

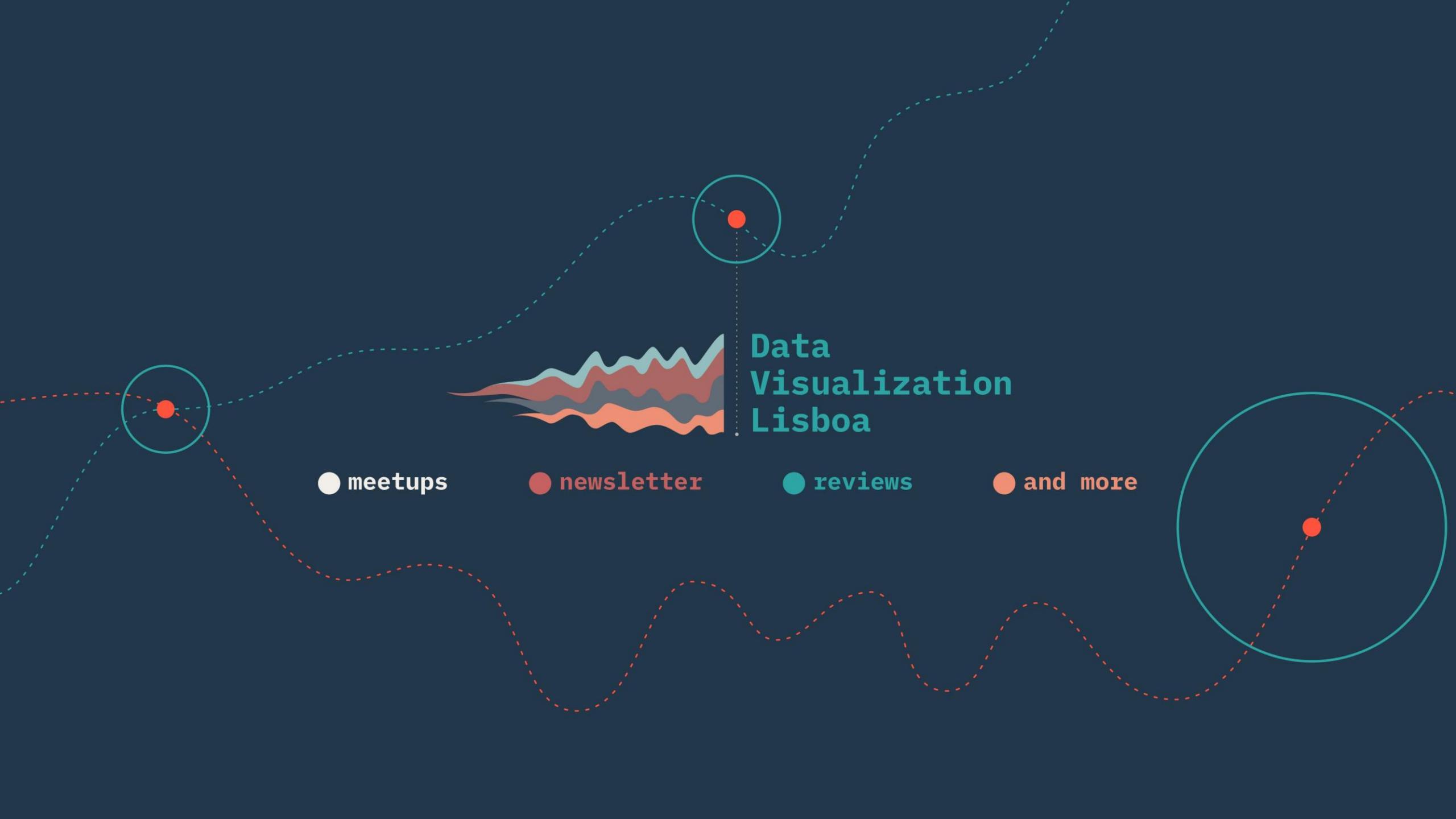
Visualização de Dados em Saúde Pública



Data
Visualization
Lisboa



AEENSP NOVA
ASSOCIAÇÃO DE ESTUDANTES
DA ESCOLA NACIONAL DE SAÚDE PÚBLICA



Data Visualization Lisboa

● meetups

● newsletter

● reviews

● and more



Sara Mesquita
Co-fundadora



Caroline Doyé
Co-fundadora



Renata Steffen



Salomé Esteves



Novo membro



Novo membro

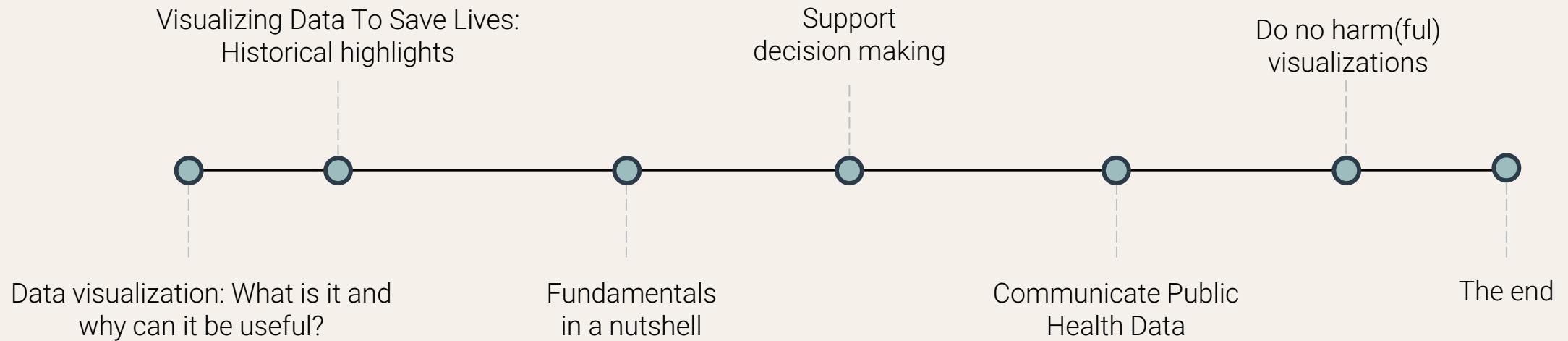


Novo membro

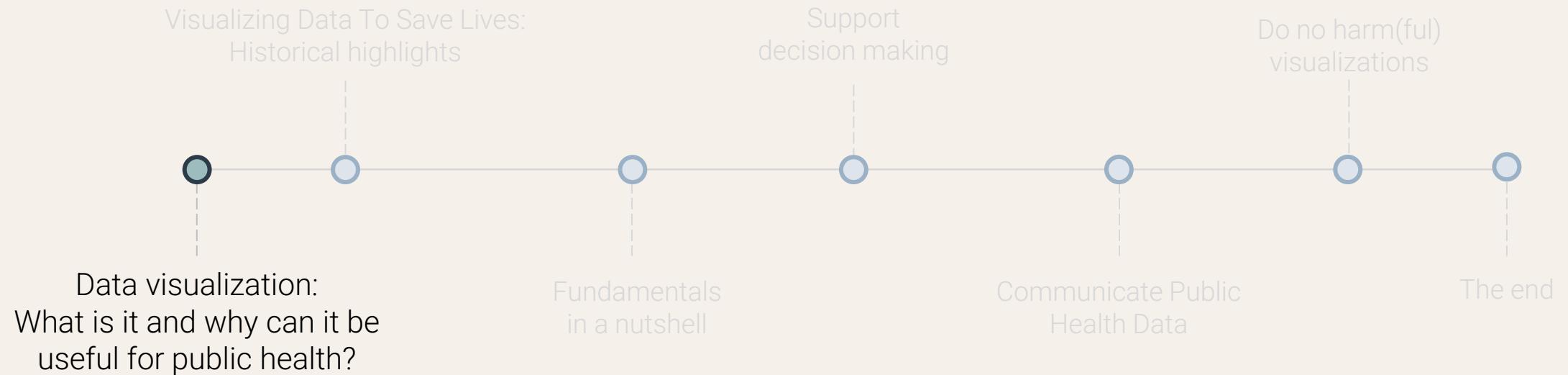


Estamos
à procura de
voluntários!

Agenda



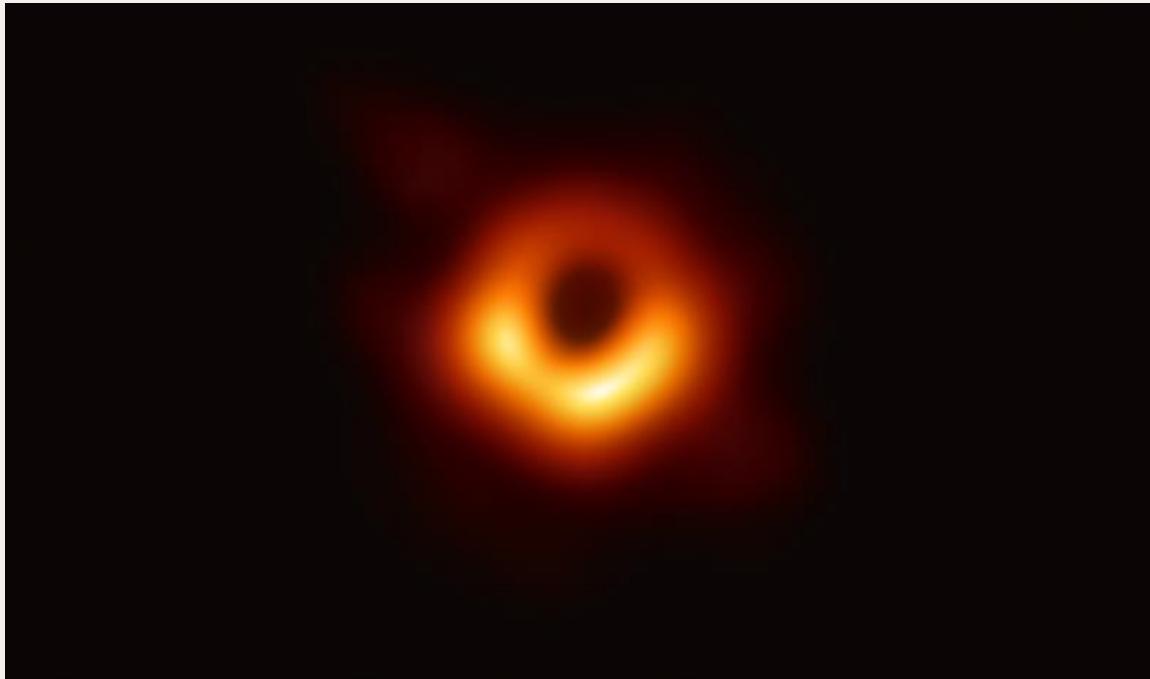
Agenda



“The visual representation and presentation of data to facilitate understanding”

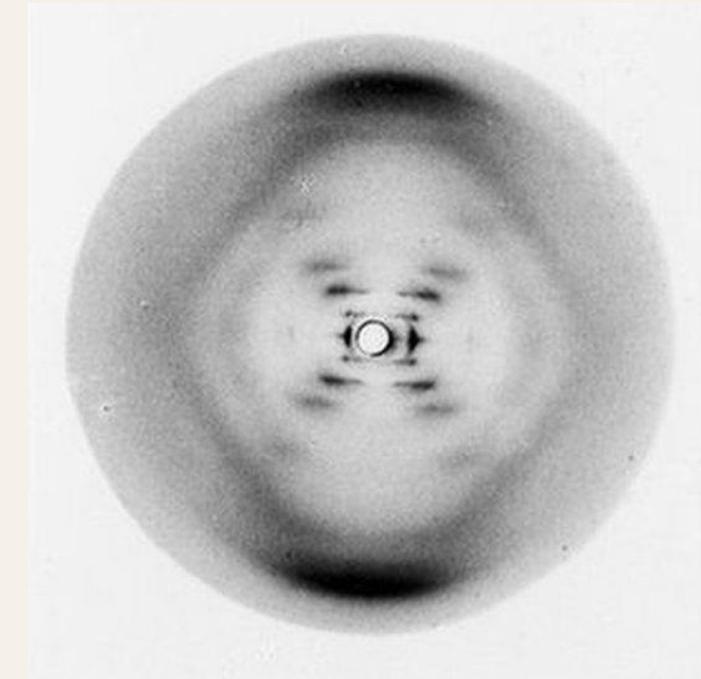
(Andy Kirk, 'Data Visualisation: A Handbook for Data Driven Design' , 2019)

The invisible becomes visible



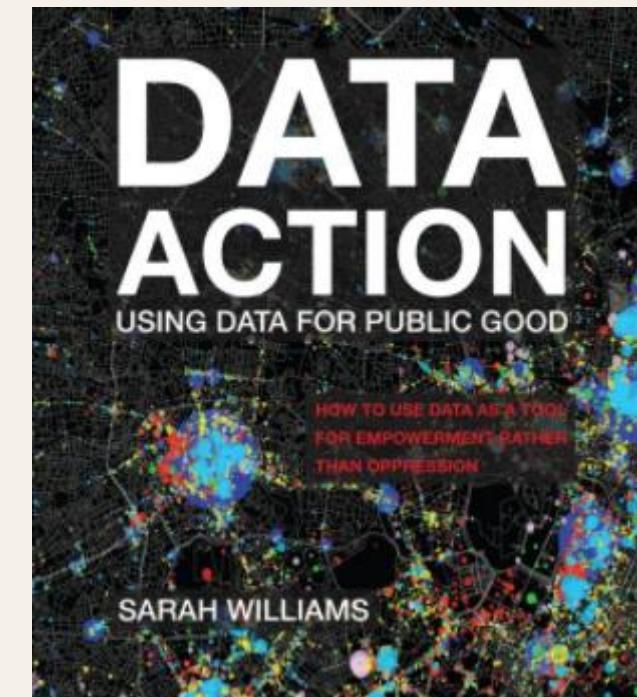
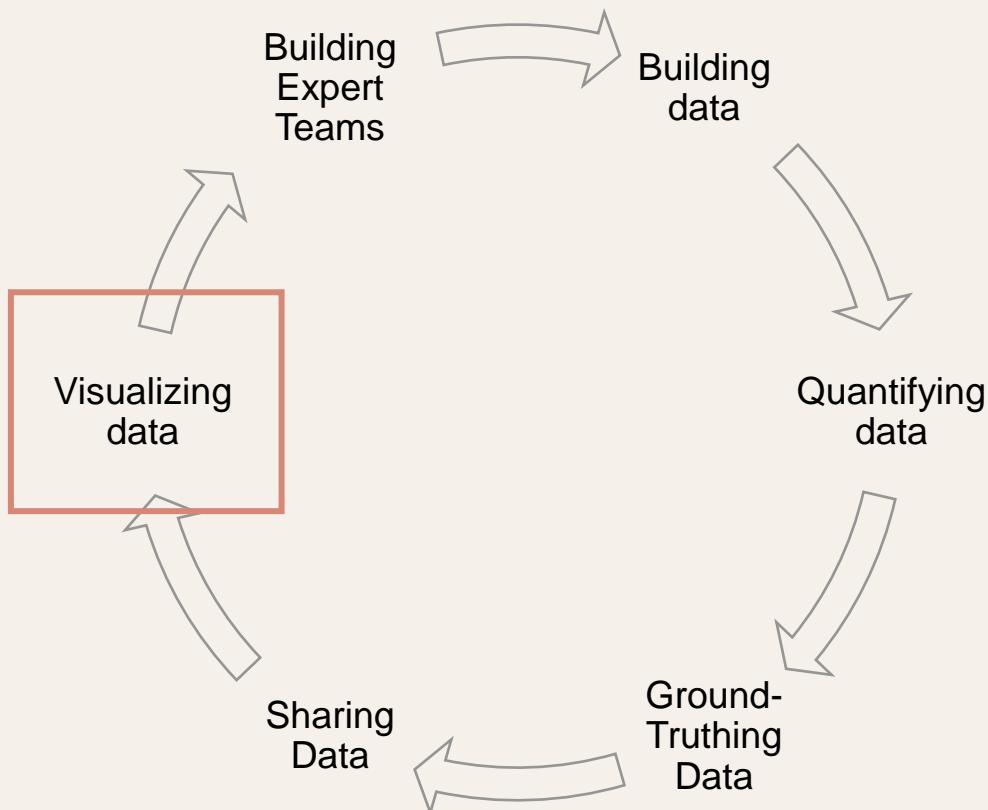
Event Horizon Telescope Collaboration translated petabytes of data from a stack of hard disks into the first-ever image of a black hole.

<https://www.atlasoftheinvisible.com/>
<https://www.youtube.com/watch?v=5kKB97170IU>



Rosalind Franklin and graduate Ray Gosling bombarded a strand of DNA with X-rays for more than sixty hours. When the rays bounced off the electrons in the molecule's atoms, they refracted into this cruciform pattern, from which the double -helix structure of DNA was inferred.

Data action cycle



<https://mitpress.mit.edu/books/data-action>

<https://www.youtube.com/watch?v=XsjdSzQMRr0&t=1972s>

Agenda

Visualizing Data To Save Lives: Historical highlights

Data visualization: What is it and
why can it be useful for public health?

Support decision making

Fundamentals
in a nutshell

Communicate Public
Health Data

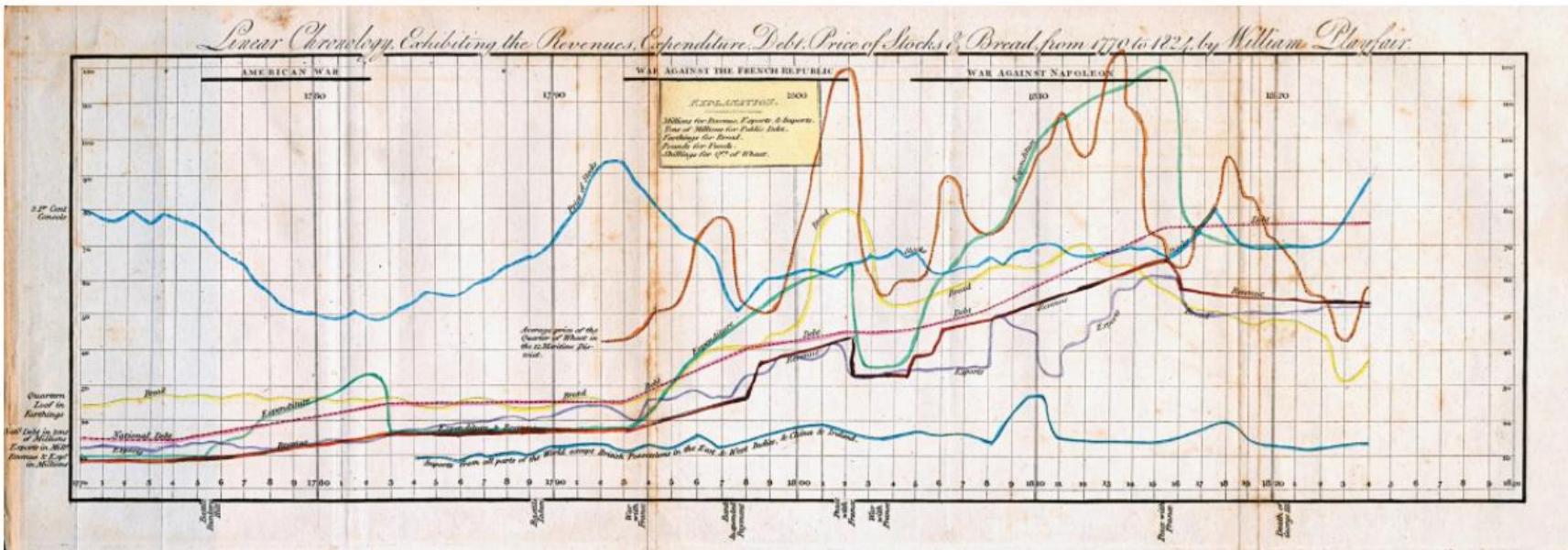
Do no harm(ul) visualizations

The end

WHEN GRAPHS ARE A MATTER OF LIFE AND DEATH

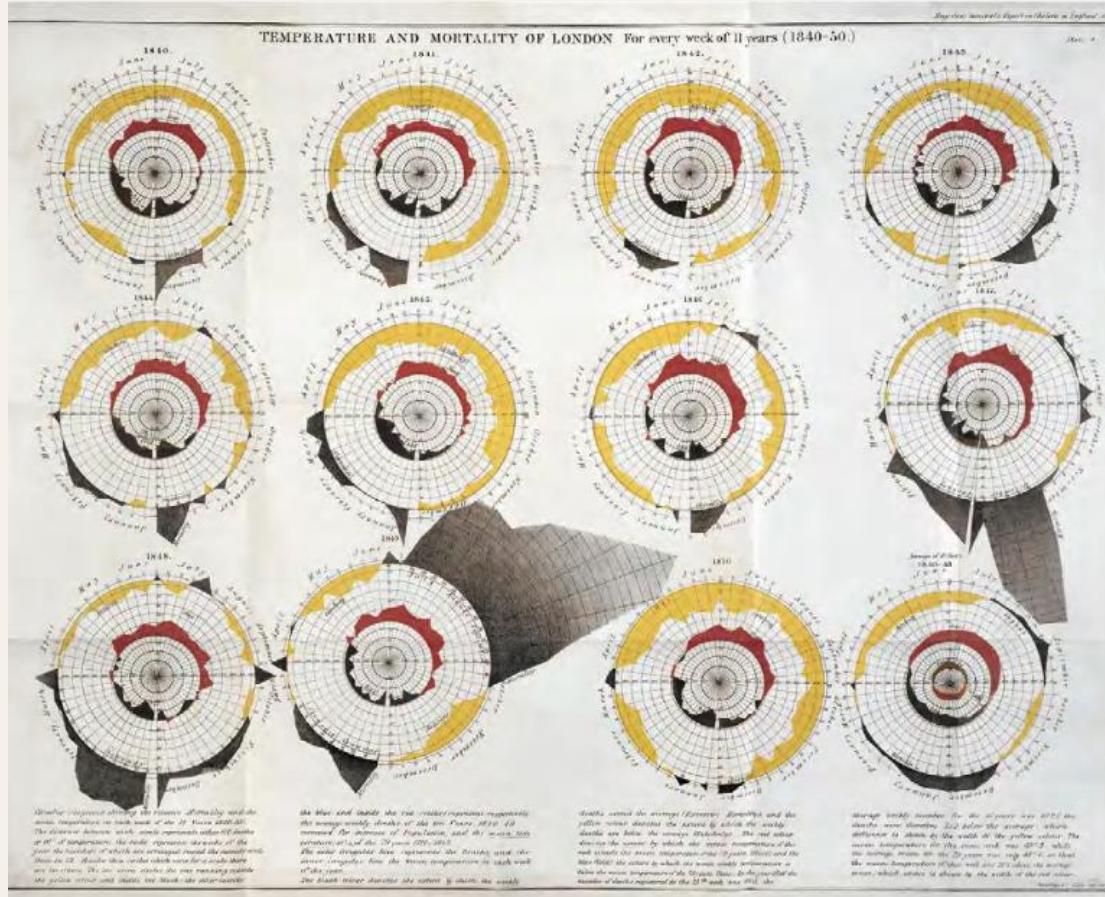
Pie charts and scatter plots seem like ordinary tools, but they revolutionized the way we solve problems.

By Hannah Fry
June 14, 2021



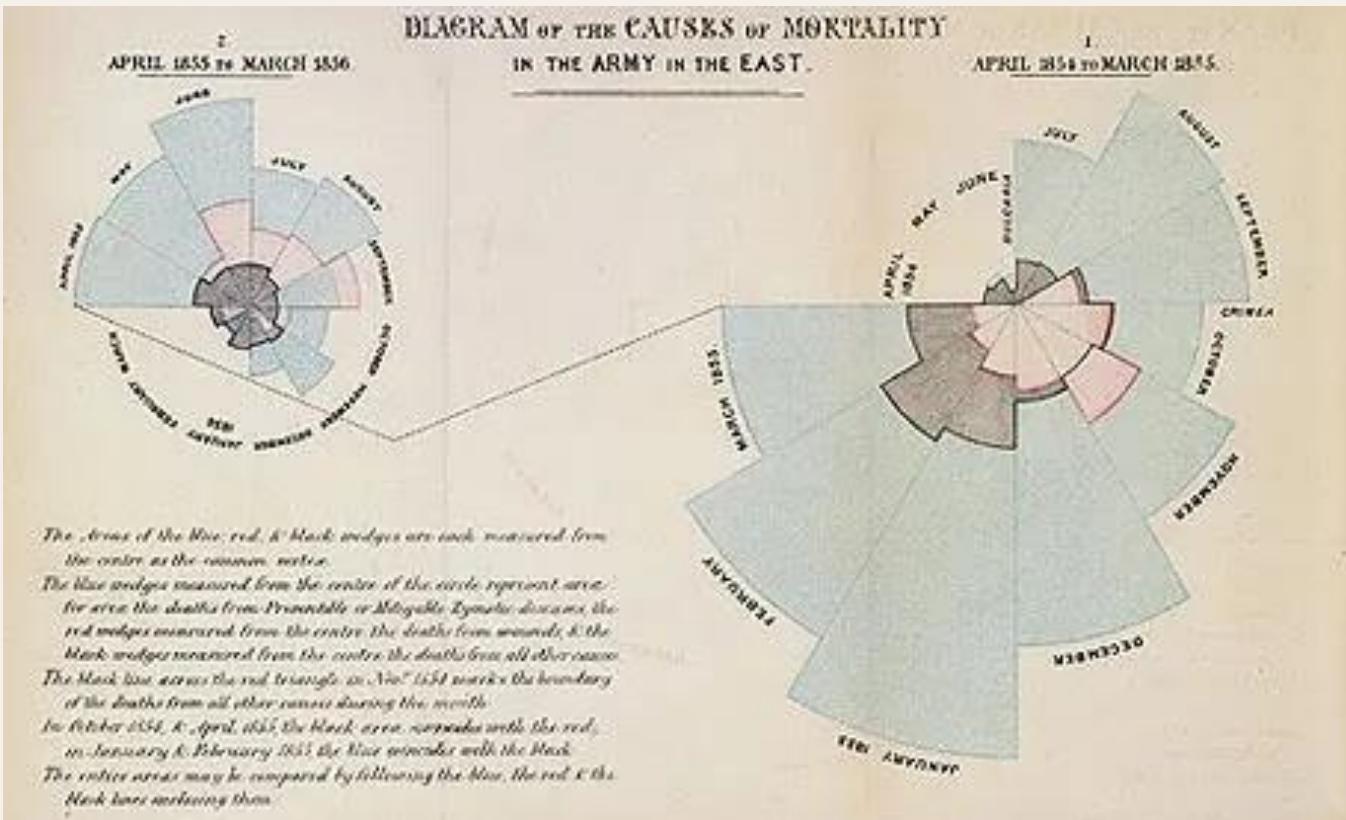
An 1824 time-series graph by William Playfair. Early viewers were bewildered by conventions we now take for granted. Image by William Playfair

William Farr



William Farr
Temperature and Mortality of London, 1852

Florence Nightingale



"Diagram of the Causes of Mortality in the Army in the East," from Florence Nightingale's "A Contribution to the Sanitary History of the British Army during the Late War with Russia"

Ignaz Semmelweis

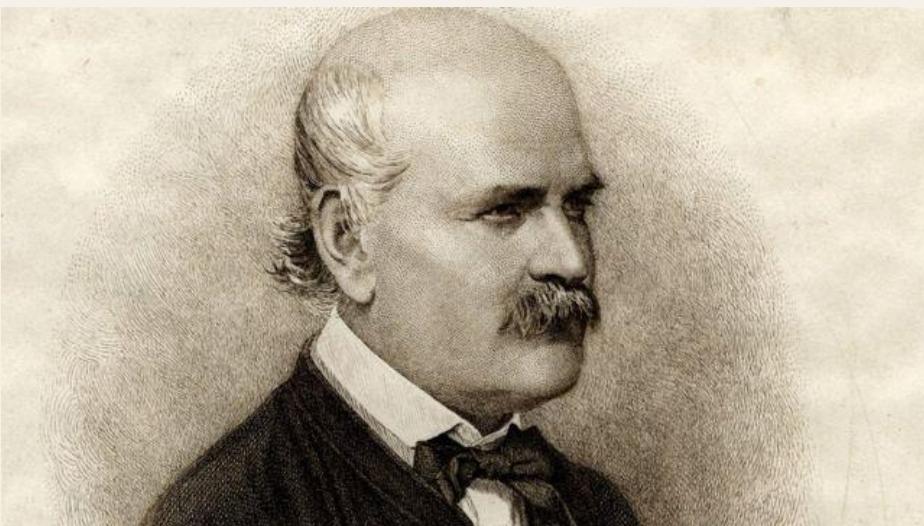
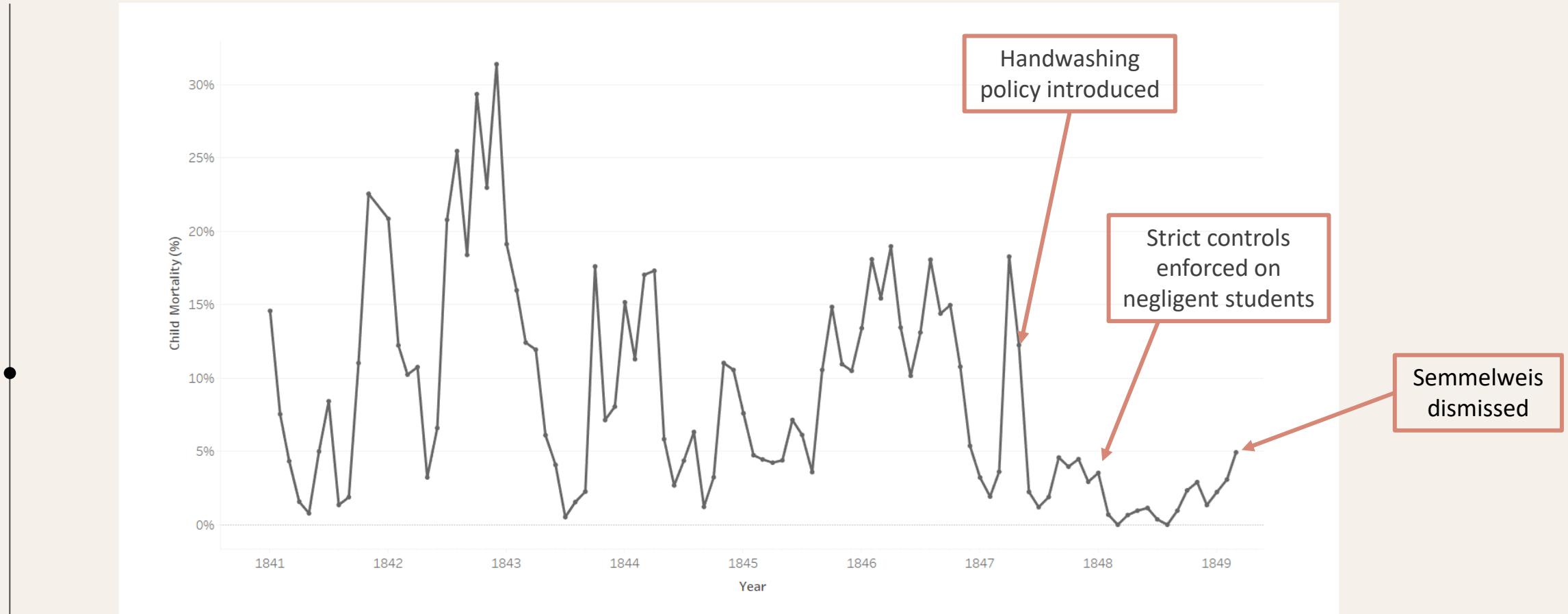


Tabelle Nr. XVII.

Standesausweis der k. k. Gebäraanstalt vom 16. August 1784
angefangen.

Jahr	Aufge-nommen	Zahl der Gestor-benen	Percent-Antheil	Jahr	Aufge-nommen	Zahl der Gestor-benen	Percent-Antheil
1784	284	6	2 . 11	1817	2735	25	0 . 91
1785	899	13	1 . 44	1818	258	56	2 . 18
1786	1151	5	0 . 43	1819	3080	154	4 . 98
1787	1407	5	0 . 35	1820	2998	75	2 . 50
1788	1425	5	0 . 35	1821	3294	55	1 . 66
1789	1246	7	0 . 56	1822	3066	26	0 . 84
1790	1326	10	0 . 75	1823	2872	214	7 . 45
1791	1395	8	0 . 57	1824	2911	144	4 . 94
1792	1574	14	0 . 89	1825	2594	229	4 . 82
1793	1684	44	2 . 61	1826	2359	192	8 . 12
1794	1768	7	0 . 39	1827	2367	51	2 . 15
1795	1198	38	2 . 11	1828	2833	101	3 . 56
1796	1904	22	1 . 16	1829	3012	140	4 . 64
1797	2012	5	0 . 24	1830	2797	111	3 . 97
1798	2046	5	0 . 24	1831	3353	222	6 . 62
1799	2067	20	0 . 96	1832	3331	105	3 . 15
1800	2070	41	1 . 98	1833	3907	205	5 . 25
1801	2106	17	0 . 80	1834	4218	355	8 . 41
1802	2346	9	0 . 38	1835	4040	227	5 . 61
1803	2215	16	0 . 72	1836	4144	331	7 . 95
1804	2022	8	0 . 39	1837	4363	375	8 . 59
1805	2112	9	0 . 40	1838	4560	179	3 . 92
1806	1875	13	0 . 71	1839	4992	248	4 . 96
1807	925	6	0 . 63	1840	5166	328	6 . 44
1808	855	7	0 . 81	1841	5454	330	6 . 65
1809	912	13	1 . 42	1842	6024	730	12 . 11
1810	744	6	0 . 80	1843	5914	457	7 . 72
1811	1050	20	1 . 90	1844	6244	336	5 . 38
1812	1419	9	0 . 63	1845	6756	313	4 . 63
1813	1945	21	1 . 08	1846	7027	567	8 . 06
1814	2062	66	3 . 70	1847	7039	210	2 . 98
1815	2591	19	0 . 72	1848	7095	91	1 . 18
1816	2410	12	0 . 49	—	—	—	—

Ignaz Semmelweis

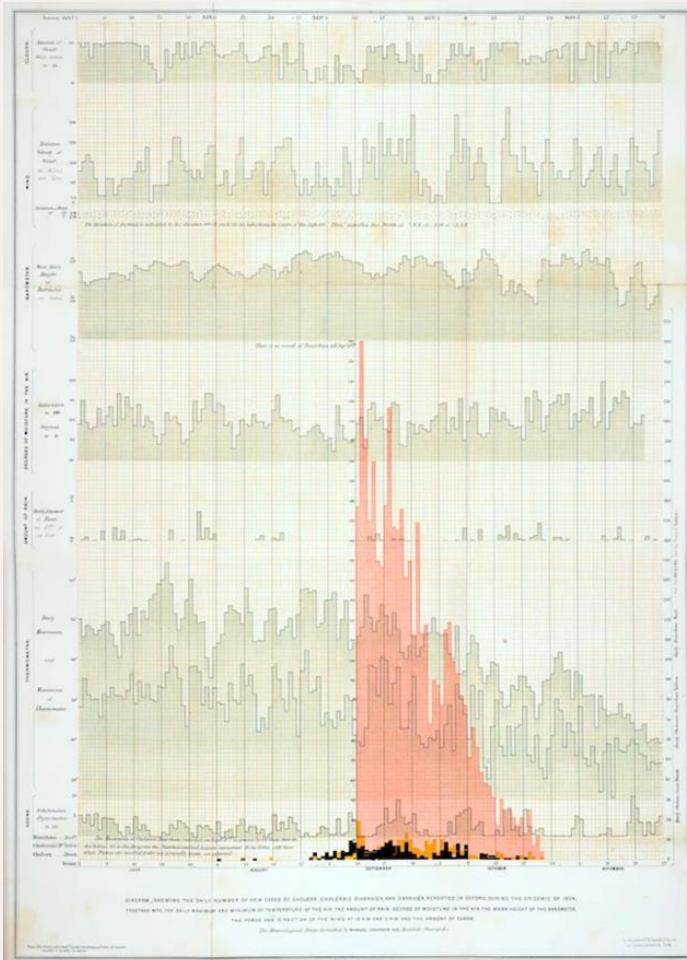


John Snow



Map of London created by John Snow in 1854. London was experiencing a deadly cholera epidemic, when Snow tracked the cases on this map. The cholera cases are highlighted in black. Using this map, Snow was able to trace the cholera outbreak to a single infected water pump.

John Snow

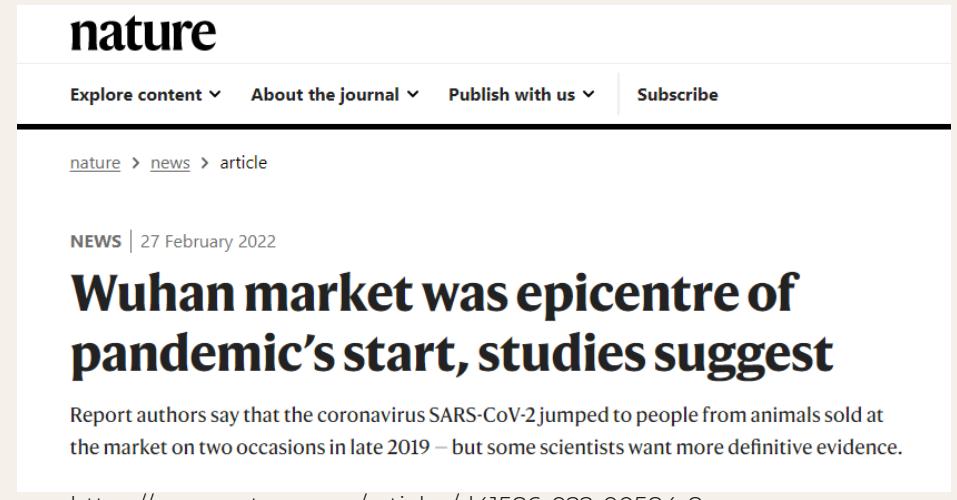


Acland's chart showing weather variables he speculated might have contributed to the Oxford outbreak.

Same problems, different data



<https://twitter.com/besttrousers/status/1497670412083380234?s=20&t=wnPdFxf7DqZ0ZrB5RIQYEQ>



nature

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature > news > article

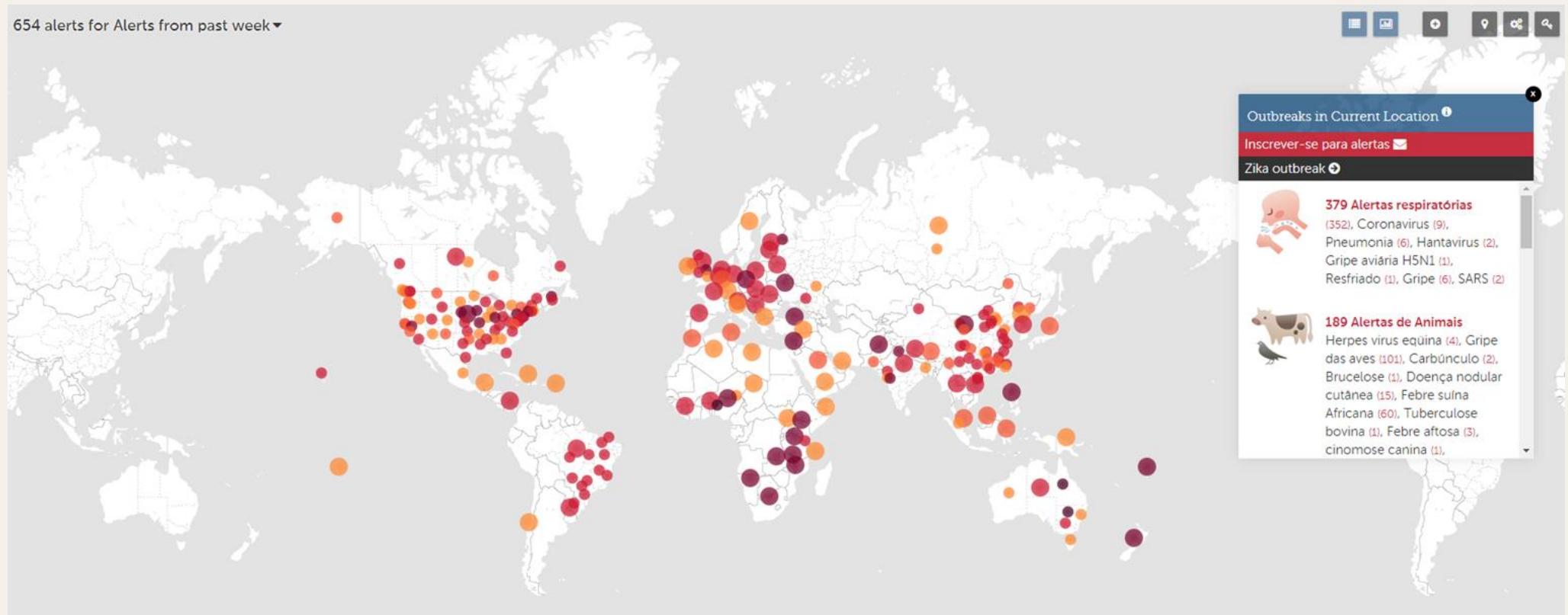
NEWS | 27 February 2022

Wuhan market was epicentre of pandemic's start, studies suggest

Report authors say that the coronavirus SARS-CoV-2 jumped to people from animals sold at the market on two occasions in late 2019 – but some scientists want more definitive evidence.

<https://www.nature.com/articles/d41586-022-00584-8>

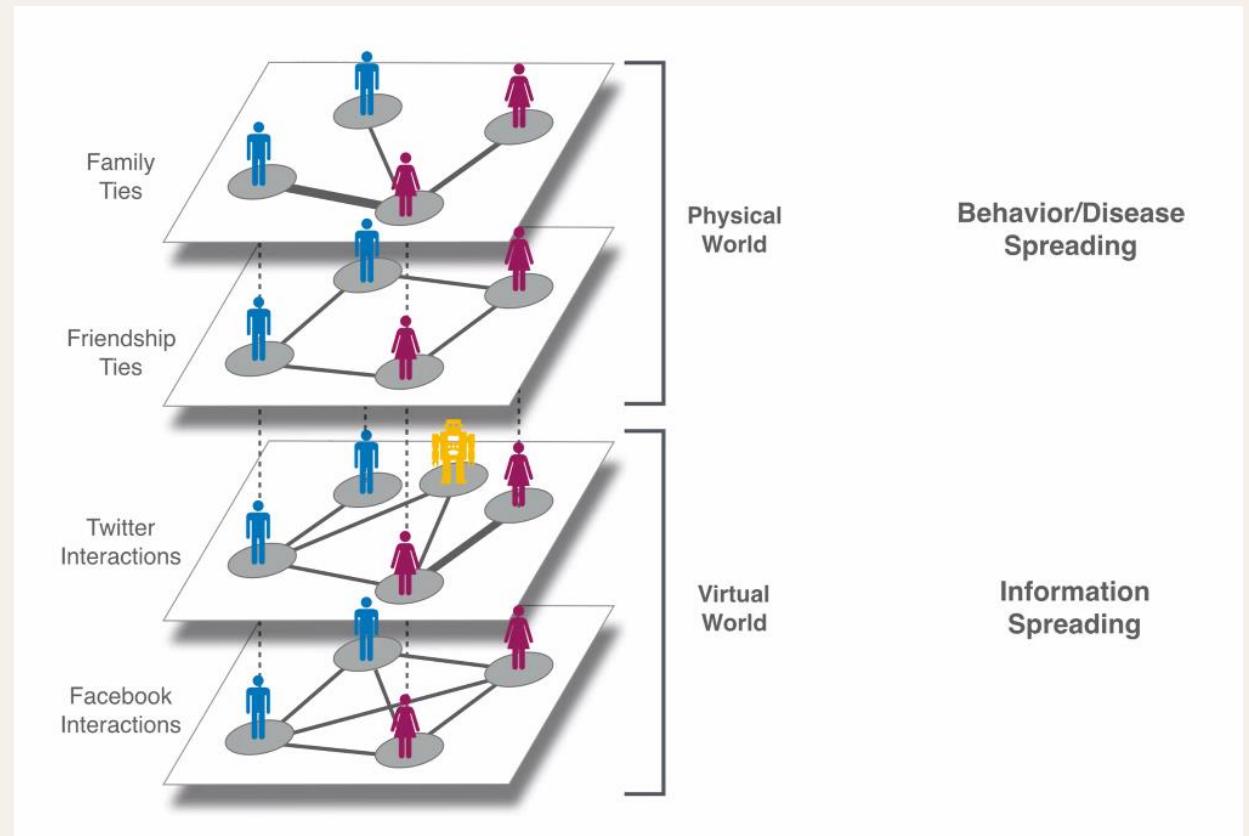
Same problems, different data



Brownstein, J. S., Freifeld, C. C., Reis, B. Y., & Mandl, K. D. (2008). Surveillance Sans Frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project. PLoS medicine, 5(7), e151.

Same problems, different (more) data

- DNA
- *omics data
- Location/ satellite data
- Activity data
- Lab tests
- Medical imaging
- Social media
- Media
- Sensor data (e.g.: pollution)
- Health history
- Census, etc.



Agenda



But first:

Will this data or data visualization help solve my problem?



<https://sketchplanations.com/looking-under-the-lamppost>

Different purposes:

- **Support decision making**
- **Communicate to different audiences**

Explanatory analysis

- Hypothesis Generation (EDA)
- Monitor progress

Exploratory analysis

Who, what, and how

Identifying the decision maker is one way of narrowing our audience.

(E.g.: public health professionals, community leaders or policy makers, healthcare providers, public or special populations, etc.)

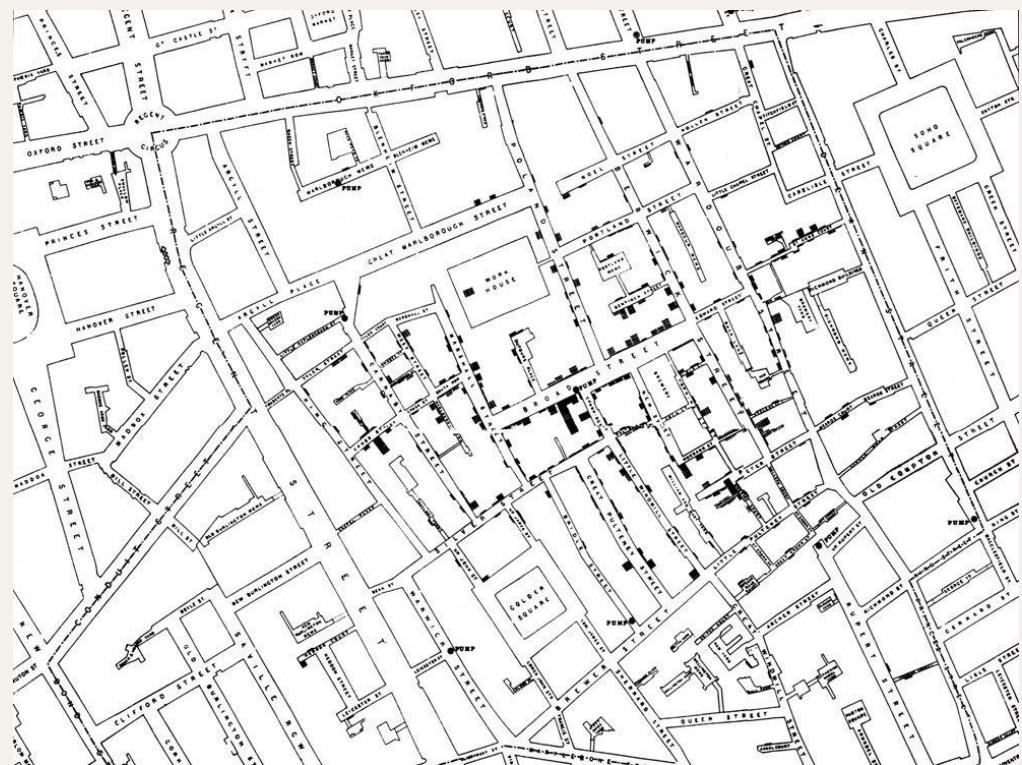
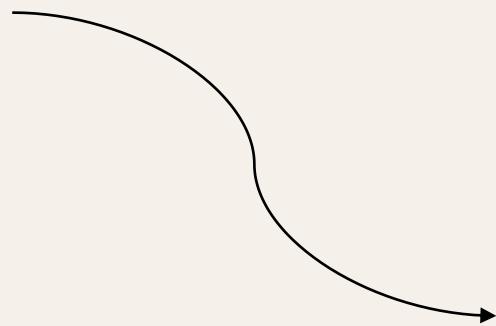
What do we need our **audience to know** or do?

What **data is available** that will help make our point? Data becomes supporting evidence of the story we will build and tell.

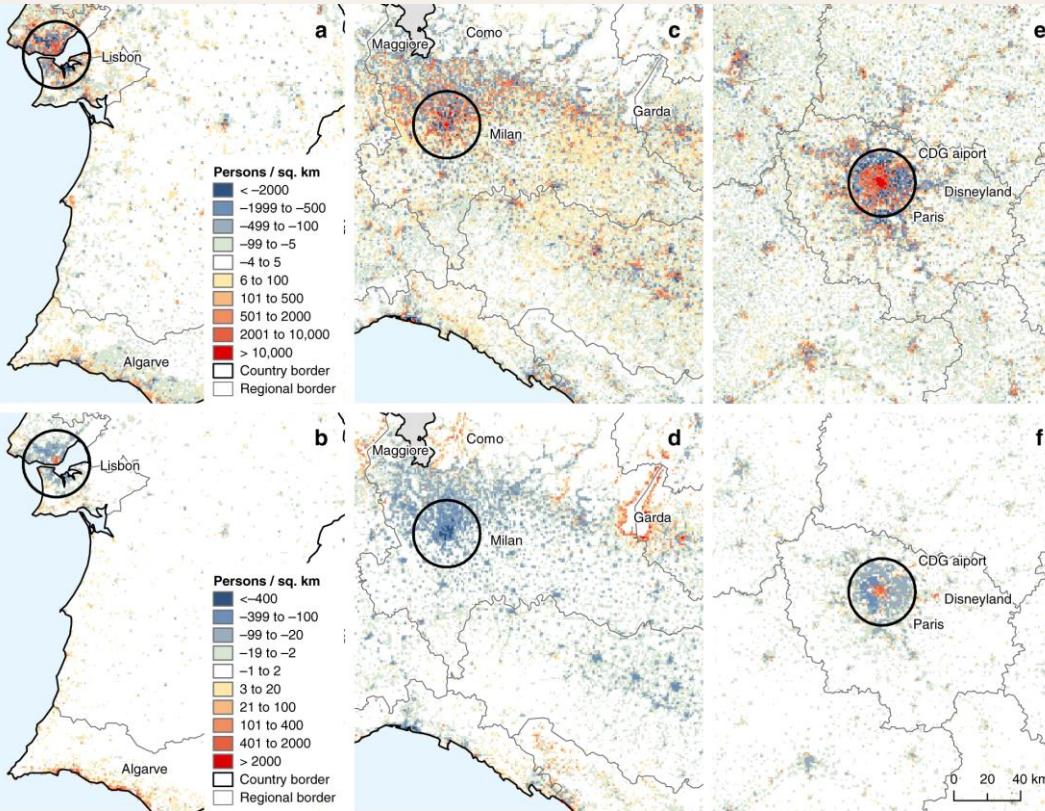
Who = Board of guardians

What = remove pump handle

How =



What makes a data visualization?



Spatio-temporal differences in population density in selected sites.

<https://www.nature.com/articles/s41467-020-18344-5>

Intent: showing temporal changes in Europe's population density patterns

Data:

- Population records
- Population at 1 km² grid-cell level between day and nighttime

Visual Choice:

- Map
- Points
- Color
- Annotations

Calculations:

- Absolute differences
- Average of all months
- Grid cell location

Data visualization process

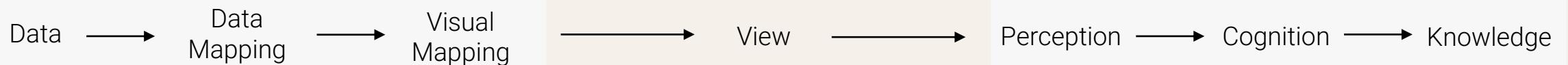
Encoding

Data Transformation, Rendering



Decoding

Perception, Cognition

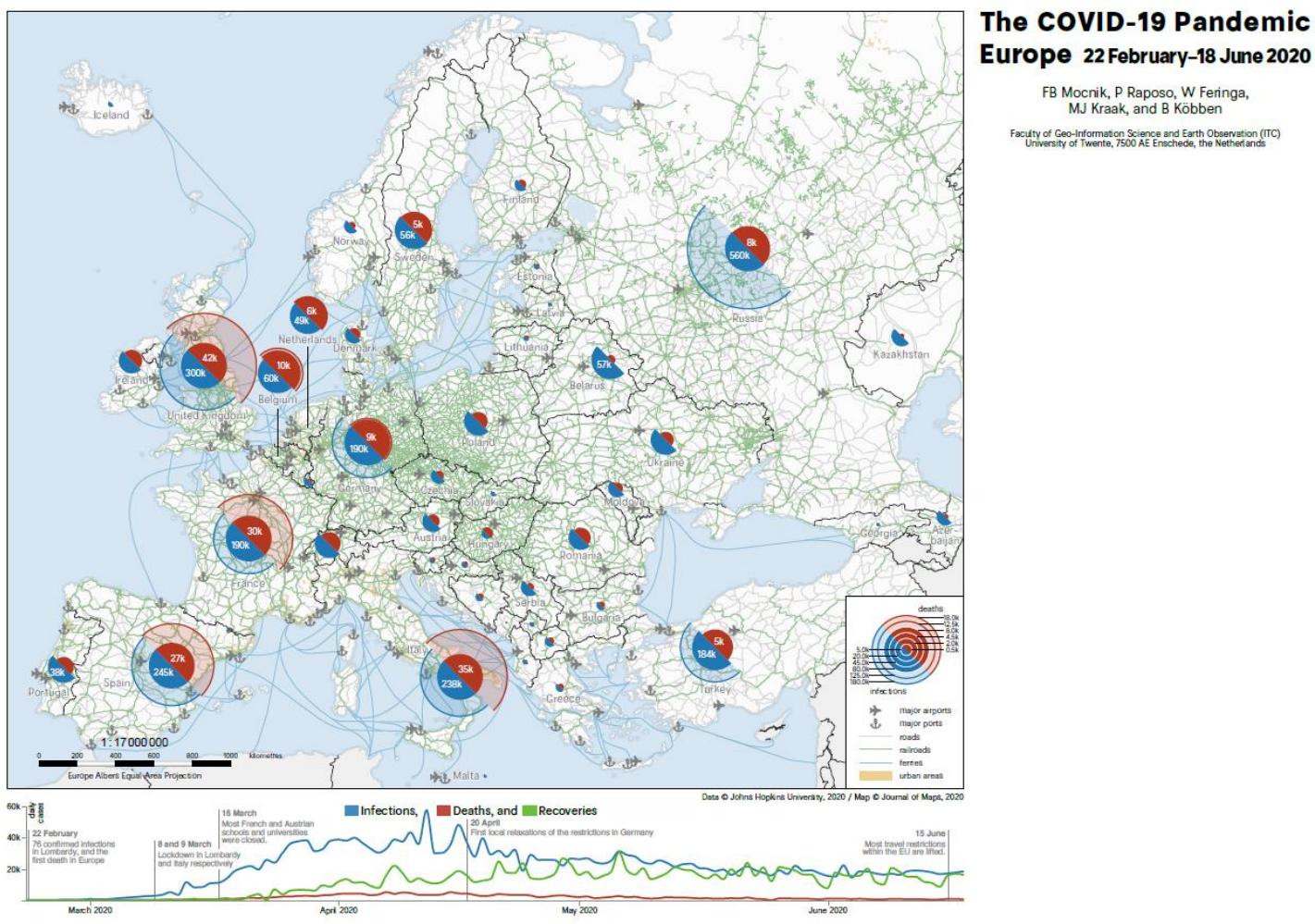


Data visualization process



"Commonly used metrics are poorly understood by most health professionals. Reporting numbers using natural frequencies and visual aids may facilitate improved understanding and better estimation of the probability of disease."

Exercise 1



Intent?

Data?

Visual Choice?

Calculations?

Chart types (some)

SIMPLE TEXT



Just because you have numbers doesn't mean you need a graph!

OFTEN THERE ARE MORE EFFICIENT WAYS

TABLE



What is the main point I want to make?

CATEGORY	A	B	C
CATEGORY 1	154	211	407
CATEGORY 2	492	363	202
CATEGORY 3	354	172	191
CATEGORY 4	107	274	582

Avoid using tables in live presentations because people stop listening & start reading

HEAT MAP



CATEGORY	A	B	C
CATEGORY 1	154	211	407
CATEGORY 2	492	363	202
CATEGORY 3	354	172	191
CATEGORY 4	107	274	582

EYES CAN EASILY PICK OUT BIG DIFFERENCES IN COLOR INTENSITY, but smaller ones don't stand out

Can work well when beginning to explore data and deciding where to dig further

SCATTER PLOT



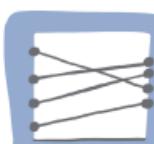
Good for encoding data simultaneously on two axes to identify what relationships exist

LINE



Rule: The lines that connect the dots have to make sense! Most effective with continuous data, often time

SLOPE GRAPH



A FANCY WORD FOR A LINE GRAPH WITH ONLY 2 POINTS

Useful to focus on change between two points in time or difference between groups

WATERFALL



BEGINNING QUANTITY ENDING QUANTITY

ADDITIONS & DEDUCTIONS

Often used in finance to show variance to budget

SQUARE AREA (AKA WAFFLE CHART)



THE GRID IS IMPORTANT BECAUSE WE TEND TO OVERESTIMATE AREAS

Good for showing numbers of very different magnitudes, or as an alternative to a pie chart

It depends on ...

Nature of the data to represent:

- Categorical (nominal)
- Ordinal
- Quantitative

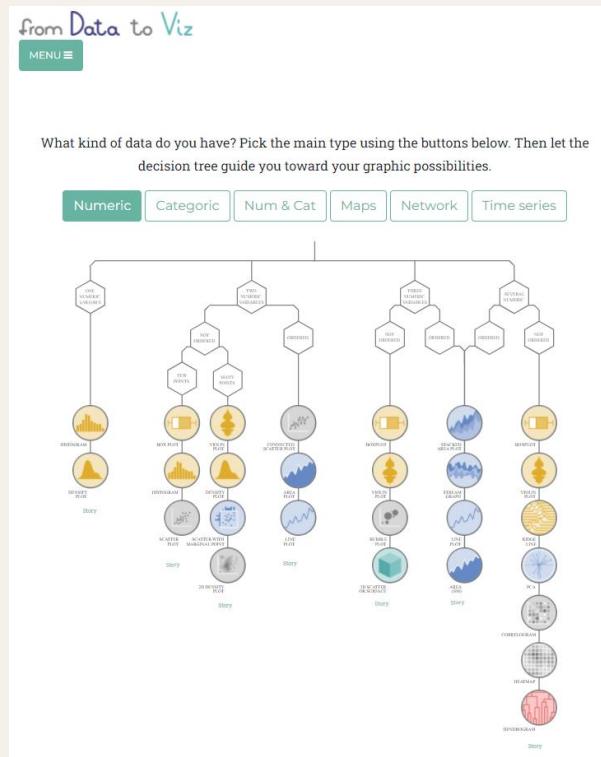
Number of attributes:

- Univariate
- Bivariate
- Multivariate

Data structures:

- Linear
- Temporal
- Spatial or geographical
- Hierarchical
- Network

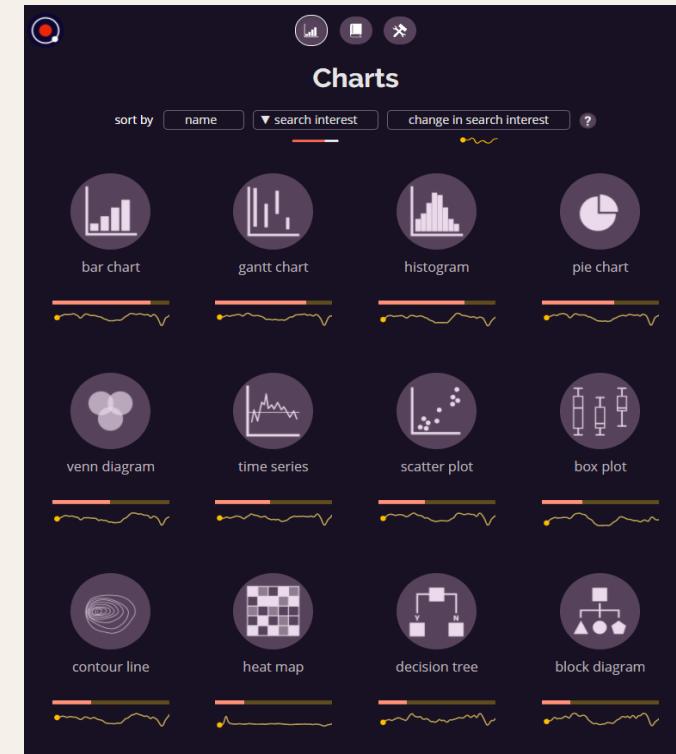
DataViz Catalogue



[.data-to-viz.com](http://data-to-viz.com)



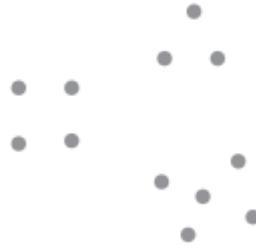
datavizproject.com



visualizationuniverse.com

While data might work with multiple chart types, it's up to us to select the one that ensures that the message is clear and accurate.

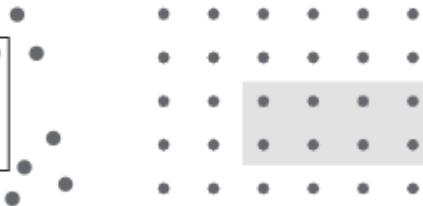
Gestalt principles of visual perception



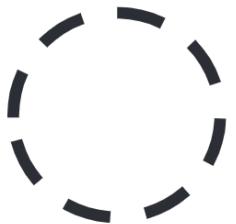
Proximity



Similarity



Enclosure



Closure



Continuity



Connection

Preattentive attributes

How many 3s?

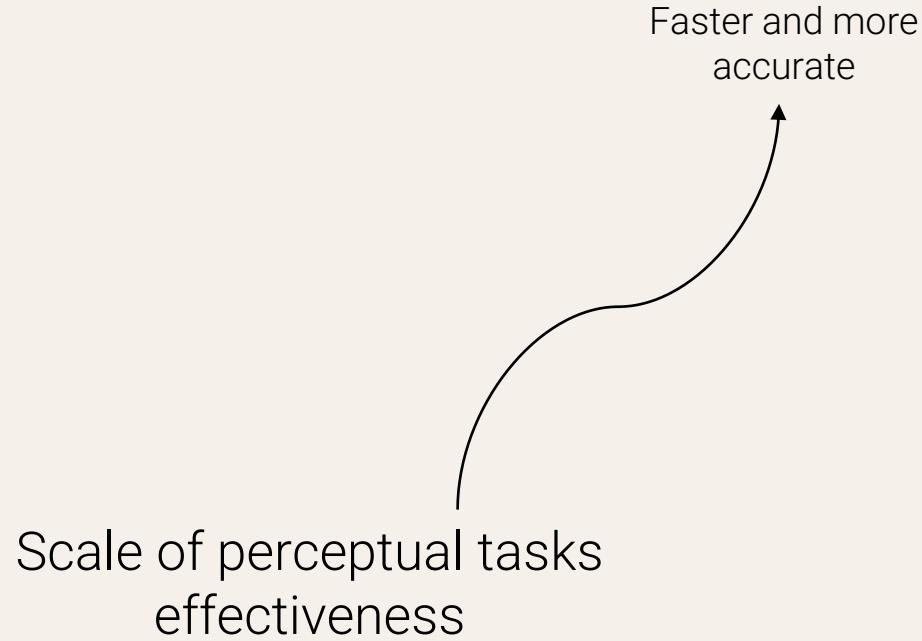
756395068473
658663037576
860372658602
846589107830

Preattentive attributes

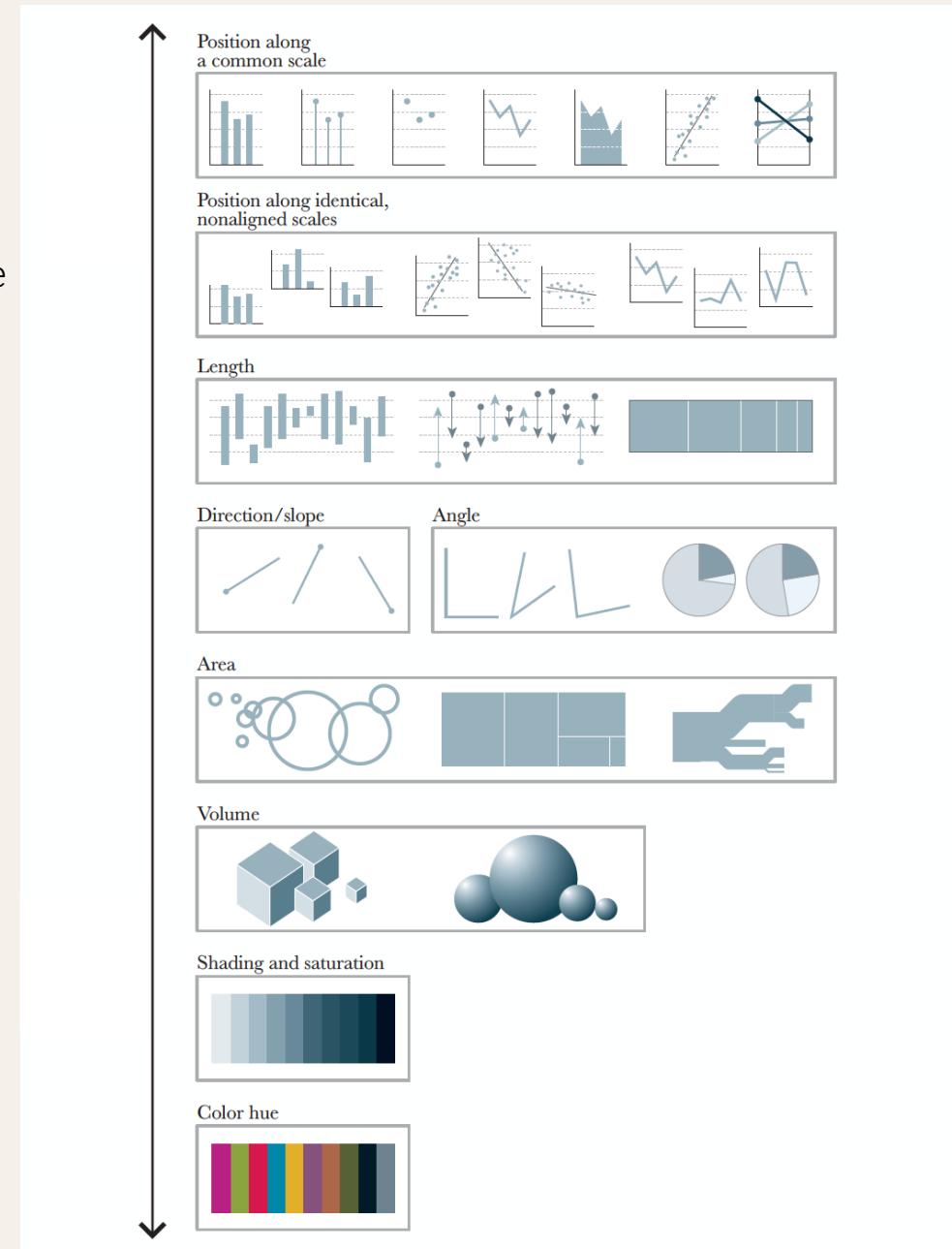
“Enable the audience to see
what we want them to see
before they even
know they’re seeing it!”

756395068473
658663037576
860372658602
846589107830

Preattentive attributes



Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association*, 79(387), 531-554.



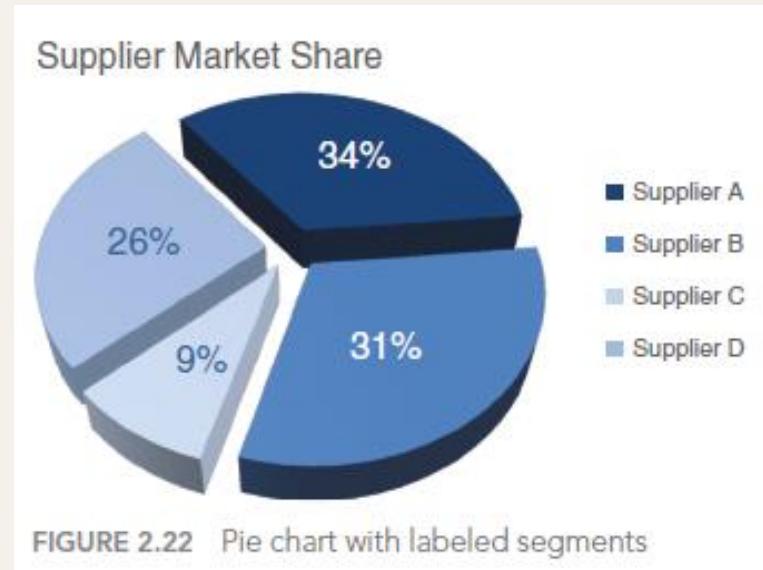
Choosing an effective visual

Which supplier is the largest based on this visual?

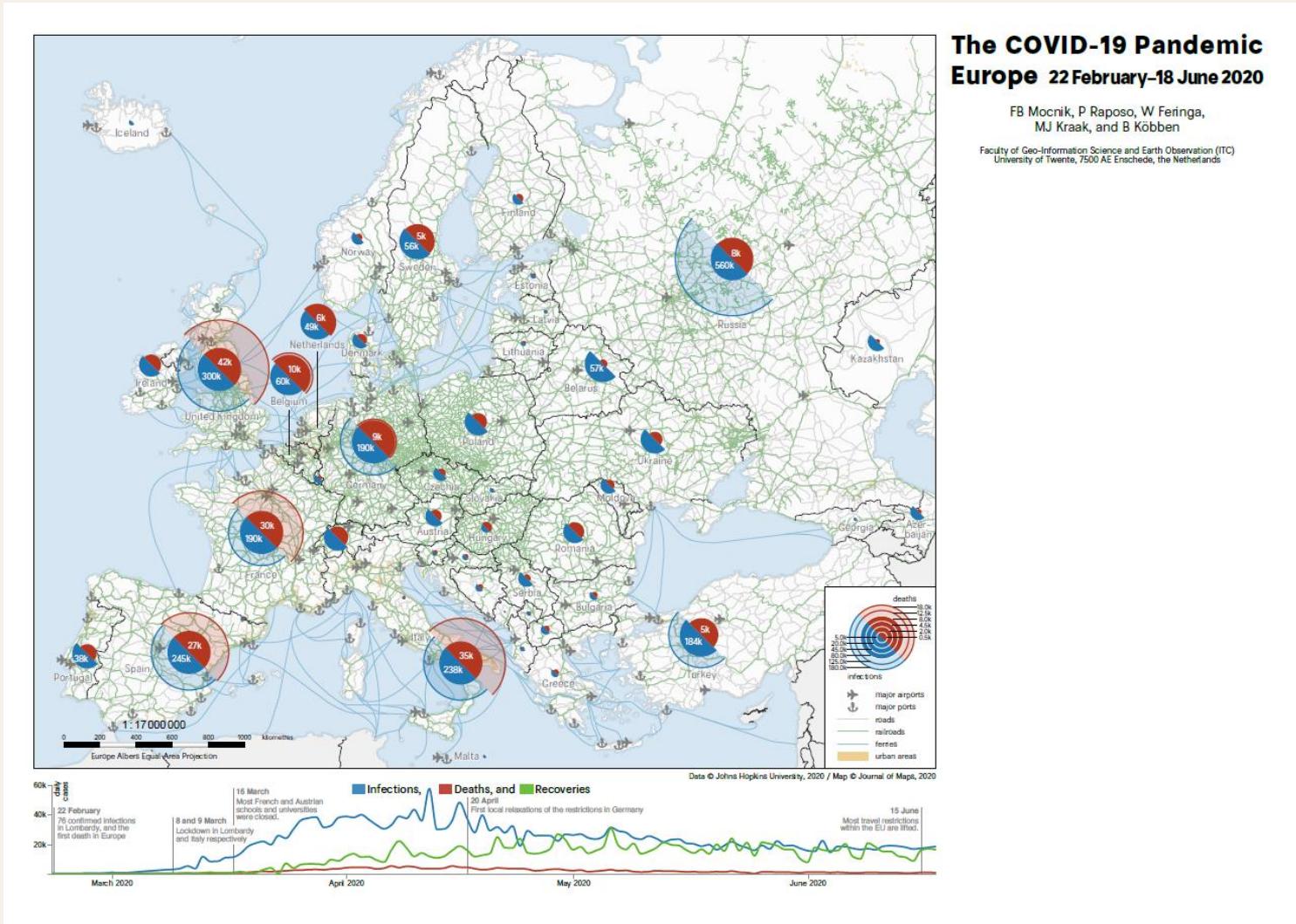


Choosing an effective visual

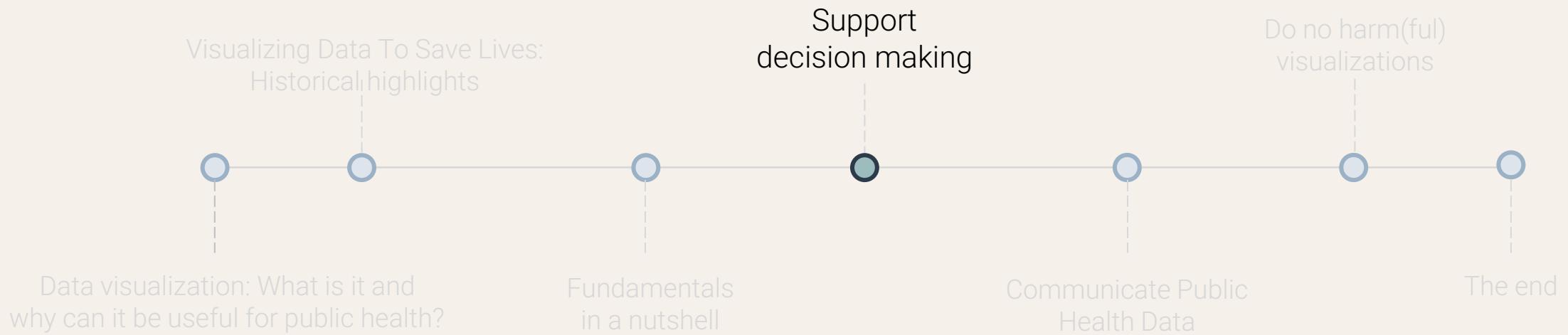
Which supplier is the largest based on this visual?



Exercise 2 – Let's discuss



Agenda



Data and Data Visualizations in Public Health

POPULATIONS AT RISK ACROSS THE
LIFESPAN-POPULATION STUDIES

PHN PUBLIC HEALTH NURSING

WILEY

Understanding data use and preference of data visualization for public health professionals: A qualitative study

Seungeun Park PhD, MPH, RN¹  | Betty Bekemeier PhD, MPH, FAAN² |
Abraham D. Flaxman PhD³

Data and Data Visualizations in Public Health

Challenges:

- Data quality issues
- Scattered, outdated and inconsistent data sources
- Limited access to data
- Lack of data related to emergent issues (opioids, tattoo, e-cigarette)
- Lack of experts

Tasks:

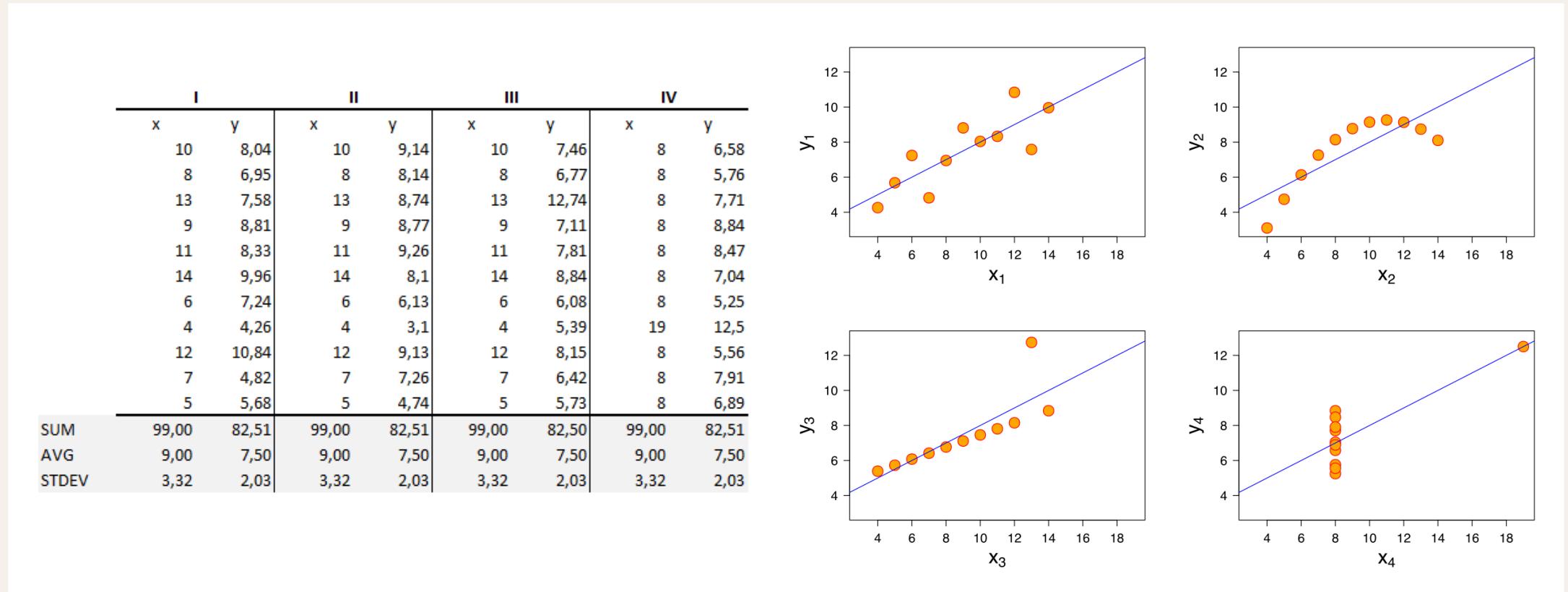
- Making comparisons with other regions, countries, over time, etc.
- Understanding the needs of the
- Setting priorities for program planning and program evaluation
- Justifying funding
- Policy development
- Educating about public health issues

“Visualization brings advantages by **increasing the amount of information delivered** and **decreasing the cognitive and intellectual burden** to interpret information for decision-making.”

Some misconceptions

- “The data should speak for itself”
- “Visualization is intuitive”
- “A picture is worth a thousand words”
- “Show, don’t tell!”

Anscombe quartet

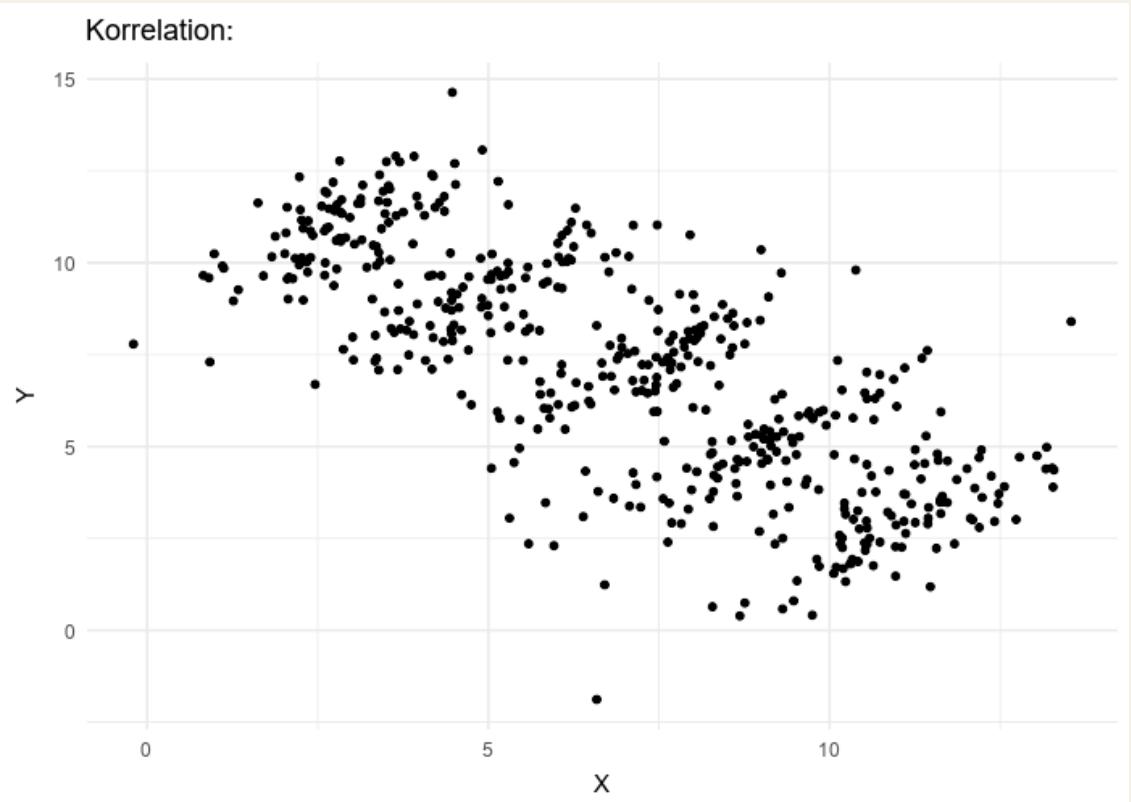


Visualization helps in situations where seeing the dataset structure in detail is better than seeing only a brief summary of it (losing information). (Munzner, 2014)

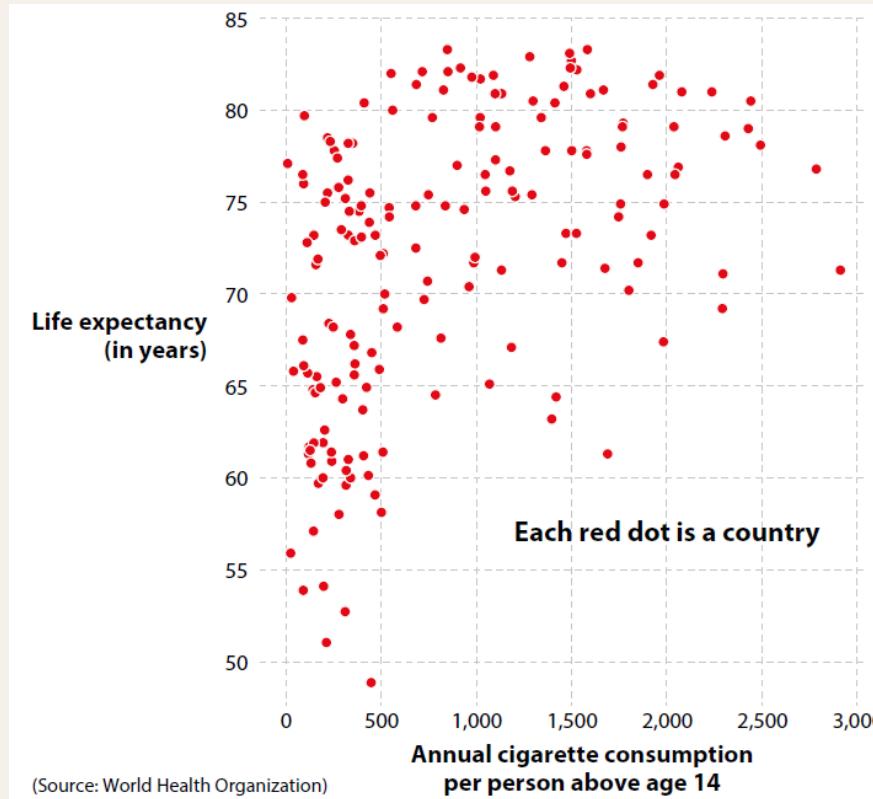
Matejka, J., & Fitzmaurice, G. (2017, May). Same stats, different graphs: generating datasets with varied appearance and identical statistics through simulated annealing. In Proceedings of the 2017 CHI conference on human factors in computing systems (pp. 1290-1294).

Simpson's Paradox

A trend appears in several groups of data but disappears or reverses when the groups are combined.

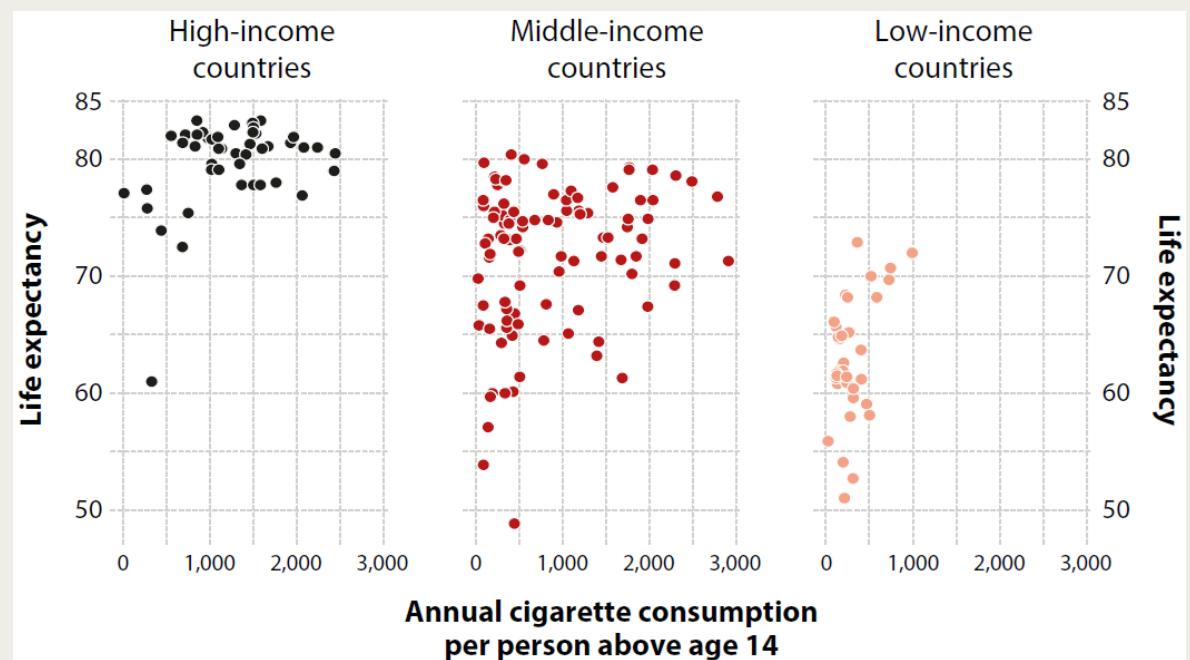
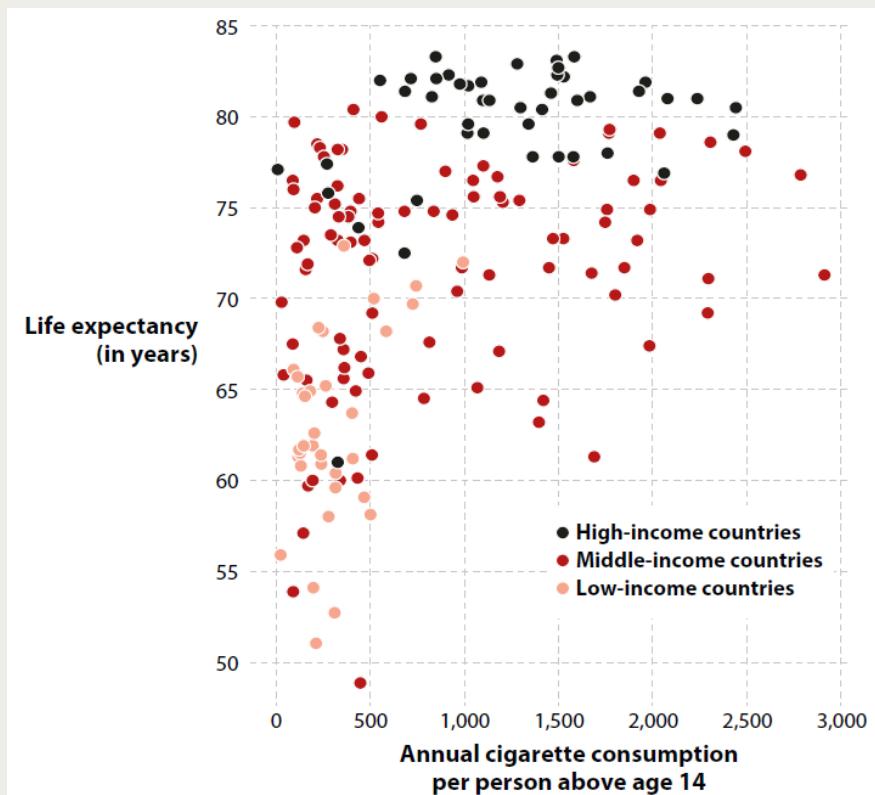


Simpson's Paradox



"A chart shows only what it shows, and nothing else"
(Alberto Cairo in *How Charts Lie*)

Simpson's Paradox

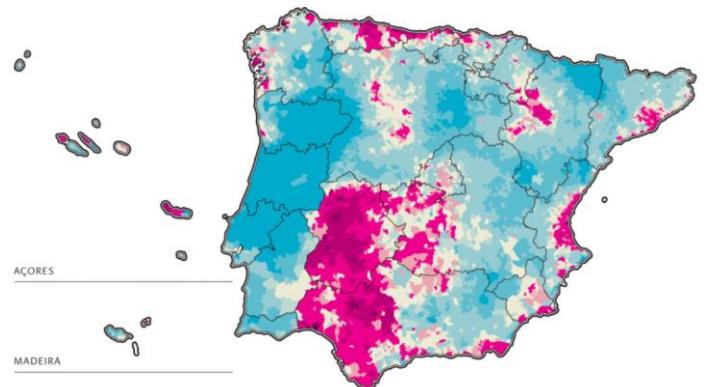


O Atlas ibérico do cancro: qual é o pior e onde é que o risco de morte é maior

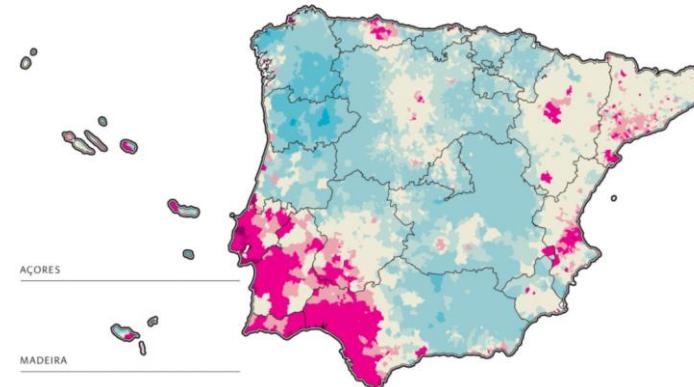
É o primeiro mapa que mostra o excesso de risco de morte por tumores em Portugal e Espanha, com mais perguntas do que respostas. Na próstata e estômago, a mortalidade é mais alta entre os portugueses. Já no pulmão e pâncreas, tumores com pior prognóstico, a vida é mais longa deste lado da fronteira



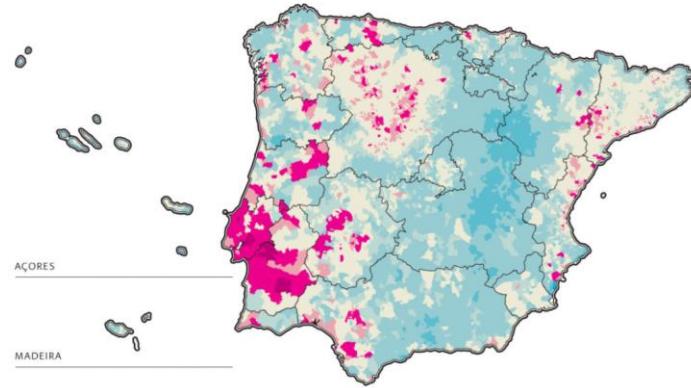
1. Pulmão



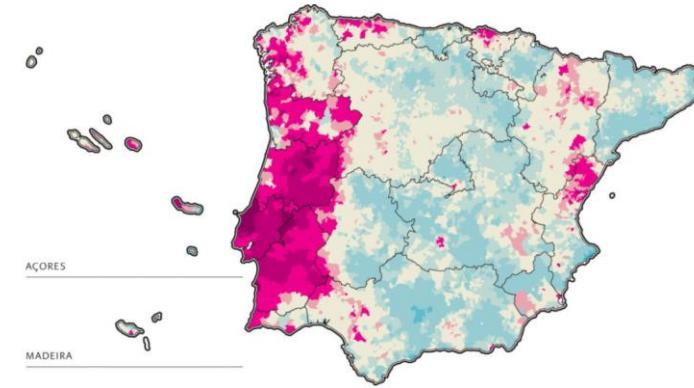
2. Mama



3. Colorretal



4. Próstata



Cáncer de pulmón

Exceso de riesgo de mortalidad

Inferior

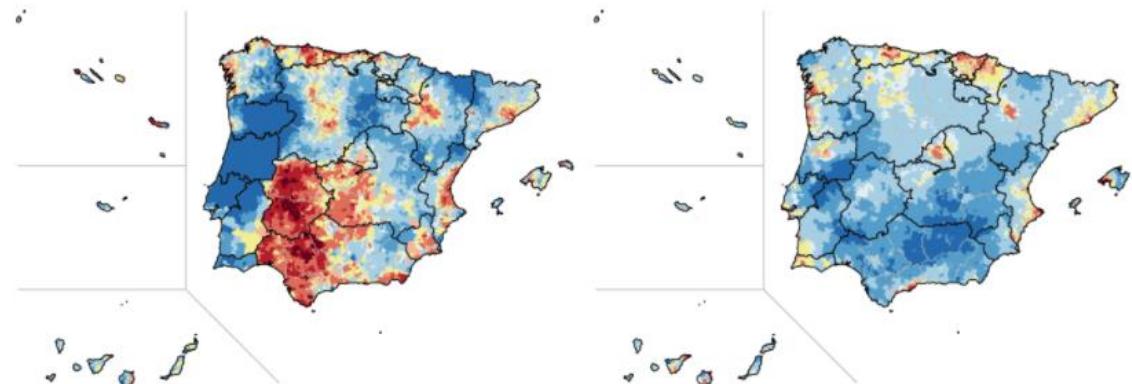
- En más del 33%
- Entre un 23% y un 33%
- Entre un 9% y un 23%
- Entre un 5% y un 9%

Promedio

Superior

- Entre un 5% y un 10%
- Entre un 10% y un 30%
- Entre un 30% y un 50%
- En más del 50%

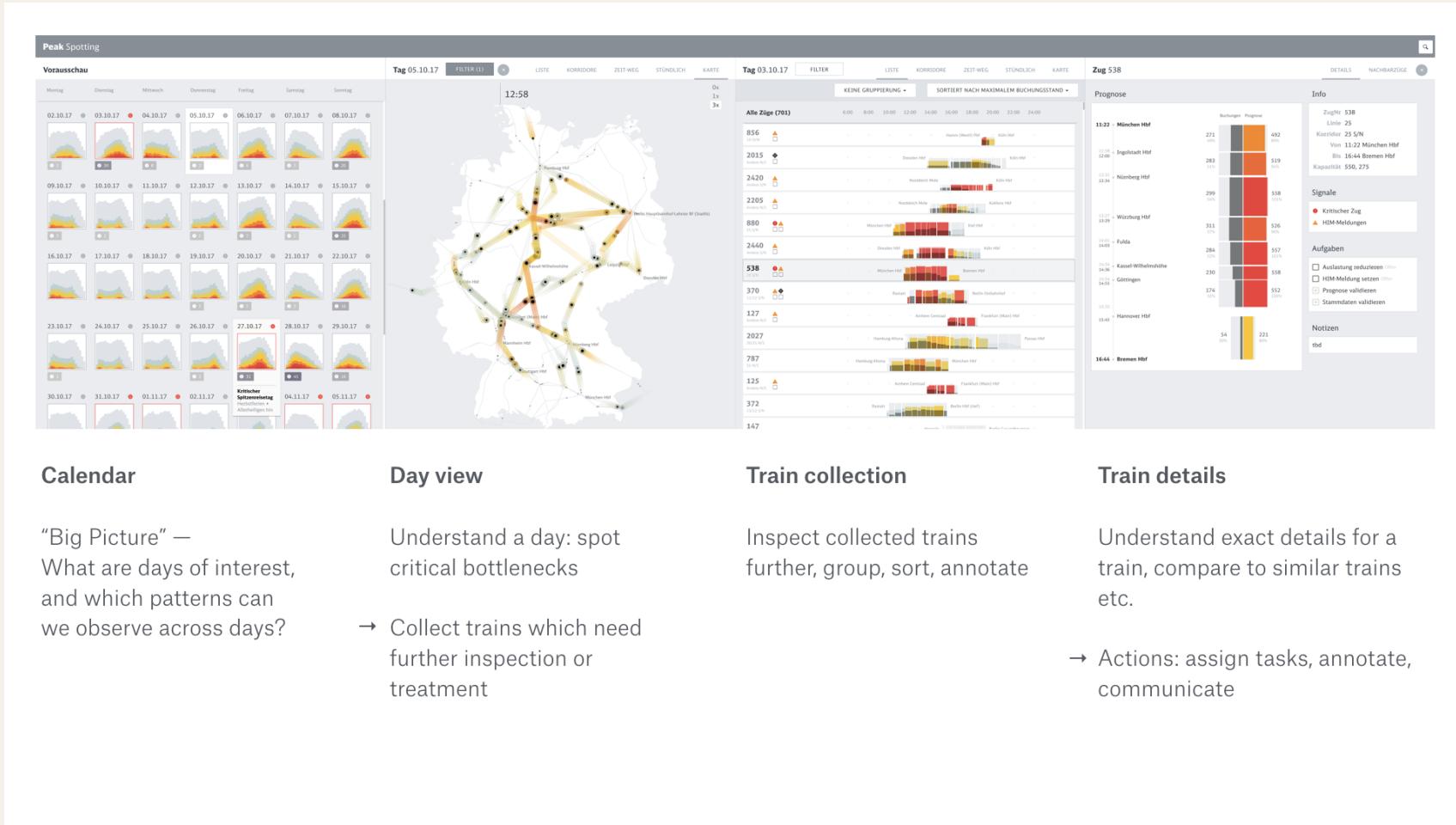
HOMBRES



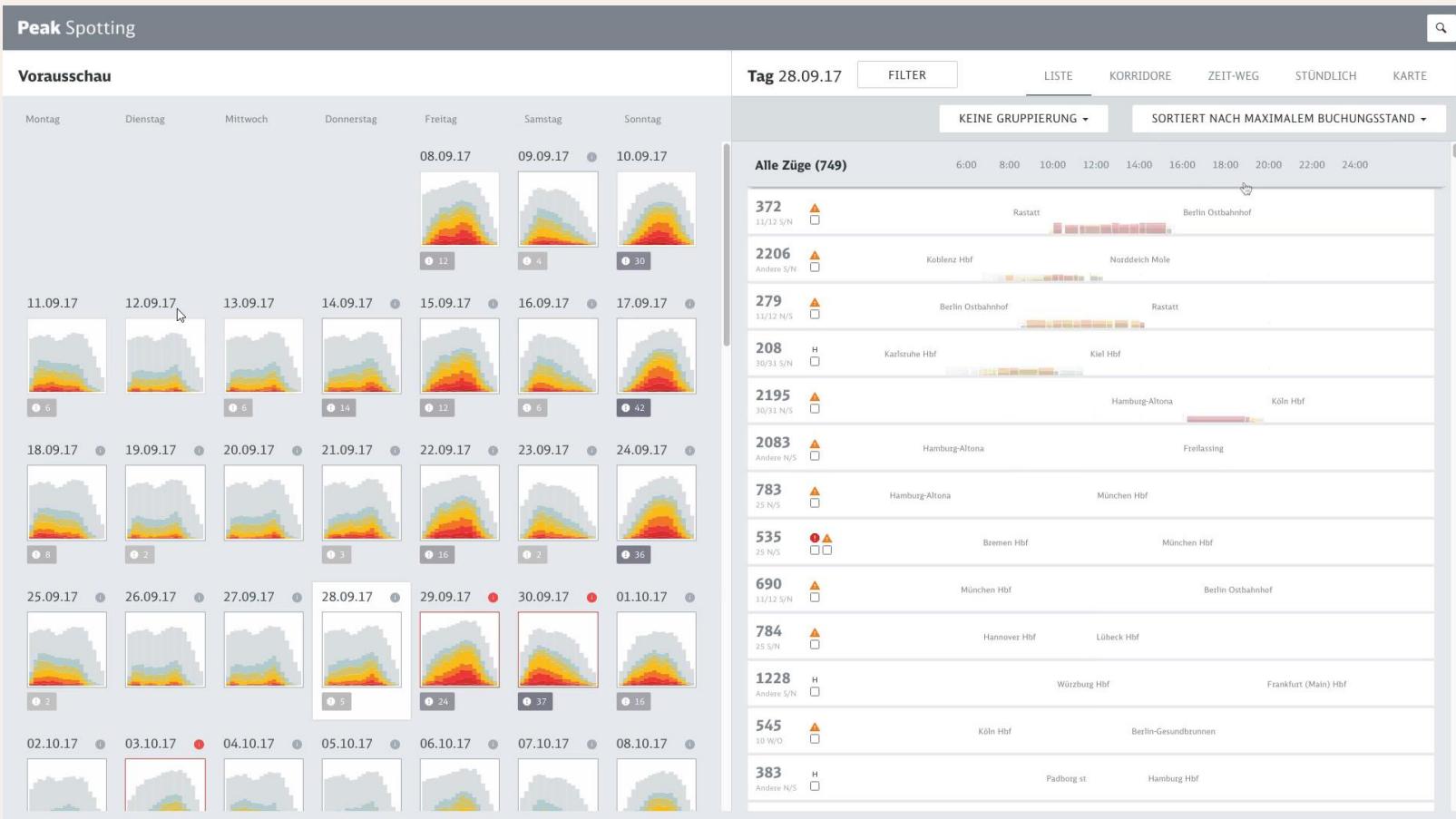
MUJERES

Fuente: Atlas de la mortalidad por cáncer en España y Portugal (2003-2012).

Peak Spotting



Peak Spotting



<https://truth-and-beauty.net/projects/peakspotting>

The COVID-NMA initiative



The COVID-NMA initiative
A living mapping and living systematic review of Covid-19 trials

COVID-NMA is an international research initiative supported by the WHO and Cochrane.

We provide a [living mapping](#) of COVID-19 trials. We are also conducting living evidence synthesis on [preventive interventions, treatments and vaccines](#) for COVID-19 to assist decision makers.

See the description of our model [here](#) and our living review protocol [here](#).

LIVING MAPPING OF TRIALS (i.e., trials registered on the WHO platform)		LIVING SYNTHESIS OF PUBLISHED STUDIES (include both articles and preprints)			
Updated weekly		Updated daily			
4075 Randomized Trials 1752 RCTs recruiting		699 Studies (RCTs or Observational studies) with complete data extraction and results included in our evidence synthesis			
3147 RCTs on treatments (1403 recruiting)	397 RCTs on prevention (134 recruiting)	531 RCTs on vaccines (215 recruiting)	458 RCTs on treatments	17 RCTs on prevention	129 RCTs on vaccines 95 observational studies on vaccines

The COVID-NMA initiative

As of February 16, 2022 the Covid-19 - living NMA initiative collected 4004 studies of treatments from the [ICTRP](#). 1731 of these trials are recruiting patients.

About us ▾

Living Mapping ▾

COVID-19 treatments ▾

Vaccines ▾

Preventive treatments

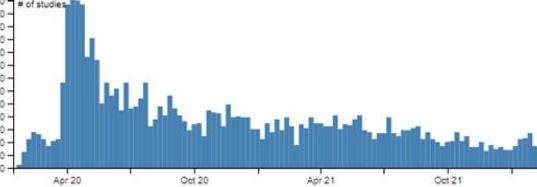
[Twitter](#) [Email](#)

Filters

All trials selected (4004) | Reset all

Search...
Ex: Interferon, antiviral, Spain, Assistance Publique, EUCTR2020...

Registration date

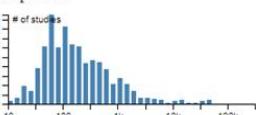


To filter by Registration dates, click and drag to create a range.

Inclusion age

- minimum 18 yo (3,683 studies)
- less than 18 yo (269 studies)
- N/A (52 studies)

Sample size



Recruitment status

- Recruiting (1,731 studies) ⓘ
- Not recruiting (1,380 studies) ⓘ
- Completed (686 studies) ⓘ
- Terminated (131 studies) ⓘ
- Withdrawn (60 studies) ⓘ
- Suspended (15 studies) ⓘ
- Not reported (1 study)

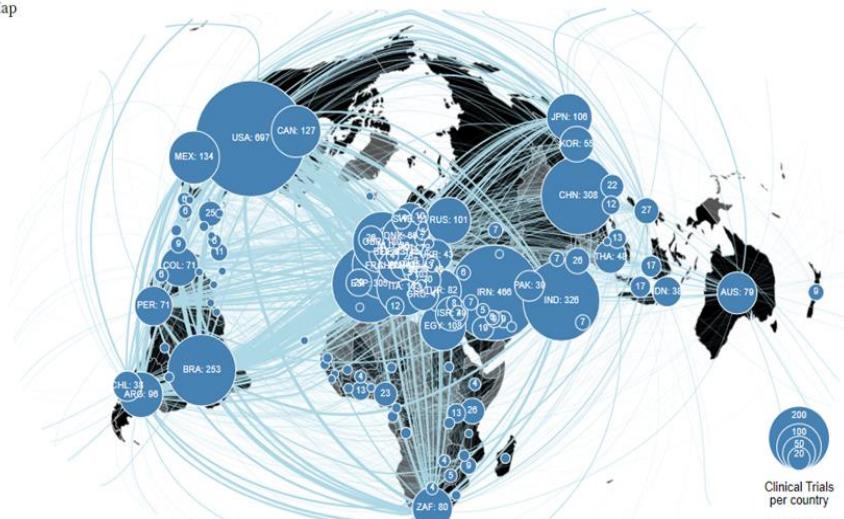
Publication status

- Not published (3,387 studies)
- Published (617 studies) ⓘ

Registry status

- No results posted (3,869 studies)
- Results posted in the registry (135 studies)

Map



Clinical Trials per country
Partnerships

The COVID-NMA initiative



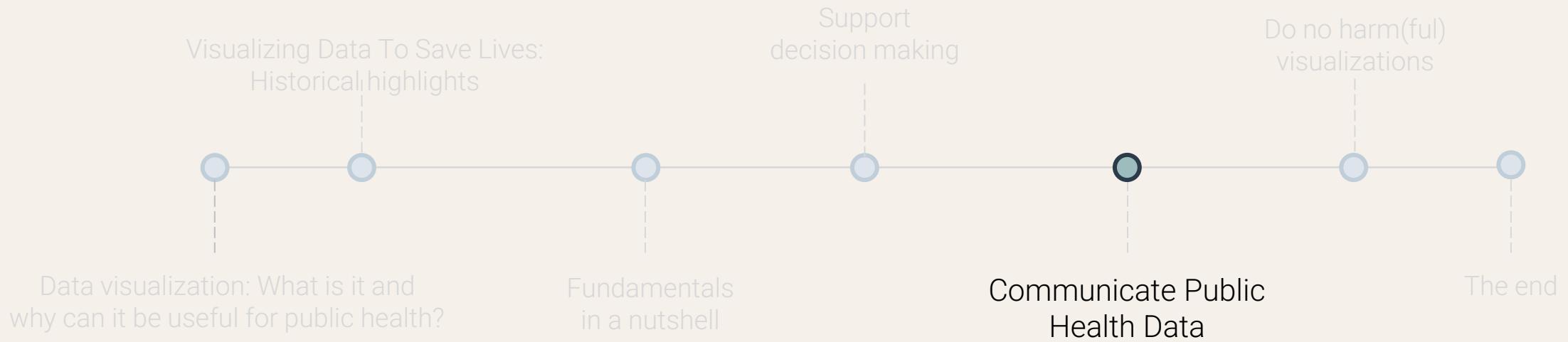
DATA VISUALIZATION
MEETS PUBLIC HEALTH:
THE COVID-NMA PROJECT

**OPEN DISCUSSION
24. MARCH**

MD, PhD, Director, Cochrane France,
Co-convenor, Cochrane Bias Methods Group

ISABELLE BOUTRON

Agenda

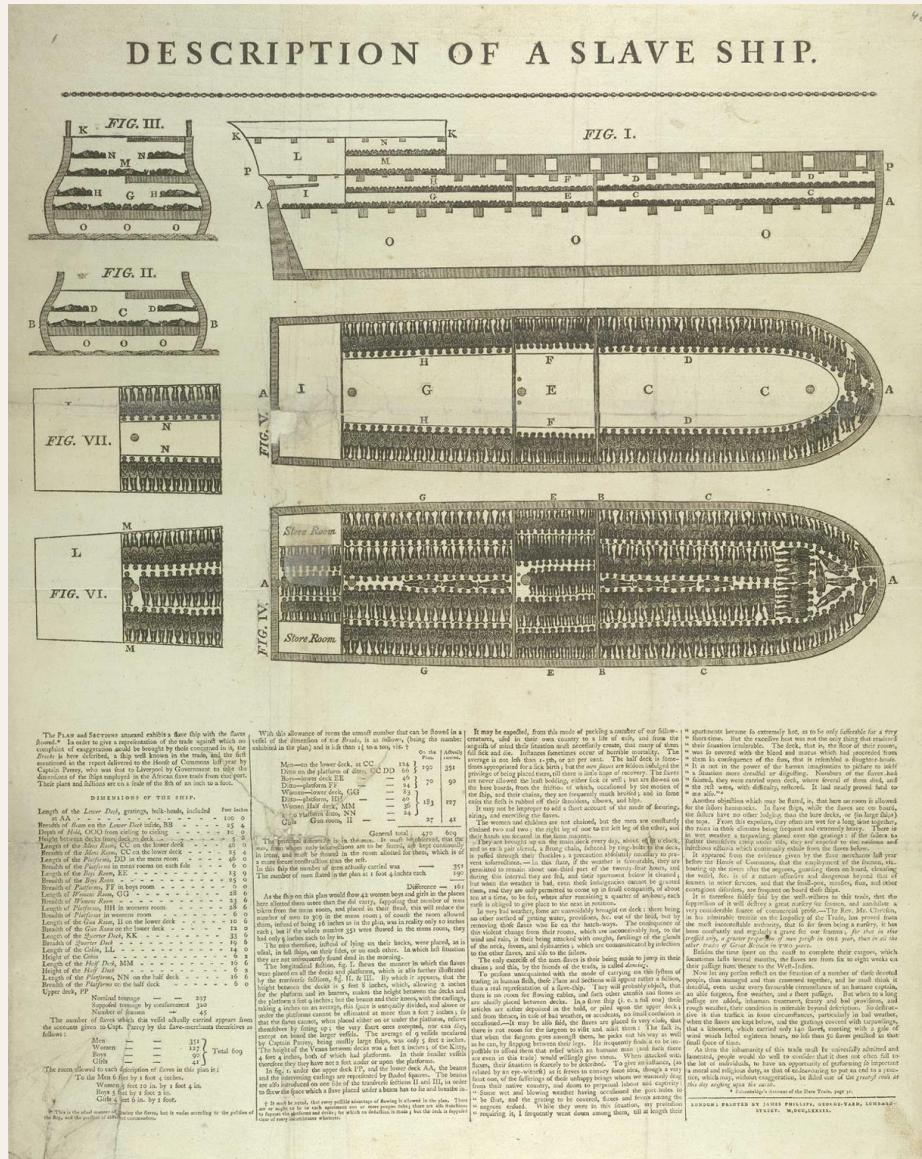


“Data visualizations help to create a narrative around an idea, and it’s the **narrative** that ultimately has the ability to **change people’s hearts and minds.**”

(Sarah Williams, *Data Action* , 2020)

A map created as an awareness campaign about the horrors of slavery and to persuade those in governmental power to abolish it.

Brookes Slave Ship Map, 1788. Source: "Diagram of a Slave Ship", 1801, <http://www.bl.uk/learning/timeline/item106661.html>.

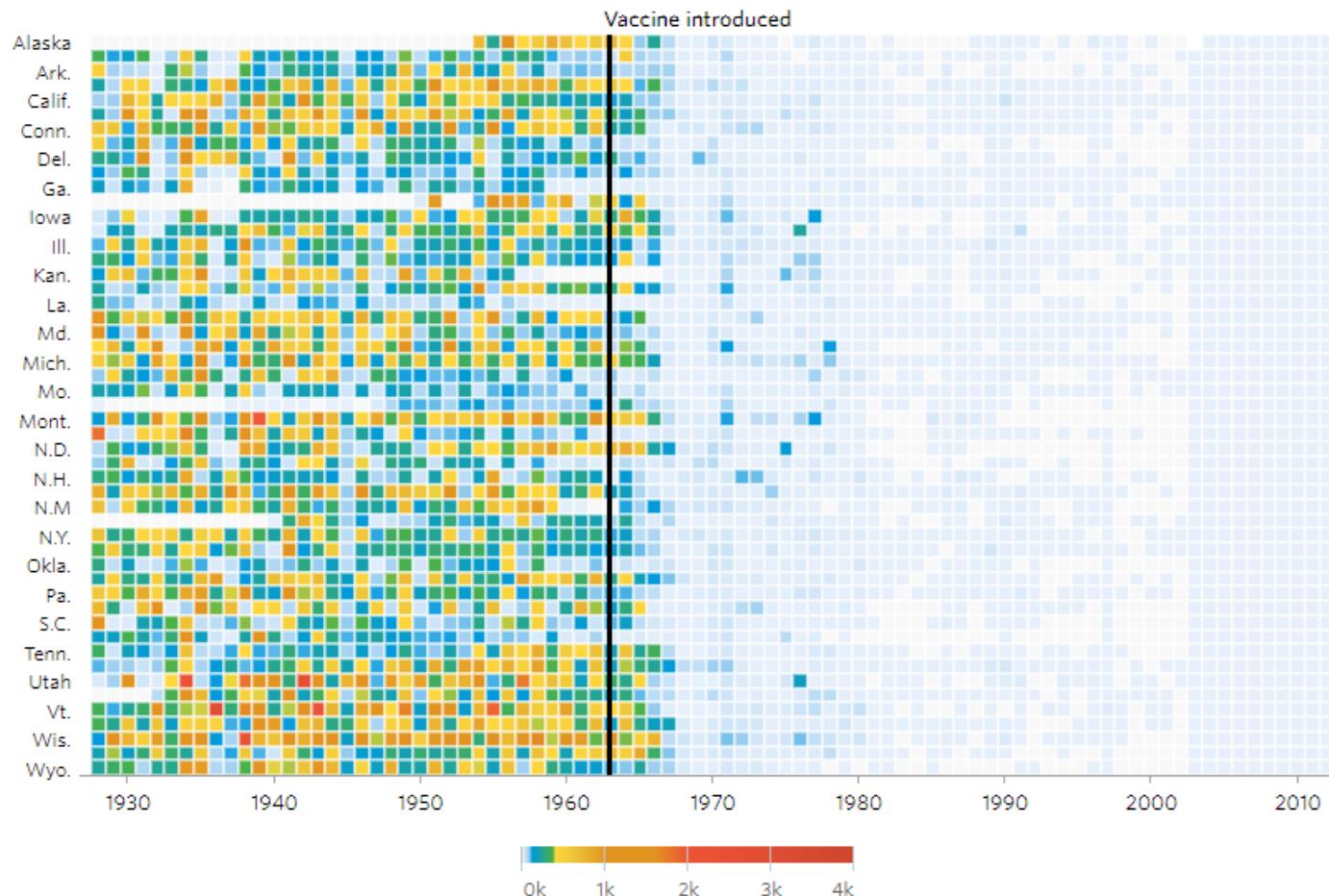


Battling Infectious Diseases in the 20th Century: The Impact of Vaccines

By [Tynan DeBold](#) and [Dov Friedman](#)

Published Feb. 11, 2015 at 3:45 p.m. ET

Measles



“Don’t just show, explain.”

“The greatest value of a picture is when it forces us to notice what we never expected to see”¹

“Good charts reveal realities that may otherwise go unnoticed”²



Hans Rosling: <https://www.gapminder.org/>

[1] – John W. Tukey; http://www.ru.ac.bd/wp-content/uploads/sites/25/2019/03/102_05_01_Tukey-Exploratory-Data-Analysis-1977.pdf

[2] - <http://www.thefunctionalart.com/p/reviews.html>



Gapminder

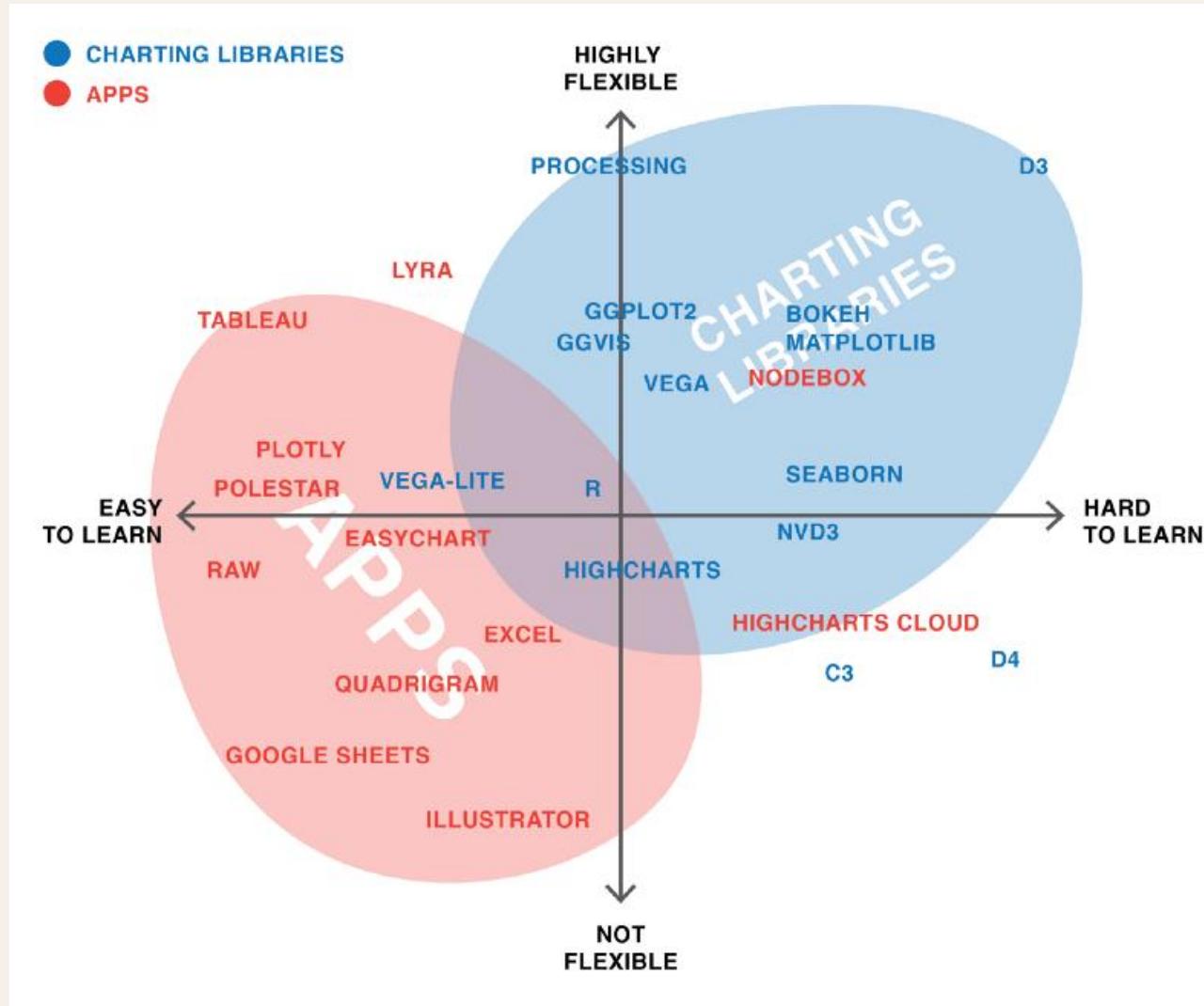
transparency = similarity

size = population

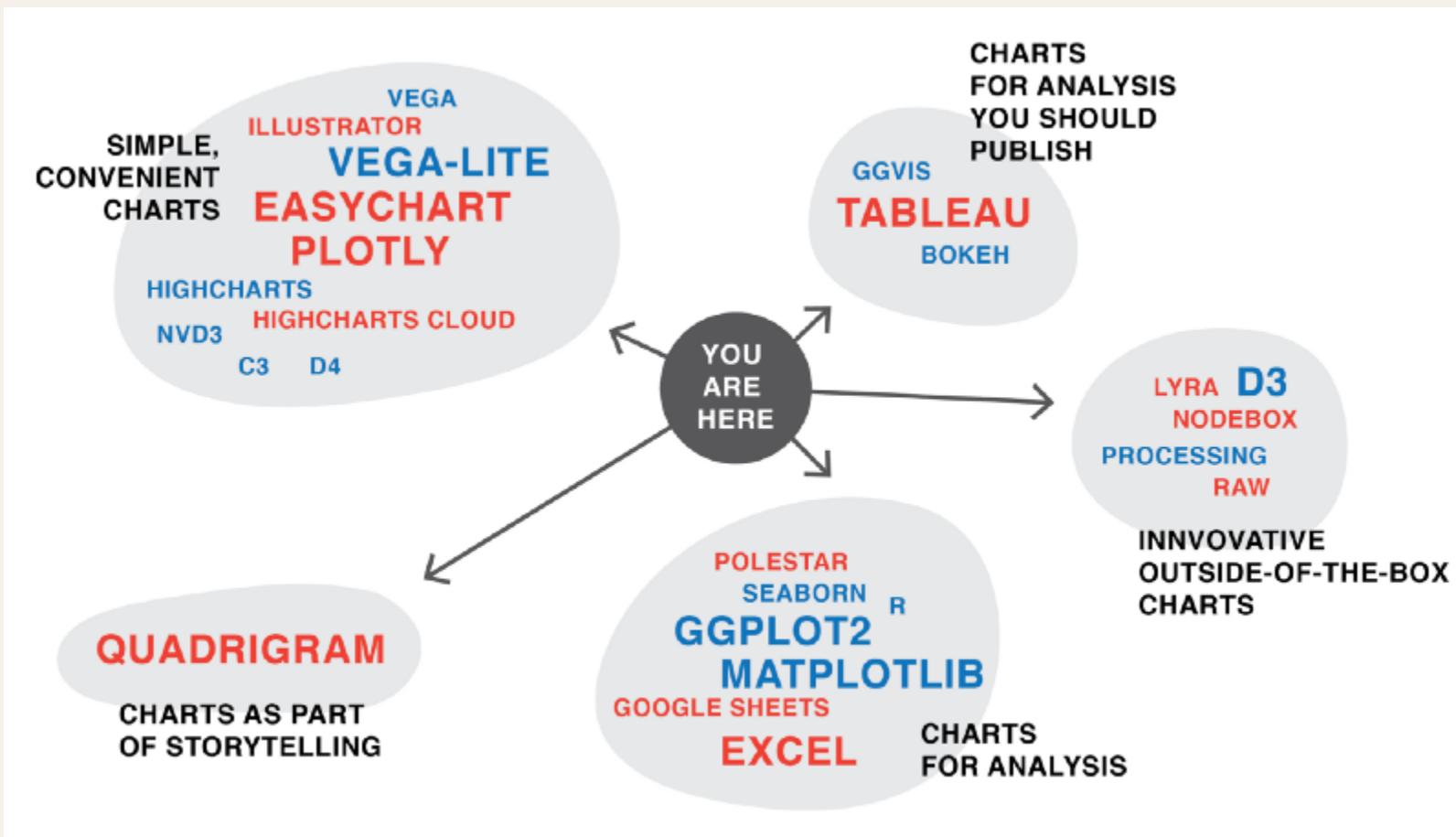
color = continent



Tools



Tools



Powerful tool to spark conversations



Community groups mapped hand-collected bicycle traffic data, community gardens, churches, magnet schools, and food banks. Groups of students mapped their schools, how they got there, and where it was or wasn't safe.

Once a map was drawn, a projection representing everything from bus lines to poverty levels to historical red-lining maps was made. The mappers were surprised to see their experiences confirmed - or challenged - by data that were invisible to them before.

it's a model that can be replicated by communities everywhere so that mapmaking can be a communal, human act.

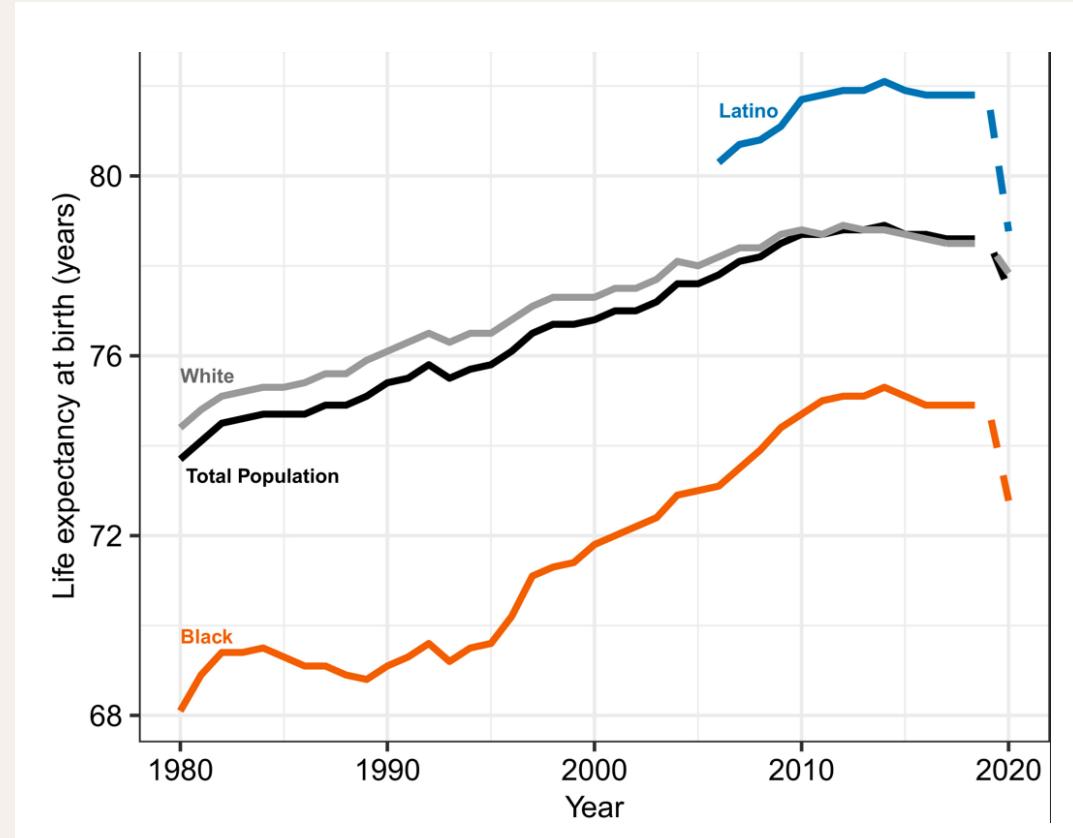
Jer Thorp, *Living in data*, 2021

“These are not just data points, and so when you add them as dots on a map, or lines on a chart, I think **you have to remember that these are true people.**”

(Tim Meko, *Washington Post*)

Practice Empathy

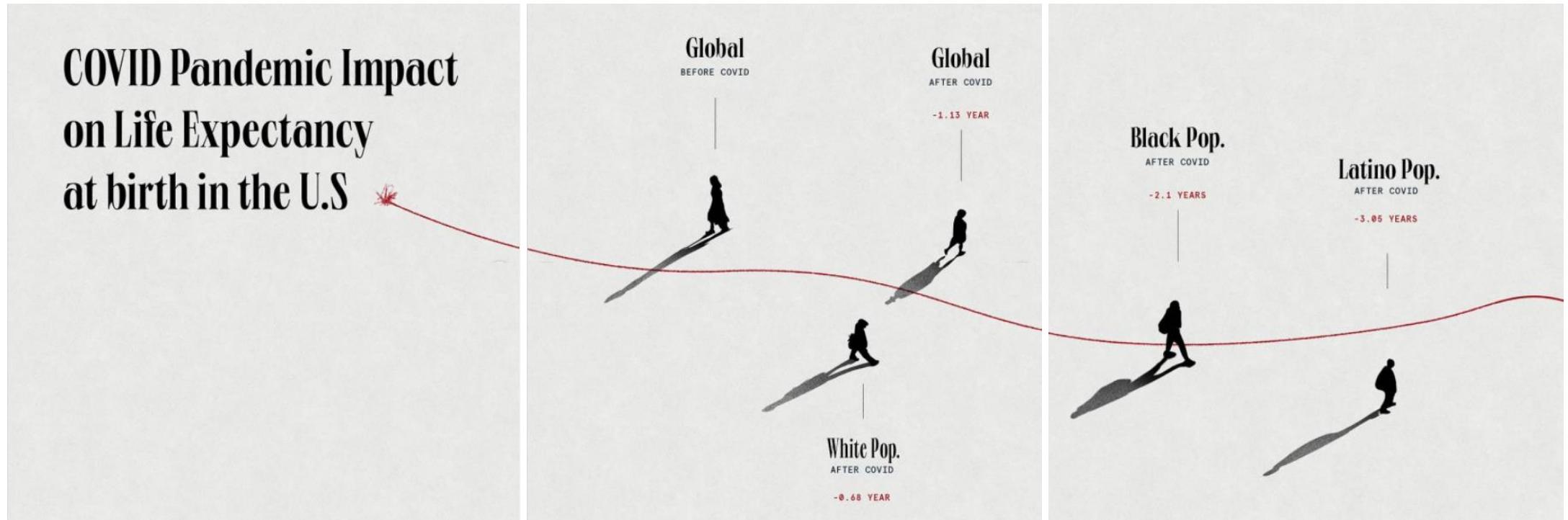
From this



Andrasfay, T., & Goldman, N. (2021). Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations. Proceedings of the National Academy of Sciences, 118(5).

Practice Empathy

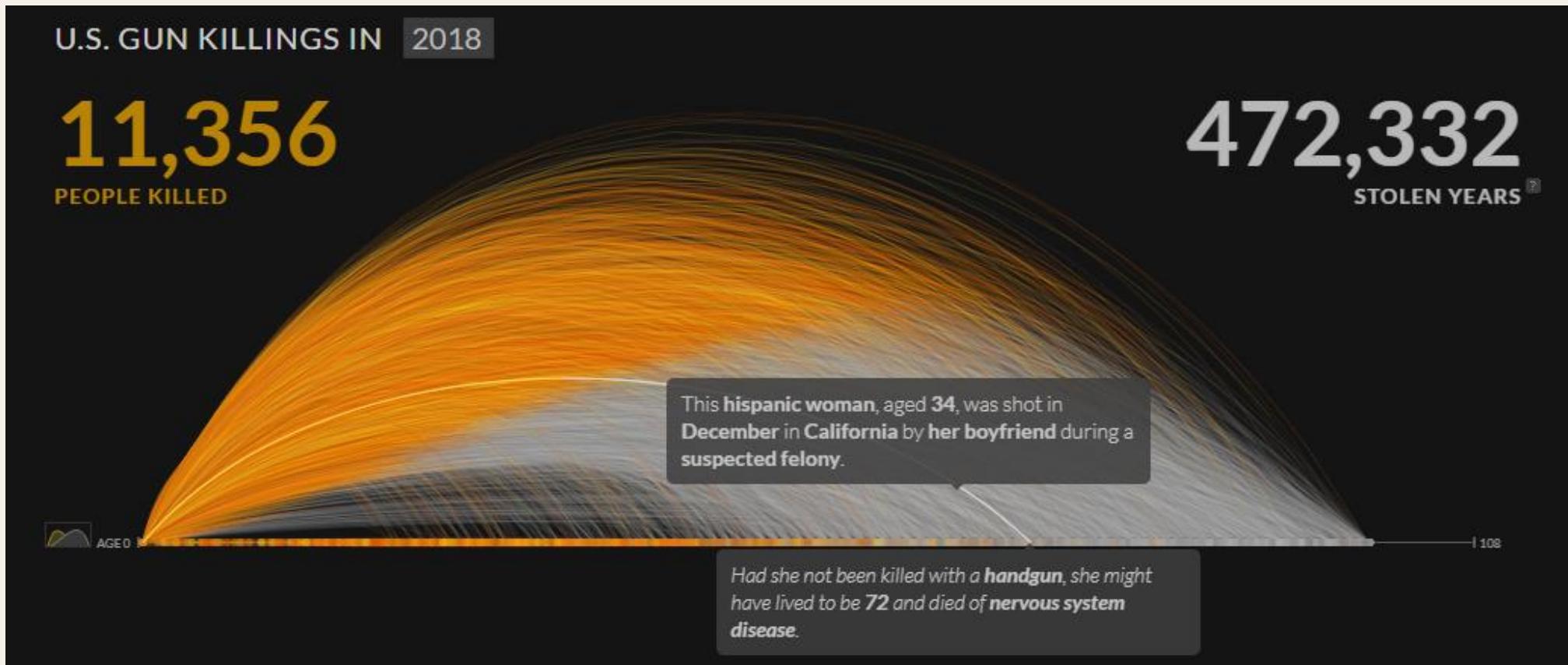
To this



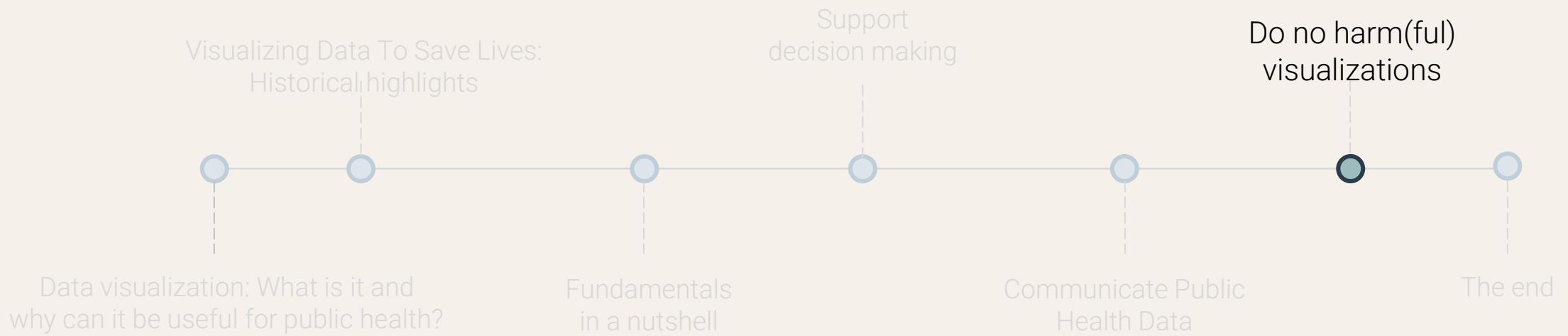
<https://www.gabriellemerte.com/>

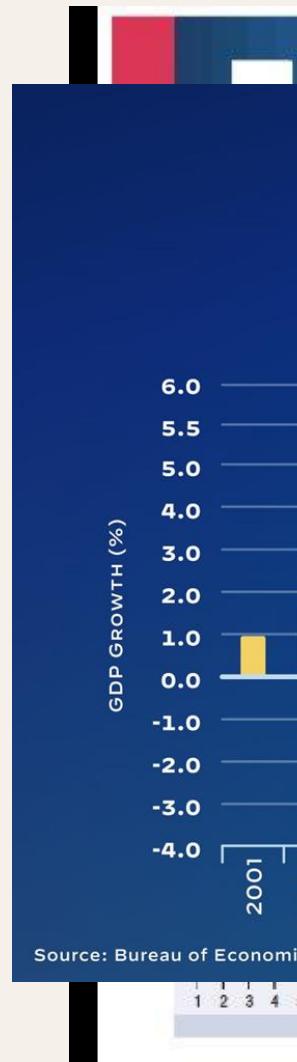
https://www.instagram.com/p/CLh_SdWBu9V/

Practice Empathy



Agenda





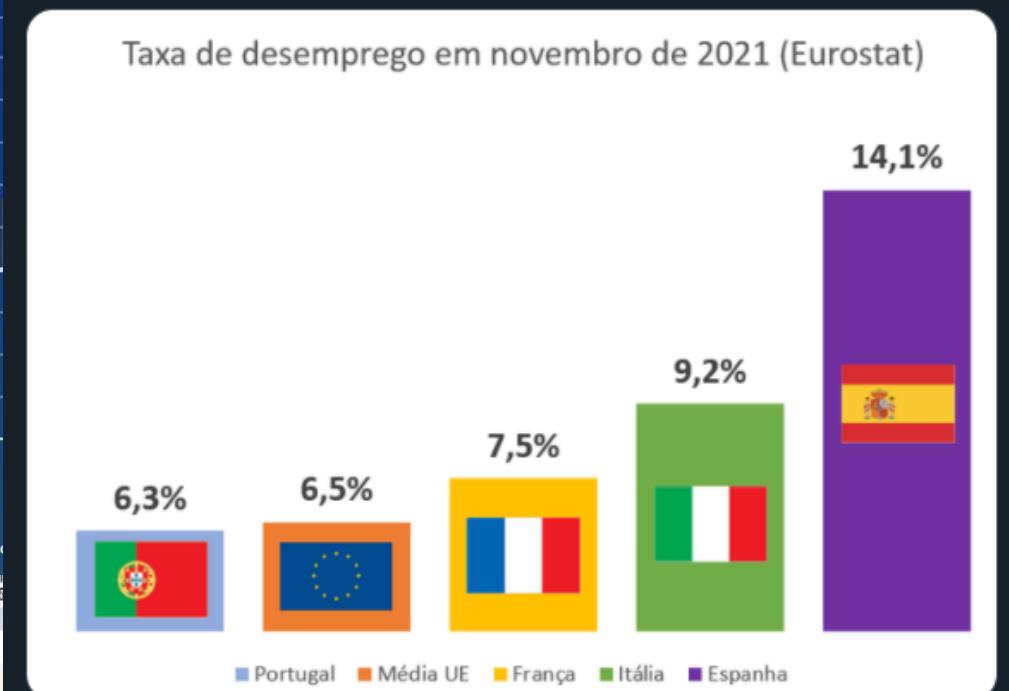
Trabalho PT ✅
@trabalho_pt

✓ A taxa de desemprego em novembro, que se cifrou em 6,3% e que atingiu o valor mais baixo neste mês desde 2001, está abaixo da média europeia, que é de 6,5%

✓ Portugal regista um desemprego inferior à maioria dos países do sul da Europa

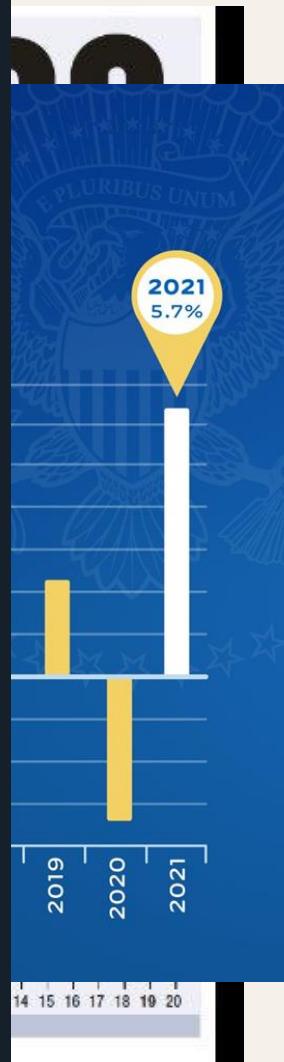
#MTSSS

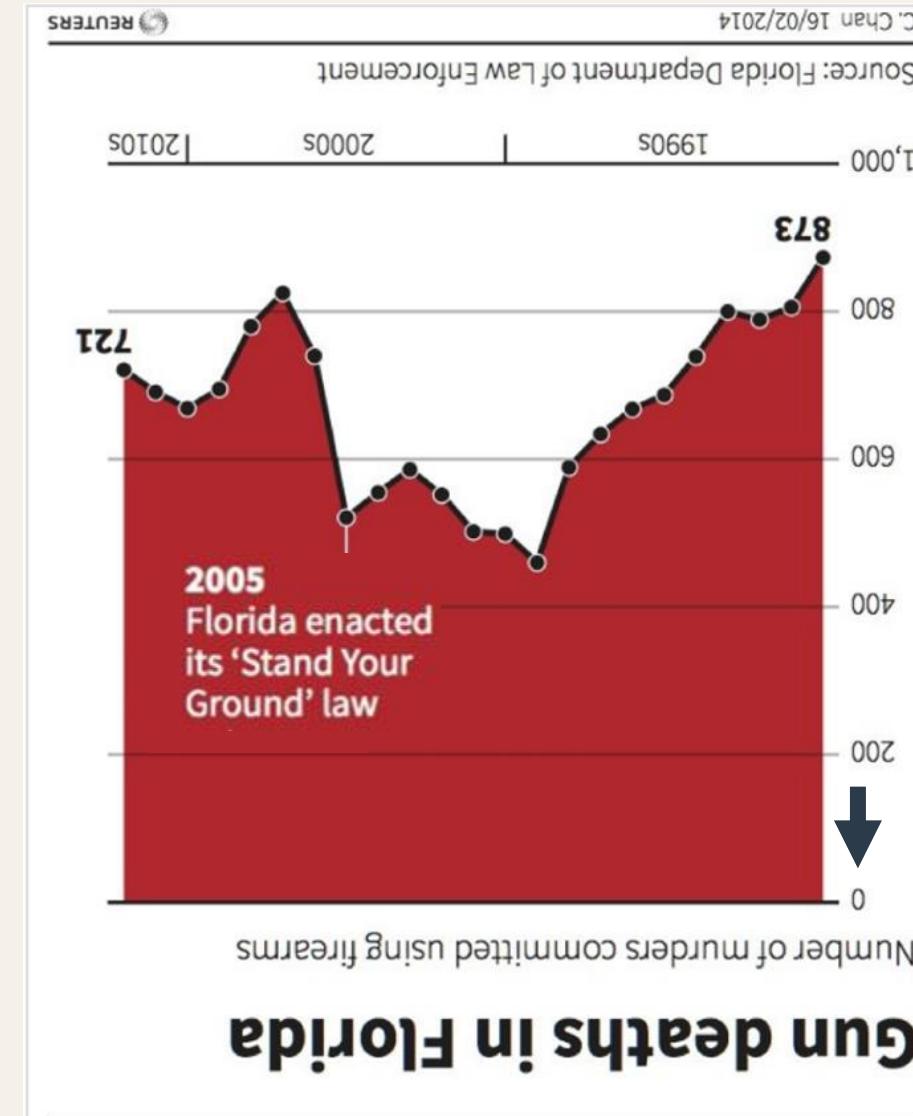
Taxa de desemprego em novembro de 2021 (Eurostat)



País	Taxa de desemprego (%)
Portugal	6,3%
Média UE	6,5%
França	7,5%
Itália	9,2%
Espanha	14,1%

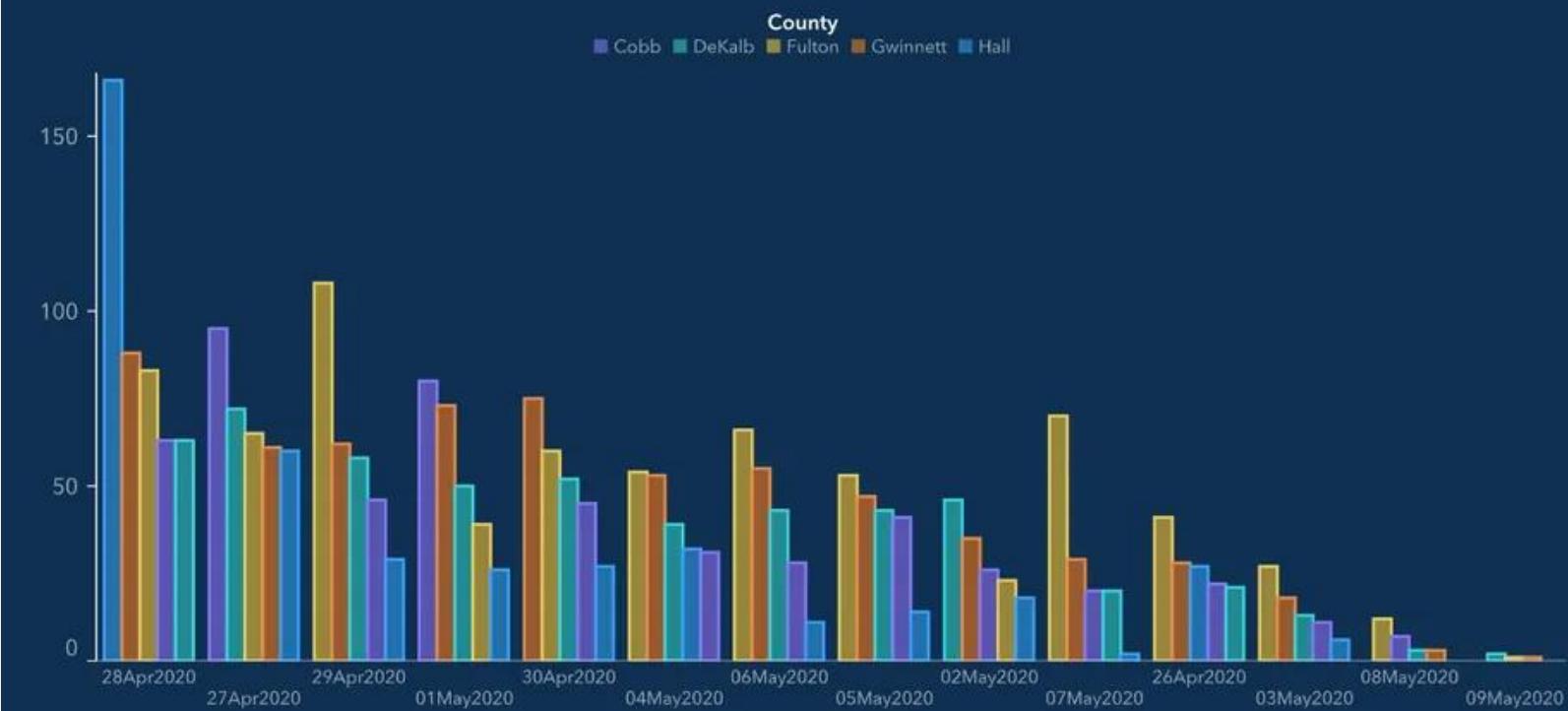
República Portuguesa





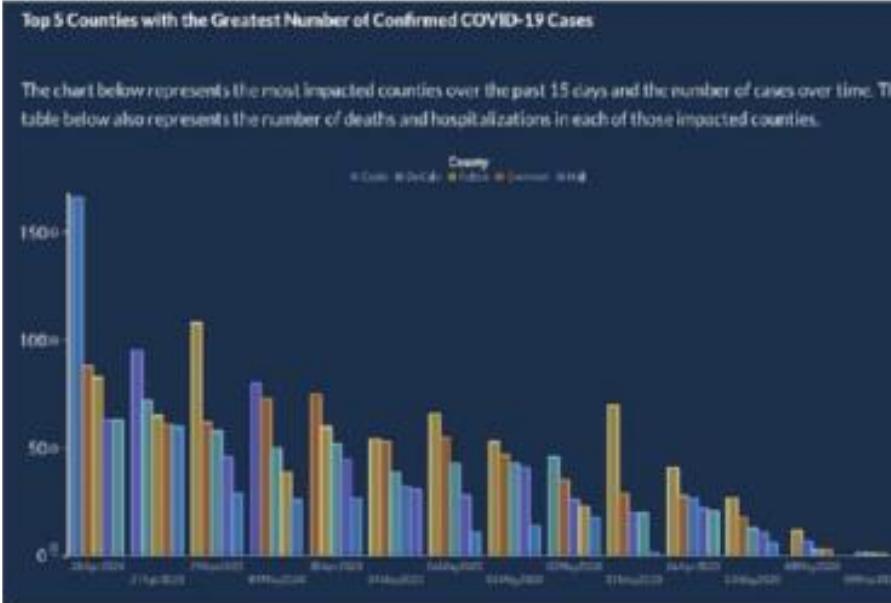
Top 5 Counties with the Greatest Number of Confirmed COVID-19 Cases

The chart below represents the most impacted counties over the past 15 days and the number of cases over time. The table below also represents the number of deaths and hospitalizations in each of those impacted counties.

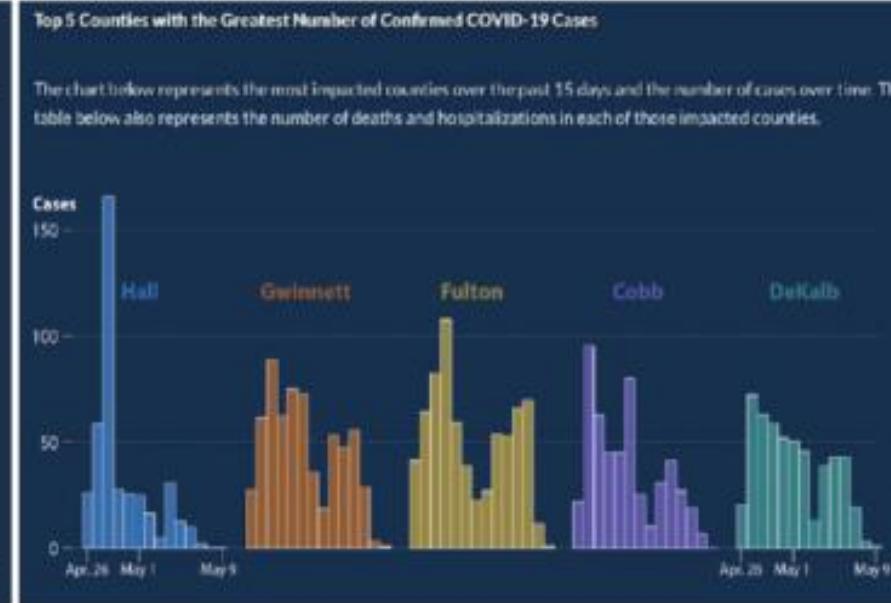


A screenshot of the chart published on Georgia's Covid-19 dashboard in May.

ORIGINAL



MAKEOVER



Always ask:

- How was this data generated?
- Am I communicating it ethically?
- Who is included and who is excluded from this data?
- Why were these data collected?
- Who stands to benefit from this data?
- Who might be harmed by the collection or publication of this data?

Additional Resources:

Books to consider:

- Alberto Cairo, How Charts Lie: Getting Smarter about Visual Information, W. W. Norton & Company; 1st edition, 2020
- Alberto Cairo, The Truthful Art: Data, Charts, and Maps for Communication, New Riders, 2016
- Cole Nussbaumer Knaflic, Story Telling with Data: A Data Visualization Guide for Business Professionals, Wiley, 2015
- Jer Thorp, Living in Data: A Citizen's Guide to a Better Information Future, 2021
- Jonathan Schwabish, Better Data Visualizations: A Guide for Scholars, Researchers, and Wonks, 2021
- Tamara Munzner, Visualization Analysis and Design: Principles, Techniques, and Practice, 2014

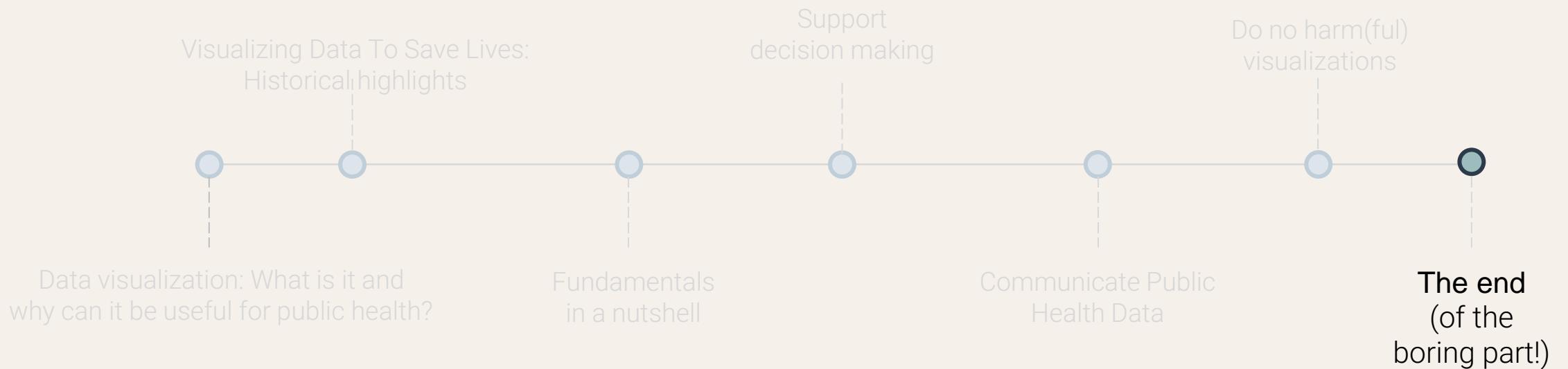
Online resources:

- Data Visualization Society: <https://www.datavisualizationsociety.org/>
- A blog about visualization research: <https://medium.com/multiple-views-visualization-research-explained>

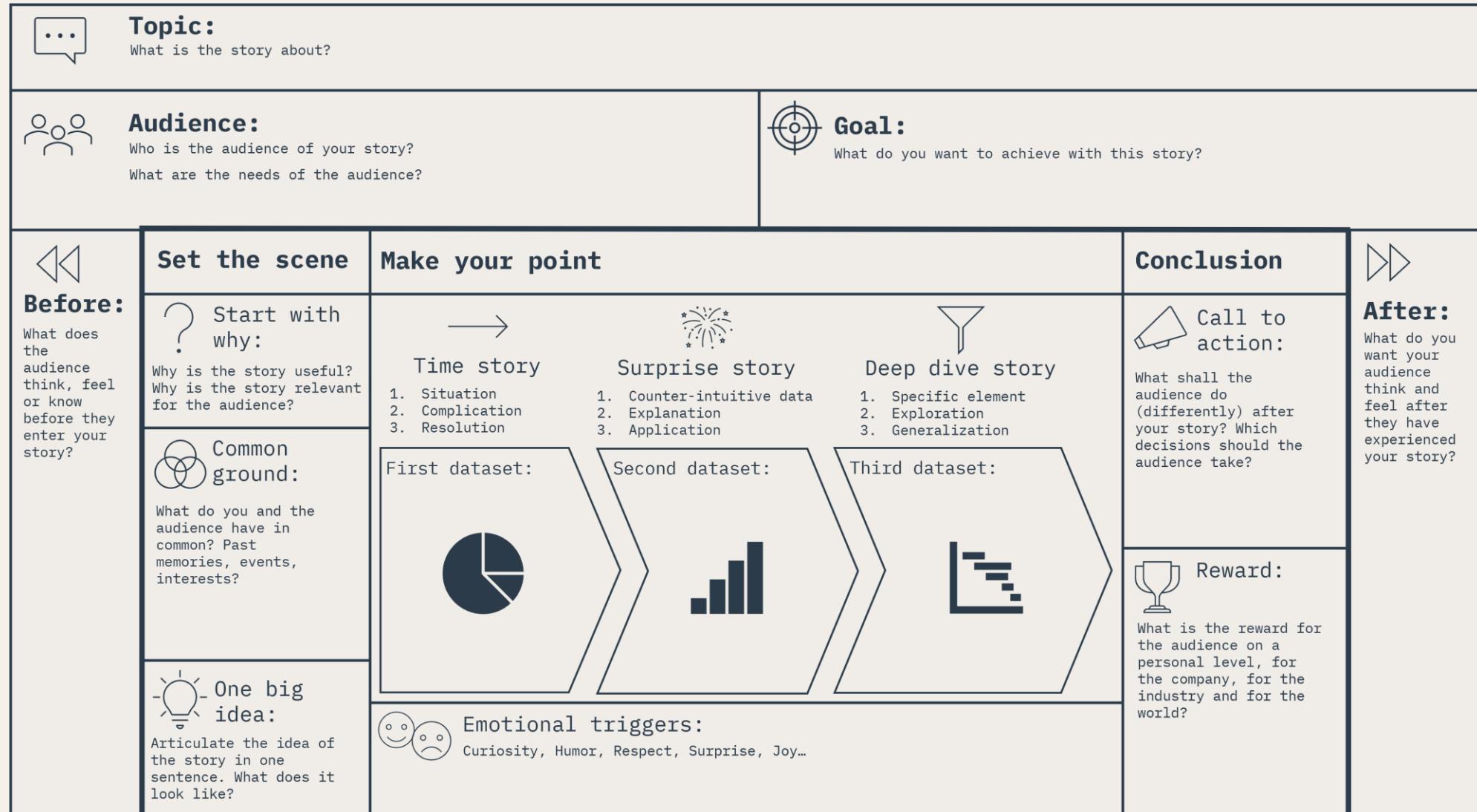
Scientific articles:

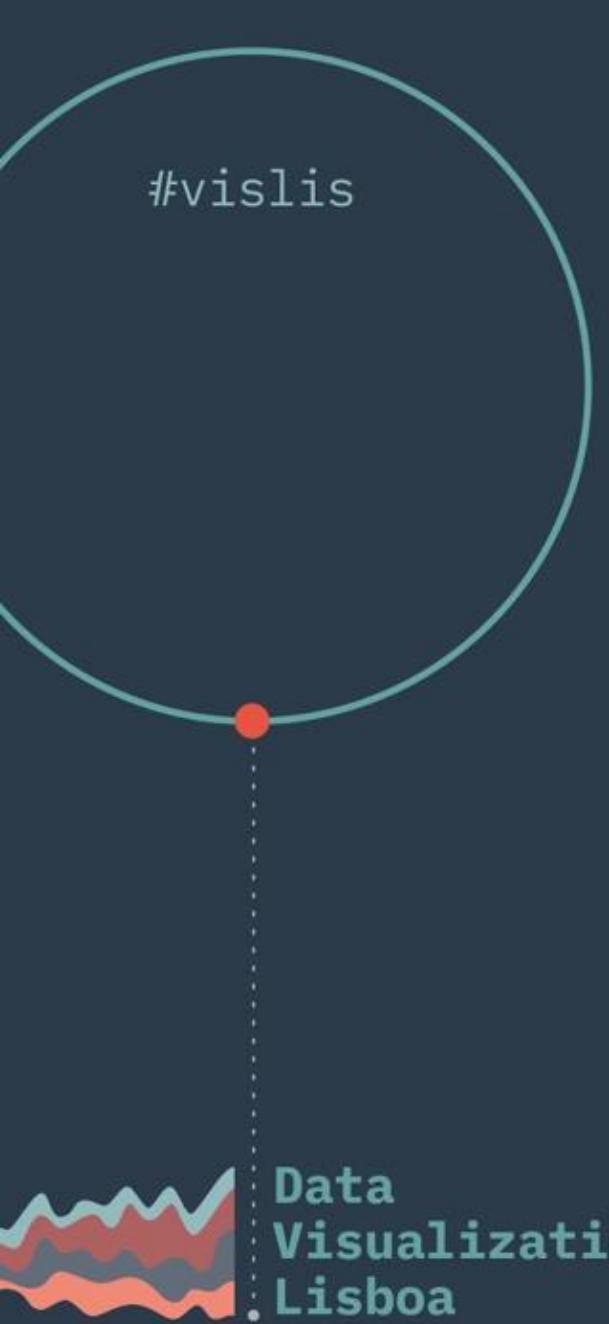
- Zook, Matthew et al. "Ten simple rules for responsible big data research." PLoS computational biology vol. 13,3 e1005399. 30 Mar. 2017, doi:10.1371/journal.pcbi.1005399
- Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The science of visual data communication: What works. Psychological Science in the Public Interest, 22(3), 110-161.

Agenda



Data Storytelling Canvas





#vislis

Thank you!

hello@datavis-lisboa.pt



Data
Visualization
Lisboa