

# Data Visualization

This class isn't mainly about how to **DESIGN** visualizations.  
It's about how to **REASON** about visualization.

**Anybody can learn visualization**

for the same reasons that anybody can learn to read and write

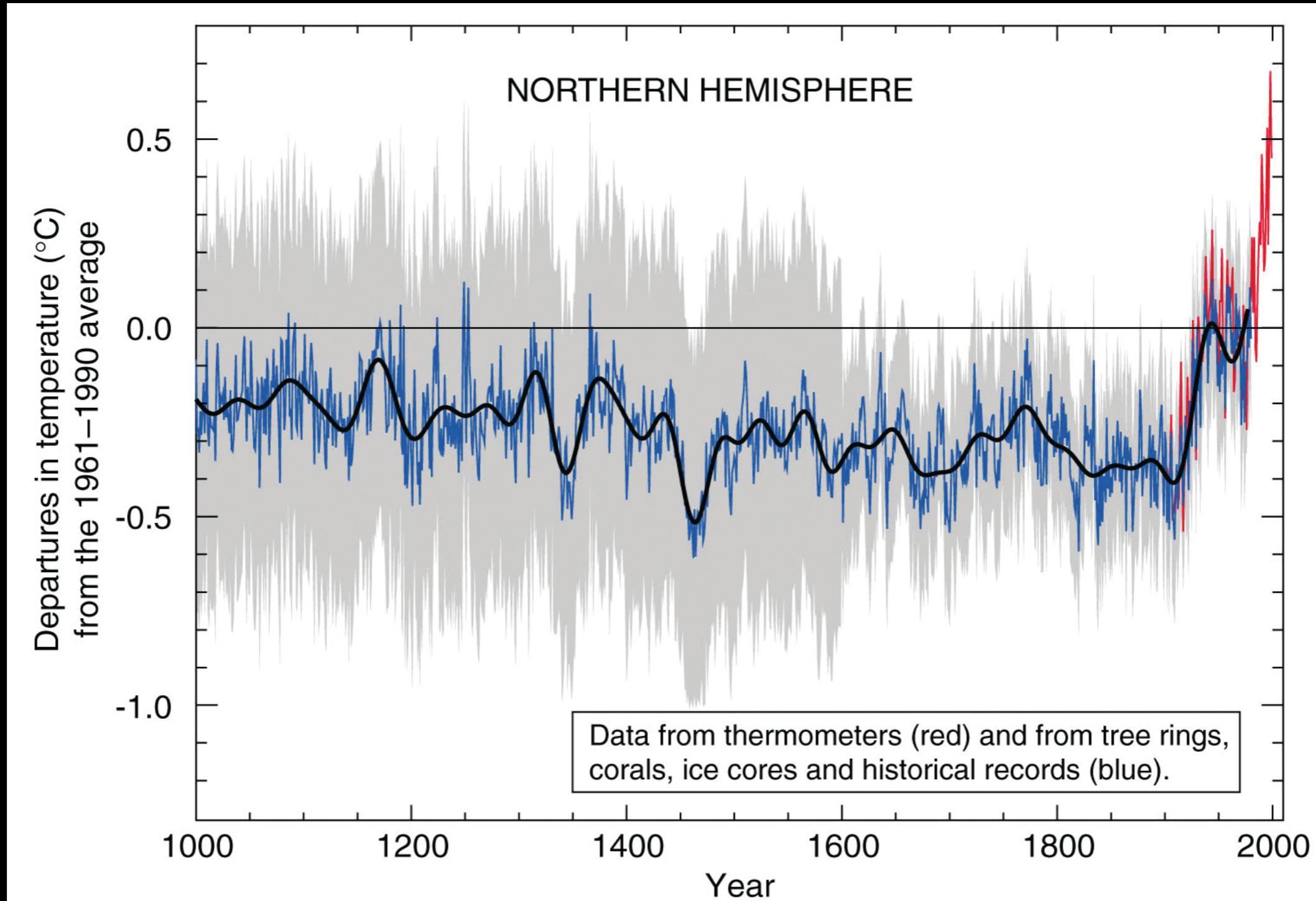
1. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles

1. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles

A **visualization** is a  
representation designed to  
enable **exploration**,  
**discovery**, or **communication**

	A	B	C	D	E	F	G
1	YEAR	TEMP	YEAR	1 SIGMA	2 SIGMA		
2	1000	0.0659	1000	0.240346	0.480693	0.206137	0.123588
3	1001	-0.1241	1001	0.240347	0.480694	0.206137	0.123589
4	1002	-0.1208	1002	0.240346	0.480692	0.206136	0.123588
5	1003	-0.1801	1003	0.240347	0.480694	0.206137	0.123589
6	1004	-0.0711	1004	0.240347	0.480693	0.206137	0.123588
7	1005	-0.1334	1005	0.240346	0.480692	0.206136	0.123588
8	1006	-0.0644	1006	0.240346	0.480693	0.206137	0.123588
9	1007	0.0042	1007	0.240347	0.480693	0.206137	0.123588
10	1008	-0.1288	1008	0.240347	0.480693	0.206137	0.123588
11	1009	-0.0296	1009	0.240347	0.480693	0.206137	0.123588
12	1010	0.1187	1010	0.240347	0.480694	0.206137	0.123589
13	1011	-0.1252	1011	0.240346	0.480692	0.206136	0.123588
14	1012	-0.1634	1012	0.240347	0.480694	0.206137	0.123588
15	1013	-0.0791	1013	0.240347	0.480693	0.206137	0.123588
16	1014	-0.1120	1014	0.240347	0.480693	0.206137	0.123588
17	1015	-0.1146	1015	0.240346	0.480692	0.206136	0.123588
18	1016	-0.1206	1016	0.240346	0.480692	0.206136	0.123588
19	1017	-0.0815	1017	0.240347	0.480693	0.206137	0.123588
20	1018	-0.2031	1018	0.240346	0.480693	0.206137	0.123588
21	1019	0.0305	1019	0.240347	0.480693	0.206137	0.123588
22	1020	0.1098	1020	0.240347	0.480694	0.206137	0.123589
23	1021	0.0244	1021	0.240347	0.480693	0.206137	0.123588
24	1022	-0.0743	1022	0.240347	0.480693	0.206137	0.123588
25	1023	-0.0323	1023	0.240347	0.480693	0.206137	0.123588
26	1024	-0.0434	1024	0.240346	0.480693	0.206137	0.123588

	A	B	C	D	E	F	G
878	1876	-0.1891	1876	0.113228	0.226456	8.25297E-02	7.75207E-
879	1877	-0.0140	1877	0.113228	0.226457	8.25299E-02	7.75209E-
880	1878	-0.0873	1878	0.113228	0.226457	8.25298E-02	7.75209E-
881	1879	-0.2959	1879	0.113229	0.226458	8.25302E-02	7.75212E-
882	1880	-0.2368	1880	0.113229	0.226457	8.25300E-02	7.75210E-
883	1881	-0.1977	1881	0.113229	0.226458	8.25302E-02	7.75212E-
884	1882	-0.2036	1882	0.113229	0.226457	8.25300E-02	7.75210E-
885	1883	-0.2489	1883	0.113228	0.226455	8.25293E-02	7.75204E-
886	1884	-0.2125	1884	0.113229	0.226457	8.25301E-02	7.75211E-
887	1885	-0.1896	1885	0.113228	0.226457	8.25299E-02	7.75210E-
888	1886	-0.1084	1886	0.113228	0.226456	8.25298E-02	7.75208E-
889	1887	-0.3265	1887	0.113228	0.226456	8.25296E-02	7.75206E-
890	1888	-0.1694	1888	0.113228	0.226457	8.25298E-02	7.75209E-
891	1889	-0.1339	1889	0.113228	0.226456	8.25298E-02	7.75208E-
892	1890	-0.3107	1890	0.113229	0.226457	8.25301E-02	7.75211E-
893	1891	-0.1754	1891	0.113229	0.226457	8.25300E-02	7.75210E-
894	1892	-0.3186	1892	0.113228	0.226456	8.25295E-02	7.75205E-
895	1893	-0.3236	1893	0.113228	0.226456	8.25297E-02	7.75207E-
896	1894	-0.1970	1894	0.113228	0.226456	8.25295E-02	7.75205E-
897	1895	-0.1578	1895	0.113228	0.226456	8.25297E-02	7.75207E-
898	1896	-0.0804	1896	0.113228	0.226456	8.25298E-02	7.75208E-
899	1897	-0.0537	1897	0.113228	0.226456	8.25298E-02	7.75208E-
900	1898	-0.2195	1898	0.113229	0.226457	8.25301E-02	7.75211E-
901	1899	-0.3486	1899	0.113228	0.226456	8.25297E-02	7.75207E-
902	1900	-0.1253	1900	0.113229	0.226457	8.25300E-02	7.75210E-
903	1901	-0.1575	1901	0.113228	0.226456	8.25296E-02	7.75206E-



**Michael E. Mann, Raymond S. Bradley, and Malcolm K. Hughes**  
Intergovernmental Panel on Climate Change (IPCC), Third Report, 2001



# A golden age of visualization in news media

**The New York Times**  
Thursday, January 23, 2018  
Today's Paper | Video | 27° | Dow +0.15%

## Trump Says He Is Willing to Speak Under Oath to Mueller

BY MARCO LACROIX AND JULIE HERSHLEIGH  
President Trump said he was "looking forward" to speaking with Robert S. Mueller III, the special counsel for the Russia inquiry.

### More than 160 women say Larry Nassar sexually abused them. Here are his accusers in their own words.

By CARLA CORREA and MEGHAN LUCYTT

**The Washington Post**  
Democracy Dies in Darkness  
January 23, 2018

## Surprises us, alarming offering with Mueller

BY RICH GOODWIN  
The news "looking forward" to the special counsel in the White House "is a study in surprise. But the headline reflects our performance that the can't speak."

### We're losing some of the classic culture's high schools and lockers no longer the right combination

Once the go-to center of the high school day, lockers long ago lost their allure. As today's students move clear that they have no interest in changing the scene, it looks as if the lockers from their hallways and buildings don't even include them in their design.

### Manhattan District Attorney Says He'll No Longer Accept Contributions From Lawyers With Cases Before Him

By Jesse Klagberg  
Jan. 23, 7:04 p.m. EST

**PROPUBLICA**  
TOPICS | SERIES | NEWS APPS | GET INVOLVED | IMPACT | ABOUT |

## The Red Cross Forced Out an Executive Over Sexual Harassment - Then Helped Him Land a Job at Save the Children

By Austin Elliott and Ariana Tablin  
Jan. 23, 9 a.m. EST  
A senior Red Cross official harassed a subordinate and was accused of raping another. The charity's now-general counsel David Melzer praised Fire on Fire way out for "transparency" and "dedication."

### Chicago Cop Under Investigation Again Over Social Media Posts

By Josh B. Cohen,  
Jan. 23, 4 a.m. CST

### Baltimore Judge Tosses Alford Plea, Rehiring Prosecutor

By Megan Rose,  
Jan. 23, 7:32 p.m. EST  
Demetrius Smith has long maintained he pleaded guilty to a shooting he did not commit. Now, over the prosecutor's objection, his conviction has been set aside.

**Standard**  
LOGIN | TODAY | MAGAZINE | WRITERS | PODCASTS | BOOKS & ARTS | NEWSLETTERS | EVENTS | STORE

## LABASH: Drone-Assisted Fishing Is Real, and It's Pathetic

Want your fish to be big? Drone-assisted fishing is real, but it's a bit pathetic. The government even has a program for it. But the drone-assisted fishing is real, and it's pathetic. The government even has a program for it. But the drone-assisted fishing is real, and it's pathetic.

**AXIOS**  
SECTIONS | SPECIAL FEATURES | MORE |

## The insane news cycle of Trump's presidency in 1 chart

If it feels like the Trump presidency has been hit by non-stop scandals and crises from day one, it's probably because it has been. The Google News Lab looked at the search trends for studies about 10 of the biggest news events of Trump's presidency from Jan. 20 to Sept. 1. You can see how we've all jumped from one four-alarm news fire to another.

Topic	Search Interest (Index)
Height is search interest in a given topic, indexed to 100	100
Color is average search interest from January to September	100

**FiveThirtyEight**  
Politics | Sports | Science & Health | Economics | Culture | Politics Podcast: The Three-Day Shutdown

## 7 Ways To Redistrict The Entire Freely Country - From Fair To Heavily Gerrymandered

By Aaron Byrnes, Ella Koski, David Wasserman and Julia Wolfe

### Significant Digits For Thursday, Jan. 24, 2018

8:00 AM  
The Atlas Of Redistricting

### Hating Gerrymandering Is Easy: Fixing It Is Harder.

5:50 AM  
We Drew 2,568 Congressional Districts By Hand. Here's How.

### Why Do Gerrymandered

3:14, 24

**Vox**  
EXPLAINERS | POLITICS & POLICY | WORLD | CULTURE | SCIENCE & HEALTH | IDENTITIES | MORE |

## Activists got Democrats to shut down the government. Can they do it again?

"We don't see it." Activists don't think Democrats won the shutdown fight.

### The sex abuse scandal surrounding USA Gymnastics team doctor Larry Nassar, explained

Gymnasts he abused are being treated as if they're not. They consider more than just the doctor.

### 3 potential problems for an obstruction of justice case against Trump

Why some experts think Mueller would need a lot more to make a case.

**THE WALL STREET JOURNAL**  
World | U.S. | Politics | Economy | Business | Tech | Markets | Opinion | Lifestyle | Real Estate

## Trump Says He Is Willing to Testify to Mueller Under Oath

President Trump said he was willing to be sworn under oath to testify to special counsel Robert S. Mueller III about his role in the Russia investigation.

### White House's Infrastructure Push Targets Private Firms

As more members of Congress push for a bill to speed up the process, the White House is targeting private firms.

### As More Companies Demand Arbitration Agreements, Sexual Harassment Claims Persist

It's not clear if the new rules will be enough to speed up the process.

**FINANCIAL TIMES**  
WORLD BUSINESS NEWSPAPER  
TUESDAY 10 MAY 2018

## Downfall How the high-flying boss of China's Anbang fell foul of the ruling party

BY BENJAMIN SVETKEY  
The Chinese government has taken a major step in dismantling the empire of the country's most powerful private banker.

### Dozens die on Gaza border as US and Israel mark embassy opening

At least 12 people had died in the border area in the night.

### Haitian populists seek more time to hold deal

The deal would allow the country to hold a referendum on the constitution.

**Berliner Morgenpost**  
Home | Berlin | Bezirk | Interaktiv | Politik | Wirtschaft | Sport | Panorama | Kultur | Wissen | Reise | Lifestyle | Abo | Specials | Service

## Zehnjährige in Brandenburg von Laster überrollt

SCHWERER UNFALL  
Ein 10-jähriges Kind wurde von einem Laster überrollt.

### Meistgelesene Artikel

- SCHWERER UNFALL 25.01.2018  
Zehnjährige in Brandenburg von Laster überrollt
- VEROFFENTLICHUNG 25.01.2018  
Bundesrat nennt AfD-Politiker Björn Höcke „Bemd“
- DREIQUARTERBAHN 25.01.2018  
Zwei neue Bahnstationen für den Berliner Süden
- LANGJÄHRIGE DEBATE 25.01.2018  
Pferdekutschen am

**ТЕКСТИ.ORG.UA**  
НАШІ ВИКОНАННЯ: ДА. ВІКІ. ВІДП. РЕДАКЦІЯ. НОВА. ВАРШТАБ. НОВА. ВІКІ. АНАЛІТИЧНИЙ ЦЕНТР

## Діло йде до хунти. Цивільні органи влади втрачають довіру, військові нарощують (ІНФОГРАФІКА)

Володарі і збройні сили України внаслідок традиційно отримують найвищу довіру українського суспільства. Читати повністю

Category	Trust	Distrust
Government	45%	55%
Military	75%	25%

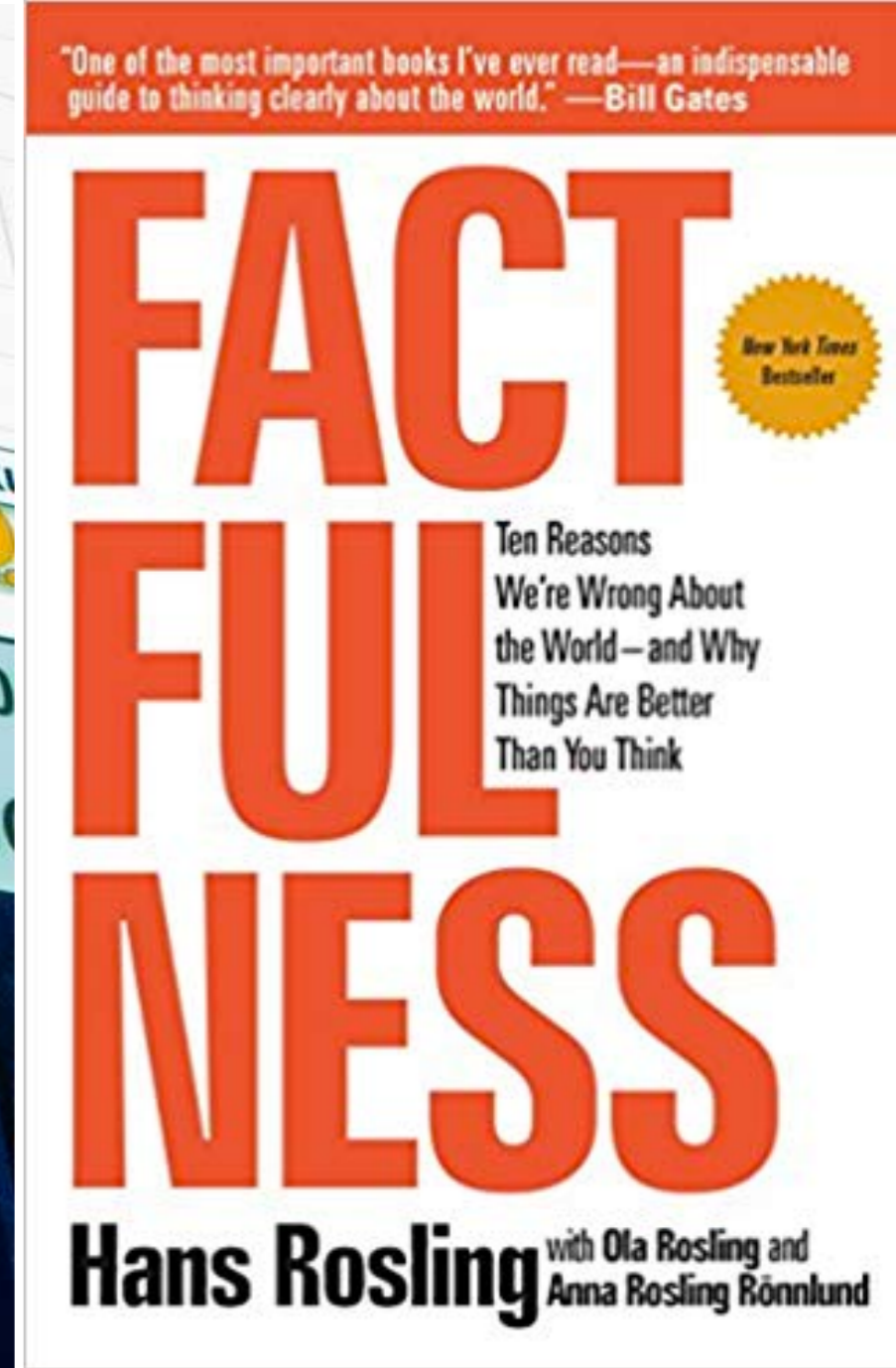
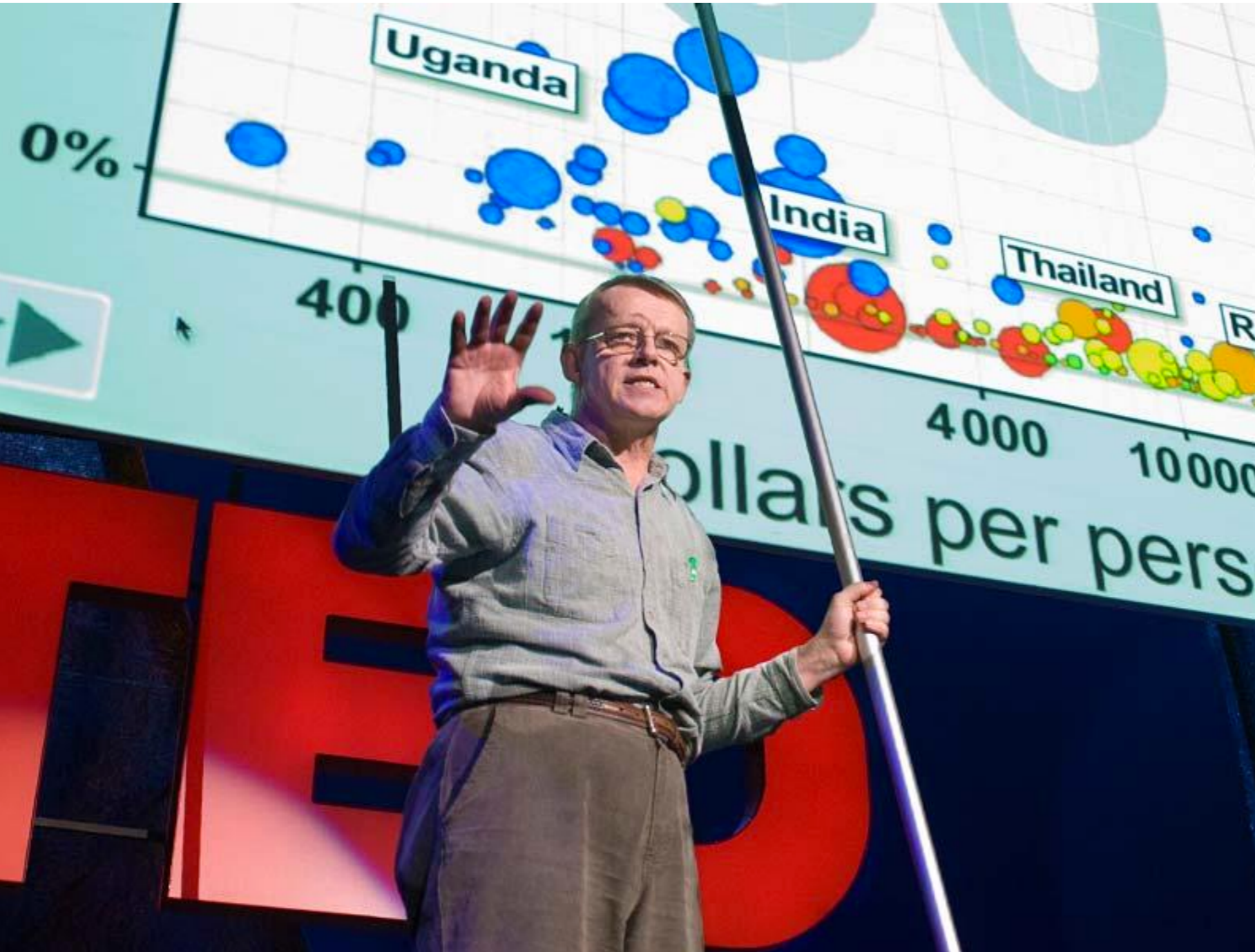
**ZEITUNG ONLINE**  
politics | society | economy | Culture | Knowledge | Digital campus | Job | Discover | Sports | time Mag

## A question of credibility

Martin Schulz loses recognition in the SPD. He has excluded entry to the cabinet Merkel. Can he still become vice-chancellor?

### Nashі інтерактивні візуалізації

Небезпечно для життя



## Popularizing visualization: Hans Rosling

<https://www.youtube.com/watch?v=hVimVzgtD6w>

[www.gapminder.org](http://www.gapminder.org)

1. Defining visualization
- 2.** The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles

# Layers of a visualization (from the bottom-up):

## 1. Scaffolding layer

Features that support the content such as legends, scales, axes, etc.

## 2. Encodings layer

The features that represent the data

## 3. Annotation layer

Words on the visualization itself that explains data points or put them in context

## 4. The “me” layer

Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?

# Layers of a visualization (from the bottom-up):

## 1. Scaffolding layer

Features that support the content such as legends, scales, axes, etc.

## 2. Encodings layer

The features that represent the data

## 3. Annotation layer

Words on the visualization itself that explains data points or put them in context

## 4. The “me” layer

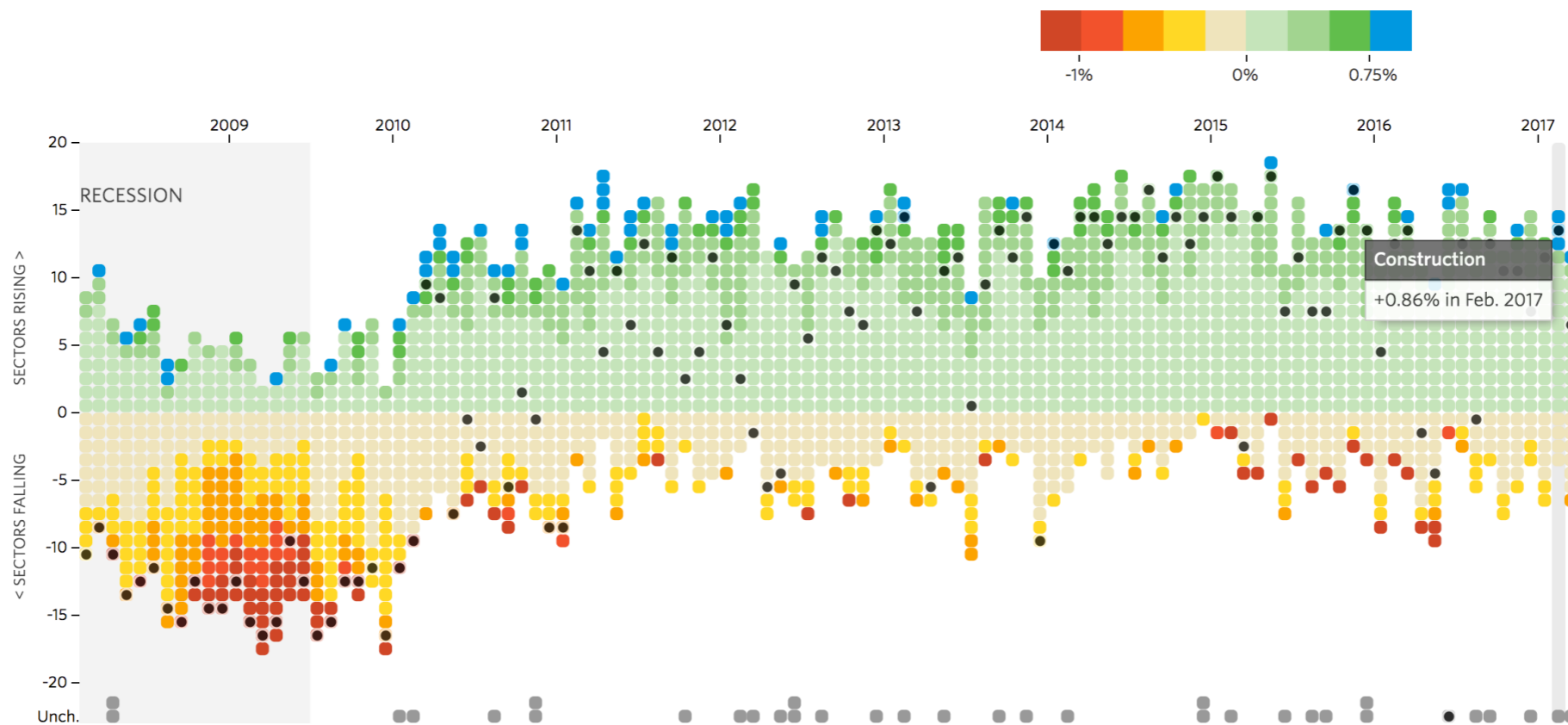
Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?

# Track National Unemployment, Job Gains and Job Losses

By [Andrew Van Dam](#) and [Renee Lightner](#)

## Winners and Losers: Job Gains and Losses [Jump to National Unemployment](#)

Track the number of sectors gaining or losing jobs each month. Boxes are shaded based on percentage change from the previous month in each sector's payrolls.



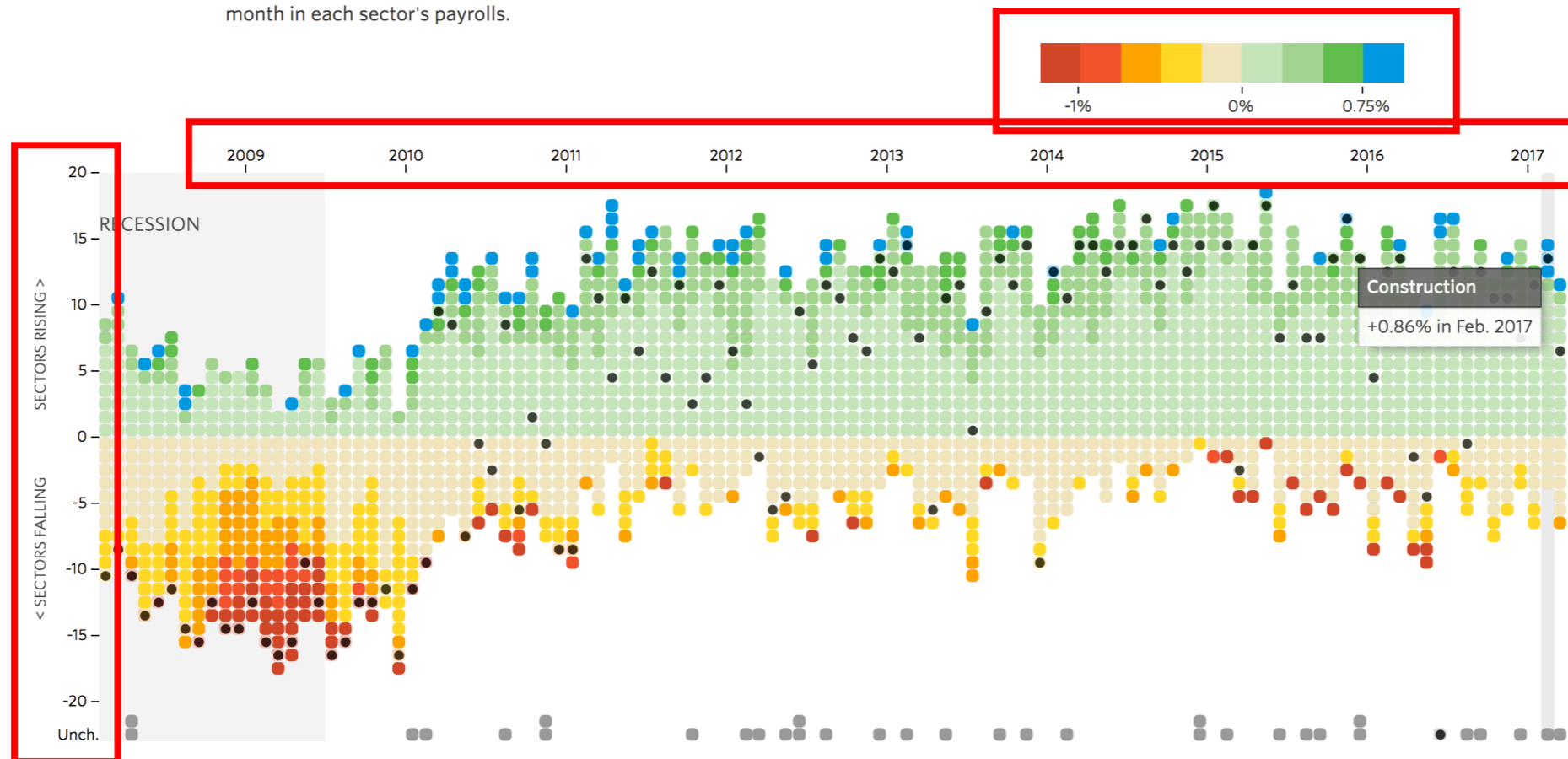
<https://graphics.wsj.com/job-market-tracker/>

# Track National Unemployment, Job Gains and Job Losses

By [Andrew Van Dam](#) and [Renee Lightner](#)

## Winners and Losers: Job Gains and Losses [Jump to National Unemployment](#)

Track the number of sectors gaining or losing jobs each month. Boxes are shaded based on percentage change from the previous month in each sector's payrolls.



<https://graphics.wsj.com/job-market-tracker/>

# Layers of a visualization (from the bottom-up):

## 1. Scaffolding layer

Features that support the content such as legends, scales, axes, etc.

## 2. Encodings layer

The features that represent the data

## 3. Annotation layer

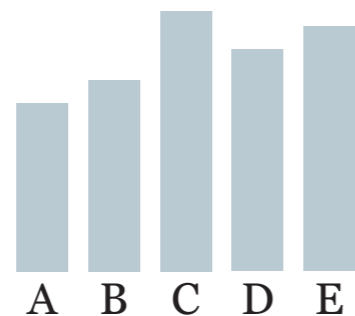
Words on the visualization itself that explains data points or put them in context

## 4. The “me” layer

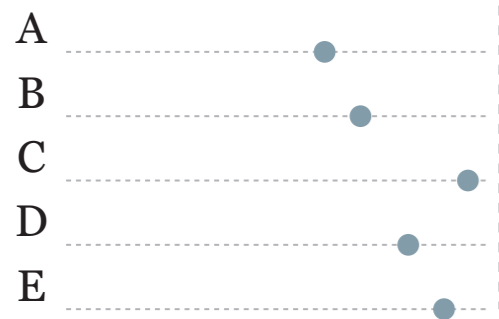
Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?



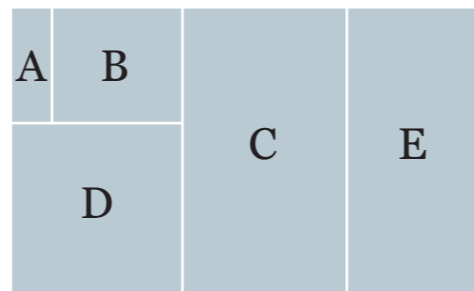
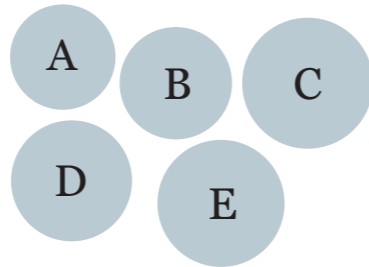
## Length or height



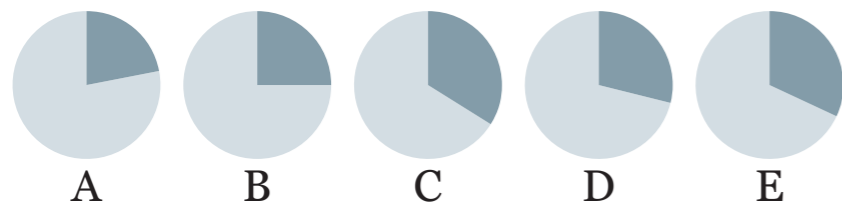
## Position



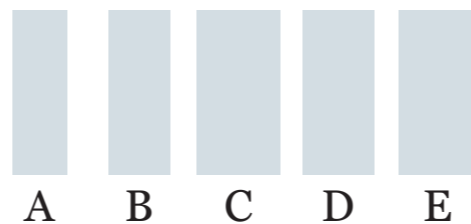
## Area



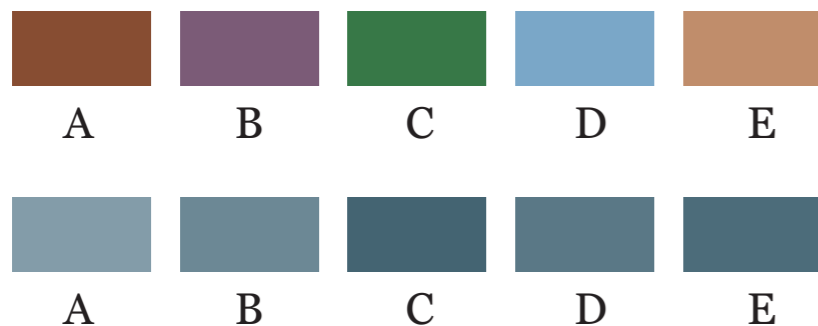
## Angle/area



## Line weight



## Hue and shade



Figures represented  
in all these graphics:  
22%, 25%, 34%, 29%, 32%

Data visualization consists of mapping data onto properties of objects (commonly abstract shapes.) These properties are called “encodings”, “methods of encoding”, or “aesthetics”

# Encodings

**Source: *The Truthful Art***

# What encodings can you spot here?

The **rate** of **violent crime** was up **236%** in **Milwaukee** between **1975** and **2015**.

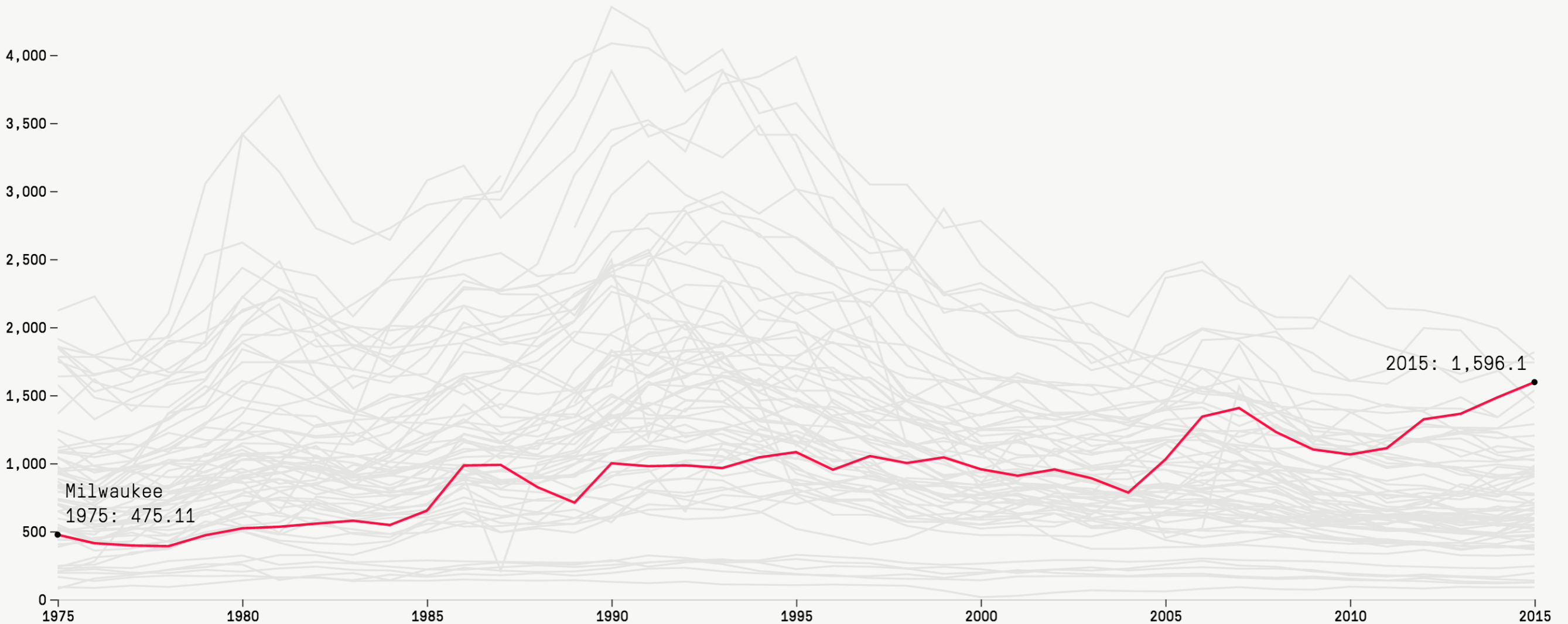
MILWAUKEE'S VIOLENT CLIMB

THE NEW YORK CITY DROP

BALTIMORE'S SPIKE

RANDOM

4,500 – VIOLENT CRIMES PER 100,000 PEOPLE



<https://www.themarshallproject.org/2016/08/18/crime-in-context#.Q2NYkV3mN>

# What encodings can you spot here?

POLICY-ISH

## Maps Show A Dramatic Rise In Health Insurance Coverage Under ACA

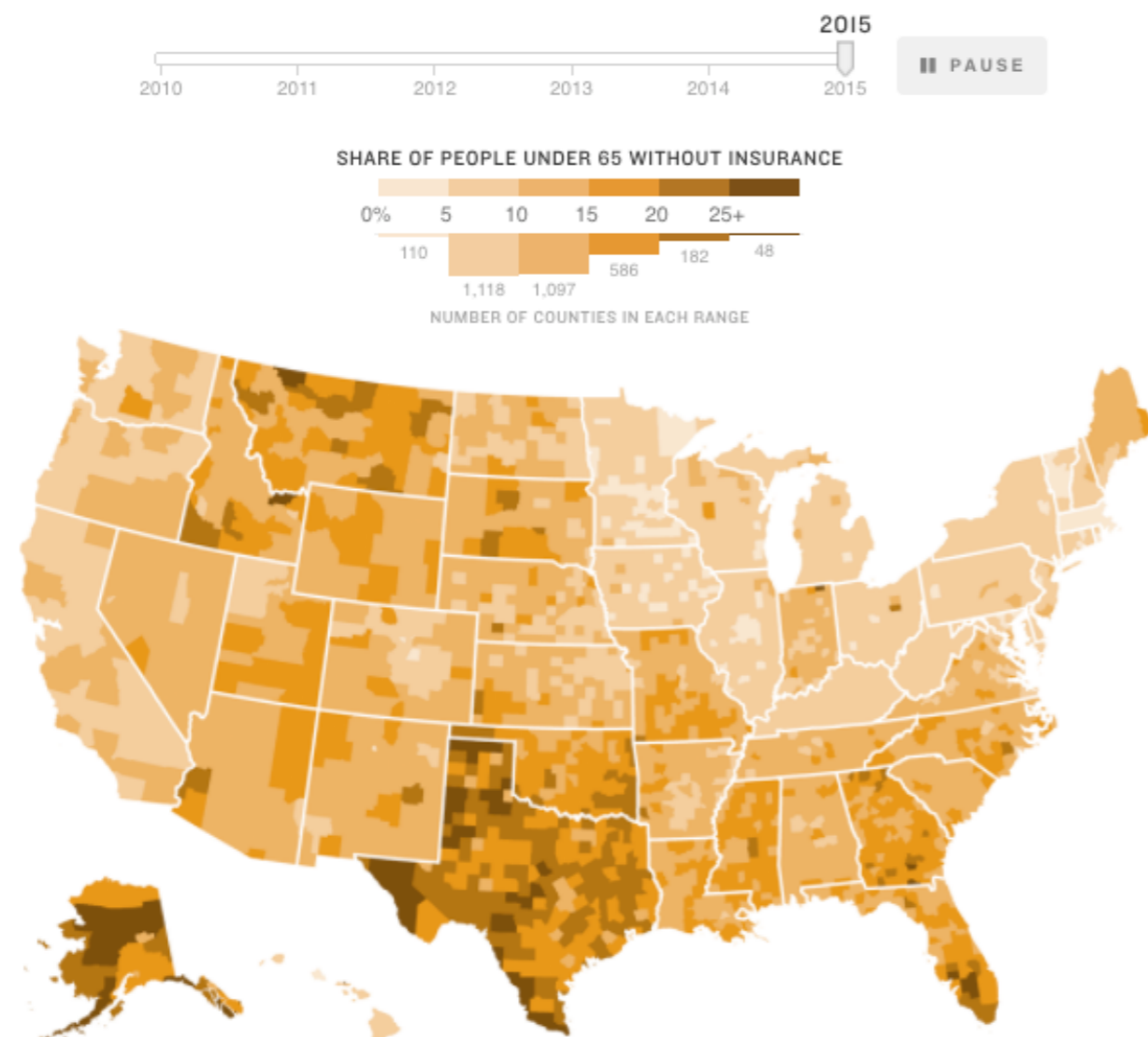
April 14, 2017 · 9:37 AM ET

ALYSON HURT



JUAN ELOSUA

REBECCA HERSHER



Source: NPR analysis of Census Bureau data

<http://www.npr.org/sections/health-shots/2017/04/14/522956939/maps-show-a-dramatic-rise-health-in-insurance-coverage-under-aca>

# Layers of a visualization (from the bottom-up):

## 1. Scaffolding layer

Features that support the content such as legends, scales, axes, etc.

## 2. Encodings layer

The features that represent the data

## 3. Annotation layer

Words on the visualization itself that explains data points or put them in context

## 4. The “me” layer

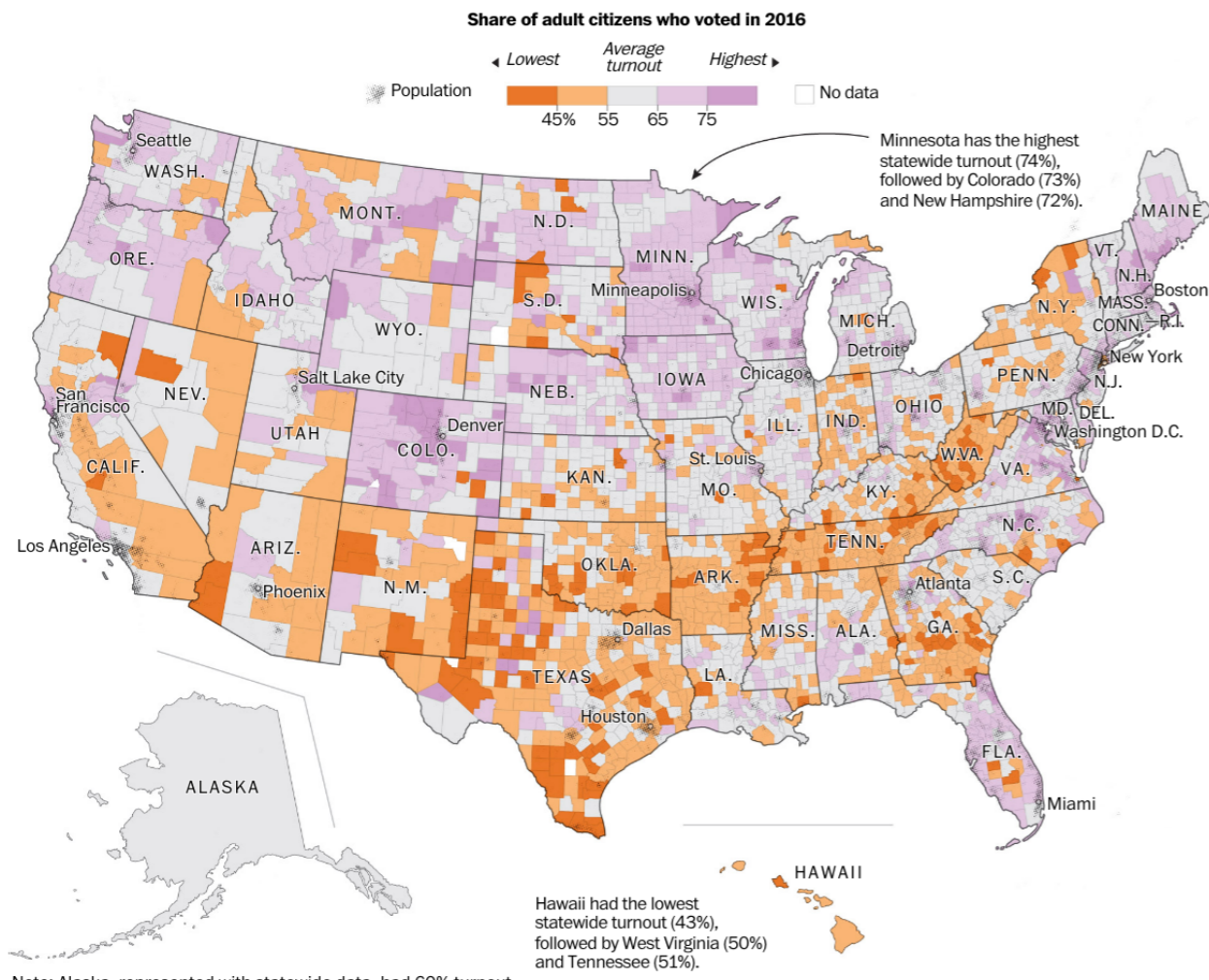
Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?

# The geography of voting — and not voting

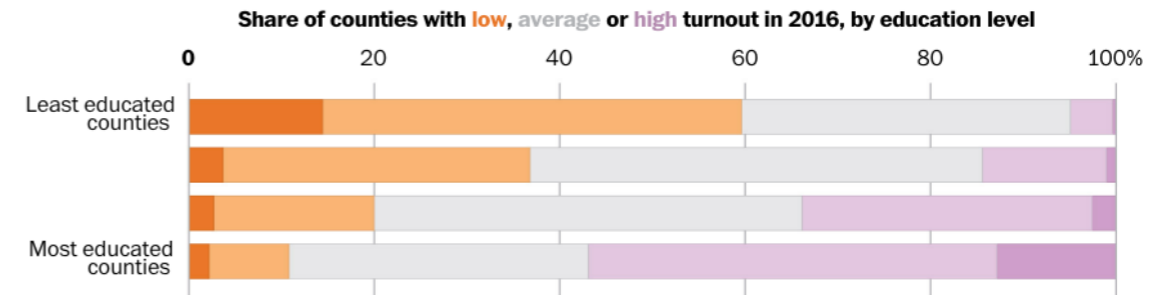
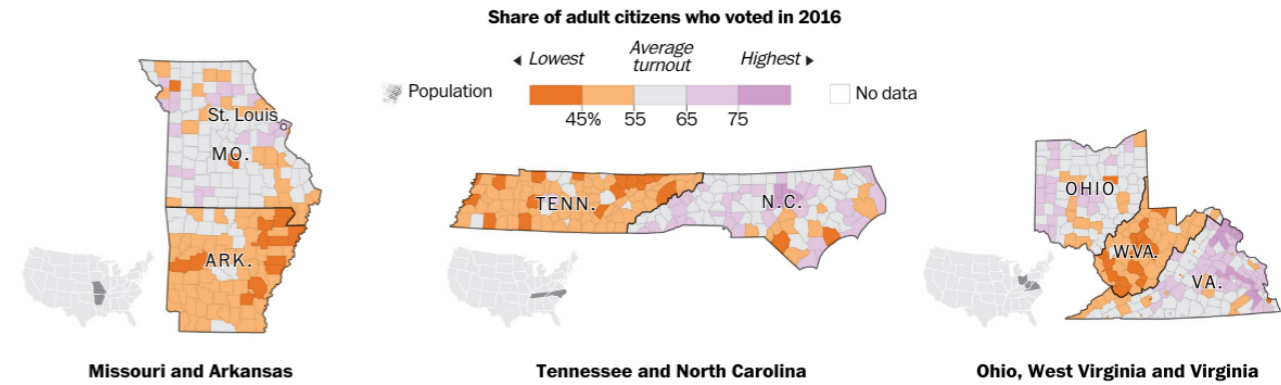
By Ted Mellnik, Lauren Tierney and Kevin Uhrmacher Oct. 23, 2018

People around the country can pass judgment on their government Nov. 6 in the first national election in two years — if they're registered to vote, and cast a ballot.

Many will not, if recent history holds true. Only about 60 percent of U.S. adult citizens voted in the last national election, a turnout that [ranked in the bottom third](#) of the world's developed countries.



Note: Alaska, represented with statewide data, had 60% turnout.



[https://www.washingtonpost.com/graphics/2018/politics/voter-turnout/?utm\\_term=.ad1e6990074e](https://www.washingtonpost.com/graphics/2018/politics/voter-turnout/?utm_term=.ad1e6990074e)

# Layers of a visualization (from the bottom-up):

## 1. Scaffolding layer

Features that support the content such as legends, scales, axes, etc.

## 2. Encodings layer

The features that represent the data

## 3. Annotation layer

Words on the visualization itself that explains data points or put them in context

## 4. The “me” layer

Where is the reader in the data? Can readers find themselves? Can they create different scenarios? Can they even generate their own data?

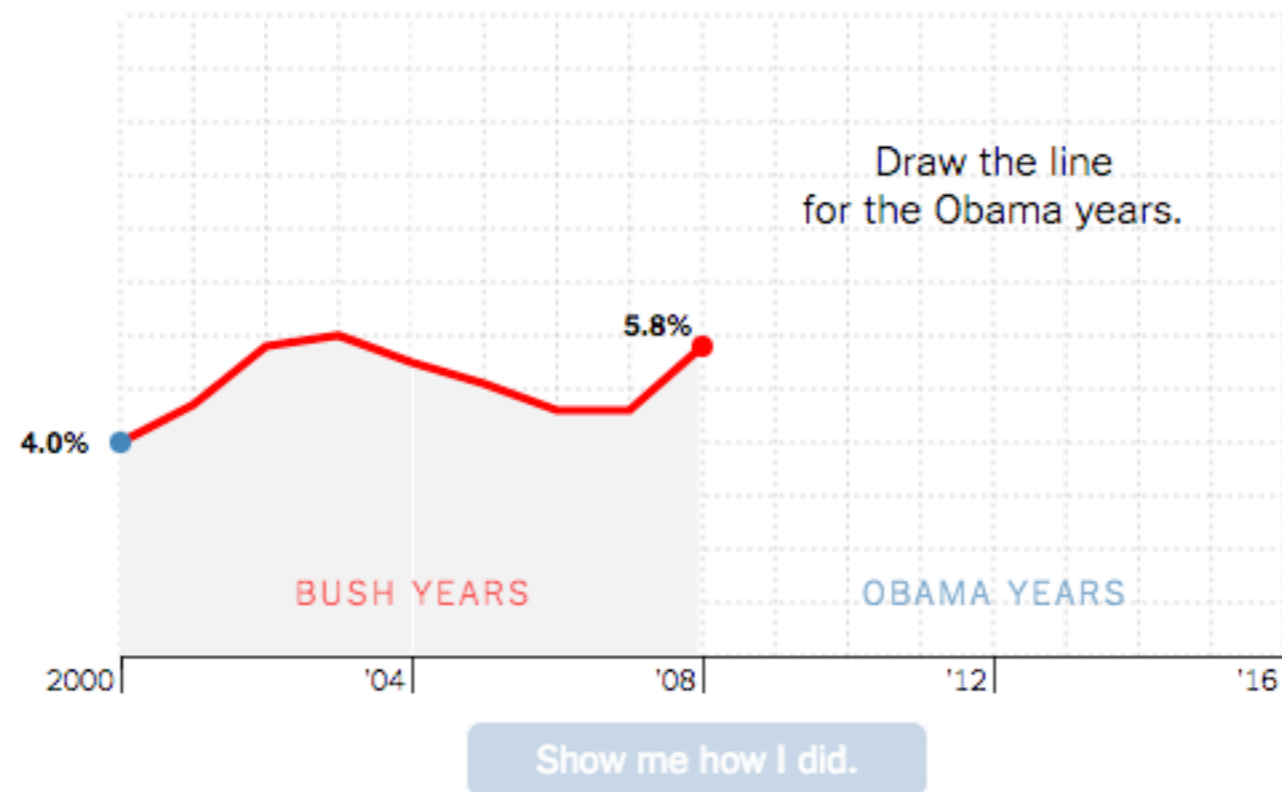
# Challenging readers before showing information

## You Draw It: What Got Better or Worse During Obama's Presidency

By LARRY BUCHANAN, HAEYOUN PARK and ADAM PEARCE JAN. 15, 2017

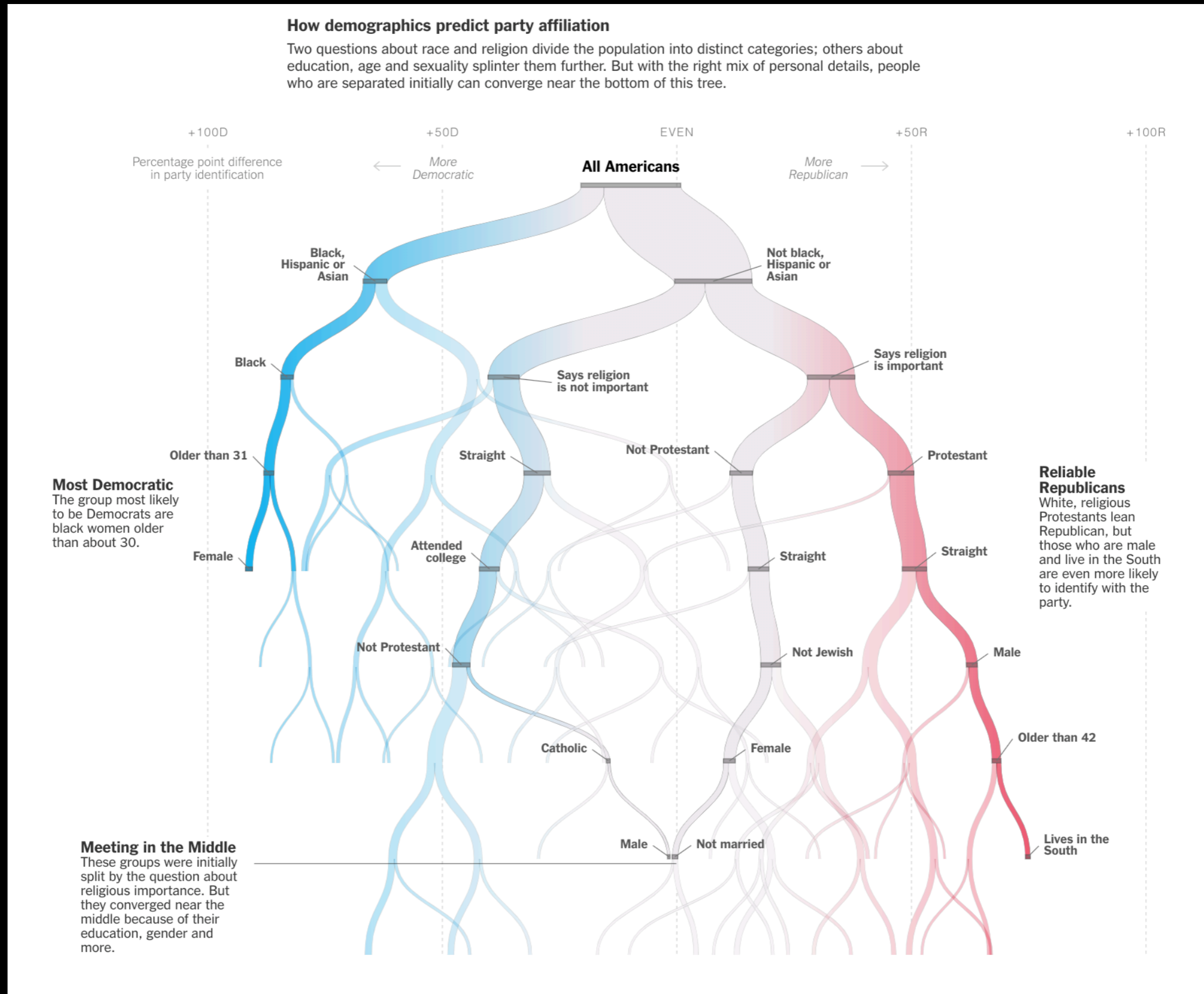
Draw your guesses on the charts below to see if you're as smart as you think you are.

Under President Obama, the **unemployment rate** ...



<https://www.nytimes.com/interactive/2017/01/15/us/politics/you-draw-obama-legacy.html>

# Letting people see themselves in the data, and compare themselves to others



<https://www.nytimes.com/interactive/2019/08/08/opinion/sunday/party-polarization-quiz.html>



# Letting readers create their own scenarios

FiveThirtyEight



SERIES The Gerrymandering Project



PUBLISHED JAN. 25, 2018 AT 6:00 AM

## The Atlas Of Redistricting

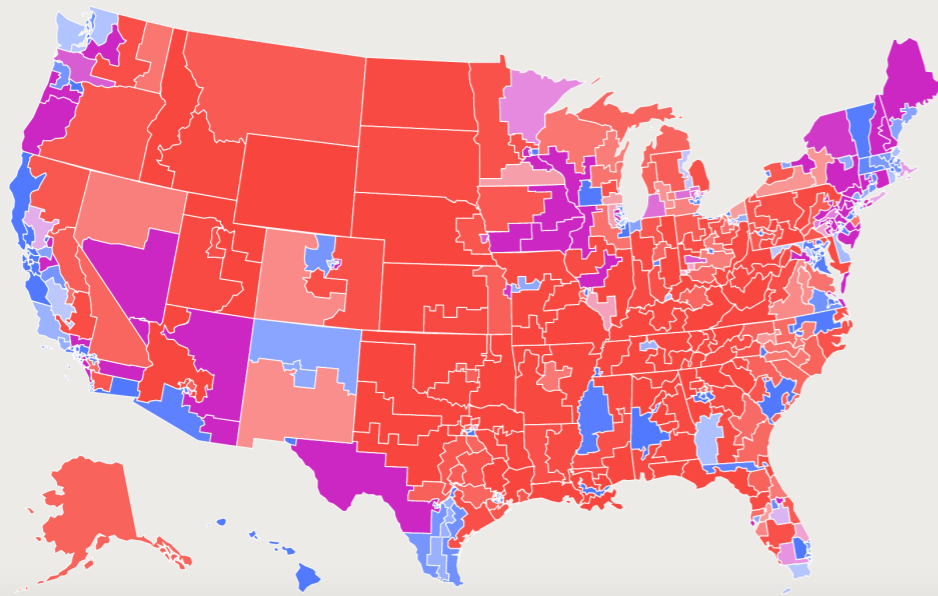
By [Aaron Bycoffe](#), [Ella Koeze](#), [David Wasserman](#) and [Julia Wolfe](#)

- Show current district boundaries
- Gerrymander districts to favor Republicans
- Gerrymander districts to favor Democrats
- Match partisan breakdown of seats to electorate
- Promote highly competitive elections
- Maximize number of majority-minority districts
- Make district shapes compact (using an algorithm)
- Make districts compact while following county borders

### Current congressional district boundaries

How often we'd expect a party to win each of the nation's 435 seats over the long term — not specifically the 2018 midterms — based on historical patterns since 2006

CHANCE OF BEING REPRESENTED BY EITHER PARTY  
100% D 100% R



### The politics of every map

The partisan breakdown of districts and the expected number of seats controlled by Democrats and Republicans, based on their long-term likelihood of winning each district

MAP	USUALLY DEM. DISTRICTS	HIGHLY COMPETITIVE	USUALLY REPUBLICAN	EXPECTED SEAT SPLIT	
				DEM.	GOP
Democratic gerrymander	263	27	145	250.6	184.4
Proportionally partisan	174	82	179	214.0	221.0
Majority minority	169	82	184	209.8	225.2
Highly competitive	94	242	99	209.4	225.6
Compact (borders)	155	99	181	203.9	231.1
Compact (algorithmic)	151	104	180	202.8	232.2
<b>Current</b>	168	72	195	200.6	234.4
Republican gerrymander	139	21	275	171.3	263.7

<https://projects.fivethirtyeight.com/redistricting-maps/>

1. Defining visualization
2. The elements of a visualization
- 3.** How visualization lies
4. Visualization for communication: essential principles

# How Charts Lie



Getting Smarter about  
Visual Information

Alberto Cairo

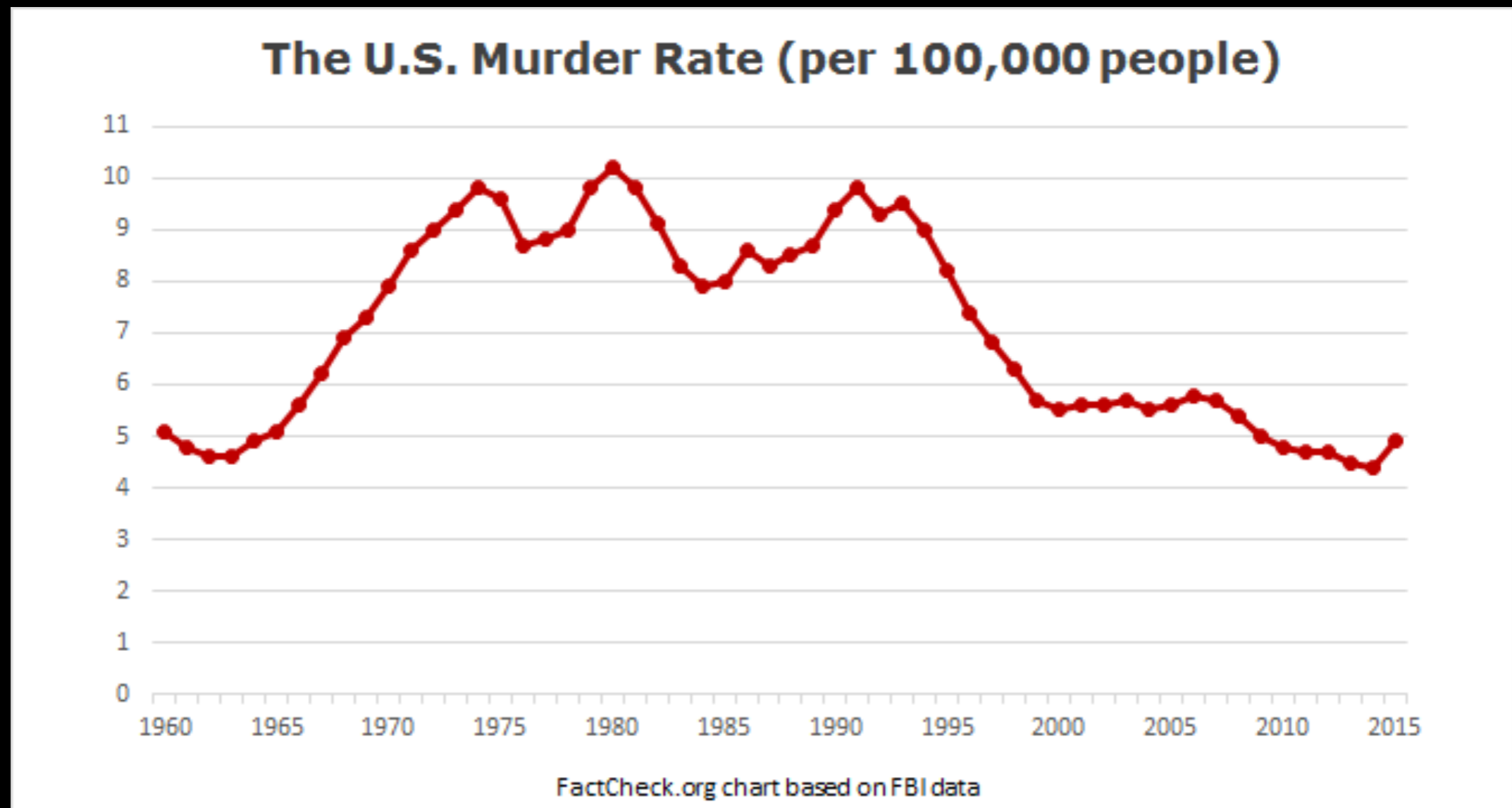
## Common misconceptions

1. “A picture is worth a thousand words”
2. “Visualization is intuitive”
3. “The data should speak for itself”

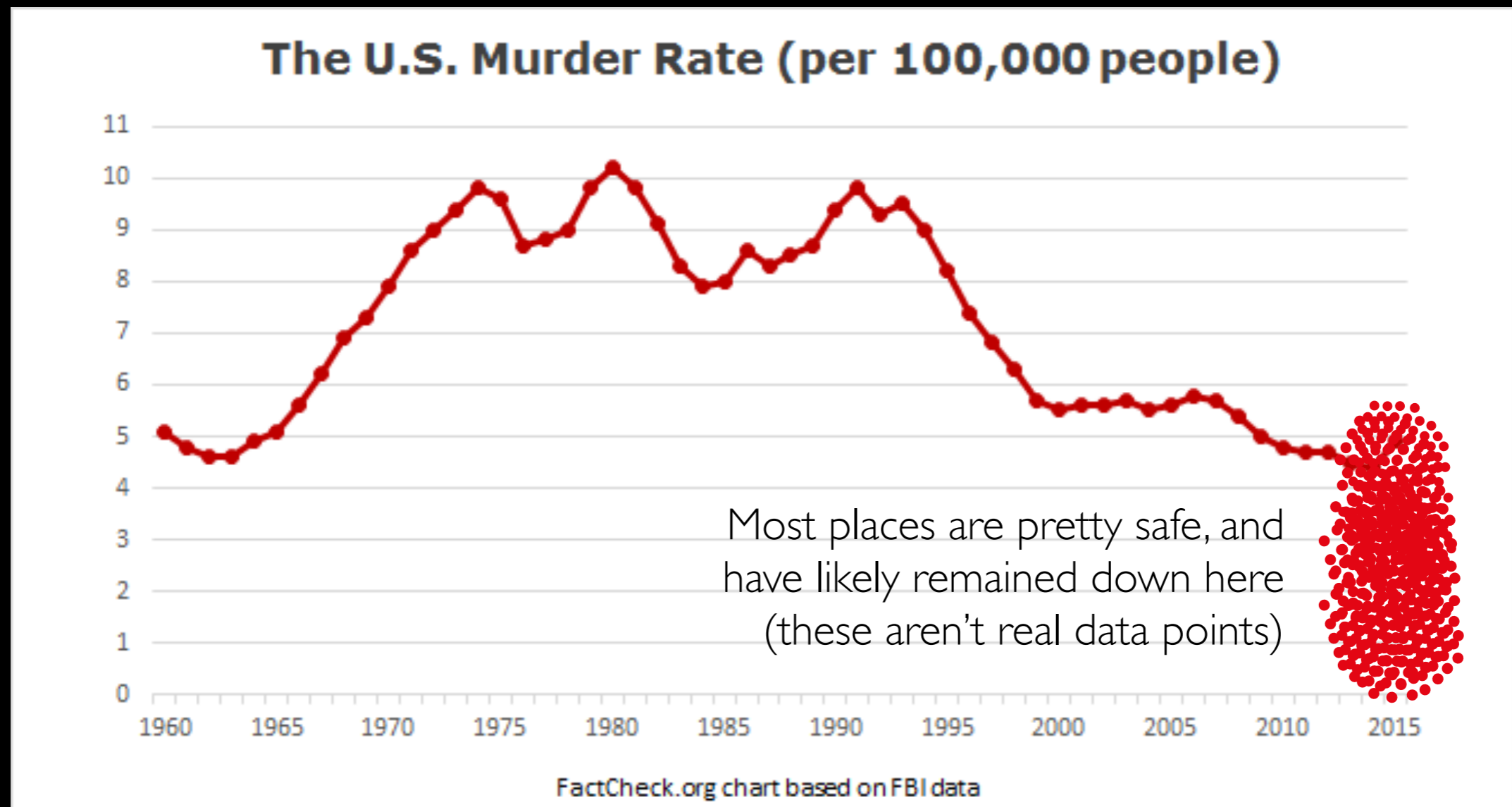
## Common misconceptions

1. “A picture is worth a thousand words”
2. “Visualization is intuitive”
3. “The data should speak for itself”

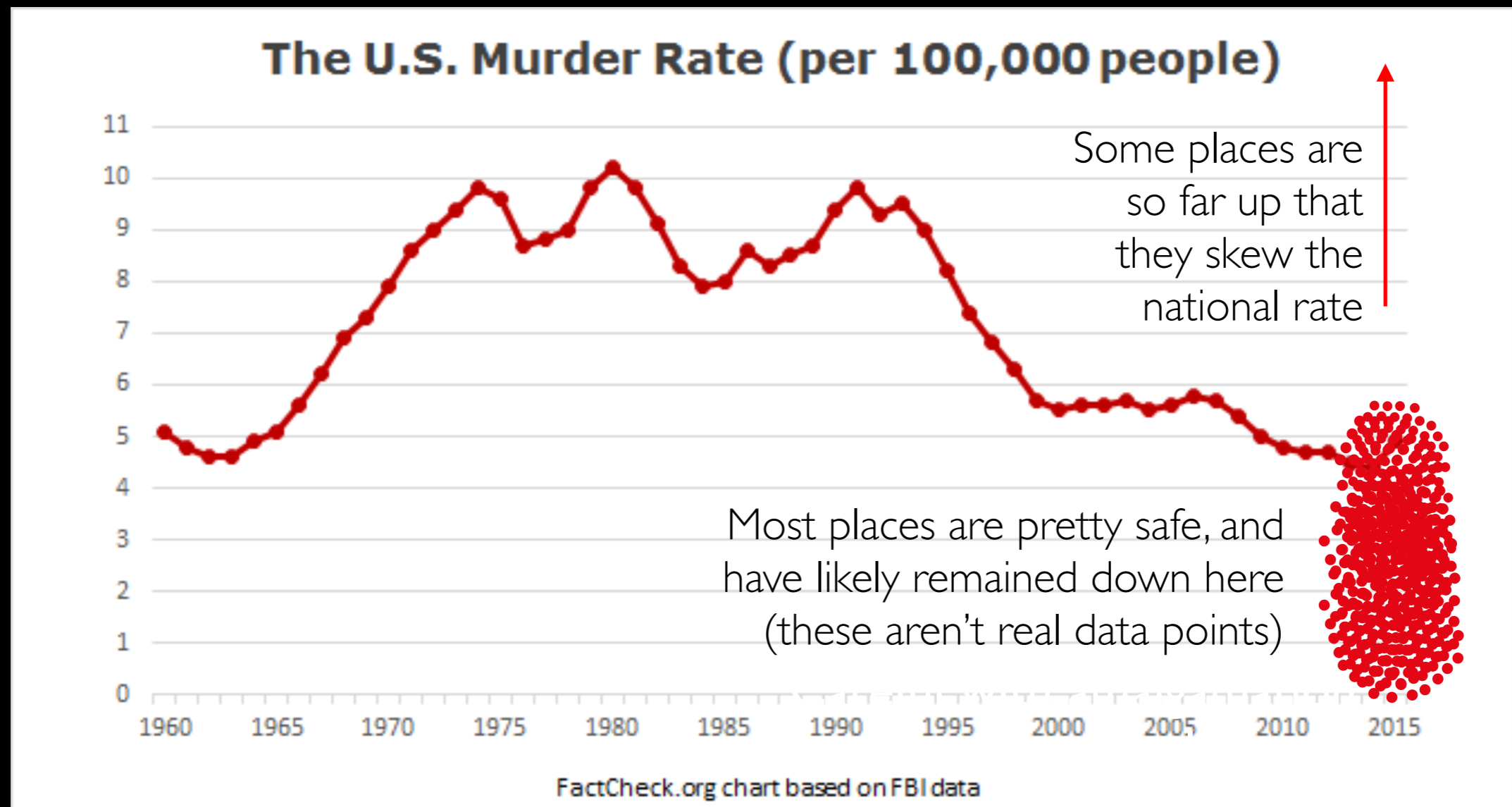
You may think that you're reading this chart correctly, but maybe you aren't



The danger of aggregating data too much, and presenting just averages and other simplistic summaries

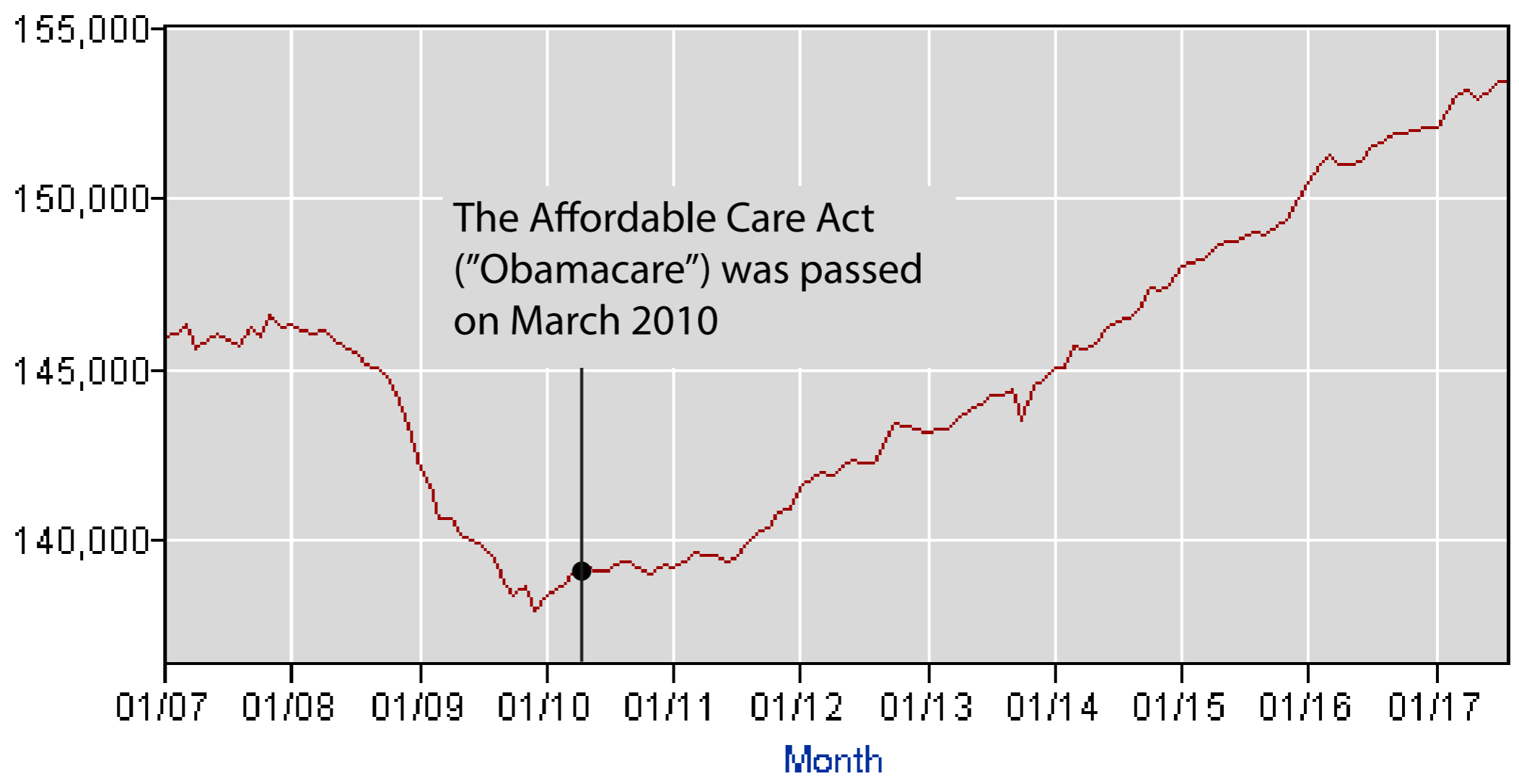


The danger of aggregating data too much, and presenting just averages and other simplistic summaries





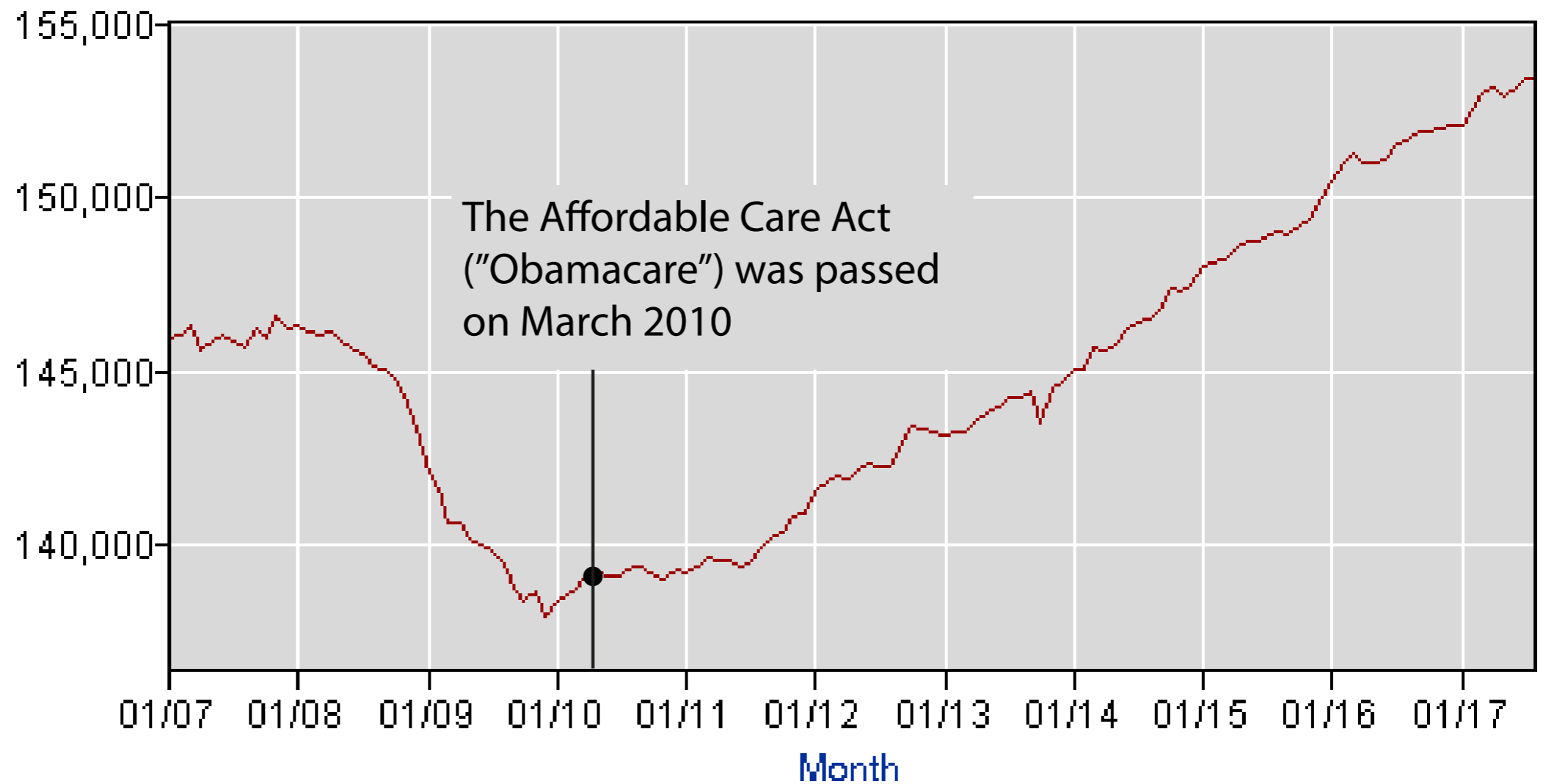
Thousands of people in the workforce



# It's easy to use charts to confirm what we already believe

**Liberal pundit:** "Obamacare isn't bad for the job market"

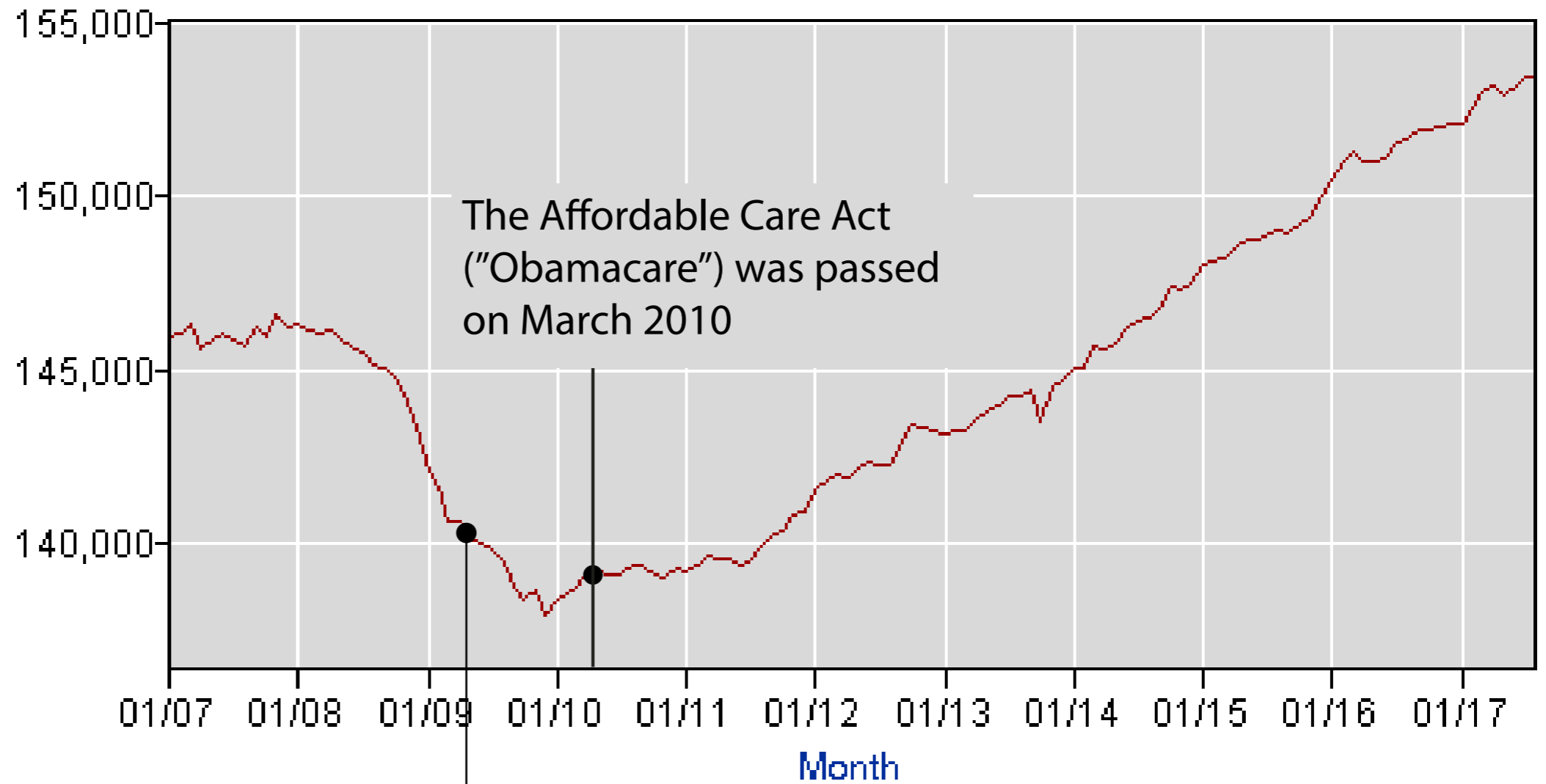
Thousands of people in the workforce



—but a chart shows only what it shows, and nothing else

**Liberal pundit:** “Obamacare isn’t bad for the job market”

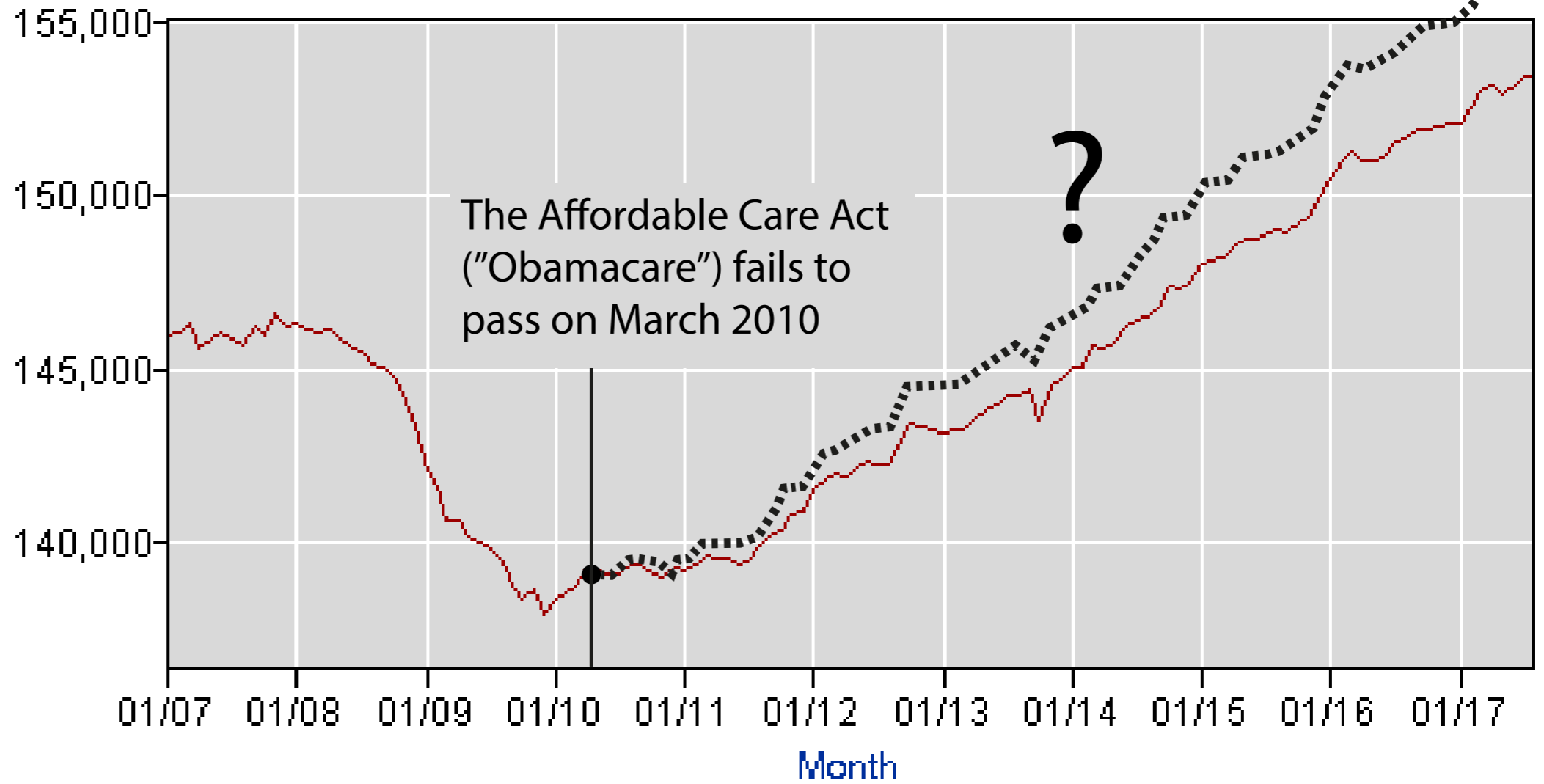
Thousands of  
people in the  
workforce



Stimulus package passed on February 2009

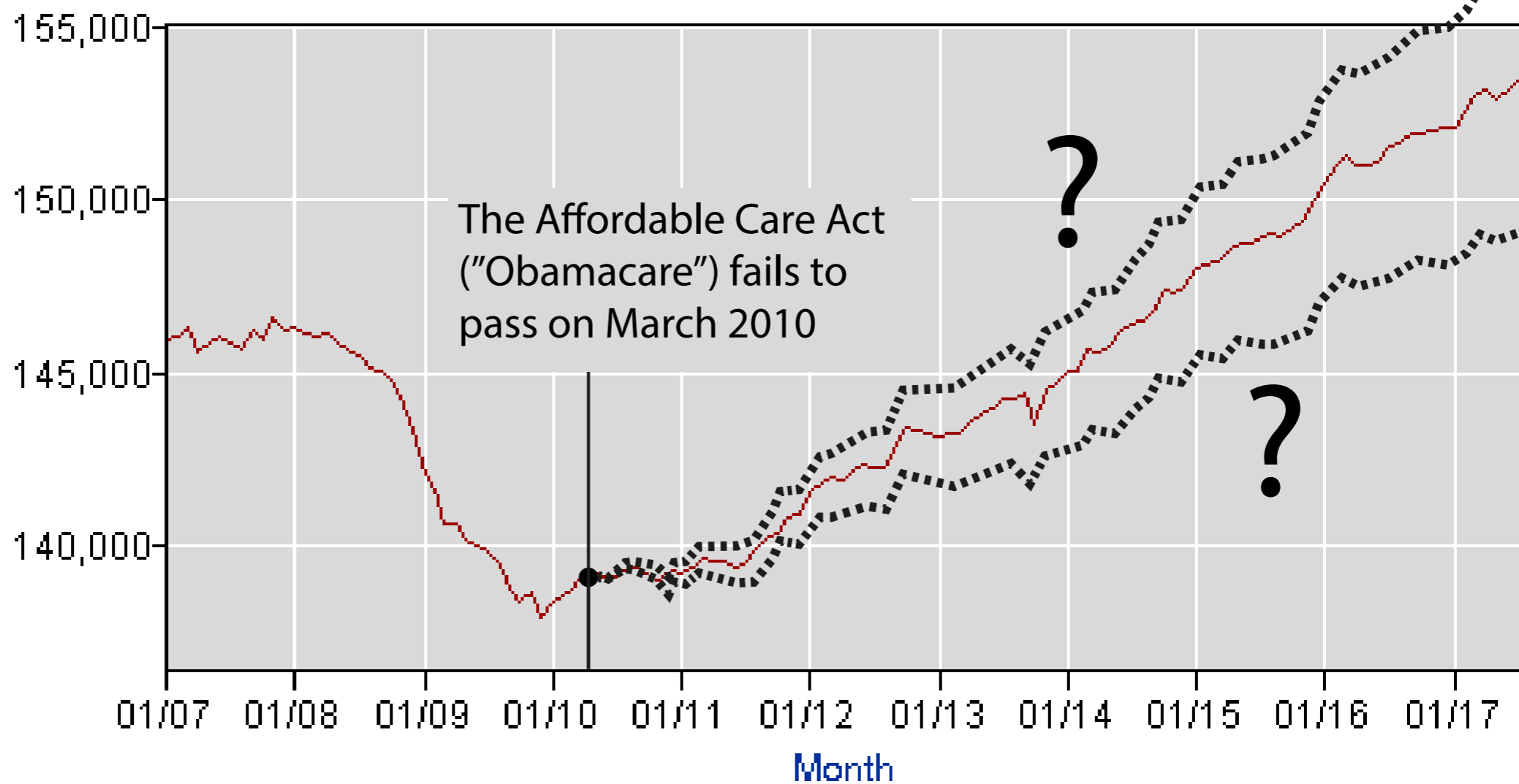
# But what if Obamacare hadn't passed?

Thousands of people in the workforce



## But what if Obamacare hadn't passed?

Thousands of people in the workforce

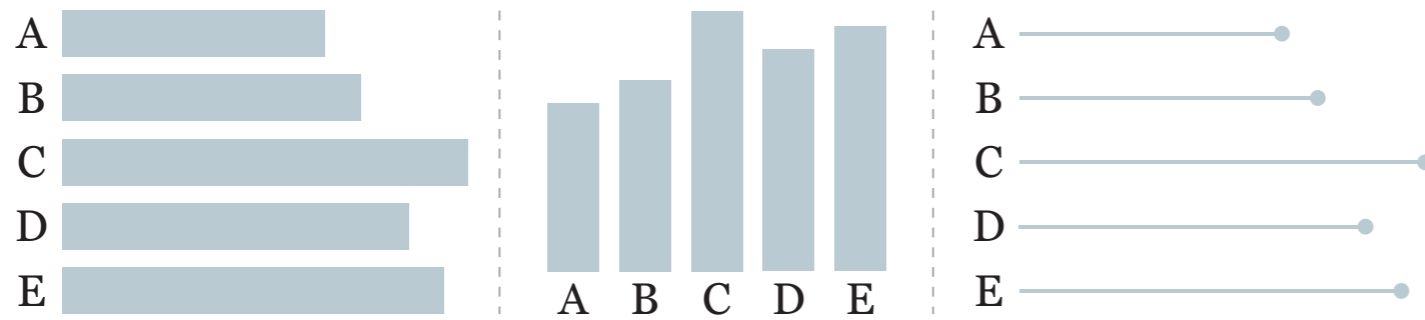


Alberto Cairo • University of Miami • [www.thefunctionalart.com](http://www.thefunctionalart.com) • Twitter: @albertocairo

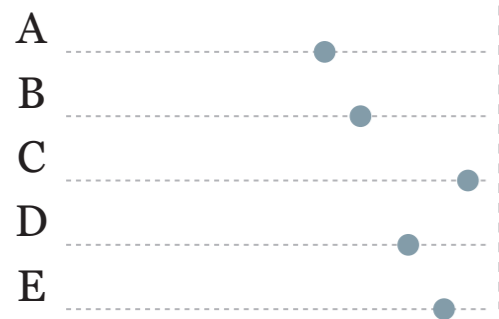
Alberto Cairo • University of Miami • [www.thefunctionalart.com](http://www.thefunctionalart.com) • Twitter: @albertocairo

1. Defining visualization
2. The elements of a visualization
3. How visualization lies
4. Visualization for communication: essential principles

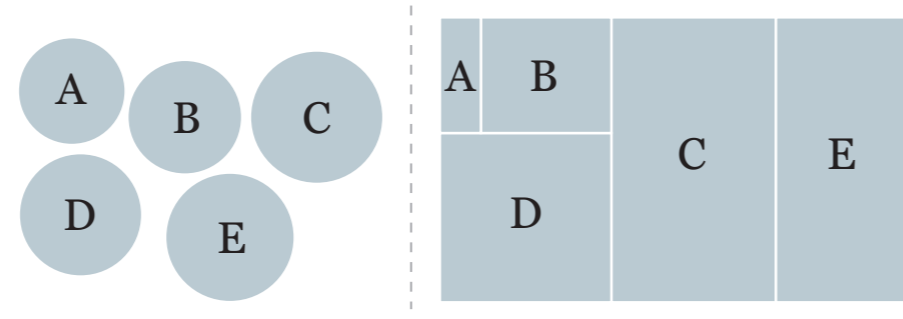
## *Length or height*



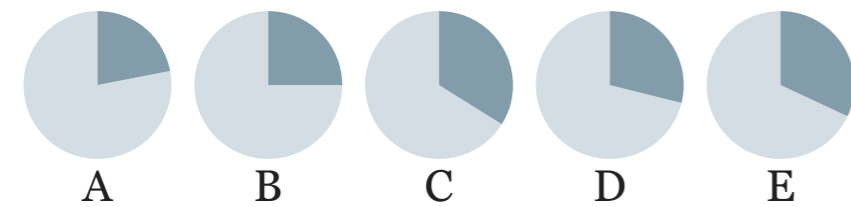
## *Position*



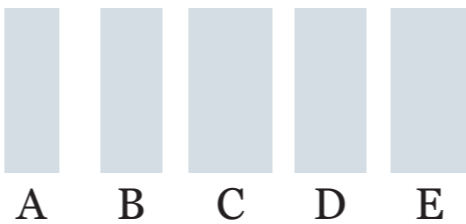
## *Area*



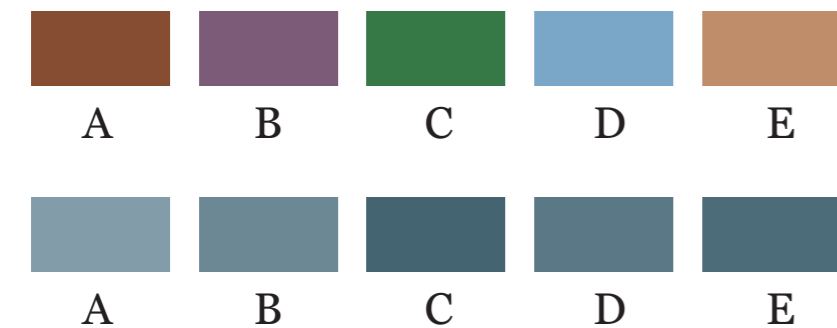
## *Angle/area*



## *Line weight*



## *Hue and shade*



Figures represented  
in all these graphics:  
22%, 25%, 34%, 29%, 32%

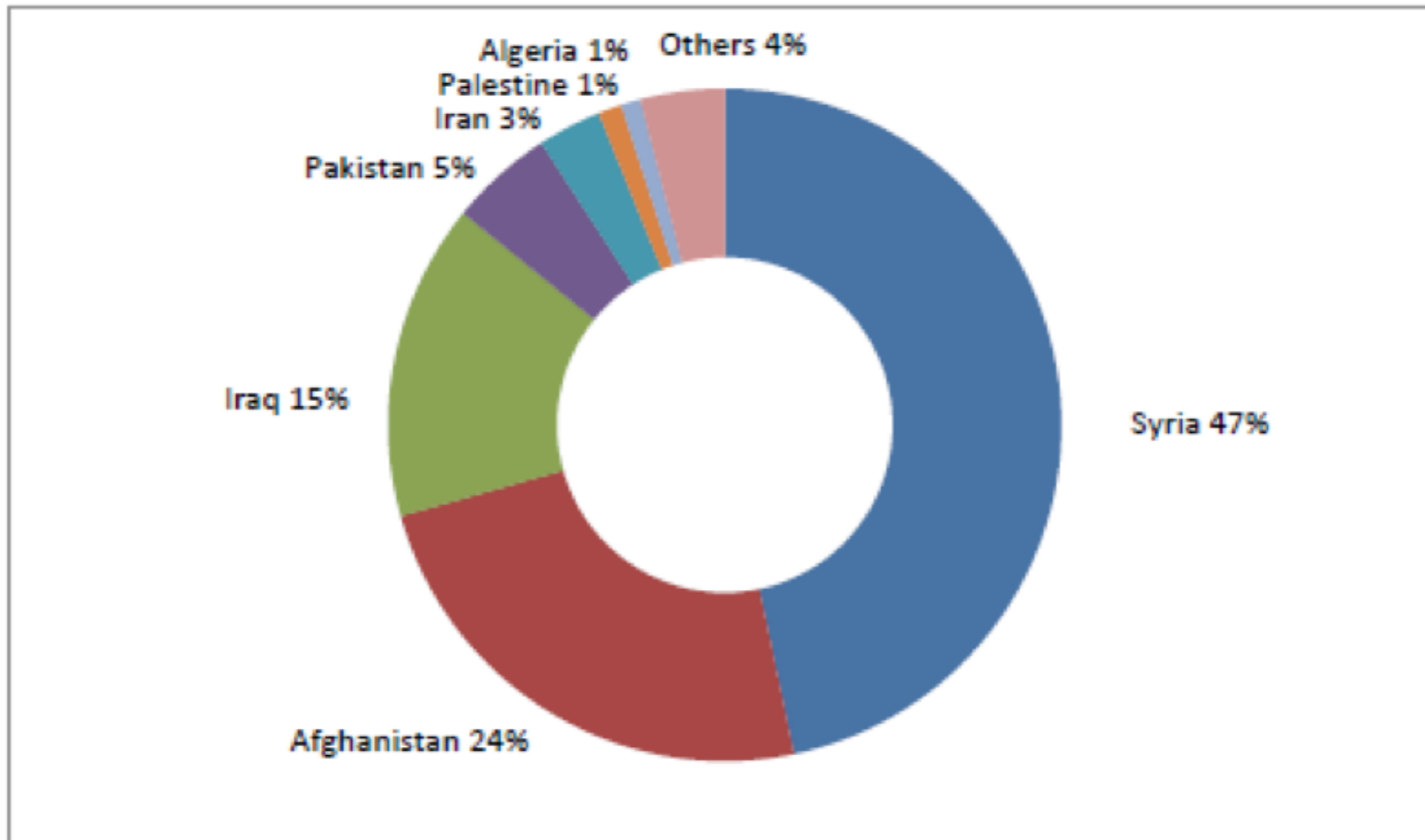
# Encodings

Source: **The Truthful Art**

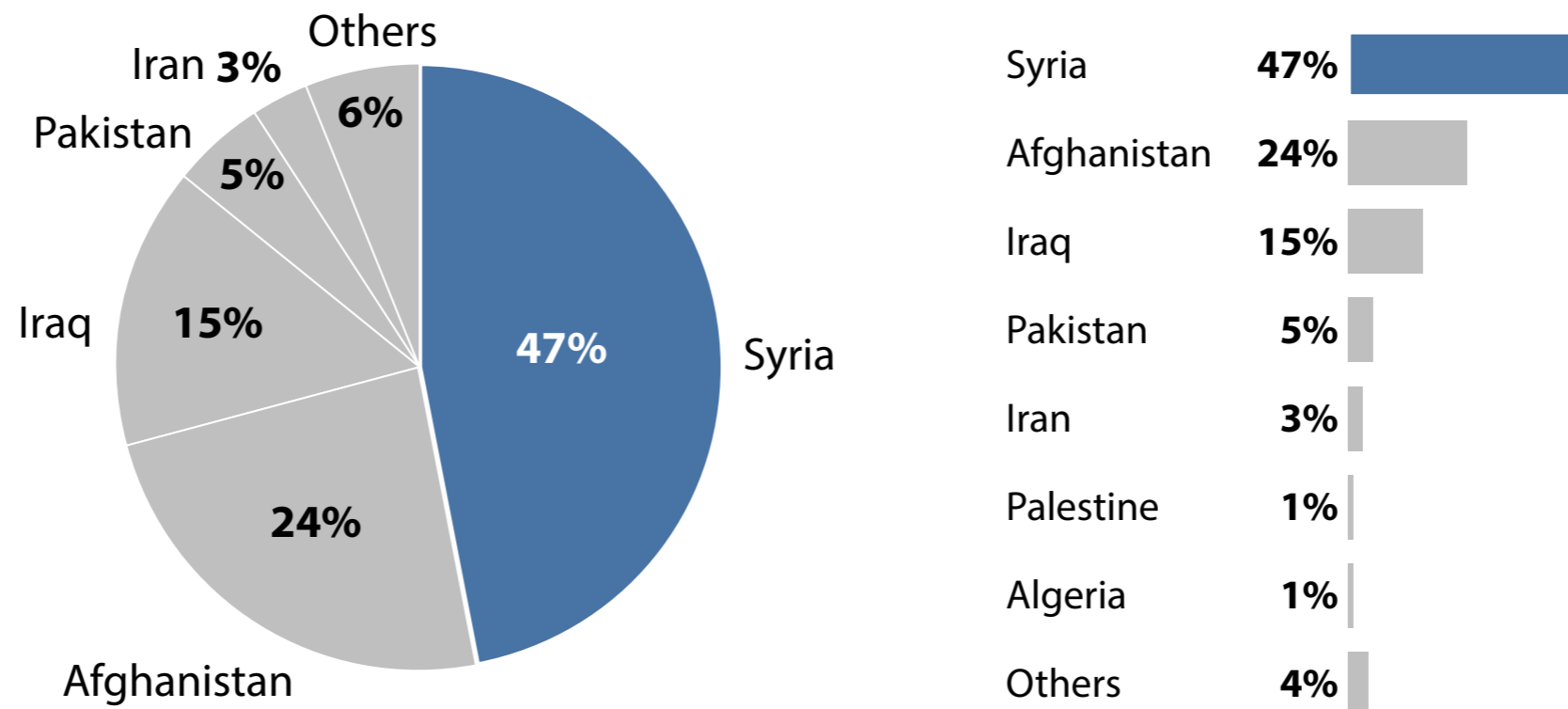
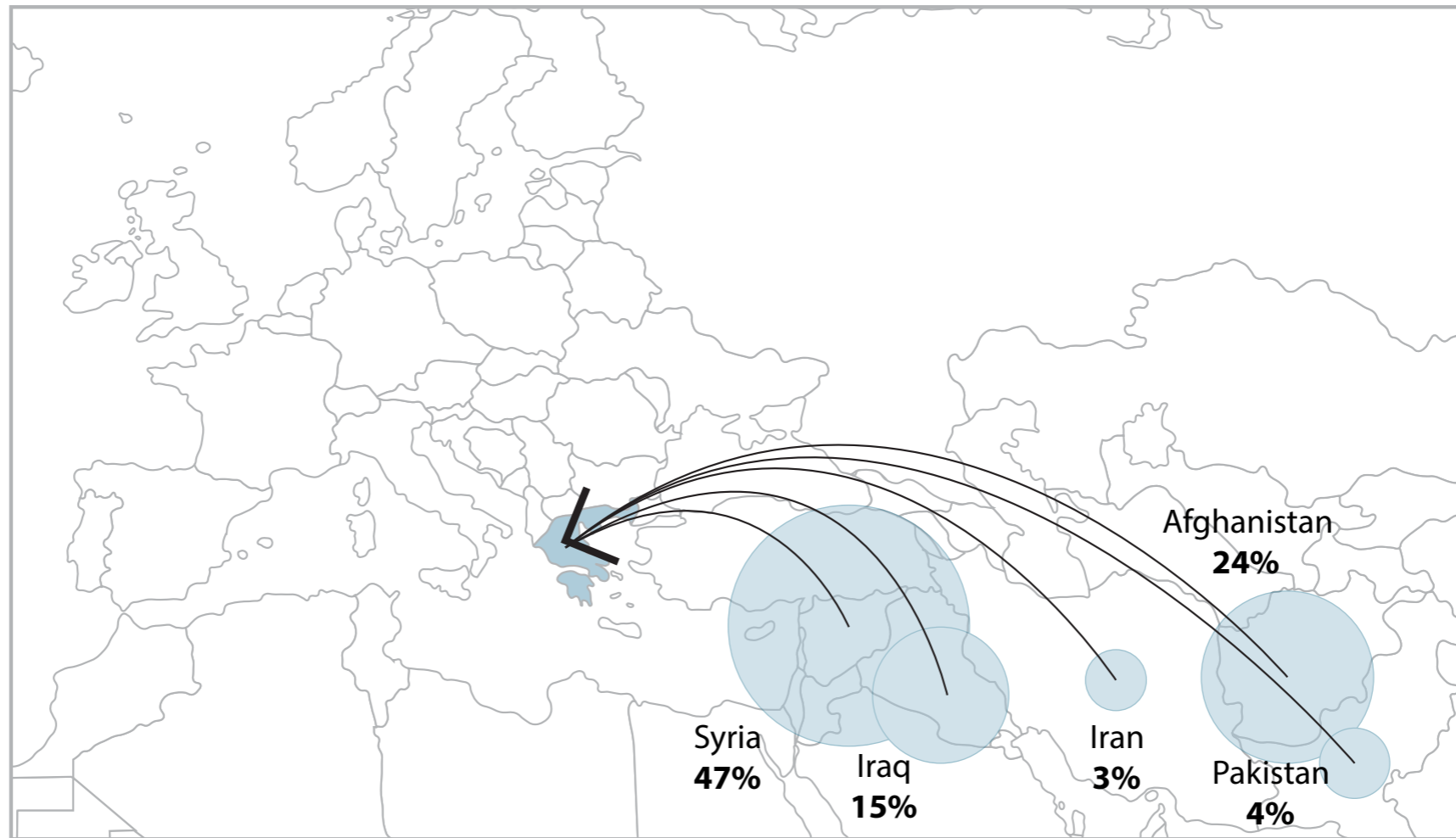
# Always think about the purpose of a graphic

Figure 2 - Main nationalities of arriving migrants – 2016

Greece







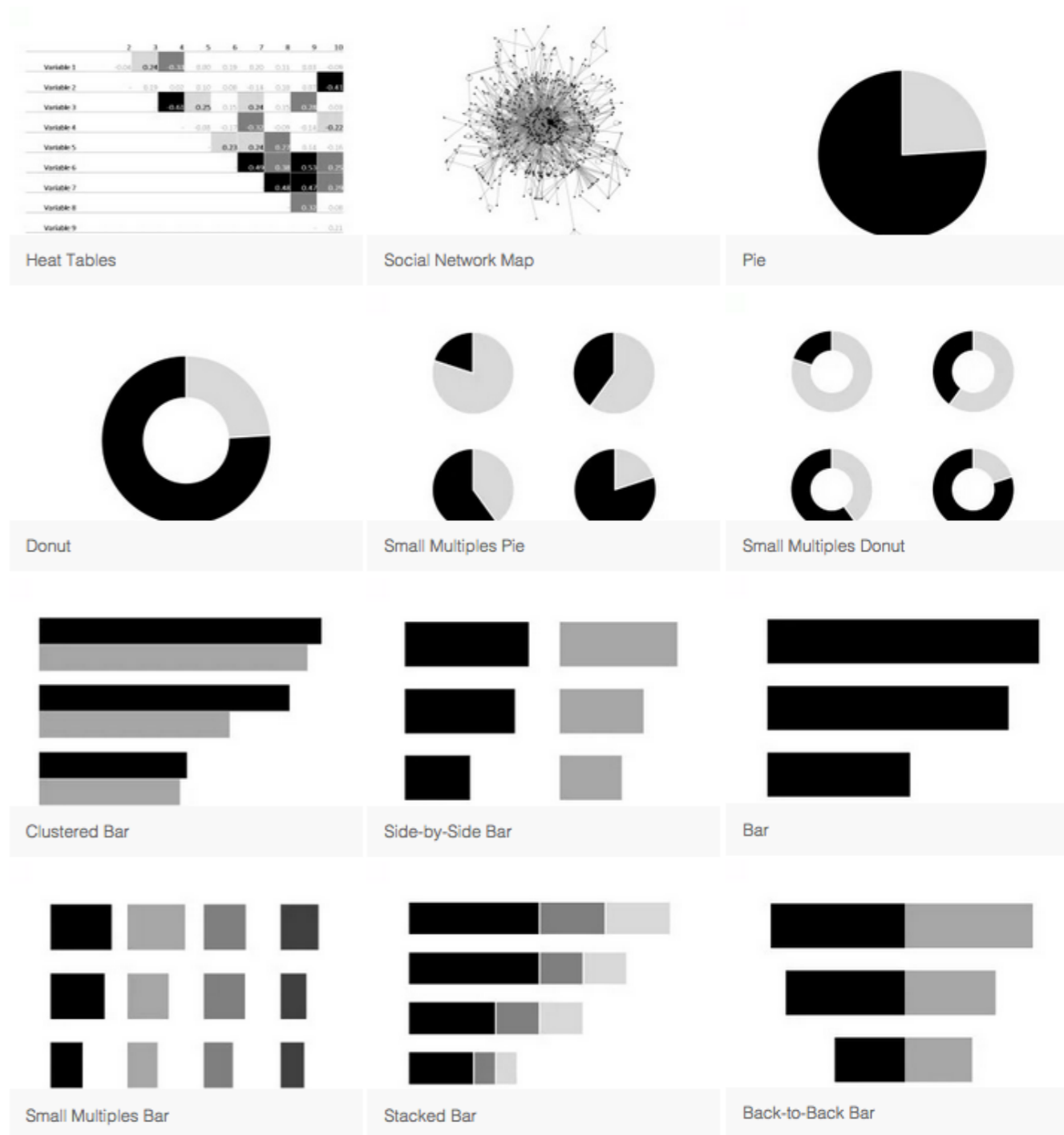
Search by Function

View by List



## EMERY'S ESSENTIALS Chart Choosing Tool

ALL / SMALL MULTIPLES / COMPARING 2 OR MORE CATEGORIES / RANGES OR DISPERSION / PART TO WHOLE / DO-ABLE IN EXCEL / GEOGRAPHIC MAPS / RELATIONSHIPS / COLLAGES / QUALITATIVE / EXPLORATORY / CORRELATION / 1 POINT IN TIME / 2 POINTS IN TIME / 3+ POINTS IN TIME



<http://www.datavizcatalogue.com/>

<http://annkemery.com/essentials/>

<h3>Deviation</h3> <p>Emphasise variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a long-term average. Can also be used to show sentiment (positive/neutral/negative).</p> <p><b>Example FT uses</b> Trade surplus/deficit, climate change</p> <p><b>Diverging bar</b> A simple standard bar chart that can handle both negative and positive magnitude values.</p> <p><b>Diverging stacked bar</b> Perfect for presenting survey results which involve sentiment (eg disagree/neutral/agree).</p> <p><b>Spline chart</b> Splits a single value into 2 contrasting components (eg Male/female).</p> <p><b>Surplus/deficit filled line</b> The shaded area of these charts allows a balance to be shown either against a baseline or between two series.</p>	<h3>Correlation</h3> <p>Show the relationship between two or more variables. Be mindful that unless you tell them otherwise, many readers will assume the relationships you show them to be causal (i.e. one causes the other).</p> <p><b>Example FT uses</b> Inflation &amp; unemployment, income &amp; life expectancy</p> <p><b>Scatterplot</b> The standard way to show the relationship between two continuous variables, each of which has its own axis.</p> <p><b>Line + Column</b> A good way of showing the relationship between an amount (columns) and a rate (line).</p> <p><b>Connected scatterplot</b> Usually used to show how the relationship between 2 variables has changed over time.</p> <p><b>Bubble</b> Like a scatterplot, but adds additional detail by sizing the circles according to a third variable.</p> <p><b>XY heatmap</b> A good way of showing the patterns between 2 categories of data, less good at showing the differences in amounts.</p>	<h3>Ranking</h3> <p>Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.</p> <p><b>Example FT uses</b> Wealth, deprivation, league tables, constituency election results</p> <p><b>Ordered bar</b> Standard bar charts display the ranks of values much more easily when sorted into order.</p> <p><b>Ordered column</b> See above.</p> <p><b>Ordered proportional symbol</b> Use when there are big variations between values and/or seeing fine differences between data is not so important.</p> <p><b>Dot strip plot</b> Dots placed in order on a strip are a space-efficient method of laying out ranks across multiple categories.</p> <p><b>Slope</b> Perfect for showing how ranks have changed over time or vary between categories.</p> <p><b>Lollipop chart</b> Lollipops draw more attention to the data value than standard bar/column and can also show rank and value effectively.</p>	<h3>Distribution</h3> <p>Show values in a dataset and how often they occur. The shape (or 'look') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.</p> <p><b>Example FT uses</b> Income distribution, population (age/sex) distribution</p> <p><b>Histogram</b> The standard way to show a statistical distribution - keep the gaps between columns small to highlight the 'shape' of the data.</p> <p><b>Boxplot</b> Summarise multiple distributions by showing the median (centre) and range of the data.</p> <p><b>Violin plot</b> Similar to a box plot but allows for complex distributions (data that cannot be summarised with simple average).</p> <p><b>Population pyramid</b> A standard way for showing the age and sex breakdown of a population distribution; effectively, back to back histograms.</p> <p><b>Dot strip plot</b> Good for showing individual values in a distribution, can be a problem when too many dots have the same value.</p> <p><b>Dot plot</b> A simple way of showing the change or range (min/max) of data across multiple categories.</p> <p><b>Barcode plot</b> Like dot strip plots, good for displaying all the data in a table; they work best when highlighting individual values.</p> <p><b>Cumulative curve</b> A good way of showing how unequal a distribution is; y axis is always cumulative frequency, x axis is always a measure.</p>	<h3>Change over Time</h3> <p>Give emphasis to changing trends. These can be short (Out for 'Year') movements or extended series (travelling decades or centuries). Choosing the correct time period is important to provide suitable context for the reader.</p> <p><b>Example FT uses</b> Share price movements, economic time series</p> <p><b>Line</b> The standard way to show a changing time series. If data are irregular, consider markers to represent data points.</p> <p><b>Column</b> Columns work well for showing change over time - but usually best with only one series of data at a time.</p> <p><b>Line + column</b> A good way of showing the relationship over time between an amount (columns) and a rate (line).</p> <p><b>Stack price</b> Usually focused on day-to-day activity, these charts show opening/closing and hi/low points of each day.</p> <p><b>Slope</b> Good for showing changing data as long as the data can be simplified into 2 or 3 points without missing a key part of story.</p> <p><b>Area chart</b> Use with care - these are good at showing changes to total, but seeing change in components can be very difficult.</p> <p><b>Fan chart (projections)</b> Use to show the uncertainty in future projections - usually this grows the further forward to projection.</p> <p><b>Connected scatterplot</b> A good way of showing changing data for two variables wherever there is a relatively clear pattern of progression.</p> <p><b>Calendar heatmap</b> A great way of showing temporal patterns (daily, weekly, monthly) - at the expense of showing precision in quantity.</p> <p><b>Priestley timeline</b> Great when date and duration are key elements of the story in the data.</p> <p><b>Circle timeline</b> Good for showing discrete values of varying size across multiple categories (eg earthquakes by continent).</p> <p><b>Seismogram</b> Another alternative to the circle timeline for showing series where there are big variations in the data.</p>	<h3>Part-to-whole</h3> <p>Show how a single entity can be broken down into its component elements. If the reader's interest is solely in the size of the components, consider a magnitude-type chart instead.</p> <p><b>Example FT uses</b> Fiscal budgets, company structures, national election results</p> <p><b>Stacked column</b> A simple way of showing part-to-whole relationships but can be difficult to read with more than a few components.</p> <p><b>Proportional stacked bar</b> A good way of showing the size and proportion of data at the same time - as long as the data are not too complicated.</p> <p><b>Pie</b> A common way of showing part-to-whole data - but be aware that it's difficult to accurately compare the size of the segments.</p> <p><b>Donut</b> Similar to a pie chart - but the centre can be a good way of making space to include more information about the data (eg total).</p> <p><b>Treemap</b> Use for hierarchical part-to-whole relationships; can be difficult to read when there are many small segments.</p> <p><b>Voronoi</b> A way of turning points into areas - any point within each area is closer to the central point than any other centroid.</p> <p><b>Sunburst</b> Another way of visualising hierarchical part-to-whole relationships. Use sparingly (if at all) for obvious reasons.</p> <p><b>Arc</b> A hemicycle, often used for visualising political results in parliaments.</p> <p><b>Gridplot</b> Good for showing % information, they work best when used on whole numbers and work well in multiple layout form.</p> <p><b>Venn</b> Generally only used for schematic representation.</p> <p><b>Waterfall</b> Can be useful for showing part-to-whole relationships where some of the components are negative.</p>	<h3>Magnitude</h3> <p>Show size comparisons. These can be relative (just being able to see larger/smaller) or absolute (need to see fine differences). Usually these show a 'counted' number (for example, barrels, dollars or people) rather than a calculated rate or per cent.</p> <p><b>Example FT uses</b> Commodity production, market capitalisation</p> <p><b>Column</b> The standard way to compare the size of things. Must always start at 0 on the axis.</p> <p><b>Bar</b> See above. Good when the data are not time series and labels have long category names.</p> <p><b>Paired column</b> As per standard column but allows for multiple series. Can become more than 2 series.</p> <p><b>Paired bar</b> See above.</p> <p><b>Proportional stacked bar</b> A good way of showing the size and proportion of data at the same time - as long as the data are not too complicated.</p> <p><b>Proportional symbol</b> Use when there are big variations between values and/or seeing fine differences between data is not so important.</p> <p><b>Isotype (pictogram)</b> Excellent solution in some instances - use only with whole numbers (do not slice off an arm to represent a decimal).</p> <p><b>Lollipop chart</b> Lollipop charts draw more attention to the data value than standard bar/column - does not HAVE to start at zero (but preferable).</p> <p><b>Radar chart</b> A space-efficient way of showing value of multiple variables - but make sure they are organised in a way that makes sense to reader.</p> <p><b>Parallel coordinates</b> An alternative to radar charts - again, the arrangement of the variables is important. Usually benefits from highlighting values.</p>	<h3>Spatial</h3> <p>Used only when precise locations or geographical patterns in data are more important to the reader than anything else.</p> <p><b>Example FT uses</b> Location risks, population density, natural resource locations, natural disaster risk/impact, catchment areas, variation in election results</p> <p><b>Basic choropleth (rate/ratio)</b> The standard approach for putting data on a map - should always be rates rather than totals and use a sensible base geography.</p> <p><b>Proportional symbol (count/magnitude)</b> Use for totals rather than rates - be wary that small differences in data will be hard to see.</p> <p><b>Flow map</b> For showing unambiguous movement across a map.</p> <p><b>Contour map</b> For showing areas of equal value on a map. Can use deviation colour schemes for showing +/- values.</p> <p><b>Equalised cartogram</b> Converting each unit on a map to a regular and equally-sized shape - good for representing strong regions with equal value.</p> <p><b>Scaled cartogram (value)</b> Stretching and shrinking a map so that each area is sized according to a particular value.</p> <p><b>Dot density</b> Used to show the location of individual events/locations - make sure to annotate any patterns the reader should see.</p> <p><b>Heat map</b> Grid-based data values mapped with an intensity colour scale. As choropleth map - but not snapped to an administrative unit.</p>	<h3>Flow</h3> <p>Show the reader volumes or intensity of movement between two or more states or conditions. These might be logical sequences or geographical locations.</p> <p><b>Example FT uses</b> Movement of funds, trade, migrants, lawsuits, information relationship graphs.</p> <p><b>Sankey</b> Shows changes in flows from one condition to at least one other, good for tracing the eventual outcome of a complex process.</p> <p><b>Waterfall</b> Designed to show the sequencing of data through a flow process, typically budgets. Can include +/- components.</p> <p><b>Chord</b> A complex but powerful diagram which can illustrate 2-way flows (and net winner) in a matrix.</p> <p><b>Network</b> Used for showing the strength and inter-connectedness of relationships of varying types.</p>
---	---	---	---	---	--	--	--	--

# Visual vocabulary

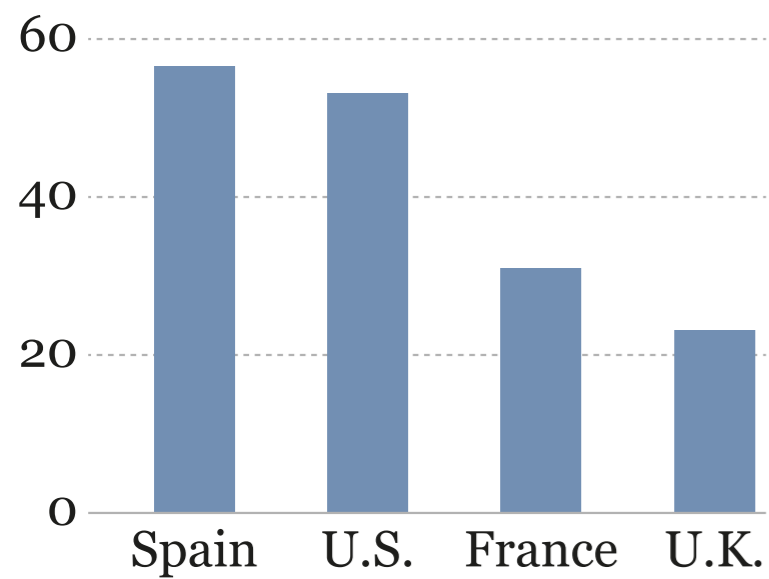
## Designing with data

There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

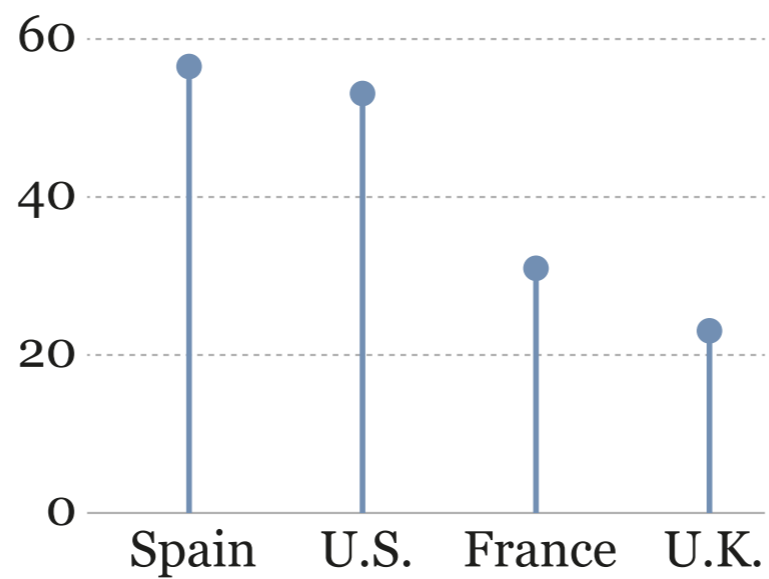
FT graphic: Alan Smith, Chris Campbell, Ian Beth, Liz Favre, Graham Parrish, Billy Ehrenberg, Paul McCallum, Martin Stabe  
Inspired by the Graphic Continuum by Jon Schwabish and Severino Ribezza

[ft.com/vocabulary](https://ft.com/vocabulary)

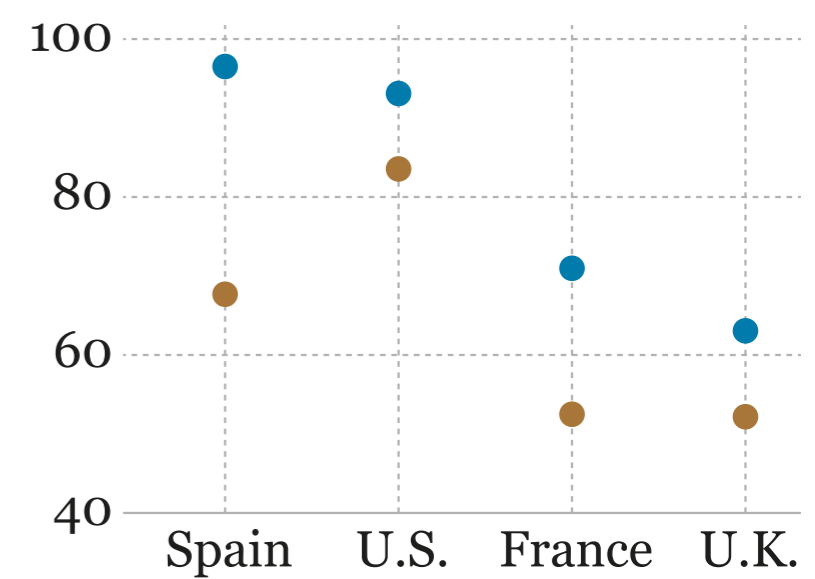




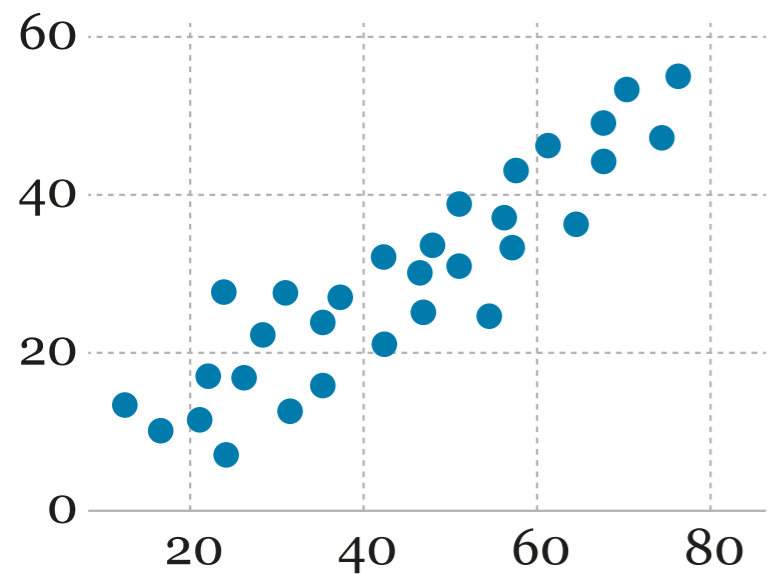
**BAR CHART**



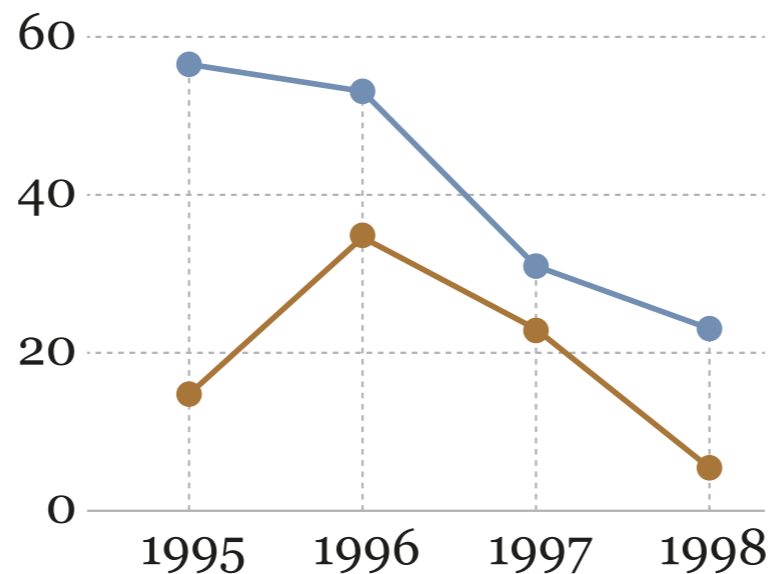
**LOLLIPOP CHART**



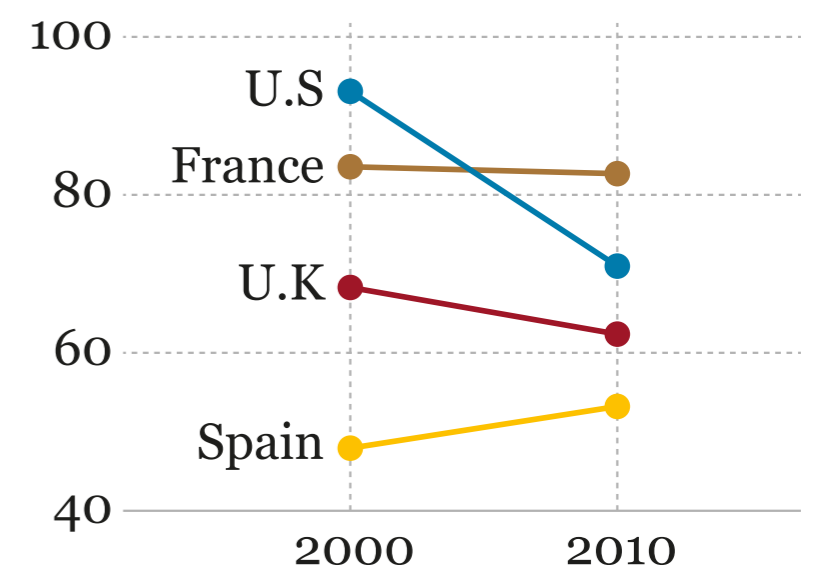
**DOT CHART  
(or dot plot)**



**SCATTER CHART  
(or scatter plot)**



**LINE CHART  
(or time series chart)**



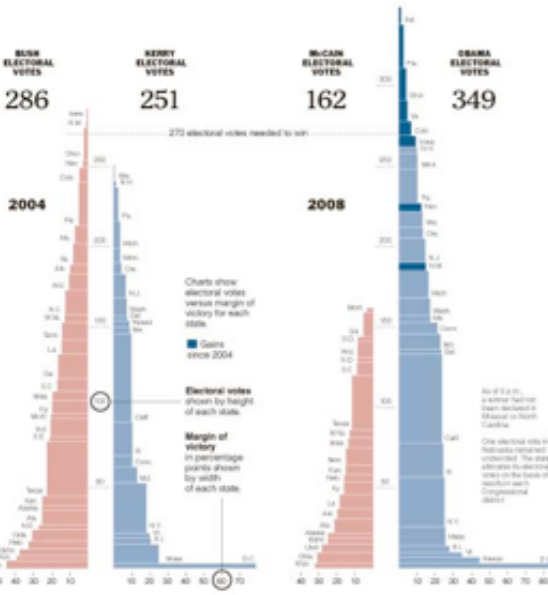
**SLOPE CHART**

Height, length, position are appropriate encodings to enable accurate judgments: Comparisons, relationships, change over time, etc.

# In a Decisive Victory, Obama Reshapes the Electoral Map

Barack Obama's historic win, with at least 349 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, especially in more industrial states, and high numbers of new voters helped flip key areas from red to blue. Even where Mr. McCain beat Mr. Obama, he won by slimmer margins, as much of the electorate — across age, race and income lines — swung toward the Democratic Party.

By Erin Aigner, Joe Burgess, Baden Copeland, Matthew Ericson, Hannah Fairfield, Ford Fausst, Hayoun Park and Archie Tate



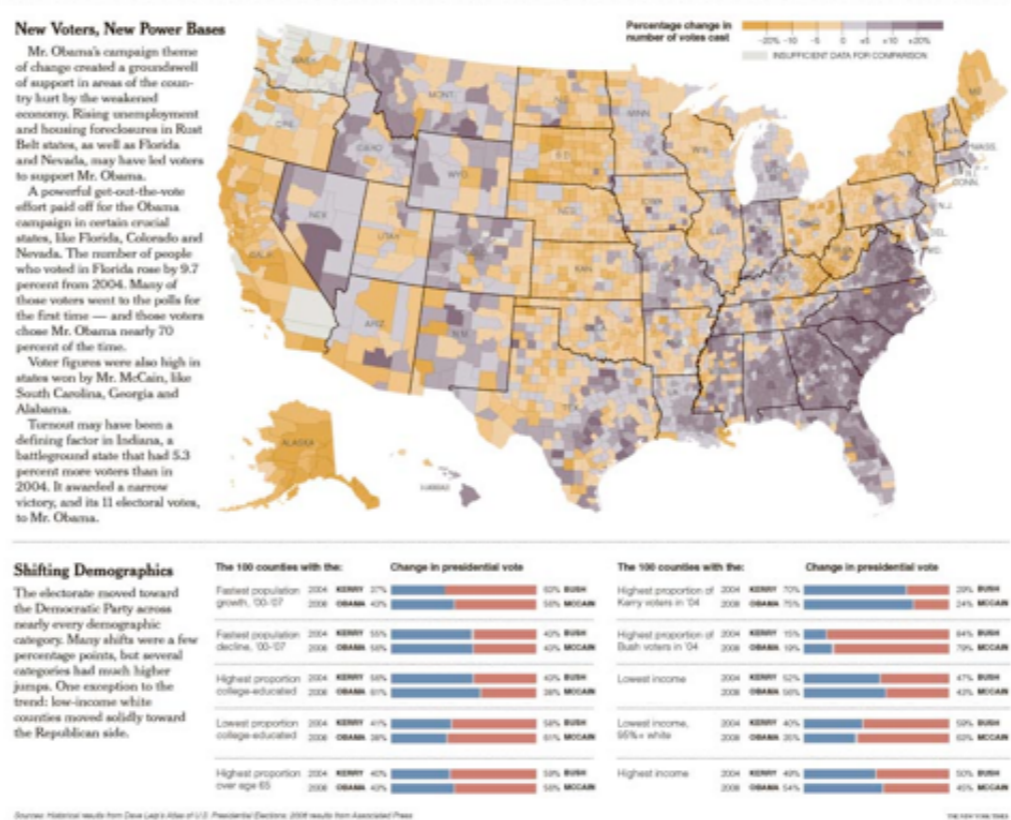
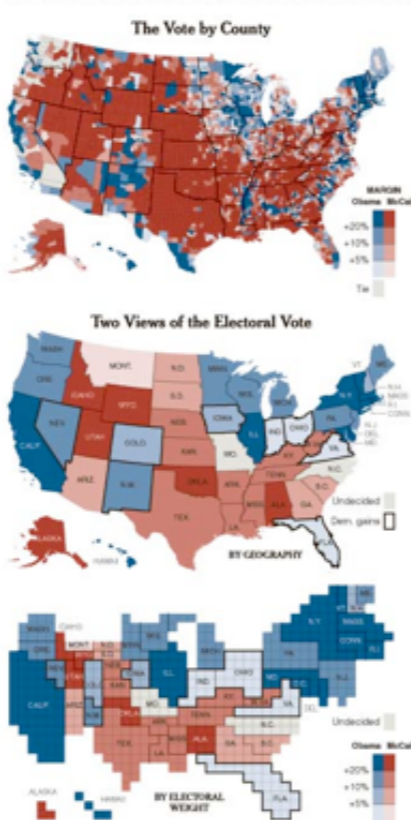
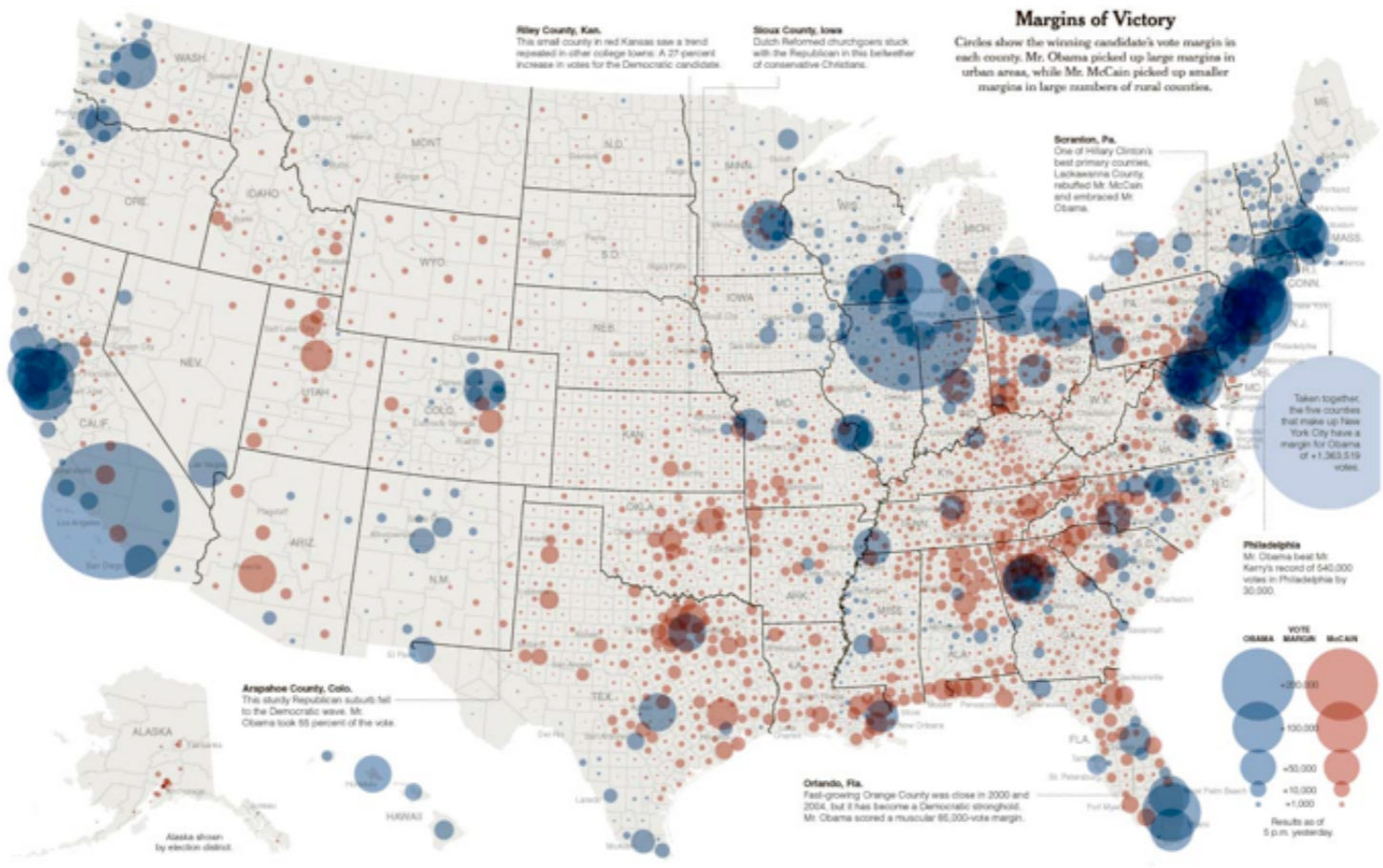
States won by Obama	349							
Hawaii	100%	3,068,621	72%	1,112,848	27%	+65	+3	4
Indiana	100%	1,367,304	50%	1,341,921	49%	+1	+21	11
Delaware	100%	247,366	61%	151,667	36%	+24	+8	3
New Mexico	100%	452,038	57%	332,979	42%	+18	+1	5
Montana	91%	201,899	67%	95,422	32%	+55	+30	3
Nevada	100%	531,884	55%	411,989	43%	+12	+3	3
Illinois	100%	3,593,343	62%	1,975,801	37%	+25	+10	21
California	37%	6,179,542	61%	3,736,444	37%	+24	+10	55
Michigan	100%	2,967,993	57%	2,544,422	47%	+15	+1	17
Virginia	99%	1,792,522	52%	1,637,327	47%	+4	+8	13
Wisconsin	100%	1,633,117	54%	1,251,853	45%	+13	+1	10
Colorado	92%	1,108,328	53%	966,957	46%	+7	+3	9
Connecticut	98%	343,819	60%	306,268	39%	+22	+10	7
Maryland	99%	1,408,150	61%	873,320	36%	+23	+13	10
Iowa	100%	918,242	54%	677,308	45%	+9	+10	7
Washington	58%	666,228	57%	690,352	41%	+16	+7	11
Oregon	70%	666,695	58%	532,376	42%	+11	+4	7
New Hampshire	92%	355,301	55%	290,344	45%	+18	+1	4
Maine	92%	390,147	58%	271,876	40%	+18	+9	4
New Jersey	100%	2,073,934	57%	1,540,907	42%	+18	+7	15
Pennsylvania	100%	3,184,807	55%	2,584,119	44%	+12	+3	21
Rhode Island	98%	275,228	63%	152,197	39%	+28	+21	4
Florida	100%	4,103,638	51%	3,908,736	46%	+2	+3	27
New York	99%	4,357,360	62%	2,573,368	37%	+22	+18	31
Minnesota	100%	1,573,246	54%	1,275,653	44%	+18	+3	10
D.C.	100%	210,423	93%	14,821	7%	+86	+80	3
Ohio	98%	2,667,468	51%	2,481,198	47%	+4	+2	20
Massachusetts	100%	1,890,193	62%	1,104,088	36%	+28	+25	12

States won by McCain	162							
North Dakota	100%	141,113	45%	388,523	53%	+9	+27	3
Nebraska	100%	315,913	41%	438,421	57%	+10	+33	4
Montana	100%	220,401	47%	236,513	50%	+3	+21	3
Utah	100%	301,771	34%	555,497	63%	+20	+46	5
South Dakota	100%	170,817	45%	232,899	53%	+1	+21	3
Idaho	100%	235,709	36%	402,098	62%	+25	+38	4
Georgia	99%	1,811,198	47%	2,522,409	52%	+5	+17	15
Texas	100%	3,521,164	44%	4,467,748	56%	+12	+23	34
Kansas	100%	499,863	41%	695,414	57%	+12	+25	6
South Carolina	100%	842,441	45%	1,038,727	54%	+9	+17	8
Wyoming	100%	95,496	33%	160,438	63%	+31	+40	3
Mississippi	100%	1,517,899	42%	232,899	53%	+11	+20	6
Alabama	100%	911,510	36%	1,263,741	60%	+22	+28	9
Kentucky	100%	746,510	41%	1,043,264	57%	+10	+20	4
Arizona	99%	861,589	45%	1,212,878	54%	+4	+10	10
Alaska	99%	80,342	36%	136,349	62%	+28	+26	3
Oklahoma	100%	502,286	34%	869,645	66%	+31	+31	7
West Virginia	100%	301,438	42%	304,278	56%	+11	+13	5
Tennessee	100%	1,561,074	42%	1,470,160	57%	+15	+14	11
Louisiana	100%	730,361	42%	1,147,823	59%	+18	+15	9
Arkansas	98%	417,314	39%	632,142	59%	+20	+10	6

No winner called	27							
Missouri	100%	1,436,745	49%	1,442,613	49%	+7	+7	11
North Carolina	100%	2,115,854	50%	2,102,761	50%	+12	+2	15



Other encodings, such as area, color, etc., aren't great to enable accurate estimates, but may work well if the goal is to see the big picture

# When possible and appropriate, let people choose

33186



CHANGE IN HOME VALUE SINCE 2004

-10% 0% +9% +18% +28% +40%



Water Conservation Area 3B

Miami

Kendall

Biscayne National Park

Homestead

Everglades National Park

**UP 15%** 33186

Homes in Zip code 33186 in Miami, Fla., are worth \$41,112 more than in 2004. This is in the Miami-Fort Lauderdale-West Palm Beach, F metro area.

+100%

+50%

0%

-50%

2004

2009

2014

33186 Metro

© Mapbox © OpenStreetMap Improve this map, Census Bureau and Black Knight Financial Services

<https://www.washingtonpost.com/graphics/business/wonk/housing/overview/>

1. The elements of visualization
2. How visualization lies
3. Visualization for communication: essential principles

**THANK YOU!**