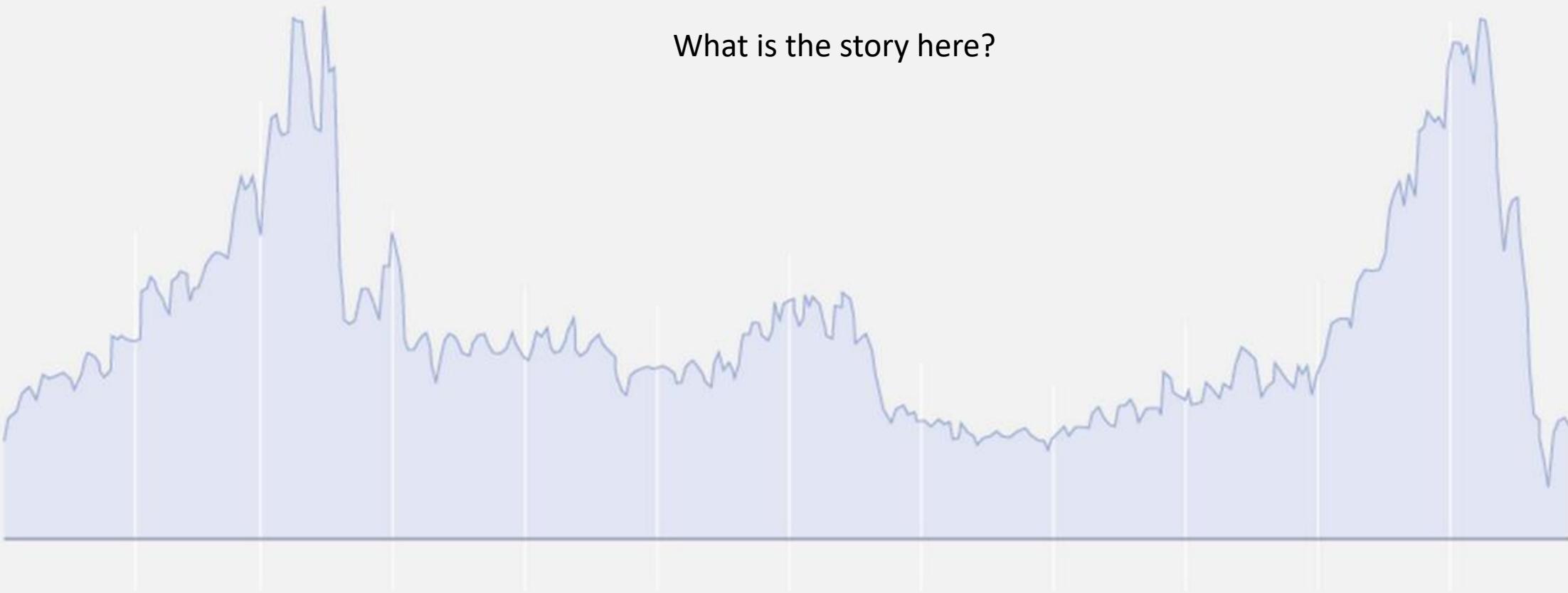


# YOUR DATA TELLS A STORY



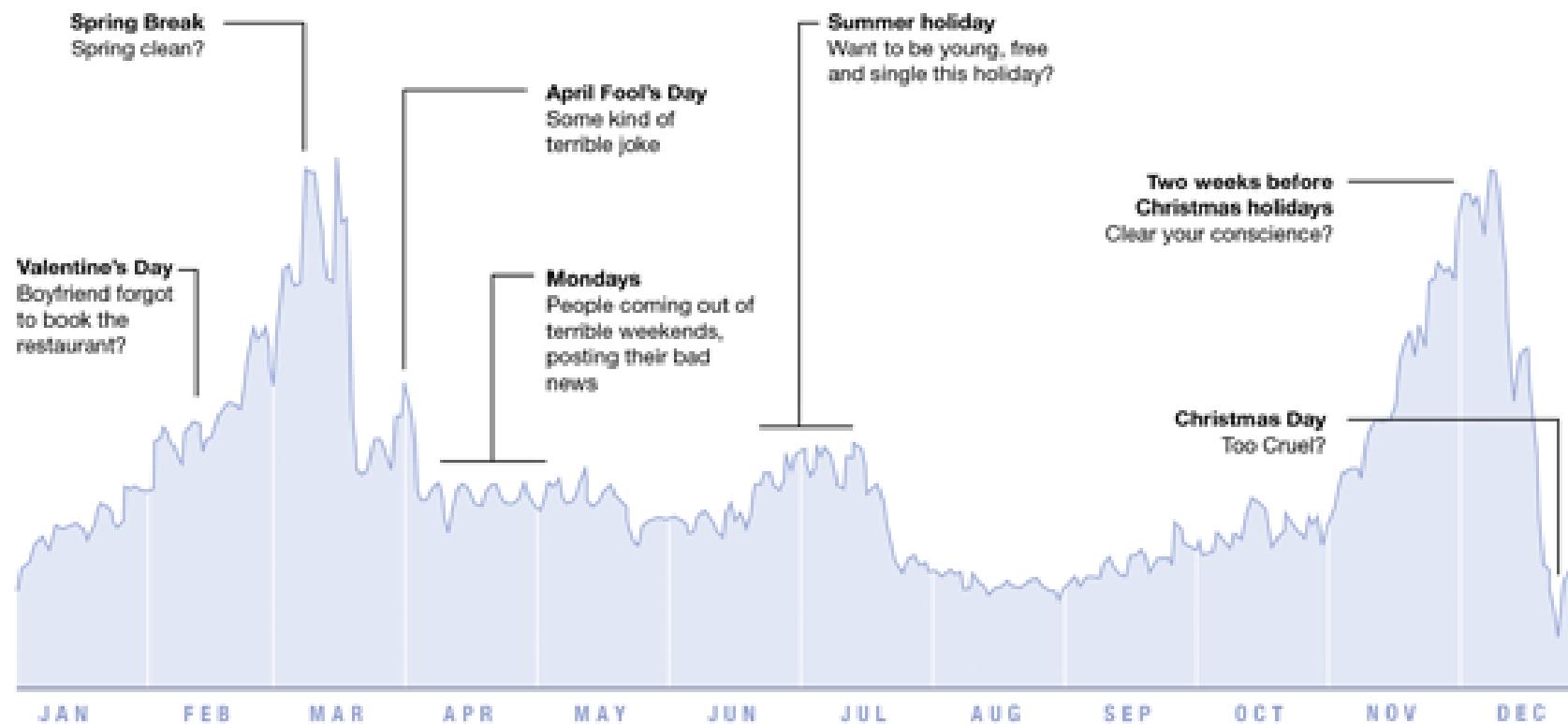


Peak Break-Up Times On Facebook

# Peak Break-up Times

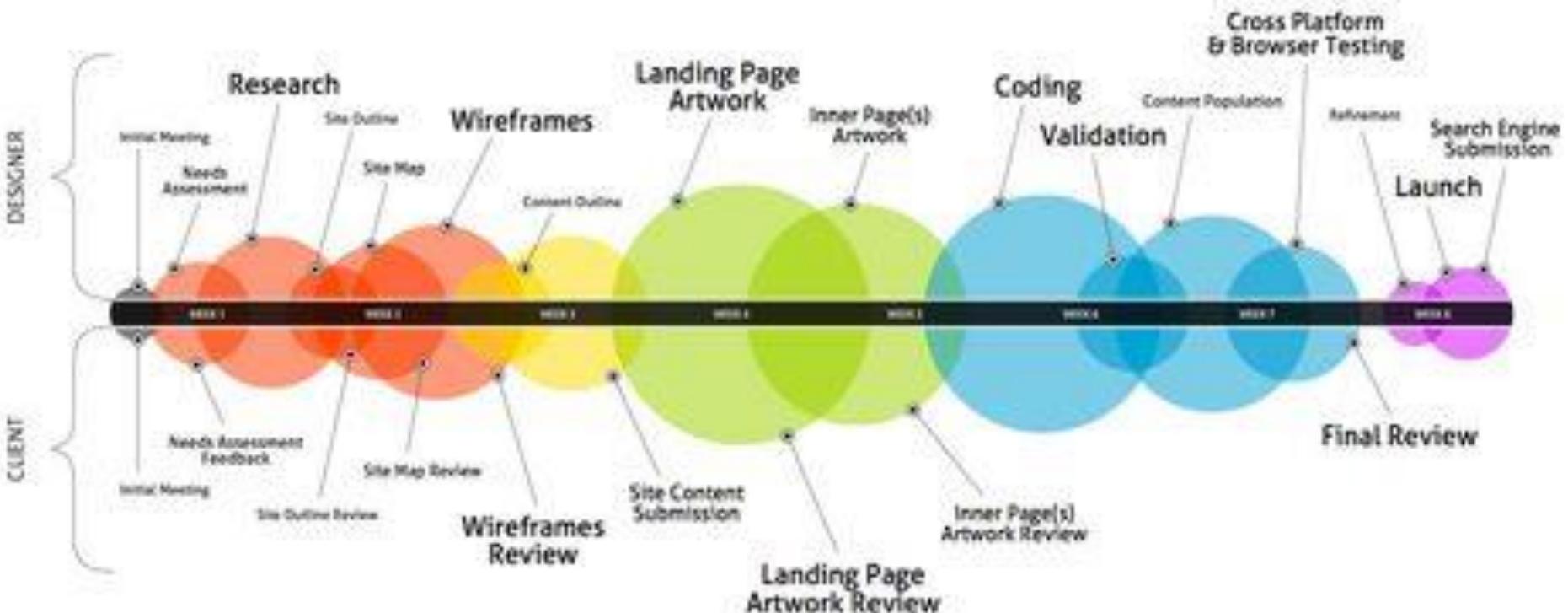
According to Facebook status updates

InformationIsBeautiful.net



# A Web Site Designed

MILESTONES, INVOLVEMENT, IMPORTANCE & TIMELINE



## MILESTONES

- Milestone Document

## INVOLVEMENT



## IMPORTANCE

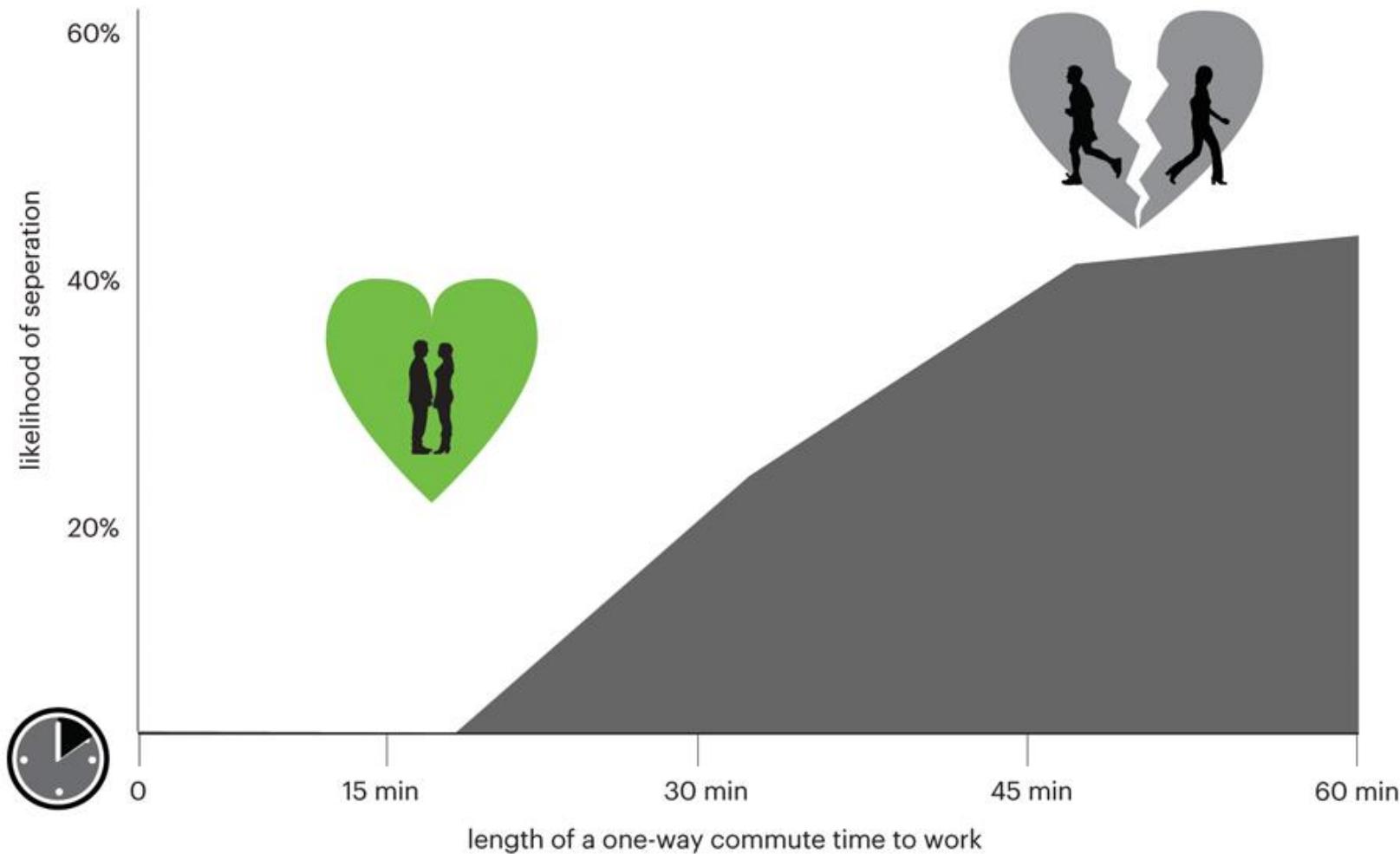
Average  
Greater

## PHASE



## ROMANCE & COMMUTING TIME

Long commutes put unnecessary strain on relationships. Commute times of over 45 minutes each way lead to a 40% higher rate of separation or divorce among couples that lived together.



# Anatomy of a Graph

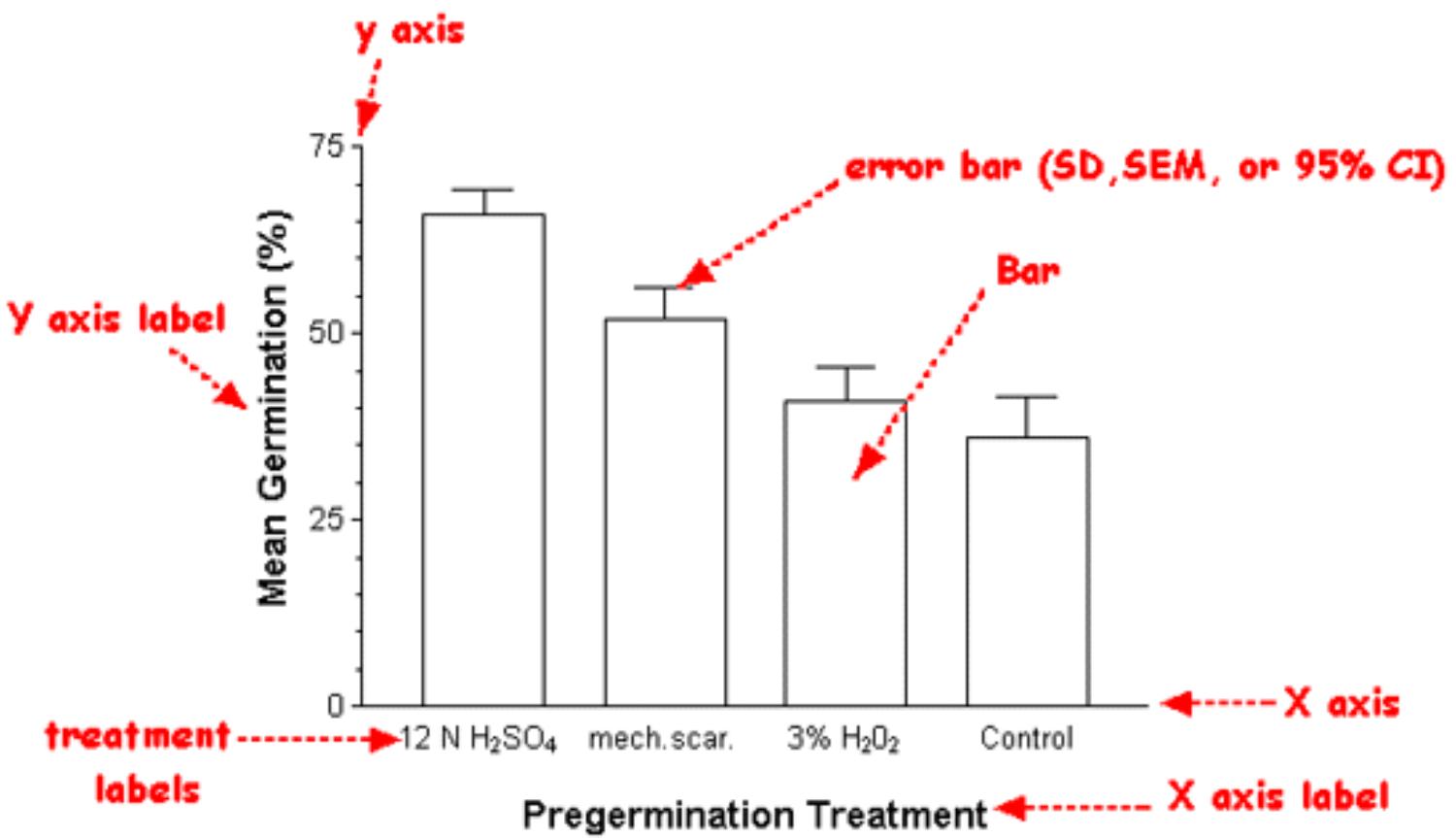


Figure 1. Mean germination (%) of gourd seeds following various pregermination treatments. N=10 groups of 100 seeds per treatment and control. Treatments: 12 hour soak in 12 N H<sub>2</sub>SO<sub>4</sub>, 90 second scarification of seed coat with 80 grit sandpaper, 6 hour soak in 3% H<sub>2</sub>O<sub>2</sub>.

figure legend

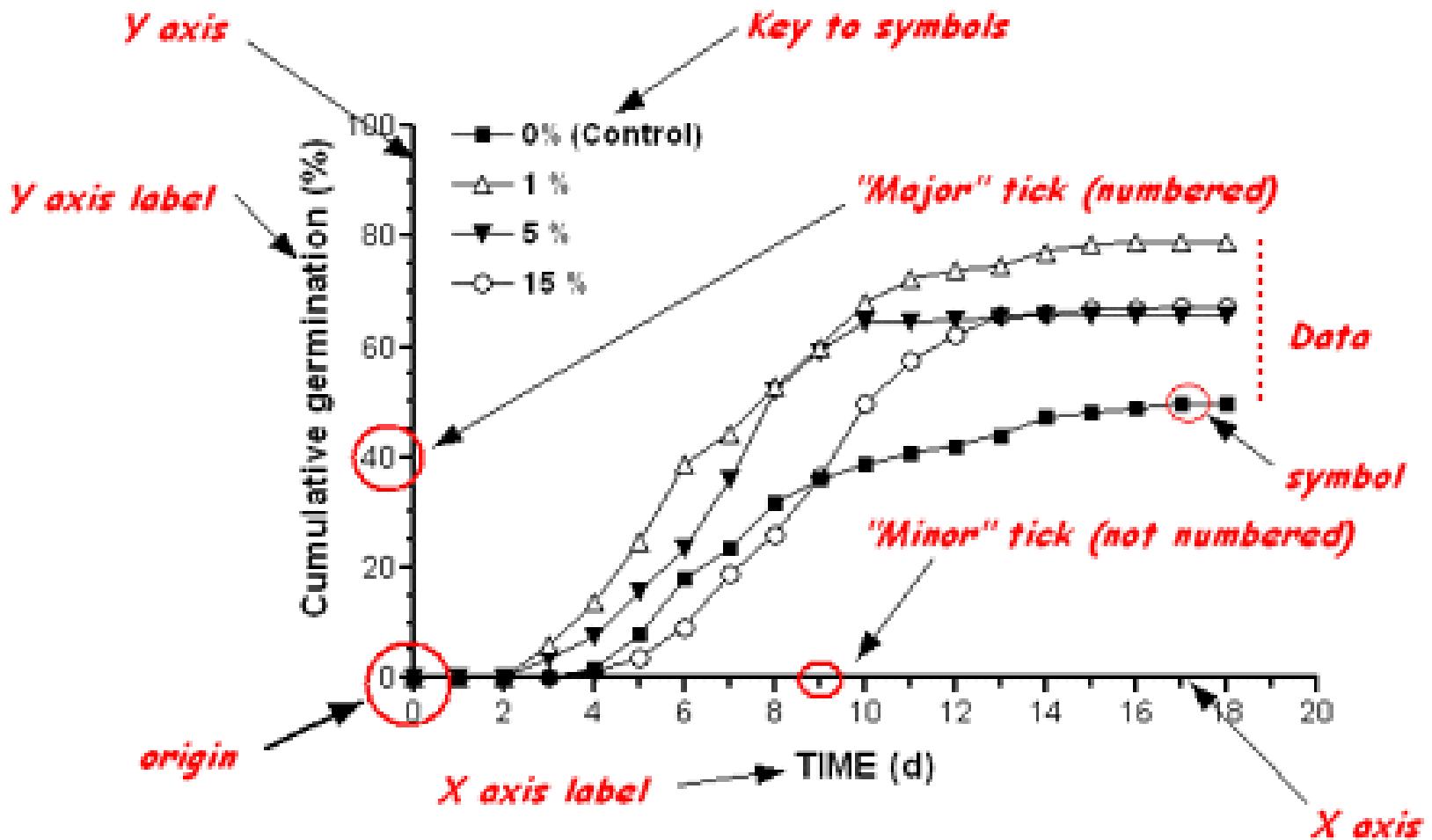


Figure 1. Cumulative germination of *Chenopodium* seeds after pregermination treatment of 2 day soak in NaCl solutions.

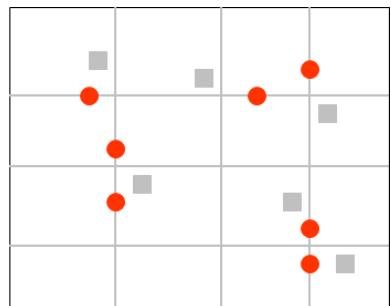
n = 1 trial per treatment group (100 seeds/trial.)

# Graphical Hierarchy

# GROUND AND FIGURE

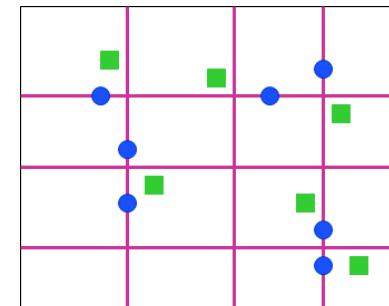
- Assign bright colors (red, orange, yellow, green, blue) to important graphic elements
- Features are known as **figure**, or the **subject** of the map.
  
- Assign drab colors to the graphic elements that provide orientation or context, especially shades of gray
- Features known as **ground**, or the **context** for the subject.

GOOD

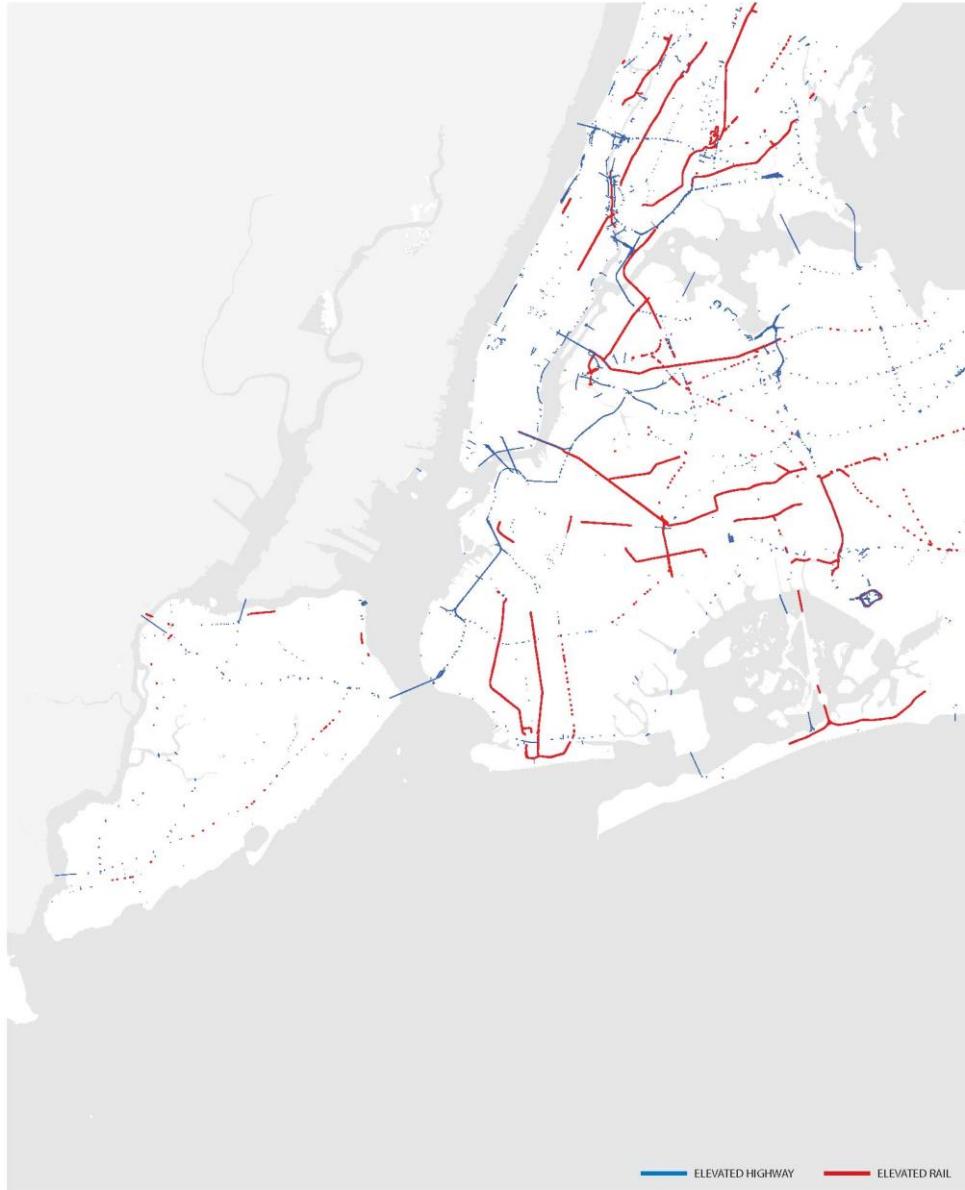


Circles in figure, squares and lines in ground

BAD



All features in figure

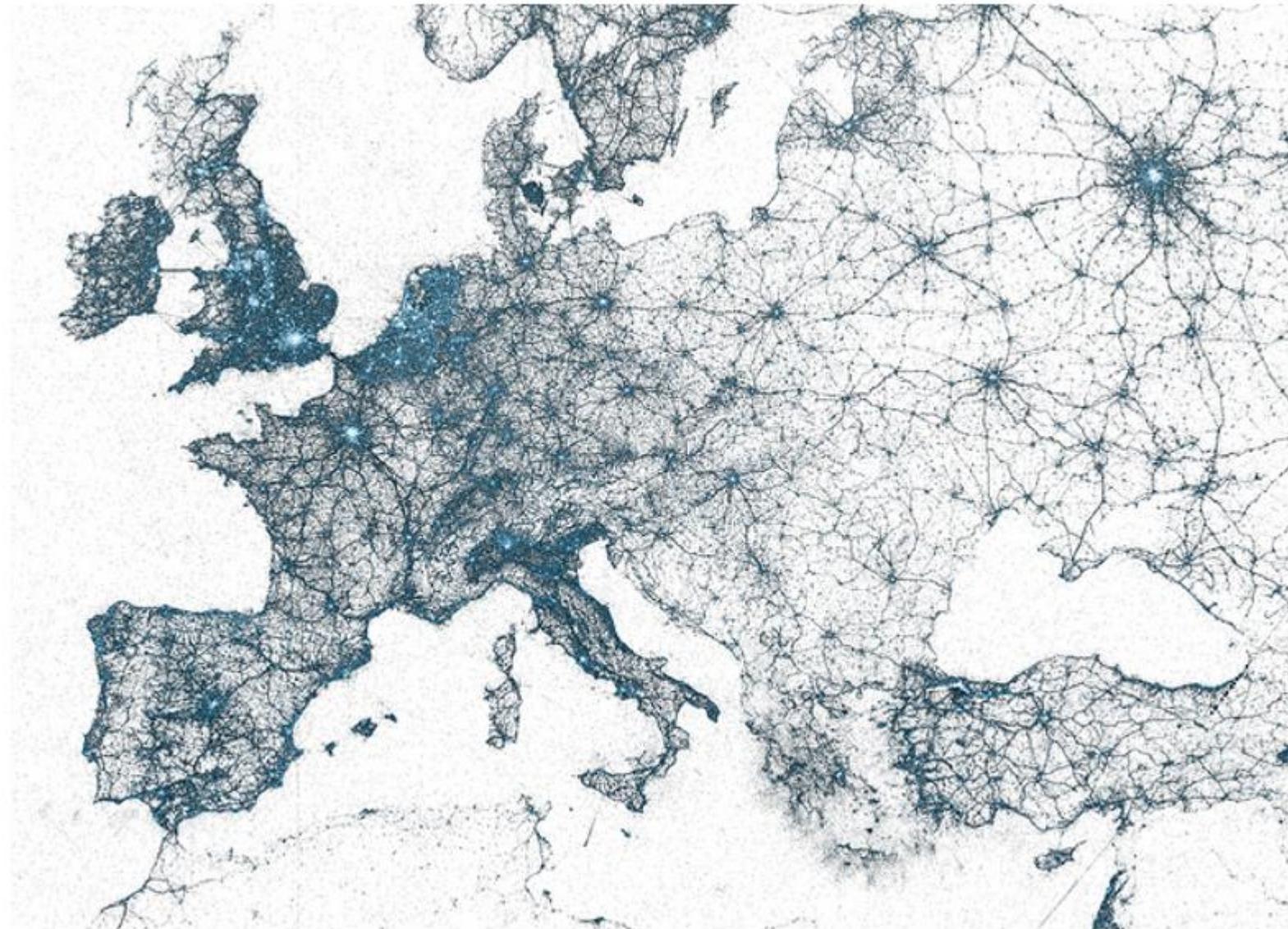


DESIGN TRUST  
FOR PUBLIC SPACE

ELEVATED INFRASTRUCTURE  
SOURCE: PLUTO 2003/DOT/T 2009



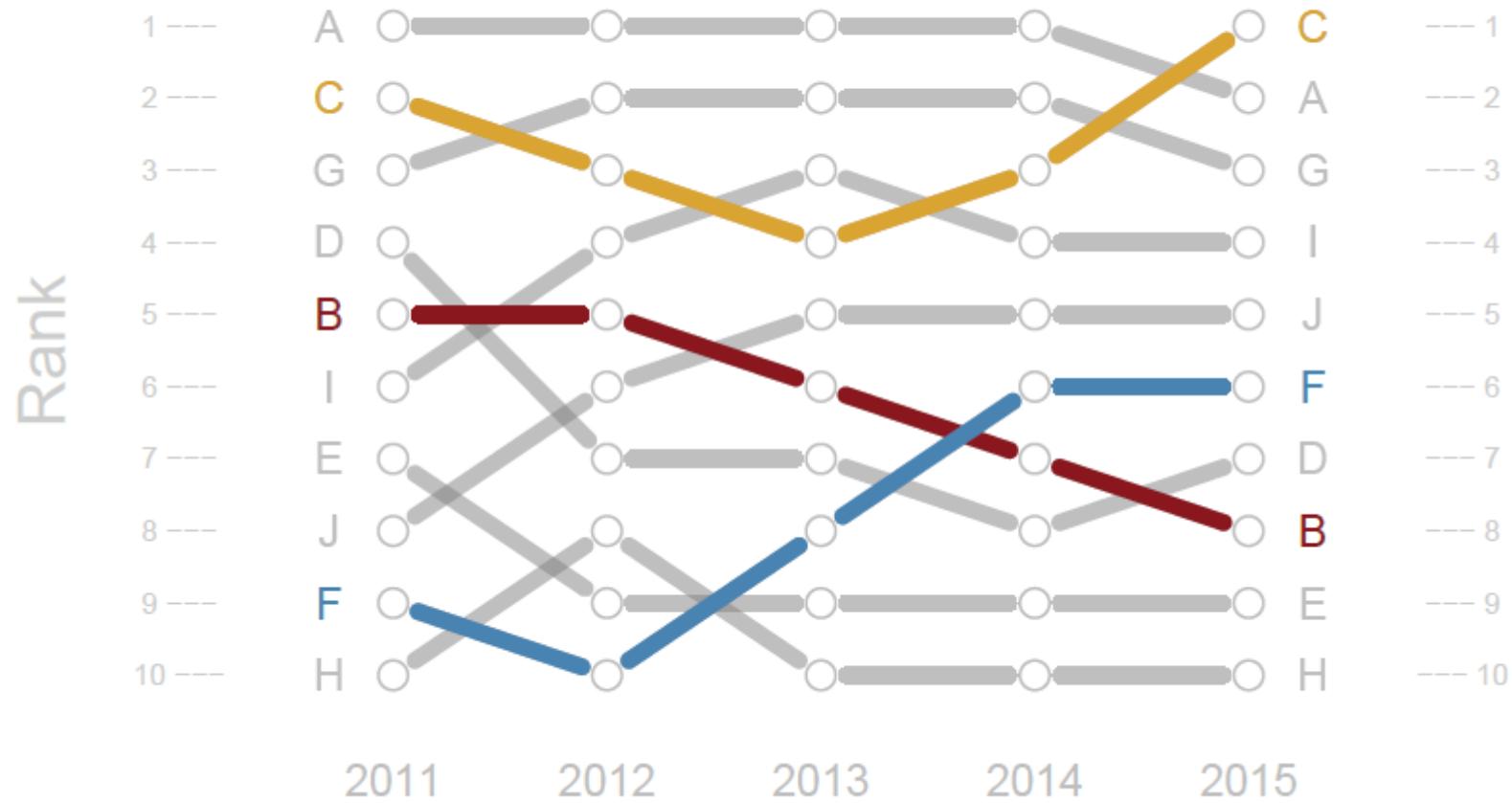
A single visualization can reveal unforeseen patterns present in truly massive data sets. For example, Twitter [created visualizations of geotagged tweets](#) to combine billions of data points into beautiful images of human activity.



# CONNECTING POCKET PARKS MAKES CITIES COHERENT

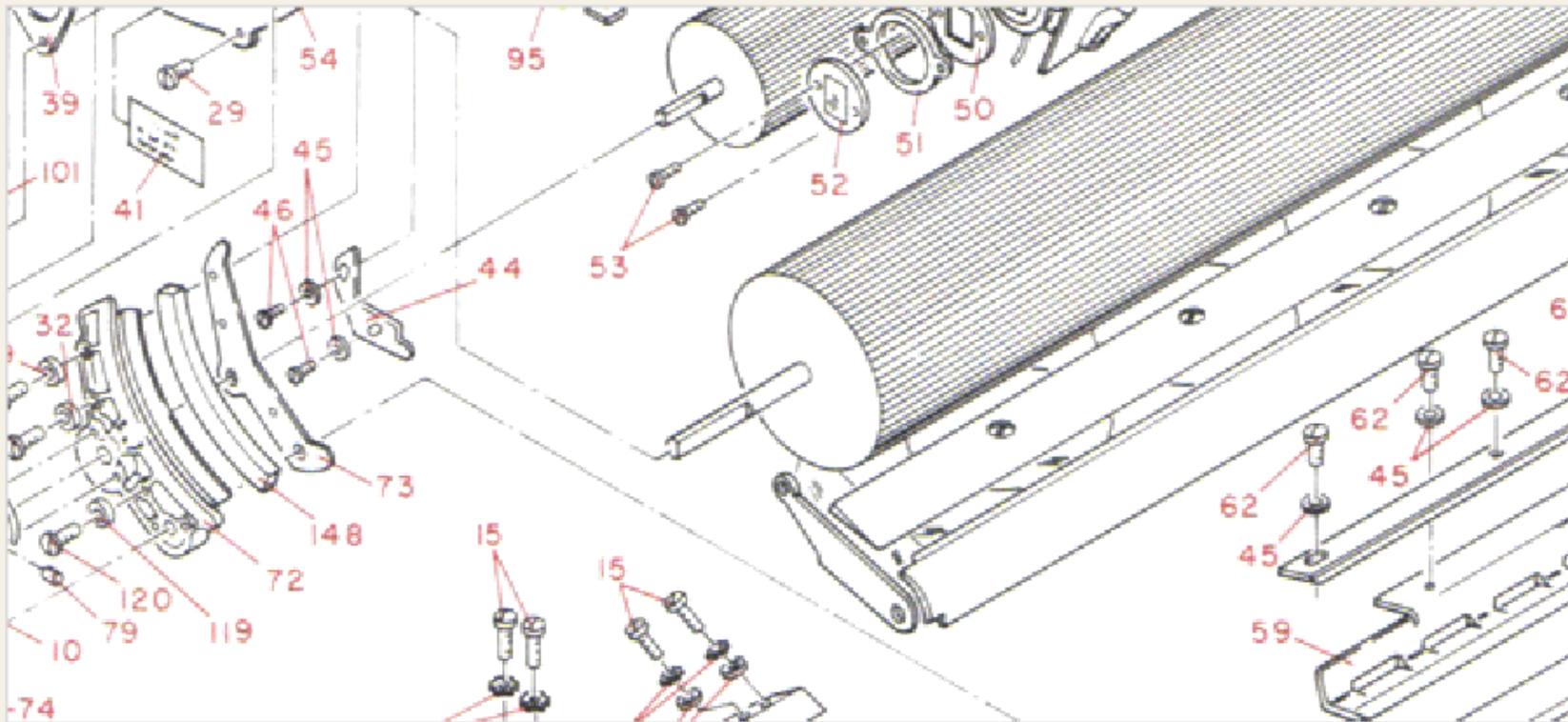


## School Rank 2011–2015



# Label Data

## 5: Use Labels



Labels help data stand out and have meaning.

[REPLAY](#)

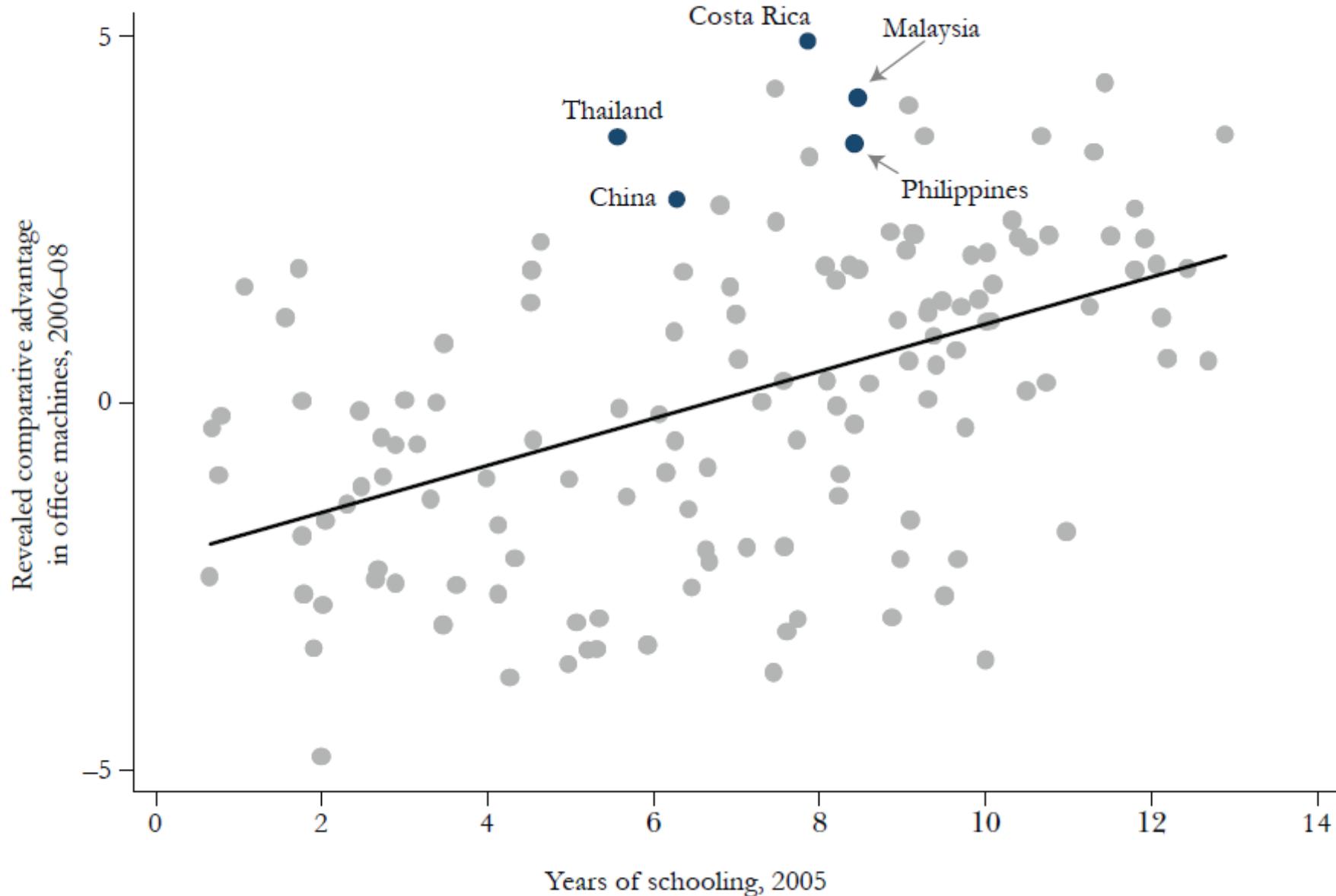
IN THIS example, the labels represent numbers that are associated with the specific parts.

Without the labels, it would be very difficult to keep track of all the parts.

Labels are informative guides and are essential in data design.

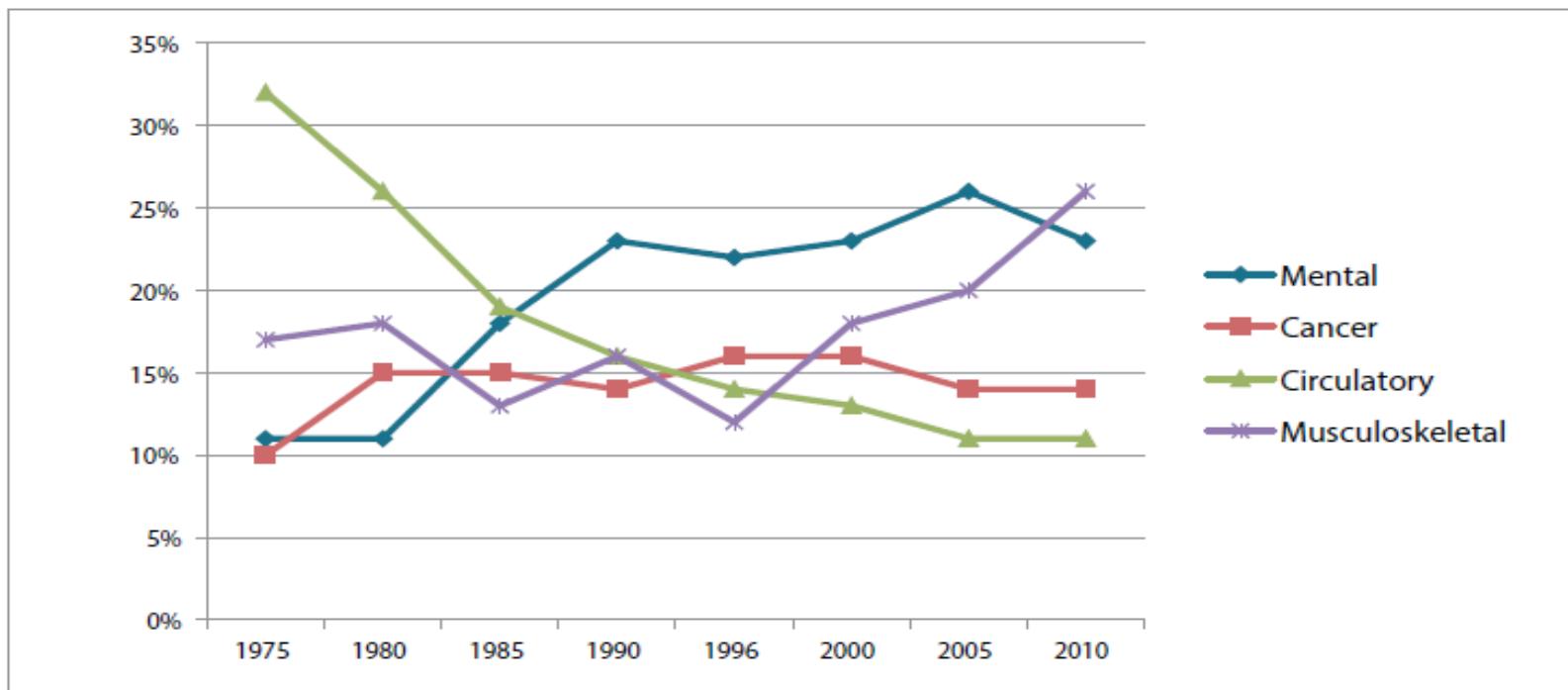
To label effectively, focus on clarity, readability and differentiation. The labels should stand out from the data.

## Education and Exports of Office Machines



*Figure 6A*  
**A Spaghetti Chart**

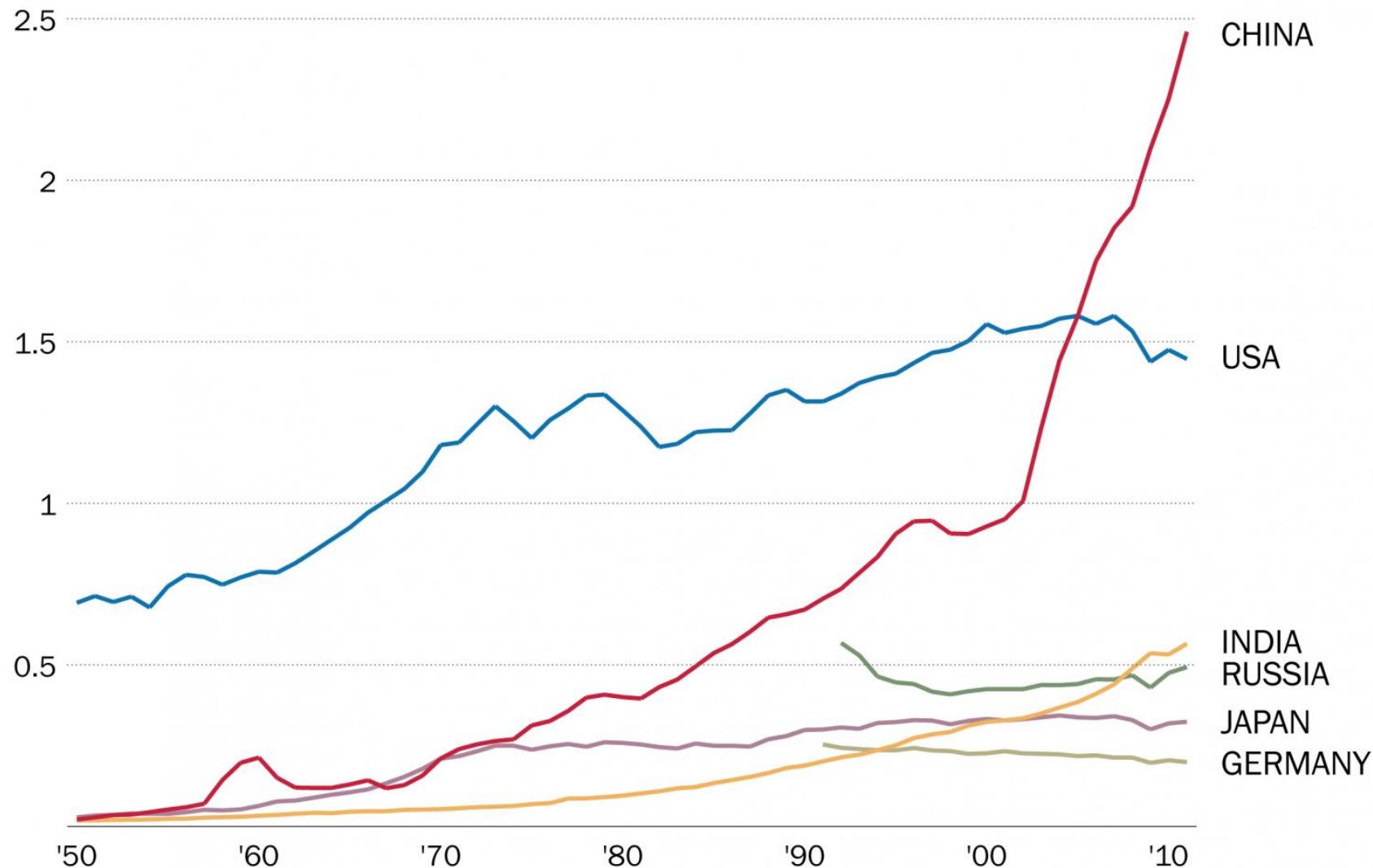
27. Initial DI Worker Awards by Major Cause of Disability—Calendar Years 1975-2010



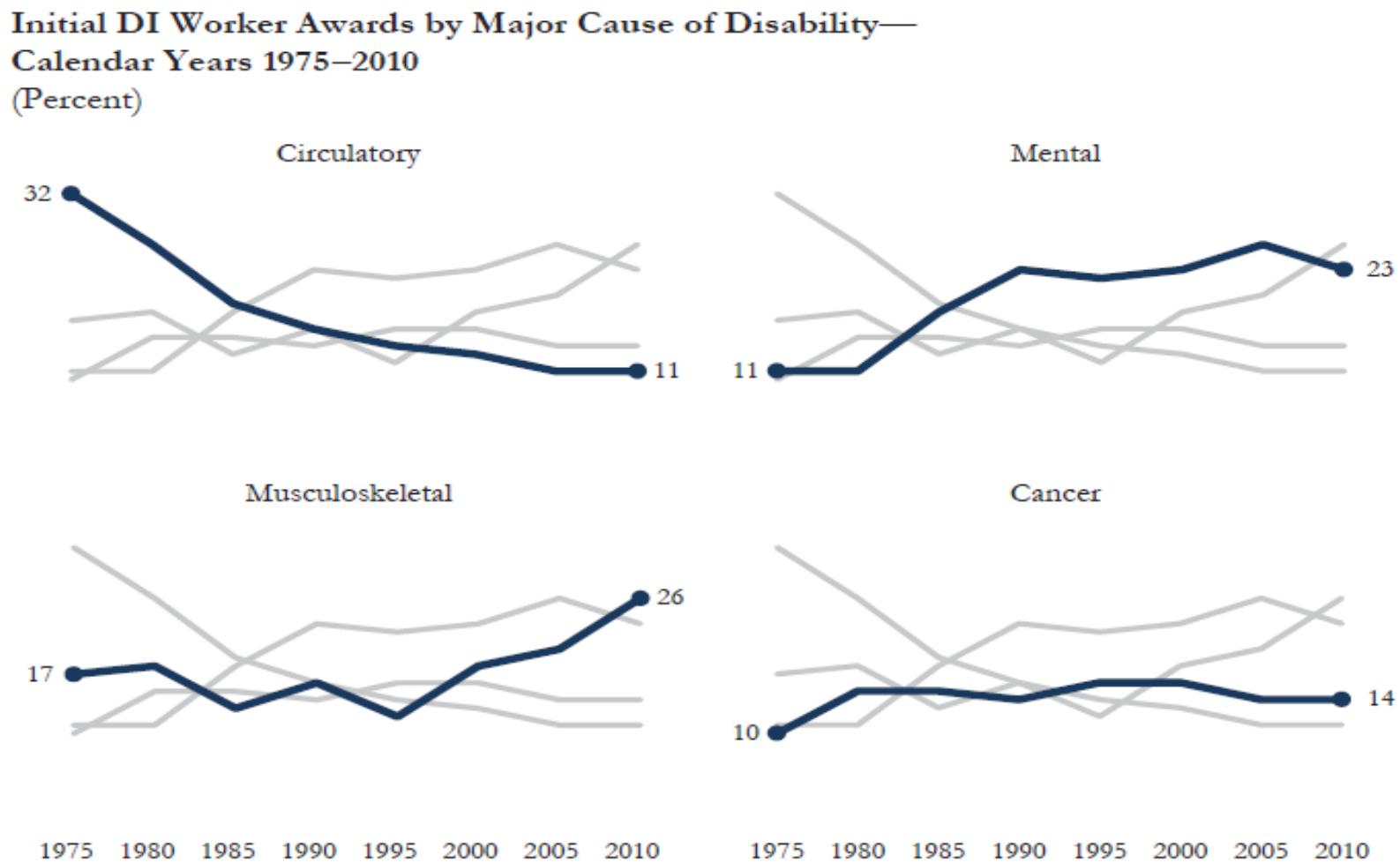
Source: Social Security Advisory Board (2012).

# Carbon dioxide emissions, by country

Data from the CDIAC. Billions of metric tons of CO<sub>2</sub>.

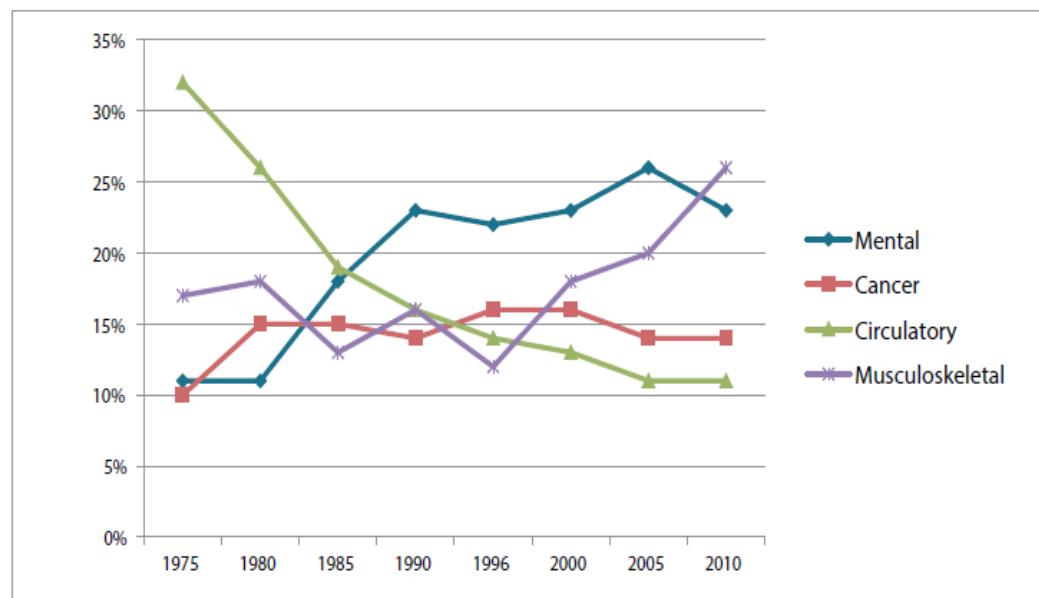


*Figure 6B*  
Revising the Spaghetti Chart



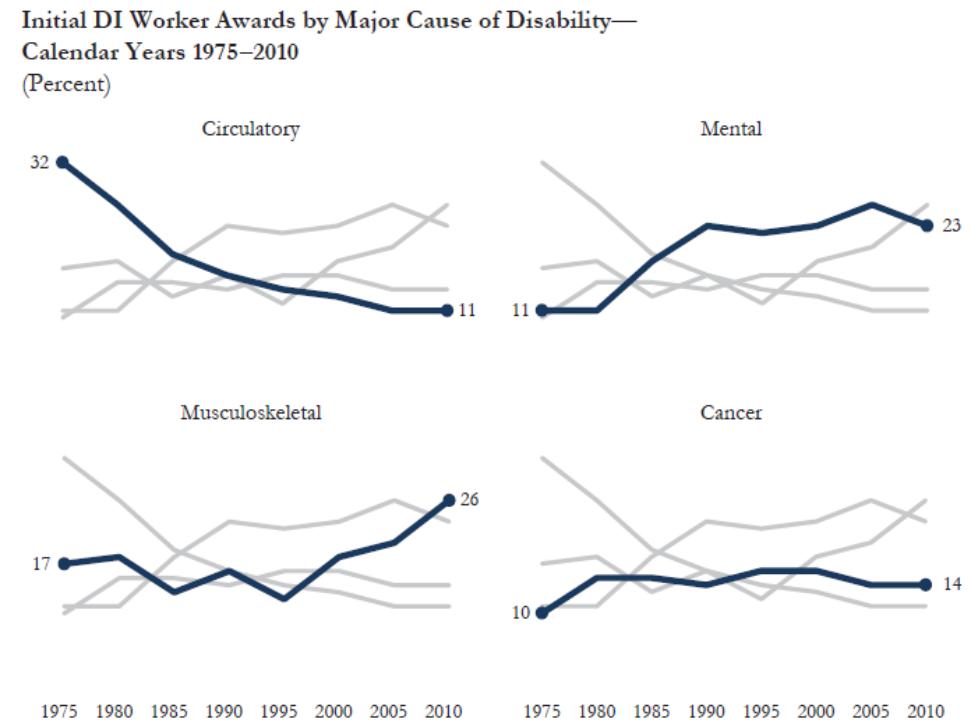
*Figure 6A*  
A Spaghetti Chart

27. Initial DI Worker Awards by Major Cause of Disability—Calendar Years 1975–2010



Source: Social Security Advisory Board (2012).

*Figure 6B*  
Revising the Spaghetti Chart

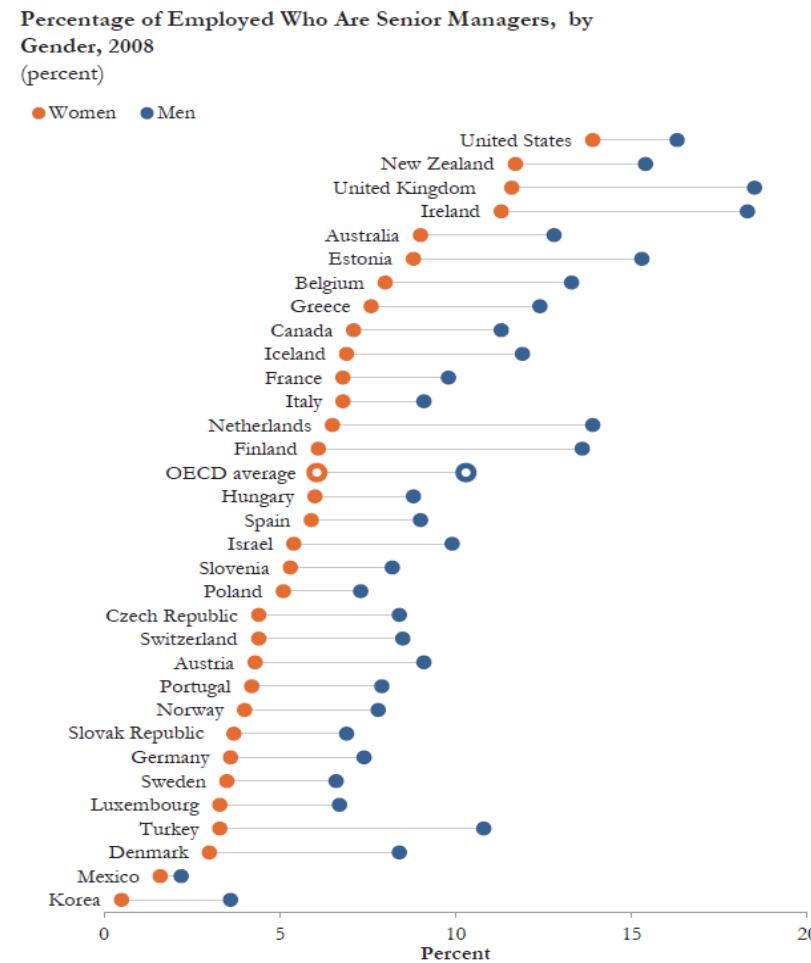


*Figure 5A*  
An Unbalanced Chart

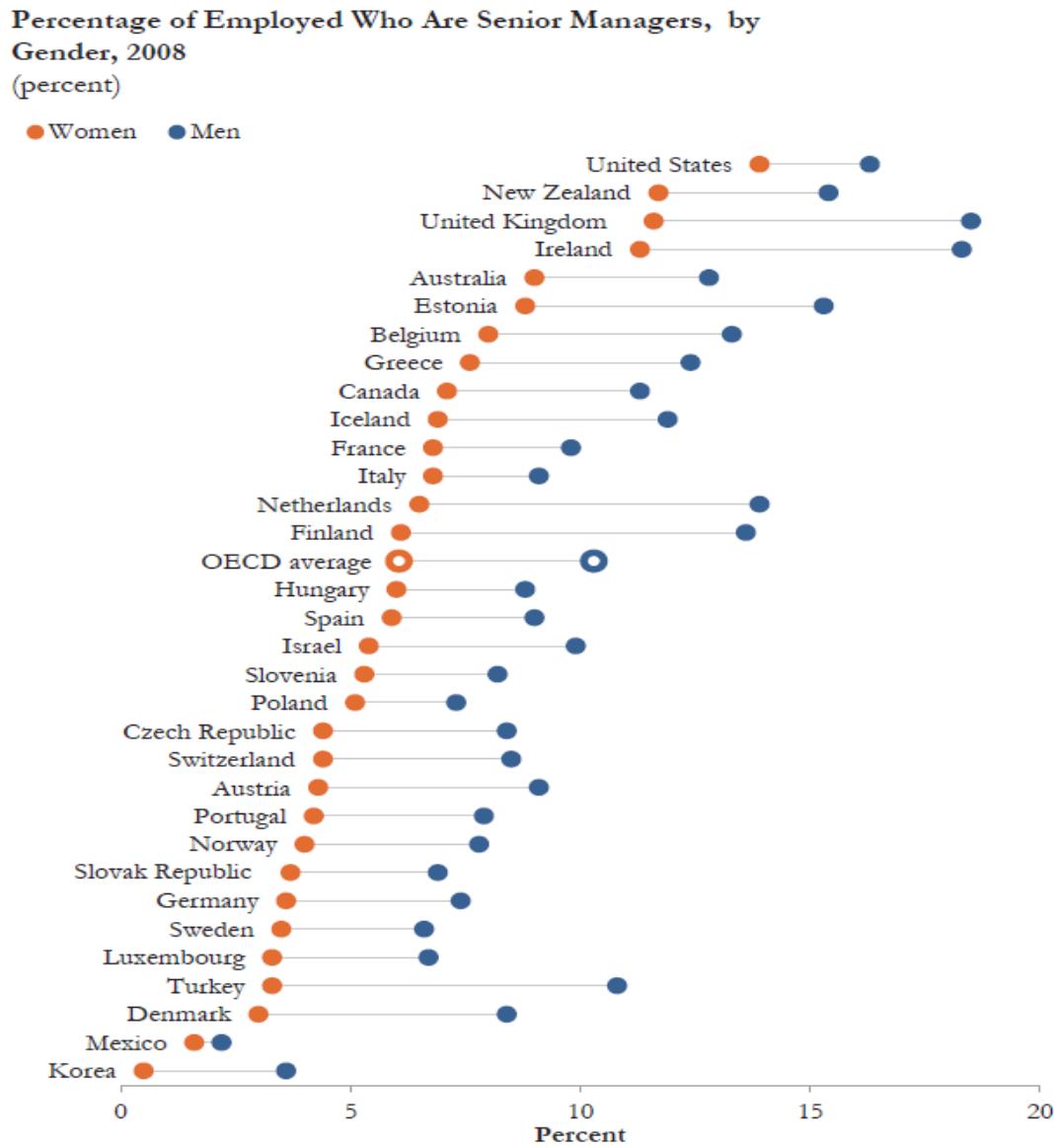


Source: Author, based on OECD (no date) and Rampell (2013).

*Figure 5B*  
Revising the Unbalanced Chart

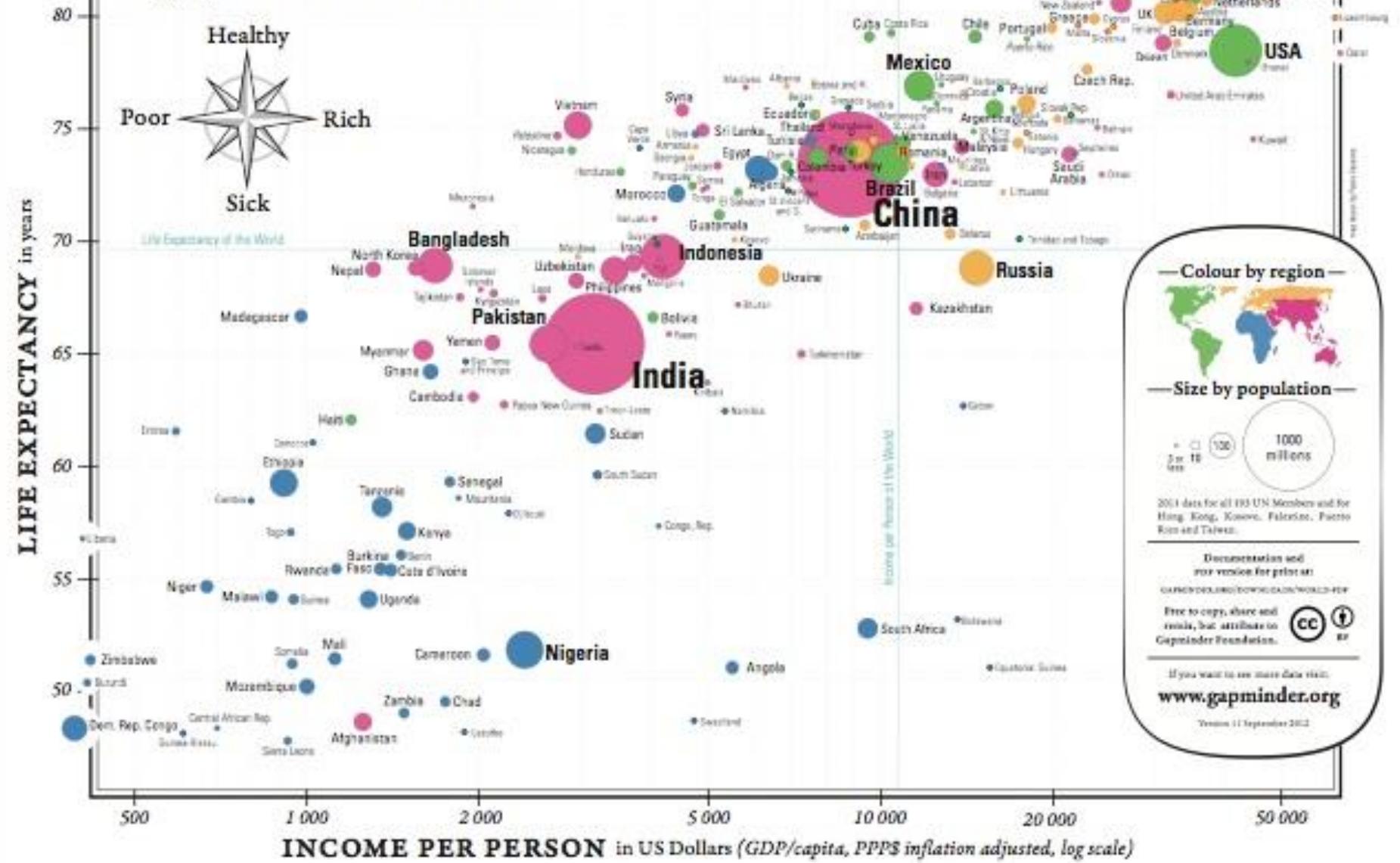


*Figure 5B*  
Revising the Unbalanced Chart



# GAPMINDER WORLD 2012

Mapping the Wealth and Health of Nations



Be Honest

Which of these maps tells the story better?

**Obama** ✓

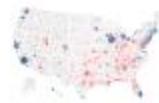
62,608,181 votes



States



Counties



Size of lead



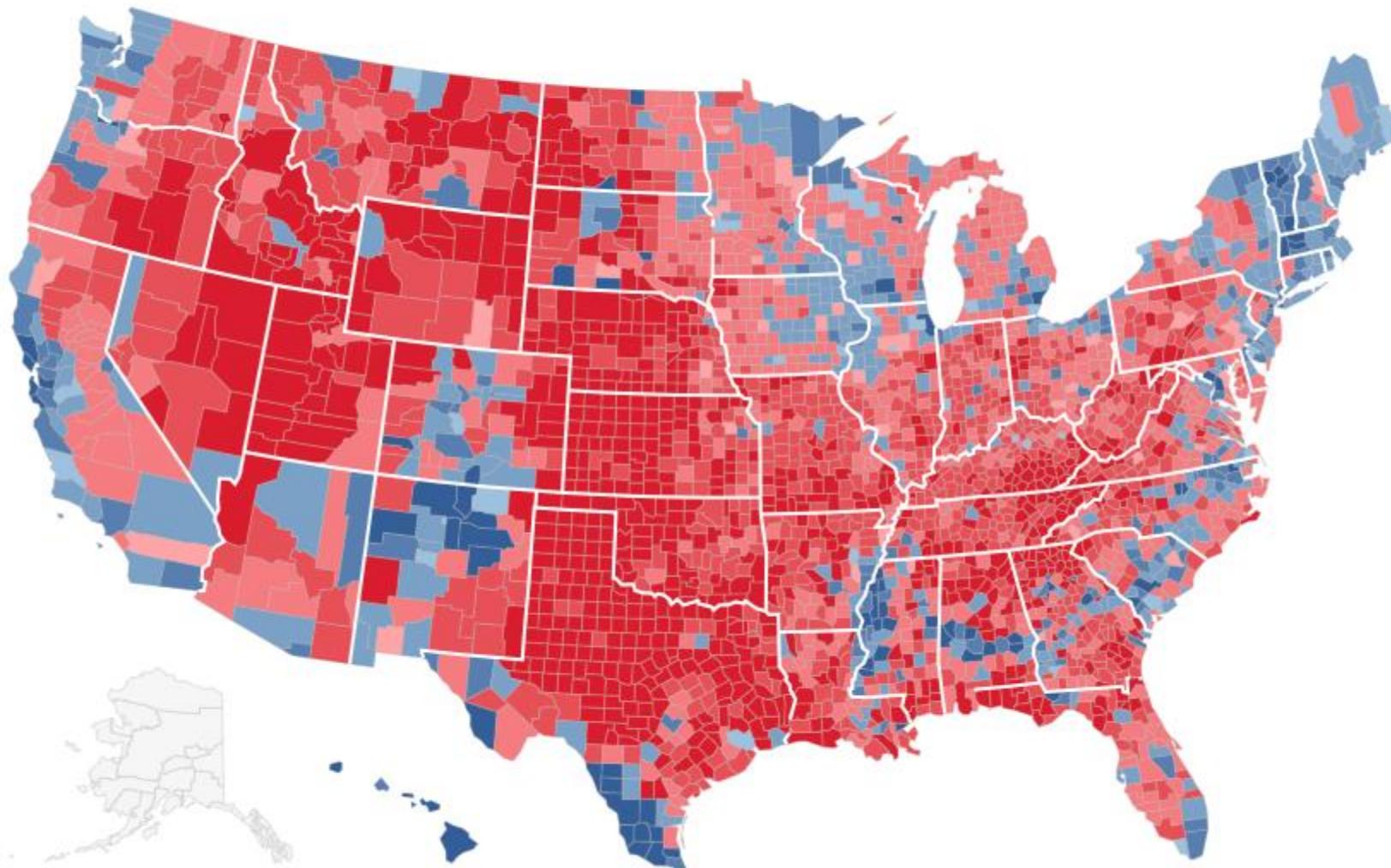
Shift from 2008

0  
undecided

270 to win

**Romney**

59,130,484 votes



TOP CANDIDATE'S  
SHARE OF VOTE

DEM. 40% 50% 60% 70%

REP. 40% 50% 60% 70%

IND. 40% 50% 60% 70%

PRESIDENT  
ELECTORAL VOTES  
**332✓ 206**

SENATE  
CANDIDATE  
**+2 -2**

HOUSE  
SEATS HIGH  
**195 ✓ 234**



STATES

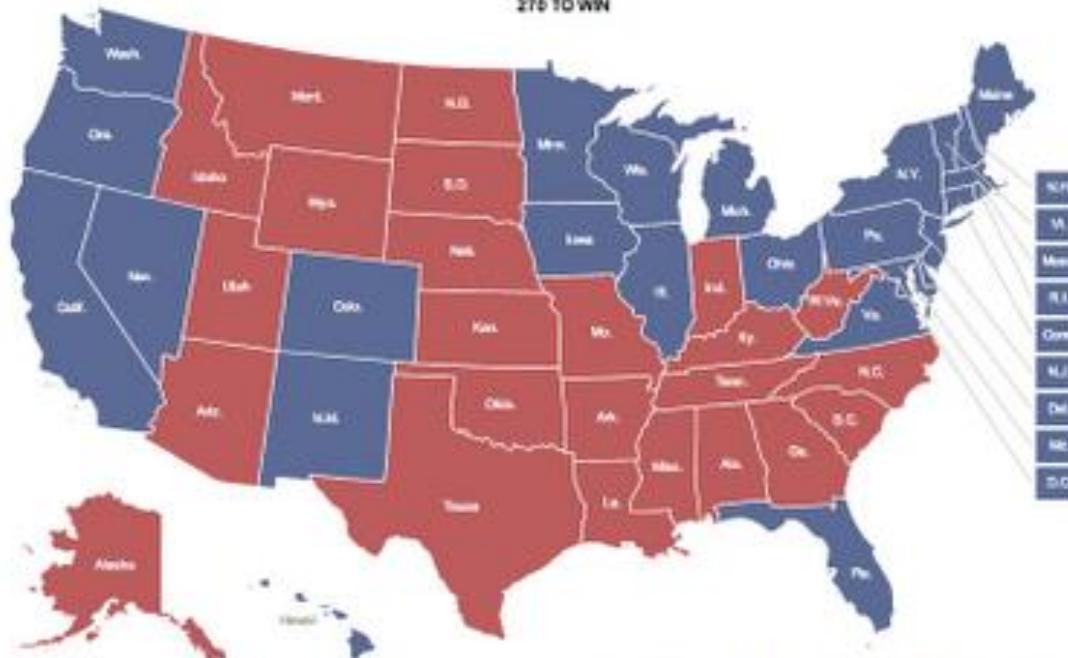


**332✓**  
ELECTORAL VOTES



**206**  
ELECTORAL VOTES

270 TO WIN



■ POLLS OPEN  
■ POLLS CLOSED

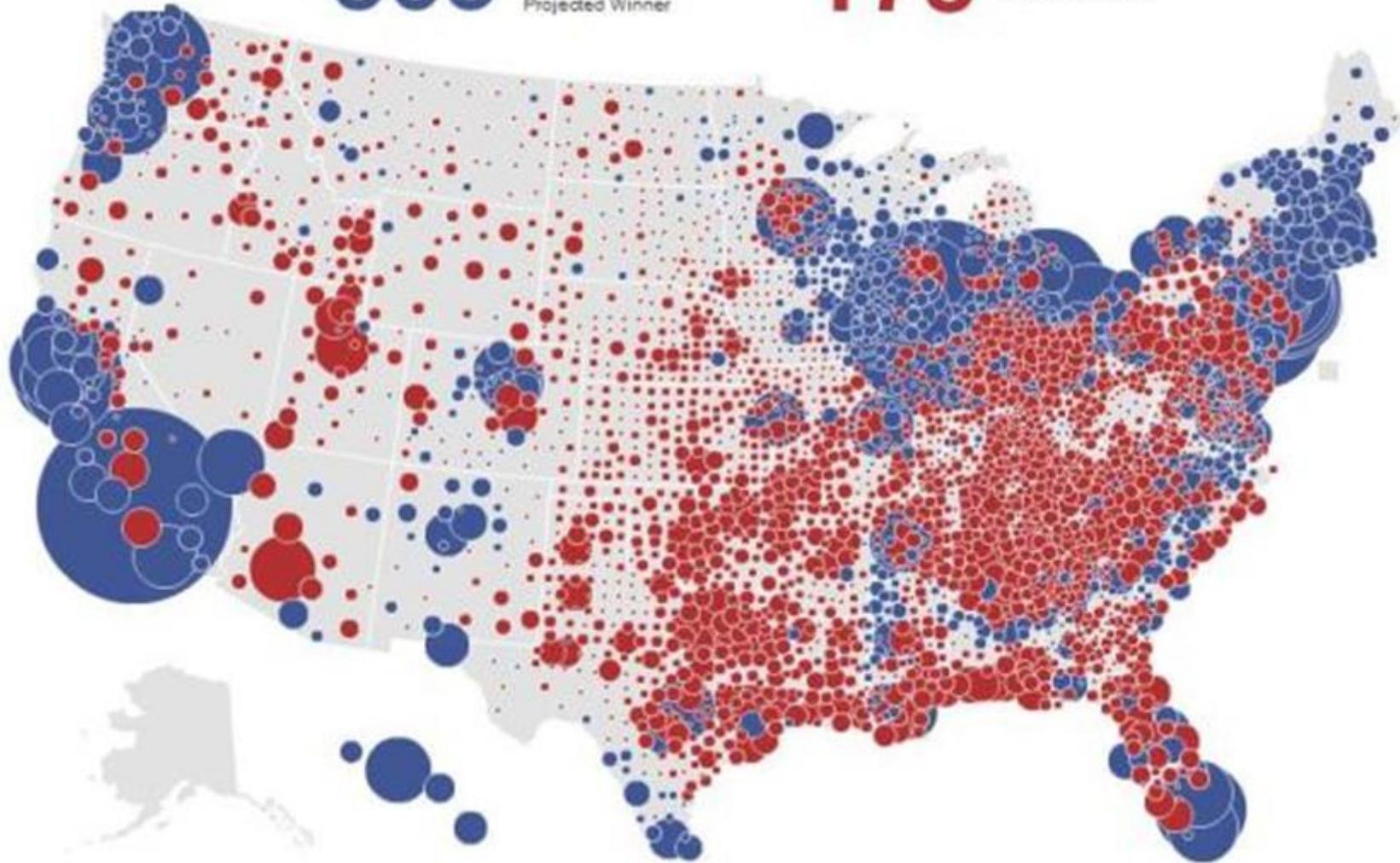
■ OBAMA LEADING  
■ OBAMA WON

■ ROMNEY LEADING  
■ ROMNEY WON

**365**

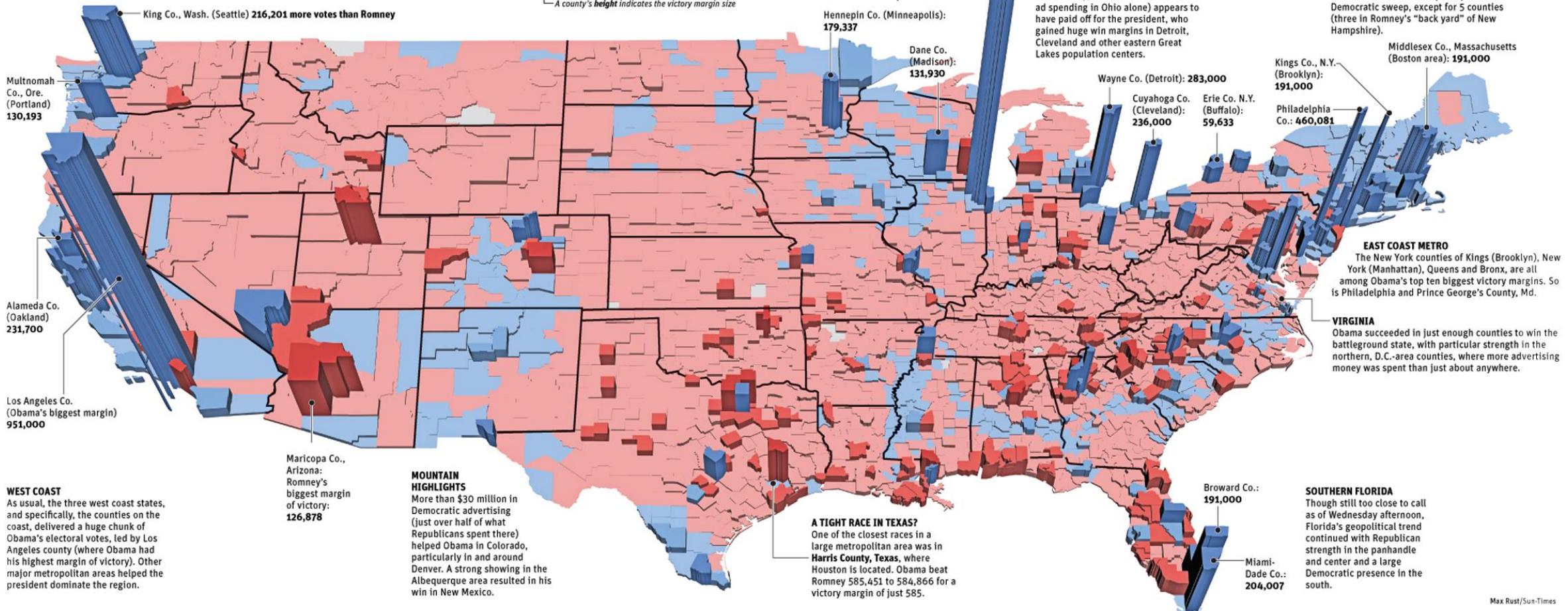
Obama  
Electoral Votes  
Projected Winner

**173** McCain  
Electoral Votes



# PLAYING THE MARGINS

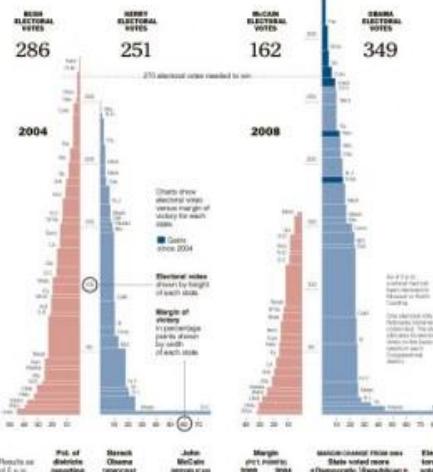
As in 2008 Barack Obama succeeded in his presidential quest by winning contests in the nation's densely populated metropolitan areas, while his opponent gained from suburban, exurban and rural parts of the country. When applied to the more than 3,000 counties, Tuesday's vote data reveals a political map of mostly flat, red Republican territory punctured by blue, Democratic spikes. Here's a look at how this geopolitical reality resulted in the president's re-election:



# In a Decisive Victory, Obama Reshapes the Electoral Map

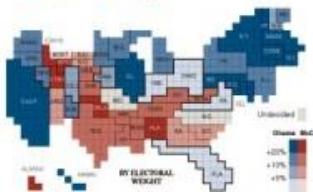
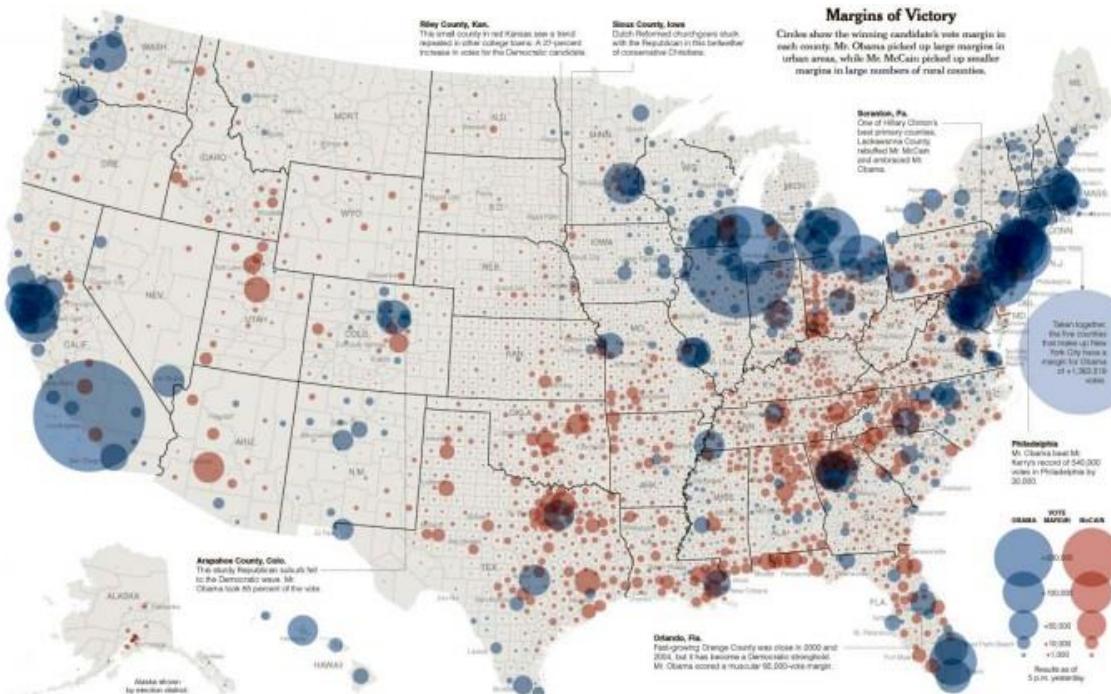
Barrack Obama's historic win, with at least 340 electoral votes to John McCain's 162, can be attributed to his victories in several high-population states, like Florida, Virginia and Ohio, that George W. Bush won handily in 2004. The struggling economy is mostly to blame.

By Erin Aigrain, Joe Ferguson, Basile Gouelard, Matthew Erickson, Hansch Fairfield, Ford Fassnender, Hayes Park and Archie Tan



States won by Obama	
Hawaii	100% 298,621 72%
Maine	100% 1,284,000 57%
Delaware	100% 247,995 67%
New Mexico	100% 452,038 57%
Vermont	97% 201,069 67%
Arkansas	100% 2,000,000 57%
Illinois	87% 8,056,943 67%
California	87% 6,735,543 61%
Michigan	100% 2,907,053 67%
Virginia	89% 1,730,938 52%
Minnesota	100% 2,073,004 57%
Colorado	87% 1,705,001 57%
Connecticut	89% 943,319 67%
Wyoming	89% 1,430,150 67%
North Dakota	100% 671,000 57%
Washington	89% 980,008 67%
Oregon	89% 866,466 67%
New Hampshire	89% 853,301 55%
Iowa	92% 3,021,070 57%
New Jersey	100% 2,073,004 57%
Pennsylvania	100% 3,785,007 57%
Rhode Island	97% 273,031 67%
Tennessee	89% 4,200,000 57%
Mississippi	89% 2,057,360 67%
Minnesota	100% 1,873,043 54%
D.C.	100% 210,403 99%
Ohio	98% 2,967,489 67%
Massachusetts	100% 1,805,181 62%

States won by McCain	
Penn-Del	100% 141,113 47%
Idaho	100% 315,913 47%
Montana	100% 120,000 47%
Utah	100% 261,771 47%
Rocky-Mtn	100% 173,977 47%
Alaska	100% 256,730 97%
Georgia	89% 1,211,194 47%
Texas	89% 3,023,404 62%
Kansas	100% 449,963 47%
South Dakota	100% 942,441 47%
Wyoming	100% 90,496 37%
Alabama	100% 113,000 47%
Ak-Sas	100% 811,935 37%
Kentucky	100% 746,513 47%
Ariz	99% 261,269 47%
Alaska	89% 80,345 36%
Arkansas	100% 1,081,074 47%
Louisiana	100% 192,985 47%
Ala-Cala	89% 417,314 38%
Total	94% 63,000,432 57%
No winner called	27
Missouri	100% 1,430,740 49%
North-Carolina	100% 2,116,954 57%
	< 1%



## The Vote by County

Mr. Obama's campaign theme of "change" was rewarded with support in areas of the country hurt by the weakened economy. Rising unemployment and housing foreclosures in Rust Belt states, as well as Florida and Nevada, may have led voters to support him.

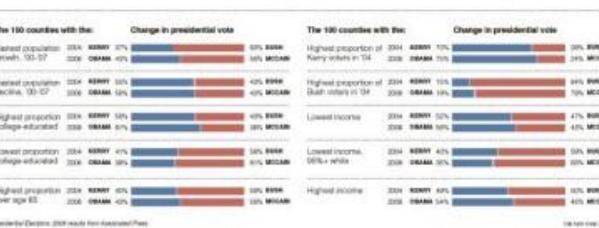
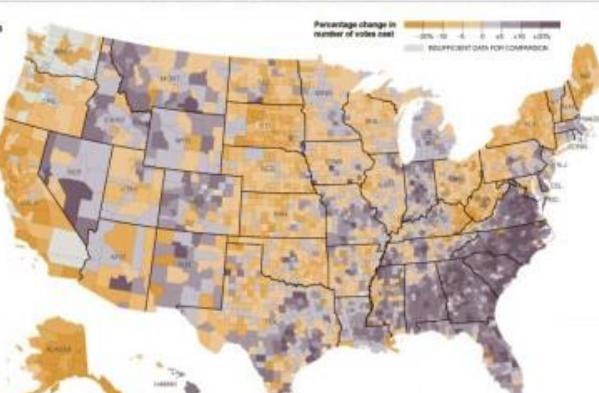
A powerful get-out-the-vote effort paid off for the Obama campaign in certain crucial states, like Florida, Colorado and Nevada. The number of people who voted in Florida rose by 9.7 percent from 2004. Many of those voters went to the polls for the first time — and those voters chose Mr. Obama nearly 70 percent of the time.

Voter figures were also high in states won by Mr. McCain, like South Carolina, Georgia and Alaska.

Turnout may have been a defining factor in Indiana, a battleground state that had 5.3 percent more voters than in 2004. It awarded a narrow victory, and its 11 electoral votes, to Mr. Obama.

## Shifting Demographics

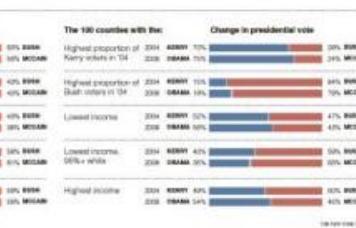
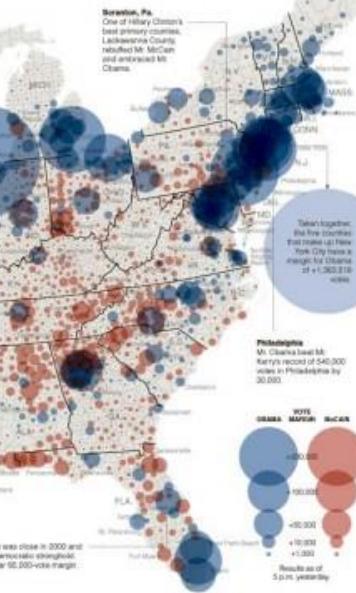
The electorate moved toward the Democratic Party in almost every demographic category. Many shifts were a few percentage points, but several categories had much higher jumps. One exception to the trend: low-income white couples moved solidly toward the Republican side.



Source: Preliminary results from Data USA at [tiny.cc/meyarw](http://tiny.cc/meyarw). Presumed election 2008 results from Associated Press.

## Margins of Victory

Circles share the winning candidate's vote margin in each county. Mr. Obama picked up large margins in urban areas, while Mr. McCain picked up smaller margins in large numbers of rural counties.



Source: Preliminary results from Data USA at [tiny.cc/meyarw](http://tiny.cc/meyarw). Presumed election 2008 results from Associated Press.

OBAMA



**269**

ELECTORAL VOTES

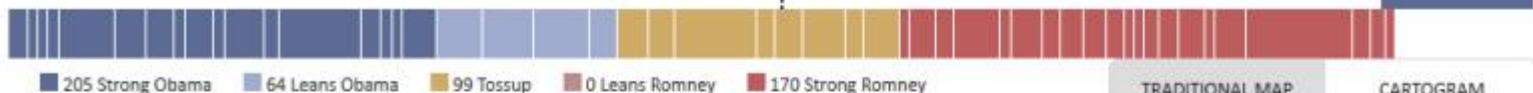
ROMNEY

**170**

ELECTORAL VOTES

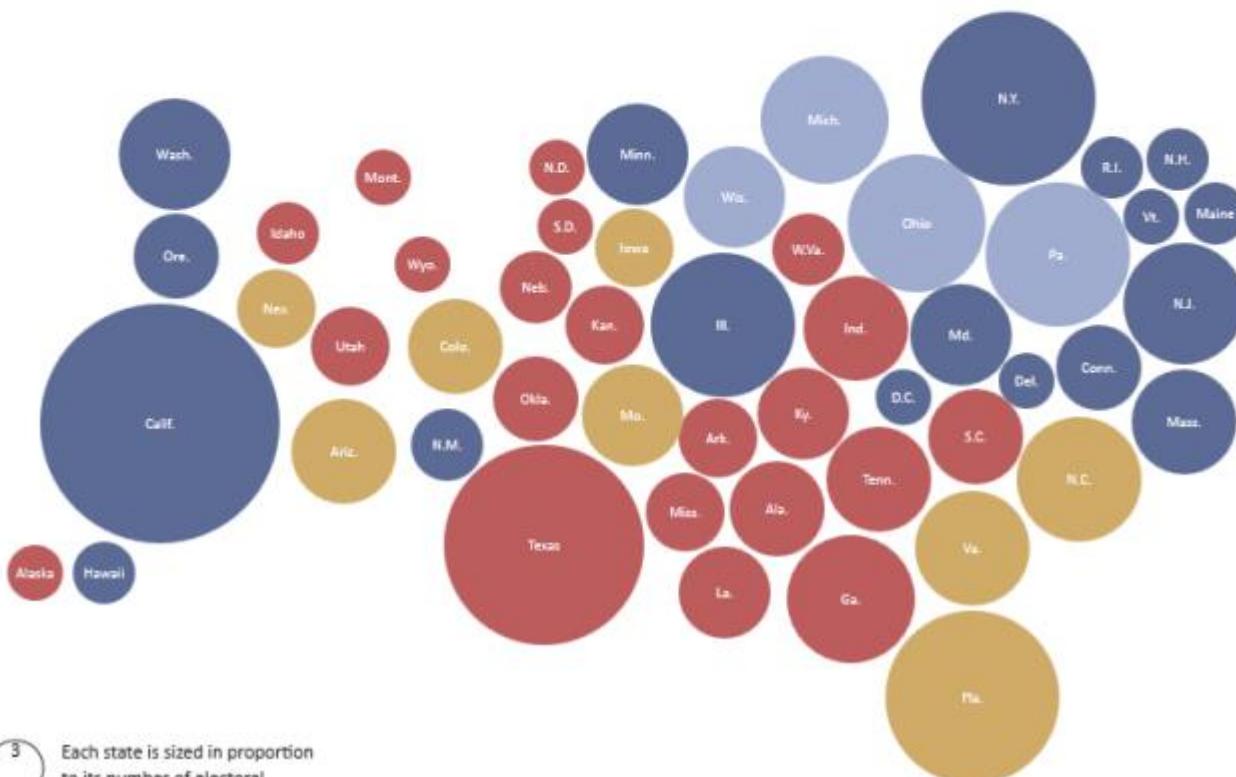


270 ELECTORAL VOTES NEEDED TO WIN



TRADITIONAL MAP

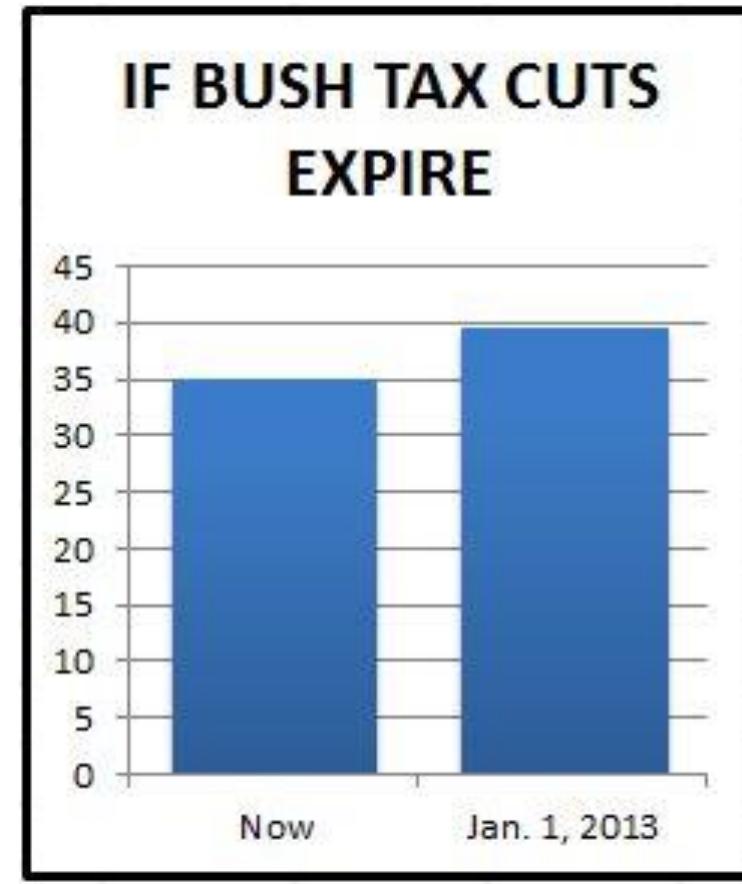
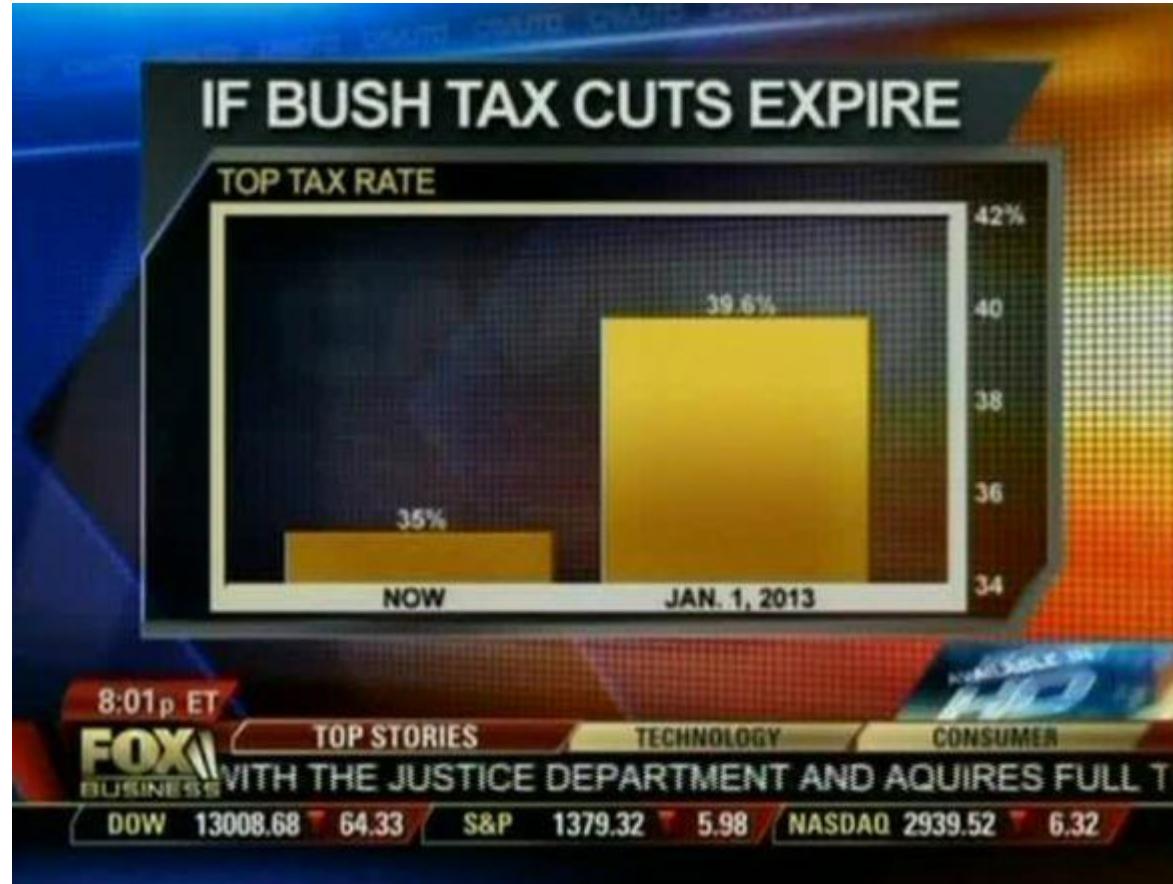
CARTOGRAM



3

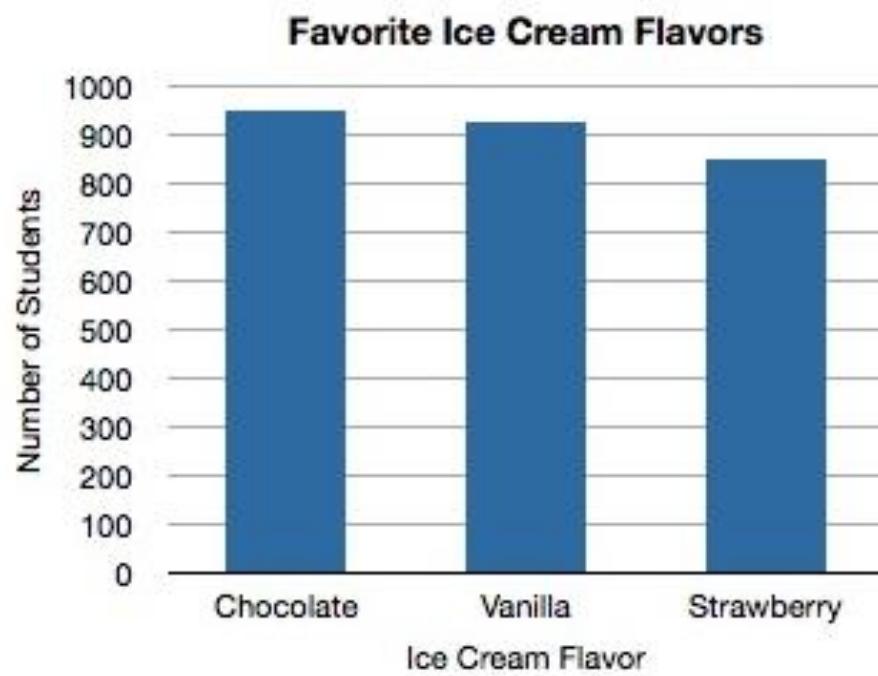
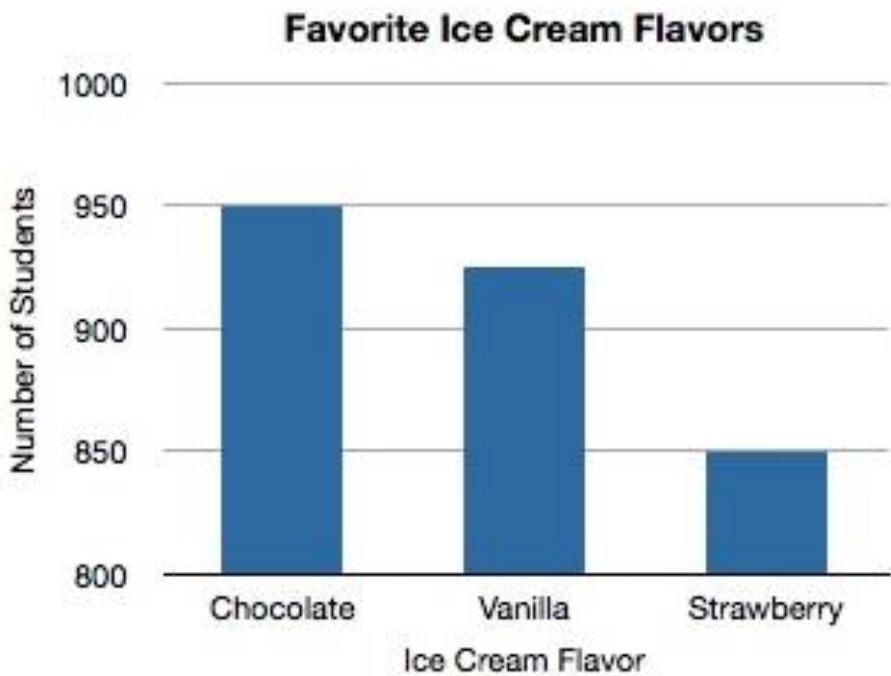
Each state is sized in proportion  
to its number of electoral  
votes.

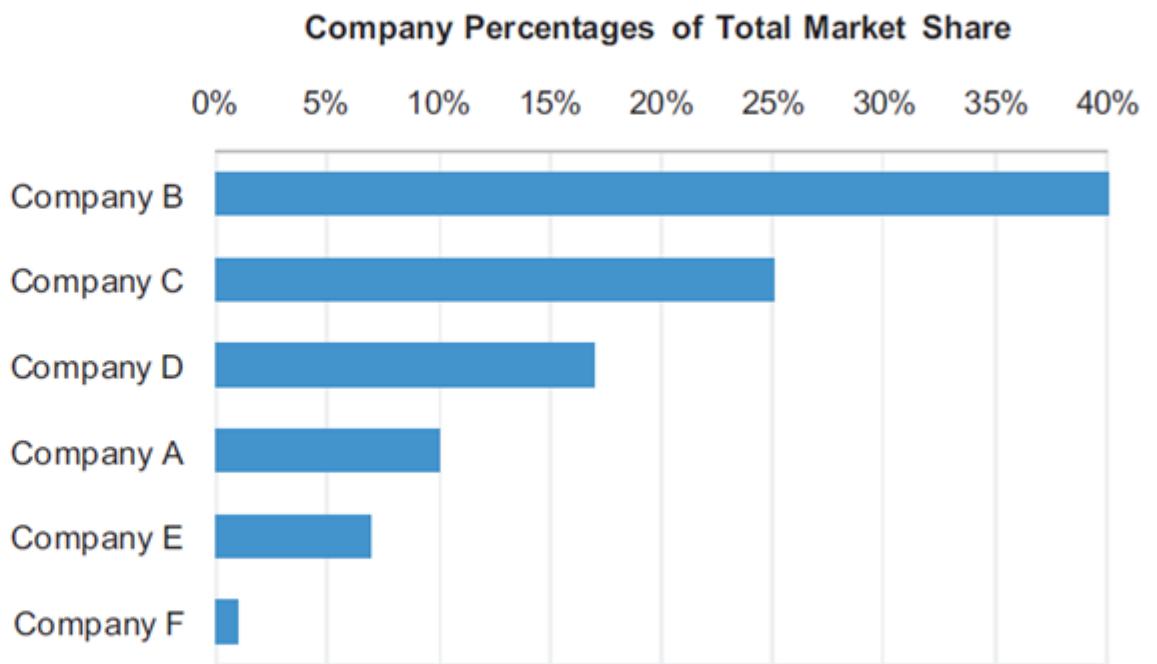
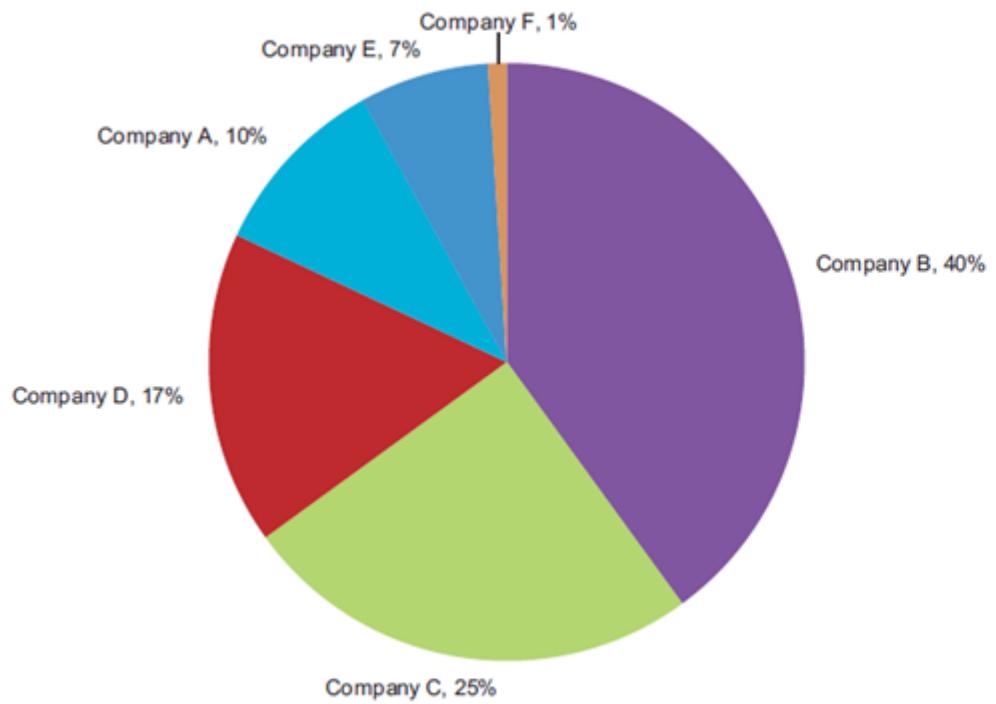
By emphasizing integrity of the wrong element (geography versus population density here), you may mislead your audience.



<http://www.vox.com/2015/11/19/9758062/y-axis-zero-chart>

Honest Visualization = Intention + Execution

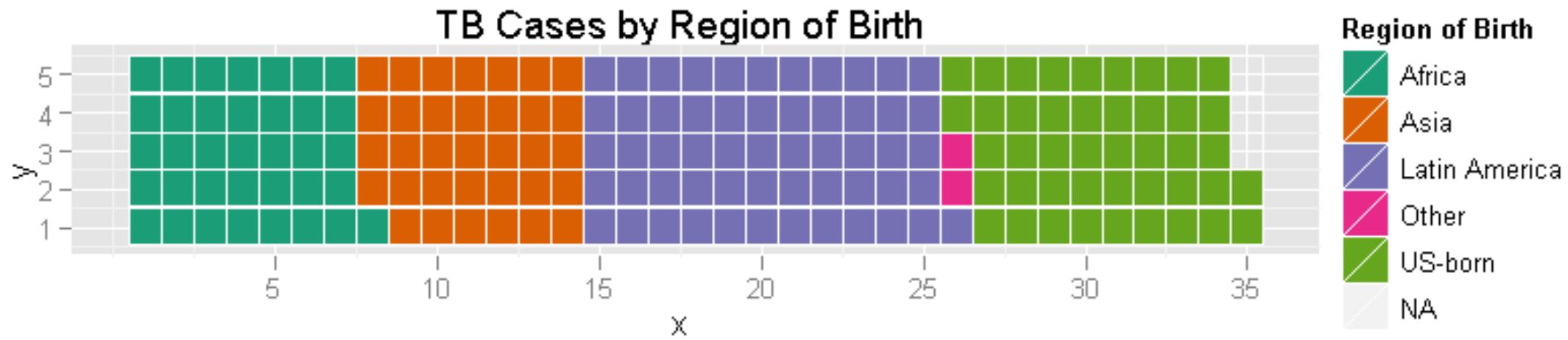




Never use pie charts

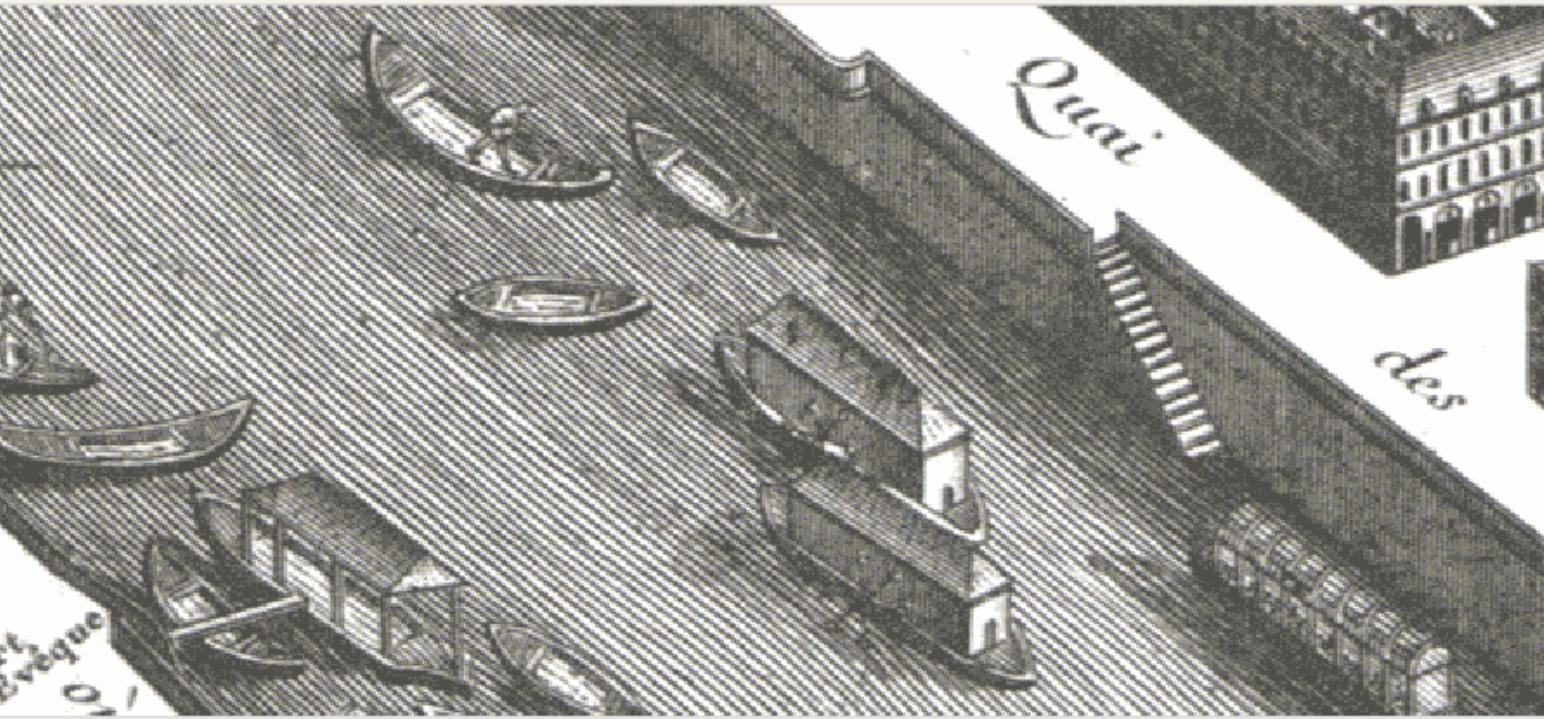
Unless it's this pie chart





# Drilling Down

## 6: Utilize Micro/Macro



Details are all in the micro-readings.

[REPLAY](#)

---

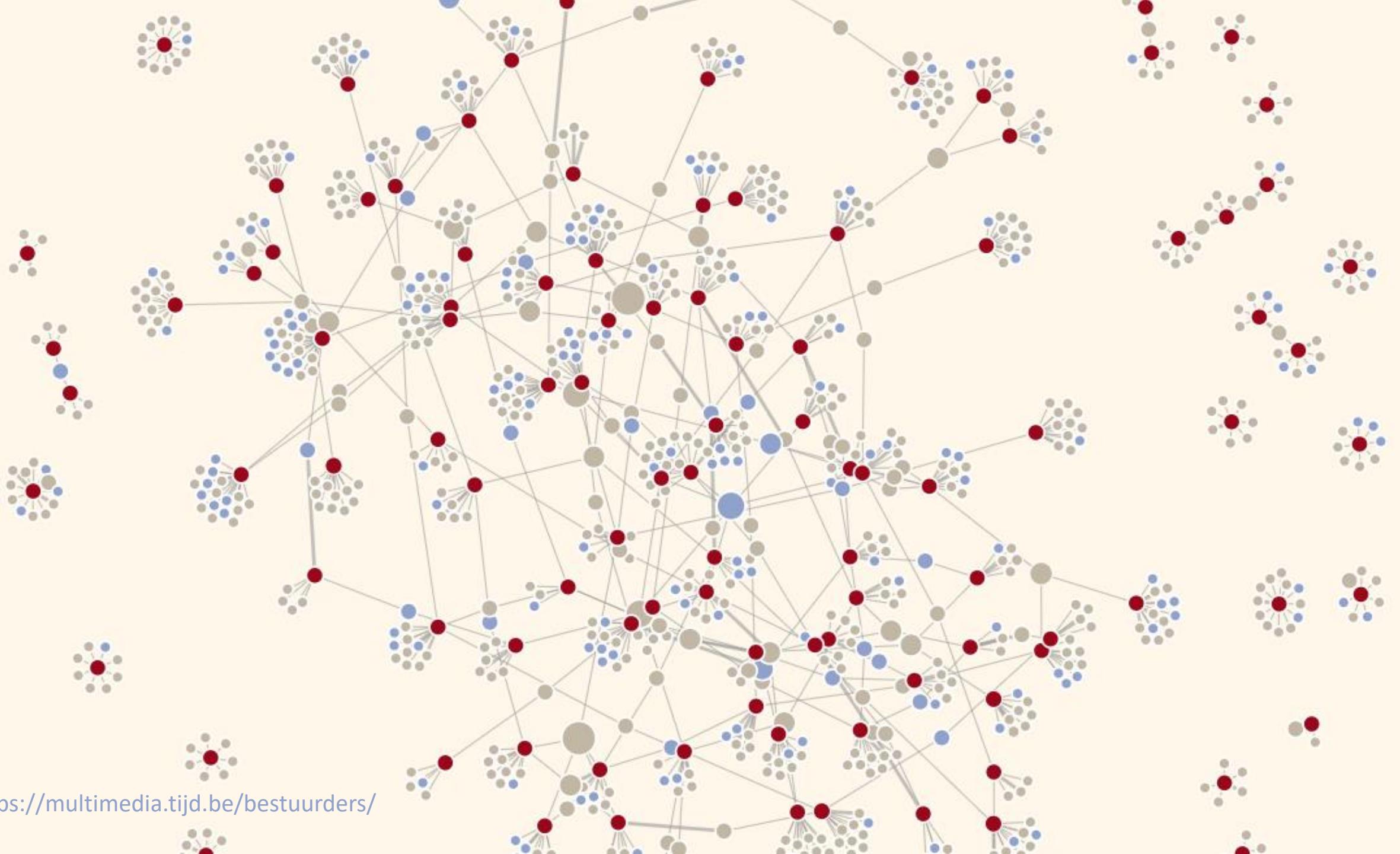
IN THE example above, the drawing of Paris is completely packed with detail.

When viewed from afar, the thousands of tiny windows and doors mesh together to form the coherent structure of buildings and neighborhoods.

Therefore, the clarity of the macro is determined by the quality and quantity of the micro.

“To clarify, add detail.”

- Edward Tufte



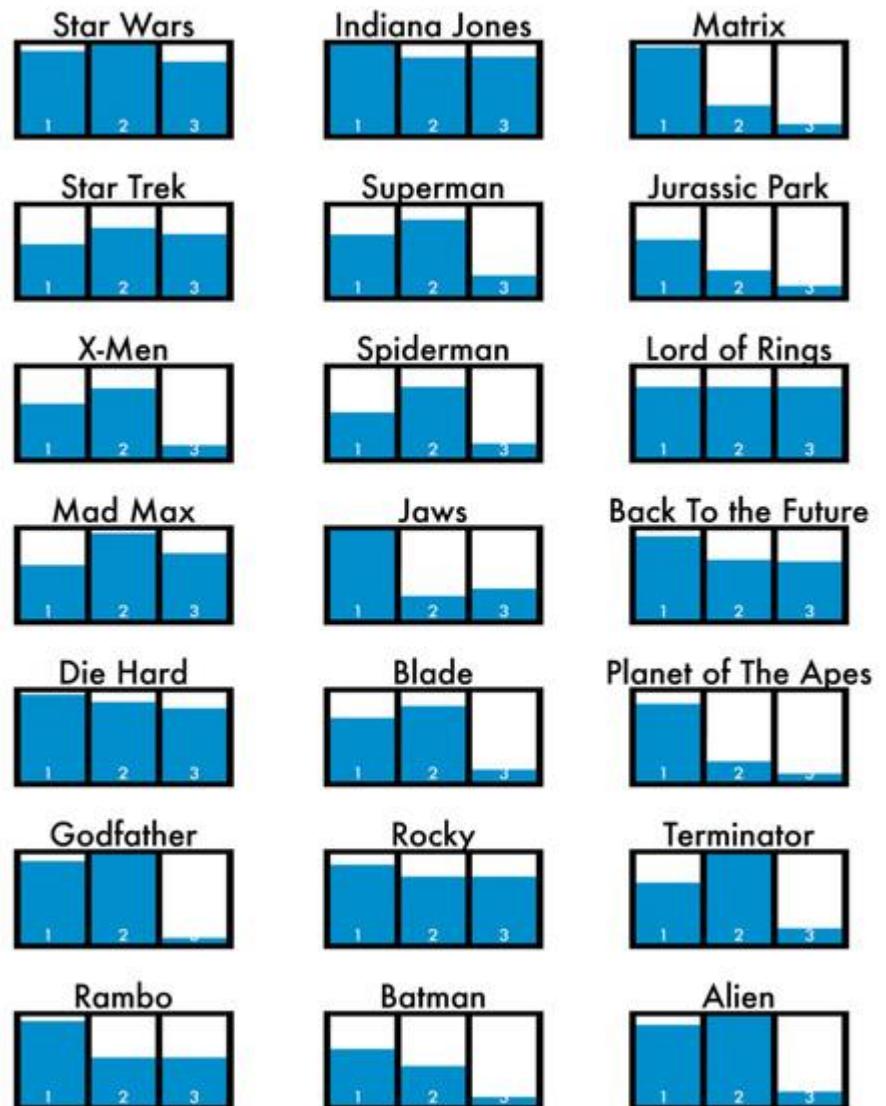
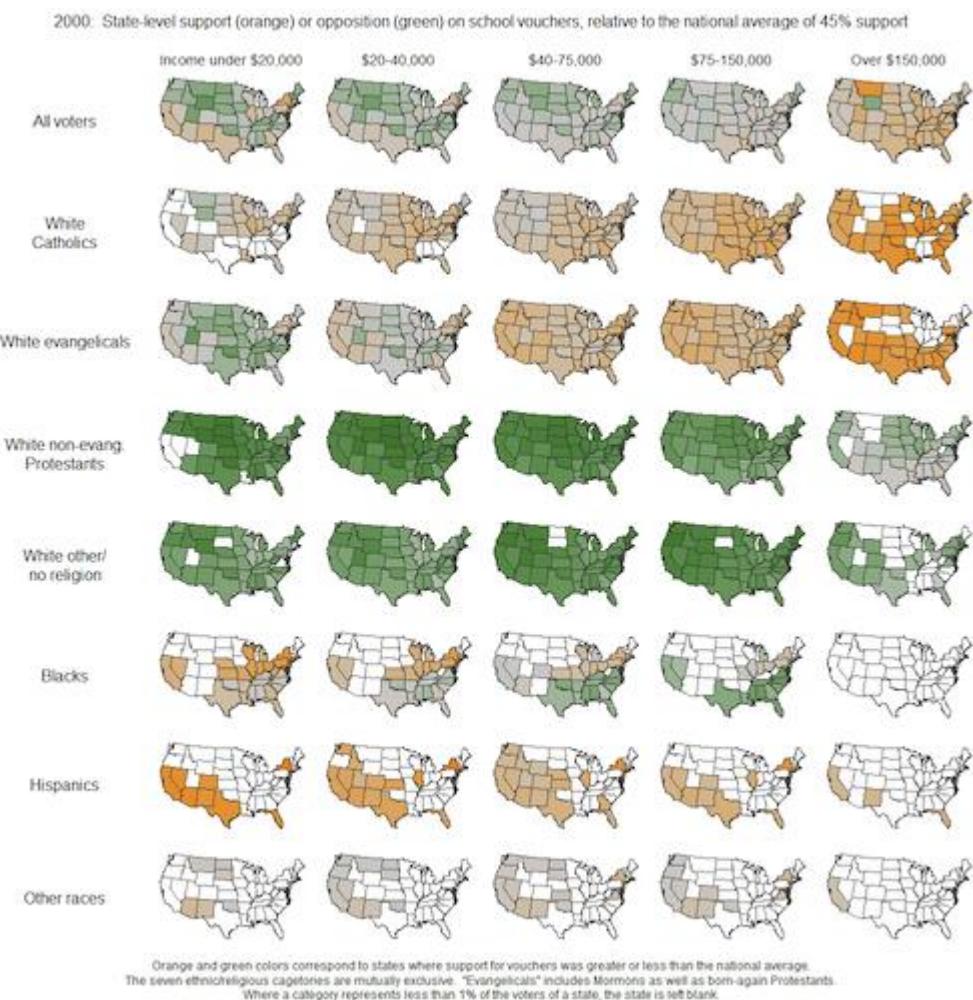
<https://multimedia.tijd.be/bestuurders/>

[http://www.sealthreinhold.com/school/tuftes-rules/rule\\_one.php](http://www.sealthreinhold.com/school/tuftes-rules/rule_one.php)

Rules 5-6

# Small Multiples

# THE TRILOGY METER



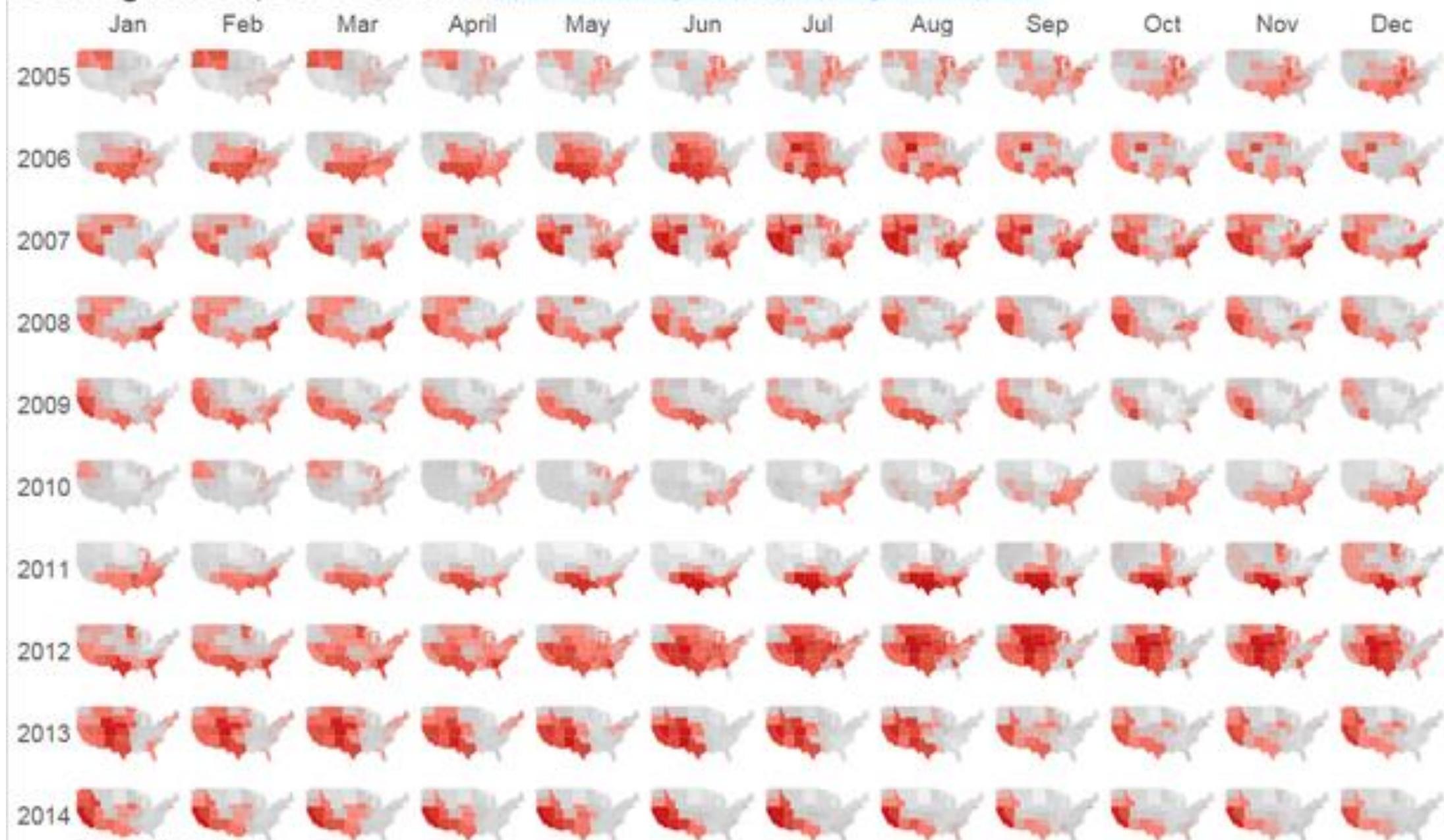
#1 In A Series of Pop Cultural Charts

DANMETH.COM

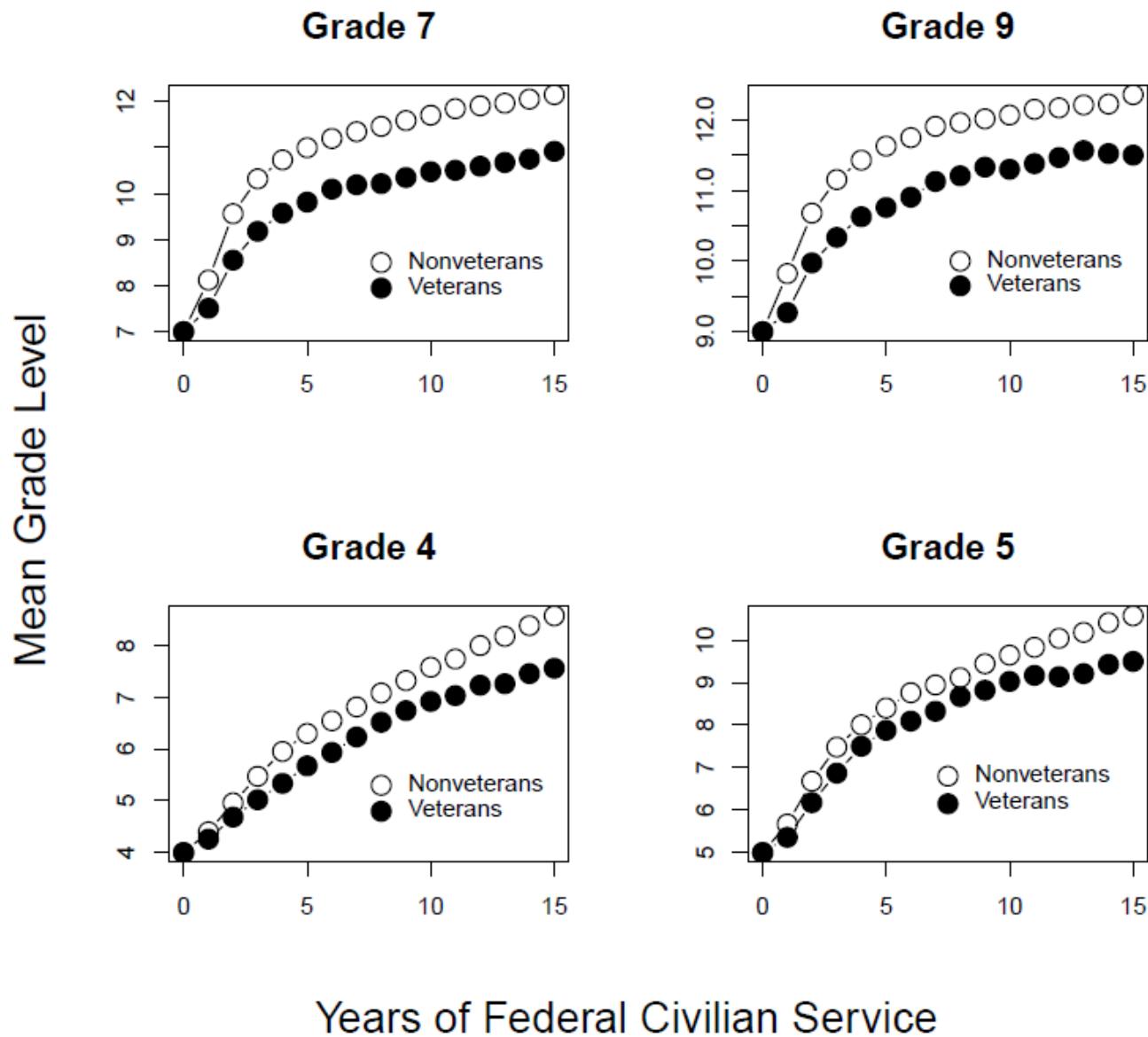
# US Drought Index, 2005-2014

source: <http://www.ncdc.noaa.gov/temp-and-precip/drought/historical-palmers/>

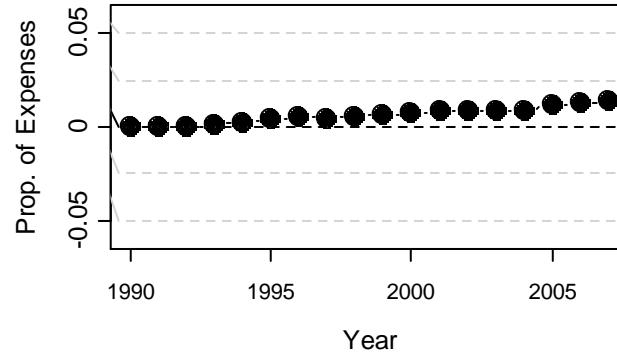
-6.00  8.06



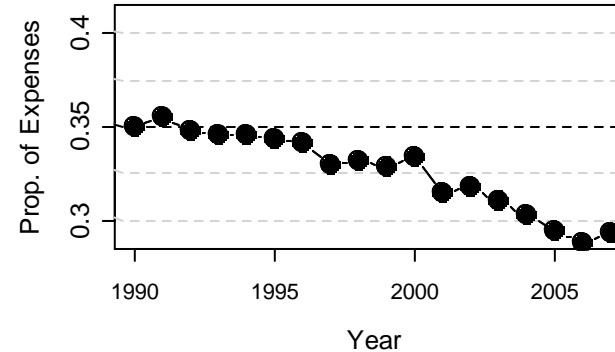
## Figure 2. Career Progression by Entry Grade



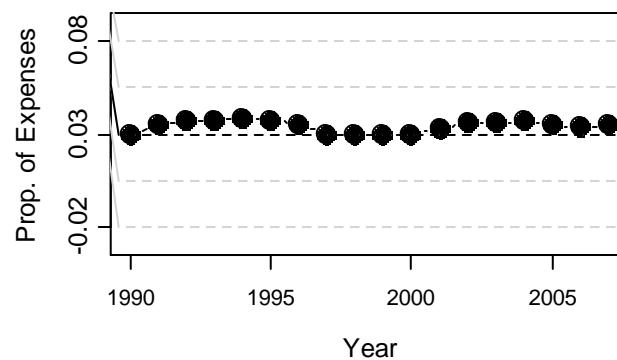
### Officer Wages



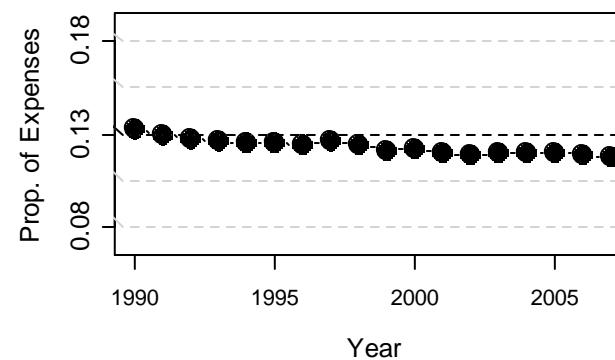
### Staff Wages



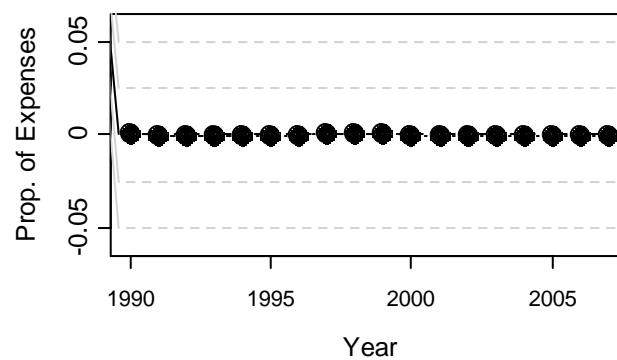
### Benefits



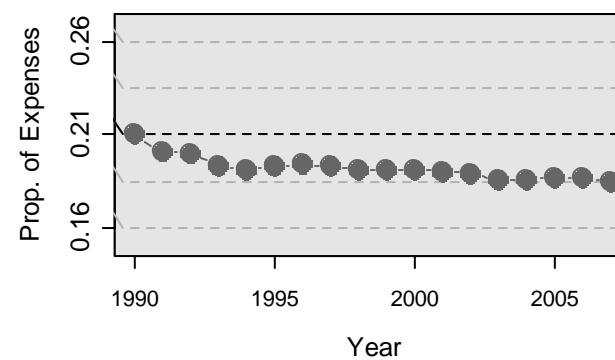
### Professional Fees



### Operations



### Overhead Ratio

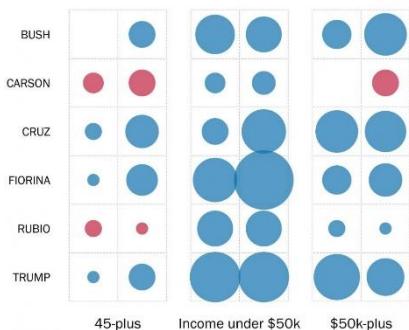
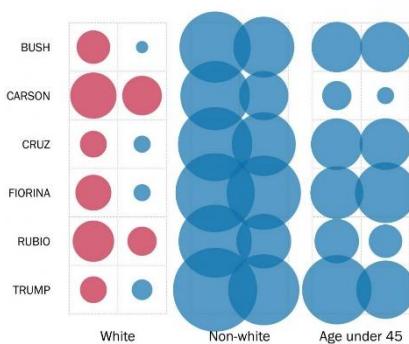
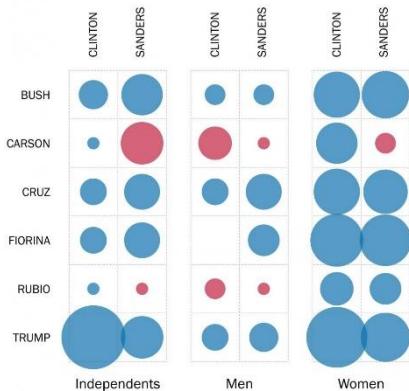


### Head-to-head match-ups, by demographic

Data from Marist/McClatchy poll of November, 2015.

KEY:  
DEMOCRAT LEADS BY 5

REPUBLICAN LEADS BY 25



## Head-to-head match-ups, by demographic

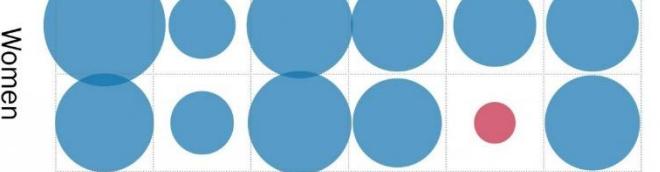
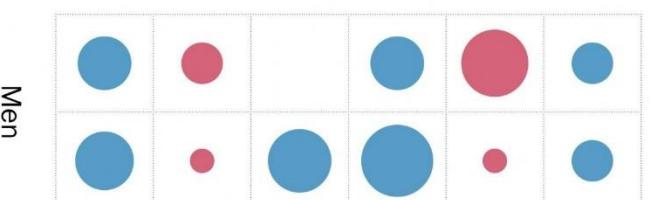
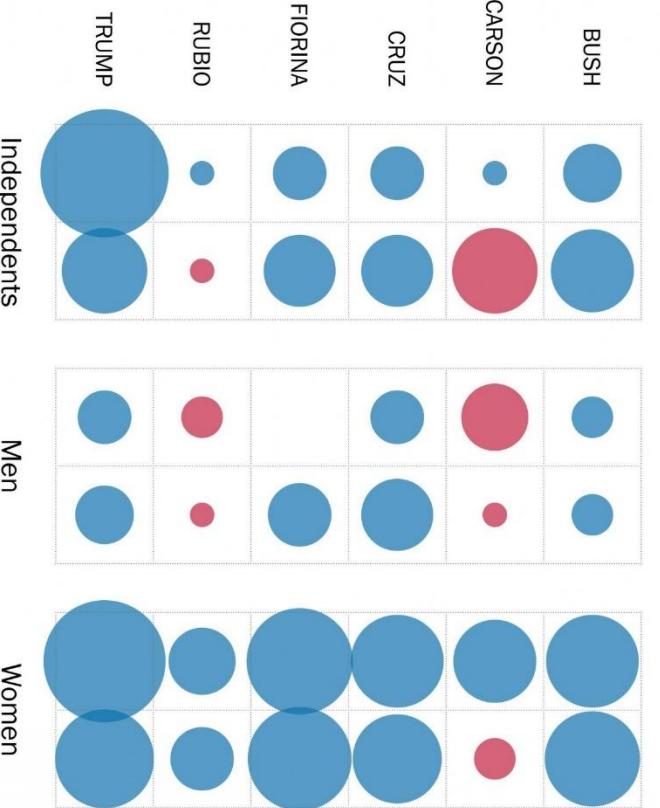
Data from Marist/McClatchy poll of November, 2015.

KEY:

 DEMOCRAT LEADS BY 5



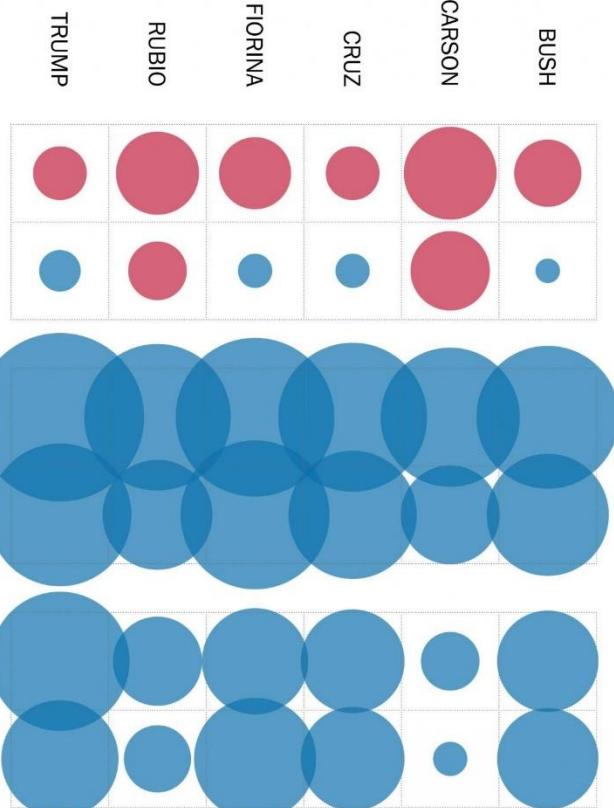
REPUBLICAN LEADS BY 25



Independents

Men

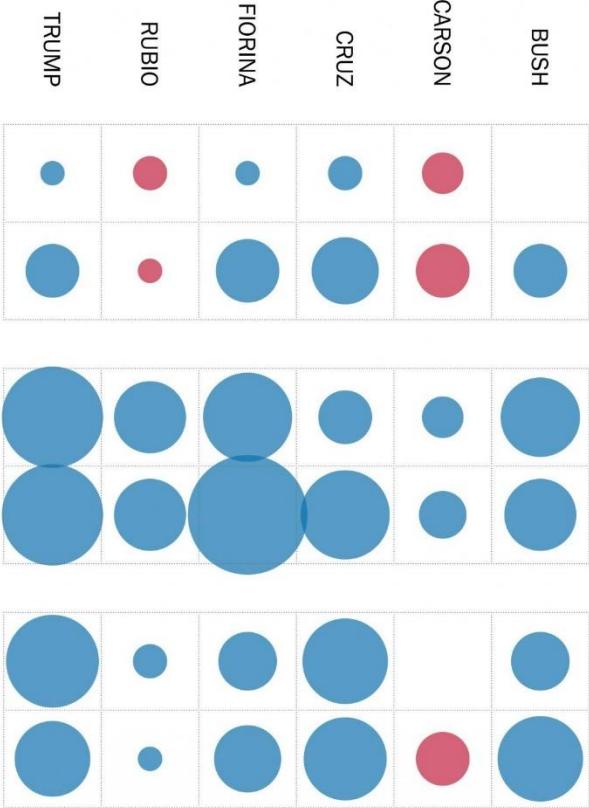
Women



White

Non-white

Age under 45

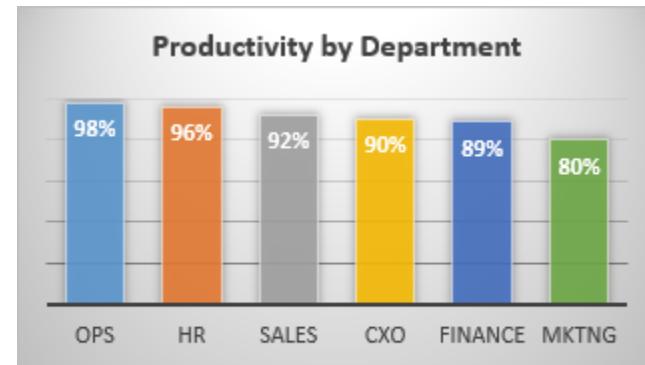
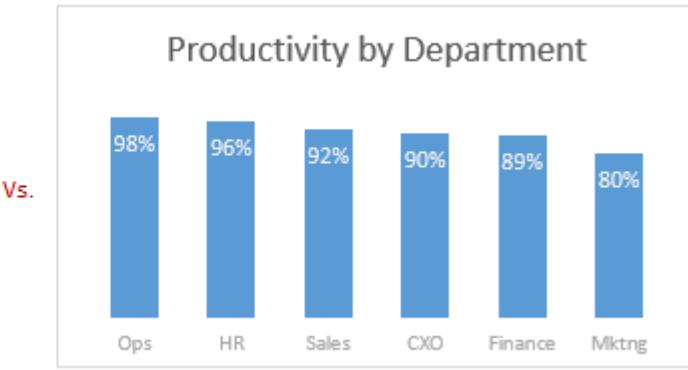
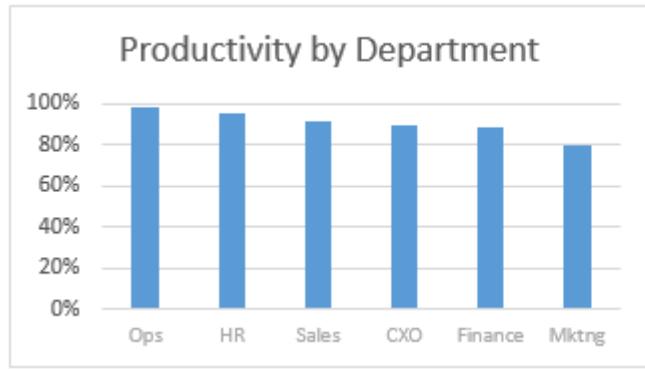


45-plus

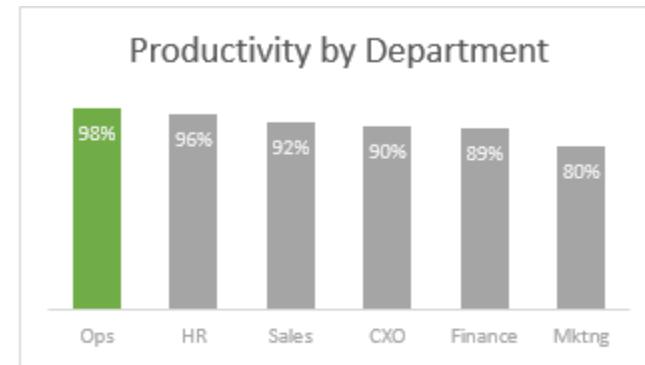
Income under \$50k

\$50k-plus

# Remove Chart Junk

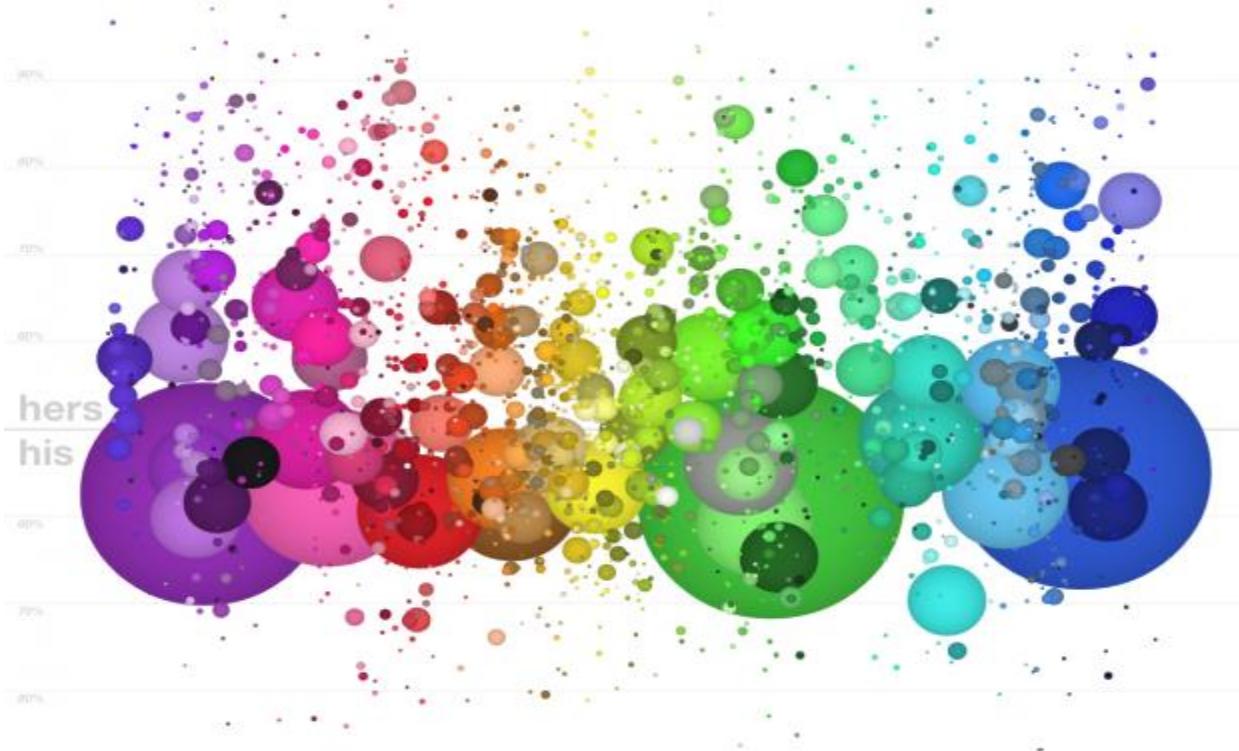


🚫 Gaudy & busy



☑ Simple & elegant

Colors Have Meaning



**Red** is the color of energy, passion, action, ambition and determination. It is also the color of anger and passion.

**Orange** is the color of social communication and optimism. From a negative color meaning it is also a sign of pessimism and superficiality.

**Yellow** is the color of the mind and the intellect. It is optimistic and cheerful. However it can also suggest impatience, criticism and cowardice.

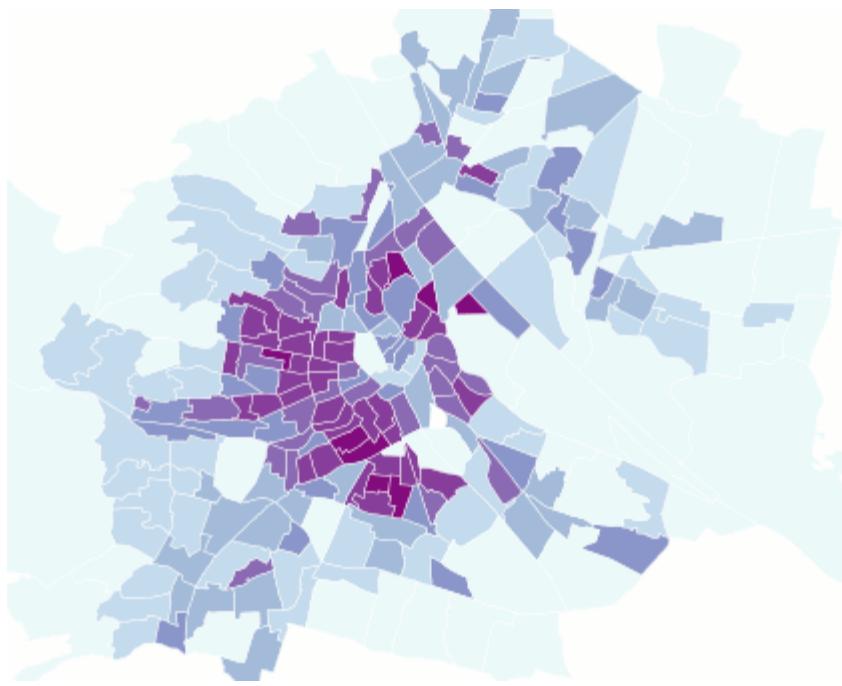
**Green** is the color of balance and growth. It can mean both self-reliance as a positive and possessiveness as a negative, among many other meanings.

**Blue** is the color of trust and peace. It can suggest loyalty and integrity as well as conservatism and frigidity.

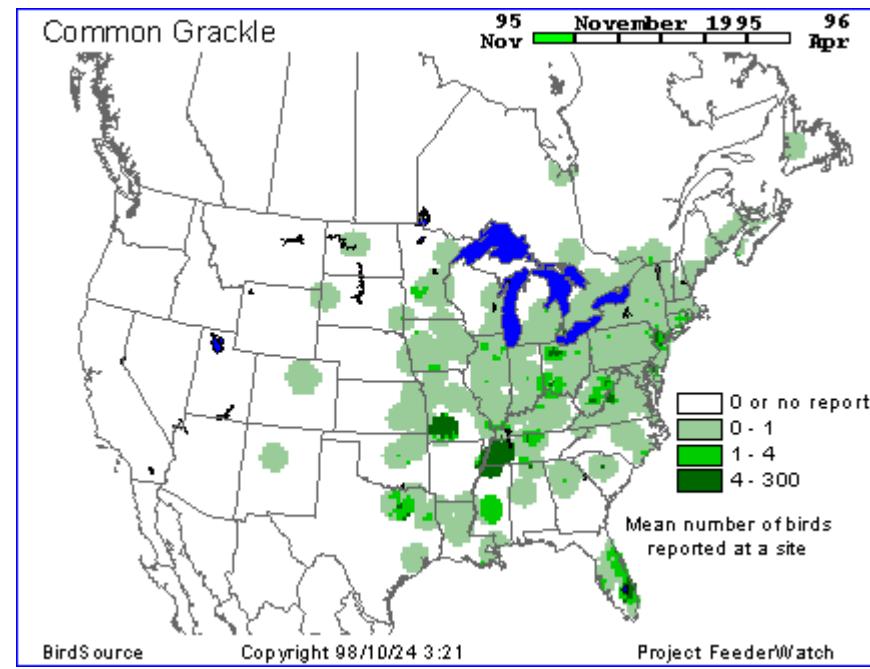
# Movement Provides Information

Data **Vi**sualization



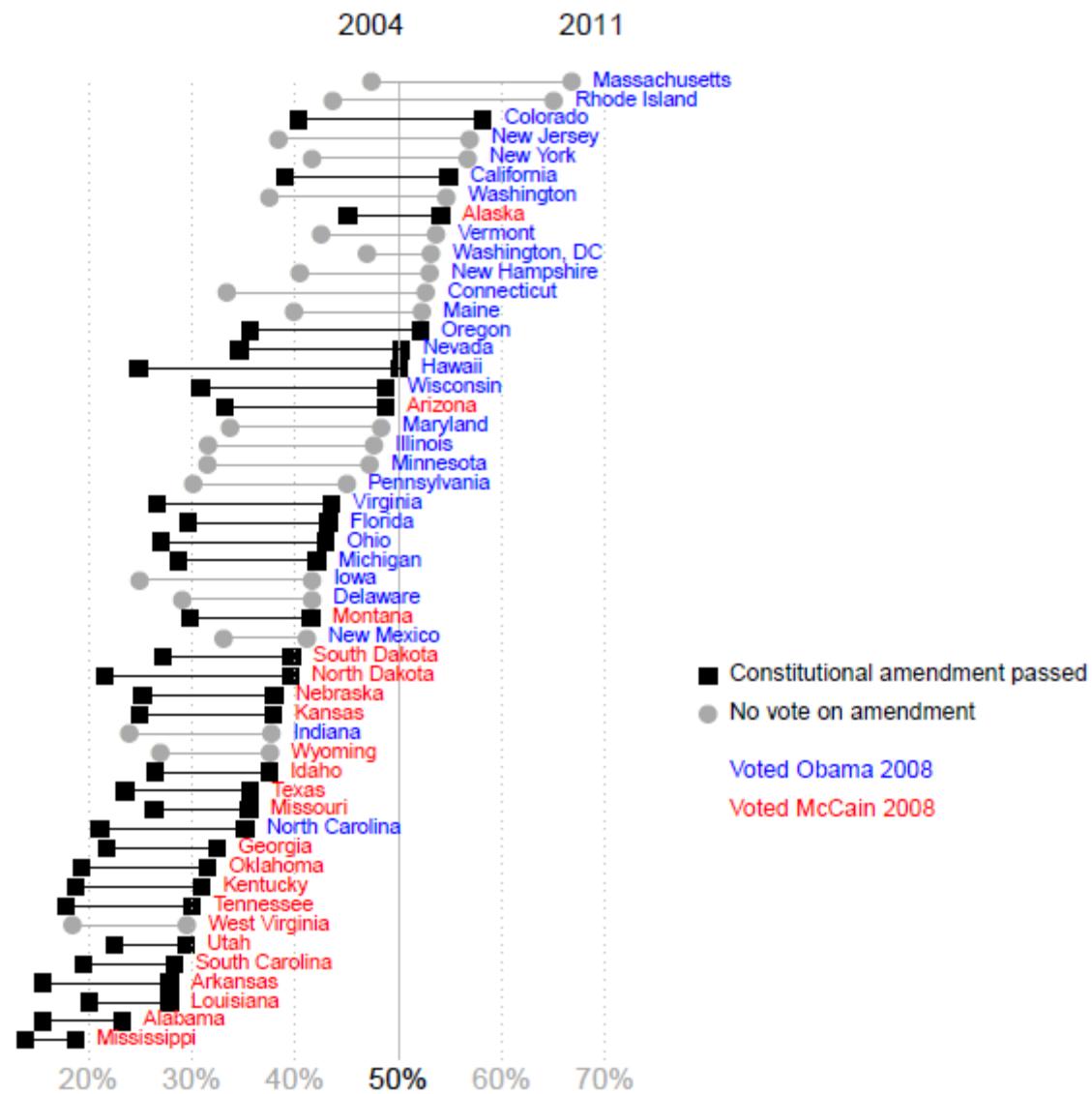


Two population density measures compared.

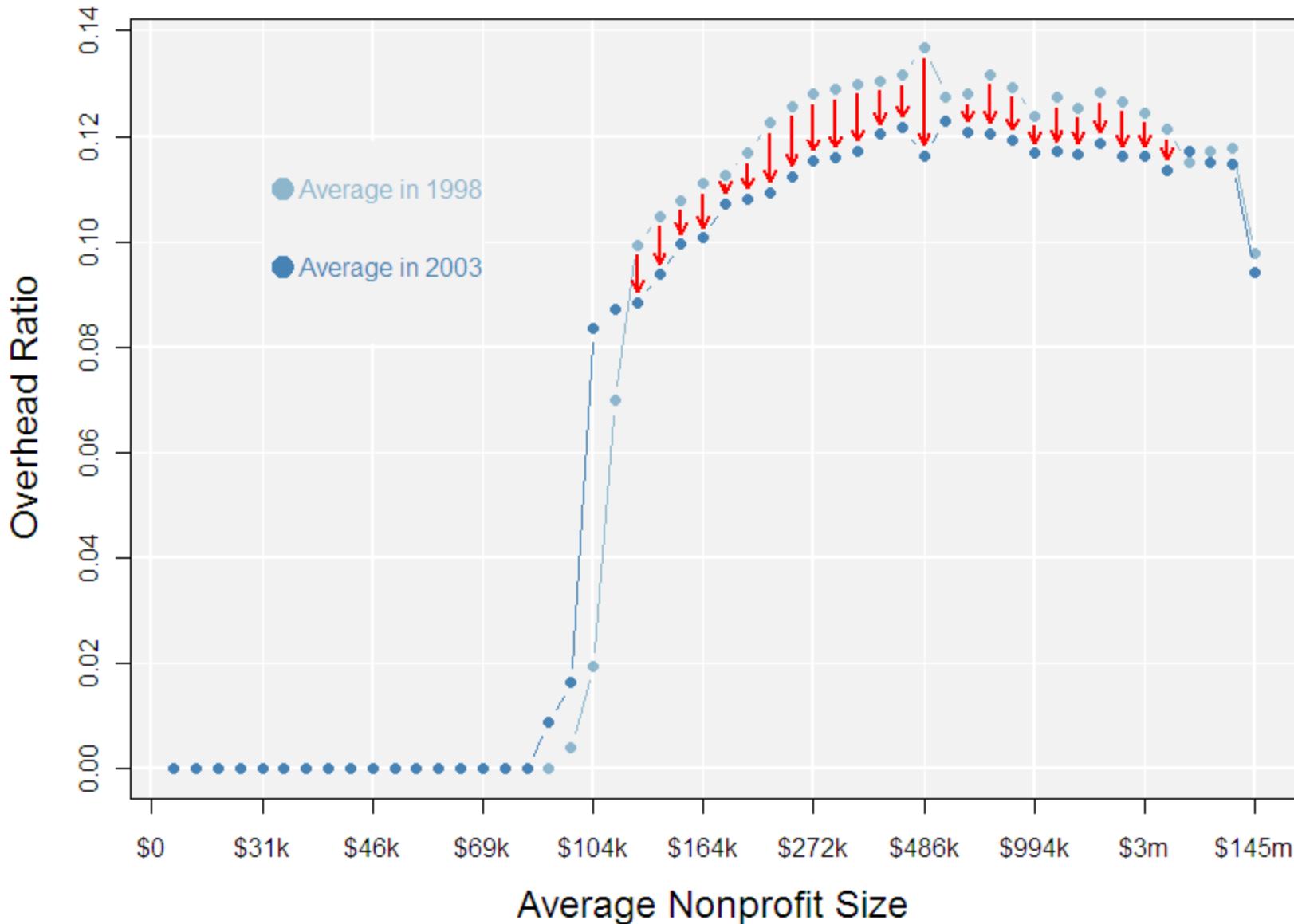


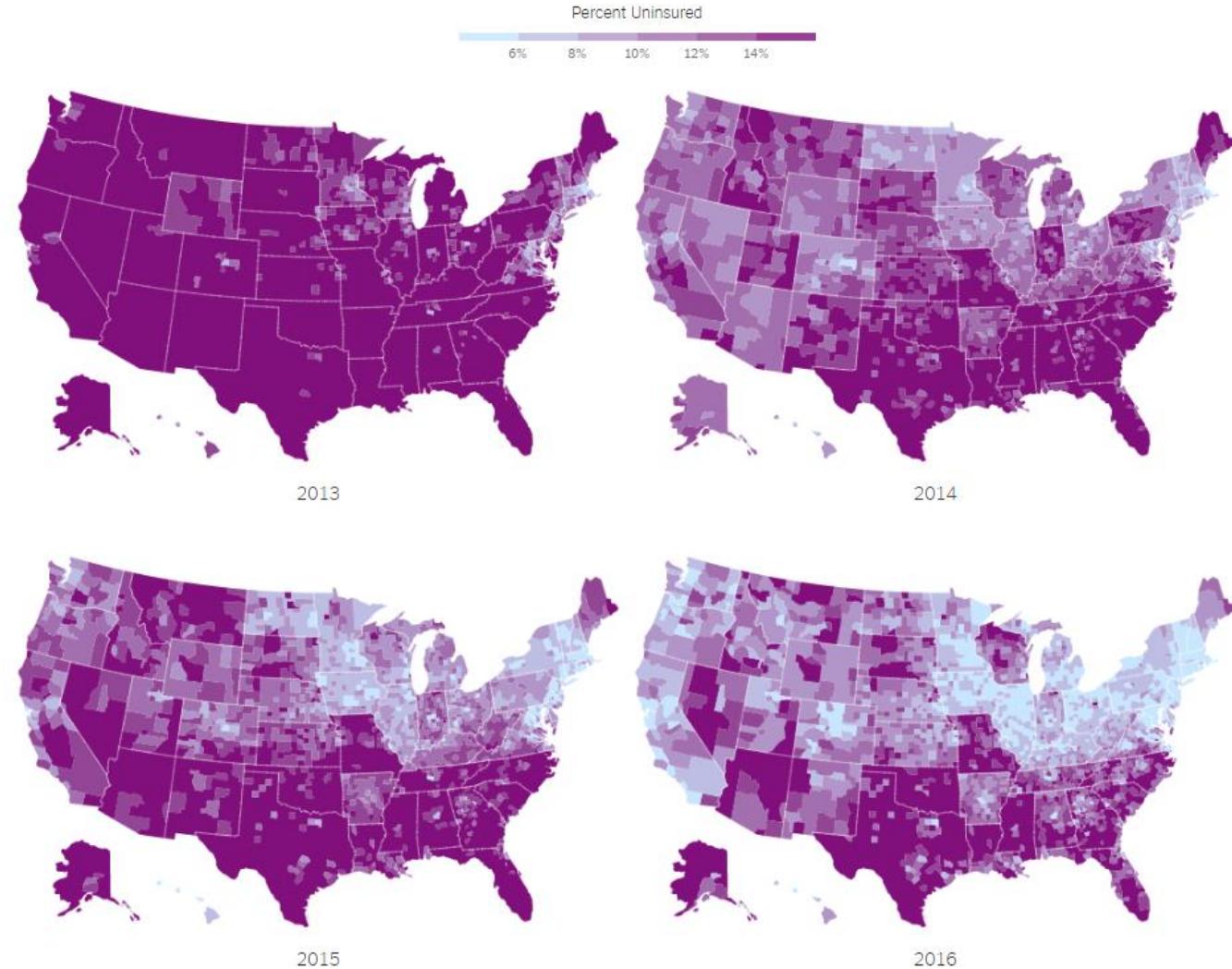
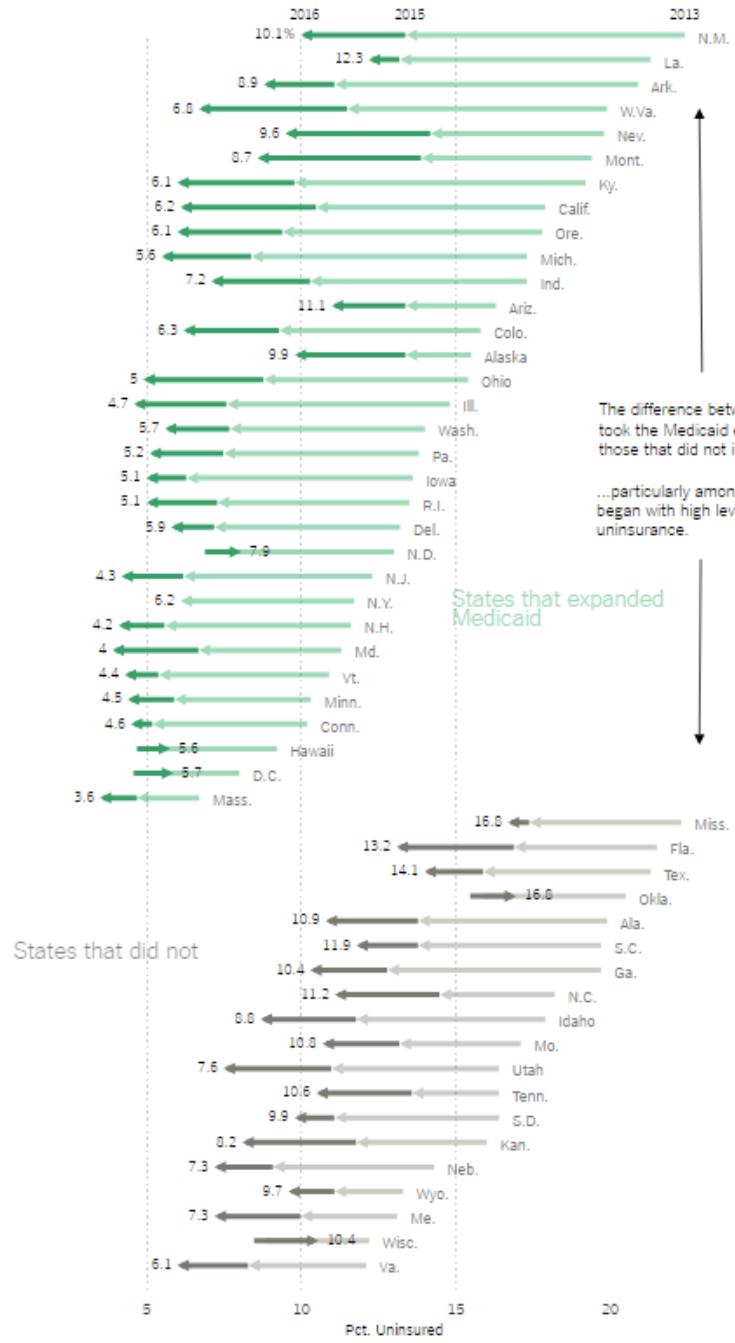
Migration patterns of birds.

## Percentage Supporting Same Sex Marriage



## Change in Overhead by Nonprofit Size 1998-2003





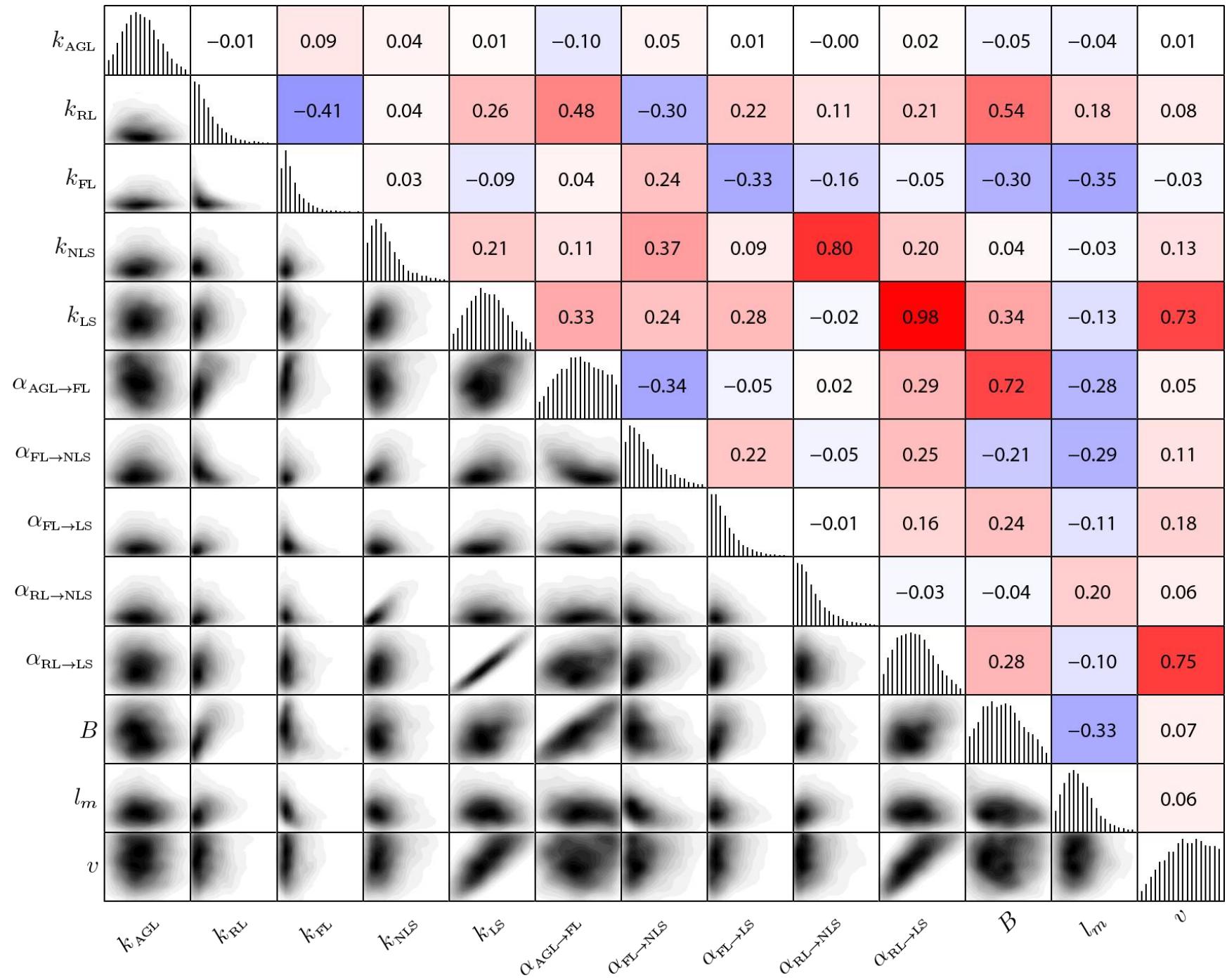
<https://www.nytimes.com/interactive/2018/03/19/upshot/race-class-white-and-black-men.html>

# Visualize Tables

**Table III** Descriptive statistics and correlation matrix for study variables – correlation matrix

	MO	FP	MP	IM	IC	FM	FE	FI	SPC	DSC	DST
MO	1.00										
FP	0.31 <sup>a</sup>	1.00									
MP	0.32 <sup>a</sup>	0.71 <sup>a</sup>	1.00								
IM	0.36 <sup>a</sup>	0.12 <sup>c</sup>	0.14 <sup>c</sup>	1.00							
IC	0.39 <sup>a</sup>	0.18 <sup>b</sup>	0.21 <sup>a</sup>	0.62 <sup>a</sup>	1.00						
FM	0.26 <sup>a</sup>	0.21 <sup>a</sup>	0.14 <sup>c</sup>	0.30 <sup>a</sup>	0.27 <sup>a</sup>	1.00					
FE	0.47 <sup>a</sup>	0.21 <sup>a</sup>	0.18 <sup>b</sup>	0.38 <sup>a</sup>	0.28 <sup>a</sup>	0.24 <sup>a</sup>	1.00				
FI	0.53 <sup>a</sup>	0.26 <sup>a</sup>	0.22 <sup>a</sup>	0.36 <sup>a</sup>	0.37 <sup>a</sup>	0.29 <sup>a</sup>	0.47 <sup>a</sup>	1.00			
SPC	0.32 <sup>a</sup>	0.22 <sup>a</sup>	0.31 <sup>a</sup>	0.51 <sup>a</sup>	0.47 <sup>a</sup>	0.32 <sup>a</sup>	0.37 <sup>a</sup>	0.35 <sup>a</sup>	1.00		
DSC	-0.12 <sup>c</sup>	0.03 <sup>c</sup>	0.05 <sup>c</sup>	0.17 <sup>b</sup>	0.08 <sup>c</sup>	0.18 <sup>b</sup>	-0.05 <sup>c</sup>	0.06 <sup>c</sup>	0.01 <sup>c</sup>	1.00	
DST	-0.02 <sup>c</sup>	-0.01 <sup>c</sup>	0.05 <sup>c</sup>	0.24 <sup>a</sup>	0.14 <sup>c</sup>	0.05 <sup>c</sup>	-0.05 <sup>c</sup>	0.05 <sup>c</sup>	0.05 <sup>c</sup>	0.56 <sup>a</sup>	1.00
DM	0.05 <sup>c</sup>	0.144	0.136 <sup>c</sup>	0.199 <sup>a</sup>	0.169 <sup>b</sup>	0.247 <sup>a</sup>	0.08 <sup>c</sup>	0.11 <sup>c</sup>	0.14 <sup>c</sup>	0.46 <sup>a</sup>	0.71 <sup>a</sup>

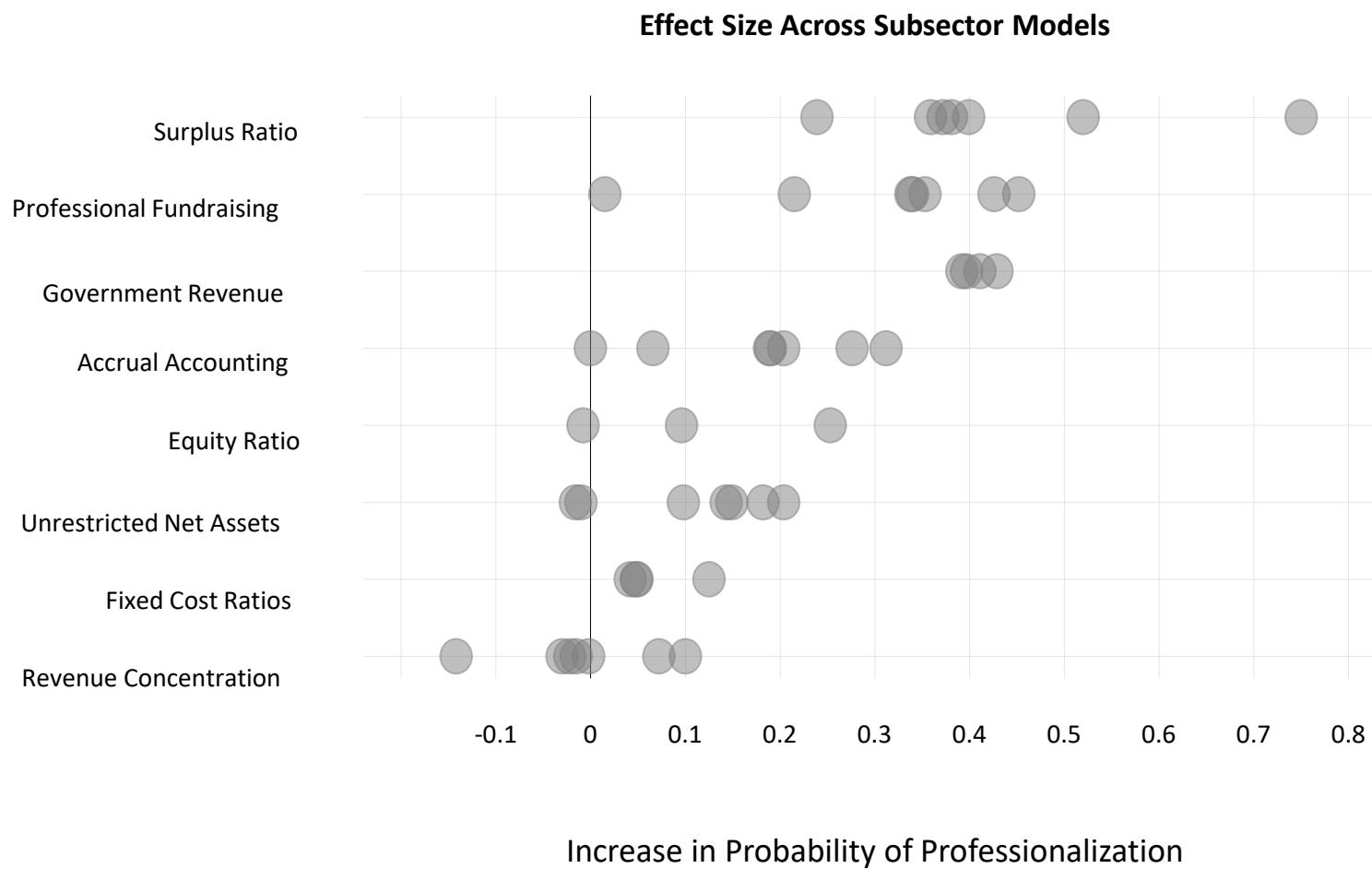
Notes: <sup>a</sup> correlation is significant at the 0.01 level (two-tailed); <sup>b</sup> correlation is significant at the 0.05 level (two-tailed); <sup>c</sup> correlation is non-significant



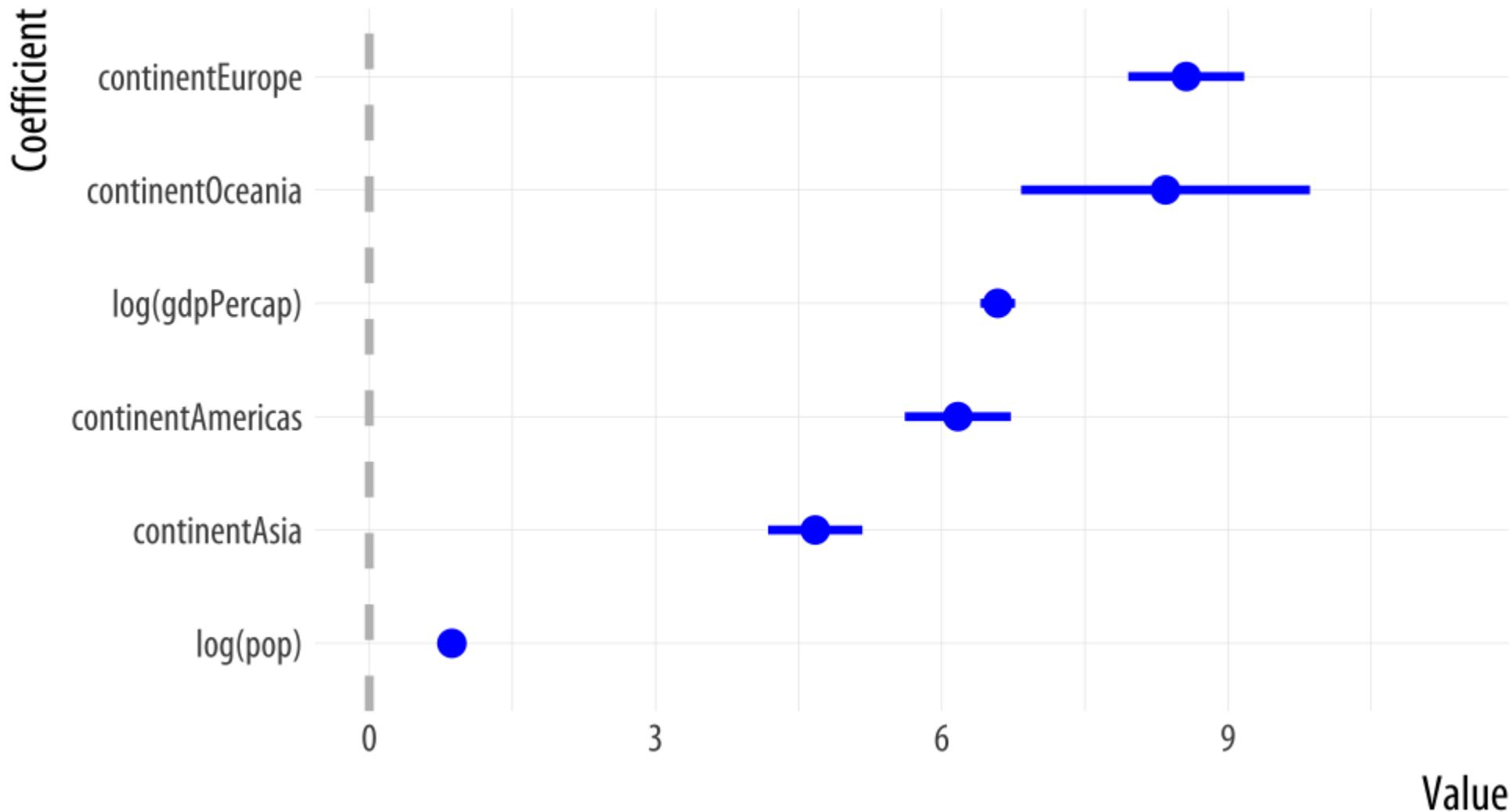
## Table of Effect Sizes

	ALL	ARTS	HEALTH	HUMAN SERVICES	PUBLIC	EDUCATION	OTHER
<b>Surplus Ratio</b>	0.381	0.520	0.239	0.359	0.372	0.399	0.750
<b>Professional Fundraising</b>	0.340	0.426	0.215	0.338	0.353	0.452	0.015
<b>Government Revenue</b>	0.392	0.365	0.270	0.397	0.411	0.429	0.014
<b>Accrual Accounting</b>	0.190	0.276	0.066	0.189	0.204	0.312	0.000
<b>Equity Ratio</b>	0.114	0.193	-0.035	0.096	0.206	0.253	-0.008
<b>Unrestricted Net Assets</b>	0.098	0.143	-0.016	0.149	0.204	0.182	-0.010
<b>Fixed Cost Ratios</b>	0.042	0.125	-0.081	0.048	0.049	0.139	-0.010
<b>Revenue Concentration</b>	-0.022	0.072	-0.142	-0.030	-0.002	0.100	-0.015

Figure X: Effects Across Models



# Coefficient Plot



# In Conclusion

# Twenty rules for good graphics

1. Use vector graphics such as eps or pdf. These scale properly and do not look fuzzy when enlarged. Do not use jpeg, bmp or png files as these will look fuzzy when enlarged, or if saved at very high resolutions will be enormous files. Jpegs in particular are designed for photographs not statistical graphics.
2. Use readable fonts. For graphics I prefer sans-serif fonts such as Helvetica or Arial. Make sure the font size is readable after the figure is scaled to whatever size it will be printed.
3. Avoid cluttered legends. Where possible, add labels directly to the elements of the plot rather than use a legend at all. If this won't work, then keep the legend from obscuring the plotted data, and make it small and neat.
4. If you must use a legend, move it inside the plot, in a blank area.
5. No dark shaded backgrounds. Excel always adds a nasty dark gray background by default, and I'm always asking authors to remove it. Graphics print much better with a white background. The ggplot for R also uses a gray background (although it is lighter than the Excel default). I don't mind the ggplot version so much as it is used effectively with white grid lines. Nevertheless, even the light gray background doesn't lend itself to printing/photocopying. White is better.
6. Avoid dark, dominating grid lines (such as those produced in Excel by default). Grid lines can be useful, but they should be in the background (light gray on white or white on light gray).
7. Keep the axis limits sensible. You don't have to include a zero (even if Excel wants you to). The defaults in R work well. The basic idea is to avoid lots of white space around the plotted data.
8. Make sure the axes are scaled properly. Another Excel problem is that the horizontal axis is sometimes treated categorically instead of numerically. If you are plotting a continuous numerical variable, then the horizontal axis should be properly scaled for the numerical variable.
9. Do not forget to specify units.
10. Tick intervals should be at nice round numbers.
11. Axes should be properly labelled.
12. Use linewidths big enough to read. 1pt lines tend to disappear if plots are shrunk.
13. Avoid overlapping text on plotting characters or lines.
14. Follow Tufte's principles by removing chart junk and keeping a high data-ink ratio.
15. Plots should be self-explanatory, so include detailed captions.
16. Use a sensible aspect ratio. I think width:height of about 1.6 works well for most plots.
17. Prepare graphics in the final aspect ratio to be used in the publication. Distorted fonts look awful.
18. Use points not lines if element order is not relevant.
19. When preparing plots that are meant to be compared, use the same scale for all of them. Even better, combine plots into a single graph if they are related.
20. Avoid pie-charts. Especially 3d pie-charts. Especially 3d pie-charts with exploding wedges. I promise all my students an instant fail if I ever see anything so appalling.