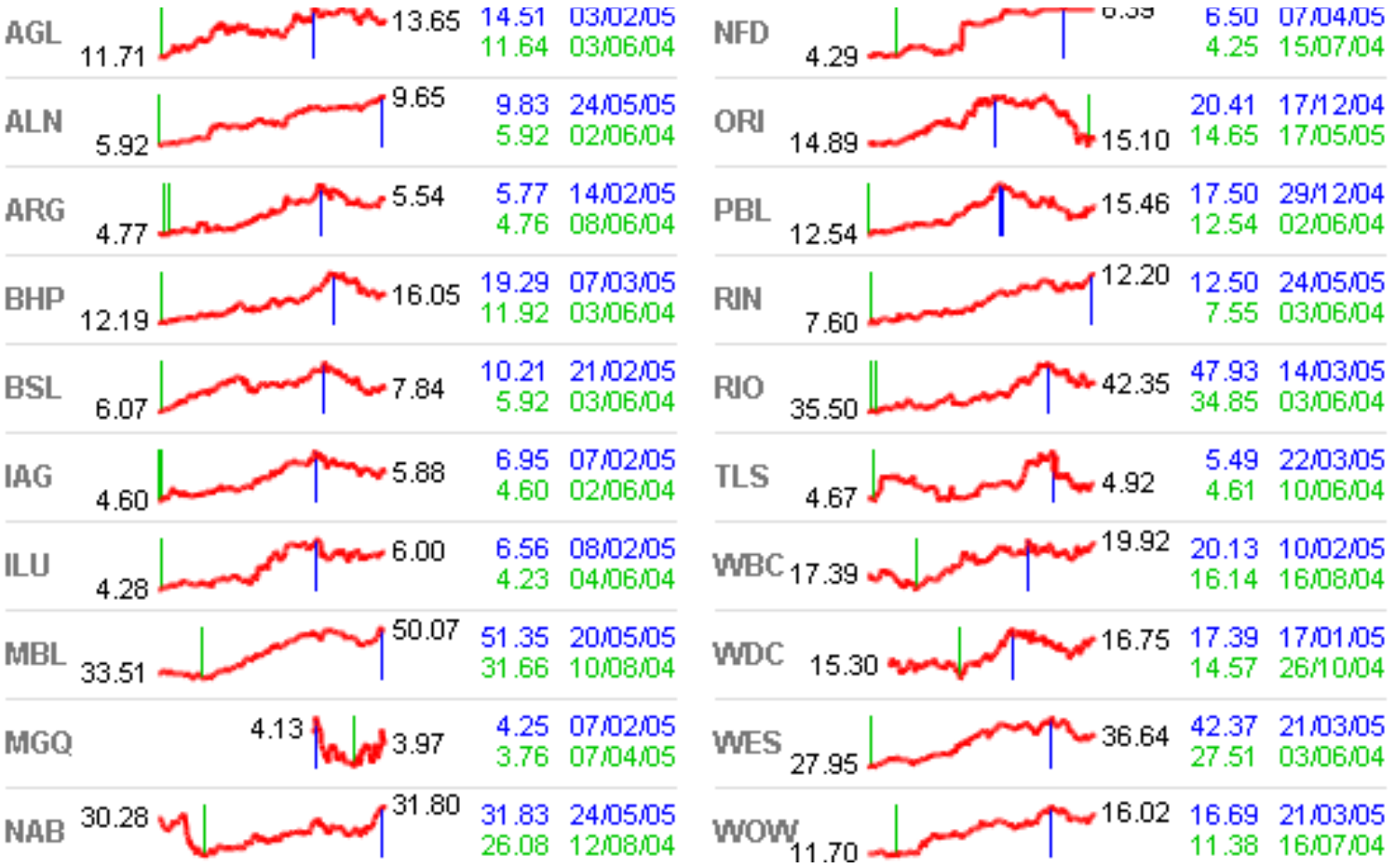
Change in Home Prices (year over year)

From New York Times Economix blog



http://informationdesignmdippold.files.wordpress.com/2011/06/good-id-1.jpg

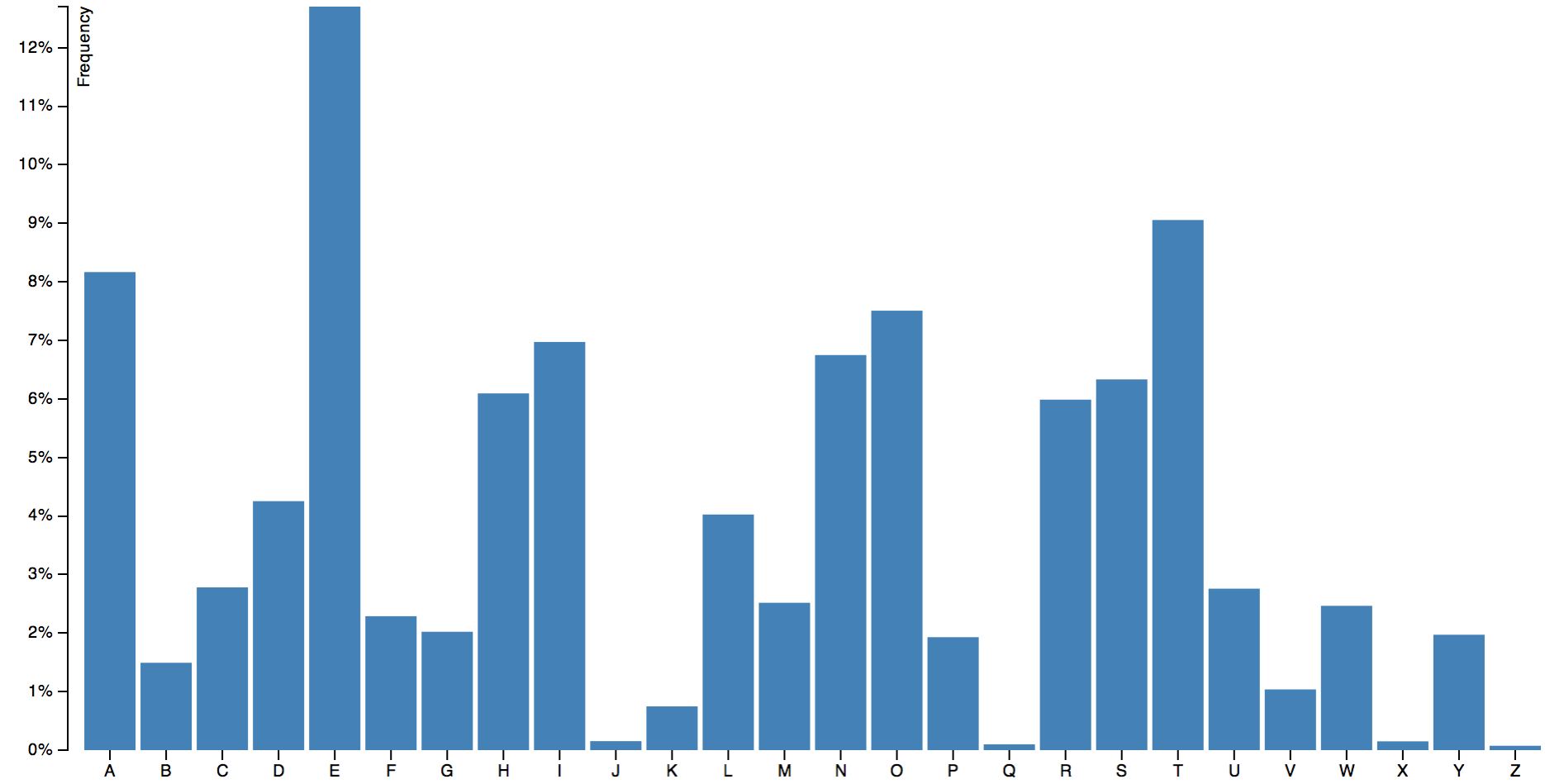
Spark Lines

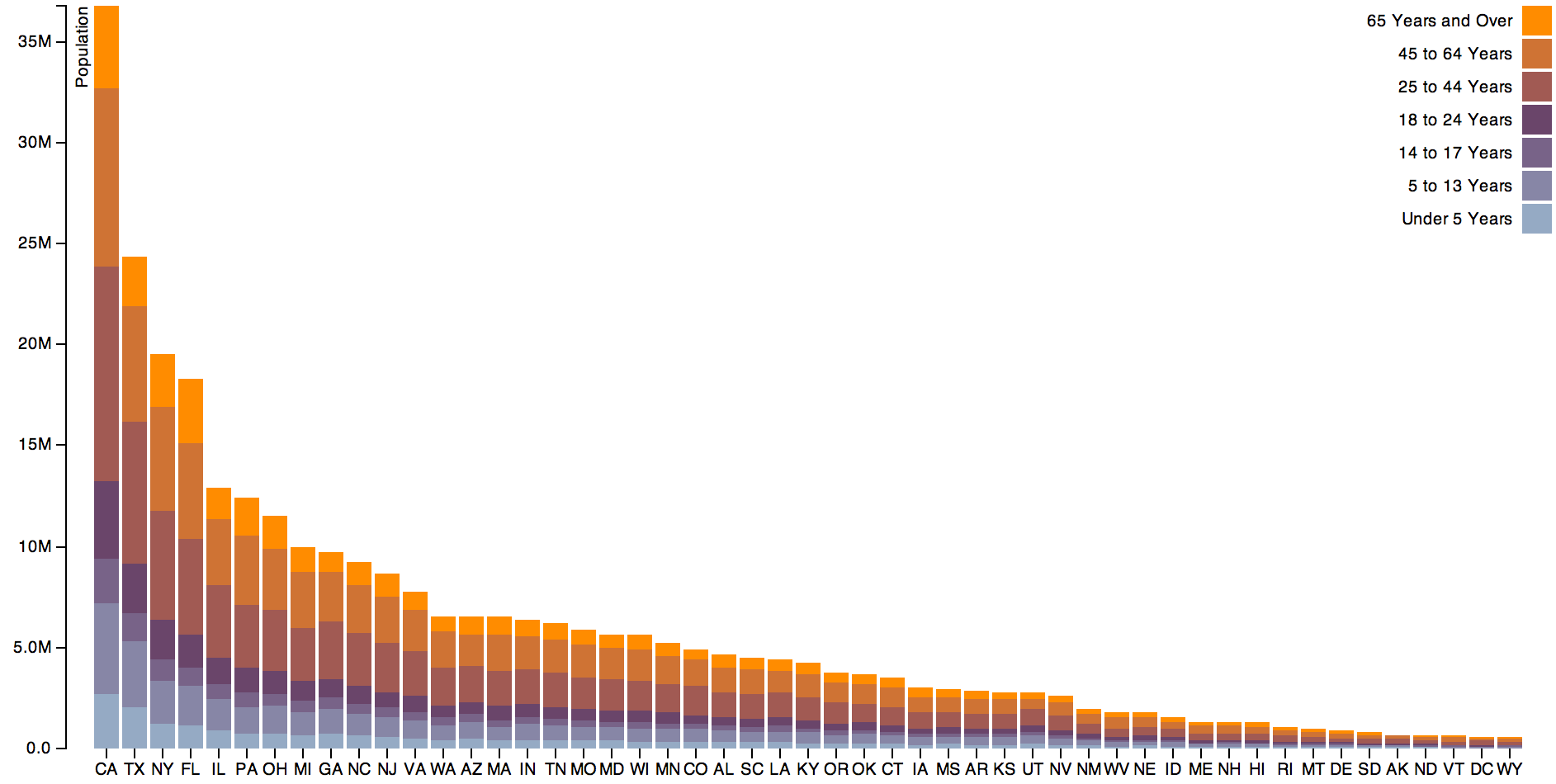


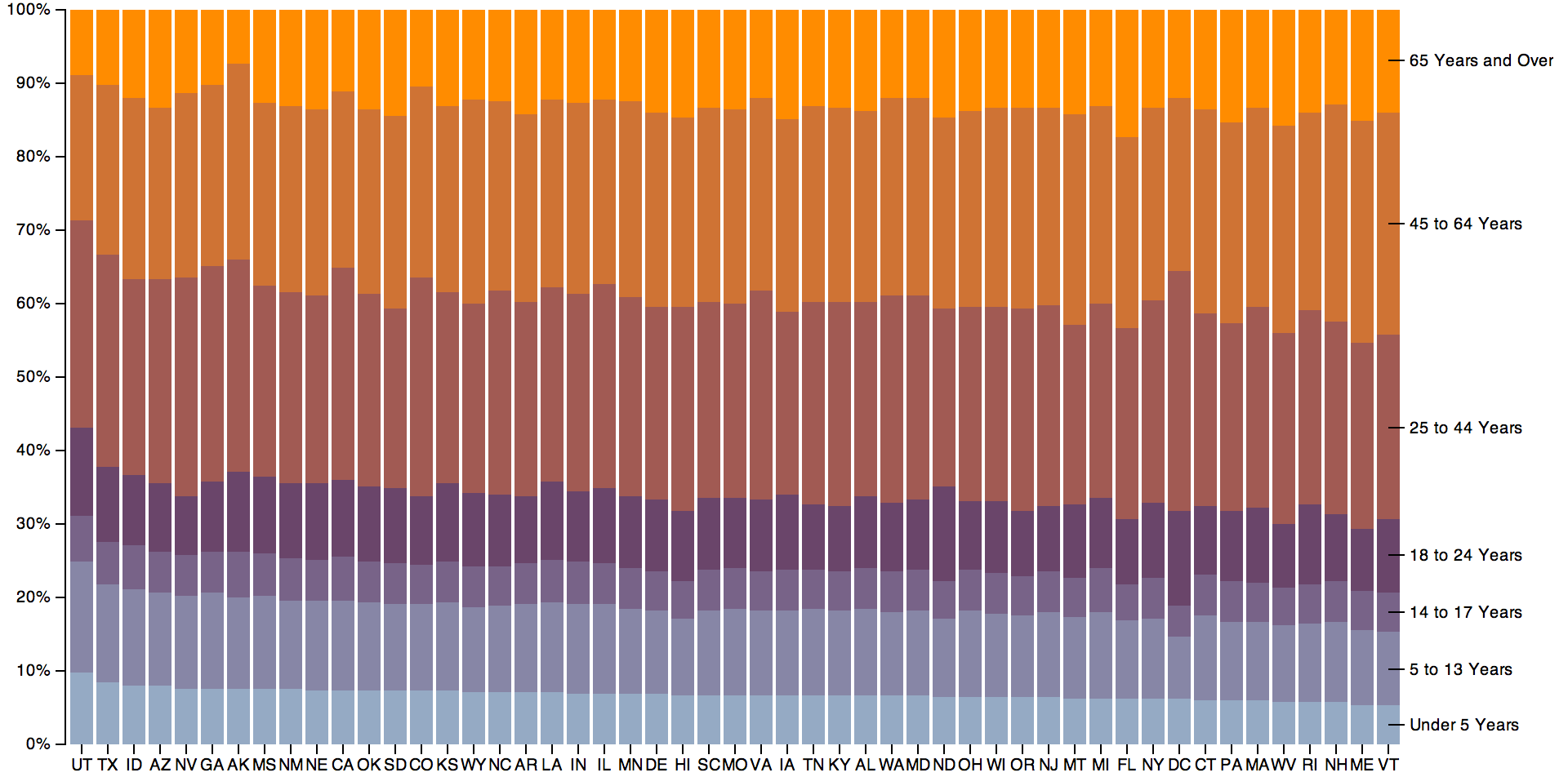
http://www.plainlanguage.com.au/sparklines.png

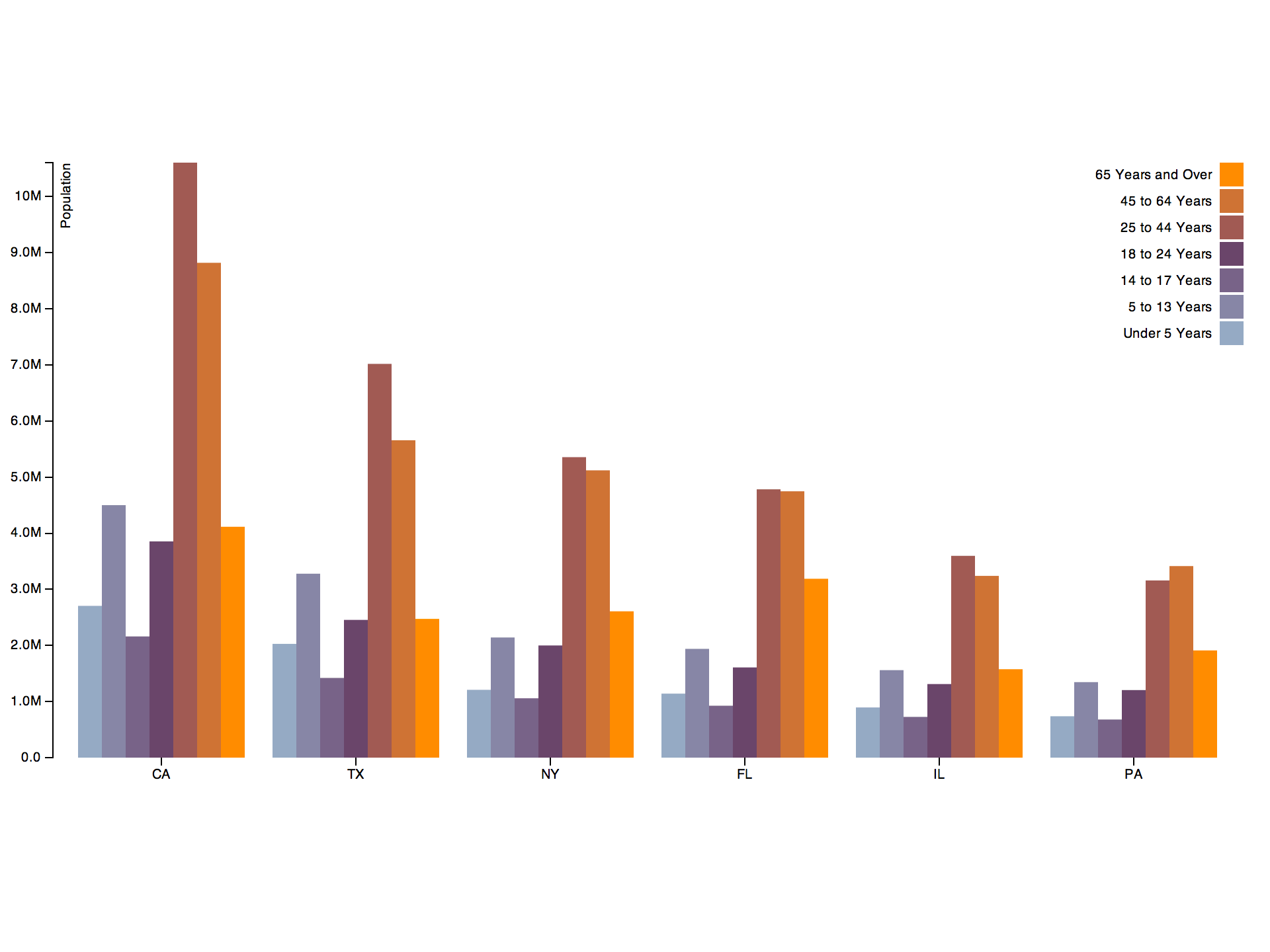
Bar Graphs (+ Variants)

- Suitable for what types of data?
- Following examples from d3.js


Frequency
12%
11%
10%
9%
8%
7%
6%
5%
4%
3%
2%
1%
0%
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

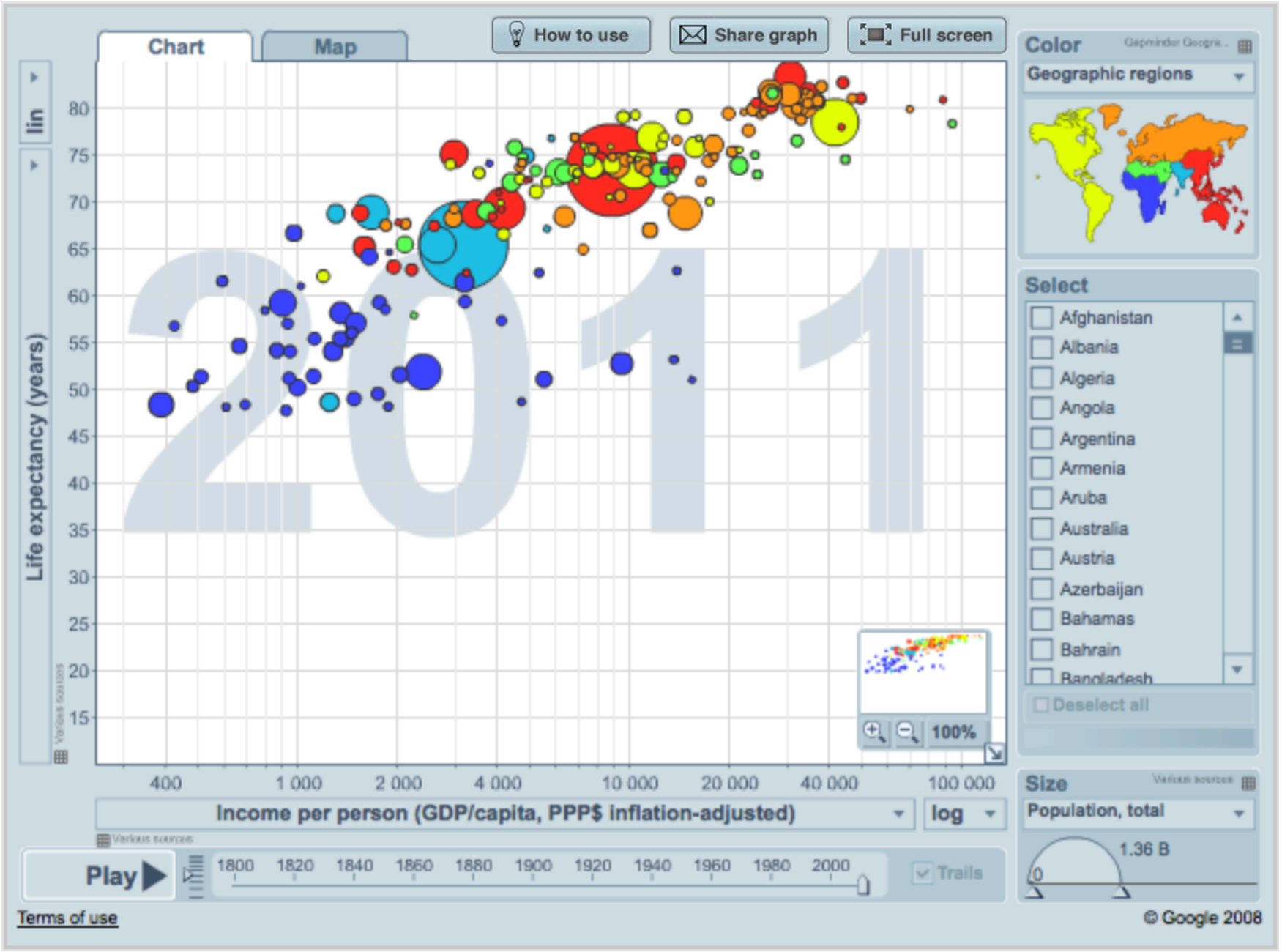

Population
35M
30M
25M
20M
15M
10M
5.0M
0.0
CA TX NY FL IL PA OH MI GA NC NJ VA WA AZ MA IN TN MO MD WI MN CO AL SC LA KY OR OK CT IA MS AR KS UT NV NM WV NE ID ME NH HI RI MT DE SD AK ND VT DC WY
65 Years and Over
45 to 64 Years
25 to 44 Years
18 to 24 Years
14 to 17 Years
5 to 13 Years
Under 5 Years


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
65 Years and Over
45 to 64 Years
25 to 44 Years
18 to 24 Years
14 to 17 Years
5 to 13 Years
Under 5 Years
UT TX ID AZ NV GA AK MS NM NE CA OK SD CO KS WY NC AR LA IN IL MN DE HI SC MO VA IA TN KY AL WA MD ND OH WI OR NJ MT MI FL NY DC CT PA MA WV RI NH ME VT


Population
10M
9.0M
8.0M
7.0M
6.0M
5.0M
4.0M
3.0M
2.0M
1.0M
0.0
CA
TX
NY
FL
IL
PA
65 Years and Over
45 to 64 Years
25 to 44 Years
18 to 24 Years
14 to 17 Years
5 to 13 Years
Under 5 Years

Scatterplots

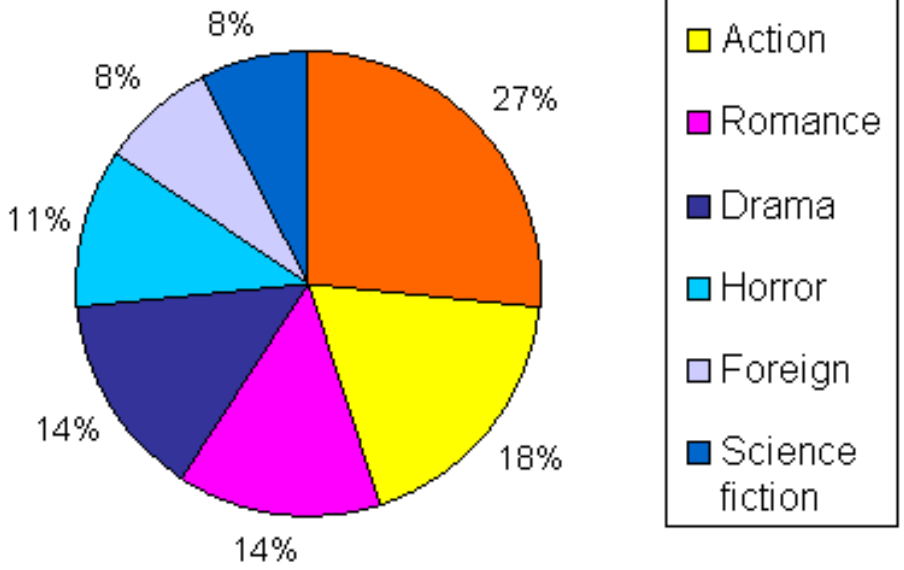
- Suitable for what types of data?
- Note availability of other ways to encode data at points

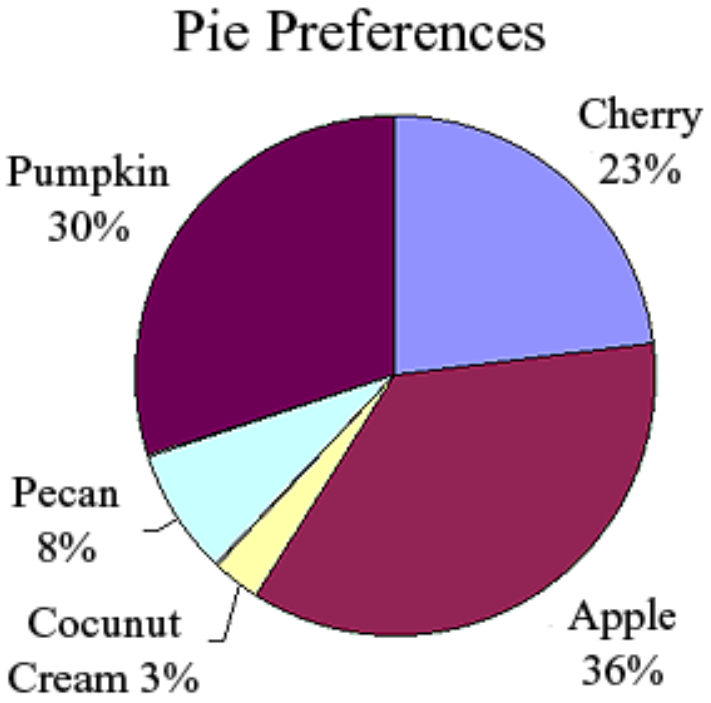


http://www.gapminder.org/

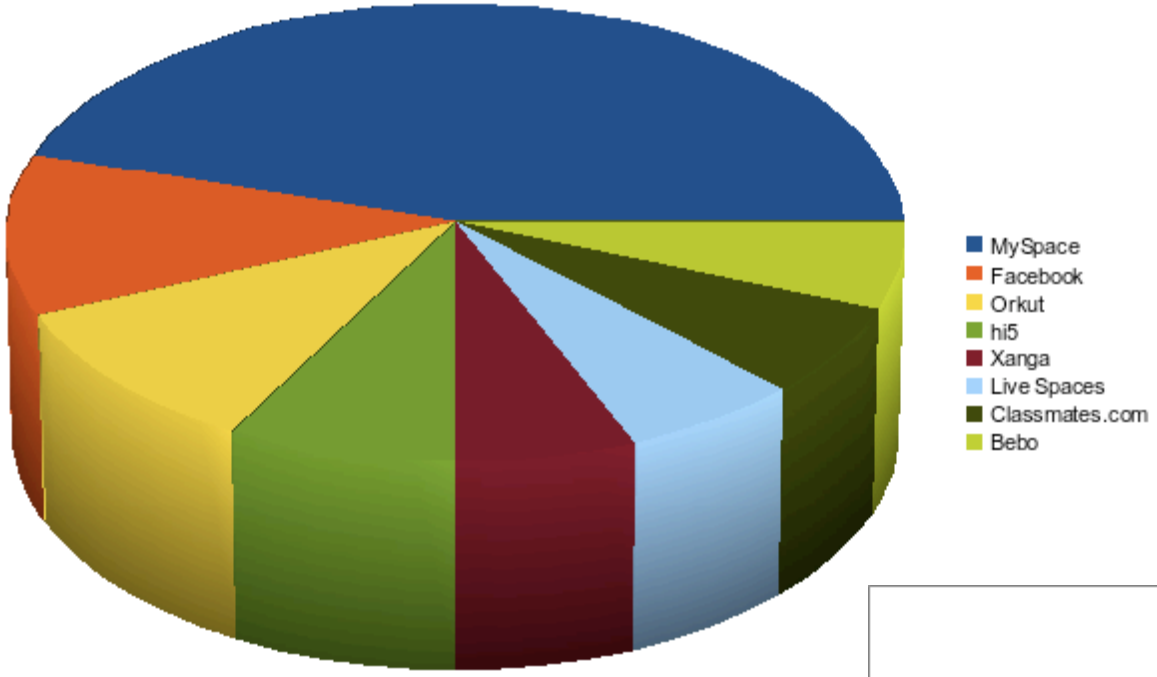
Pie Charts (+ Variants)

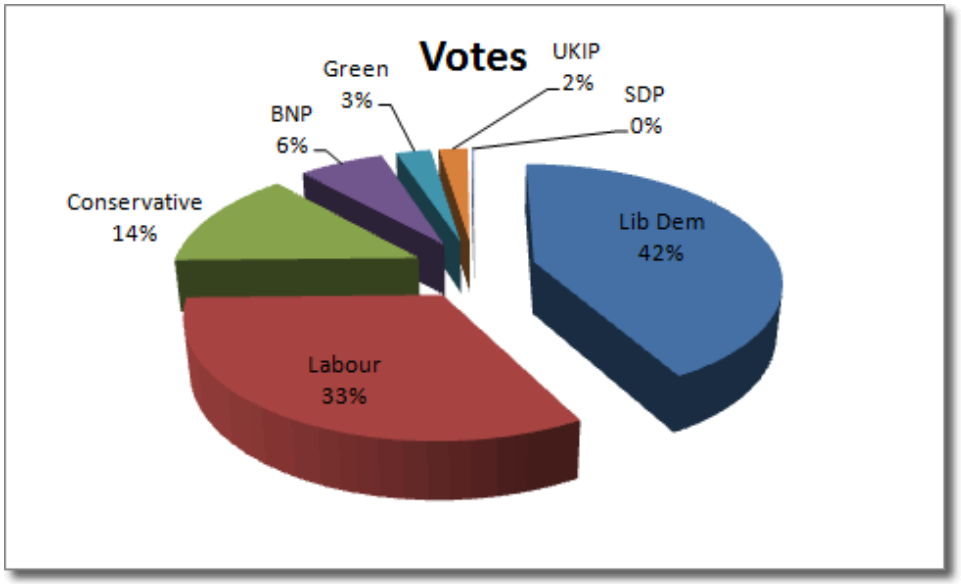
- Suitable for what types of data?
- Different ways for labeling



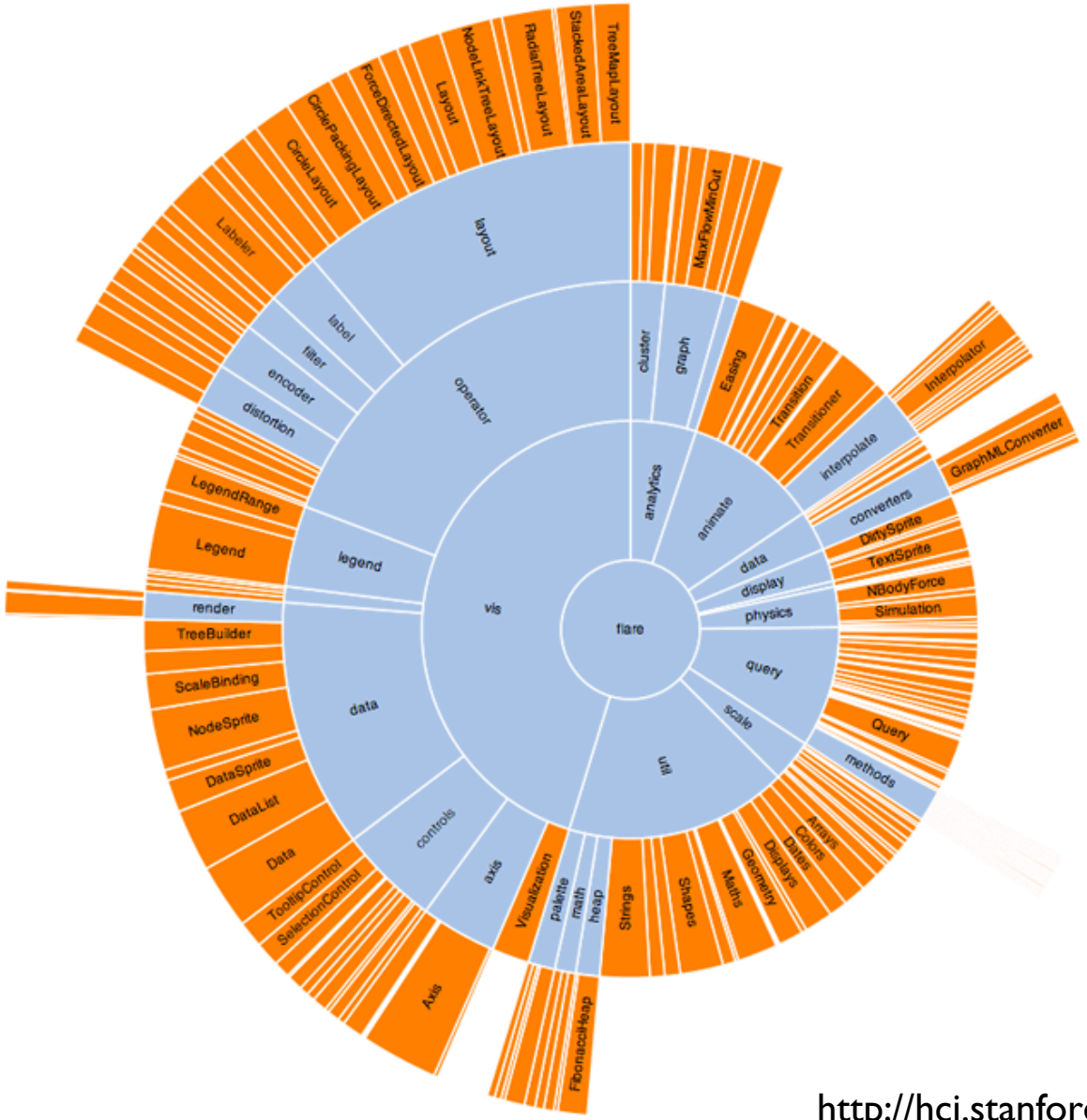


More Chart Junk



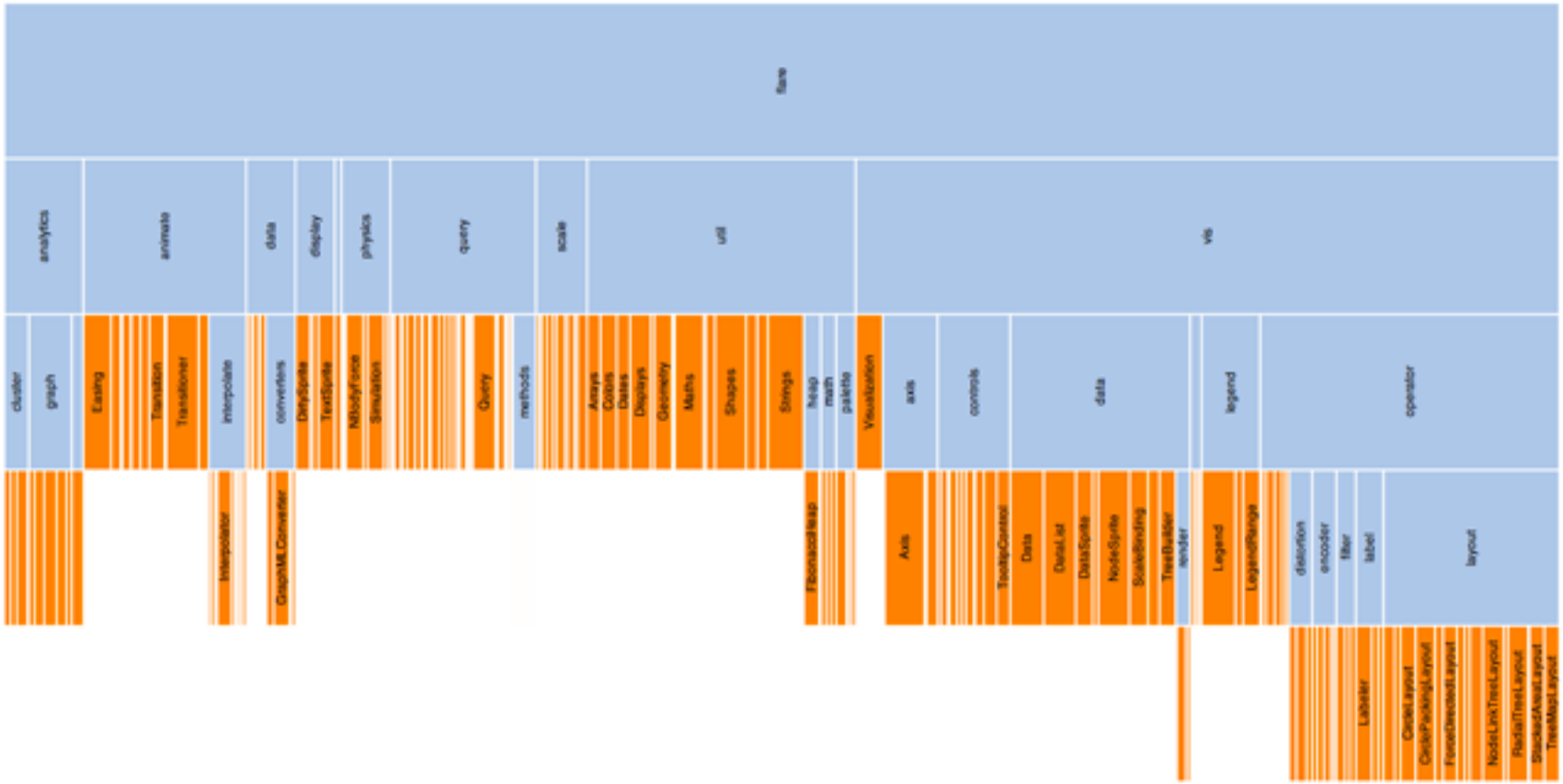


Sunburst Trees



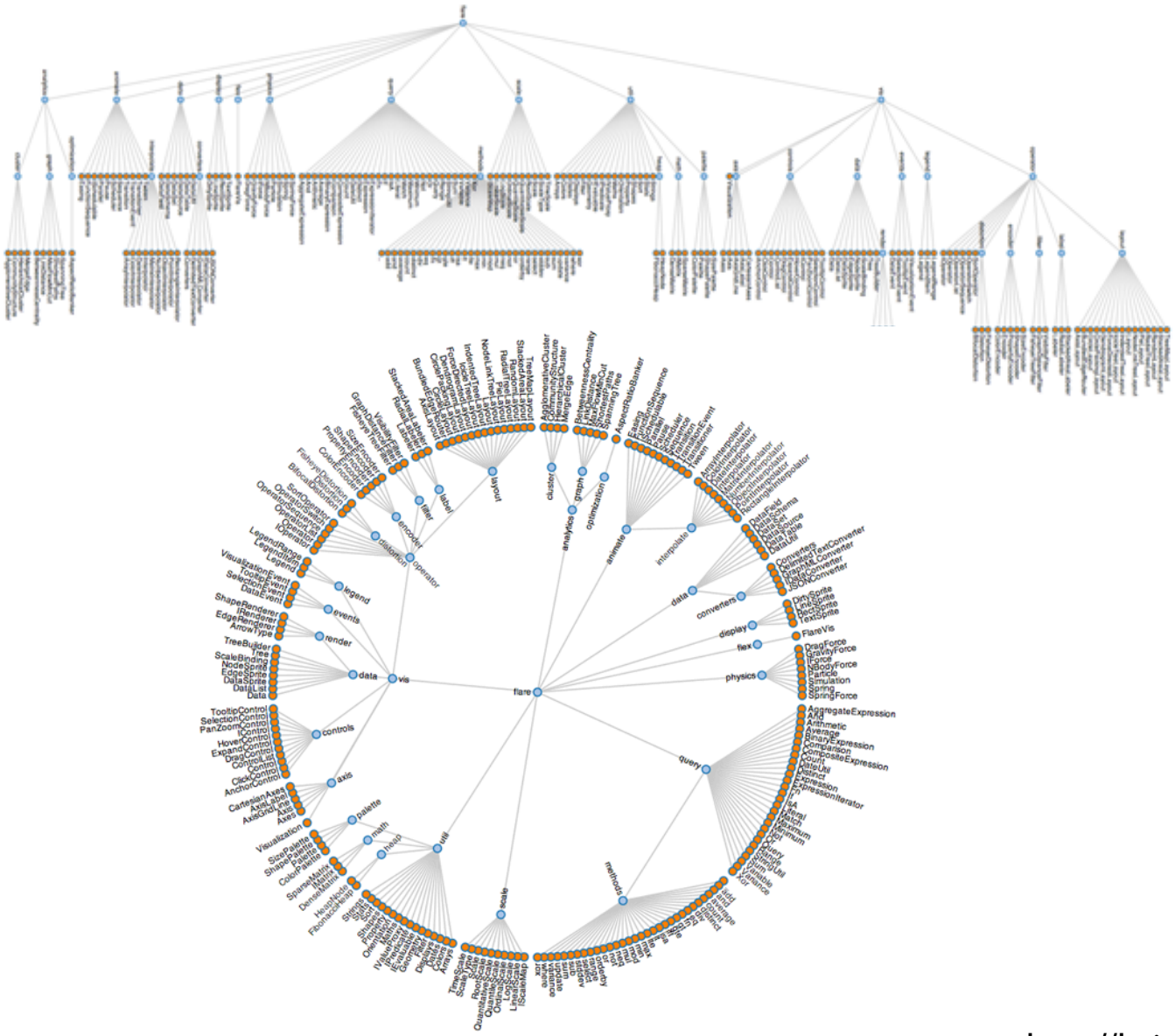
http://hci.stanford.edu/jheer/files/zoo/

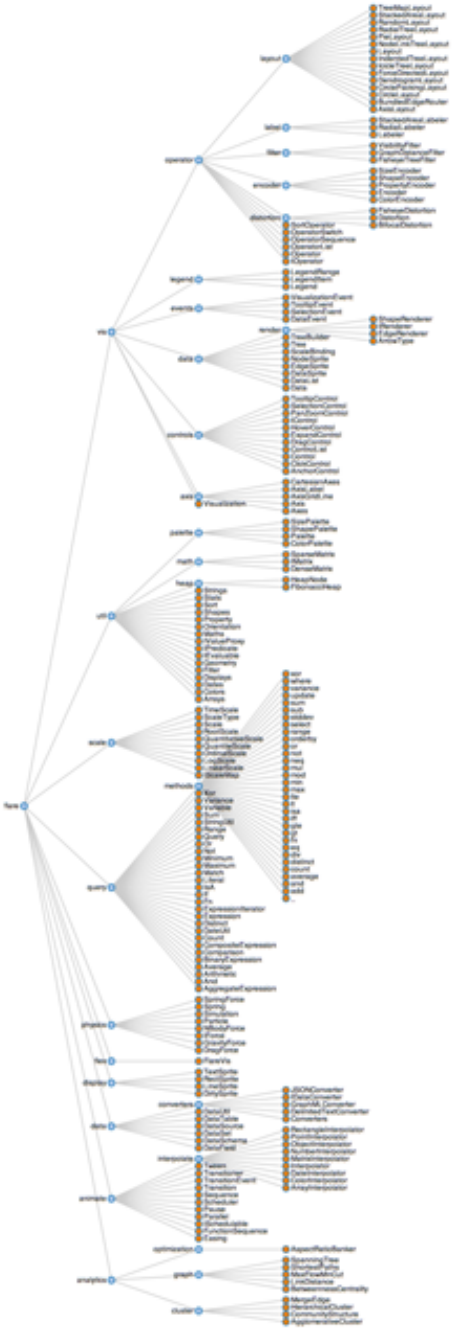
Icicle Diagrams



http://hci.stanford.edu/jheer/files/zoo/

Visualizing Hierarchies

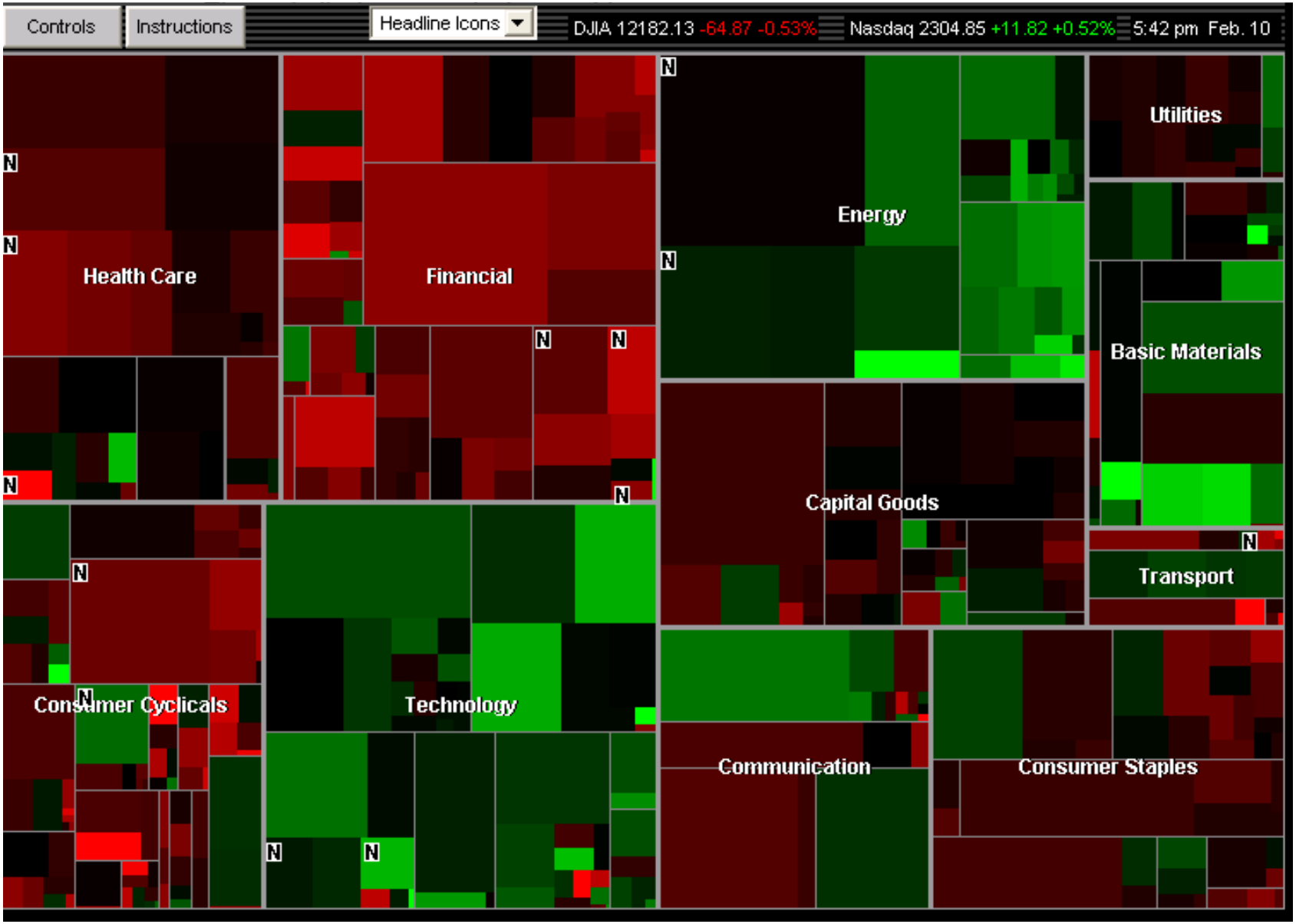




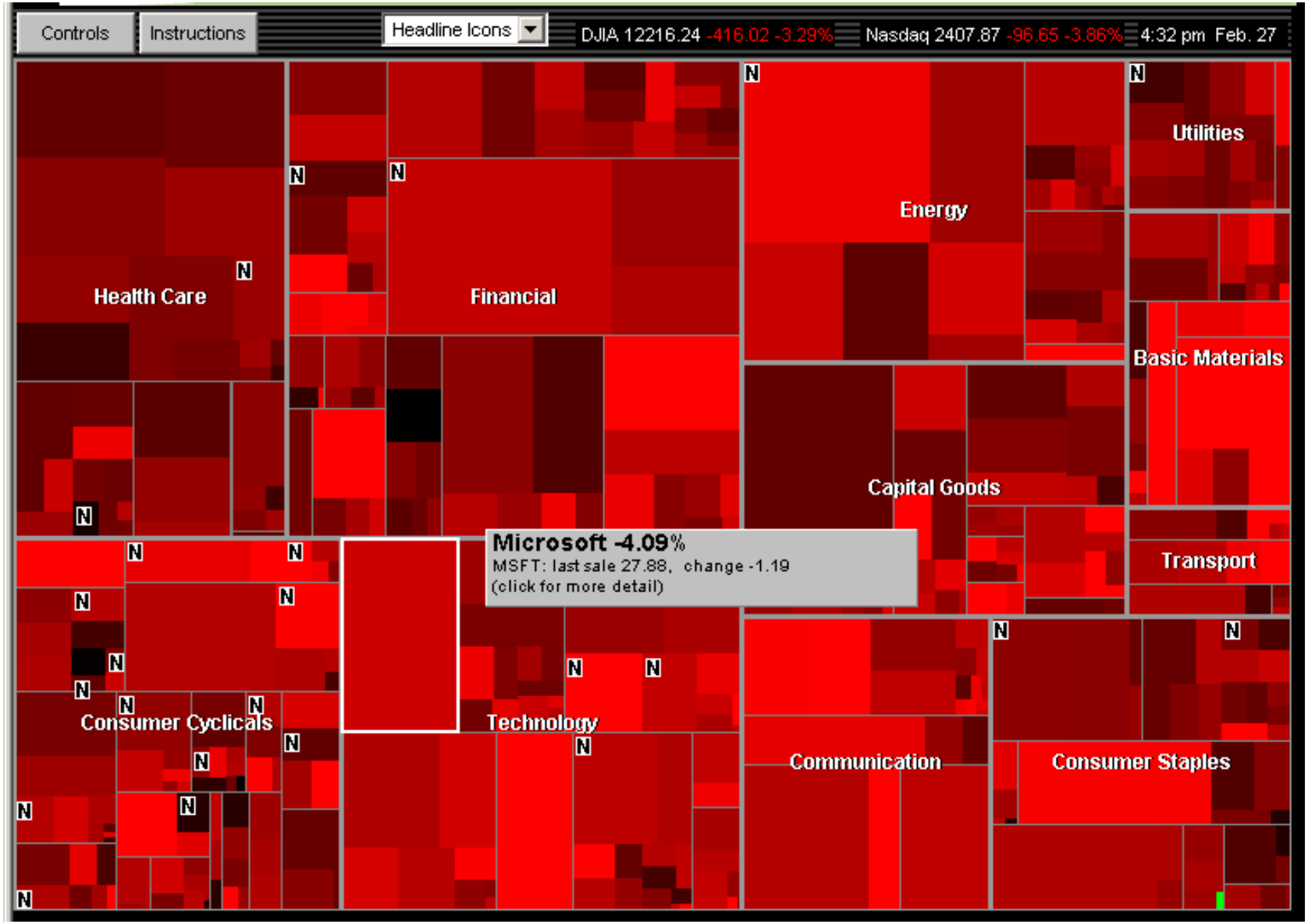
http://hci.stanford.edu/jheer/files/zoo/

TreeMaps

- Useful for representing hierarchical relationships
- Two mechanisms for coding information:
  - Size and Color
  - Physical location of boxes not usually meaningful
- Requires learning to read

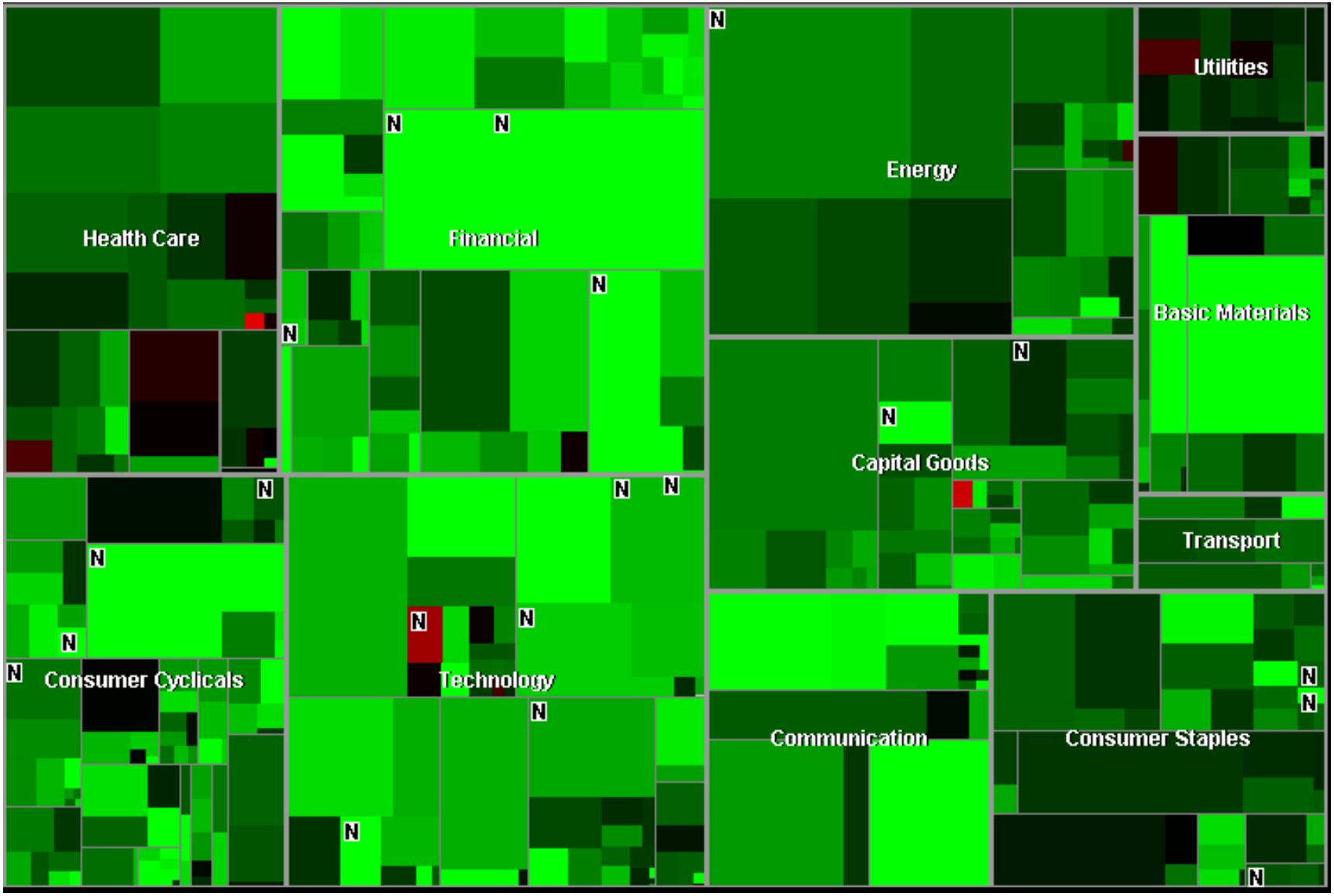


http://www.smartmoney.com/map-of-the-market/



Market mixed, February 8, 2008:

Energy & Technology up Financial & Health Care down

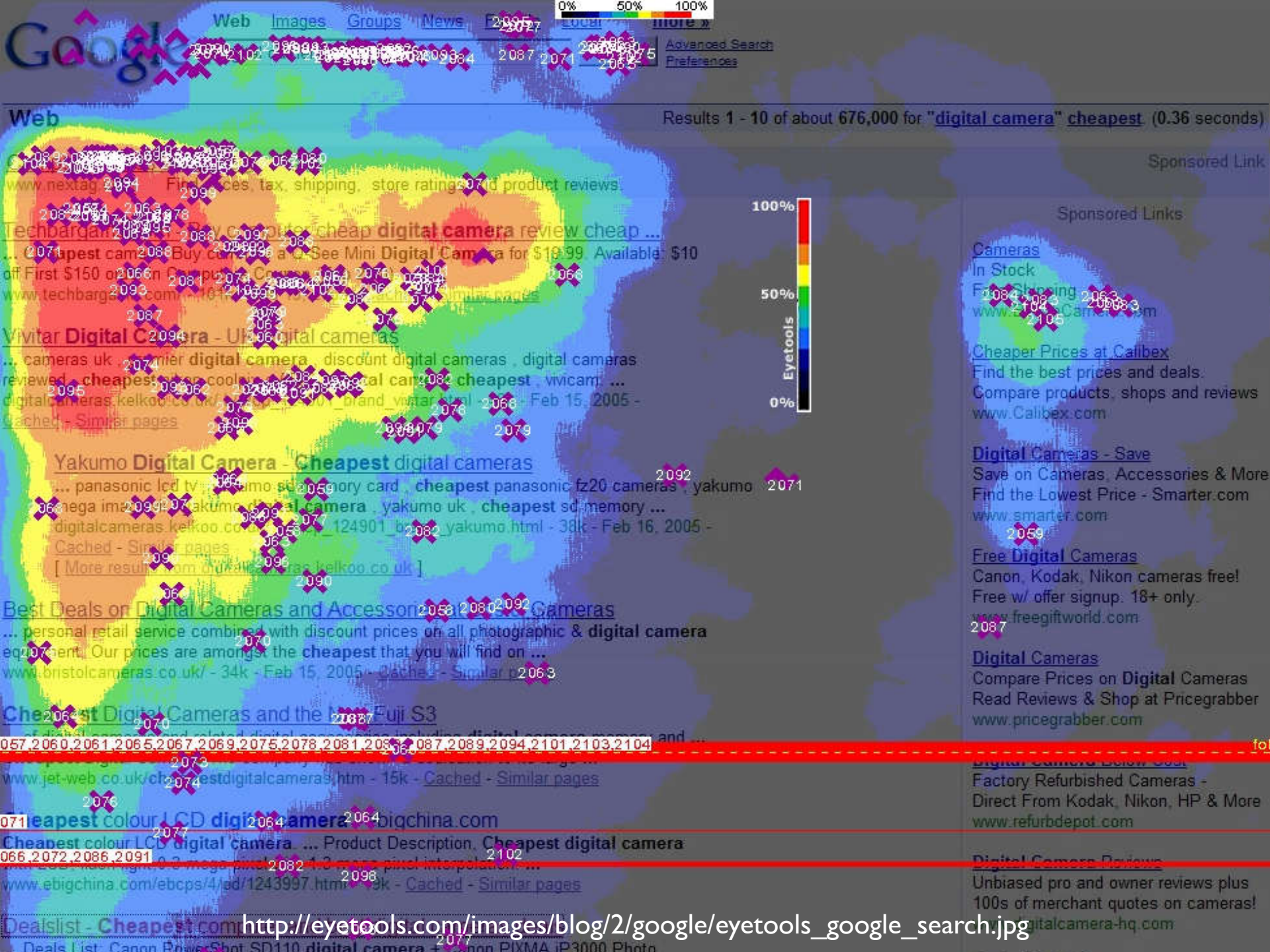


Market rises 319 points, November 13, 2007:

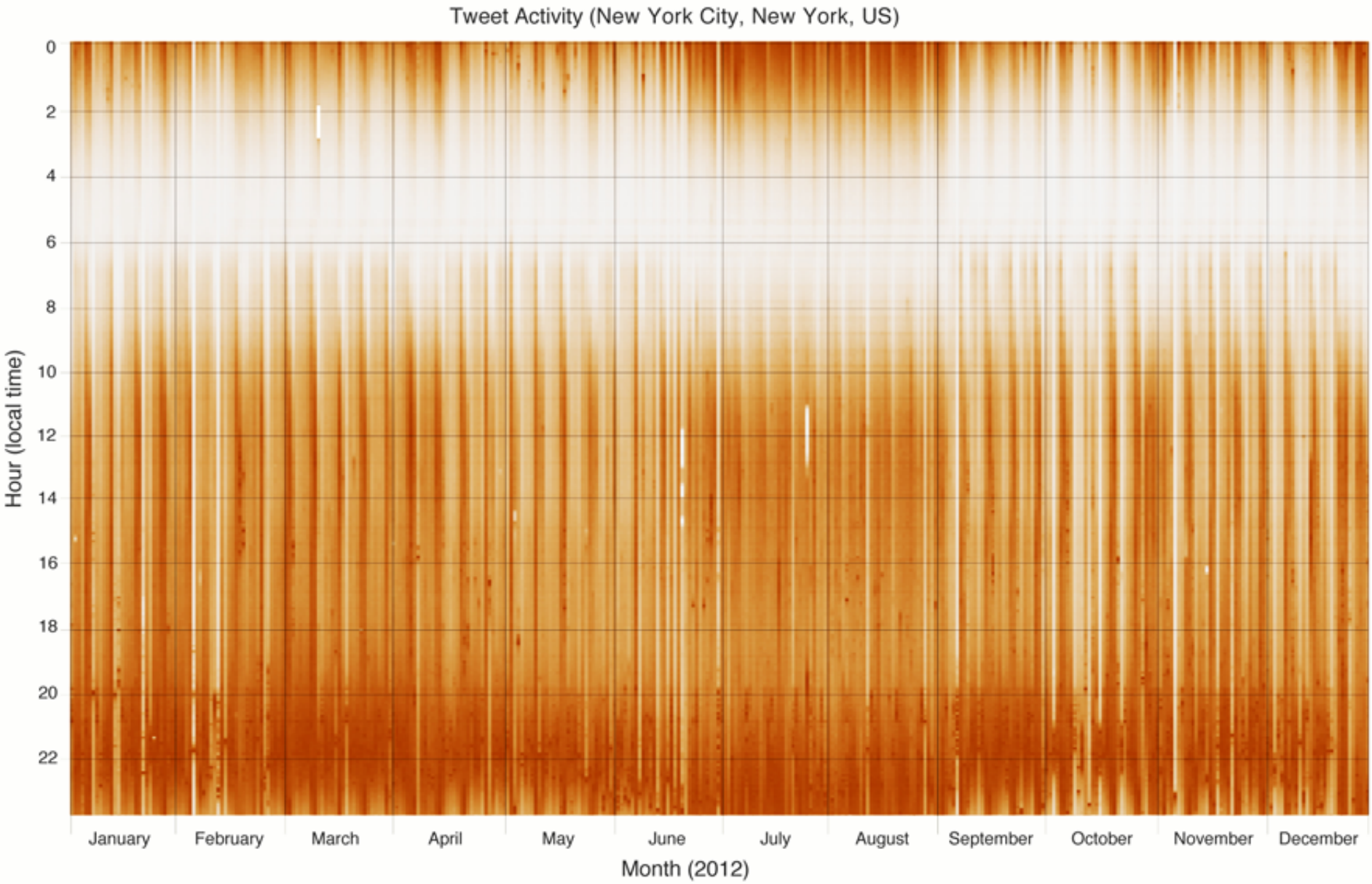
Five exceptions

Heatmaps

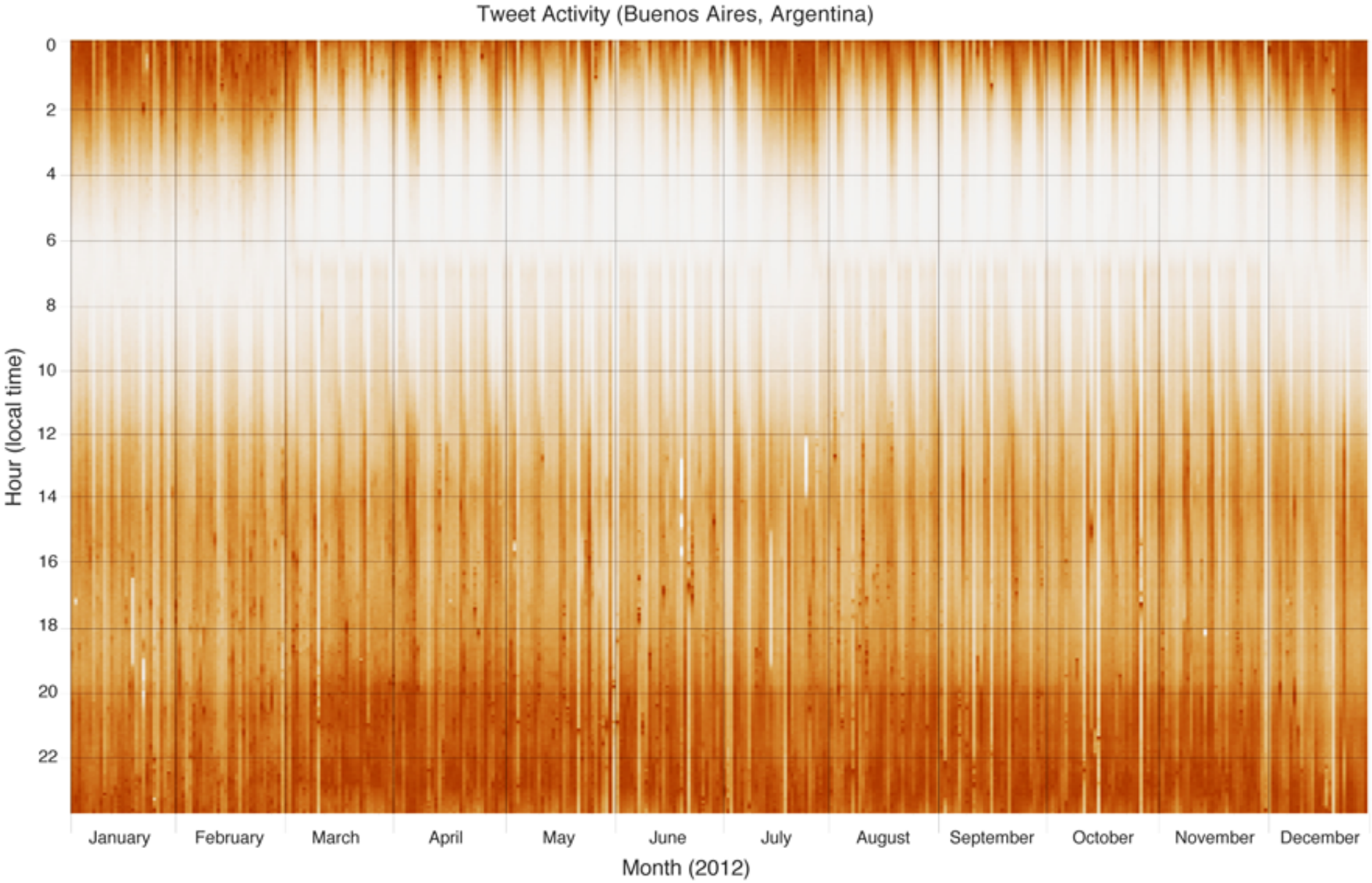
- Data represented in a matrix
- Matrix values are code coded



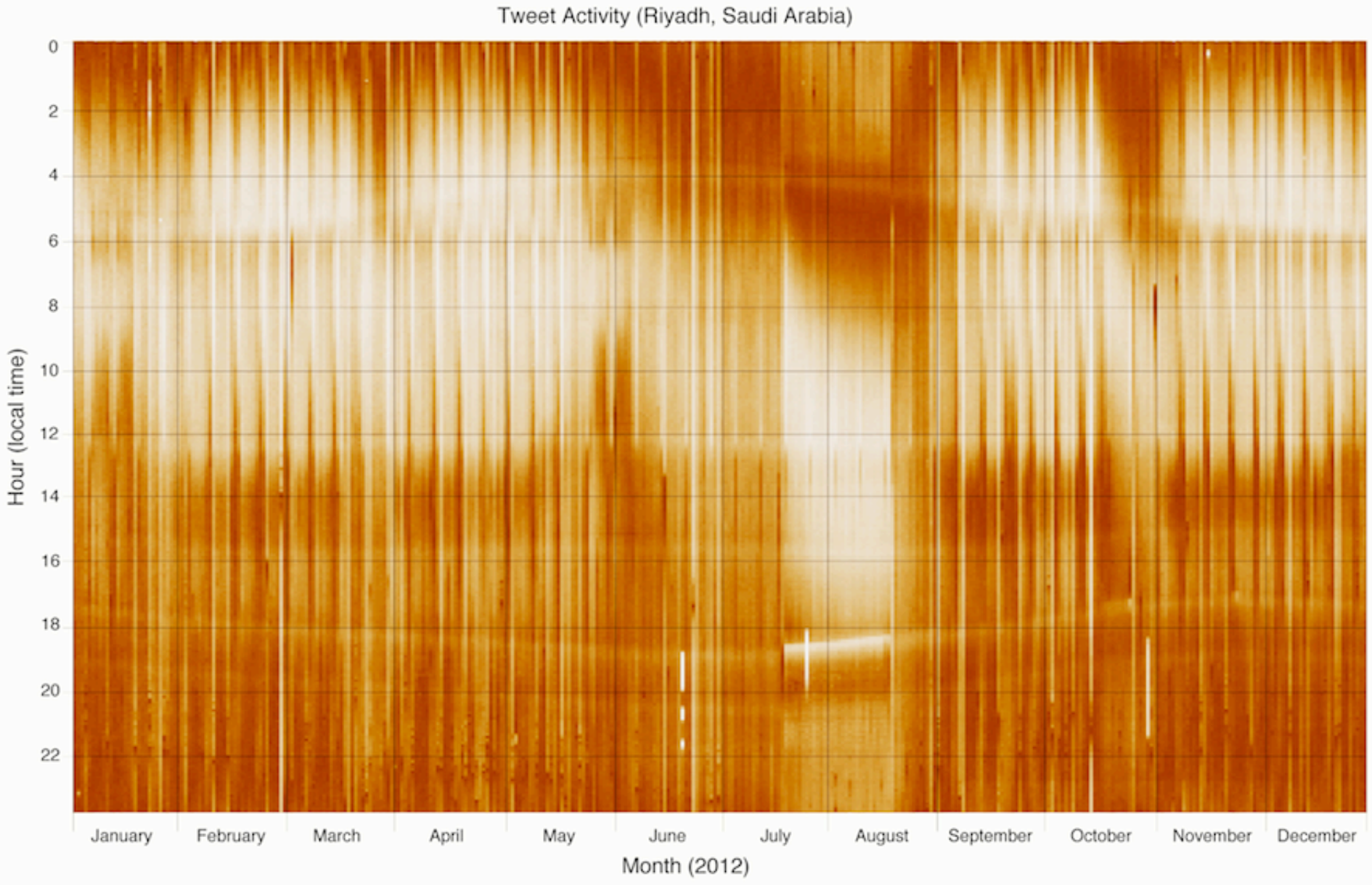
http://eyetools.com/images/blog/2/google/eyetools_google_search.jpg



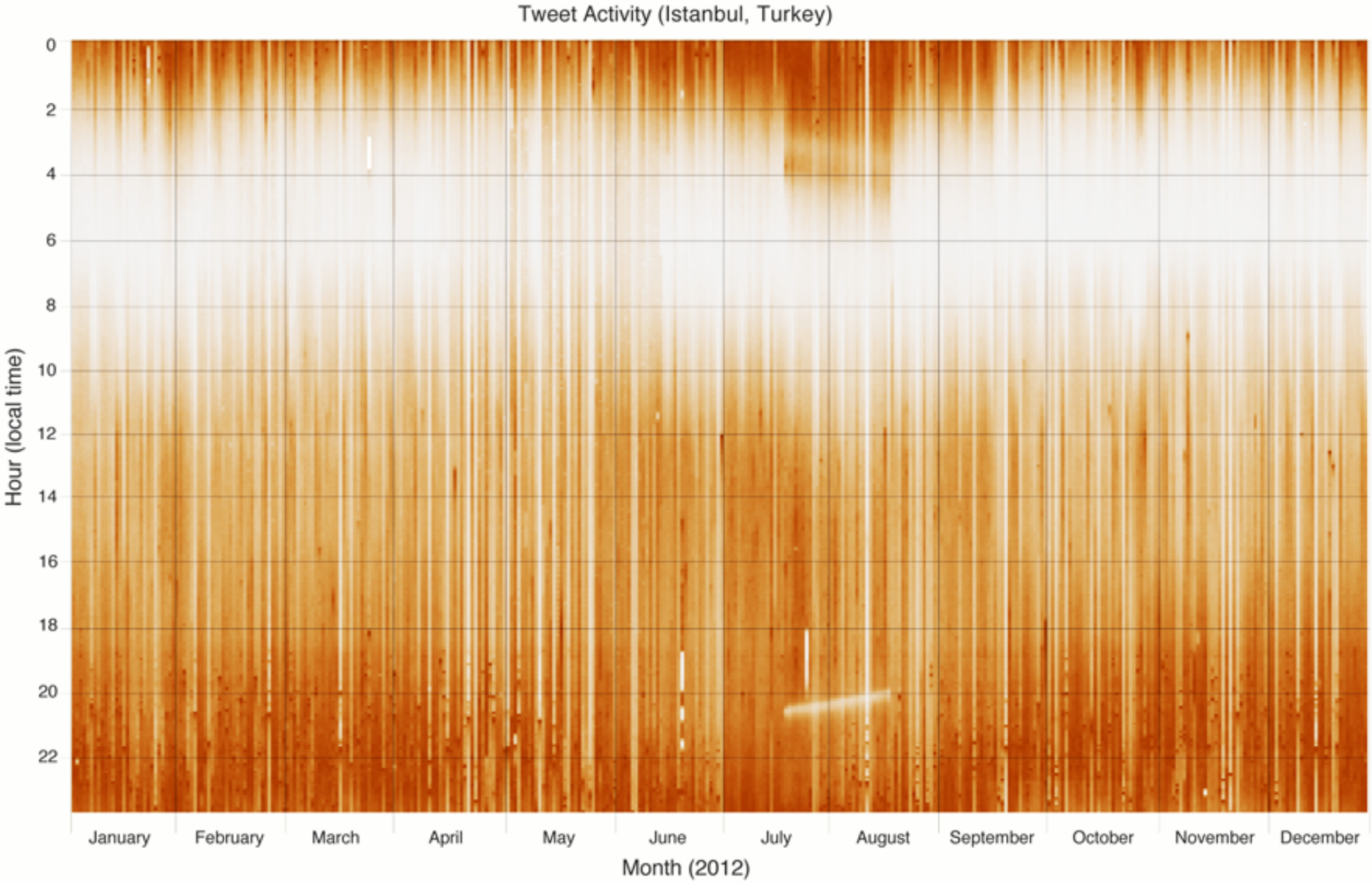
Miguel Rios and Jimmy Lin. Visualizing the "Pulse" of World Cities on Twitter. ICWSM 2013.



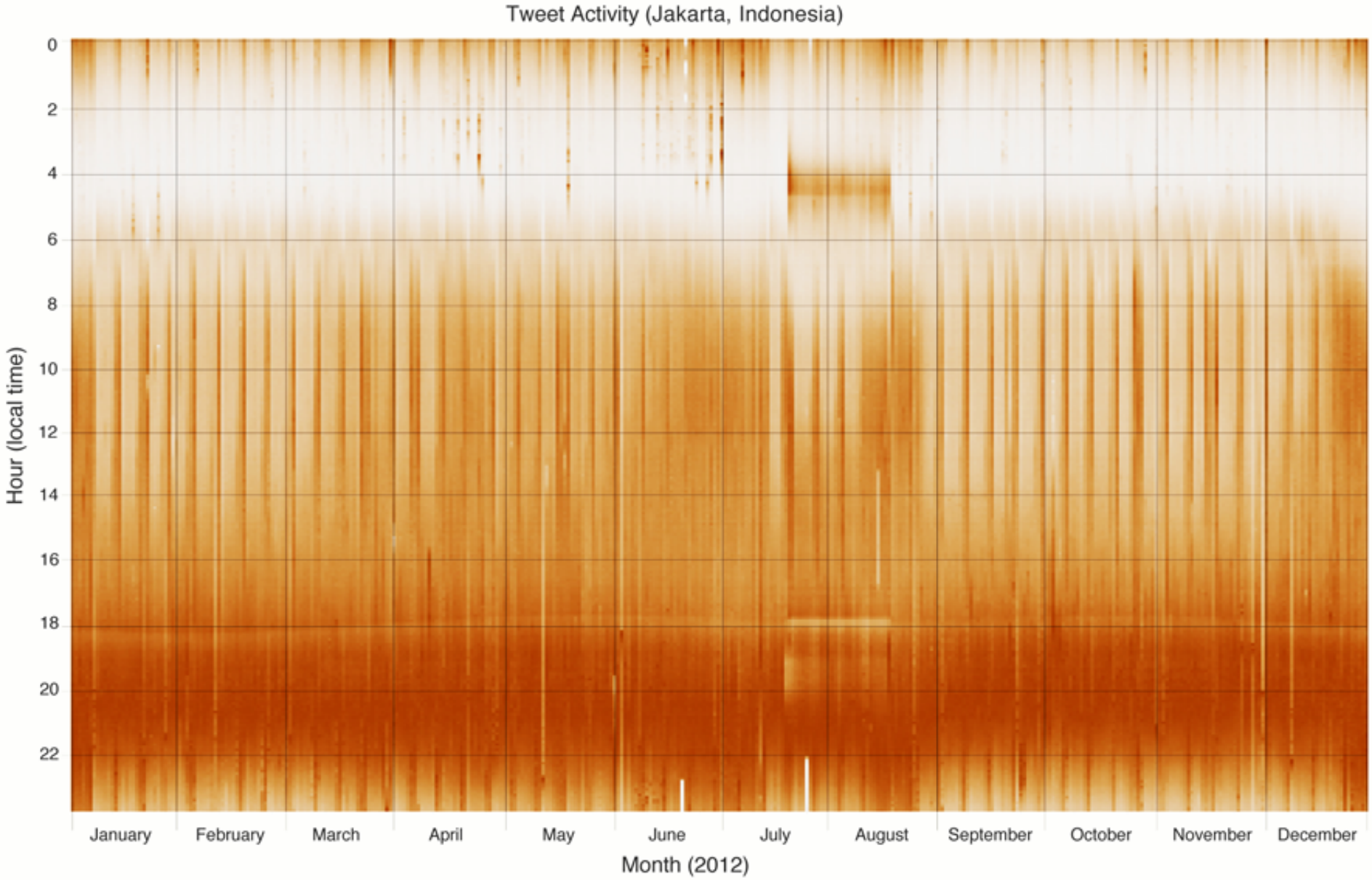
Miguel Rios and Jimmy Lin. Visualizing the "Pulse" of World Cities on Twitter. ICWSM 2013.



Miguel Rios and Jimmy Lin. Visualizing the "Pulse" of World Cities on Twitter. ICWSM 2013.

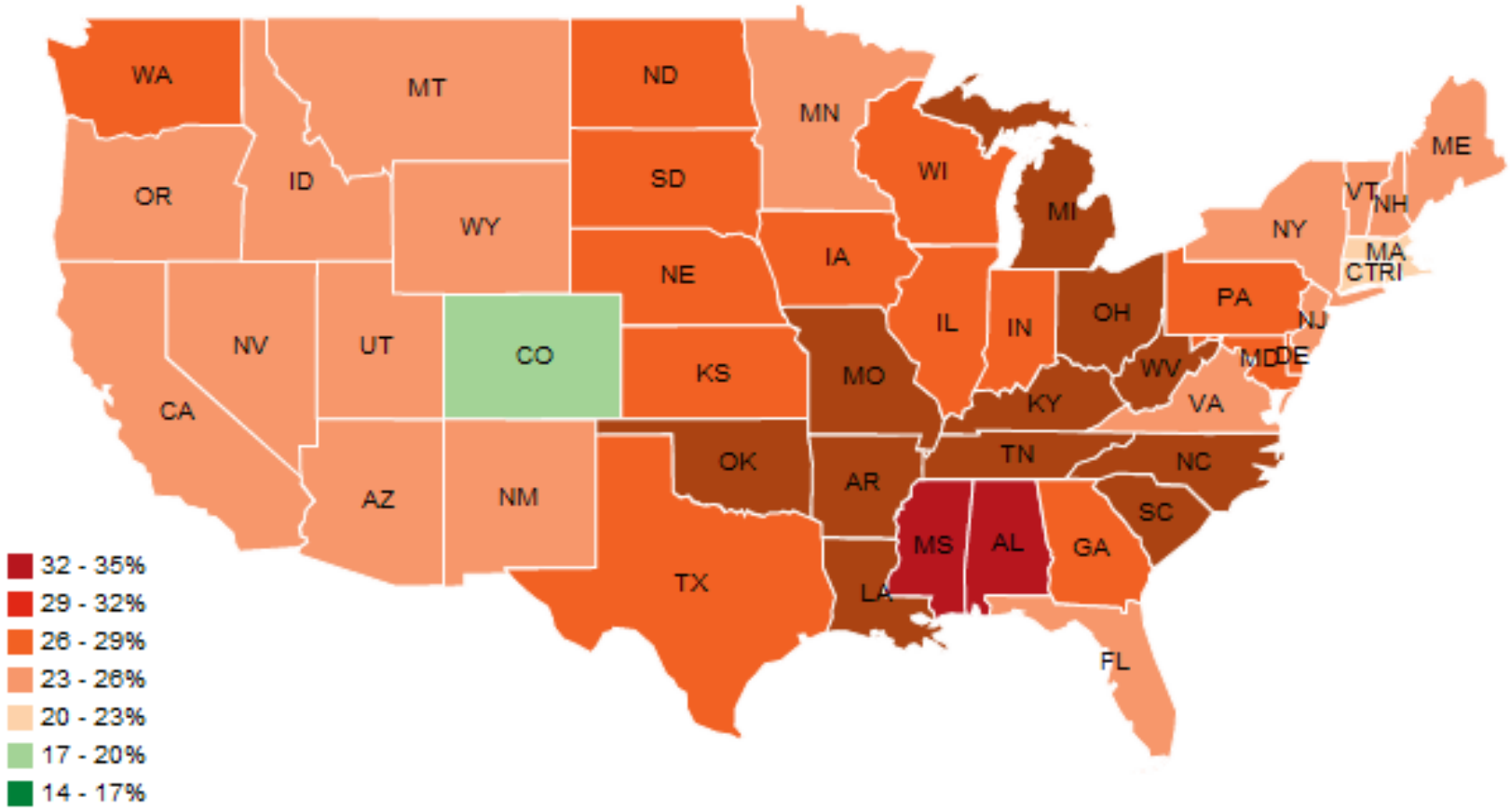


Miguel Rios and Jimmy Lin. Visualizing the "Pulse" of World Cities on Twitter. ICWSM 2013.

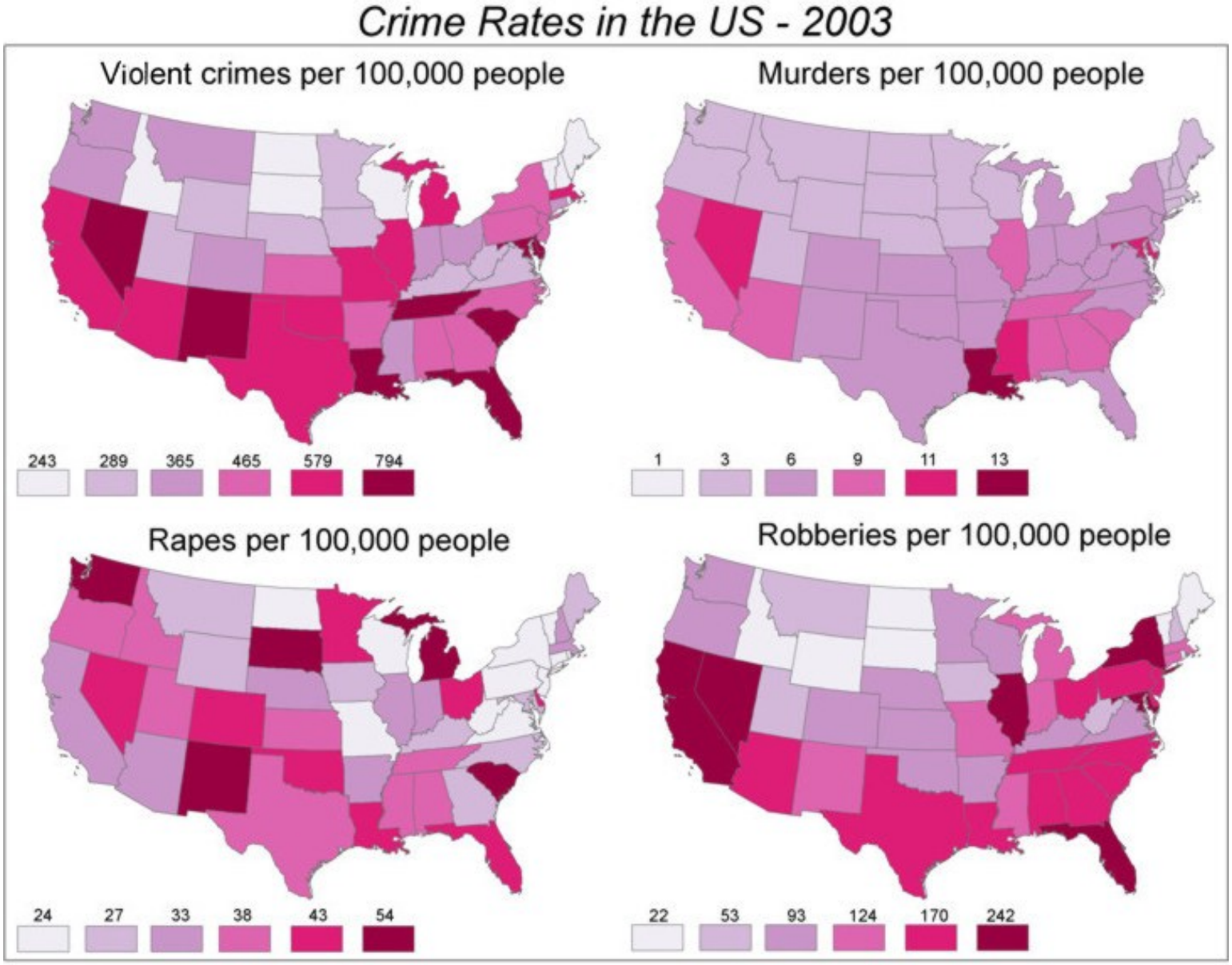


Miguel Rios and Jimmy Lin. Visualizing the "Pulse" of World Cities on Twitter. ICWSM 2013.

Choropleth Maps

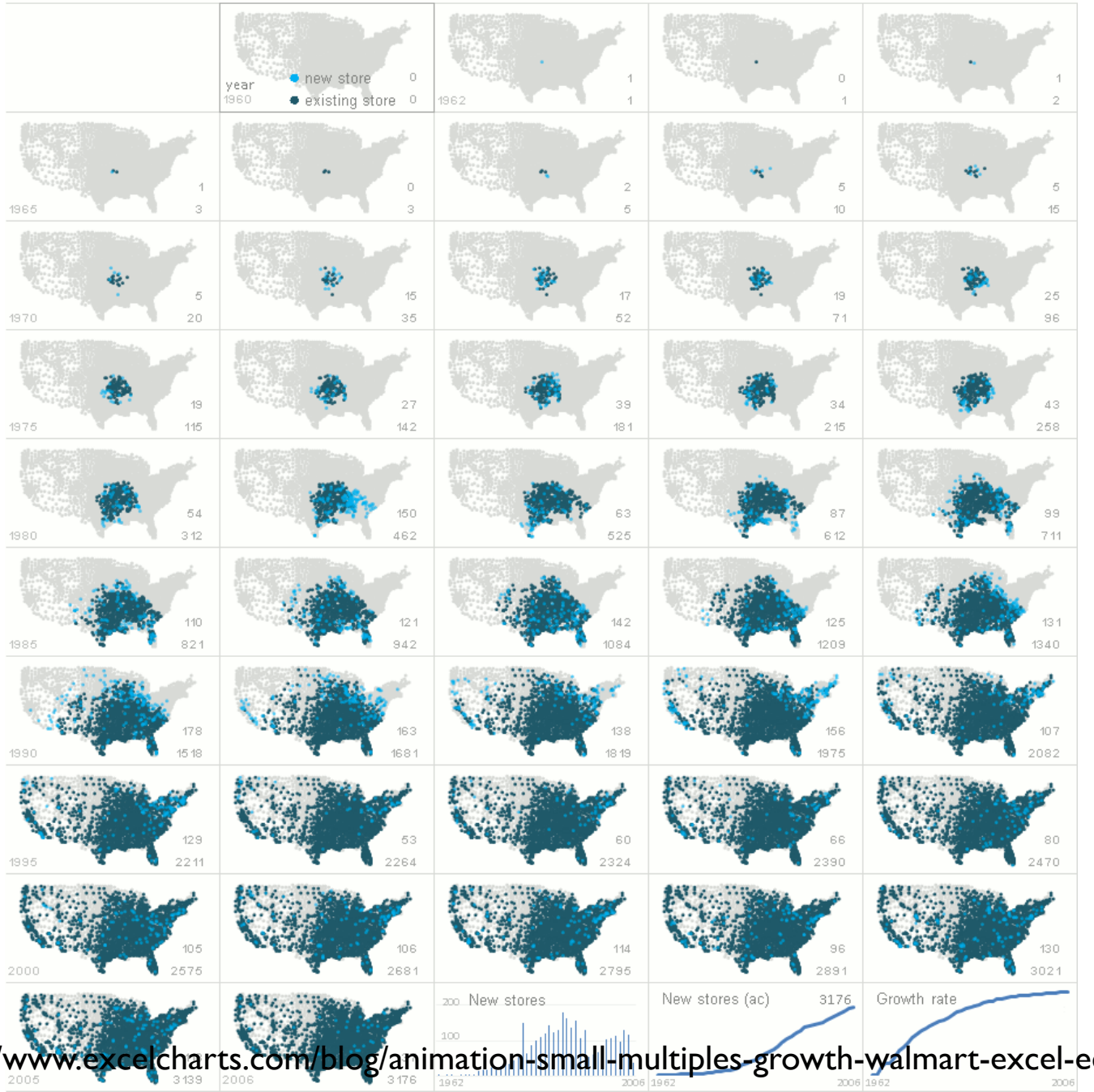


http://hci.stanford.edu/jheer/files/zoo/



http://www.geog.ucsb.edu/~jeff/gis/choropleth_maps.html

Growth of Walmart



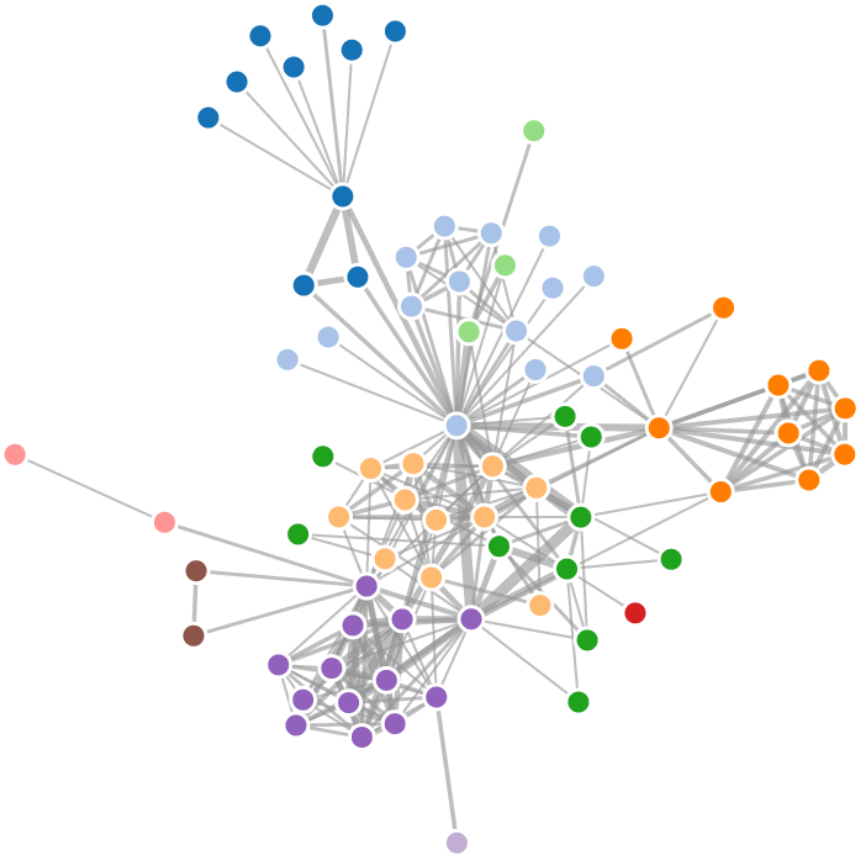
http://www.excelcharts.com/blog/animation-small-multiples-growth-walmart-excel-edition/

Small Multiples vs. Animations

- Advantages and disadvantages?

Visualizing Networks

- Types of networks: social networks, interaction networks, biological networks, computer networks, etc.
- Standard algorithm uses force-directed layout



SAN DIEGO
ZOO
Panda Viewing
9am - 5:30pm