

Engaging Data Visualizations for Graphical Storytelling

Principles of Data Visualization Design

Dr Cédric Scherer

Fannie Mae | September 2024



Data Visualization

is any graphical representation
of information and data.



Data Visualization

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



Data Visualization

converts information into visual forms
as quantifiable features.



Data Visualization

is part art and part science.



Typology of Information Graphics

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

→ Type of information:

depict information schematically <> convert information into visual forms

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

→ Purpose of the graphic:

facilitate discovery <> communicate information



TYPOLOGY OF INFORMATION GRAPHICS

Type of information →

Conceptual

Purely conceptual information cannot be visualized, only illustrated.

Purpose of the graphic ↓

Explanatory

Measurable

Exploratory

Visualization

Infographics

Information graphics

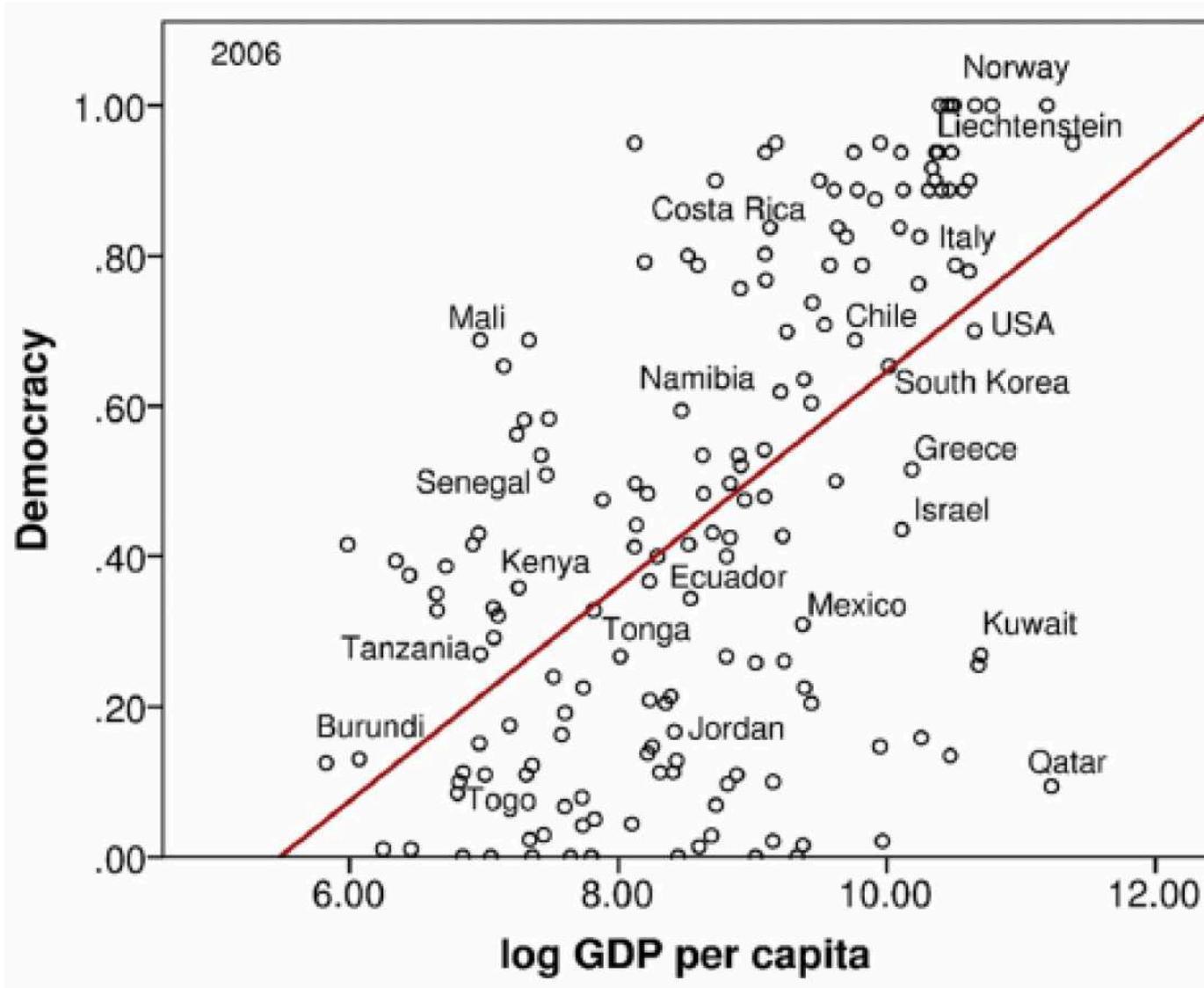
Source: Koponen & Hildén, "Data Visualization Handbook" (2020), page 25





Source: eazybi

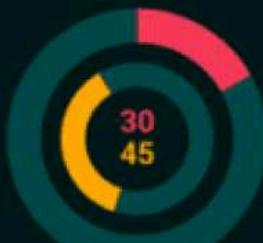




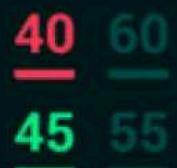
Source: Ranganathan et al. 2014

Cédric Scherer Data Visualization & Information Design





*Lorum ipsum dolor sit amet,
consectetur adipiscing elit.
Integer ante elit.*

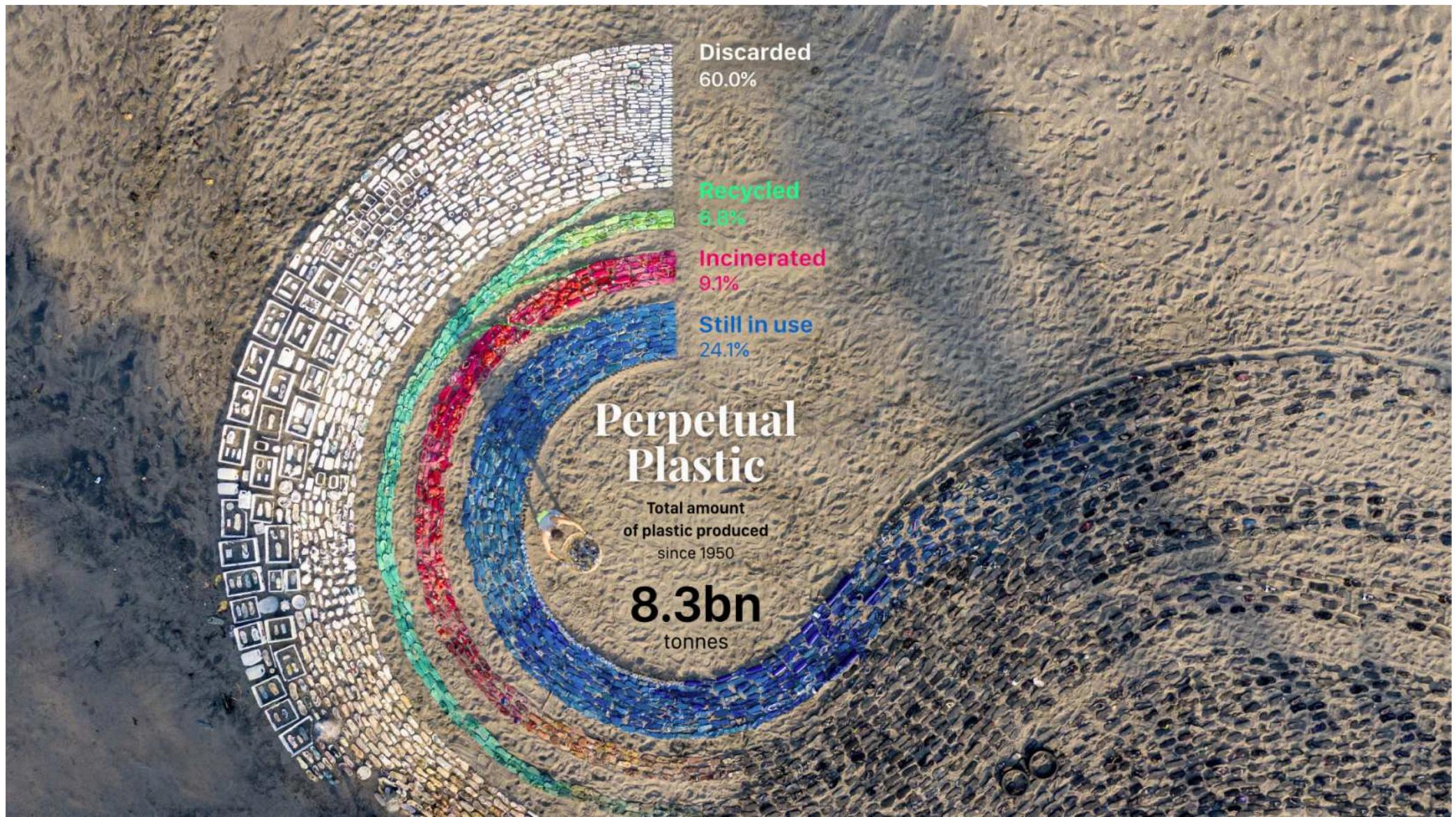


*Lorum ipsum dolor sit amet,
consectetur adipiscing elit.
Integer ante elit.*



Source: datameer.com





Source: "Perpetual Plastic" by Liina Klauss, Skye Morét & Moritz Stefaner

Cédric Scherer Data Visualization & Information Design





Source: "Patchwork Kingdoms" by Nadieh Bremer

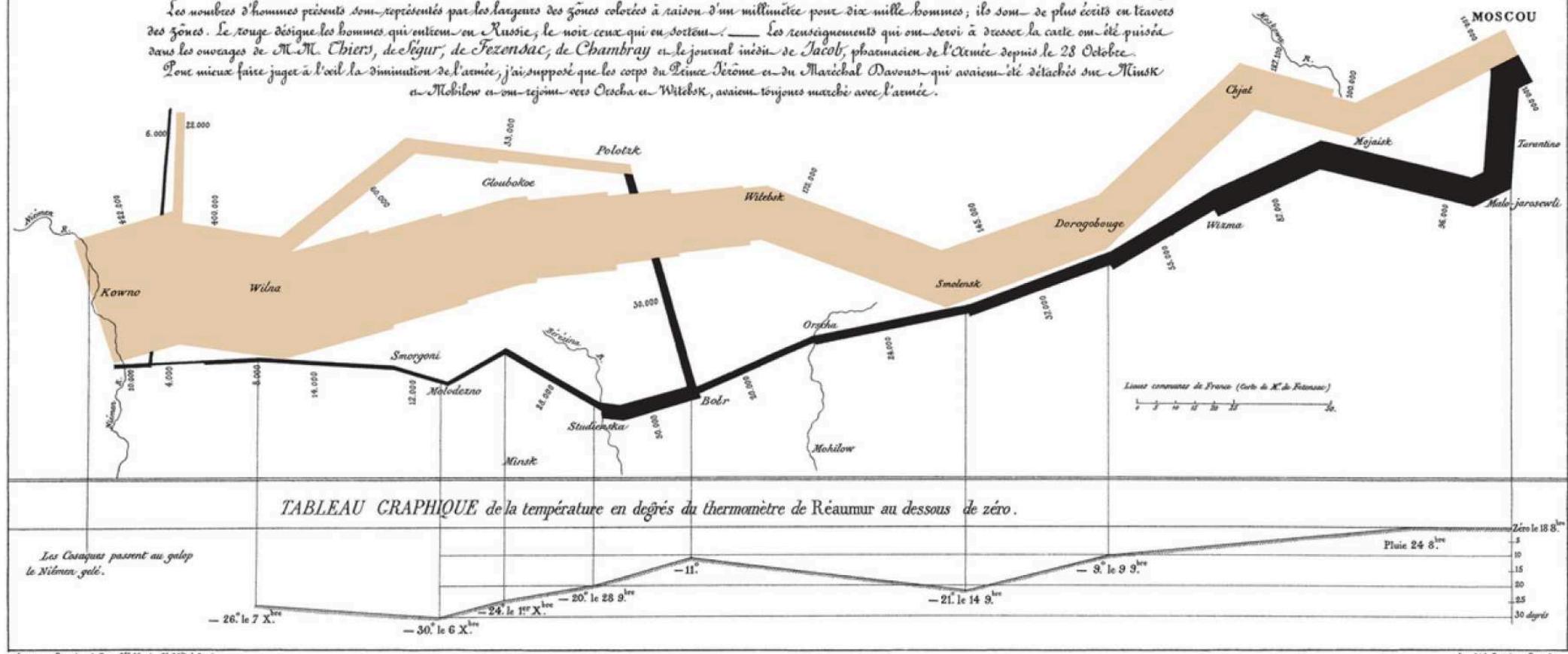


Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

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

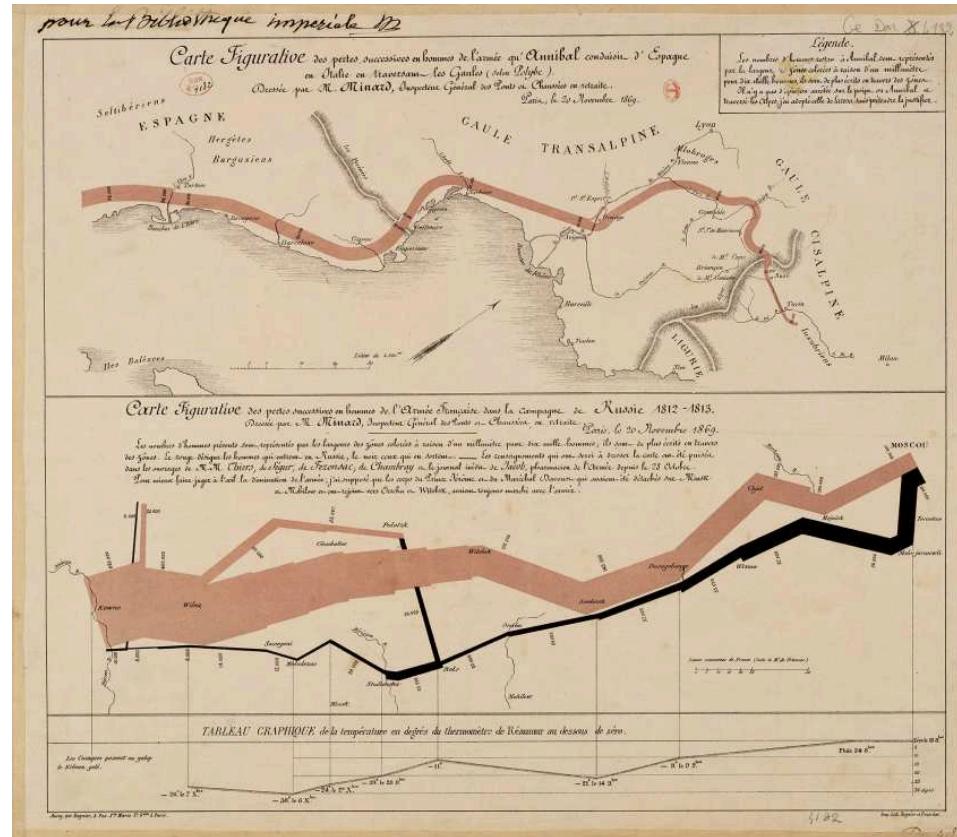
Les nombres d'hommes perdus sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres de zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été pris dans les ouvrages de M. M. Chier, de Segur, de Fezensac, de Chambray 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 Nérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mogilow et qui rejoignirent Ossaka en Wilensk, avaient toujours marché avec l'armée.



"Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813" by Charles Joseph Minard (1869)

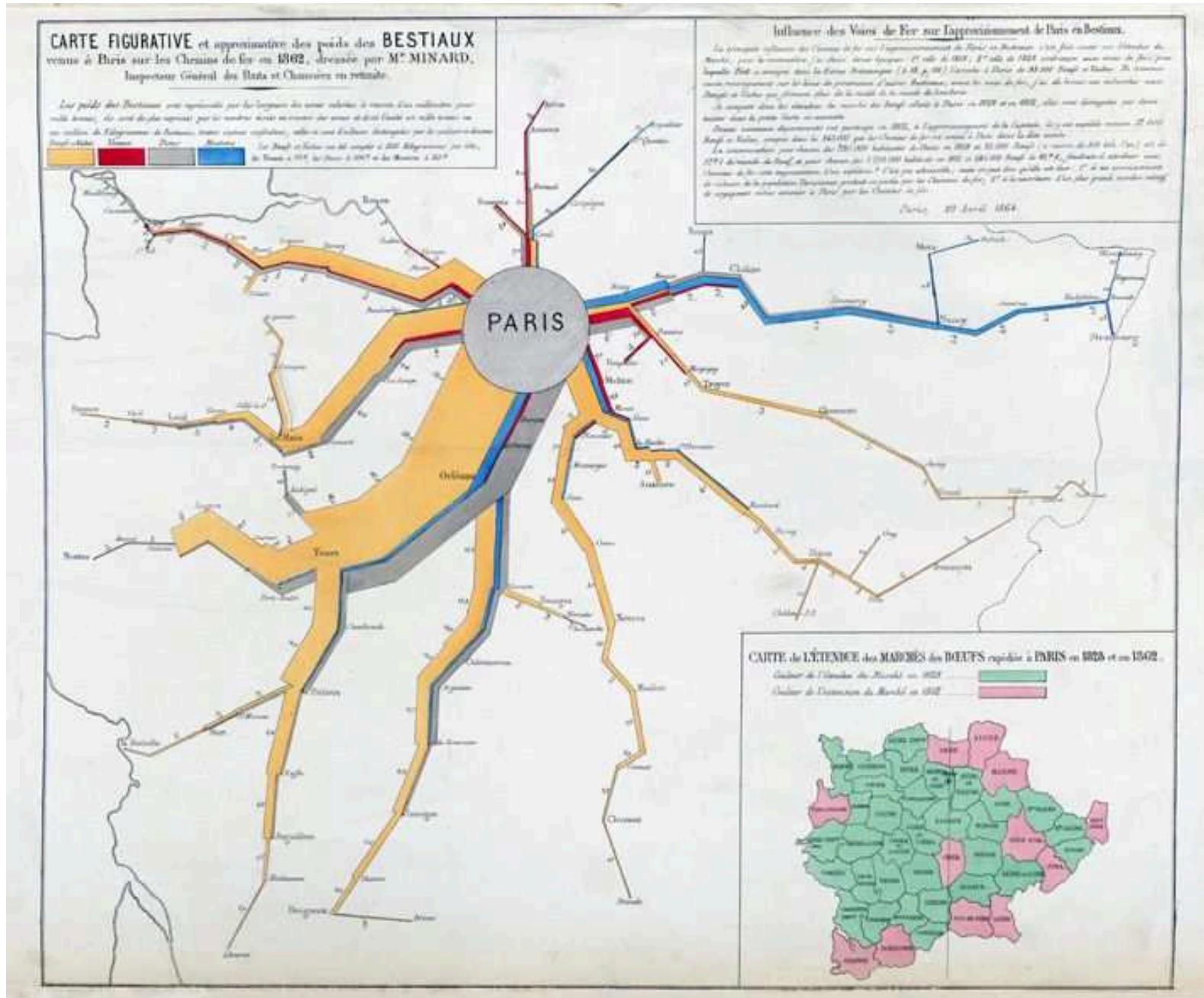




"Carte figurative des pertes successives en hommes de l'Armée qu'Annibal conduisit d'Espagne en Italie en traversant les Gaules" (top) and "Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812–1813" (bottom) by Charles Joseph Minard (1869)

- shows the force levels of the armies of Hannibal (218 BC) and Napoleon (1812-1813), respectively
- some data visualization practitioners call it (one of) **the best statistical drawings ever created**





"Carte figurative et approximative des poids des bestiaux venus à Paris sur les chemins de fer en 1862" by Charles Joseph Minard (1864)



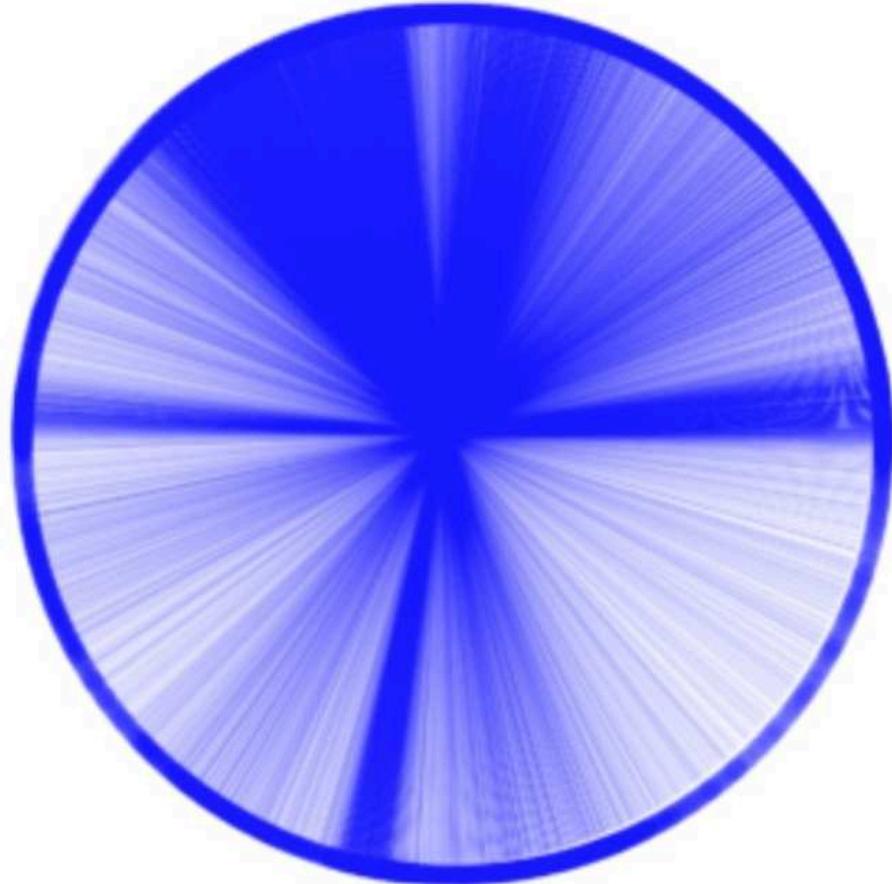
Mental Exercise



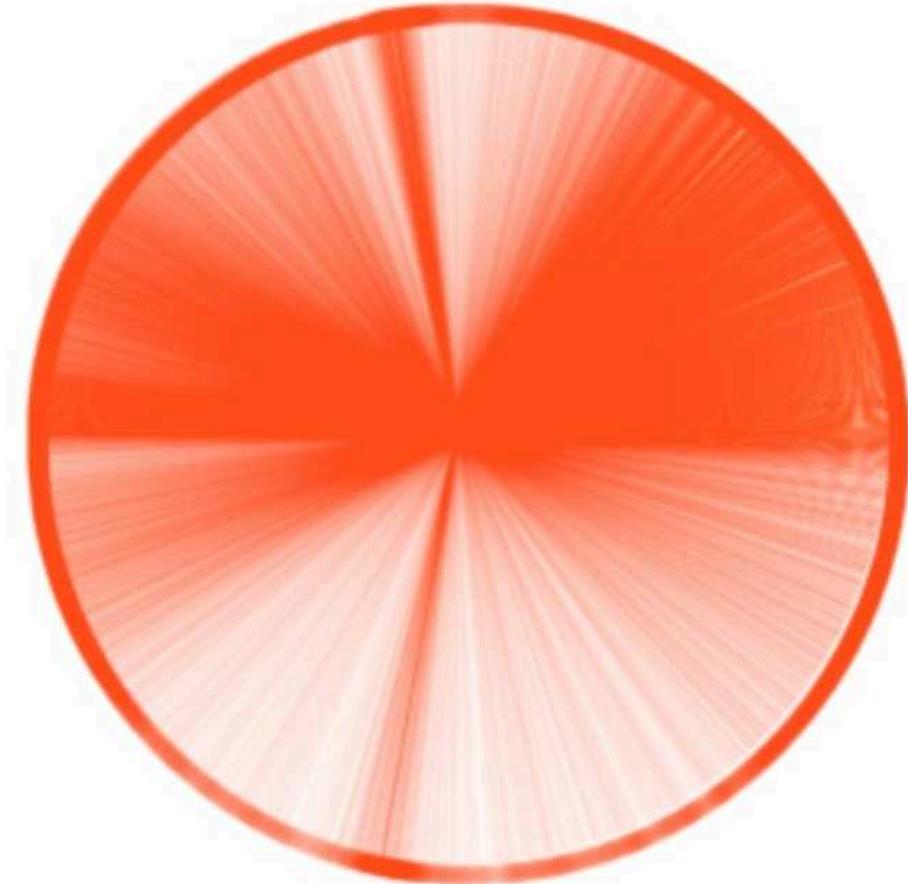
If the year is a circle—where's March and December in your mind?

- Create a polar representation of the year: draw a circle and indicate the position of March and December.
- Use an arrow to illustrate the direction of time.
- Note if you've ever thought about this before.
- Compare the results with your neighbors.





December

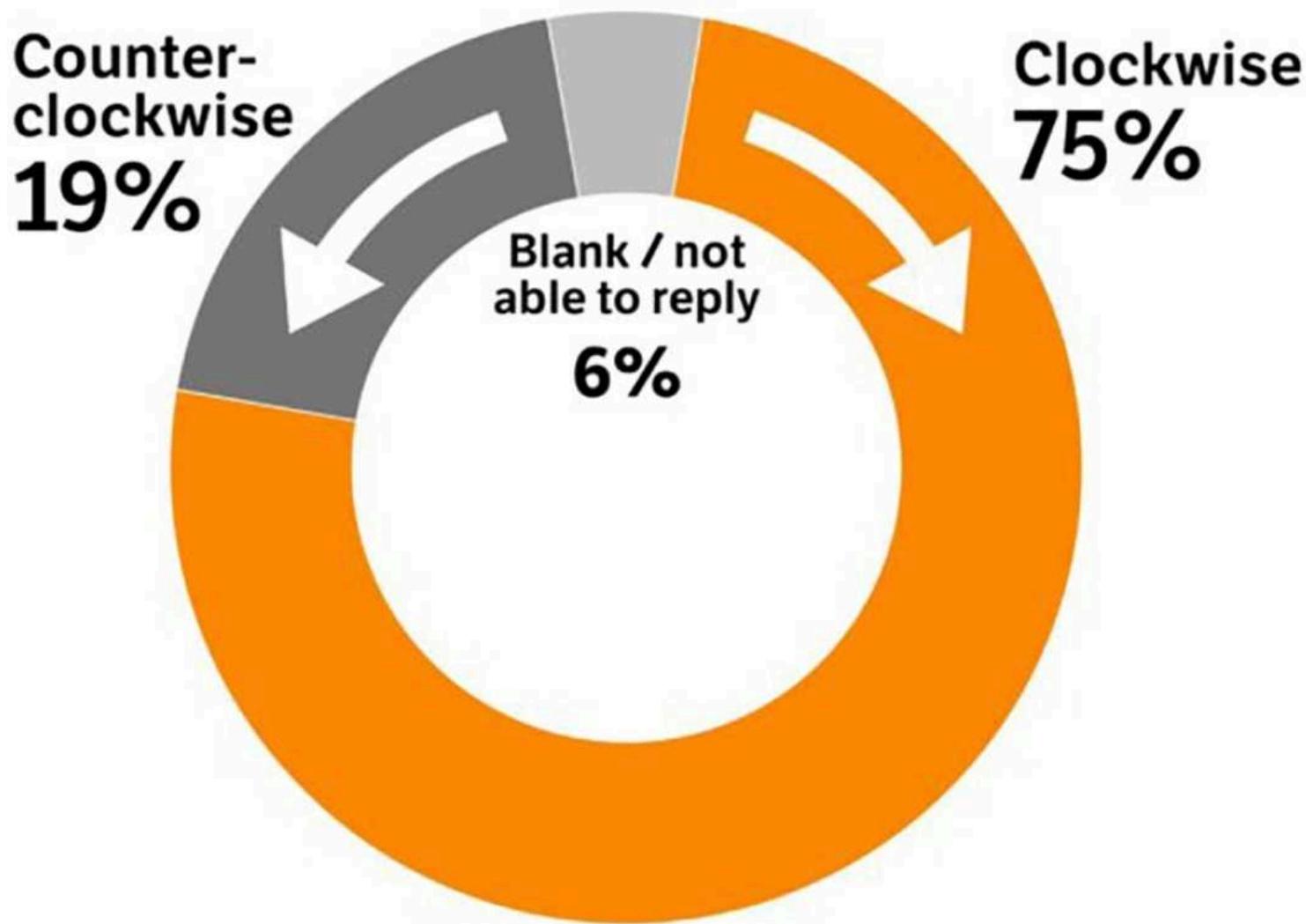


March

Wheel diagram of 76,922 placements of the months of December and March on the circumference of an empty circle.

Graphics by Henrik Lied at NRKbeta. [Laeng & Hofseth, *Front Psychol.* 2019](#)

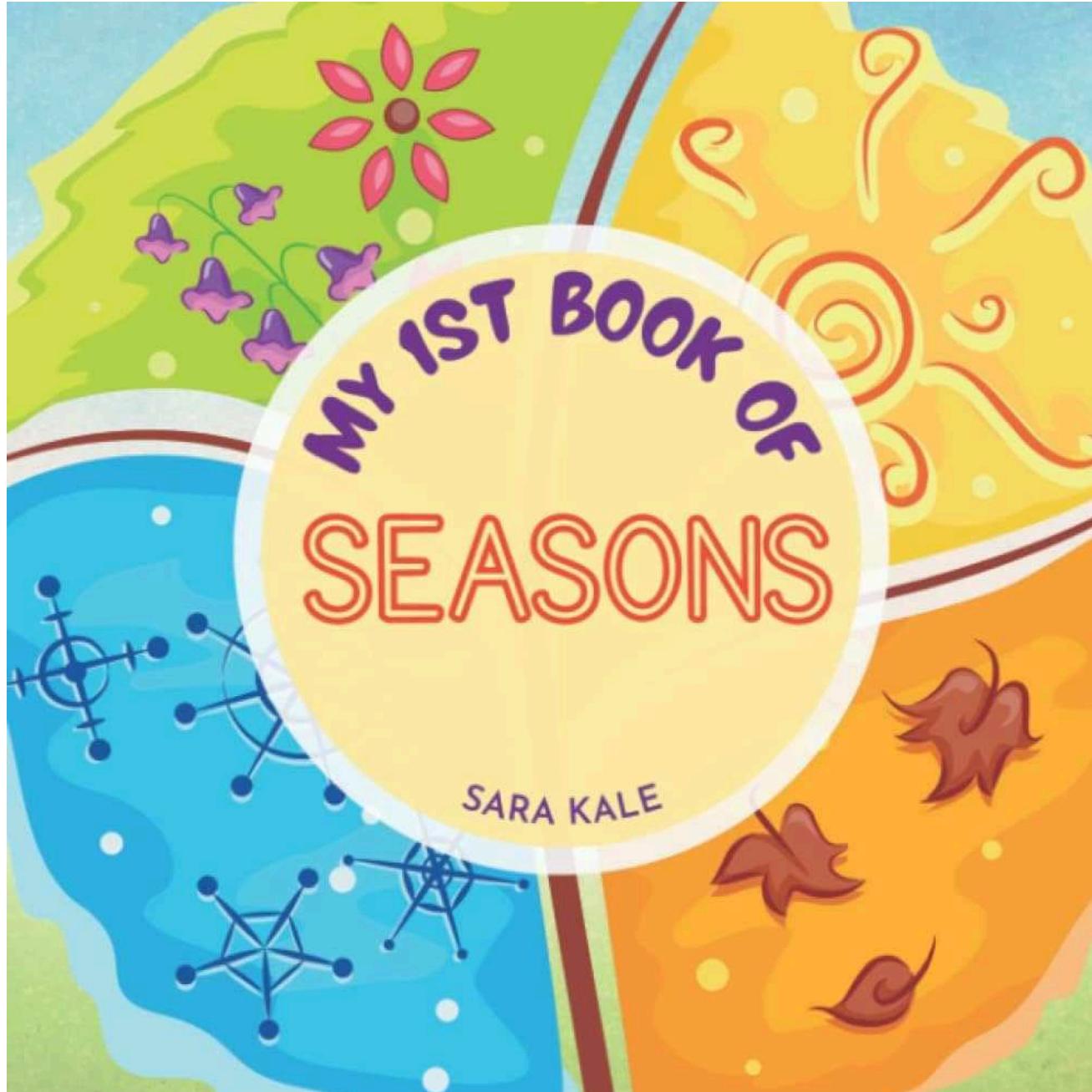




Proportion of respondents choosing opposite direction of time on the year's wheel.

Graphics by Vidar Kvien, NRK. [Laeng & Hofseth, Front Psychol. 2019](#)



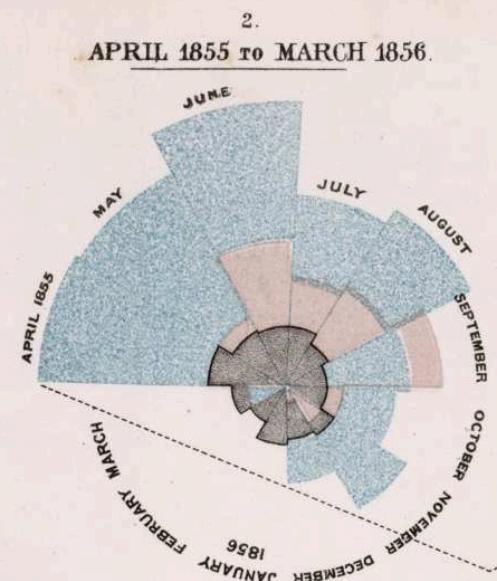


Book cover of the toddler and kids book "My 1st Book of Seasons" by Sara Kale

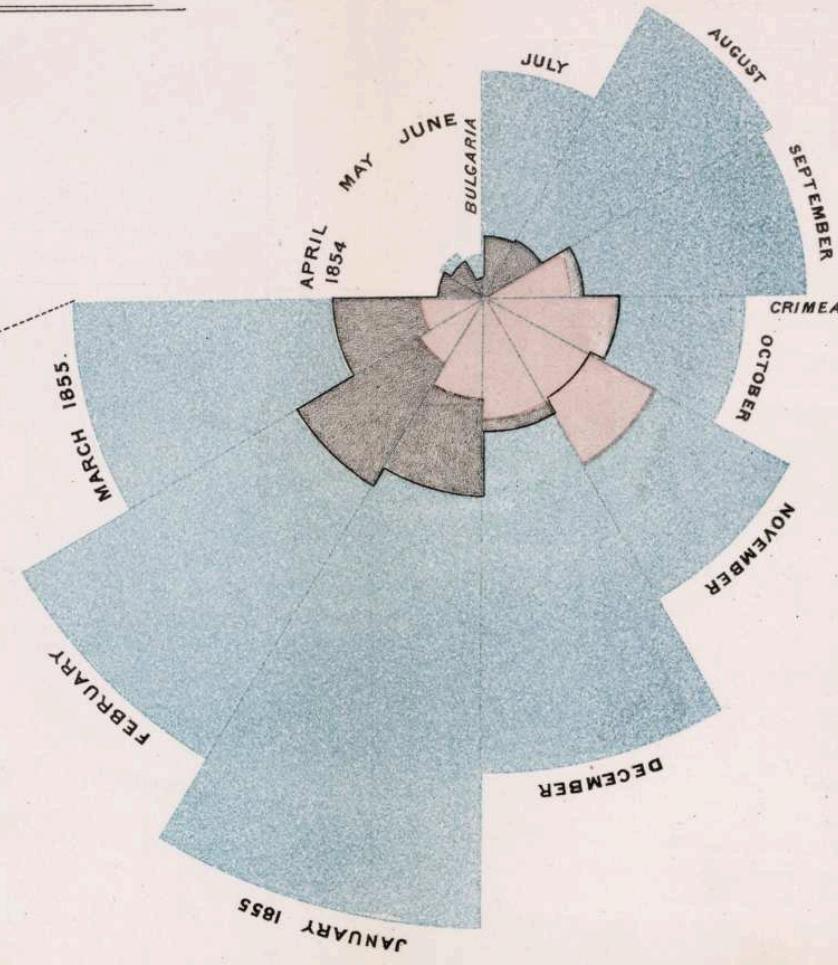
Cédric Scherer Data Visualization & Information Design



DIAGRAM OF THE CAUSES OF MORTALITY
IN THE ARMY IN THE EAST.



1.
APRIL 1854 TO MARCH 1855.



The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.

The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.

The black line across the red triangle in Nov? 1854 marks the boundary of the deaths from all other causes during the month.

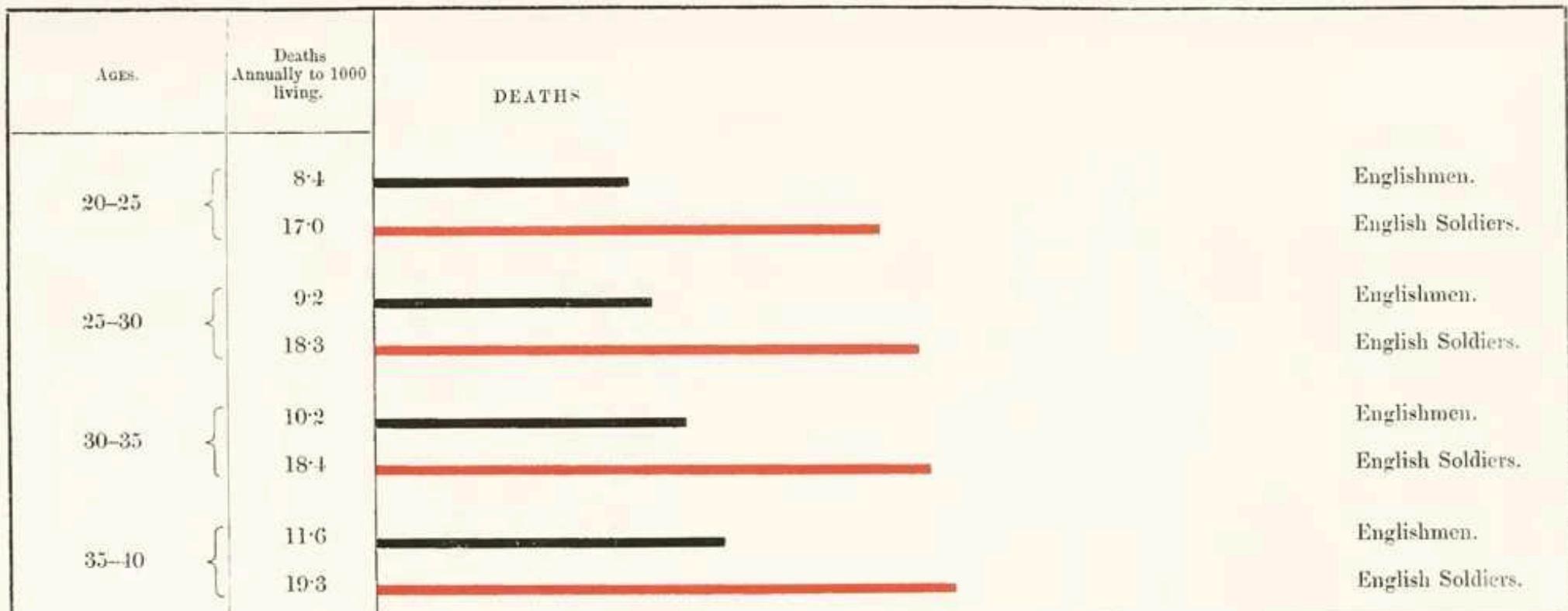
In October 1854, & April 1855, the black area coincides with the red; in January & February 1856, the blue coincides with the black.

The entire areas may be compared by following the blue, the red & the black lines enclosing them.

"Diagram of the causes and mortality in the army in the East" (a so-called *coxcomb diagram*) by Florence Nightingale (1858)



Representing the Relative Mortality of the Army at Home and of the English Male Population at corresponding Ages.



JAMES LEWIS, del.

"Relative mortality of the army at home and of the English male population at corresponding ages" by Florence Nightingale (1858)



Anscombe's Quartet

	I	II	III	IV			
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89



Anscombe's Quartet

	I	II	III	IV			
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.10	4	5.39	19	12.50
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89



Anscombe's Quartet

	I	II	III	IV
4	4.26	4 3.10	4 5.39	8 5.25
5	5.68	5 4.74	5 5.73	8 5.56
6	7.24	6 6.13	6 6.08	8 5.76
7	4.82	7 7.26	7 6.42	8 6.58
8	6.95	8 8.14	8 6.77	8 6.89
9	8.81	9 8.77	9 7.11	8 7.04
10	8.04	10 9.14	10 7.46	8 7.71
11	8.33	11 9.26	11 7.81	8 7.91
12	10.84	12 9.13	12 8.15	8 8.47
13	7.58	13 8.74	13 12.74	8 8.84
14	9.96	14 8.10	14 8.84	19 12.50



Anscombe's Quartet

I

$$\begin{aligned}\mu_x &= 9 \\ \sigma_x &= 3.16\end{aligned}$$

$$\begin{aligned}\mu_y &= 7.5 \\ \sigma_y &= 1.94\end{aligned}$$

$$r = 0.82$$

II

$$\begin{aligned}\mu_x &= 9 \\ \sigma_x &= 3.16\end{aligned}$$

$$\begin{aligned}\mu_y &= 7.5 \\ \sigma_y &= 1.94\end{aligned}$$

$$r = 0.82$$

III

$$\begin{aligned}\mu_x &= 9 \\ \sigma_x &= 3.16\end{aligned}$$

$$\begin{aligned}\mu_y &= 7.5 \\ \sigma_y &= 1.94\end{aligned}$$

$$r = 0.82$$

IV

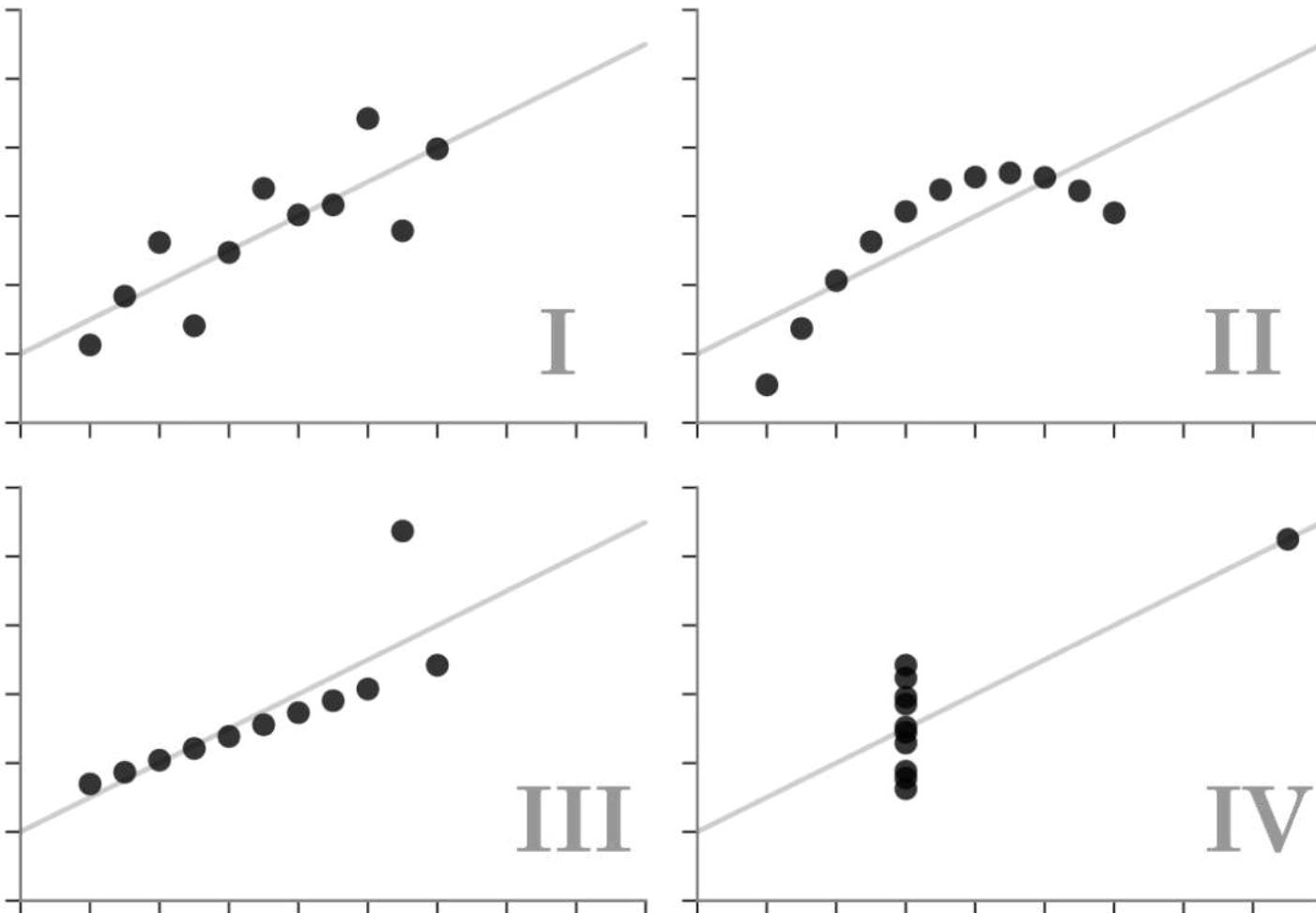
$$\begin{aligned}\mu_x &= 9 \\ \sigma_x &= 3.16\end{aligned}$$

$$\begin{aligned}\mu_y &= 7.5 \\ \sigma_y &= 1.94\end{aligned}$$

$$r = 0.82$$



Anscombe's Quartet



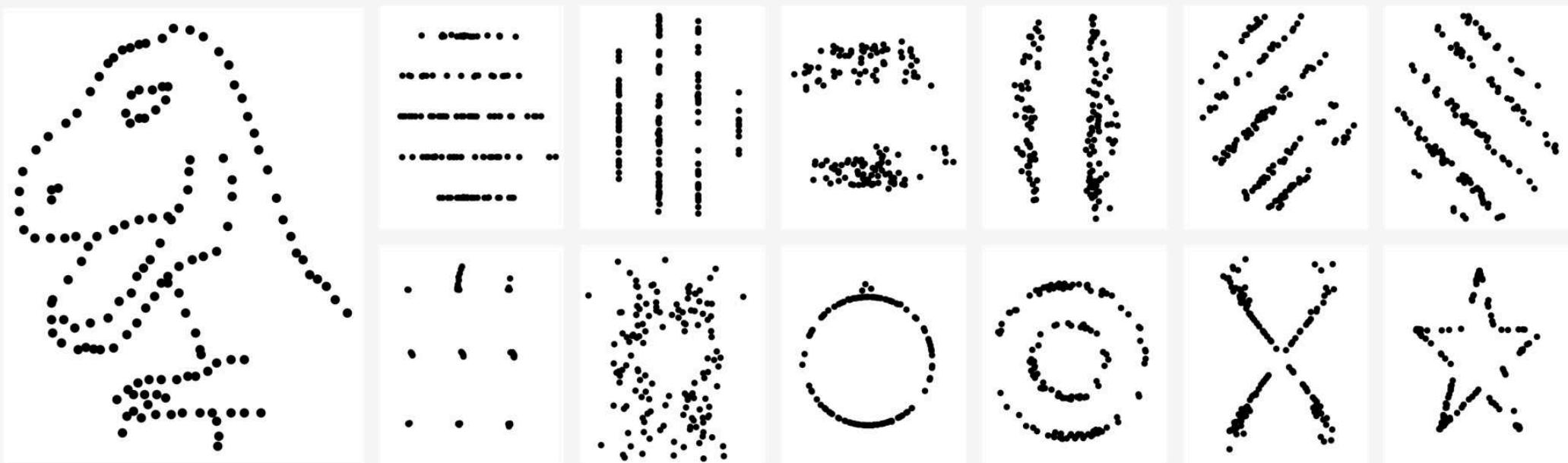
Source: Matejka & Fitzmaurice (2017)

Cédric Scherer Data Visualization & Information Design



The Datasaurus Dozen

A set of 13 datasets with nigh-identical summary statistics, which could lead one to believe the datasets are quite similar. After visualizing the data, it becomes clear that they are clearly different and visually distinct.

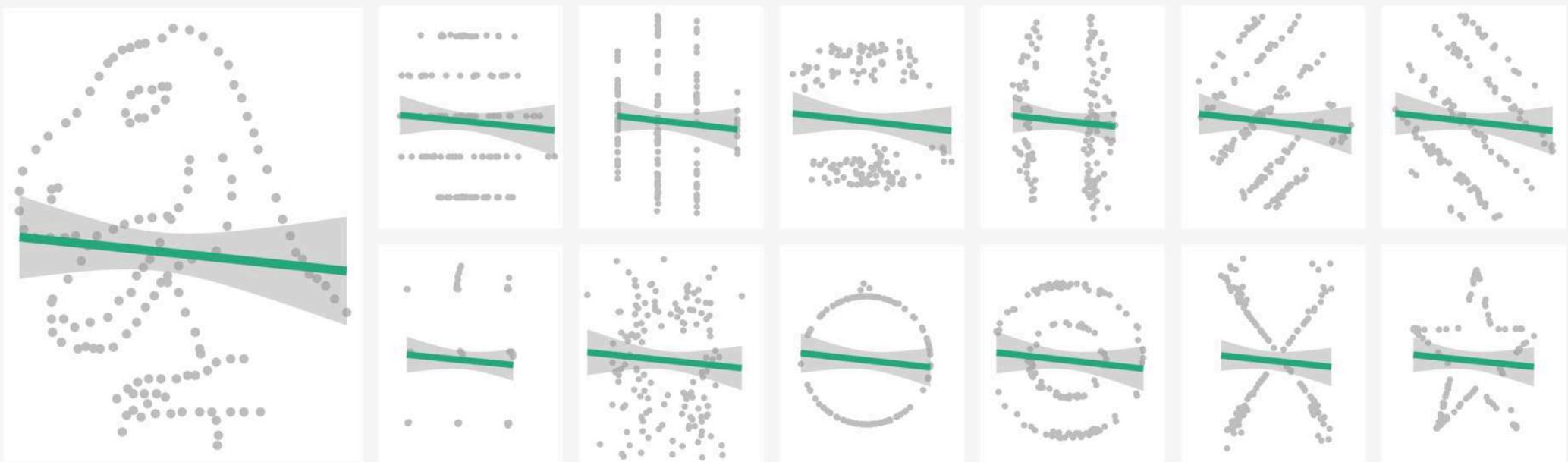


"Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing" by Justin Matejka & George Fitzmaurice (2017)



The Datasaurus Dozen

A set of 13 datasets with nigh-identical summary statistics, which could lead one to believe the datasets are quite similar. After visualizing the data, it becomes clear that they are clearly different and visually distinct.



"Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing" by Justin Matejka & George Fitzmaurice (2017)



*“... make both **calculations** and **graphs**.
Both sorts of output should be studied;
each will **contribute** to **understanding**.”*

F. J. Anscombe (1973)



*A good data visualization can mean the difference between **success** and **failure**.*

- Revealing the underlying patterns in data
- Communicating complex findings and phenomena
- Raising money for an organization, event or department
- Helping people to make informed arguments and decisions
- Providing guidance for improvement
- ...
- **Getting your point across!**





Cédric Scherer Data Visualization & Information Design





Good vs Bad



What Makes a Good Data Visualization?

- » **Integrity** (information)
- » **Story** (interestingness)
- » **Goal** (usefulness)
- » **Visual Form** (beauty)



Integrity (information)



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



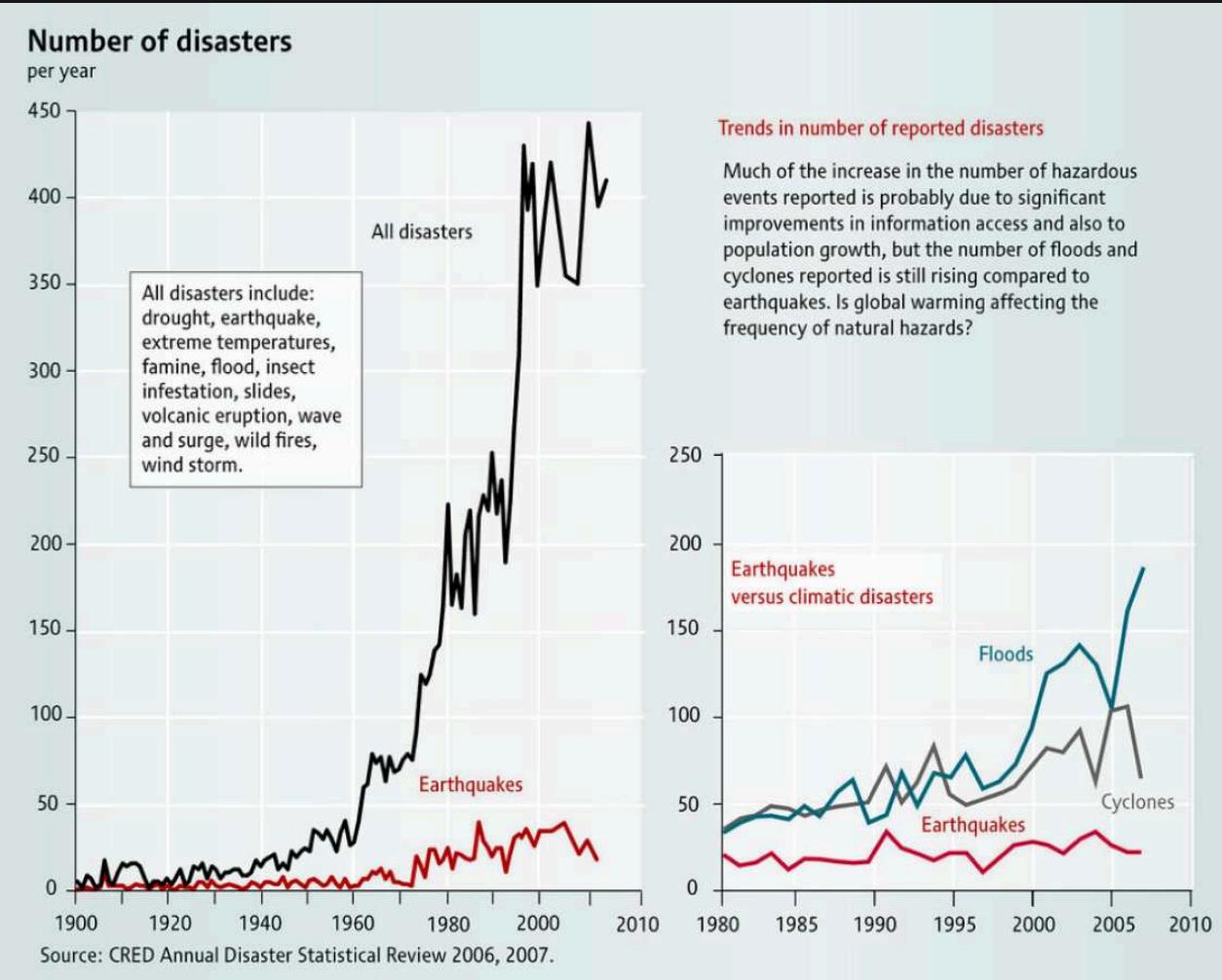
Data Integrity

- **data quality:**
 - guesstimation, precision, and failures
 - miscalculations and errors
 - incomplete data and missing values
 - summaries and aggregations
- **only a subset:**
 - not crime but reported crime*
 - historical or present state

* or rats, UFO sightings, ...

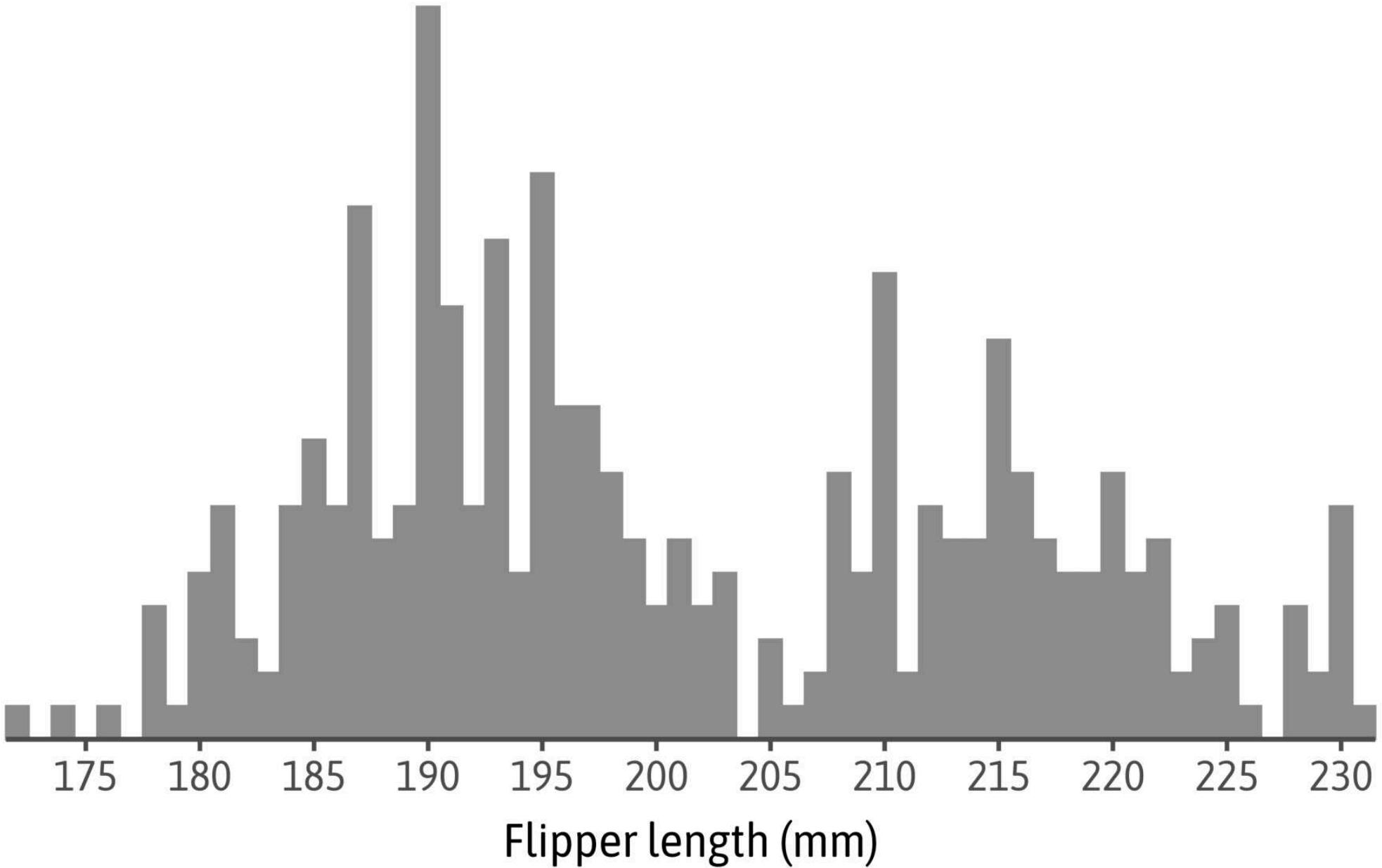


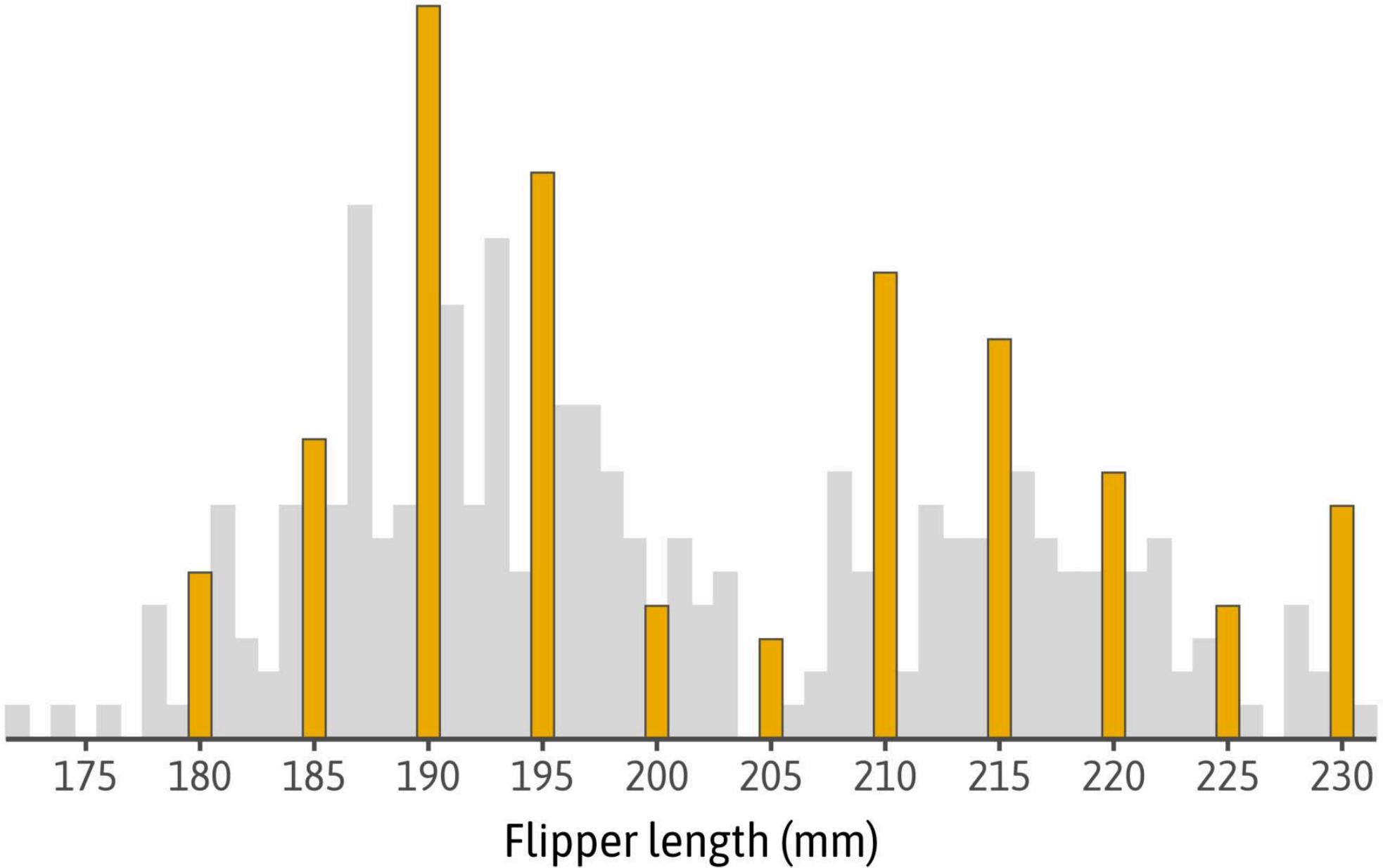


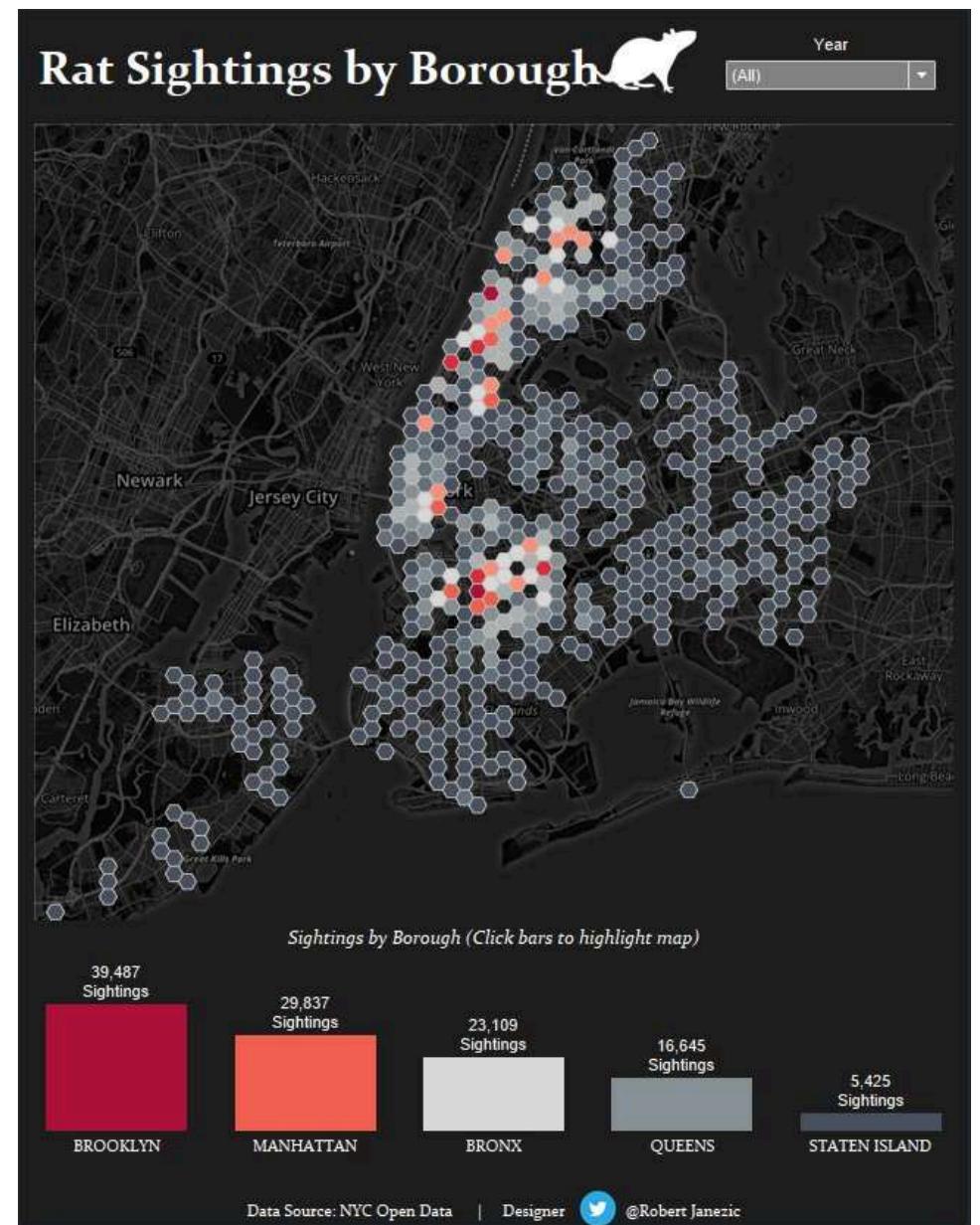
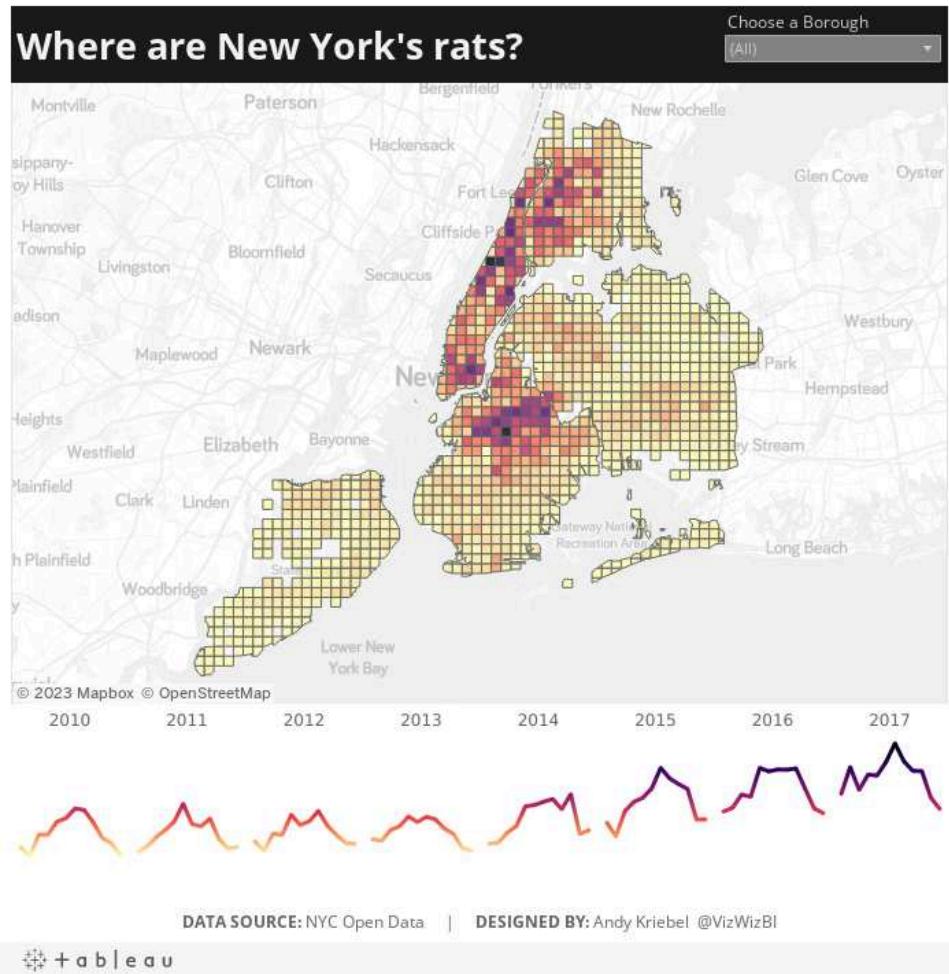


"Much of the increase of hazardous events reported is probably due to **significant improvements in information access."**









Story

(interestingness)



Story

- **relevant topic:**
 - identify the main message
 - address your audience
 - adjust the level of detail

- **meaningful data:**
 - condense information
 - emphasize important insights



The 3-Minute Story

Imagine that you have only **three minutes** to tell your audience what they **need to know or do**

—what would you say?



The Big Idea

Imagine you have to condense your story to a **single sentence** to convey **what's unique** about your point and **what's at stake**

—what would you say?



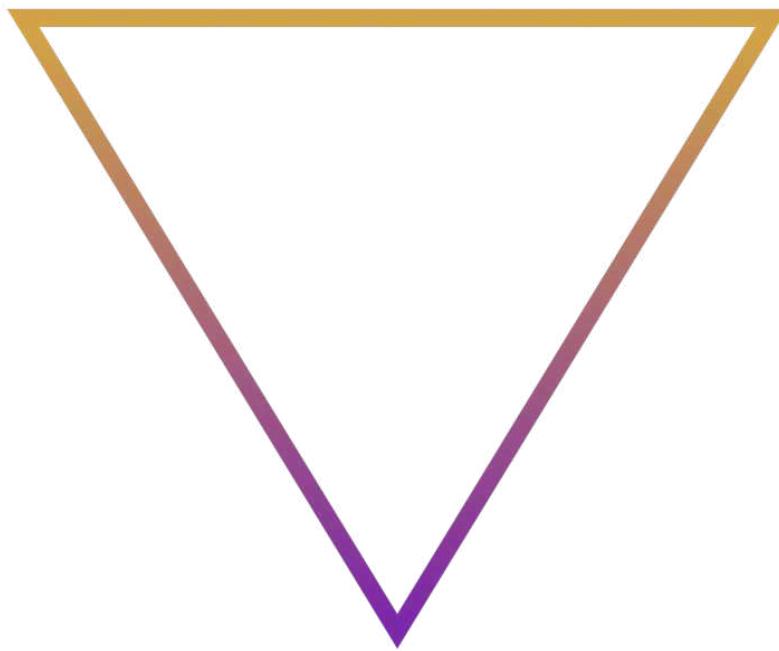
The Importance of Context



Exploratory
Discovery

Explanatory
Communication

Affective
Emotion



Priority:
efficient + effective

Response:
functional



Priority:
creative + novel

Response:
emotional

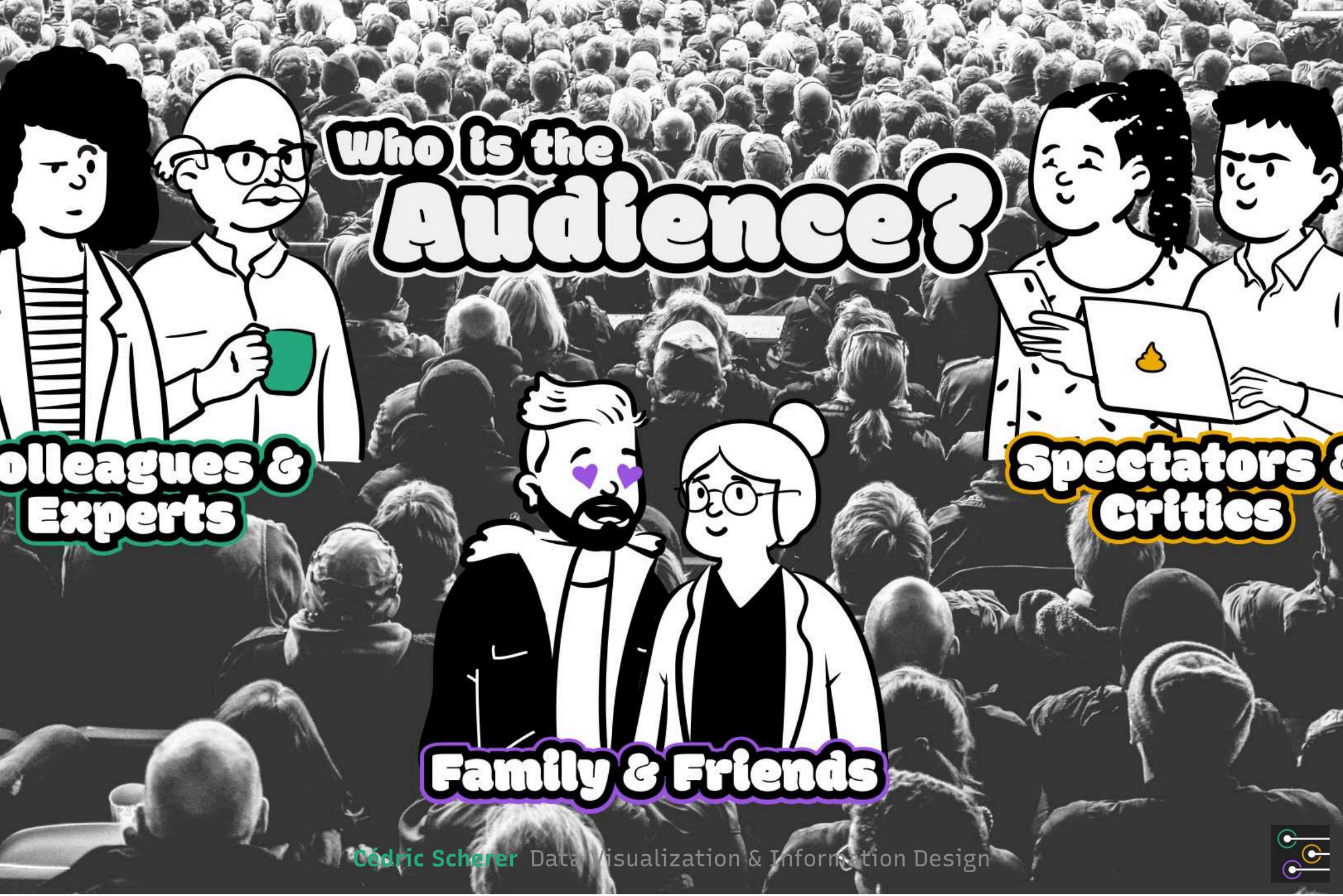
"Vertices of Visualization" by Alberto Cairo, personal communication (modified version)



Audience (who)

- To whom are you communicating?
- What do they already know?
- What is your position and relationship?





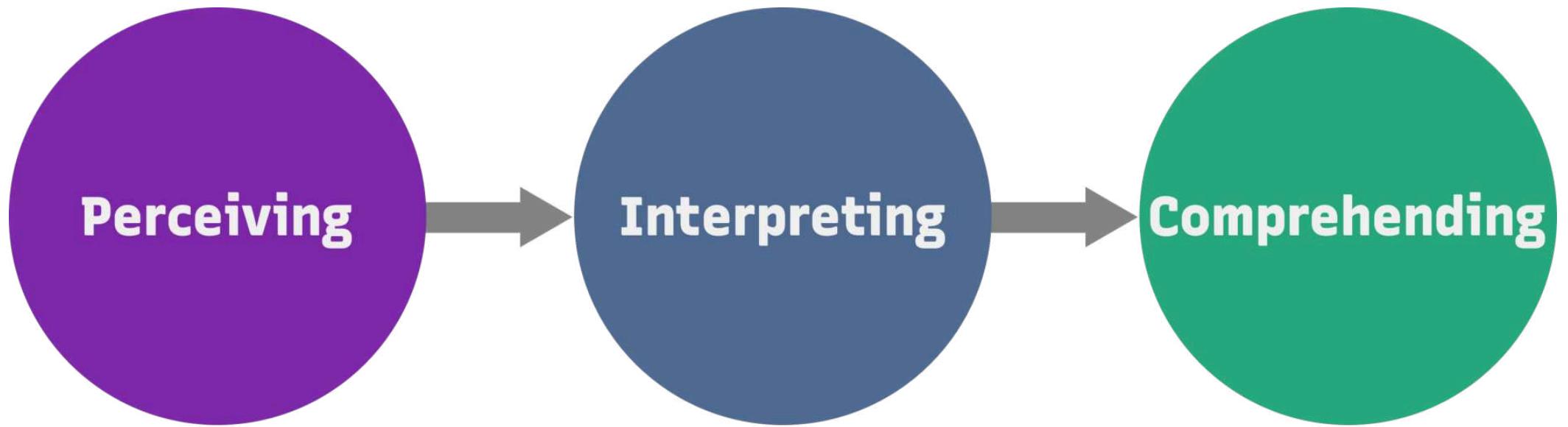
Audience (who)

- To whom are you communicating?
- What do they already know?
- What is your position and relationship?

Content (what)

- What do you want them to know or do?
- How will you communicate with them?
- What tone do you want your communication to set?





Visualiser Control

Viewer Control

Scheme by Andy Kirk (modified)





Joshua Stevens @jscarto · 10m
Visualizing data vs Data Visualization

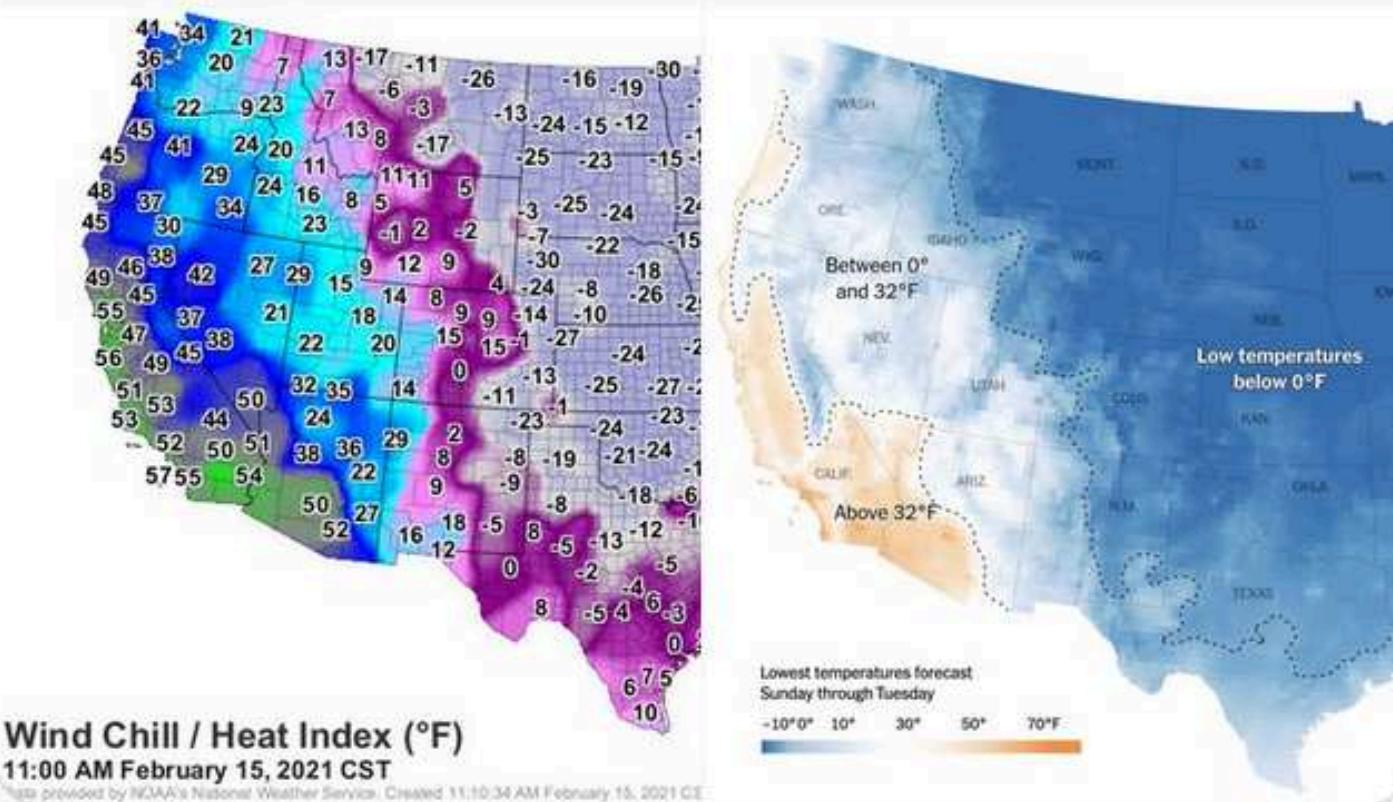
...



Ben Jones @DataRemixed · 3h

Same variable, same territory, same week, two different color palettes & two different styles. @okmesonet & @nytimes.

[Show this thread](#)



4

8

17



Cédric Scherer Data Visualization & Information Design



SETTING

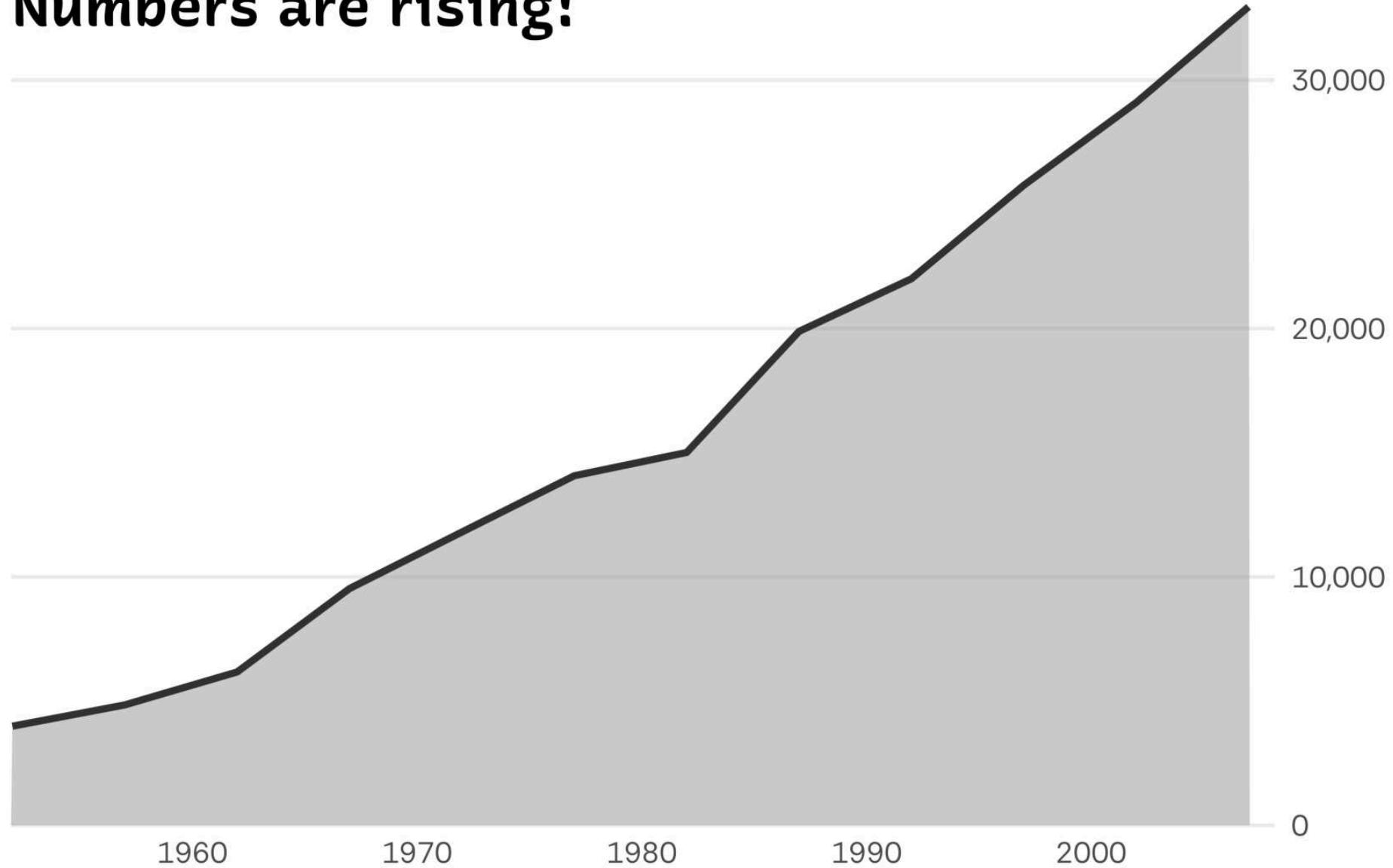
inspirational
celebrating
spicy
warning
joyful
serious

TONE

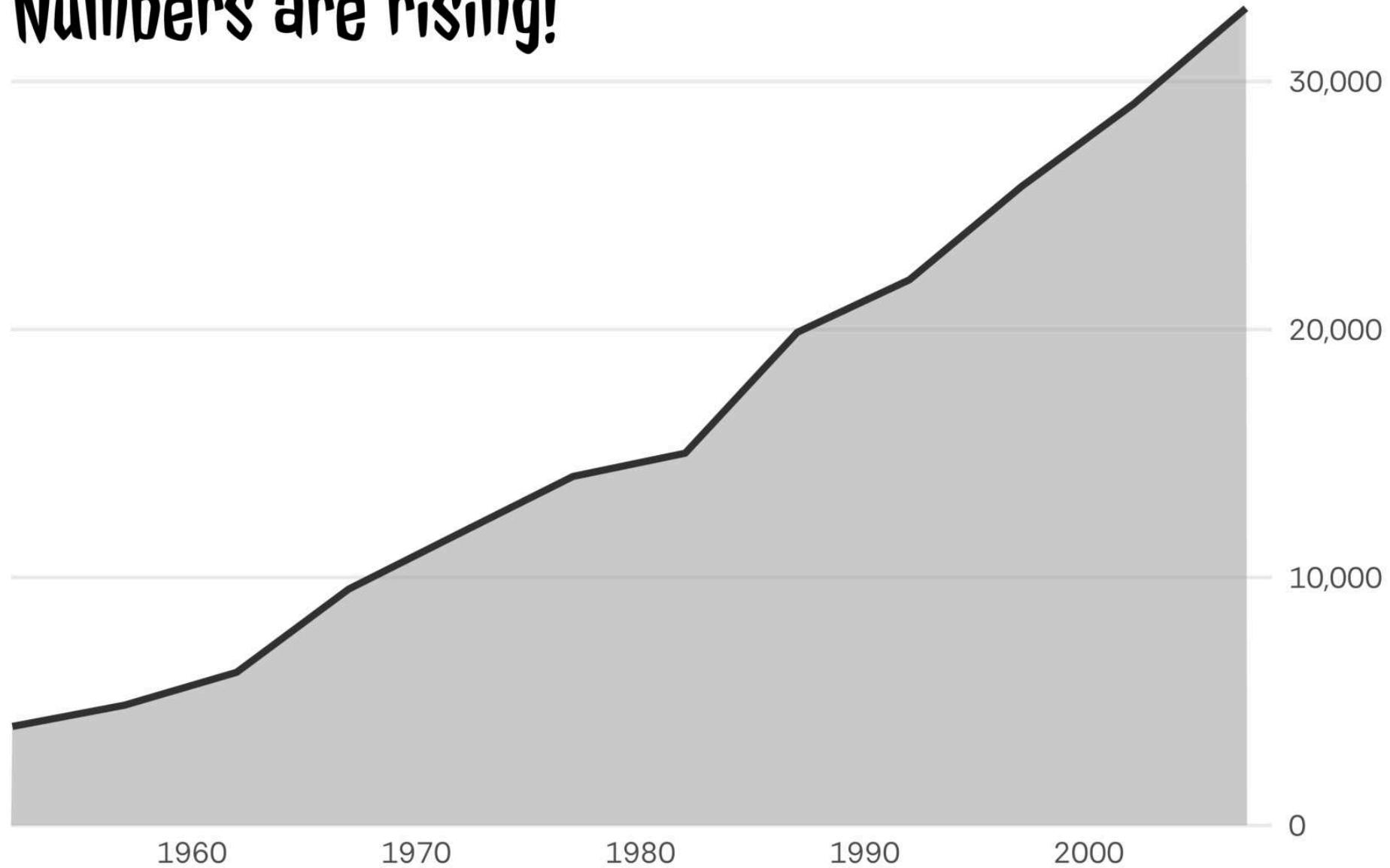
cheering
DRIVING
lighthearted
educational
urgent



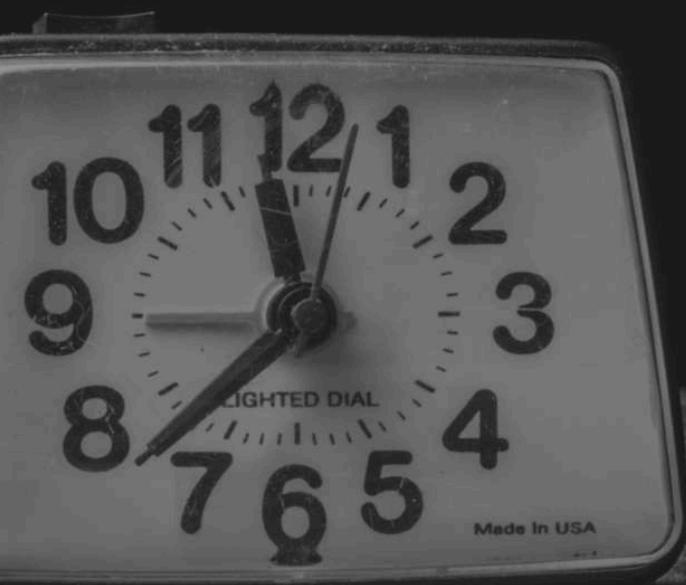
Numbers are rising!



Numbers are rising!



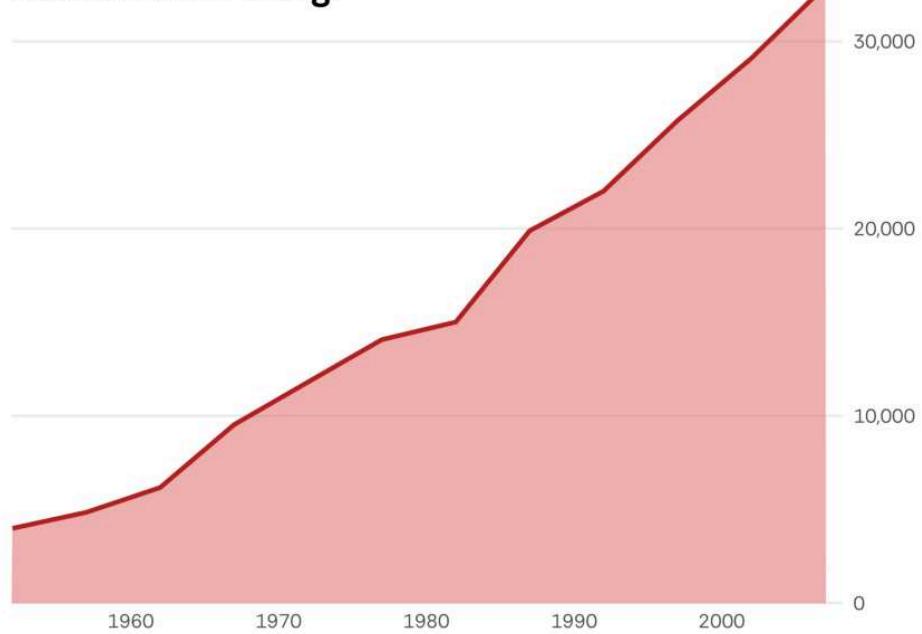
I'll be waiting for you!



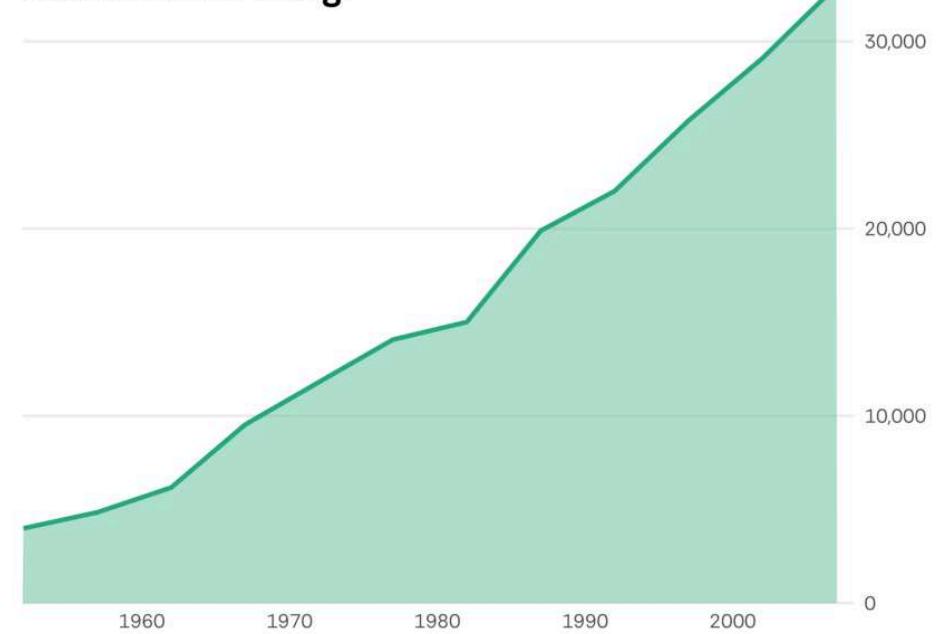
I'LL BE WAITING FOR YOU

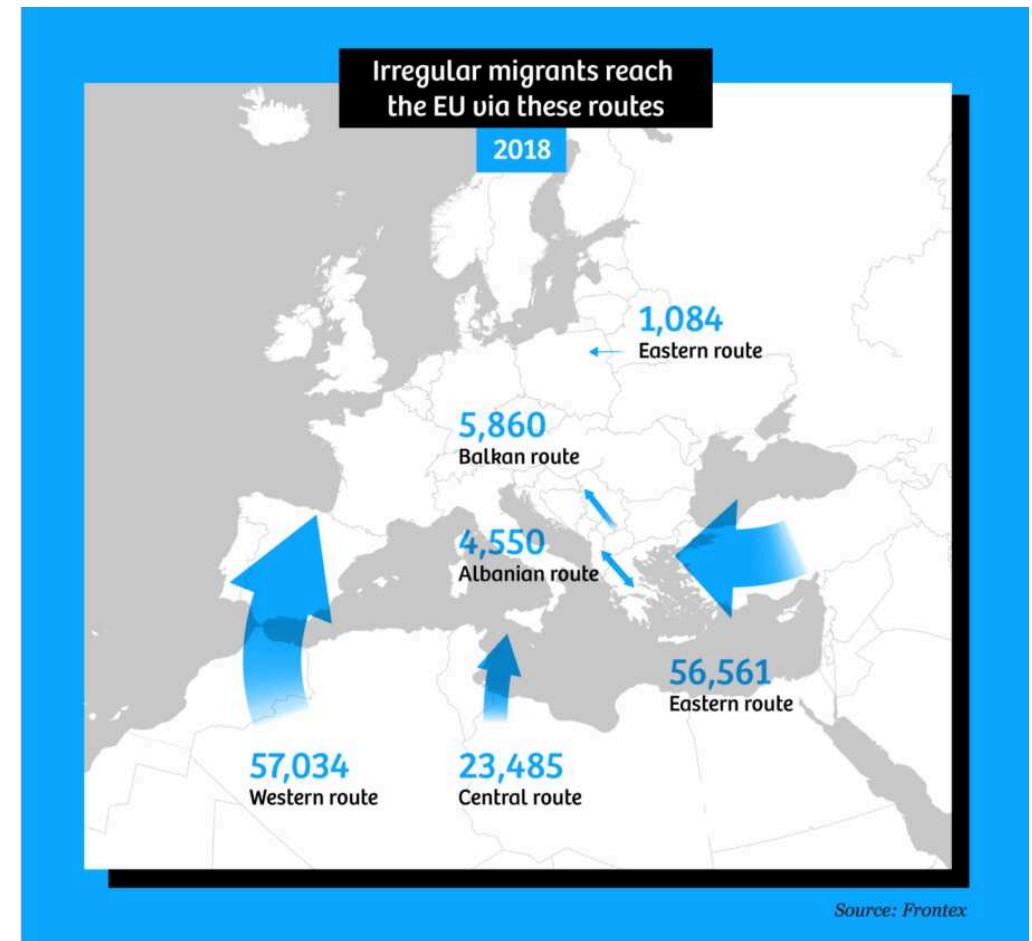


Numbers are rising!



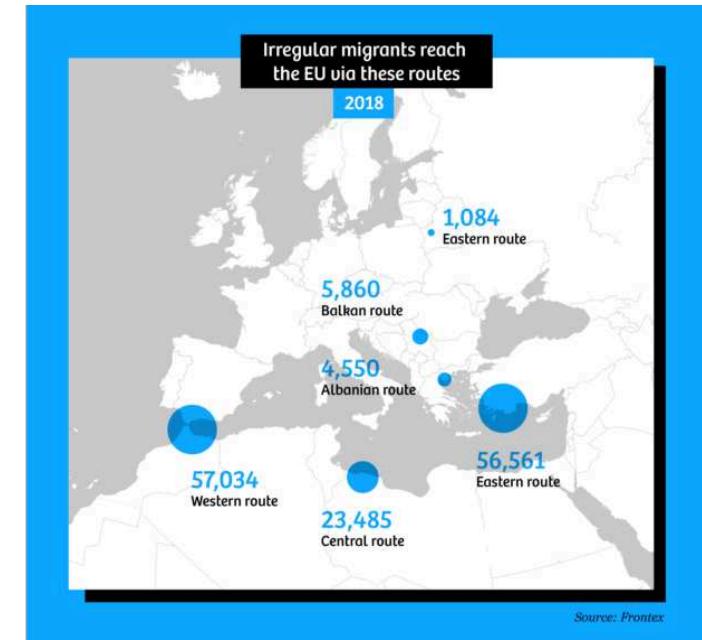
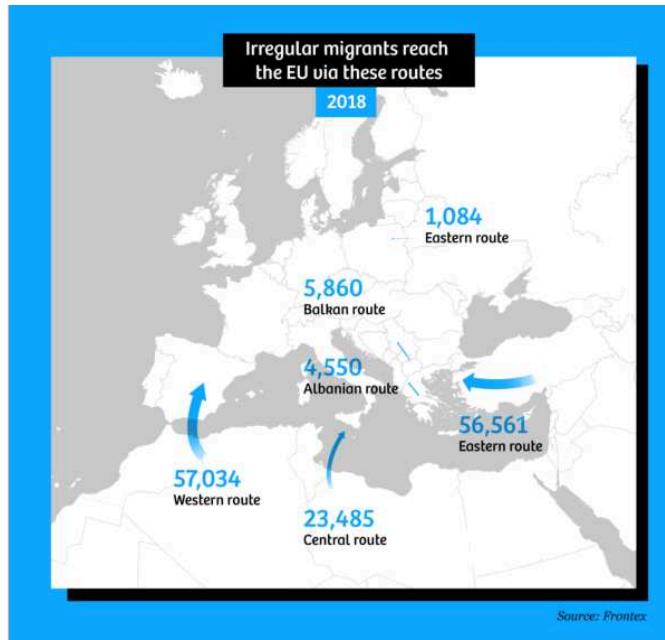
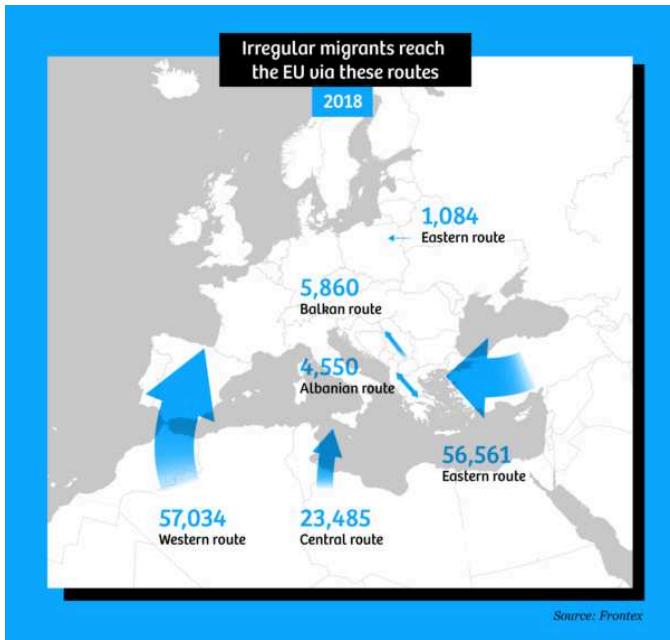
Numbers are rising!





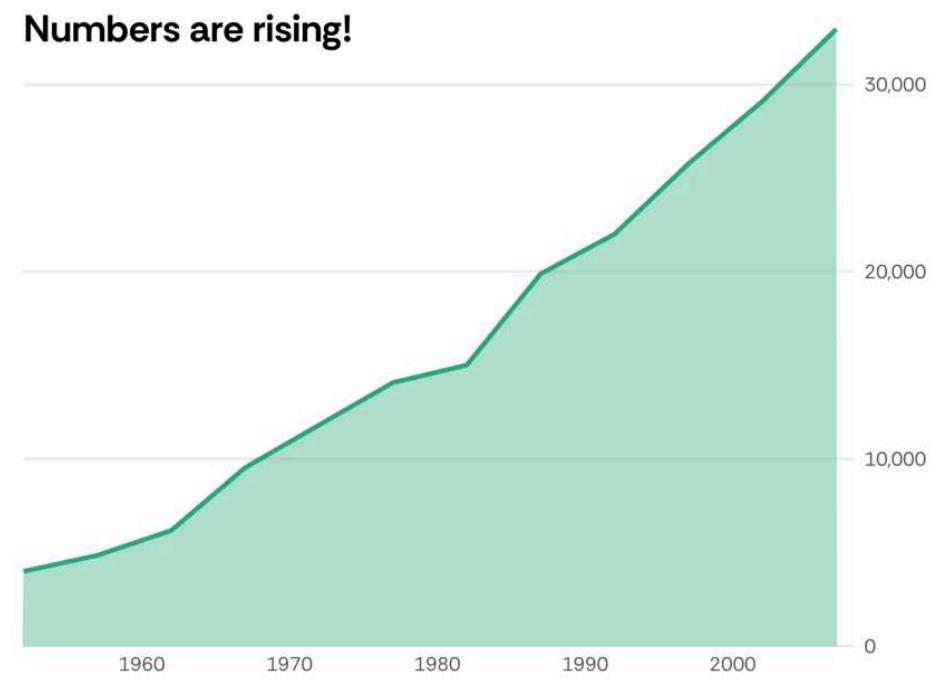
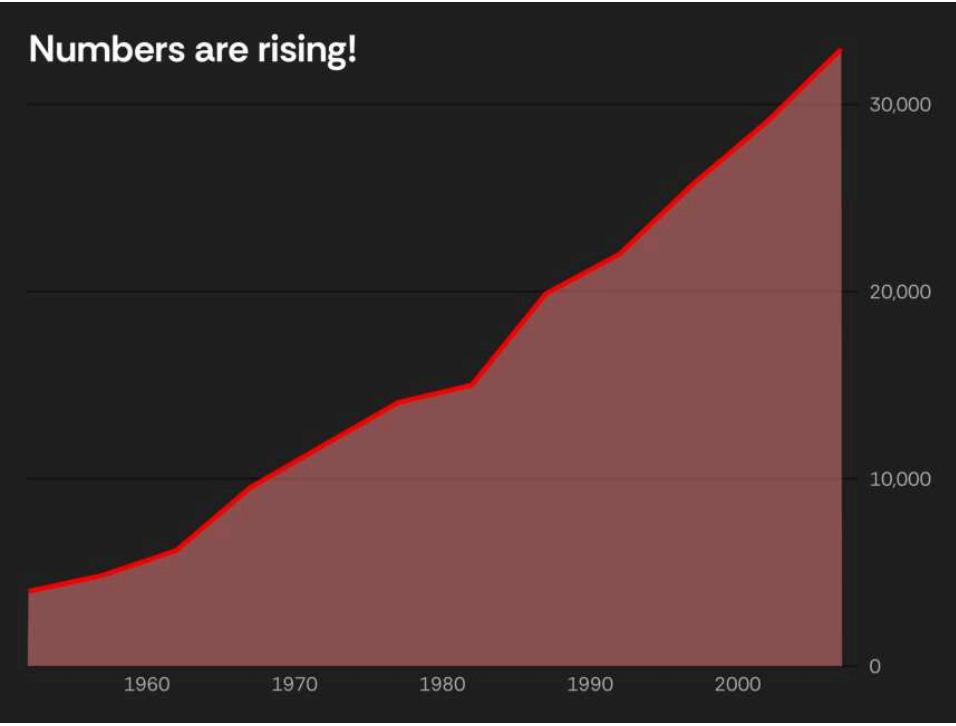
"How maps in the media make us more negative about migrants" by Maite Vermeulen, Leon de Korte & Henk van Houtum

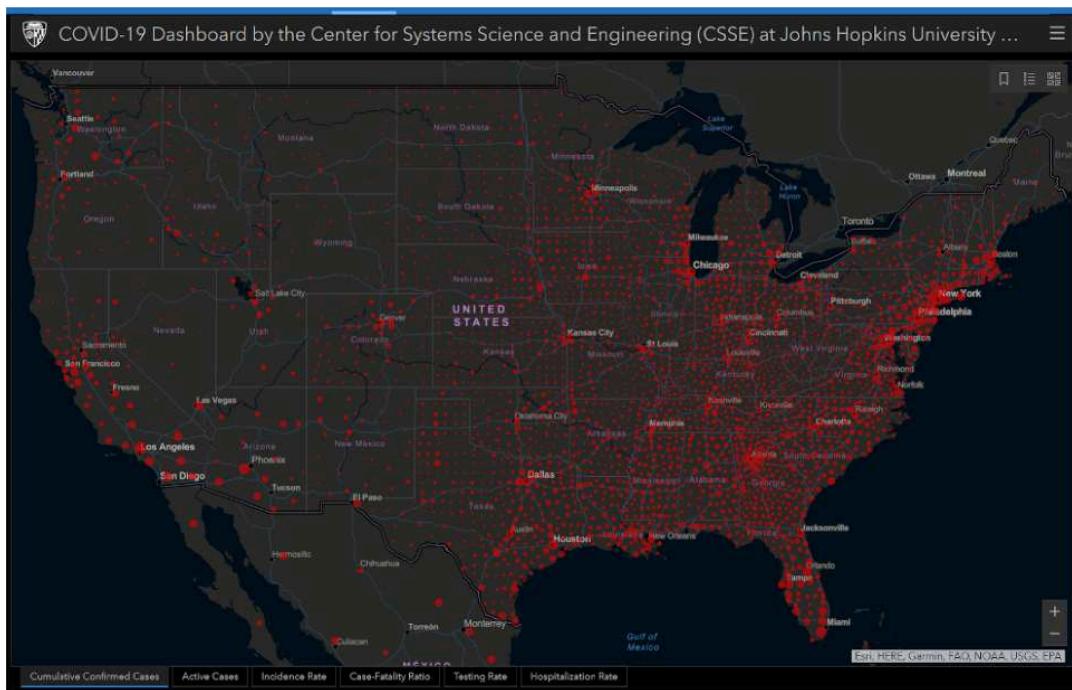




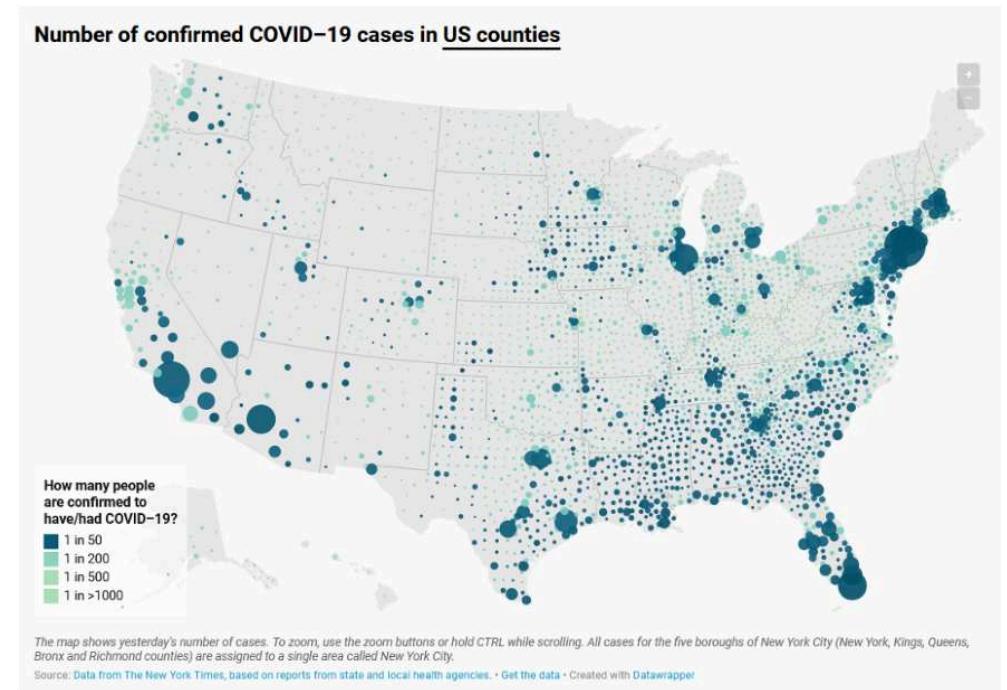
"How maps in the media make us more negative about migrants" by Maite Vermeulen, Leon de Korte & Henk van Houtum







The famous [COVID-19 tracker by the John Hopkins University](#) using an “alien-death-blood visual identity” (Screenshot 2020-07-27) next to a [map of confirmed COVID-19 cases by Datawrapper](#) using more decent colors to avoid an emotion of fear (Screenshot 2020-07-27).



Audience (who)

- To whom are you communicating?
- What do they already know?
- What is your position and relationship?

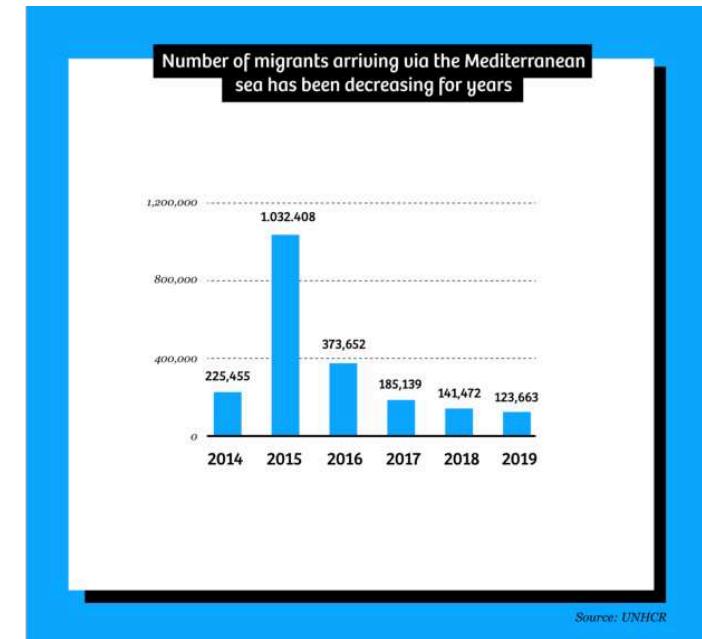
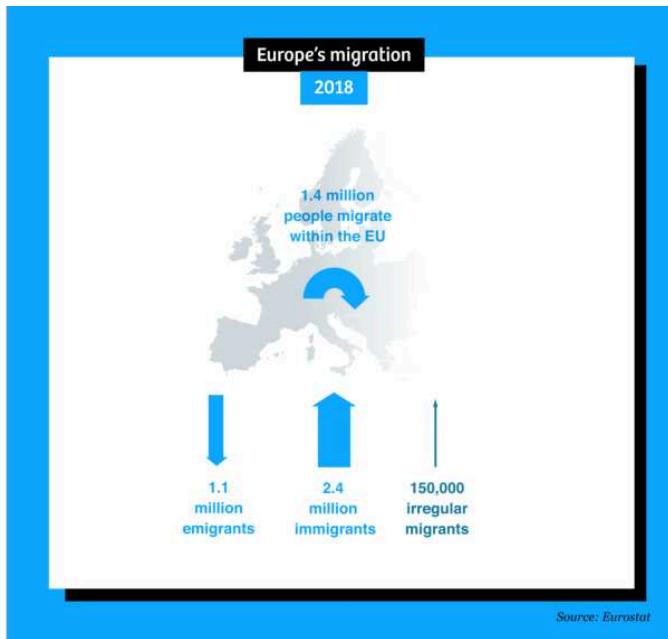
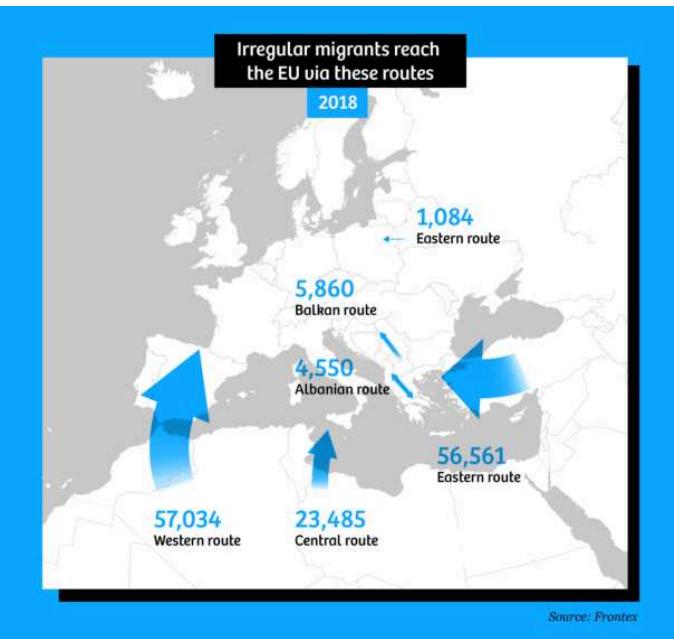
Content (what)

- What do you want them to know or do?
- How will you communicate with them?
- What tone do you want your communication to set?

Evidence (how)

- What data is available to make my point?





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



Goal

(usefulness)



Goal

- **usability:**
 - choose a suitable chart
 - respect data literacy skills
- **efficiency:**
 - use encodings that transport your message
 - guide the view(er)

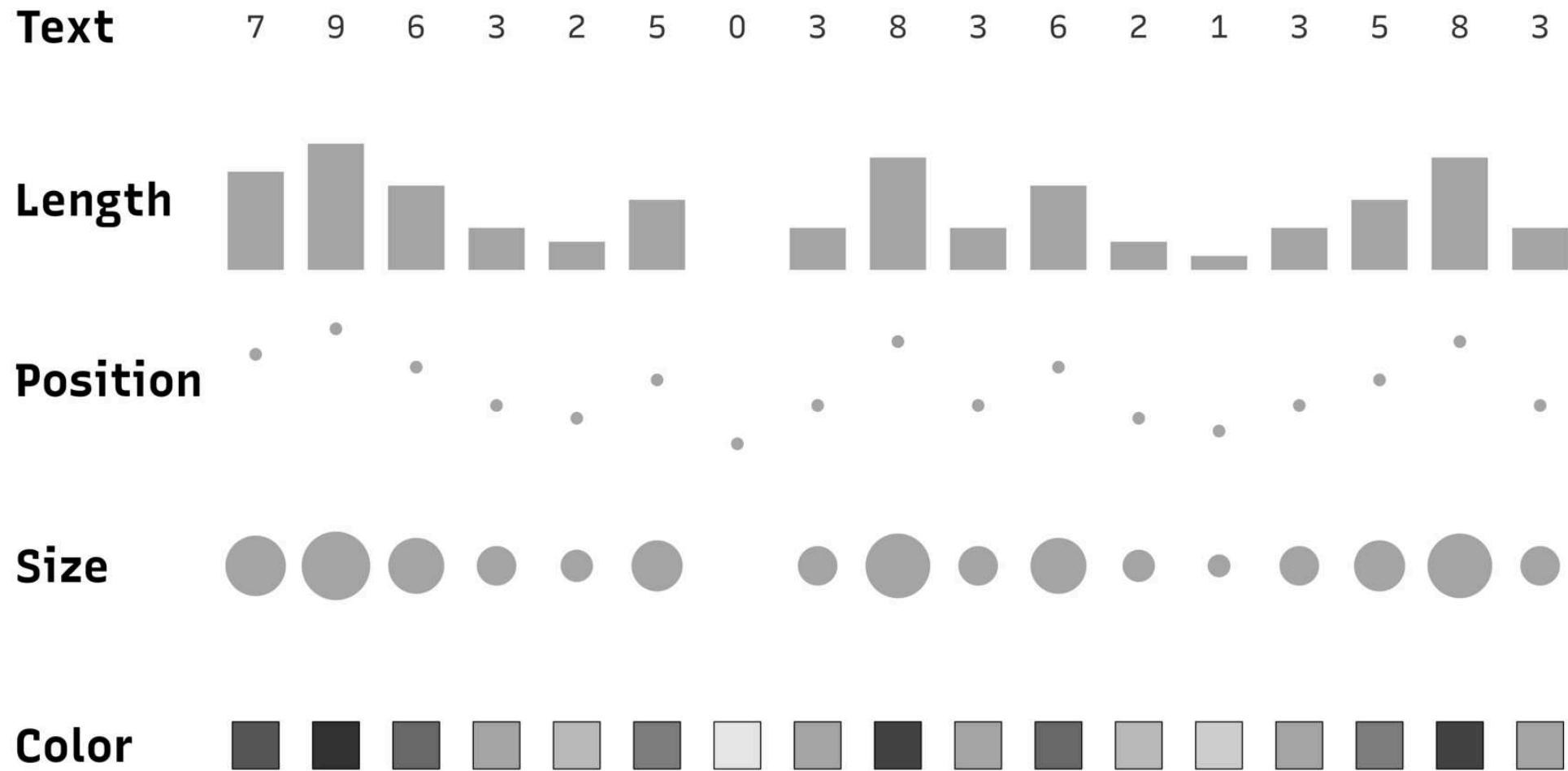


Encoding

Mapping data values to visual attributes.

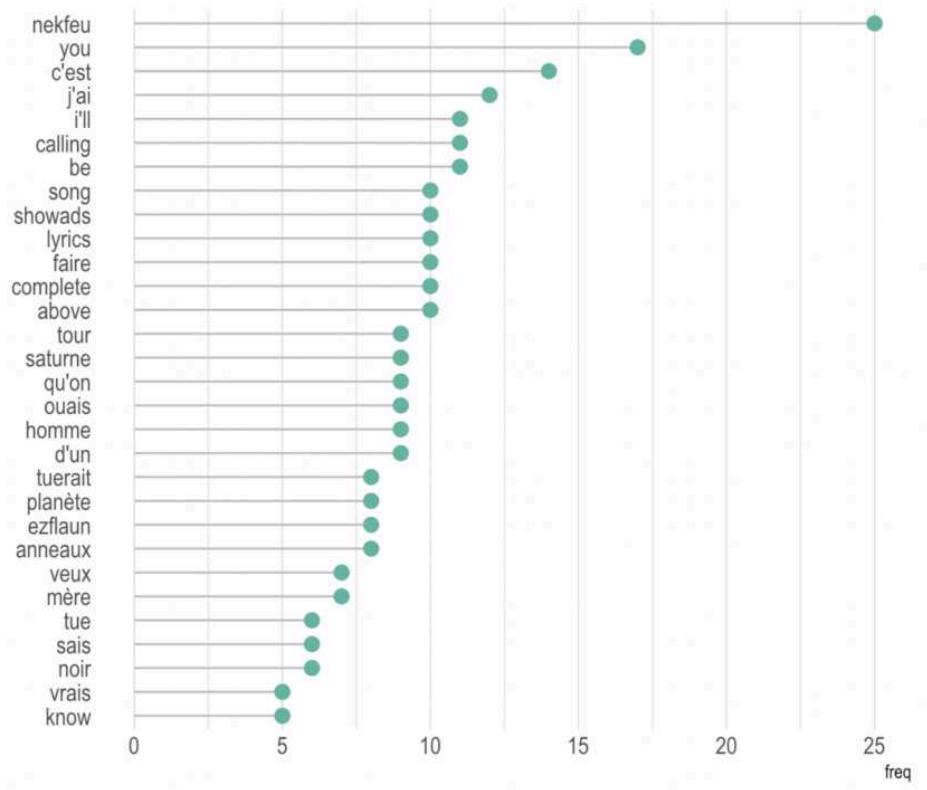


Same Data, Different Encodings



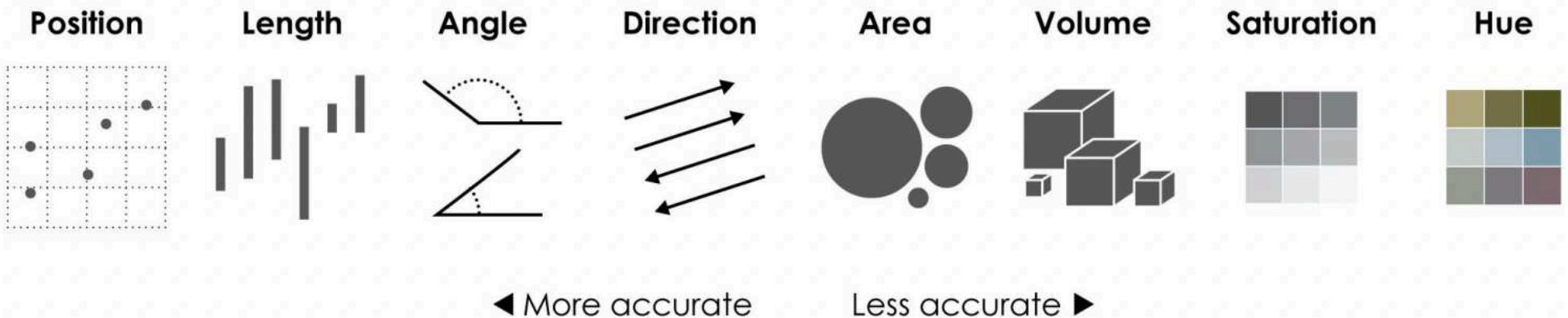


Source: [Data to Viz](#)



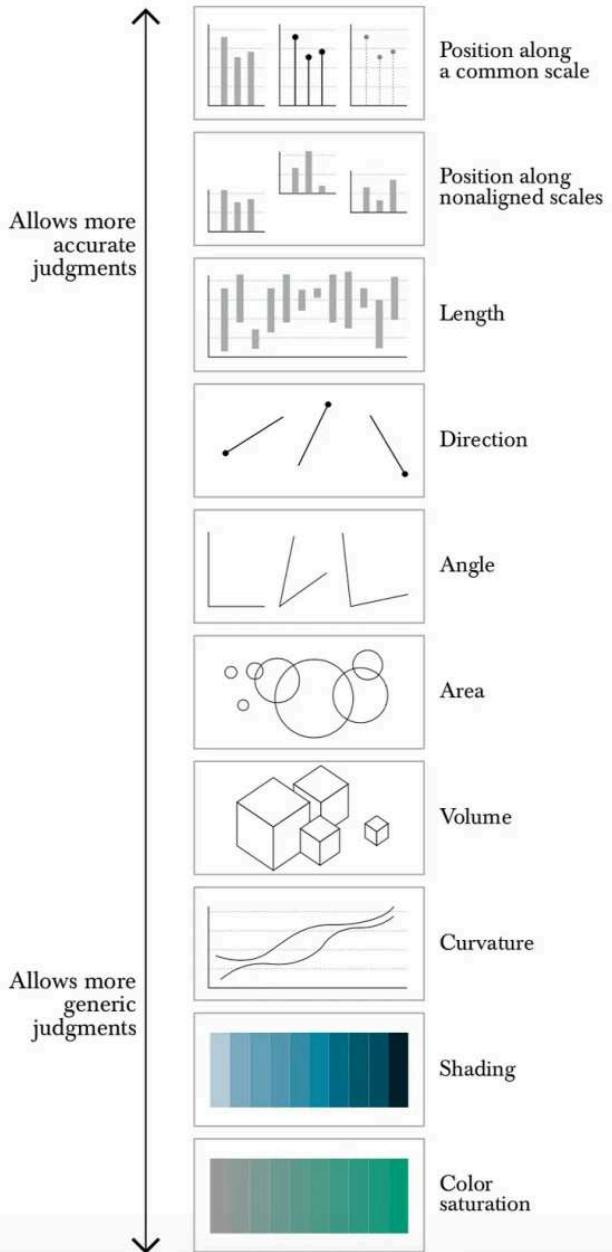
Source: [Data to Viz](#)





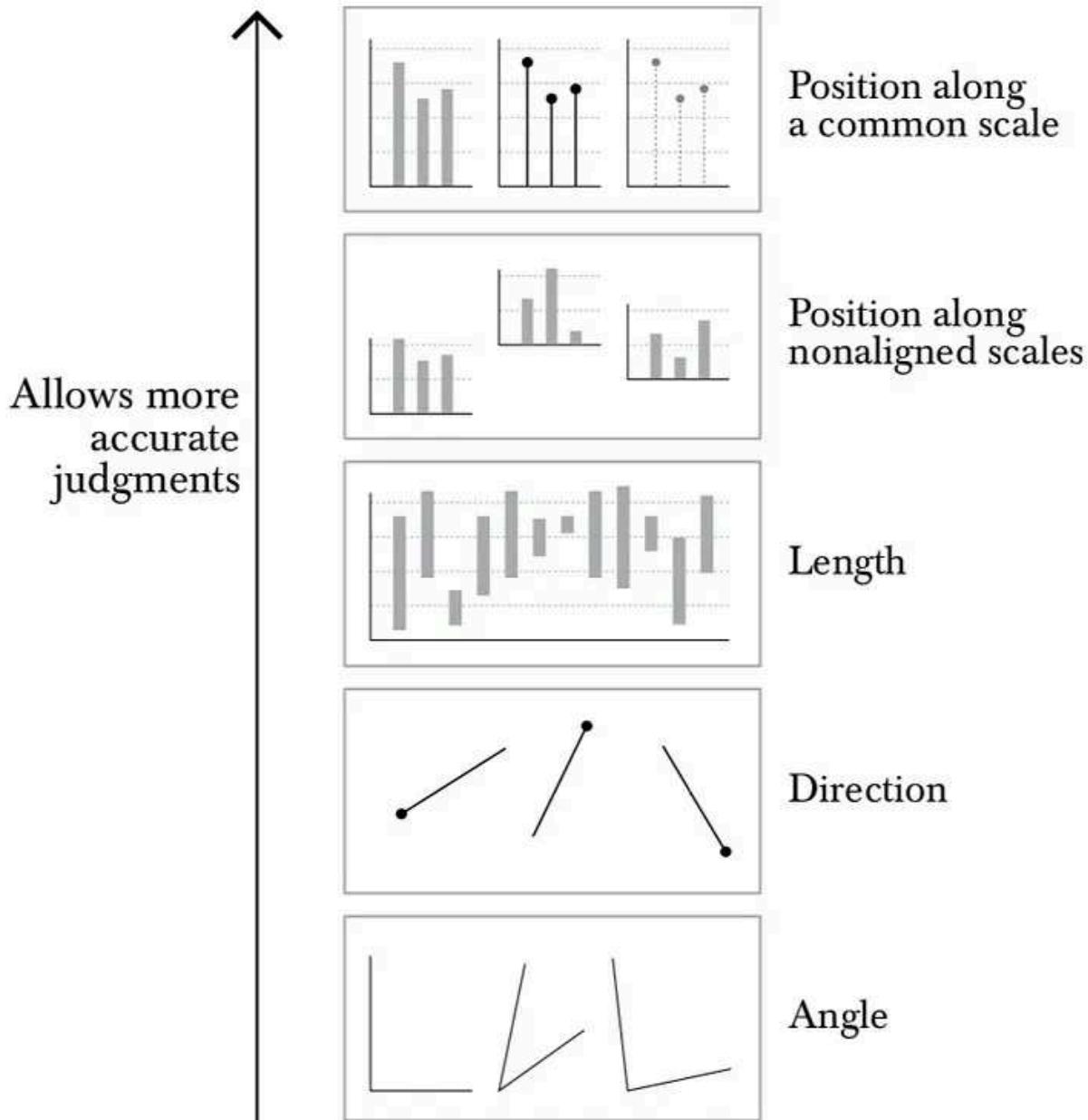
Source: "Data Points" by Nathan Yau (p. 104)





Cleveland's and McGill's Scale of Perpetual Elementary Tasks
 Source: "The Functional Art" by Alberto Cairo (p. 120)

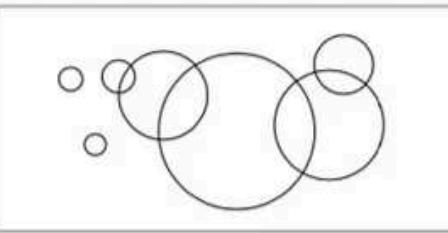




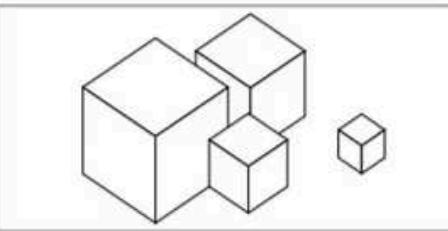
Cleveland's and McGill's Scale of Perpetual Elementary Tasks
Source: "The Functional Art" by Alberto Cairo (p. 120)



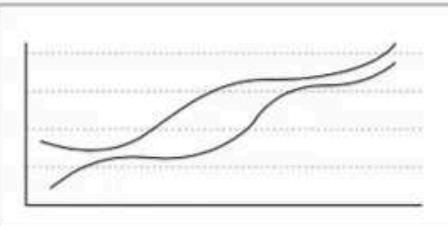
Allows more generic judgments



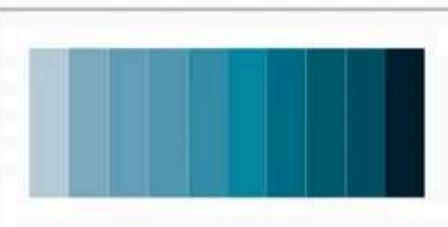
Area



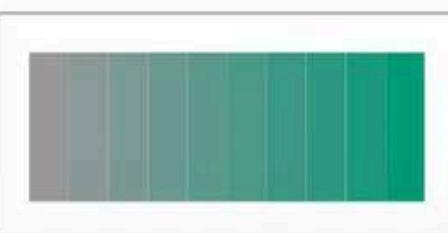
Volume



Curvature



Shading

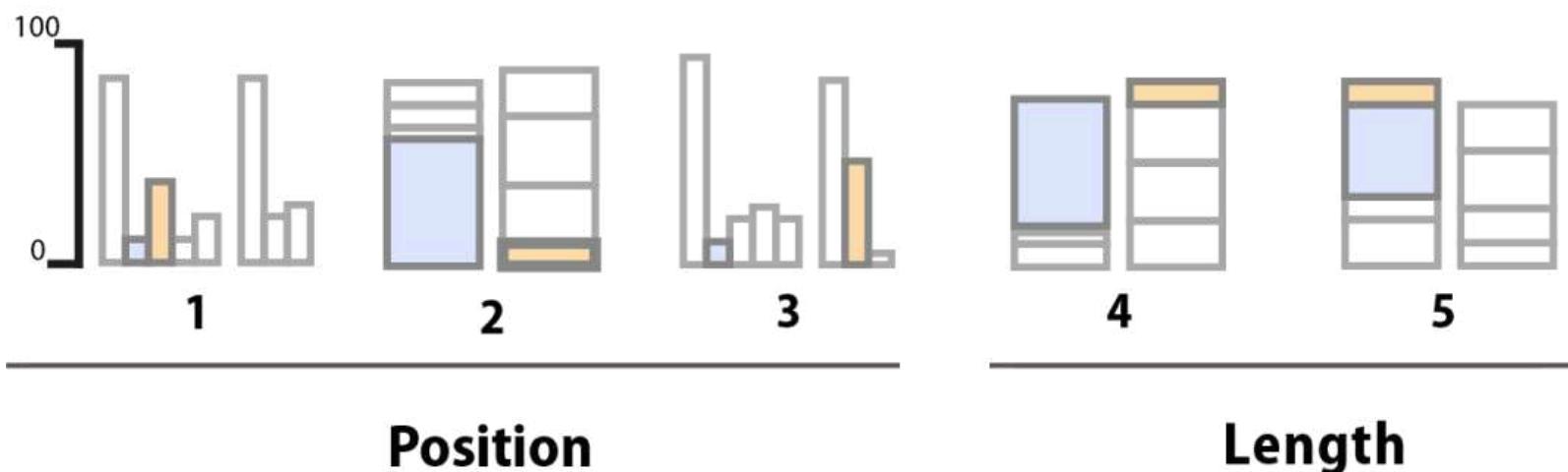
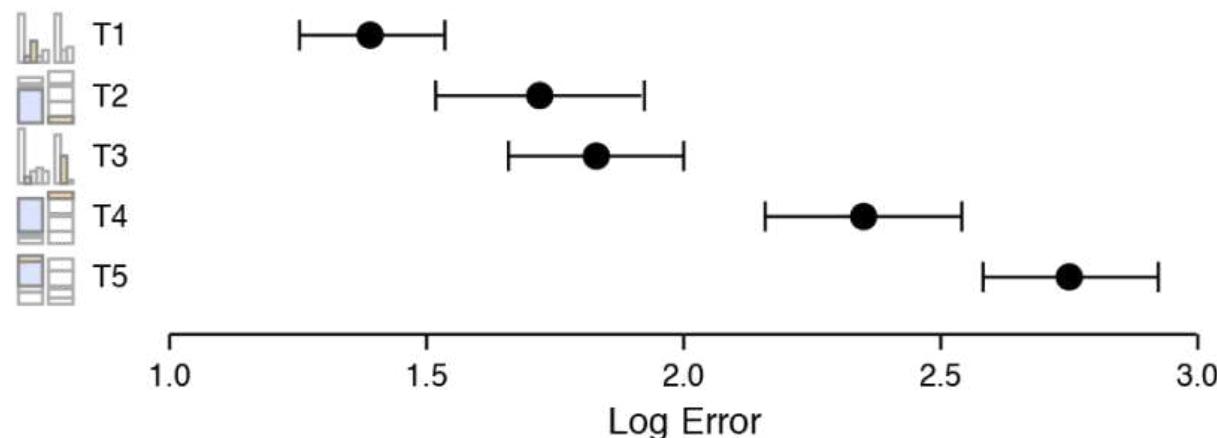


Color saturation

Cleveland's and McGill's Scale of Perceptual Elementary Tasks
Source: "The Functional Art" by Alberto Cairo (p. 120)



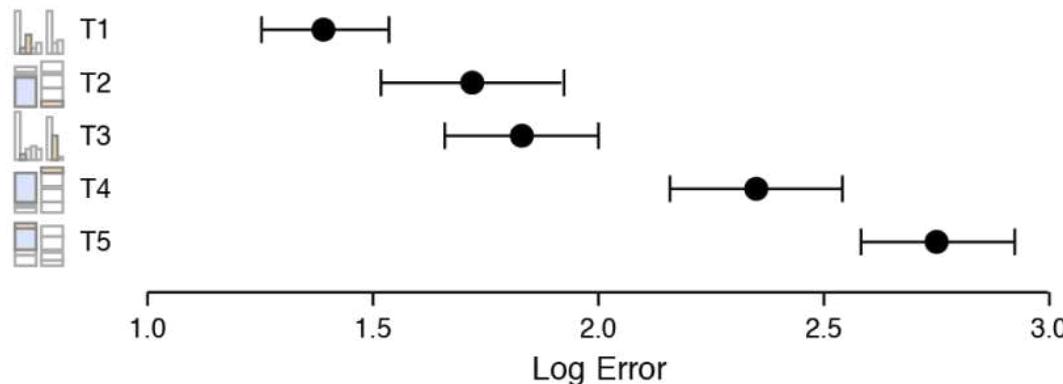
Cleveland & McGill's Results



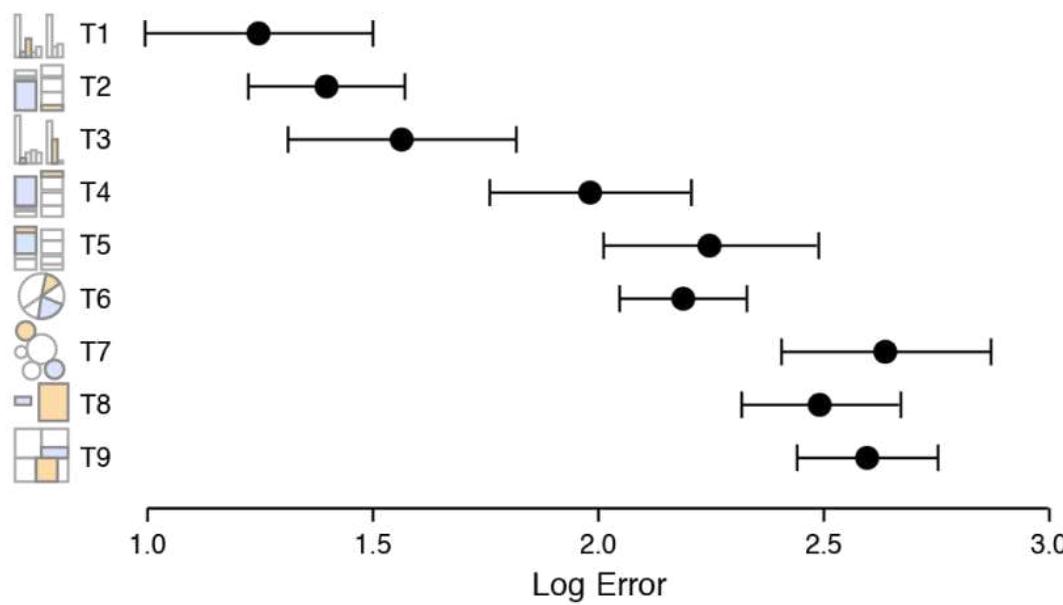
Source: "Data Visualization: A Practical Introduction" by Kieran Healy
Results based on Cleveland and McGill (1984, 1987)



Cleveland & McGill's Results



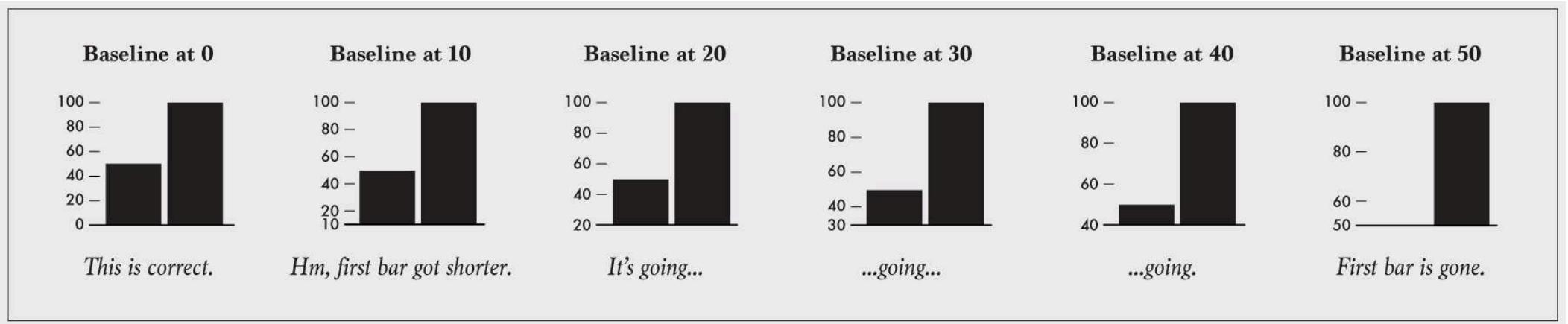
Crowdsourced Results



Source: "Data Visualization: A Practical Introduction" by Kieran Healy
Results based on Cleveland and McGill (1984, 1987) and Heer & Bostock (2010)



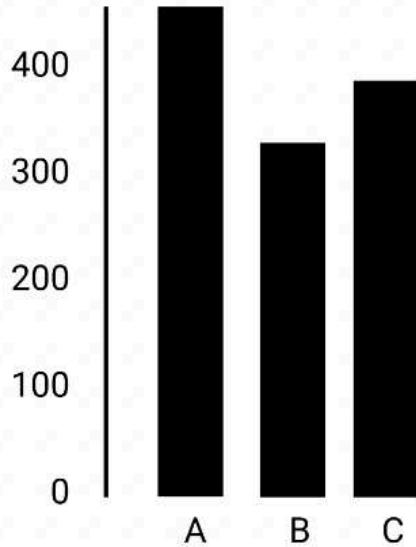
Always Start at Zero?



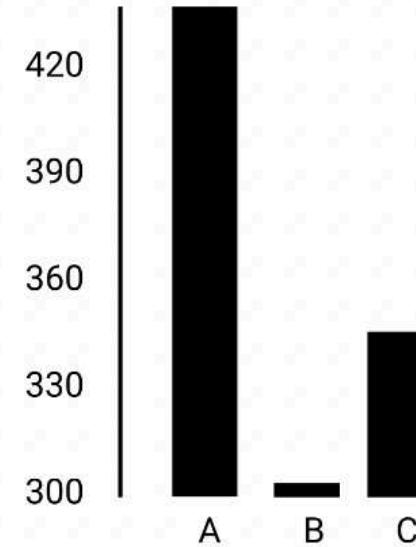
Source: Nathan Yau (flowingdata.com)



Always Start at Zero?



*Good. Y-axis starts at 0
and has natural increments*



*Bad. Y-axis starts at 300,
exaggerating the difference
among columns. Intervals are
awkward*

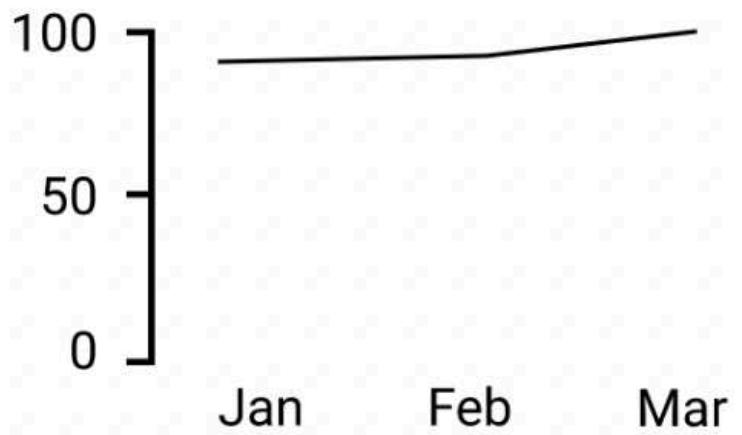
Source: "Hands-On Data Visualization" by Jack Dougherty & Ilya Ilyankou

→ more on the topic: e.g. Correl, Bertini & Franconeri 2020 and Witt (2019)

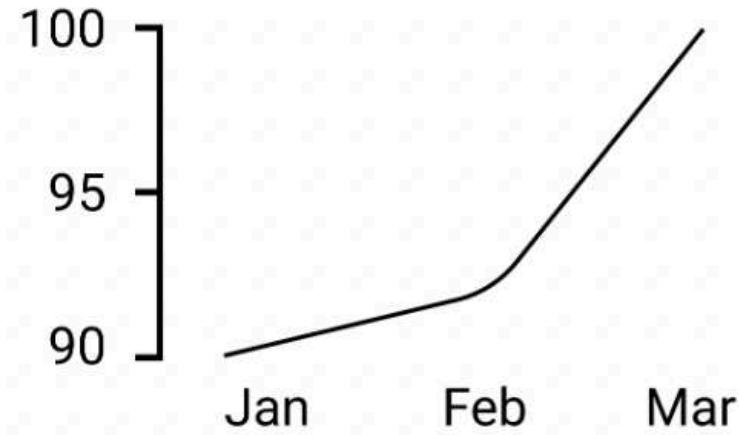
Cédric Scherer Data Visualization & Information Design



Always Start at Zero?



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

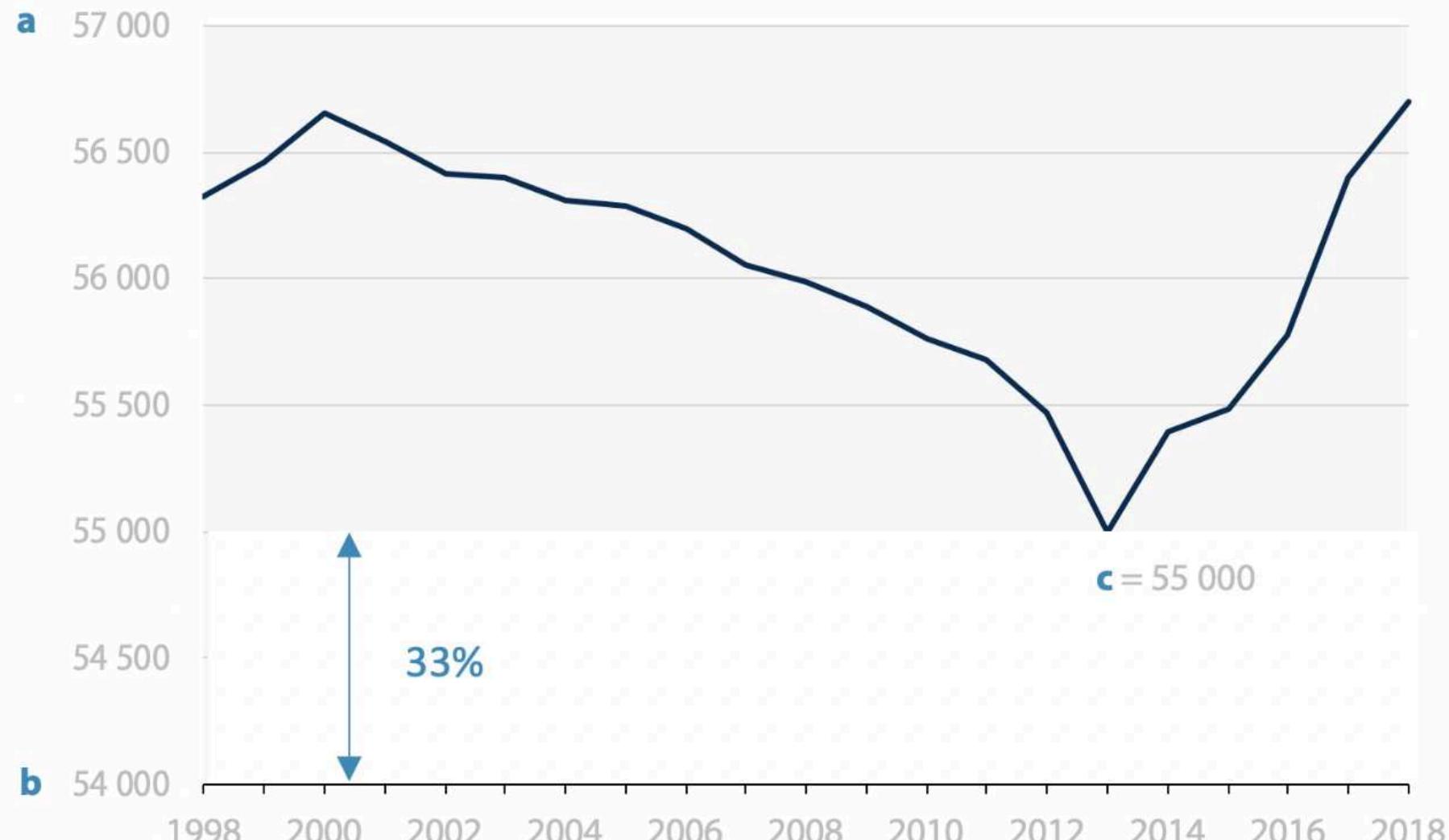


Better: By reducing the vertical axis to match the values, we see change more clearly

Source: "Hands-On Data Visualization" by Jack Dougherty & Ilya Ilyankou



Sales of widgets

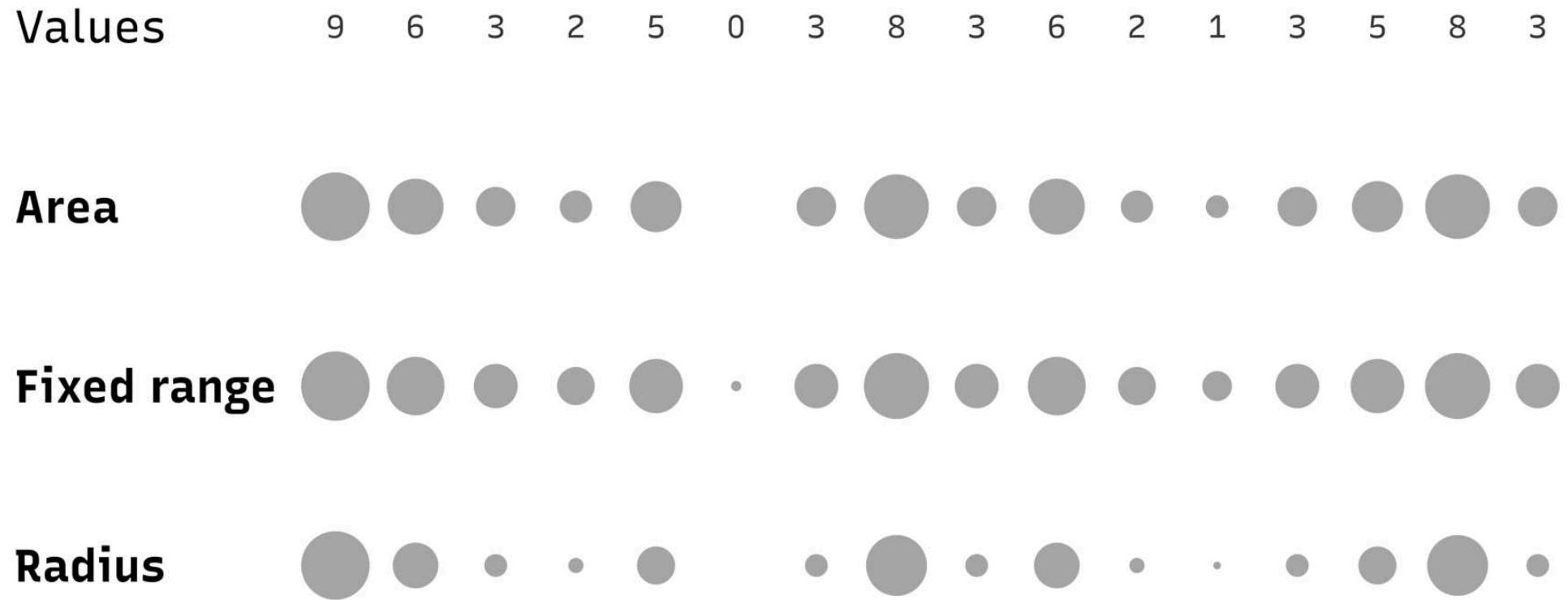


Source: Francis Gagnon

Cédric Scherer Data Visualization & Information Design

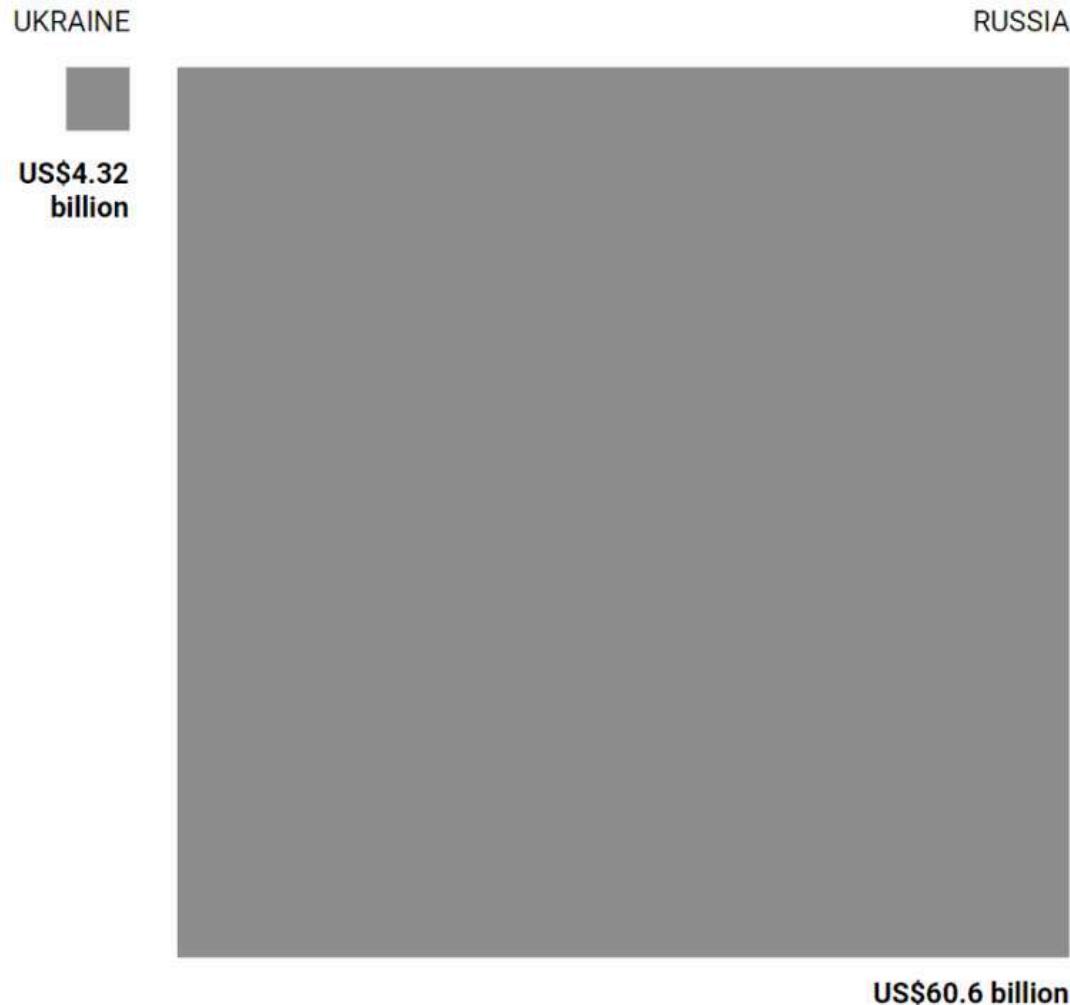


Same Data, Different Scaling



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.



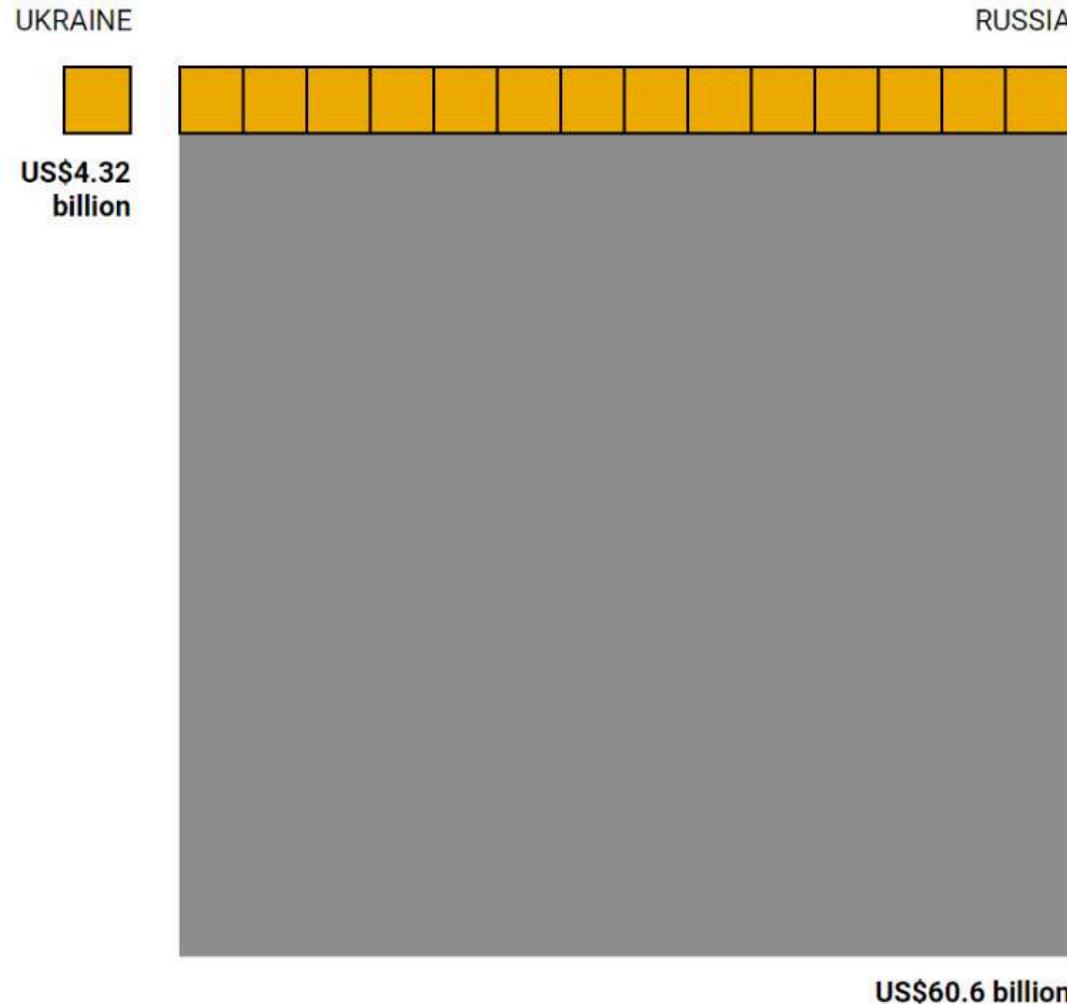
"Russia attacks Ukraine" by SCMP Graphic (South China Morning Post)

Cédric Scherer Data Visualization & Information Design



DEFENCE BUDGETS: RUSSIA VS UKRAINE (2020)

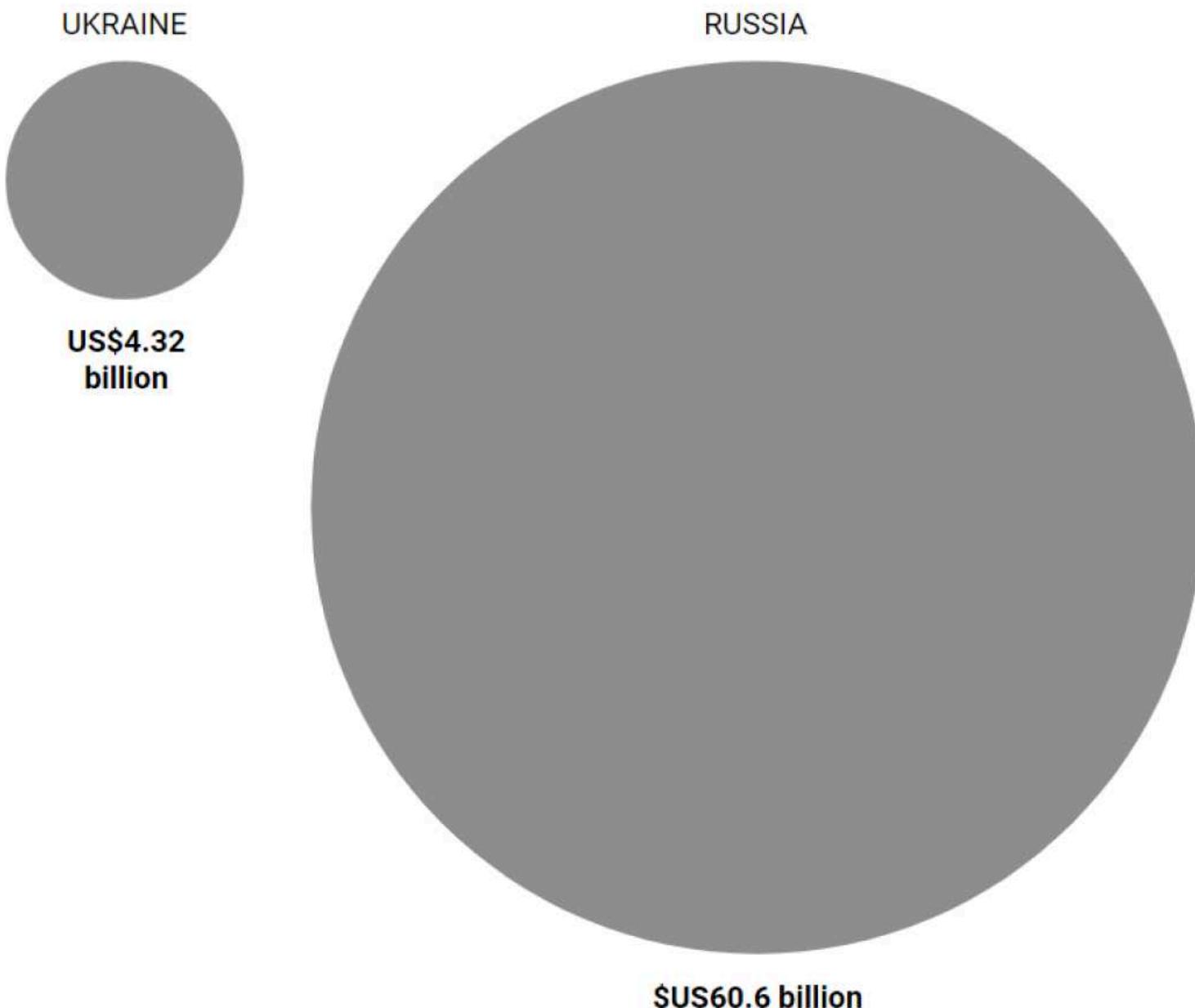
The national balance of forces is overwhelmingly in Russia's favour. Russian military spending in 2020 amounted to US\$60.6 billion in 2020. Ukraine's was less than a 10th of that amount.



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

Cédric Scherer Data Visualization & Information Design





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

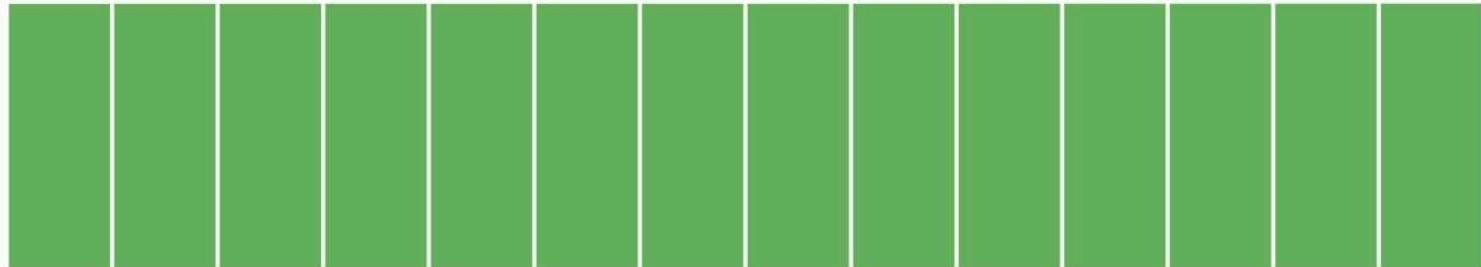
Cédric Scherer Data Visualization & Information Design





UKRAINE

US\$4.3 billion



RUSSIA

US\$60.6 billion



**Ukraine's military spending was
less than a 14th of Russia's in 2020.**



Same Data, Different Palettes

Values

5 0 3 8 3 6 2 1 3 5 8 3

Sequential (desaturated)



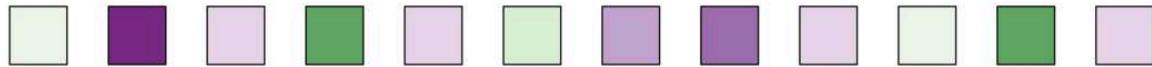
Sequential (single-hue)



Sequential (multi-hue)

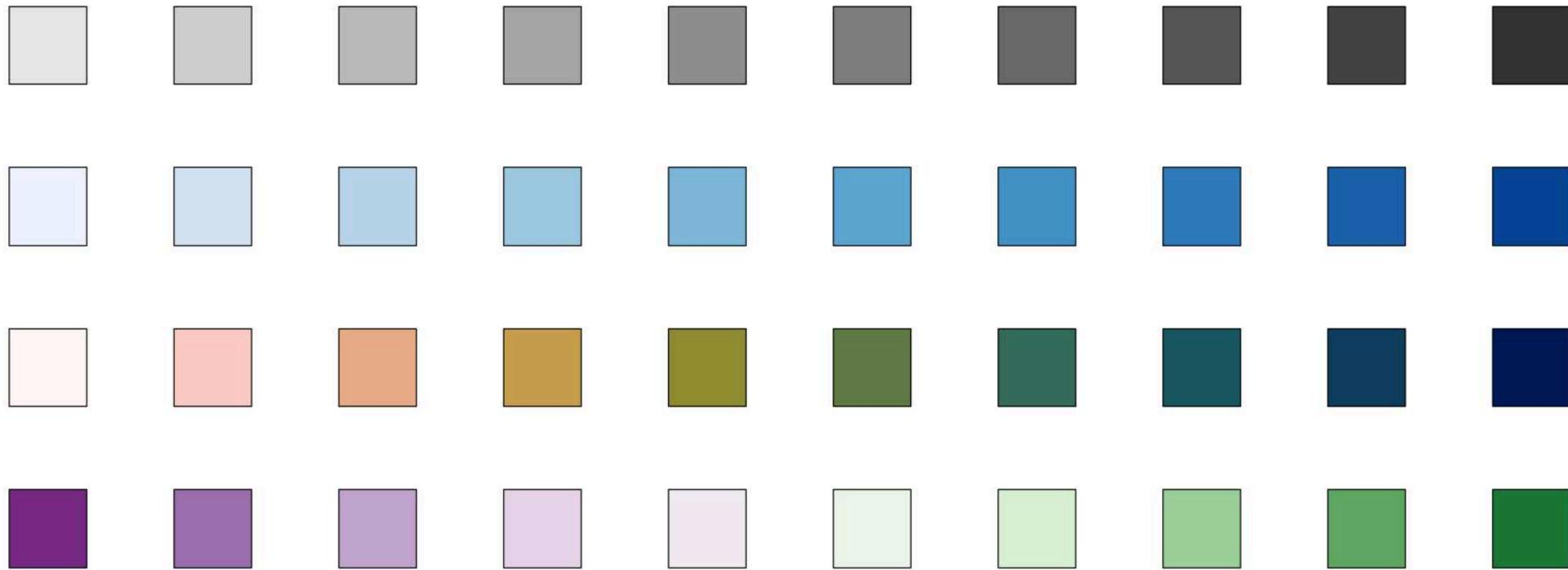


Diverging



Same Data, Different Palettes

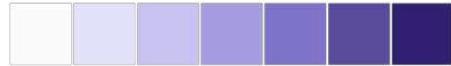
0 1 2 3 4 5 6 7 8 9



Palette Types

Sequential —————

Palette



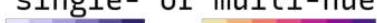
Desaturated



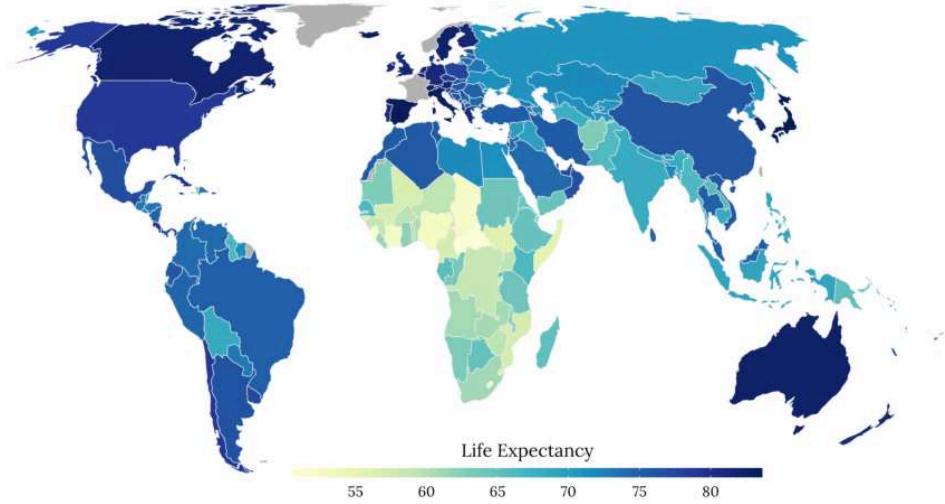
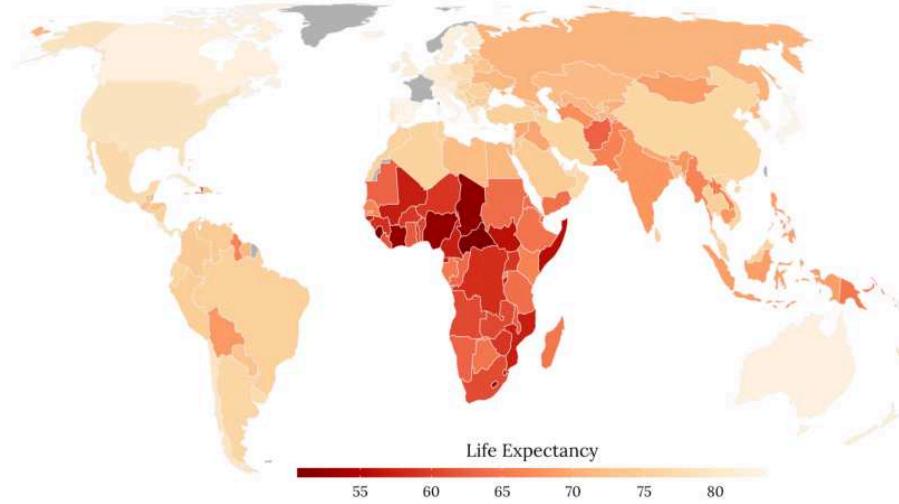
Use to encode
**numerical information
with order**

use highest contrast for
most important information

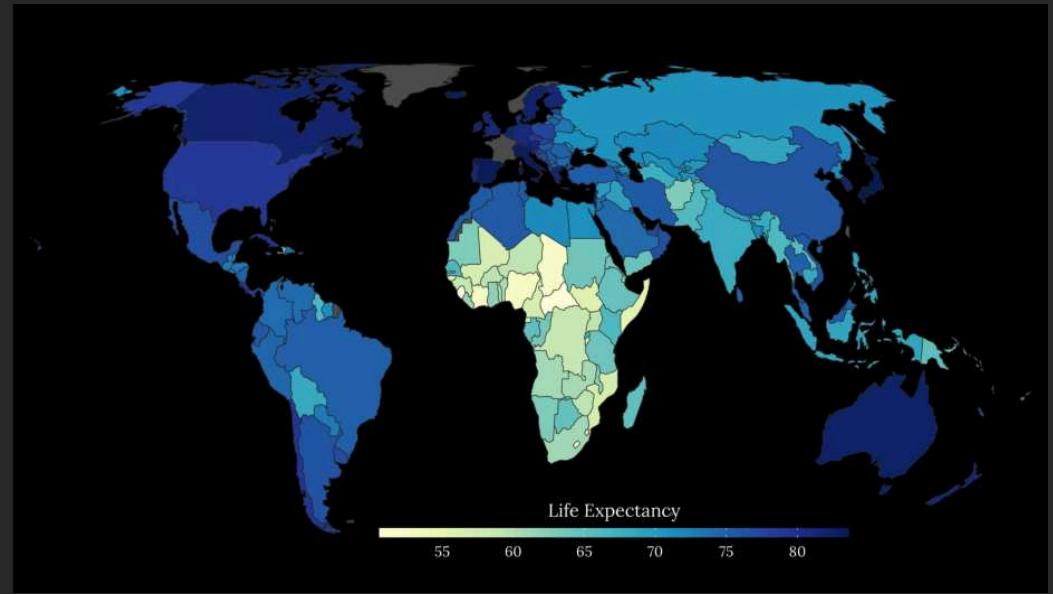
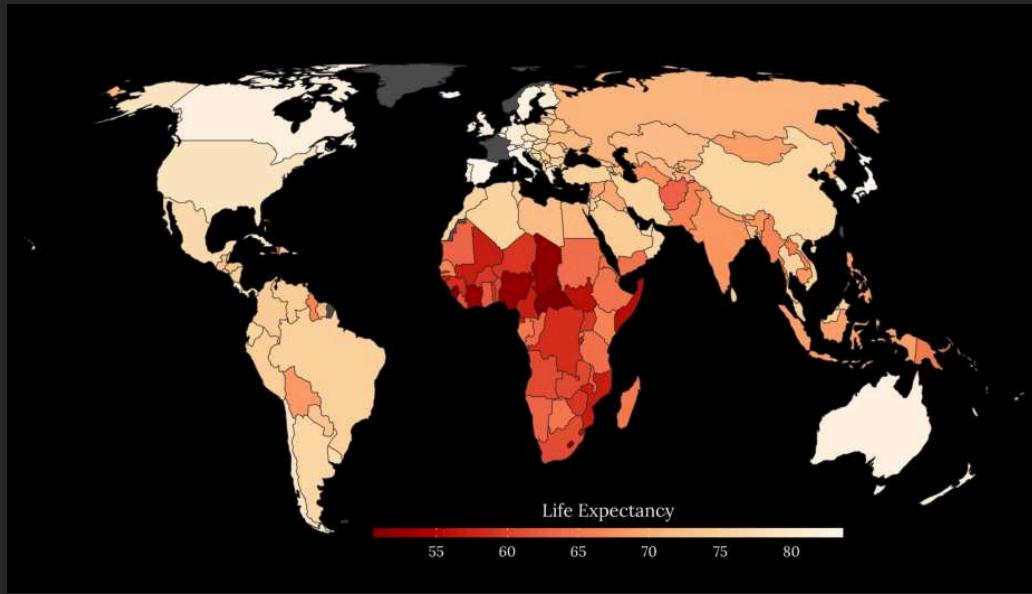
either single- or multi-hue



Sequential: Dark Equals More (?)



Sequential: Dark Equals More (?)



Palette Types

Sequential



Use to encode
**numerical information
with order**

use highest contrast for
most important information

either single- or multi-hue



Diverging



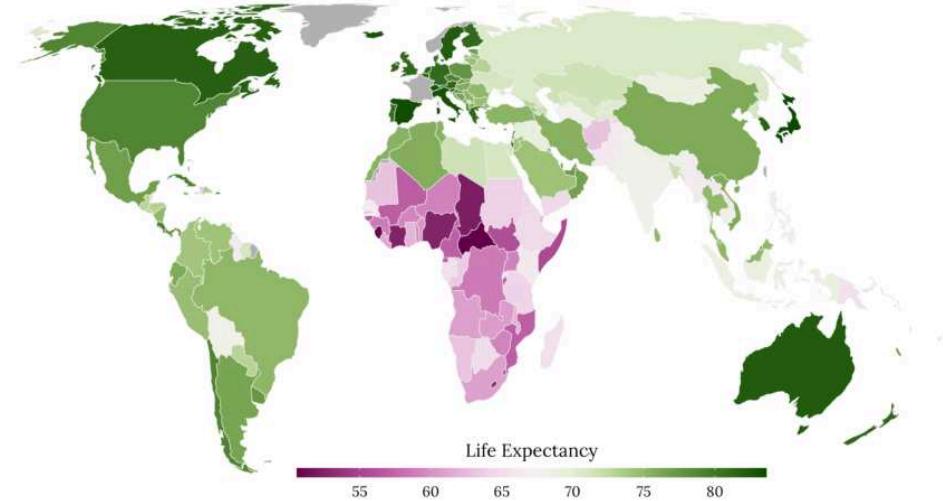
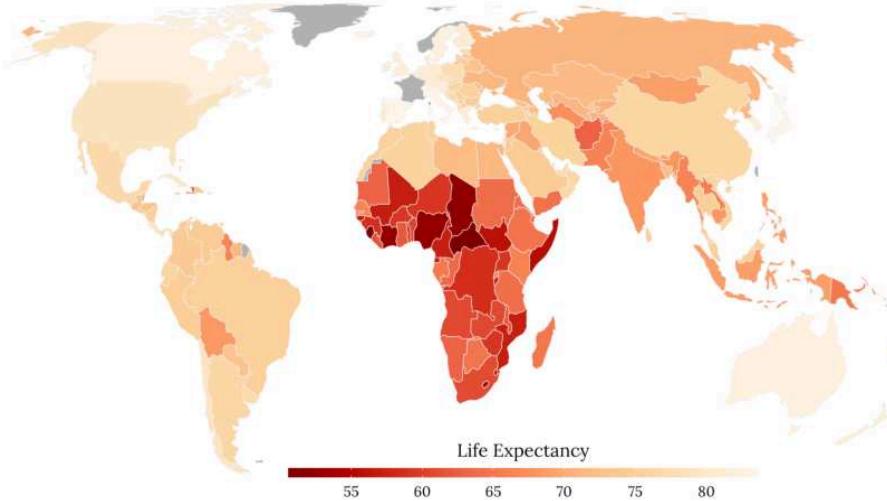
Use to encode
**numerical information
with critical midpoint**

ensure a meaningful midpoint value
and use balanced extremes

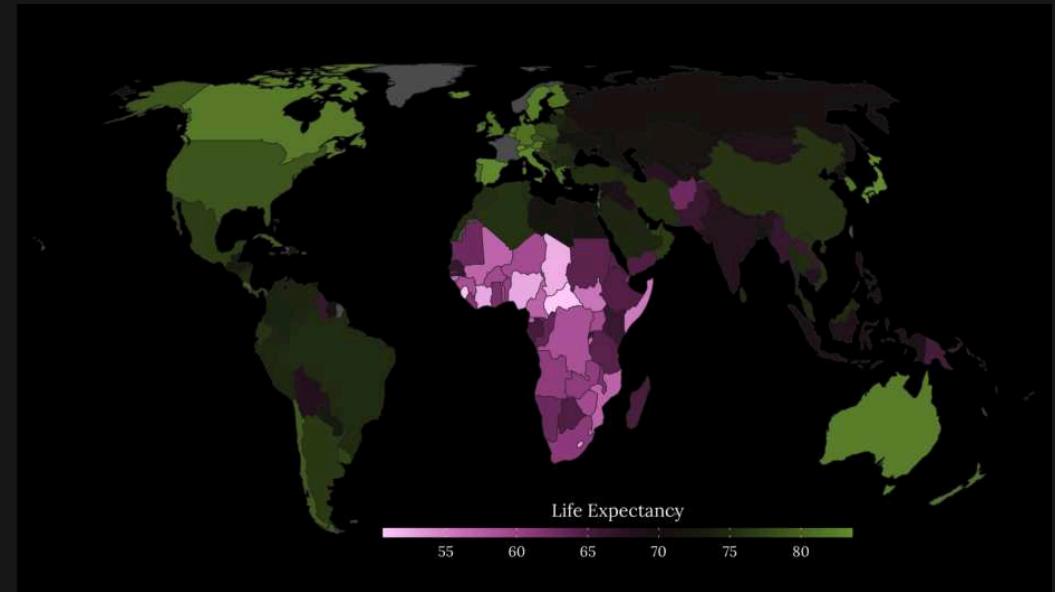
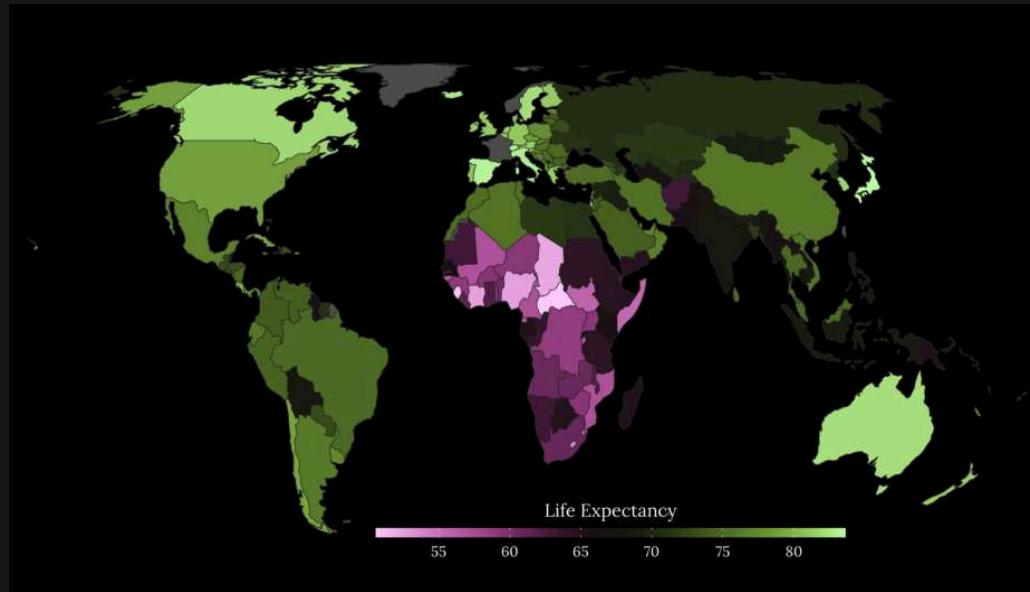
combination of
two sequential palettes



Sequential or Diverging?



Midpoint of Diverging Palettes



Palette Types

Sequential

Palette



Desaturated



Use to encode
**numerical information
with order**

use highest contrast for
most important information

either single- or multi-hue

Diverging

Palette



Desaturated



Use to encode
**numerical information
with critical midpoint**

ensure a meaningful midpoint value
and use balanced extremes

combination of
two sequential palettes

Qualitative

Palette



Desaturated



Use to encode
categorical information

pick distinct colors with
the same perceptual weight

limit categories to 5-8



Sequential or Qualitative!

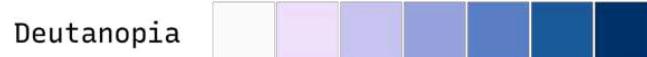
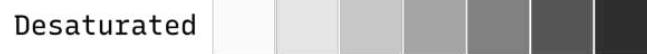
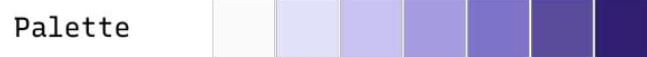


Source: "What to consider when choosing colors for data visualization" by Lisa Charlotte Muth, Datawrapper Blog



Palette Types & Colorblindness

Sequential



Use to encode
**numerical information
with order**

use highest contrast for
most important information

either single- or multi-hue

Diverging



Use to encode
**numerical information
with critical midpoint**

ensure a meaningful midpoint value
and use balanced extremes

combination of
two sequential palettes

Qualitative



Use to encode
categorical information

pick distinct colors with
the same perceptual weight

limit categories to 5-8



Design with the Colorblind in Mind

original



deuteranomaly



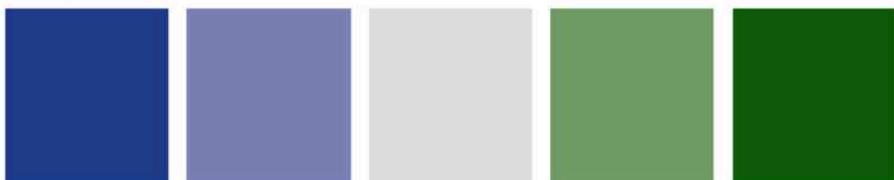
protanomaly



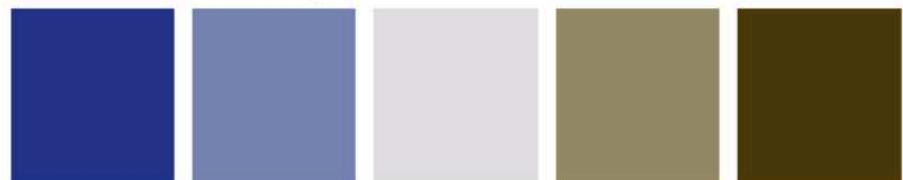
tritanomaly



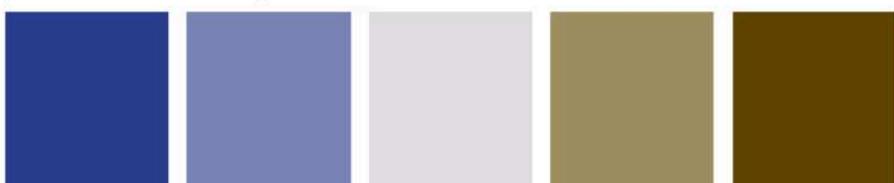
original



deuteranomaly



protanomaly



tritanomaly



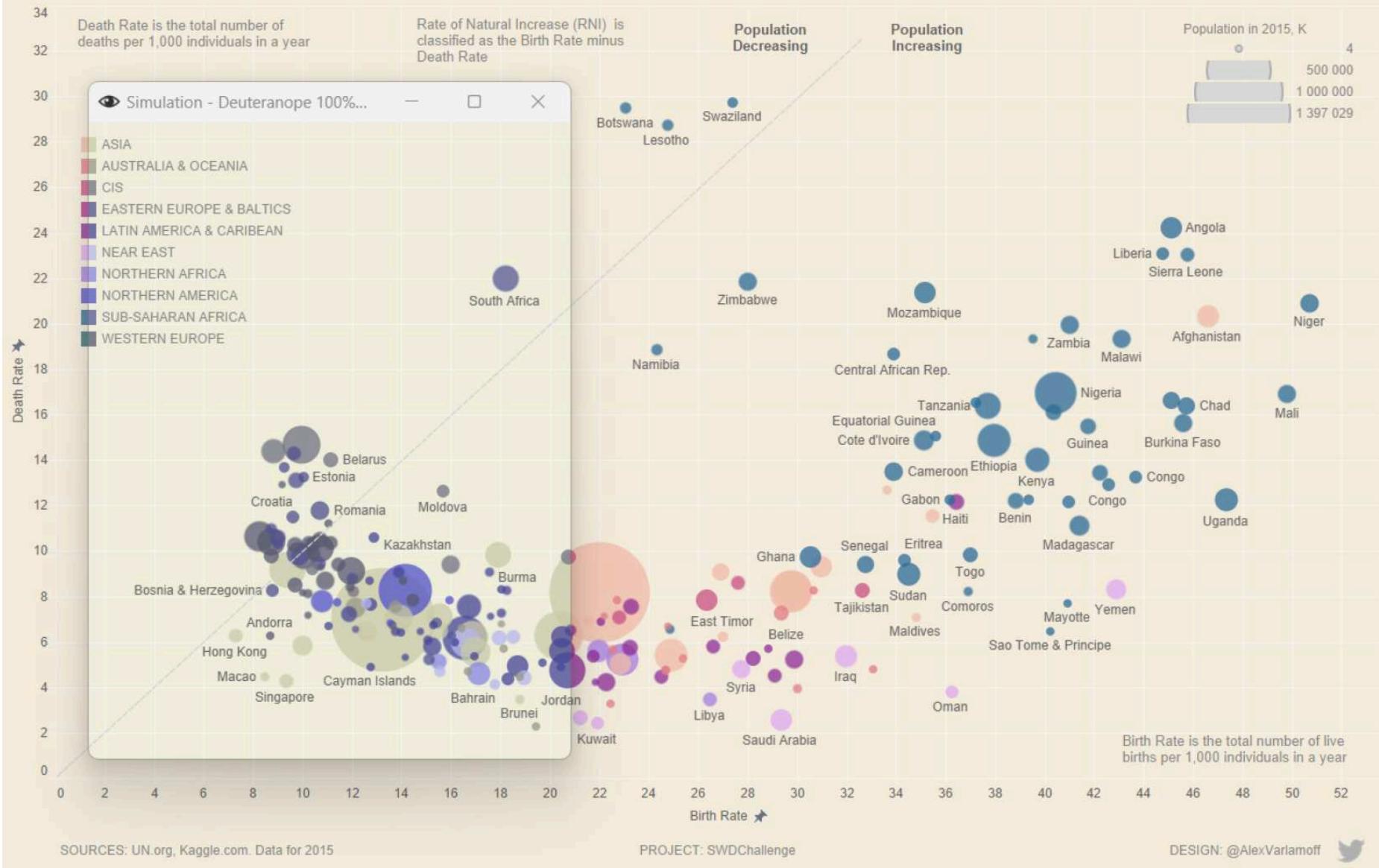
Source: "Fundamentals of Data Visualization" by Claus O. Wilke



Natural Increase in the World

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

Natural Increase
35,65

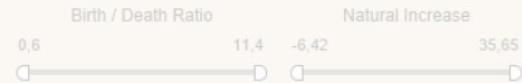


Contribution to the SWDchallenge by Alex Varlamoff – tested with a color-blindness simulator

Cédric Scherer Data Visualization & Information Design



Natural Increase in the World



Death Rate is the total number of deaths per 1,000 individuals in a year

Rate of Natural Increase (RNI) is classified as the Birth Rate minus Death Rate

Population Decreasing

Population Increasing

Population in 2015, K

4

500 000

1 000 000

1 397 029

Simulation - Deutanope 100%

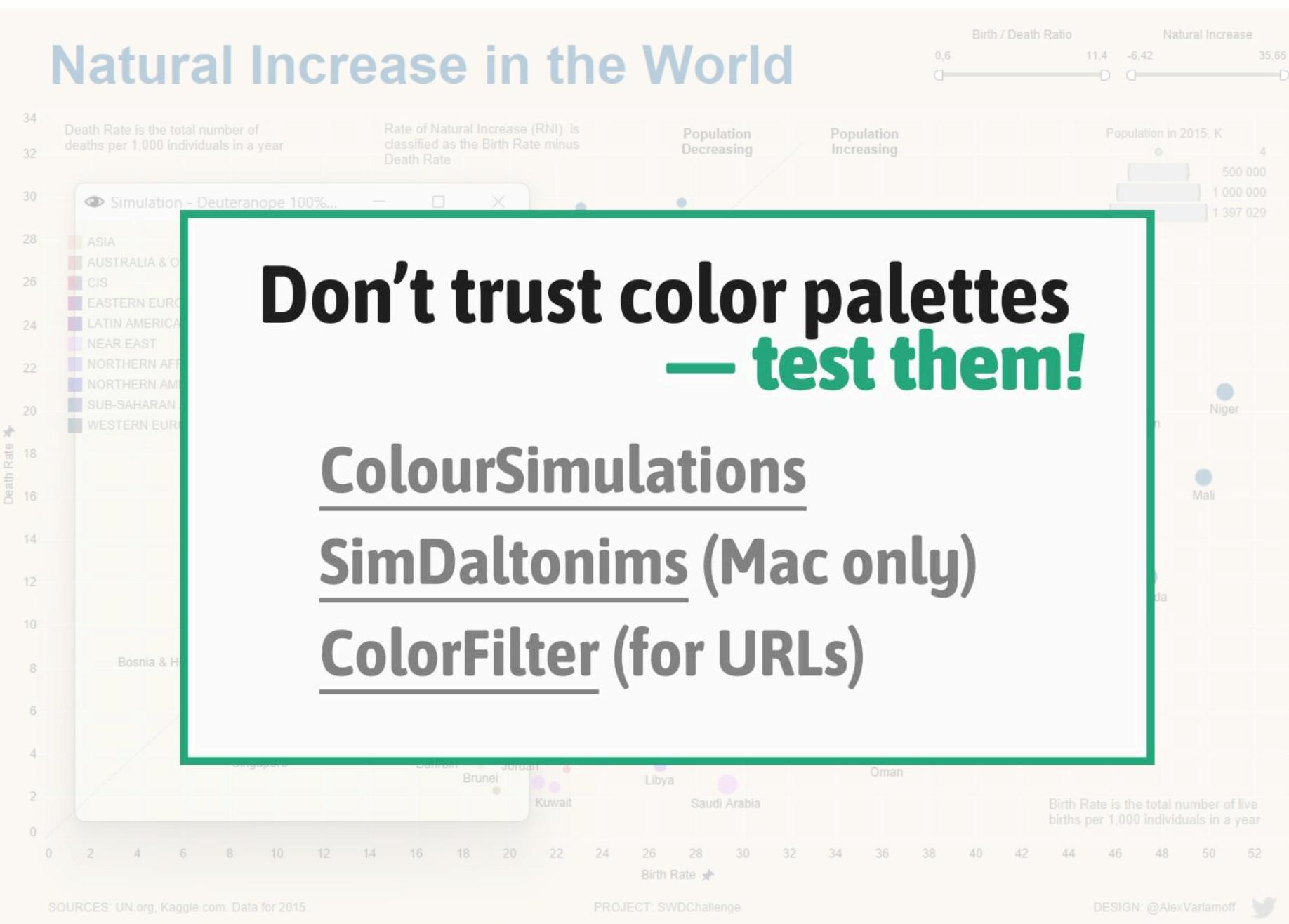
ASIA
AUSTRALIA & O.
CIS
EASTERN EURO.
LATIN AMERICA
NEAR EAST
NORTHERN AFR.
NORTHERN AM.
SUB-SAHARAN
WESTERN EU.

Don't trust color palettes — test them!

ColourSimulations

SimDaltonims (Mac only)

ColorFilter (for URLs)



Contribution to the SWDchallenge by Alex Varlamoff – tested with a color-blindness simulator

Cédric Scherer Data Visualization & Information Design



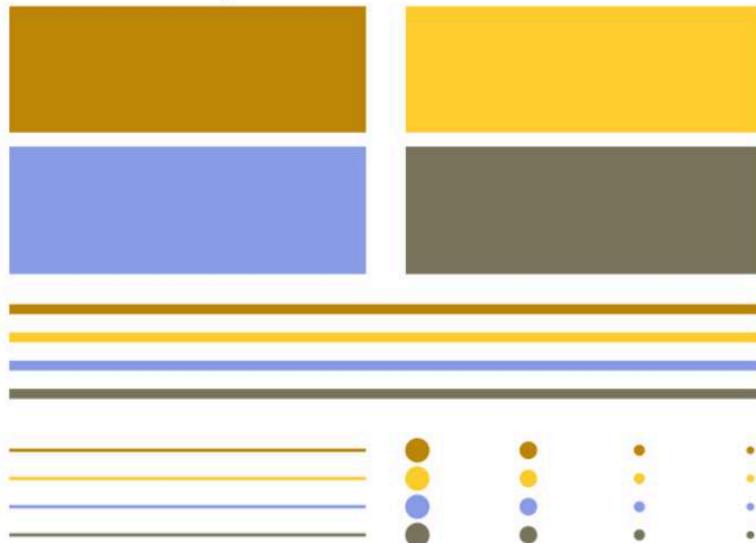
original



deuteranomaly



protanomaly



tritanomaly



Source: "Fundamentals of Data Visualization" by Claus O. Wilke

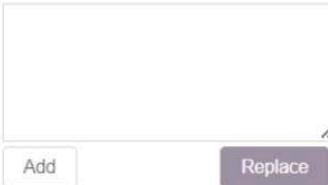


VIZ PALETTE

By: Elijah Meeks
& Susie Lu

PICK

Use Chroma.js



Use Colorgorical

Use ColorBrewer

EDIT

4 Colors

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

hex rgb

hsl

GET

hex rgb

hsl

- String quotes
 Object with metadata

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

COLORS IN ACTION

Background color: #eeeeee [edit](#)

Font color: ● #212121 [edit](#)

Charts made with [Semiotic](#)

Color Population:

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



The [VizPalette](#) tool tests color palettes for legibility for different chart types, line weights, backgrounds, and font colors.

Cédric Scherer Data Visualization & Information Design

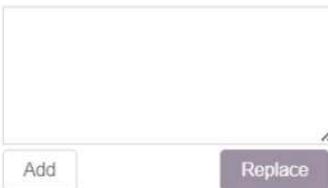


VIZ PALETTE

By: Elijah Meeks
& Susie Lu

PICK

Use Chroma.js



Use Colorgorical

Use ColorBrewer

EDIT

4 Colors

- ≡ 1 ● #2a9571
- ≡ 2 ● #8fb9bf
- ≡ 3 ● #dfb468
- ≡ 4 ● #4b8cd8

Add

hex rgb

hsl

GET

hex rgb

hsl

- String quotes
 Object with metadata

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

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



The [VizPalette](#) tool tests color palettes for legibility for different chart types, line weights, backgrounds, and font colors.

Cédric Scherer Data Visualization & Information Design



VIZ PALETTE

By: Elijah Meeks
& Susie Lu

PICK

Use Chroma.js



Use Colorgorical

Use ColorBrewer

EDIT

4 Colors

- ≡ 1 ● #2a9571
- ≡ 2 ● #8fb9bf
- ≡ 3 ● #dfb468
- ≡ 4 ● #4b8cd8

Add

hex rgb

hsl

GET

hex rgb

hsl

String quotes
 Object with metadata
["#2a9571",
"#8fb9bf",
"#dfb468",
"#4b8cd8"]

COLORS IN ACTION

Background color: #eeeeee

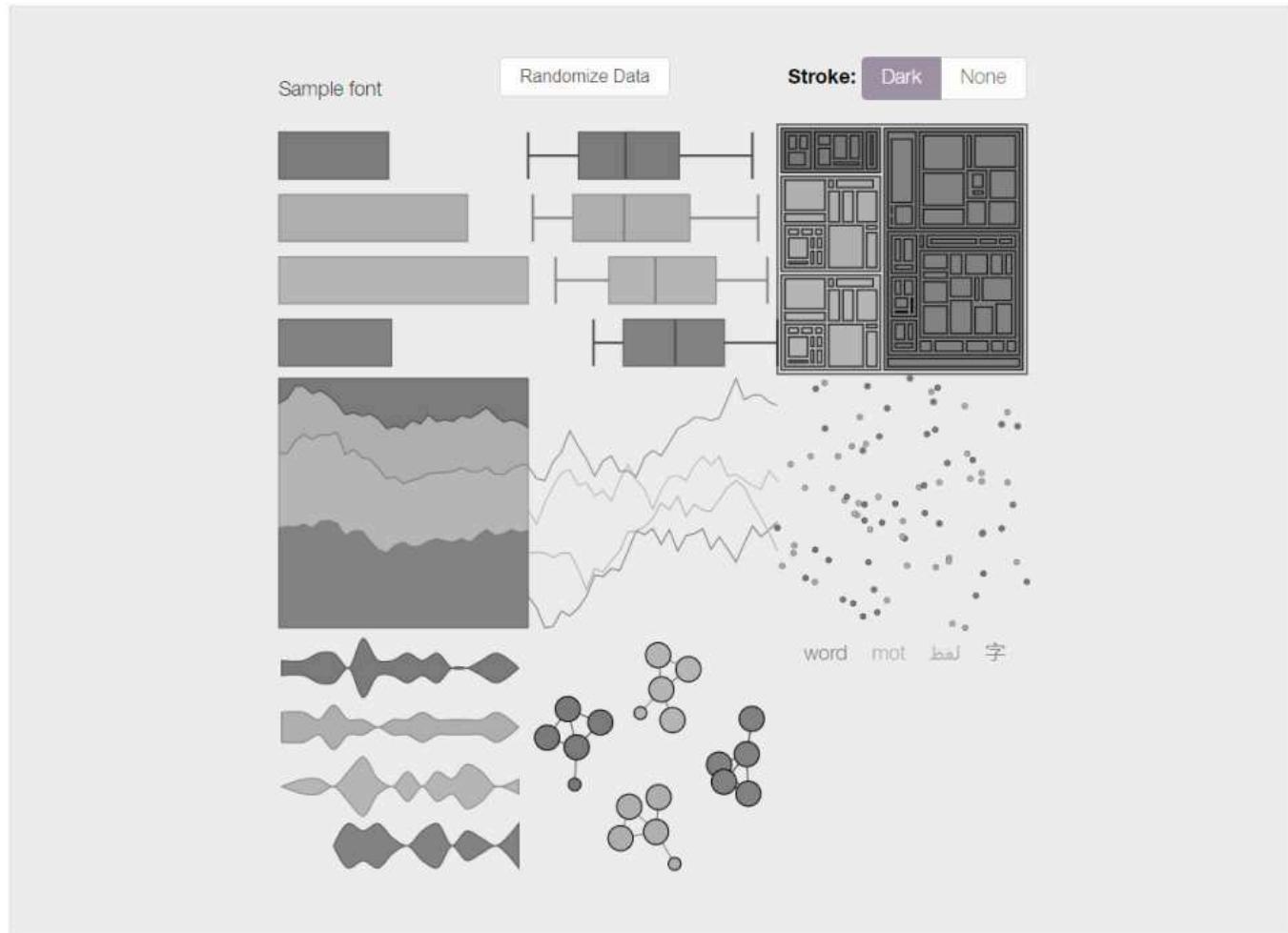
Font color: #212121

Charts made with [Semiotic](#)

Color Population:

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

Greyscale



The [VizPalette](#) tool tests color palettes for legibility for different chart types, line weights, backgrounds, and font colors.

Cédric Scherer Data Visualization & Information Design



VIZ PALETTE

By: Elijah Meeks
& Susie Lu

PICK

Use Chroma.js



Use Colorgorical

Use ColorBrewer

EDIT

4 Colors

- ≡ 1 ● #2a9571 [edit](#) x
- ≡ 2 ● #8fb9bf [edit](#) x
- ≡ 3 ● #dfb468 [edit](#) x
- ≡ 4 ● #4b8cd8 [edit](#) x

hex rgb

hsl

GET

hex rgb

hsl

- String quotes
- Object with metadata

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

COLORS IN ACTION

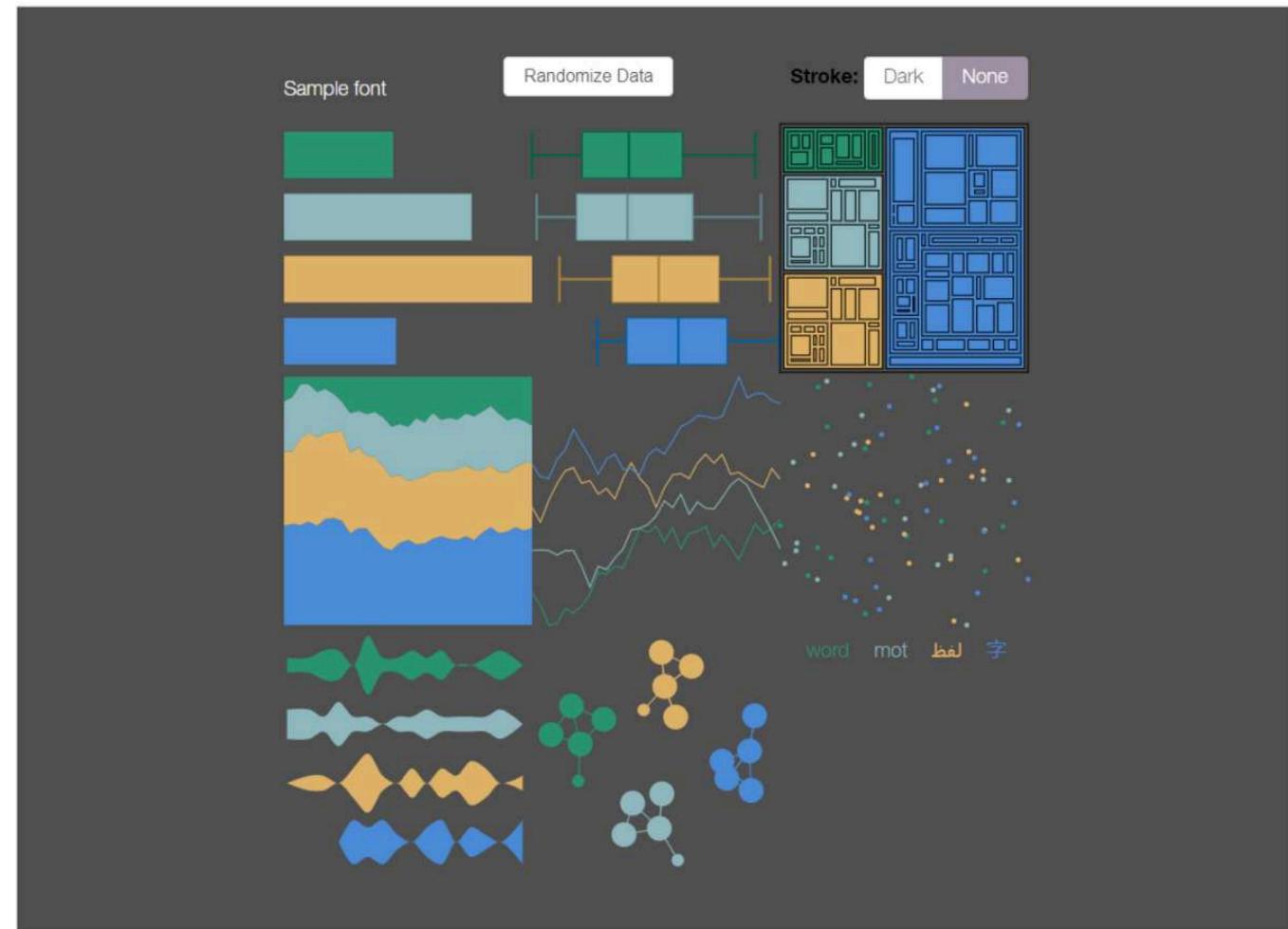
Background color: ● #505050 [edit](#)

Font color: #fefefe [edit](#)

Charts made with [Semiotic](#)

Color Population:

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



The [VizPalette](#) tool tests color palettes for legibility for different chart types, line weights, backgrounds, and font colors.

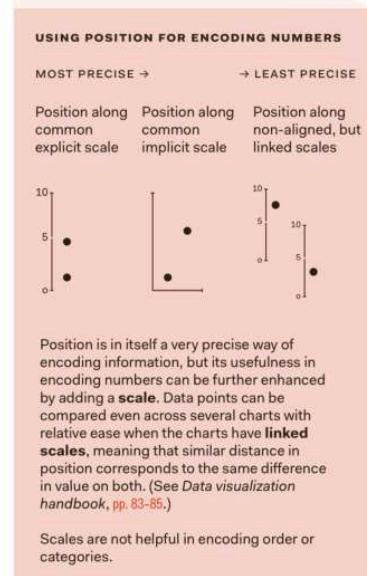
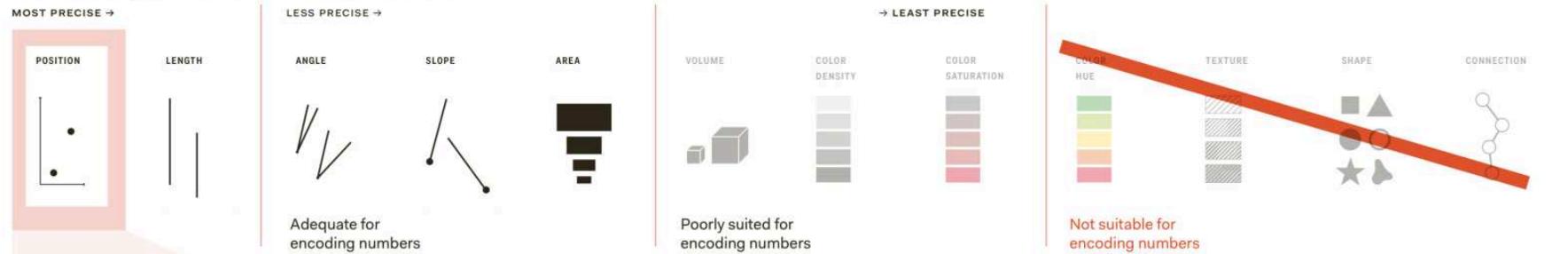
Cédric Scherer Data Visualization & Information Design



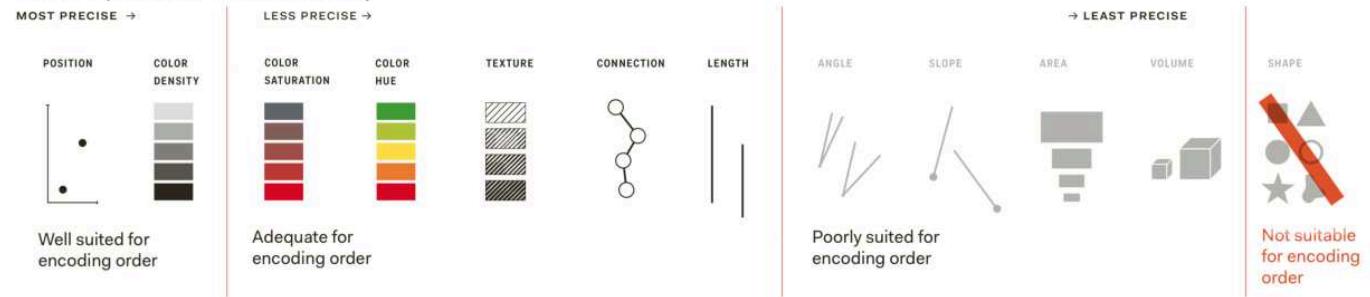
VISUAL VARIABLES

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

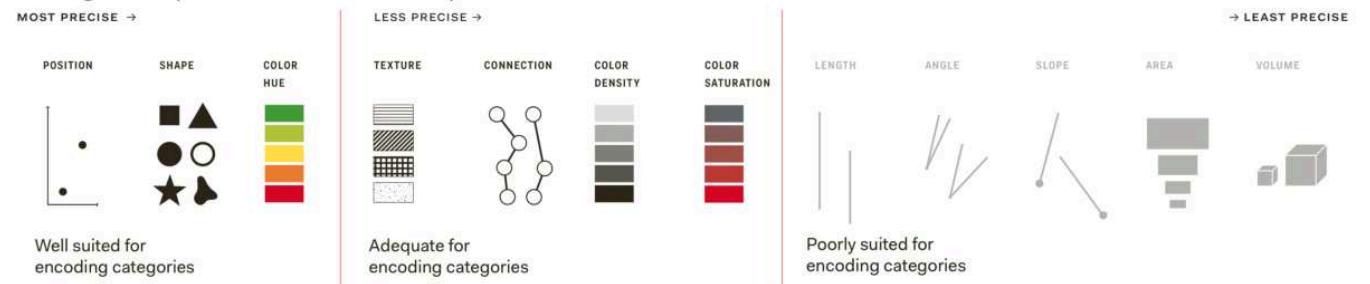
Numbers (data on ratio or interval scale)



Order (data on ordinal scale)



Categories (data on nominal scale)



Source: Koponen & Hildén, "Data Visualization Handbook" (2020)

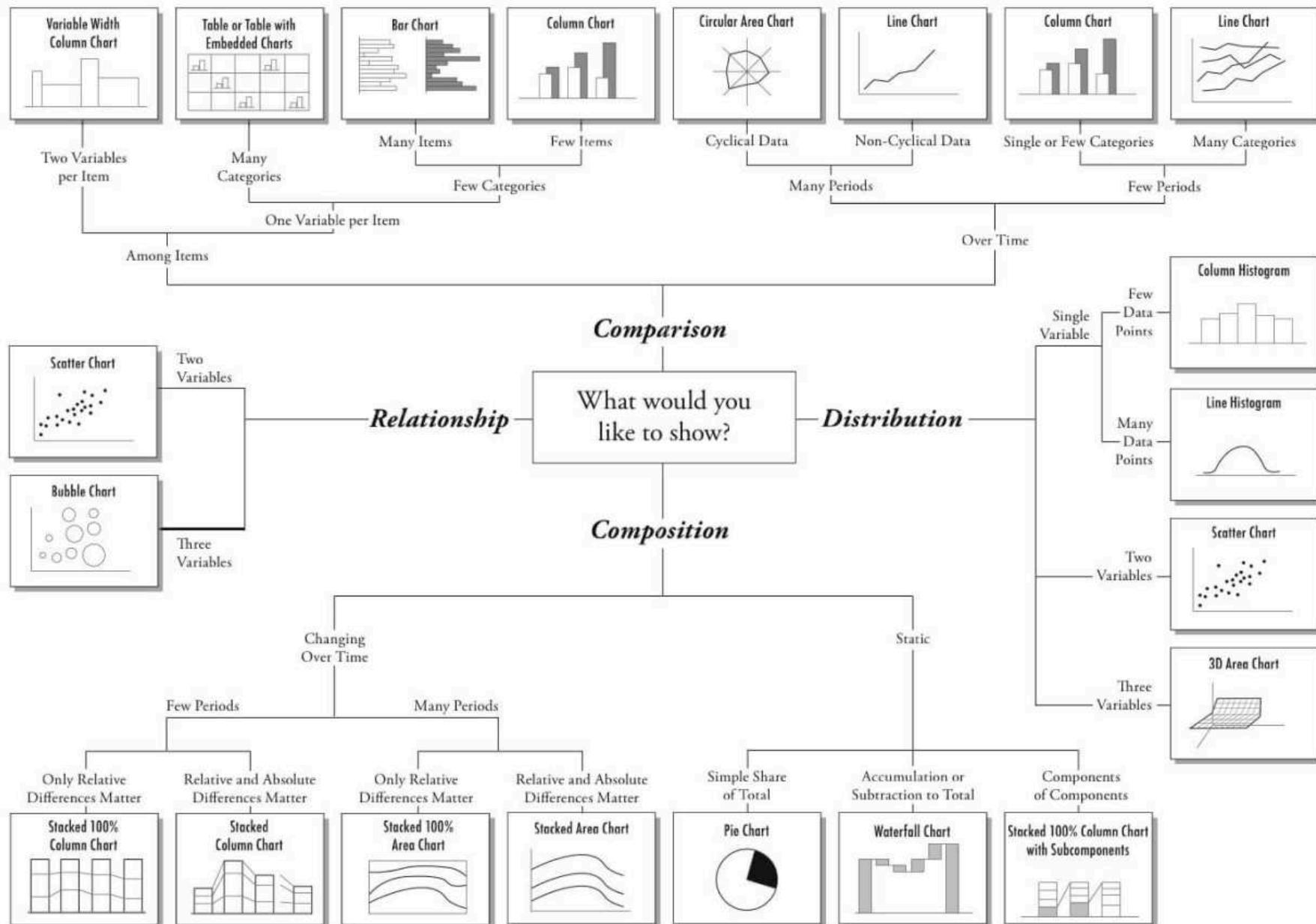


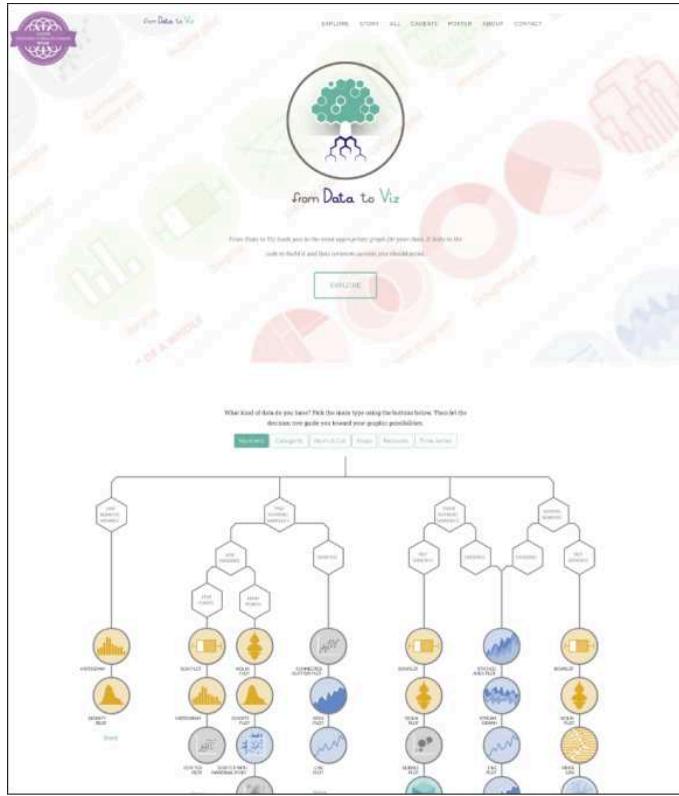
chart Type choice



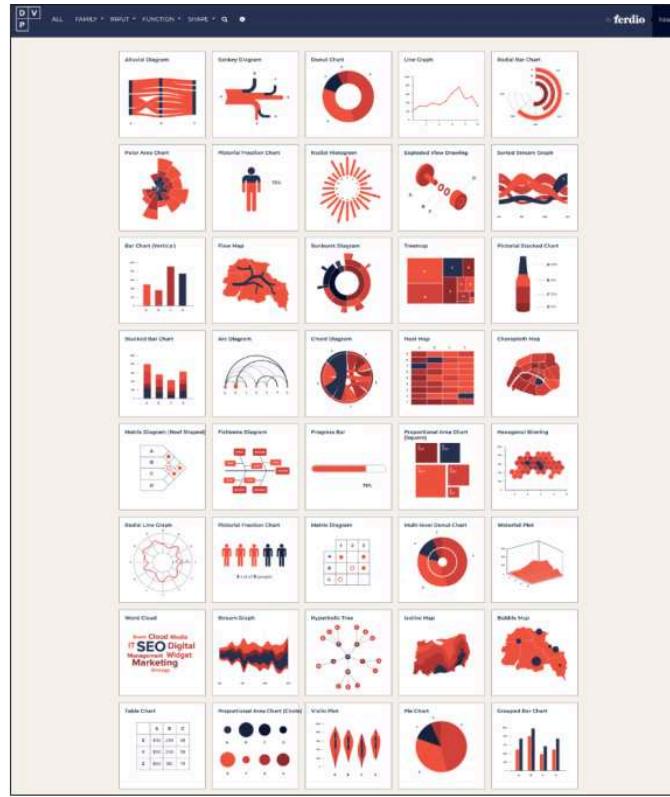
Chart Suggestions—A Thought-Starter

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





data-to-viz.com



datavizproject.com



visualizationuniverse.com





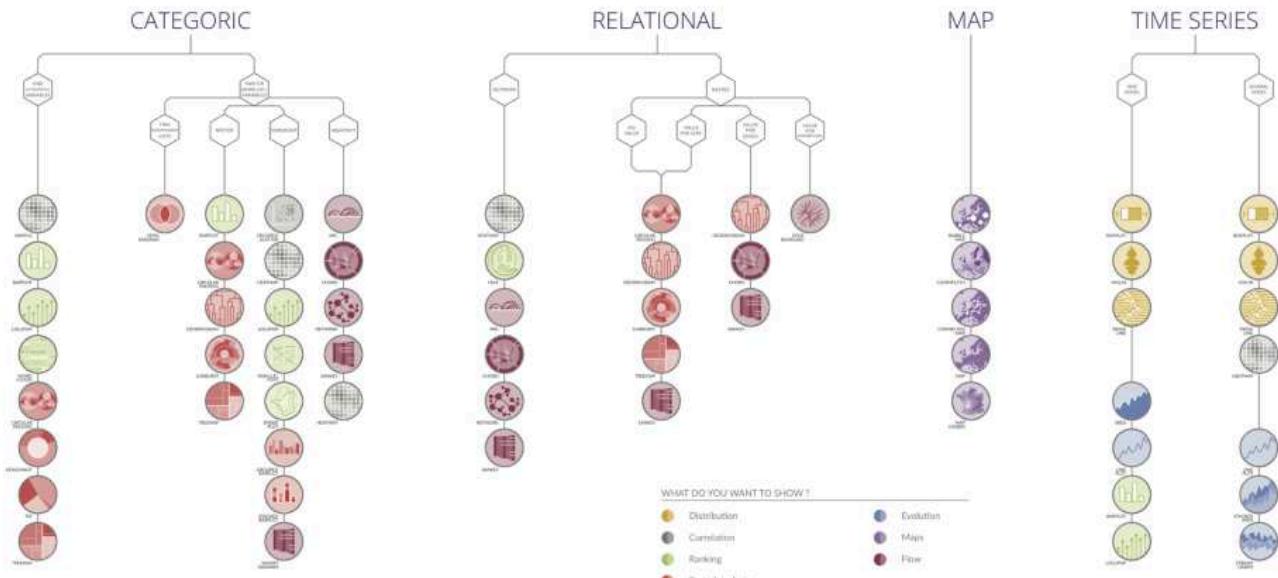
from Data to Viz

'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.

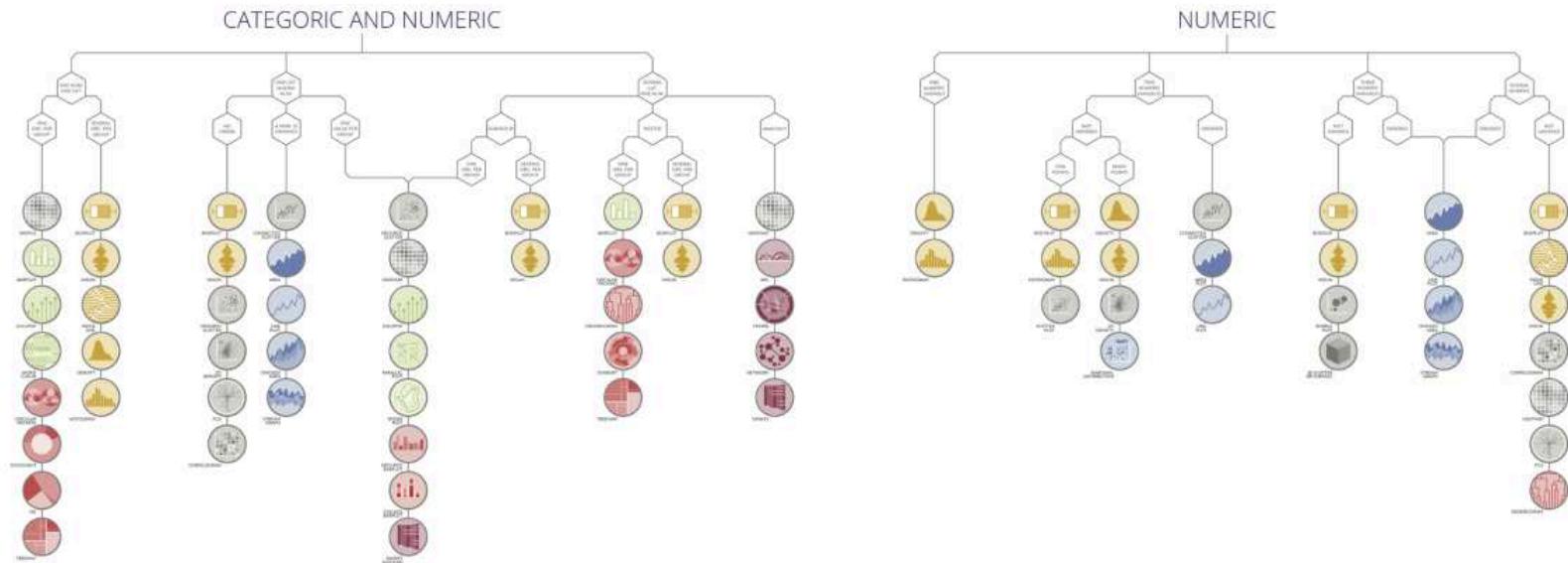
DataViz is a work in progress and this project does not claim to be exhaustive. However it should provide you with a good starting point. For an interactive version and much more, visit:

data-to-viz.com



WHAT DO YOU WANT TO SHOW ?

- | | |
|---|--|
| <ul style="list-style-type: none"> ● Distribution ● Correlation ● Ranking ● Part of a whole | <ul style="list-style-type: none"> ● Evolution ● Map ● Flow |
|---|--|



© 2014 by Ivan Kavcic & Cedric Scherer for www.data-to-viz.com

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

Cédric Scherer Data Visualization & Information Design



The screenshot shows a modal window on the Data to Viz website. At the top left is a yellow circular icon containing a boxplot. Below it is the word "BOXPLOT" in green. A gray "X" button is in the top right corner. To the right of the modal, the main page header "VISUALIZATIONS" is visible, along with a sub-header "presented in this website." and four categories: "Part of a whole", "Evolution", "Map", and "Flow".

BOXPLOT

Summarize the distribution of numeric variables

About

A boxplot gives a nice summary of one or several numeric variables. The line that divides the box into 2 parts represents the median of the data. The end of the box shows the upper and lower quartiles. The extreme lines show the highest and lowest value excluding outliers.

Common Mistakes

- 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.
- [Order your boxplot by median](#) can make it more insightful.

Code

[R graph gallery](#) [Python gallery](#) [D3.js gallery](#) [Flourish](#)

Read More

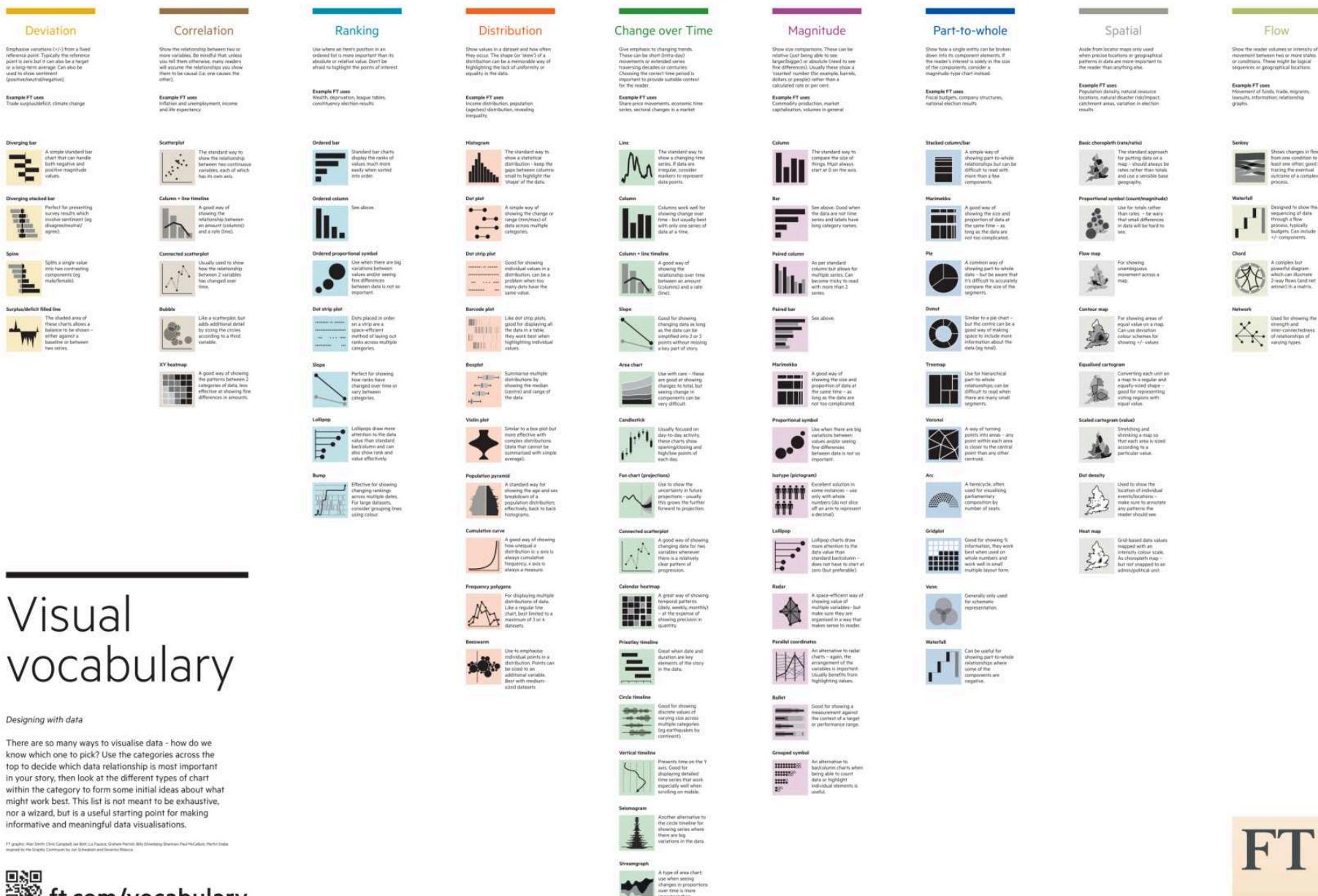
See the [dedicated page](#).

Below the modal, there is a grid of 12 circular icons representing different data visualization types:

- Boxplot
- Ridgeline
- Scatter
- Connected scatter
- Density 2d
- Barplot
- Lollipop
- Circular Barplot
- Treemap
- Dendrogram
- Circular packing
- Sunburst

Source: [Data to Viz](#)





Visual vocabulary

Designing with data

There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

FT graphic: Alan Smith; Chris Campbell; Ian Boni; Liz Fawcett; Graham French; Billy Giovannini; Sharmeen; Paul McCullagh; The Art of Data
Inspired by the Graphics Committee for Jon Schwabish and Jennifer Moshita

ft.com/vocabulary

"FT Visual Vocabulary" by Alan Smith et al. ([Posters](#))

Cédric Scherer Data Visualization & Information Design



© Financial Times 2010-2015
This work is licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License



THE GRAPHIC CONTINUUM

The Graphic Continuum shows several ways that data can be illustrated individually or combined to show relationships. Use of various shapes, chart types, and colors can help identify patterns, tell stories, and reveal relationships between different sets and types of data. Bar charts, or histograms, for example, can illustrate a distribution of data over time, but they also can show categorical or geographic differences. Scatterplots can illustrate data from a single instance or for a period, but they also can be used to identify a distribution around a mean.

This set of charts does not constitute an exhaustive list, nor do the connections represent every possible pathway for linking data and ideas; instead, the Graphic Continuum identifies some presentation methods, and it illustrates some of the connections that can bind different representations together. The six groups do not define all possibilities. Many other useful, overlapping data types and visualization techniques are possible.

This chart can guide graphic choices, but your imagination can lead the way to other effective ways to present data.

COMPARING CATEGORIES

Compare values across categories



DISTRIBUTION

Graphical representations of the distribution of data



A Histogram presents relative frequency of data points with respect to their value.

A Box & Whisker illustrates a box containing the median and quartiles, with whiskers extending to minimum and maximum values.

TIME

Track changes over time



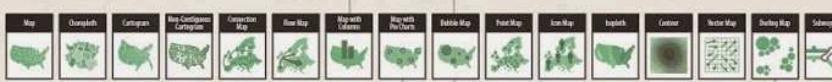
A Line Chart illustrates the relationship between two variables over time.

A Sparkline chart compresses multiple data series into a single line of data points.

A Connected Scatterplot illustrates the relationship between variables over time.

GEOSPATIAL

Relate data to its geography



PART-TO-WHOLE

Visualizations that relate the part of a variable to its total



RELATIONSHIP

Illustrates correlations or relationships between variables



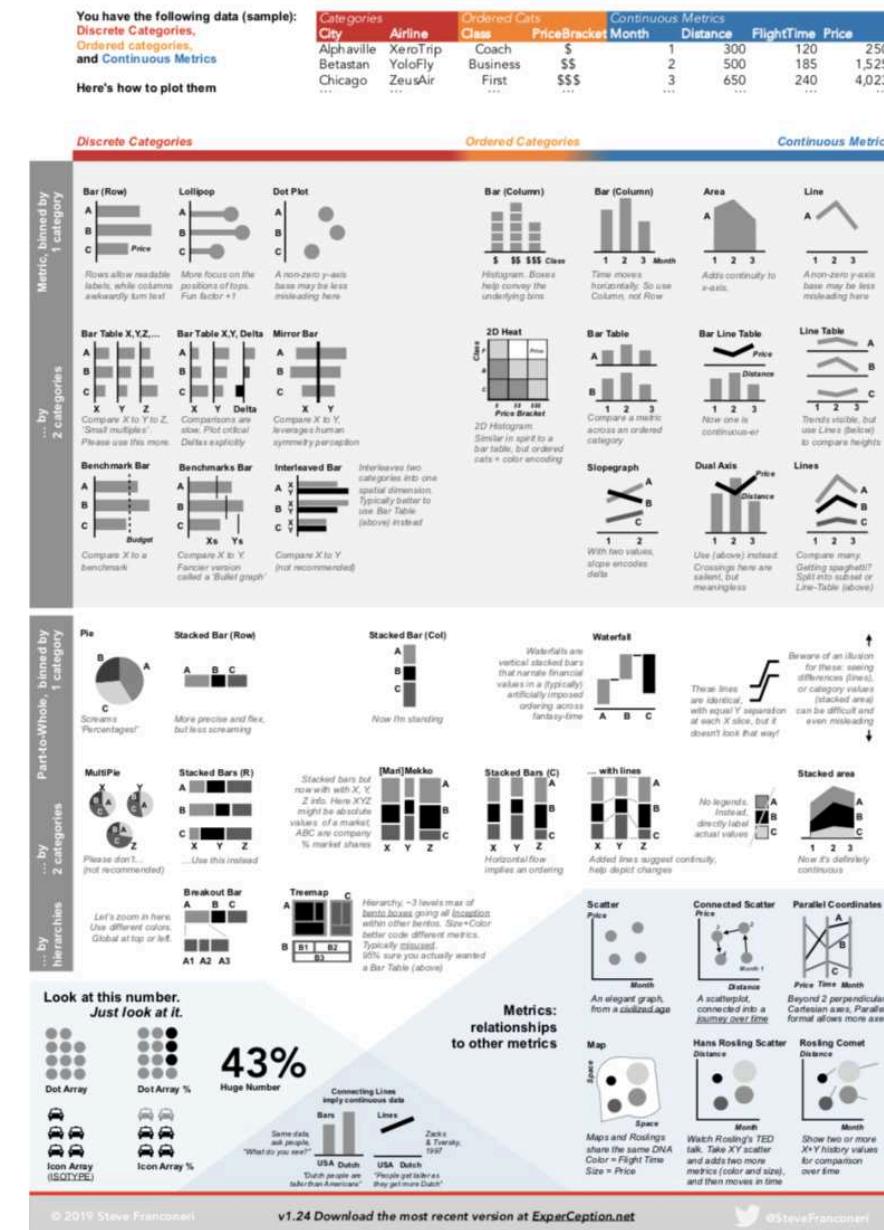
© Jonathan Schwabish & Sverino Ribecca
@jonschwabish QSR_Vizual_Inf

"The Graphic Continuum" by Jon Schwabish & Sverino Ribecca

Cédric Scherer Data Visualization & Information Design



Which Visualization? A Quick Reference



"Which Visualization? A Quick Reference" by Steven Franconeri

Cédric Scherer Data Visualization & Information Design



Choosing a chart type to show the breakdown of a total

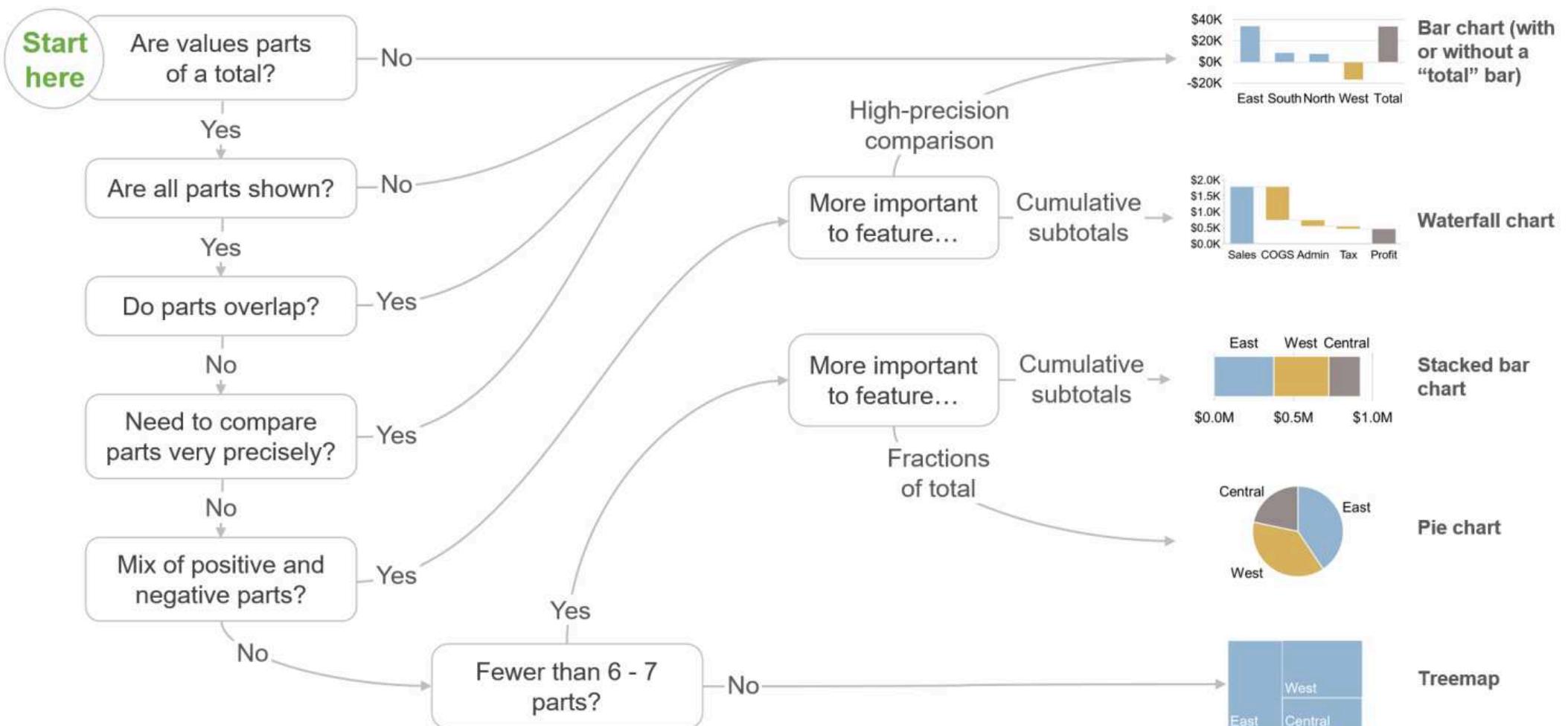


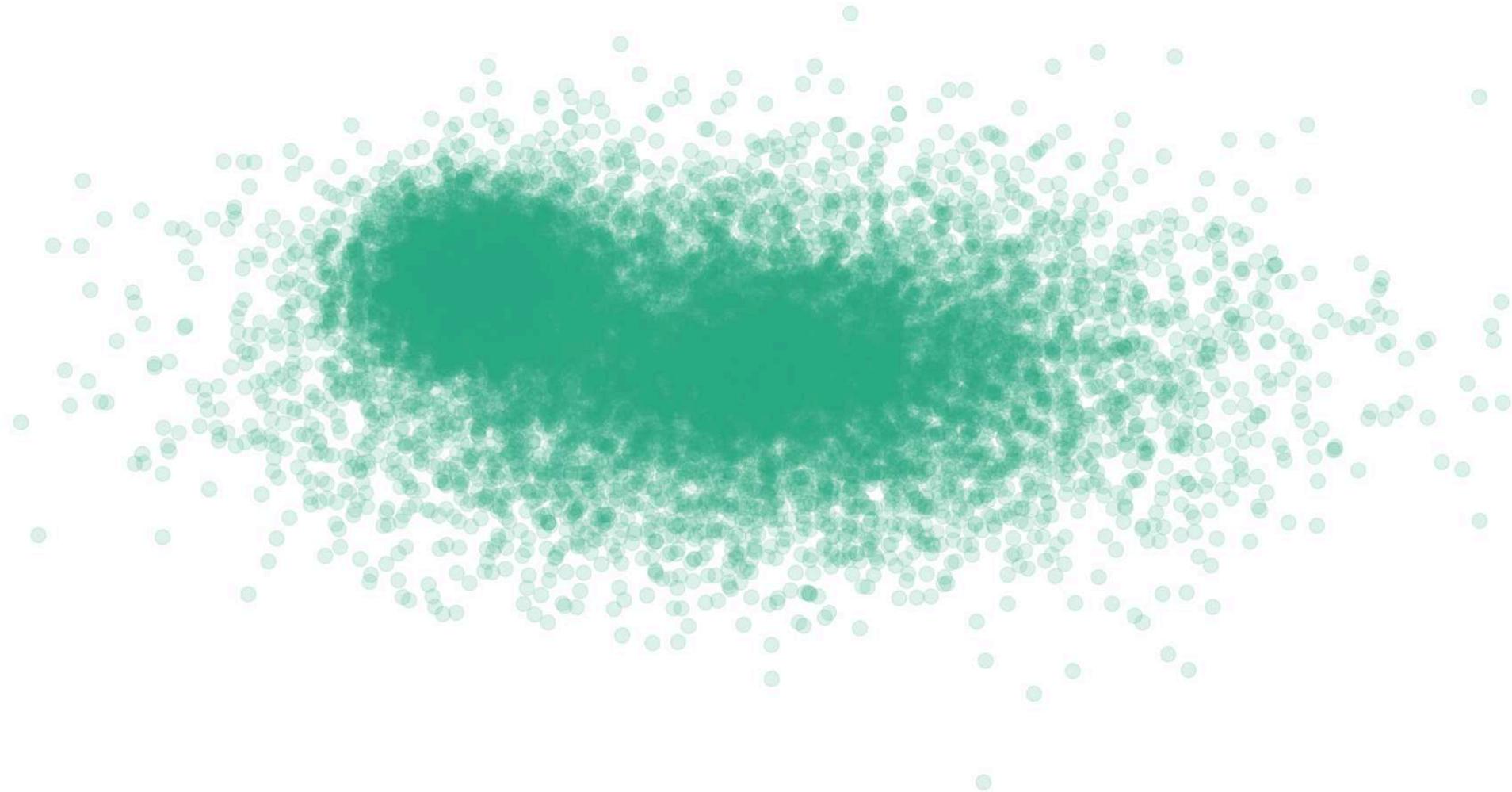
Chart Choice Decision Trees by Nick Desbarats



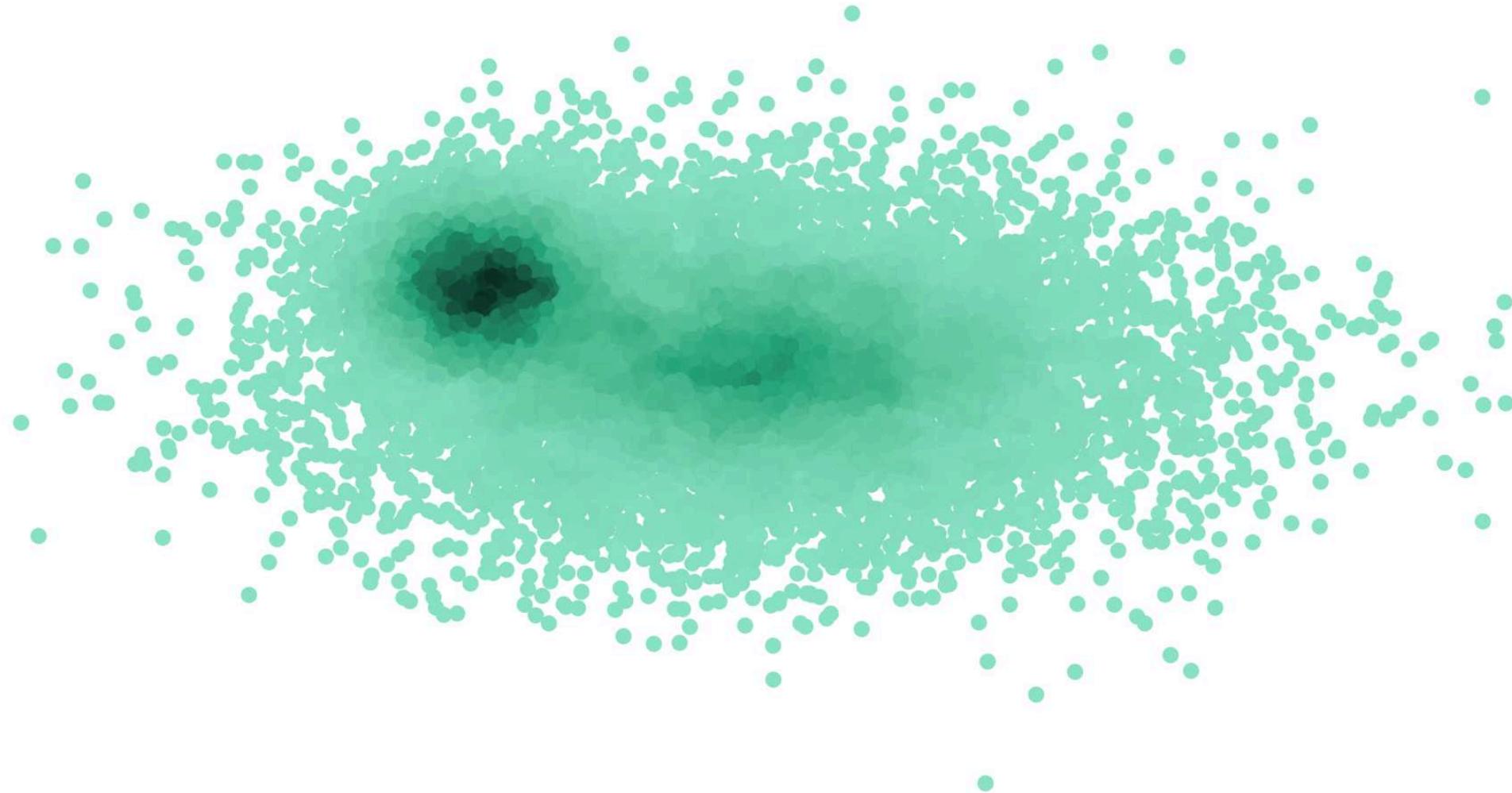
Solving Common Chart Issues



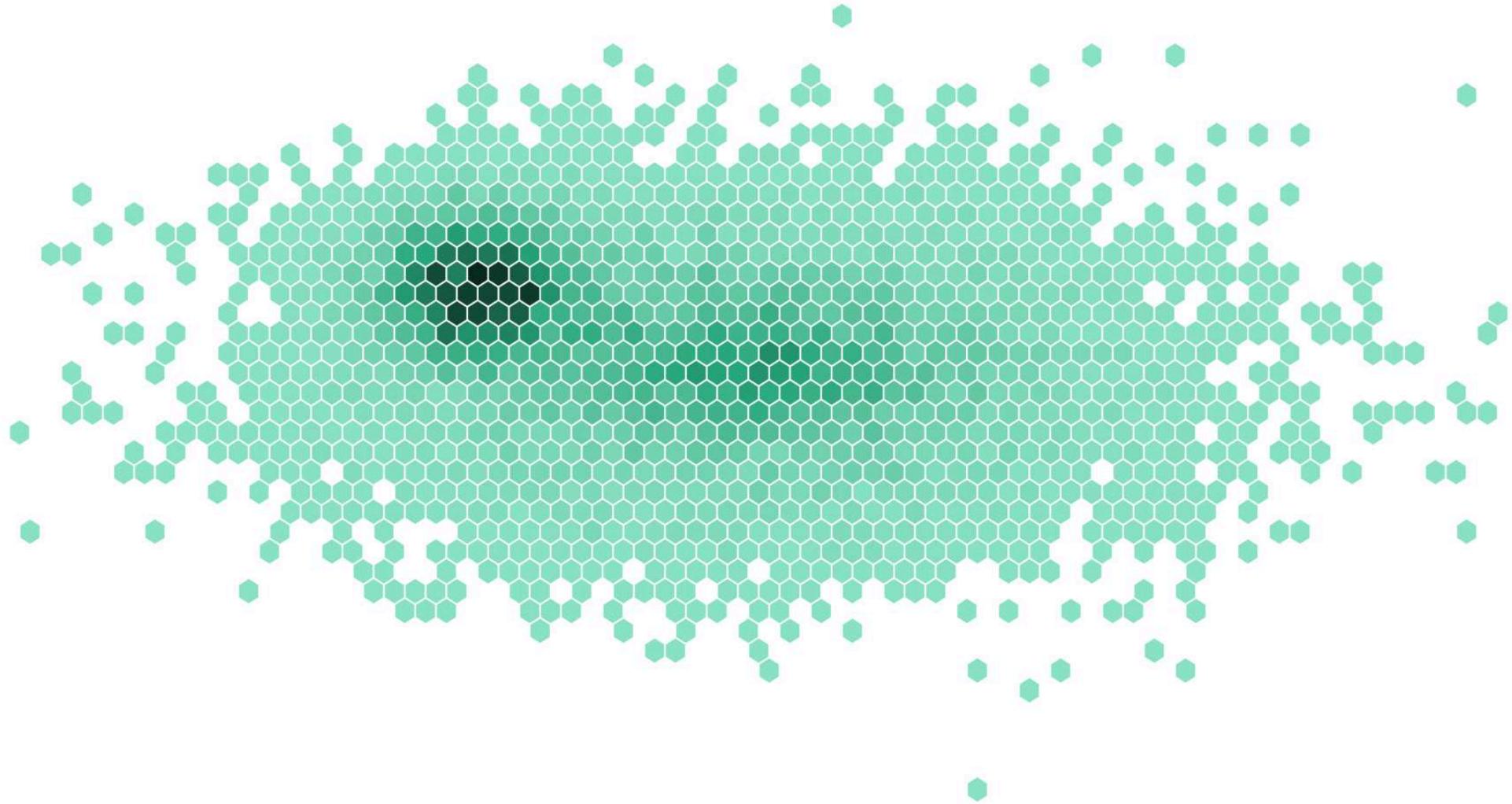
Overplotting Issue



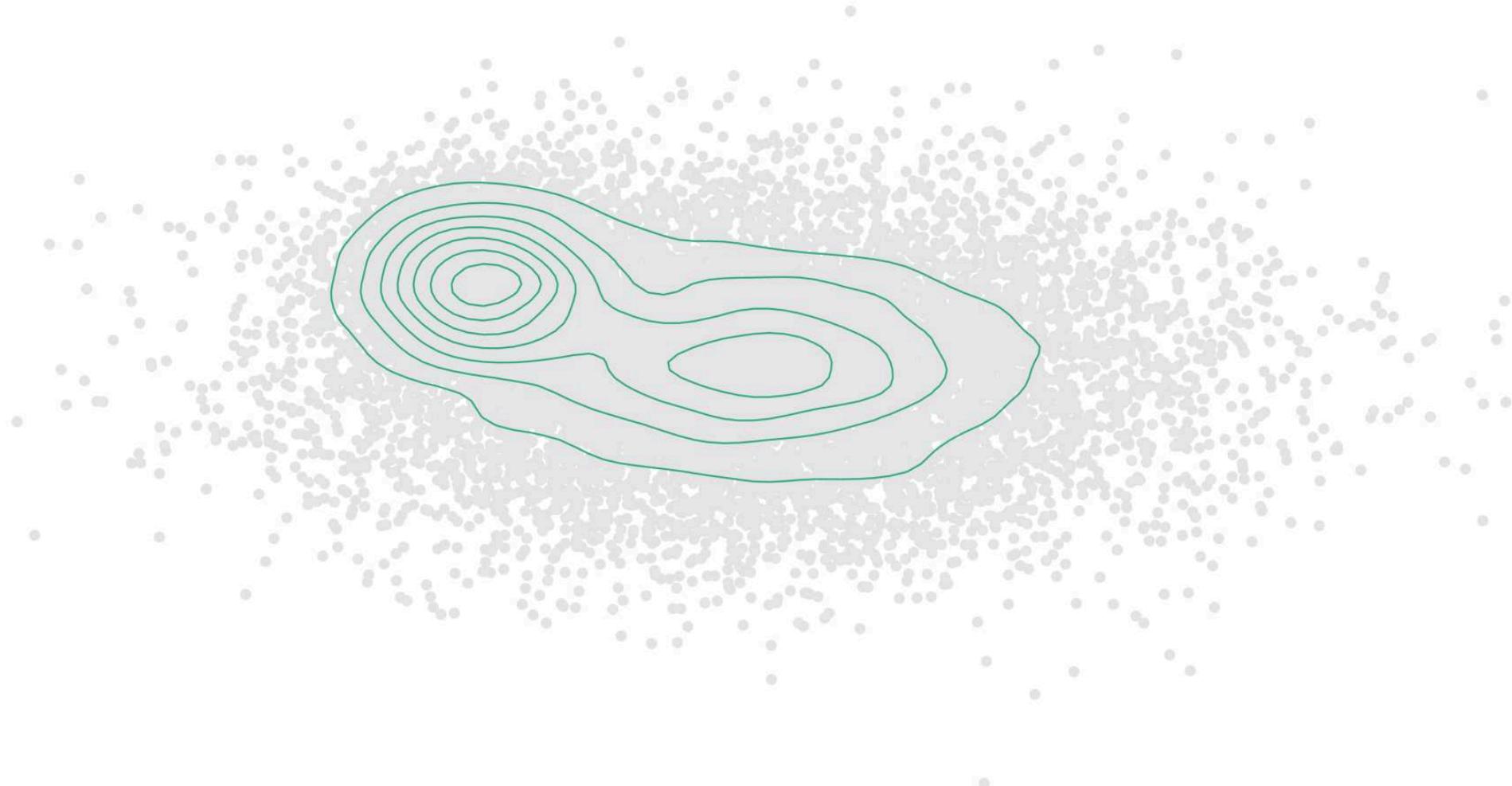
Overplotting Issue



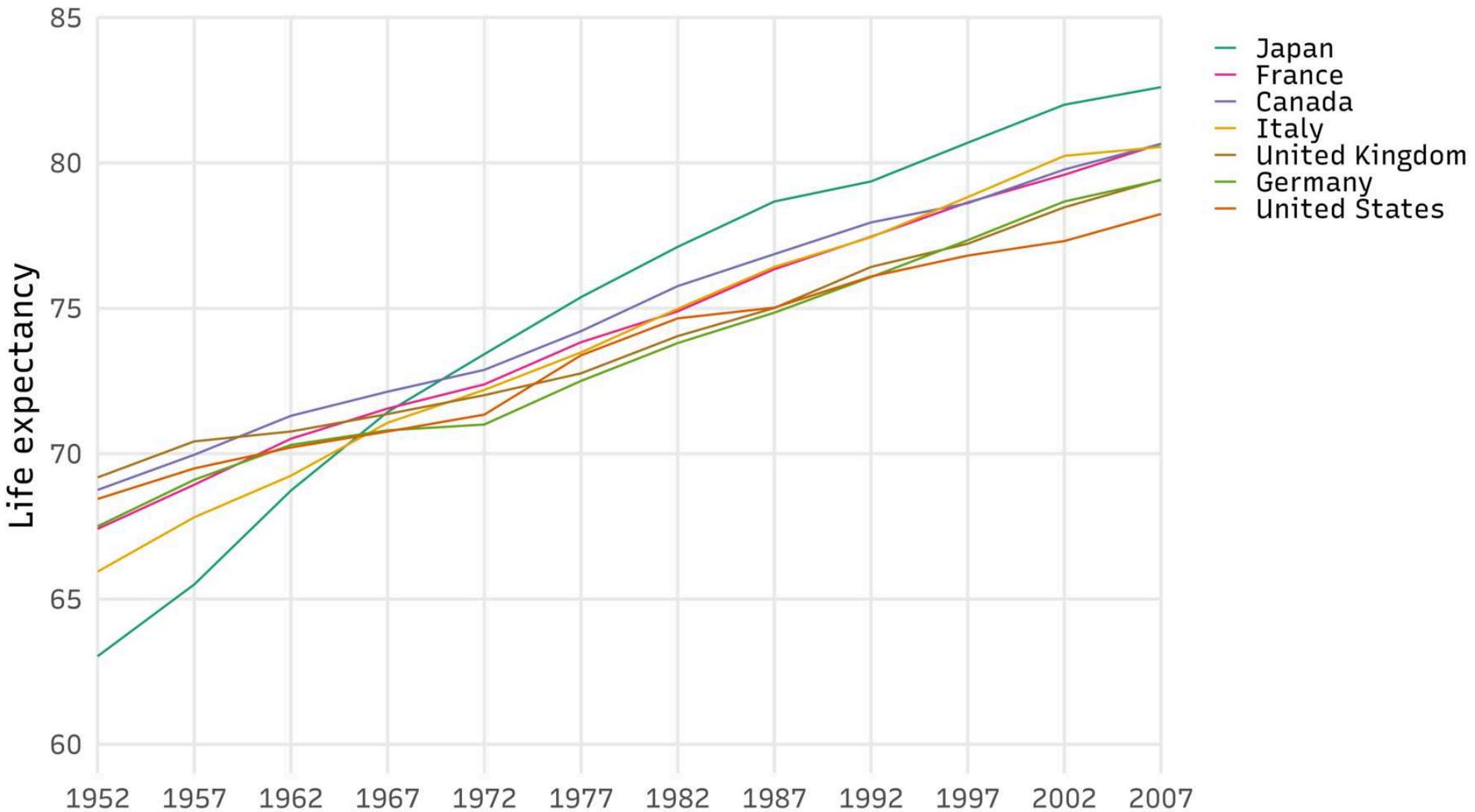
Overplotting Issue



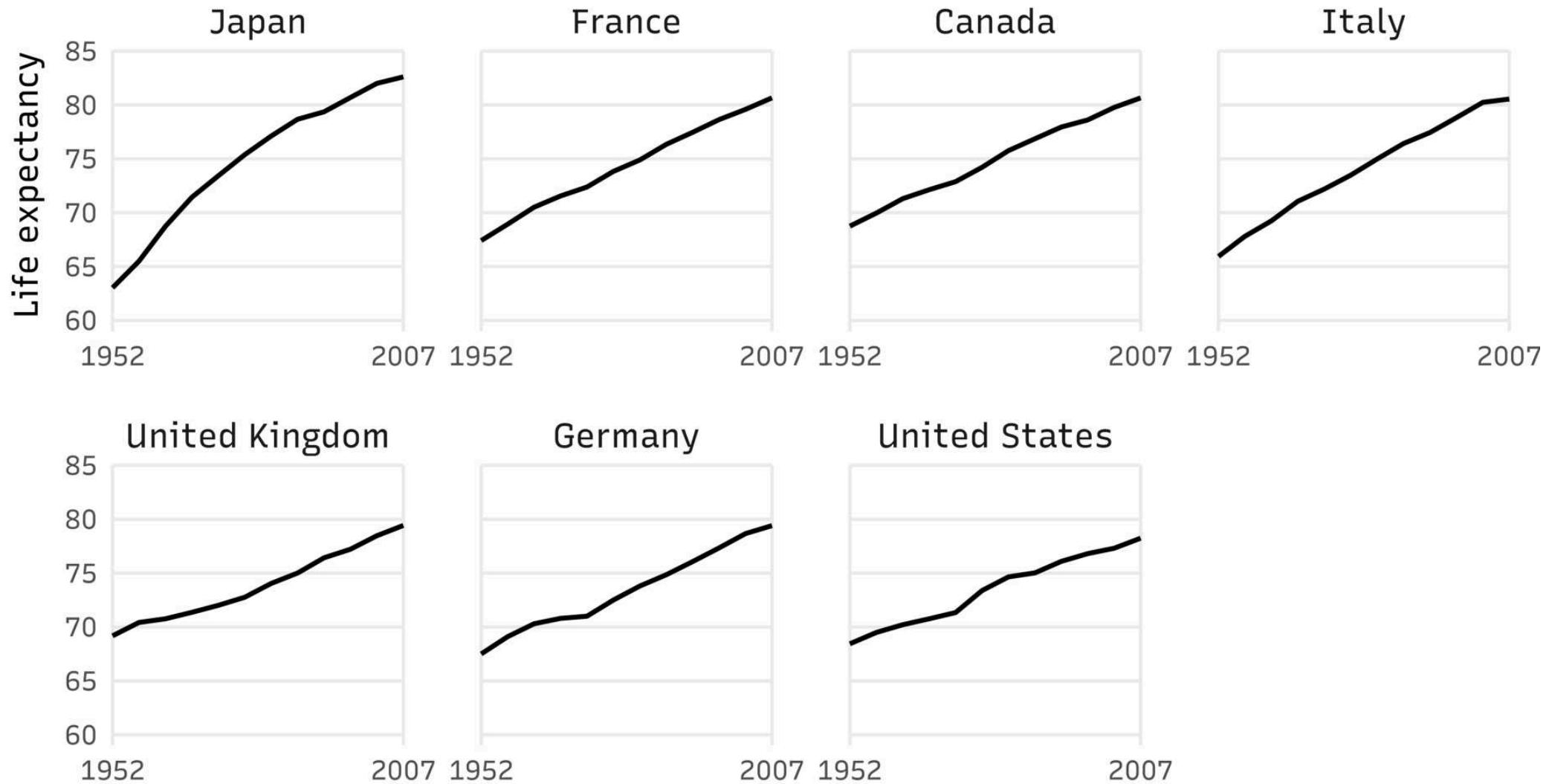
Overplotting Issue



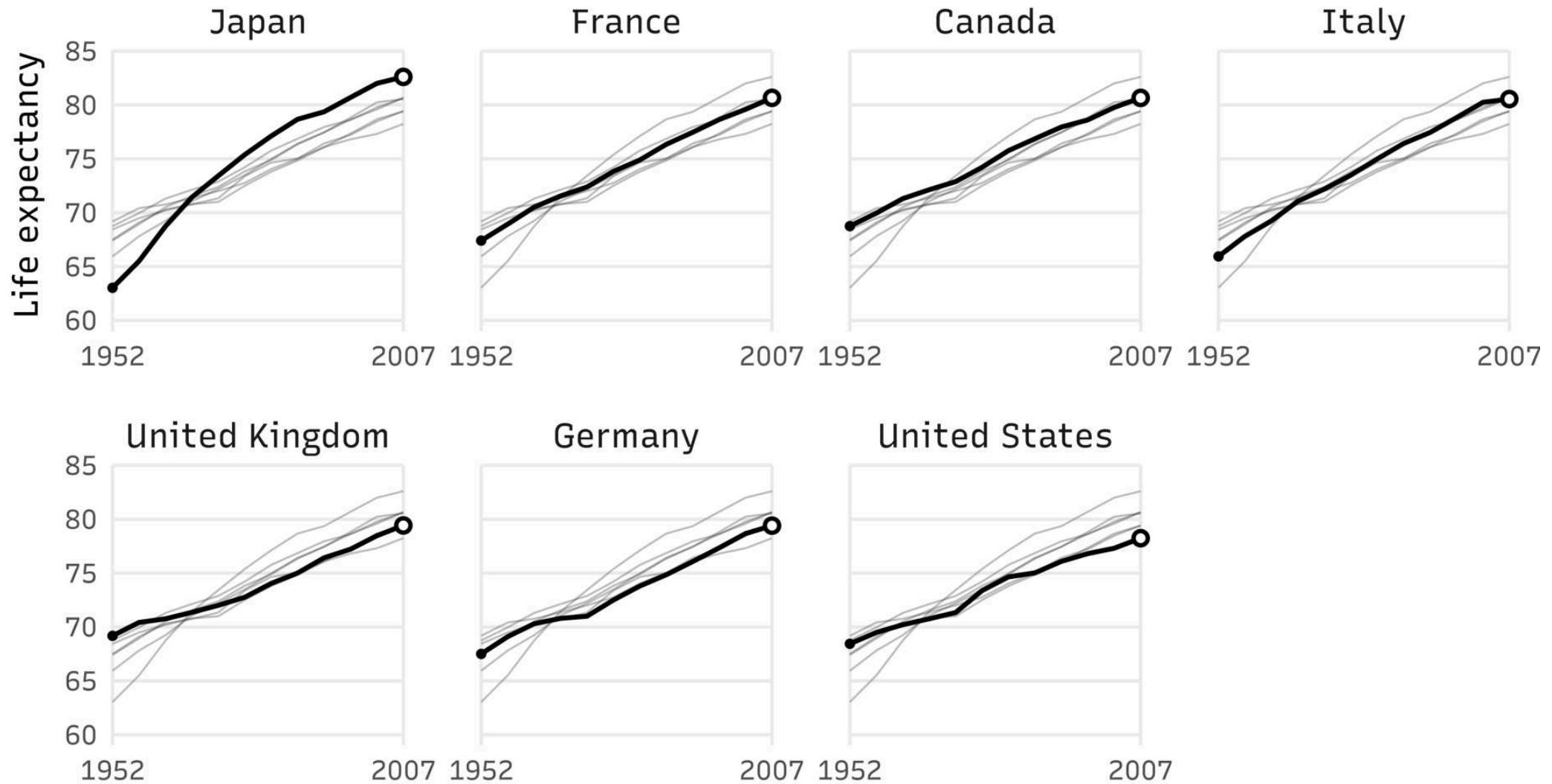
Spaghetti Plots



Spaghetti Plots



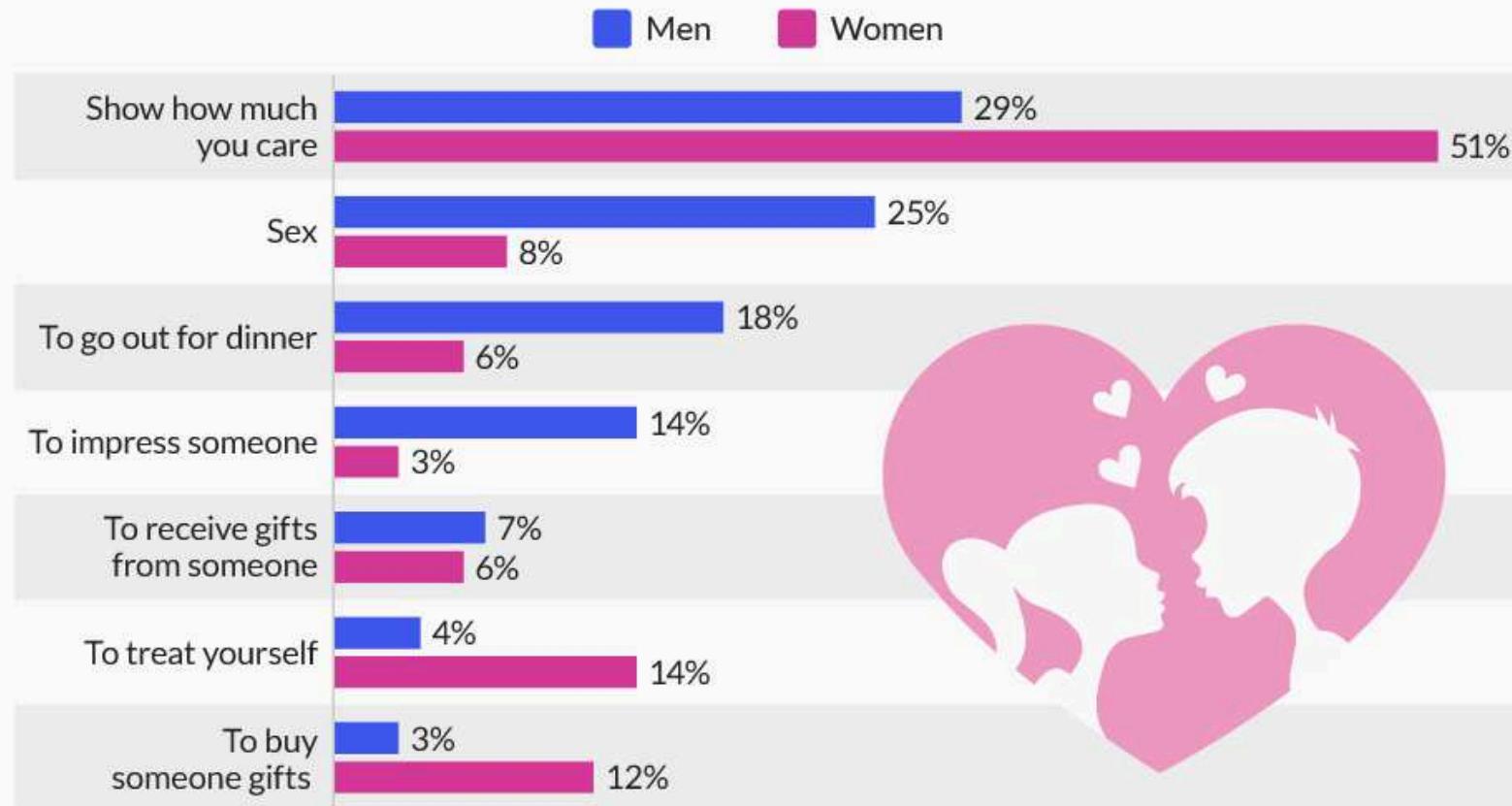
Spaghetti Plots



Grouped Bars Issue

What's the point of Valentine's Day?

% of men and women in the UK saying the following are the point of Valentine's Day



@StatistaCharts

Source: Chillisauce.co.uk

i100

from
The INDEPENDENT

statista

Source: Statista

Cédric Scherer Data Visualization & Information Design



Grouped Bars Issue

For 25% of surveyed men, the point of Valentine's Day is to “*get laid*” —more than the half of women responded “*show how much you care*”



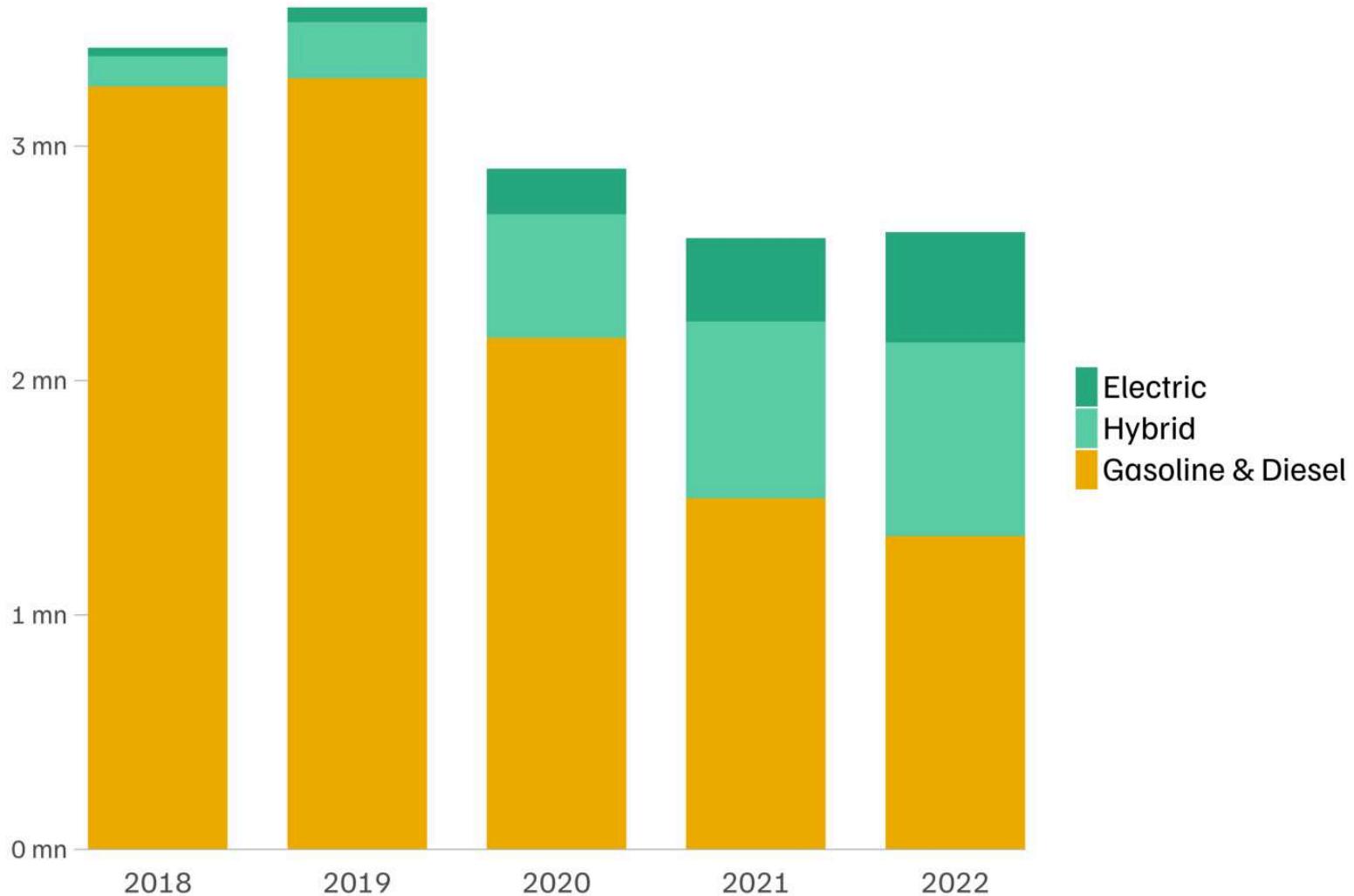
Visualization: Cédric Scherer • Data: UK survey in 2015 by chilisauce.co.uk

Cédric Scherer Data Visualization & Information Design



Stacked Bars Issue

New vehicle registrations in Germany



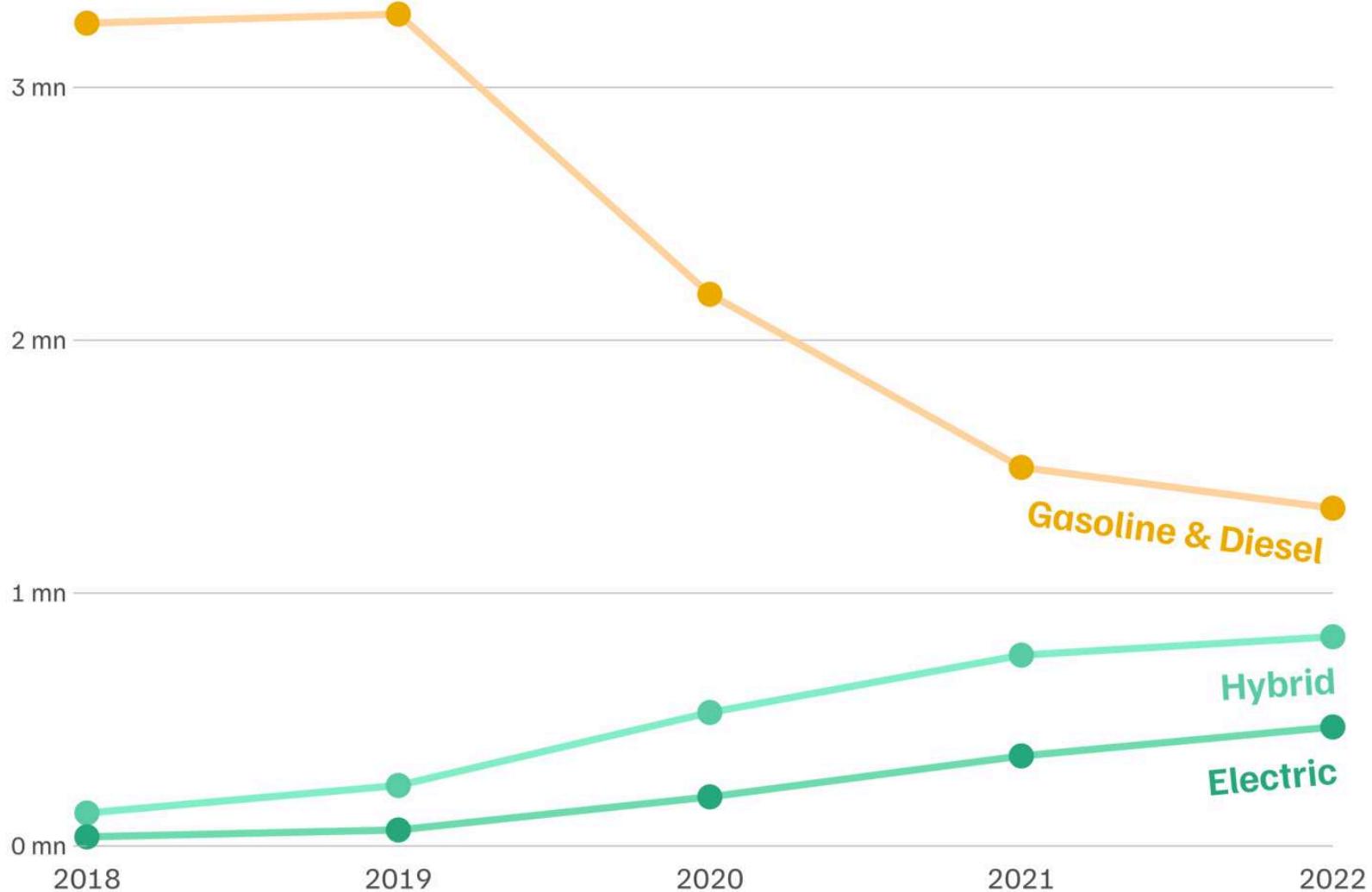
Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



Stacked Bars Issue

New vehicle registrations in Germany



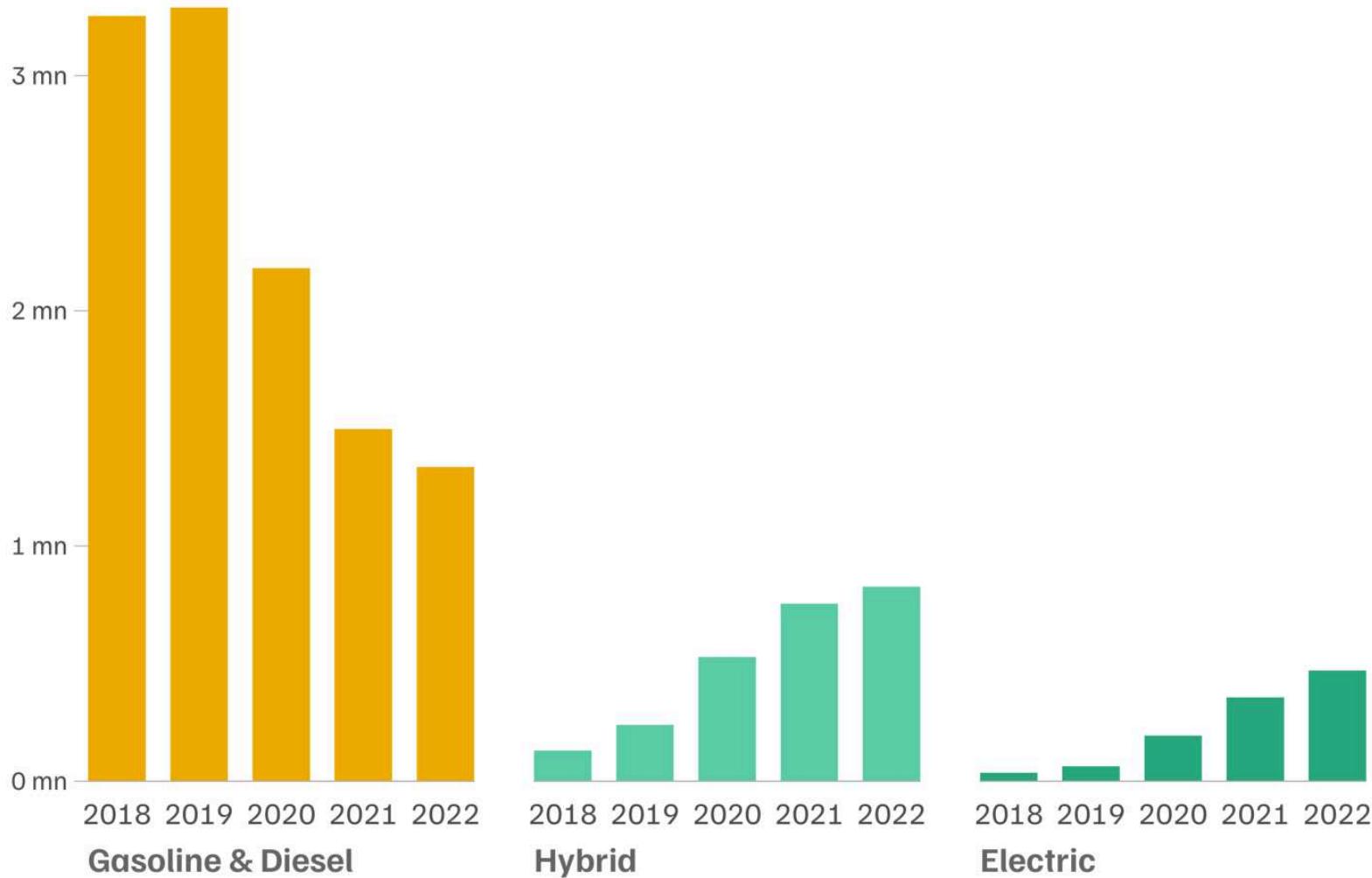
Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



Stacked Bars Issue

New vehicle registrations in Germany



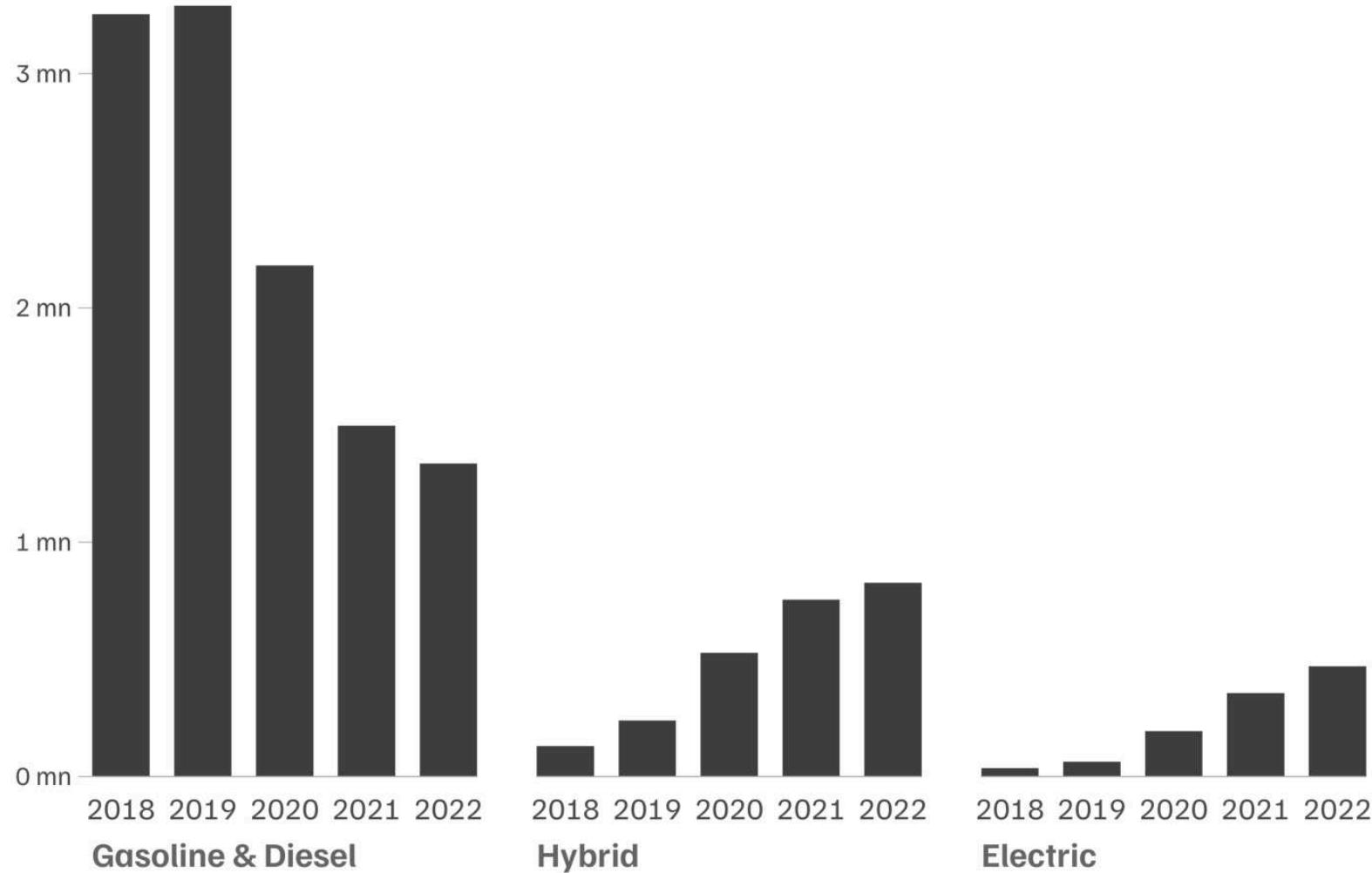
Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



Stacked Bars Issue

New vehicle registrations in Germany



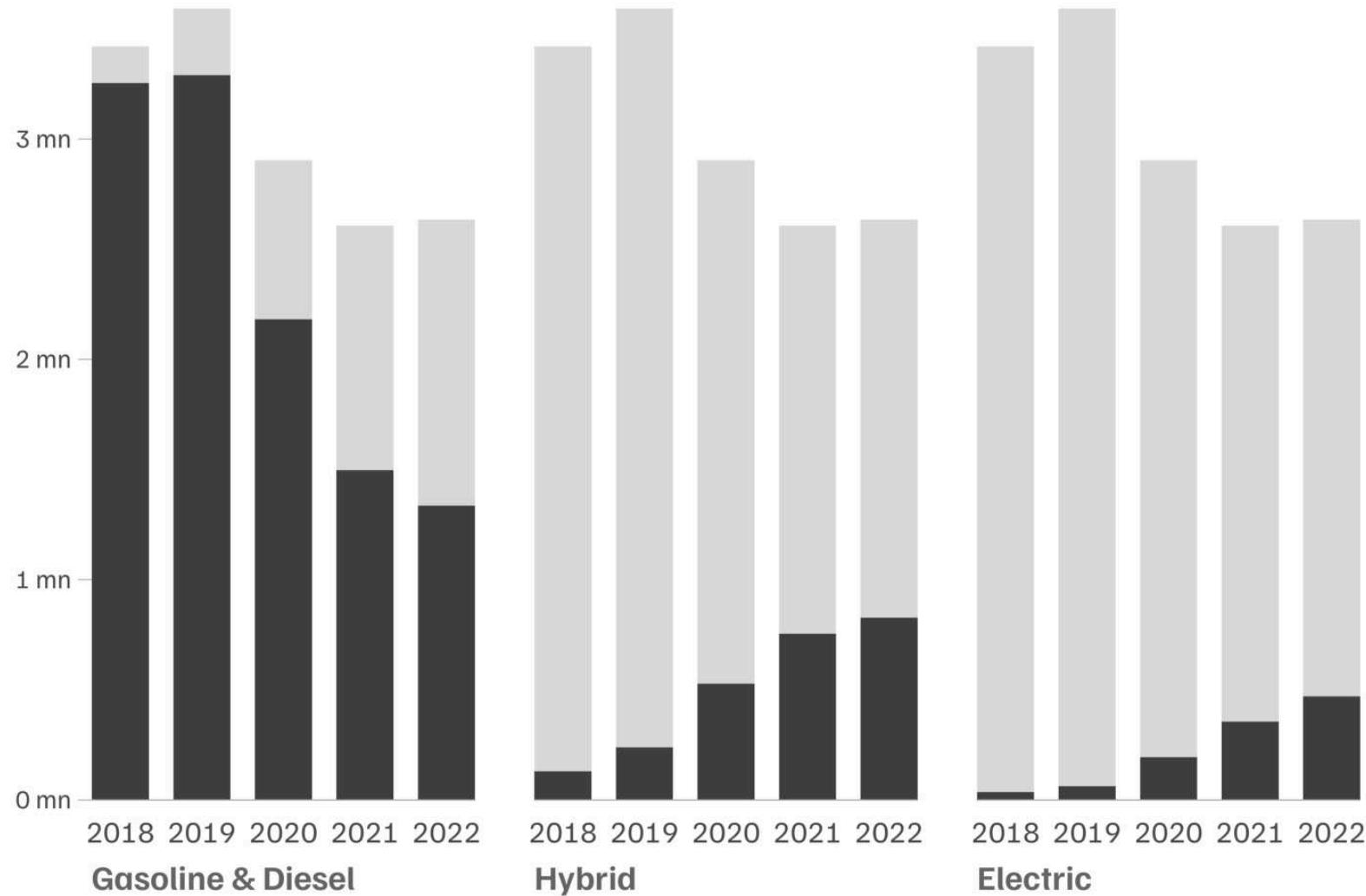
Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



Stacked Bars Issue

New vehicle registrations in Germany



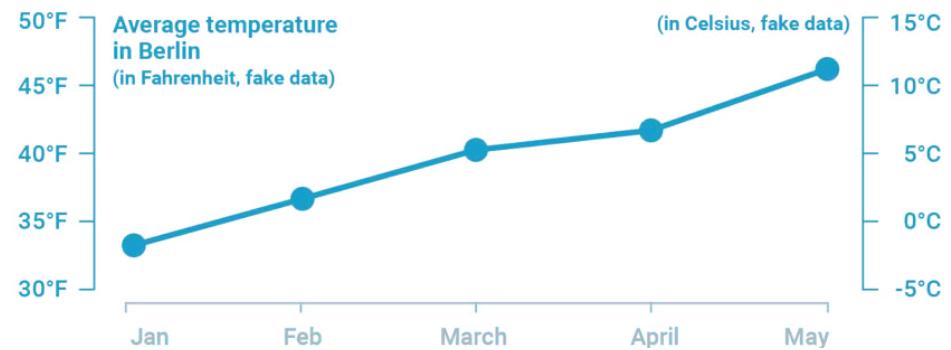
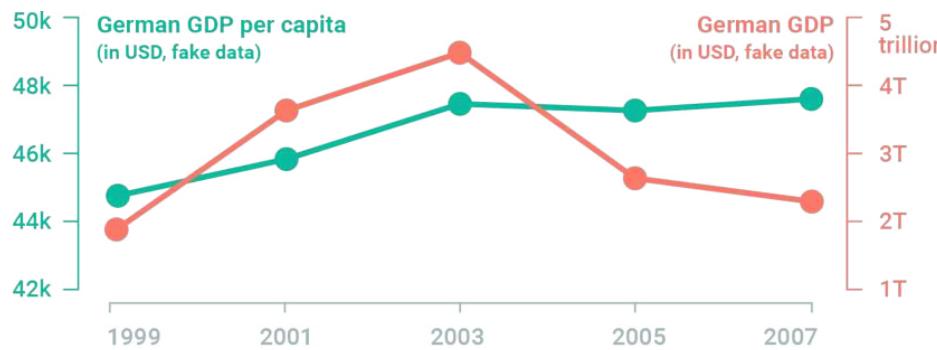
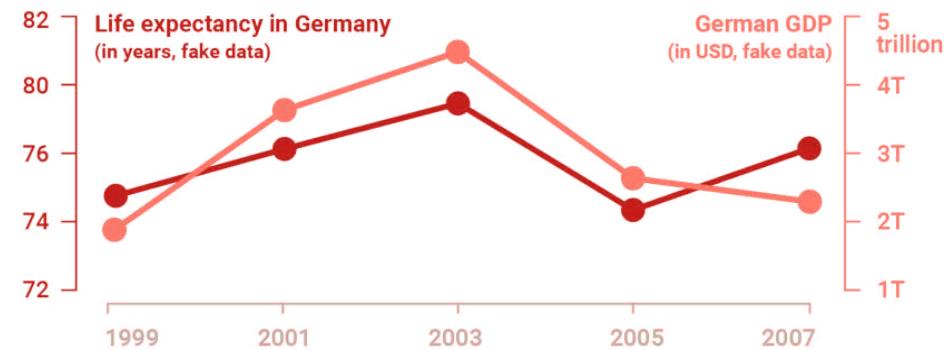
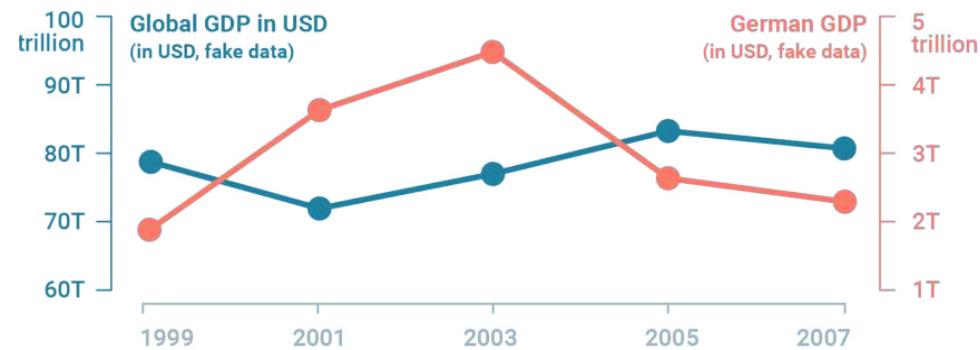
Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



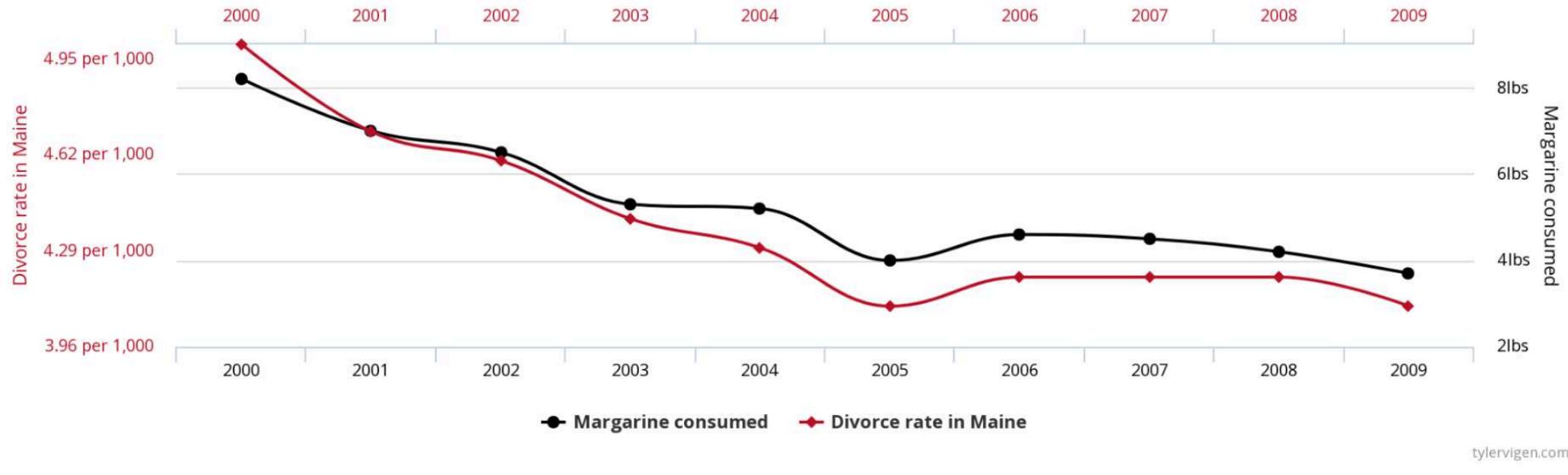
Dual Axis Issue

Examples by Lisa Charlotte Muth, [Datawrapper Blog](#)



Dual Axis Issue

Divorce rate in Maine
correlates with
Per capita consumption of margarine



tylervigen.com

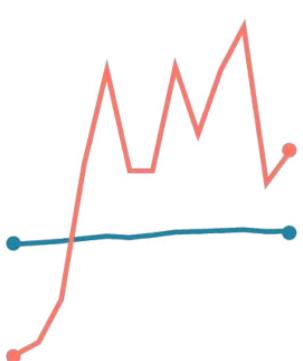
Source: [Spurious Correlations](#)



Dual Axis Issue



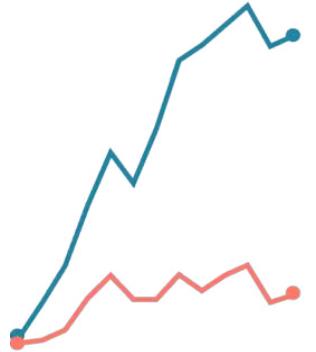
Orange steady,
Blue massively increasing.



Blue steady,
Orange increasing.



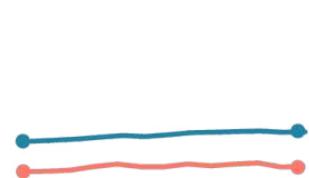
Both started at the same
level, but Orange increased
far more than Blue.



Both started at the same
level, but Blue increased far
more than Orange.



Both started with the
same increase, then Blue
raced to the top.

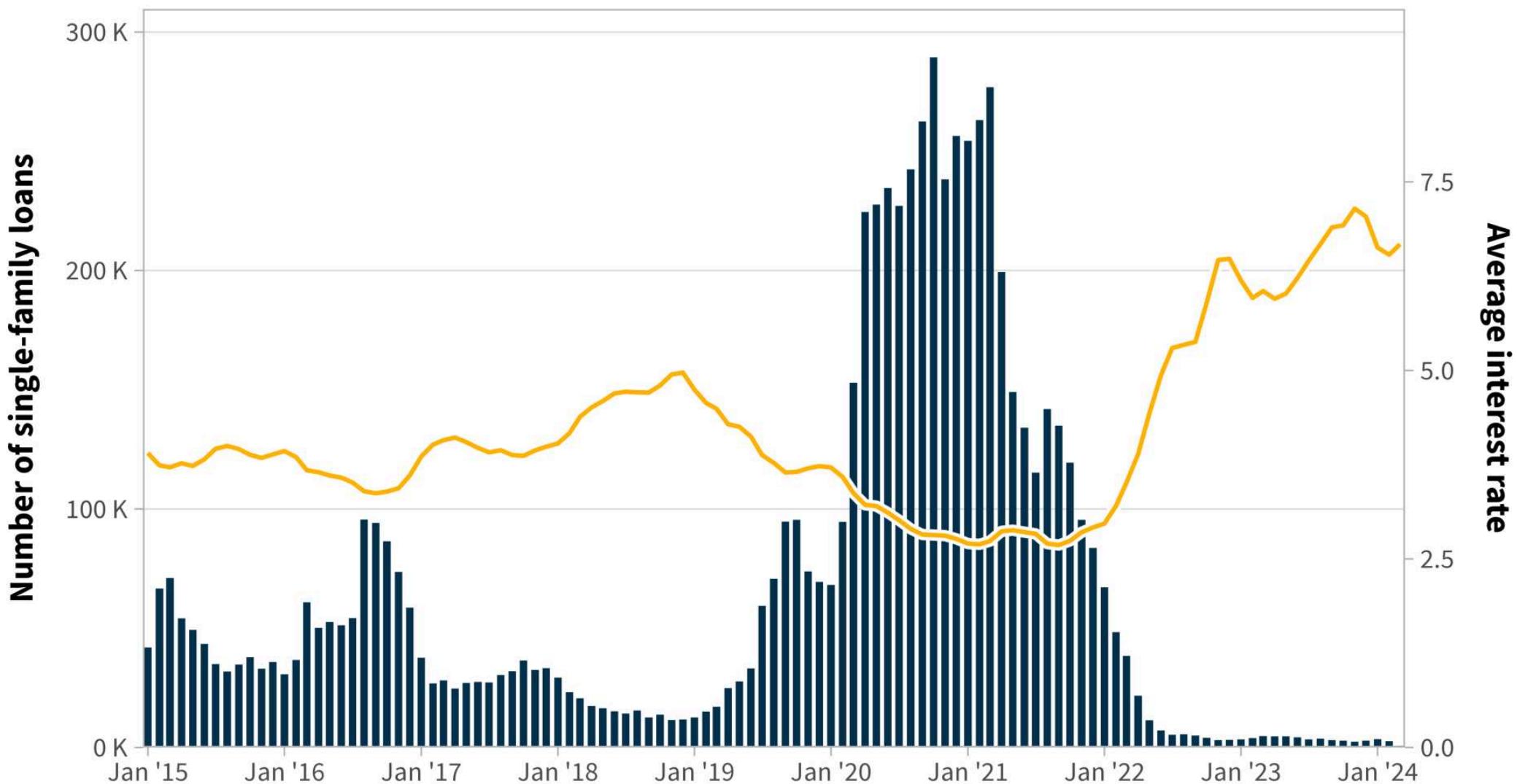


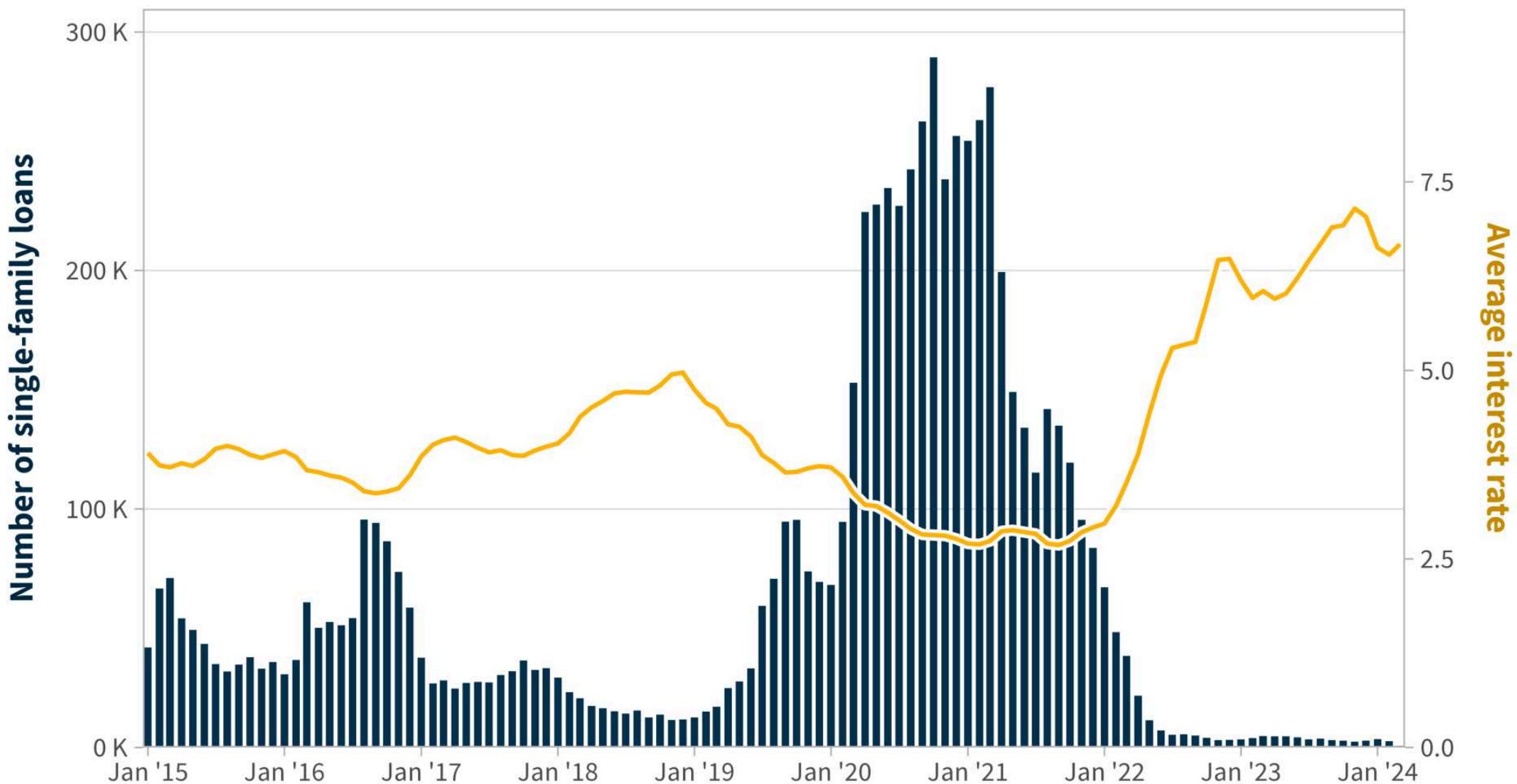
Both steady.

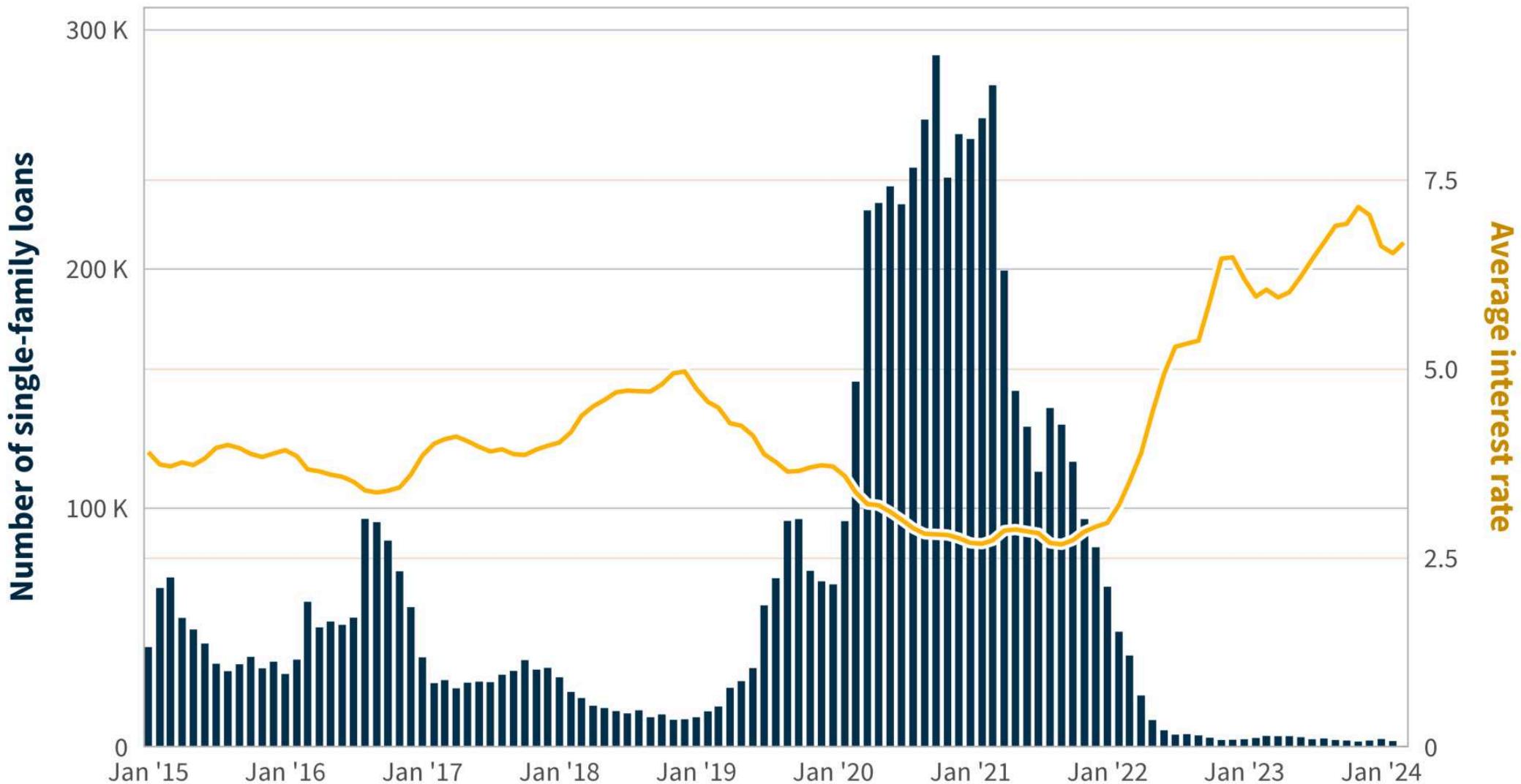
Source: Lisa Charlotte Muth, [Datawrapper Blog](#)

Cédric Scherer Data Visualization & Information Design

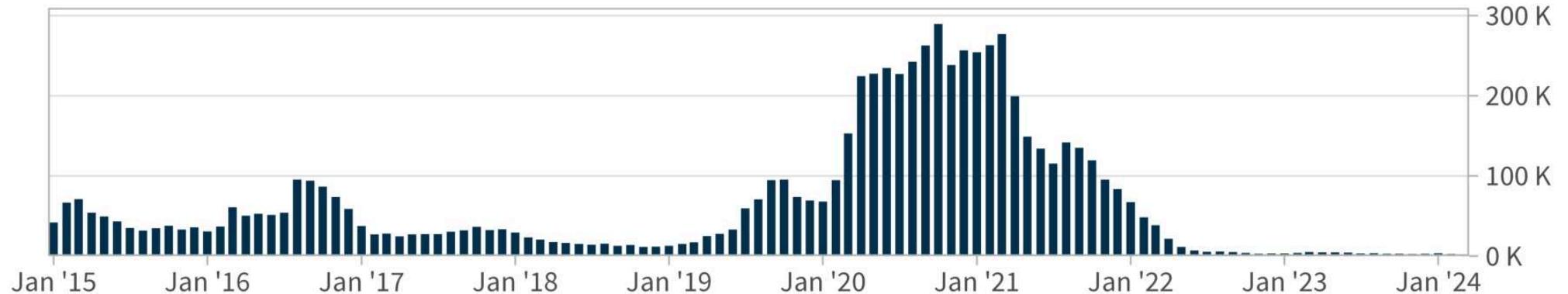




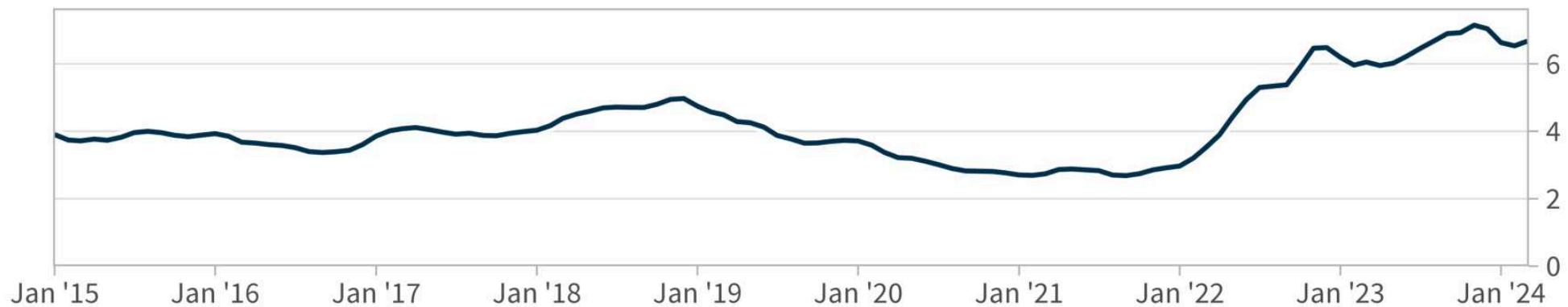




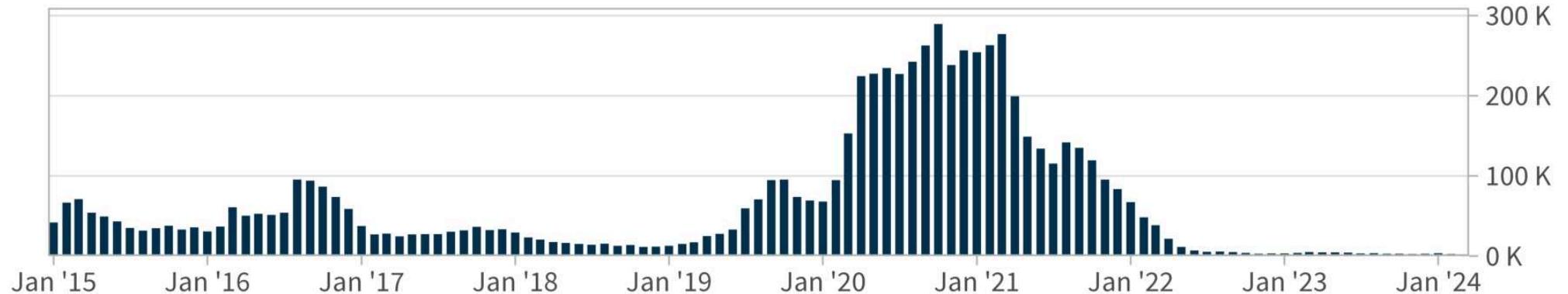
Number of single-family loans



Average interest rate

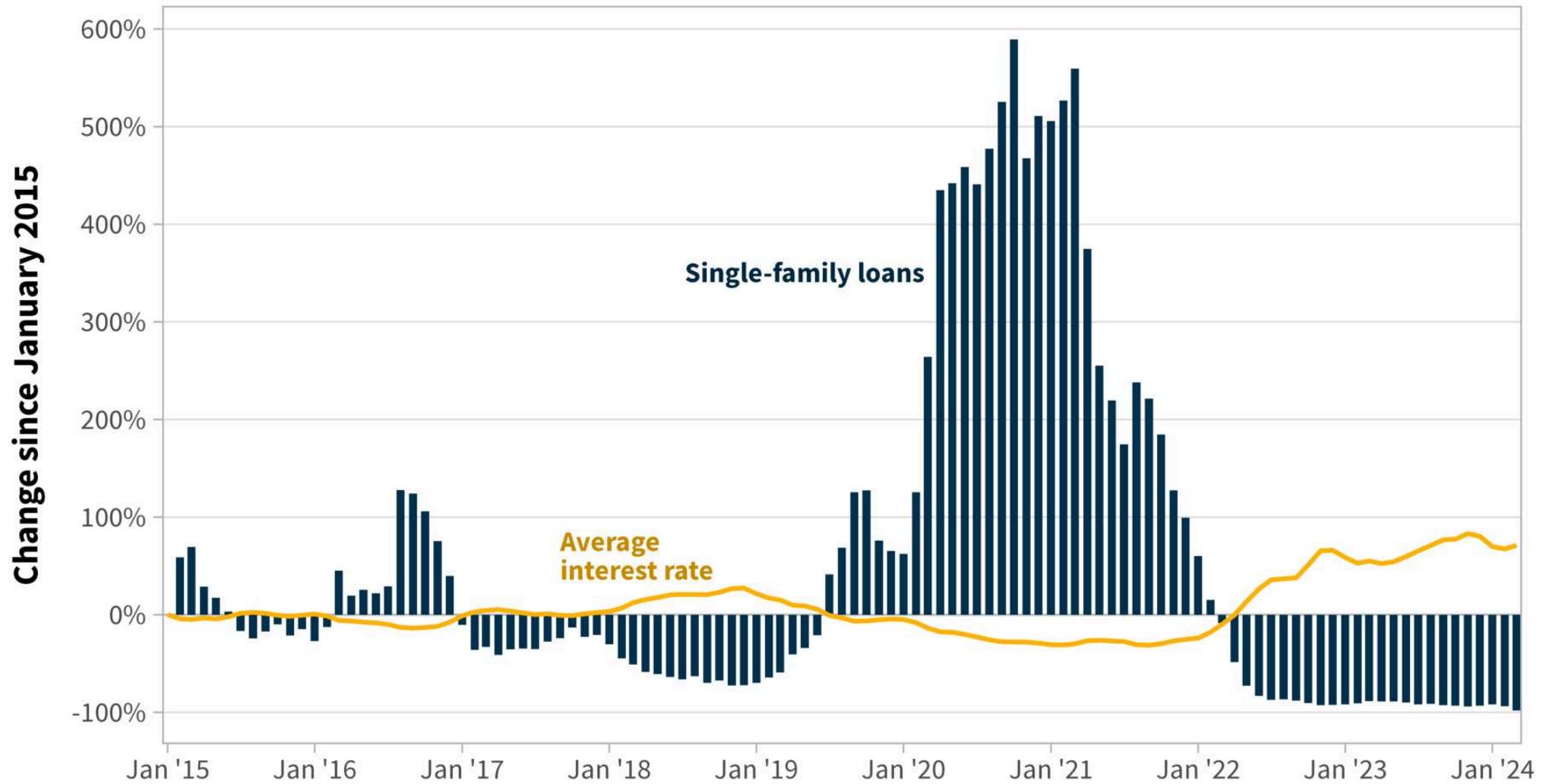


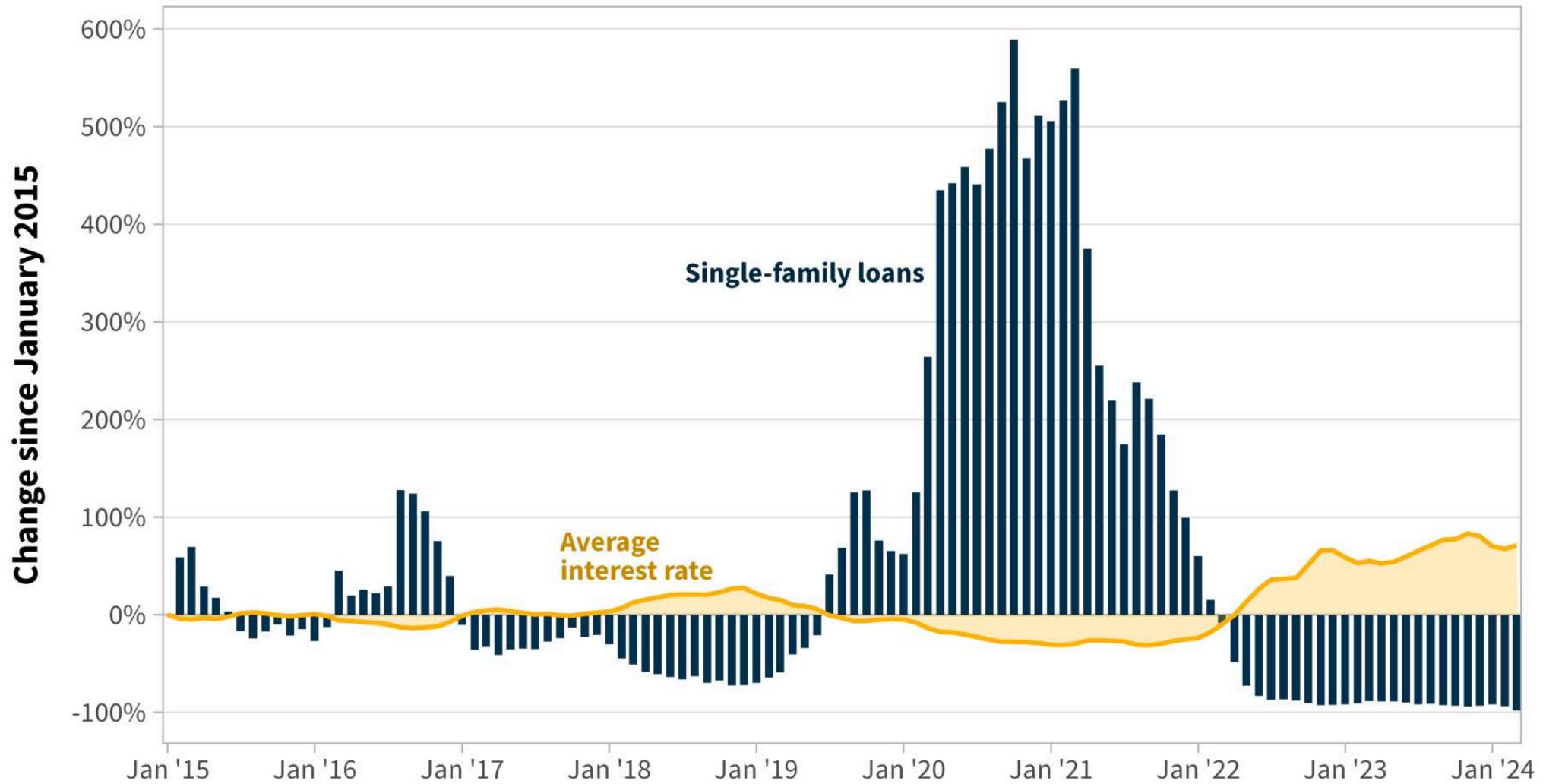
Number of single-family loans

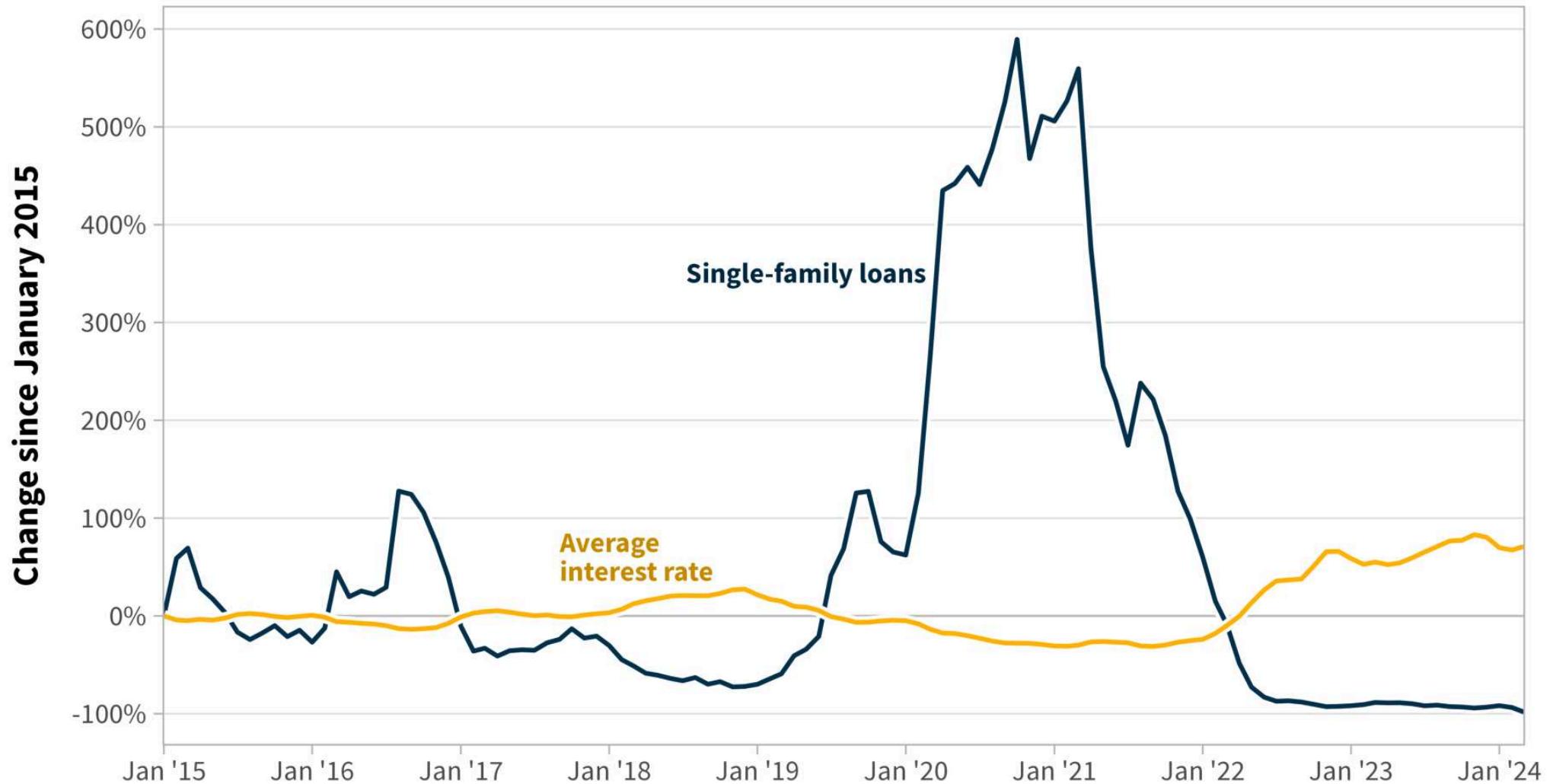


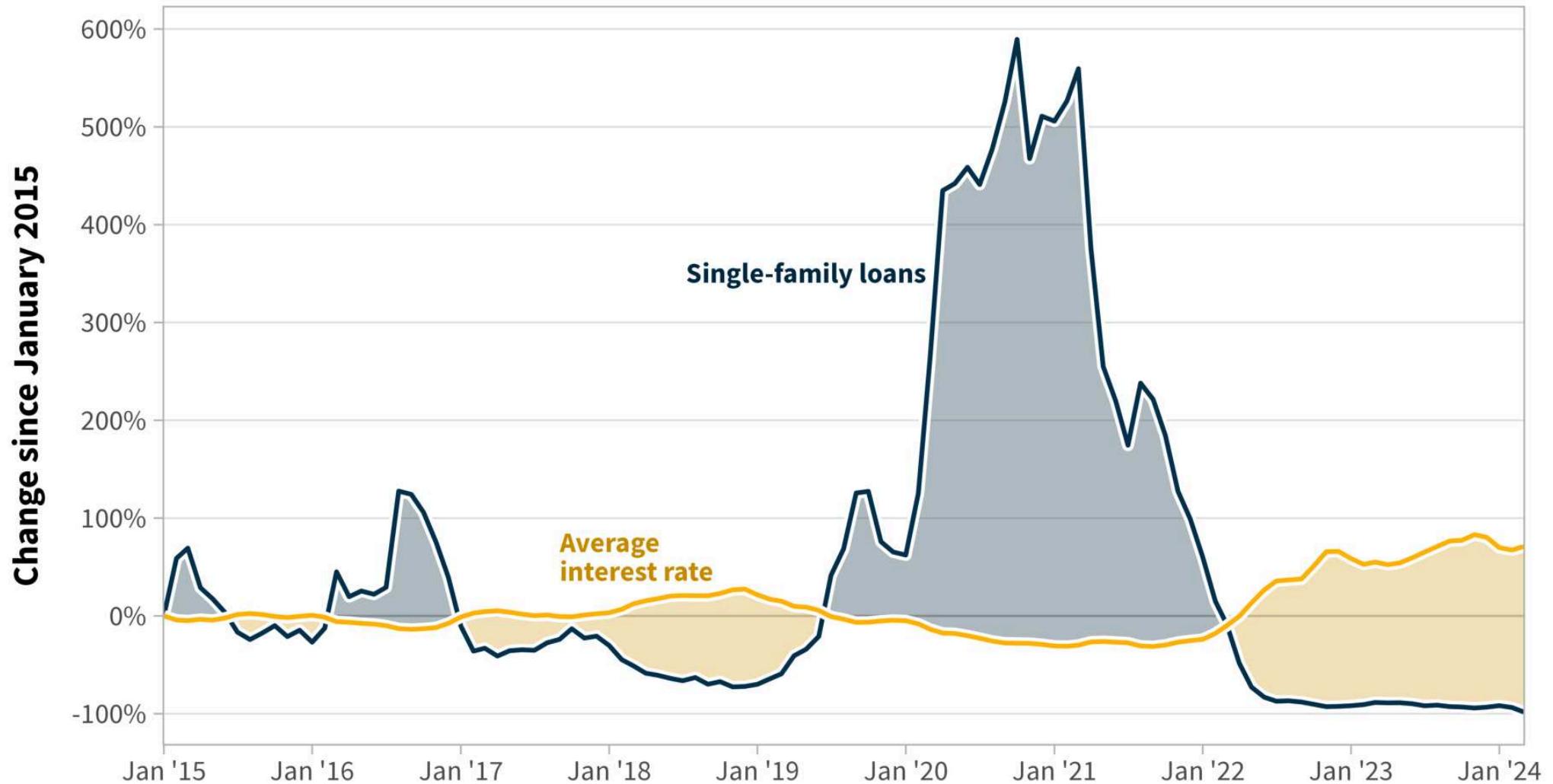
Average interest rate

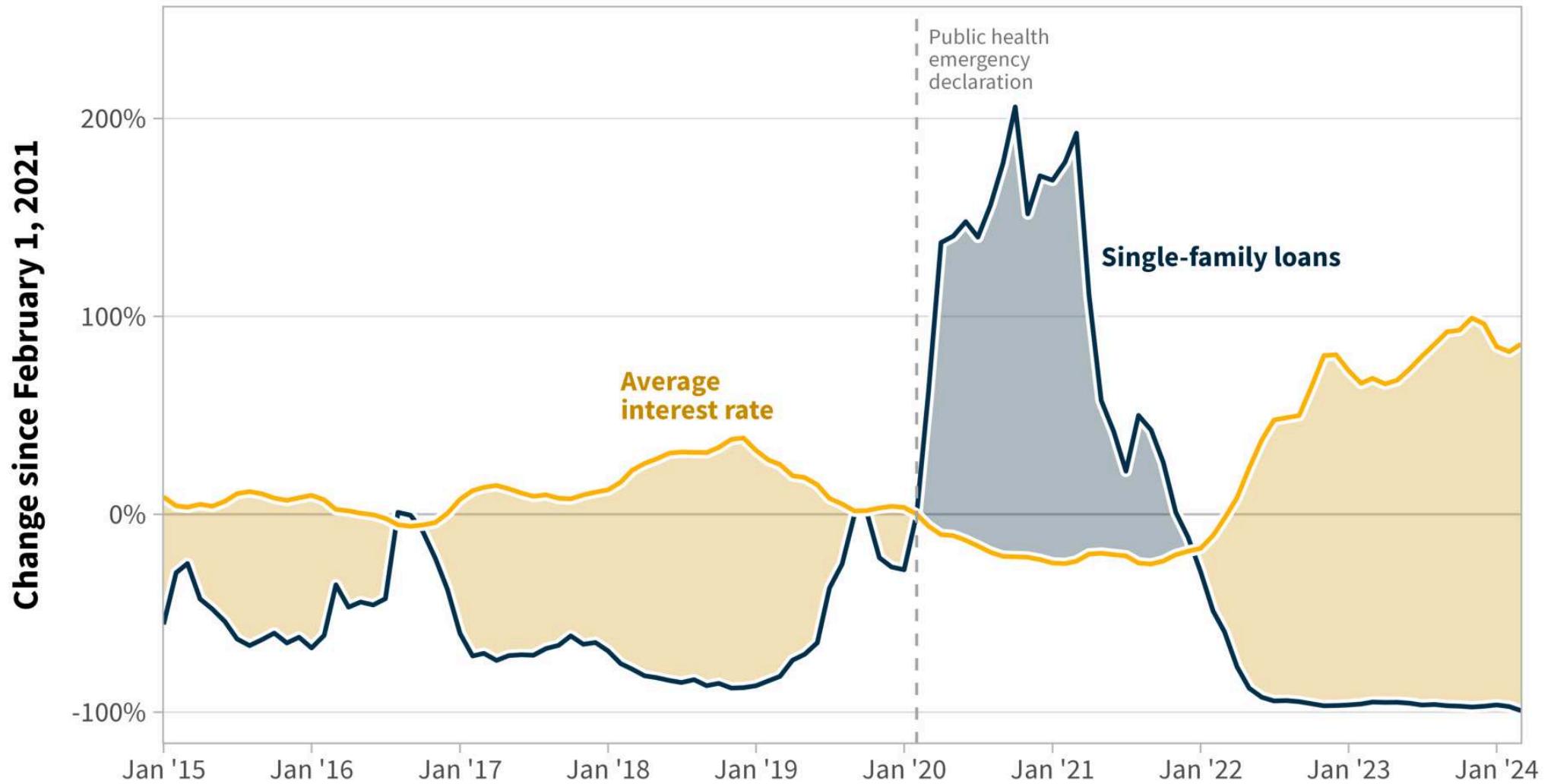




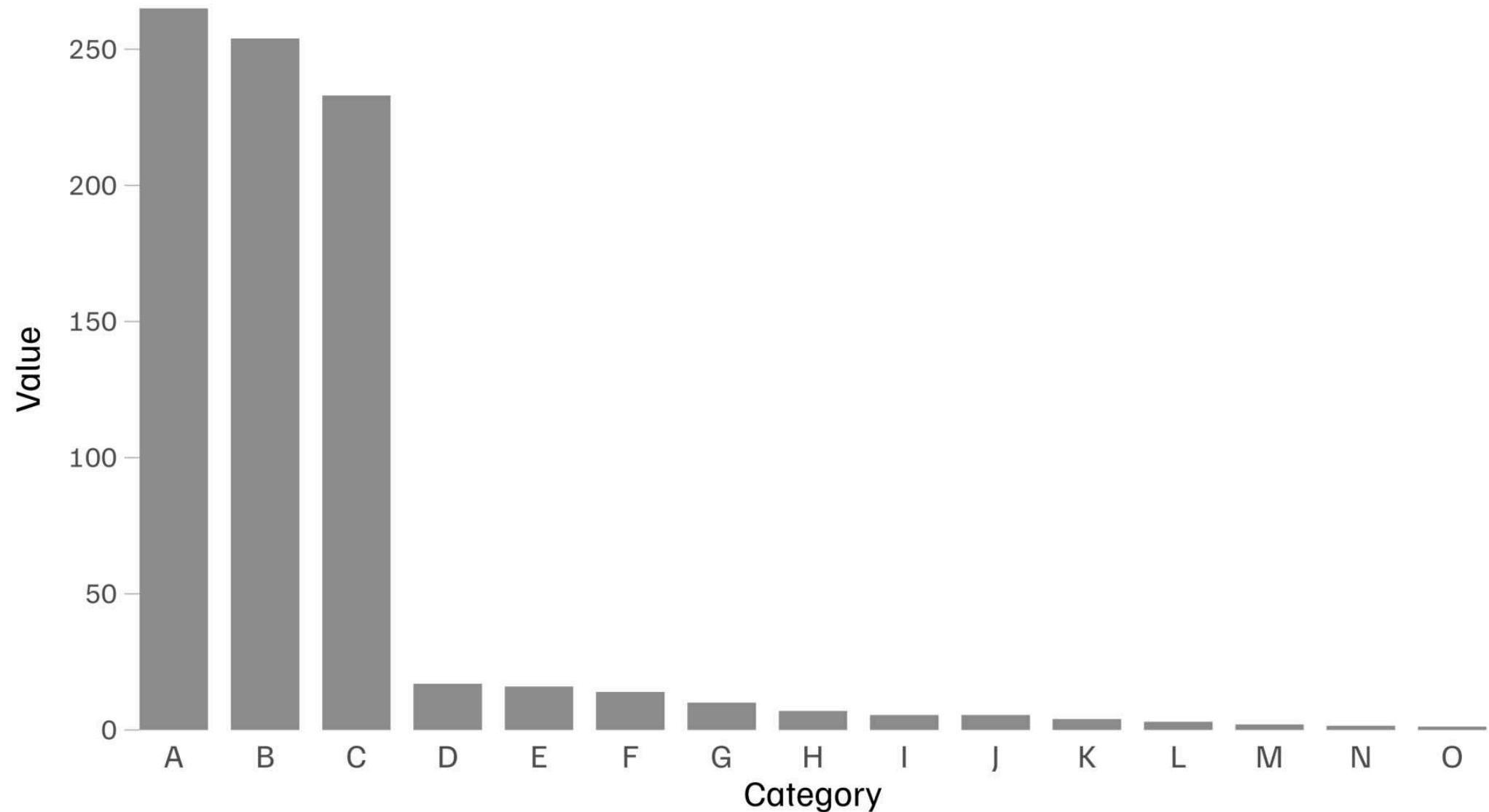




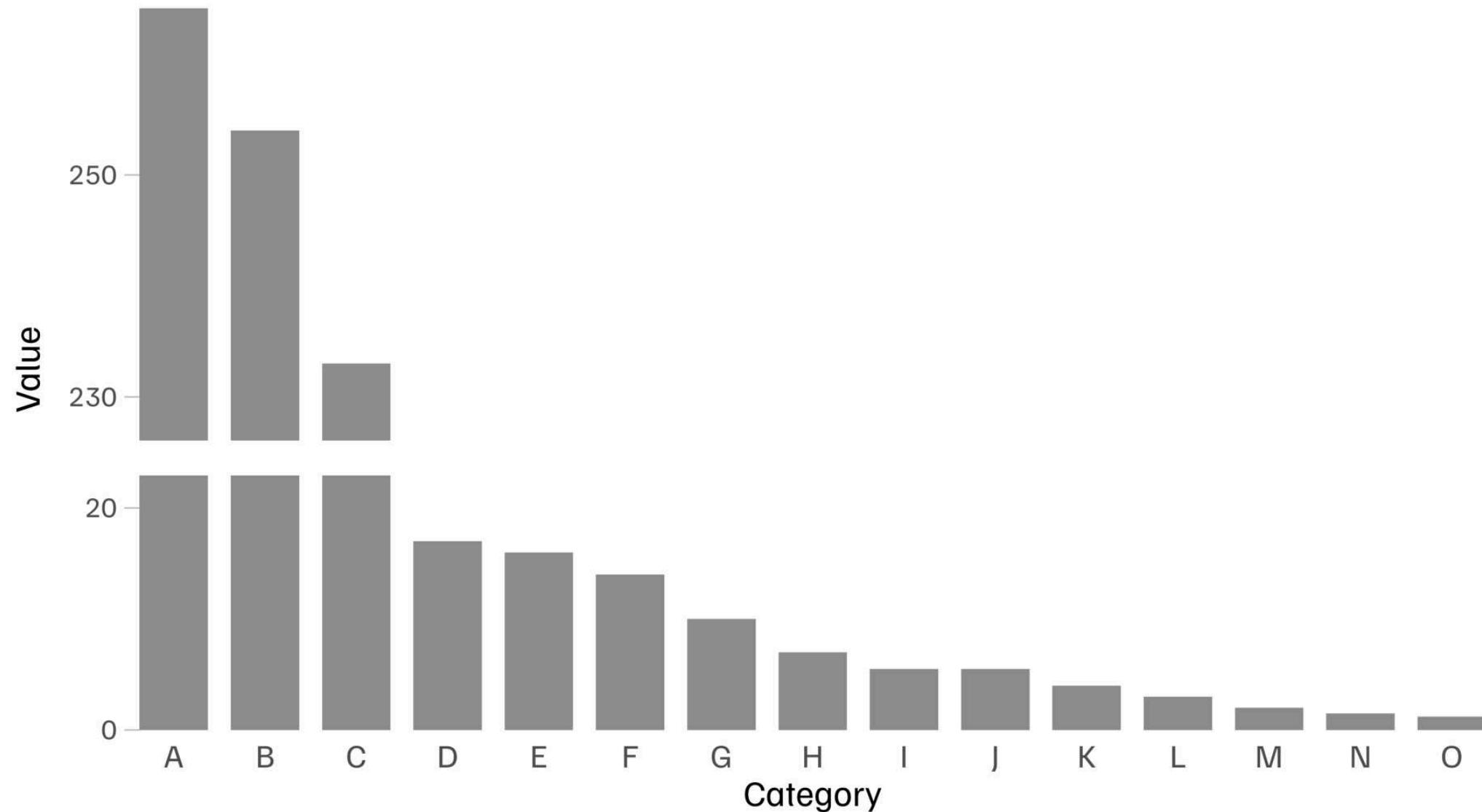




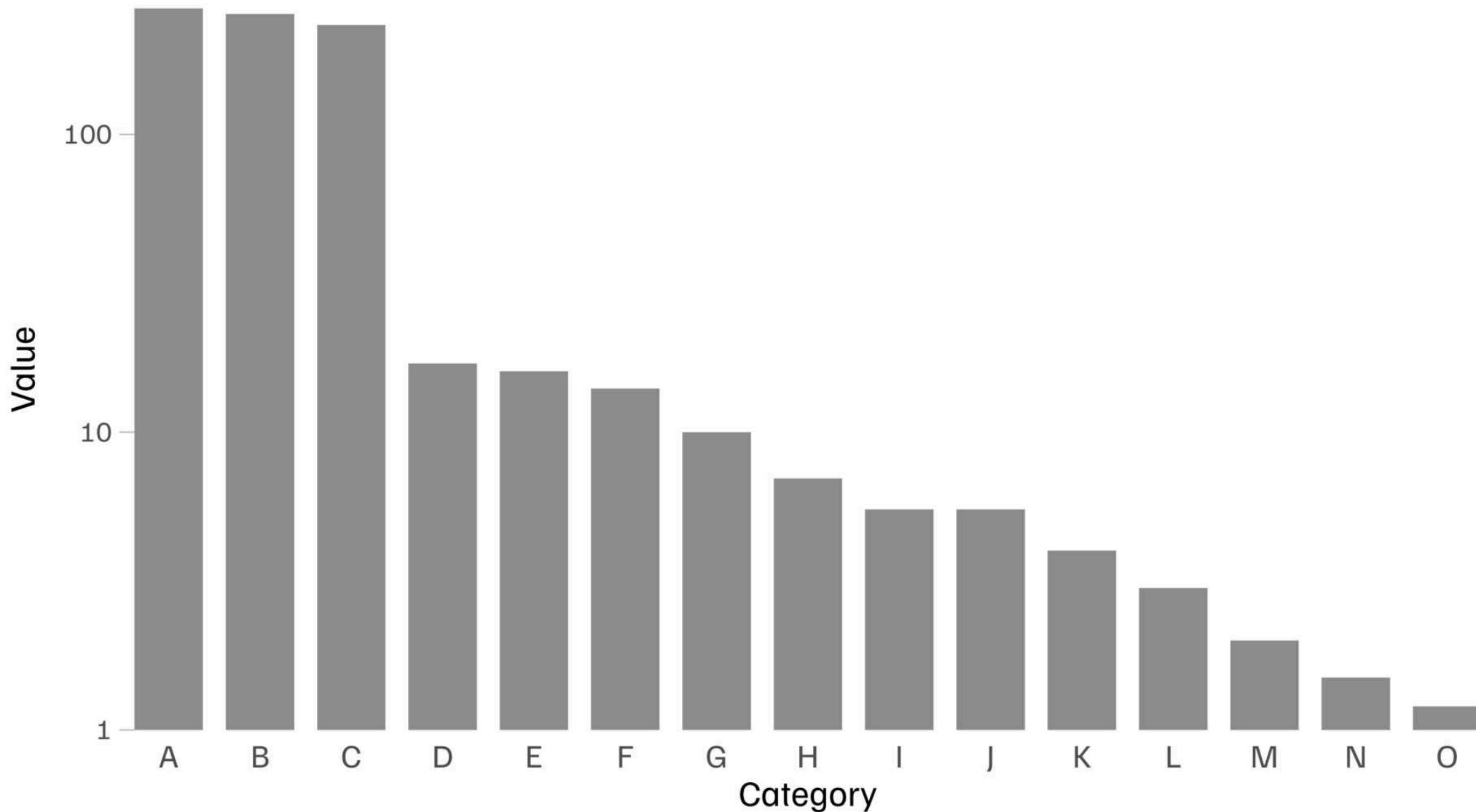
Skewed Data



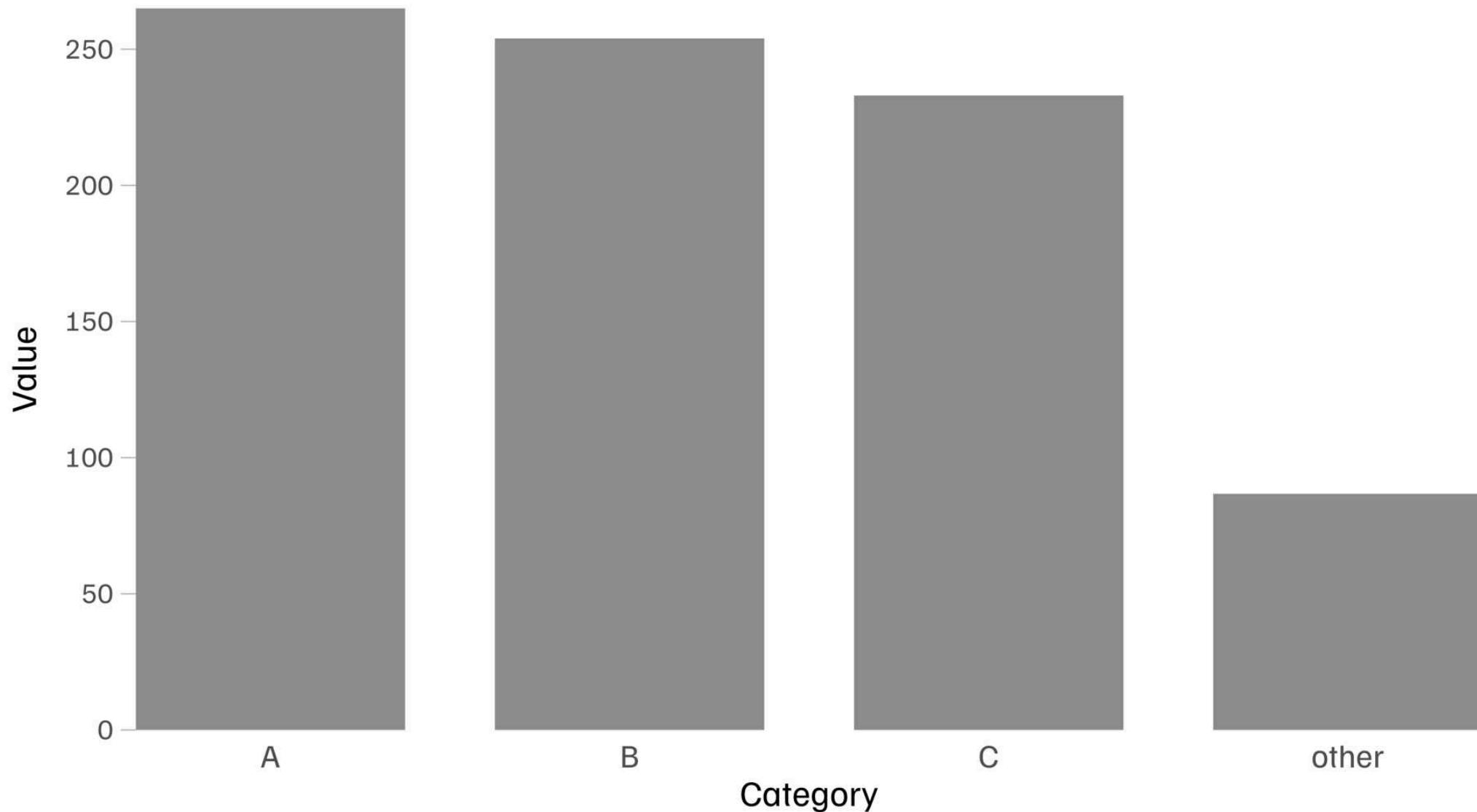
Skewed Data



Skewed Data

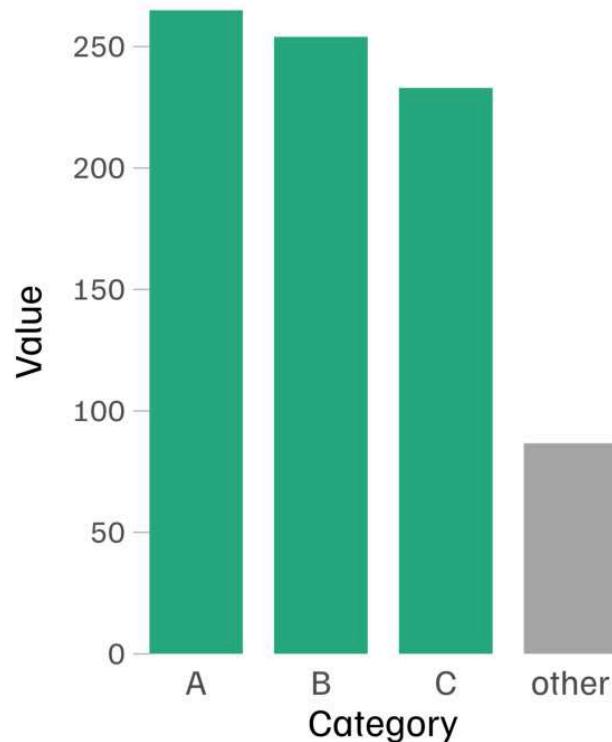


Skewed Data

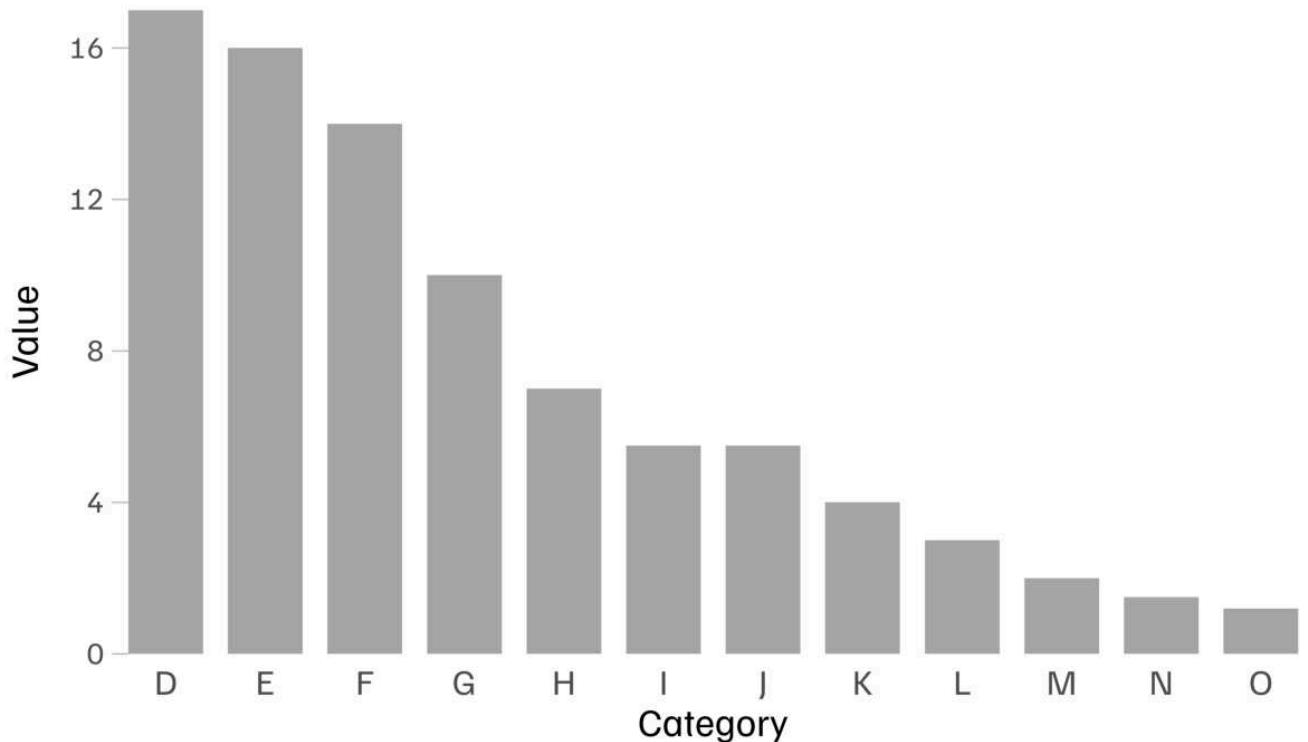


Skewed Data

All categories



Other categories



visual Form

(beauty)



Visual Form

- **design principles:**
 - structure the layouts of visualizations
 - balance unity and variety
 - create hierarchy

- **appearance:**
 - declutter your graphics
 - add meaningful cues



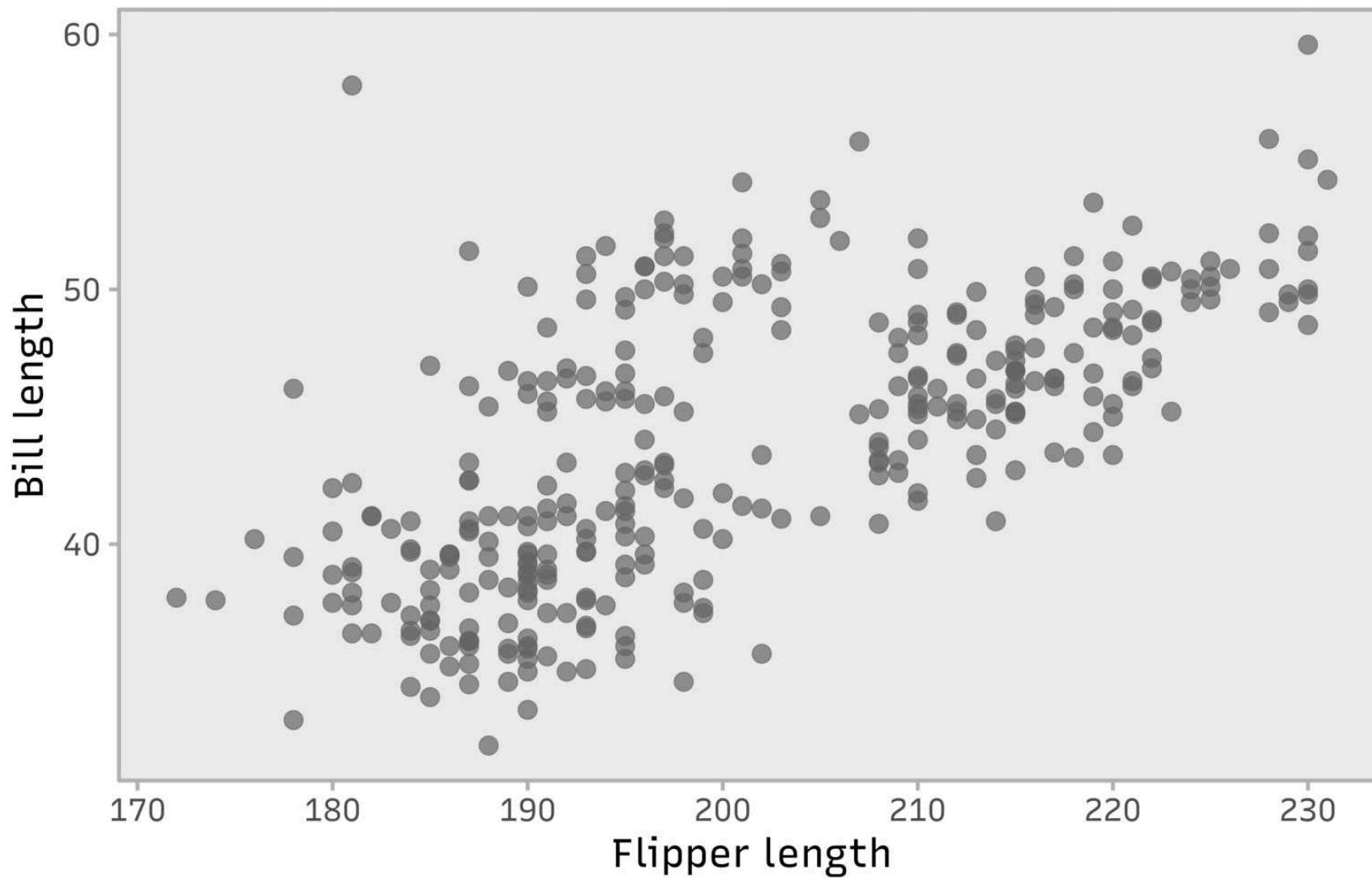
Gestalt Principles

A set of (varying) principles that describe how humans perceive and make sense of visual information and stimuli.

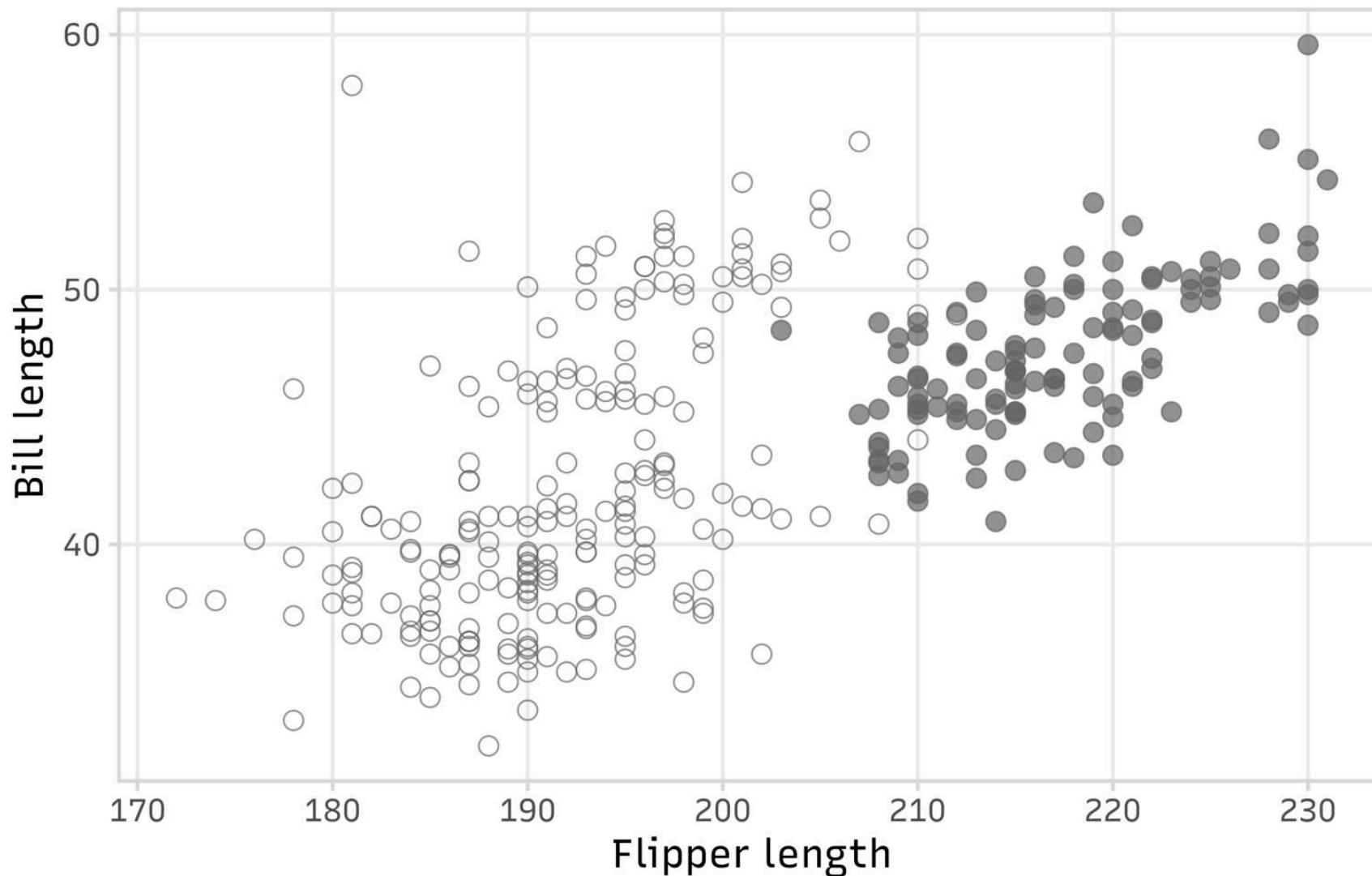
How they apply to data visualizations →



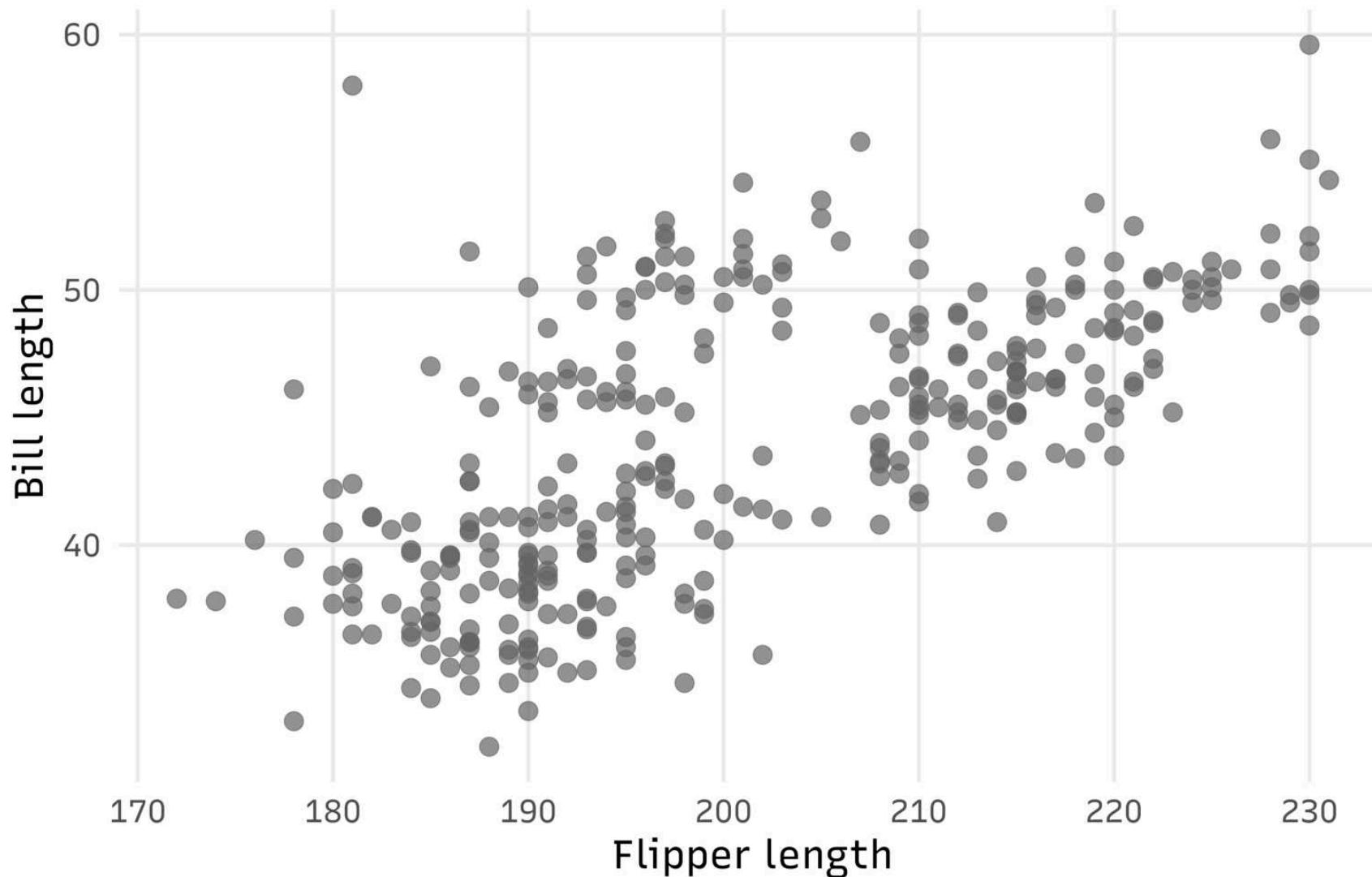
Principle of Figure and Ground



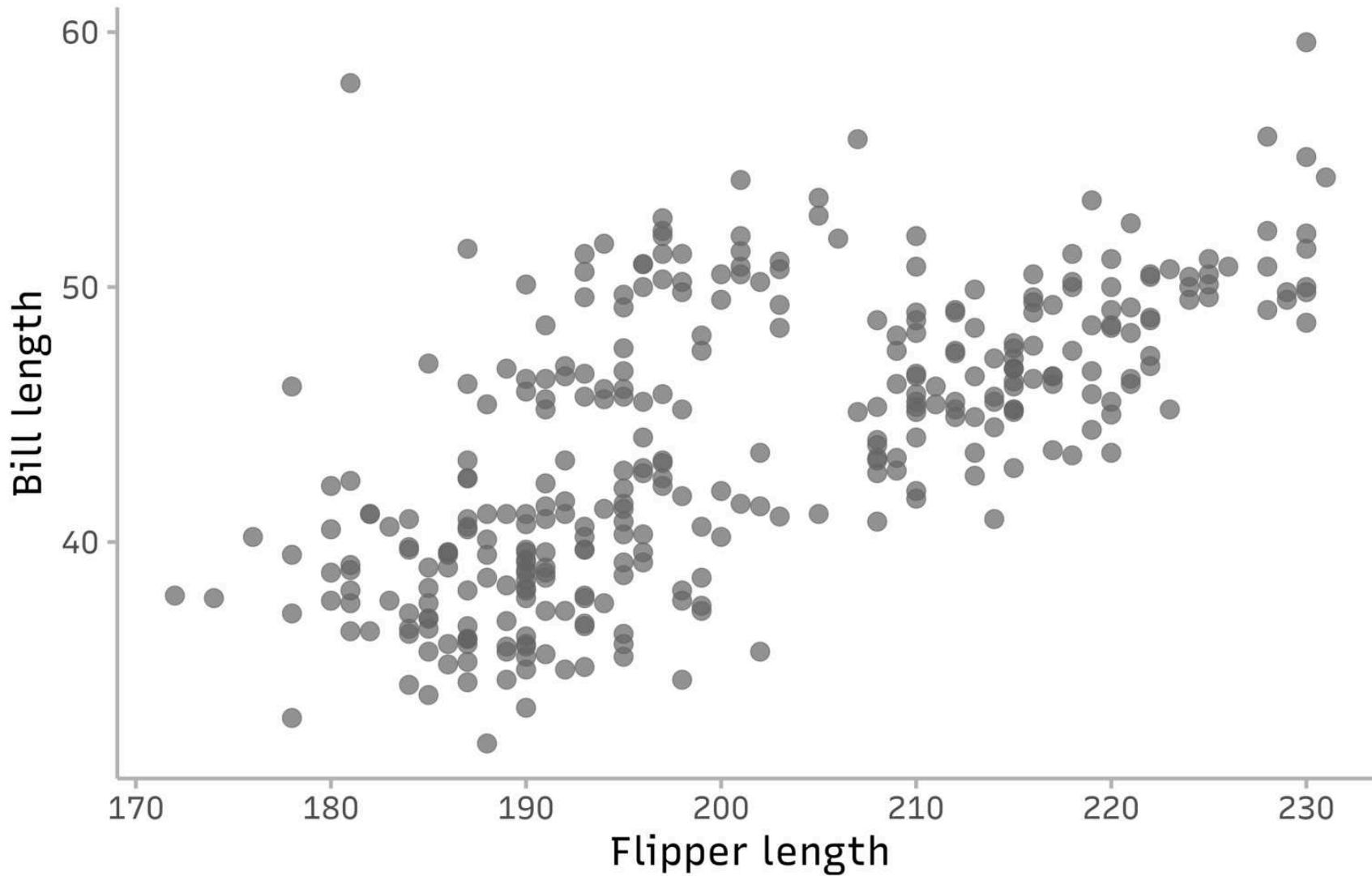
Principle of Figure and Ground



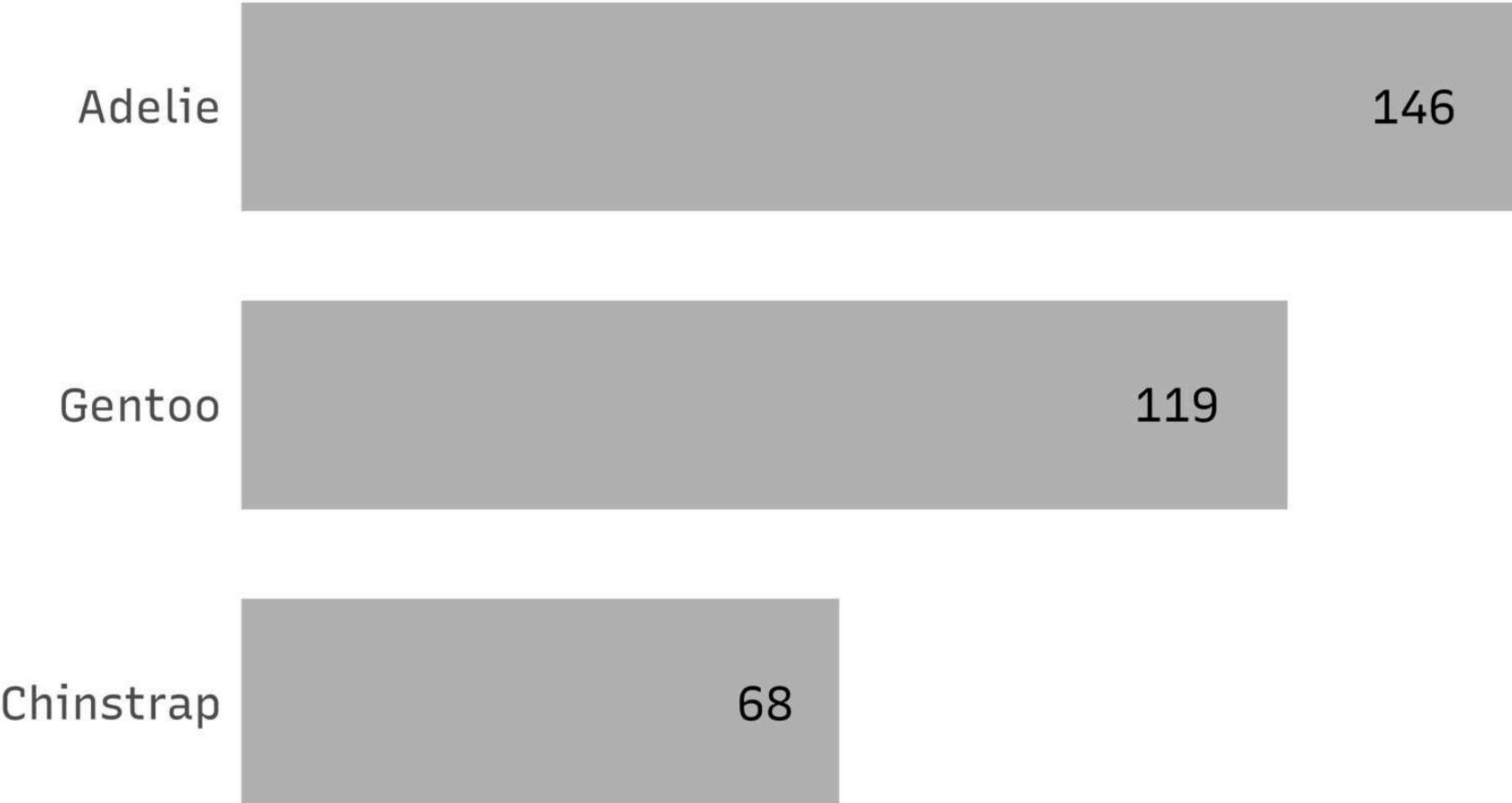
Principle of Closure



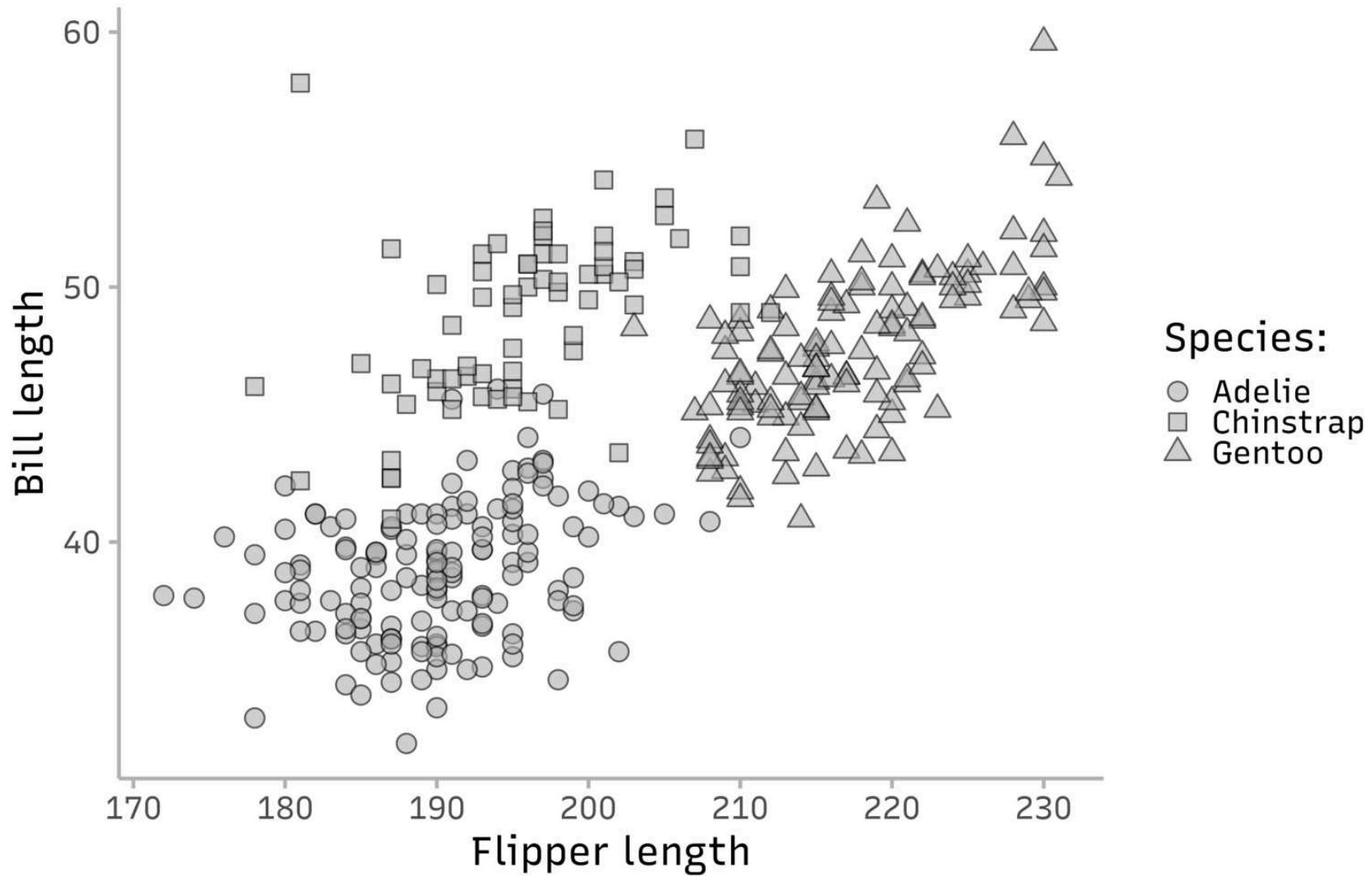
Principle of Closure



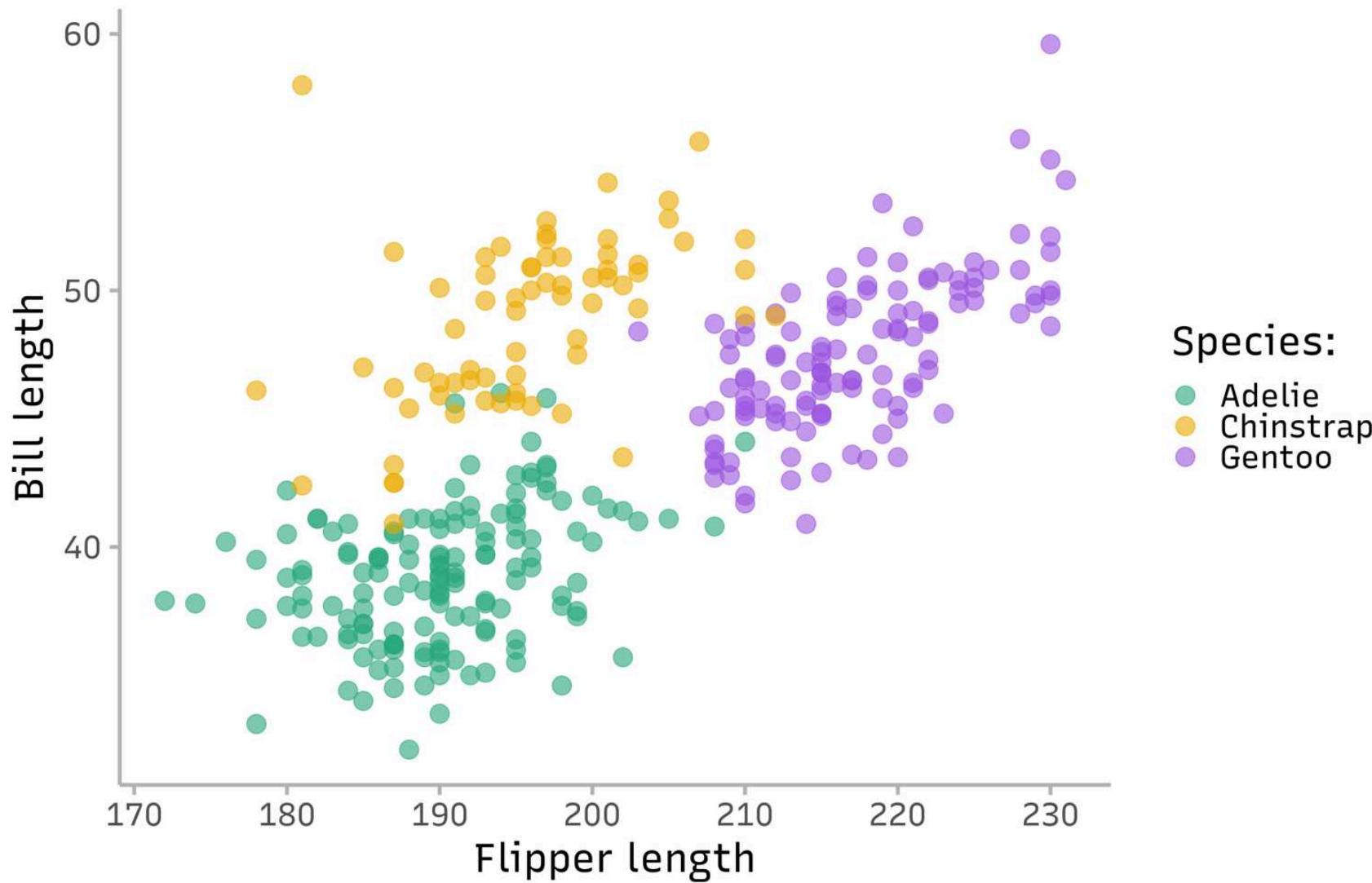
Principle of Closure



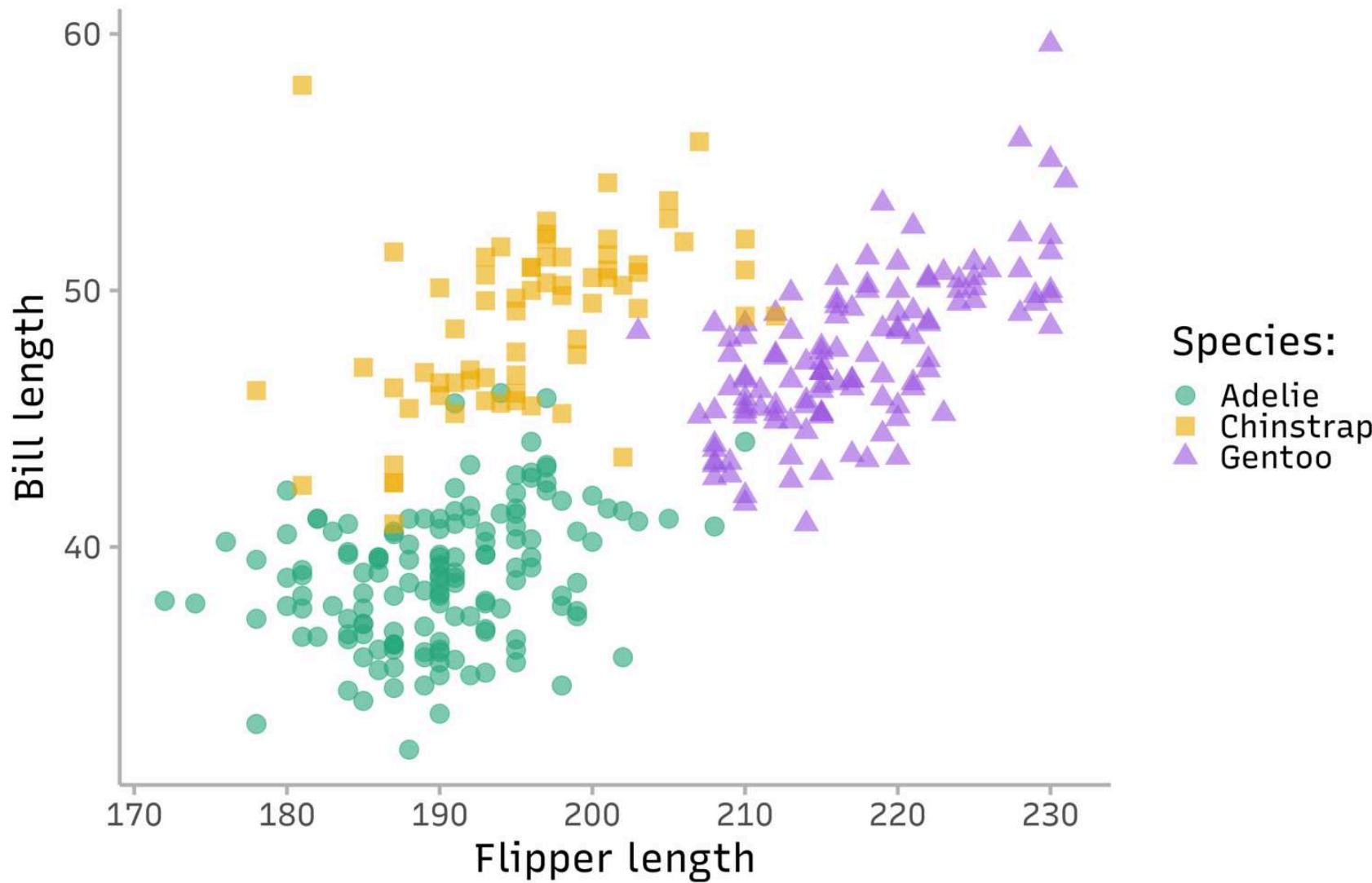
Principle of Similarity



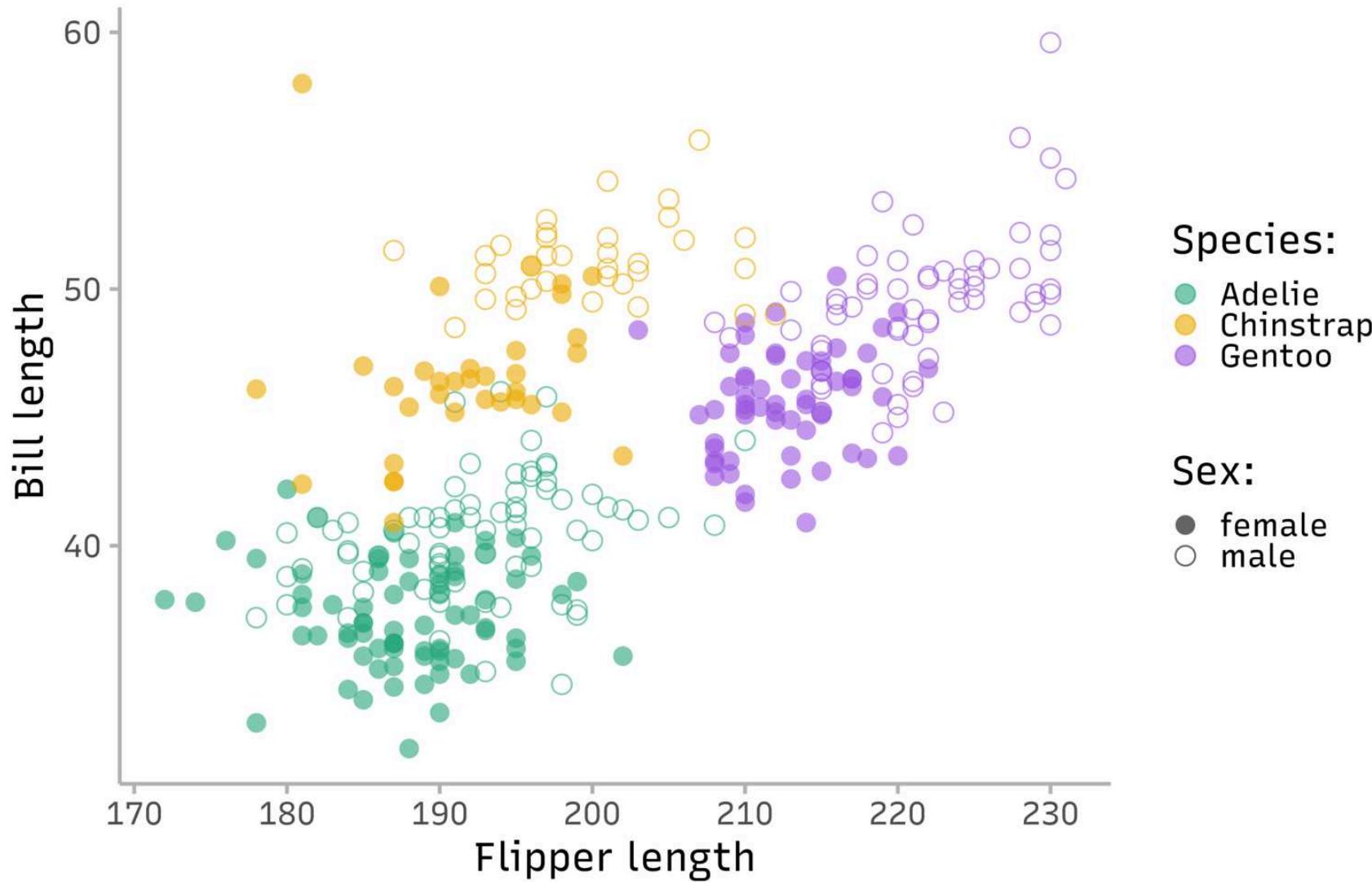
Principle of Similarity



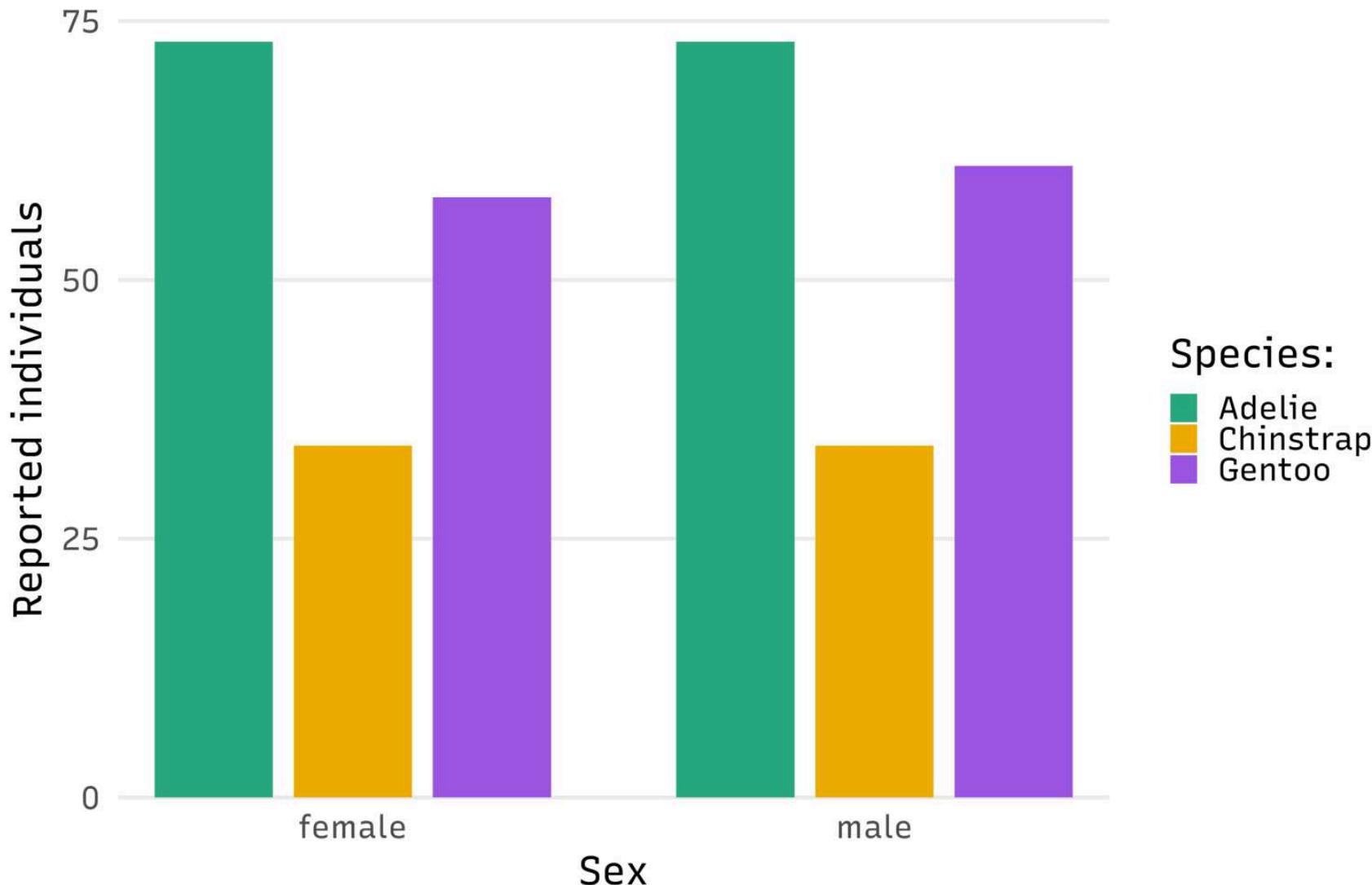
Principle of Similarity



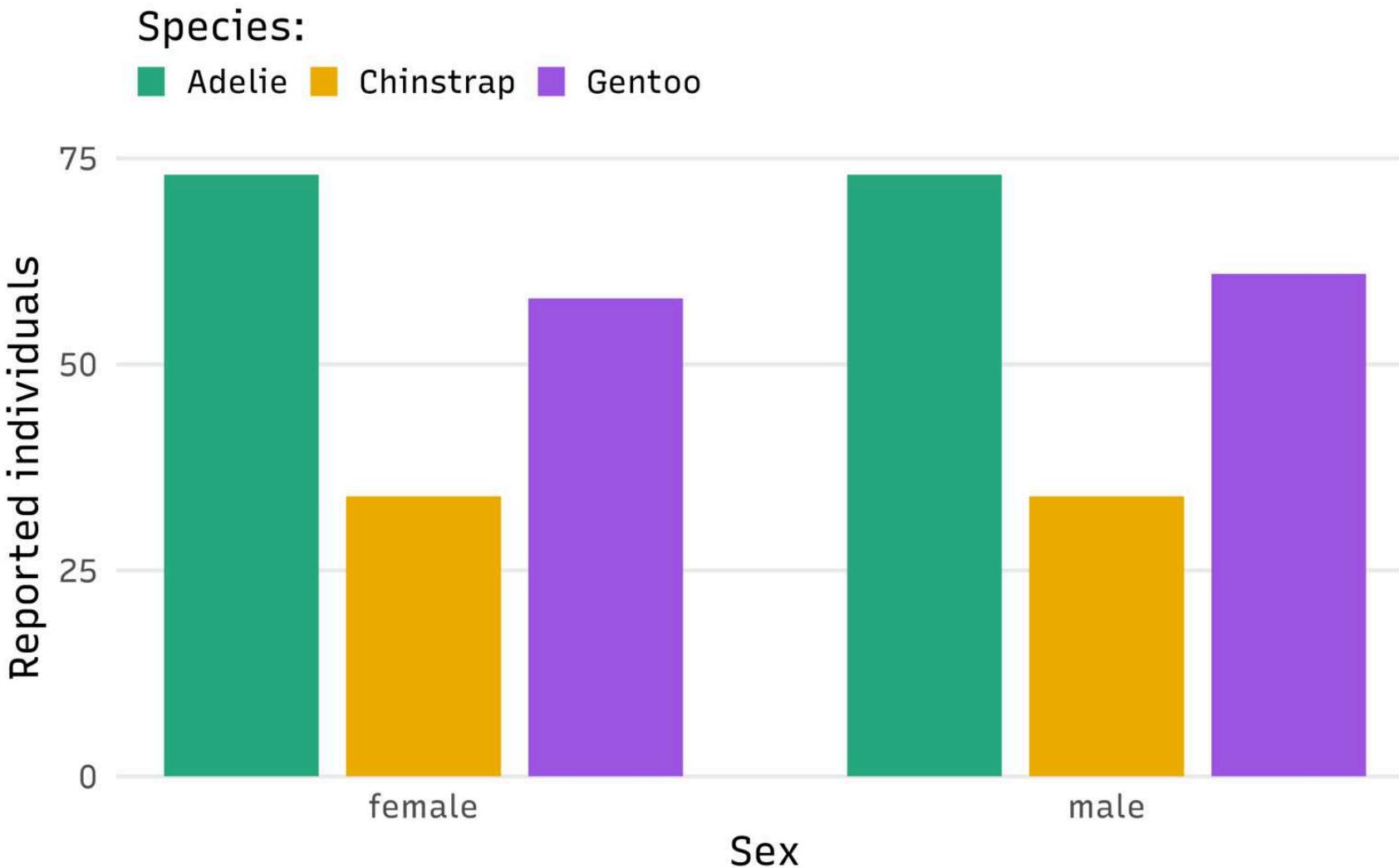
Principle of Similarity



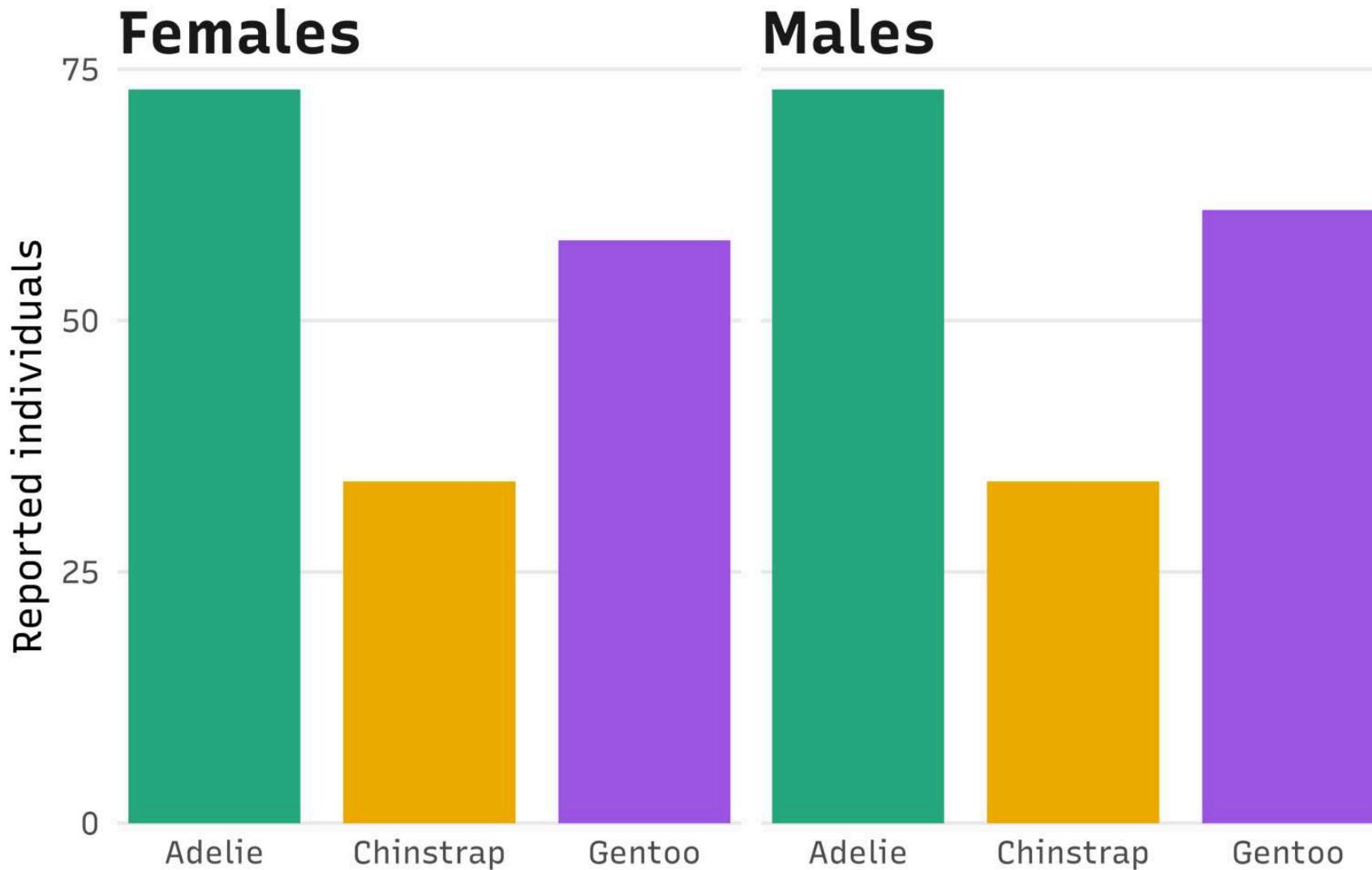
Principle of Proximity



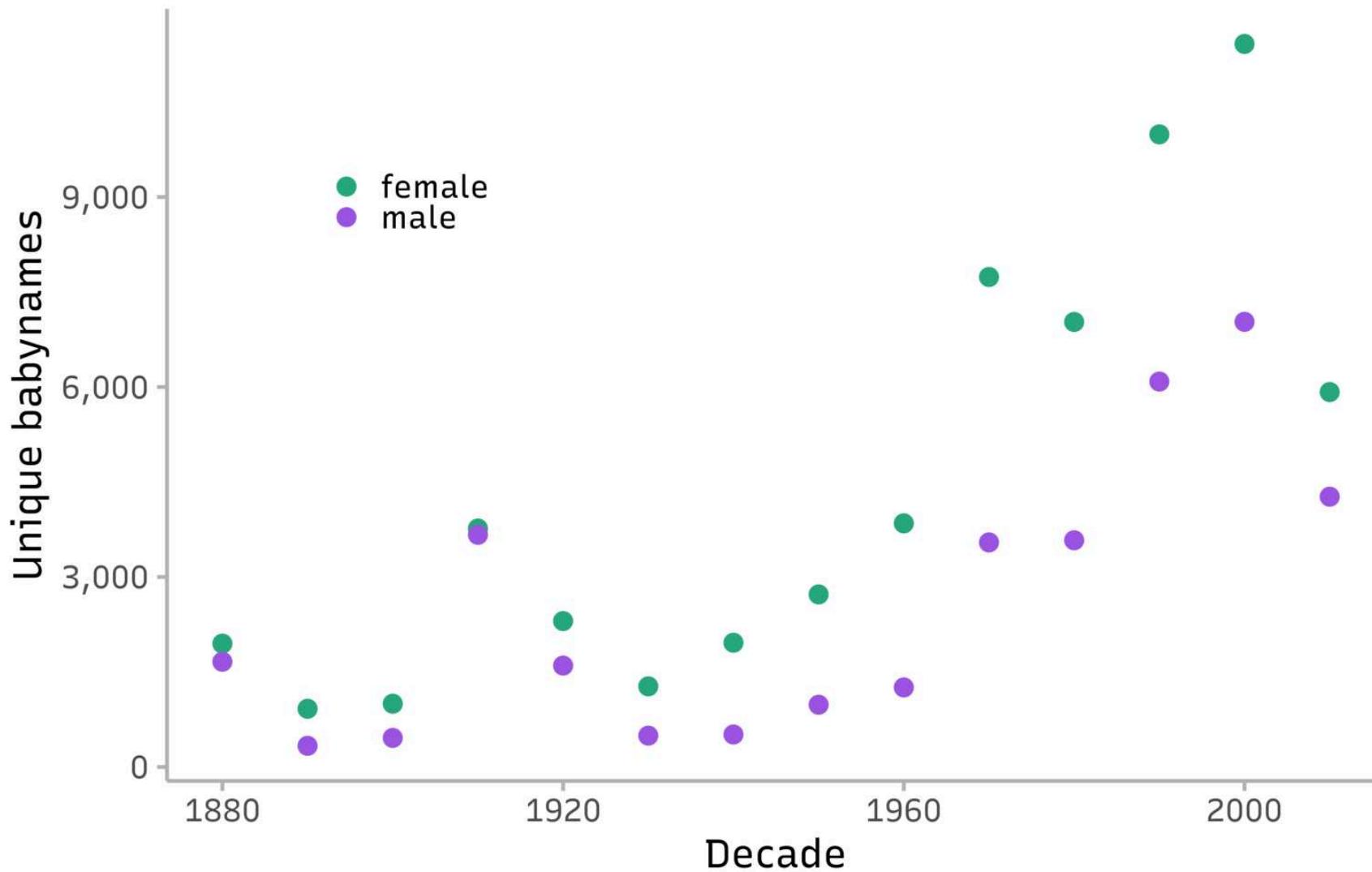
Principle of Proximity



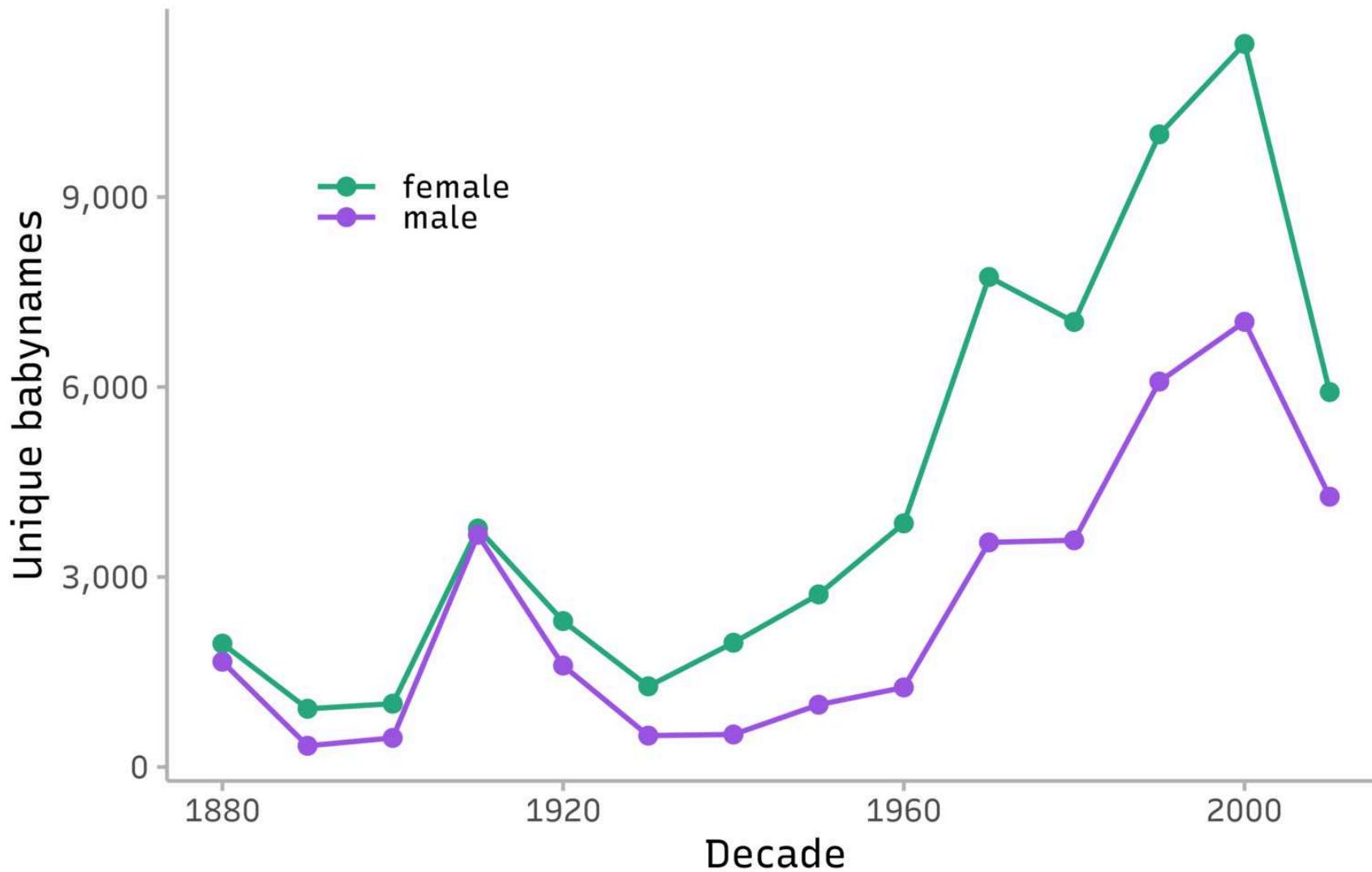
Principle of Proximity



Principle of Connectivity



Principle of Connectivity



Principles of Visual Perception

**"Make your audience look
where you want them to look."**



Get ready!

How many 4's are shown on the next slide?



5 3 7 2 0 9 4 8 1 3
9 4 5 4 3 5 2 5 4 6
1 0 5 8 2 4 6 8 5 0
7 3 2 9 6 0 2 5 9 1
4 8 1 2 0 7 0 2 4 3





Source: "Where's Waldo?" by Martin Handford © Kilburn & Strode LLP

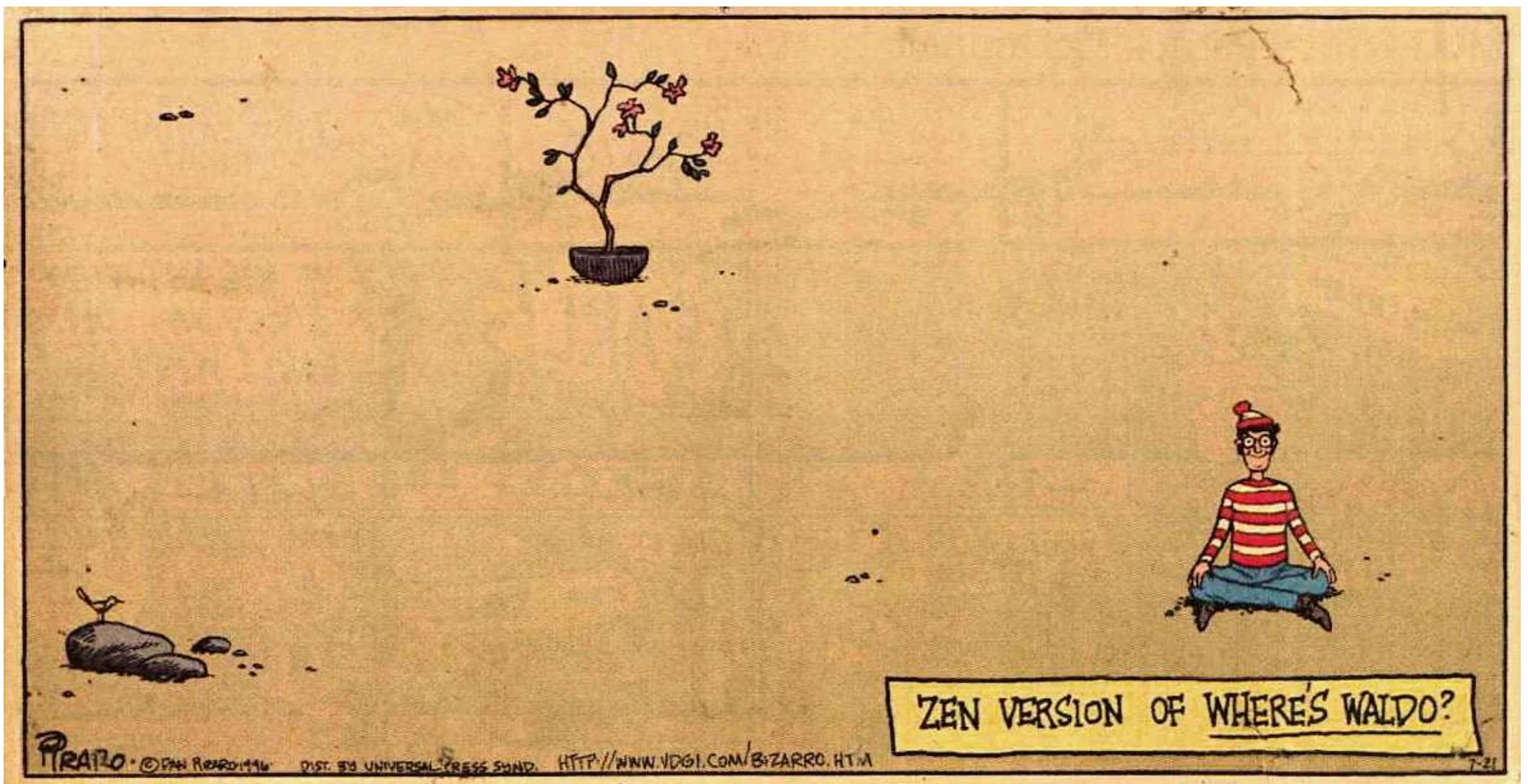




Source: "Where's Waldo?" by Martin Handford © Kilburn & Strode LLP

Cédric Scherer Data Visualization & Information Design





Source: Zen Version of "Where's Waldo?" by Dan Piraro



Iconic Memory

is a type of **sensory memory** that briefly retains visual information for a very short period.

Iconic memory operates at a pre-attentive level, meaning that it functions before conscious attention is directed toward the visual stimuli.



Preattentive Attributes

Visual properties that the human brain can instantly and subconsciously perceive.



Keep it simple

In his book [Information Visualization: Perception for Design](#), Colin Ware states:

“It is easy to spot a hawk in a sky full of pigeons, but as the variety of birds increases the hawk becomes harder to pick out.”

In other words, **the more things are made different, the less any of them stand out.**

So, it is *good practice* to start with figuring out an item of interest you want to emphasise, and then trying to make it the one thing that is different, thus leveraging your contrast strategically.

Source: Deya Milcheva (5rdata.com)



Keep it simple

← size, weight, type + spacing

color + added mark

In his book Information Visualization: Perception for Design, Colin Ware states:

added mark

↑ spacing

“It is easy to spot a hawk in a sky full of pigeons, but as the variety of birds increases the hawk becomes harder to pick out.”

↔ position

↑ spacing

type + style

In other words, the more things are made different, the less any of them stand out.

↑ weight

style

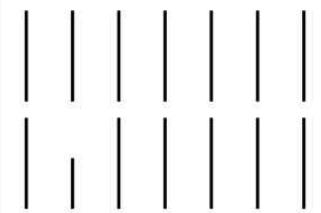
↑ spacing

So, it is good practice to start with figuring out an item of interest you want to emphasise, and then trying to make it the one thing that is different, thus leveraging your contrast strategically.

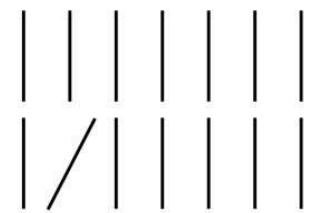
Source: Deya Milcheva (5rdata.com) ← size, color + spacing



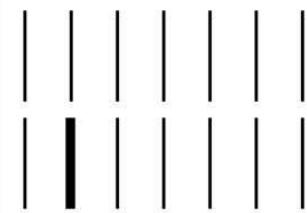
Length



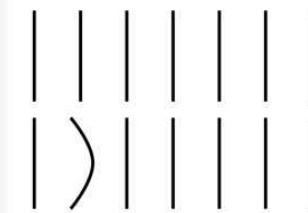
Orientation / Direction



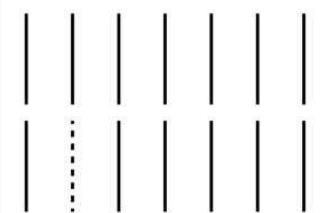
Width / Size



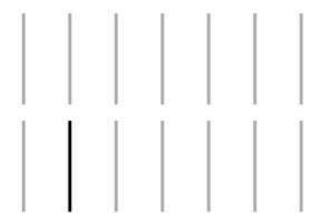
Curvature



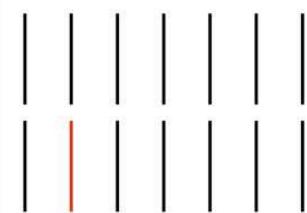
Shape / Linetype



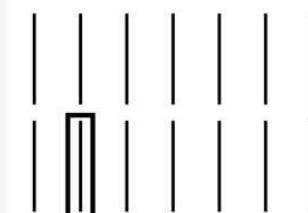
Intensity



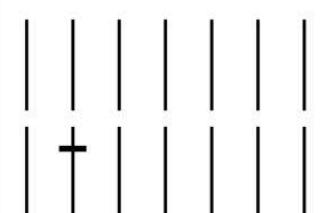
Hue



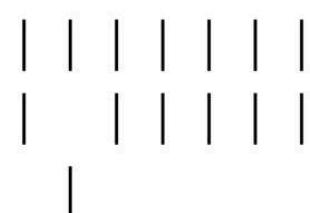
Enclosure



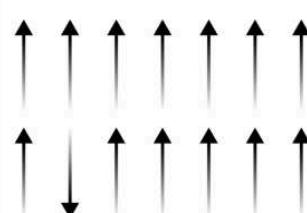
Added Marks



Position



Direction of Motion



Adapted from Stephen Few and others



5 3 7 2 0 9 4 8 1 3
9 4 5 4 3 5 2 5 4 6
1 0 5 8 2 4 6 8 5 0
7 3 2 9 6 0 2 5 9 1
4 8 1 2 0 7 0 2 4 3

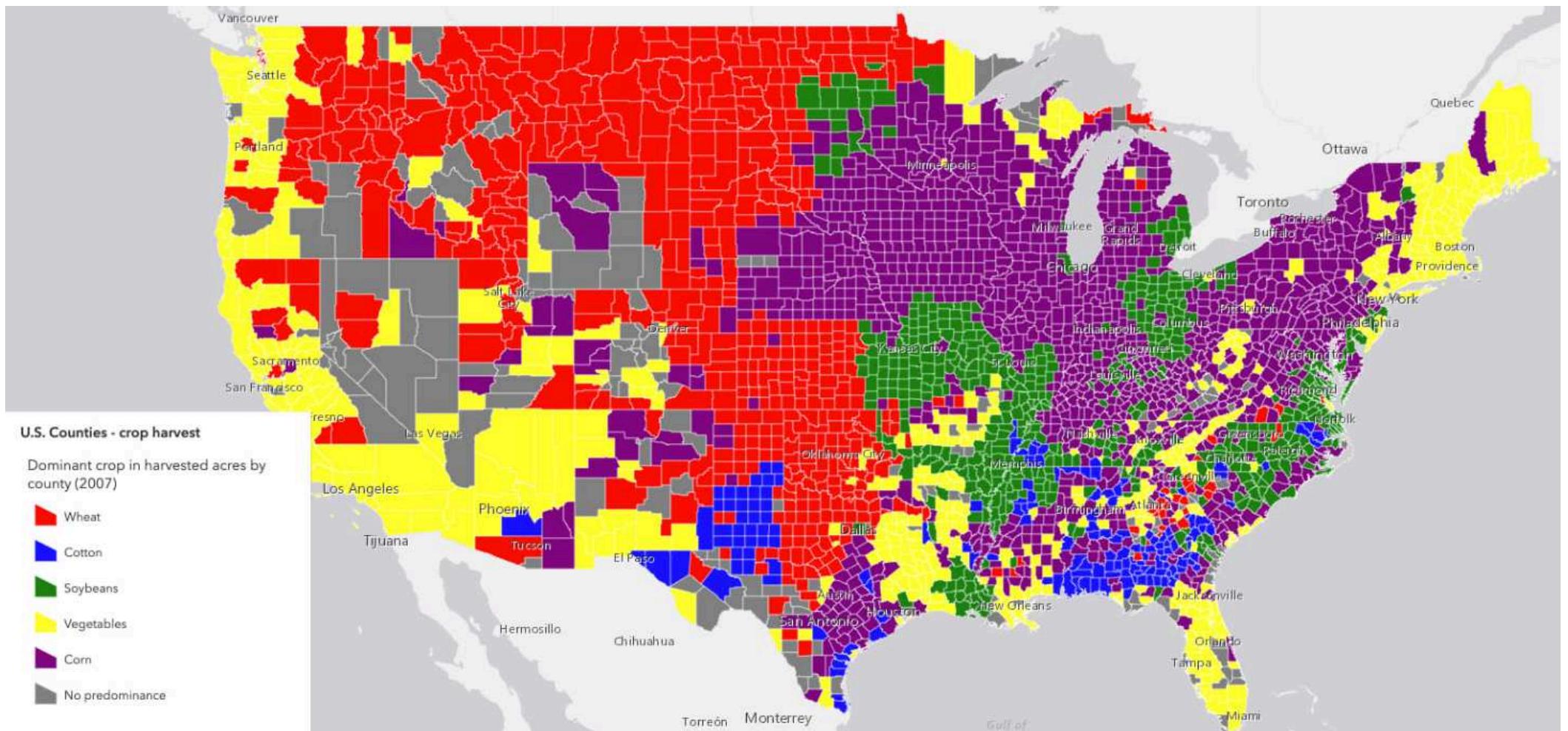


5 3 7 2 0 9 4 8 1 3
9 4 5 4 3 5 2 5 4 6
1 0 5 8 2 4 6 8 5 0
7 3 2 9 6 0 2 5 9 1
4 8 1 2 0 7 0 2 4 3



5	3	7	2	0	9	4	8	1	3
9	4	5	4	3	5	2	5	4	6
10	5	8	2	4	6	8	5	0	
7	3	2	9	6	0	2	5	9	1
4	8	1	2	0	7	0	2	4	3





Source: ESRI



Eliminating Distractions

Cédric Scherer Data Visualization & Information Design



clutter

also known as

"chart Junk"

are visual elements that take up space
but do not increase understanding.

Even worse, they increase cognitive load.



“The larger the share of a graphic's ink devoted to data, the better.”

E. Tufte (1983)



Chart junk



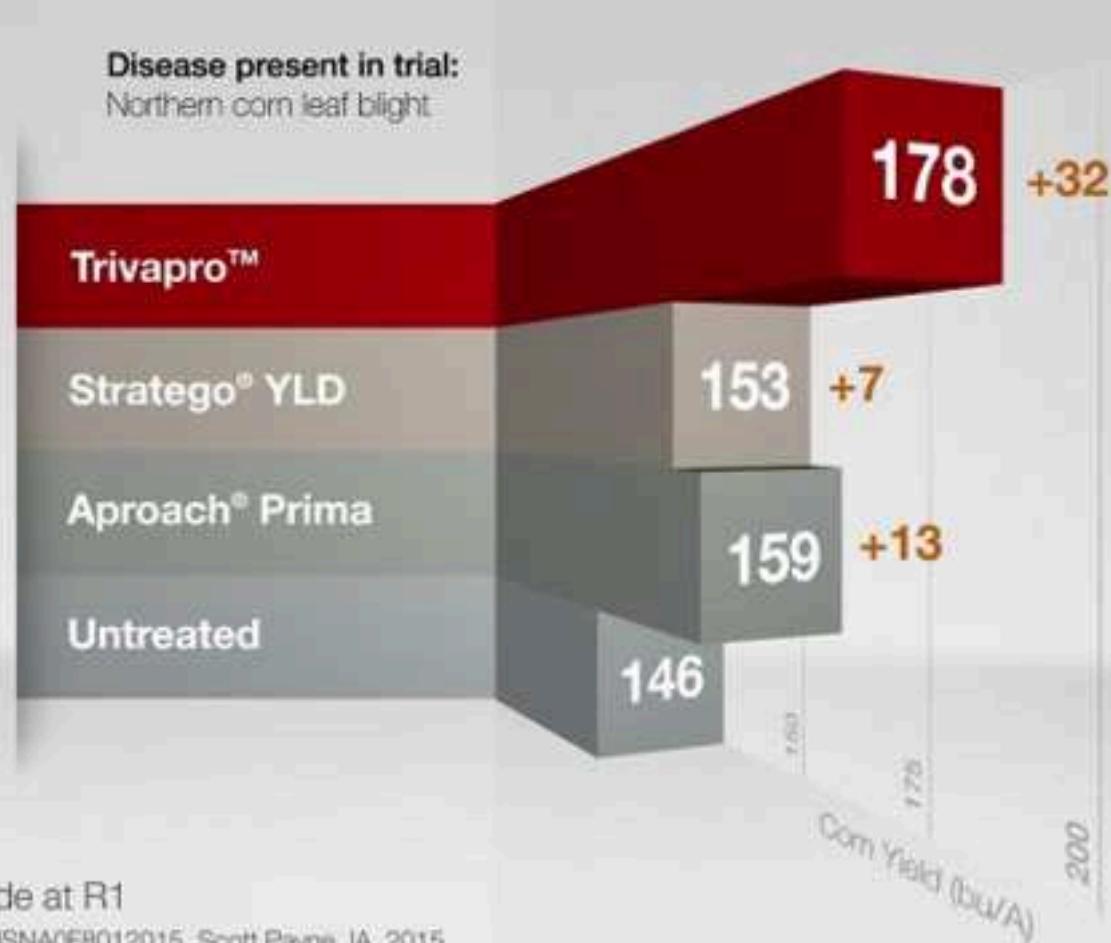
Found in "News that fit to print" by Arturo Perez-Reyes

Cédric Scherer Data Visualization & Information Design



Trivapro corn yield response

In Boone, IA



All applications made at R1

FSF001A4-2015JS, Trial USNA0F8012015, Scott Payne, IA, 2015

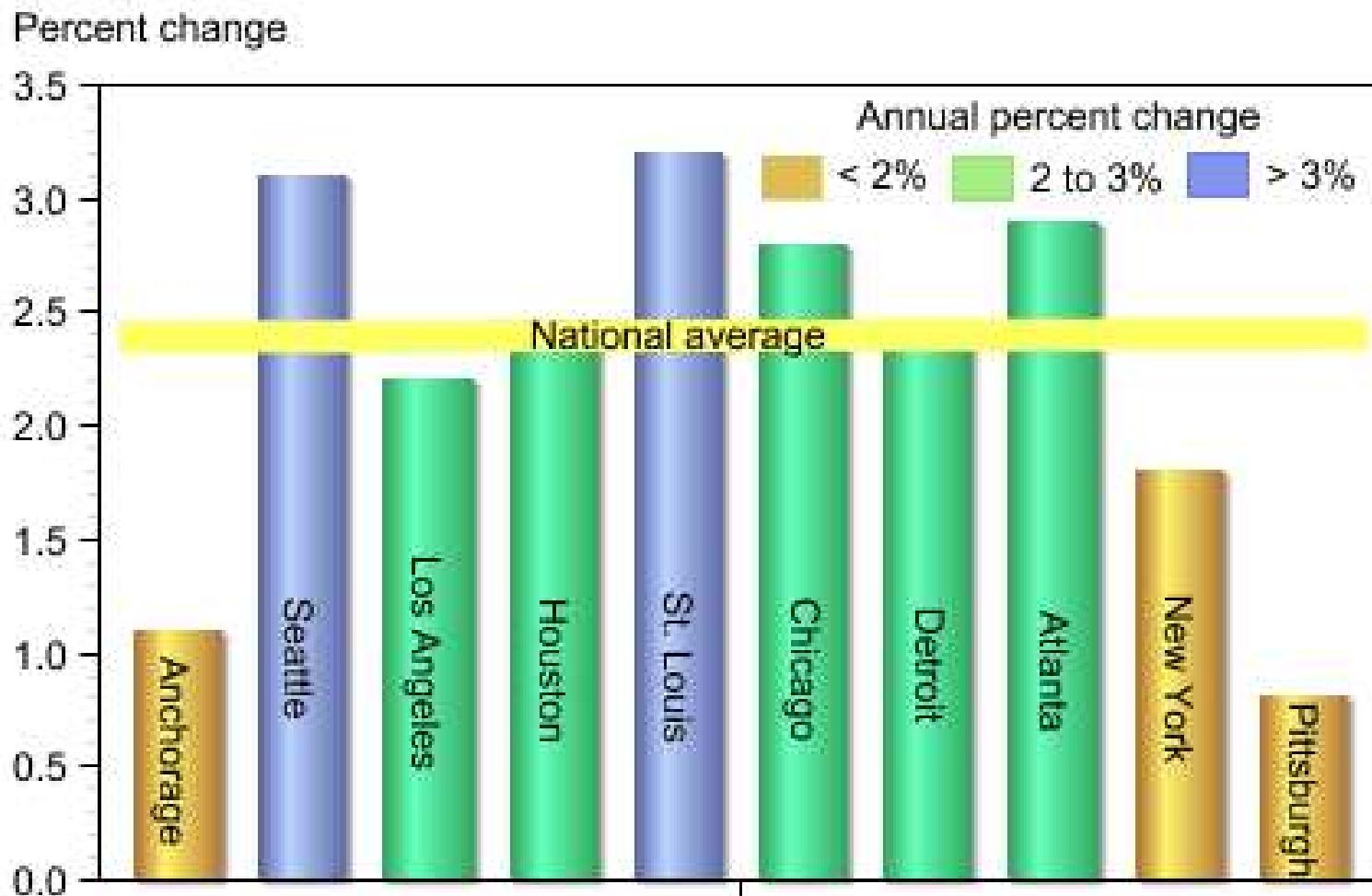
Source: Syngenta

Cédric Scherer Data Visualization & Information Design



Annual grocery store inflation by city, 2014

In 2014, food price inflation was higher in Seattle, St. Louis, Chicago, and Atlanta than in other metropolitan areas



Source: Calculated by ERS, USDA, using Bureau of Labor Statistics (BLS) data.



MONSTROUS COSTS

Total House and Senate campaign expenditures,
in millions



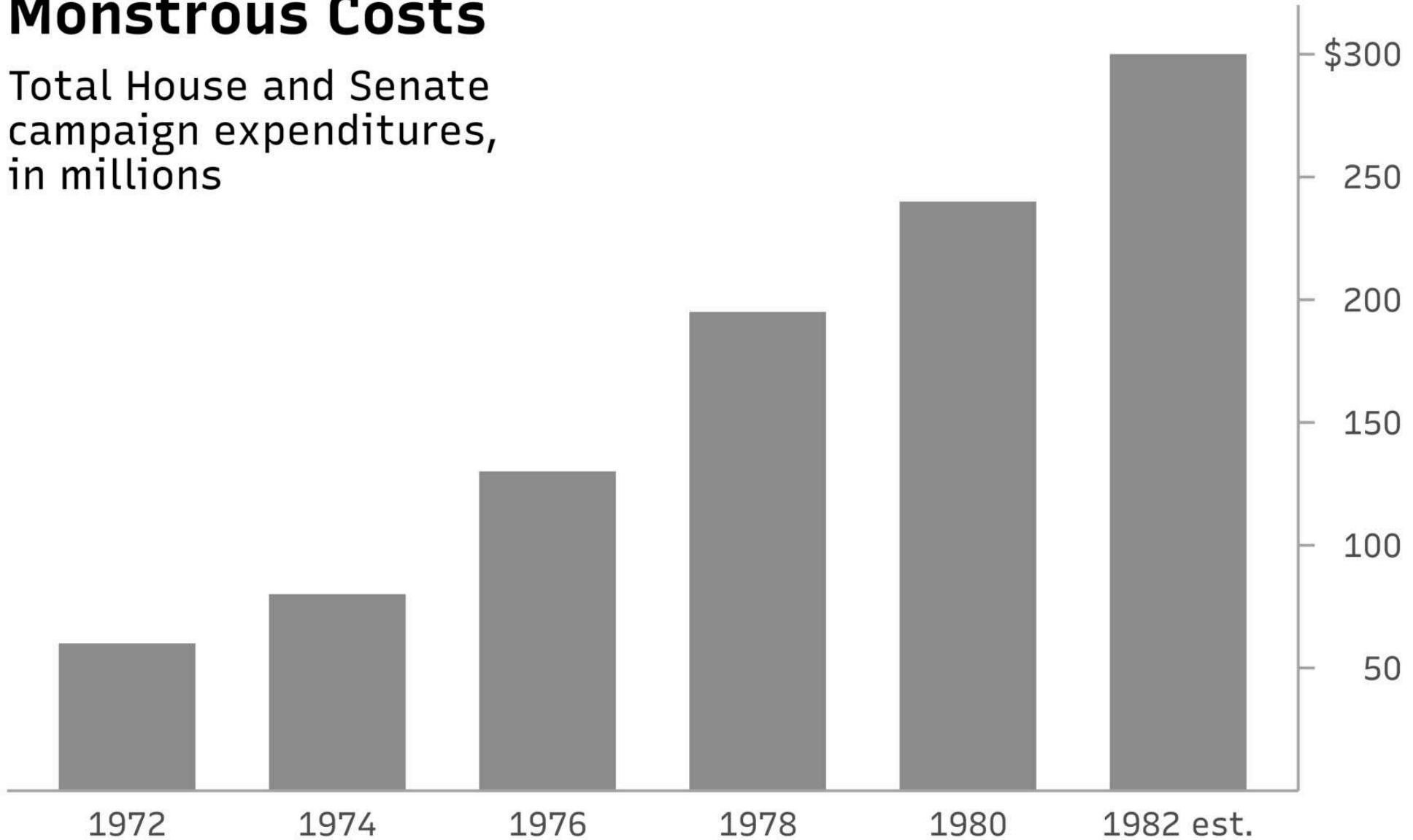
Source: "Monstrous Costs" by Nigel Holmes

Cédric Scherer Data Visualization & Information Design



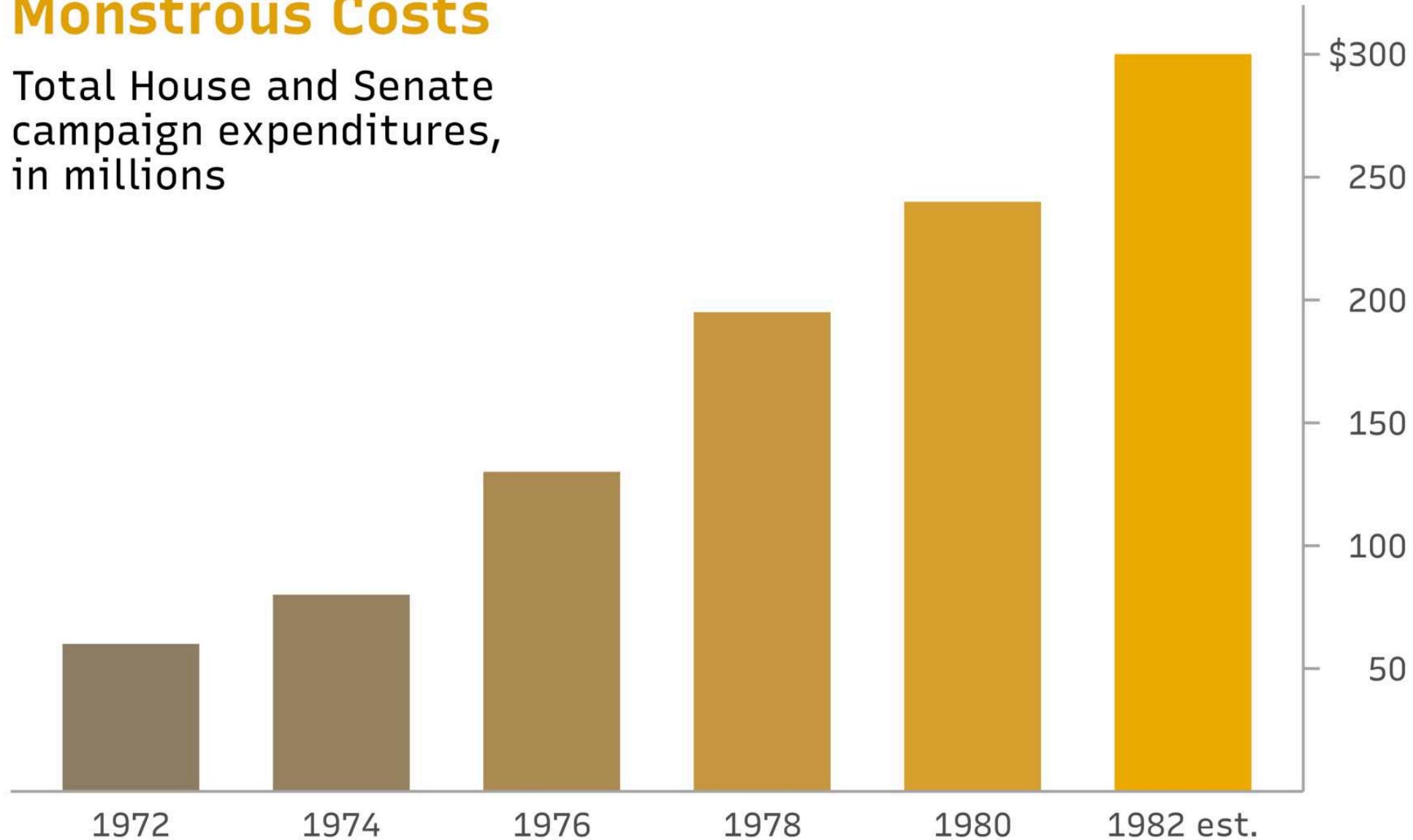
Monstrous Costs

Total House and Senate
campaign expenditures,
in millions



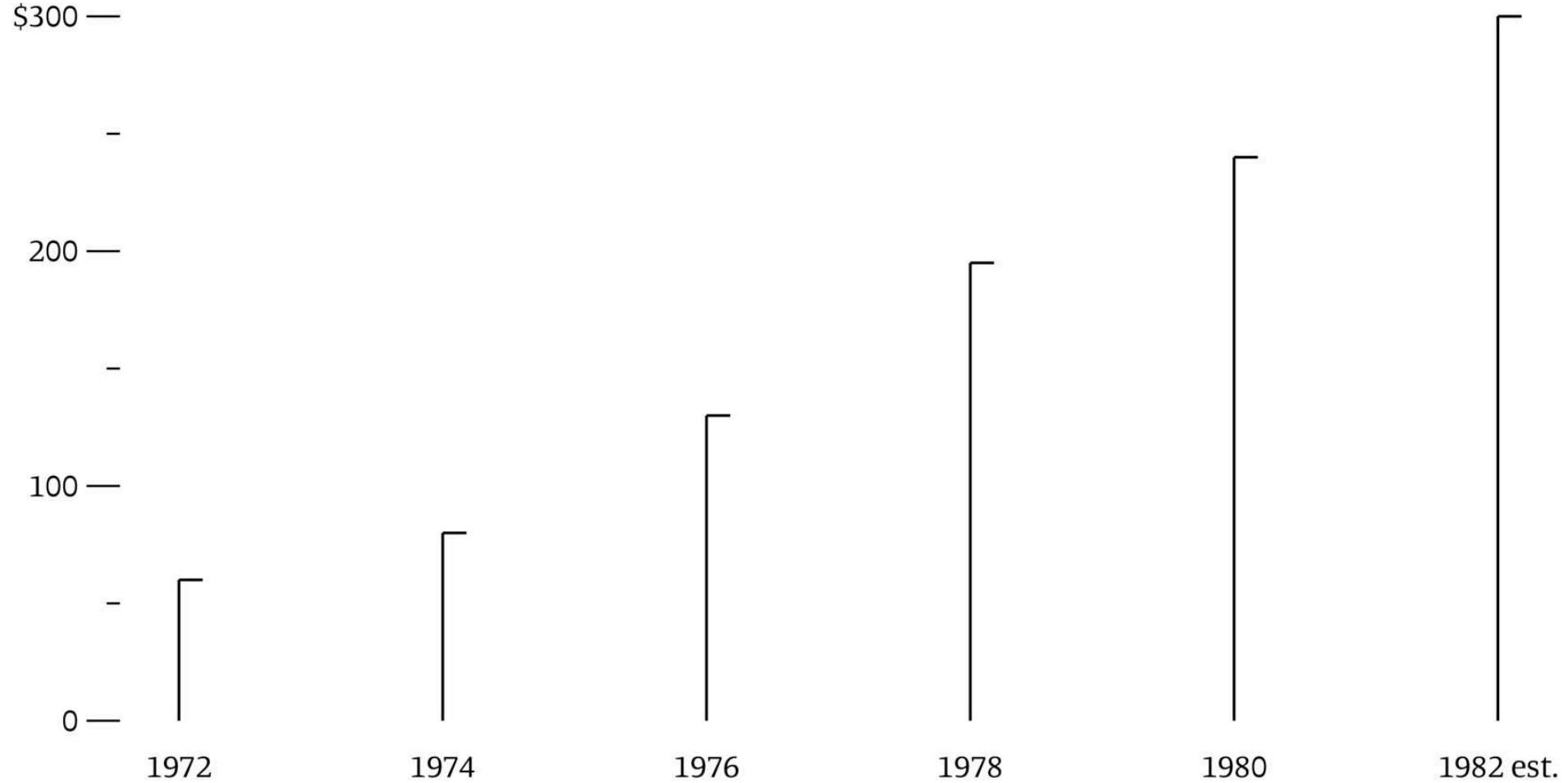
Monstrous Costs

Total House and Senate
campaign expenditures,
in millions



Monstrous Costs

Total House and Senate campaign expenditures, in millions



*"People's accuracy in
describing the embellished charts was
no worse than for plain charts, and [...]
**their recall after a two-to-three week
gap was significantly better."***

S. Bateman et al. (2010)

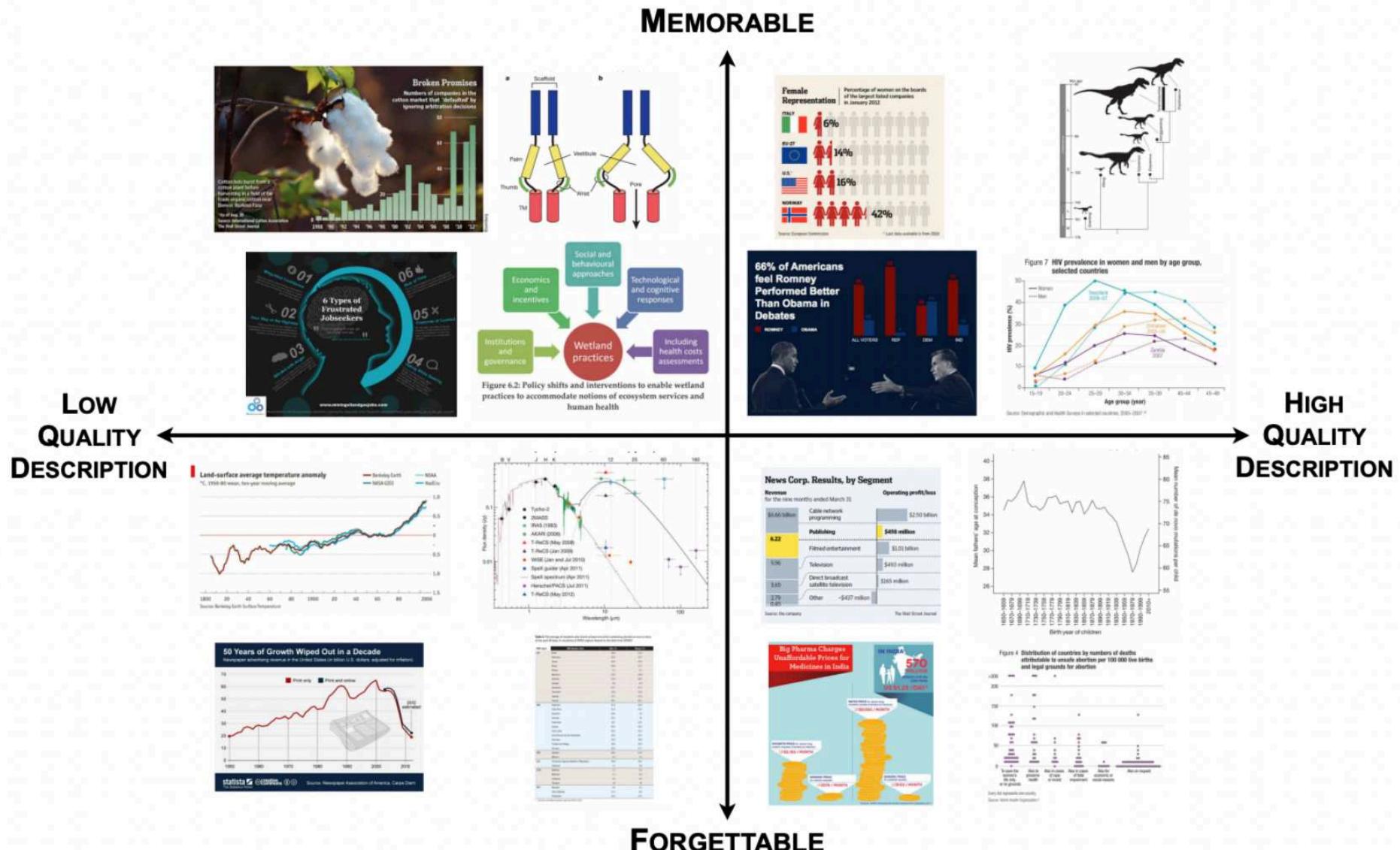


*“It appears that **novel and unexpected visualizations can be better remembered than the visualizations with limited variability that we are exposed to since elementary school.**”*

M. A. Borkin et al. (2013)



MEMORABILITY VERSUS DESCRIPTION QUALITY



Source: Borkin et al. (2013)

Cédric Scherer Data Visualization & Information Design



creating Layouts



Main Principles of Graphic Design

- » **Unity**
- » **Variety**
- » **Hierarchy**

Good balance between **unity** and **variety** creates **hierarchy**.



Table 1: Change in household wealth 2018–19 by region

	Total wealth	Change in total wealth		Wealth per adult	Change in wealth per adult	Change in financial assets		Change in non-financial assets		Change in debts	
	2019	2018-19	2018-19	2019	2018-19	2018-19	2018-19	2018-19	2018-19	2018-19	2018-19
		USD bn	USD bn	%	USD	%	USD bn	%	USD bn	%	USD bn
Africa	4,119	190	3.9	6,488	0.4	1	0.1	164	6.6	35	7.7
Asia-Pacific	64,778	825	1.3	54,211	-0.3	599	1.5	672	1.9	386	4.2
China	63,827	1,889	3.1	58,544	2.6	88	0.2	2,273	7.5	471	10.9
Europe	90,752	1,093	1.2	153,973	1.2	127	0.8	1,156	2.0	190	1.4
India	12,614	625	5.2	14,569	3.3	37	1.4	708	6.9	120	11.5
Latin America	9,906	469	4.9	22,502	3.2	198	4.0	340	5.7	70	5.0
North America	114,607	4,061	3.7	417,694	2.7	9,394	3.6	1,355	3.8	626	3.8
World	360,608	9,087	2.6	70,849	1.2	4,819	2.0	6,666	3.7	1,696	4.0

Source: James Davies, Rodrigo Lluberas and Anthony Shorrocks, Global wealth databook 2019



Table 1: Change in household wealth 2018–2019 by region

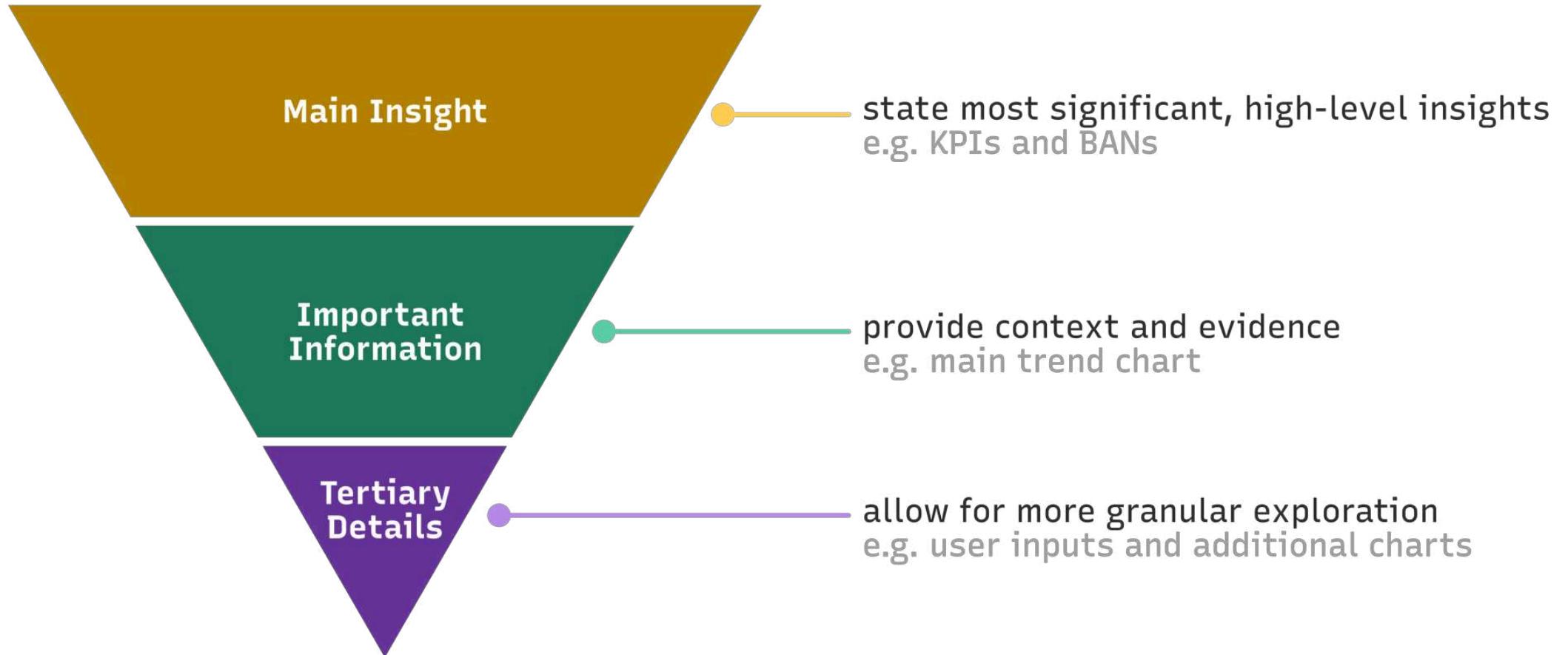
Overall total wealth and mean wealth per adult refer to 2019 only. Triangles pointing upwards indicate the highest proportional change per category triangles pointing downwards the lowest change.

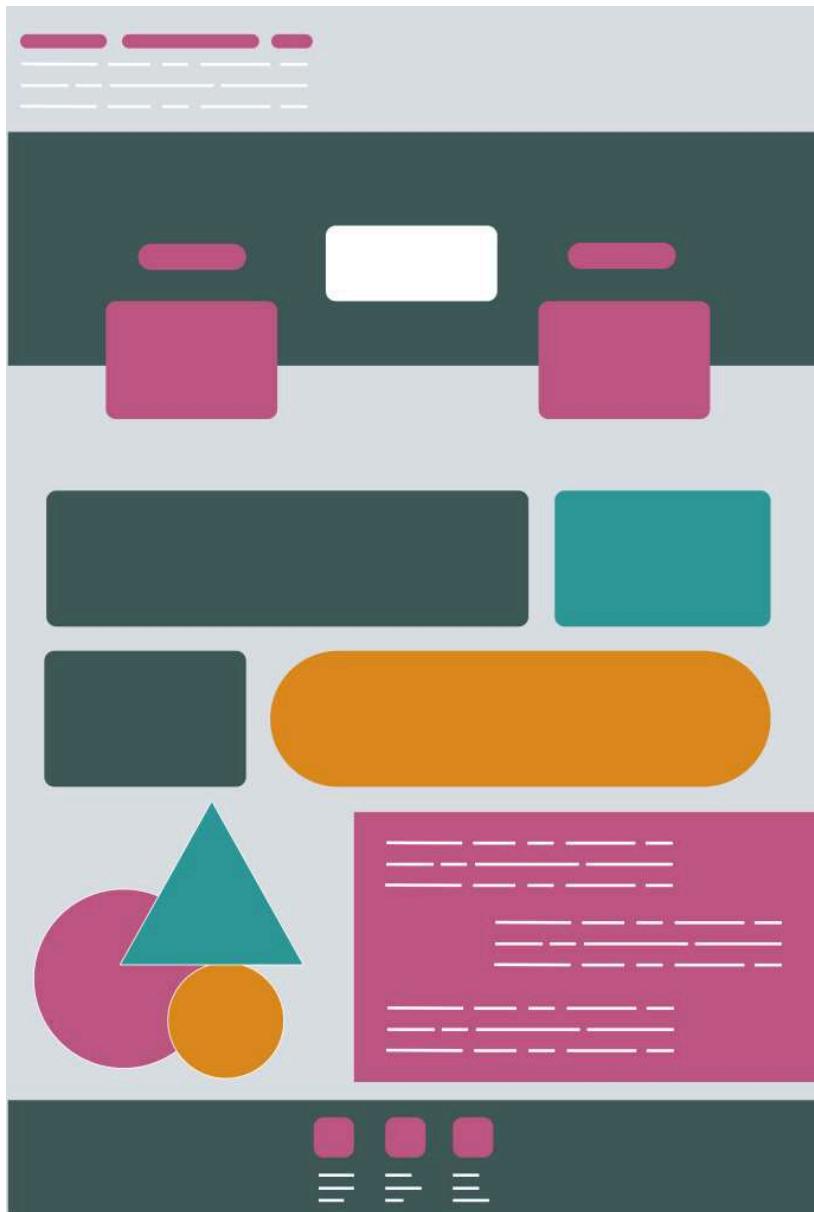
	TOTAL WEALTH			WEALTH PER ADULT		FINANCIAL ASSETS		NON-FINANCIAL ASSETS		DEBTS	
	OVERALL (USD bn)	CHANGE^ (USD bn)	(prop.)	MEAN (USD)	CHANGE (prop.)	(USD bn)	CHANGE (prop.)	(USD bn)	(prop.)	(USD bn)	CHANGE (prop.)
World	360,603	9,087	2.6%	70,849	1.2%	4,319	2.0%	6,666	3.7%	1,898	4.0%
North America	114,607	4,061	3.7%	417,694	2.7%	3,334	3.6%	1,353	3.8%	626	3.8%
Europe	90,752	1,093	1.2% ▽	153,973	1.2%	127	3.0%	1,156	2.0%	190	1.4% ▽
Asia-Pacific	64,778	825	1.3%	54,211	-0.3% ▽	539	1.5%	672	1.9% ▽	386	4.2%
China	63,827	1,889	3.1%	58,544	2.6%	88	2.0%	2,273	7.5% △	471	10.9%
India	12,614	615	5.2% △	14,569	3.3% △	37	1.4%	708	6.9%	120	11.5% △
Latin America	9,906	463	4.9%	22,502	3.2%	193	4.0% △	340	5.7%	70	5.0%
Africa	4,119	130	3.3%	6,488	0.4%	1	1.0% ▽	164	6.6%	35	7.7%

Source: James Davies, Rodrigo Lluberas and Anthony F. Shorrocks, Global Wealth Databook, 2019 | Makeover: Cedric Scherer, Frontpage Data

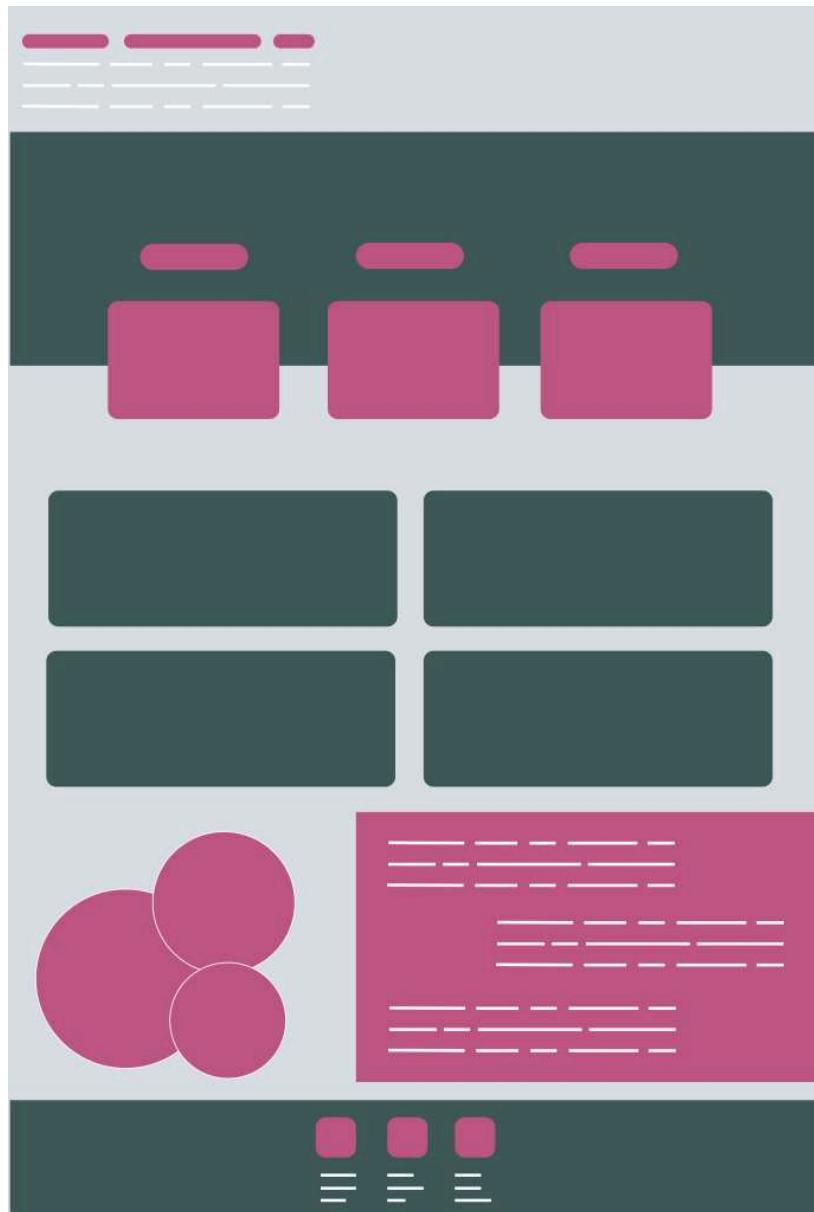


Hierarchy





Created by Lindsay Betzendahl

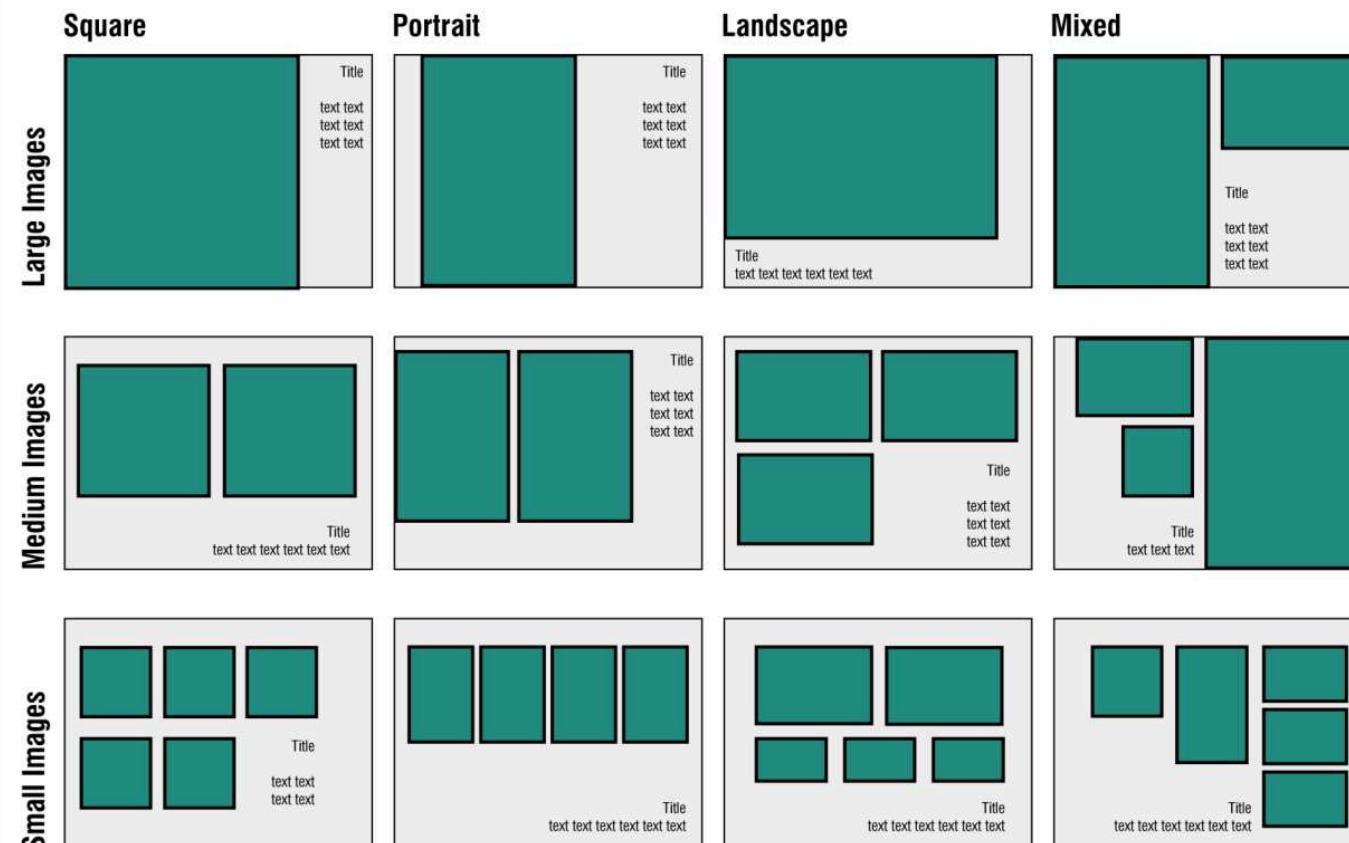


Created by Lindsay Betzendahl



Modular Design

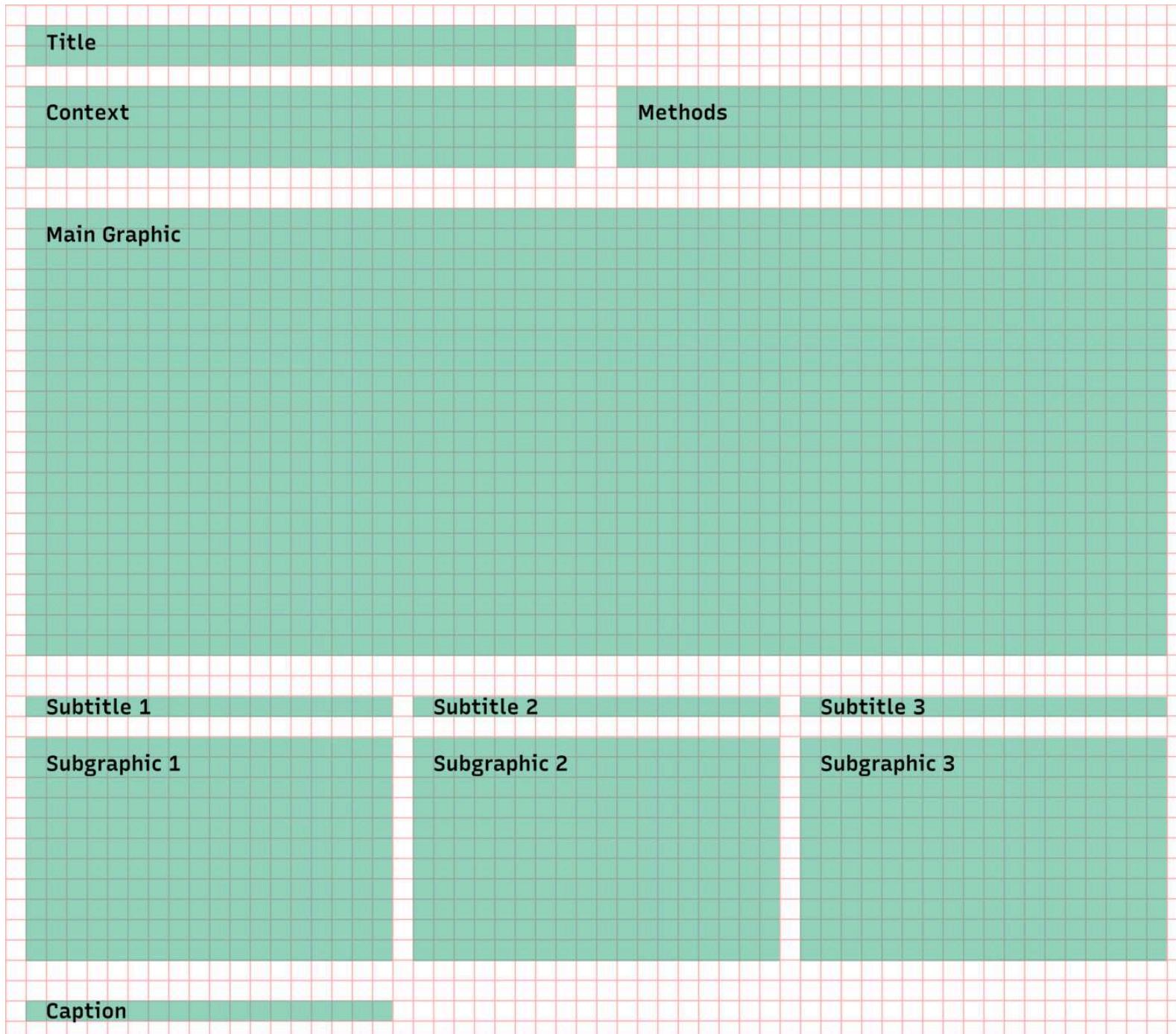
Table showing samples of how large and small your images could look on a single page in the book

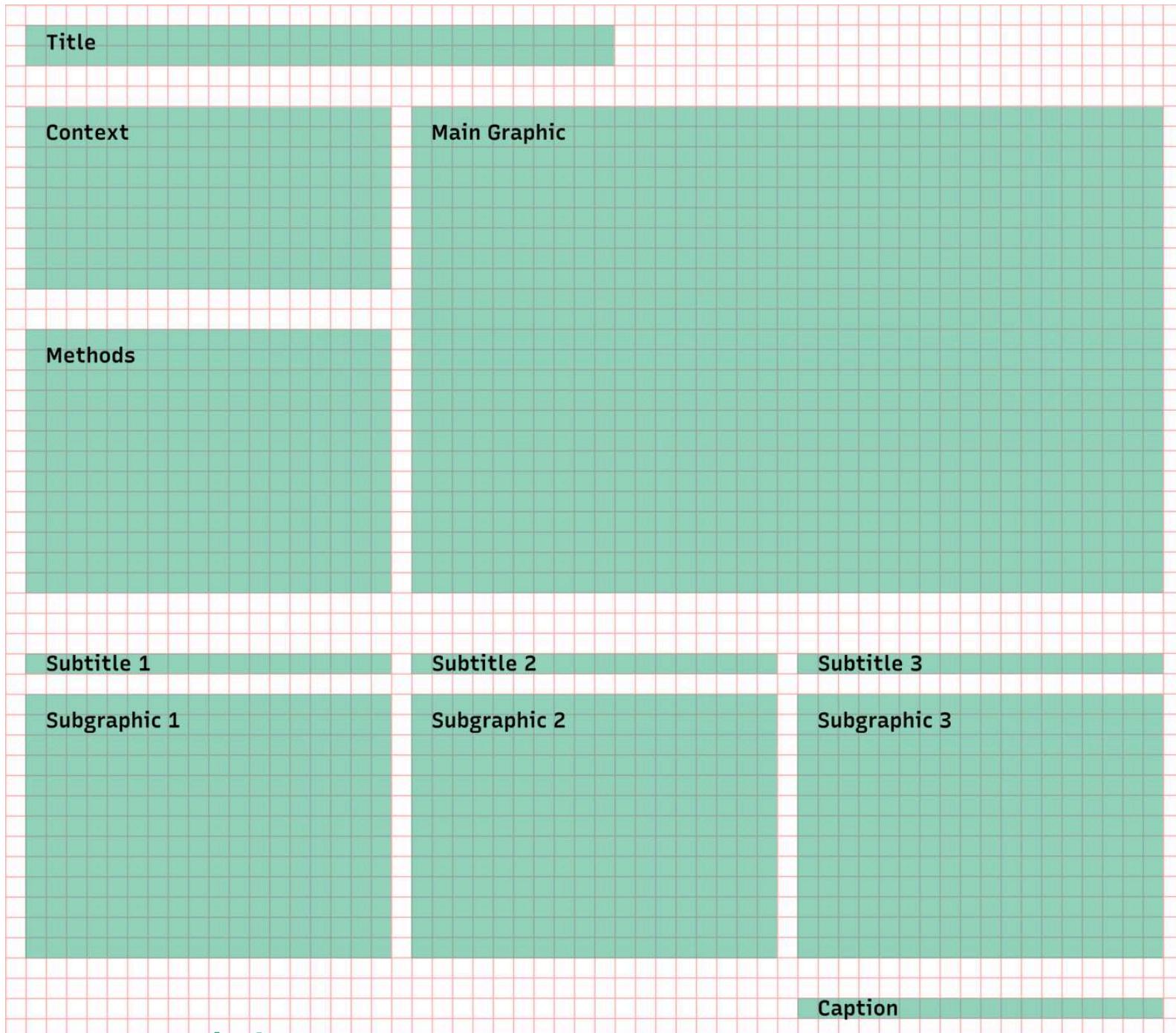


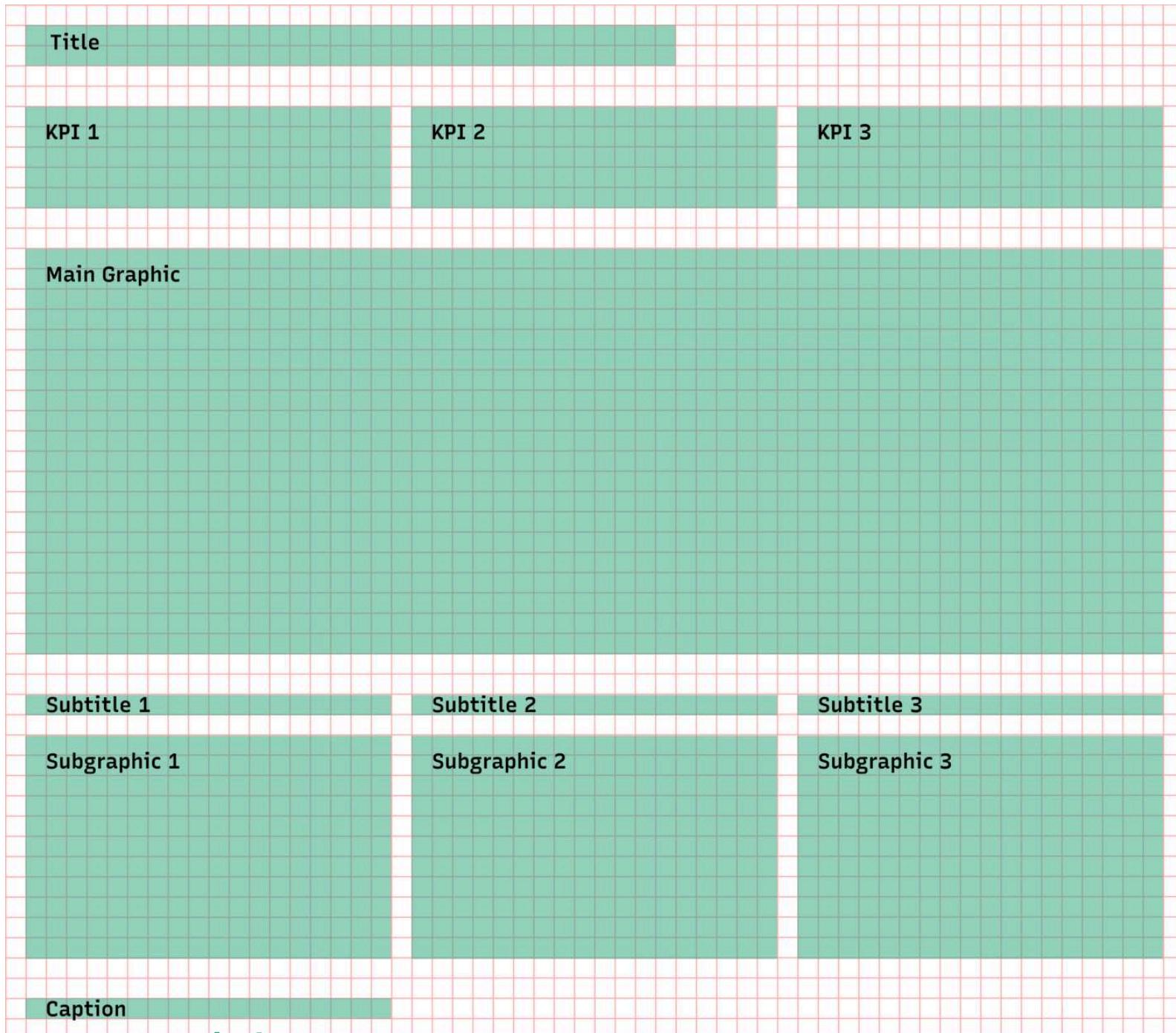
Source: londonphotography.org.uk

Cédric Scherer Data Visualization & Information Design







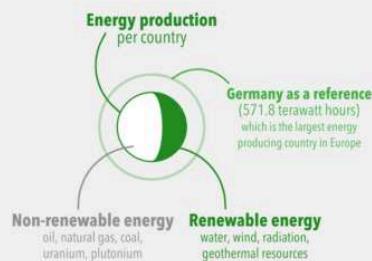


How European countries generated electricity in 2018

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.

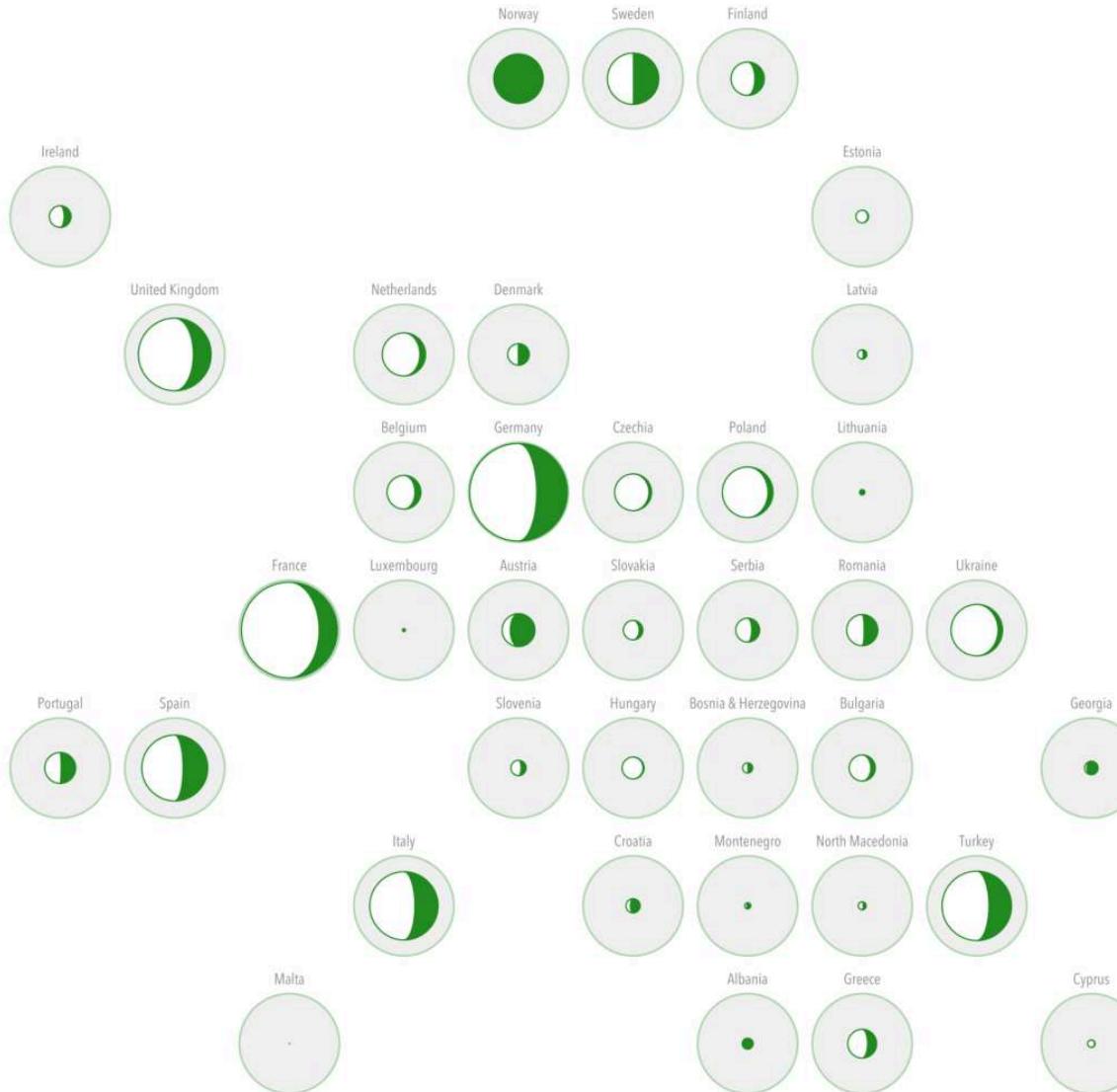


Renewable energy is energy that comes from resources that are naturally replenished such as sunlight, wind, water, and geothermal heat. Unlike fossil fuels, such as oil, natural gas and coal, or nuclear power sources such as uranium and plutonium, renewable energy regenerates naturally in a short period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

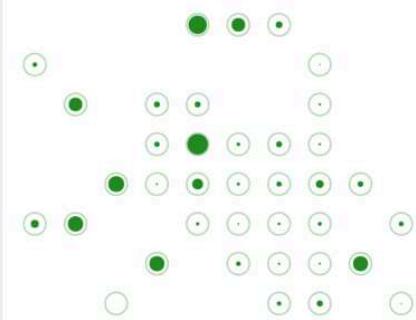


Norway had an electricity production almost entirely made up of renewable energy (98%). This makes Norway the second largest producer of this energy type in Europe. Interestingly, most of the renewable energy is produced by hydro power that take up 95% and only 3% by wind. In contrast, twelve European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

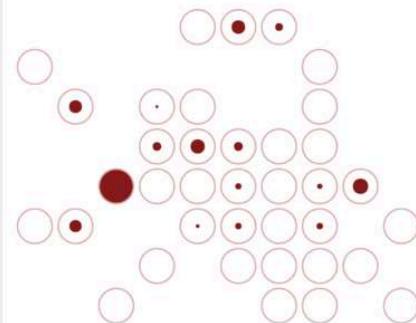
Note: Energy production is mapped to the area of the circles.
Visualization by Cédric Scherer • Data by Eurostat



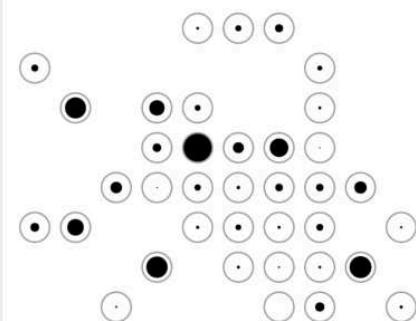
Renewable energy



Nuclear energy



Conventional thermal energy

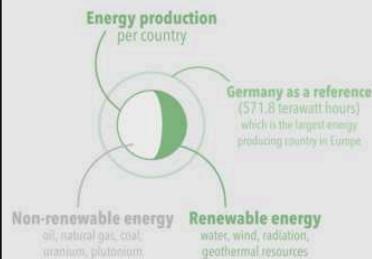


How European countries generated electricity in 2018

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.

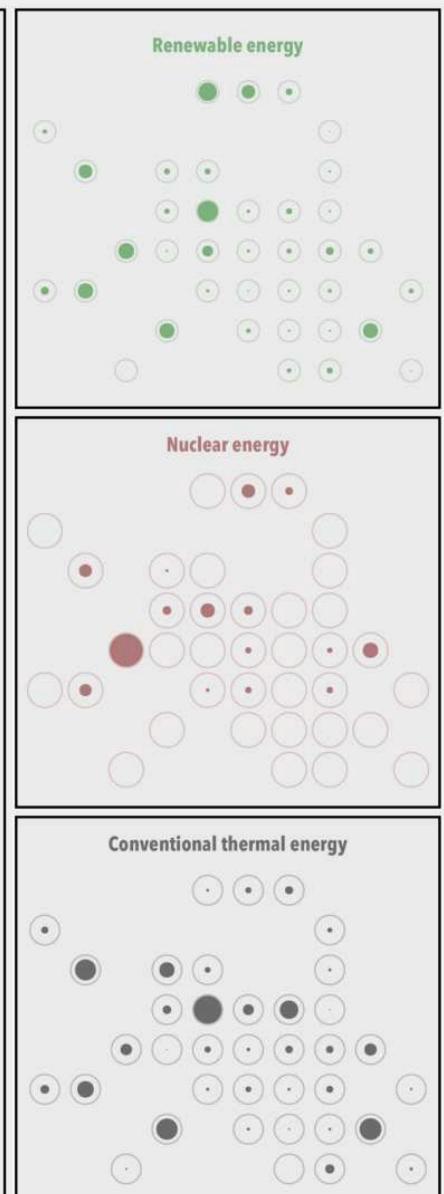
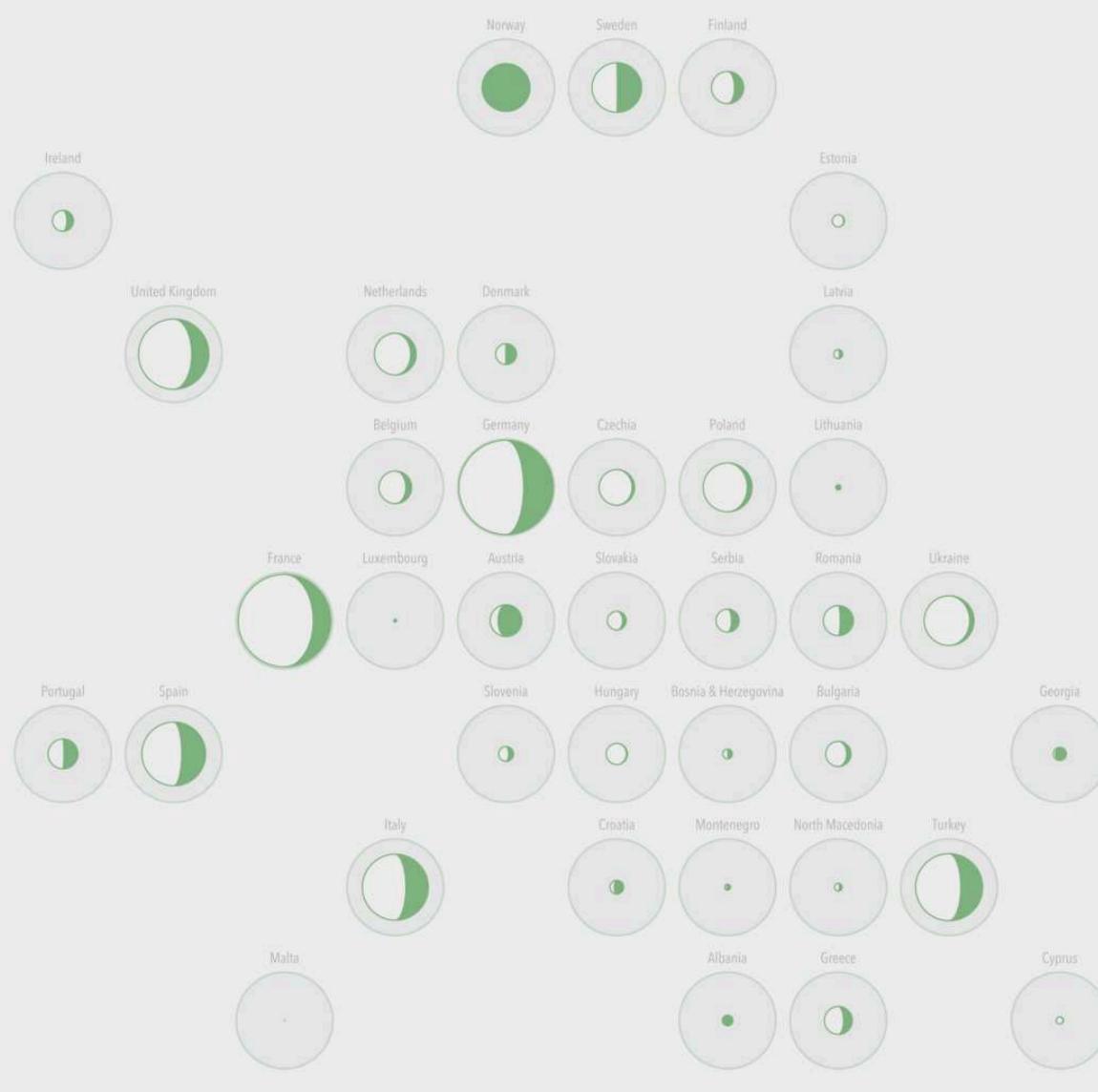


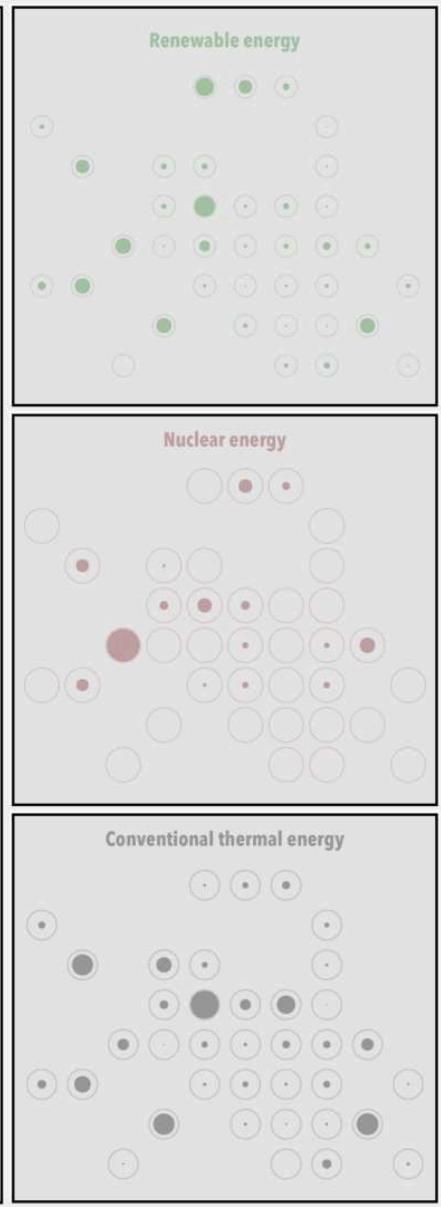
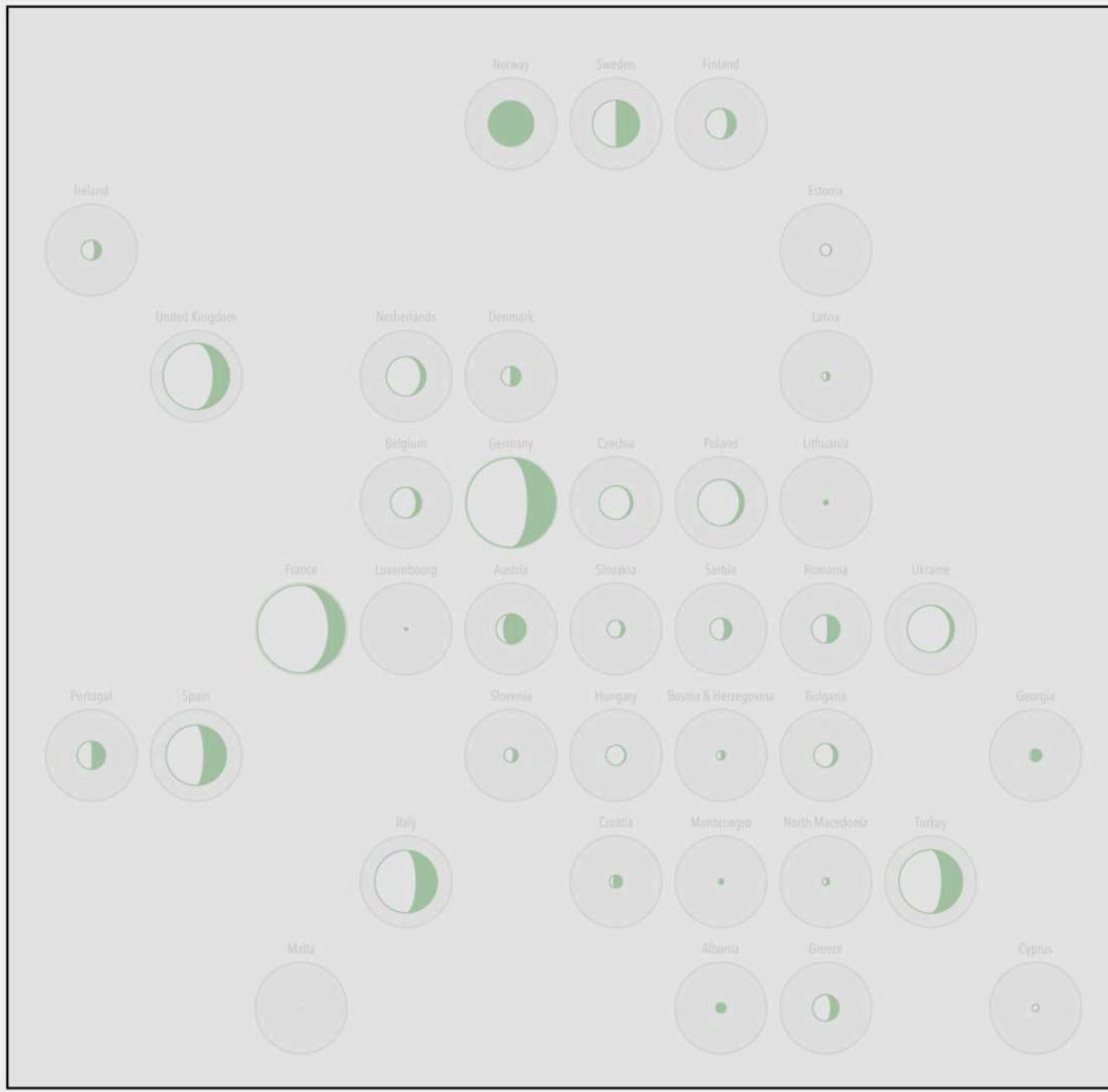
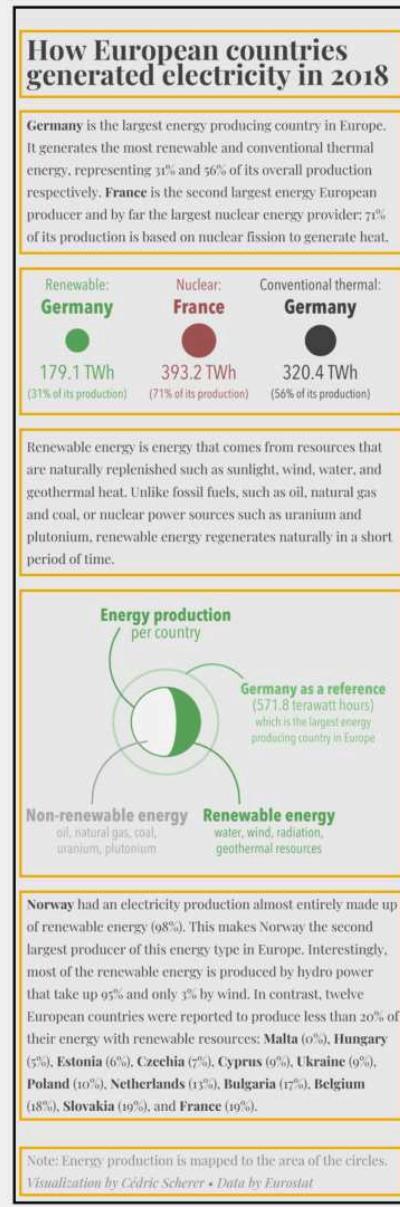
Renewable energy is energy that comes from resources that are naturally replenished such as sunlight, wind, water, and geothermal heat. Unlike fossil fuels, such as oil, natural gas and coal, or nuclear power sources such as uranium and plutonium, renewable energy regenerates naturally in a short period 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 (2%), Estonia (6%), Czechia (7%), Cyprus (9%), Ukraine (9%), Poland (10%), Netherlands (11%), 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

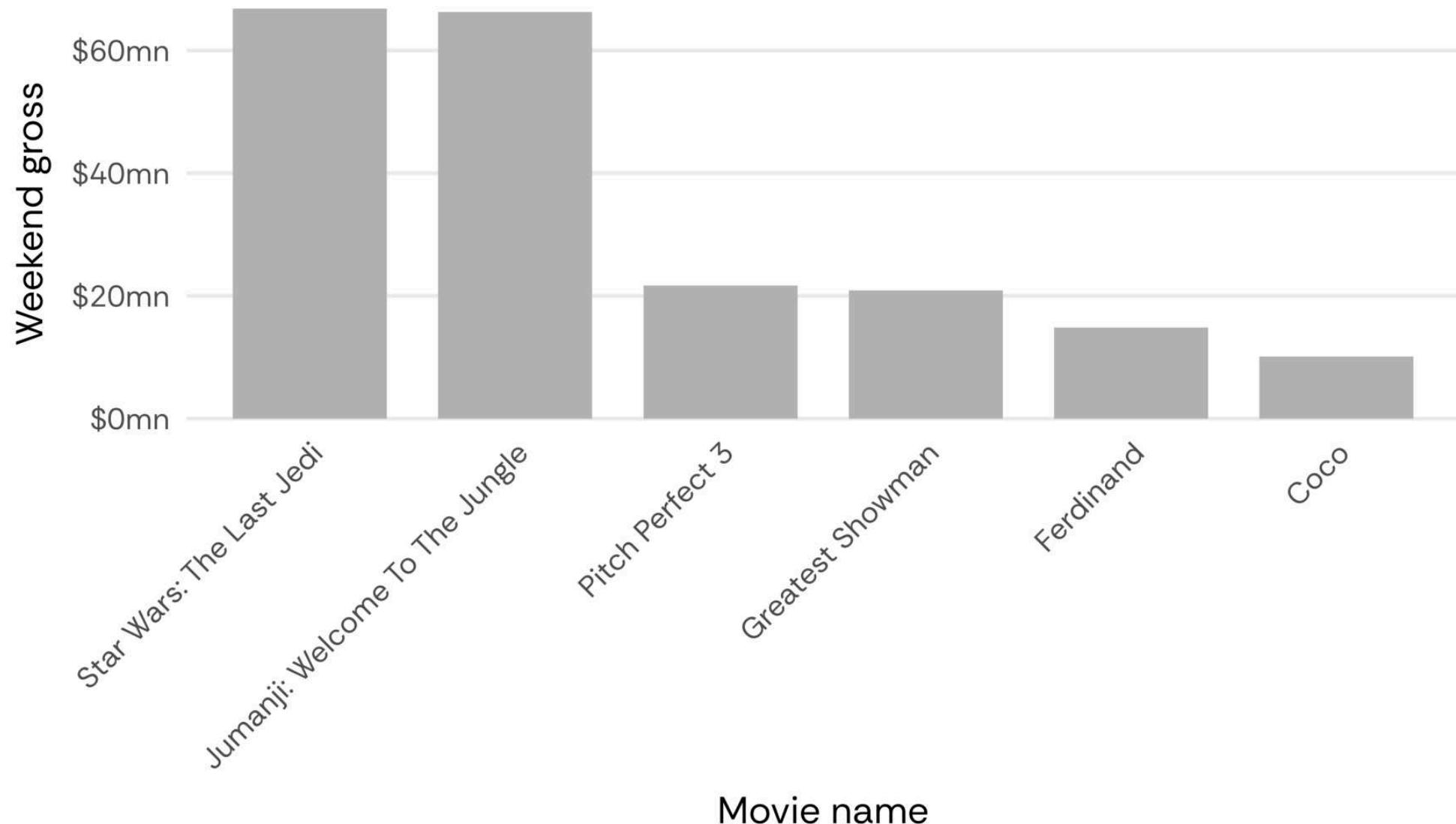




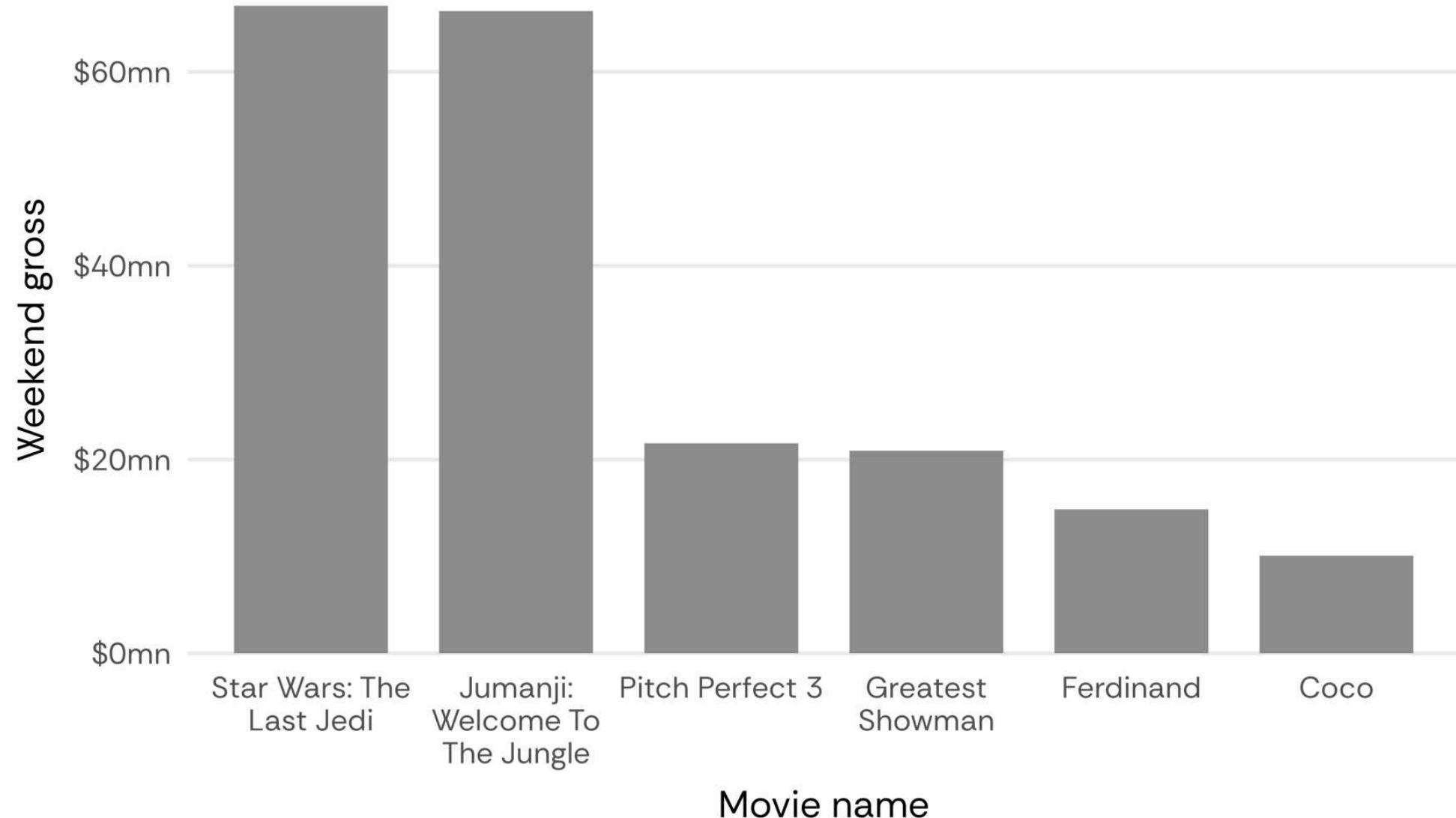
Alignment



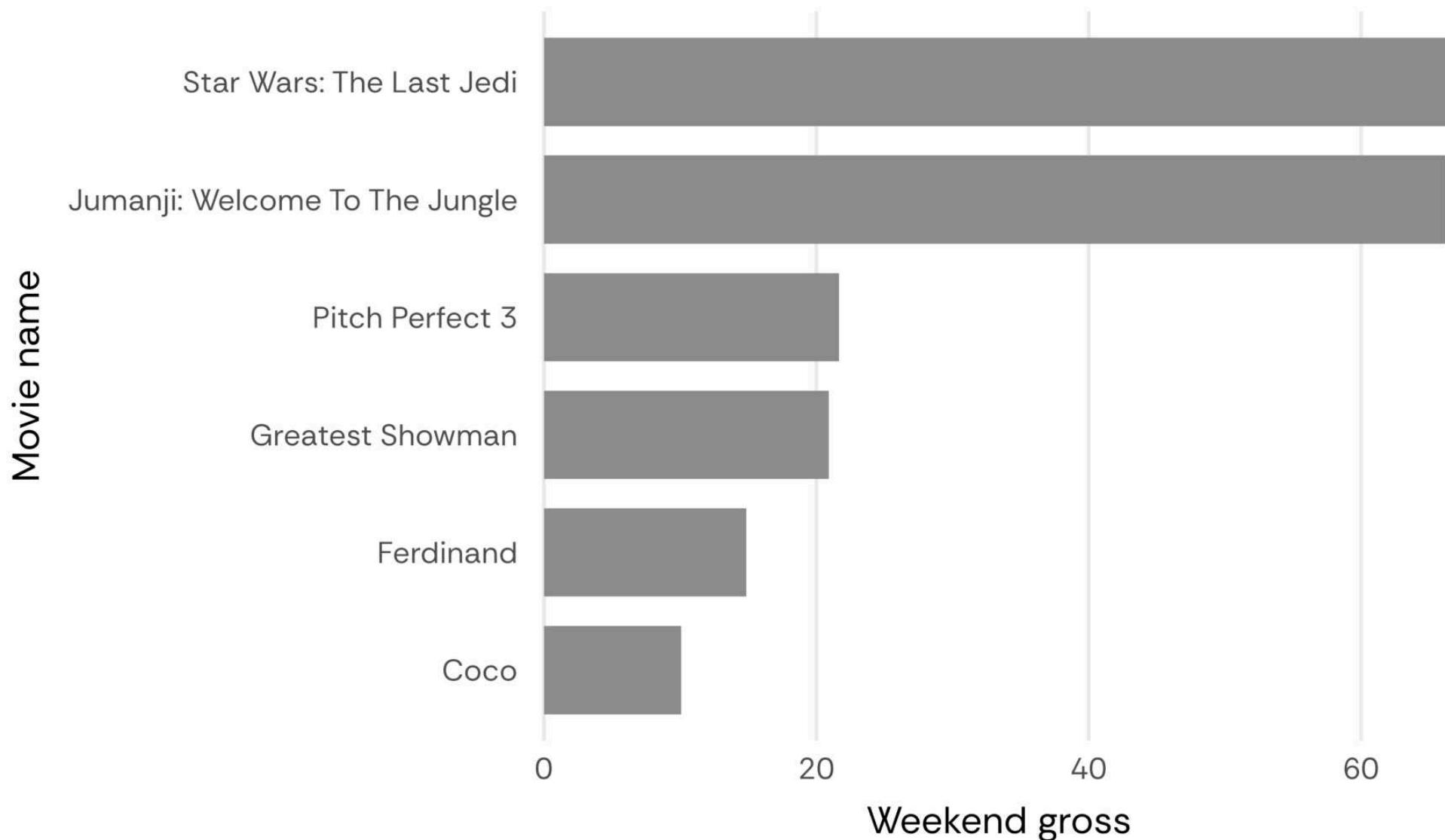
Alignment of Text Elements



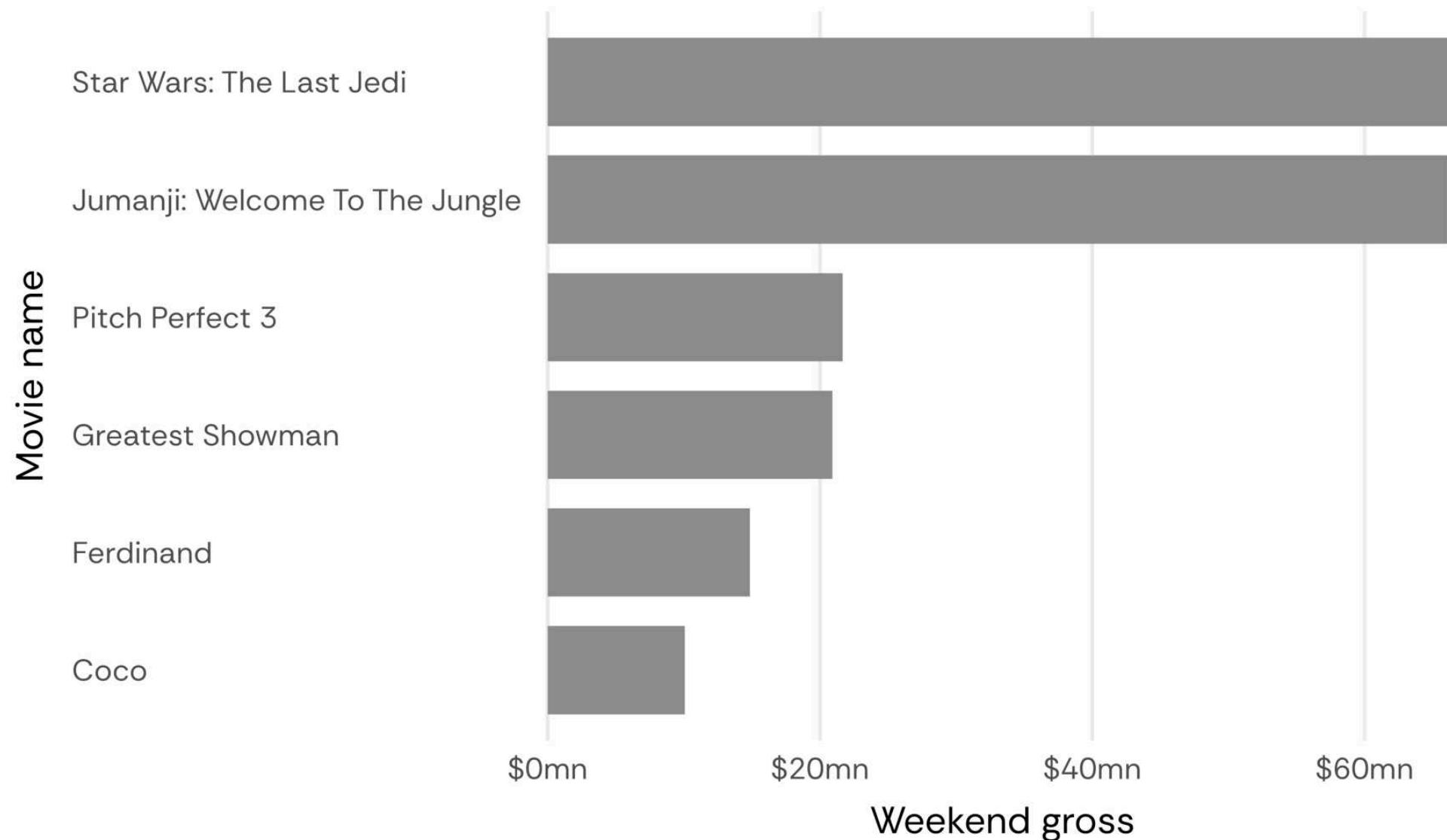
Alignment of Text Elements



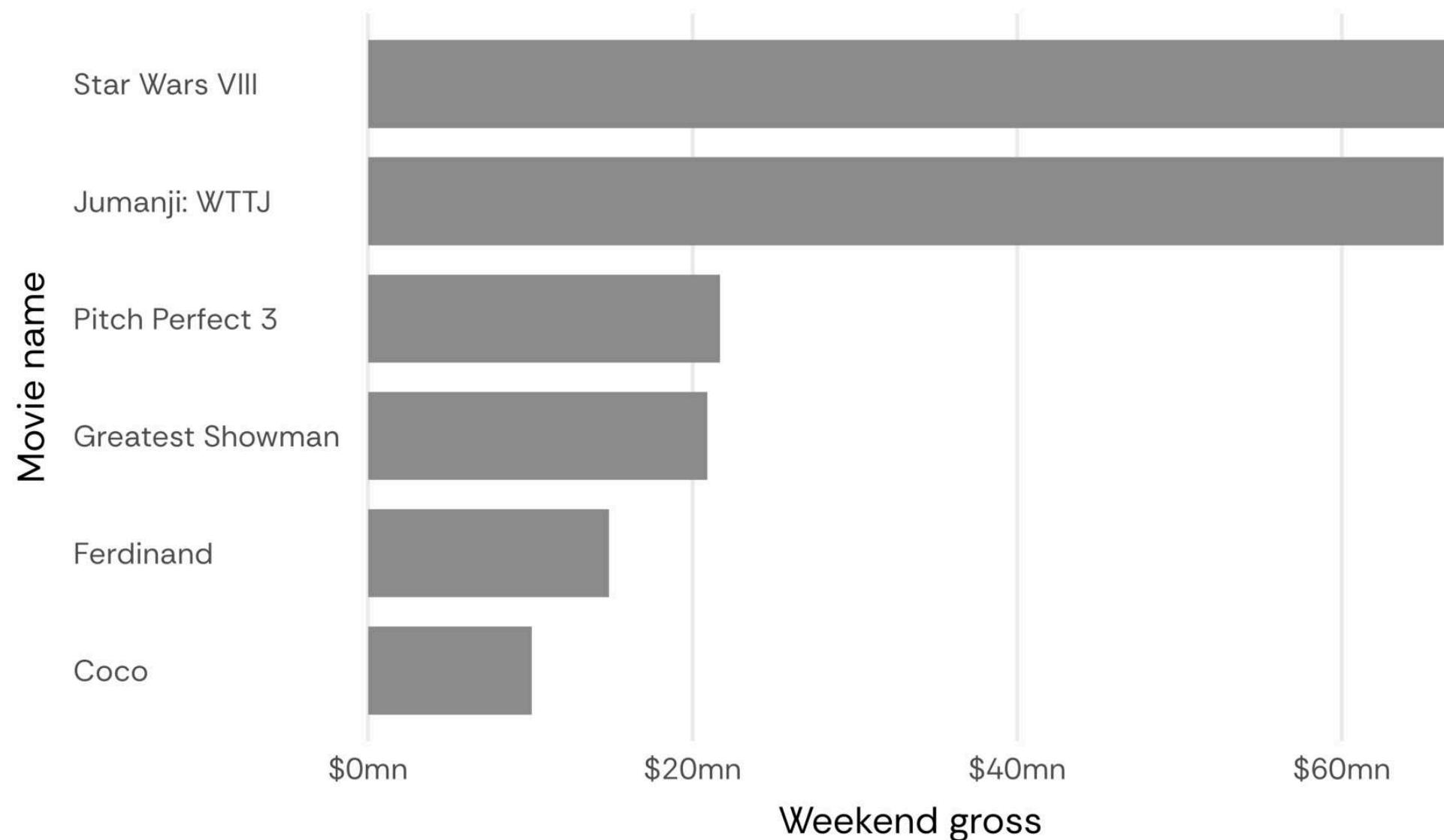
Alignment of Text Elements



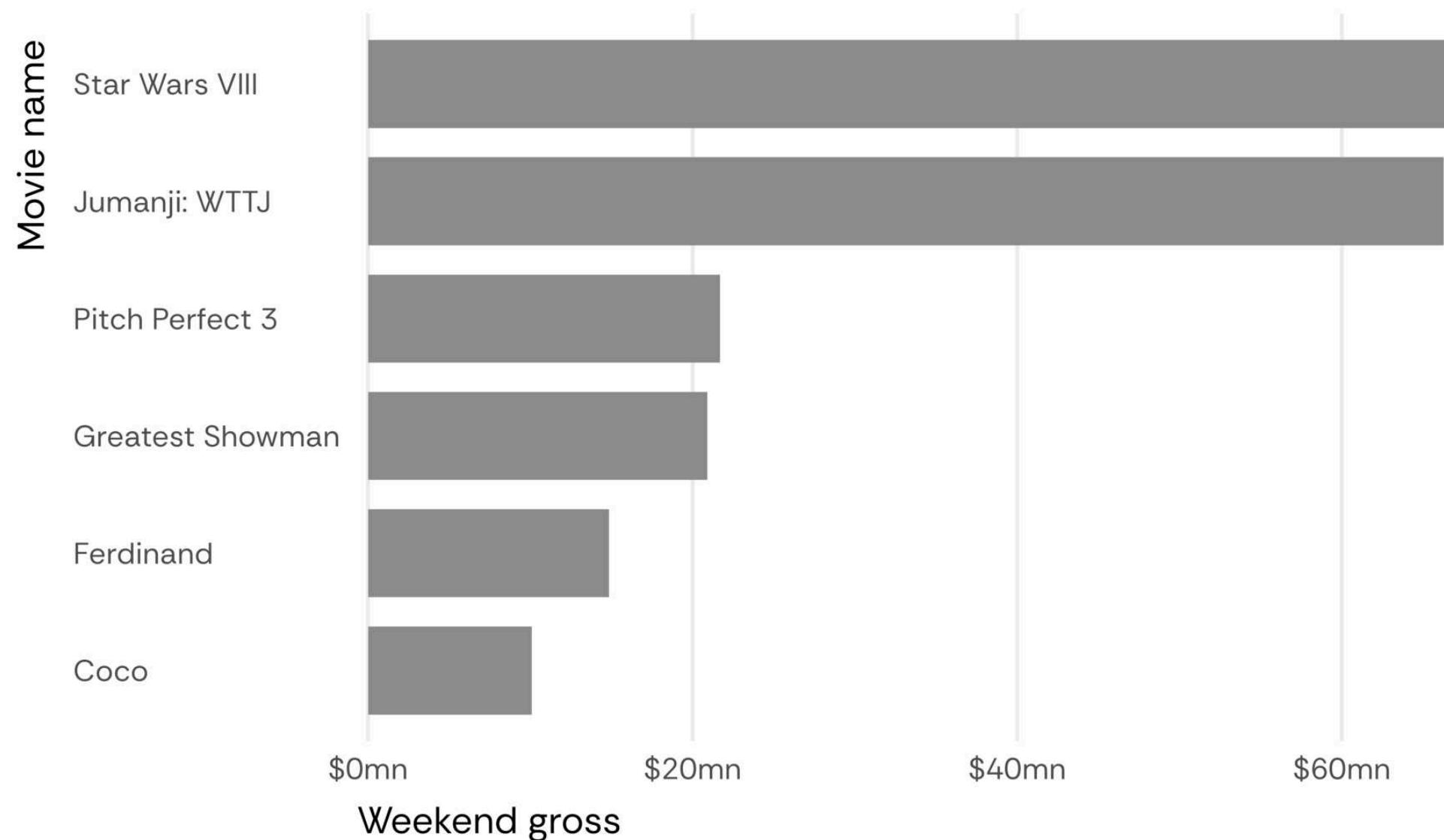
Alignment of Text Elements



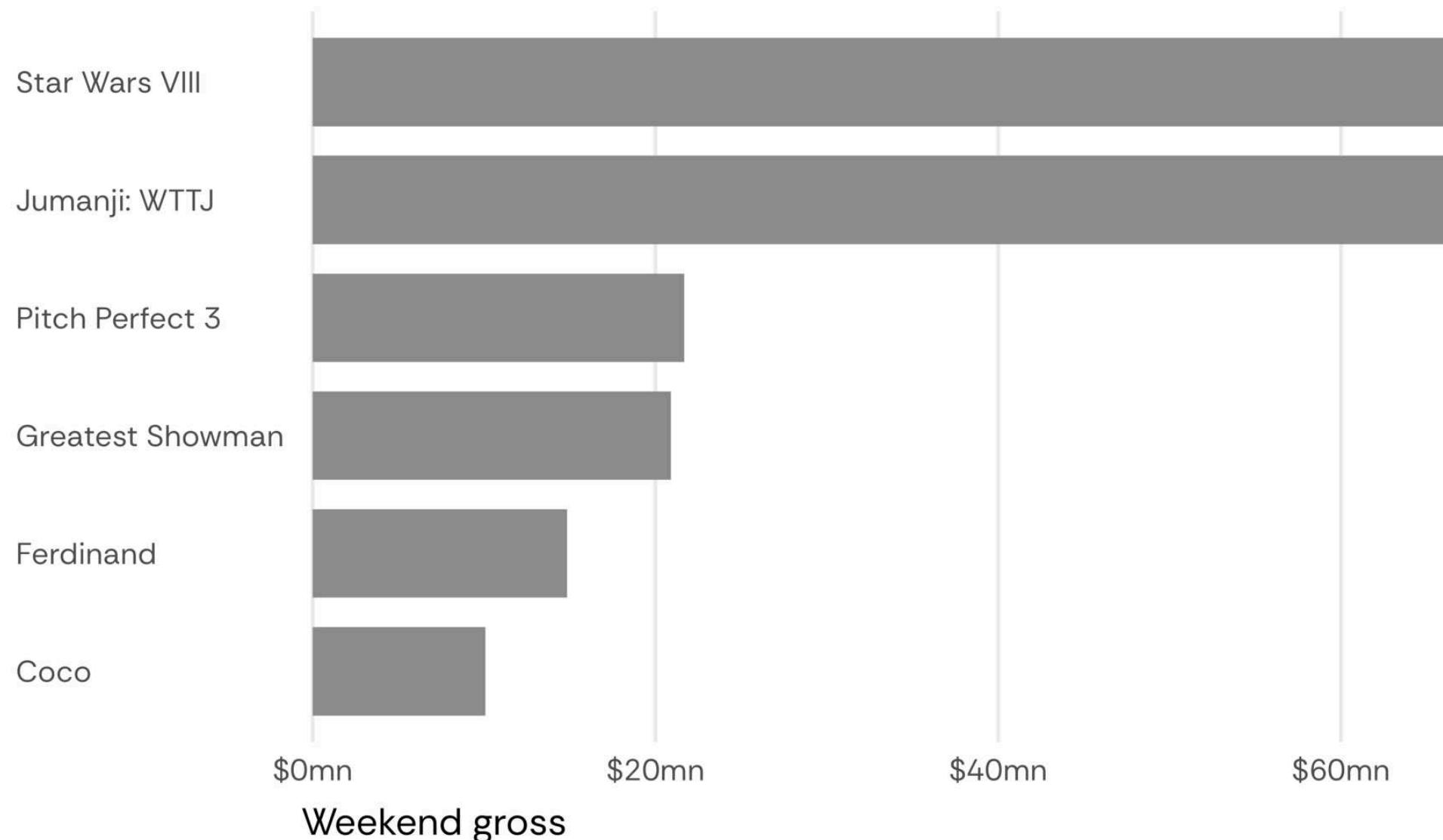
Alignment of Text Elements



Alignment of Text Elements

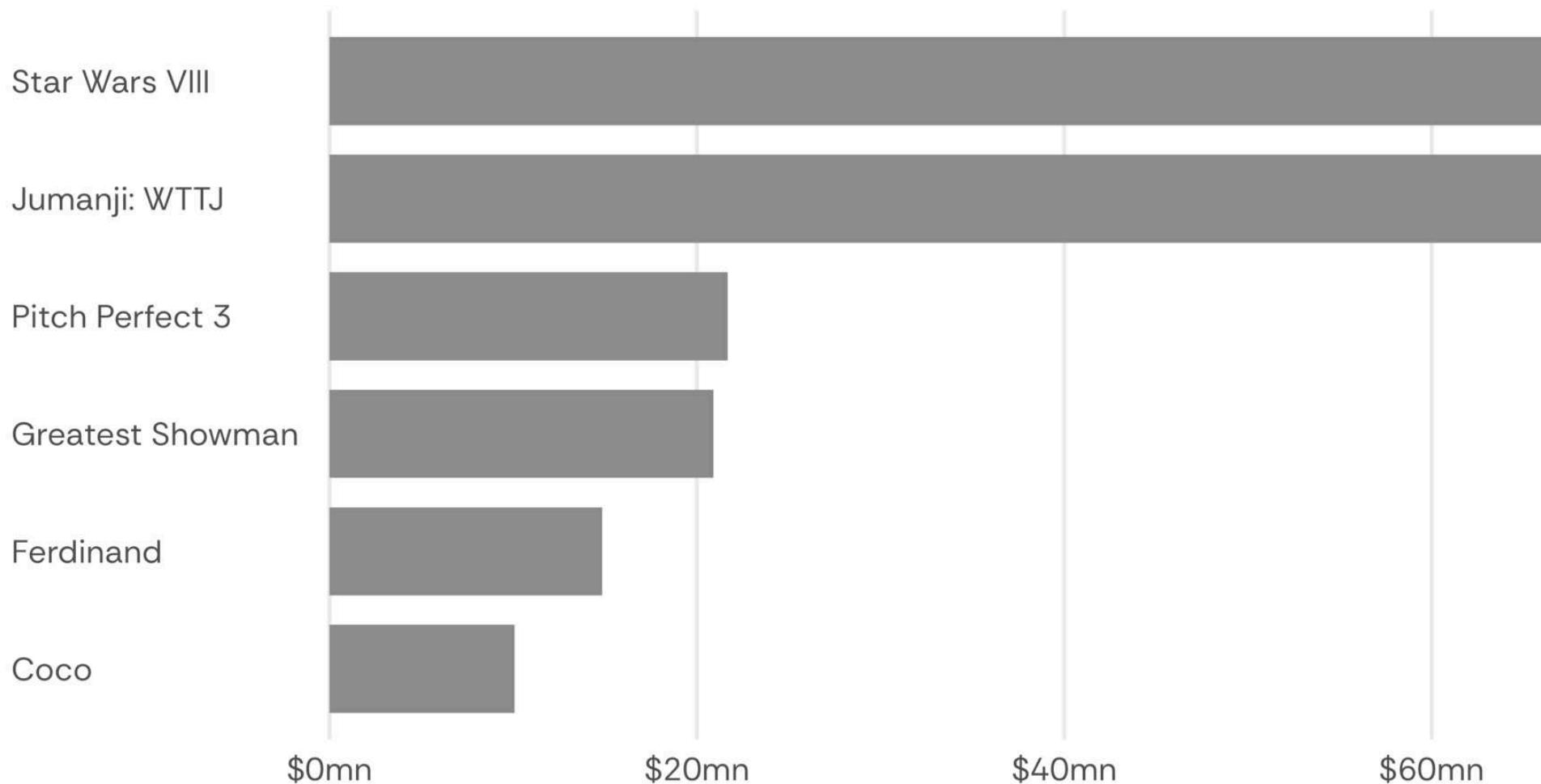


Alignment of Text Elements



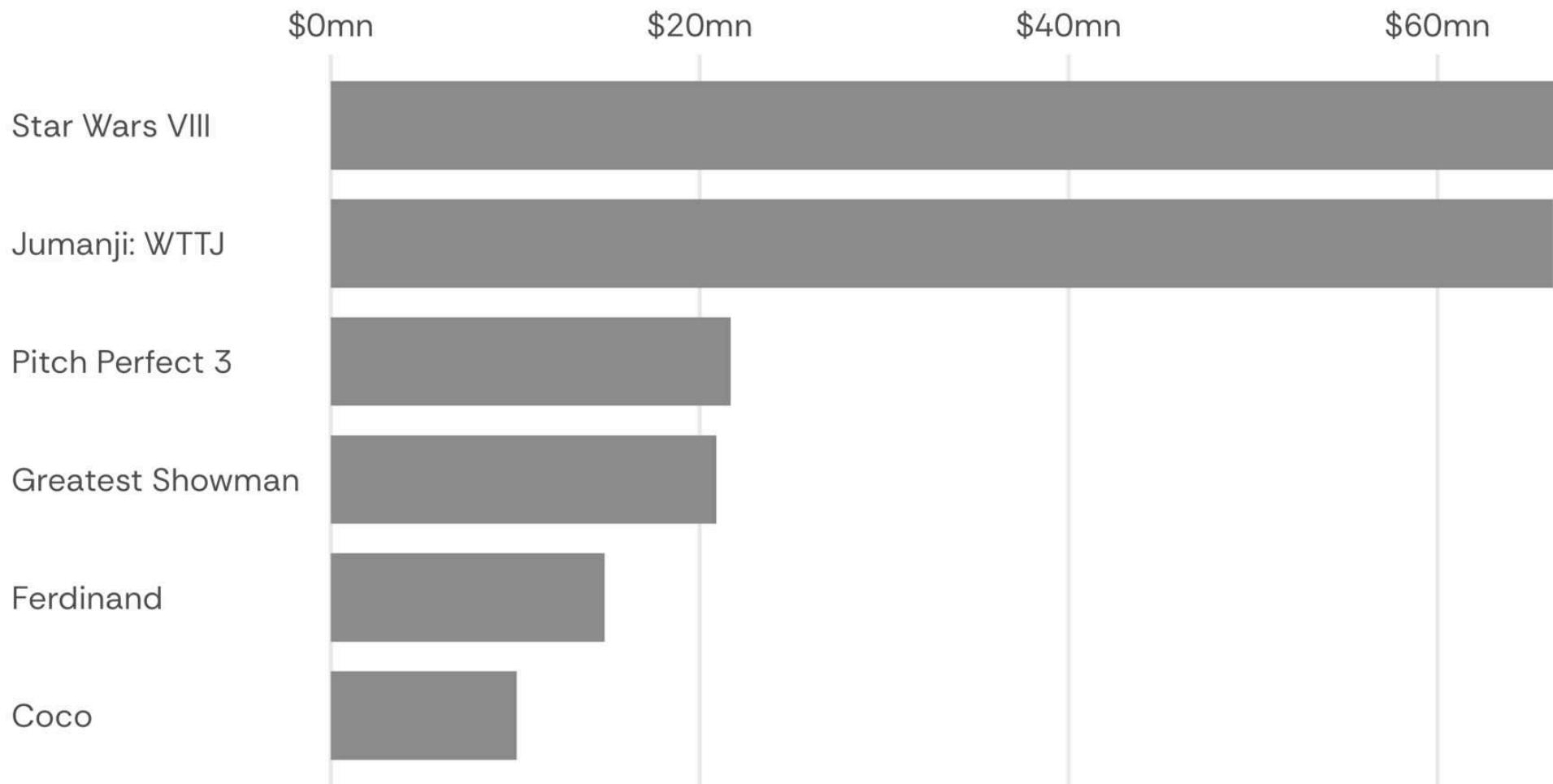
Alignment of Text Elements

"Stars Wars: The Last Jedi" just barely topped the box office for the third weekend in a row, facing stiff competition from "Jumanji: Welcome to the Jungle"



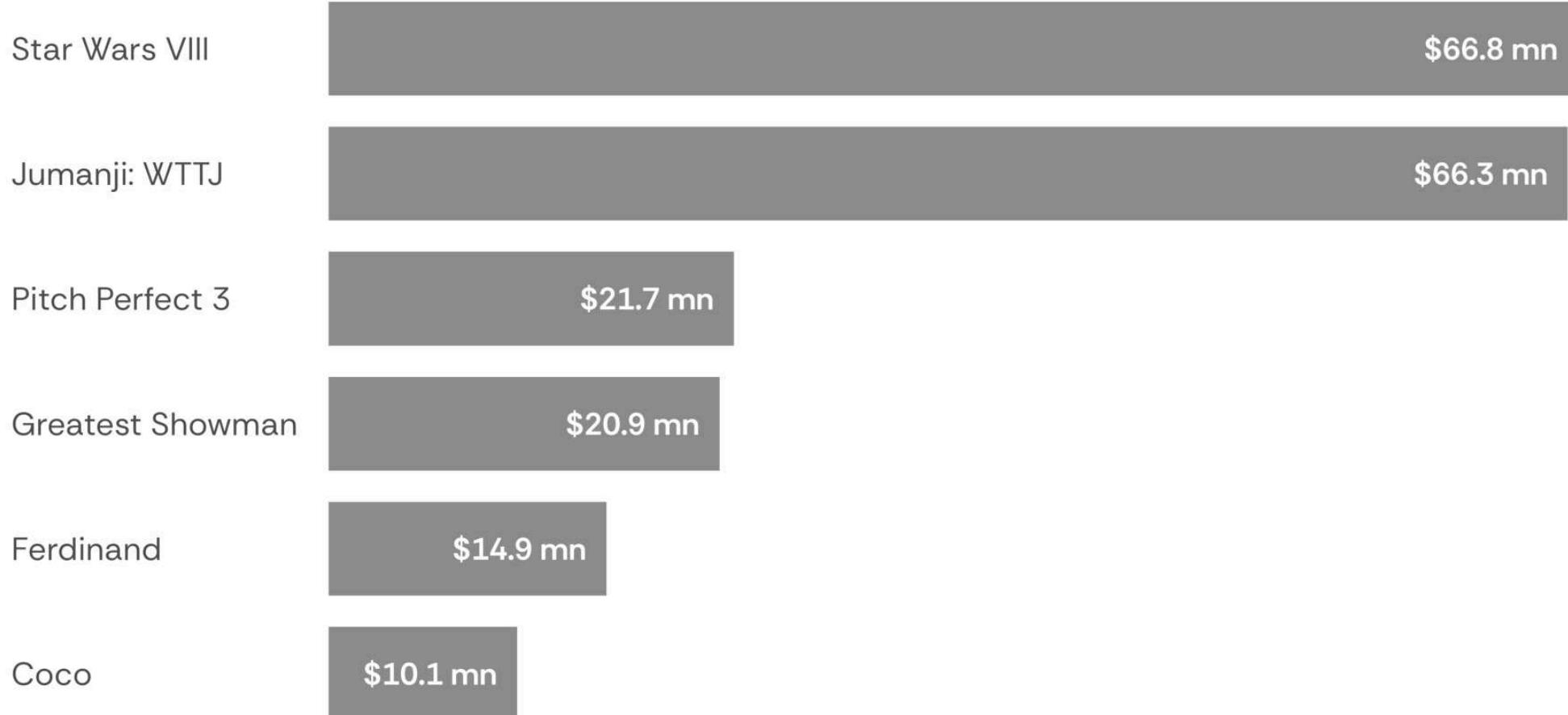
Alignment of Text Elements

"Stars Wars: The Last Jedi" just barely topped the box office for the third weekend in a row, facing stiff competition from "Jumanji: Welcome to the Jungle"



Alignment of Text Elements

"Stars Wars: The Last Jedi" just barely topped the box office for the third weekend in a row, facing stiff competition from "Jumanji: Welcome to the Jungle"



Proportional Numbers

123.45
678.90

Tabular Numbers

123.45
678.90



Proportional Numbers

1	1	1	1	1	1	1
6	5	4	3	2	1	
1	0	0	0			
7	8	9				

Tabular Numbers

1	1	1	1	1	1	1
6	5	4	3	2	1	
1	0	0	0			
7	8	9				





Peak Break-Up Times

According to Facebook status updates



David McCandless & Lee Byron, [Information is Beautiful](#) (modified)

Cédric Scherer Data Visualization & Information Design



Peak Break-Up Times

According to Facebook status updates



David McCandless & Lee Byron, [Information is Beautiful](#)

Cédric Scherer Data Visualization & Information Design



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

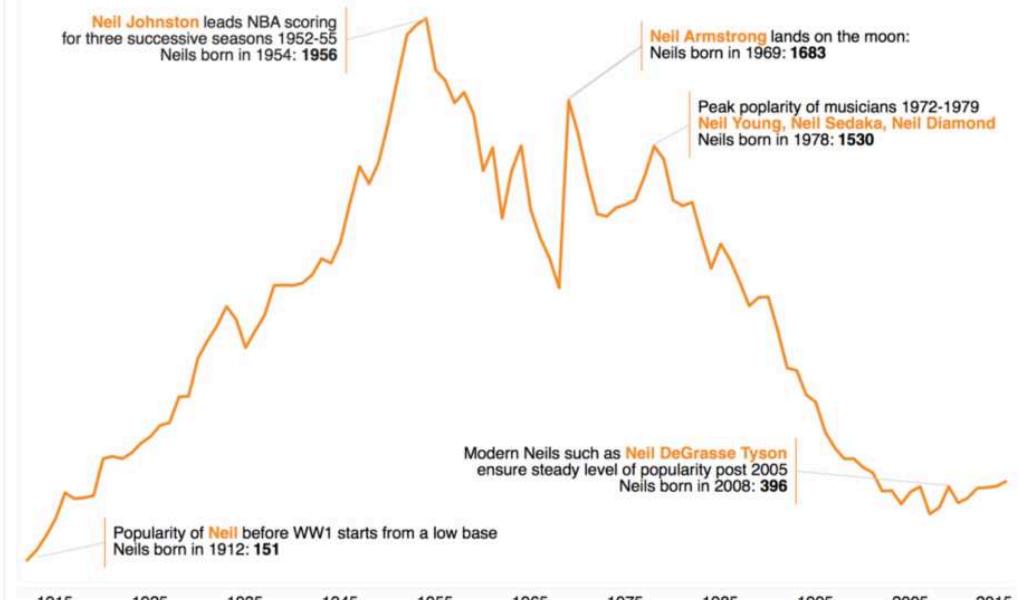
Source: data.gov



Visualisation: @theneilrichards

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

Source: data.gov



Visualisation: @theneilrichards

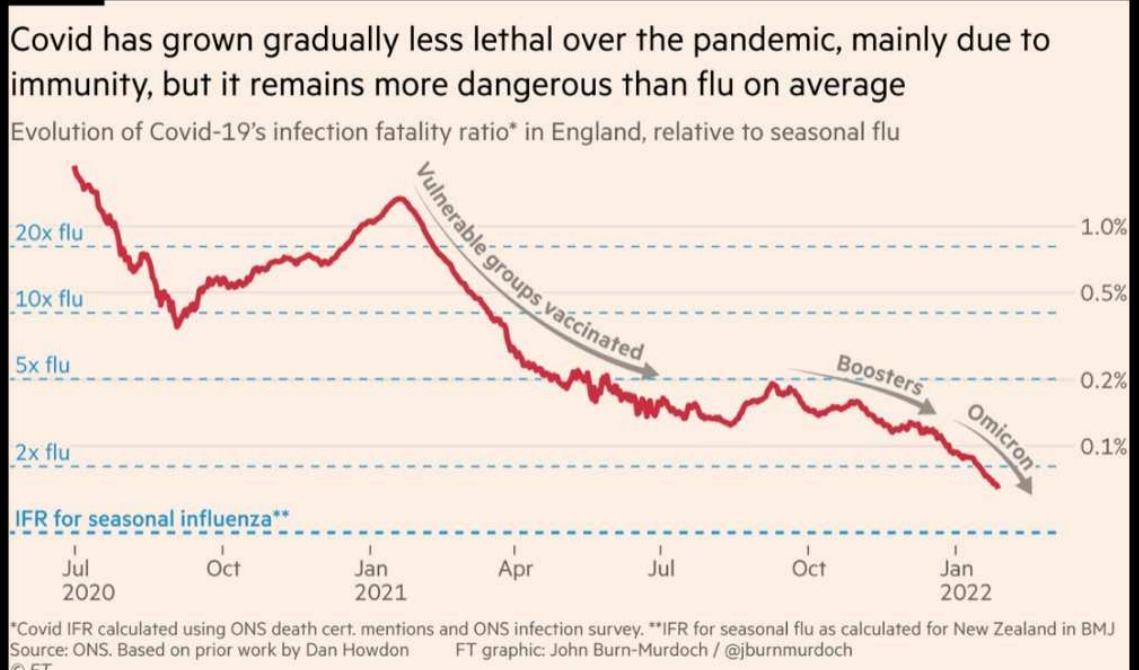
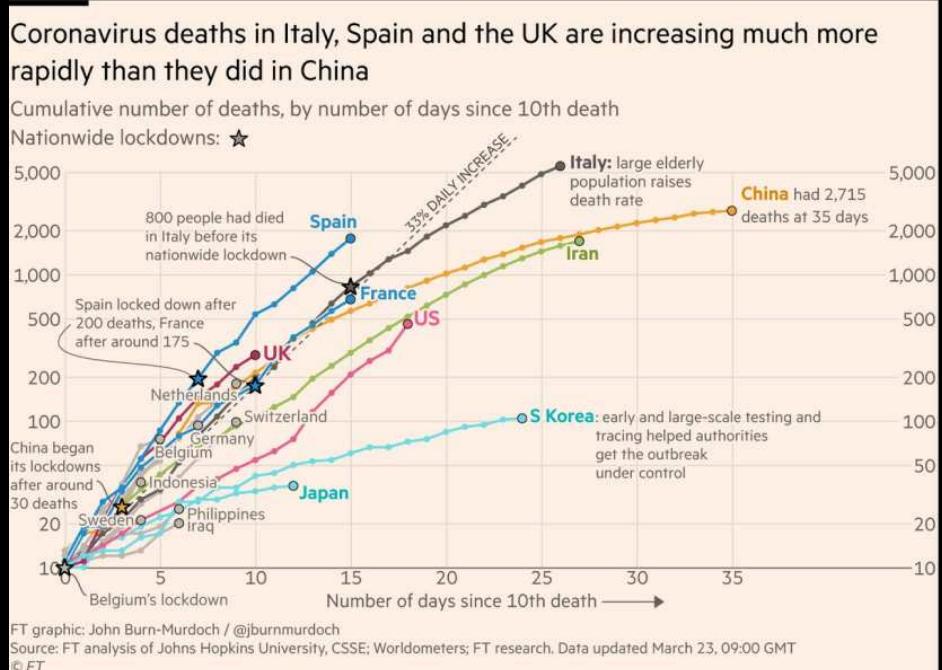
#SWDChallenge

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

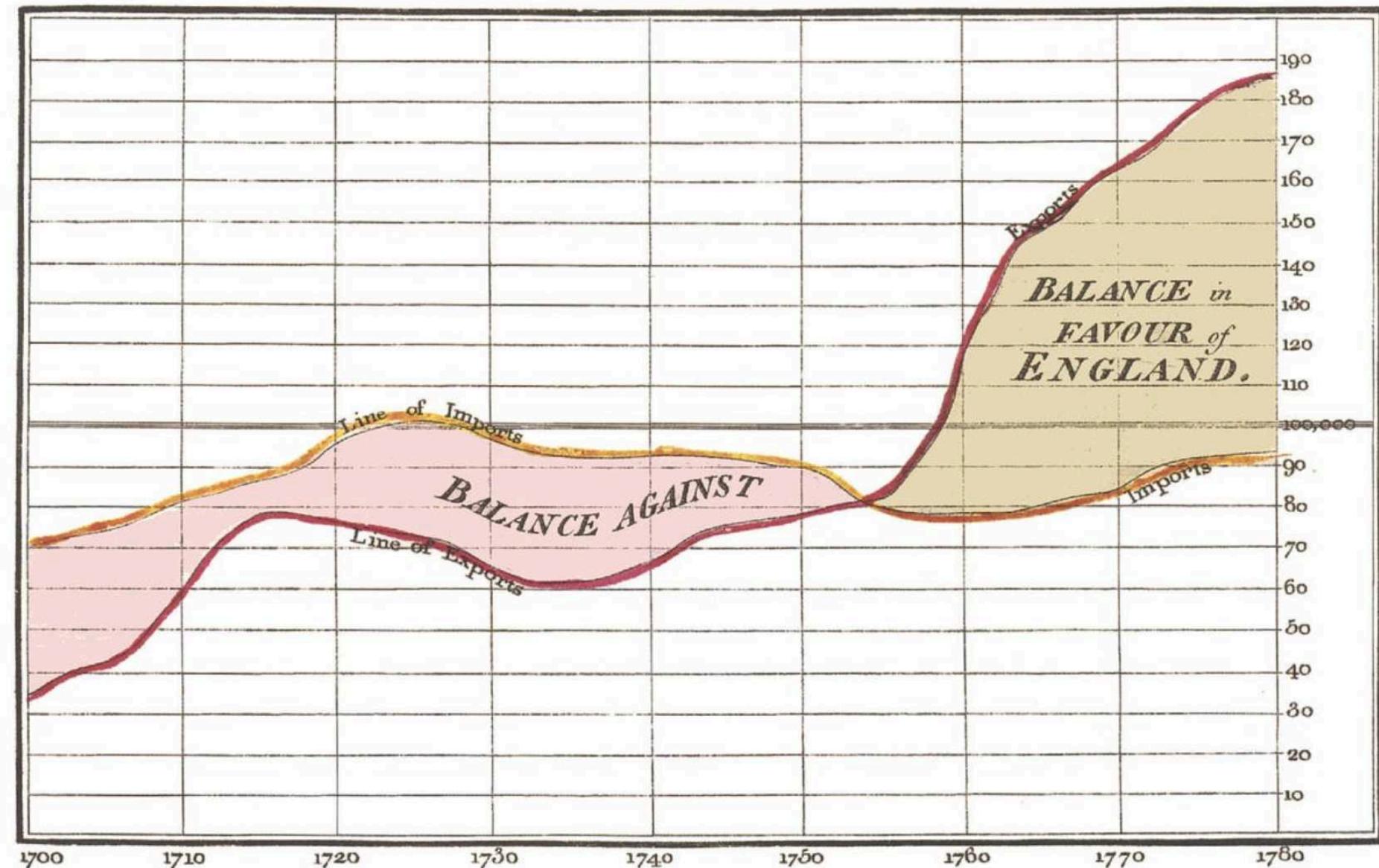


"The key thing we do is to add a title to the chart, as an entry point. [...] Text and other annotations add enormous value for non-chart people."

John Burn-Murdoch, Financial Times



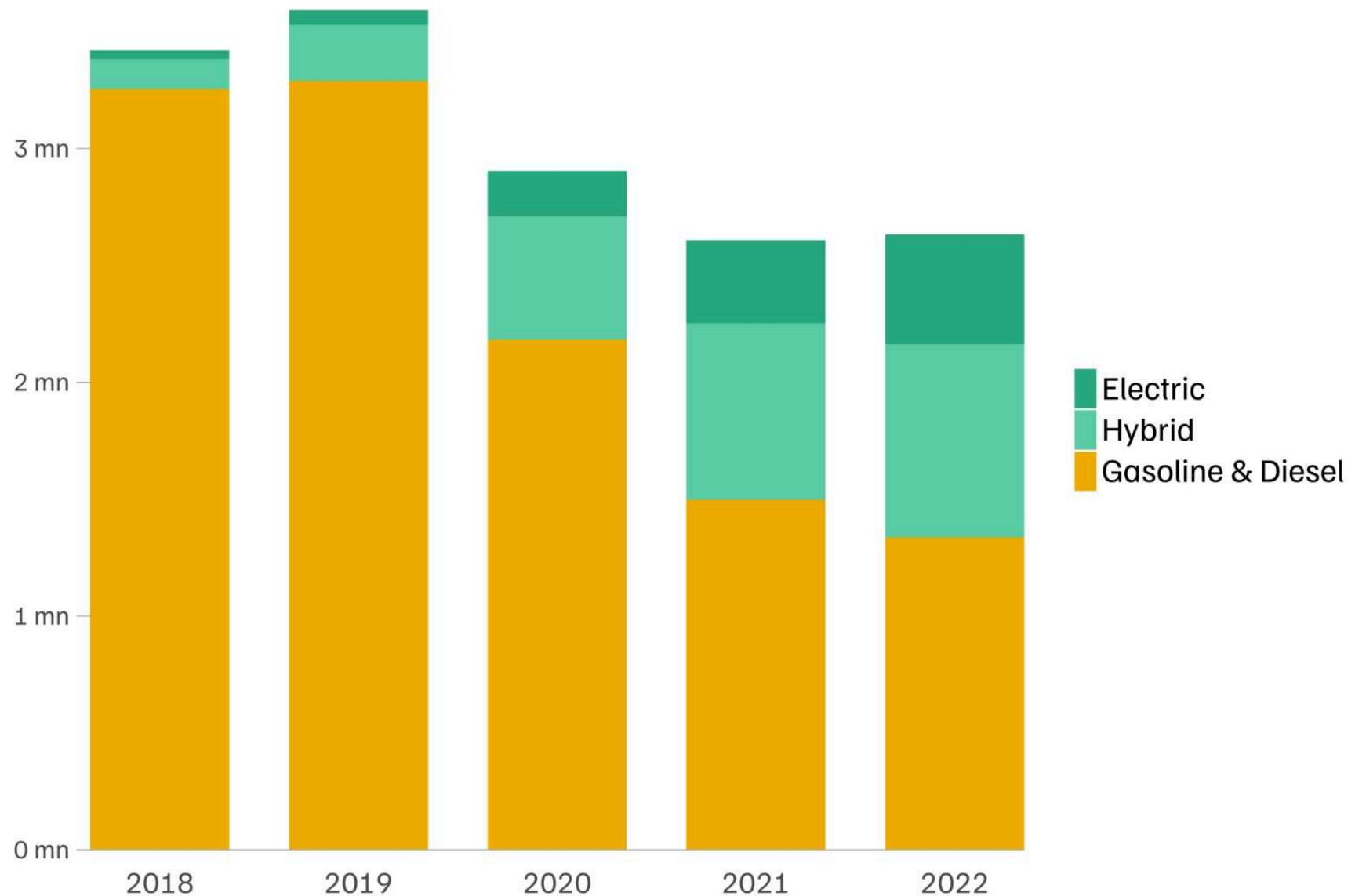
Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.



Time series with annotations by William Playfair from "The Commercial and Political Atlas and Statistical Breviary" (1786)



New vehicle registrations in Germany

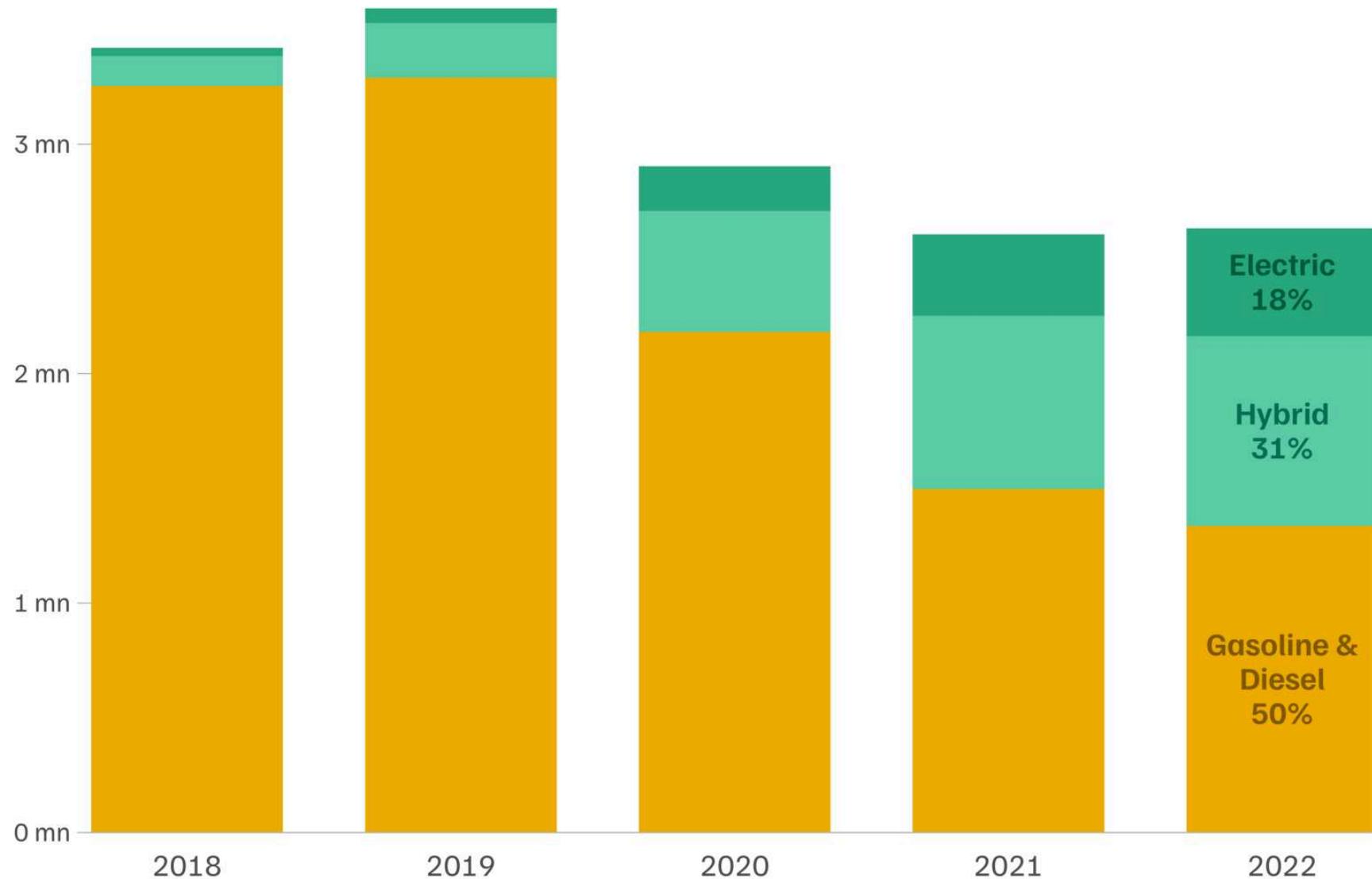


Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



In 2022, gasoline and diesel vehicles account for only half of all new registrations in Germany – hybrid and electric continue to increase.



Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



In 2022, **gasoline and diesel vehicles** account for only half of all new registrations in Germany – **hybrid** and **electric** continue to increase.

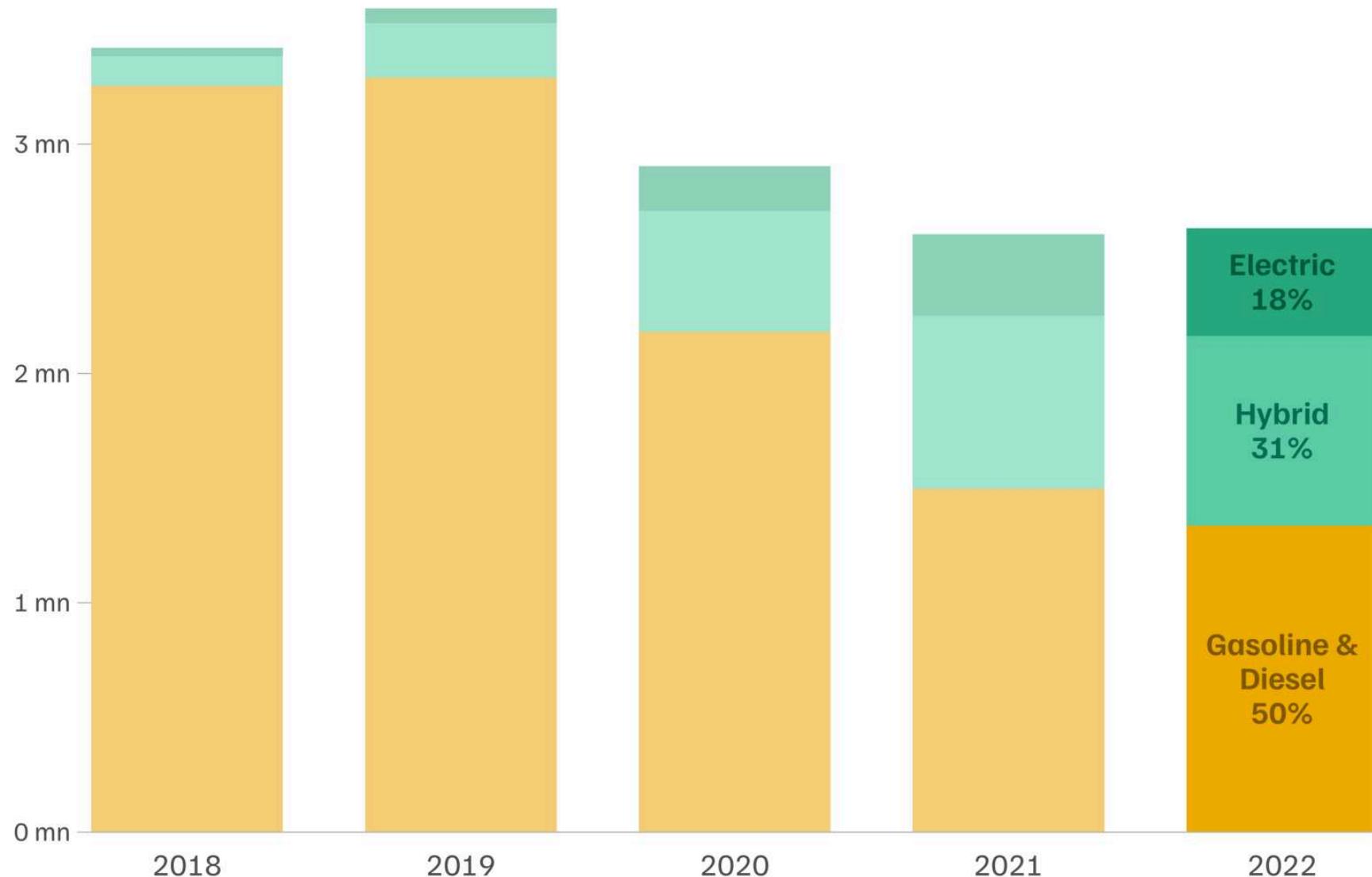


Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



In 2022, **gasoline and diesel vehicles** account for only half of all new registrations in Germany – **hybrid** and **electric** continue to increase.

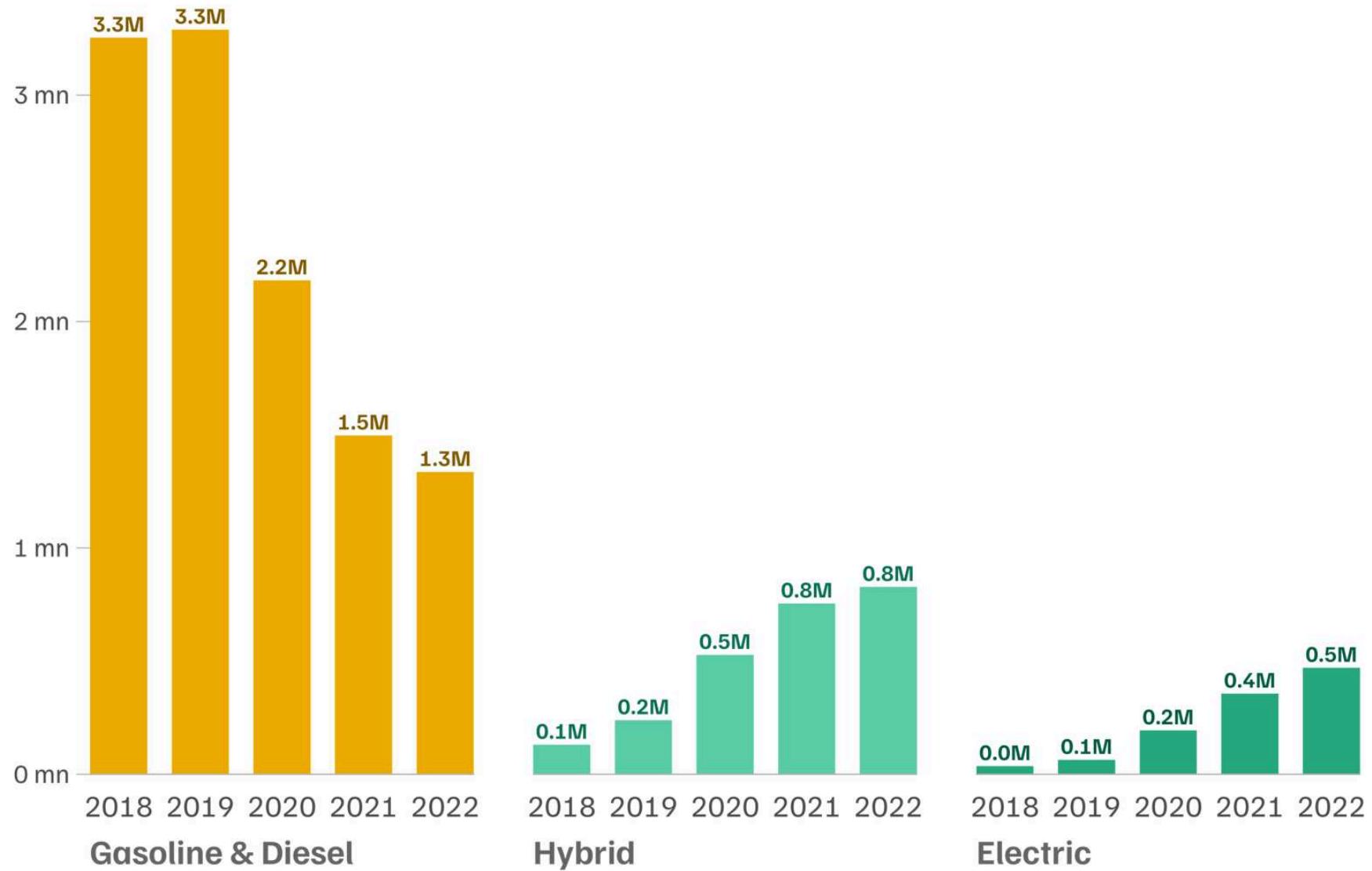


Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



Registrations of **gasoline and diesel vehicles** drop in Germany while those of new **hybrid and electric** cars are steadily increasing.

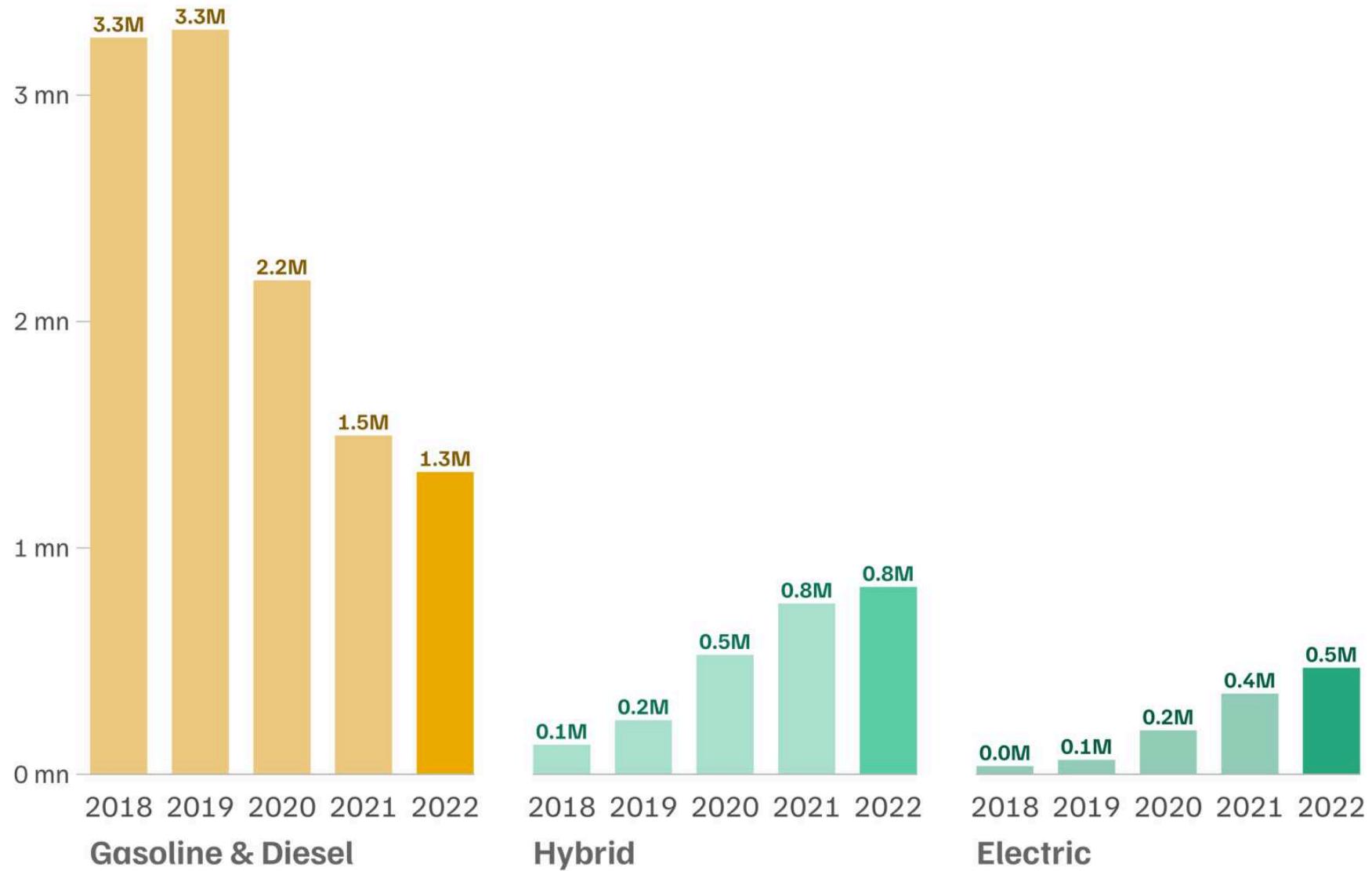


Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

Cédric Scherer Data Visualization & Information Design



Registrations of **gasoline and diesel vehicles** drop in Germany while those of new **hybrid and electric** cars are steadily increasing.

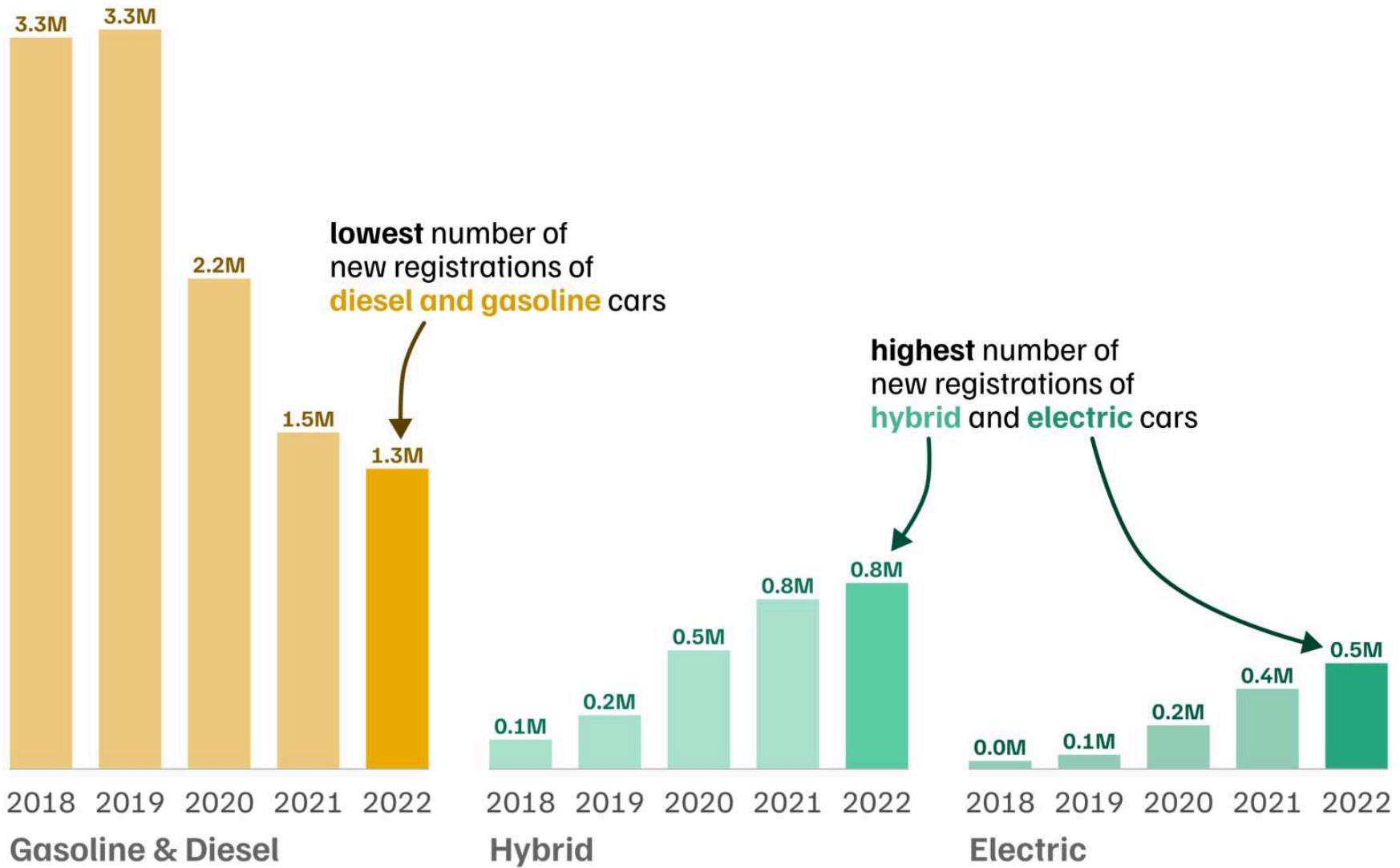


Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

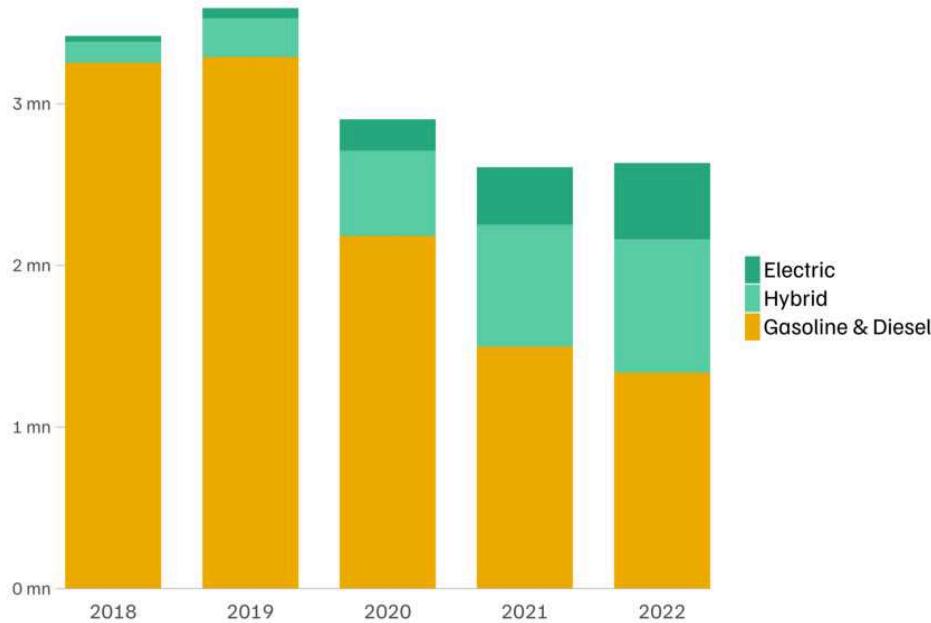
Cédric Scherer Data Visualization & Information Design



Registrations of **gasoline and diesel vehicles** drop in Germany while those of new **hybrid and electric** cars are steadily increasing.

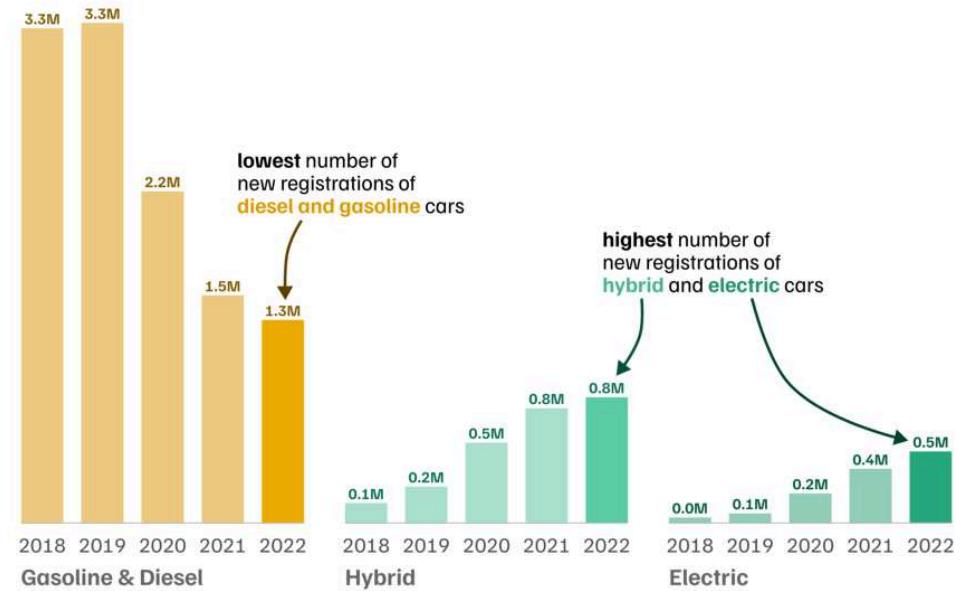


New vehicle registrations in Germany



Source: Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

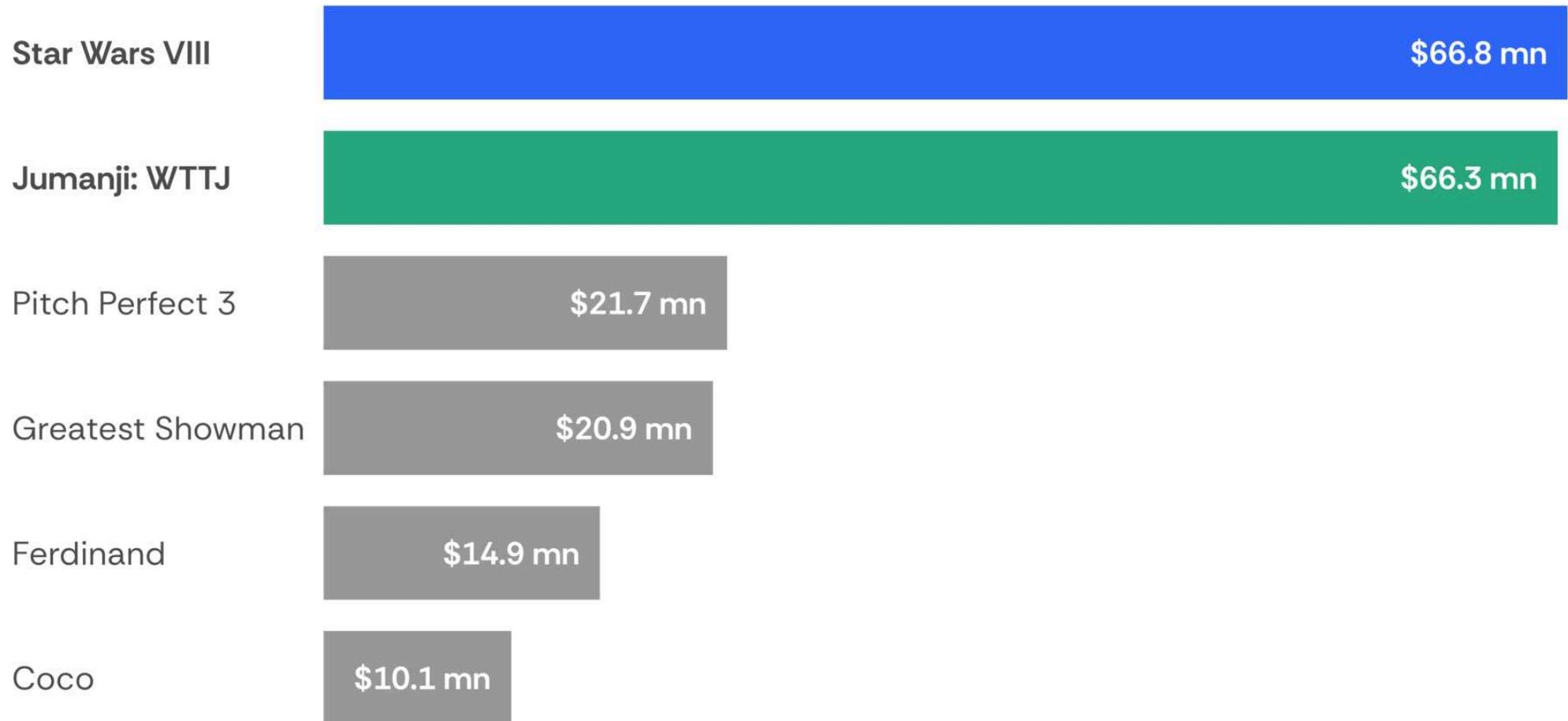
Registrations of **gasoline and diesel vehicles** drop in Germany while those of new **hybrid** and **electric** cars are steadily increasing.



Federal ministry for digital and transport / Kraftfahrt-Bundesamt • Graphic: Cédric Scherer

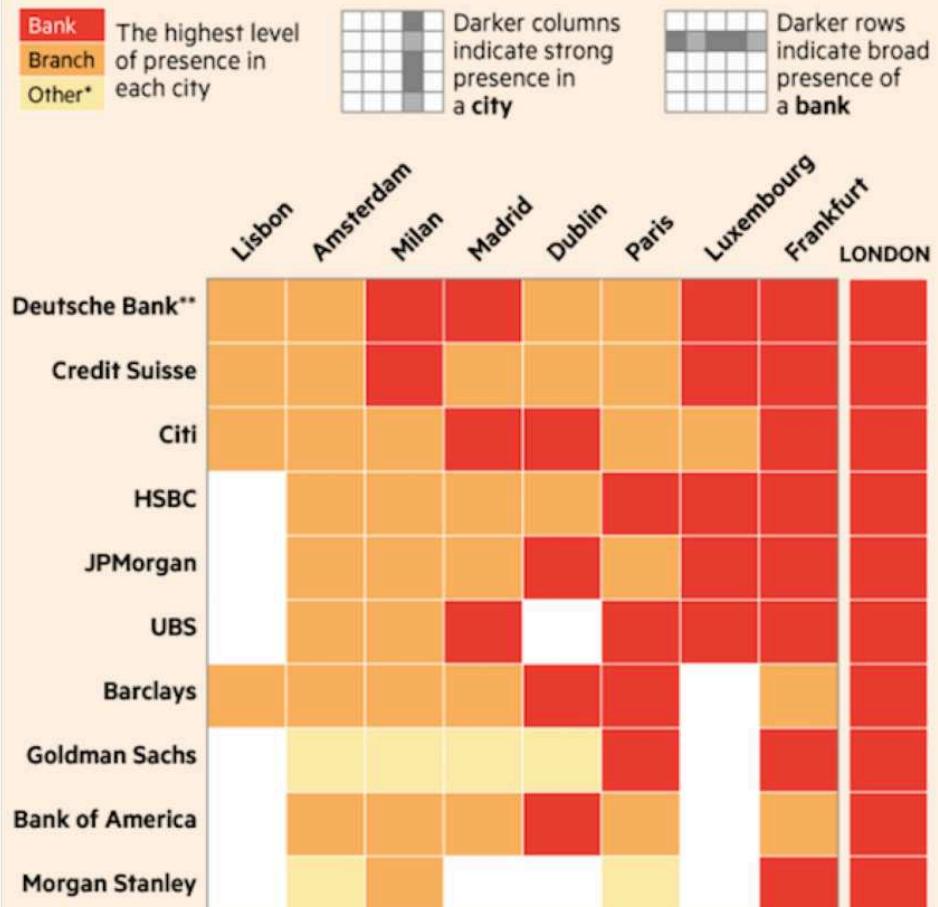


“Star Wars VIII – The Last Jedi” just barely topped the box office this weekend, facing stiff competition from “Jumanji: Welcome to the Jungle”.



Assist the Viewer

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



* Broker dealer branches are included for Morgan Stanley and Goldman Sachs as they are a significant part of their European network

** Deutsche Bank has a London subsidiary but its main entity is a branch

FT graphic Alan Smith, Laura Noonan Source: FT research

FT

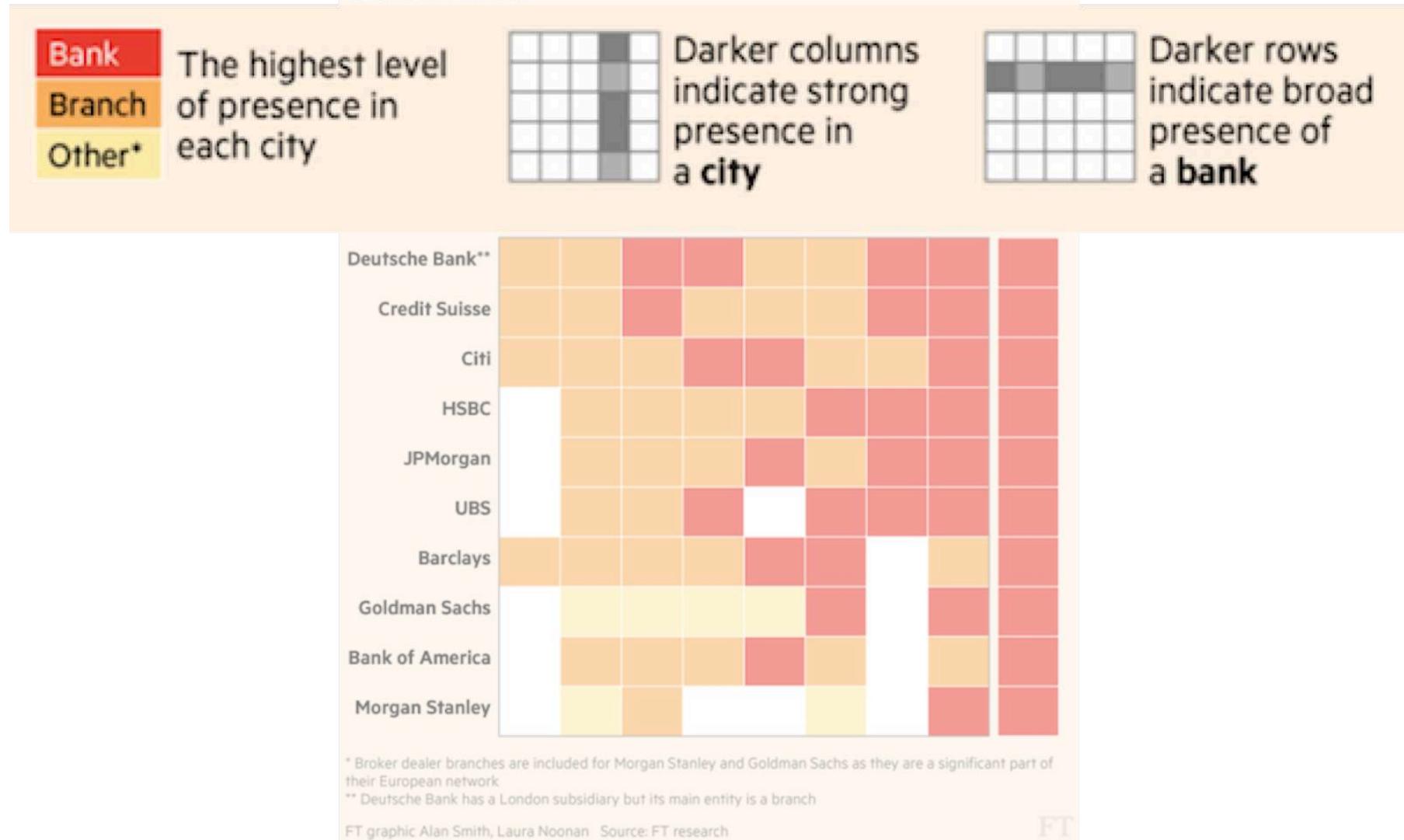
"Frankfurt vies for UK banking jobs post-Brexit" by Alan Smith & Laura Noonan (Financial Times)

Cédric Scherer Data Visualization & Information Design



Assist the Viewer

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



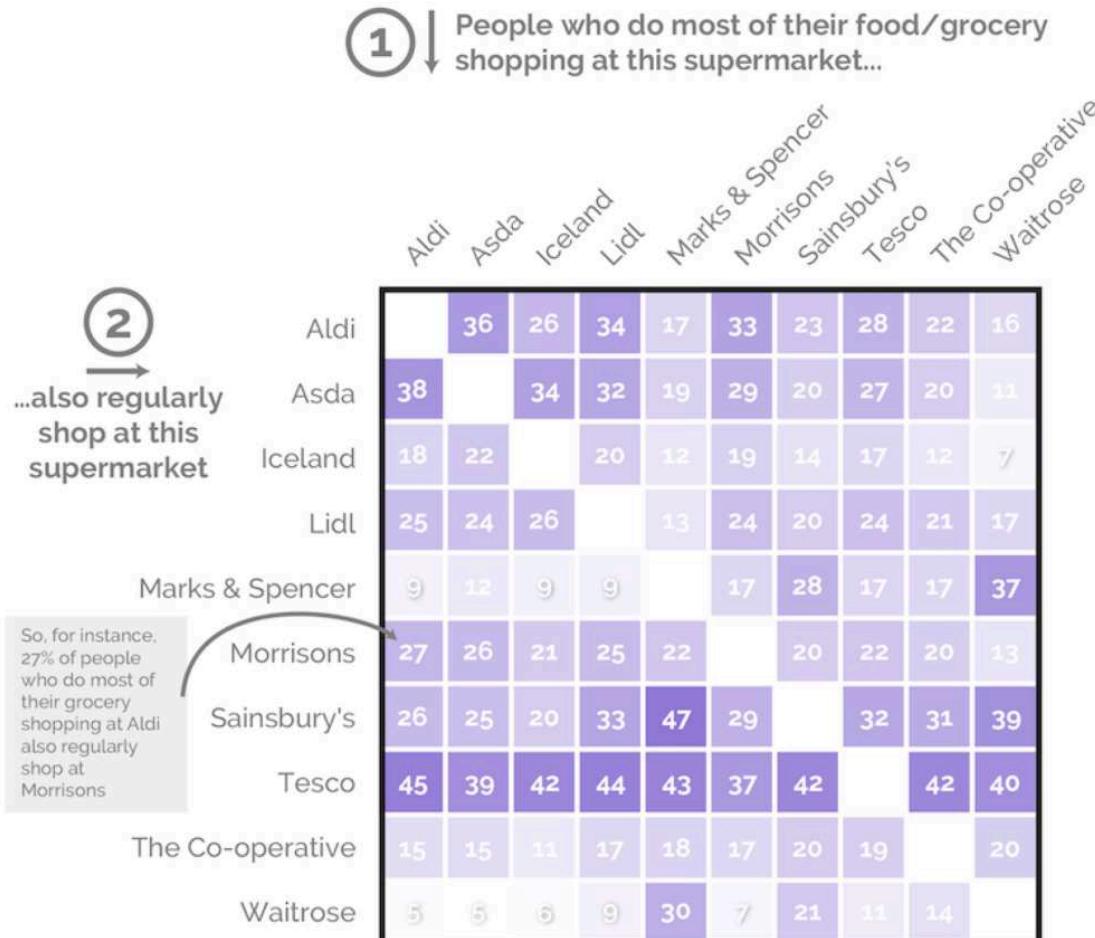
"Frankfurt vies for UK banking jobs post-Brexit" by Alan Smith & Laura Noonan (Financial Times)

Cédric Scherer Data Visualization & Information Design



Assist the Viewer

Supplementary supermarket shopping



YouGov | yougov.com

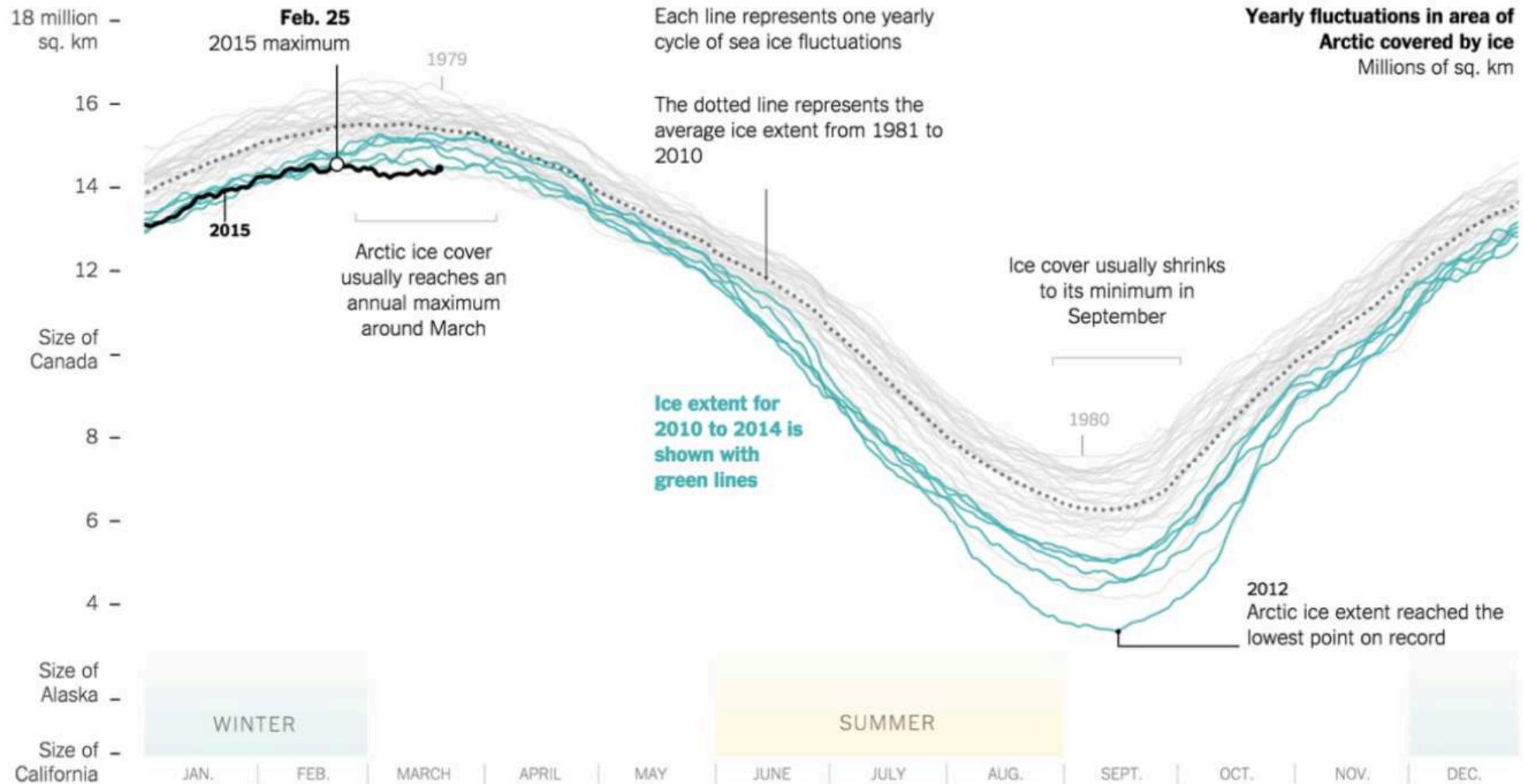
YouGov Profiles December 2018

"Tesco is the nation's primary AND secondary supermarket" by Matthew Smith (YouGov)

Cédric Scherer Data Visualization & Information Design



Assist the Viewer



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

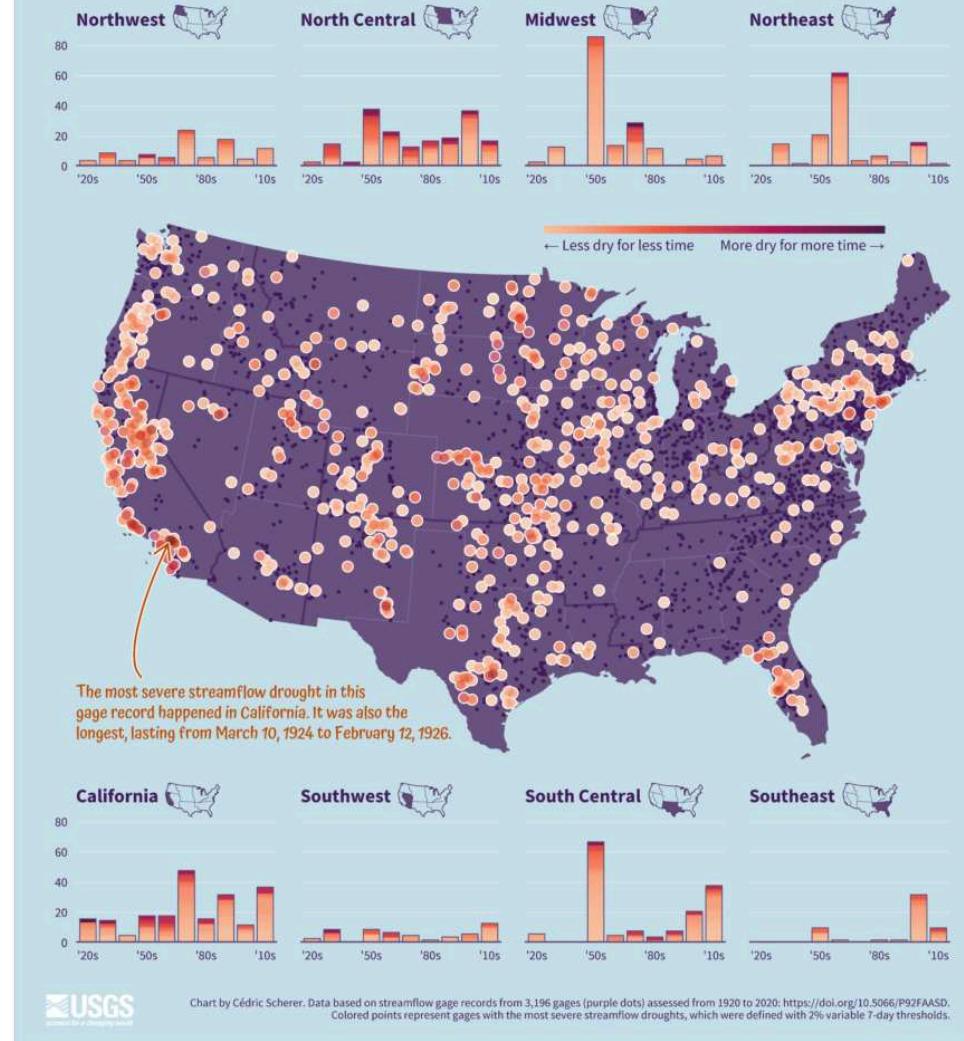
Cédric Scherer Data Visualization & Information Design



Assist the Viewer

100 Years of Streamflow Droughts

These are the 1000 most severe streamflow droughts at gages from 1920 to 2020 by region and decade.

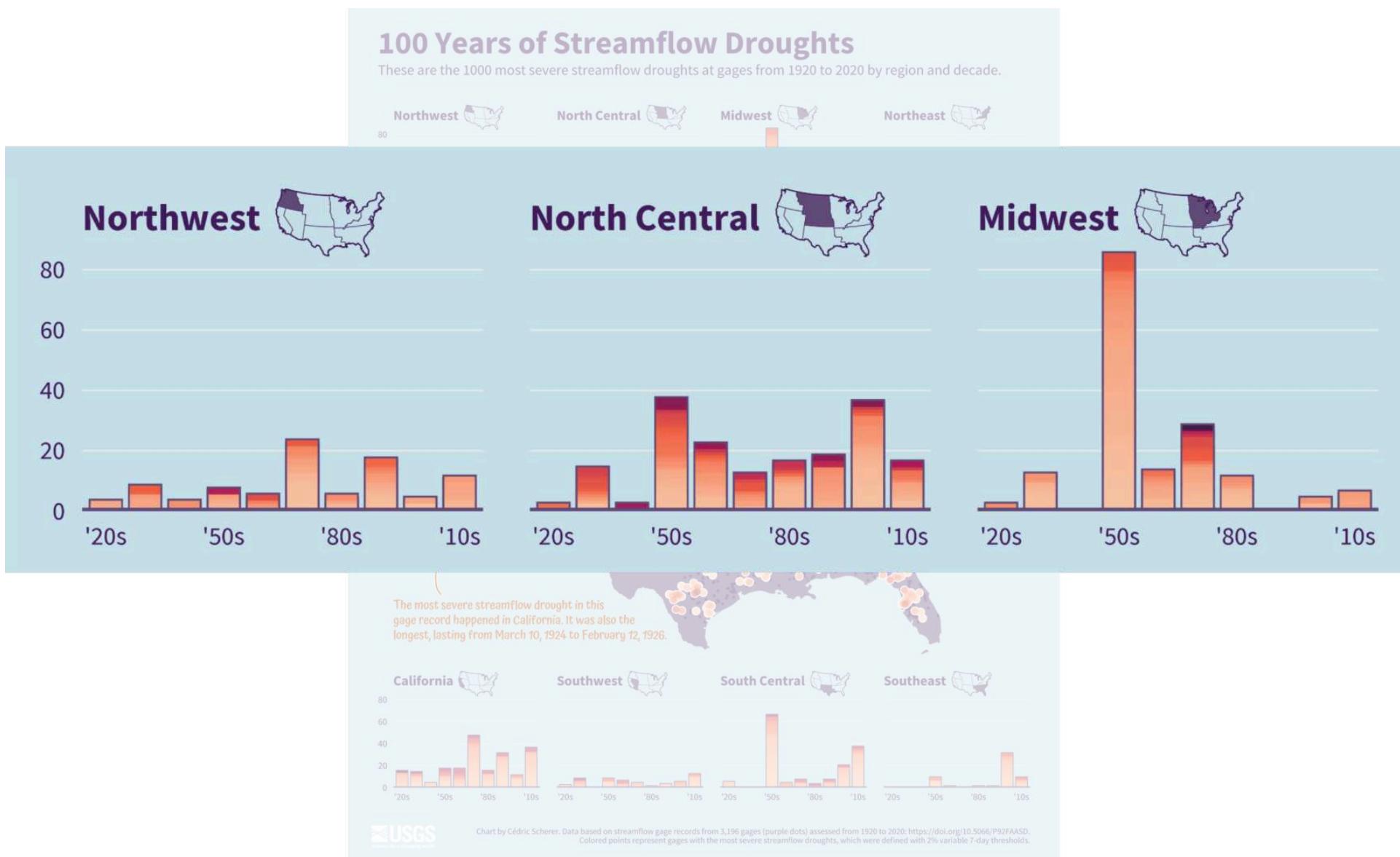


"100 Years of Streamflow Drought", in collaboration with USGS

Cédric Scherer Data Visualization & Information Design



Assist the Viewer

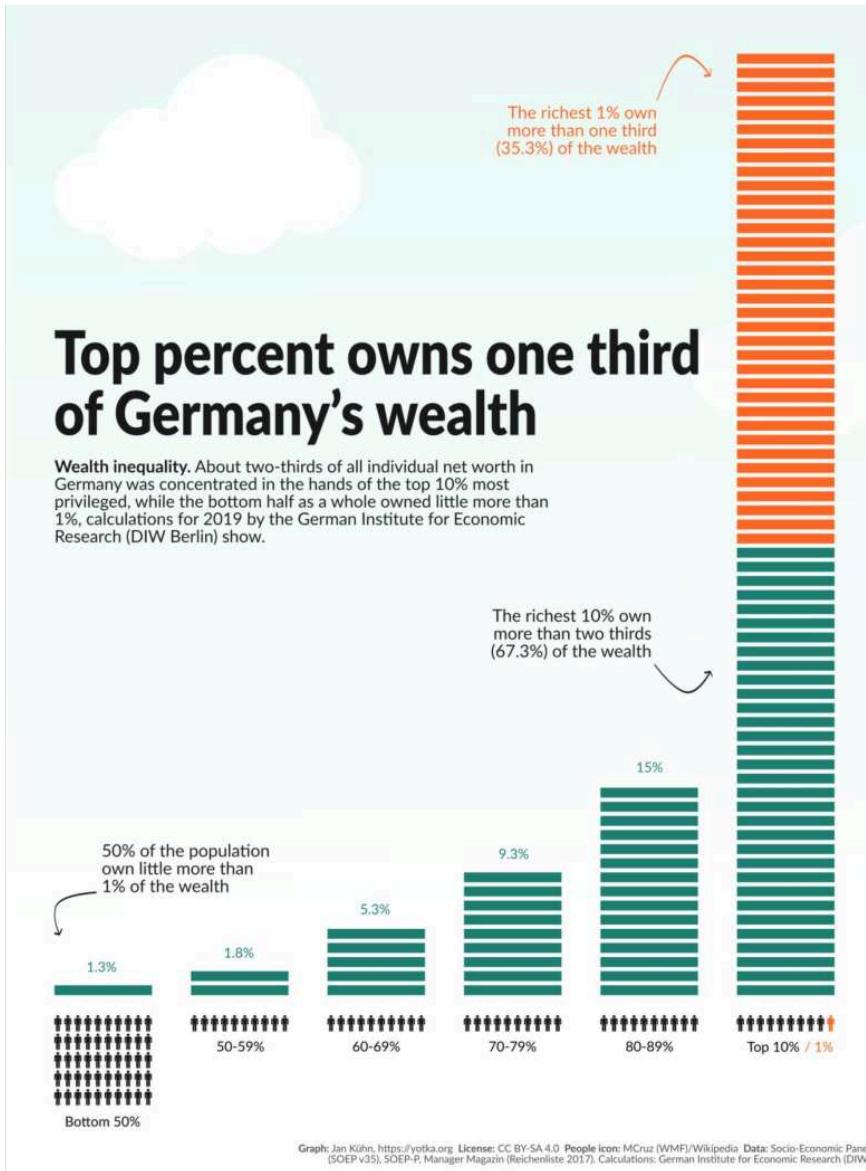


"100 Years of Streamflow Drought", in collaboration with USGS

Cédric Scherer Data Visualization & Information Design



Assist the Viewer

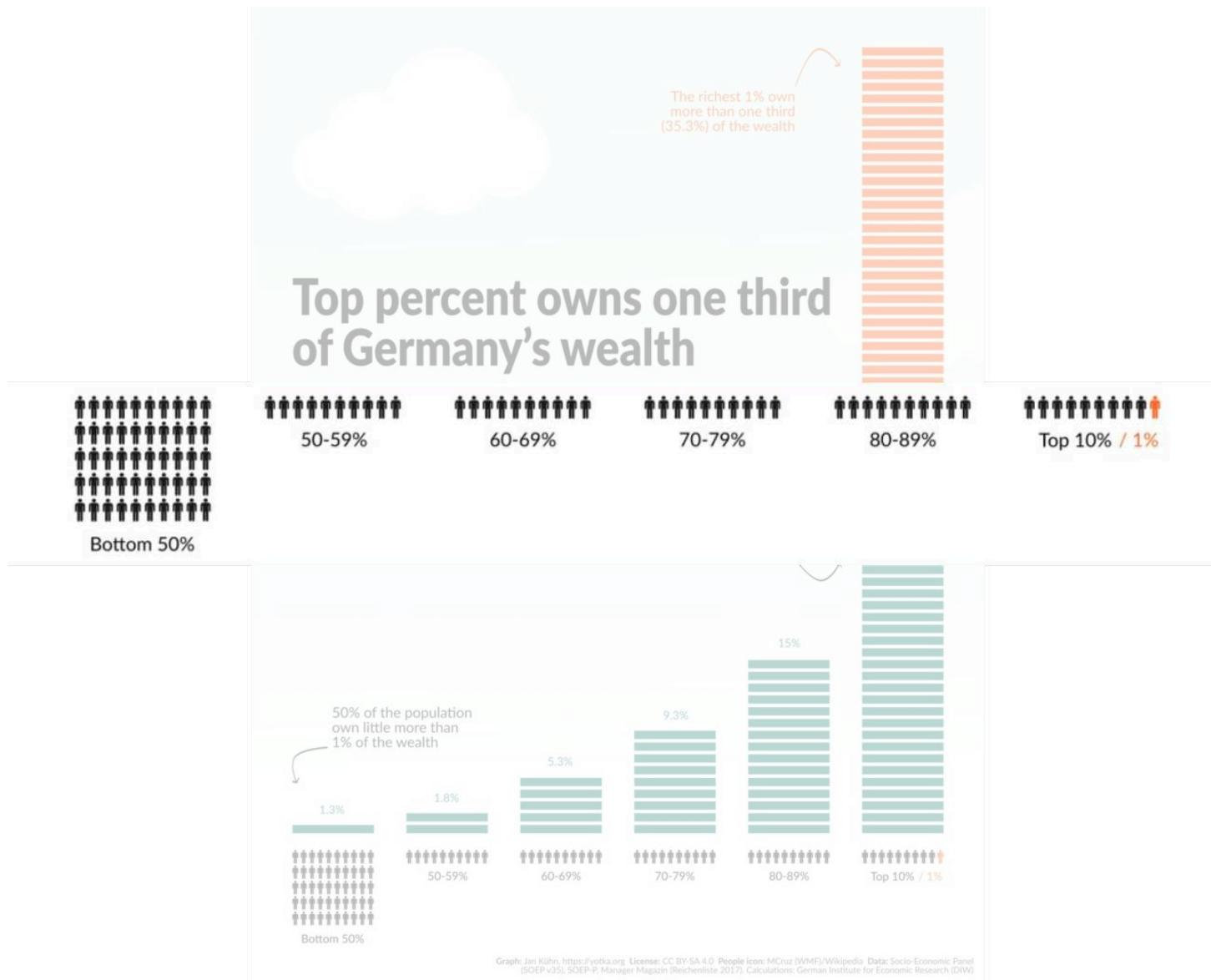


Source: Jan Kühn

Cédric Scherer Data Visualization & Information Design



Assist the Viewer



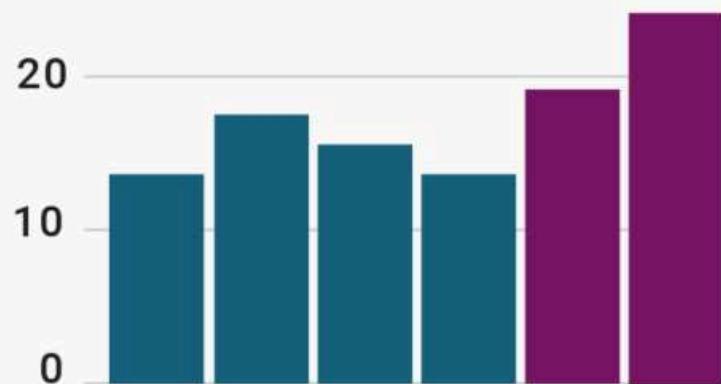
Source: Jan Kühn

Cédric Scherer Data Visualization & Information Design

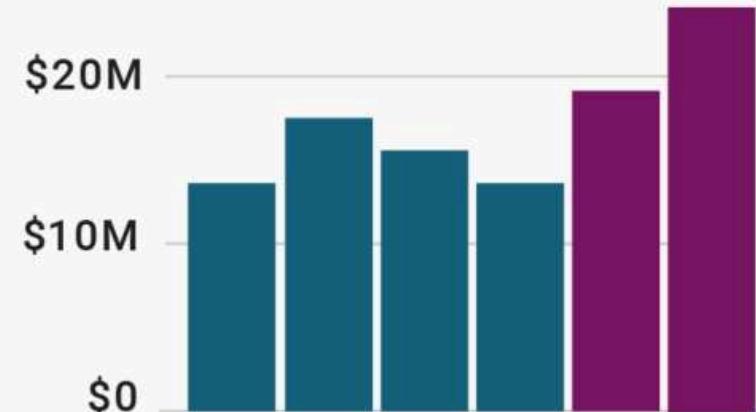


Assist the Viewer

REVENUE IN U.S. DOLLAR, IN MILLION,
2020-2025



REVENUE IN U.S. DOLLAR, 2020-2025



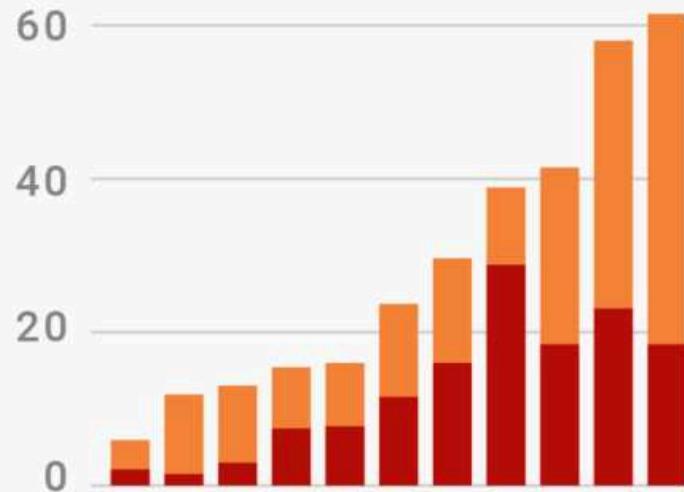
NOT IDEAL

BETTER

Source: Lisa Charlotte Muth, [Datawrapper Blog](#)



Assist the Viewer



NOT IDEAL

BETTER

Source: Lisa Charlotte Muth, [Datawrapper Blog](#)

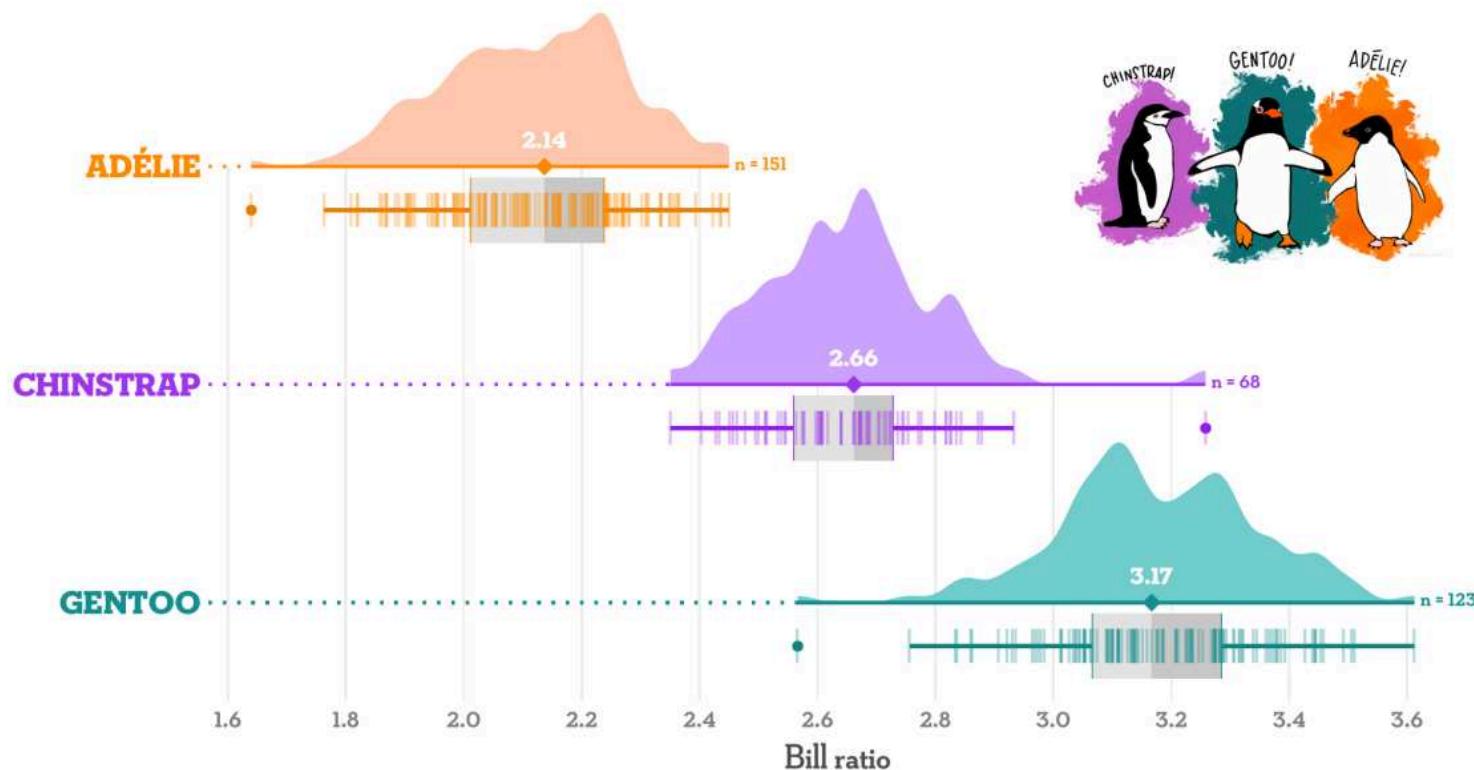
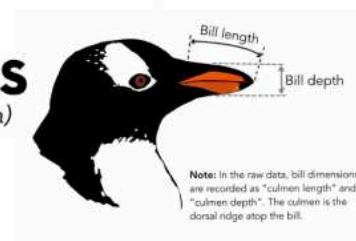


Illustrate with Images

BILL DIMENSIONS OF BRUSH-TAILED PENGUINS

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

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](https://doi.org/10.1371/journal.pone.0090081) • Illustrations: Allison Horst

Modified #TidyTuesday Contribution | Images: Allison Horst

Cédric Scherer Data Visualization & Information Design

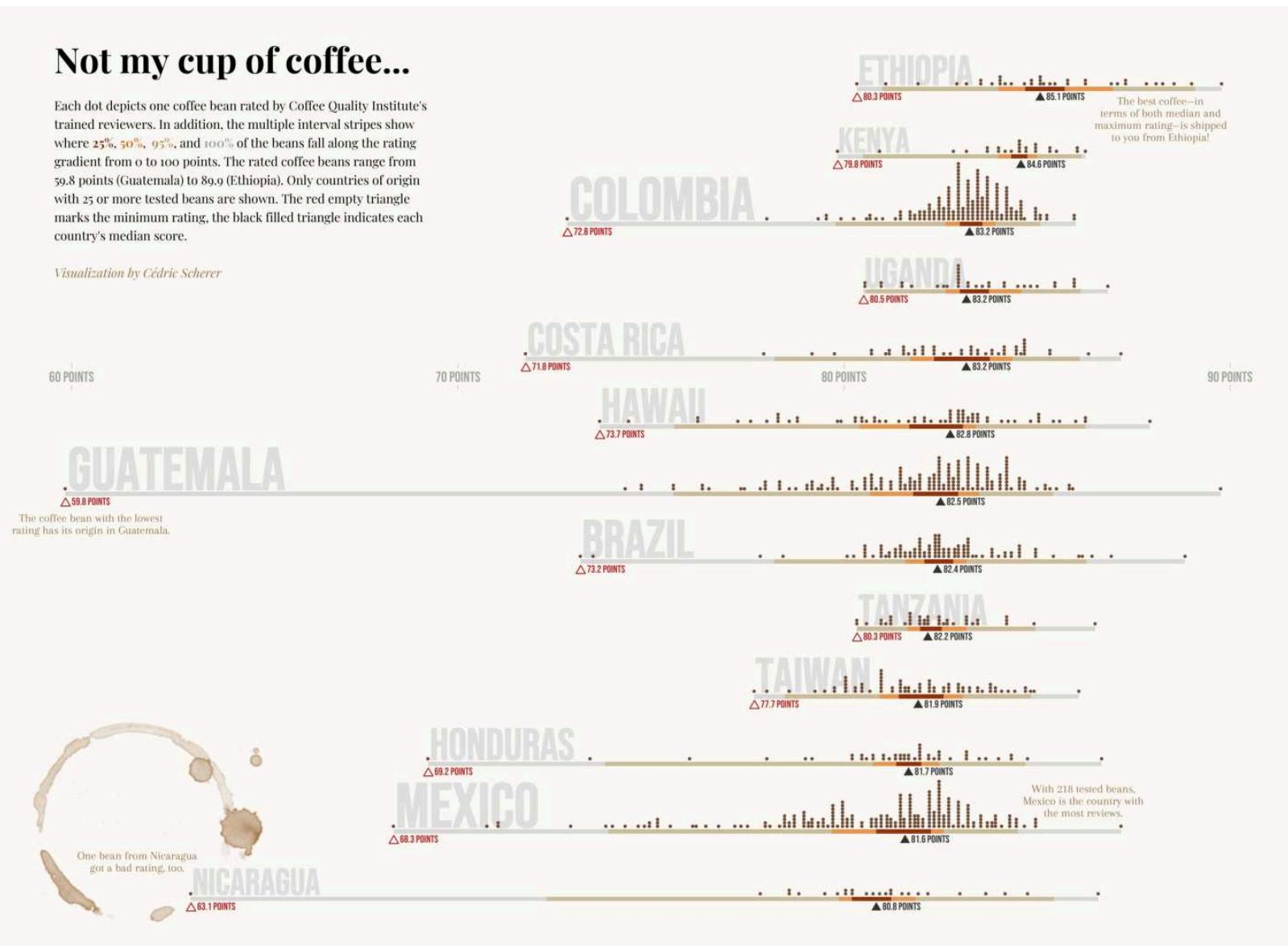


Illustrate with Images

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



"Not my Cup of Coffee" (#TidyTuesday Contribution)

Cédric Scherer Data Visualization & Information Design

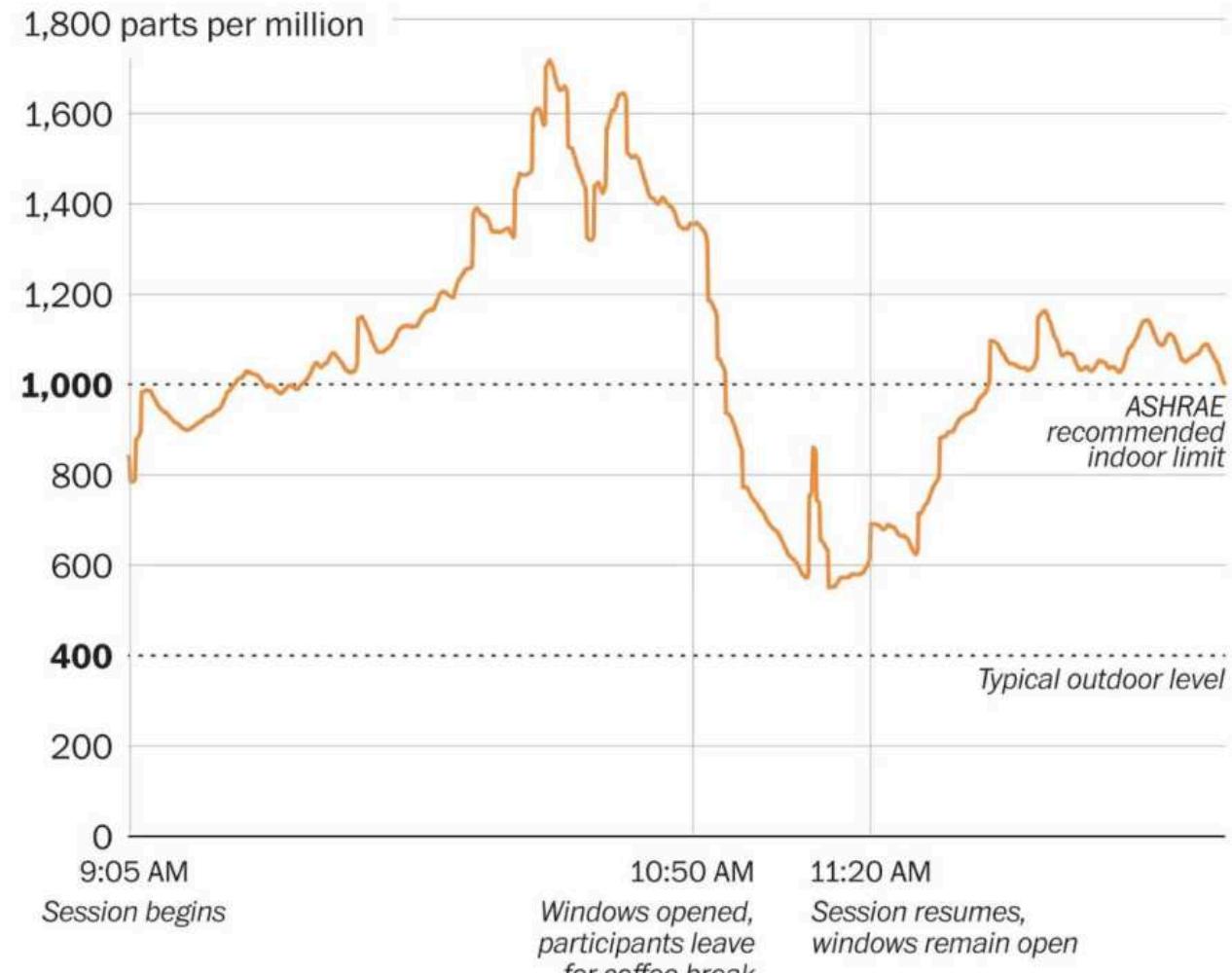


Wrap-Up



Clearing the air

CO₂ levels in an occupied conference room on June 4, 2019



Source: Adam Ginsburg

THE WASHINGTON POST

Source: "Clearing the Air" by Adam Ginsburg (Washington Post)

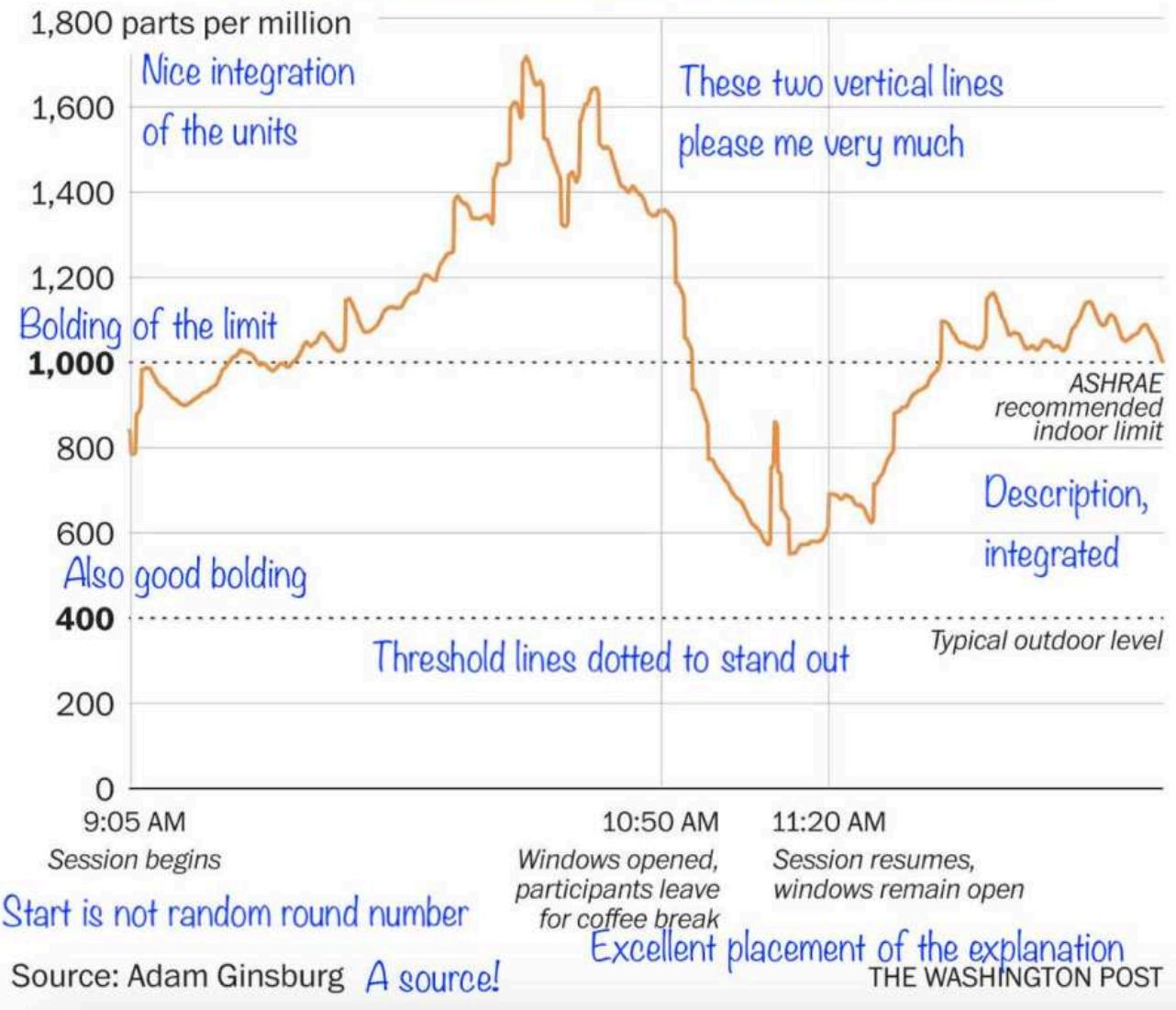
Cédric Scherer Data Visualization & Information Design



Clearing the air

Fun and helpful title

CO₂ levels in an occupied conference room on June 4, 2019
Units and metho in a subtitle, NOT in vertical text on the side



Notes by Francis Gagnon (Voilà)



What Makes a Good Data Visualization?

- » **Integrity** (information)
- » **Story** (interestingness)
- » **Goal** (usefulness)
- » **Visual Form** (beauty)



How to Craft Effective Data Visualizations

- understand visual perception
- evaluate the context
- define meaningful objectives
- condense information
- guide the viewer



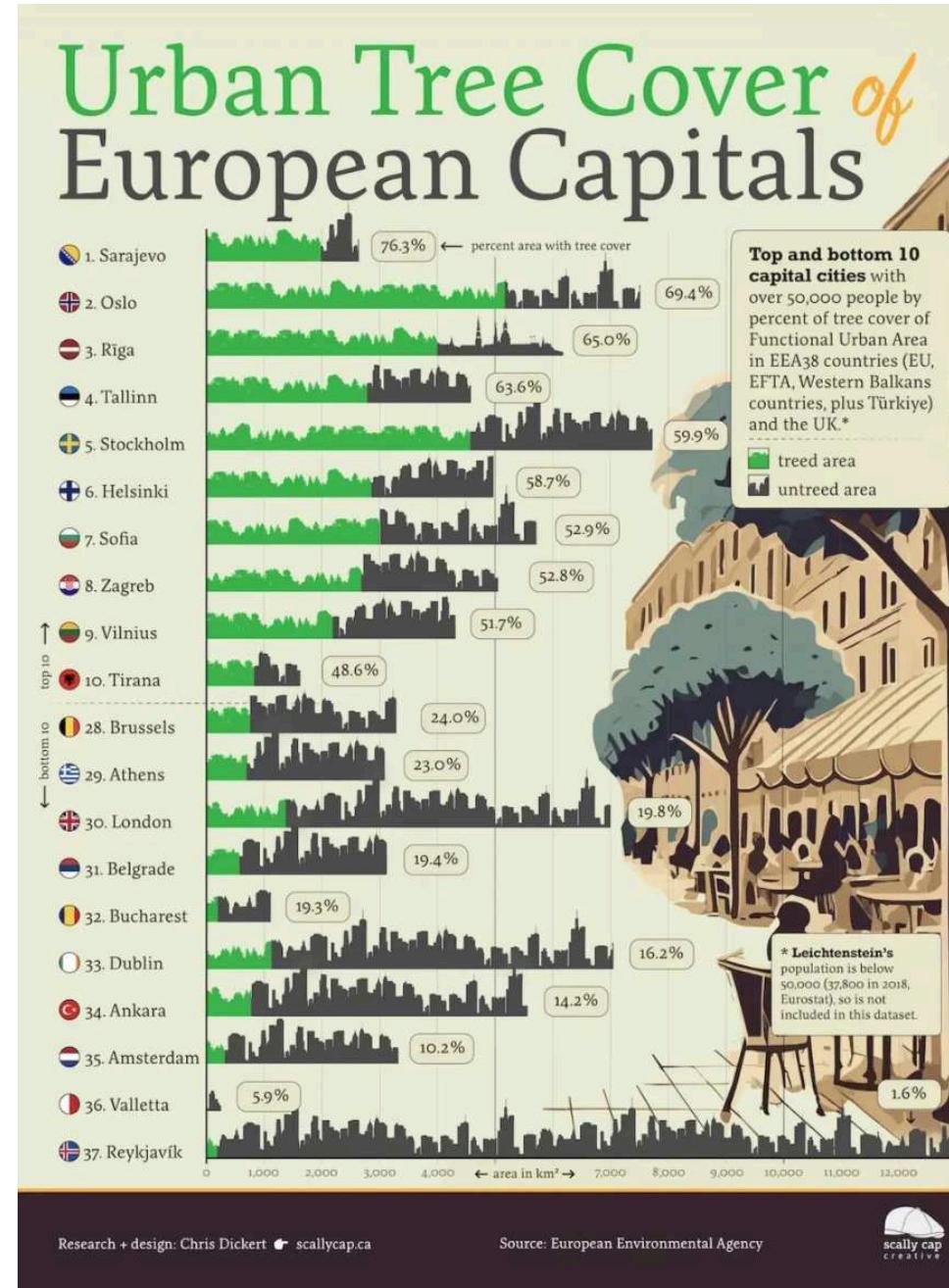
Exercises



Exercise 1

- **Take a closer look at (one of) the following three visualizations and address these questions:**
 - What is the main message you learn from the graphic?
 - What is the purpose of the visualization?
 - Who is the audience?
- **Rate the graphics according to the four levels:** information, story, goal, and visual form.
- **Collect three things you notice, no matter if positive or negative.**
 - How could you fix the details you dislike?



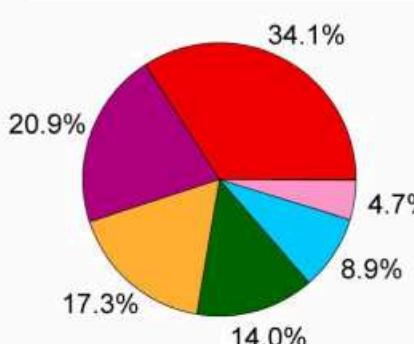
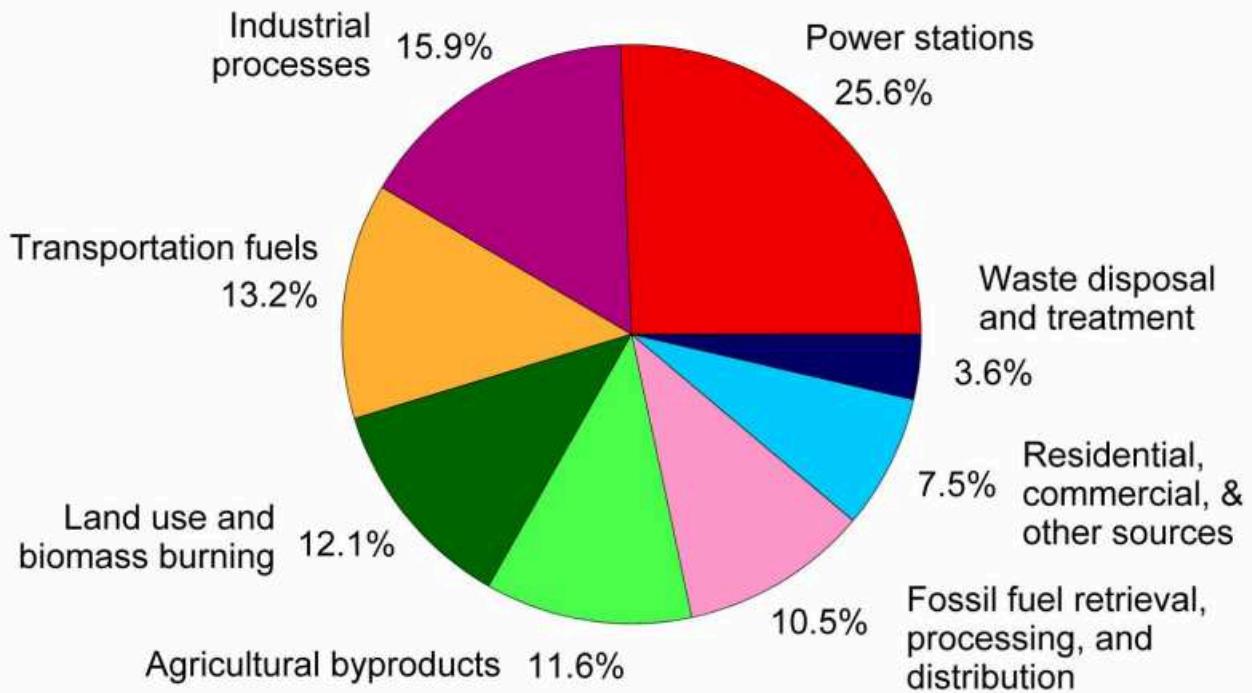


"Urban Tree Cover of European Capital Cities" by Chris Dickert

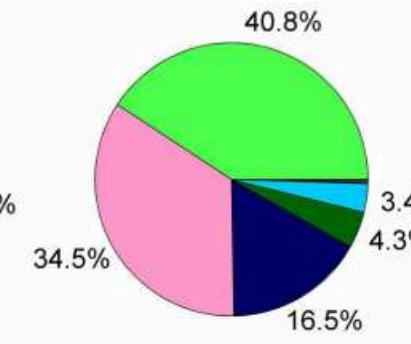
Cédric Scherer Data Visualization & Information Design



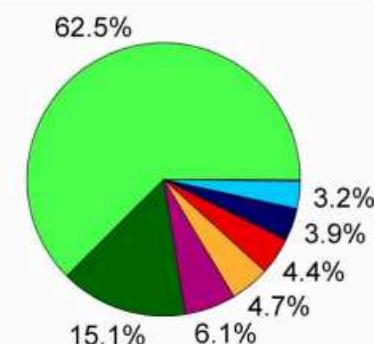
Annual Greenhouse Gas Emissions by Sector



Carbon Dioxide
(72% of total)



Methane
(20% of total)



Nitrous Oxide
(5% of total)

Emissions data for the year 2010 derived from the Emissions Database for Global Atmospheric Research (EDGAR) v4.2 FT2010

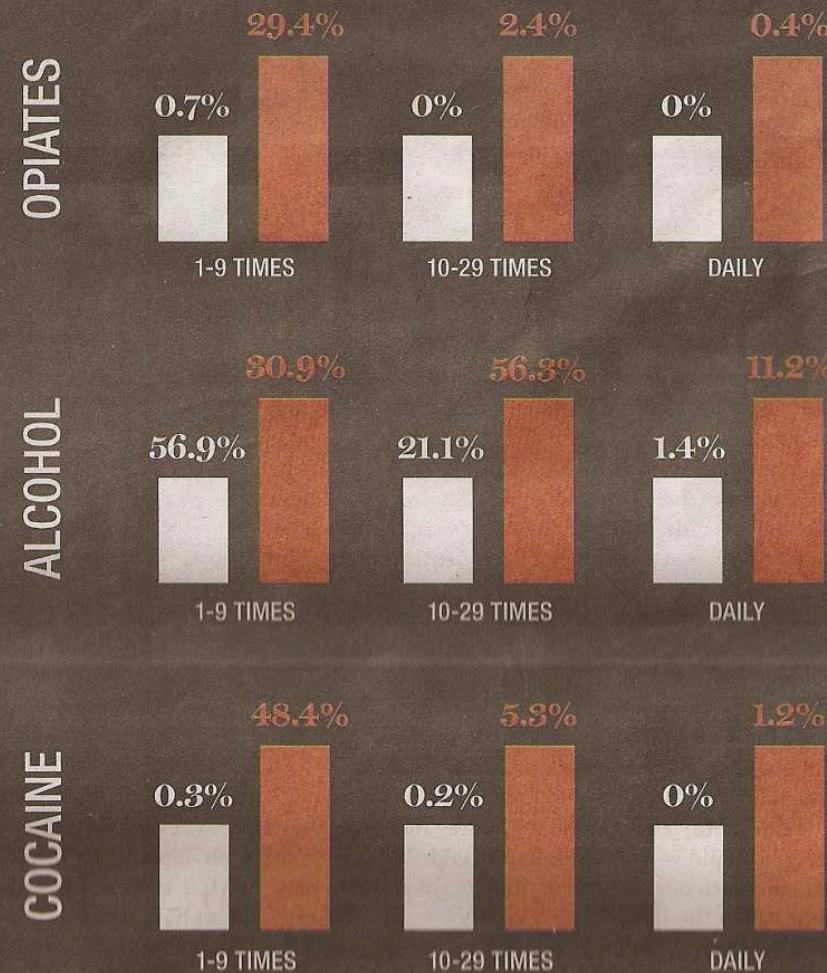
Source: Dr. Robert Rohde ([Tweet](#))

Cédric Scherer Data Visualization & Information Design



BY THE NUMBERS

The National Collegiate Health Assessment was taken by 1,000 UCSB students in Spring 2009. Participants were asked how frequently they used substances over the past 30 days. Numbers in white reflect actual student use, while red numbers indicate perceived substance use. The average age of participants was 20 years and approximately 99 percent were full-time students.

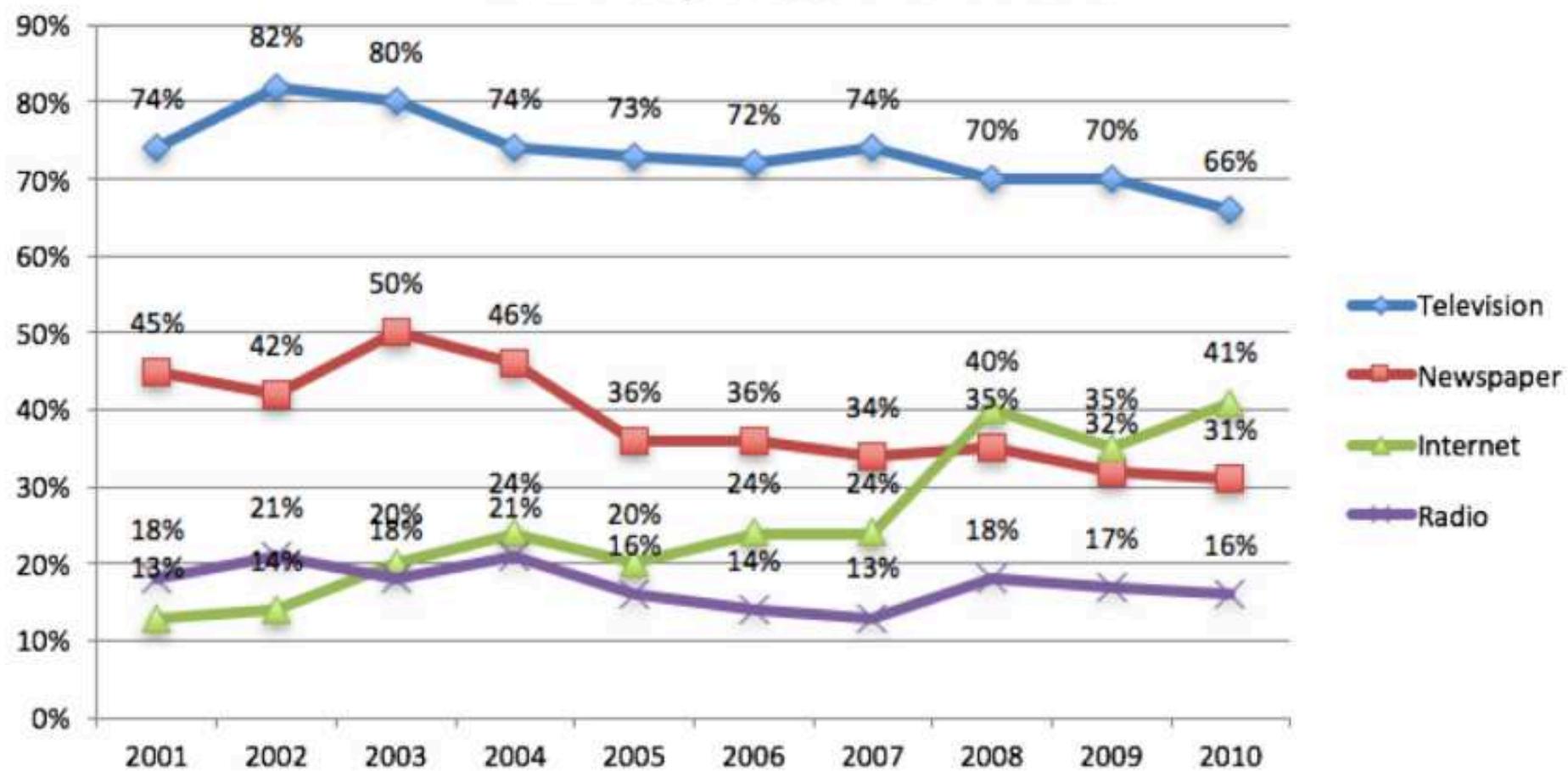


Exercise 2

- **Inspect the line chart “How People Get Their News”.**
 - Identify the main message of the chart.
 - What are details that distract you as a viewer?
 - Is all the information relevant? Or could you condense the data?
- **Improve the graphic to focus attention.**
 - Make use of preattentive attributes to emphasize the main message.
 - Eliminate the clutter from the graphic.
 - Apply Gestalt principles if possible.
 - How could you improve the interpretation and comprehension?



How People Get Their News



AN INCREASING PROPORTION CITE THE INTERNET AS THEIR PRIMARY NEWS SOURCE.

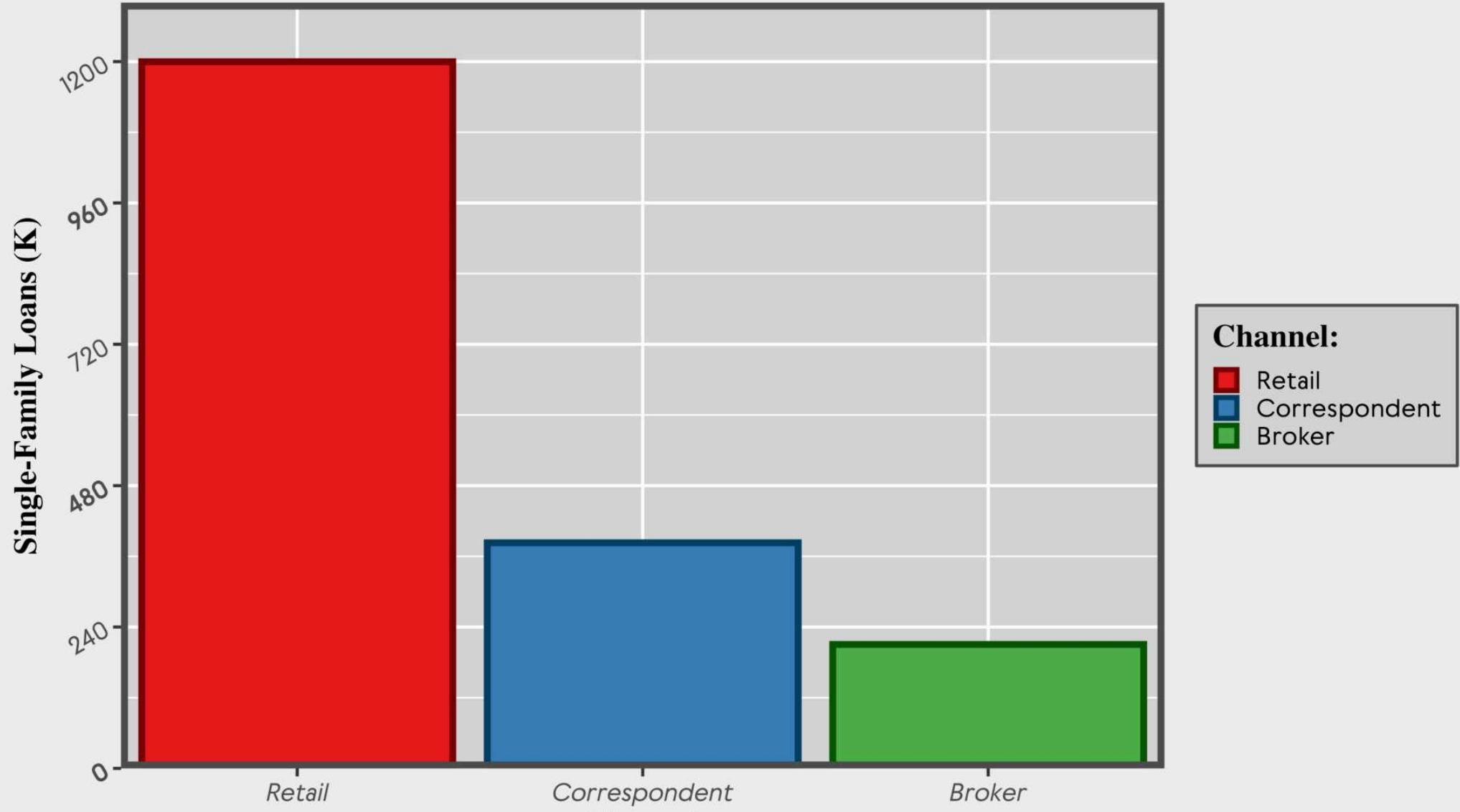
DATA SOURCE: PEW RESEARCH CENTER. REPRESENTS RESPONSES TO THE QUESTION, "WHERE DO YOU GET MOST OF YOUR NEWS ABOUT NATIONAL AND INTERNATIONAL ISSUES? FIGURES SUM TO MORE THAN 100% BECAUSE RESPONDENTS COULD VOLUNTEER UP TO TWO MAIN SOURCES.



Exercise 3

- **Declutter the following visualization.**
 - List at least five components that can be simplified or eliminated.
 - Identify any redundant elements.
 - Suggest ways to improve data labels for clarity and readability
 - **Bonus:** What might be reasons to include colors here?



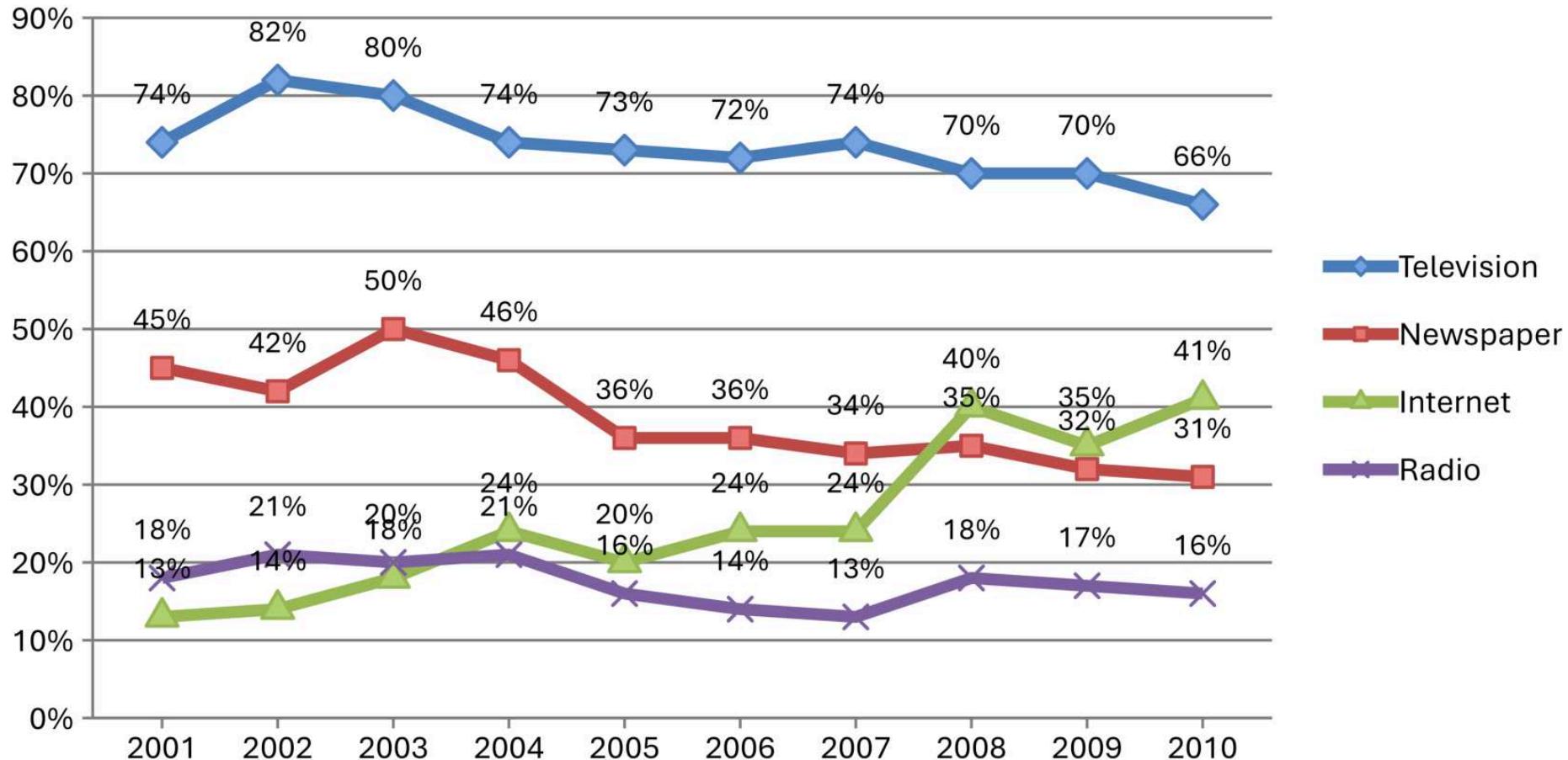


Exercise 2

Suggested Solutions



How People Get Their News

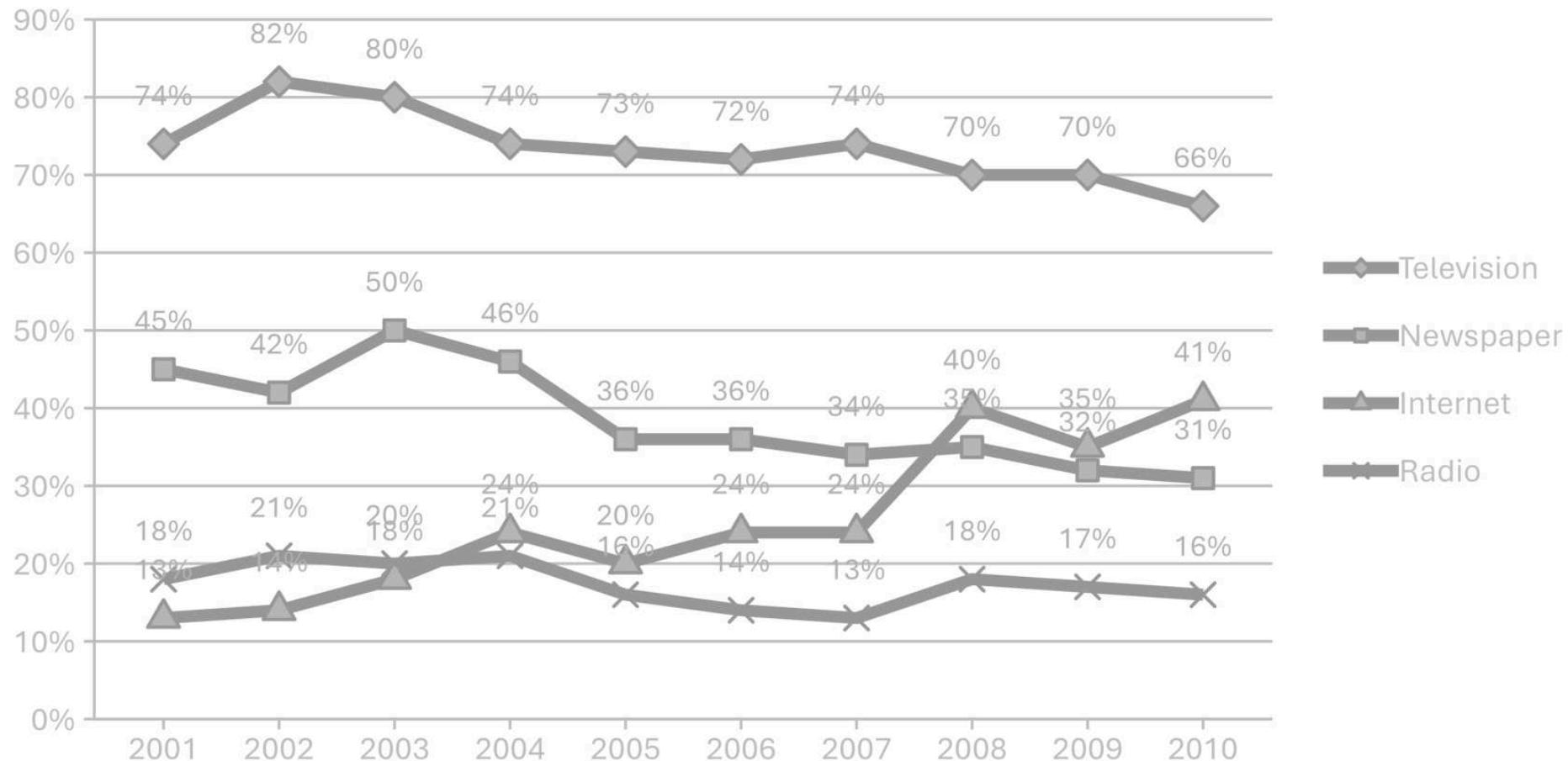


AN INCREASING PROPORTION CITE THE INTERNET AS THEIR PRIMARY NEWS SOURCE.

DATA SOURCE: PEW RESEARCH CENTER. REPRESENTS RESPONSES TO THE QUESTION "WHERE DO YOU GET MOST OF YOUR NEWS ABOUT NATIONAL AND INTERNATIONAL ISSUES? FIGURES SUM TO MORE THAN 100% BECAUSE RESPONDENTS COULD VOLUNTEER UP TO TWO MAIN SOURCES.



How People Get Their News

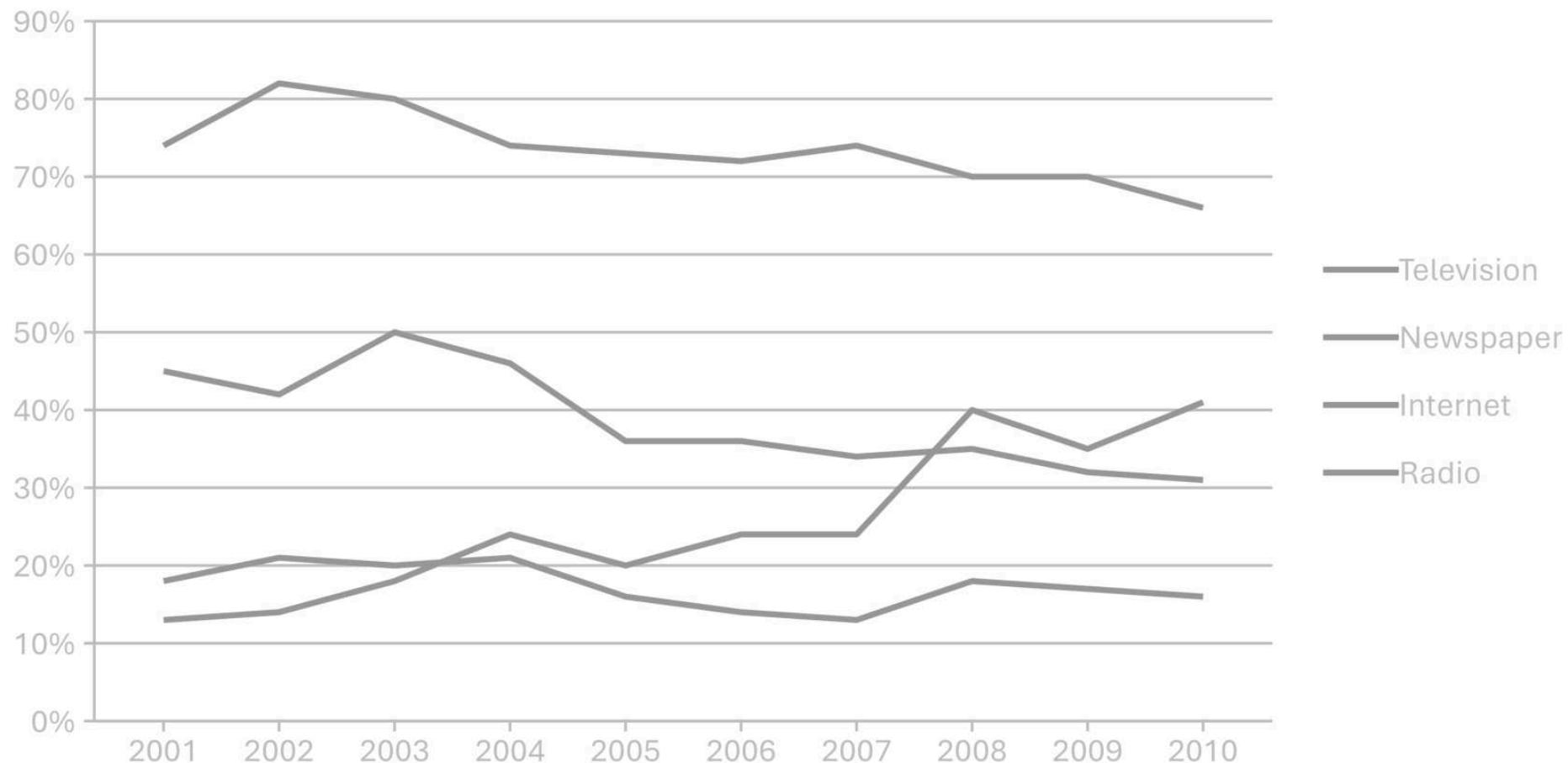


AN INCREASING PROPORTION CITE THE INTERNET AS THEIR PRIMARY NEWS SOURCE.

DATA SOURCE: PEW RESEARCH CENTER. REPRESENTS RESPONSES TO THE QUESTION "WHERE DO YOU GET MOST OF YOUR NEWS ABOUT NATIONAL AND INTERNATIONAL ISSUES? FIGURES SUM TO MORE THAN 100% BECAUSE RESPONDENTS COULD VOLUNTEER UP TO TWO MAIN SOURCES.



How People Get Their News

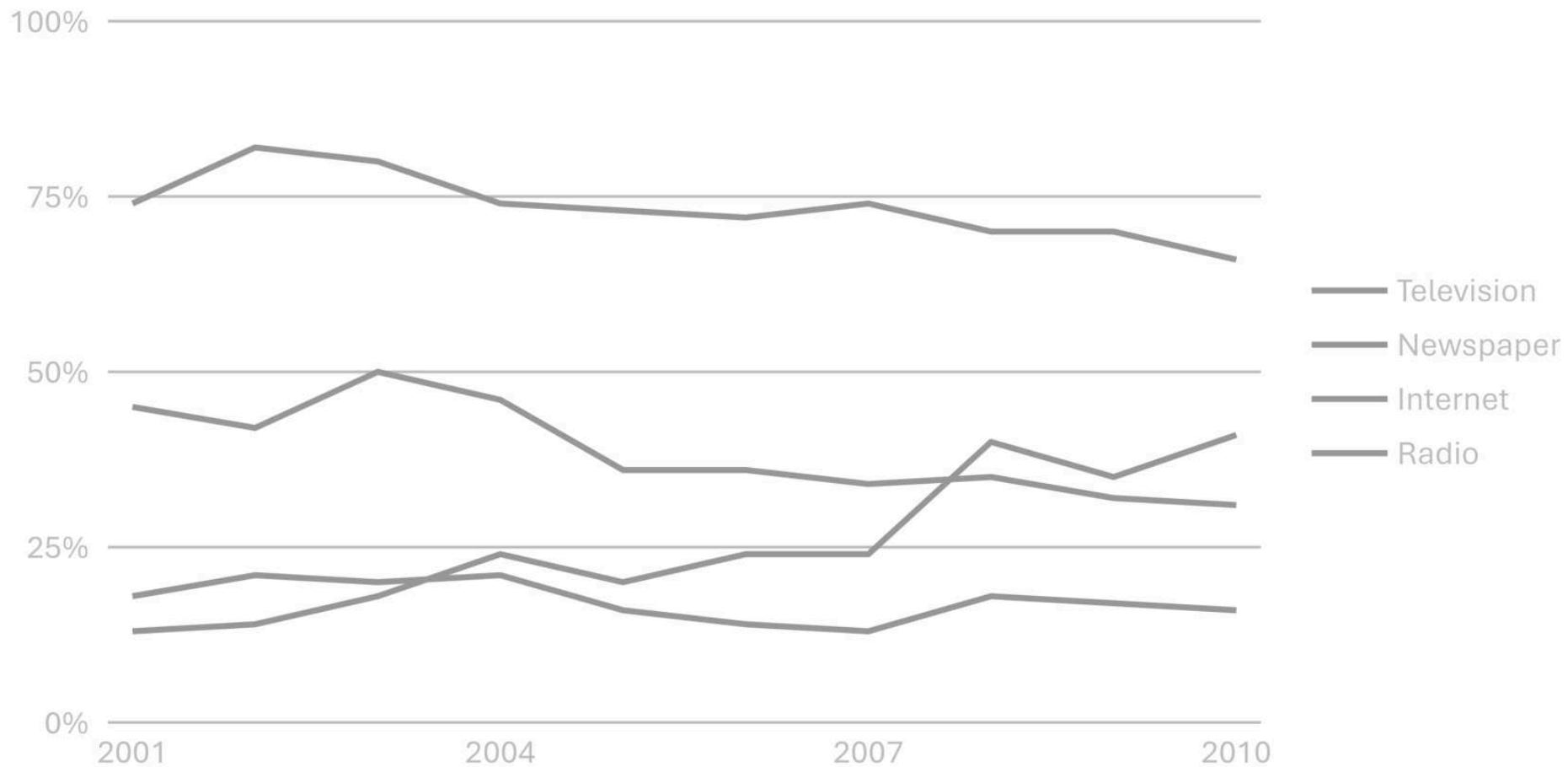


AN INCREASING PROPORTION CITE THE INTERNET AS THEIR PRIMARY NEWS SOURCE.

DATA SOURCE: PEW RESEARCH CENTER. REPRESENTS RESPONSES TO THE QUESTION "WHERE DO YOU GET MOST OF YOUR NEWS ABOUT NATIONAL AND INTERNATIONAL ISSUES? FIGURES SUM TO MORE THAN 100% BECAUSE RESPONDENTS COULD VOLUNTEER UP TO TWO MAIN SOURCES.



How People Get Their News

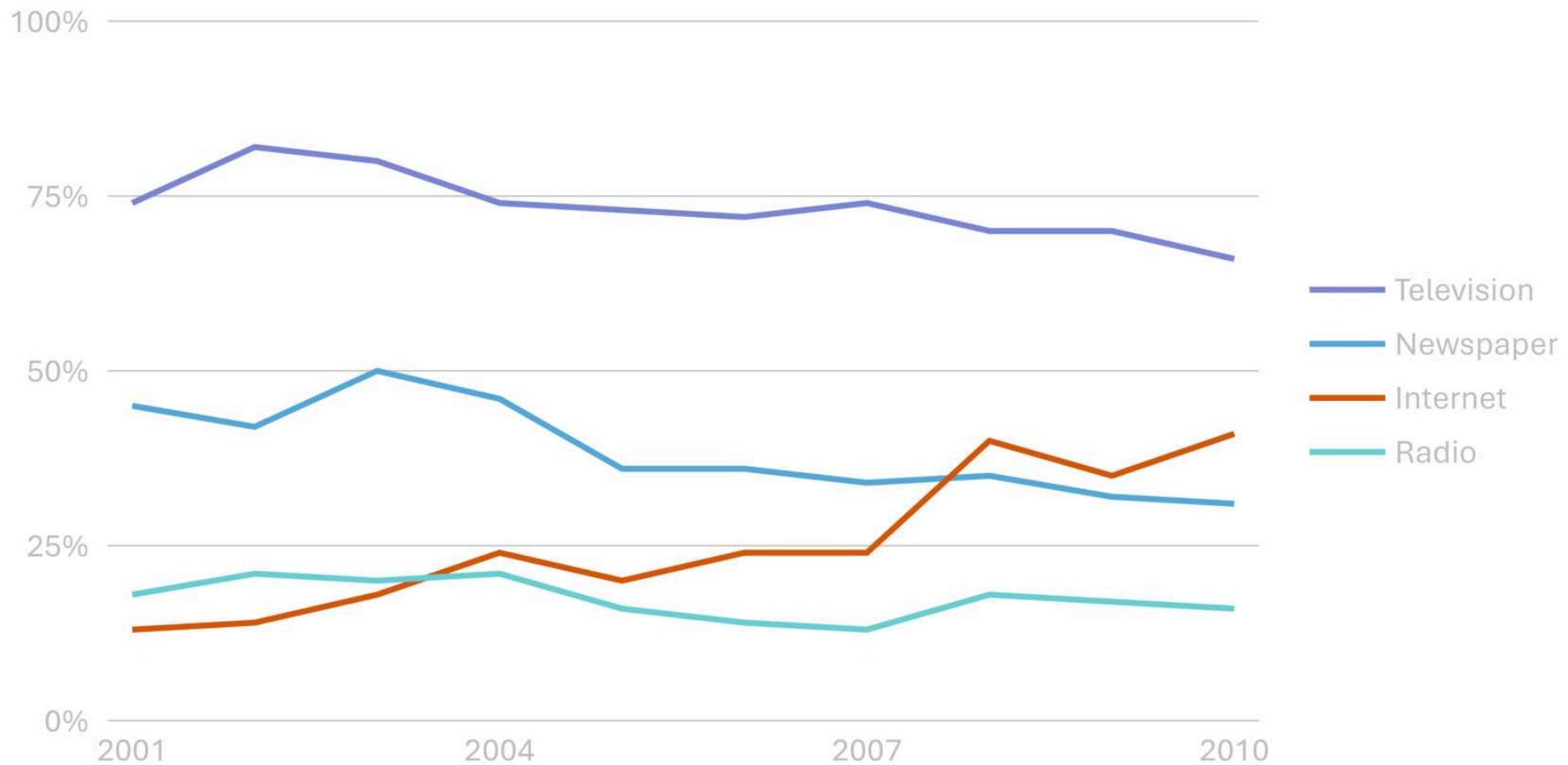


An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



How People Get Their News

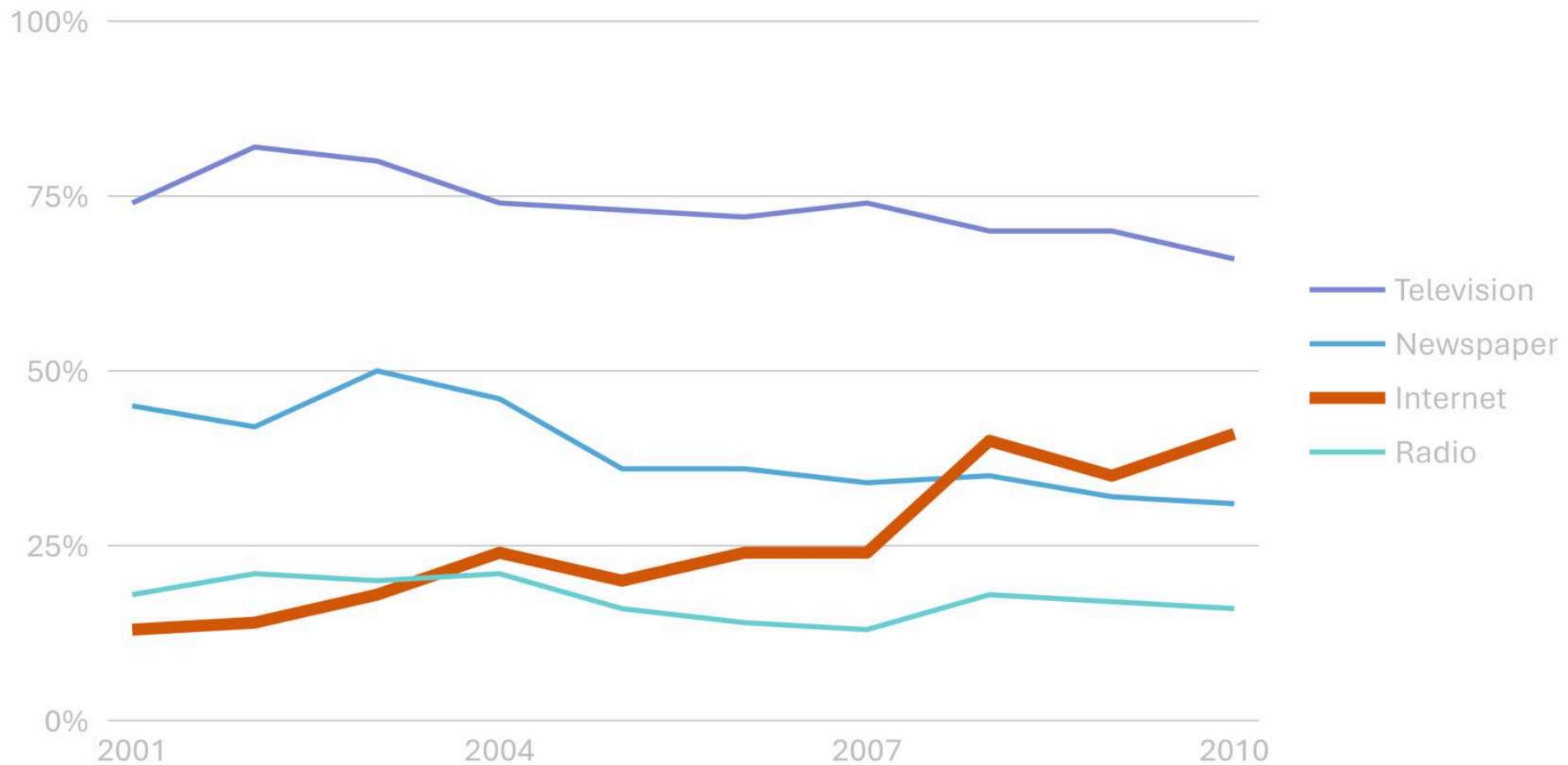


An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



How People Get Their News

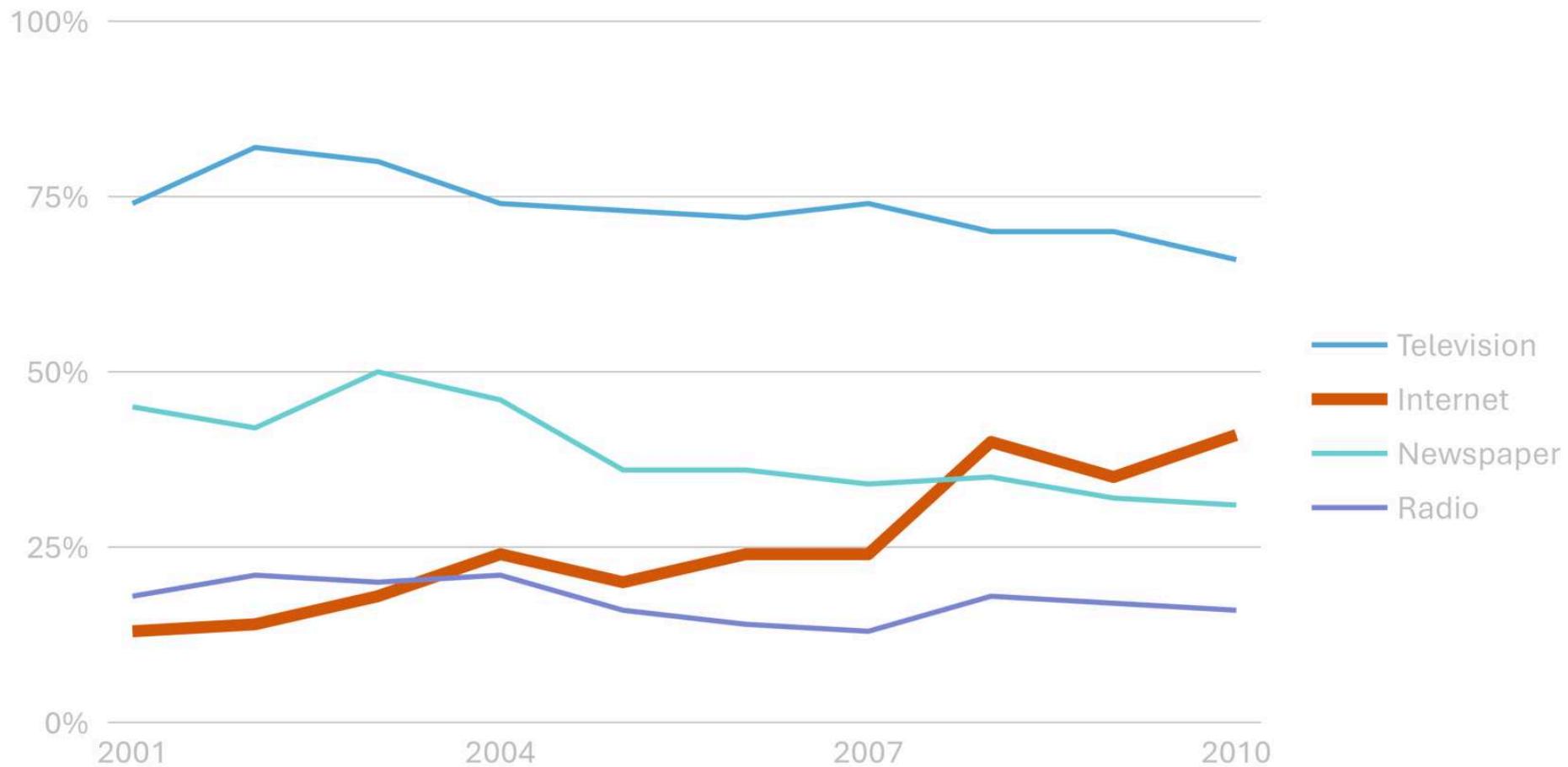


An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



How People Get Their News

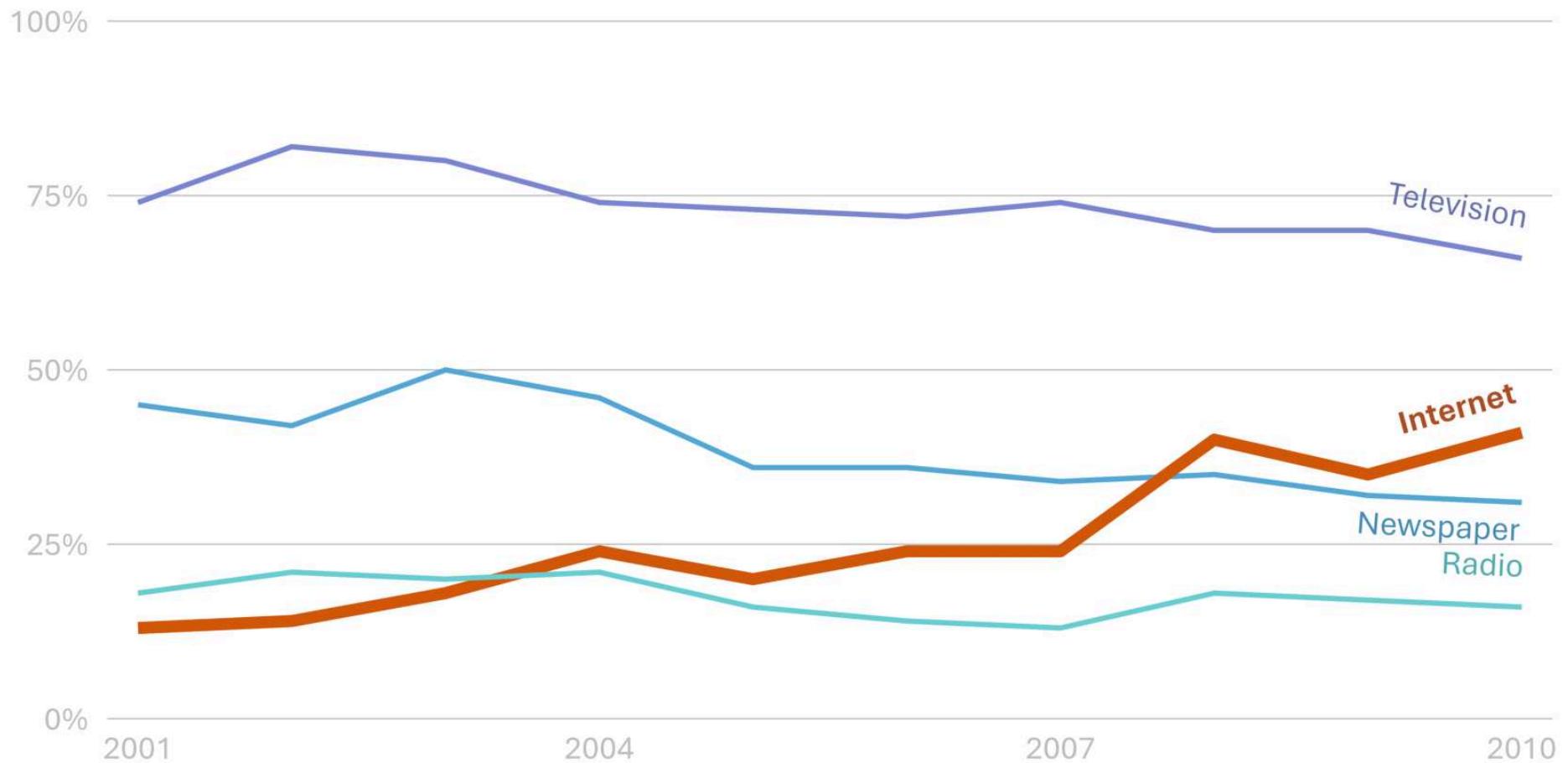


An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



How People Get Their News

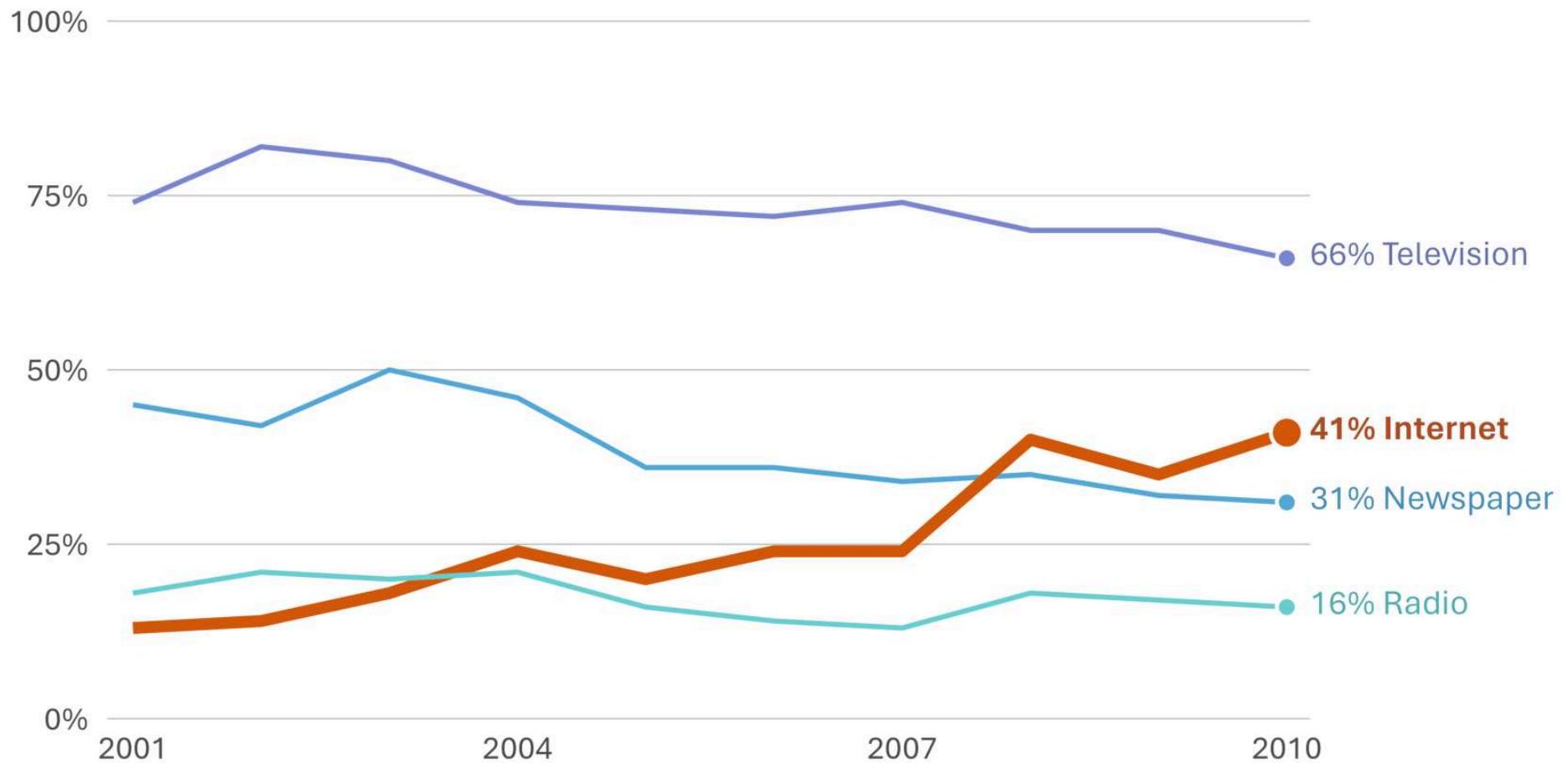


An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



How People Get Their News

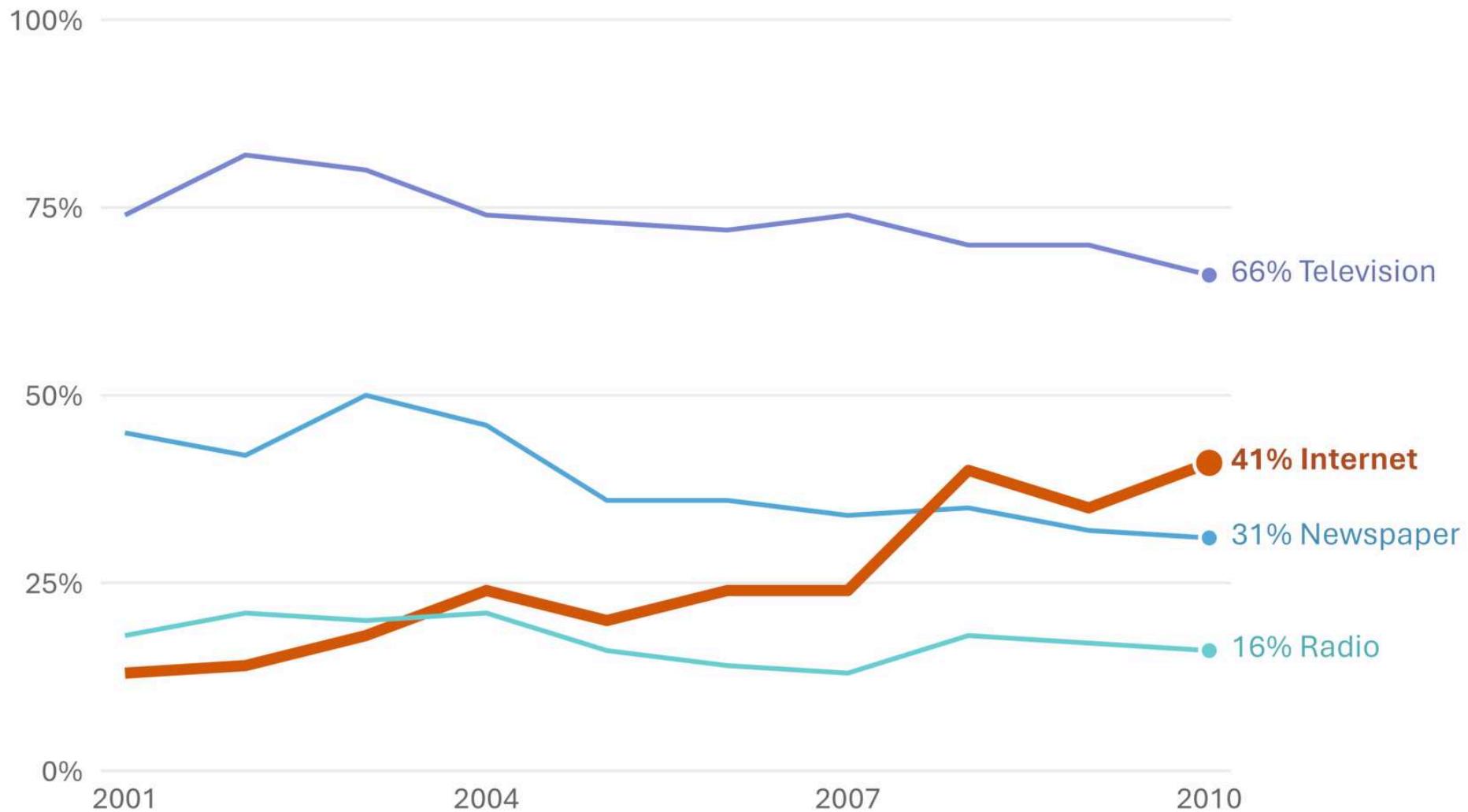


An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



How People Get Their News

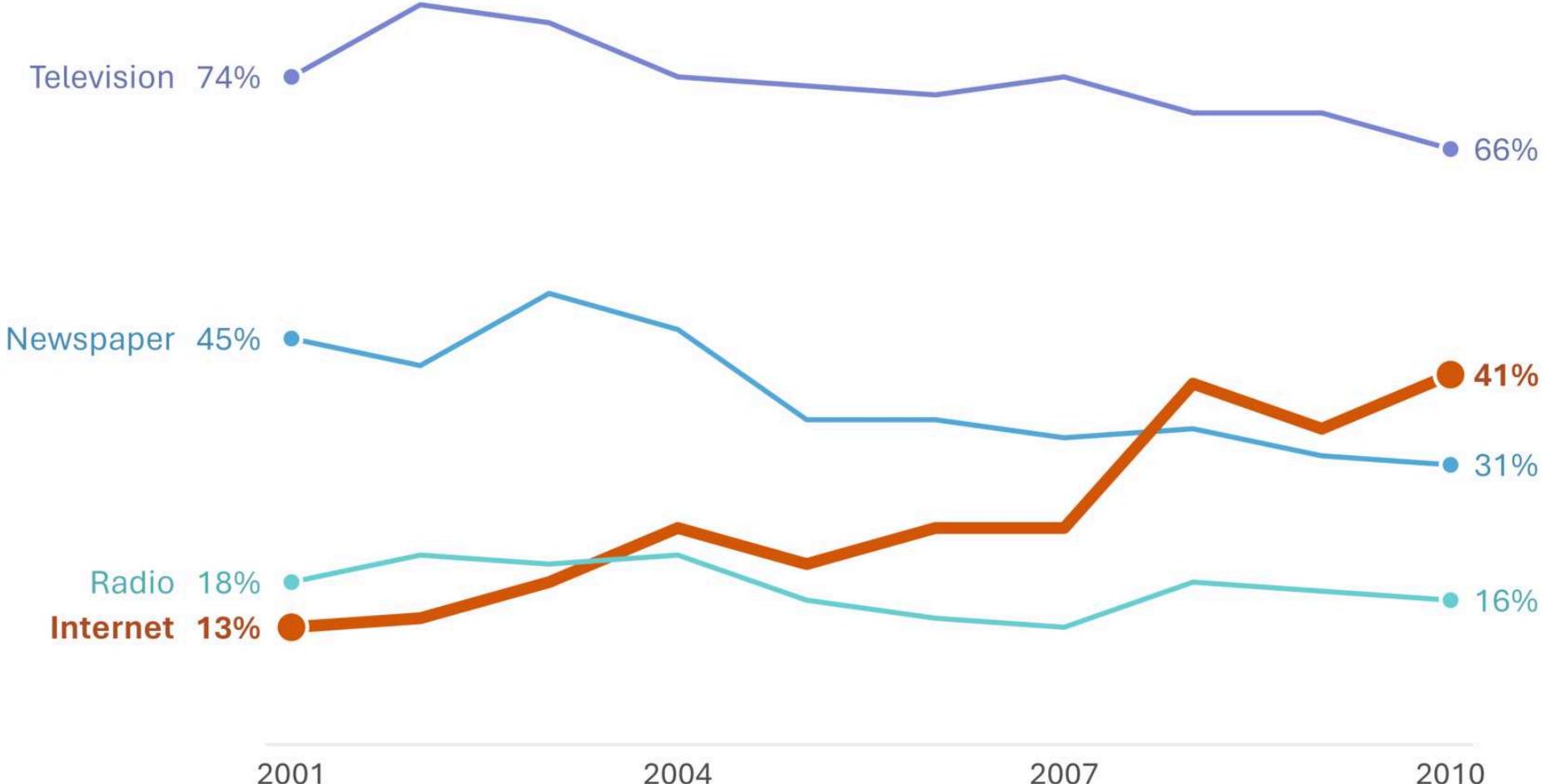


An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



How People Get Their News



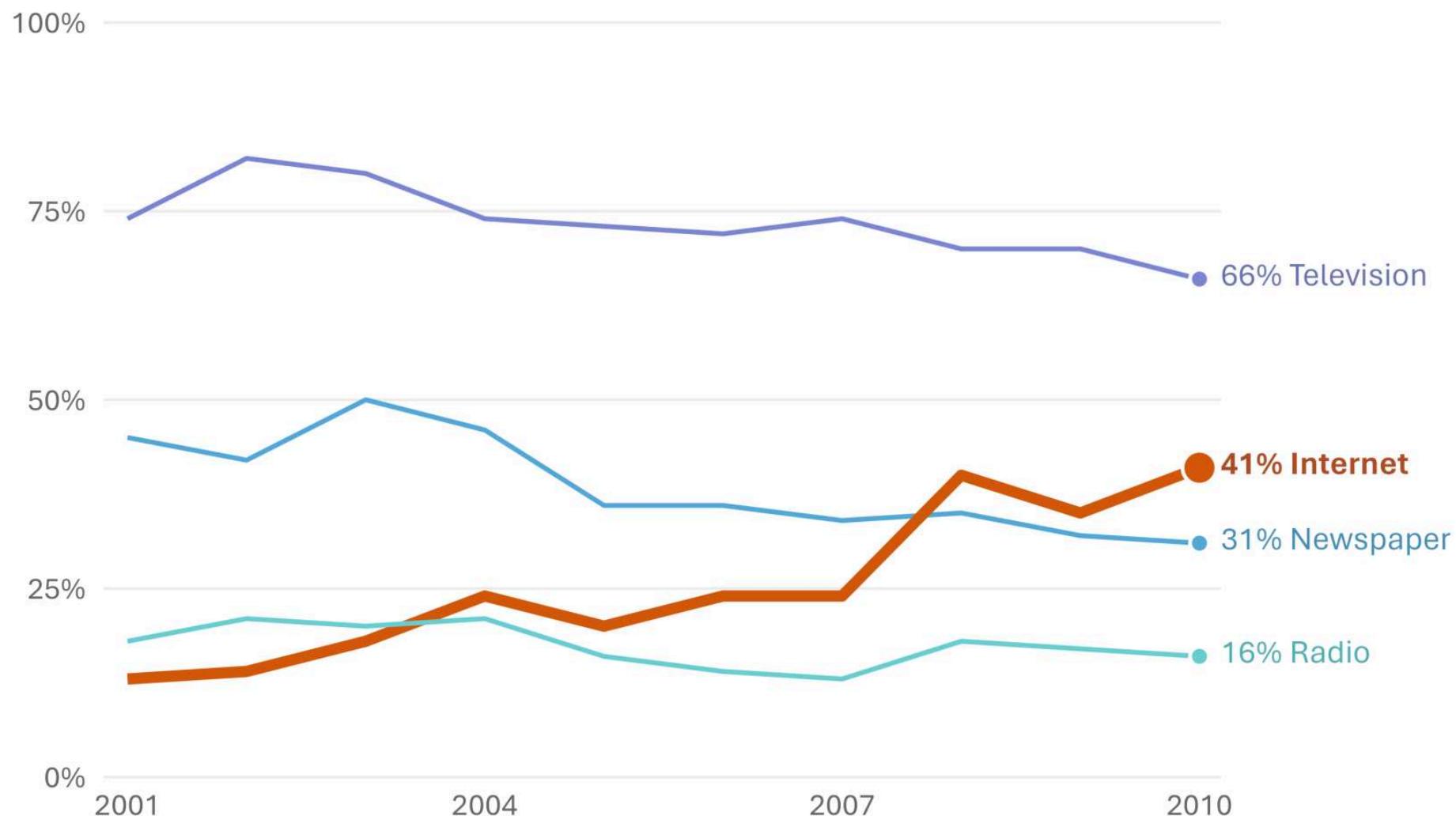
An increasing proportion cite the internet as their primary news source.

Data source: Pew Research Center. Represents responses to the question “where do you get most of your news about national and international issues?” Figures sum to more than 100% because respondents could volunteer up to two main sources.



An increasing proportion cite the internet as their primary news source.

Responses to the question “where do you get most of your news about national and international issues?”



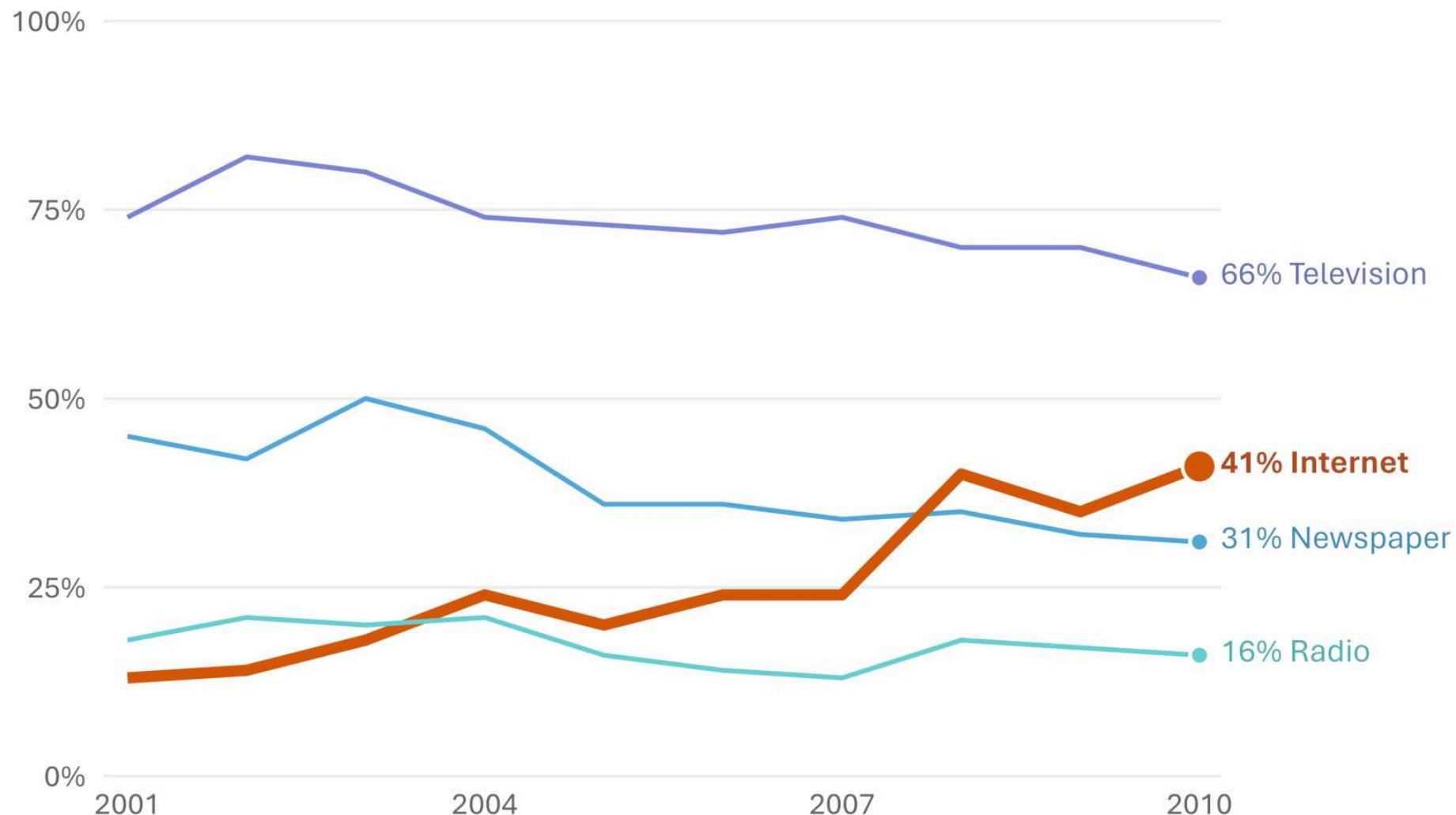
Data source: Pew Research Center

Figures sum to more than 100% because respondents could volunteer up to two main sources.



An increasing proportion cite the **internet** as their primary news source.

Responses to the question “where do you get most of your news about national and international issues?”



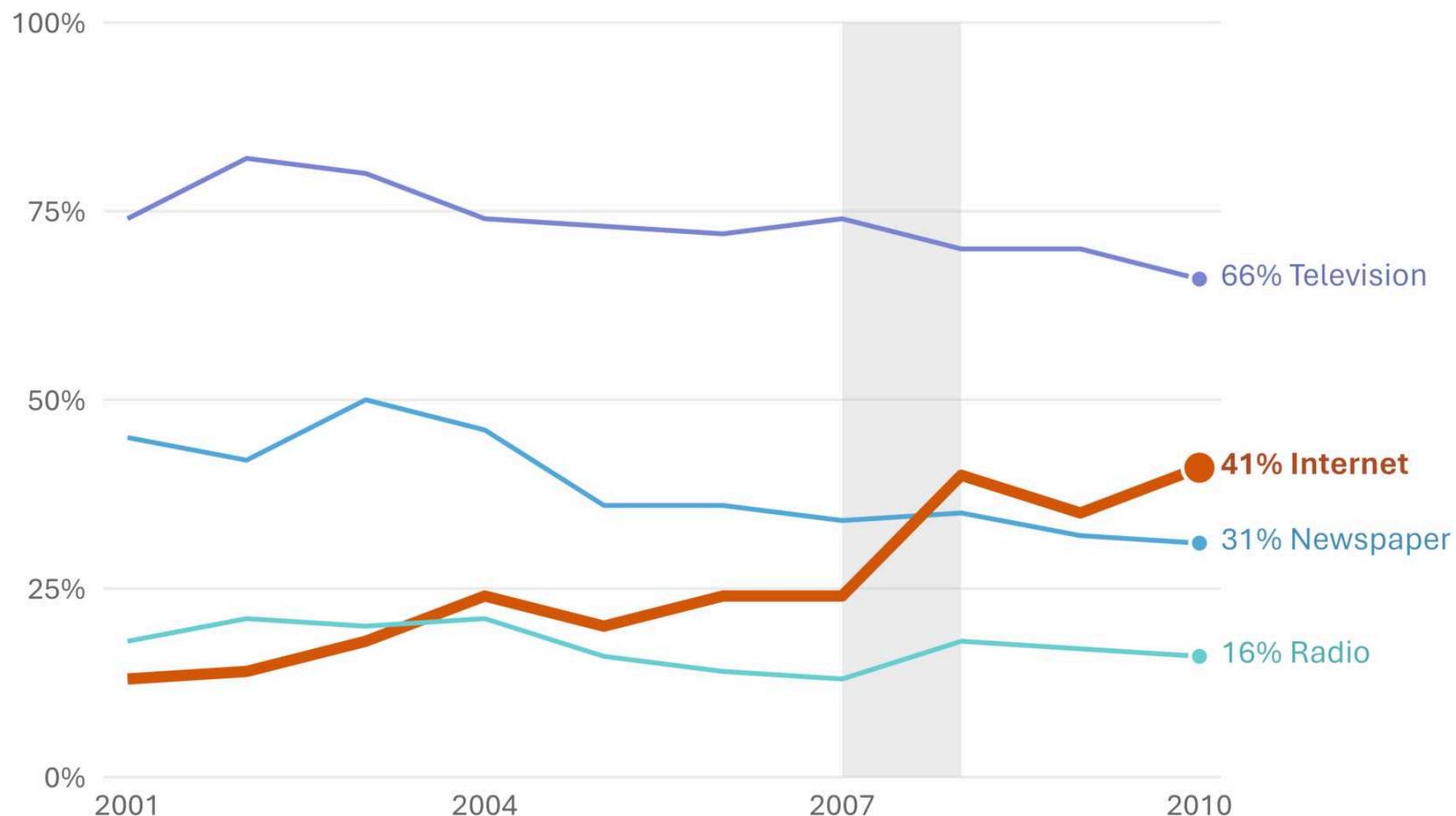
Data source: Pew Research Center

Figures sum to more than 100% because respondents could volunteer up to two main sources.



An increasing proportion cite the **internet** as their primary news source.

Responses to the question “where do you get most of your news about national and international issues?”



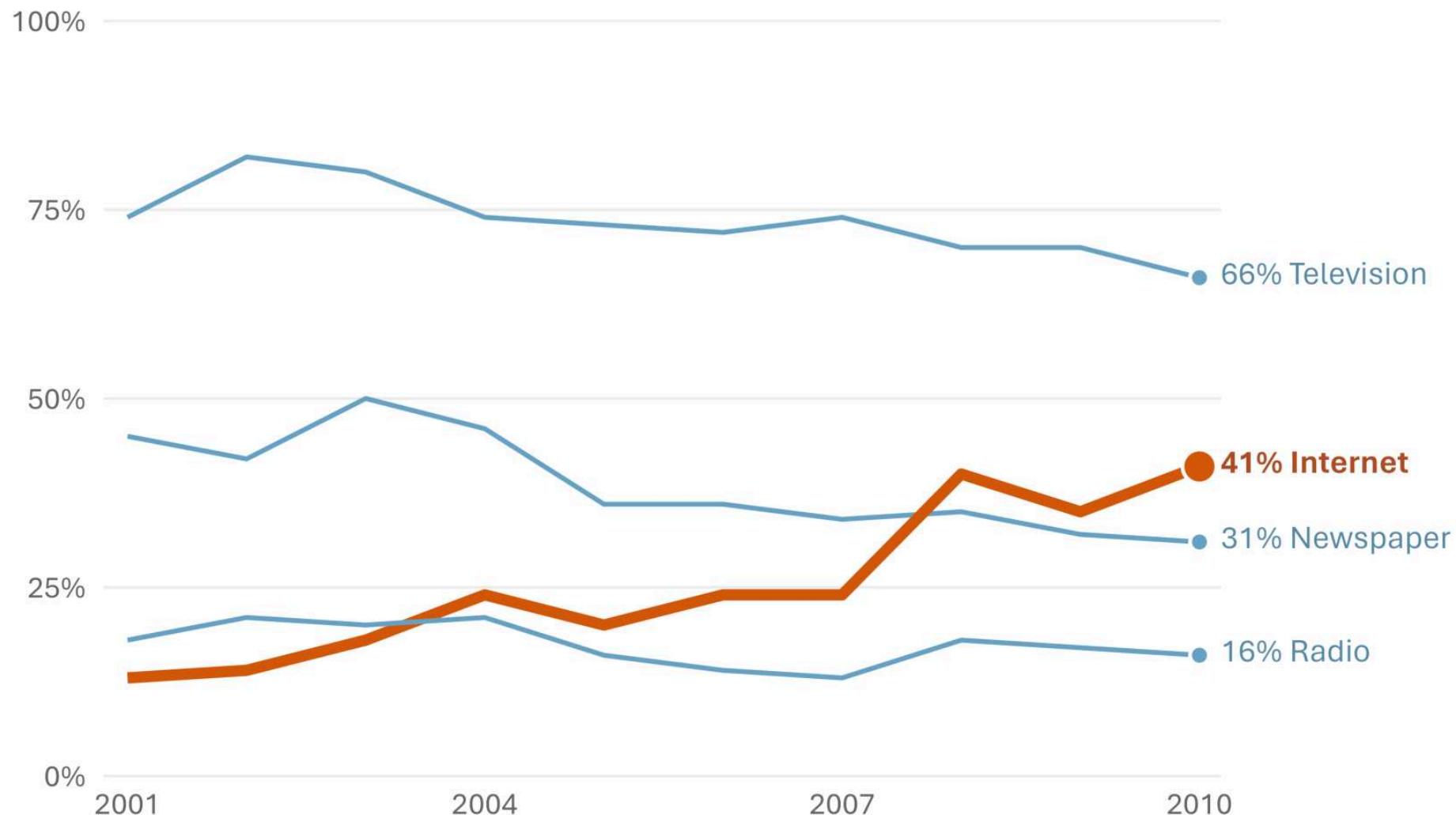
Data source: Pew Research Center

Figures sum to more than 100% because respondents could volunteer up to two main sources.



An increasing proportion cite the **internet** as their primary news source.

Responses to the question “where do you get most of your news about national and international issues?”



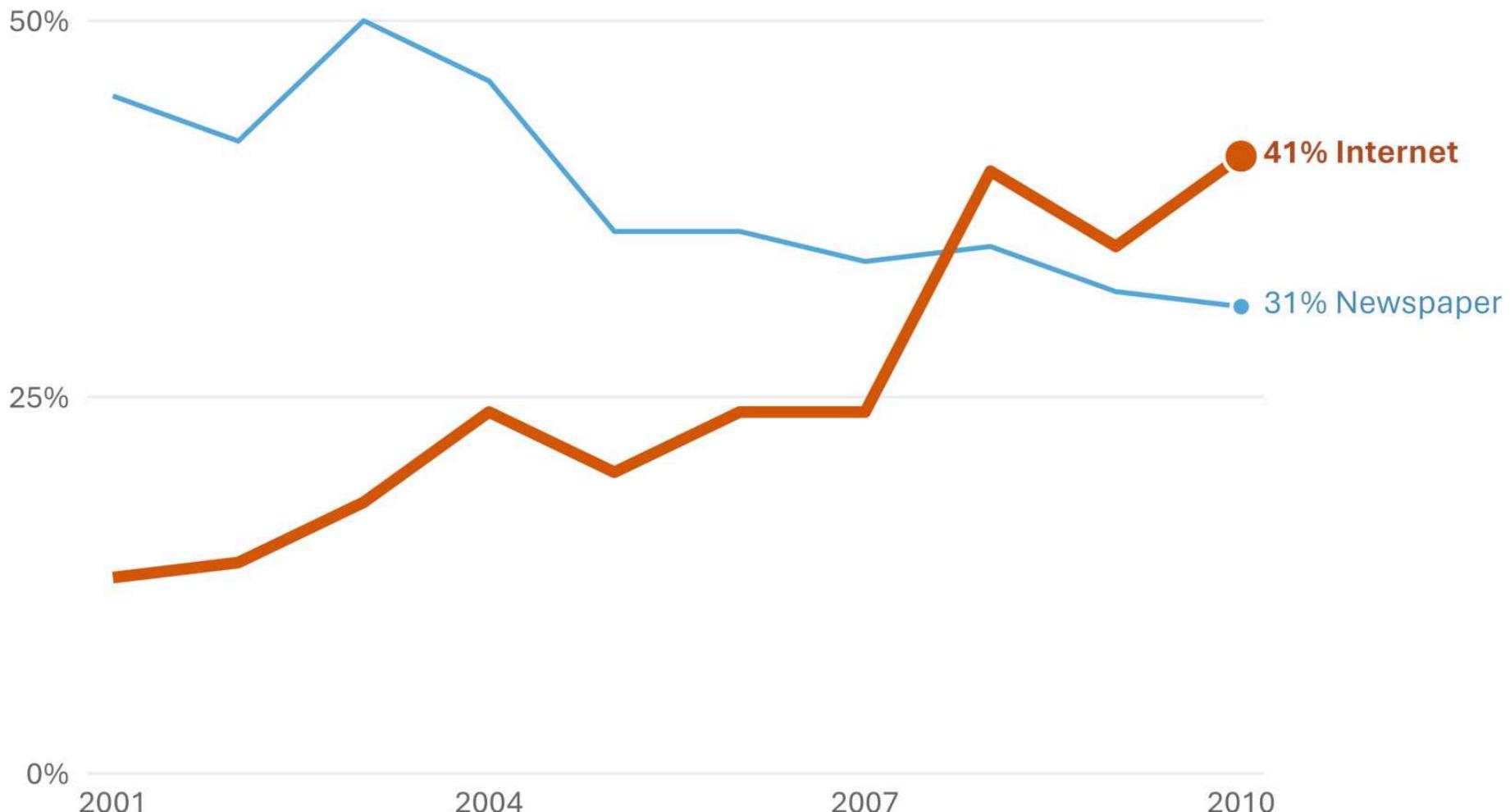
Data source: Pew Research Center

Figures sum to more than 100% because respondents could volunteer up to two main sources.



Internet has surpassed newspapers as the main source of news.

Responses to the question “where do you get most of your news about national and international issues?”



Data source: Pew Research Center

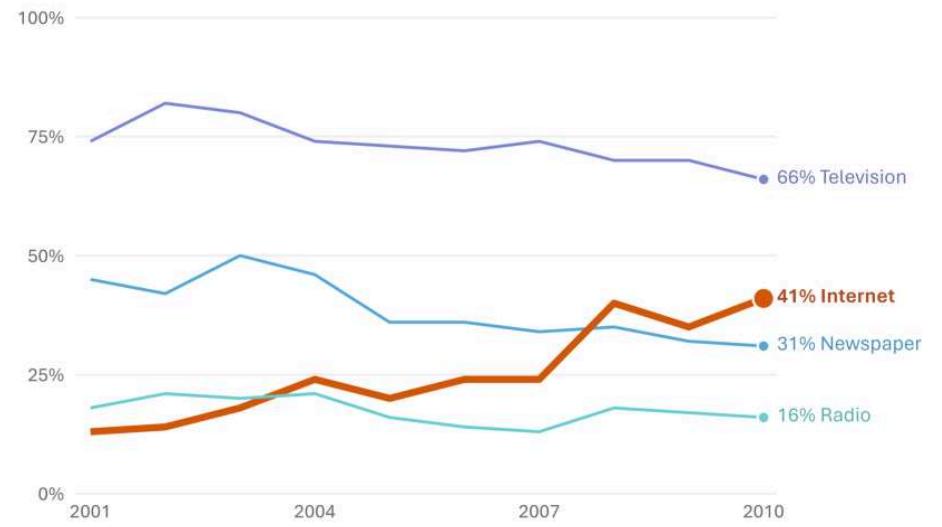
Figures sum to more than 100% because respondents could volunteer up to two main sources.





An increasing proportion cite the **internet** as their primary news source.

Responses to the question "where do you get most of your news about national and international issues?"



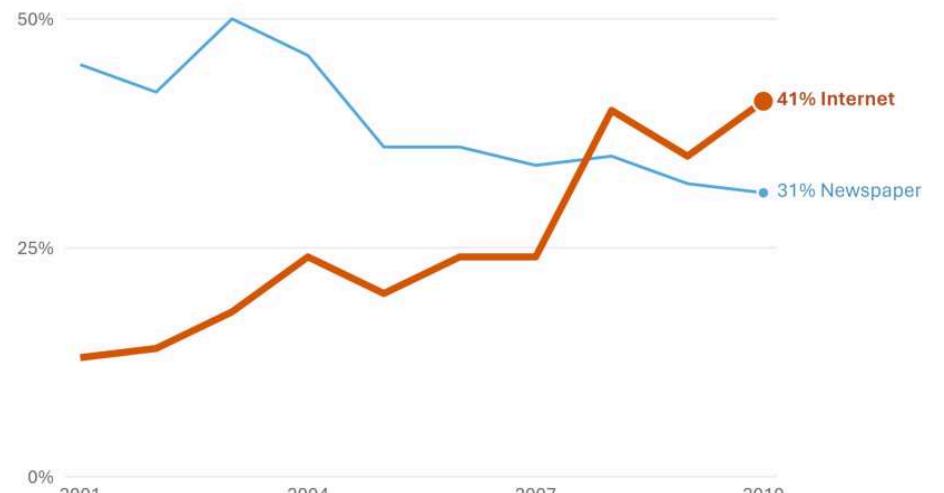
Data source: Pew Research Center
Figures sum to more than 100% because respondents could volunteer up to two main sources.





Internet has surpassed newspapers as the main source of news.

Responses to the question "where do you get most of your news about national and international issues?"

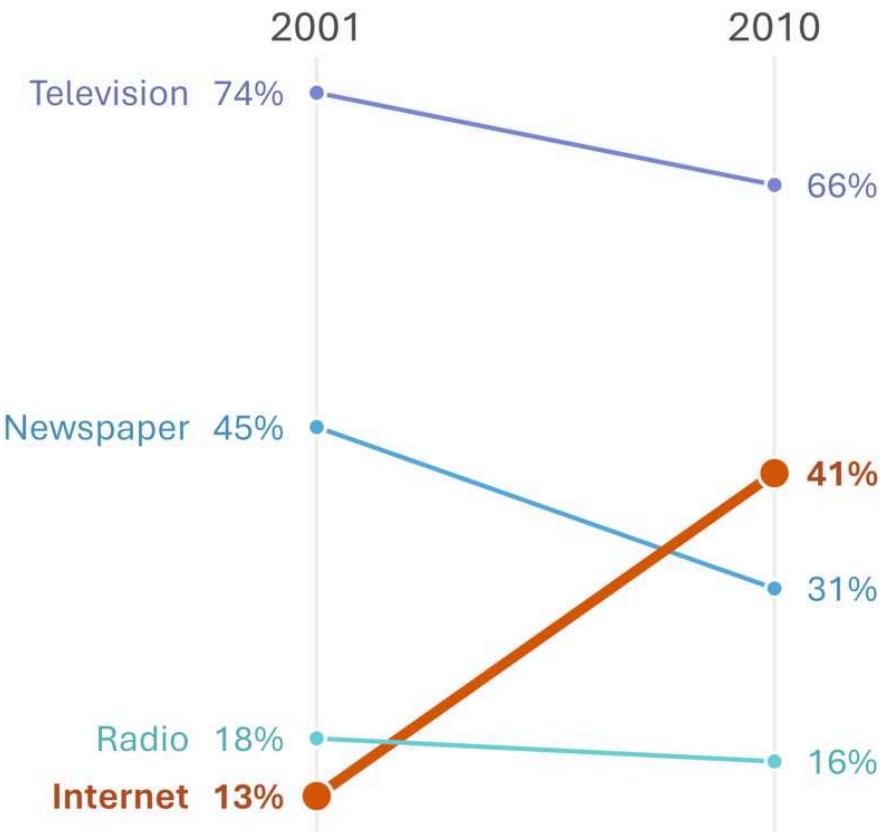


Data source: Pew Research Center
Figures sum to more than 100% because respondents could volunteer up to two main sources.



An increasing proportion cite the internet as their primary news source.

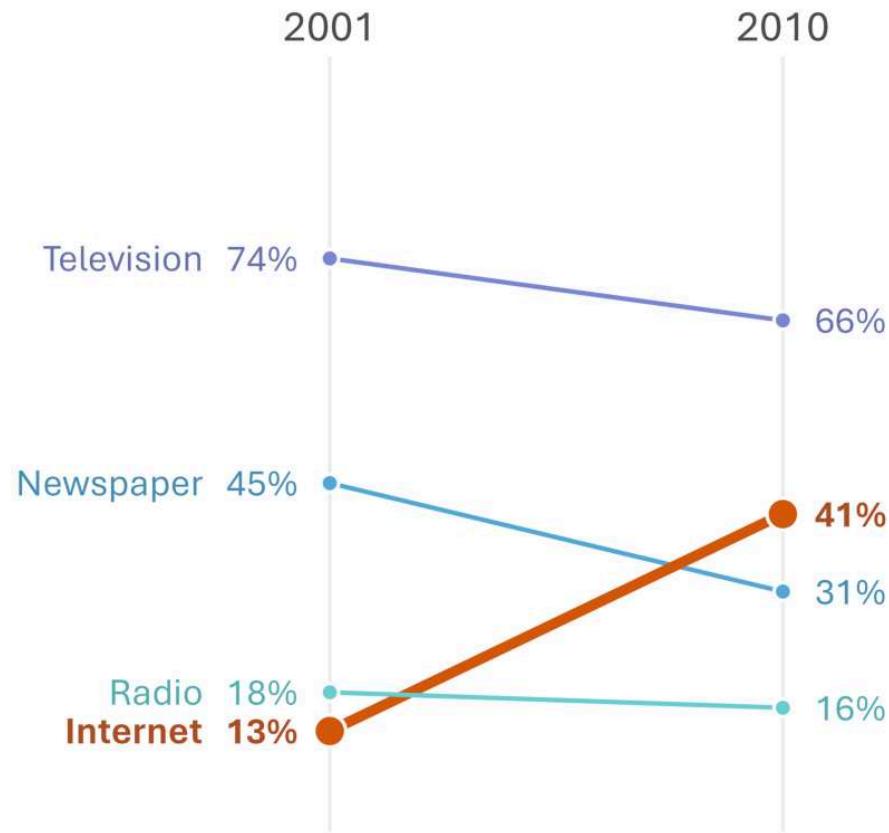
Responses to the question “where do you get most of your news about national and international issues?”



An increasing proportion cite the internet as their primary news source.
Data source: Pew Research Center. Represents responses to the question
“where do
you get most of your news about national and international issues?” Figures sum
to
more than 100% because respondents could volunteer up to two main sources.

An increasing proportion cite the internet as their primary news source.

Responses to the question “where do you get most of your news about national and international issues?”



An increasing proportion cite the internet as their primary news source.
Data source: Pew Research Center. Represents responses to the question
“where do
you get most of your news about national and international issues?” Figures sum
to
more than 100% because respondents could volunteer up to two main sources.

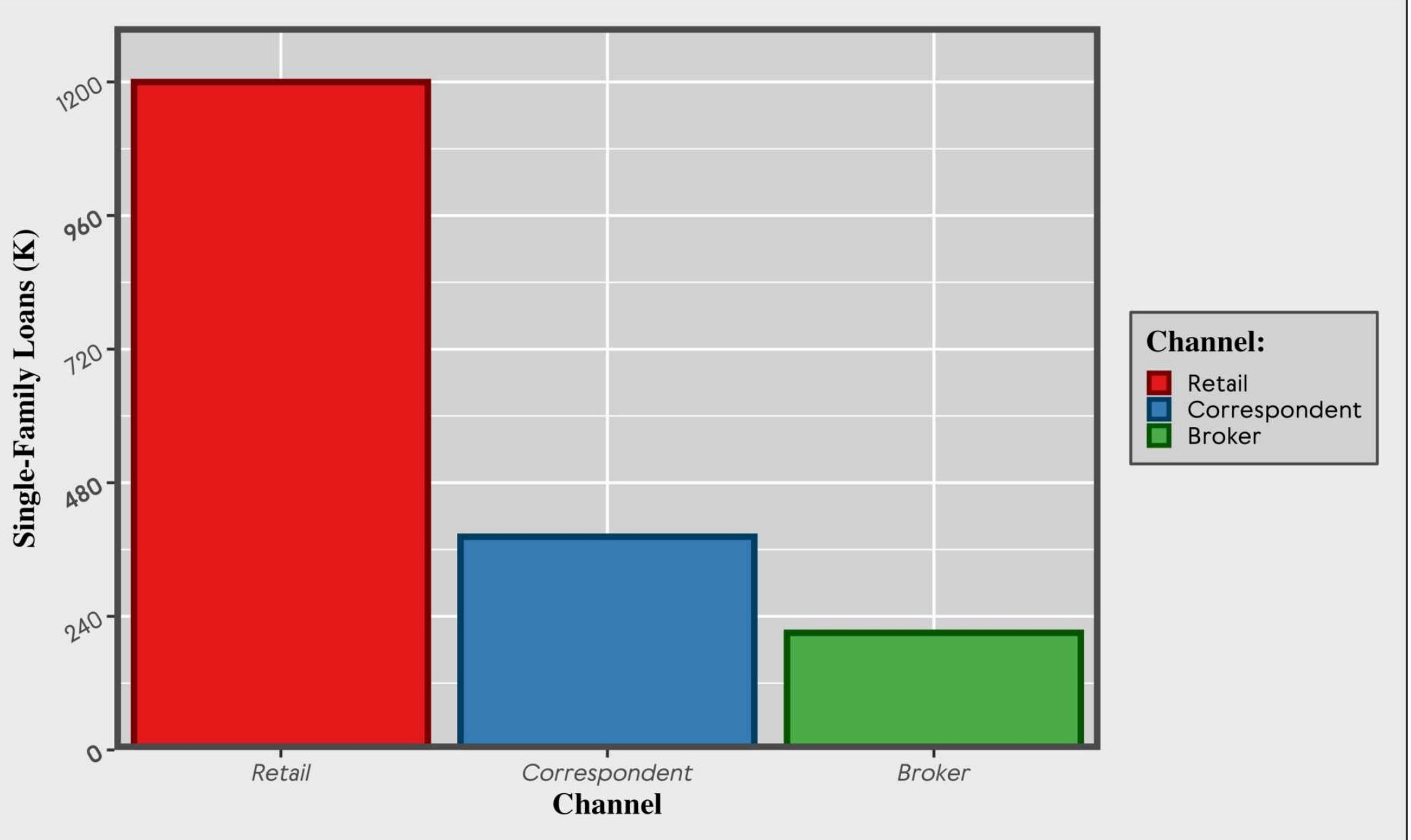


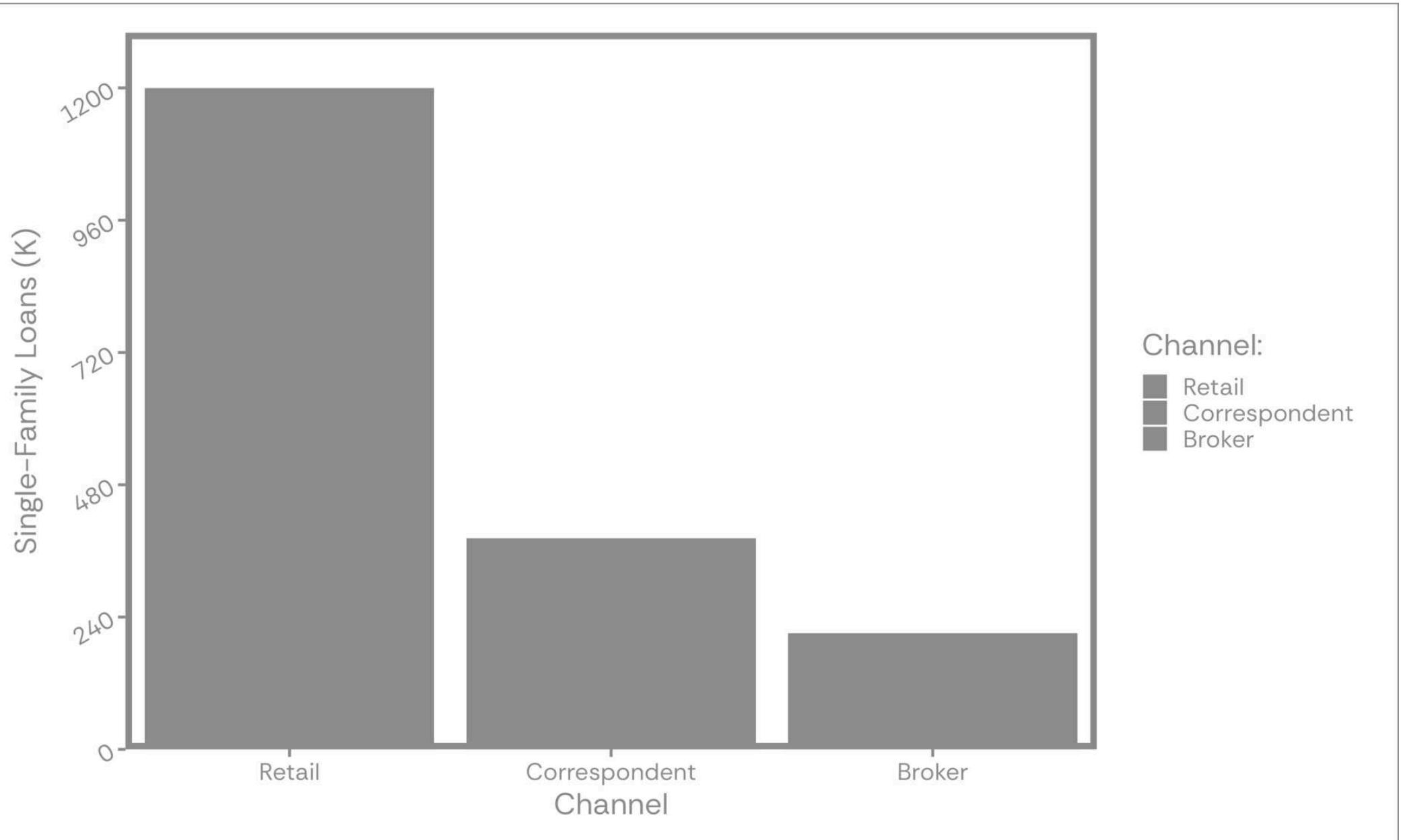


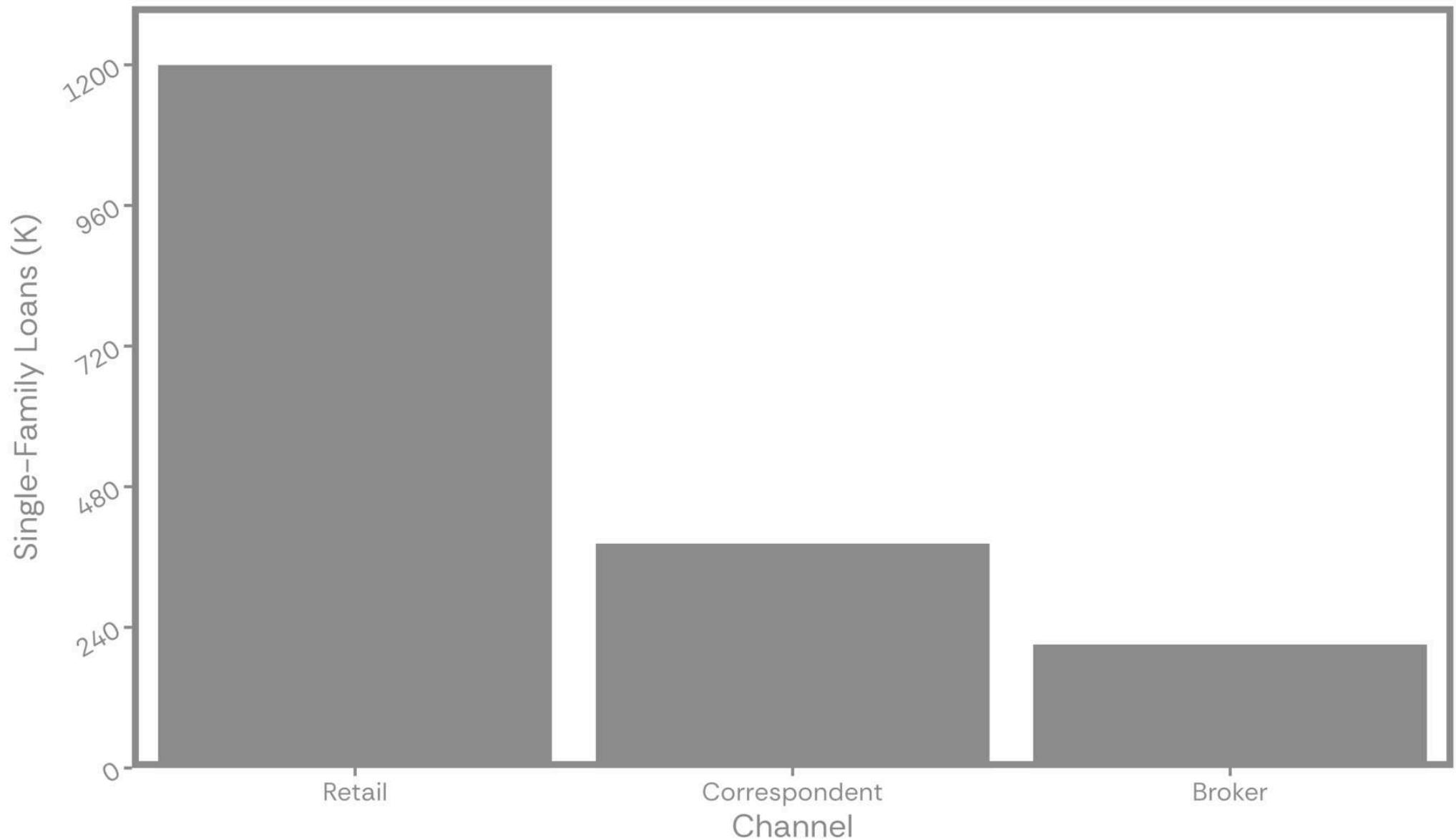
Exercise 3

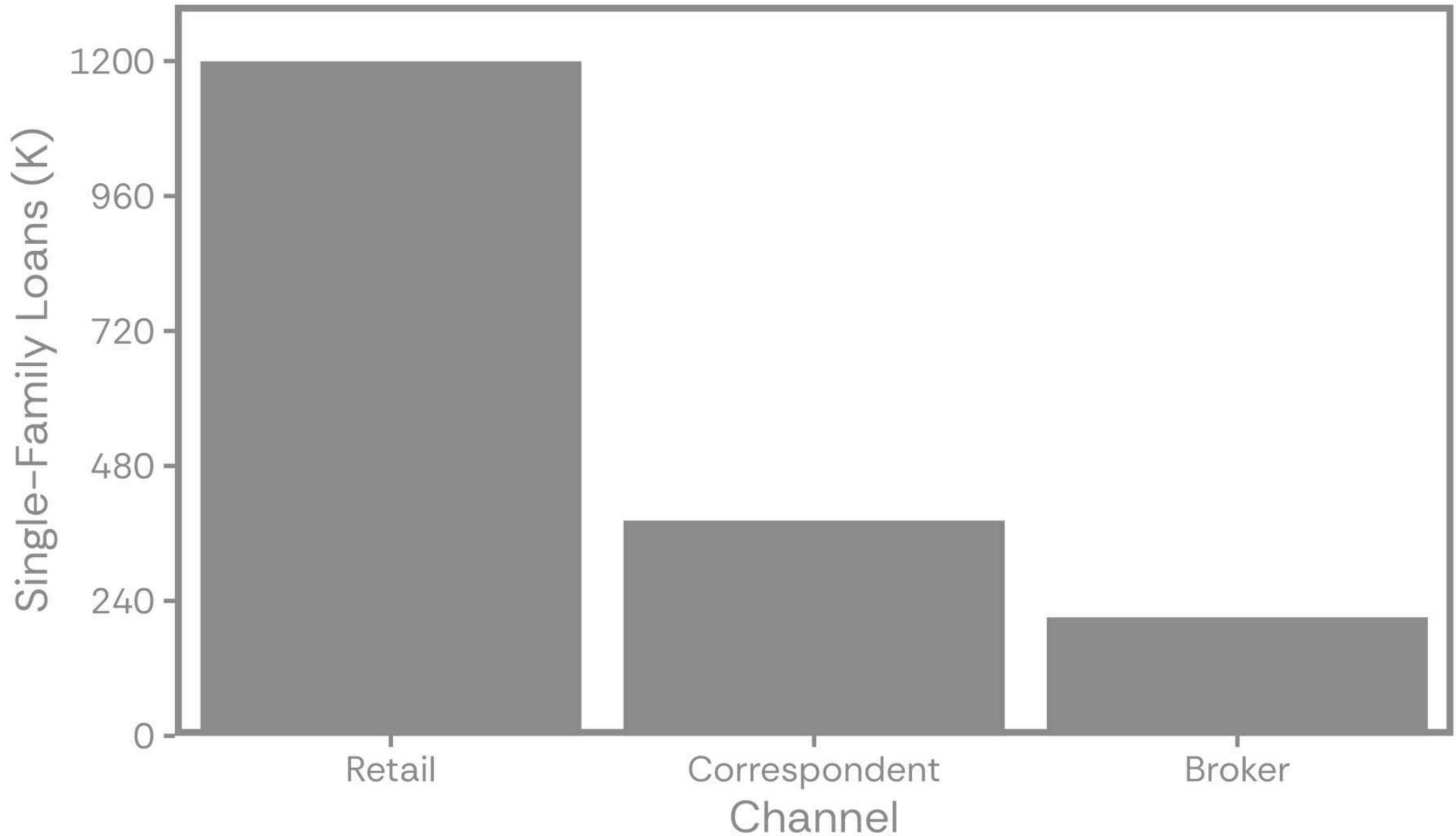
Suggested Solutions

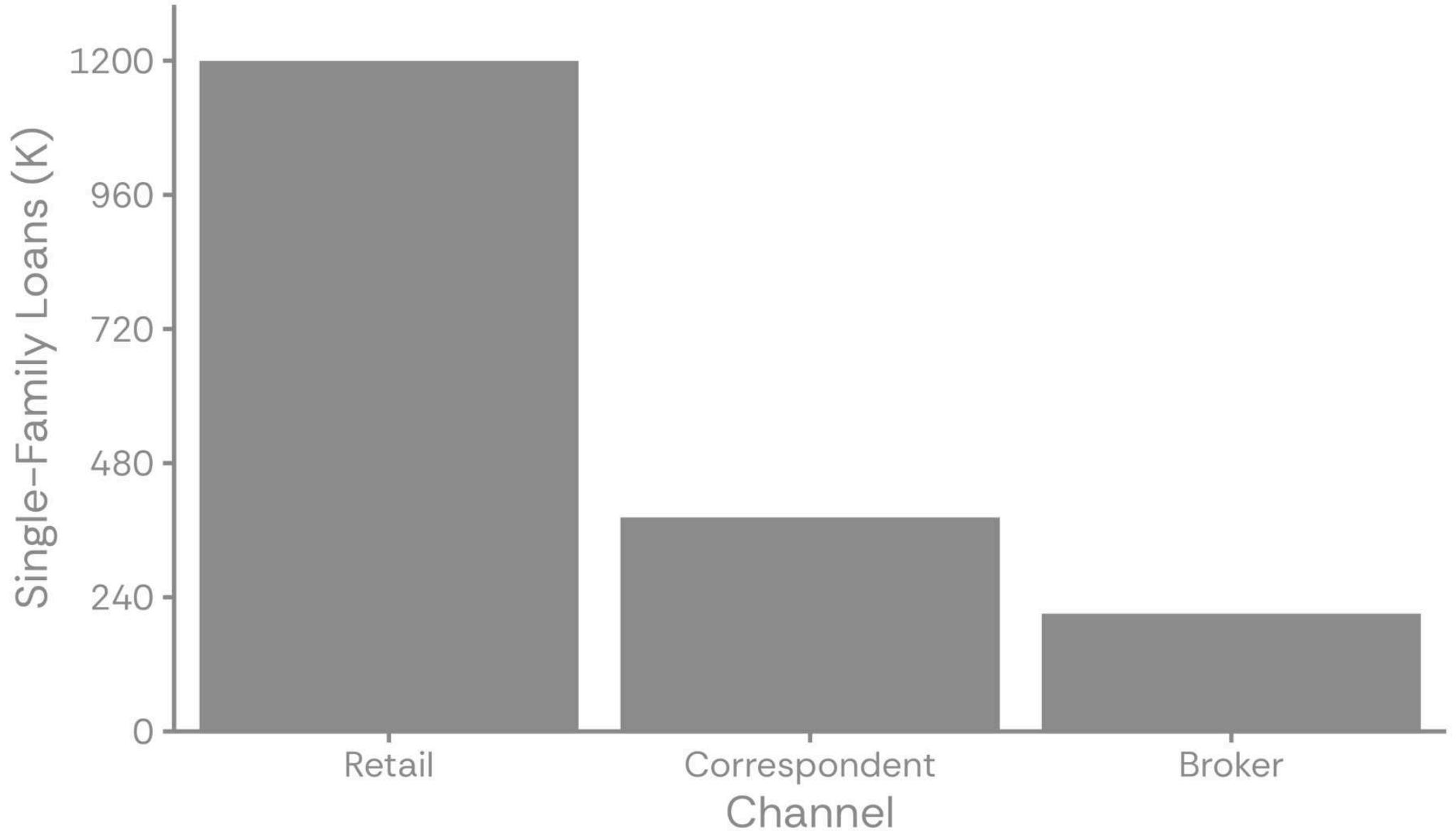


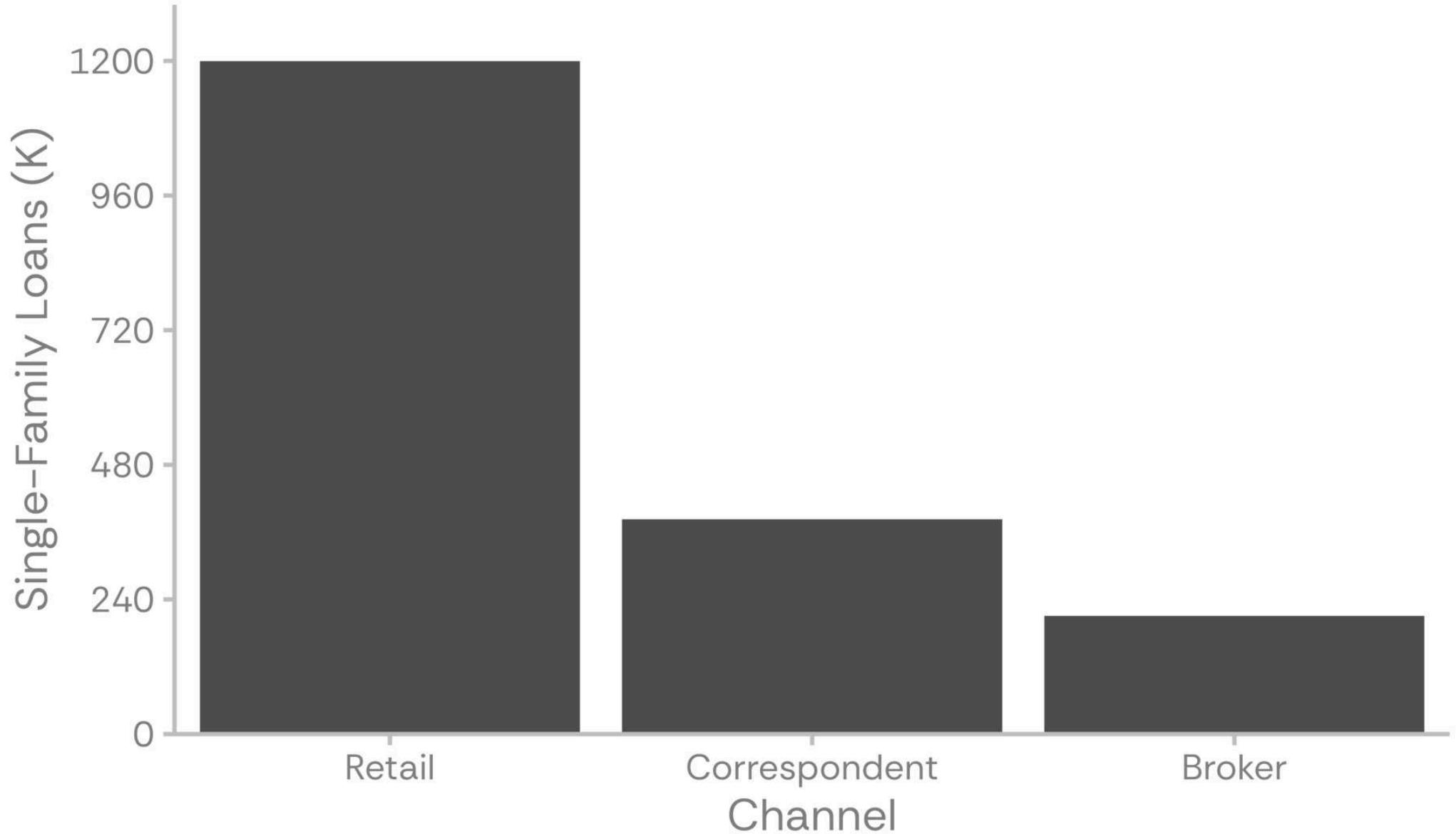


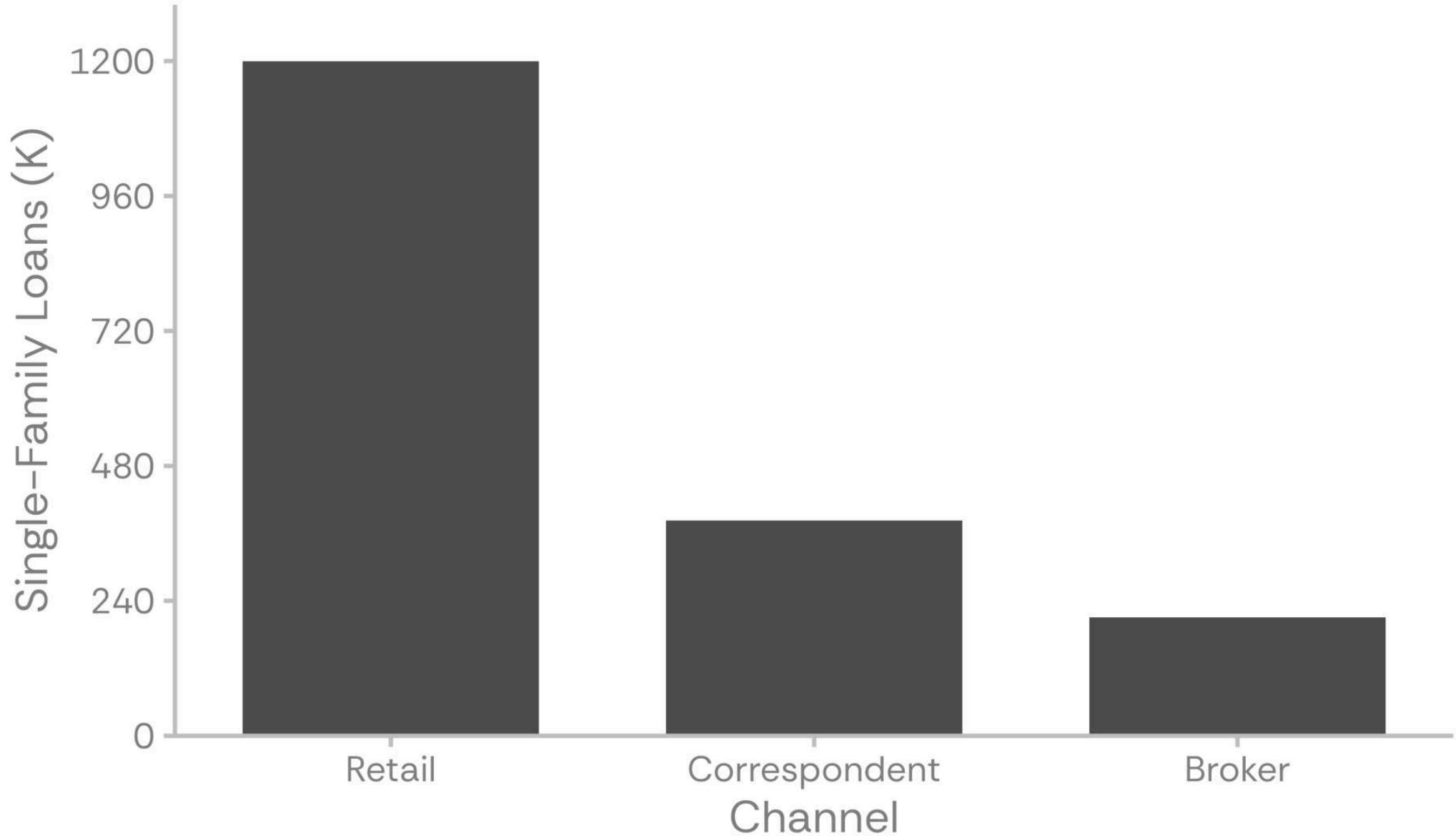


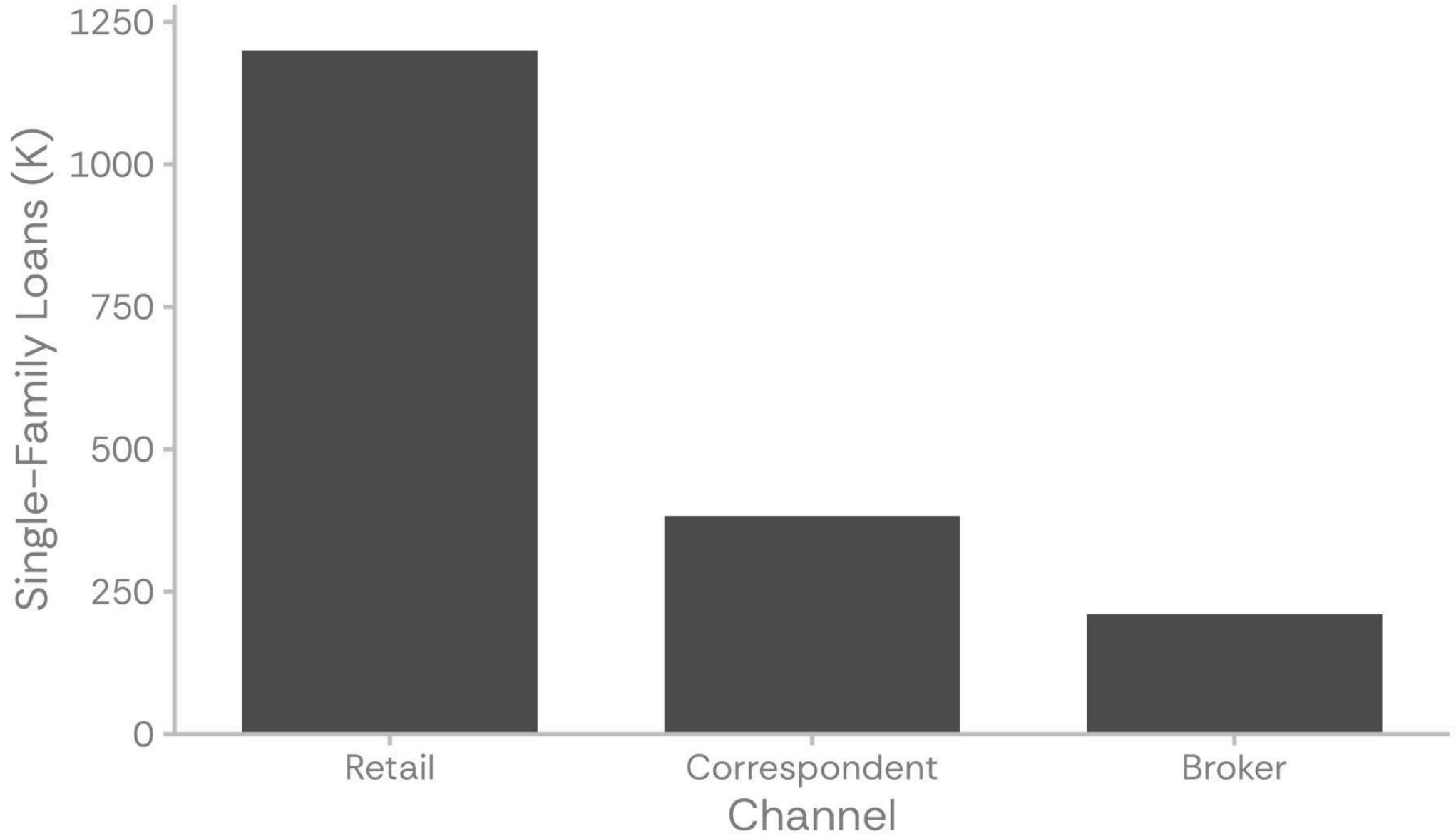


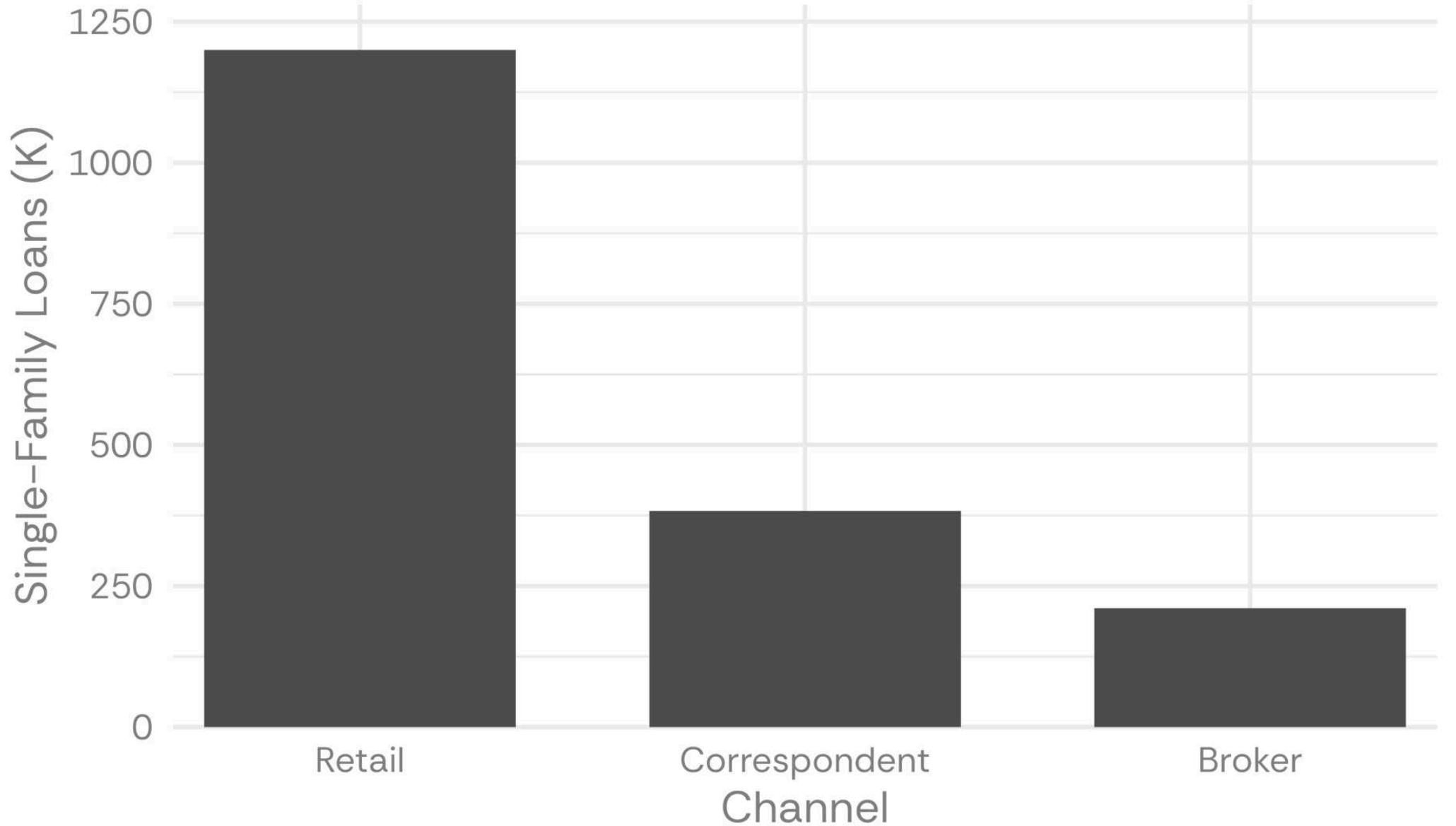


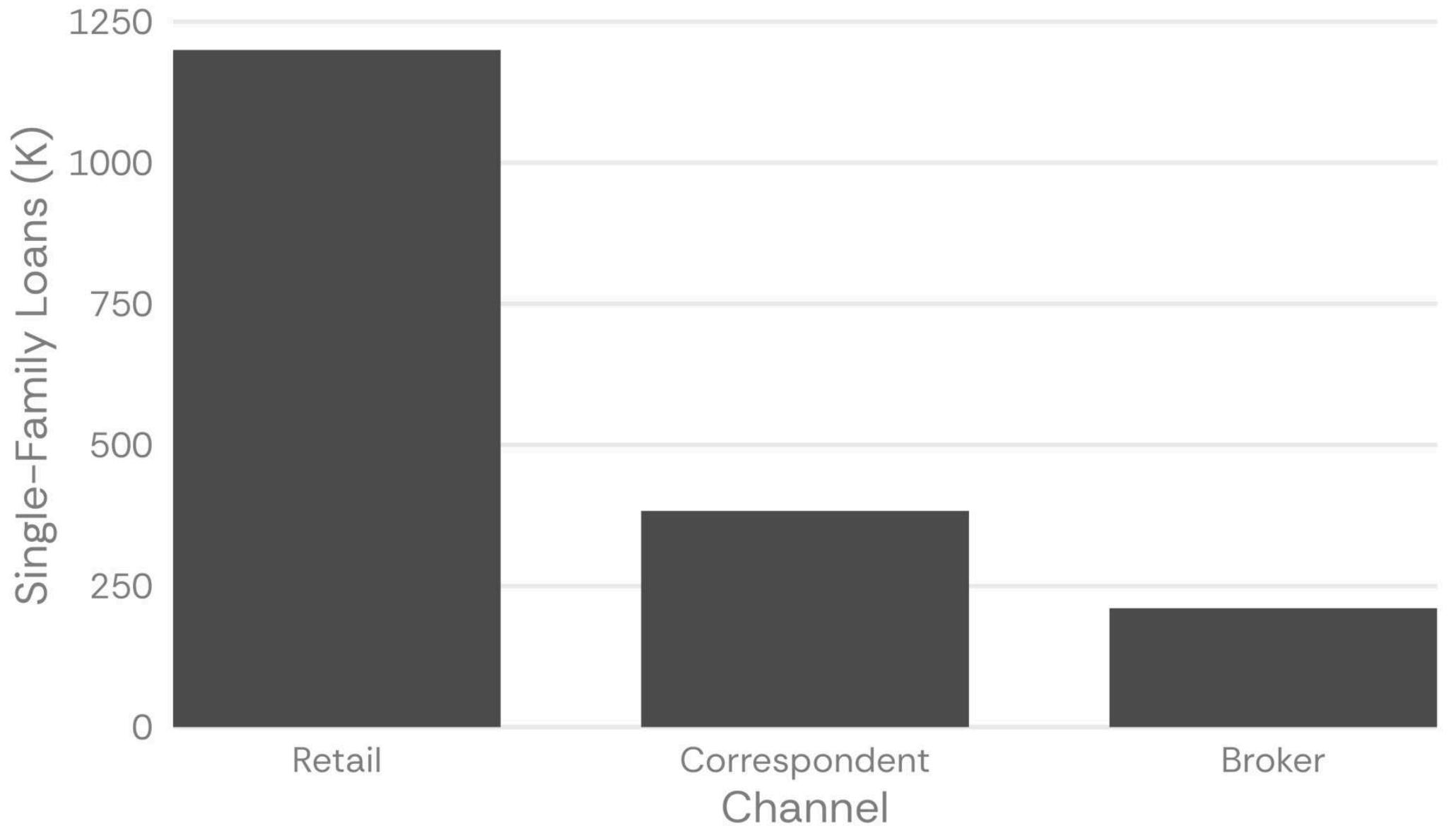


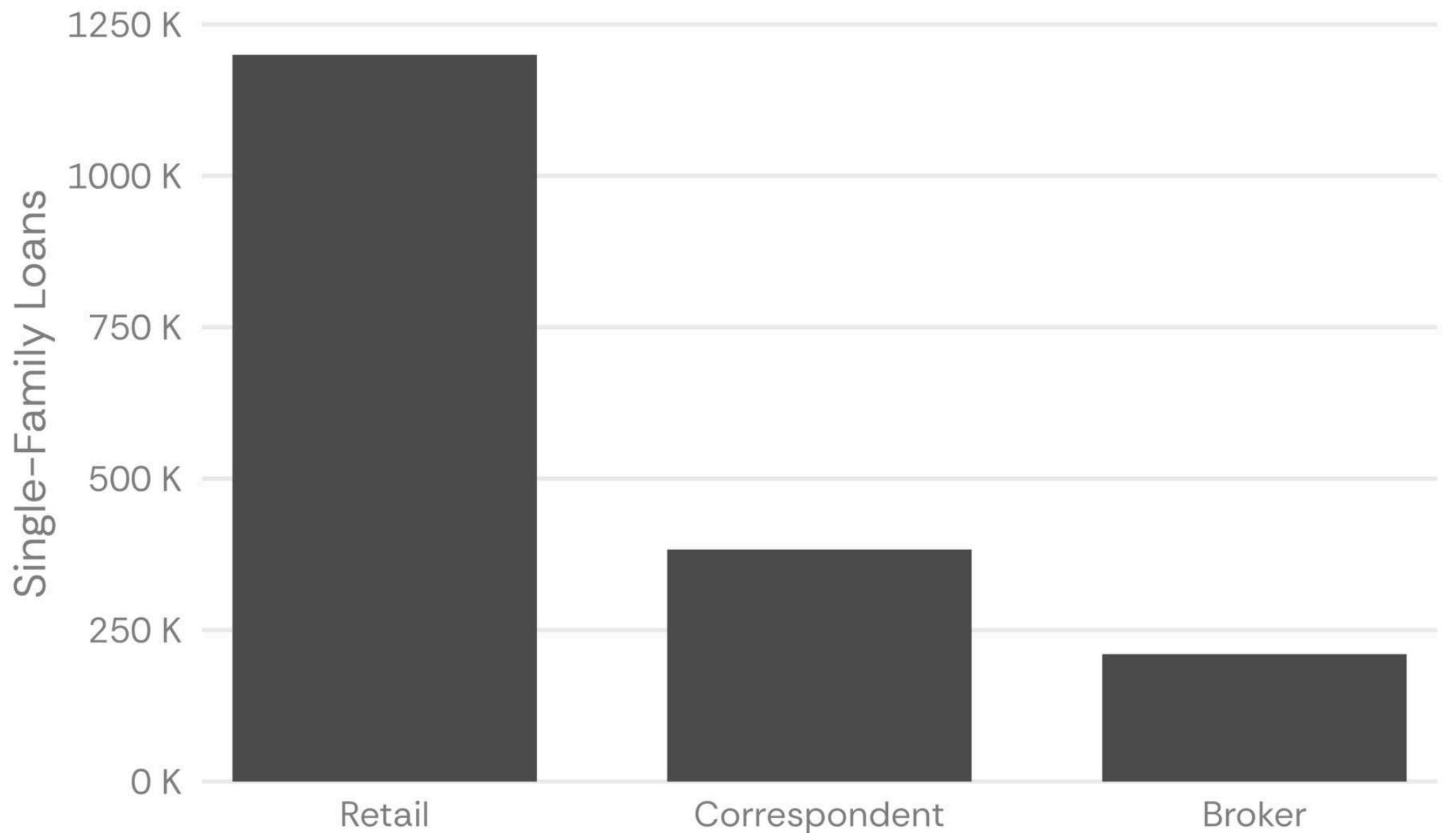




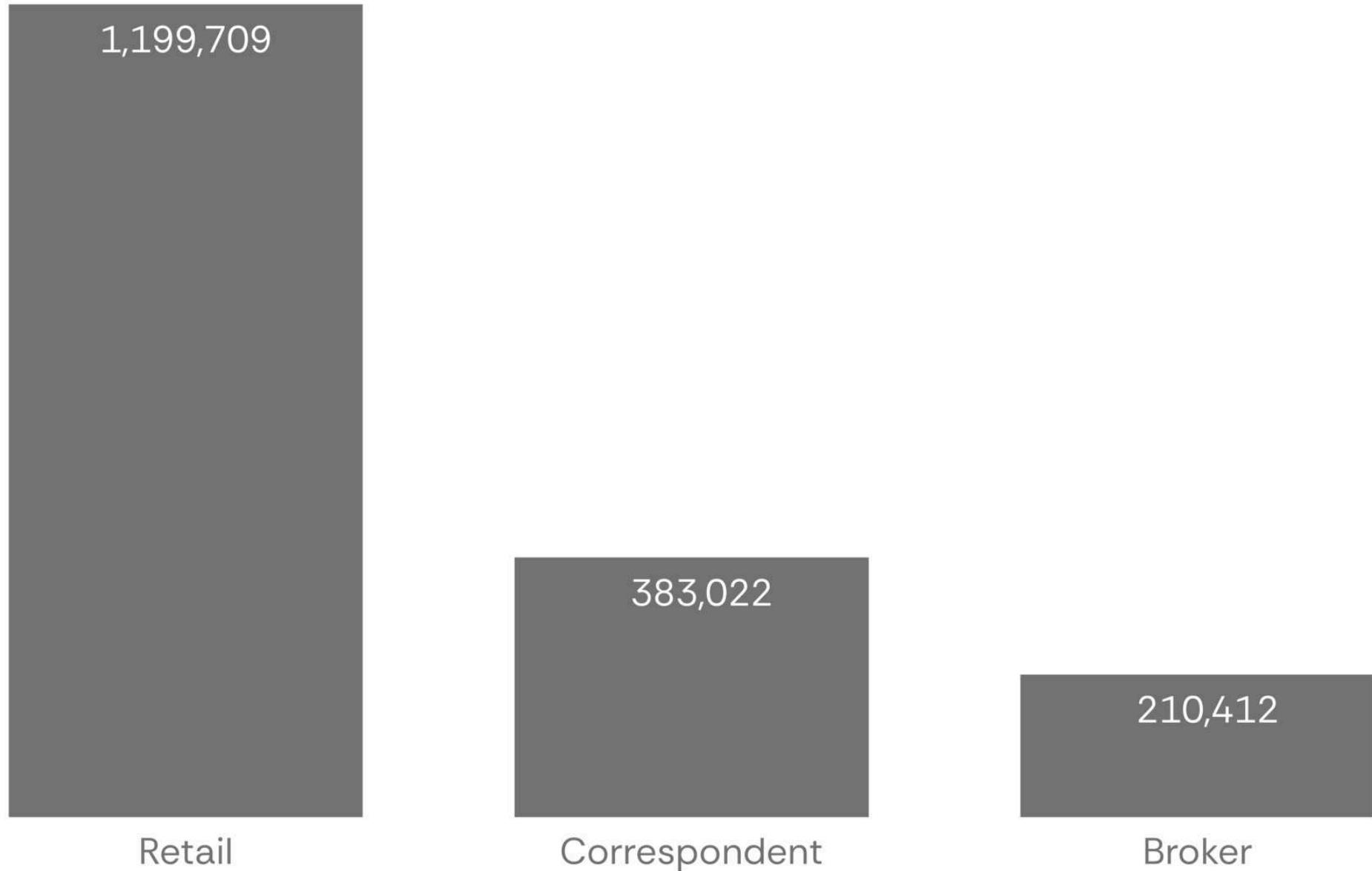




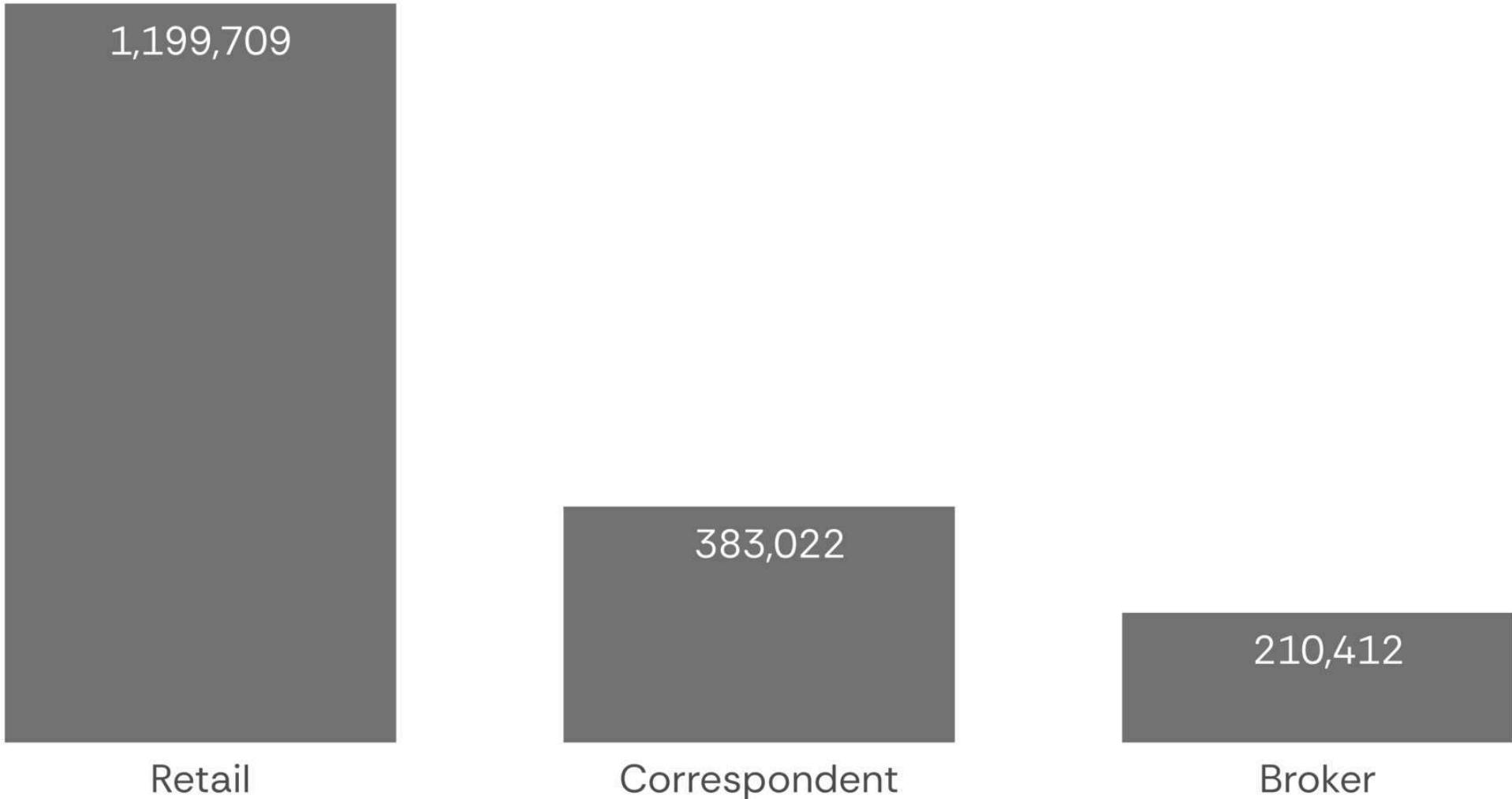




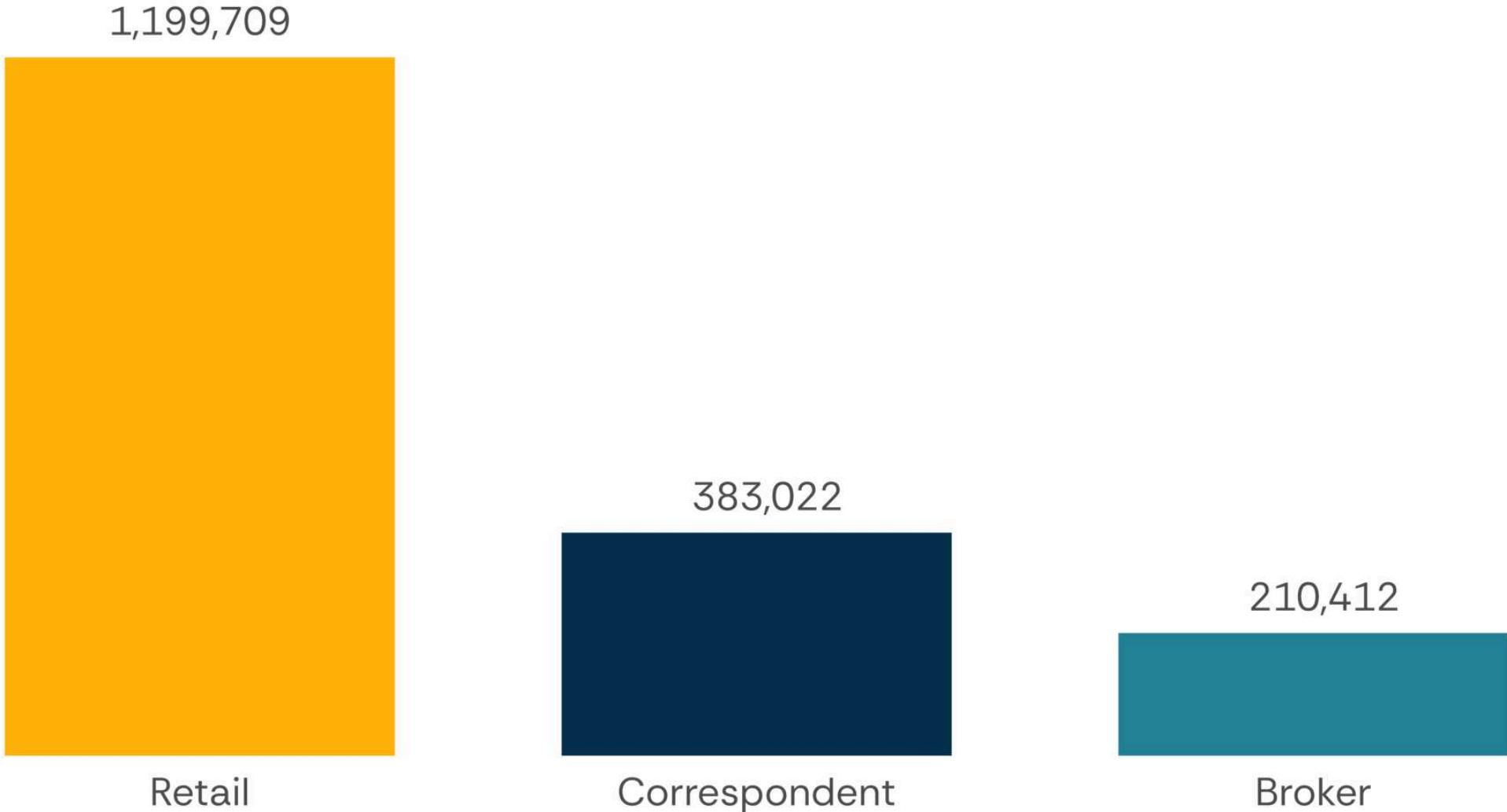
Single-Family Loans



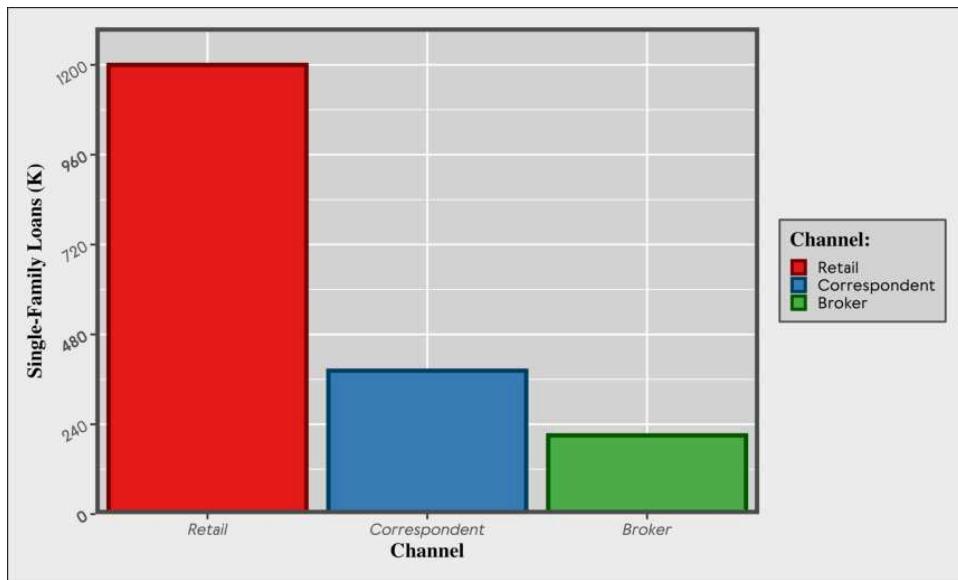
Number of single-family loans in 2022



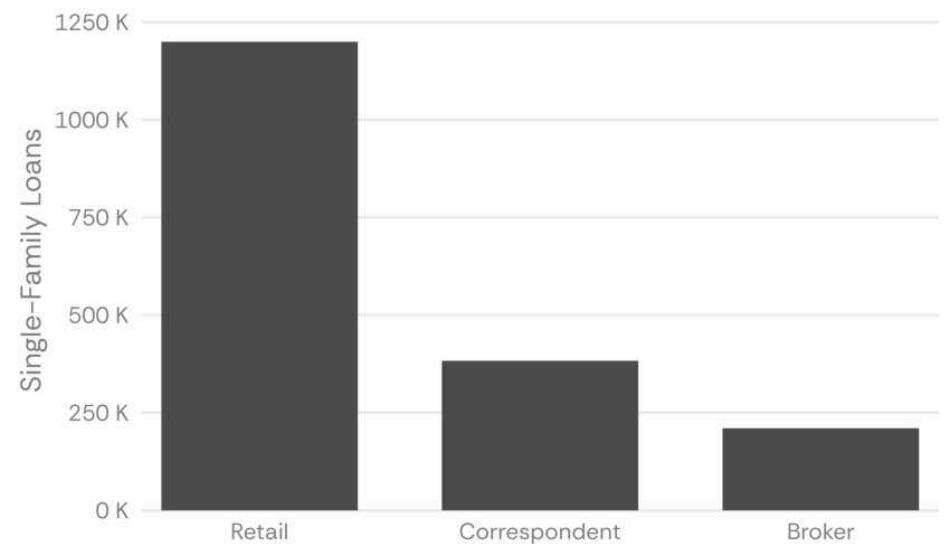
Number of single-family loans in 2022



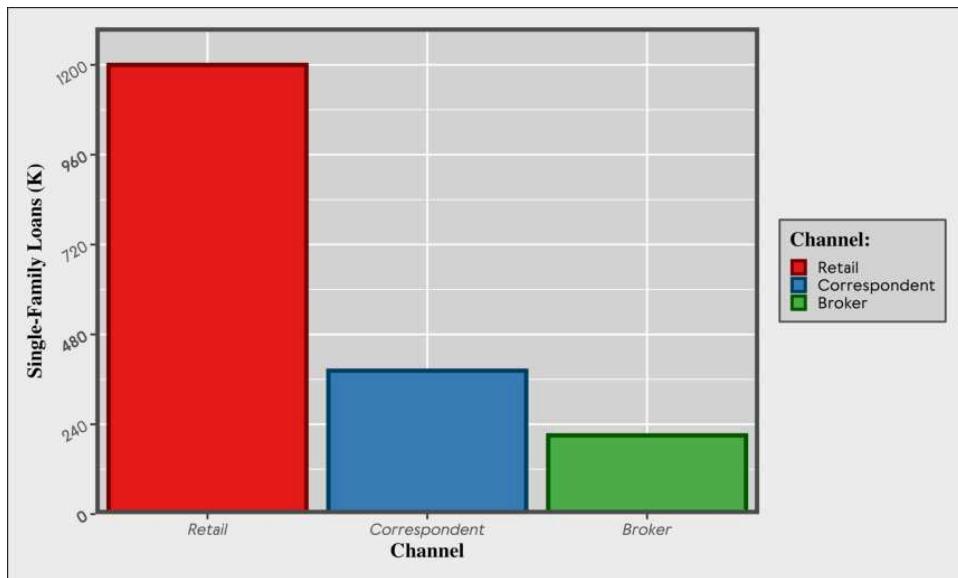
Before



After

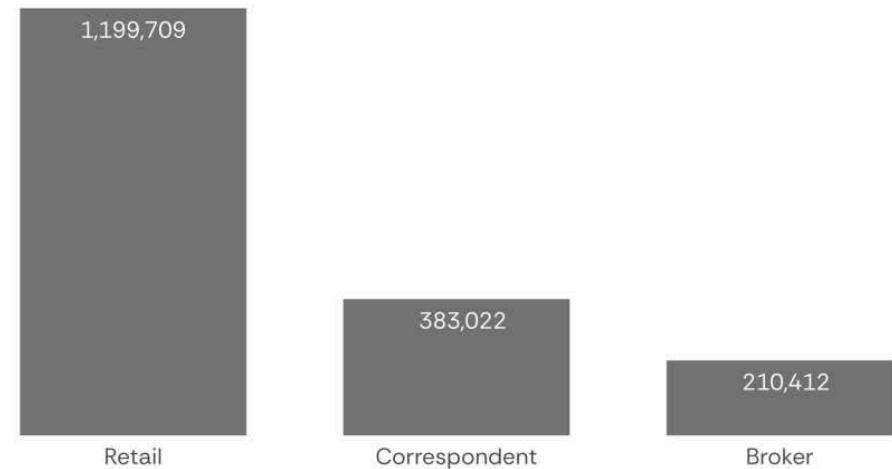


Before

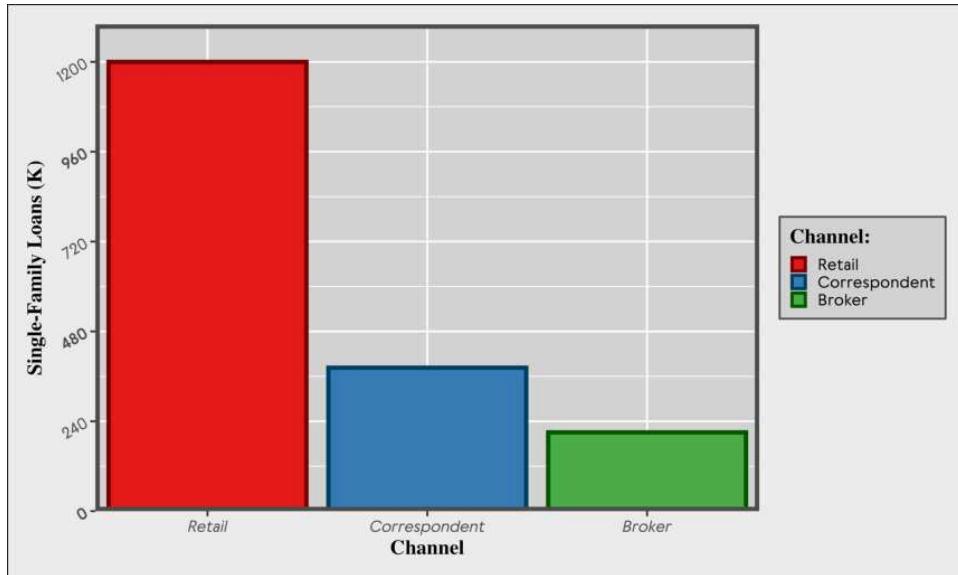


After

Number of single-family loans in 2022

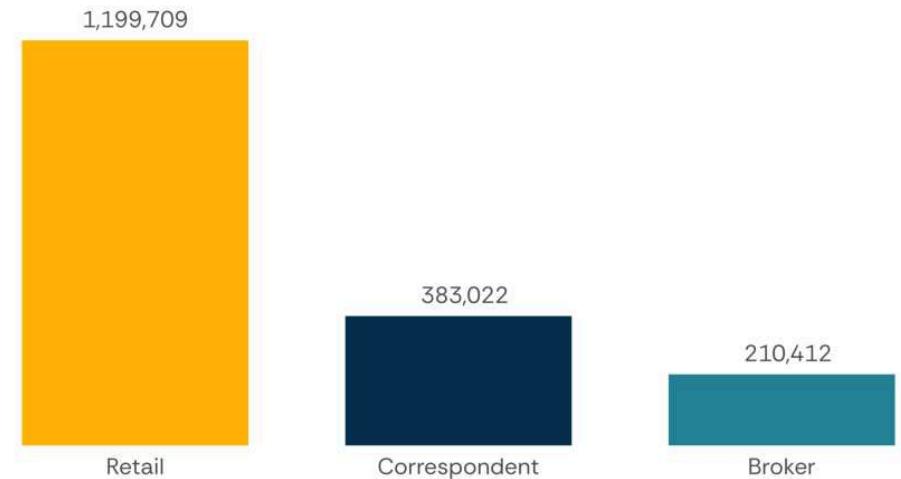


Before



After

Number of single-family loans in 2022



That's all, Folks!

