

HANDS ON DATENVISUALISIERUNG

LETTY, KAREN OTTE

AGENDA HEUTE

- ▶ 9:30 - 10:30 Vorstellungsrunde/ Erwartungen
- ▶ 10:30 - 11:30 Kurze Einführung ins Thema I
Kurze Pause
- ▶ 11:45 - 12:30 Kurze Einführung ins Thema II
- ▶ 12:30 - 13:00 Praktischen Tools und weiteres Vorgehen

WER SIND WIR?

VORSTELLUNGSRUNDE – KATJA DITTRICH

► Medizininformatik

► Visual Computing

► selbstständige Softwareentwicklerin & Dozentin

Schwerpunkte: App Entwicklung, Visualisierungen für Web Anwendungen, (Datenbanken + Anbindungen)

Programmiersprachen: JavaScript, Python

VORSTELLUNGSRUNDE – KAREN OTTE

- ▶ Medizininformatikerin
- ▶ Promotionsstudentin (rer. medic.), Charité
Universitätsmedizin Berlin
- ▶ Wissenschaftliche Leitung bei Motognosis GmbH

Medizinische Forschung: Bewegungsanalyse,
Signalverarbeitung, Algorithmen- und Software Entwicklung,
Statistische Auswertung

Programmiersprachen: C#, Python, MATLAB, (R)

ZIELE DES SEMINARS

- ▶ Hands on!
- ▶ Visualisierungen für das eigene Projekt entwickeln

Bonus:

- ▶ Verstehen was gute Visualisierungen ausmacht
- ▶ Praktische Tipps für das Erstellen von Visualisierungen
- ▶ Gemeinsam Probleme überwinden

TERMINE

- ▶ 26.10. - Auftakt(heute)
- ▶ 02.11. - Grundlagen von Visualisierungen
- ▶ 09.11. - Praktische Anwendung und Best Practices
- ▶ 16.11. - "Künstlerische Methoden"
- ▶ 23.11. - Vorstellung und Besprechung der Ergebnisse

Zwischenzeitlich: Arbeiten am eigenen Datensatz;
Gern Fragen!

WER SEID IHR?

VORSTELLUNGSRUNDE – UND DU?

- ▶ Wer bist du?
- ▶ Woher kommst du?
Institut/Einrichtung
- ▶ Mit welchen Daten arbeitest du?
Z.B. Big/Small, Zeitreihen/Ortsdaten
- ▶ Kannst du programmieren?
Ja/Nein, Sprache?
- ▶ Welche Tools/Programmenutzt du?
Excel/SPSS

ERWARTUNGEN?

Existierende Probleme?

Wie stellt man mehrdimensionale Daten in 2D dar?
Wie erstellt man 3D Visualisierungen?
Statistische Kennwerte in Plots darstellen?

Zusammenhänge in Netzwerkgraphiken verständlich machen?

Darstellung von linearen Modellen? Darstellung von Einfluss von einzelnen Kennwerten unter Berücksichtigung von Ko-Variablen

Fragen?

Wie vergleicht man mehrdimensionale Graphiken?
Wie stellt man Interaktionen zwischen Einheiten dar?

Wie stellt man Zeitreihen / zeitlichen Verlauf in Daten dar?
(z.B. für Publikationen, explorative Datenanalyse)

Wie stellt man Korrelationen dar?
Wie vergleicht Signale?

Wie kann man Visualisierungspipelines wiederverwenden?

Wünsche?

Eingehen auf Grundlagen von Wahrnehmung (Buch: Beautiful Evidence?)

Einführung in Programmierung (R und Python)

Hilfe für schnelle, schöne Standardvisualisierungen. Was gibt es an Templates? Evtl Tooling?

**WARUM
VISUALISIERUNG?**

WARUM IST DATENVISUALISIERUNG WICHTIG??

“Children vaccinated with MMR had significantly higher mental BSID-II scores in the 36th month than those vaccinated with single measles vaccine (103.8 ± 10.3 vs. 97.2 ± 11.2 , $p=0.004$) [...]. Neither results of Raven test nor WISC-R were significantly different between groups of children exposed to MMR and single measles vaccine. The results of MDI in the 36th month in children vaccinated with MMR vs. vaccinated with single measles vaccine became non-significant ($\beta=4.7$, $p=0.056$) after standardization to child's gender, maternal education, family economical status, maternal IQ, birth order and passive tobacco smoking. Results of MDI in the 24th month as well as WISC-R and Raven in MMR and monovalent group didn't differ significantly.”

Quelle: Mrozek-Budzyn, D. et al. *Vaccine* vol. 31,22 (2013): 2551-7. doi:10.1016/j.vaccine.2013.03.057

WARUM IST DATENVISUALISIERUNG WICHTIG??

“Children vaccinated with MMR had significantly higher mental BSID-II scores in the 36th month than those vaccinated with single measles vaccine (103.8 ± 10.3 vs. 97.2 ± 11.2 , $p=0.004$) [...]. Neither results of Raven test nor WISC-R were significantly different between groups of children exposed to MMR and single measles vaccine. The results of MDI in the 36th month in children vaccinated with MMR vs. vaccinated with single measles vaccine became non-significant ($\beta=4.7$, $p=0.056$) after standardization to child's gender, maternal education, family economical status, maternal IQ, birth order and passive tobacco smoking. Results of MDI in the 24th month as well as WISC-R and Raven in MMR and monovalent group didn't differ significantly.”

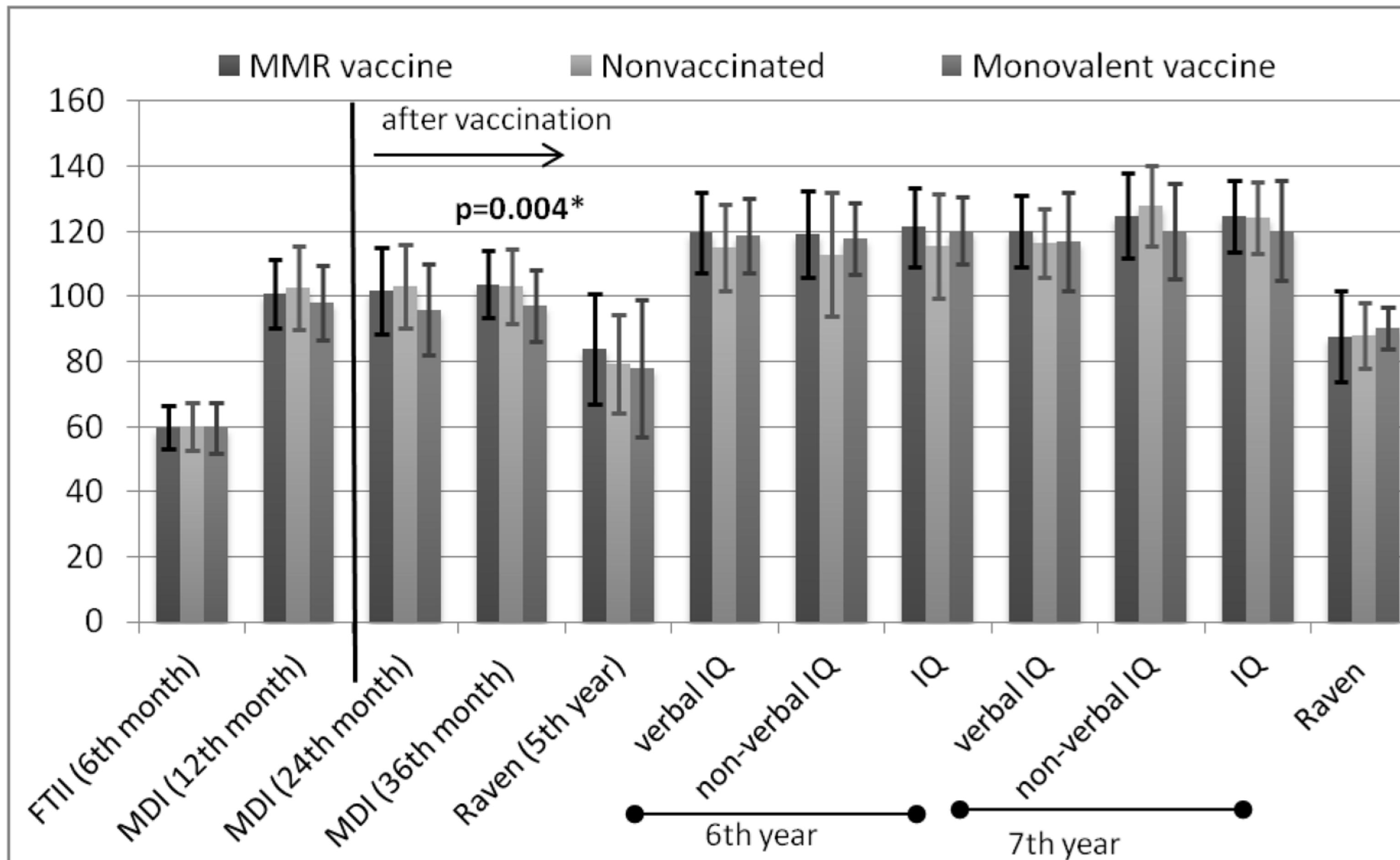
Quelle: Mrozek-Budzyn, D. et al. Vaccine vol. 31,22 (2013): 2551-7. doi:10.1016/j.vaccine.2013.03.057

WARUM IST DATENVISUALISIERUNG WICHTIG??

test	age	MMR vs. non-vaccinated			MMR vs. monovalent vaccine		
		β^*	95% CI	p	β^*	95% CI	p
MDI of BSID II	24 th month of life	-3.7	-9.6; 2.1	0.212	1.0	-5.1; 7.0	0.749
MDI of BSID II	36 th month of life	-3.4	-8.2; 1.5	0.172	4.7	-0.1; 9.5	0.056
Raven (centilles)	5 th year of life	-3.0	-13.5; 7.5	0.574	3.1	-7.8; 14.0	0.572
WISC-R	6 th year of life						
verbal IQ **		-2.5	-11.1; 6.1	0.568	-3.5	-11.8; 4.7	0.394
non-verbal IQ **		-2.8	-11.1; 5.5	0.502	-0.3	-8.4; 7.7	0.936
IQ **		-3.0	-11.0; 5.1	0.464	-2.23	-10.1; 5.6	0.572
WISC-R	7 th year of life						
verbal IQ **		3.5	-4.2; 11.2	0.370	-0.7	-9.6; 8.1	0.867
non-verbal IQ **		-0.9	-10.1; 8.2	0.843	3.4	-7.2; 14.1	0.523
IQ **		1.5	-6.0; 9.1	0.688	1.3	-7.4; 10.1	0.761
Raven (centilles) **	8 th year of life	-2.3	-13.9; 9.3	0.694	0	-12.8; 12.7	0.997

Quelle: Mrozek-Budzyn, D. et al. Vaccine vol. 31,22
(2013): 2551-7. doi:10.1016/j.vaccine.2013.03.057

WARUM IST DATENVISUALISIERUNG WICHTIG??



* MMR vs. monovalent vaccine

Quelle: Mrozek-Budzyn, D. et al. Vaccine vol. 31,22 (2013): 2551-7. doi:10.1016/j.vaccine.2013.03.057

ZIELE VON VISUALISIERUNGEN

- ▶ Kommunikation / Präsentation
- ▶ Explorative Analyse
- ▶ Konfirmative Analyse

- ▶ Mehr dazu nächsten Montag

**WAS SIND
VISUALISIERUNGEN?**



esa

1986

Quelle: ESA

1995

Quelle: ESA

2010

Quelle: ESA



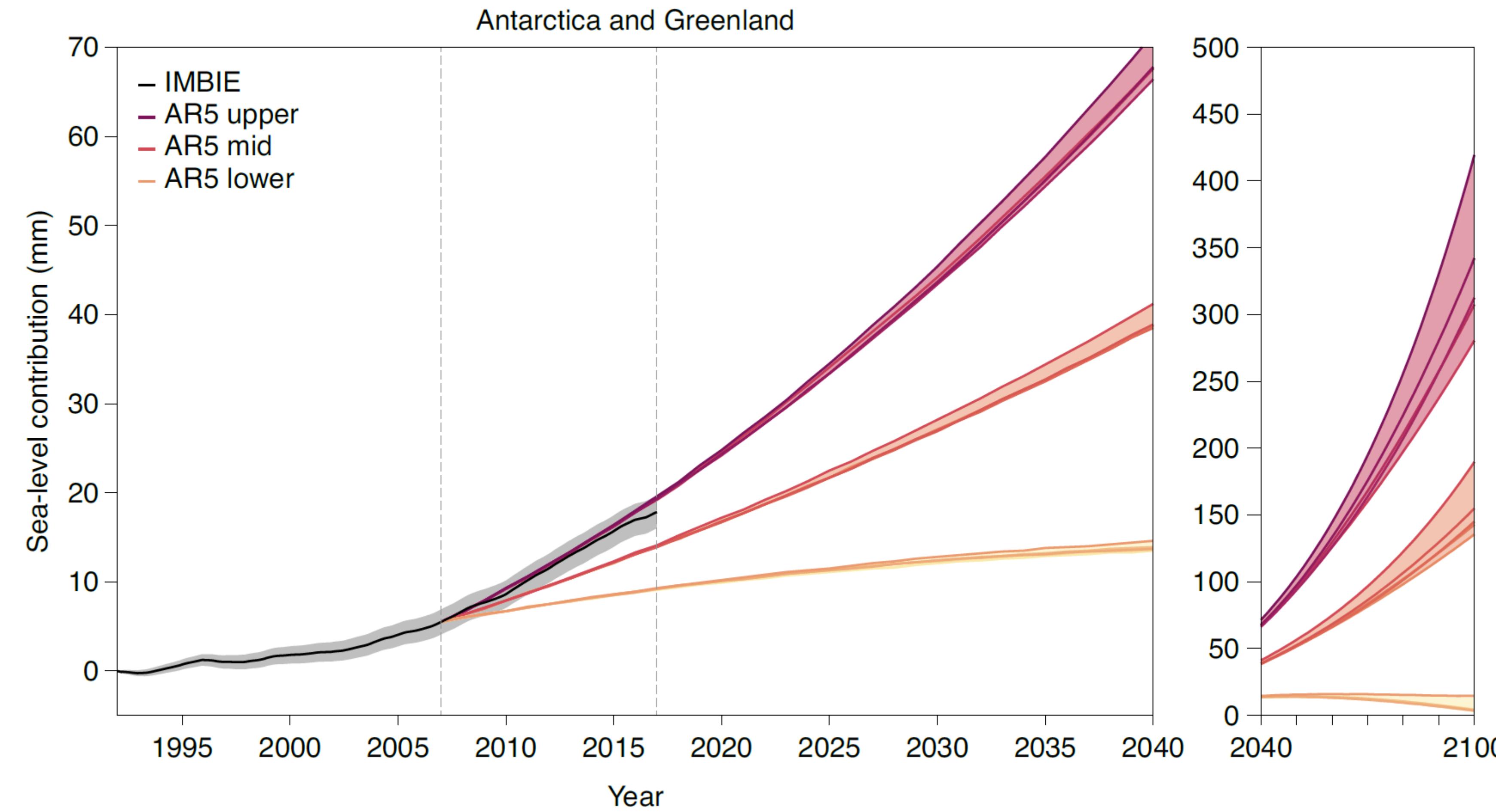
1986



2010

Quelle: ESA

SIGNALE UND ZEITREIHEN



Quelle: ESA

DIAGRAMME

ICE

The ice sheets that cover Greenland and Antarctica hold about 99% of Earth's freshwater, with the potential to raise sea level dramatically.

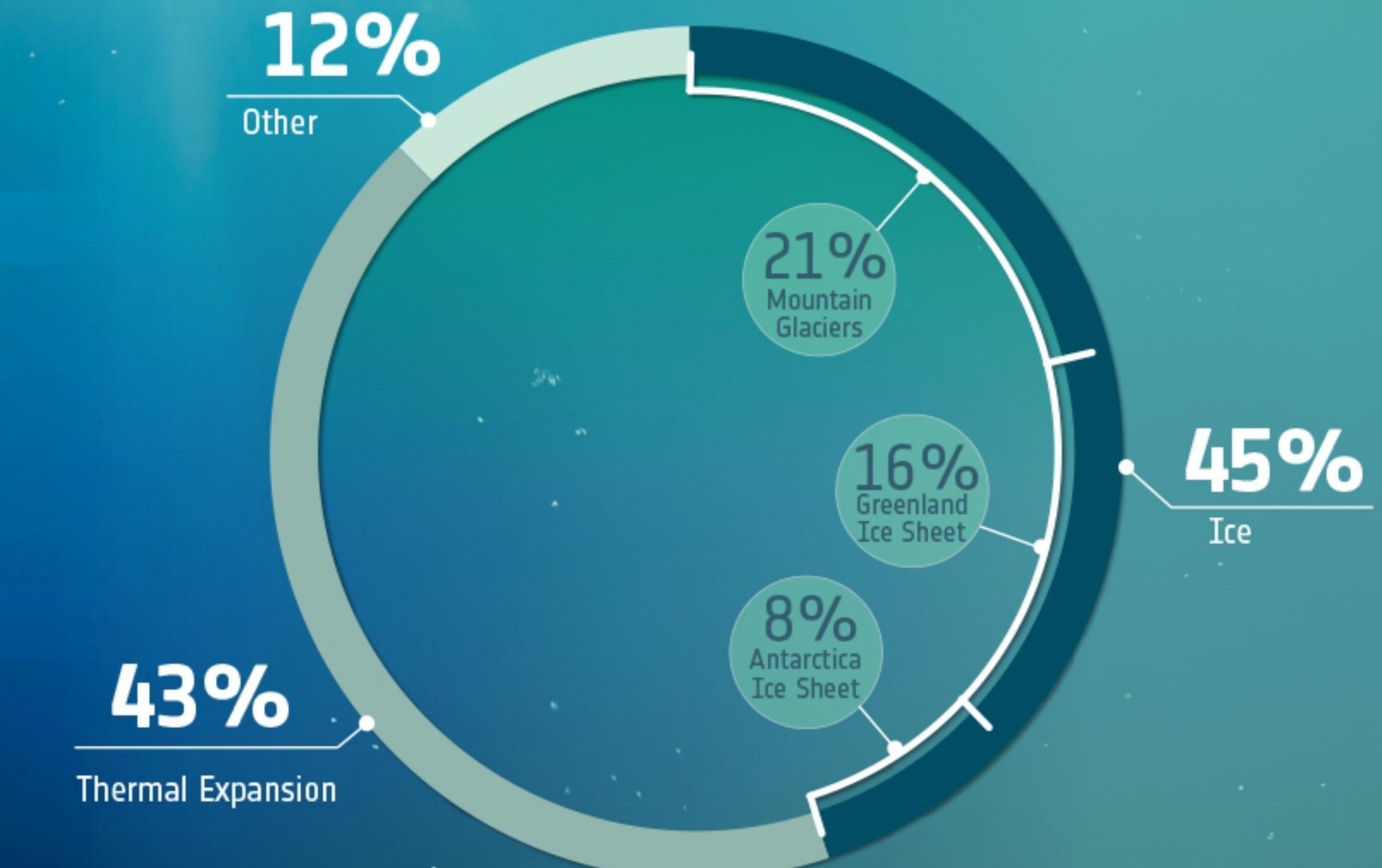
WARMING

The average global temperature has increased by about 1°C over the last 150 years. This is largely because of greenhouse gases being emitted into the atmosphere.

THERMAL EXPANSION

As the ocean warms in response to climate change, seawater expands and, as a result, sea level rises.

CONTRIBUTORS TO SEA-LEVEL RISE
(1993-2015)



PIKTOGRAMME



ICE

The ice sheets that cover Greenland and Antarctica hold about 99% of Earth's freshwater, with the potential to raise sea level dramatically.



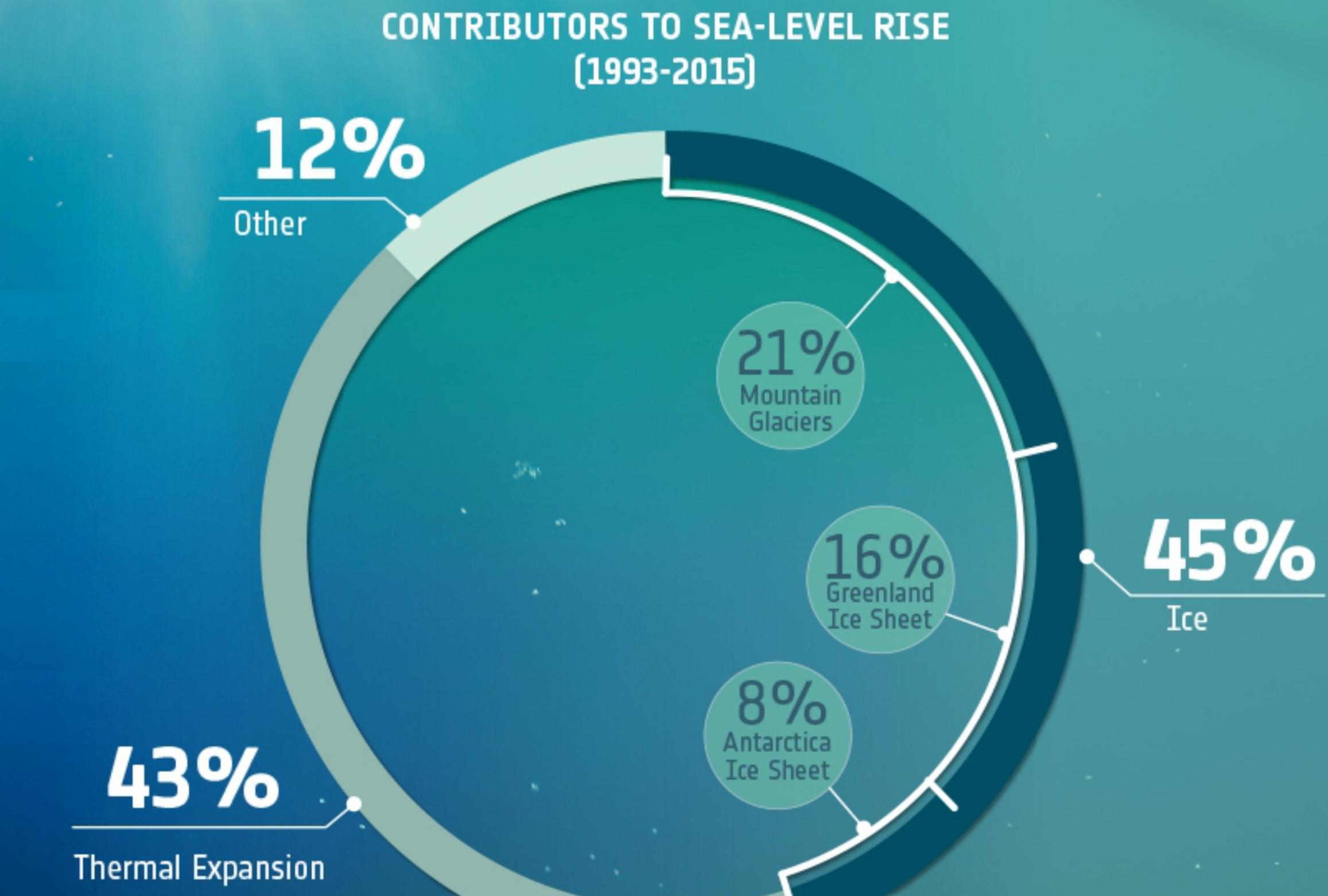
WARMING

The average global temperature has increased by about 1°C over the last 150 years. This is largely because of greenhouse gases being emitted into the atmosphere.



THERMAL EXPANSION

As the ocean warms in response to climate change, seawater expands and, as a result, sea level rises.



WORTWOLKEN

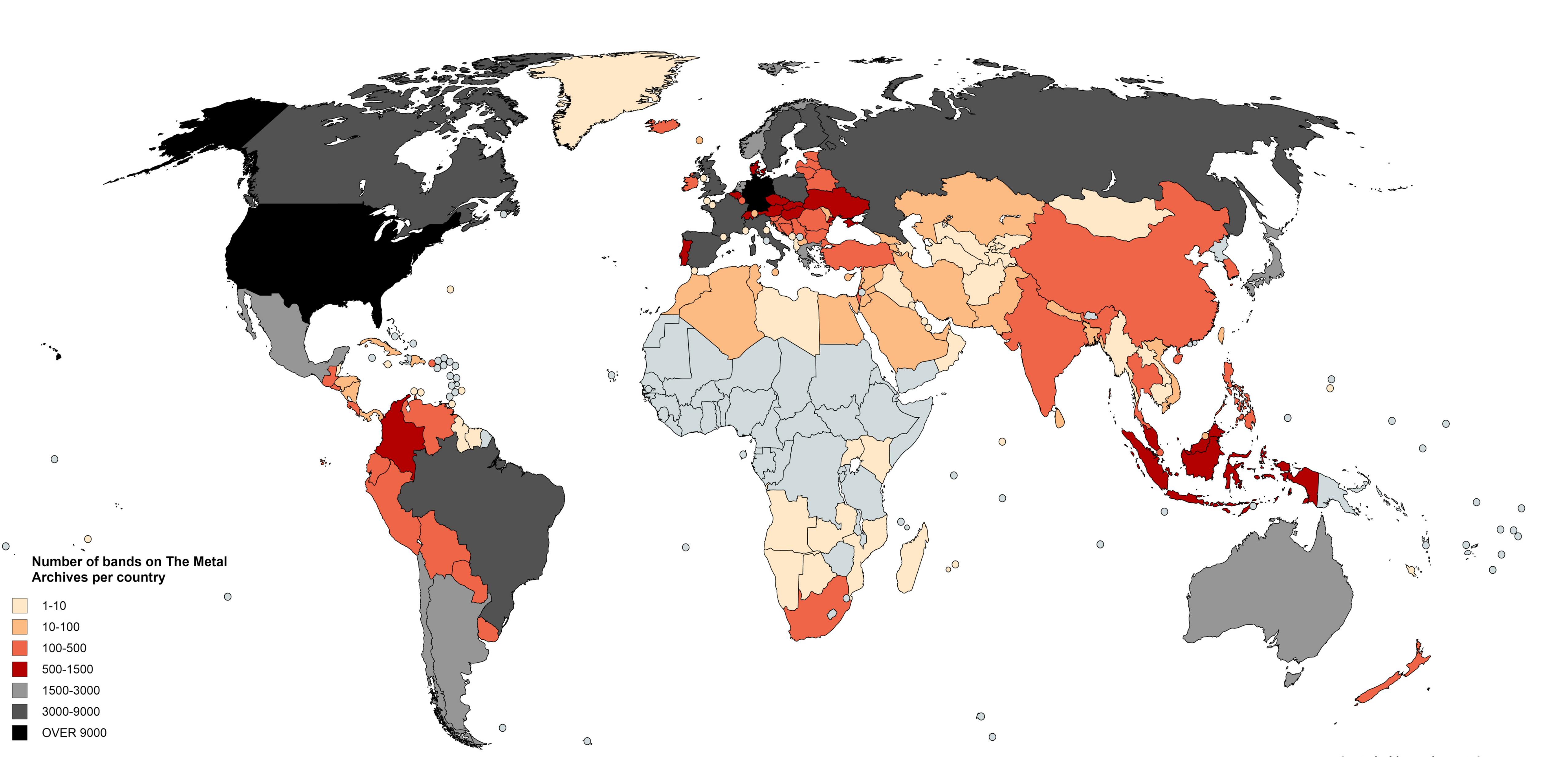


Quelle: The Data Journalism Handbook, O'reilly

TABELLEN

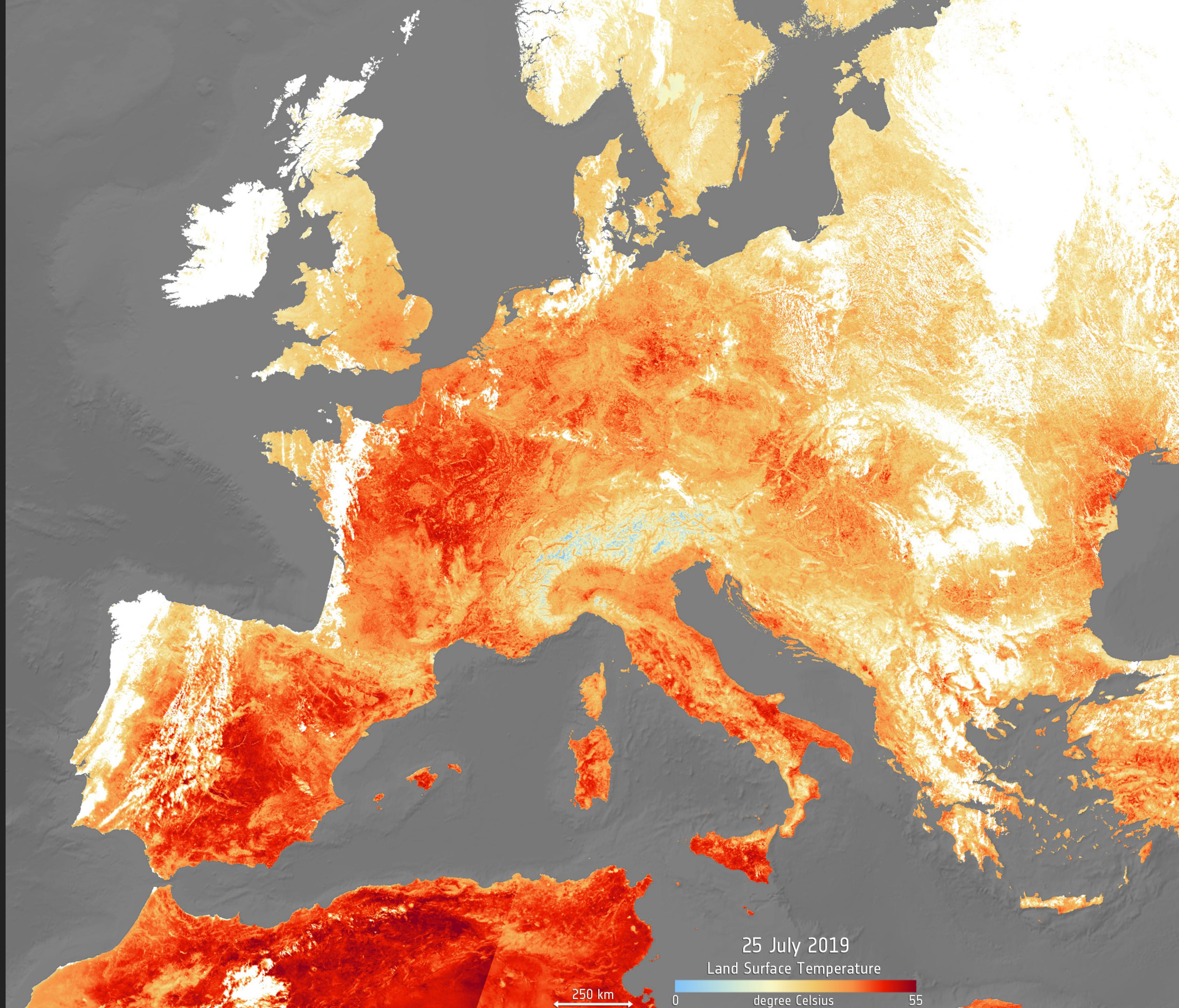
test	age	MMR vs. non-vaccinated			MMR vs. monovalent vaccine		
		β^*	95% CI	p	β^*	95% CI	p
MDI of BSID II	24 th month of life	-3.7	-9.6; 2.1	0.212	1.0	-5.1; 7.0	0.749
MDI of BSID II	36 th month of life	-3.4	-8.2; 1.5	0.172	4.7	-0.1; 9.5	0.056
Raven (centilles)	5 th year of life	-3.0	-13.5; 7.5	0.574	3.1	-7.8; 14.0	0.572
WISC-R	6 th year of life						
verbal IQ **		-2.5	-11.1; 6.1	0.568	-3.5	-11.8; 4.7	0.394
non-verbal IQ **		-2.8	-11.1; 5.5	0.502	-0.3	-8.4; 7.7	0.936
IQ **		-3.0	-11.0; 5.1	0.464	-2.23	-10.1; 5.6	0.572
WISC-R	7 th year of life						
verbal IQ **		3.5	-4.2; 11.2	0.370	-0.7	-9.6; 8.1	0.867
non-verbal IQ **		-0.9	-10.1; 8.2	0.843	3.4	-7.2; 14.1	0.523
IQ **		1.5	-6.0; 9.1	0.688	1.3	-7.4; 10.1	0.761
Raven (centilles) **	8 th year of life	-2.3	-13.9; 9.3	0.694	0	-12.8; 12.7	0.997

Quelle: Mrozek-Budzyn, D. et al. Vaccine vol. 31,22 (2013): 2551-7. doi:10.1016/j.vaccine.2013.03.057



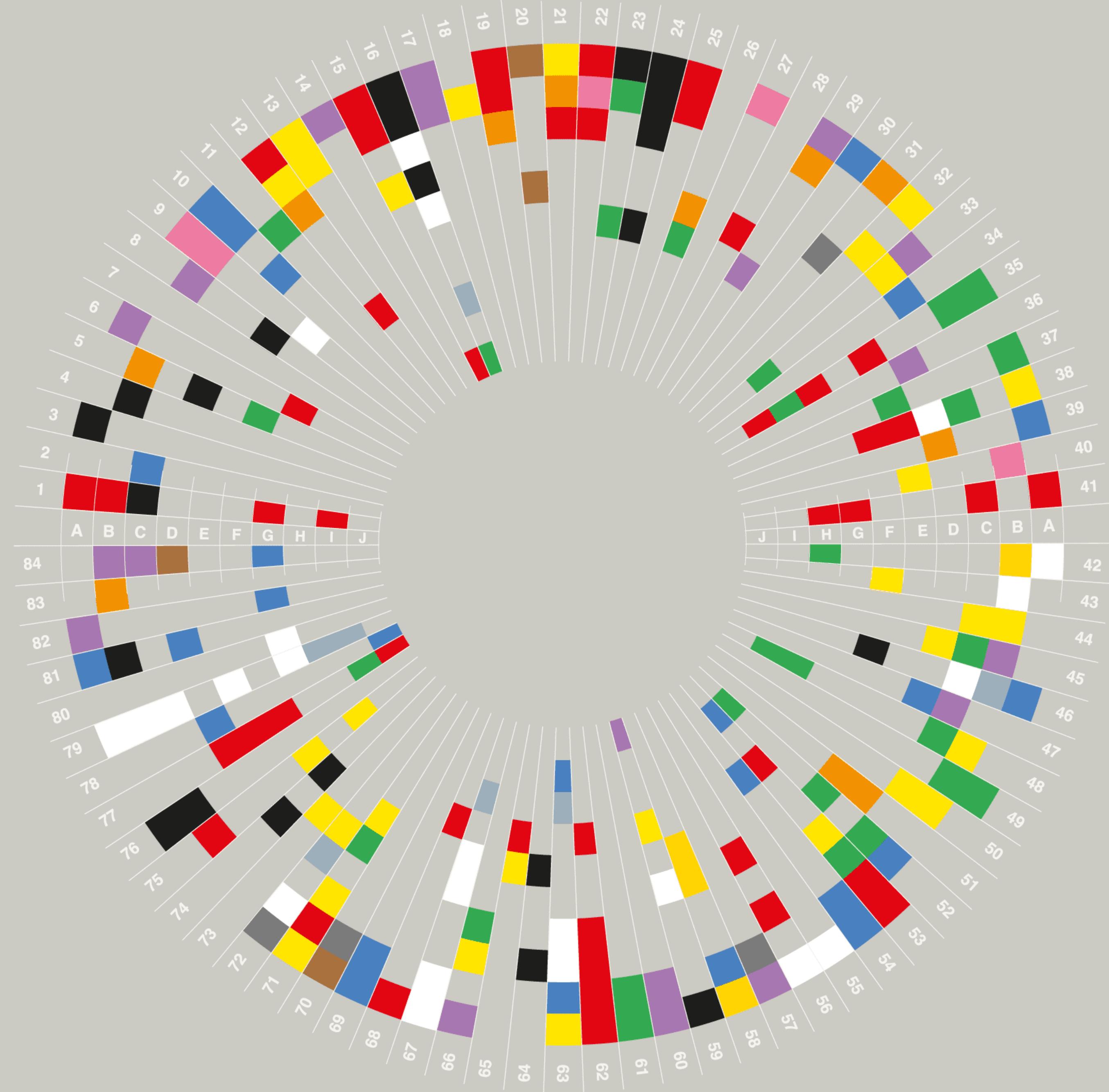
Created with [mapchart.net](#) ©

Quelle: imgur

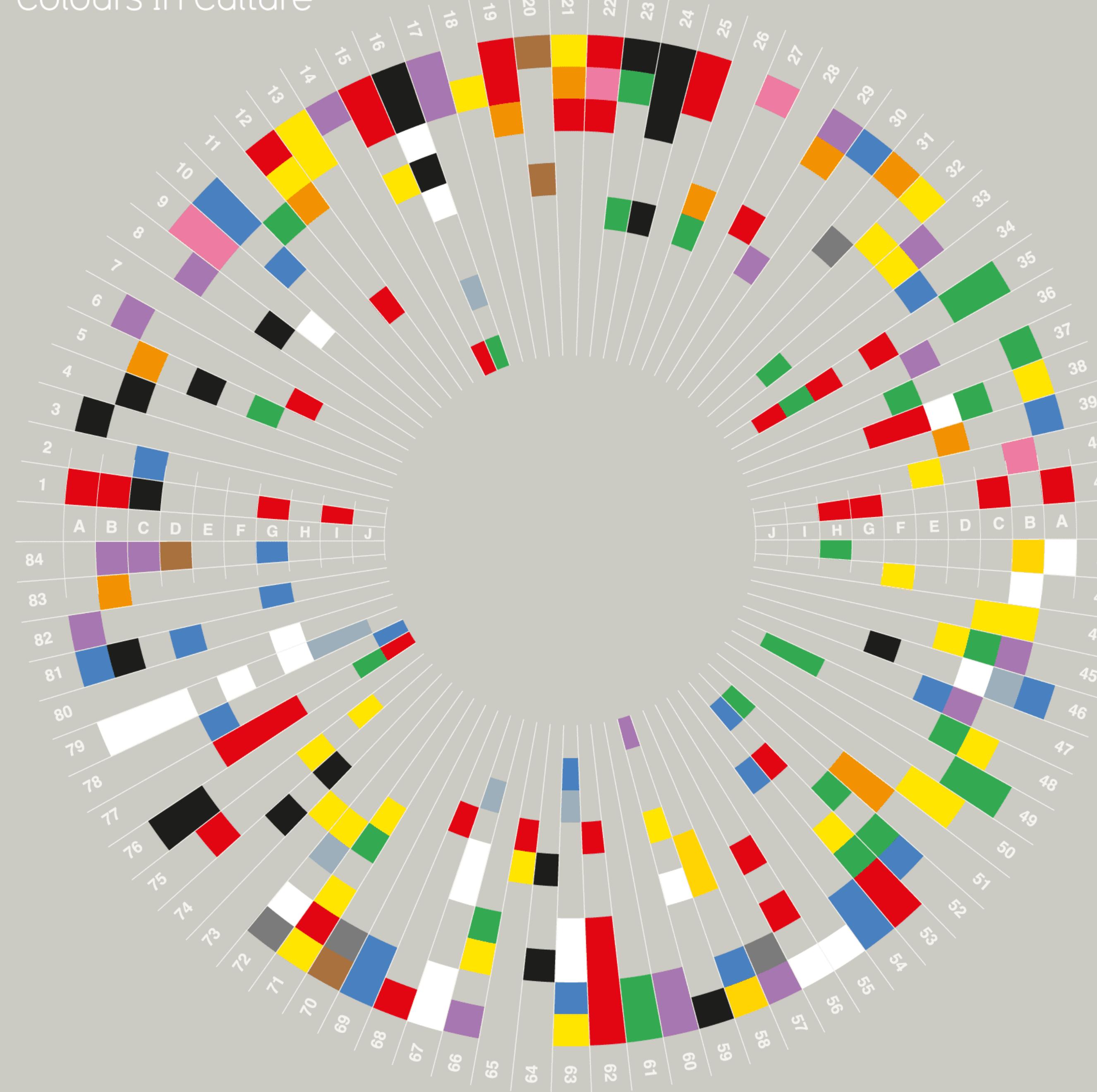


Quelle: ESA





Colours In Culture



A Western / American
B Japanese
C Hindu
D Native American
E Chinese
F Asian
G Eastern European
H Arab
I African
J South American

1 Anger
2 Art / Creativity
3 Authority
4 Bad Luck
5 Balance
6 Beauty
7 Calm
8 Celebration
9 Children
10 Cold
11 Compassion
12 Courage
13 Cowardice
14 Cruelty
15 Danger
16 Death
17 Decadence
18 Deceit
19 Desire
20 Earthy
21 Energy
22 Erotic
23 Eternity
24 Evil
25 Excitement
26 Family
27 Femininity
28 Fertility
29 Flamboyance
30 Freedom
31 Friendly
32 Fun
33 God
34 Gods
35 Good Luck
36 Gratitude

37 Growth
38 Happiness
39 Healing
40 Healthy
41 Heat
42 Heaven
43 Holiness
44 Illness
45 Insight
46 Intelligence
47 Intuition
48 Religion
49 Jealousy
50 Joy
51 Learning
52 Life
53 Love
54 Loyalty
55 Luxury
56 Marriage
57 Modesty
58 Money
59 Mourning
60 Mystery
61 Nature
62 Passion
63 Peace
64 Penance
65 Power
66 Personal power
67 Purity
68 Radicalism
69 Rational
70 Reliable
71 Repels Evil
72 Respect

73 Royalty
74 Self-cultivation
75 Strength
76 Style
77 Success
78 Trouble
79 Truce
80 Trust
81 Unhappiness
82 Virtue
83 Warmth
84 Wisdom

Yellow
Gold
Grey
Silver

Quelle: Davic McCandless // InformationIsBeautiful.net

source: Pantone, ColorMatters & web sources

**WAS SIND
VISUALISIERUNGEN?**

A GRAPHIC IS NOT 'DRAWN'
ONCE AND FOR ALL

Bertin(1981)

WAS SIND VISUALISIERUNGEN?

... es gibt verschiedene Definitionen.

Veranschaulichung von [abstrakte] Daten
und ihren Zusammenhänge in einer
graphischen bzw. visuellen Form.

WAS SIND VISUALISIERUNGEN?

„Ziel der Visualisierung ist es, abstrakte Daten und Informationen graphisch so zu präsentieren, dass ihre relevanten Charakteristika intuitiv erfasst werden können.“

- Heidrun Schuhmann

WAS SIND VISUALISIERUNGEN?

„Bilder sollen so aufgebaut sein, dass der Betrachter in der Lage ist, nicht nur zu sehen, sondern auch zu erkennen, zu verstehen und zu bewerten.“

- Heidrun Schuhmann

WAS SIND VISUALISIERUNGEN?

„Bilder sollen so aufgebaut sein, dass der Betrachter in der Lage ist, nicht nur zu sehen, sondern auch zu **erkennen**, zu **verstehen** und zu **bewerten**.“

- Heidrun Schuhmann

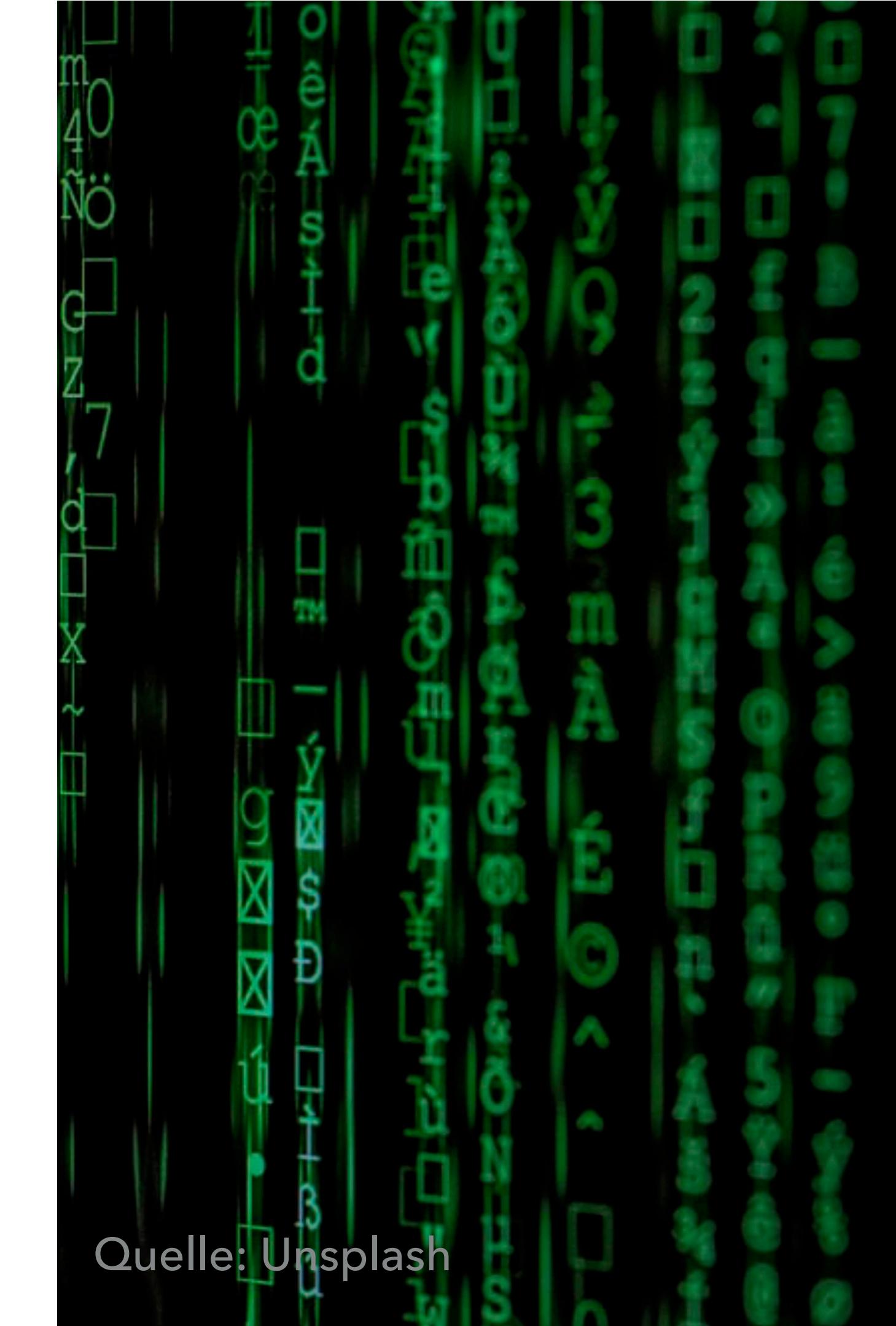
**WAS
FÜR DATEN?**

WAS FÜR DATEN GIBT ES?

... und welche verwendet ihr?

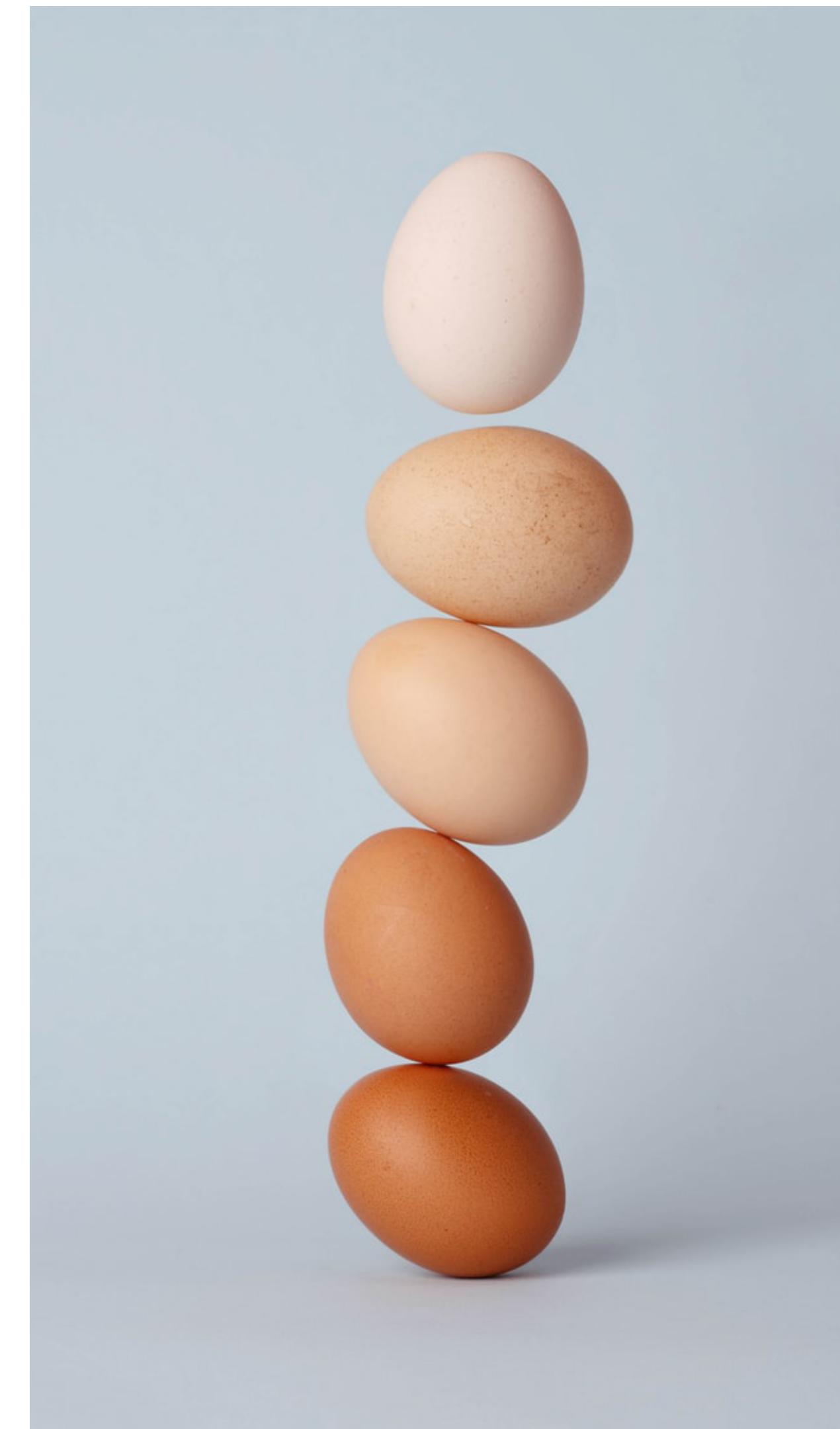
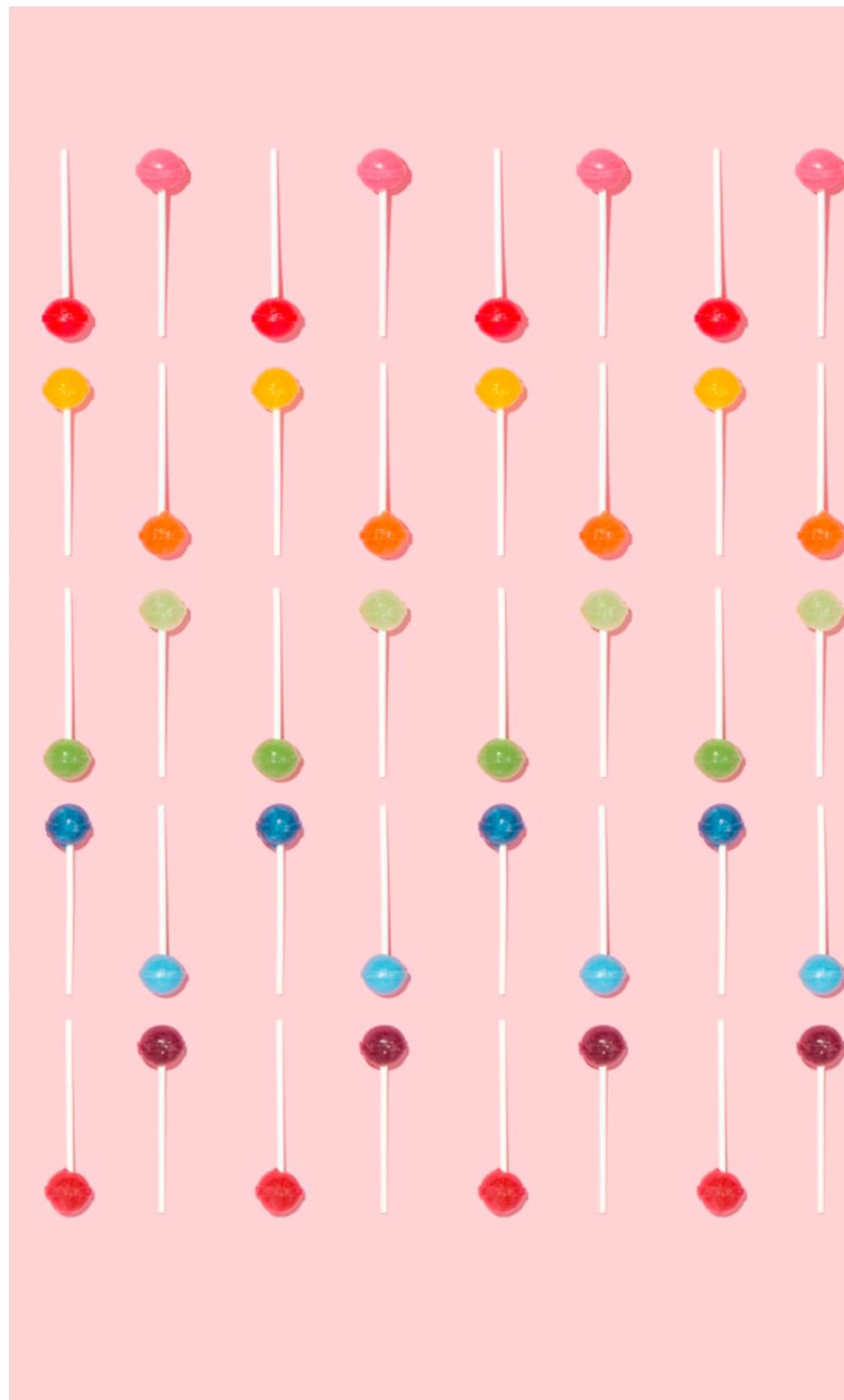
- ▶ Daten dienen als Grundlage für Visualisierungen
- ▶ Datenquellen sind vielfältig
- ▶ Gemeinsame Nomenklaturschaffen

UNSTRUKTURIERT VS. SEMIS- VS. STRUKTURIERT



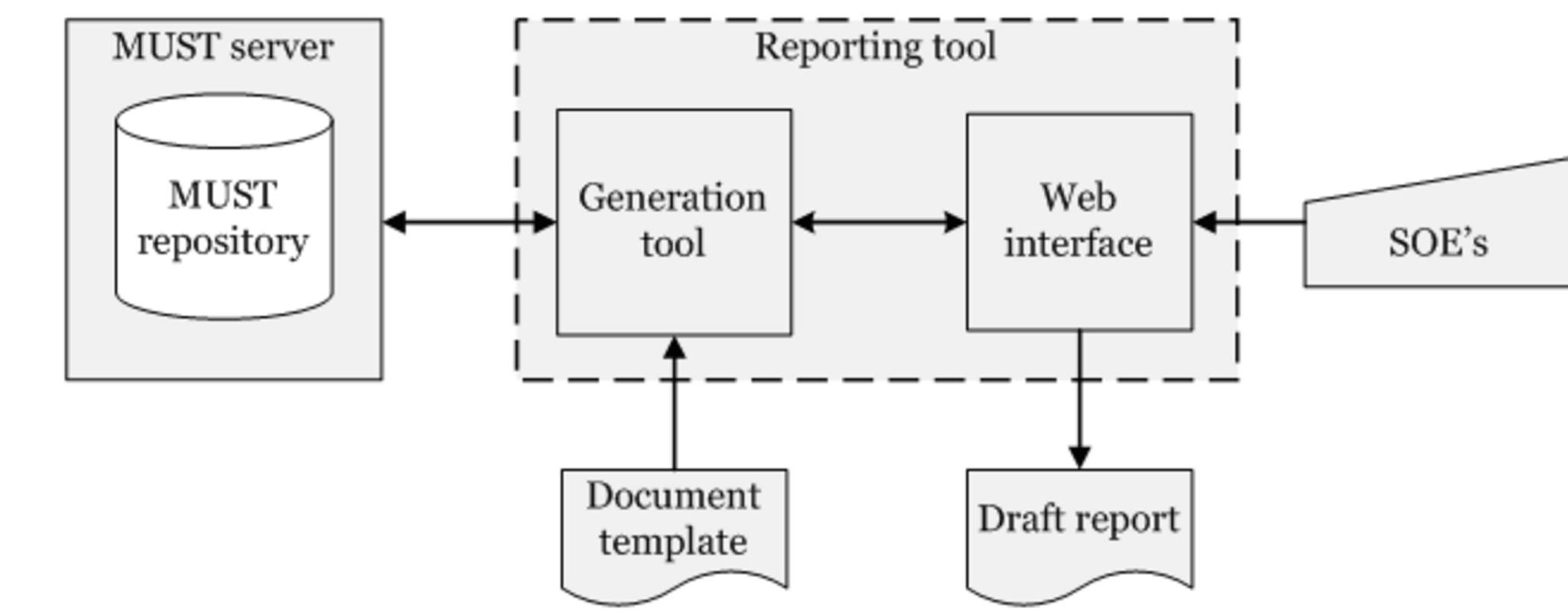
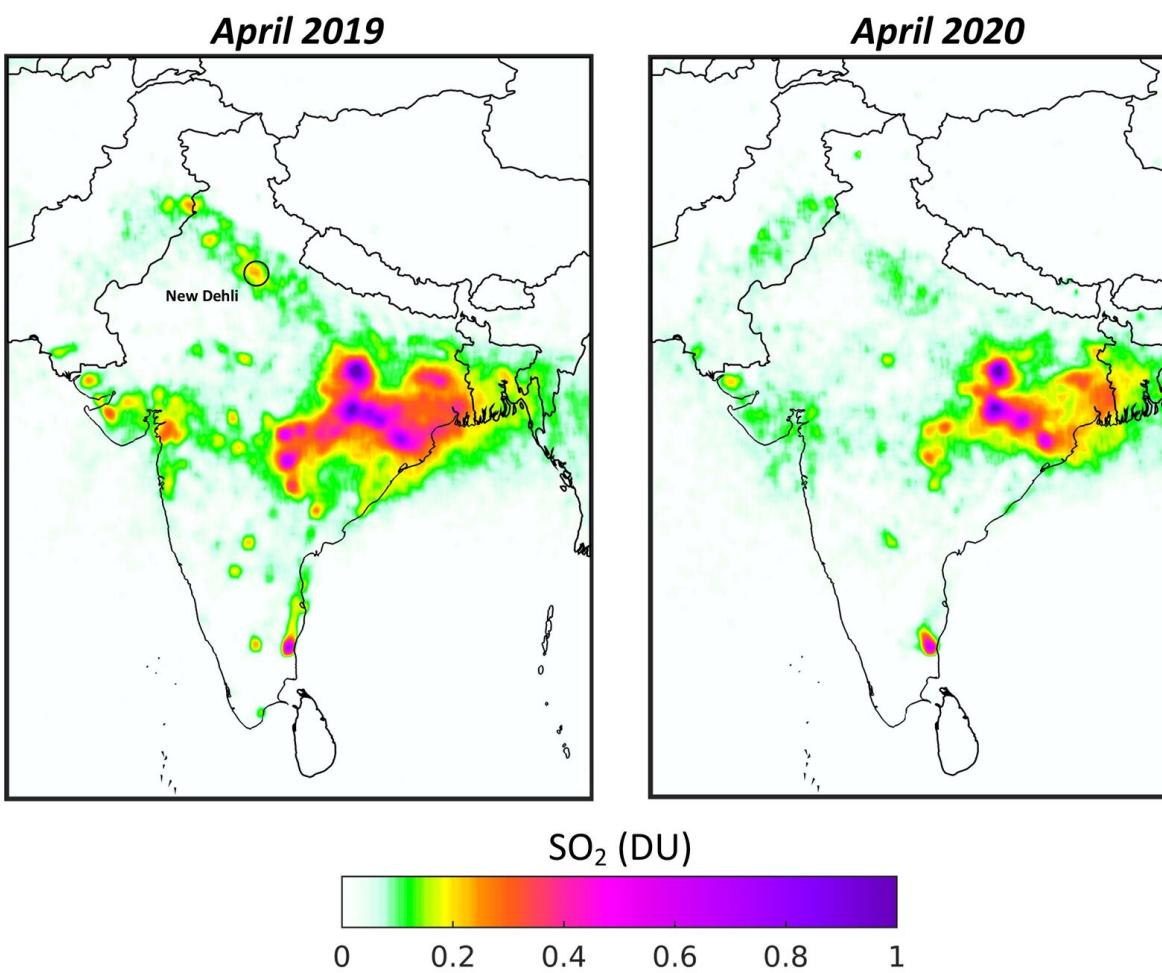
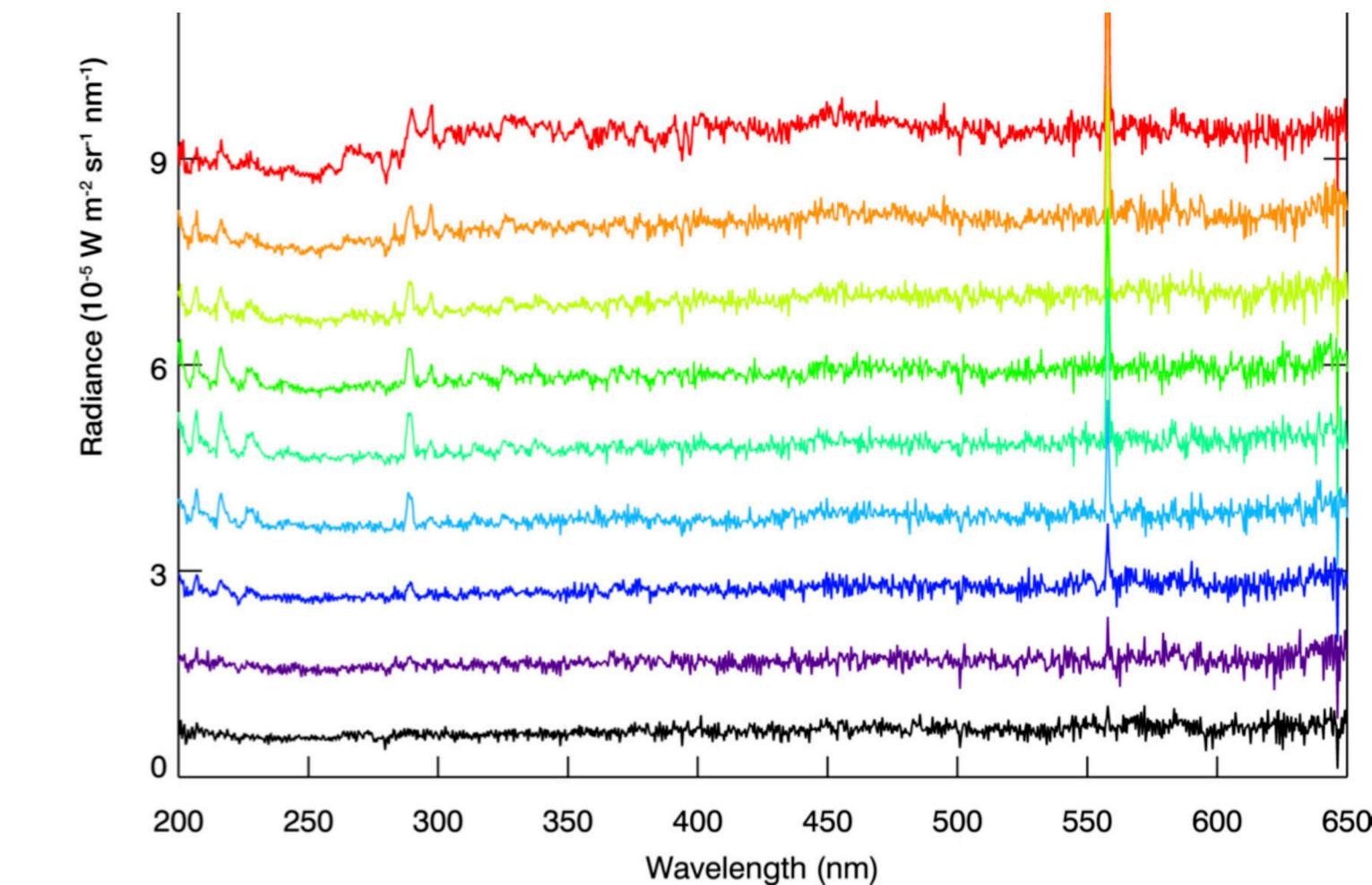
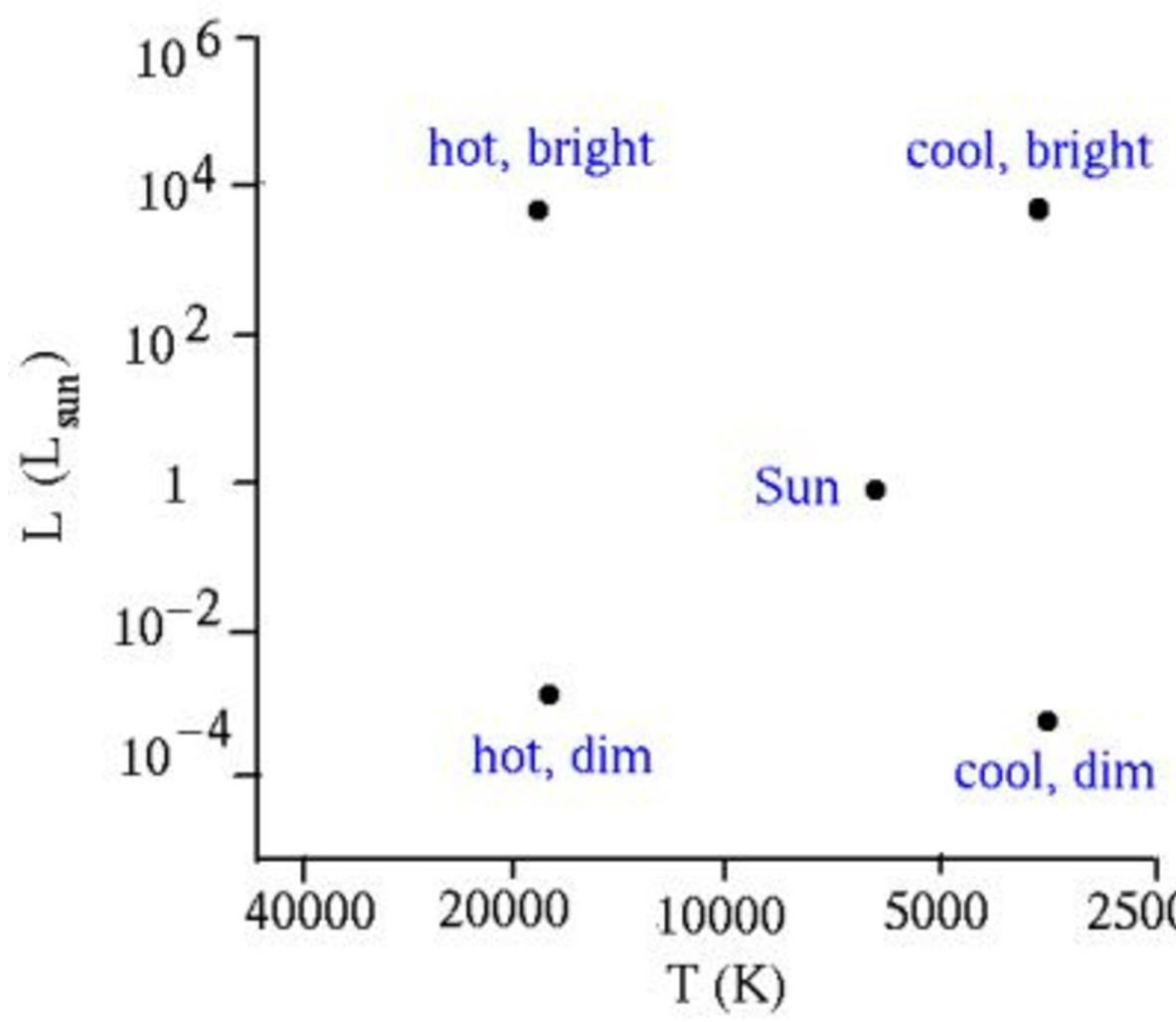
Quelle: Unsplash

NOMINAL VS. ORDINAL VS. INTERVAL/RATIO SCALE



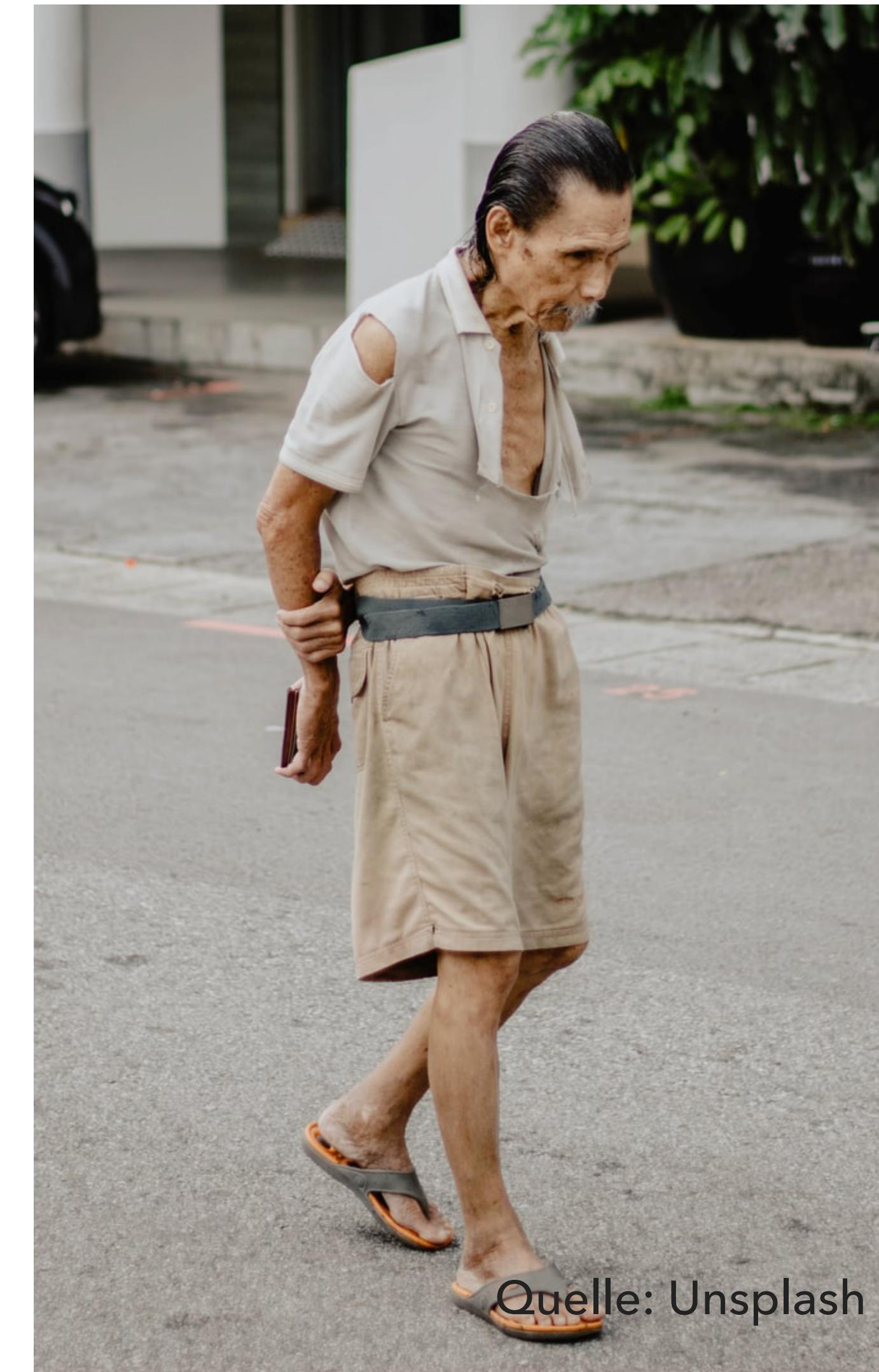
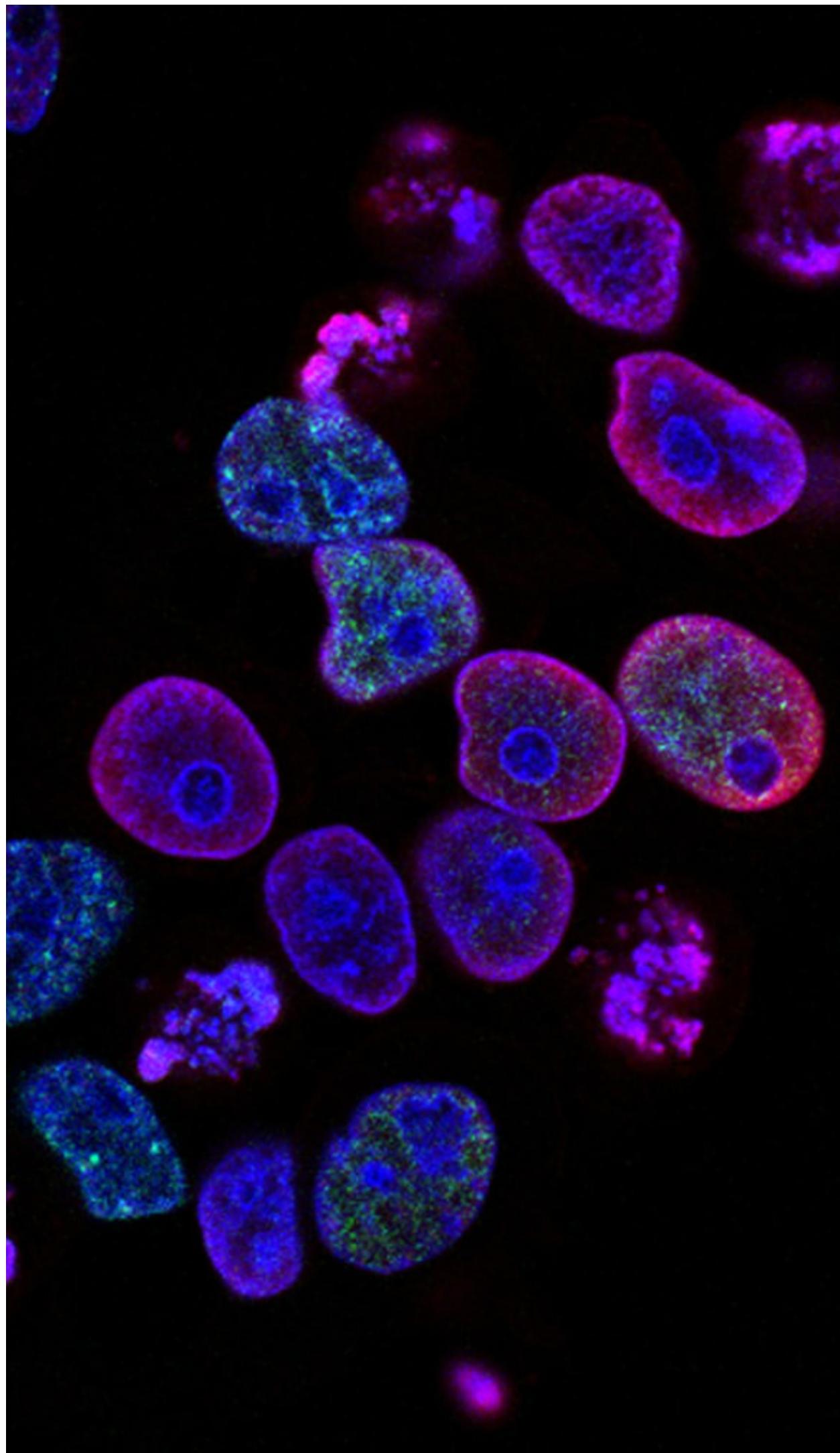
Quelle: Unsplash

EINZELWERTE / ZEITREIHEN / KOORDINATEN / ABLÄUFE



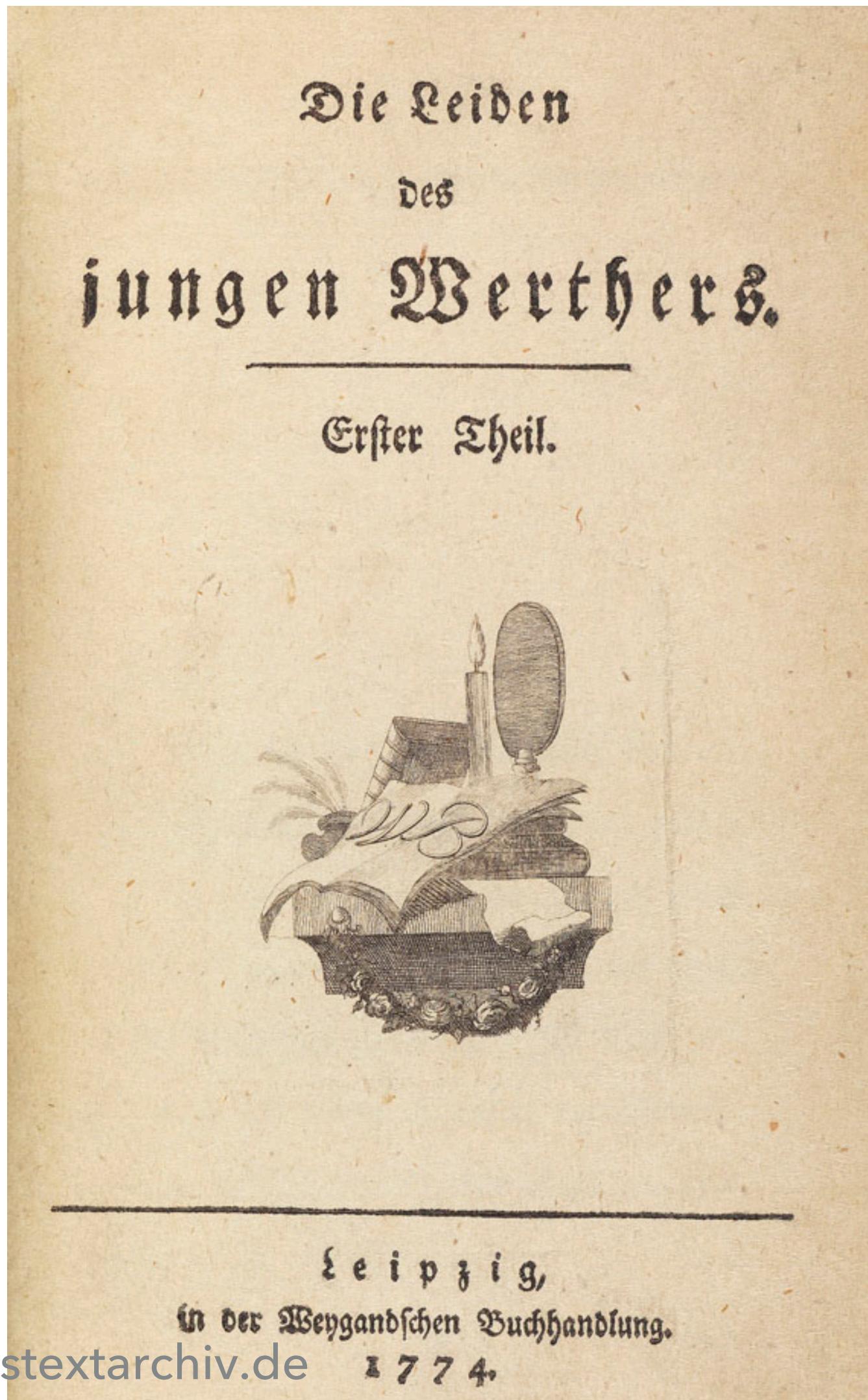
Quelle: ESA

BILDER / VIDEO



Quelle: Unsplash

KULTURDATEN



Quelle:
deutschestextarchiv.de

	DEUTSCHE NATIONAL BIBLIOTHEK
	Link zu diesem Datensatz
Titel	
Person(en)	
Verlag	
Zeitliche Einordnung	
Umfang/Format	
ISBN/Einband/Preis	
Sprache(n)	
Beziehungen	
Sachgruppe(n)	
Literarische Gattung	

<http://d-nb.info/gnd/4099202-0>

Goethe, Johann Wolfgang von

Die Leiden des jungen Werthers

Die Leiden des jungen Werther

Werther

Ālām aš-šab Firtir (Arabisch)

Zenc Verterin Iztirablary (Aserbeidschanisch)

Werther gaztearen arrangurak (Baskisch)

Shao-nien Wei-t'e ti fan-nao (Chinesisch)

Den unge Werthers lidelser (Dänisch)

The sorrows of young Werther (Englisch)

La suferoj de la junia Werther (Esperanto)

Les passions du jeune Werther (Französisch)

Les souffrances du jeune Werther (Französisch)

As mágoas do mozo Werther (Galizisch)

I dolori del giovane Werther (Italienisch)

Les tribulacions del jove Werther (Katalanisch)

Jeolm eun be leu te leu ui seul peum (Koreanisch)

Jauna Wertera ciešanas (Lettisch)

Het Lyden van den jongen Werther (Niederländisch)

Het lijden van de jonge Werther (Niederländisch)

Rang hā-i Wirtir-i g̊awān (Persisch)

Cierplenia młodego Wertera (Polnisch)

A paixao do jovem werther (Portugiesisch)

Os sofrimentos do jovem Werther (Portugiesisch)

Los sufrimientos del joven Werther (Spanisch)

Las penas del joven Werther (Spanisch)

Utrpenie mladého Werthera (Tschechisch)

Az ifju Werther gyötrelmei (Ungarisch)

Naujavān Varthar kī dāstān-i ġam (Urdu)

Kindler (3. Aufl., online)

Goethes Werke im WWW

DATEN ...

- ▶ ... sind extrem vielfältig.
- ▶ ... existieren in einem Kontext.
- ▶ ... haben vielseitige Nutzungsanforderungen.

**WIE
VIESSUALISIEREN?**

HILFREICHE TOOLS UND PROGRAMME

- ▶ Datenhaltung
Excel / SPSS / MATLAB
- ▶ Programmiersprachen
Python / R / Javascript / MATLAB
- ▶ Vektorgraphik Programme
InkScape (freeware) / Affinity Designer / Adobe Illustrator / PowerPoint
- ▶ Explorative Datenanalyse und Inforgraphiken
Tableau / Infogram / Datawrapper / Power BI / ...

DATENHALTUNG

Microsoft Excel:

- Teil von Microsoft Office, Tabellen basiert, kann simple Datenanalyse und erstellen von Graphiken

Libre Office Calc (<https://www.libreoffice.org/>):

- Open Source Alternative für Microsoft Excel

SPSS:

- Statistik Software , Tabellen basiert, Goldstandard für statistische Analysen und erstellen von Graphiken

PROGRAMMIERSPRACHEN

MATLAB (<https://www.mathworks.com>):

- Kommerzielle, mächtige Software und Programmiersprache zur Datenanalyse, Visualisierung, Programmierung

R (<https://www.r-project.org/>):

- Freie Programmiersprache zur Datenanalyse und Visualisierung, weit verbreitet im mathematischen Bereich

Python (<https://www.python.org/>):

- Freie Programmiersprache u.a. zur Datenanalyse und Visualisierung, weit verbreite und sehr vielseitig

Javascript:

- Programmiersprache, ursprünglich für das Web, einige Frameworks zur (interaktiven) Visualisierung

VEKTORGRAPHIK PROGRAMME

Adobe Illustrator (<https://www.adobe.com/>):

- Industriestandard für das Erzeugen von Vektorgraphiken

Affinity Designer (<https://affinity.serif.com/>) :

- Kostengünstigere Alternative zu Adobe Illustrator

InkScape (<https://inkscape.org/>):

- Kostenlose Alternative zu Adobe Illustrator, u.U. höhere Lernkurve

Microsoft PowerPoint:

- Unterstützt einiges an Vektorgraphikoperationen

EXPLORATIVE DATENANALYSE UND INFOGRAPHIKEN

Tableau (<https://public.tableau.com>):

- ▶ Verschiedene Tools/Produkte zur explorativen Datenanalyse und Erstellung von Infographiken

Infogram (<https://infogram.com/>):

- ▶ Online Tool zum erstellen von Graphiken, simpler als Tableau

Weitere Alternativen:

- ▶ Datawrapper (<https://www.datawrapper.de/>)

- ▶ Microsoft Power BI (<https://powerbi.microsoft.com/>)

**WIE
GEHTS WEITER?**

TERMINE

- ▶ 26.10. - Auftakt(heute)
- ▶ 02.11. - Grundlagen von Visualisierungen
- ▶ 09.11. - Praktische Anwendung und Best Practises
- ▶ 16.11. - "Künstlerische Methoden"
- ▶ 23.11. - Vorstellung und Besprechung der Ergebnisse

Zwischenzeitlich: Arbeiten am eigenen Datensatz;
Gern Fragen!

EIGENES PROJEKT

- Wer hat schon Daten mit denen Sie arbeiten wollen?
- Wer braucht Daten um an dem Beispiel zu arbeiten?

AUFGABEN

- Beschreibung der eigenen Daten:
Formate? Skalen?
- Was ist deine Fragestellung?
- Was ist für dich eine gute Graphik?
(such dir Beispiele)

VIELEN DANK

CONTACT@LETTY.IO
KAREN.OTTE@CHARITE.DE