

Data Visualization Basics

Telling your story in charts, graphs, and maps

Shane Pekny

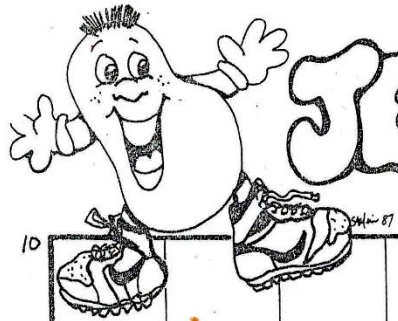
NOVEMBER 10, 2017



What: The symbolic representation of numbers and relationships.

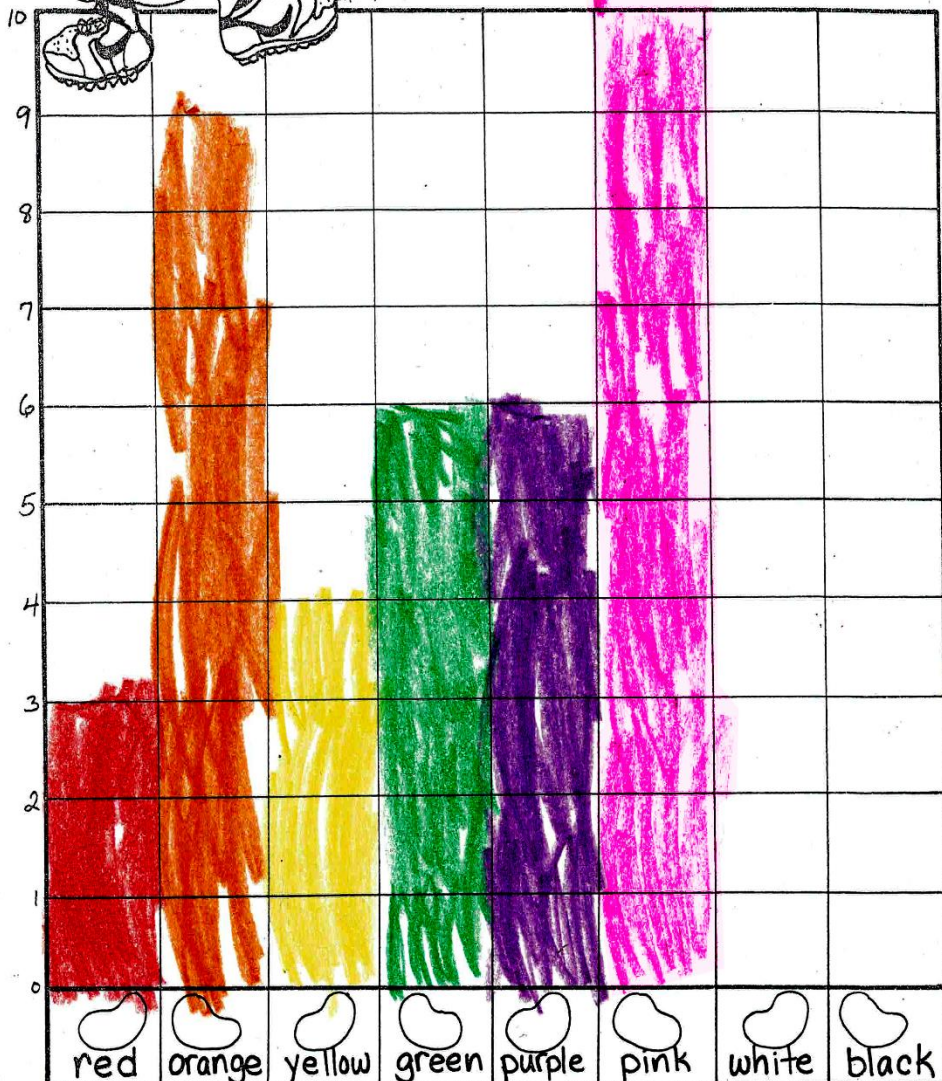
AKA: Data visualization, dataviz, informational graphics, infographics, visual communication, plots, charts, statistical graphics, graphs, or maps.

Why: Concise, precise, and authoritative.



JELLY BEANS

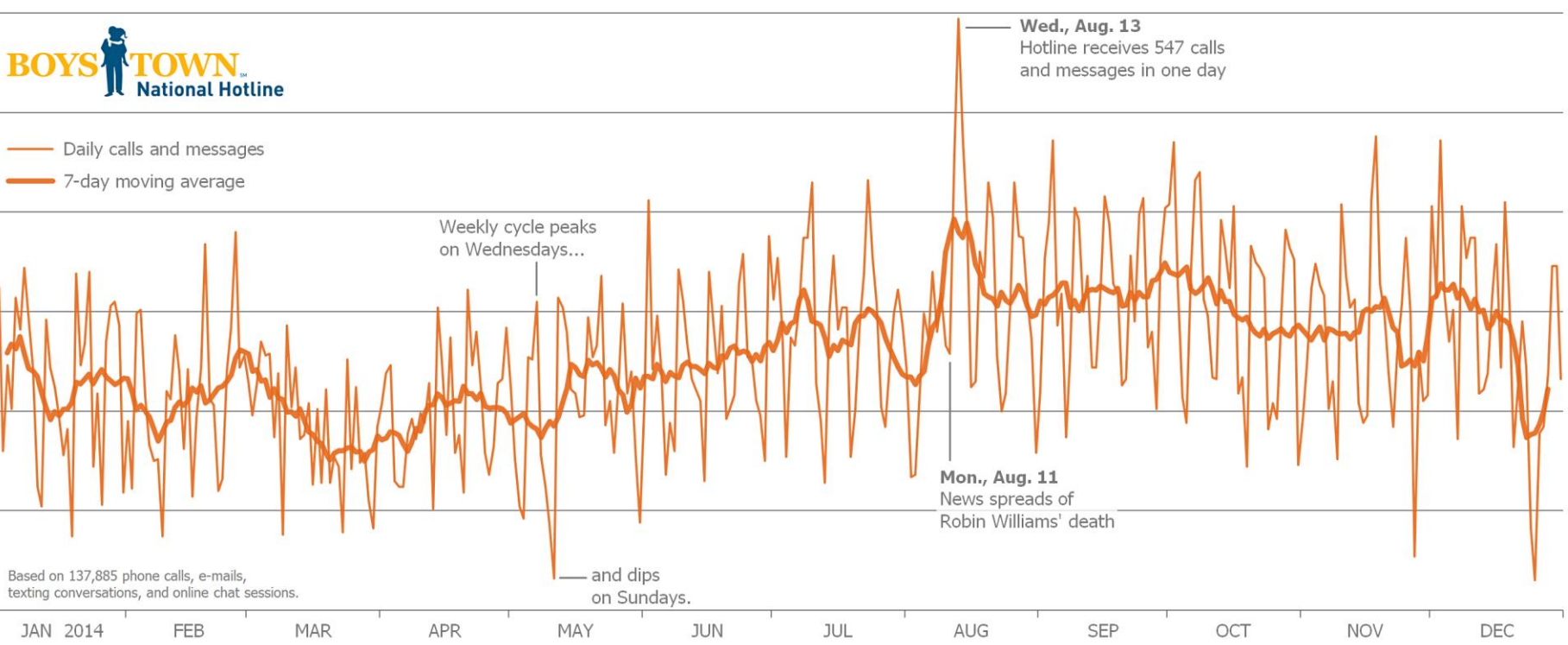
Max Name



Source: Maxwell Pekny's kindergarten homework (2016)



— Daily calls and messages
— 7-day moving average



Data Source: Boys Town National Hotline contacts database (2014)

North Carolina

2014 NATIONAL HOTLINE CONTACTS

In 2014, the Boys Town National Hotline received calls, e-mails, online messages, and mobile texts from all across North Carolina. Counselors helped teens and parents in communities large and small.

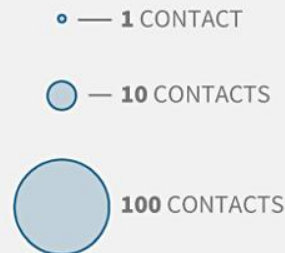
2,694

TOTAL CONTACTS

151

CITIES & TOWNS

Contact origins determined by caller ID and caller-provided information. Map symbols exclude 632 contacts where city of origin could not be determined. Statewide numbers may exclude additional contacts, many of them online, where states of origin could not be determined.



Raleigh
170 CONTACTS

Charlotte
327 CONTACTS

Fayetteville
283 CONTACTS



Watch videos about the people we help:
www.boystown.org/hotline

About me

Past 14 years of experience:

- Habitat for Humanity
- Goodwill Industries
- University of Nebraska at Omaha
- Boys Town
- Part-time teaching, magazine writing
- Prior experience and training in journalism

Most important lesson:

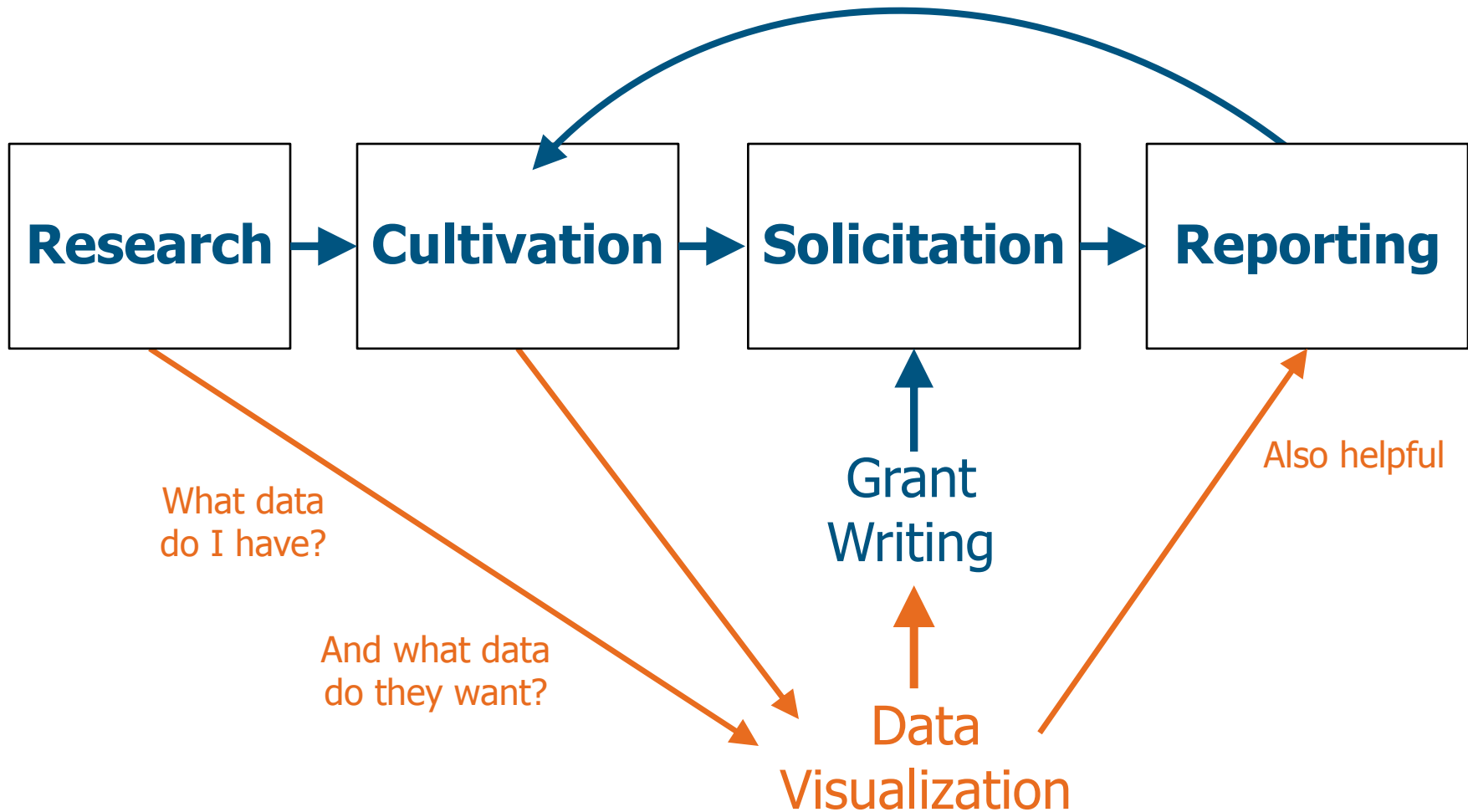
- Focus on the audience.

My goals:

- Be clear, concise, and compelling.

The context

Grant writing is just part of the process



The goals for today

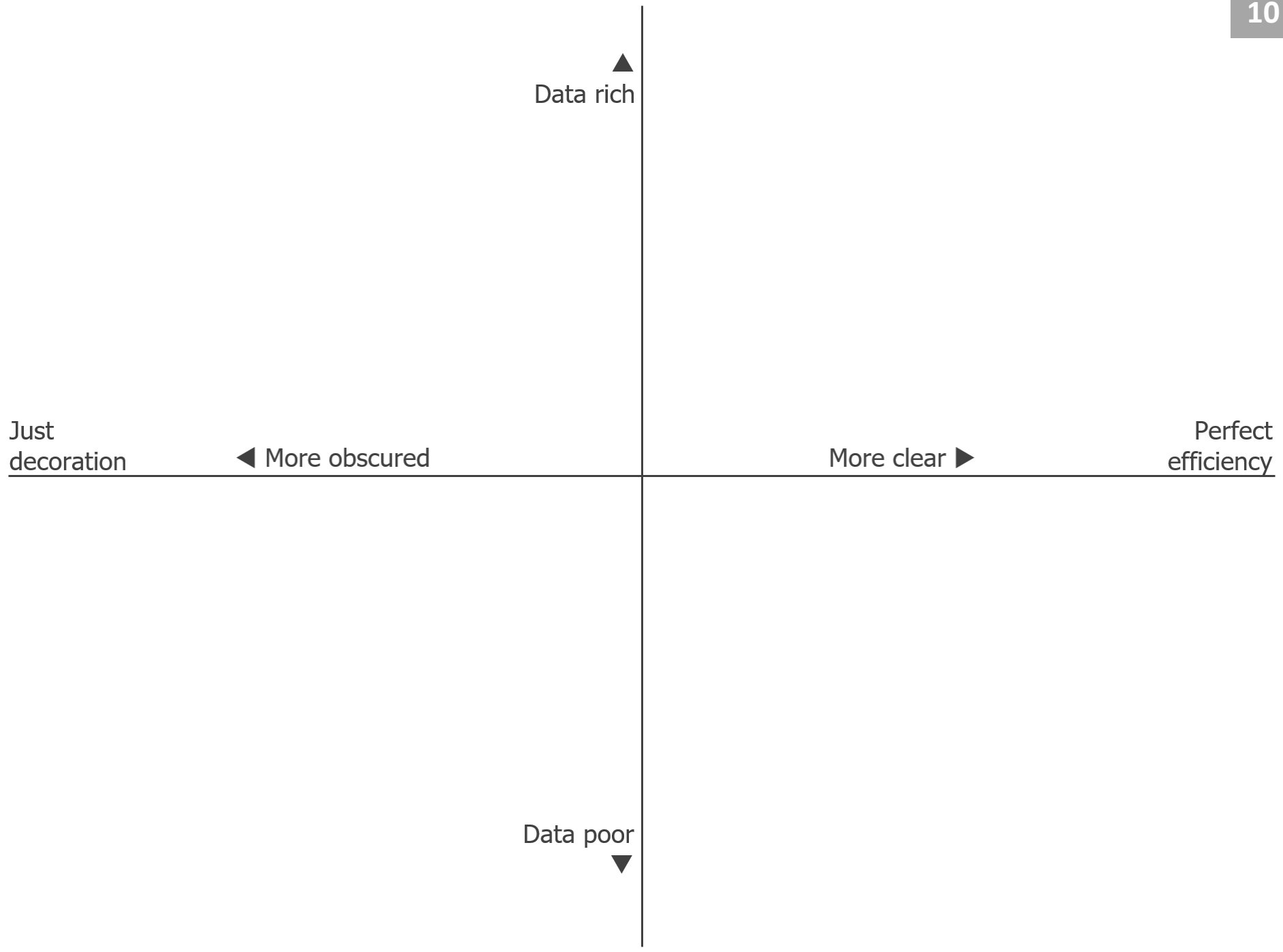
In data visualization:

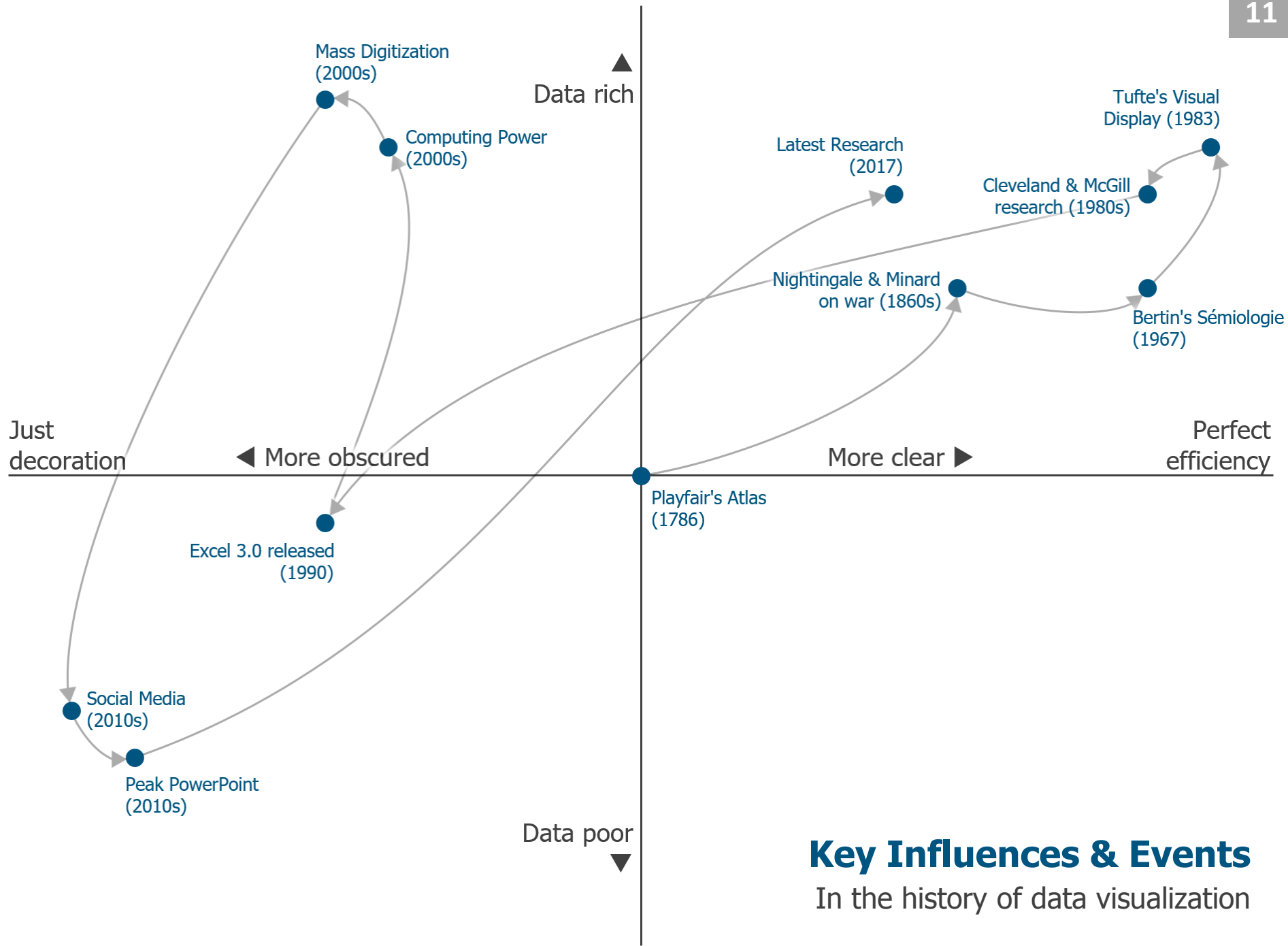
- Clarify purpose
- Increase clarity and depth
- Avoid pitfalls
- Identify next steps

Key Influences & Events

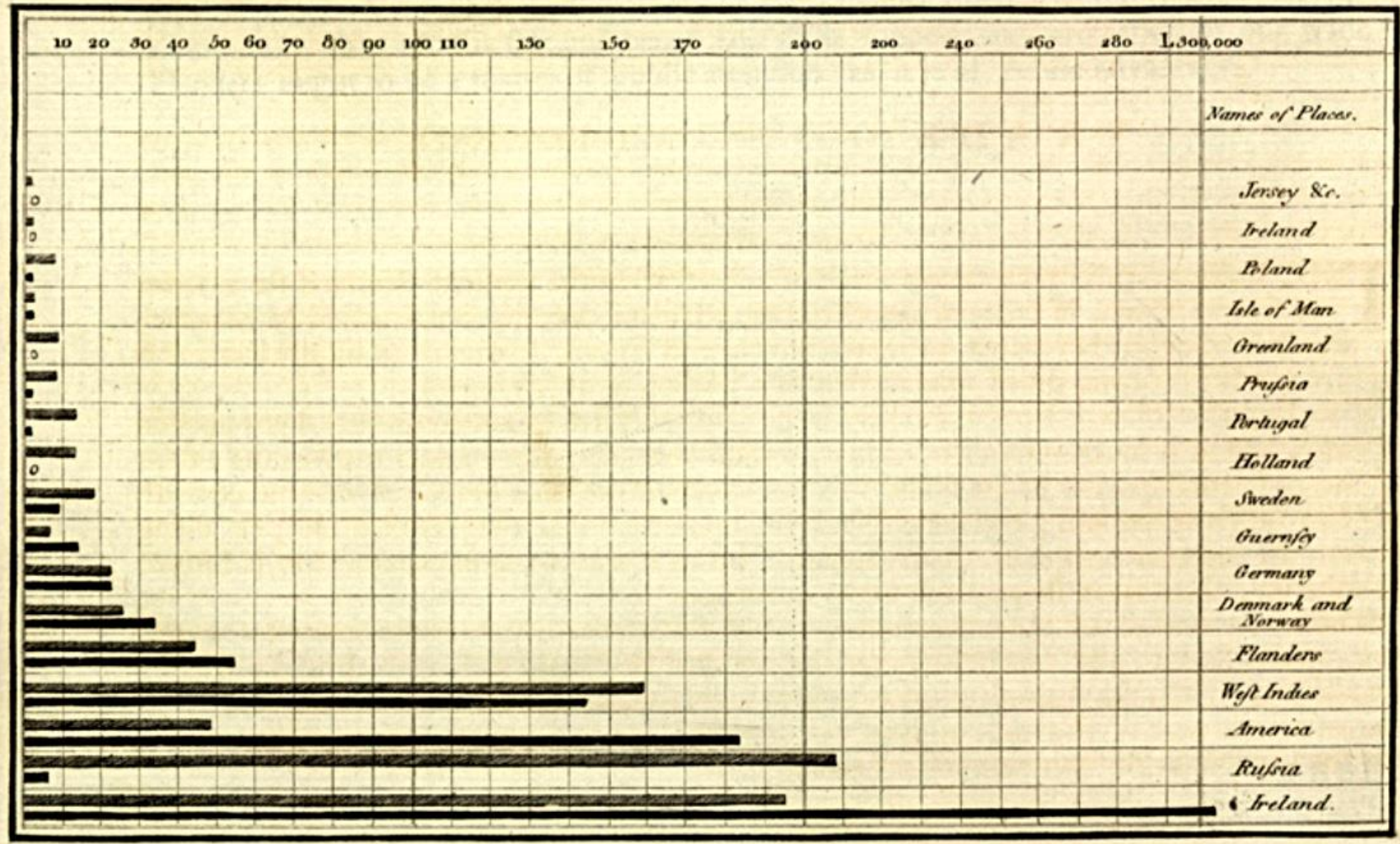
In the history of data visualization

Influence or event	When	Impact
Playfair's <i>Atlas</i>	1786	Invents line charts and bar charts
Nightingale & Minard on war	1860s	Reveals true causes and effects
Bertin's <i>Sémiologie</i>	1967	Builds theory and stresses efficiency
Tufte's <i>Visual Display</i>	1983	Manifesto on clarity and simplicity
Cleveland & McGill research	1980s	Studies perception, kills pie charts
Excel 3.0 released	1990	Puts 3D charts in everyone's hands
Computing power	2000s	Special forms for specialized fields
Mass digitization	2000s	Data more accessible, dizzying
Social media	2010s	Shares and likes are paramount
Peak PowerPoint	2010s	Impress in the briefing or board room
Latest research	2017	Old theories need new context





Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.



The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbedlines Imports.

Source: *The Commercial and Political Atlas*, William Playfair (1786)



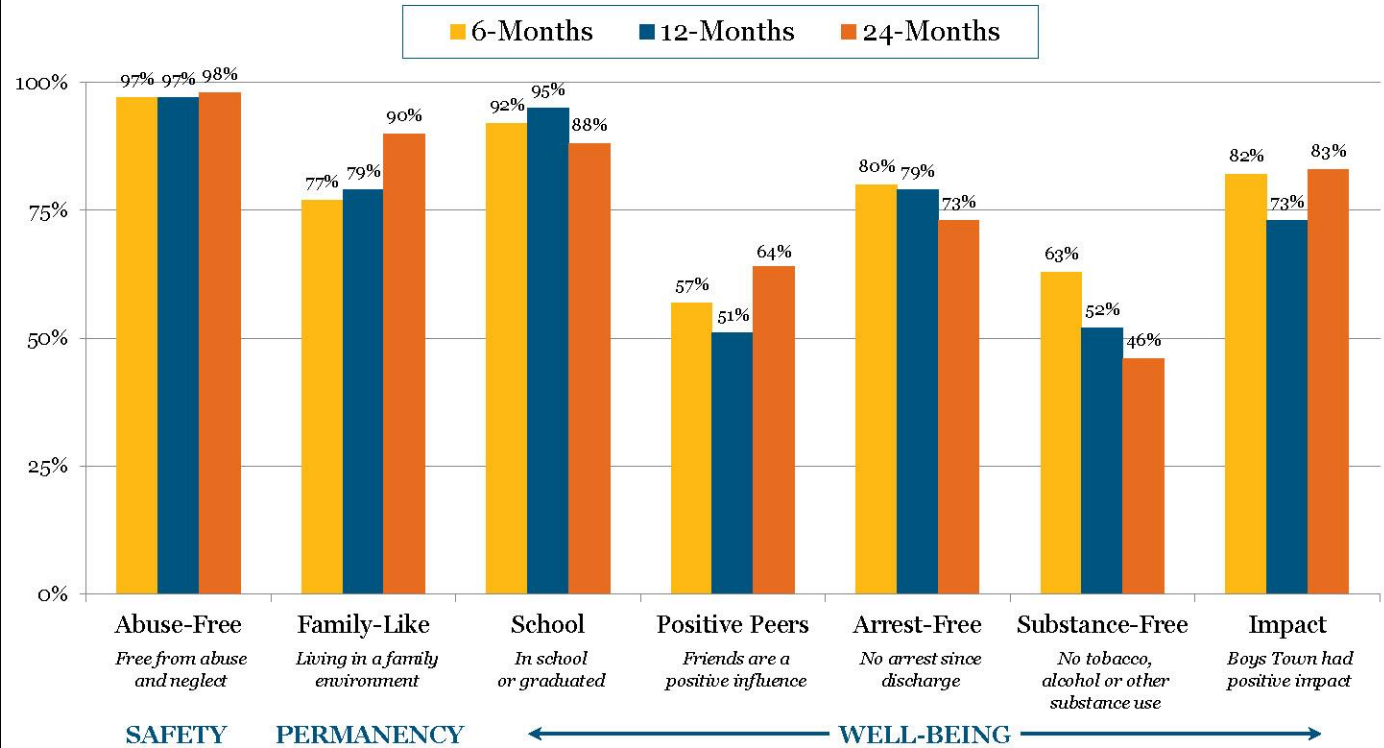
**Standard data
and clarity**



Youth Follow-Up Outcomes

National Family Home Program

Data collected January 1, 2016 – December 31, 2016

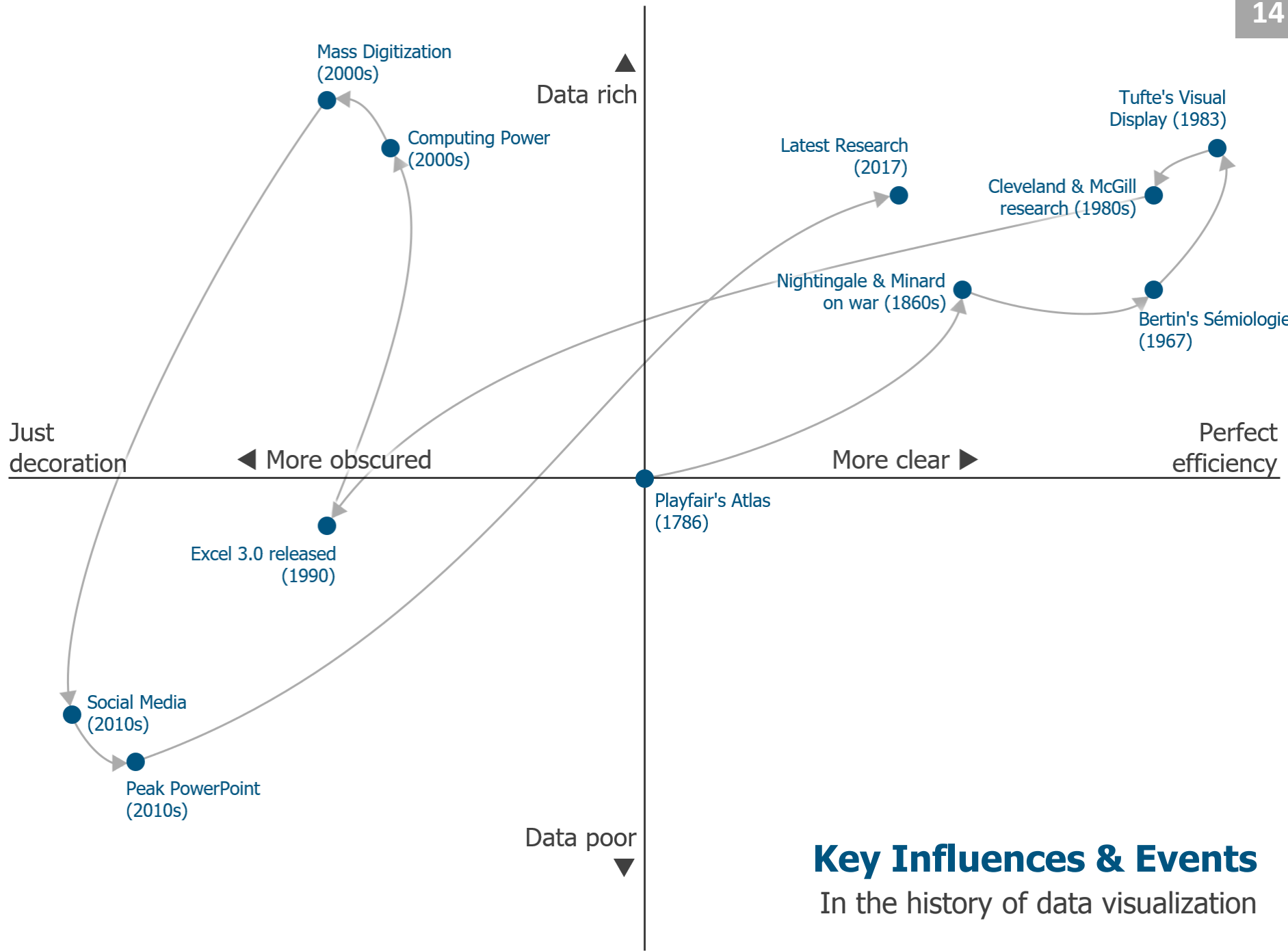


Standard data and clarity

Discharge Program: Family Home Program for Adolescents

Number of interviews completed (response rate) by timeframe:
 6-mo=352 (66%); 12-mo=246 (68%); 24-mo=226 (67%). 'N's vary by indicator.

Source: Family Home Program Follow-up Outcomes, Boys Town National Research Institute (2017)



Key Influences & Events

In the history of data visualization

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.
Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mohilow n'ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

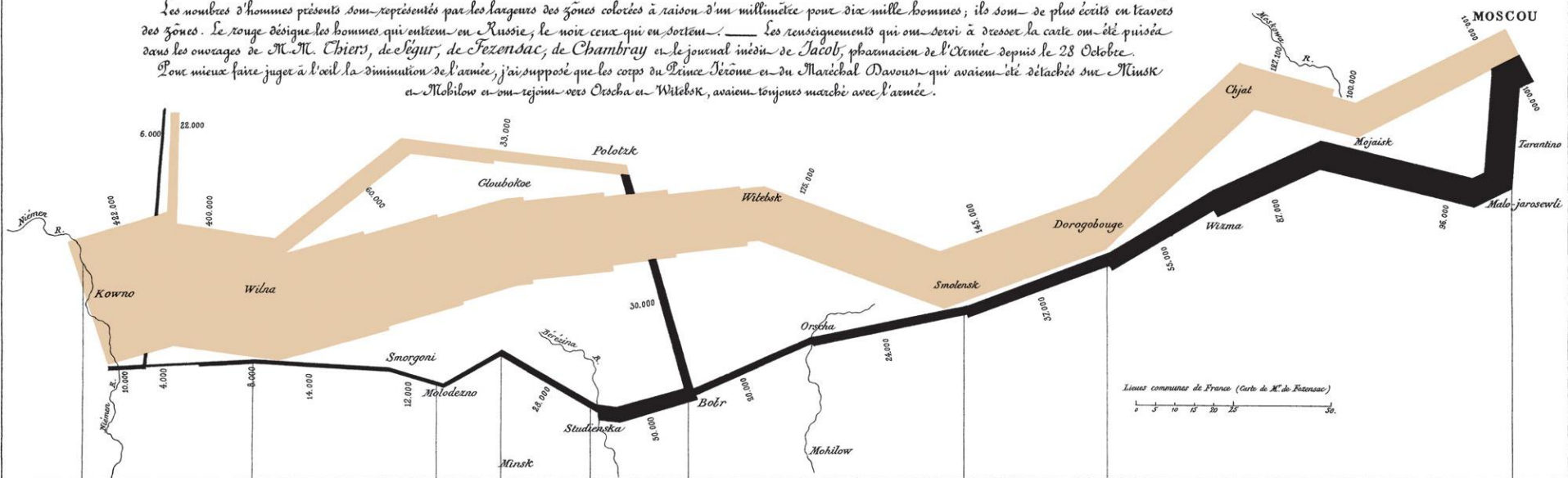
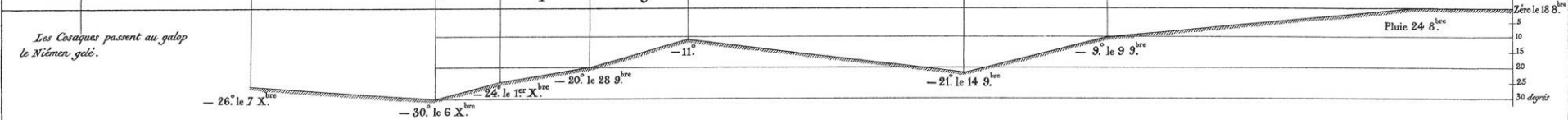


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



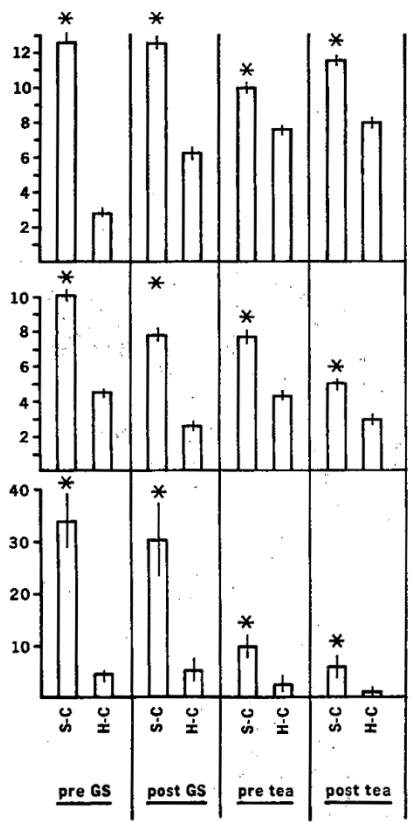
Autog. par Regnier, 8. Par. 5^{te} Marie St O^{me} à Paris.

Imp. Lit. Regnier et Douv. det.

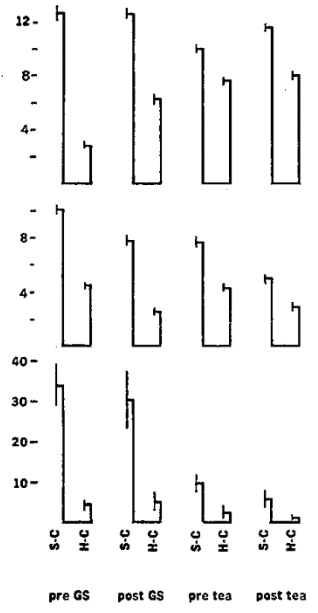
Source: "Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812-1813", Charles Minard (1869)



Consider this display, which compares each long bar with the adjacent short bar to show the viewer that, under the various experimental conditions, the long bar is longer:

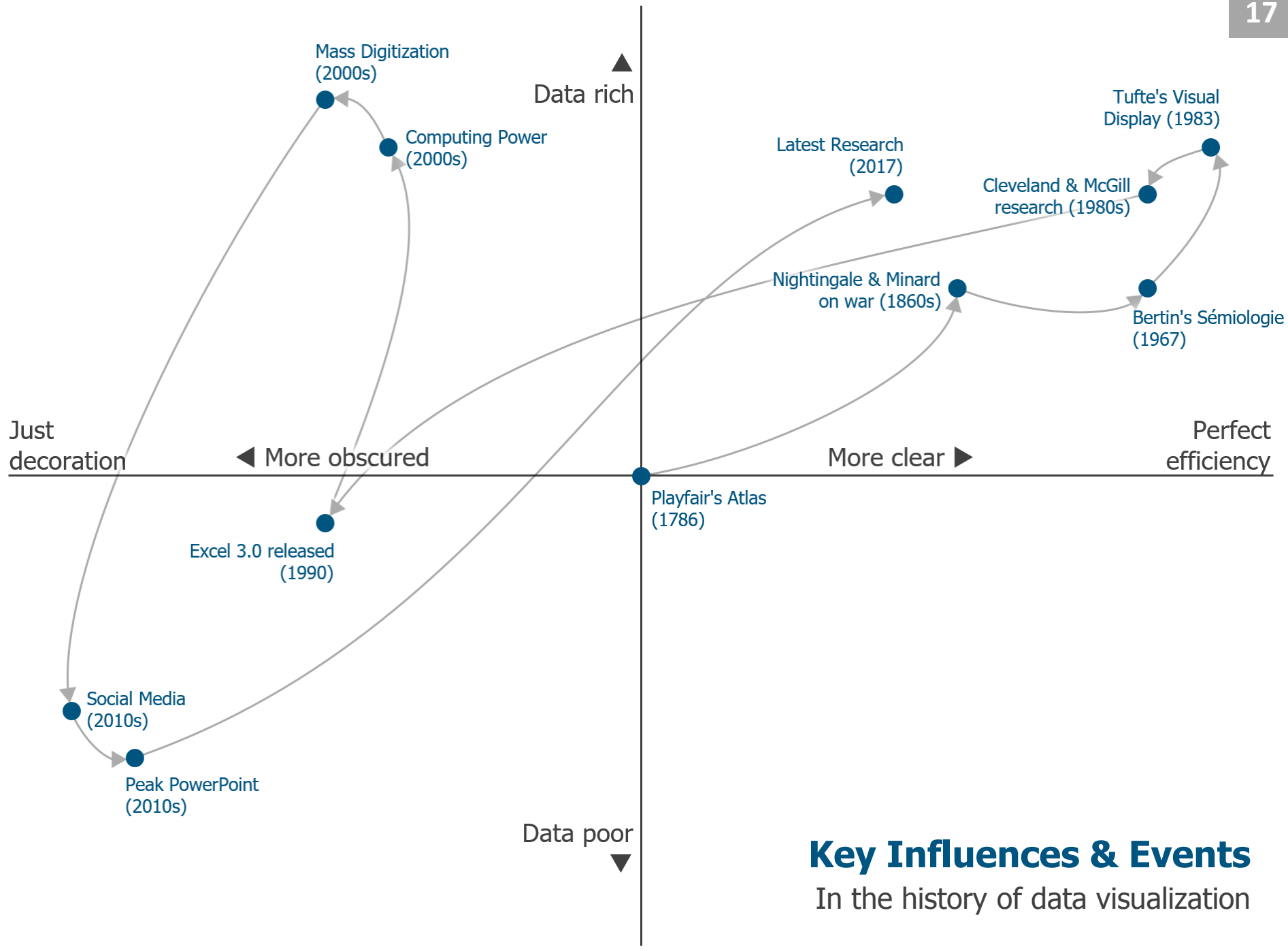


Vigorous pruning improves the graphic immensely, while still retaining all the original data. It is remarkable that erasing alone can work such a transformation:



Source: *The Visual Display of Quantitative Information*, Edward Tuft (1983)

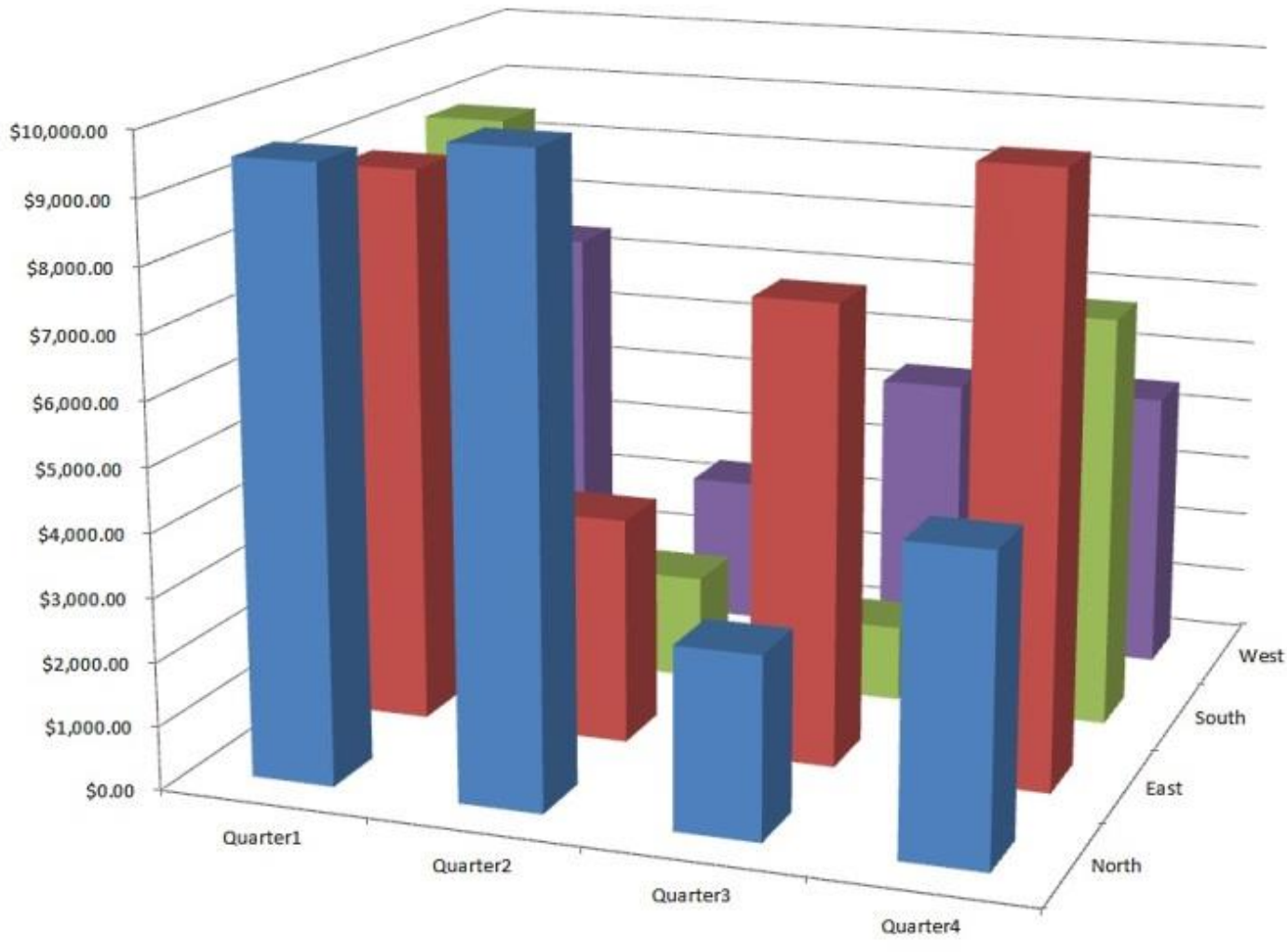




Key Influences & Events

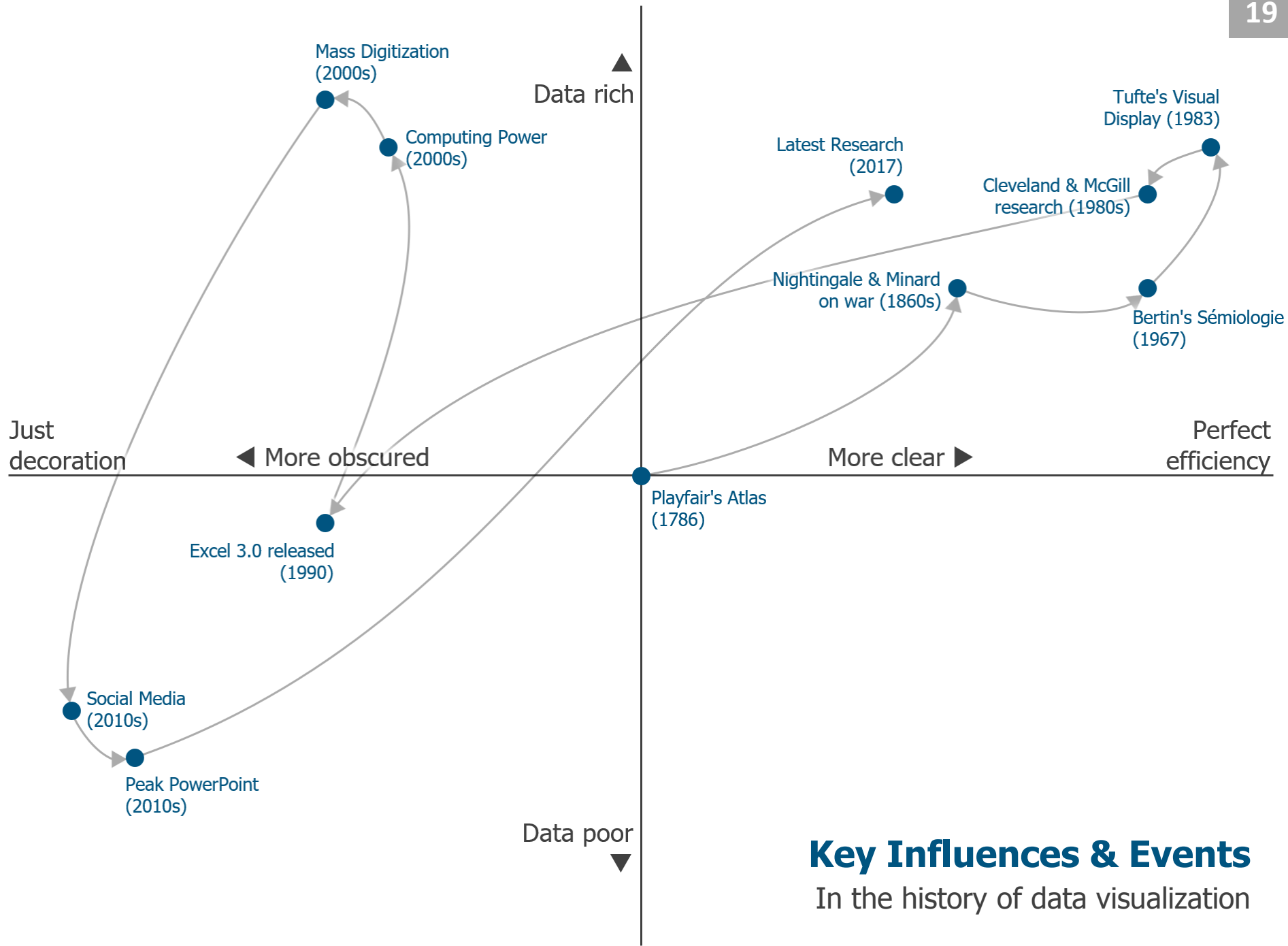
In the history of data visualization

Quarterly Report by Region



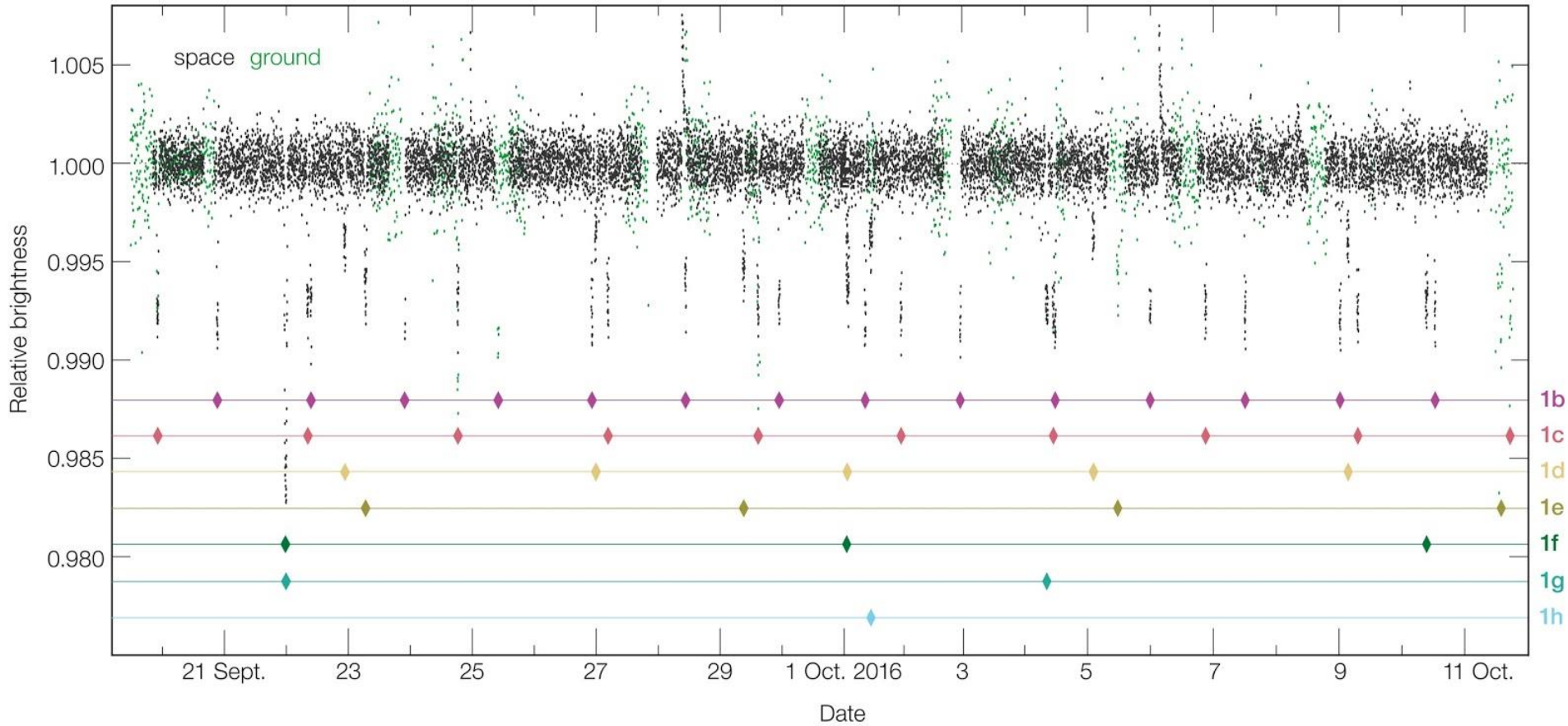
**Data poor,
More obscured**

Source: [https://msdn.microsoft.com/en-us/library/office/hh243933\(v=office.14\).aspx](https://msdn.microsoft.com/en-us/library/office/hh243933(v=office.14).aspx) (retrieved June 2017)



Key Influences & Events

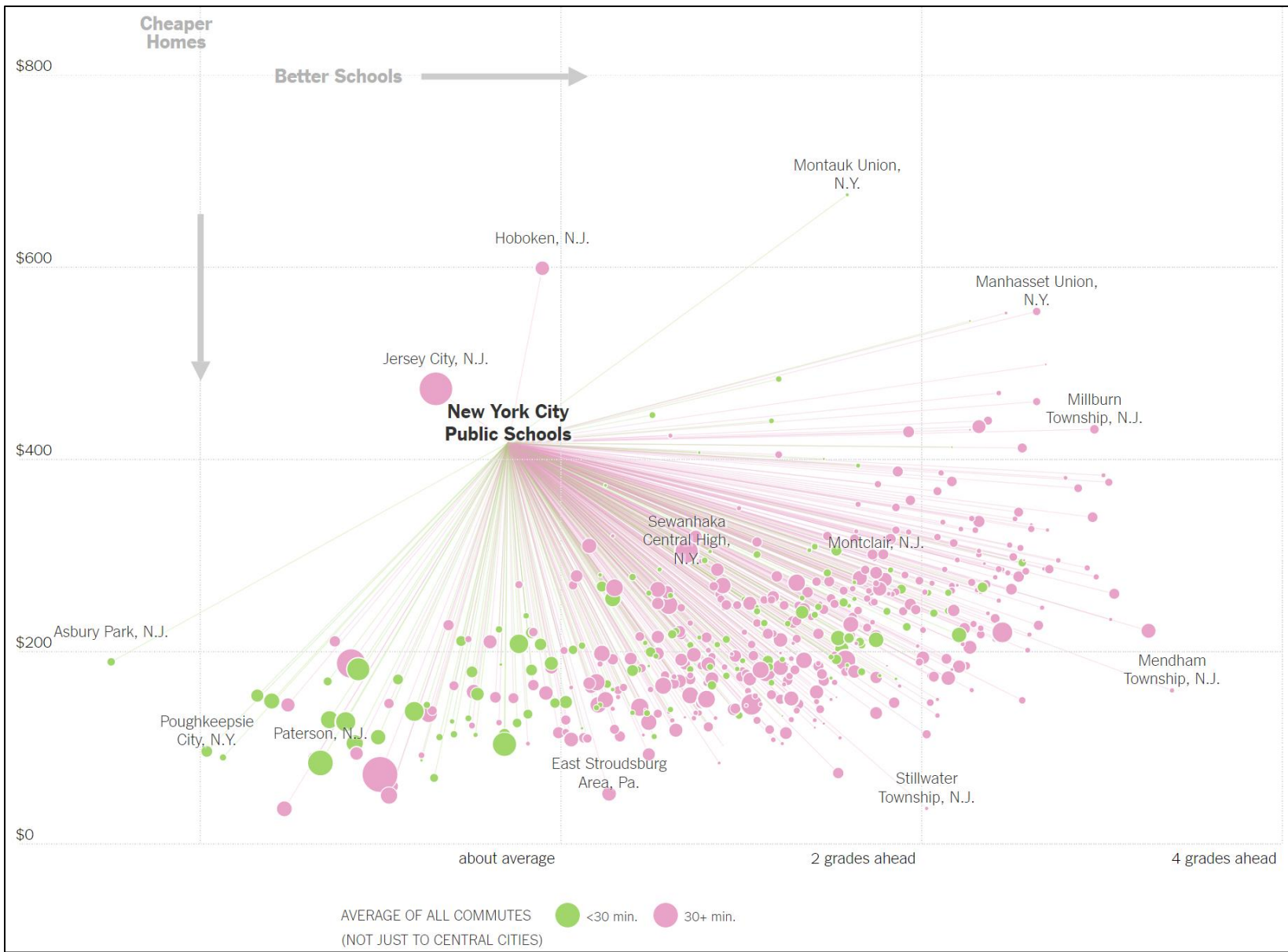
In the history of data visualization



Source: "A seven-planet resonant chain in TRAPPIST-1",
Nature Astronomy (May 2017)

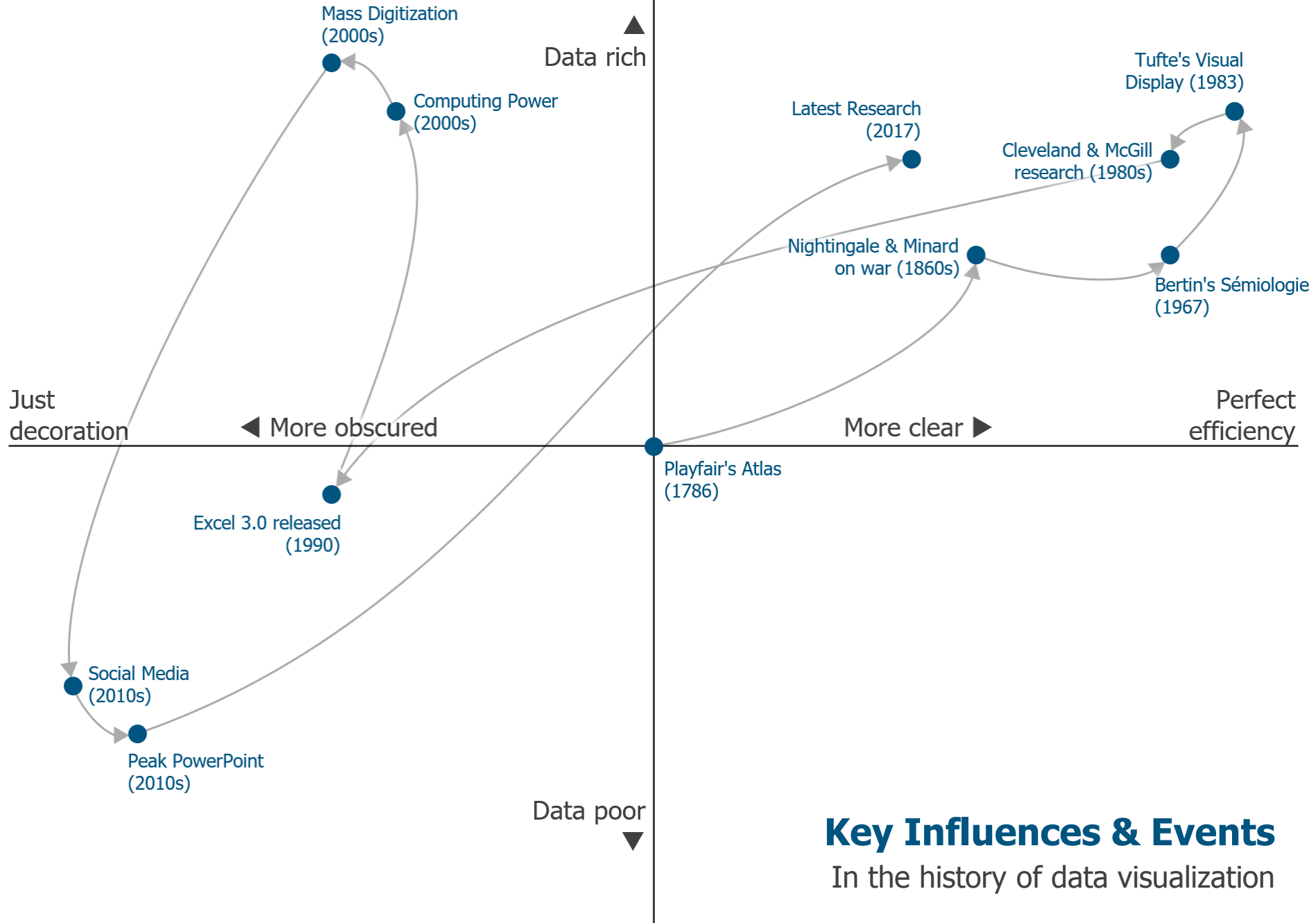


**Data rich,
 More obscure**



Source: "Good Schools, Affordable Homes: Finding Suburban Sweet Spots", *The New York Times/Upshot* (March 2017)





Key Influences & Events


In the history of data visualization

BOND KISSES

○ TOTAL OF KISSES AND
 ◐ AVERAGE KISSES PER FILM

Not surprisingly, **Roger Moore** was the luckiest Bond, kissing **24 women** over **7 movies** and winning an average of **3.4 kisses per film**.



45  Bond smooches once every **45 mins**



Timothy Dalton lacked the charm of the other Bonds kissing only **2** women per movie



**Data poor,
More obscured**

Source:
<http://www.gbshowplates.co.uk/battle-of-the-bonds-infographic/>
 (retrieved May 2017)

Land bank
70 000 ha

120 000 m²
 agricultural premises

1 300
 agricultural
 technical units

1 200
 sheep



2 500
 forage cows



3 630 ha
 crops & legumes



4 840 ha feeds

2 530 ha
 vegetables and
 perennial herbs



**Data poor,
 More obscured**

Source: Anton Egorov's page at www.coroflot.com (retrieved May 2017)

ALIEN MOVIES BY STATE

Which alien movie does your state love most?



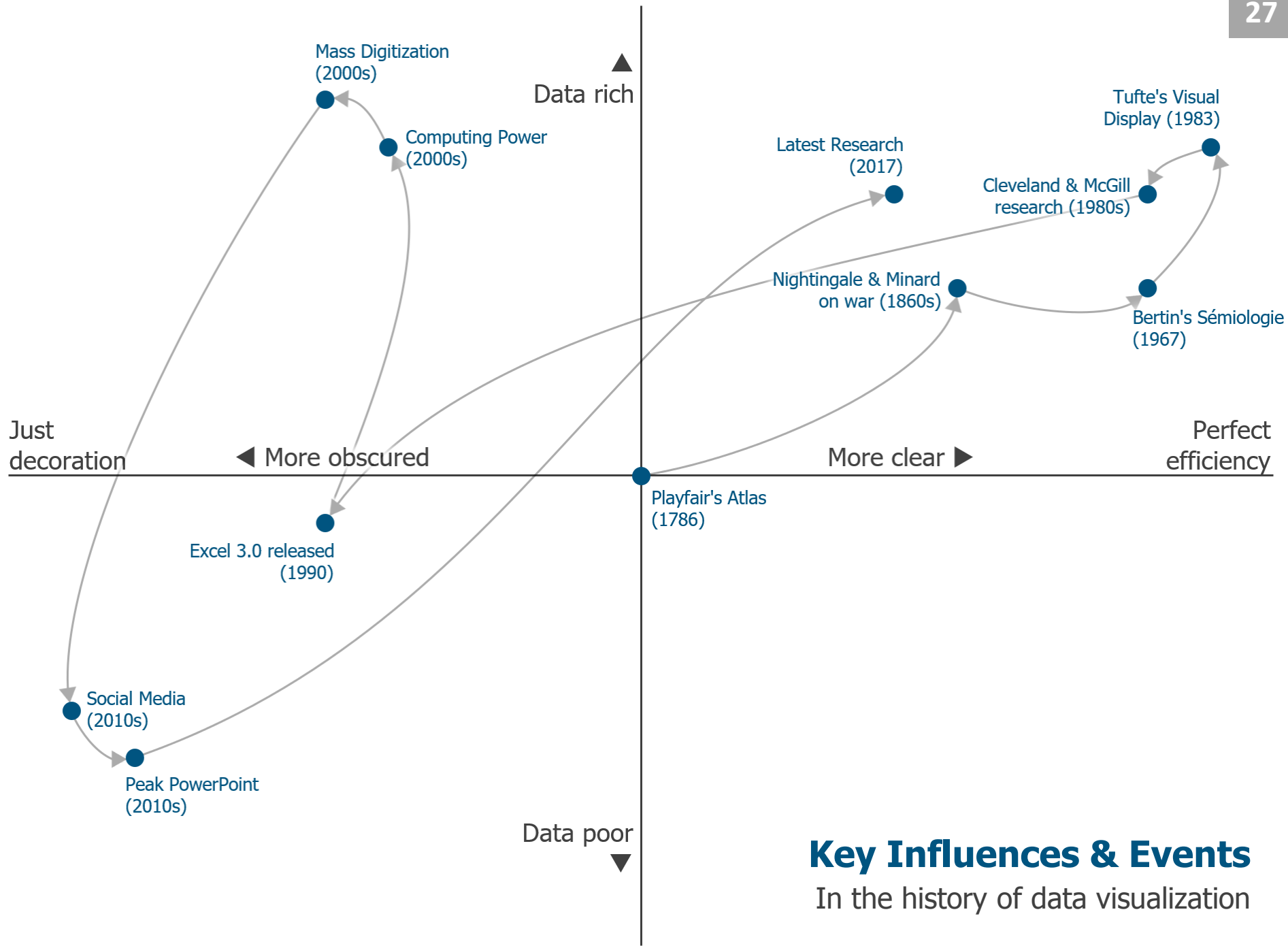
Data poor,
More obscured

Source: decluttr (<http://www.decluttr.com/blog/2017/05/10/what-is-americas-favorite-alien-movie/>)
(retrieved June 2017)



**Data poor,
More obscured**

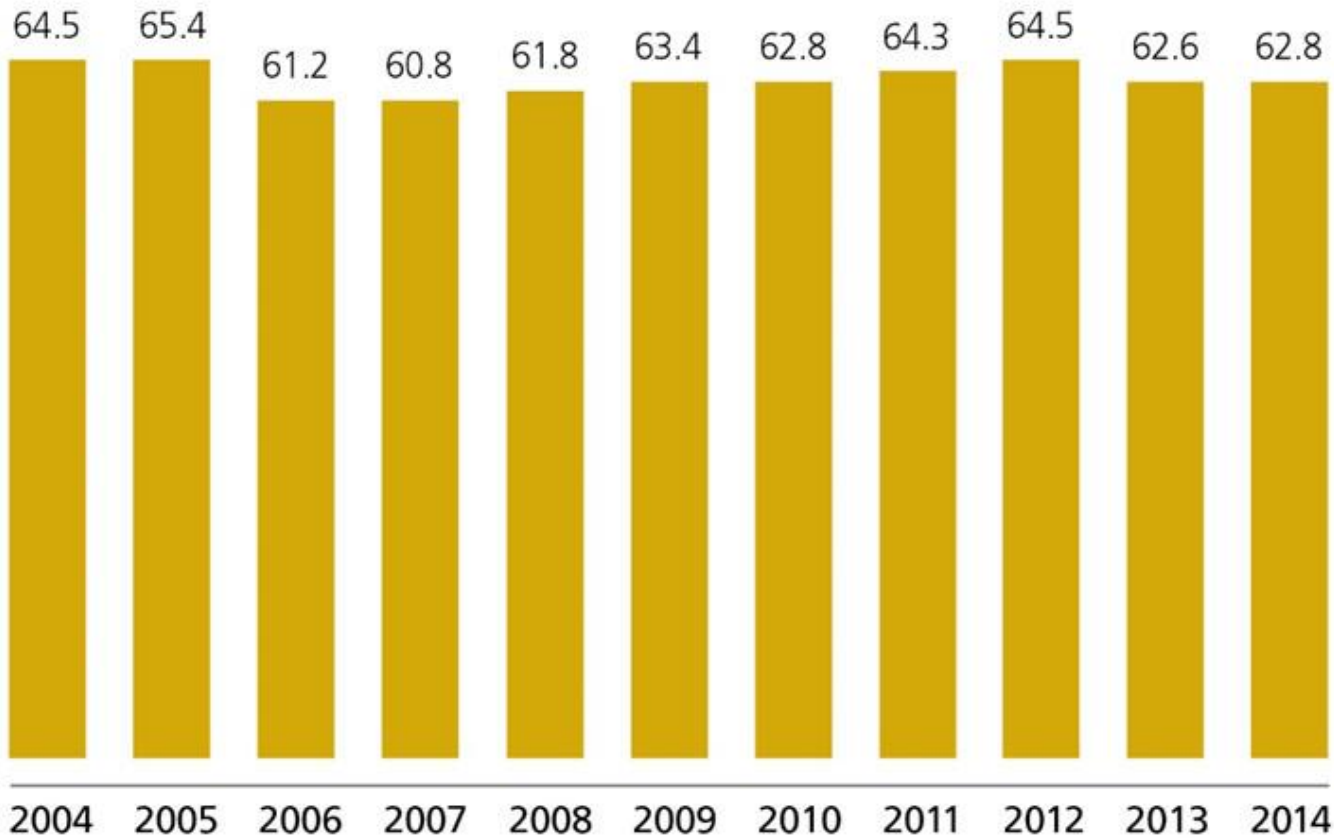
Source: <https://www.jpmorganchase.com/corporate/Corporate-Responsibility/corporate-responsibility.htm>
(retrieved May 2017)



Key Influences & Events

In the history of data visualization

Number of volunteers, in millions of people, 2004–2014

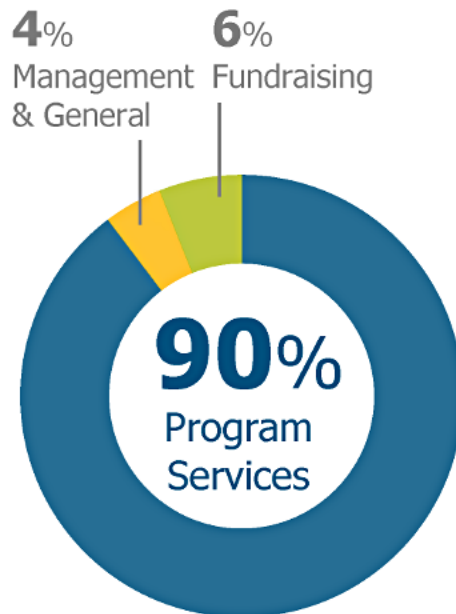


**Data poor,
More clear**

Source: *Giving USA 2016*, Giving USA Foundation 2016

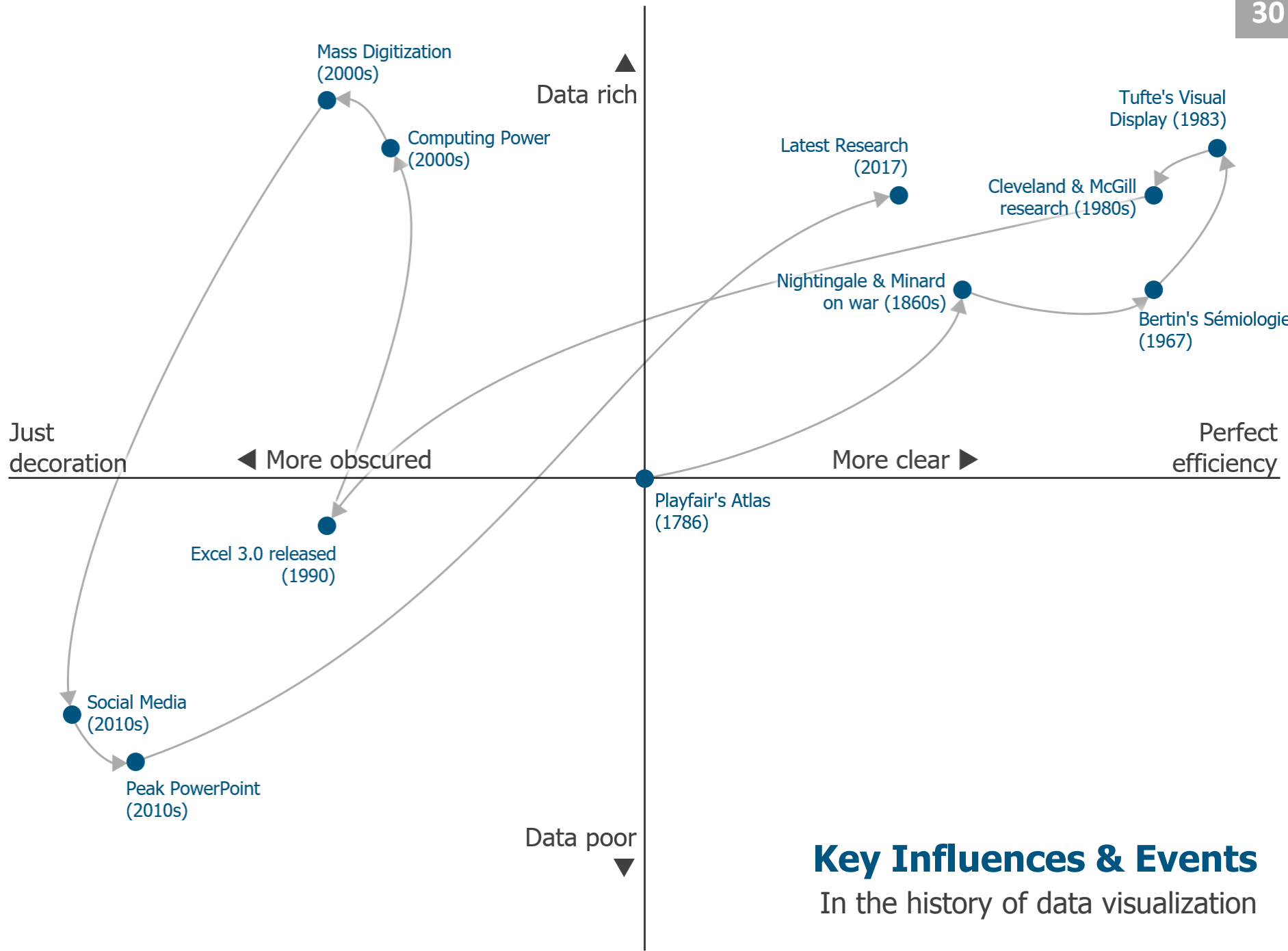
Cost Efficiency

Boys Town spends the vast majority of its revenue on research and the direct care of children & families.



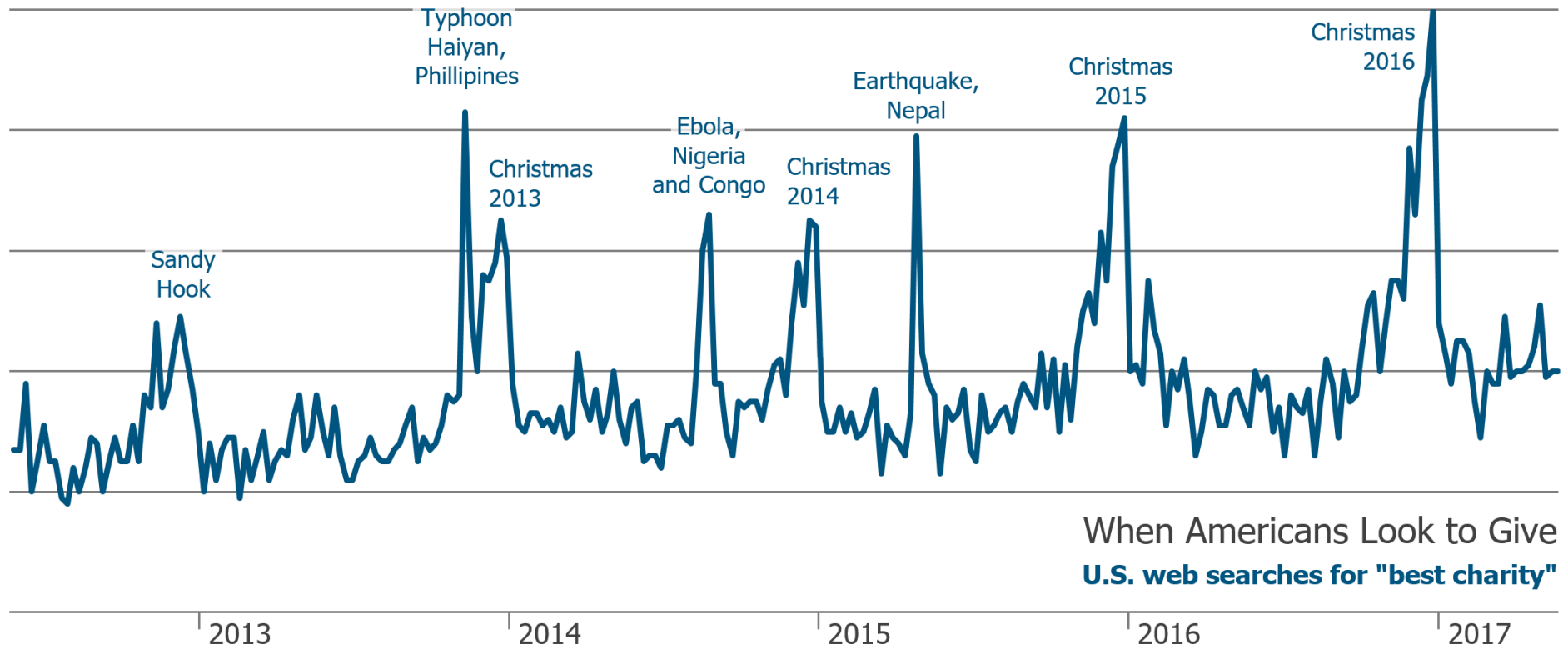
**Data poor,
More clear**

Data Source: *Boys Town Consolidated
Financial Statements December 31, 2014*



Key Influences & Events

In the history of data visualization



When Americans Look to Give
U.S. web searches for "best charity"

Data source: Google Trends

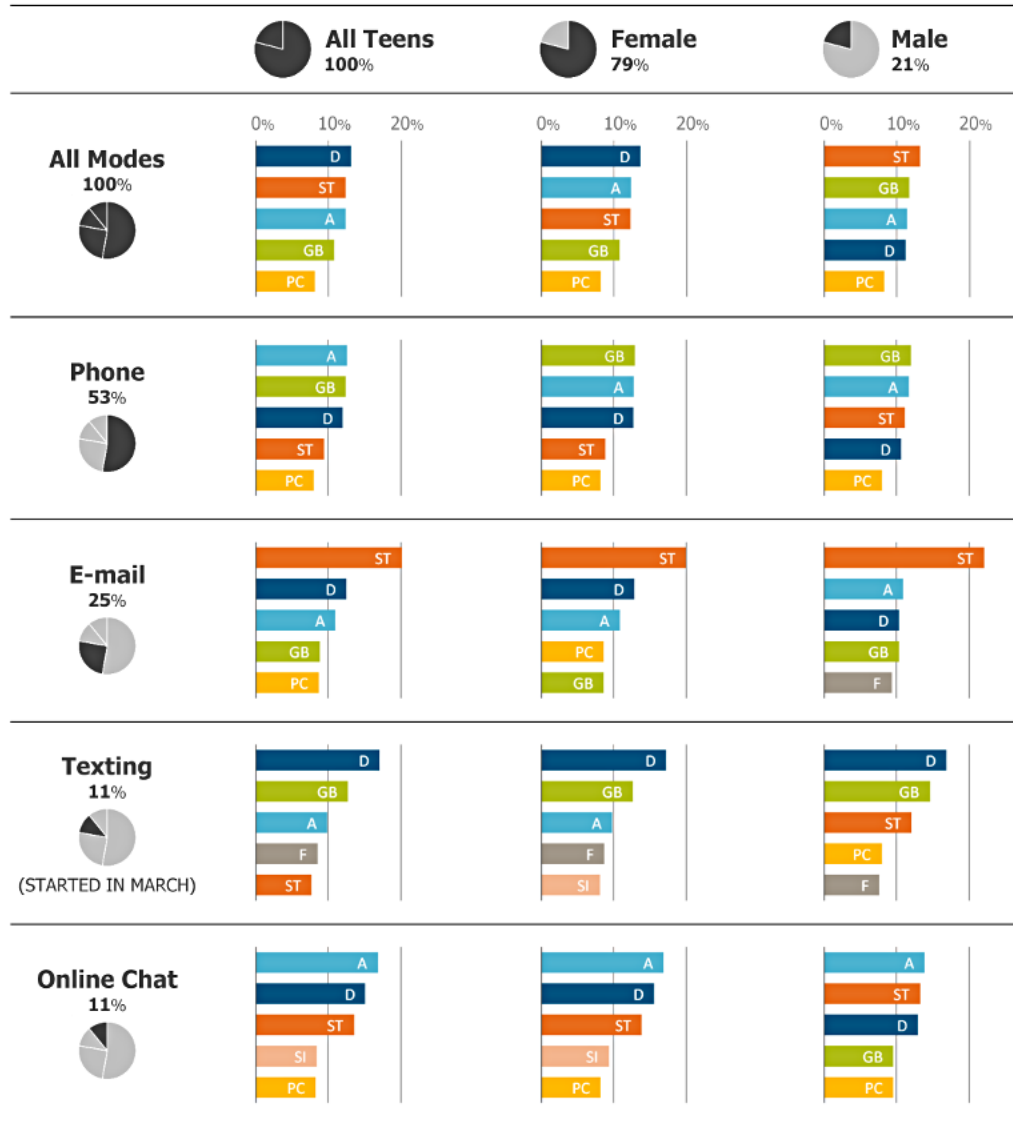


Top Issues Faced by Teens

BY GENDER & MODE OF COMMUNICATION

NATIONWIDE CONTACTS IN 2014
FROM YOUTH AGES 19 OR YOUNGER

- D – Depression
- ST – Suicidal thoughts
- A – Anxiety
- GB – Girlfriend or boyfriend issues
- PC – Parent-child issues
- F – Friendship issues
- SI – Self-injury



Data source: Boys Town National Hotline contact database (2014)

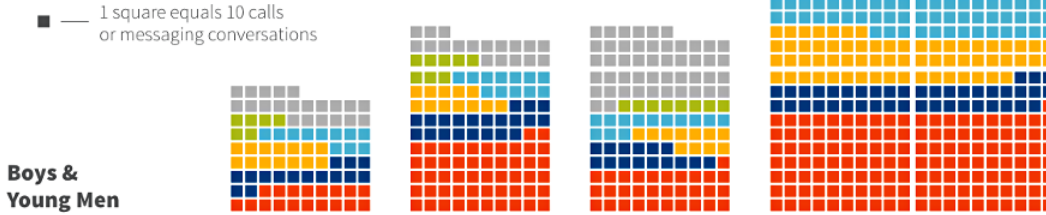
Note: Only the top five issues in each category are shown. Counselors use a total of 47 codes to classify issues.

Top Issues Faced by Teens

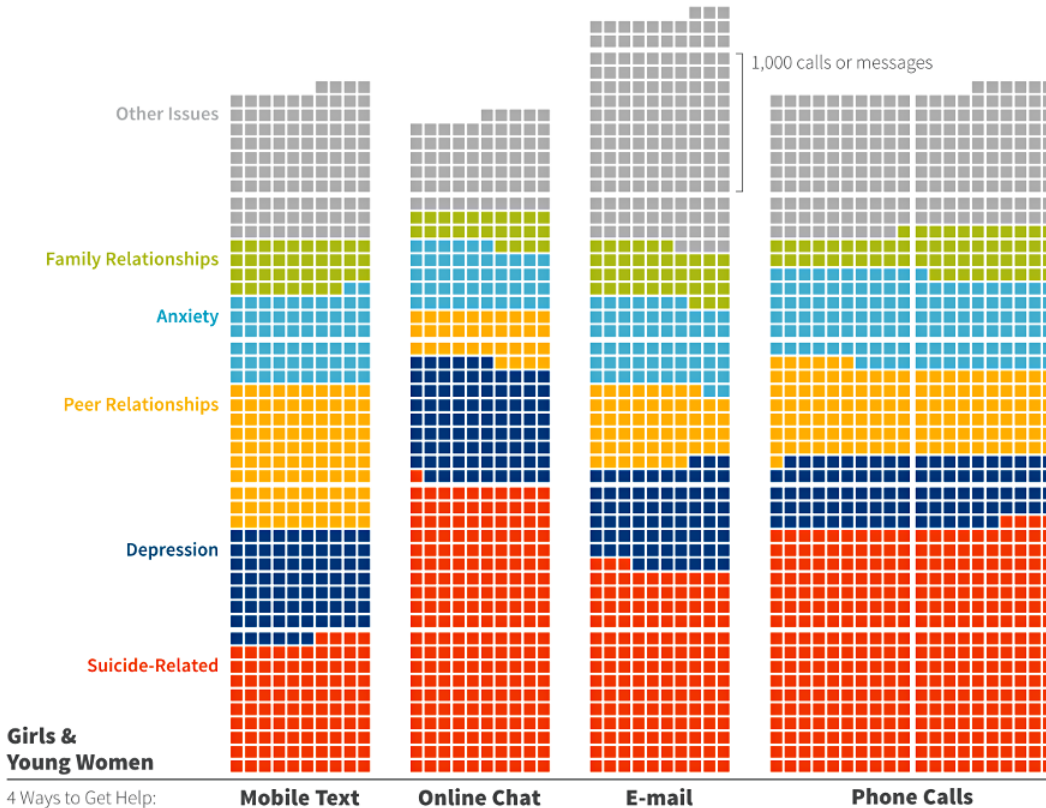
BY GENDER & MODE OF COMMUNICATION

Hotline counselors received more than 30,000 calls and messages from teens in 2016. The most common issues were **suicide-related**, including suicidal thoughts, concerns about a friend or family member's suicidal behavior, and even suicides in progress. Mental health issues, primarily **depression** and **anxiety**, were top concerns as well, along with **family relationships**. Issues related to **peer relationships** included gangs, dating, online harassment, and bullying. Other issues were diverse – questions about sex, addictions, school failure, loss of loved ones, and general loneliness, among others.

Visualization based on 31,879 phone calls, e-mails, and messaging conversations from youth ages 19 and under in 2016.



Data source: Boys Town National Hotline contact database (2016)



Educational attainment in each school district in the U.S.

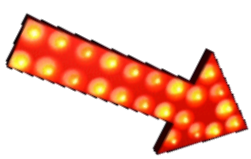
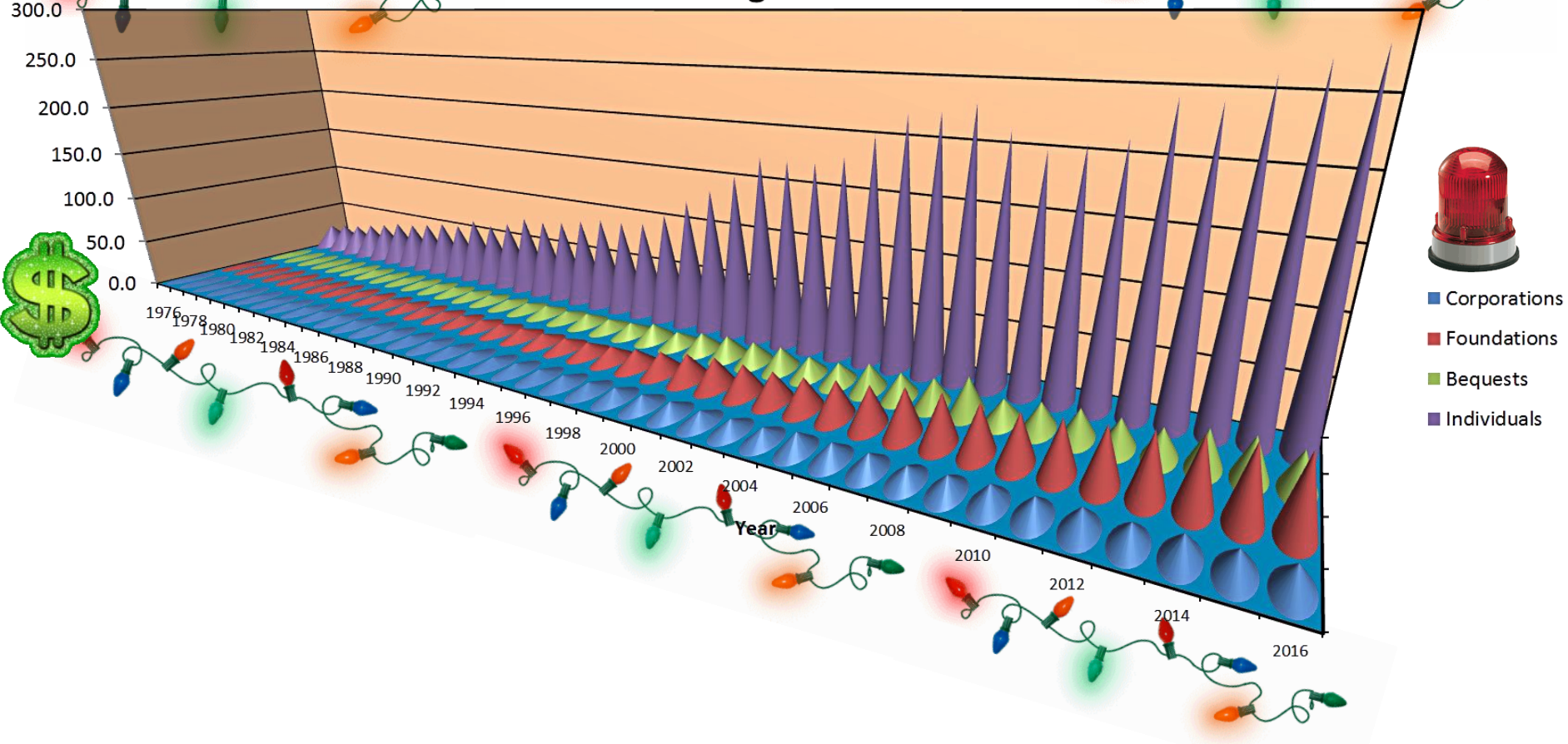


Each circle represents one school district. Larger circles represent districts with more students.

Source: "Money, Race and Success: How Your School District Compares",
The New York Times/Upshot (April 2016)



Charitable Giving Sources 1976-2016



Data source: *Giving USA 2017*, Giving USA Foundation



If this is 10...




Is this 5? 3?


Change Chart Type [?] [X]

- Templates
 - Column
 - Line
 - Pie
 - Bar
 - Area
 - XY (Scatter)
 - Stock
 - Surface
 - Doughnut
 - Bubble
 - Radar


Column




Line




Pie




Bar




Area




XY (Scatter)




Stock




Surface




Doughnut



Bubble



Radar




Manage Templates... Set as Default Chart OK Cancel


Change Chart Type [?] [X]

- Templates
 - Column
 - Line
 - Pie
 - Bar
 - Area
 - XY (Scatter)
 - Stock
 - Surface
 - Doughnut
 - Bubble
 - Radar


Column




Line




Pie




Bar




Area




XY (Scatter)




Stock




Surface



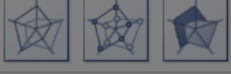
Doughnut



Bubble



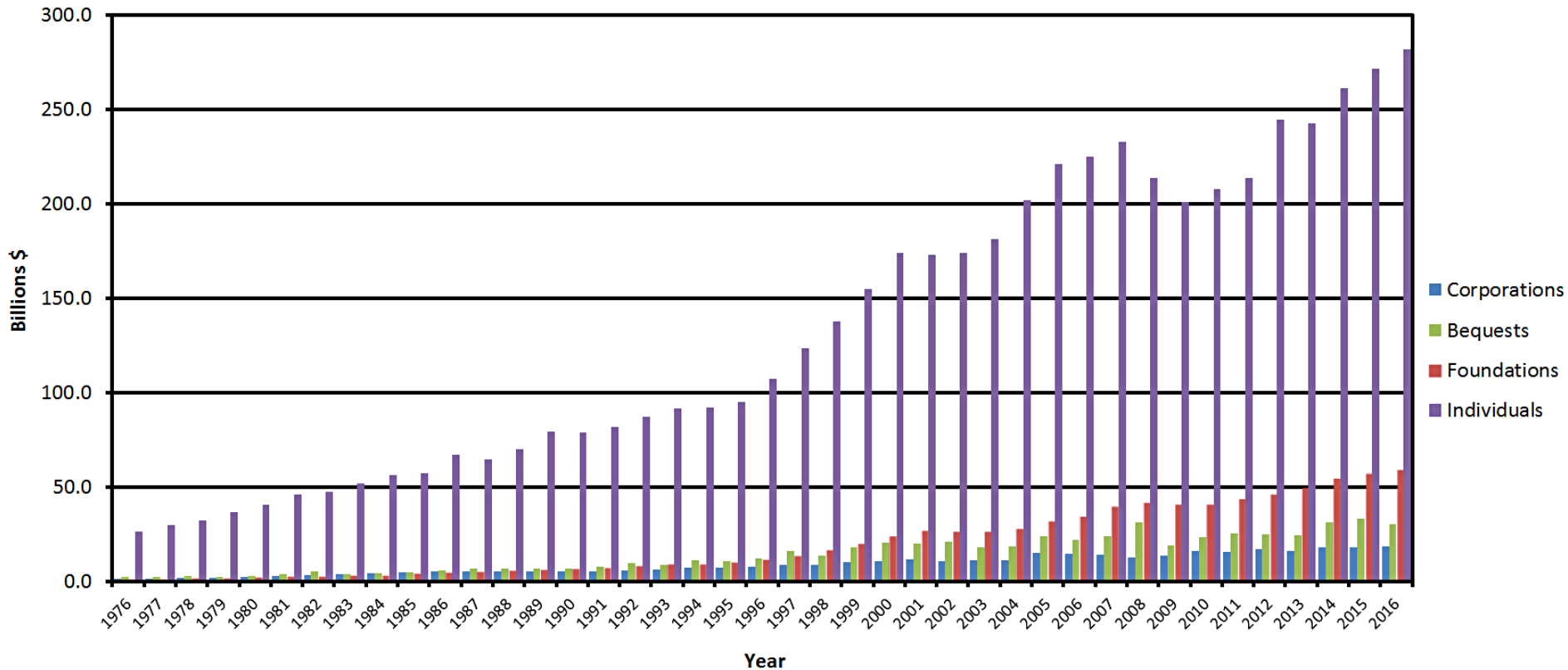
Radar



Manage Templates... Set as Default Chart [OK] [Cancel]

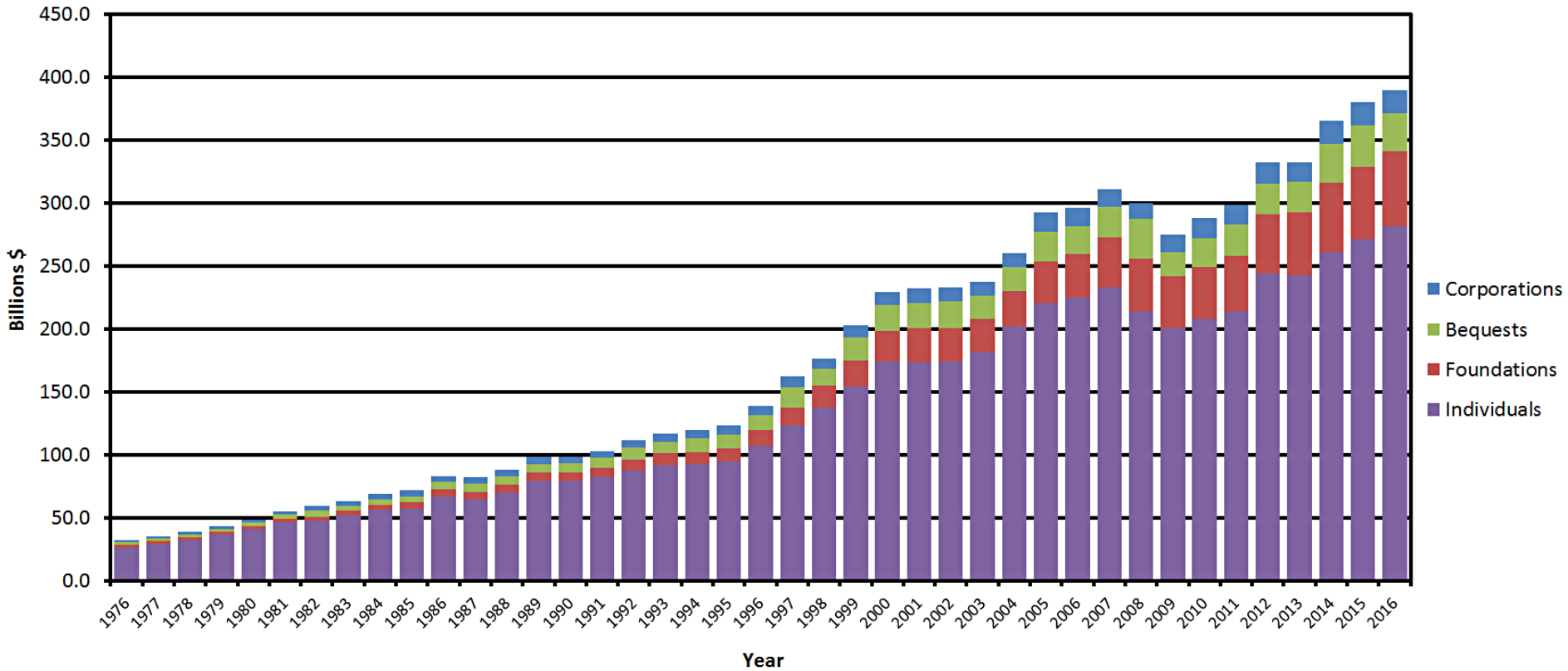
First, go 2D, clear background, and generally undecorate.

Charitable Giving Sources 1976-2016



Next, choose right chart type, scale, and resolution.

Charitable Giving Sources 1976-2016



Hotline Contacts 2014 by Date Time State City.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Design **Format** Tell me what you want to do

Series " 7-day moving av" Format Selection Reset to Match Style

Current Selection Insert Shapes Shape Styles WordArt Styles Arrange Size

Chart 6 =SERIES(Sheet1!\$EC\$1,Sheet1!\$DZ\$2:\$DZ\$366,Sheet1!\$EC\$2:\$EC\$363,2)

	DN	DO	DP	DQ	DR	DS	DT	DU	DV	DW	DX	DY	DZ	EA	EB	EC	ED
28												Jan	1/27/2014	2	385	367.1429	
29												Jan	1/28/2014	3	403	365.2857	
30												Jan	1/29/2014	4	405	363.2857	
31												Jan	1/30/2014	5	393	364.4286	
32												Jan	1/31/2014	6	309	366.4286	
33												Feb	2/1/2014	7	345	366.1429	
34												Feb	2/2/2014	1	311	359.2857	
35												Feb	2/3/2014	2	399	350.7143	
36												Feb	2/4/2014	3	401	353	
37												Feb	2/5/2014	4	357	350.2857	
38												Feb	2/6/2014	5	333	346.8571	
39												Feb	2/7/2014	6	325	341.2857	
40												Feb	2/8/2014	7	326	334.8571	
41												Feb	2/9/2014	1	287	339.2857	
42												Feb	2/10/2014	2	360	344.5714	
43												Feb	2/11/2014	3	356	345.4286	
44												Feb	2/12/2014	4	388	351.8571	
45												Feb	2/13/2014	5	370	354.7143	
46												Feb	2/14/2014	6	331	352.8571	
47												Feb	2/15/2014	7	371	355.1429	
48												Feb	2/16/2014	1	307	361.7143	
49												Feb	2/17/2014	2	347	359.7143	
50												Feb	2/18/2014	3	372	362.8571	

Format Data Series

Series Options

Line Marker

Line

- No line
- Solid line
- Gradient line
- Automatic

Color

Transparency 10%

Width 2.25 pt

Compound type

Dash type

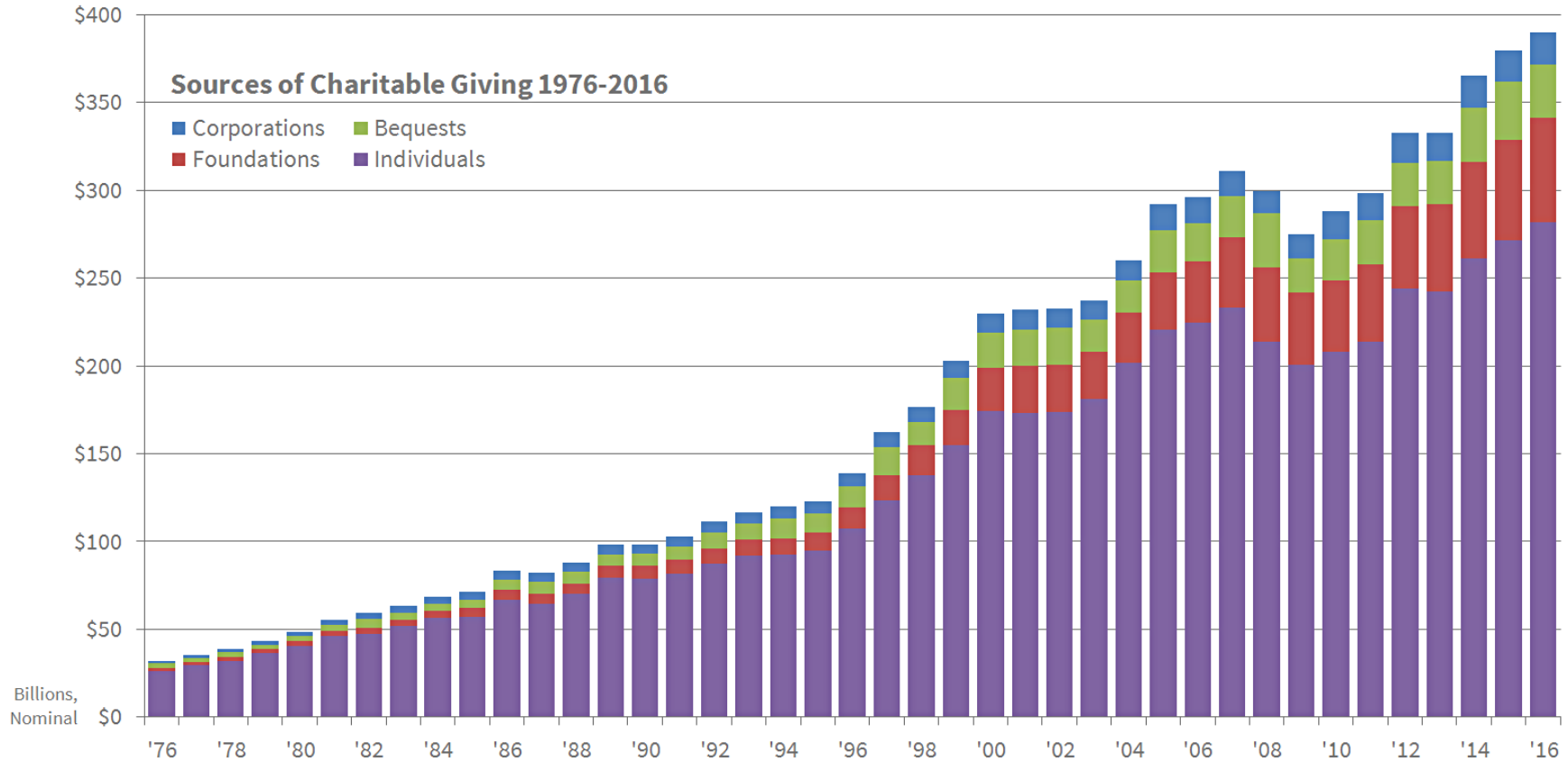
Cap type Round

Join type Round

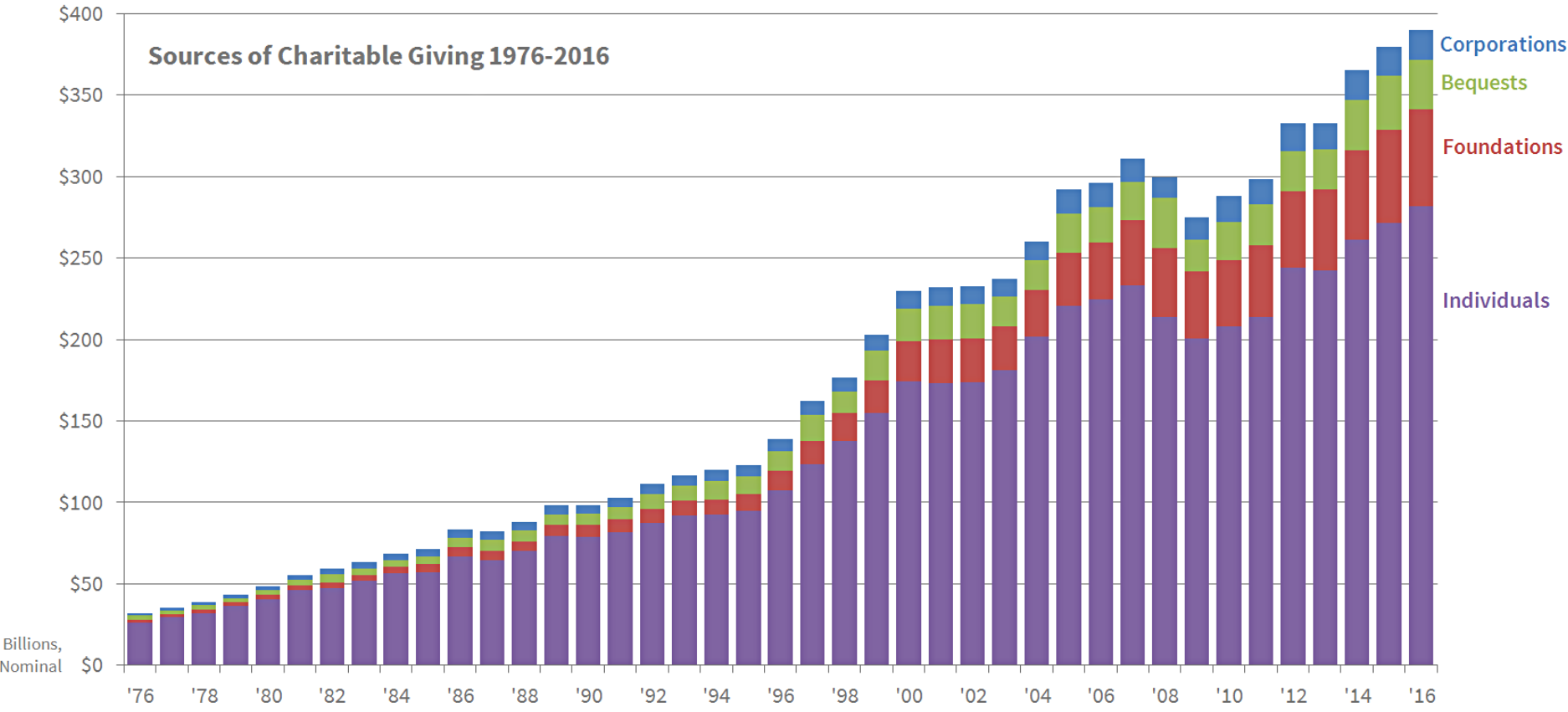
▲
**Right click
 on something**

▲
**Change its color, width,
 or make it disappear**

Establish a hierarchy of information, prune, and de-emphasize non-data.



Avoid keys or legends; put labels right next to data.



Data source: *Giving USA 2017*, Giving USA Foundation

Our Colors

Just as important as correct use of our logo are the colors we use.

By using specific colors, we can adapt the Boys Town image to nearly every application, while maintaining high recognition value.

Our color palette consists of two primary colors (blue and gold), six secondary colors (warm red, medium gray, light blue, soft yellow, light gray, green) and white.

Primary colors



Blue
PANTONE® 302
C:100 M:25 Y:0 K:50
R:0 G:84 B:128



Gold
PANTONE® 130
C:0 M:30 Y:100 K:0
R:253 G:185 B:19

Secondary colors



Warm Red
PANTONE® 173
C:0 M:70 Y:100 K:5
R:232 G:108 B:31



Medium Gray
PANTONE® 403
C:0 M:7 Y:17 K:43
R:161 G:151 B:137



Light Blue
PANTONE® 7459
C:70 M:0 Y:6 K:15
R:77 G:179 B:208



Soft Yellow
PANTONE® 7402
C:0 M:6 Y:90 K:0
R:255 G:238 B:188



Light Gray
PANTONE®
WARM GRAY 3
C:0 M:4 Y:87 K:17
R:216 G:209 B:201



Green
PANTONE® 583
C:25 M:0 Y:80 K:5
R:175 G:189 B:33



WHITE
C:0 M:0 Y:0 K:0
R:255 G:255 B:255

Use your brand guide, or try an online tool:

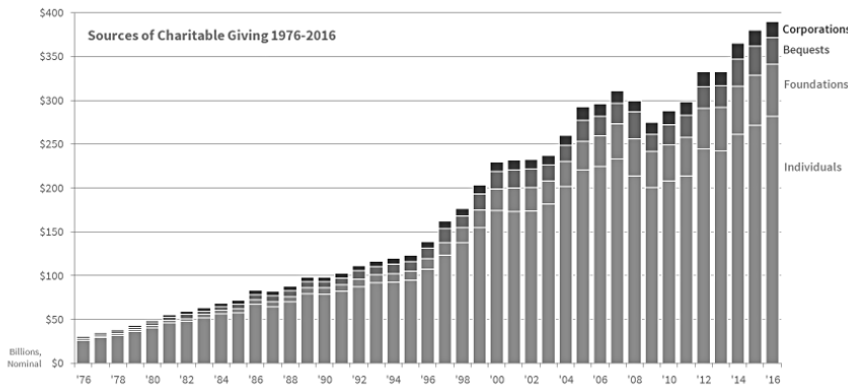
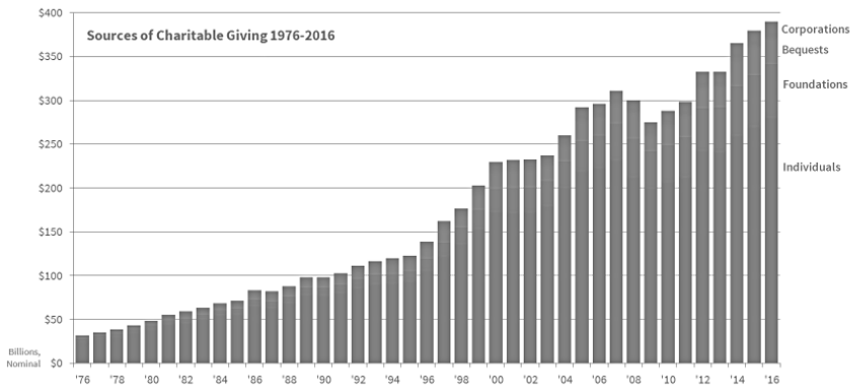
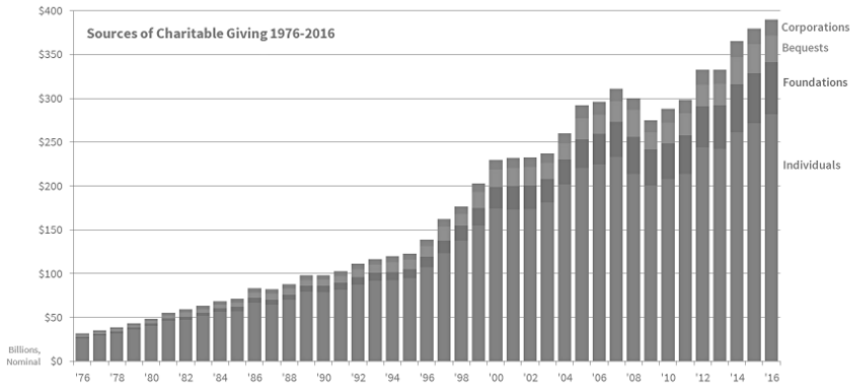
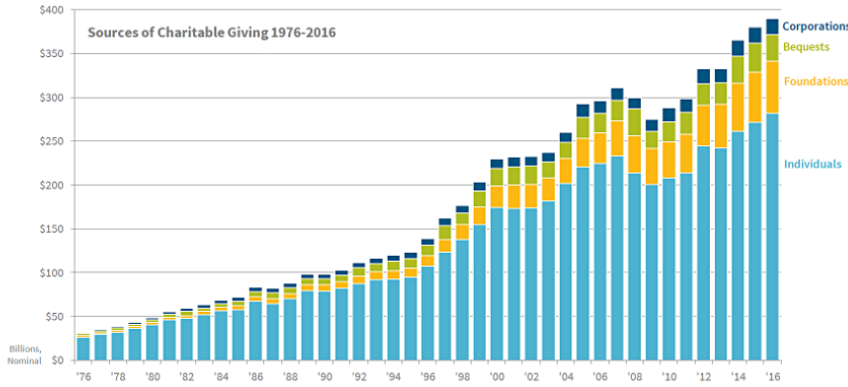
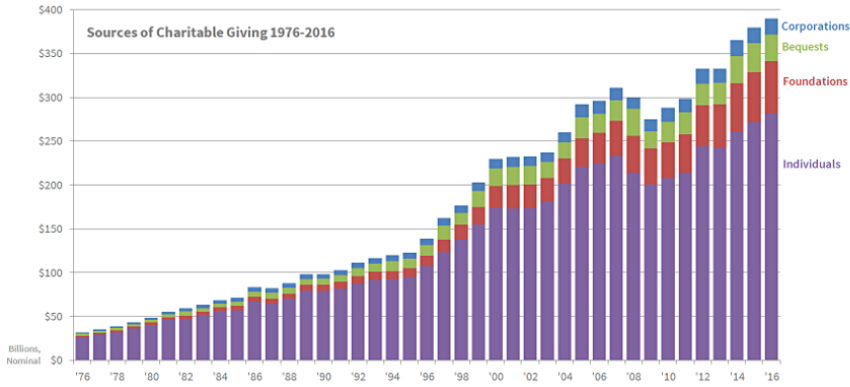
- www.colorbrewer2.org
- www.sessions.edu/color-calculator/
- color.adobe.com

Color rules:

- **Avoid red-green proximity**
- **Avoid colors of equal weight**
- **5 is the max, maybe 6**
- **Use the right associations**

Colors of equal weight

Colors of unequal weight



Giving USA data 2017.xlsx - Excel

File Home Insert Page Layout Formulas Data Review View Design Format Tell me what you want to do

Series "Corporations" Format Selection Reset to Match Style Current Selection Insert Shapes Shape Styles WordArt Styles Arrange Size

Chart 7: =SERIES(Sheet1!\$D\$2, Sheet1!\$M\$4:\$M\$44, Sheet1!\$D\$4:\$D\$44, 4)

Sources of Charitable Giving 1976-2016

- Corporations
- Bequests
- Foundations
- Individuals

Billions, Nominal

Format Data Series

Series Options

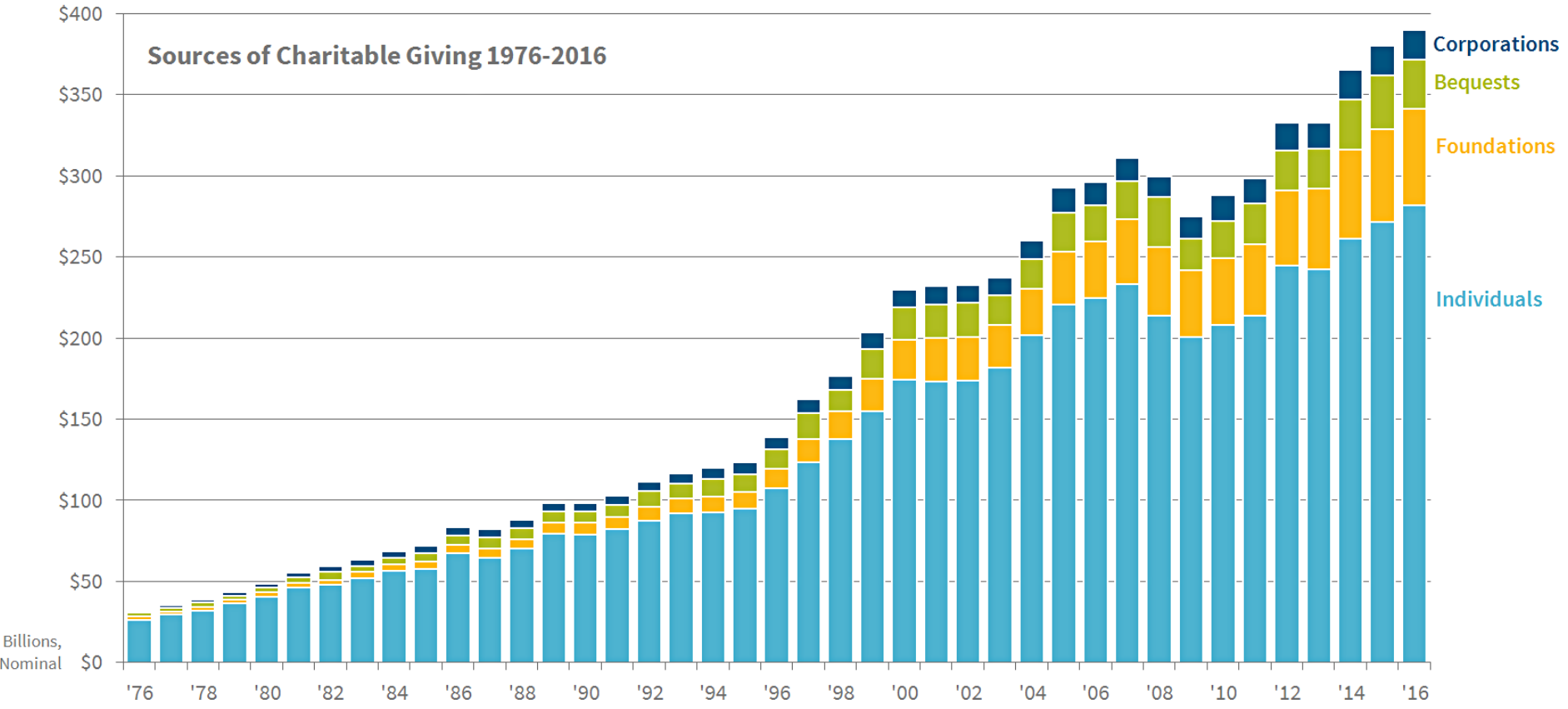
- Fill
- Border
 - No line
 - Solid line
 - Gradient line
 - Automatic
- Color
- Transparency 0%
- Width 1 pt
- Compound type
- Dash type
- Cap type Flat
- Join type Round

Ready

▲
**Right click
 on something**

▲
**Change its color, width,
 or make it appear**

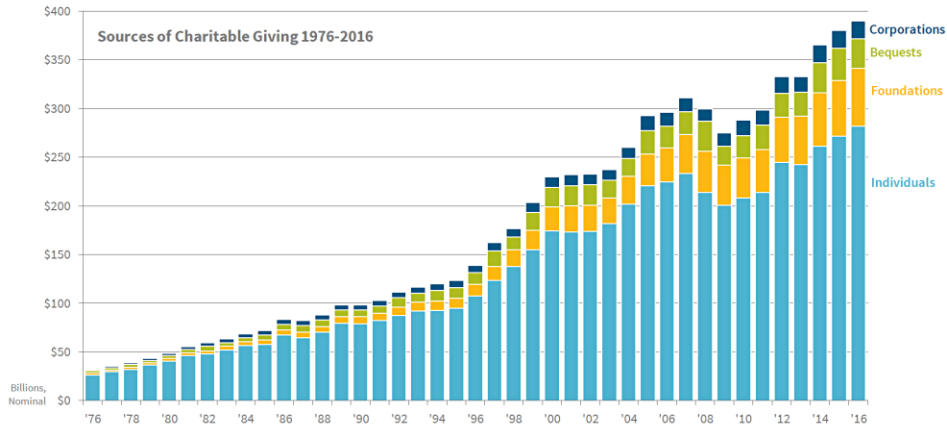
Choose the right color scheme; clearly distinguish among series.



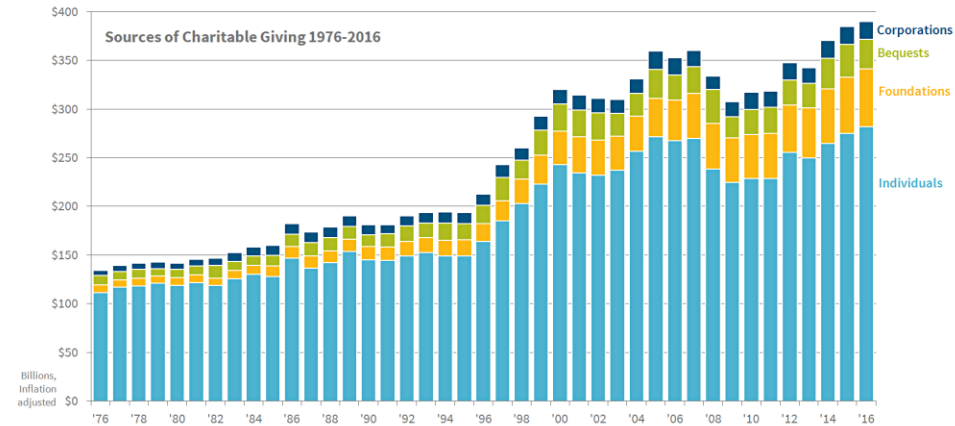
Data source: *Giving USA 2017*, Giving USA Foundation

Ask whether the data require adjustment.

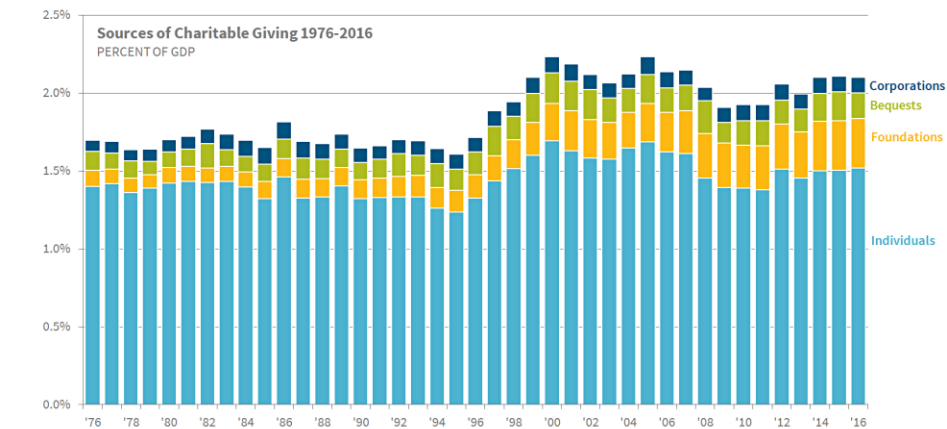
Nominal or current



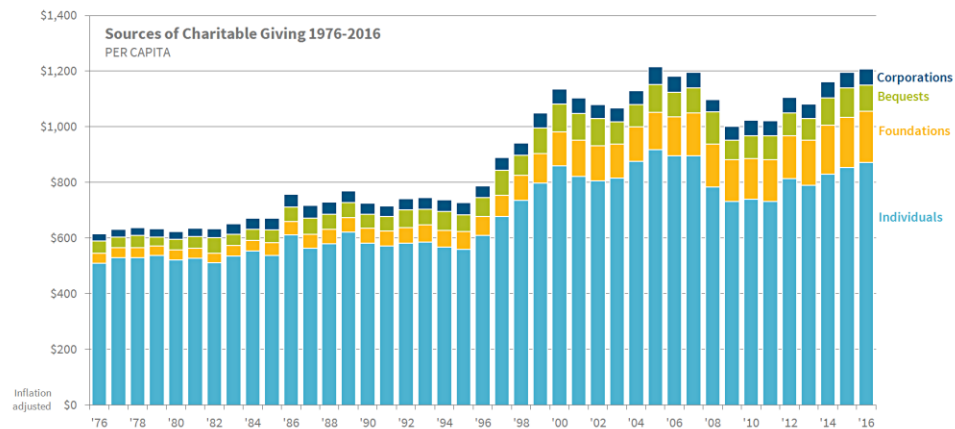
Inflation adjusted or real

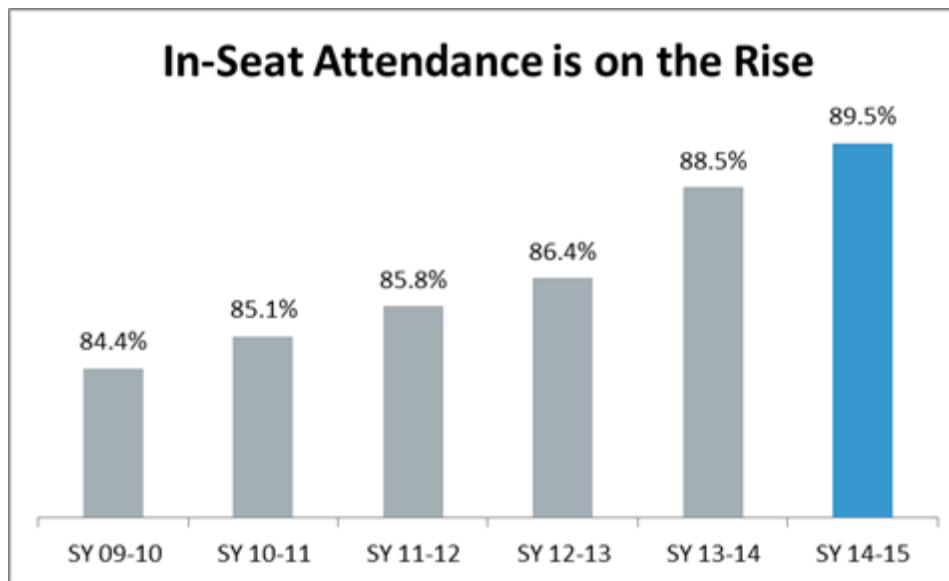
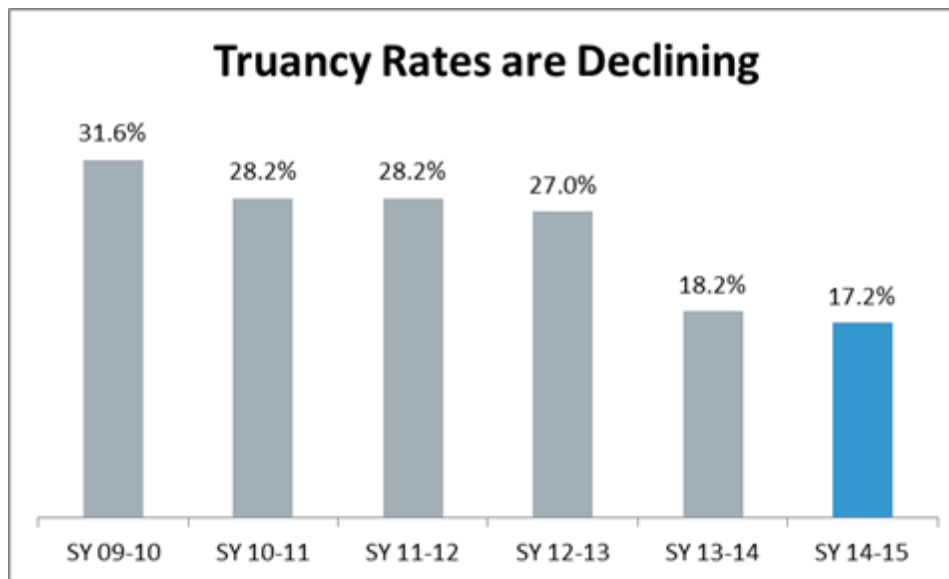


Percent of GDP



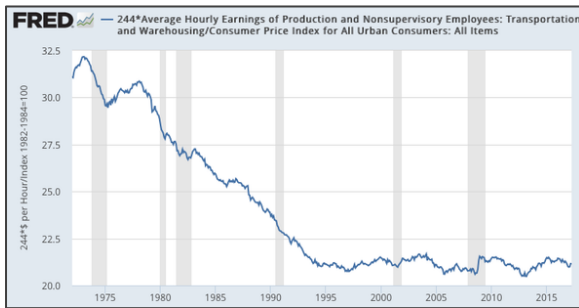
Per capita





Source: "DCPS at a Glance: Attendance and Truancy", District of Columbia Public Schools website (<https://dcps.dc.gov/page/dcps-glance-attendance>) (retrieved August 2017)

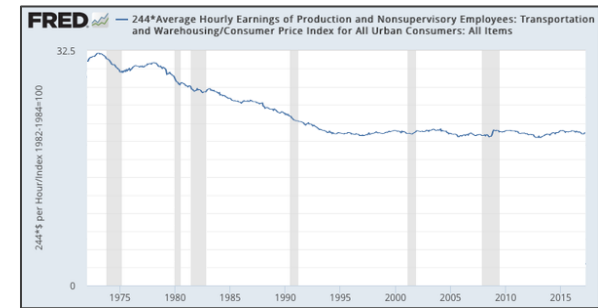
X axis = 20



X axis = 0



X axis = 0

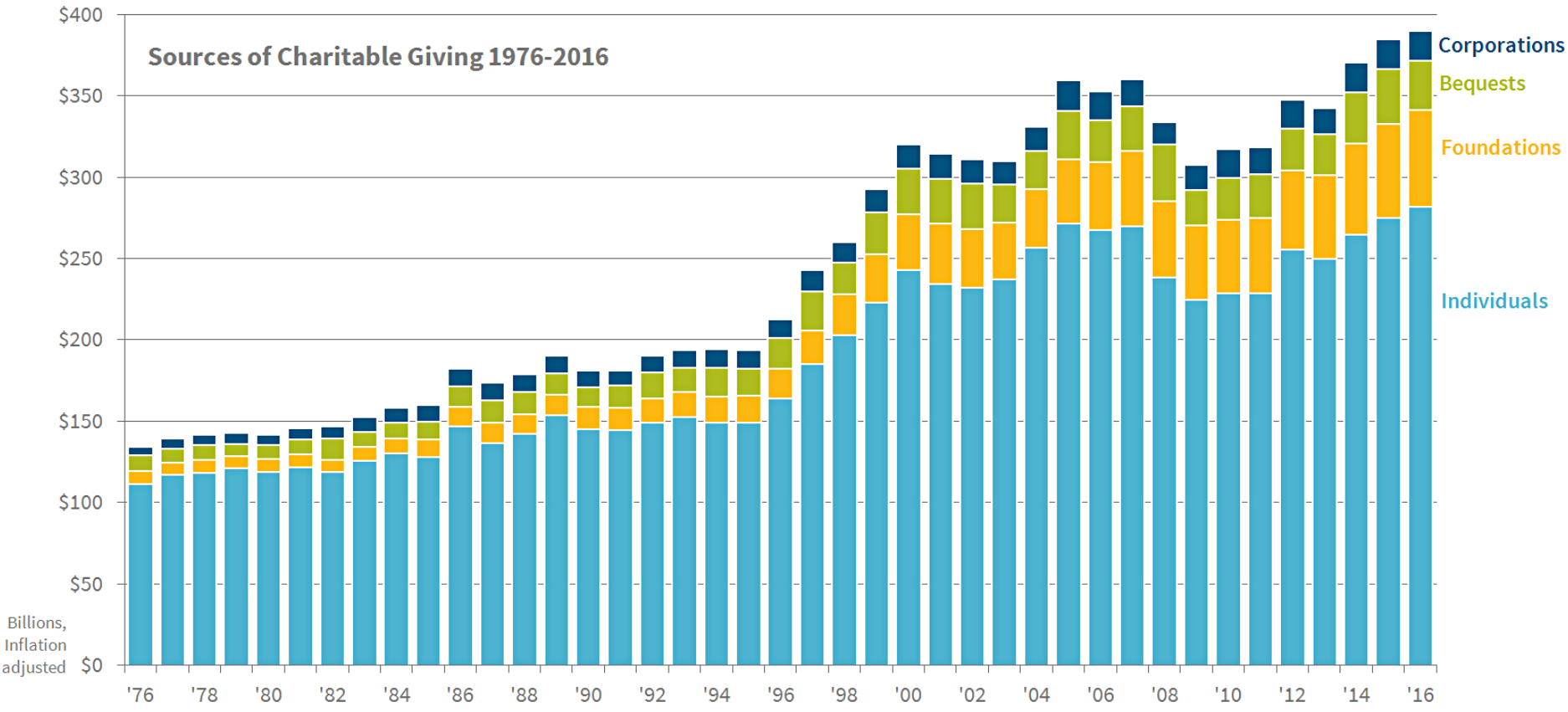


Source: "Trucking and Blue-Collar Woes",
*The New York Times/The Conscience of
a Liberal blog* (May 2017)

For bar charts, the X axis should always be at 0.

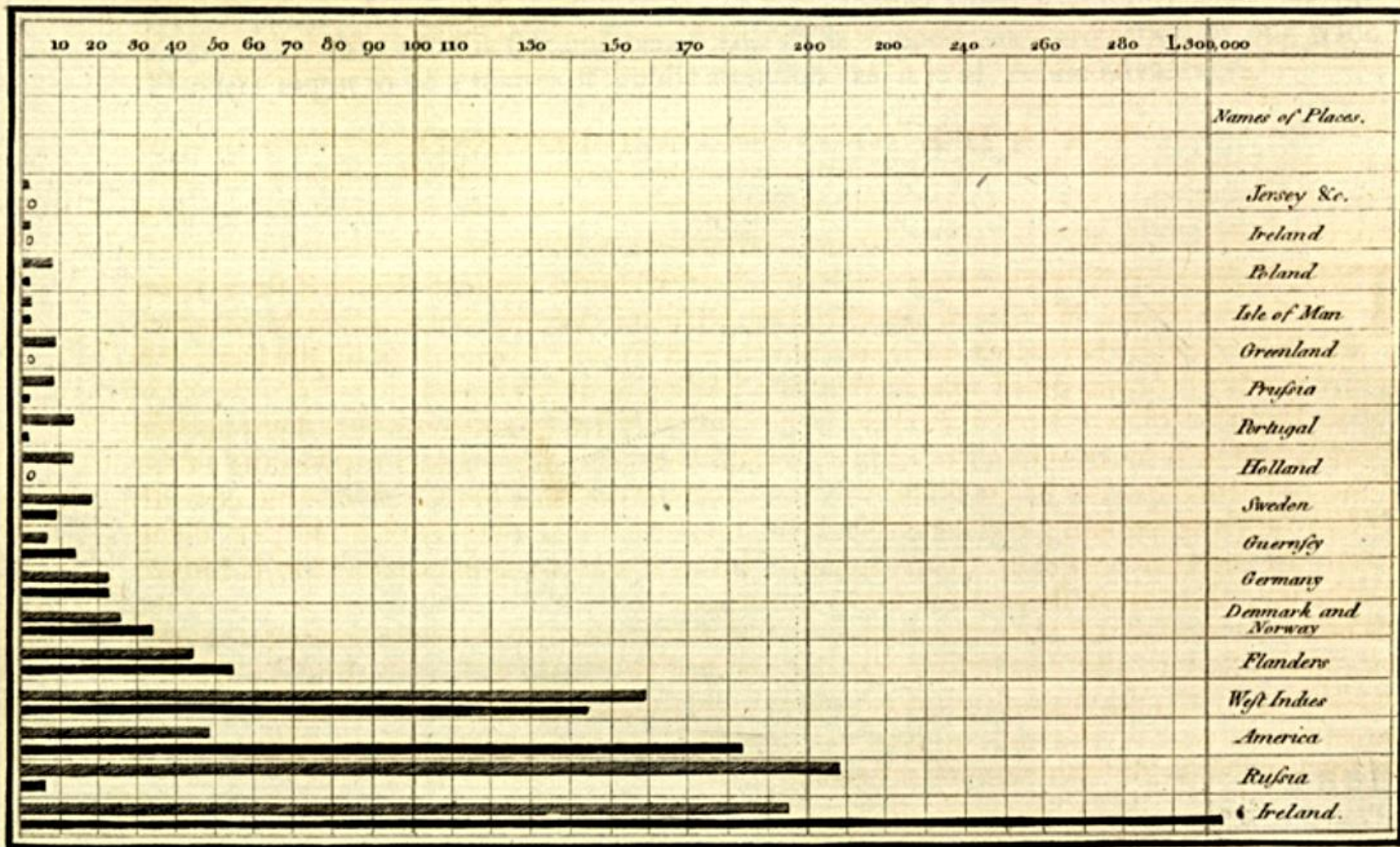
Same for line graphs, with the following exceptions:

- **For specialized audiences.**
- **When comparing two trend lines.**
- **When absolute level doesn't matter.**



Data source: *Giving USA 2017*, Giving USA Foundation

Exports and Imports of **SCOTLAND** to and from different parts for one Year from Christmas 1780 to Christmas 1781.



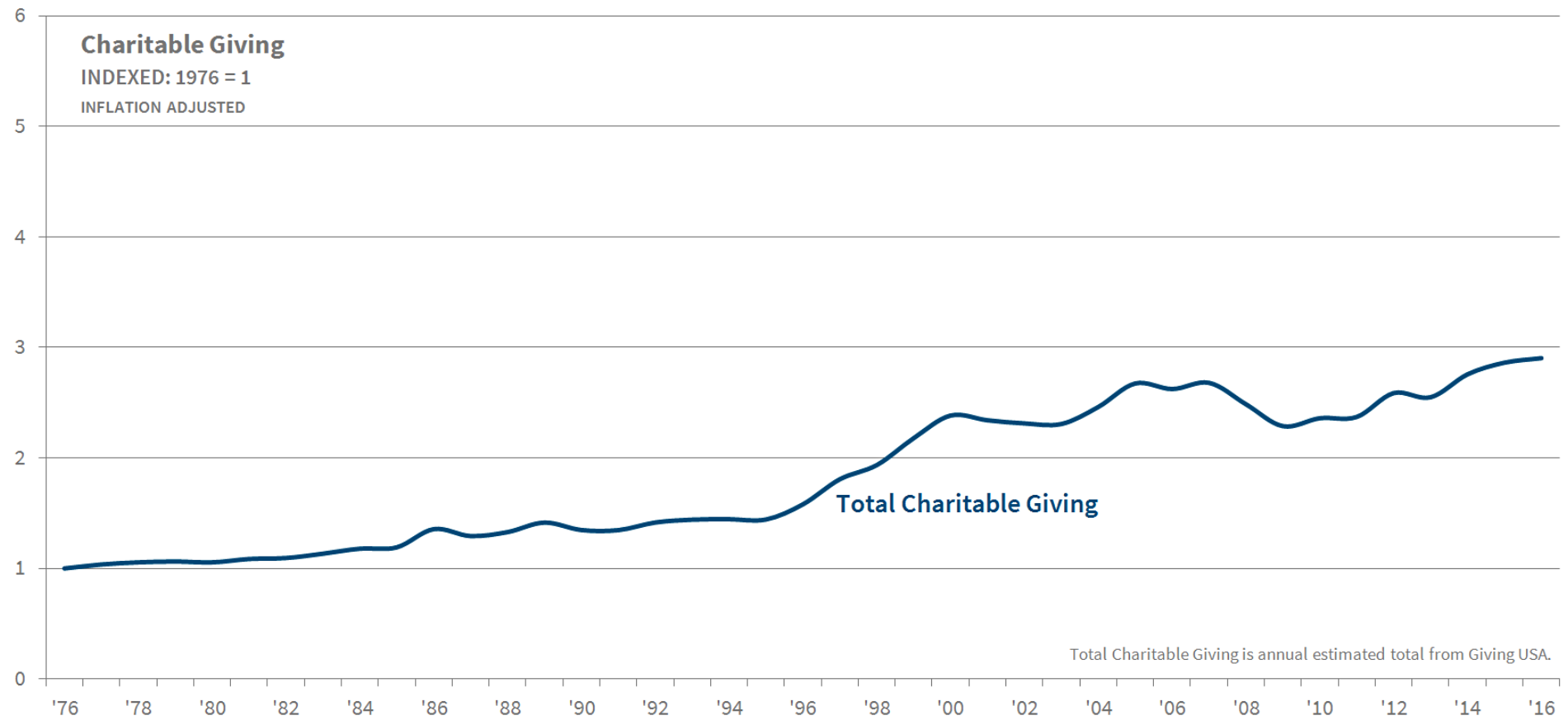
The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbedlines Imports.

Source: *The Commercial and Political Atlas*, William Playfair (1786)

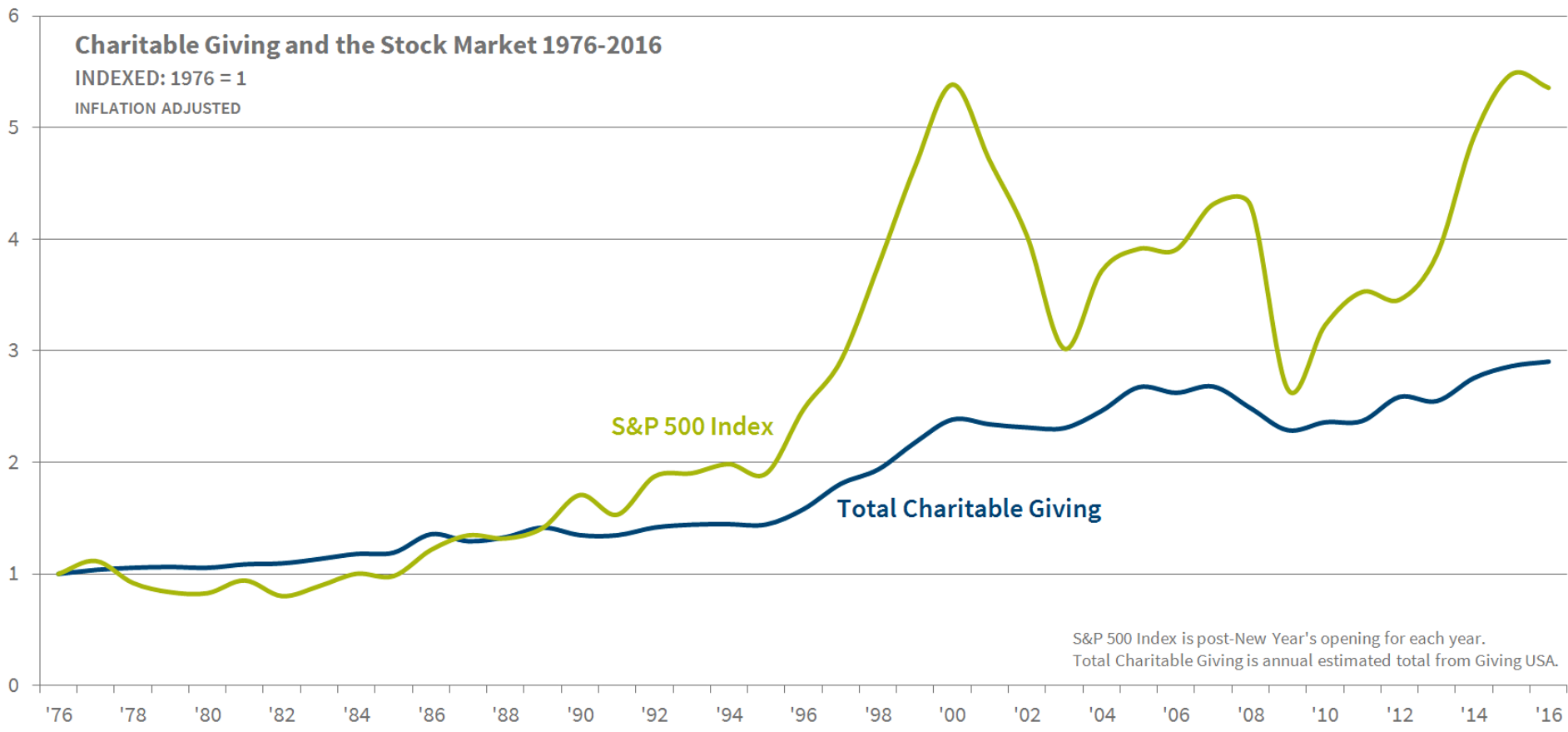


**Standard data
and clarity**

Now, see if you can add meaning by adding data.



Now, see if you can add meaning by adding data.

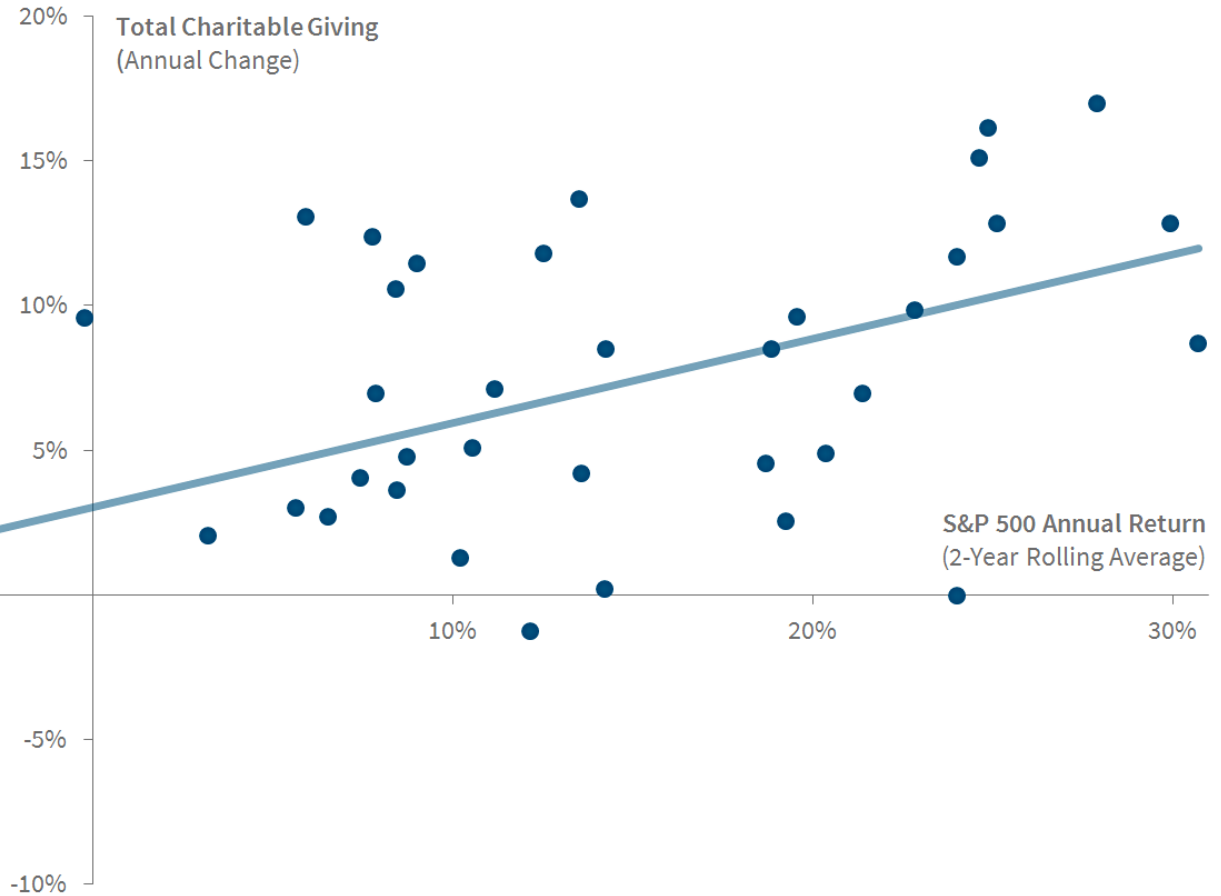


Data sources: *Giving USA 2017*, Giving USA Foundation; Historical Data, Yahoo! Finance

Finally, try a more relational chart type, like a scatterplot.

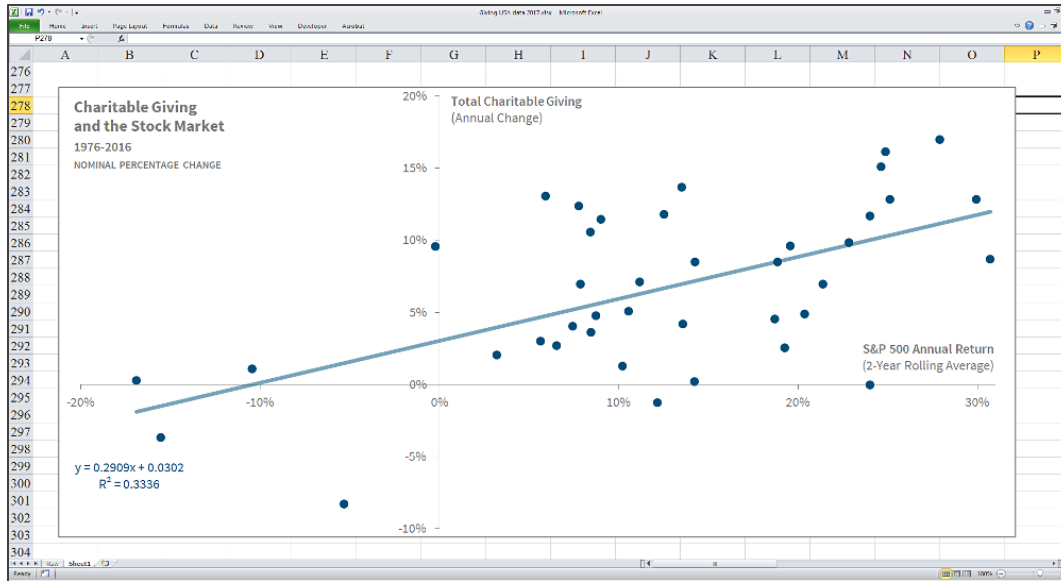
Charitable Giving and the Stock Market

1976-2016
NOMINAL PERCENTAGE CHANGE



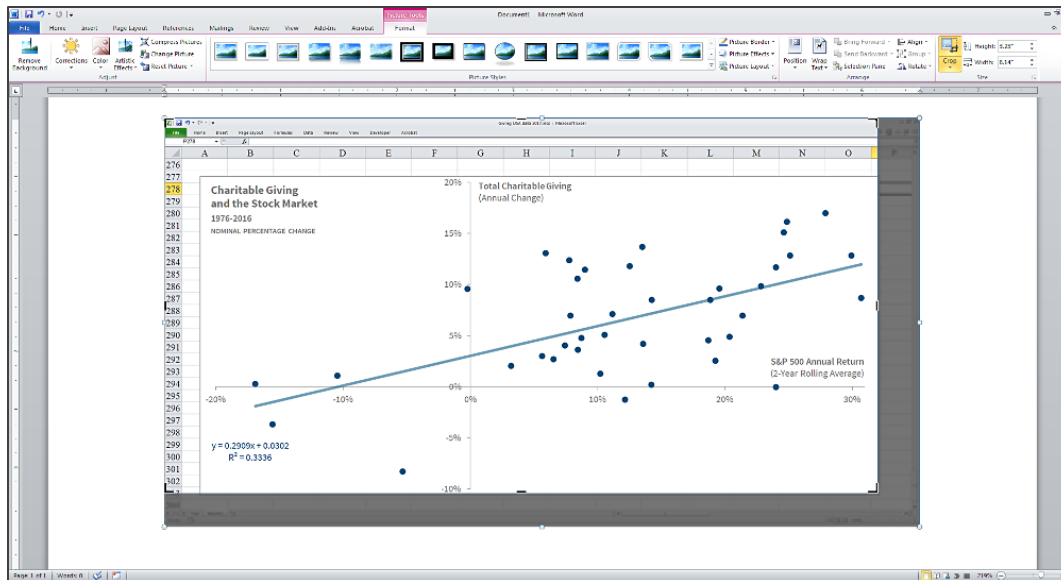
Data sources: *Giving USA 2017*, Giving USA Foundation; Historical Data, Yahoo! Finance

Last thing: Take a screenshot (Alt + PrnScn) and paste the chart in a document.



In Excel:

- **Maximize window**
- **Zoom in ~200%**
- **Alt + Print Screen**



In Word:

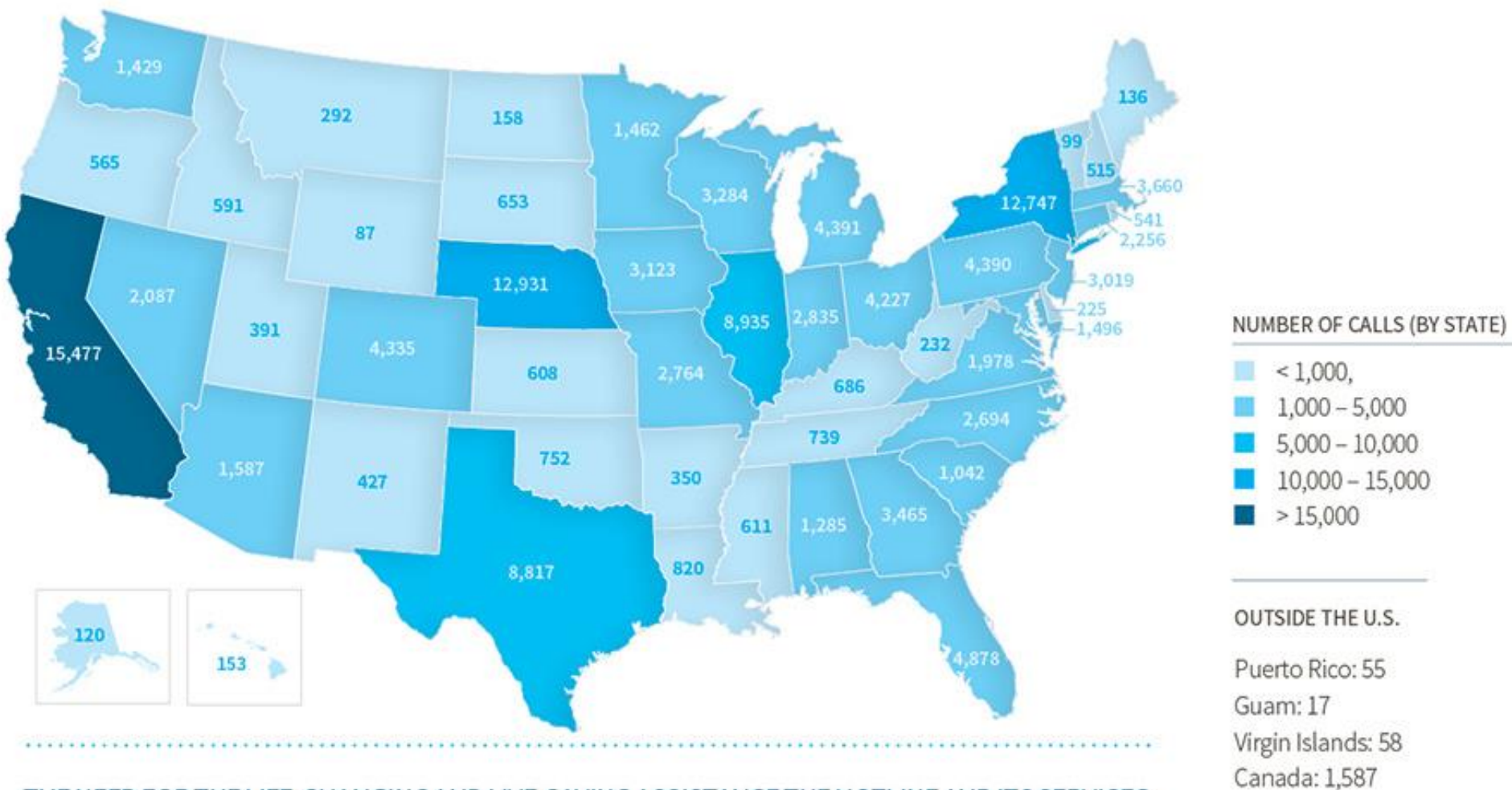
- **Paste**
- **Crop**
- **Adjust size**

Your next steps:

1. Choose your key message (urgency or effectiveness).
2. Do you have data to support this message?
3. Is there enough interesting data to make a chart?
4. Make a simple chart in Excel (bar, line, or scatterplot).
5. Prune, prioritize, choose right colors, adjust data.
6. Do you have more data (show causes or relationships)?
7. Add data or labels to strengthen key message.
8. Prune, prioritize, and test. Do readers understand?

131,959 Hotline Calls* Received in 2014

* Calls only; does not include emails, texts or chats



THE NEED FOR THE LIFE-CHANGING AND LIVE-SAVING ASSISTANCE THE HOTLINE AND ITS SERVICES PROVIDE IS CRITICAL, FAR-REACHING AND MORE COMMON THAN YOU MIGHT IMAGINE.



In 2015, counselors at the Boys Town National Hotline received calls, e-mails, instant messages, and mobile texts from teens and parents in all 50 states and the Nation's Capital.

5,564

MOBILE TEXT CONVERSATIONS

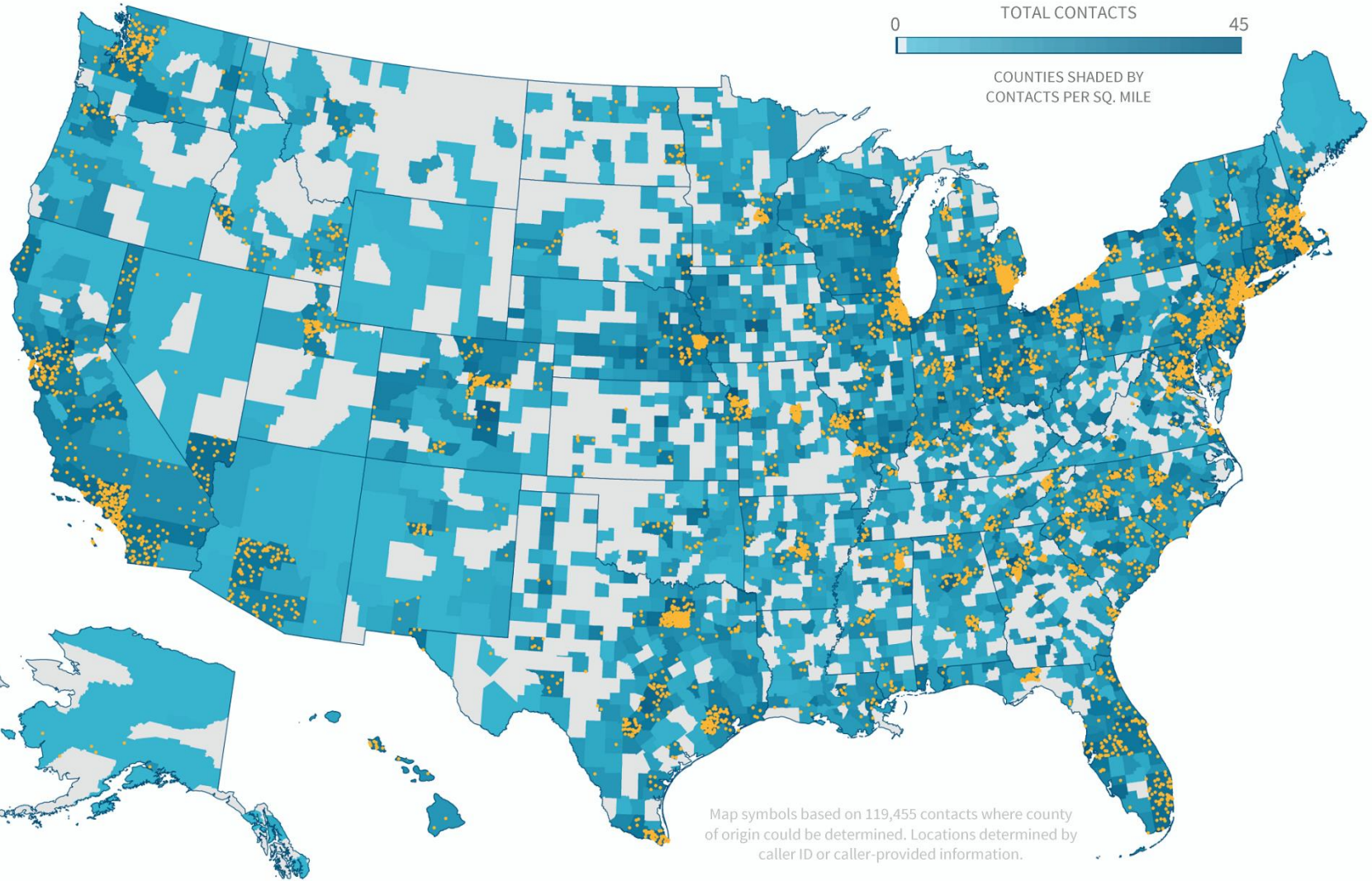
••• - 1 DOT PER CONVERSATION

184,222

TOTAL CONTACTS



COUNTIES SHADED BY CONTACTS PER SQ. MILE



Map symbols based on 119,455 contacts where county of origin could be determined. Locations determined by caller ID or caller-provided information.

Families Served

OMAHA METRO
2013-2016

Boys Town's offices in North and South Omaha make services more accessible for the city's most distressed families. In South Omaha, school partnerships reach more families with preventive, family-strengthening services. Concentrated services promise to make a big, enduring impact on family stability and child well-being.

Families served by:

- — Care Coordination
- — Parent Connectors
- — Common Sense Parenting®
- — In-Home Family Services®

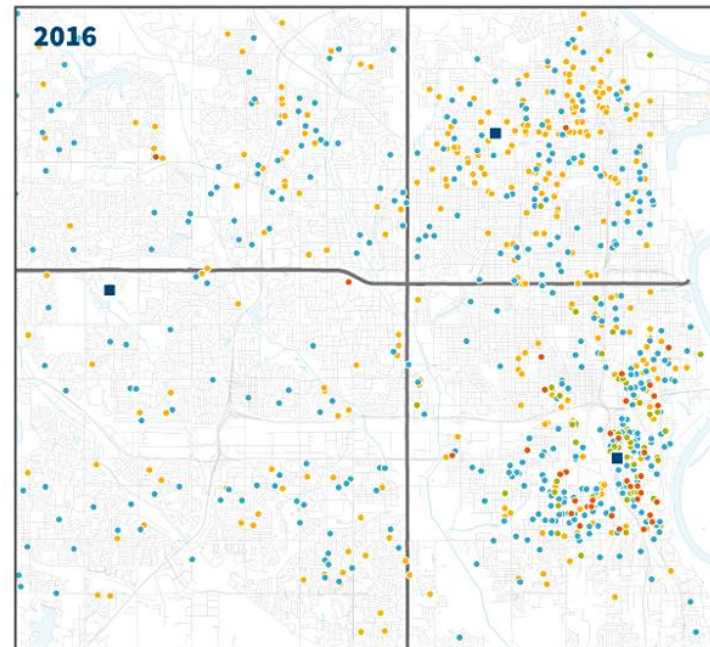
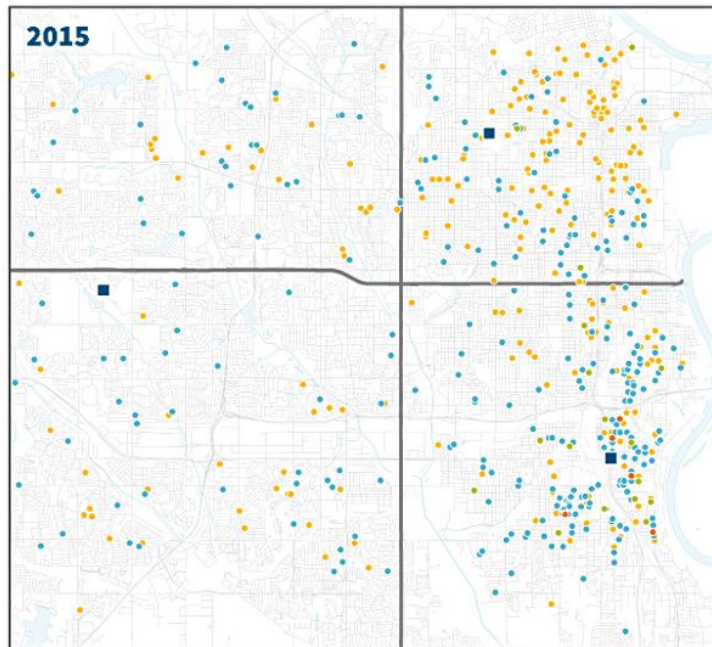
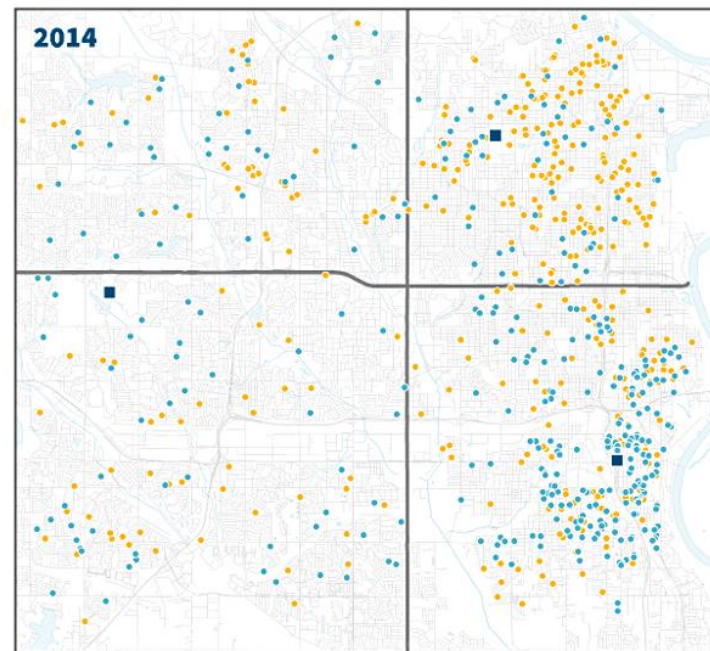
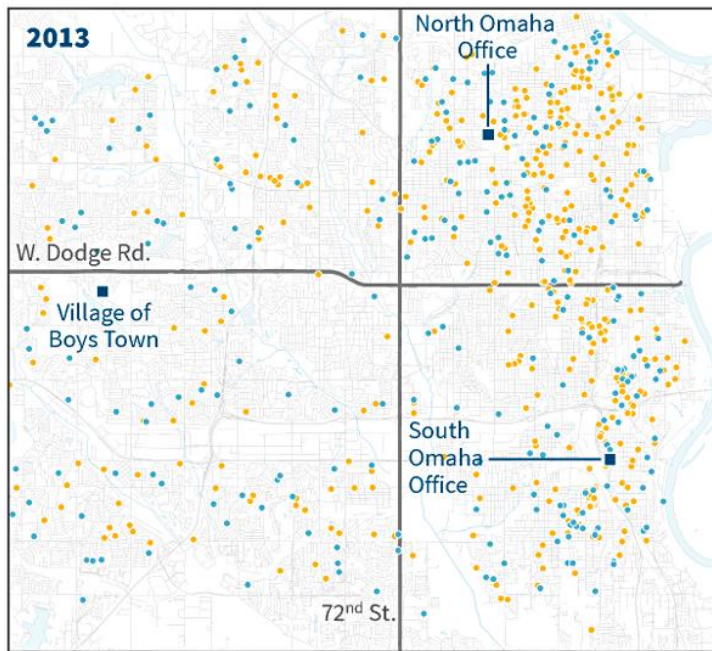
North Omaha Office

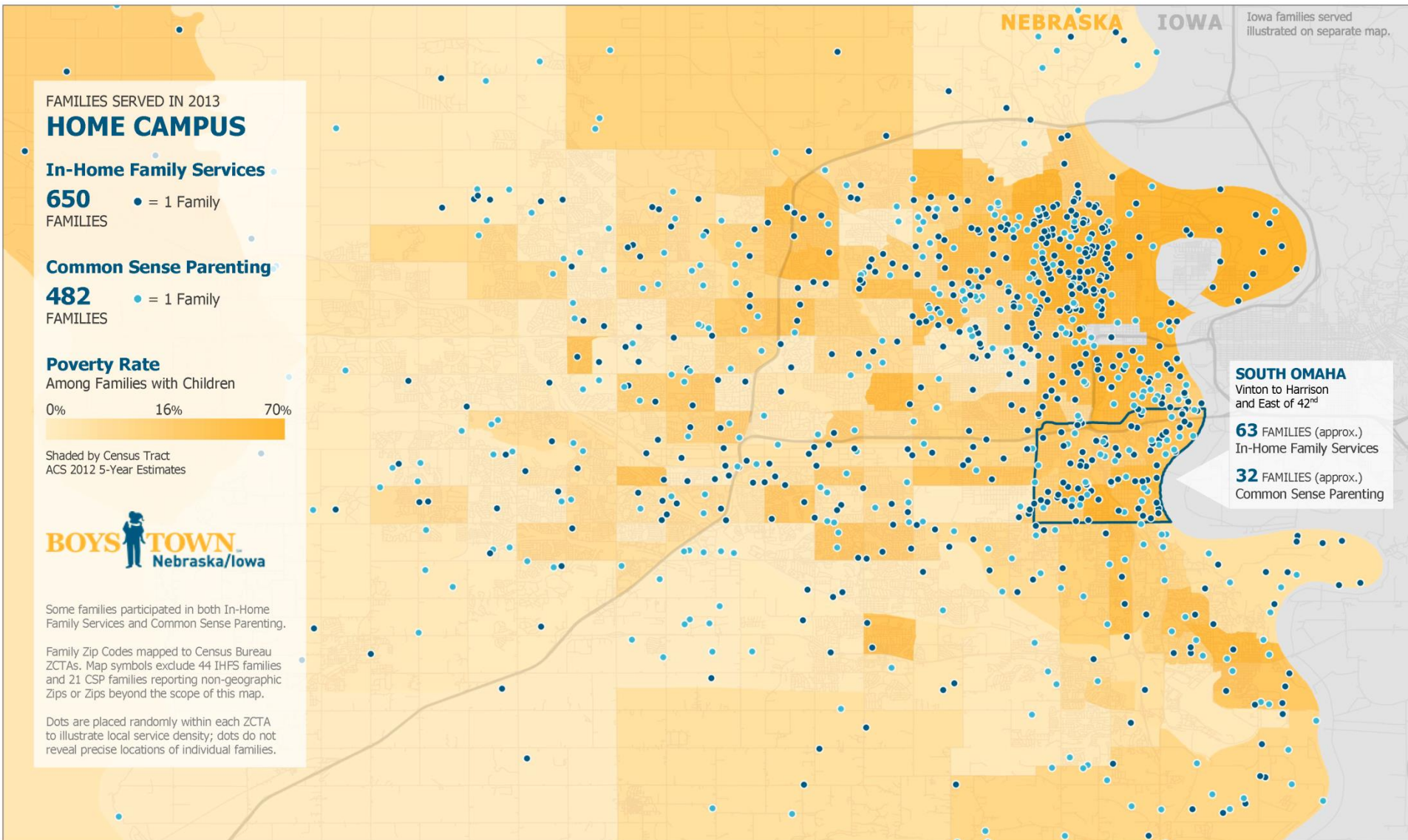
5076 Ames Avenue
(Walmart frontage)

South Omaha Office

4612 S. 25th Street
(25th and L streets)

Note: Visualization based on 3,341 residences of families served from 2013 to 2016.





Data Sources: Boys Town National Database (2013);
 U.S. Census Bureau, American Community Survey 5-Year Estimates (2012)



South Omaha Continuum of Care

OUTREACH

South High School

School staff help Boys Town engage the families of struggling students.

Walk-ins

Parents hear about Continuum services and come to Boys Town's office at 25th and L.

Helpline

Boys Town takes crisis calls through this statewide hotline. South Omaha families are connected to local services.

Partnerships

Reach more families through partners, including the Learning Community, OneWorld Health, the City of Omaha, and gang-prevention specialists.

SCHOOL PARTNERSHIPS

Boys Town also trains teachers in classroom behavior management and provides an alternative learning environment for students who have been suspended from South High.

TRIAGE

Rapid Assessment

Check-in and overview of less-intensive services.

Family Advocacy & Support Tool

Asses family's stability, needs, and resources. Prioritize goals and select the right services.

Behavioral Health Assessments

Determine if child or family could benefit from clinical behavioral health treatment.

MORE
INTENSIVE

CONTINUUM SERVICES

Common Sense Parenting Classes

Convenient help and support for all parents.

Parent Connectors

Peer mentoring for parents of struggling students.

Care Coordination

Consultation and connections to increase family stability.

Ways to Work

Financial coaching and help in securing reliable transportation.

In-Home Family Services

Consultation and intensive training in parenting skills.

Behavioral Health Care

Clinical treatment for childhood behavioral health problems.

IMPACT

Strong Families

Children are living safely at home, where they belong.

Thriving Children

Children are attending school, progressing toward graduation, and staying out of trouble.

Healthy Communities

Build infrastructure and relationships, partner with more schools, and measure the project's collective impact.

Recommended Reading

Longer,
more
in-depth

The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures

By Dona M. Wong

(2010) (160 pages)

The Visual Display of Quantitative Information

By Edward R. Tufte

(2001, 2nd Edition) (197 pages)

Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations

By Scott Berinato

(2016) (264 pages)

The Functional Art: An Introduction to Information Graphics and Visualization

By Alberto Cairo

(2013) (350 pages)

The Truthful Art: Data, Charts, and Maps for Communication

By Alberto Cairo

(2016) (400 pages)

Training & Tools

Microsoft Excel help center

<http://support.office.com/excel>

Lynda.com Excel training

(Check your local library)

Datawrapper web-based chart creator

<https://www.datawrapper.de/>

Tableau desktop data visualization application

<https://www.tableau.com/>

My Maps in Google

<https://www.google.com/mymaps>

MapWindow open-source GIS mapping

<http://www.mapwindow.org/>

Esri ArcGIS desktop mapping applications

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

More
advanced

