

Storytelling with Data

Nam Wook Kim

Mini-Courses — January @ GSAS
2018

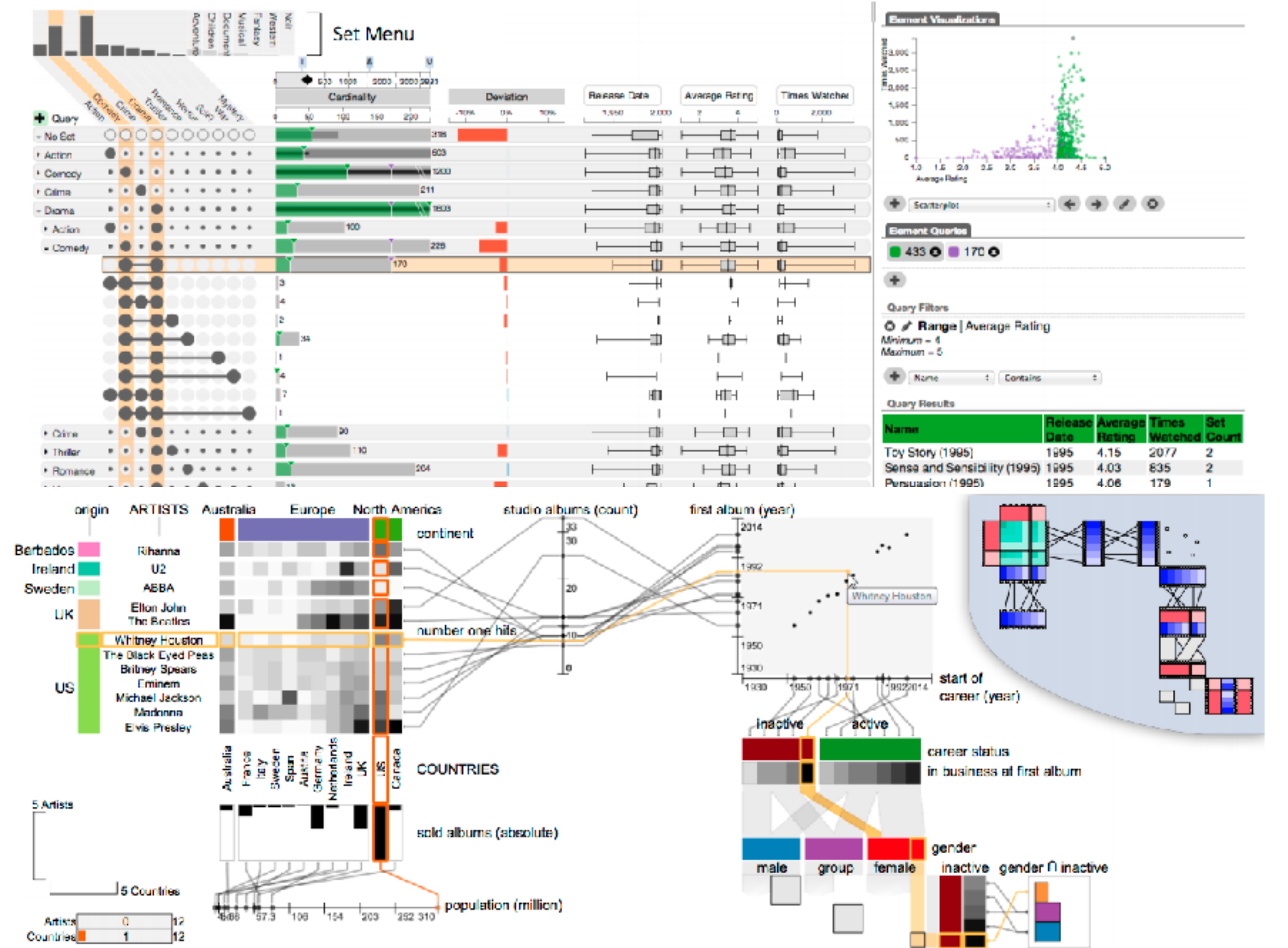
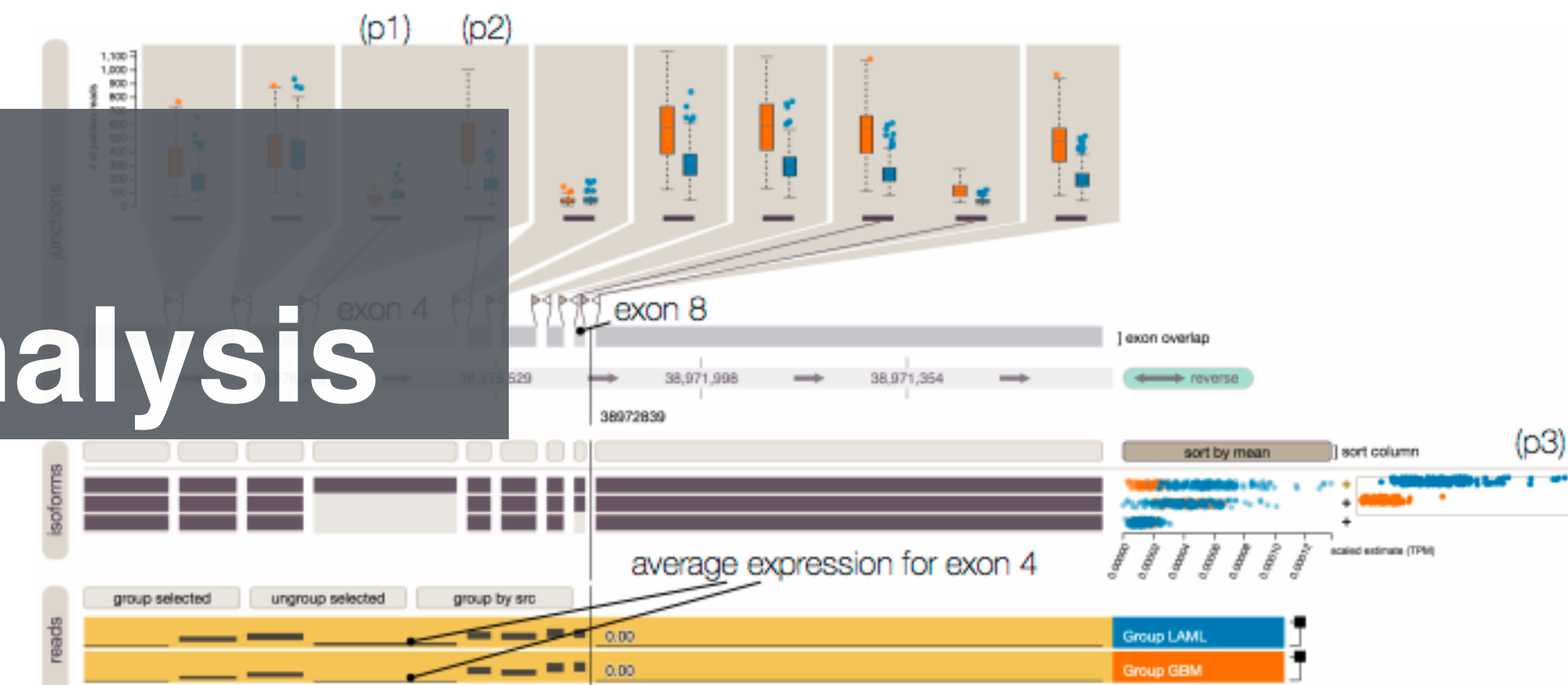
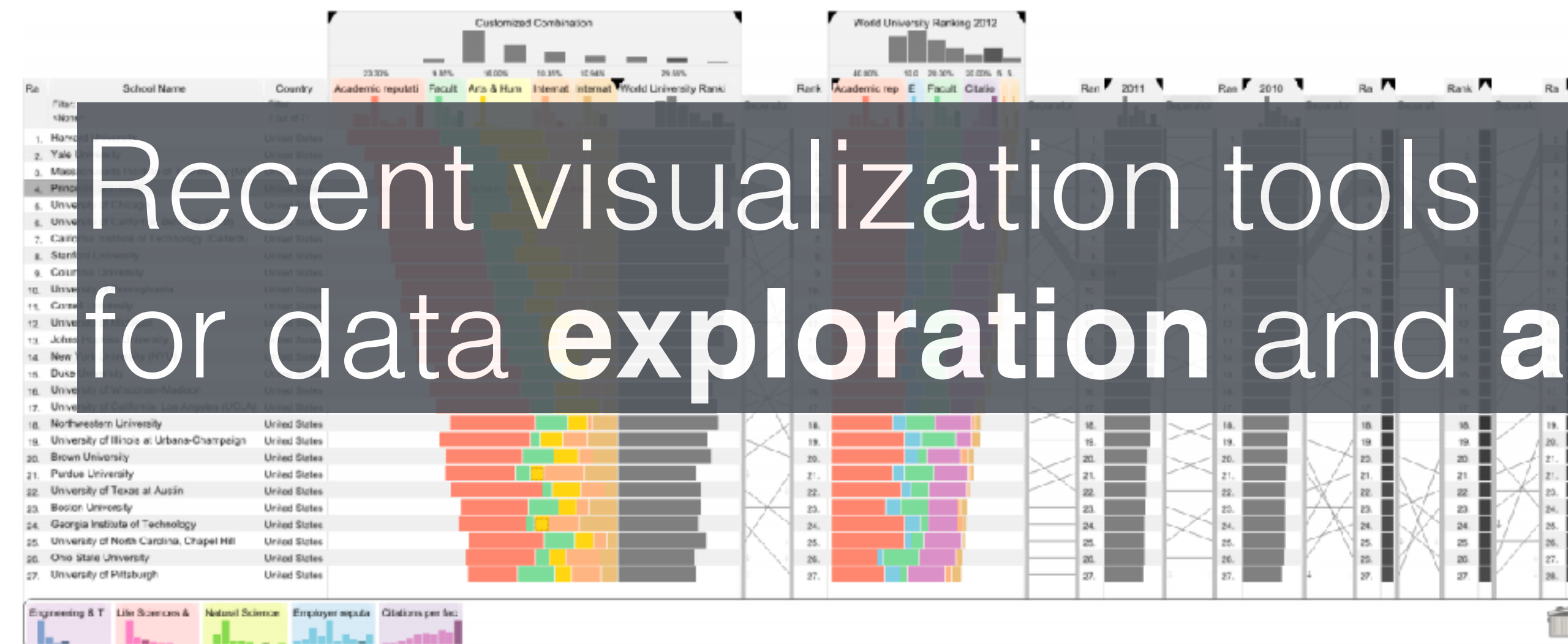
Goal

From exploratory analysis
to communication

Topics

- Visualization for communication
- Data-Driven Storytelling
- Presentation & Storytelling in Tableau

Recent visualization tools for data exploration and analysis



Data → Explore →

?

- Data Centered
- Domain Experts
- Discovering Insights

Need for **communicating** large, complex, and diverse data.



Open Government



Health & Medicine

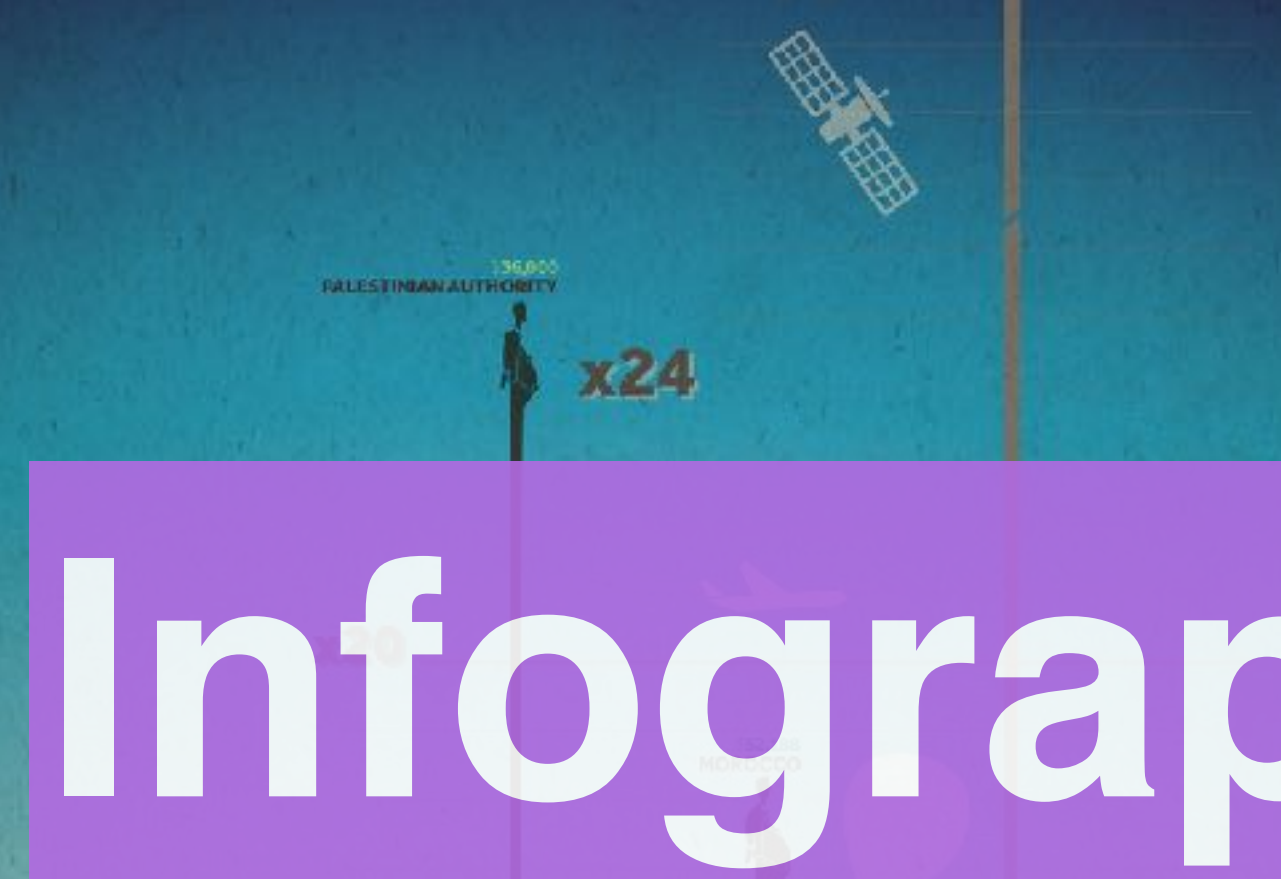


Sensor & Devices



Social Media

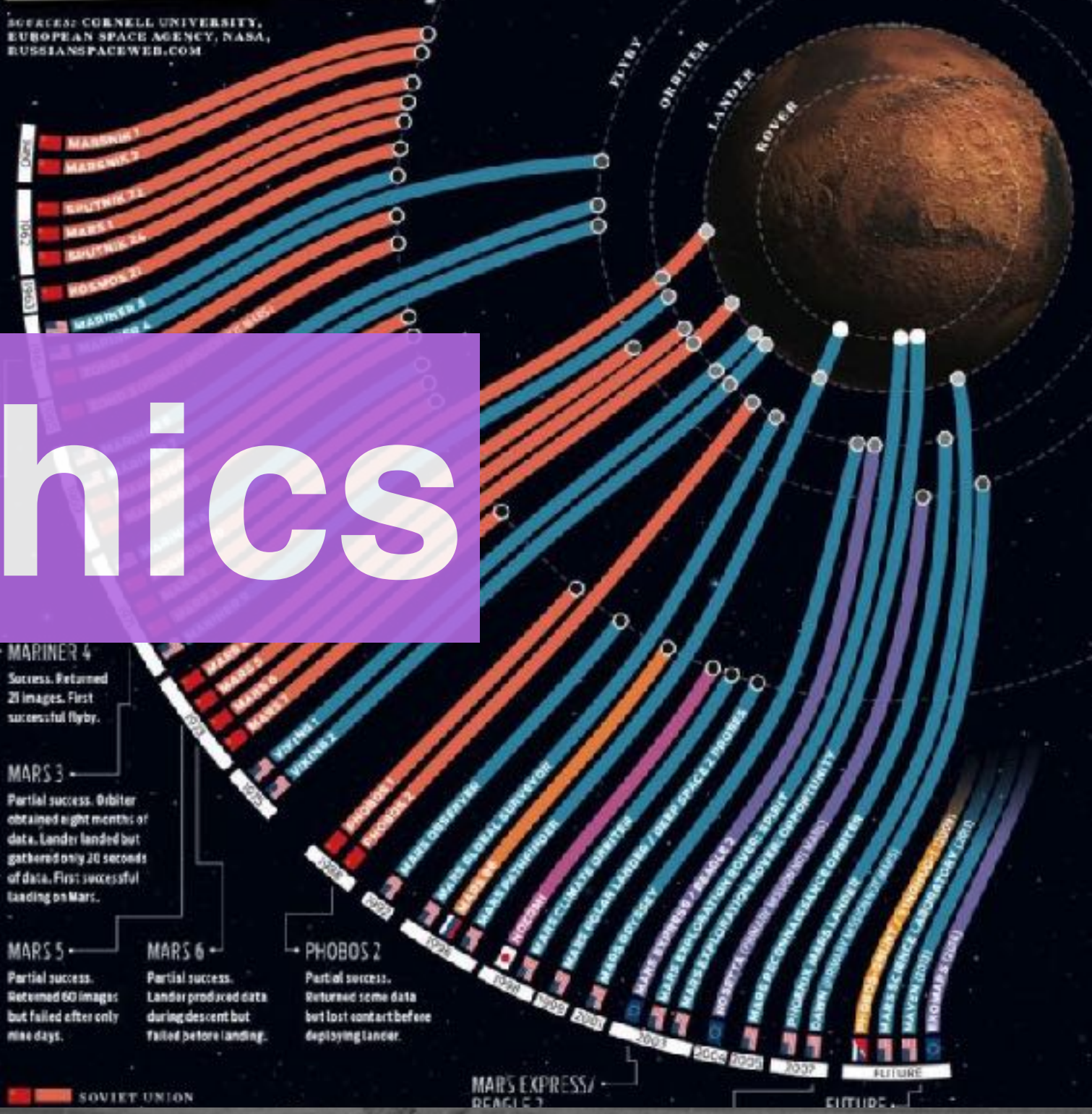
POLITICIANS' SALARIES & INCOME INEQUALITY



Infographics

MISSION(S) TO MARS

Though most missions to the red planet have failed, the success rate is definitely improving + + +



U.S. HISTORY OF TAXATION ON CIGARETTES

IN BILLIONS OF DOLLARS

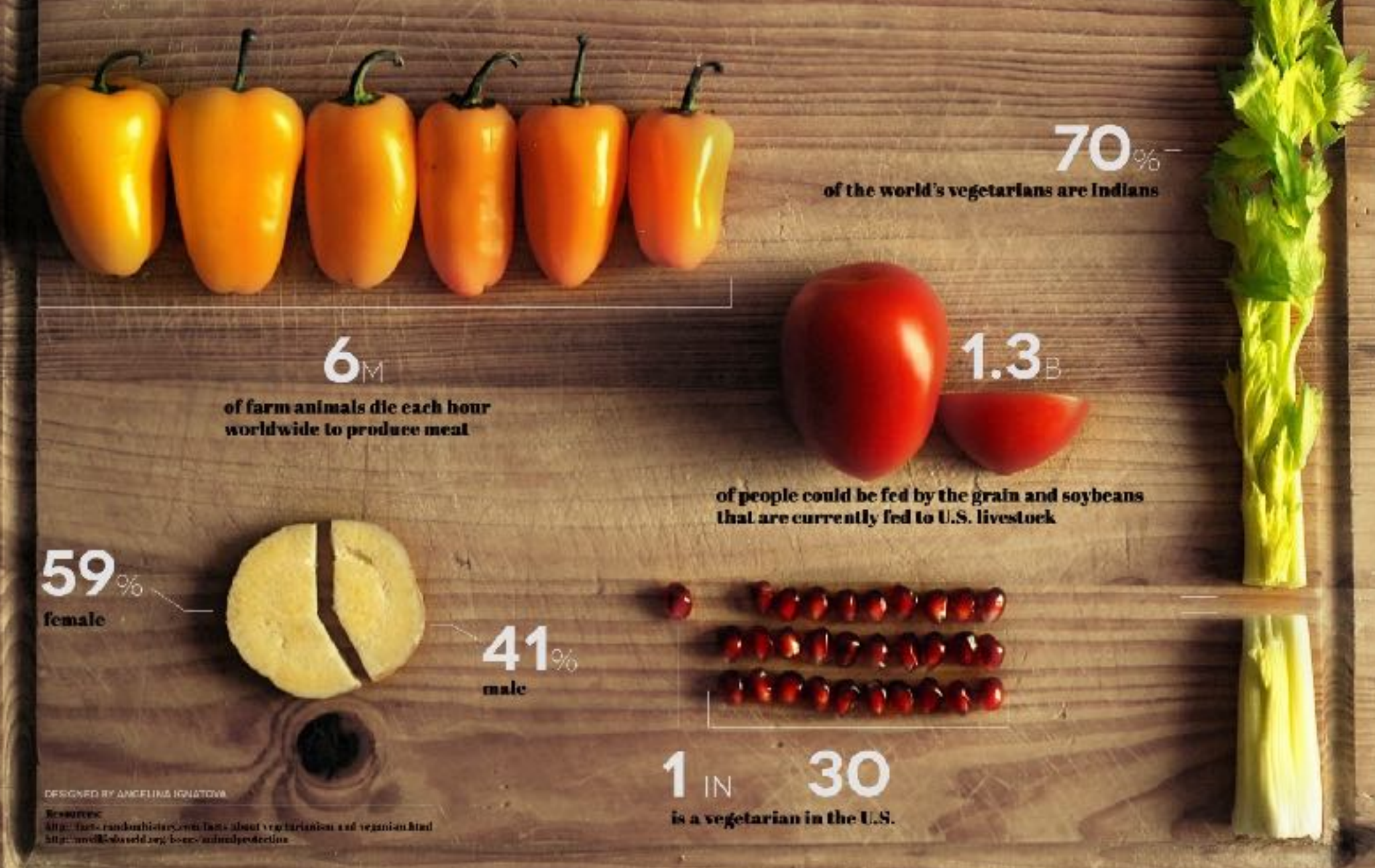


42.5% Thinks it is unethical if schoolteachers and educators go to work wearing an Islamic headscarf
46% Thinks it is unethical if nurses and doctors go to work wearing an Islamic headscarf
66% Thinks it is unethical if judges go to work wearing an Islamic headscarf



38% Thinks it is unethical if house cleaners go to work wearing an Islamic headscarf

VEGETARIANISM IN NUMBERS



VISUALIZINGIMPACT
 SOURCES:
 Politician salary data for 100 countries or regions are available from the Transparency International website. GDP per capita data is from the CIA World Factbook.

DESIGNED BY ANGELINA IGATOVA
 RESEARCH:
 http://www.randushinets.com/facts-about-vegetarianism-and-veganism.html
 http://www.world.org/body-vegetarianism



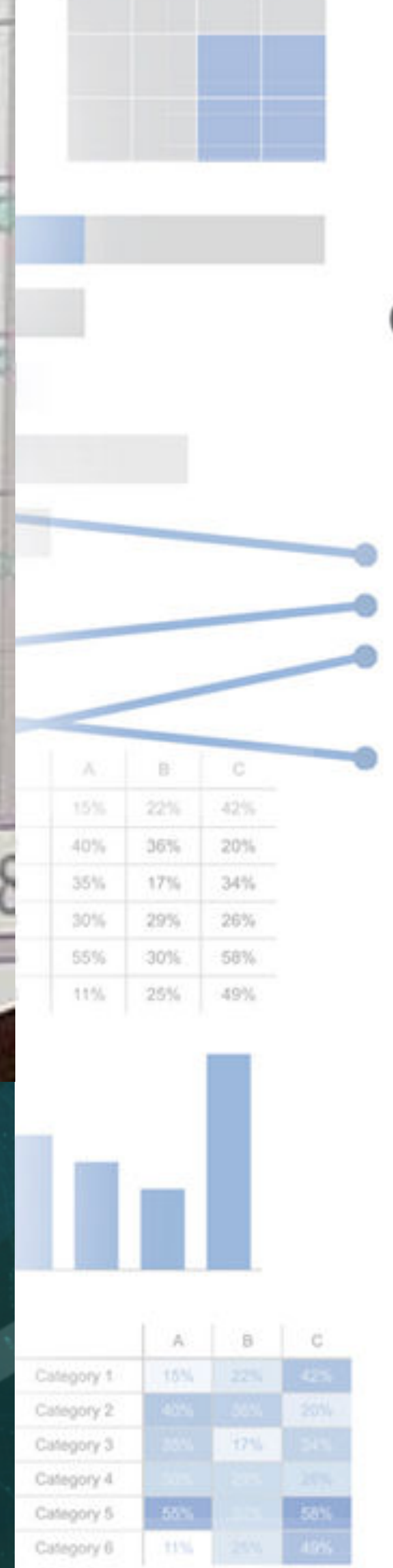
Storytelling

Hans Rosling

tapestry

Storytelling is the most powerful way to put ideas into the world today.

Robert McKee



cole nussbaumer knaflic

storytelling with data

a data visualization guide for business professionals

WILEY

How Far Is Europe Swinging to the Right?

by GREGOR ADEZI, ADAM PEARCE, and BRYAN BURTON | Updated December 1, 2016

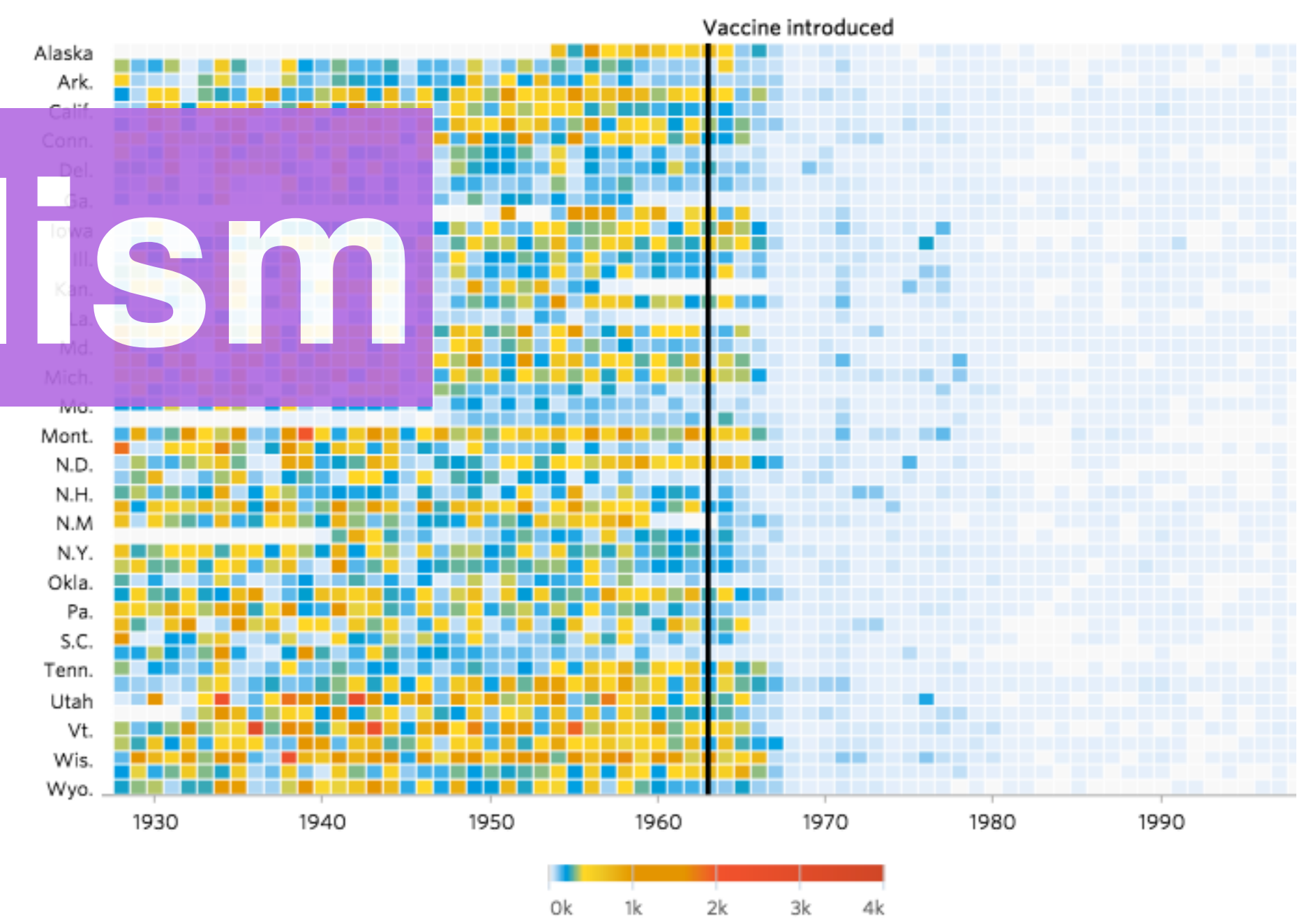
Avoid a migrant crisis, sluggish economic growth and growing disillusionment with the European Union, right-wing parties in a growing number of European countries have made electoral gains. The right-wing parties included below range across a wide policy spectrum, from populist and nationalist to far-right neo-fascist.

Party ideology in parliamentary elections*
 Center-left, center-right
 Other parties
 Right-wing and far-right

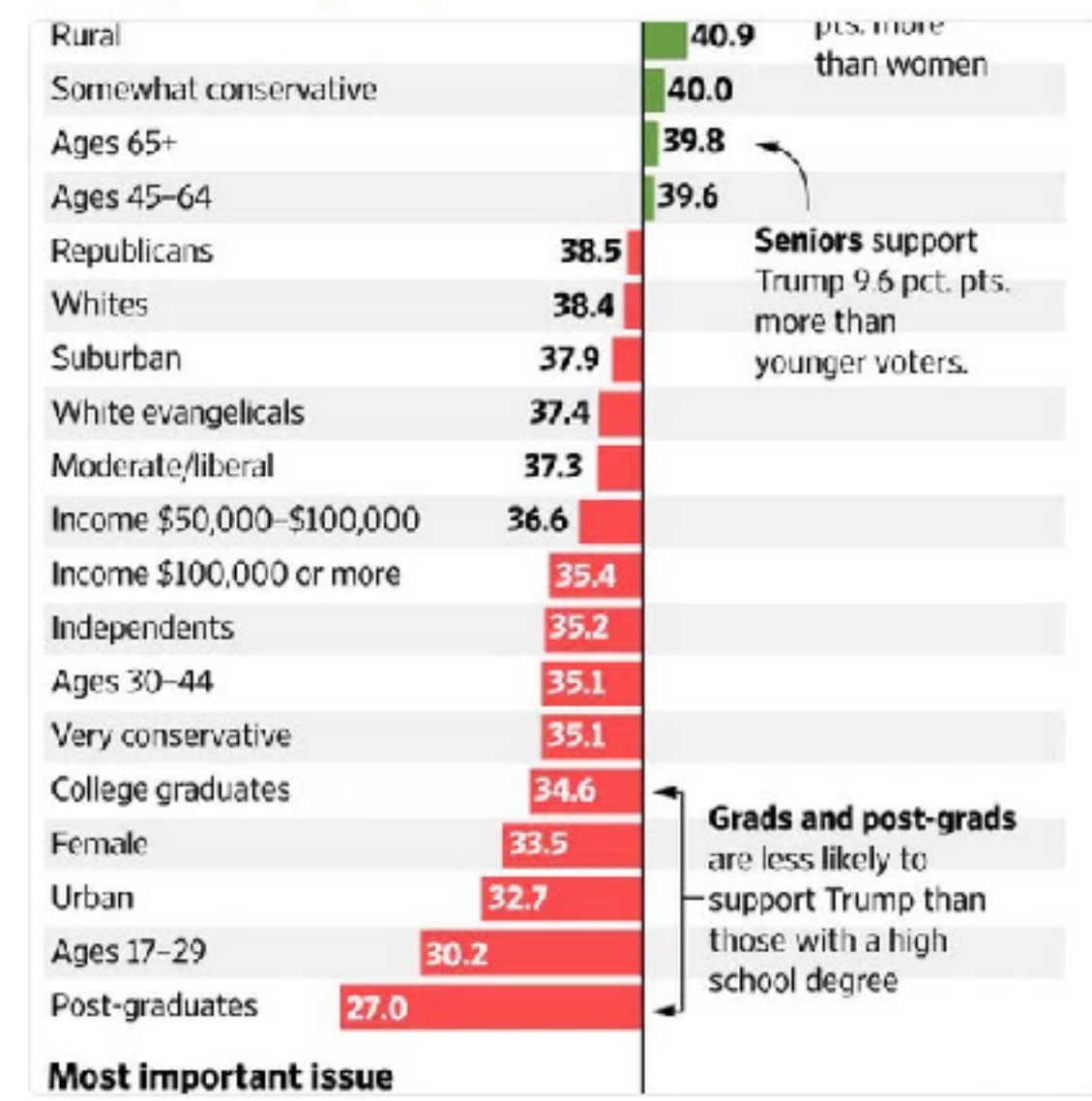
Data Journalism



Measles



Inside the Trump coalition: How he performs among voter groups

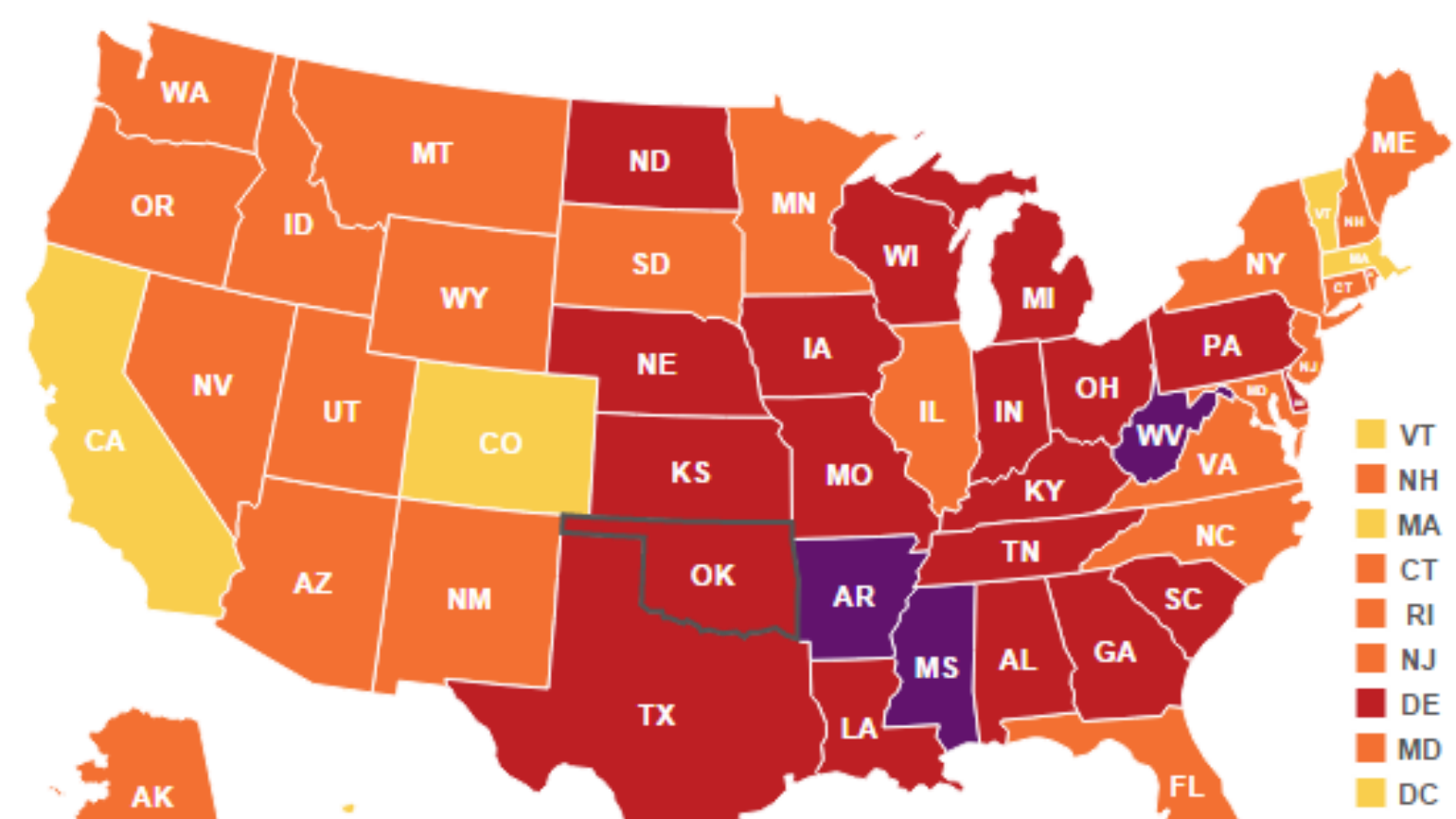


Adult Obesity Rate by State, 2014

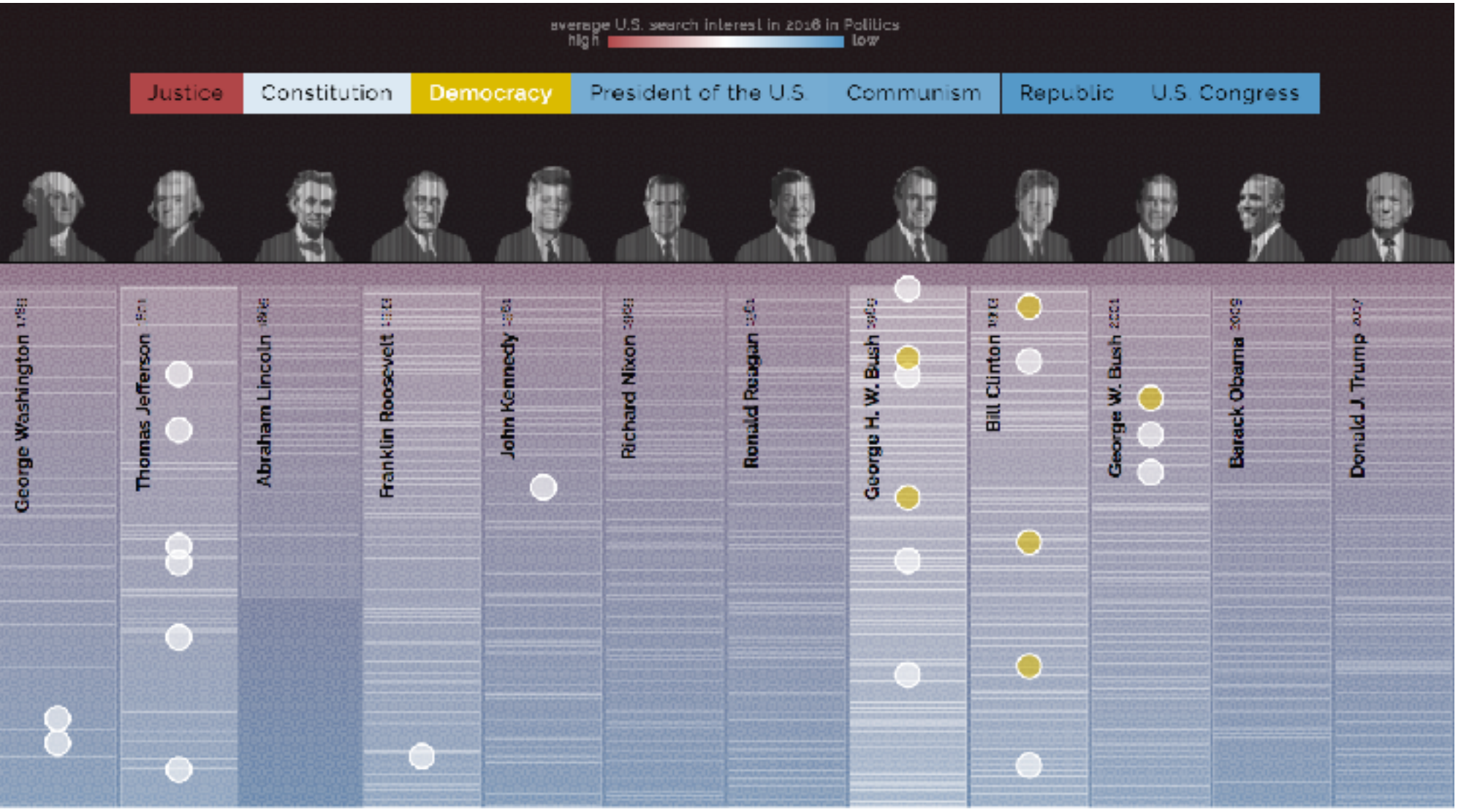
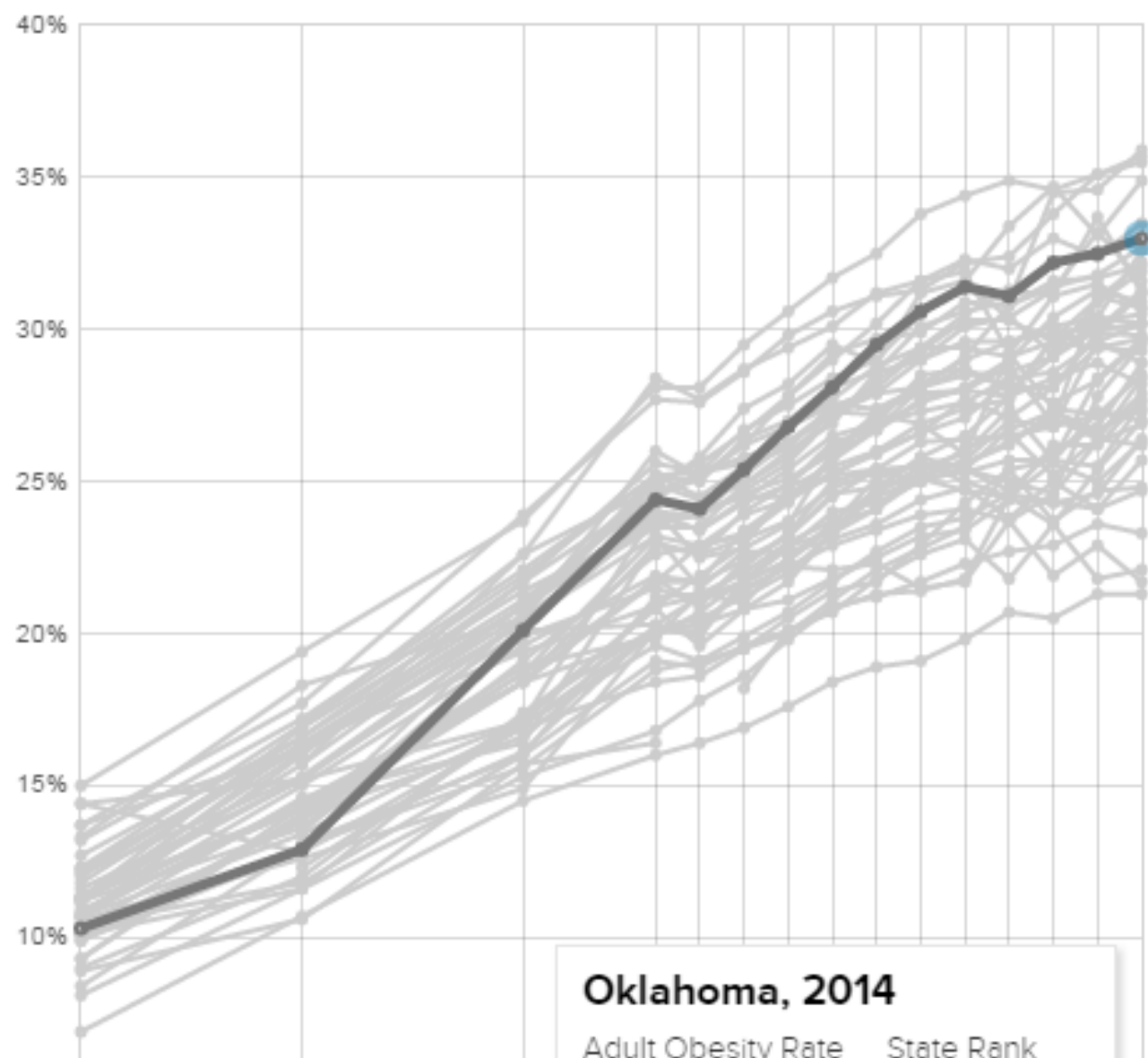
Select years with the slider to see historical data. Hover over states for more information. Click a state to lock the selection. Click again to unlock.

Percent of obese adults (Body Mass Index of 30+)

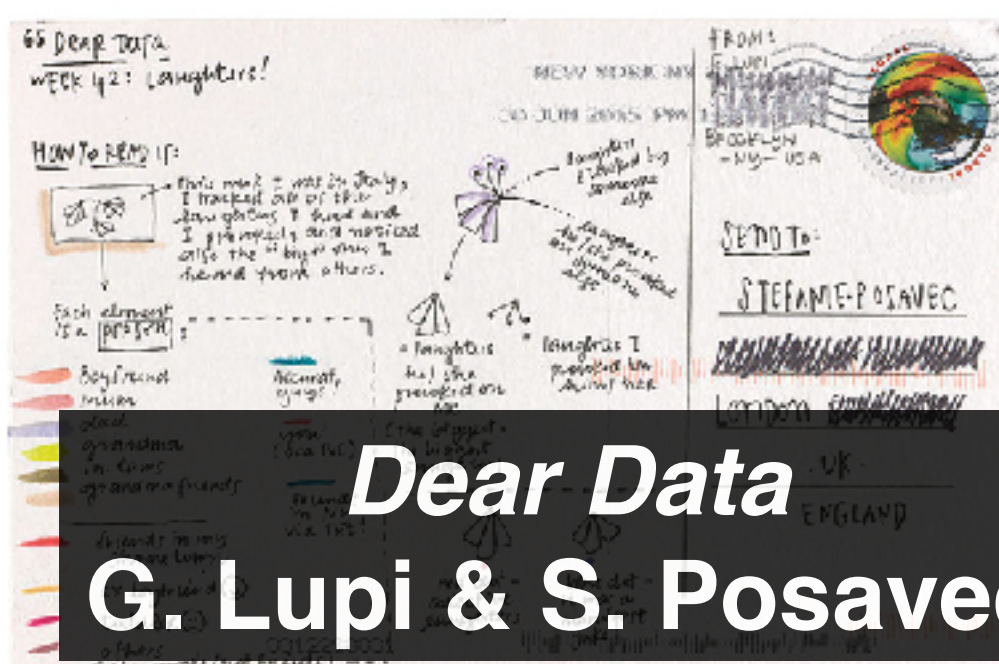
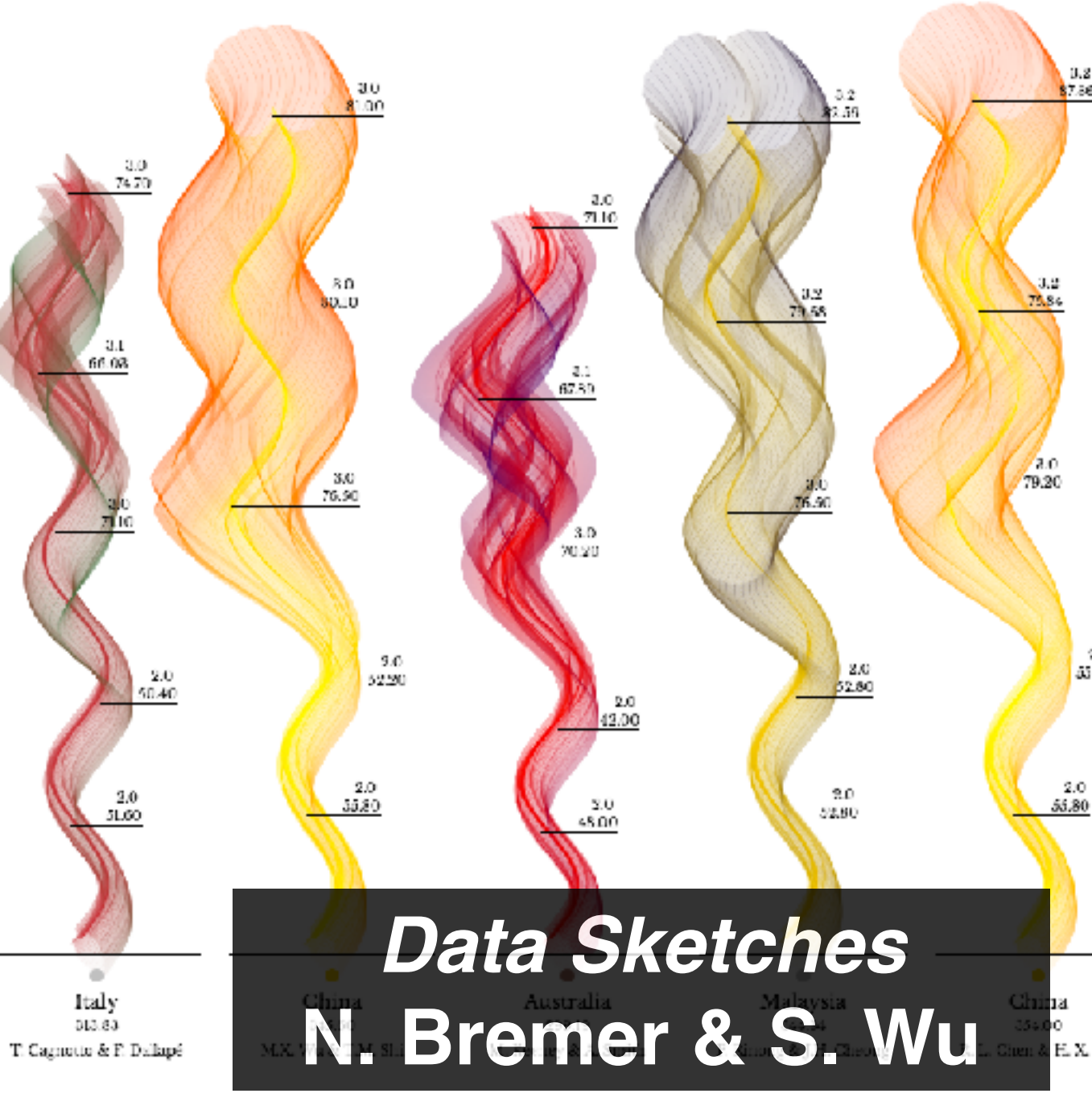
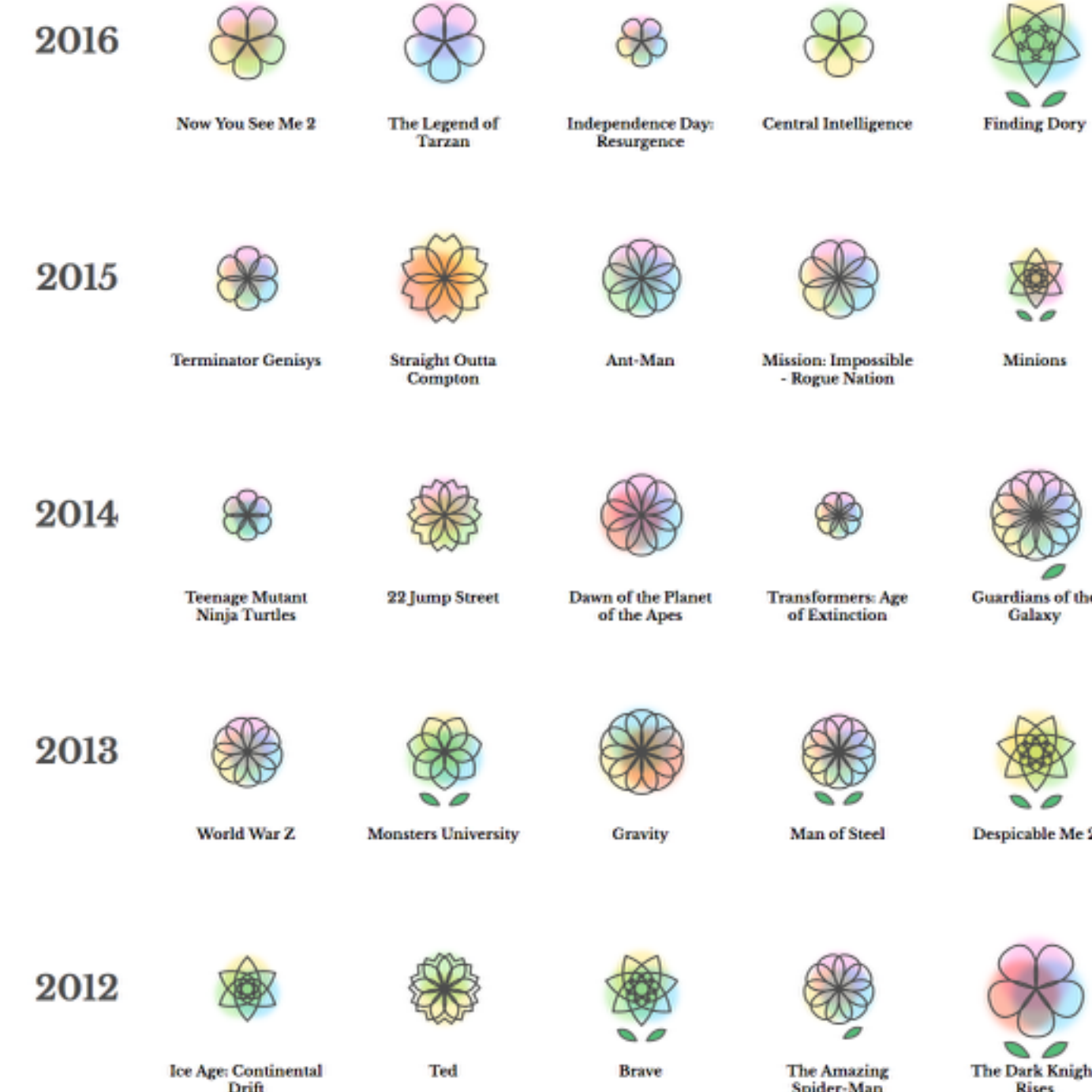
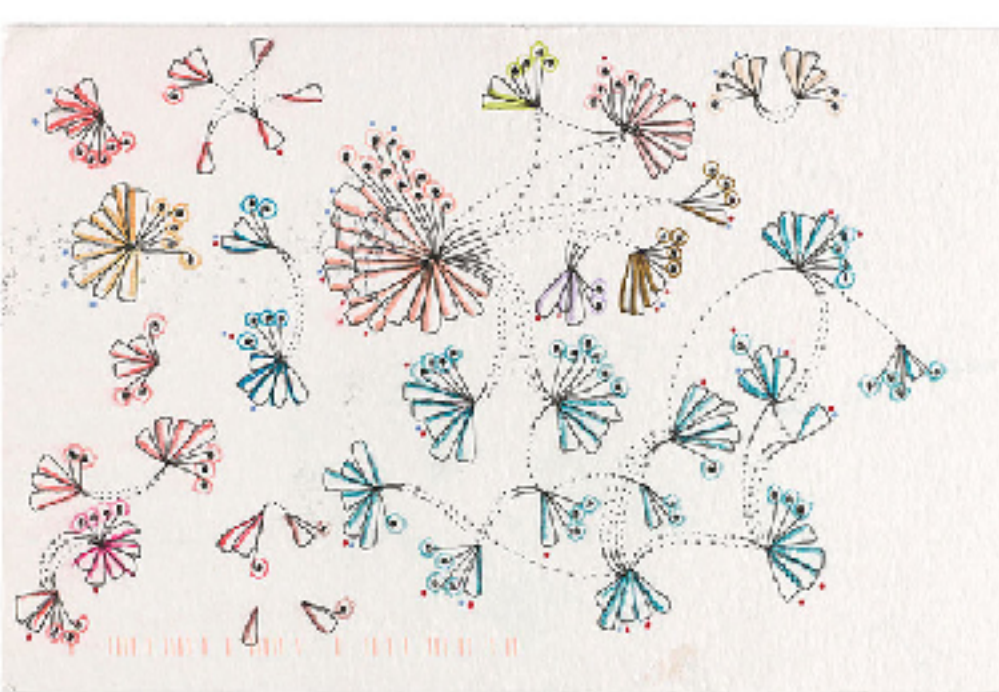
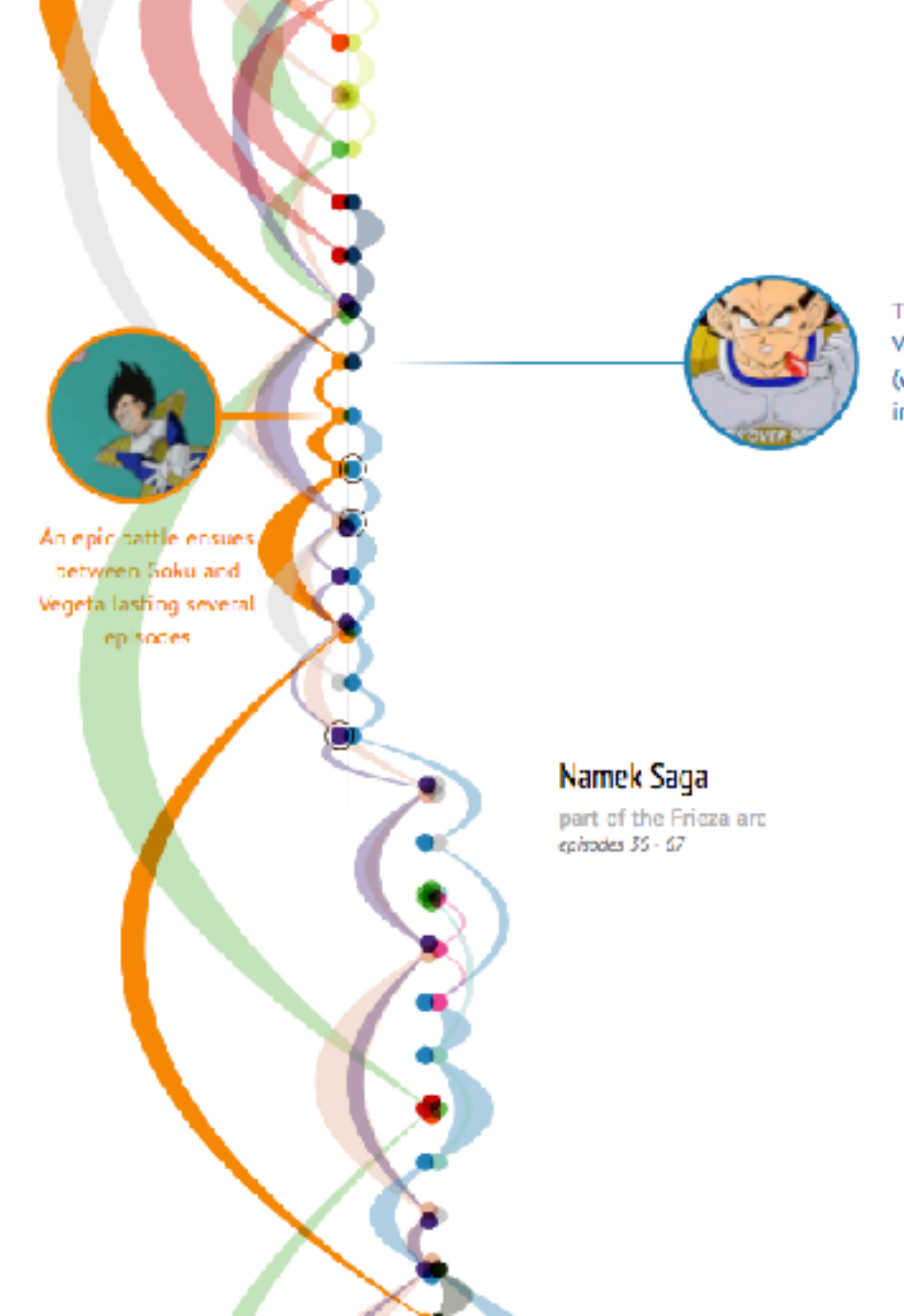
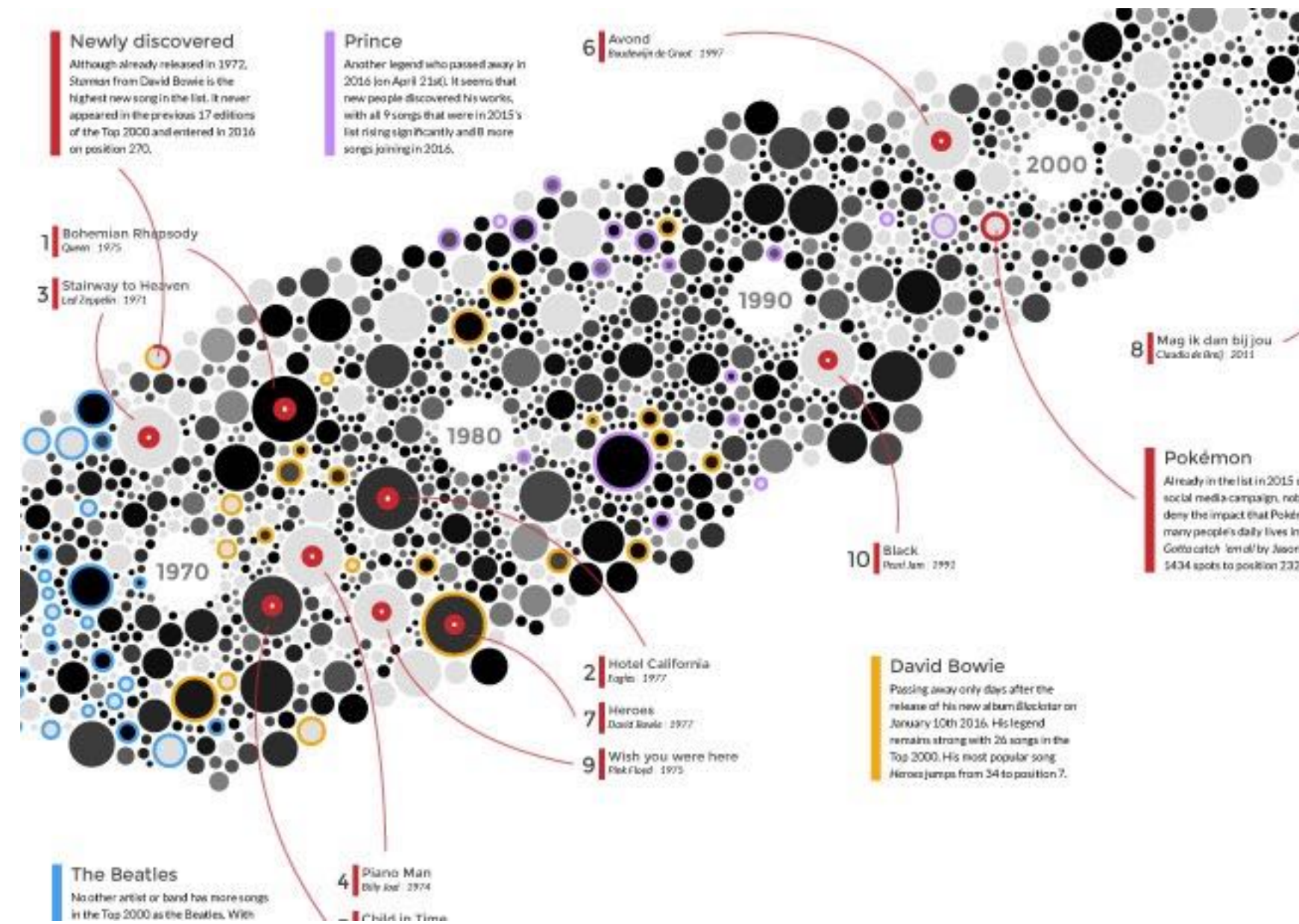
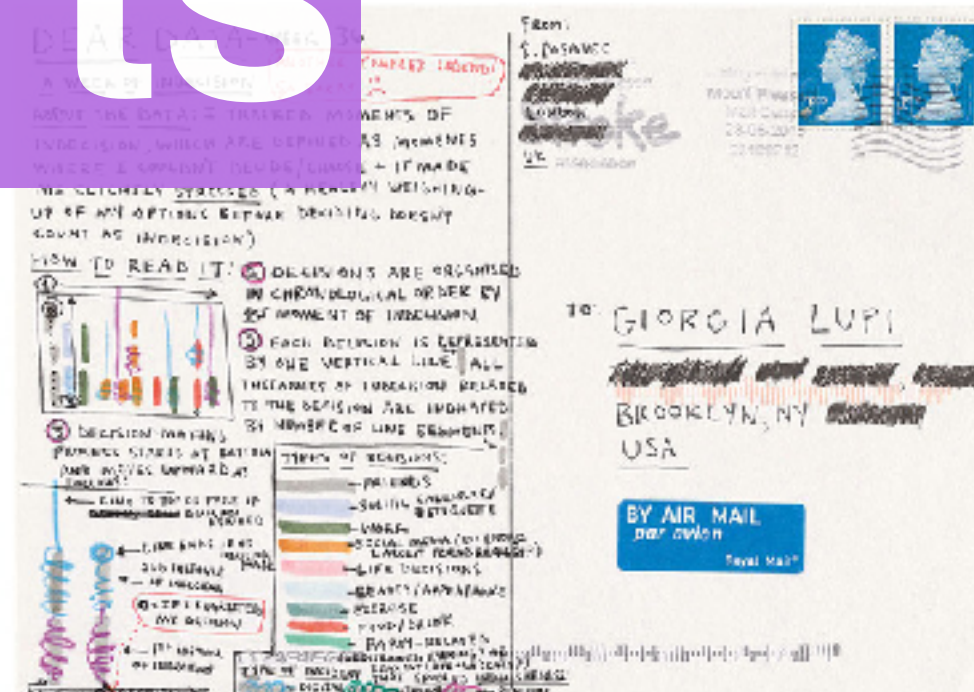
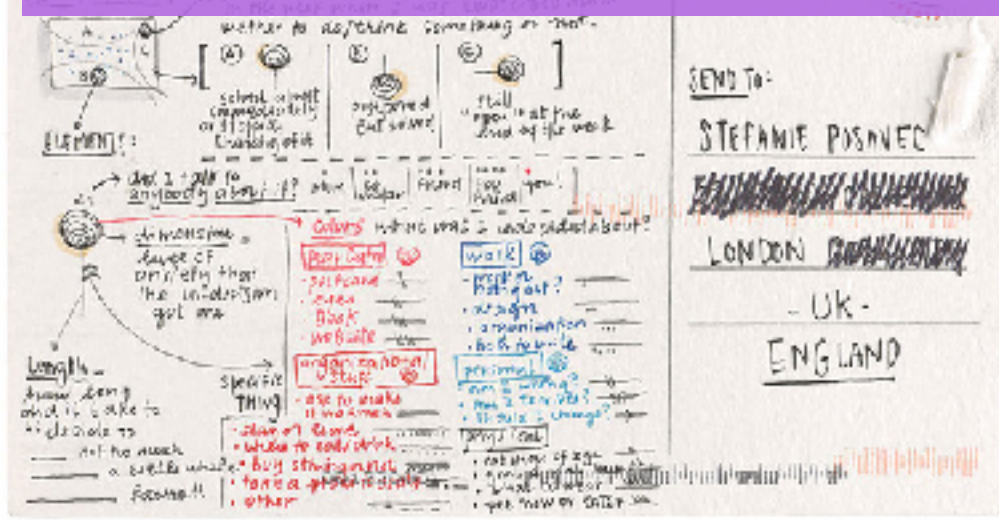
- 0 - 9.9%
- 10 - 14.9%
- 15 - 19.9%
- 20 - 24.9%
- 25 - 29.9%
- 30 - 34.9%
- 35%+



Adult obesity rates, 1990 to 2014



Data Arts



Dear Data
G. Lupi & S. Posavec

Data Sketches
N. Bremer & S. Wu

Data → Explore → Communicate

- Data Centered
- Domain Experts
- Discovering Insights

- Human Centered
- General Audience
- Conveying Messages

New Challenges

What is beyond data exploration?

What is a communication-oriented visualization?

How can we better support communication of data?

How can we tell compelling stories with data?

“

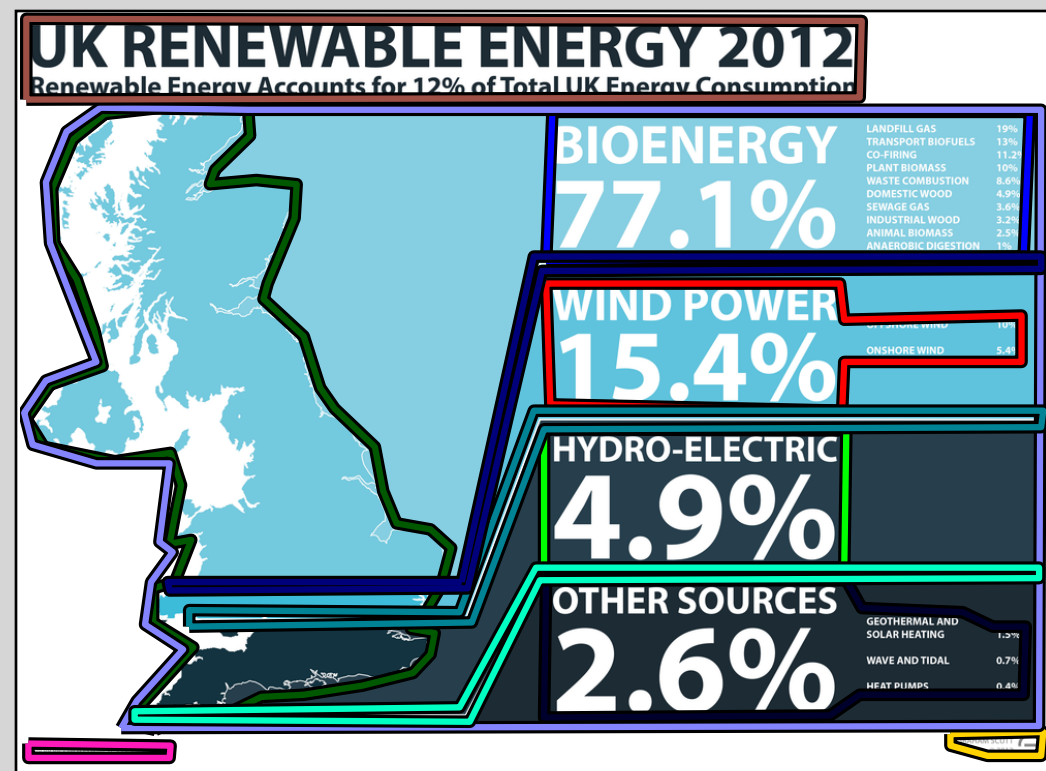
Presentation—specifically, its **use of elements from storytelling**—is the next logical step in visualization research and should be a focus of at least **equal importance** with **exploration** and **analysis**

— Kosara & Mackinlay 2013

Research:

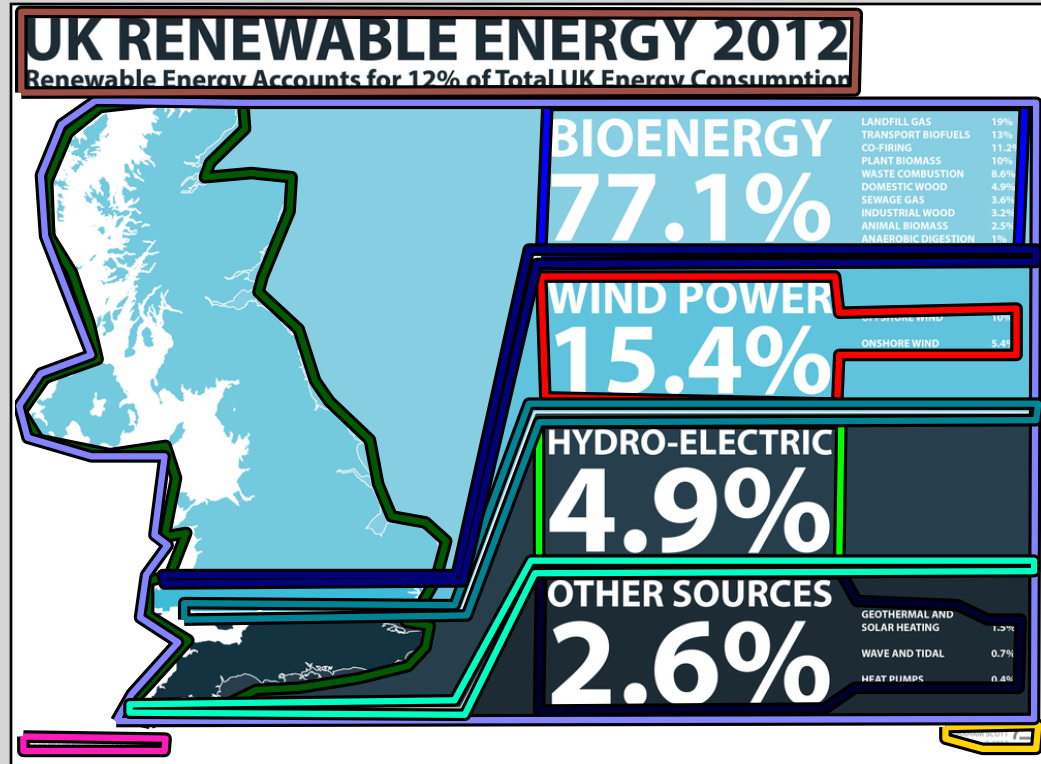
What makes a visualization **communicative**?

LABELED VISUALIZATION DATABASE




393 visualizations

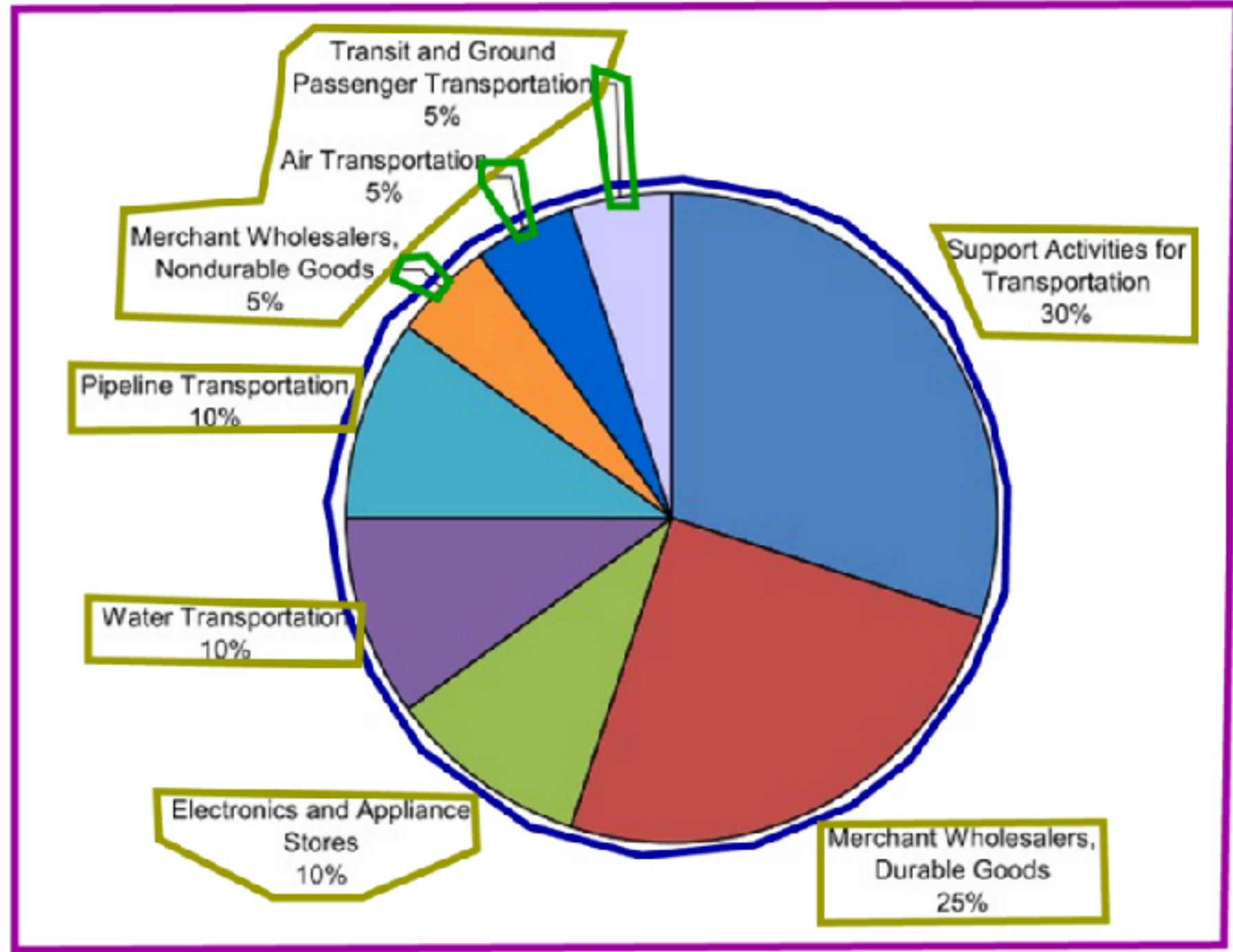
LABELLED VISUALIZATION DATABASE



393 visualizations



Total Wholesale, Retail, and Transportation 2009-2011



Category	Percentage
Support Activities for Transportation	30%
Merchant Wholesalers, Durable Goods	25%
Water Transportation	10%
Pipeline Transportation	10%
Electronics and Appliance Stores	10%
Merchant Wholesalers, Nondurable Goods	5%
Air Transportation	5%
Transit and Ground Passenger Transportation	5%

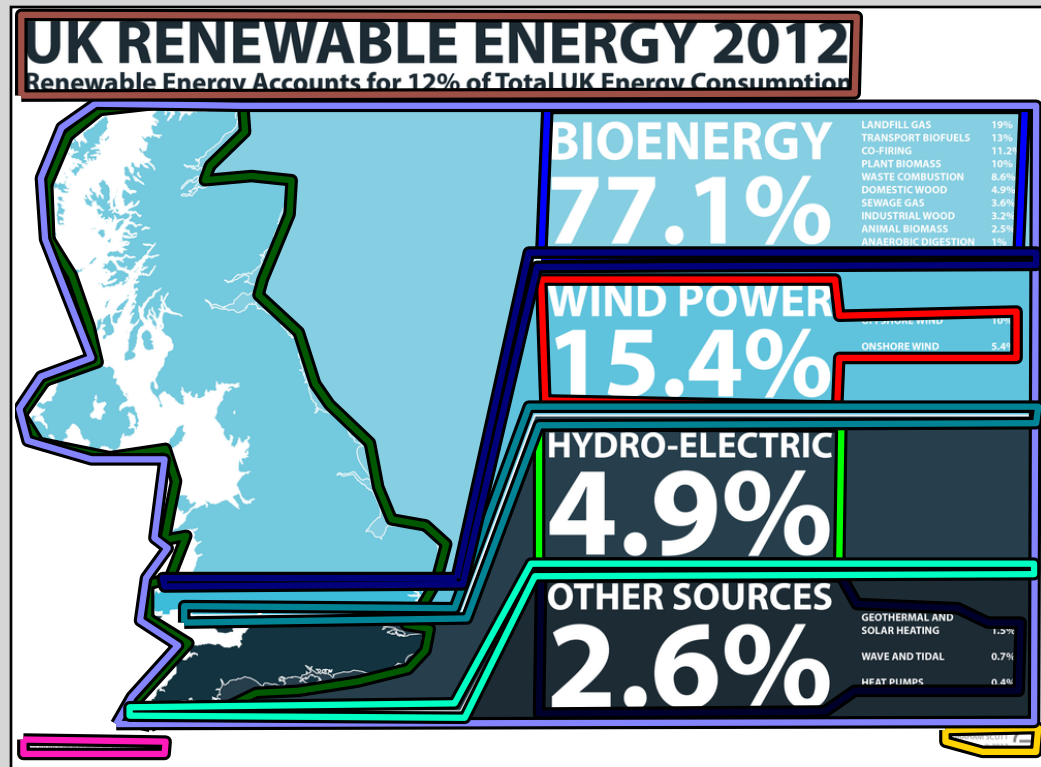
Covered Transactions from the Wholesale, Retail, and Transportation Sector

Polygons in this image (13)

[Hide all polygons](#)

- [text \(title\)](#)
- [text \(paragraph\)](#)
- [data \(circle\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [data](#)
- [annotation](#)
- [annotation](#)
- [annotation](#)

LABELLED VISUALIZATION DATABASE



393 visualizations

LabelMe [IMG] [XML] [ALL] [?] [»]

Total Wholesale, Retail, and Transportation 2009-2011

Category	Percentage
Support Activities for Transportation	30%
Merchant Wholesalers, Durable Goods	25%
Electronics and Appliance Stores	10%
Pipeline Transportation	10%
Water Transportation	10%
Merchant Wholesalers, Nondurable Goods	5%
Air Transportation	5%
Transit and Ground Passenger Transportation	5%

Polygons in this image (13)

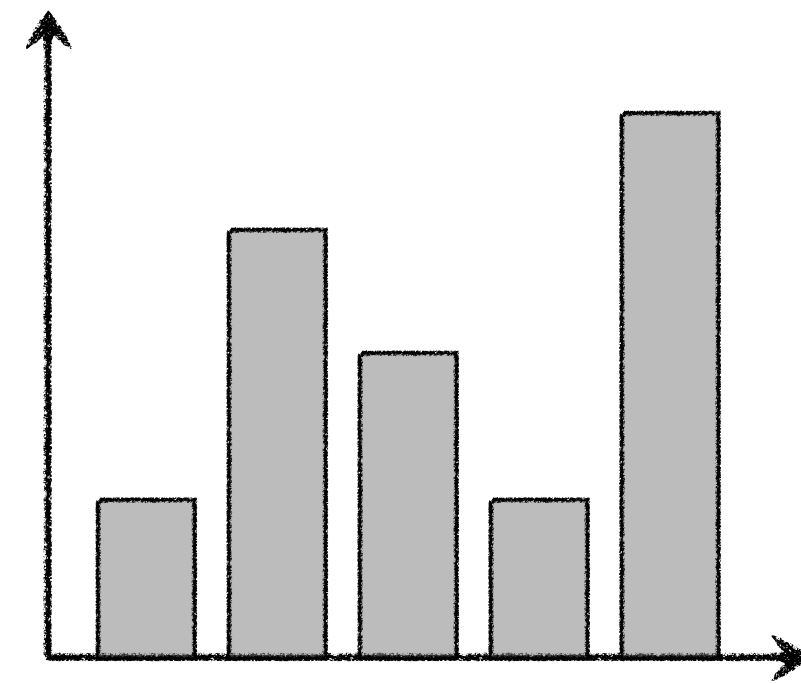
[Hide all polygons](#)

- [text \(title\)](#)
- [text \(paragraph\)](#)
- [data \(circle\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [text \(label\)](#)
- [data](#)
- [annotation](#)
- [annotation](#)
- [annotation](#)

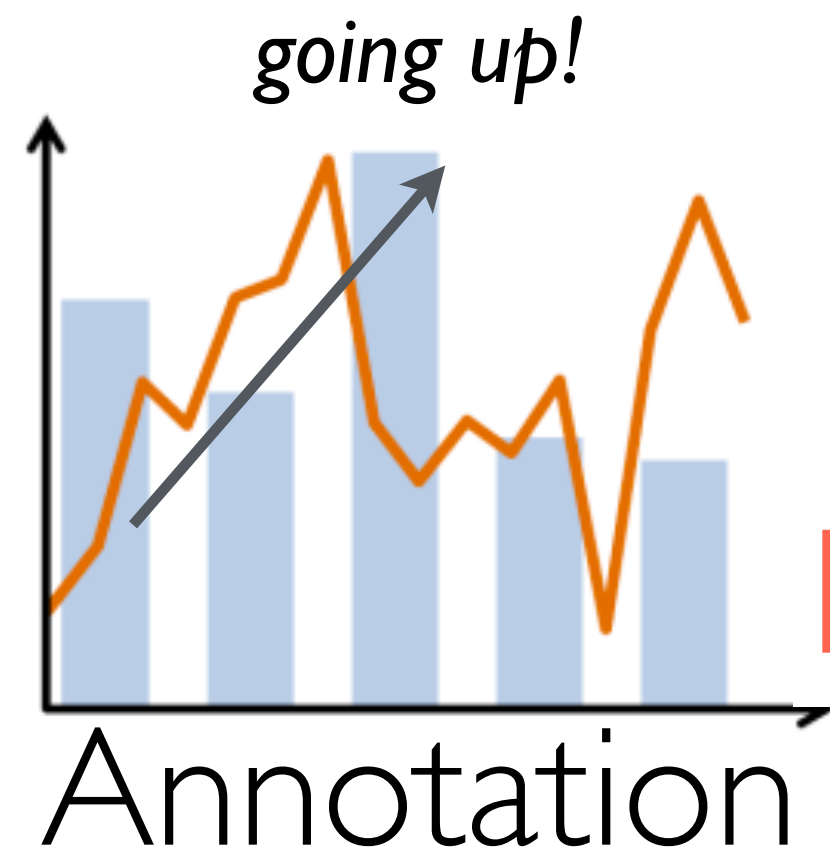
Covered Transactions from the Wholesale, Retail, and Transportation Sector

+ Data-Ink Ratio and Visual Density
Human Recognizable Objects
Data and Message Redundancy

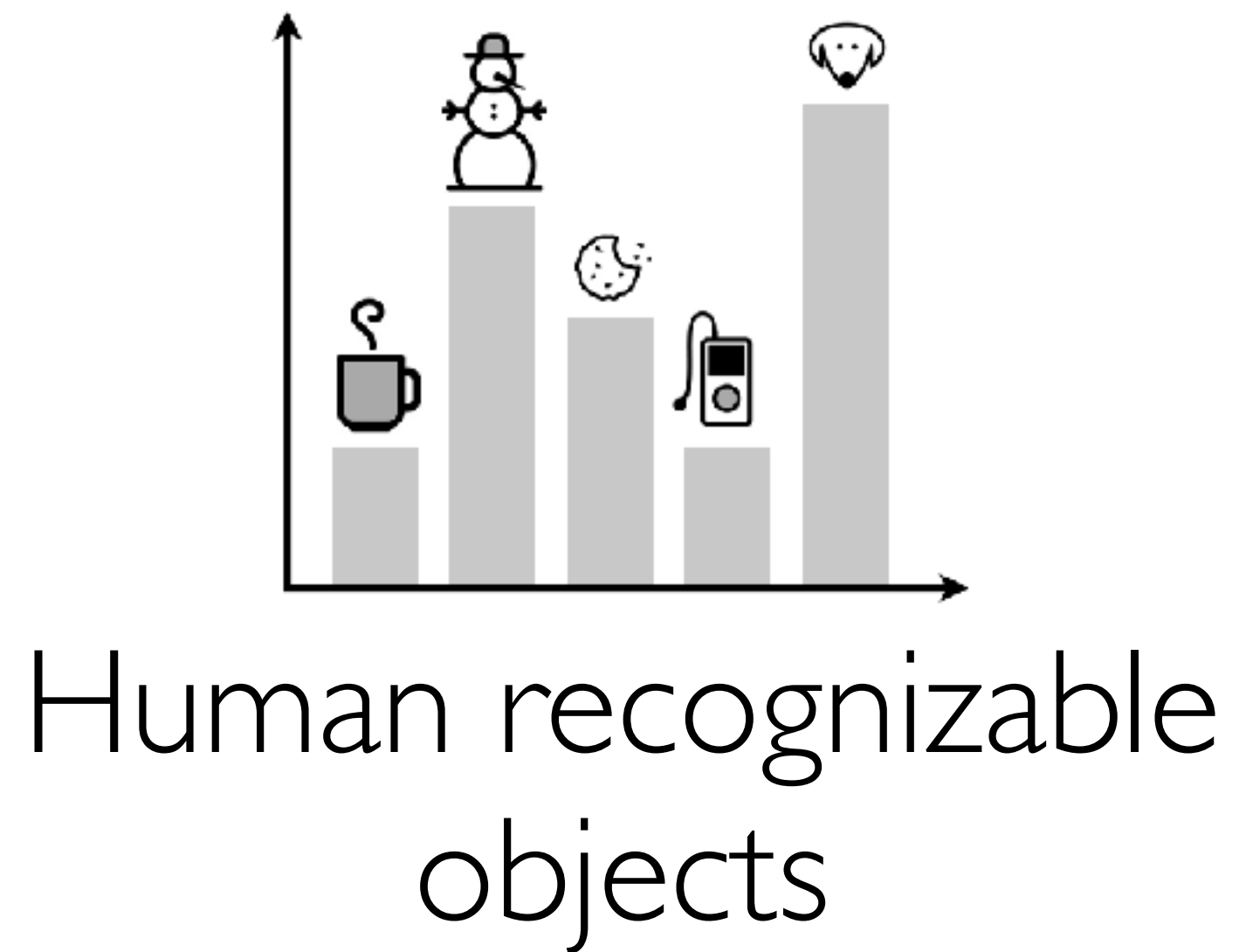
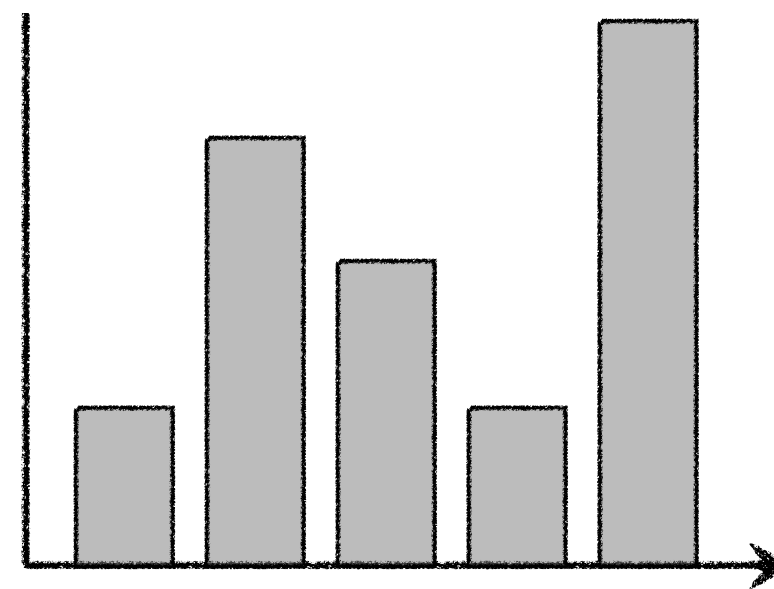
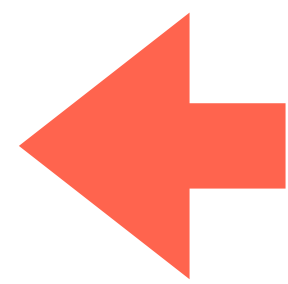
Examples of Redundant Encoding



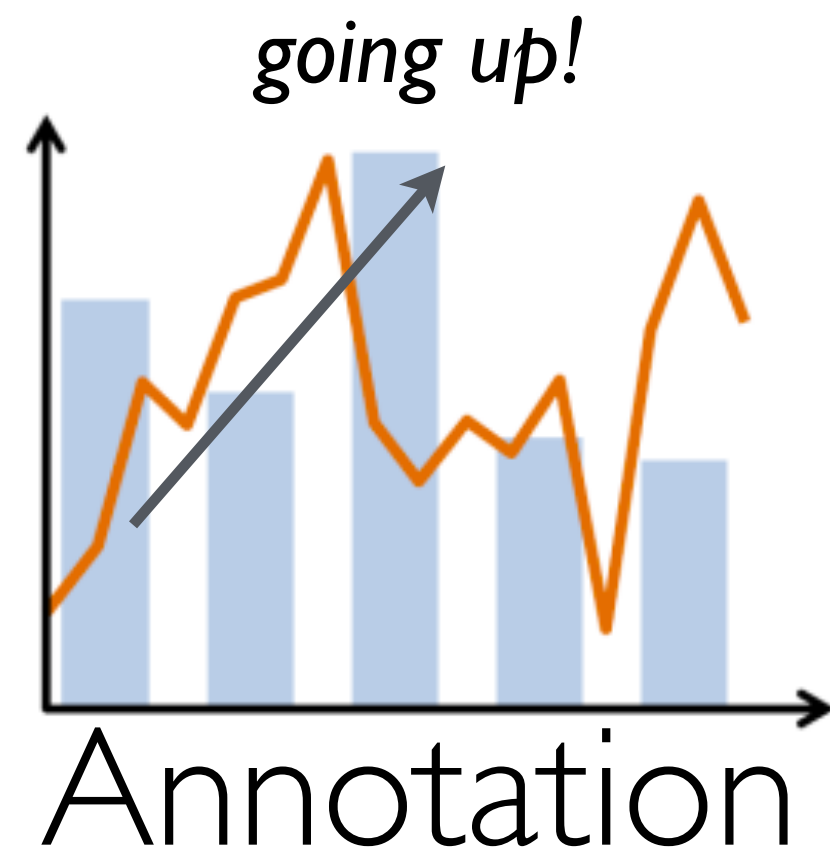
Examples of Redundant Encoding



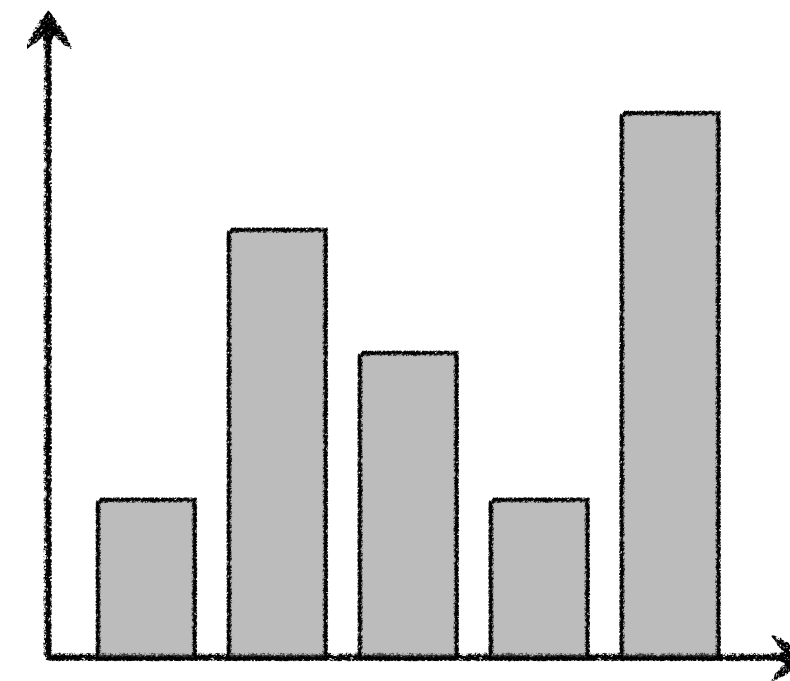
Message
Redundancy



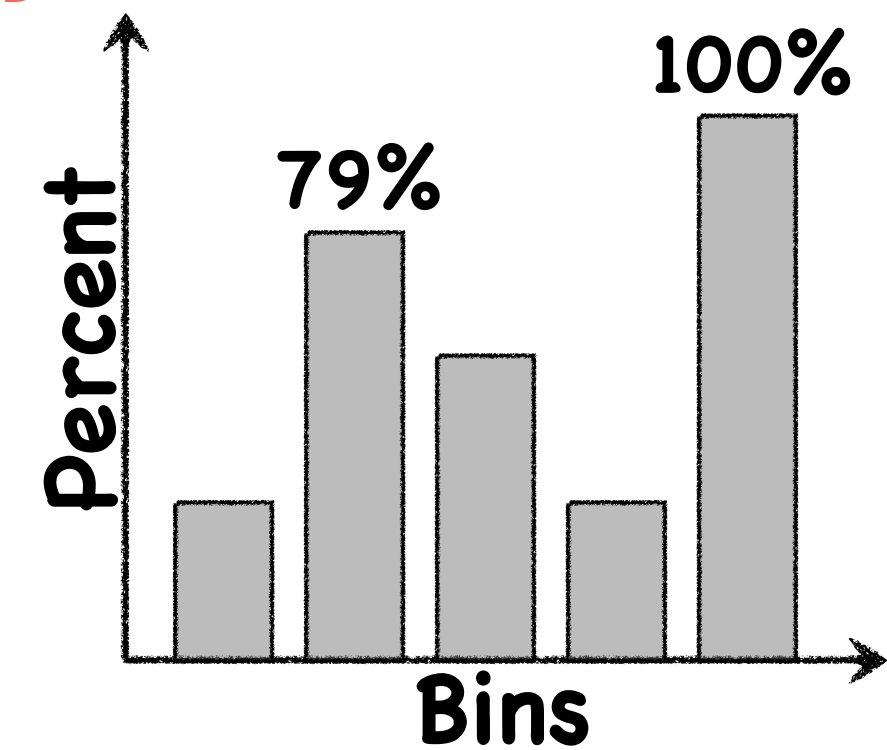
Examples of Redundant Encoding



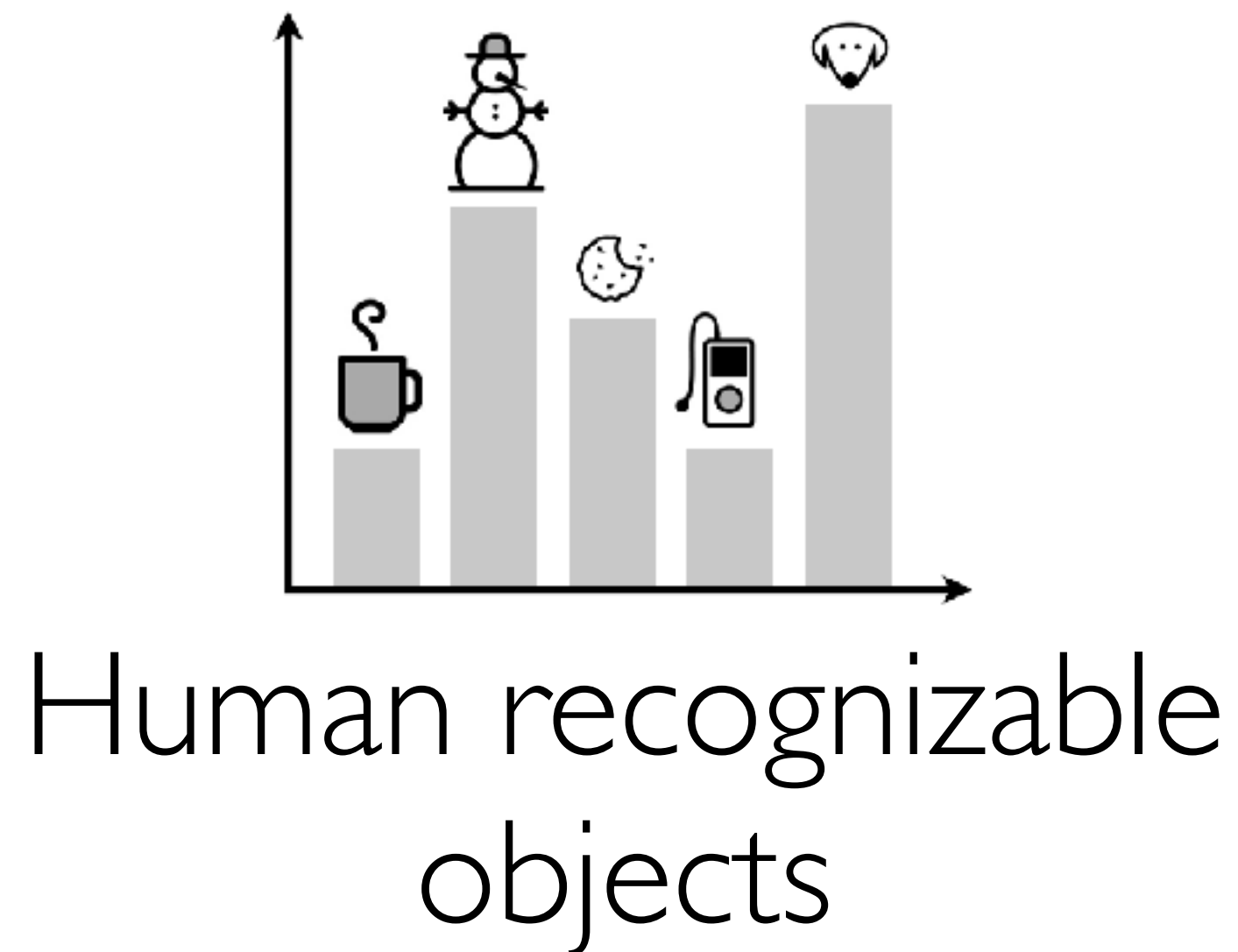
**Message
Redundancy**



**Data
Redundancy**

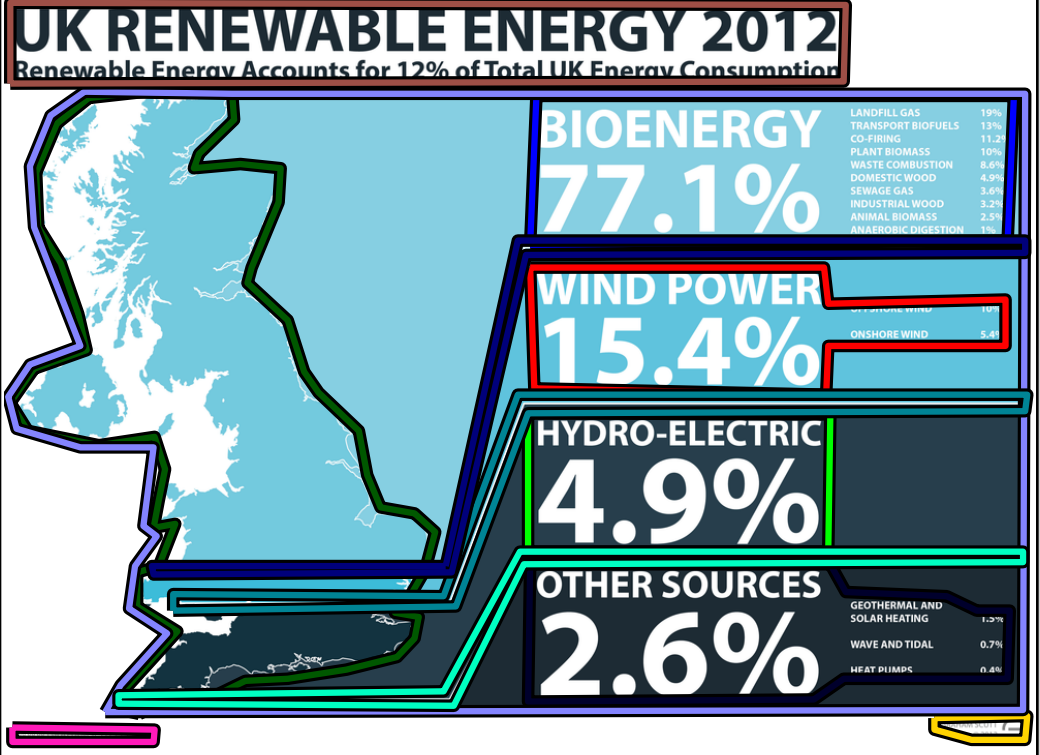


Data Labels



Experiment Procedure

LABELED VISUALIZATION DATABASE



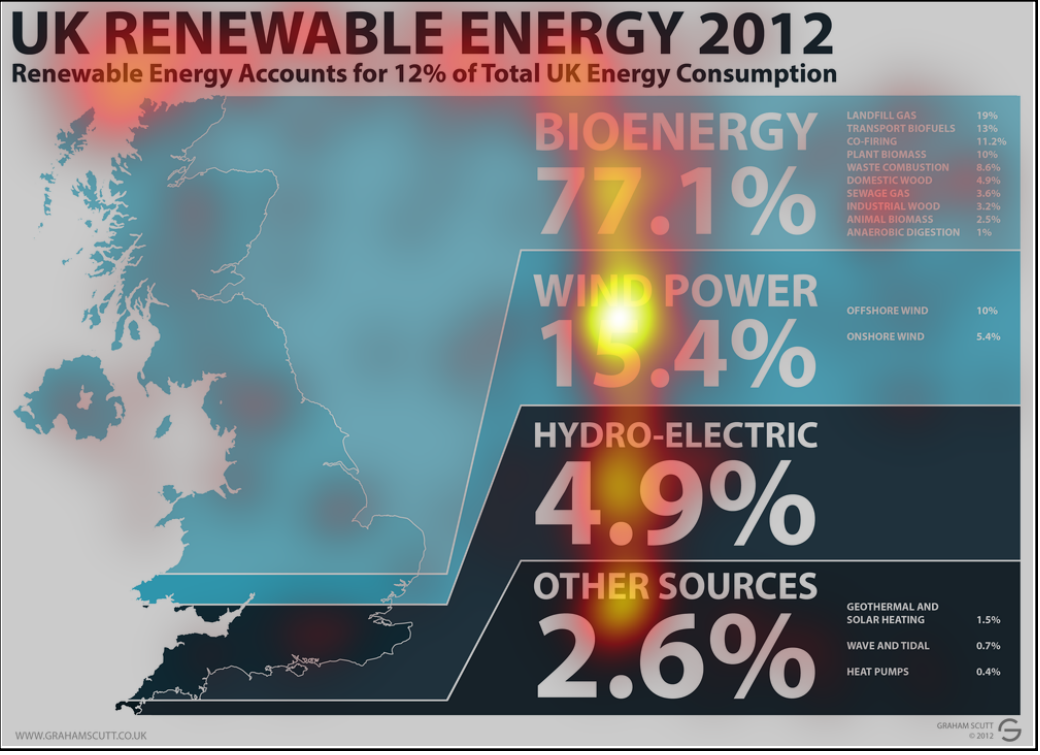
UK RENEWABLE ENERGY 2012
Renewable Energy Accounts for 12% of Total UK Energy Consumption

BIOENERGY	77.1%
WIND POWER	15.4%
HYDRO-ELECTRIC	4.9%
OTHER SOURCES	2.6%

393 visualizations

33 PARTICIPANTS

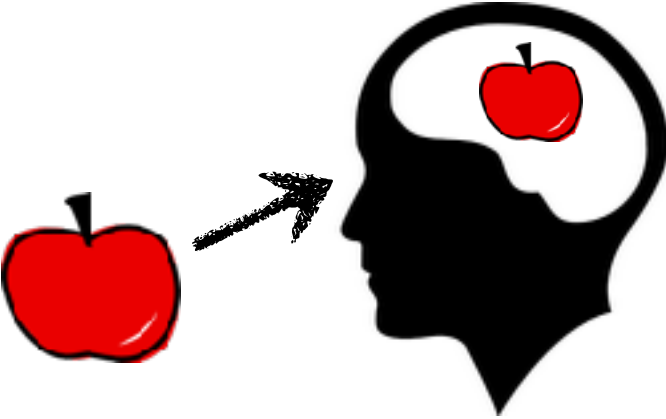
“ENCODING”
10 seconds / image



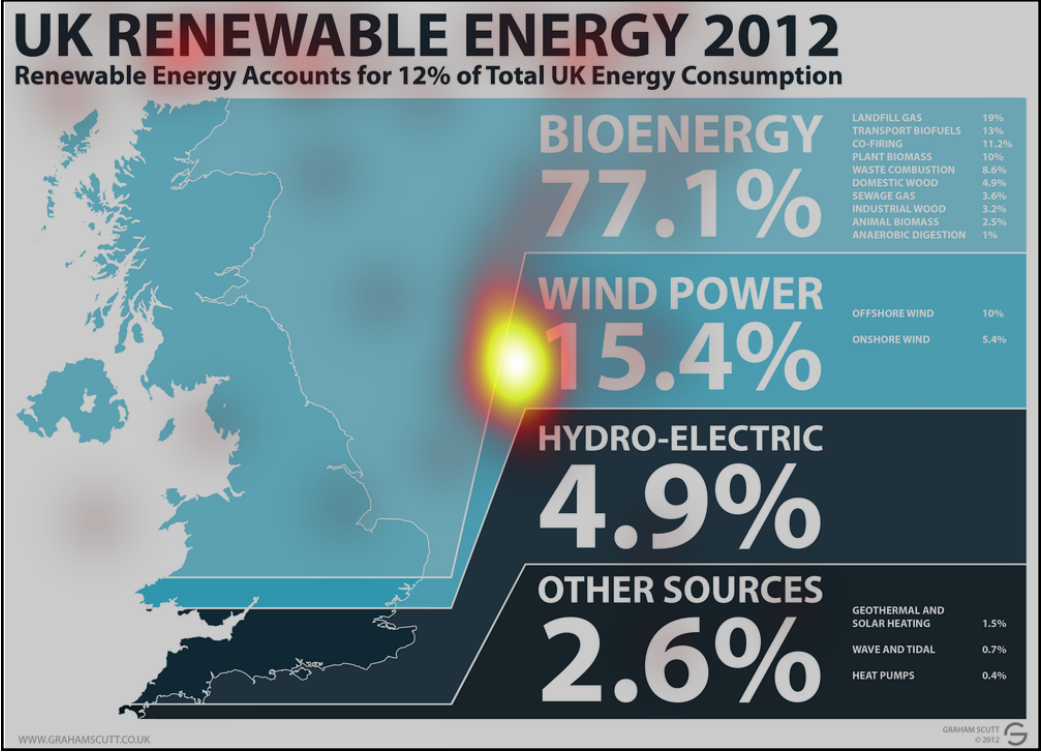
UK RENEWABLE ENERGY 2012
Renewable Energy Accounts for 12% of Total UK Energy Consumption

BIOENERGY	77.1%
WIND POWER	15.4%
HYDRO-ELECTRIC	4.9%
OTHER SOURCES	2.6%

EYE-TRACKING DATA



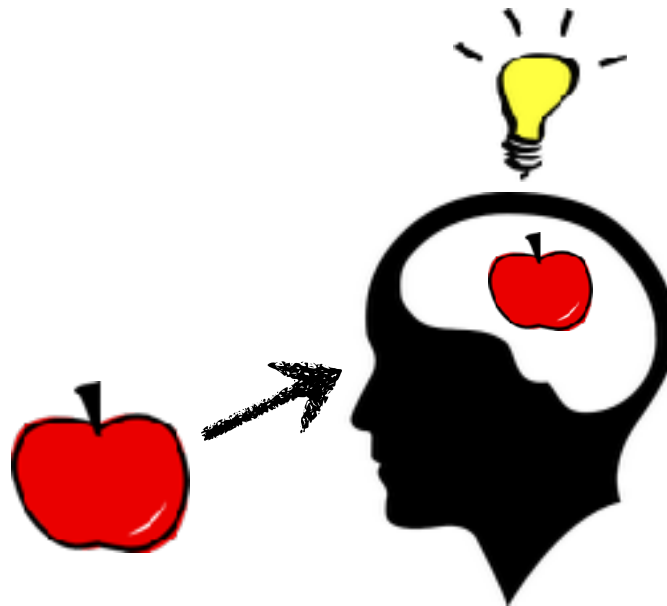
“RECOGNITION”
2 seconds / image



UK RENEWABLE ENERGY 2012
Renewable Energy Accounts for 12% of Total UK Energy Consumption

BIOENERGY	77.1%
WIND POWER	15.4%
HYDRO-ELECTRIC	4.9%
OTHER SOURCES	2.6%

EYE-TRACKING DATA



“RECALL”
20 minutes



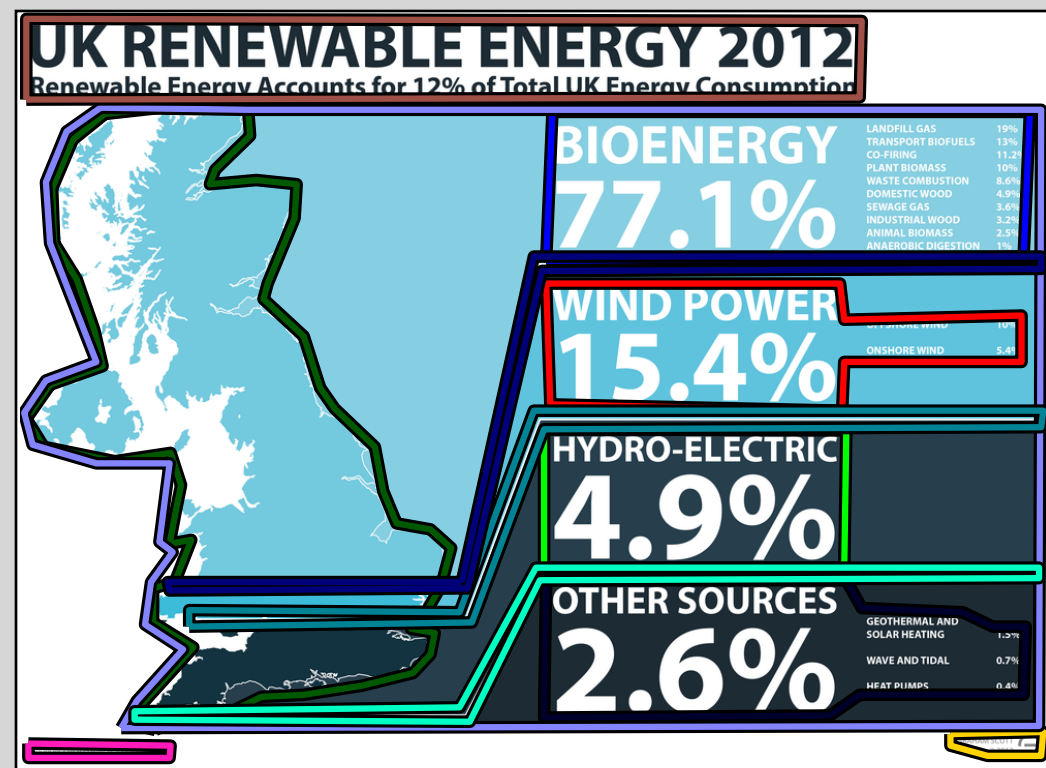
TEXT DESCRIPTIONS



~60 minutes

Experiment Procedure

LABELED VISUALIZATION DATABASE

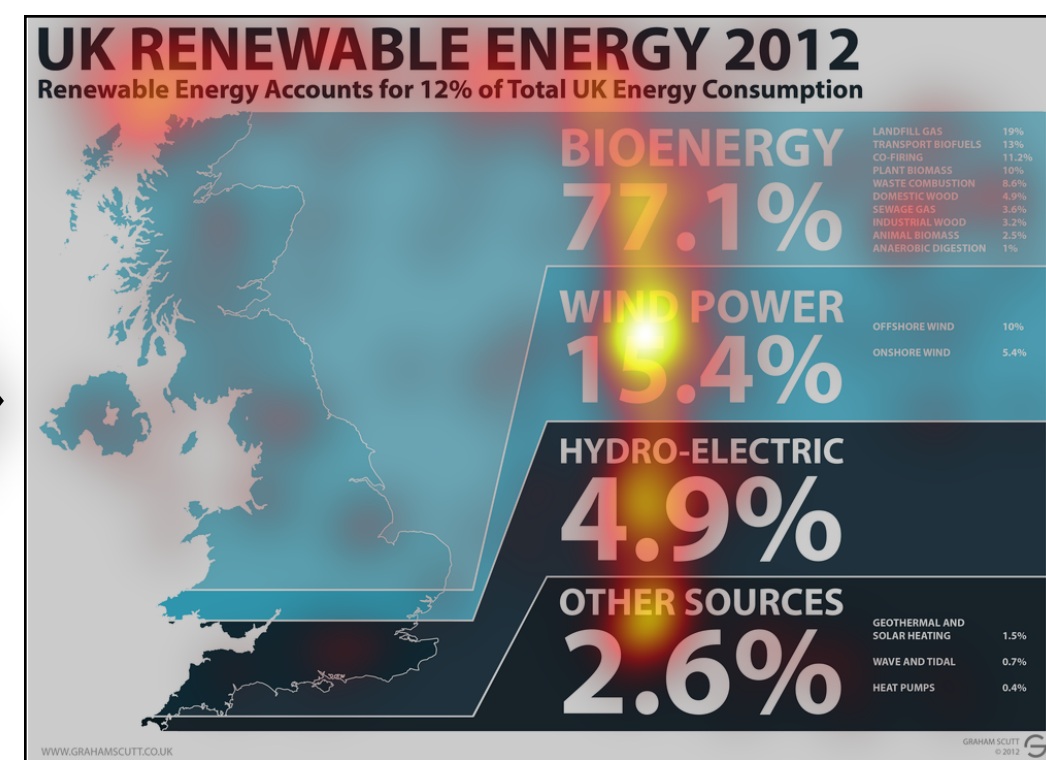


393 visualizations

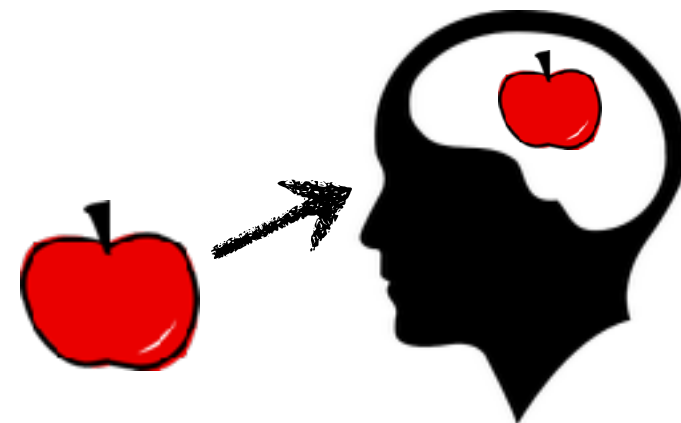
33 PARTICIPANTS

“ENCODING”

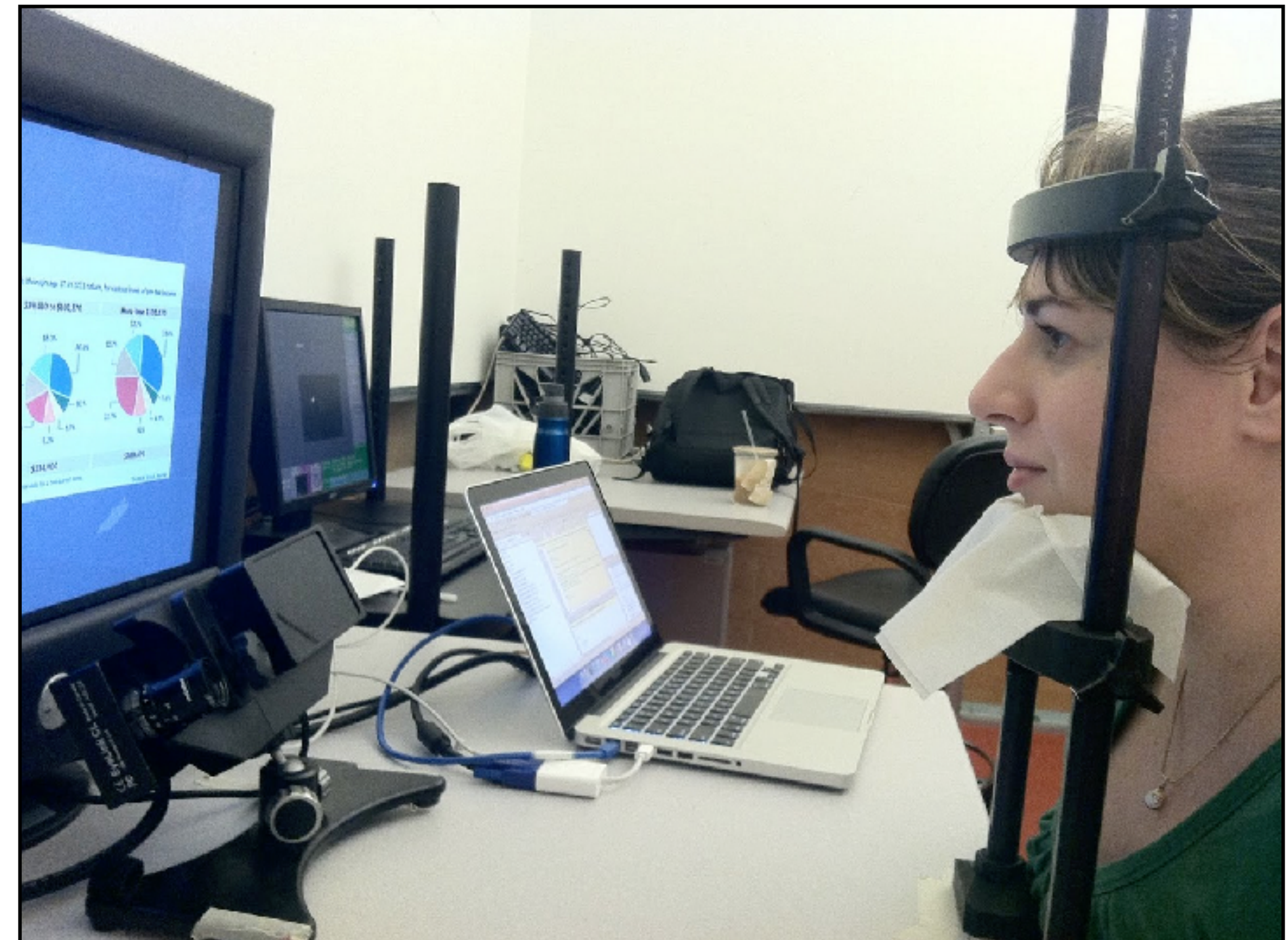
10 seconds / image



EYE-TRACKING DATA

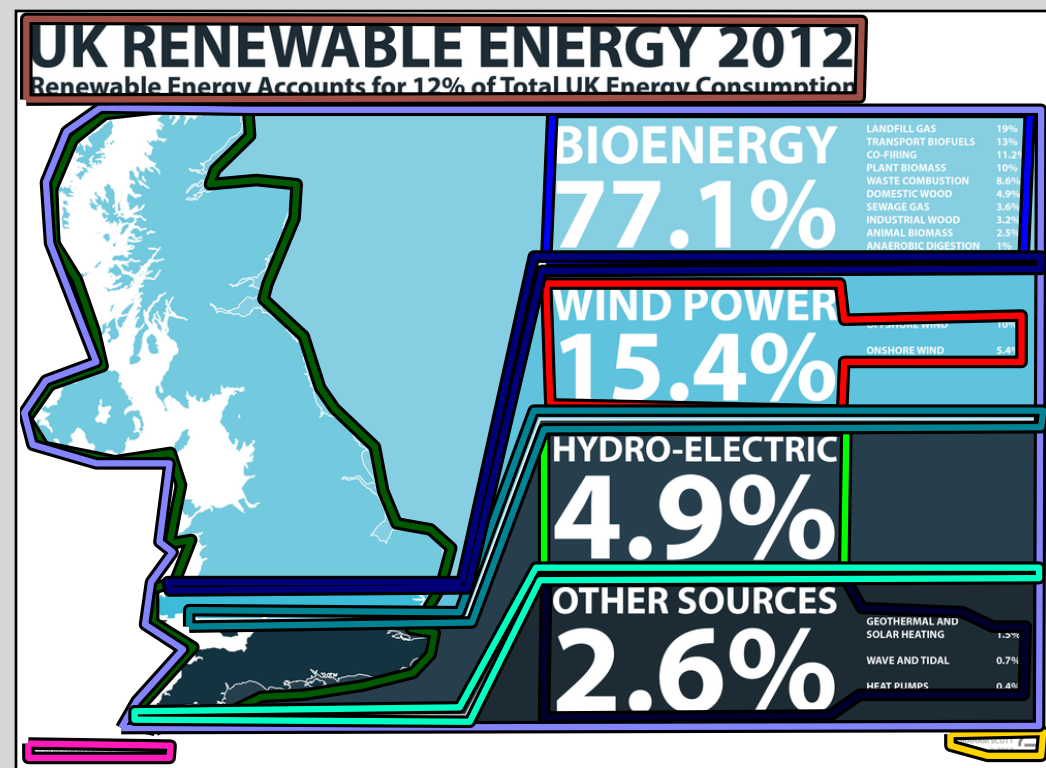


20 minutes



Experiment Procedure

LABELED VISUALIZATION DATABASE

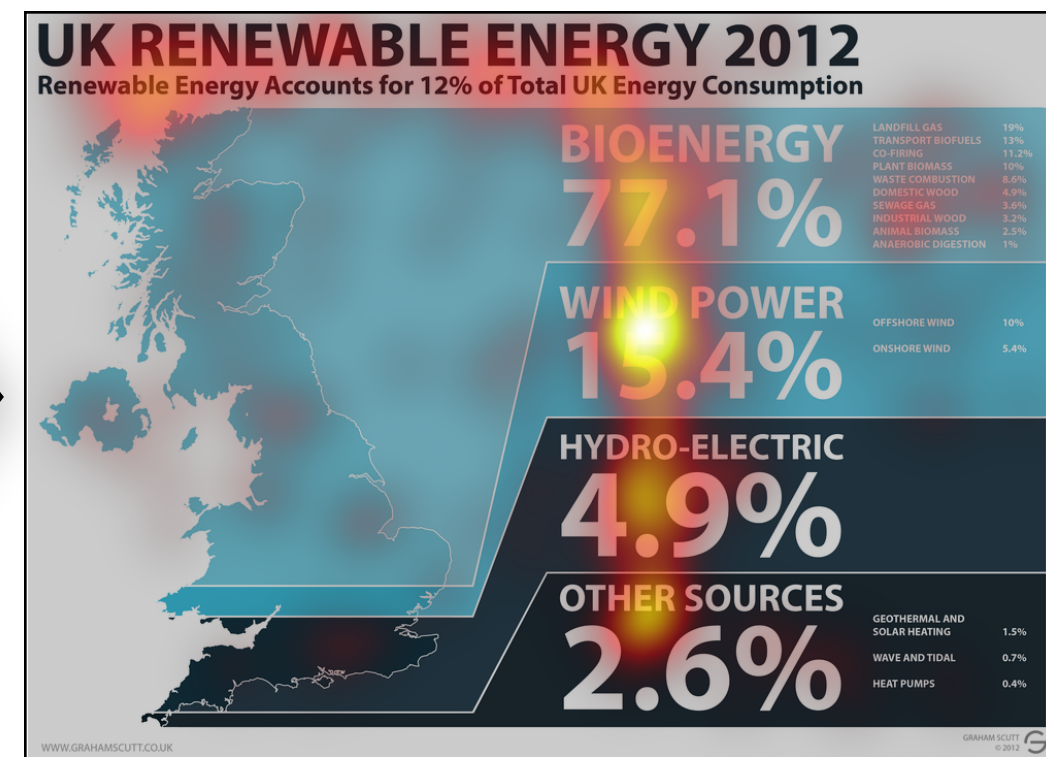


393 visualizations

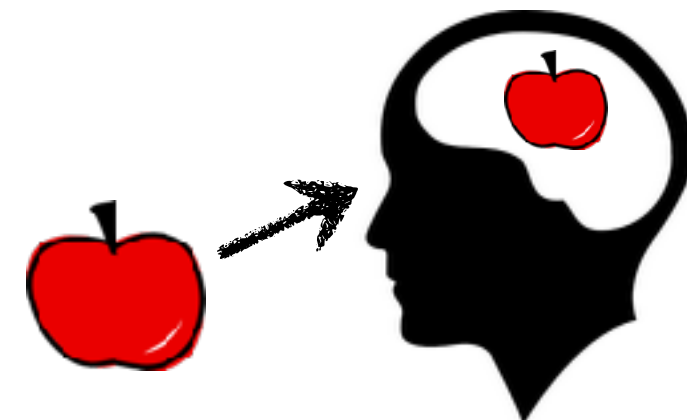
33 PARTICIPANTS

“ENCODING”

10 seconds / image

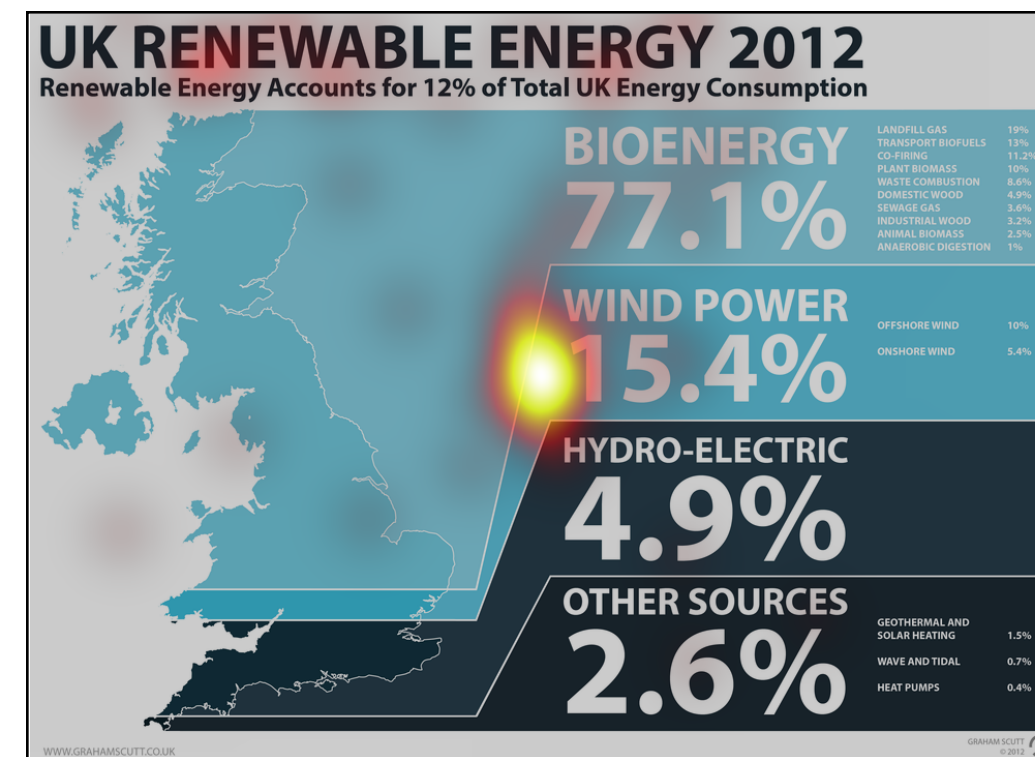


EYE-TRACKING DATA

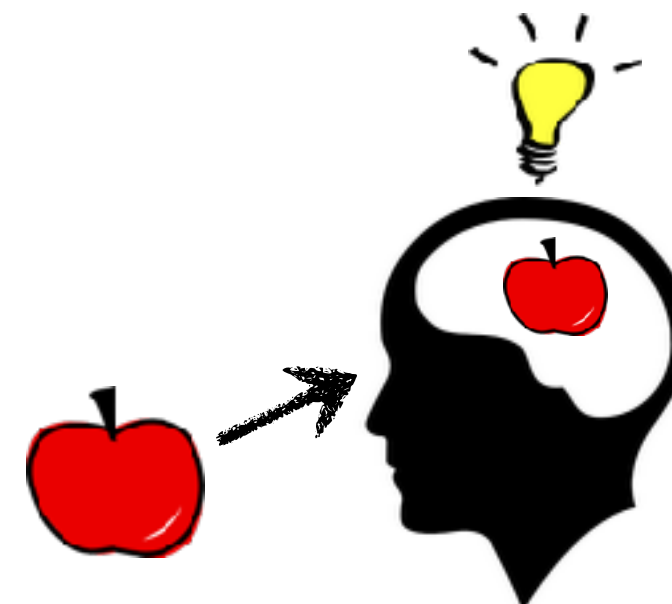


“RECOGNITION”

2 seconds / image



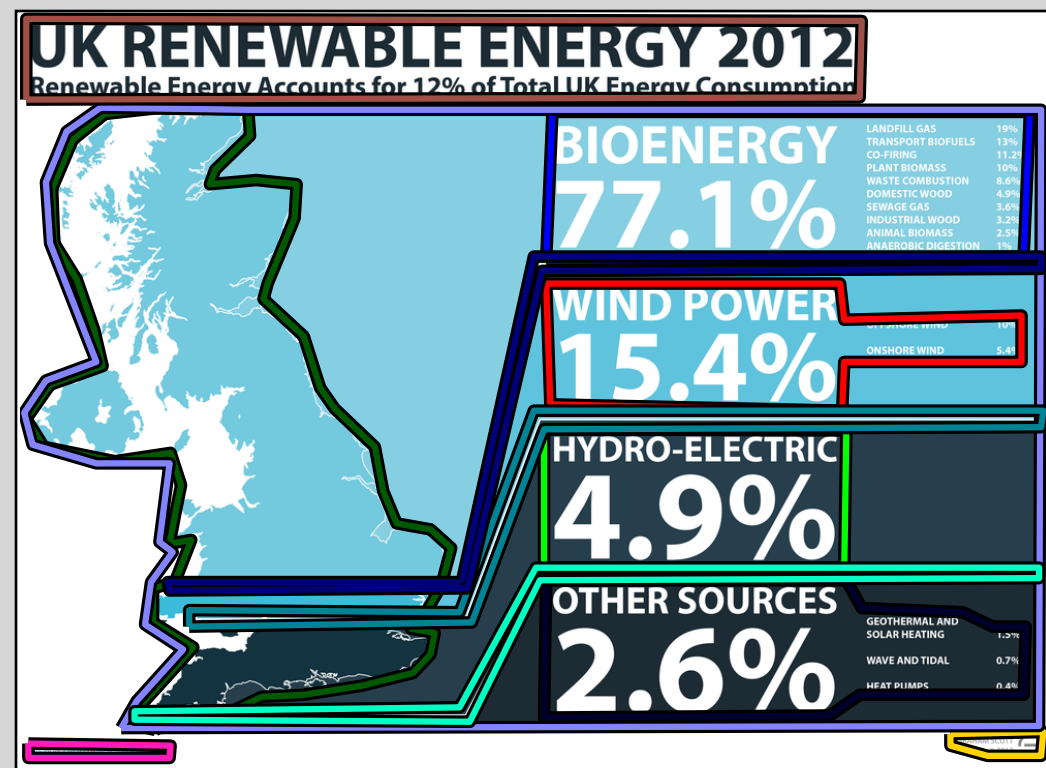
EYE-TRACKING DATA



10 minutes

Experiment Procedure

LABELED VISUALIZATION DATABASE

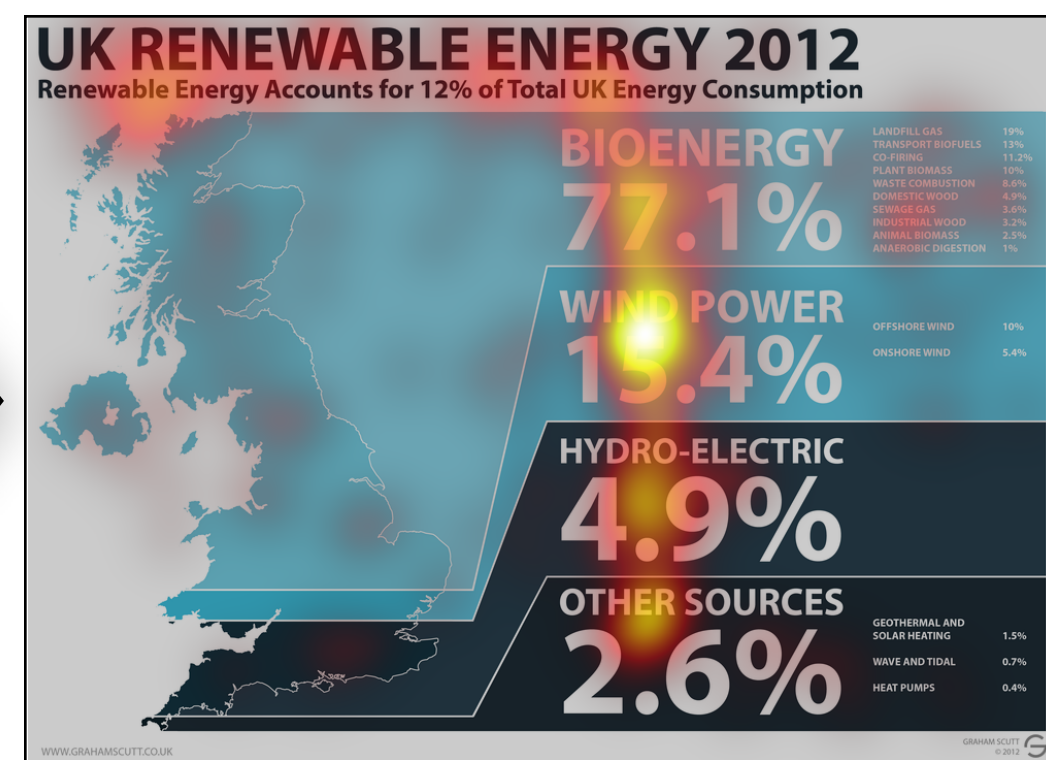


393 visualizations

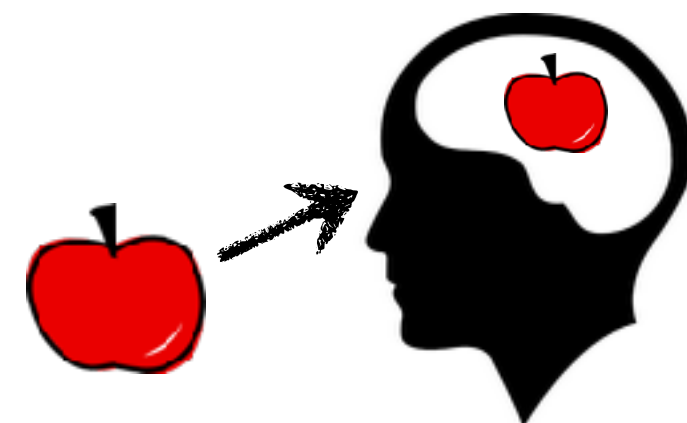
33 PARTICIPANTS

“ENCODING”

10 seconds / image

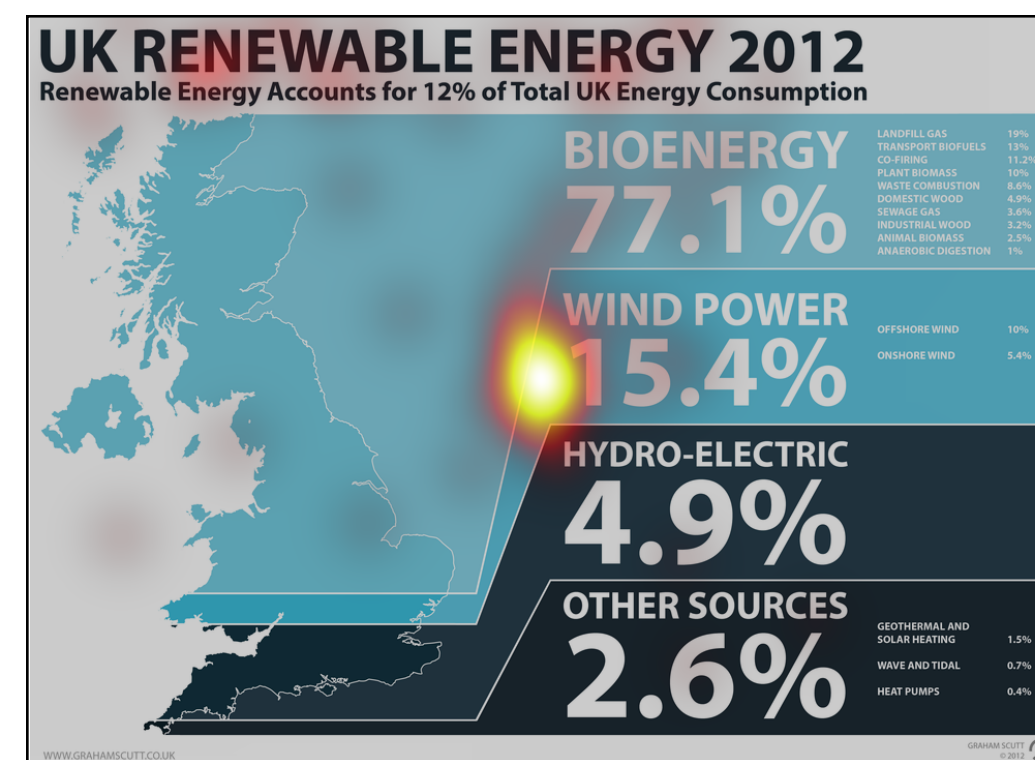


EYE-TRACKING DATA

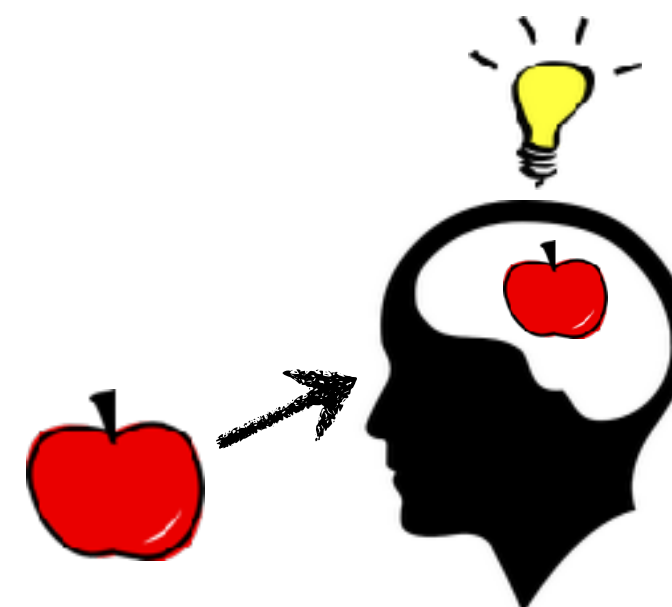


“RECOGNITION”

2 seconds / image



EYE-TRACKING DATA



“RECALL”

20 minutes



TEXT DESCRIPTIONS



20 minutes

Results

Encoding (10 sec)

Most Memorable

The iPhone evolution (based on iPhone 2G original data)

Model	Processor	Frequency	RAM	Chipset	Resolution	Screen Size	Camera	Connectivity	Release Date	Approx. Price
iPhone	ARM 11	117 MHz	128 MB	ARM	320x480 pixels	3.5-inch	2 MP fixed-focus	EDGE	January 9, 2007	~500
iPhone 2G	ARM 11	412 MHz	256 MB	ARM	320x480 pixels	3.5-inch	2 MP fixed-focus	HSDPA 3.6 Mbps	June 9, 2008	50,000
iPhone 3GS	ARM Cortex A8	600 MHz	512 MB	ARM	320x480 pixels	3.5-inch	3.2 MP auto-focus	HSDPA 7.2 Mbps	June 8, 2009	225,000+
iPhone 4	Apple A4	1 GHz	512 MB	ARM	640x960 pixels	3.5-inch Retina	5 MP auto-focus	HSPA & EV-DO Rev.A	June 7, 2010	500,000+
iPhone 5	Apple A5	1.3 GHz	1 GB	ARM	640x960 pixels	4-inch Retina	8 MP auto-focus	HSPA+ & EV-DO Rev.A	October 2011	500,000+

Least Memorable

Summary data from respondent countries indicate that median government expenditure on sanitation and drinking-water is one-fifth of that for health and one-sixth of that for education (Table 3.4).

TABLE 3.4 Government expenditure on health, education and WASH

Country	Government expenditure on health (% of GDP)	Government expenditure on education (% of GDP)	Government expenditure on sanitation and drinking water (% of GDP)
Algeria	1.9	2.4	0.4
Armenia	5.4	4.7	0.5
Burkina Faso	1.5	2.9	0.5
Cameroon	4.5	—	1.2
India	1.3	—	0.2
Iran (Islamic Republic of)	2.3	4.7	0.7
Kenya	2.1	6.9	1.5
Lesotho	2.4	—	3.5
Madagascar	2.7	3.2	0.4
Nepal	1.7	4.7	0.8
Pyramid	6.1	3.9	0.5
Thailand	3.7	4.1	0.8
Yemen	1.4	—	0.7
Yemen	1.1	3.2	0.2
Yemen	8.4	14.9	3.5
Yemen	2.3	5.5	0.7

Government expenditure (from taxes and transfers) on sanitation and drinking-water ranged from 0.37% to 3.5% of GDP (Figure 3.2).

FIGURE 3.2 Public spending (from funds obtained through domestic taxes and external transfers) on sanitation and drinking-water as a percentage of GDP (2010 data)

FIGURE 3.11 Average operation and minor maintenance costs for utilities covered by utilities?

FIGURE 3.12 Hobbled Hiring

Waterfront Estimated Eligible Homeowners

Not all 60+ day delinquent loans are currently eligible for HAMP. Other characteristics may preclude homeowner eligibility. Based on the estimates, of the 4.1 million homeowners who are currently 60+ days delinquent, about 700,000 homeowners are eligible for HAMP. As this represents a point-in-time snapshot of the delinquency population and estimated HAMP eligibility, we expect that as homeowners become seriously delinquent between now and the end of 2013, some of those homeowners will also be eligible for HAMP.

Other exclusions include: no longer owner-occupied, unemployed borrowers, investor's pending and servicing agreement precludes modification, manufactured housing loans with title-guarantee issues that exclude them from HAMP, and 1st and permanent modifications disqualified from HAMP. Note: Chart refers only to loans eligible for the Tier 1 first lien modification program.

Sources: Fannie Mae, monthly survey of participating servicers for July 31, 2012; Total 60+ day delinquency figure derived from 2nd Quarter 2012 MBA National Delinquency Survey. Excluded loans are as reported by large servicers by survey who have signed a servicer participation agreement for HAMP.

Encoding (10 sec)

Most Memorable

The iPhone evolution

Model	Processor	Display	Connectivity	Camera	Announced	Units Sold
iPhone	ARM 11	3.5-inch	EDGE	2 MP fixed-focus	January 9, 2007	~500
iPhone 3G	ARM 11	3.5-inch	HSDPA	2 MP fixed-focus	June 9, 2008	50,000
iPhone 3GS	ARM Cortex A8	3.5-inch	HSDPA	3.2 MP auto-focus	June 8, 2009	225,000+
iPhone 4	Apple A4	3.5-inch Retina	HSPA & EV-DO Rev.A	5 MP auto-focus	June 7, 2010	500,000+
iPhone 5	Apple A5	4-inch Retina	HSPA+ & EV-DO Rev.A	8 MP auto-focus	October, 2011	500,000+

Least Memorable

Summary data from respondent countries indicate that median government expenditure on sanitation and drinking-water is one-third of that for health and one-sixth of that for education (Table 3.4).

Country	Government expenditure on health (% of GDP)	Government expenditure on education (% of GDP)	Government expenditure on sanitation and drinking water (% of GDP)
Burkina Faso	1.1	2.4	0.4
Comoros	5.4	4.7	0.5
Egypt	1.9	2.9	0.8
Guatemala	4.6	—	1.2
India	1.3	—	0.2
Iran (Islamic Republic of)	2.3	4.7	0.7
Kenya	2.1	6.9	1.1
Lesotho	8.4	—	3.5
Madagascar	2.7	3.2	0.4
Nepal	1.7	4.7	0.8
Paraguay	6.1	3.9	0.5
Thailand	3.7	4.1	0.8
Yemen	1.4	—	0.7
Minimum	1.1	3.2	0.2
Maximum	8.4	14.9	3.5
Median	2.3	5.5	0.7

FIGURE 3.2 Public spending (from funds obtained through domestic taxes and external transfers) on sanitation and drinking-water as a percentage of GDP (2010 data).

FIGURE 3.11 Are operation and minor maintenance costs for utilities covered by user fees?

FIGURE 3.12 Monthly change in jobs

Waterfall of Estimated Eligible Homeowners For the First Lien Modification Program

FIGURE 3.13 Hobbled Hiring

Recognition

The iPhone evolution

Model	Processor	Display	Connectivity	Camera	Announced	Units Sold
iPhone	ARM 11	3.5-inch	EDGE	2 MP fixed-focus	January 9, 2007	~500
iPhone 3G	ARM 11	3.5-inch	HSDPA	2 MP fixed-focus	June 9, 2008	50,000
iPhone 3GS	ARM Cortex A8	3.5-inch	HSDPA	3.2 MP auto-focus	June 8, 2009	225,000+
iPhone 4	Apple A4	3.5-inch Retina	HSPA & EV-DO Rev.A	5 MP auto-focus	June 7, 2010	500,000+
iPhone 5	Apple A5	4-inch Retina	HSPA+ & EV-DO Rev.A	8 MP auto-focus	October, 2011	500,000+

Summary data from respondent countries indicate that median government expenditure on sanitation and drinking-water is one-third of that for health and one-sixth of that for education (Table 3.4).

Country	Government expenditure on health (% of GDP)	Government expenditure on education (% of GDP)	Government expenditure on sanitation and drinking water (% of GDP)
Burkina Faso	1.1	2.4	0.4
Comoros	5.4	4.7	0.5
Egypt	1.9	2.9	0.8
Guatemala	4.6	—	1.2
India	1.3	—	0.2
Iran (Islamic Republic of)	2.3	4.7	0.7
Kenya	2.1	6.9	1.1
Lesotho	8.4	—	3.5
Madagascar	2.7	3.2	0.4
Nepal	1.7	4.7	0.8
Paraguay	6.1	3.9	0.5
Thailand	3.7	4.1	0.8
Yemen	1.4	—	0.7
Minimum	1.1	3.2	0.2
Maximum	8.4	14.9	3.5
Median	2.3	5.5	0.7

FIGURE 3.2 Public spending (from funds obtained through domestic taxes and external transfers) on sanitation and drinking-water as a percentage of GDP (2010 data).

FIGURE 3.11 Are operation and minor maintenance costs for utilities covered by user fees?

FIGURE 3.12 Monthly change in jobs

Waterfall of Estimated Eligible Homeowners For the First Lien Modification Program

FIGURE 3.13 Hobbled Hiring

Encoding (10 sec)

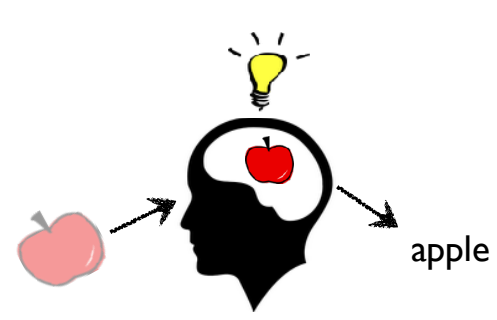
Most Memorable

Least Memorable

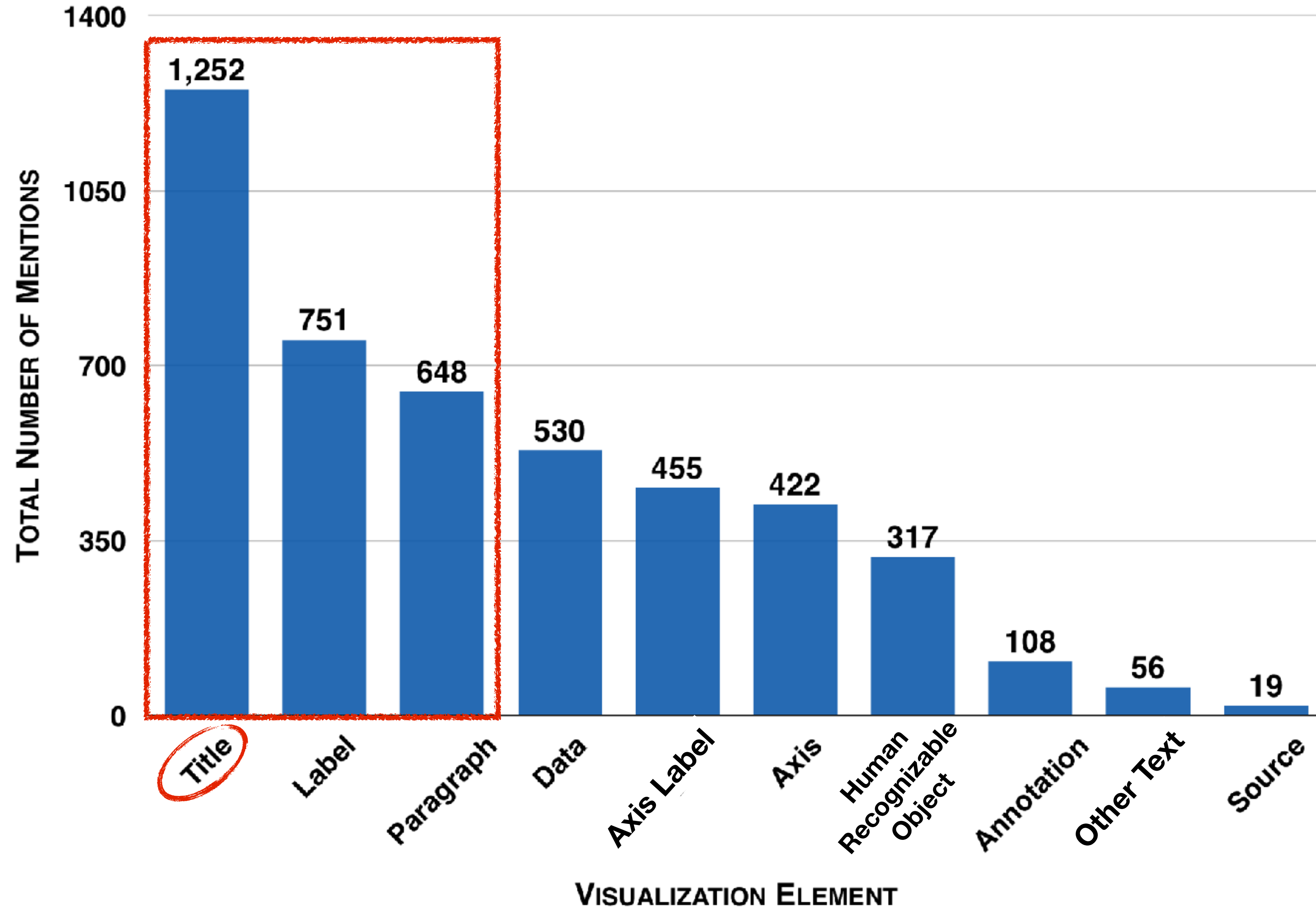
Visual Associations

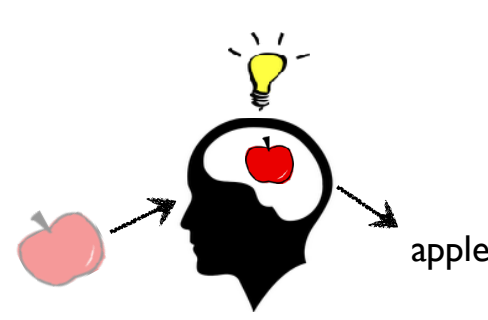
Semantic Associations

Recognition

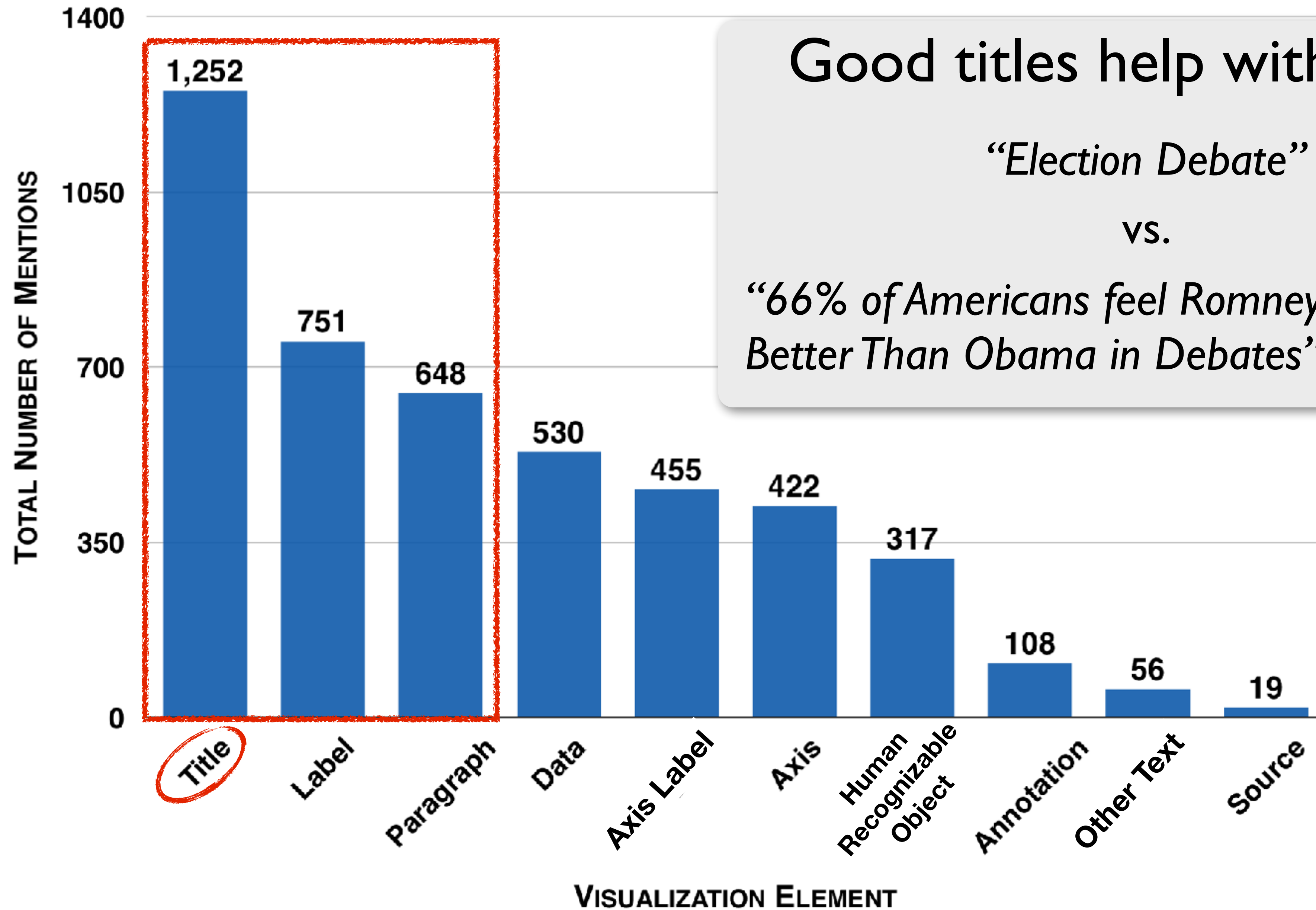


RECALL





RECALL

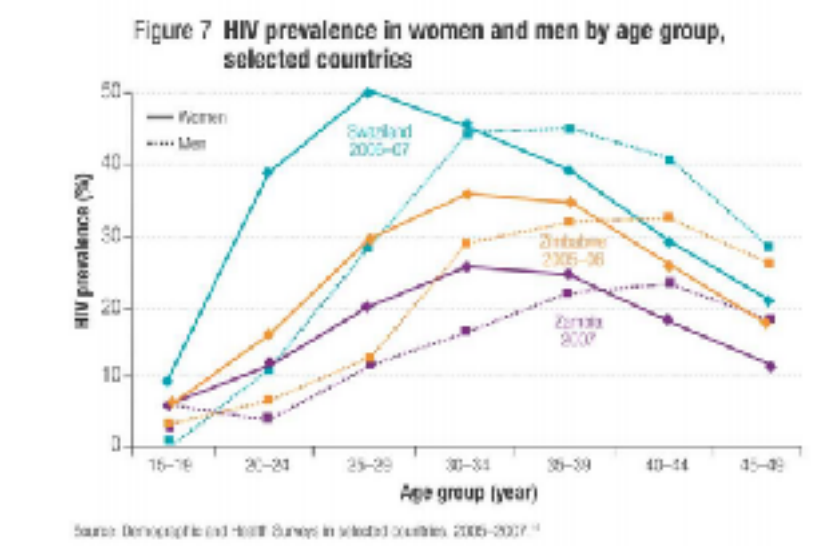
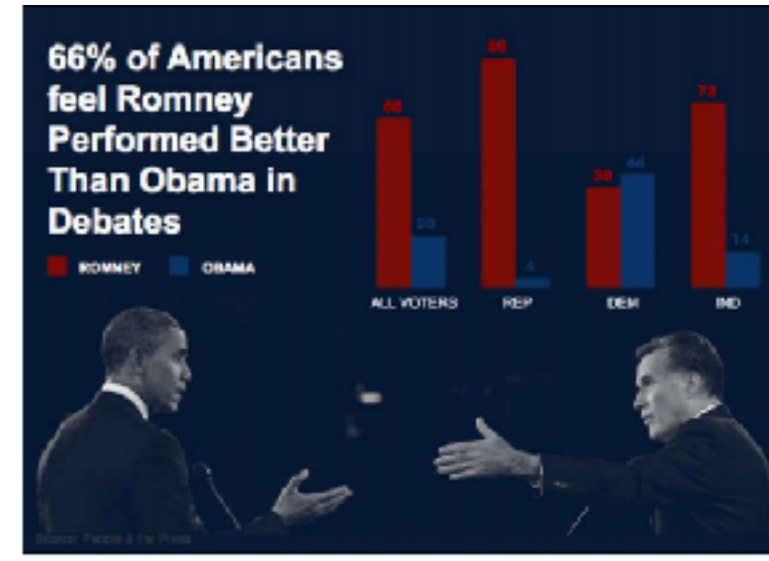
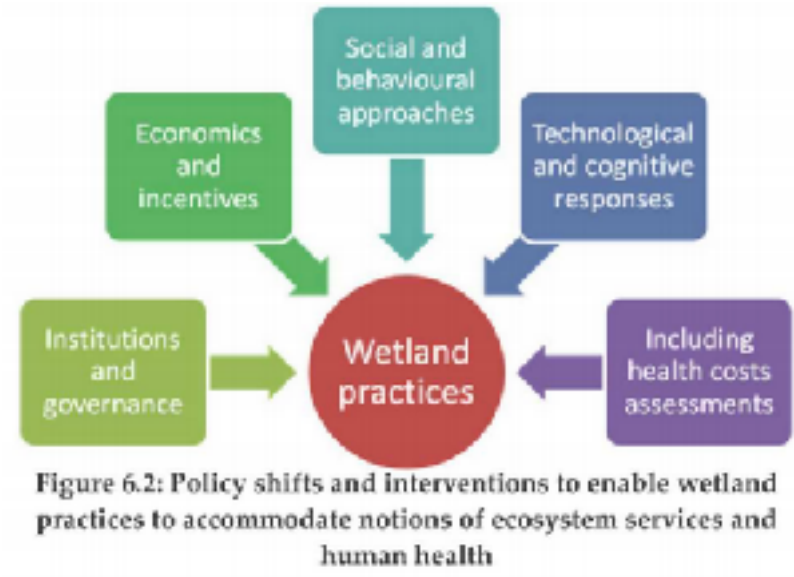
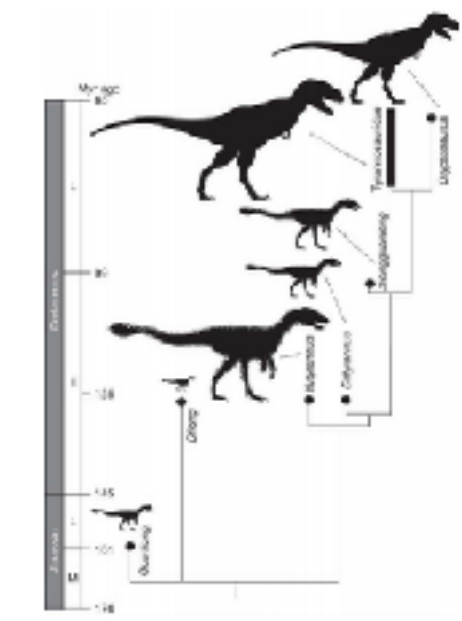
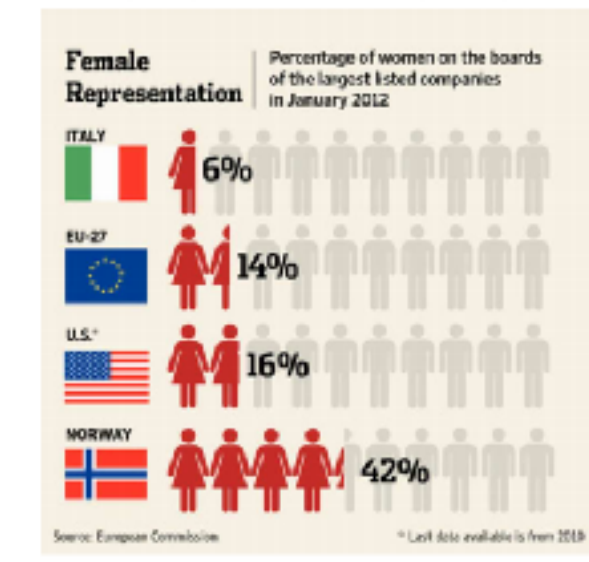
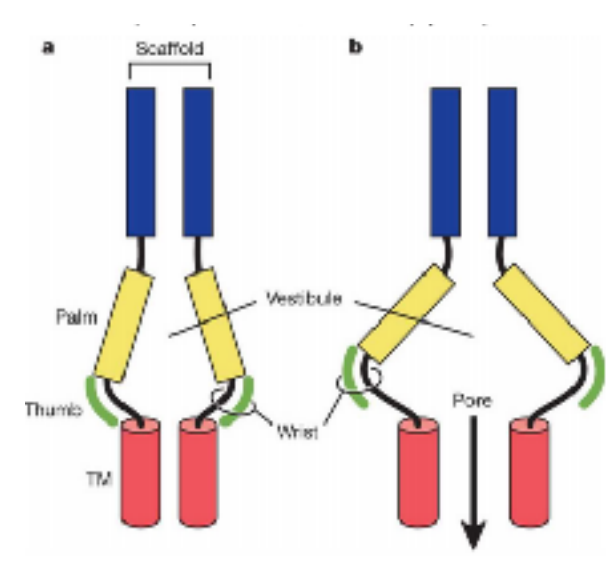


Good titles help with recall!

“Election Debate”

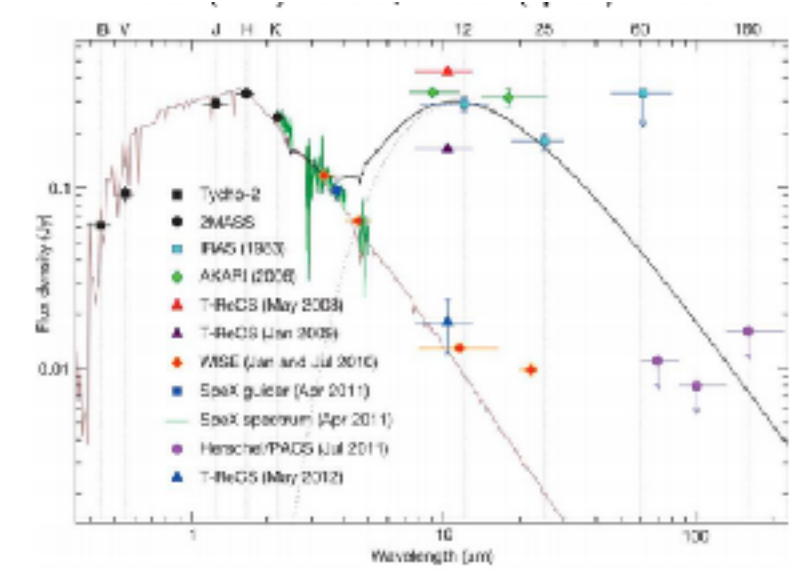
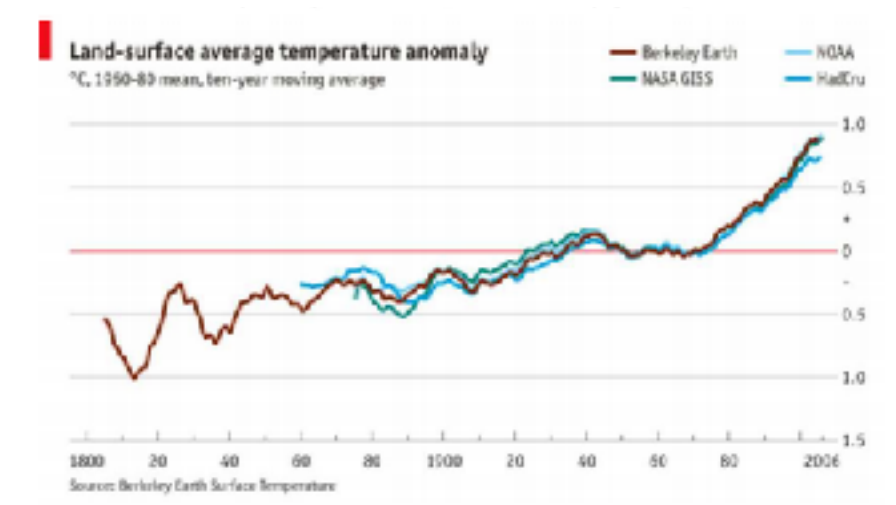
vs.

“66% of Americans feel Romney Performed Better Than Obama in Debates”



LOW QUALITY DESCRIPTION

HIGH QUALITY DESCRIPTION



News Corp. Results, by Segment
for the nine months ended March 31

Revenue	Operating profit/loss
\$566 million	Cable network programming \$250 million
6.22	Publishing \$458 million
5.56	Filmed entertainment \$1.01 billion
3.65	Television \$493 million
2.79	Direct broadcast satellite television \$365 million
0.45	Other -\$437 million

Source: The company
The Wall Street Journal

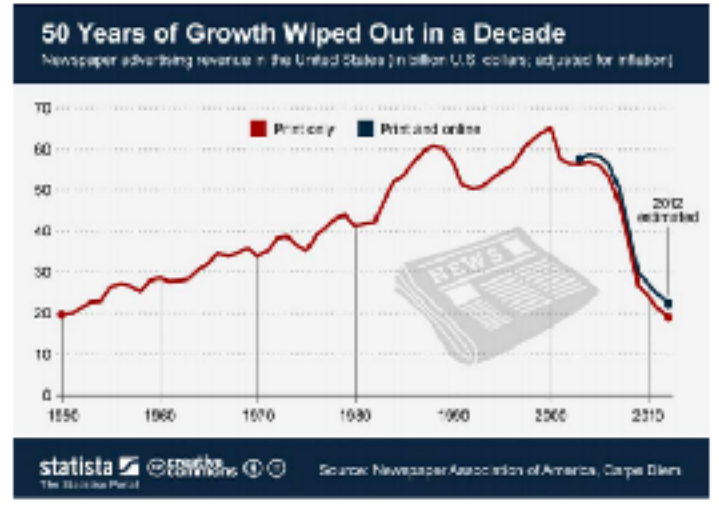
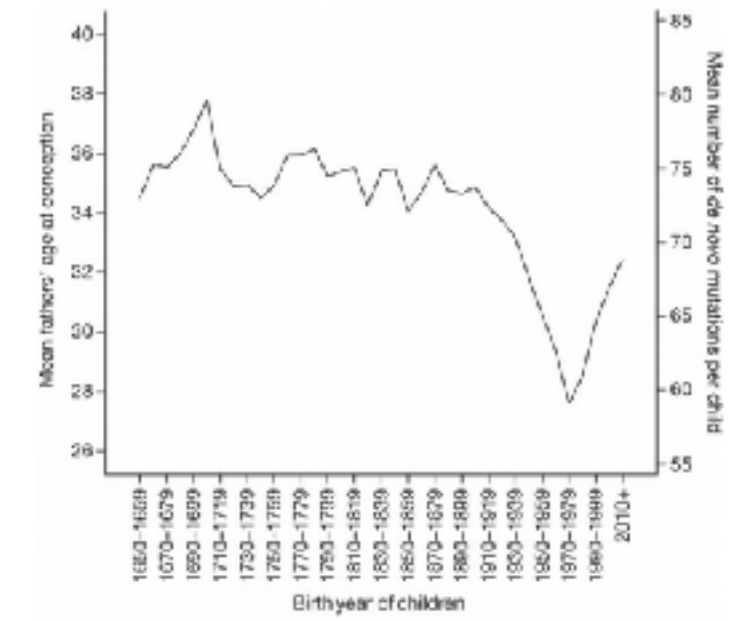
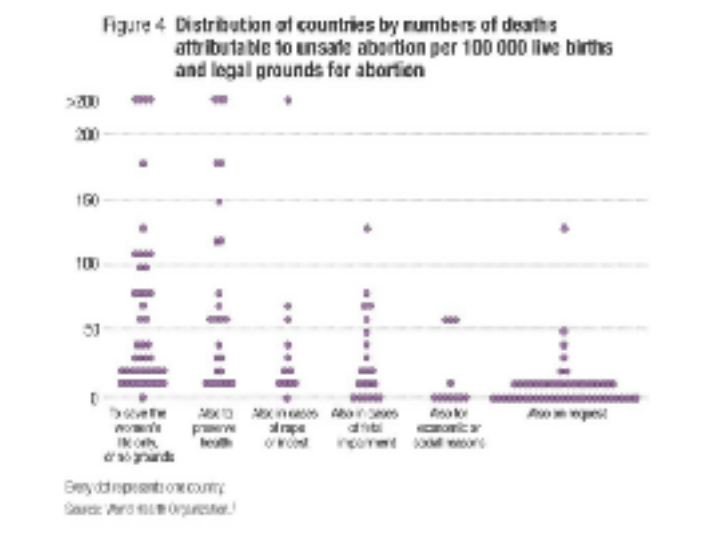
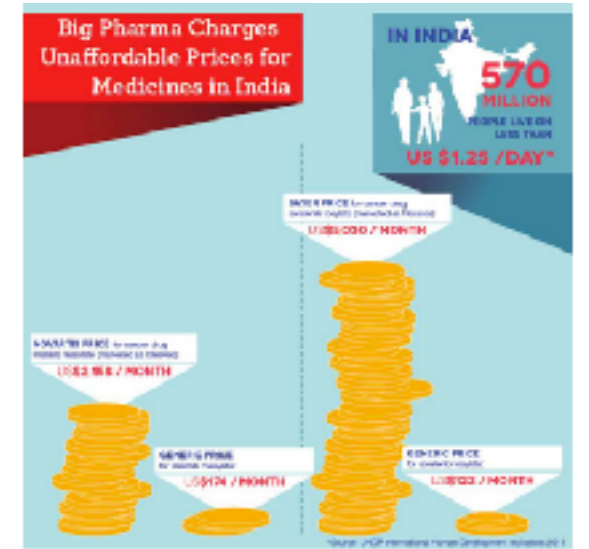


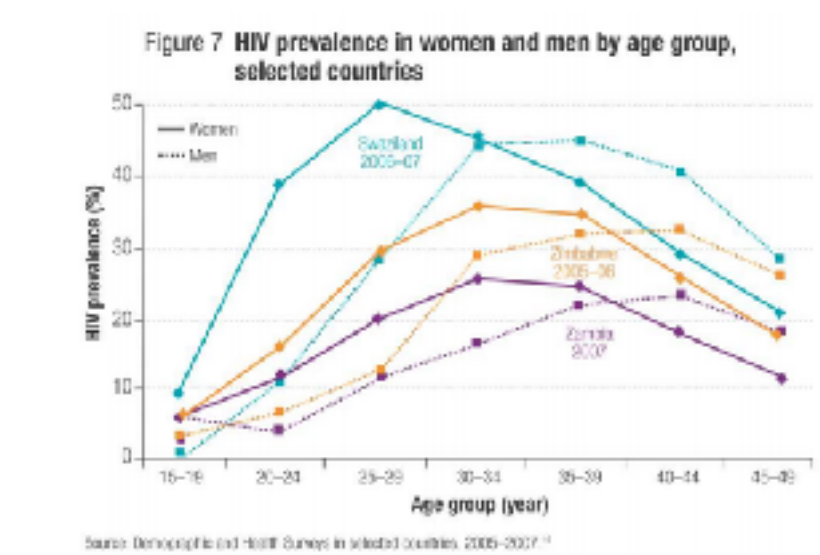
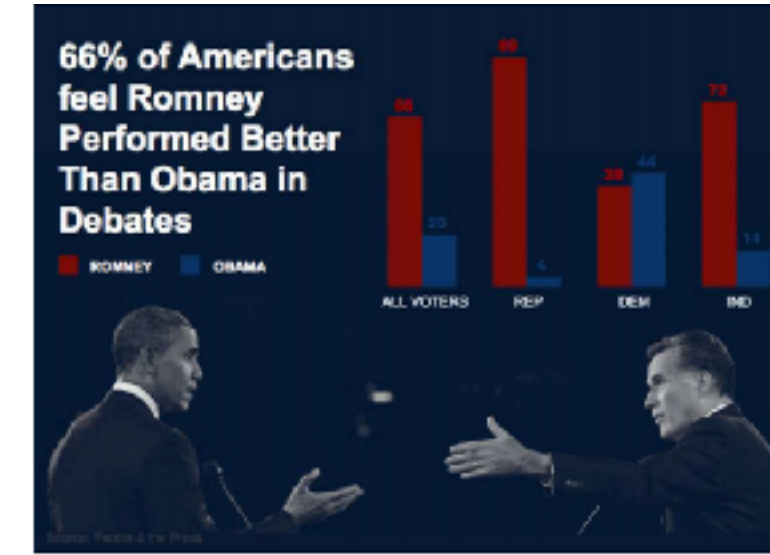
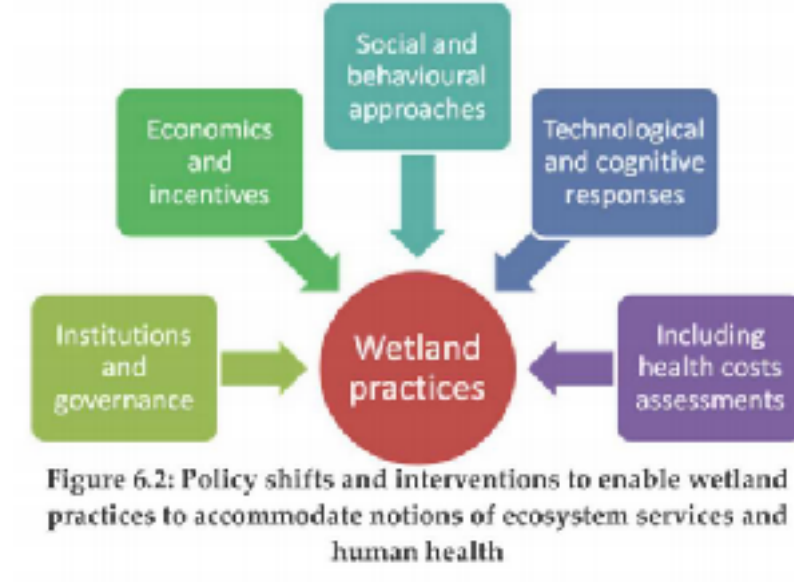
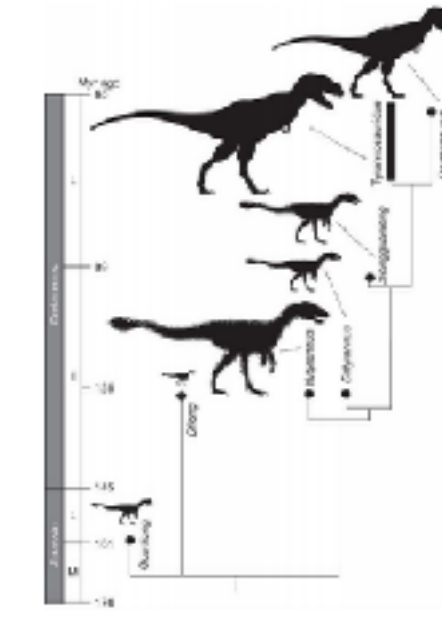
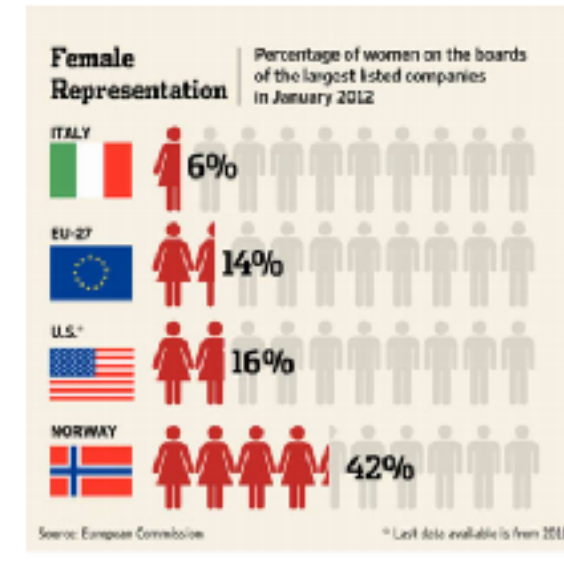
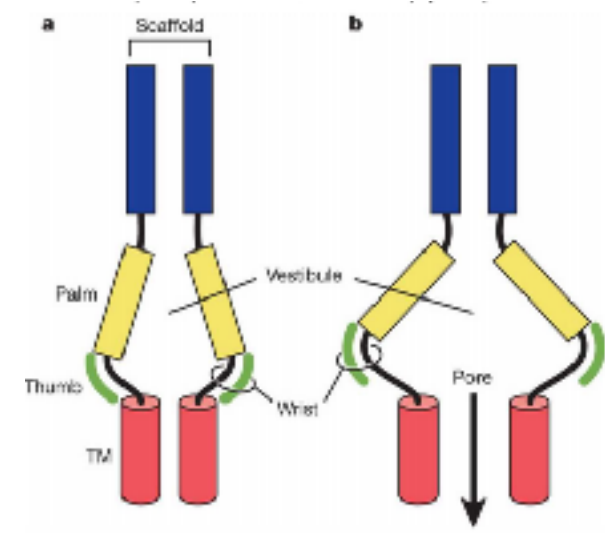
Table 3: Percentage of medicines that are for use in the developing world

Country	2008	2009
India	21	22
China	12	12
USA	12	12
Japan	12	12
Germany	12	12
France	12	12
UK	12	12
Italy	12	12
Spain	12	12
Sweden	12	12
Denmark	12	12
Norway	12	12
Finland	12	12
Poland	12	12
Czech Republic	12	12
Slovakia	12	12
Hungary	12	12
Slovenia	12	12
Lithuania	12	12
Latvia	12	12
Estonia	12	12
Other	12	12

Source: WHO, WHOCC, WHOCC, WHOCC

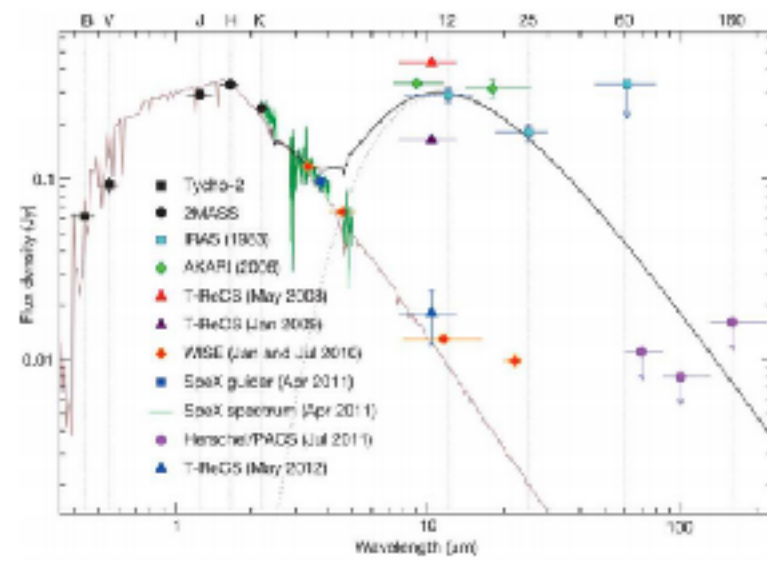
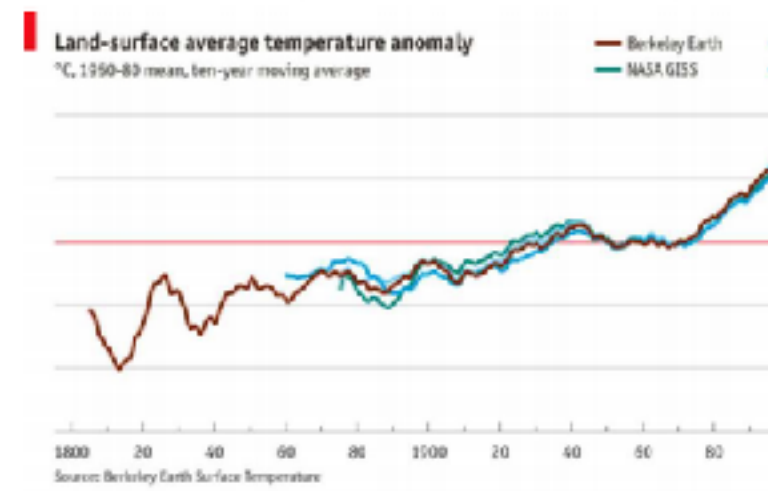


MEMORABLE



LOW QUALITY DESCRIPTION

HIGH QUALITY DESCRIPTION



News Corp. Results, by Segment
for the nine months ended March 31

Revenue	Operating profit/loss
\$566 million	\$250 million
6.22	\$458 million
5.56	\$1.01 billion
3.65	\$493 million
2.79	\$365 million
0.45	-\$437 million

Source: The company
The Wall Street Journal

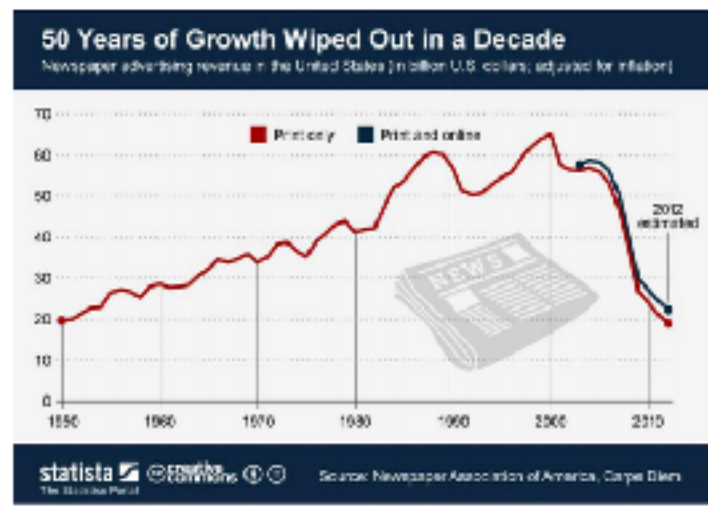
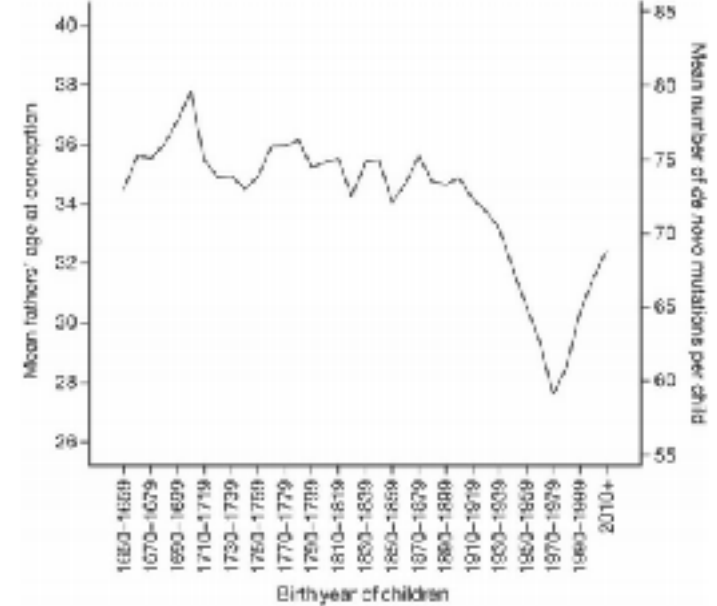
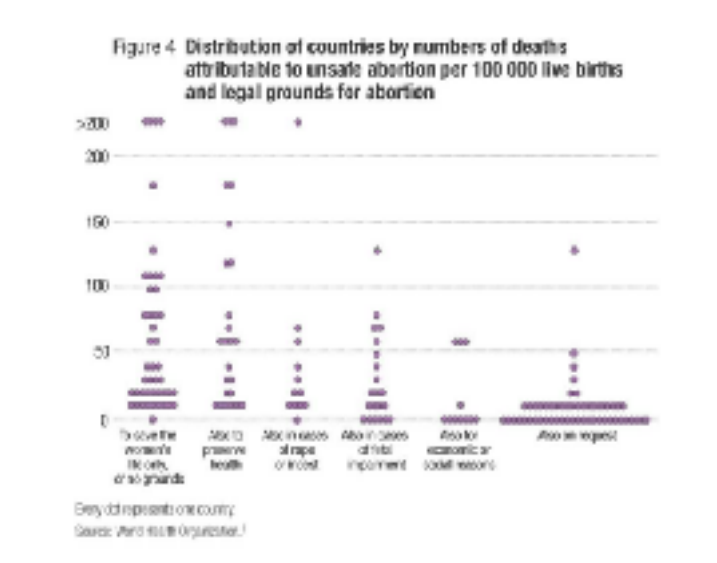
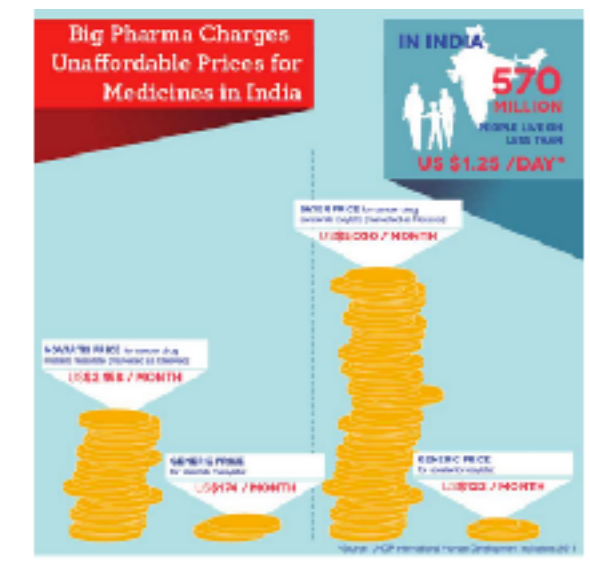


Table 3: Percentage of median household income from non-employment (total income)

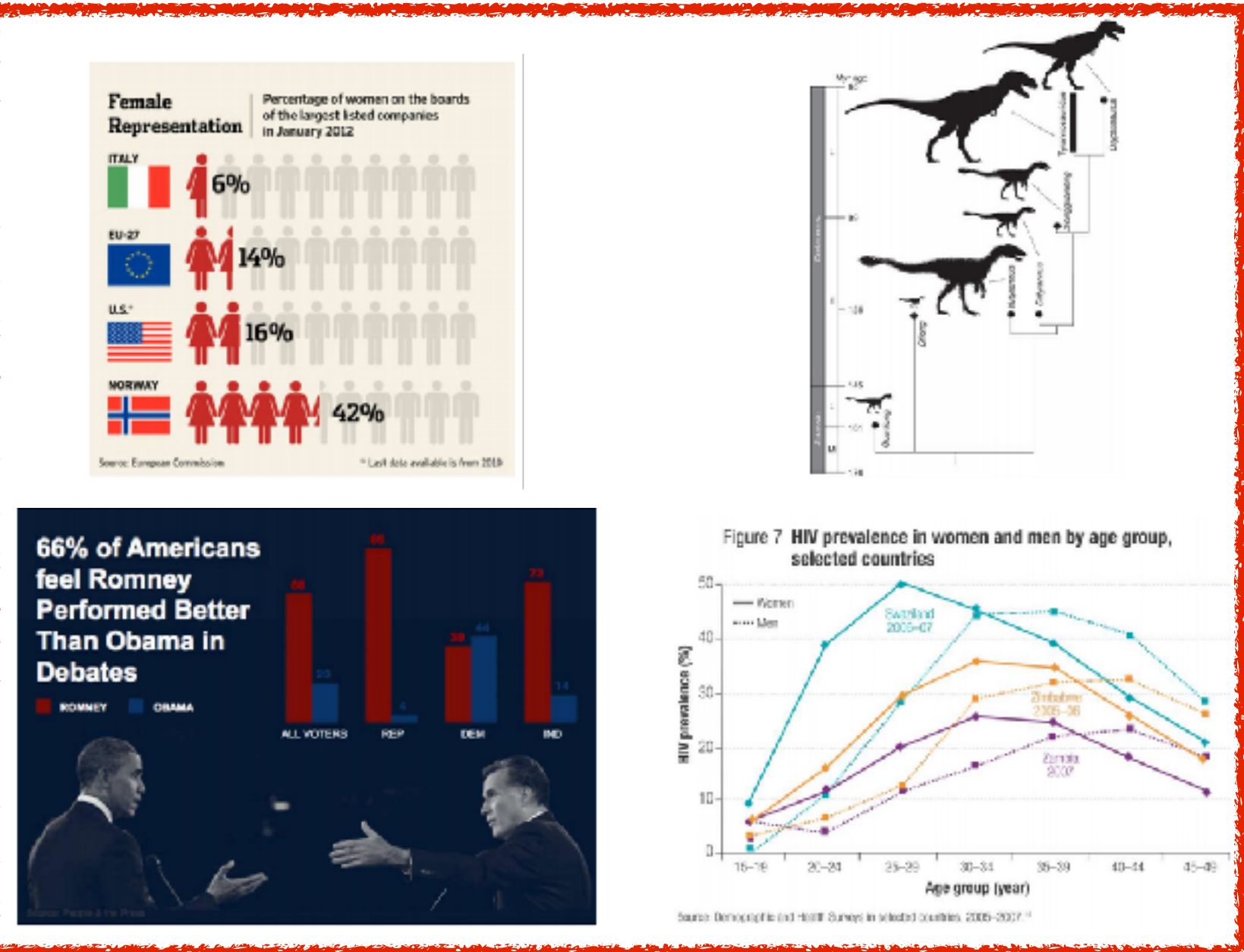
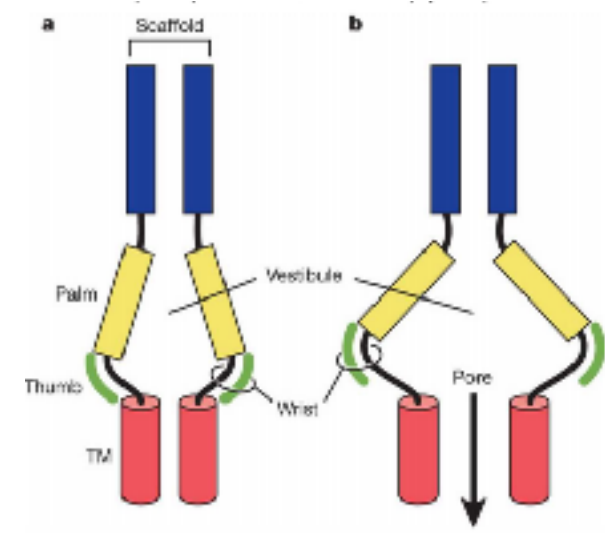
Country	2000	2005
Australia	11.1	11.1
Canada	11.1	11.1
France	11.1	11.1
Germany	11.1	11.1
Italy	11.1	11.1
Japan	11.1	11.1
UK	11.1	11.1
USA	11.1	11.1



FORGETTABLE

MEMORABLE

“Effective”



Visual Associations

Semantic Associations

Message redundancy

Data redundancy

Titles/annotations

HIGH QUALITY DESCRIPTION

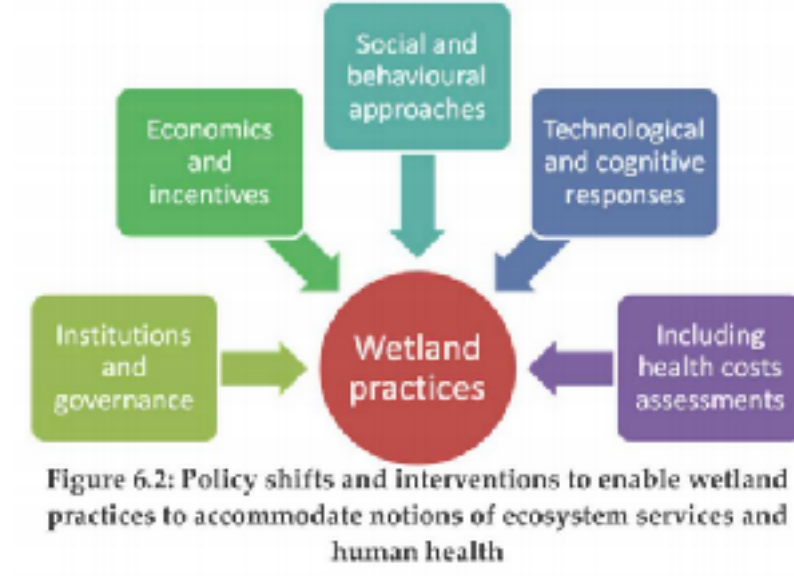
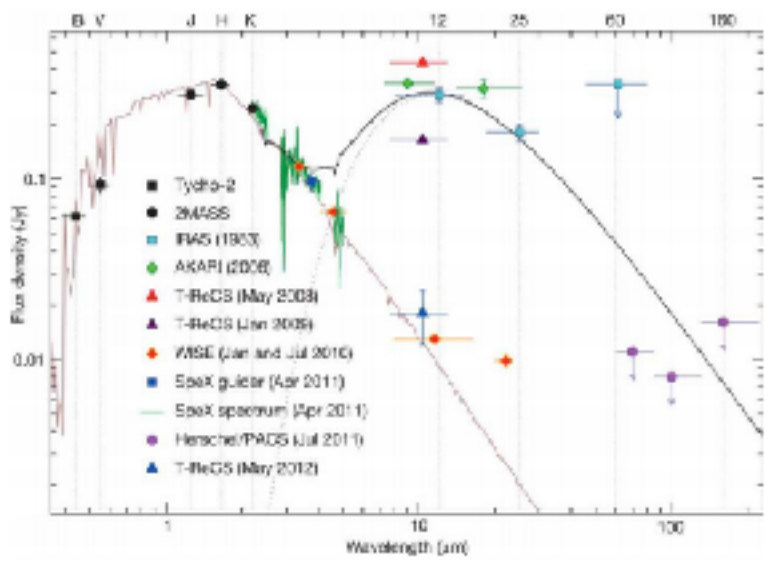
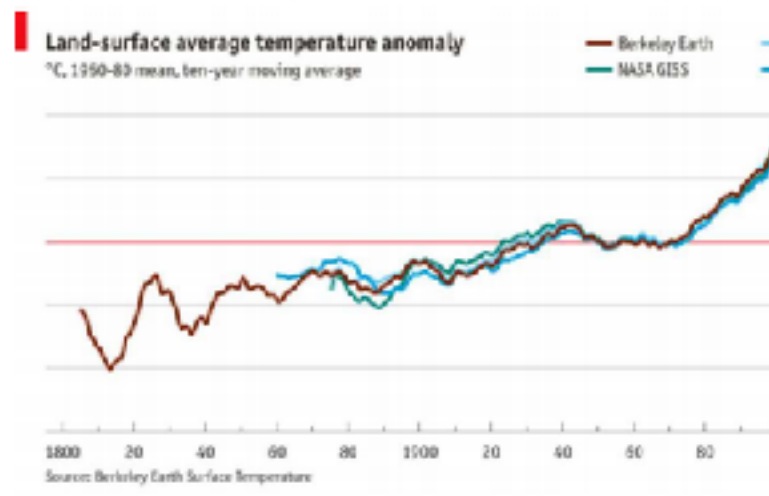


Figure 6.2: Policy shifts and interventions to enable wetland practices to accommodate notions of ecosystem services and human health

LOW QUALITY DESCRIPTION



News Corp. Results, by Segment
for the nine months ended March 31

Revenue	Operating profit/loss
\$566 billion	\$250 billion
6.22	\$458 million
5.56	\$1.01 billion
3.65	\$493 million
2.79	\$365 million
0.45	-\$437 million

Source: the company
The Wall Street Journal

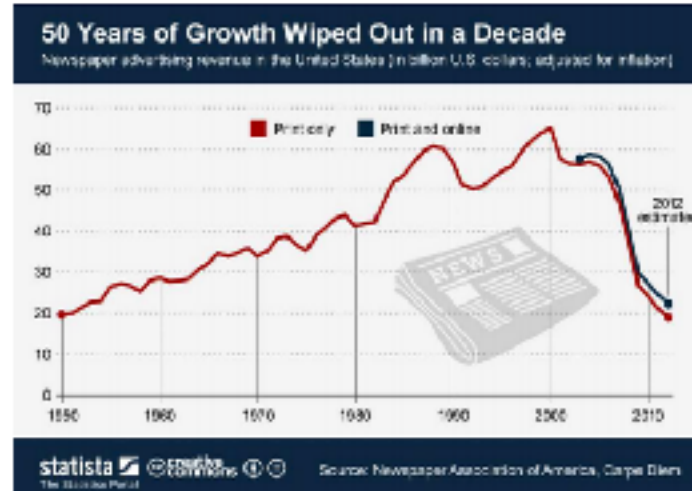
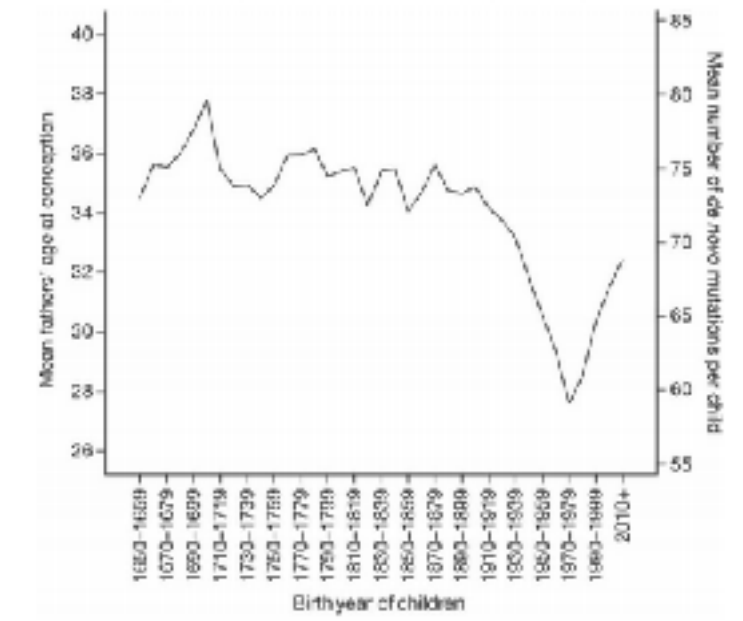
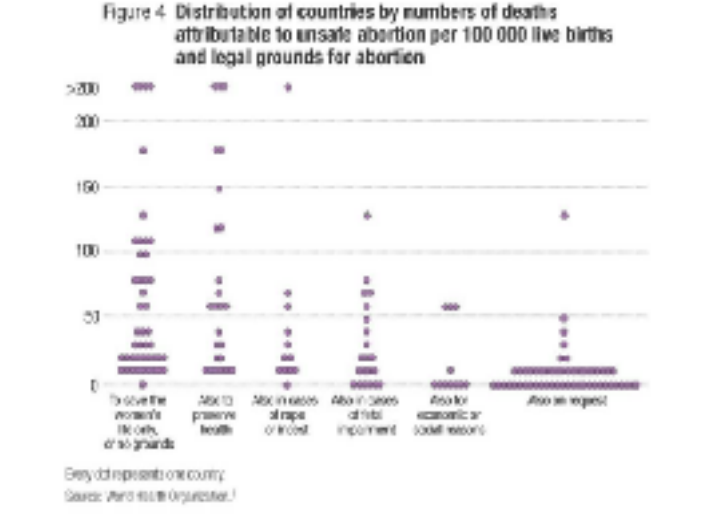
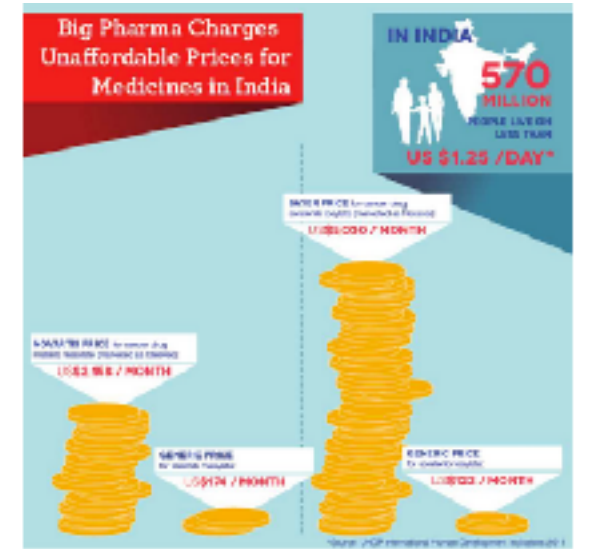


Table 3: Percentage of median household income spent on health care in 2010

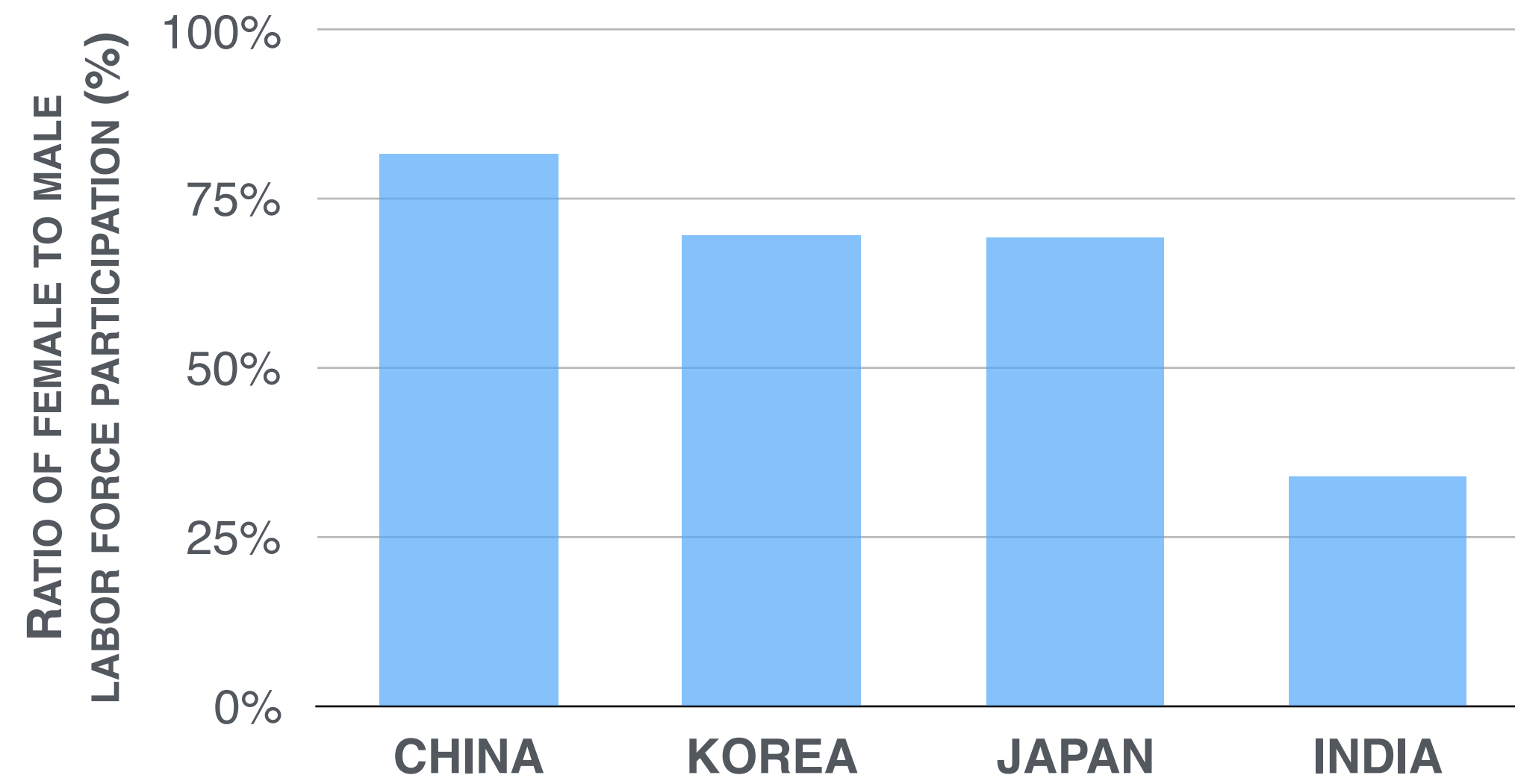
Country	Percentage
Australia	10.1
Austria	10.1
Belgium	10.1
Canada	10.1
Denmark	10.1
France	10.1
Germany	10.1
Italy	10.1
Japan	10.1
South Korea	10.1
Spain	10.1
Sweden	10.1
Switzerland	10.1
United Kingdom	10.1
United States	10.1



FORGETTABLE

Examples

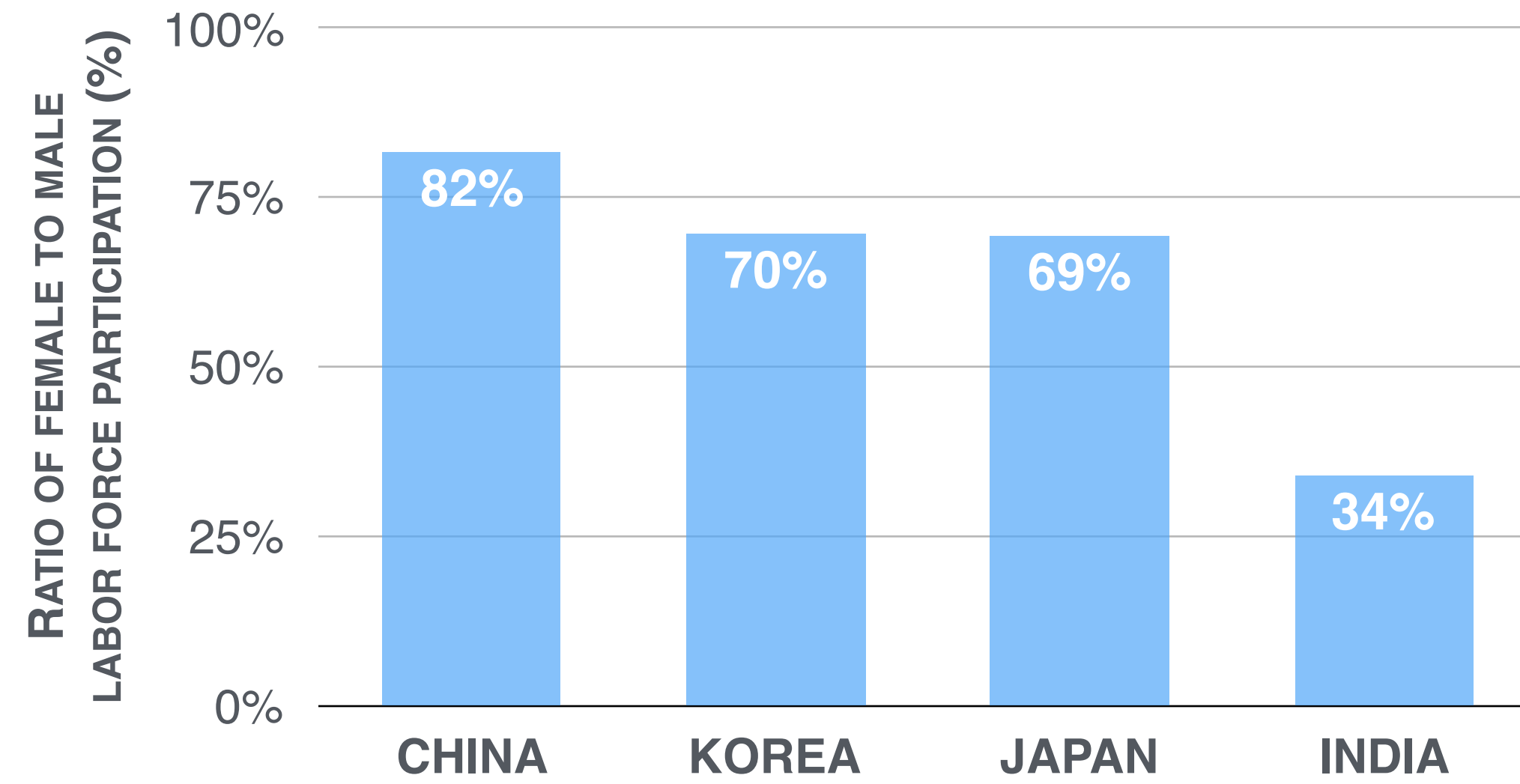
GENDER EQUALITY IN LABOR FORCE PARTICIPATION



Source: Gender Statistics 2013, World Bank

ORIGINAL

GENDER EQUALITY IN LABOR FORCE PARTICIPATION

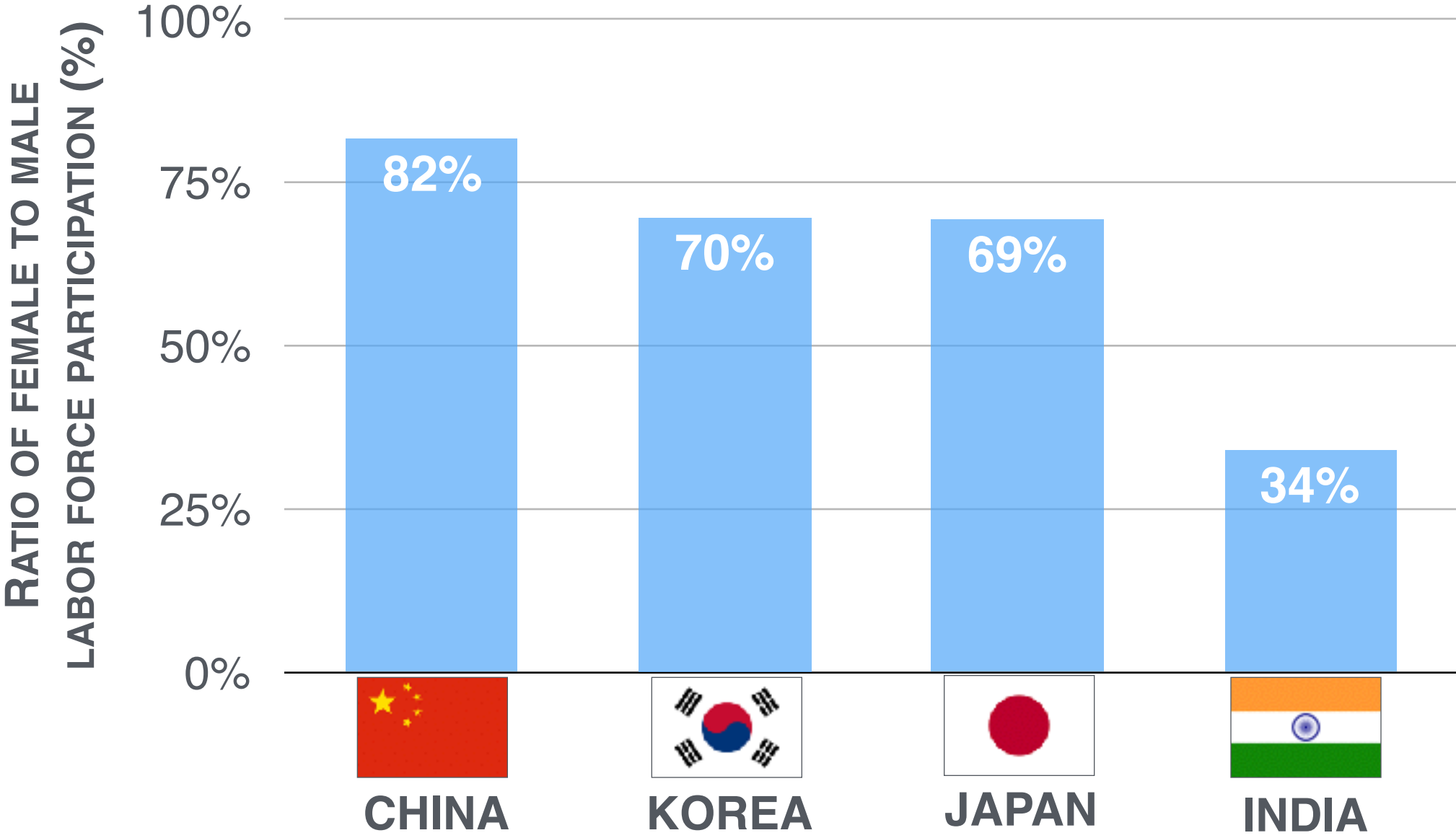


Source: Gender Statistics 2013, World Bank

DATA REDUNDANCY

GENDER EQUALITY IN LABOR FORCE PARTICIPATION

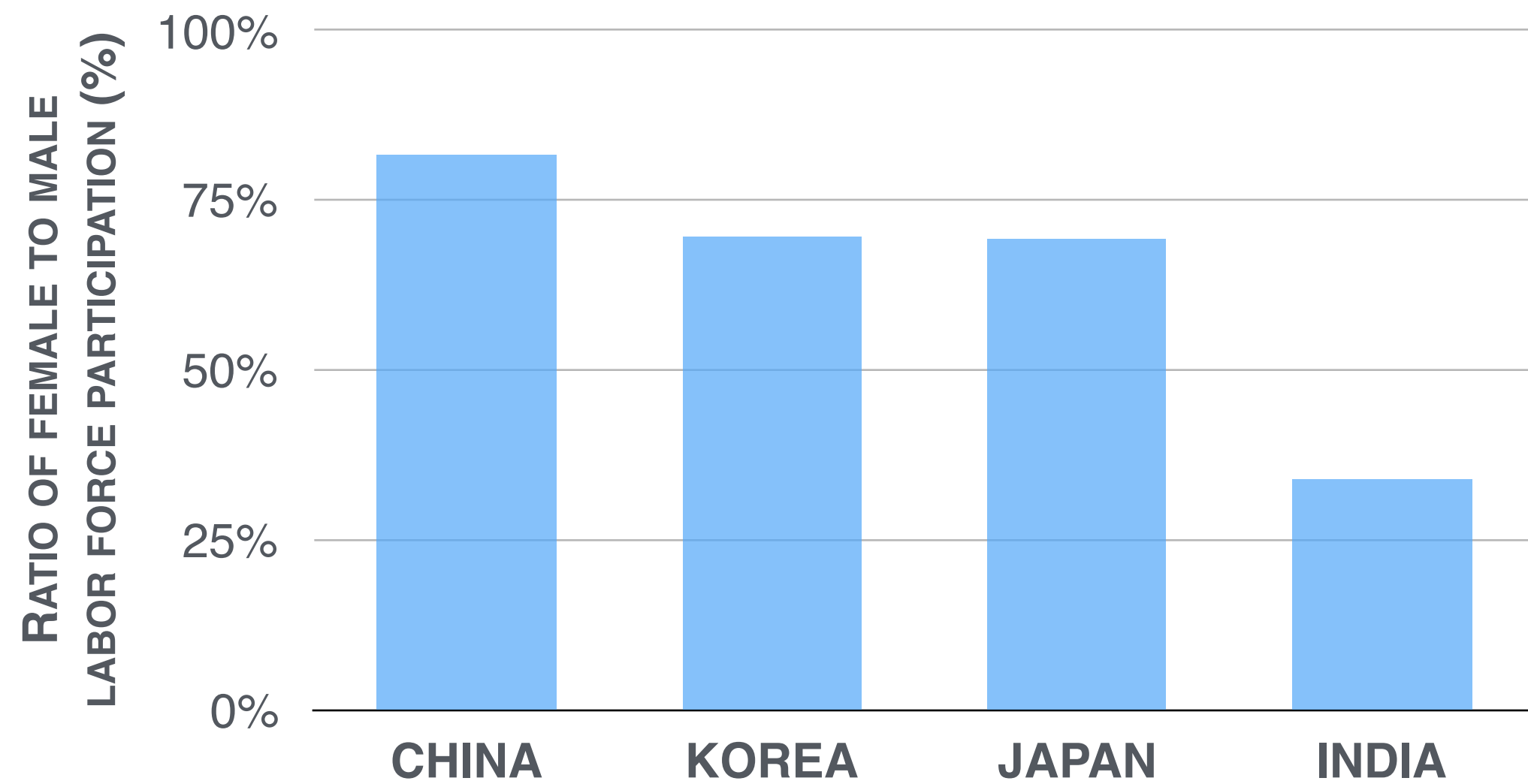
CHINA LEADS IN FEMALE LABOR FORCE PARTICIPATION WHEREAS INDIA LAGS SIGNIFICANTLY BEHIND AMONG OTHER ASIAN-PACIFIC COUNTRIES IN 2013.



Source: Gender Statistics 2013, World Bank

DATA & MESSAGE REDUNDANCY

GENDER EQUALITY IN LABOR FORCE PARTICIPATION

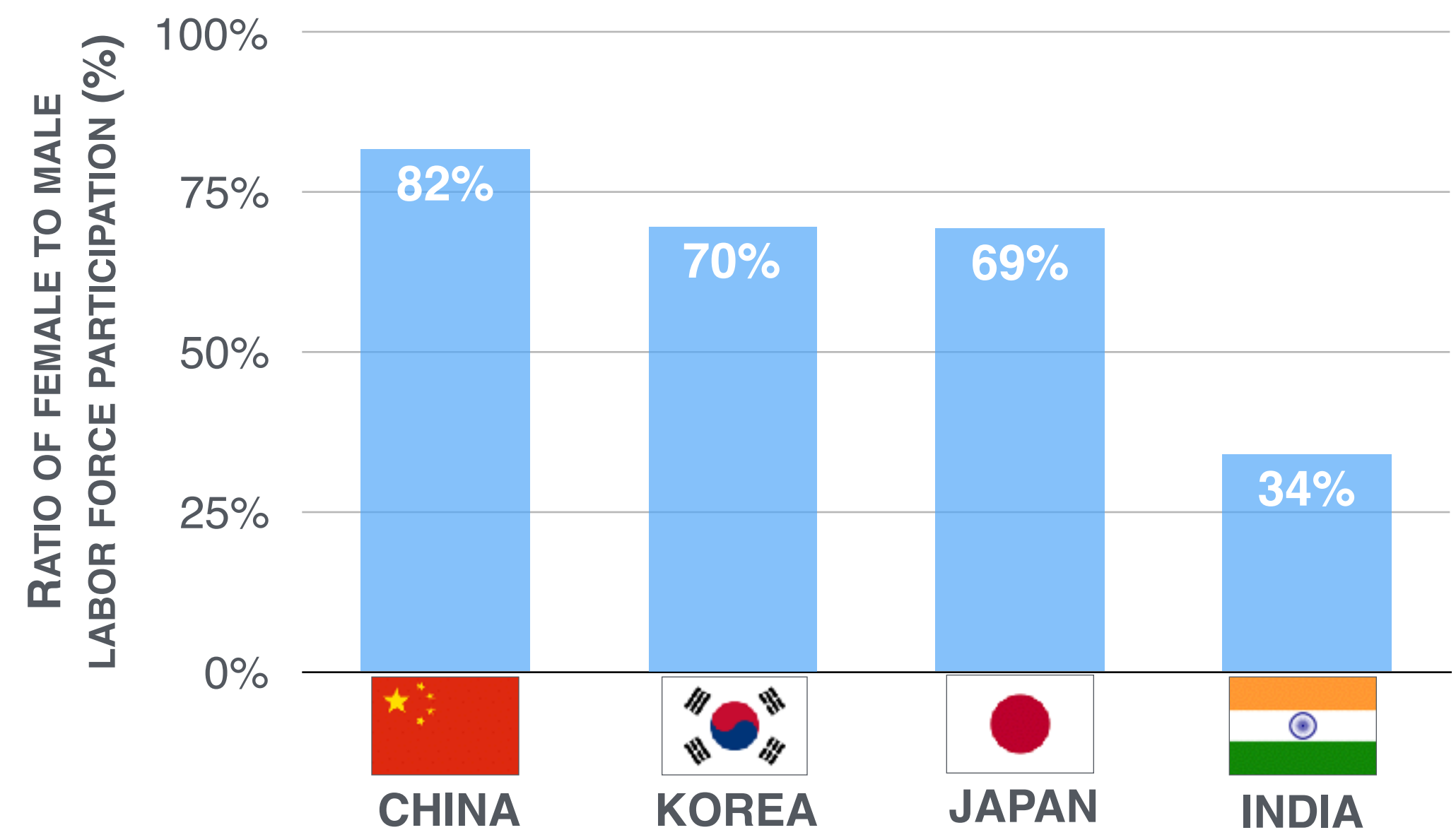


Source: Gender Statistics 2013, World Bank

ORIGINAL

GENDER EQUALITY IN LABOR FORCE PARTICIPATION

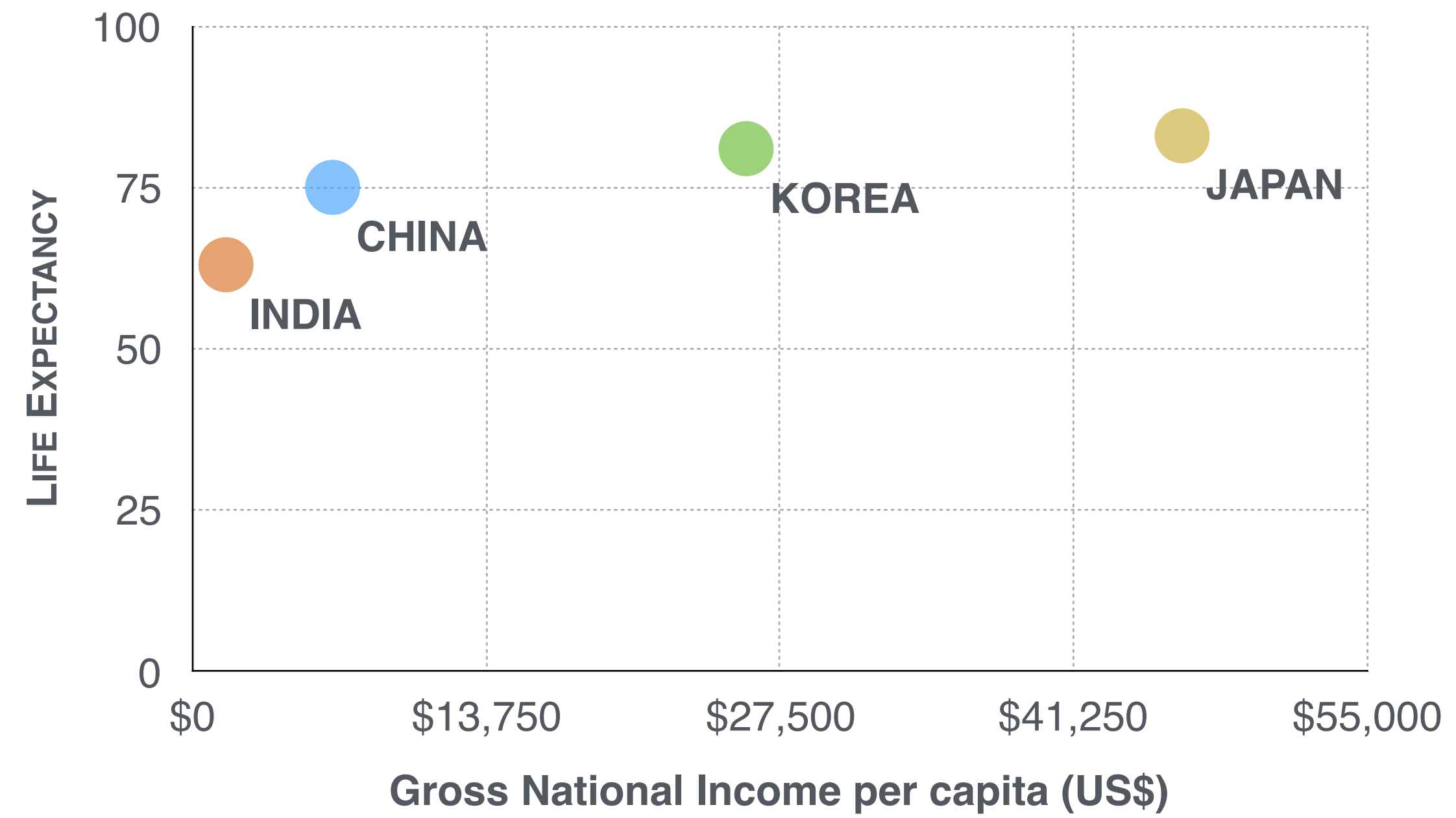
CHINA LEADS IN FEMALE LABOR FORCE PARTICIPATION WHEREAS INDIA LAGS SIGNIFICANTLY BEHIND AMONG OTHER ASIAN-PACIFIC COUNTRIES IN 2013.



Source: Gender Statistics 2013, World Bank

DATA & MESSAGE REDUNDANCY

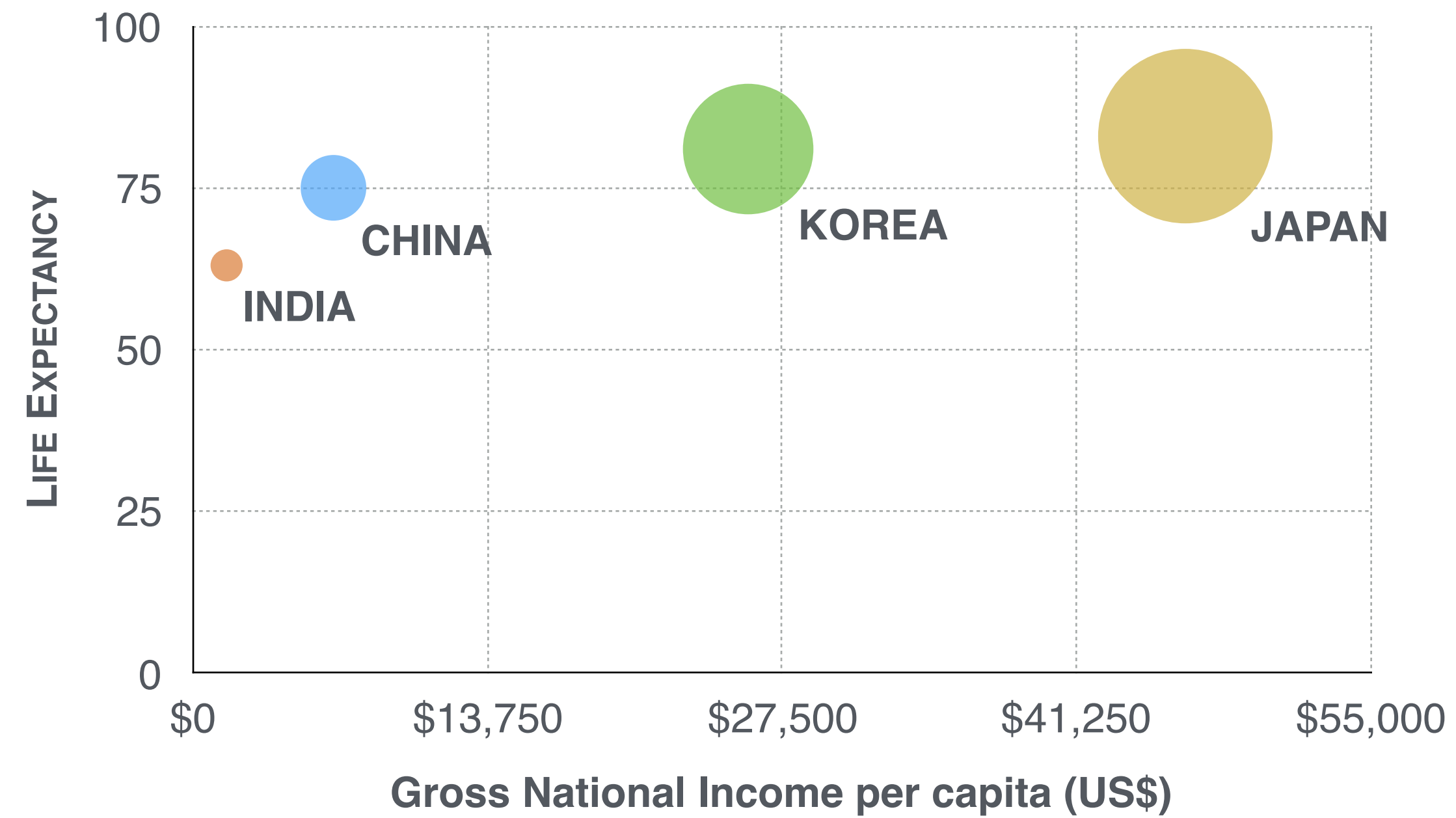
LIFE EXPECTANCY VS ECONOMIC GROWTH



Source: World Bank national accounts data, and
OECD National Accounts data files (2013).

ORIGINAL

LIFE EXPECTANCY VS ECONOMIC GROWTH

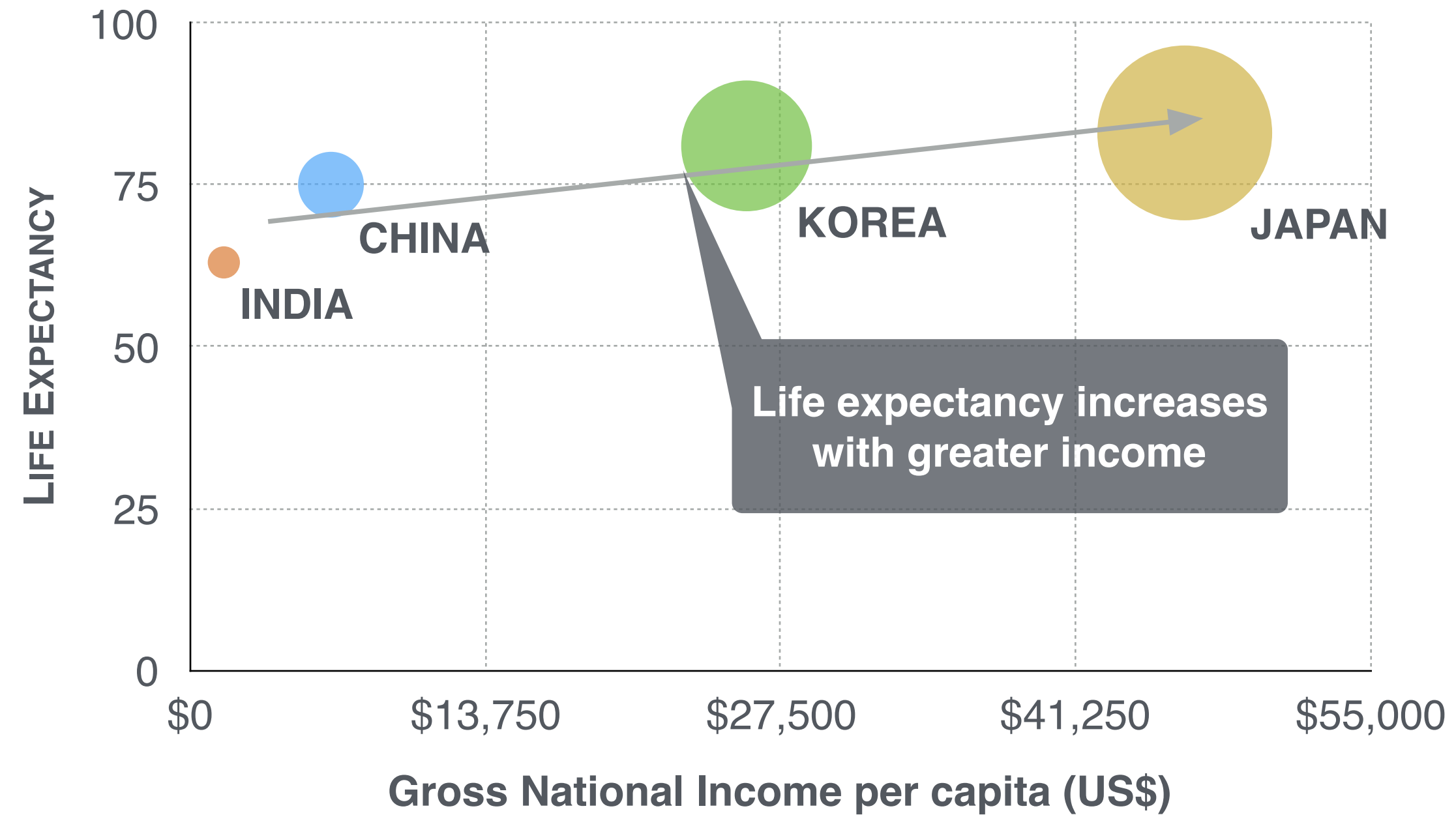


Source: World Bank national accounts data, and
OECD National Accounts data files (2013).

DATA REDUNDANCY

LIFE EXPECTANCY VS ECONOMIC GROWTH

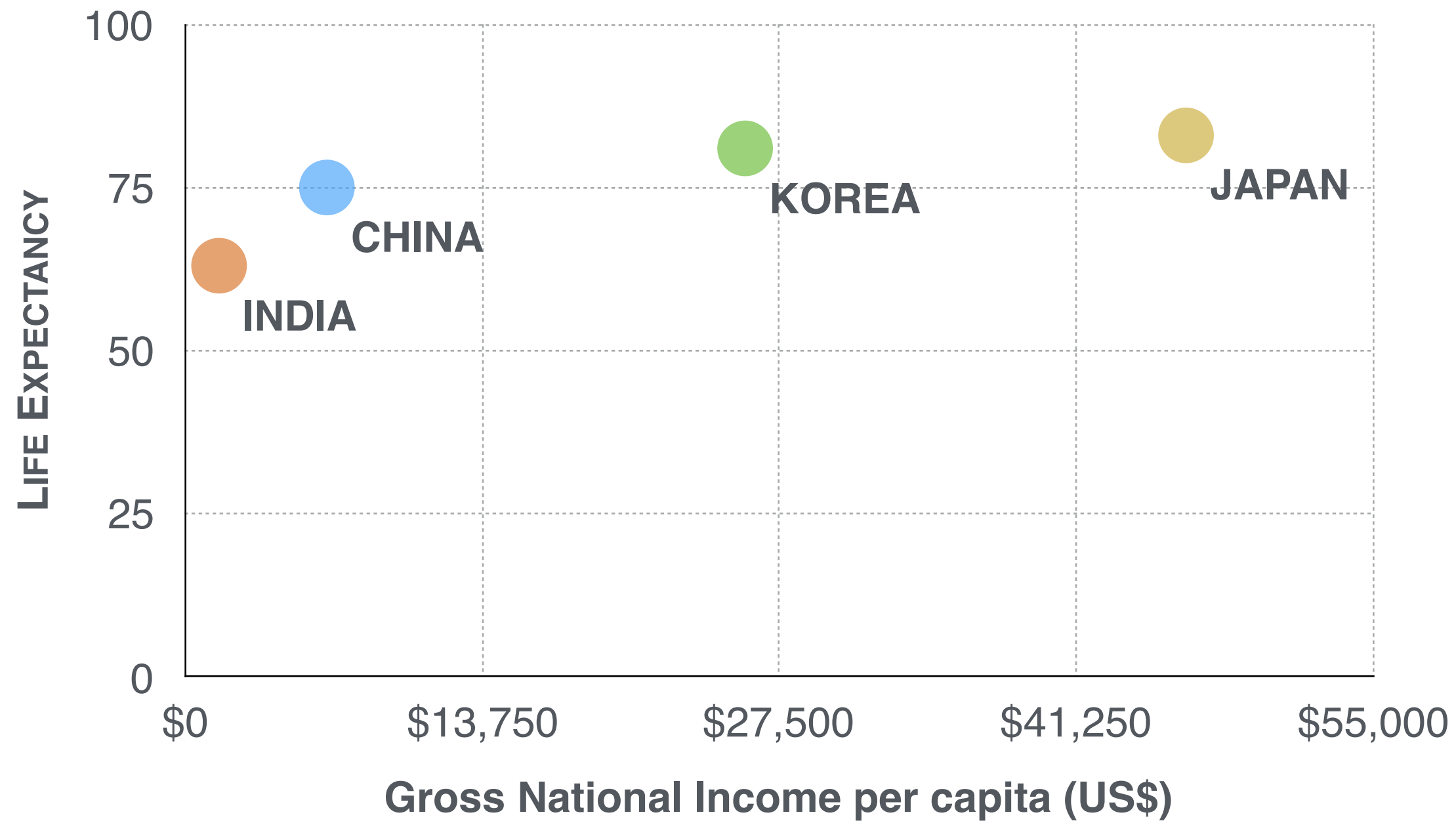
RELATION BETWEEN LIFE EXPECTANCY AND GROSS NATIONAL INCOME OF ASIAN-PACIFIC COUNTRIES IN 2013



Source: World Bank national accounts data, and OECD National Accounts data files (2013).

DATA & MESSAGE REDUNDANCY

LIFE EXPECTANCY VS ECONOMIC GROWTH

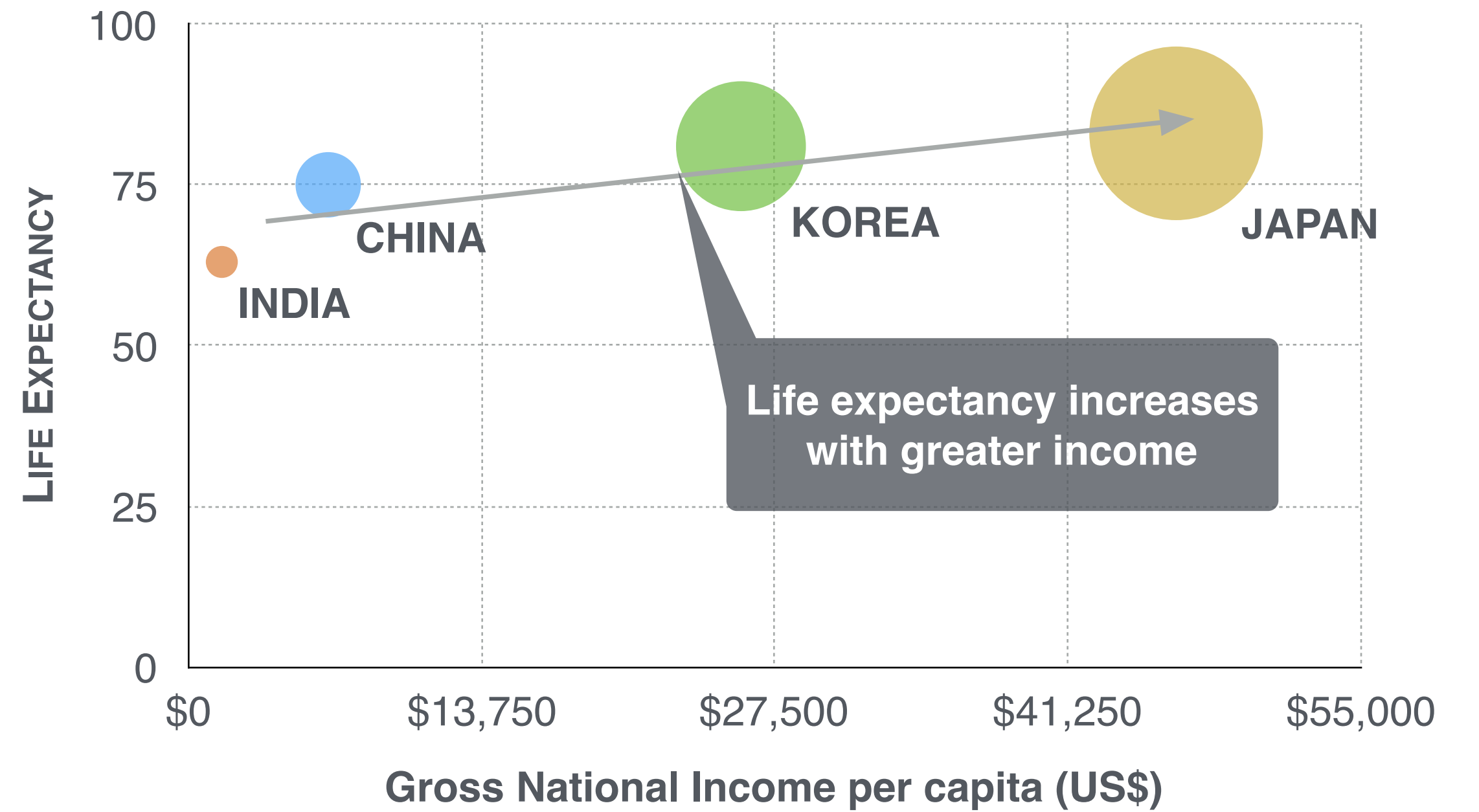


Source: World Bank national accounts data, and OECD National Accounts data files (2013).

ORIGINAL

LIFE EXPECTANCY VS ECONOMIC GROWTH

RELATION BETWEEN LIFE EXPECTANCY AND GROSS NATIONAL INCOME OF ASIAN-PACIFIC COUNTRIES IN 2013

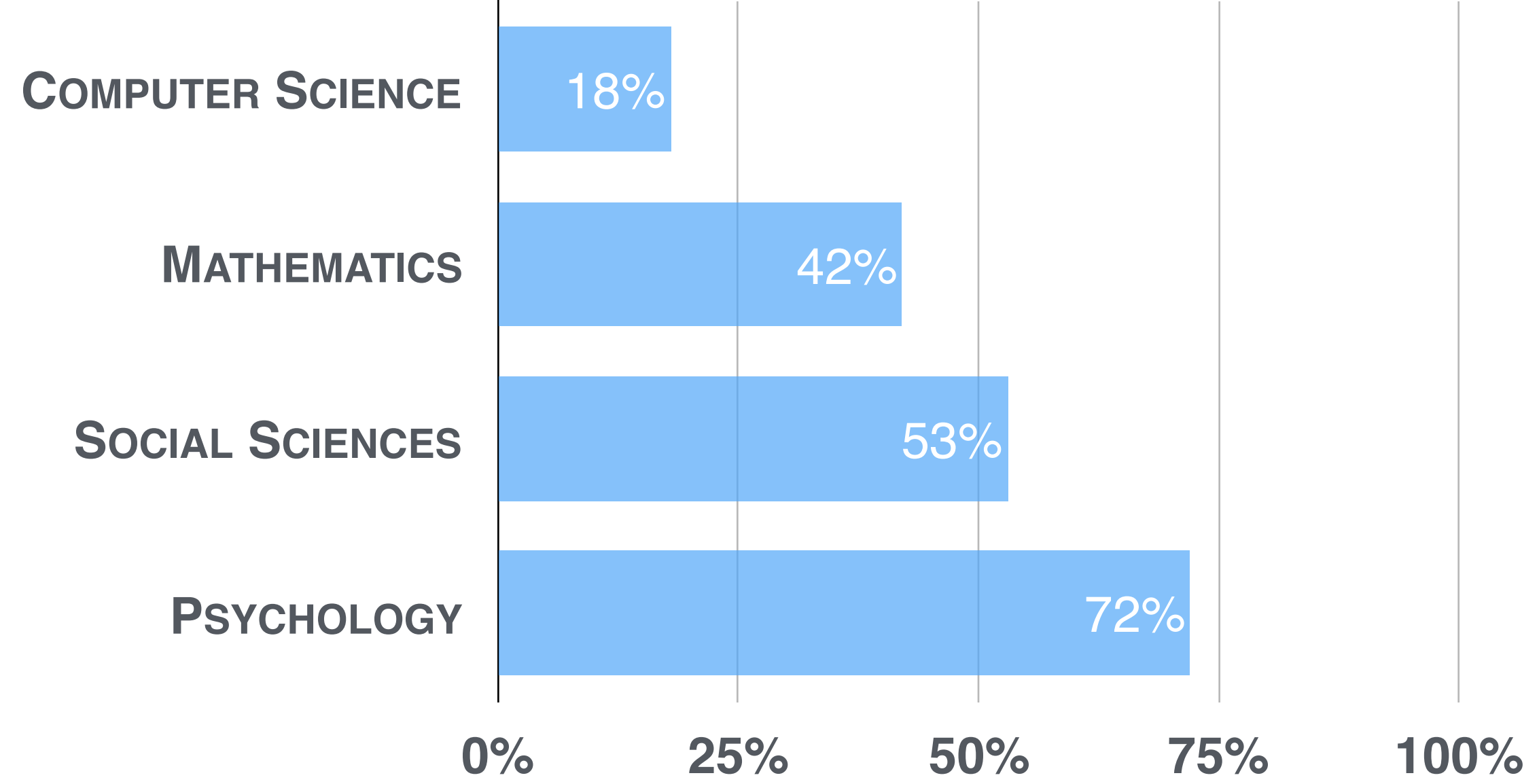


Source: World Bank national accounts data, and OECD National Accounts data files (2013).

DATA & MESSAGE REDUNDANCY

VERY FEW WOMEN ARE STUDYING COMPUTER SCIENCE

% of bachelor degrees awarded for women

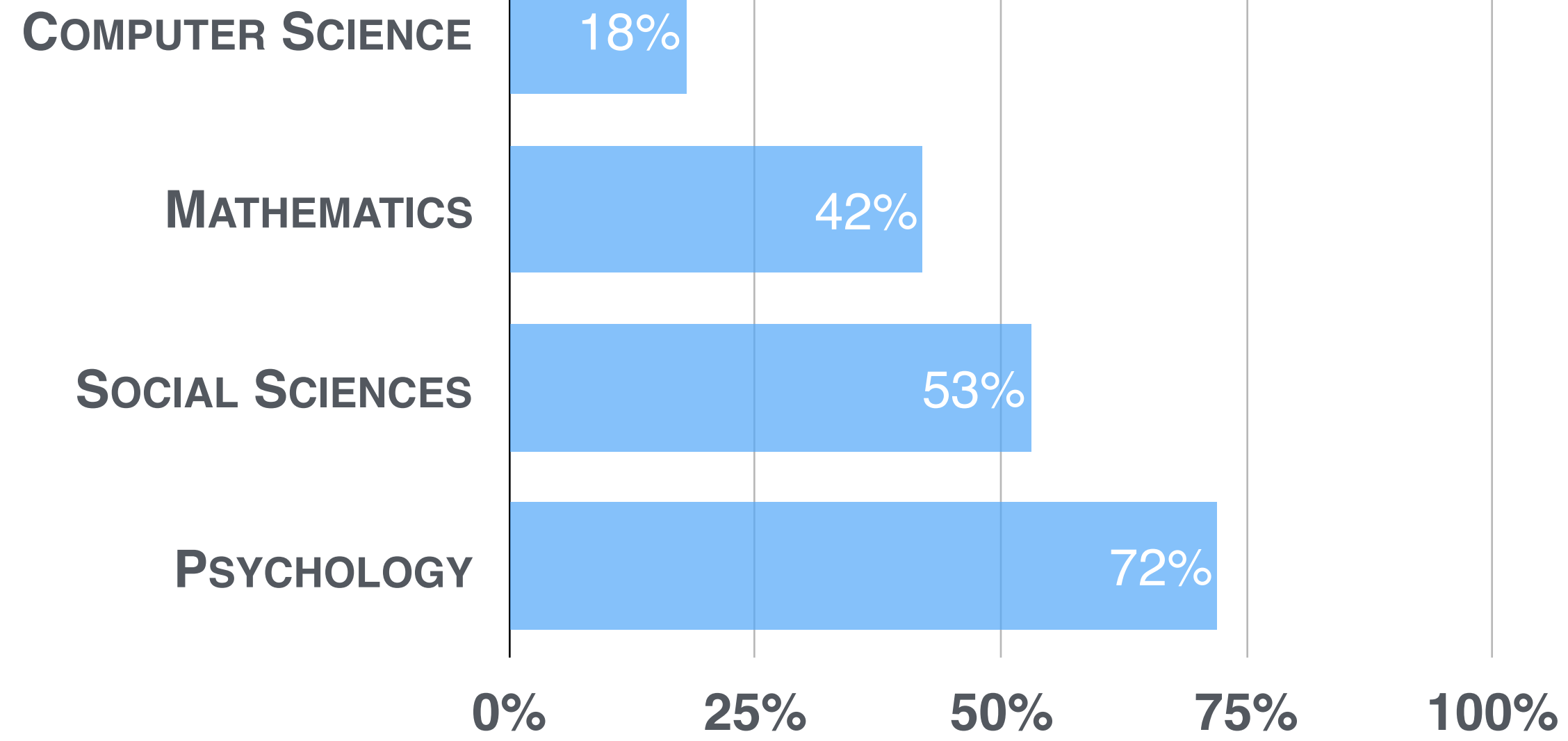


Source: National Student Clearinghouse Research Center (2014).

WITHOUT PICTOGRAM

VERY FEW WOMEN ARE STUDYING COMPUTER SCIENCE

% of bachelor degrees awarded for women

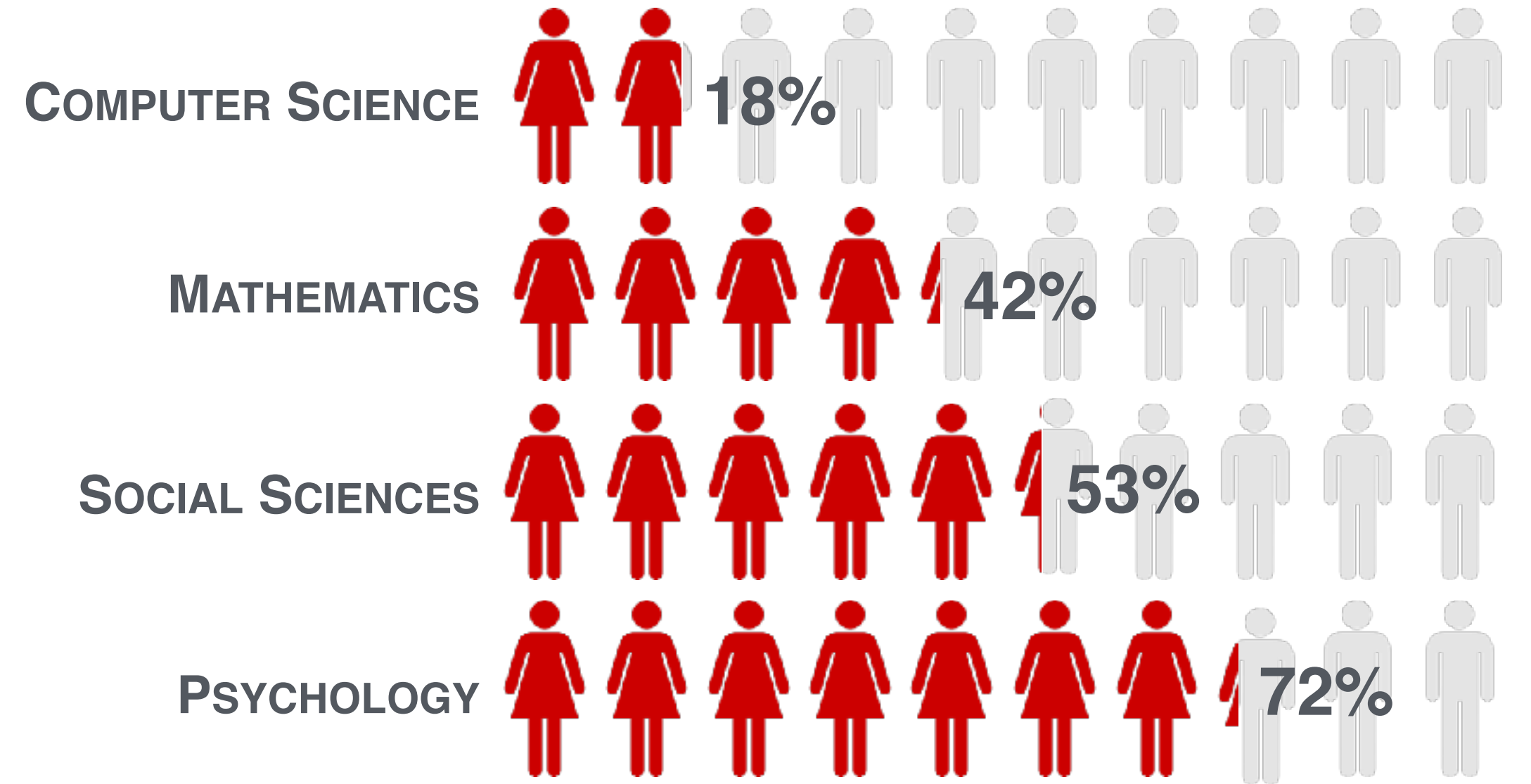


Source: National Student Clearinghouse Research Center (2014).

WITHOUT PICTOGRAM

VERY FEW WOMEN ARE STUDYING COMPUTER SCIENCE

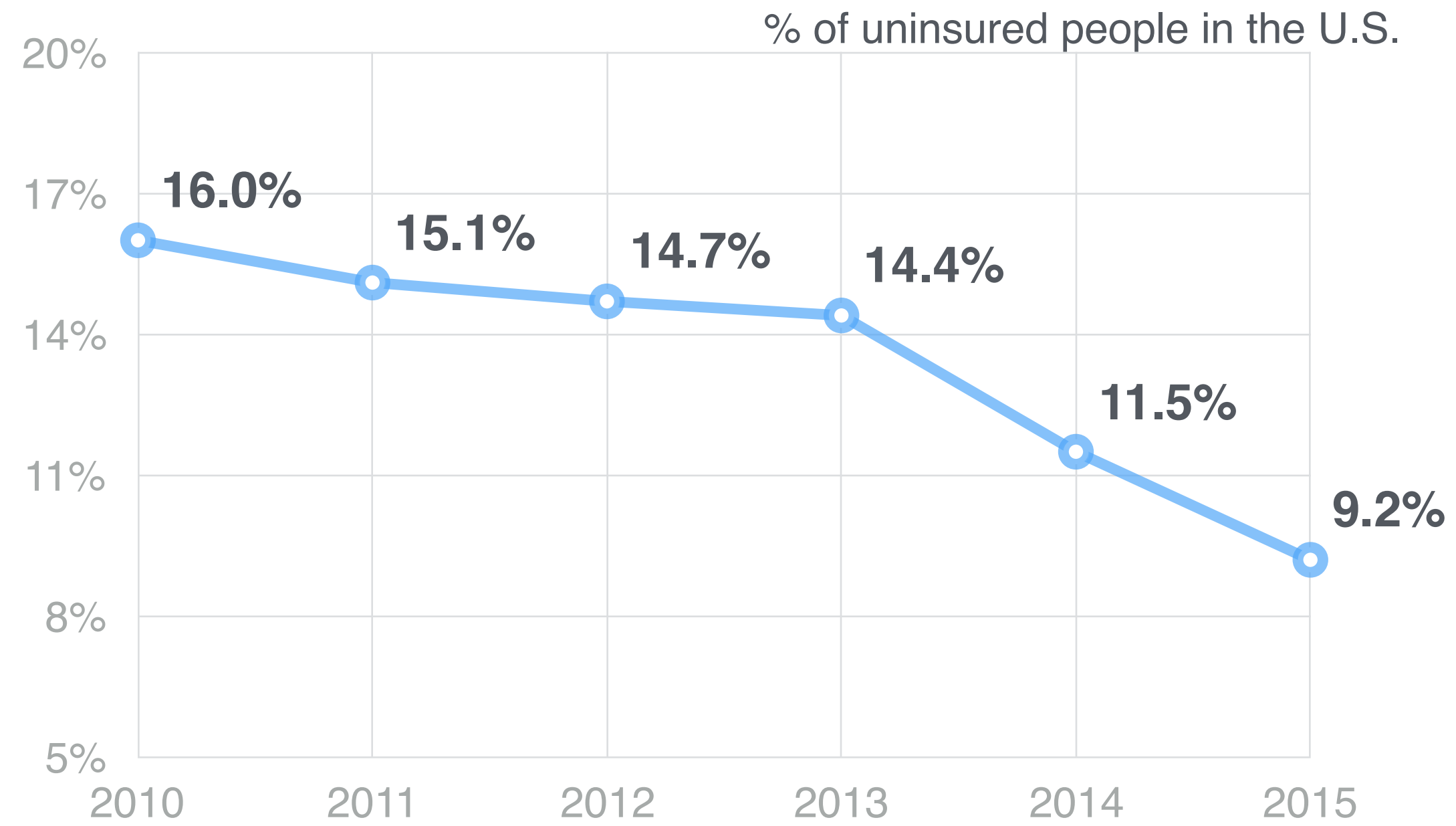
% of bachelor degrees awarded for women



Source: National Student Clearinghouse Research Center (2014).

WITH PICTOGRAM

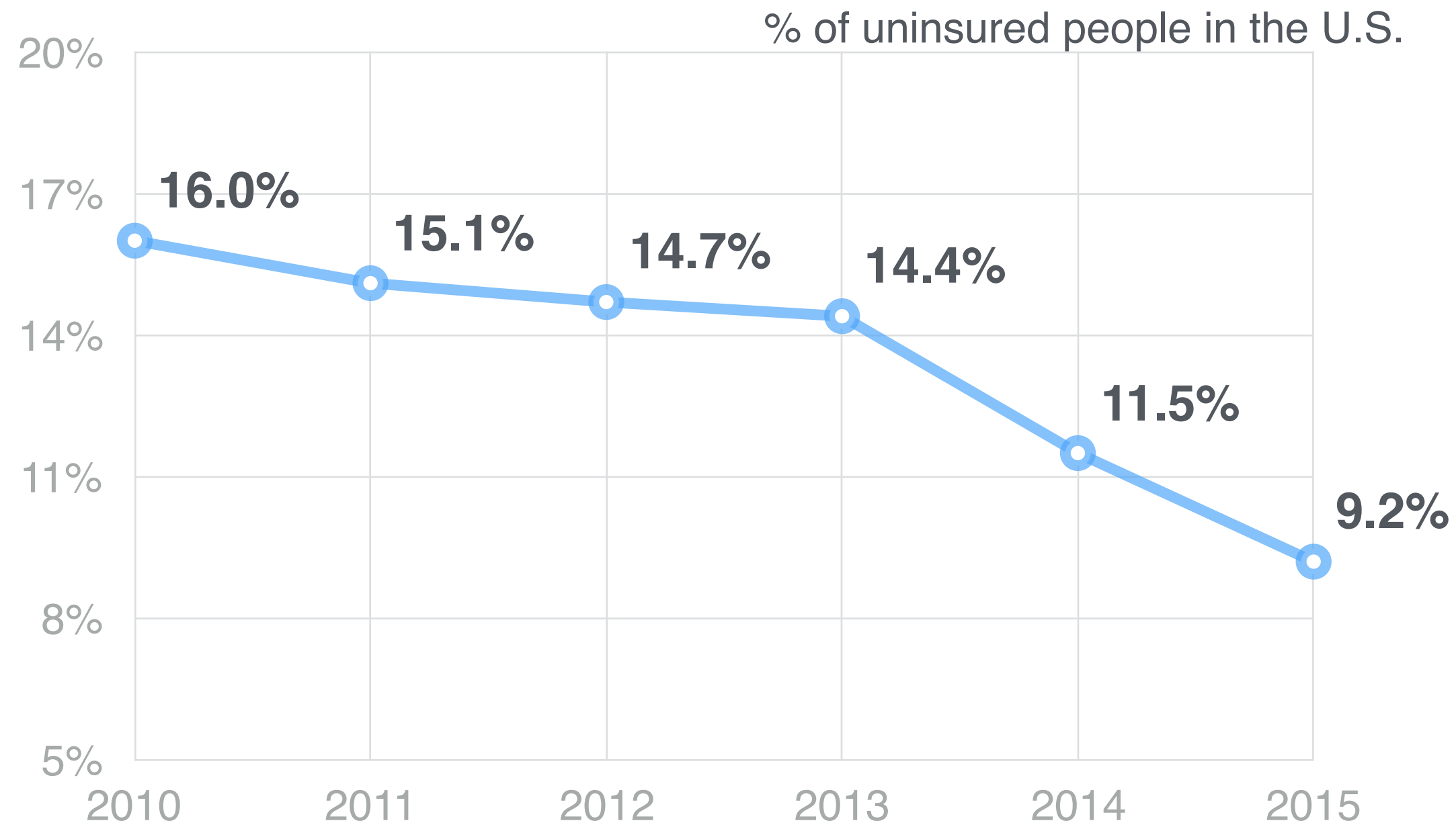
PERCENTAGE OF UNINSURED AMERICANS



Source: CDC/NCHS, National Health Interview Survey, 2010–2015

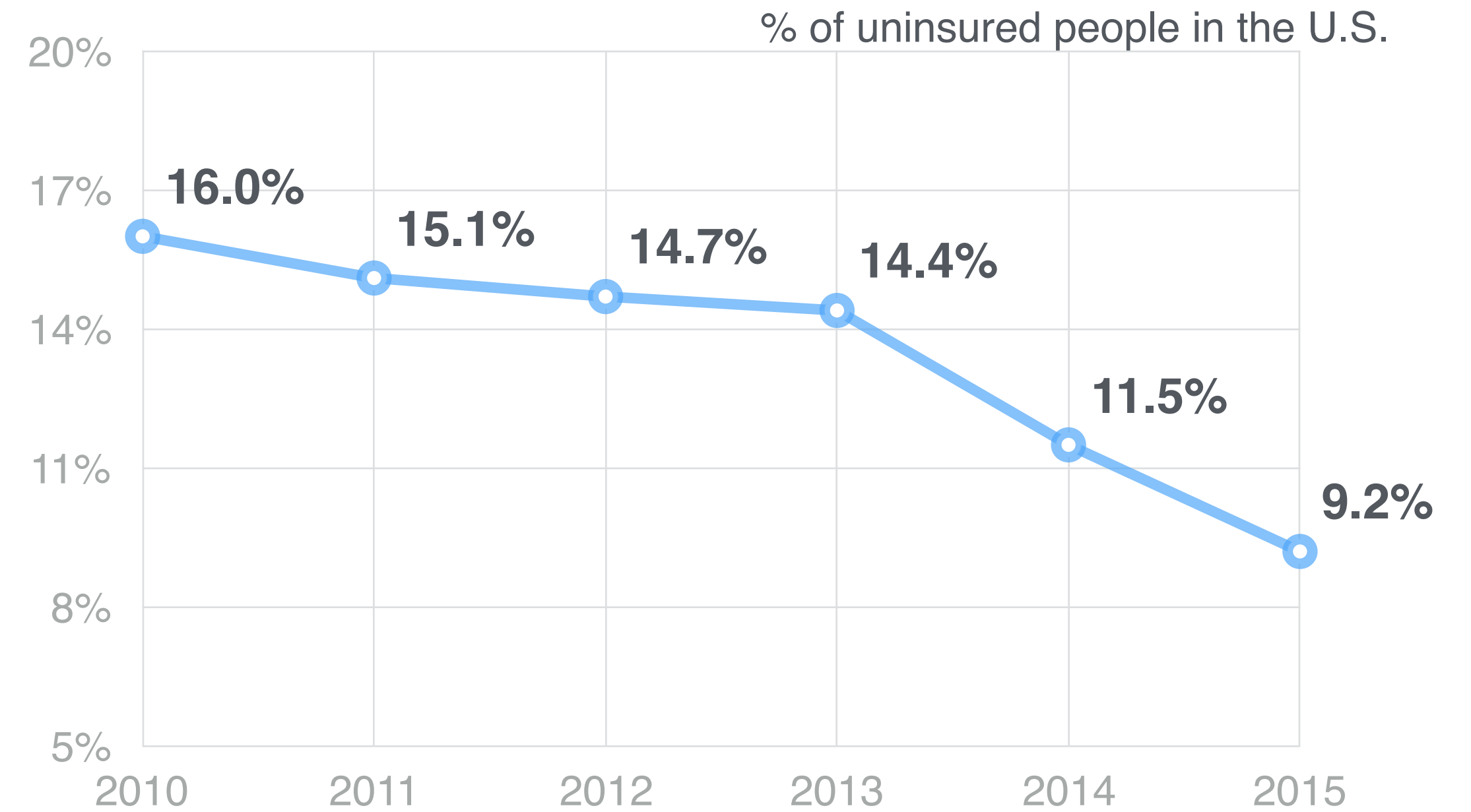
BAD TITLE

PERCENTAGE OF UNINSURED AMERICANS



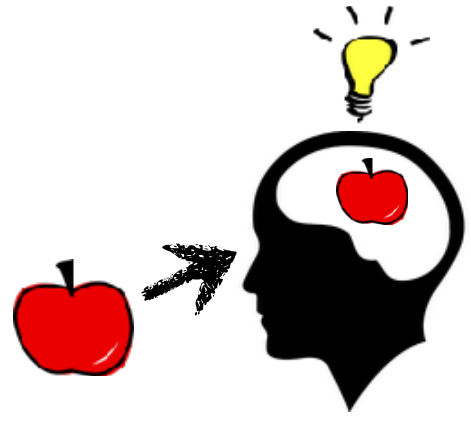
BAD TITLE

AMERICA'S UNINSURED RATE DIPS BELOW 10%



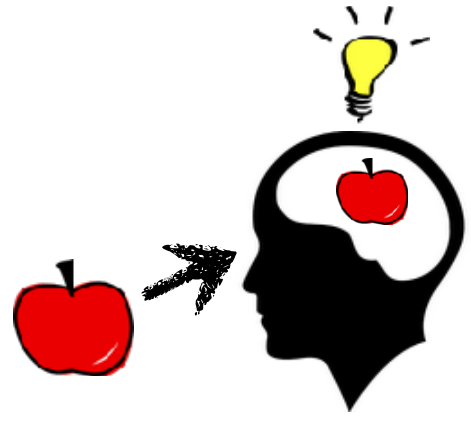
GOOD TITLE

Summary

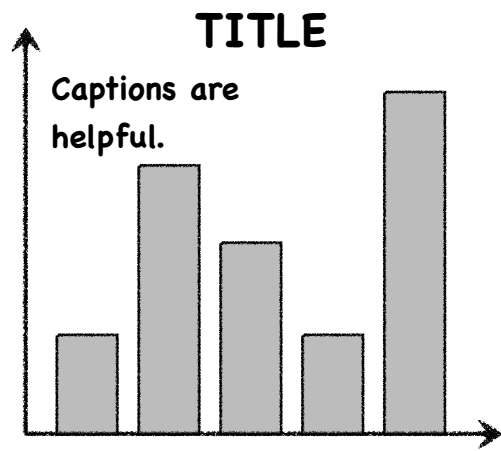


Visualizations that are memorable “at-a-glance” have memorable content.

Summary

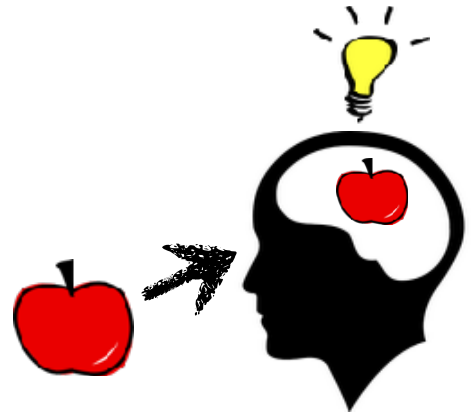


Visualizations that are memorable “at-a-glance” have memorable content.

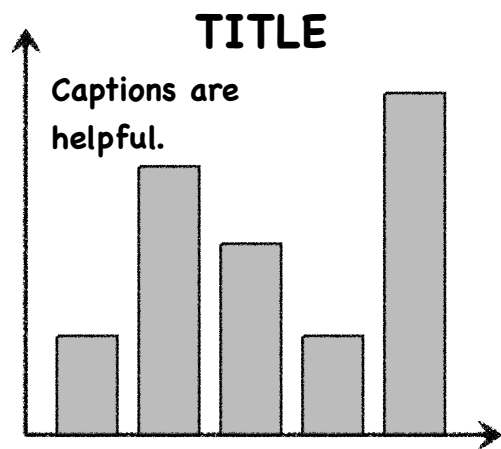


Titles and text are key elements in a visualization and help recall the message.

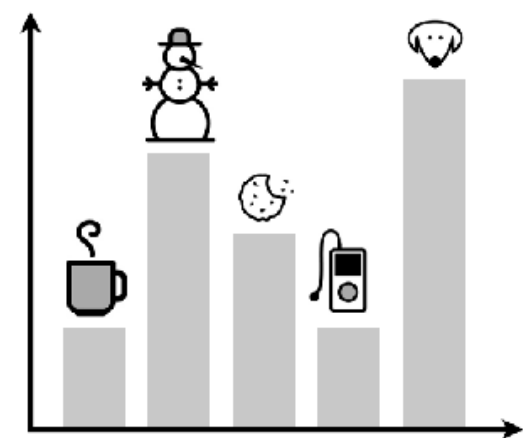
Summary



Visualizations that are memorable “at-a-glance” have memorable content.

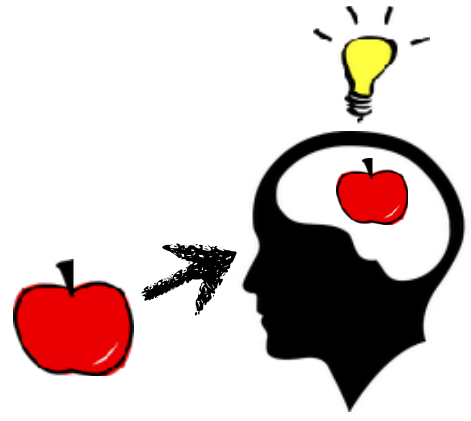


Titles and text are key elements in a visualization and help recall the message.

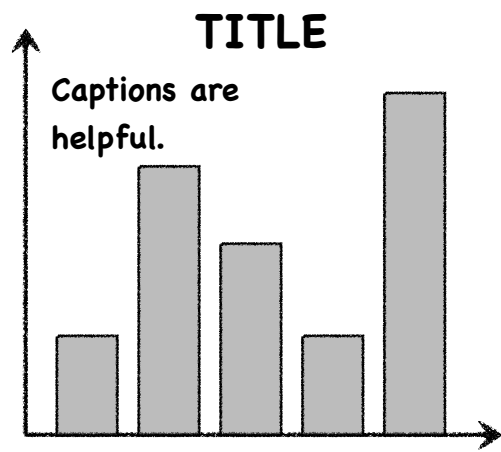


Human recognizable objects (e.g., pictograms) can help with the recognition or recall of a visualization.

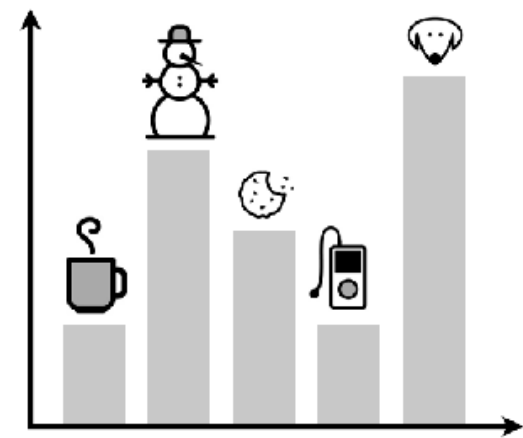
Summary



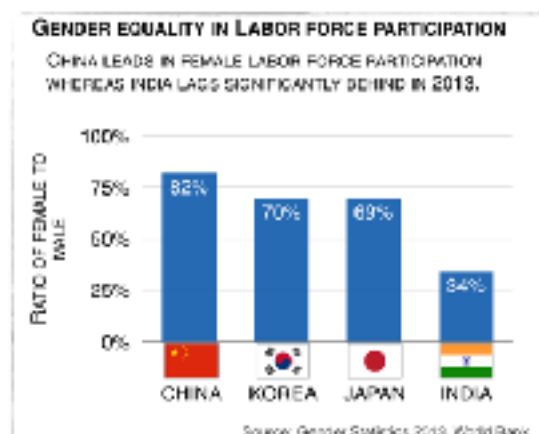
Visualizations that are memorable “at-a-glance” have memorable content.



Titles and text are key elements in a visualization and help recall the message.



Human recognizable objects (e.g., pictograms) can help with the recognition or recall of a visualization.

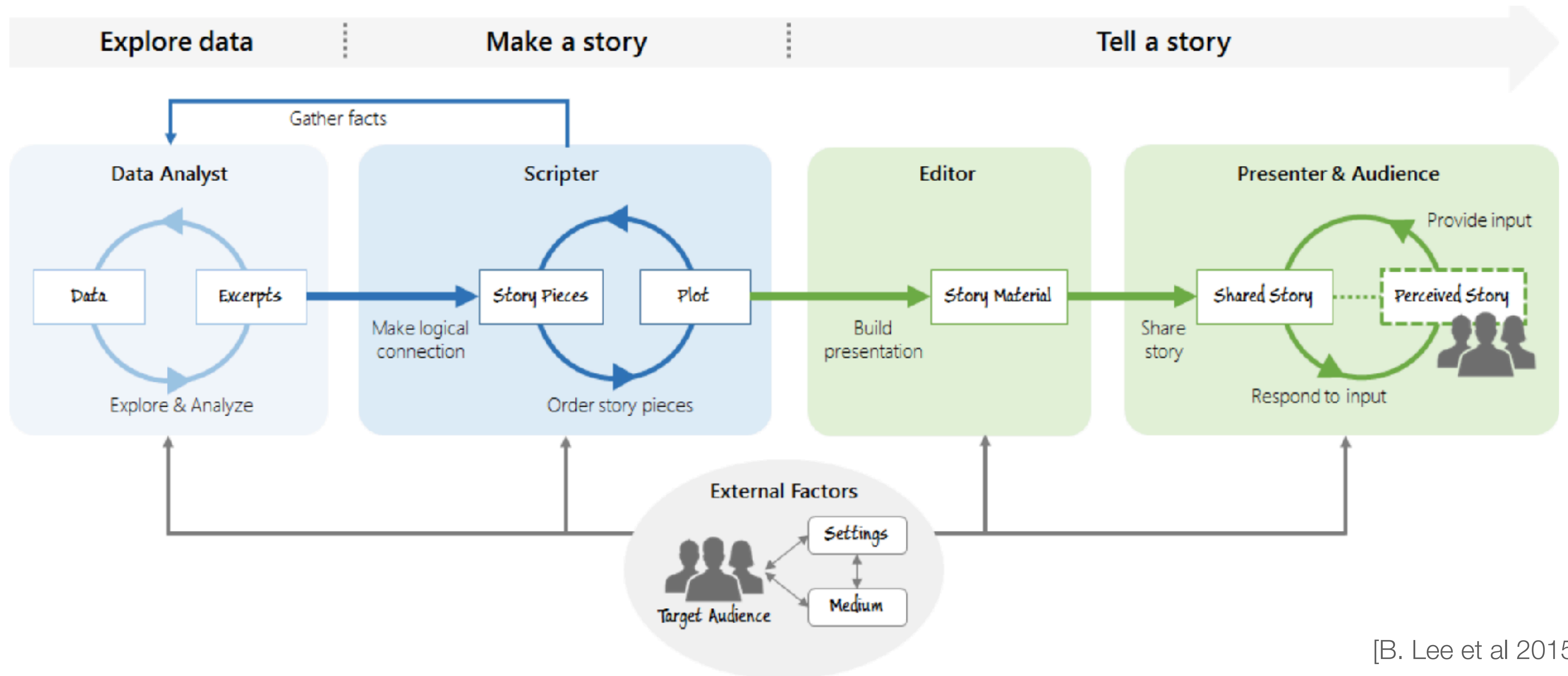


Redundancy helps with visualization recall and understanding.

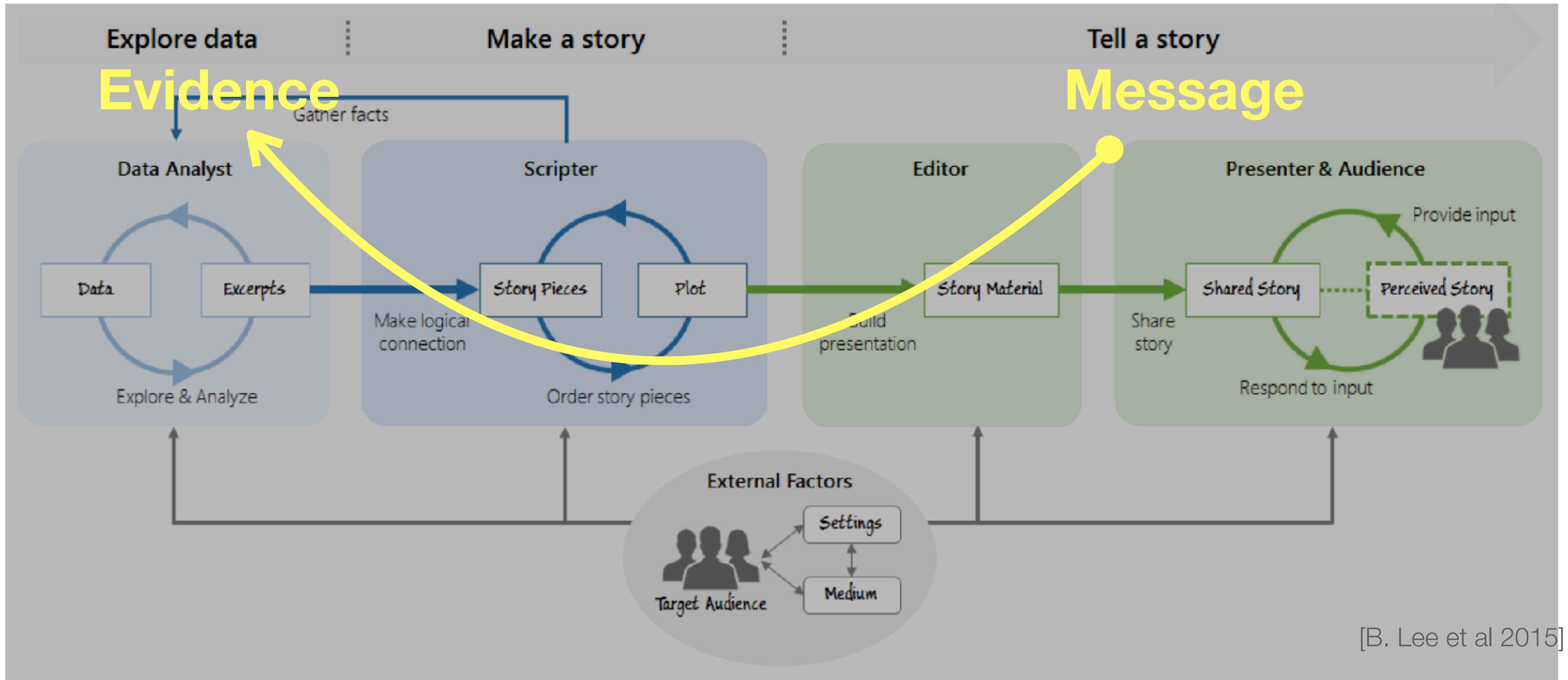
A single picture cannot tell compelling stories

Data-Driven Storytelling

Storytelling process: transforming data into visually shared stories.



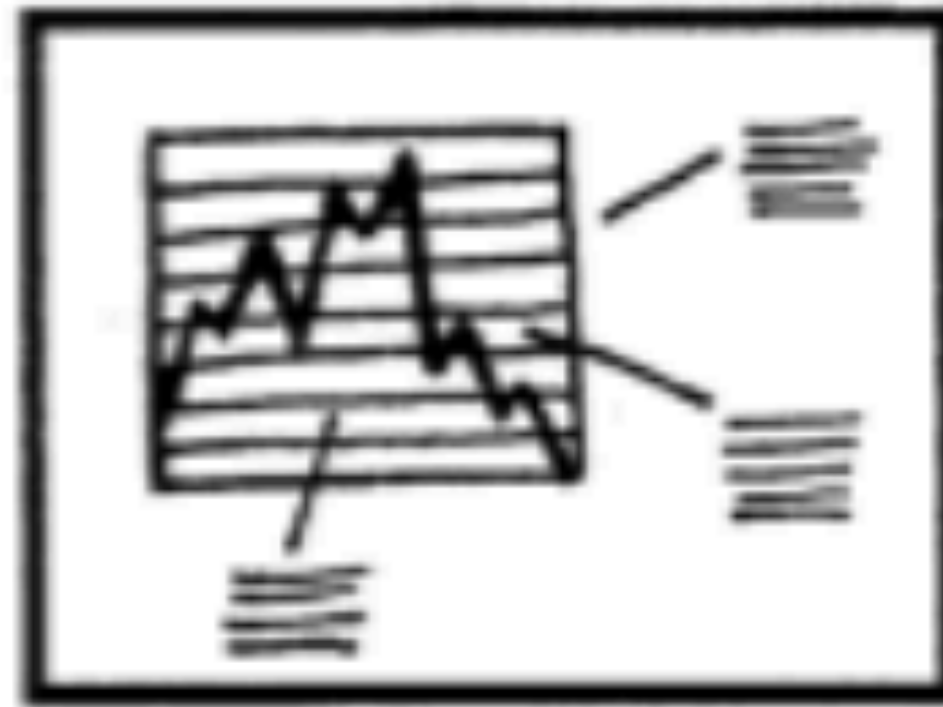
In reality...



Seven
Genres



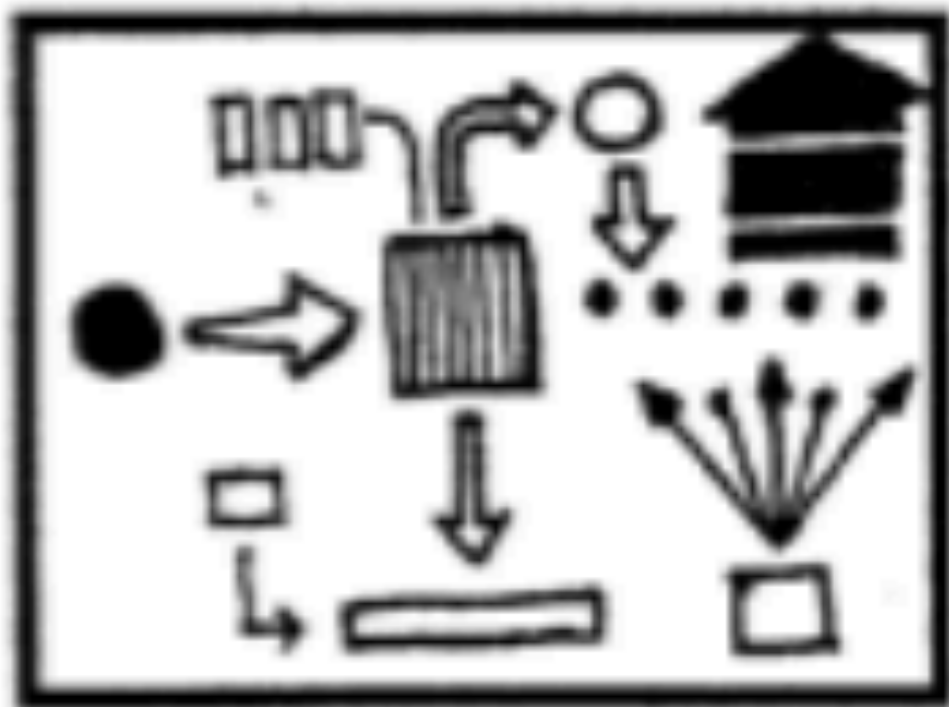
Magazine Style



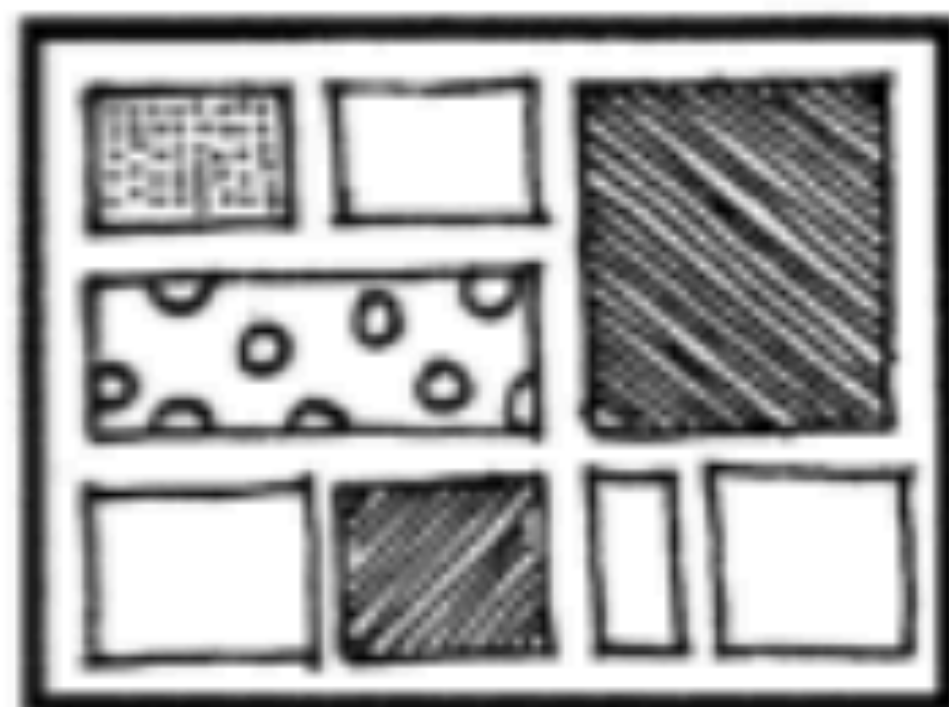
Annotated Chart



Partitioned Poster



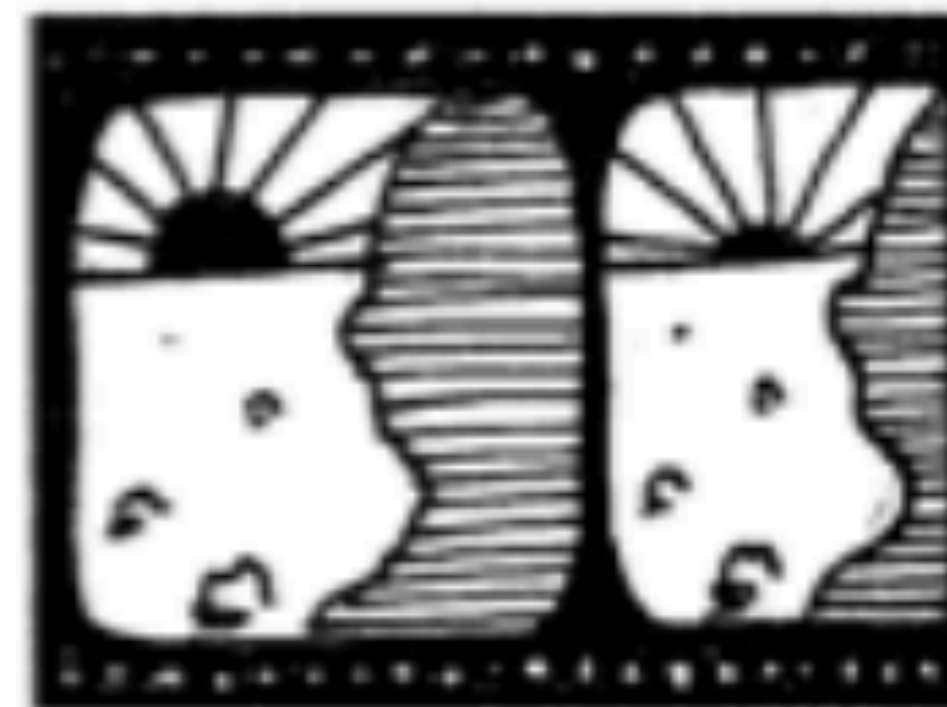
Flow Chart



Comic Strip



Slide Show



Film/Video/Animation

Mostly combined...

[Segel & Heer 2010]

Examples

Magazine

755

Steroids or Not, the Pursuit Is On

Barry Bonds is taking aim at the career home run record. He needs only six more to tie Babe Ruth and 47 to equal Hank Aaron.



Lines are cumulative home runs.

Hank Aaron
755 homers
23 seasons



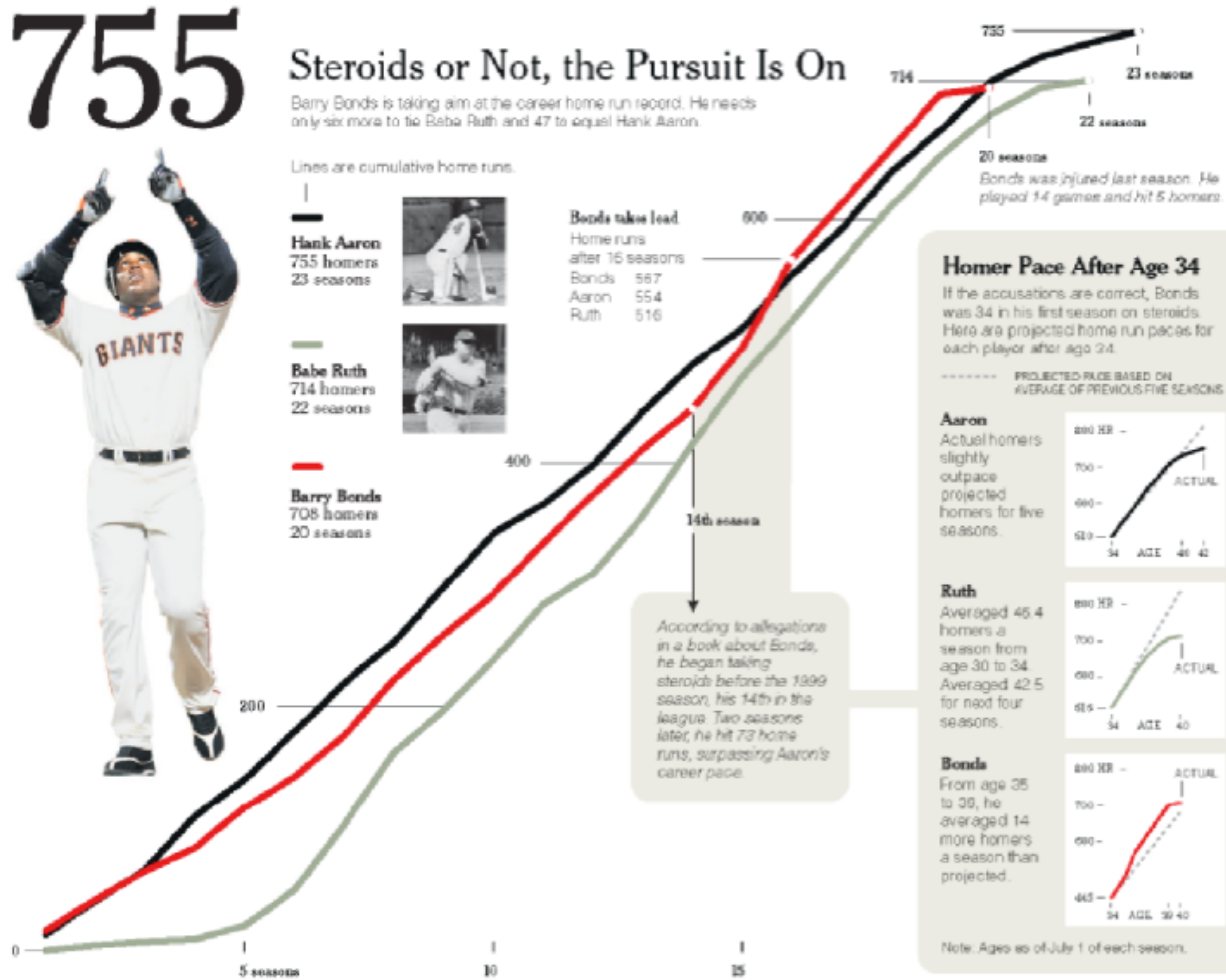
Babe Ruth
714 homers
22 seasons



Barry Bonds
708 homers
20 seasons

Bonds takes lead
Home runs
after 16 seasons
Bonds 567
Aaron 554
Ruth 516

Bonds was injured last season. He played 14 games and hit 5 homers.

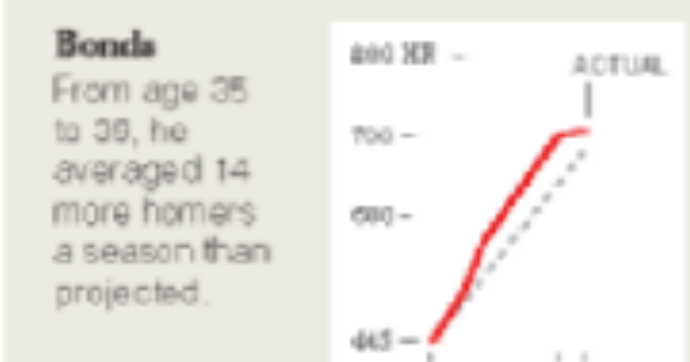
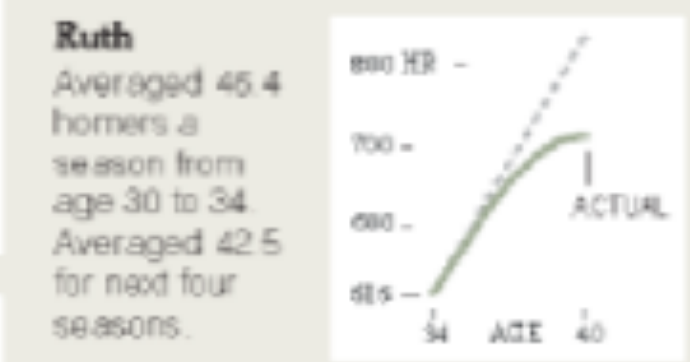
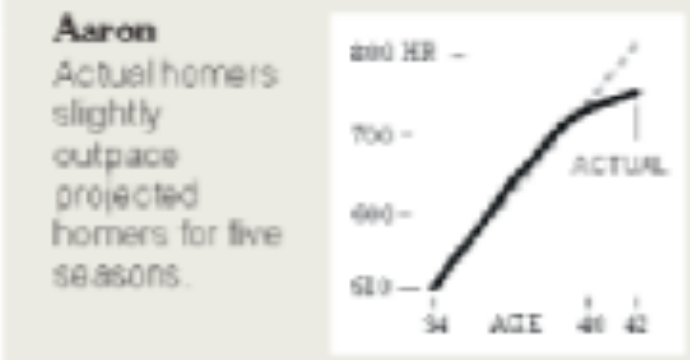


According to allegations in a book about Bonds, he began taking steroids before the 1999 season, his 14th in the league. Two seasons later, he hit 73 home runs, surpassing Aaron's career pace.

Homer Pace After Age 34

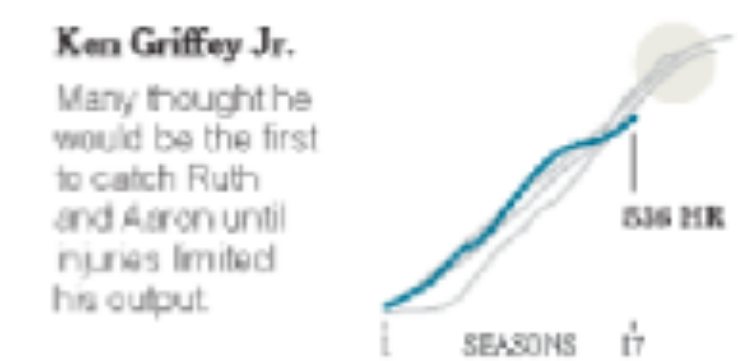
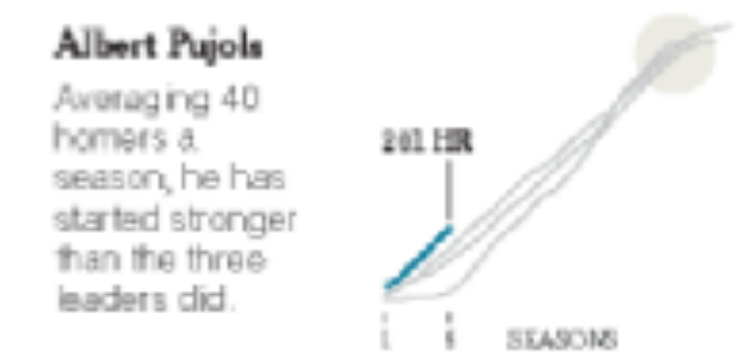
If the accusations are correct, Bonds was 34 in his first season on steroids. Here are projected home run paces for each player after age 34.

----- PROJECTED PACE BASED ON AVERAGE OF PREVIOUS FIVE SEASONS



Note: Ages as of July 1 of each season.

Others Taking Aim



Differing Paths to the Top of the Charts

The top seven players on the career home run list, along with a look at Griffey (12th), Rodriguez (37th) and Pujols (tied 257th).

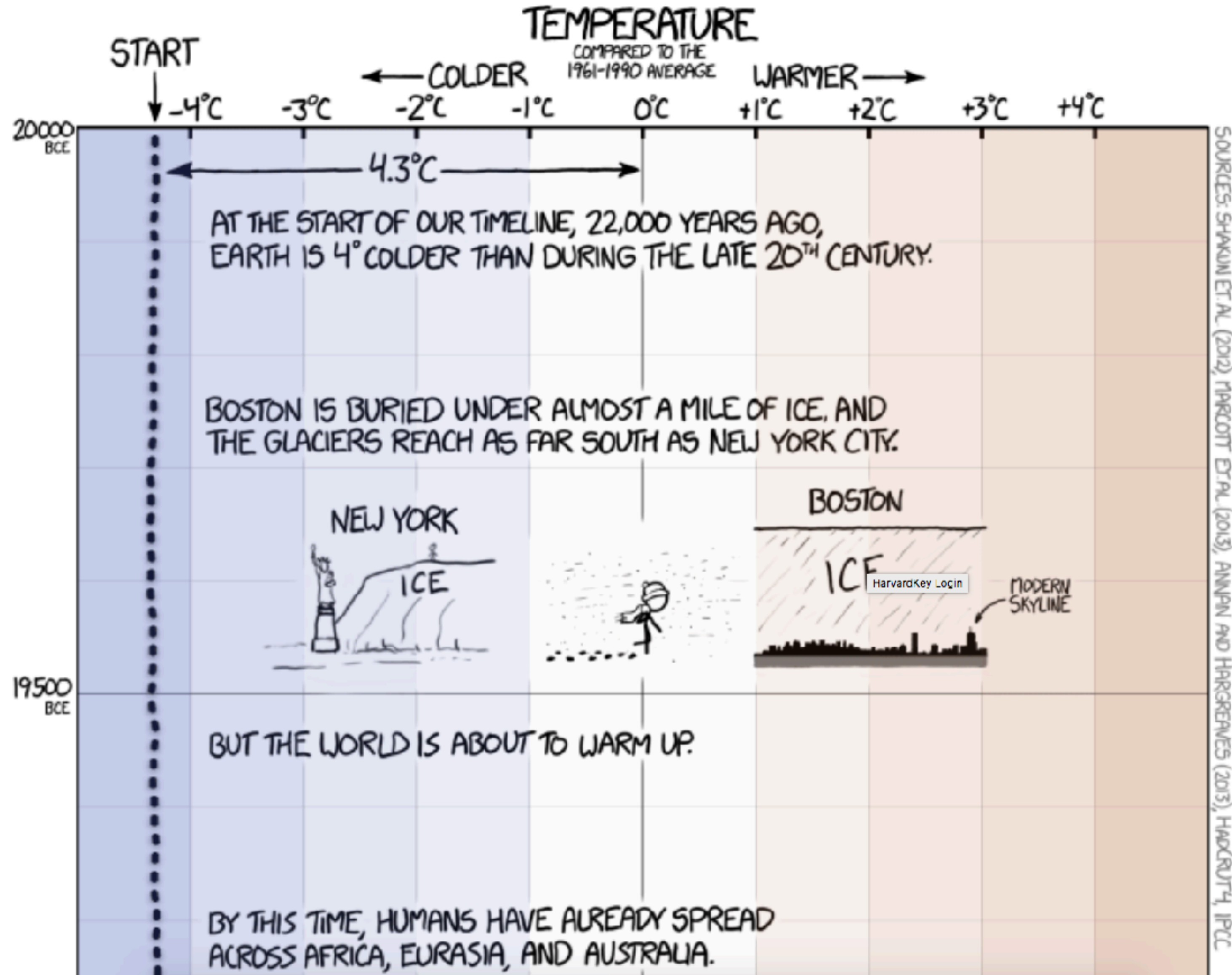


New York Times 2006

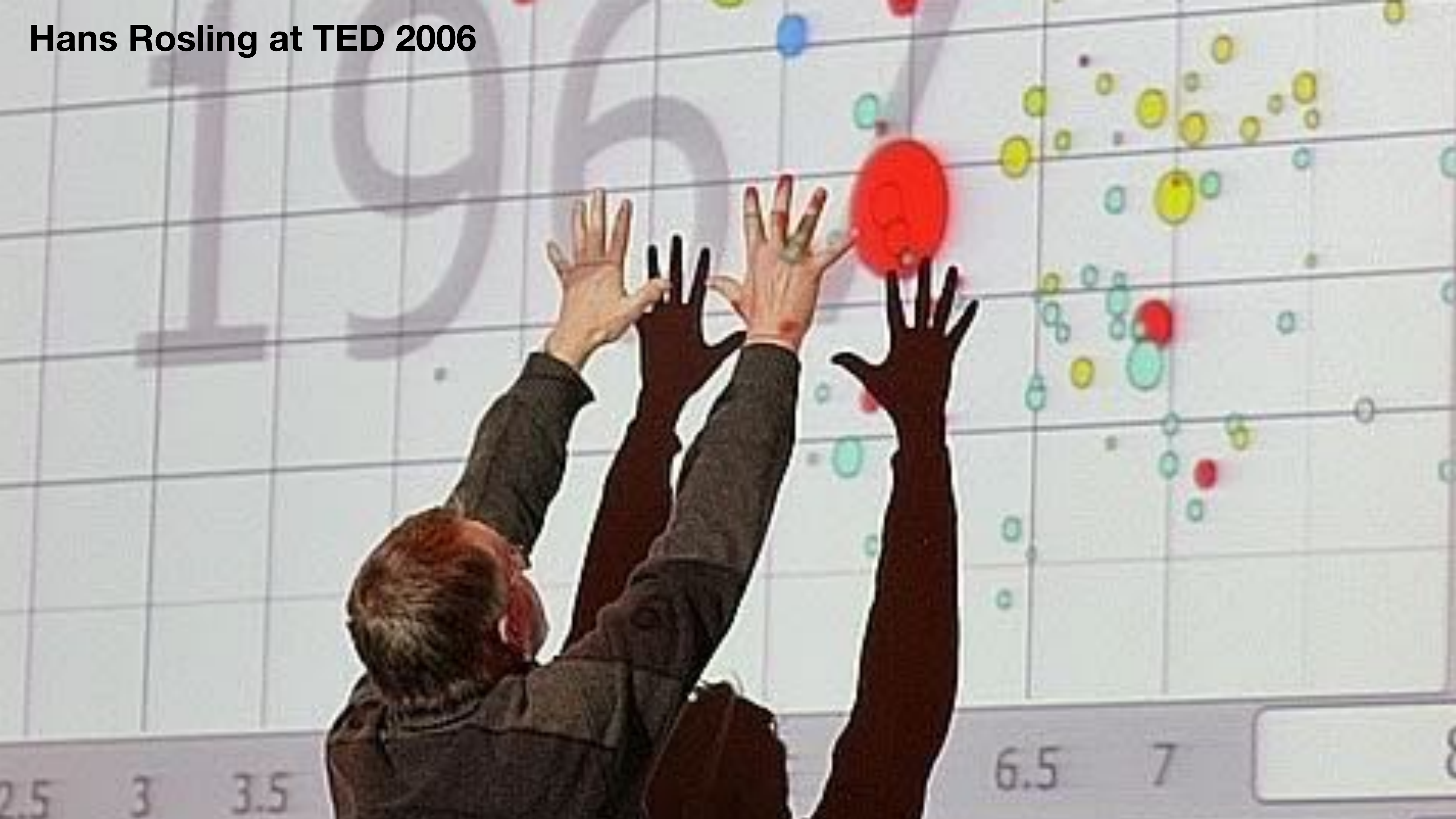
Comics

A TIMELINE OF EARTH'S AVERAGE TEMPERATURE SINCE THE LAST ICE AGE GLACIATION

WHEN PEOPLE SAY "THE CLIMATE HAS CHANGED BEFORE,"
THESE ARE THE KINDS OF CHANGES THEY'RE TALKING ABOUT.

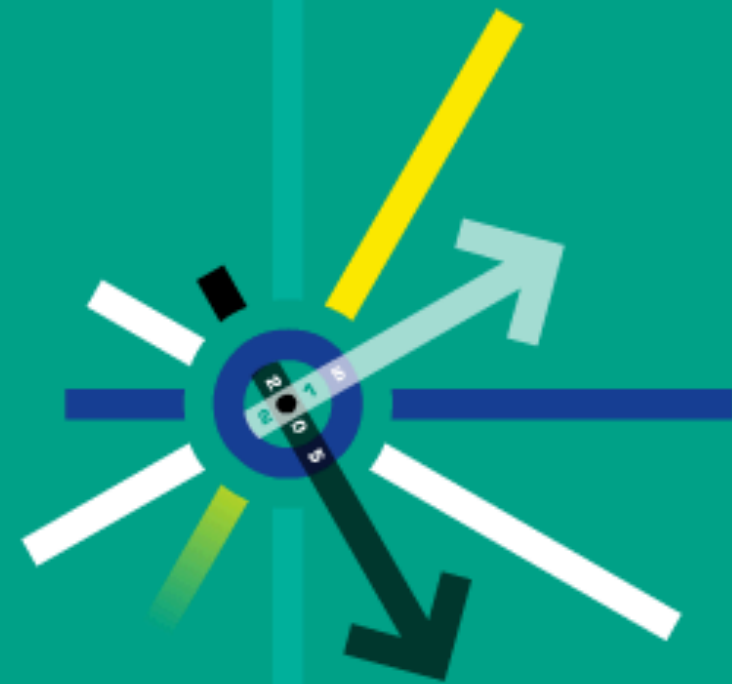


Hans Rosling at TED 2006



Human Development Trends 2005

Interactive presentation of some of the messages in the Human Development Report 2005

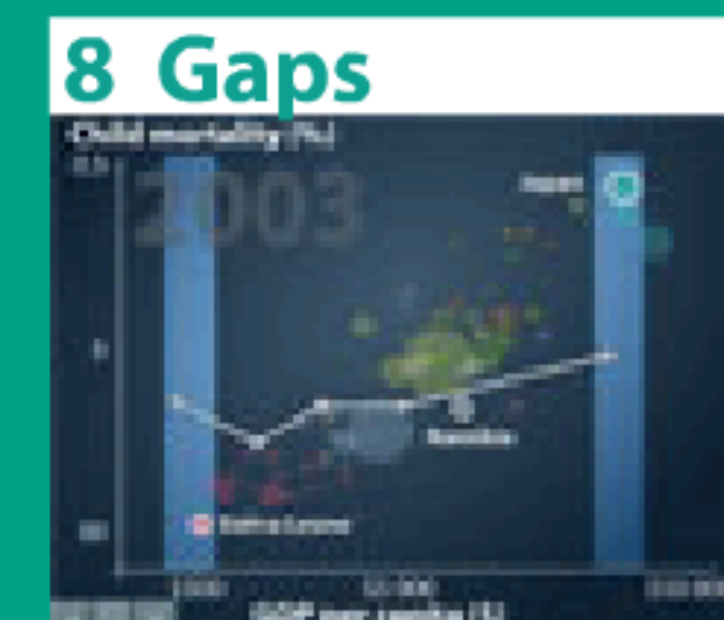
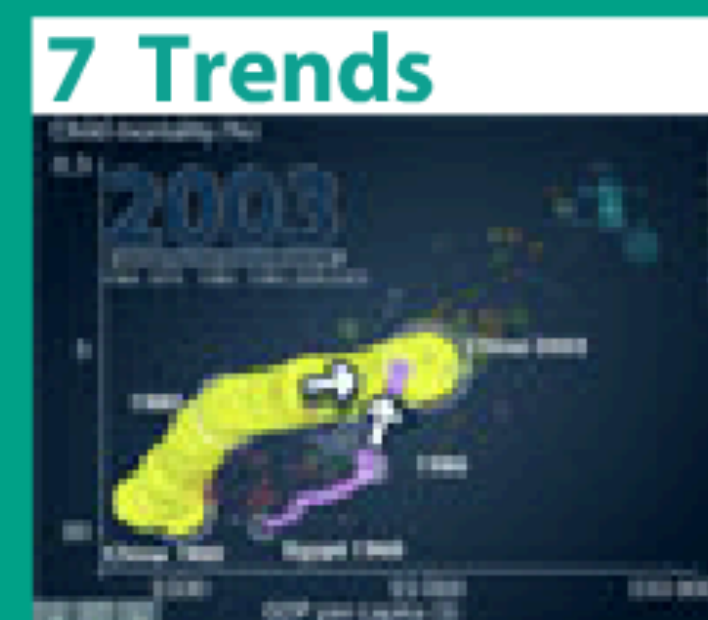
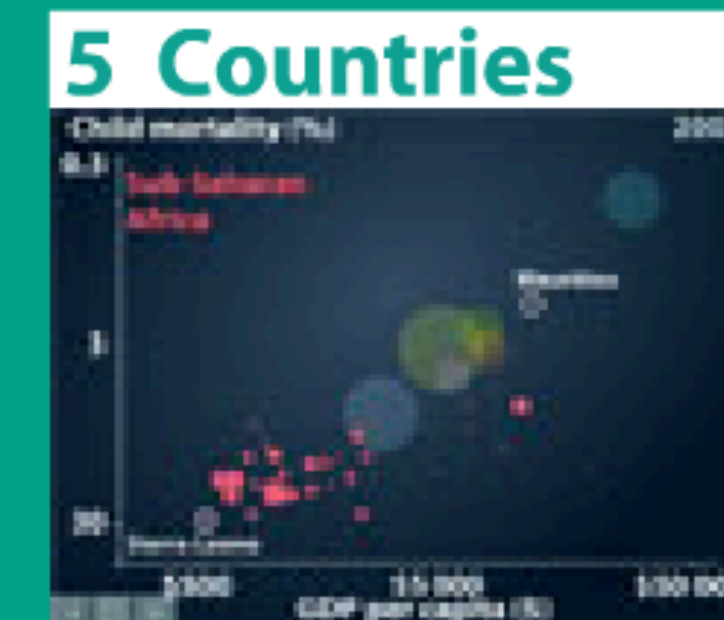
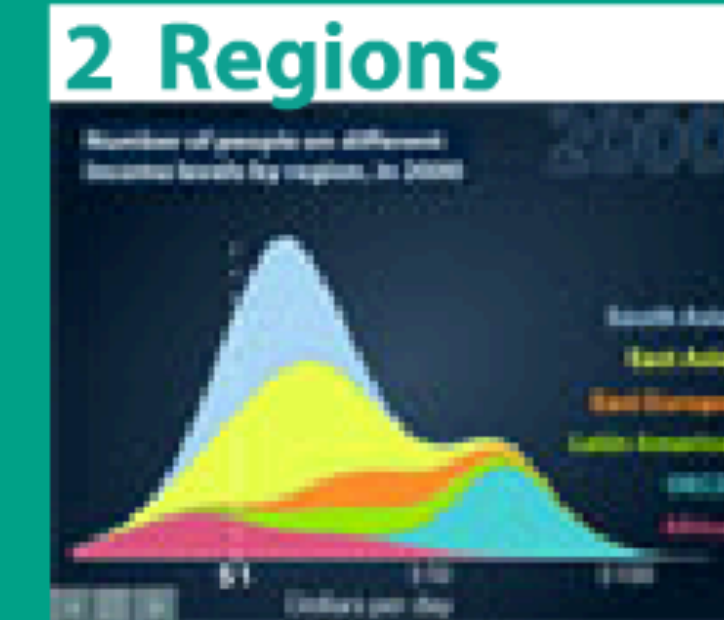


- English
- Dansk
- Portuguese
- Suomi
- Français
- Deutsch

Produced in collaboration with:

GAPMINDER
www.gapminder.org

English translation: Claes Johansson, UNDP



Budget Forecasts, Compared With Reality

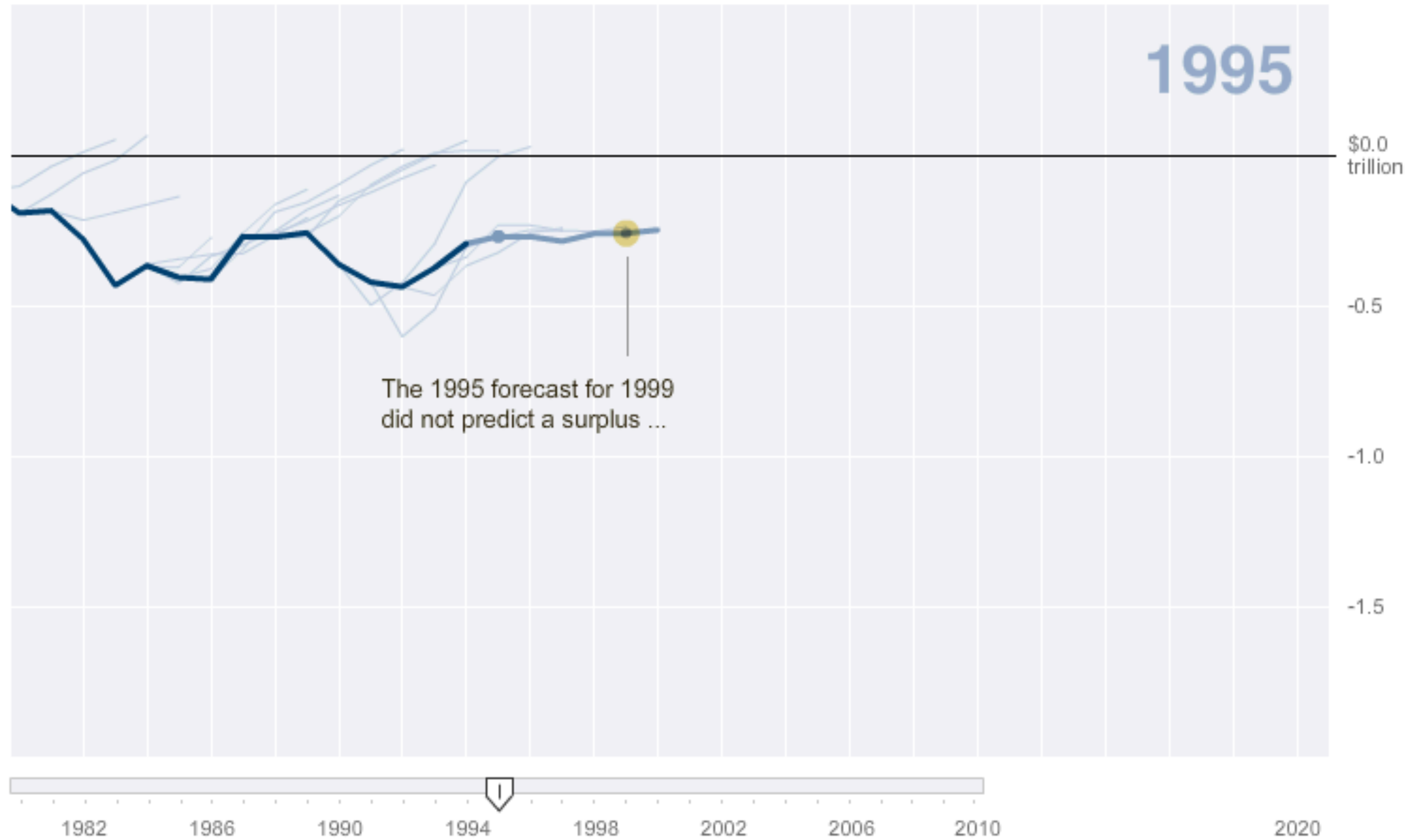
Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

1 2 3 4 5 6 NEXT ▶

Past forecasts

Even that may be an understatement. In the last 30 years, about 80 percent of four-year deficit forecasts have been too optimistic.

The early Clinton budgets — which failed to predict the surpluses that were generated, in part, by a stock market bubble — are the only major exception.



By AMANDA COX | [Send Feedback](#)

Source: Office of Management and Budget

[TWITTER](#) [LINKEDIN](#) [SHARE](#)

U.S. GUN DEATHS IN

2013 2010

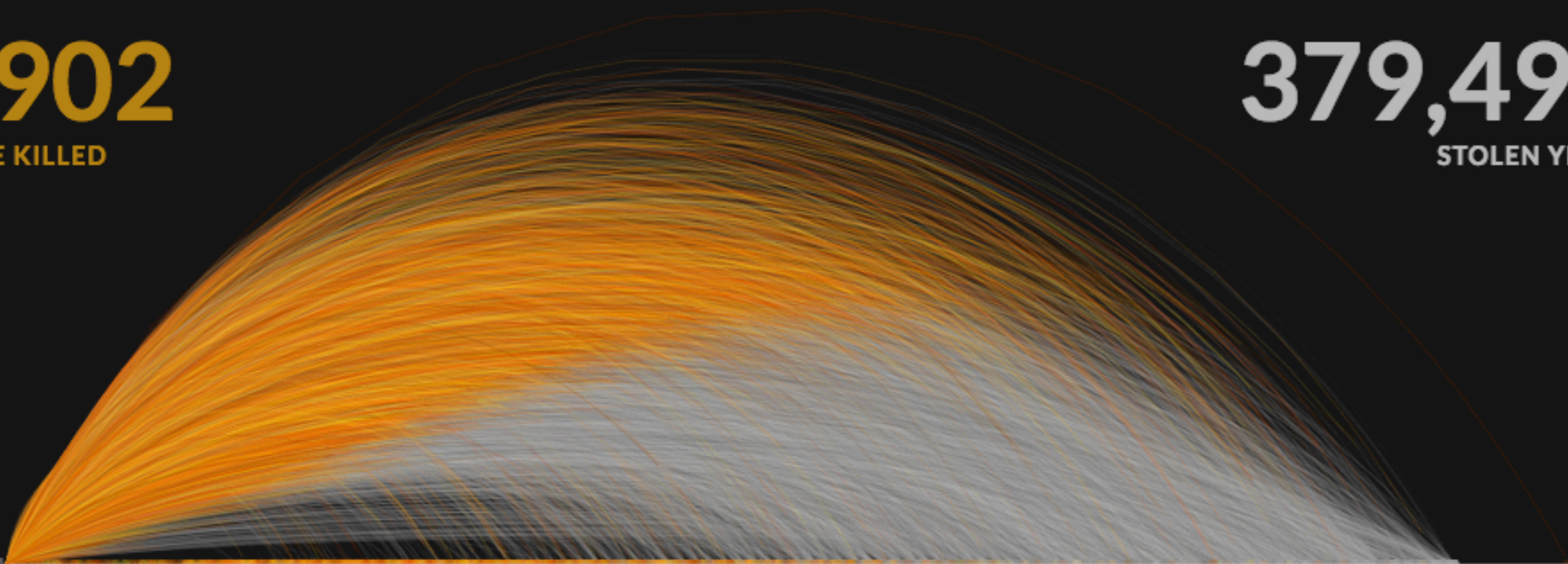
OCTOBER

8,902

PEOPLE KILLED

379,491

STOLEN YEARS



AGEO |

| 108

THE
Fallen
OF
World War II



Seven
Genres



Magazine Style

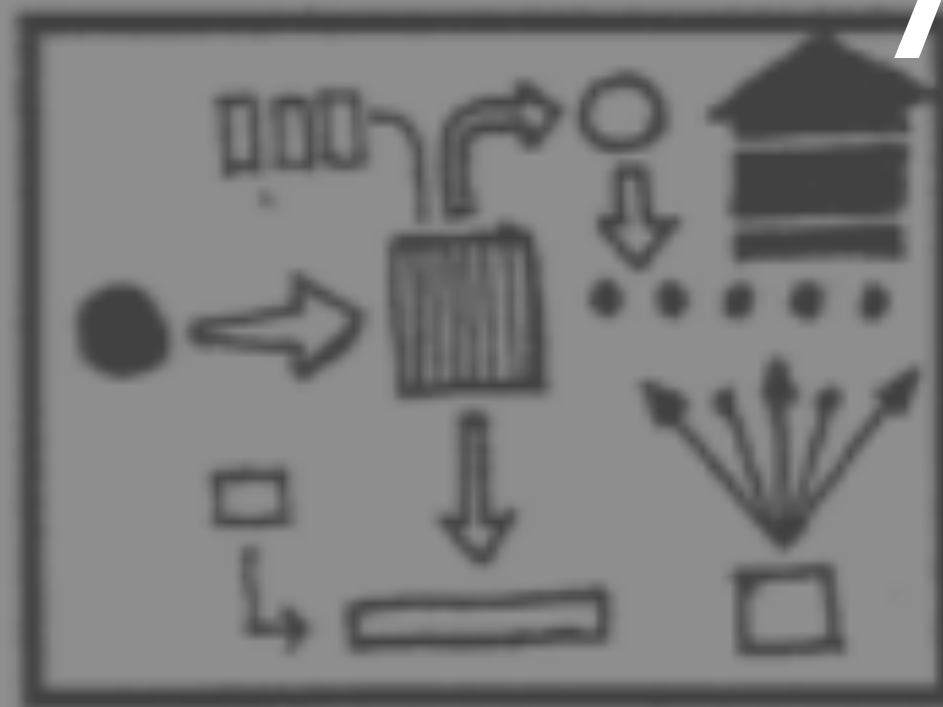


Annotated Chart



Partitioned Poster

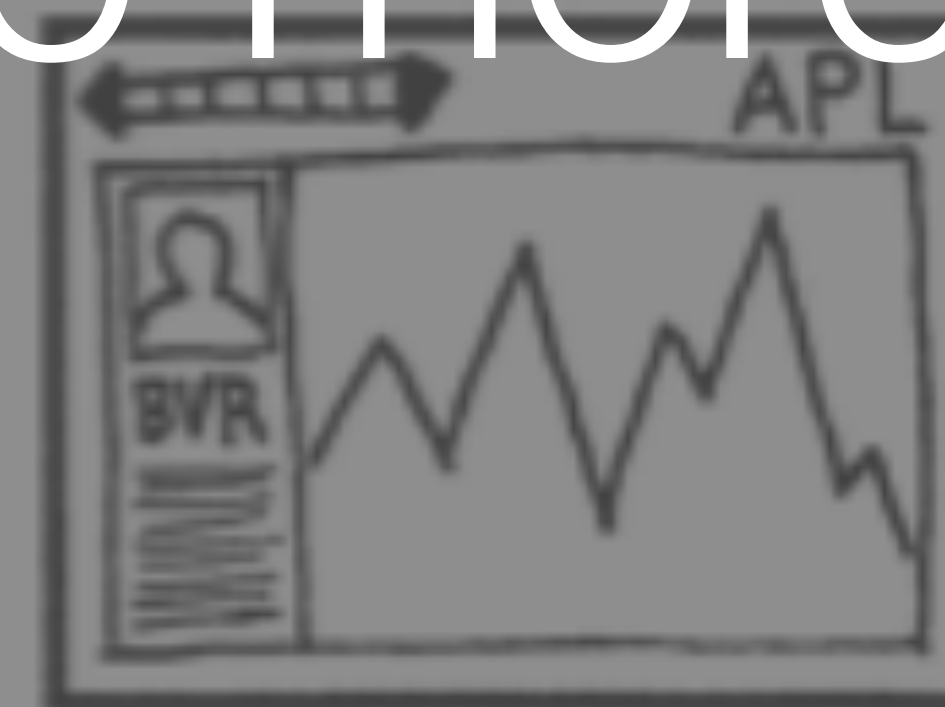
Are there more?



Flow Chart



Comic Strip



Slide Show



Film/Video/Animation

R2
D3

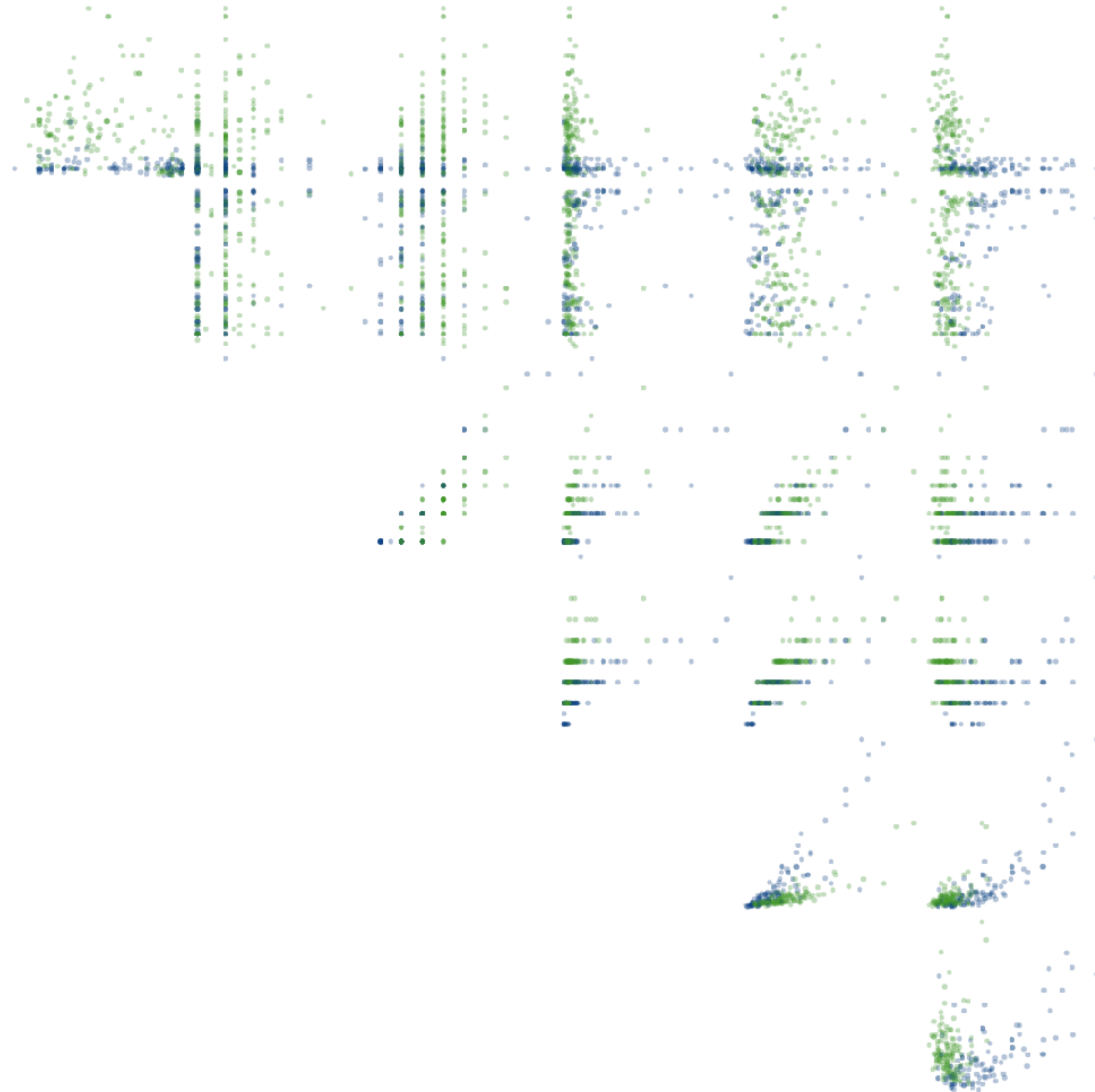
A visual introduction to machine learning

🌐 English

In machine learning, computers apply **statistical learning** techniques to automatically identify patterns in data. These techniques can be used to make highly accurate predictions.

Keep scrolling. Using a data set about homes, we will create a machine learning model to distinguish homes in New York from homes in San Francisco.

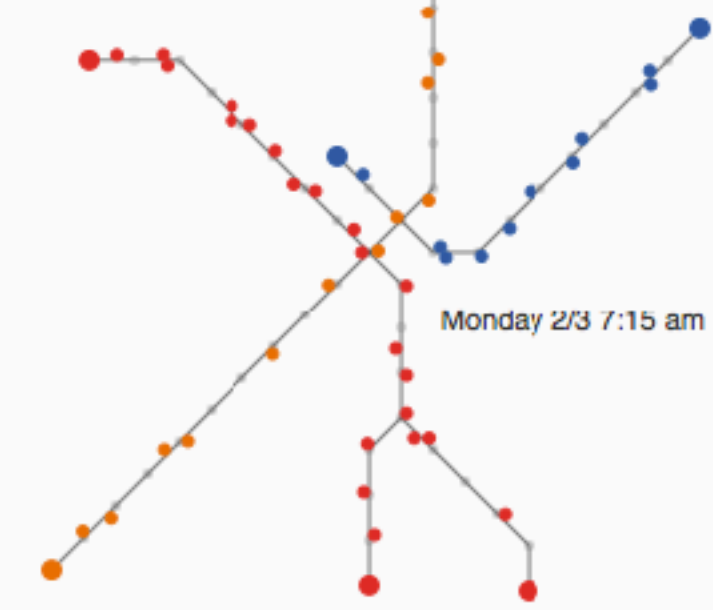
SCROLL



Visualizing MBTA Data

An interactive exploration of Boston's subway system

Mike Barry and Brian Card - June 10, 2014



Star 676 Share 6.9K Tweet G+ Share

Boston's Massachusetts Bay Transit Authority (MBTA) operates the 4th busiest subway system in the U.S. after New York, Washington, and Chicago. If you live in or around the city you have probably ridden on it. The MBTA recently began publishing substantial amount of subway data through its public APIs. They provide the full schedule in General Transit Feed Specification (GTFS) format which powers Google's transit directions. They also publish realtime train locations for the Red, Orange, Blue, and Green lines. The following visualizations use data captured from these feeds for the entire month of February, 2014. [Green Line data became available](#) in October, 2014 so is not shown here. Also, working with the MBTA, we were able to acquire per-minute entry and exit counts at each station measured at the turnstiles used for payment.

We attempt to present this information to help people in Boston better understand the trains, how people use the trains, and how the people and trains interact with each other.

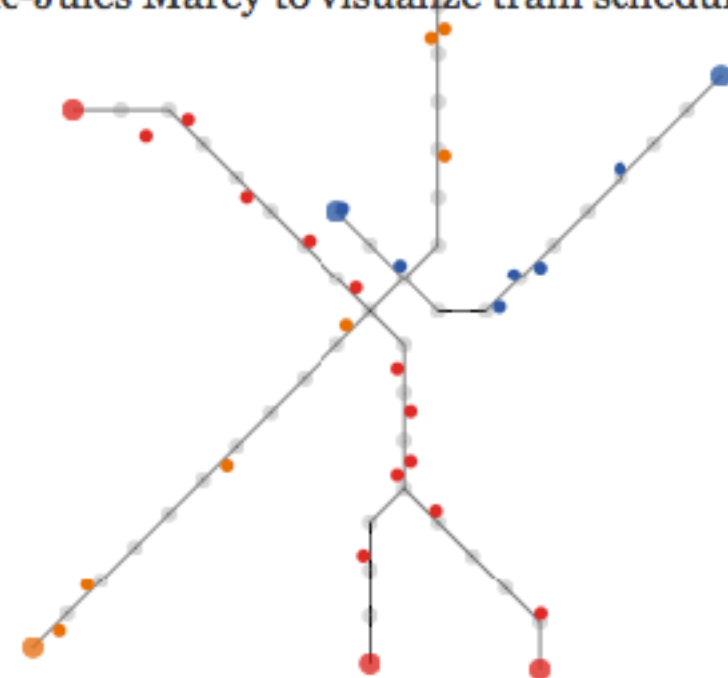
The Trains

In a typical weekday, trains make approximately 1150 trips on the red, orange, and blue lines starting at 5AM and continuing through 1AM the next morning. On Saturdays trains make 870 trips and on Sundays they make 760.

To better understand how the trains operate on a typical day, below are all trips that trains took on the red, orange, and blue lines on Monday February 3 2014. Each vertical line represents a station, and time extends from top to bottom. Steeper lines indicate slower trains. This visualization was first used by Étienne-Jules Marey to visualize train schedules and is typically called a "Marey Diagram."

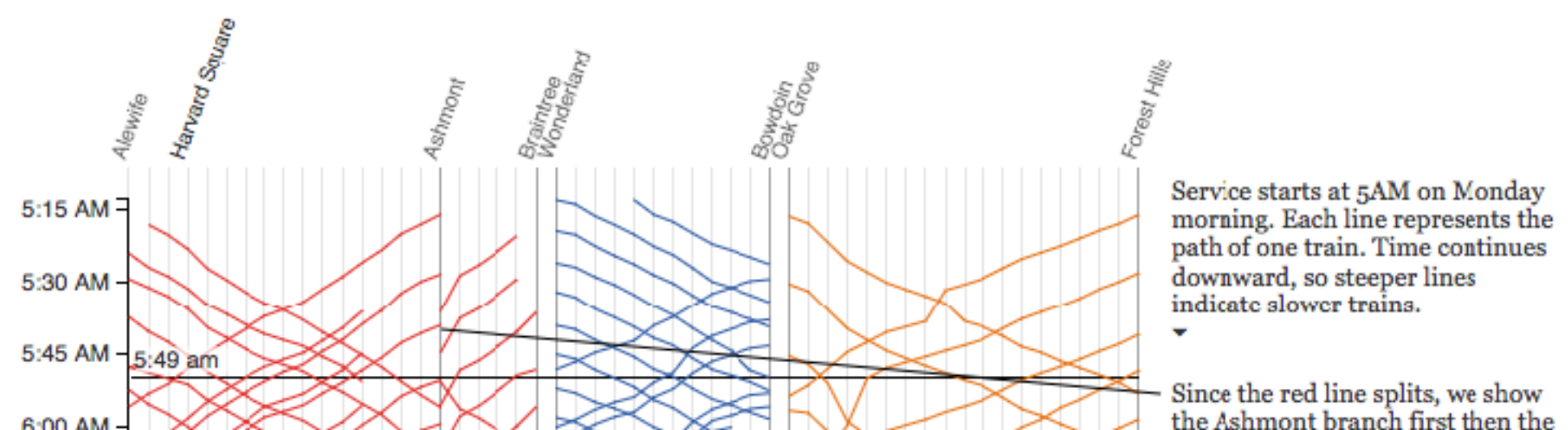
Average Number of Trips per Day

	Weekdays	Saturdays	Sundays
Red	450	350	300
Orange	320	260	220
Blue	380	260	240
Total	1150	870	760



Locations of each train on the red, blue, and orange lines at 5:49

Subway Trips on Monday February 3, 2014

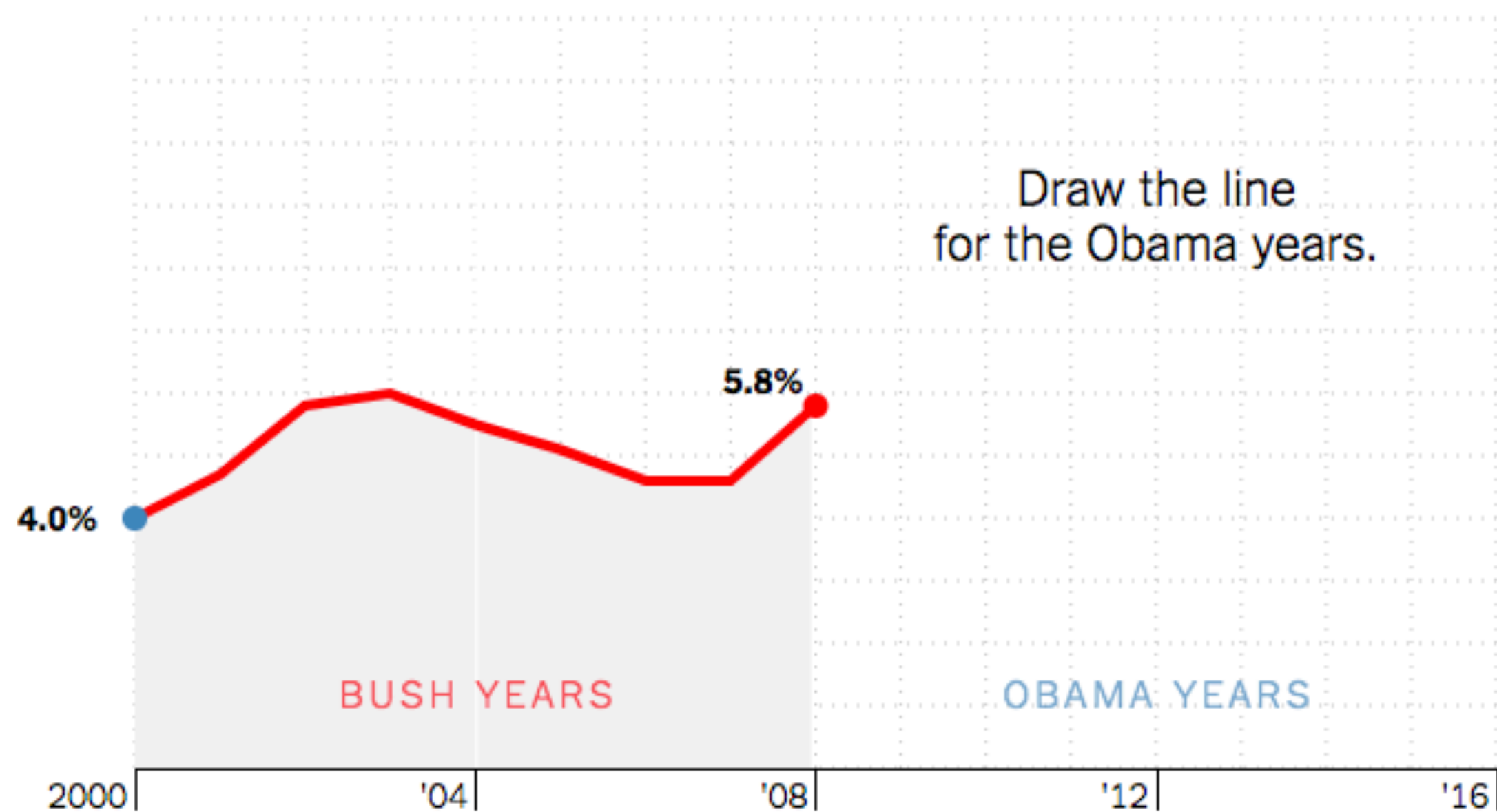


You Draw It: What Got Better or Worse During Obama's Presidency

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** ...



Show me how I did.

Salubrious Nation: a game-y look at U.S. health ?



Level (1 / 8)

Adult Smoking
(Percentage rate)

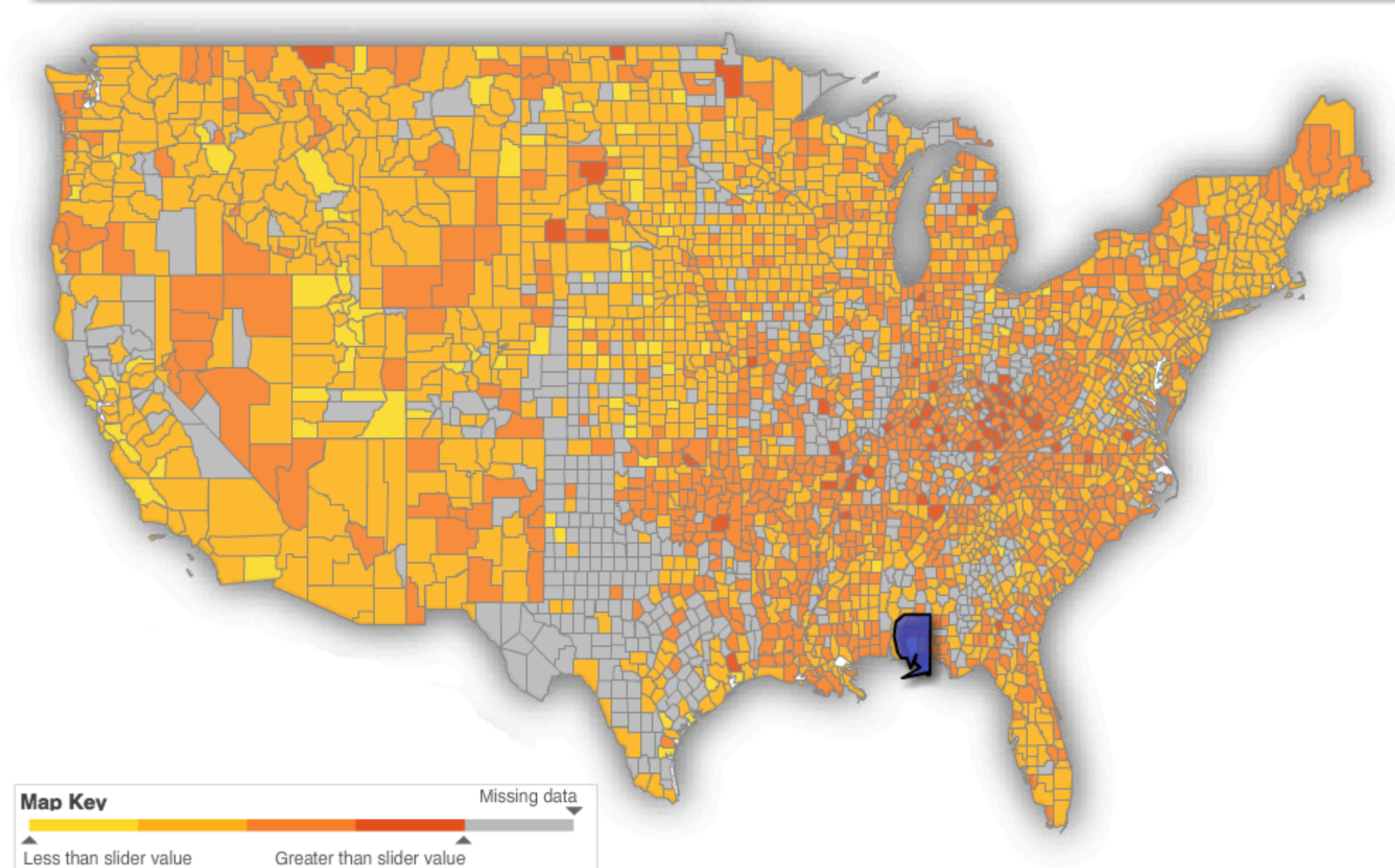
Guess the health parameters for **Santa Rosa, FL.**
Hover your mouse over a county for demographic hints.

Guess:

Submit

Your Score: 0

High Score: 86





digital **dreams**

0 minutes
49 seconds

X 547
Y 320ç

0

0



START RIDE



Epoch
000,000

Learning rate
0.03

Activation
Tanh

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

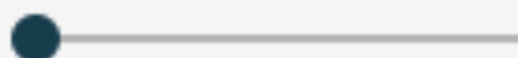
Which dataset do you want to use?



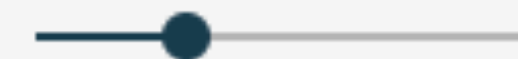
Ratio of training to test data: 50%



Noise: 0



Batch size: 10



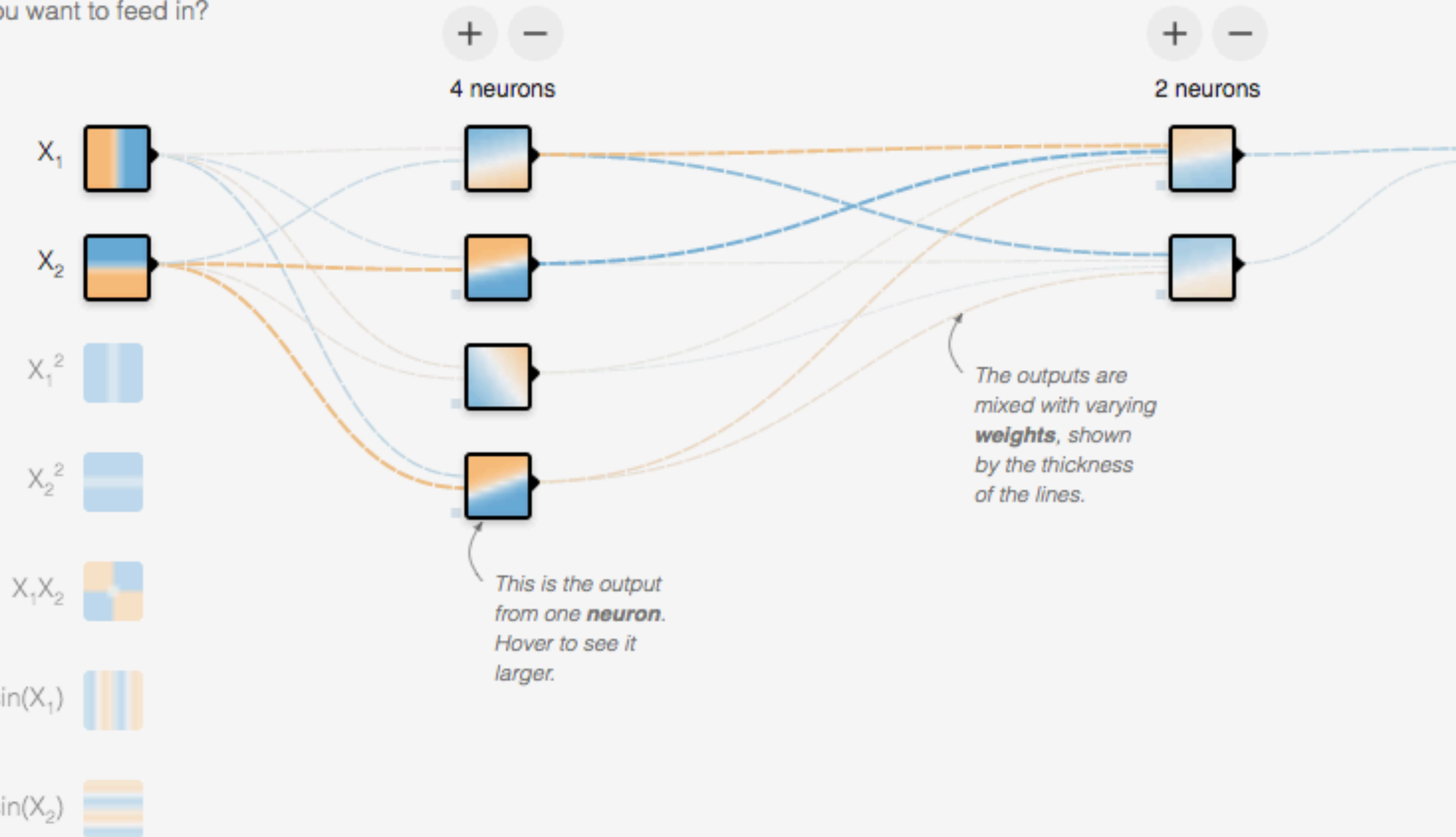
REGENERATE

FEATURES

Which properties do you want to feed in?

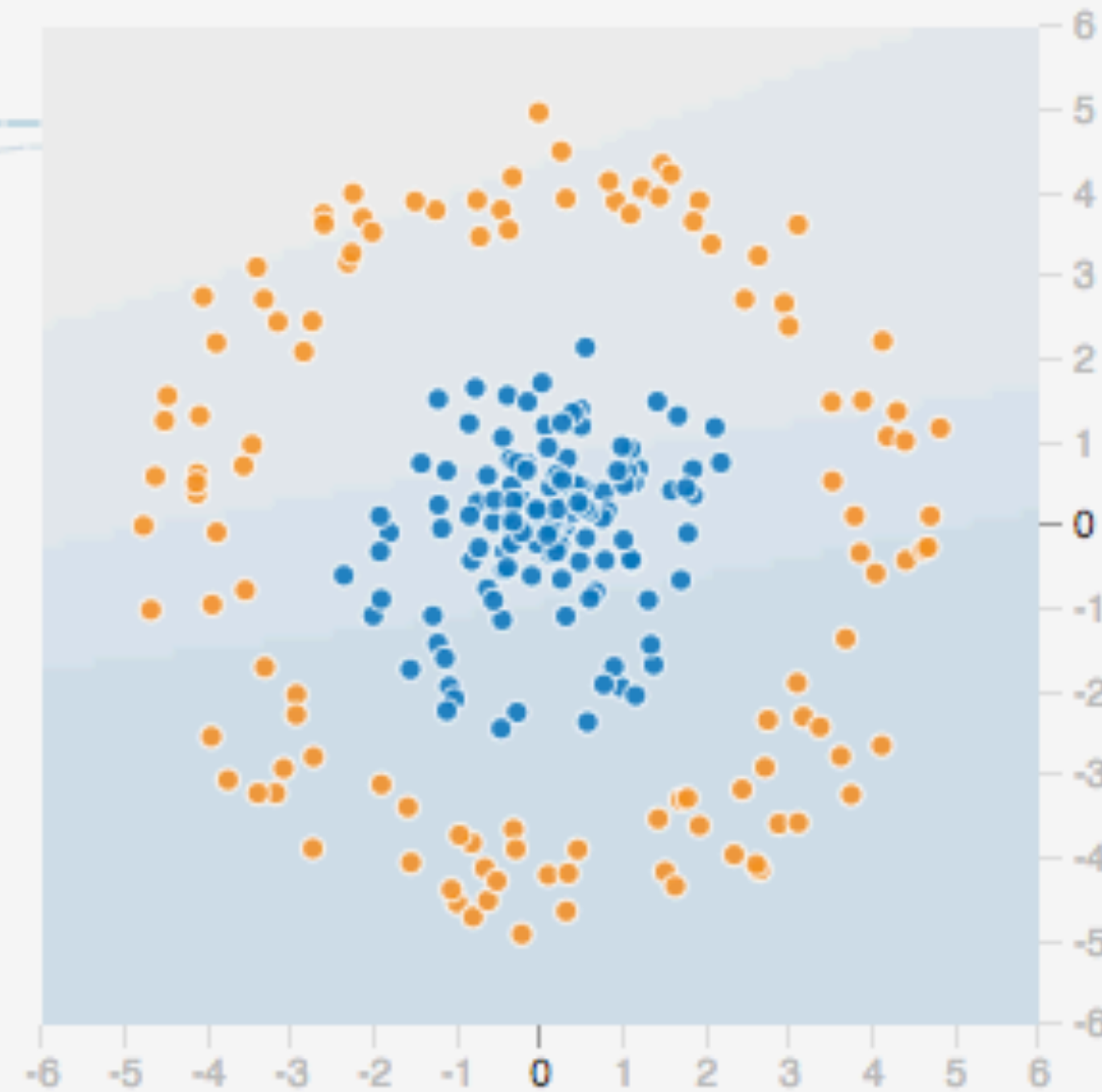
- X_1
- X_2
- X_1^2
- X_2^2
- $X_1 X_2$
- $\sin(X_1)$
- $\sin(X_2)$

+ - 2 HIDDEN LAYERS



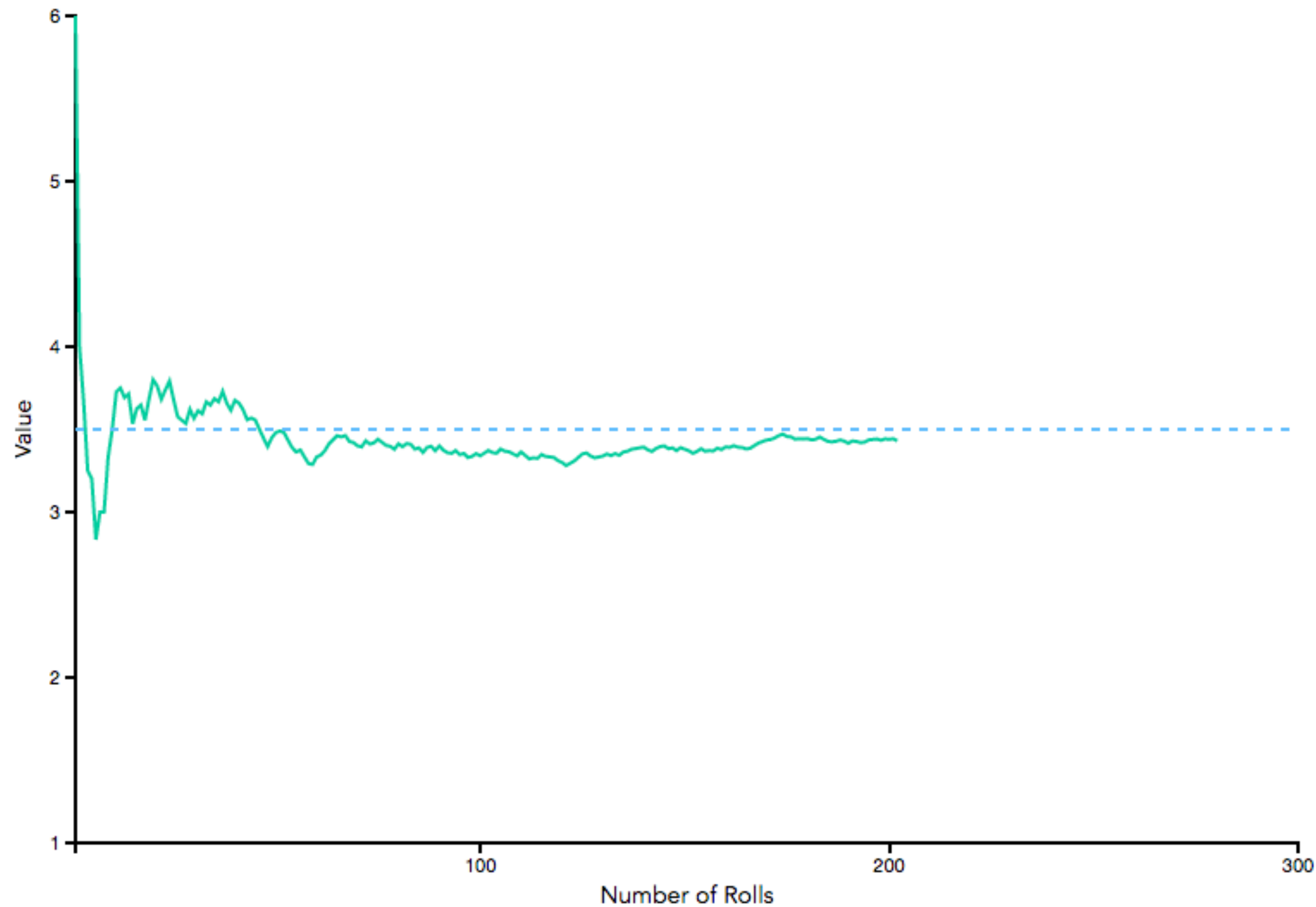
OUTPUT

Test loss 0.513
Training loss 0.500



Colors shows data, neuron and weight values.

Show test data Discretize output



Expectation

The expectation is a statistic that attempts to summarize the average of a random experiment. It can be interpreted as the long term average outcome of repeating the experiment many times. Mathematically, it is defined as the probability weighted sum of all outcomes in the sample space, denoted,

$$E[X] = \sum_{x \in \Omega} xP(x)$$

Consider the random experiment of rolling a fair die and watch as the running average converges to the expectation of 3.5.



Roll the Die

Roll 100 times

Make the die unfair by adjusting the blue bars below and observe how this changes the expectation.



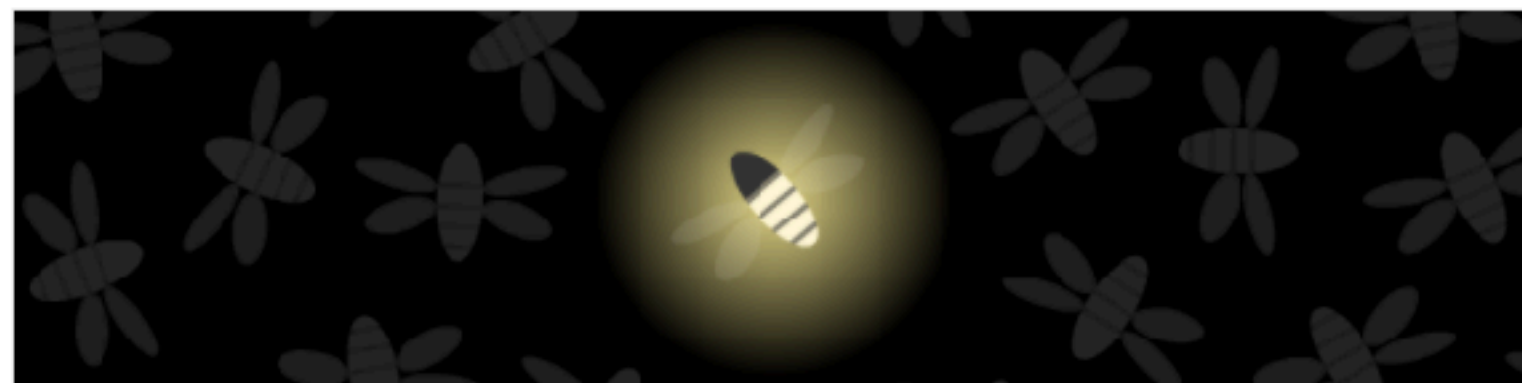
EXPLORABLE EXPLANATIONS

Lion cubs play-fight to learn hunting skills. Rats play to learn social & emotional skills. Monkeys play to learn cognitive skills, to practice problem-solving and creativity.

And yet, in the last century, we humans have convinced ourselves that play is useless, and learning is *supposed* to be boring. Gosh, no wonder we're all so miserable.

Welcome to **Explorable Explanations**, a hub for learning through play! We're a disorganized "movement" of artists, coders & educators who want to *reunite play and learning*.

Let's get started! Check out these **3 random Explorables**:



★ **Fireflies** biology

a small, serene simulation of self-synchronization

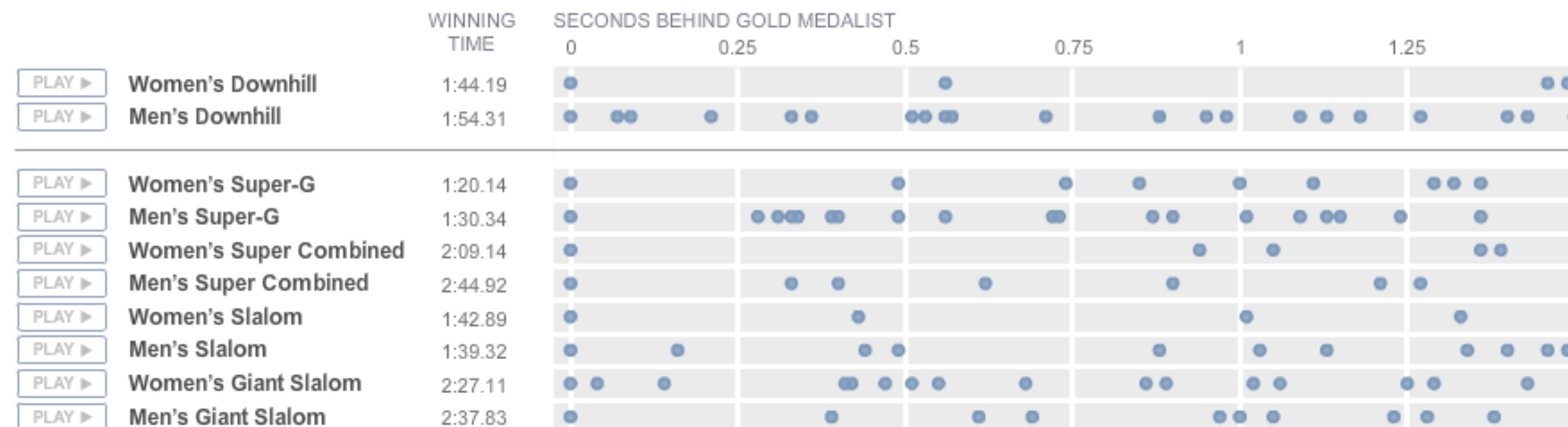
Fractions of a Second: An Olympic Musical

At the Olympics, the blink of an eye can be all that separates the gold medalist from the 10th-place finisher. In some events, this is obvious. But in others, with athletes racing one by one, the closeness of the race is harder to perceive. Listen to the differences below.



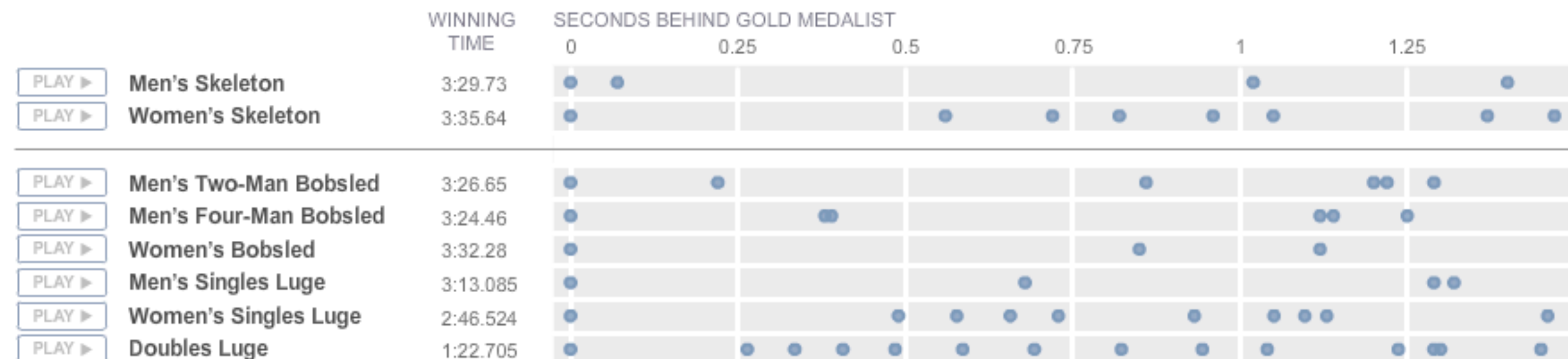
Alpine skiing

The women's downhill course was extremely tiring, and, because it was more challenging than the men's course, it ended up separating the skiers by much larger margins. This pattern appears in the two speed events: the downhill and the super-G.



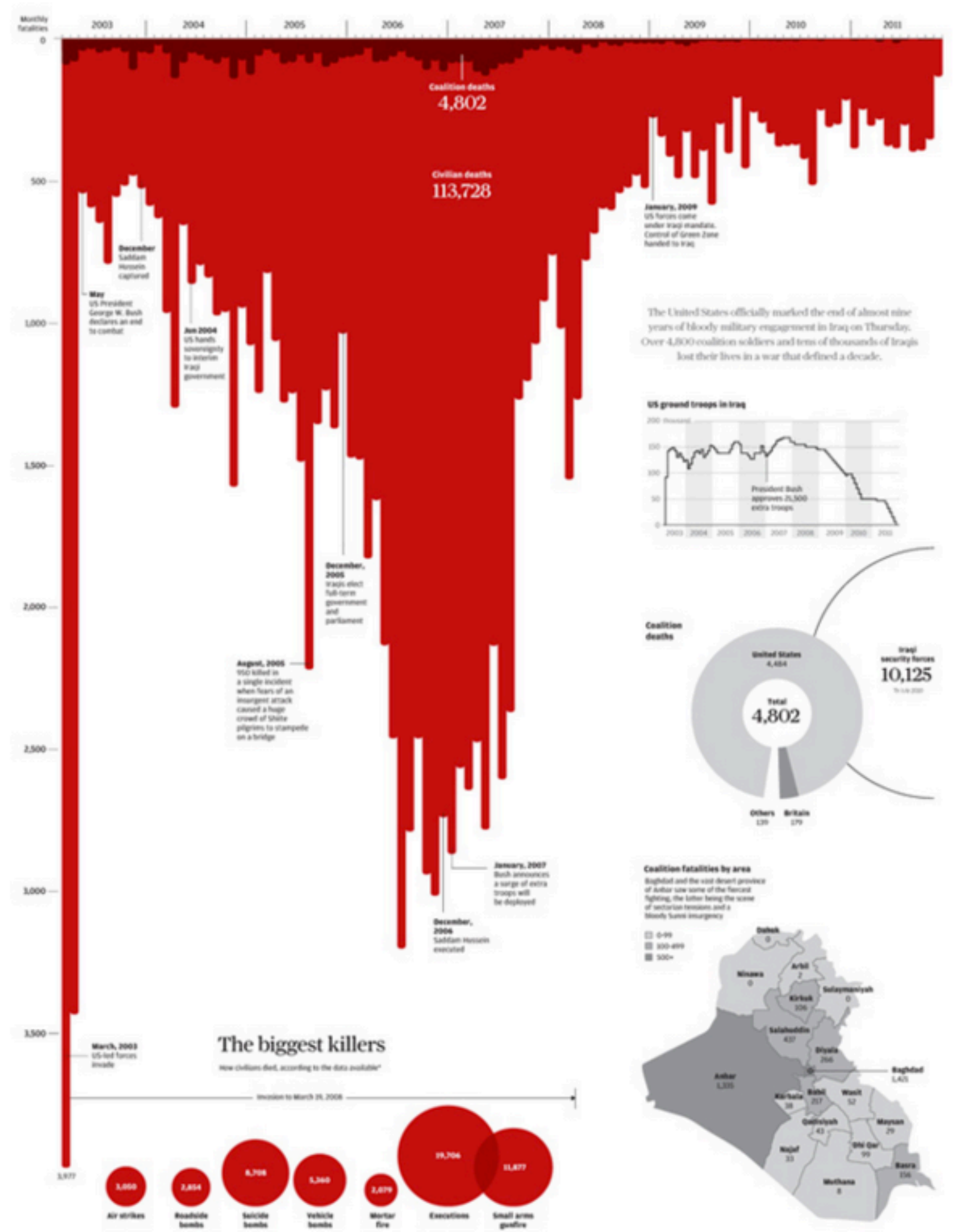
Skeleton, Bobsled and Luge

In percentage terms, the men's skeleton had one of the tightest finishes in Vancouver, with only .07 of a second separating the top two finishers across a three-and-a-half-minute run. But the difficult track produced speeds higher than expected, and many sliding events had relatively large gaps between gold and silver.



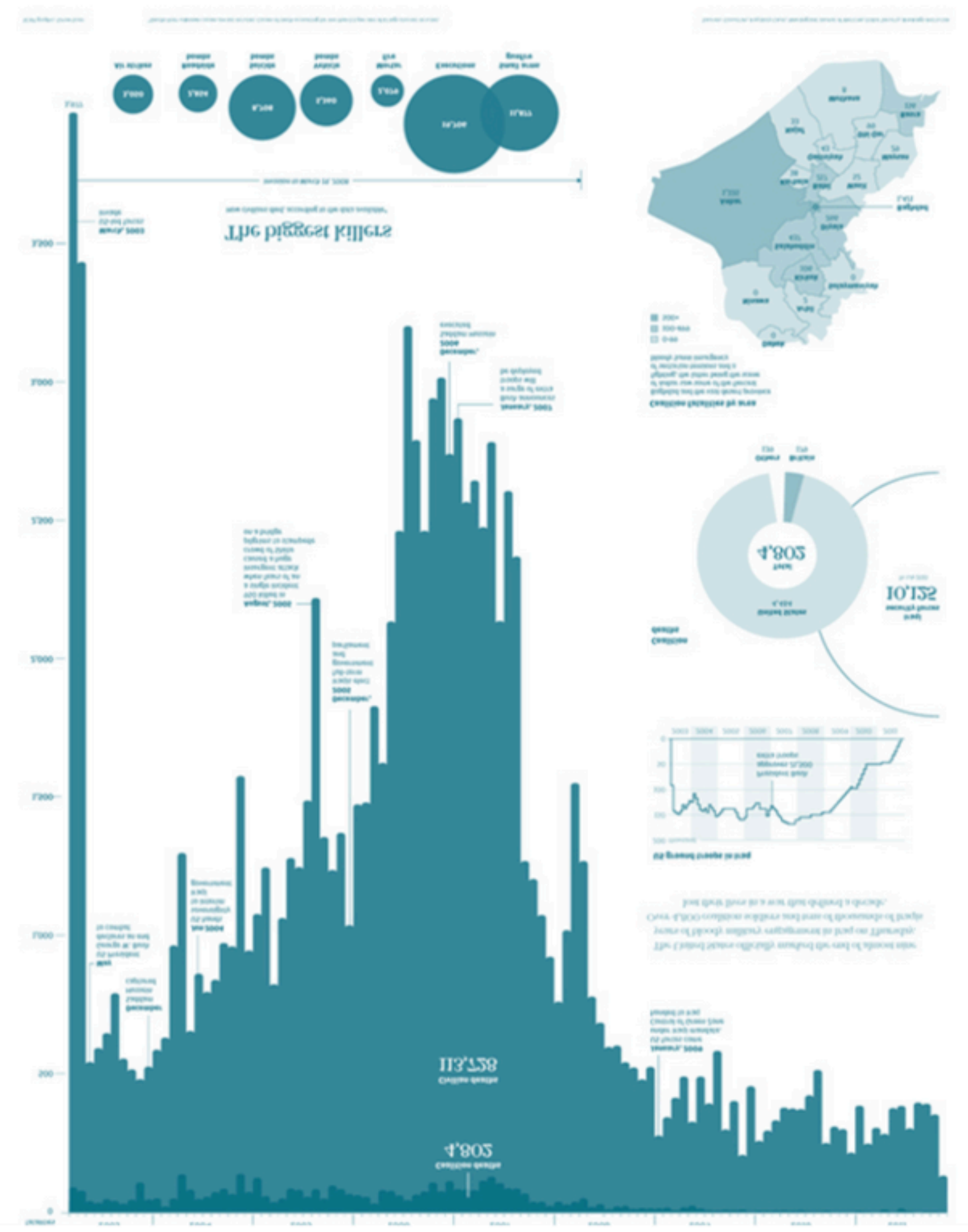
Storytelling, Double-edged Sword?

Iraq's bloody toll



[South China Morning Post 2011]

Iraq: Deaths on the decline

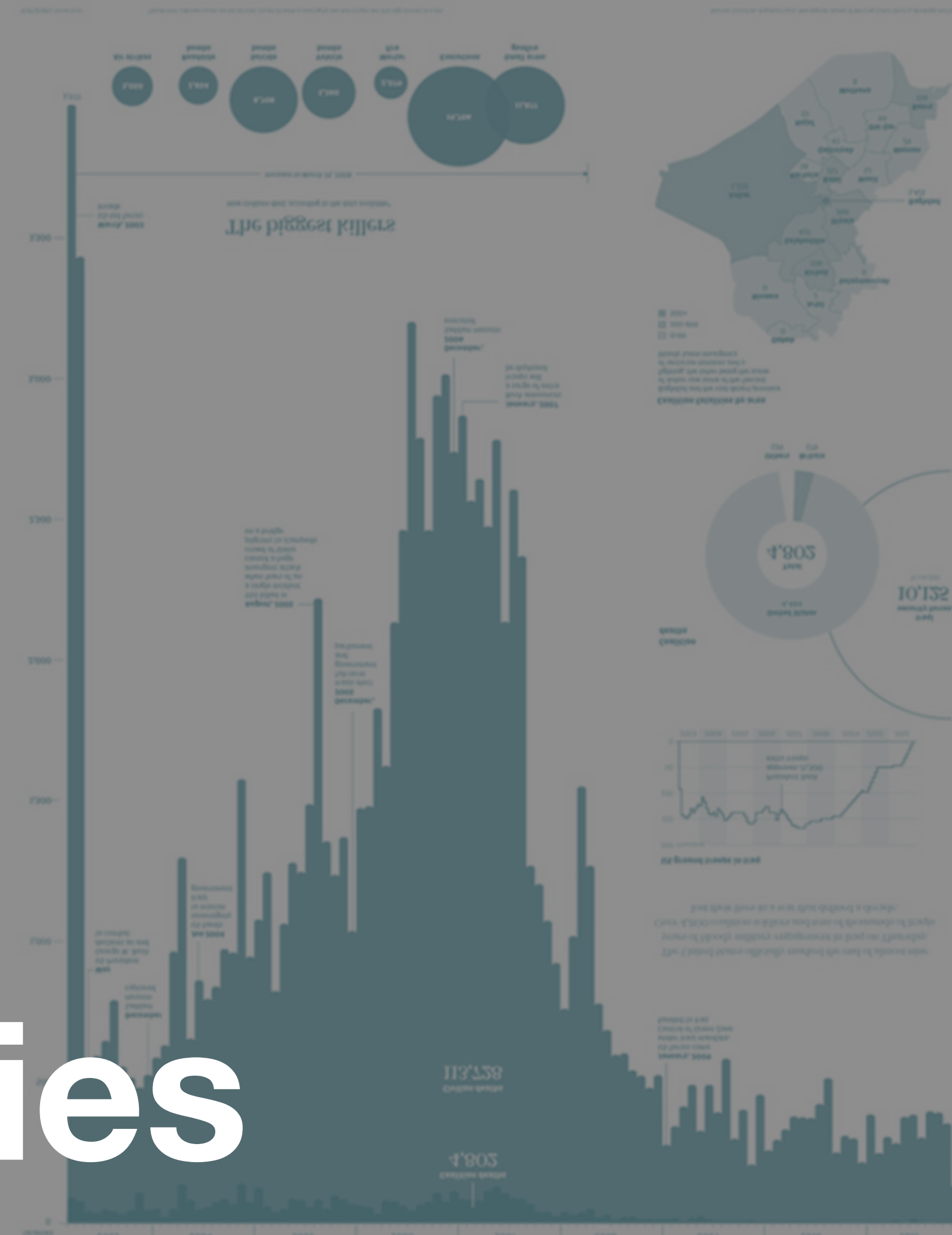


Flipped upside down....

Iraq's bloody toll



Iraq: Deaths on the decline



Same Data, Different Stories

[South China Morning Post 2011]

Flipped upside down....

Presentation & Storytelling in Tableau

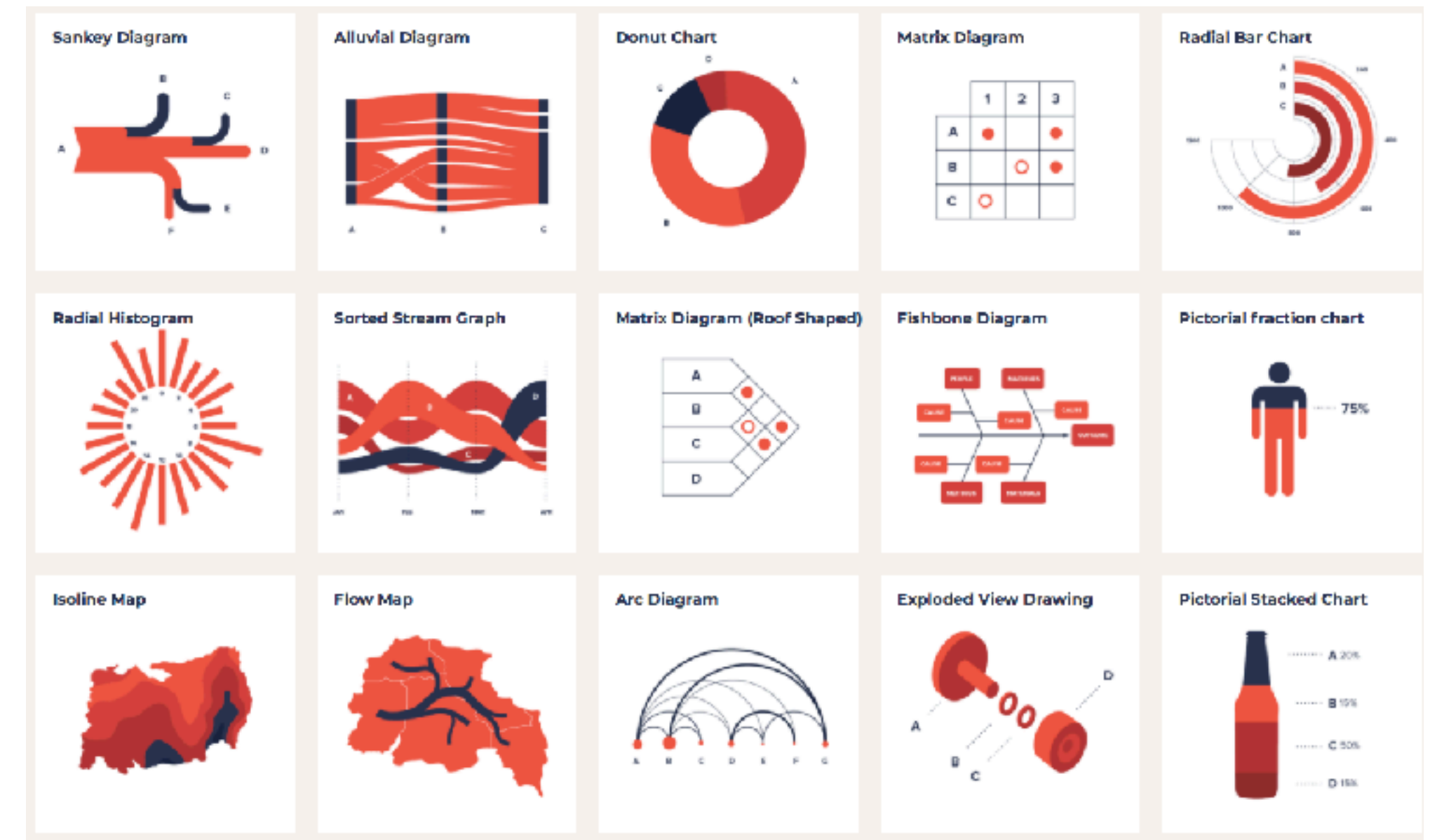
Dashboard

Annotation

Story Points

Next

Advanced visualizations



10 min break