



# CS-GY 6313 B: Information Visualization

11/13/2024

# Logistics

- Assignment 4 due Nov 20 @ 11:59pm
- Guest lecture moved to Dec 5:
  - Dr. David Saffo and Dr. Benjamin Lee @ JPMC
  - 3D + immersive viz
  - Bring your laptop!
- Today:
  - Misinformation + ethics
  - Narrative storytelling
  - User study crash course



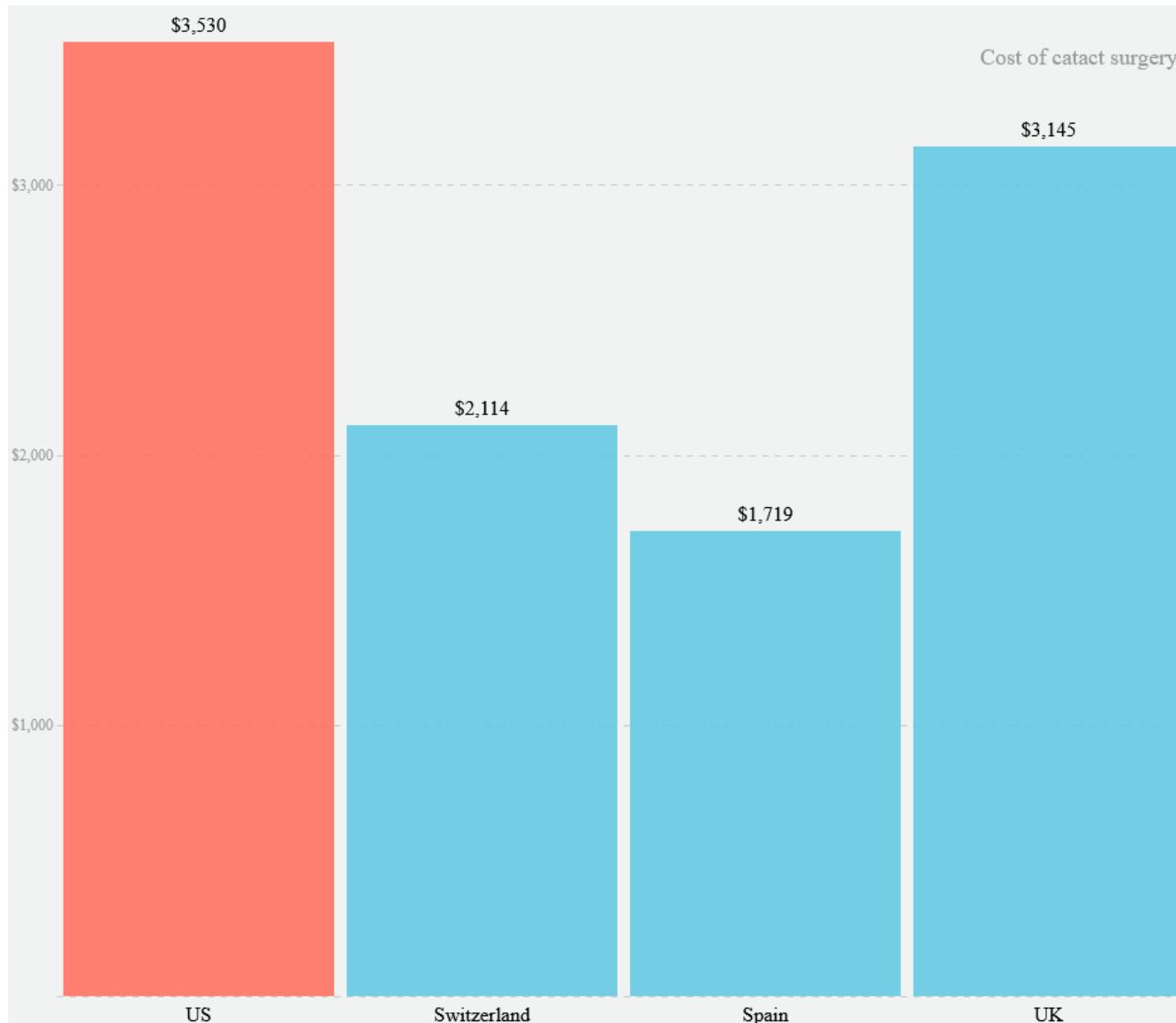
# Misinformation & Ethics

# Misinformation

- 5 rules to not get mislead:
  - Is the author using the correct data and disclosing the source?
  - Are you reading too much into the graphic?
  - Are the data represented accurately?
  - Is the visualization showing an appropriate amount of data?
  - Is uncertainty relevant? If yes, is it shown?

# Misinformation

- Is the author using the correct data and disclosing the source?



<https://www.vox.com/a/health-prices#chart/9>

# Misinformation

- Is the author using the correct data and disclosing the source?



International Federation  
of Health Plans

2015 Comparative Price Report  
Variation in Medical and Hospital Prices by Country

## 2015 Survey Overview

This year's survey has been conducted in much the same way as the previous study, which includes pricing for several specialty prescription drugs, other prescription drugs and a selection of typical medical procedures. Prices for each country were submitted by participating federation member plans, and are drawn from public or commercial sectors as follows:

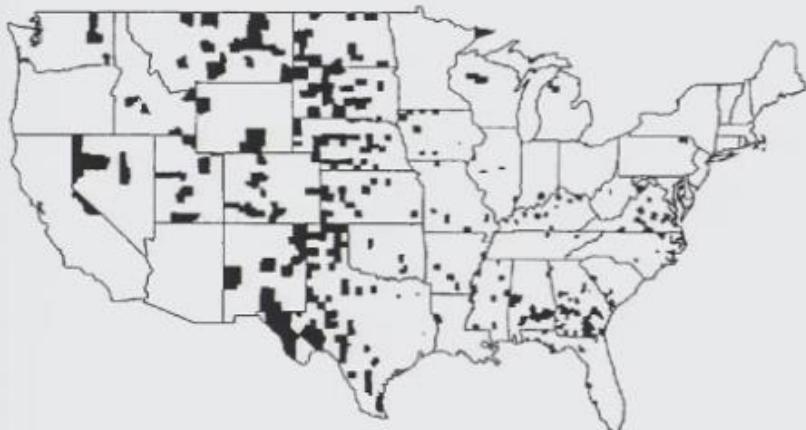
- Prices for the United States were derived from over 370 million medical claims and over 170 million pharmacy claims that reflect prices negotiated and paid to health care providers.
- Prices for Australia, New Zealand, Spain, South Africa, Switzerland and the UK are from the private sector, with data provided by one private health plan in each country.

# Misinformation

- Are you reading too much into the graphic?

Always tell yourself:

“A chart shows as much as it hides —think about what might be hidden”



Counties with the LOWEST  
kidney cancer death rates  
(1980-1989)

# Misinformation

- Are the data represented accurately?

The screenshot shows the homepage of the Venezuelan television network's website. The top navigation bar includes links for 'INICIO', 'NOTICIAS', 'PROGRAMAS', 'MULTIMEDIA', 'OPINIÓN', 'EL CANAL', 'ELECCIONES', and 'HUGO CHÁVEZ'. A banner at the top reads '¡ide verdad!' and 'VENEZUELA'. The main content area features a large graphic titled 'ELECCIÓN PRESIDENCIAL 2013' with two large 3D bars representing percentages. The red bar is labeled 'Nicolás Maduro' and '50.66%', while the blue bar is labeled 'Henrique Capriles' and '49.07%'. To the right, a headline states 'Ganó el hijo de Chávez: Nicolás Maduro es el Presidente Electo de Venezuela (+Video)'.

Candidato	Porcentaje
Nicolás Maduro	50.66%
Henrique Capriles	49.07%

**Ganó el hijo de Chávez: Nicolás Maduro es el Presidente Electo de Venezuela (+Video)**

La tendencia es irreversible, señaló Tibisay Lucena, presidenta del ente comicial / Maduro obtuvo 50.66% v Capriles 49.07%

[Leer más](#)

**Nicolás Maduro**

**Henrique Capriles**

**ELECCIONES**

**Hackedadas cuentas Twitter del Presidente Nicolás Maduro y del Pdvsa**

**Nicolás Maduro: Hoy es un día de Historia, vemos con Cristo Redentor (+Fotos)**

**Telvisión venezolana**

# Misinformation

- Are the data represented accurately?

The screenshot shows the homepage of the website of Venezolana de Televisión. The top navigation bar includes links for INICIO, NOTICIAS, PROGRAMAS, MULTIMEDIA, OPINIÓN, EL CANAL, ELECCIONES, and HUGO CHÁVEZ. The date 'Lunes, 15 de Abril 2013' is displayed at the top right. A search bar is also present.

The main content area features a large graphic for the 'ELECCIÓN PRESIDENCIAL 2013'. The graphic includes two large portraits: one of Nicolás Maduro on the left and one of Capriles Radonski on the right. Between them is a red 3D bar chart showing the percentage results: 50,66% for Maduro and 49,07% for Capriles. The text 'PORCENTAJES' is written above the chart.

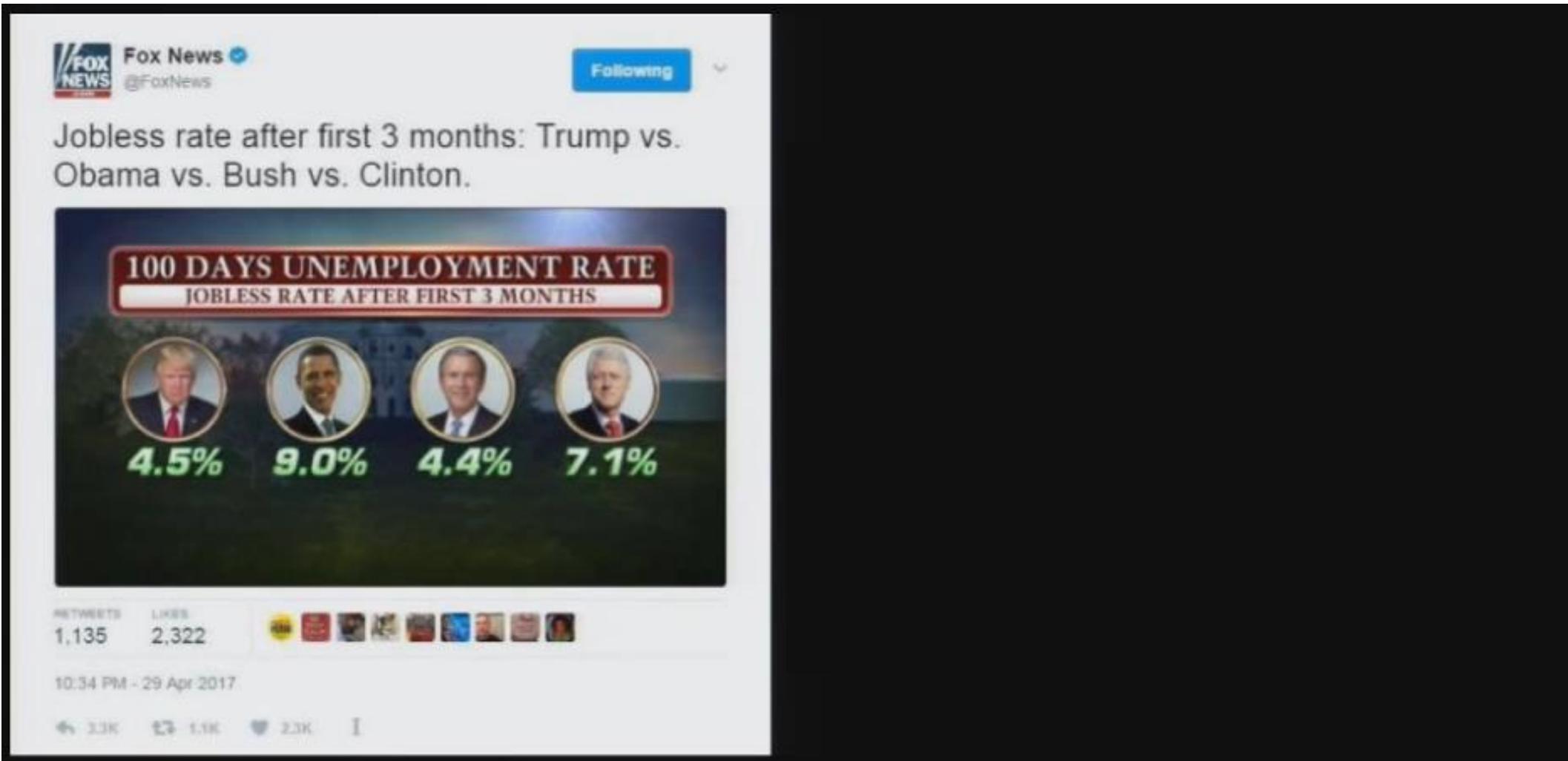
To the right of the graphic, a large headline reads: 'Ganó el hijo de Chávez: Nicolás Maduro es el Presidente Electo de Venezuela (+Video)'. Below this, a sub-headline says 'La tendencia es irreversible, señaló Tibisay Lucena, presidenta del ente comicial / Maduro obtuvo 50.66% y Capriles 49.07%'.

At the bottom of the page, there are several smaller news cards and a footer section.

Candidato	Porcentaje
Nicolás Maduro	50,66%
Capriles Radonski	49,07%

# Misinformation

- Is the visualization showing an appropriate amount of data?



# Misinformation

- Is uncertainty relevant? If yes, is it shown?

The screenshot shows a news article from EL PAÍS IN ENGLISH. The header includes the EL PAÍS logo, a navigation menu, and links for POLITICS, ARTS, SPORTS, OPINION, and SPANISH WAY OF LIFE. The article is categorized under REGIONAL POLITICS. The main headline reads: "Catalan public opinion swings toward ‘no’ for independence, says survey". A subtitle below the headline states: "Growing numbers of citizens support federal model that falls short of secession from Spain". The author is PERE RÍOS, and the date is 19 DIC 2014 - 10:23 EST. Social sharing icons for Facebook, Twitter, and LinkedIn are at the top, along with a comment icon. At the bottom, there are icons for liking, sharing, and saving. The URL of the article is [http://ccaa.elpais.com/ccaa/2014/12/19/catalunya/1418984873\\_128596.html](http://ccaa.elpais.com/ccaa/2014/12/19/catalunya/1418984873_128596.html).

# Misinformation

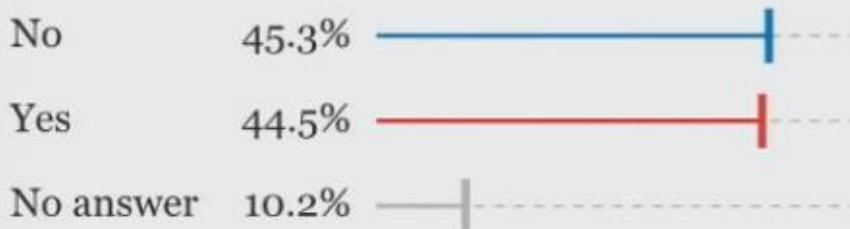
- Is uncertainty relevant? If yes, is it shown?

For the first time since Catalan leader Artur Mas began his ongoing independence drive in 2012, a survey shows that a majority in the region would reject secession if a referendum were held now.

[http://www.thefunctionalart.com/  
2014/12/adventures-in-margin-of-  
error.html](http://www.thefunctionalart.com/2014/12/adventures-in-margin-of-error.html)

The latest poll by the Catalan executive's Opinion Studies Center (CEO) shows that 45.3 percent of citizens would vote no to the question: "Would you like Catalonia to become an independent state?" compared with 44.5 percent who would support the move.

*Do you want Catalonia  
to become an independent state?*

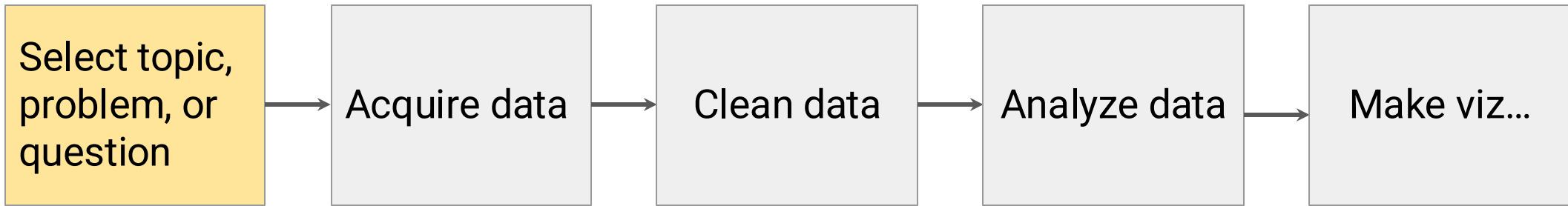


Margin of error: +/-2.95 at 95% confidence level

# Ethics

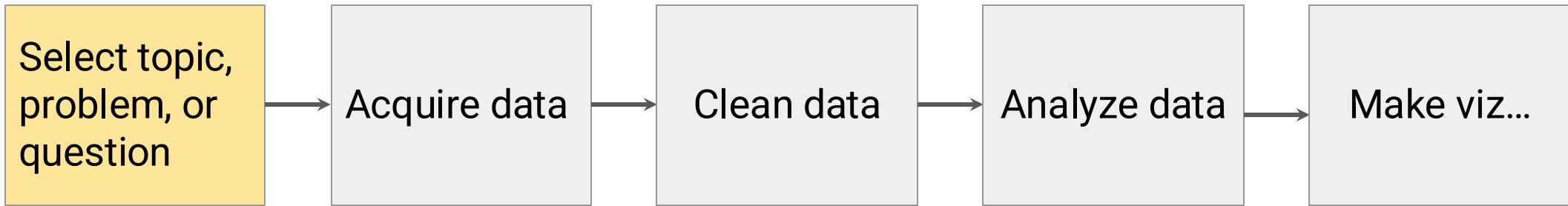
- What is ethics?
  - “a set of moral principles, especially ones relating to or affirming a specified group, field, or form of conduct.” –Google
- Ethics are not fixed
  - If possible, be up-front about your ethics
  - Morals are determined individually, laws are geopolitical
- No single right way to do data viz
  - Several wrong ways!
  - Must avoid these wrong ways
- Data are **not** neutral!

# Ethics: Data Viz Process



- Why is this study being done?
- Who benefits from the outcome of this study?
- Who might be harmed?
- Am I the right person to do this research?
- If appropriate, have I preregistered my research?

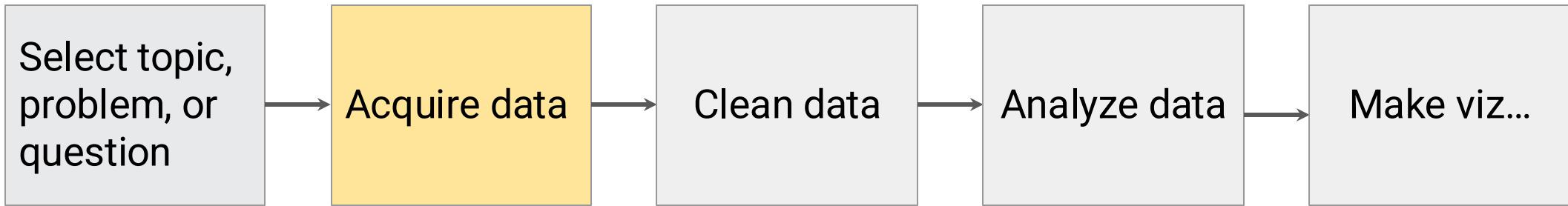
# Ethics: Data Viz Process



- **Exclusion because of research scope:**

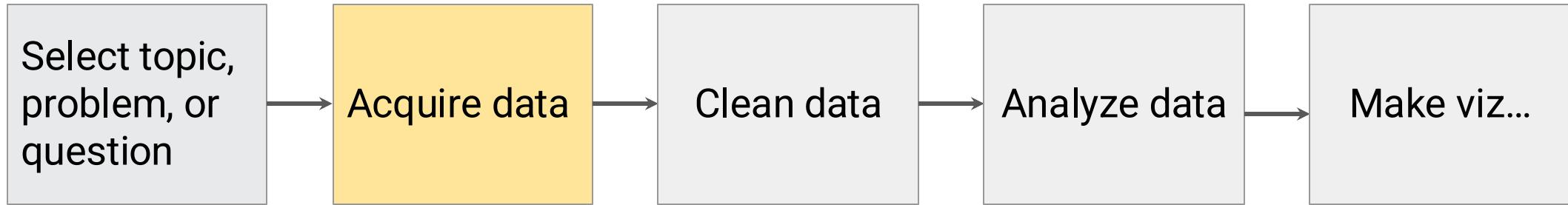
- Focusing a study of heart attack symptoms on symptoms that are more common for men, like chest pain
- Focusing a study of child cereal preference on healthy cereals but not accounting for nut allergies
- Focusing a study of library satisfaction on physical spaces, ignoring distance students

# Ethics: Data Viz Process



- What choices were made about collection method, participant selection, questions included, etc.? What influenced those choices?
- Are the people represented by the data being treated with dignity? Have they been included in the decision-making process? Are you minimizing the burden and risk placed on them?
- Will the data you have answer your research question?
- Have you cited the source of any secondary data?

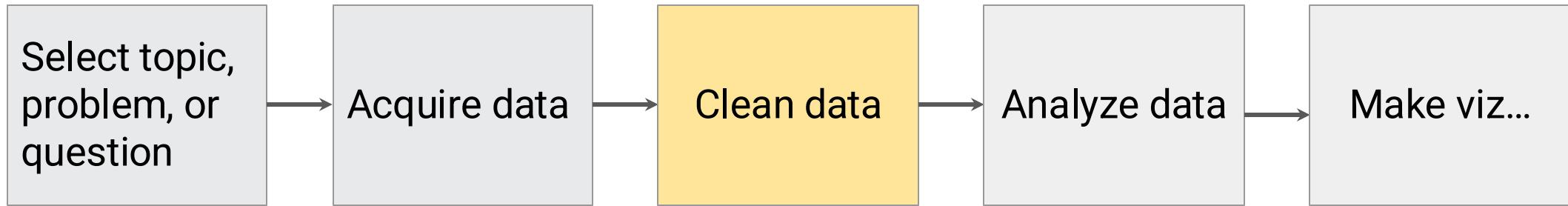
# Ethics: Data Viz Process



- **Bias introduced by taking shortcuts:**

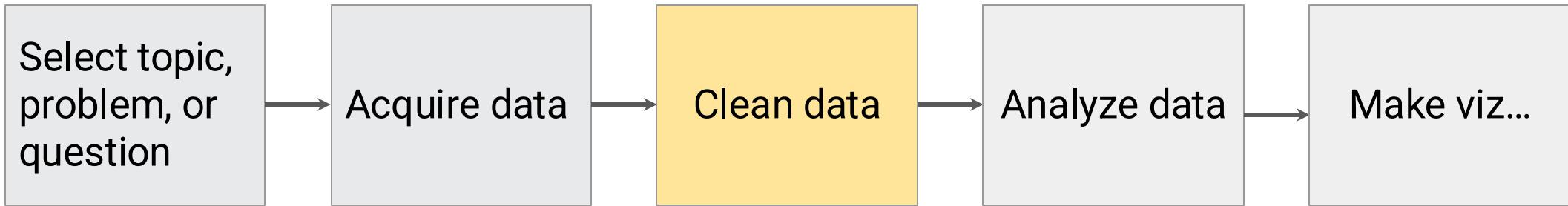
- Conducting psychology experiments only with students enrolled in introductory psychology courses, ignoring the bias that introduces in the population (age, gender, educational background, race/ethnicity, etc.)
- Participant exclusion because focus group is held in the evening when nontraditional students might have caretaking responsibilities
- Participant exclusion by advertising study only to certain listservs
- Exploring community dynamics on Twitter but only looking at public tweets

# Ethics: Data Viz Process



- What assumptions are you making about the data? Are there other interpretations?
- Are you removing any data from the analysis? Will that introduce systematic bias into the analysis?
- Are you simplifying data for the analysis? How does the loss of that complexity influence your results? Are your participants still being well represented by the data?
- Have you given appropriate credit to anyone who helped with data cleaning?

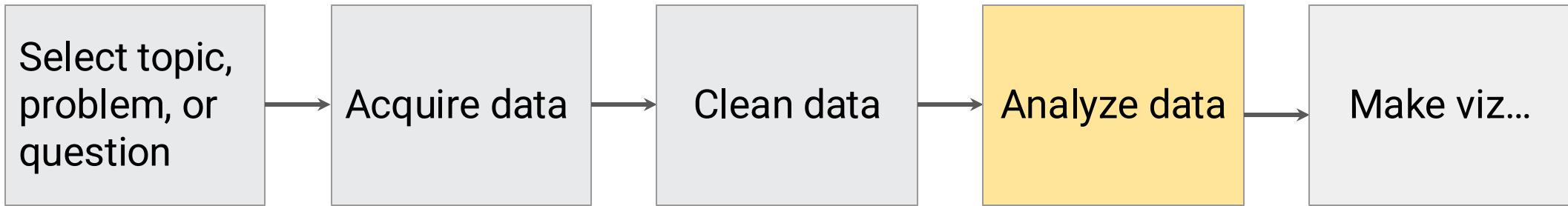
# Ethics: Data Viz Process



- **Problems related to data aggregation:**

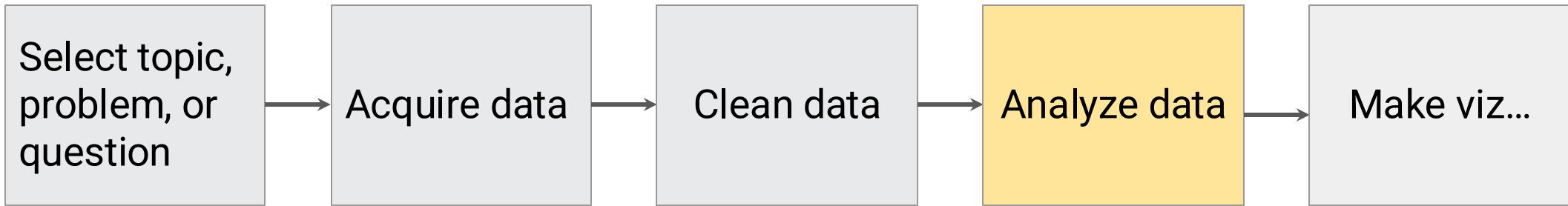
- Survey respondents from underrepresented minorities are often grouped together to avoid identification, but this decision also leads to the erasure of these groups.
- Students with low GPAs are grouped together in a category (“<3.5 GPA”) even though over 85% of students in this category have a GPA under 2.5

# Ethics: Data Viz Process



- Are you applying appropriate analysis methods / statistical tests to the data you have?
- Are you fully exploring data interactions and alternative explanations?
- Do you know enough about the data to interpret the results?
- Are you explicitly declaring all of the limitations of this analysis?
- Have you given appropriate credit to anyone who helped with data analysis?

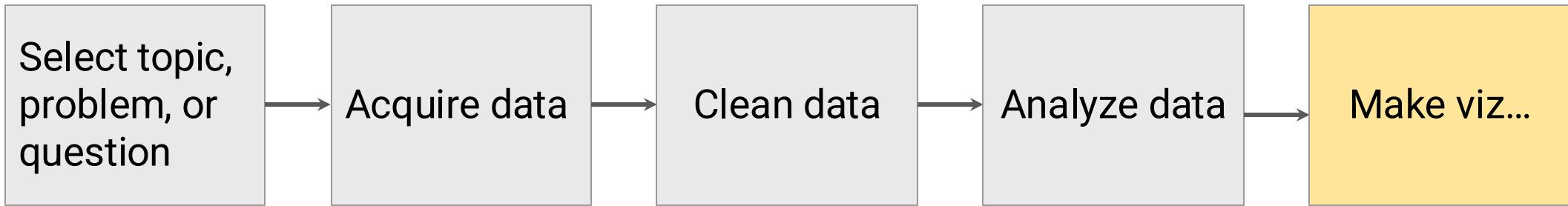
# Ethics: Data Viz Process



- **Bias introduced by assumptions:**

- Misunderstanding a hashtag that is meant to indicate sarcasm and, thus, coding a series of tweets as having positive sentiment instead of negative
- Ignoring gender because it wasn't expected to be an issue, only later to realize that gender has an effect on your question-of-interest
- Seeing an anomalous 50% drop in counts as a data collection problem, when really there was a global pandemic that year

# Ethics: Data Viz Process



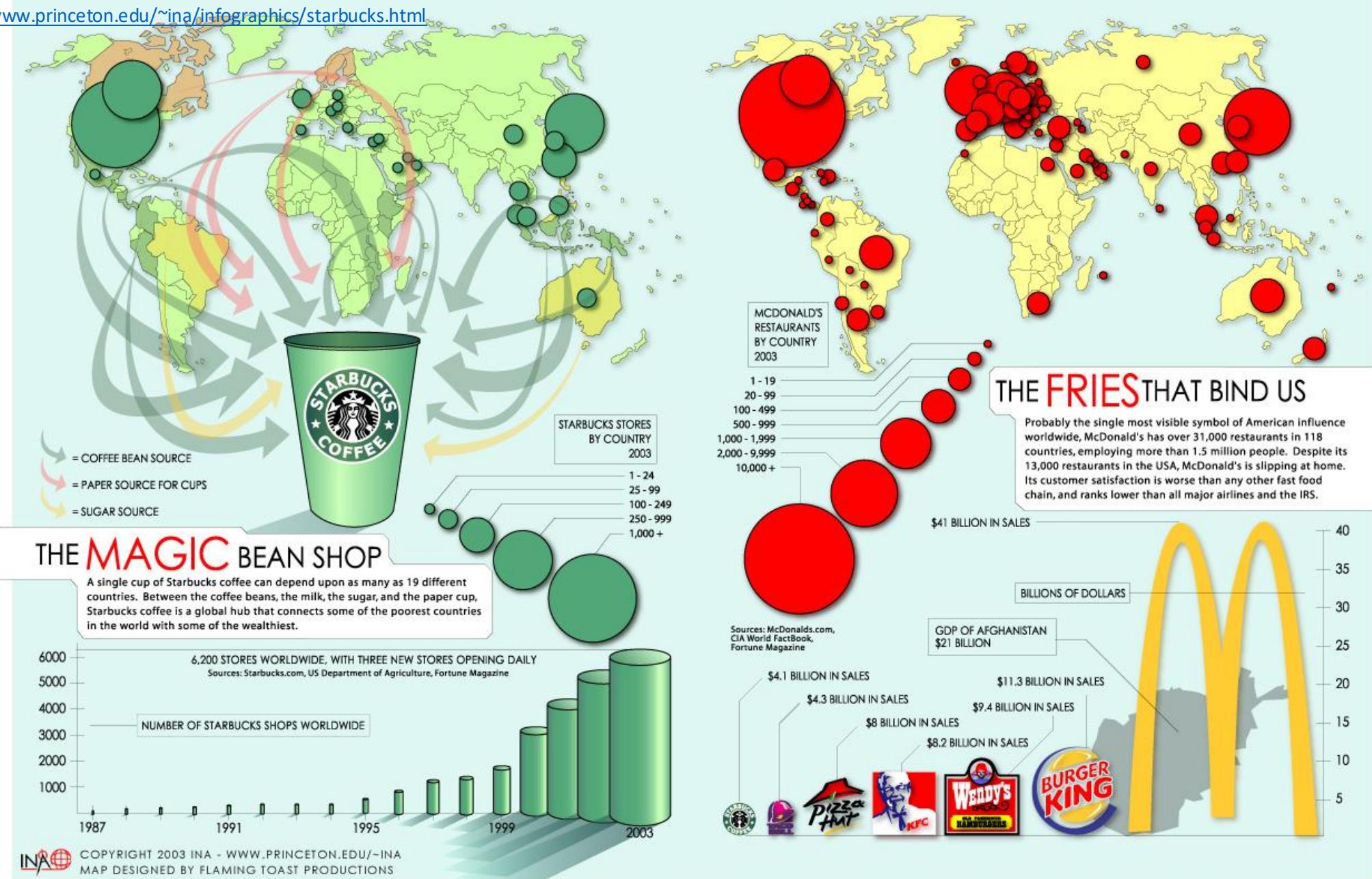
- Ask these questions of collaborators or yourself for each step of the process
- Show the process
- Give credit
- Fill in personal knowledge gaps by consulting literature, subject experts, and, if human-subject data is involved, work put out by members of the community being visualized

# Ethics: Know your audience

- Both intended and unintended audiences
- Conscious of historical context and what may come after ([datapractices.org](http://datapractices.org))
  - People can feel recognized, misrecognized, and unrecognized in data visualization (Naerland, 2020)
- Work towards increasing benefit and preventing harm ([datapractices.org](http://datapractices.org))
- What gaps might our audience have and what narrative needs to be added? (Cogley, 2018a)
- Does the visualization empower the audience? (D'Ignazio & Klein, 2016)

# The Fries that Bind Us and The Magic Bean Shop

- <https://www.princeton.edu/~ina/infographics/starbucks.html>



# Mapping Census Data: Poverty and Gender

## People with Income Below Poverty (Men)

Estimate; Income in the past 12 months below poverty level:

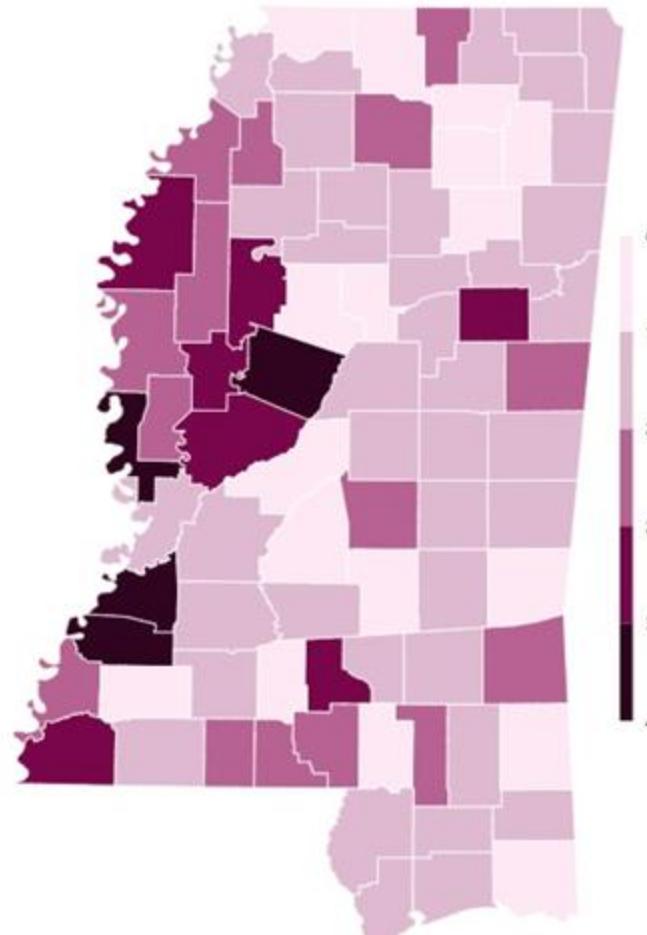
**Note:** Maps display percent of each demographic group below poverty by county. Poverty thresholds vary depending on the size of the family unit and the number of children under 18 years in the family. In 2018, the poverty threshold for an individual under 65 years is \$12,784. For a family of four (two parents, two children) the threshold is \$25,465. For each additional child, the threshold increases by between \$3,000 and \$5,000. The thresholds for each year since 1978 can be found here:

<https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>

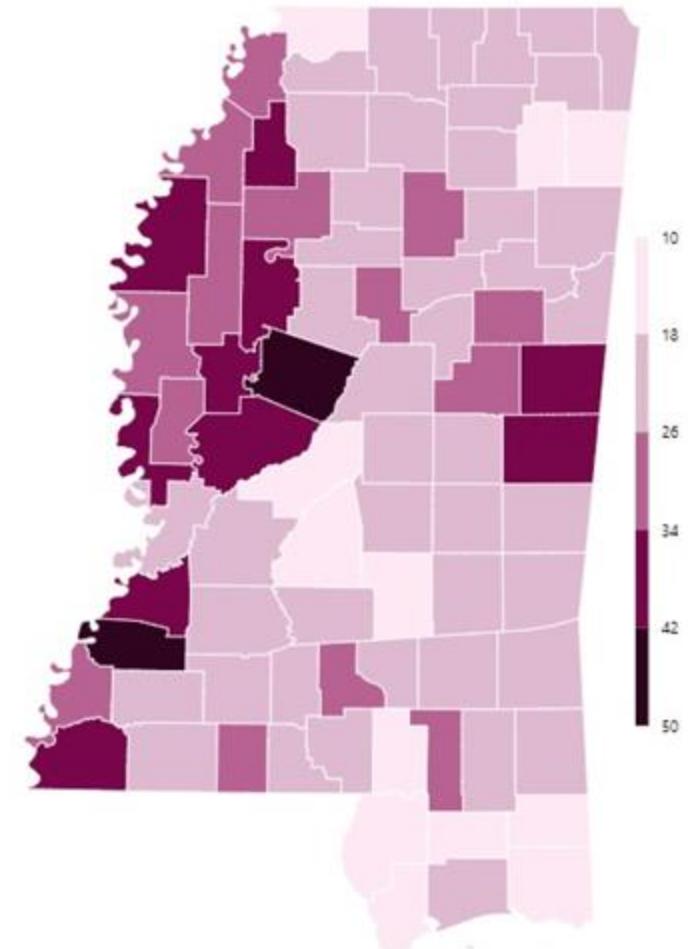
## People with Income Below Poverty (Women)

Estimate; Income in the past 12 months below poverty level:

Spatial Distribution

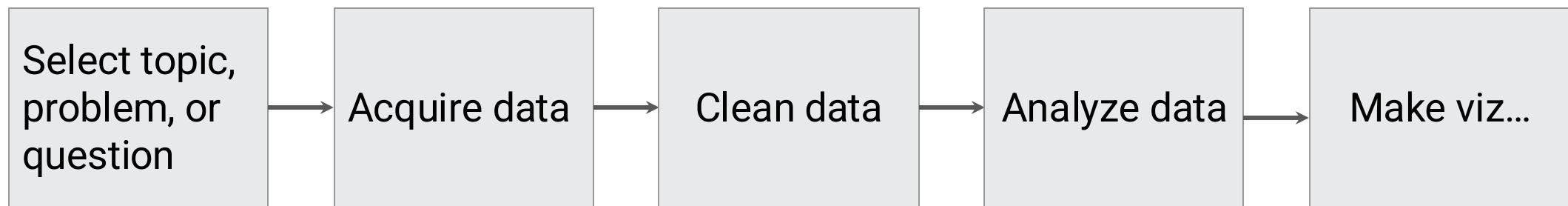


Spatial Distribution



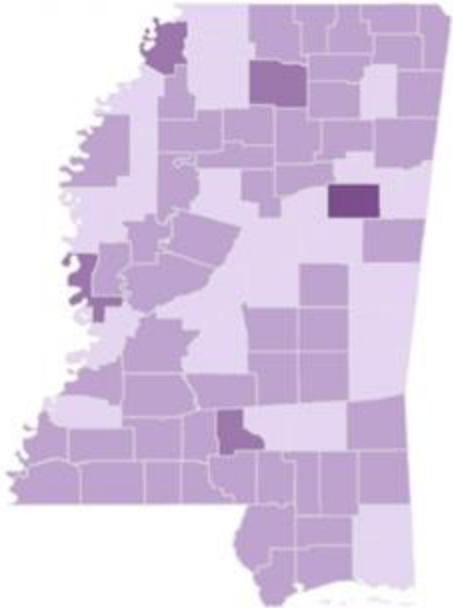
# Mapping Census Data: Poverty and Gender

- Challenging assumptions:
  - Did we know enough about the history of Mississippi to be able to interpret the relationship between poverty and gender?
  - Should the map creator(s) have investigated the intersection of race and gender as it relates to poverty in Mississippi?
  - Examining positionality: Who is benefitting from this work?
    - “Mapping inequality is not impactful in and of itself. Must continuously ask ourselves: are our practices creating the knowledge by which communities can build power?” (Koli 2019)

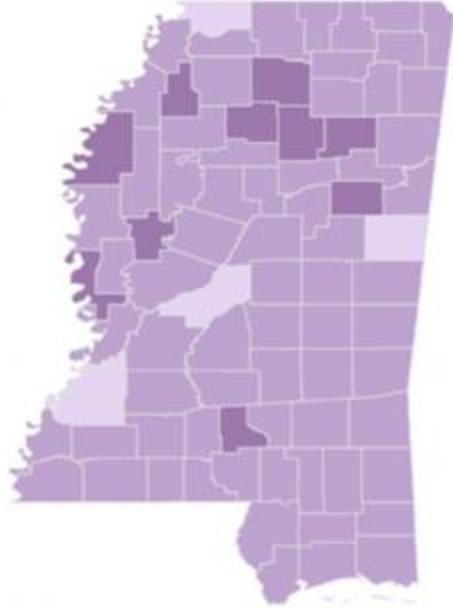


# Mapping Census Data: Poverty and Gender

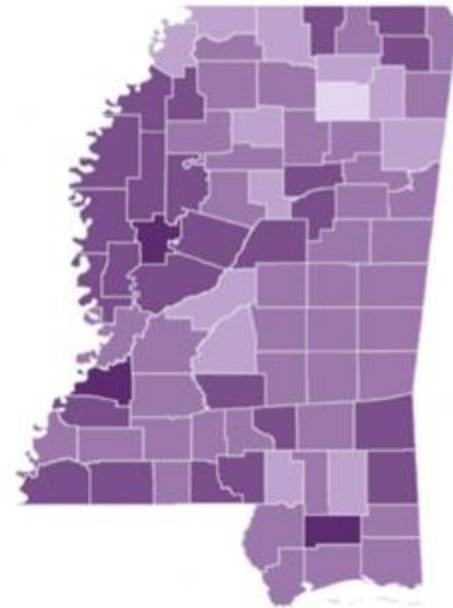
White Men



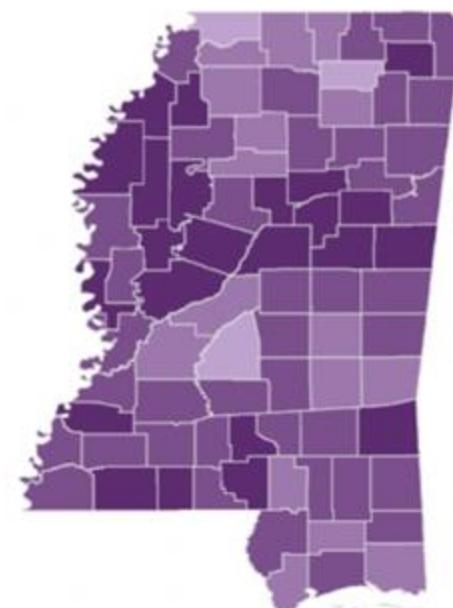
White Women



Black Men



Black Women

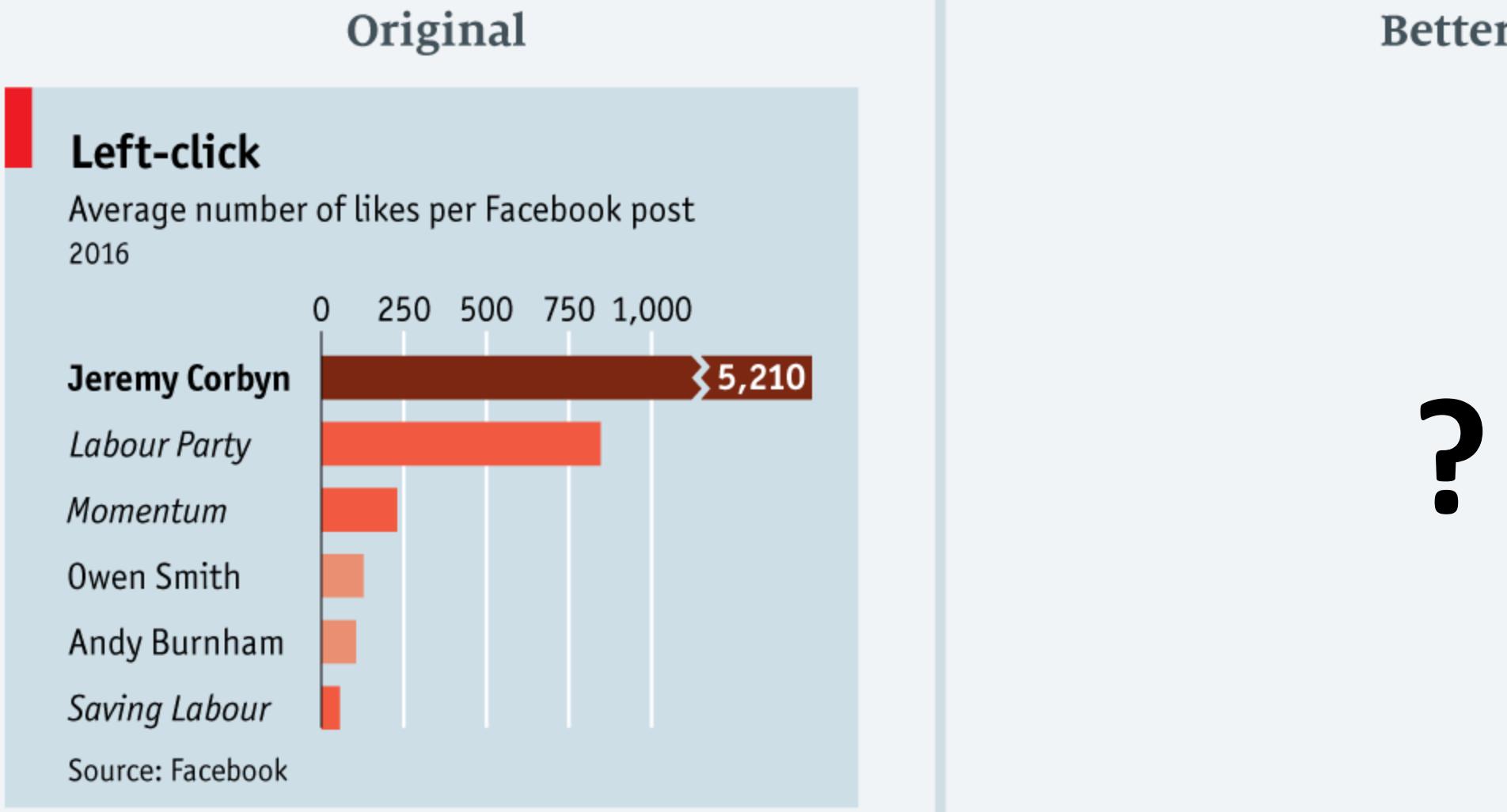


**Percent of Demographic Group Below Poverty**

0% - 10%
10.01% - 20%
20.01% - 30%
30.01% - 40%
40.01% - 51%

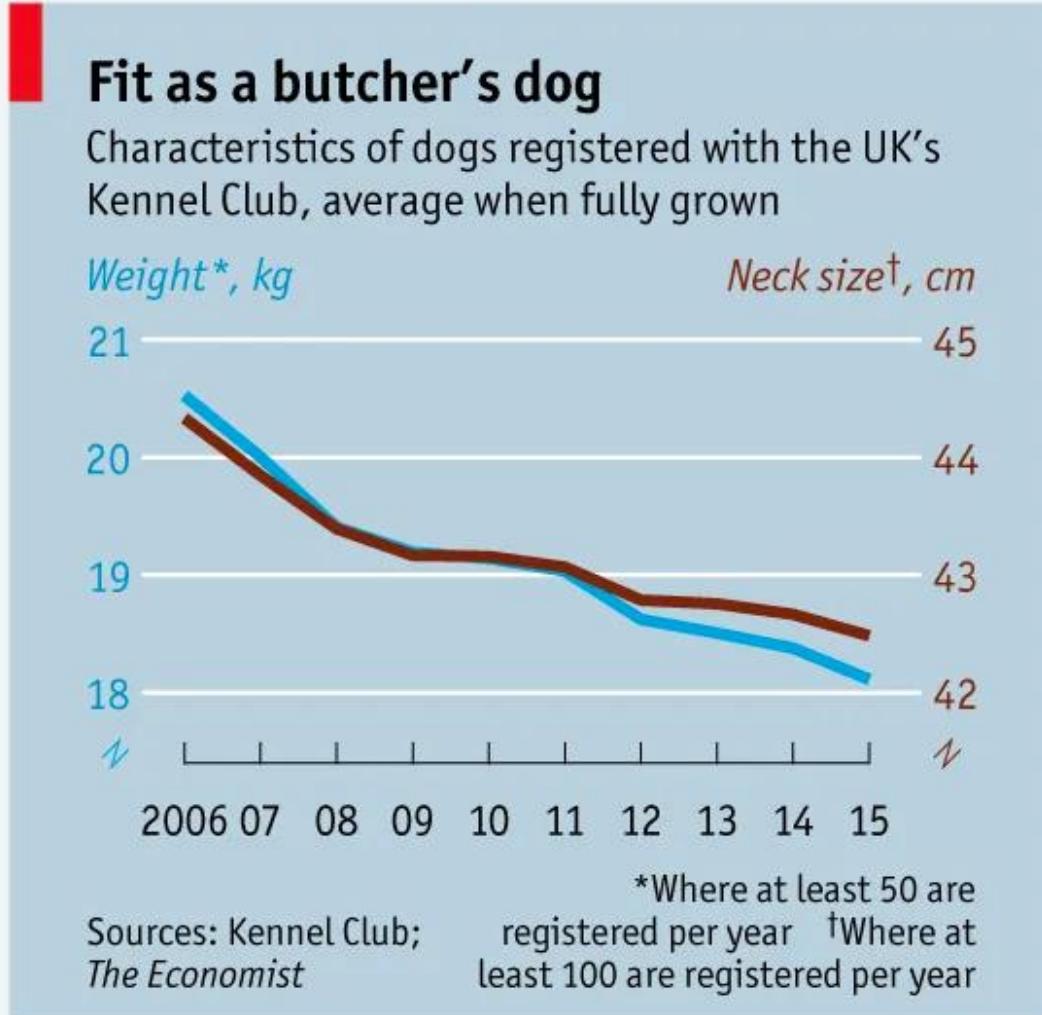
Source: 2012 ACS 5-year estimates. Koli, F. "(Un)Privileging the Map: A Community Collaboration in Understanding Economic Security." Paper presented at the **MAPPING (IN)JUSTICE SYMPOSIUM**: Digital Theory + Praxis For Critical Scholarship. Fordham University / November 7-9 2019

# The Economist visualizations



# The Economist visualizations

Original

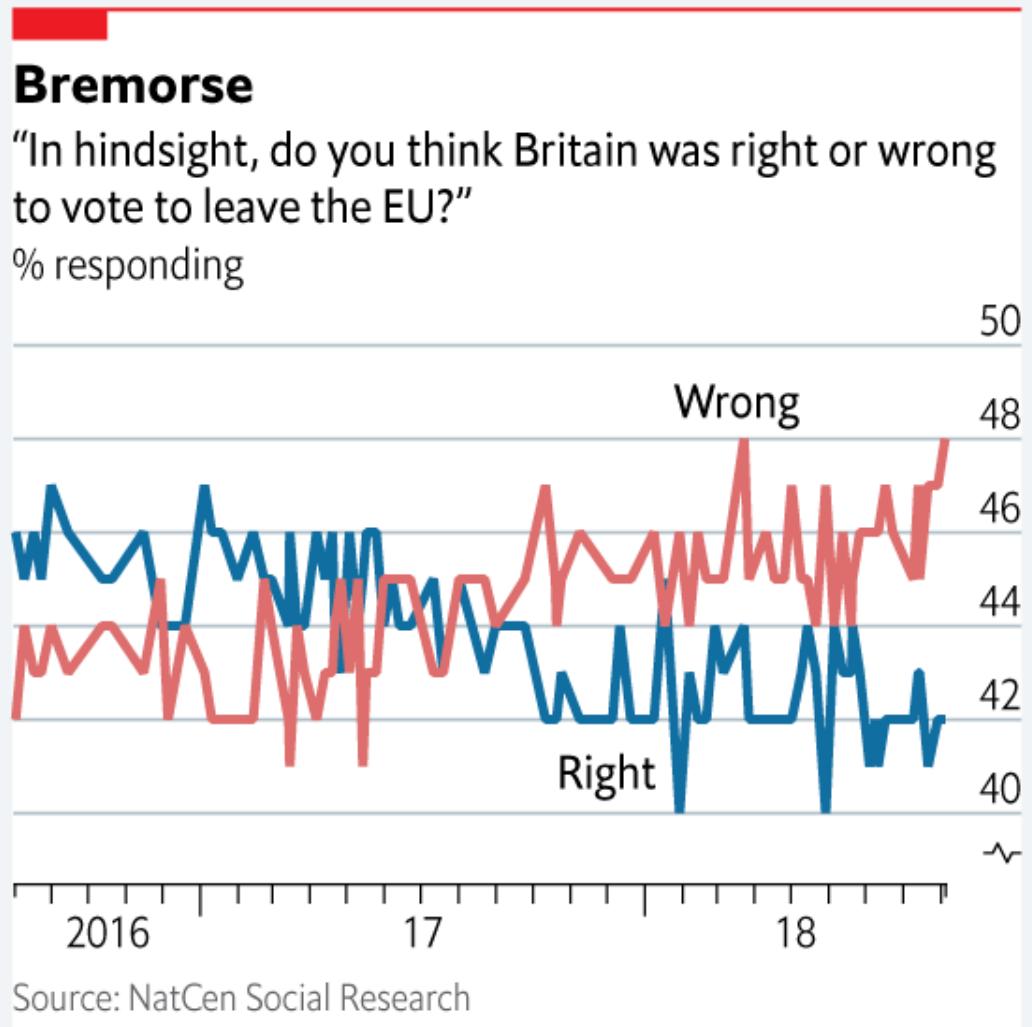


Better

?

# The Economist visualizations

Original



Better

?

# The Economist visualizations

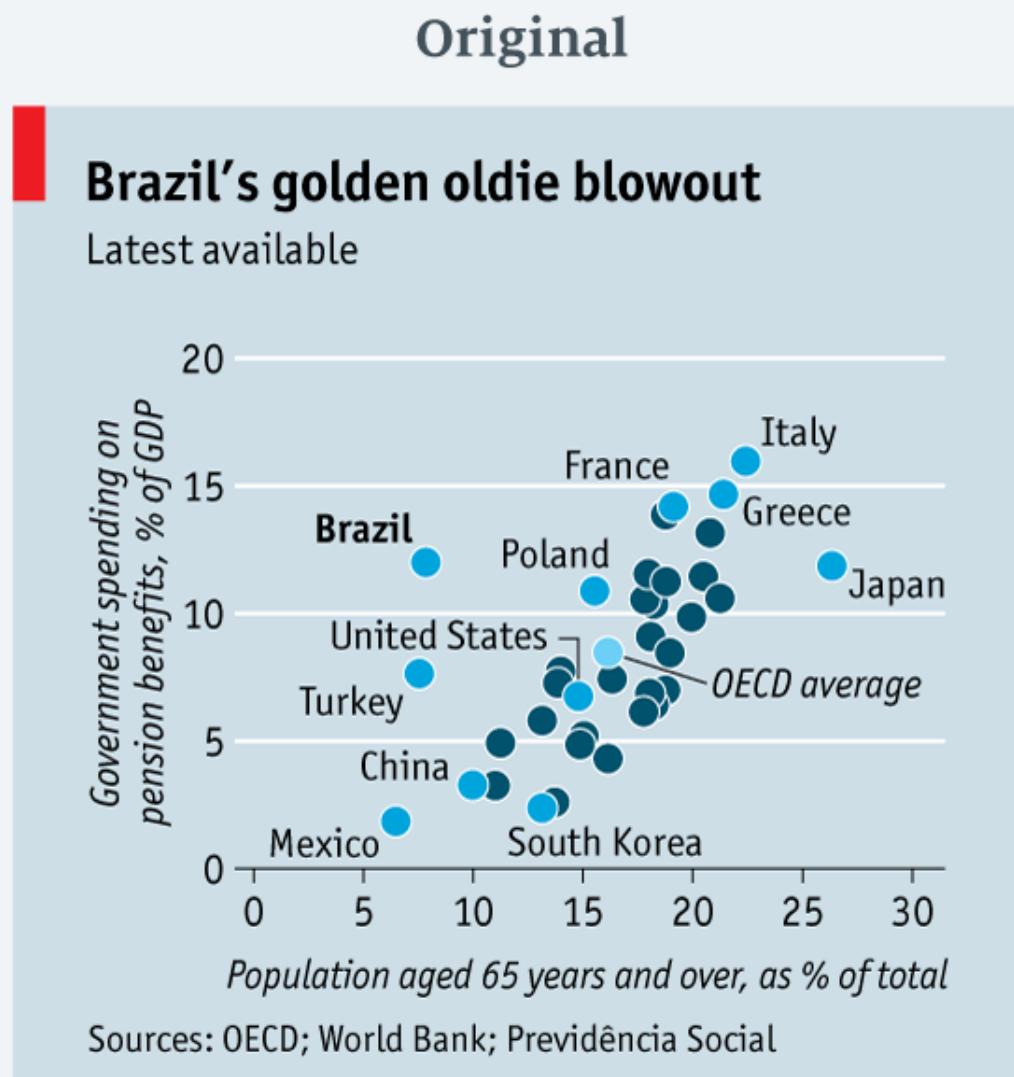
Original



Better

?

# The Economist visualizations



Better

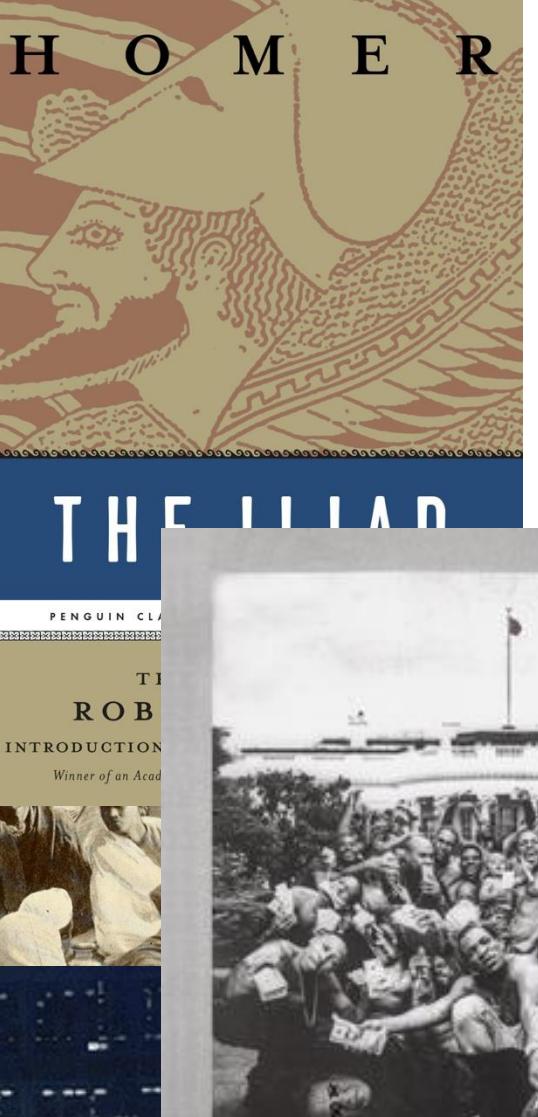
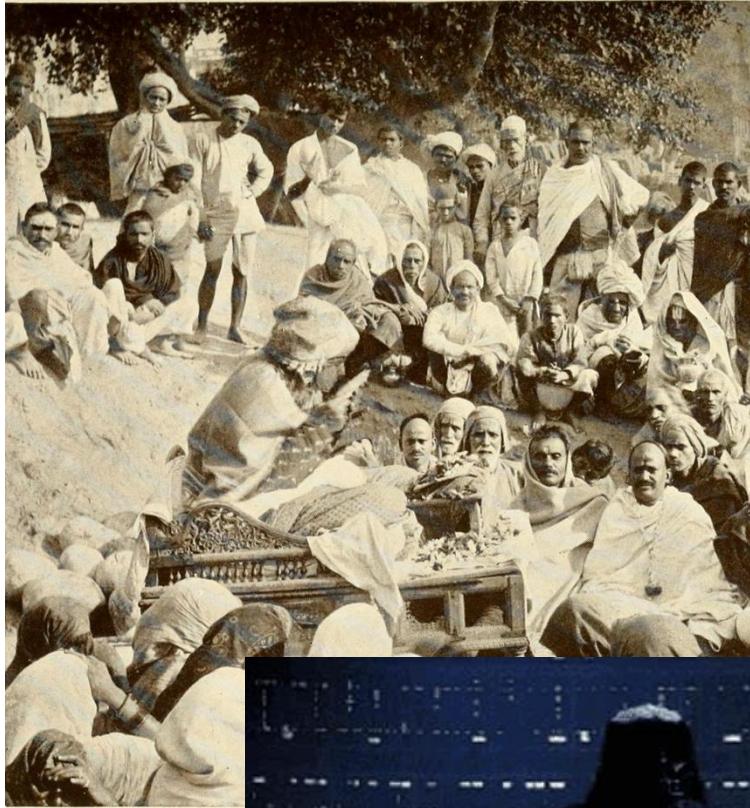
?



# Storytelling and narratives

# Storytelling

- As old as mankind...



# Storytelling

- People tell stories
- Words tell stories
- Images tell stories
- Comics tell stories
- Movies tell stories

# Narrative storytelling

- **narrative (n):** An account of a series of events, facts, etc., given in order and with the establishing of connections between them.
- Effective story-telling “require[s] skills like those familiar to movie directors, beyond a technical expert’s knowledge of computer engineering and science.” - Gershon & Page ‘01

## Generals Wary of Move to Cut Their Ranks

By GINGER THOMPSON and THOM SHANKER

WASHINGTON — Maj. Gen. Paul D. Eaton, a retired Army officer, is familiar with the perks and pitfalls of power, having commanded tens of thousands of troops at Fort Benning, Ga., managed budgets exceeding \$2 billion in Iraq, and overseen layers upon layers of staff members who helped manage both his professional duties and his personal life.

He has experienced the full range of lifestyles that come with military leadership, living at one point in an elegant antebellum mansion, and at another, with eight other officers crowded in a marble bathhouse behind one of Saddam Hussein's old palaces.

When he traveled, he was occasionally able to justify the use of military aircraft, but most times, he said, he flew coach. And today he lives on a pension worth 75 percent of his military salary, with health benefits that cover everything except dental and eye care for himself and his wife.

"We are well compensated, and we live very comfortable lives," General Eaton said, referring to the military's most senior leaders. "But when you look at all the things going on around a general, the nation is getting a very, very high return on its money."

Not everyone at the Pentagon agrees. Two weeks ago, Defense Secretary Robert M. Gates announced a sweeping effort to improve efficiency that, among other things, takes aim at the military's sacrosanct corps of generals and admirals, ordering his staff to cut at least 50 positions, and making clear that he would be happier if they cut more.

# Storytelling

August 26, 2010

## Generals Wary of Move to Cut Their Ranks

By GINGER THOMPSON and THOM SHANKER

WASHINGTON — Maj. Gen. Paul D. Eaton, a retired Army officer, is familiar with the perks and pitfalls of power, having commanded tens of thousands of troops at Fort Benning, Ga., managed budgets exceeding \$2 billion in Iraq, and overseen layers upon layers of staff members who helped manage both his professional duties and his personal life.

### Anecdotal lead

He has experienced the full range of lifestyles that come with military leadership, living at one point in an elegant antebellum mansion, and at another, with eight other officers crowded in a marble bathhouse behind one of Saddam Hussein's old palaces.

When he traveled, he was occasionally able to justify the use of military aircraft, but most times, he said, he flew coach. And today he lives on a pension worth 75 percent of his military salary, with health benefits that cover everything except dental and eye care for himself and his wife.

"We are well compensated, and we live very comfortable lives," General Eaton said, referring to the military's most senior leaders. "But when you look at all the things going on around a general, the nation is getting a very, very high return on its money."

Not everyone at the Pentagon agrees. Two weeks ago, Defense Secretary Robert M. Gates announced a sweeping effort to improve efficiency that, among other things, takes aim at the military's sacrosanct corps of generals and admirals, ordering his staff to cut at least 50 positions, and making clear that he would be happier if they cut more.

## Generals Wary of Move to Cut Their Ranks

By GINGER THOMPSON and THOM SHANKER

WASHINGTON — Maj. Gen. Paul D. Eaton, a retired Army officer, is familiar with the perks and pitfalls of power, having commanded tens of thousands of troops at Fort Benning, Ga., managed budgets exceeding \$2 billion in Iraq, and overseen layers upon layers of staff members who helped manage both his professional duties and his personal life.

He has experienced the full range of lifestyles that come with military leadership, living at one point in an elegant antebellum mansion, and at another, with eight other officers crowded in a marble bathhouse behind one of Saddam Hussein's old palaces.

When he traveled, he was occasionally able to justify the use of military aircraft, but most times, he said, he flew coach. And today he lives on a pension worth 75 percent of his military salary, with health benefits that cover everything except dental and eye care for himself and his wife.

"We are well compensated, and we live very comfortable lives," General Eaton said, referring to the military's most senior leaders. "But when you look at all the things going on around a general, the nation is getting a very, very high return on its money."

Not everyone at the Pentagon agrees. Two weeks ago, Defense Secretary Robert M. Gates announced a sweeping effort to improve efficiency that, among other things, takes aim at the military's sacrosanct corps of generals and admirals, ordering his staff to cut at least 50 positions, and making clear that he would be happier if they cut more.

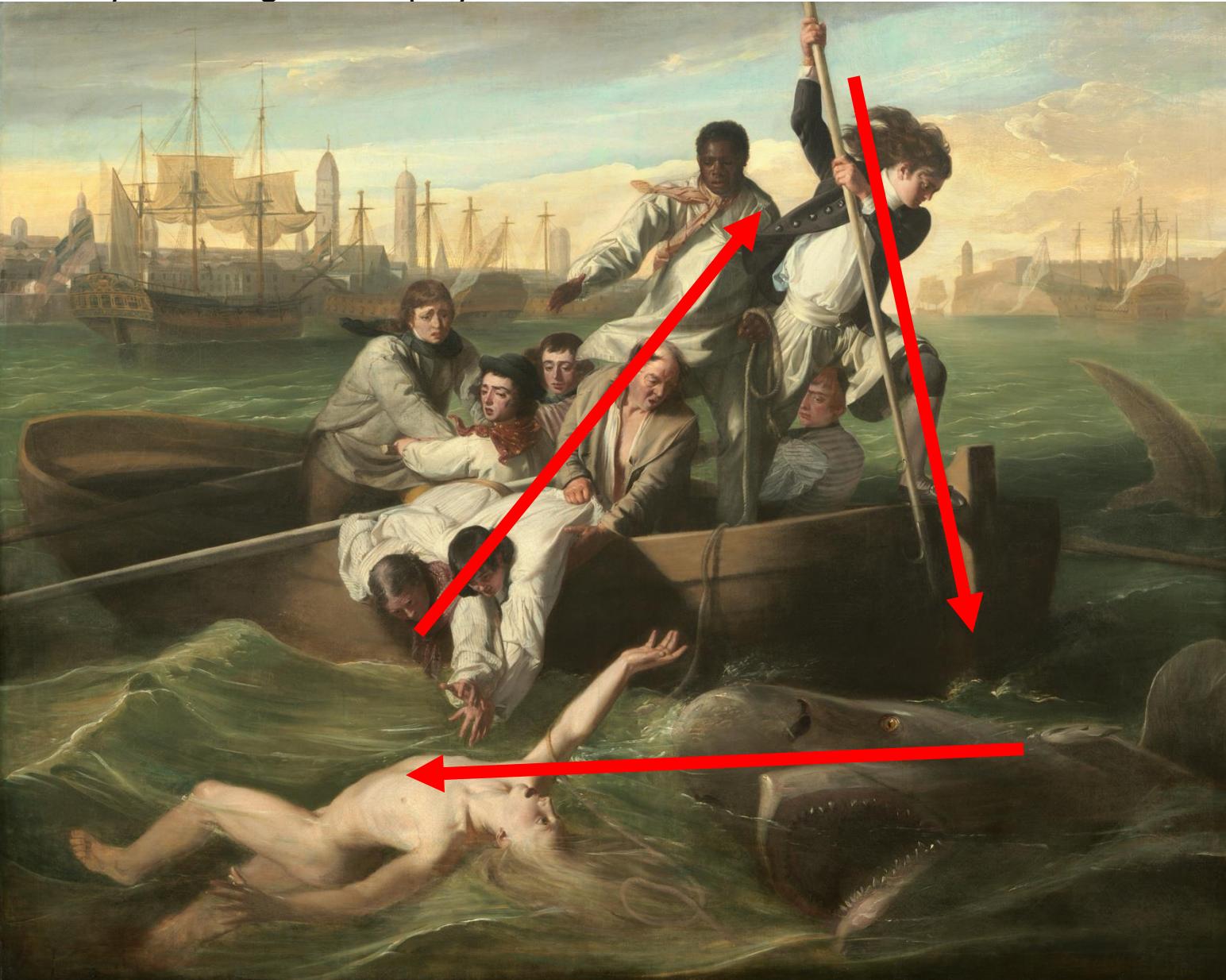
Nutshell paragraph

# Visual narration

- How do we use visuals to guide the viewer in a similar way as is done in text?

# Visual narration

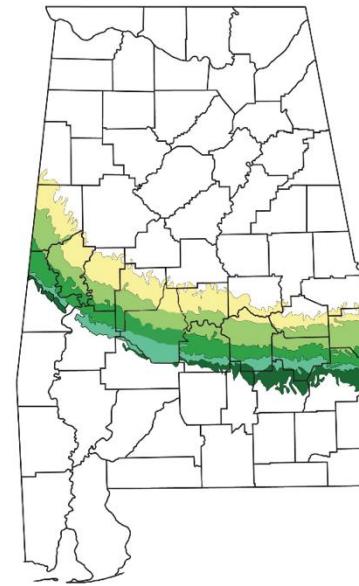
- Watson and the Shark by John Singleton Copley



# Storytelling with data

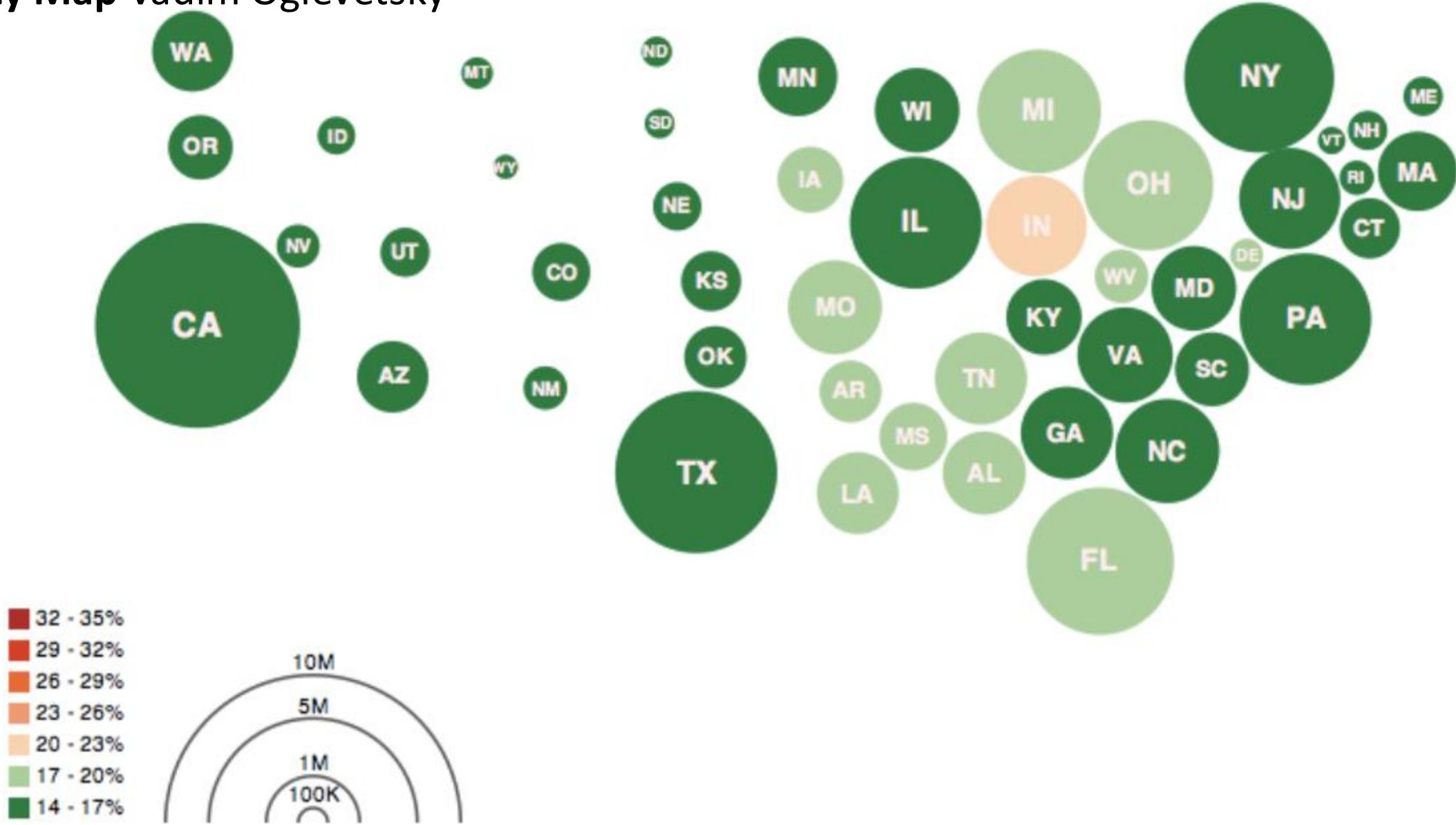
- <https://starkeycomics.com/2021/06/11/how-a-coastline-100-million-years-ago-influences-modern-election-results-in-alabama/>
- <https://encyclopediaofalabama.org/article/cretaceous-period-in-alabama/>
- [https://en.wikipedia.org/wiki/Black\\_Belt\\_\(region\\_of\\_Alabama\)](https://en.wikipedia.org/wiki/Black_Belt_(region_of_Alabama))

Cretaceous Sediments



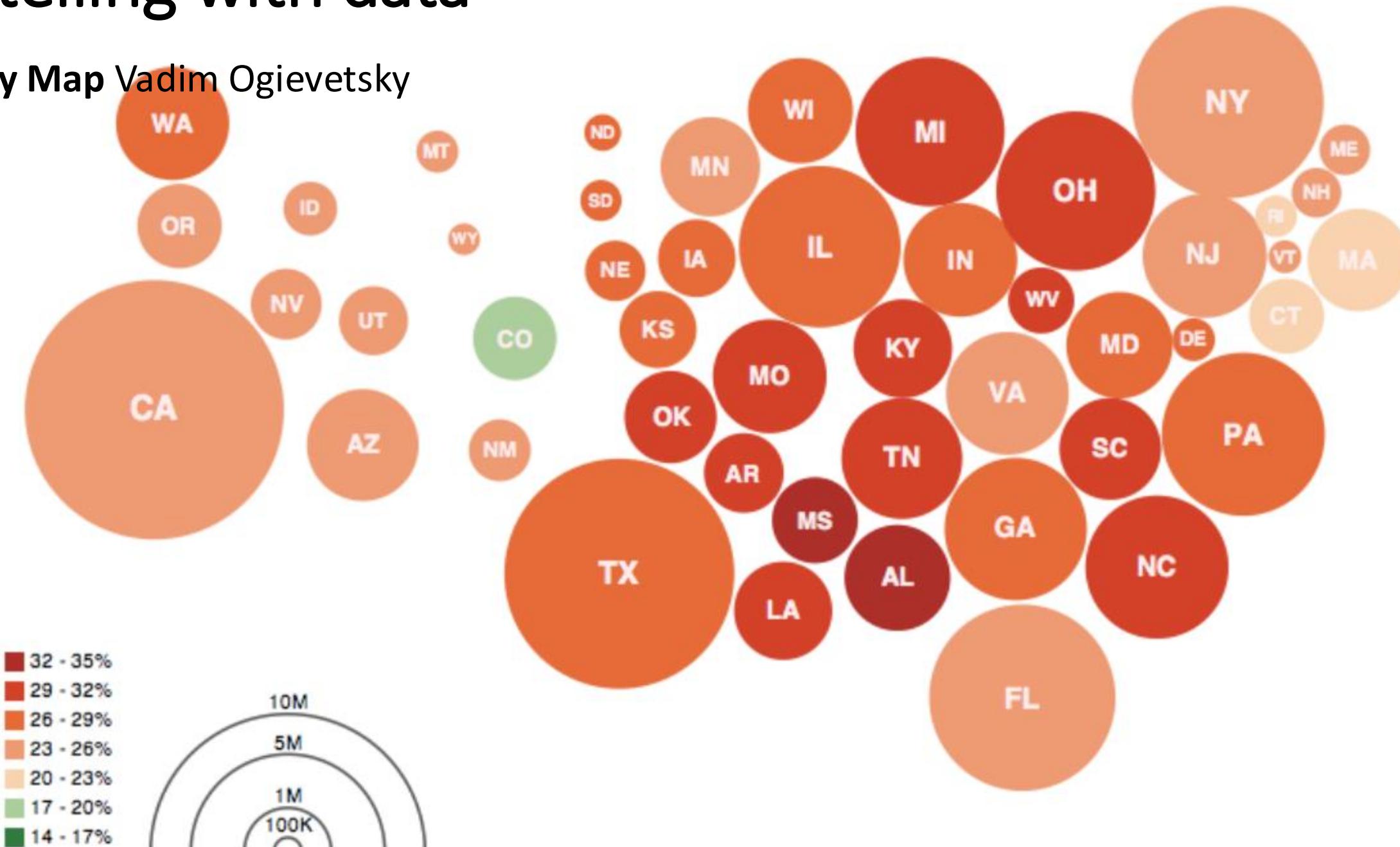
# Storytelling with data

- **Obesity Map** Vadim Ogrevetsky



# Storytelling with data

- Obesity Map Vadim Ogievetsky

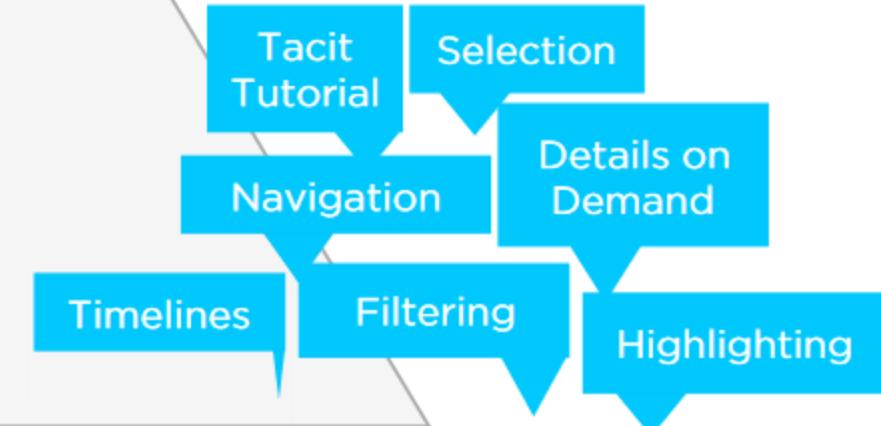
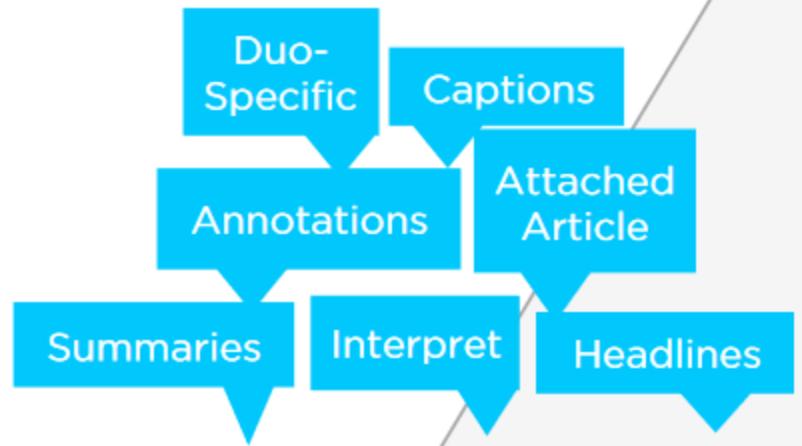
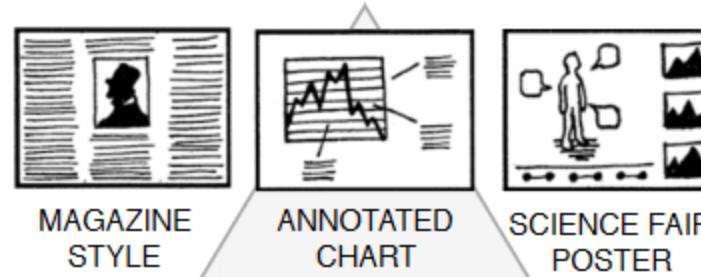


# Narrative Visualization Techniques

Segel, Edward, and Jeffrey Heer.

"Narrative visualization: Telling stories with data." *IEEE transactions on visualization and computer graphics* 16.6 (2010): 1139-1148.

## Visual Design



## Messaging

## Interactivity

# Magazine Style

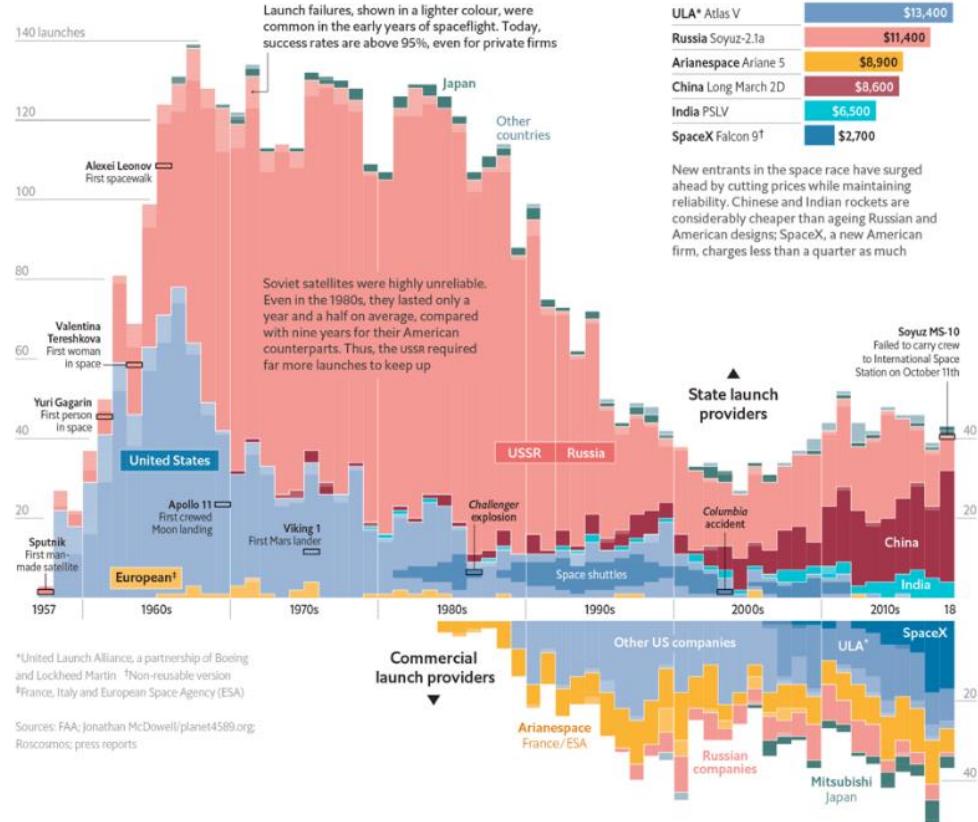
- Single frame
- Lots of text!
- Usually static



## Graphic detail

### Space launches

To Earth orbit or higher, at October 17th 2018



### A modern space race

## The next generation

Private businesses and rising powers are replacing the cold-war duopoly

OME 4,500 satellites circle Earth, providing communications services and navigational tools, monitoring weather, observing the universe, spying and doing more besides. Getting them there was once the business of the superpowers' armed forces and space agencies. Now it is mostly done by companies and the governments of developing countries.

During the early years of the space race

reaching orbit was hard. Between 1957 and 1962, 32% of American launches and 30% of Soviet ones failed. Accidents still happen: on October 11th a Russian rocket aborted its ascent shortly after launch (both crew members landed unharmed). Only states could assume such risks—and even if American firms had wanted to bear them, its government would not let them on national-security grounds. Companies eager to send objects into space, including telecoms firms, had to hitch a ride with NASA.

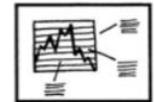
This changed when European countries started building launchers through a mostly state-owned company called Arianespace, which touted for custom among satellite-makers around the world. When the space shuttle *Challenger* exploded in 1986, NASA got out of the satellite-launching business. It and, later, the Pentagon be-

came new customers for private launch firms, alongside the satellite operators.

In the past decade the West's space-launch market has become more competitive thanks to an innovative new entrant, SpaceX. But state-run programmes still lead the way in emerging markets. In 2003 China became the third country to put a person into orbit; India plans to follow suit in 2022. Both sell launch services to private clients. China did legalise private space flight in 2014, but no companies based there have yet reached orbit on their own.

Like their cold-war predecessors, these Asian titans have strategic goals as well as a thirst for publicity. They need independent access to space for communication, intelligence and navigation. However commercialised space gets, the competition will never be solely economic. ■

# Annotated Chart



- Text overlayed on the viz
- Highlight key events and points of interest
  - Don't obscure other parts of the data!
- Brief, informative descriptions
- Annotations create a reading order
  - Depends on the viz type
- Don't include too many annotations



# Partitioned poster

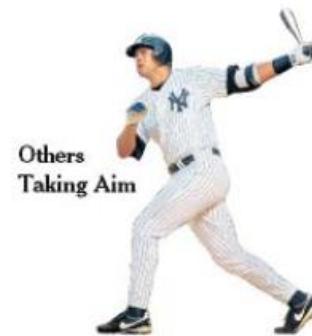
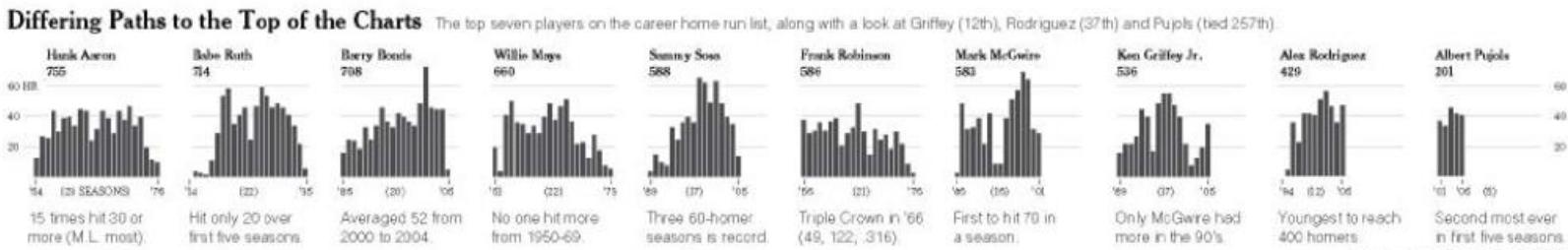


- Partitions define different aspects of the data
- Consistent ordering of elements in each partition
  - Less overhead when moving to a new partition
- Reading order defined by size/color/bolding/image content of elements
  - Smaller charts on bottom are less important
  - Arrows help to guide the reading order

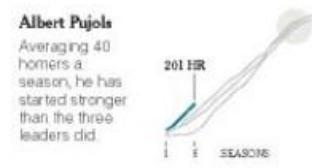
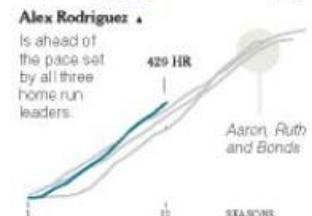
755



## Differing Paths to the Top of the Charts

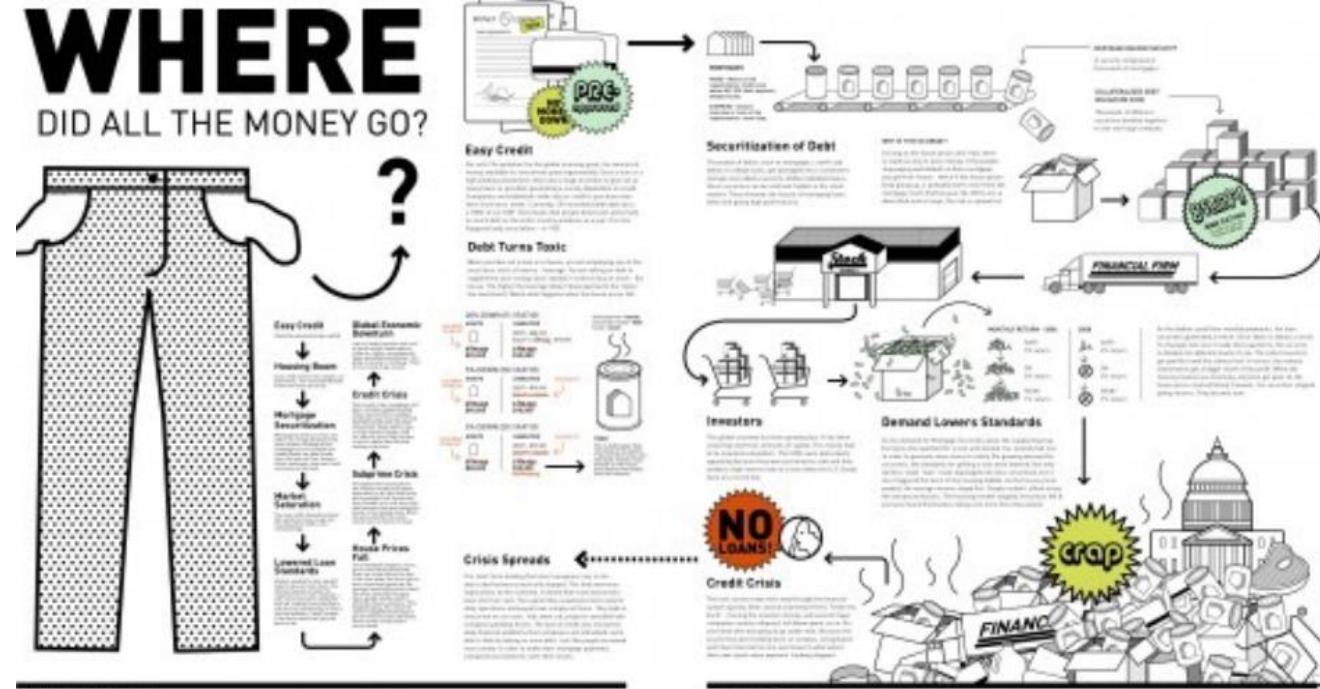


**Others Taking Aim**



# Flow chart

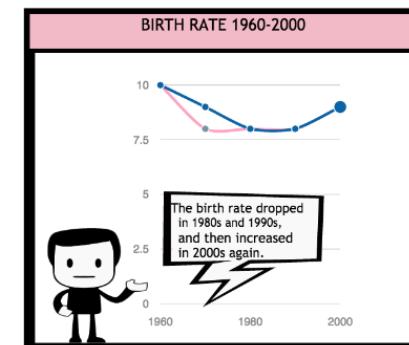
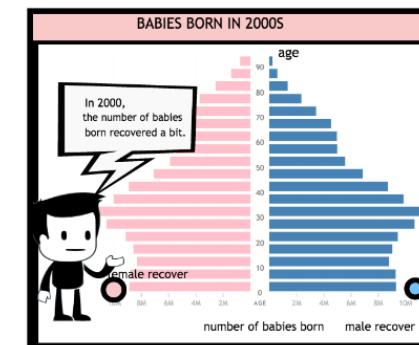
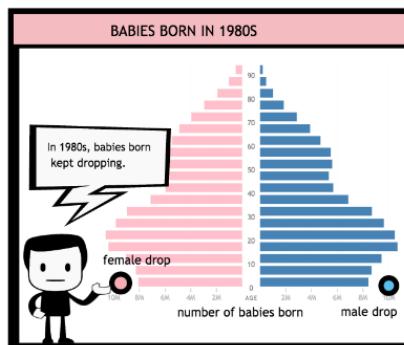
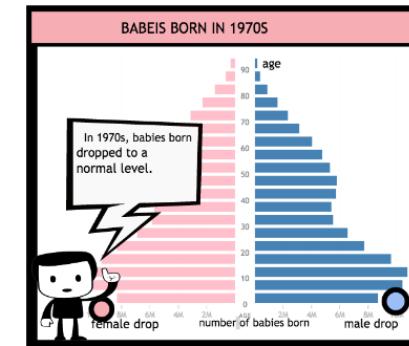
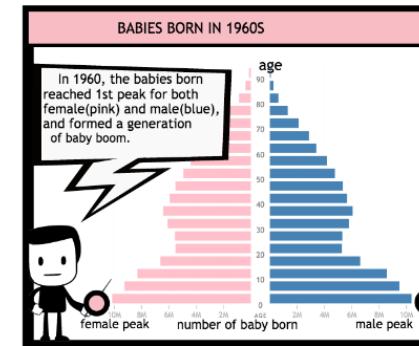
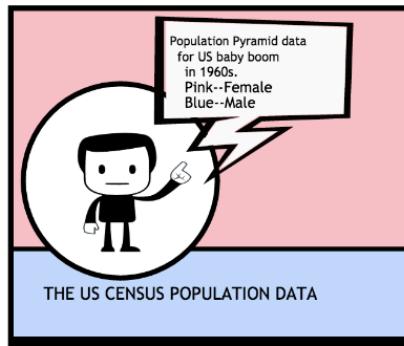
- Maps a complex process/system into digestible components
- Clear ordering
  - Obvious entry point (title, large icons)
  - Clear progression (arrows)
  - Top-bottom or left-right



# Comic Strip

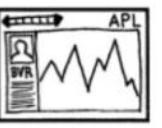


- Directly uses a comic strip format
- Very familiar, easy to understand, and memorable
  - Salient visual elements
  - Comics are usually simple
  - Direct narration on each panel (act like annotations)
  - More approachable for a wide audience
- Can be difficult to create
  - Requires some artistic skills
  - Not suitable for strict professional settings
  - Can be distracting
  - Can oversimplify data

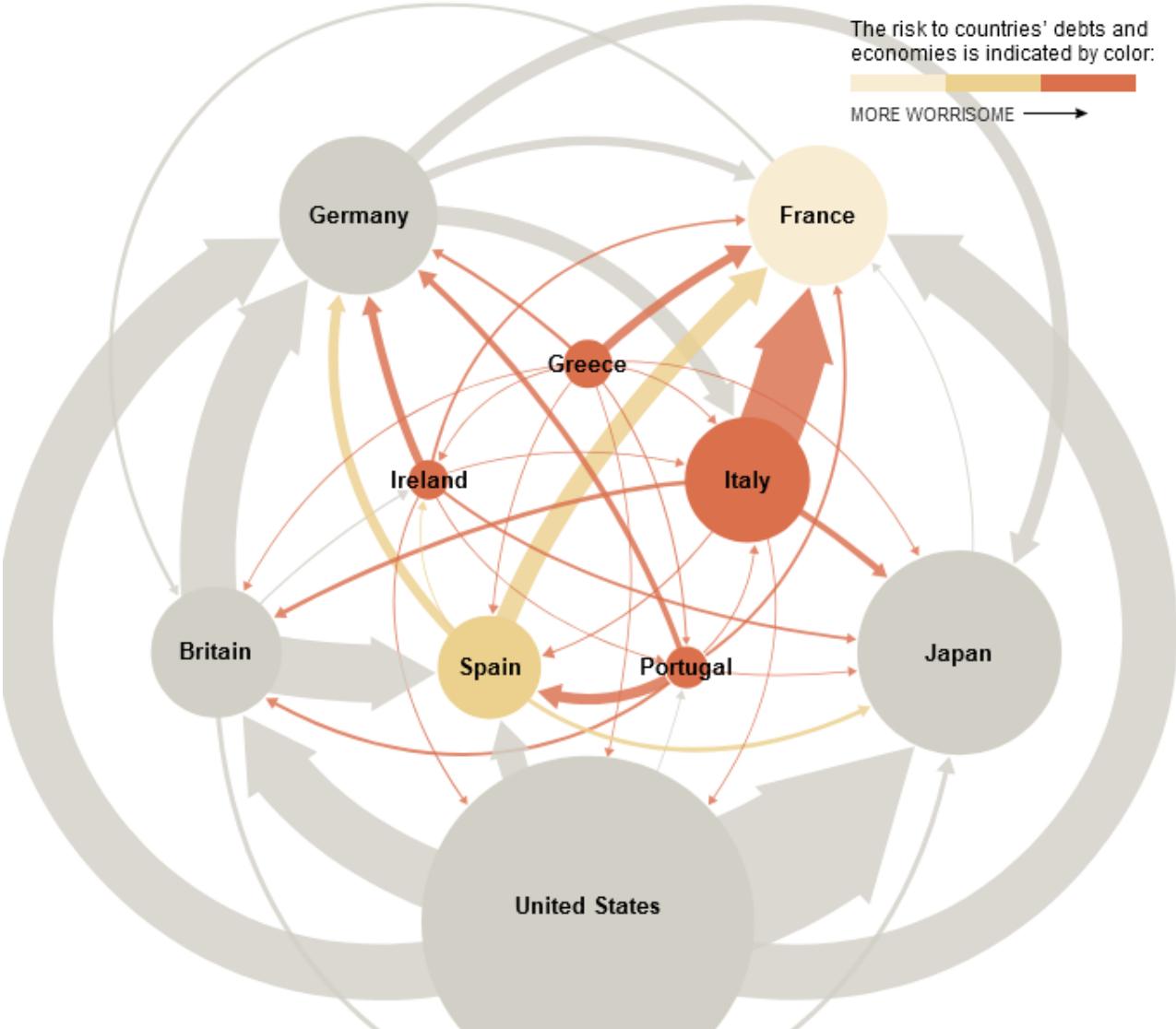


Zhao, Zhenpeng, Rachael Marr, and Niklas Elmquist. "Data comics: Sequential art for data-driven storytelling." *tech. report* (2015).

# Slide show



- Utilizes discrete stages of viz
  - Move through the stages one by one
- Very clear sequence (forced, usually)
- More appropriate for professional settings
- Each slide (view) can implement other storytelling techniques (e.g. partitioned poster)
- Can enable user-driven exploration (user decides which slides to look at for how long)
- Especially good for cause-and-effect relationships



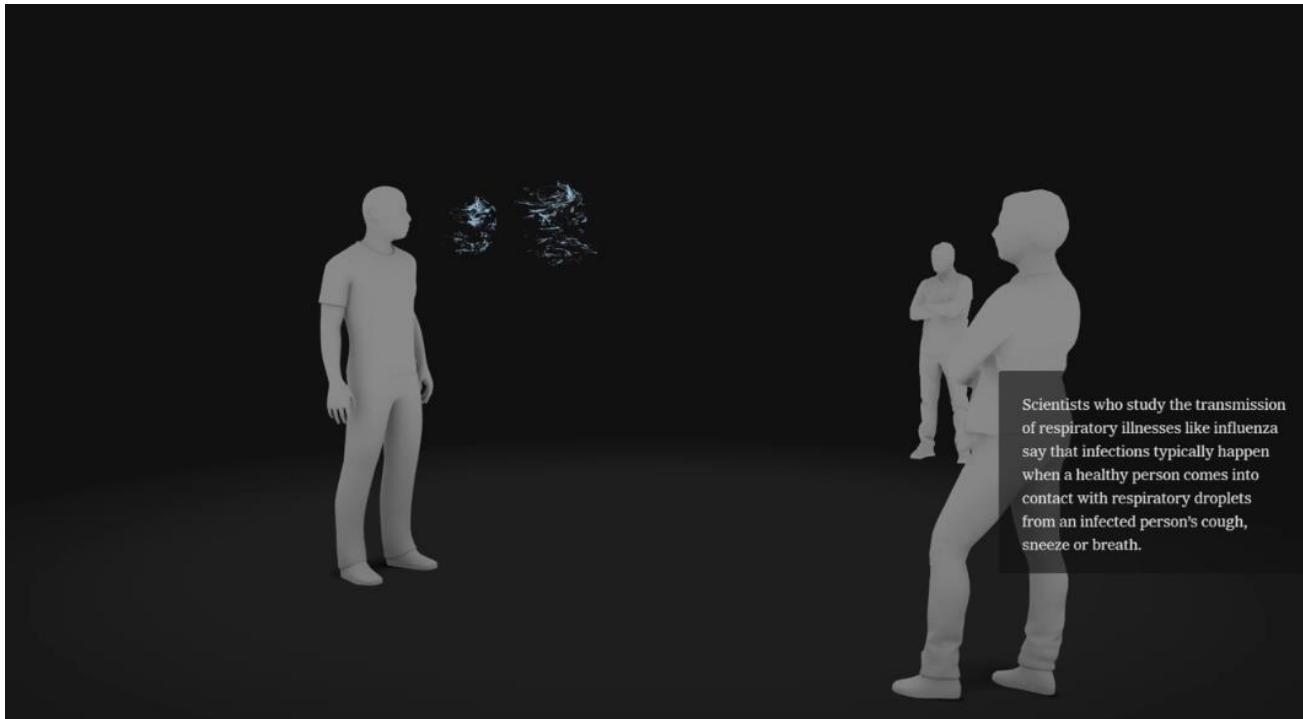
Arrows show imbalances of debt exposure between borrowers in one country and banks in another; arrows point from debtors to their bank creditors. Arrow widths are proportional to the balance of money owed. For example, French borrowers owe Italian banks \$50.6 billion; Italian borrowers owe French banks \$416.4 billion. The difference — their imbalance — shows France's banking system more exposed to Italian debtors by about \$365.8 billion.

<https://archive.nytimes.com/www.nytimes.com/interactive/2011/10/23/sunday-review/an-overview-of-the-euro-crisis.html>

# Film/video/animation

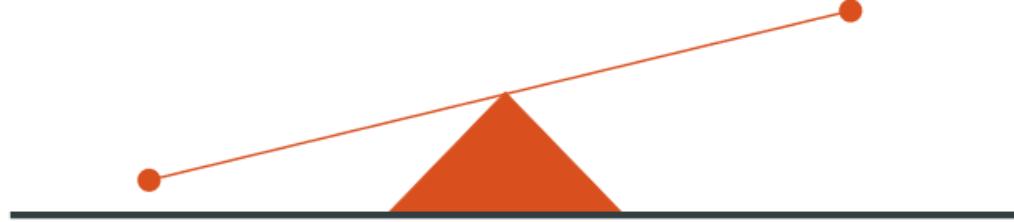


- Uses motion!
- Especially good for temporal data
- Can be very engaging/immersive for viewers
- Also allows for audio!
- Electronic only...



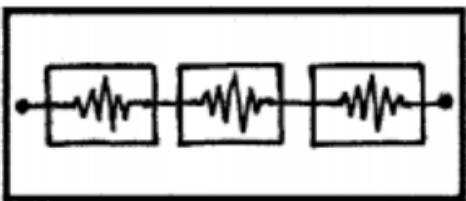
<https://www.nytimes.com/interactive/2020/04/14/science/coronavirus-transmission-cough-6-feet-ar-ul.html>

# Level of Guidance Control



## **Author driven**

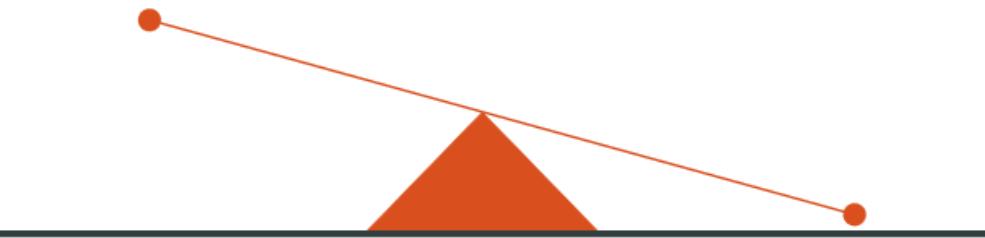
Prescribed ordering  
Stronger messaging  
Limited interactivity



Interactive Slideshow

<https://archive.nytimes.com/www.nytimes.com/interactive/2013/10/09/us/yellen-fed-chart.html>

# Level of Guidance Control

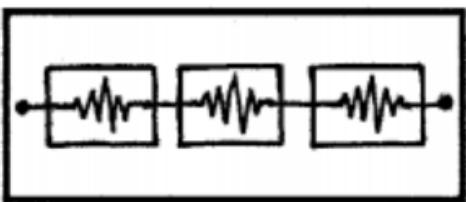


## Author driven

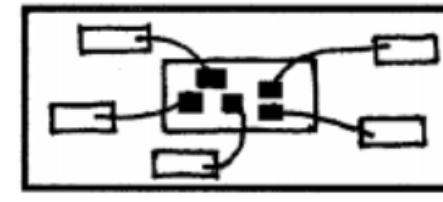
Prescribed ordering  
Stronger messaging  
Limited interactivity

## Reader driven

Multiple orderings  
Less messaging  
More open interactivity



Interactive Slideshow



Drill down

<https://www.washingtonpost.com/wp-srv/special/world/north-korean-prison-camps-2009/>  
(need Flash)

# Level of Guidance Control

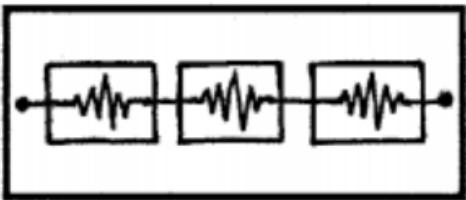


## Author driven

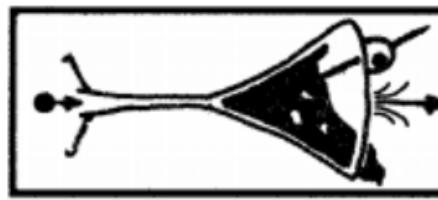
Prescribed ordering  
Stronger messaging  
Limited interactivity

## Reader driven

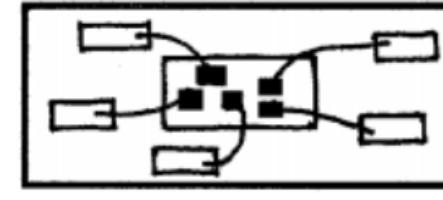
Multiple orderings  
Less messaging  
More open interactivity



Interactive Slideshow



Martini Glass



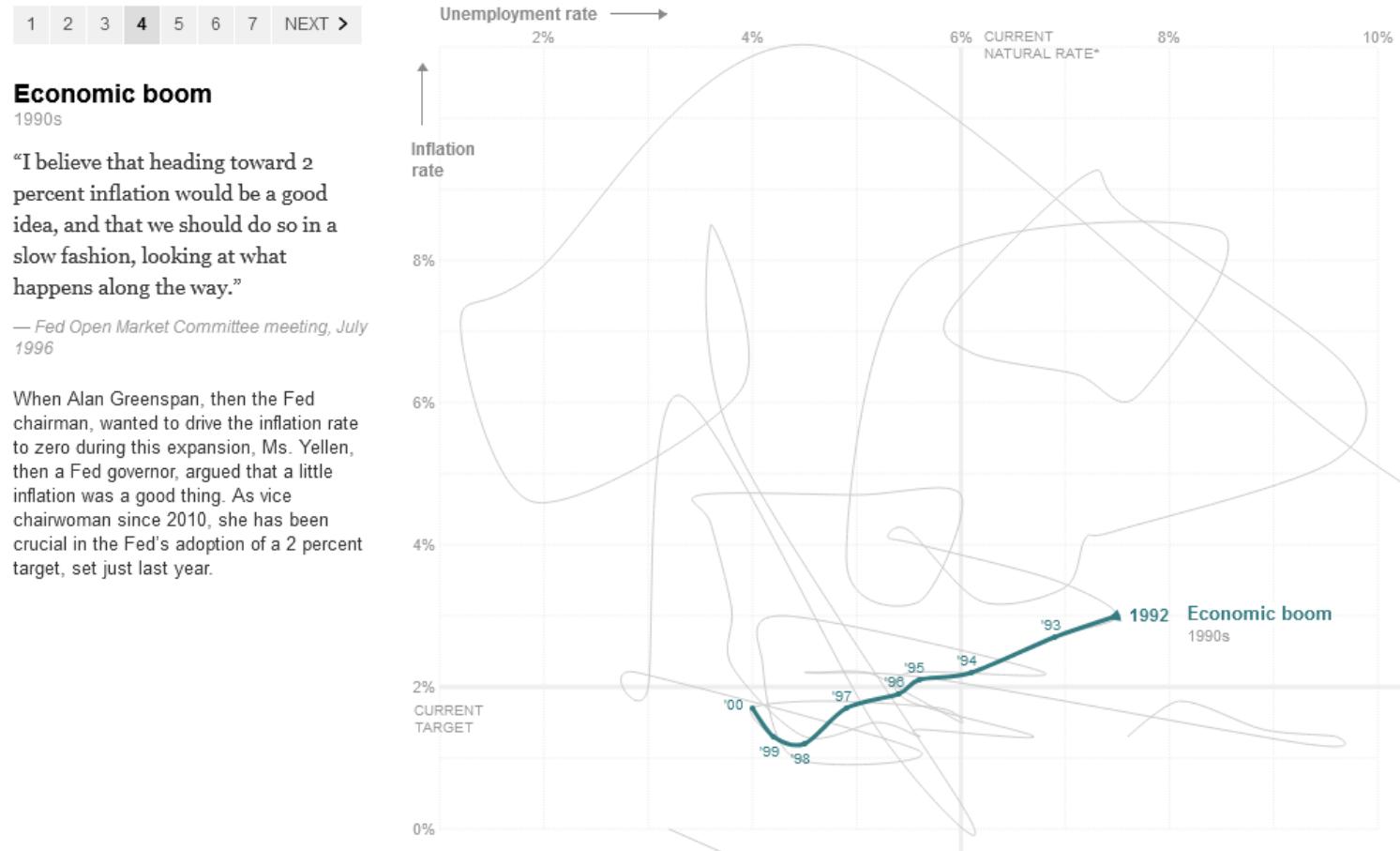
Drill down

# Narrative Visualization Techniques: Stepper

- **Steppers**

- Controlled trajectory through the data
- Carefully controlled views of the data

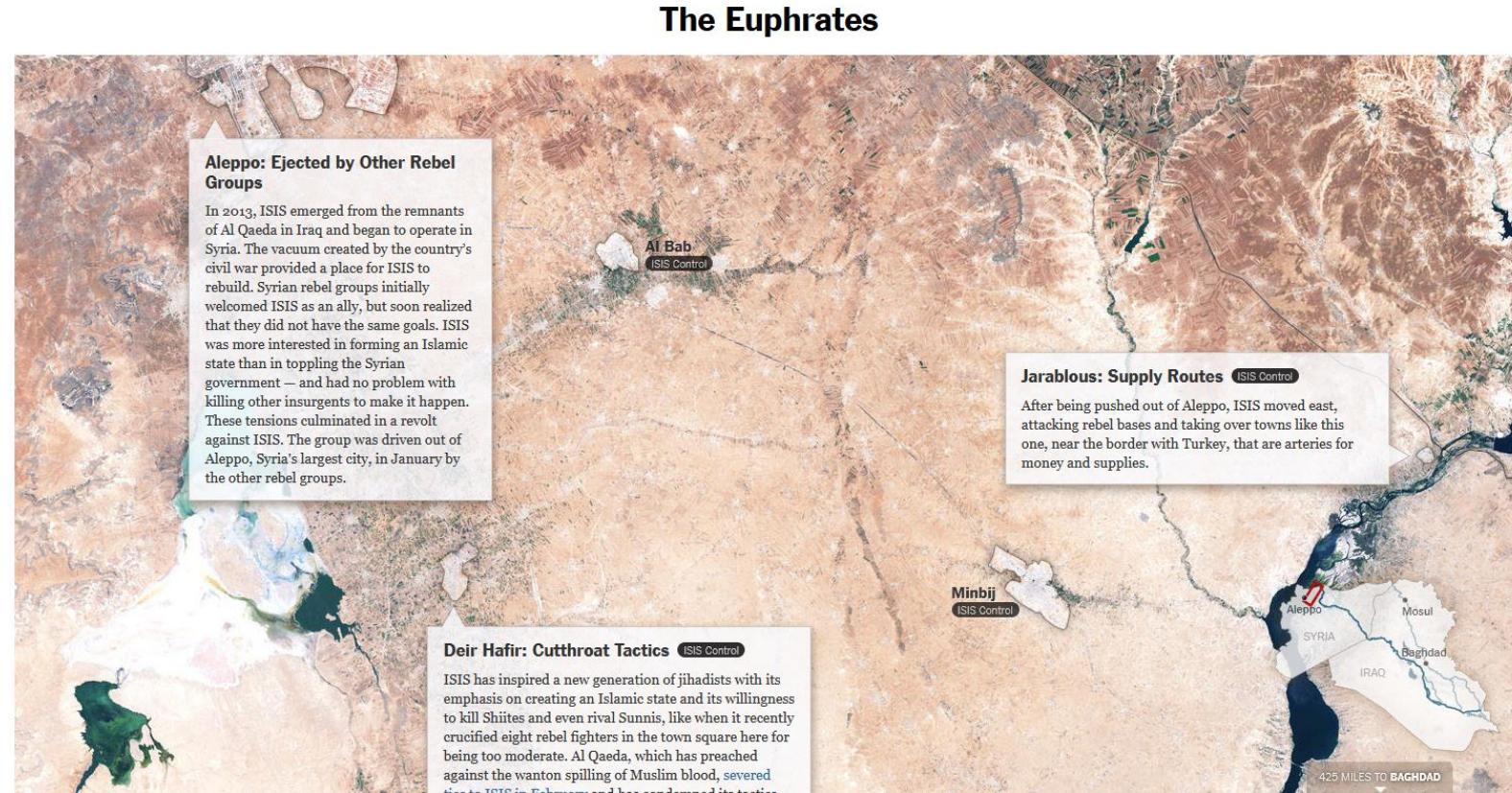
Janet L. Yellen, on the Economy's Twists and Turns



<https://archive.nytimes.com/www.nytimes.com/interactive/2013/10/09/us/yellen-fed-chart.html>

# Narrative Visualization Techniques: Scrollytelling

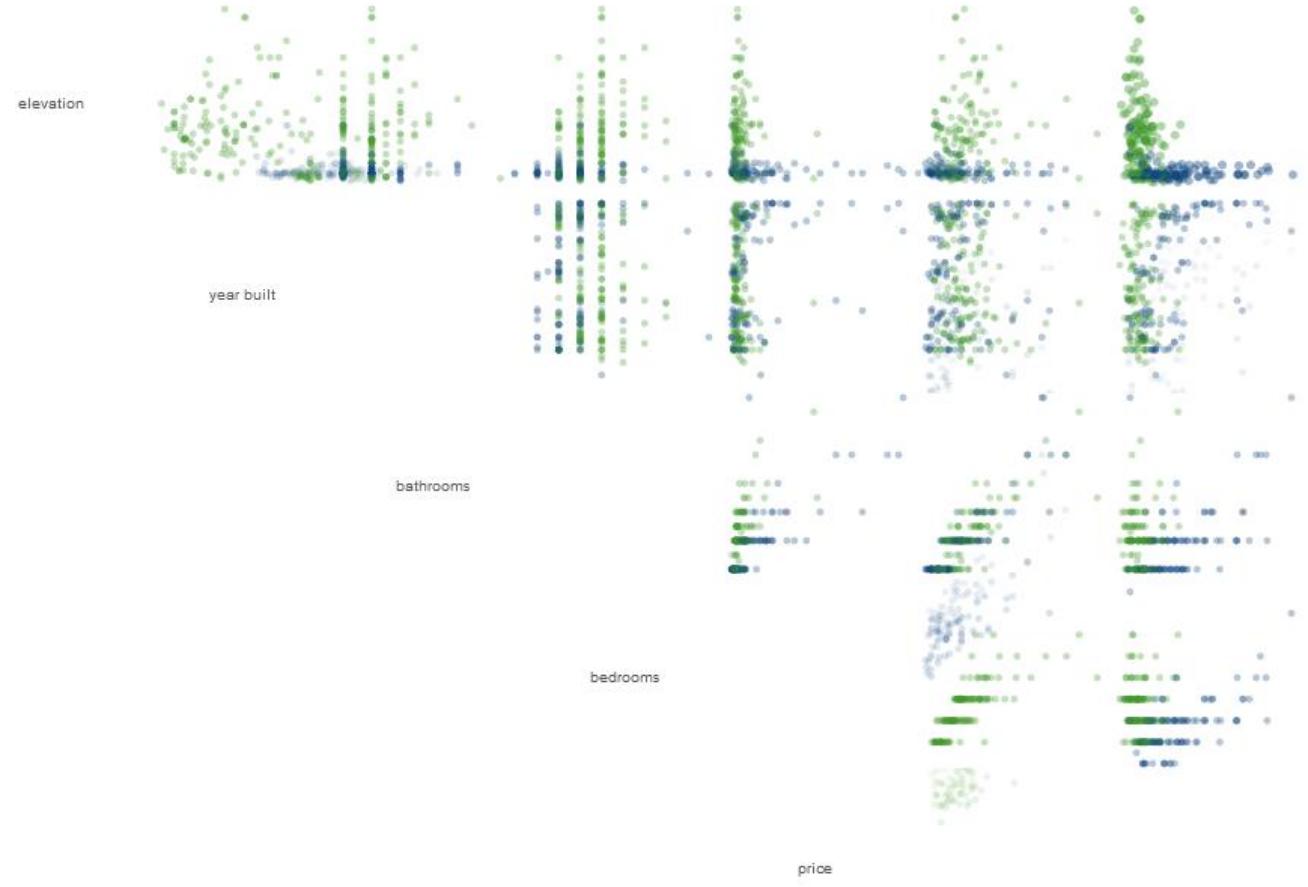
- Scrollytelling
  - Similar to steppers
  - use scroll wheel instead



<https://www.nytimes.com/interactive/2014/07/03/world/middleeast/syria-iraq-isis-rogue-state-along-two-rivers.html>

# Narrative Visualization Techniques: Scrollytelling

- Scrollytelling
  - Similar to steppers
  - use scroll wheel instead



<http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

# Narrative Visualization Techniques

- Discrete vs continuous
- Stepper pros:
  - Very clear
  - Very precise control
- Scrollytelling pros:
  - Lightweight
  - Easy to keep progressing
  - More engaging (?)
- Implementation:
  - Do it yourself: monitor scroll position → element positions
    - <https://bostocks.org/mike/scroll/>
    - [https://vallandingham.me/think\\_you\\_can\\_scroll.html](https://vallandingham.me/think_you_can_scroll.html)
  - Libraries:
    - <https://idyll-lang.org/gallery>

# Narrative Visualization Examples

- <https://mathisonian.github.io/trig/etymology/>

## The Etymology of Trig Functions

By: Matthew Conlen

*Ed. Note: Mouseover the blue links to see concepts demonstrated visually.*

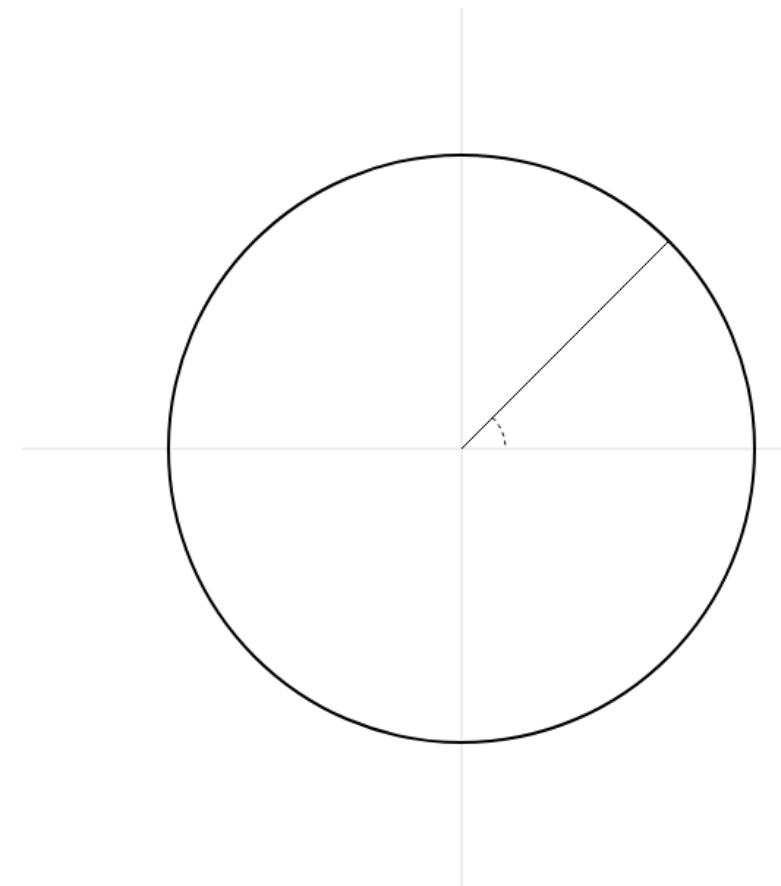
You may have learned about trigonometric functions such as **sine** and **cosine** as being defined by the ratios of sides of a triangle (SOHCAHTOA), or in terms of points and lines related to the unit circle. For me, it didn't totally click until I started to think visually about inscribing a triangle inside of a circle.

For example, you can think about the sine function as measuring [the distance from the x-axis](#) of a point on the unit circle at a particular angle. The sign (+/-) of that value indicates if the point lies above or below the axis. Similarly the cosine can be thought of as measuring [the distance from the y-axis](#) of that same point.

It is useful to note that the cosine of an angle is the same as the sine of the complement of the angle. In other words, it is the same operation as sine, just with respect to the y-axis instead of the x-axis.

The word *sine* originally came from the latin *sinus*, meaning "bay" or "inlet". However, it had a long path to get there. The earliest known reference to the sine function is from Aryabhata the Elder, who used both *ardha-jya* (half-chord) and *jya* (chord) to mean *sine* in Aryabhatiya, a Sanskrit text finished in 499 CE.

Jya, meaning chord, became *jiba* in Arabic, and was abbreviated as just *jb*. When the term was translated to latin in the twelfth century, *jb* was incorrectly



$$\sin(0.79) = 0.7071$$

$$\cos(0.79) = 0.7071$$

$$\tan(0.79) = 1.0000$$

$$\cotan(0.79) = 1.0000$$

$$\sec(0.79) = 1.4142$$

$$\cosec(0.79) = 1.4142$$

CLICK TO SELECT ANGLE OR FUNCTION

# Narrative Visualization Examples

- <https://www.nytimes.com/interactive/2015/05/28/upshot/you-draw-it-how-family-income-affects-childrens-college-chances.html>

 TheUpshot



## You Draw It: How Family Income Predicts Children's College Chances

By GREGOR AISCH, AMANDA COX and KEVIN QUEALY MAY 28, 2015

How likely is it that children who grow up in very poor families go to college? How about children who grow up in very rich families?

We'd like you to **draw your guess** for every income level on the chart below.

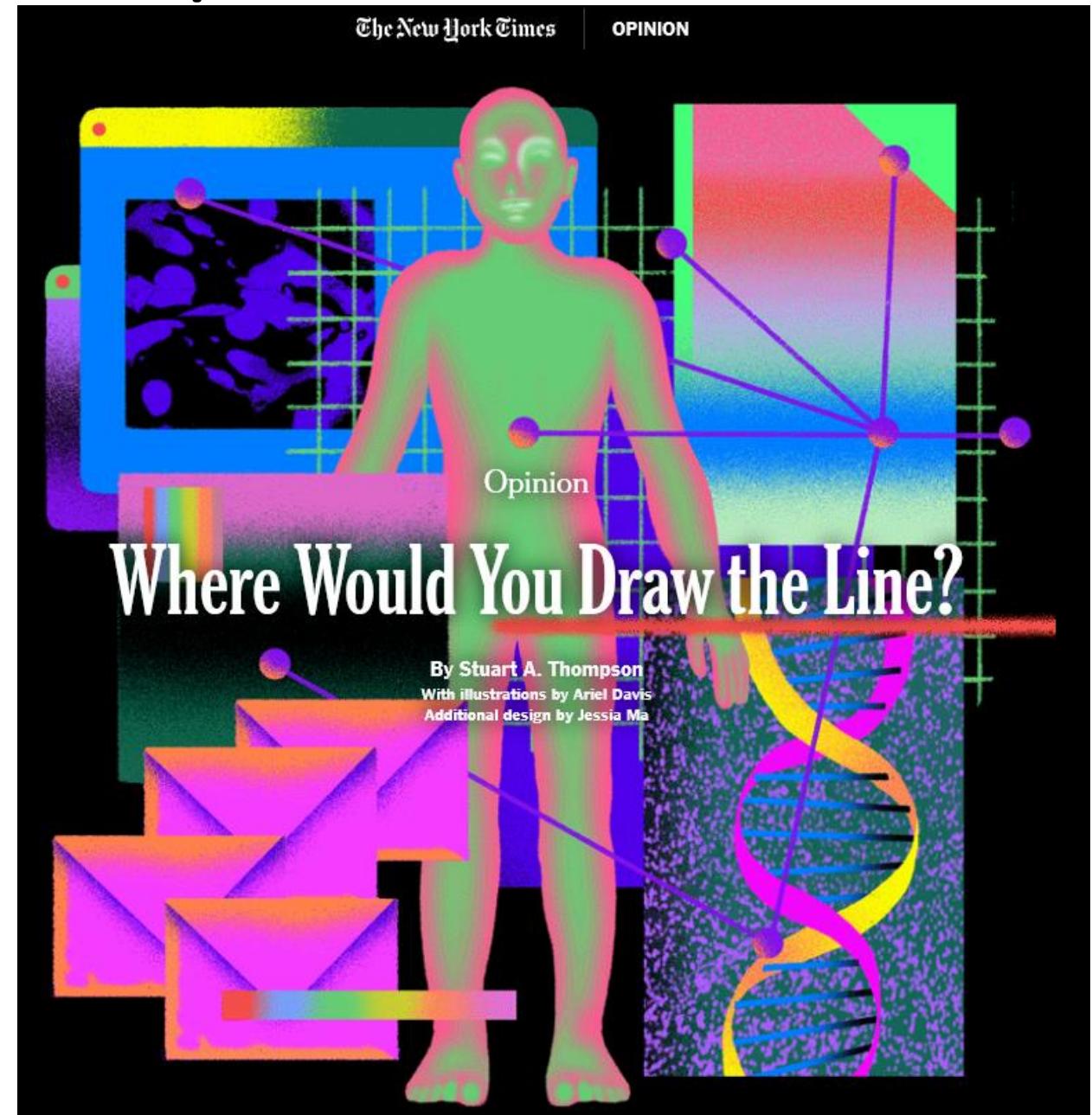
If you think the chances of enrolling in college (or vocational school) are about the same for everyone, you should draw something like this:

— . If you think the odds are especially harsh for children from the poorest families, but higher for middle- and higher-income children, your drawing would instead look like this: ↗ . Or here is one for a situation in which chances level off after a certain income threshold: ↘ . Or for one that spikes ↗ or dips ↘ for the very richest.

When you've finished drawing, we'll compare your line to the reality for children born in the early 1980s, based on research by a team of

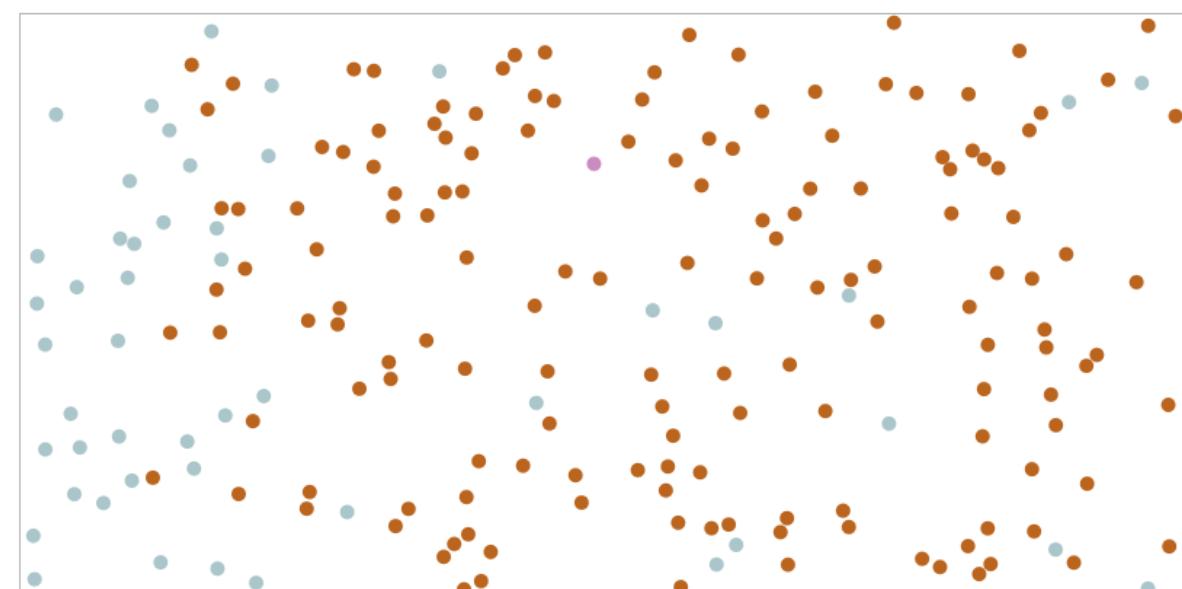
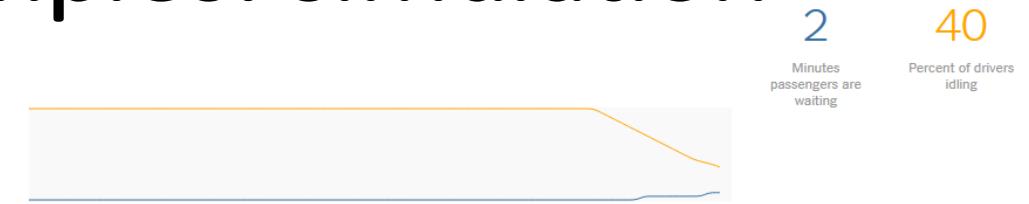
# Narrative Visualization Examples

- <https://www.nytimes.com/interactive/2019/04/10/opinion/privacy-survey.html>



# Narrative Visualization Examples: simulation

- <https://www.nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html>
- <https://www.washingtonpost.com/graphics/2020/world/corona-simulator/>
- <https://projects.fivethirtyeight.com/parole-assessment-simulator/>





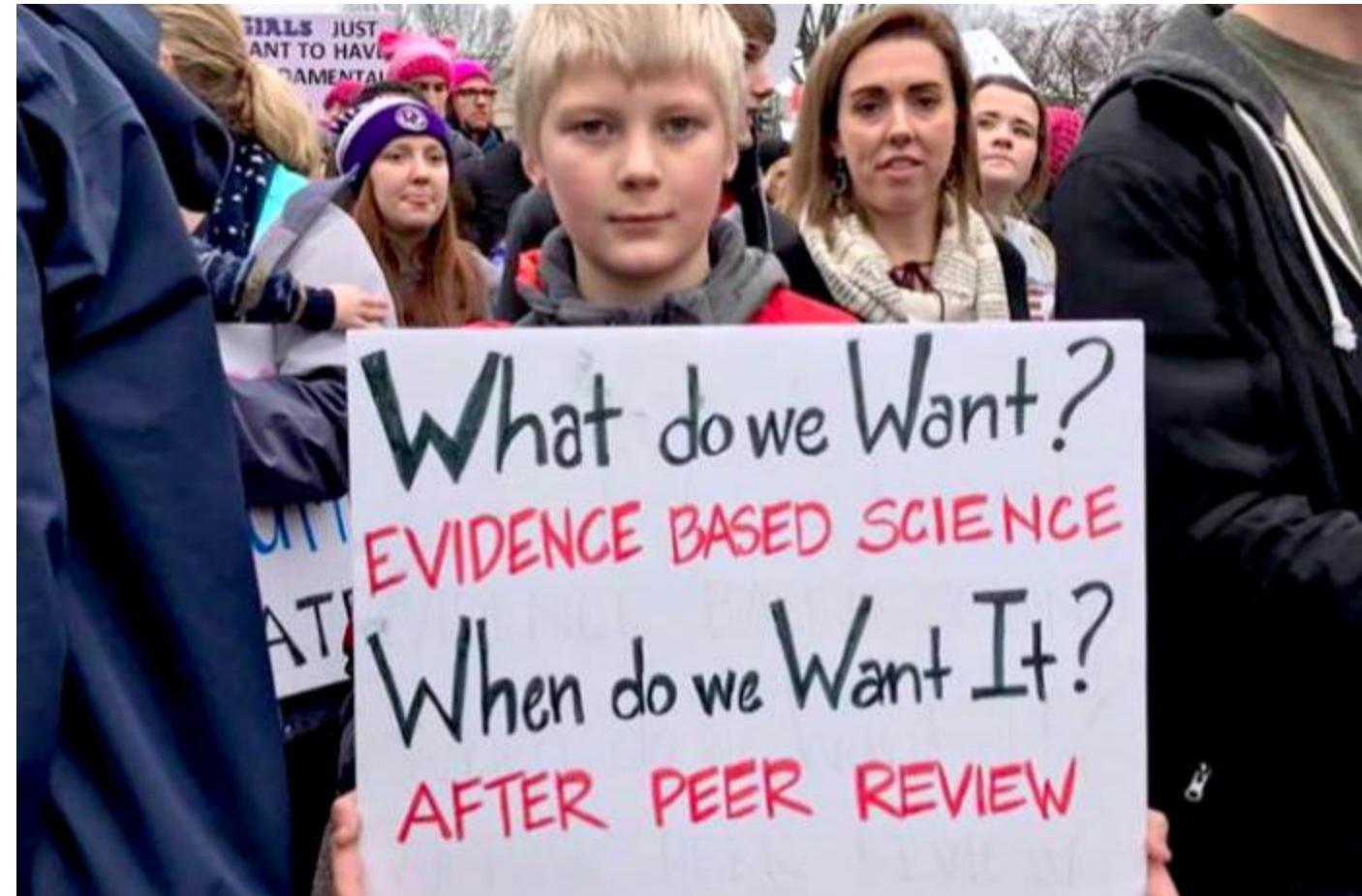
Break  
Back at 1:07pm



# User Study Crash Course

# Empirical Research

- We want to find an answer
- Need data!
- Run empirical studies to collect this data



<https://twitter.com/LDMay/status/822922129637330944>

# What is science?

- Both *knowledge* and a *process*
- Establish and organize facts → body of knowledge
- Facts should be **falsifiable**
  - Make predictions
  - Validate models

# Is computer science science?

- Sometimes!
  - Math
  - Engineering
  - Art
  - Psychology
- When appropriate, we would like it to be science
  - Pretty efficient way of getting things done
- We are historically not great at this
  - Random sample of ACM papers in 1993: 40% with claims that needed empirical evidence had none
    - Vs. 15% in optical engineering, etc. – (Tichy et al., 1995)

# Empirical Research for viz

- **Why?**
  - Validate visualization effectiveness
  - Identify usability issues
  - Gather user feedback
- **How?**
  - Systematic and controlled
  - Different kinds of user studies
    - Exploratory studies
    - Observational studies
    - Hypothesis testing
- **Goals:**
  - Representation and modeling: learn something true about a phenomenon
  - Generalizability: do your findings apply beyond the specific study?

# Empirical Research for viz

- **Define a research question.**
  - What do you want an answer to?
- **Identify a target audience**
  - This defines your participant pool that you will recruit
  - Does your target audience have specialized knowledge or unique characteristics?
- **What are your outcome variable(s)?**
  - Task completion time
  - Information retention
  - User satisfaction
- **How will you measure outcome(s)?**
  - “Think aloud” study
  - Task-based study
  - Interviews
  - Surveys/questionnaires

# Empirical Research for viz

- **Design your experiment**
  - Give users a task that will generate the data (outcome variables) you need
  - Task should usually be specific and easy-to-understand
    - Users do unpredictable things, so you want to minimize likelihood of this happening
  - Obtain **informed consent**
  - Write a script for your procedure, follow it for each participant
    - Do not want to bias the results via different instructions and procedures!
- **Data collection (with consent)**
  - Log keys and mouse movements
  - Record screen
  - Record audio for think-aloud and interview paradigms
  - **Use the same system setup for all participants**

# Interviews

- **Structured interview**
  - Specific list of questions that every participant answers
  - Do not deviate from this list
  - Pros:
    - Easier to compare across participants
    - Less noise
- **Unstructured interview**
  - No predetermined questions, more like a natural conversation
  - Pros:
    - Good for exploring subjective data (attitudes, feelings, task-completion strategies)
  - Cons:
    - People are bad at introspecting
- **Semi-structured interview**
  - Mix of structured and unstructured
  - Allows for follow-up questions if participants say something interesting
  - Pros:
    - Lots of flexibility
    - Middle-ground

# Questionnaires

- Likert scale
  - “Rate X on a scale from 1 – 7”
  - Usually has a middle point
- Be careful with wording
  - "How great was our hard-working customer service team?"
  - "How would you describe your experience with the customer service team?"
- Avoid loaded questions that ask multiple things
  - Be specific
- If possible, probably best to use an existing questionnaire that has been validated
- English is not everyone's first language