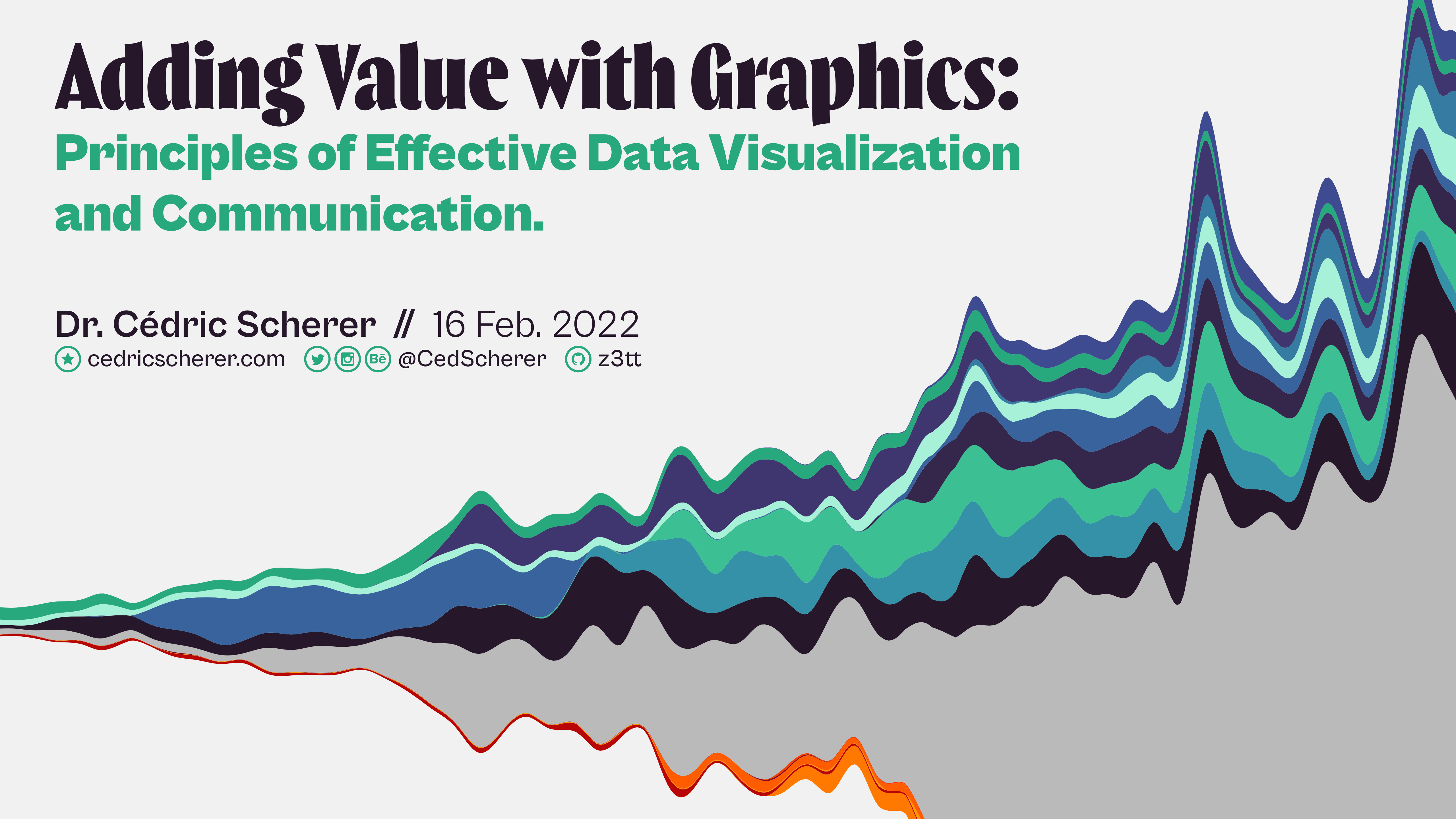


# Adding Value with Graphics:

## Principles of Effective Data Visualization and Communication.

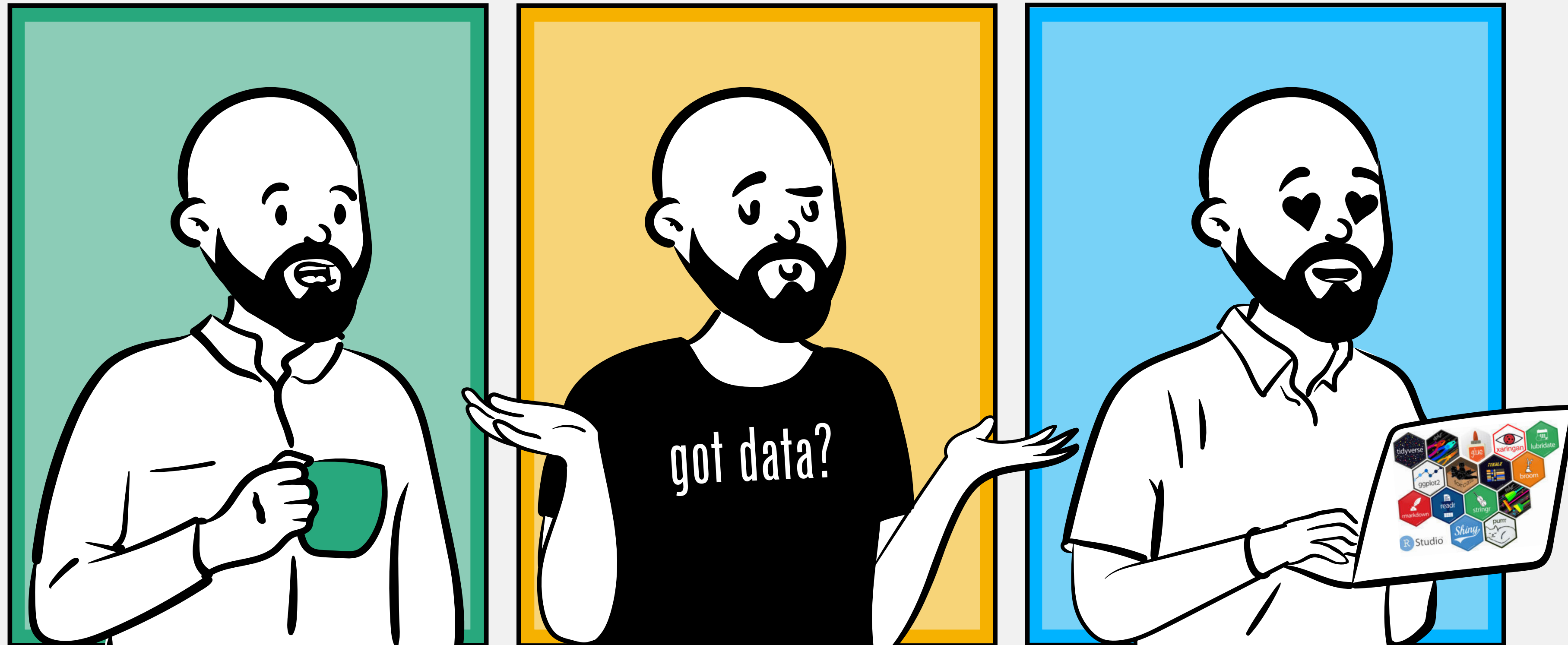
Dr. Cédric Scherer // 16 Feb. 2022

★ [cedricscherer.com](http://cedricscherer.com)    @CedScherer  z3tt



# Cédric Scherer

Independent Data Visualization Specialist  
Computational Ecologist at IZW Berlin

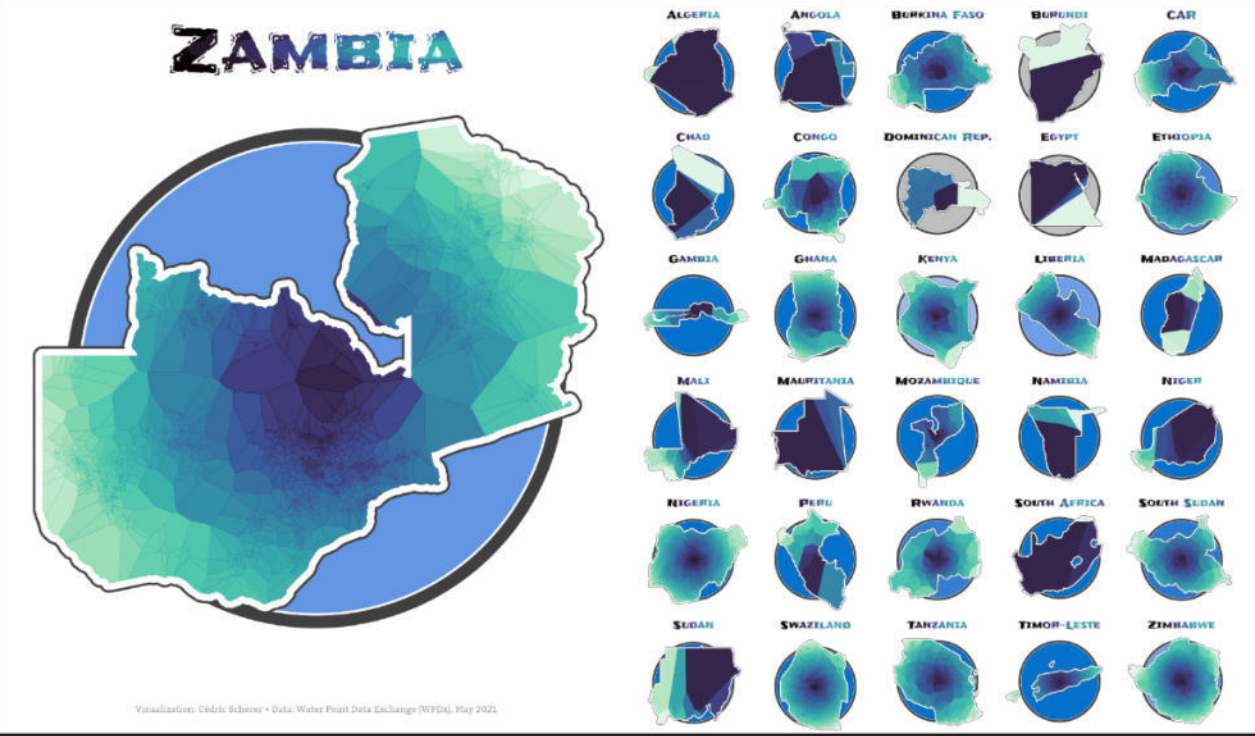
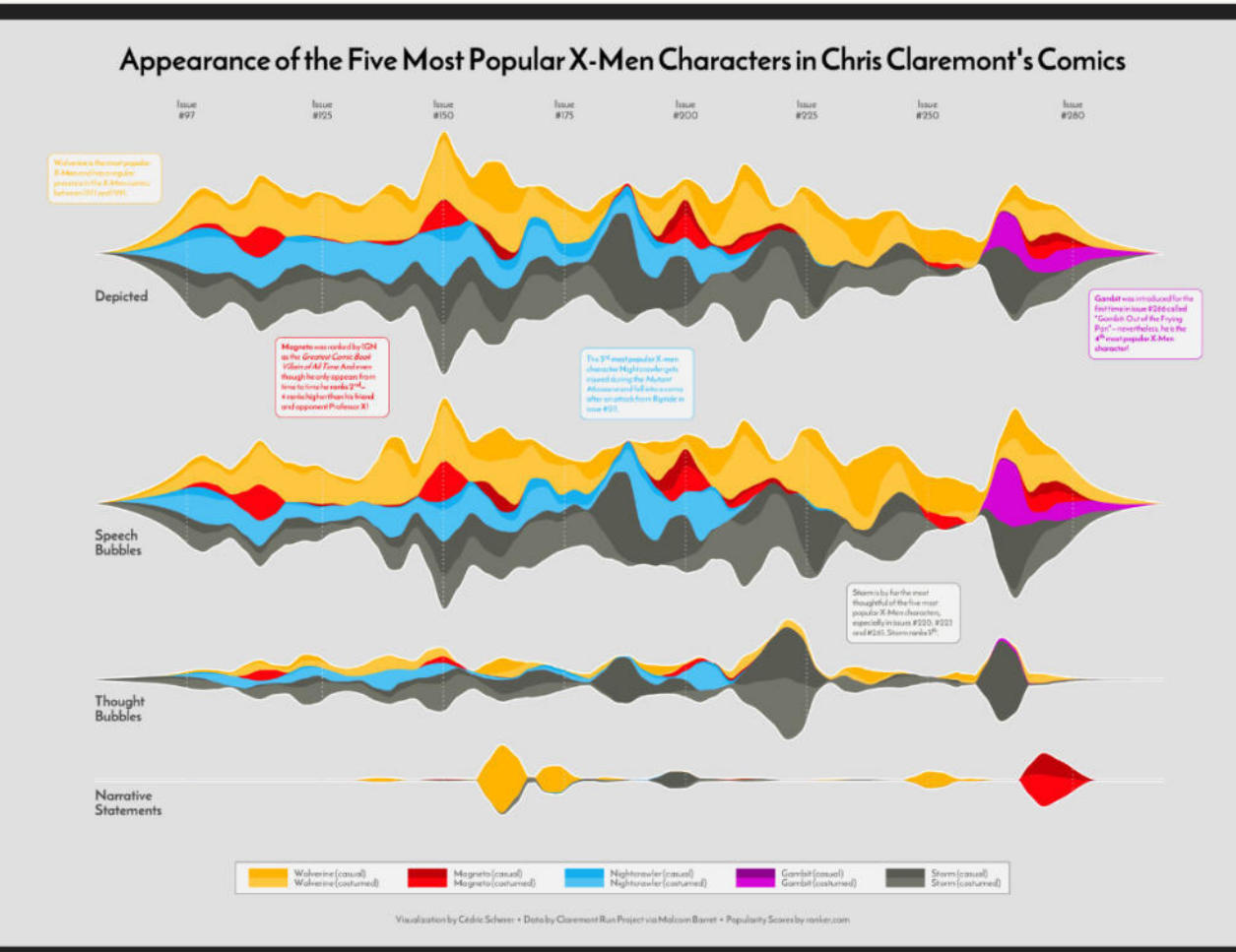
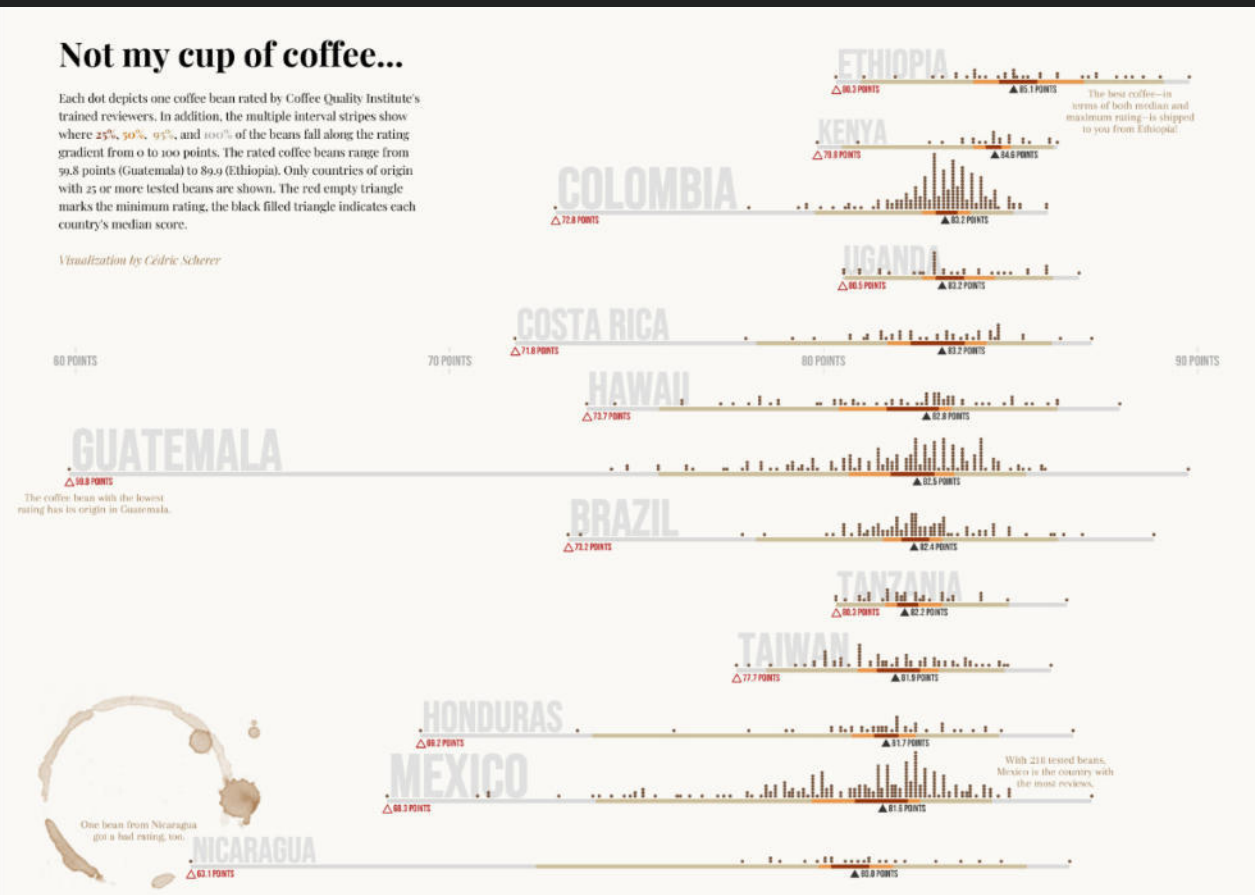
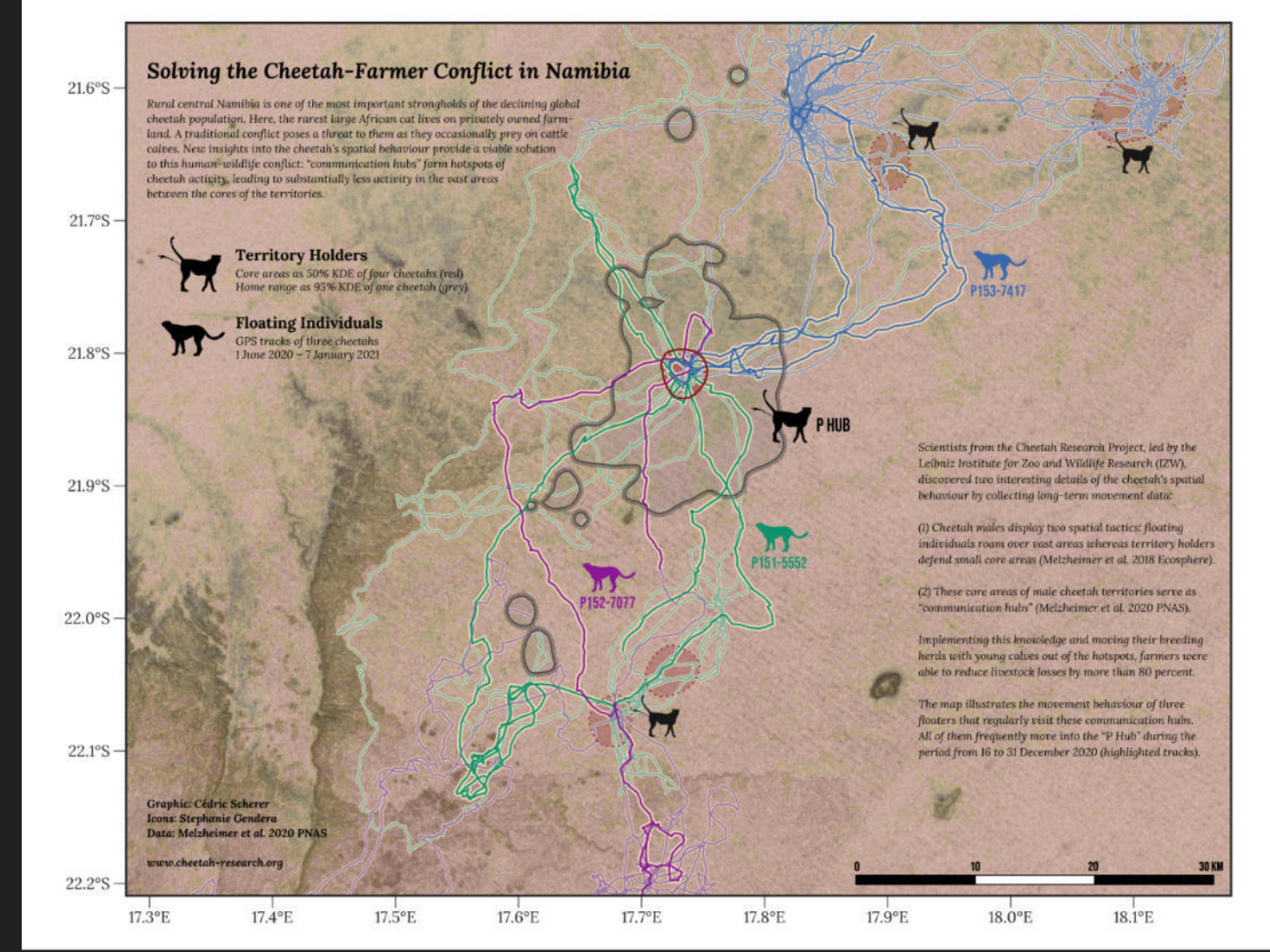
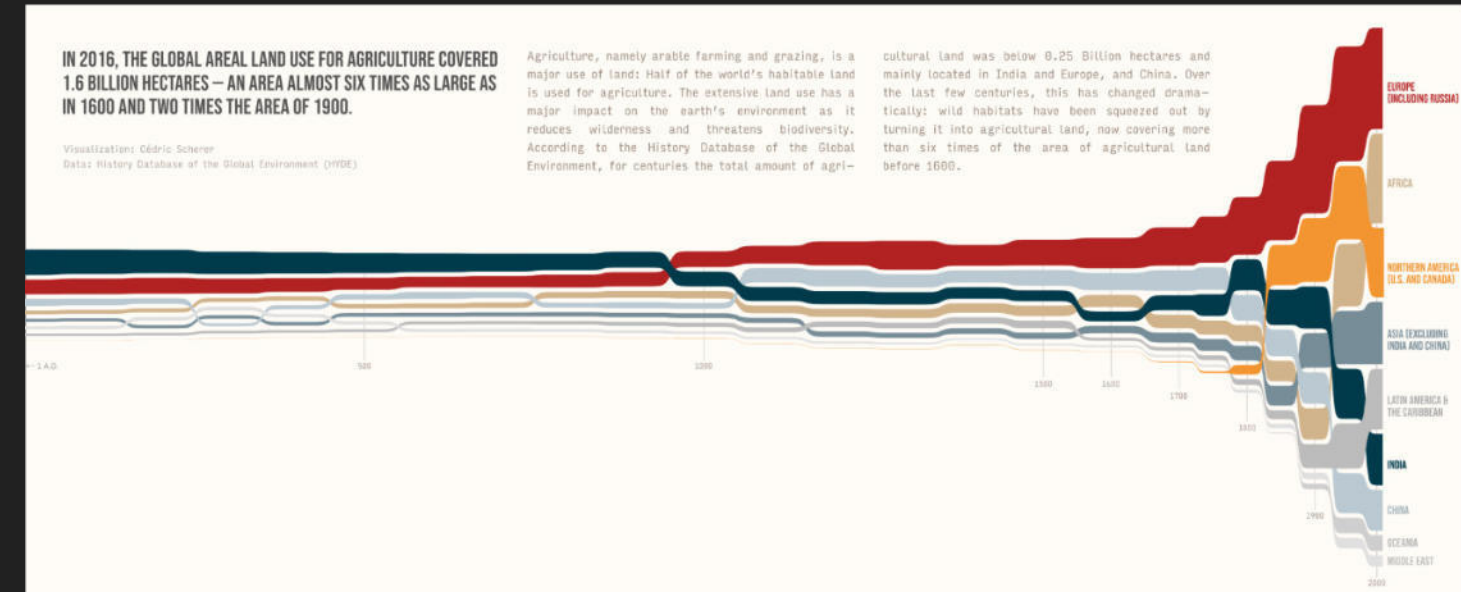
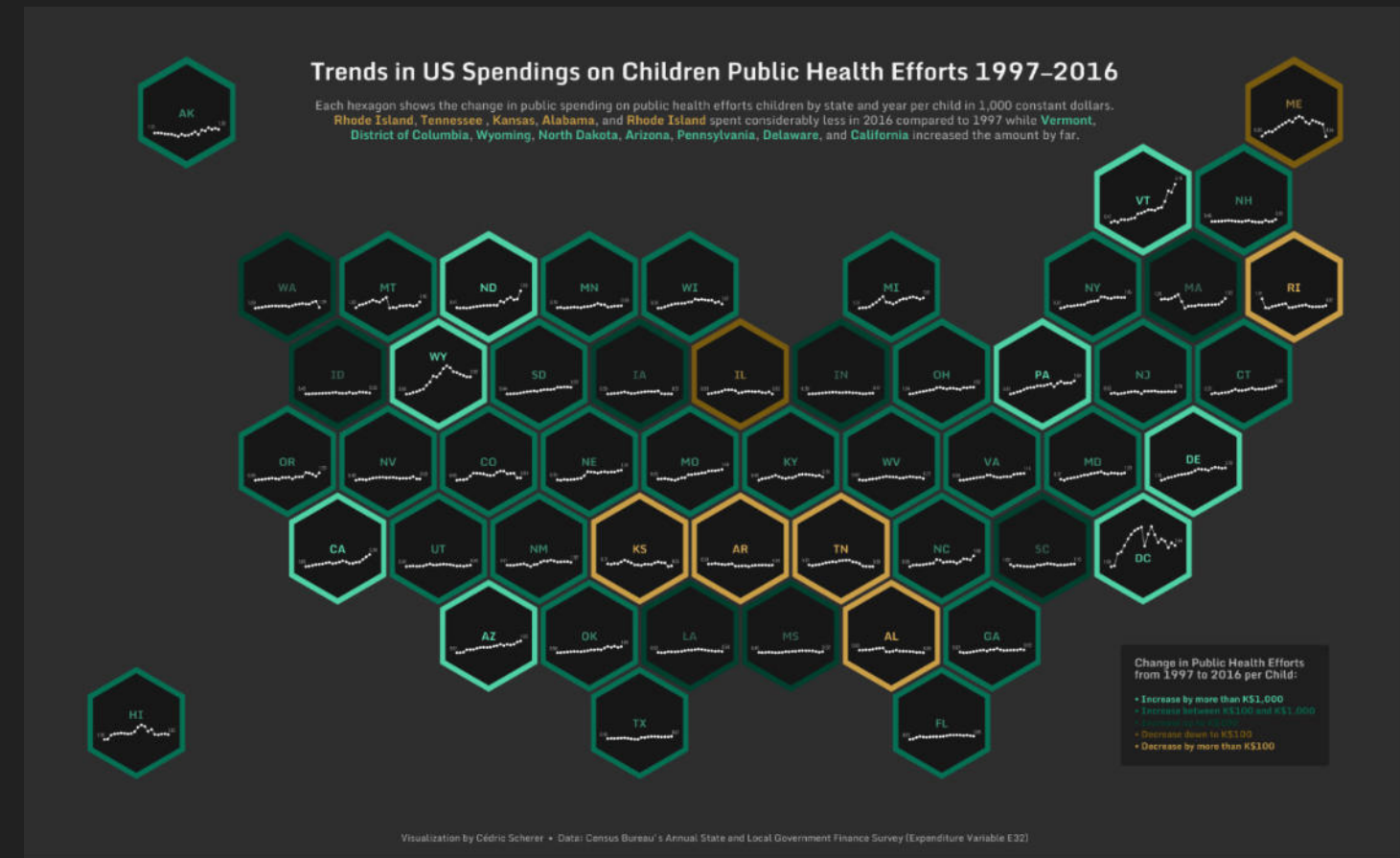


Consulting

Coaching

Coding



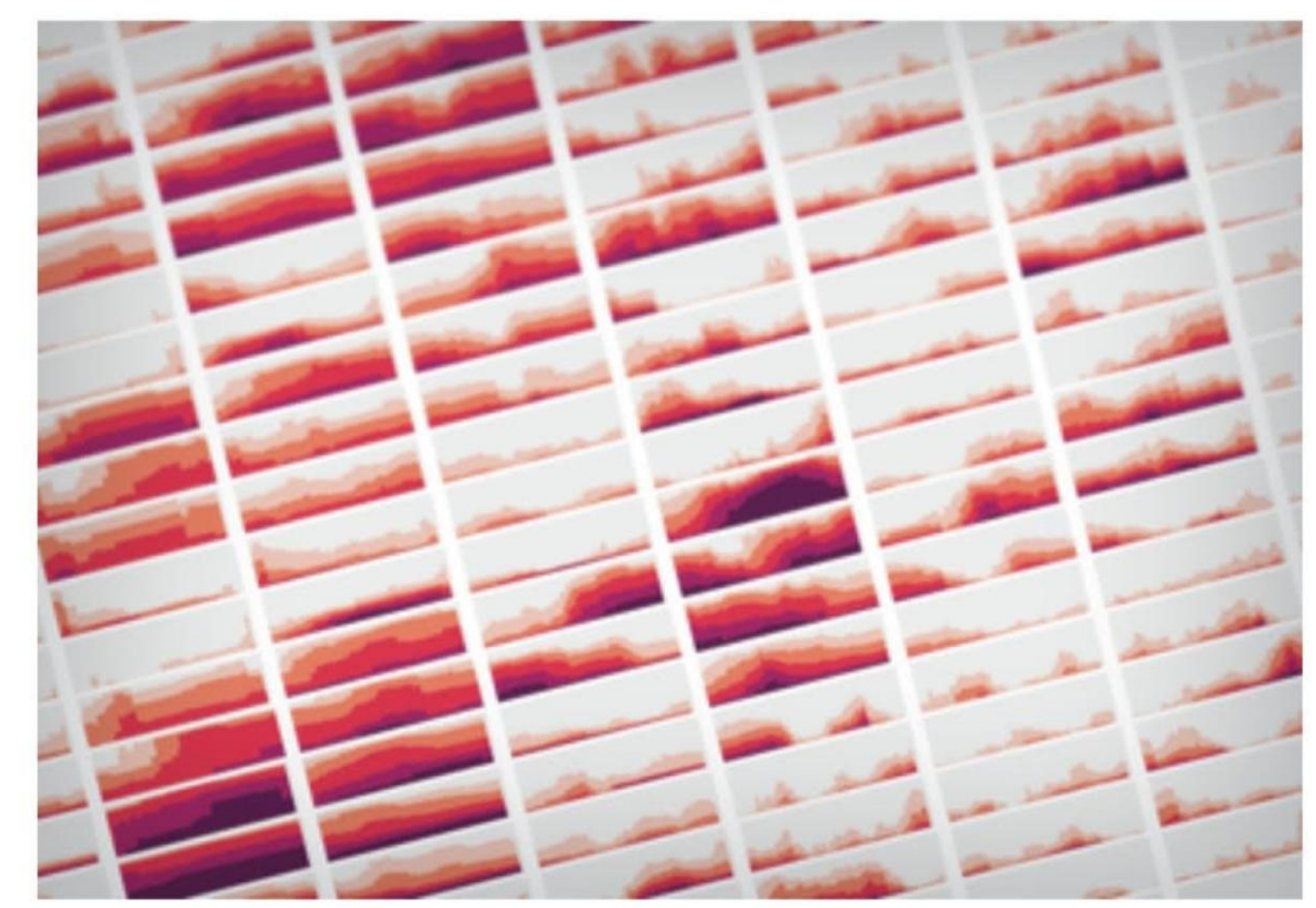


# CLIMATE CHANGE

## Climate Change Drives Escalating Drought

The past two decades have seen some of the most extreme dry periods in U.S. history

By Clara Moskowitz, Cédric Scherer, Georgios Karamanis  
| Scientific American November 2021 Issue



Credit: Cédric Scherer and Georgios Karamanis



# CÉDRIC SCHERER

*Data Visualization & Computational Ecology*

## The World's Countries Colored by Their First Letter

While preparing the mapping section for a Pearson–O'Reilly training, I got the idea to visualize the first letter of each country. And got especially curious about how much landmass each letter covers. Turns out: A, C and R are covering the largest areas!

Posted by Cédric • Friday, August 27, 2021

## A Quick How-to on Labelling Bar Graphs in ggplot2

Bar charts are likely the most common chart type out there and come in several varieties. Most notably, Direct labels can increase accessibility of a bar graph. I got

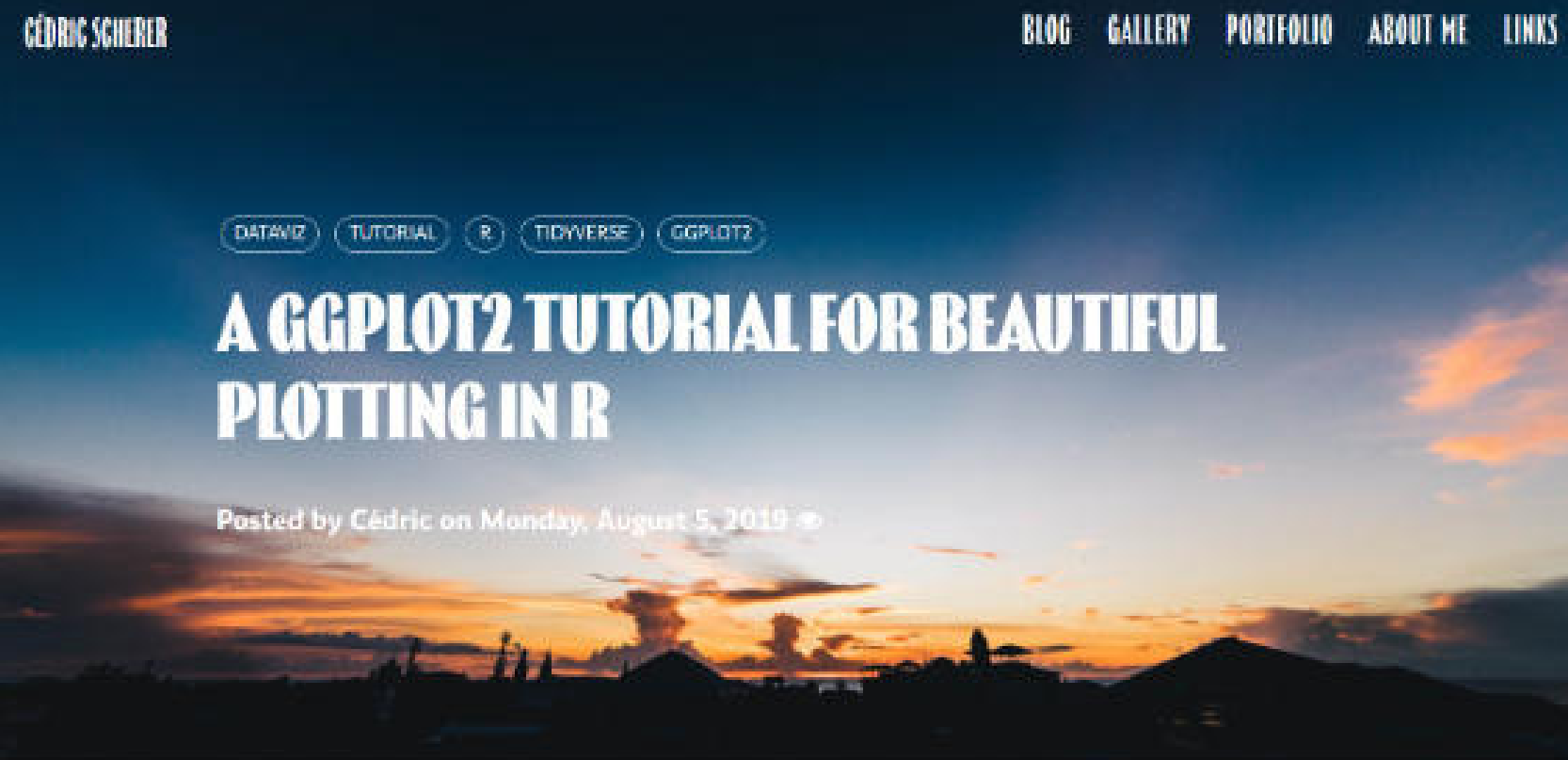


Always coding. Passionate about design. Worried about nature.

Proud dad.







[DATAVIZ](#) [TUTORIAL](#) [R](#) [TIDYVERSE](#) [GGPLOT2](#)

# A GGLOT2 TUTORIAL FOR BEAUTIFUL PLOTTING IN R

Posted by Cédric on Monday, August 5, 2019

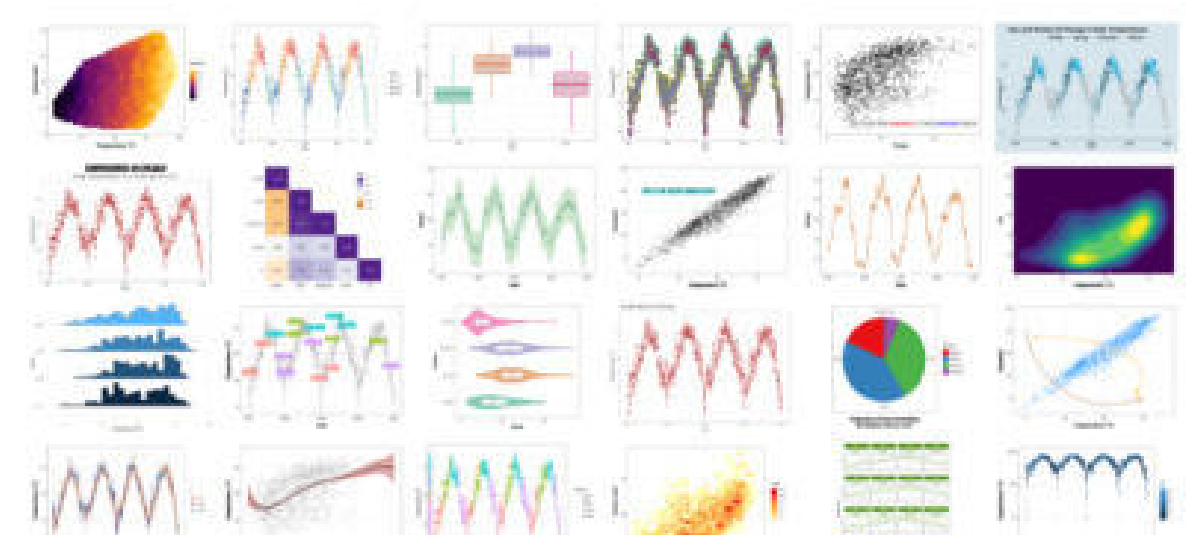
Last update: 2021-02-09

## Introductory Words

I don't care, just show me the content!

Back in 2016, I had to prepare my PhD introductory talk and I started using `ggplot2` to visualize my data. I never liked the syntax and style of base plots in R, so I was quickly in love with `ggplot`. Especially useful was its faceting utility. But because I was short on time, I plotted these figures by trial and error and with the help of lots of googling. The resource I came always back to was a blog entry called [Beautiful plotting in R: A ggplot2 cheatsheet](#) by Zev Ross, updated last in January 2016. After giving the talk which contained some decent plots thanks to the blog post, I decided to go through this tutorial step-by-step. I learned so much from it and directly started modifying the codes and over the time I added additional code snippets, chart types and resources.

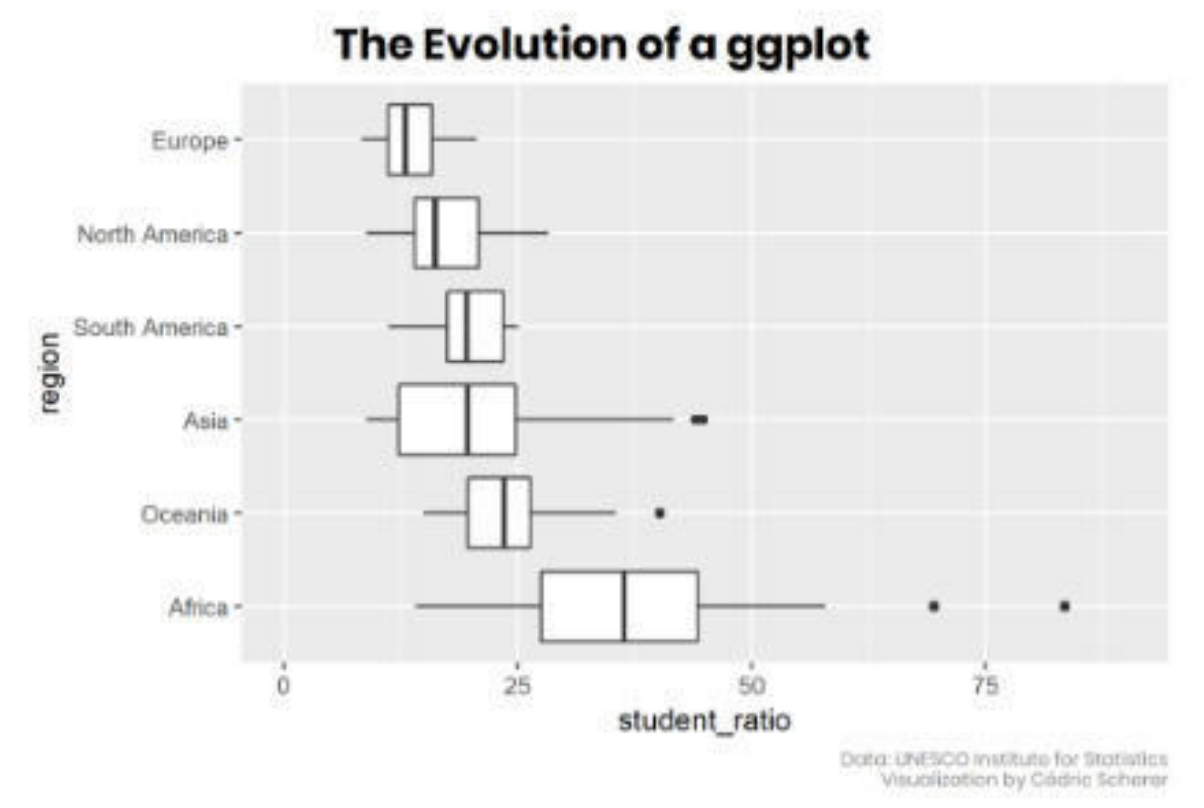
Since the blog entry by Zev Ross was not updated for some years and step by step this became a unique version of a tutorial, I decided to host the updated version on my GitHub. Now it finds its proper place on this homepage! (Plus I added a ton of other updates—just to name a few: The fantastic `patchwork`, `ggtext` and `ggforce` packages. How to deal with custom fonts and colors. A collection of R packages tailored to create interactive charts. And several other chart types including pie charts because everyone looooves pie charts!)



[DATAVIZ](#) [TUTORIAL](#) [ANIMATIONS](#) [GGPLOT EVOLUTION](#) [R](#) [GGPLOT2](#) [TIDYVERSE](#) [TIDYTUESDAY](#)

# THE EVOLUTION OF A GGLOT (EP. 1)

Posted by Cédric on Friday, May 17, 2019



- 🚩 Aim of this Tutorial
- 🛠️ Data Preparation
- 📊 The Default Boxplot
- 📦 Sort Your Data!
- 💡 Let Your Plot Shine—Get Rid of the Default Settings
- 📄 The Choice of the Chart Type
- 📦 More Geoms, More Fun, More Info!
- 💬 Add Text: Boxes to Let The Plot Speak for Itself
- 🗺️ Bonus: Add a Tile Map as Legend
- 🏆 The Final Evolved Visualization
- 📄 Complete Code for Final Plot
- 📄 Post Scriptum: Mean versus Median

## 🚩 Aim of this Tutorial

In this series of blog posts, I aim to show you how to turn a default ggplot into a plot that visualizes information in an appealing and easily understandable way. The goal of each blog post is to provide a step-by-step tutorial explaining how my visualization have evolved from a typical



# Data Visualization

---

is any graphical representation  
of information and data.



# Data Visualization

---

**converts information into visual forms as quantifiable features.**



# Data Visualization

---

helps to amplify cognition, gain insights,  
discover, explain, and make decisions.



# Data Visualization

---

is part art and part science.





Source: [eazybi.com](http://eazybi.com)

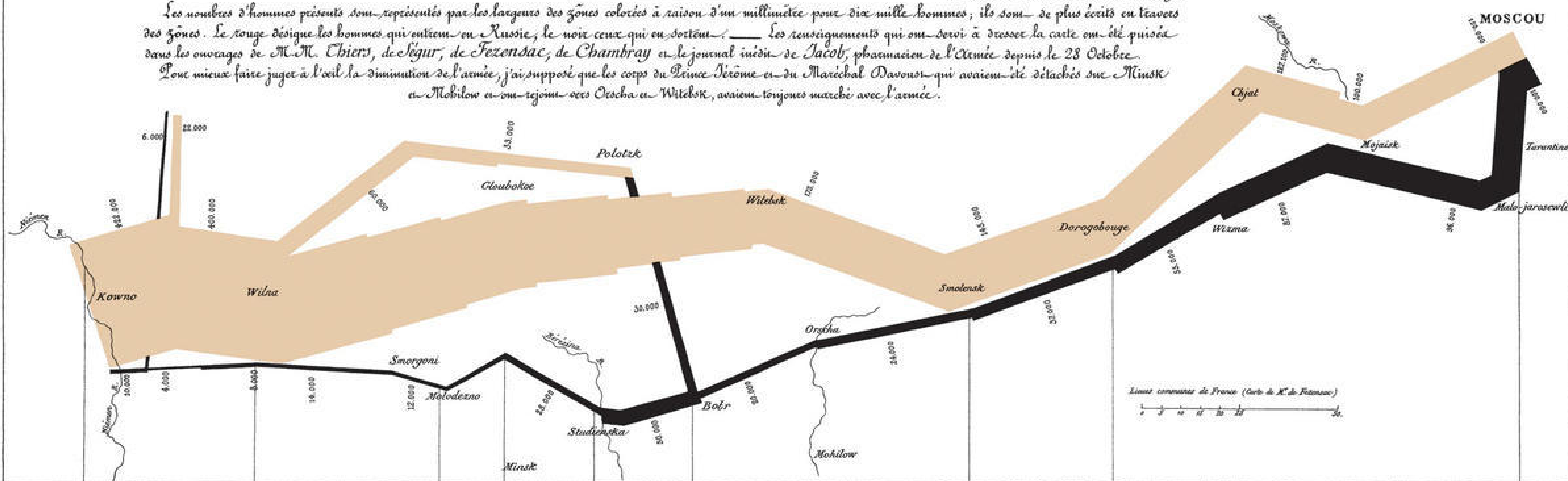


# Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

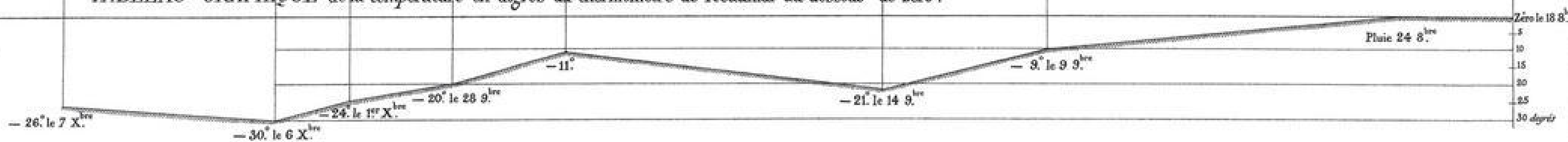
Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui ont été en Russie, le noir ceux qui en sont sortis. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et ont rejoint vers Oescha et Witebsk, avaient toujours marché avec l'armée.



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

Les Cosaques passent au galop le Niémen gelé.

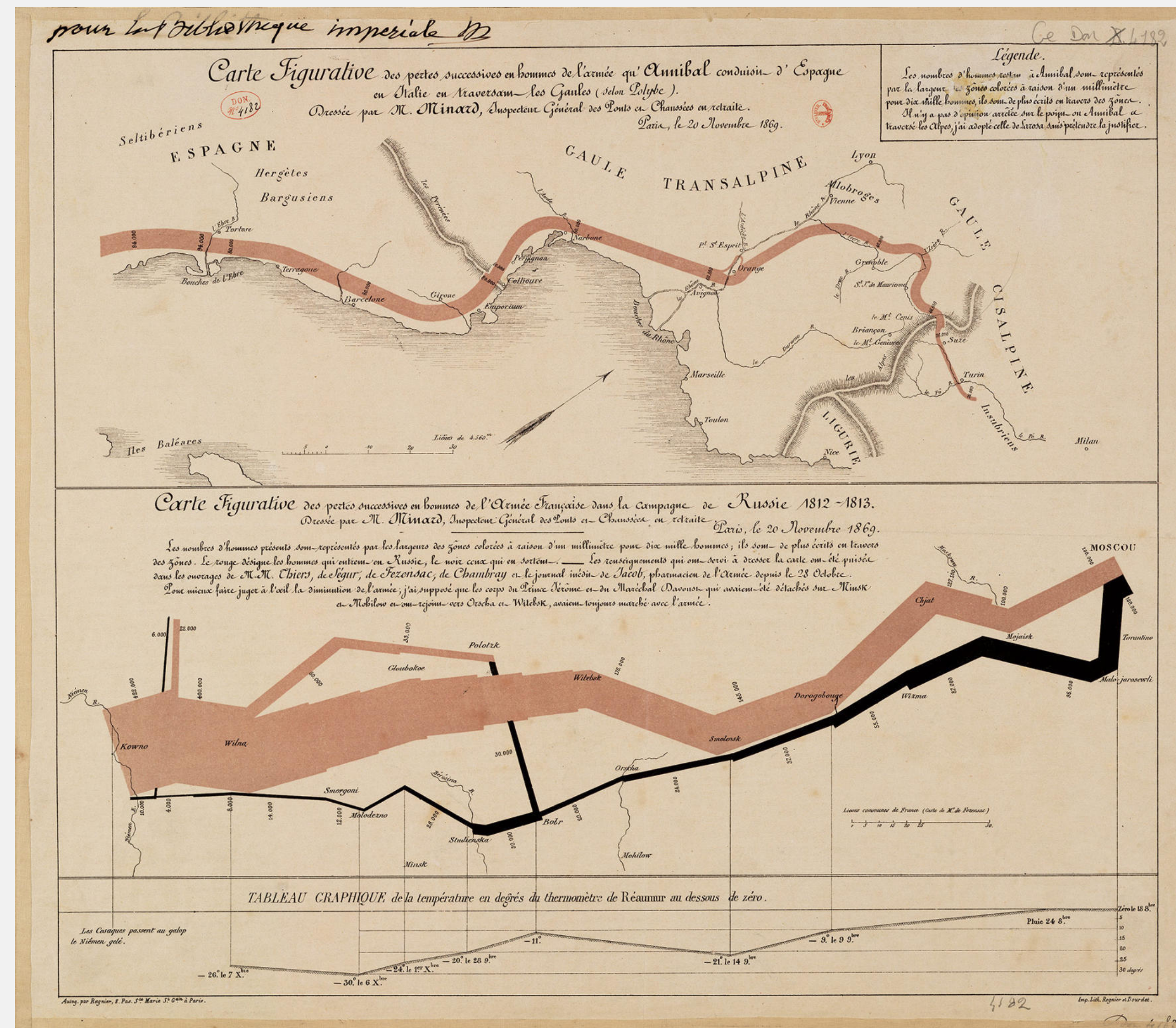


Autog. par Regnier, 1. Pas. 5<sup>e</sup> Marie 5<sup>e</sup> G<sup>de</sup> à Paris.

Imp. Lith. Regnier et Dourdet.

Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813 by Charles Joseph Minard



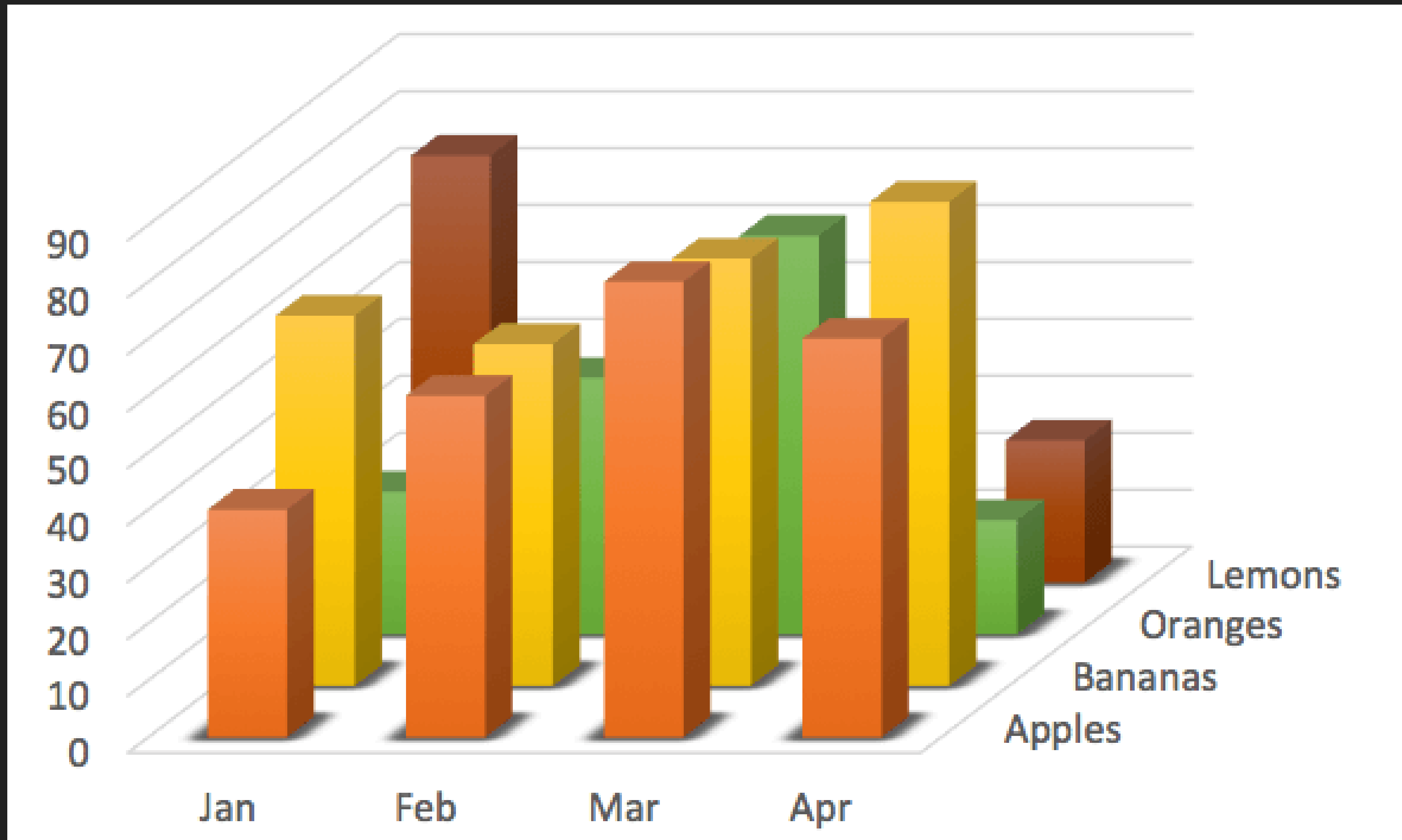


Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813 and Carte figurative des pertes successives en hommes de l'Armée qu'Annibal conduisit d'Espagne en Italie en traversant les Gaules (selon Polybe) by Charles Joseph Minard

- displays the progress of the troops of **Hannibal** (218 BC) and **Napoleon** (1812-1813)
- often considered as the **best statistical graphic ever** drawn



# What makes it a bad data visualization?

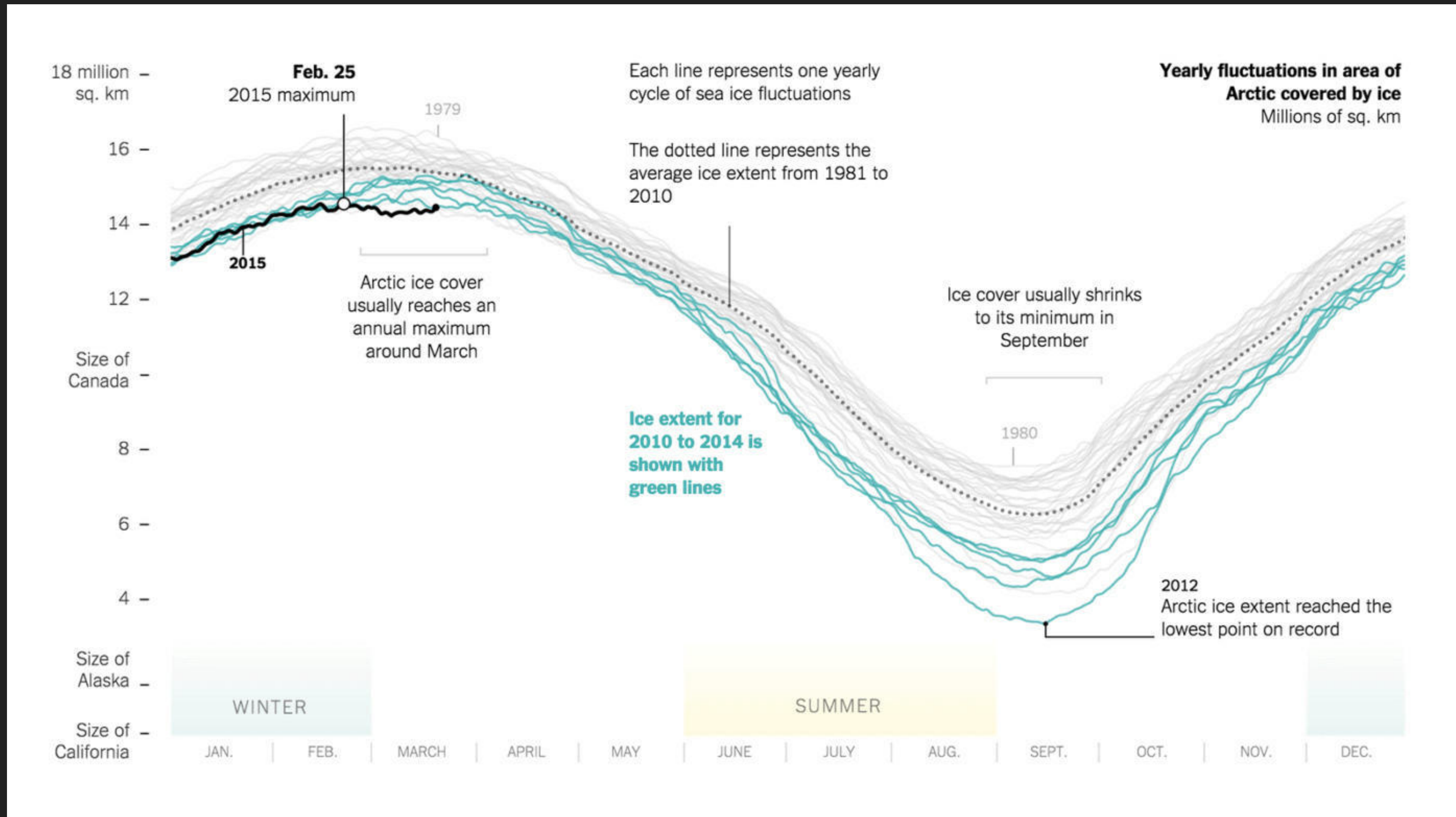


# What makes it a bad data visualization

- ☞ **substantive problems** (bad data / story)
- ☞ **aesthetic problems** (bad design)
- ☞ **perceptual problems** (bad encoding)



# What makes it a good data visualization?



*“Yearly Fluctuations in Area of Arctic Covered by Ice” by Derek Watkins (New York Times)*

# What makes it a good data visualization

👉 **INFORMATION** (integrity)

👉 **STORY** (interestingness)

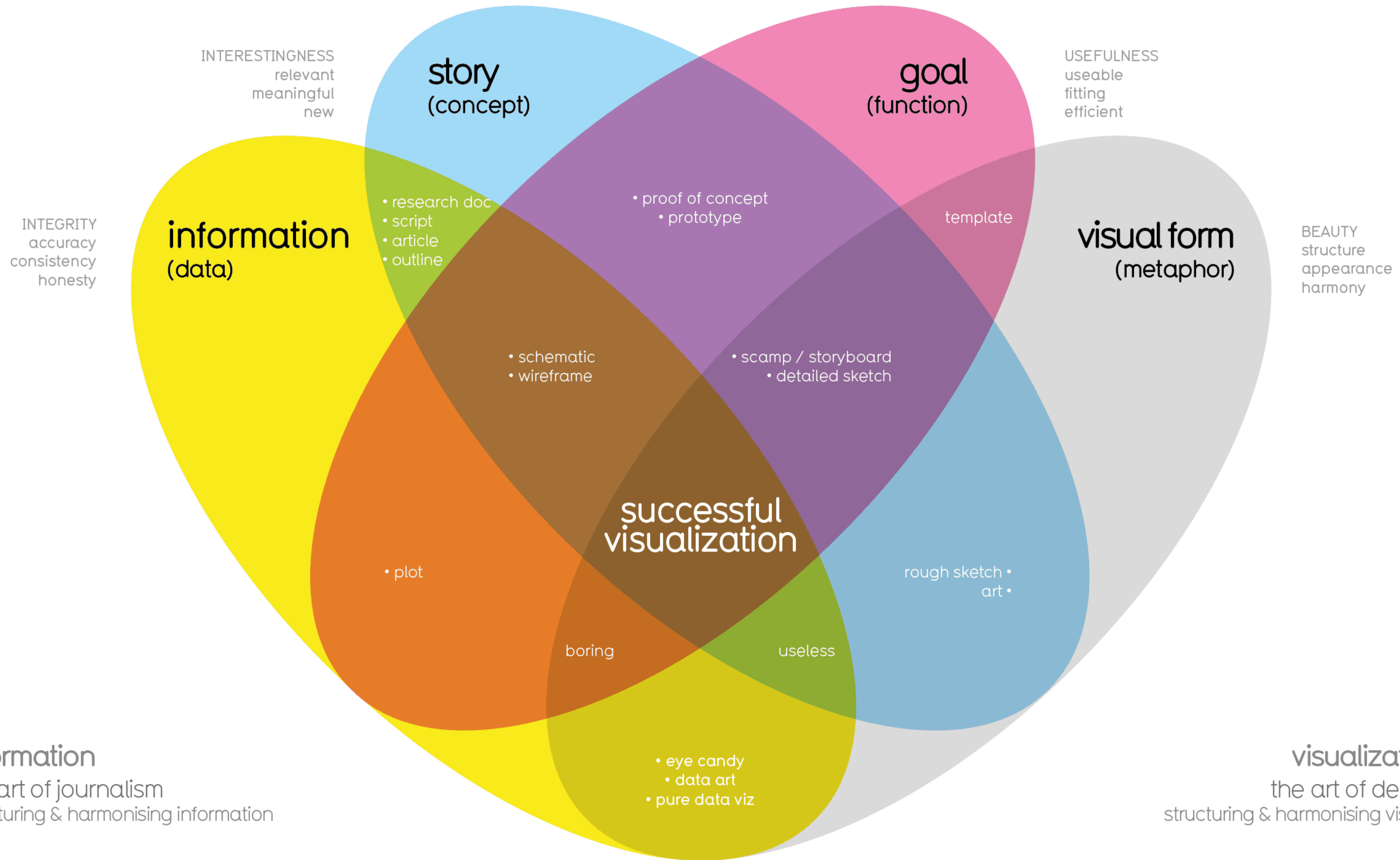
👉 **GOAL** (usefulness)

👉 **VISUAL FORM** (beauty)



# What Makes a Good Visualization?

explicit (implicit)



David McCandless  
InformationisBeautiful.net

taken from new book  
Knowledge is Beautiful

find out more  
[bit.ly/KIB\\_Books](http://bit.ly/KIB_Books)

# INFORMATION

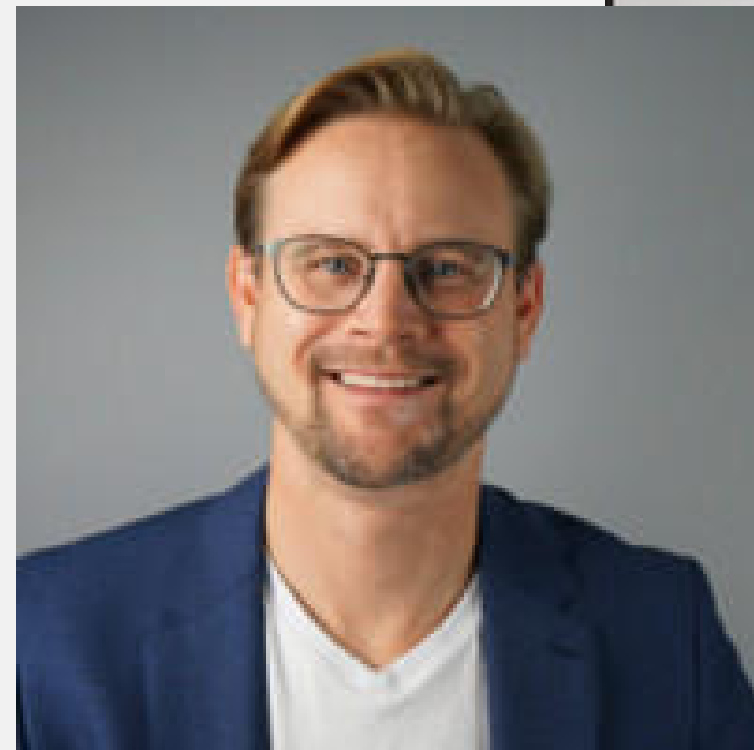
---

Understand your data and be accurate



How to Steer Clear of Common Blunders When Working  
with Data and Presenting Analysis and Visualizations

# AVOIDING DATA PITFALLS



**BEN JONES**

Founder and CEO, Data Literacy

**WILEY**

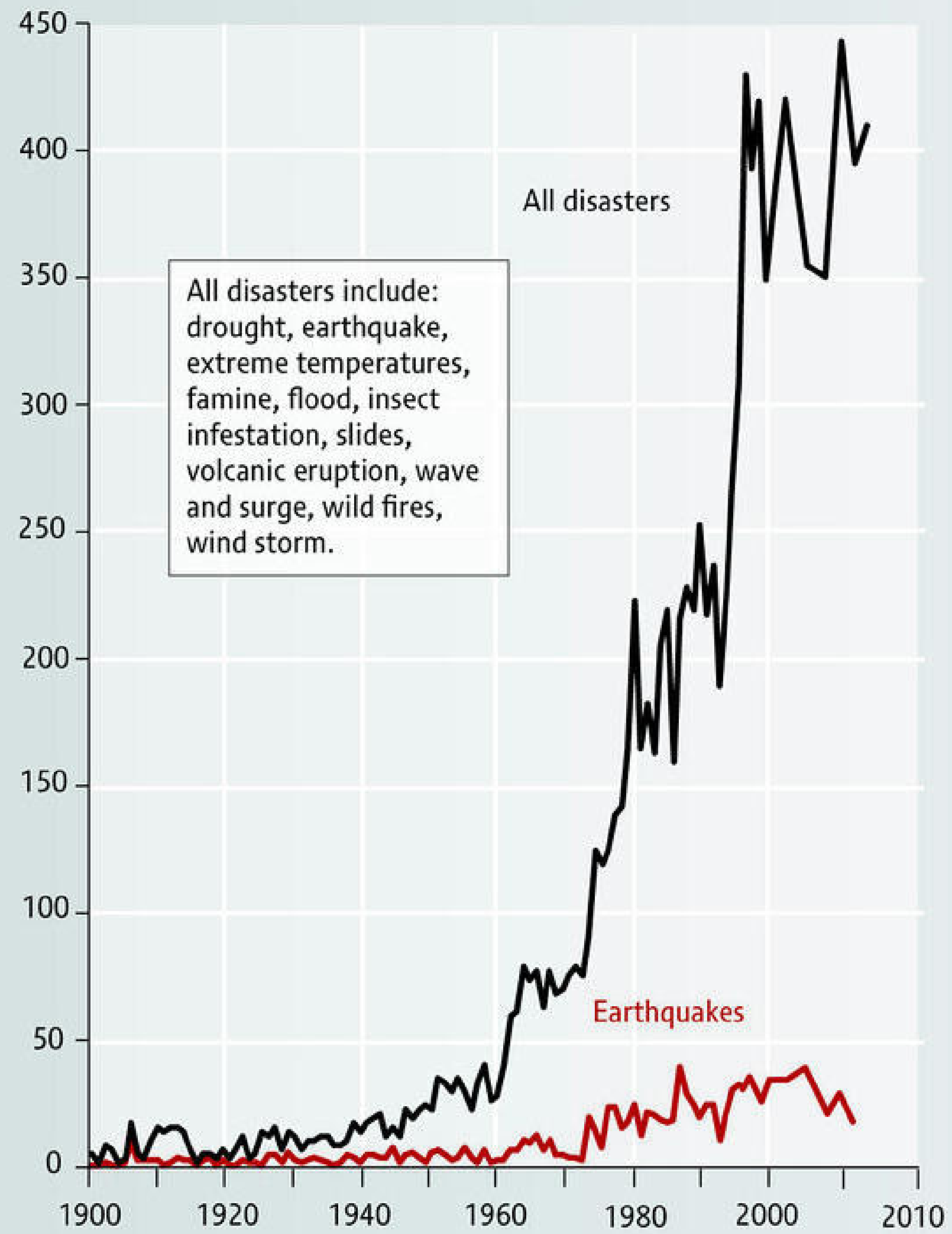
# Our data is never a perfect reflection of the real world.

- **only a subset:** not crime but reported crime
- **collected by humans:** guesstimation, precision and errors
- **collected by machines:** precisions and errors



# Number of disasters

per year

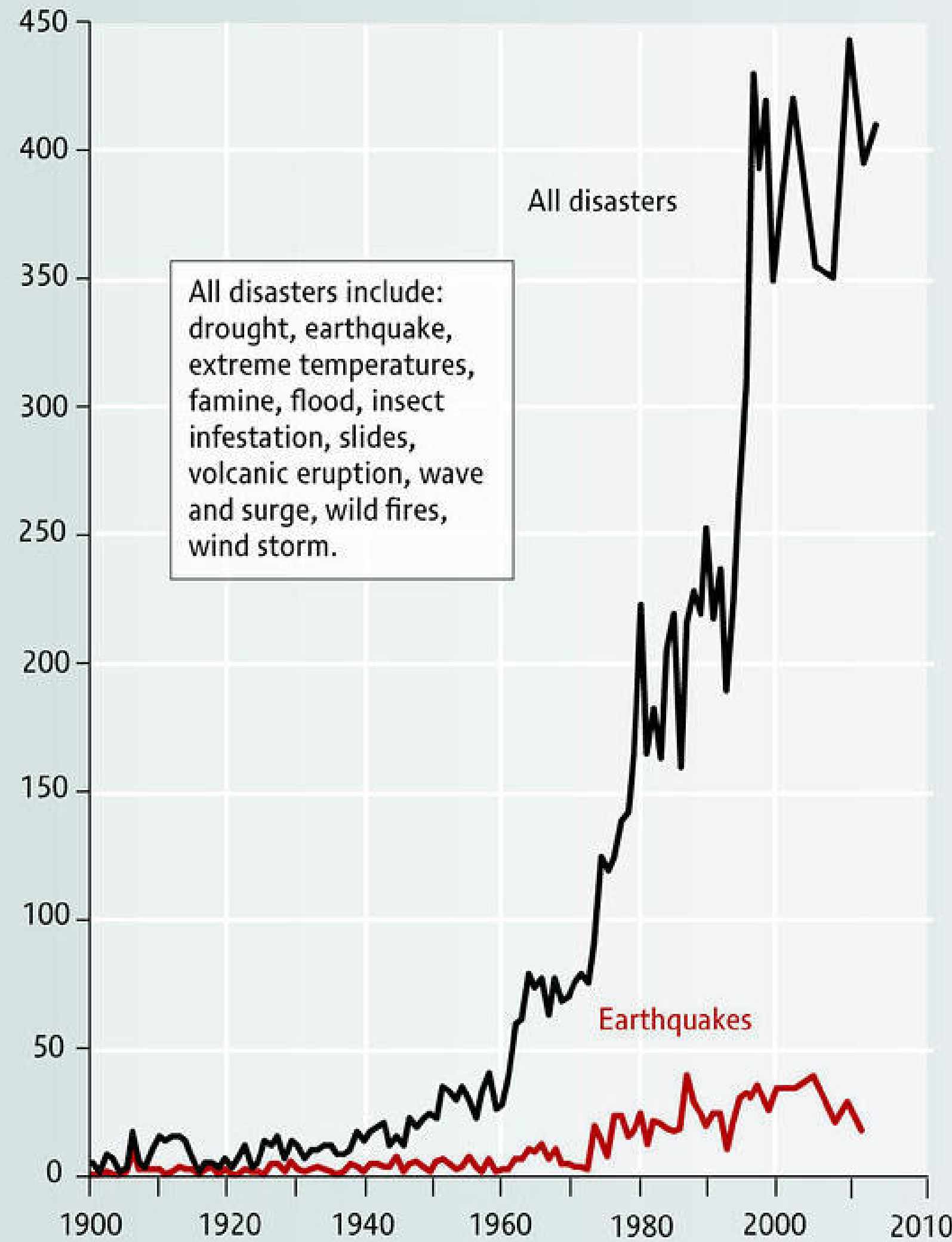


Source: CRED Annual Disaster Statistical Review 2006, 2007.

“Much of the increase of **hazardous events reported** is probably due to significant **improvements in information access**”

### Number of disasters

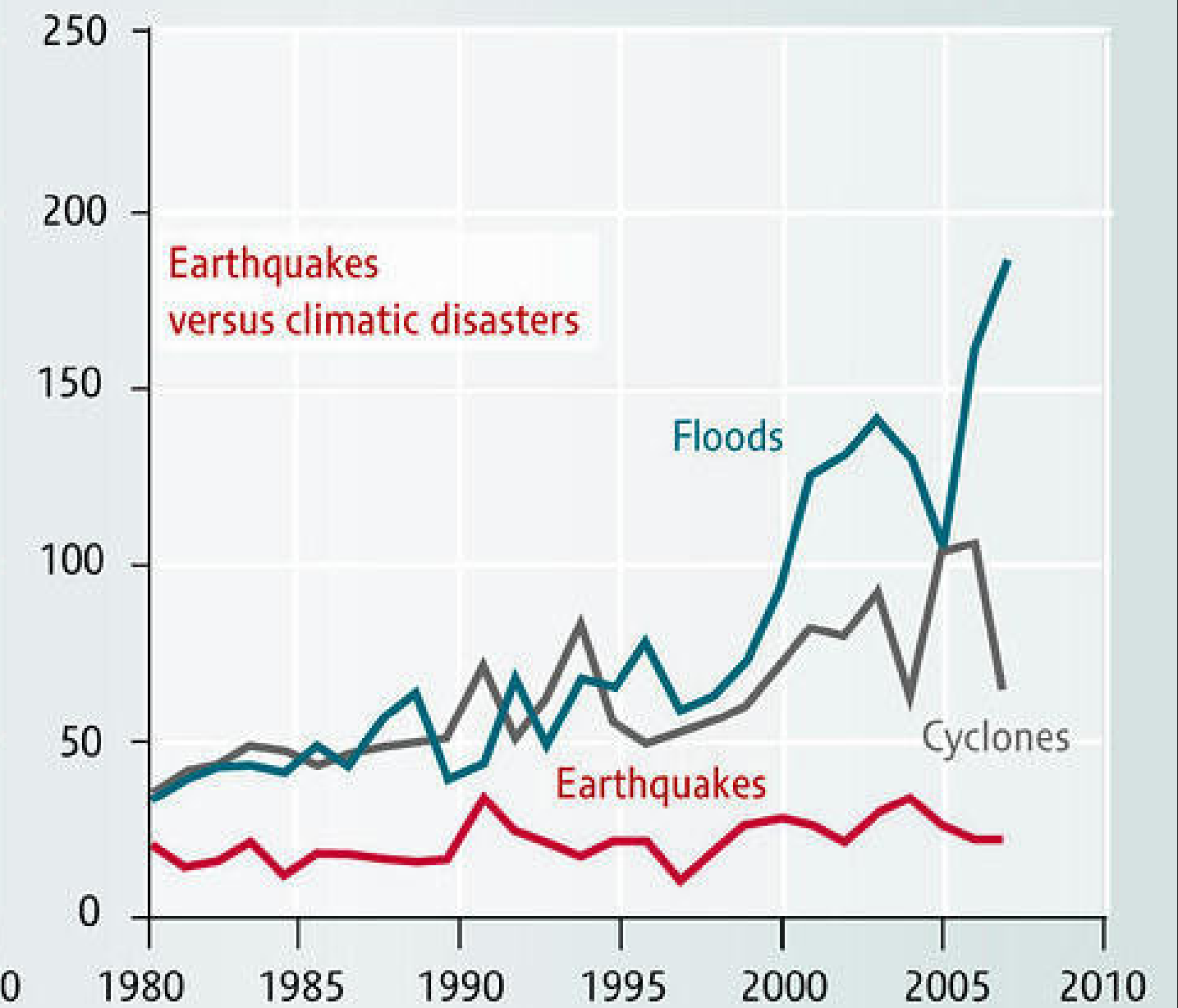
per year



Source: CRED Annual Disaster Statistical Review 2006, 2007.

### Trends in number of reported disasters

Much of the increase in the number of hazardous events reported is probably due to significant improvements in information access and also to population growth, but the number of floods and cyclones reported is still rising compared to earthquakes. Is global warming affecting the frequency of natural hazards?





The best use of data is to  
teach us what **isn't true.**

# The best use of data is to teach us what isn't true.

- don't formulate a single statement
- confront yourself with a **falsifiable universal statement**



# The best use of data is to teach us what isn't true.

- single statement: “The swan is white.”
- **falsifiable universal statement: “All swans are white.”**

# STORY

---

**Be clear about the message of your visualization**



# Who is my audience?

Which story is **interesting** for them?

# Who is my audience?

Which story is **interesting** for them?

What are **relevant** details to include?



# Who is my audience?

Which story is **interesting** for them?

What are **relevant** details to include?

Which variables are **meaningful** to them?

# Who is my audience?

Which story is **interesting** for them?

What are **relevant** details to include?

Which variables are **meaningful** to them?

How will they **encounter** the visualization?

# Who is my audience?

Which story is **interesting** for them?

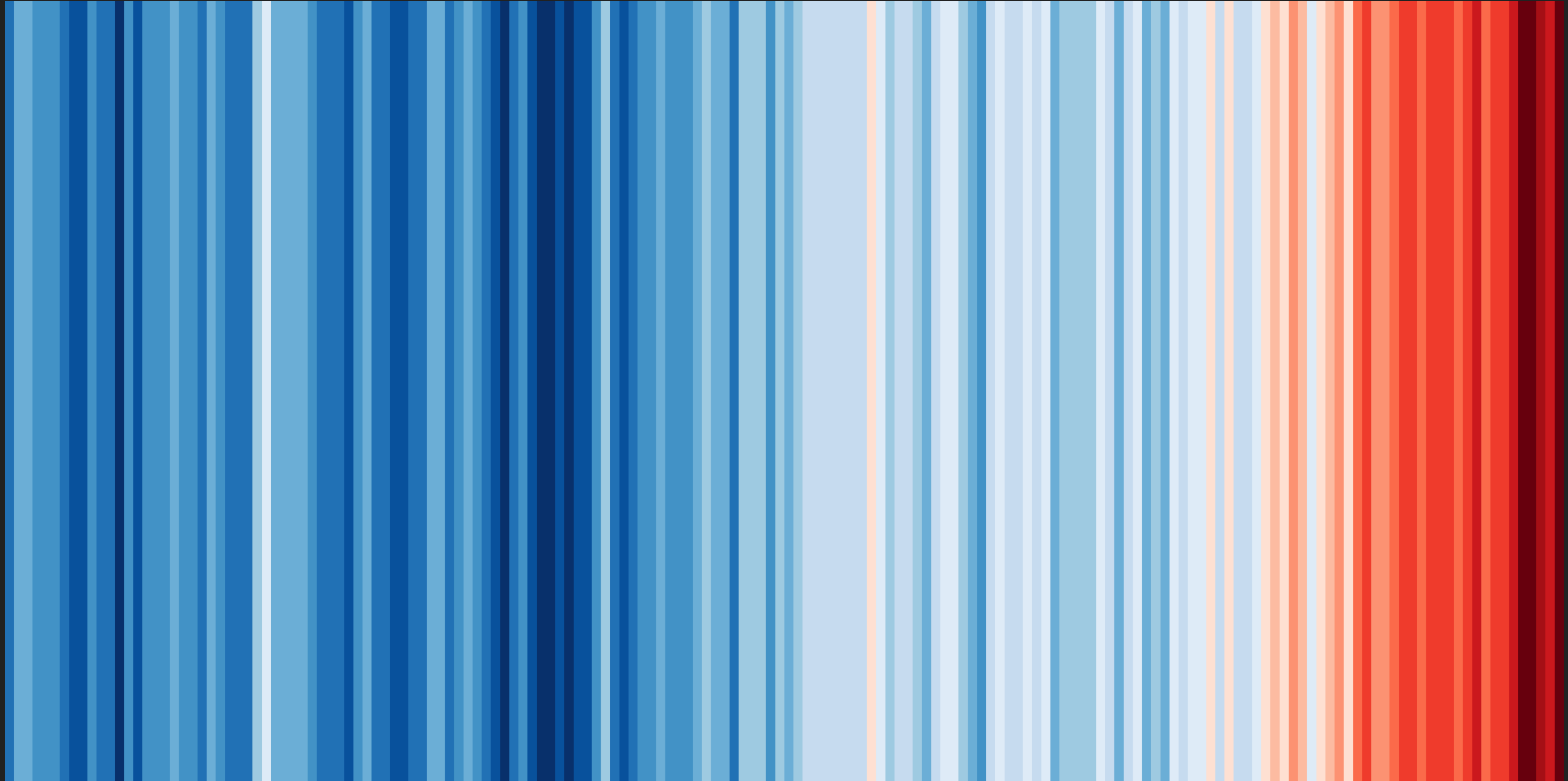
What are **relevant** details to include?

Which variables are **meaningful** to them?

How will they **encounter** the visualization?

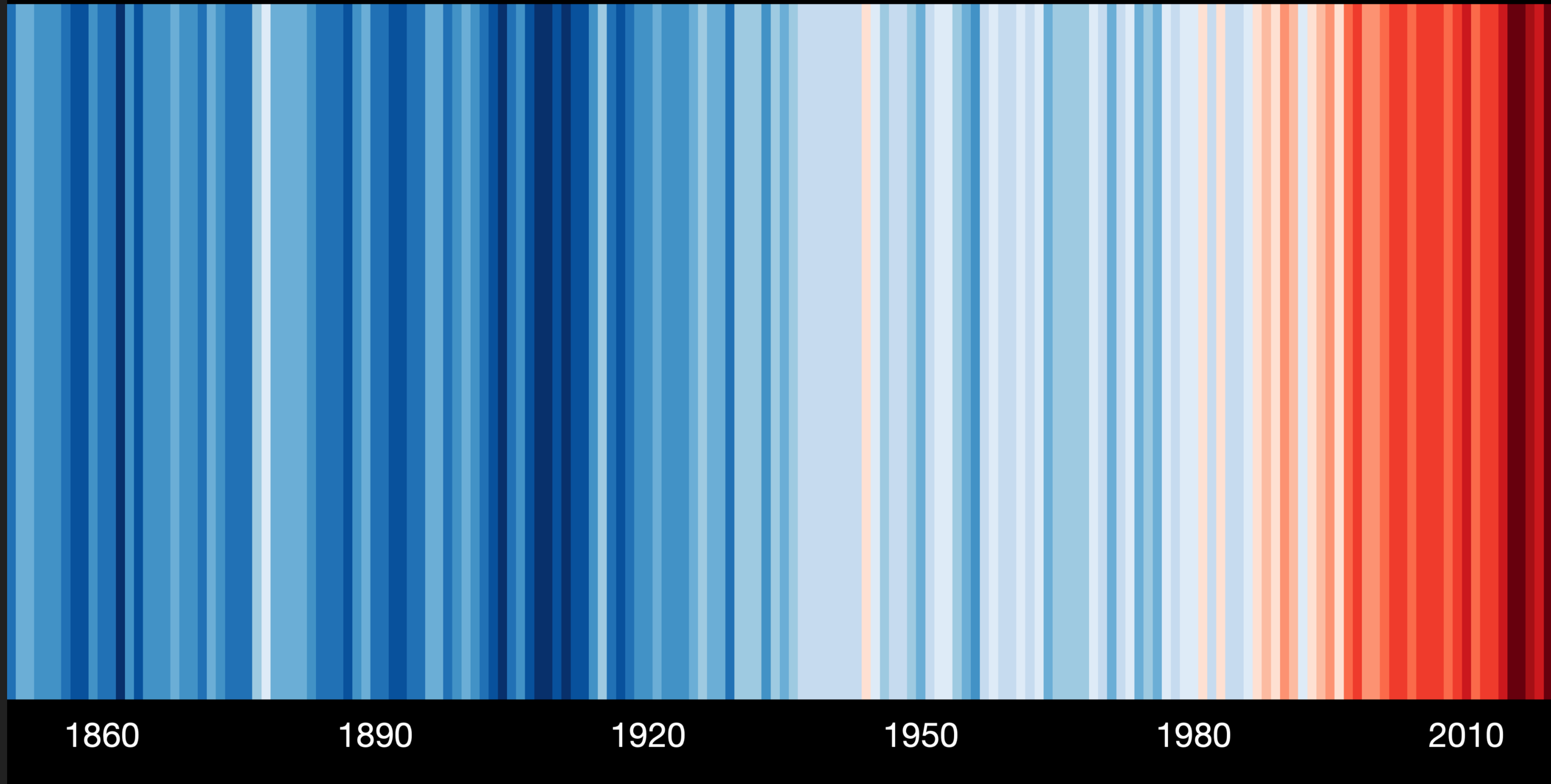
**Do I need a visualization at all??**





*Warming Stripes* by Ed Hawkins


# Global temperature change (1850-2019)



*Warming Stripes* by Ed Hawkins

## FAQ : Frequently asked questions

What are these graphics? 

What do the graphics show? 

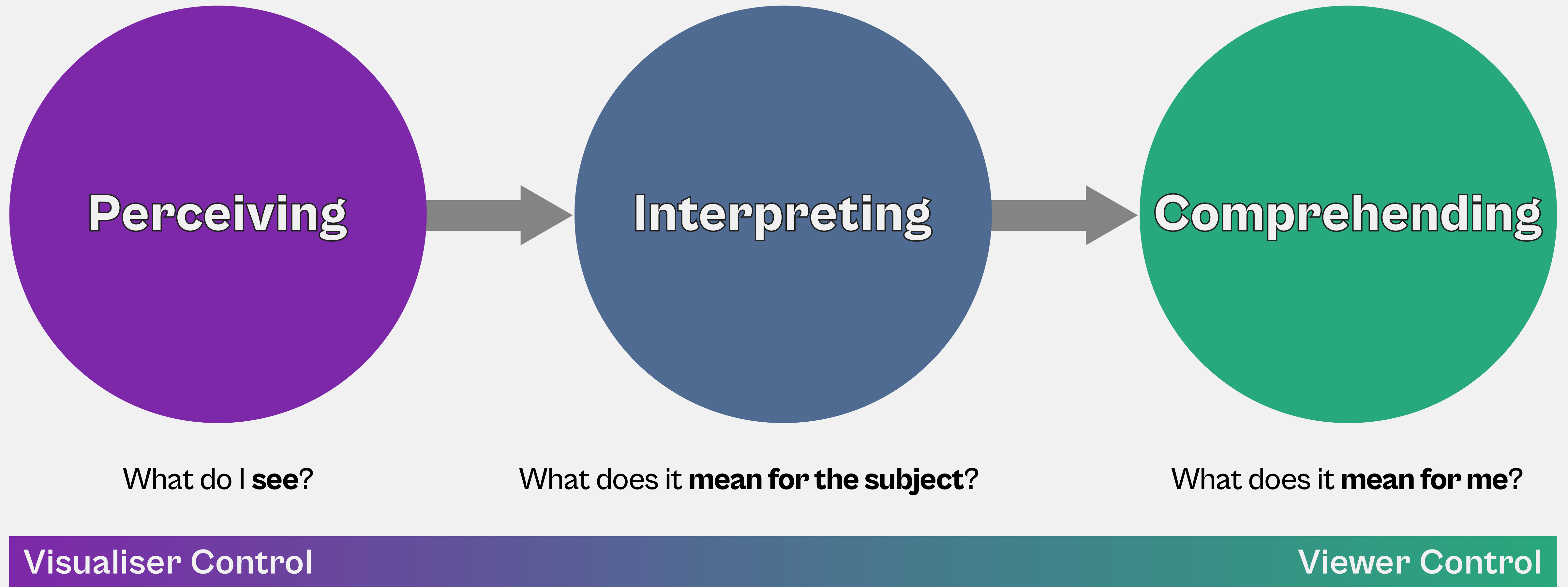
Why are there no numbers on the graphics? 

»» These graphics are specifically designed to be as simple as possible, and to start conversations about our warming world and the risks of climate change. There are numerous sources of information which provide more specific details about how temperatures have changed, so these graphics fill a gap and enable communication with minimal scientific knowledge required to understand their meaning.



These graphics are specifically **designed to [...] start conversations** about our warming world and the risks of climate change.

» These graphics are specifically designed to be as simple as possible, and to start conversations about our warming world and the risks of climate change. There are numerous sources of information which provide more specific details about how temperatures have changed, so these graphics fill a gap and enable communication with minimal scientific knowledge required to understand their meaning.



*Scheme by Andy Kirk*

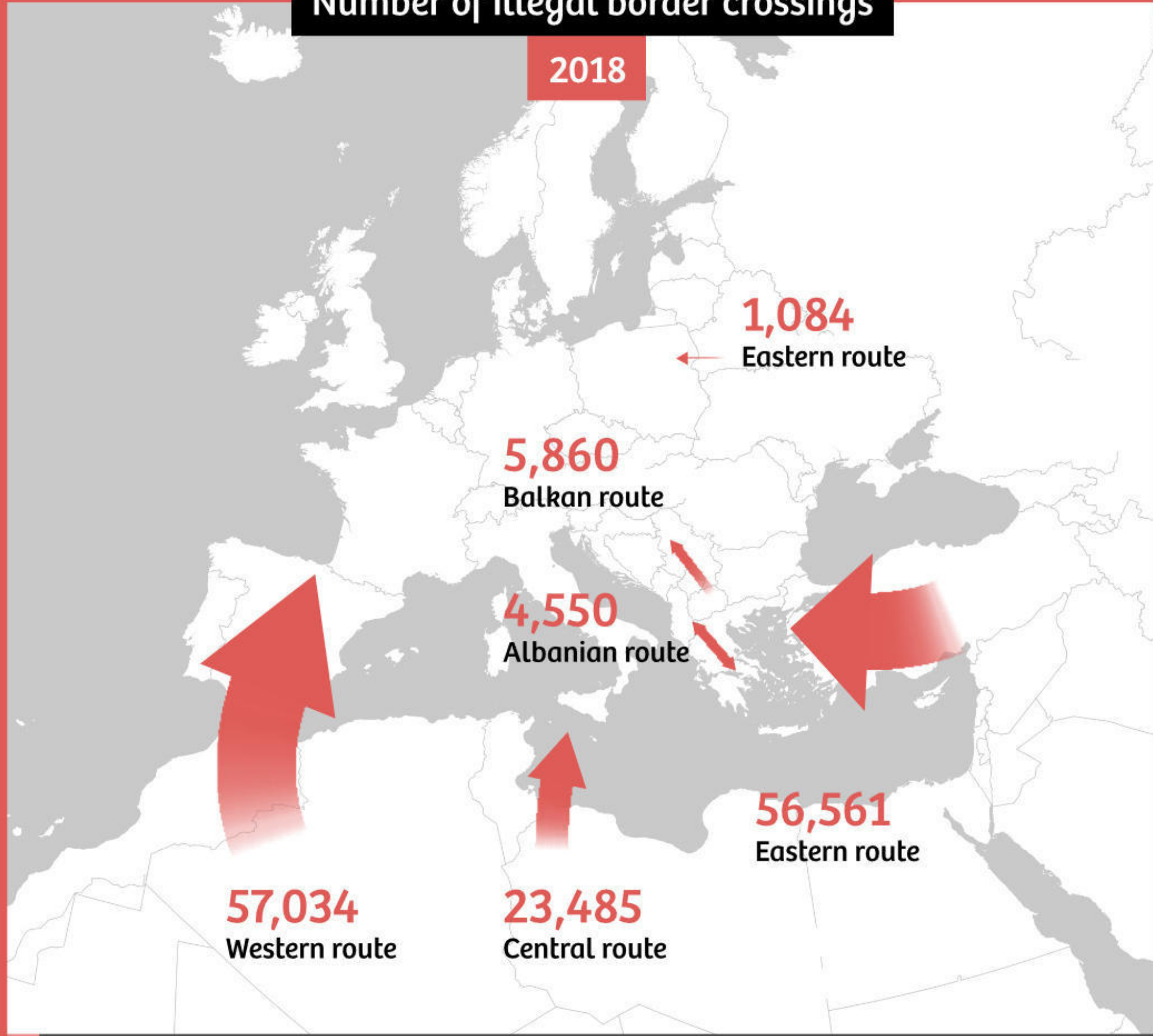


*“How maps in the media make us more negative about migrants” by Maite Vermeulen, Leon de Korte & Henk van Houtum*



### Number of illegal border crossings

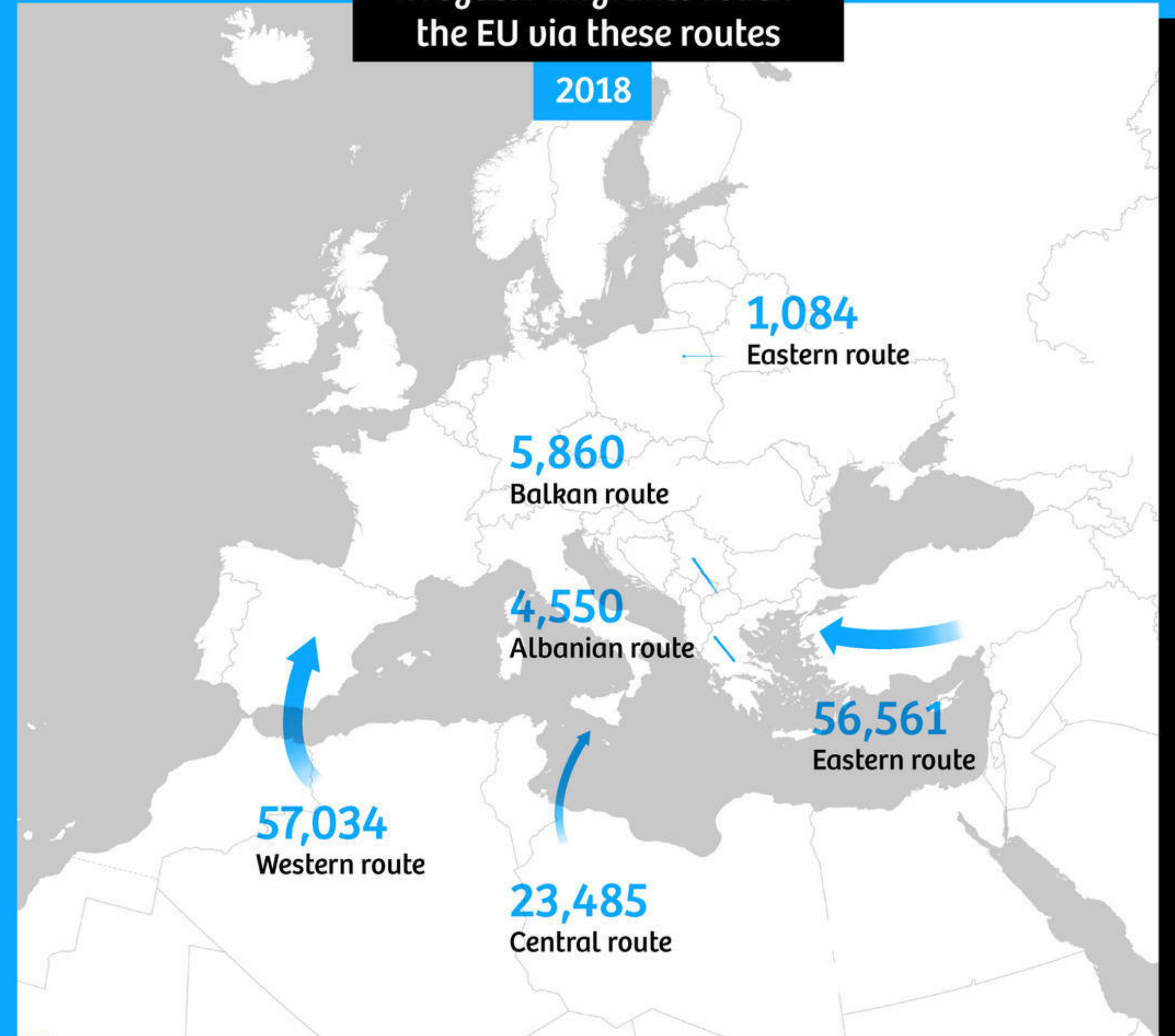
2018



Source: Frontex

### Irregular migrants reach the EU via these routes

2018



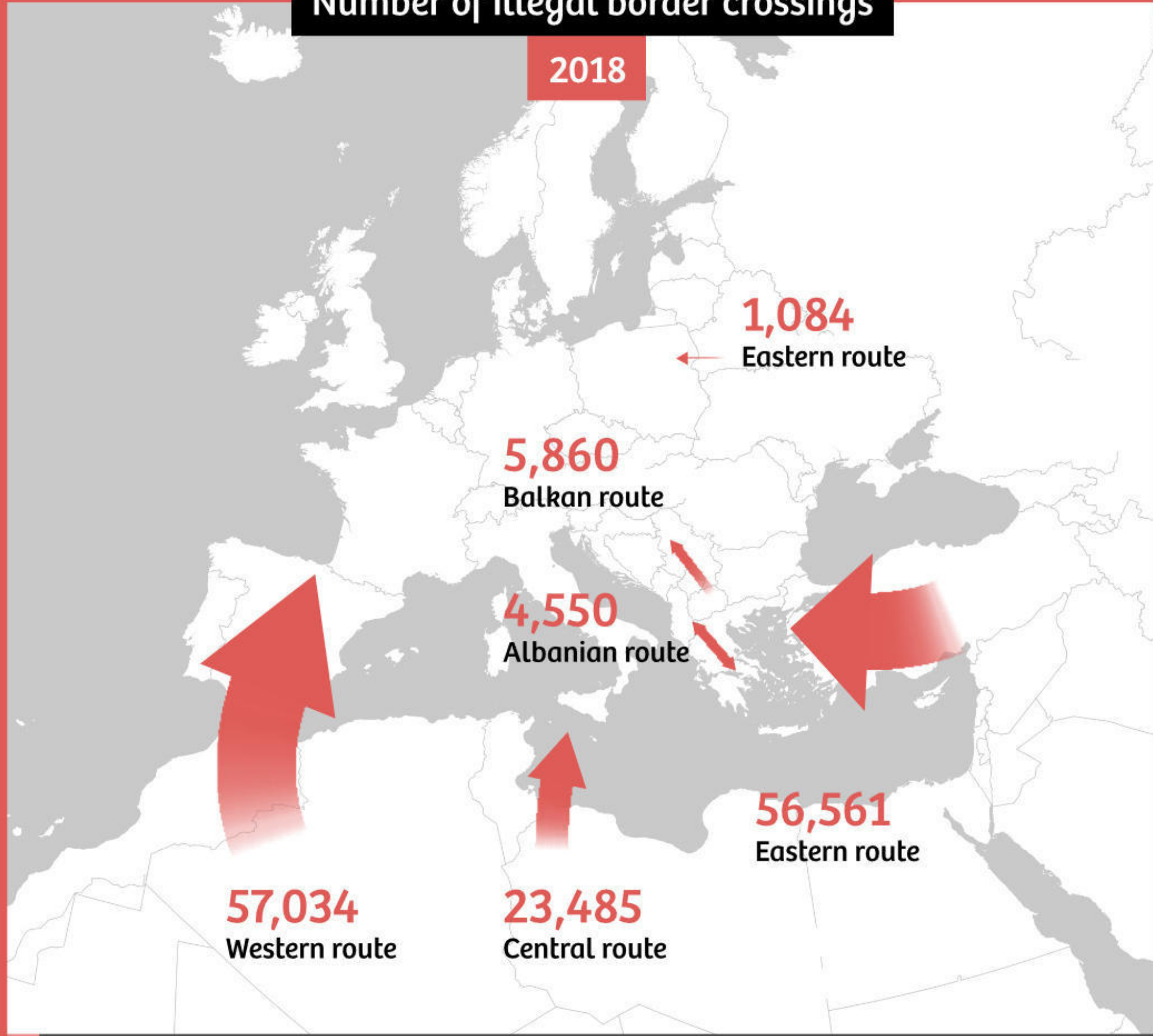
Source: Frontex

*“How maps in the media make us more negative about migrants” by Maite Vermeulen, Leon de Korte & Henk van Houtum*



### Number of illegal border crossings

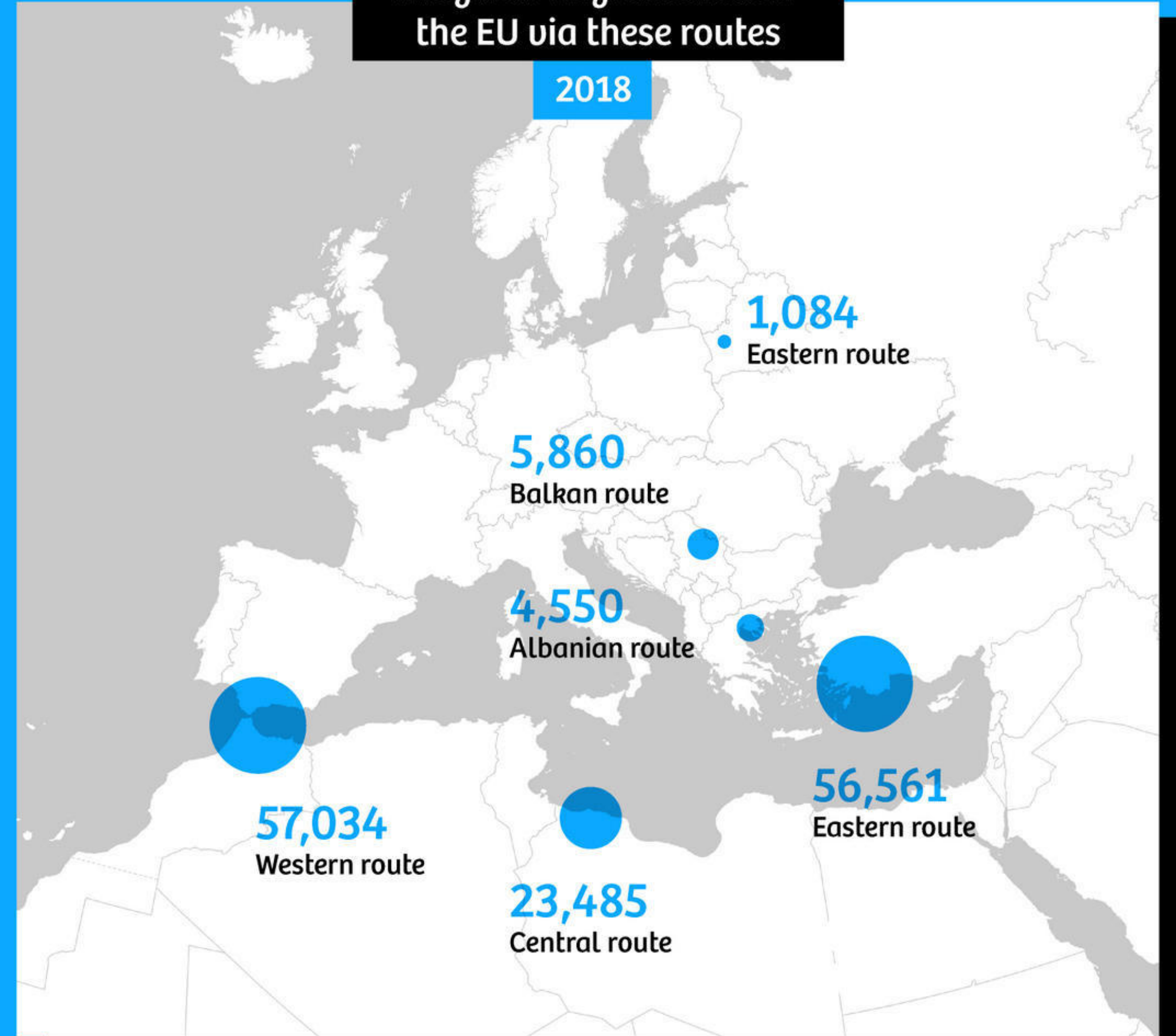
2018



Source: Frontex

### Irregular migrants reach the EU via these routes

2018

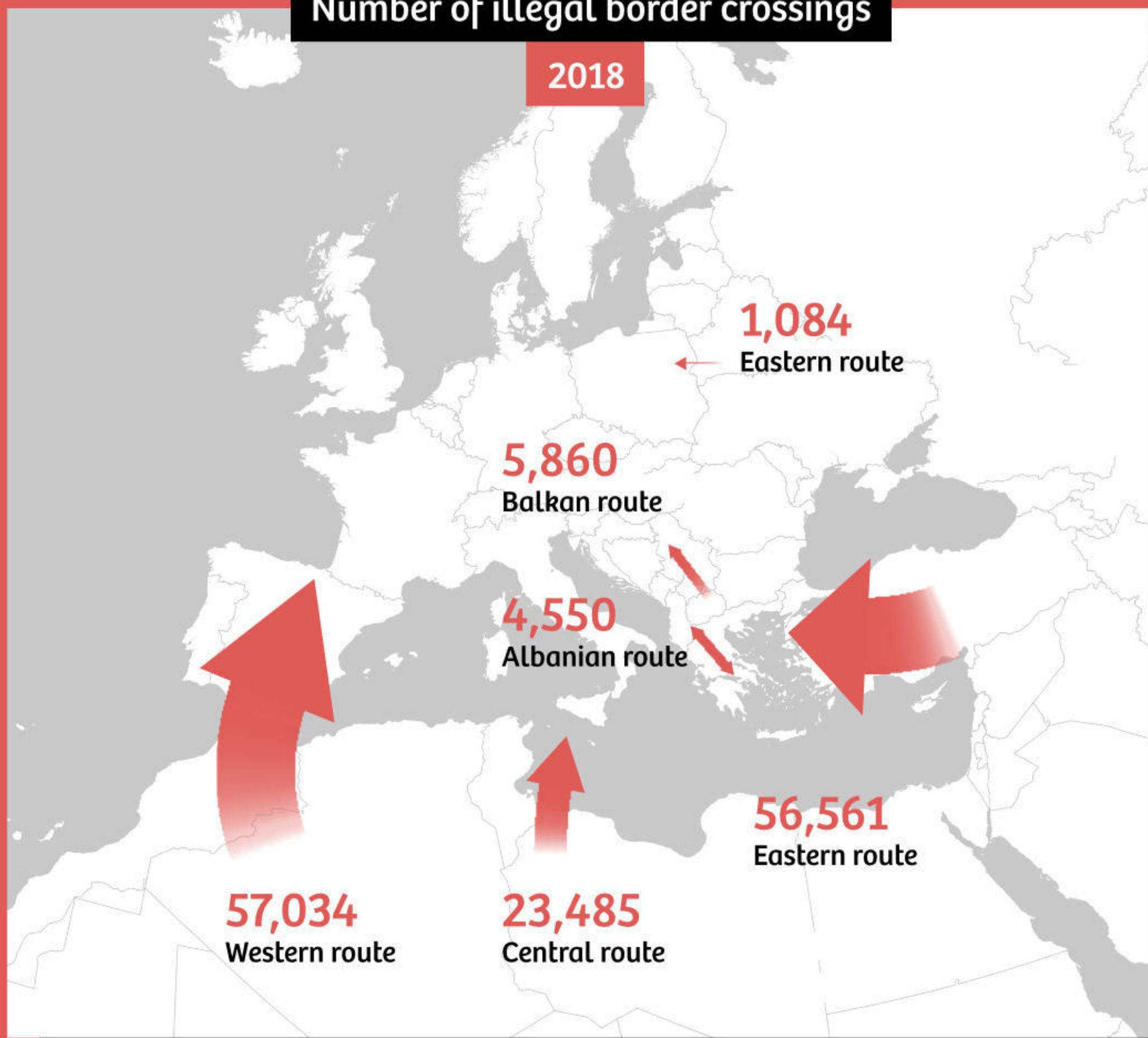


Source: Frontex

*“How maps in the media make us more negative about migrants” by Maite Vermeulen, Leon de Korte & Henk van Houtum*

### Number of illegal border crossings

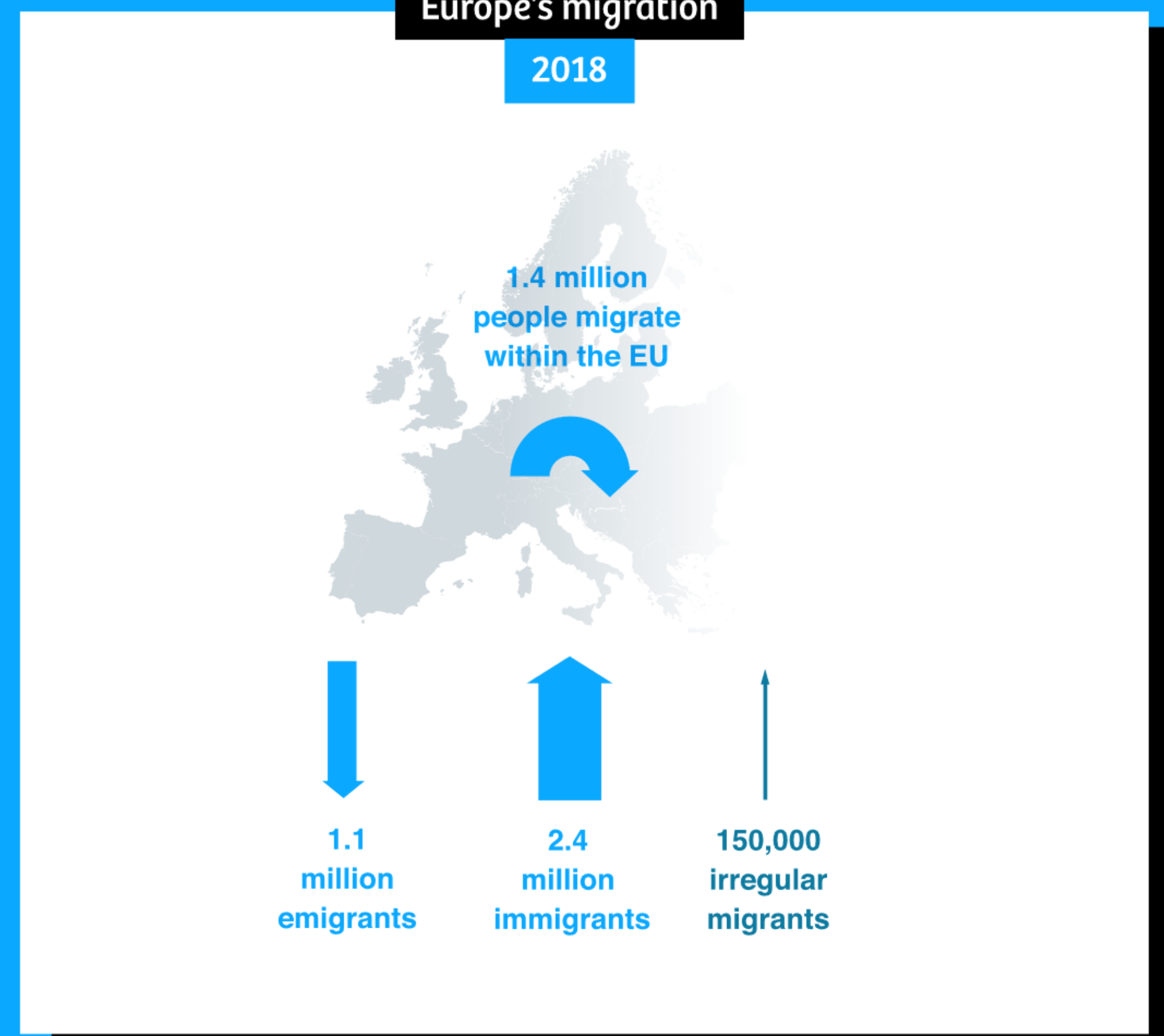
2018



Source: Frontex

### Europe's migration

2018

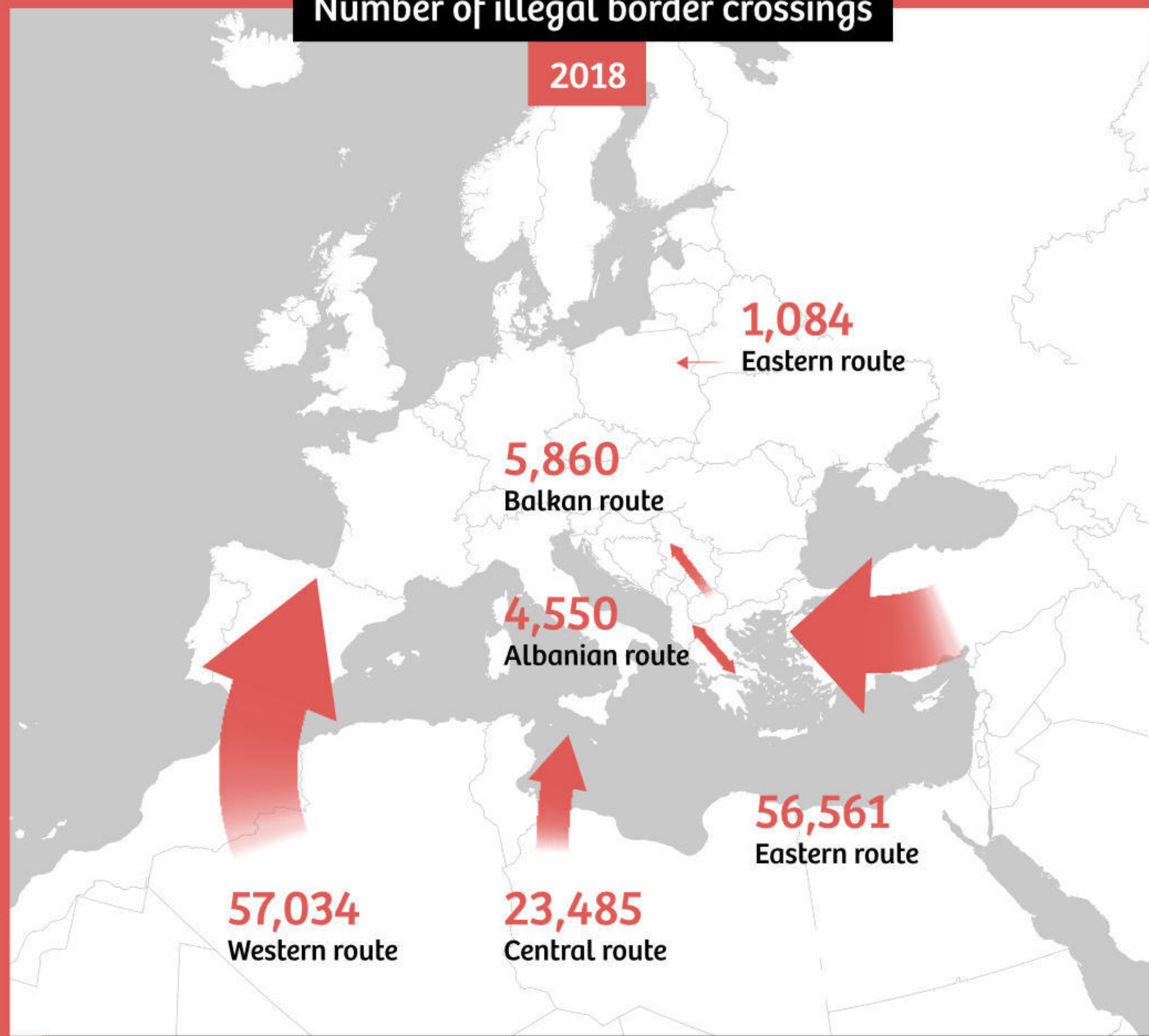


Source: Eurostat

*“How maps in the media make us more negative about migrants” by Maite Vermeulen, Leon de Korte & Henk van Houtum*

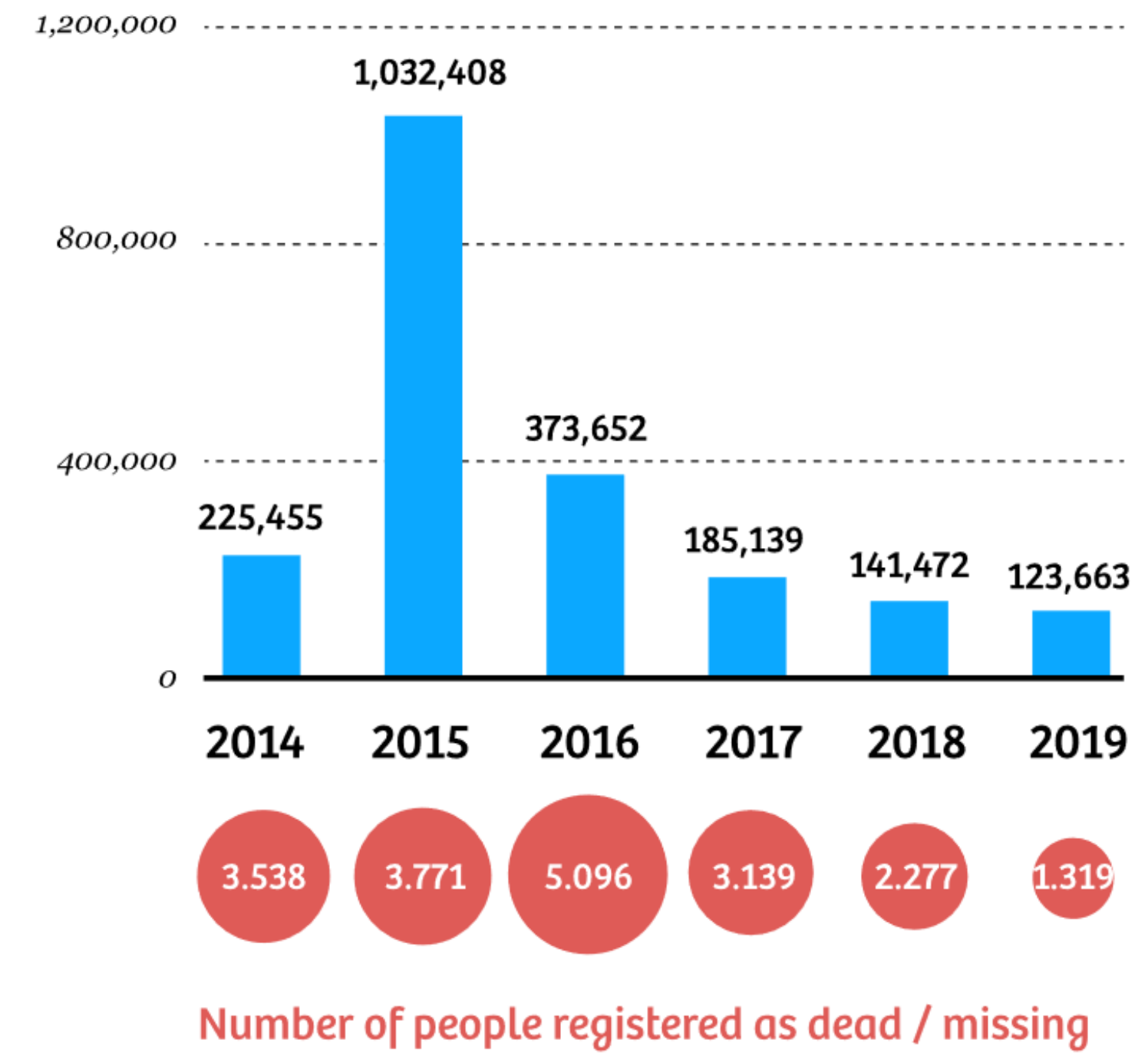


### Number of illegal border crossings



Source: Frontex

### Number of migrants arriving via the Mediterranean sea has been decreasing for years



Source: UNHCR

*“How maps in the media make us more negative about migrants” by Maite Vermeulen, Leon de Korte & Henk van Houtum*

# GOAL

---

Select charts that successfully transport your story

# Typology of Information Graphics

by Juuso Koponen & Jonatan Hildén, "Data Visualization Handbook" (2020), p. 25

Is the **information** **conceptual** or **measurable**?

☞ **Type of information:** depict conceptual information <> convert information into visual forms

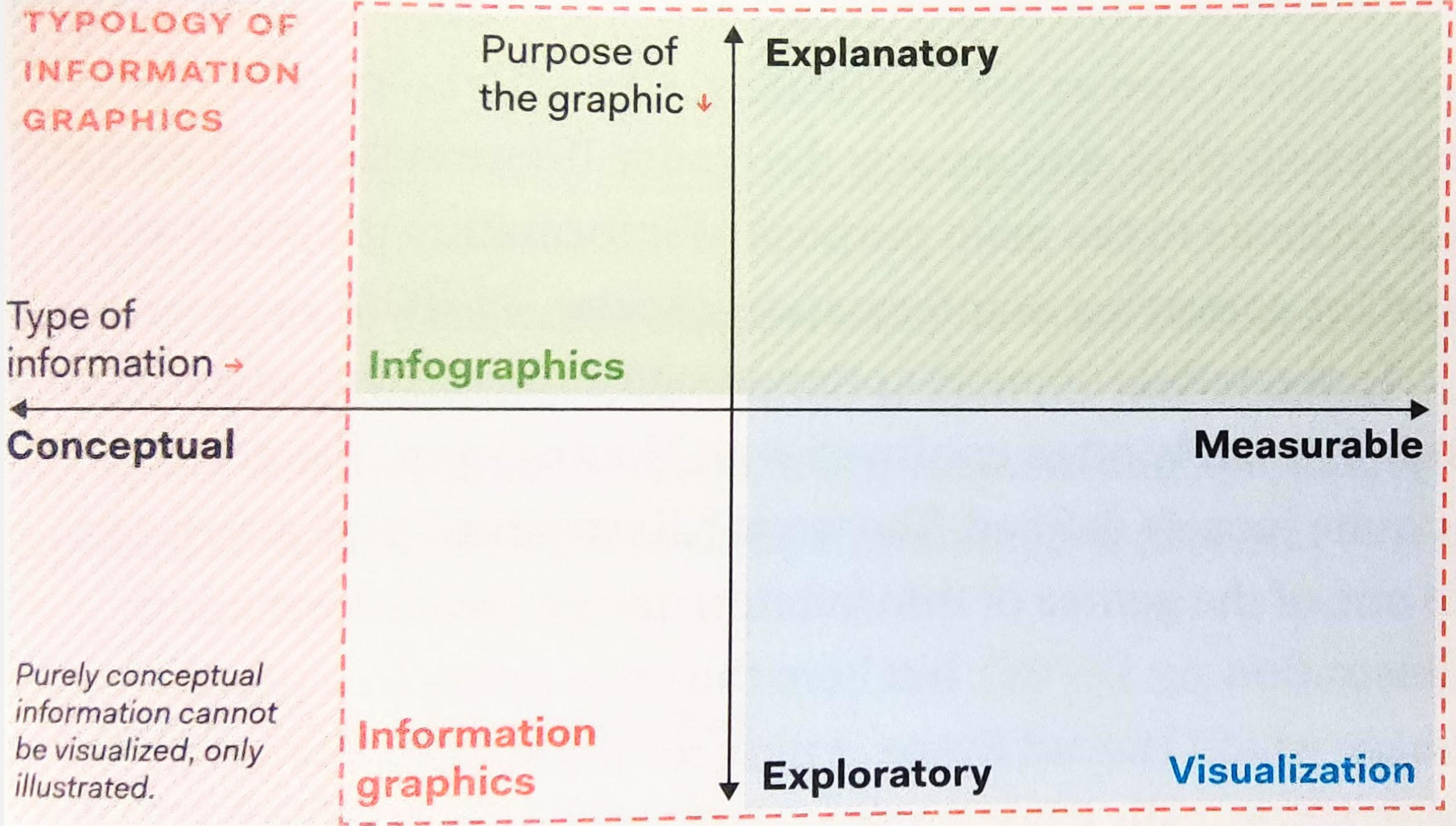
Is the **purpose** to **explore** or to **explain** the information?

☞ **Purpose of the graphic:** facilitate discovery <> communicate information



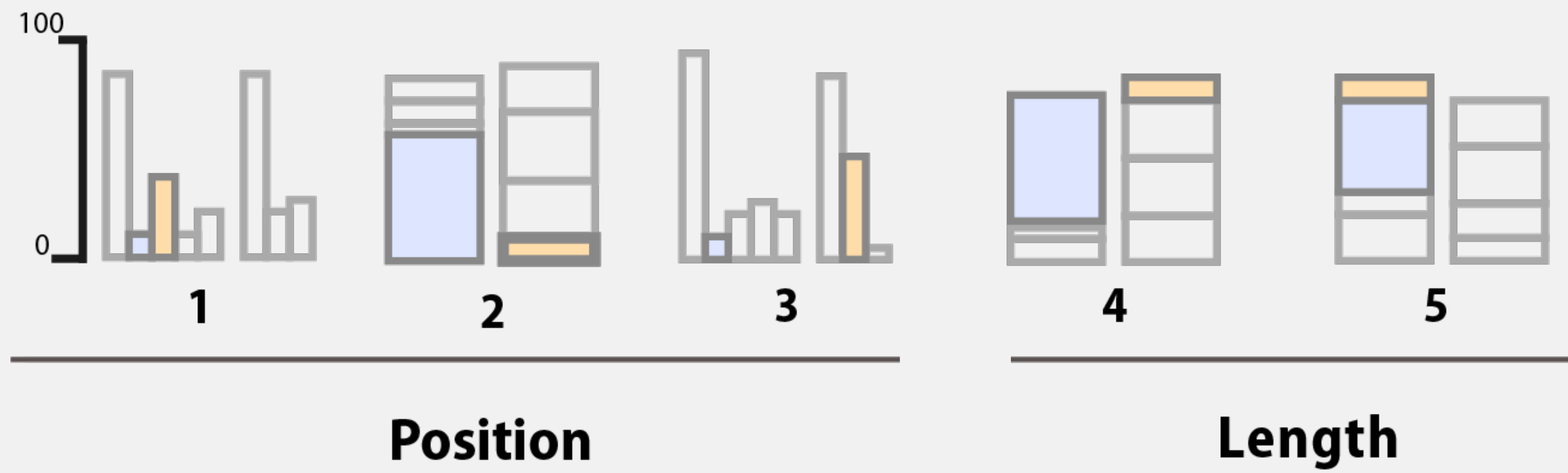
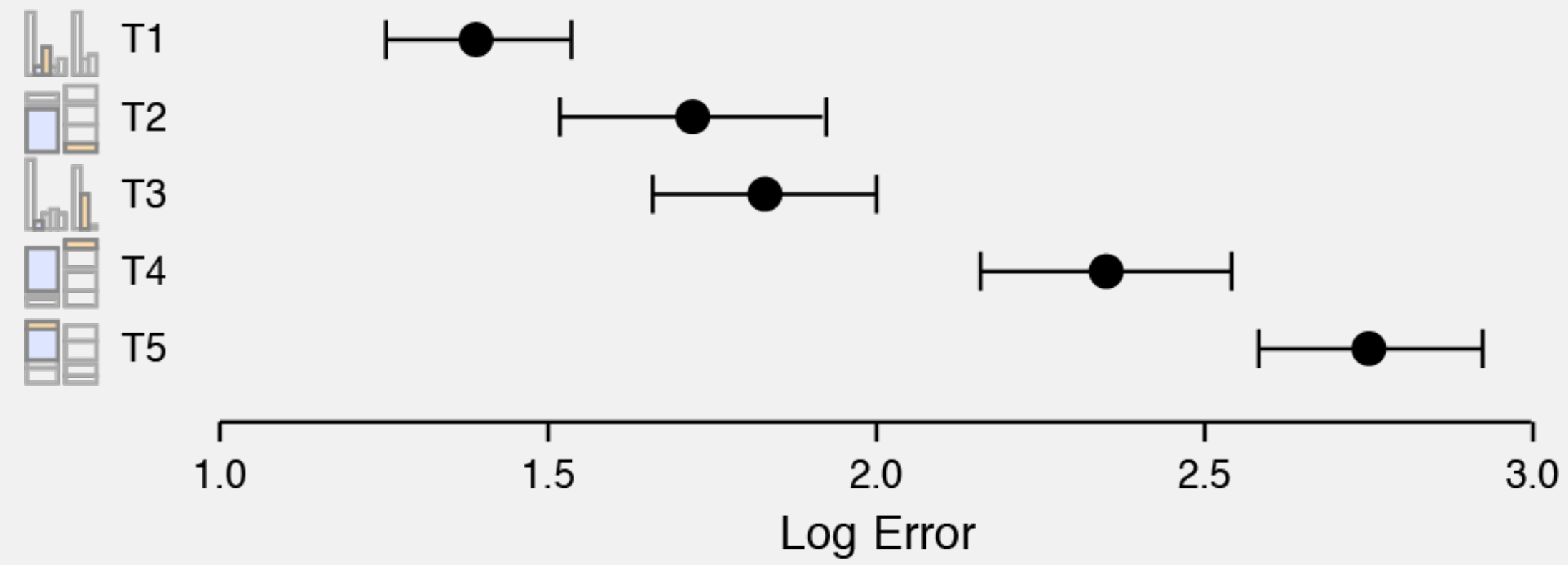
# Typology of Information Graphics

by Juuso Koponen & Jonatan Hildén, "Data Visualization Handbook" (2020), p. 25



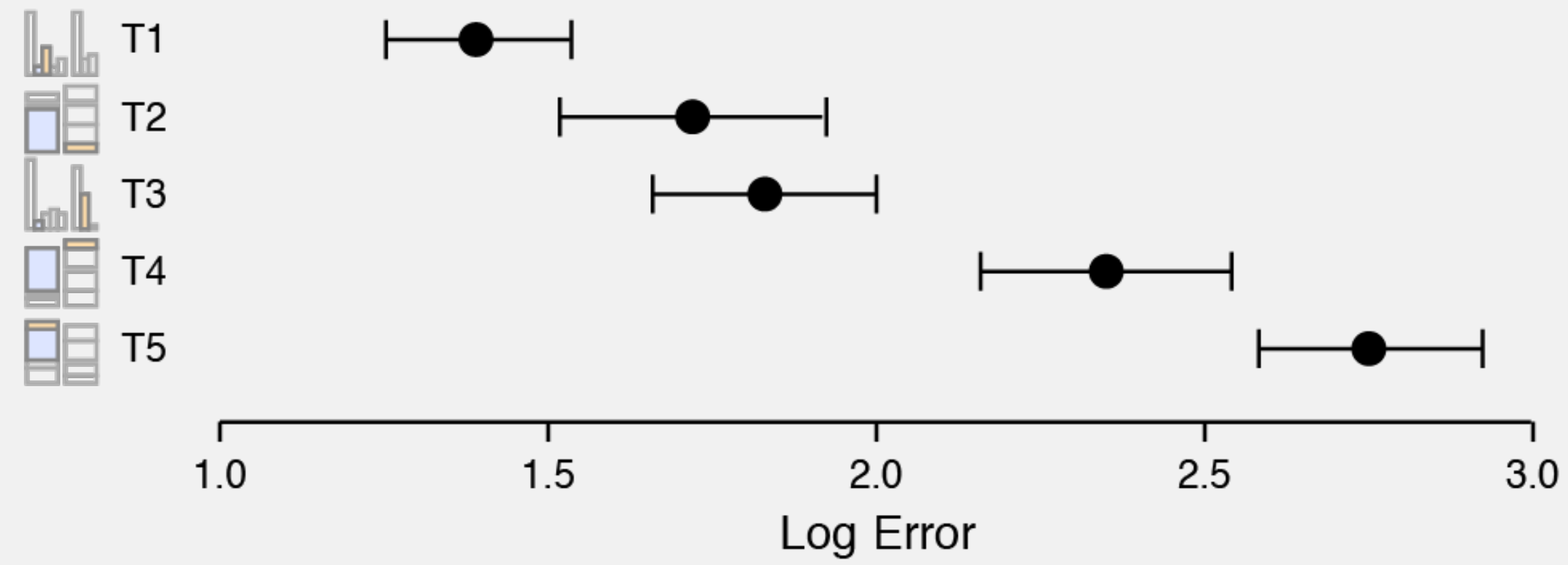


### Cleveland & McGill's Results

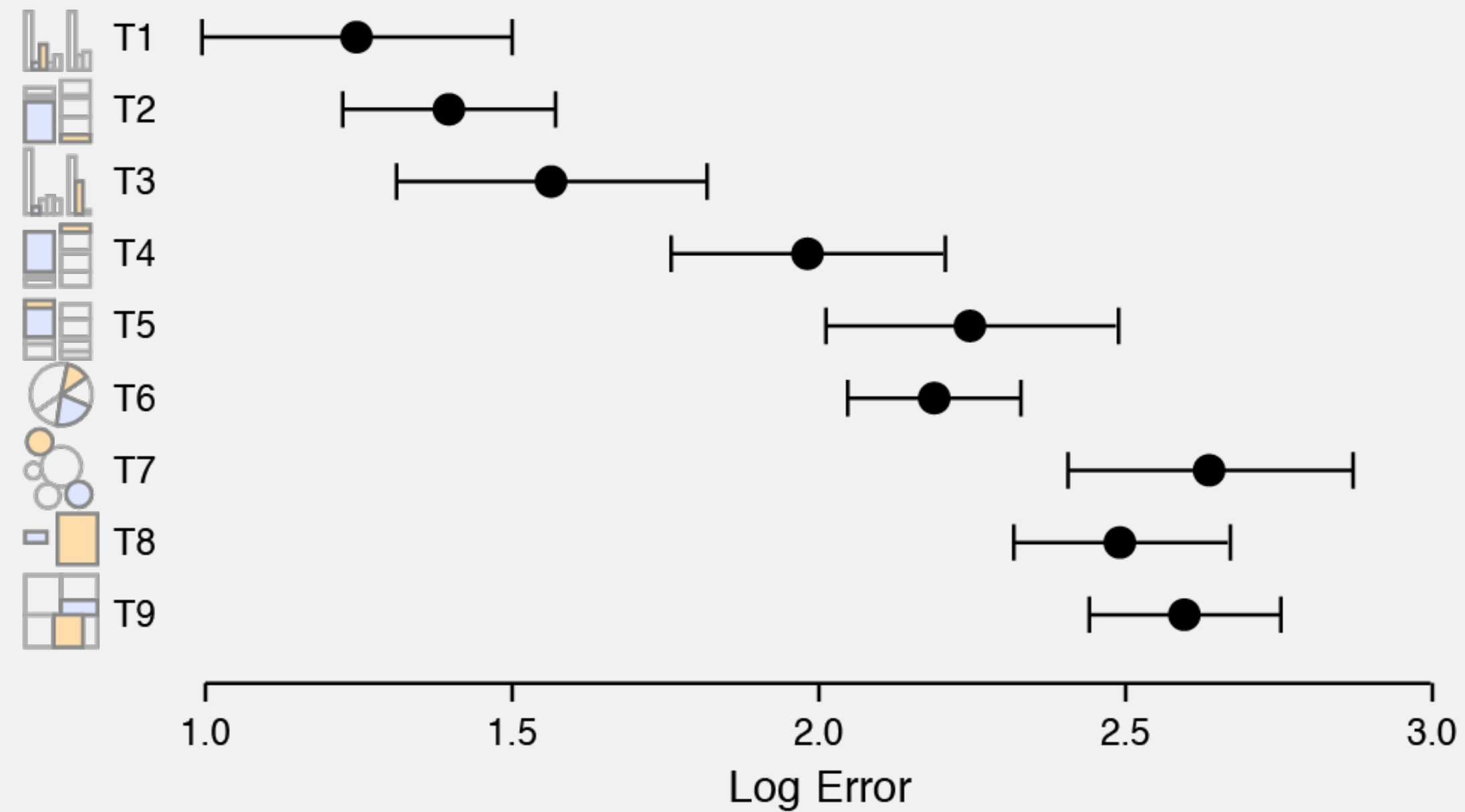


Source: Kieran Healy's ["Data Visualization: A Practical Introduction"](#); results based on Heer and Bostock, following Cleveland and McGill

### Cleveland & McGill's Results



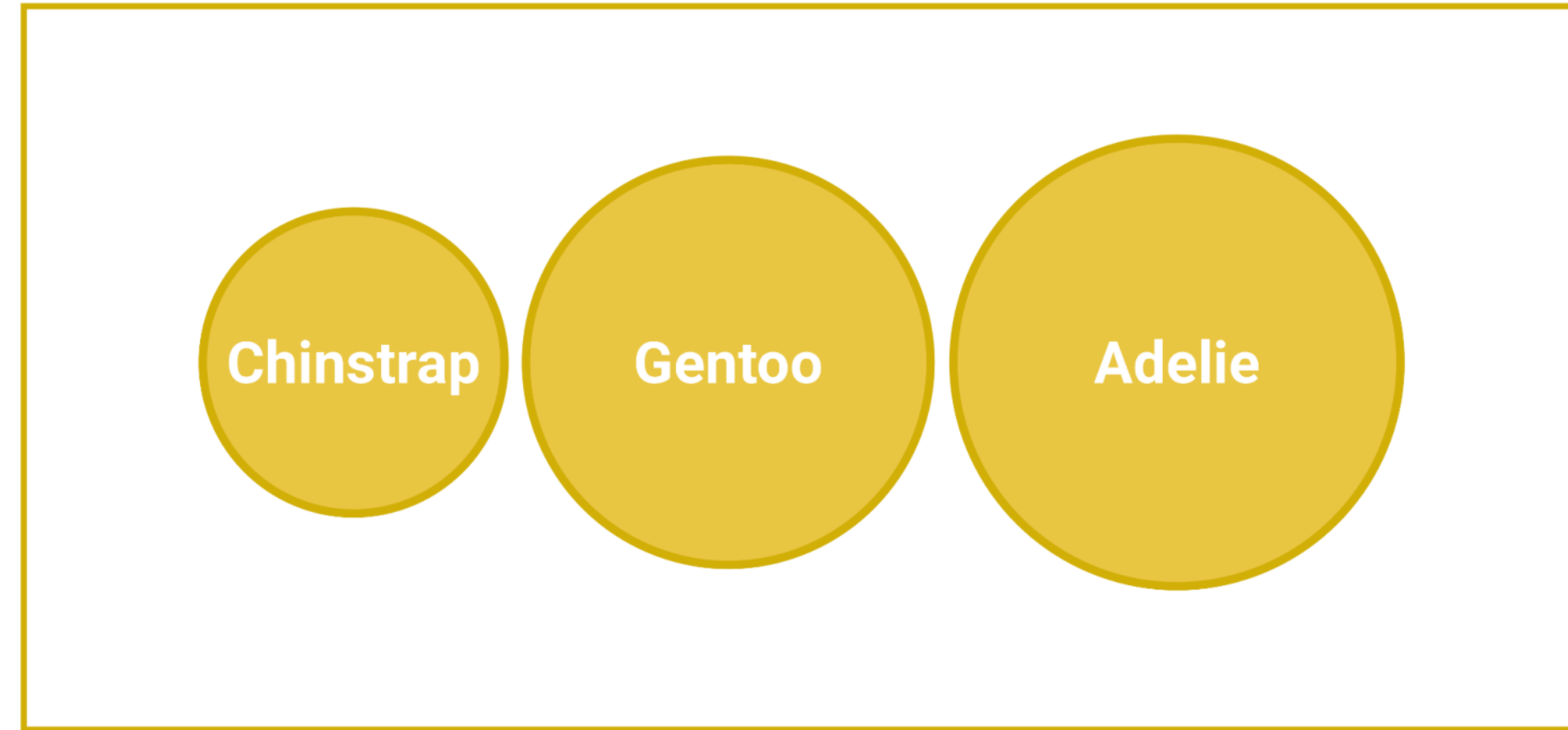
### Crowdsourced Results



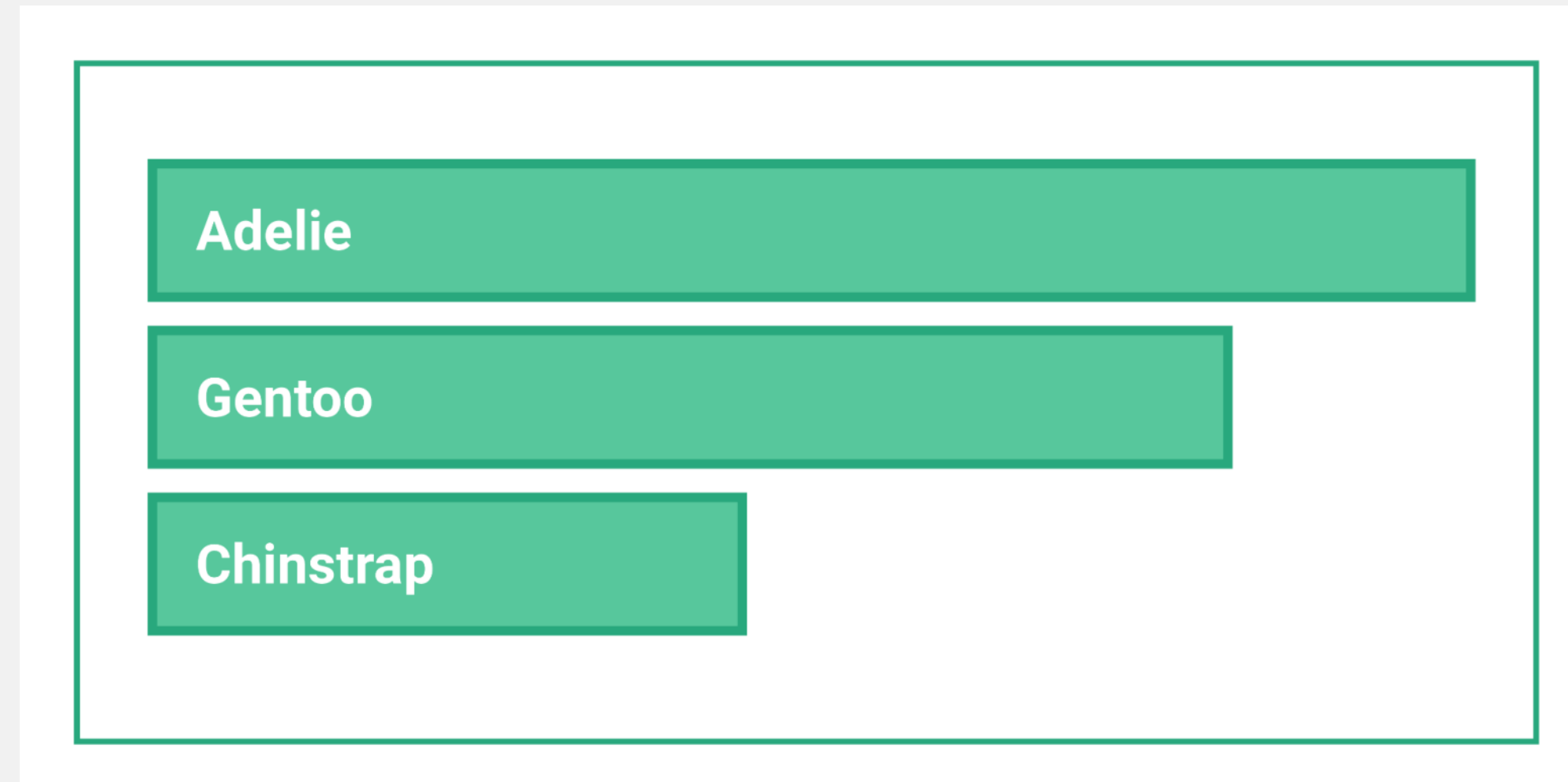
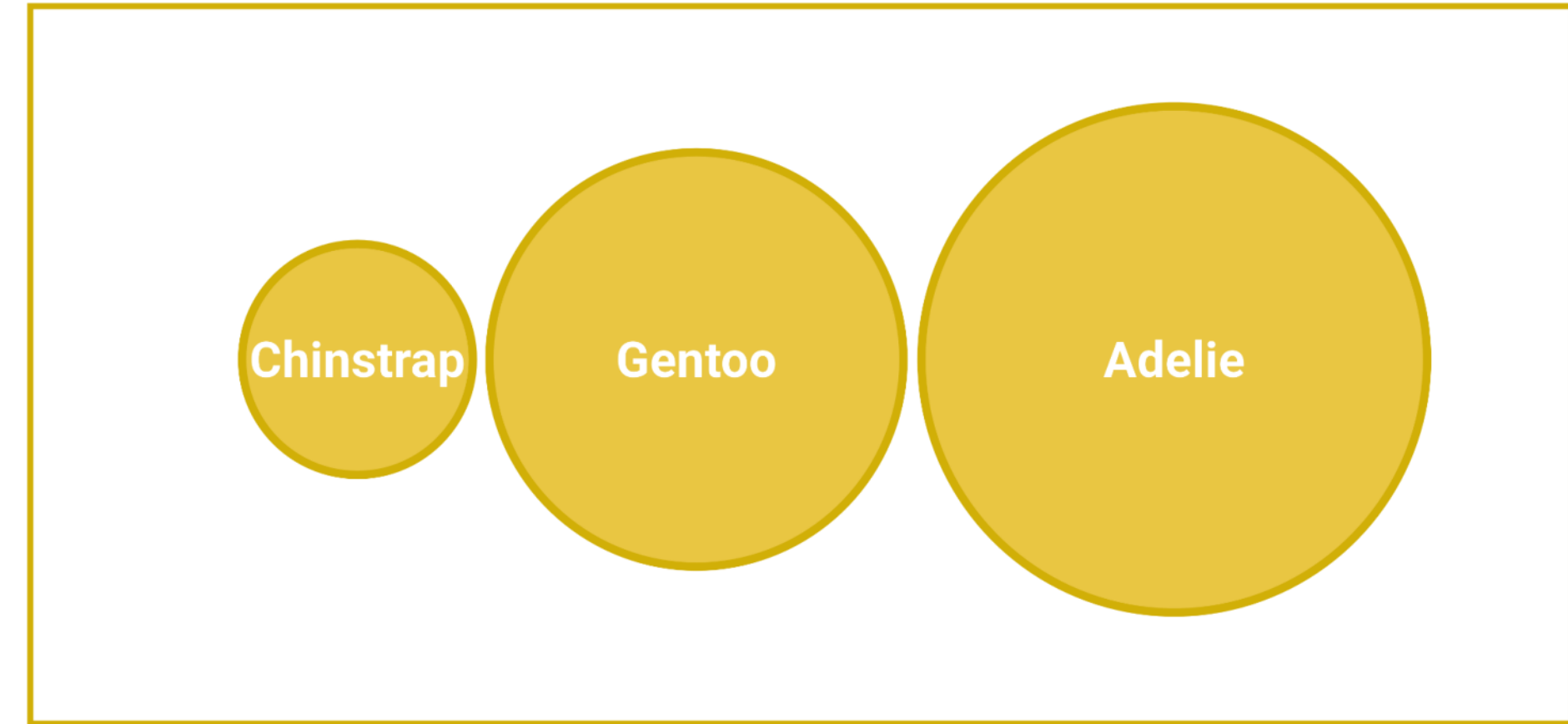
Source: Kieran Healy's ["Data Visualization: A Practical Introduction"](#); results based on Heer and Bostock, following Cleveland and McGill



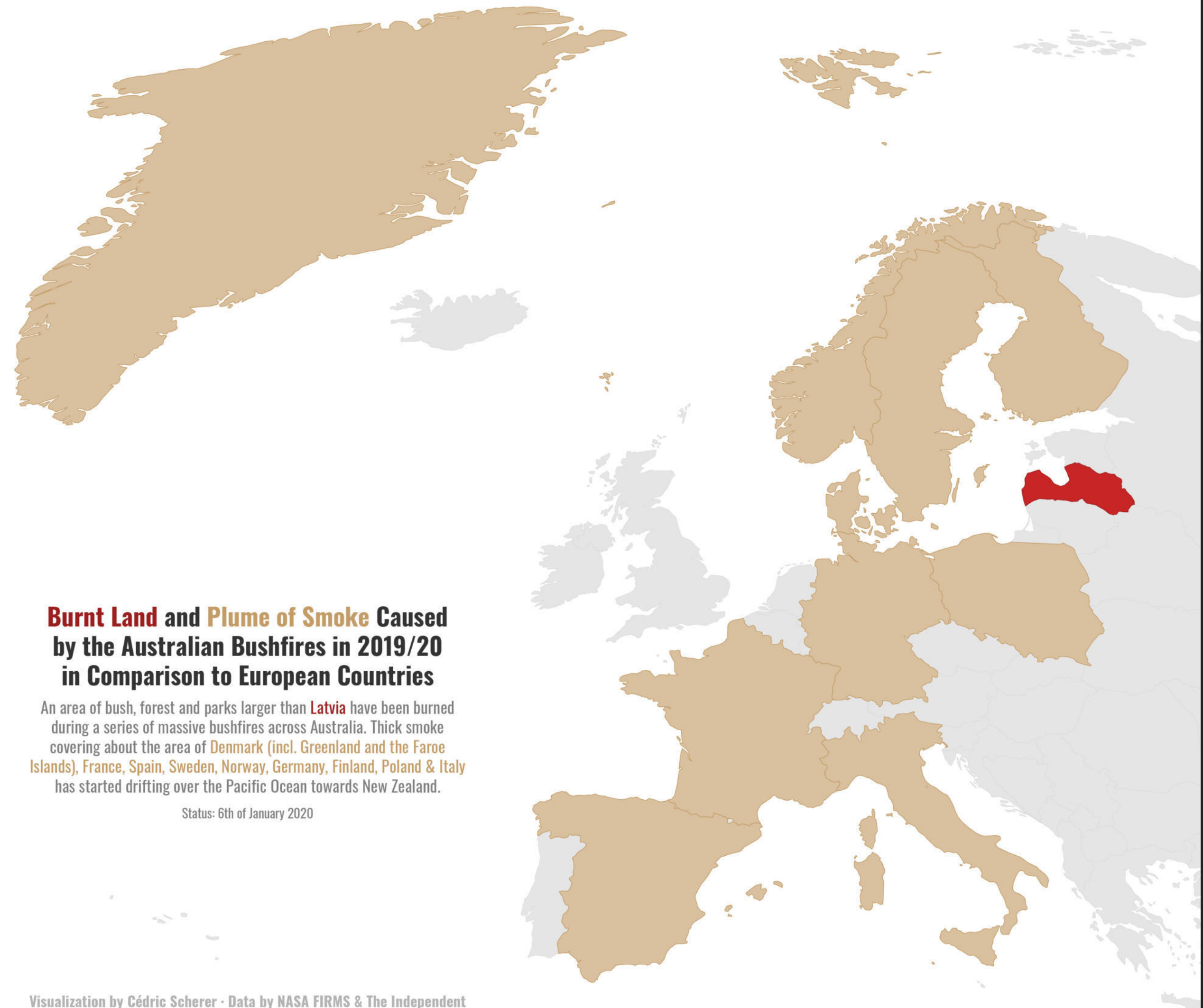
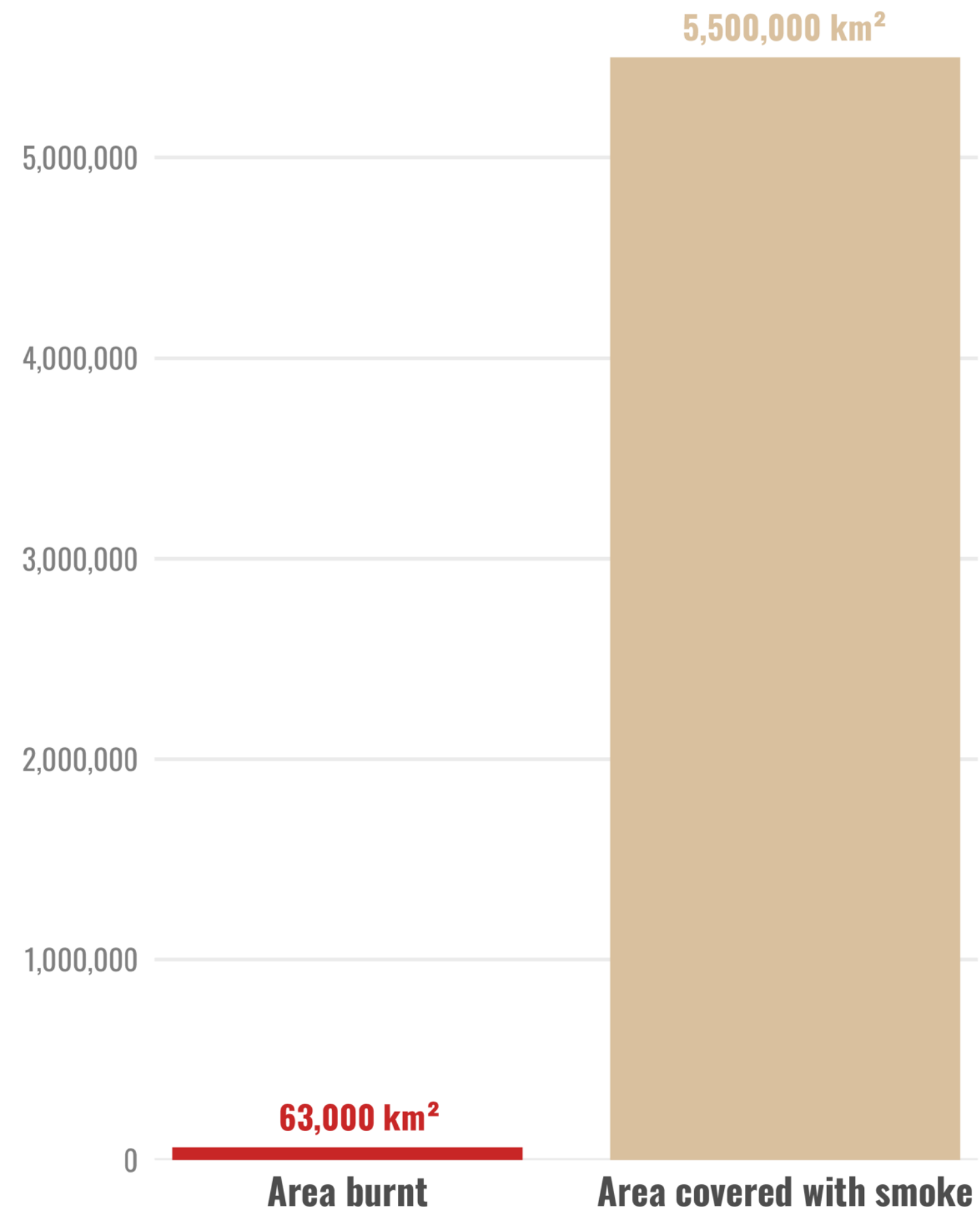
Always use area.



Never use radius!



**Burnt land and plume of smoke caused by the Australian bushfires in 2019/20**  
(as of 6<sup>th</sup> of January 2020)



PERSPECTIVE

# Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm

**Tracey L. Weissgerber<sup>1\*</sup>, Natasa M. Milic<sup>1,2</sup>, Stacey J. Winham<sup>3</sup>, Vesna D. Garovic<sup>1</sup>**

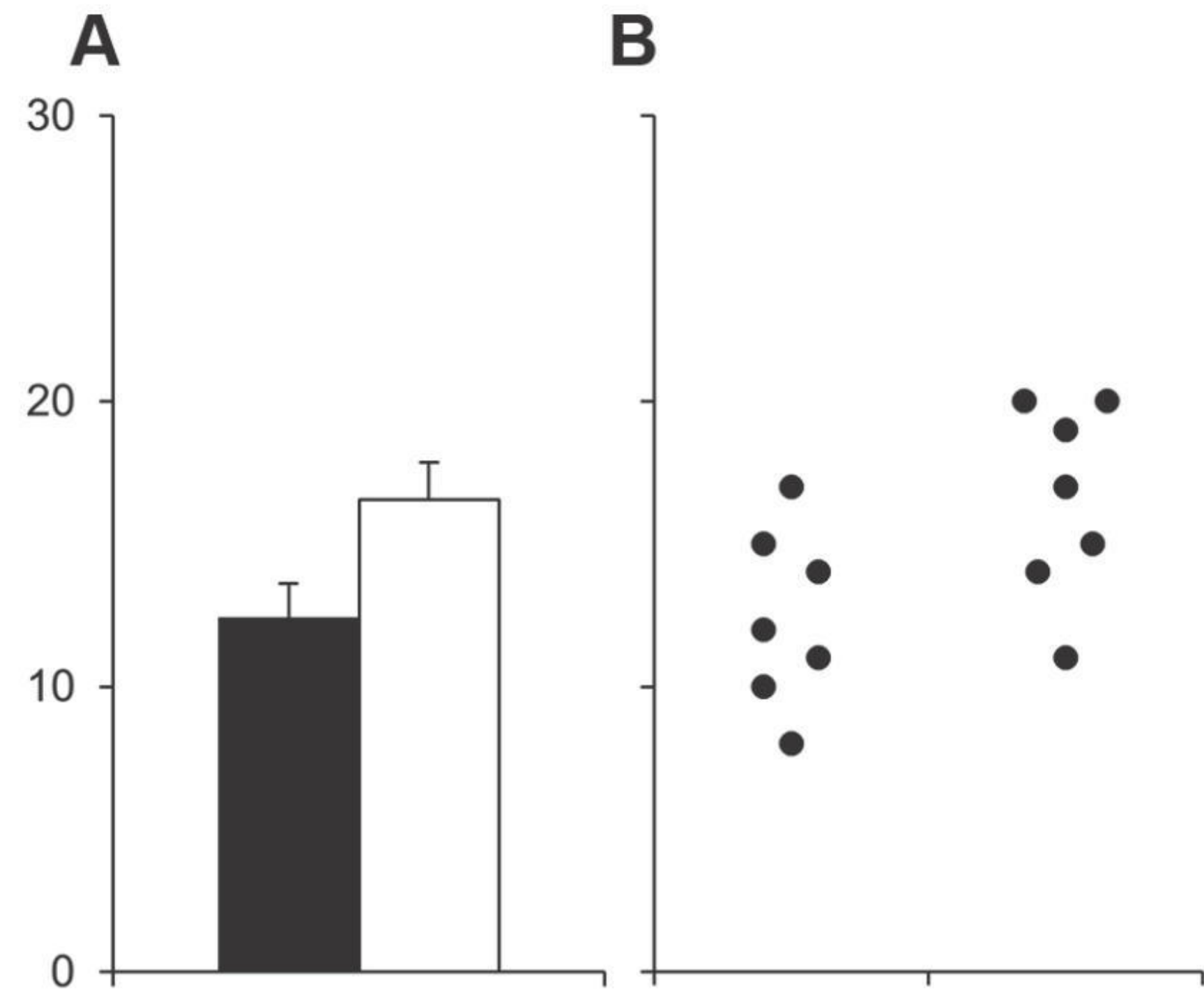
**1** Division of Nephrology & Hypertension, Mayo Clinic, Rochester, Minnesota, United States of America, **2** Department of Biostatistics, Medical Faculty, University of Belgrade, Belgrade, Serbia, **3** Division of Biomedical Statistic and Informatics, Mayo Clinic, Rochester, Minnesota, United States of America

\* [weissgerber.tracey@mayo.edu](mailto:weissgerber.tracey@mayo.edu)

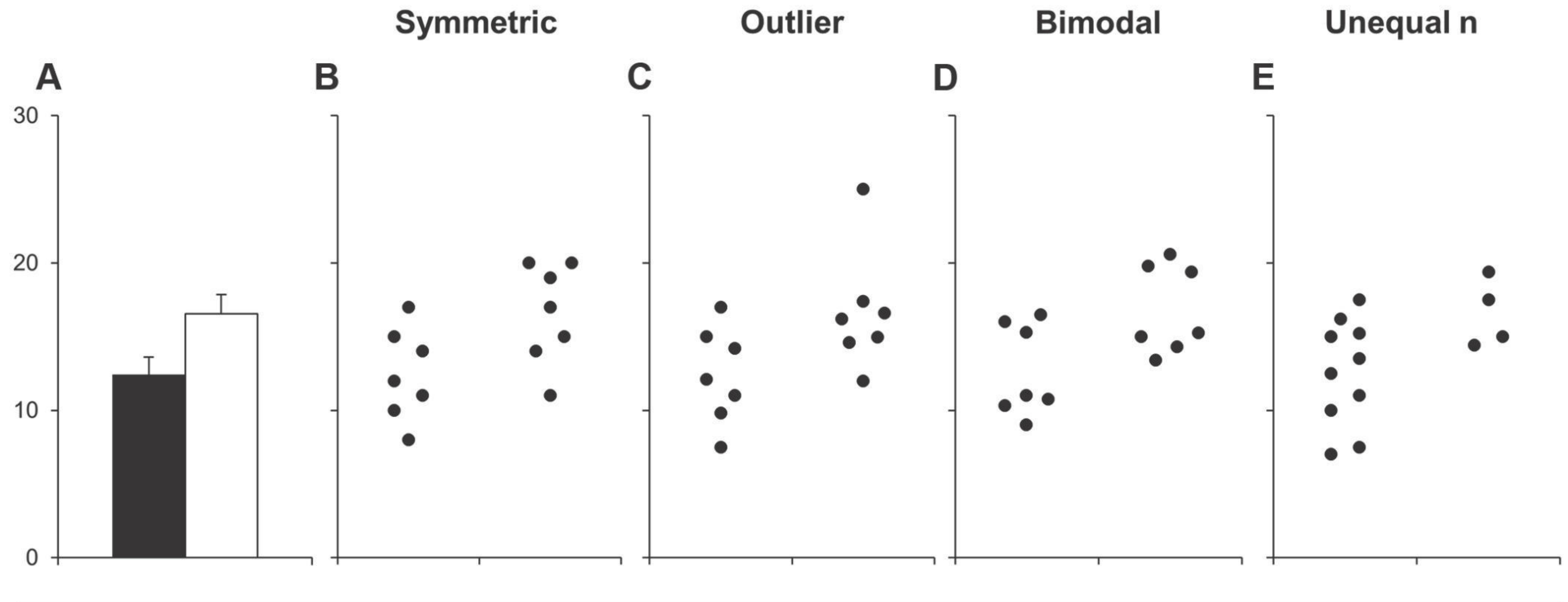
*Weissgerber et al. (2015) PLoS Biology*



# Symmetric

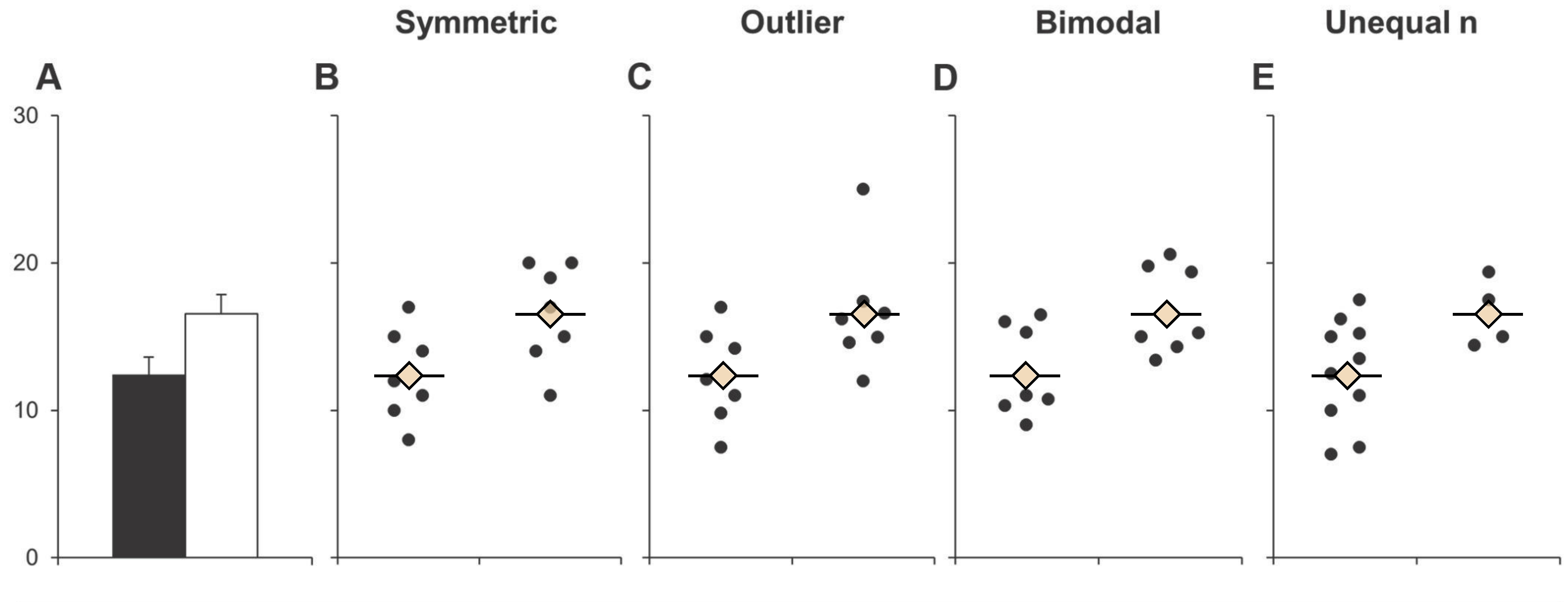


*Weissgerber et al. (2015) PLoS Biology*



Test	p value			
T-test: Equal var.	0.035	0.050	0.026	0.063
T-test: Unequal var.	0.035	0.050	0.026	0.035
Wilcoxon	0.054	0.073	0.128	0.103

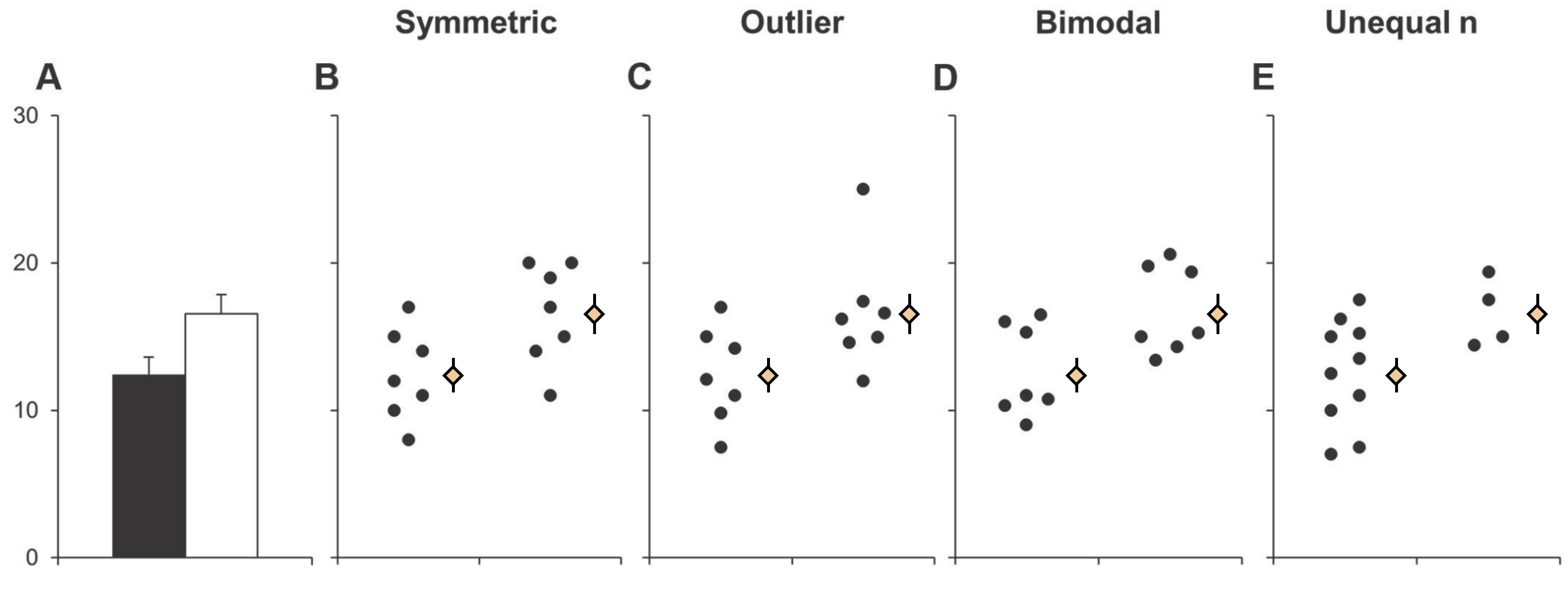
Weissgerber et al. (2015) PLoS Biology



Test	p value			
T-test: Equal var.	0.035	0.050	0.026	0.063
T-test: Unequal var.	0.035	0.050	0.026	0.035
Wilcoxon	0.054	0.073	0.128	0.103

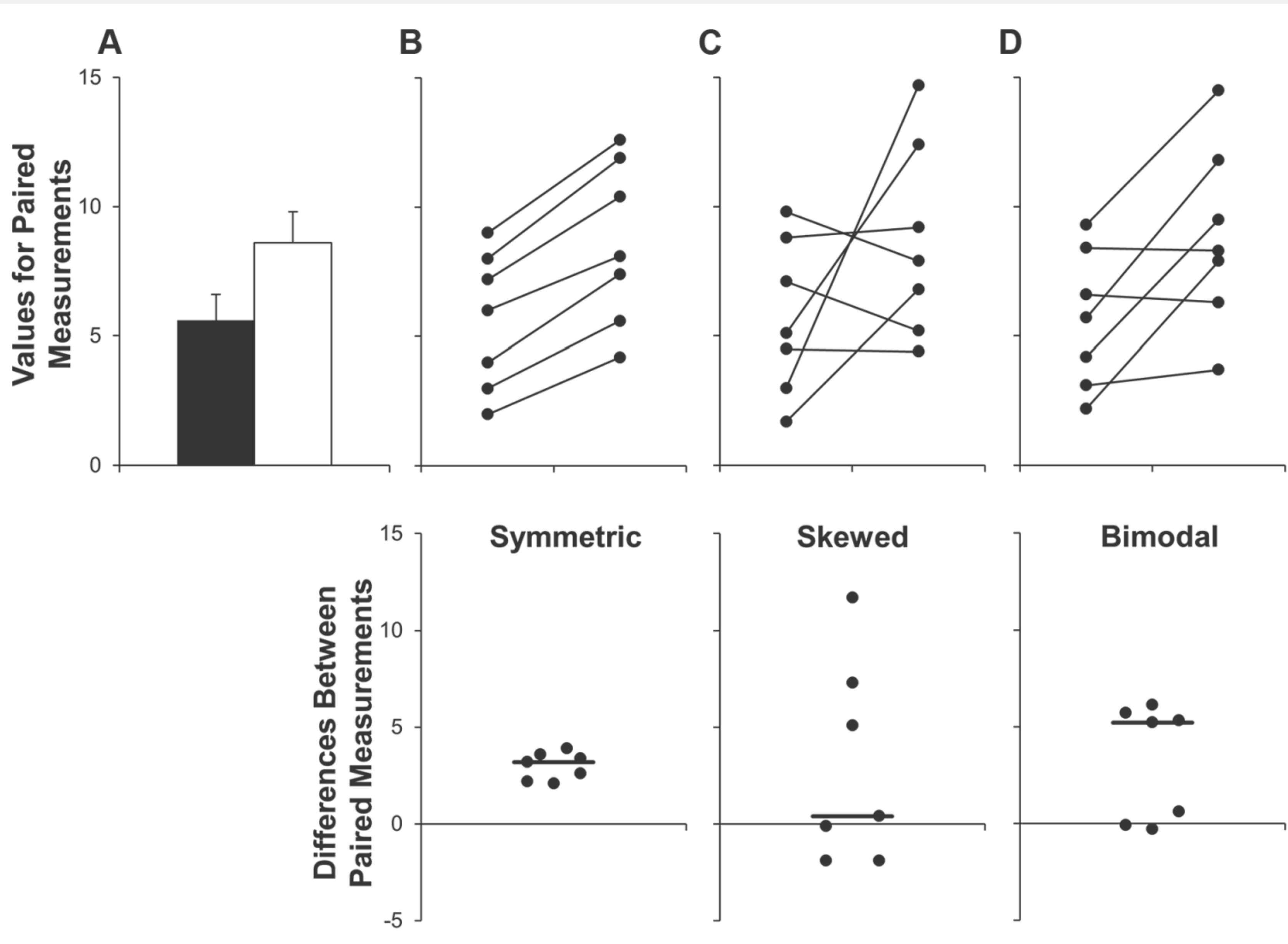
Modified from Weissgerber et al. (2015) PLoS Biology



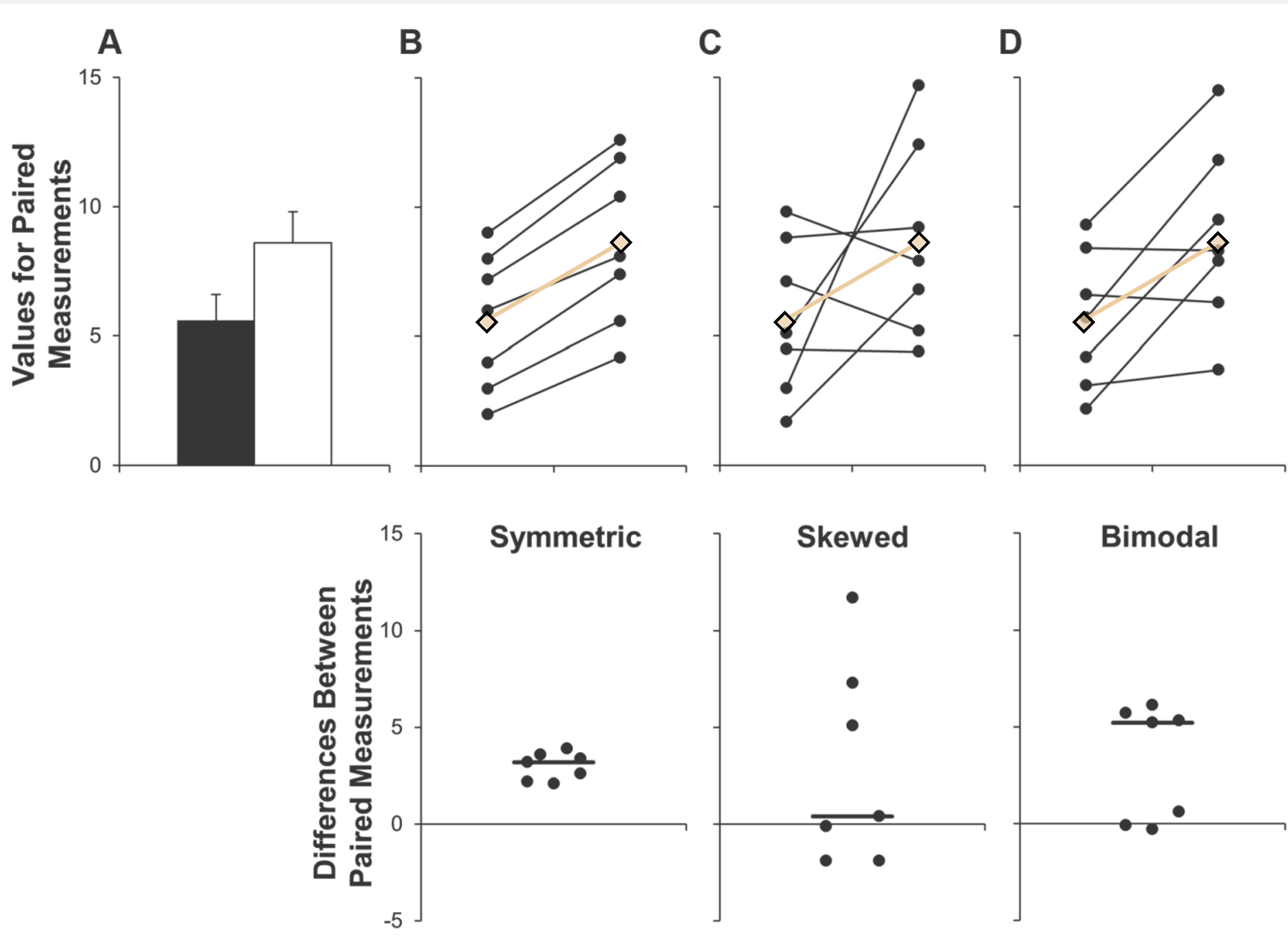


Test	p value			
T-test: Equal var.	0.035	0.050	0.026	0.063
T-test: Unequal var.	0.035	0.050	0.026	0.035
Wilcoxon	0.054	0.073	0.128	0.103

Modified from Weissgerber et al. (2015) PLoS Biology



Weissgerber et al. (2015) PLoS Biology



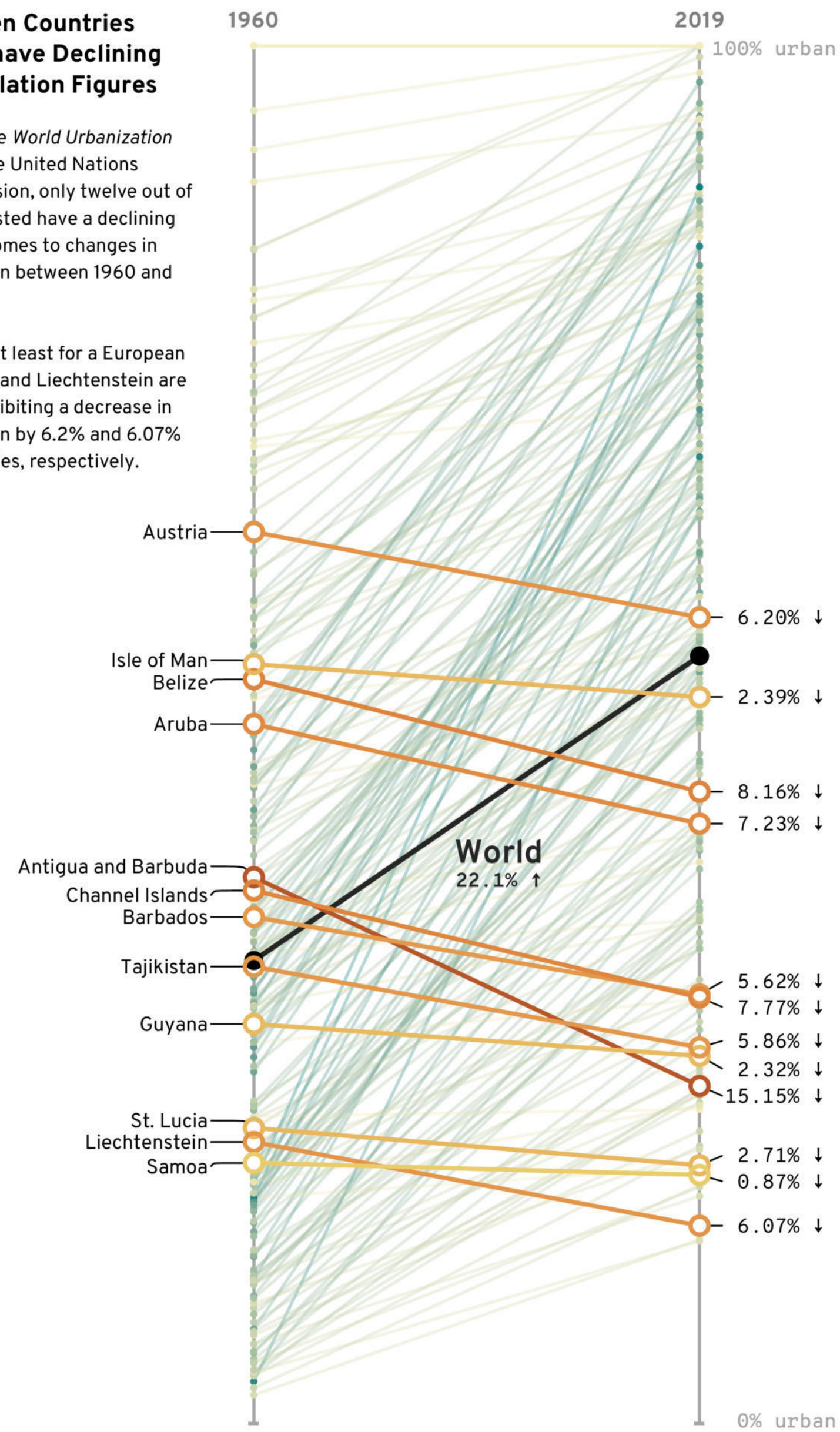
Weissgerber et al. (2015) PLoS Biology



## Only a Dozen Countries Worldwide have Declining Urban Population Figures

According to the *World Urbanization Prospects* by the United Nations Population Division, only twelve out of 218 countries listed have a declining trend when it comes to changes in urban population between 1960 and 2019.

Surprisingly—at least for a European citizen—Austria and Liechtenstein are among these exhibiting a decrease in urban population by 6.2% and 6.07% within six decades, respectively.



Visualization by Cédric Scherer | Data: United Nations Population Division, World Urbanization Prospects: 2018 Revision. | #30DayChartChallenge 2021 Day 5: Slope



# Not my cup of coffee...

Each dot depicts one coffee bean rated by Coffee Quality Institute's trained reviewers. In addition, the multiple interval stripes show where 25%, 50%, 95%, and 100% of the beans fall along the rating gradient from 0 to 100 points. The rated coffee beans range from 59.8 points (Guatemala) to 89.9 (Ethiopia). Only countries of origin with 25 or more tested beans are shown. The red empty triangle marks the minimum rating, the black filled triangle indicates each country's median score.

Visualization by Cédric Scherer

60 POINTS

70 POINTS

80 POINTS

90 POINTS

## GUATEMALA

▲ 59.8 POINTS

The coffee bean with the lowest rating has its origin in Guatemala.



One bean from Nicaragua got a bad rating, too.

▲ 63.1 POINTS

## NICARAGUA

## COLOMBIA

▲ 72.8 POINTS

## COSTA RICA

▲ 71.8 POINTS

## HAWAII

▲ 73.7 POINTS

## BRAZIL

▲ 73.2 POINTS

## HONDURAS

▲ 69.2 POINTS

## MEXICO

▲ 68.3 POINTS

## ETHIOPIA

▲ 80.3 POINTS

▲ 85.1 POINTS

The best coffee—in terms of both median and maximum rating—is shipped to you from Ethiopia!

## KENYA

▲ 79.8 POINTS

▲ 84.6 POINTS

## UGANDA

▲ 80.5 POINTS

▲ 83.2 POINTS

▲ 83.2 POINTS

▲ 82.8 POINTS

▲ 82.5 POINTS

▲ 82.4 POINTS

▲ 80.3 POINTS

▲ 82.2 POINTS

## TAIWAN

▲ 77.7 POINTS

▲ 81.9 POINTS

▲ 81.7 POINTS

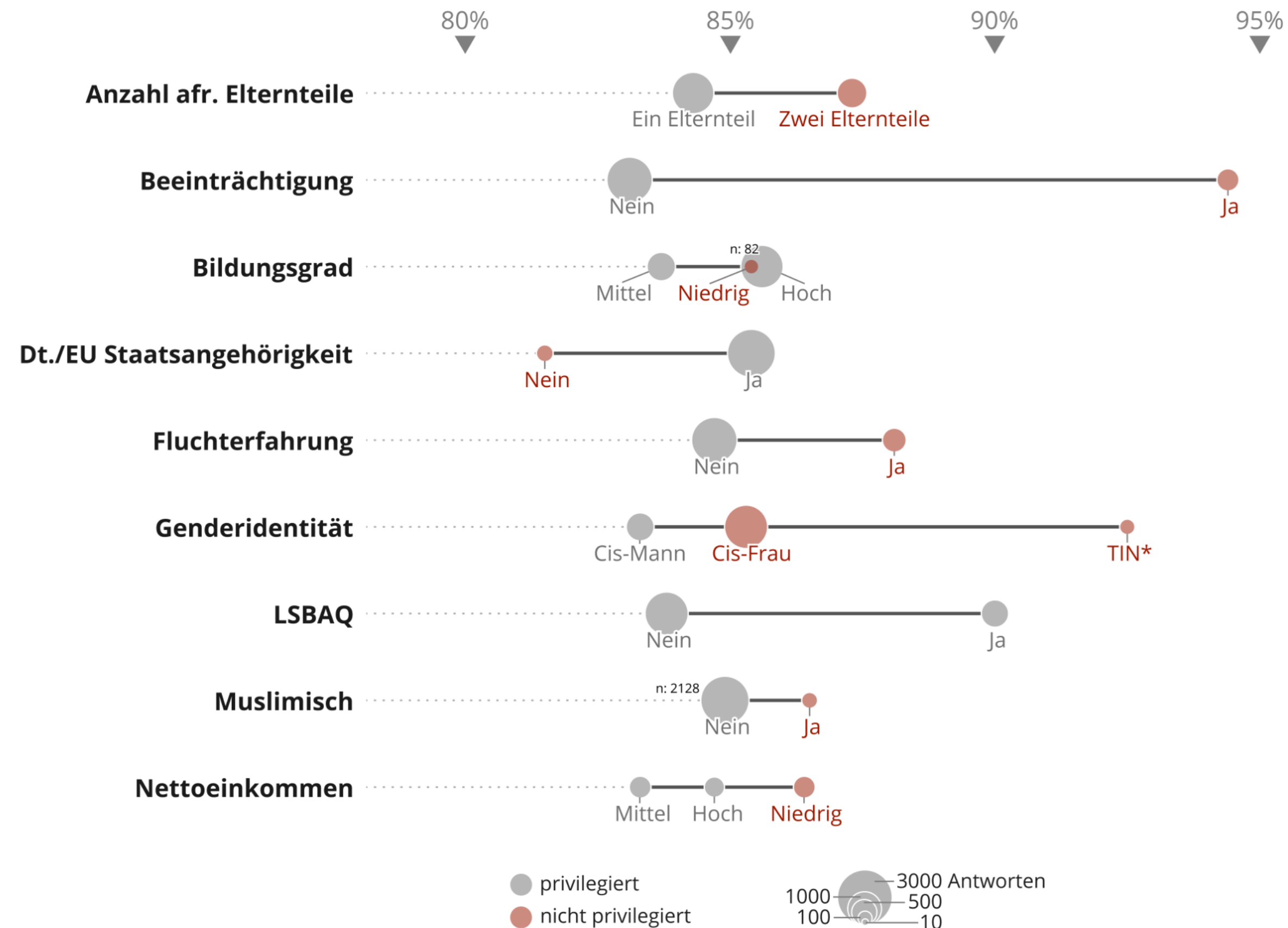
▲ 81.6 POINTS

▲ 80.8 POINTS

With 218 tested beans, Mexico is the country with the most reviews.

“Not my cup of coffee”, #TidyTuesday Contribution

## Häufigkeit von Diskriminierungserfahrungen entlang ausgewählter Vielfaltdimensionen im Lebensbereich „Medien und Internet“



**Lesebeispiel:** LSBAQ-Befragte des Afrozensus geben im Vergleich mit heterosexuellen Afrozensus-Befragten häufiger an, im Lebensbereich „Medien und Internet“ in den letzten zwei Jahren Diskriminierung erlebt zu haben.

Quelle: Abb. 46 in Aikins, M A; Bremberger, T; Aikins, J K; Gyamerah, D; Yıldırım-Calıman, D (2021): Afrozensus 2020 | Datenteam: Reiber, L; Vivanco, J | Design: Scherer, C  
 Lizenz: CC-BY-NC by EOTO & CFE | afrozensus.de

Abb. 46 “Afrozensus 2020” by Citizens For Europe & EOTO e.V.



From Data to Viz

EXPLORE STORY ALL CAVEATS POSTER ABOUT CONTACT

From Data to Viz leads you to the most appropriate graph for your data. It links to the code to build it and lists common caveats you should avoid.

EXPLORE

What kind of data do you have? Pick the main type using the buttons below. Then let the decision tree guide you toward your graphic possibilities.

NUMERIC CATEGORICAL NUM & CAT MAPS NETWORK TIME SERIES

ONE NUMERIC VARIABLE

TWO NUMERIC VARIABLES

NOT ORDERED

ORDERED

ONE CATEGORICAL VARIABLE

TWO CATEGORICAL VARIABLES

THREE NUMERIC VARIABLES

SEVERAL NUMERIC

HISTOGRAM

DENSITY PLOT

SCATTER PLOT

SCATTER PLOT WITH REGRESSION LINE

BOX PLOT

VIOLIN PLOT

AREA PLOT

LINE PLOT

PIE CHART

DONUT CHART

SANKEY DIAGRAM

WATERFALL PLOT

CHORD DIAGRAM

HEAT MAP

CHOROPLETH MAP

MAP

TABLE CHART

PROPORTIONAL AREA CHART (CIRCLE)

VIOLIN PLOT

PIE CHART

GROUPED BAR CHART

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

D V P ALL FAMILY INPUT FUNCTION SHAPE Q

ferdio

Alluvial Diagram

Sankey Diagram

Donut Chart

Line Graph

Radial Bar Chart

Polar Area Chart

Pictorial Fraction Chart

Radial Histogram

Exploded View Drawing

Sorted Stream Graph

Bar Chart (Vertical)

Flow Map

Sunburst Diagram

Treemap

Pictorial Stacked Chart

Stacked Bar Chart

Arc Diagram

Chord Diagram

Heat Map

Choropleth Map

Matrix Diagram (Roof Shaped)

Fishbone Diagram

Progress Bar

Proportional Area Chart (Squares)

Hexagonal Binning

Radial Line Graph

Pictorial Fraction Chart

Matrix Diagram

Multi-level Donut Chart

Waterfall Plot

Word Cloud

Stream Graph

Hyperbolic Tree

Isoline Map

Bubble Map

Table Chart

Proportional Area Chart (Circle)

Violin Plot

Pie Chart

Grouped Bar Chart

[datavizproject.com](http://datavizproject.com)

Charts

sort by name search interest change in search interest

bar chart

gantt chart

histogram

pie chart

venn diagram

time series

scatter plot

box plot

contour line

heat map

decision tree

block diagram

ishikawa diagram

candlestick chart

word cloud

line chart

bubble chart

sankey diagram

radar chart

waterfall chart

tree diagram

tree map

sparkline

self-organizing map

density plot

flow map

dendrogram

choropleth map

cartogram

area chart

violin plot

tape diagram

doughnut chart

chord diagram

timeline

sunburst chart

stream graph

small multiples

radial tree

pyramid diagram

proportional symbol map

polar chart

pictorial chart

parallel coordinates

onion diagram

marimekko chart

kagi chart

hyperbolic tree

hive plot

funnel chart

dot distribution map

dasymetric map

chernoff face

bullet chart

arc diagram

alluvial diagram

Are we missing any good basic/chart? Send us suggestions.

Data

This data is updated daily for all the 10,000+ data points shown here. If you'd like to know more about what data was used and what choices made, read on.

Team

The Visualization Universe was created by Anna Vital and Mark Vital from Adoma and Alexander Vasilakis with the assistance of the Google News Lab, Simon Rogers and Alberto Casio for art direction.

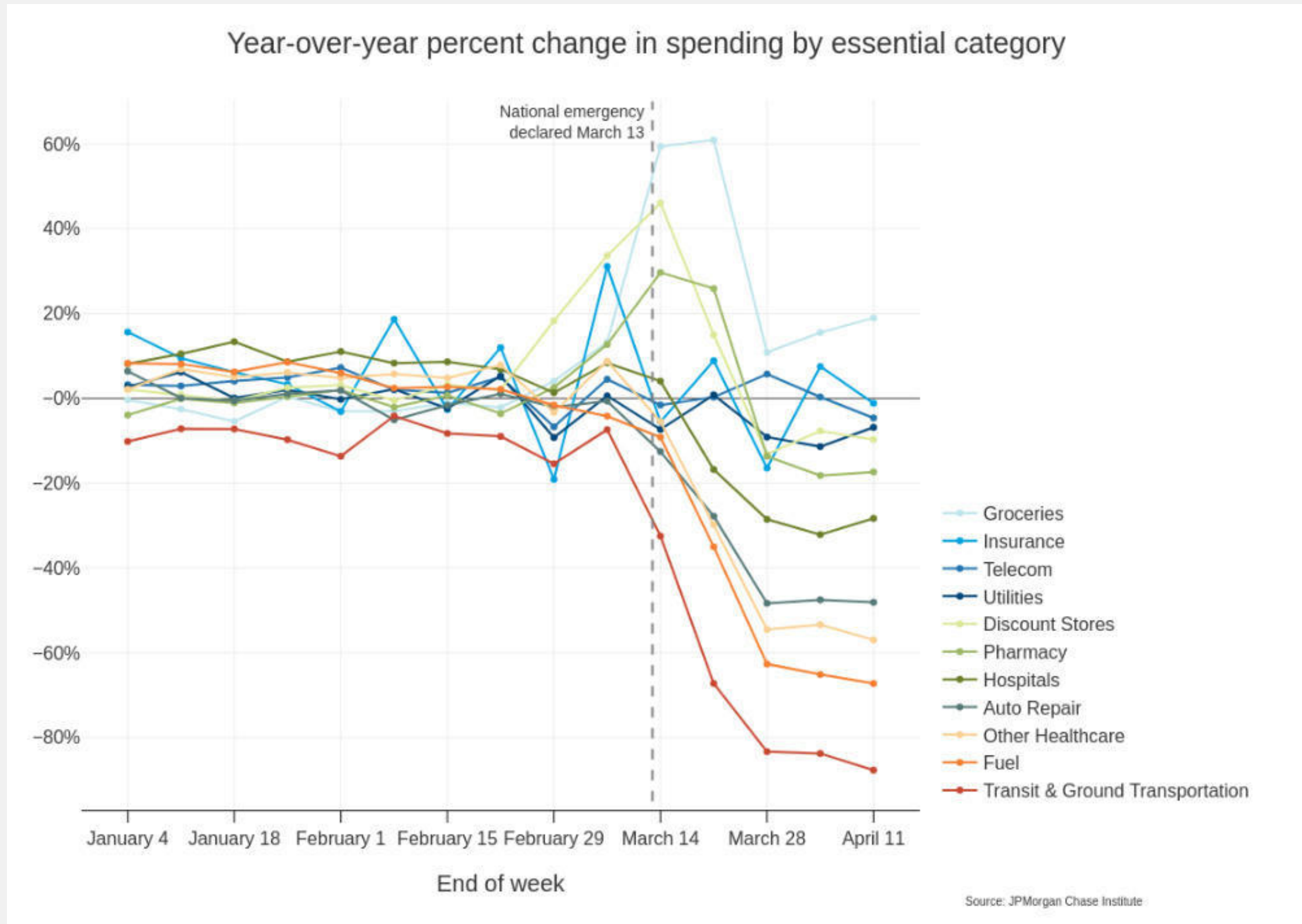
Contact

Contact us at visualization@adoma.com for questions about this project.

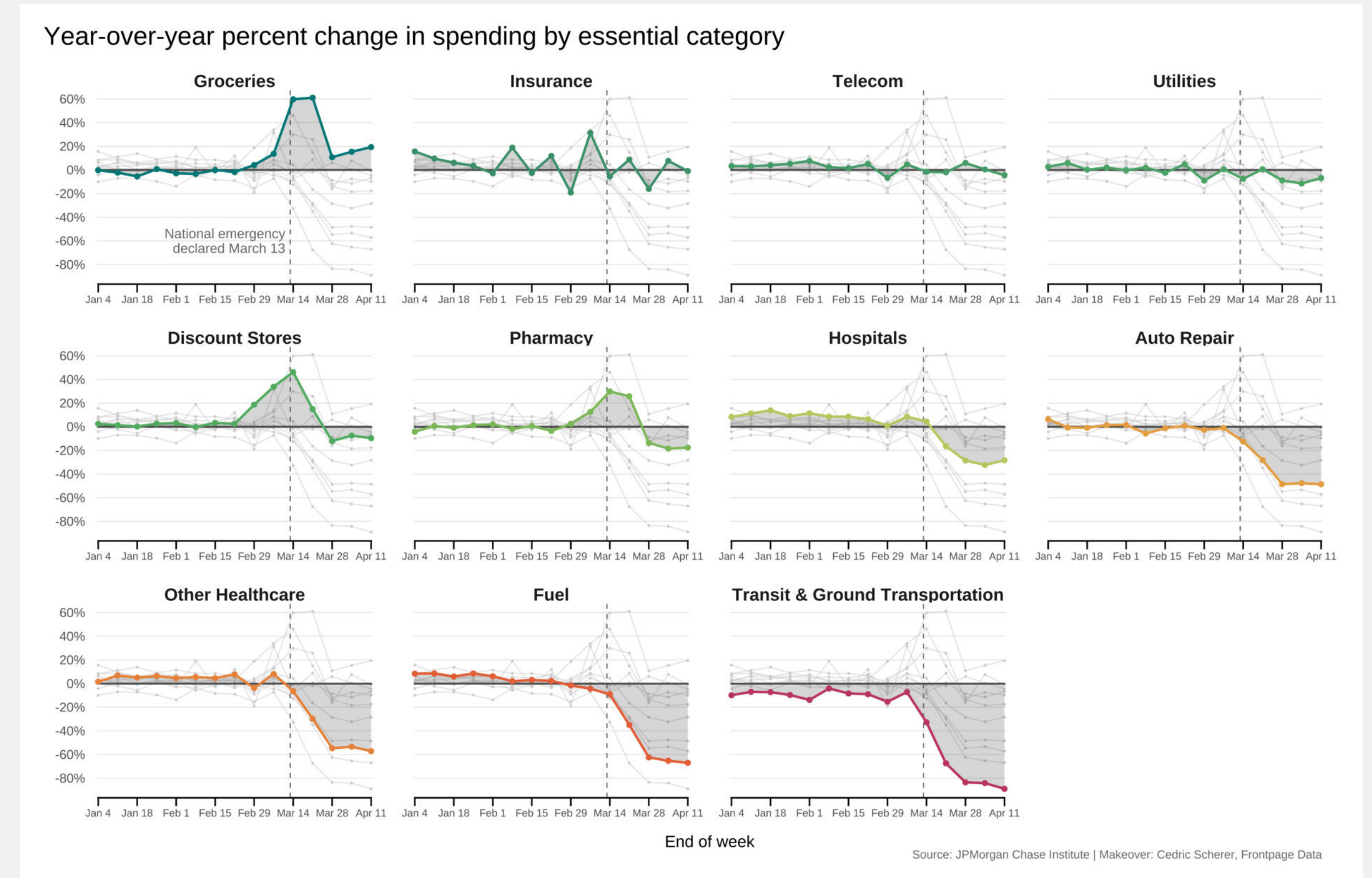
[visualizationuniverse.com](http://visualizationuniverse.com)



# The Power of Small Multiples



Original graphic by JPMorgan Chase Institute



Makeover using small multiples



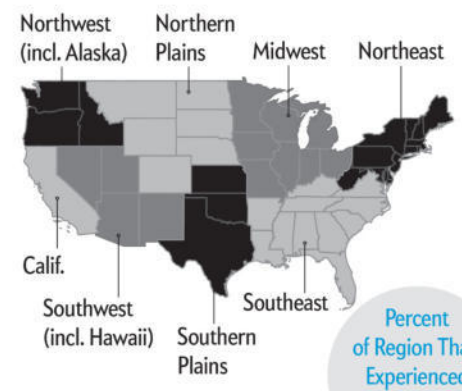
**GRAPHIC SCIENCE**

Text by Clara Moskowitz | Graphic by Cédric Scherer and Georgios Karamanis

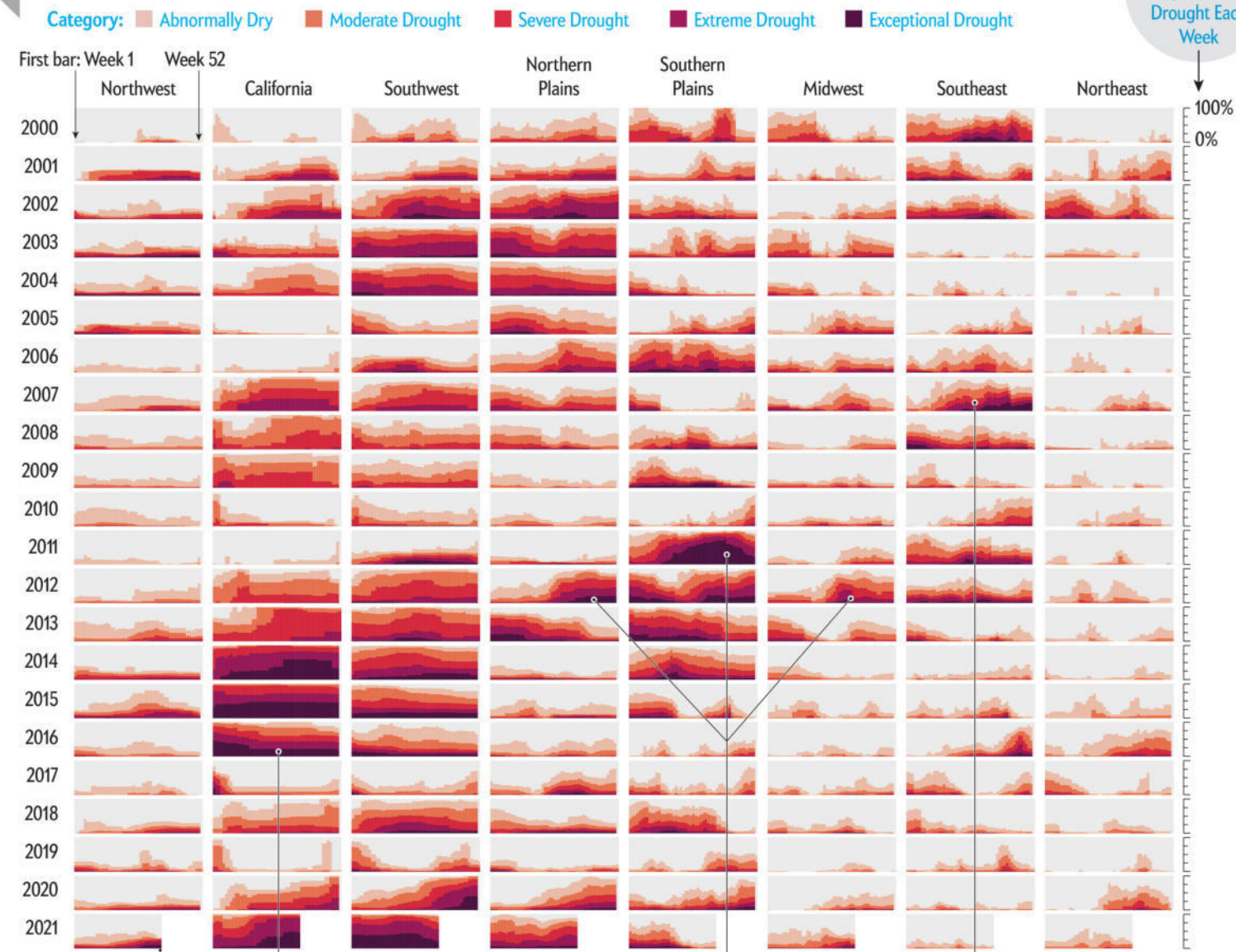
# Escalating Drought

Climate change is intensifying periods of extreme dryness, particularly in the U.S. West

For more than 20 years the National Drought Mitigation Center (NDMC) has been monitoring dozens of indices of drought around the country, including satellite measurements of evaporation and color in vegetation, soil-moisture sensors, rainfall estimates, and river and streamflow levels. Although the agency's weekly assessments have identified periods of exceptional drought before, lately dryness has been ramping up. "The changing climate is definitely contributing to more natural disasters, drought being one of them," says Brian Fuchs, a climatologist who oversees the weekly report at the NDMC. "We're seeing more frequent and high-intensity episodes. This year some of these areas in the West have been in drought more than they have been without drought."



**Drought Extent and Intensity by Region over Time**

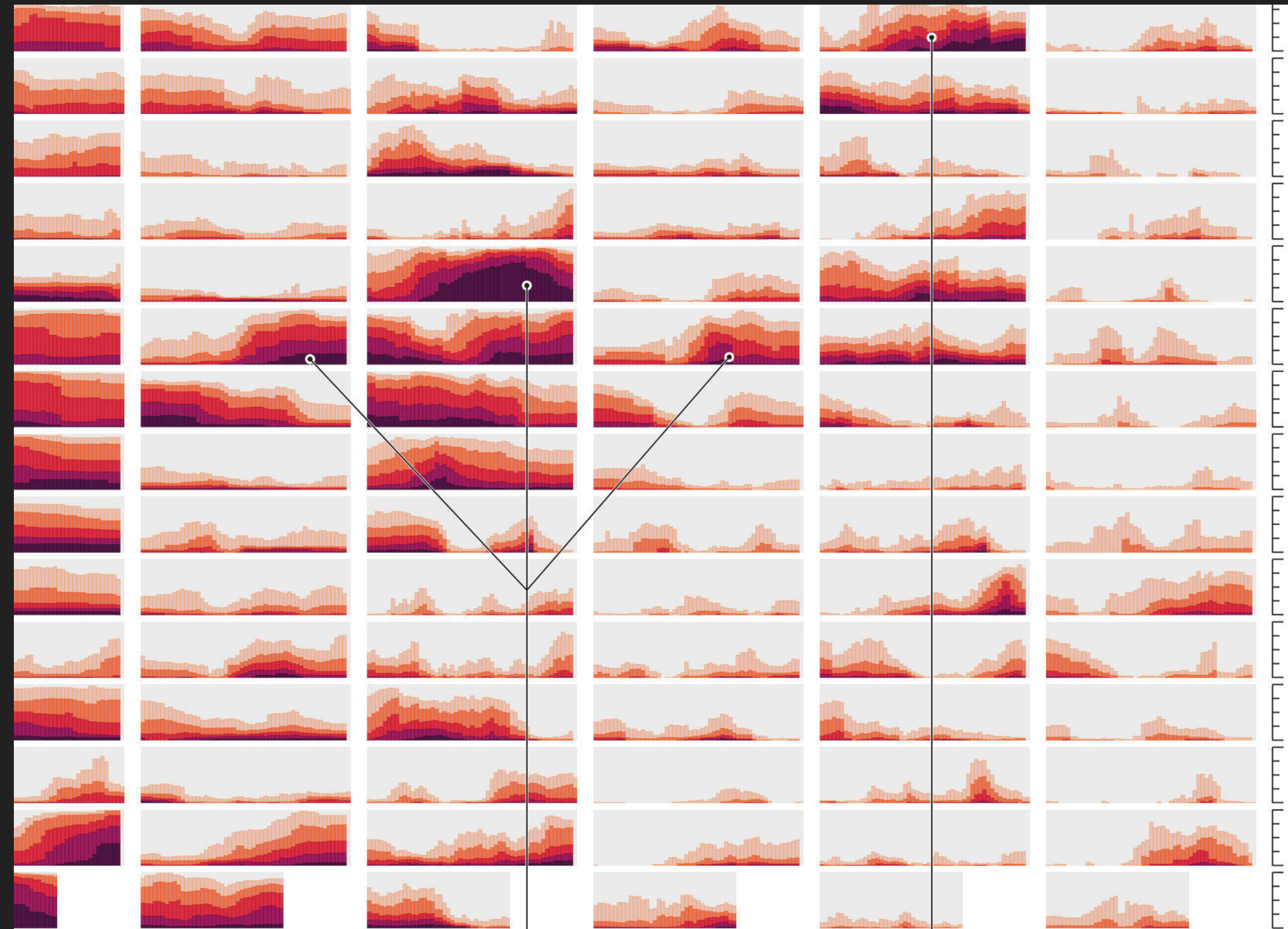


California experienced its hottest drought in recorded history from 2012 to 2016. A warming climate makes the atmosphere thirstier, which increases evaporation and boosts drought.

A drought that originated in the Southern Plains in 2011 eventually spread to the Midwest and Northern Plains when the moisture coming in from the Gulf of Mexico was absorbed by the parched South before it could reach the North.

The Southeast's driest year to date was 2007, when only 31.85 inches of rain fell in Atlanta, 62 percent of its average yearly rainfall.

Sources: U.S. Drought Monitor, jointly produced by the National Drought Mitigation Center at the University of Nebraska-Lincoln, U.S. Department of Agriculture, and National Oceanic and Atmospheric Administration (data)



A drought that originated in the Southern Plains in 2011 eventually spread to the Midwest and Northern Plains when the moisture coming in from the Gulf of Mexico was absorbed by the parched South before it could reach the North.

The Southeast's driest year to date was 2007, when only 31.85 inches of rain fell in Atlanta, 62 percent of its average yearly rainfall.

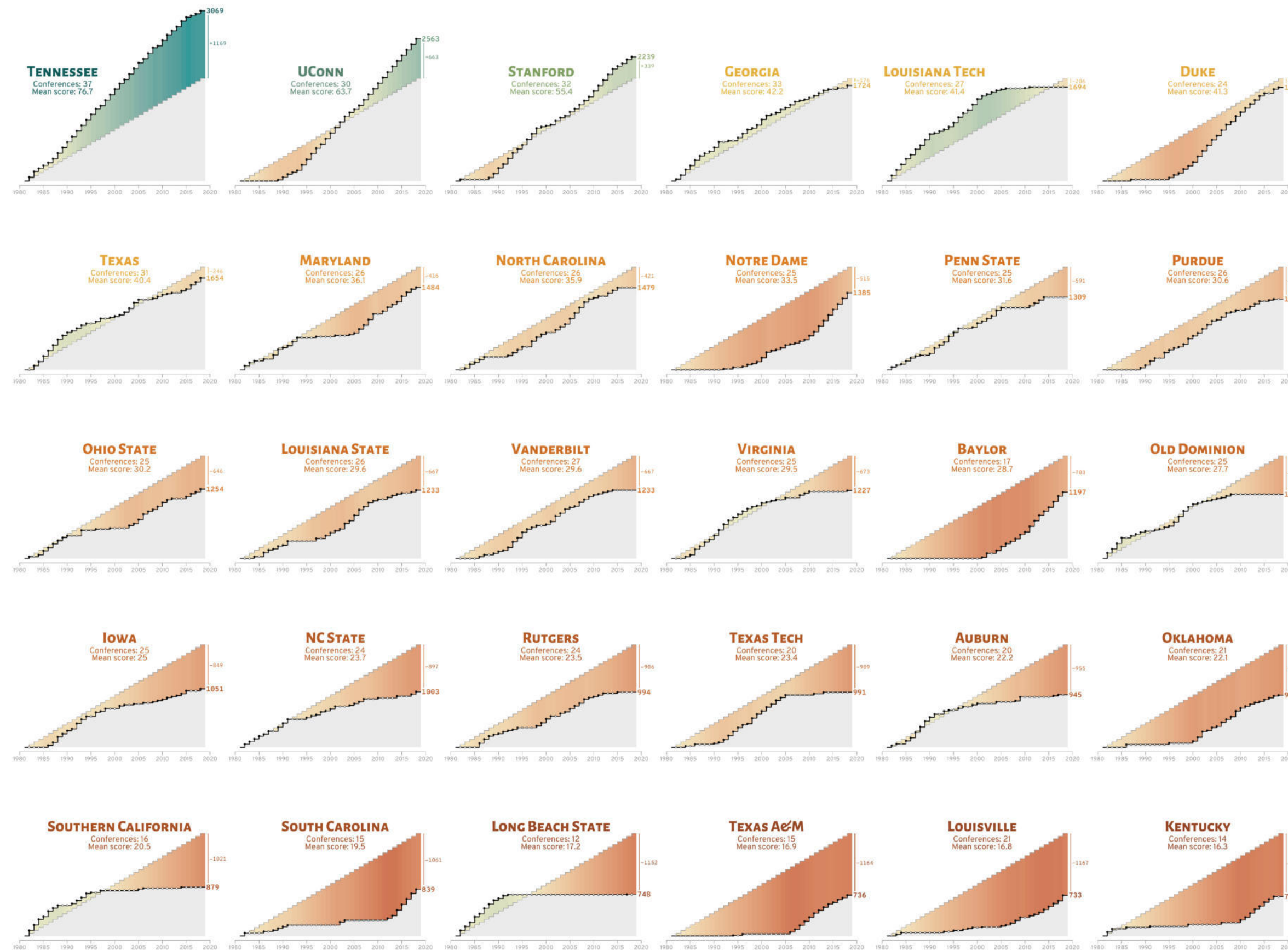
*“Escalating Drought”, together with Georgios Karamanis for Scientific American, Issue Nov 2021*



# — THE RISE & FALL OF WOMEN'S COLLEGE BASKETBALL DYNASTIES —

A number of teams that were the titans of the early NCAA women's basketball tournament have struggled in recent decades. And in their place, a new ruling class of schools has emerged to become the defining programs of the modern age. FiveThirtyEight estimated the team strength over time based on NCAA Tournament seeds as a proxy in the absence of game-level data. To measure this, FiveThirtyEight awarded "seed points" in proportion to a given seed number's expected wins in the tournament, calibrated to a 100-point scale where the No. 1 seed gets 100 points, No. 2 gets 70 points, and so forth.

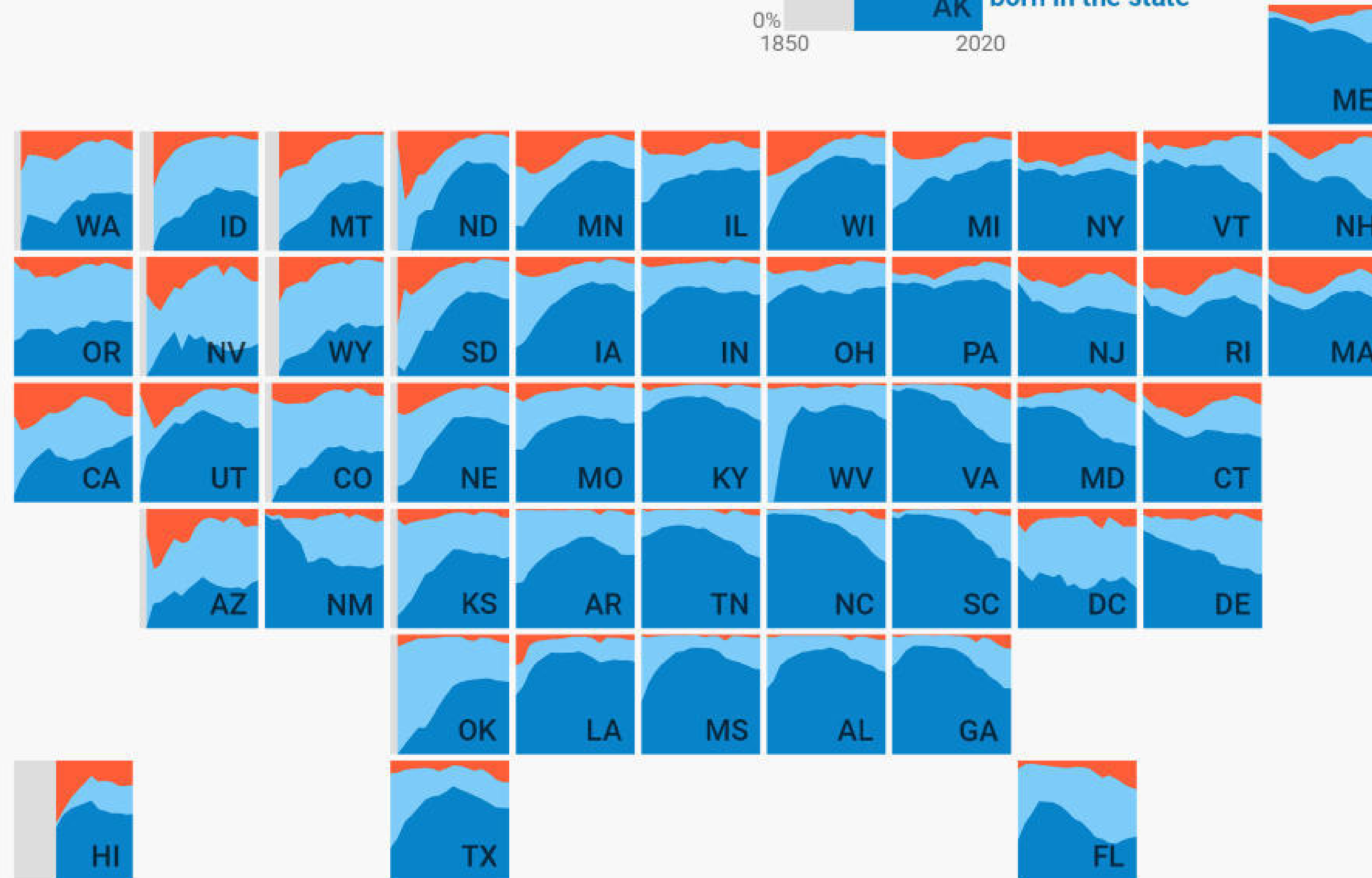
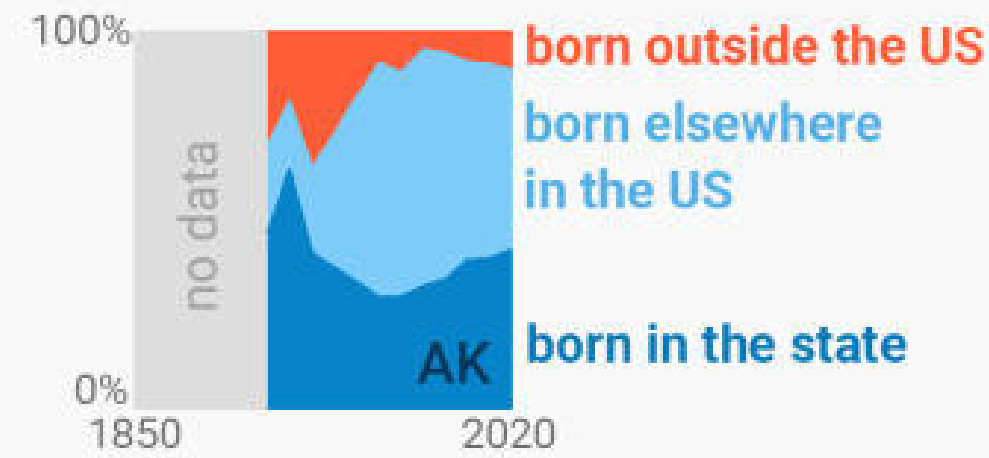
The visualization shows the cumulative sum of awarded seed points on a 100-point scale from the very first women's NCAA basketball tournaments in 1982 until 2018 in comparison to a hypothetical team that participated in all of the 37 conferences and gained half of the points each time (grey line). The curves highlight the fall of yesterday's women's basketball powerhouses such as **Louisiana Tech**, **Long Beach State**, **Southern California**, and **Old Dominion** that have been very good throughout the history of the women's tournament but have experienced big drop-offs in seed points over the last years. At the same time, schools such as **UConn**, **Stanford**, **Notre Dame**, **Baylor**, and **Duke** started slow but picked up steam into the present day. Some teams, such as **Tennessee**, have been relatively consistent throughout the NCAA era gathering always more seed points than an average team. *Shown are the top 30 college teams that participated in at least ten conferences between 1982 and 2018, sorted by the cumulative sum of seed points.*



VISUALIZATION BY CÉDRIC SCHERER - DATA BY FIVETHIRTYEIGHT

“The Rise and Fall of Women’s College Basketball Dynasties”, #TidyTuesday Contribution

# Where are Americans born?



1POINT21  
INTERACTIVE

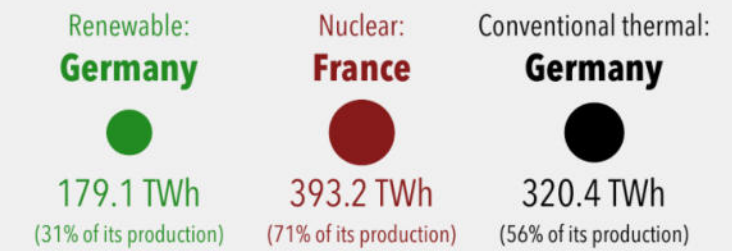
source: Steven Ruggles, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler and Matthew Sobek.  
IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. <https://doi.org/10.18128/D010.V11.0>

“Where are Americans born?” by @ErinDataViz

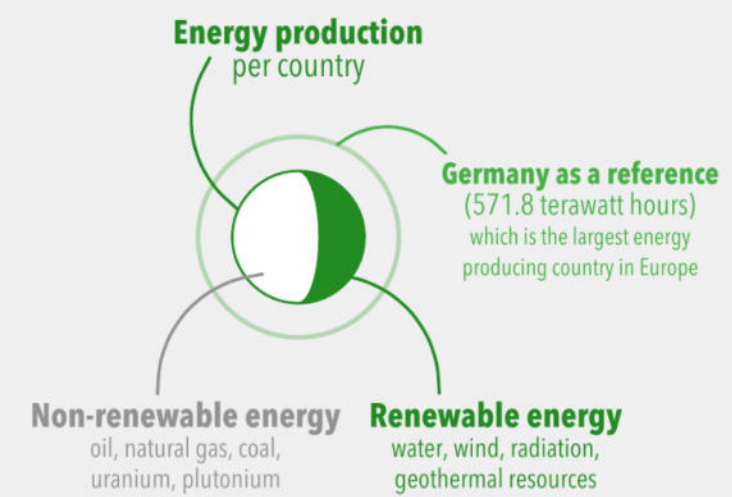


# How European countries generated electricity in 2018

Germany is the largest energy producing country in Europe. It generates the most renewable and conventional thermal energy, representing 31% and 56% of its overall production respectively. France is the second largest energy European producer and by far the largest nuclear energy provider: 71% of its production is based on nuclear fission to generate heat.

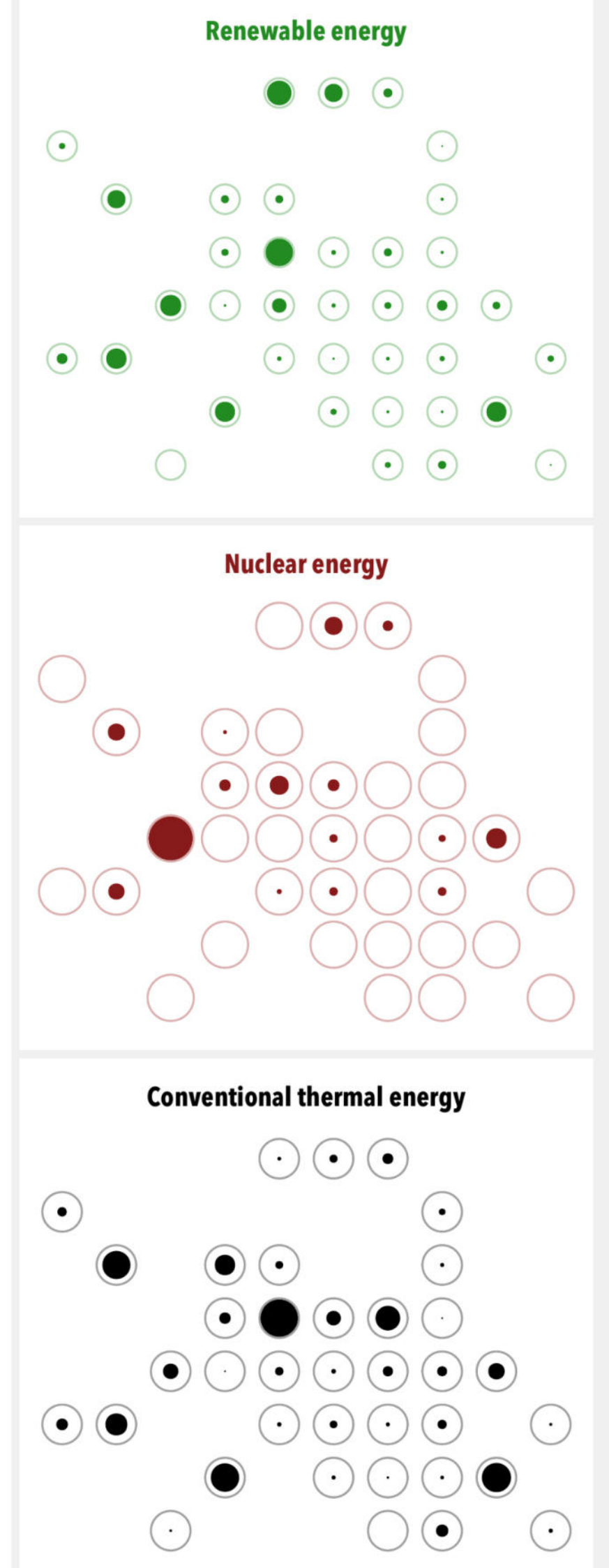
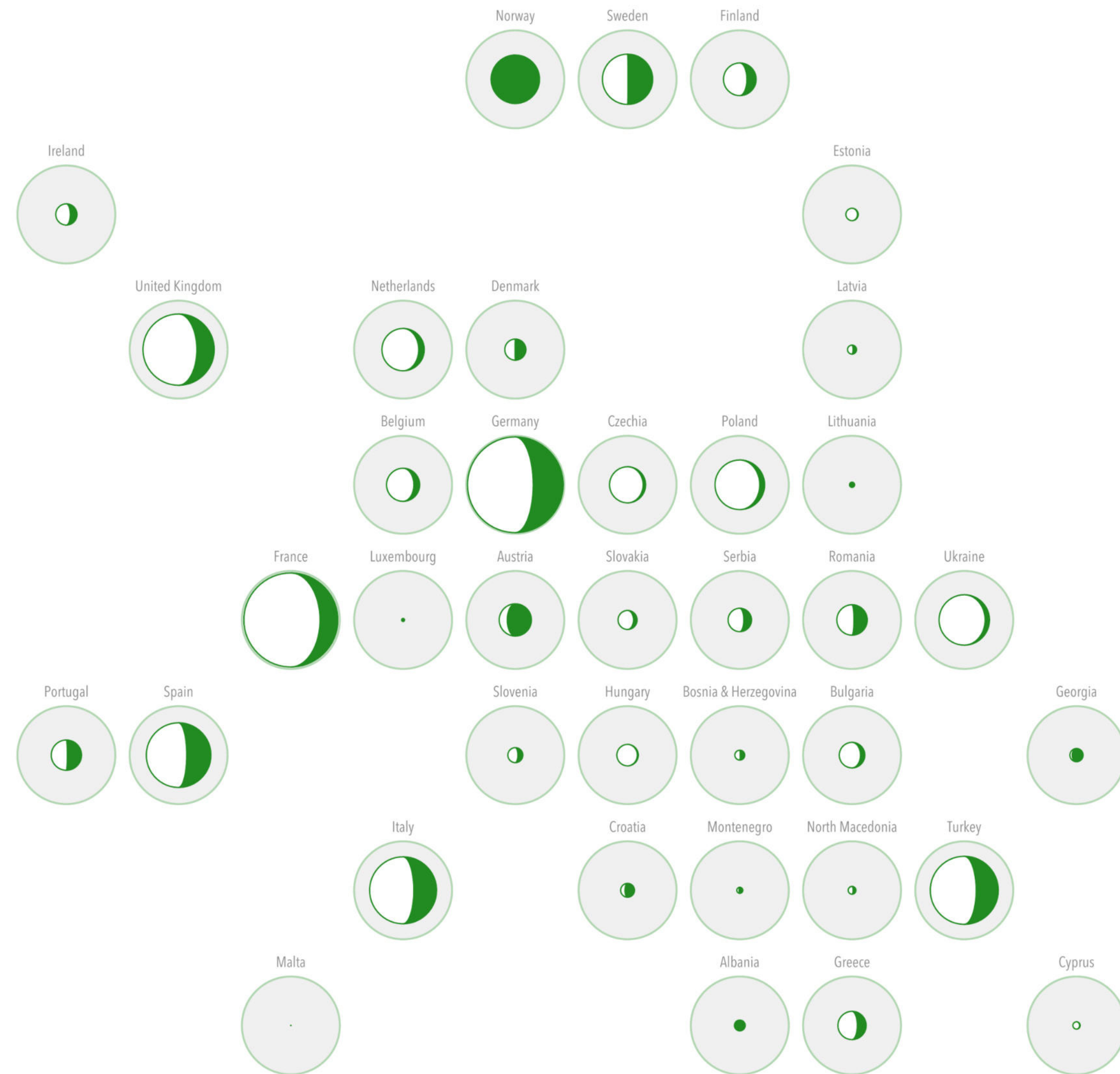


Renewable energy is energy that comes from resources that are naturally replenished such as sunlight, wind, water, and geothermal heat. Unlike fossil fuels, such as oil, natural gas and coal, or nuclear power sources such as uranium and plutonium, renewable energy regenerates naturally in a short period of time.



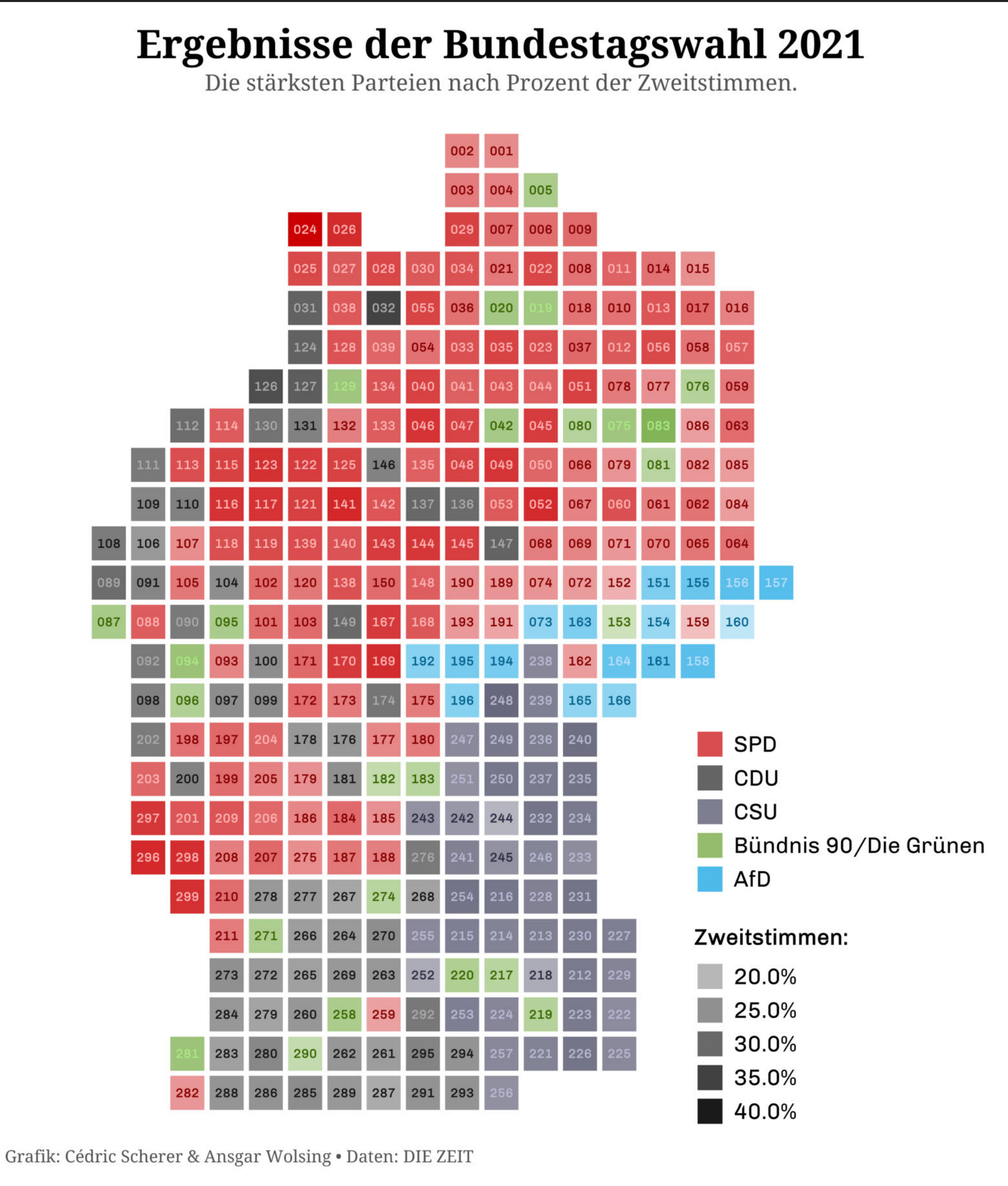
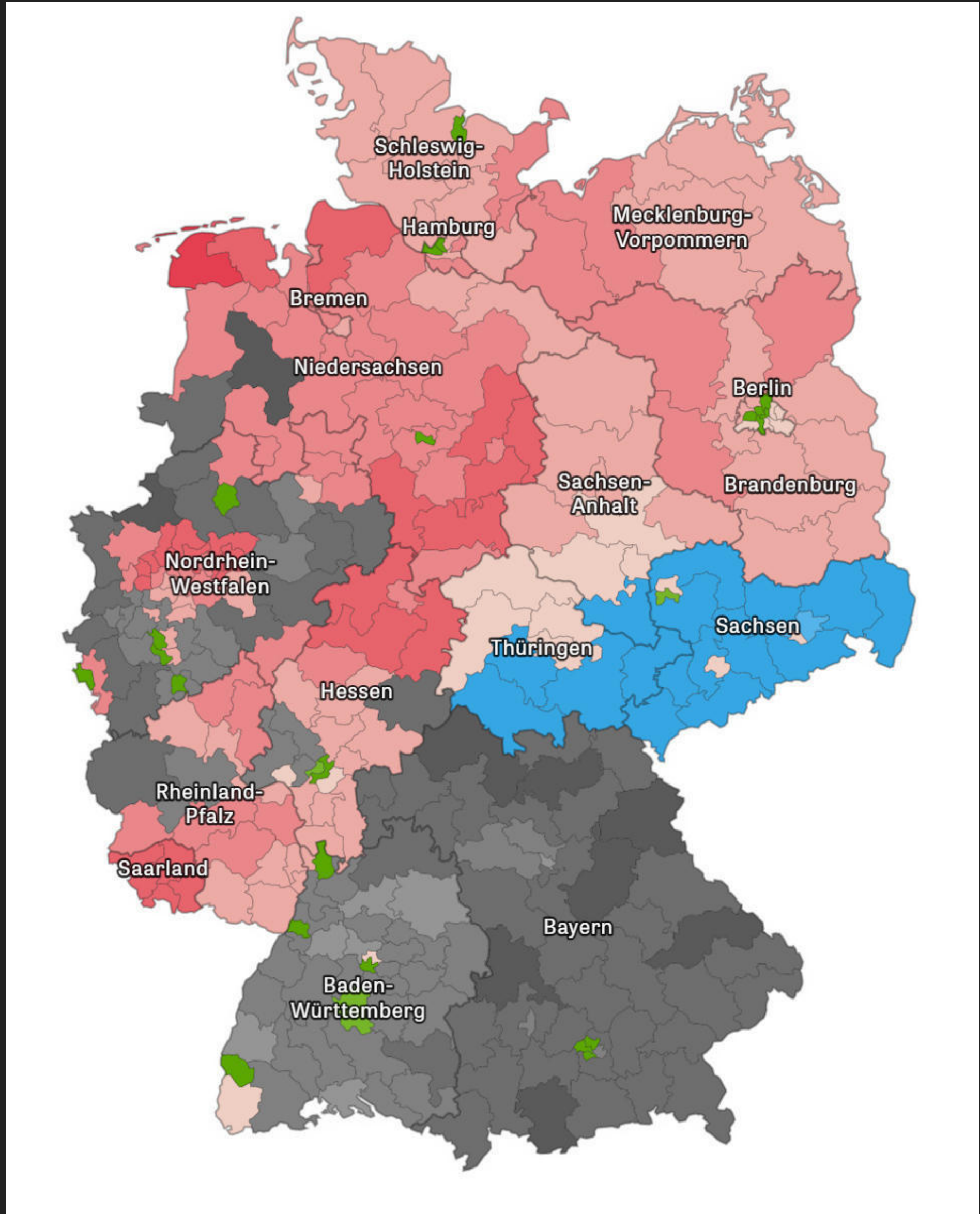
Norway had an electricity production almost entirely made up of renewable energy (98%). This makes Norway the second largest producer of this energy type in Europe. Interestingly, most of the renewable energy is produced by hydro power that take up 95% and only 3% by wind. In contrast, twelve European countries were reported to produce less than 20% of their energy with renewable resources: **Malta** (0%), **Hungary** (5%), **Estonia** (6%), **Czechia** (7%), **Cyprus** (9%), **Ukraine** (9%), **Poland** (10%), **Netherlands** (13%), **Bulgaria** (17%), **Belgium** (18%), **Slovakia** (19%), and **France** (19%).

Note: Energy production is mapped to the area of the circles.  
Visualization by Cédric Scherer • Data by Eurostat



“How European countries generated electricity in 2018”, #TidyTuesday Contribution





Grafik: Cédric Scherer & Ansgar Wolsing • Daten: DIE ZEIT

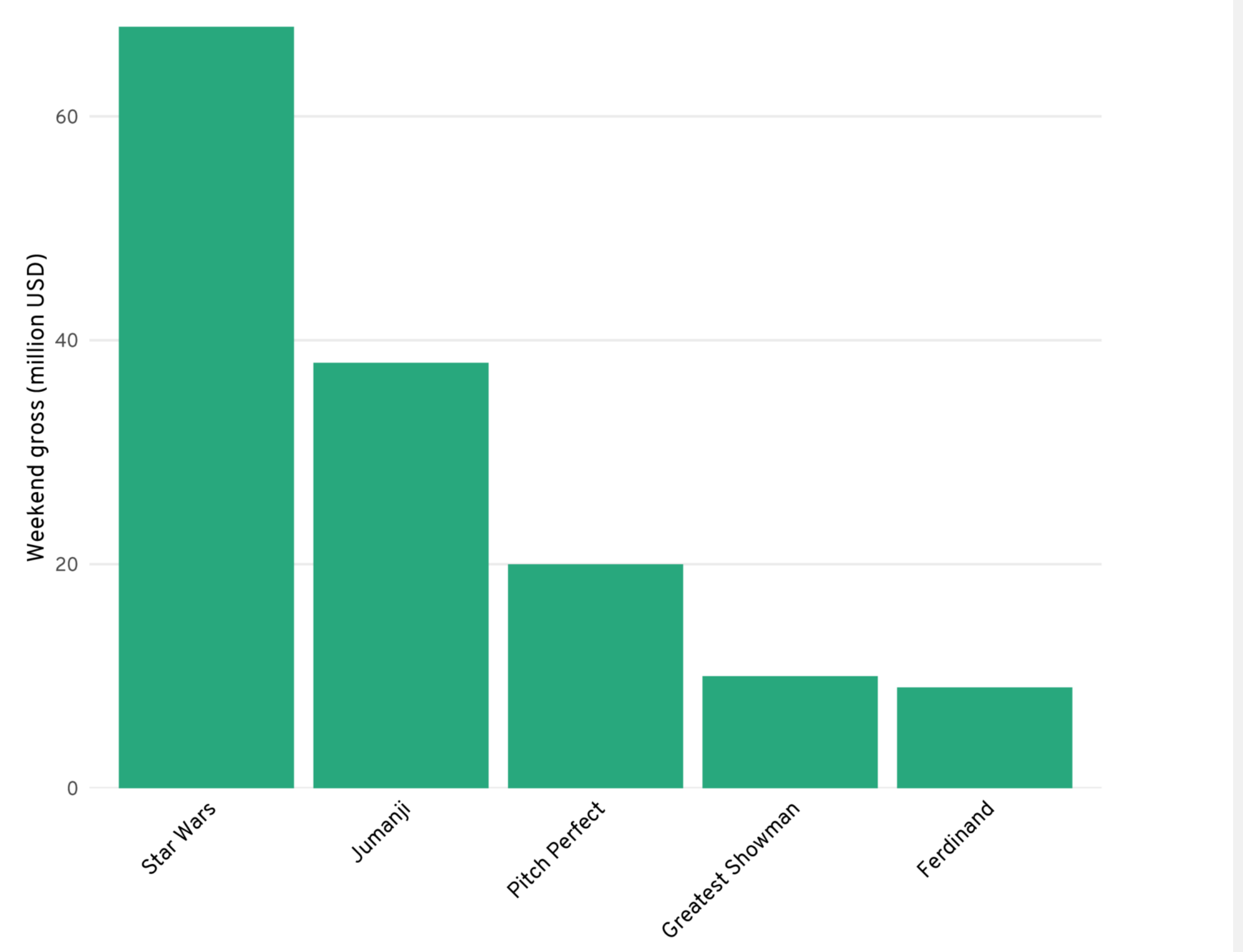
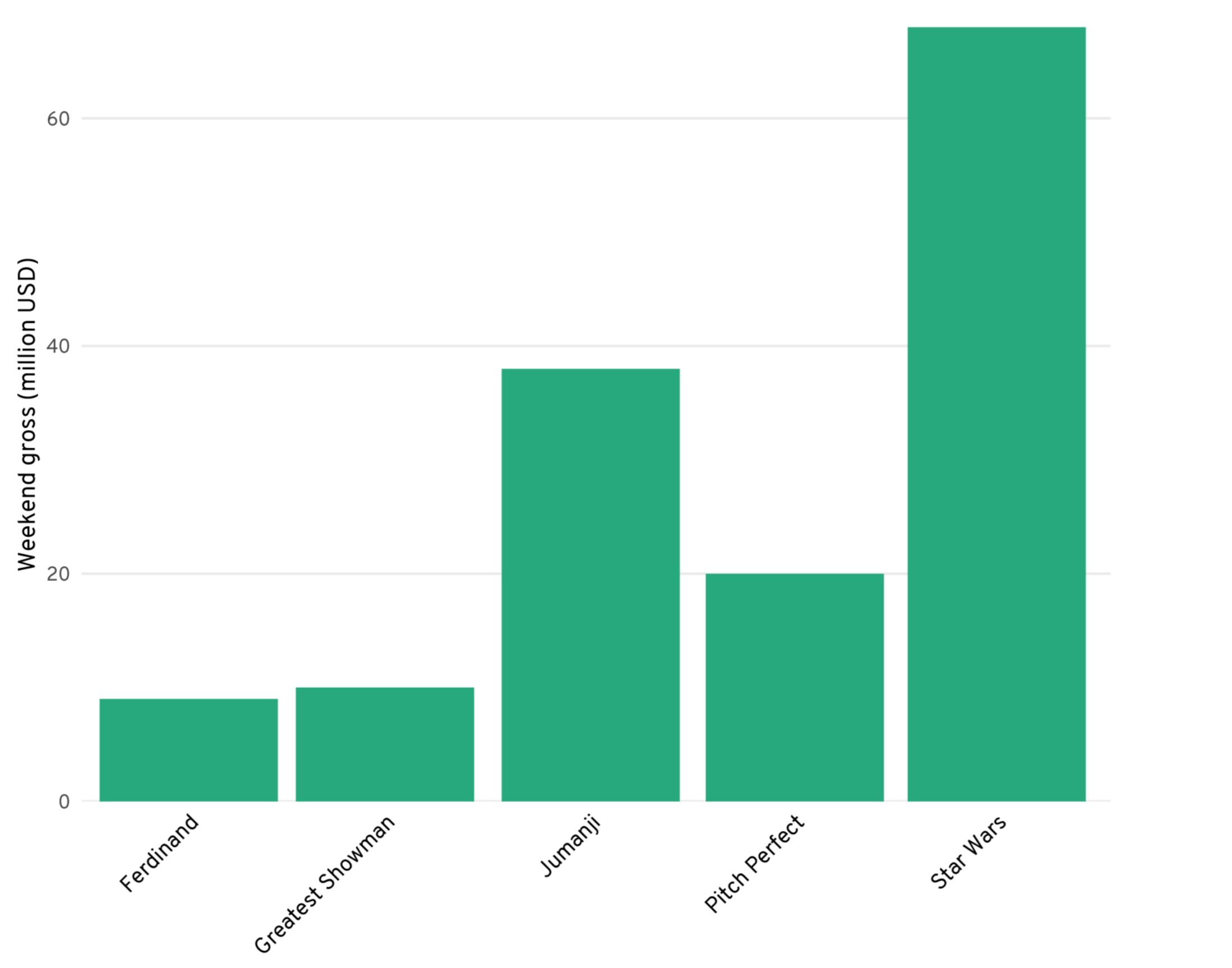
Left: Choropleth Map by Die Zeit | Right: Tile Grid Map by Cédric Scherer & Ansgar Wolsing

# VISUAL FORM

---

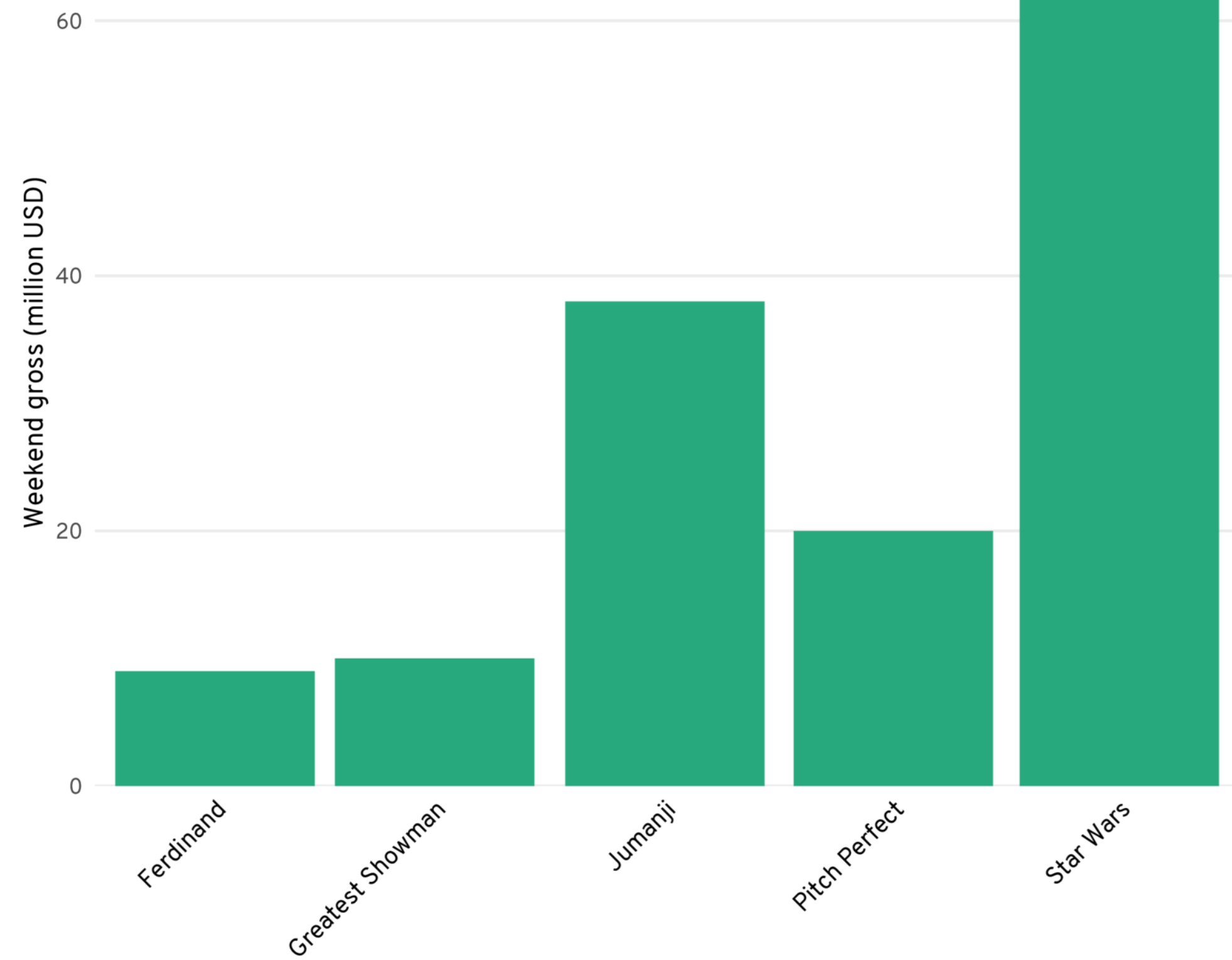
Follow design rules and data visualization principles

# Order your data

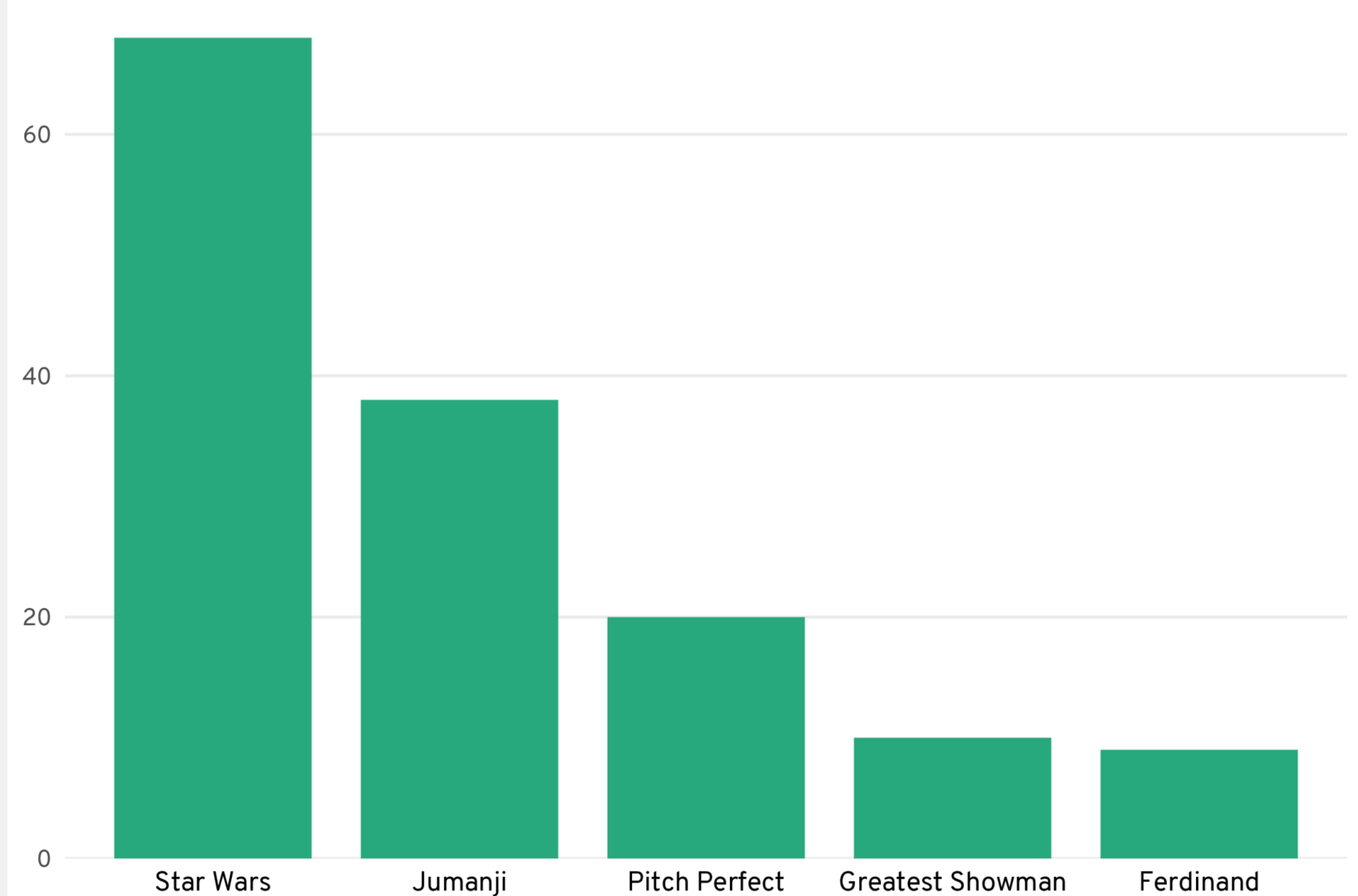




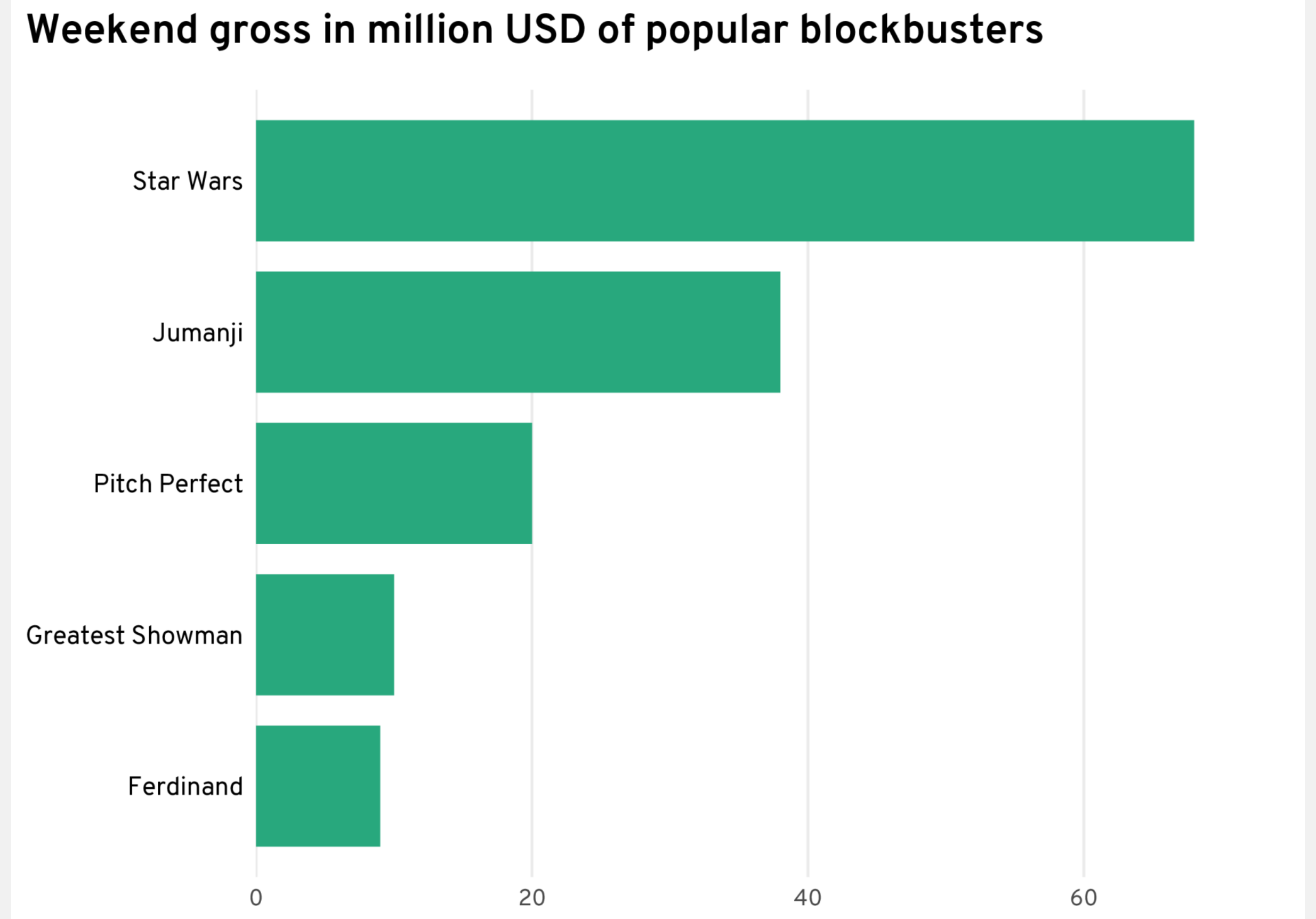
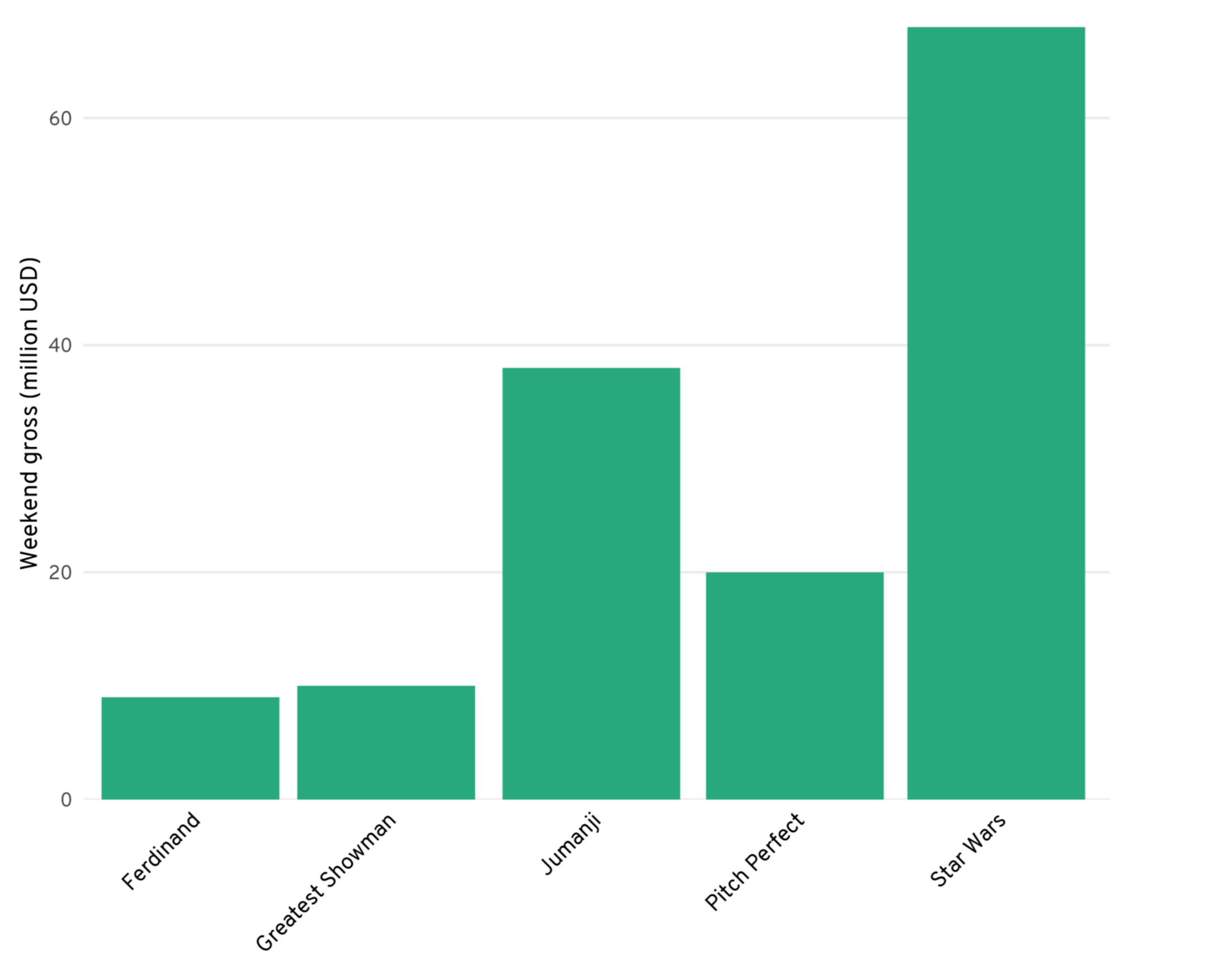
# Don't rotate your text



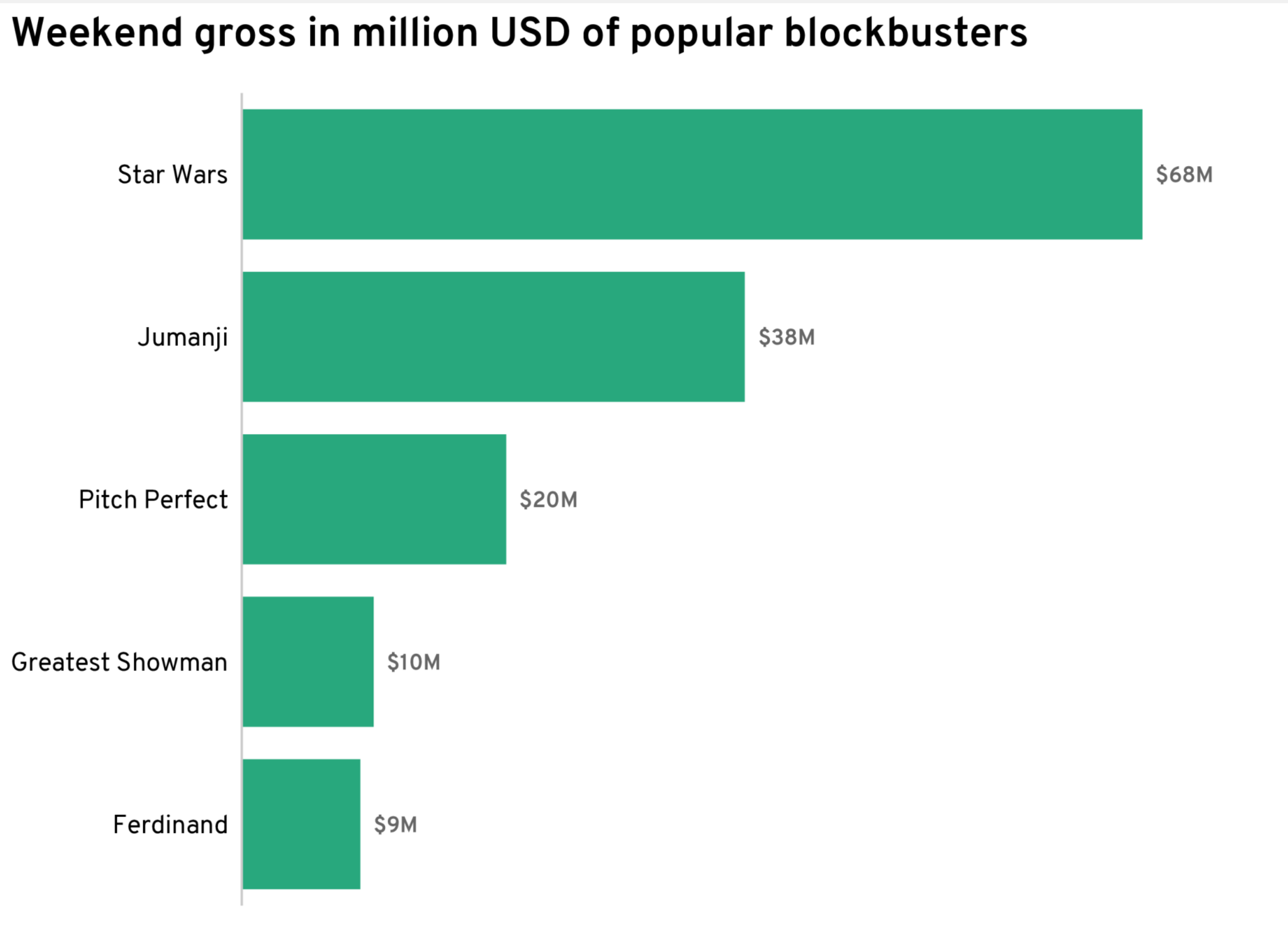
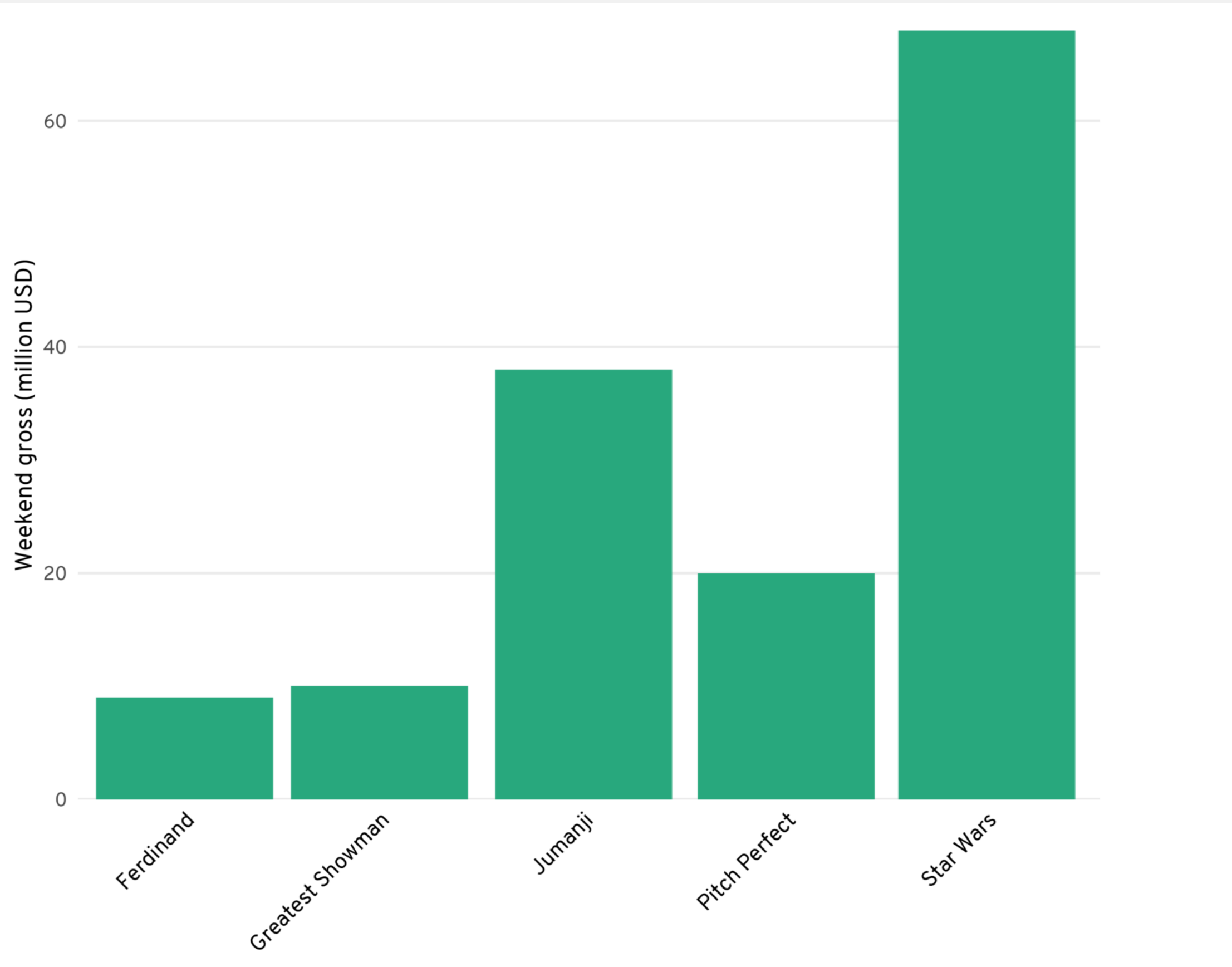
Weekend gross in million USD of popular blockbusters



# Don't rotate your text

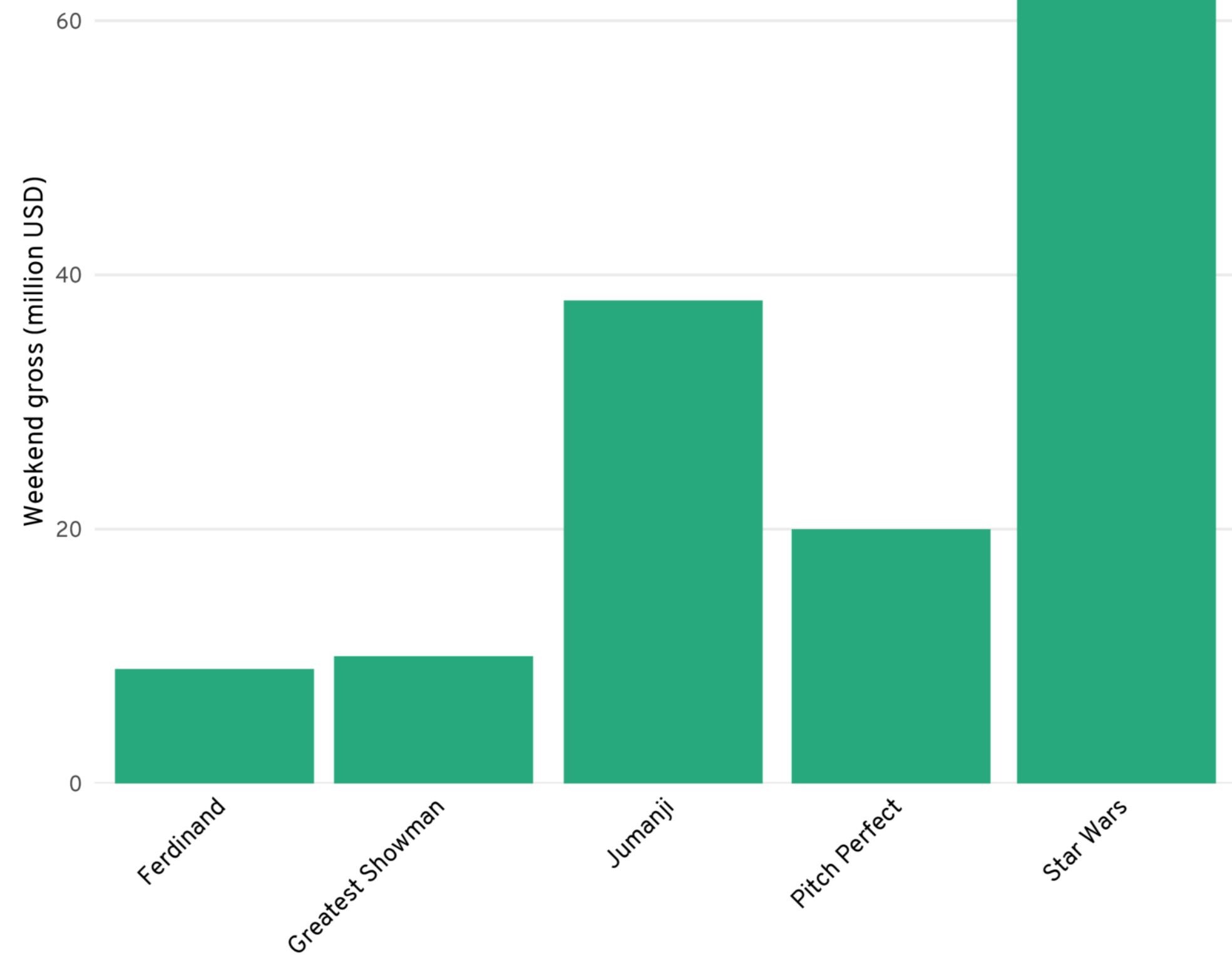


# Add direct labels

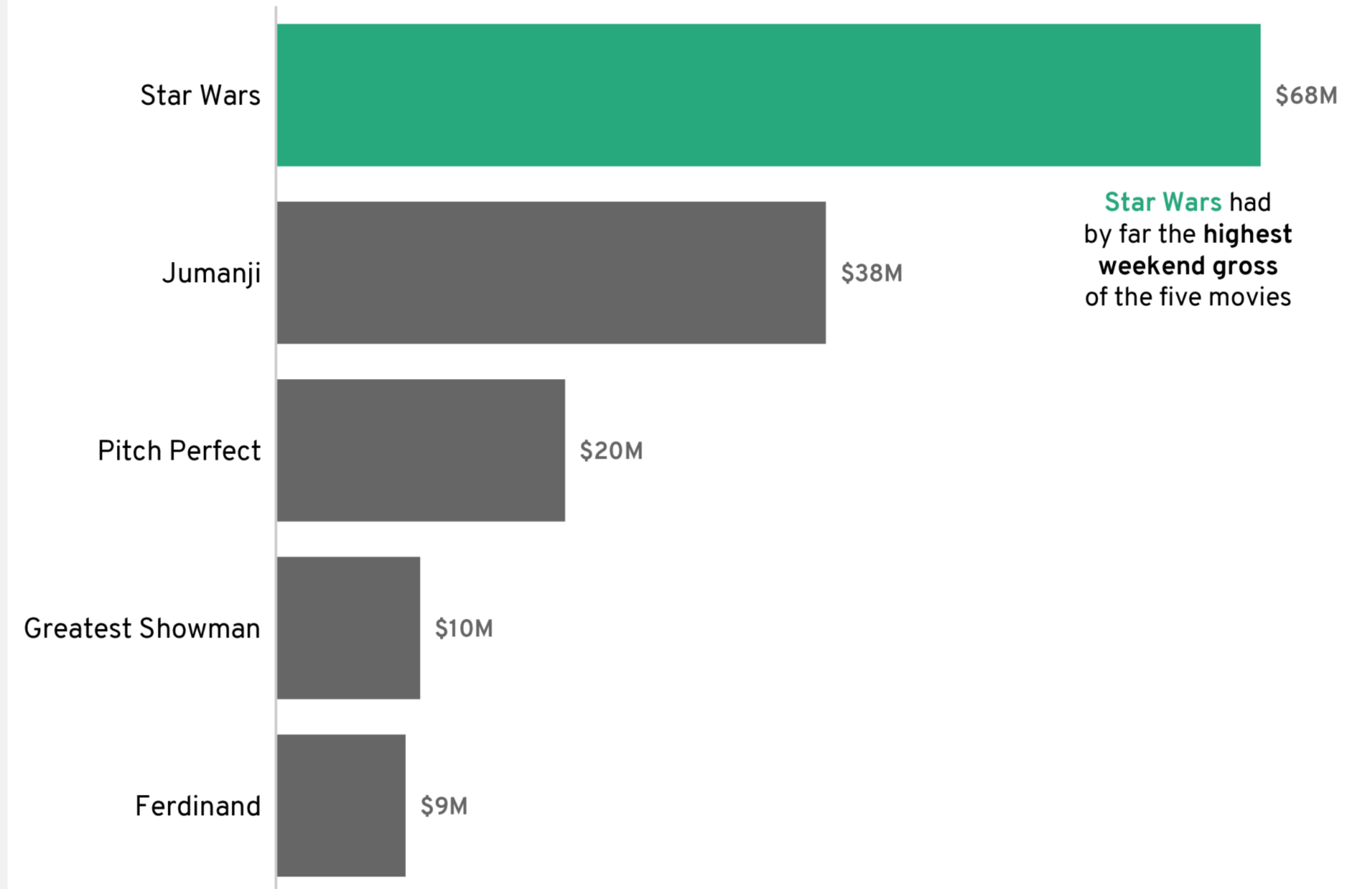




# Use colors + annotations wisely



Weekend gross in million USD of popular blockbusters

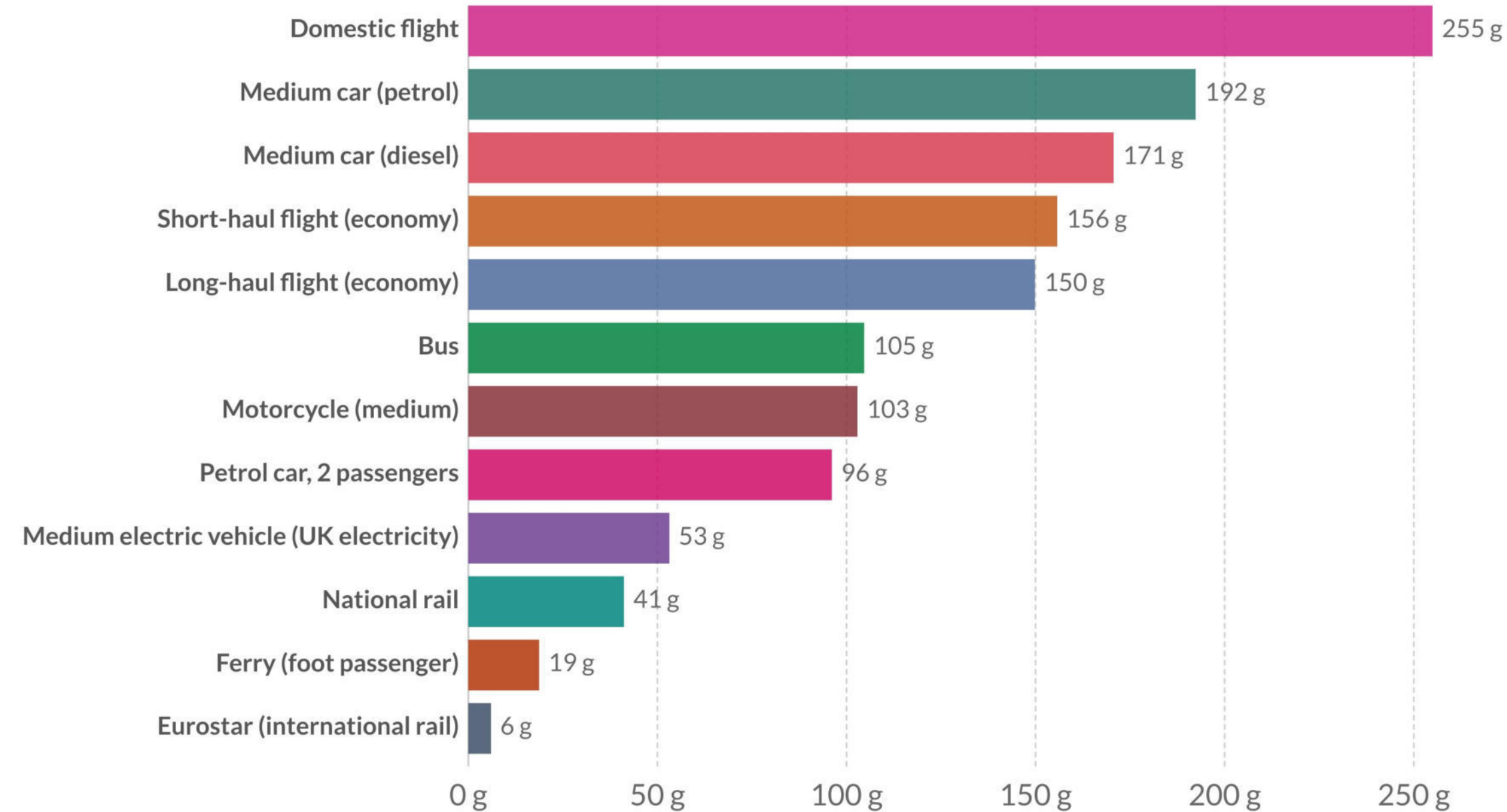


# Use color wisely (and correctly)

## Carbon footprint of travel per kilometer, 2018

The carbon footprint of travel is measured in grams of carbon dioxide equivalents per passenger kilometer. This includes carbon dioxide, but also other greenhouse gases, and increased warming from aviation emissions at altitude.

Our World in Data



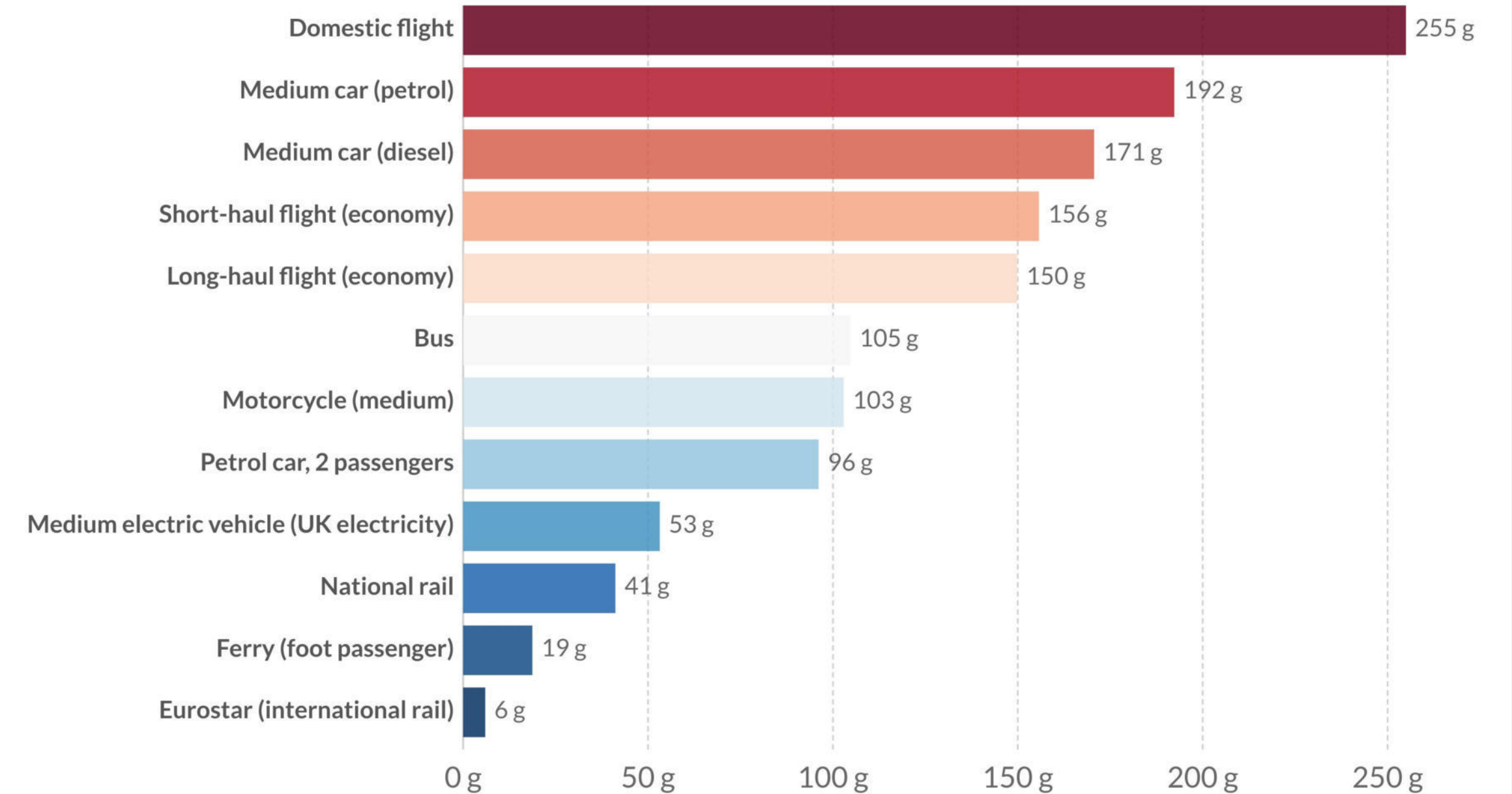
Source: UK Department for Business, Energy & Industrial Strategy. Greenhouse gas reporting: conversion factors 2019.  
Note: Data is based on official conversion factors used in UK reporting. These factors may vary slightly depending on the country.

CC BY

## Carbon footprint of travel per kilometer, 2018

The carbon footprint of travel is measured in grams of carbon dioxide equivalents per passenger kilometer. This includes carbon dioxide, but also other greenhouse gases, and increased warming from aviation emissions at altitude.

Our World in Data



Source: UK Department for Business, Energy & Industrial Strategy. Greenhouse gas reporting: conversion factors 2019.  
Note: Data is based on official conversion factors used in UK reporting. These factors may vary slightly depending on the country.

CC BY

Original graphic with a random categorical palette

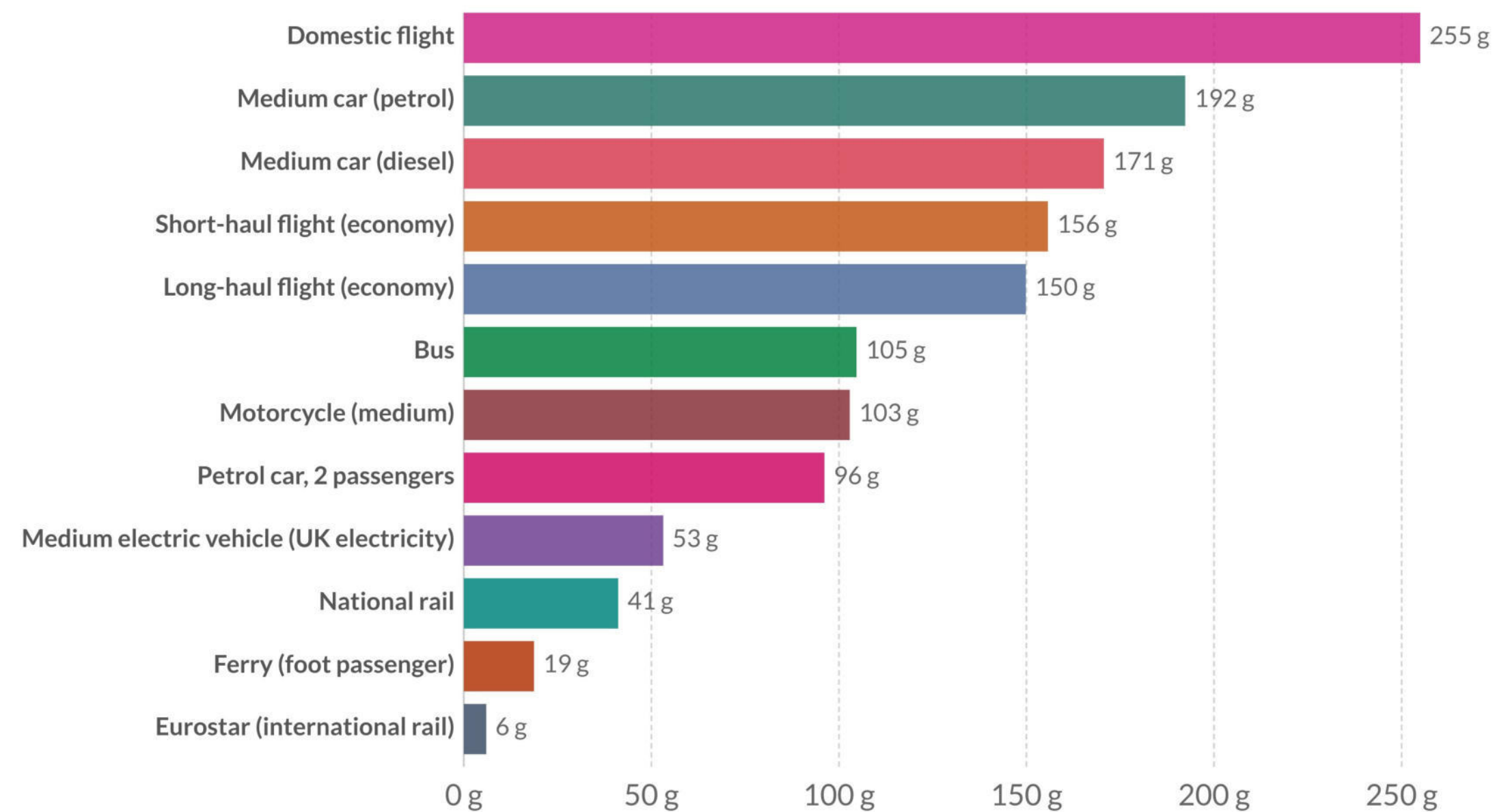
Reworked graphic using a diverging palette

# Use color wisely (and correctly)

## Carbon footprint of travel per kilometer, 2018

The carbon footprint of travel is measured in grams of carbon dioxide equivalents per passenger kilometer. This includes carbon dioxide, but also other greenhouse gases, and increased warming from aviation emissions at altitude.

Our World  
in Data

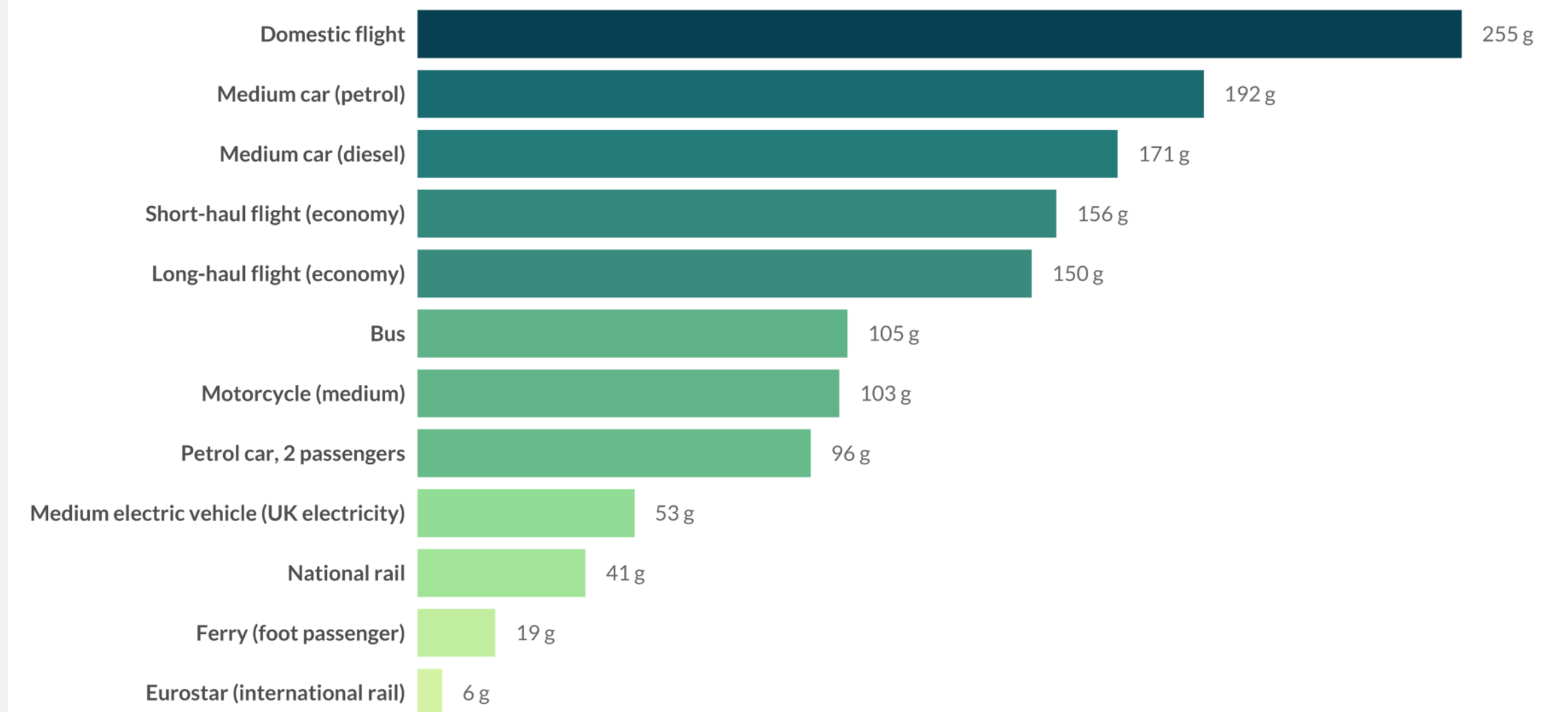


Source: UK Department for Business, Energy & Industrial Strategy. Greenhouse gas reporting: conversion factors 2019.  
Note: Data is based on official conversion factors used in UK reporting. These factors may vary slightly depending on the country.

CC BY

## Carbon footprint of travel per kilometer, 2018

The carbon footprint of travel is measured in grams of carbon dioxide equivalents per passenger kilometer. This includes carbon dioxide, but also other greenhouse gases, and increased warming from aviation emissions at altitude.



Source: UK Department for Business, Energy & Industrial Greenhouse gas reporting: conversion factors 2019.  
Note: Data is based on official conversion factors used in UK reporting. These factors may vary slightly depending on the country.

Original visualization by Hannah Ritchie, OurWorldInData.org | Makeover by Cédric Scherer

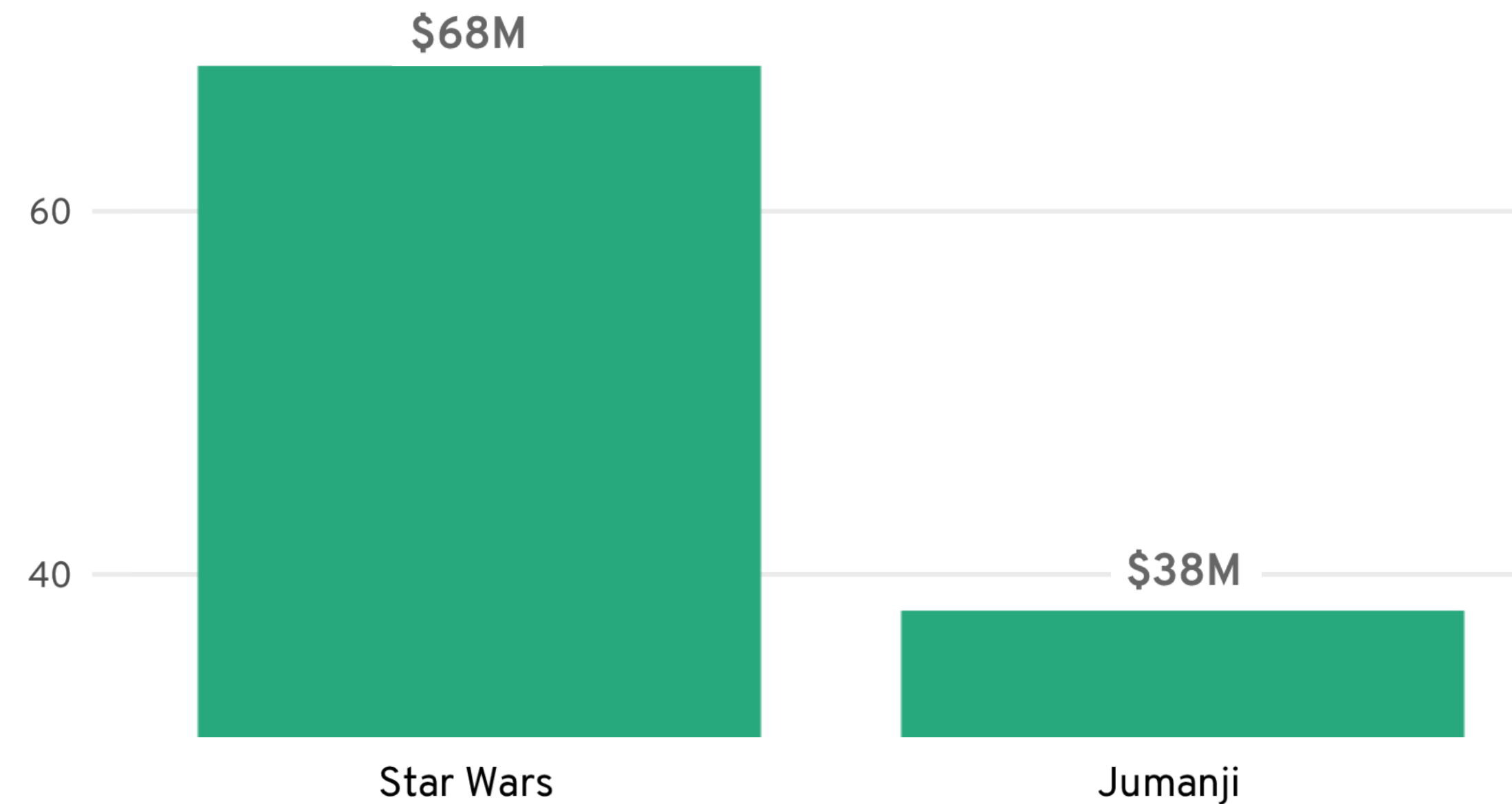
*Original graphic with a random categorical palette*

*Makeover using a continuous palette*

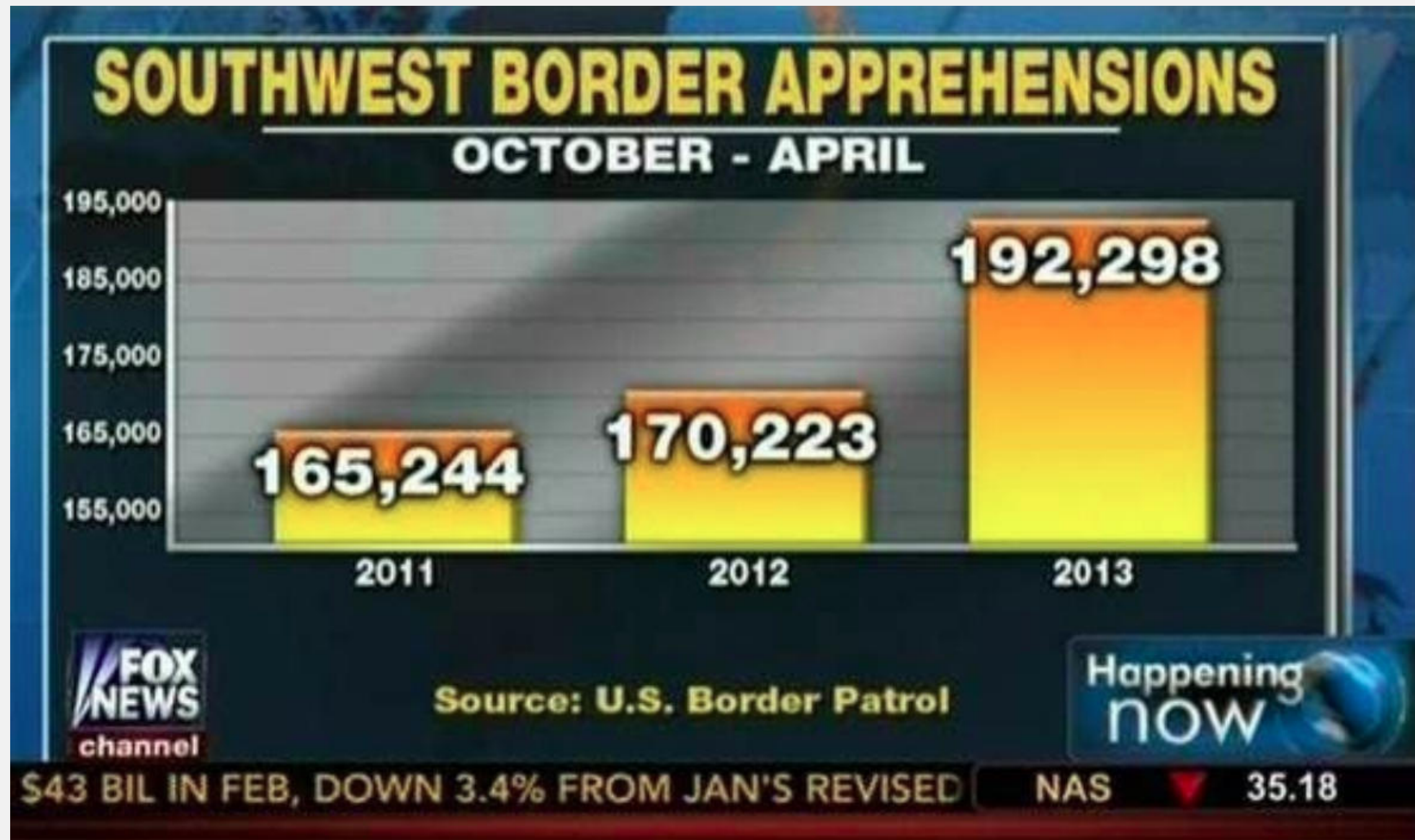


# Always start at zero

Weekend gross in million USD of popular blockbusters



# Always start at zero

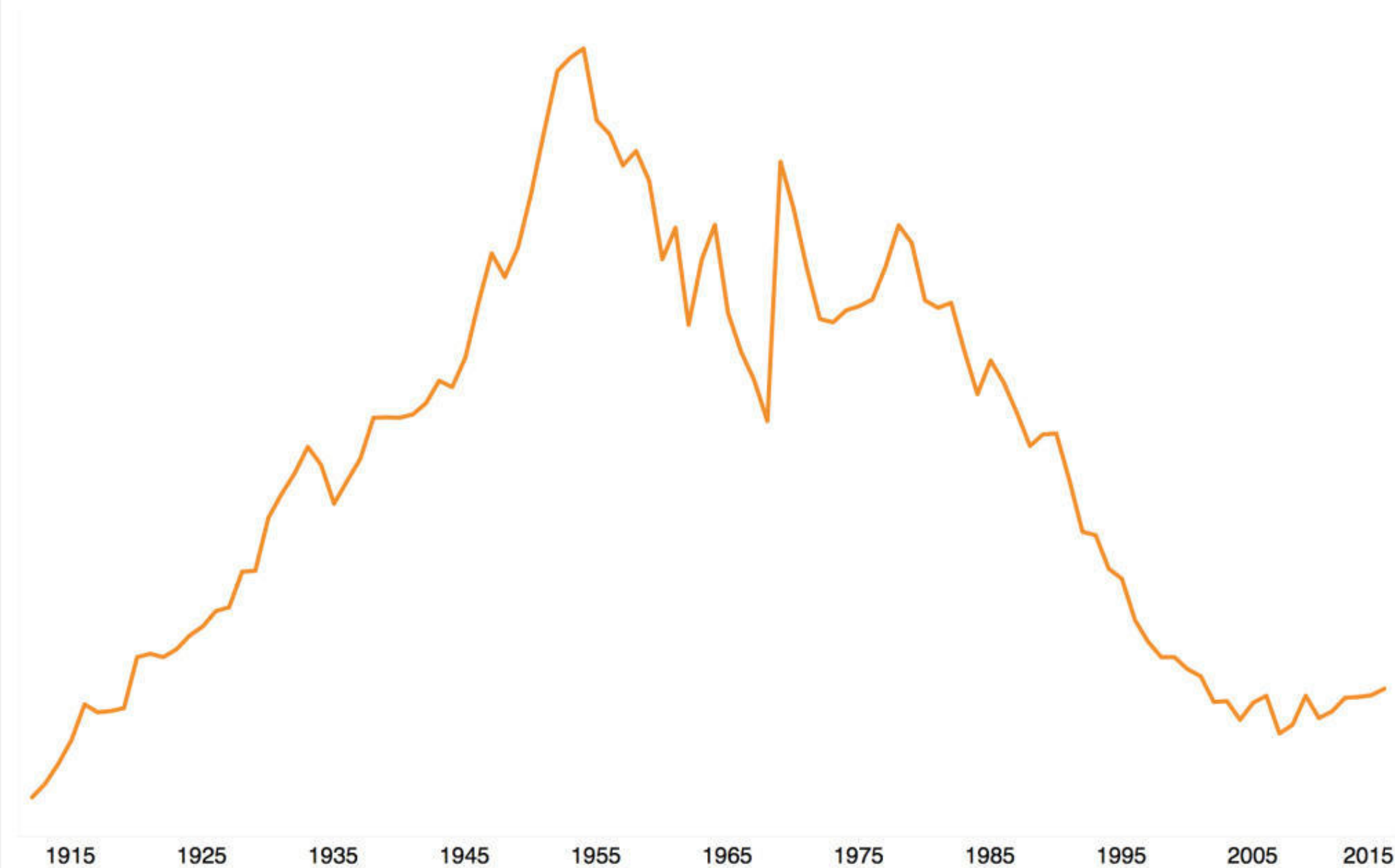




# The Power of Annotations

Rise and Fall of the name **Neil** in the USA  
Births 1912-2015

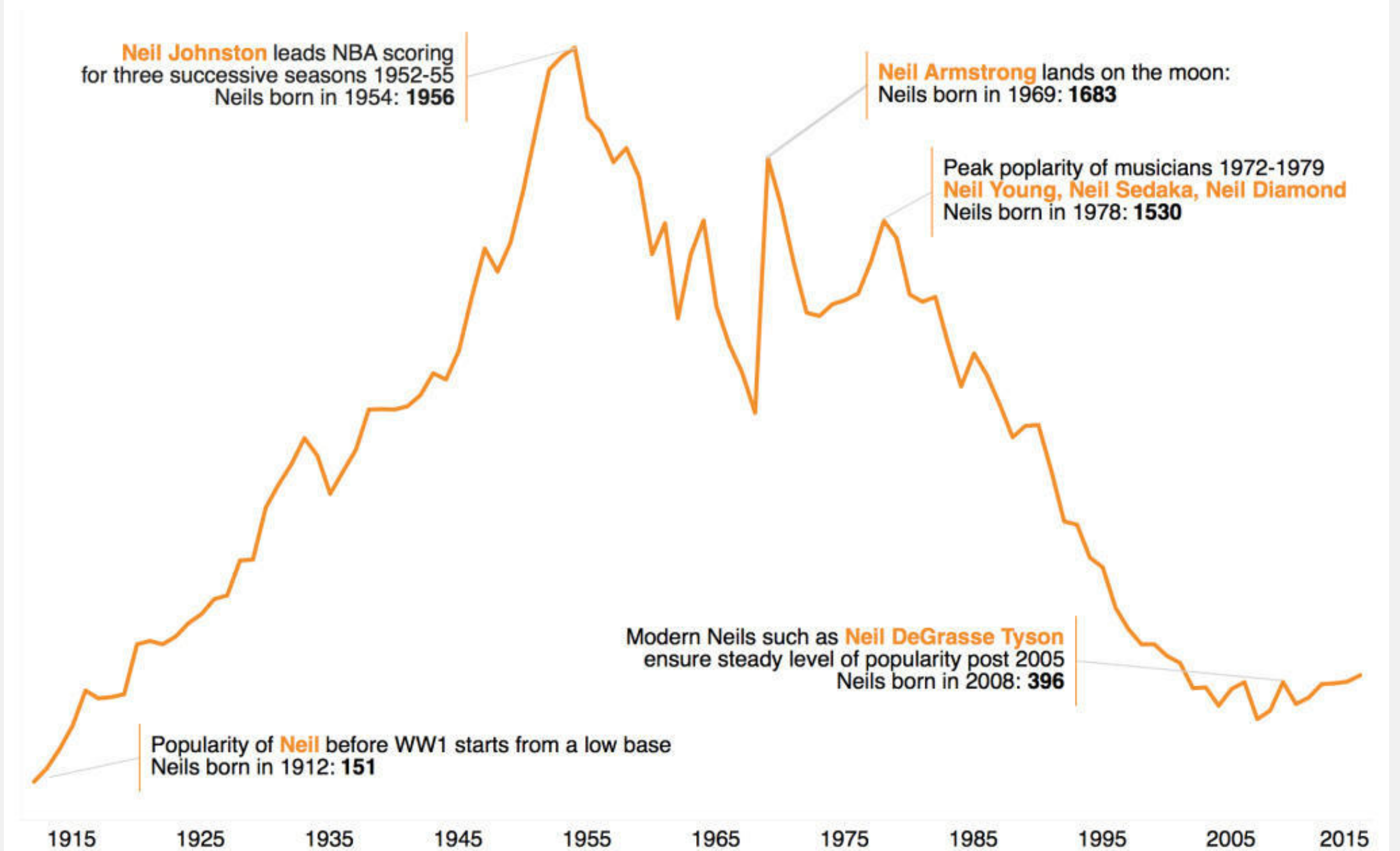
Source: data.gov



Visualisation: @theneilrichards

Rise and Fall of the name **Neil** in the USA  
Births 1912-2015

Source: data.gov



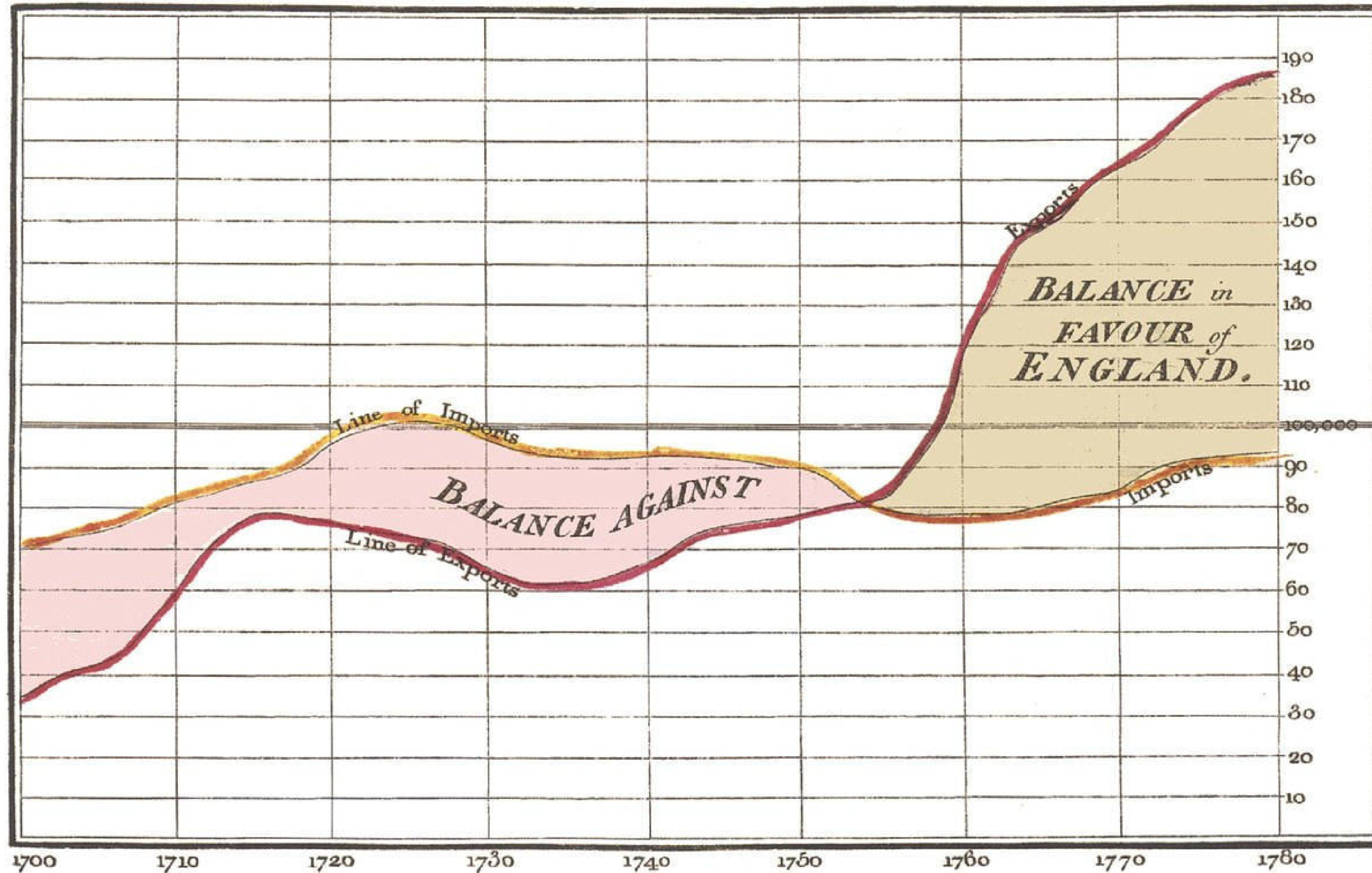
Visualisation: @theneilrichards

#SWDChallenge

*“Is white space always your friend?” by Neil Richards*



Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.



Annotated time-series chart by William Playfair from "The Commercial and Political Atlas and Statistical Breviary" (1786)



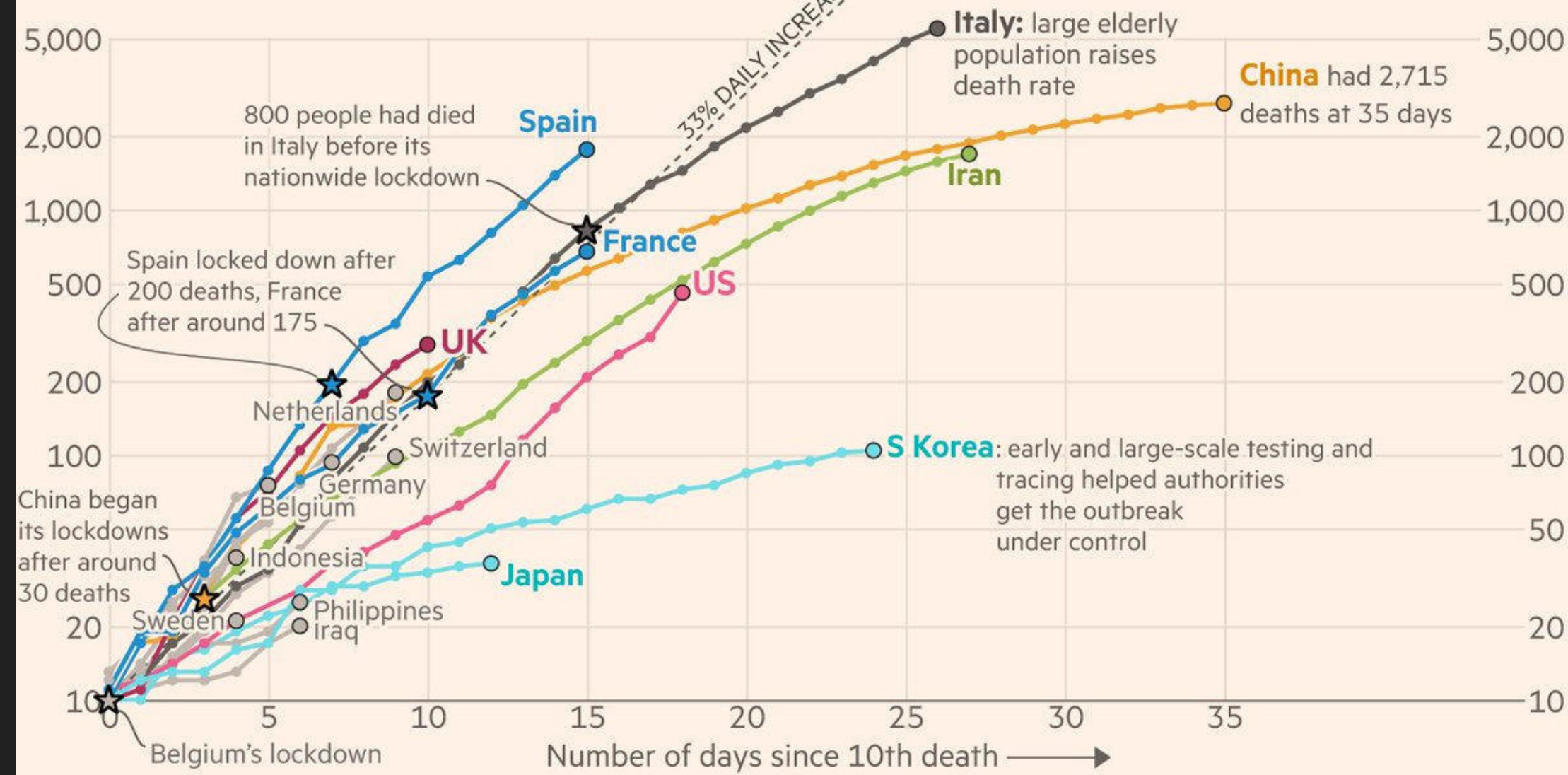
“The key thing we do is to add a title to the chart, as an entry point and to explain what is going on. **Text and other annotations add enormous value for non-chart people.**”

~ John Burn-Murdoch, Financial Times

Coronavirus deaths in Italy, Spain and the UK are increasing much more rapidly than they did in China

Cumulative number of deaths, by number of days since 10th death

Nationwide lockdowns: ★



FT graphic: John Burn-Murdoch / @jburnmurdoch  
 Source: FT analysis of Johns Hopkins University, CSSE; Worldometers; FT research. Data updated March 23, 09:00 GMT  
 © FT

Covid has grown gradually less lethal over the pandemic, mainly due to immunity, but it remains more dangerous than flu on average

Evolution of Covid-19's infection fatality ratio\* in England, relative to seasonal flu

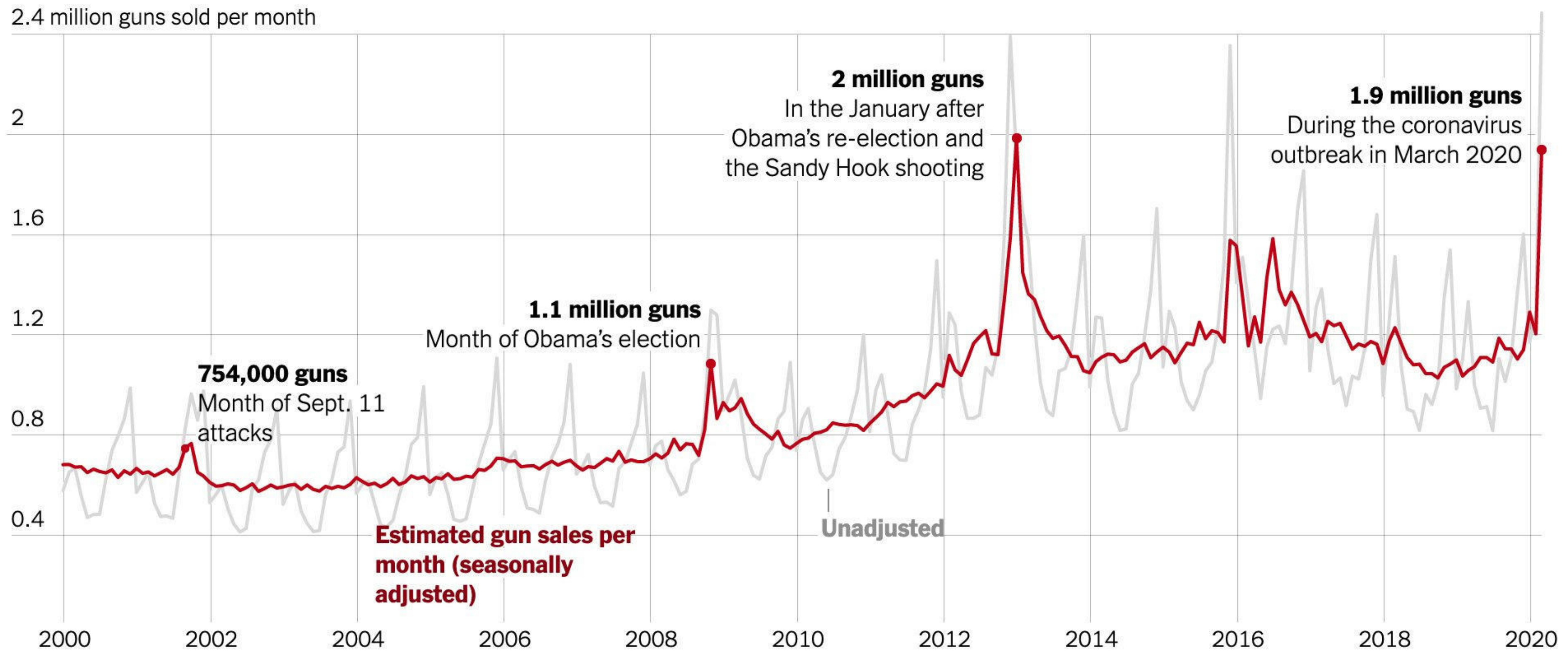


\*Covid IFR calculated using ONS death cert. mentions and ONS infection survey. \*\*IFR for seasonal flu as calculated for New Zealand in BMJ  
 Source: ONS. Based on prior work by Dan Howdon FT graphic: John Burn-Murdoch / @jburnmurdoch  
 © FT



# About 2 Million Guns Were Sold in the U.S. as Virus Fears Spread

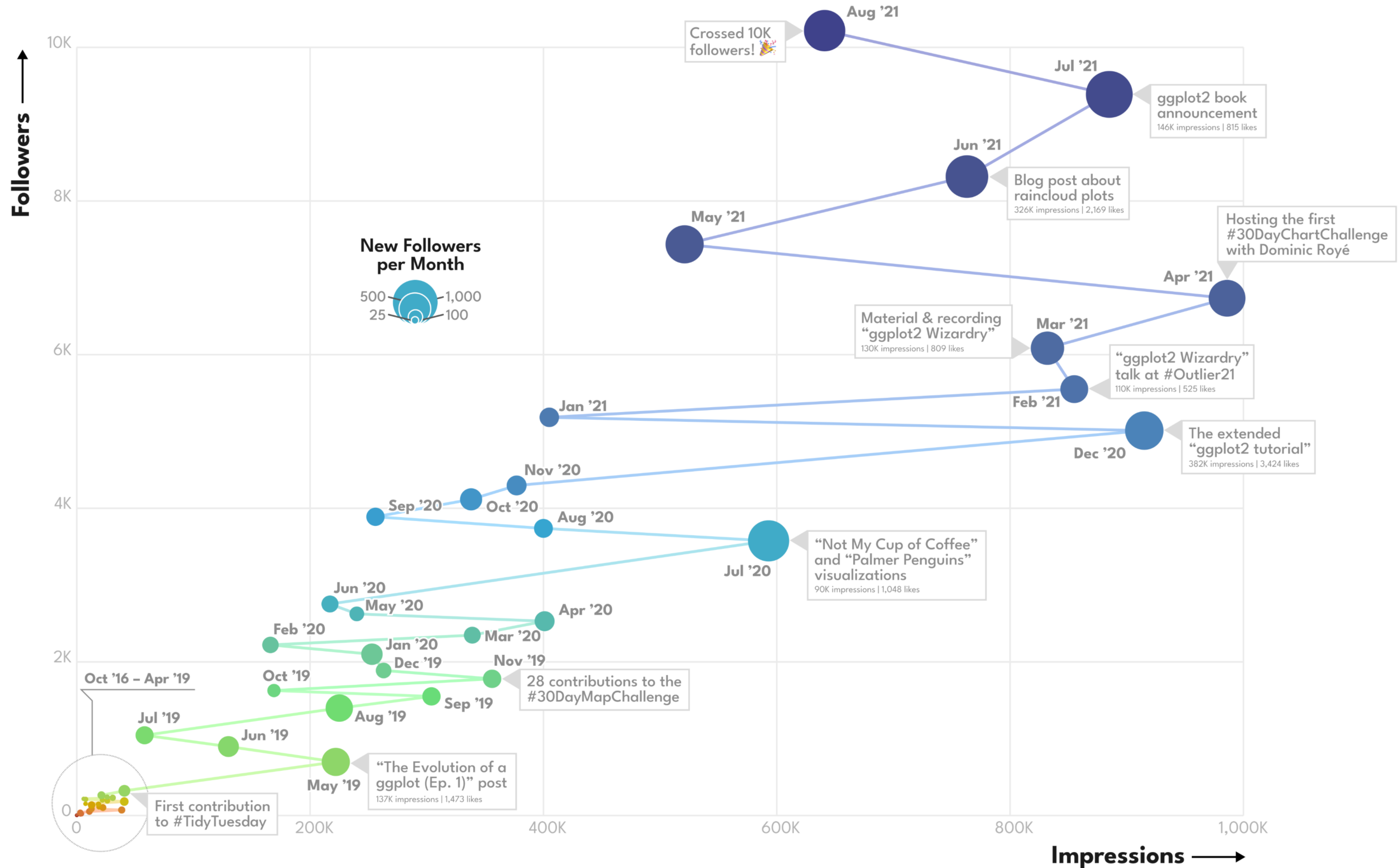
By Keith Collins and David Yaffe-Bellany April 1, 2020



*"About 2 Million Guns Were Sold in the U.S. as Virus Feats Spread" by Keith Collins and David Yaffe-Bellany (New York Times)*



# My Road to 10K on Twitter – Thank You All for Following!



Graphic: Cédric Scherer • Data: Twitter Analytics for @CédScherer

# WRAP UP

---

# Information .....

Understand your data and be accurate.

# Story .....

Be clear about the message of your visualization.

# Goal .....

Select charts that successfully transport your story.

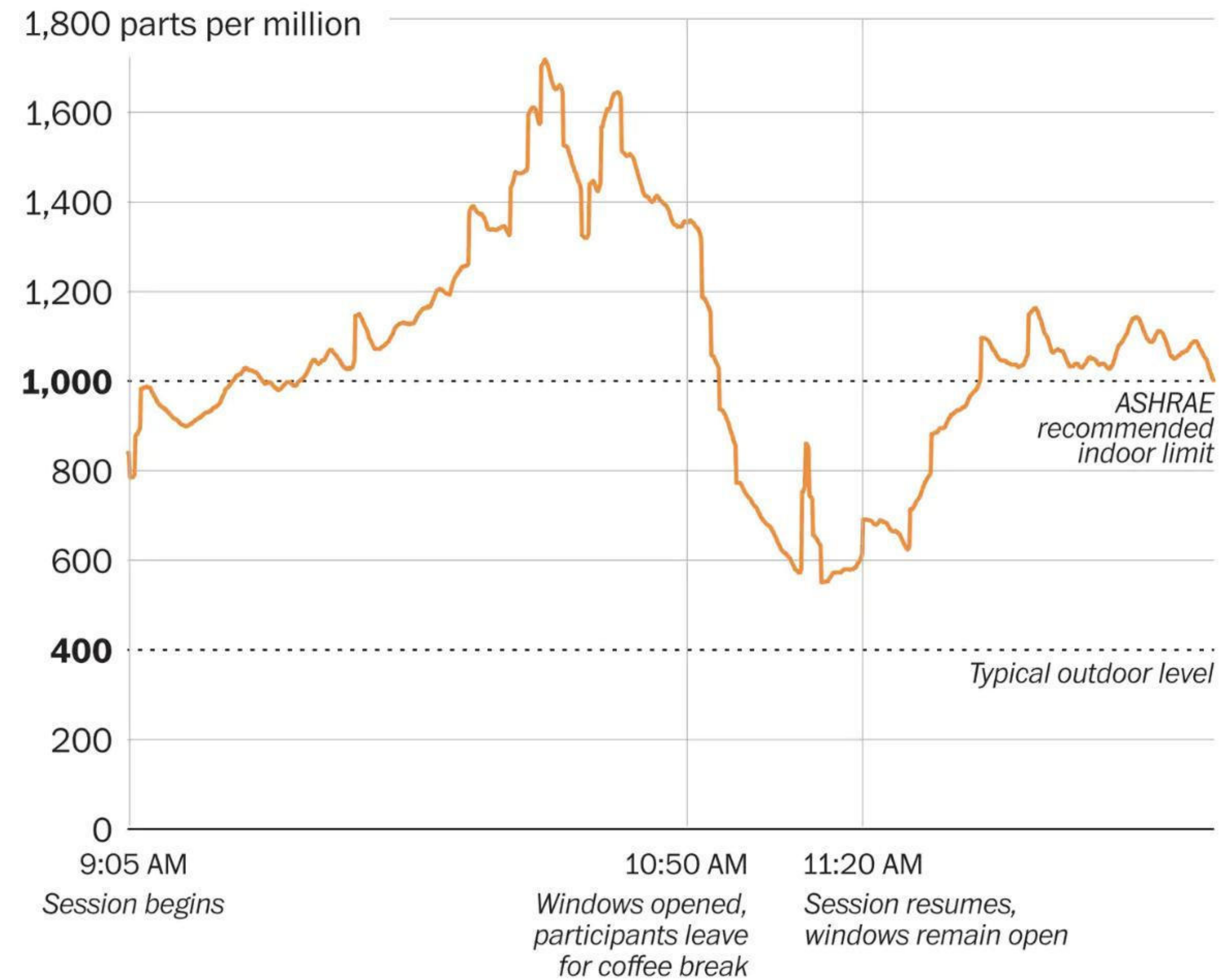
# Visual Form .....

Follow design rules and data visualization principles.



# Clearing the air

CO<sub>2</sub> levels in an occupied conference room on June 4, 2019



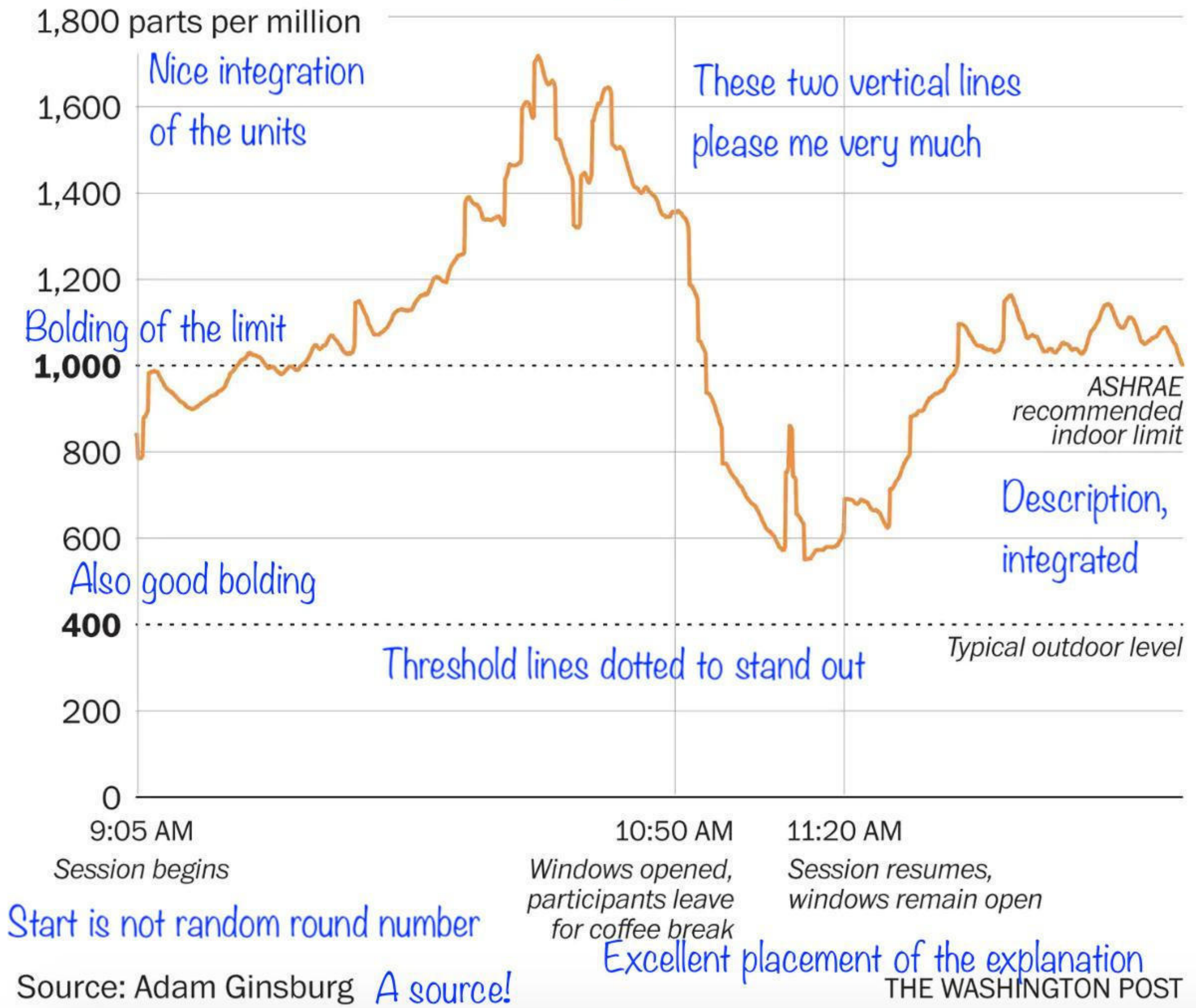
Source: Adam Ginsburg

THE WASHINGTON POST

*“Clearin the Air” by Adam Ginsburg (Washington Post)*

# Clearing the air *Fun and helpful title*

CO<sub>2</sub> levels in an occupied conference room on June 4, 2019  
*Units and metho in a subtitle, NOT in vertical text on the side*



*Notes by Francis Gagnon (Voilà)*

# Design for your audience.

- Choose charts based on your goal not tradition or novelty (only).
- Make sure your visualizations are accesible for everyone (colors, readability, ALT text).
- Use visual contrast to highlight important information.
- Provide meaningful labels.



# Design for your audience.

- Choose charts based on your goal not tradition or novelty (only).
- Make sure your visualizations are accesible for everyone (colors, readability, ALT text).
- Use visual contrast to highlight important information.
- Provide meaningful labels.

# Be honest.

- Show the distribution of the raw data if possible.
- Don't truncate bar charts, add spacing to truncated axes.
- Be consistent with axis scaling (especially in case of small multiples).

# Design for your audience.

- Choose charts based on your goal not tradition or novelty (only).
- Make sure your visualizations are accesible for everyone (colors, readability, ALT text).
- Use visual contrast to highlight important information.
- Provide meaningful labels.

# Be honest.

- Show the distribution of the raw data if possible.
- Don't truncate bar charts, add spacing to truncated axes.
- Be consistent with axis scaling (especially in case of small multiples).

# Lend a helping hand.

- Use annotations and direct labels instead / in addition to captions and legends.
- Order your data, either by value or intrinsic ranking.
- Focus on the main message and reduce data complexity.
- Reveal information step by step (if applicable).

# Thank you!

 [cedricscherer.com](http://cedricscherer.com)    @CedScherer  z3tt



*Buy me a coffee*

[www.buymeacoffee.com/z3tt](http://www.buymeacoffee.com/z3tt)