Charts

What are they?

What are they?

Battle tested encodings

Chart choosers

the tree of charts

× Line chart ×

link tags by OR
link tags by AND

refresh

what data | what scale | why | how | what type | data · scale · usage | action · scale · type | scale · action

Number of values the visual can show → What do we want to find → What do we want to do

Line chart

Type: Chart · Family: Line chart

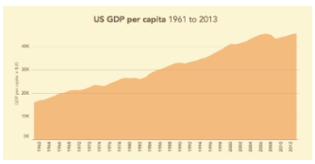
browse tree | more info | web · source

A line chart or line graph is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments. It is a basic type of chart common in many fields. It is similar to a scatter plot except that the measurement points are ordered (typically by their x-axis value) and joined with straight line segments. Line charts can represent different categories of one variable (Multi line chart). Paired with interactivity like selection/highlighting a line chart can show a large number of catgeories (hundreds). However, in its static from we should not exceed 8 categories.

Area chart

Type: Chart · Family: Line chart

browse tree | more info | web · source



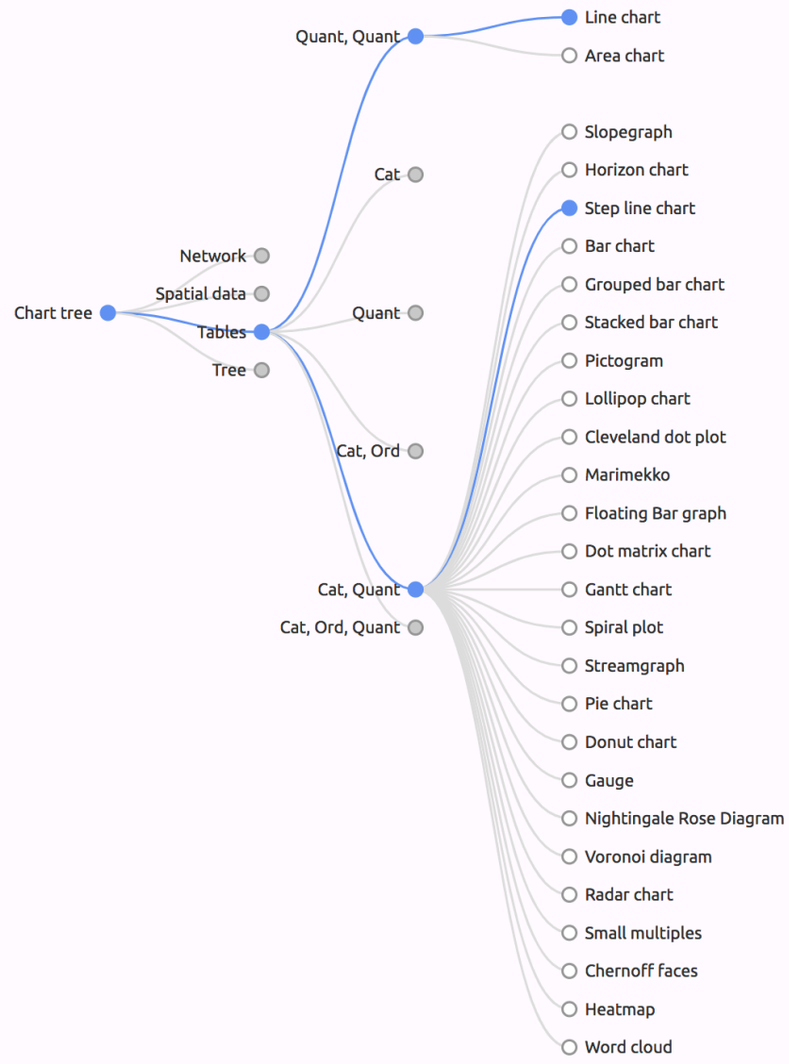
An area chart or area graph displays quantitive data graphically. It is based on the line chart. The area between axis and line are commonly emphasized with colours, textures and/or hatchings. Commonly one compares two or more quantities with an area chart in a stacked design. Possibly not the best of decisions as the eye is forced to calculate each categroy's quantity. It should be noted that it's not very good at that.

Slopegraph

Type: Chart · Family: Line chart

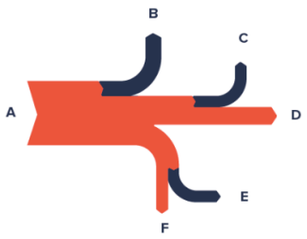
browse tree | more info | web · source

Slopegraphs compare changes usually over time for a list of nouns located on an ordinal or interval scale.

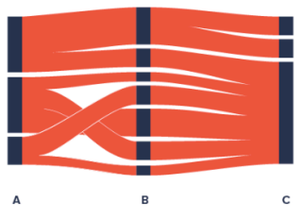


datamake.io/project/tree-of-charts

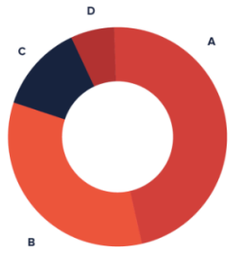
Sankey Diagram



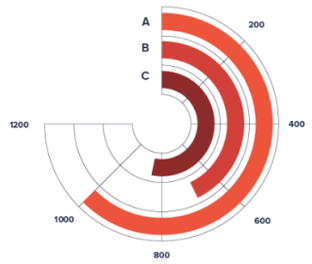
Alluvial Diagram



Donut Chart

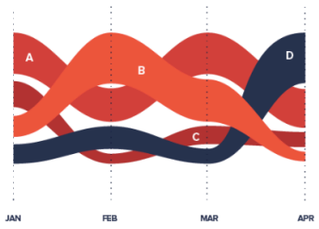


Radial Bar Chart



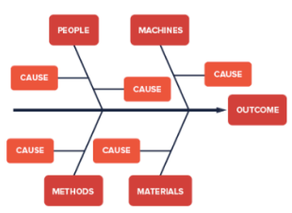
Radial Histogram

Sorted Stream Graph

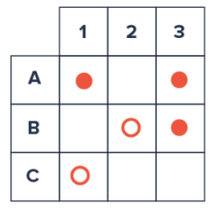


Pictorial fraction chart

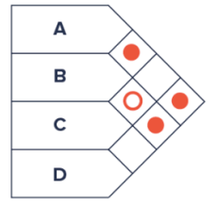
Fishbone Diagram



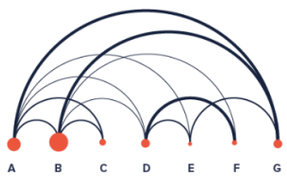
Matrix Diagram



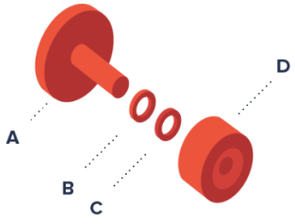
Matrix Diagram (Roof Shaped)



Arc Diagram



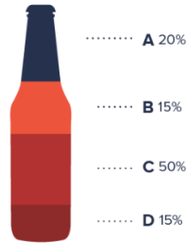
Exploded View Drawing



Flow Map

Polar Area Chart

Pictorial Stacked Chart



https://datavizproject.com

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 Bott; Liz Faunce; Graham Parrish; Billy Ehrenberg-Shannon; Paul McCallum; Martin Stabe
Inspired by the Graphic Continuum by Jon Schwabish and Severino Ribecca

ft.com/vocabulary

Deviation

Emphasise variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a long-term average. Can also be used to show sentiment (positive/neutral/negative).

Example FT uses
Trade surplus/deficit, climate change

Diverging bar

A simple standard bar chart that can handle both negative and positive magnitude values.

Diverging stacked bar

Perfect for presenting survey results which involve sentiment (eg disagree/neutral/agree).

Spine

Splits a single value into two contrasting components (eg male/female).

Surplus/deficit filled line

The shaded area of these charts allows a balance to be shown – either against a baseline or between two series.

Correlation

Show the relationship between two or more variables. Be mindful that, unless you tell them otherwise, many readers will assume the relationships you show them to be causal (i.e. one causes the other).

Example FT uses
Inflation and unemployment, income and life expectancy

Scatterplot

The standard way to show the relationship between two continuous variables, each of which has its own axis.

Column + line timeline

A good way of showing the relationship between an amount (columns) and a rate (line).

Connected scatterplot

Usually used to show how the relationship between 2 variables has changed over time.

Bubble

Like a scatterplot, but adds additional detail by sizing the circles according to a third variable.

XY heatmap

A good way of showing the patterns between 2 categories of data, less effective at showing fine differences in amounts.

Ranking

Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.

Example FT uses
Wealth, deprivation, league tables, constituency election results

Ordered bar

Standard bar charts display the ranks of values much more easily when sorted into order.

Ordered column

See above.

Ordered proportional symbol

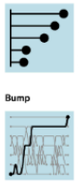
Use when there are big variations between values and/or seeing fine differences between data is not so important.

Dot strip plot

Dots placed in order on a strip are a space-efficient method of laying out ranks across multiple categories.

Slope

Perfect for showing how ranks have changed over time or vary between categories.

Lollipop

Lollipops draw more attention to the data value than standard bar/column and can also show rank and value effectively.

Bump

Effective for showing changing rankings across multiple dates. For large datasets, consider grouping lines using colour.

Distribution

Show values in a dataset and how often they occur. The shape (or 'skew') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.

Example FT uses
Income distribution, population (age/sex) distribution, revealing inequality

Histogram

The standard way to show a statistical distribution - keep the gaps between columns small to highlight the 'shape' of the data.

Dot plot

A simple way of showing the change or range (min/max) of data across multiple categories.

Dot strip plot

Good for showing individual values in a distribution, can be a problem when too many dots have the same value.

Barcode plot

Like dot strip plots, good for displaying all the data in a table, they work best when highlighting individual values.

Boxplot

Summarise multiple distributions by showing the median (centre) and range of the data

Violin plot

Similar to a box plot but more effective with complex distributions (data that cannot be summarised with simple average).

Population pyramid

A standard way for showing the age and sex breakdown of a population distribution; effectively, back to back histograms.

Cumulative curve

A good way of showing how unequal a distribution is: y axis is always cumulative frequency, x axis is always a measure.

Frequency polygons

For displaying multiple distributions of data. Like a regular line chart, best limited to a maximum of 3 or 4 datasets.

Beeswarm

Use to emphasise individual points in a distribution. Points can be sized to an additional variable. Best with medium-sized datasets

Change over Time

Give emphasis to changing trends. These can be short (intra-day) movements or extended series traversing decades or centuries: Choosing the correct time period is important to provide suitable context for the reader.

Example FT uses
Share price movements, economic time series, sectoral changes in a market

Line

The standard way to show a changing time series. If data are irregular, consider markers to represent data points.

Column

Columns work well for showing change over time – but usually best with only one series of data at a time.

Column + line timeline

A good way of showing the relationship over time between an amount (columns) and a rate (line).

Slope

Good for showing changing data as long as the data can be simplified into 2 or 3 points without missing a key part of story.

Area chart

Use with care – these are good at showing changes to total, but seeing change in components can be very difficult.

Candlestick

Usually focused on day-to-day activity, these charts show opening/closing and high/low points of each day.

Fan chart (projections)

Use to show the uncertainty in future projections – usually this grows the further forward to projection.

Connected scatterplot

A good way of showing changing data for two variables whenever there is a relatively clear pattern of progression.

Calendar heatmap

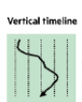
A great way of showing temporal patterns (daily, weekly, monthly) – at the expense of showing precision in quantity.

Priestley timeline

Great when date and duration are key elements of the story in the data.

Circle timeline

Good for showing discrete values of varying size across multiple categories (eg earthquakes by continent).

Vertical timeline

Presents time on the Y axis. Good for displaying detailed time series that work especially well when scrolling on mobile.

Seismogram

Another alternative to the circle timeline for showing series where there are big variations in the data.

Streamgraph

A type of area chart; use when seeing changes in proportions over time is more important than individual values

Magnitude

Show size comparisons. These can be relative (just being able to see larger/bigger) or absolute (need to see fine differences). Usually these show a 'counted' number (for example, barrels, dollars or people) rather than a calculated rate or per cent.

Example FT uses
Commodity production, market capitalisation, volumes in general

Column

The standard way to compare the size of things. Must always start at 0 on the axis.

Bar

See above. Good when the data are not time series and labels have long category names.

Paired column

As per standard column but allows for multiple series. Can become tricky to read with more than 2 series.

Paired bar

See above.

Marimekko

A good way of showing the size and proportion of data at the same time – as long as the data are not too complicated.

Proportional symbol

Use when there are big variations between values and/or seeing fine differences between data is not so important.

Isotype (pictogram)

Excellent solution in some instances – use only with whole numbers (do not slice off an arm to represent a decimal).

Lollipop

Lollipop charts draw more attention to the data value than standard bar/column – does not have to start at zero (but preferable).

Radar

A space-efficient way of showing value of multiple variables– but make sure they are organised in a way that makes sense to reader.

Parallel coordinates

An alternative to radar charts – again, the arrangement of the variables is important. Usually benefits from highlighting values.

Bullet

Good for showing a measurement against the context of a target or performance range.

Grouped symbol

An alternative to bar/column charts when being able to count data or highlight individual elements is useful.

Part-to-whole

Show how a single entity can be broken down into its component elements. If the reader's interest is solely in the size of the components, consider a magnitude-type chart instead.

Example FT uses
Fiscal budgets, company structures, national election results

Stacked column/bar

A simple way of showing part-to-whole relationships but can be difficult to read with more than a few components.

Marimekko

A good way of showing the size and proportion of data at the same time – as long as the data are not too complicated.

Pie

A common way of showing part-to-whole data – but be aware that it's difficult to accurately compare the size of the segments.

Donut

Similar to a pie chart – but the centre can be a good way of making space to include more information about the data (eg total).

Treemap

Use for hierarchical part-to-whole relationships; can be difficult to read when there are many small segments.

Voronoi

A way of turning points into areas – any point within each area is closer to the central point than any other centroid.

Arc

A hemicycle, often used for visualising parliamentary composition by number of seats.

Gridplot

Good for showing % information, they work best when used on whole numbers and work well in small multiple layout form.

Venn

Generally only used for schematic representation.

Waterfall

Can be useful for showing part-to-whole relationships where some of the components are negative.

Spatial

Aside from locator maps only used when precise locations or geographical patterns in data are more important to the reader than anything else.

Example FT uses
Population density, natural resource locations, natural disaster risk/impact, catchment areas, variation in election results

Basic choropleth (rate/ratio)

The standard approach for putting data on a map – should always be rates rather than totals and use a sensible base geography.

Proportional symbol (count/magnitude)

Use for totals rather than rates – be wary that small differences in data will be hard to see.

Flow map

For showing unambiguous movement across a map.

Contour map

For showing areas of equal value on a map. Can use deviation colour schemes for showing +/- values

Equalised cartogram

Converting each unit on a map to a regular and equally-sized shape – good for representing voting regions with equal value.

Scaled cartogram (value)

Stretching and shrinking a map so that each area is sized according to a particular value.

Dot density

Used to show the location of individual events/locations – make sure to annotate any patterns the reader should see.

Heat map

Grid-based data values mapped with an intensity colour scale. As choropleth map – but not snapped to an admin/political unit.

Flow

Show the reader volumes or intensity of movement between two or more states or conditions. These might be logical sequences or geographical locations.

Example FT uses
Movement of funds, trade, migrants, lawsuits, information; relationship graphs.

Sankey

Shows changes in flows from one condition to at least one other; good for tracing the eventual outcome of a complex process.

Waterfall

Designed to show the sequencing of data through a flow process, typically budgets. Can include +/- components.

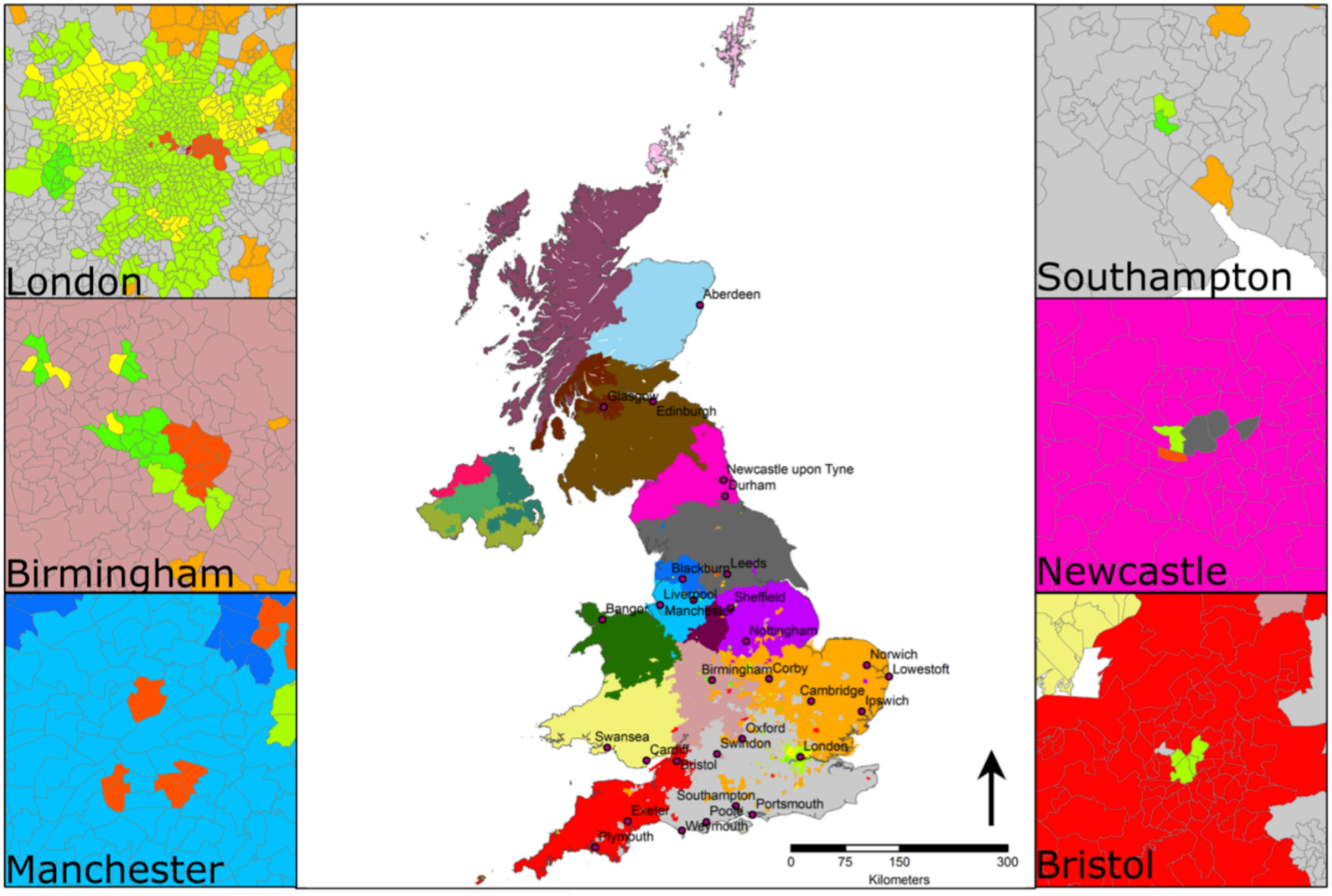
Chord

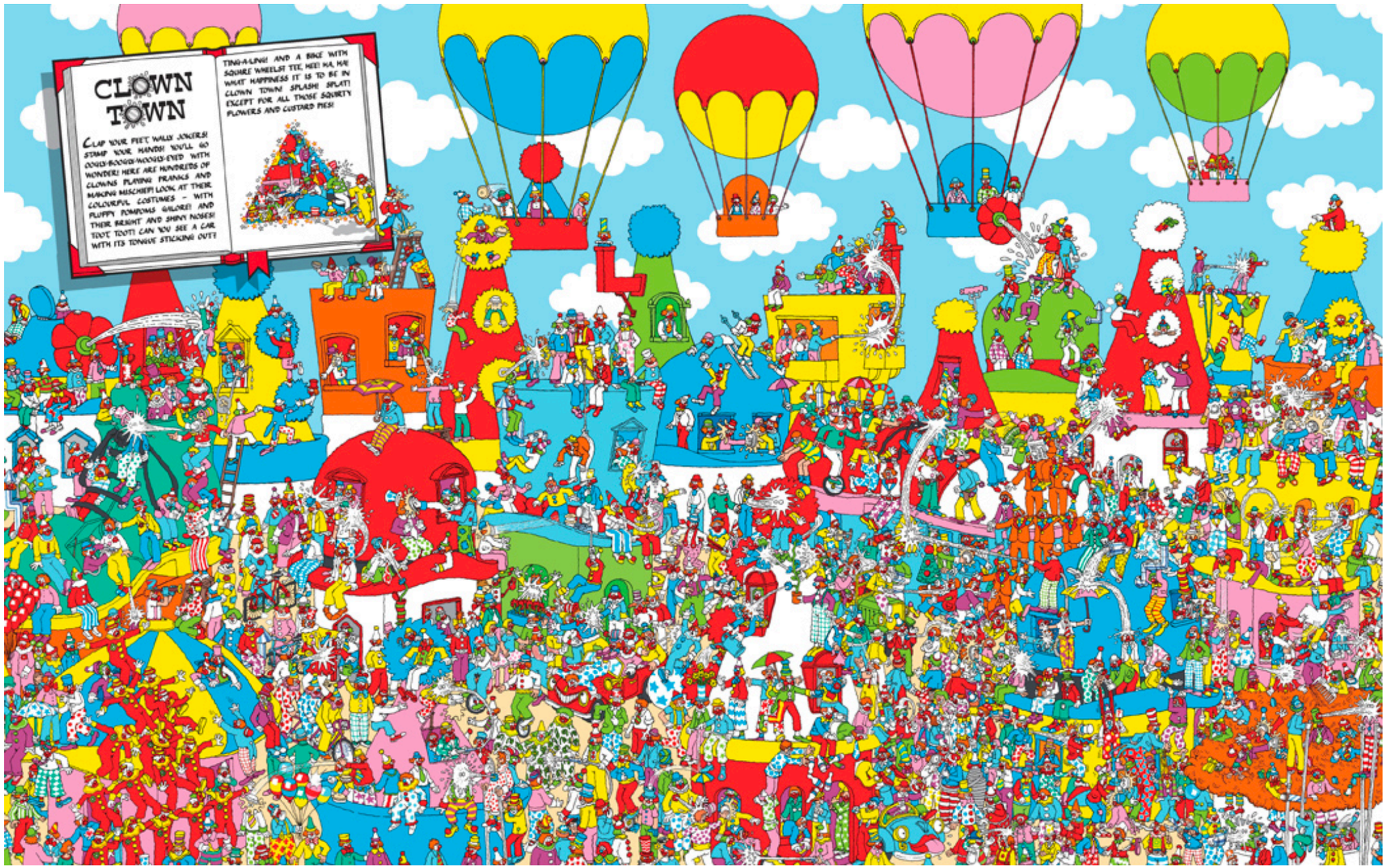
A complex but powerful diagram which can illustrate 2-way flows (and net winner) in a matrix.

Network

Used for showing the strength and inter-connectedness of relationships of varying types.

10 Pitfalls


London
Birmingham
Manchester
Southampton
Newcastle
Bristol
Aberdeen
Glasgow
Edinburgh
Newcastle upon Tyne
Durham
Blackburn
Leeds
Liverpool
Sheffield
Bangor
Manchester
Nottingham
Norwich
Lowestoft
Birmingham
Corby
Cambridge
Ipswich
Oxford
Swansea
Swindon
London
Cardiff
Bristol
Southampton
Portsmouth
Exeter
Poole
Weymouth
Plymouth
0
75
150
300
Kilometers

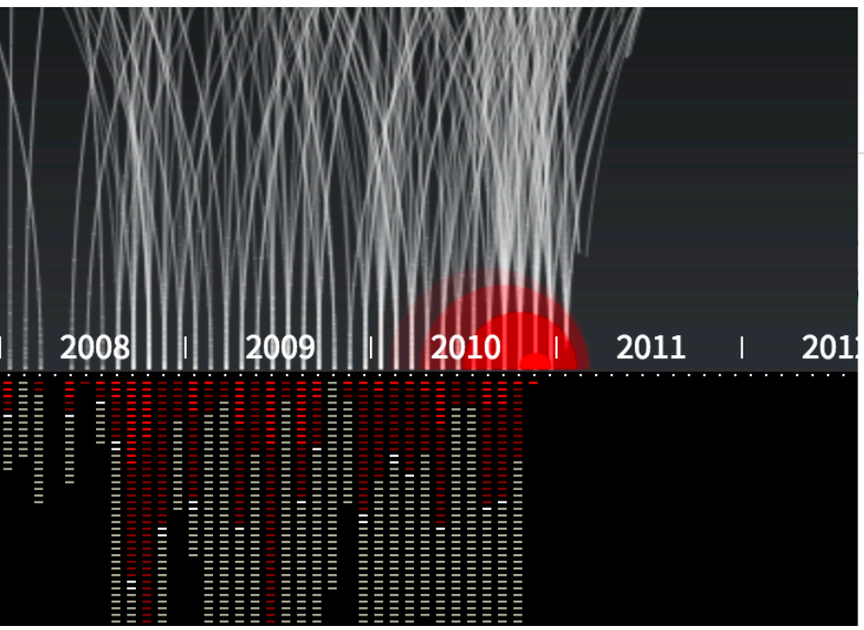

CLOWN TOWN
CLAP YOUR FEET WILLY JOGERS! STAMP YOUR HANDS! YOU'LL GO OOGLY-BOOGLY-WOOGLY-EYED WITH WONDER! HERE ARE HUNDREDS OF CLOWNS PLAYING PRANKS AND MAKING MISCHIEF! LOOK AT THEIR COLOURFUL COSTUMES – WITH FLUFFY POMPOMS GALORE! AND THEIR BRIGHT AND SHINY NOSES! TOOT TOOT! CAN YOU SEE A CAR WITH ITS TONGUE STICKING OUT?
TING-A-LING! AND A BIKE WITH SQUARE WHEELS! TEE HEE! HA, HA! WHAT HAPPINESS IT IS TO BE IN CLOWN TOWN! SPLASH! SPLAT! EXCEPT FOR ALL THOSE SQUIRTY FLOWERS AND CUSTARD PIES!

1. Make grey your best friend

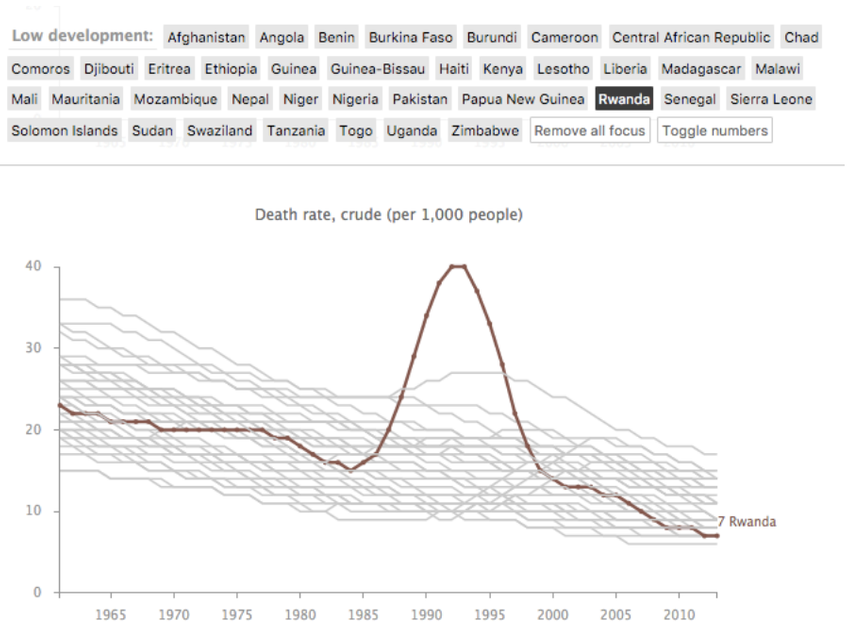
New York Times, March 2016

3 uses of colour in visualisation

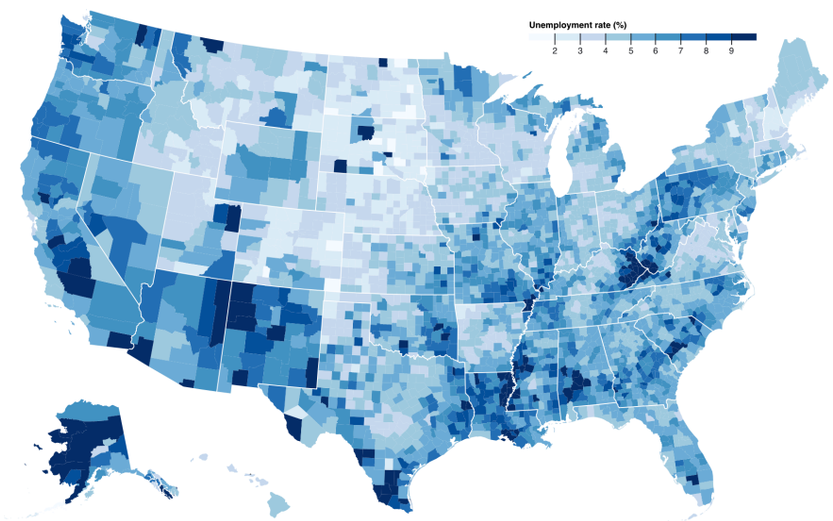
Set the mood



Highlight | Tell Story



Encode data

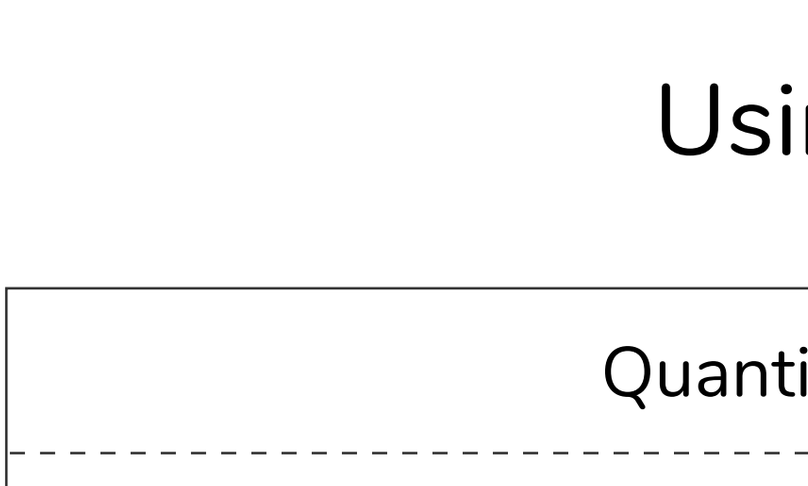


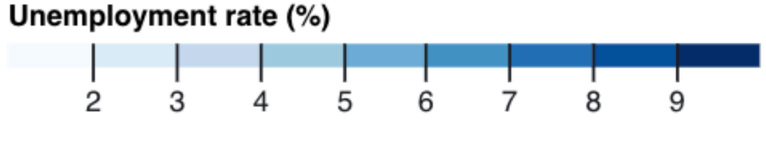
Using colour to encode data

Quantitative data

Sequential scales

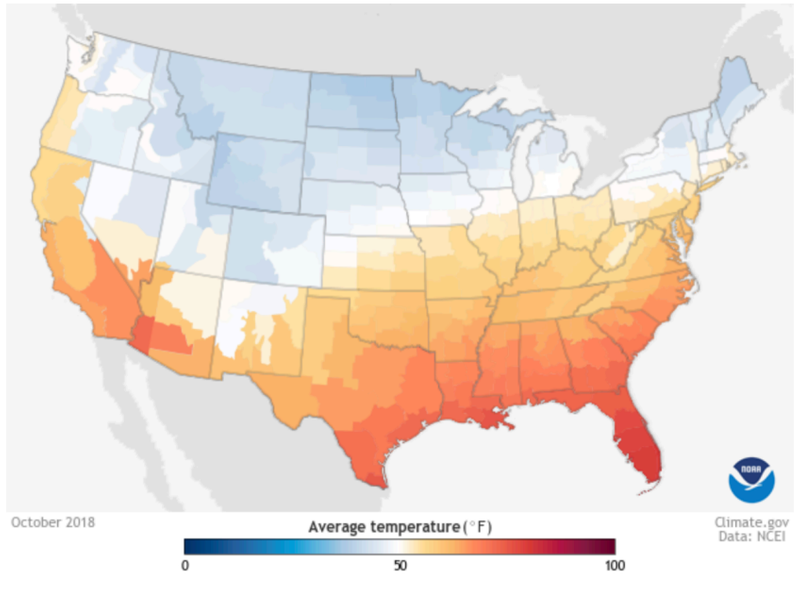
Map low to high **values**
to low to high **luminance**

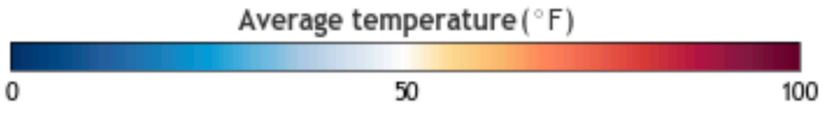




Divergent scales

Map **values with a mid-point**
to **diverging luminance**

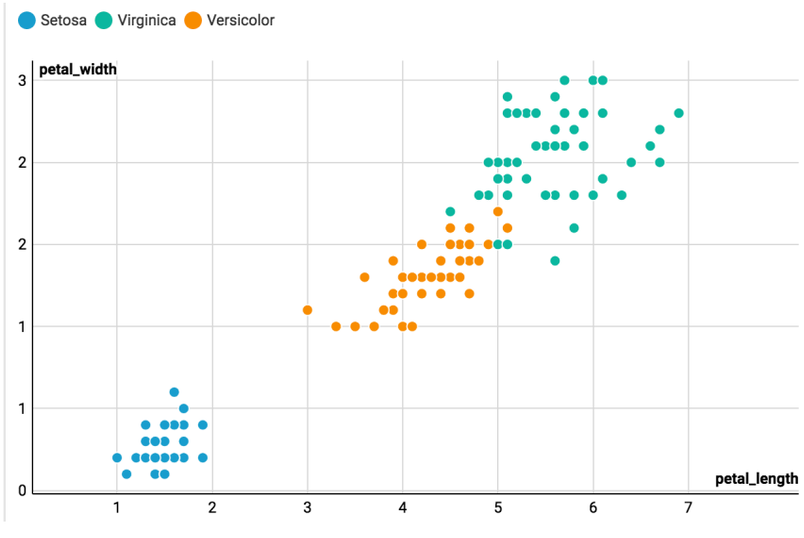




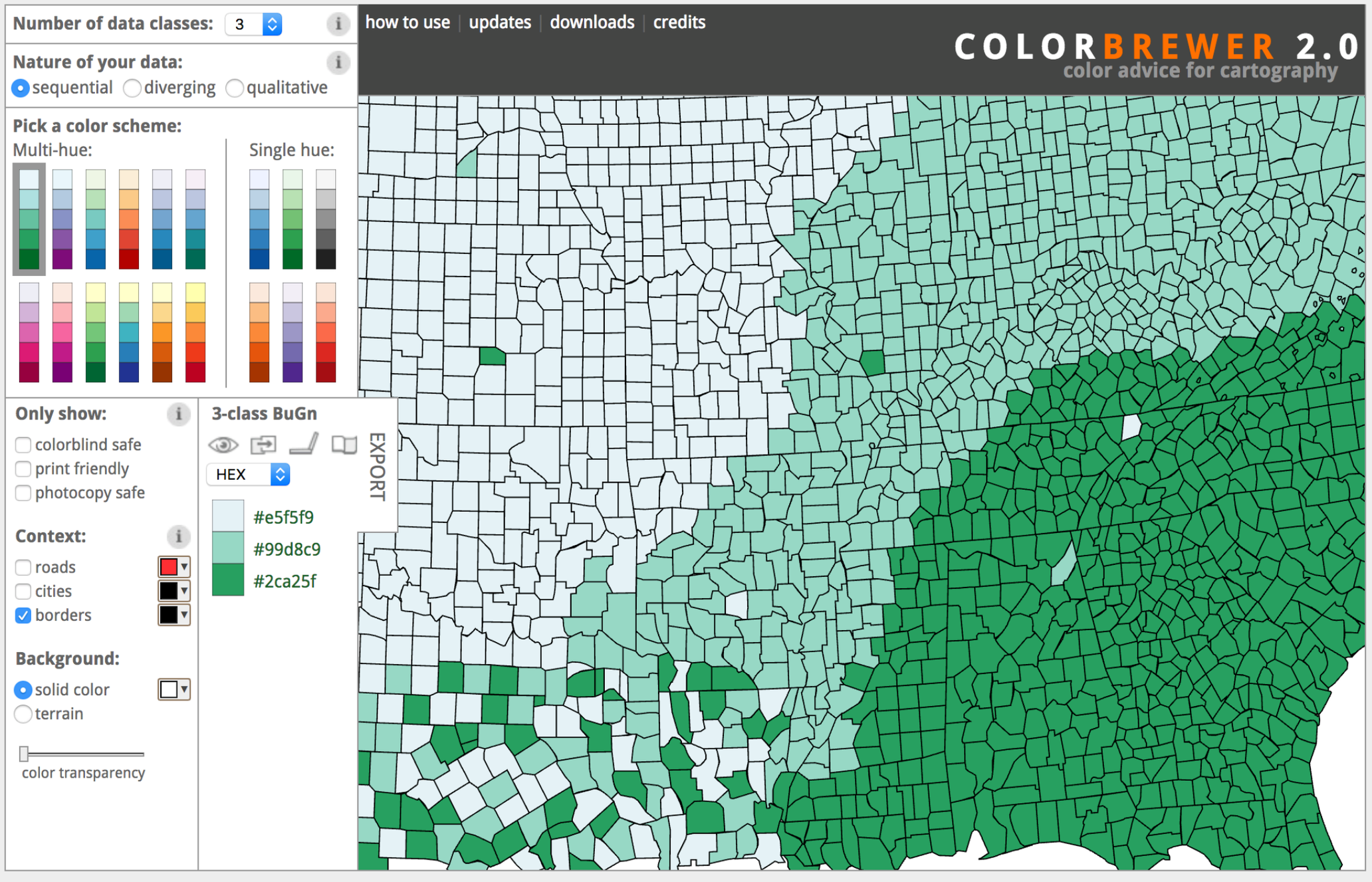
Qualitative data

Qualitative scales

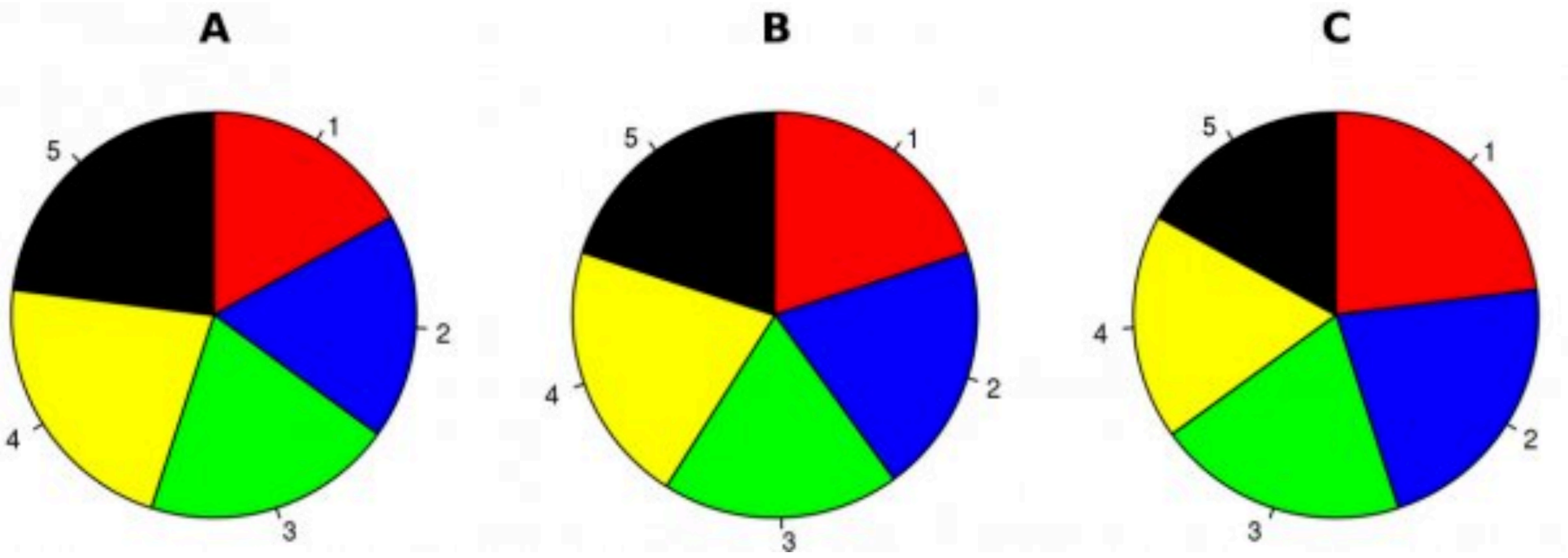
Map **categories**
to **distinct hues**



Help is nigh



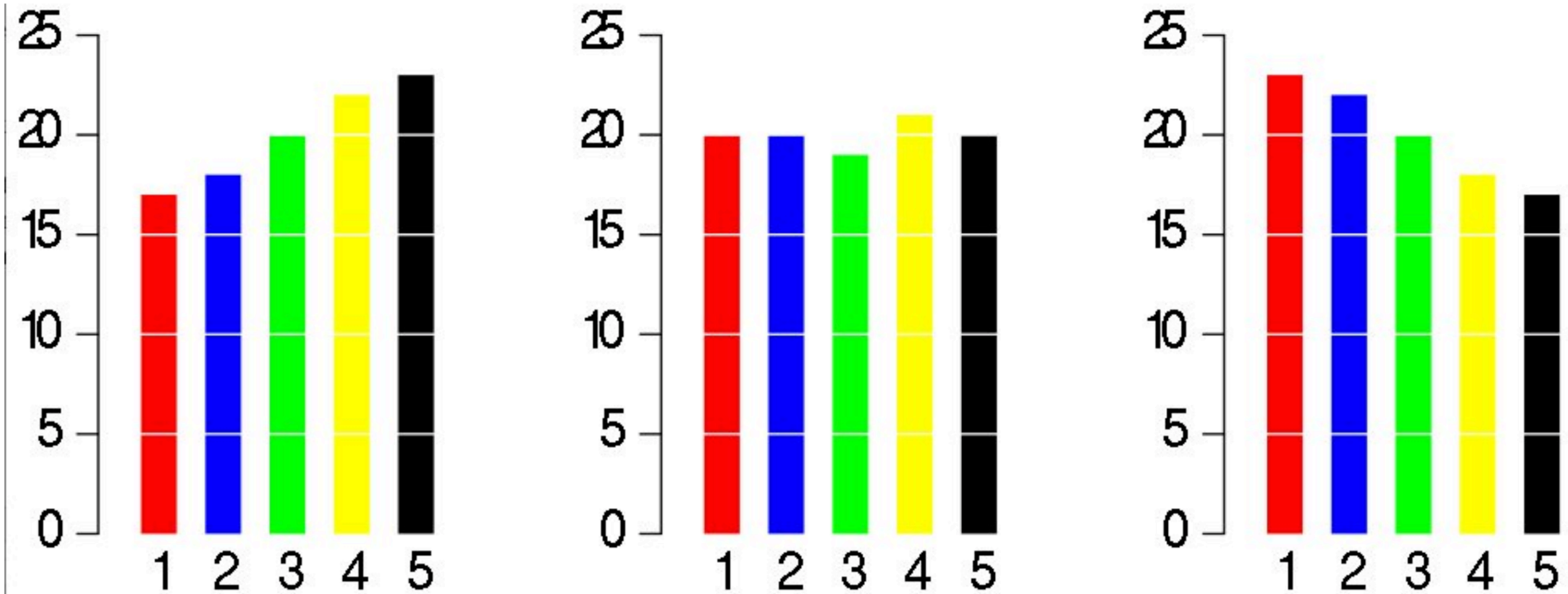
More tools: blog.datawrapper.de/colorguide



In A: which category is 1st, 2nd, 3rd?

Is B5 bigger or smaller than C5?

Which 3 is the smallest?



In A: which category is 1st, 2nd, 3rd?

Is B5 bigger or smaller than C5?

Which 3 is the smallest?

Position on common scale

Position on unaligned scale

Length (1D size)

Tilt/angle

Area (2D size)

Depth (3D position)

Color luminance

Color saturation

Curvature

Volume (3D size)

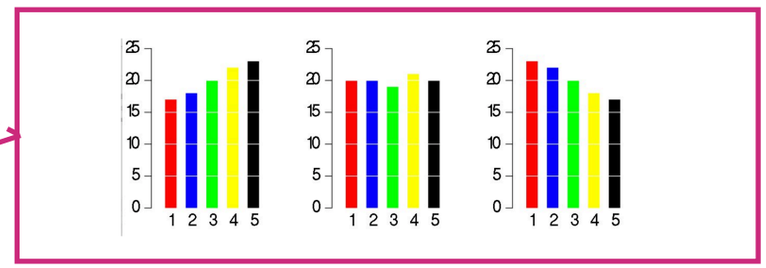
Most

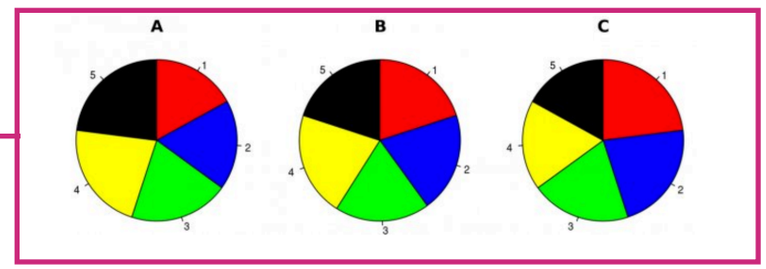
Same

Same

Least

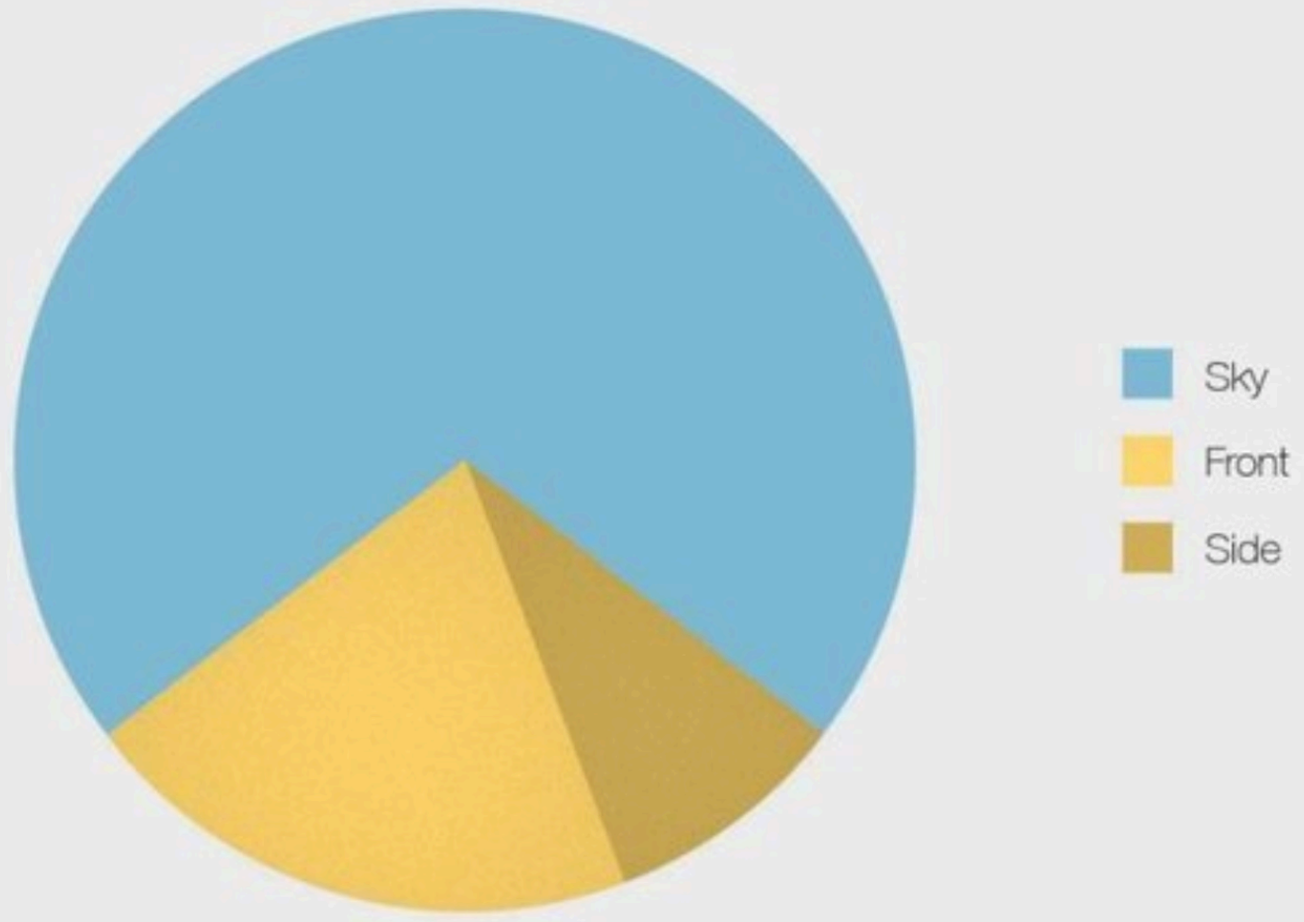
Effectiveness

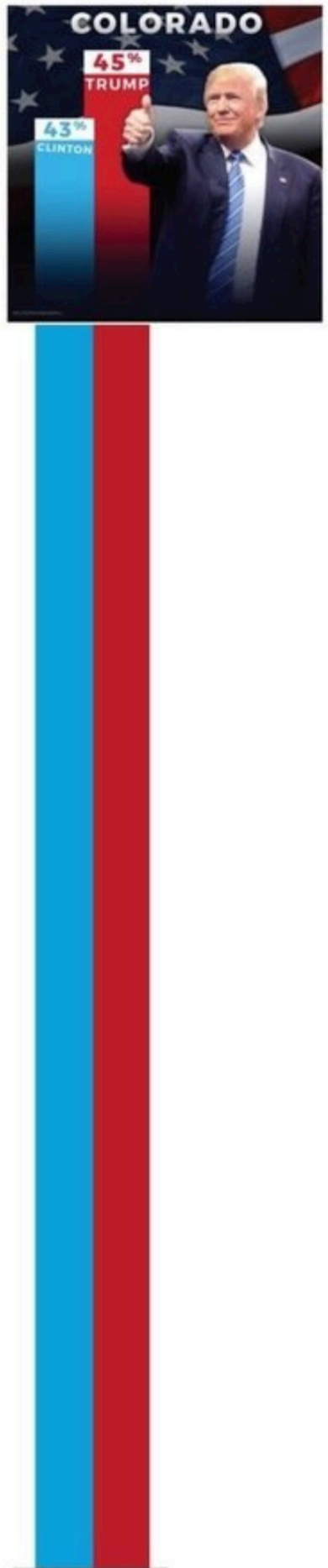




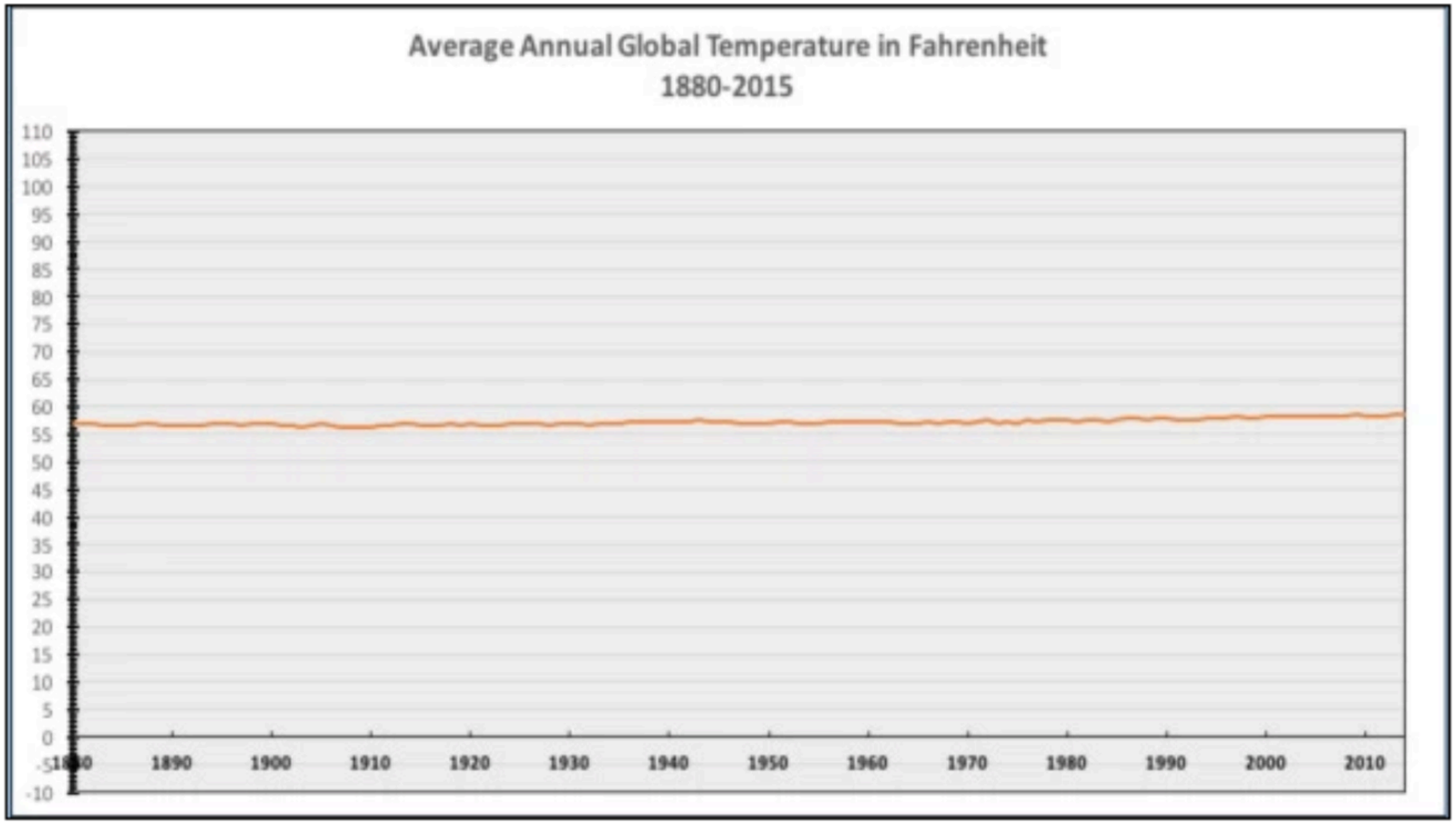
2. Save the pies for dessert

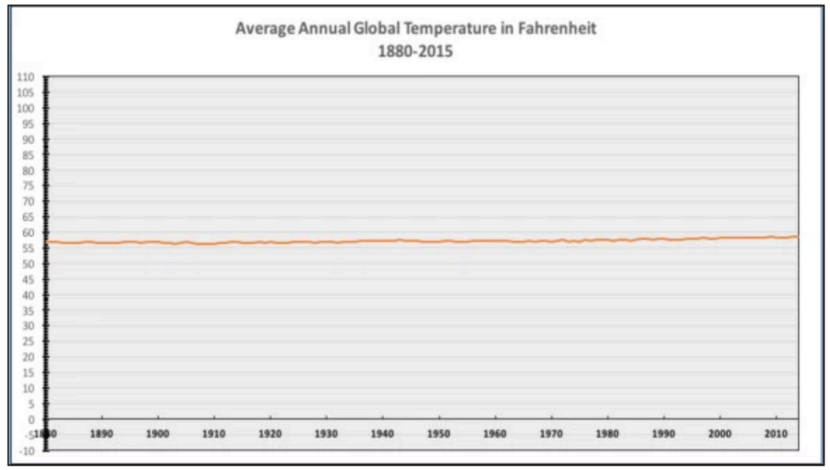
2. Save the pies for the desert





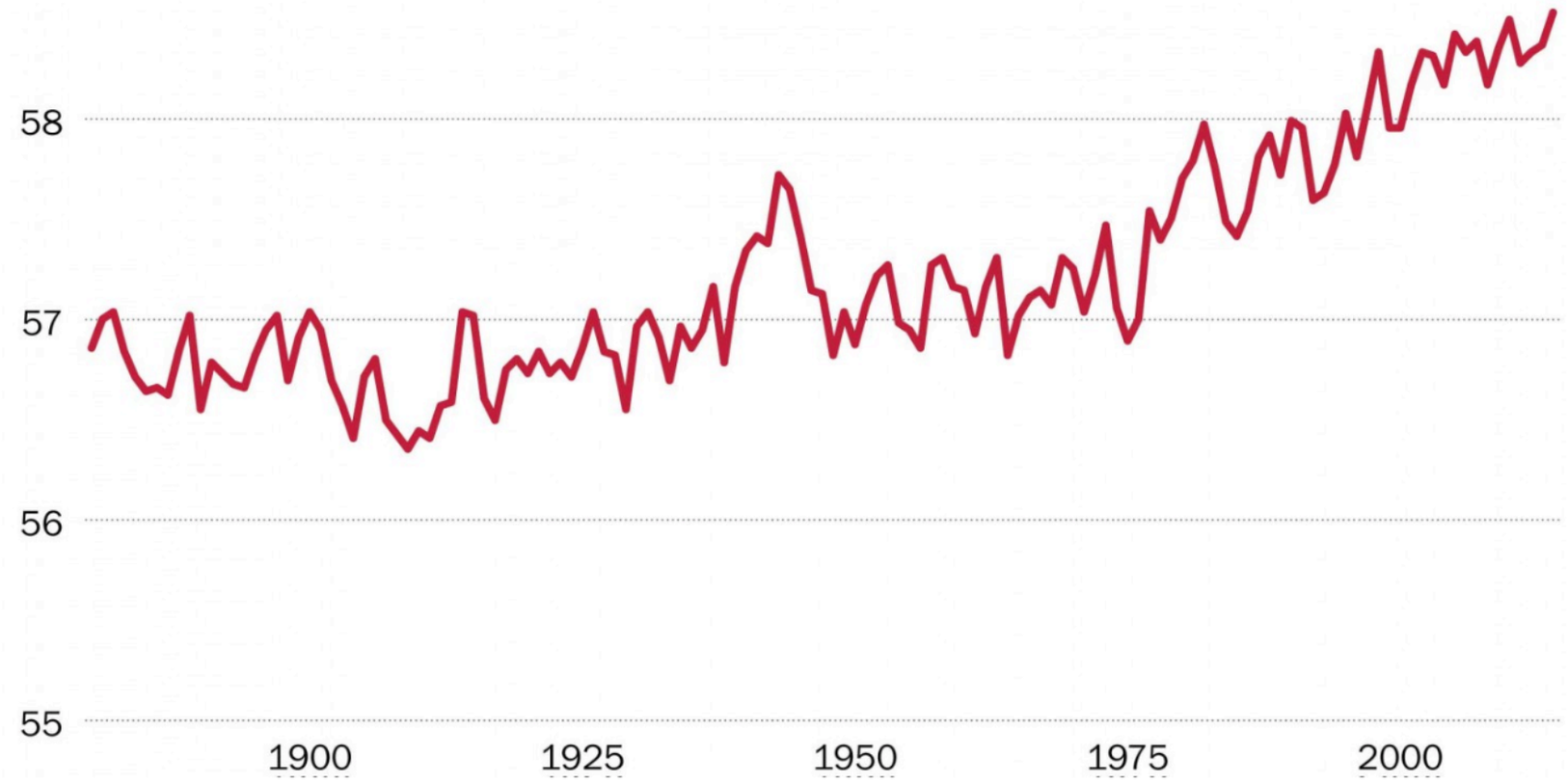
3. Don’t cut bars


Average Annual Global Temperature in Fahrenheit
1880-2015
110
105
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
-5
-10
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010



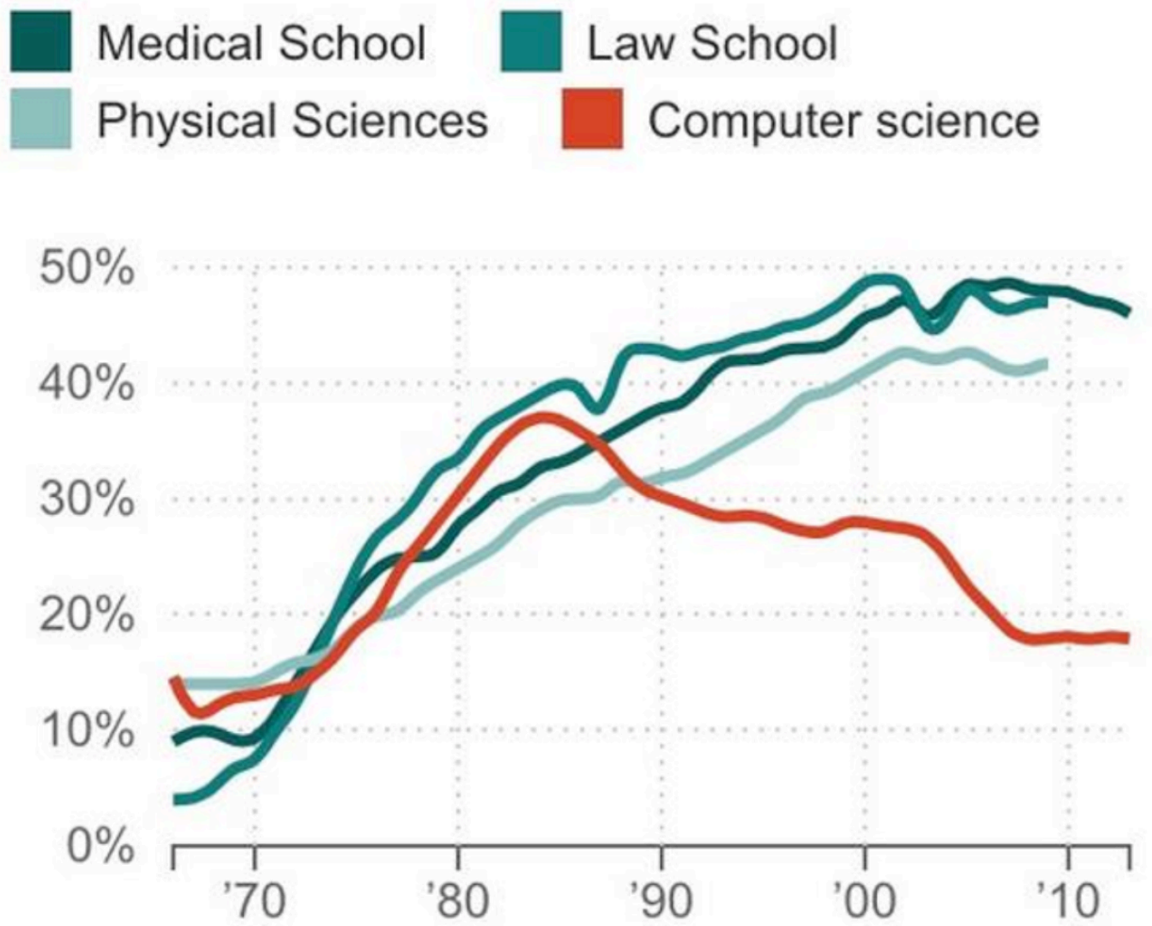
Average global temperature by year

Data from NASA/GISS.



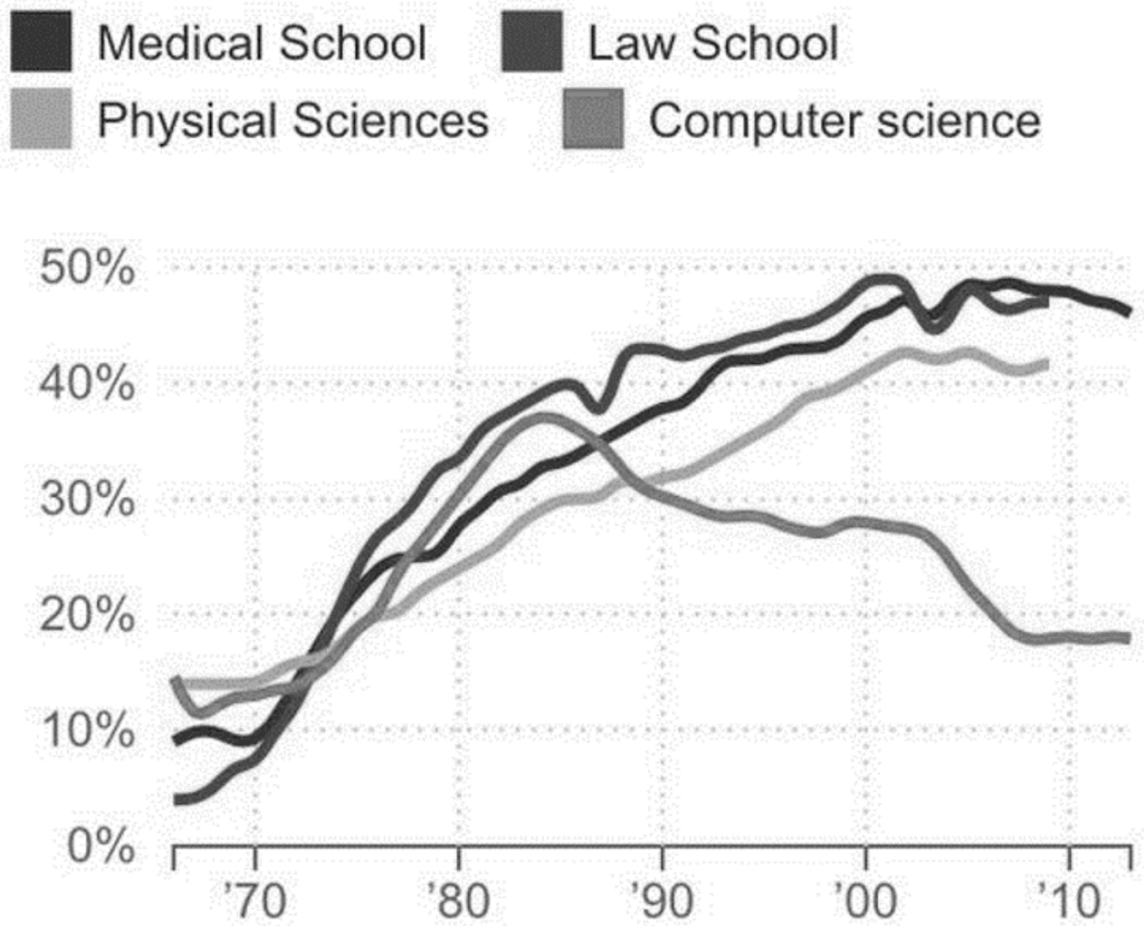
What Happened To Women In Computer Science?

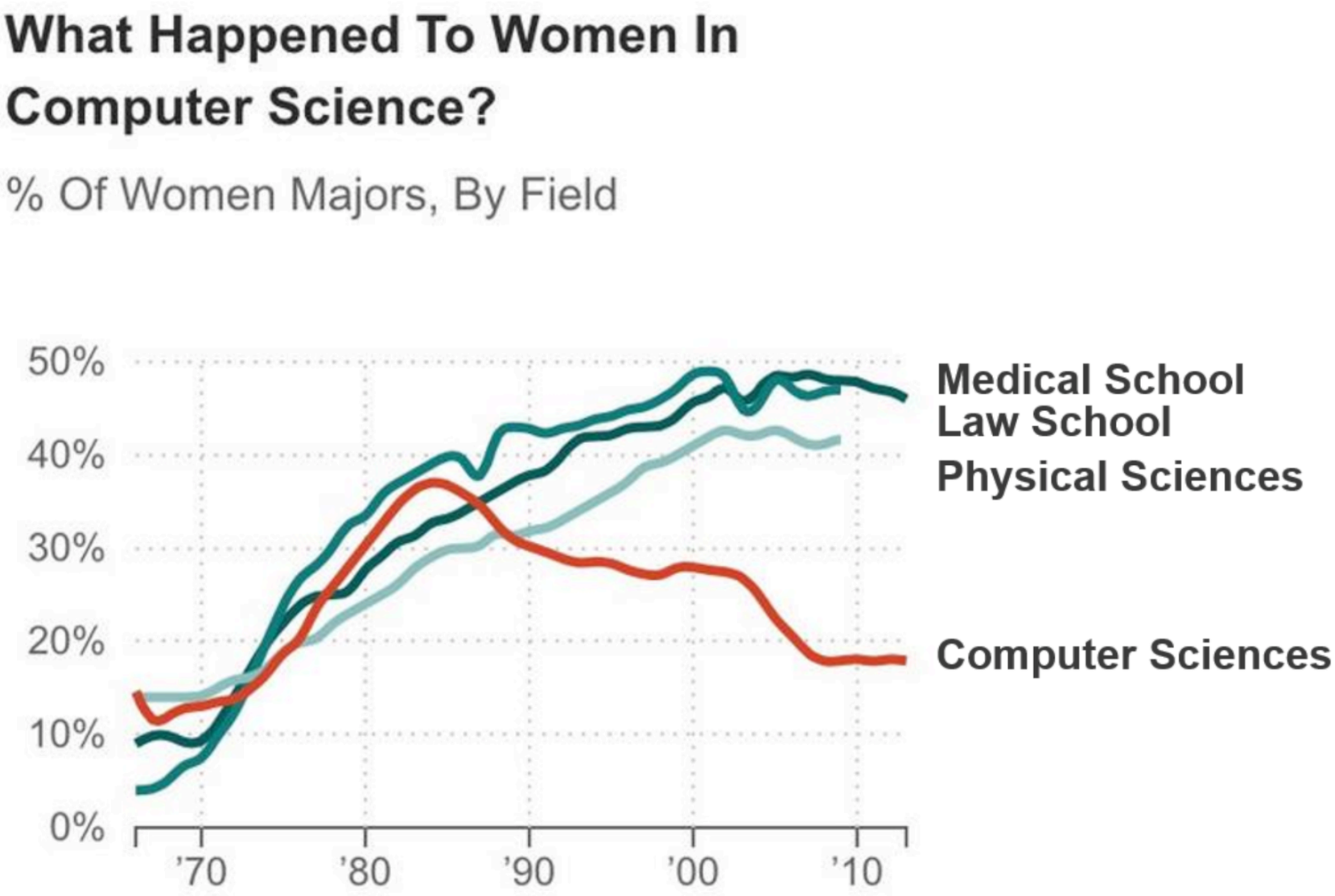
% Of Women Majors, By Field



What Happened To Women In Computer Science?

% Of Women Majors, By Field





4. Label directly

MONSTROUS COSTS

Total House and Senate campaign expenditures, in millions

$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

$$= \text{proportion of a graphic's ink devoted to the non-redundant display of data-information}$$

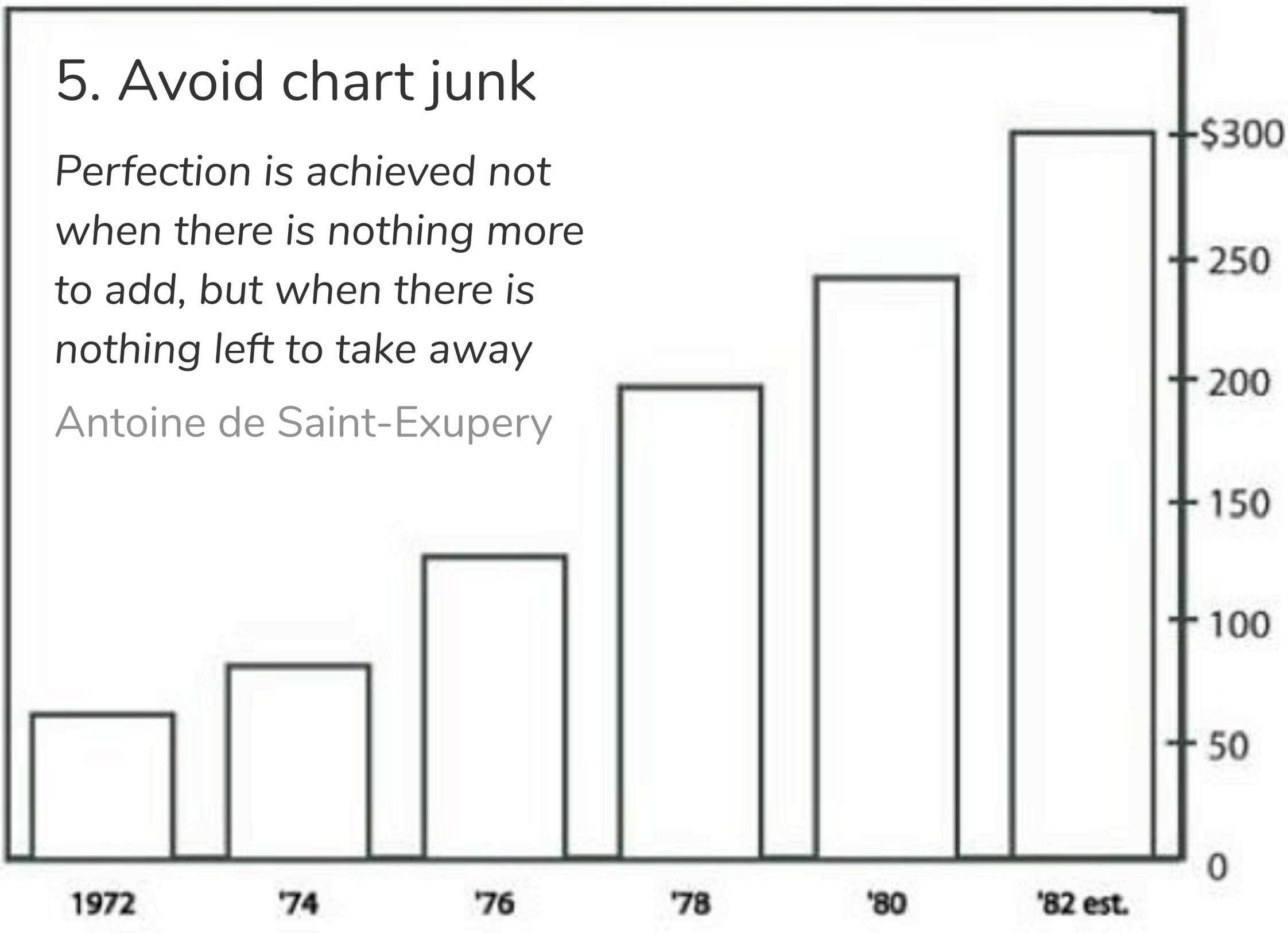
$$= \text{1.0 – proportion of a graphic that can be erased}$$

Above all else show the data, Edward Tufte, 1983

5. Avoid chart junk

Perfection is achieved not when there is nothing more to add, but when there is nothing left to take away

Antoine de Saint-Exupery



2014

69,2%

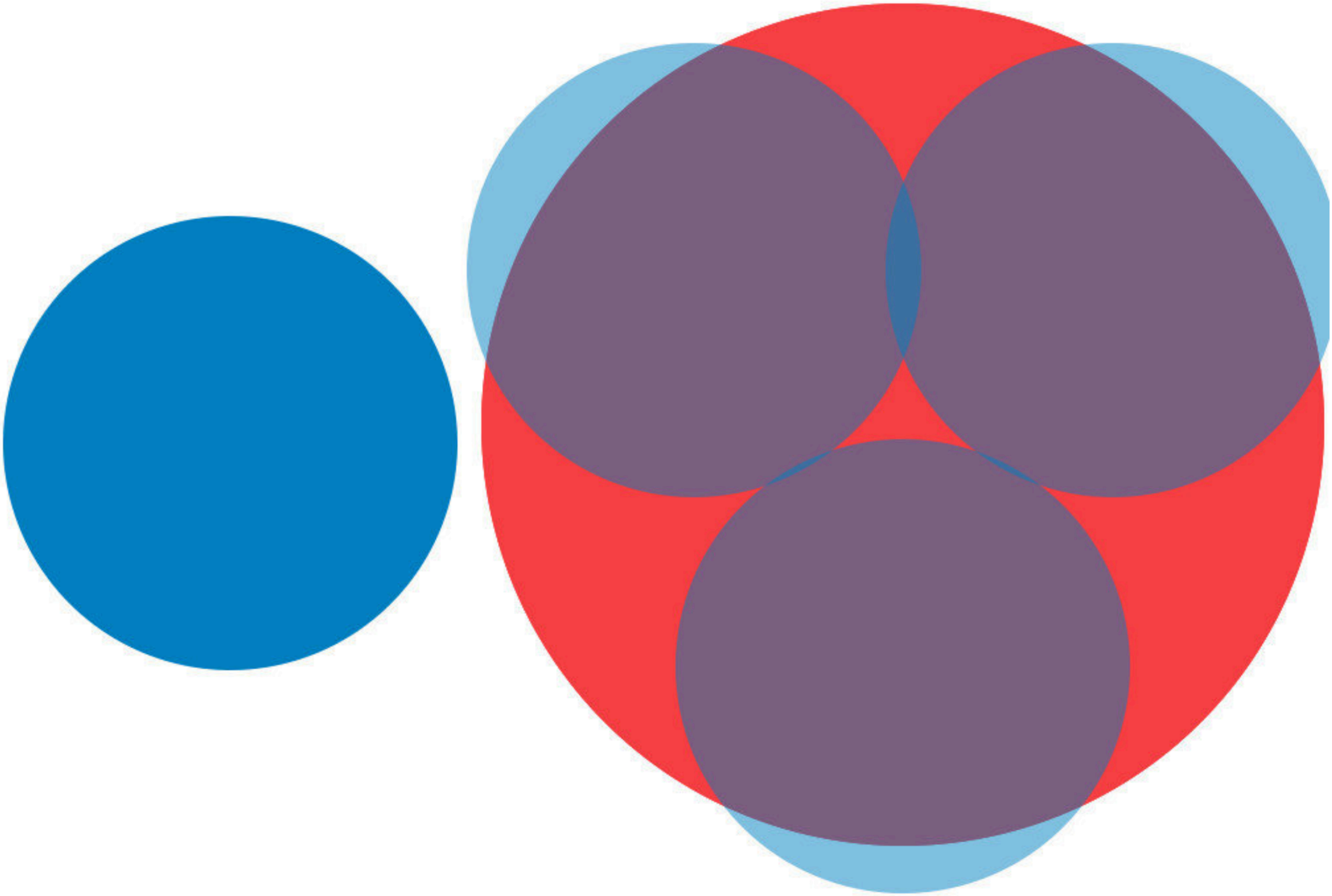
gemeenten besparen op sport

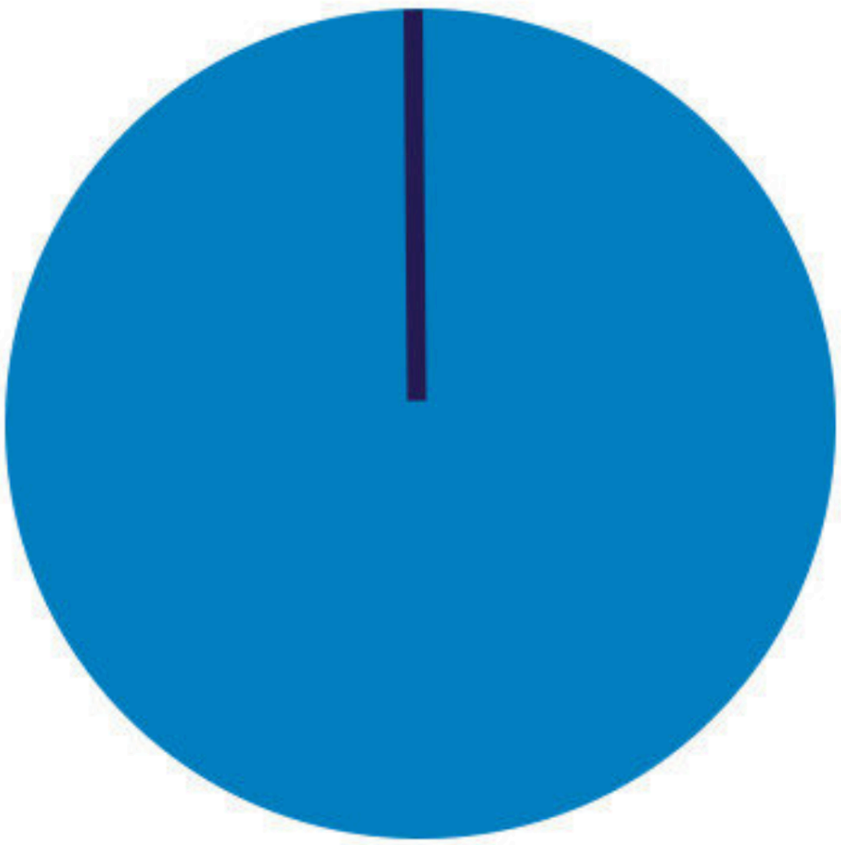
2013

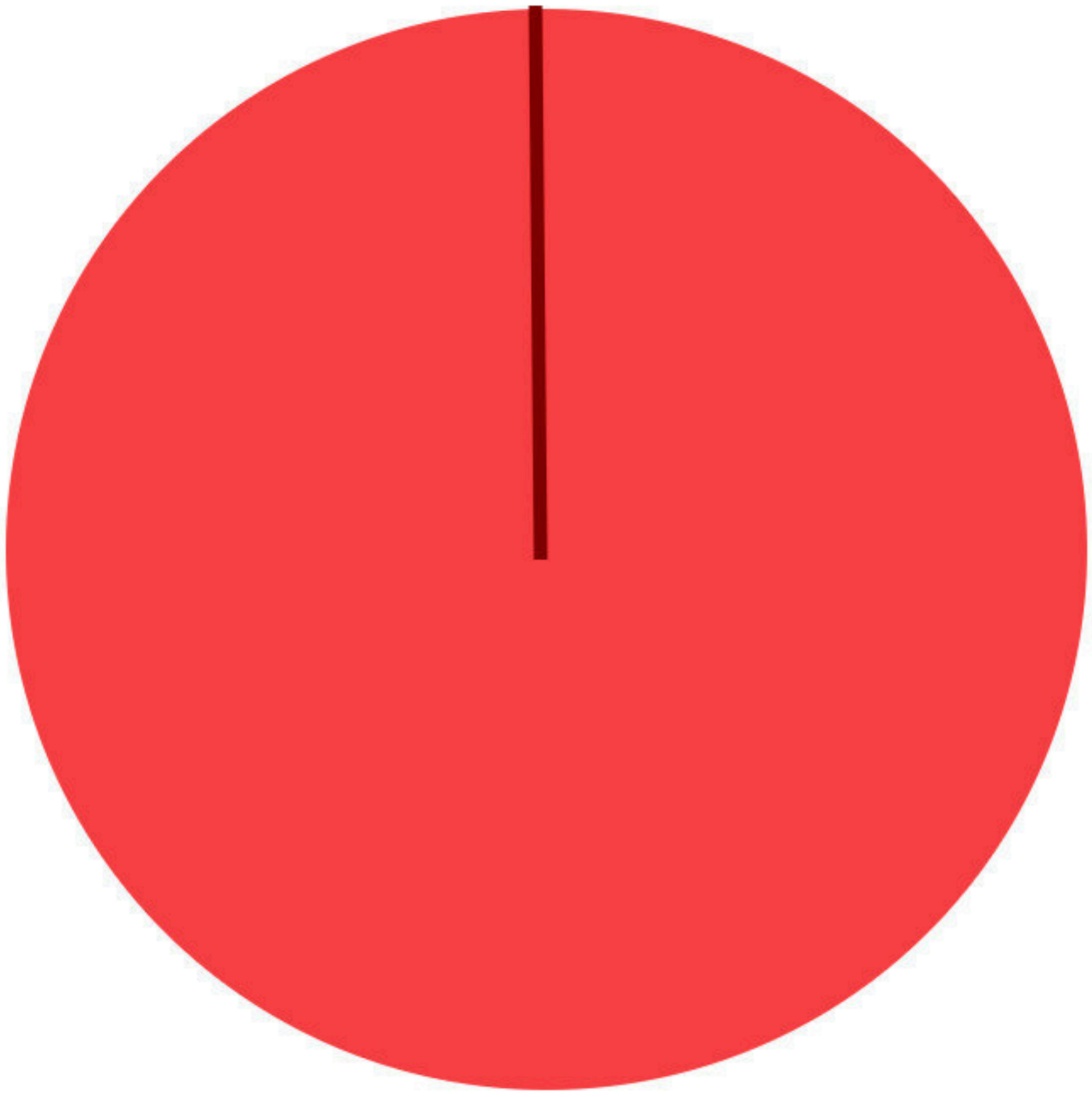
37,7%

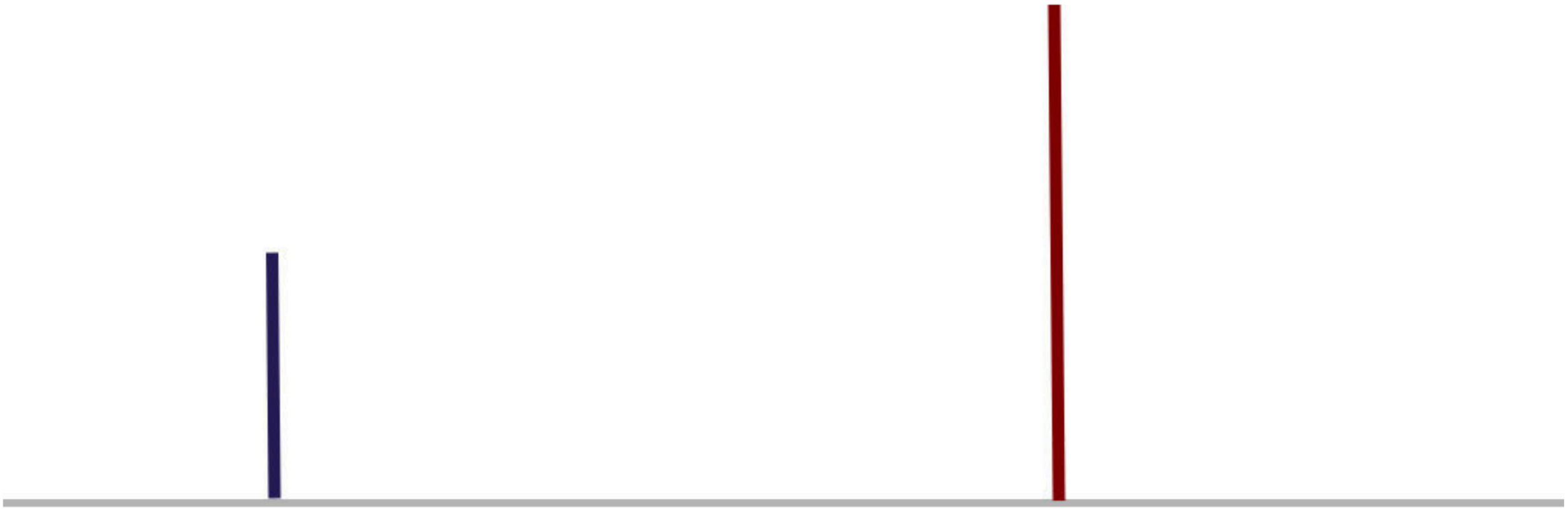
gemeenten besparen op sport

130 van de 308 gemeenten stuurden een antwoord (responsgraad van 42,2%)

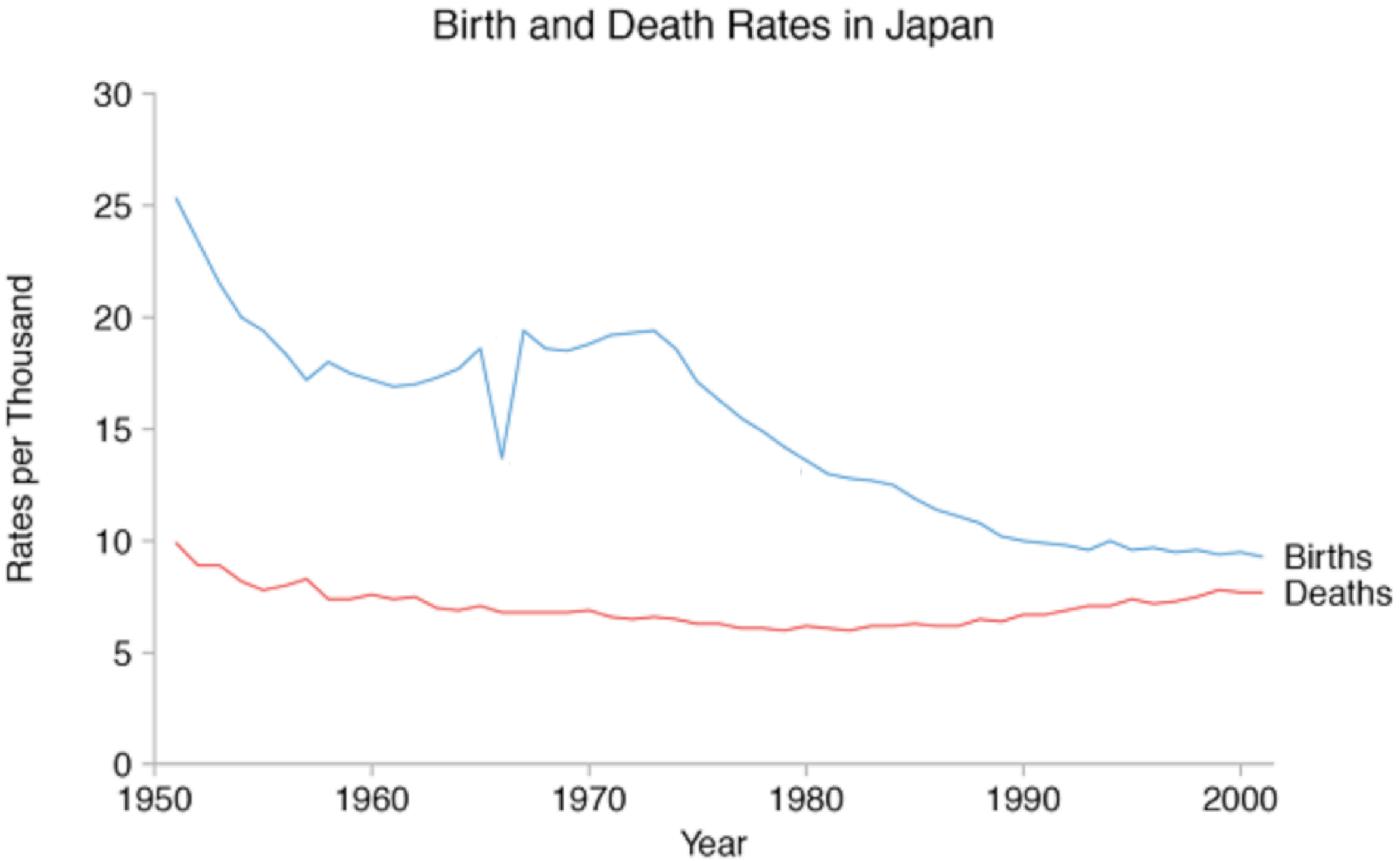




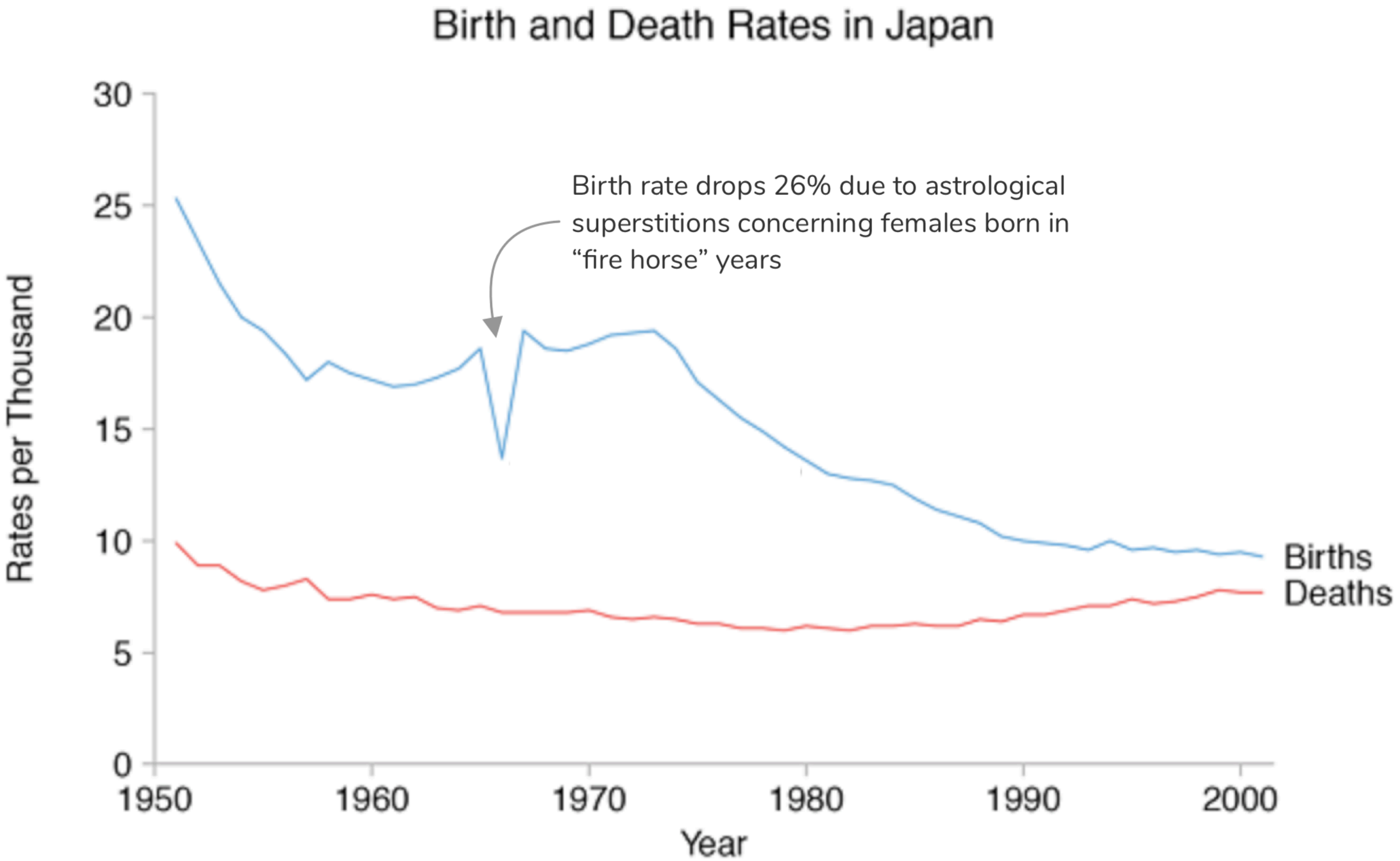




6. Scale circles by area not by radius



7. Tell the story

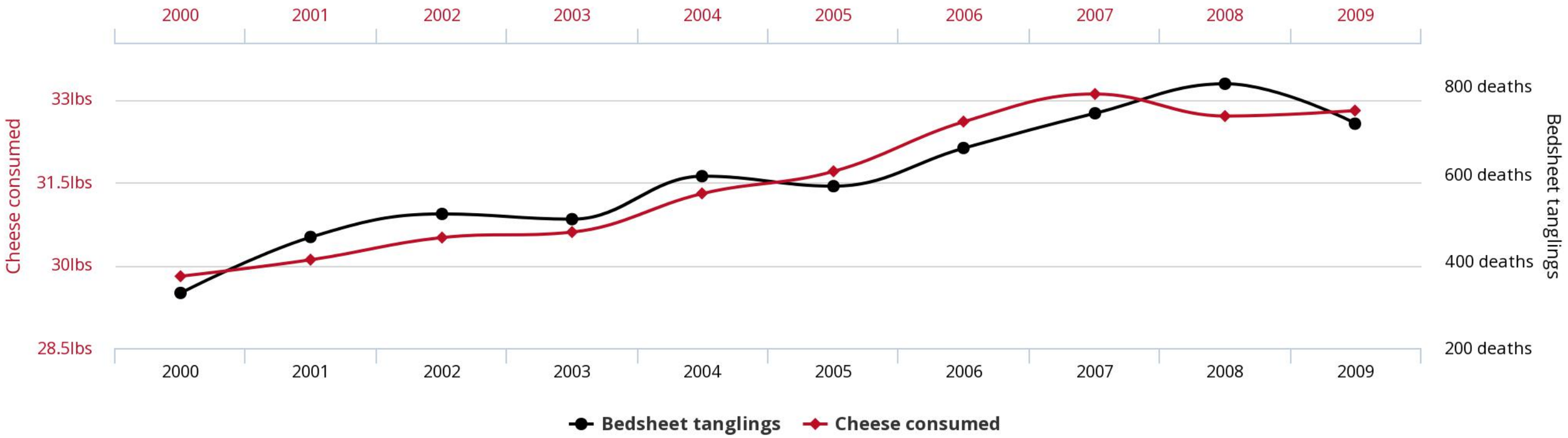


7. Tell the story

Per capita cheese consumption

correlates with

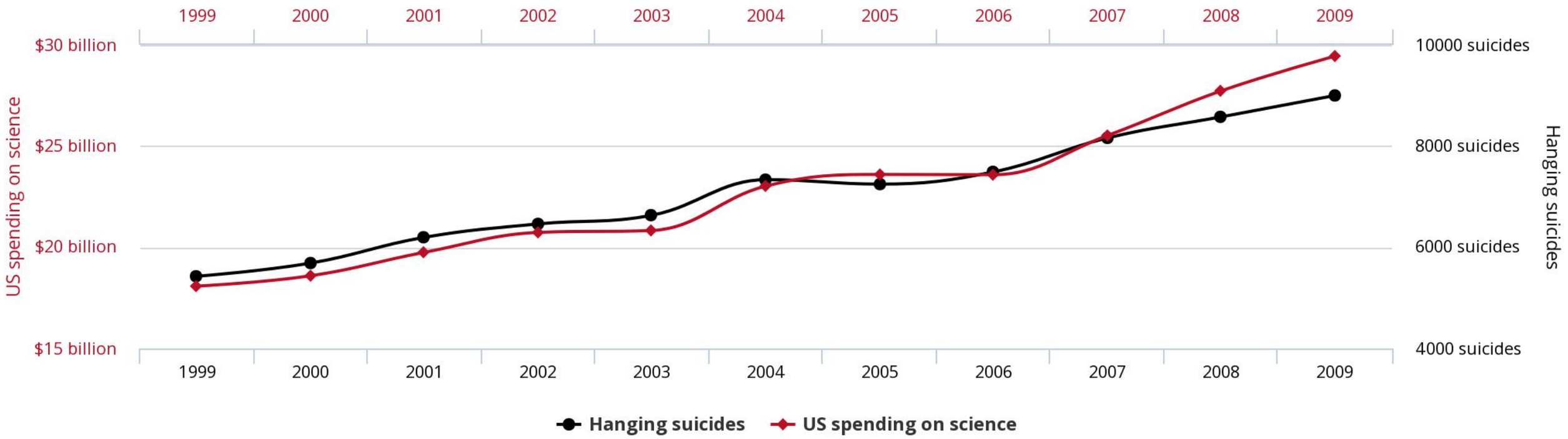
Number of people who died by becoming tangled in their bedsheets



US spending on science, space, and technology

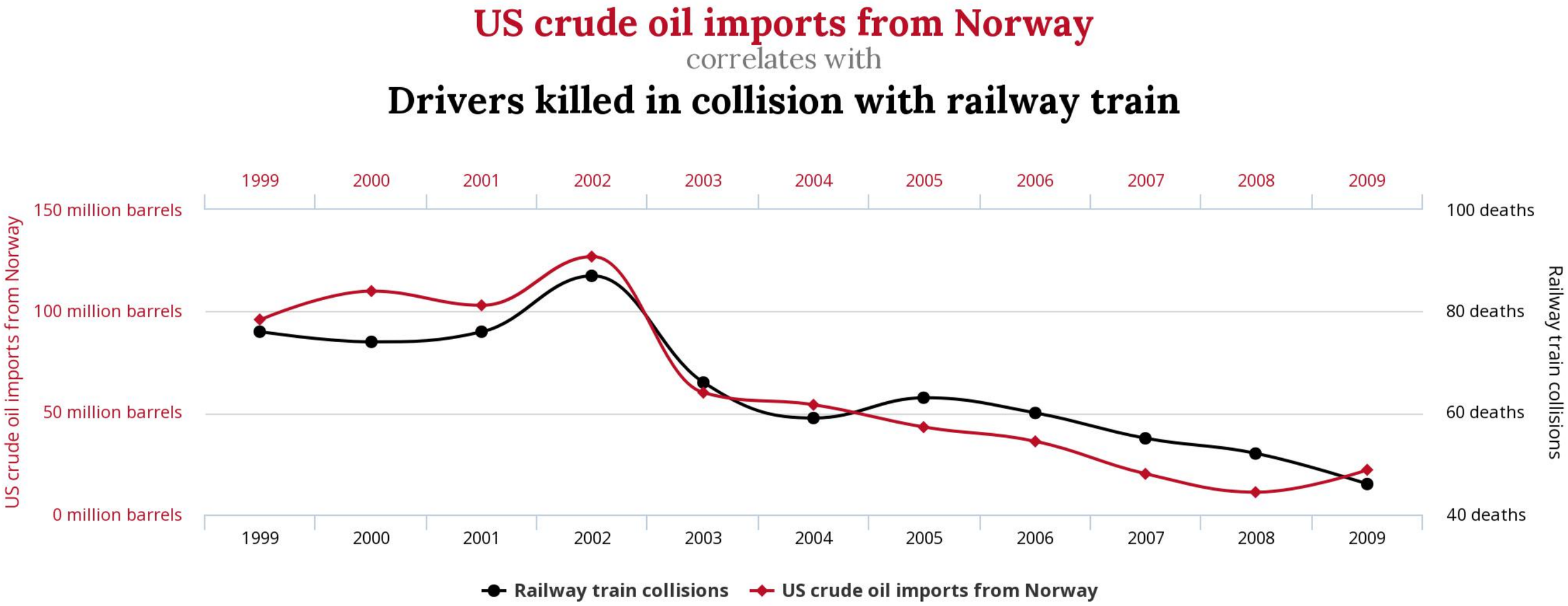
correlates with

Suicides by hanging, strangulation and suffocation



tylervigen.com

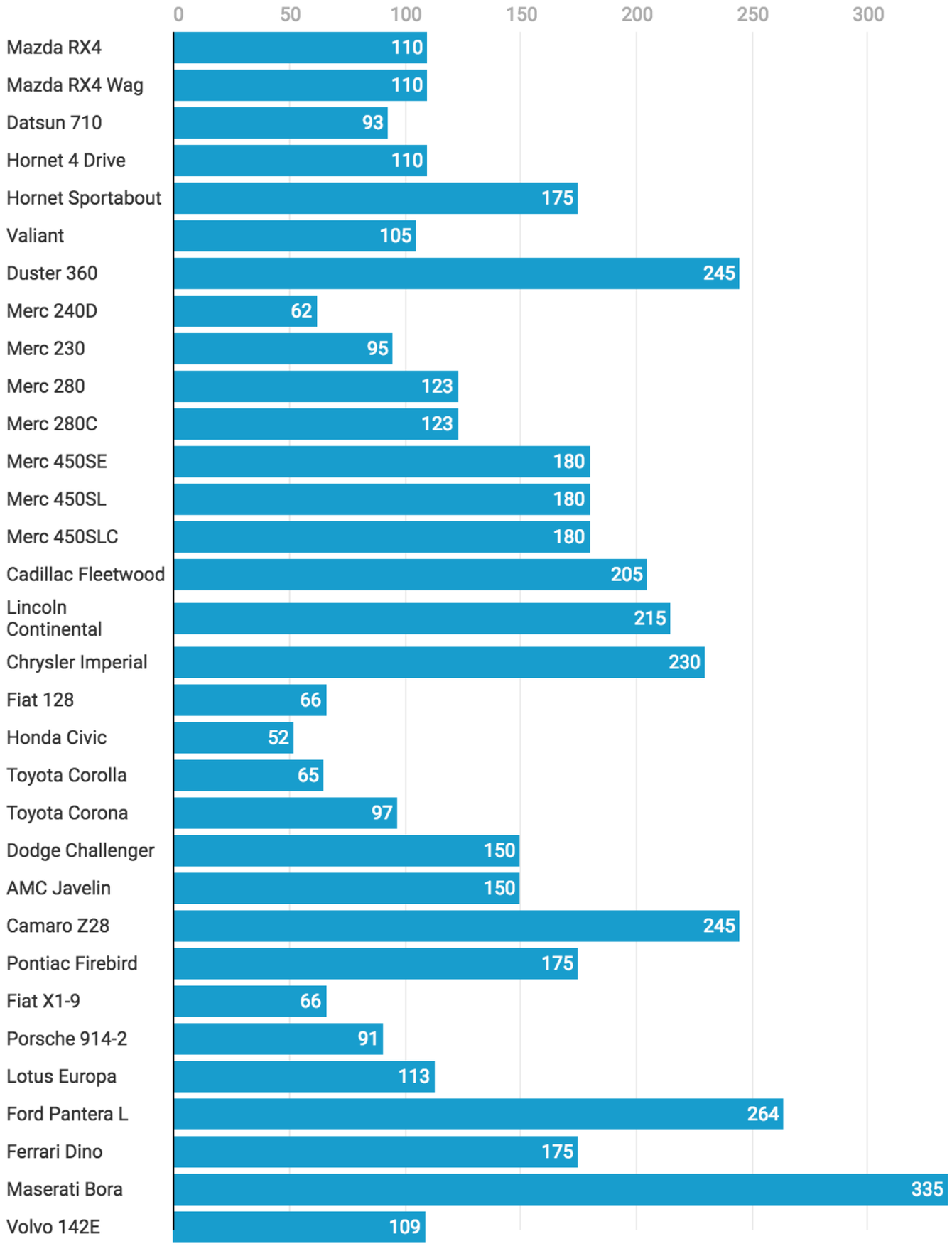
8. Correlation isn't causation



Never show a secondary axis.

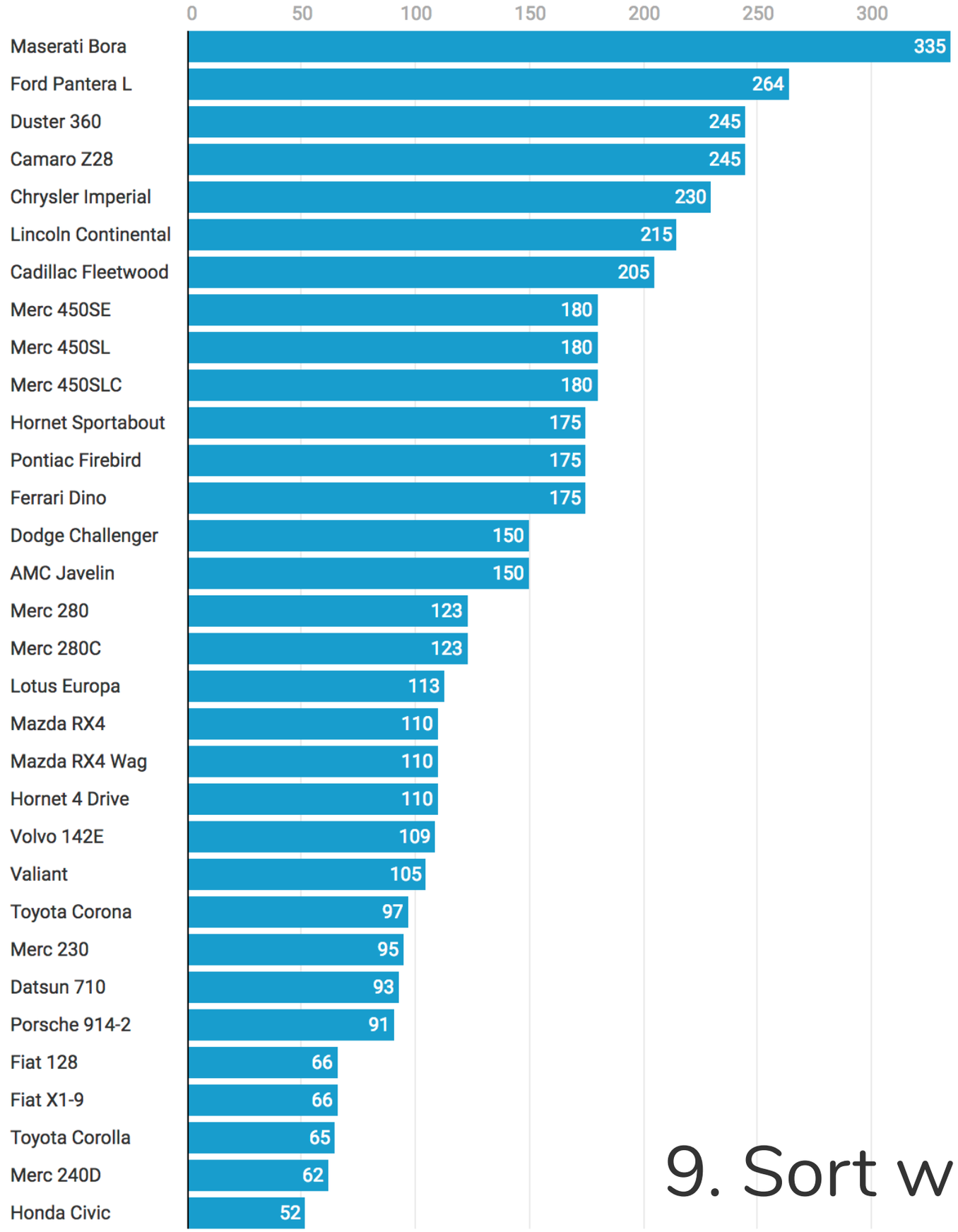
Horsepower from selected cars 1974

What?



Horsepower from selected cars 1974 • Sorted

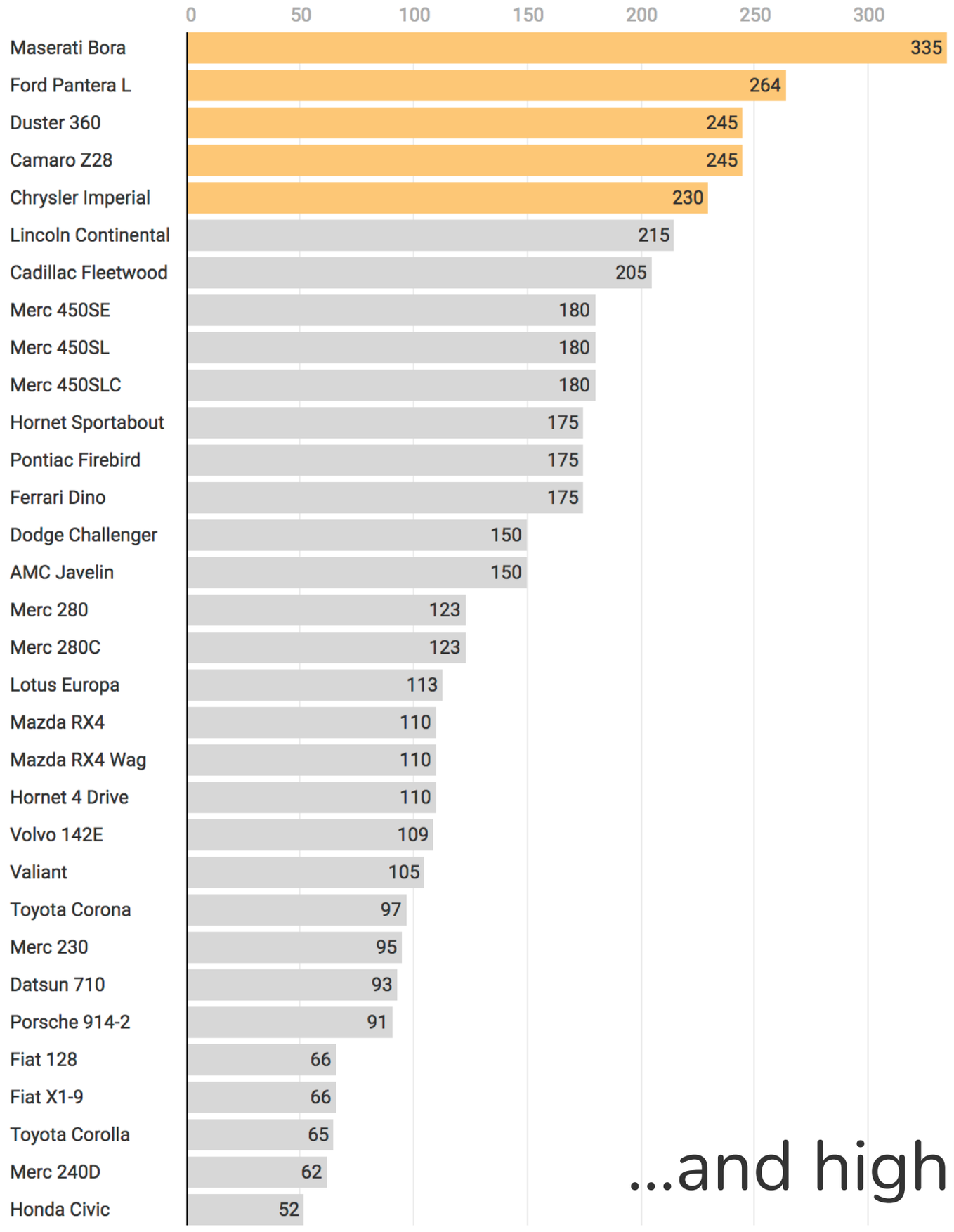
Ah ok.



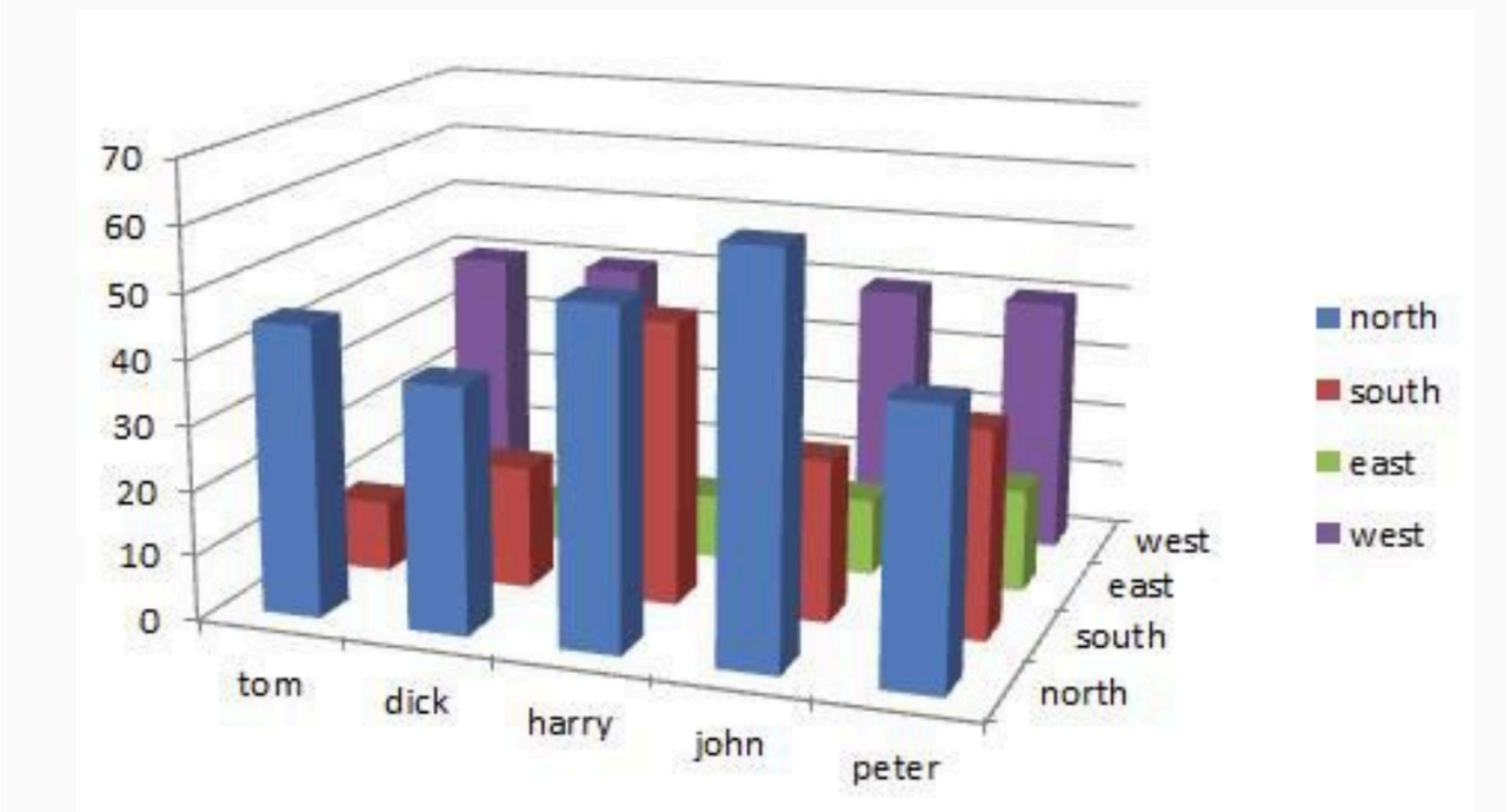
9. Sort when you can

Horsepower from selected cars 1974 • Sorted

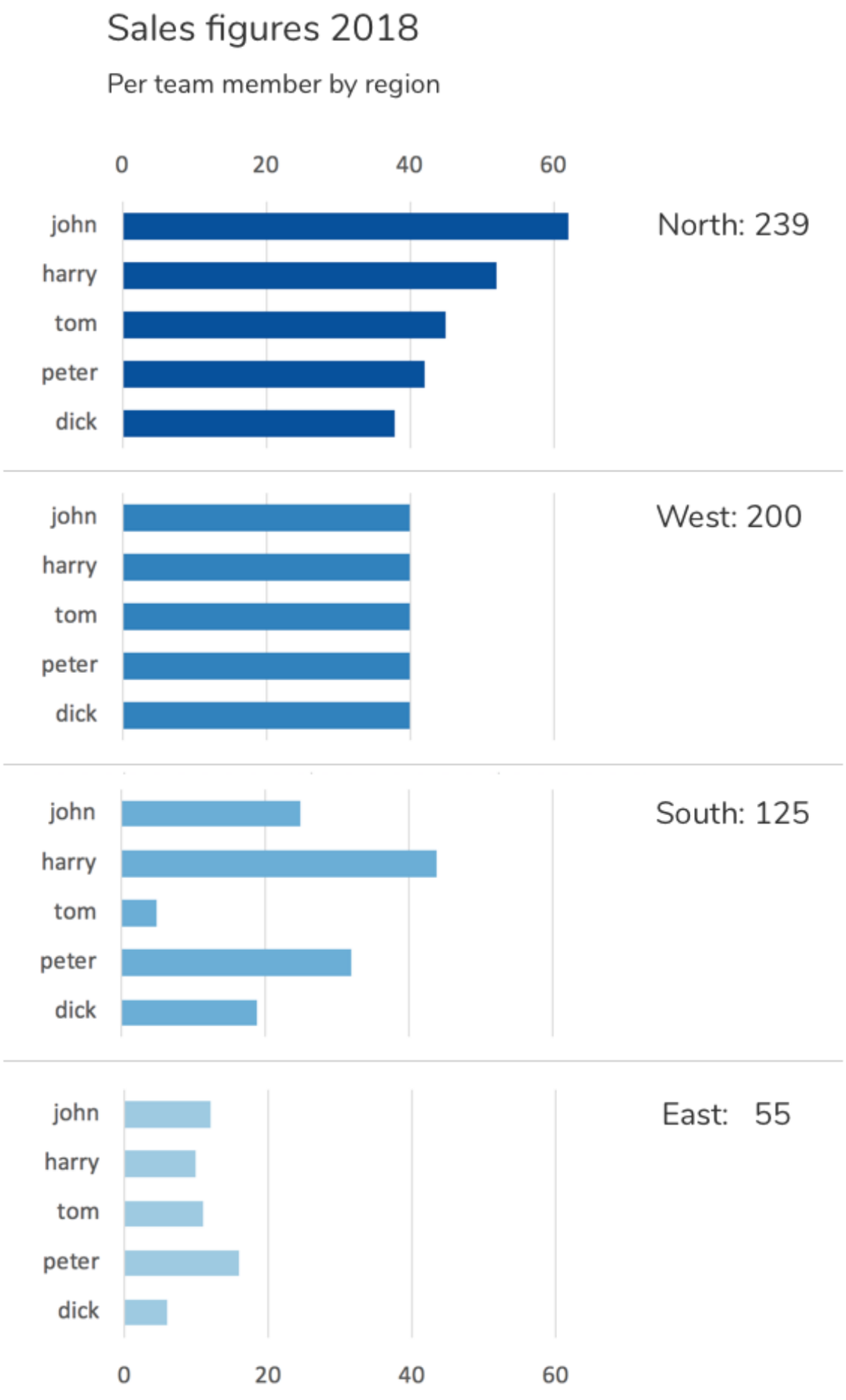
Ah ok.



...and highlight sections



10. Don’t do 3D, please


Sales figures 2018
Per team member by region
0
20
40
60
john
harry
tom
peter
dick
North: 239
West: 200
South: 125
East: 55

The pitfalls in one place

1. Use colour sparingly and conscious
2. Avoid pie charts
3. Don't cut bars
4. Label directly
5. Avoid chart junk
6. Scale circles by area
7. Tell the story / Annotate
8. Correlation isn't causation
9. Order!
10. Don't do 3D

Or:

datavizchecklist.stephanieevergreen.com

Data Visualization Checklist

by Stephanie Evergreen & Ann K. Emery
February 2018

This checklist is meant to be used as a guide for the development of high impact data visualizations. Rate each aspect of the data visualization by circling the most appropriate number, where 2 points means the guideline was fully met, 1 means it was partially met, and 0 means it was not met at all. n/a should not be used frequently, but reserved for when the guideline truly does not apply. For example, a pie chart has no axes lines or tick marks to rate. If the guidelines has been broken intentionally to make a point, rate it n/a and deduct those points from the total possible. Refer to the Data Visualization Anatomy Chart on the last page for guidance on vocabulary and the Resources at the end for more details.

	Guideline	Rating
Text	**6-12 word descriptive title is left-justified in upper left corner** Short titles enable readers to comprehend takeaway messages even while quickly skimming the graph. Rather than a generic phrase, use a descriptive sentence that encapsulates the graph's finding or "so what?" Western cultures start reading in the upper left, so locate the title there.	2 1 0 n/a
Graphs don't contain much text, so existing text must encapsulate your message and pack a punch.	**Subtitle and/or annotations provide additional information** Subtitles and annotations (call-out text within the graph) can add explanatory and interpretive power to a graph. Use them to answer questions a viewer might have or to highlight specific data points.	2 1 0 n/a
	Text size is hierarchical and readable Titles are in a larger size than subtitles or annotations, which are larger than labels, which are larger than axis labels, which are larger than source information. The smallest text - axis labels - are at least 9 point font size on paper, at least 20 on screen.	2 1 0 n/a
	Text is horizontal Titles, subtitles, annotations, and data labels are horizontal (not vertical or diagonal). Line labels and axis labels can deviate from this rule and still receive full points. Consider switching graph orientation (e.g., from column to bar chart) to make text horizontal.	2 1 0 n/a
	Data are labeled directly Position data labels near the data rather than in a separate legend (e.g., on top of or next to bars and next to lines). Eliminate/embed legends when possible because eye movement back and forth between the legend and the data can interrupt the brain's attempts to interpret the graph.	2 1 0 n/a
	Labels are used sparingly Focus attention by removing the redundancy. For example, in line charts, label every other year on an axis	2 1 0 n/a

Summary

1. Charts are **tested** concoctions of individual encodings

2. **Relationship typology**: Diagram, Network, Map

3. Consider chart choosers based on **objectives** or on **data shape**

4. Use the **datavis checklist**

The summaries' summary

Visualisation: Data → precisely defined graphic

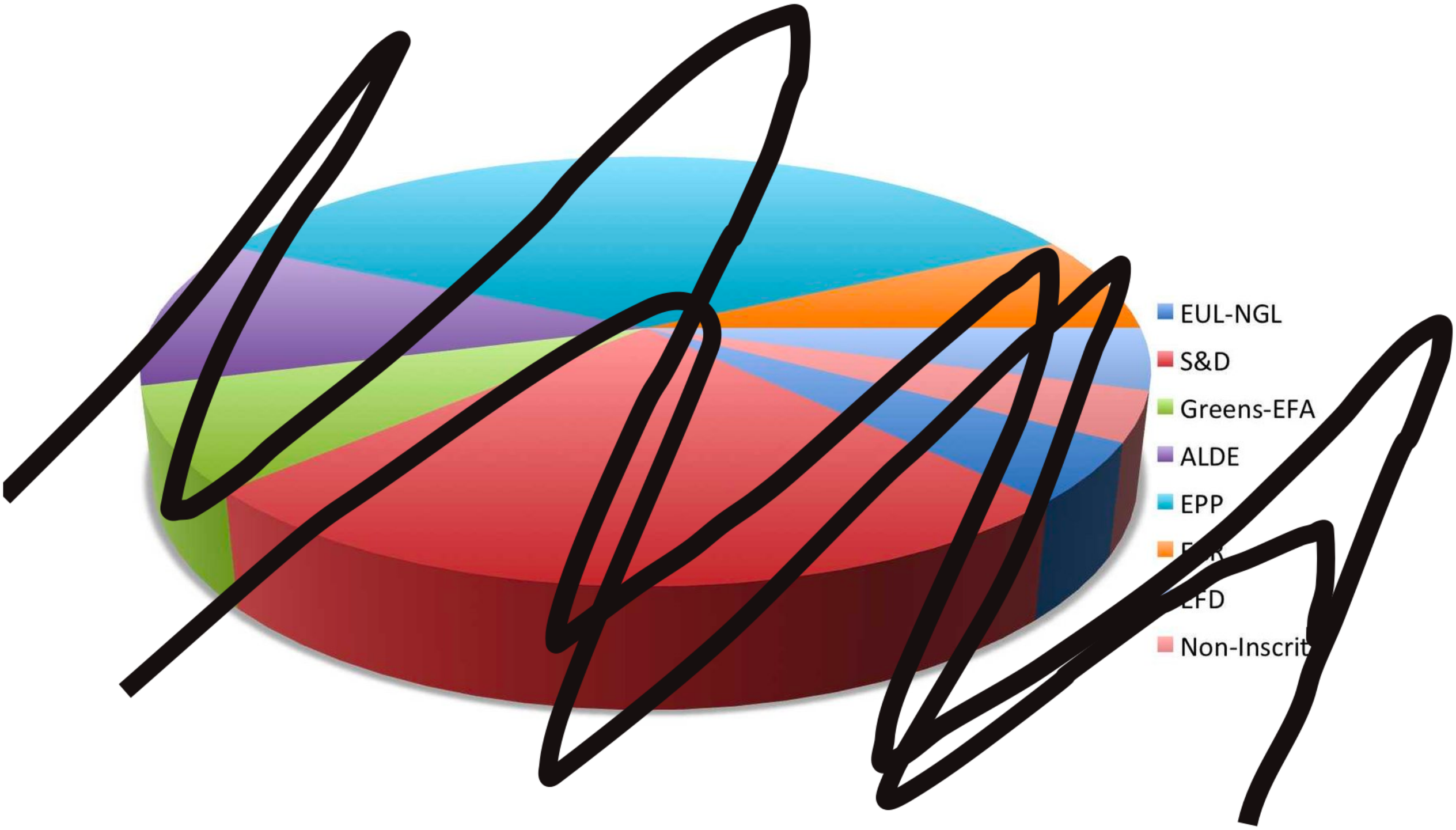
Designer creates meaning by picking the message

Make good use of our visual system. It's fast and experienced.

Follow the Visualisation Recipe:
Analyse the information → Express the data types → Encode effectively

Design tools are here to help:
Chart choosers, Colour pickers, Datavis checklist

Please don't make unsorted 3D Pie charts


EUL-NGL
S&D
Greens-EFA
ALDE
EPP
EFD
Non-Inscrit