Datenvisualisierung in der Wissenschaft

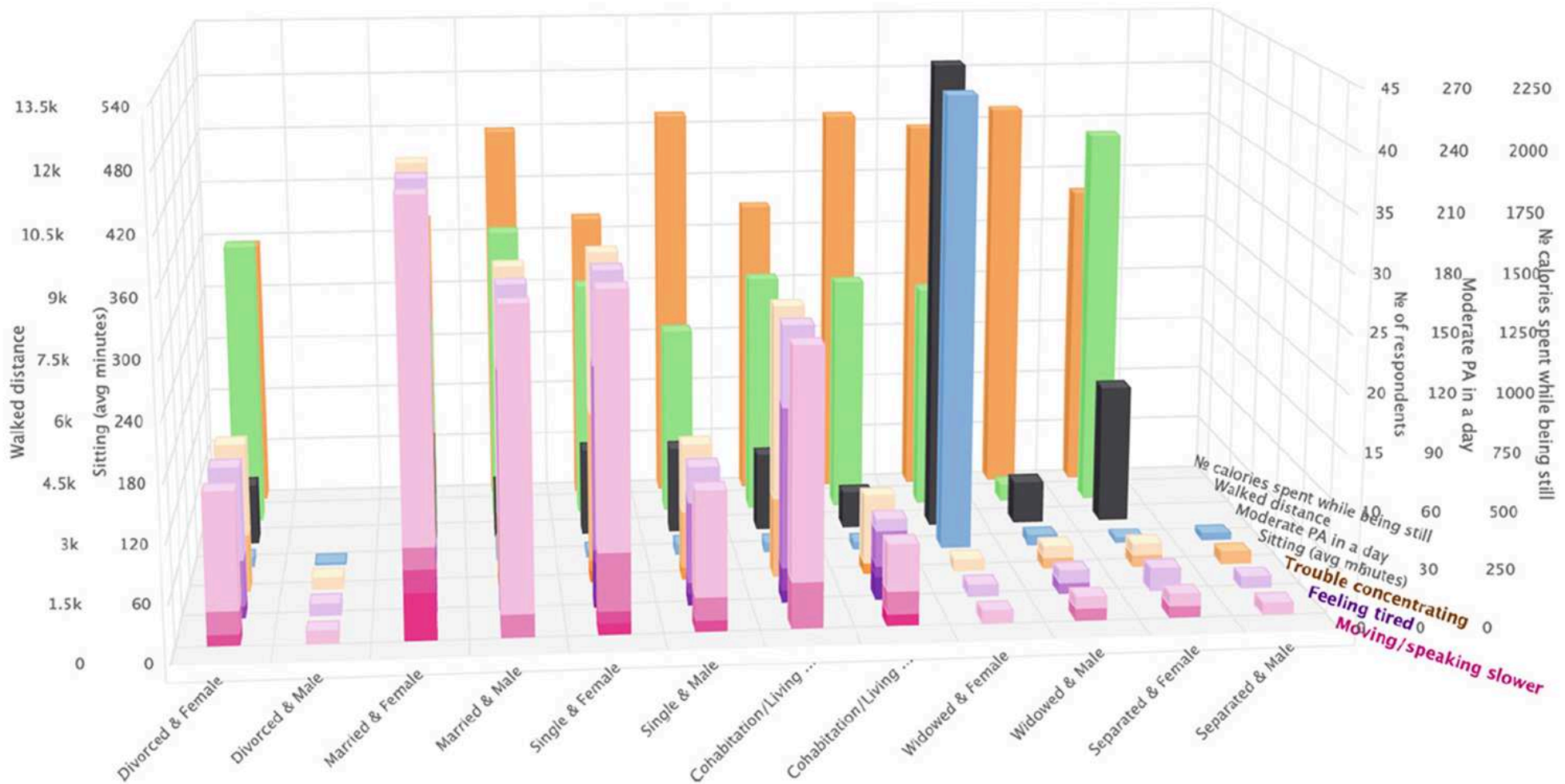
Zweckmäßigkeit der Darstellung (Teil 1)

Dr. Cédric Scherer

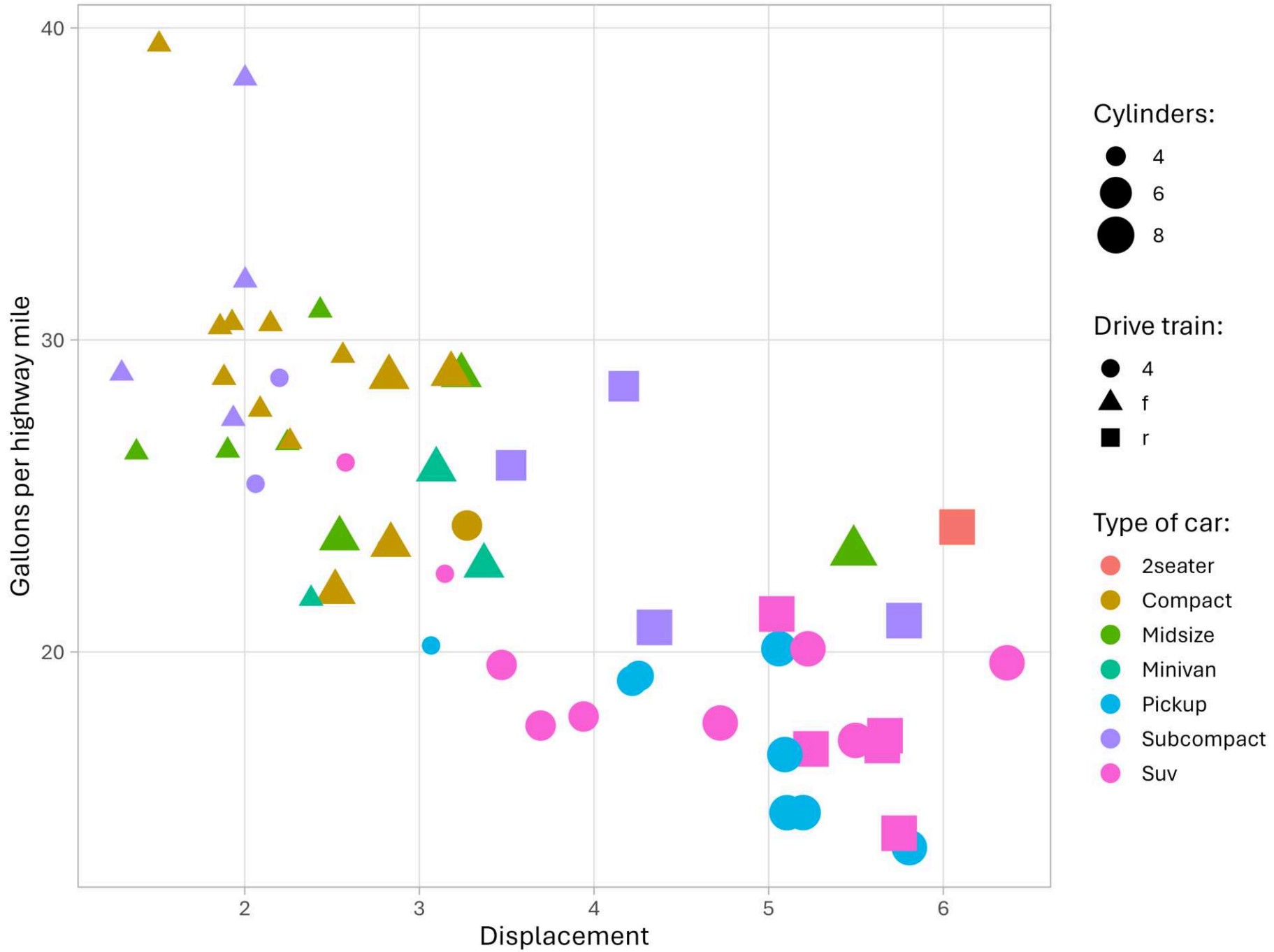
RPTU Nachwuchsring
3., 10. und 17. Juli 2025

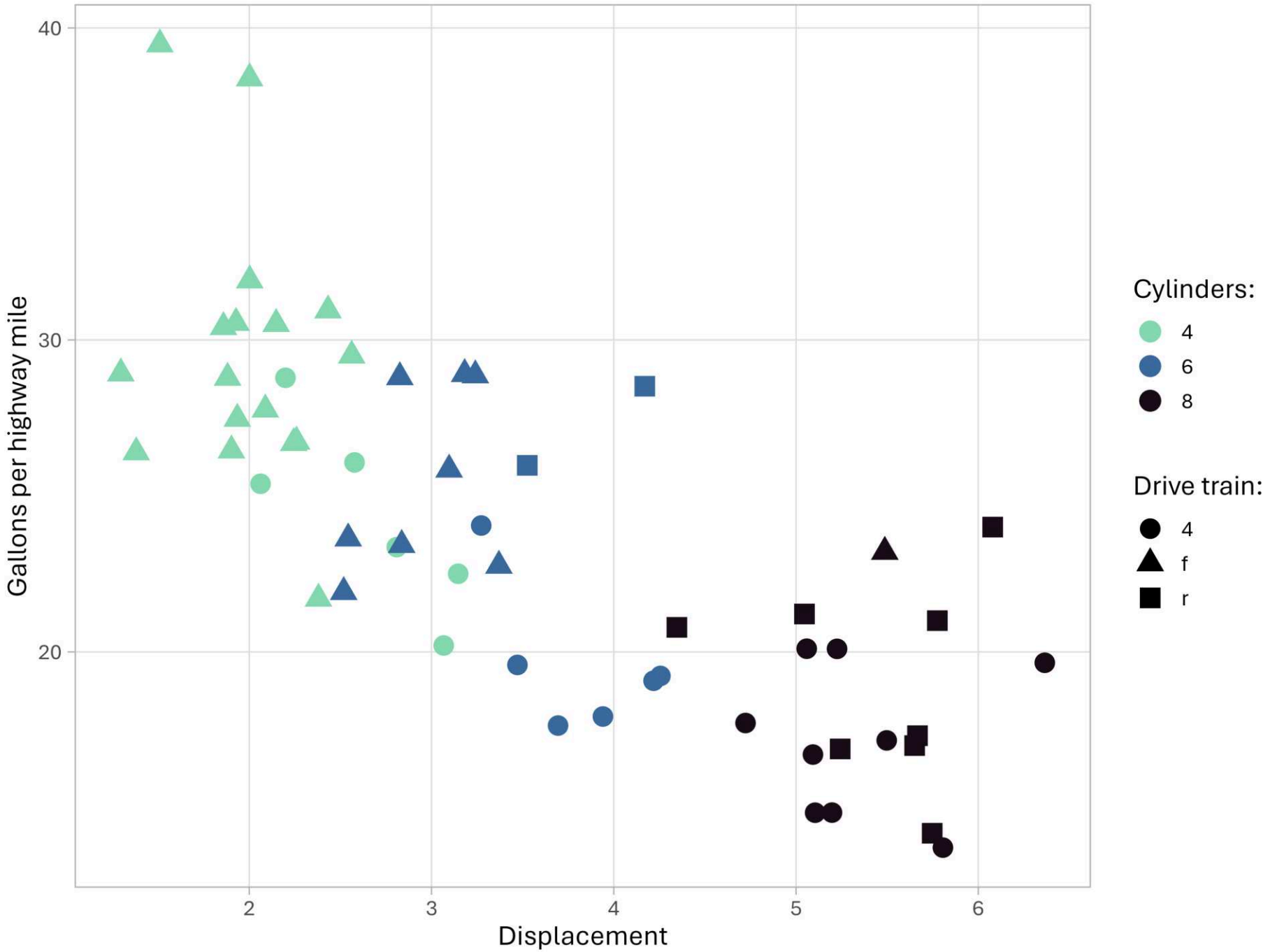
Ziel

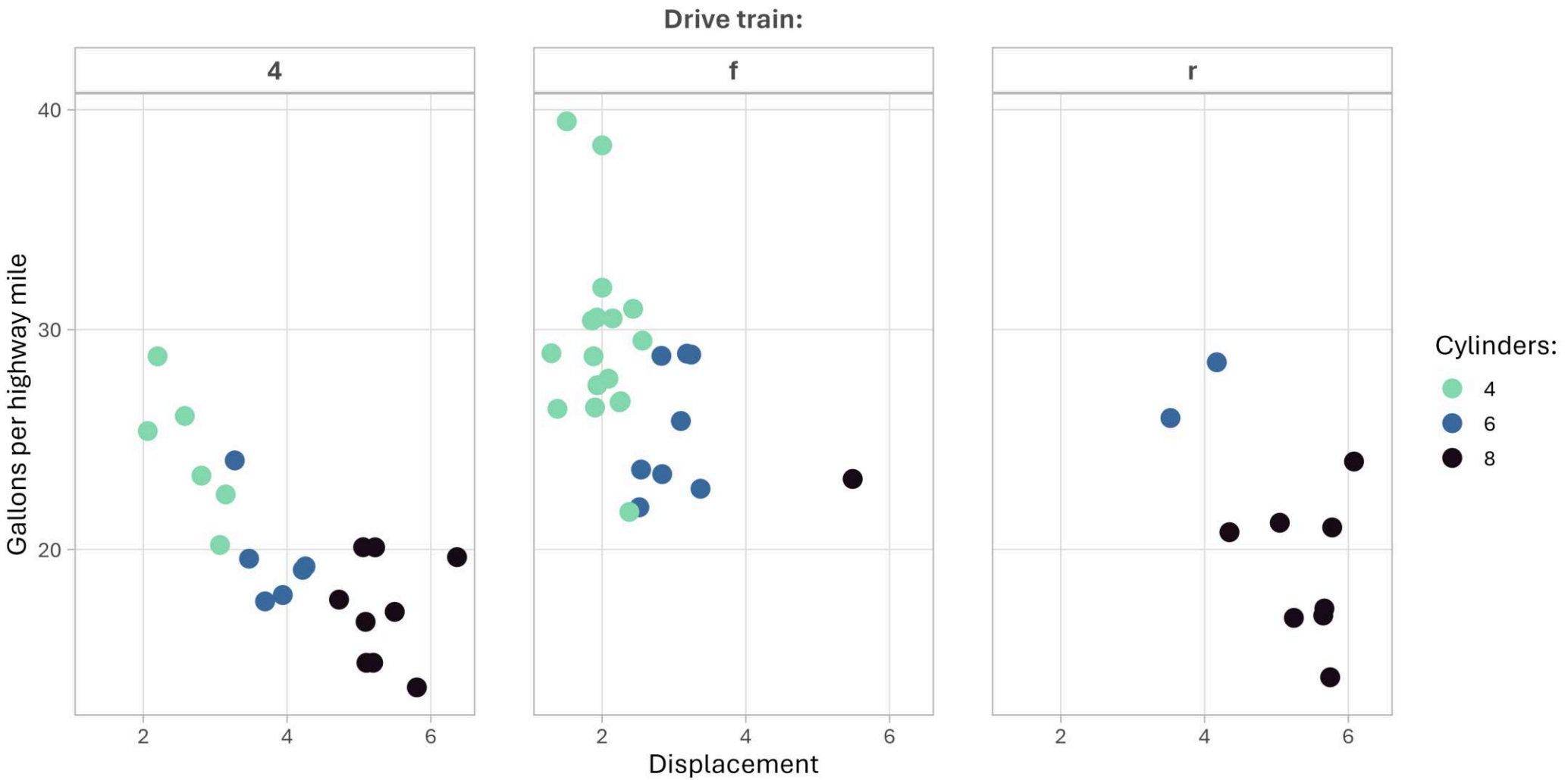
Zweckmäßigkeit der Darstellung

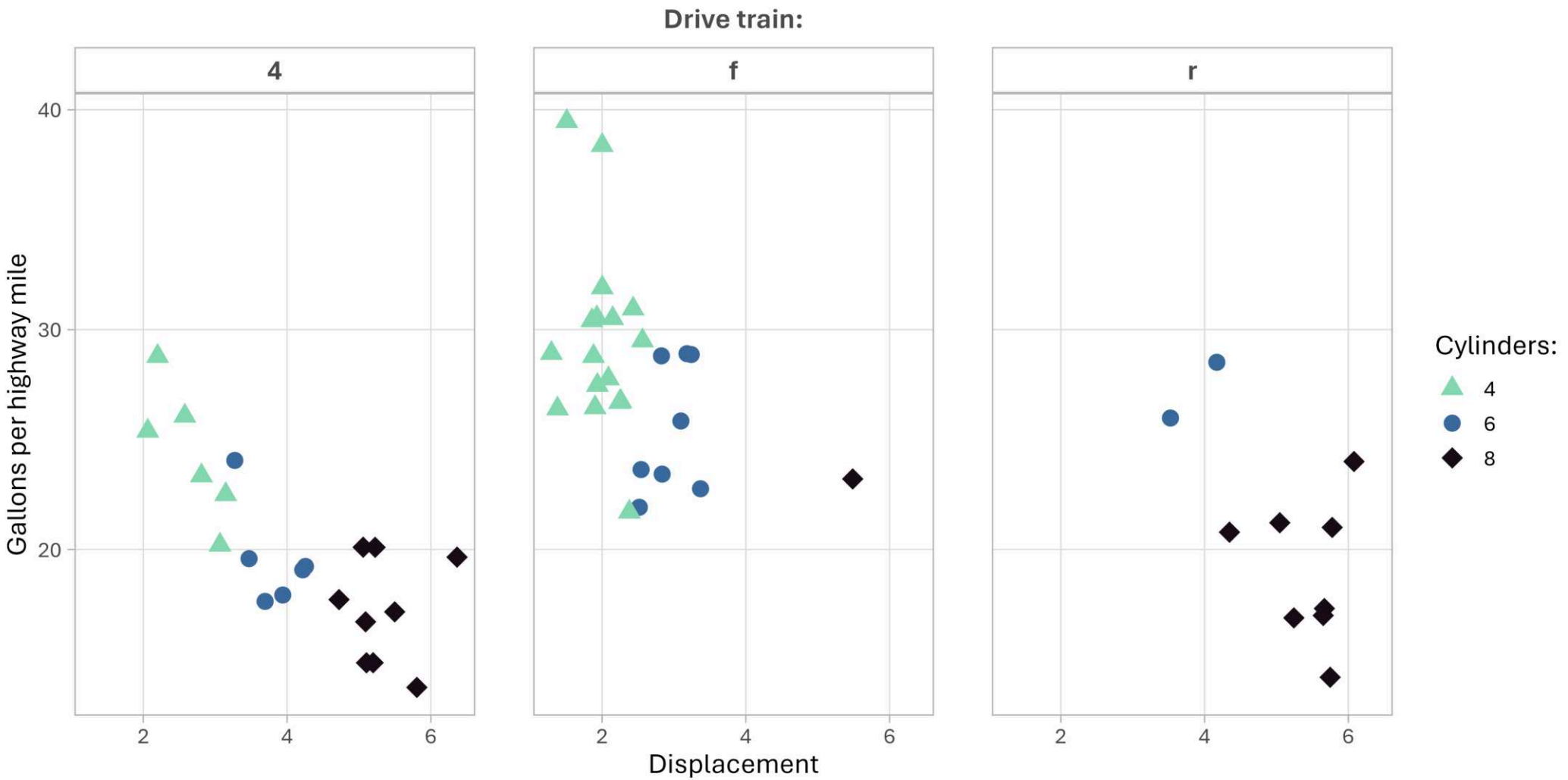


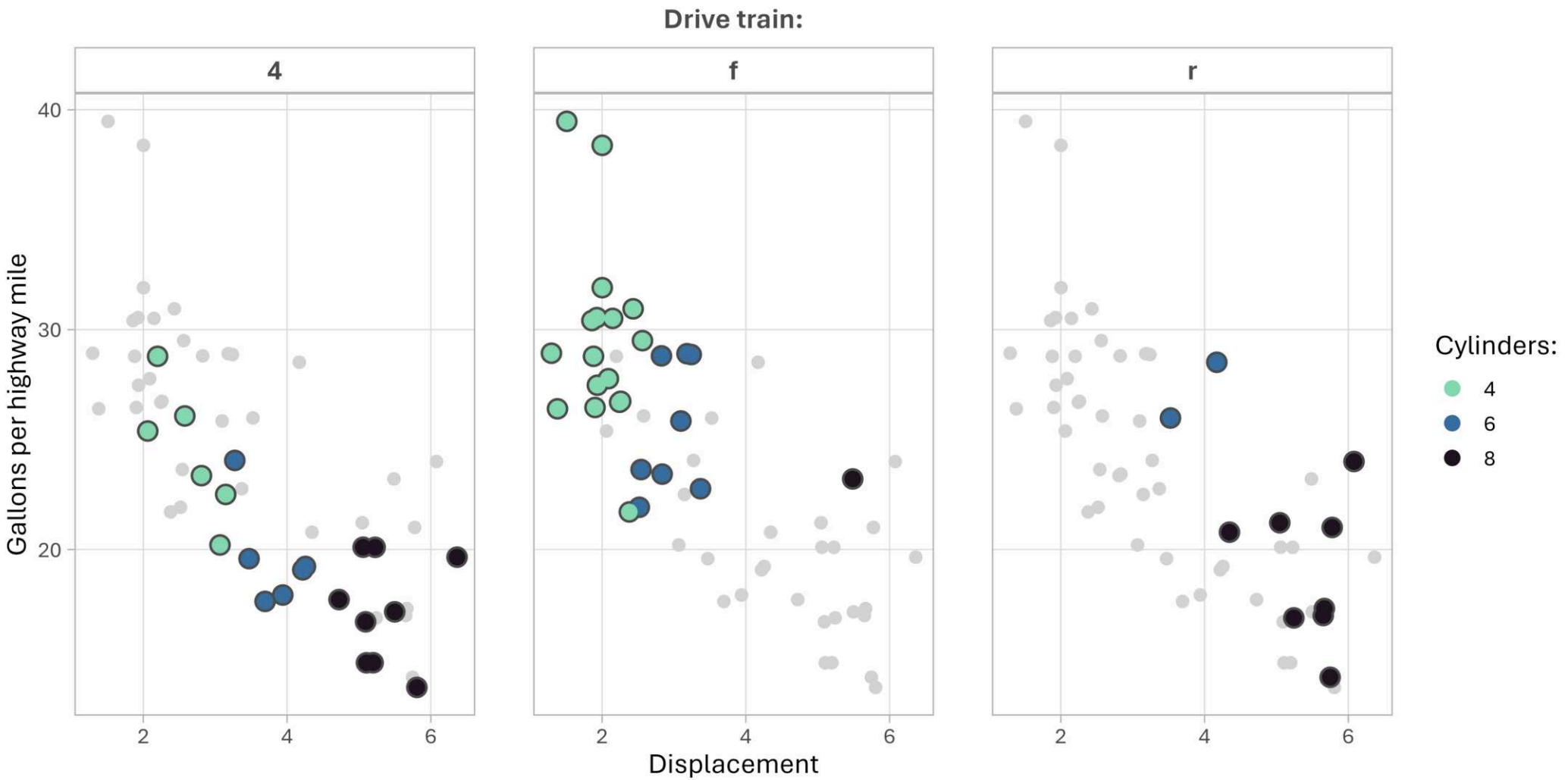
Quelle: Gonzalez-Martinez et al. (2022)

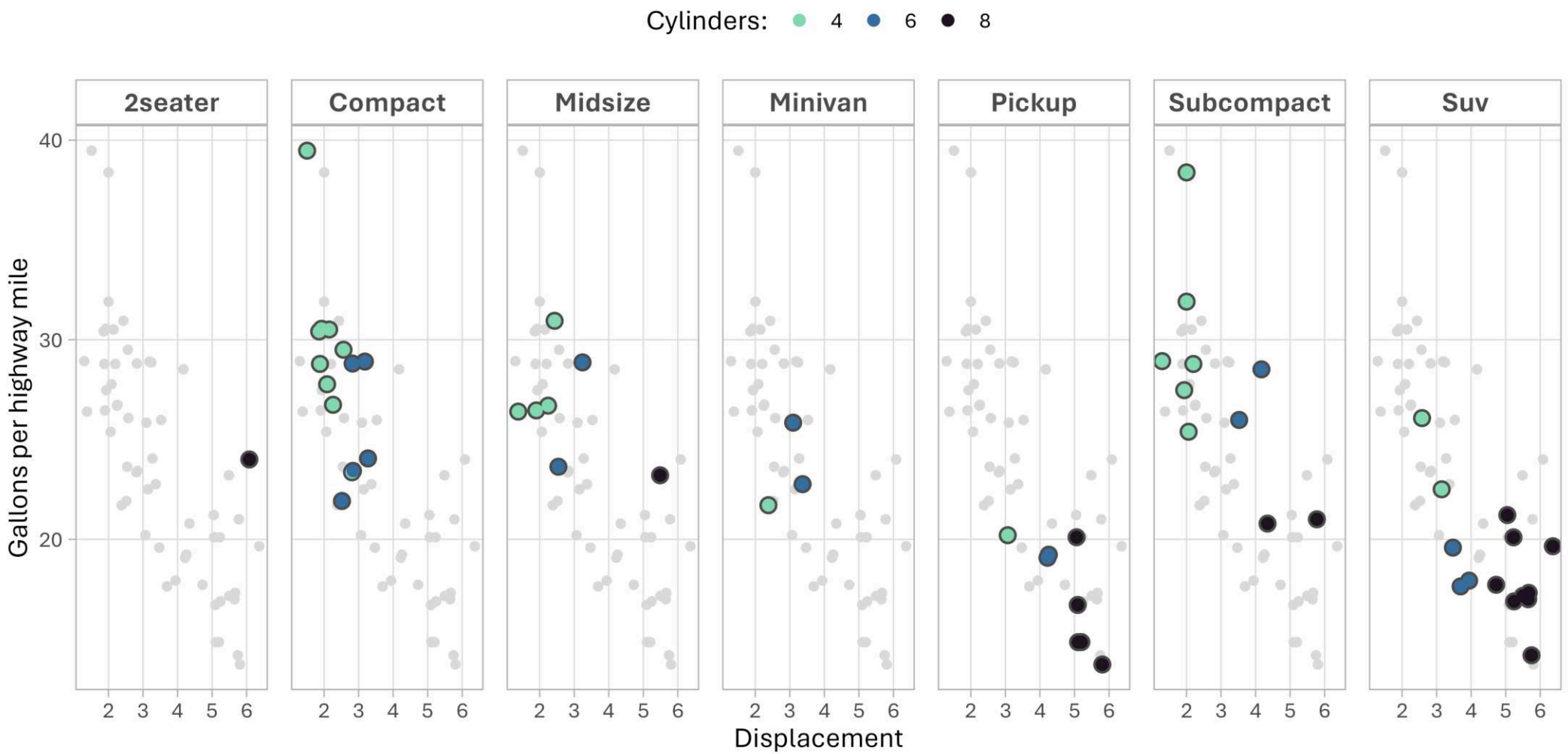

40
30
20
Gallons per highway mile
2
3
4
5
6
Displacement
Cylinders:
4
6
8
Drive train:
4
f
r
Type of car:
2seater
Compact
Midsize
Minivan
Pickup
Subcompact
Suv

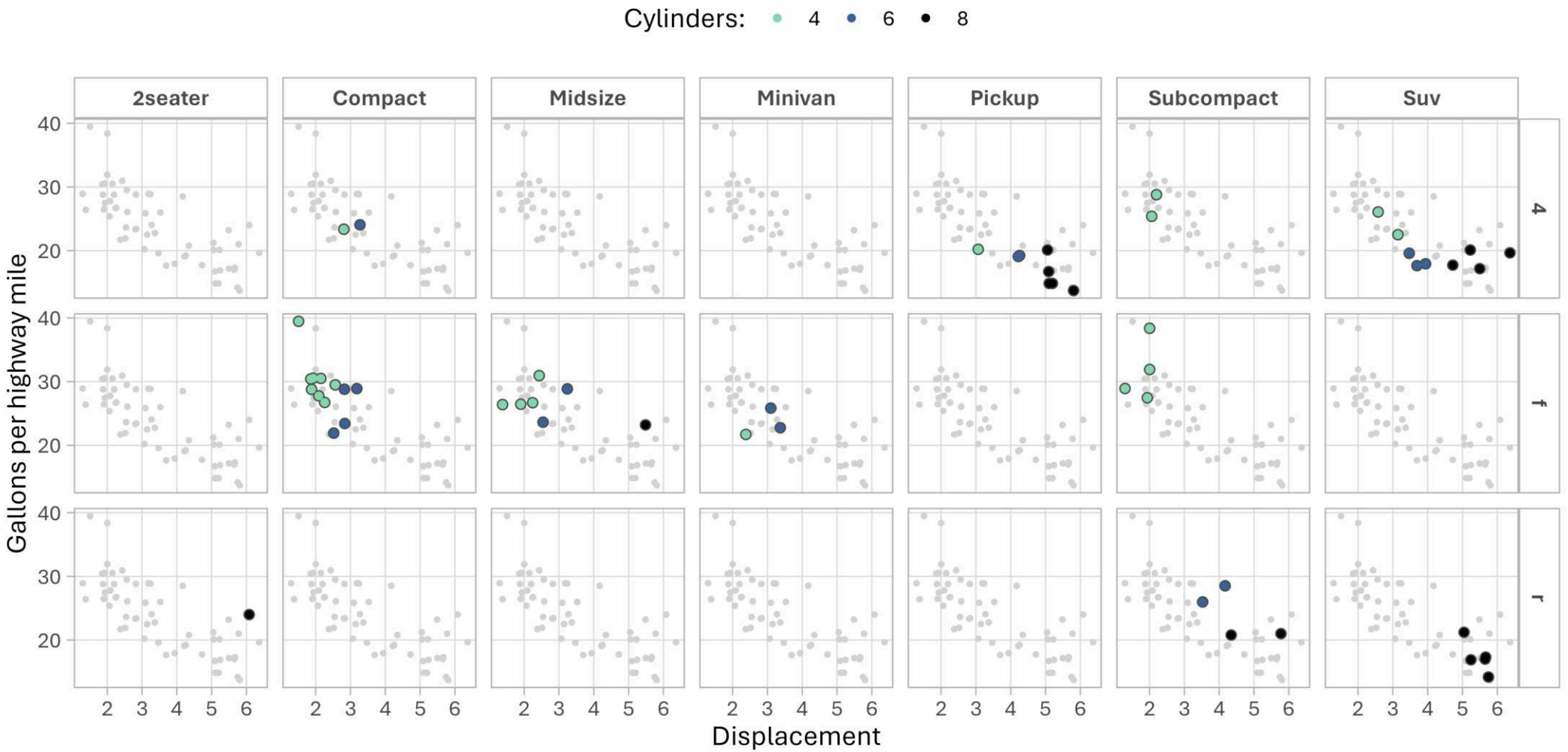

40
30
20
Gallons per highway mile
2
3
4
5
6
Displacement
Cylinders:
4
6
8
Drive train:
4
f
r


Drive train:
4
f
r
40
30
20
Gallons per highway mile
2
4
6
Displacement
Cylinders:
4
6
8


Drive train:
4
f
r
40
30
20
Gallons per highway mile
2
4
6
Displacement
Cylinders:
4
6
8


Drive train:
4
f
r
Gallons per highway mile
40
30
20
2
4
6
Displacement
Cylinders:
4
6
8

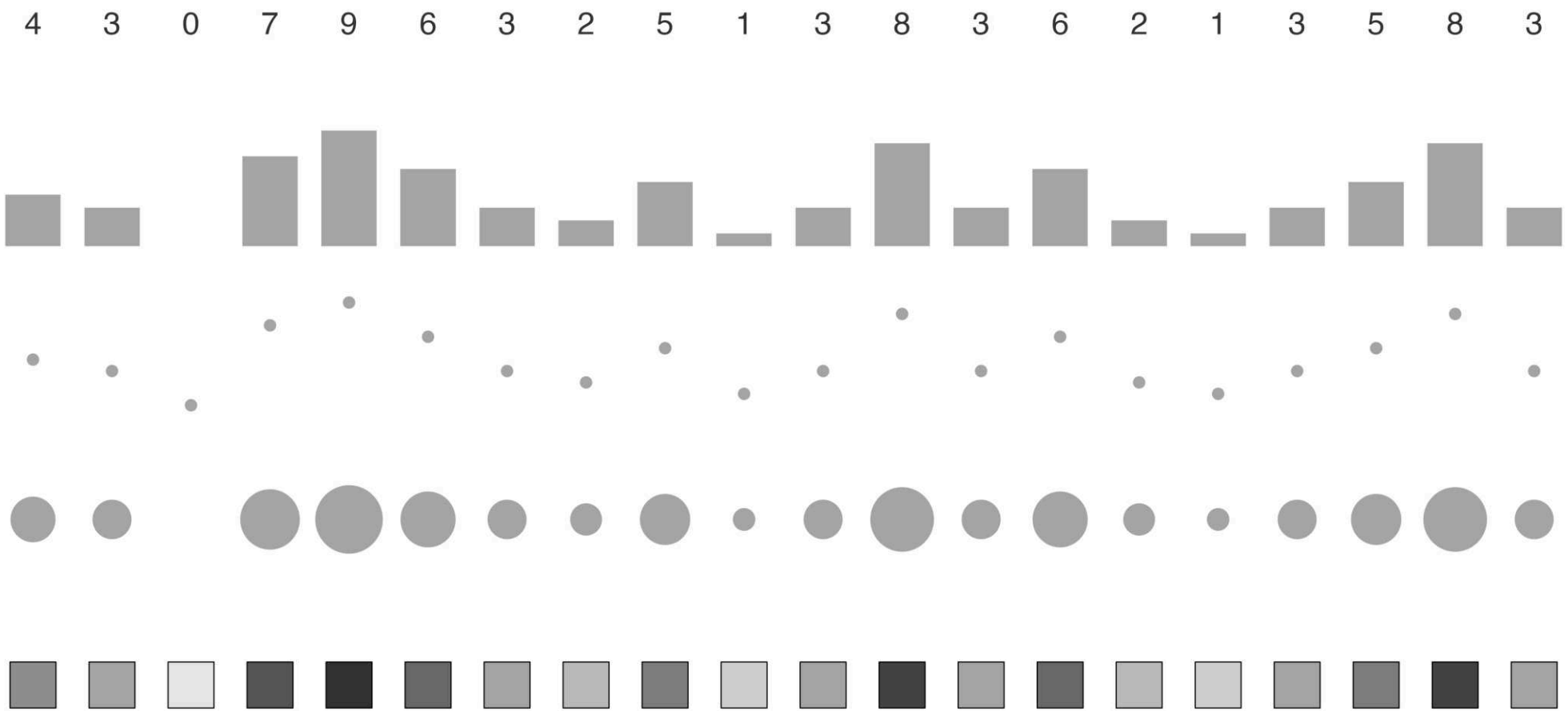

Cylinders: 4 6 8
2seater
Compact
Midsize
Minivan
Pickup
Subcompact
Suv
Gallons per highway mile
40
30
20
2 3 4 5 6
Displacement


Cylinders: 4 6 8
2seater
Compact
Midsize
Minivan
Pickup
Subcompact
Suv
4
f
r
Gallons per highway mile
Displacement
40
30
20
2 3 4 5 6

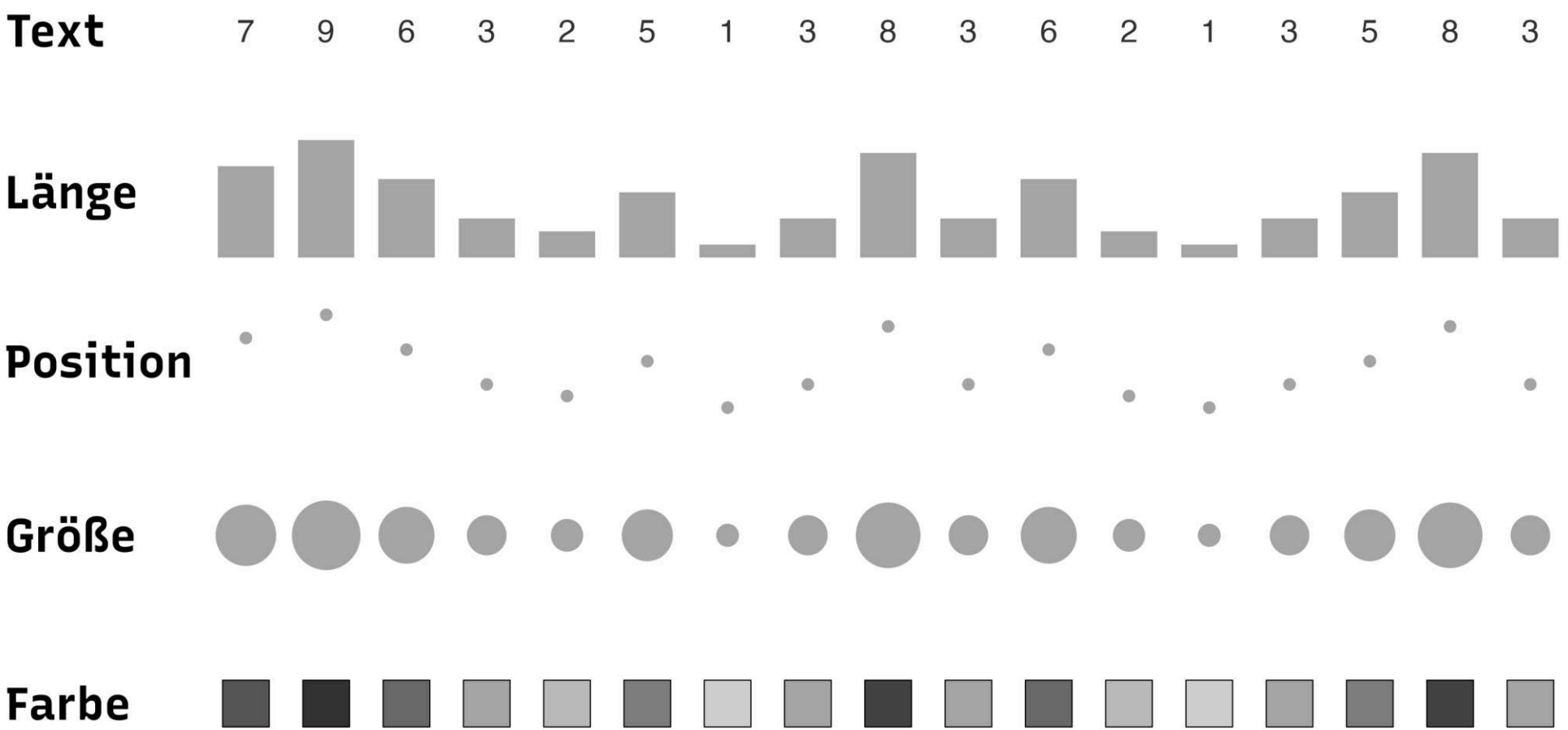
Kodierungen

Zuordnung von Datenwerten zu visuellen Eigenschaften.

Gleiche Werte, andere Kodierung



Gleiche Werte, andere Kodierung



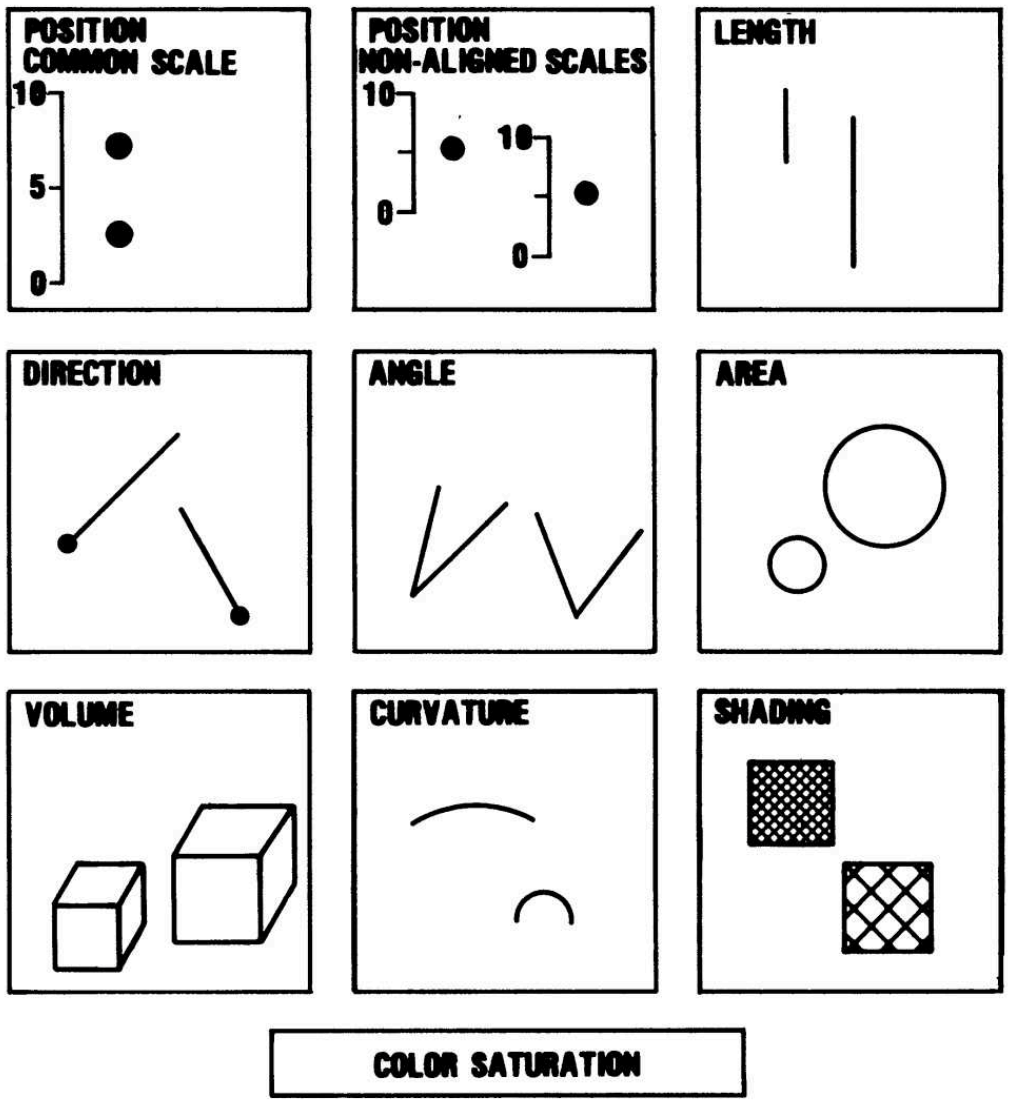
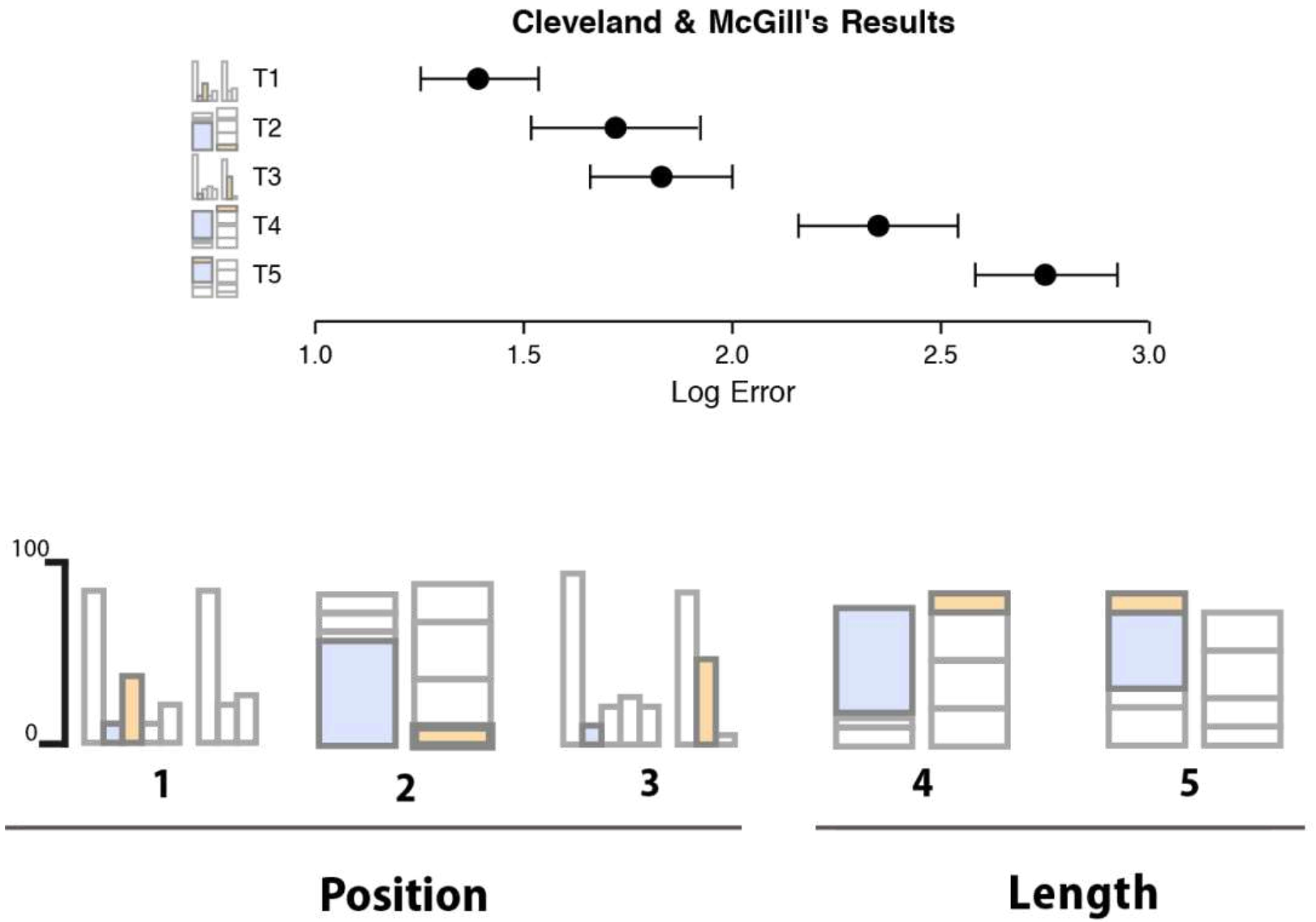
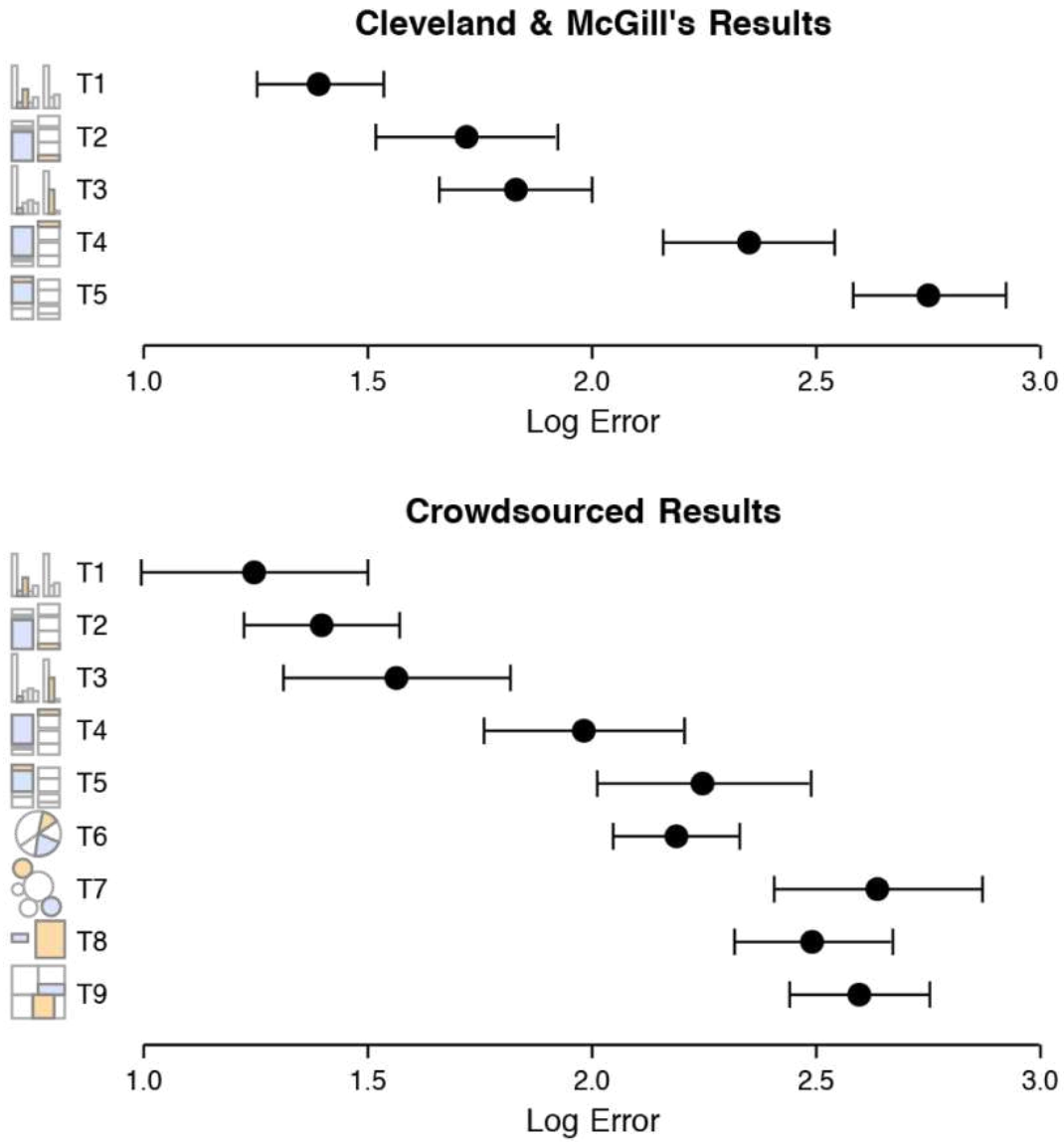


Figure 1. Elementary perceptual tasks.

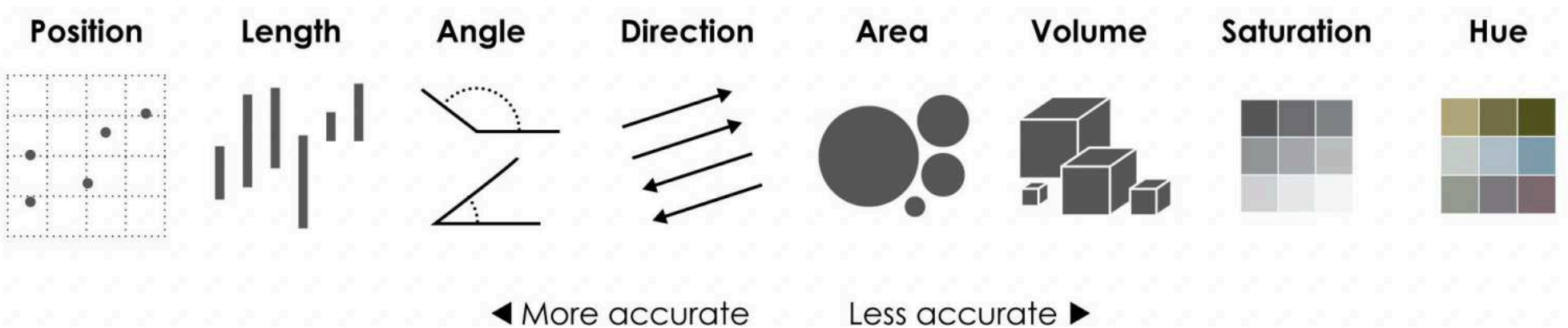
Cleveland & McGill (1984)



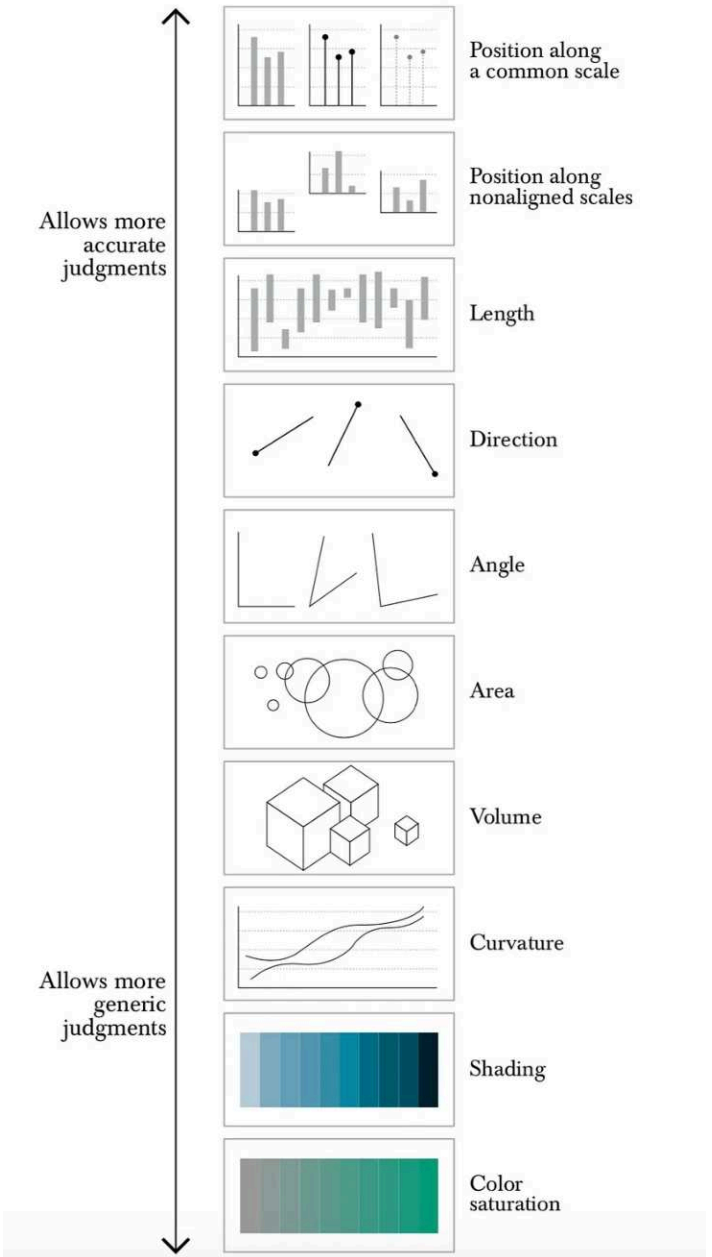
Quelle: “Data Visualization: A Practical Introduction” von Kieran Healy
Ergebnisse basierend auf Cleveland & McGill (1984, 1987) und Heer & Bostock (2010)



Quelle: “Data Visualization: A Practical Introduction” von Kieran Healy
Ergebnisse basierend auf Cleveland & McGill (1984, 1987) und Heer & Bostock (2010)



Quelle: “Data Points” von Nathan Yau (S. 104)



Cleveland's and McGill's Scale of Perpetual Elementary Tasks
Quelle: "The Functional Art" von Alberto Cairo (S. 120)

VISUAL VARIABLES

Organized by how well they are suited for representing data measured on each type of scale

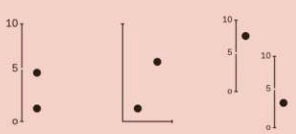
Data visualization handbook, pp. 58–62 datavizhandbook.info

Numbers (data on ratio or interval scale)

USING POSITION FOR ENCODING NUMBERS

MOST PRECISE → → LEAST PRECISE

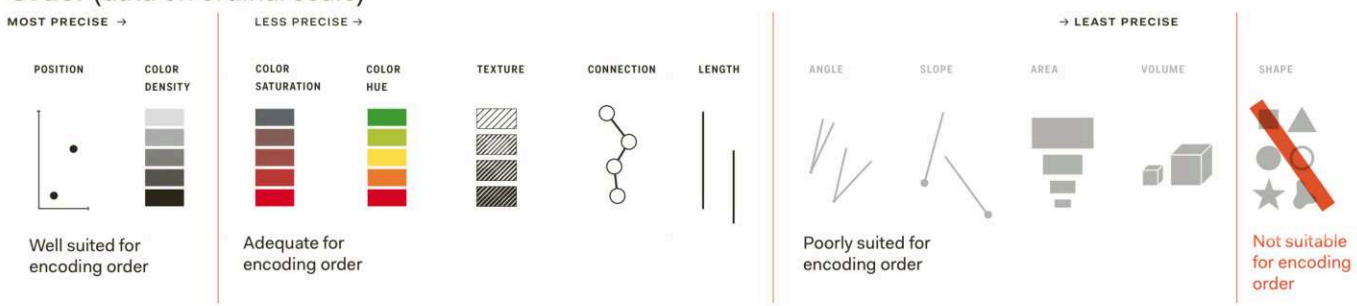
Position along common explicit scale	Position along common implicit scale	Position along non-aligned, but linked scales



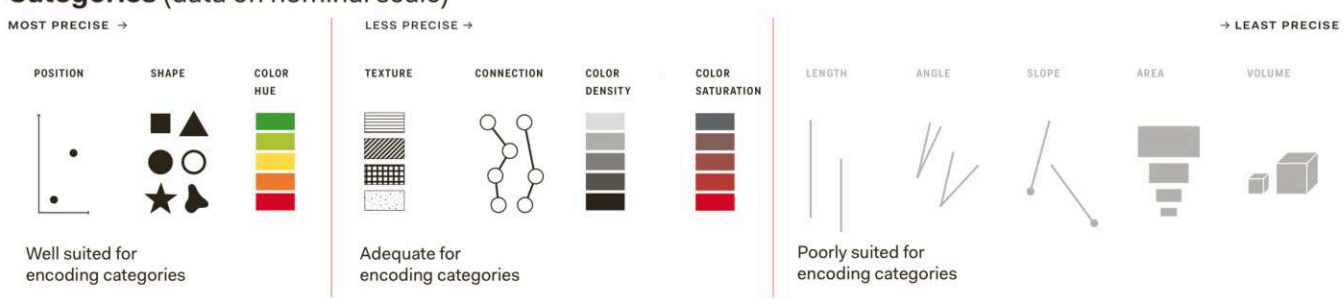
Position is in itself a very precise way of encoding information, but its usefulness in encoding numbers can be further enhanced by adding a **scale**. Data points can be compared even across several charts with relative ease when the charts have **linked scales**, meaning that similar distance in position corresponds to the same difference in value on both. (See *Data visualization handbook*, pp. 83–85.)

Scales are not helpful in encoding order or categories.

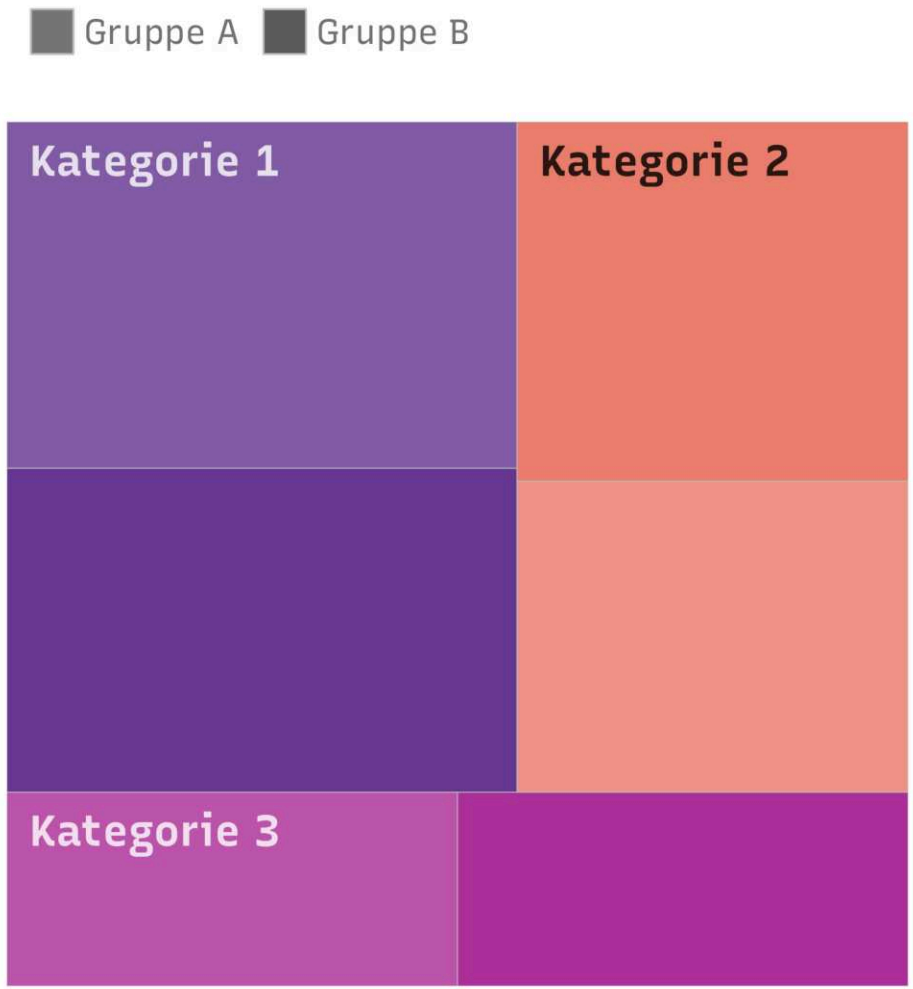
Order (data on ordinal scale)

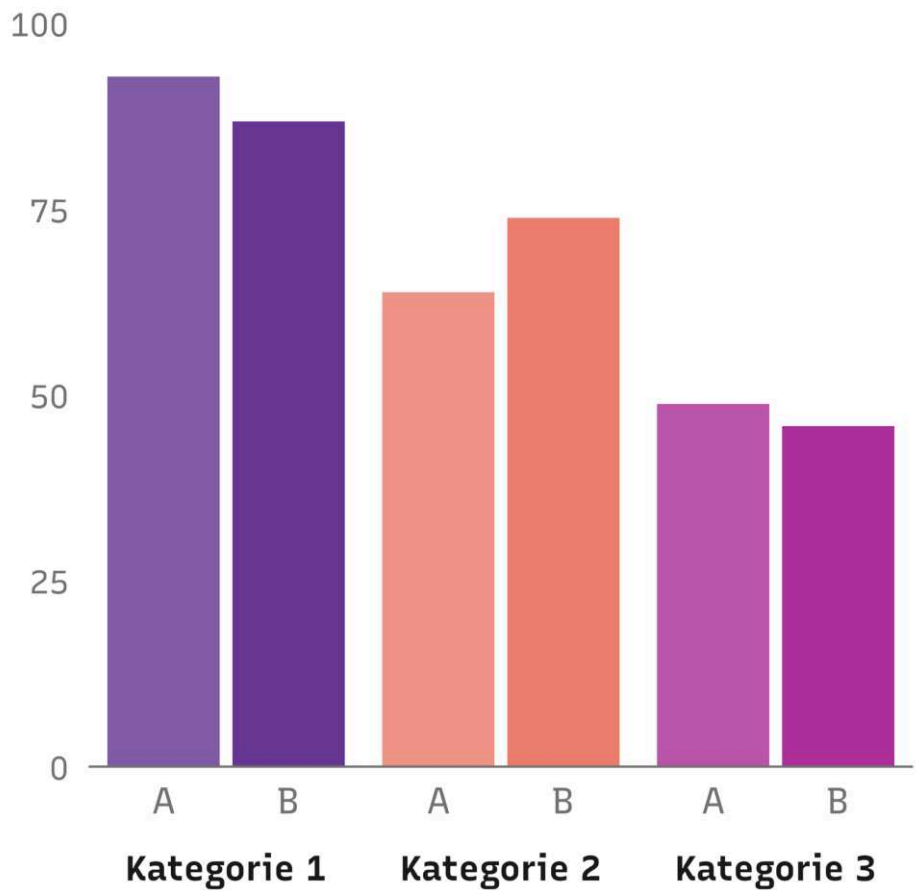


Categories (data on nominal scale)



Quelle: “Data Visualization Handbook” von Juuso Koponen & Jonatan Hildén (2020), Seite 58–62 (verfügbar als Poster

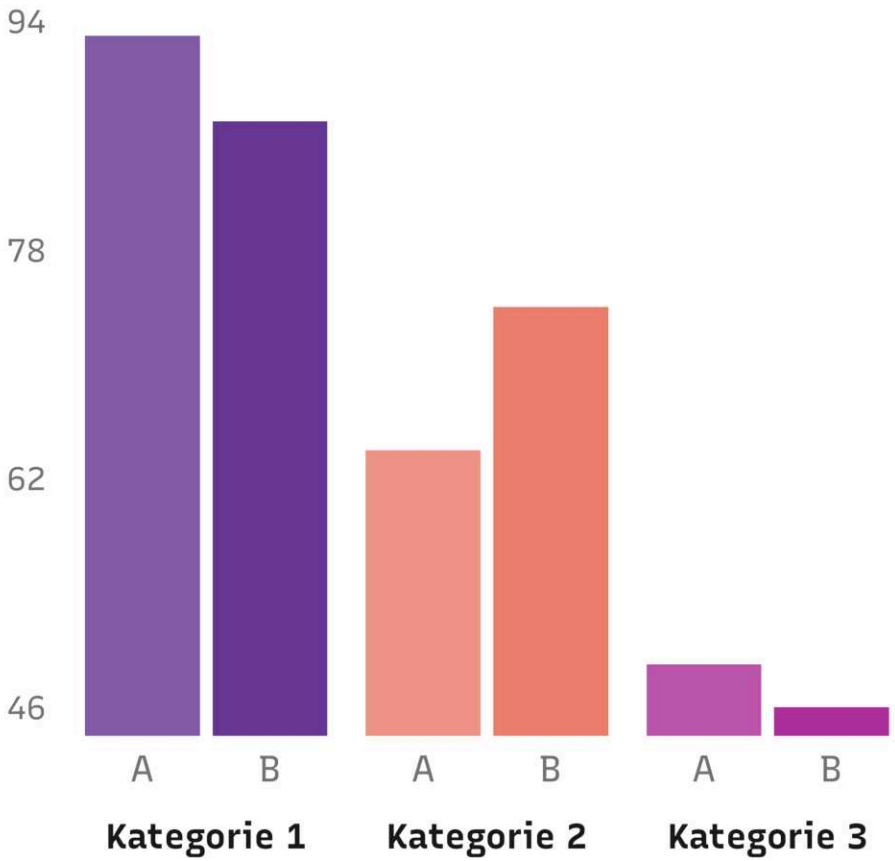

Gruppe A
Gruppe B
Kategorie 1
Kategorie 2
Kategorie 3

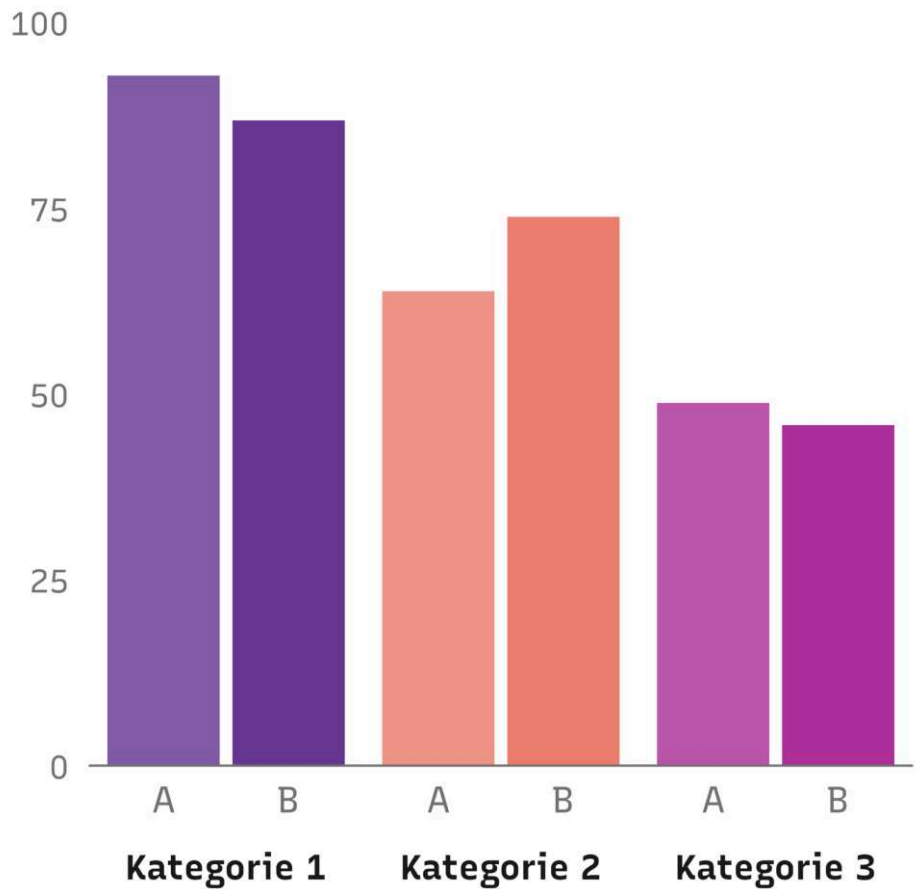

100
75
50
25
0
A
B
A
B
A
B
Kategorie 1
Kategorie 2
Kategorie 3

“There’s a strand of the data viz world that argues that **everything could be a bar chart**.

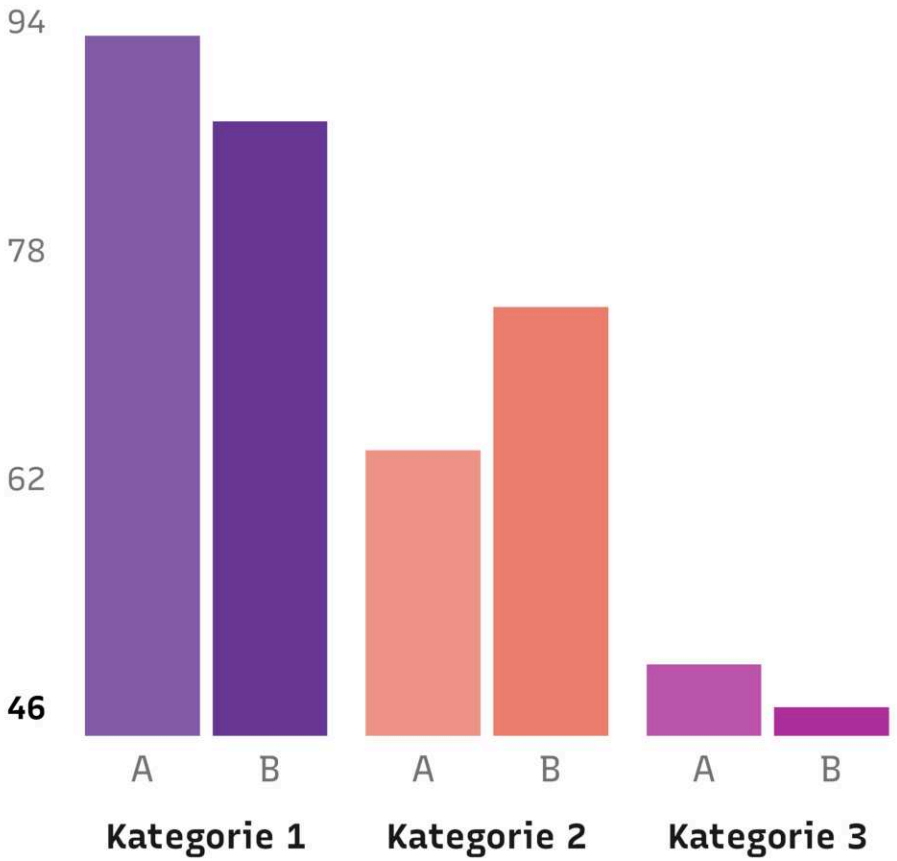
That’s possibly true but also possibly **a world without joy**.”

Amanda Cox (2013)

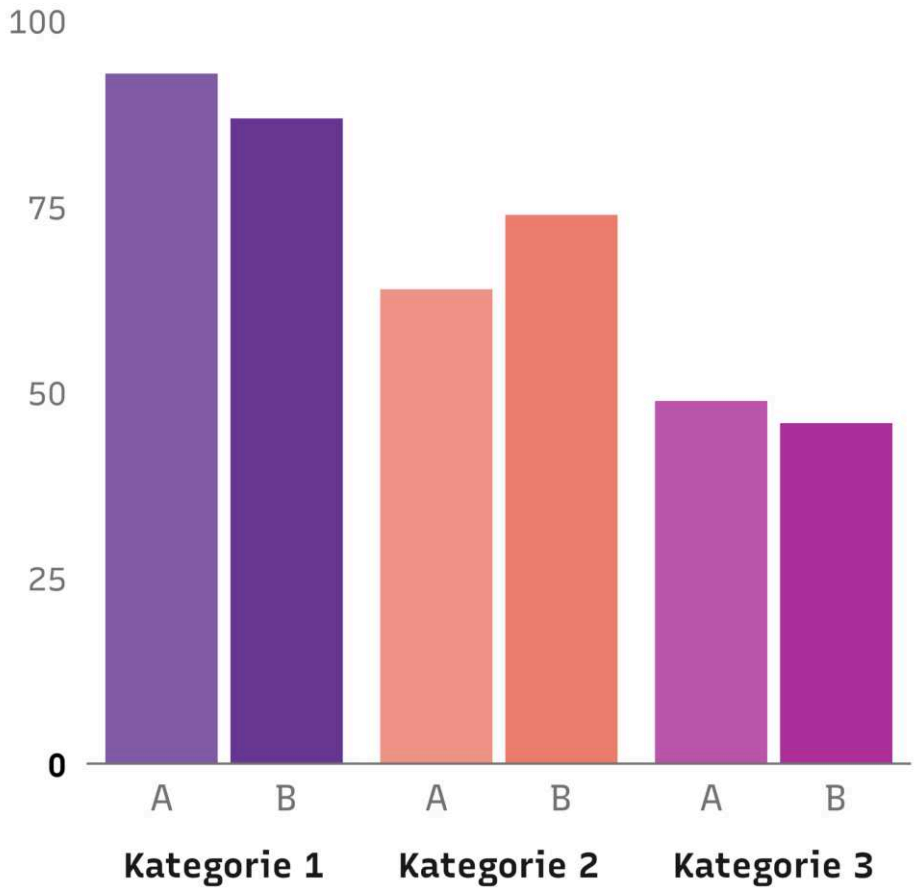

94
78
62
46
A
B
A
B
A
B
Kategorie 1
Kategorie 2
Kategorie 3

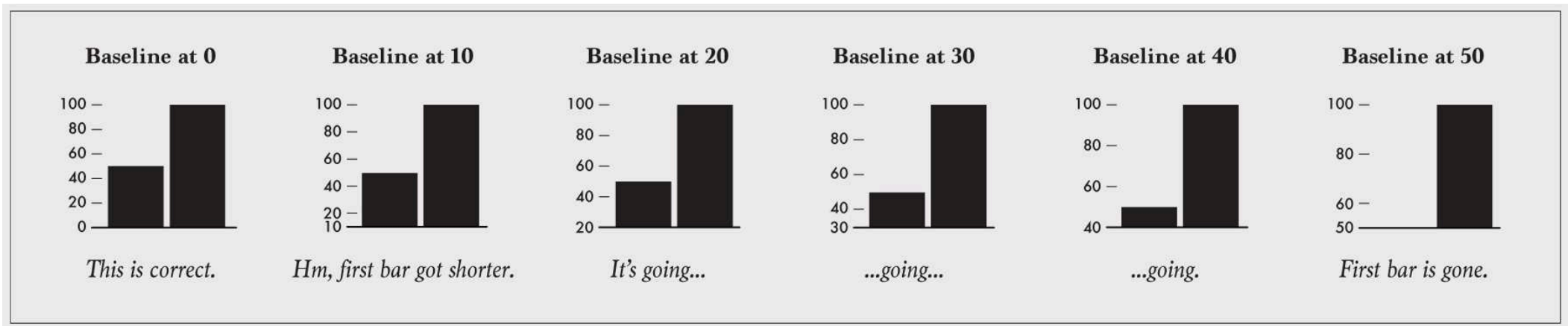

100
75
50
25
0
A
B
A
B
A
B
Kategorie 1
Kategorie 2
Kategorie 3

IRREFÜHREND ❌



UNPROBLEMATISCH ✅

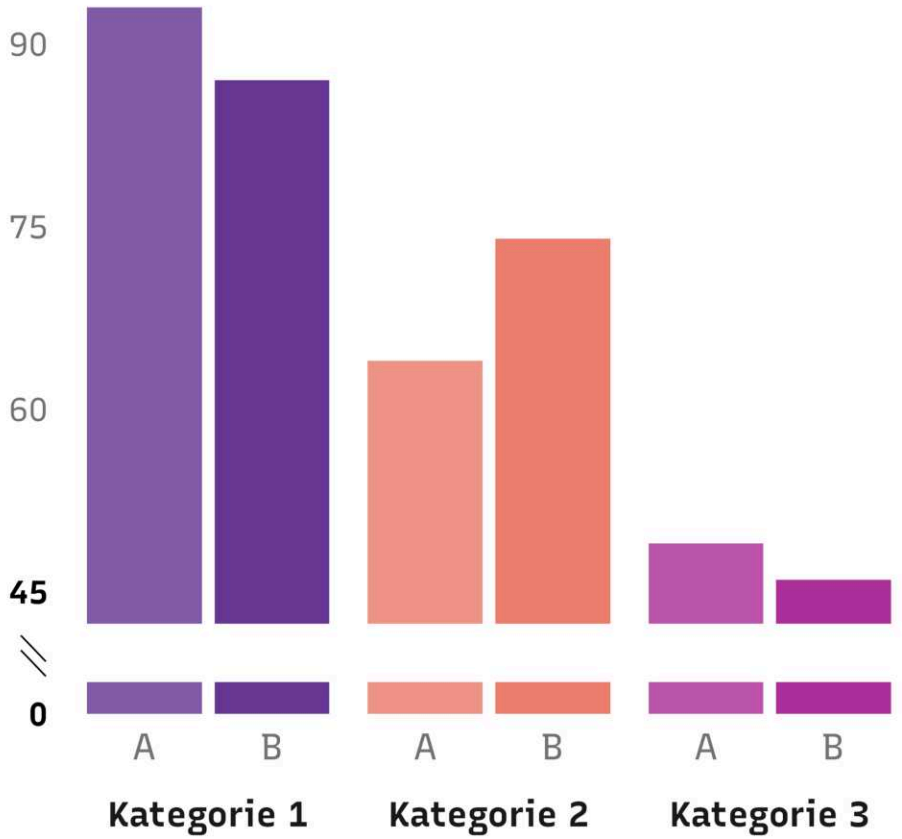




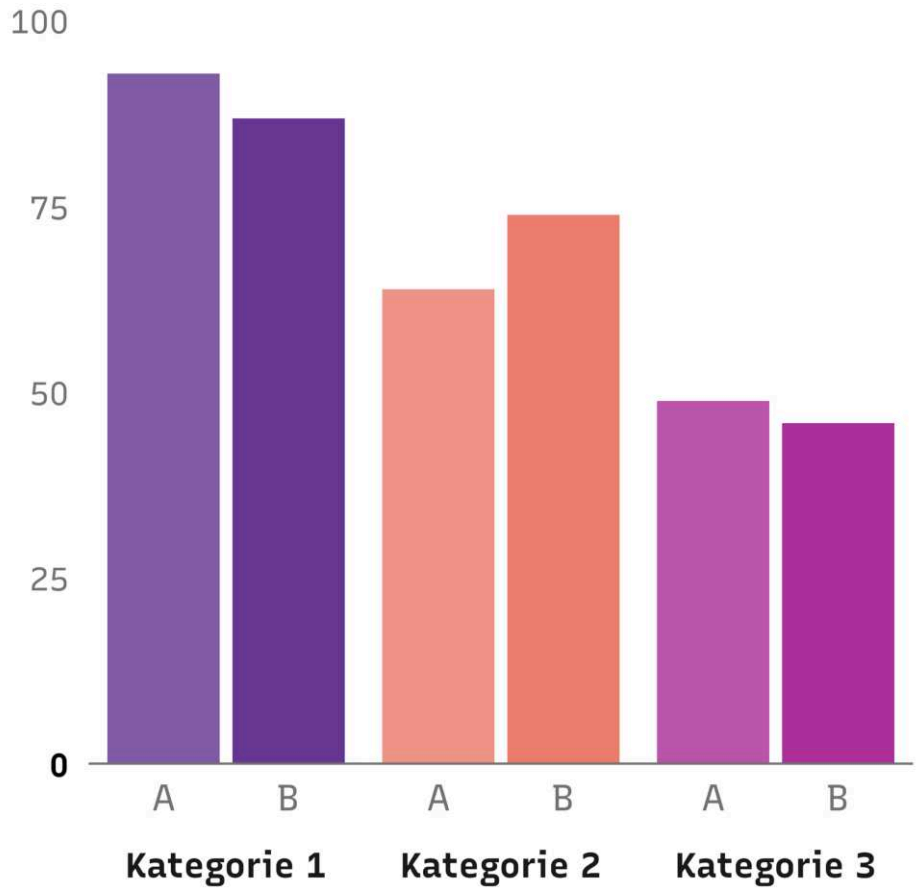
Quelle: Nathan Yau

mehr dazu: z.B. Witt (2019) und Correl, Bertini & Francoeri (2020)

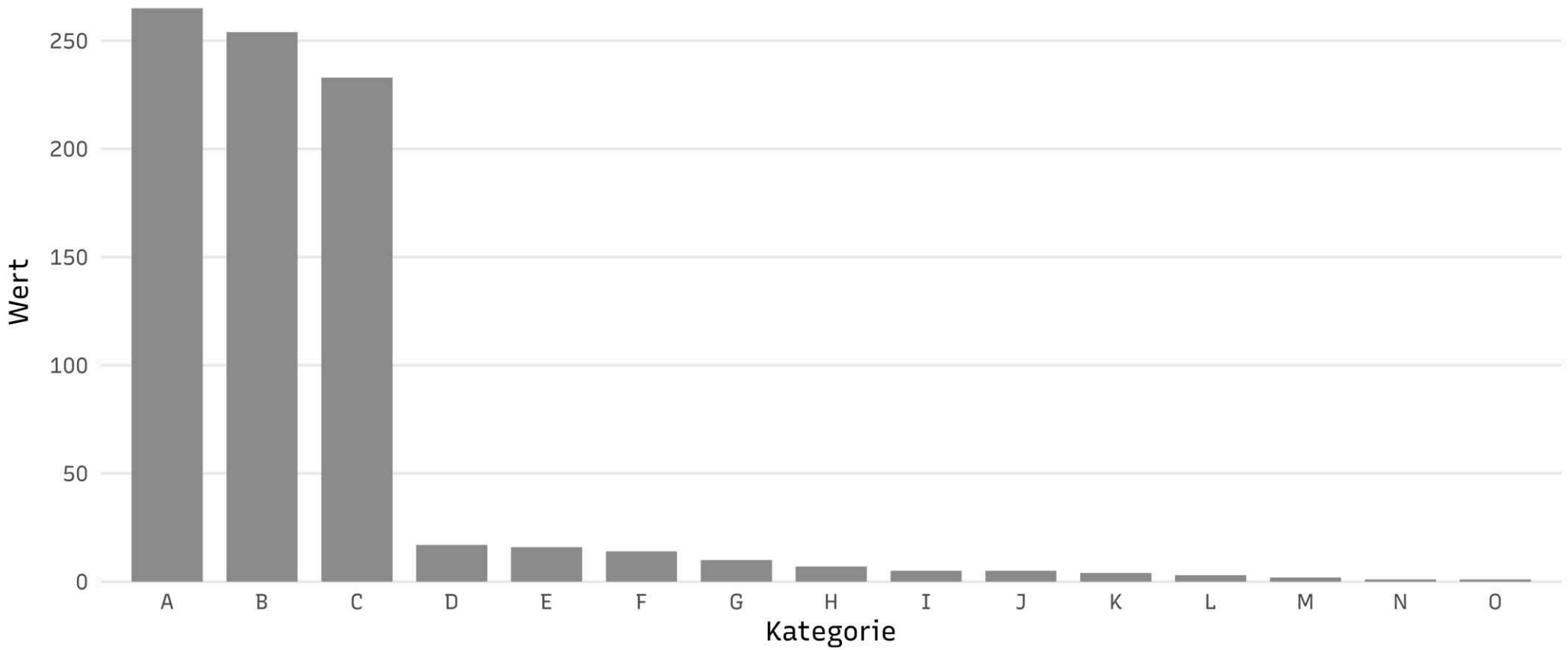
IRREFÜHREND ❌



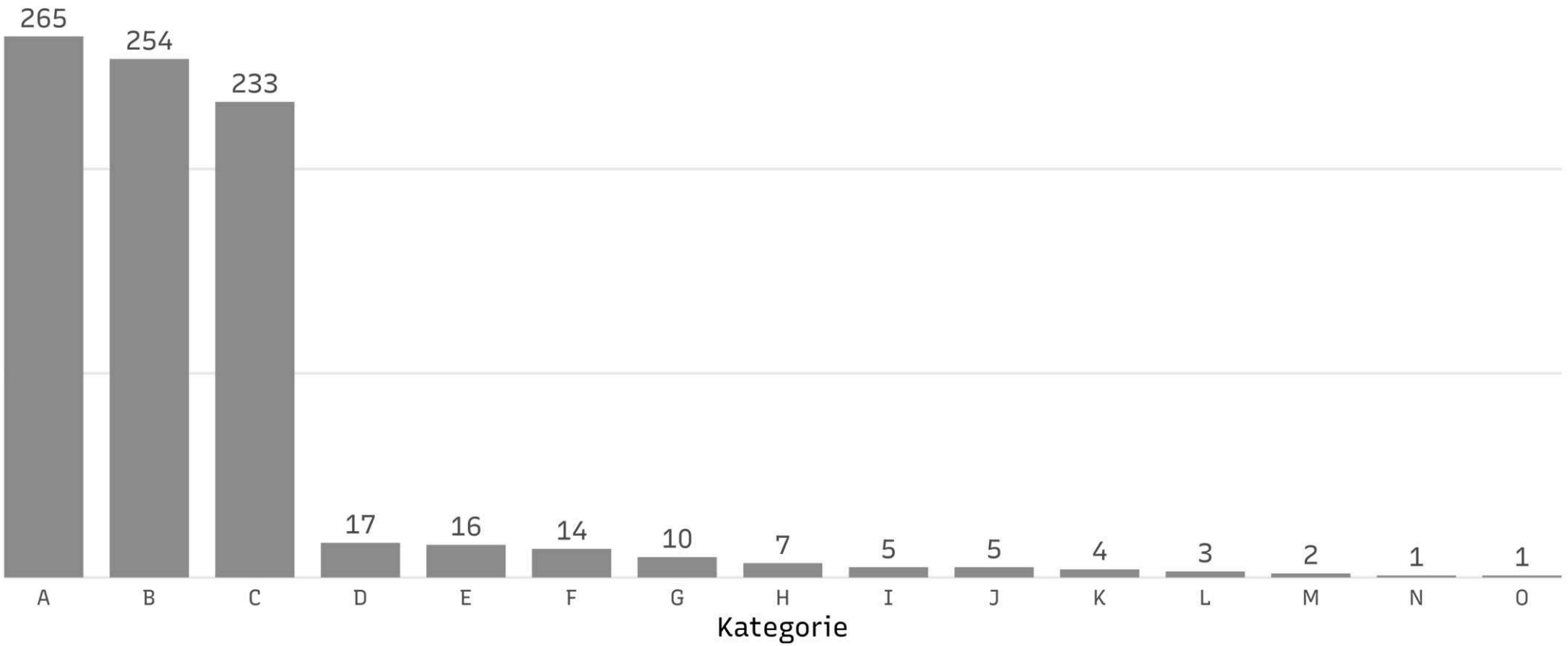
UNPROBLEMATISCH ✅



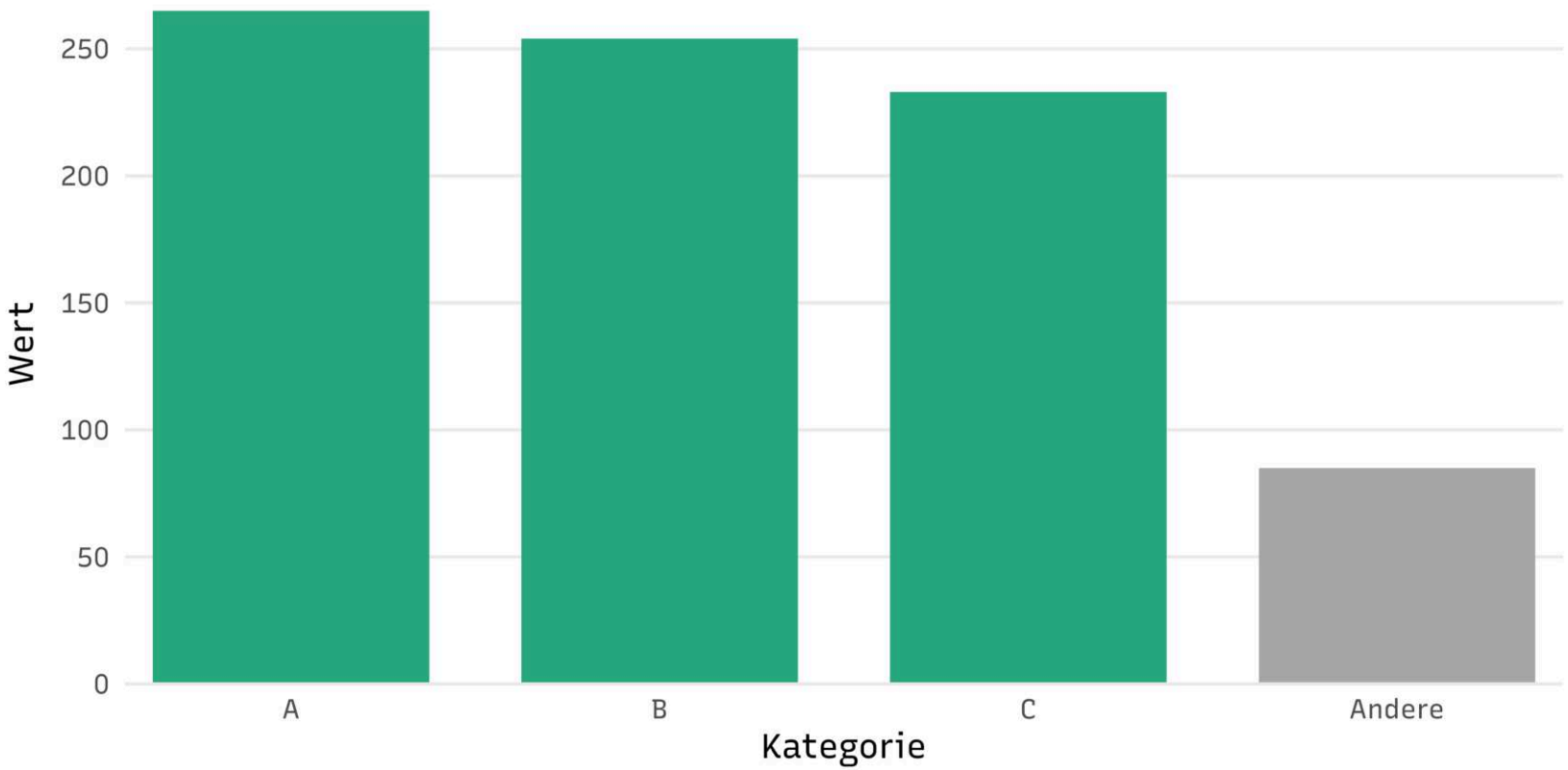
Schiefe Datenverteilung?



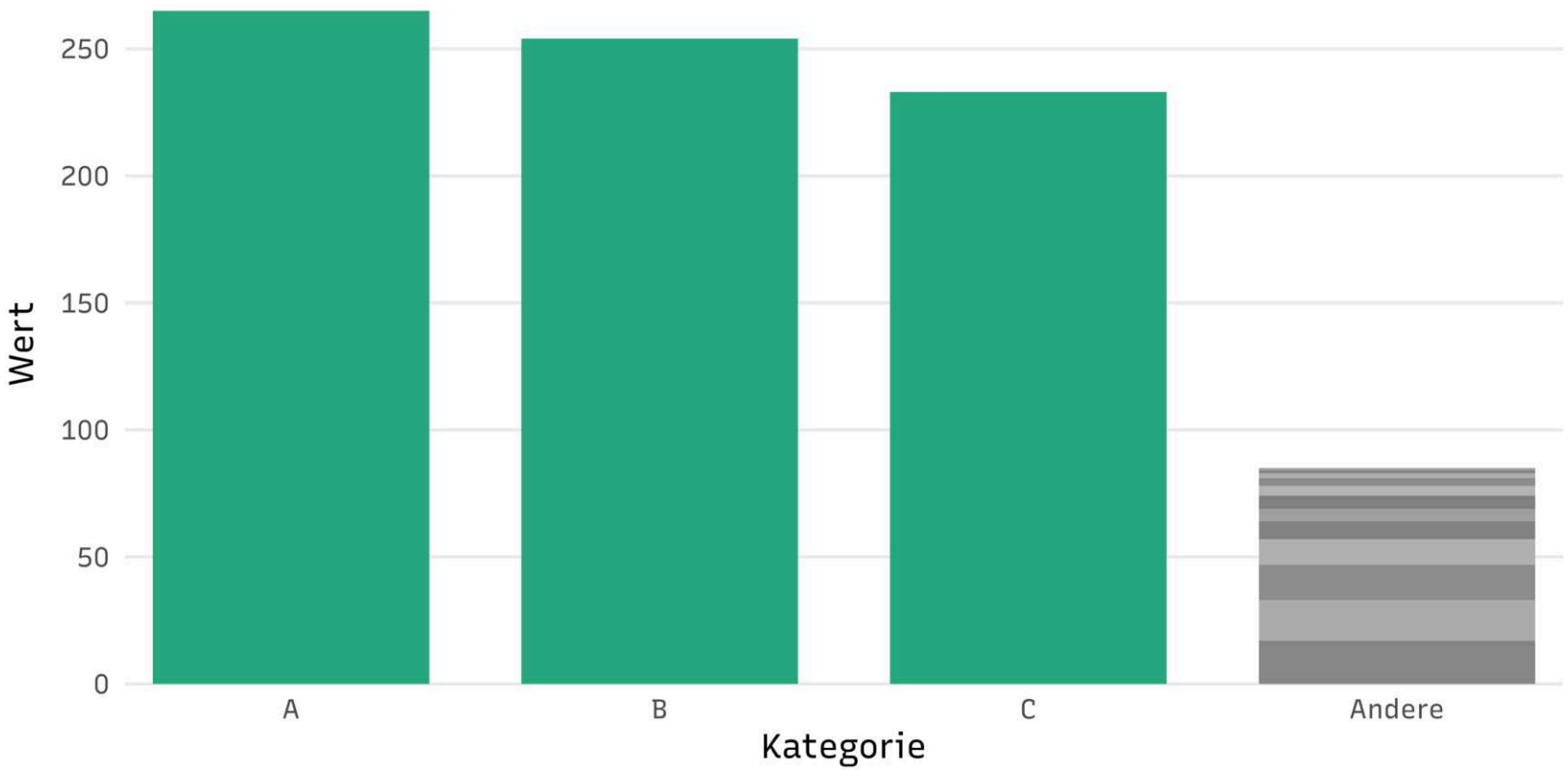
Schiefe Datenverteilung?



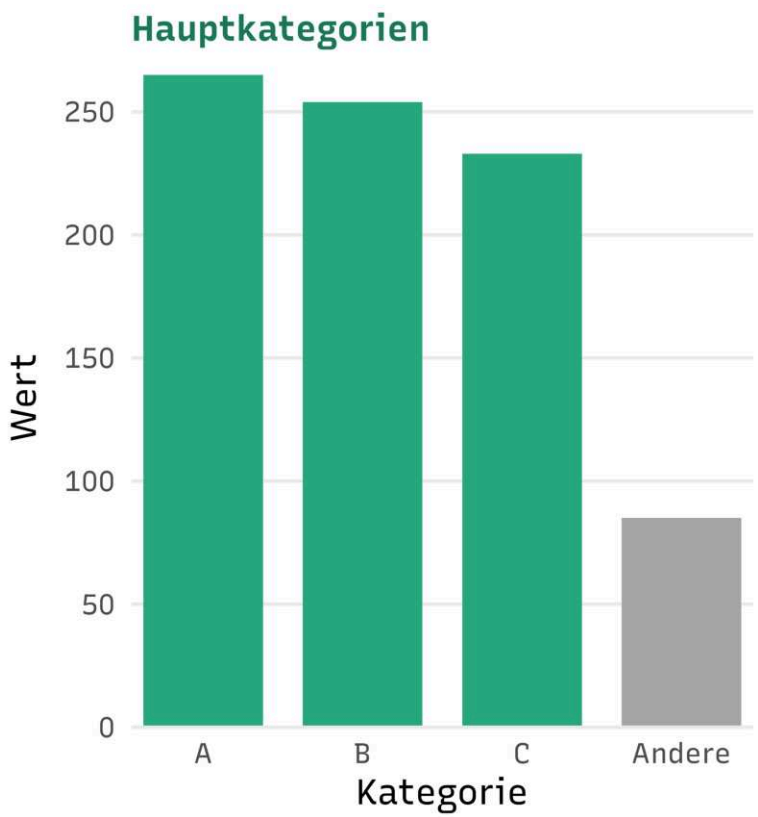
Schiefe Datenverteilung?

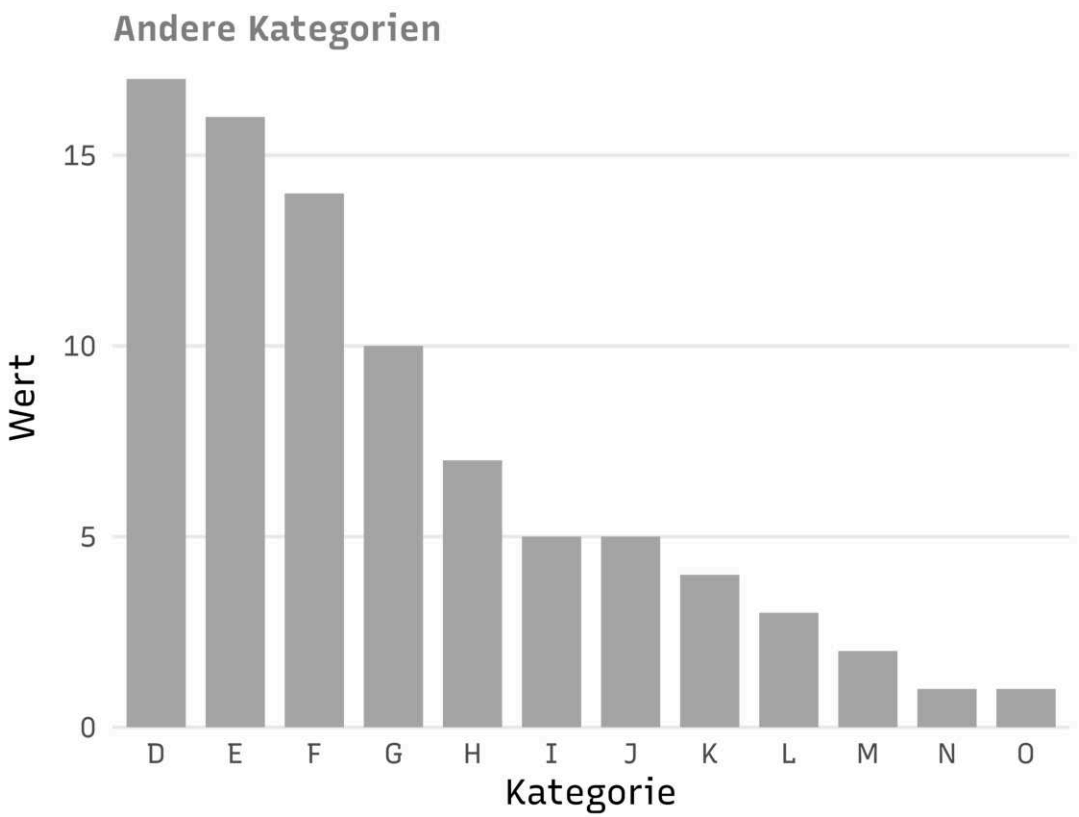


Schiefe Datenverteilung?

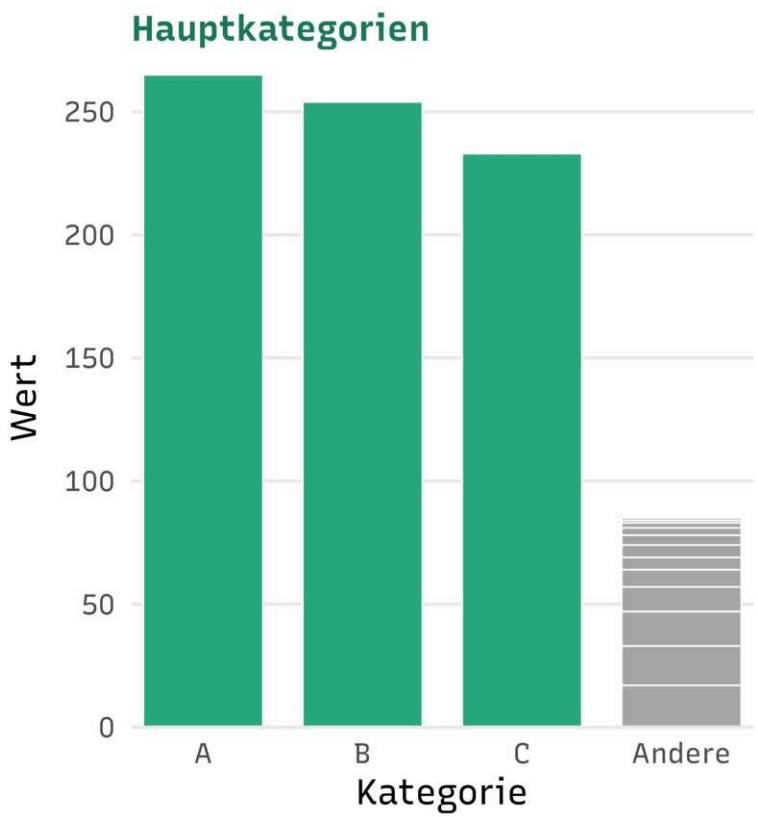


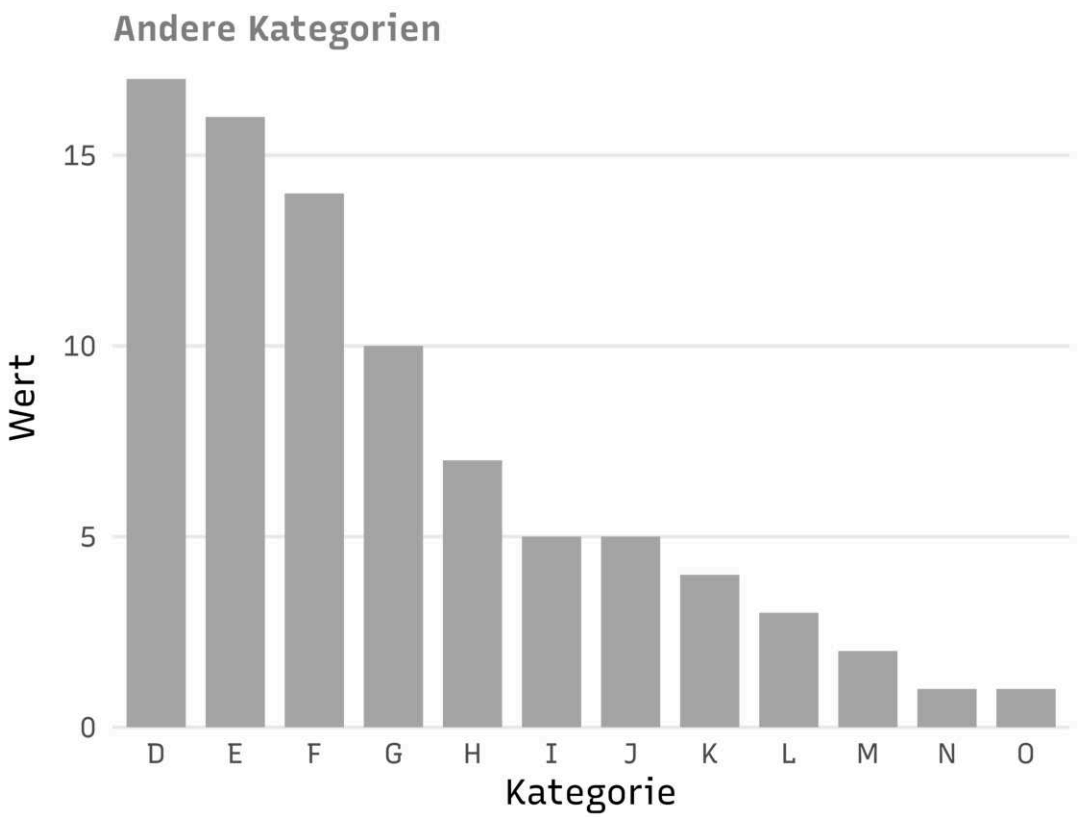
Schiefe Datenverteilung?





Schiefe Datenverteilung?





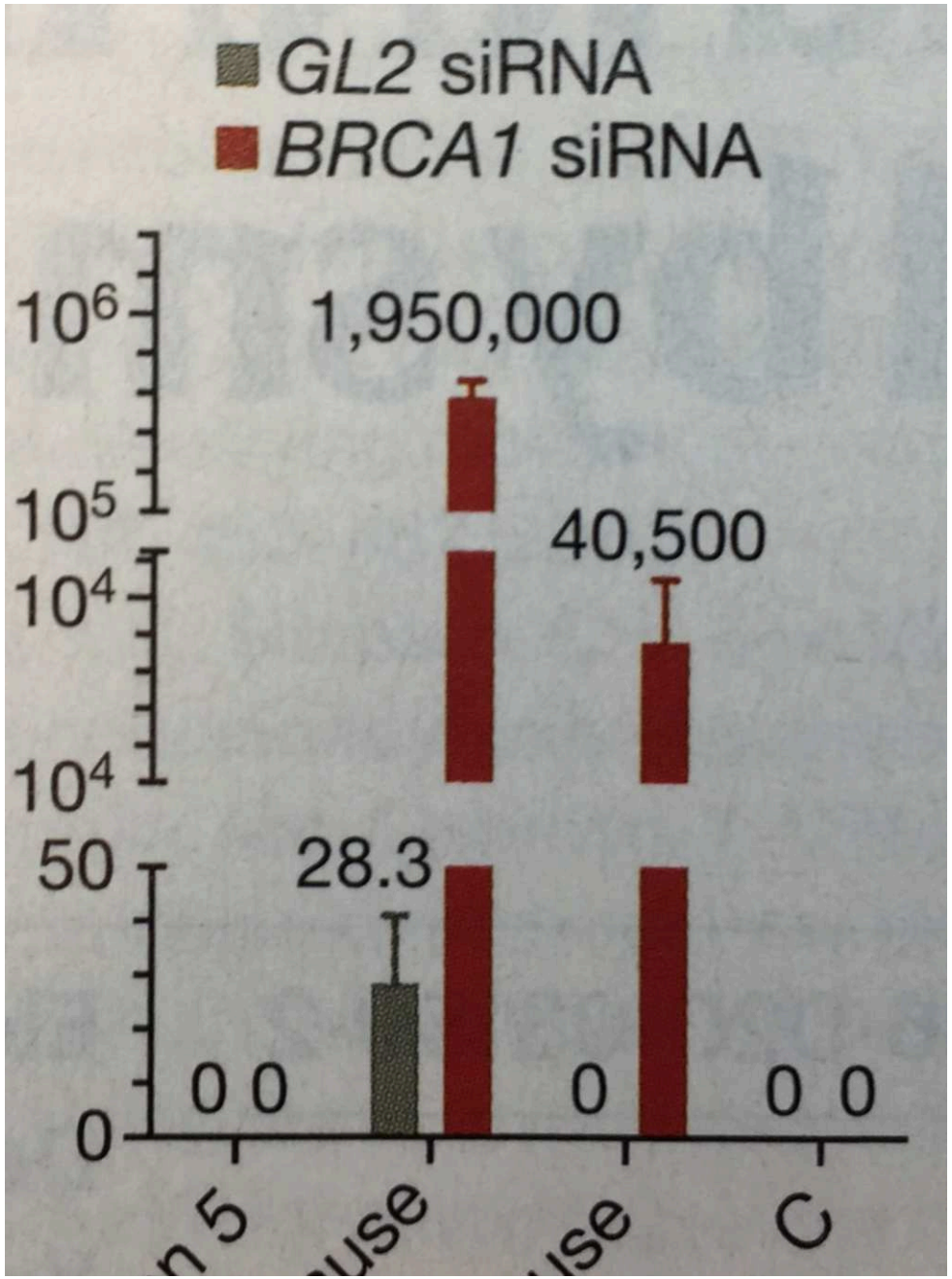
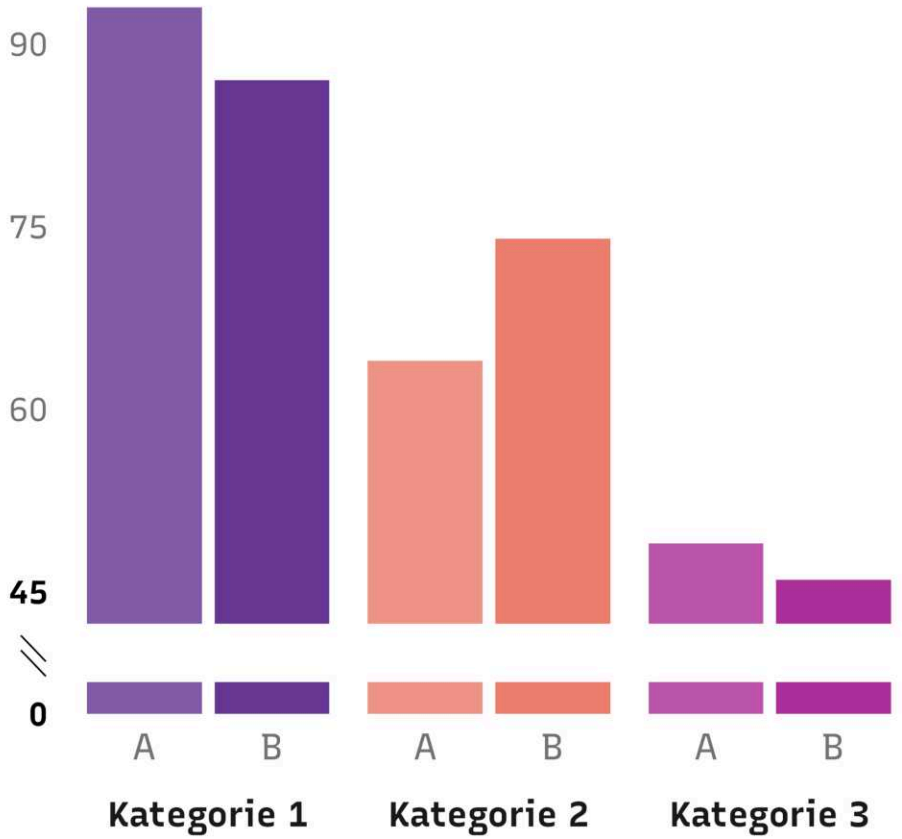
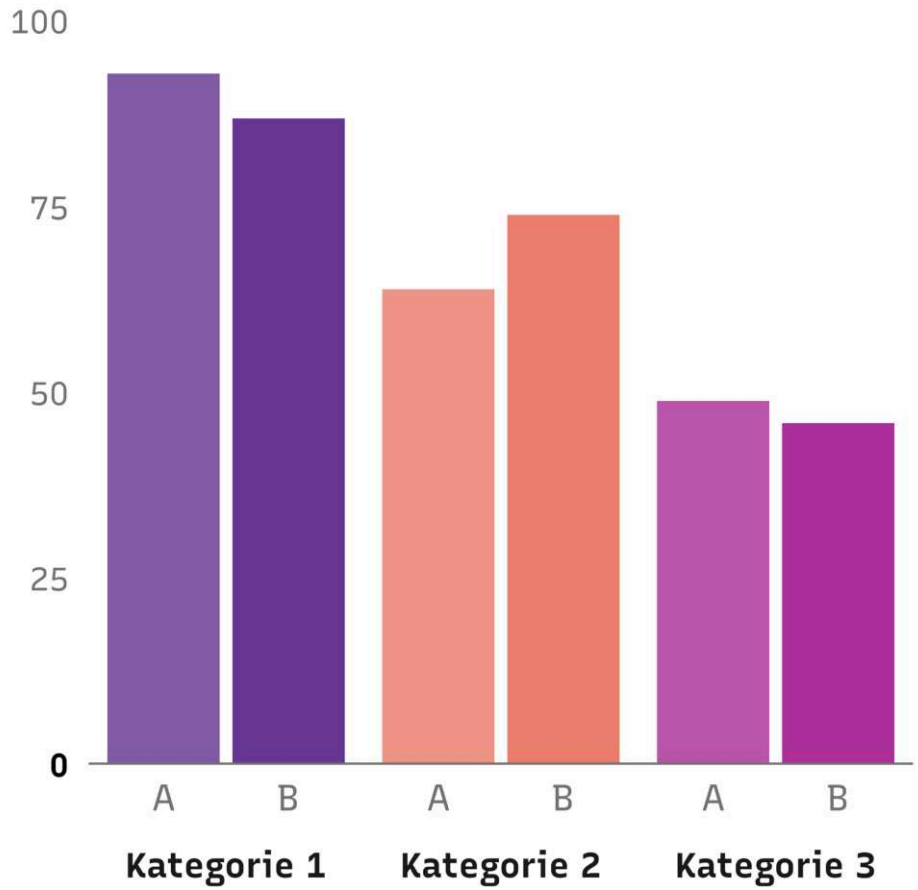


Photo Credit: Helena Jambor

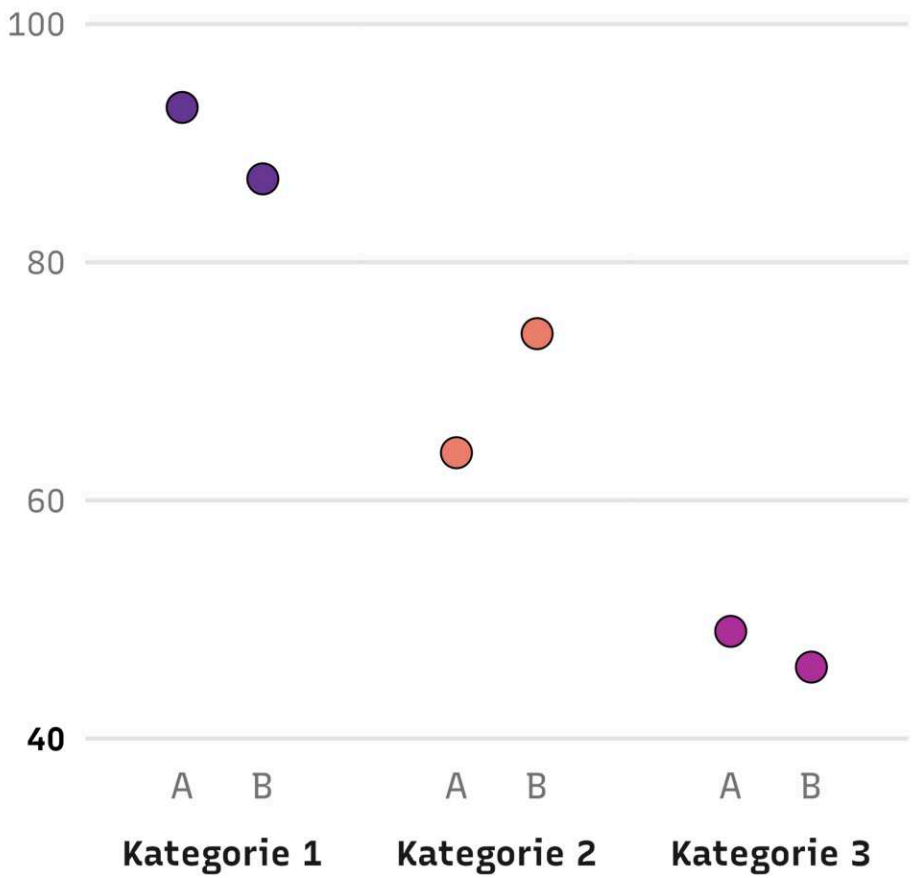
IRREFÜHREND ❌



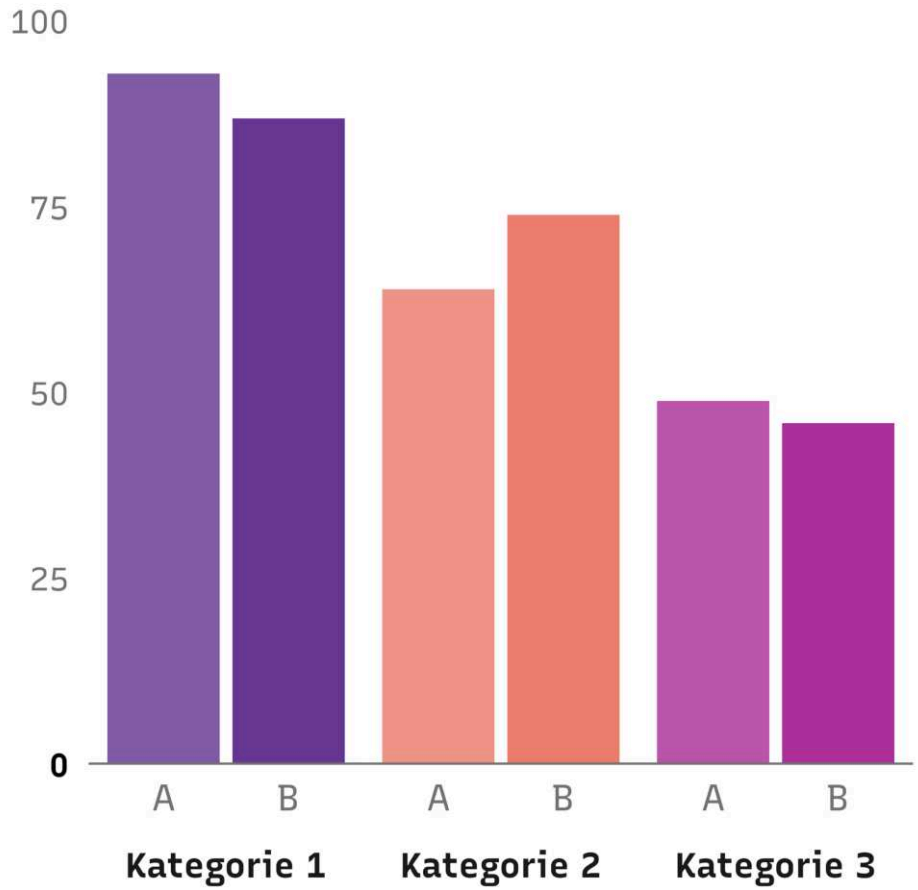
UNPROBLEMATISCH ✅



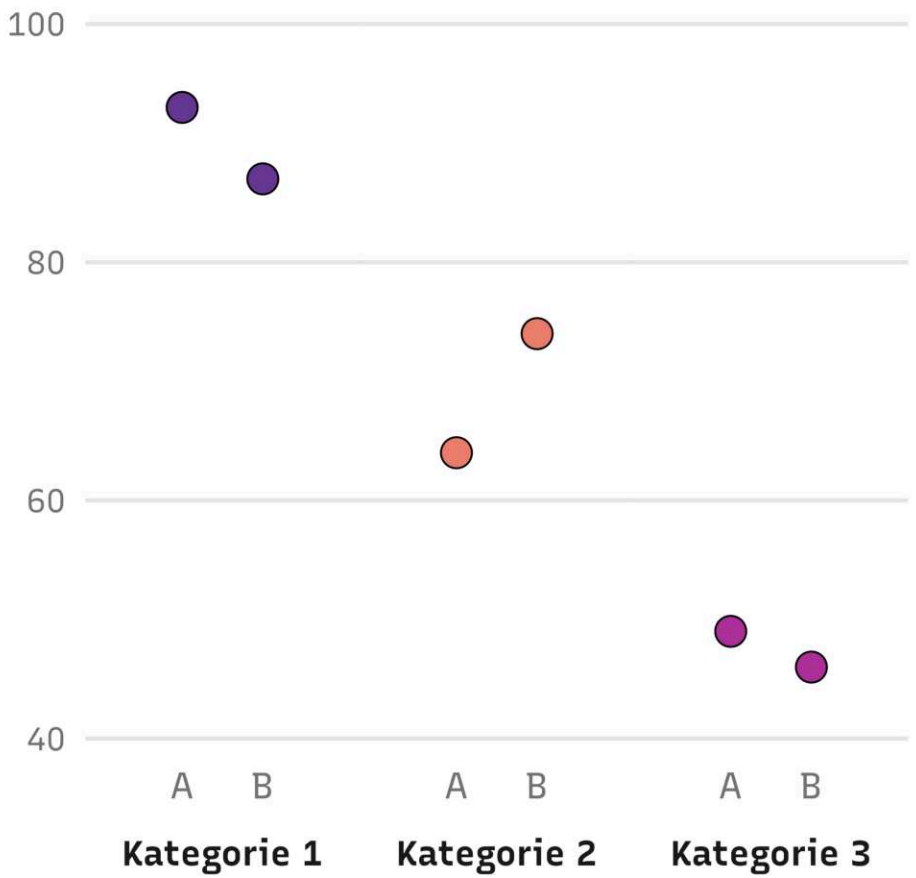
UNBEDENKLICH ✅



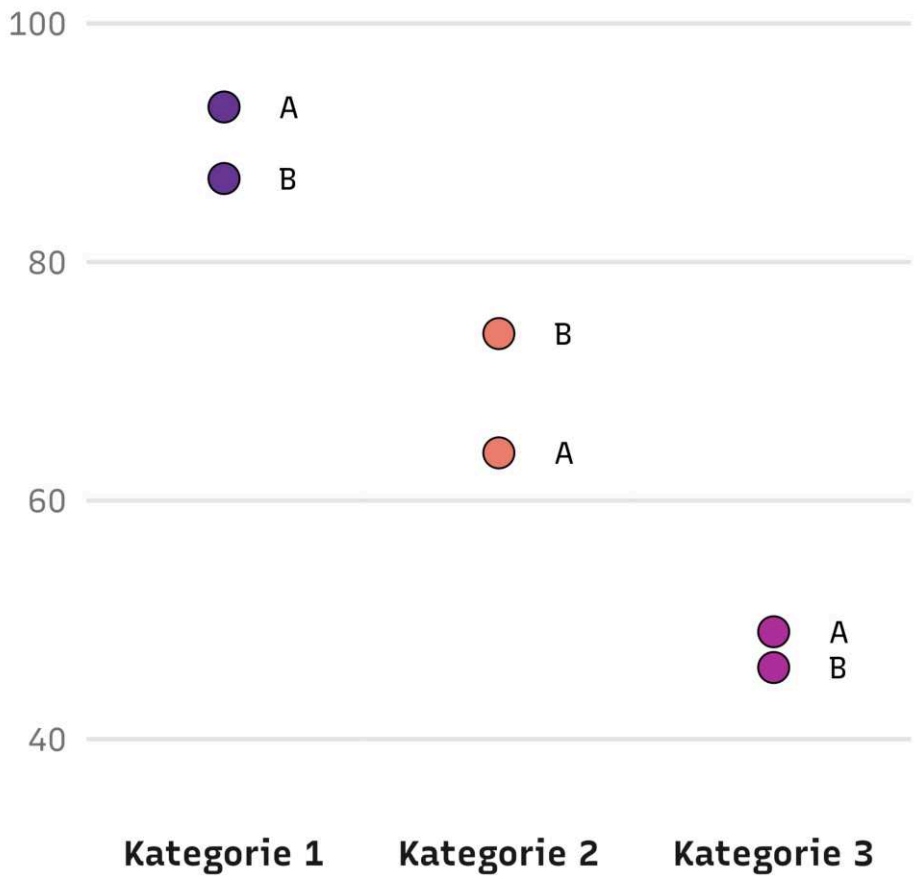
UNPROBLEMATISCH ✅



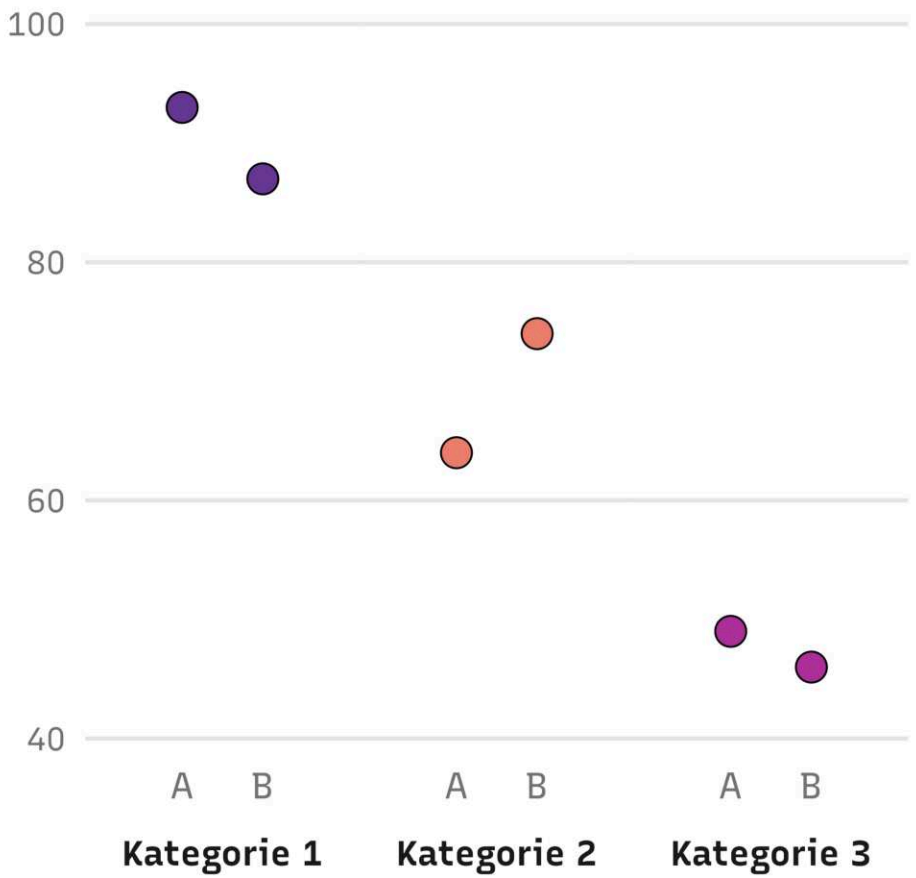
UNBEDENKLICH ✅



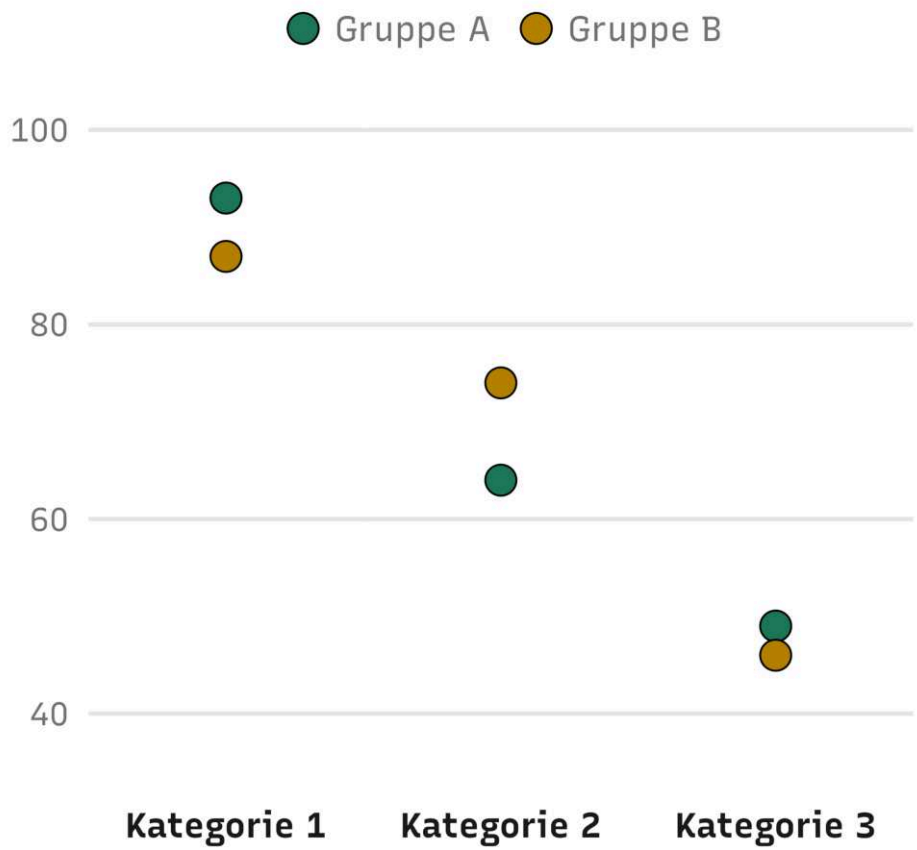
UNBEDENKLICH ✅



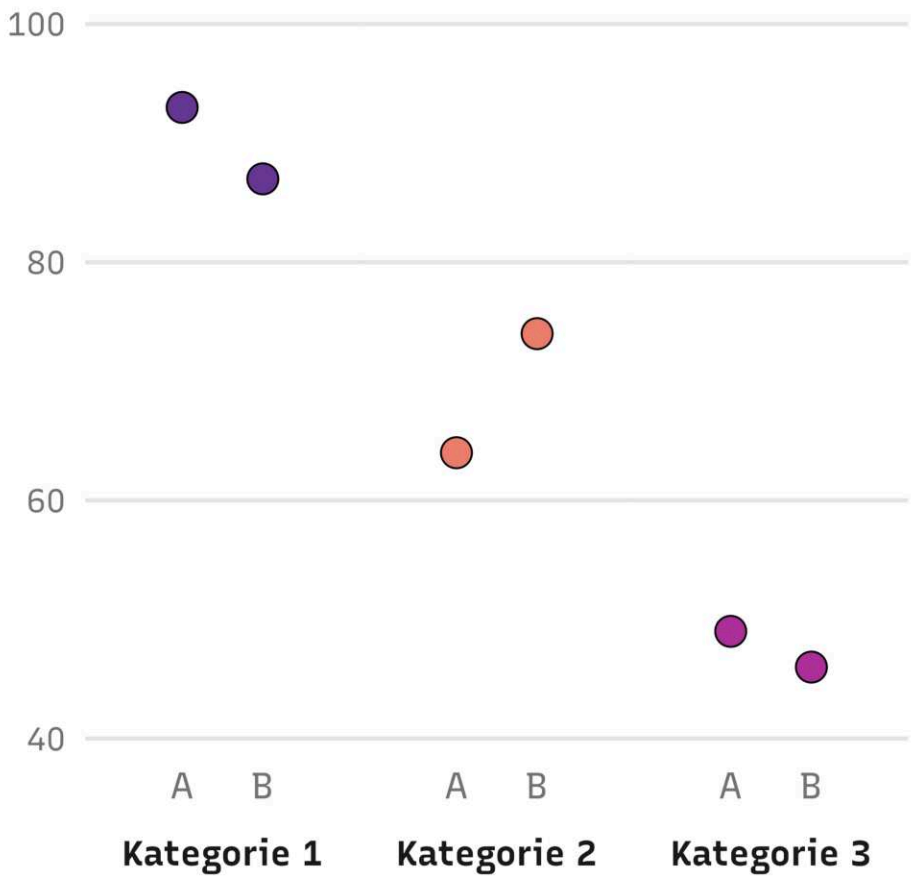
UNBEDENKLICH ✅



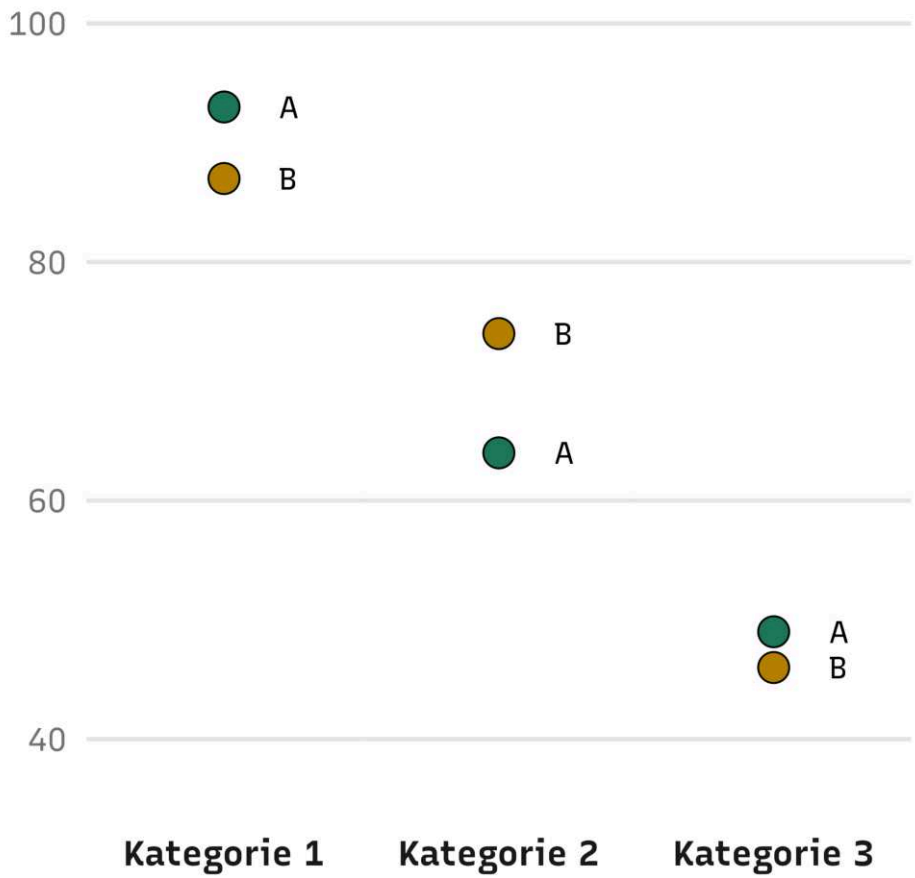
VERSTÄNDLICHER 👌



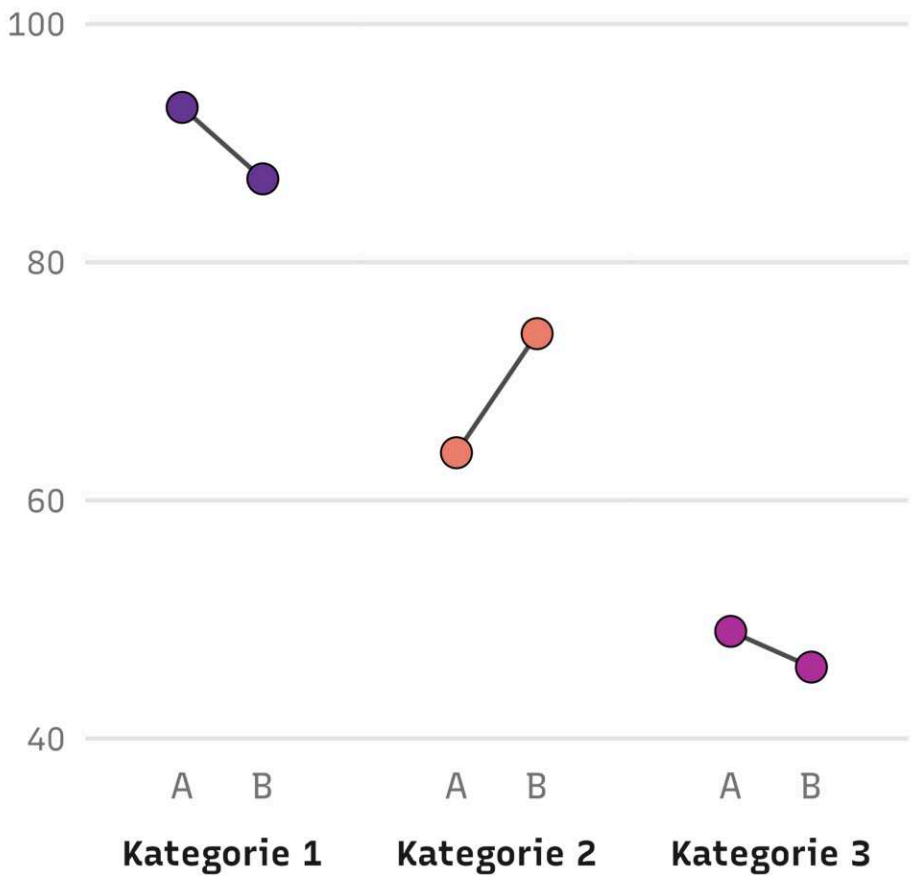
UNBEDENKLICH ✅



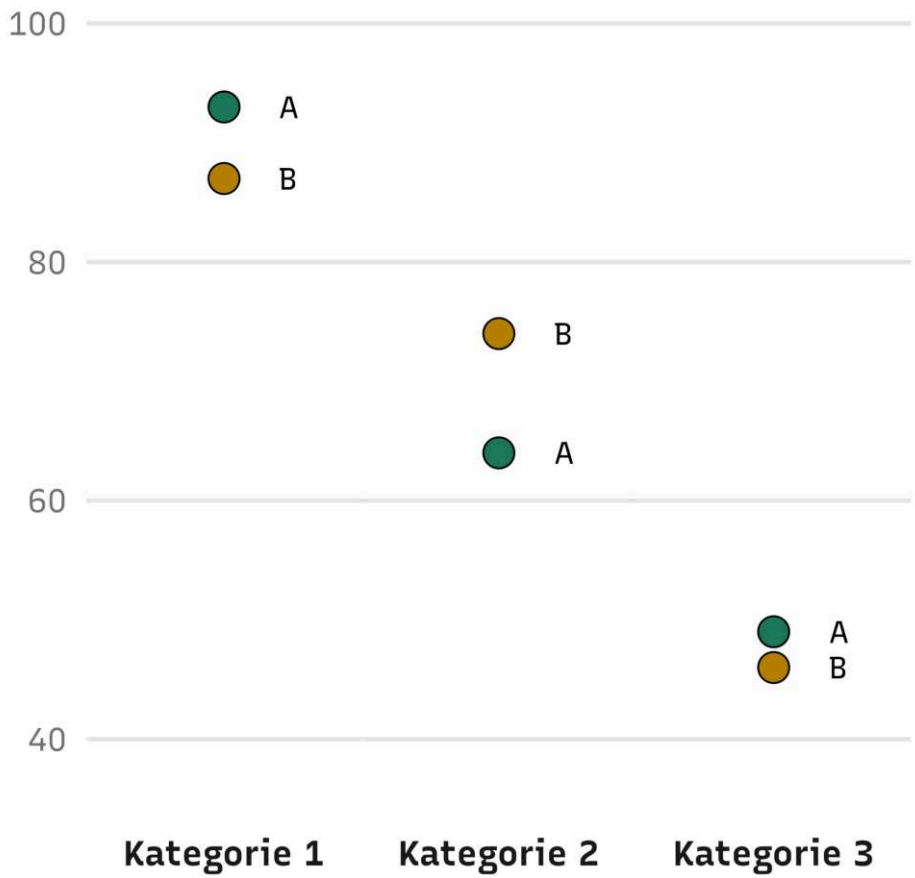
NOCH BESSER 🙌



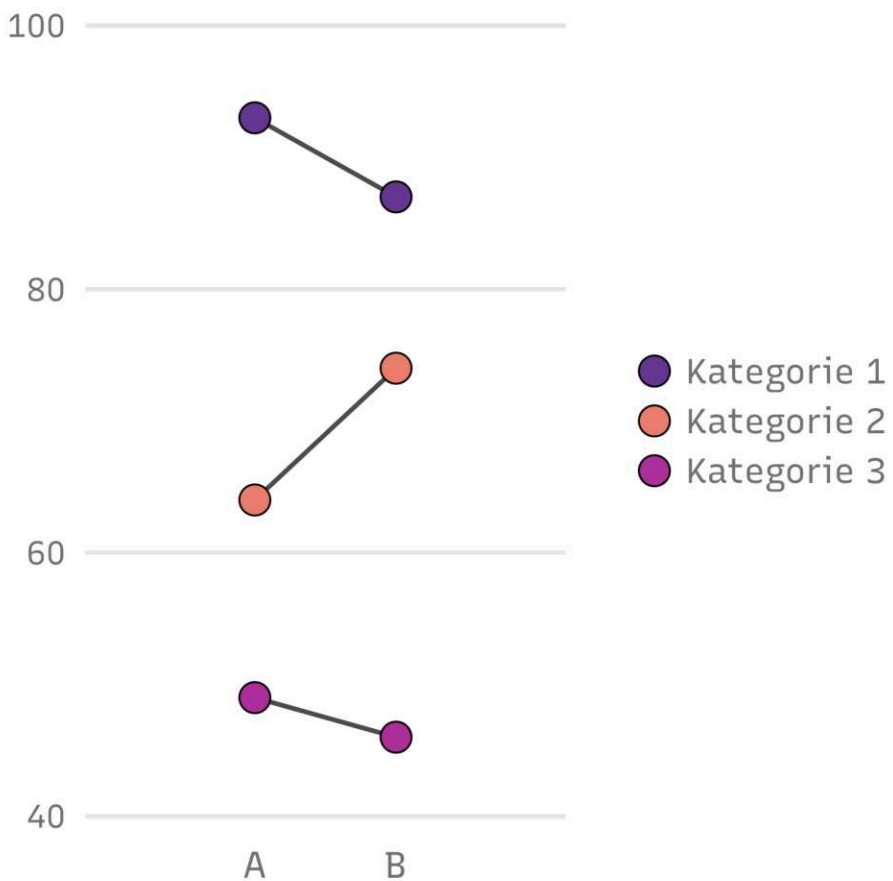
OFT SINNVOLL 💡



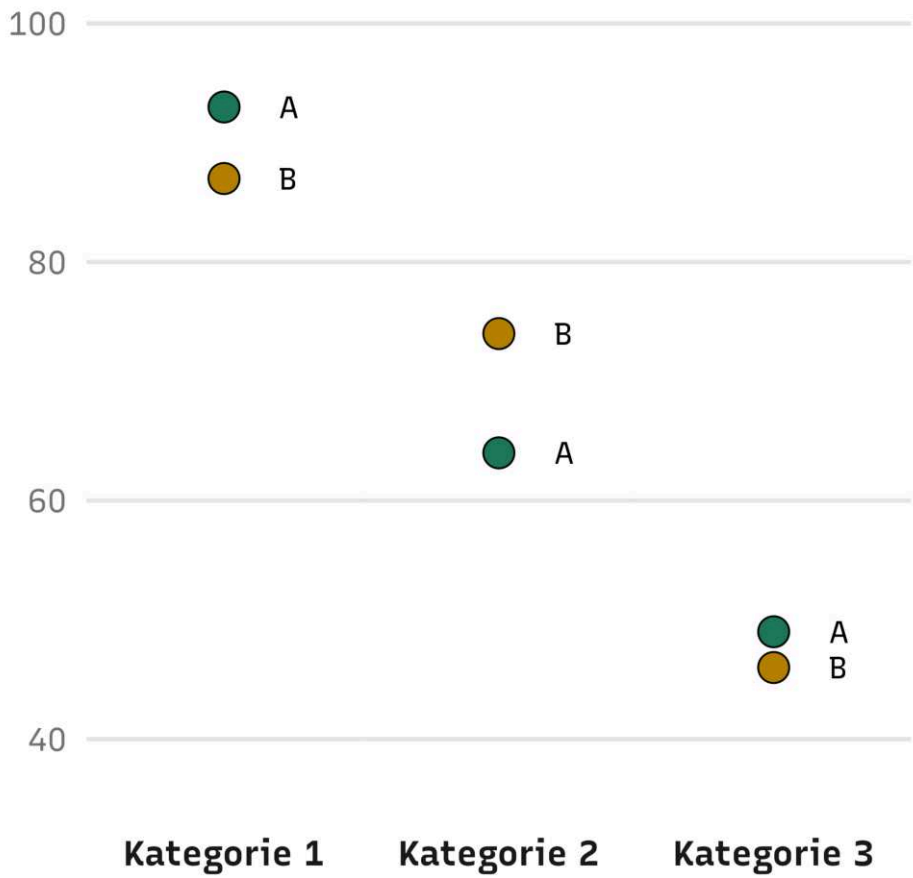
NOCH BESSER 🙌



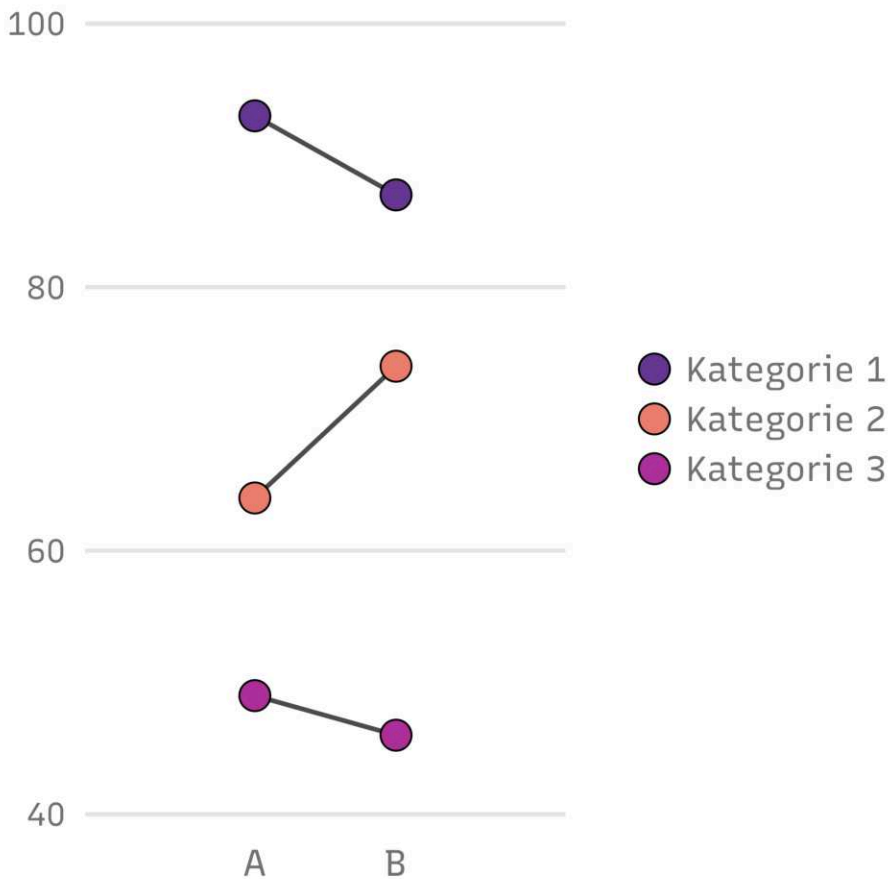
OFT SINNVOLL 💡



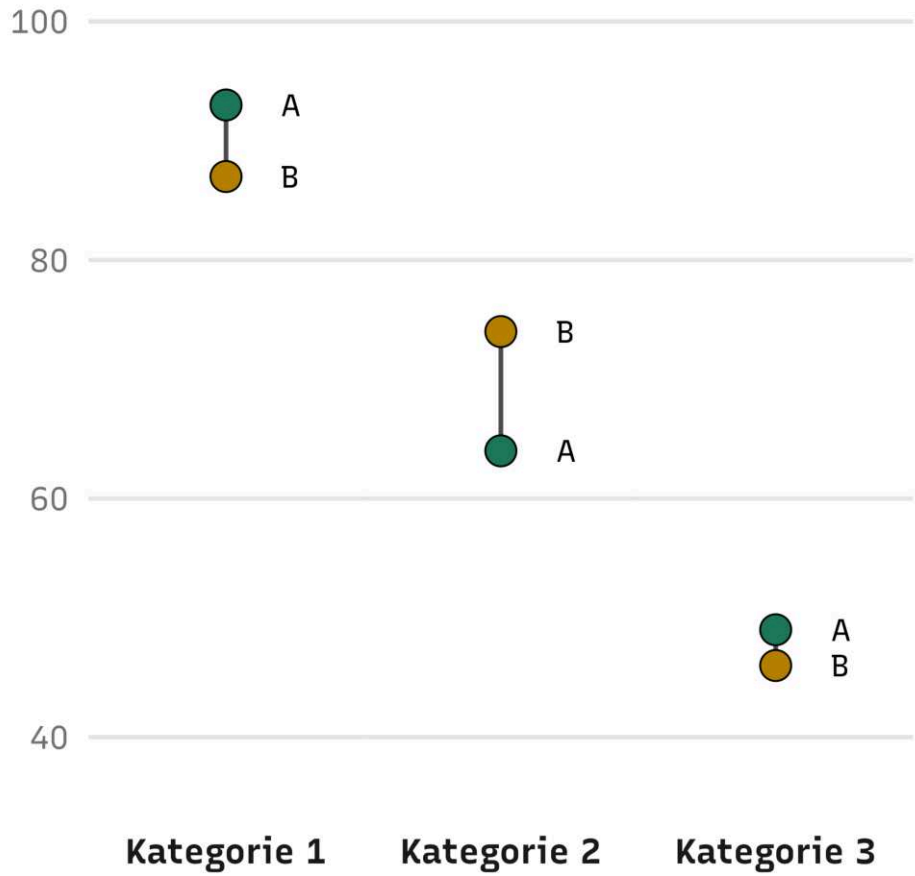
NOCH BESSER 🙌



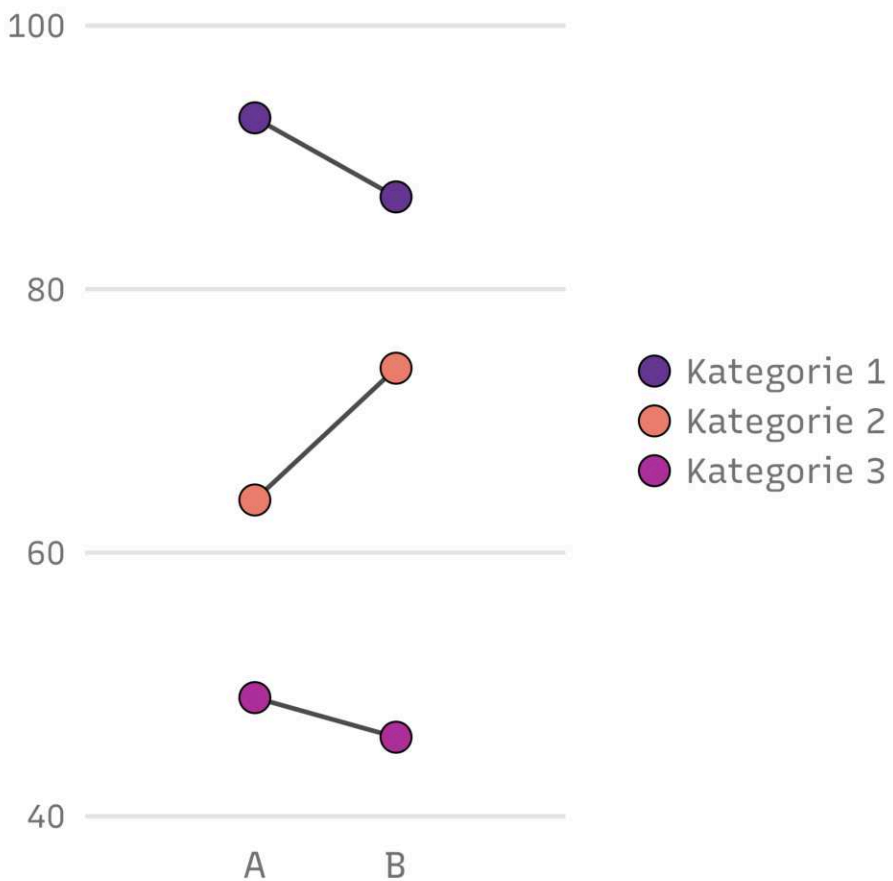
OFT SINNVOLL 💡



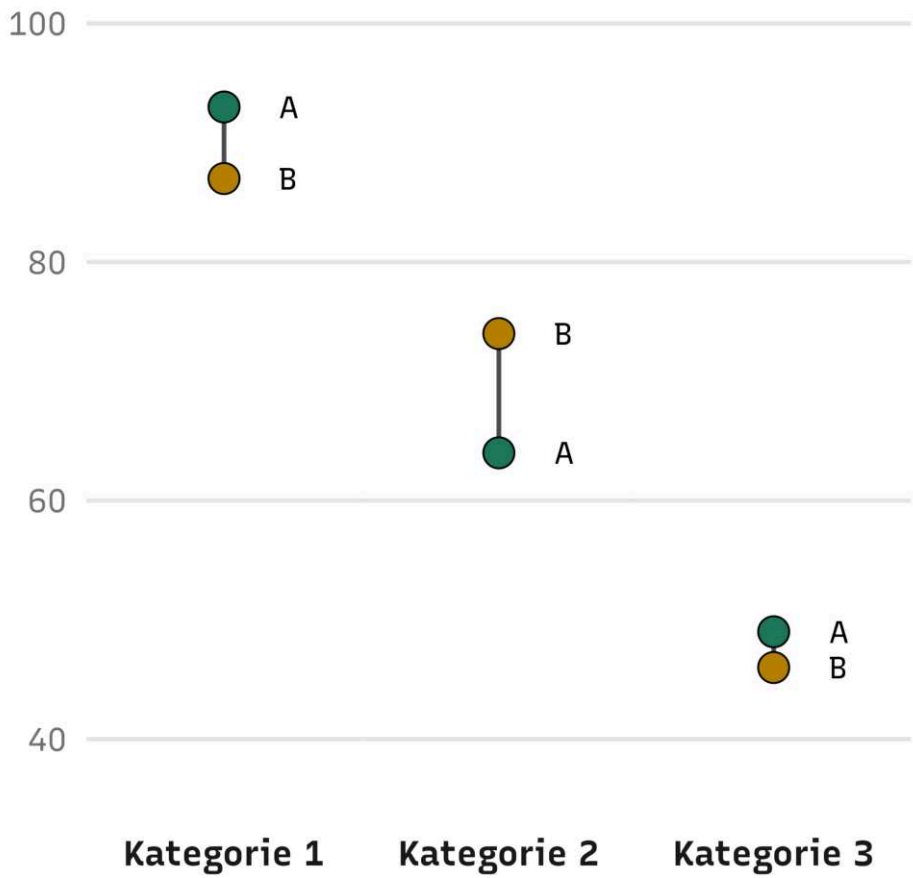
OFT SINNVOLL 💡



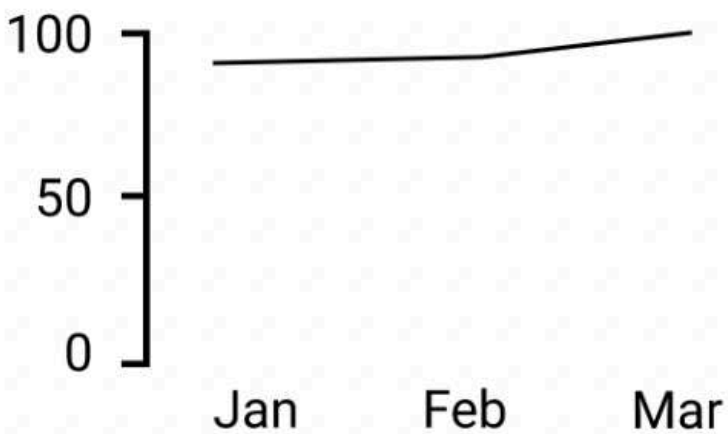
SLOPE CHART



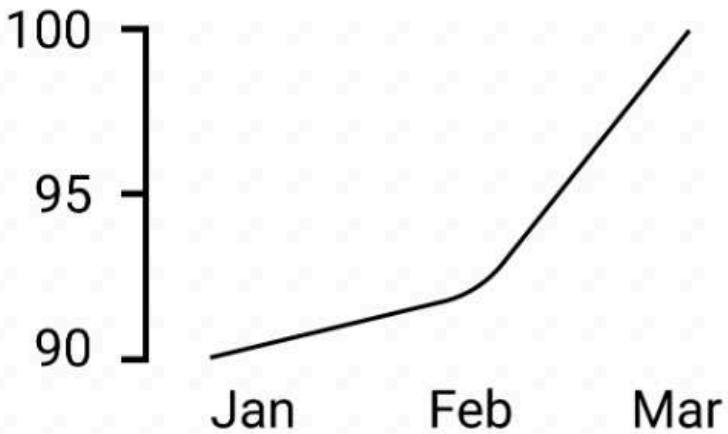
DUMBBELL PLOT



Immer bei Null beginnen?

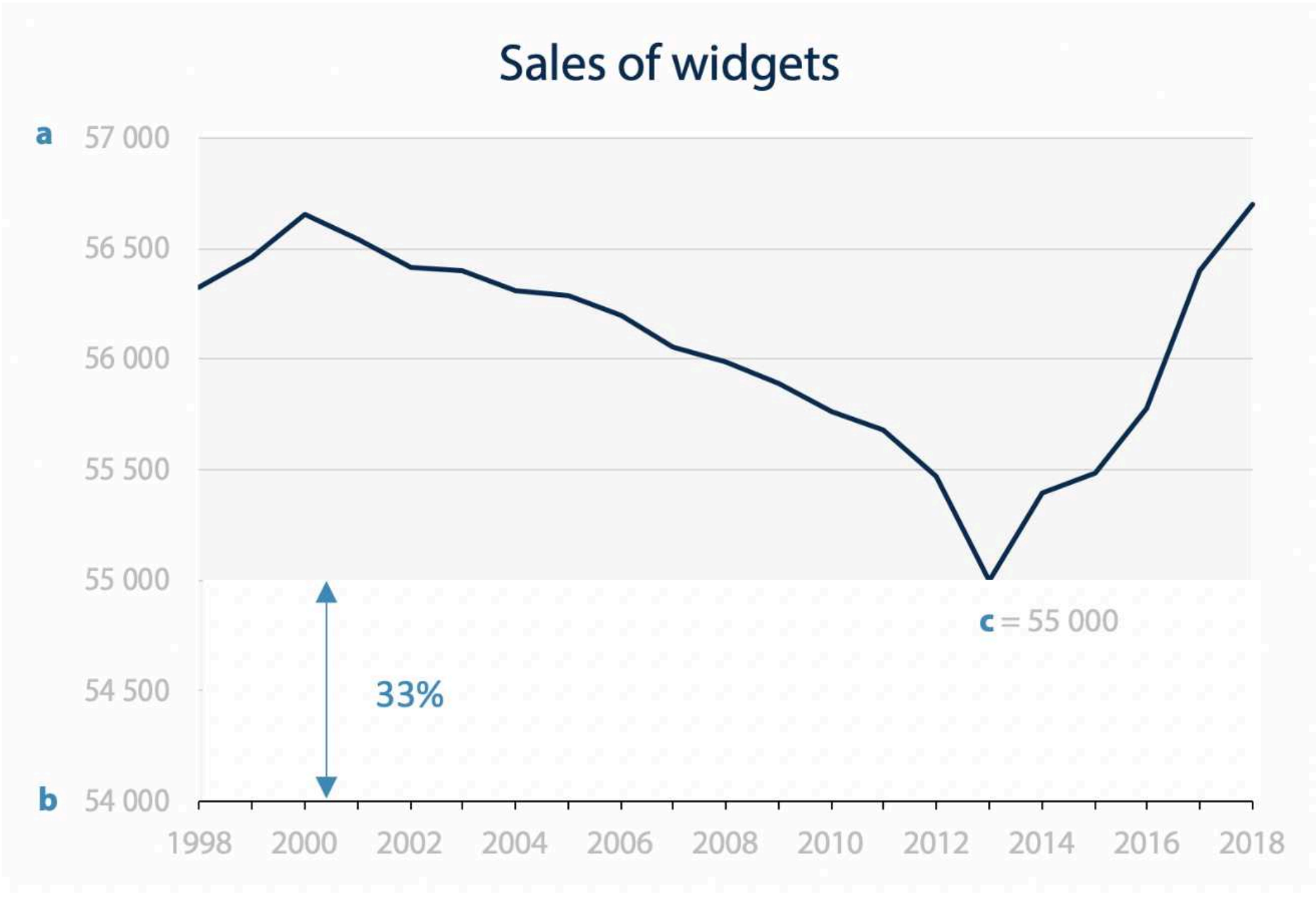


Acceptable, but starting the vertical axis at zero obscures changes in values



Better: By reducing the vertical axis to match the values, we see change more clearly

Quelle: "Hands-On Data Visualization" von Jack Dougherty & Ilya Ilyankou



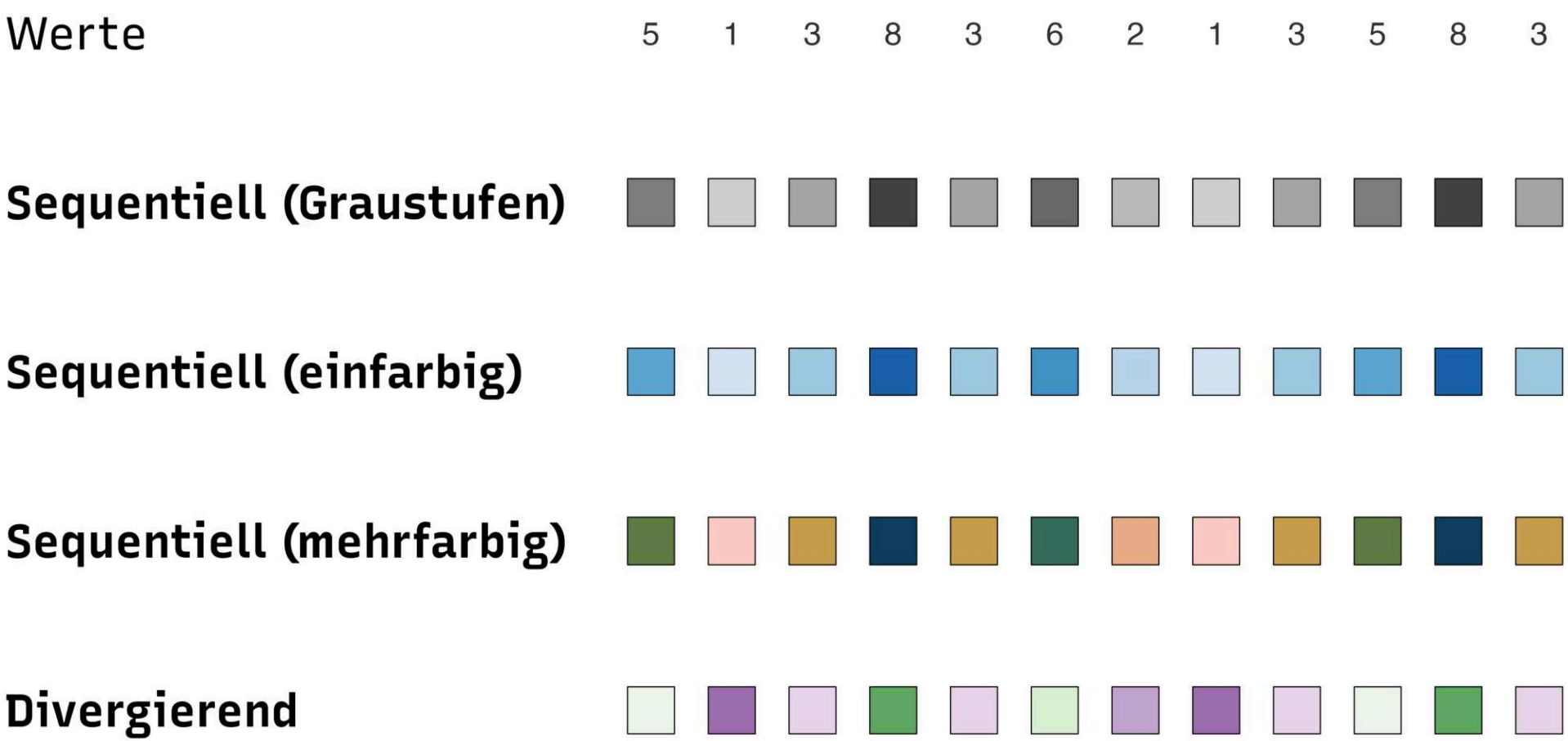
Quelle: Francis Gagnon

Kodieren mit Farben

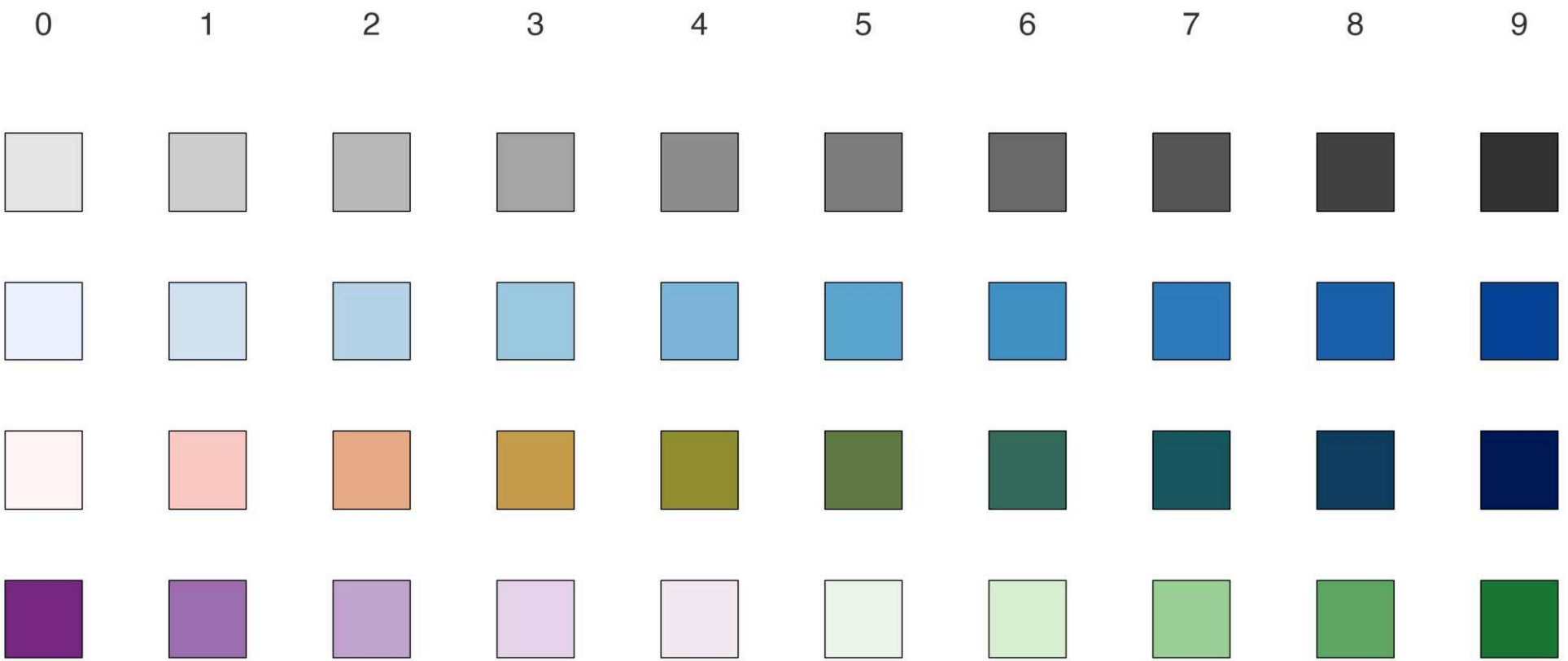
Informative und leicht lesbare Farbschlüssel nutzen (und gestalten)

Cédric Scherer Data Visualization & Information Design

Gleiche Werte, andere Farben



Gleiche Werte, andere Farben



Typen von Farbpaletten

Sequentiell

Beispiel 

Grautöne 

numerische Informationen mit geordneter Skala

verwende den höchsten Kontrast für die wichtigsten Informationen

entweder einfarbige oder mehrfarbige Farbpaletten

Typen von Farbpaletten

Sequentiell

Beispiel 

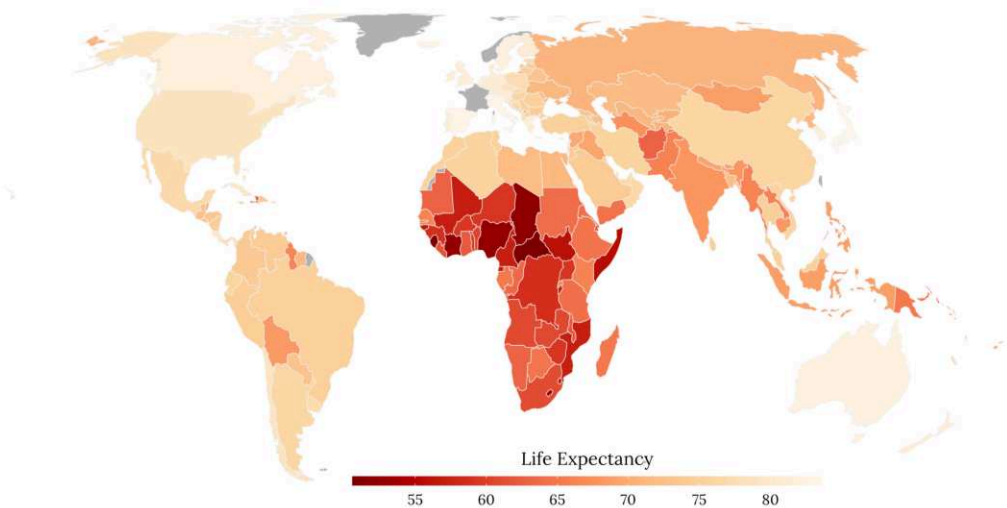
Grautöne 

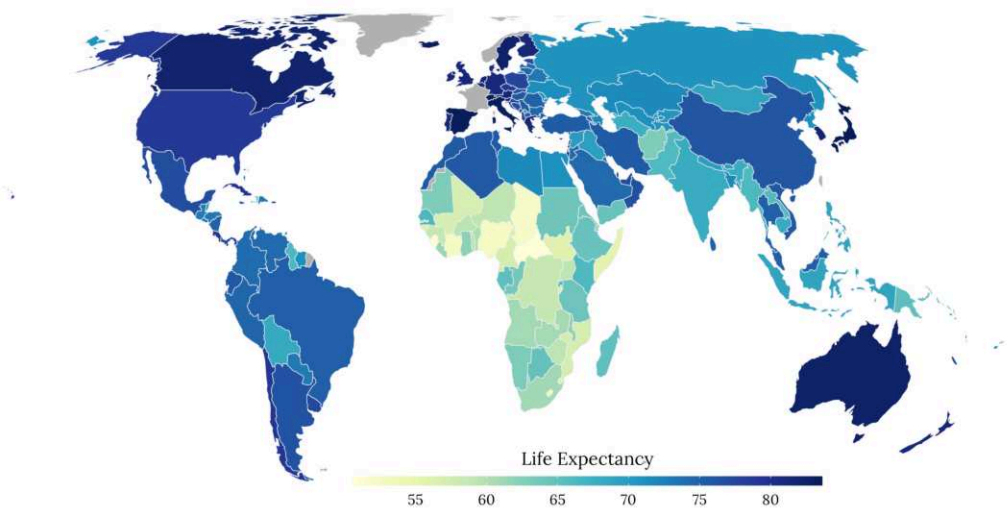
Werte mit geordneter Skala
(numerisch oder ordinal)

verwende den höchsten Kontrast
für die wichtigsten Informationen

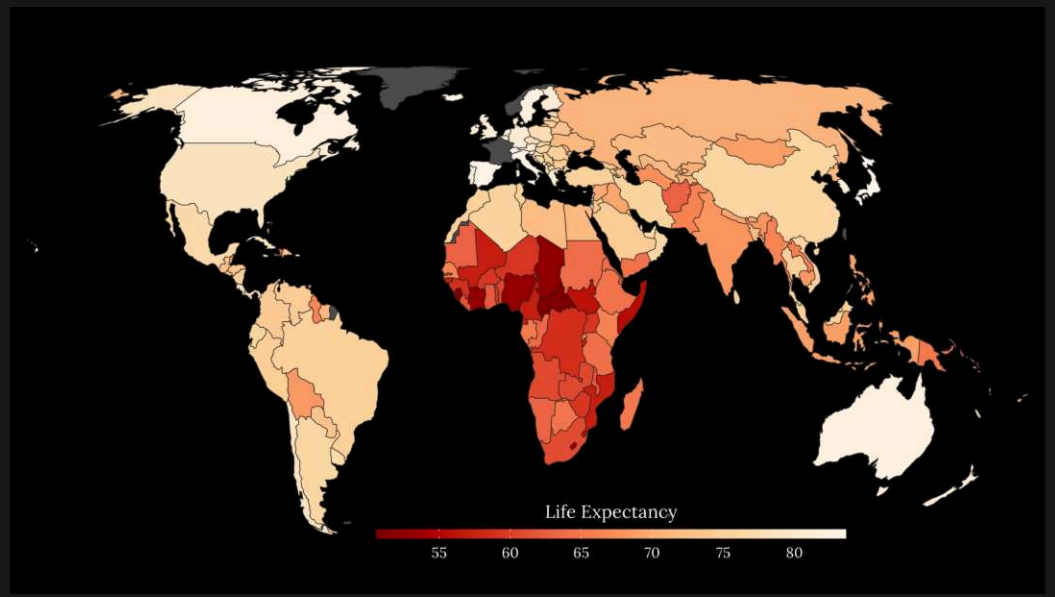
entweder einfarbige oder
mehrfarbige Farbpaletten

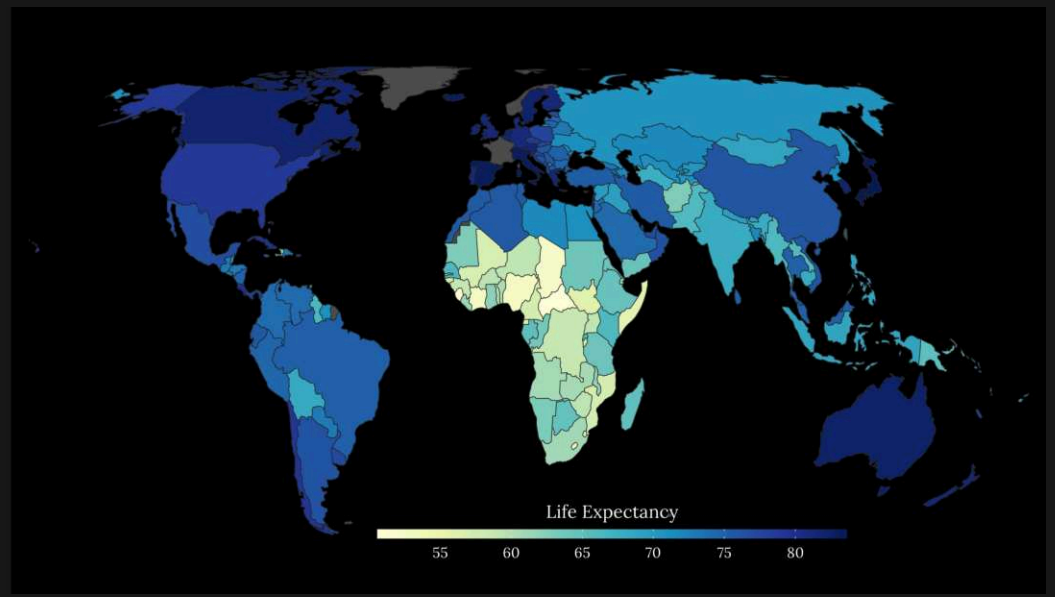
Sequentiell: Dunkel ist mehr (?)





Sequentiell: Dunkel ist mehr (?)





Typen von Farbpaletten

Sequentiell

Beispiel 

Grautöne 

Werte mit geordneter Skala
(numerisch oder ordinal)

verwende den höchsten Kontrast für die wichtigsten Informationen

entweder einfarbige oder mehrfarbige Farbpaletten

Divergierend

Beispiel 

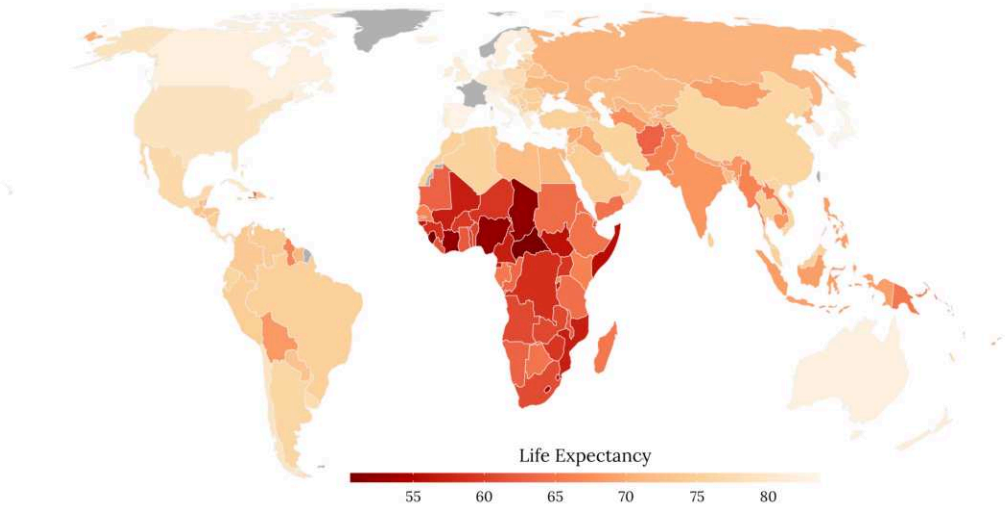
Grautöne 

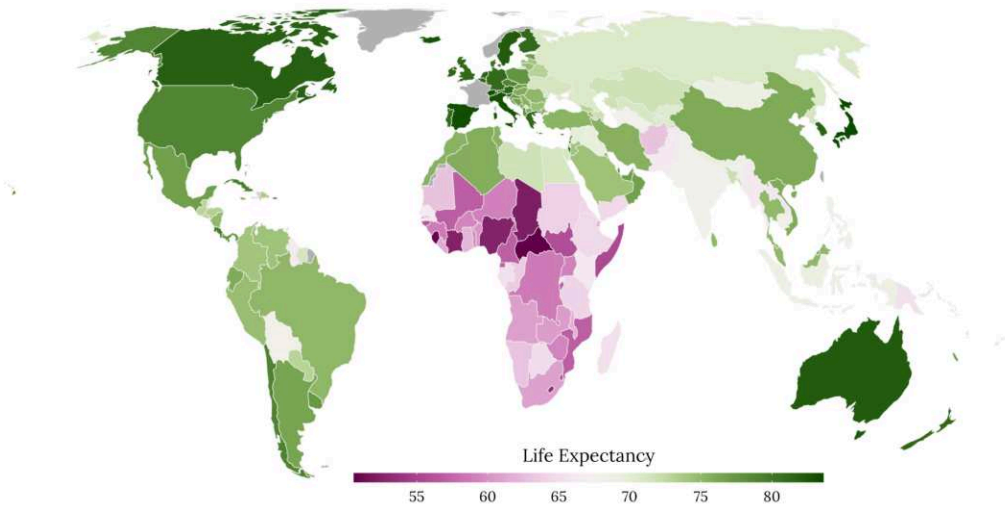
Werte mit bedeutsamen Mittelpunkt
(numerisch oder ordinal)

verwende einen bedeutsamen Mittelpunkt und nutze ausgewogene Extremwerte

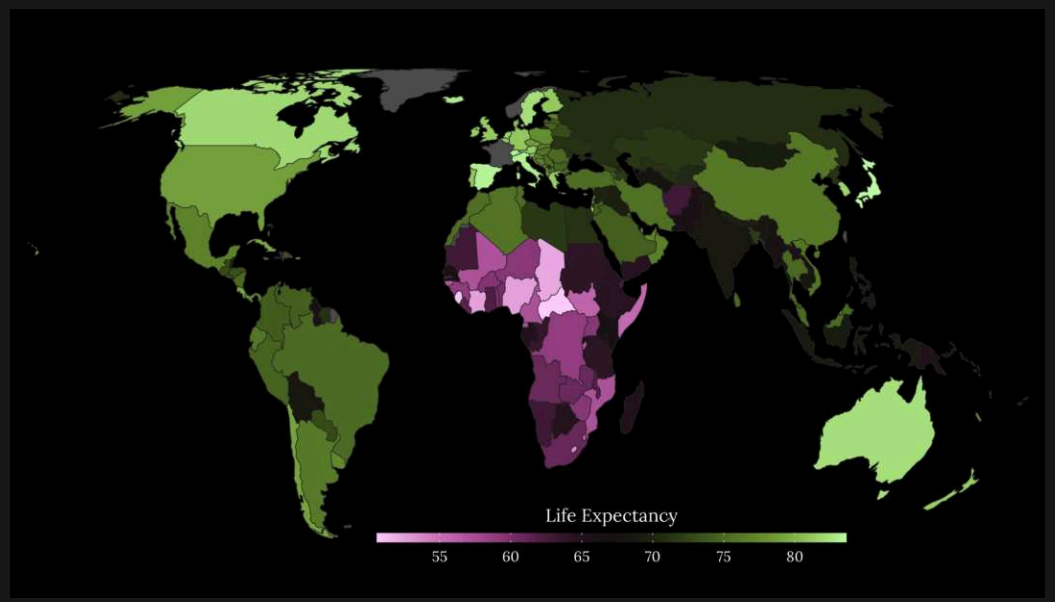
Kombination von zwei sequentiellen Farbpaletten

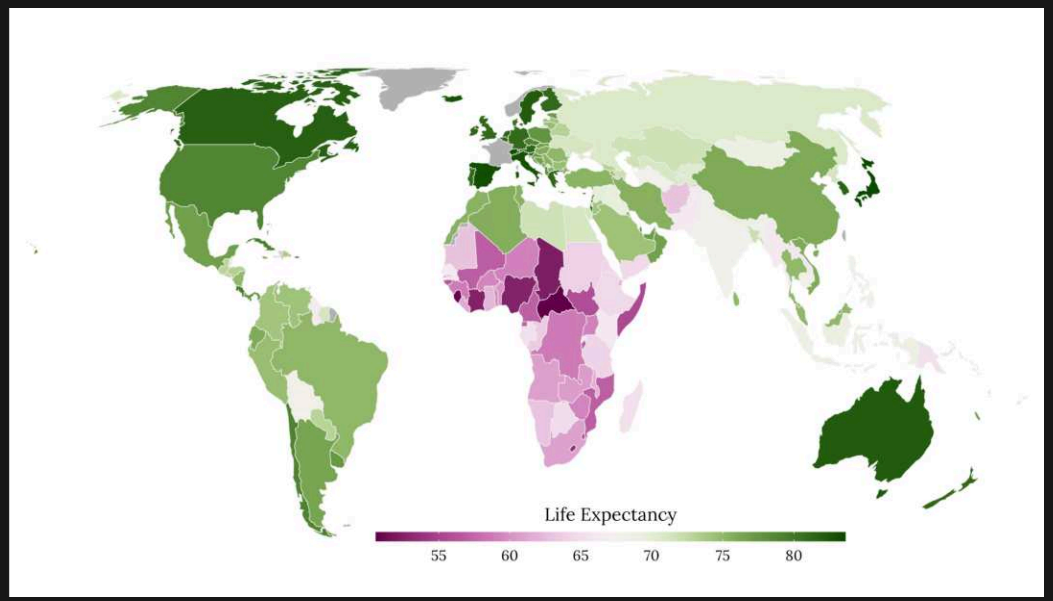
Sequentiell oder divergierend?



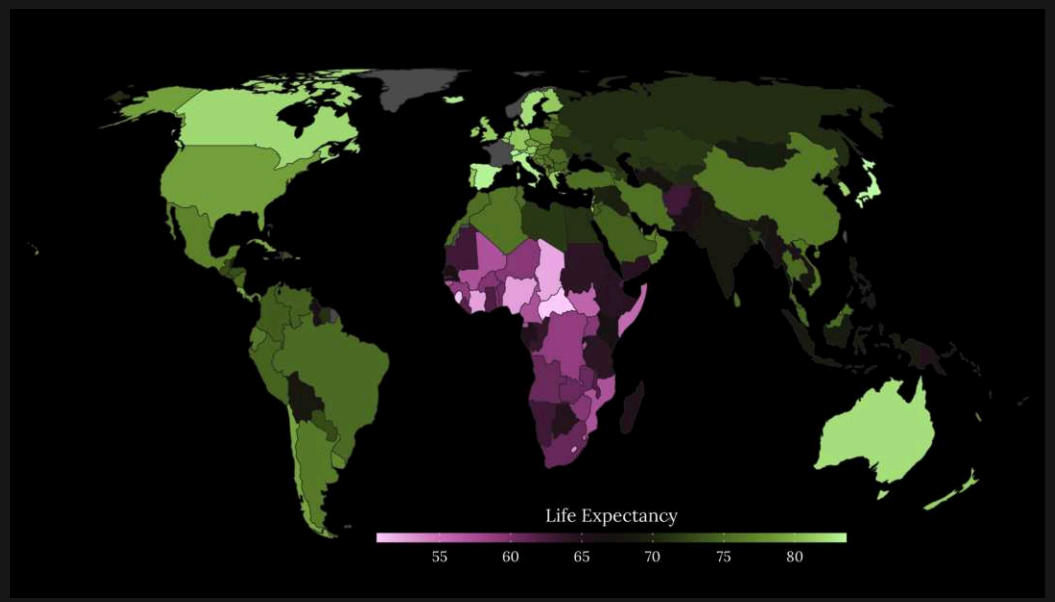


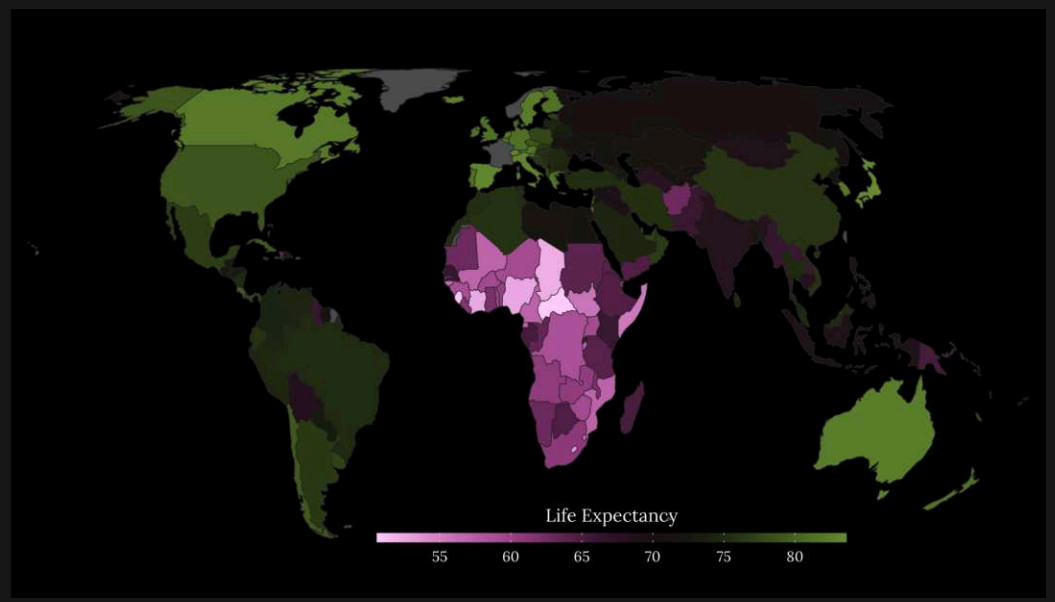
Divergierend: Dark Mode



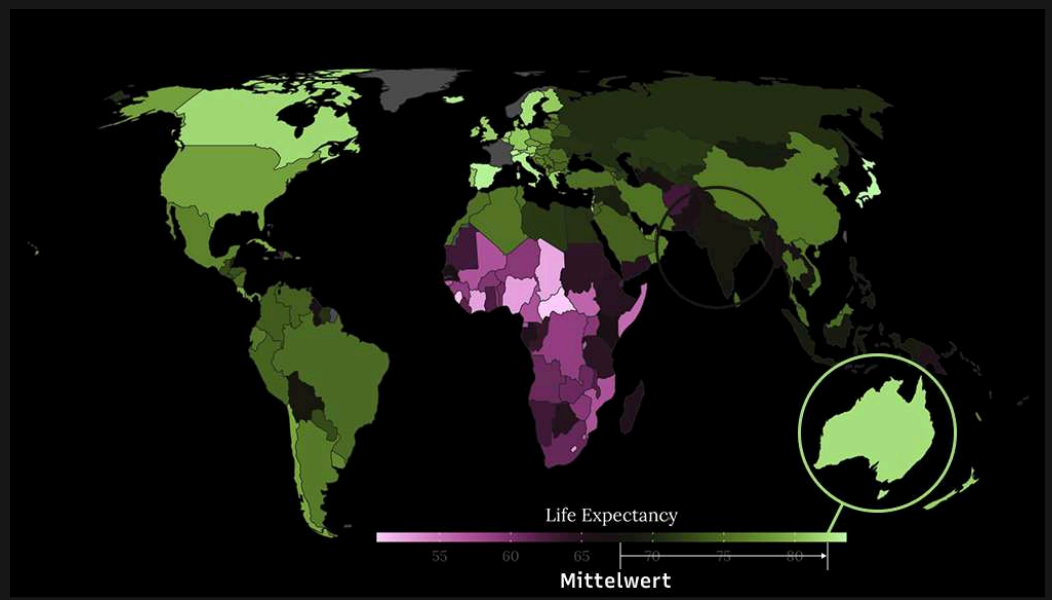


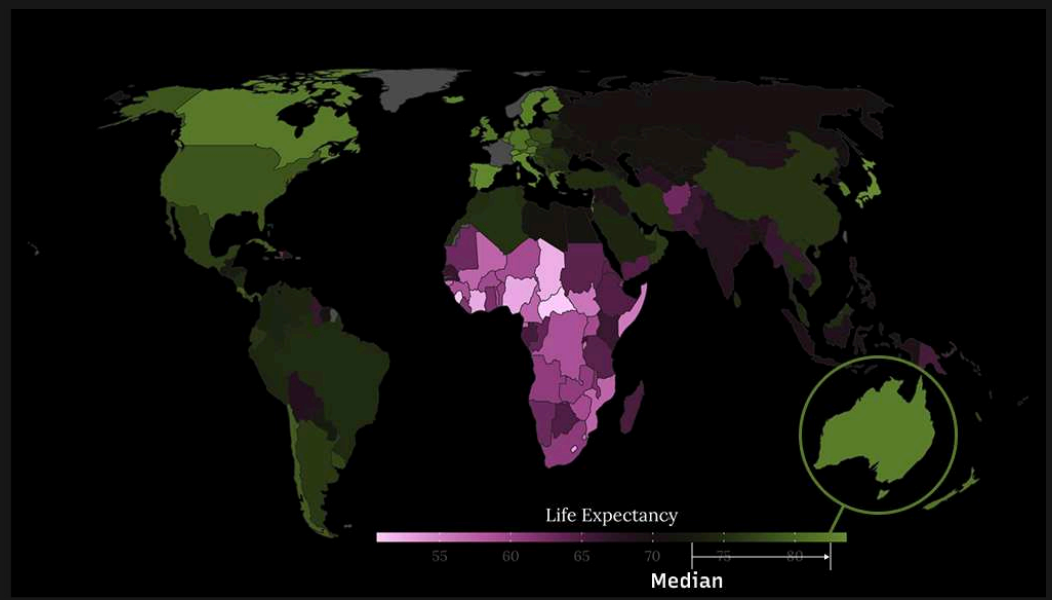
Der "richtige" Mittelpunkt



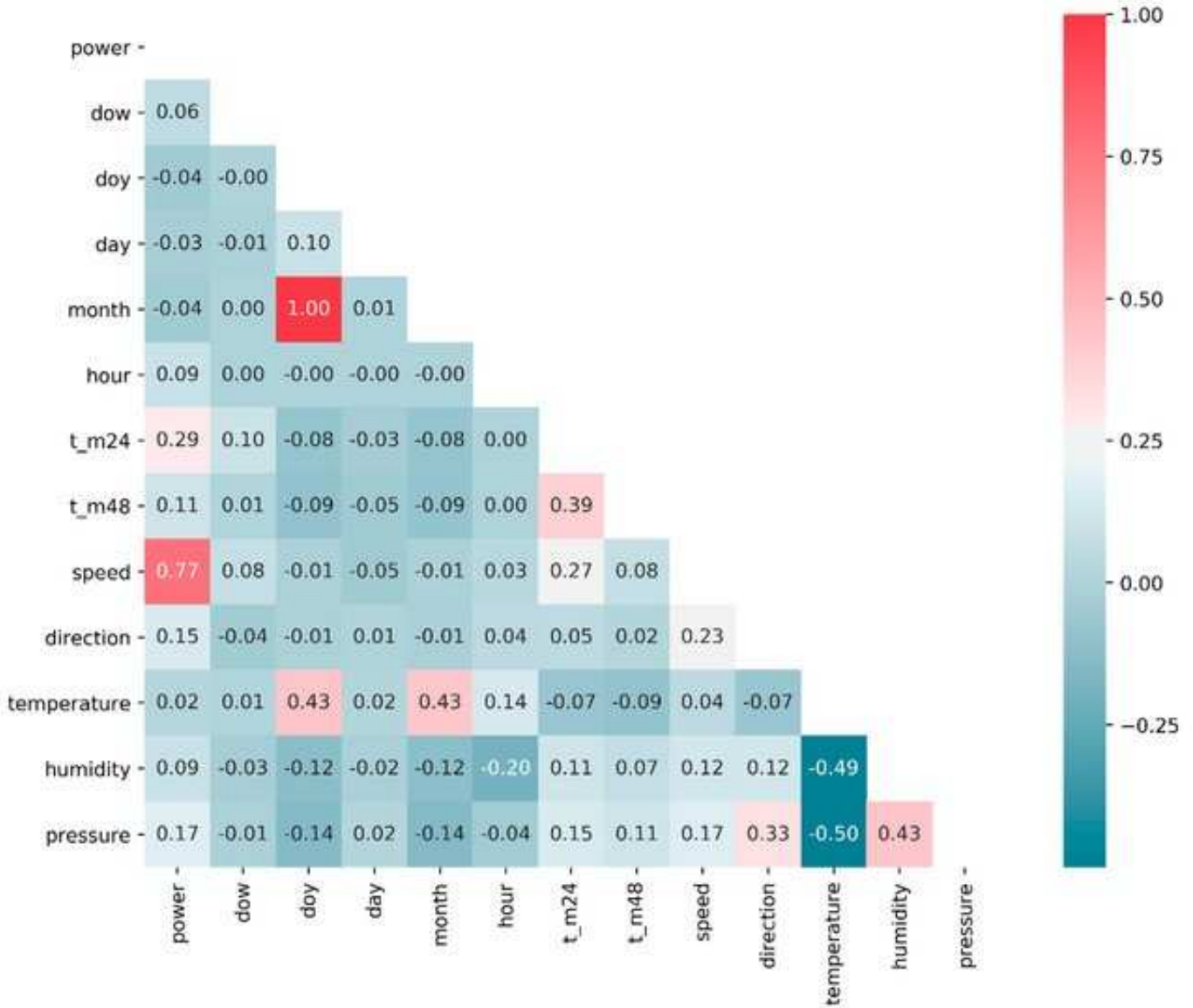


Der "richtige" Mittelpunkt



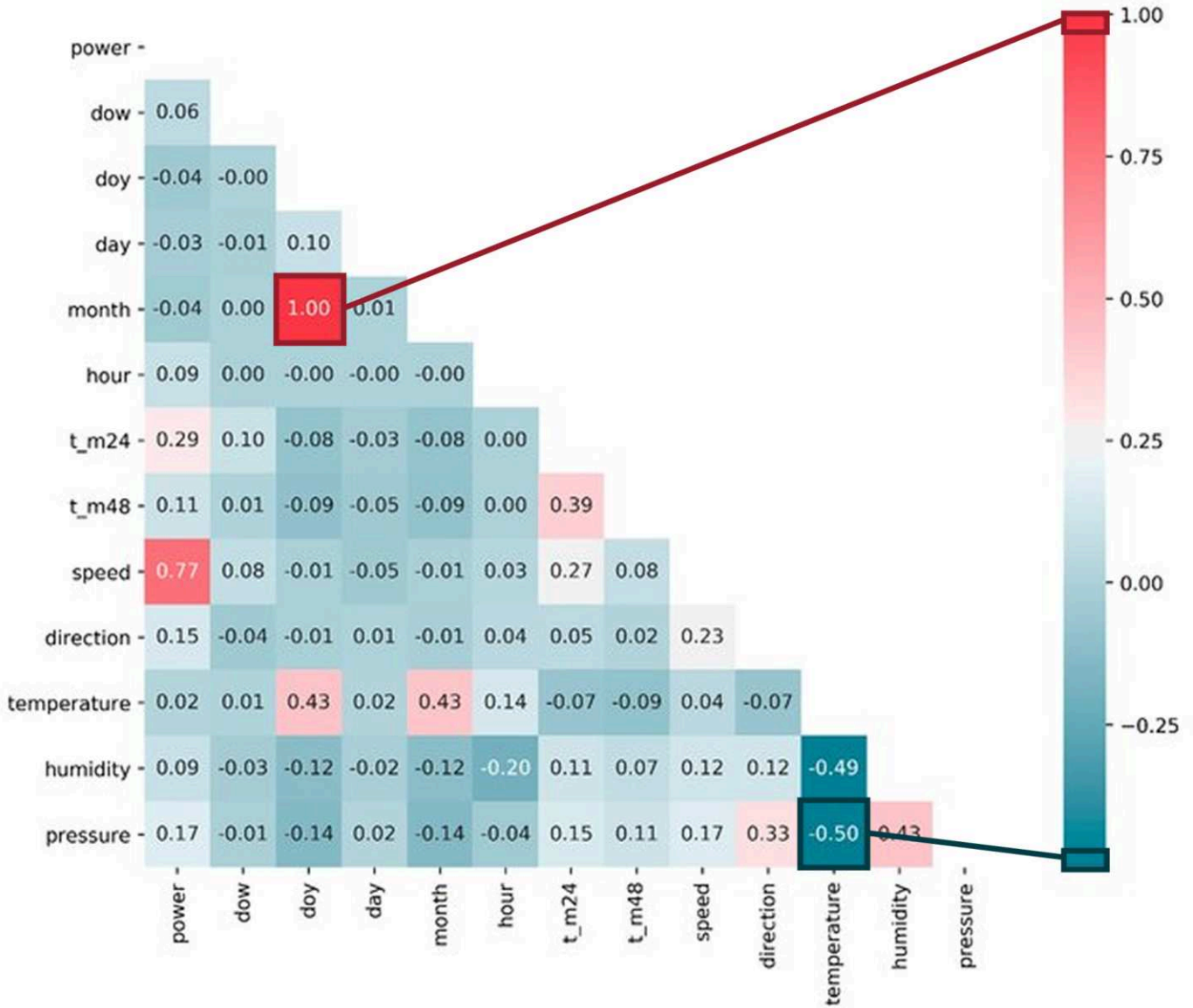


Der falsche Mittelpunkt



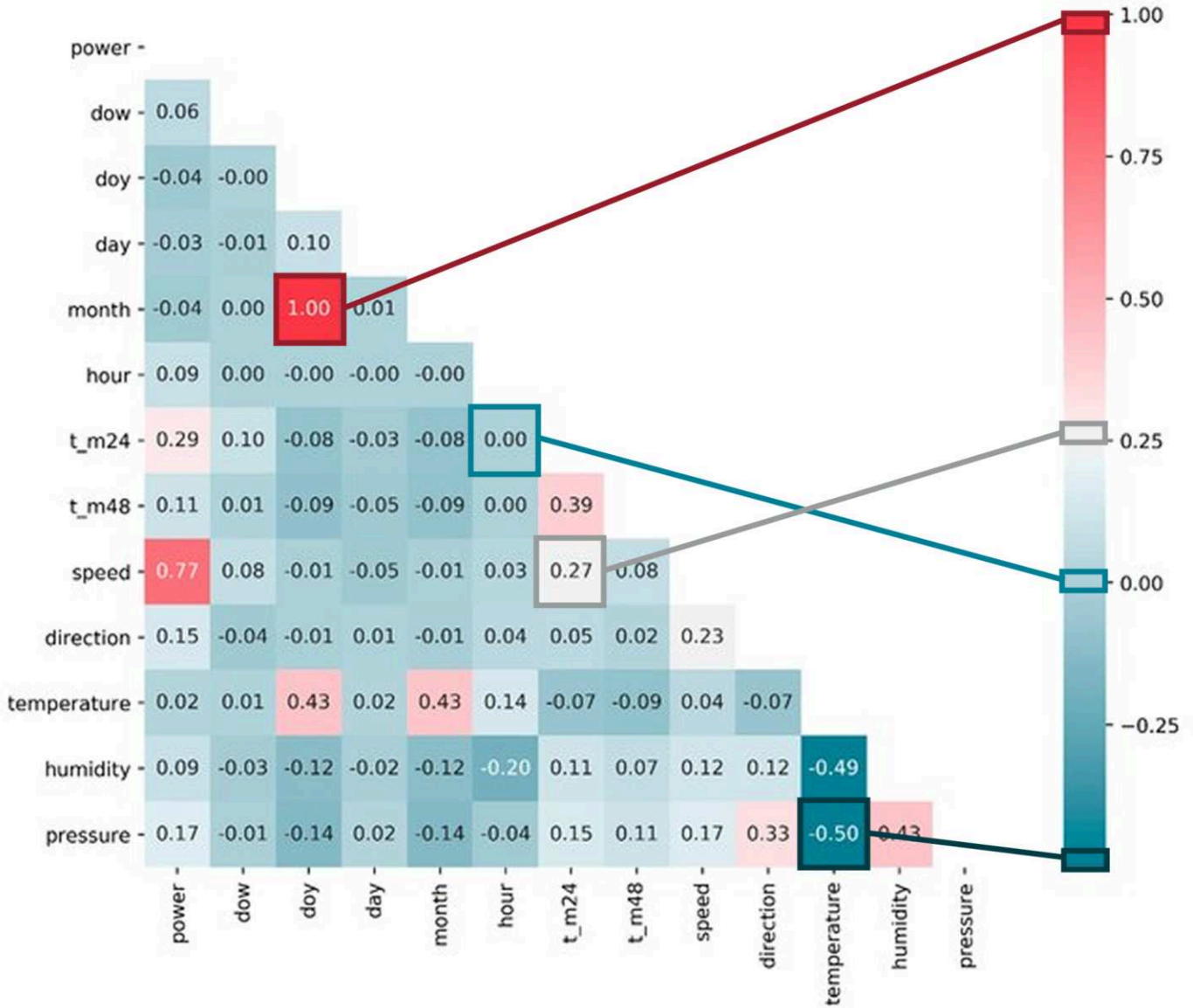
Zheng & Wu (2019) doi: 10.3390/app9153019

Der falsche Mittelpunkt



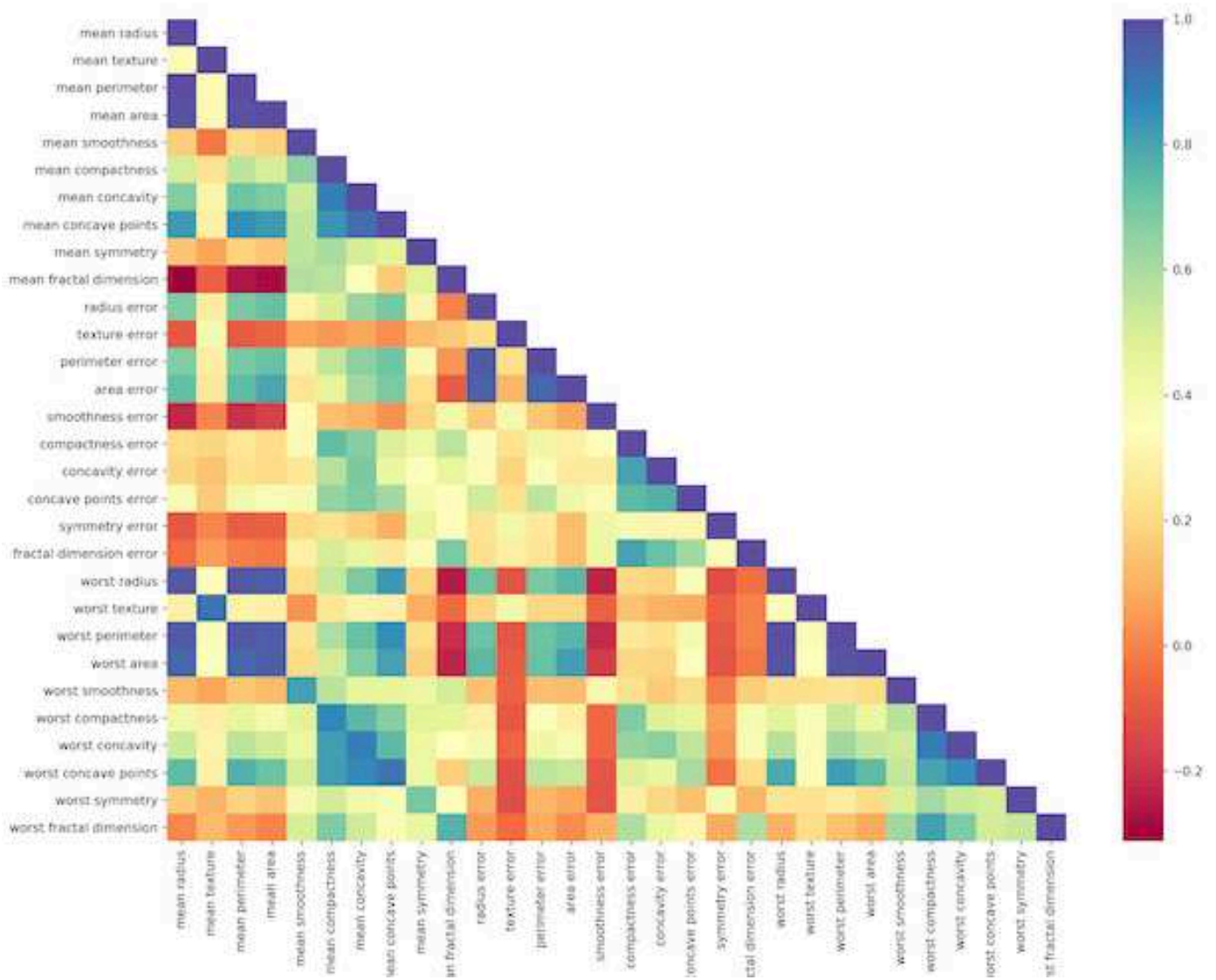
Zheng & Wu (2019) doi: 10.3390/app9153019

Der falsche Mittelpunkt



Zheng & Wu (2019) doi: 10.3390/app9153019

Der falsche Mittelpunkt



Quelle: cmdlinetips.com

Typen von Farbpaletten

Sequentiell

Beispiel 

Grautöne 

Werte mit geordneter Skala
(numerisch oder ordinal)

verwende den höchsten Kontrast für die wichtigsten Informationen

entweder einfarbige oder mehrfarbige Farbpaletten

Divergierend

Beispiel 

Grautöne 

Werte mit bedeutsamen Mittelpunkt
(numerisch oder ordinal)

verwende einen bedeutsamen Mittelpunkt und nutze ausgewogene Extremwerte

Kombination von zwei sequentiellen Farbpaletten

Qualitativ

Beispiel 

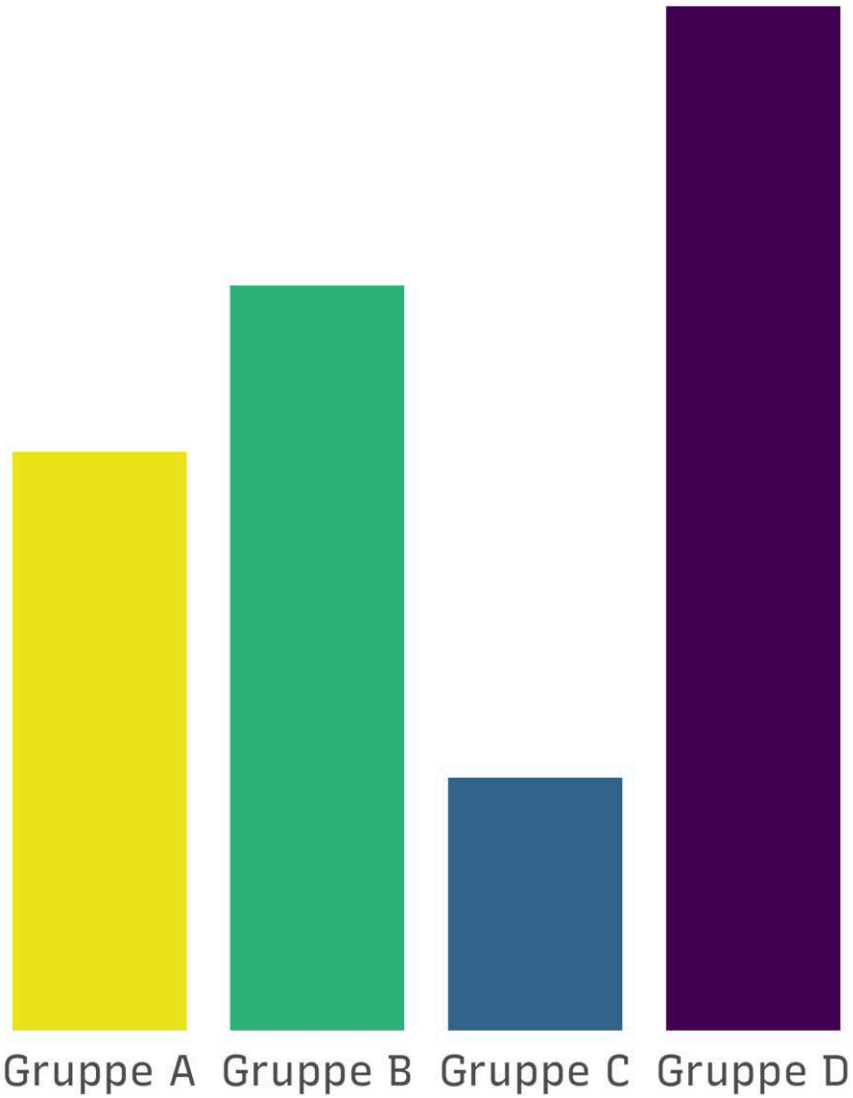
Grautöne 

Gleichwertige Gruppen
(kategoriell)

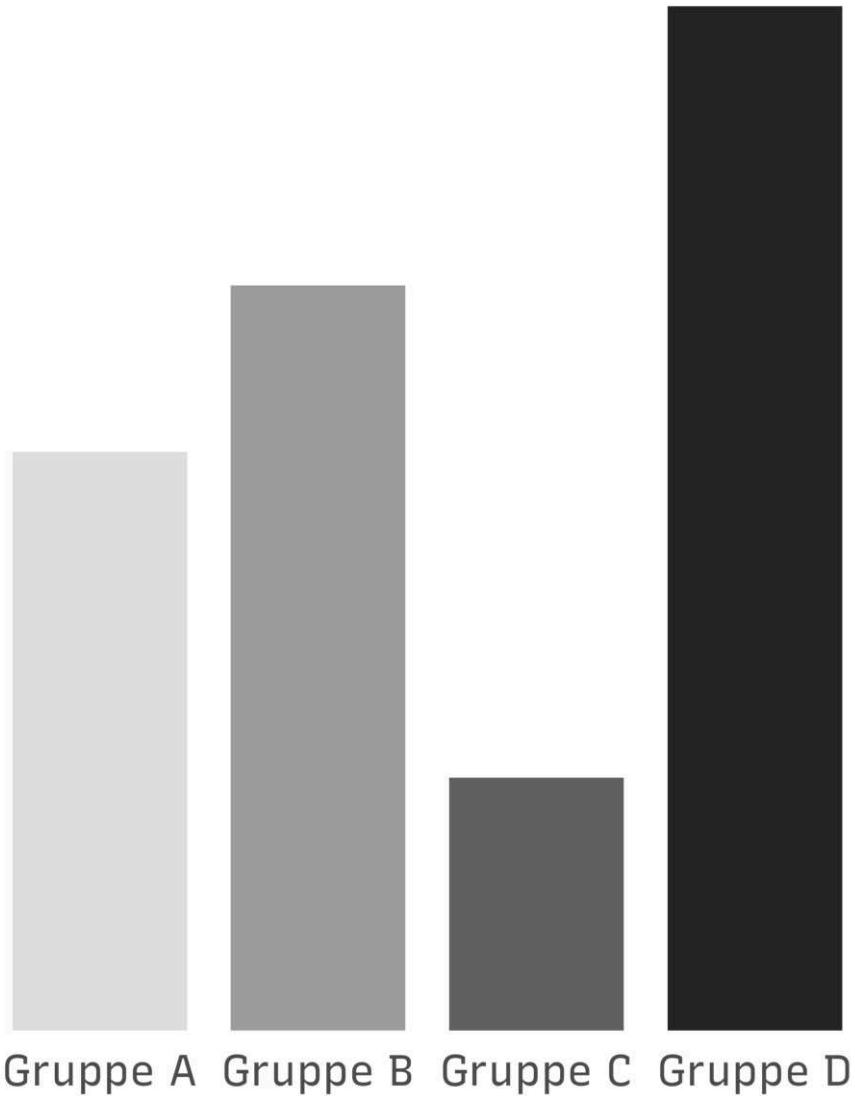
verwende verschiedene Farben mit ähnlichem "Wahrnehmungsgewicht"

Anzahl auf 5 bis 8 Kategorien begrenzen

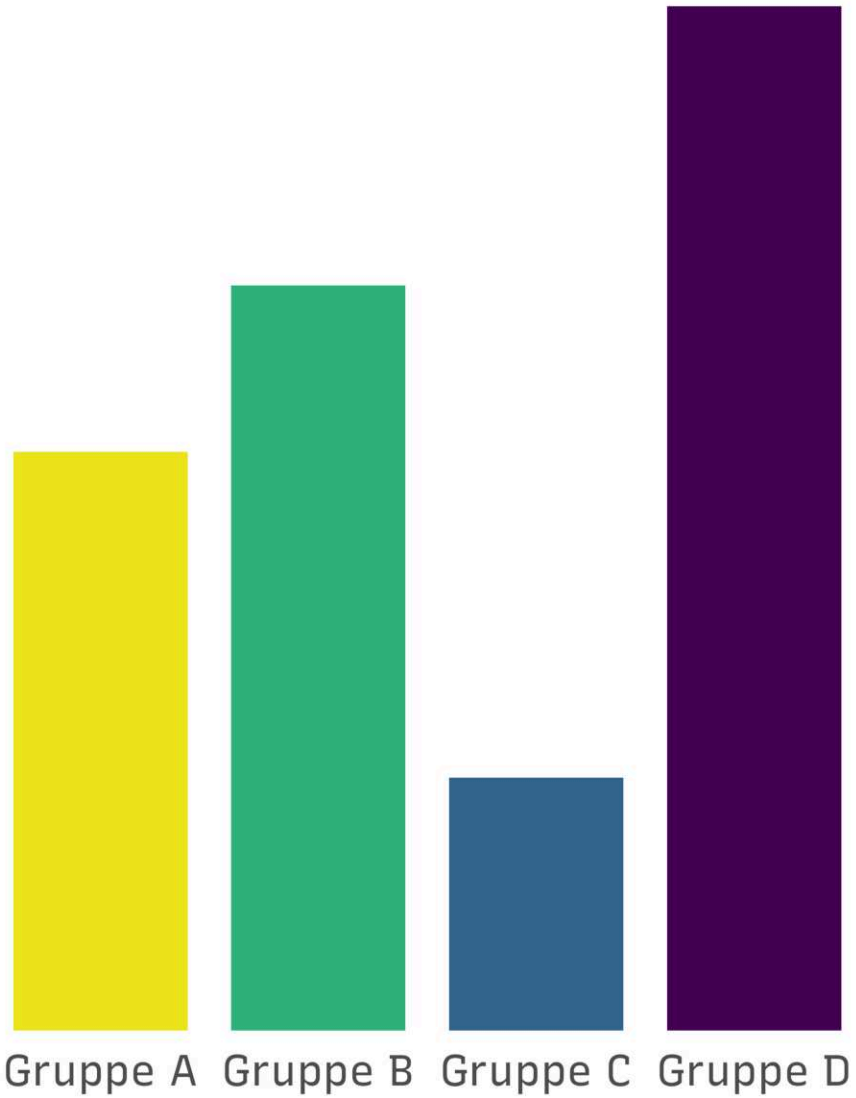
Viridis



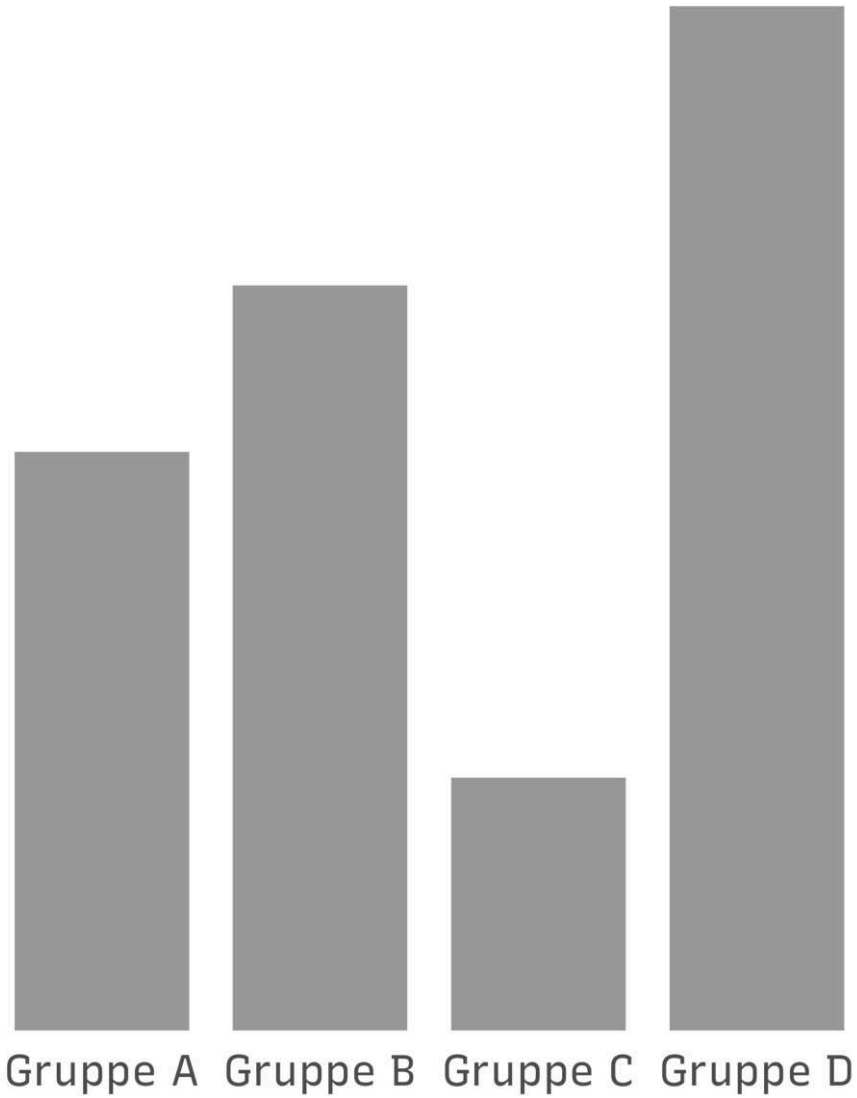
… in S/W

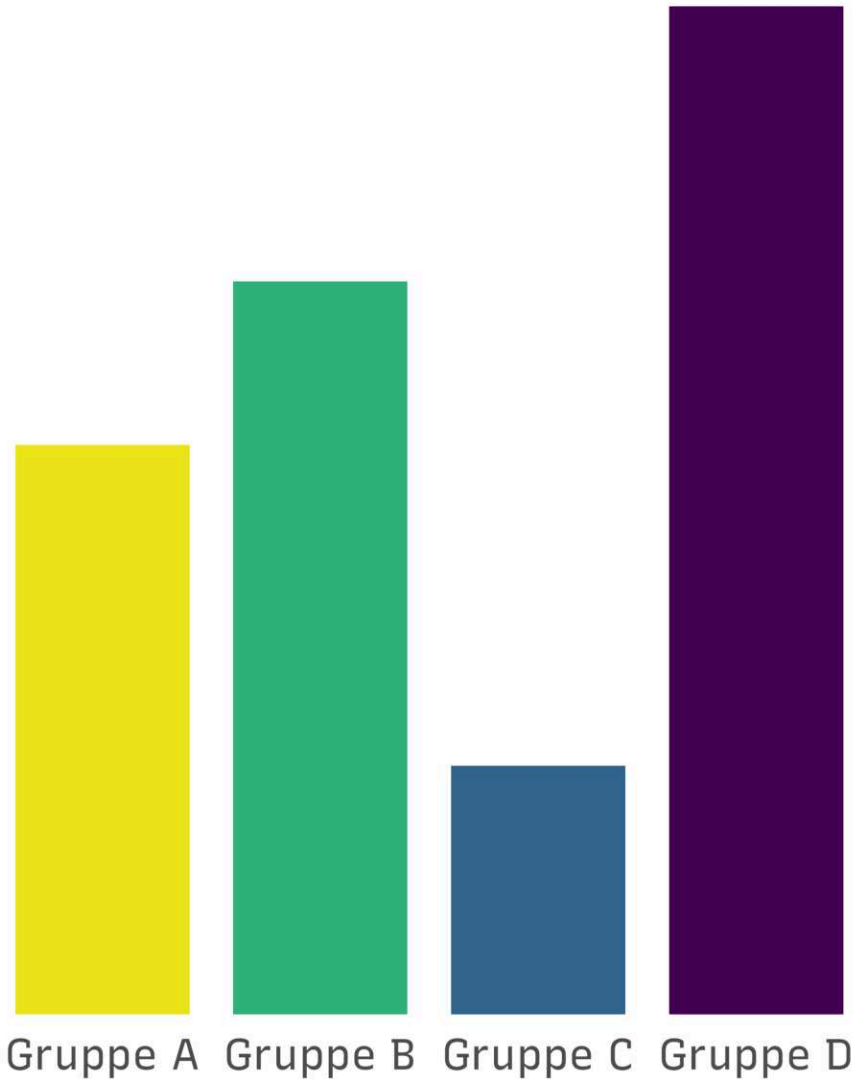


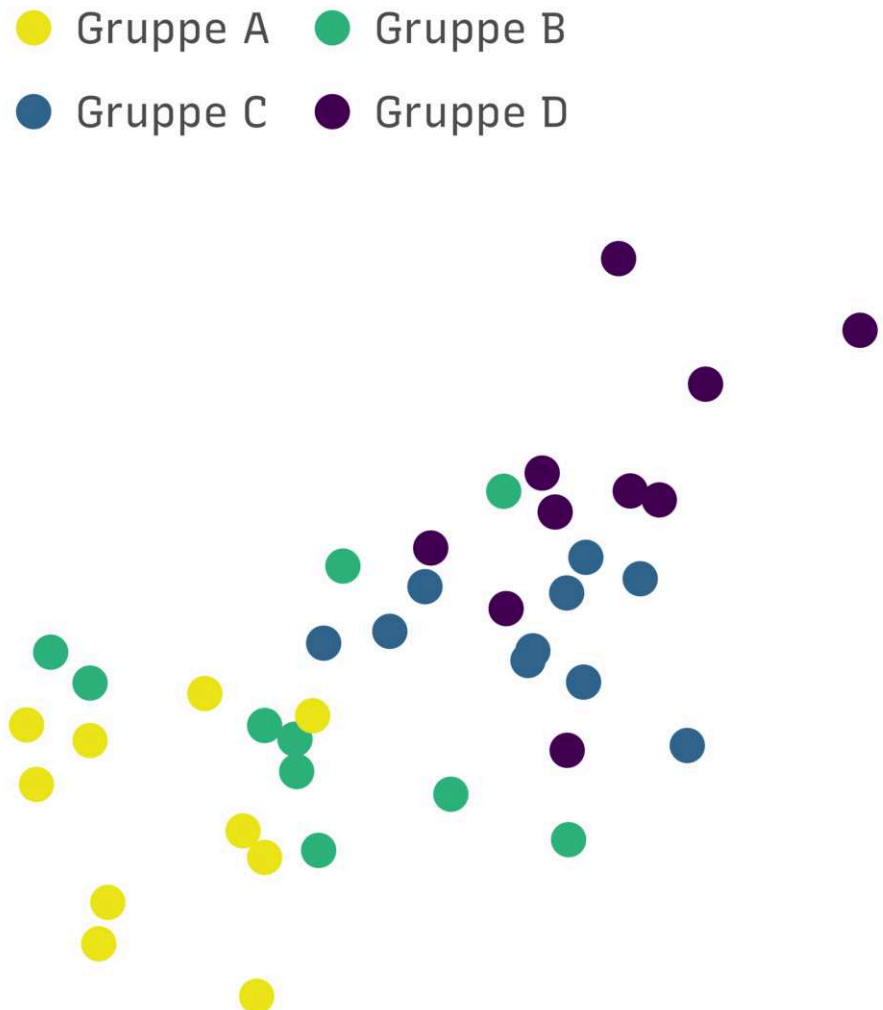
❌ Diskret

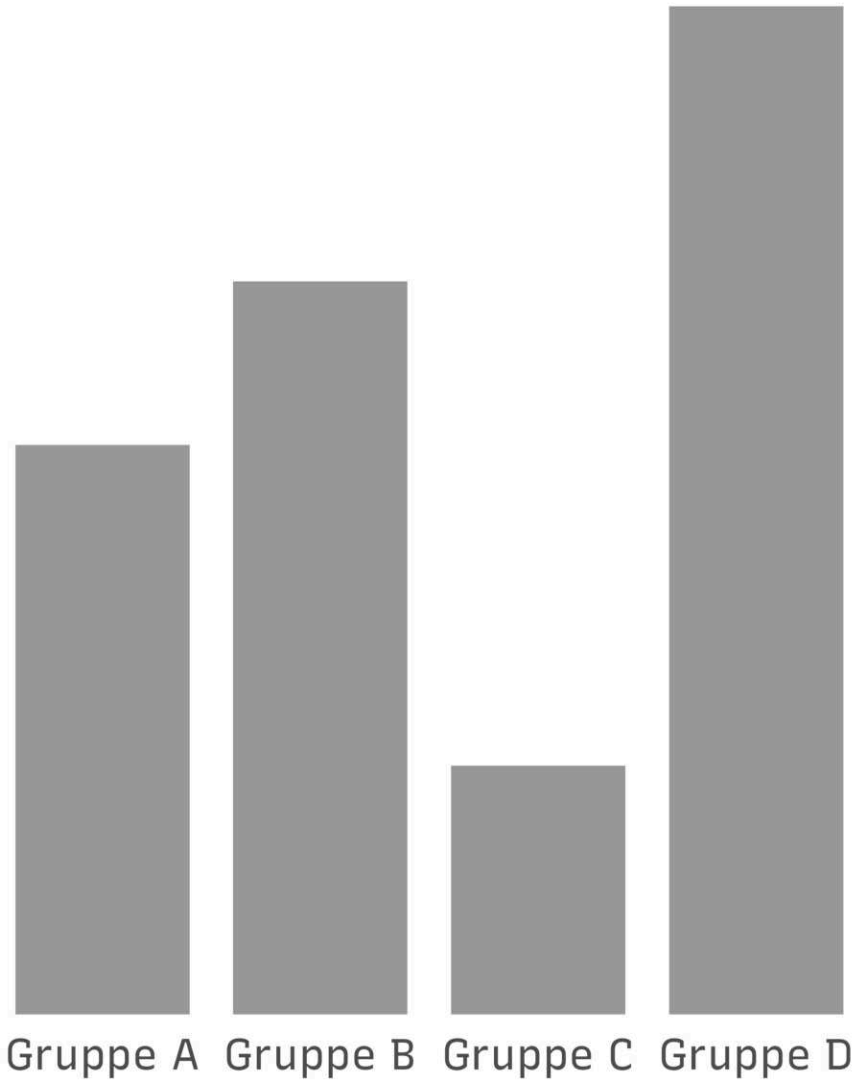


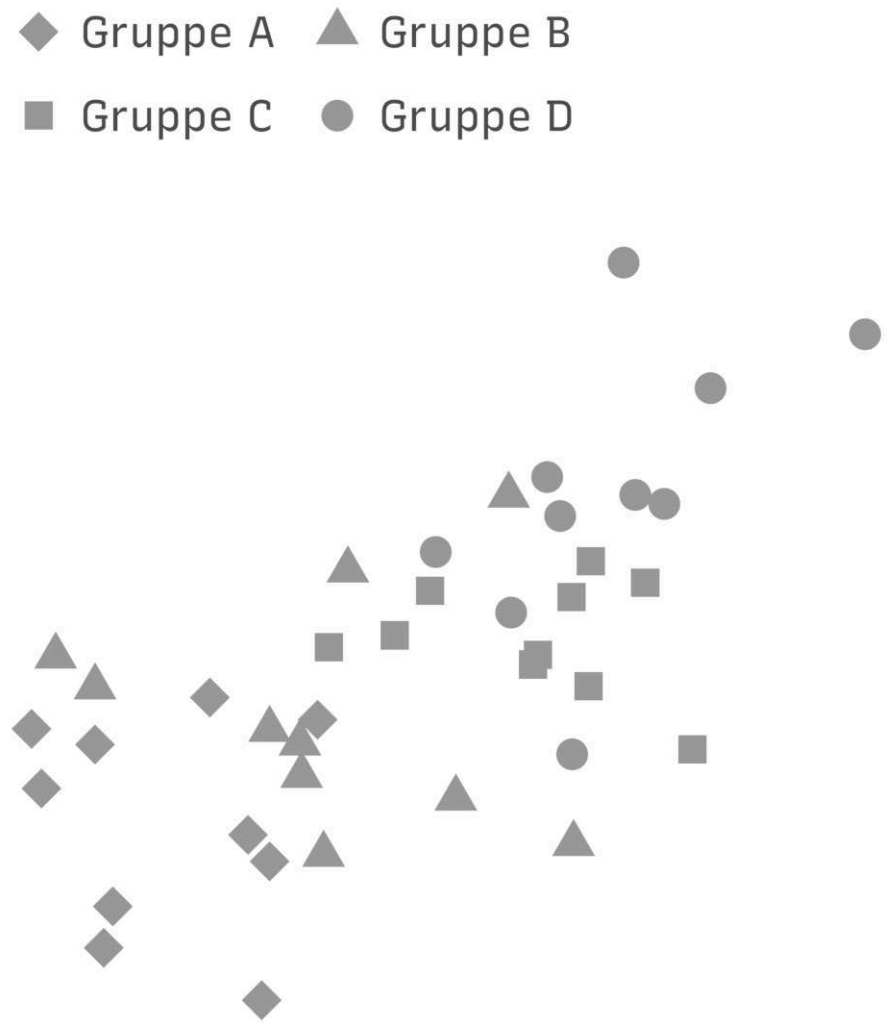
✅ keine Farben

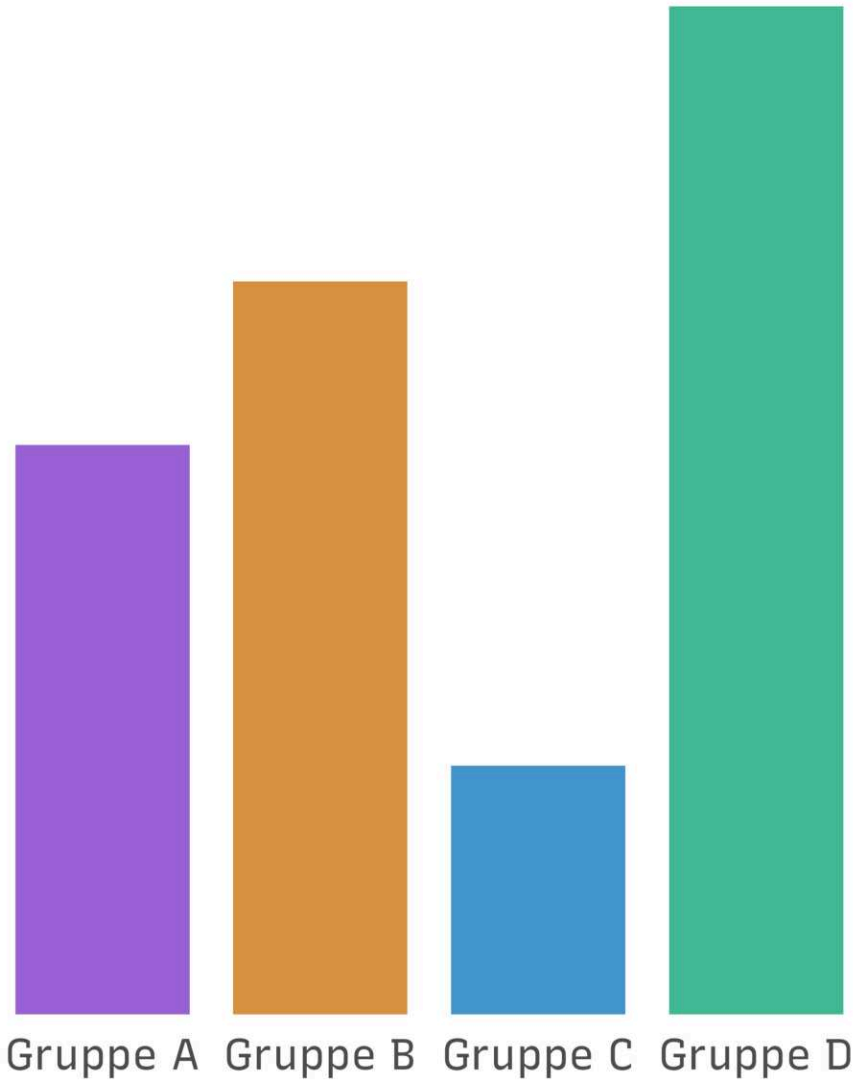


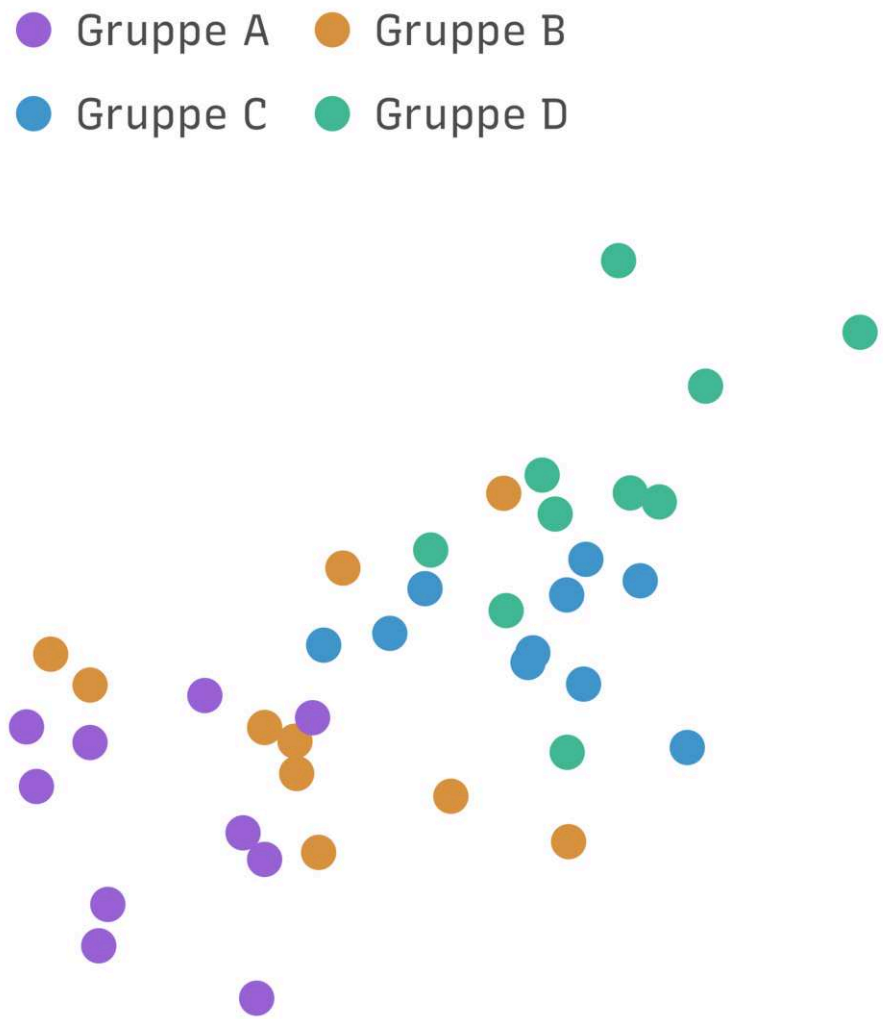

Gruppe A
Gruppe B
Gruppe C
Gruppe D

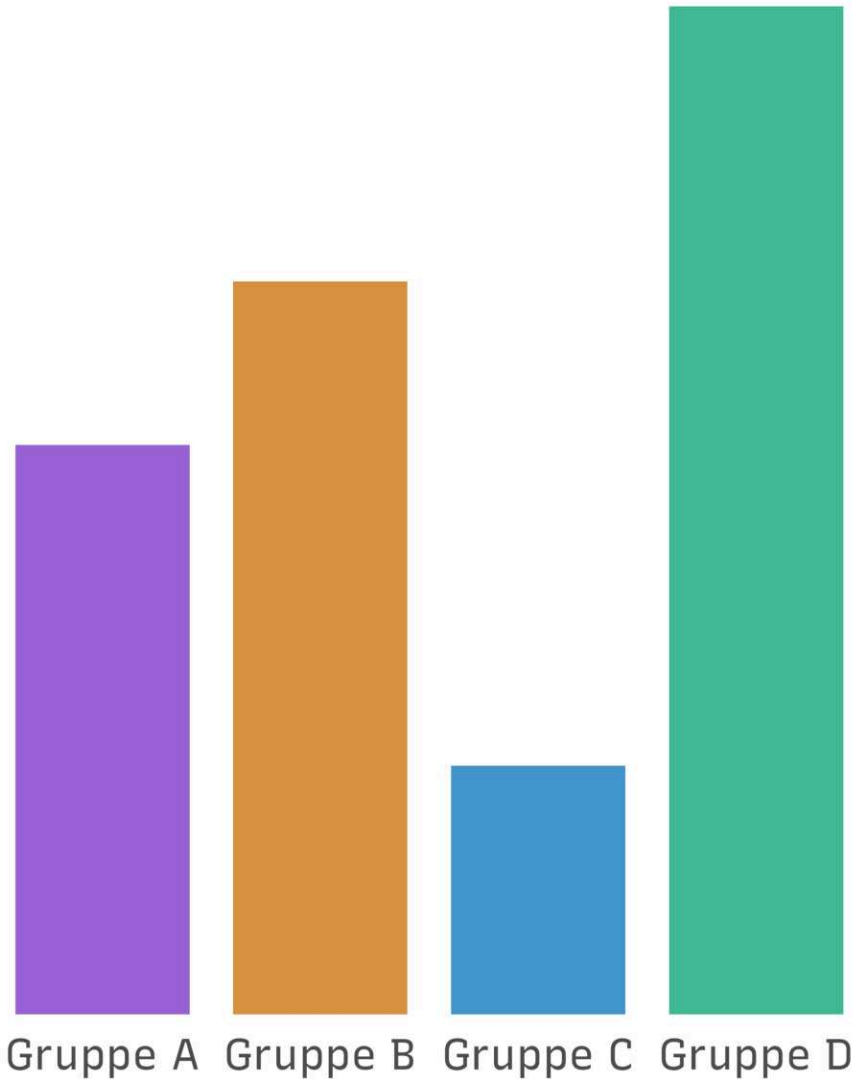

Gruppe A
Gruppe B
Gruppe C
Gruppe D

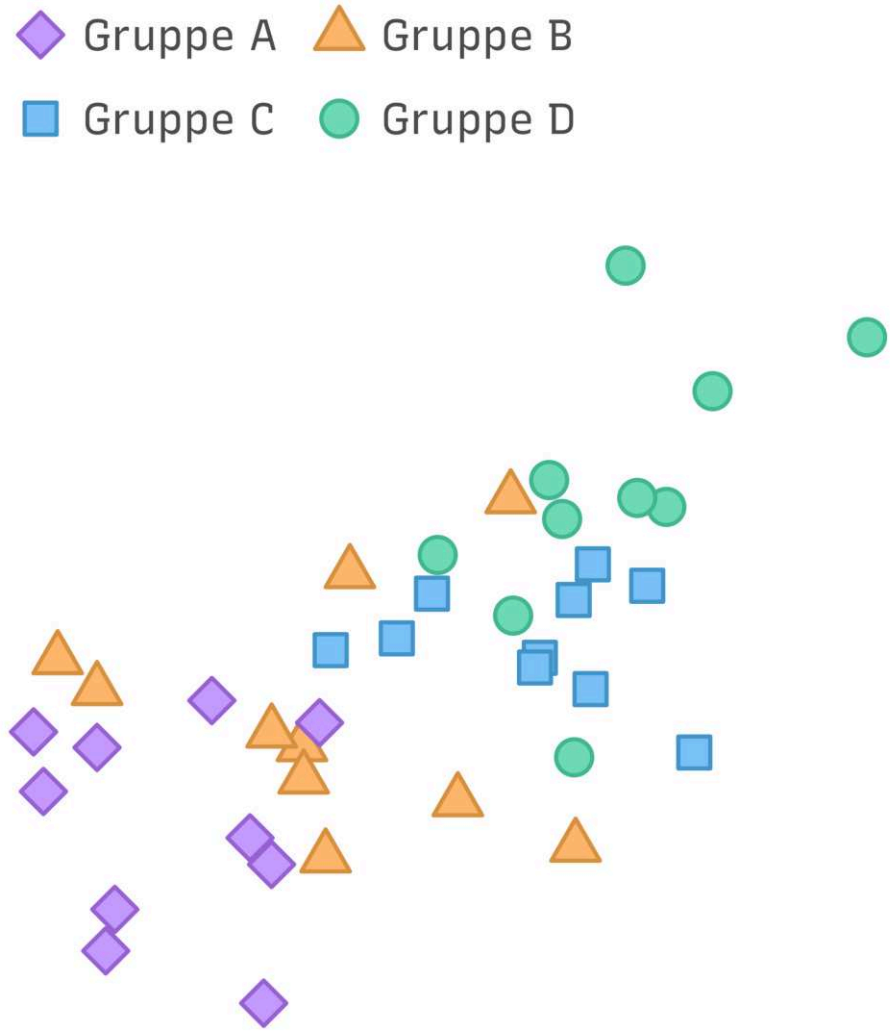

Gruppe A
Gruppe B
Gruppe C
Gruppe D


Gruppe A
Gruppe B
Gruppe C
Gruppe D


Gruppe A
Gruppe B
Gruppe C
Gruppe D


Gruppe A
Gruppe B
Gruppe C
Gruppe D


Gruppe A
Gruppe B
Gruppe C
Gruppe D


Gruppe A
Gruppe B
Gruppe C
Gruppe D

Farbwahl

VIZ PALETTE

By: Elijah Meeks & Susie Lu

PICK

Use Chroma.js

Use Colorgorical

Use ColorBrewer

Add | Replace

EDIT

6 Colors

Add

#hex | rgb | hsl

1 #1dabe6
2 #1c366a
3 #c3ced0
4 #e43034
5 #fc4e51
6 #af060f

GET

#hex | rgb | hsl

String quotes
Object with metadata

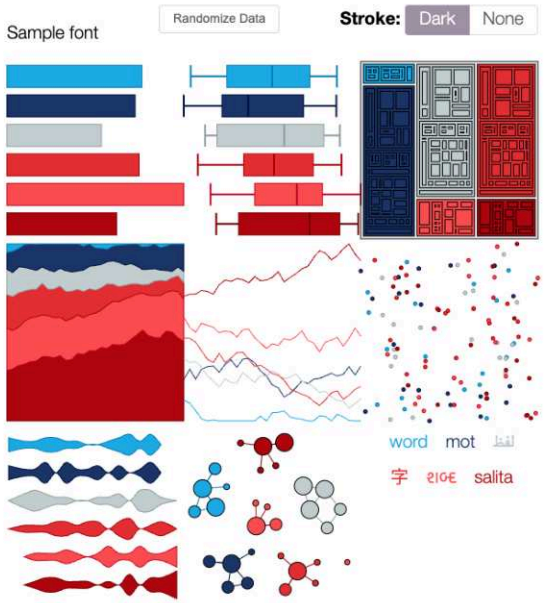
```
["#1dabe6",
"#1c366a",
"#c3ced0",
"#e43034",
"#fc4e51",
"#af060f"]
```

COLORS IN ACTION

Background color: #ffffff
Font color: #000000

Charts made with Semiotic

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



COLOR REPORT

projects.susielu.com/viz-palette

Konsistent

Quelle: Lisa Charlotte Muth, DataWrapper Blog

Intuitiv

Quelle: Lisa Charlotte Muth, DataWrapper Blog

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

PMCID: PMC5148121 | PMID: 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[a)]

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

Authors

David Borland, University of North Carolina at Chapel Hill

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

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

Education and communication

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

Michael Stoelzle[1] and Lina Stein[2]

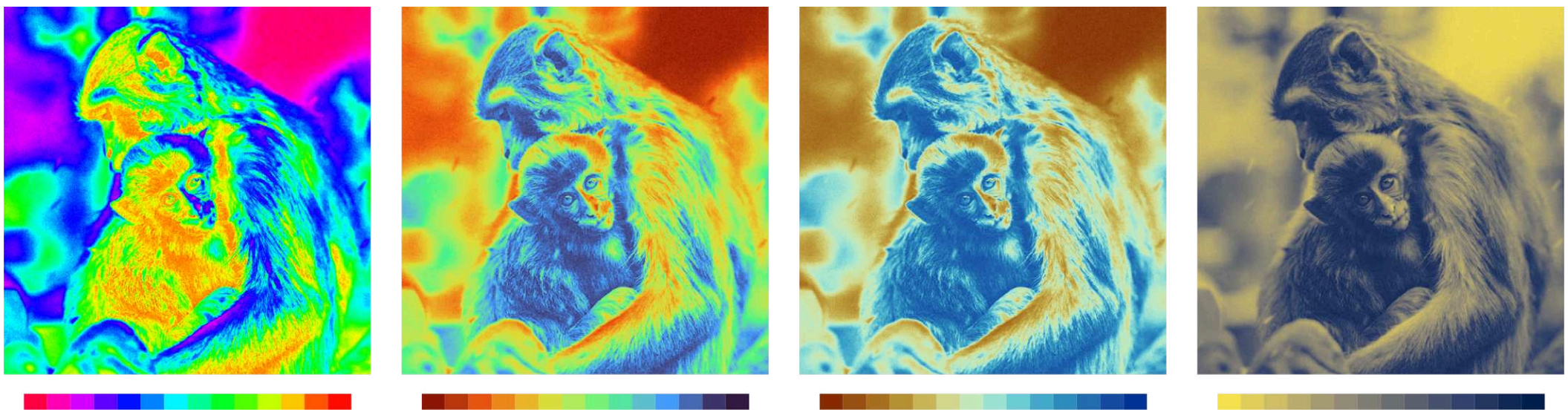
[1]Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

[2]Department of Civil Engineering, University of Bristol, Bristol, UK

Non Uniform Distances

Cédric Scherer Data Visualization & Information Design

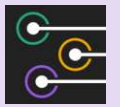
Hinter dem Regenbogen



Inspiriert von Fabio Crameri | Original Fotografie von Richard Strozynski

Cédric Scherer Data Visualization & Information Design

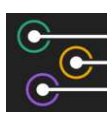
Übungsteil



Übung

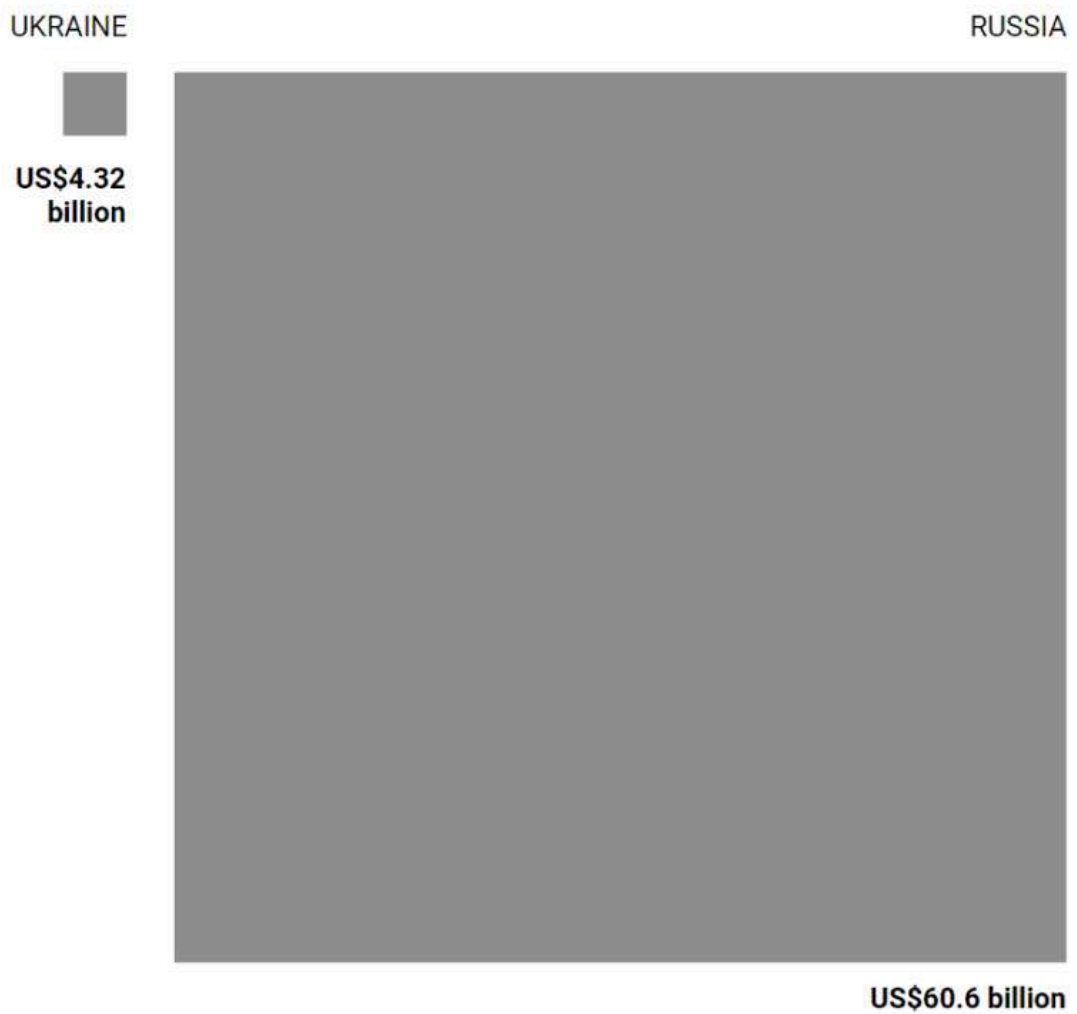
Schaue dir die folgenden Datenvisualisierungen an.

- Wieso sind die Kodierungen problematisch?
- Wie kann man die Probleme beheben?

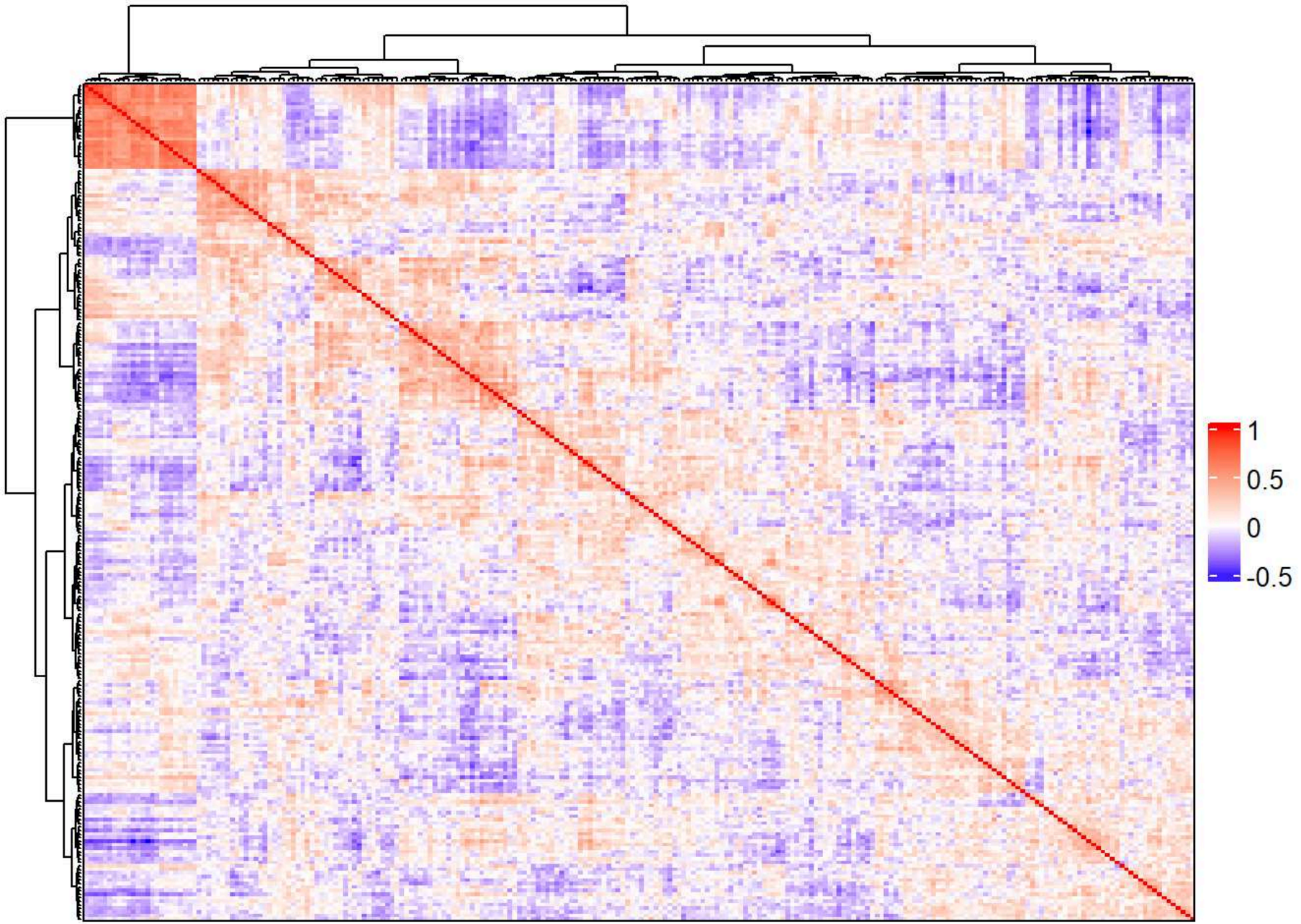


DEFENCE BUDGETS: RUSSIA VS UKRAINE (2020)

The national balance of forces is overwhelmingly in Russia's favour. Russian military spending in 2020 amounted to US$60.6 billion in 2020. Ukraine's was less than a 10th of that amount.

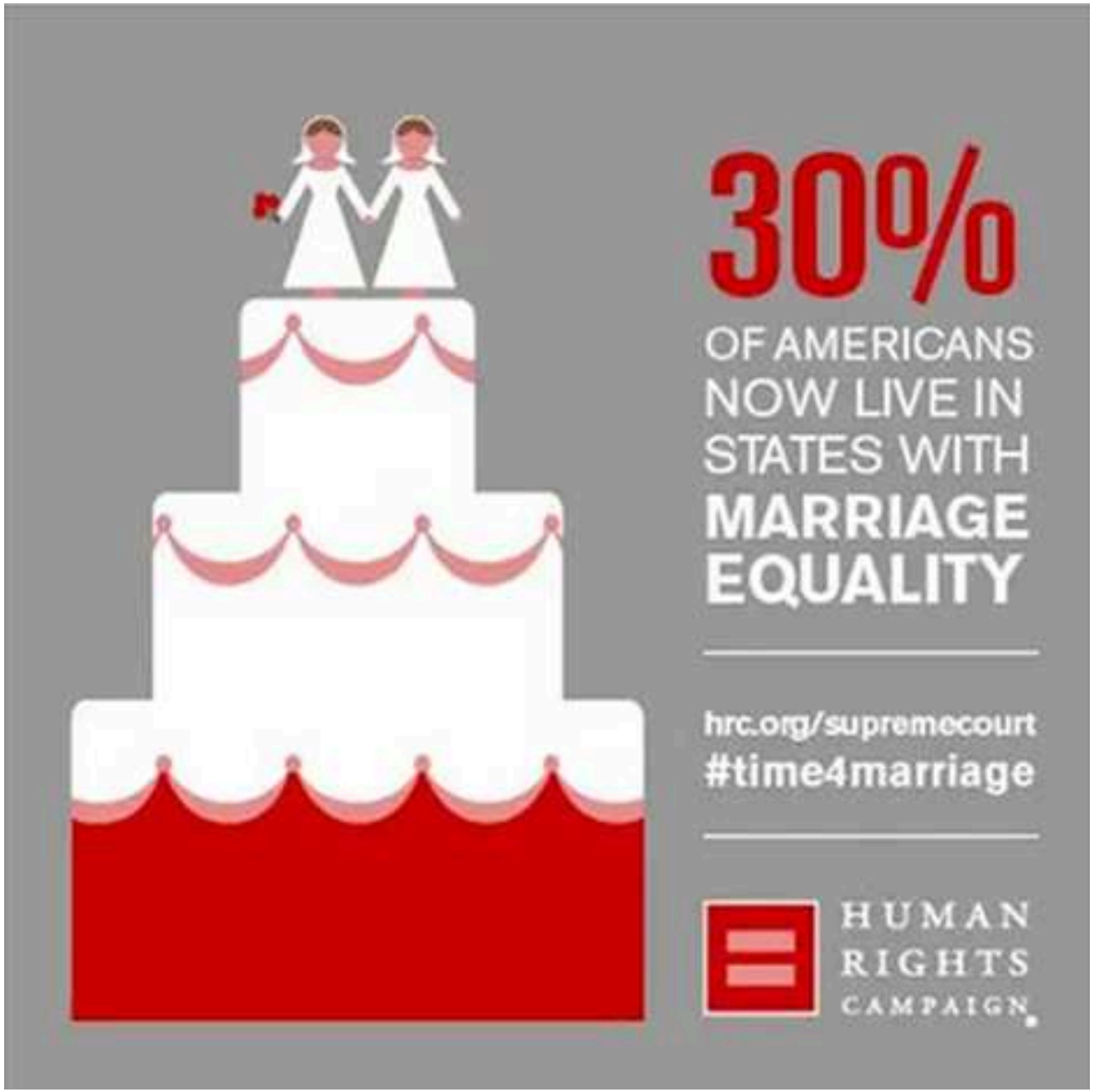


"Russia attacks Ukraine" von SCMP Graphic (South China Morning Post)



Quelle: "Proteomics Data Analysis in R/Bioconductor"

Quelle: Fox News

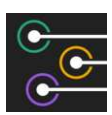


Grafik: Human Rights Campaign

Übung

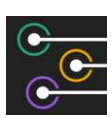
Reflektiere und recherchiere:

- Begutachte und bewerte eigene frühere Visualisierungen hinsichtlich der eingesetzten Kodierungen.
- Finde jeweils ein Beispiel (vorzugsweise aus der wissenschaftlichen Literatur), das eine besonders problematische Größen- oder Farbkodierung nutzt.

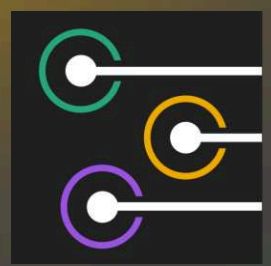


Projektarbeit

- Notiere für deine Variablen die Datentypen.
  (numerisch – absolut / Interval / Verhältnis – ordinal, kategoriell)
- Liste mögliche visuelle Kodierungen für jede Variable auf.
- Bewerte die visuelle Wirksamkeit dieser Kodierungen.
- Unterstreiche Kodierungen, die dein Ziel sinnvoll unterstützen und betonen.
- Prüfe, ob es Einschränkungen bei den gewählten Kodierungen gibt.
- Wie kannst du diese gestalterisch oder datenbasiert lösen?
- Überlege: Welche Kodierungen möchtest du gezielt vermeiden – und warum?



Dankeschön!



CÉDRIC SCHERER

Data Visualization & Information Design

www.cedricscherer.com