

# Datenvisualisierung

## Wie mache ich das richtig?

Dr. Cédric Scherer // 2. Juni 2023

 cedricscherer.com    @CedScherer  z3tt

@ JuGa Summer School 2023

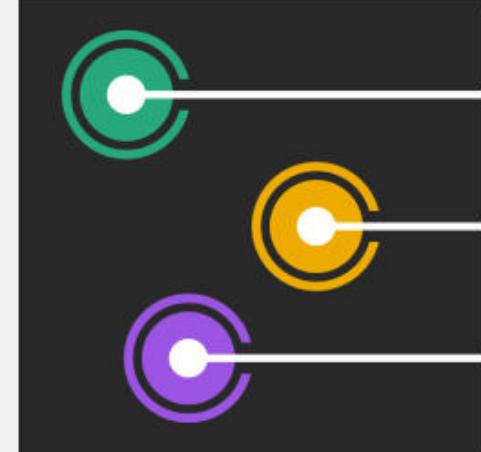
# Datenvisualisierung

## Wie mache ich das richtig (gut)!

Dr. Cédric Scherer // 2. Juni 2023

 cedricscherer.com    @CedScherer  z3tt

@ JuGa Summer School 2023



# CÉDRIC SCHERER

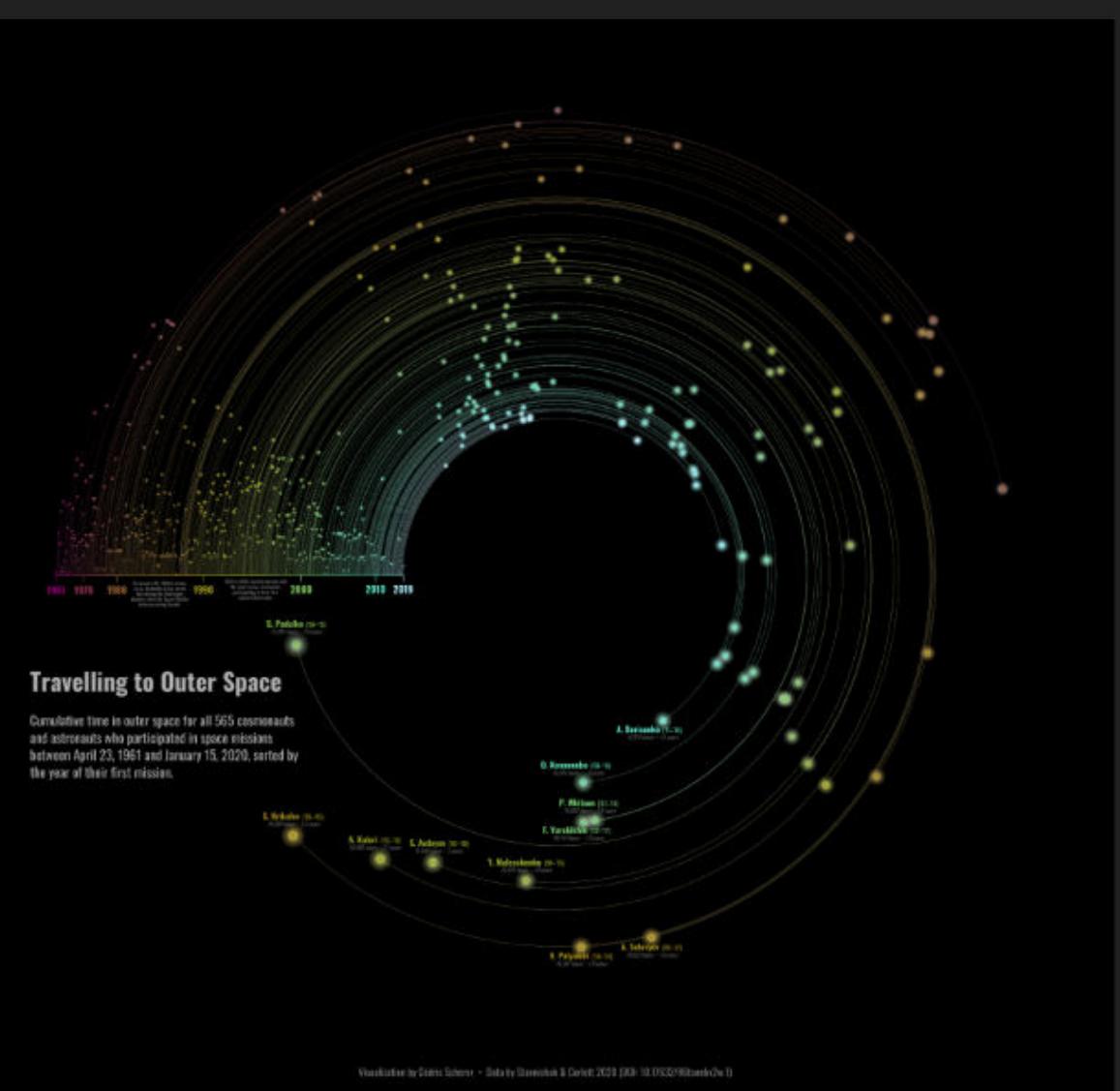
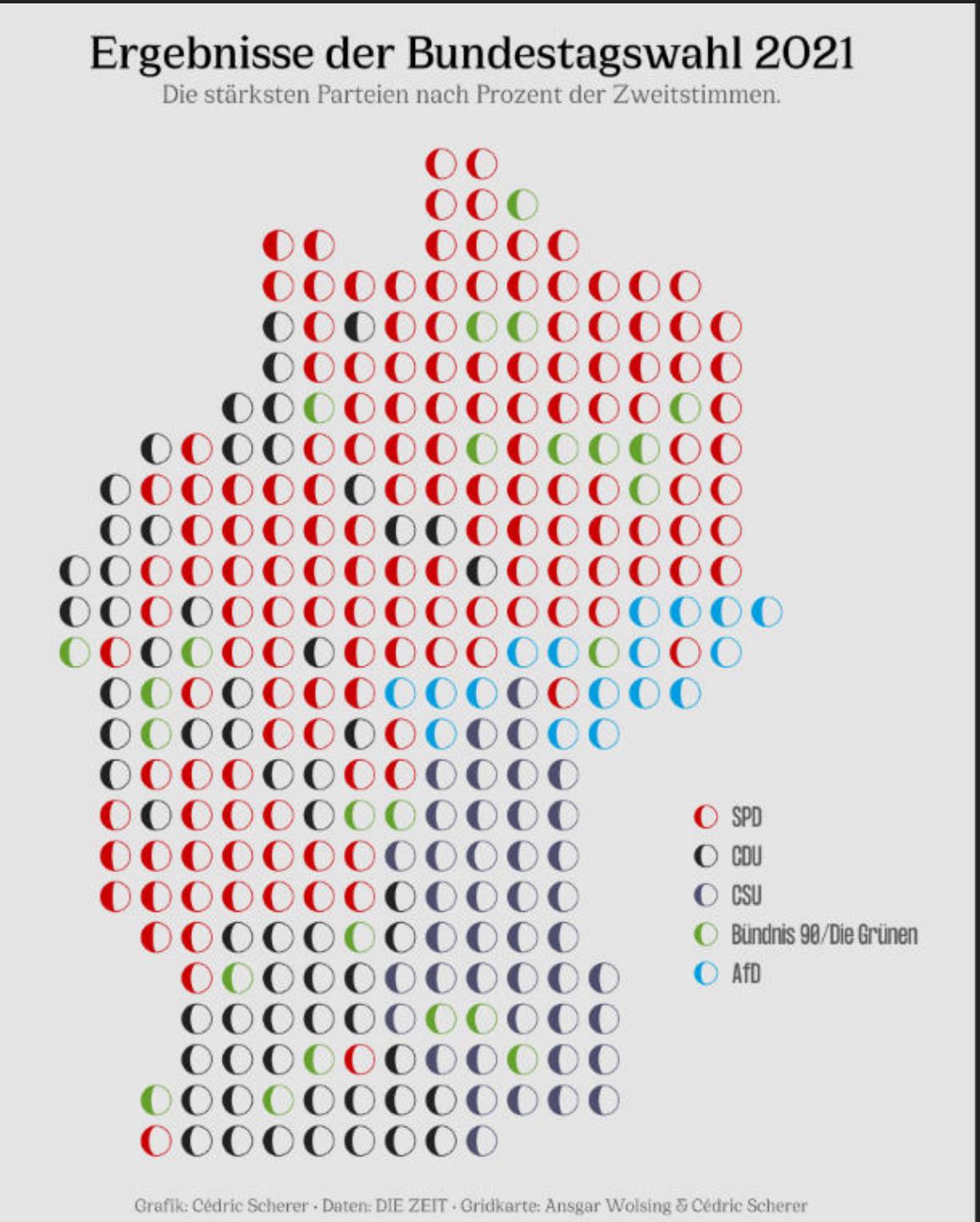
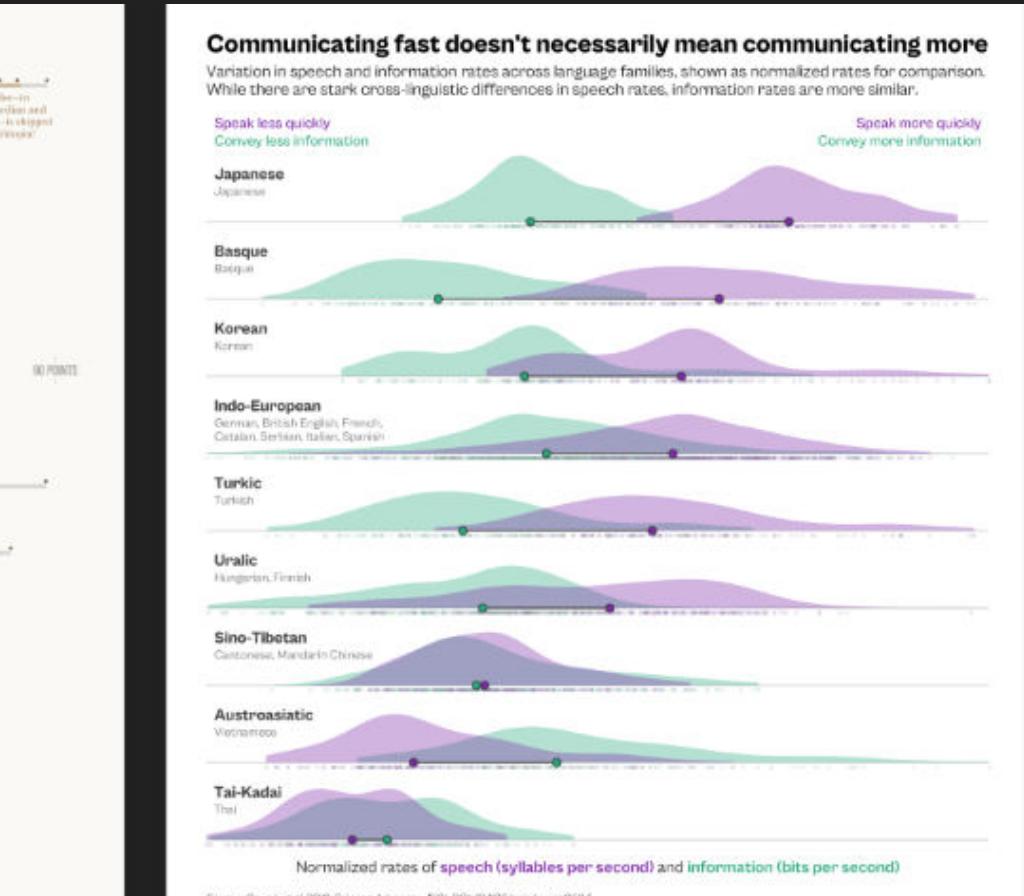
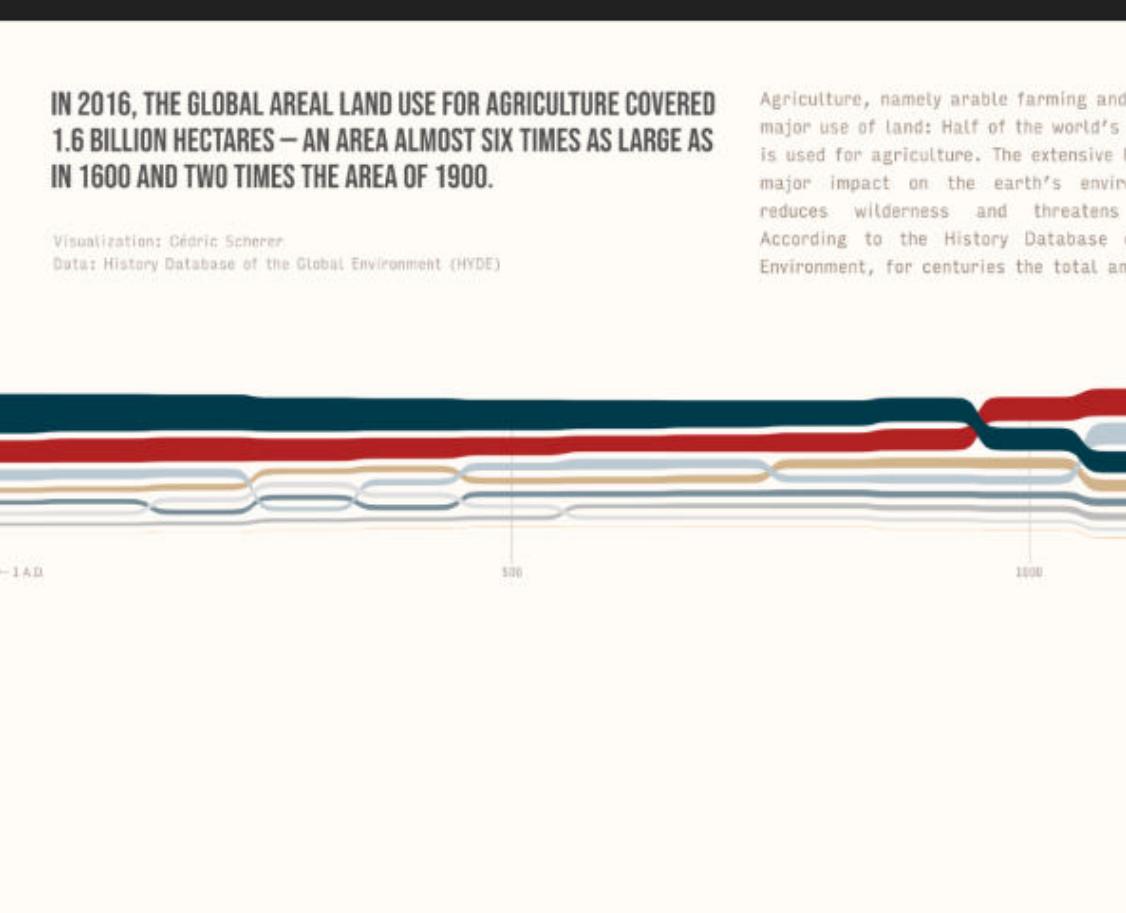
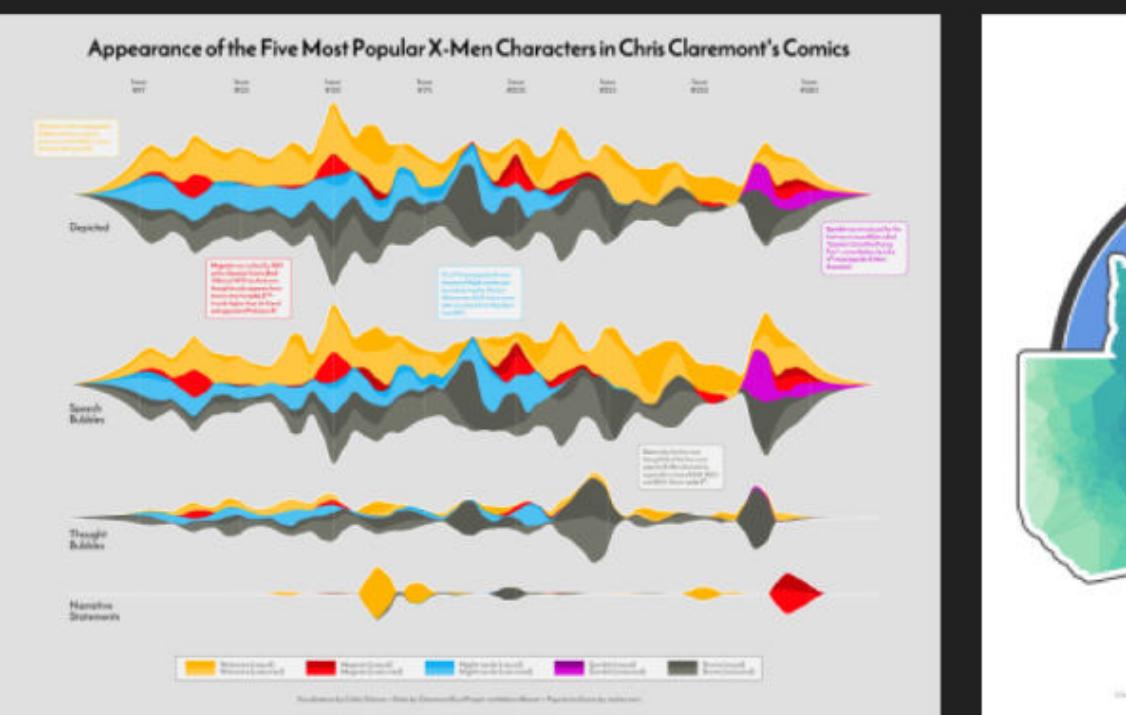
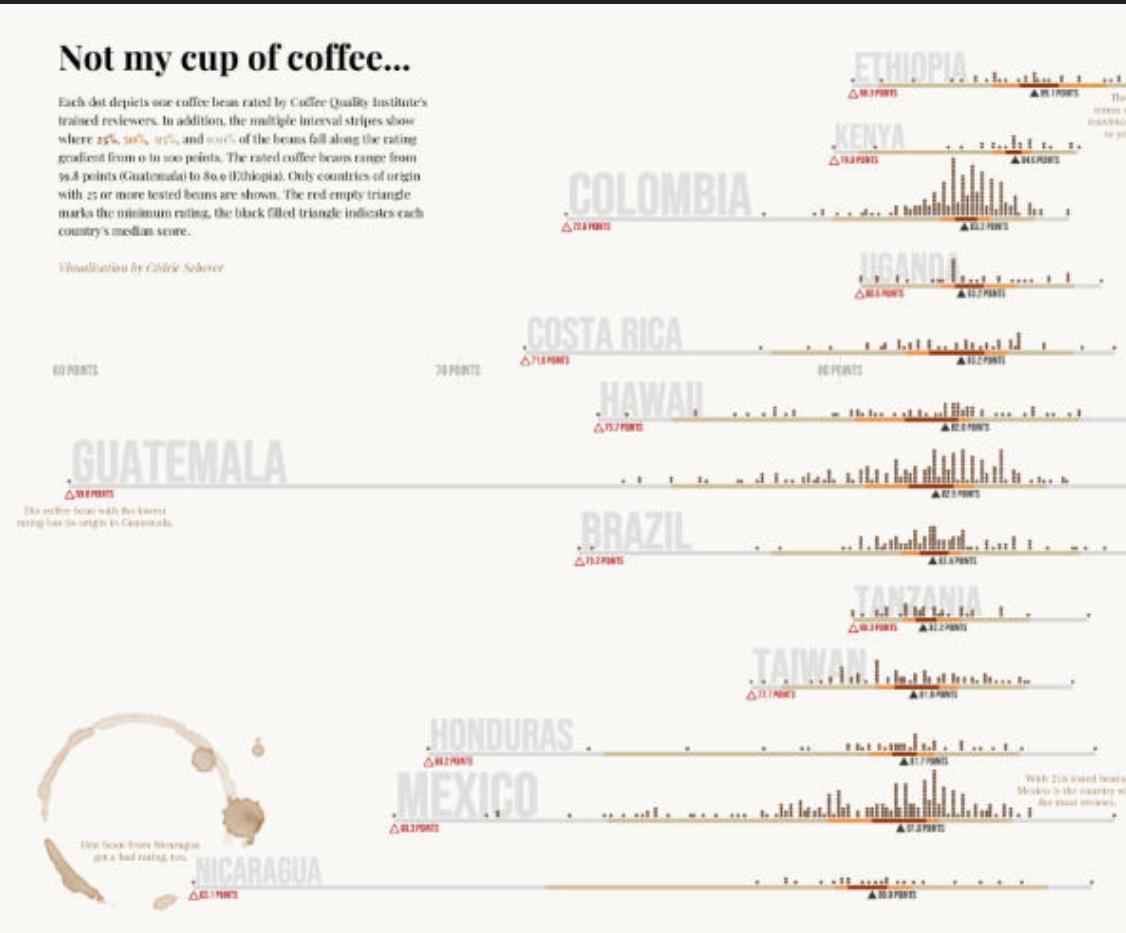
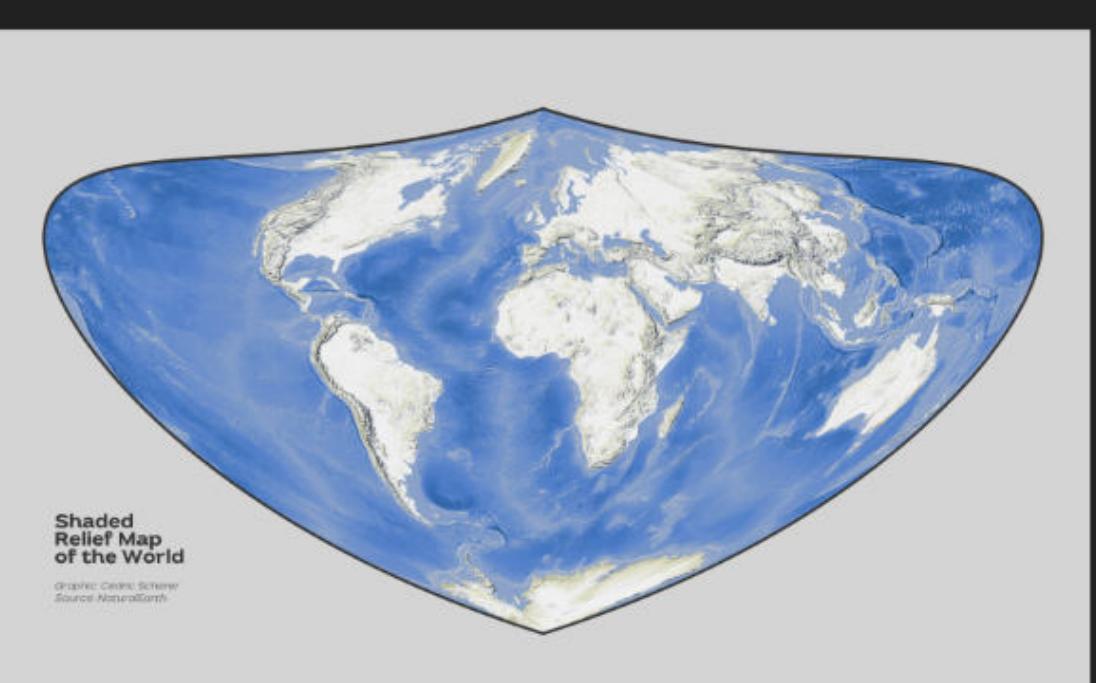
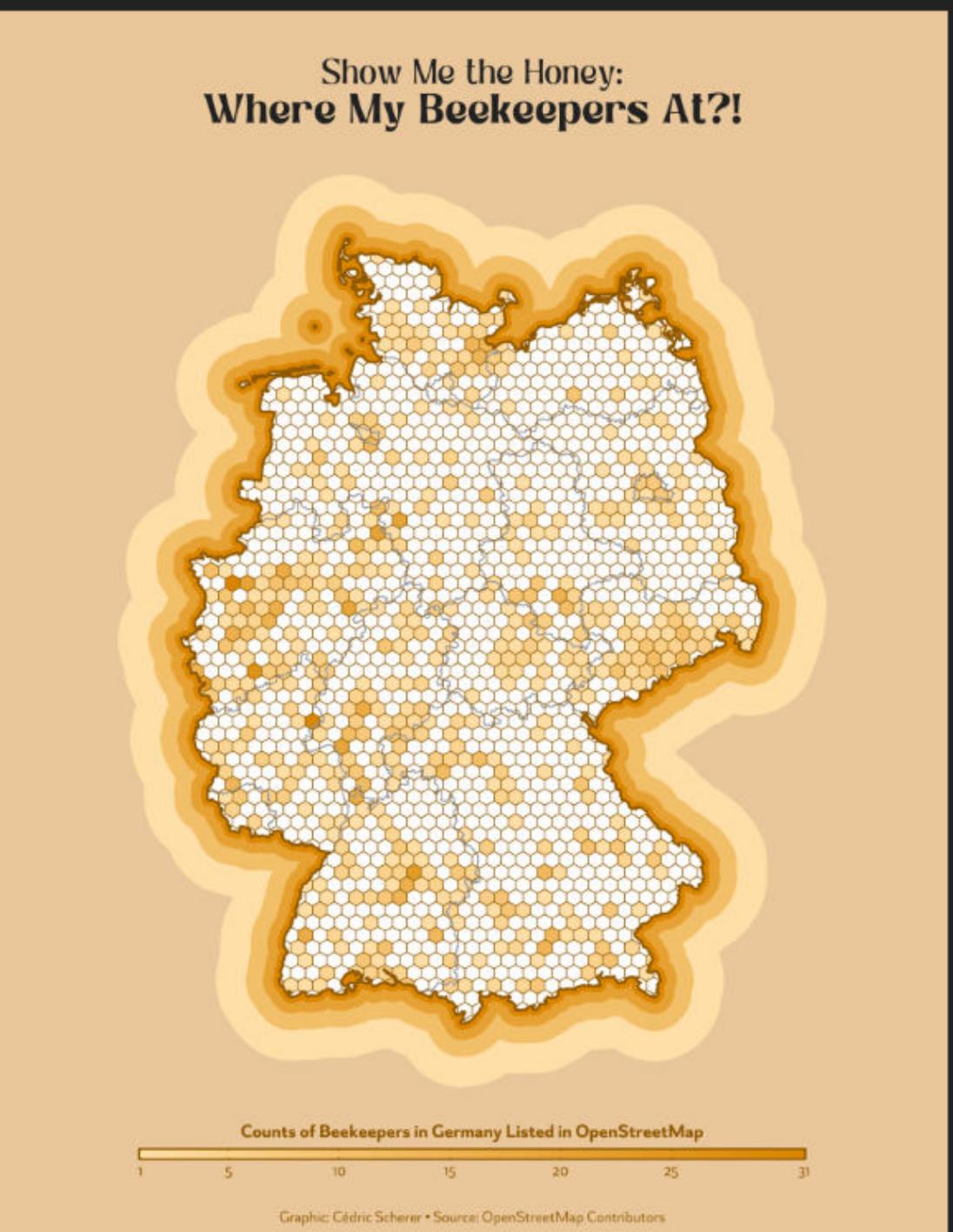
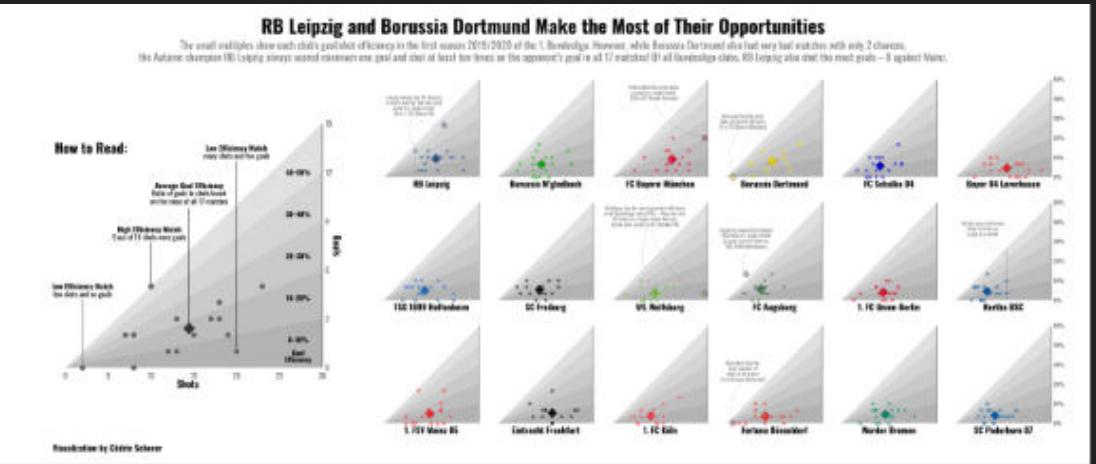
Data Visualization & Information Design

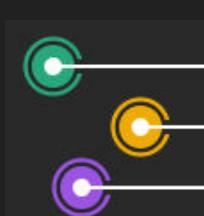
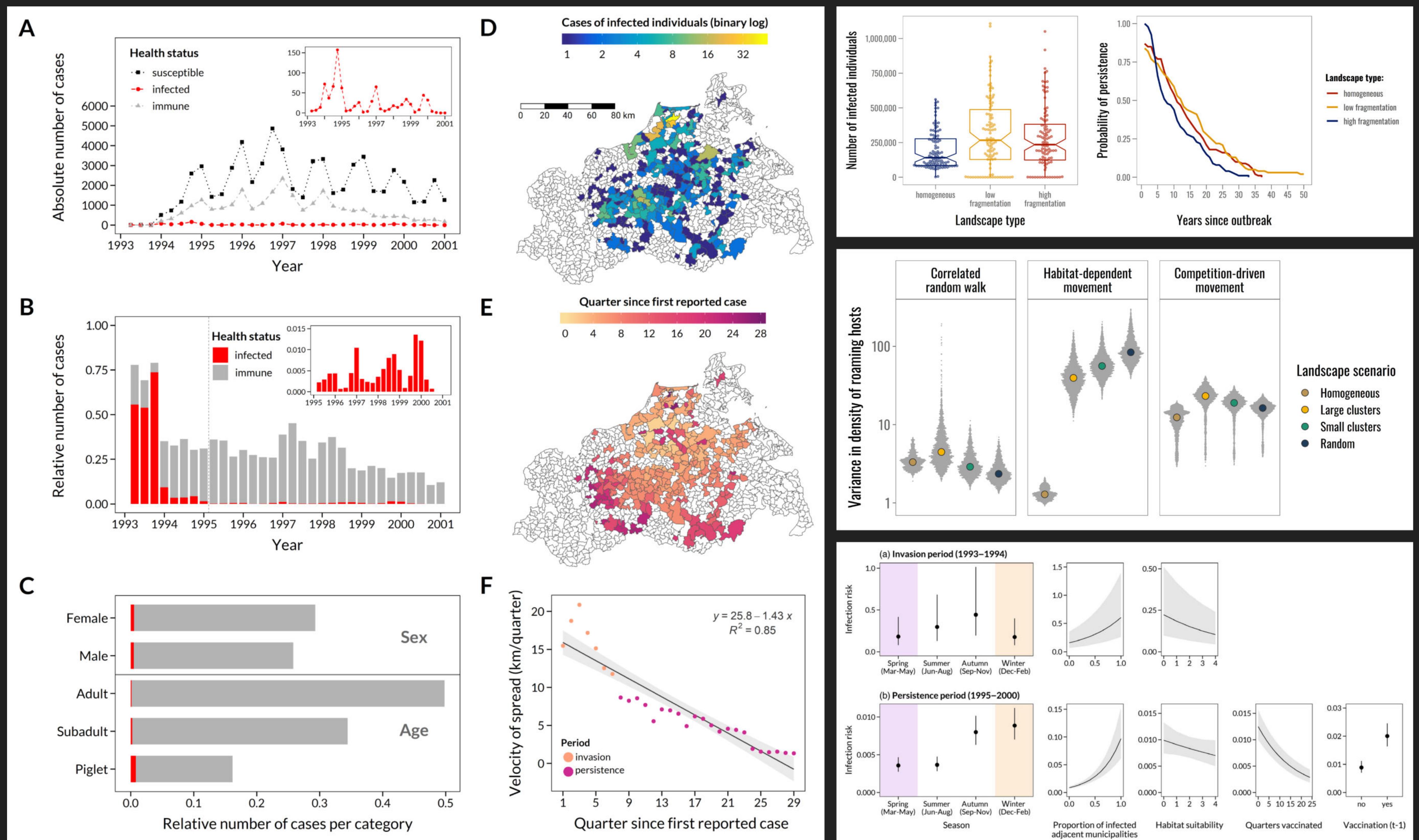


**Consulting**

**Coaching**

**Coding**





# CÉDRIC SCHERER

Data Visualization & Information Design

Hi, I am Cédric 

Data Visualization Designer, Consultant and Instructor  
for Engaging and Effective Graphical Storytelling.

[» Read more about me](#)   [» Schedule a discovery call](#)



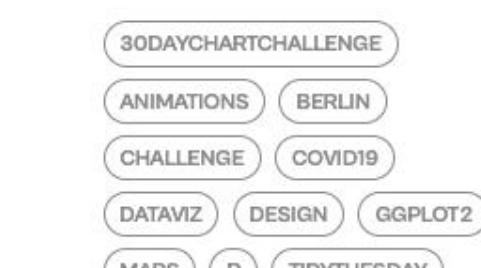
Always coding. Passionate about  
design. Worried about nature and  
inequality. Proud dad.



## Quick Links

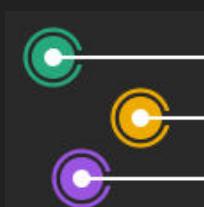
[ggplot2 Tutorial](#)  
[Evolution of a ggplot](#)  
[rstudio::conf Workshop](#)

## Featured Tags



## 2-Day Workshop on "Graphic Design with ggplot2" at rstudio::conf 2022

End of July, I had the honor to teach an in-person ggplot2 workshop at the rstudio::conf in Washington DC. All course resources are available on the course webpage featuring slides, hands-on R codes, recap notes, and exercises including prepared scripts and step-by-step



# Datenvisualisierung

---

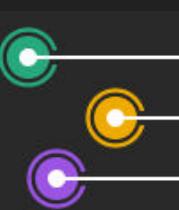
ist die grafische Darstellung von  
Informationen und Daten.



# Datenvisualisierung

---

wandelt Daten in visuelle Formen als  
quantifizierbare Merkmale um.



# Datenvisualisierung

---

hilft Einsichten zu gewinnen, zu entdecken, erklären und zu entscheiden.



# Datenvisualisierung

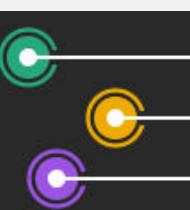
---

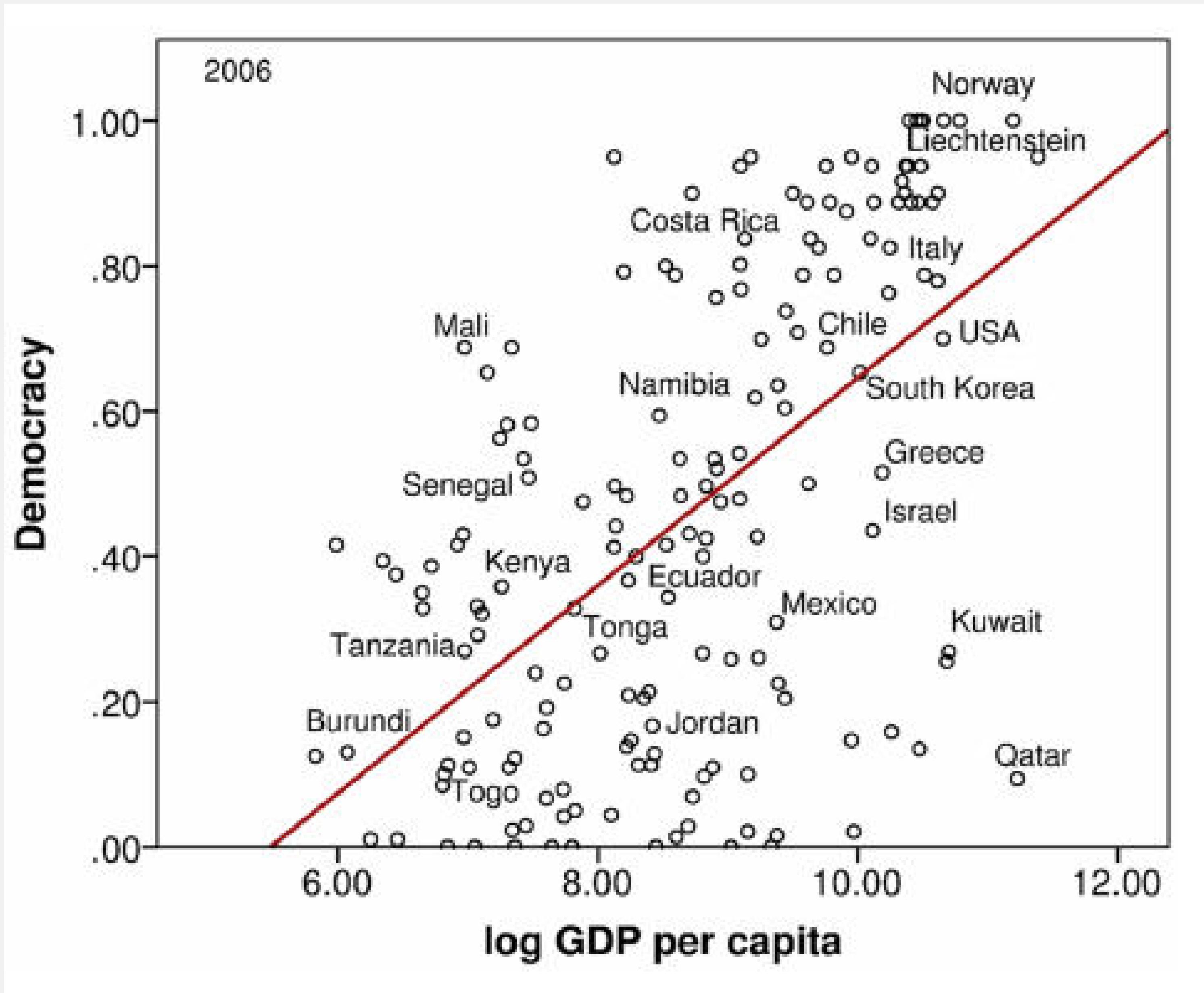
ist einerseits Kunst und  
andererseits Wissenschaft.



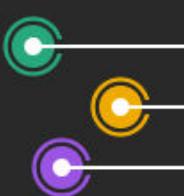


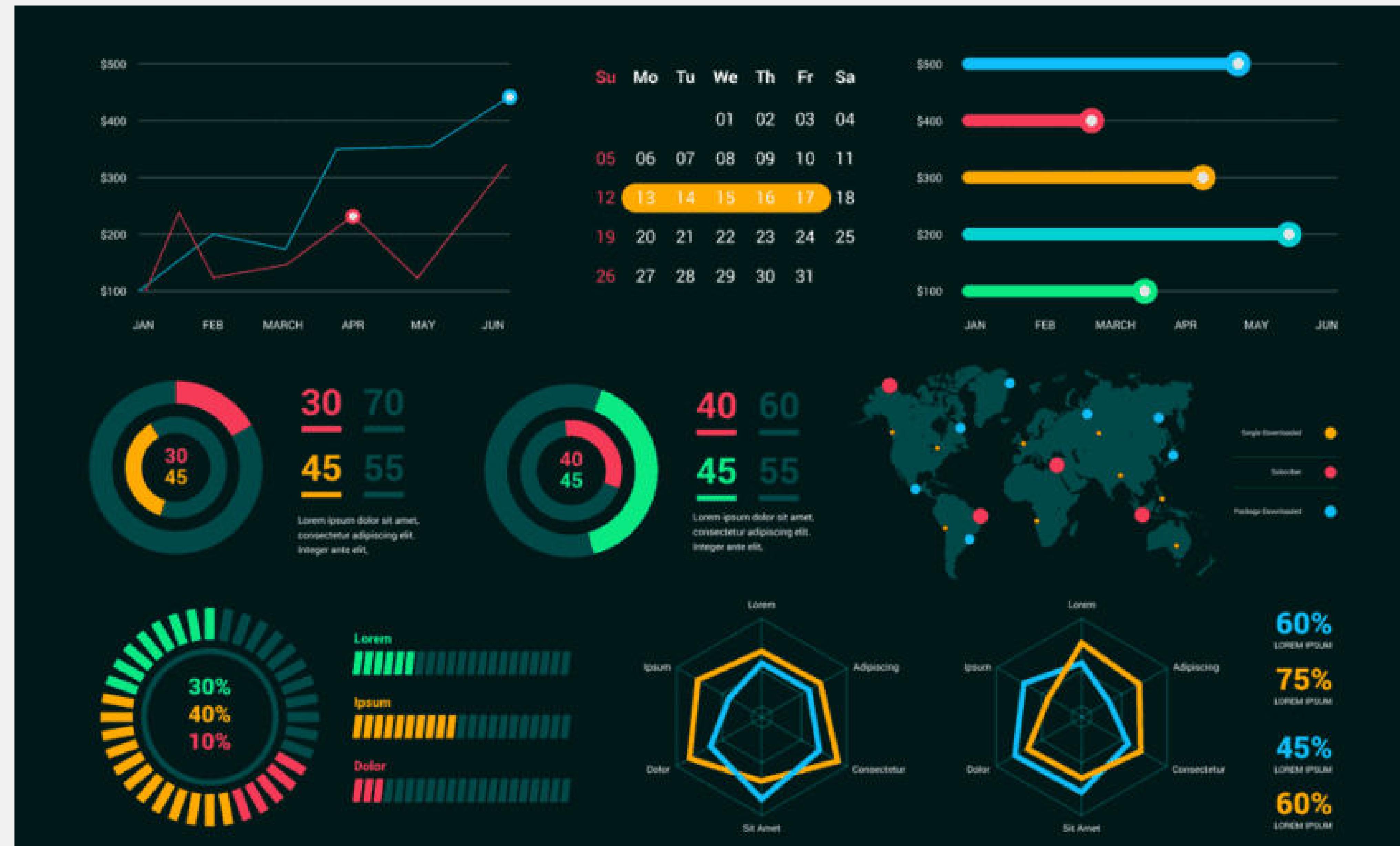
Quelle: eazybi.com



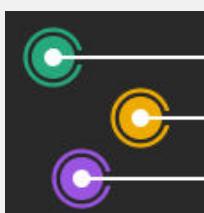


Quelle: *Ranganathan et al. 2014*





Quelle: [datameer.com](http://datameer.com)

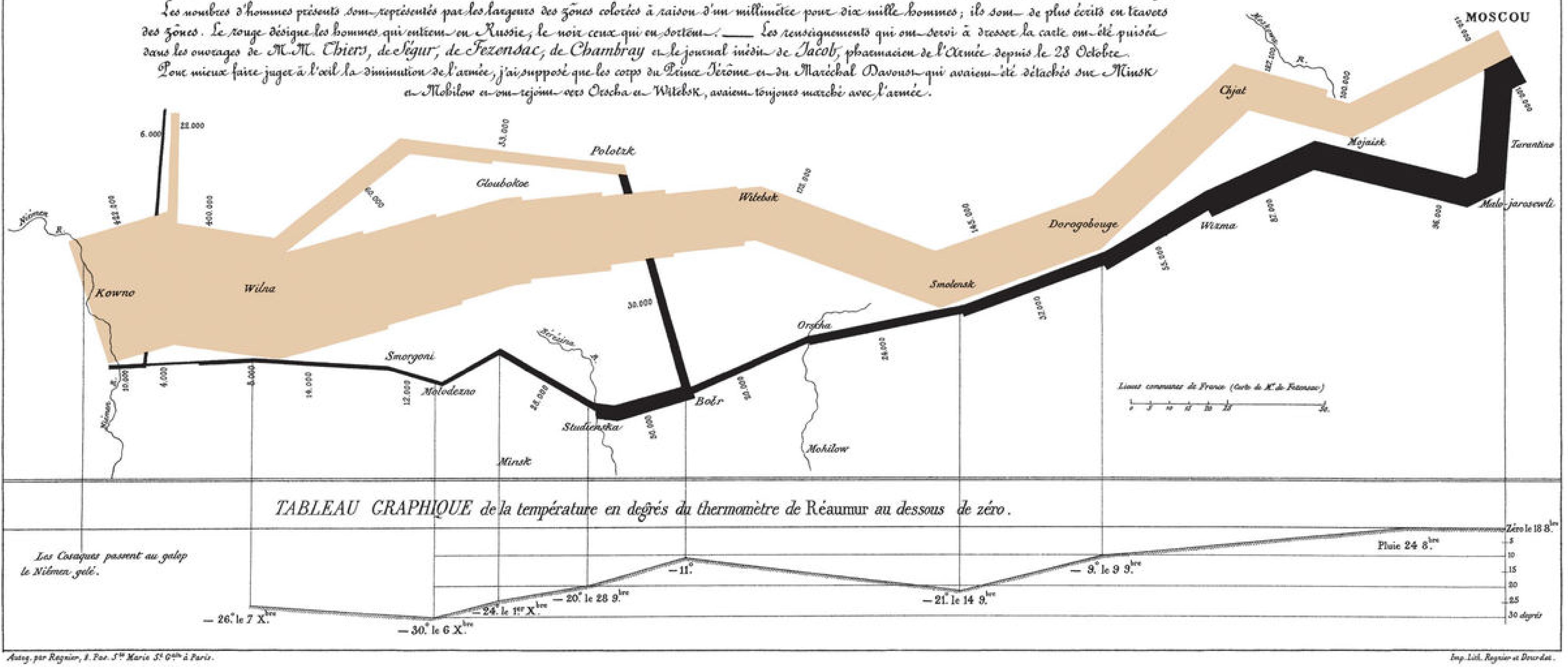


# Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

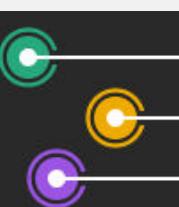
Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

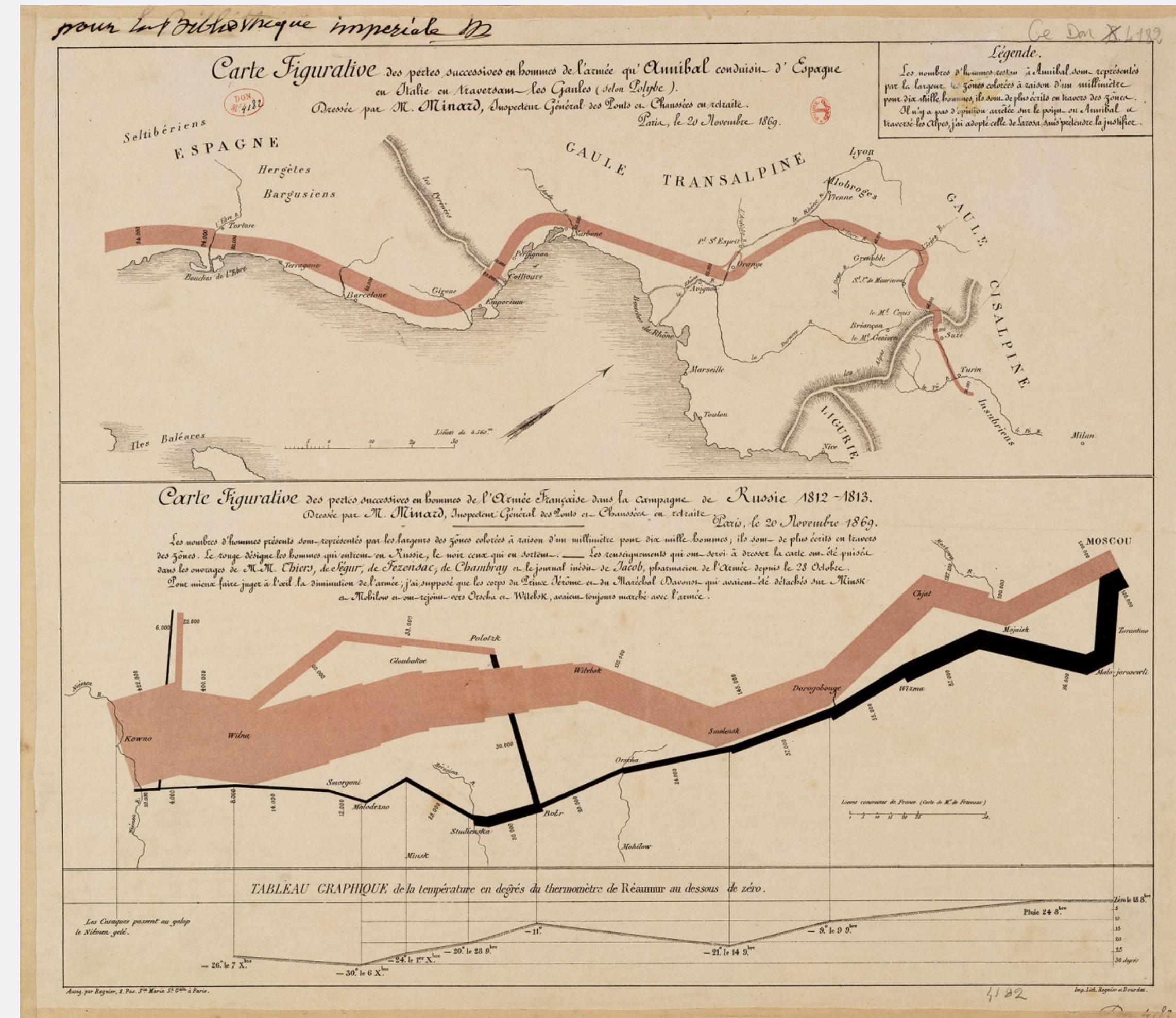
Les nombres d'hommes perdus sont représentés par les larges des zones colorées à raison d'un millimètre pour six mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chier, de Segur, de Fezensac, de Charnbray et le journal intitulé 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 Mogilow et qui rejoignirent Ossaka et Wilcok, avaient toujours marché avec l'armée.



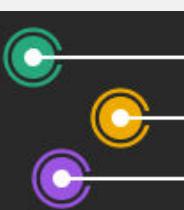
Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813 von Charles Joseph Minard

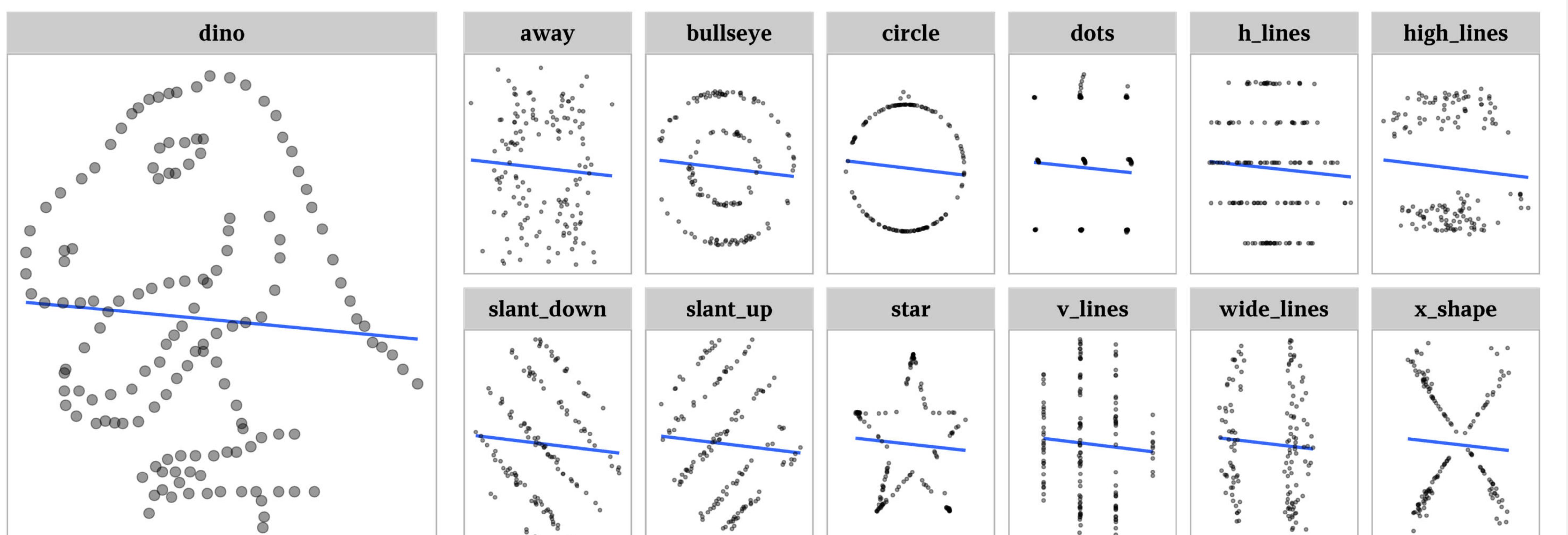




Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813 und Carte figurative des pertes successives en hommes de l'Armée qu'Annibal conduisit d'Espagne en Italie en traversant les Gaules (selon Polybe) von Charles Joseph Minard

- zeigt das Vorrücken der Truppen von **Hannibal** (218 v. Chr.) und **Napoleon** (1812-1813)
- wird oft als **die beste jemals gezeichnete statistische Grafik** bezeichnet



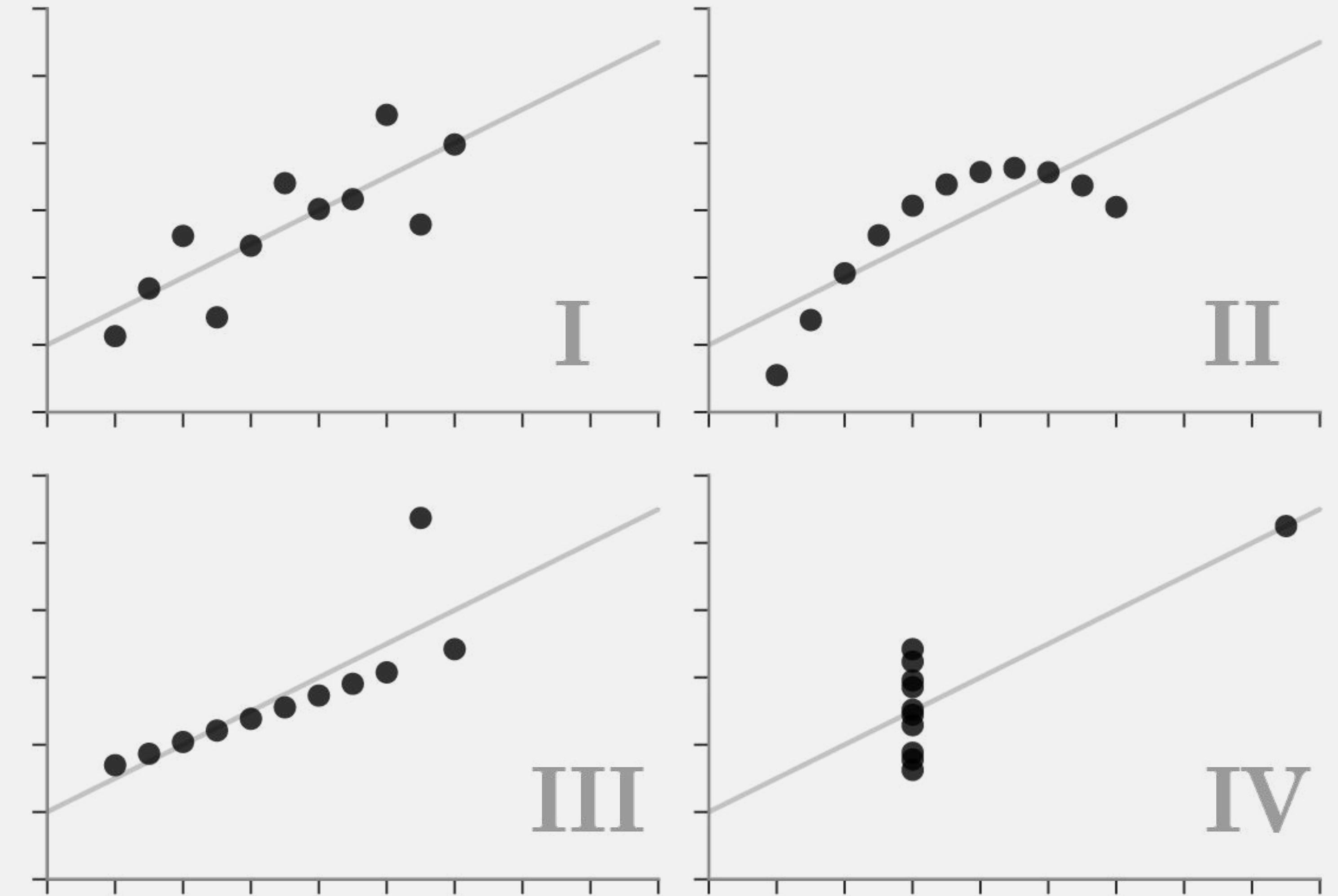


*“Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing”*  
 von Justin Matejka & George Fitzmaurice, ACM SIGCHI Conference on Human Factors in Computing Systems 2017

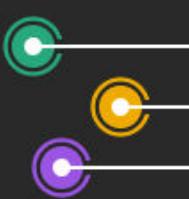


# Anscombe's Quartet

**Each dataset has the same summary statistics  
but are visually distinct.**  
mean, standard deviation, and correlation

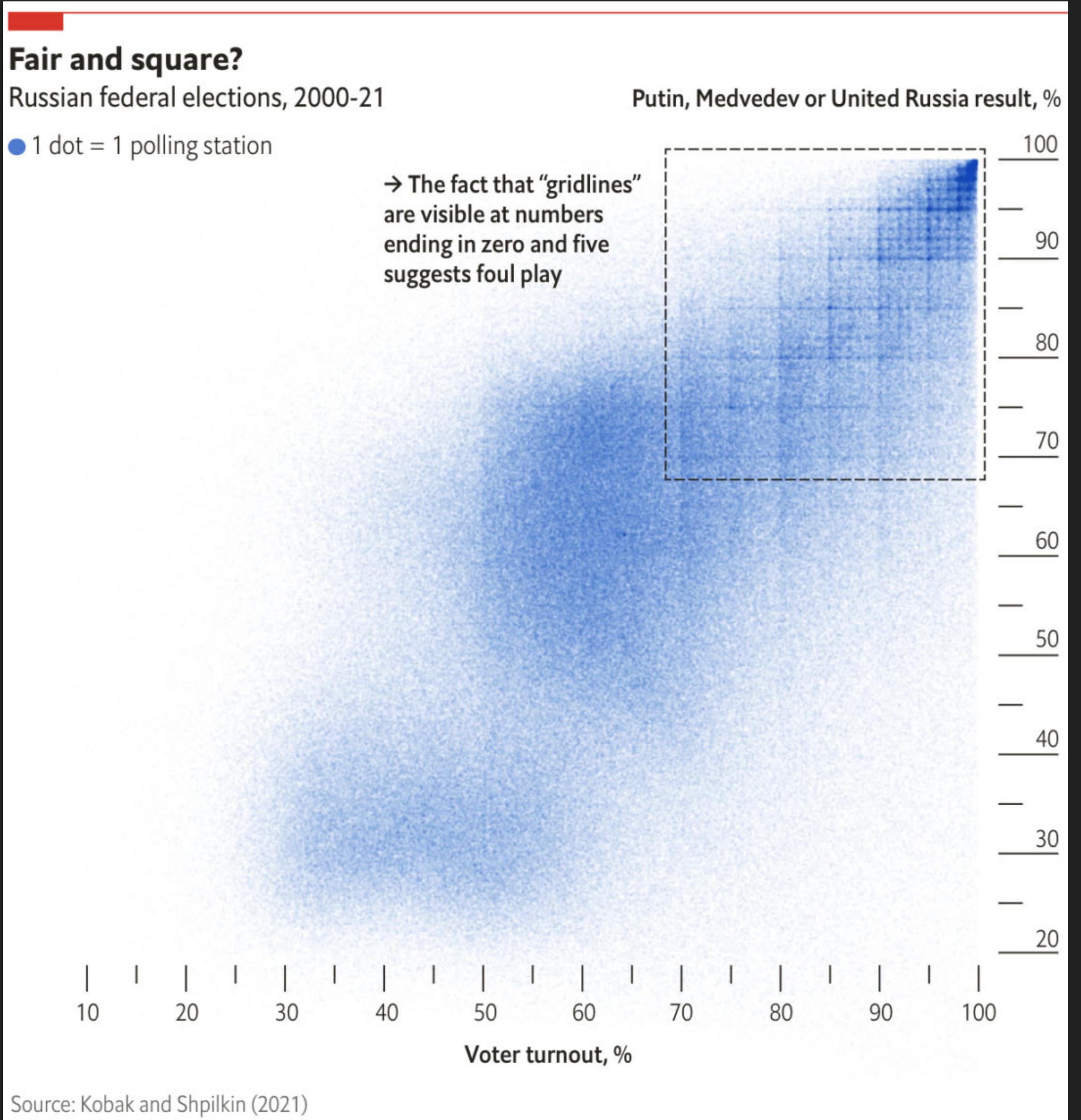


*“Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing”*  
von Justin Matejka & George Fitzmaurice, ACM SIGCHI Conference on Human Factors in Computing Systems 2017



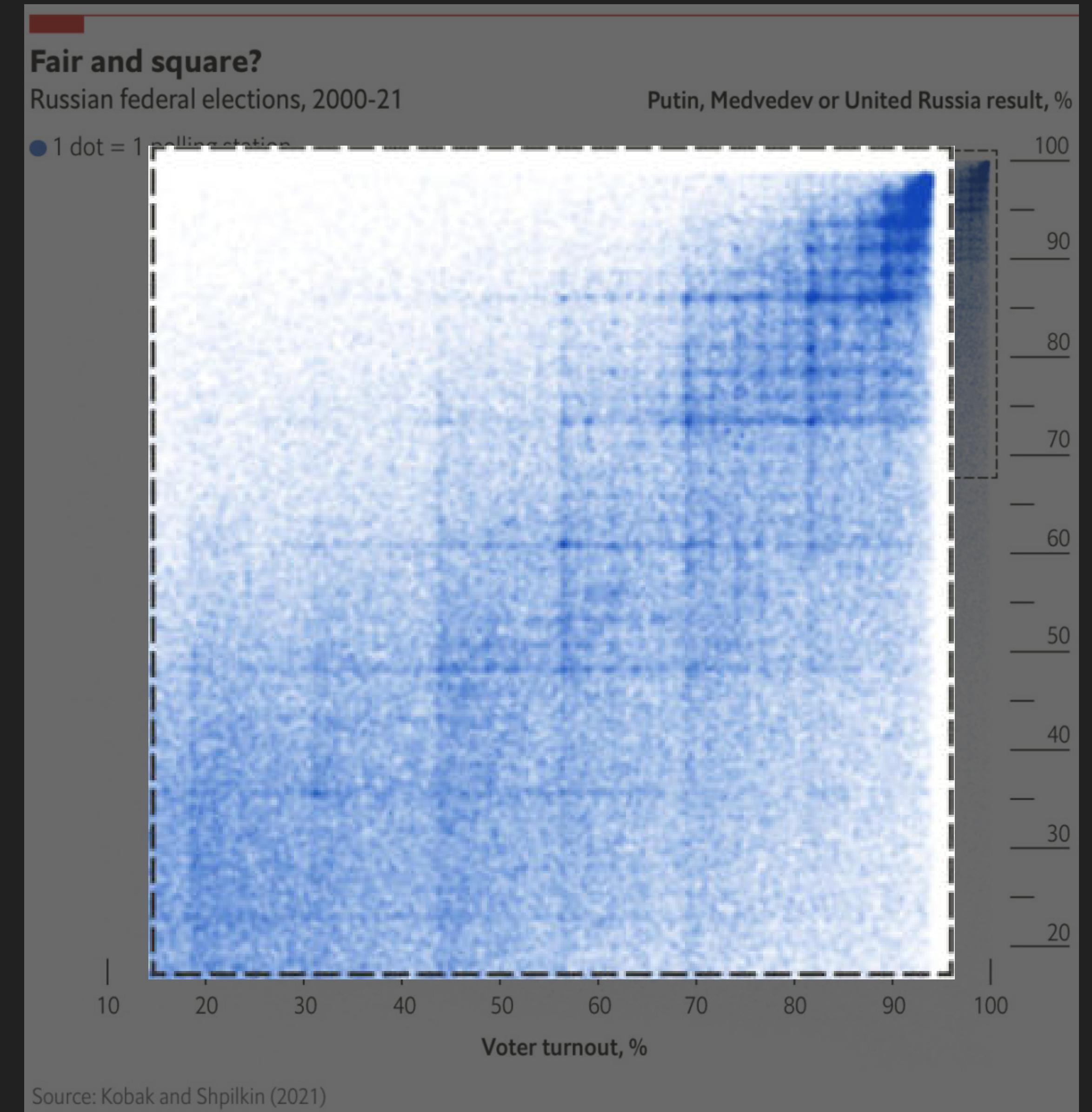
# Visualize Your Data

“When Dmitry Kobak and Sergey Shpilkin [...] analysed the results, they found that **an unusually high number of turnout and vote-share results were multiples of five** (eg, 50%, 55%, 60%), a tell-tale **sign of manipulation.**”

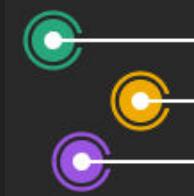
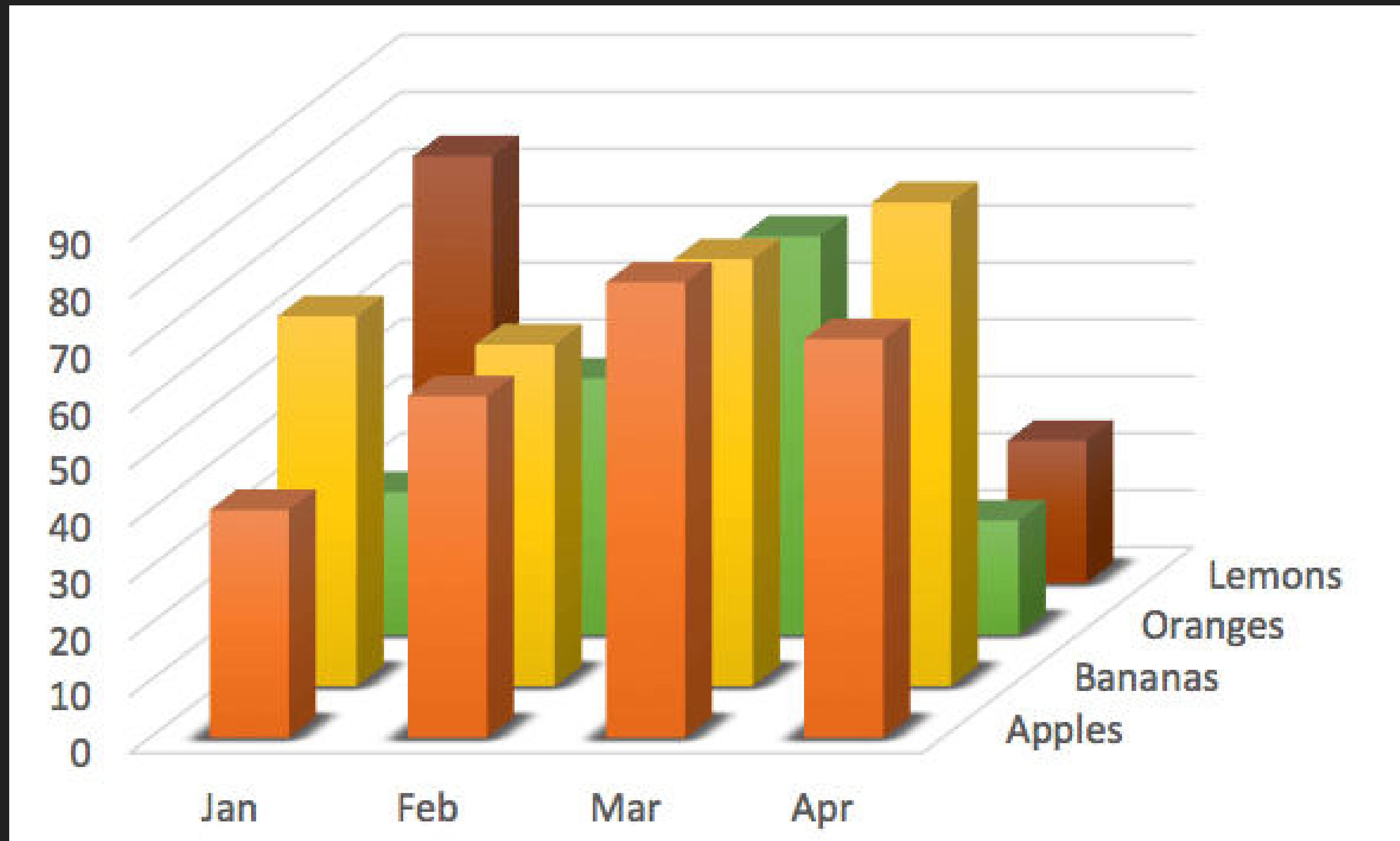


# Visualize Your Data

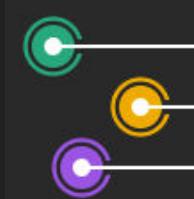
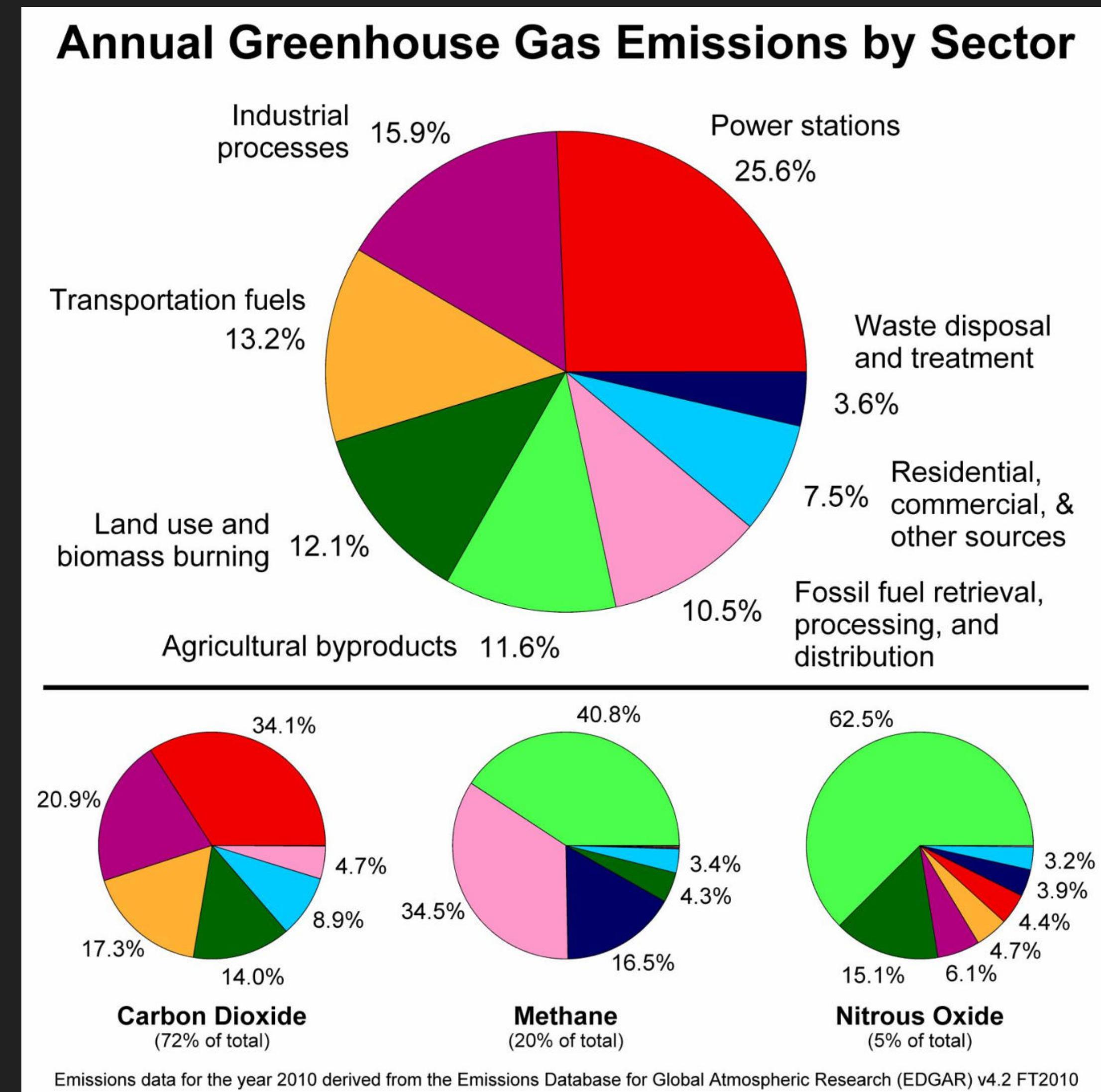
“When Dmitry Kobak and Sergey Shpilkin [...] analysed the results, they found that **an unusually high number of turnout and vote-share results were multiples of five** (eg, 50%, 55%, 60%), a tell-tale **sign of manipulation.**”



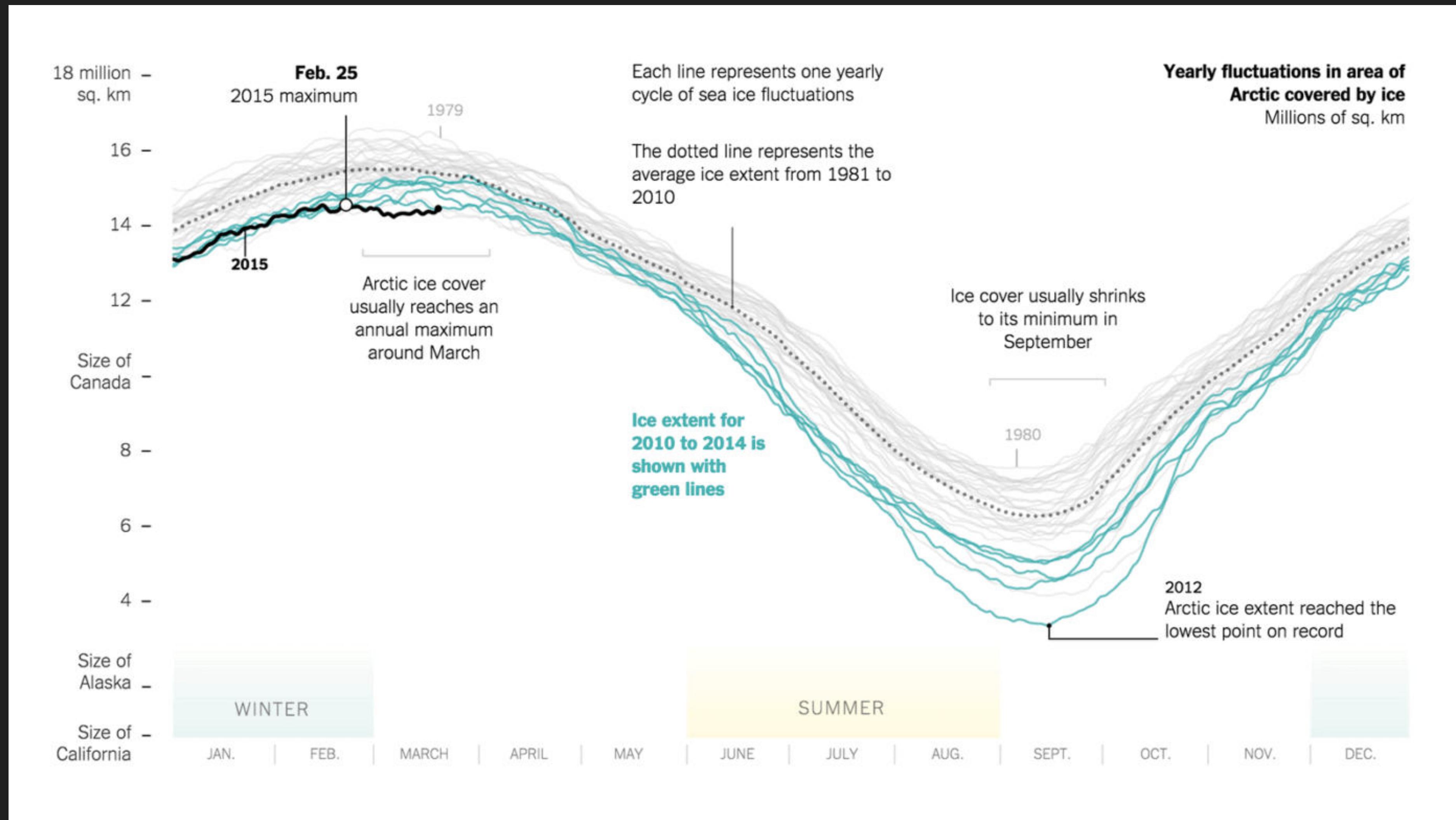
# Was macht es zu einer schlechten Grafik?



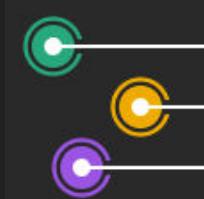
# Was macht es zu einer schlechten Grafik?



# Was macht es zu einer guten Grafik?

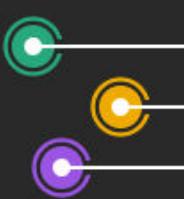


*"Yearly Fluctuations in Area of Arctic Covered by Ice"* by Derek Watkins (New York Times)



# Was macht es zu einer guten Grafik?

- ➔ **INFORMATION** (Korrektheit)
- ➔ **STORY** (Bedeutsamkeit)
- ➔ **GOAL** (Zweckmäßigkeit)
- ➔ **VISUAL FORM** (Schönheit)



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**



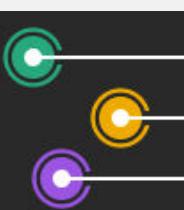
[cedricscherer.com](http://cedricscherer.com)



@CedScherer



z3tt



Unsere Daten sind nie ein  
perfektes Abbild der realen Welt.



cedricscherer.com



@CedScherer



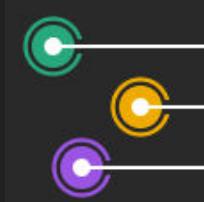
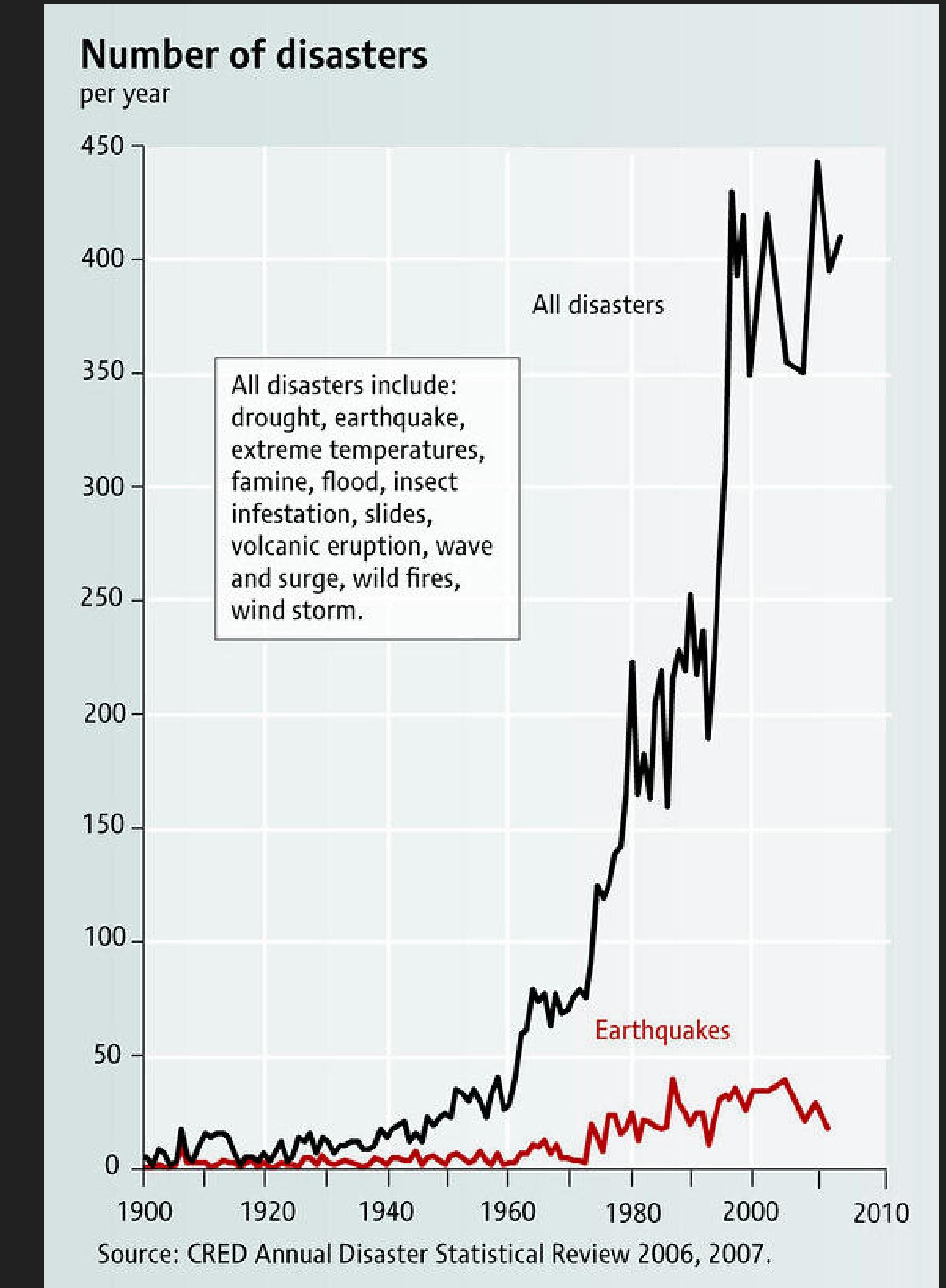
z3tt



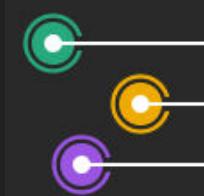
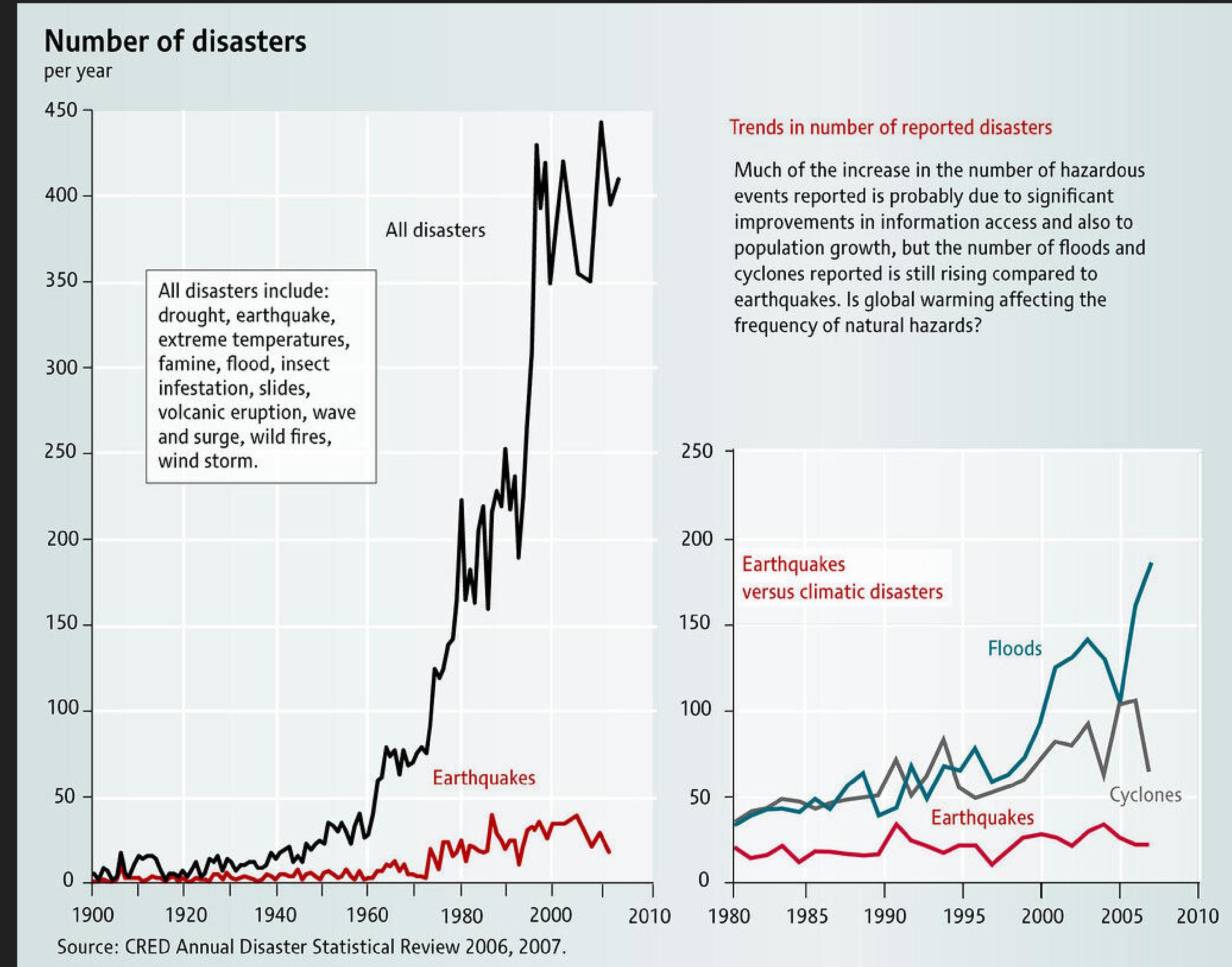
# Unsere Daten sind nie ein perfektes Abbild der realen Welt.

- nur eine Teilmenge: gemeldete Verbrechen
- von Menschen erhoben: Schätzungen, Genauigkeit & Fehler
- maschinell erfasst: Genauigkeiten & Fehler





“Ein Großteil des Anstiegs der **gemeldeten gefährlichen Ereignisse** ist wahrscheinlich auf den erheblich **verbesserten Zugang zu Informationen** zurückzuführen.”



# Beispiel: Beobachtungsstudien

**Postmenopausale Hormontherapie (HRT)**



**Risiko koronare Herzkrankheit (CVD)**



# Beispiel: Beobachtungsstudien

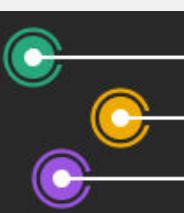
## Postmenopausale Hormontherapie (HRT)

Frauen, die HRT anwenden,  
sind in der Regel **wohlhabender,**  
**schlanker** und **gebildeter**,  
treiben **häufiger Sport** und  
trinken tendenziell **mehr Alkohol**

(Wilson et al. 1985, Persson et al. 1997)

**Confounding  
variables**

## Risiko koronare Herzkrankheit (CVD)



# Typologie von Informationsgrafiken

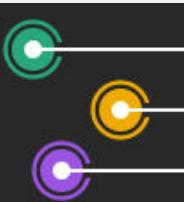
nach Juuso Koponen & Jonatan Hildén, "Data Visualization Handbook" (2020), Seite 25

Sind die **Informationen** konzeptionell oder messbar?

☞ **Art der Information:** Darstellung von Konzepten <> Umwandlung von Daten in visuelle Formen

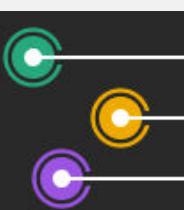
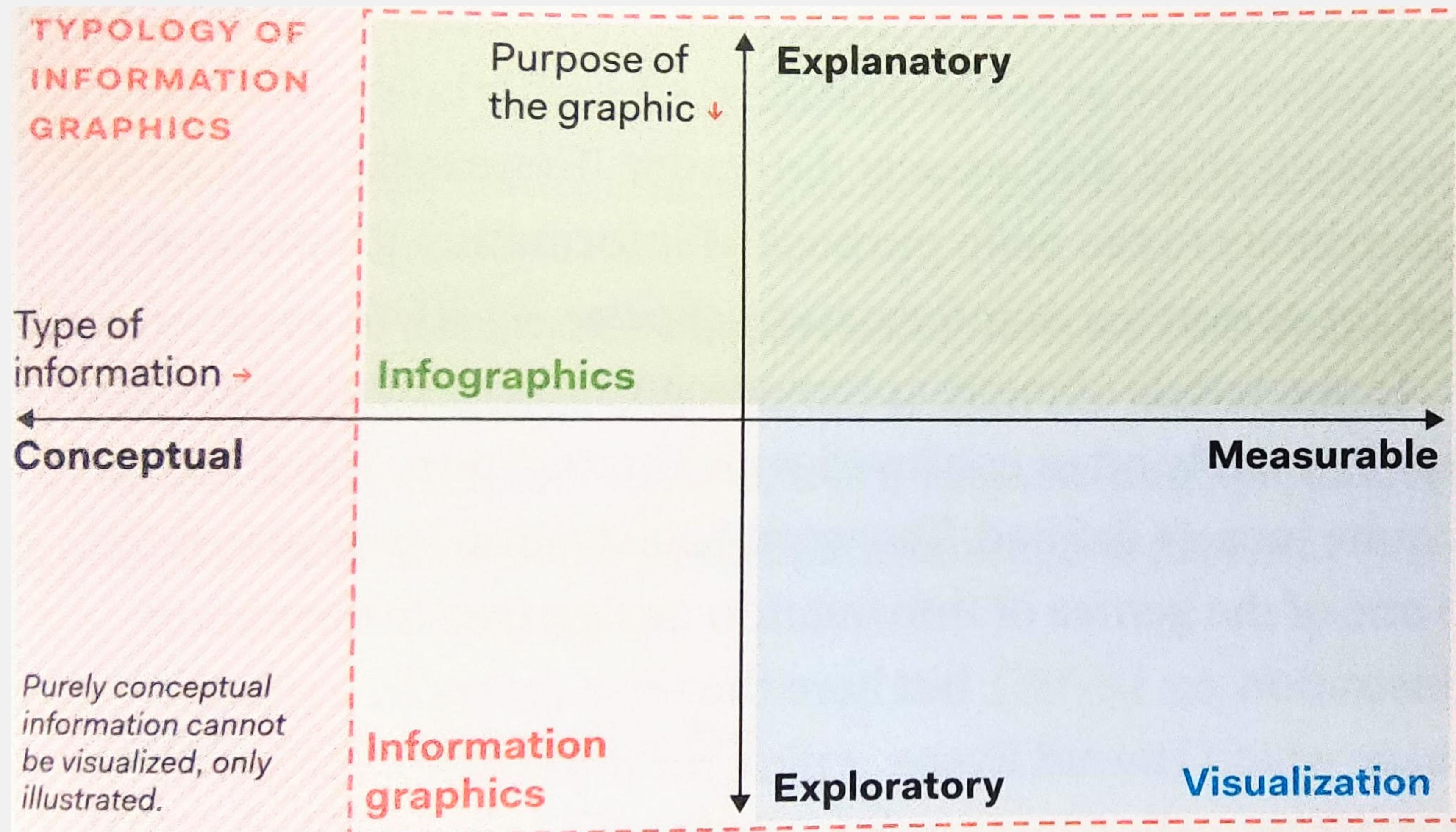
Ist der **Zweck** Informationen zu erkunden oder zu erklären?

☞ **Zweck der Grafik:** Erleichterung der Entdeckung <> Vermittlung von Informationen



# Typologie von Informationsgrafiken

nach Juuso Koponen & Jonatan Hildén, "Data Visualization Handbook" (2020), Seite 25



"Visualizations can be designed and experienced in various ways, by people of various backgrounds, and in various circumstances. That's why **reflecting on the purpose of a visualization is paramount before we design it—or before we critique it.**"

*Alberto Cairo*

Auszug aus dem Vorwort zu "Data Sketches" von Nadieh Bremer & Shirley Wu (CRC Press 2021)



cedricscherer.com



@CedScherer



z3tt





Visualiser Control

Viewer Control

Adaptiert von Andy Kirk



cedricscherer.com



@CedScherer



z3tt





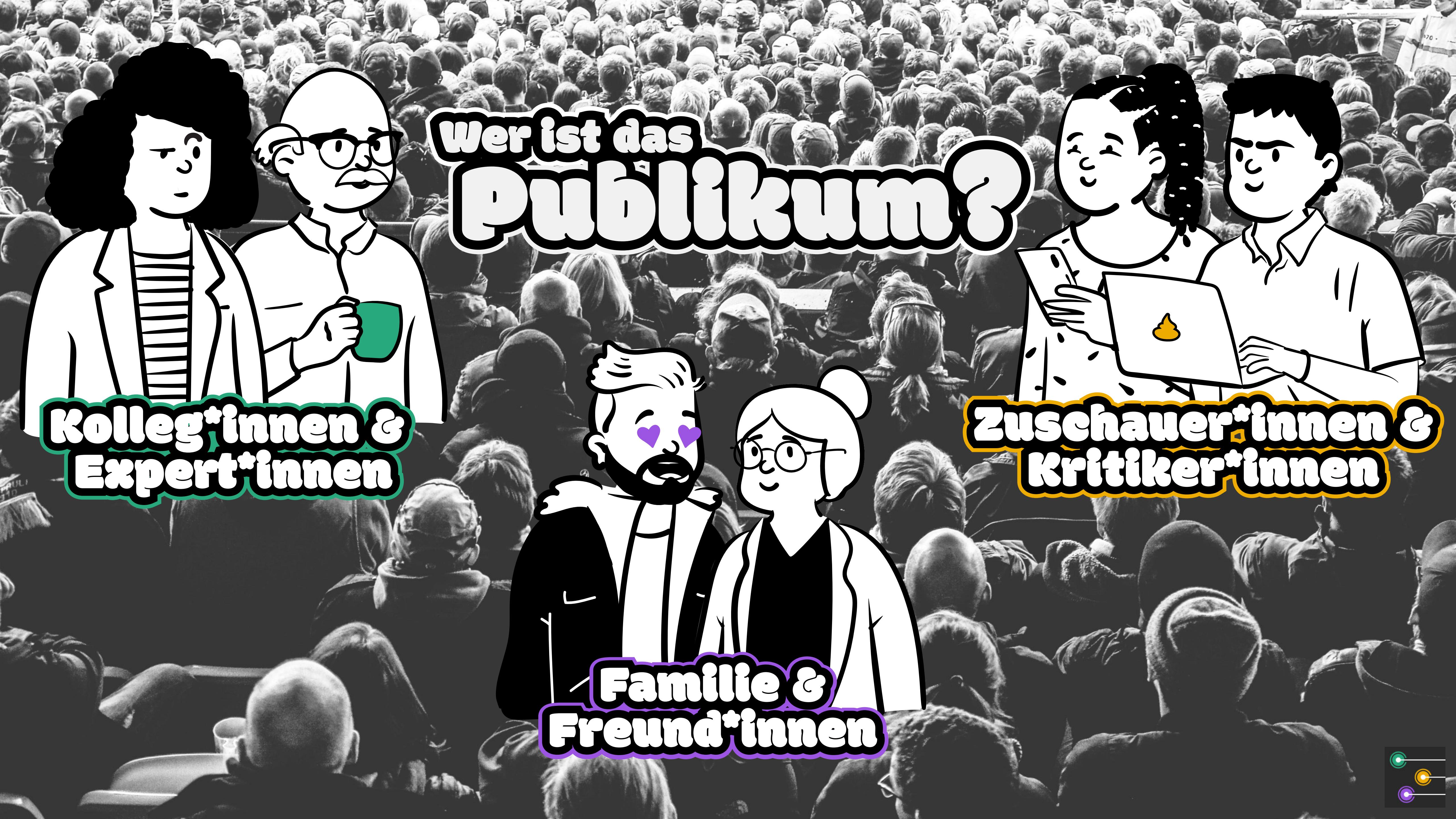
# **das Publikum**





Wer ist das  
**Publikum?**

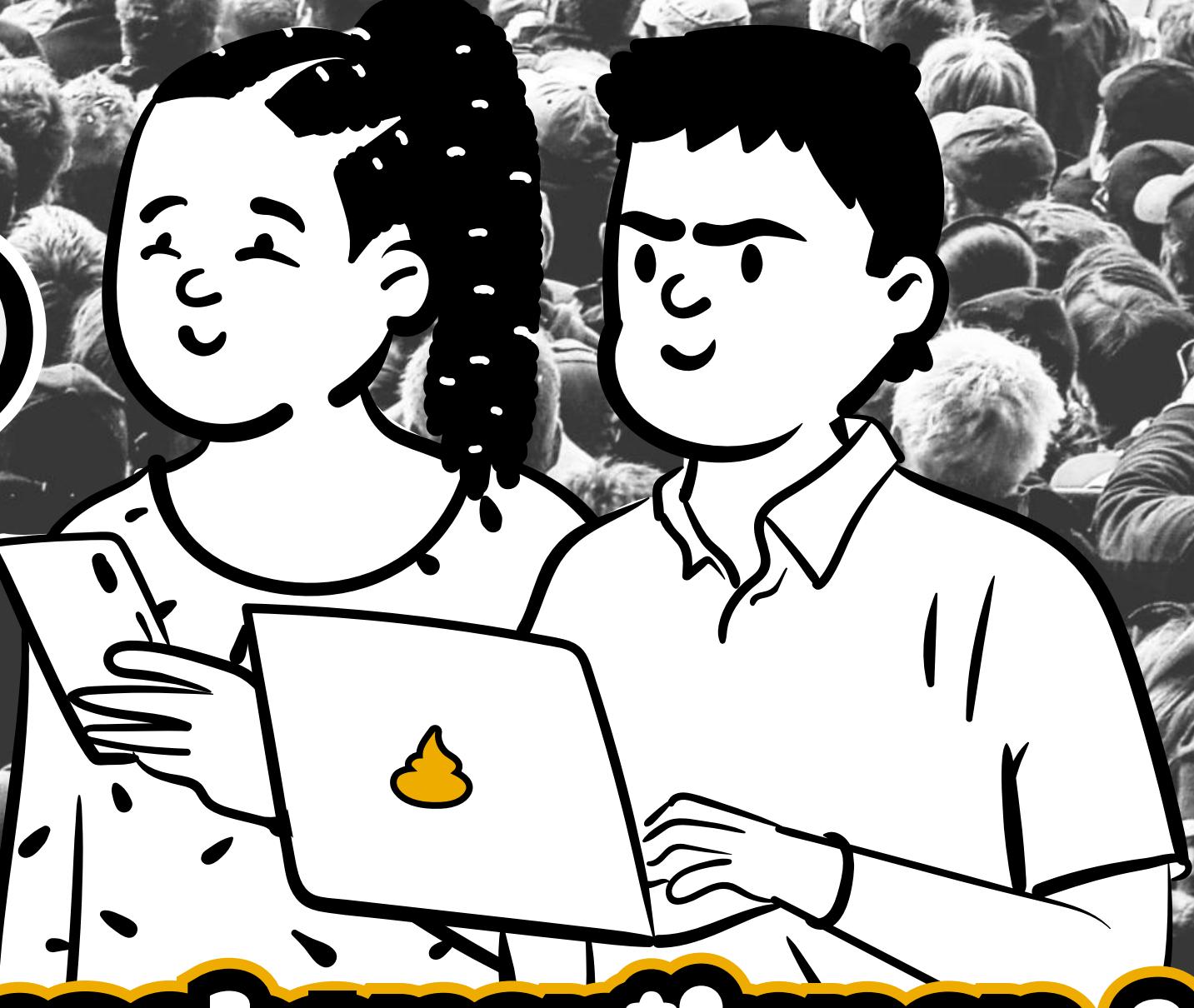




# Wer ist das Publikum?



Kolleg\*innen &  
Expert\*innen



Zuschauer\*innen &  
Kritiker\*innen



Familie &  
Freund\*innen



# Wer ist mein Zielpublikum?



Welche Geschichte ist für sie interessant?



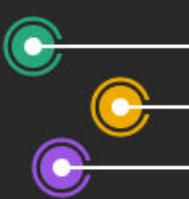
[cedricscherer.com](http://cedricscherer.com)



@CedScherer



z3tt



# Wer ist mein Zielpublikum?

- 🌟 Welche Geschichte ist für sie interessant?
- 🤔 Was sind wirklich relevante Details?



# Wer ist mein Zielpublikum?

- 🌟 Welche Geschichte ist für sie interessant?
- 🤔 Was sind wirklich relevante Details?
- 🤓 Welche Variablen sind für sie bedeutsam?



# Wer ist mein Zielpublikum?

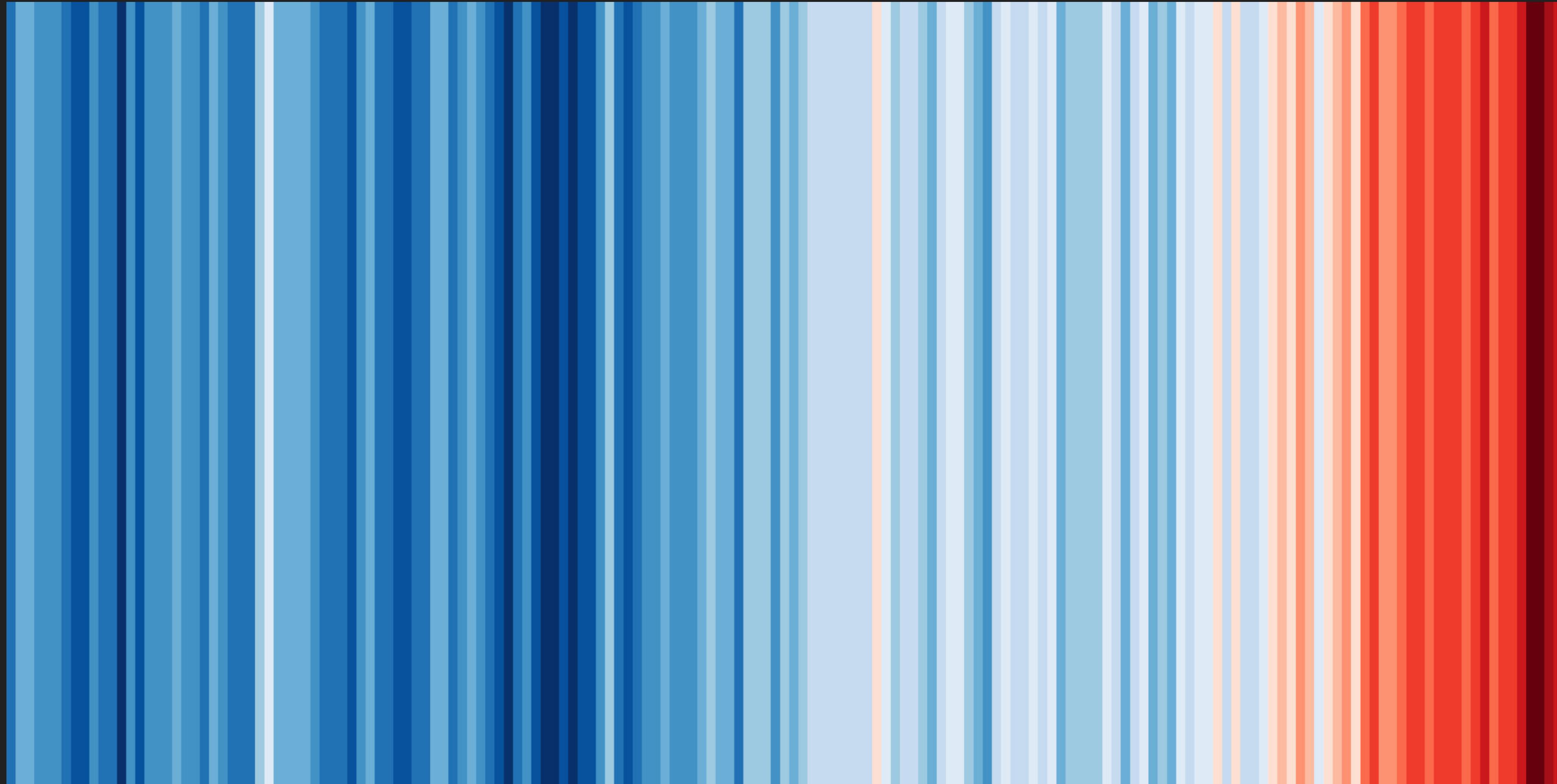
- 🌟 Welche Geschichte ist für sie interessant?
- 🤔 Was sind wirklich relevante Details?
- 🤓 Welche Variablen sind für sie bedeutsam?
- 👀 Wie und wo werden sie der Grafik begegnen?



# Wer ist mein Zielpublikum?

- 🤩 Welche Geschichte ist für sie interessant?
- 🤔 Was sind wirklich relevante Details?
- 🤓 Welche Variablen sind für sie bedeutsam?
- 👀 Wie und wo werden sie der Grafik begegnen?
- 😱 Brauche ich überhaupt eine Grafik??

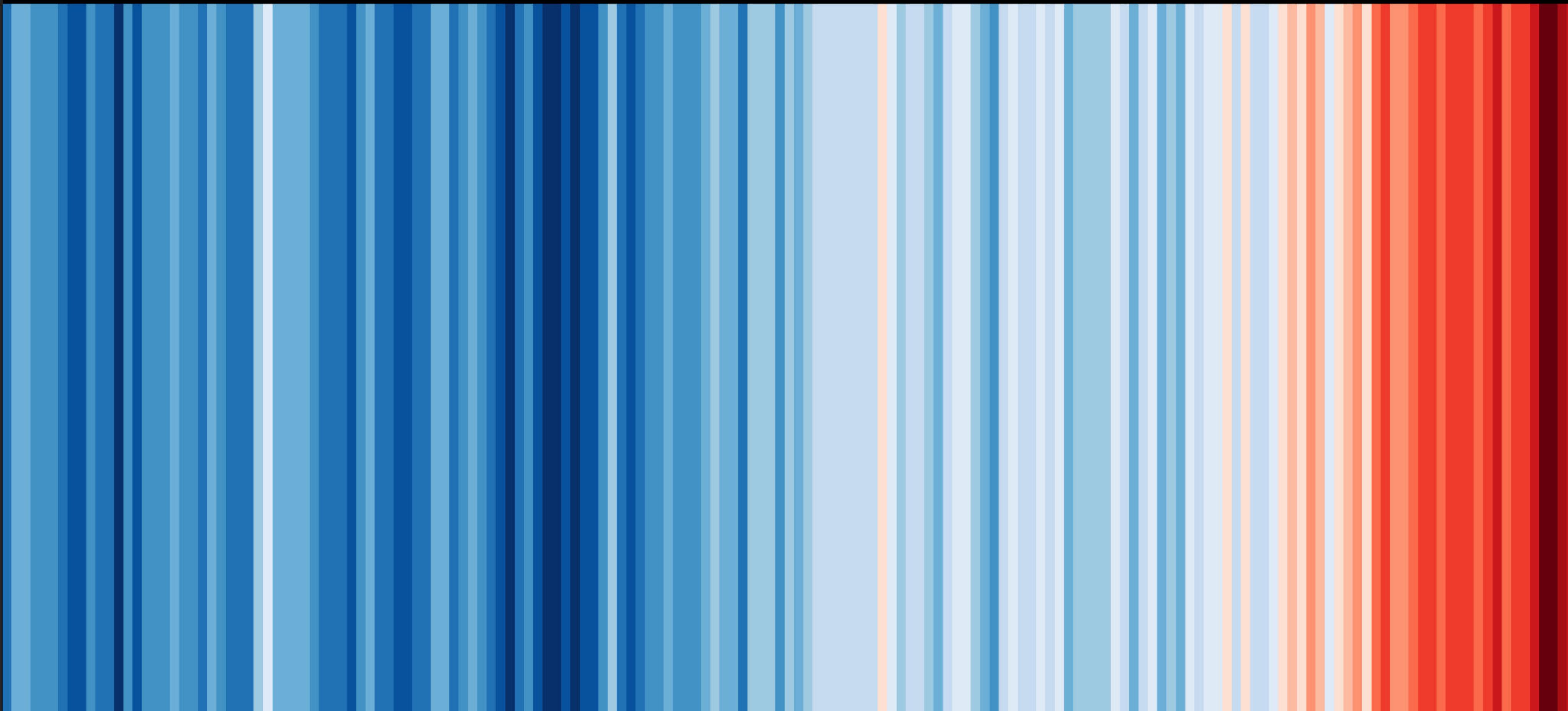




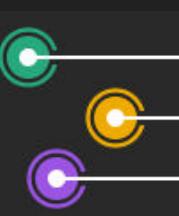
***Warming Stripes*** von Ed Hawkins



# Global temperature change (1850-2019)



*Warming Stripes* von 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.

1

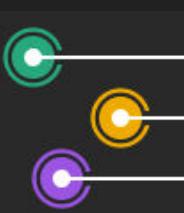
0

[showyourstripes.info/faq](http://showyourstripes.info/faq)

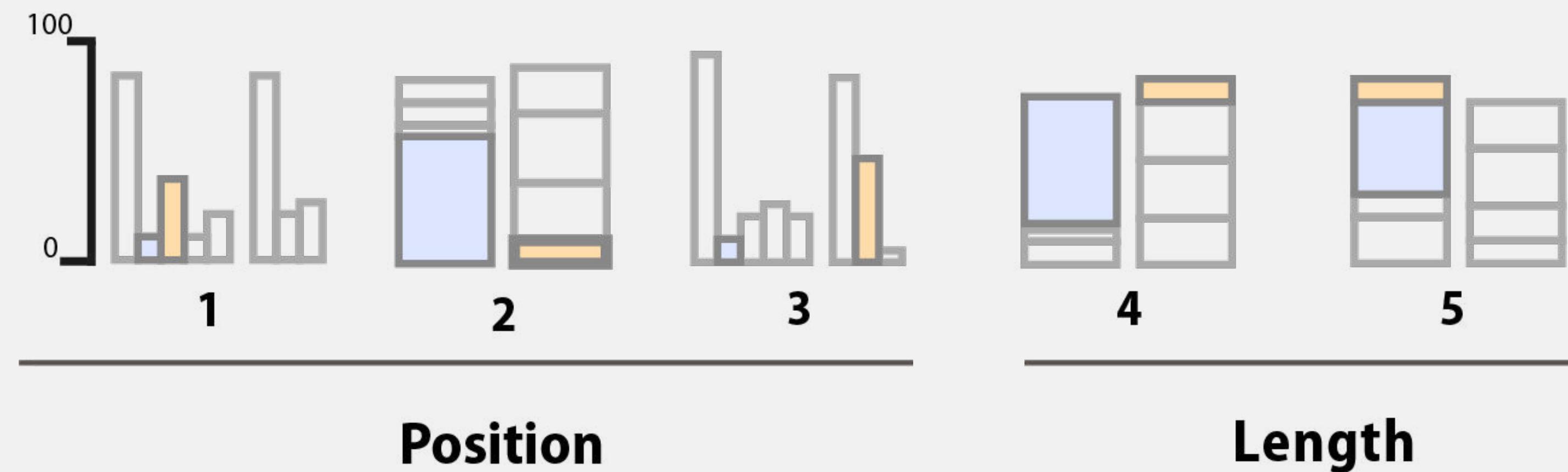
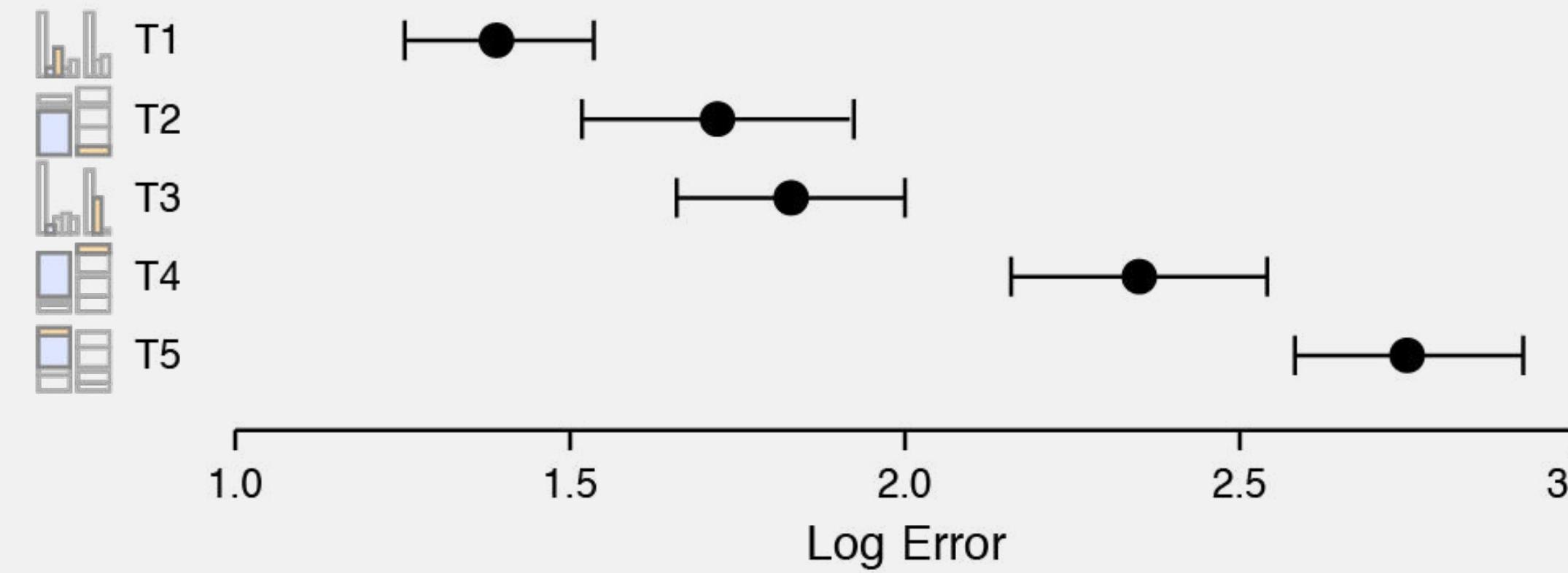


Diese Grafiken sind speziell dafür  
**konzipiert um Gespräche anzuregen**  
über die sich erwärmende Welt und  
die Risiken des Klimawandels.

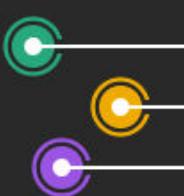
» These graphics are specifically designed to be as simple as possible, and to start conversations about our warming world and the risks of climate change. There are numerous sources of information which provide more specific details about how temperatures have changed, so these graphics fill a gap and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



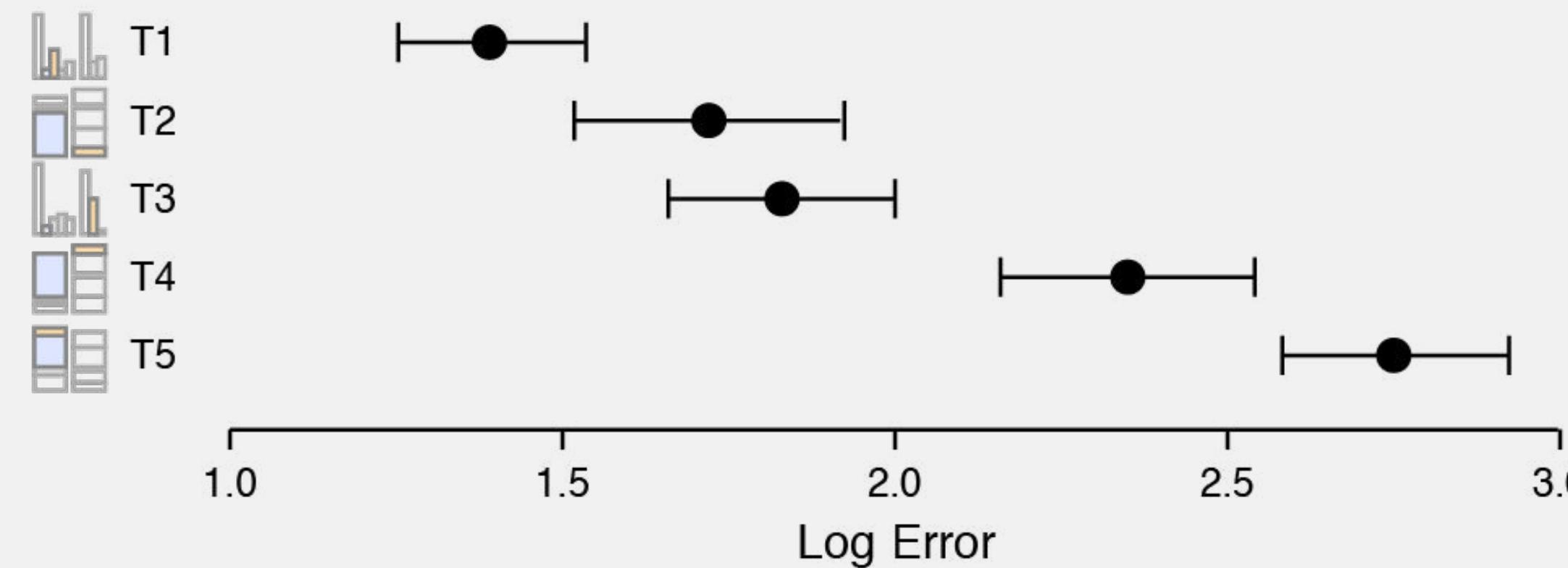
### Cleveland & McGill's Results



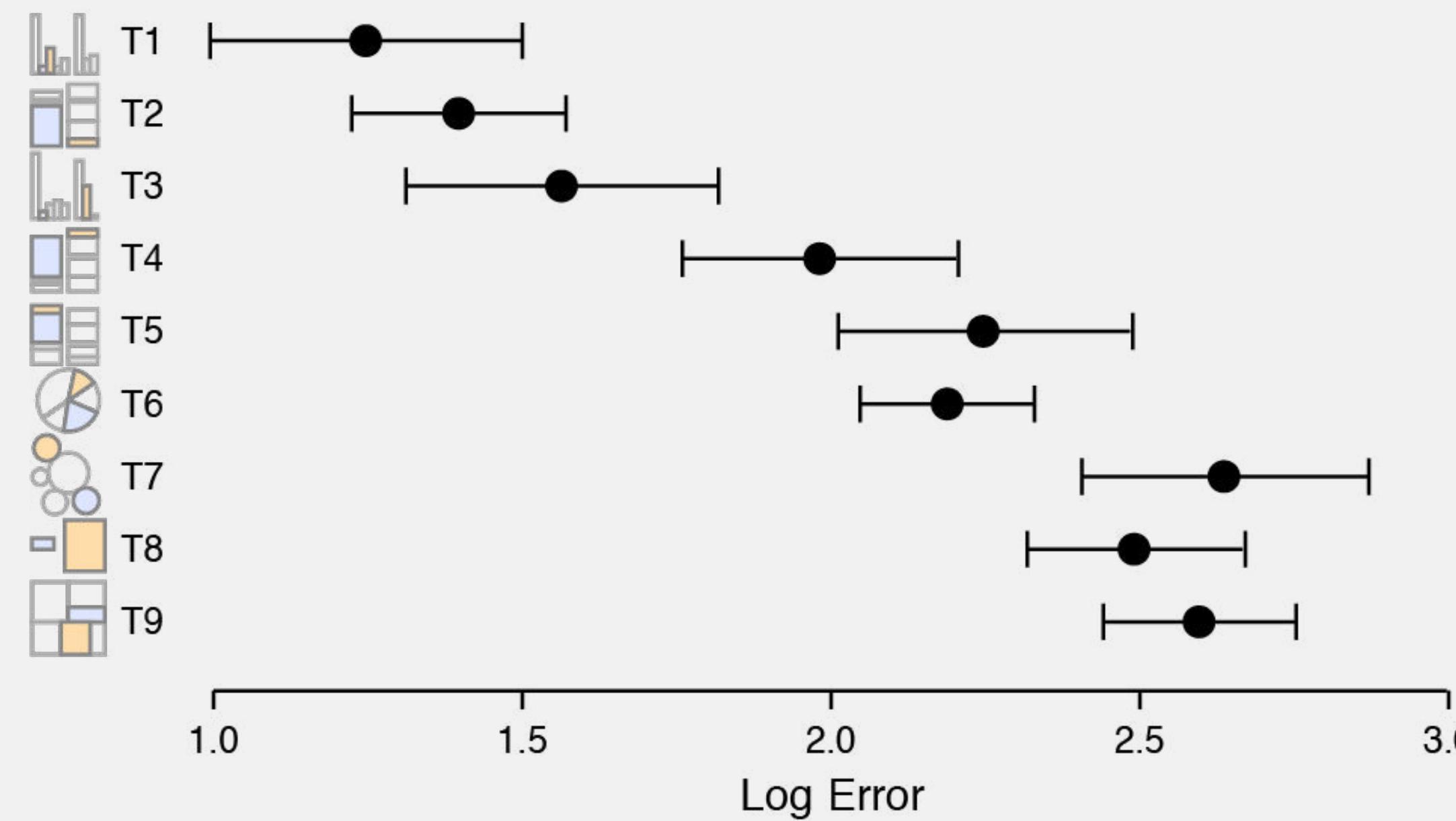
Quelle: Kieran Healy's ["Data Visualization: A Practical Introduction"](#)  
Ergebnisse basierend auf Heer & Bostock sowie Cleveland & McGill



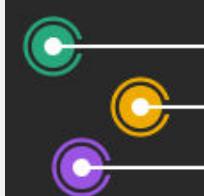
### Cleveland & McGill's Results



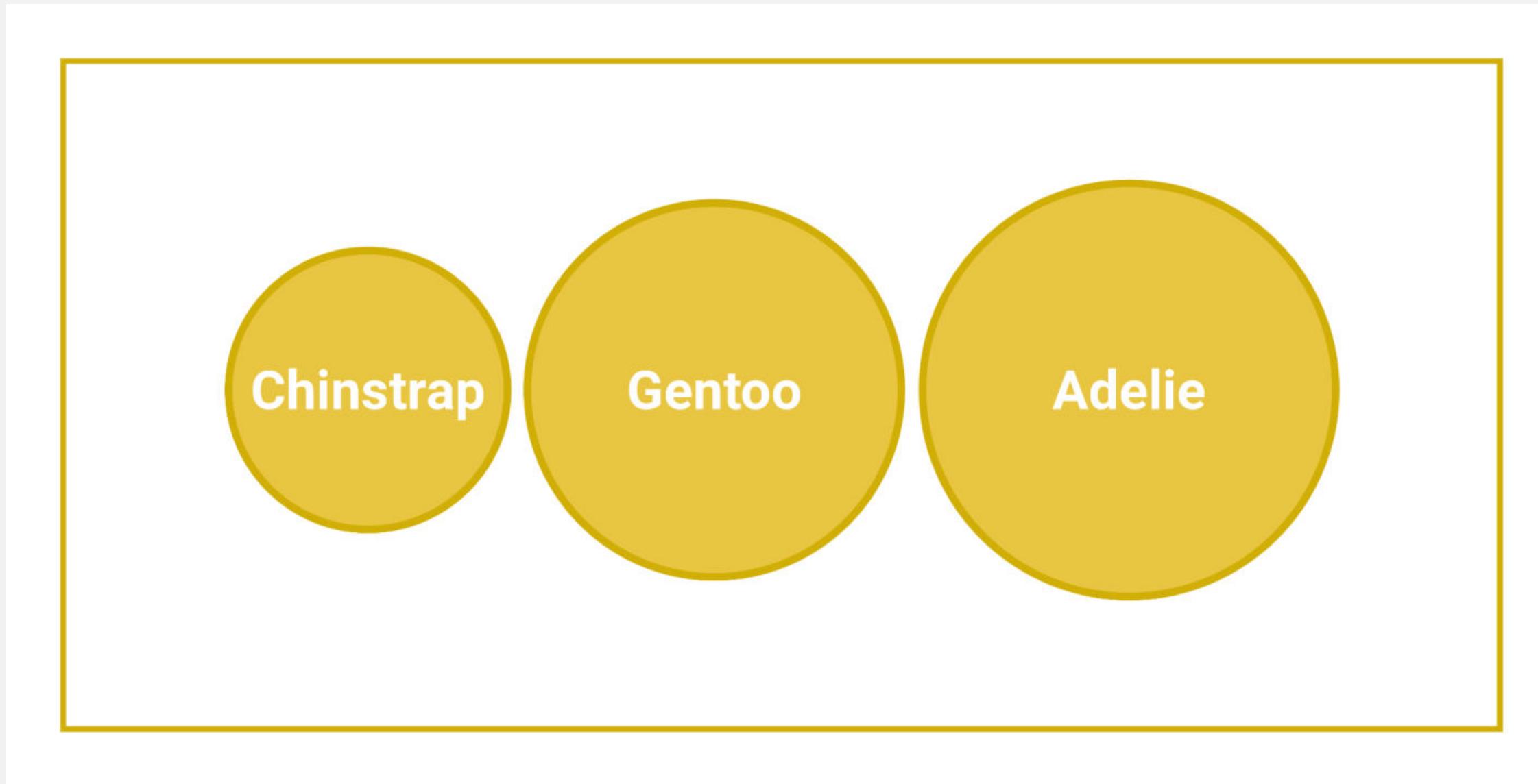
### Crowdsourced Results



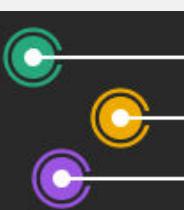
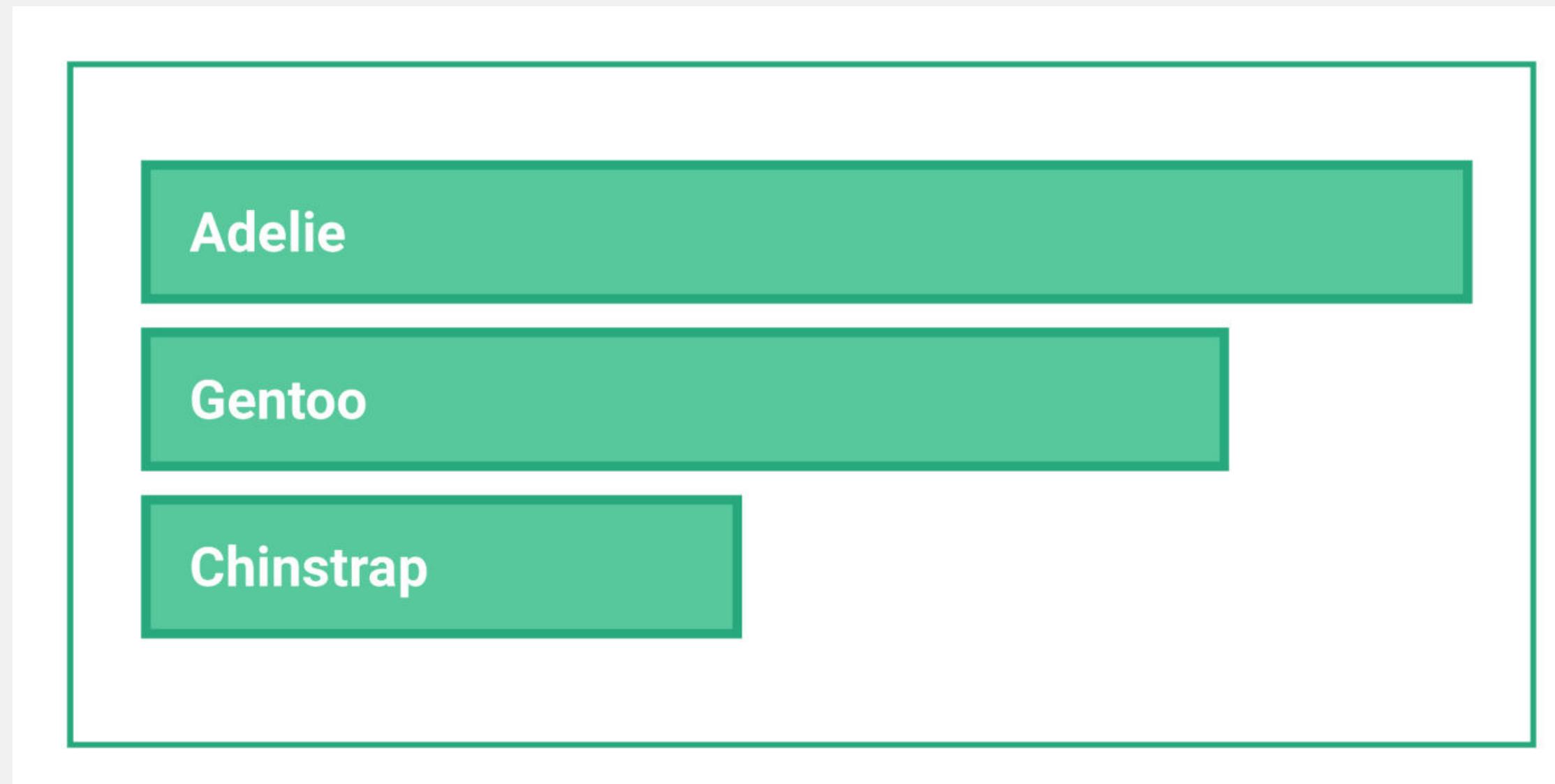
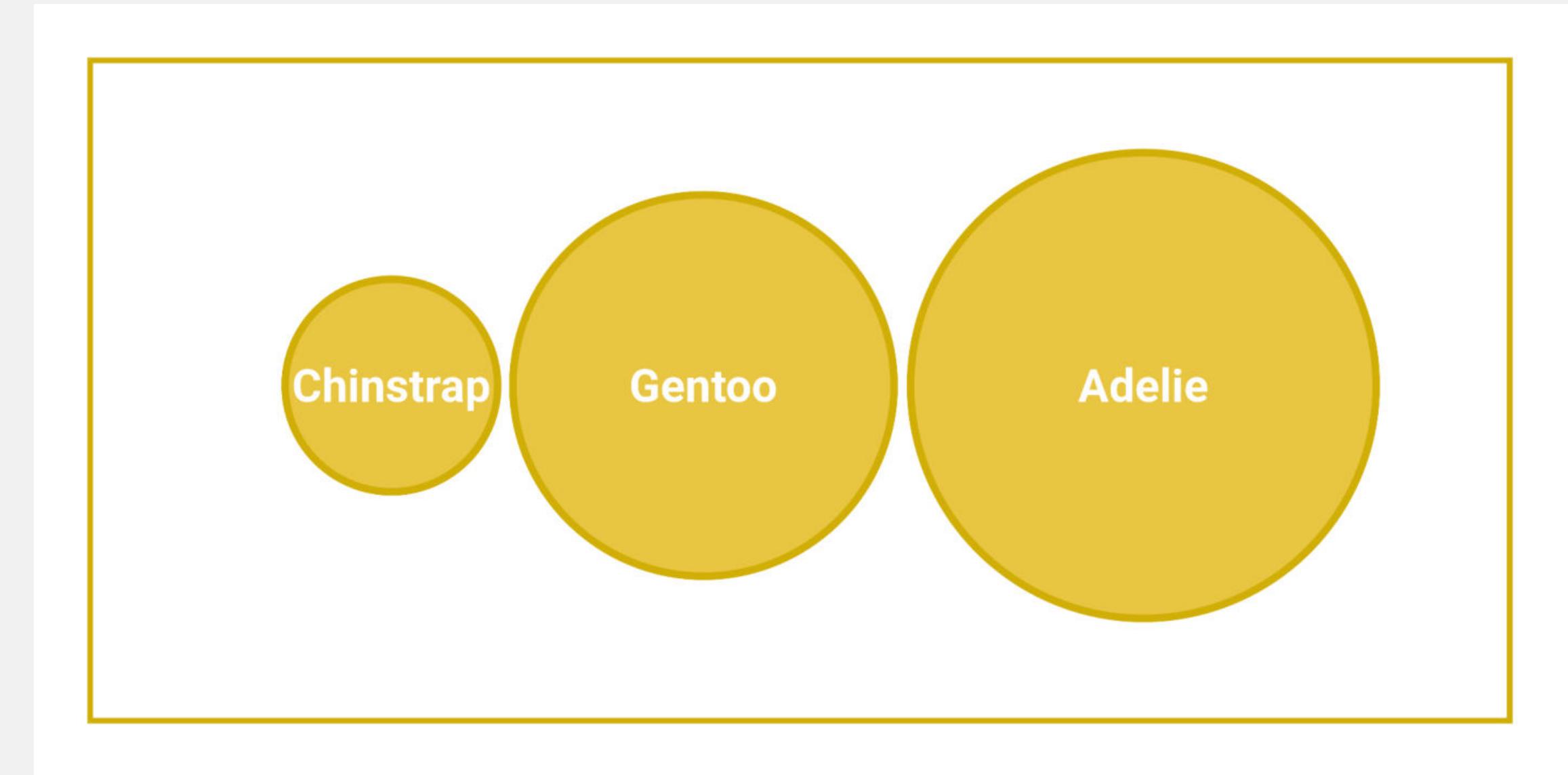
Quelle: Kieran Healy's [Data Visualization: A Practical Introduction](#)  
Ergebnisse basierend auf Heer & Bostock sowie Cleveland & McGill



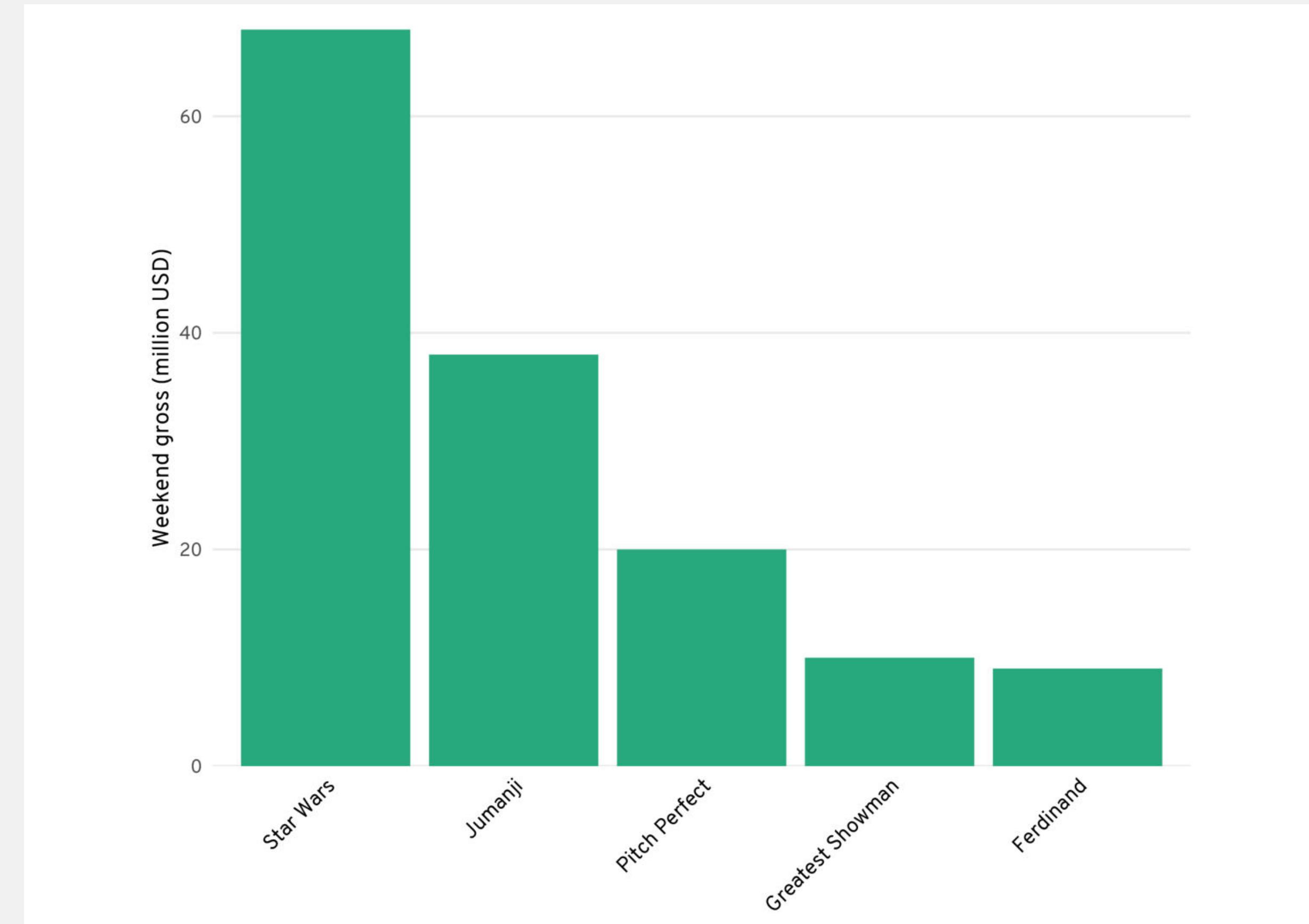
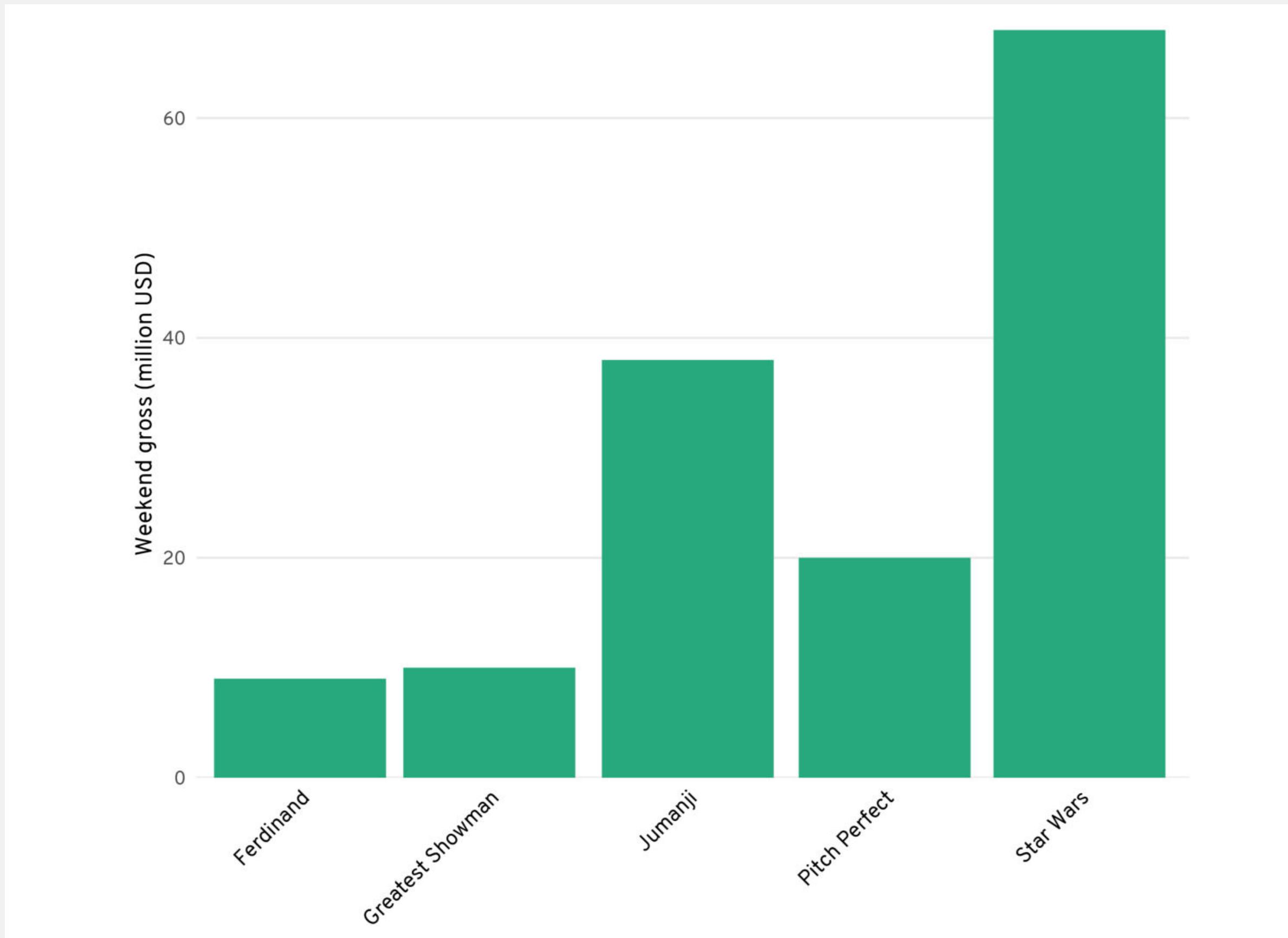
# Nutze immer Fläche.



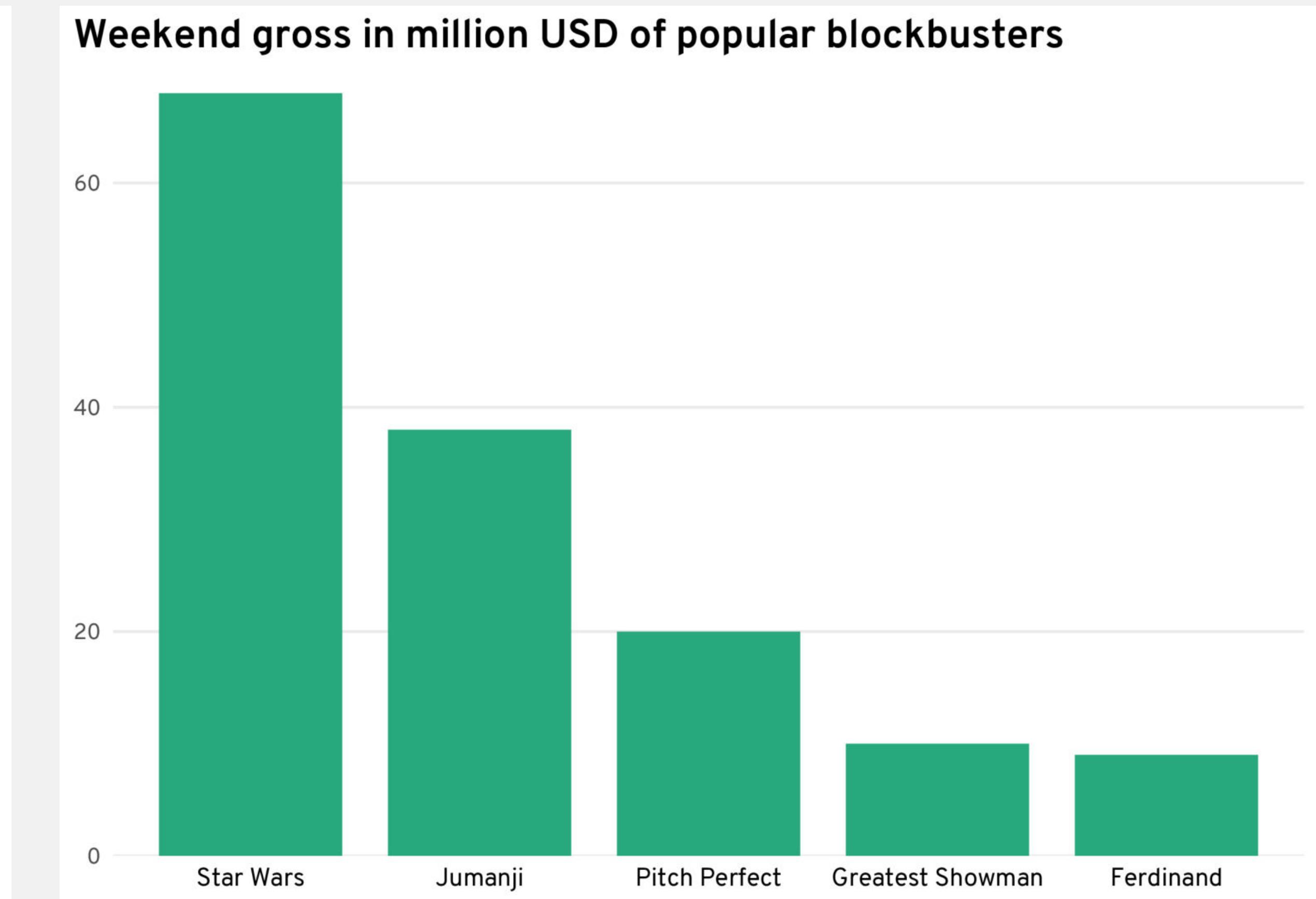
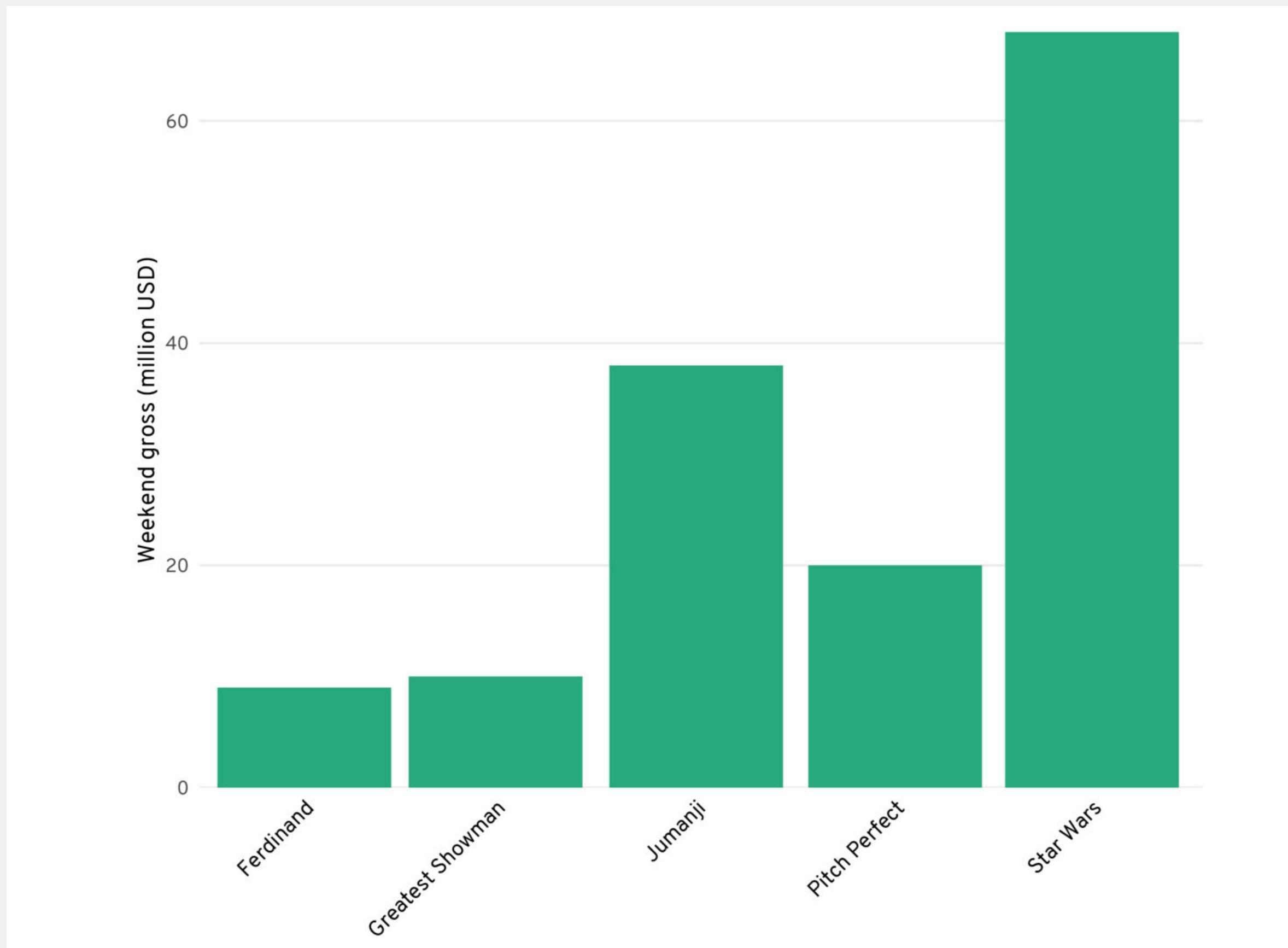
# Nutze niemals Radius!



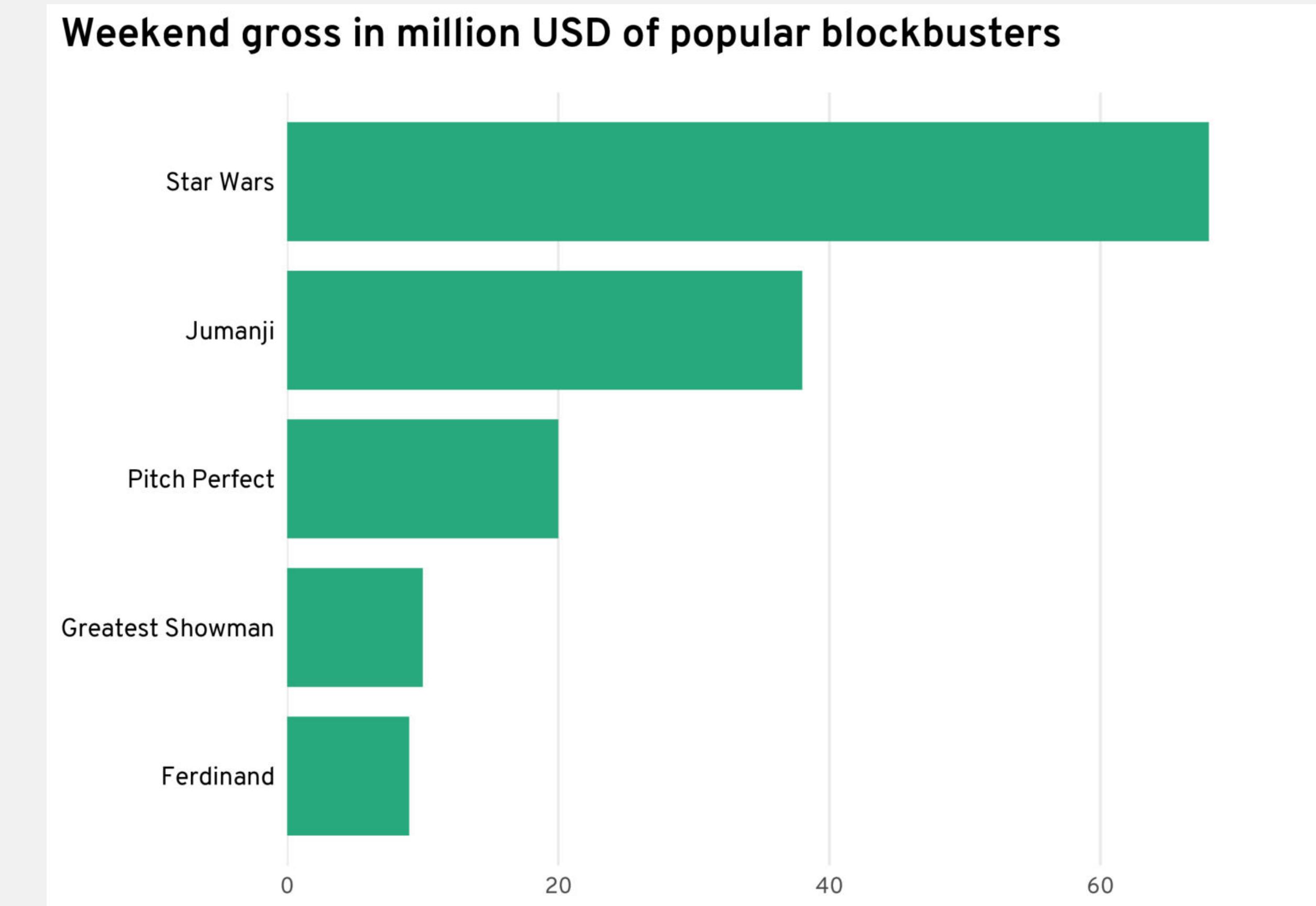
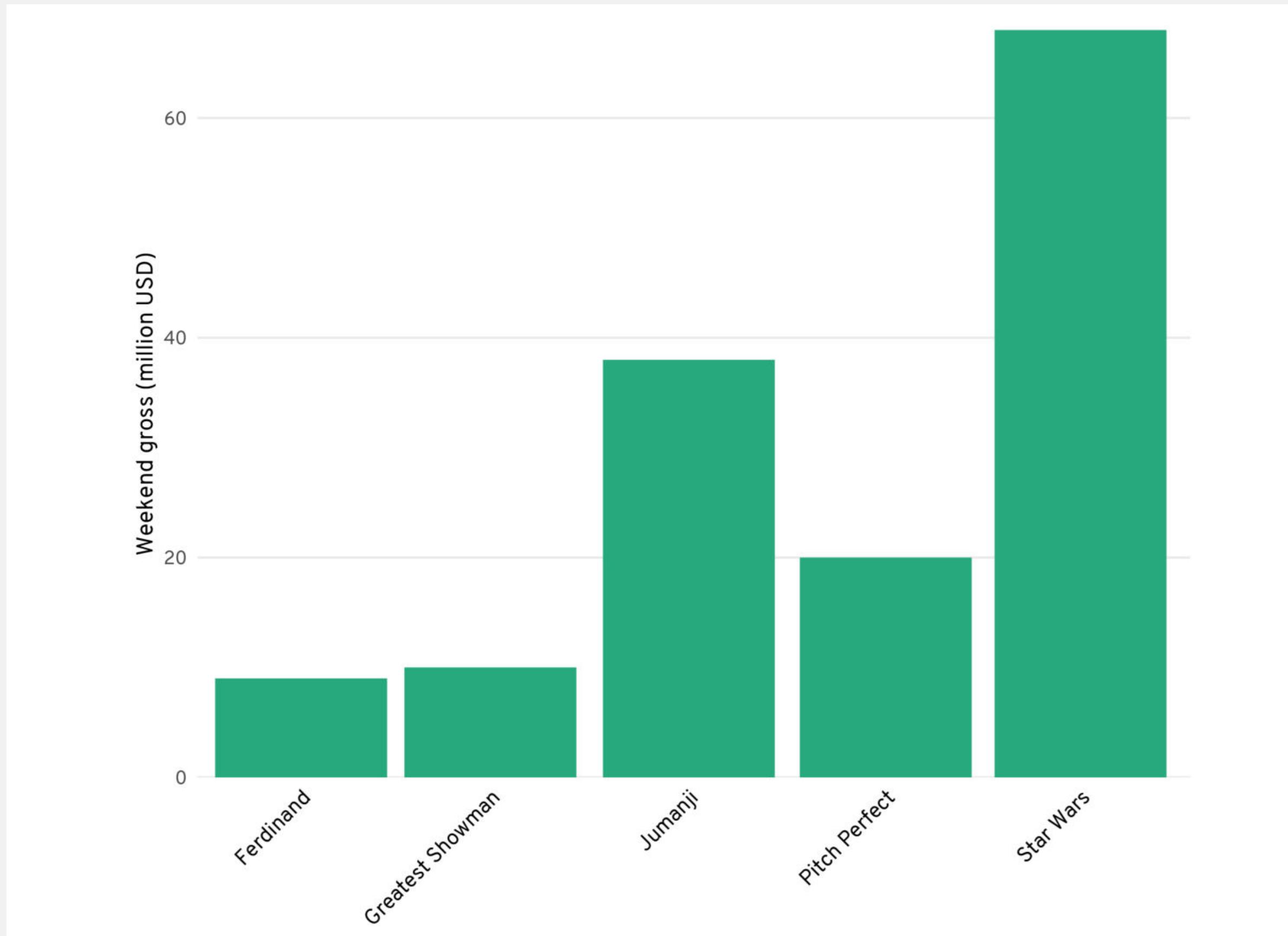
# Daten sortieren



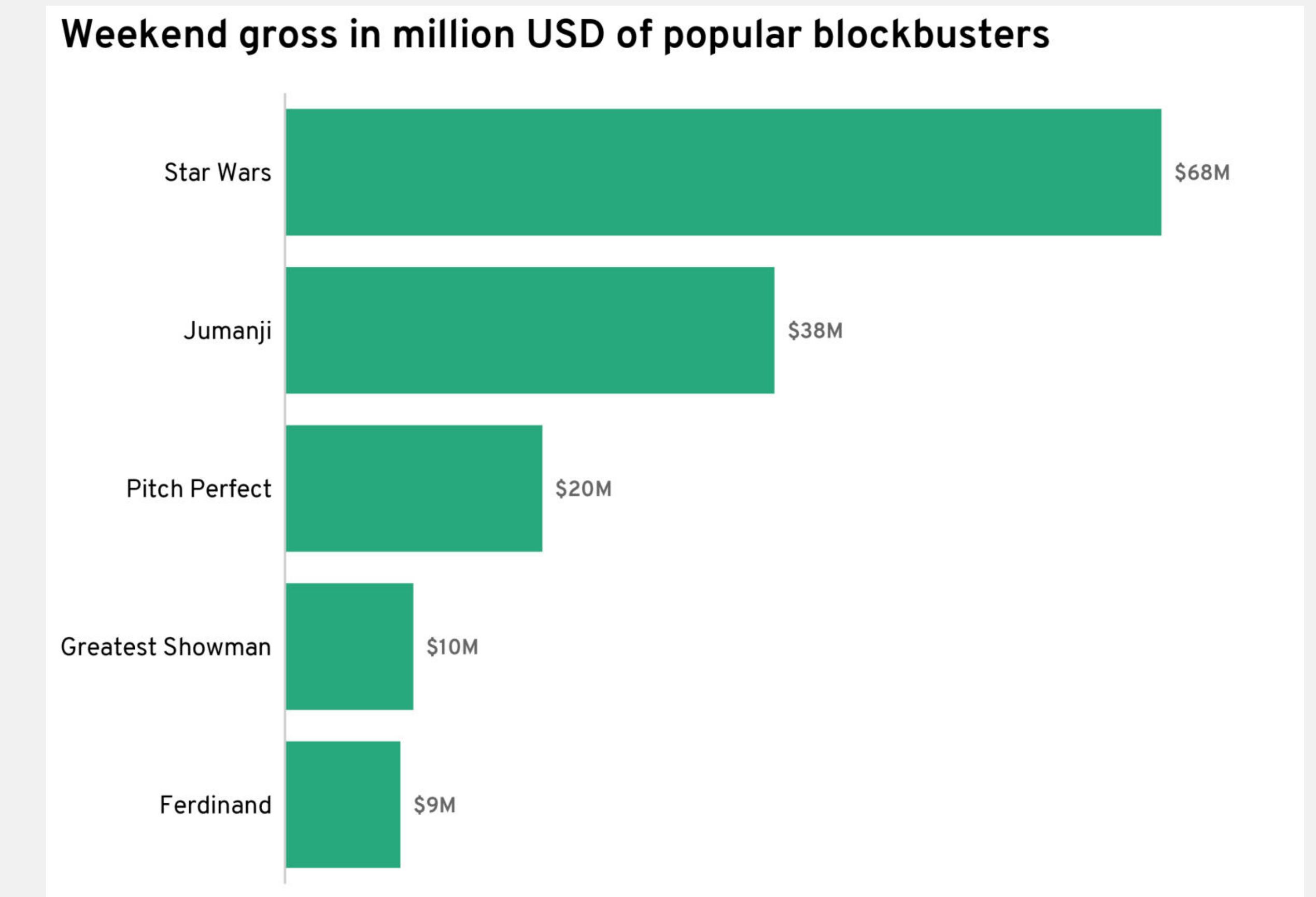
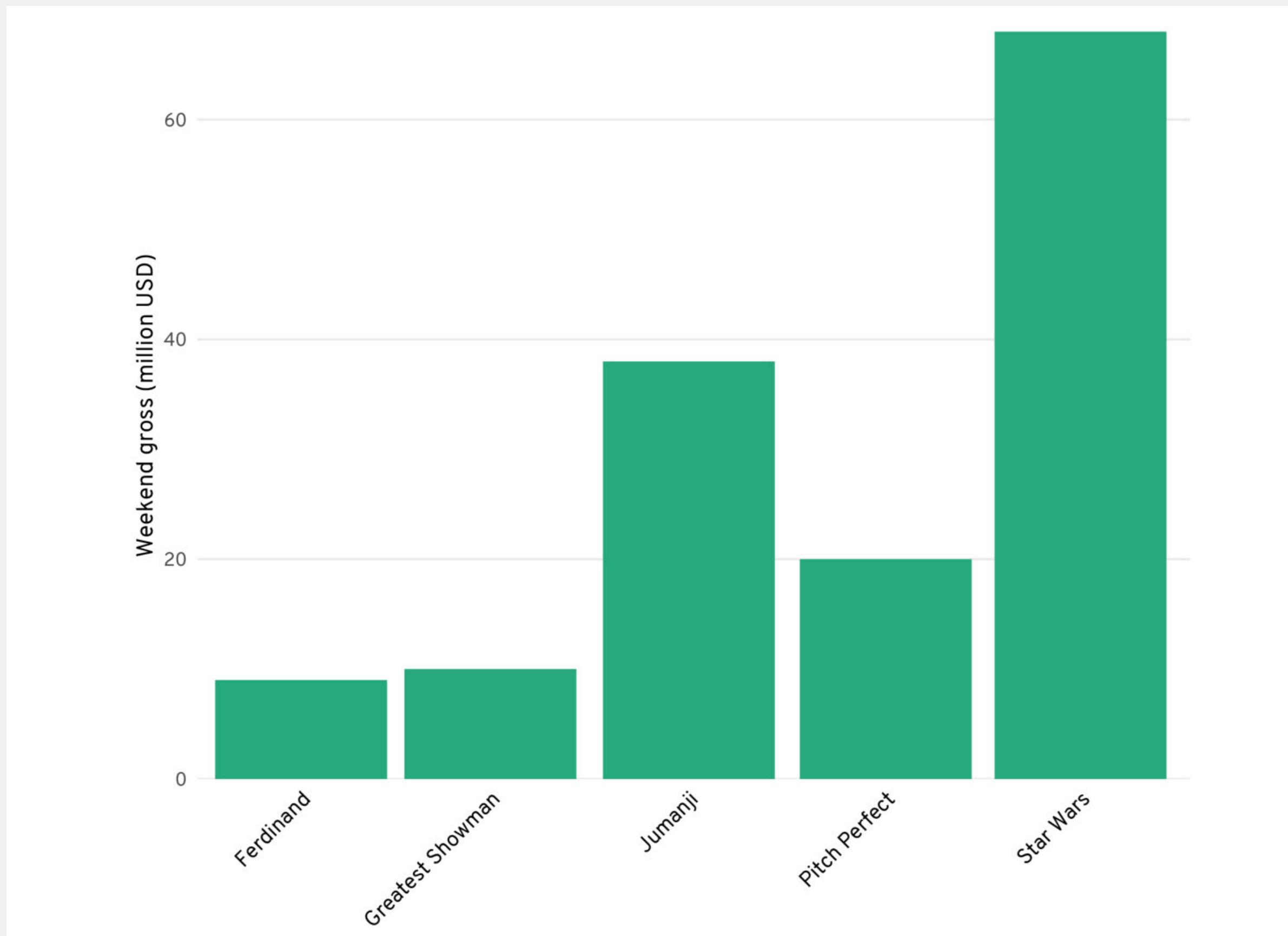
# Textdrehung vermeiden



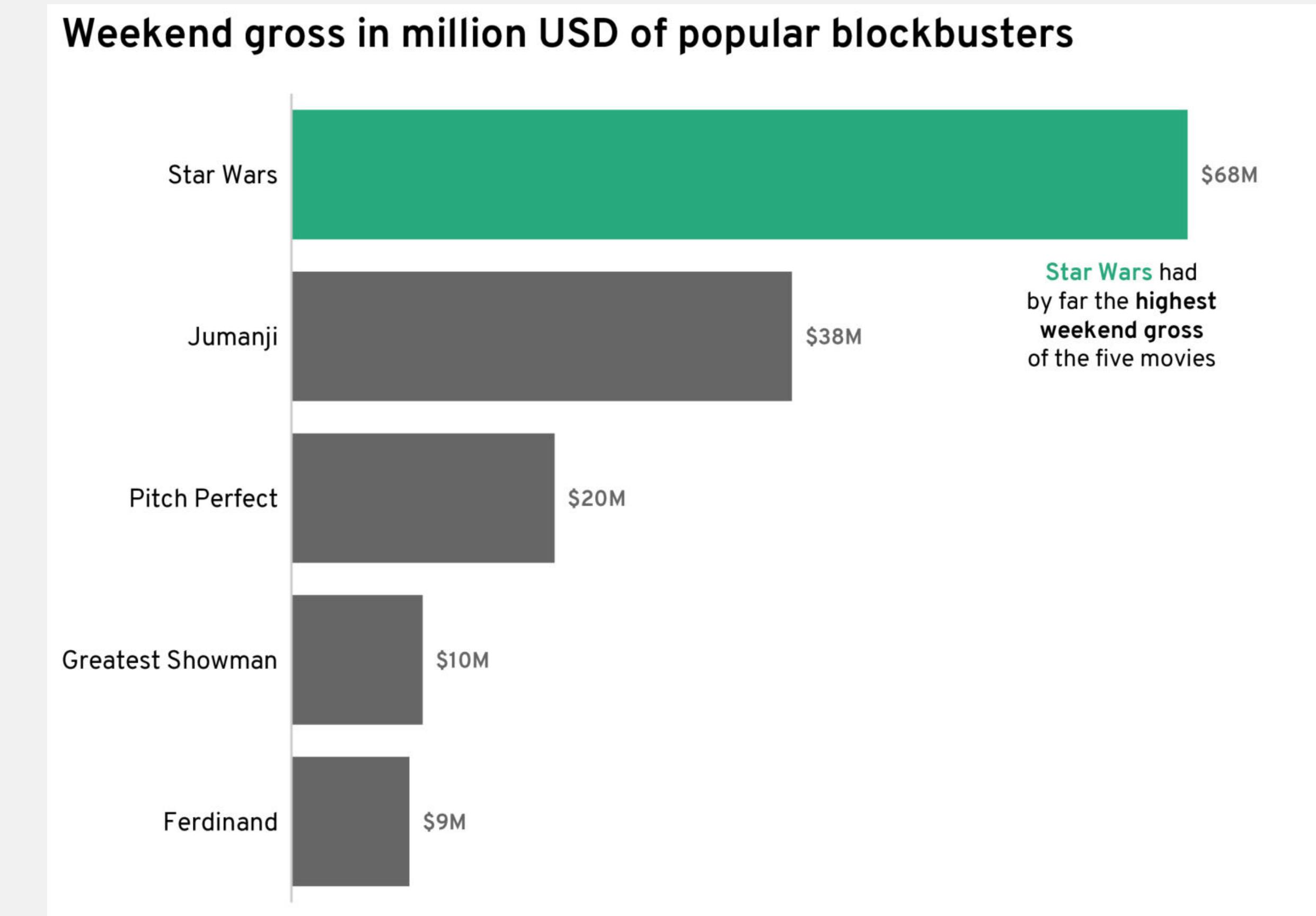
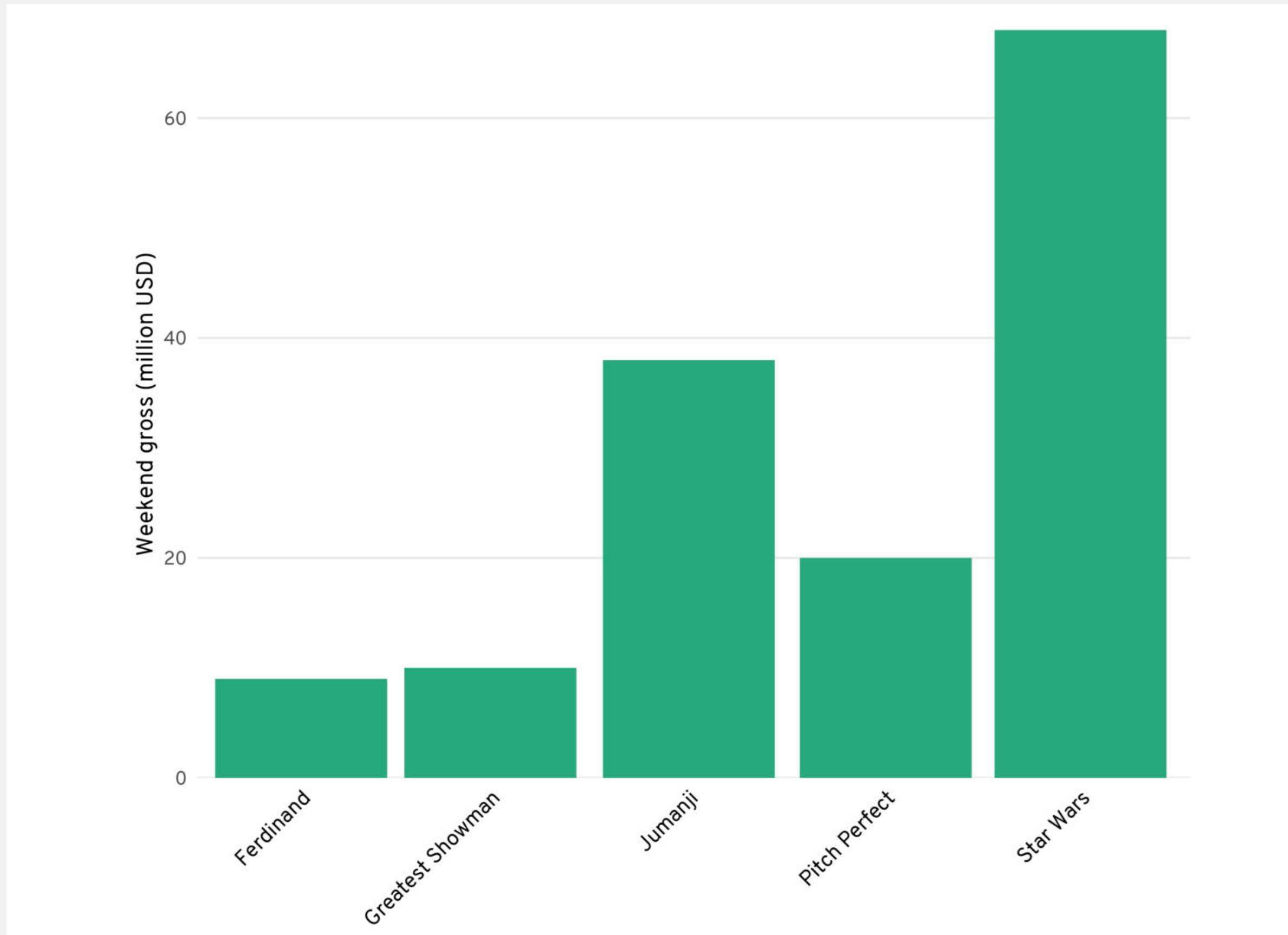
# Textdrehung vermeiden



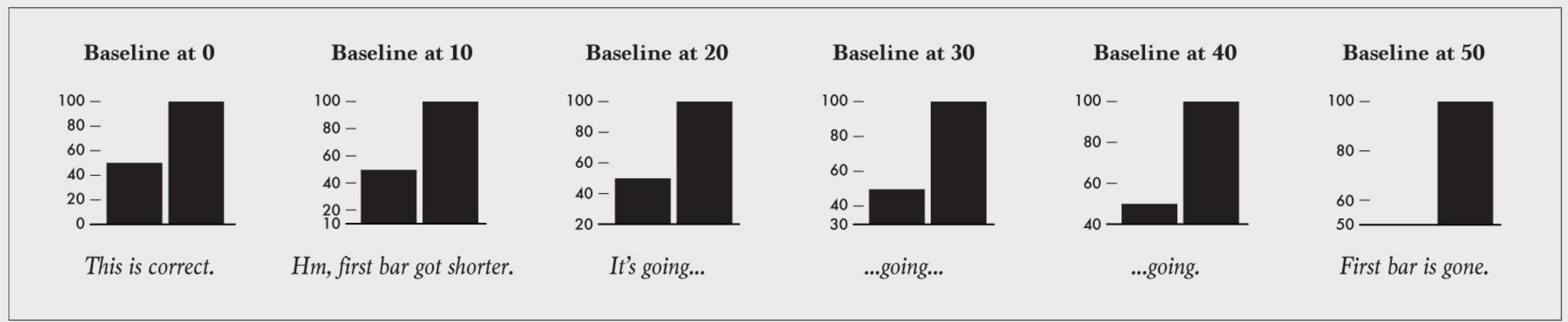
# Direktbeschriftungen hinzufügen



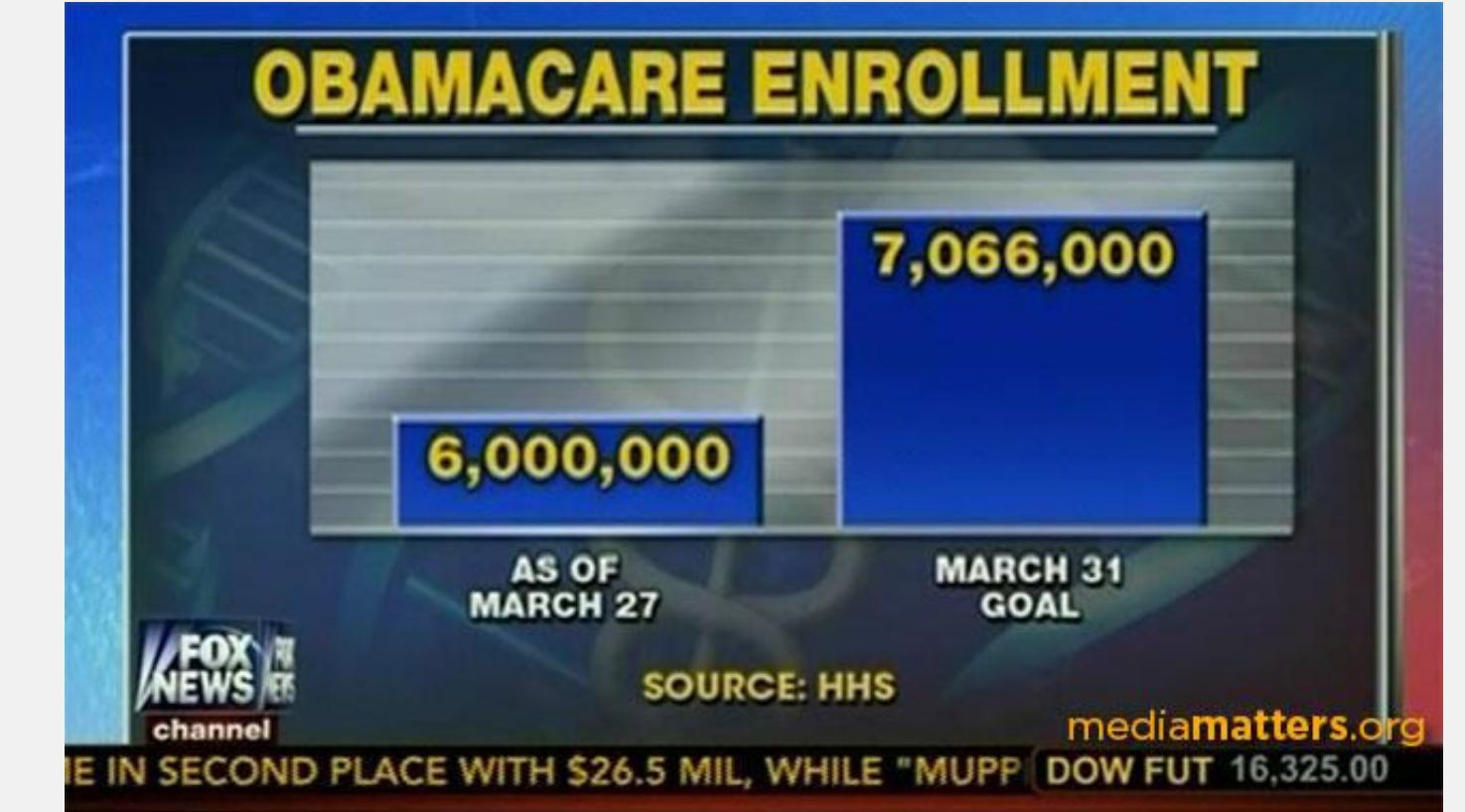
# Highlights setzen

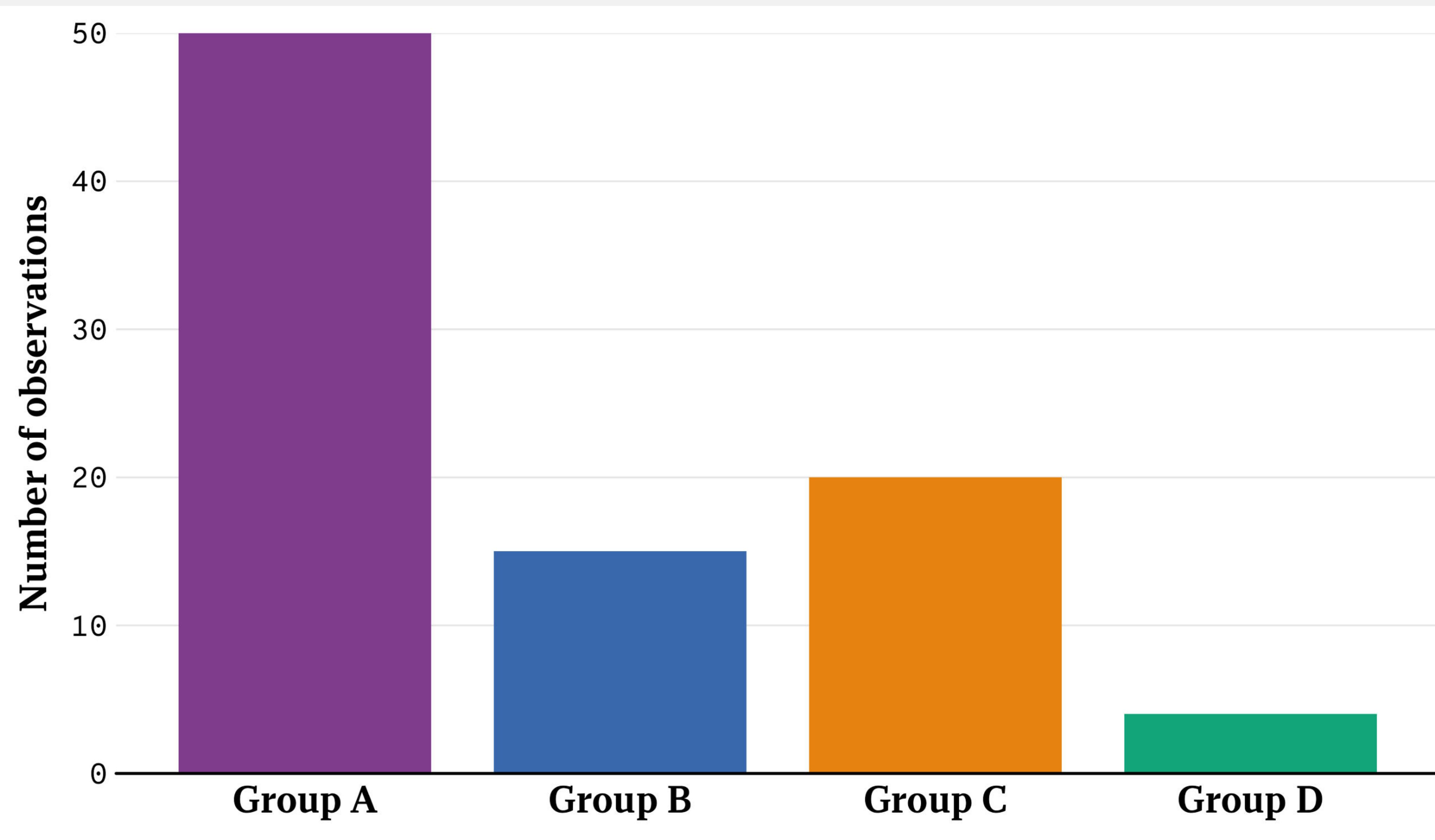


# Immer bei „Null“ starten



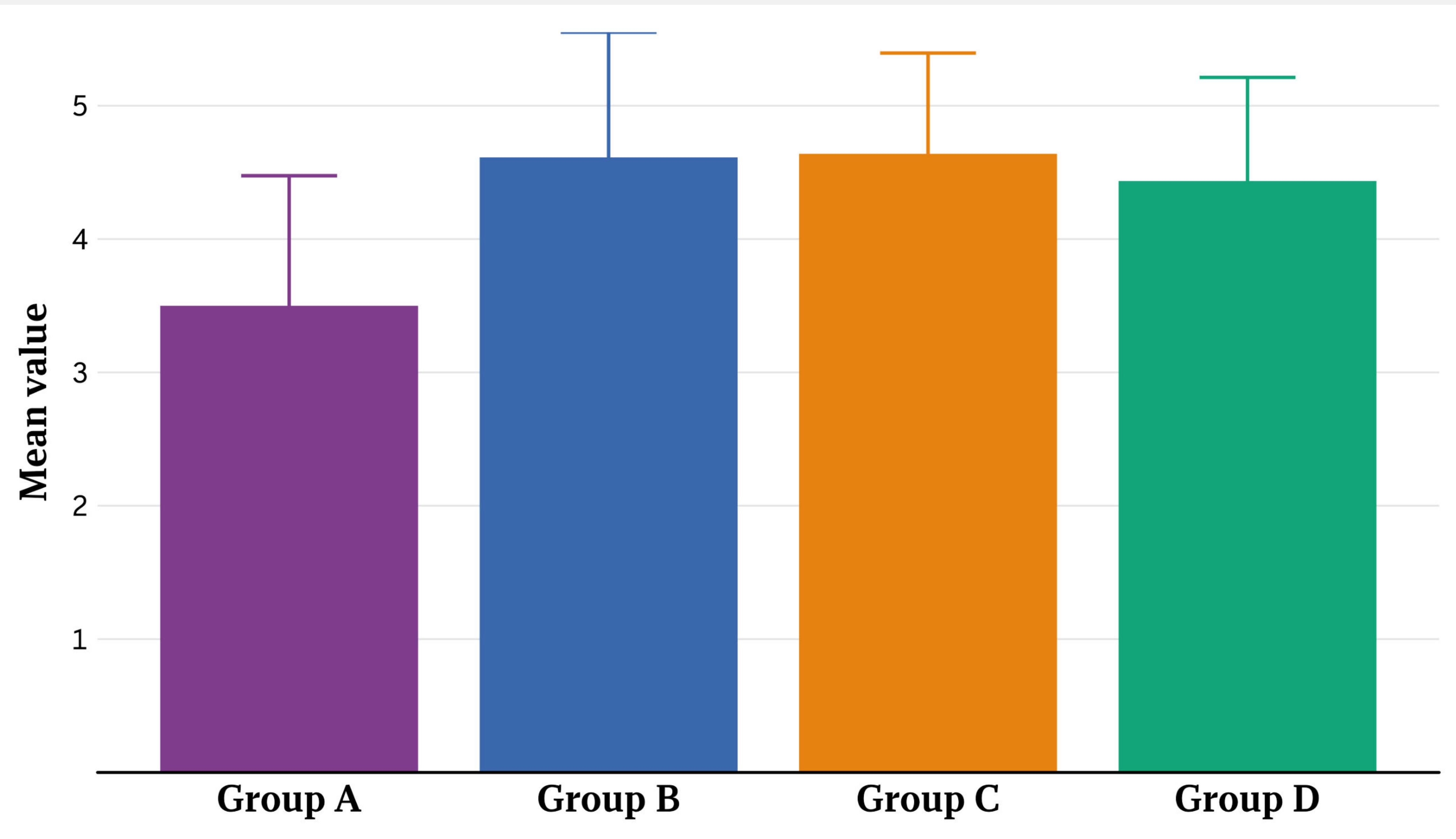
Quelle: Nathan Yau ([flowingdata.com](http://flowingdata.com))



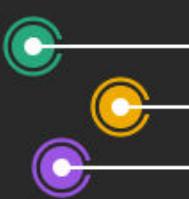


Mein Webinar für USGS Data Science “Beyond Bar and Box Plots” [[Slides](#) | [Recording](#)]





Mein Webinar für USGS Data Science “Beyond Bar and Box Plots” [[Slides](#) | [Recording](#)]



PERSPECTIVE

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

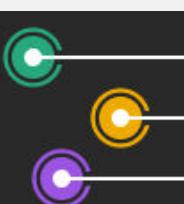
**Tracey L. Weissgerber<sup>1\*</sup>, Nataša M. Milic<sup>1,2</sup>, Stacey J. Winham<sup>3</sup>, Vesna D. Garovic<sup>1</sup>**

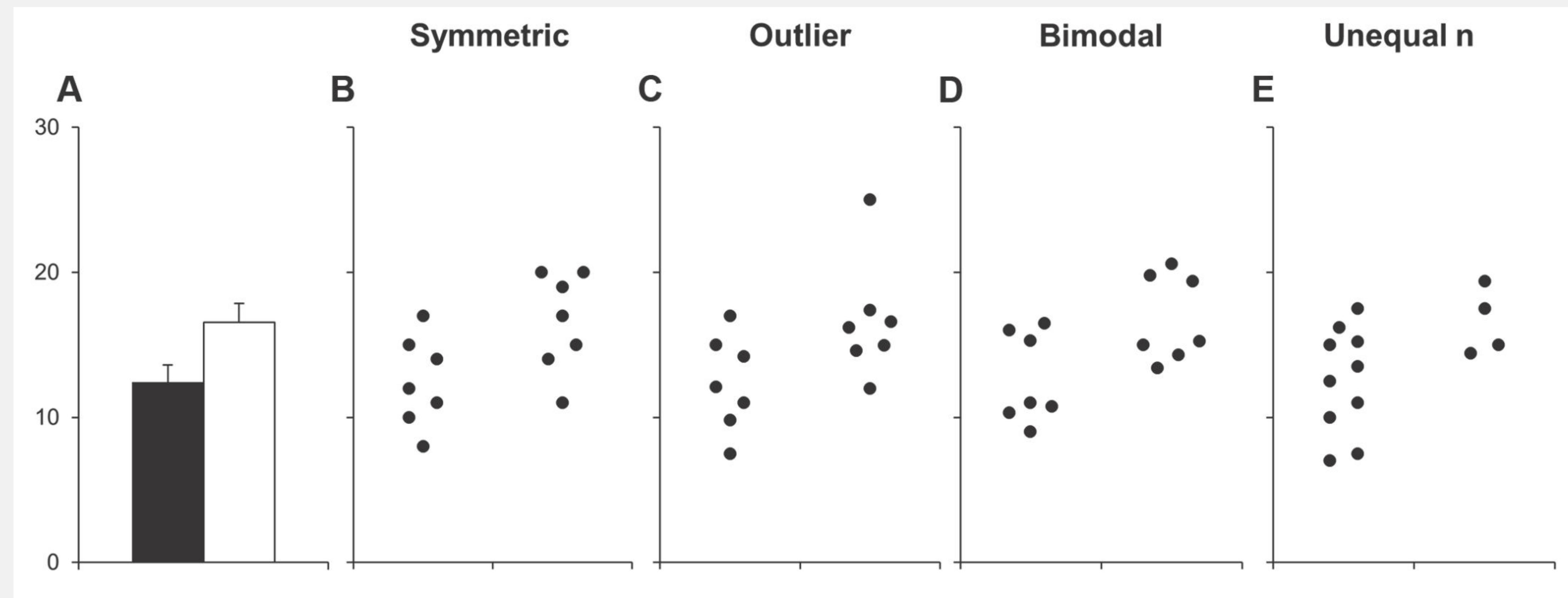
**1** Division of Nephrology & Hypertension, Mayo Clinic, Rochester, Minnesota, United States of America,

**2** Department of Biostatistics, Medical Faculty, University of Belgrade, Belgrade, Serbia, **3** Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota, United States of America

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

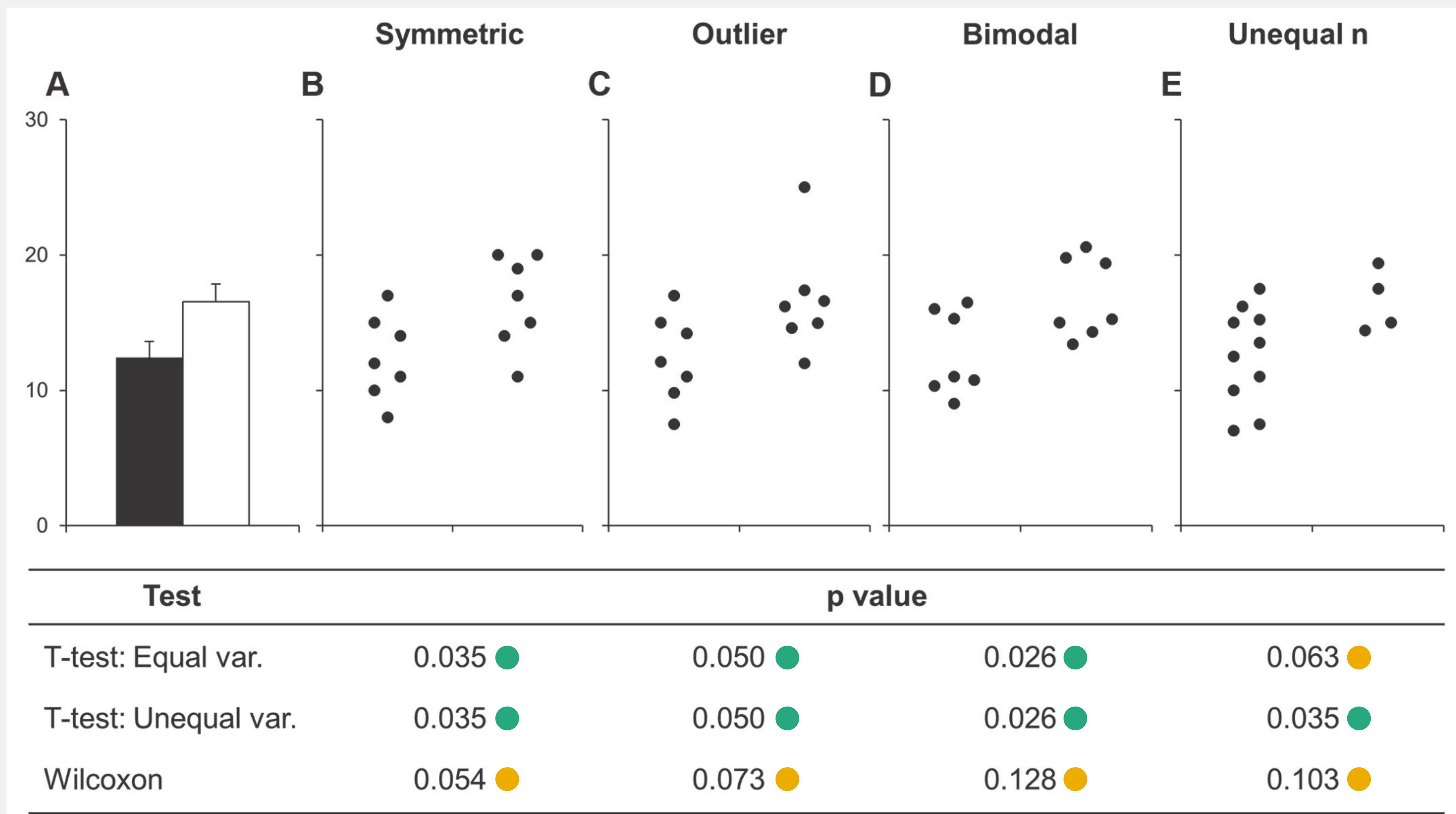
*Weissgerber et al. (2015) PLoS Biology*



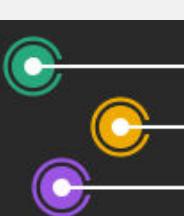


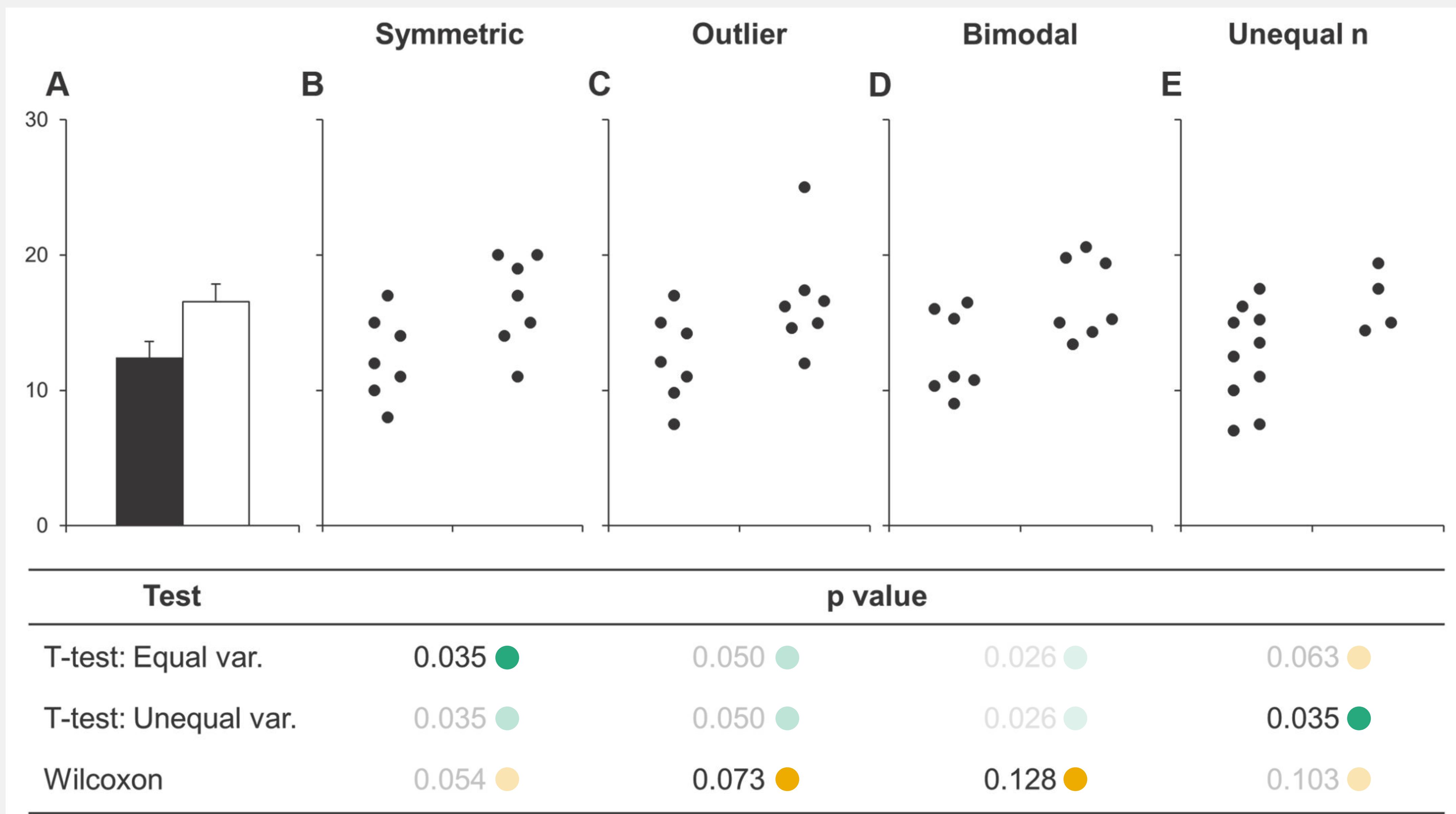
Weissgerber et al. (2015) PLoS Biology (modifiziert)



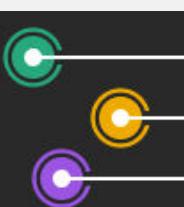


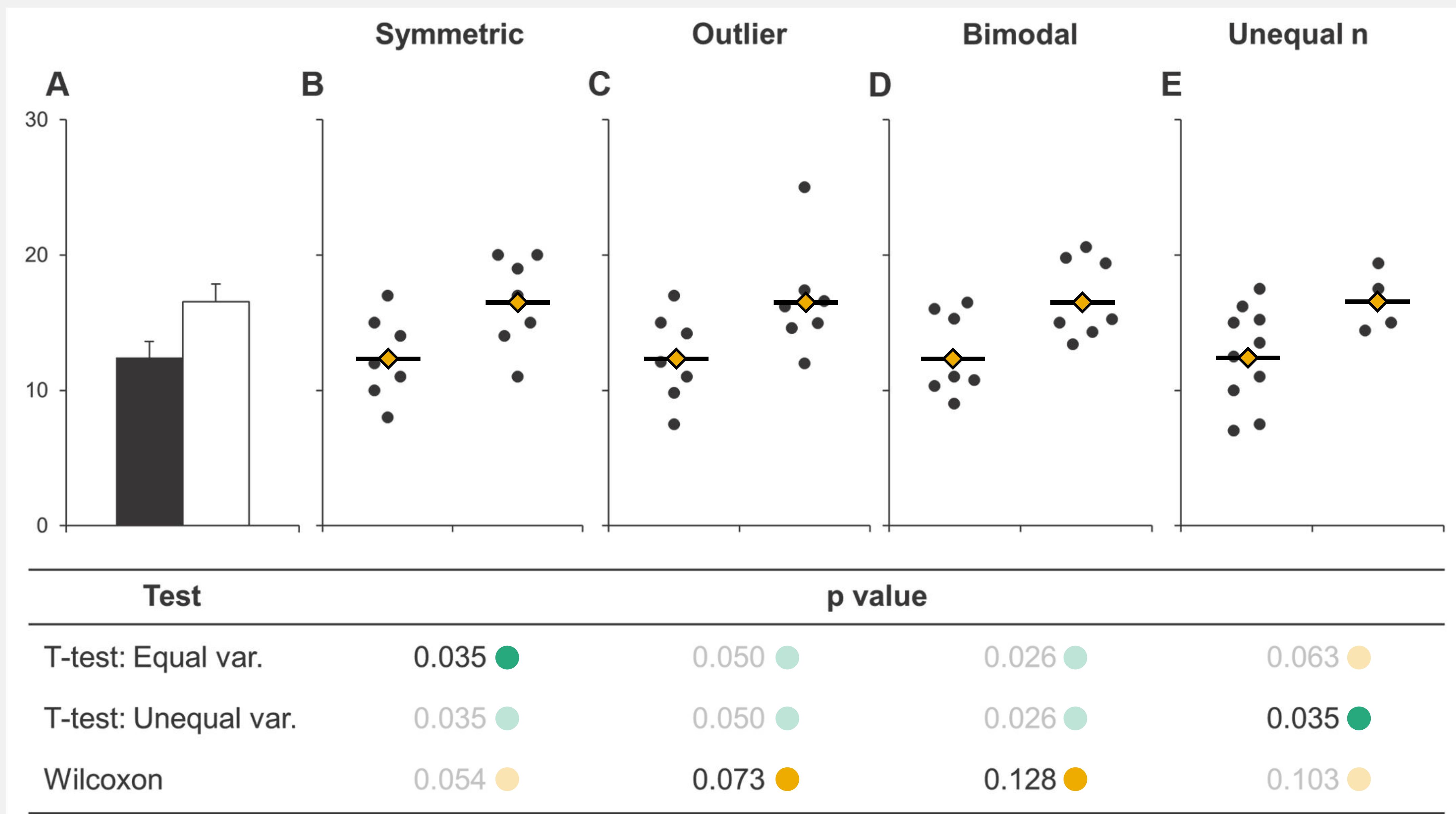
Weissgerber et al. (2015) PLoS Biology (modifiziert)



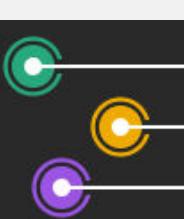


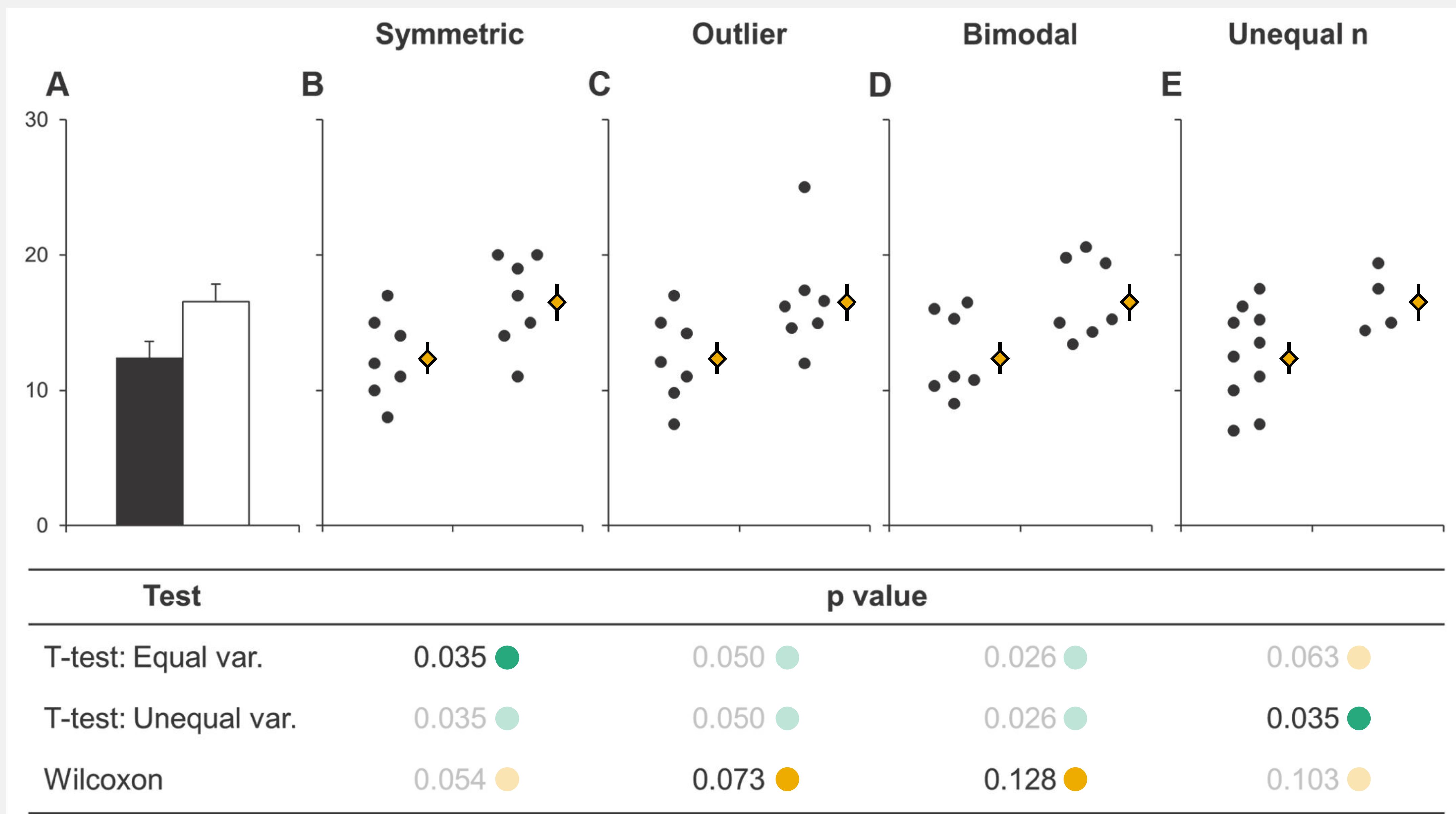
Weissgerber et al. (2015) PLoS Biology (modifiziert)



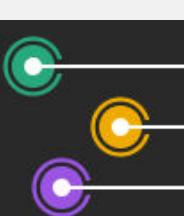


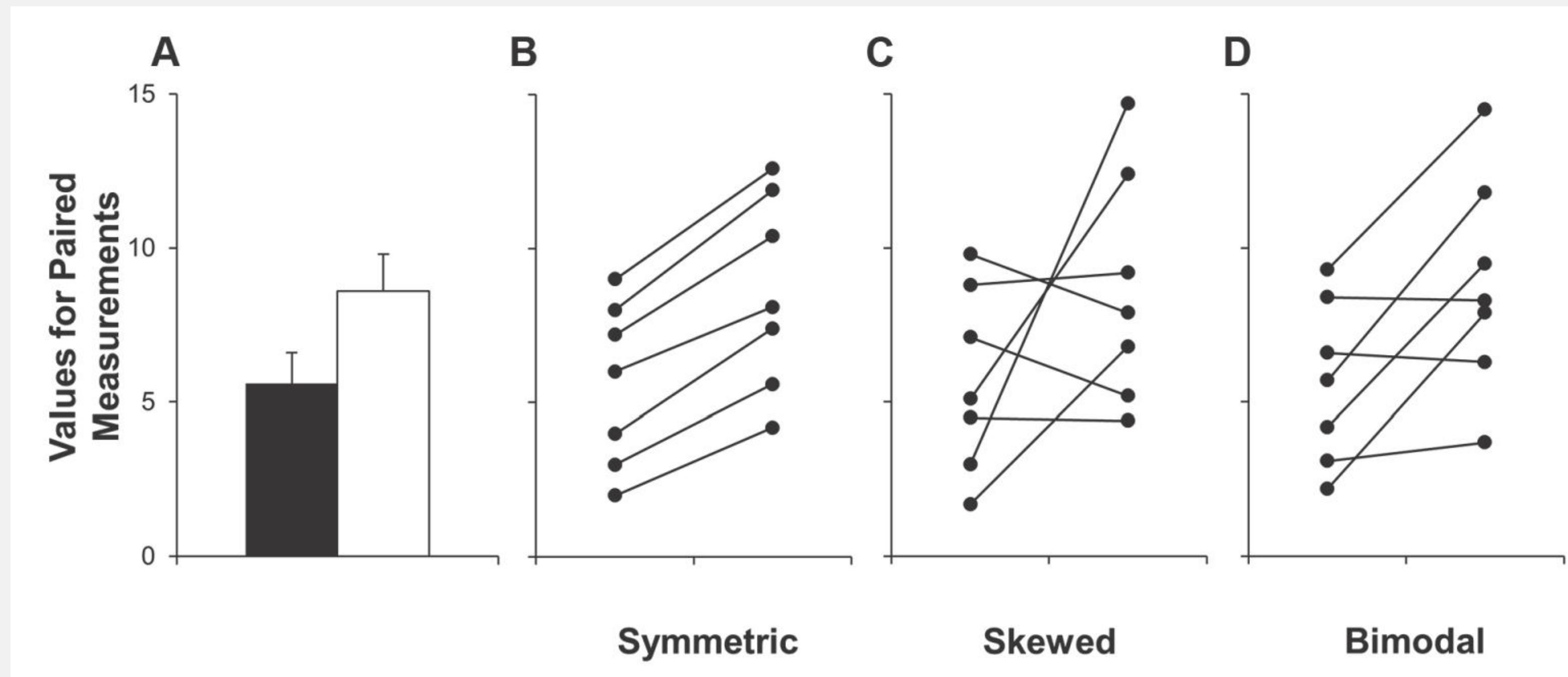
Weissgerber et al. (2015) PLoS Biology (modifiziert)





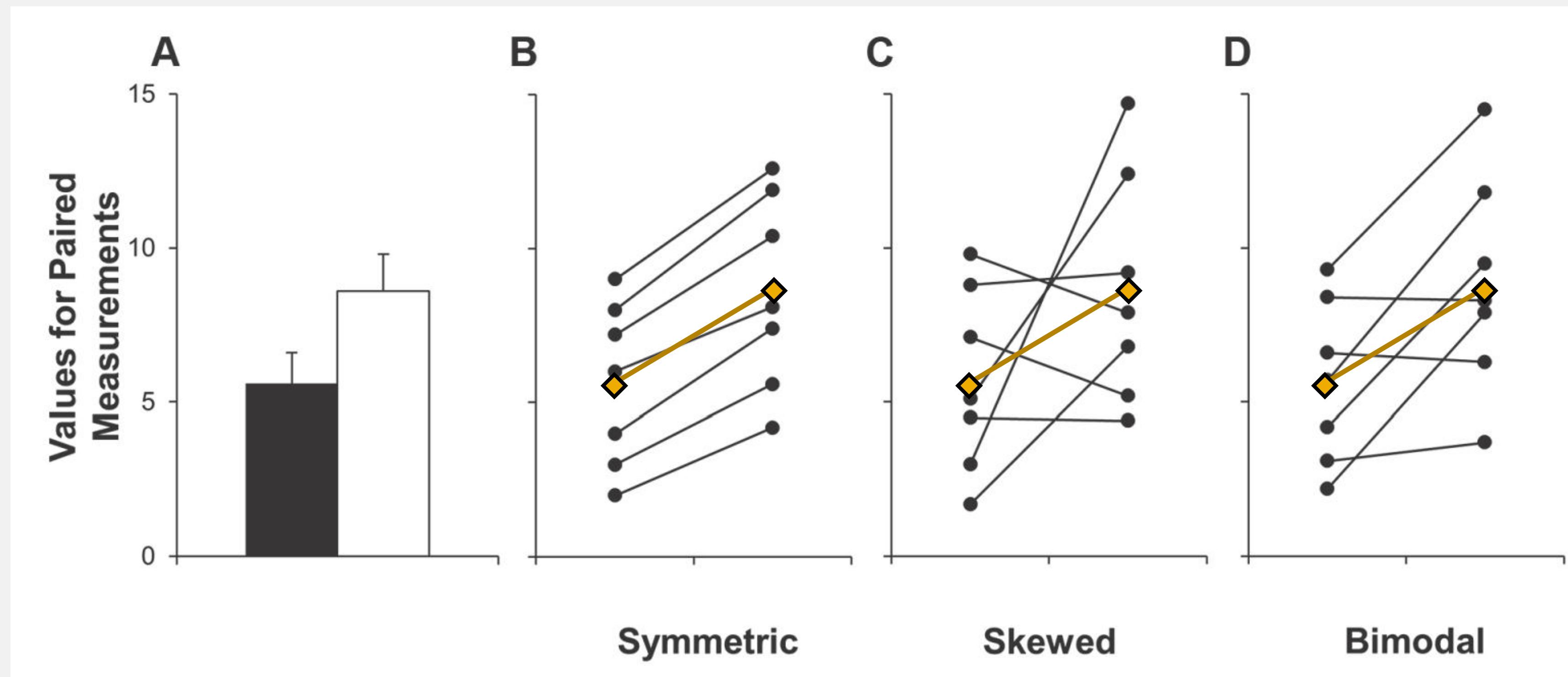
Weissgerber et al. (2015) PLoS Biology (modifiziert)





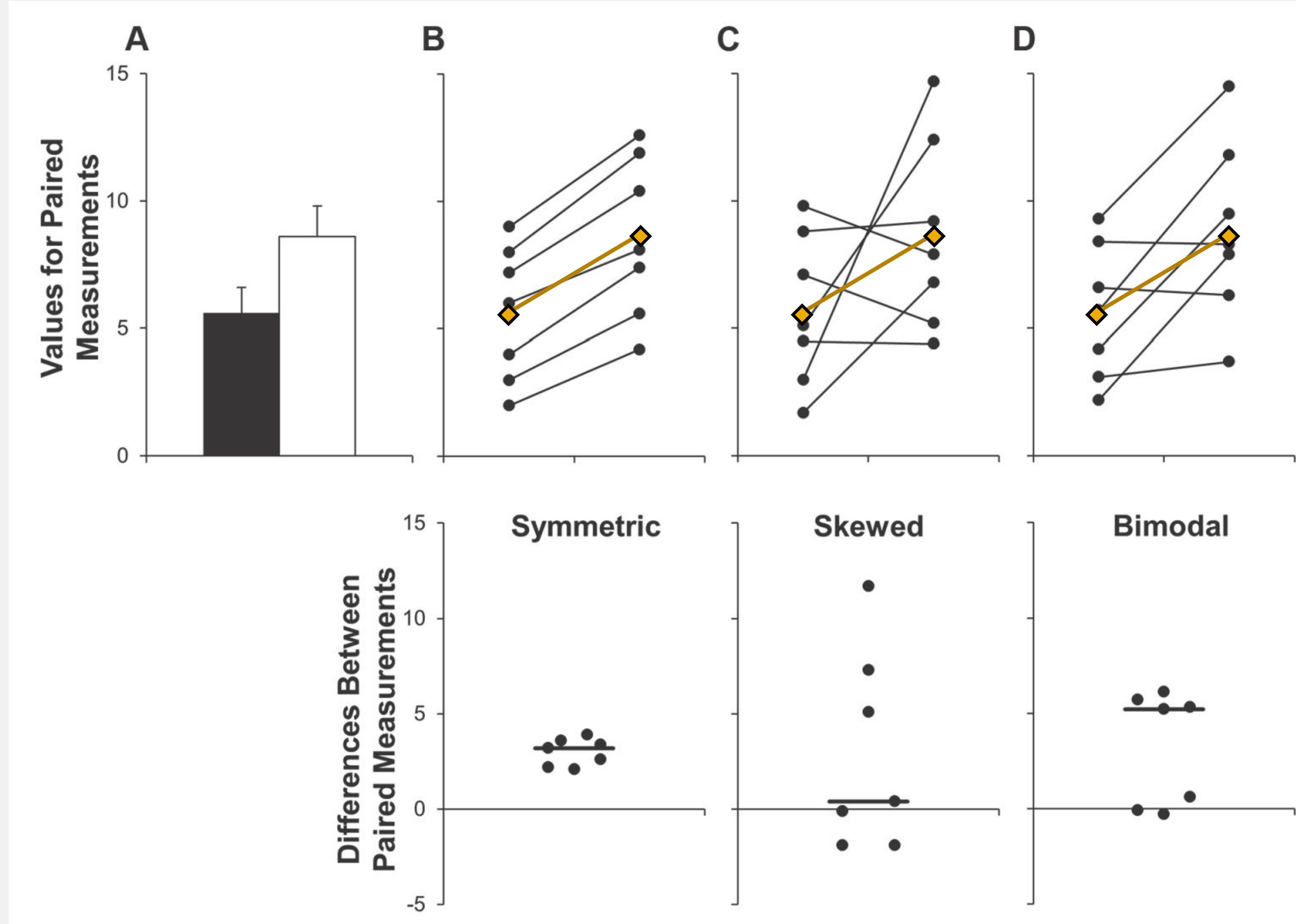
Weissgerber et al. (2015) PLoS Biology (modifiziert)



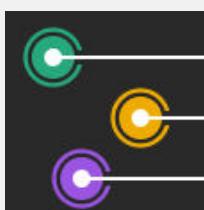


Modified from Weissgerber et al. (2015) PLoS Biology





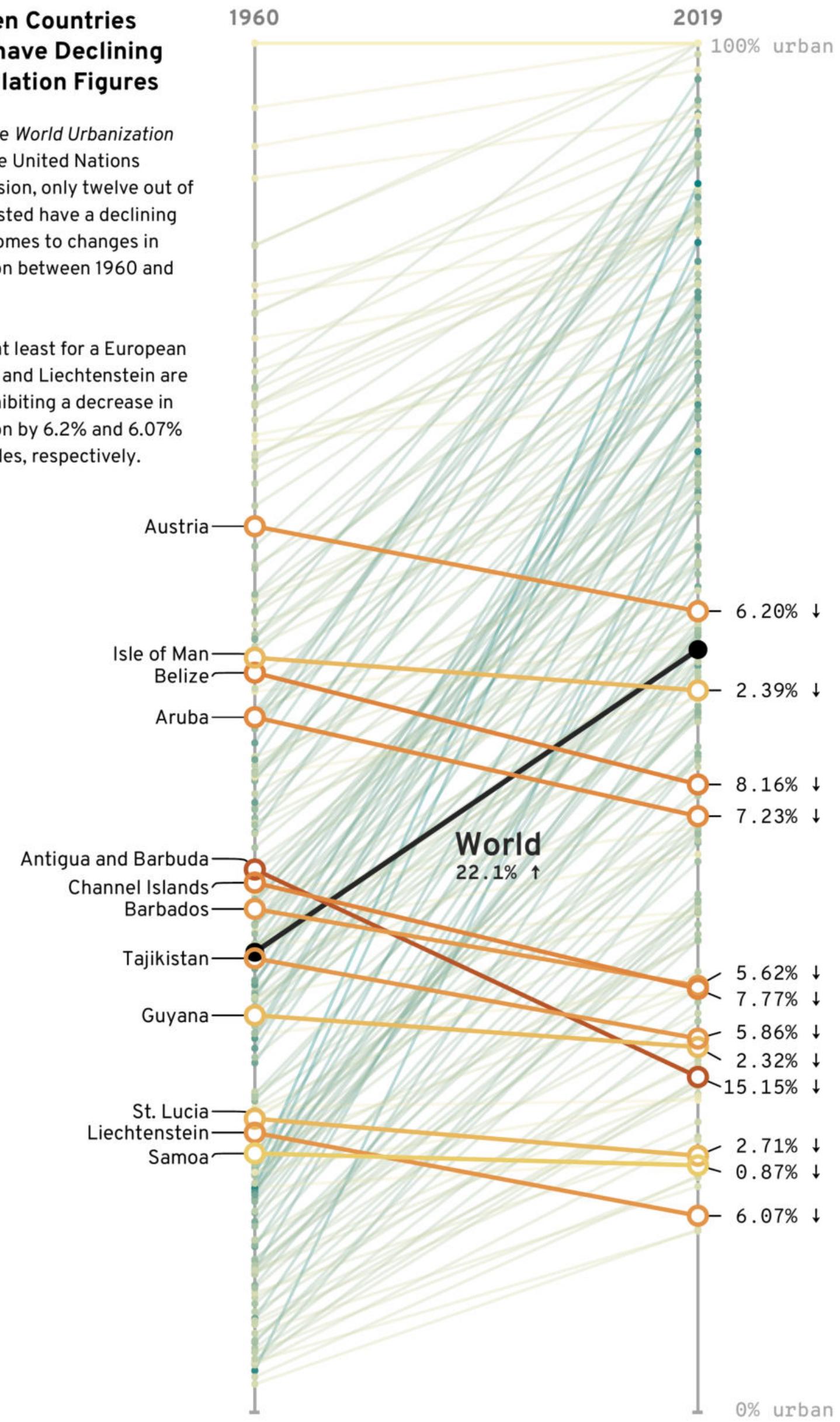
Weissgerber et al. (2015) PLoS Biology (modifiziert)



## **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 those exhibiting a decrease in urban population by 6.2% and 6.07% within six decades, respectively.

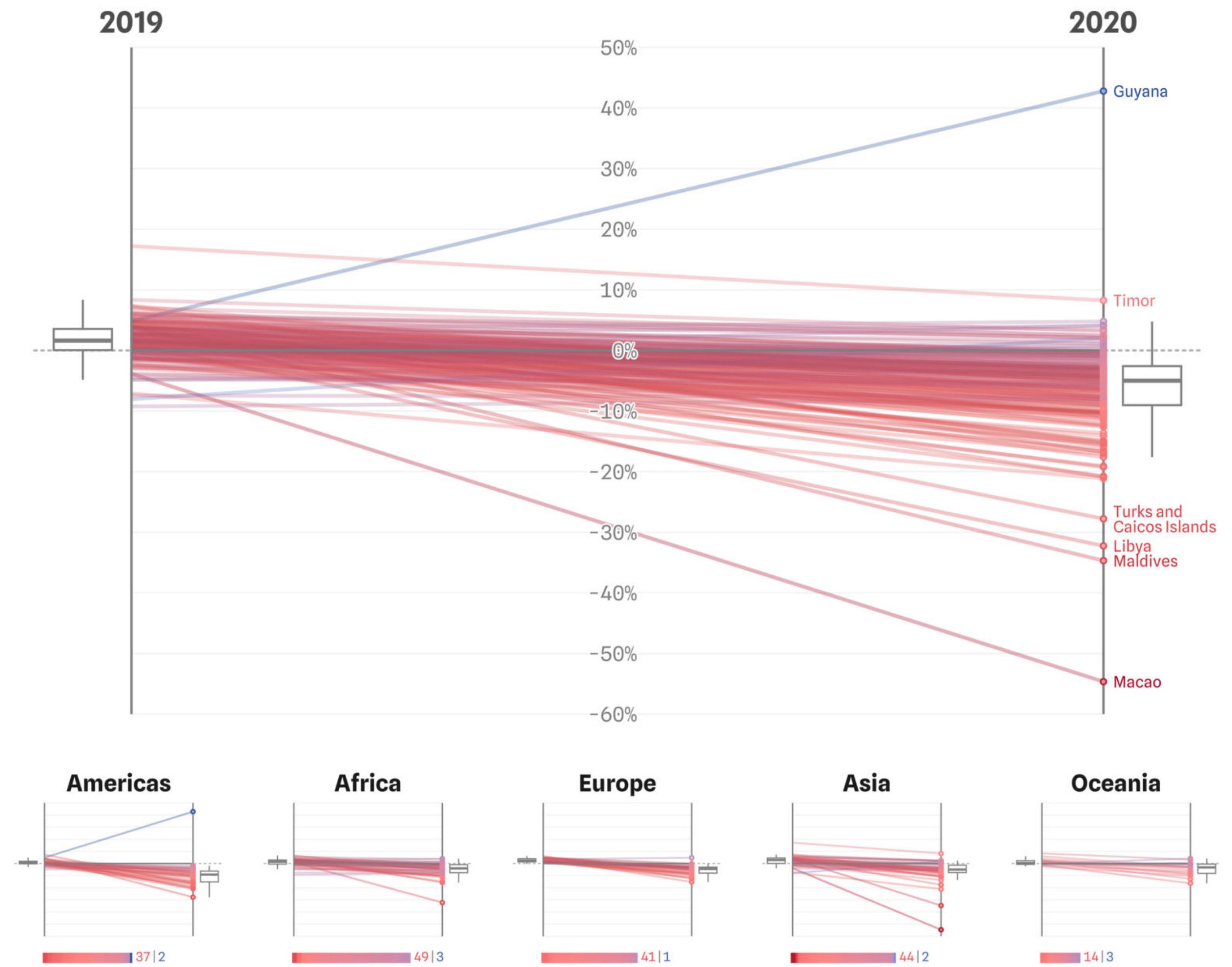


Visualization by Cédric Scherer | Data: United Nations Population Division. World Urbanization Prospects: 2018 Revision. | #30DayChartChallenge 2021 Day 5: Slope



# Average income has decreased in 185 countries from 2019 to 2020, with GDP per capita growth turning negative in 72% of these cases.

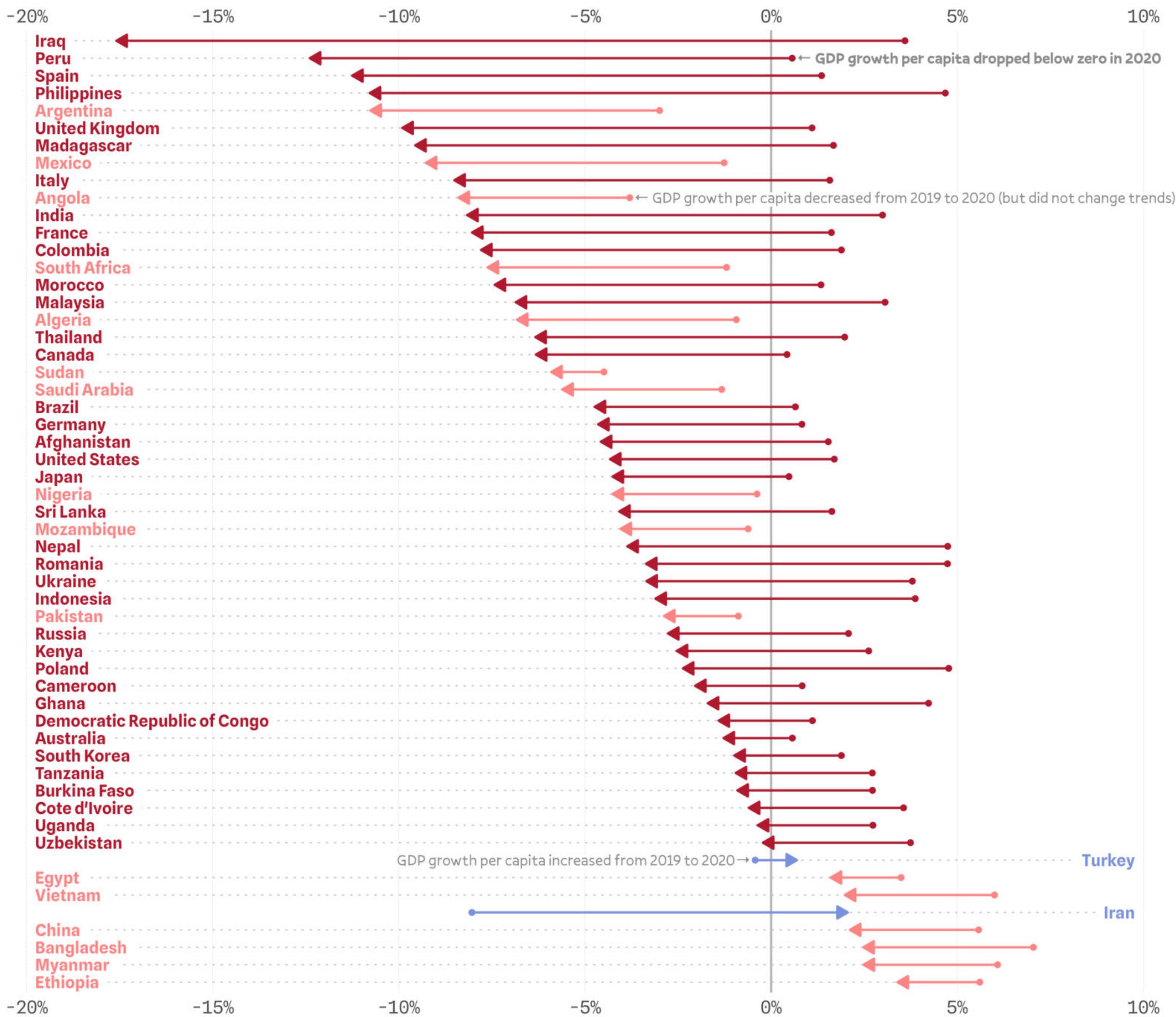
185 | 11  
Countries that experienced a **decline** or **increase** in annual change in GDP per capita growth (inflation adjusted)



Graphic: Cédric Scherer • Data: World Bank and OECD via OurWorldInData



## Average income has decreased in almost all countries from 2019 to 2020, with GDP per capita growth turning from positive to negative in most cases.

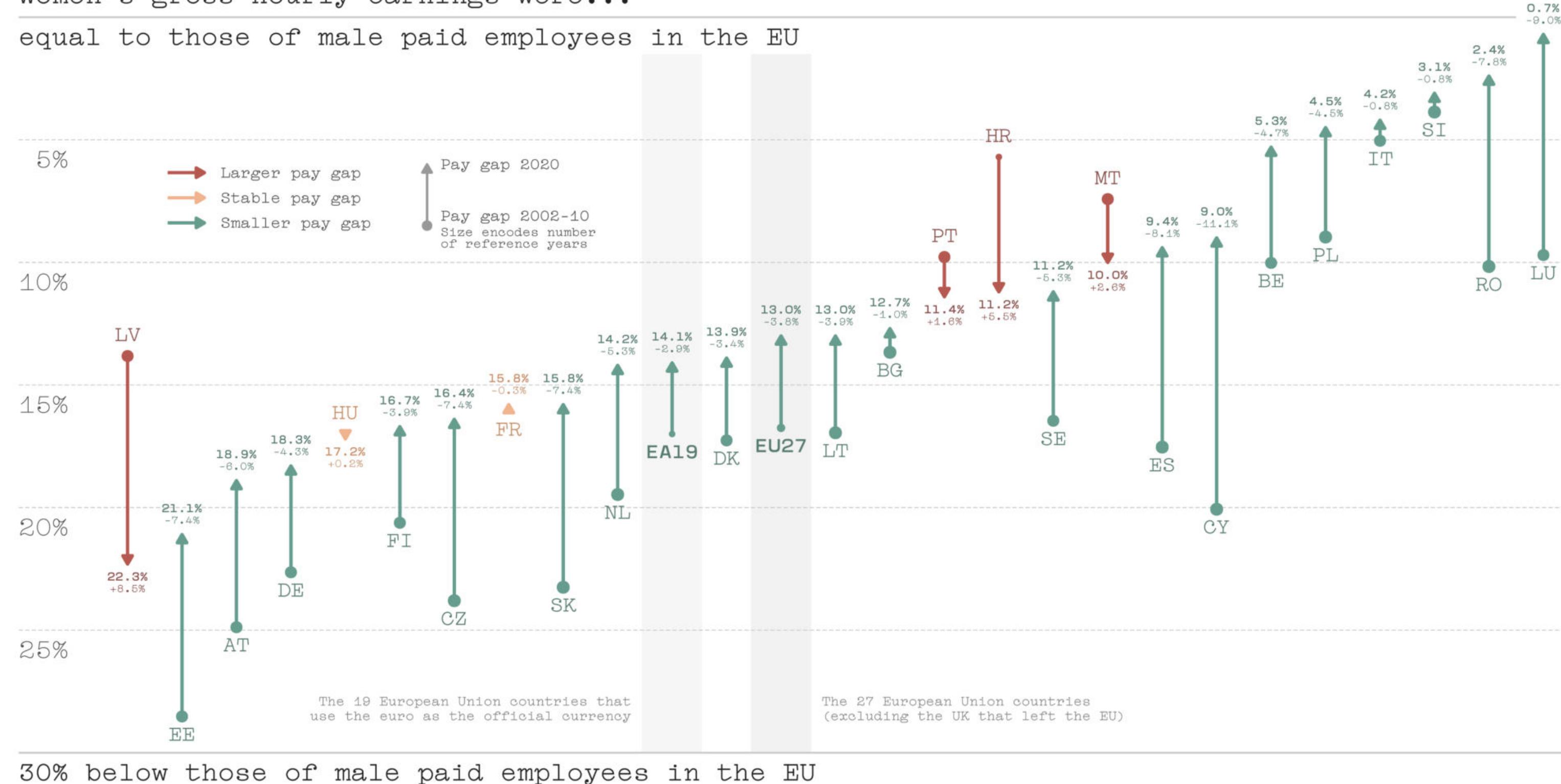


Graphic: Cédric Scherer • Data: World Bank and OECD via OurWorldInData • The graphic contains countries that inhabit more than 20M people.



Women's gross hourly earnings were...

equal to those of male paid employees in the EU



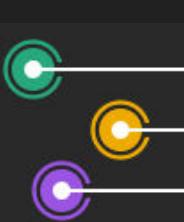
30% below those of male paid employees in the EU

In 2020, women's gross hourly earnings were on average 13% below those of men in the EU. Many countries were able to decrease the pay gap between male and female employees: most notably Luxembourg with the smallest gap in 2020 of 0.7% and Romania with a decrease of 13.6% within 18 years to 2.4%, ranking second. However, in several countries the pay gap was almost **stable** (Hungary and France) or even increased (Latvia, Portugal, Croatia, and Malta) compared to the gap in the reference period from 2002 to 2010 for all years available via Eurostat.

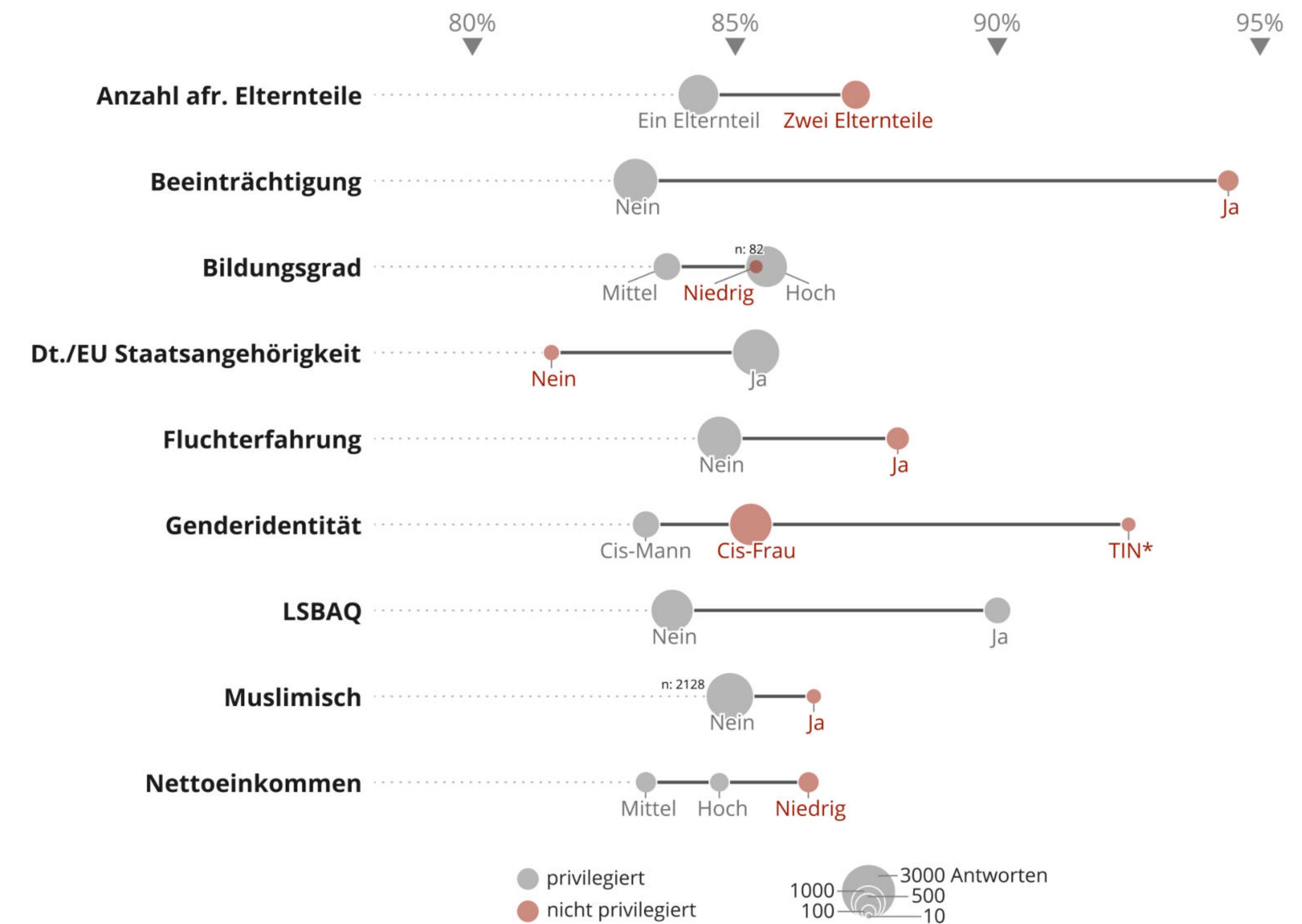
As an unadjusted indicator, the gender pay gap gives an overall picture of the differences between men and women in terms of earnings and measures – a concept which is broader than discrimination in the sense of: *equal pay for work of equal value*.

Graphic: Cédric Scherer • Data: Eurostat (SDG\_05\_20; no data for Greece and Ireland) • #InternationalWomensDay2022

*“The Pay Gap in Europe”, persönliches Projekt zum Internationalen Frauentag 2022*



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

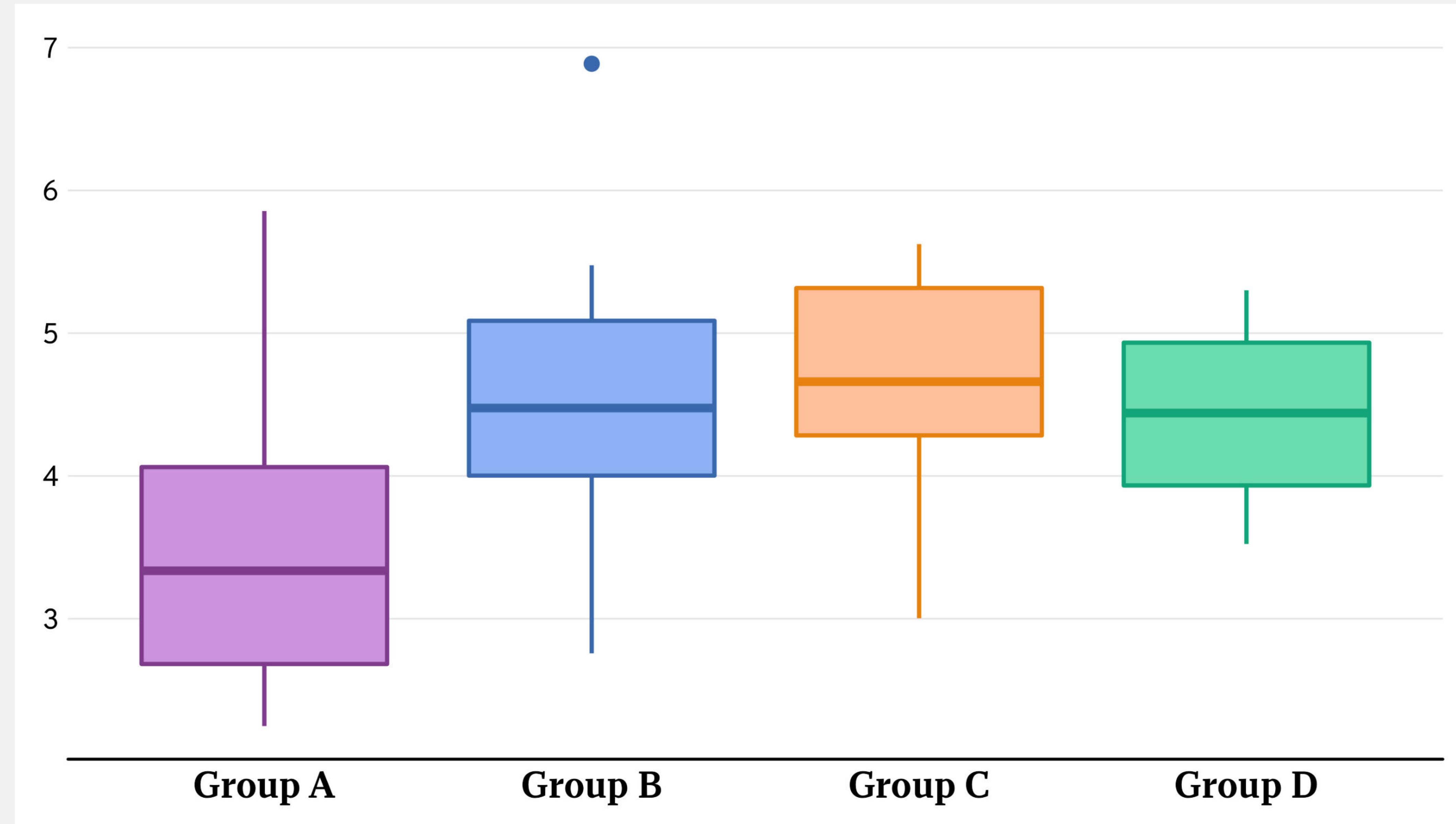


**Lesebeispiel:** LSBAQ-Befragte des Afrozensus geben im Vergleich mit heterosexuellen Afrozensus-Befragten häufiger an, im Lebensbereich „Medien und Internet“ in den letzten zwei Jahren Diskriminierung erlebt zu haben.

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

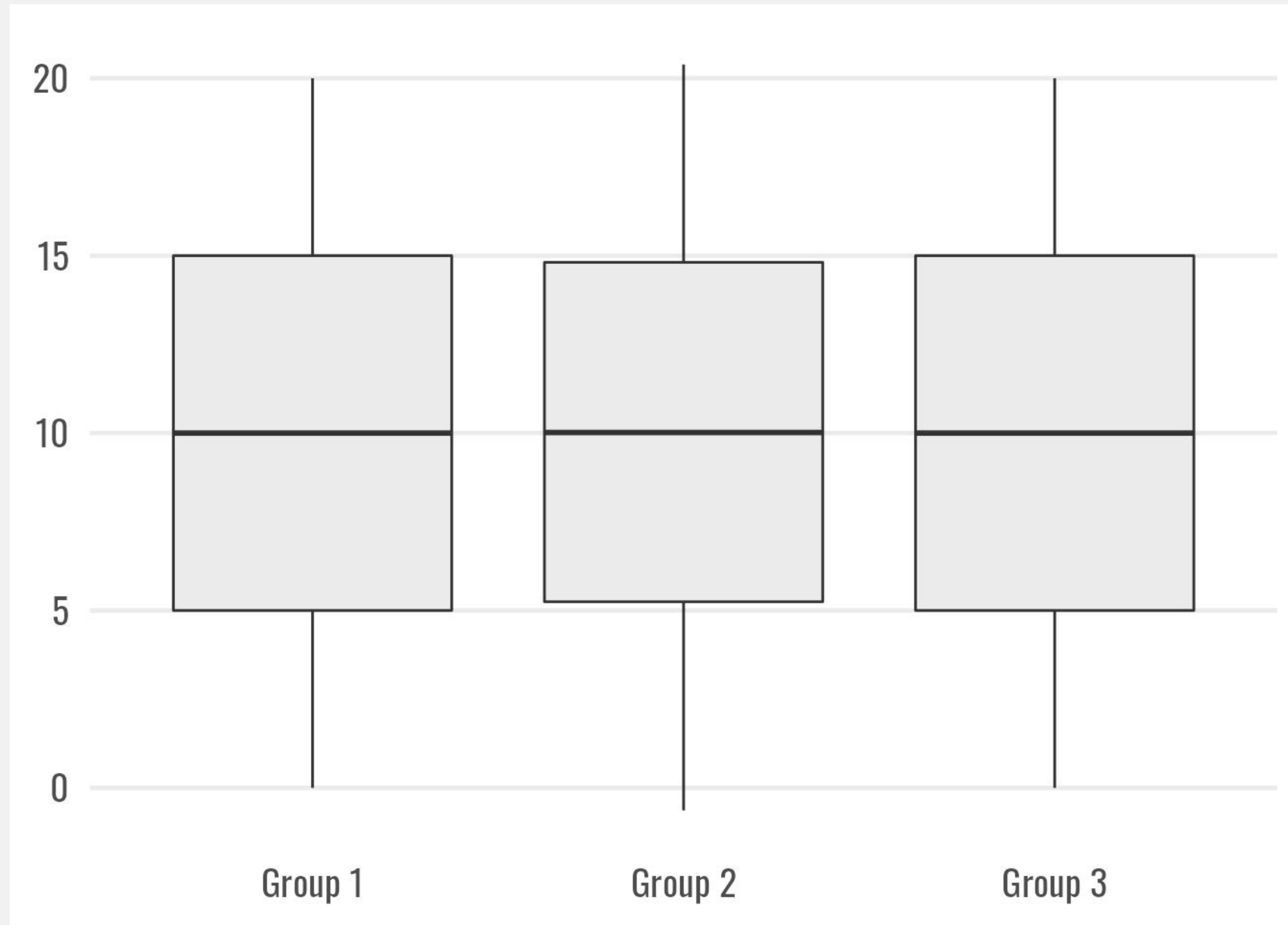
**Abb. 46 “Afrozensus 2020” von Citizens For Europe & EOTO e.V.**





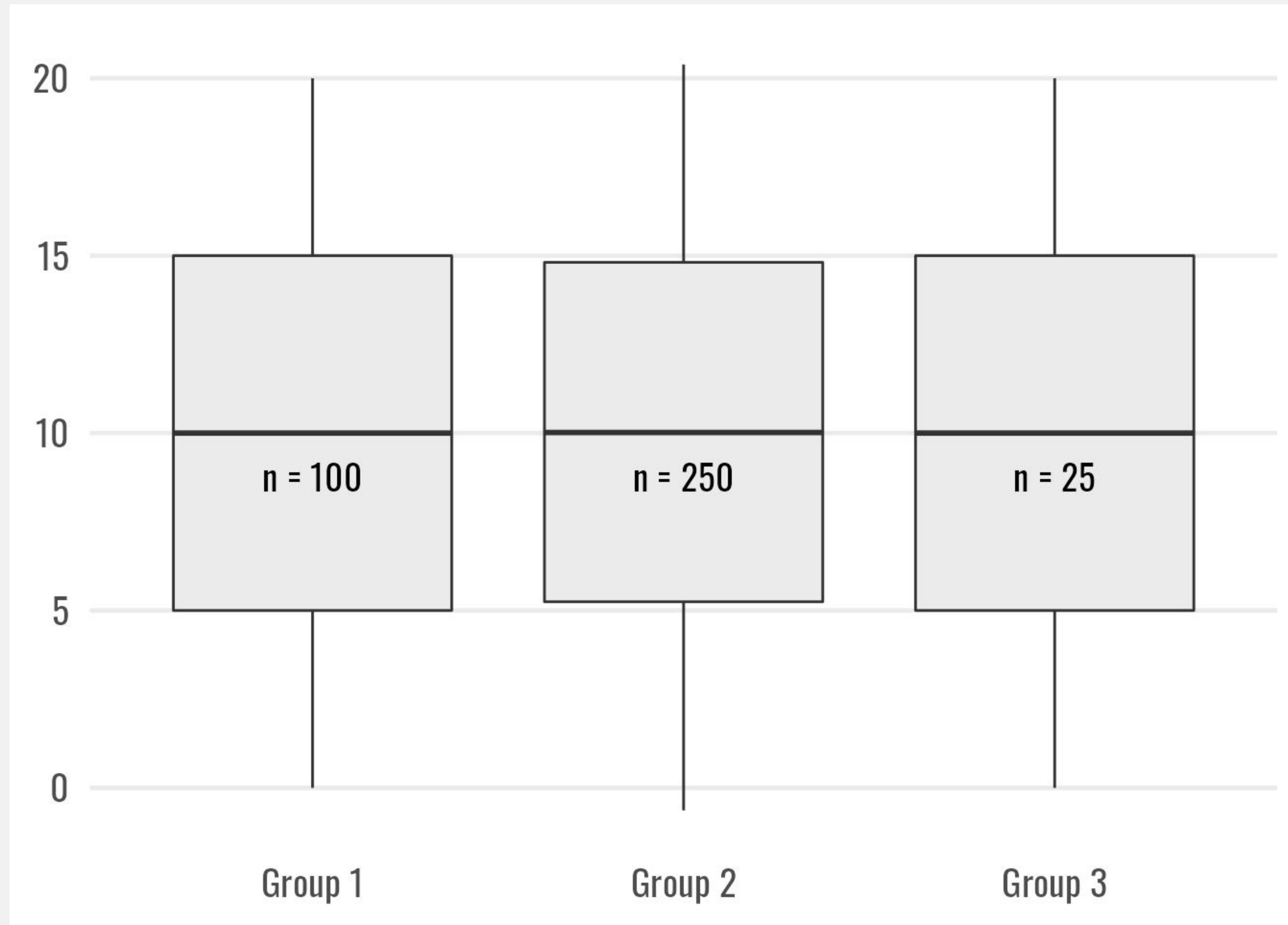
Mein Webinar für USGS Data Science “Beyond Bar and Box Plots” [[Slides](#) | [Recording](#)]



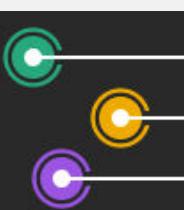


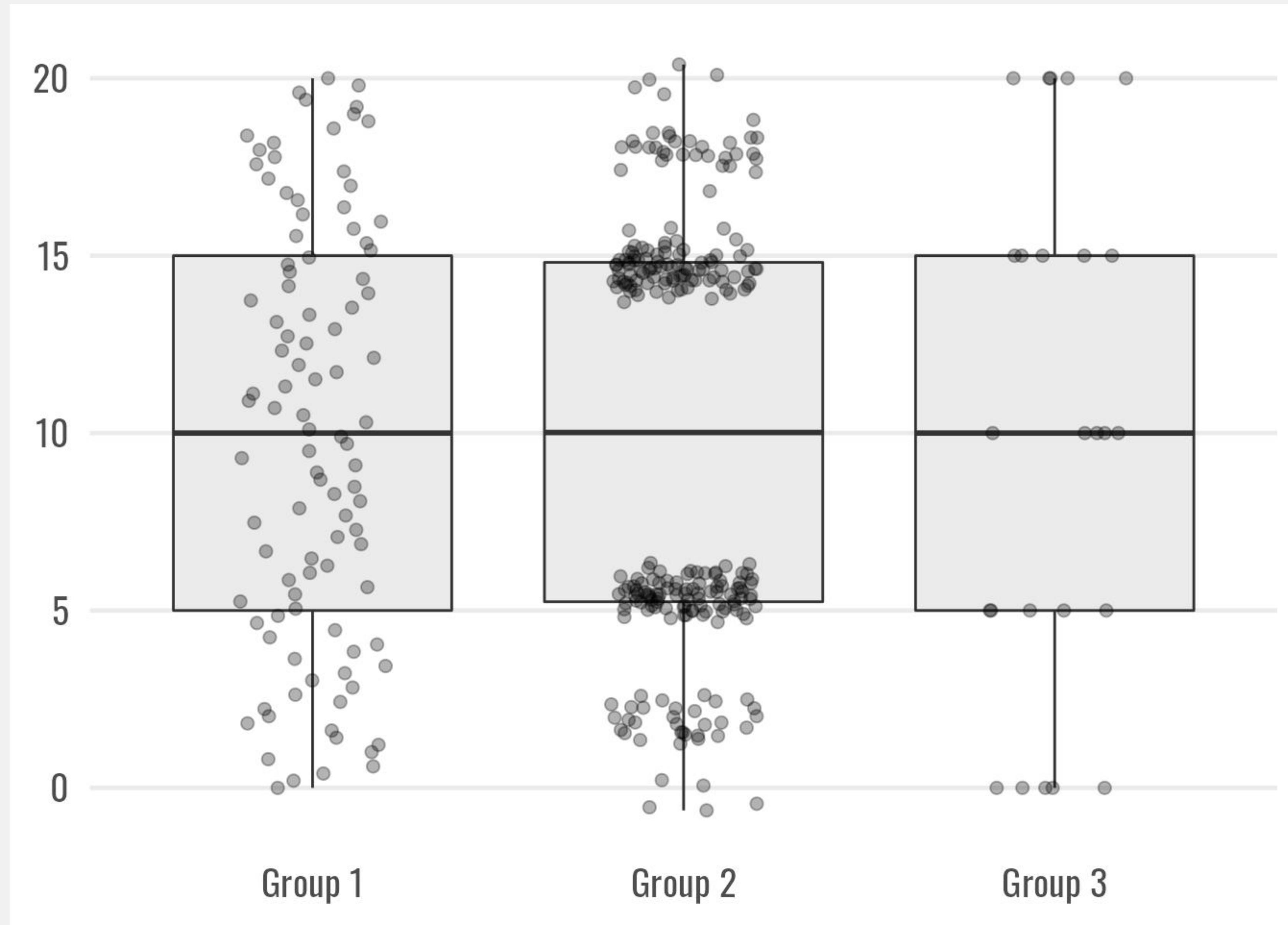
von meinem Blogpost [“Visualizing Distributions with Raincloud Plots \(and How to Create Them with ggplot2\)”](#)





von meinem Blogpost “[Visualizing Distributions with Raincloud Plots \(and How to Create Them with ggplot2\)](#)”



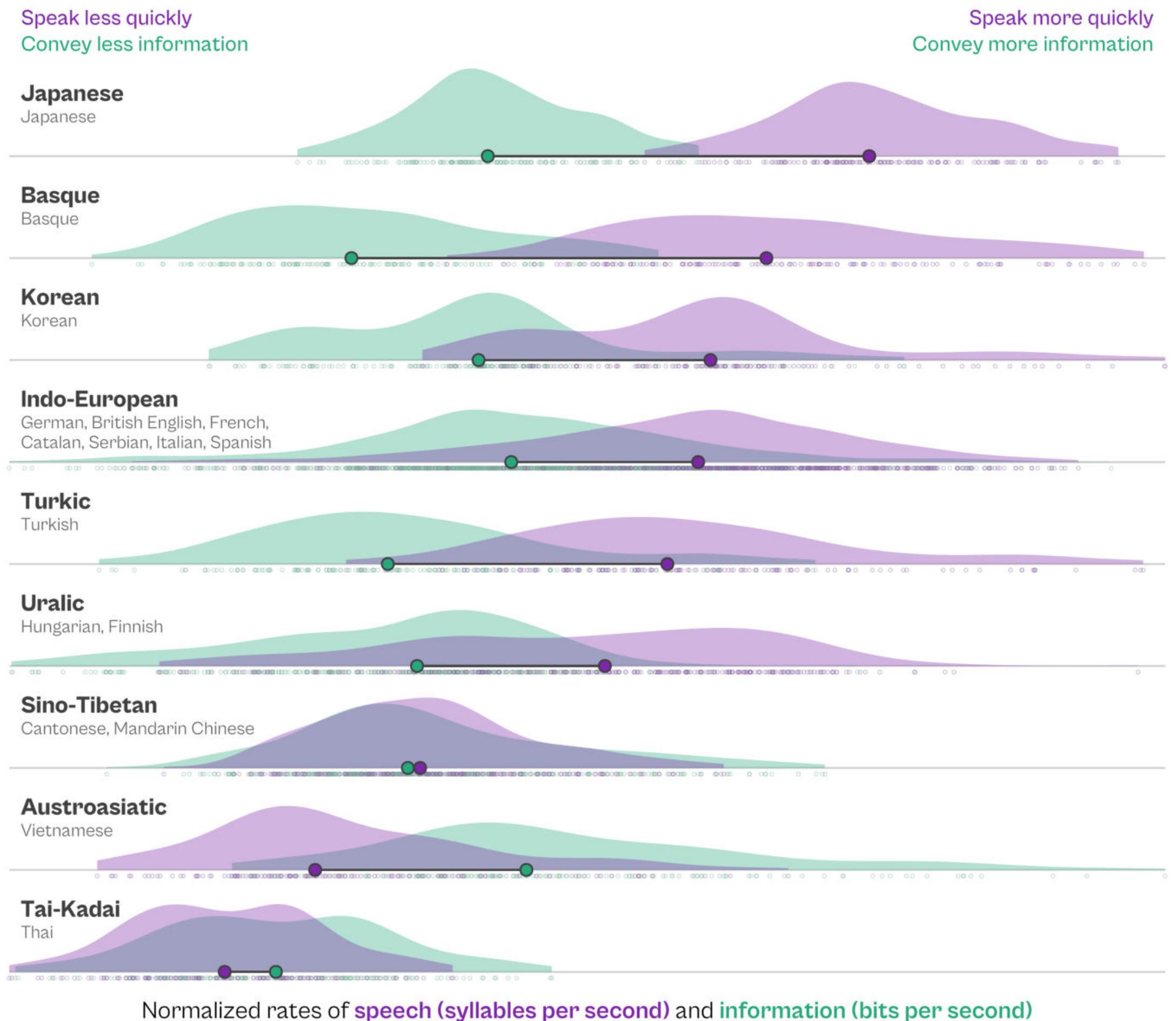


von meinem Blogpost “[Visualizing Distributions with Raincloud Plots \(and How to Create Them with ggplot2\)](#)”



## Communicating fast doesn't necessarily mean communicating more

Variation in speech and information rates across language families, shown as normalized rates for comparison. While there are stark cross-linguistic differences in speech rates, information rates are more similar.

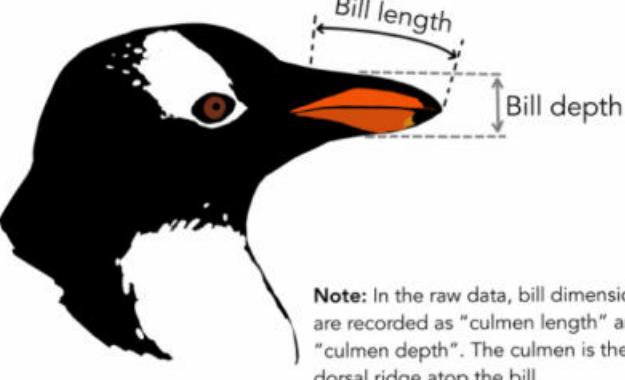


“Communicating fast doesn't necessarily mean communicating more”, #30DayChartChallenge Contribution



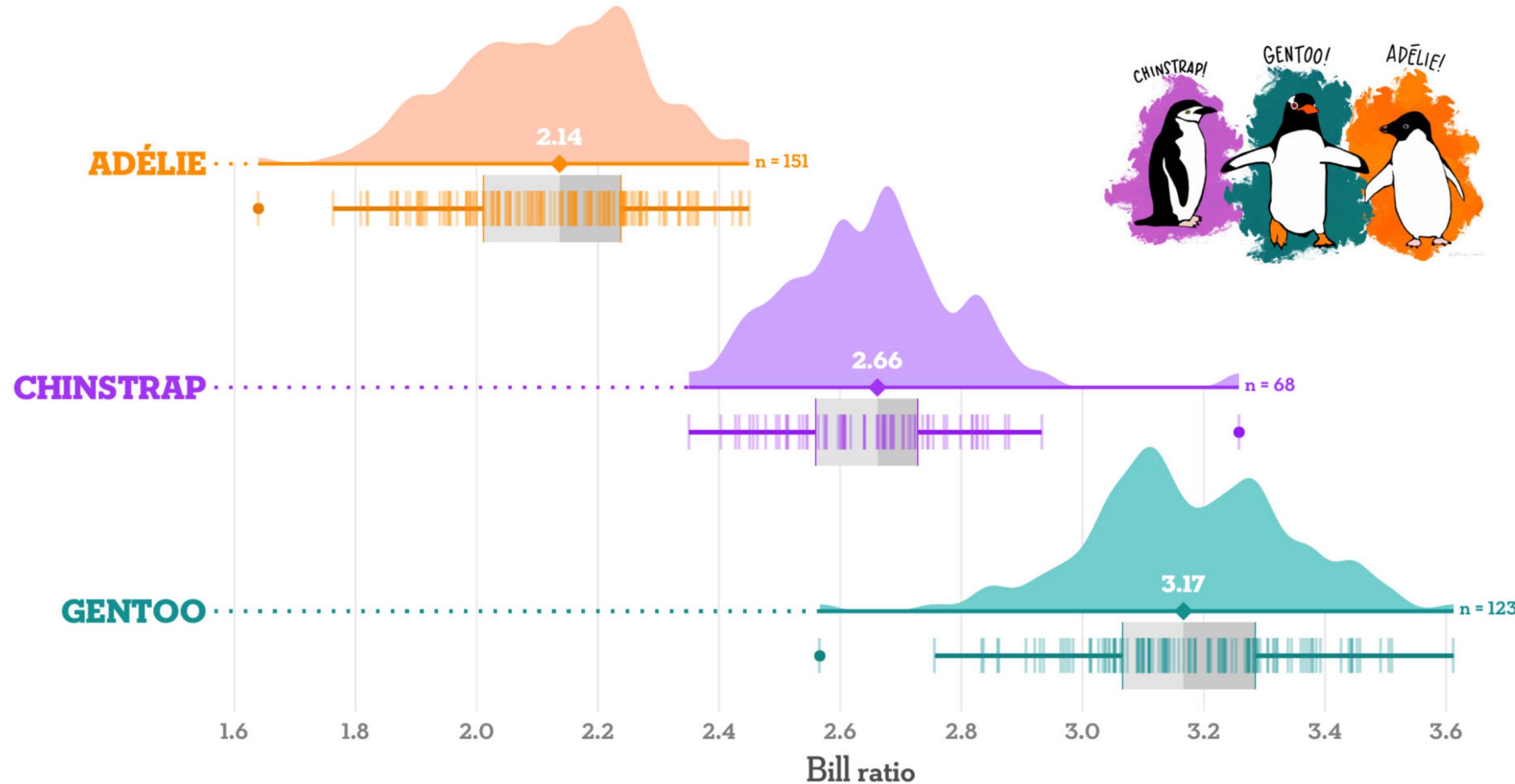
# BILL DIMENSIONS OF BRUSH-TAILED PENGUINS

*Pygoscelis adélieae* (Adélie penguin) • *P. antarctica* (Chinstrap penguin) • *P. papua* (Gentoo penguin)



Note: In the raw data, bill dimensions are recorded as "culmen length" and "culmen depth". The culmen is the dorsal ridge atop the bill.

Distribution of the bill ratio, estimated as bill length divided by bill depth

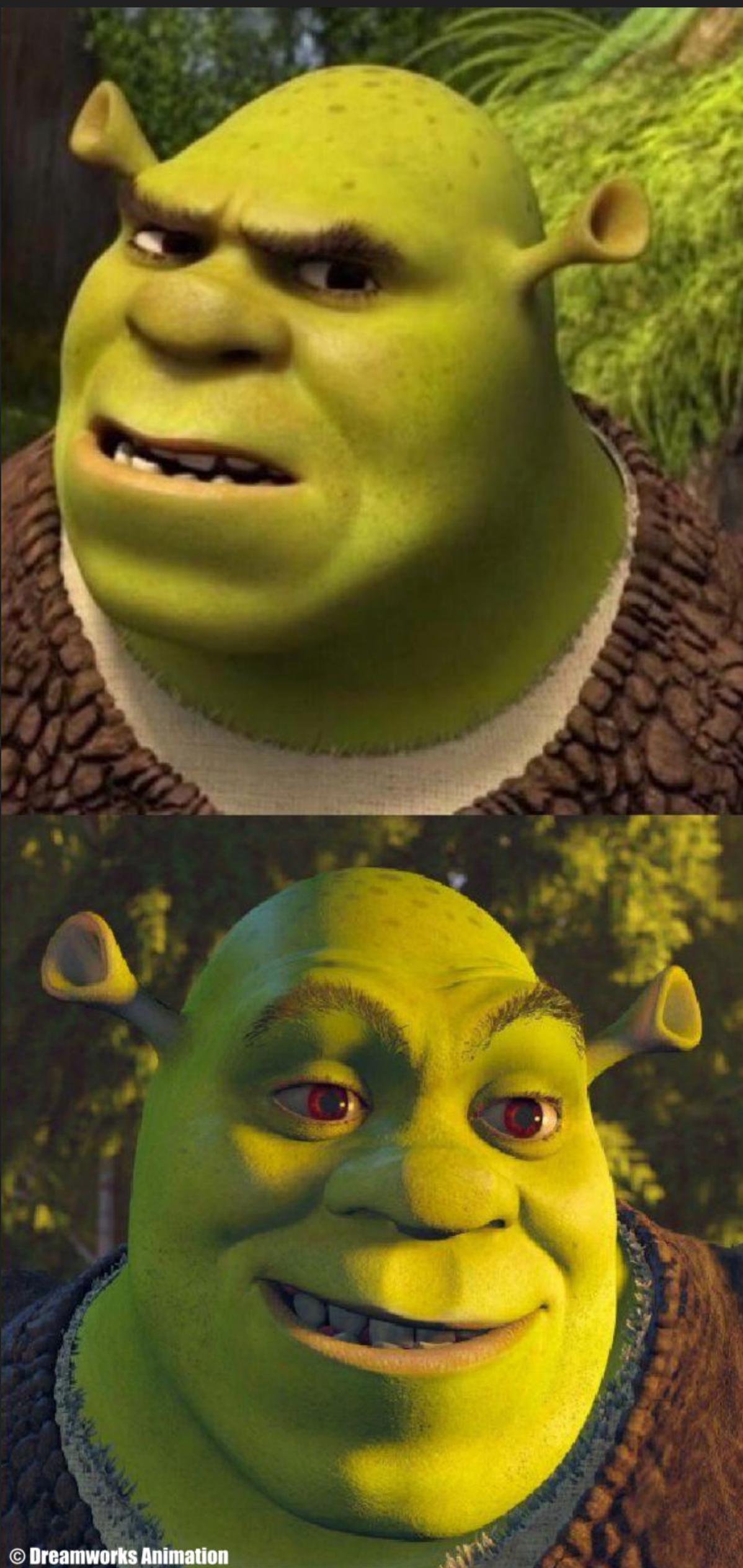


Note: In the original data, bill dimensions are recorded as "culmen length" and "culmen depth". The culmen is the dorsal (upper) ridge of a bird's bill.

Visualization: Cédric Scherer • Data: Gorman, Williams & Fraser (2014) DOI: 10.1371/journal.pone.0090081 • Illustrations: Allison Horst

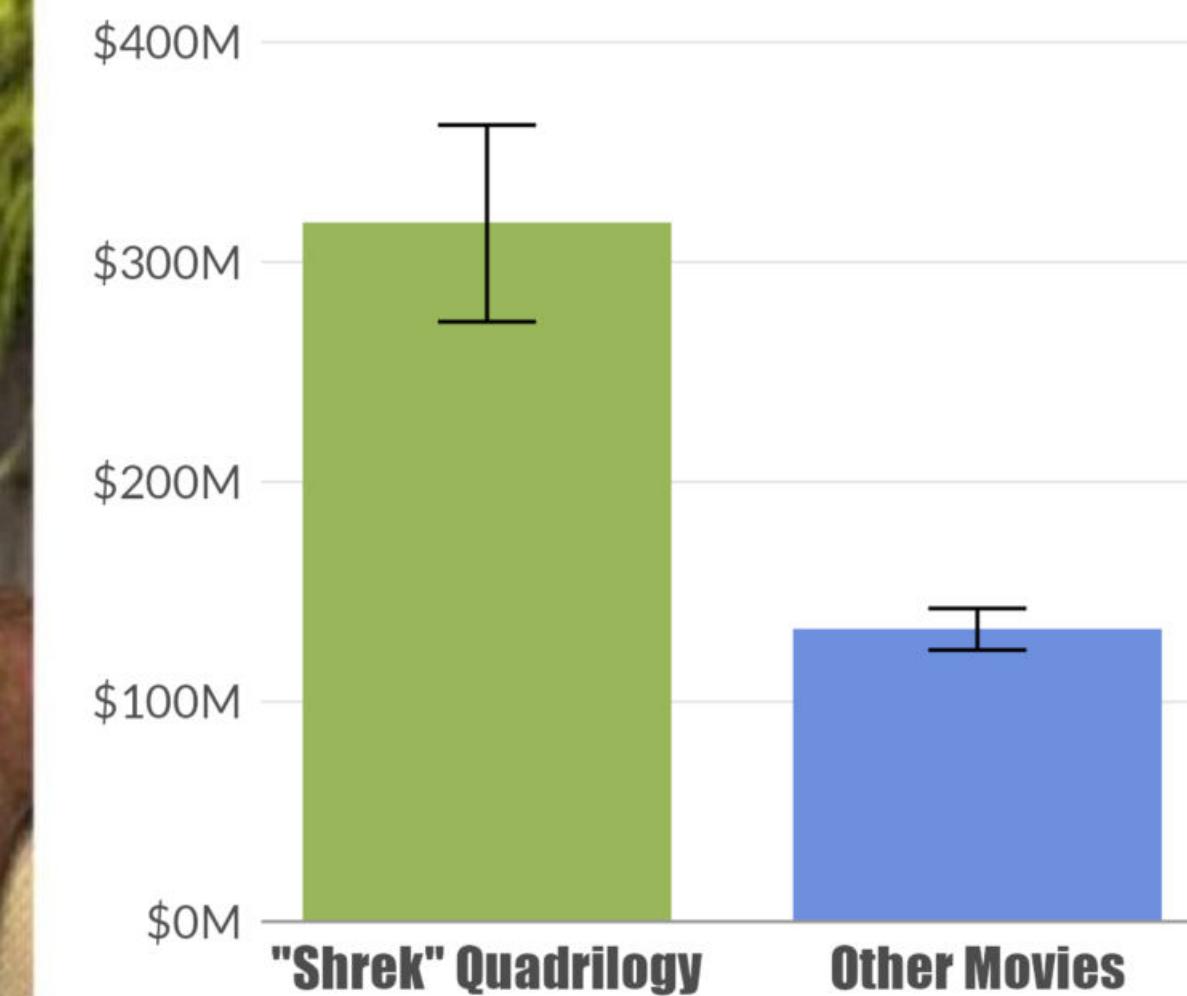
von meinem Blogpost “[Visualizing Distributions with Raincloud Plots \(and How to Create Them with ggplot2\)](#)”



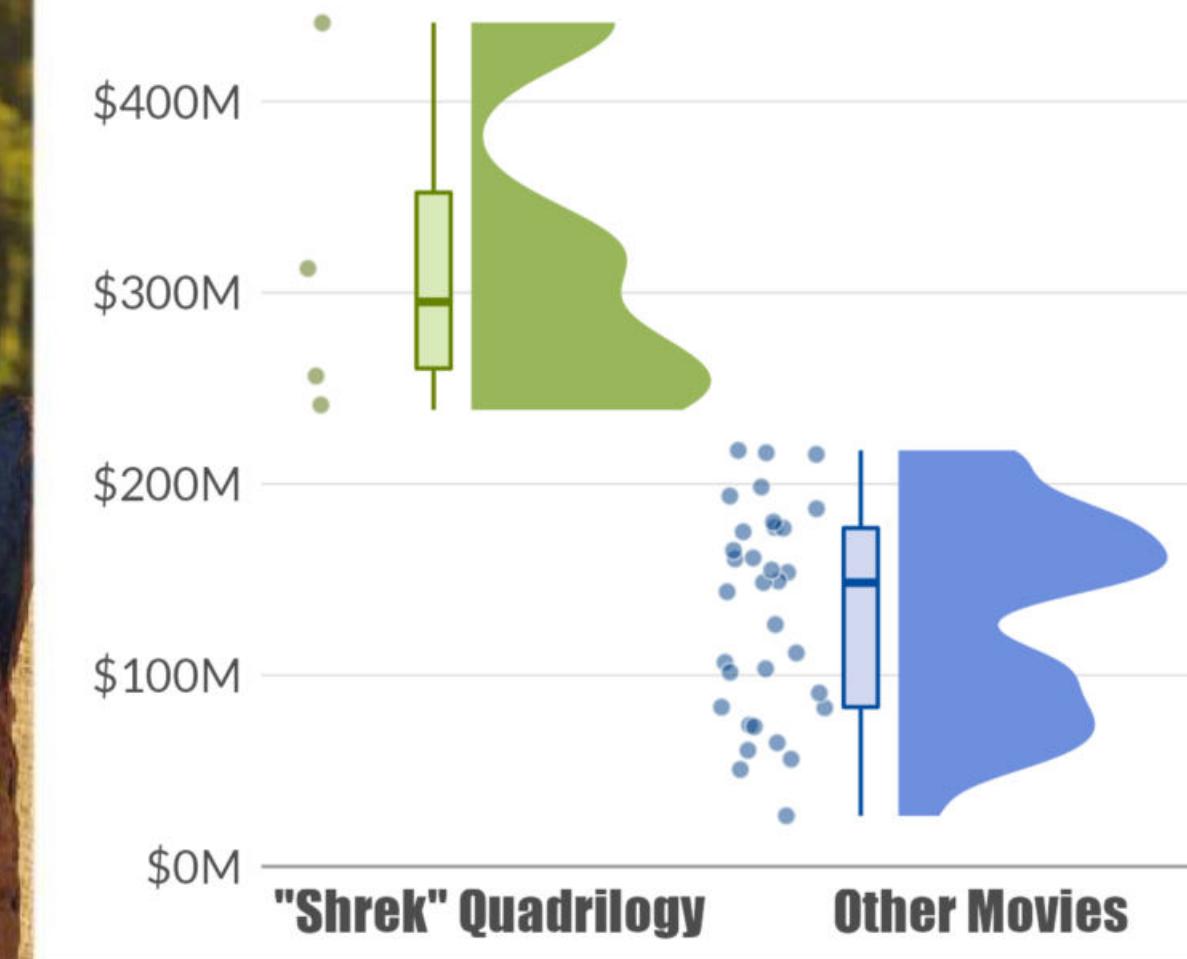


© Dreamworks Animation

## Domestic Box Office of DreamWorks Movies



## Domestic Box Office of DreamWorks Movies



Why Dynamite Plots Are Terrible—and Why You Should Use Something Else | Cédric Scherer | #30DayChartChallenge 2021 | Day 27: Educational



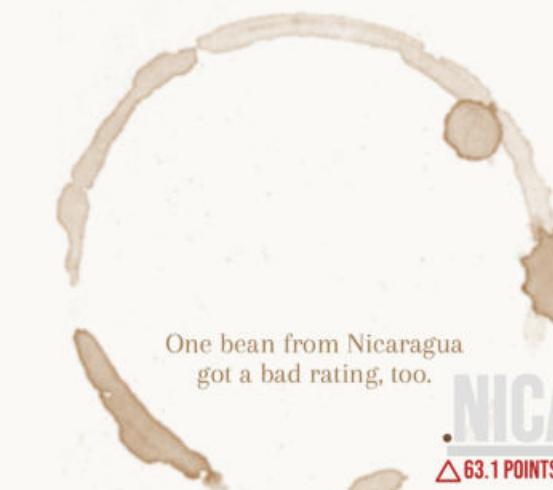
# Not my cup of coffee...

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



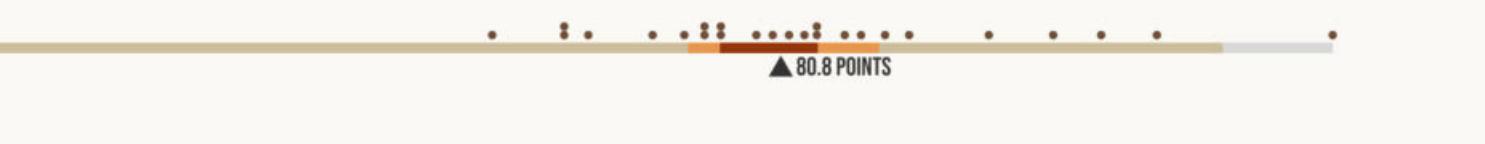
The coffee bean with the lowest rating has its origin in Guatemala.



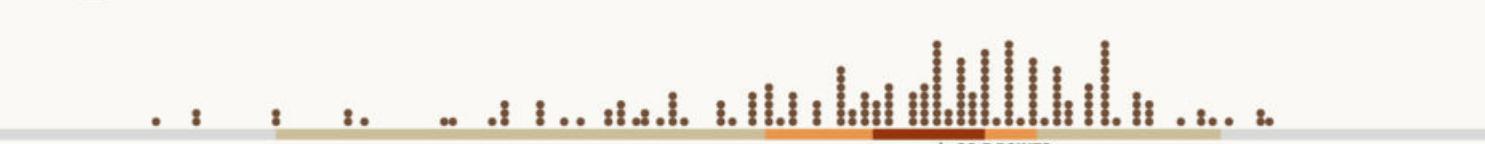
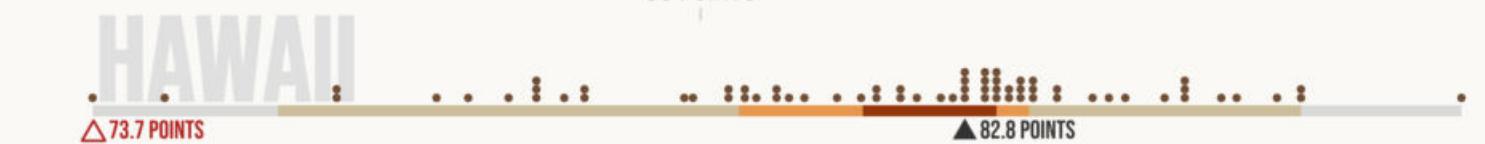
One bean from Nicaragua got a bad rating, too.



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



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

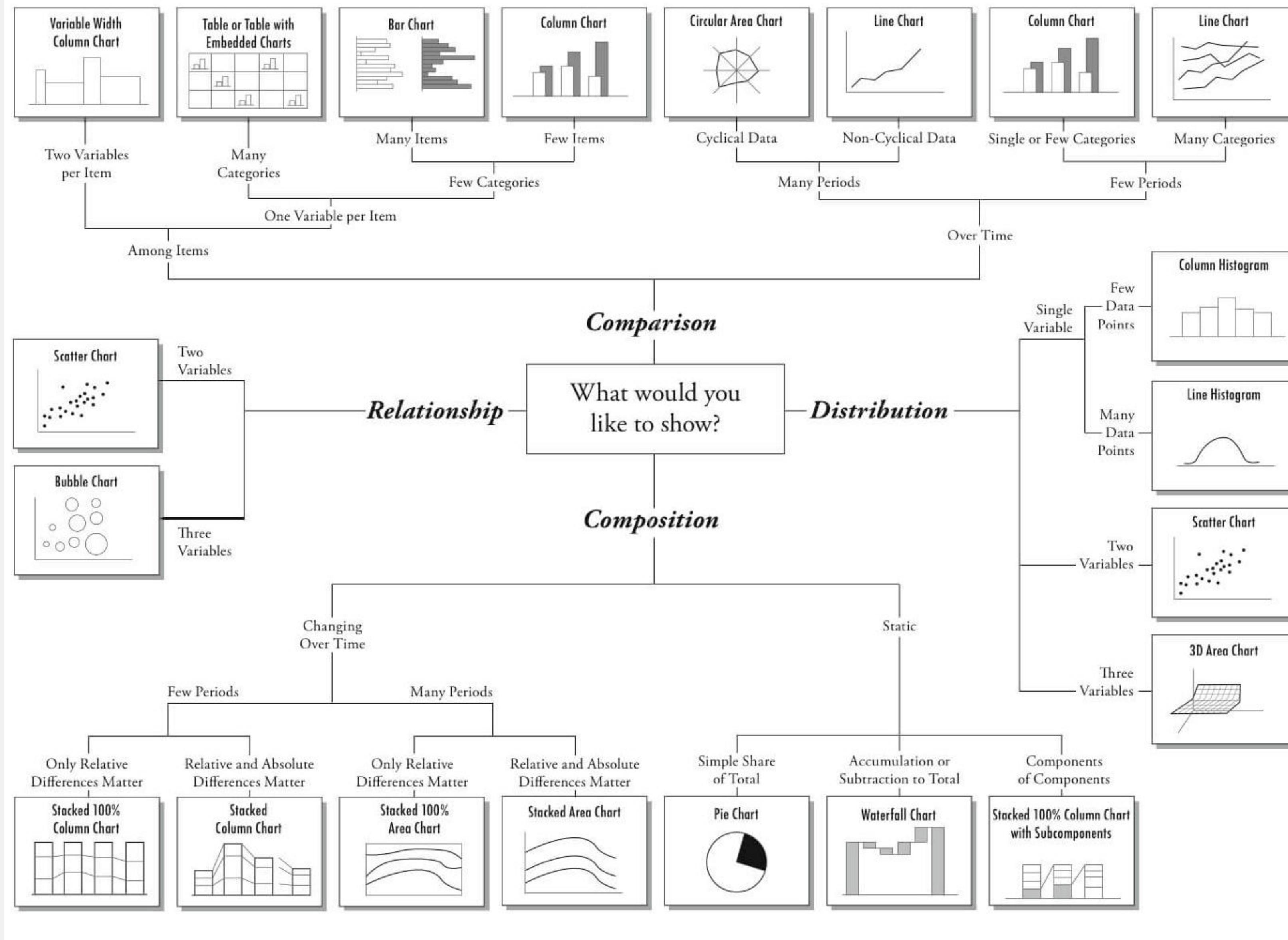


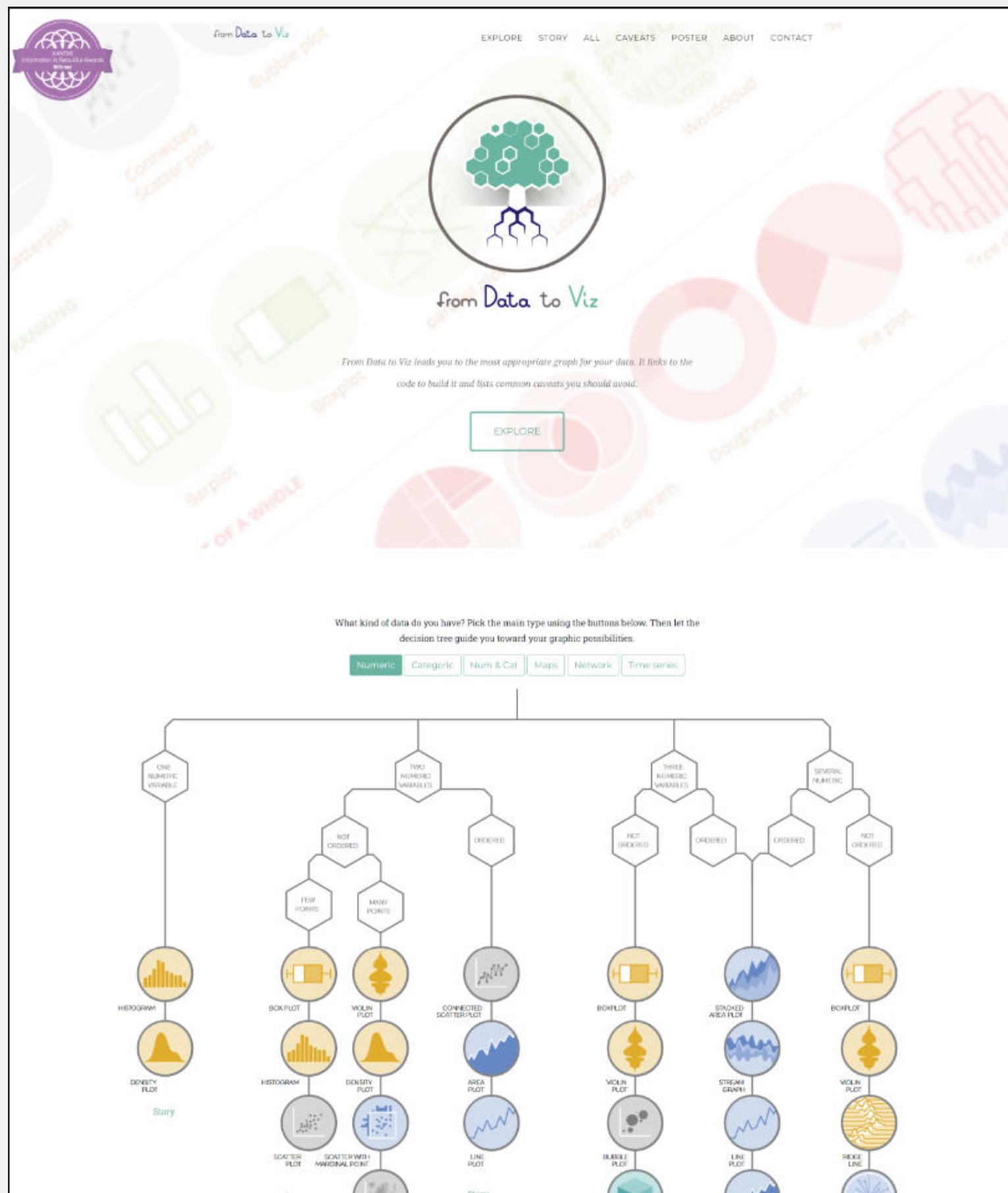
*“Not my cup of coffee”, #TidyTuesday Contribution*



# Chart Suggestions—A Thought-Starter

www.ExtremePresentation.com  
© 2009 A. Abela — a.v.abela@gmail.com

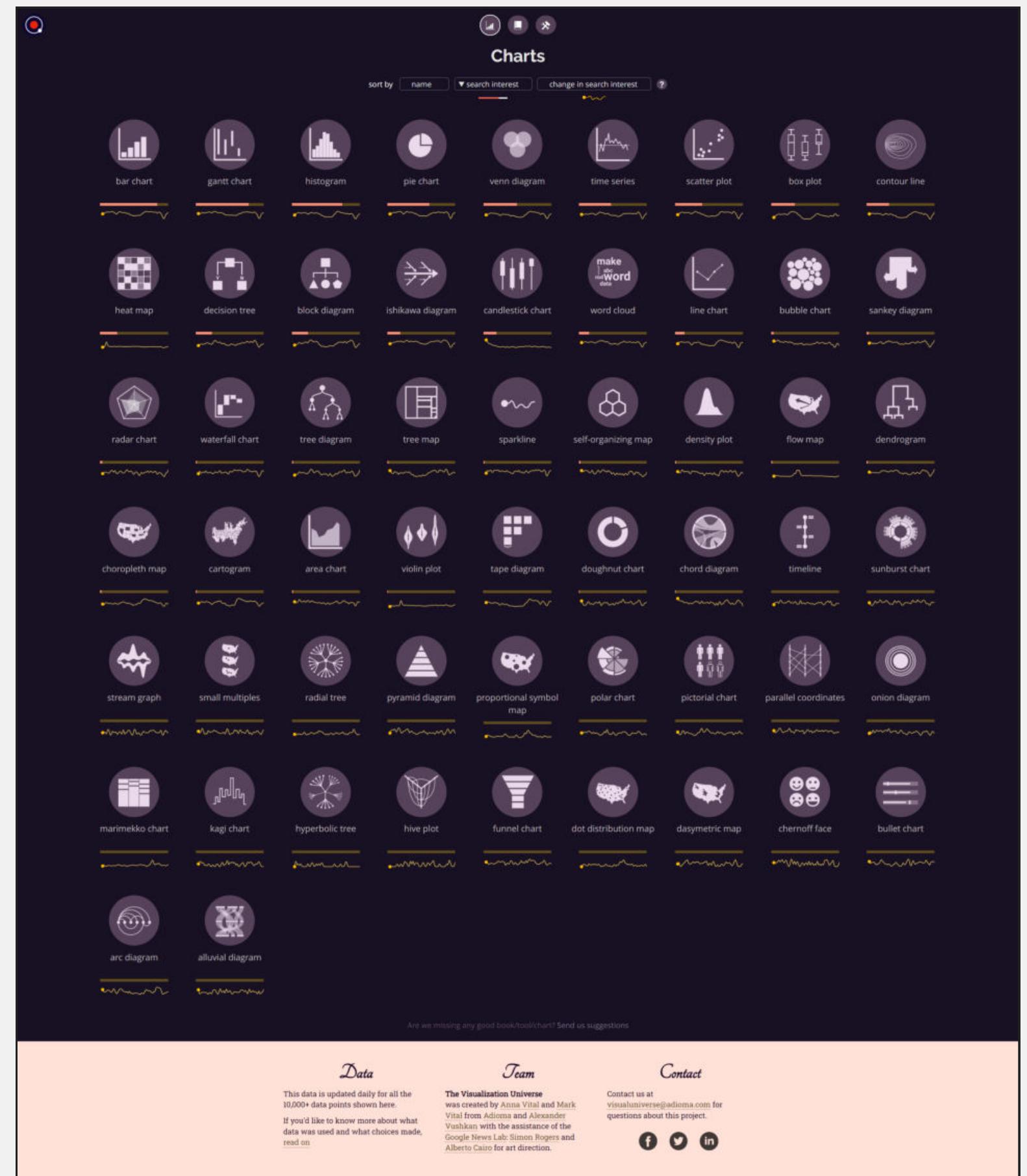




[data-to-viz.com](http://data-to-viz.com)



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



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





# from Data to Viz

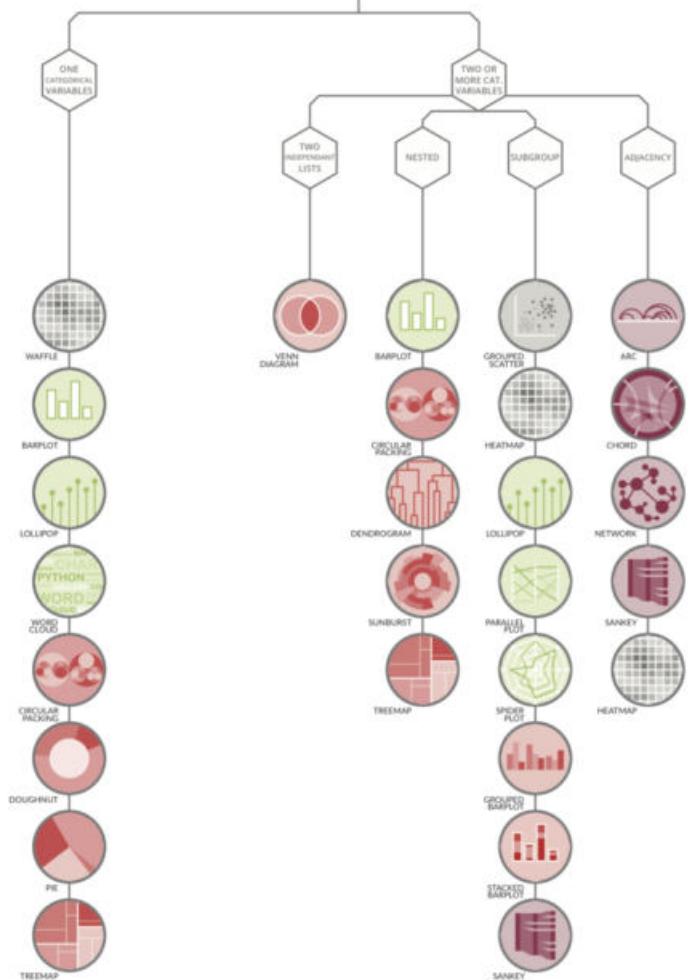
'From Data to Viz' is a classification of chart types based on input data format. It will help you find the perfect chart in three simple steps:

- 1 Identify what type of data you have.
- 2 Go to the corresponding decision tree and follow it down to a set of possible charts.
- 3 Choose the chart from the set that will suit your data and your needs best.

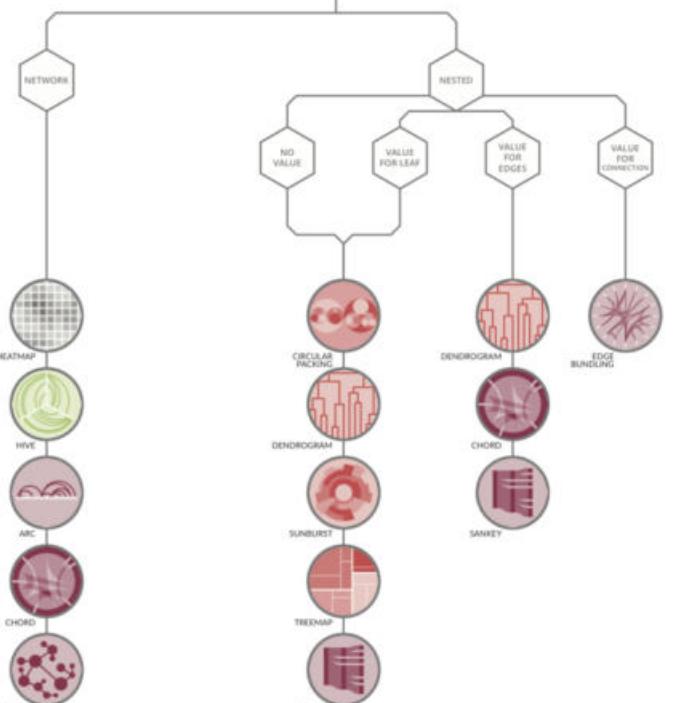
Dataviz is a world with endless possibilities and this project does not claim to be exhaustive. However it should provide you with a good starting point. For an interactive version and much more, visit:

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

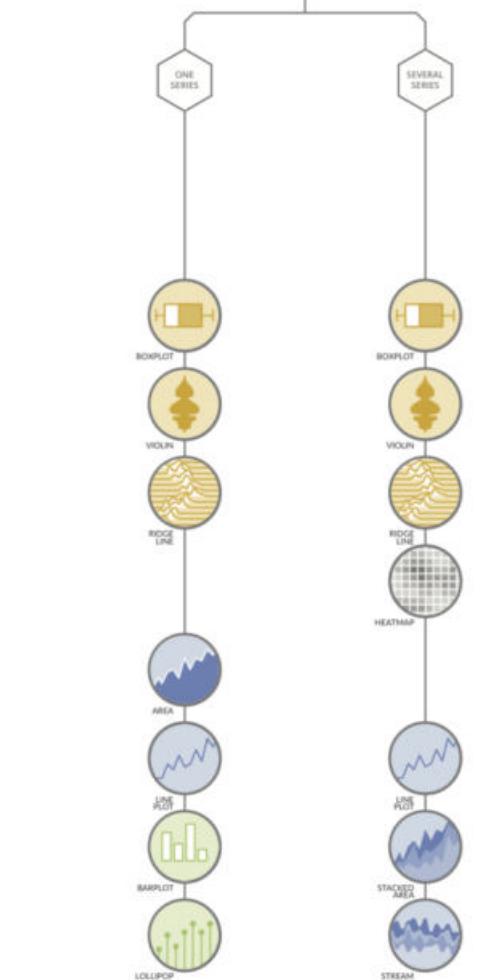
## CATEGORIC



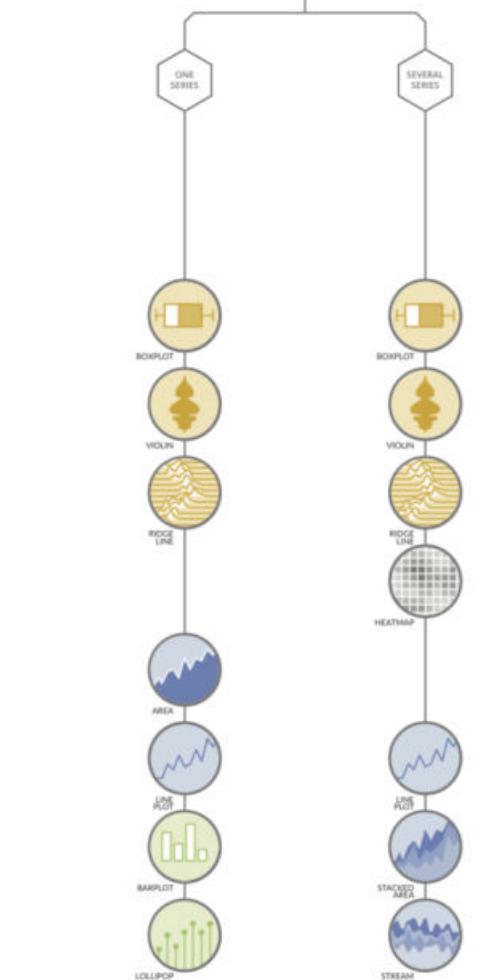
## RELATIONAL



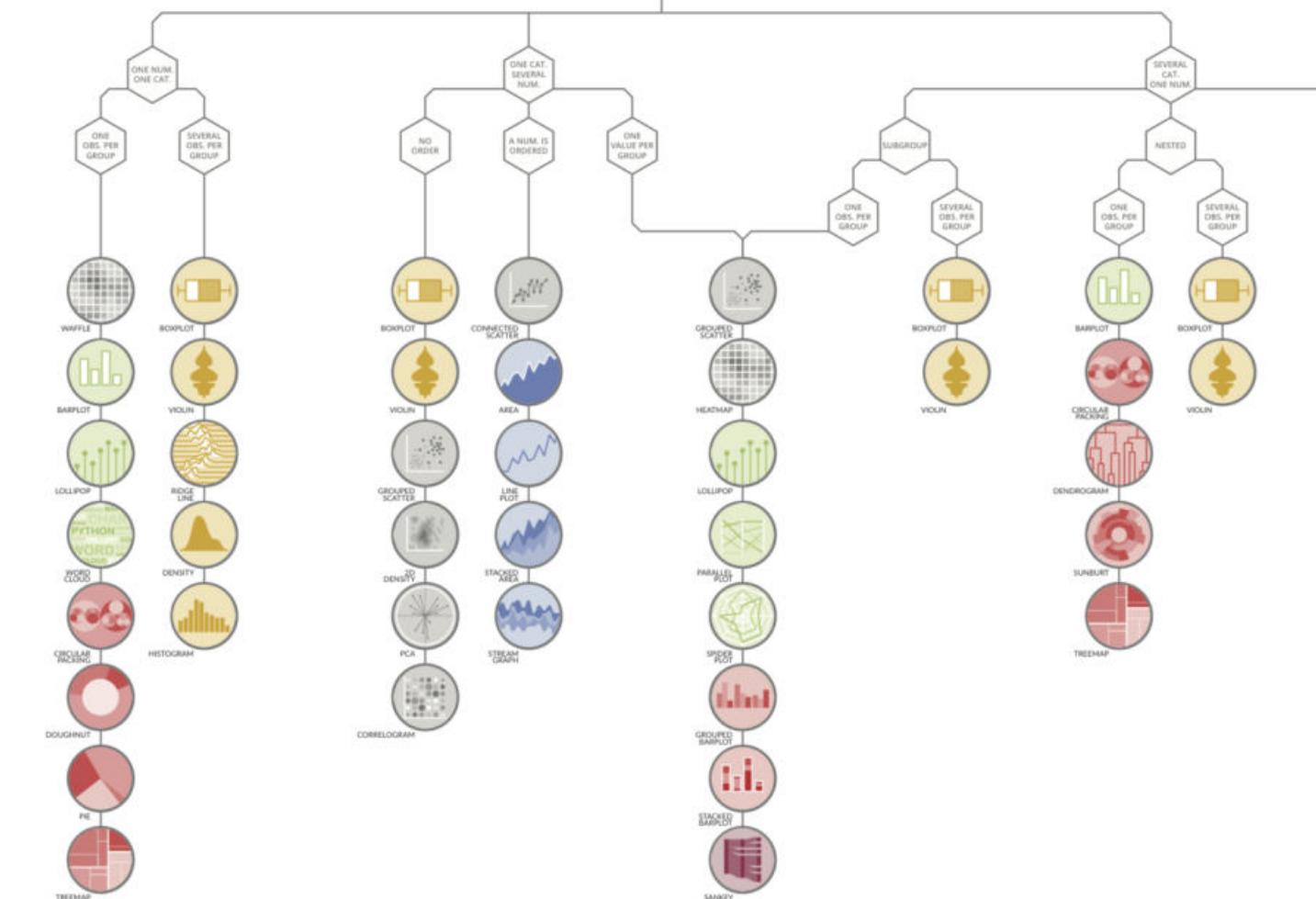
## MAP



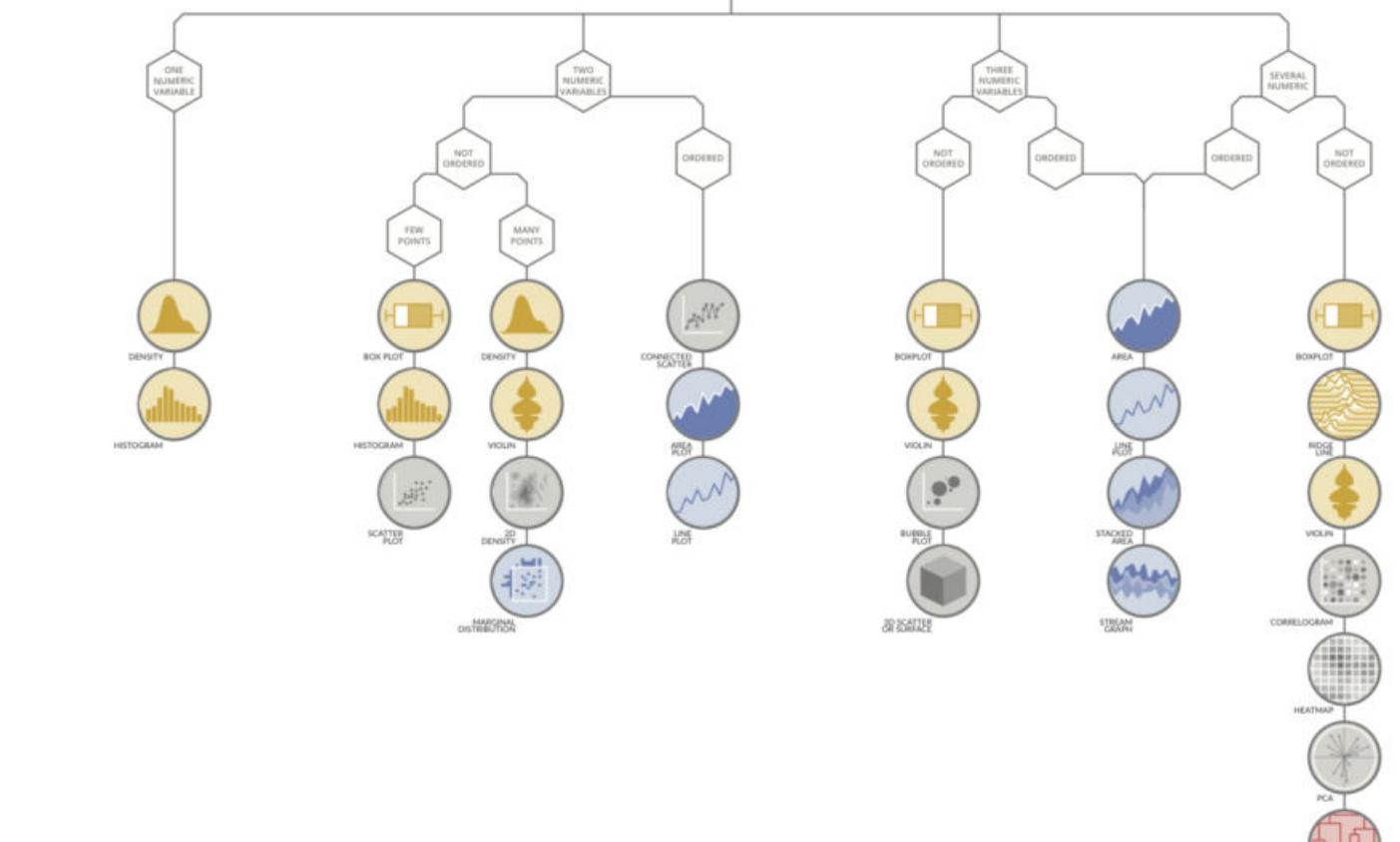
## TIME SERIES



## CATEGORIC AND NUMERIC



## NUMERIC



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



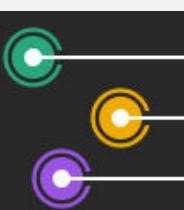
[cedricscherer.com](http://cedricscherer.com)



@CedScherer



z3tt



The screenshot shows a modal window from [data-to-viz.com](https://data-to-viz.com) with the title "BOXPLOT". At the top left is a yellow circular icon containing a boxplot. To the right is a close button (X). Below the icon is the title "BOXPLOT" in bold green capital letters. A subtitle "Summarize the distribution of numeric variables" follows. A "About" section provides a brief explanation of what a boxplot is. A "Common Mistakes" section lists three items: "Boxplot hides the sample size of each group, show it with annotation or box width.", "Boxplot hides the underlying distribution. Use jitter if low number of data points, or use violin with bigger data.", and "Order your boxplot by median can make it more insightful.". A "Code" section includes links to "R graph gallery", "Python gallery", "D3.js gallery", and "Flourish". A "Read More" section links to a dedicated page. Below the modal is a grid of 12 chart types, each with an icon and a name: Boxplot, Ridgeplot, Scatter, Connected scatter, Density 2d, Barplot, Lollipop, Circular Barplot, Treemap, Dendrogram, Circular packing, and Sunburst.

SSIBILITIES

presented in this website.

Part of a whole Evolution Map Flow

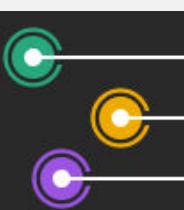
Boxplot Ridgeplot Scatter

Connected scatter Density 2d Barplot

Lollipop Circular Barplot Treemap

Dendrogram Circular packing Sunburst

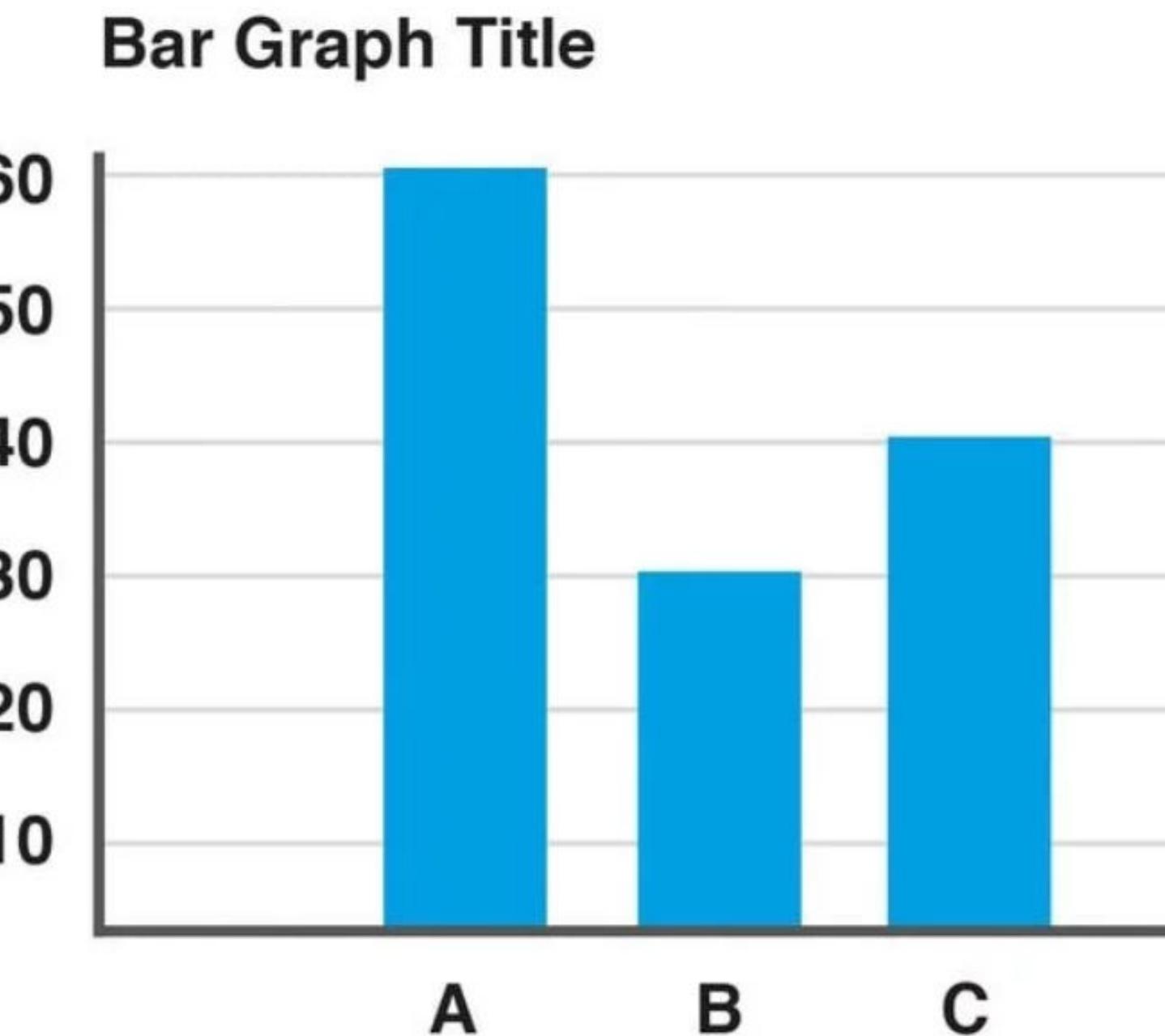
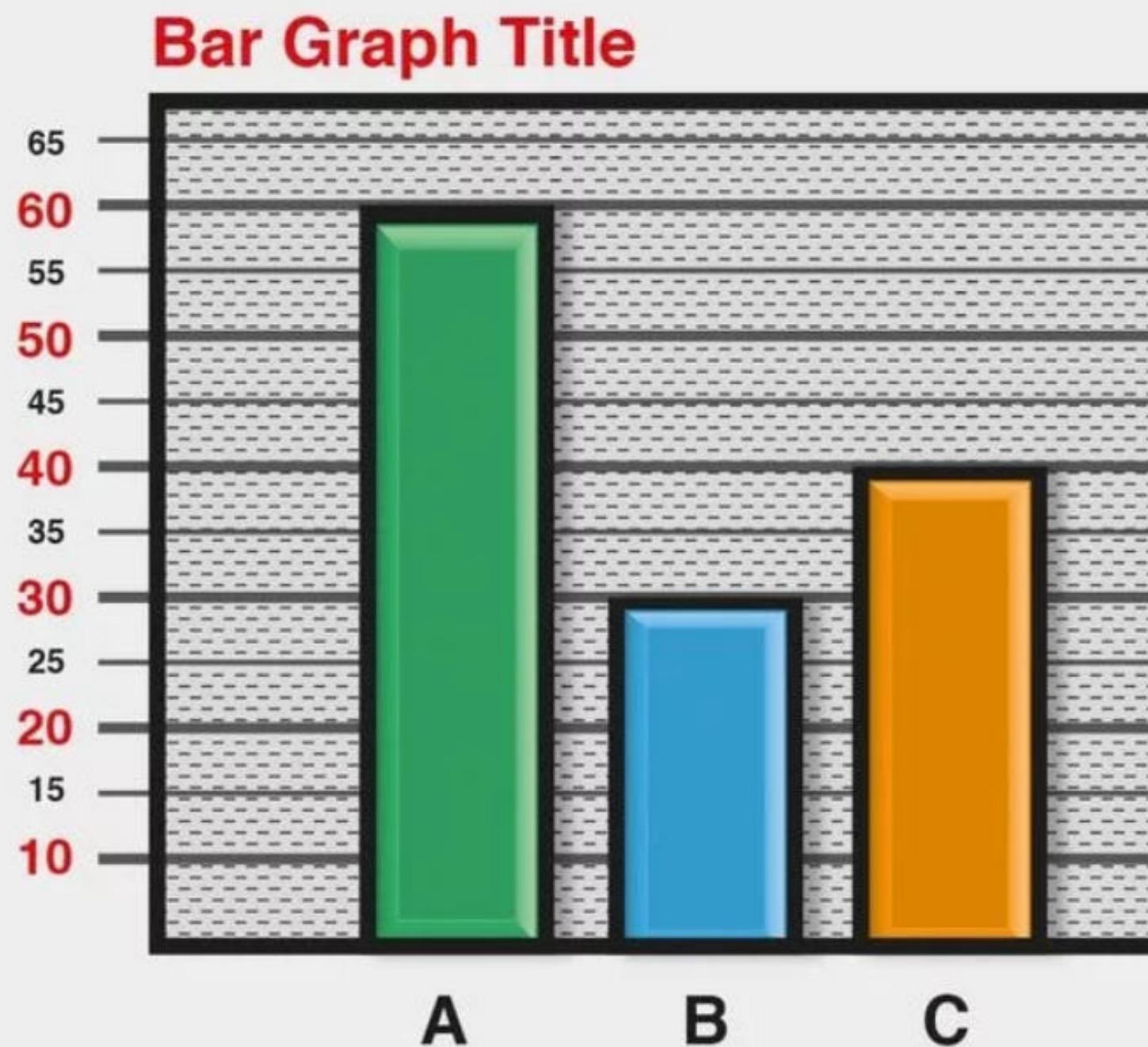
Quelle: [data-to-viz.com](https://data-to-viz.com)





Guide the  
View (e8)

# „Chart-Entrümpelung“ (Decluttering)



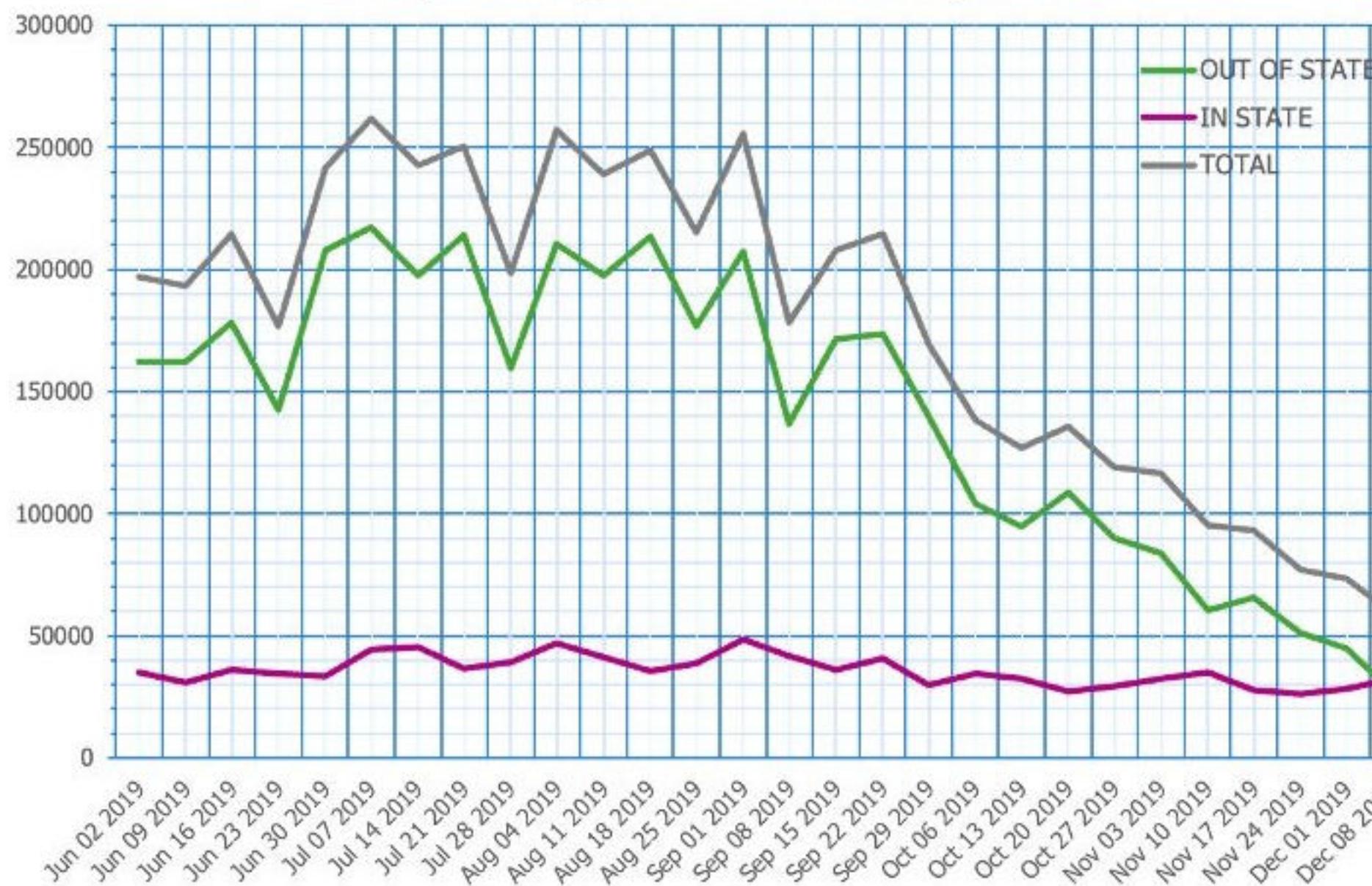
Quelle: [canva.com](https://canva.com)



# „Chart-Entrümpelung“

## BEFORE: SPOOKY SKELETON

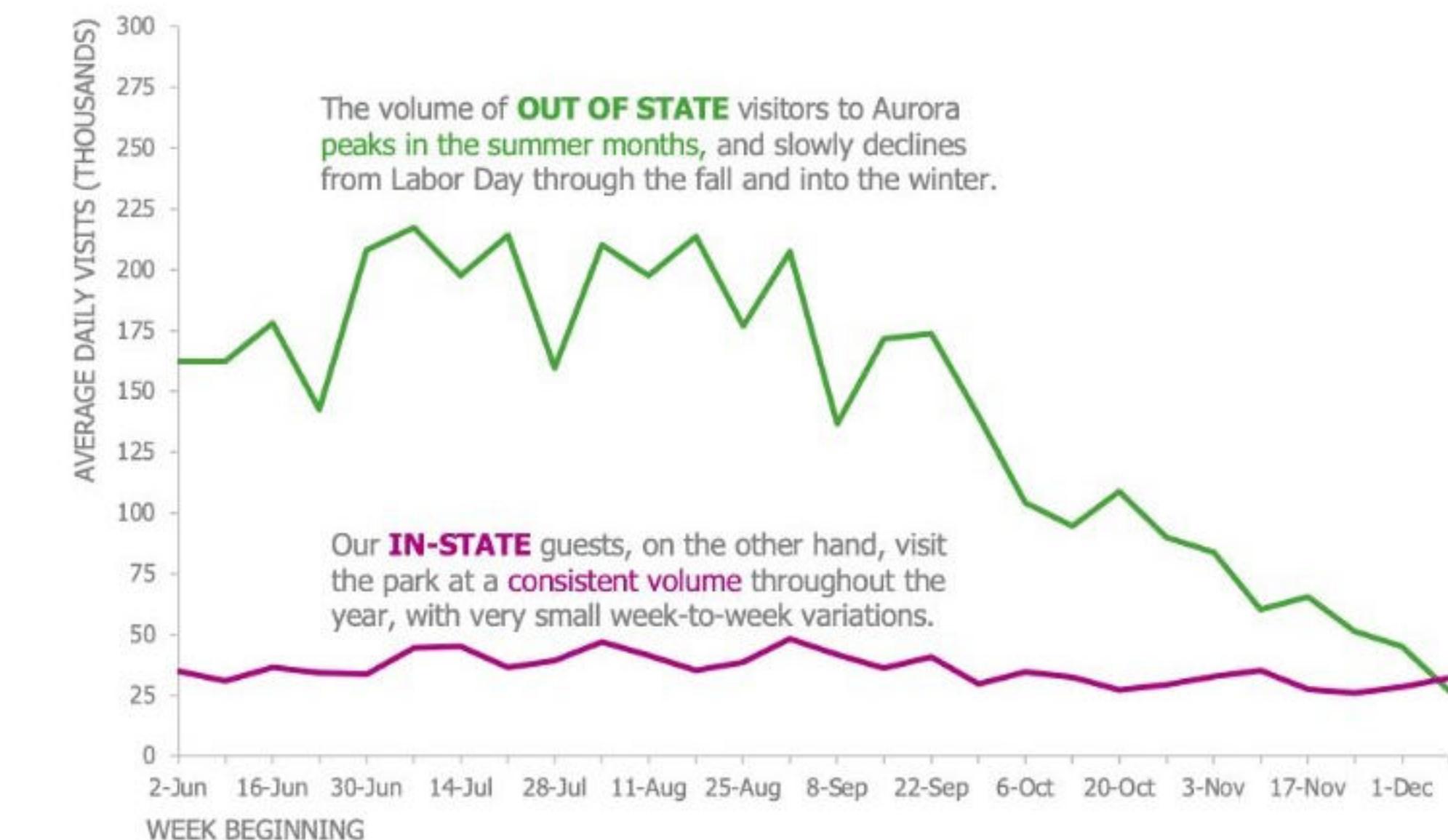
Daily Average Park Visitors By Week



## AFTER: GOOD BONES

Daily visitors to Aurora Park in summer/fall 2019

VALUES ARE CALCULATED WEEKLY AS A 7-DAY AVERAGE

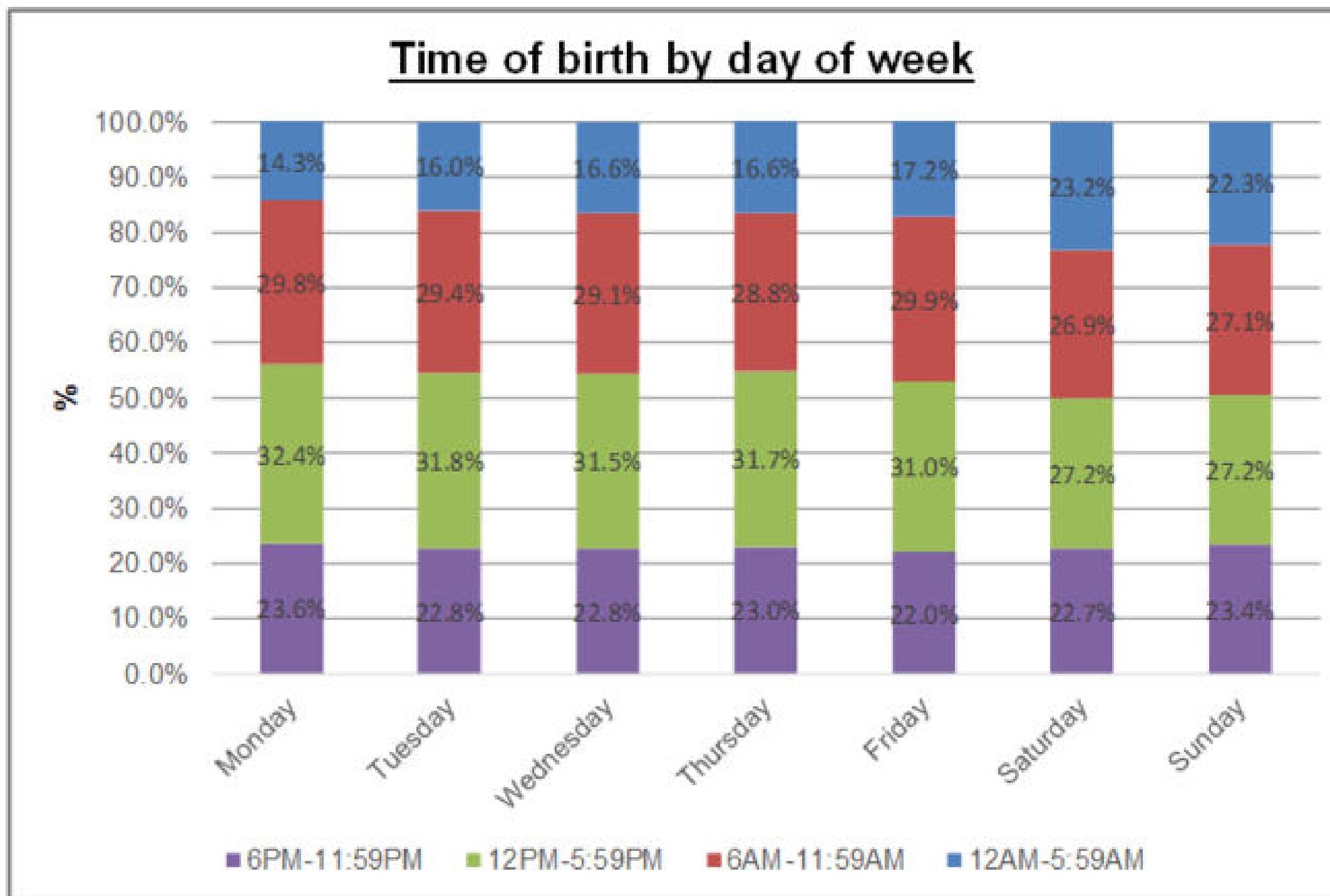


Quelle: „Your graph Skeleton Shouldn't Be Scary“ von Mike Cisneros (Storytelling With Data)



# „Chart-Entrümpelung“

BEFORE

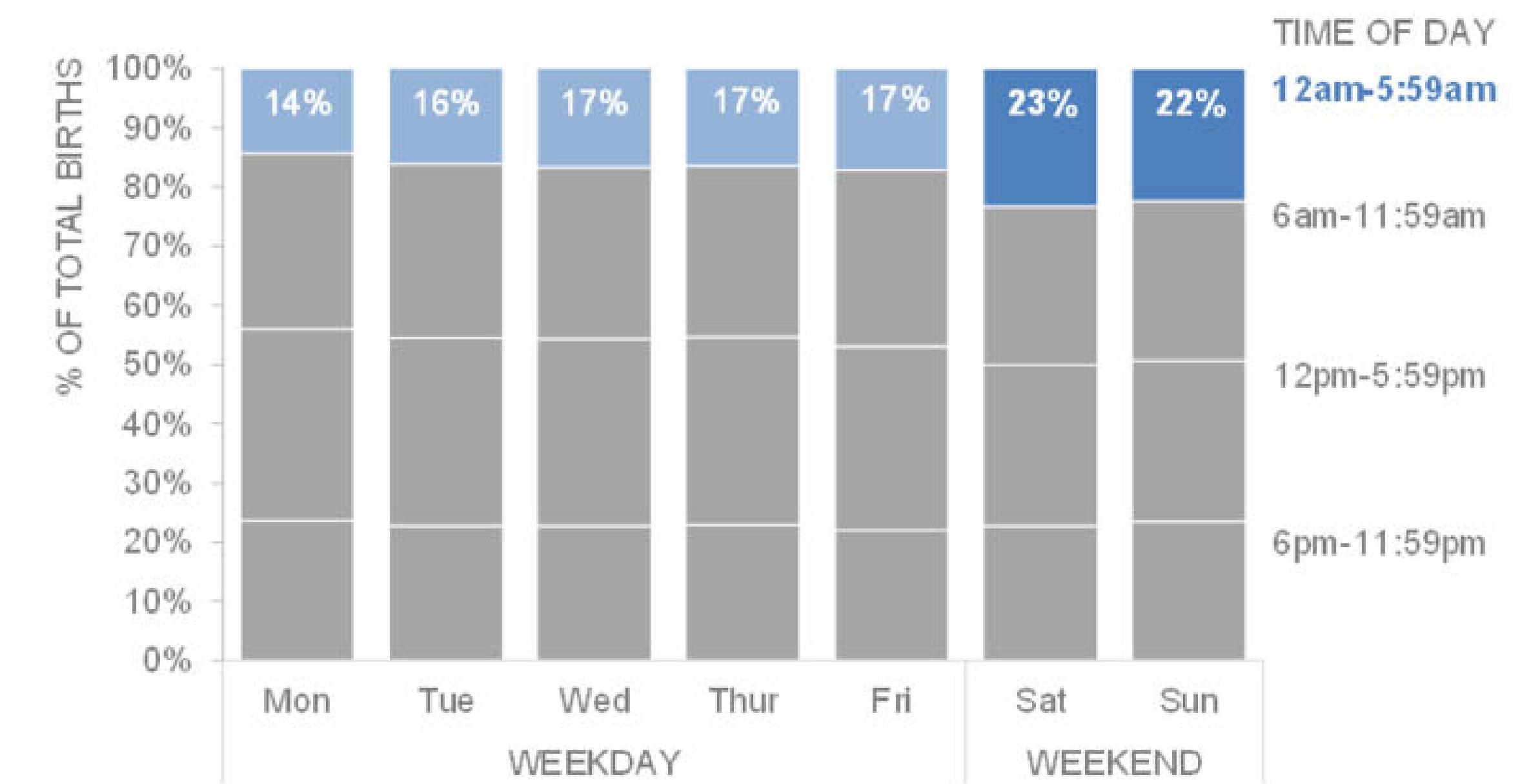


Data source: CDC (National Vital Statistics Reports, Vol. 67, No. 1, January 31, 2018)

AFTER

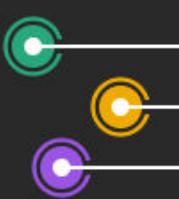
**When babies are born**

**Weekend deliveries are more likely to be in early morning, compared to weekdays**



Data source: CDC (National Vital Statistics Reports, Vol. 67, No. 1, January 31, 2018)

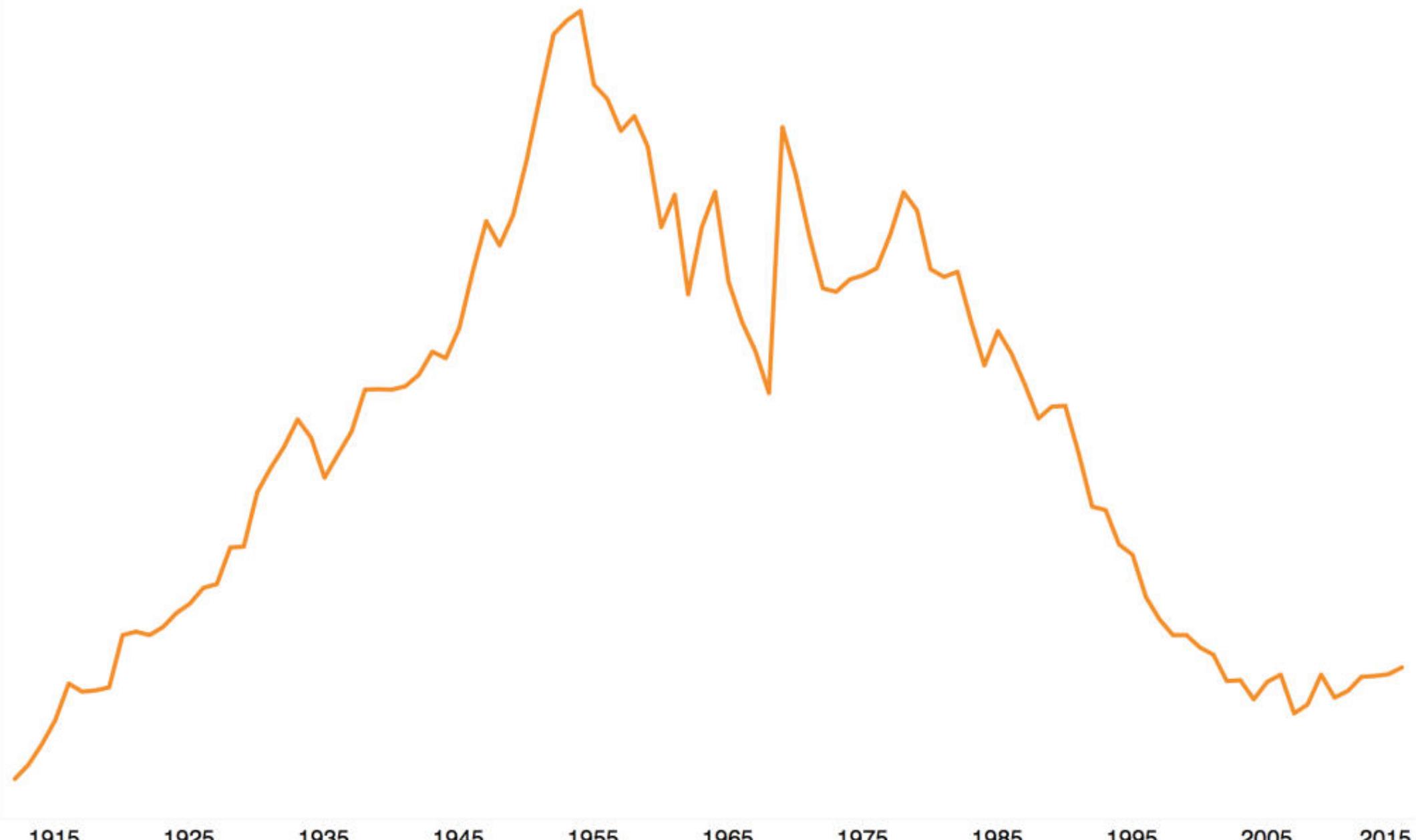
Quelle: „*Declutter! (and Question Default Settings)*“ von Elizabeth Ricks (*Storytelling With Data*)



# Die Wirkung von Annotationen

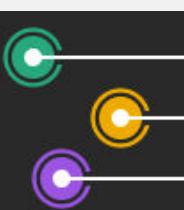
Rise and Fall of the name **Neil** in the USA  
Births 1912-2015

Source: data.gov



Visualisation: [@theneilrichards](#)

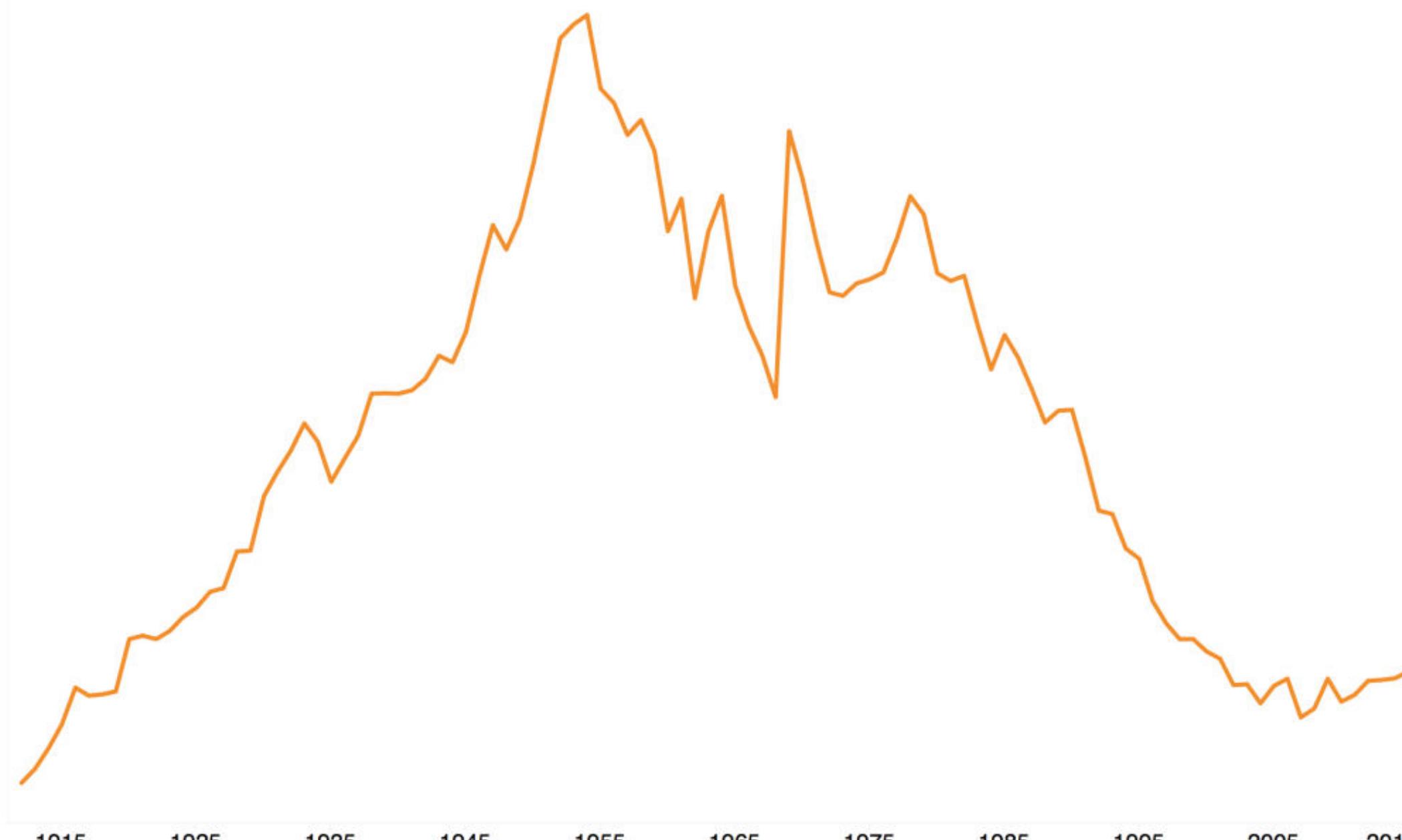
“Is white space always your friend?” von Neil Richards



# Die Wirkung von Annotationen

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

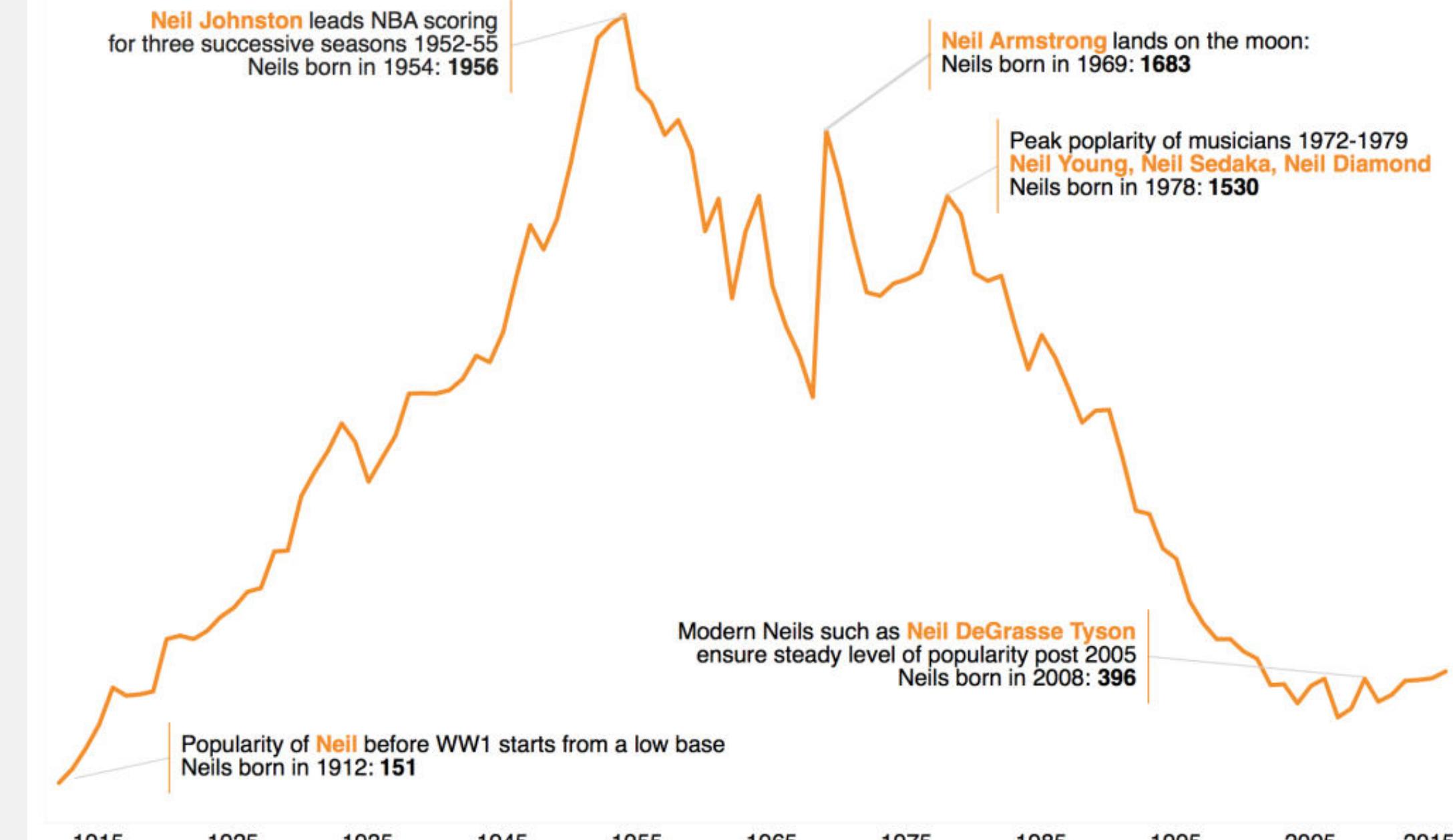
Source: data.gov



Visualisation: [@theneilrichards](#)

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

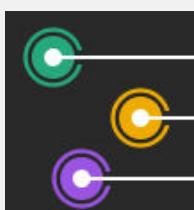
Source: data.gov



Visualisation: [@theneilrichards](#)

#SWDChallenge

"Is white space always your friend?" von Neil Richards



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

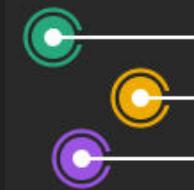
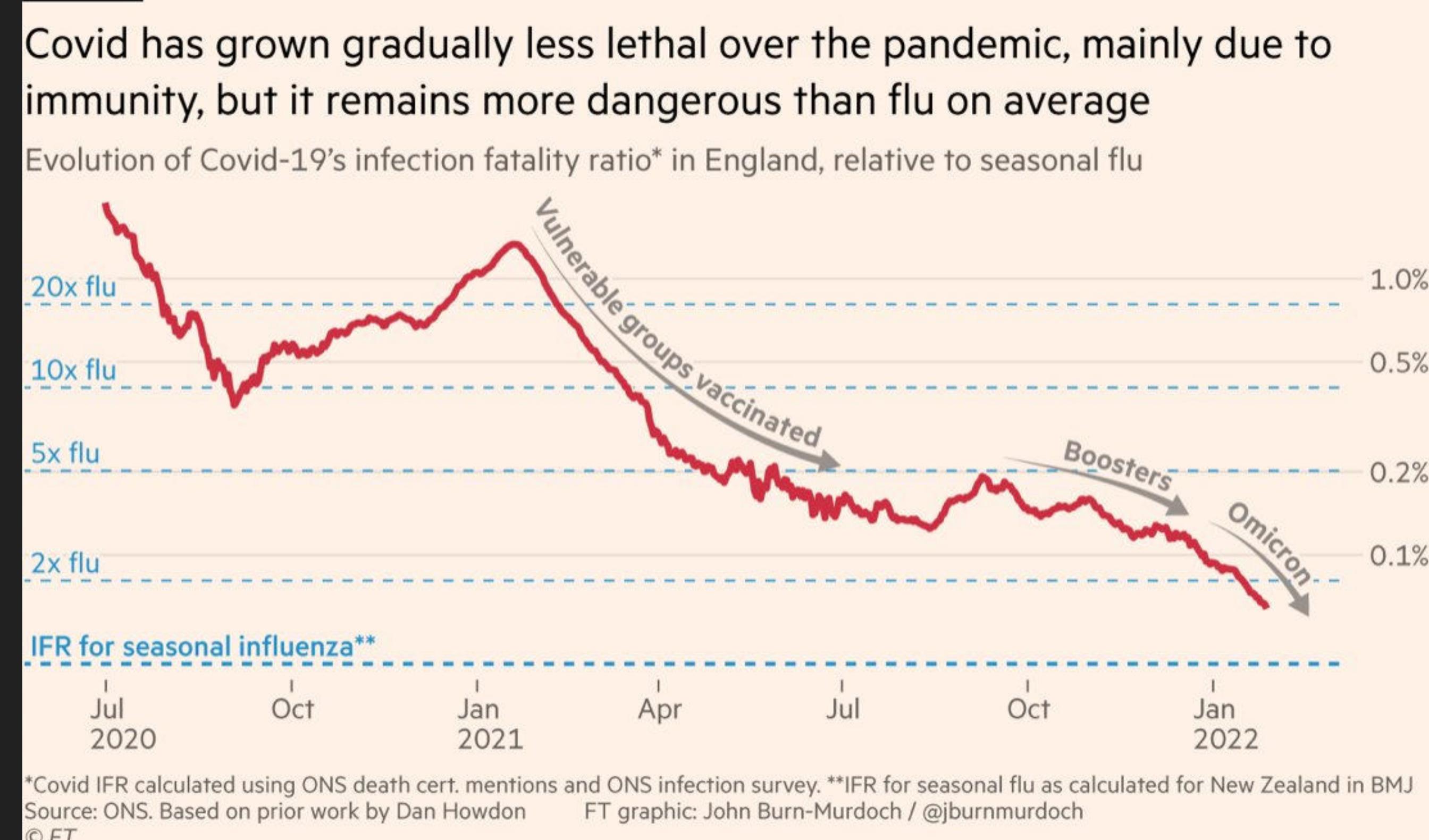
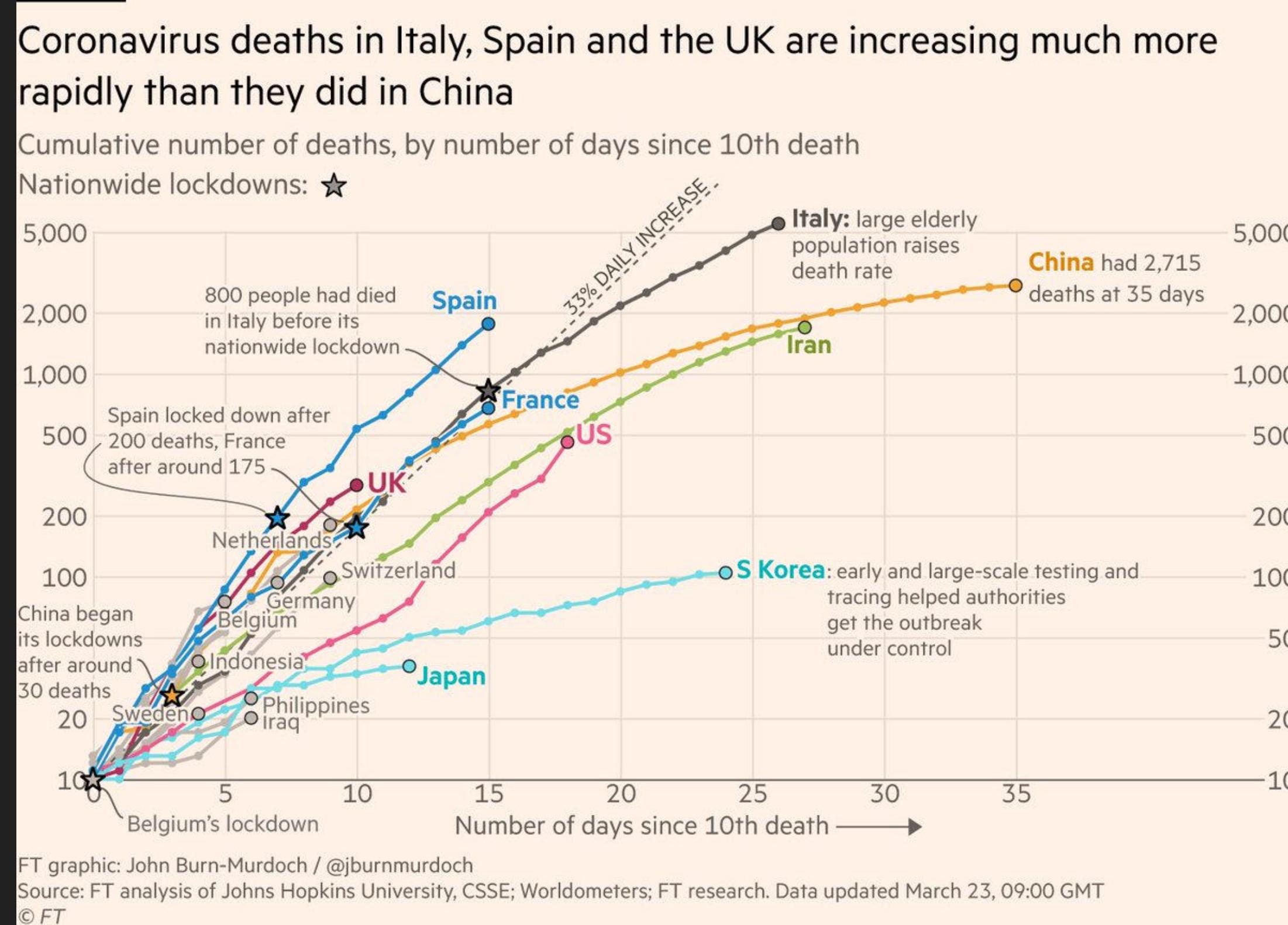


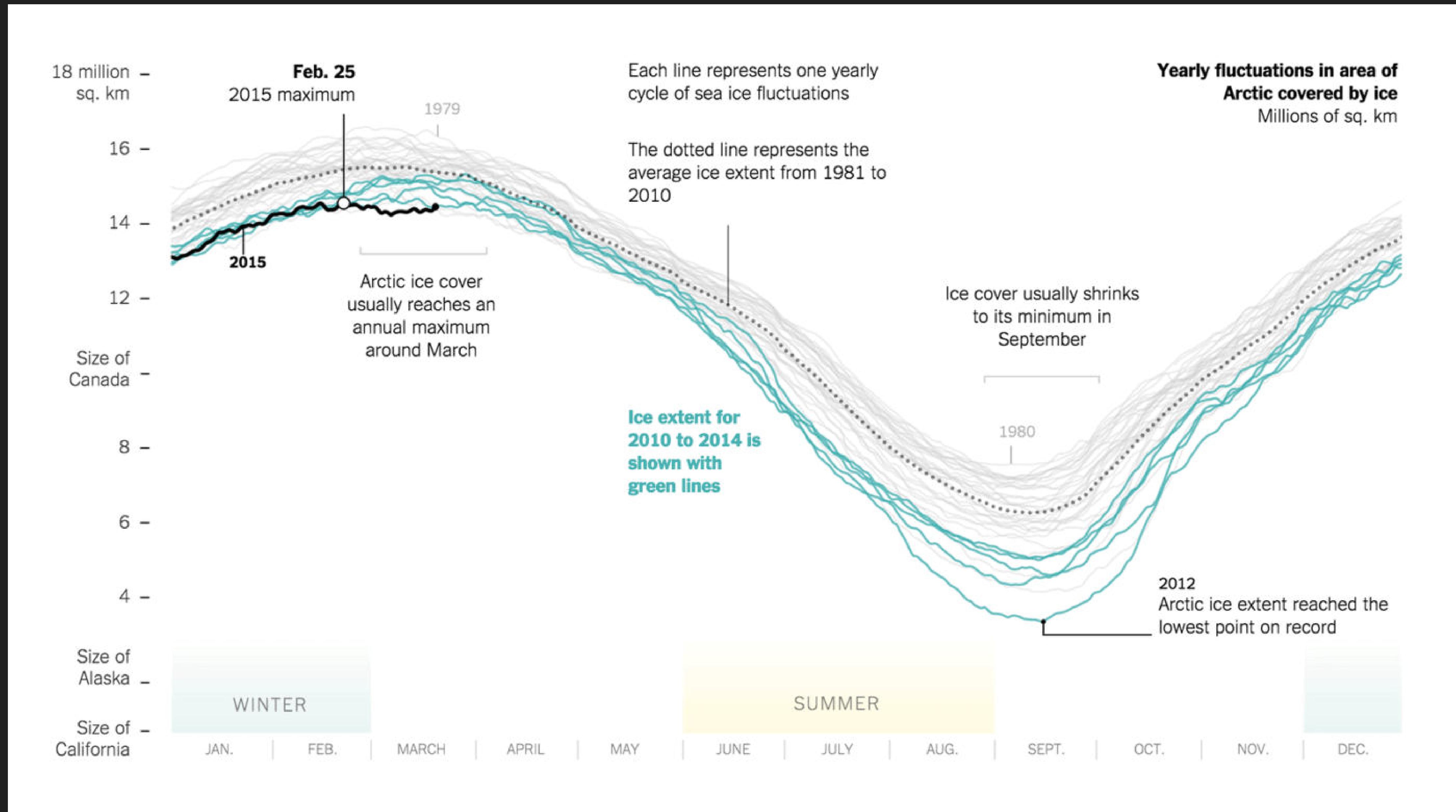
Zeitreihe mit Beschriftungen von William Playfair from "The Commercial and Political Atlas and Statistical Breviary" (1786)



“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

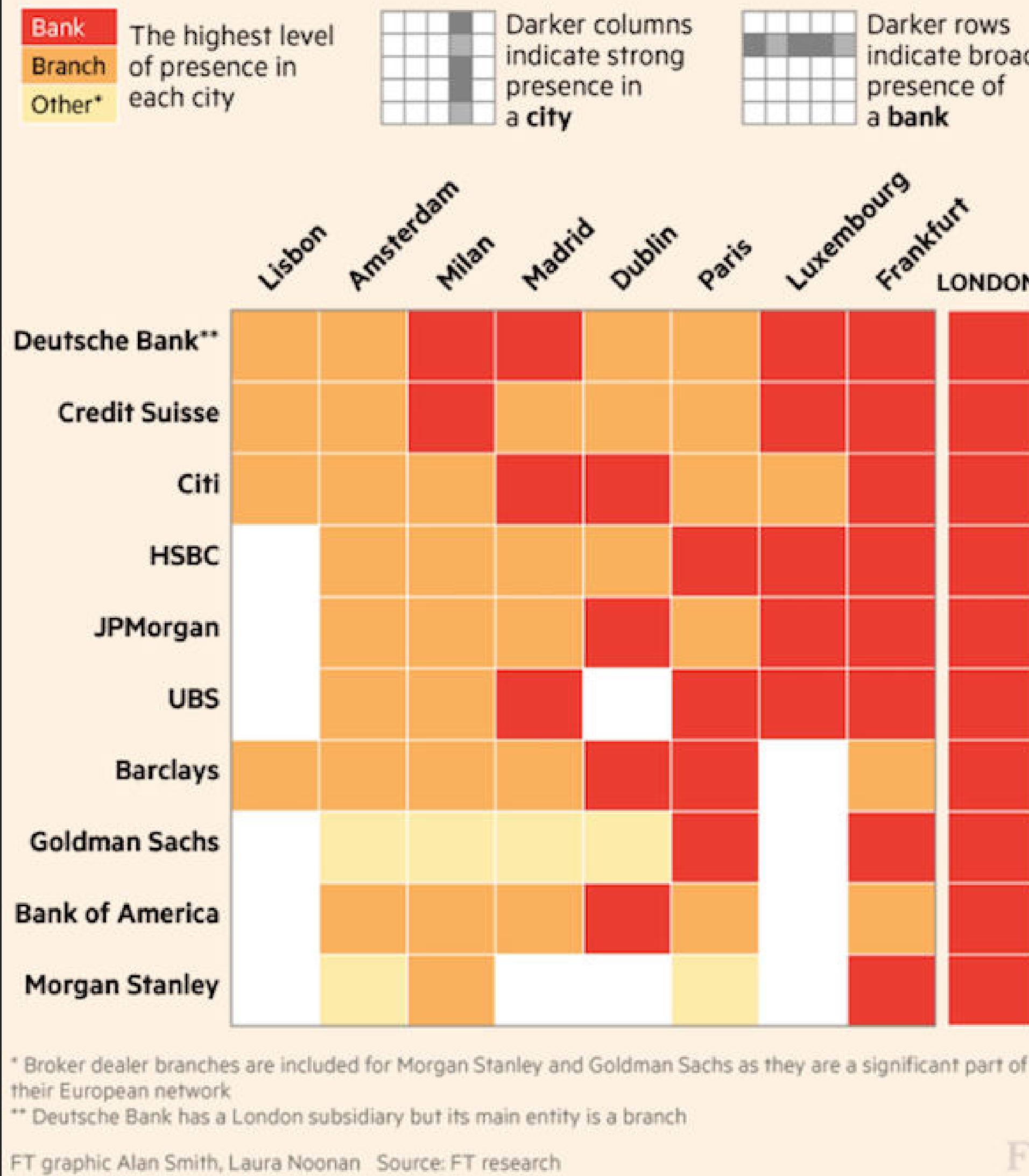




*“Yearly Fluctuations in Area of Arctic Covered by Ice” von Derek Watkins (New York Times)*



## The Brexit banking matrix: The contenders lining up for London's crown



*“Frankfurt vies for UK banking jobs post-Brexit” von Alan Smith und Laura Noonan (Financial Times)*



# Supplementary supermarket shopping

1

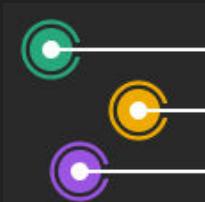
People who do most of their food/grocery shopping at this supermarket...

2

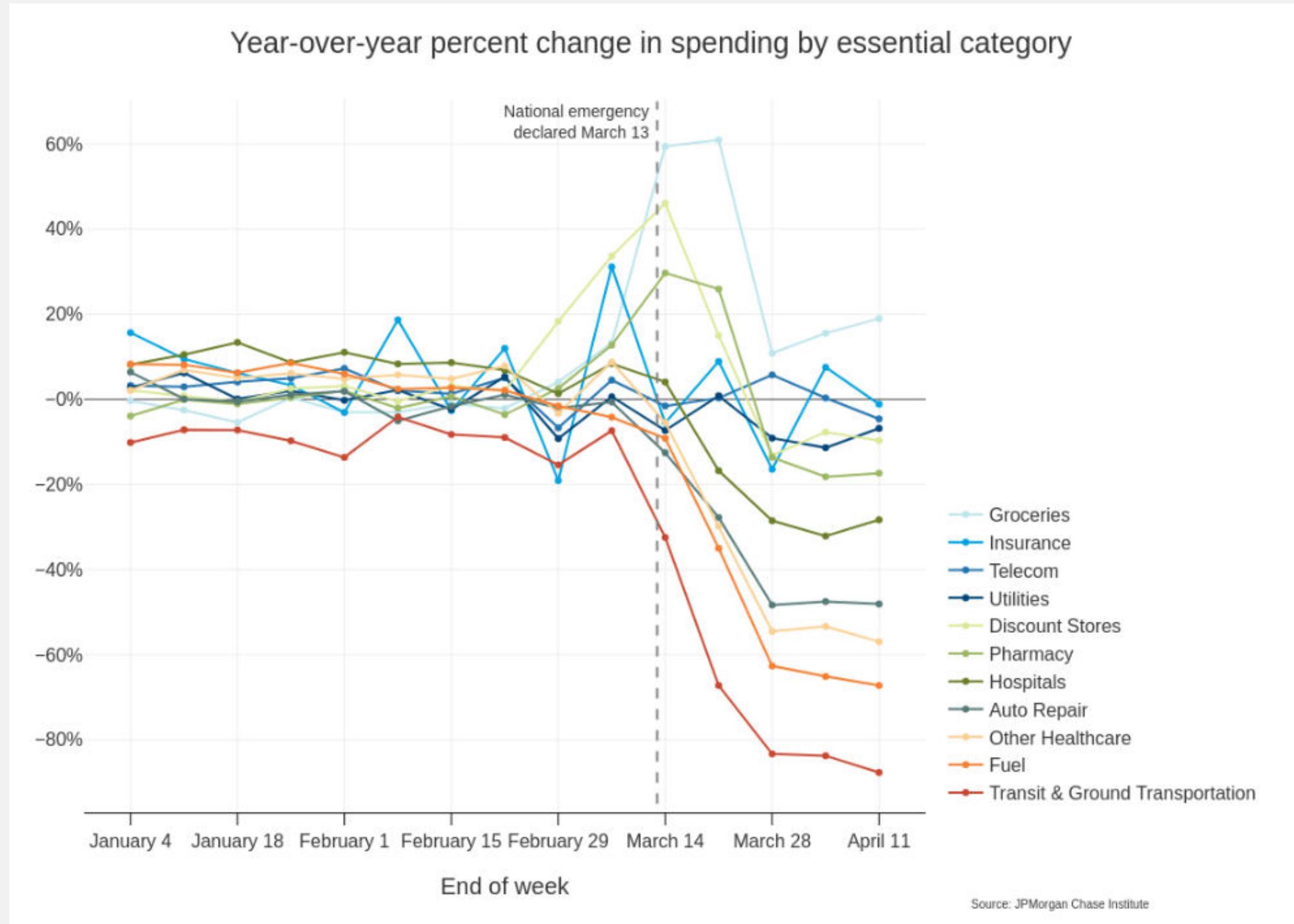
...also regularly shop at this supermarket

So, for instance,  
27% of people  
who do most of  
their grocery  
shopping at Aldi  
also regularly  
shop at  
Morrisons.

	Aldi	Asda	Iceland	Lidl	Marks & Spencer	Morrisons	Sainsbury's	Tesco	The Co-operative	Waitrose
Aldi		36	26	34	17	33	23	28	22	16
Asda	38		34	32	19	29	20	27	20	11
Iceland	18	22		20	12	19	14	17	12	7
Lidl	25	24	26		13	24	20	24	21	17
Marks & Spencer	9	12	9	9		17	28	17	17	37
Morrisons	27	26	21	25	22		20	22	20	13
Sainsbury's	26	25	20	33	47	29		32	31	39
Tesco	45	39	42	44	43	37	42		42	40
The Co-operative	15	15	11	17	18	17	20	19		20
Waitrose	5	5	6	9	30	7	21	11	14	



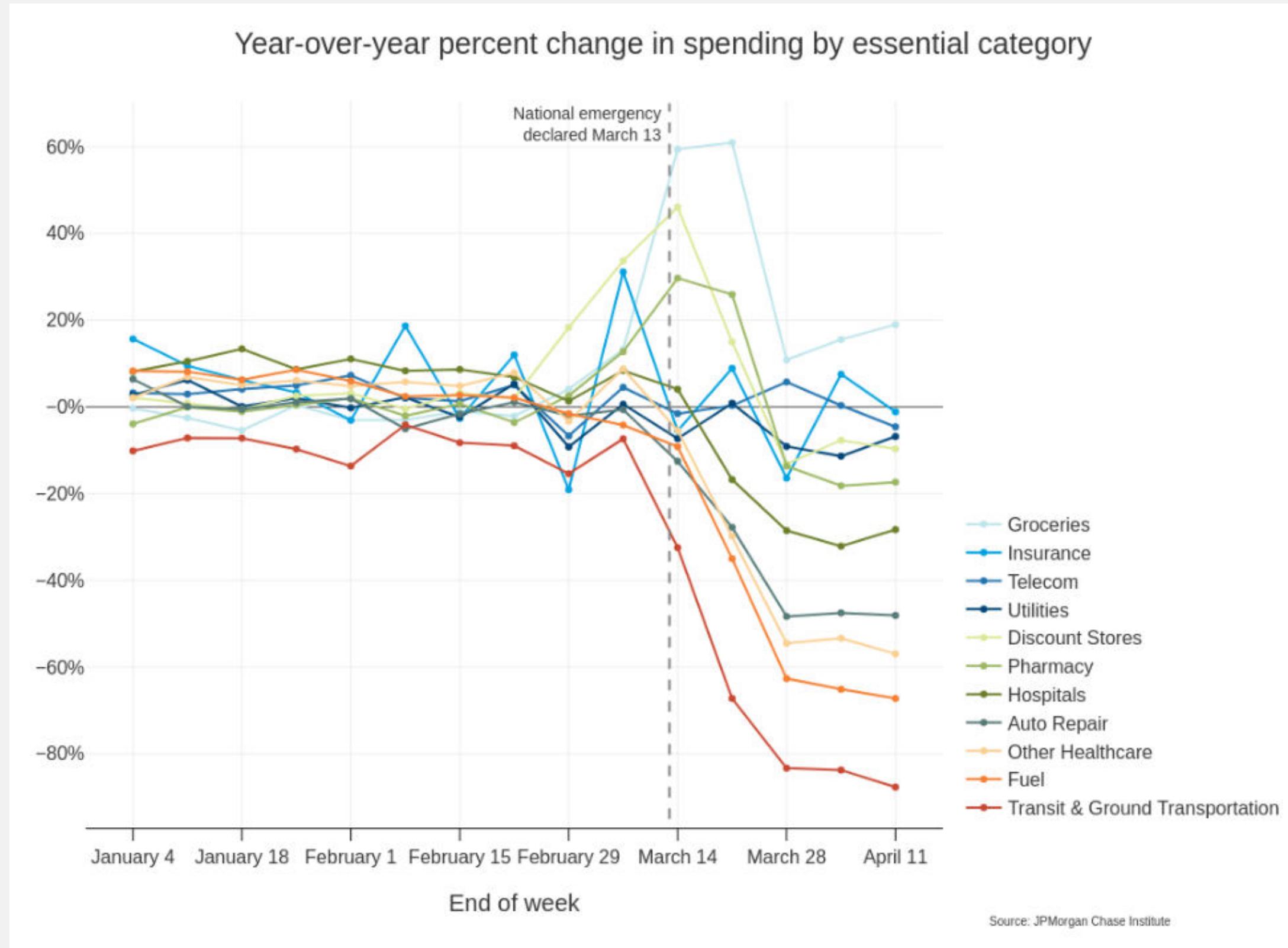
# Die Wirkung von „Small Multiples“



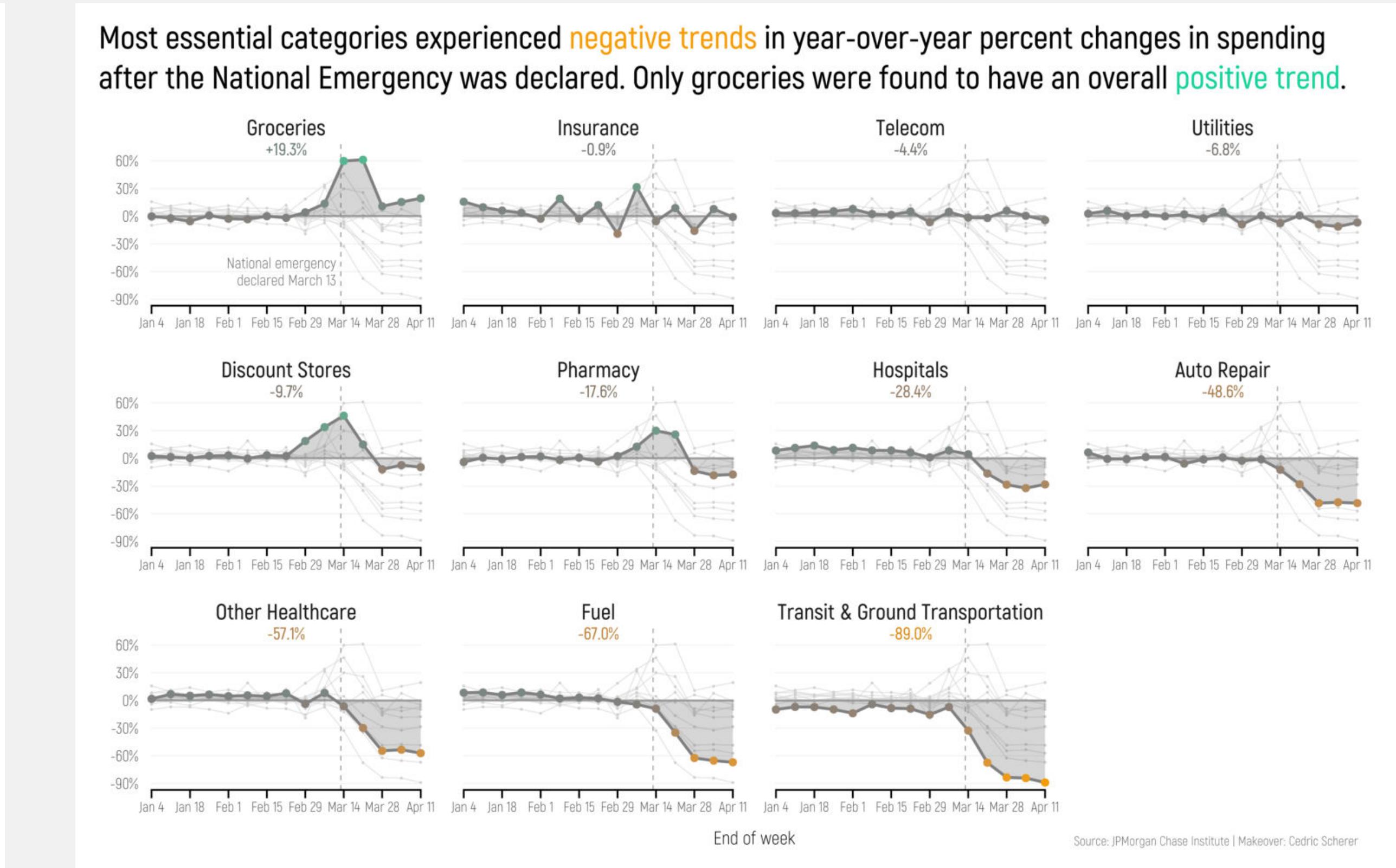
Original Grafik vom JPMorgan Chase Institut



# Die Wirkung von „Small Multiples“



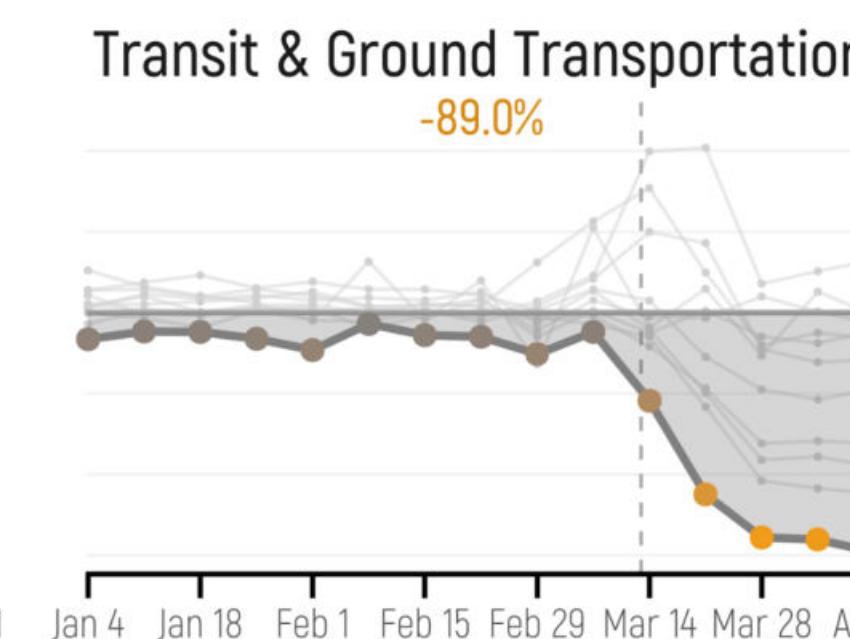
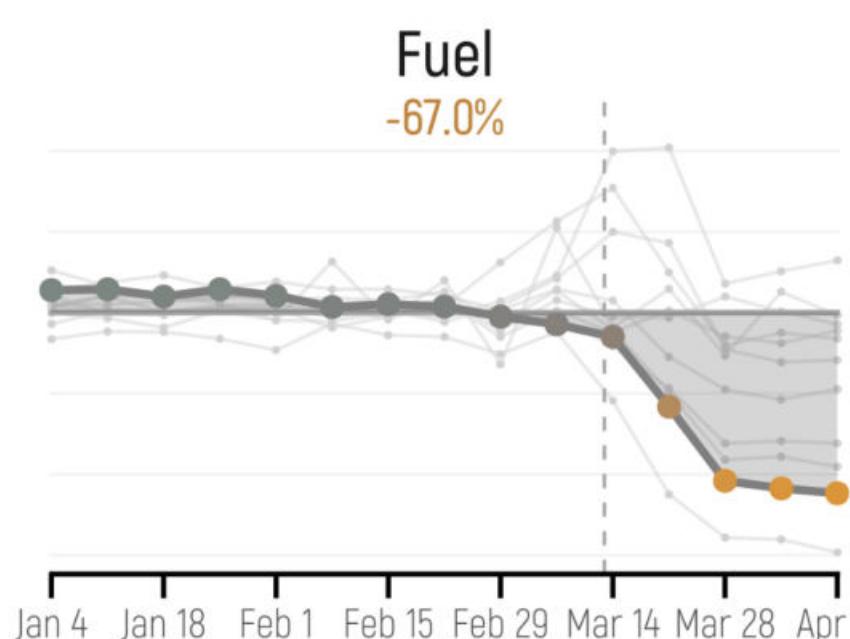
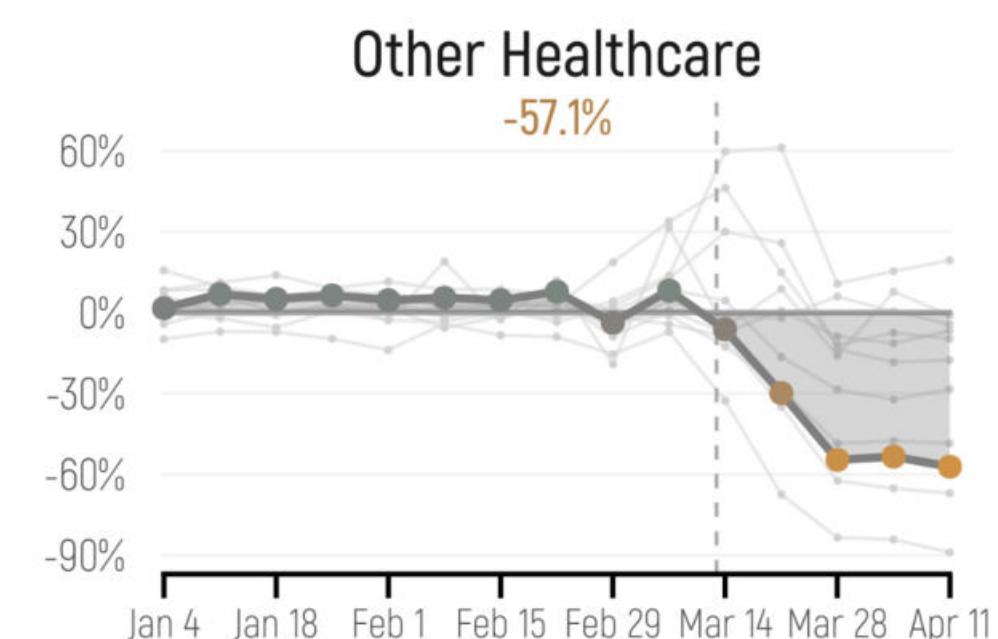
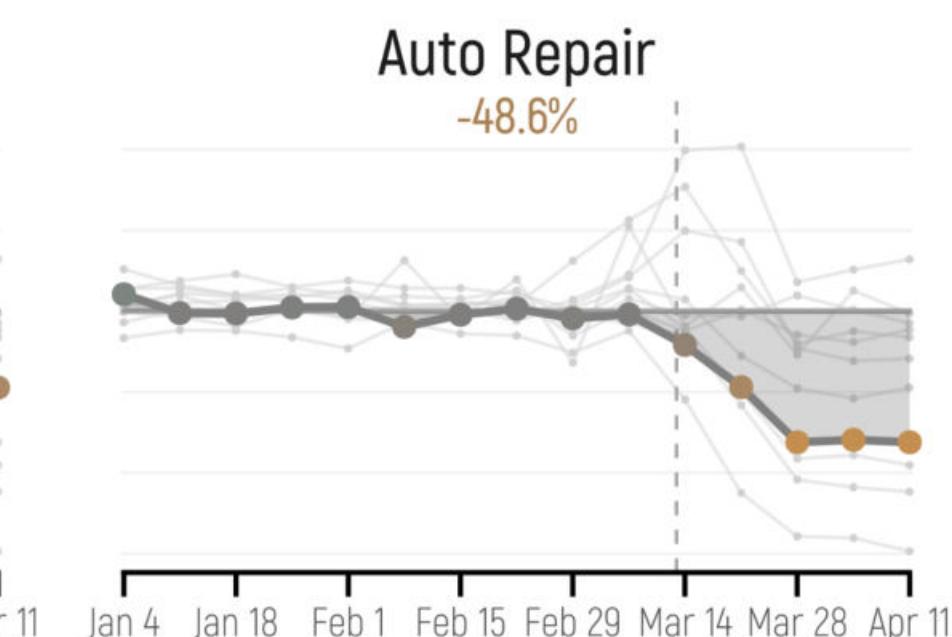
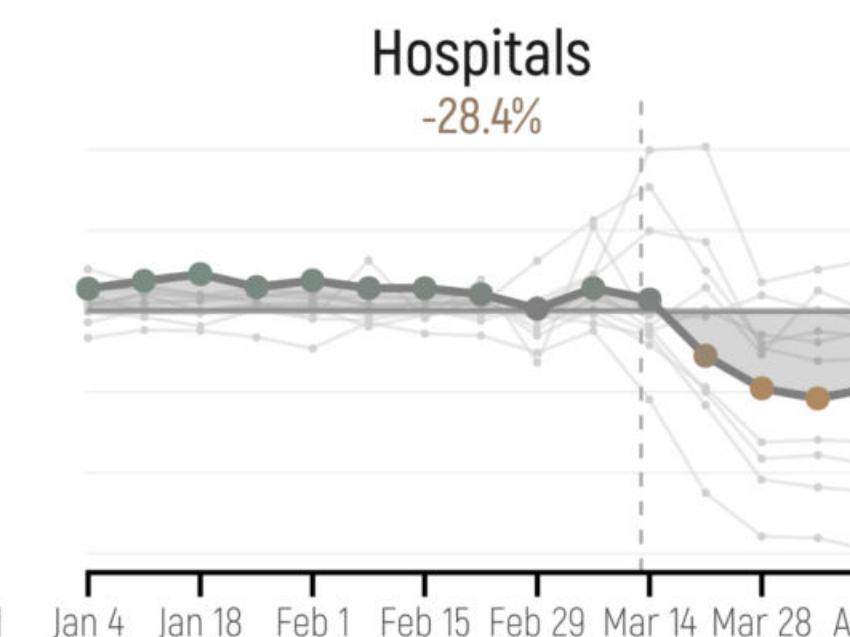
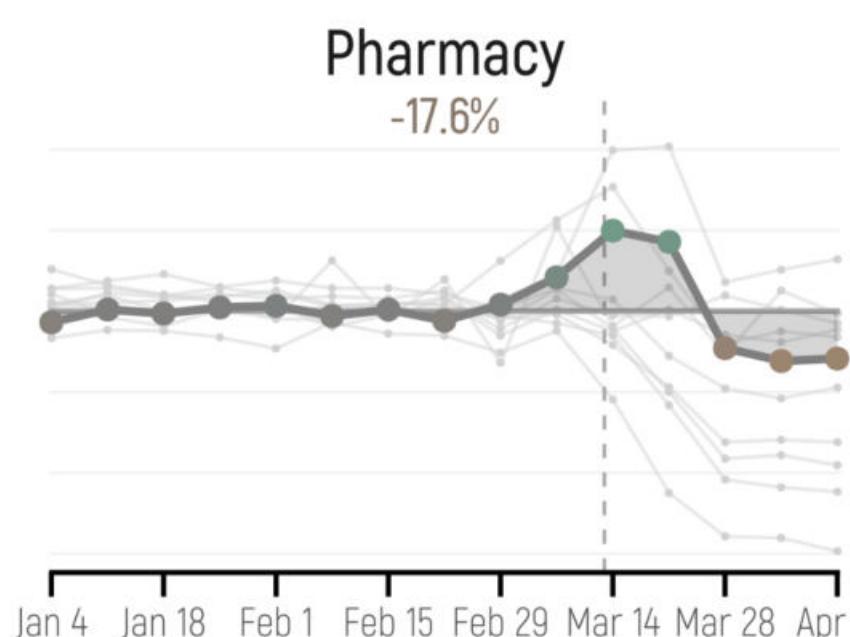
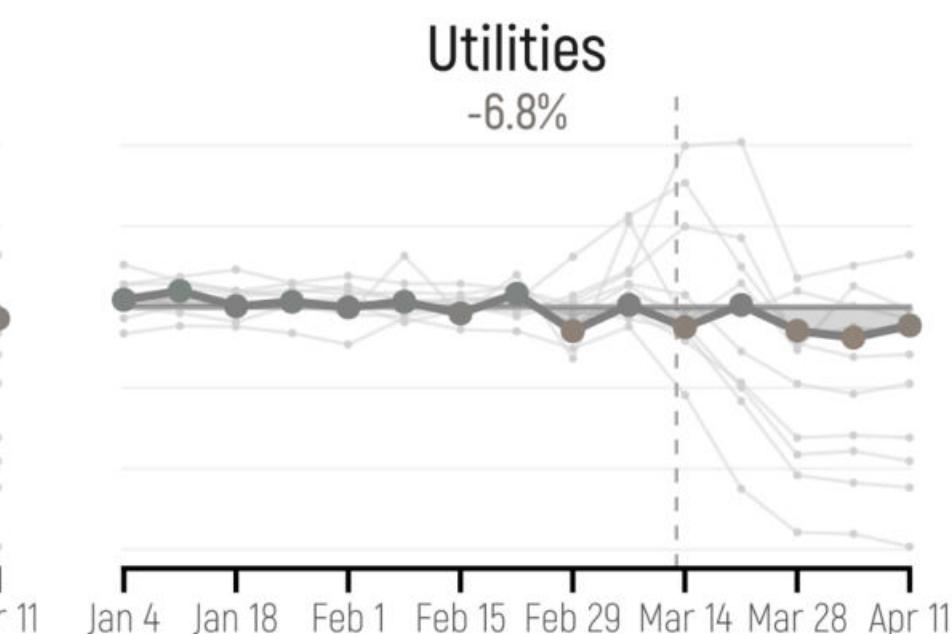
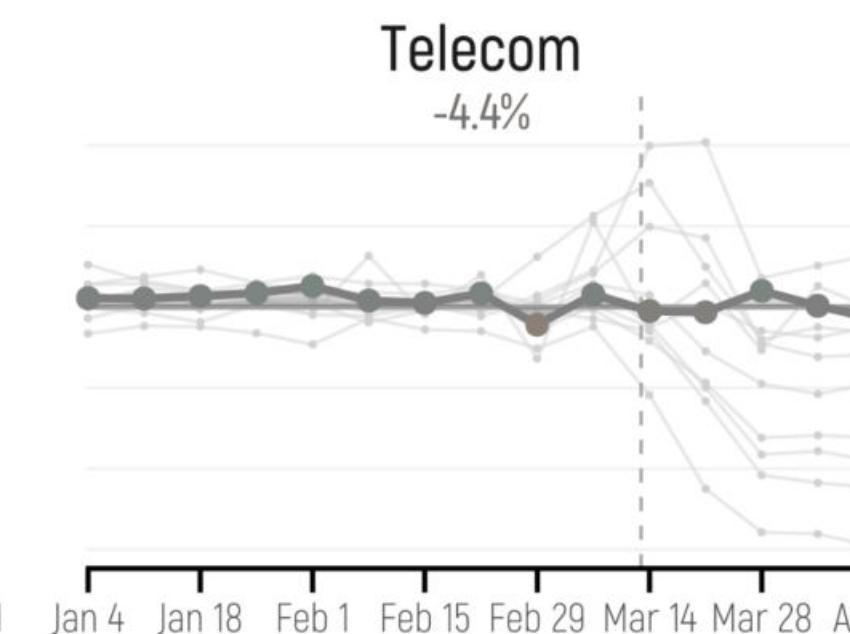
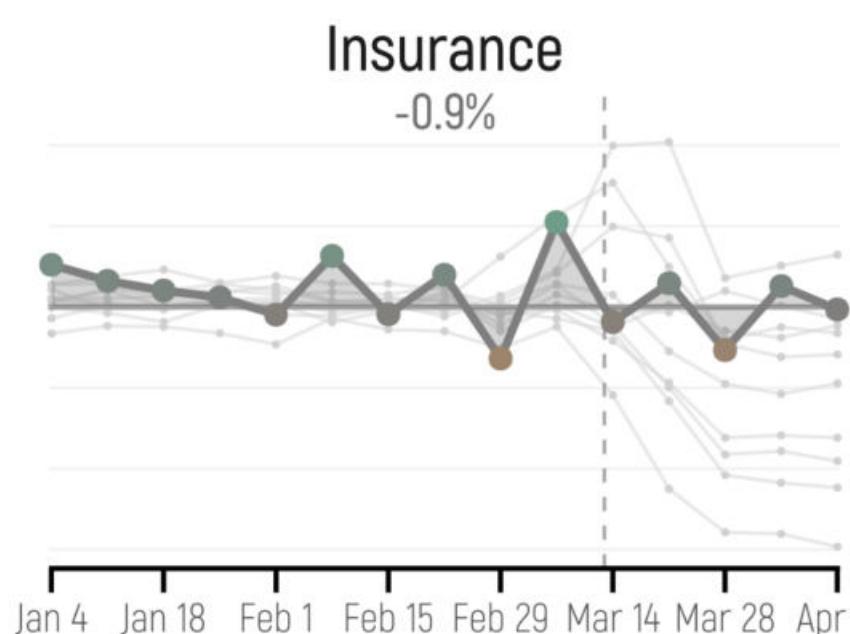
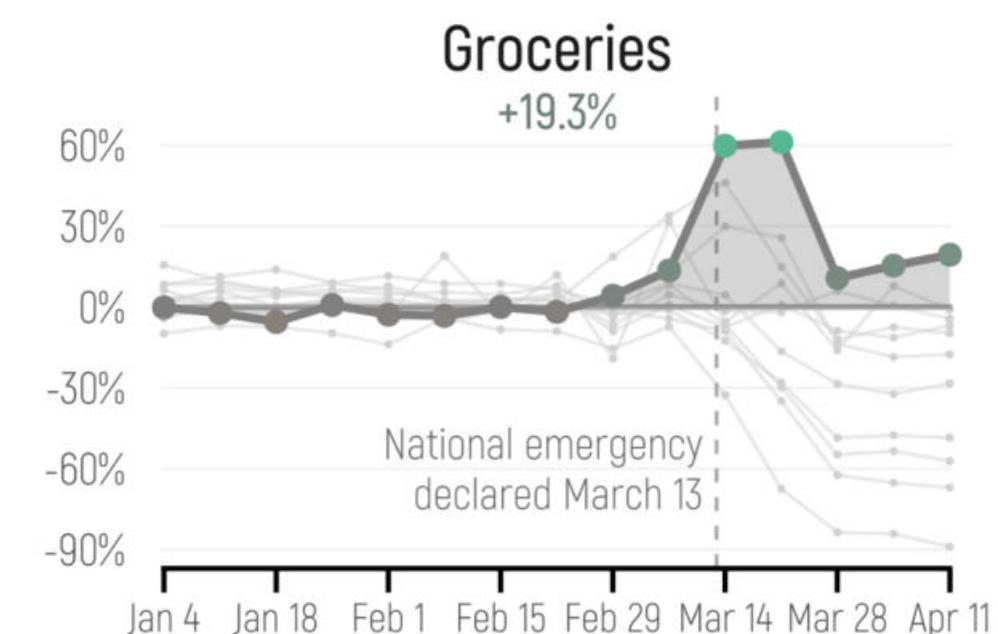
Originalgrafik vom JPMorgan Chase Institut



Umgestaltung mit “Small Multiples”

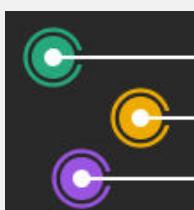


Most essential categories experienced **negative trends** in year-over-year percent changes in spending after the National Emergency was declared. Only groceries were found to have an overall **positive trend**.



End of week

Source: JPMorgan Chase Institute | Makeover: Cedric Scherer



## 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."



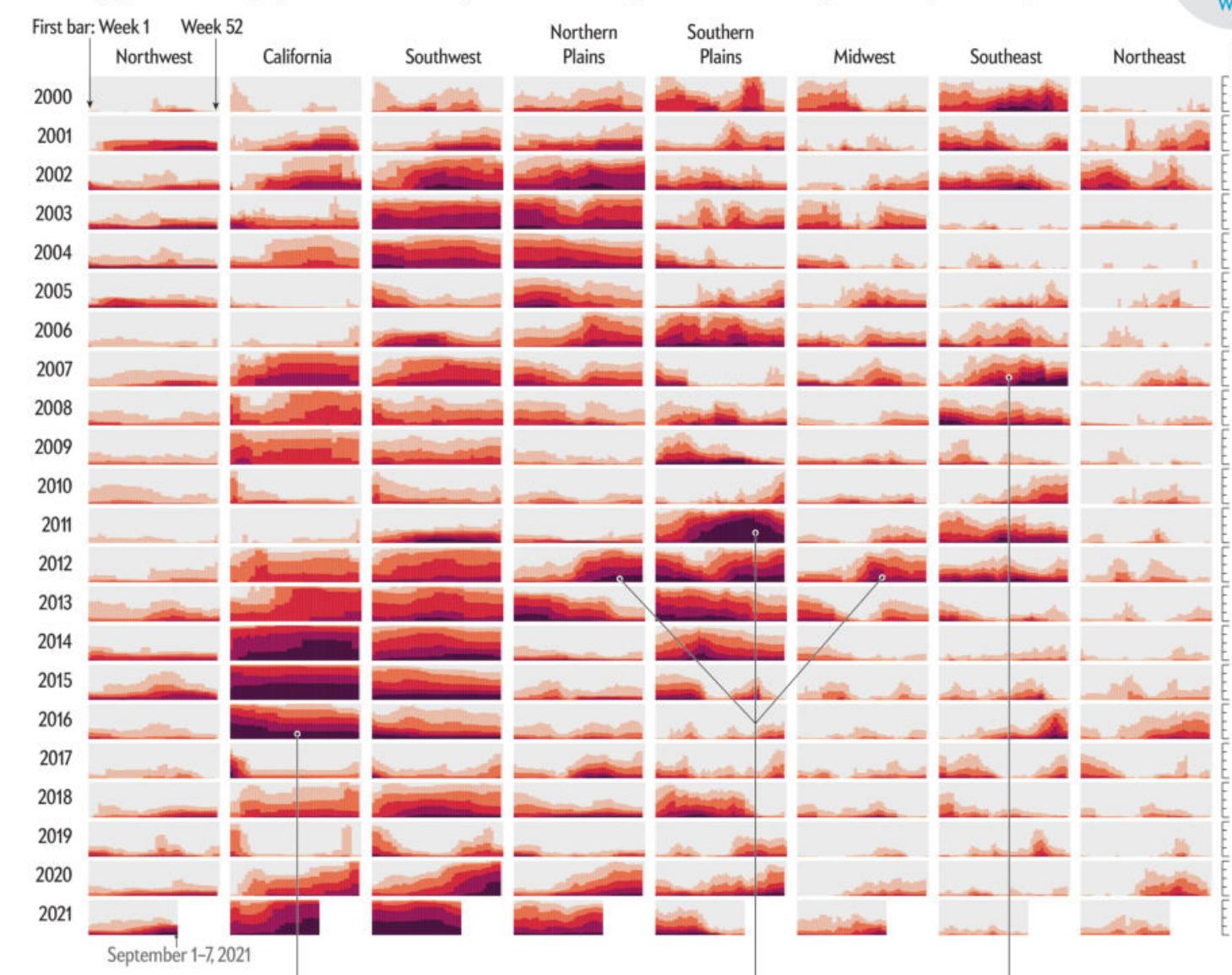
Percent of Region That Experienced Drought Each Week

100%

0%

### Drought Extent and Intensity by Region over Time

Category: Abnormally Dry Moderate Drought Severe Drought Extreme Drought Exceptional Drought

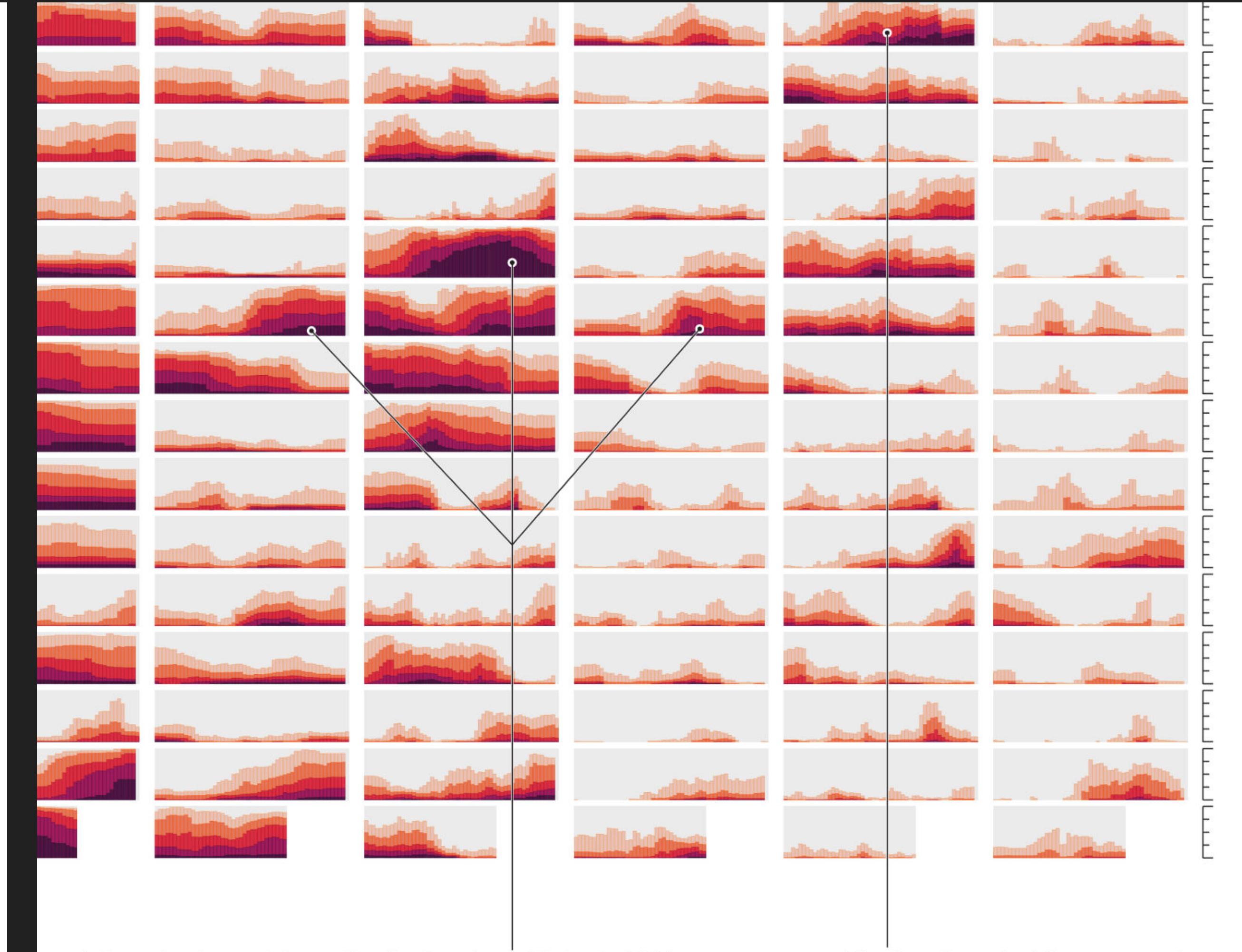


California experienced its hottest drought in recorded history from 2012 to 2016. A warming climate makes the atmosphere thirstier, which increases evaporation and boosts drought.

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.

Source: 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.

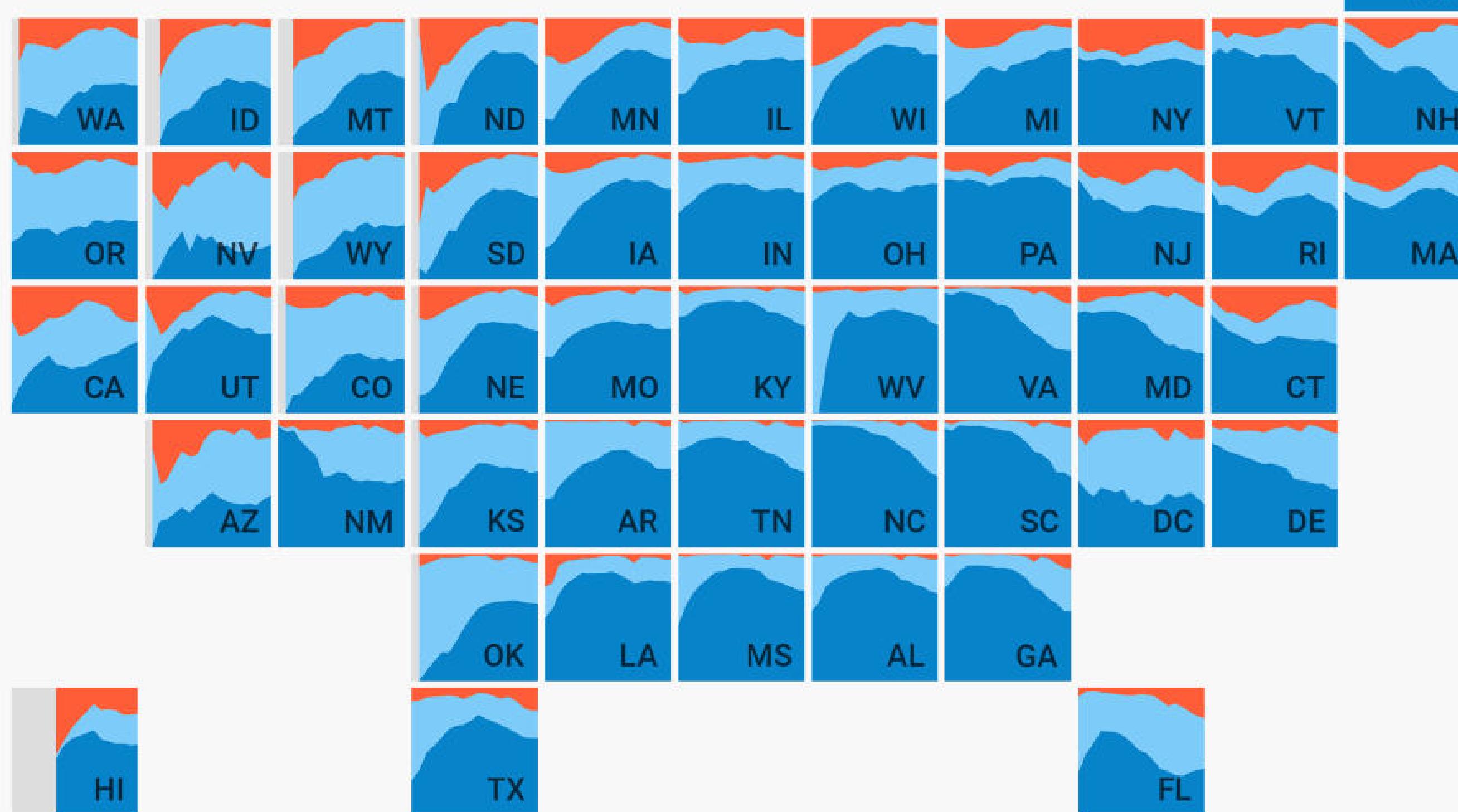
The Southeast's driest year to date was 2007, when only 31.85 inches of rain fell in Atlanta, 62 percent of its average yearly rainfall.

74 Scientific American, November 2021

**“Escalating Drought”, zusammen mit Georgios Karamanis für Scientific American, Ausgabe Nov 2021**



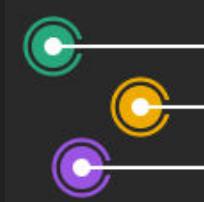
# Where are Americans born?



1POINT21  
INTERACTIVE

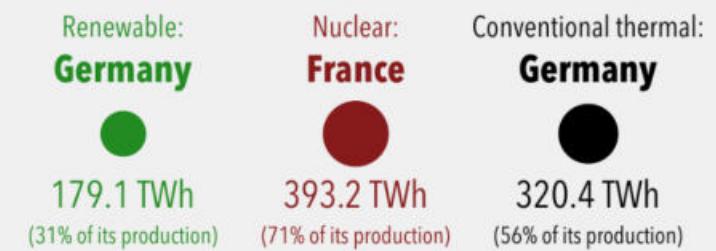
source: Steven Ruggles, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler and Matthew Sobek.  
IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. <https://doi.org/10.18128/D010.V11.0>

"Where are Americans born?" von @ErinDataViz

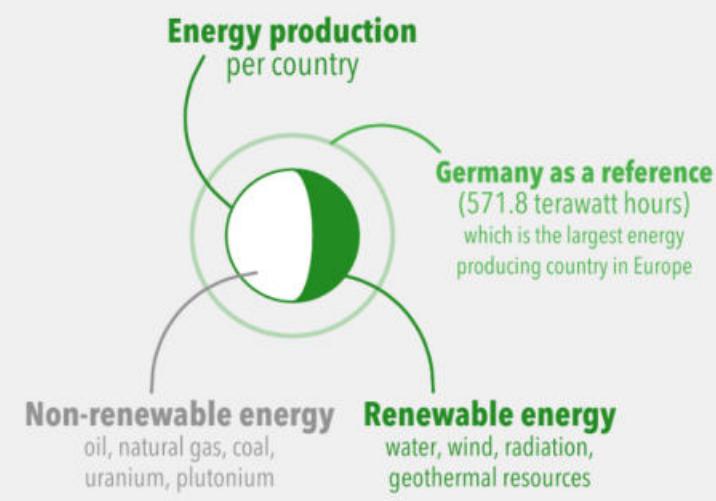


## How European countries generated electricity in 2018

**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.

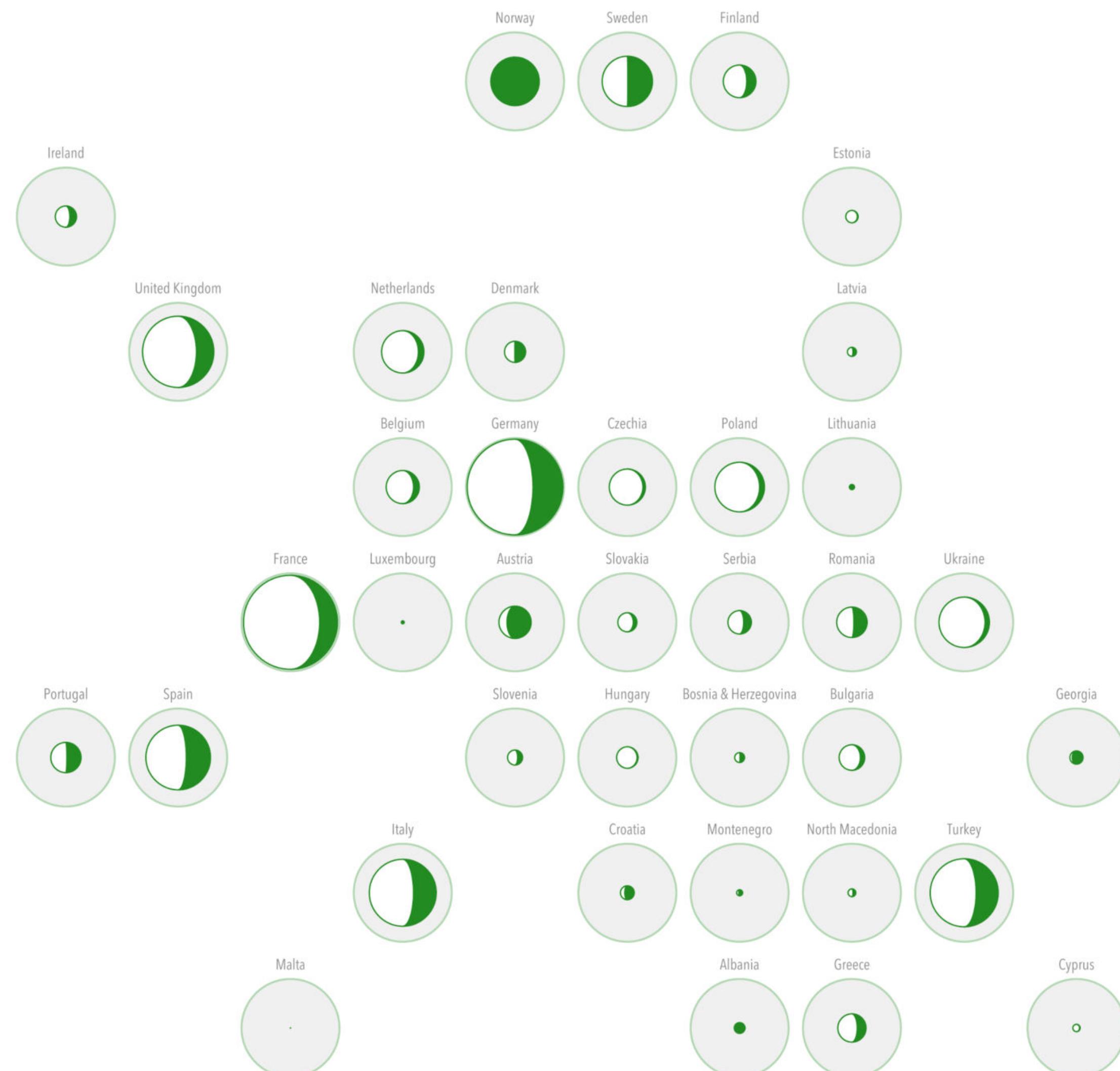


Renewable energy is energy that comes from resources that are naturally replenished such as sunlight, wind, water, and geothermal heat. Unlike fossil fuels, such as oil, natural gas and coal, or nuclear power sources such as uranium and plutonium, renewable energy regenerates naturally in a short period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

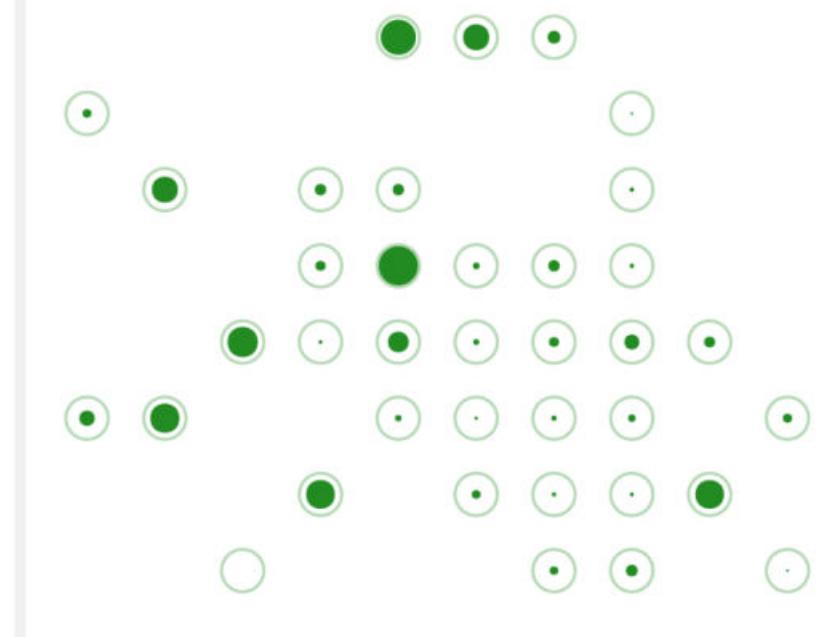


**Norway** had an electricity production almost entirely made up of renewable energy (98%). This makes Norway the second largest producer of this energy type in Europe. Interestingly, most of the renewable energy is produced by hydro power that take up 95% and only 3% by wind. In contrast, twelve European 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%).

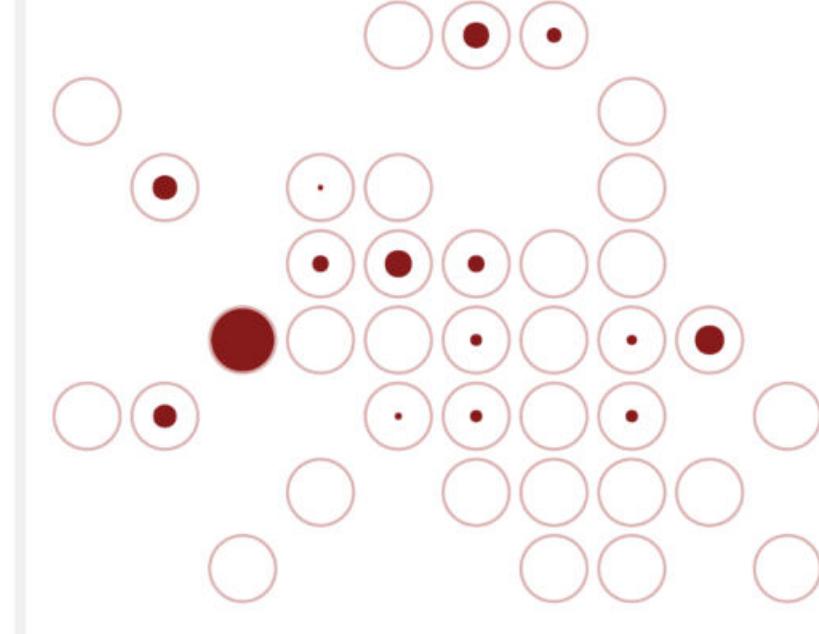
Note: Energy production is mapped to the area of the circles.  
Visualization by Cédric Scherer • Data by Eurostat



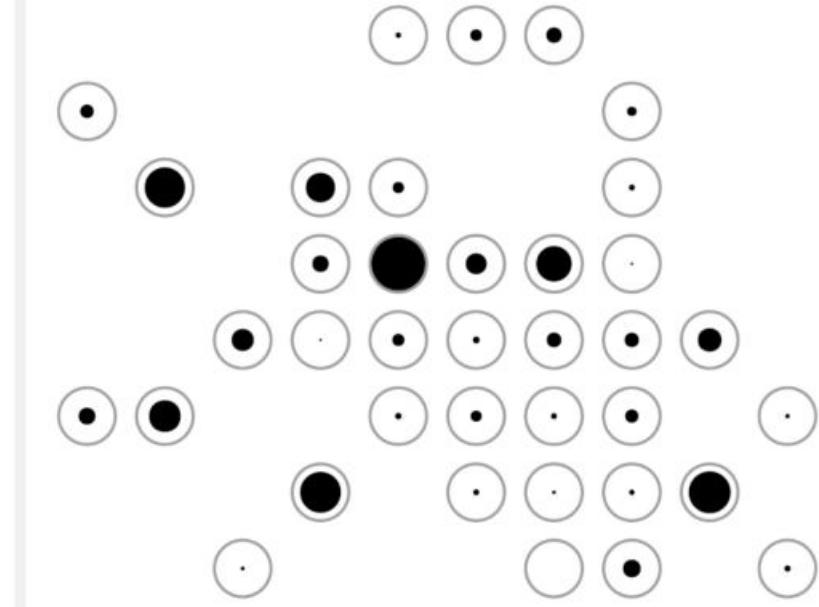
### Renewable energy



### Nuclear energy

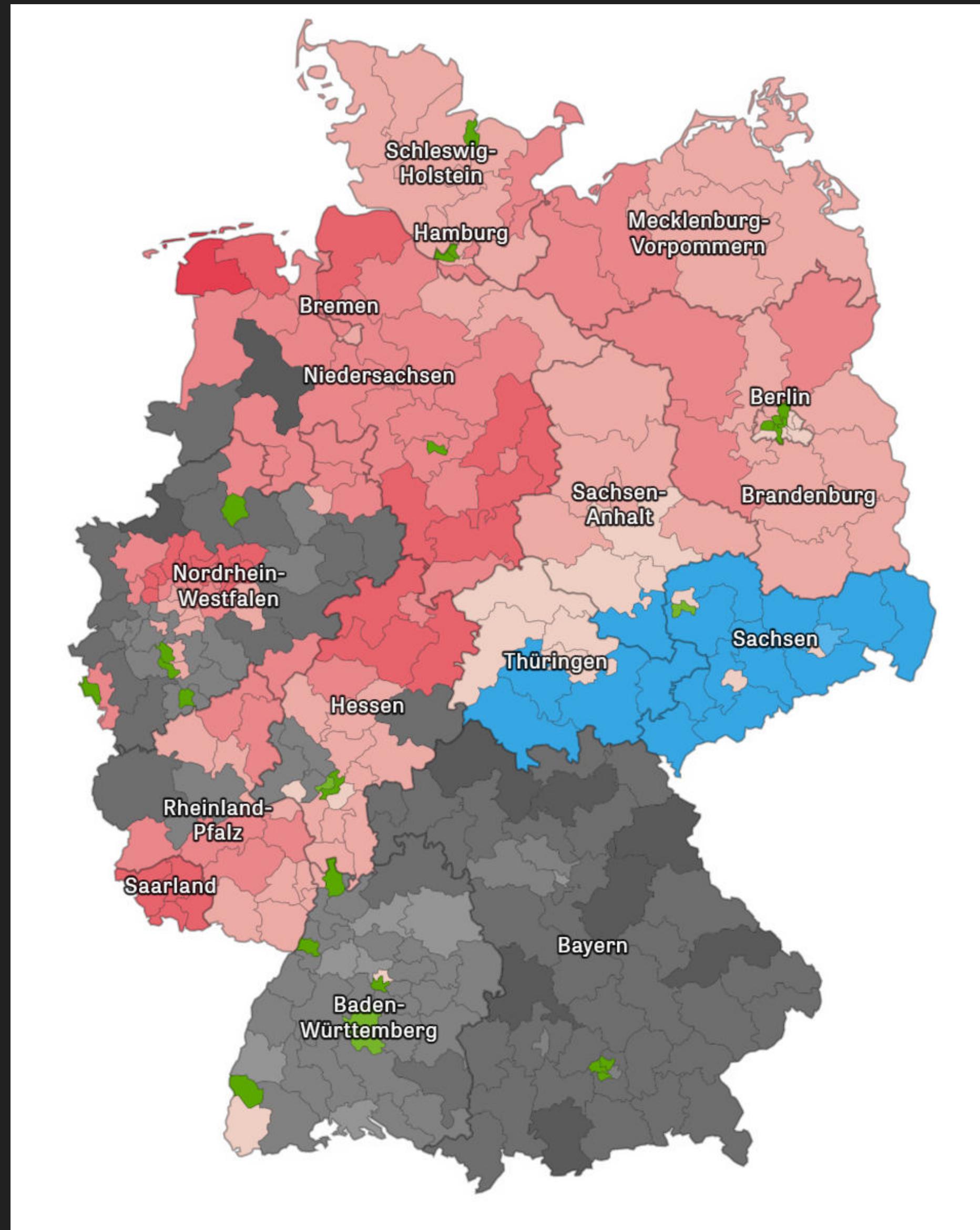


### Conventional thermal energy



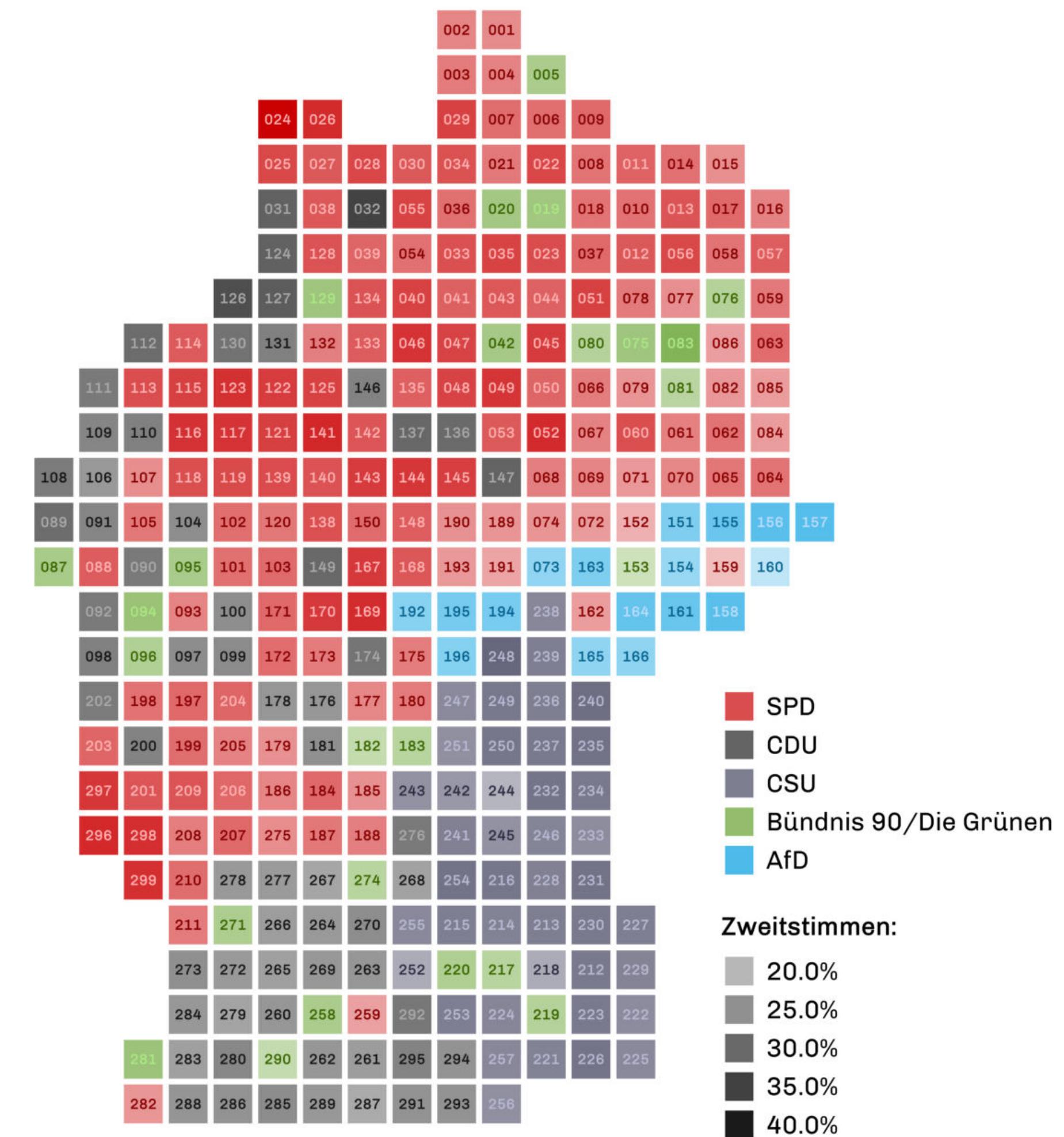
**“How European countries generated electricity in 2018”, #TidyTuesday Contribution**





## Ergebnisse der Bundestagswahl 2021

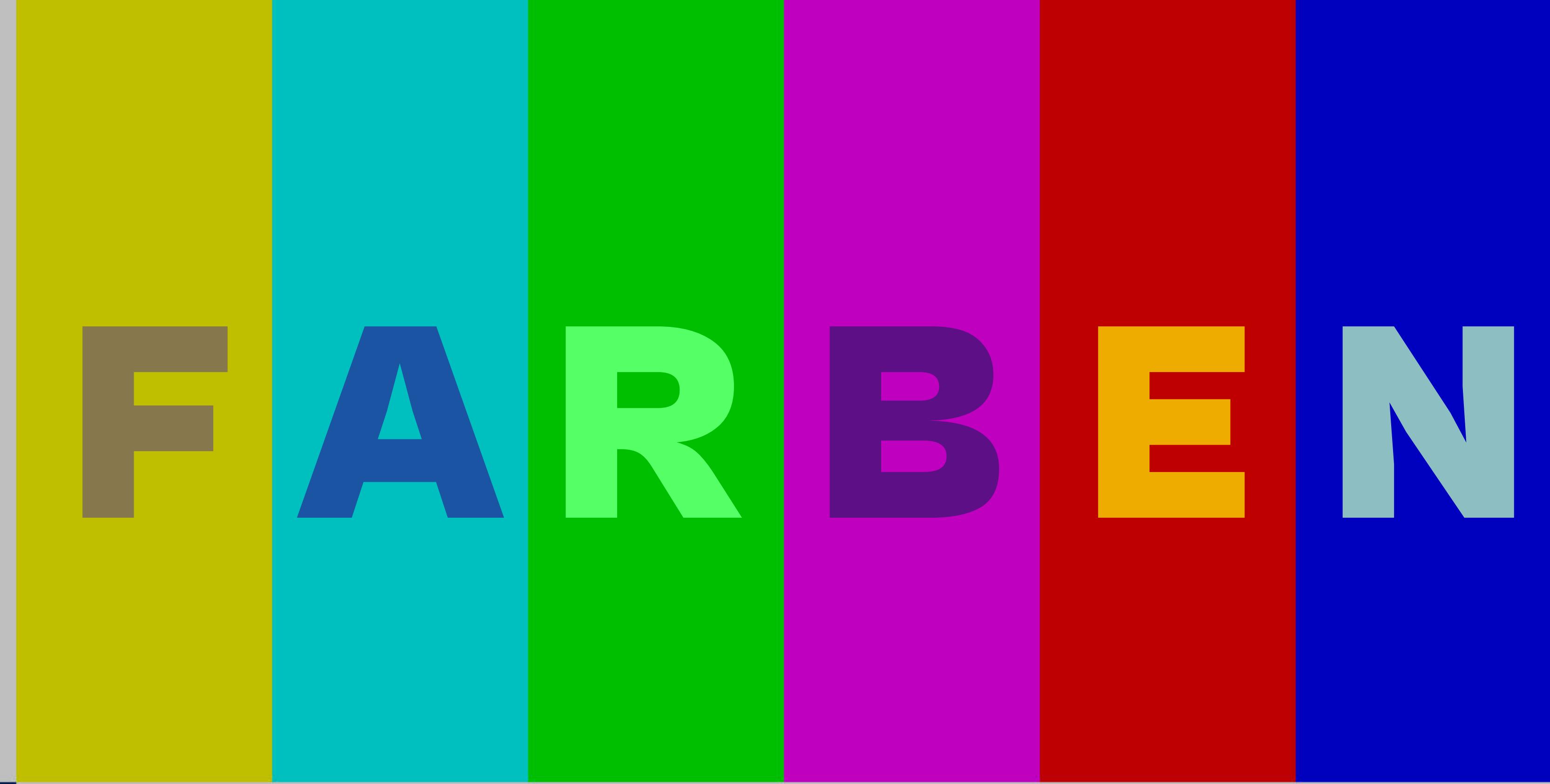
Die stärksten Parteien nach Prozent der Zweitstimmen.



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

Links: Choropleth-Karte von Die Zeit | Rechts: Kachel-Karte (“Tile Grid Map”) von Cédric Scherer & Ansgar Wolsing





FARBEN

und ihre Tücken



cedricscherer.com



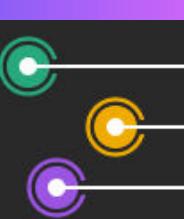
@CedScherer



@CedScherer



z3tt



# NON Uniform Distances



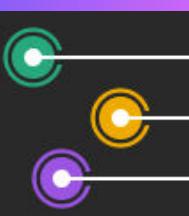
[cedricscherer.com](http://cedricscherer.com)



@CedScherer



[z3tt](https://github.com/z3tt)



[Med Phys.](#) 2015 Jun; 42(6): 2942–2954. Published online 2015 May 20. doi: [10.1118/1.4921125](https://doi.org/10.1118/1.4921125)

PMCID: PMC5148121 | PMID: [26127048](https://pubmed.ncbi.nlm.nih.gov/26127048/)

## Effect of color visualization and display hardware on the visual assessment of pseudocolor medical images

Silvina Zabala-Travers, Mina Choi, Wei-Chung Cheng, and Aldo Badano<sup>a)</sup>

10 March 2017

## Interpretation of the rainbow color scale for quantitative medical imaging: perceptually linear color calibration (CSDF) versus DICOM GSDF

[Frédérique Chesterman](#), [Hannah Manssens](#), [Céline Morel](#), [Guillaume Serrell](#), [Bastian Piepers](#), [Tom Kimpe](#)

[Author Affiliations +](#)

[Proceedings Volume 10136, Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment; 101360R \(2017\)](#) <https://doi.org/10.1117/12.2253885>

Event: [SPIE Medical Imaging](#), 2017, Orlando, Florida, United States

*IEEE Computer Graphics and Applications*

## Rainbow Color Map (Still) Considered Harmful

March/April 2007, pp. 14-17, vol. 27

DOI Bookmark: [10.1109/MCG.2007.46](https://doi.org/10.1109/MCG.2007.46)

Authors

David Borland, University of North Carolina at Chapel Hill

Russell M. Taylor II, University of North Carolina at Chapel Hill

Education and communication

Rainbow color map distorts and misleads research in hydrology – guidance for better visualizations and science communication

Michael Stoelze<sup>1</sup> and Lina Stein<sup>2</sup>

<sup>1</sup>Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

<sup>2</sup>Department of Civil Engineering, University of Bristol, Bristol, UK



[Med Phys.](#) 2015 Jun; 42(6): 2942–2954. Published online 2015 May 20. doi: [10.1118/1.4921125](https://doi.org/10.1118/1.4921125)

PMCID: PMC5148121 | PMID: [26127048](https://pubmed.ncbi.nlm.nih.gov/26127048/)

Effect of color visualization and display hardware on the visual assessment of pseudocolor medical images

[Silvina Zabala-Travers](#), [Mina Choi](#), [Wei-Chung Cheng](#), and [Aldo Badano<sup>a\)</sup>](#)

*IEEE Computer Graphics and Applications*

**Rainbow Color Map (Still) Considered Harmful**

March/April 2007, pp. 14-17, vol. 27

DOI Bookmark: [10.1109/MCG.2007.46](https://doi.org/10.1109/MCG.2007.46)

Authors

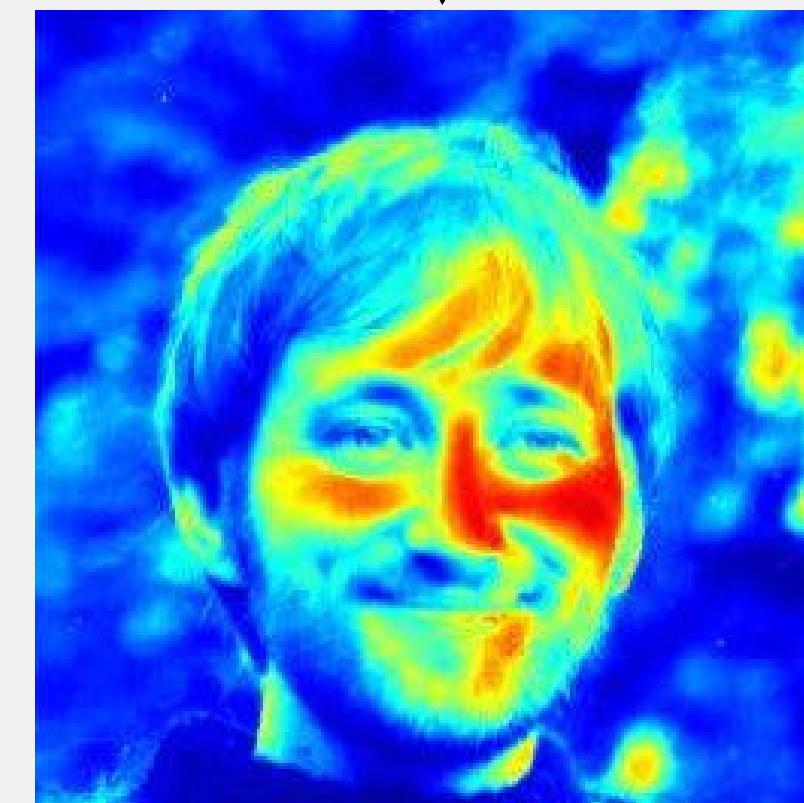
**“The ad hoc manner in which color is handled and the lack of standard approaches have been associated with suboptimal and inconsistent diagnostic decisions with a negative impact on patient treatment and prognosis.”**

*Zabala-Travers, Choi, Cheng & Badano 2015 Med Phys.*

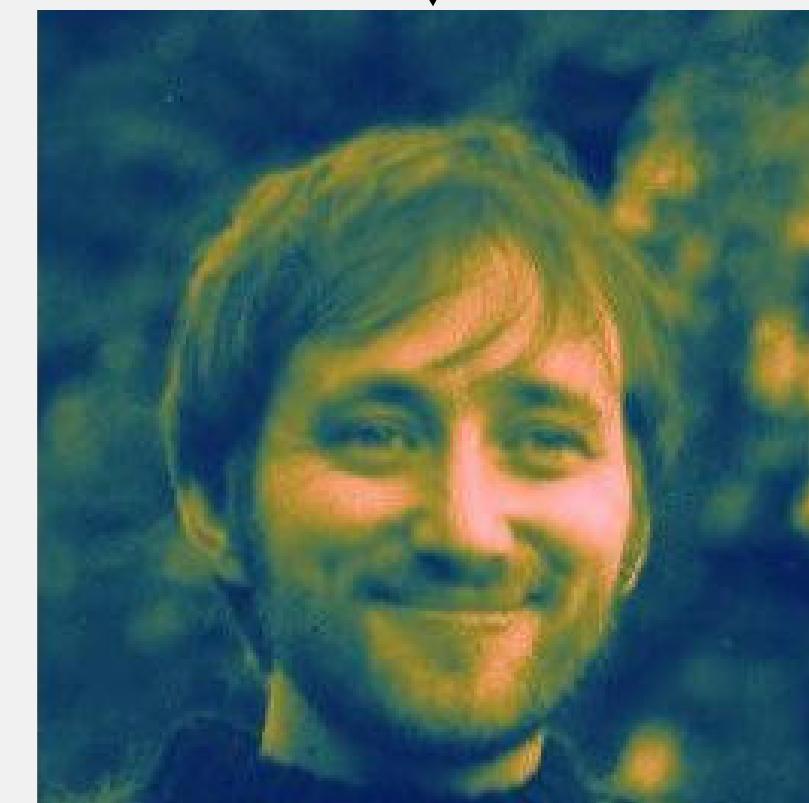




**true-color Phil**



**rainbow Phil**



**batlow Phil**

*Quelle: [Fabio Cramer](#) (modifiziert)*



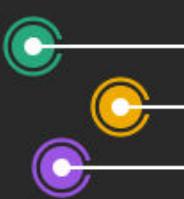
[cedricscherer.com](http://cedricscherer.com)



@CedScherer



z3tt



# Turbo

# Rainbow



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



@CedScherer



[z3tt](https://github.com/z3tt)



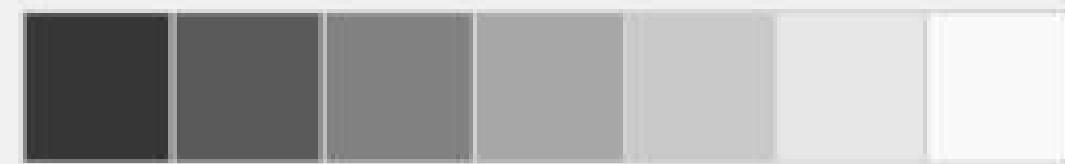
# Farbpaletten (richtig) auswählen

## Sequenziell

Beispiel



Graustufen

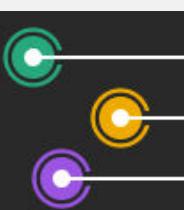


*numerische Informationen  
mit steigender Reihenfolge*

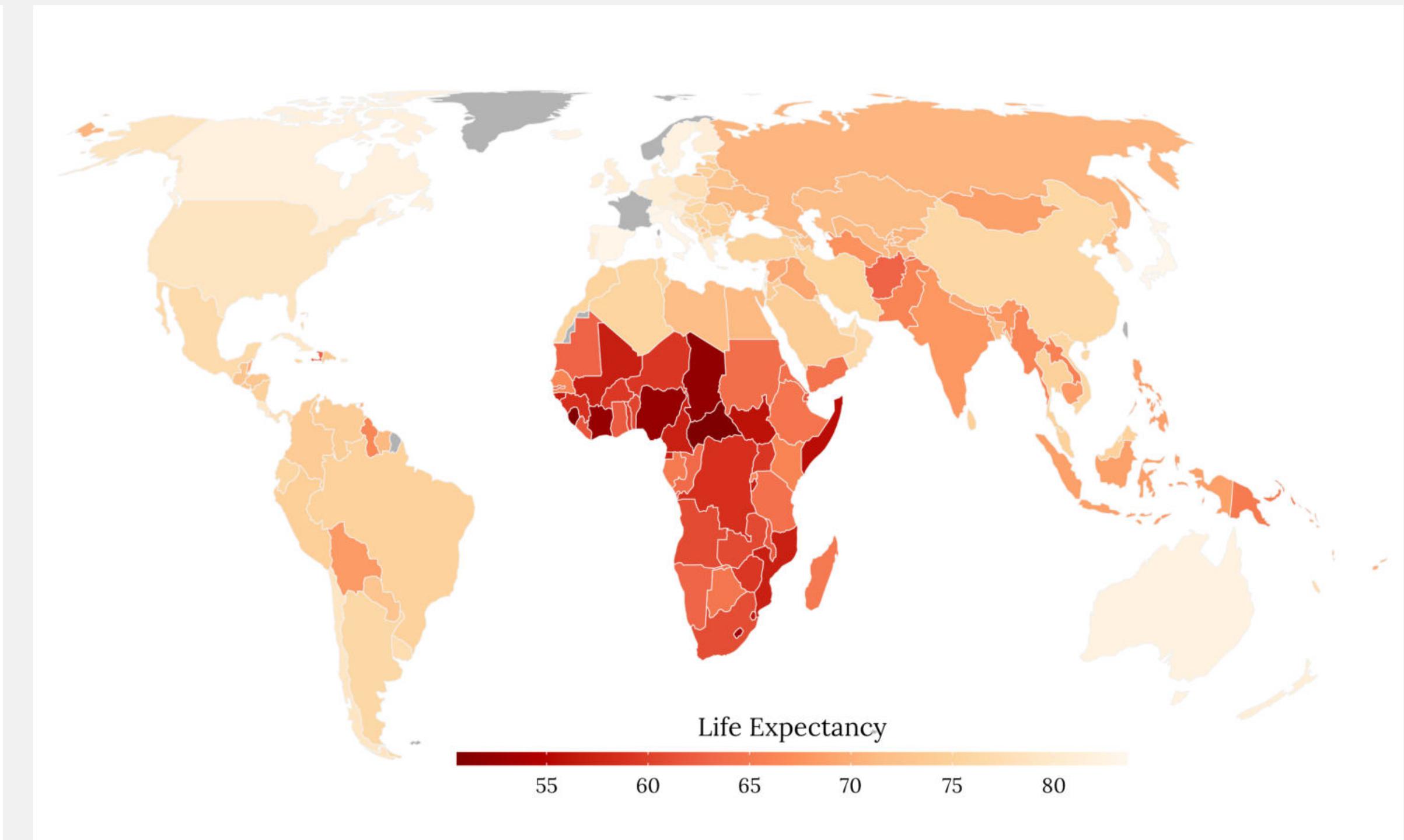
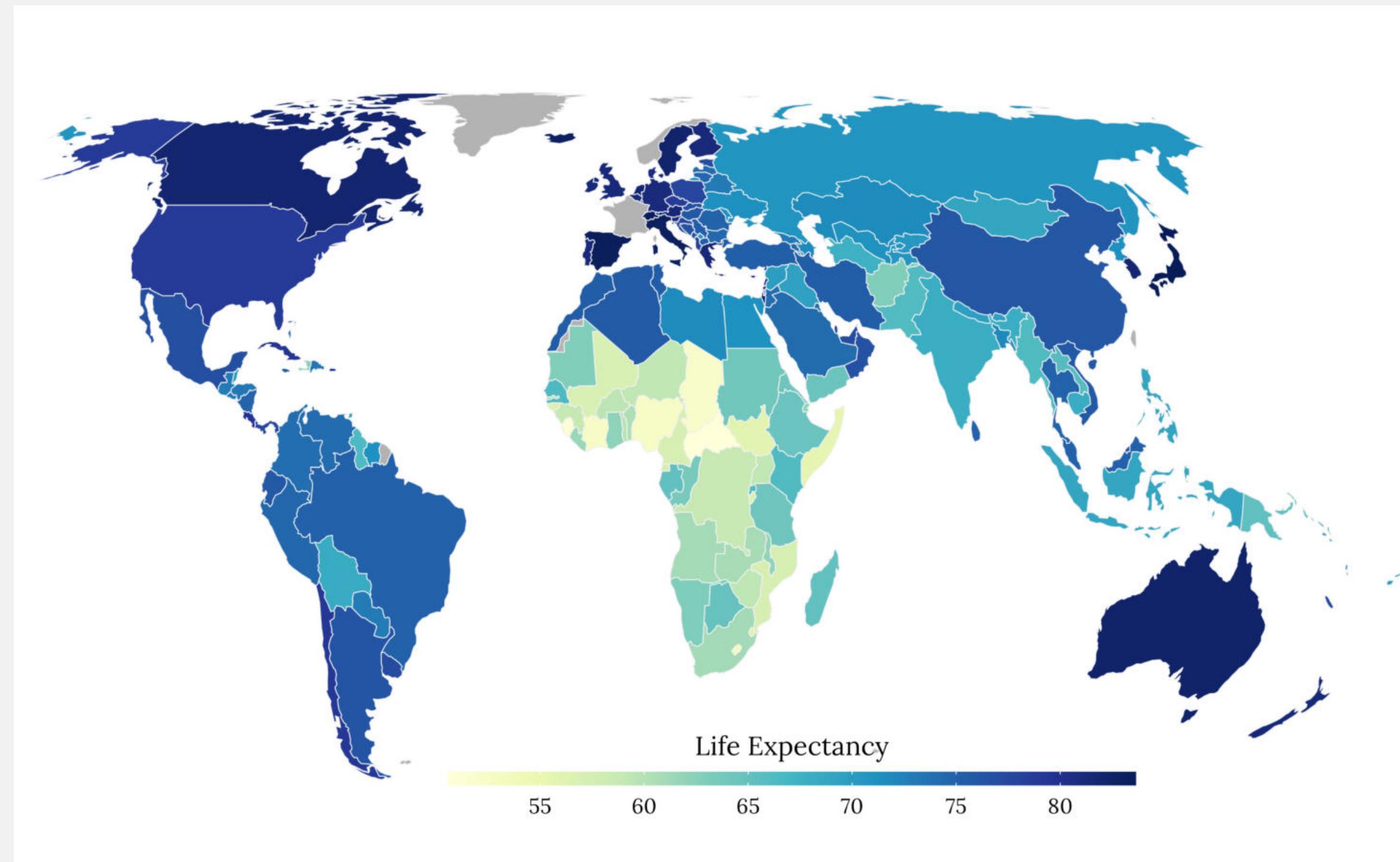
*verwende den höchsten Kontrast  
für die wichtigsten Informationen*

*entweder single- oder  
multi-hue Farbpaletten*

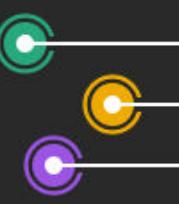
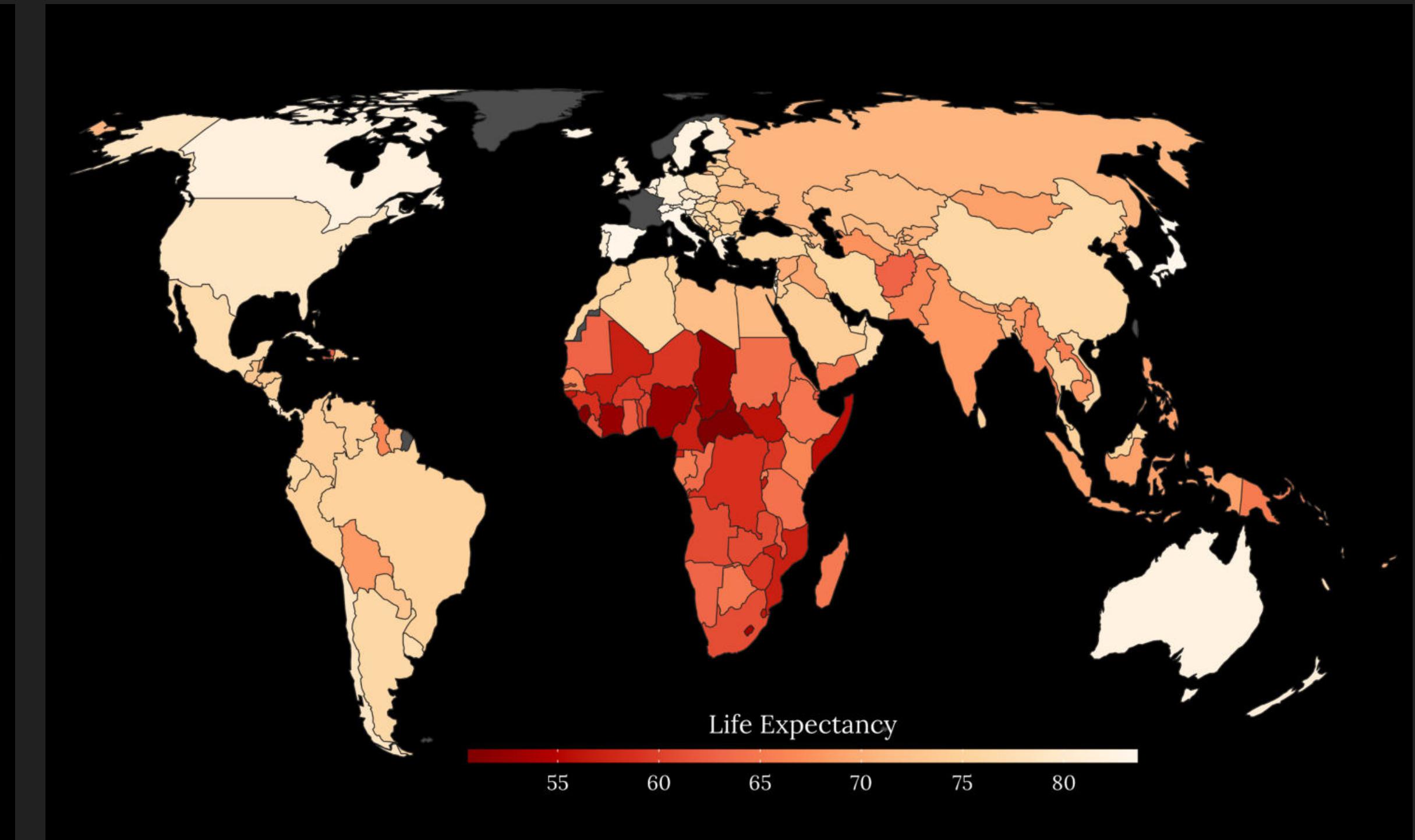
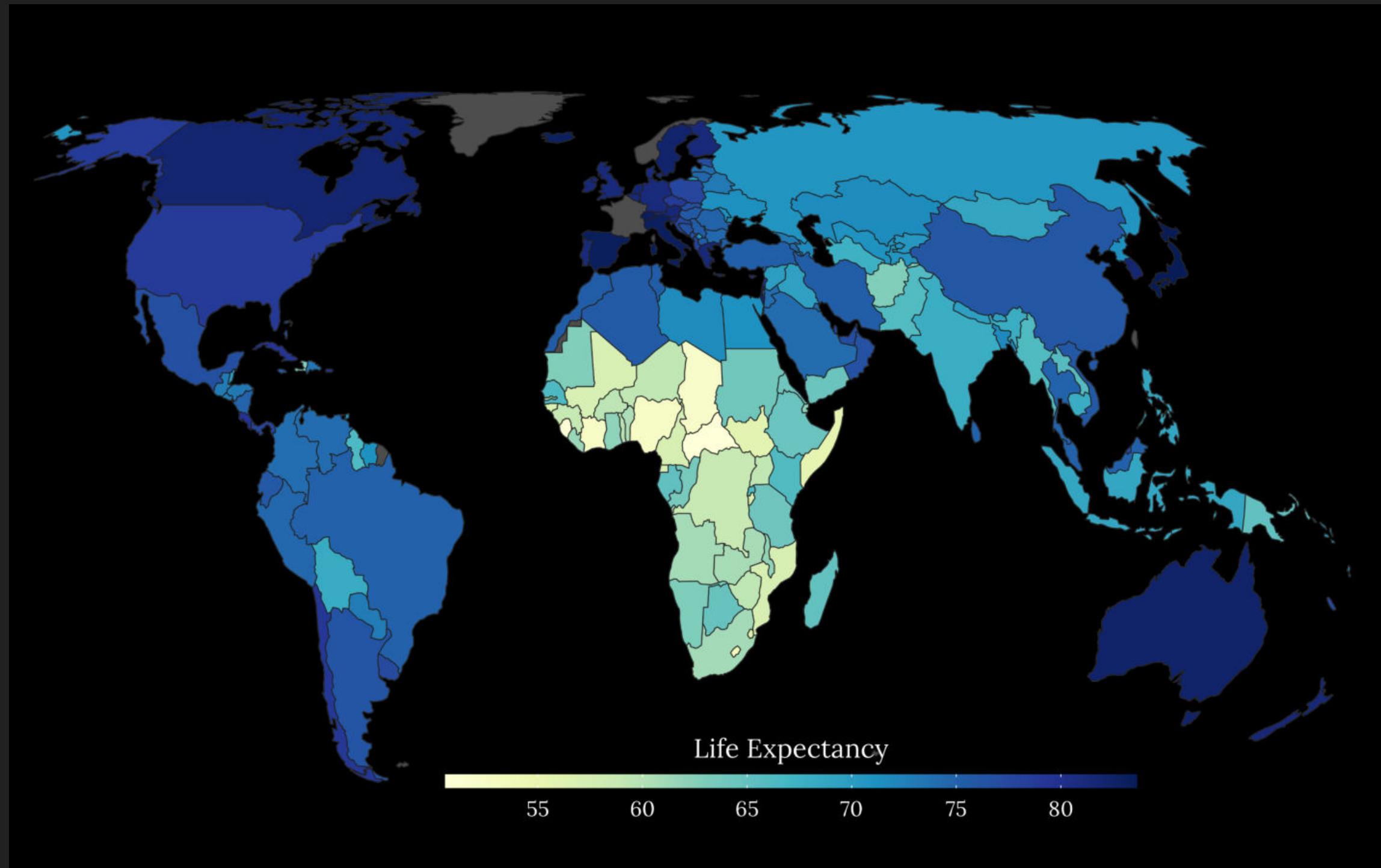
*Modifiziert von der `{colorspace}` R Package Vignette*



# Mehr gleich dunkel, richtig?



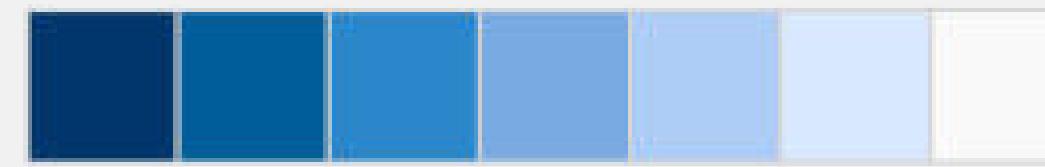
# Mehr gleich dunkel, richtig?



# Farbpaletten (richtig) auswählen

## Sequenziell

Beispiel



Graustufen



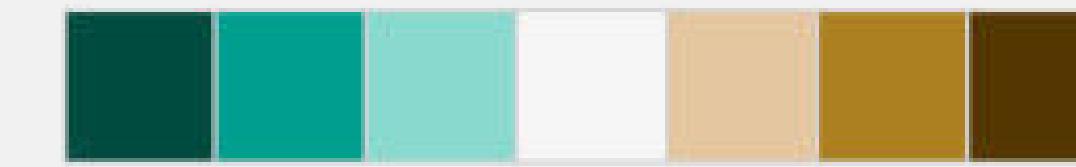
***numerische Informationen  
mit steigender Reihenfolge***

*verwende den höchsten Kontrast  
für die wichtigsten Informationen*

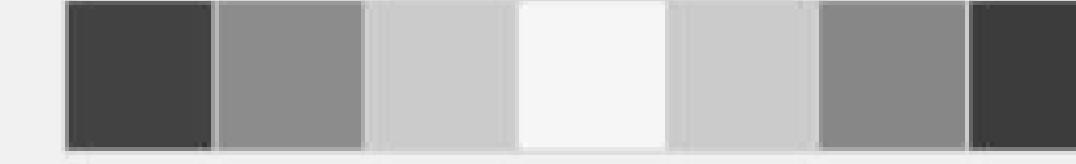
*entweder single- oder  
multi-hue Farbpaletten*

## Divergierend

Beispiel



Graustufen



***numerische Informationen  
mit kritischem Mittelpunkt***

*verwende einen bedeutsamen Mittelpunkt  
und nutze ausgewogene Extremwerte*

*Kombination von zwei  
sequenziellen Farbpaletten*

*Modifiziert von der `{colorspace}` R Package Vignette*



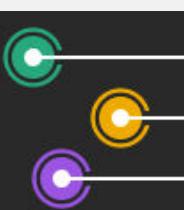
cedricscherer.com



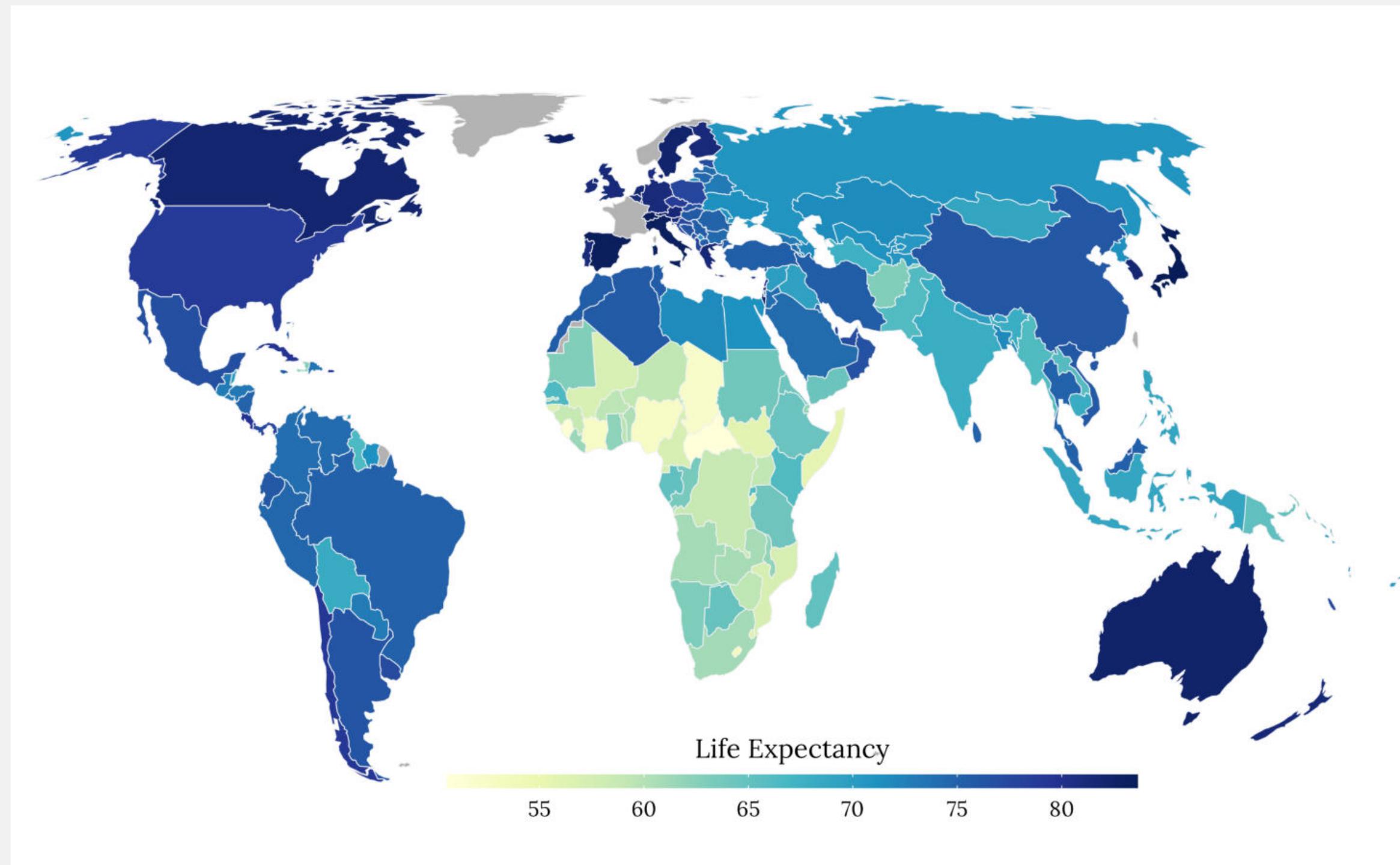
@CedScherer



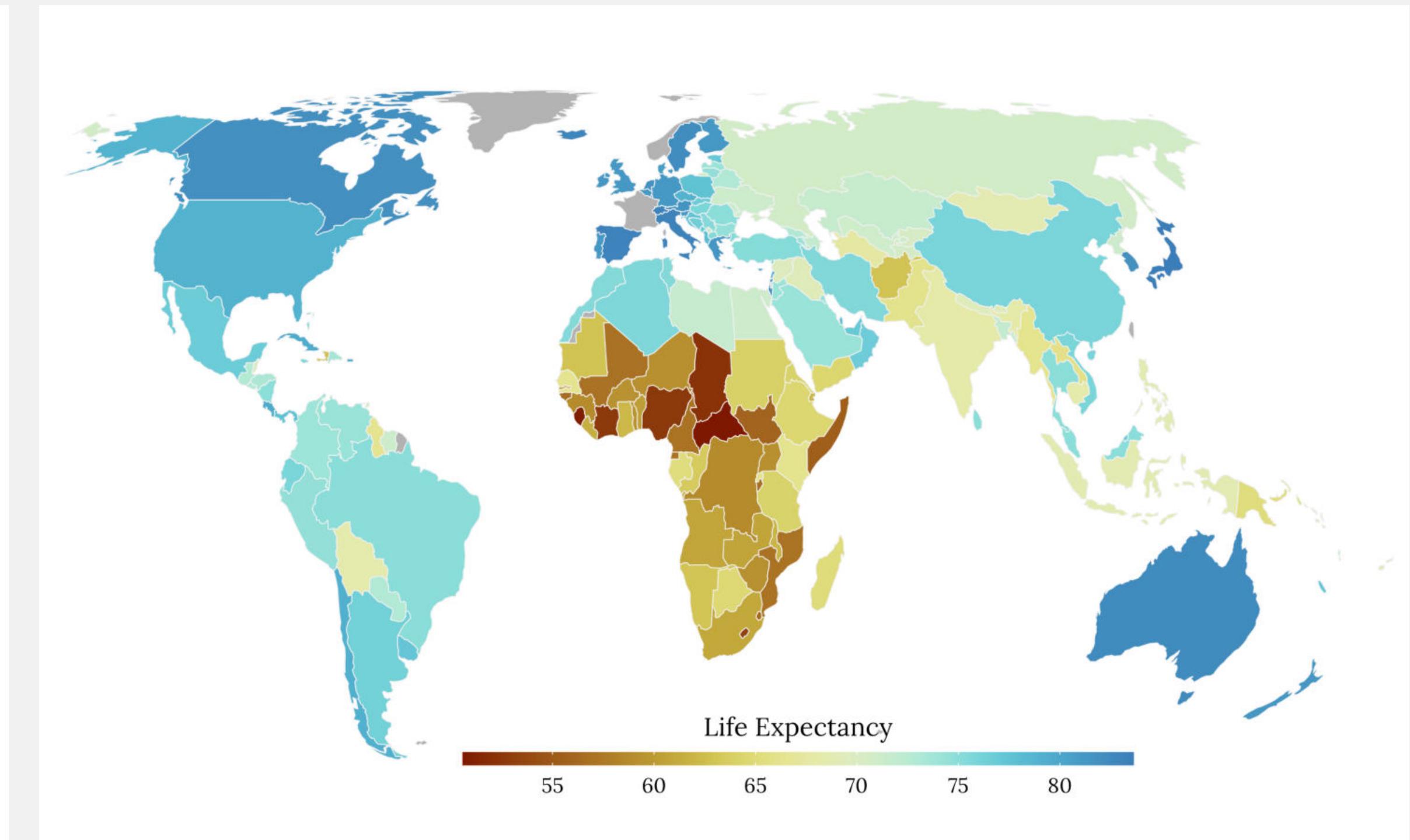
z3tt



# Was ist denn nun richtig?



sequenziell



divergierend



# Farbpaletten (richtig) auswählen

# Sequenziel

## *Beispiel*



# Graustufer



# *numerische Informationen mit steigender Reihenfolge*

*verwende den höchsten Kontrast  
für die wichtigsten Informationen*

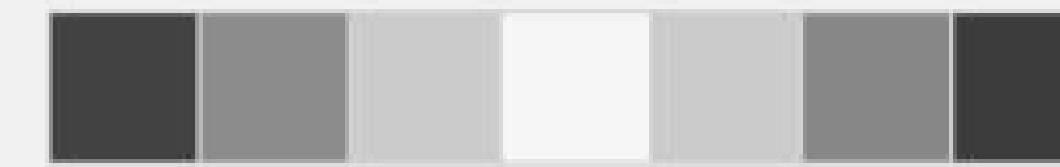
*entweder single- oder  
multi-hue Farbpaletten*

# Divergierend

## *Beispiel*



# Graustufen



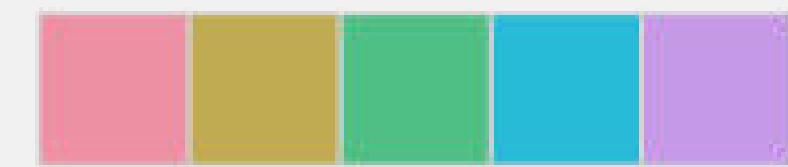
# *numerische Informationen mit kritischem Mittelpunkt*

*verwende einen bedeutsamen Mittelpunkt  
und nutze ausgewogene Extremwerte*

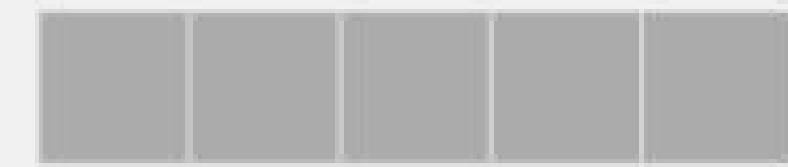
# *Kombination von zwei sequenziellen Farbpaletten*

# Qualitativ

## *Beispiel*



# Graustufen

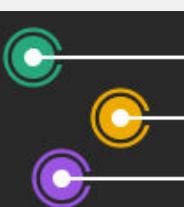


# ***kategorische Informationen*** *(ohne Reihenfolge)*

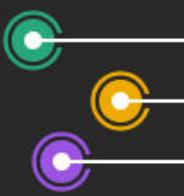
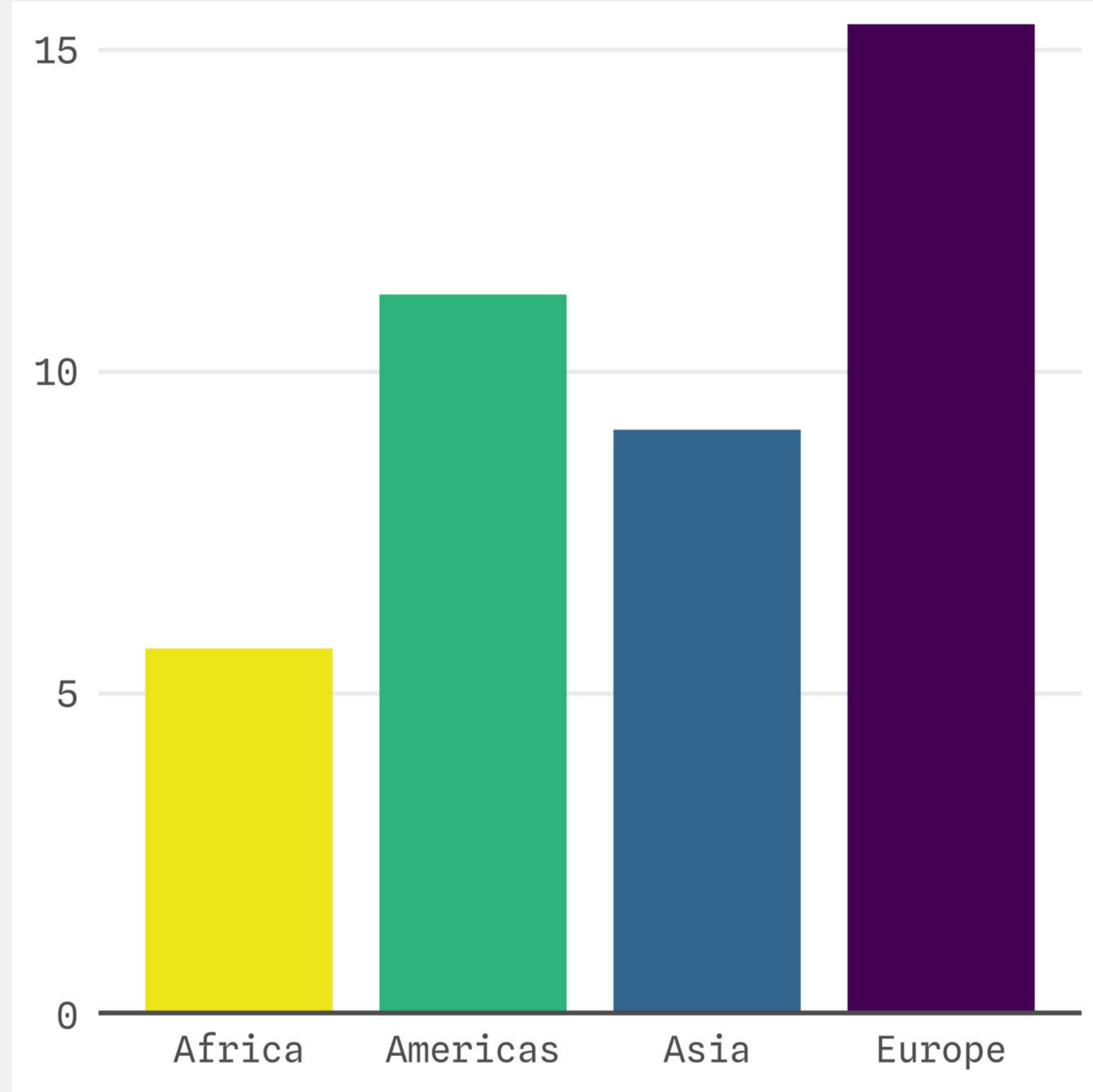
wähle verschiedene Farben mit demselben Wahrnehmungsgewicht

Anzahl auf 6 bis 8  
Kategorien begrenzen

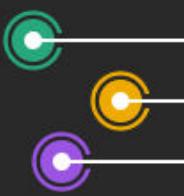
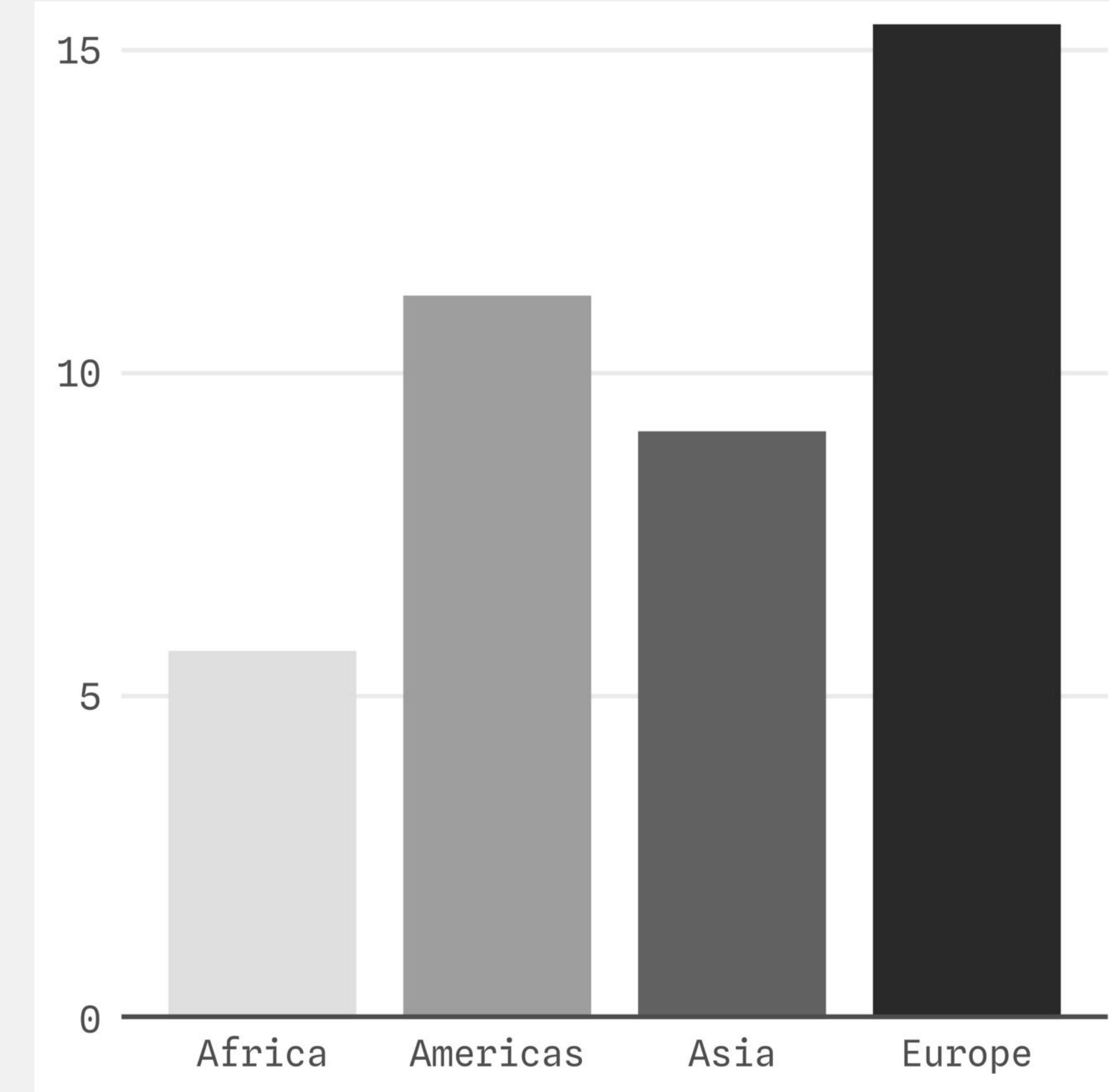
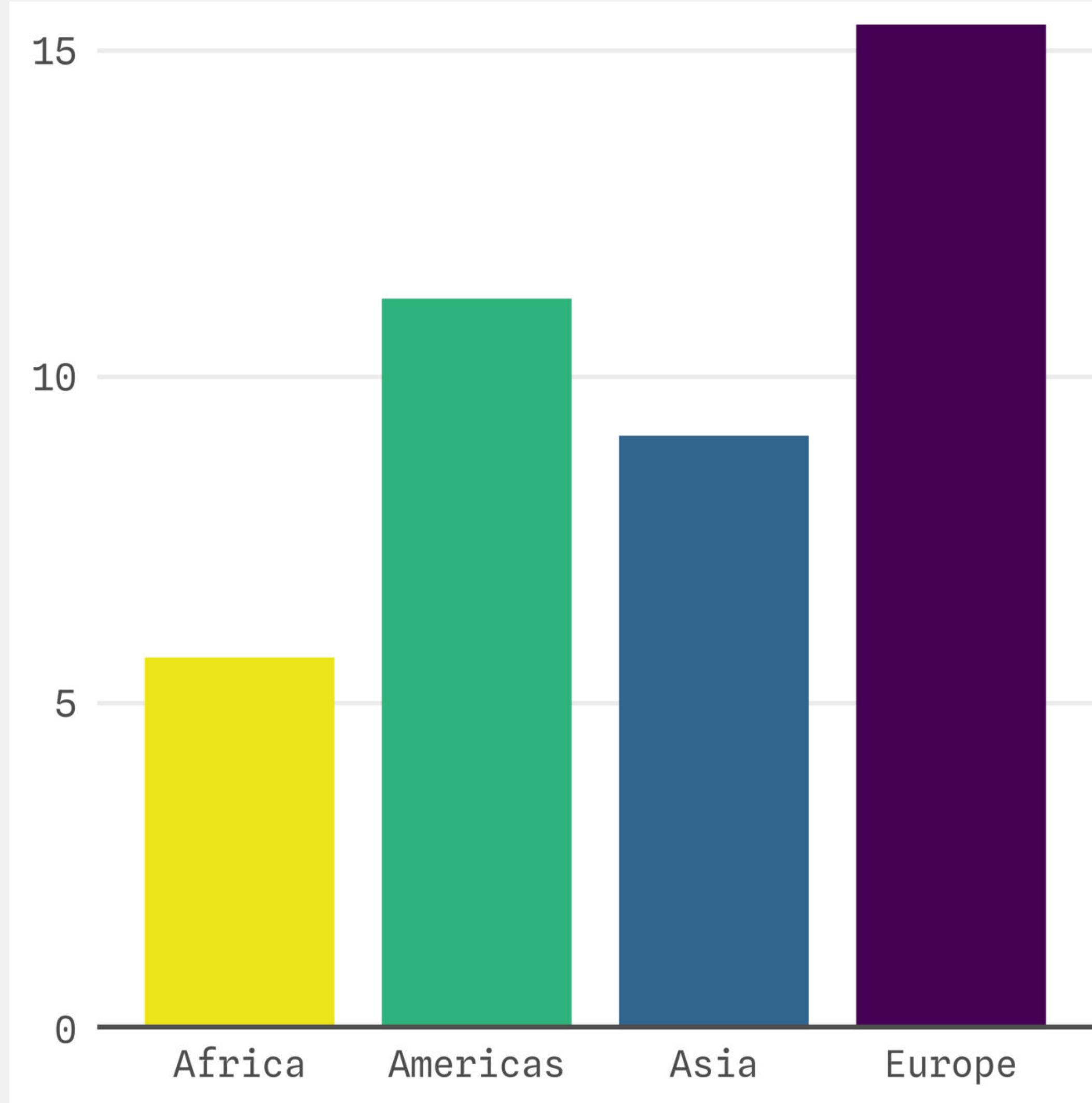
*Modifiziert von der {colorspace} R Package Vignette*



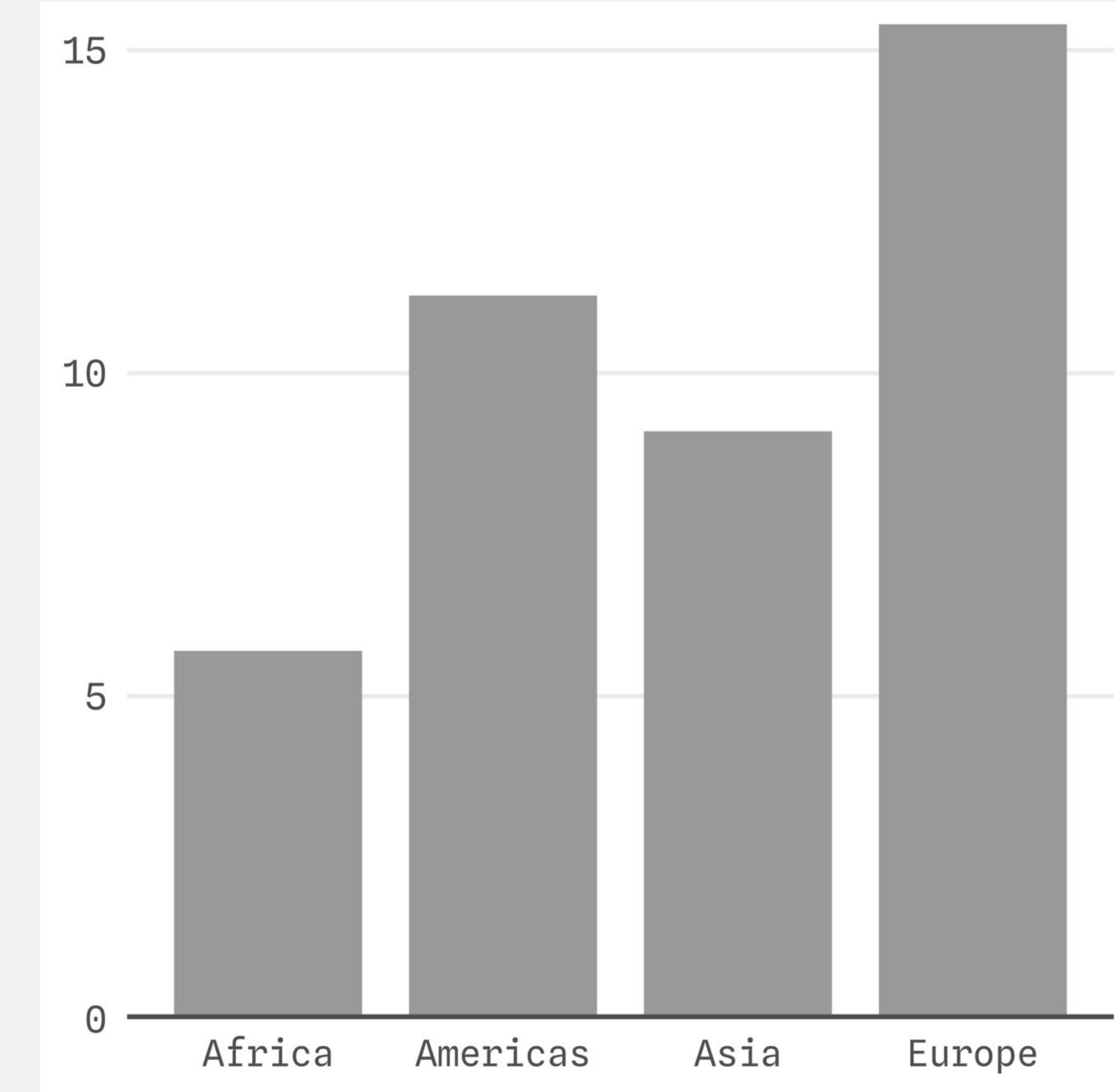
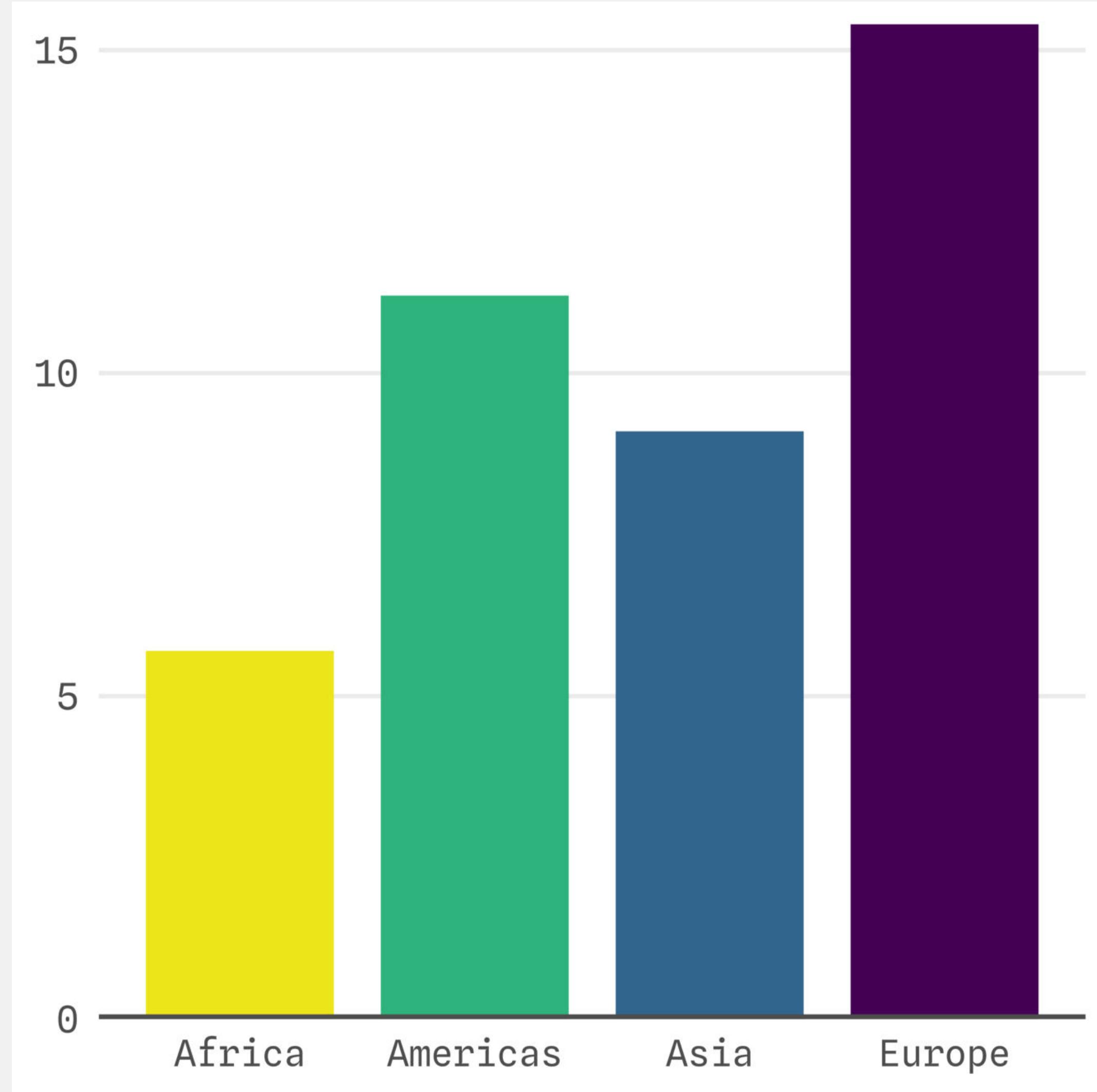
# „Viridis ist die perfekte Farbpalette!“



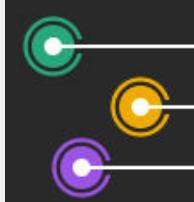
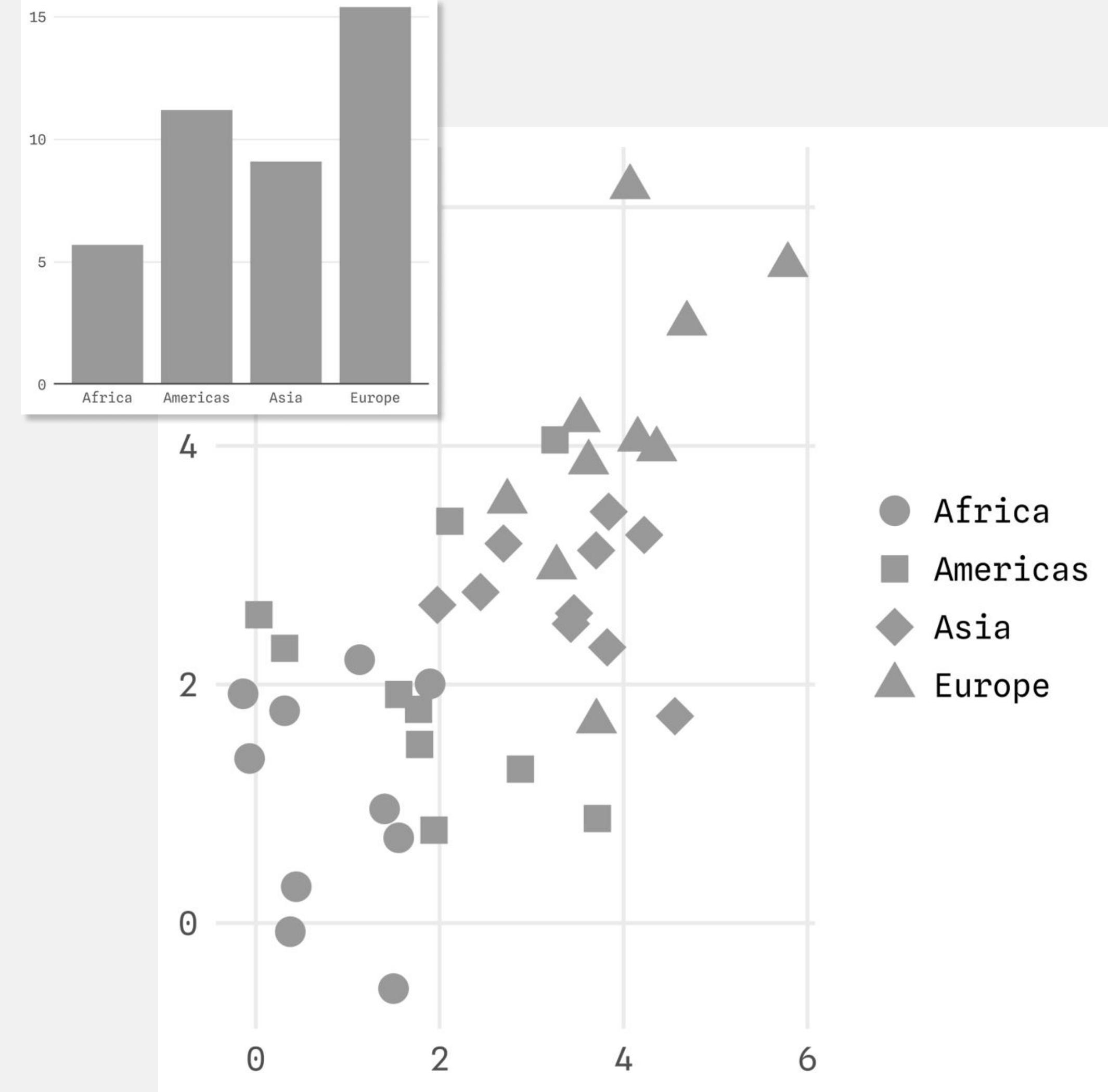
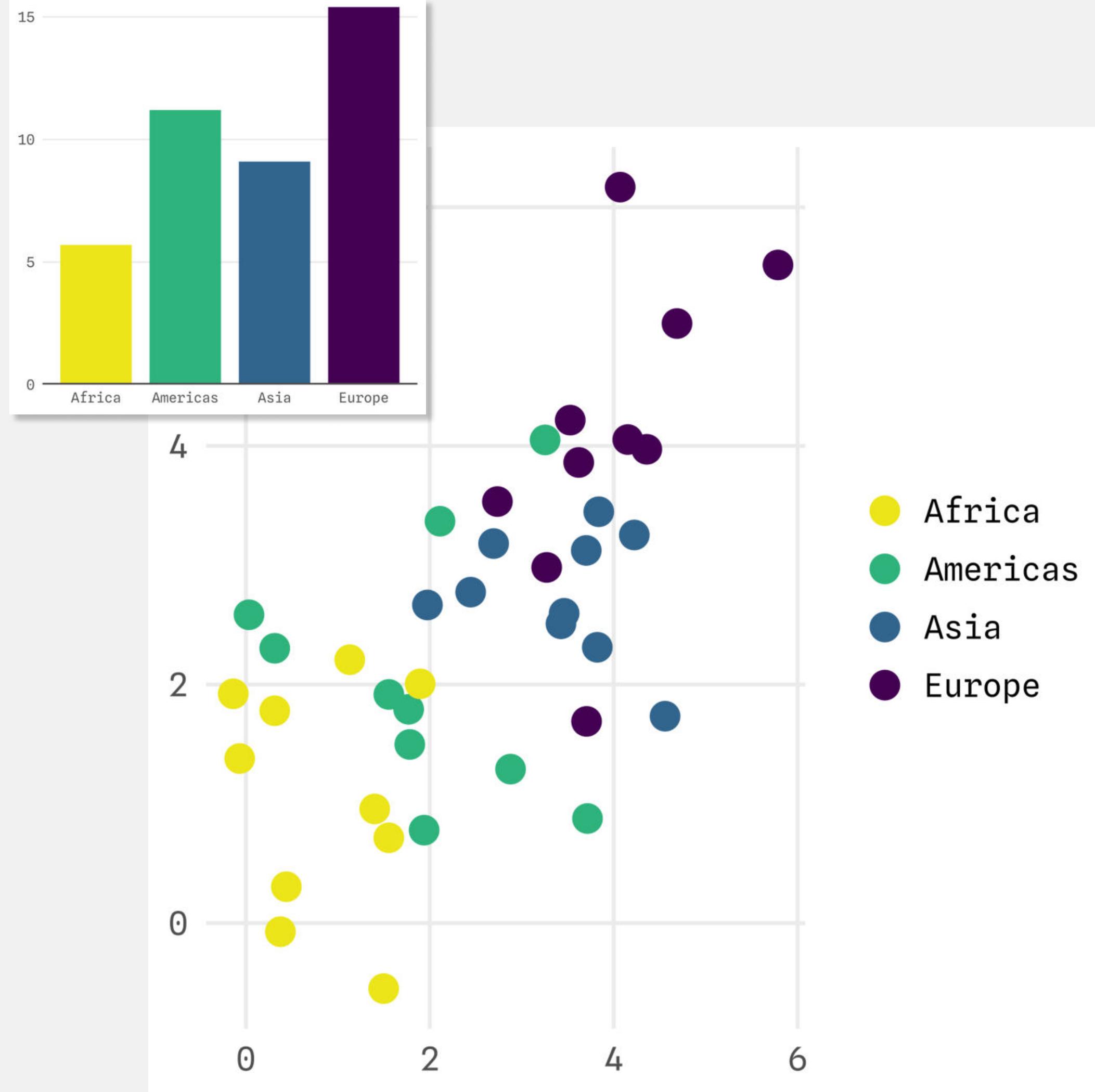
# „Viridis ist die perfekte Farbpalette!“



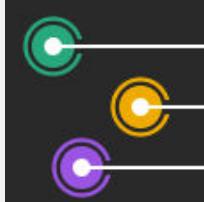
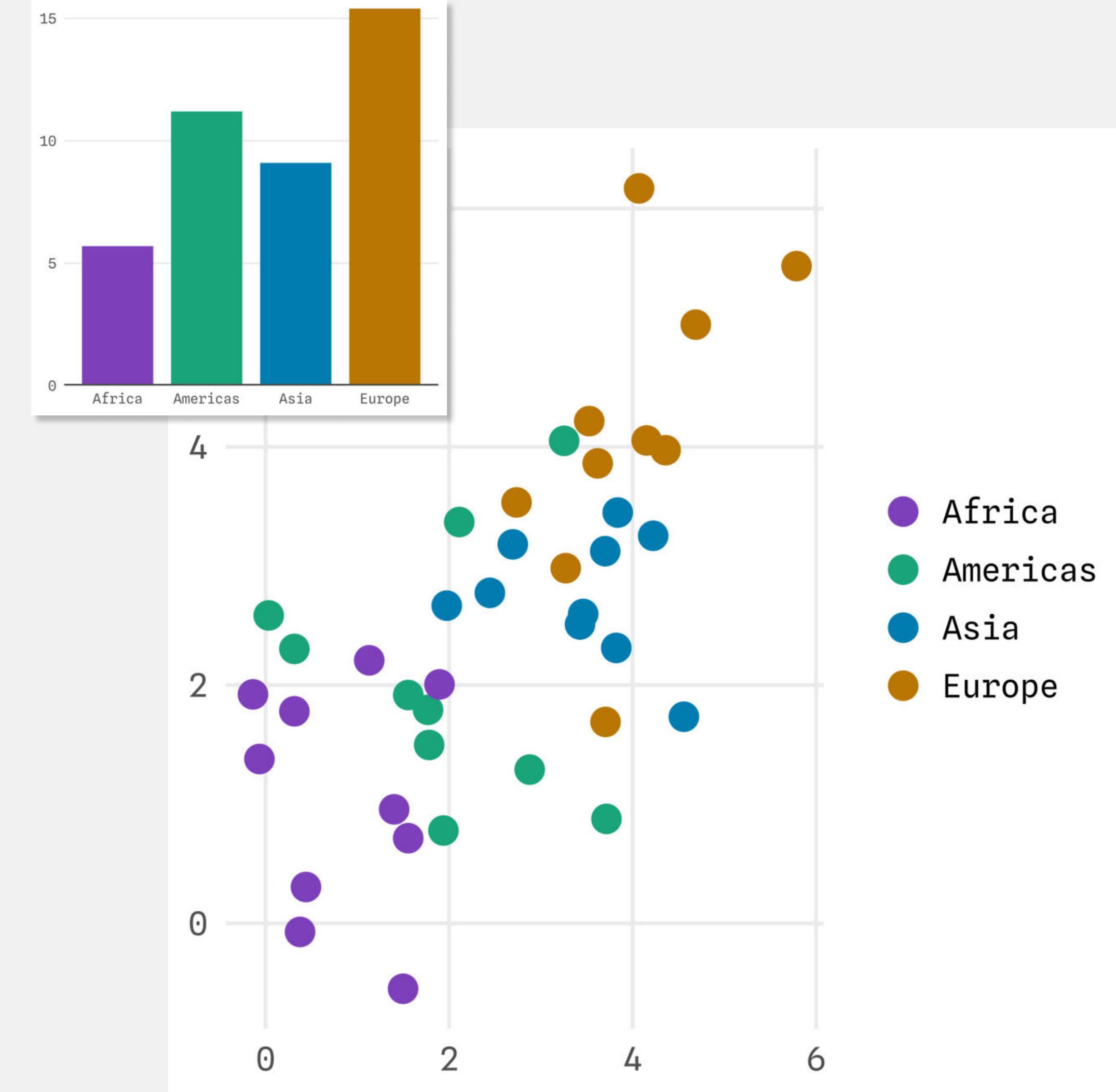
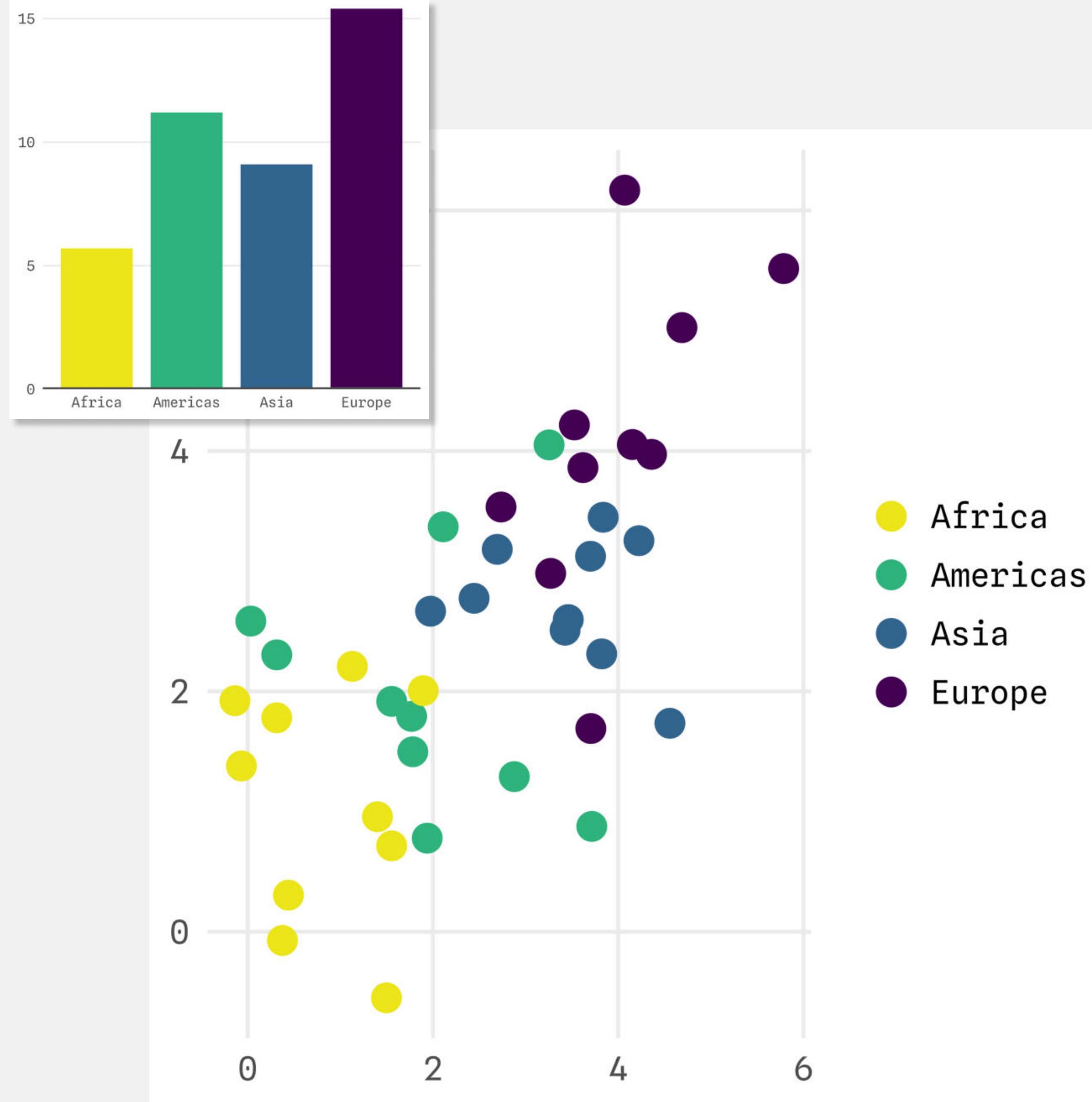
# „Viridis ist die perfekte Farbpalette!“



# „Viridis ist die perfekte Farbpalette!“



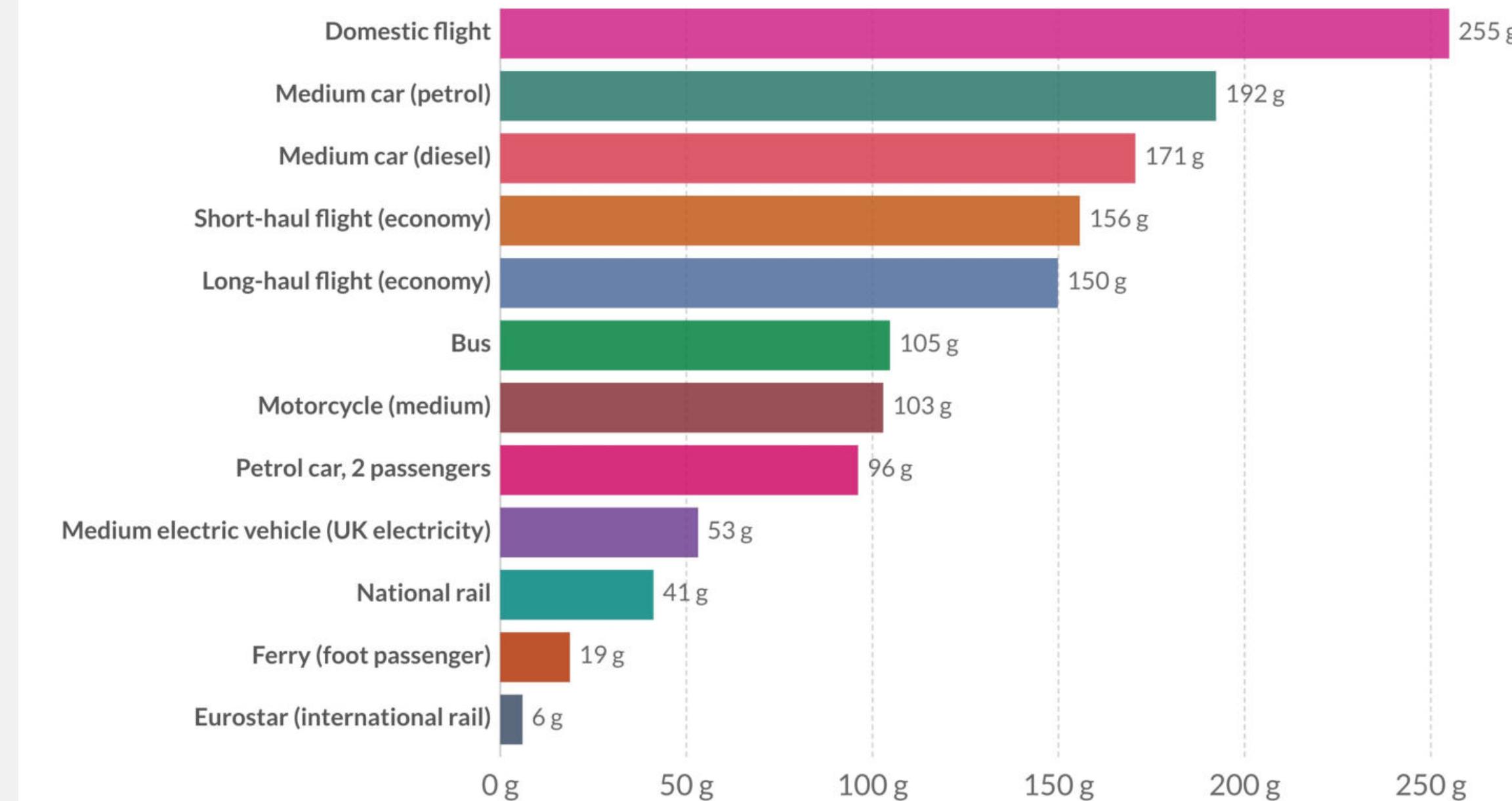
# „Viridis ist die perfekte Farbpalette!“



# Farben (richtig) nutzen

Carbon footprint of travel per kilometer, 2018

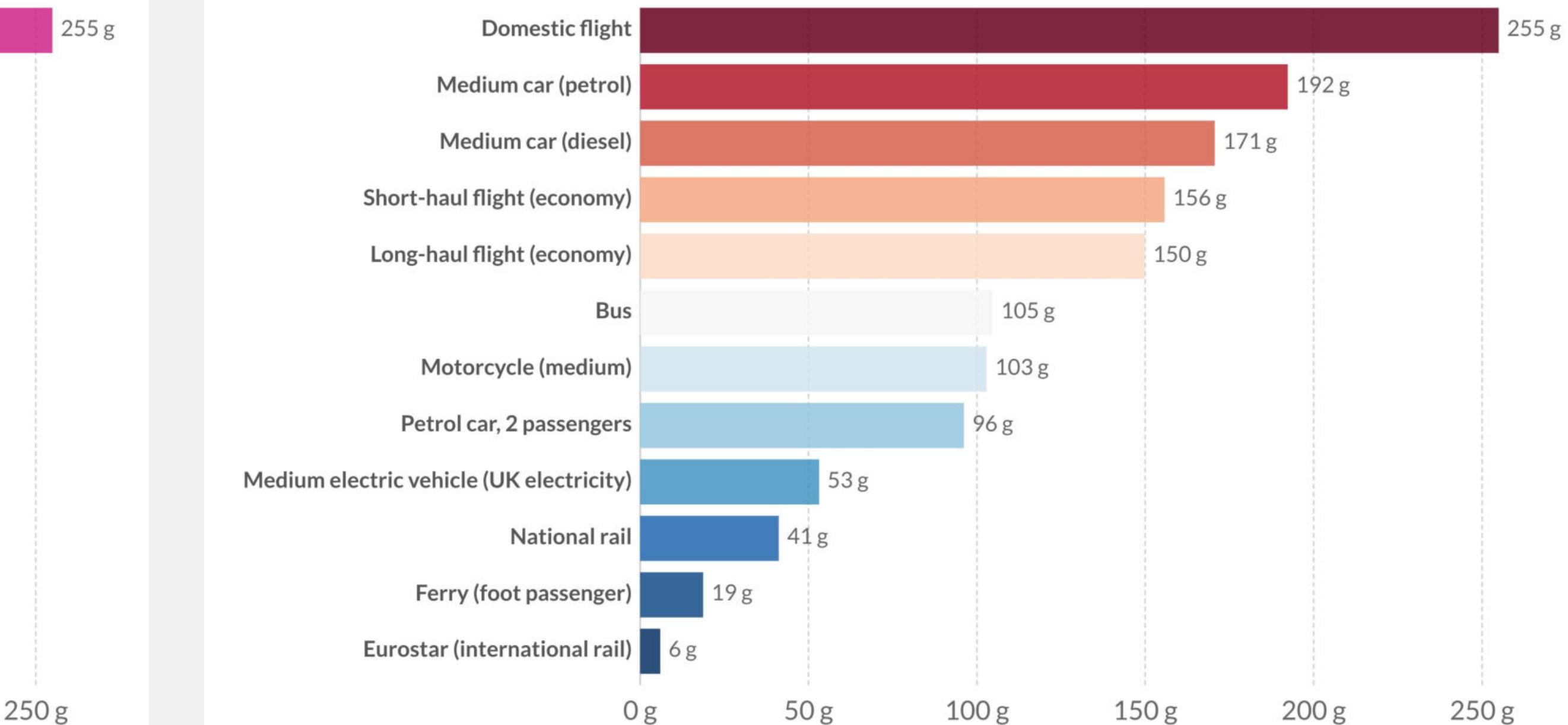
The carbon footprint of travel is measured in grams of carbon dioxide equivalents per passenger kilometer. This includes carbon dioxide, but also other greenhouse gases, and increased warming from aviation emissions at altitude.



Source: UK Department for Business, Energy & Industrial 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.

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 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.

Originalgrafik mit einer zufälligen kategoriellen Farbpalette

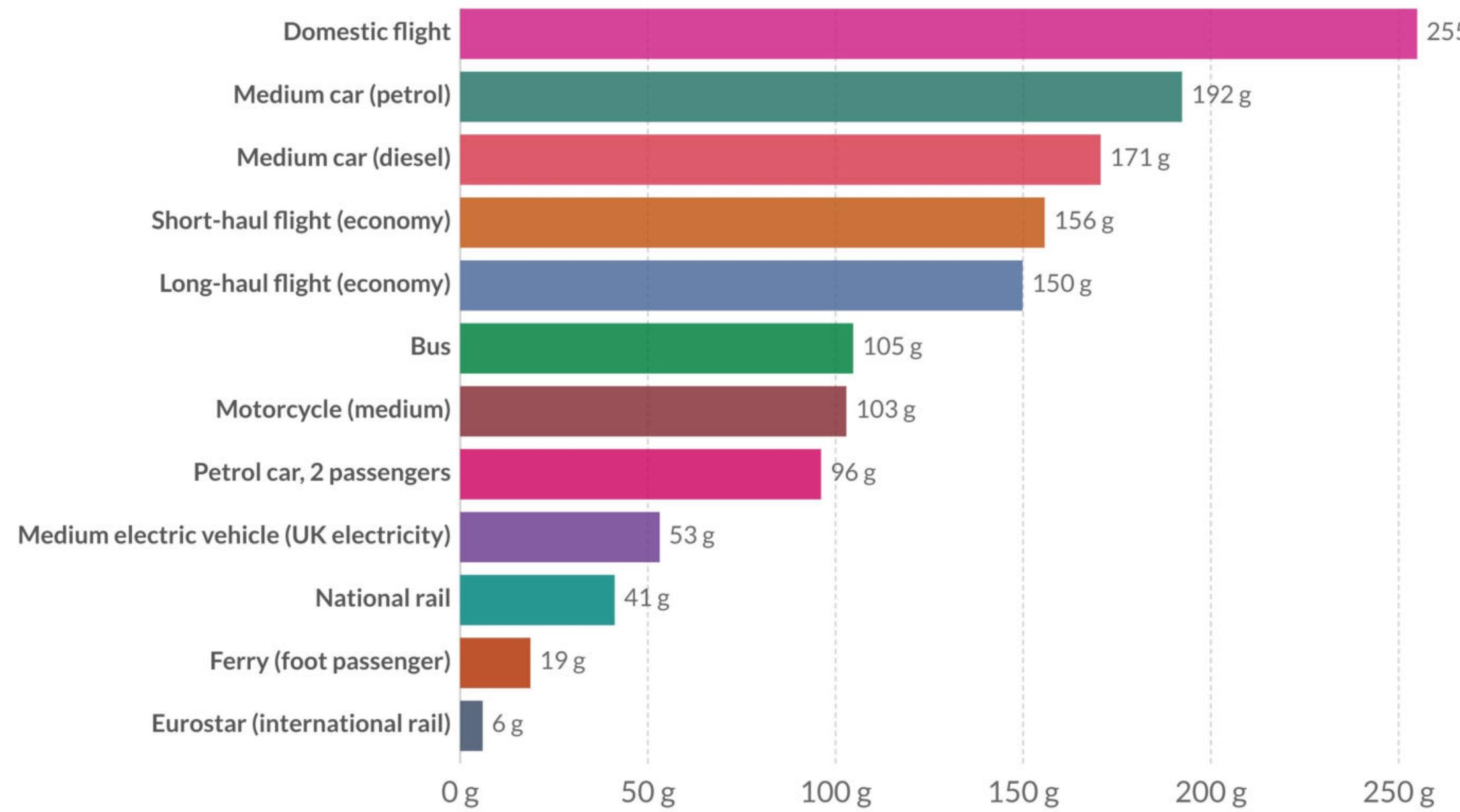
Überarbeitete Grafik mit einer divergierenden Farbpalette



# Farben (richtig) nutzen

## Carbon footprint of travel per kilometer, 2018

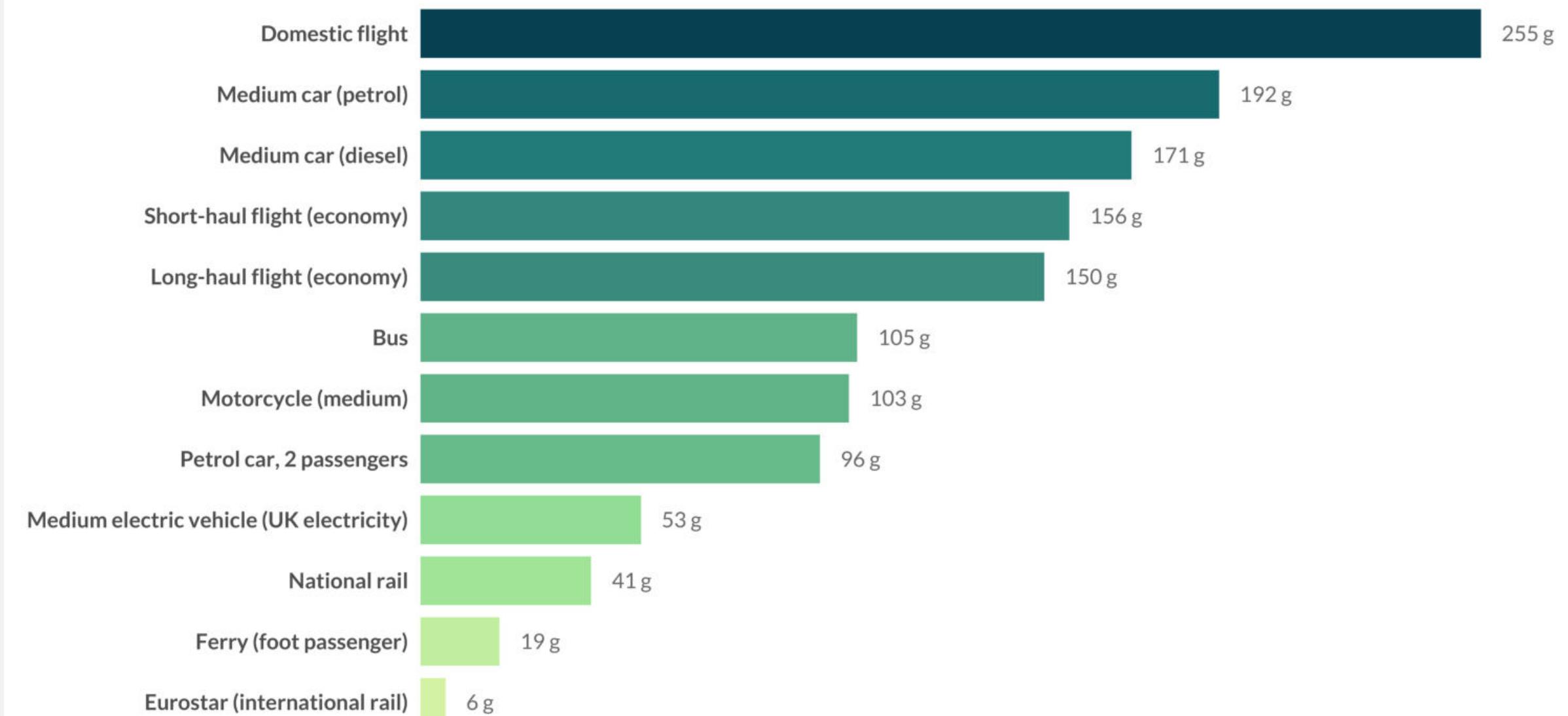
The carbon footprint of travel is measured in grams of carbon dioxide equivalents per passenger kilometer. This includes carbon dioxide, but also other greenhouse gases, and increased warming from aviation emissions at altitude.



Source: UK Department for Business, Energy & Industrial 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.

## 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.

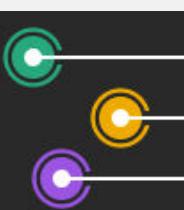


Source: UK Department for Business, Energy & Industrial Grenhouse gas reporting: conversion factors 2019.  
Note: Data is based on official conversion factors used in UK reporting. These factors may vary slightly depending on the country.

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

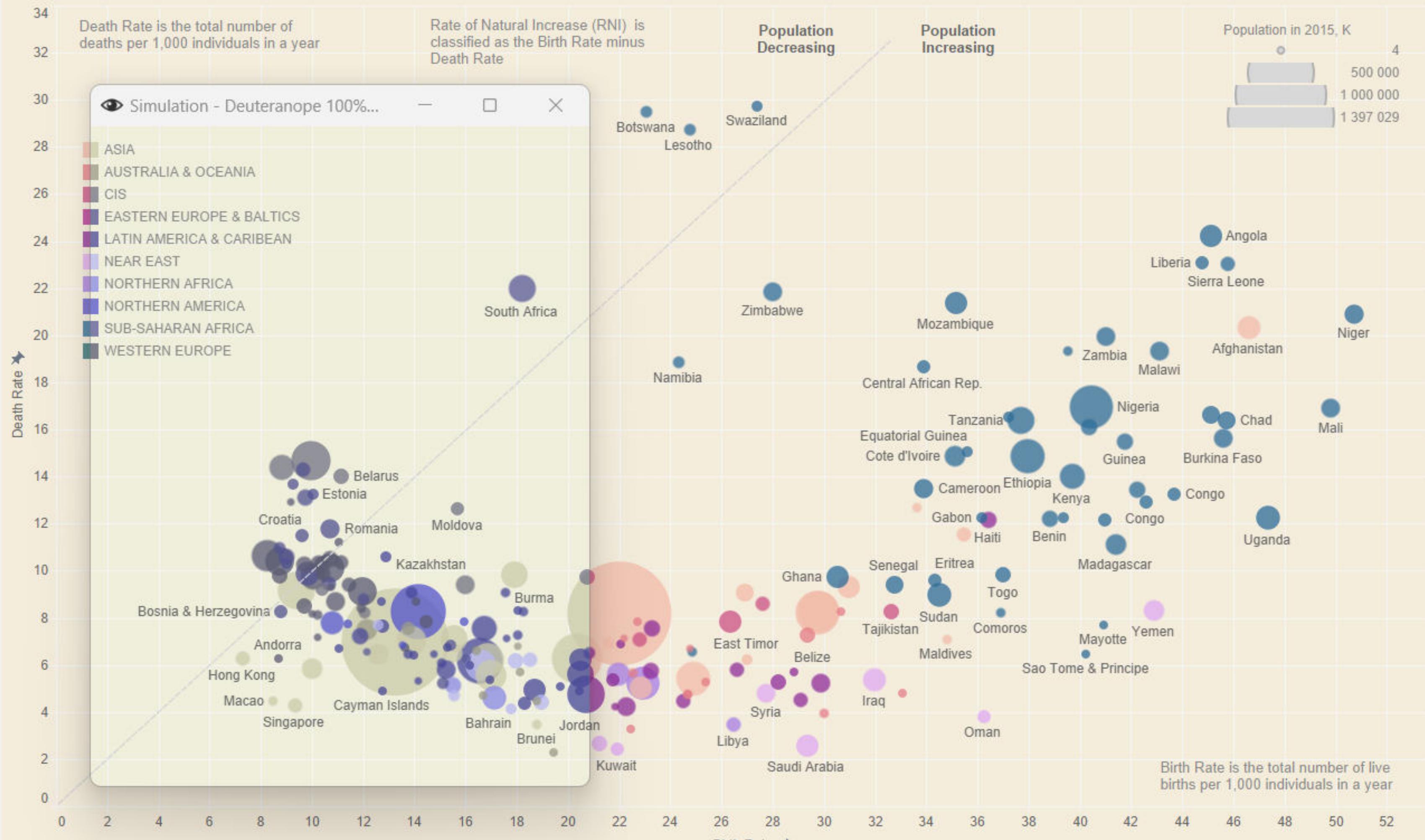
Originalgrafik mit einer zufälligen kategoriellen Farbpalette

Überarbeitete Grafik mit einer kontinuierlichen Farbpalette



# Natural Increase in the World

Birth / Death Ratio  
0,6 11,4 -6,42 35,65  
Natural Increase  
D D D D



# Natural Increase in the World



Population in 2015, K

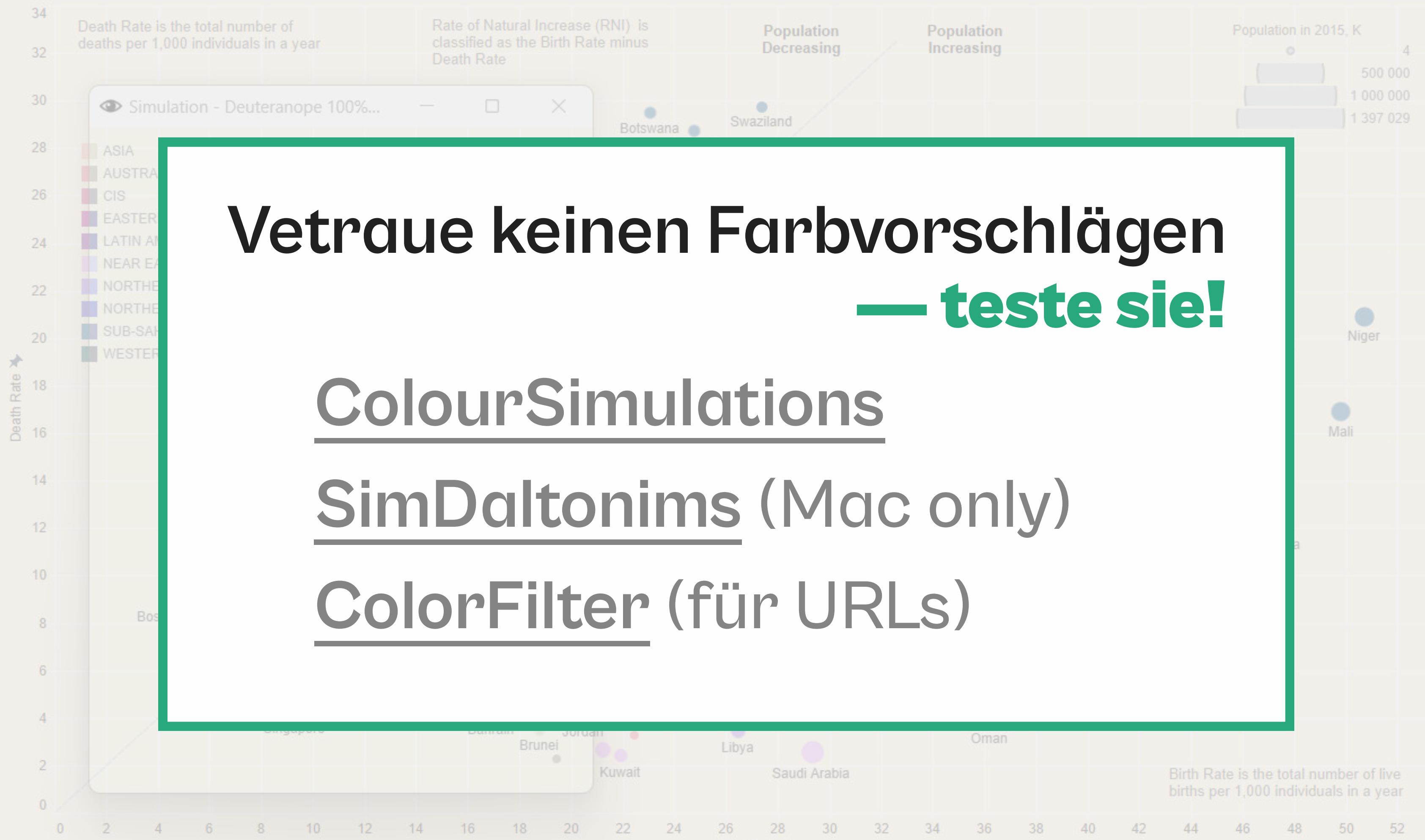


Vetraue keinen Farbvorschlägen  
— teste sie!

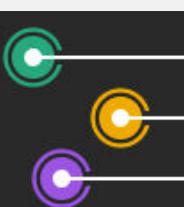
ColourSimulations

SimDaltonims (Mac only)

ColorFilter (für URLs)



Beitrag zur SWDchallenge von Alex Varlamoff — getestet mit einem Farbblindheitssimulator



# VIZ PALETTE

By: Elijah Meeks  
& Susie Lu

## PICK

Use Chroma.js



Add

Replace

Use Colorgorical

Use ColorBrewer

## EDIT

- ≡ 1 ● #2a9571 [🔗](#) ×
- ≡ 2 ● #8fb9bf [🔗](#) ×
- 4 Colors      ≡ 3 ● #dfb468 [🔗](#) ×
- ≡ 4 ● #4b8cd8 [🔗](#) ×

Add

hex  rgb

hsl

## GET

hex  rgb

hsl

String quotes  
 Object with metadata

```
[ "#2a9571",
  "#8fb9bf",
  "#dfb468",
  "#4b8cd8" ]
```

Background color: #eeeeee [🔗](#)

Font color: #212121 [🔗](#)

Charts made with [Semiotic](#)

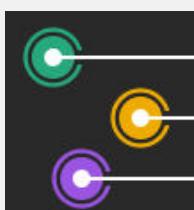
# COLORS IN ACTION

## Color Population:

No Color Deficiency - 96% Deuteranomaly - 2.7% Protanomaly - 0.66% Protanopia - 0.59% Deuteranopia - 0.56% Greyscale



[Viz Palette](#)

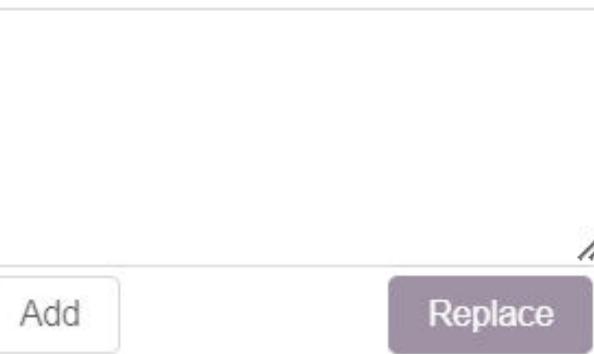


# VIZ PALETTE

By: Elijah Meeks  
& Susie Lu

## PICK

Use Chroma.js



Add

Replace

Use Colorgorical

Use ColorBrewer

## EDIT

≡ 1 ● #2a9571 ↗

x

≡ 2 ● #8fb9bf ↗

x

4 Colors

≡ 3 ● #dfb468 ↗

x

≡ 4 ● #4b8cd8 ↗

x

#hex  rgb

hsl

## GET

String quotes  
 Object with metadata

```
[ "#2a9571",
  "#8fb9bf",
  "#dfb468",
  "#4b8cd8" ]
```

#hex  rgb

hsl

# COLORS IN ACTION

Background color: ● #eeeeee ↗

Font color: ● #212121 ↗

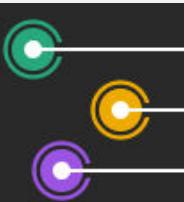
Charts made with [Semiotic](#)

### Color Population:

No Color Deficiency - 96% Deuteranomaly - 2.7% Protanomaly - 0.66% Protanopia - 0.59% Deuteranopia - 0.56% Greyscale



[Viz Palette](#)



# Wrap-Up

---



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



@CedScherer



z3tt



# Information .....

Verstehe deine Daten und sei genau.

# Story .....

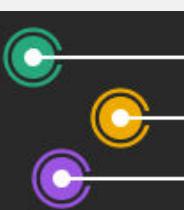
Sei dir über die Botschaft der Daten im Klaren.

# Goal .....

Wähle geeignete Grafiken, um die Geschichte zu erzählen.

# Visual Form .....

Folge Grundsätzen aus Design und Datenvisualisierung.



# Leite das Publikum (richtig)



erzähle eine Geschichte



nutze einen geeigneten Diagrammtyp



setze Farben korrekt und sinnvoll ein



verwende Annotierungen und “visual cues”

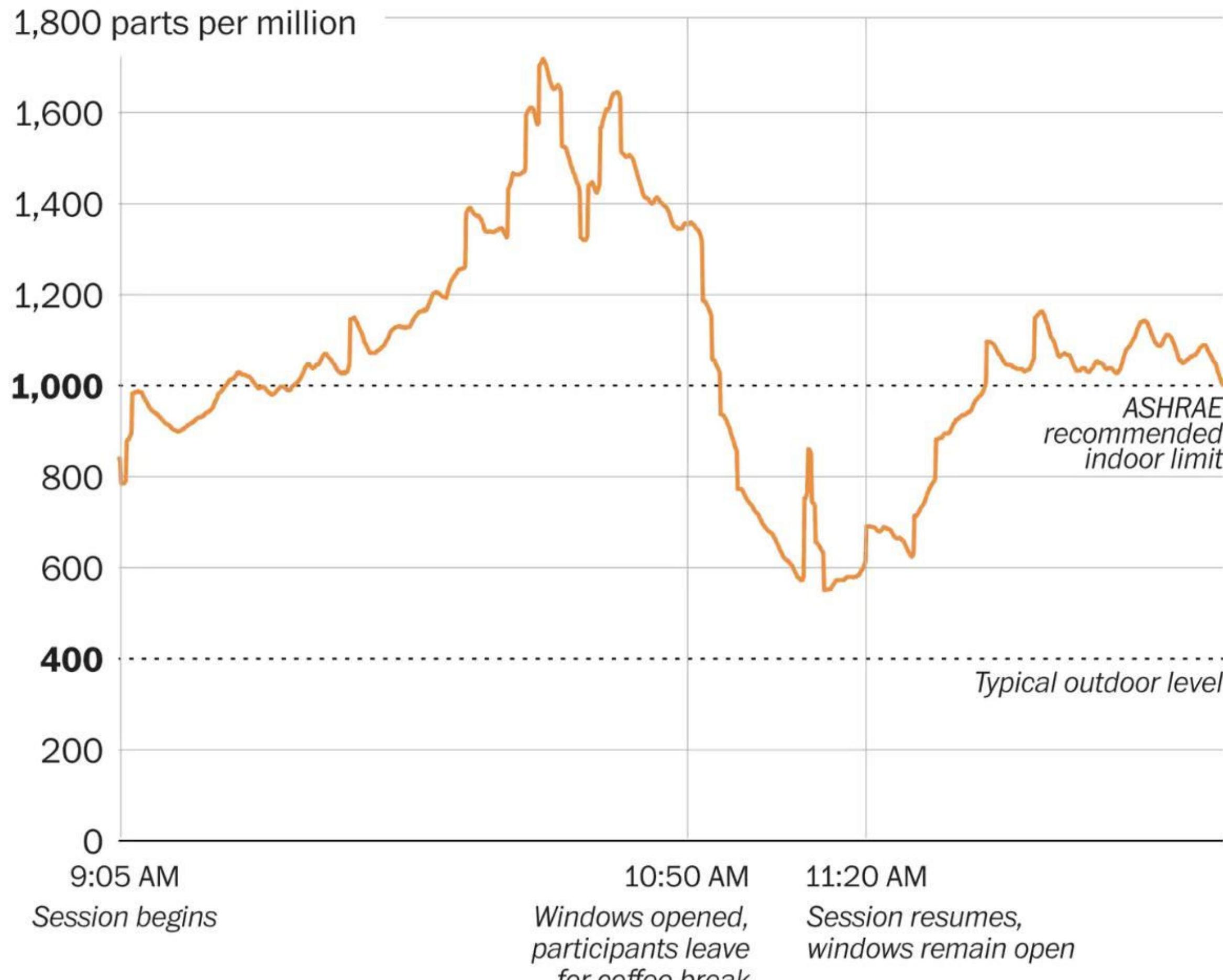


versuche Daten direkt zu beschriften



# Clearing the air

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



Source: Adam Ginsburg

THE WASHINGTON POST

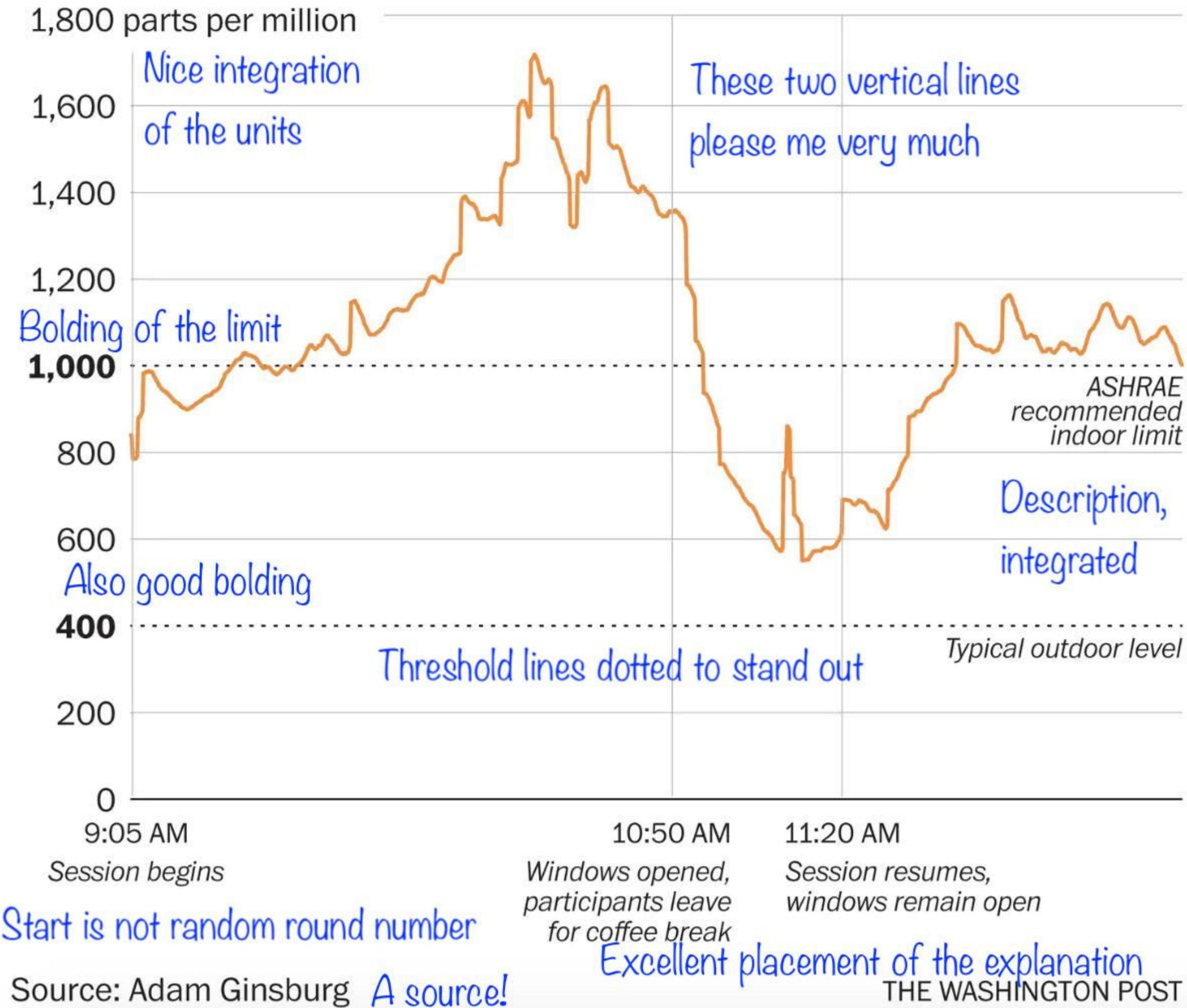
[“Clearin the Air” von Adam Ginsburg \(Washington Post\)](#)



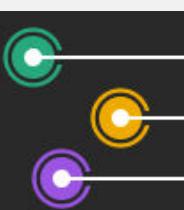
# Clearing the air

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



... mit Gedanken von Francis Gagnon (Voilà)



# vielen Dank!



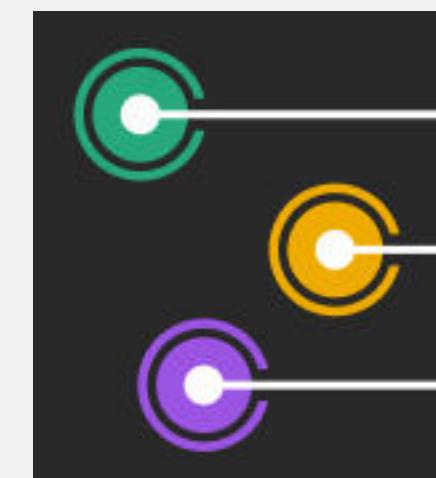
[cedricscherer.com](http://cedricscherer.com)



@CedScherer



[z3tt](https://github.com/z3tt)



**CÉDRIC SCHERER**  
Data Visualization & Information Design