

CSCI 1377

Tools for Thought

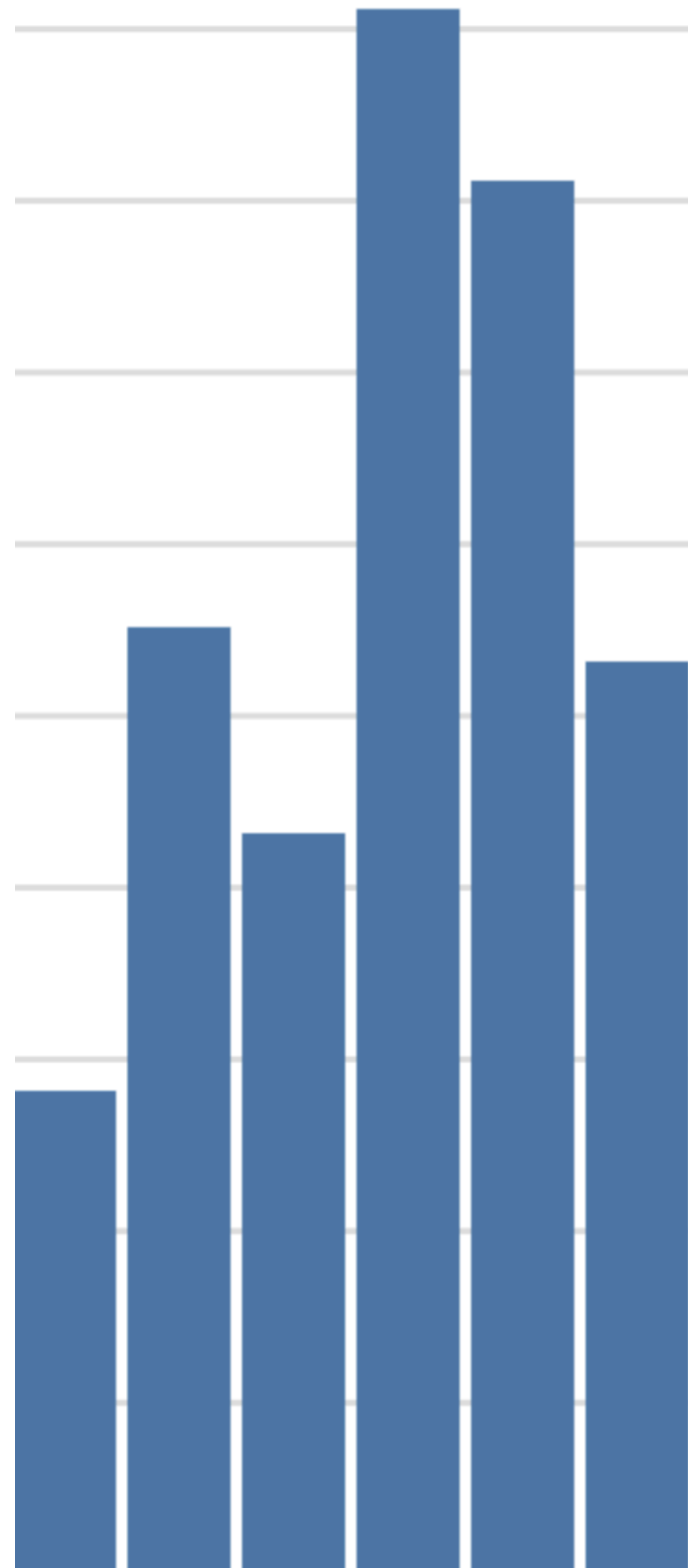
Visualization III

Expressive and Effective Visualization

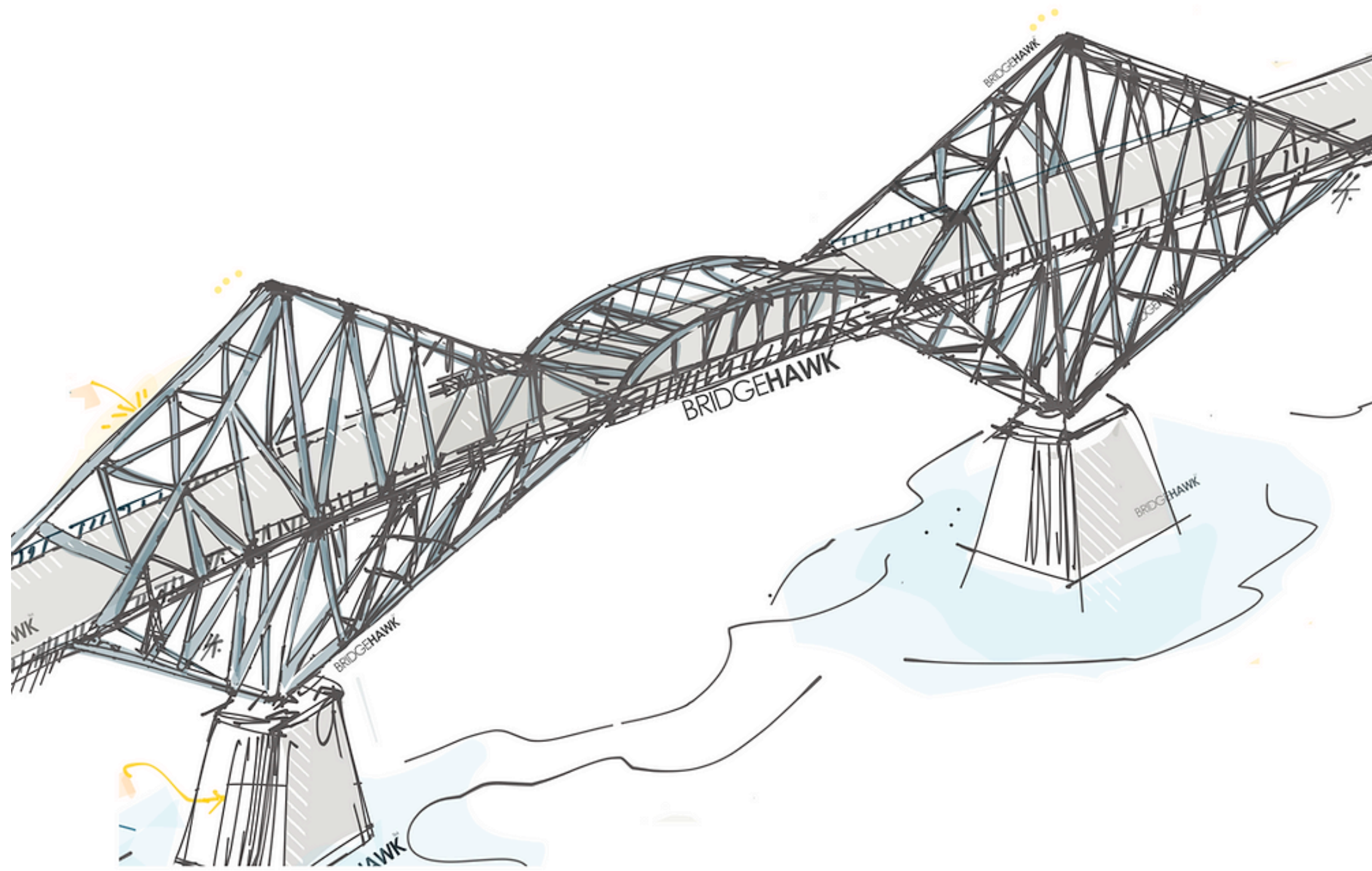
“According to *Bartlett’s Quotations*, “a picture is worth 10,000 words” is a Chinese proverb. On inquiry, we find that the Chinese seem not to have heard of it, but the proverb is certainly widely known and widely believed in our culture.”

— Larkin and Simon, “Why a Diagram is (Sometimes) Worth Ten Thousand Words” (1987)

Which of these visualizations is better?

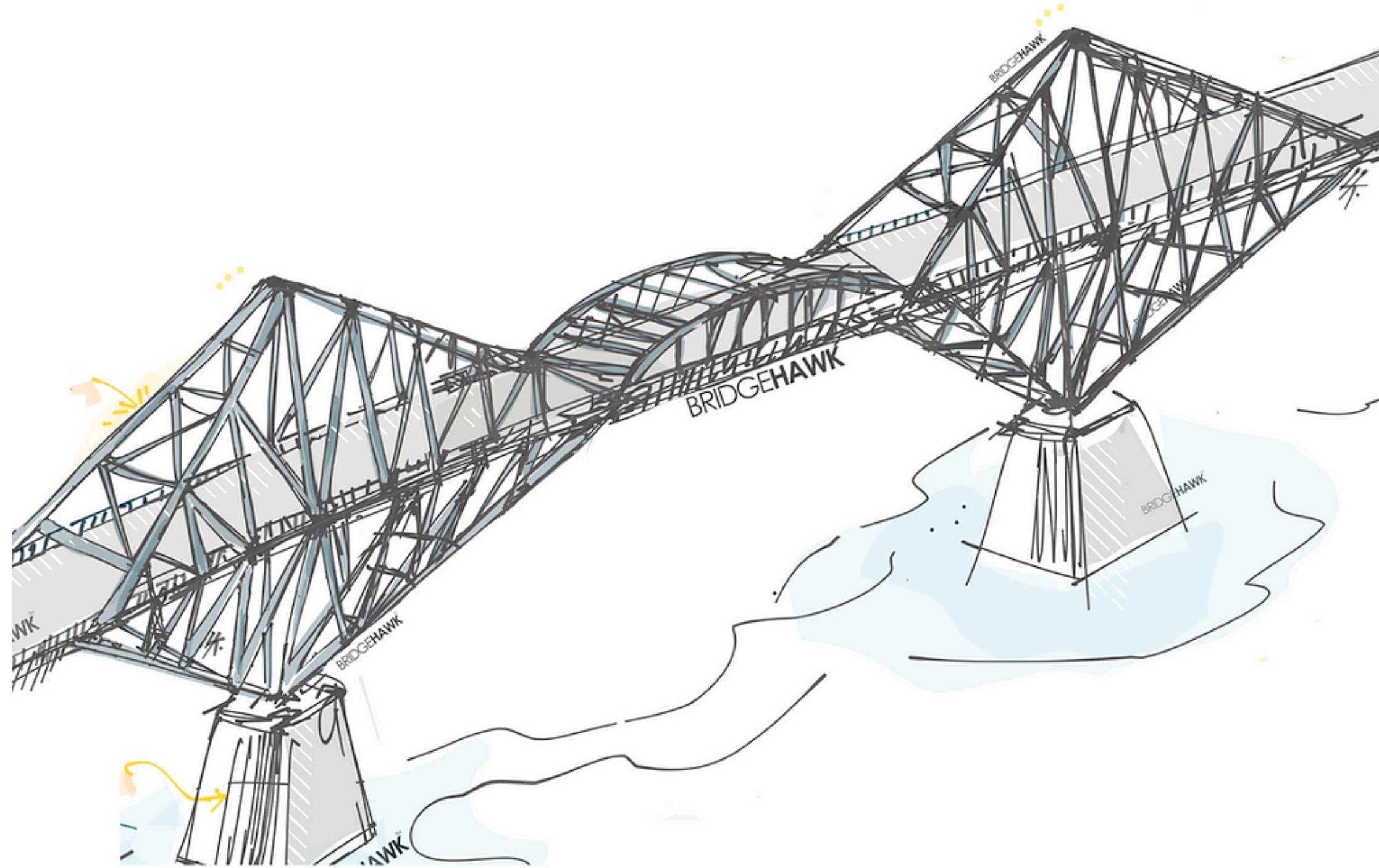


Which of these bridges is better?



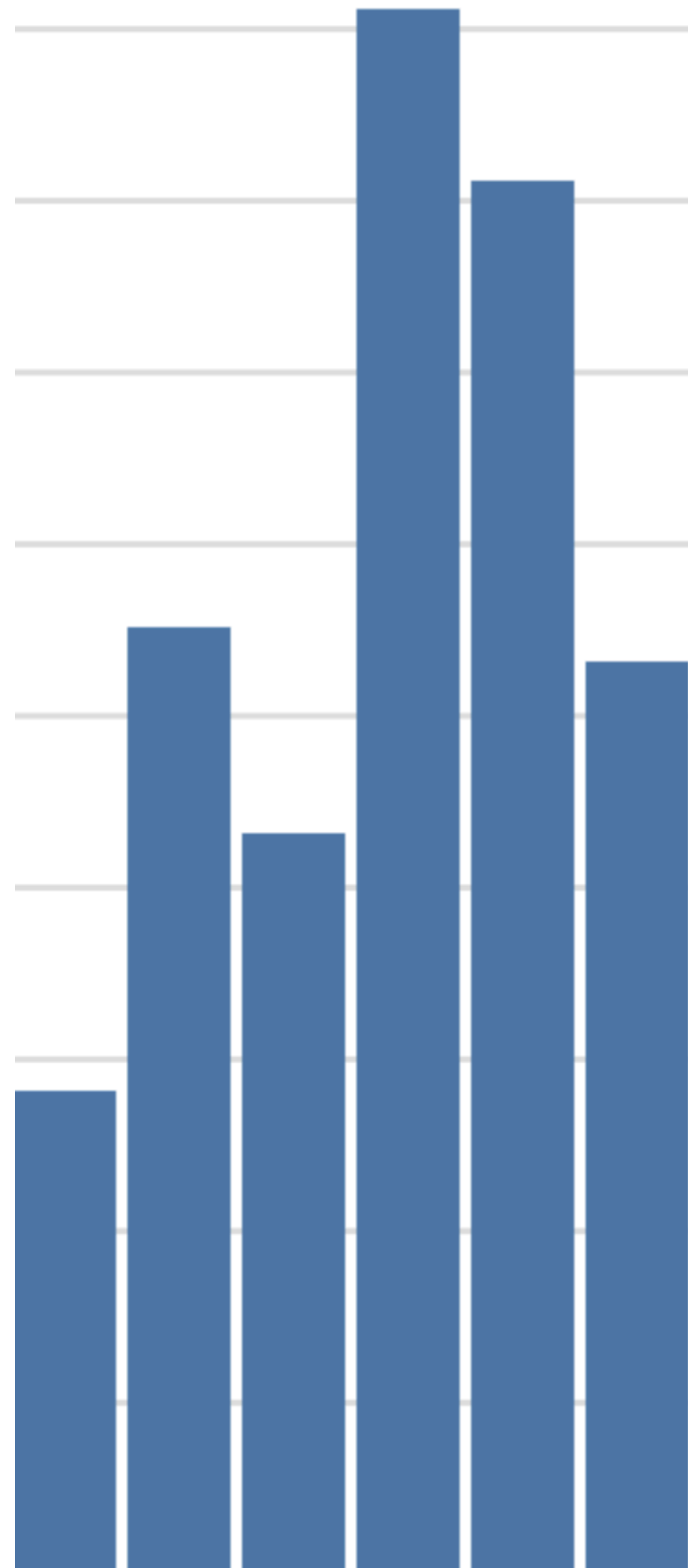
~~Which of these bridges is better?~~

Bad question!



How many cars can it carry? How much will it cost to build and maintain? Where are we building it? How well will it take the local weather?

Which of these visualizations is better?



~~Which of these visualizations is better?~~

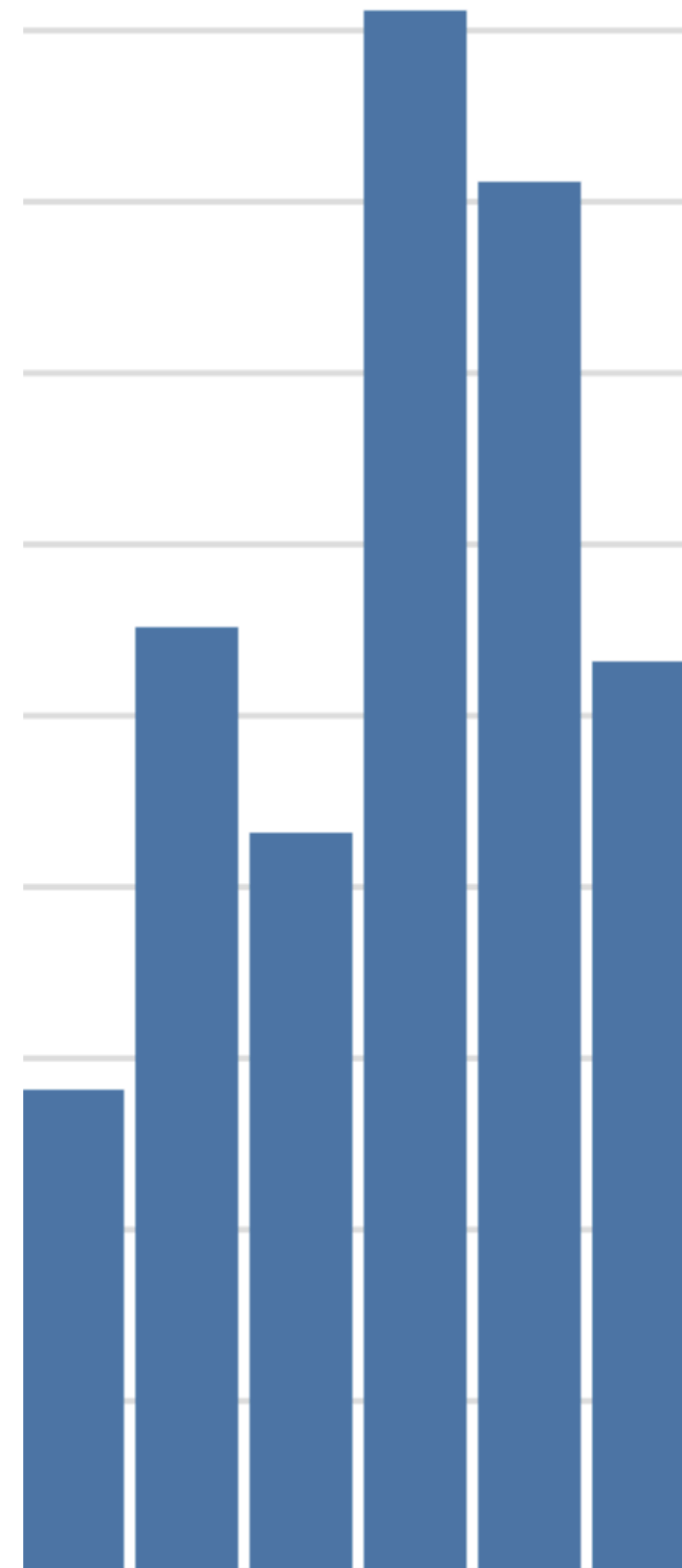
Bad question!

Which visualization helps people quickly see information about the data?

Which people?

Which information?

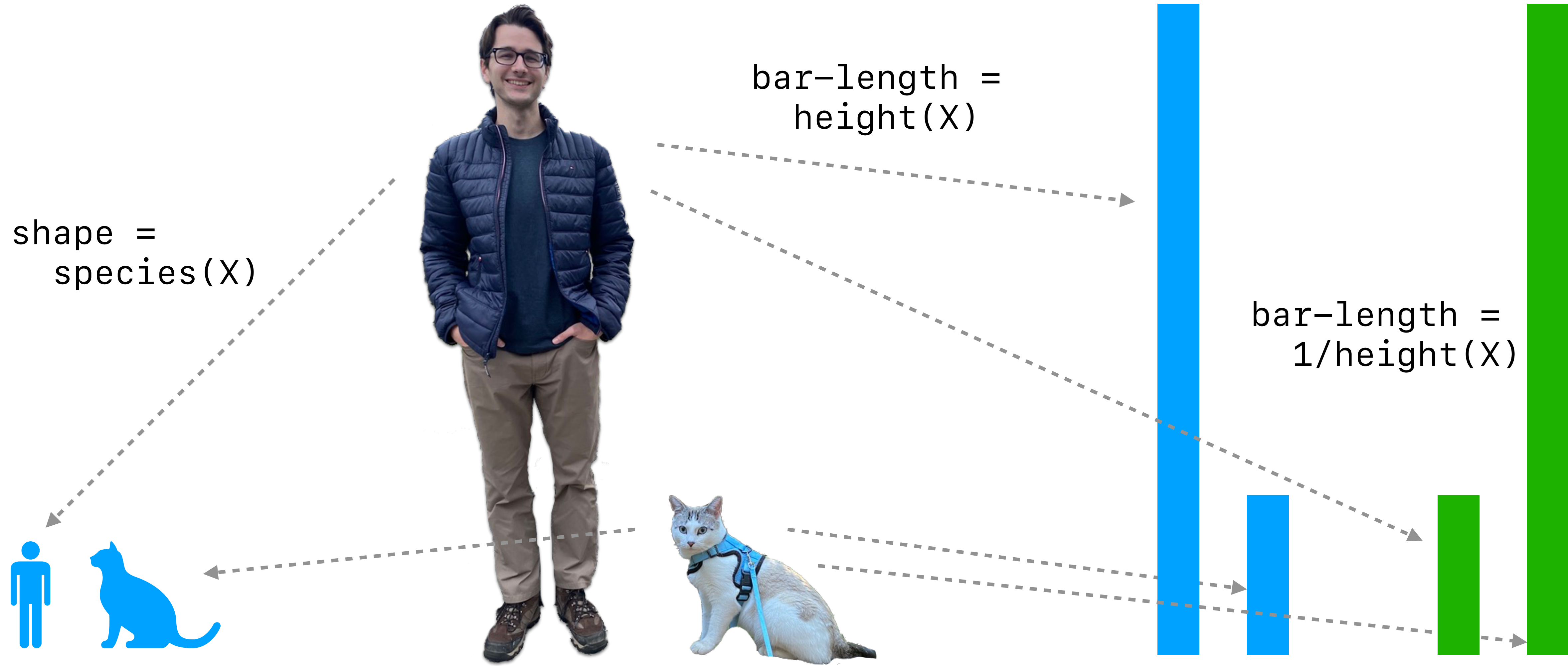
See how?



Representations and expressiveness

Which visualization helps people quickly see
information about the data?

Representational systems map information from the world to a representation



Good representational systems preserve relations between domains

$\text{height}(X)$
 $> \text{height}(Y)$

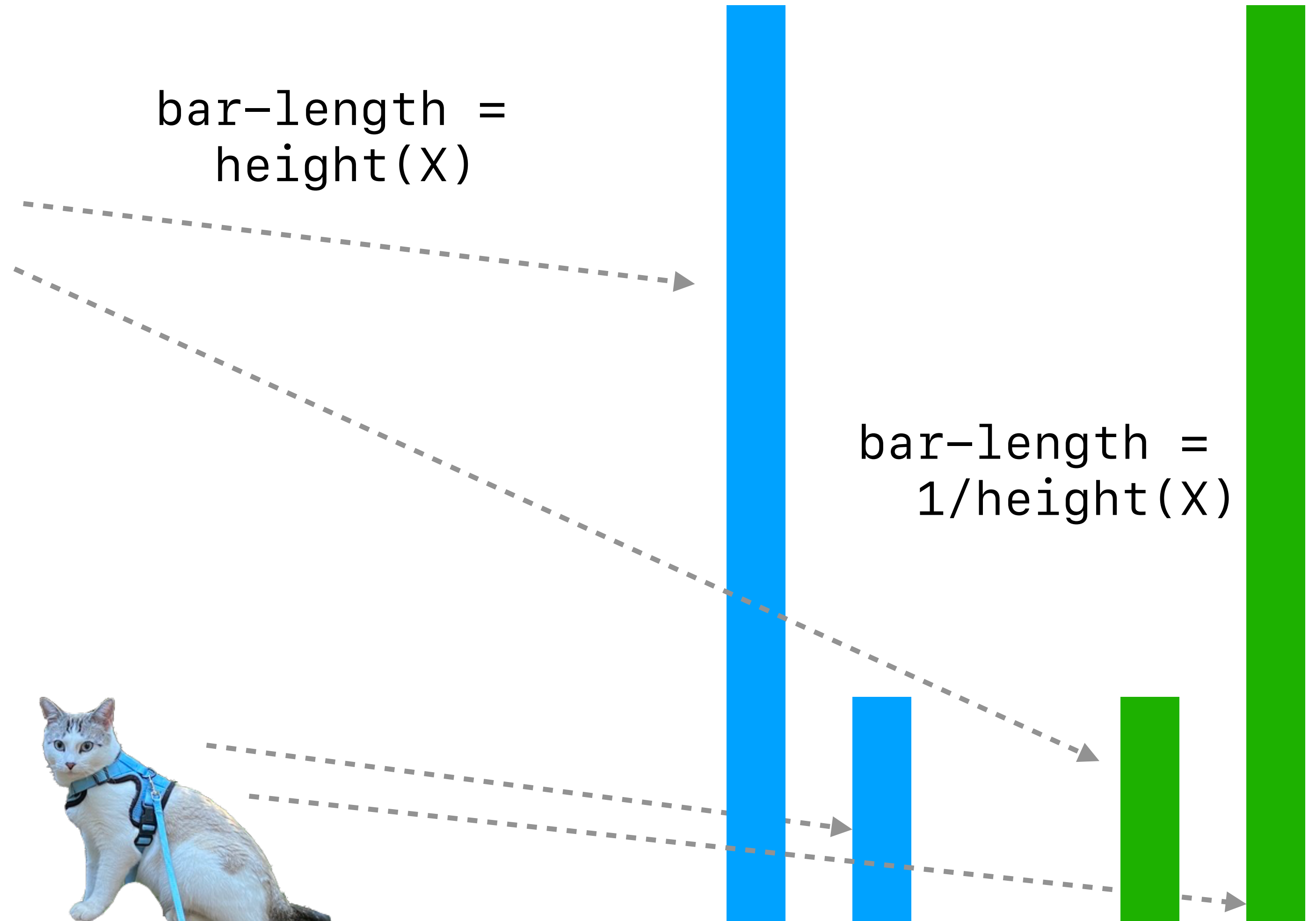


$\text{bar-length}(R(X))$
 $> \text{bar-length}(R(Y))$

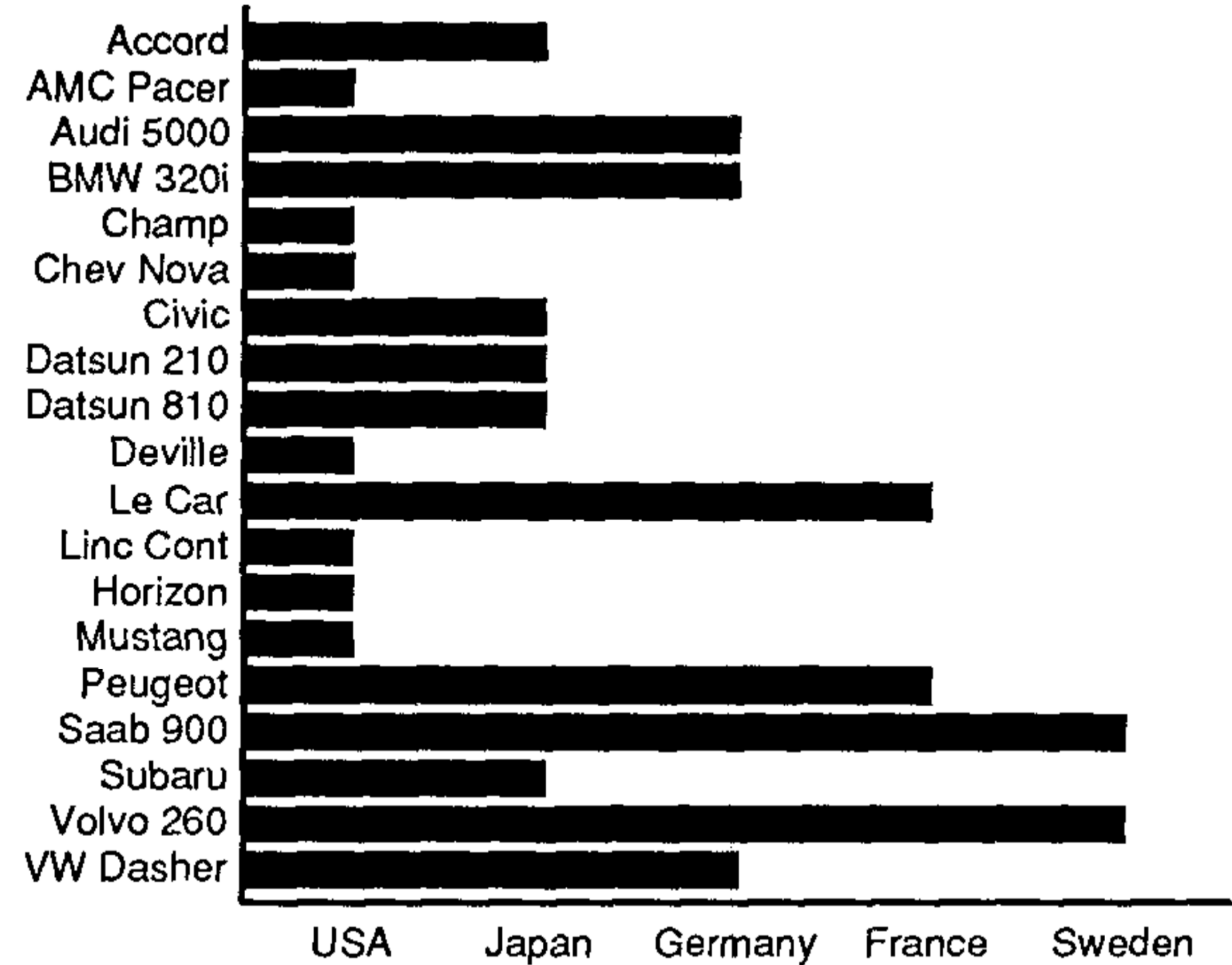
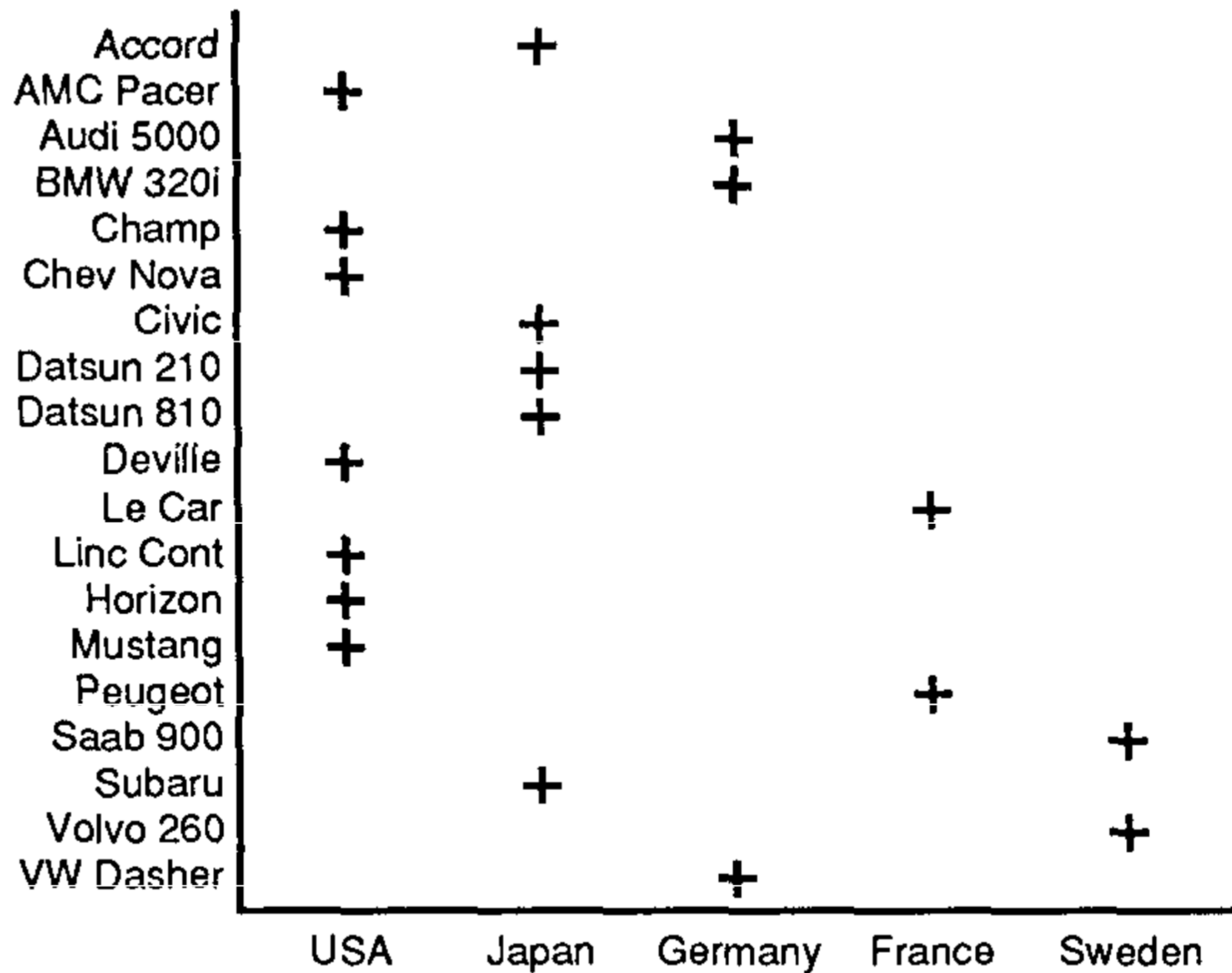


$\text{bar-length} = \text{height}(X)$

$\text{bar-length} = 1/\text{height}(X)$



Which mark best encodes the data?



Expressiveness: “a set of facts is expressible in a visual language if there is a visualization that encodes *all* and *only* the facts in the set.”

What is the mapping from data to mark?

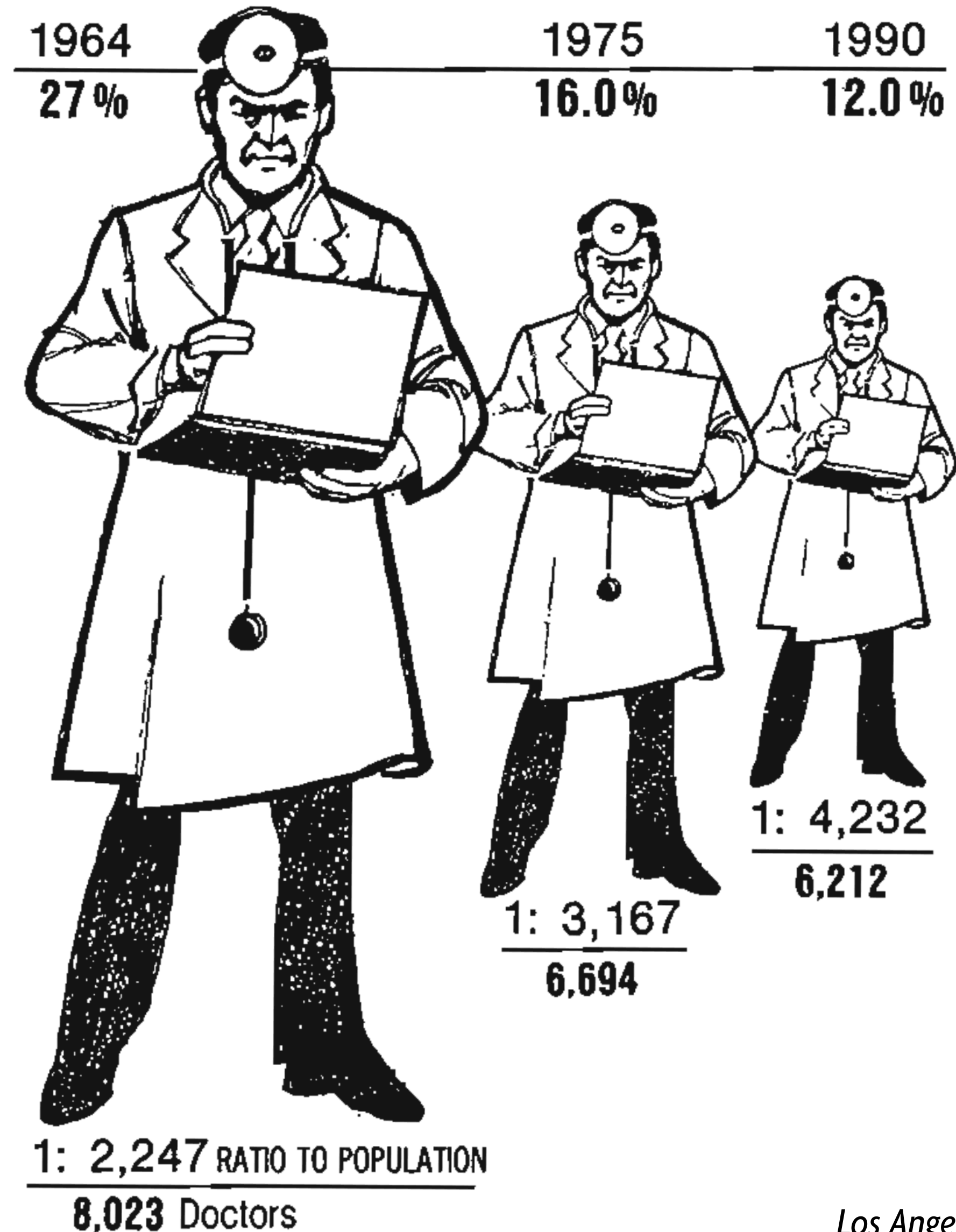
Why does it mislead?

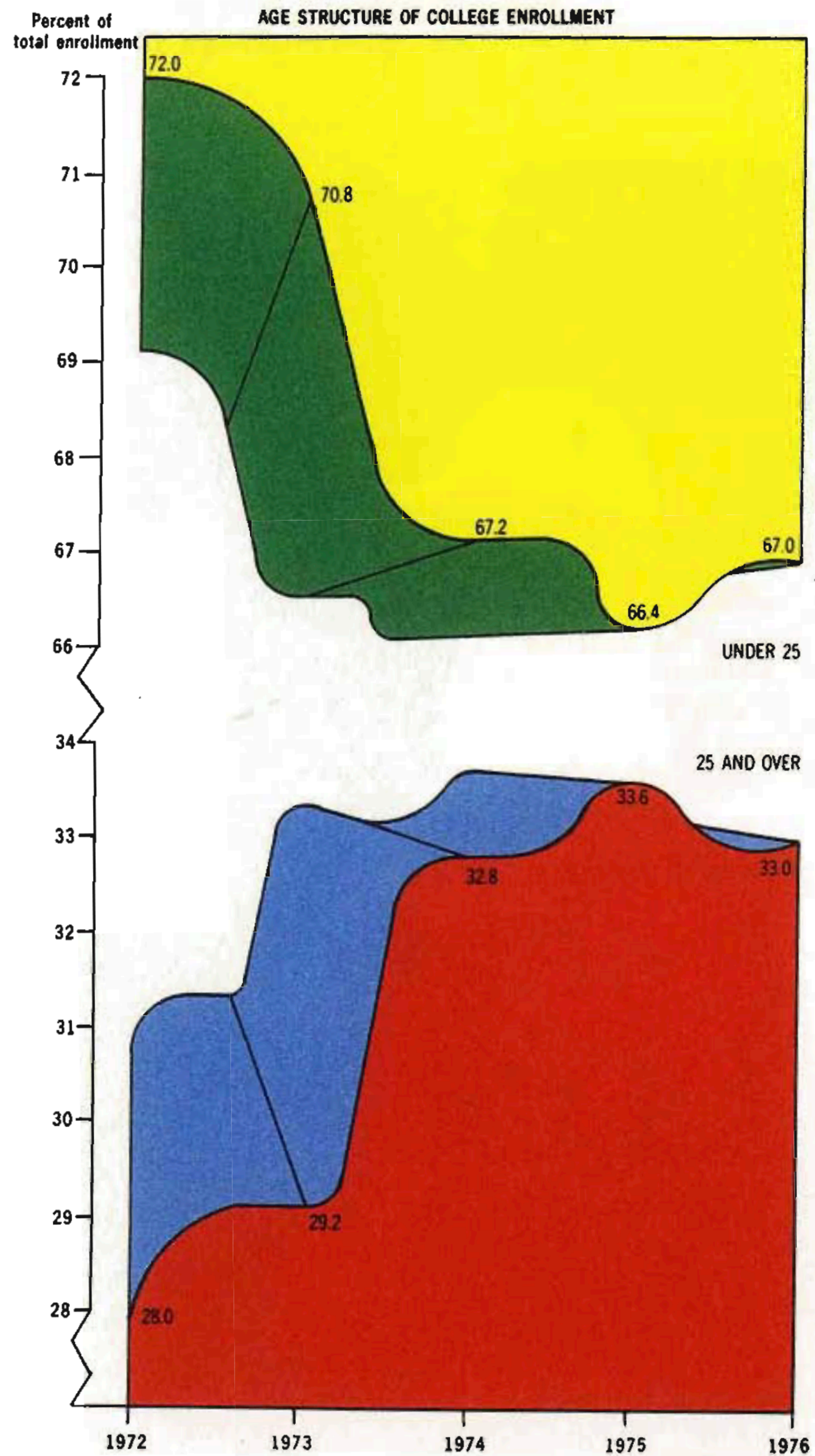
Data	Width	Area
8000	650	1,000,000
7000	400	400,000
6000	275	200,000

THE SHRINKING FAMILY DOCTOR In California

Percentage of Doctors Devoted Solely to Family Practice

1964	1975	1990
27%	16.0%	12.0%





“This may well be the worst graphic ever to find its way into print” — Tufte

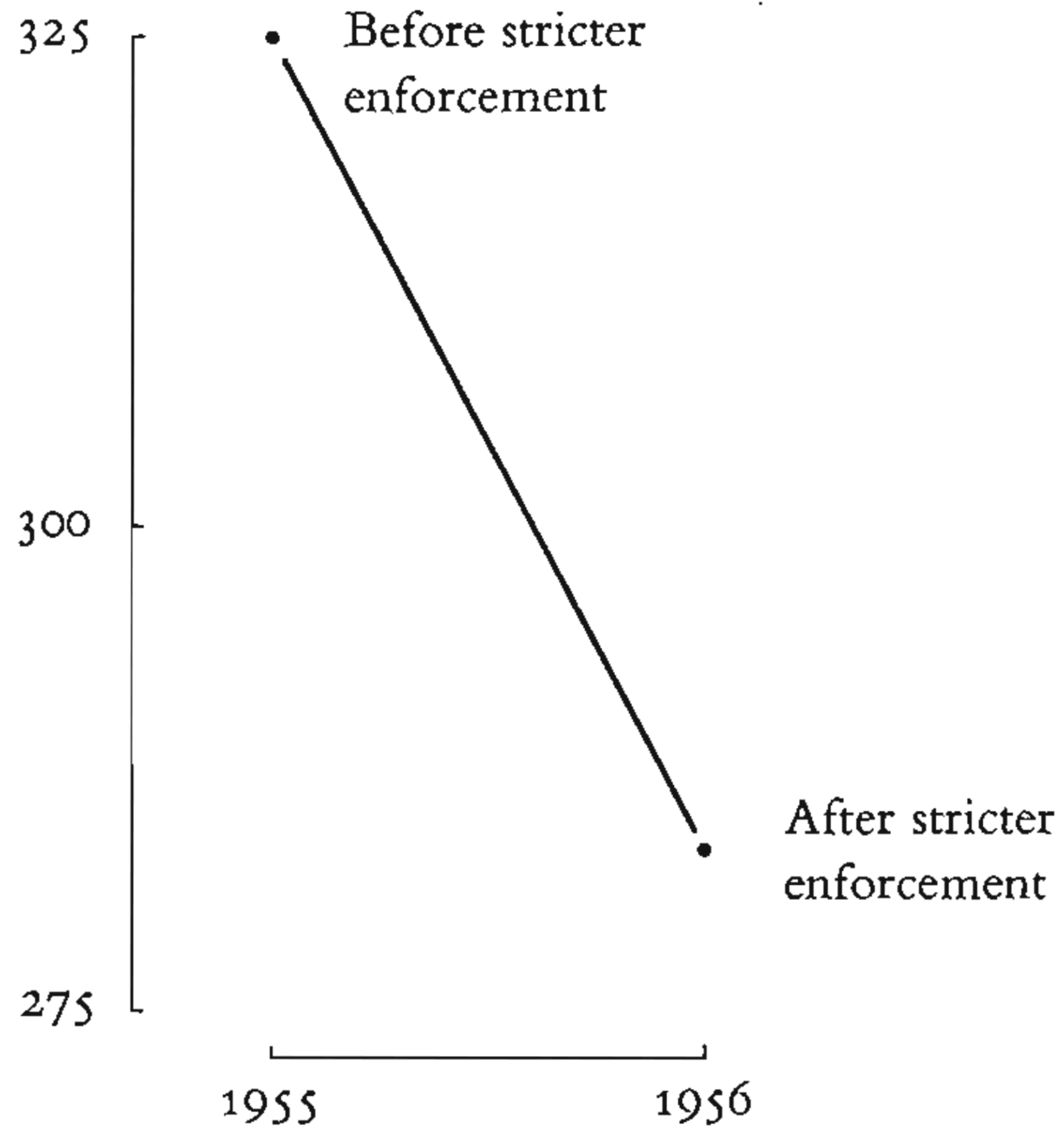
“The emaciated, data-thin design should always provoke suspicion, for graphics often lie by omission, leaving out data sufficient for comparisons. The principle:

Graphics must not quote data out of context.”

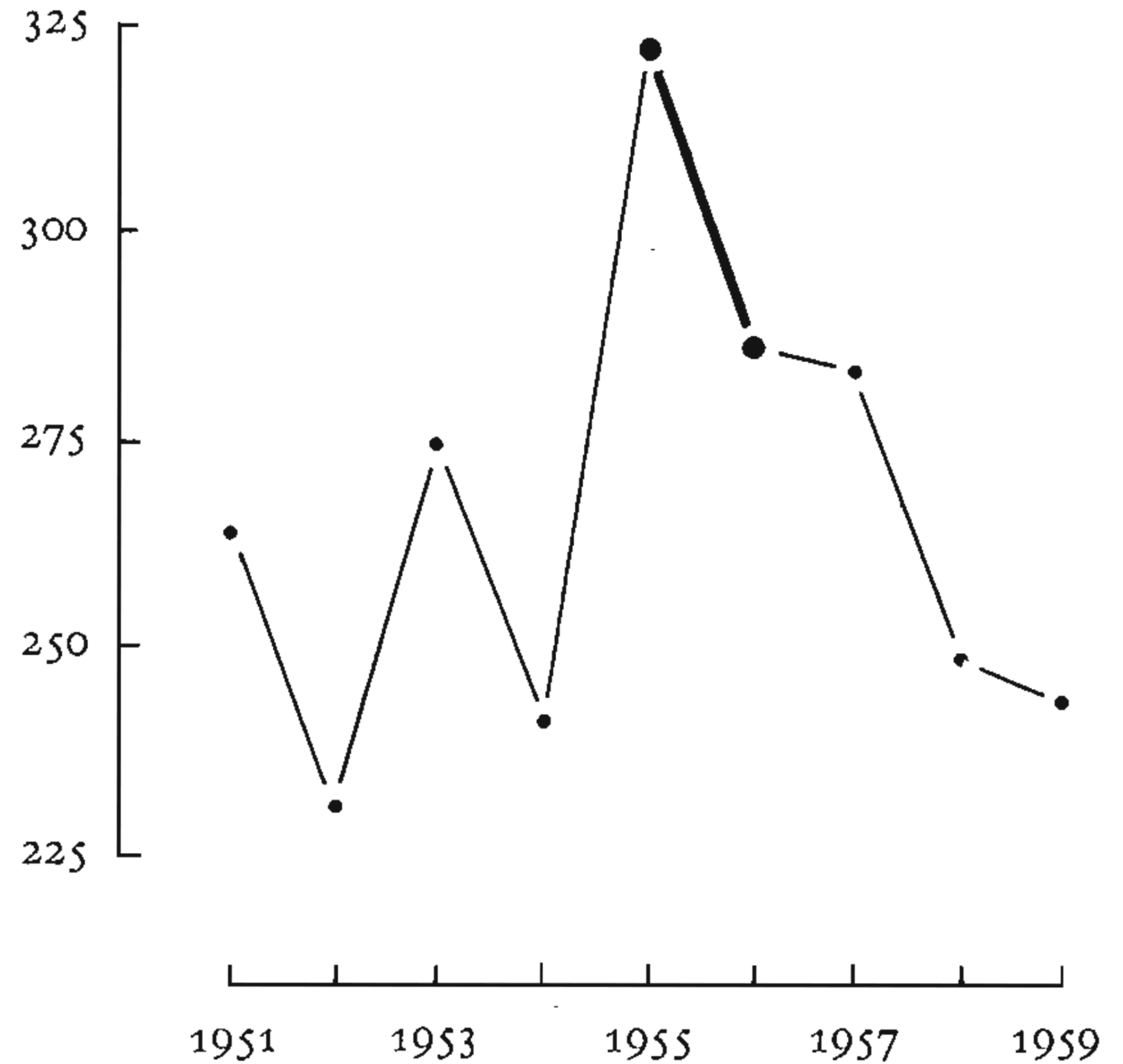


— Edward Tufte, *The Visual Display of Quantitative Information* (1983)

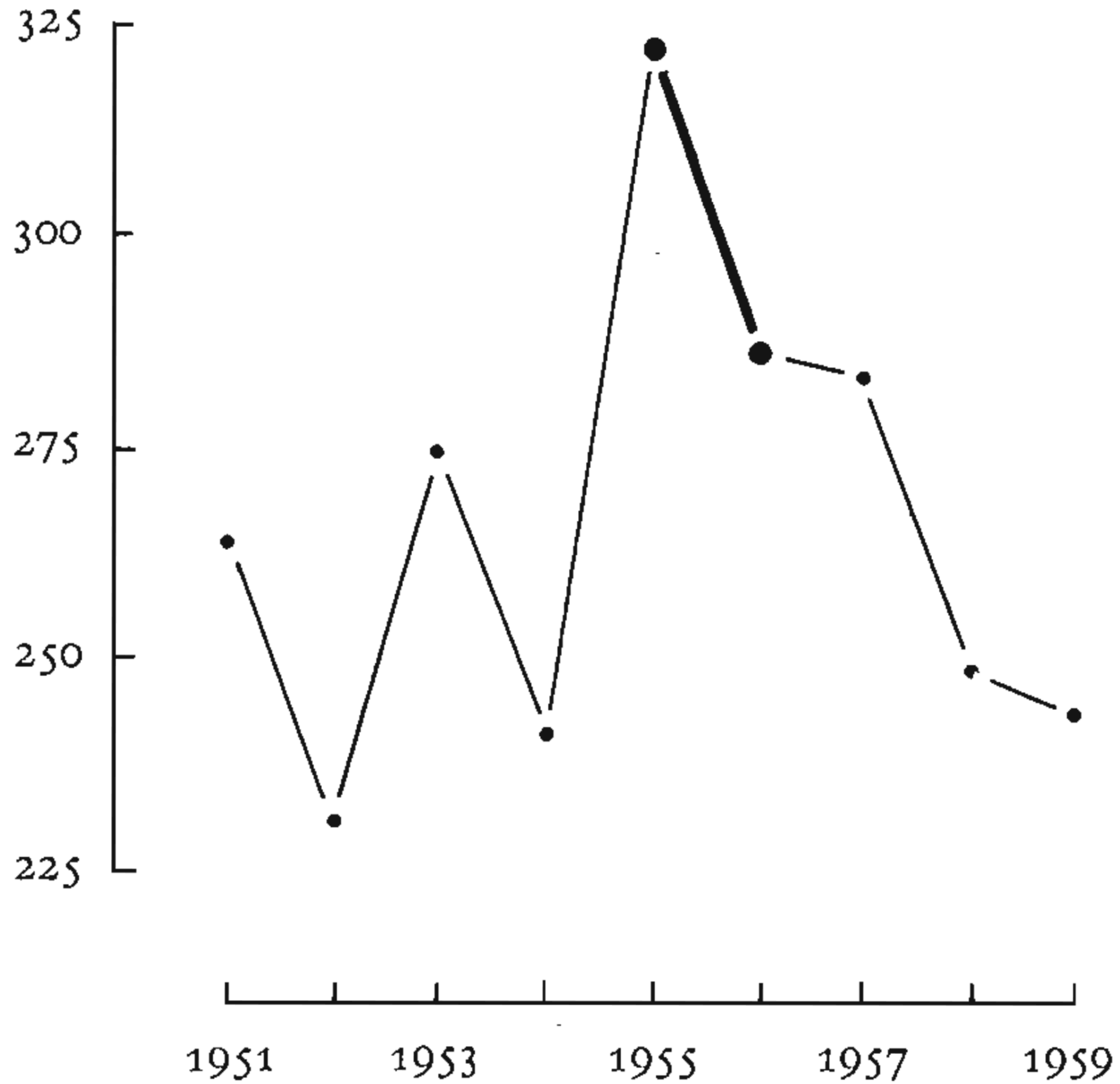
Connecticut Traffic Deaths,
Before (1955) and After (1956)
Stricter Enforcement by the Police
Against Cars Exceeding Speed limit



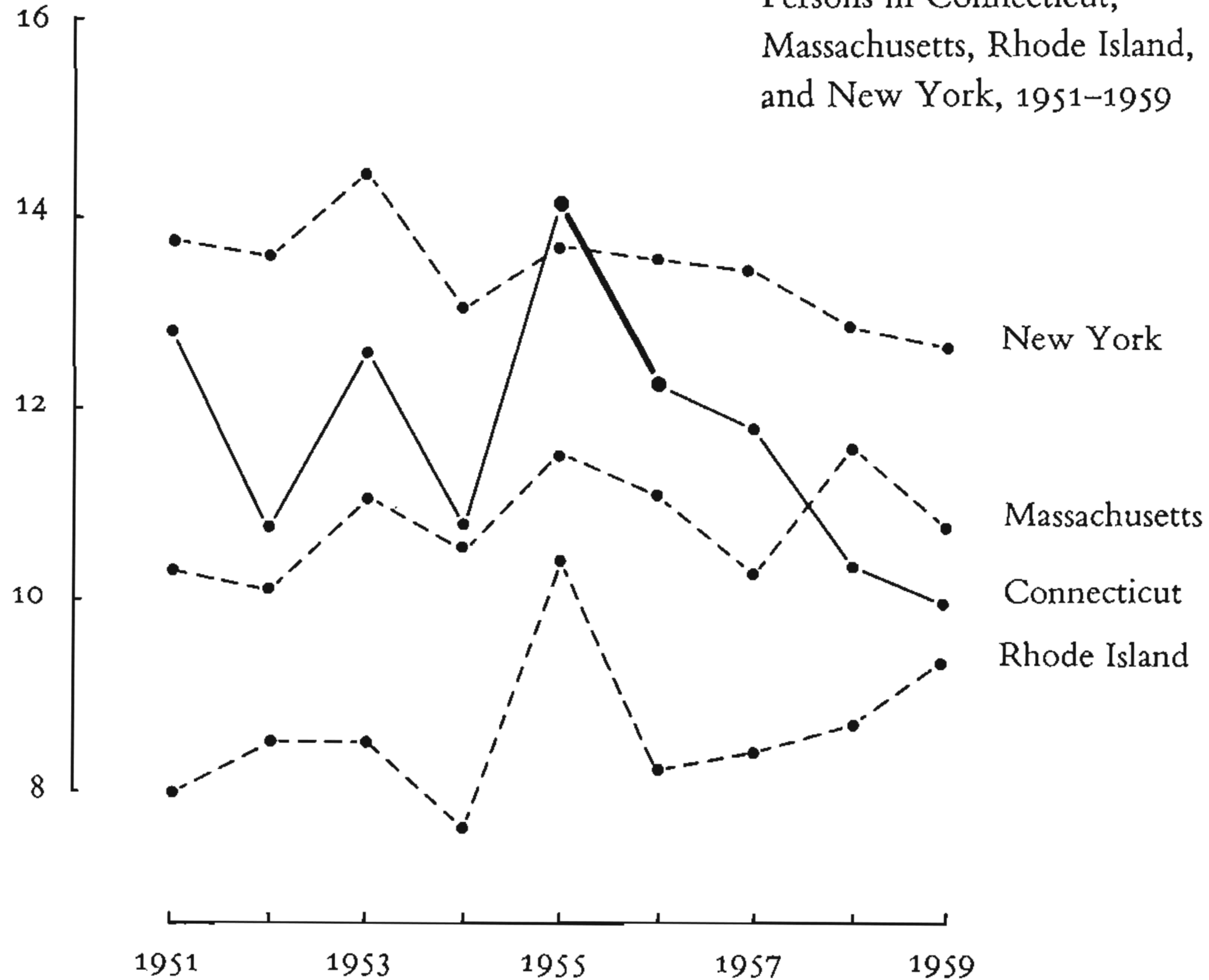
Connecticut Traffic Deaths,
1951-1959



Connecticut Traffic Deaths,
1951-1959



Traffic Deaths per 100,000
Persons in Connecticut,
Massachusetts, Rhode Island,
and New York, 1951-1959



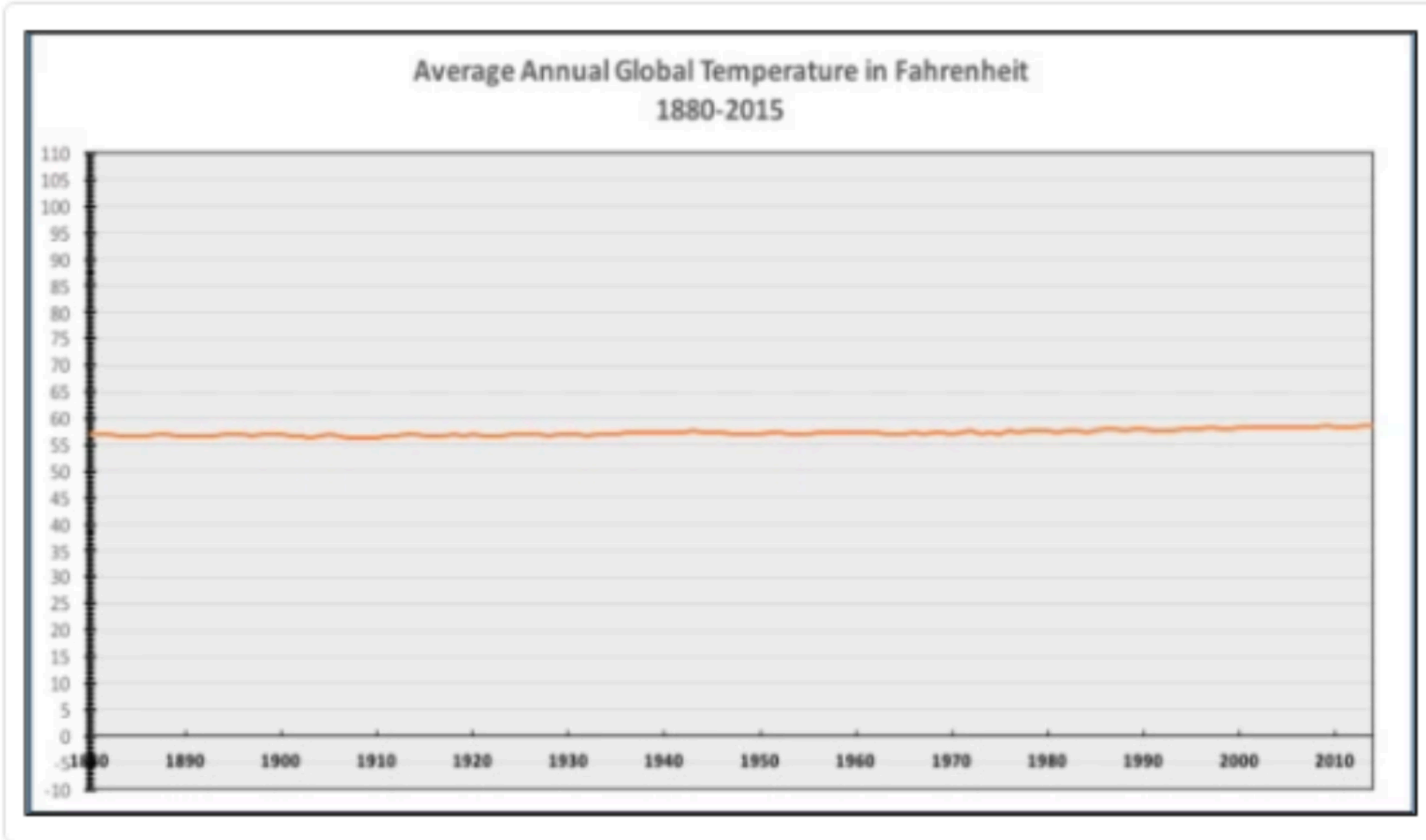
Scale is context, too!



Follow

The only [#climatechange](#) chart you need to see. natl.re/wPKpro

(h/t [@powerlineUS](#))

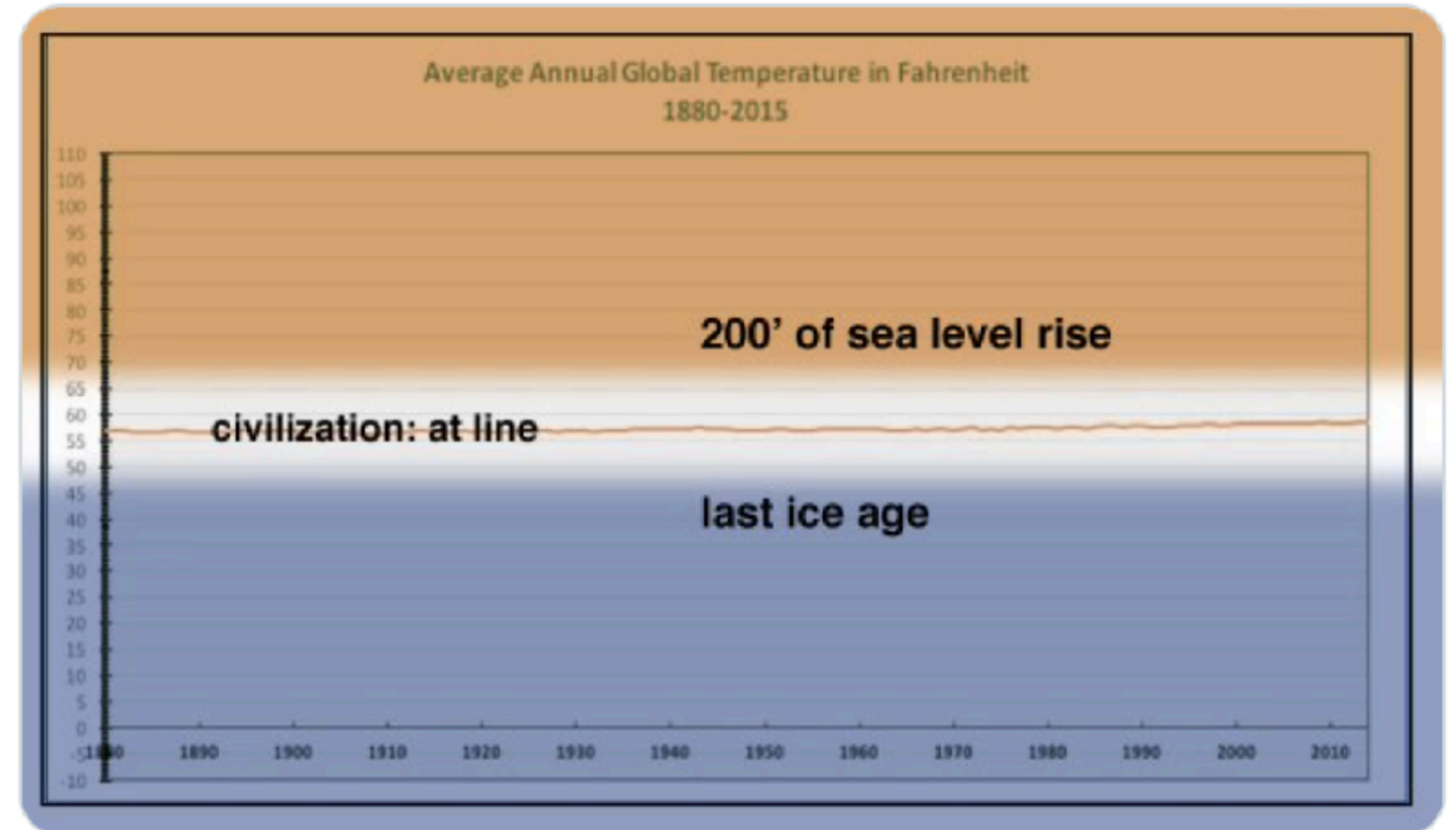


12:36 PM - 14 Dec 2015



Replying to [@NRO](#)

[.@NRO](#) [@powerlineUS](#) [@bradplumer](#) I'm sure someone else has fixed this for you, but here you go. Great idea, thx --



5:28 PM · Dec 14, 2015

78 Retweets 1 Quote Tweet 208 Likes

Encodings and effectiveness

Which visualization helps **people** **quickly see** information about the data?

Signal Detection

Discriminability: how easy is it to tell two things apart?

Magnitude Estimation

Accuracy: how correctly can we read off values?

Preattentive Processing

Pop out: how easy is it to spot some values from the rest?

Selective Attention

Separability: how much interaction occurs between attributes?

➔ **Magnitude Channels: Ordered Attributes**



Same

Same

Same

➔ **Identity Channels: Categorical Attributes**



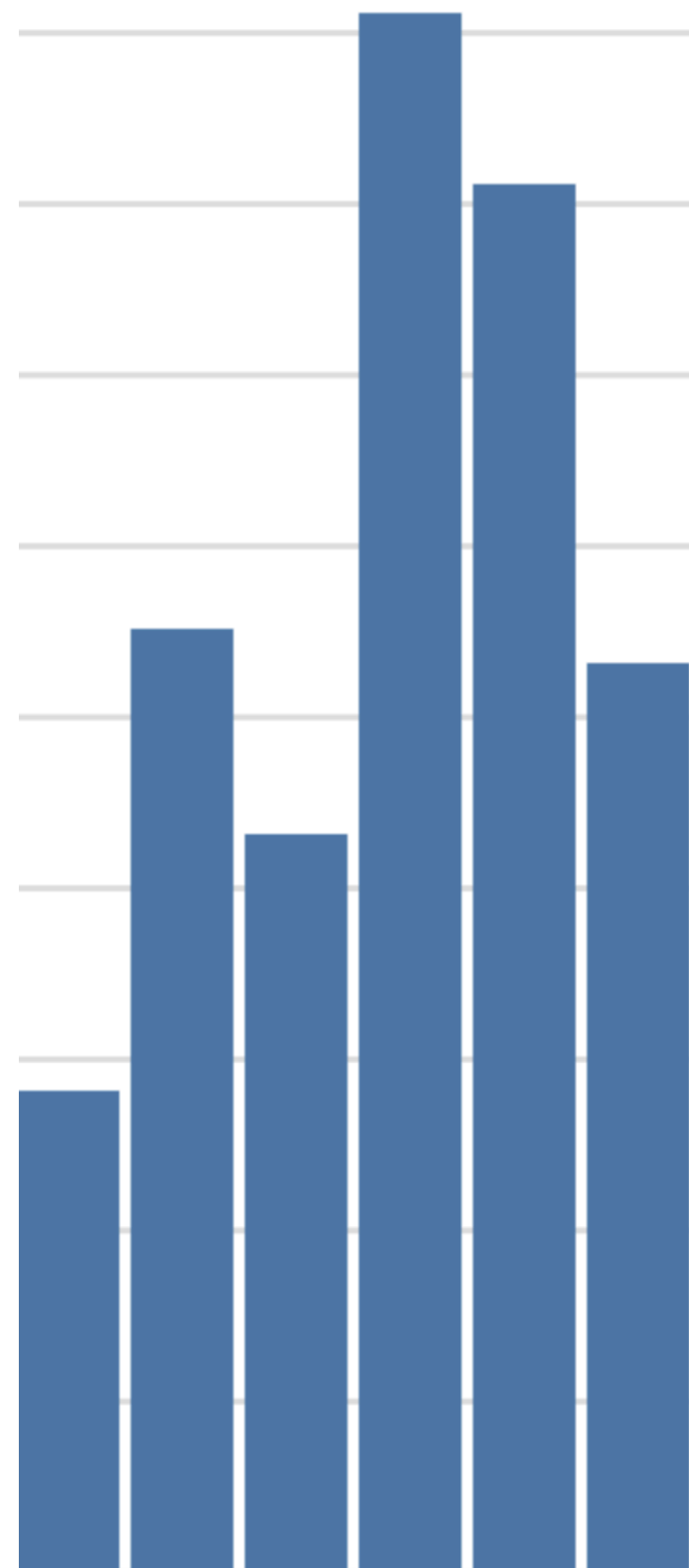
Most

Effectiveness

Least

Which of these visualizations is better?

Which encoding of nominal \times quantitative data facilitates comparisons across categories?



Position on common scale



Position on unaligned scale



Length (1D size)



Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance

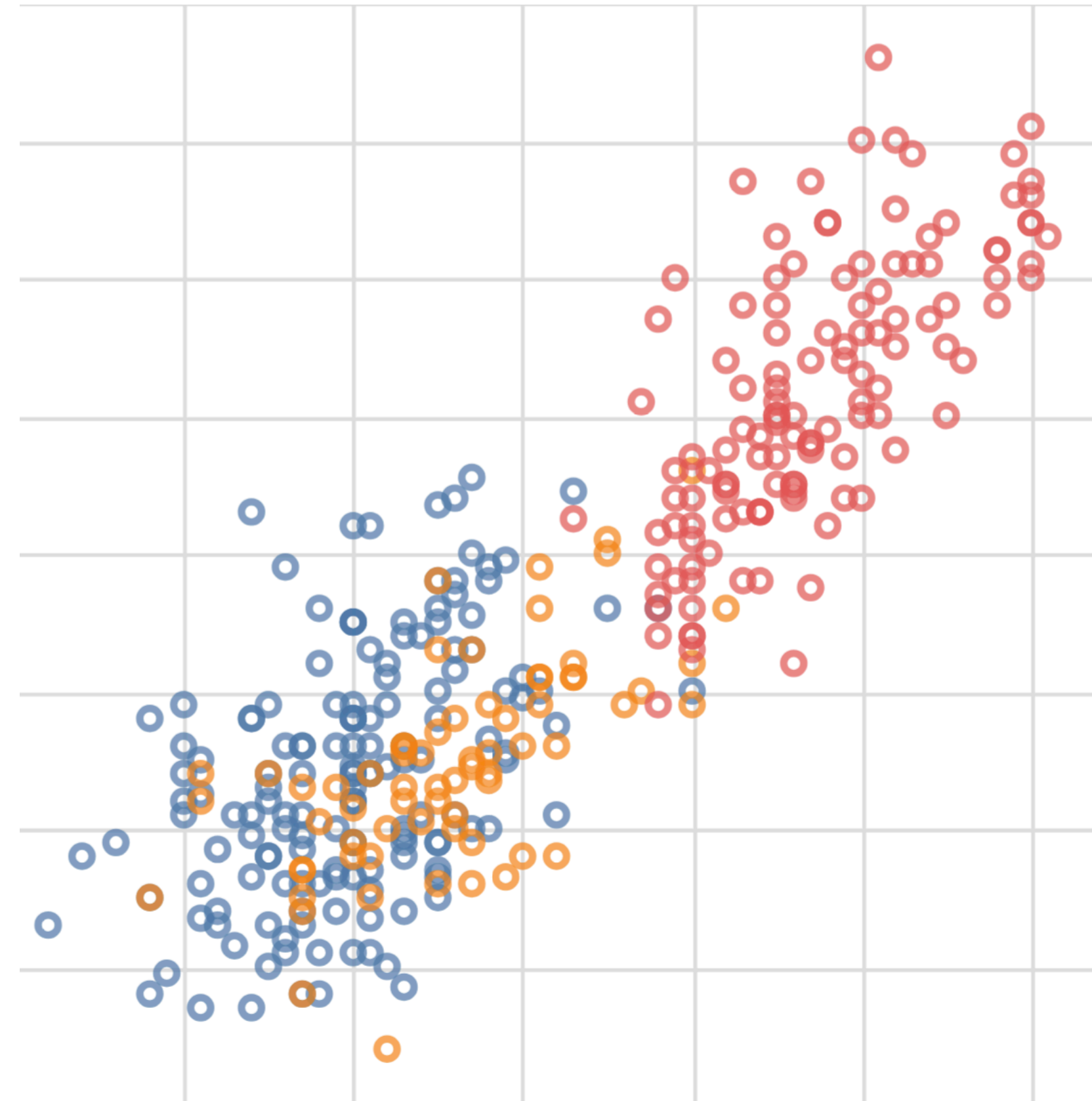
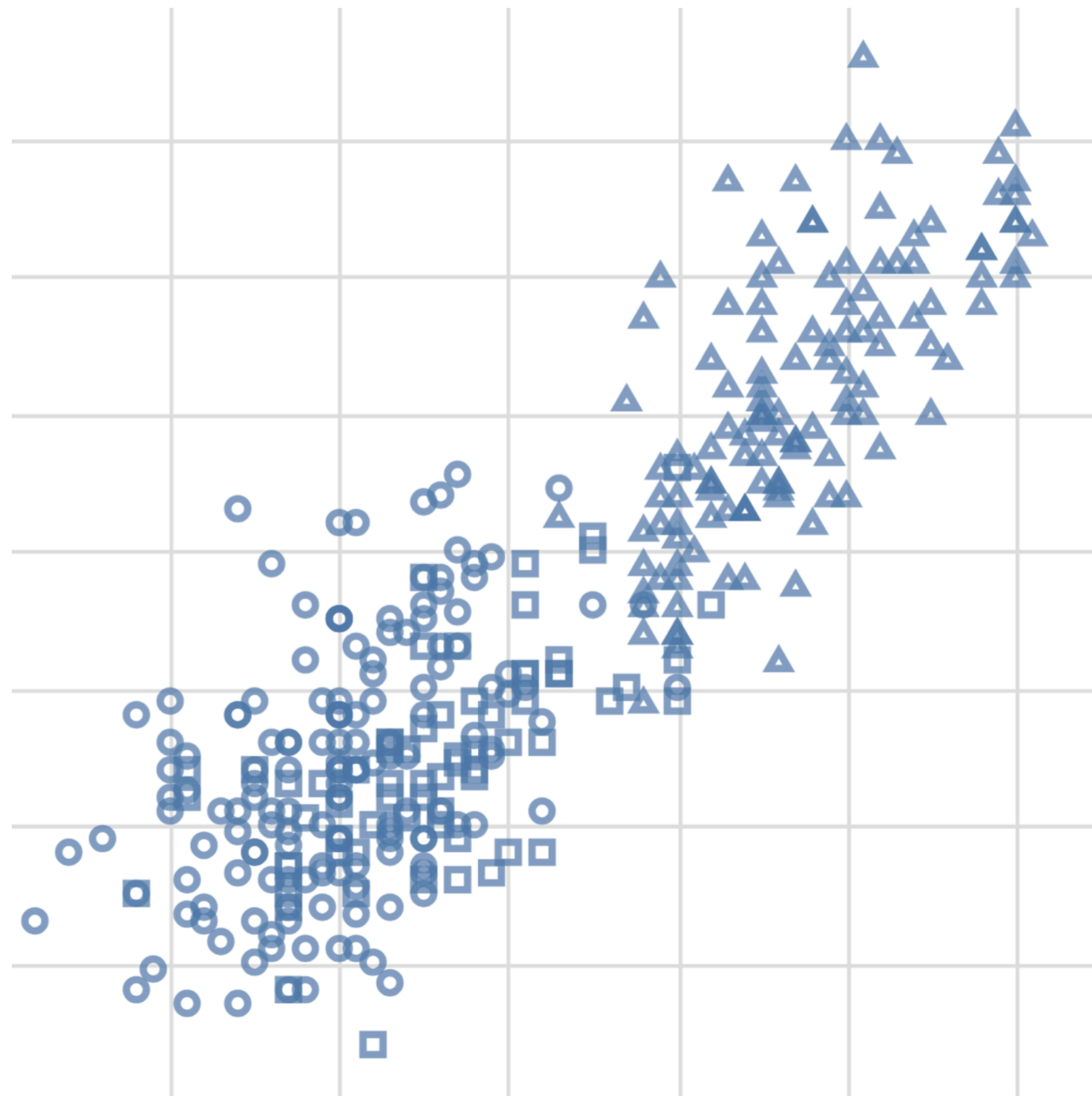


How does Z compare across X and Y ?

Color

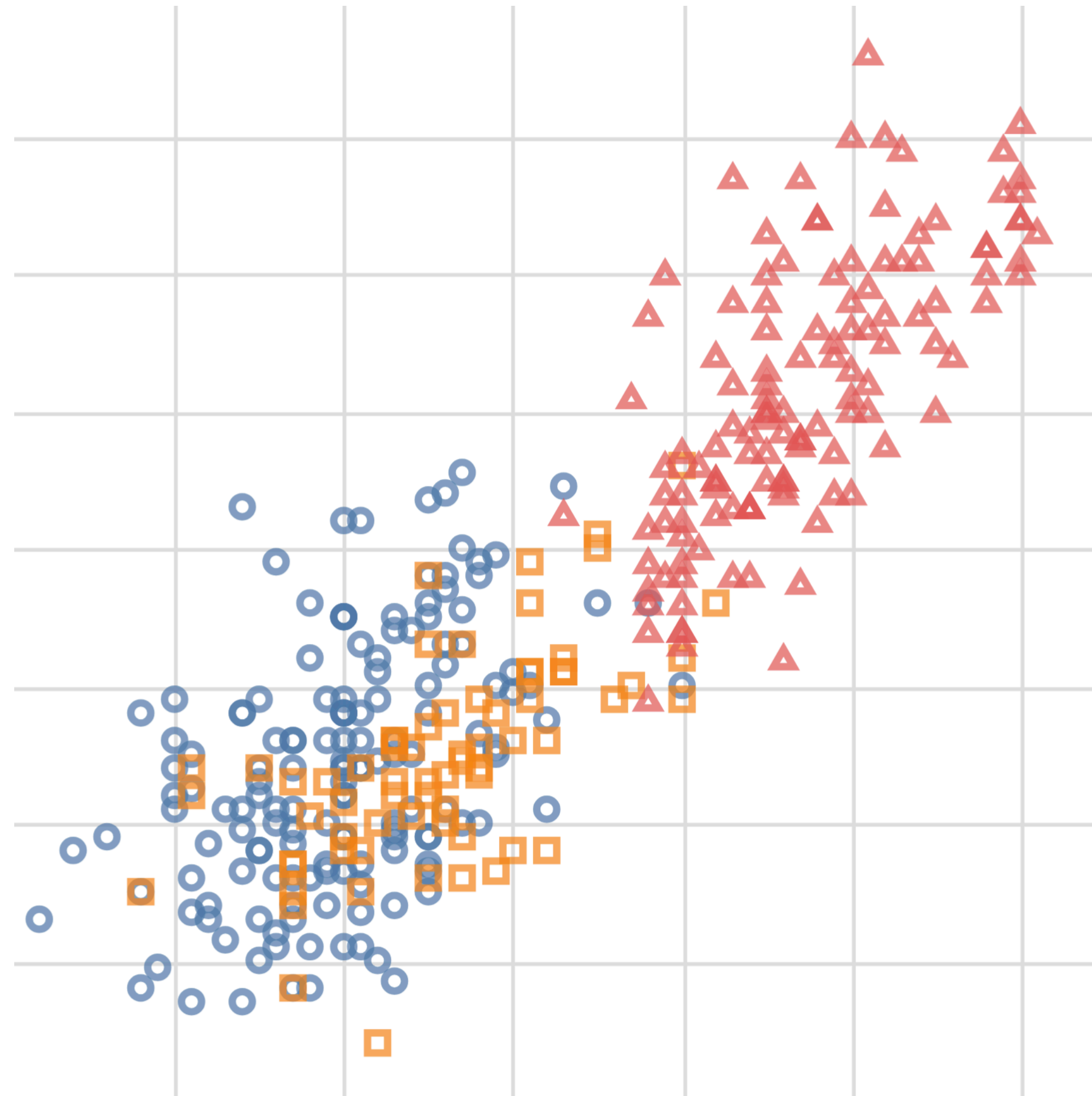


Quantitative Y



Quantitative X

Por que no los dos? Use color + shape for redundant encoding



For a given X , how does Y compare across Z ?

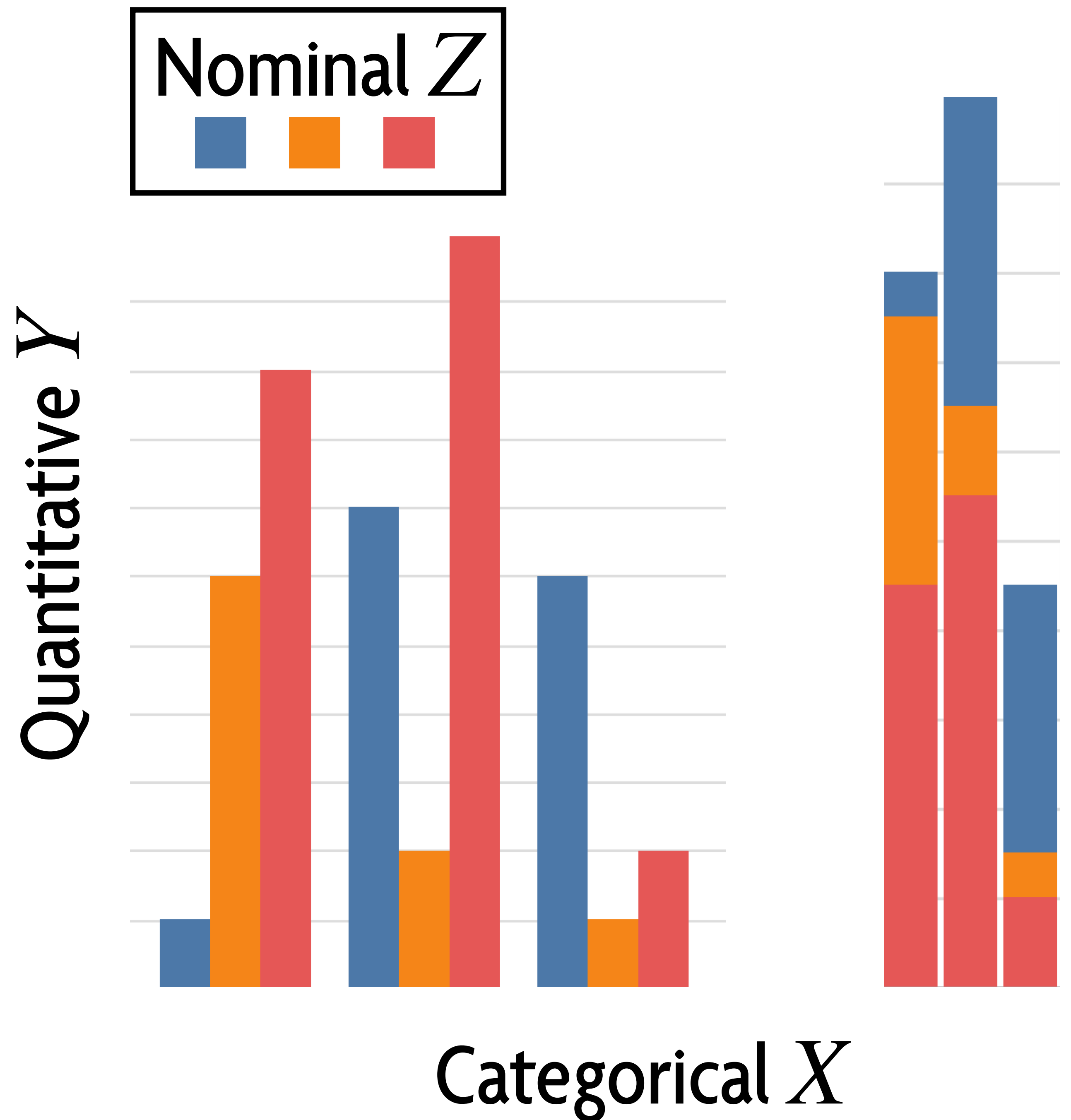
Grouped bar

Combining Z , how does Y compare across X ?

Stacked bar

For a given Z , how does Y compare across X ?

Grouped bar

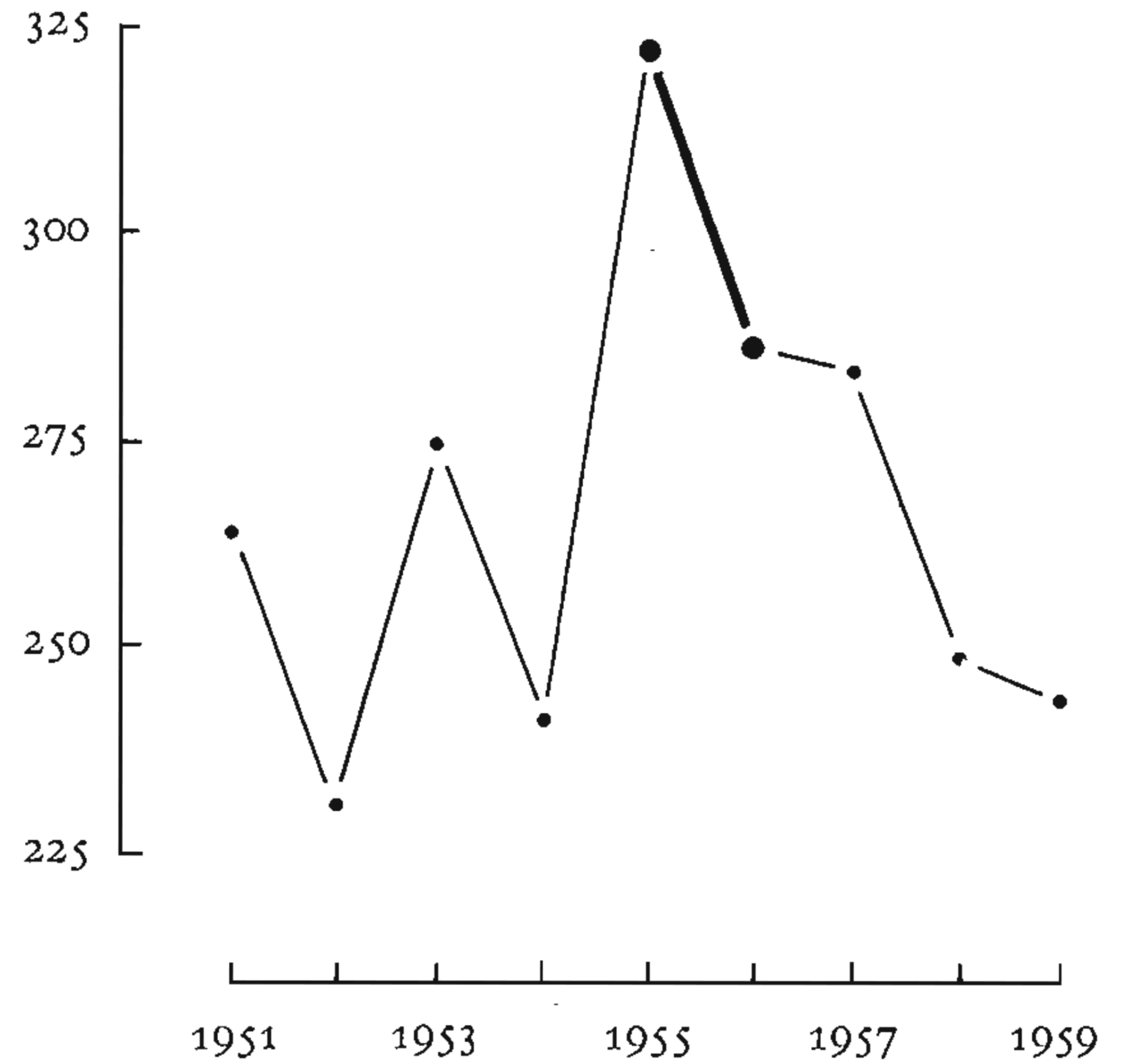


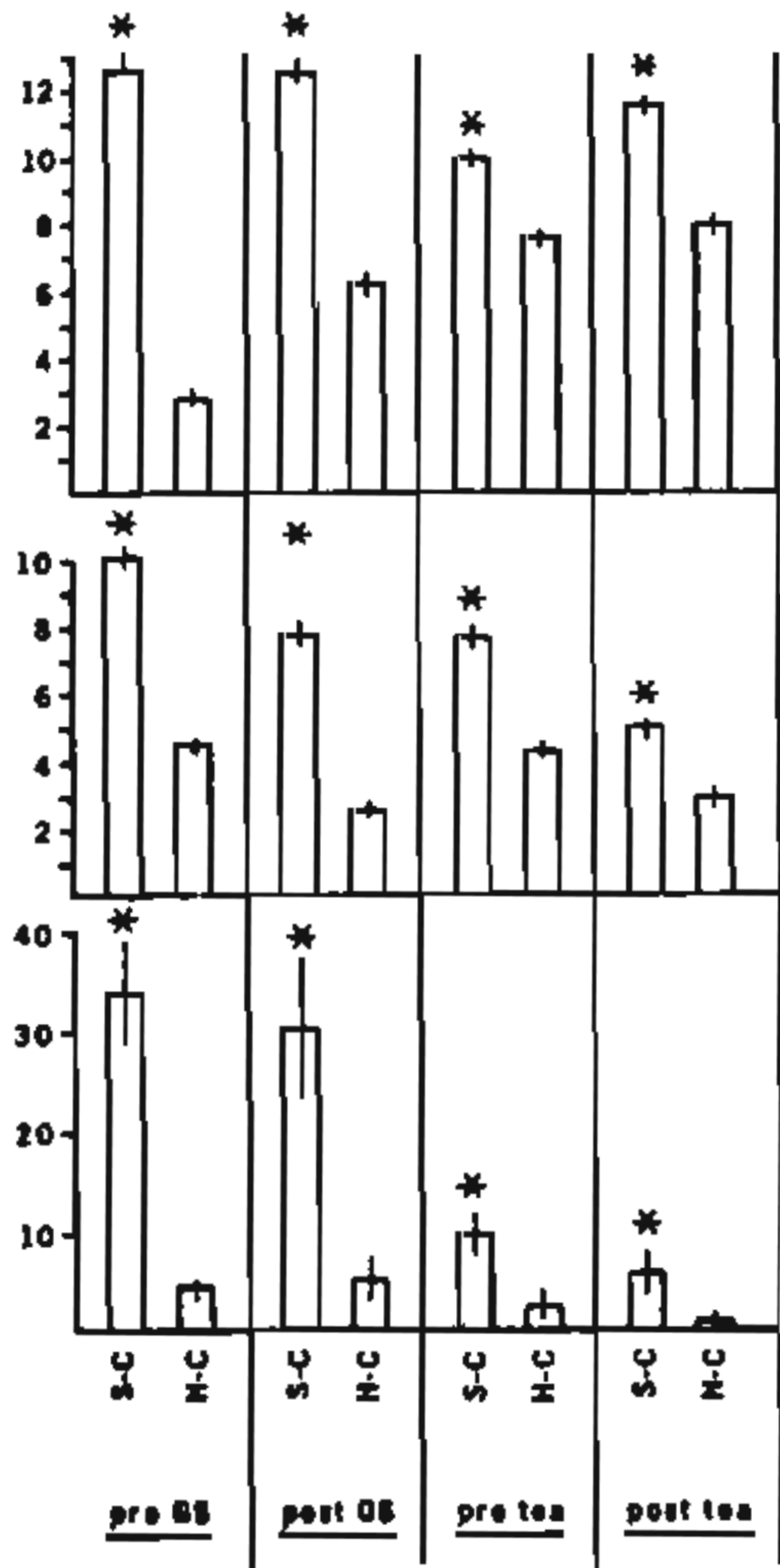
Additional rules of thumb

Data-ink ratio:

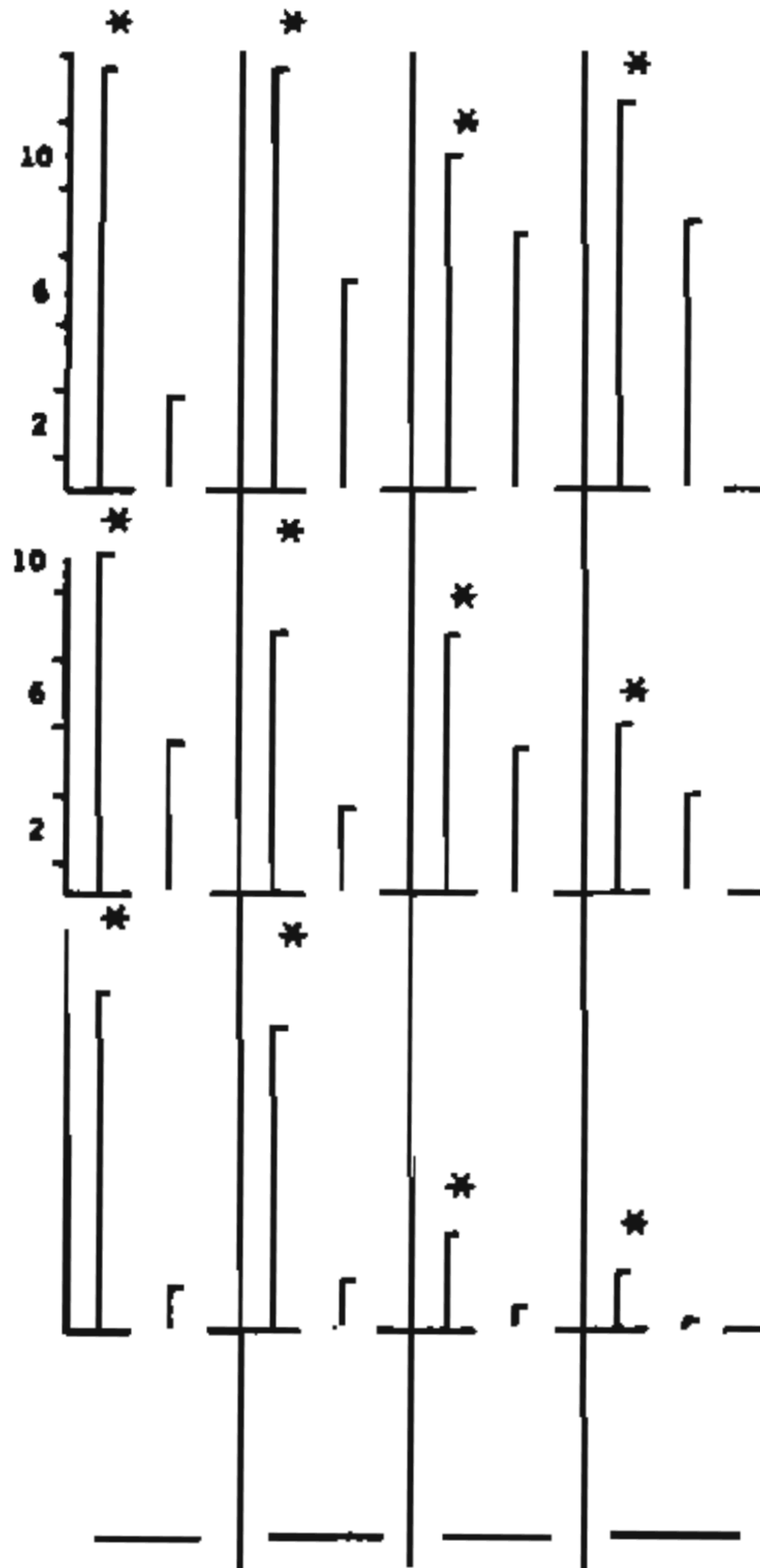
data-ink

total ink used to print the graphics

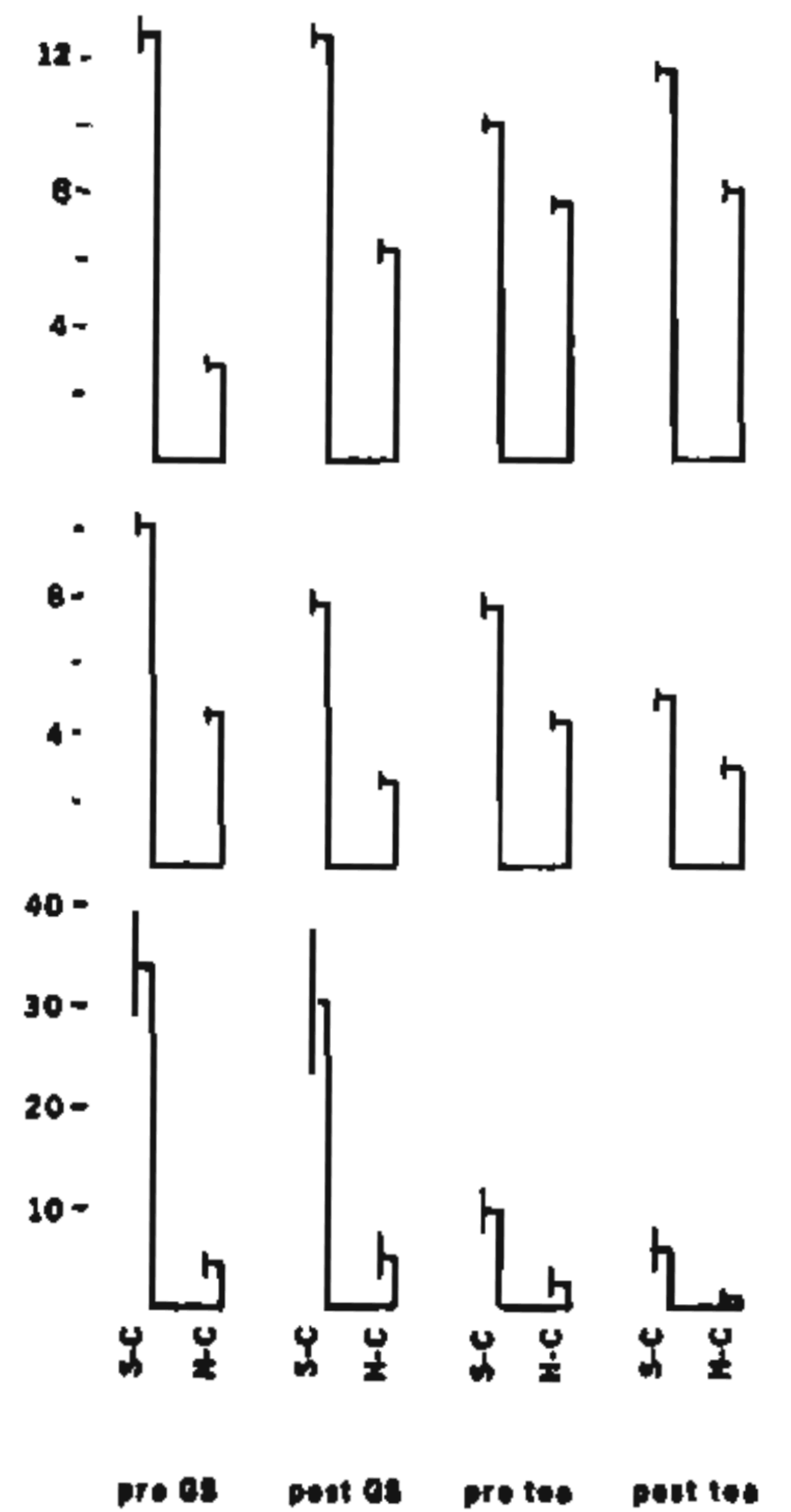




|||

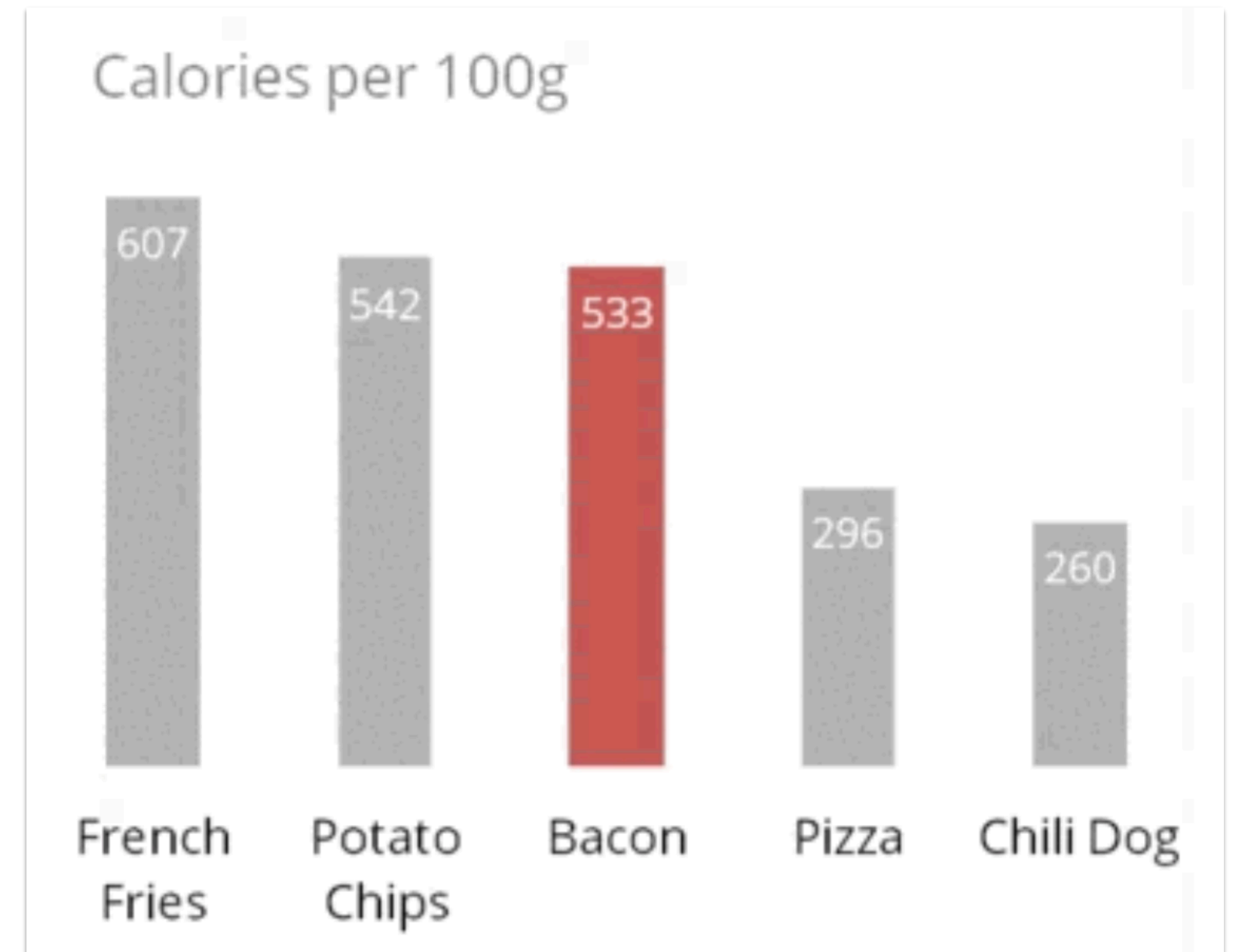
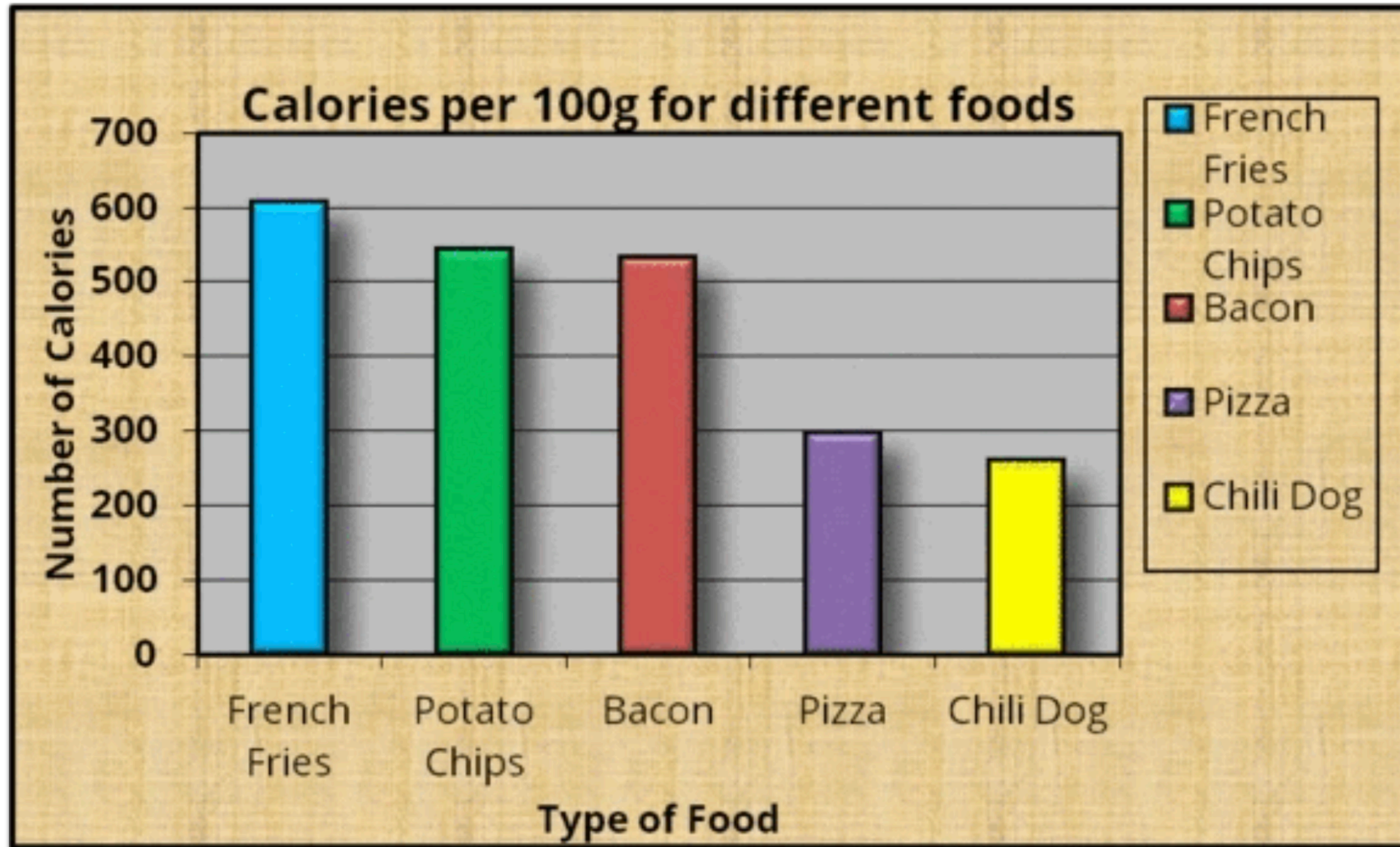


+



Remove
to improve
(the **data-ink** ratio)

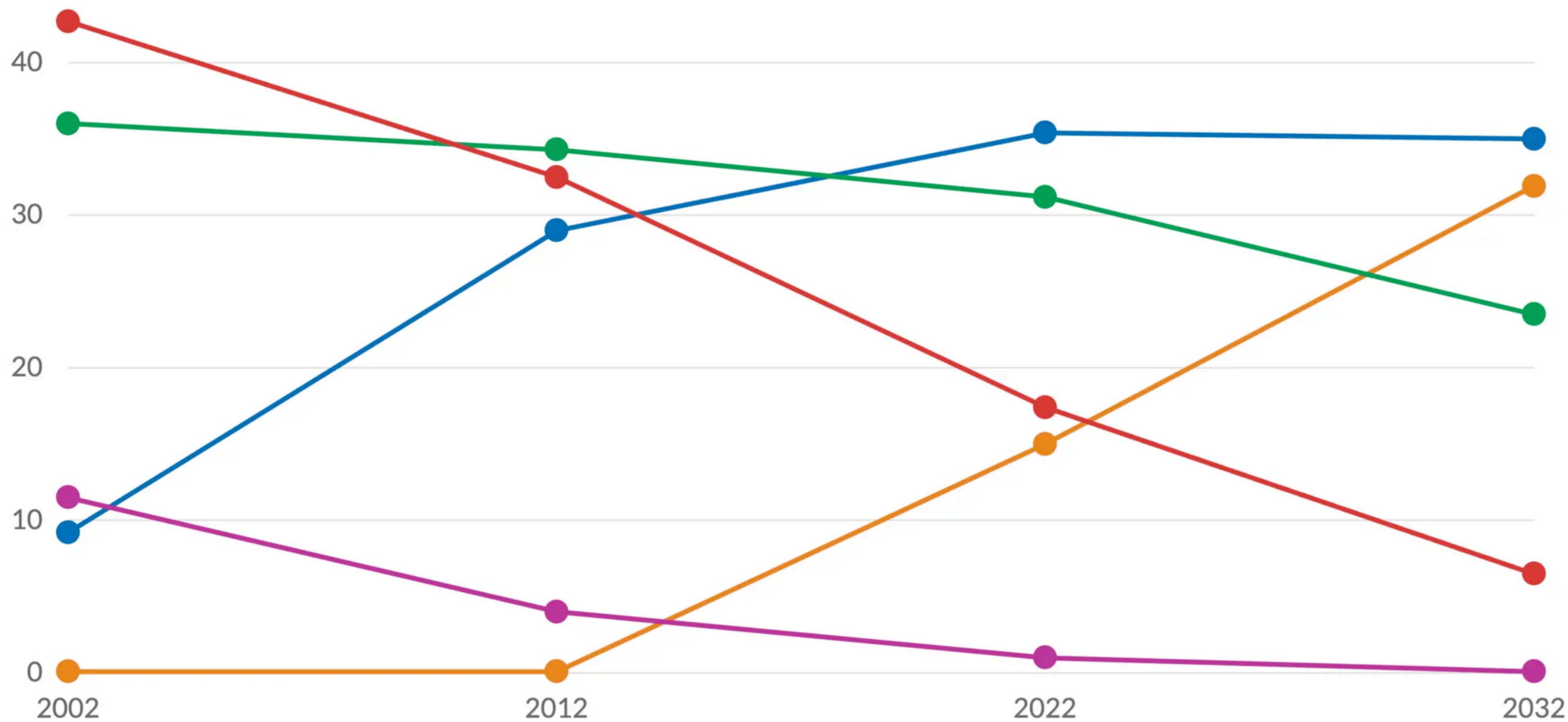
How much non-data-ink do you need?



Generations as percent of labor force

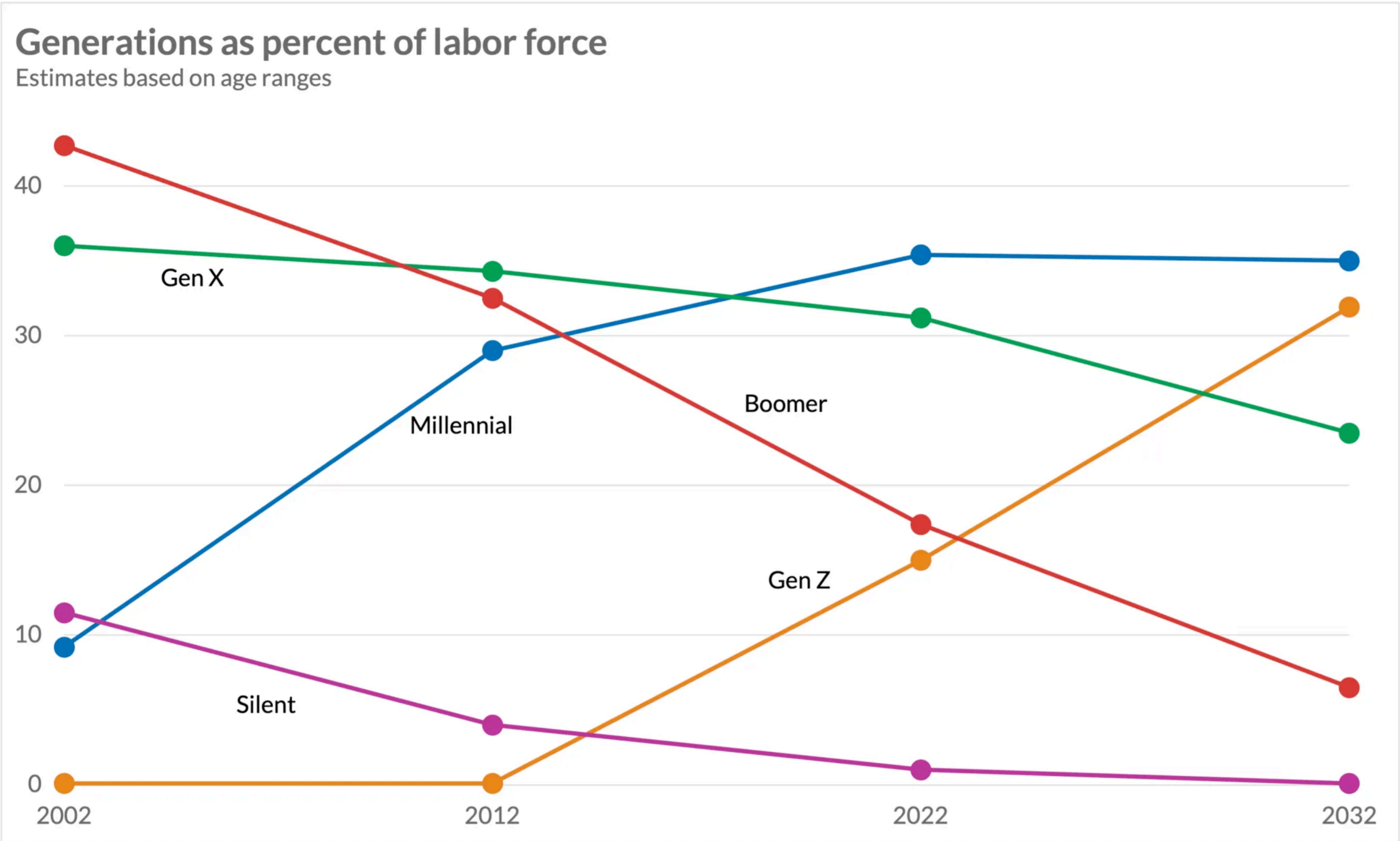
Estimates based on age ranges

Gen Z Millennial Gen X Boomer Silent



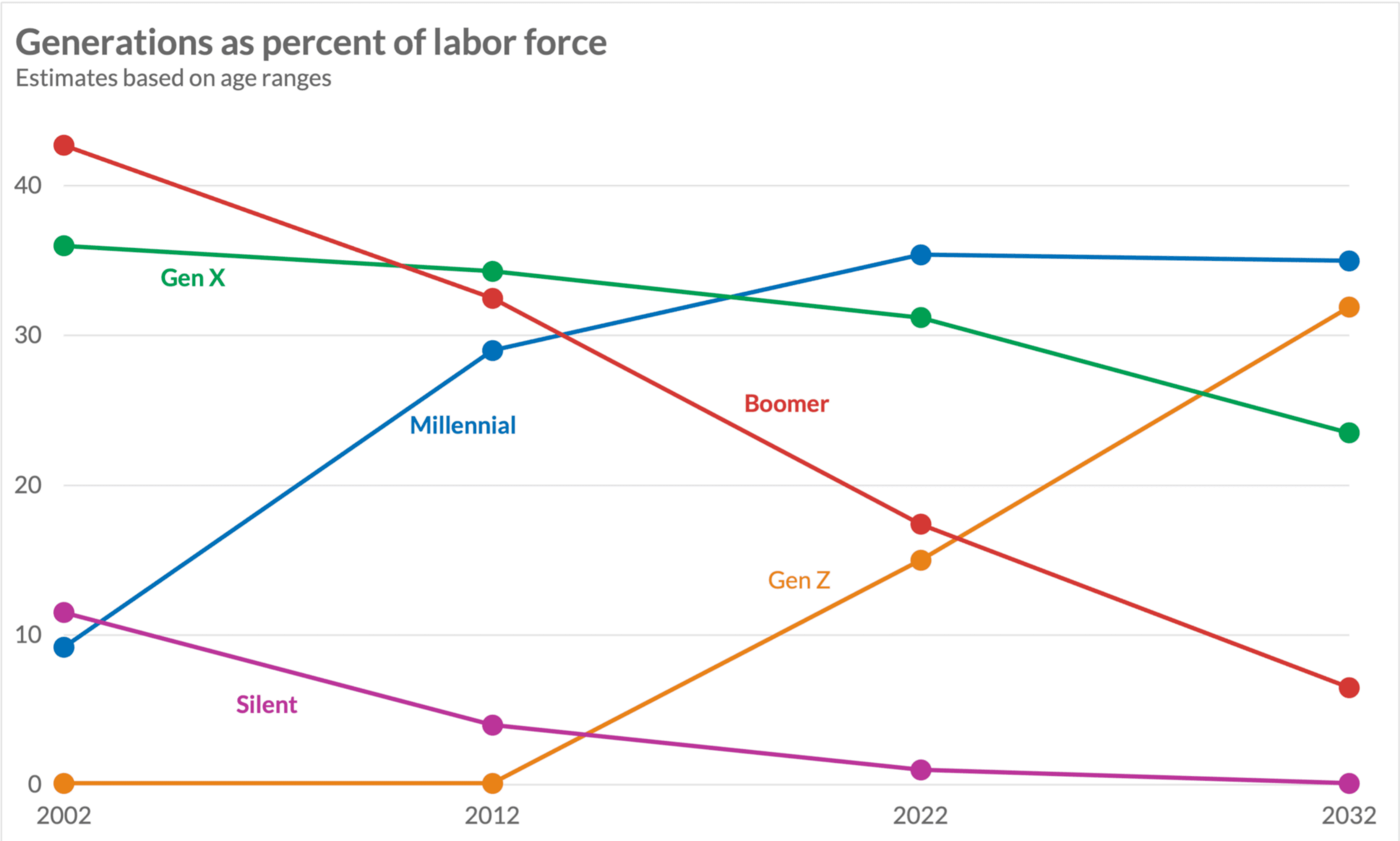
The default legend for line charts.

Can we do better?



✓ **Direct labeling:** reader doesn't need to split attention between the legend and chart

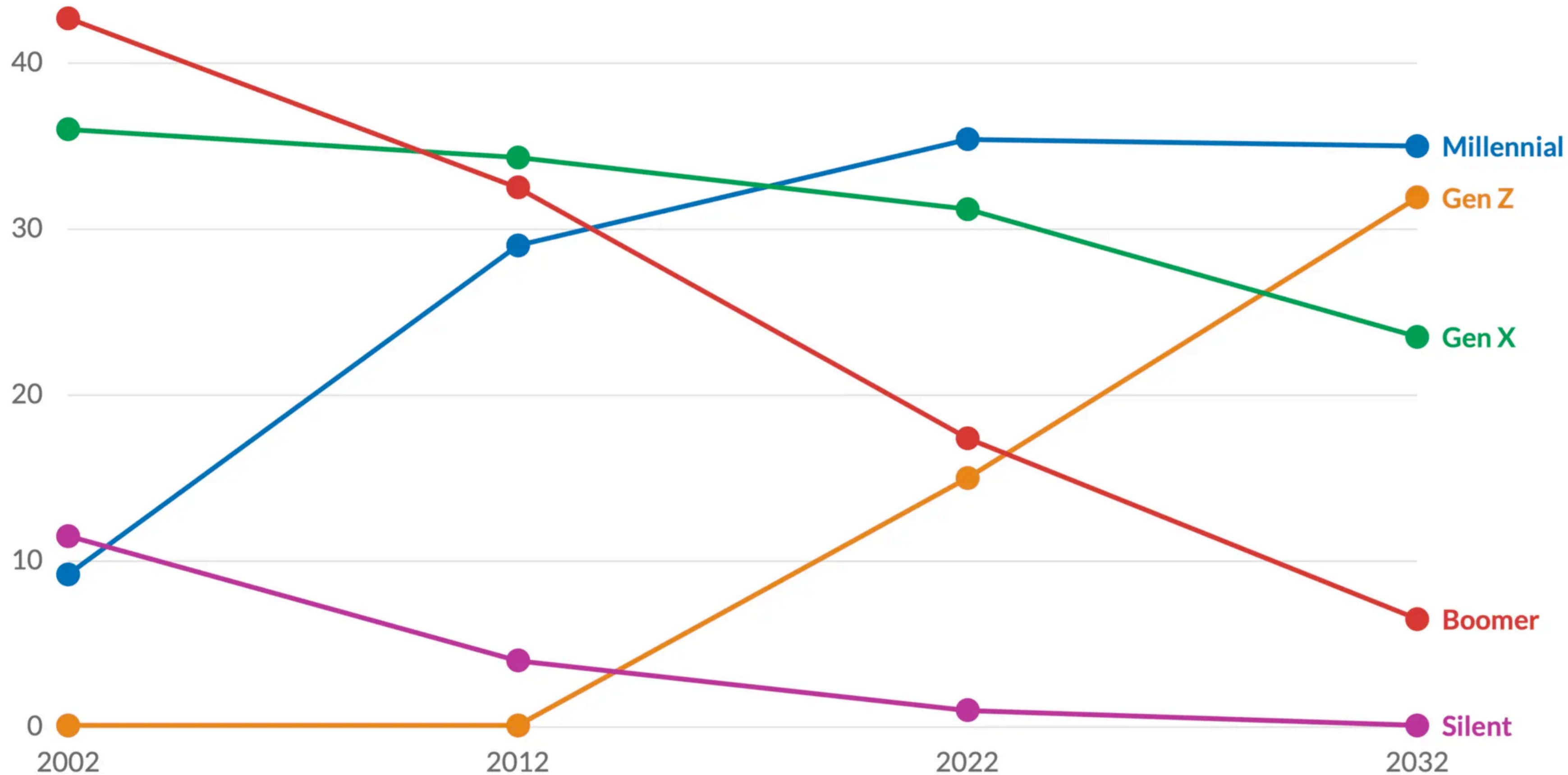
✗ **Label positions are scattered and ambiguous.**



- ✓ **Direct labeling:** reader doesn't need to split attention between the legend and chart
- ✓ **Redundant color encoding** reinforces the relationship between label and line.

Generations as percent of labor force

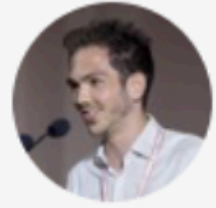
Estimates based on age ranges



✓ **Direct labeling:** reader doesn't need to split attention between the legend and chart

✓ **Redundant color encoding** reinforces the relationship between label and line.

✓ **Spatially align labels** to facilitate reading and skimming.



John Burn-Murdoch @jburnmurdo · Dec 30, 2023 ...
A good chart doesn't just show, it tells.

“Here’s what this means”.

You do this using text. Title, subtitle, labels, annotations. You lose **nothing** by doing this, but add **enormous** value. Your chart is now useful to approximately 1,000x more people than it was before.

4 9 40 4.6K



John Burn-Murdoch @jburnmurdo · Dec 30, 2023 ...
People who think the purpose of charts is to look nice/colourful/cool/beautiful make a very similar mistake.

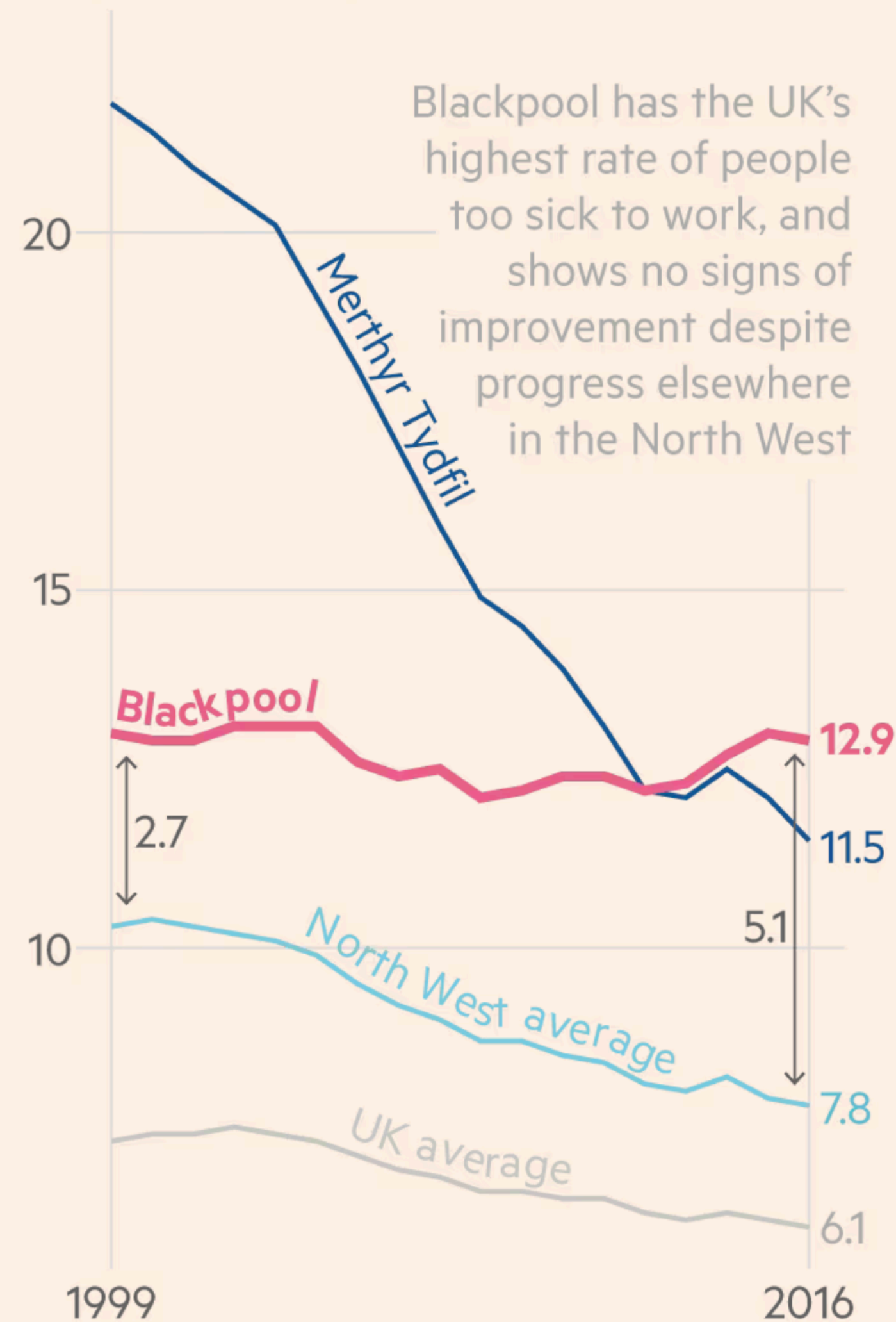
Great that looks lovely but just FYI most people’s takeaway from looking at it is “ooh lovely picture”, not “ah interesting I see”.

I used to make this mistake a lot!

2 16 1.5K

Economically and physically, **Blackpool** is unwell

Percentage of working-age population in receipt of incapacity benefits (IB) or employment support allowance (ESA)



Source: Department for Work and Pensions, Public Health England
Graphic by John Burn-Murdoch / @jburnmurdoch © FT

Blackpool has some of England’s highest rates of obesity and smoking, and the top mortality rate from alcoholic liver disease

