

Designing With Integrity

Alberto Cairo

OpenVisualizationAcademy.com



European data

data.europa.eu The official portal for European data

Home Data Academy Open data maturity Community Publications

Home > News & events > Events > Training series 'Visualising data with impact...

data.europa.eu event

Training series 'Visualising data with impact': foundations of effective data visualisation

14 October 2025

Online



European data

data.europa.eu The official portal for European data

Home Data Academy Open data maturity Community Publications

Home > News & events > Events > Training series 'Visualising data with impact...

data.europa.eu event

Training series 'Visualising data with impact': designing with integrity

15 October 2025

Online



European data

data.europa.eu The official portal for European data

Home Data Academy Open data maturity Community Publications

Home > News & events > Events > Training series 'Visualising data with impact...

data.europa.eu event

Training series 'Visualising data with impact': data storytelling and creating your narrative

16 October 2025

Online

Episode ONE - October 14

<https://data.europa.eu/en/news-events/events/training-data-visualisation-session-1-foundations-effective-data-visualisation>

Episode TWO - October 15

Today's webinar

<https://data.europa.eu/en/news-events/events/data-visualisation-training-session-2-designing-integrity>

Episode THREE - October 16

<https://data.europa.eu/en/news-events/events/data-visualisation-training-session-3-data-storytelling-and-creating-your>

EPISODE ONE: Foundations of effective data visualization

Defining information design and data visualization

Exploration, exposition, explanation, expression in visualization

Why do we visualize? The basics

The grammar of graphics

“Rules” versus decision-making

EPISODE TWO: Designing with integrity

Myths of visualization: From “A picture is worth a thousand words” to “the data should speak for itself”

The role of mental models

Why and when do charts “lie”, and what to do about it?

A structured way for thinking about visualization and minimizing misunderstanding

EPISODE THREE: Data storytelling and creating your narrative

Structuring a layout

Building a narrative

Considering visual design

Creative visualization

What comes next in visualization?

EPISODE ONE: Foundations of effective data visualization

Defining information design and data visualization

Exploration, exposition, explanation, expression in visualization

Why do we visualize? The basics

The grammar of graphics

“Rules” versus decision-making

EPISODE TWO: Designing with integrity

Myths of visualization: From “A picture is worth a thousand words” to “the data should speak for itself”

The role of mental models

Why and when do charts “lie”, and what to do about it?

A structured way for thinking about visualization and minimizing misunderstanding

EPISODE THREE: Data storytelling and creating your narrative

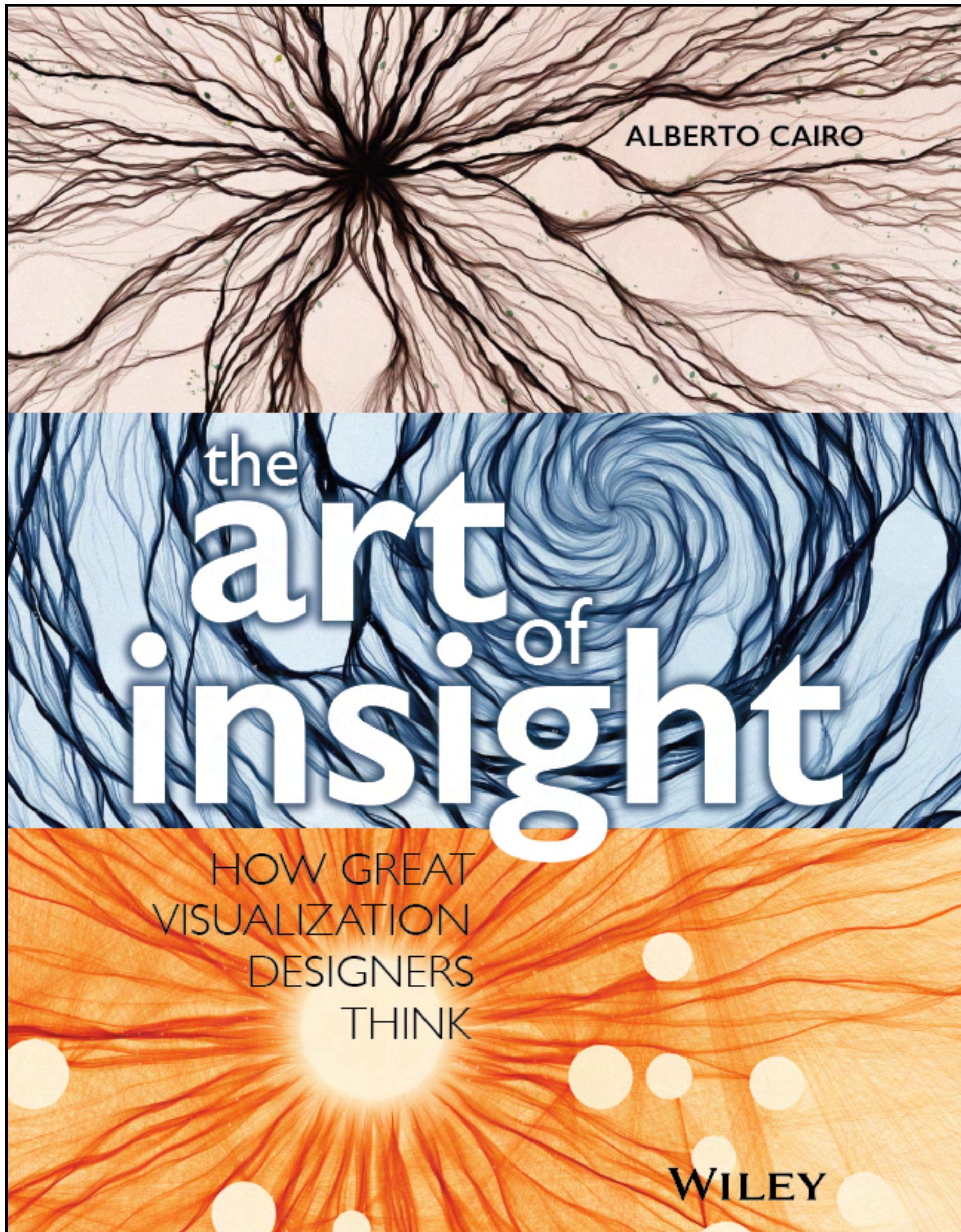
Structuring a layout

Building a narrative

Considering visual design

Creative visualization

What comes next in visualization?



Visualization design consists of **reasoning** about possible **choices** by considering the interplay between:

- 1. Content:** The nature, origin, and limitations of the data.
- 2. People:** Your audience.
- 3. Intent:** The purpose(s) that we define.
- 4. Constraints:** The limitations that we may face.
- 5. Outcomes:** How the graphic is received.

Every design choice must be **deliberate**.

It is inevitably **subjective**, but it should never be **arbitrary**.

Myths about data visualization

“The data should speak for itself!”

“A picture is worth a thousand words”

“Show, don’t tell!”

...



11:00 PM

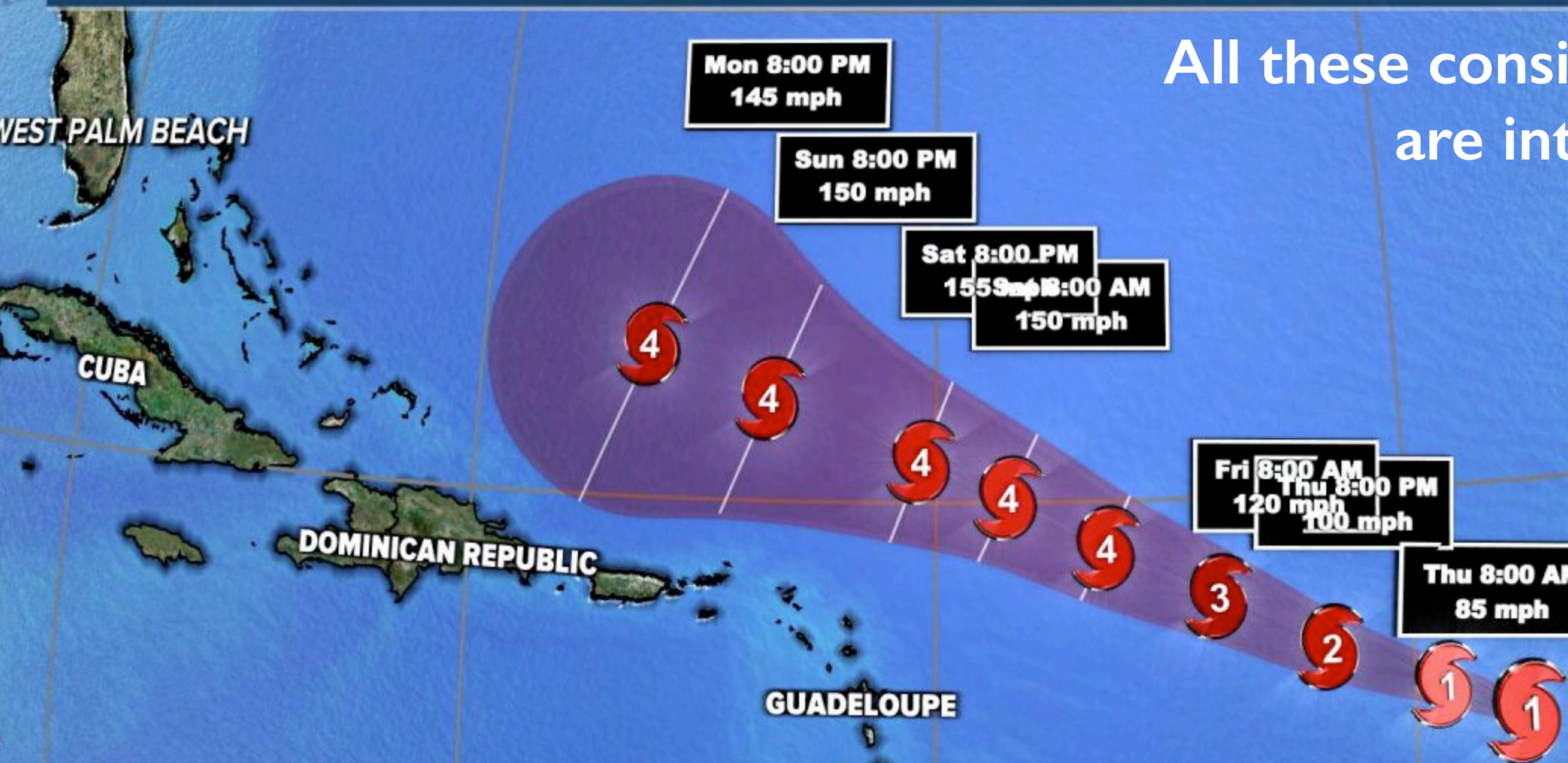
Max Winds: **80 mph**

Hurricane Lee

15.4°N, 47.7°W

Moving: **WNW at 14 mph**

Pressure: **989 mb**



All these considerations
are interrelated

Opinion

Those Hurricane Maps Don't Mean What You Think They Mean

We use hurricane forecasts to warn people. Why do we misinterpret them so often?

By Alberto Cairo

With Tala Schlossberg

<https://www.nytimes.com/interactive/2019/08/29/opinion/hurricane-dorian-forecast-map.html>

A designer might want to communicate something...

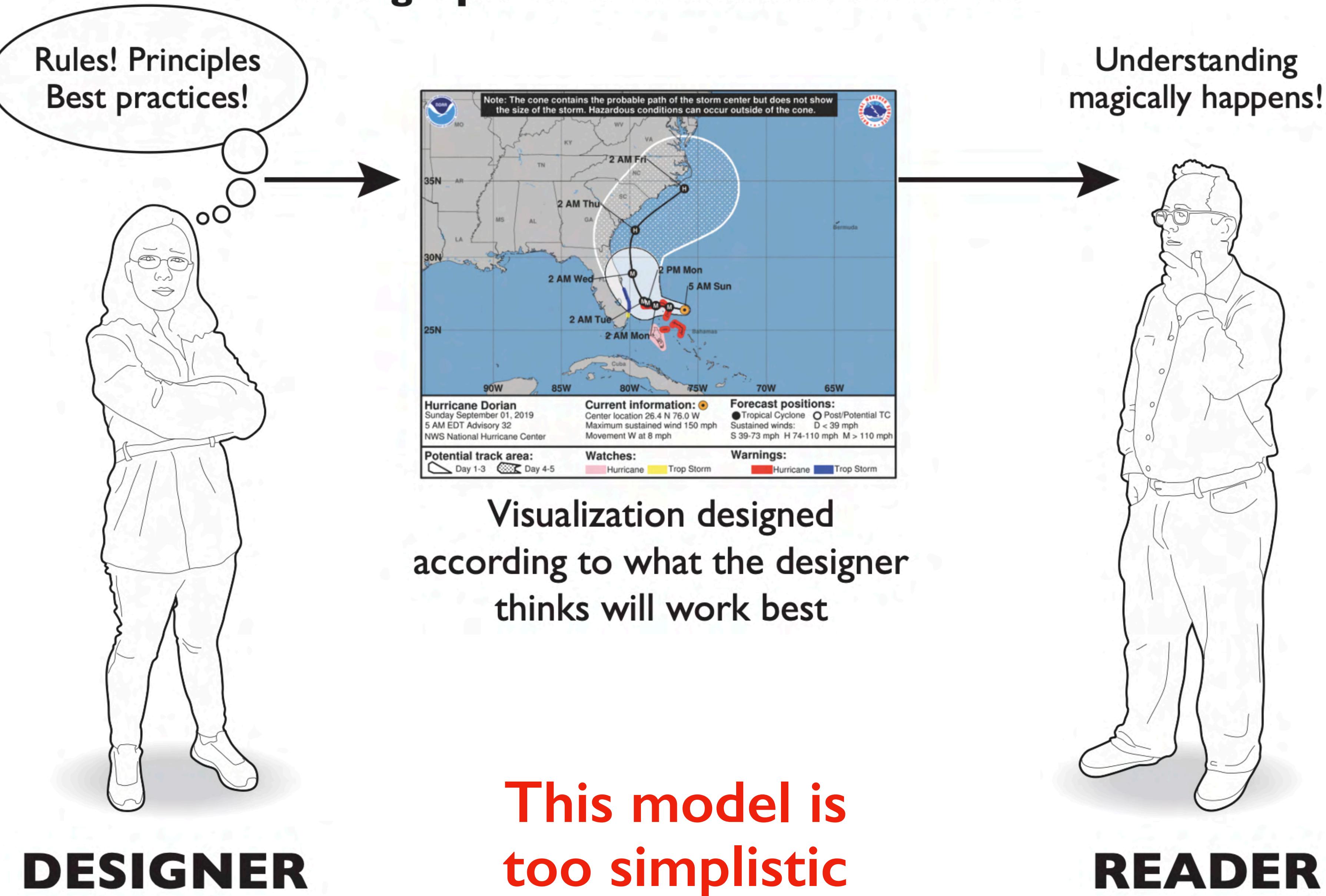


A designer might want to communicate something...

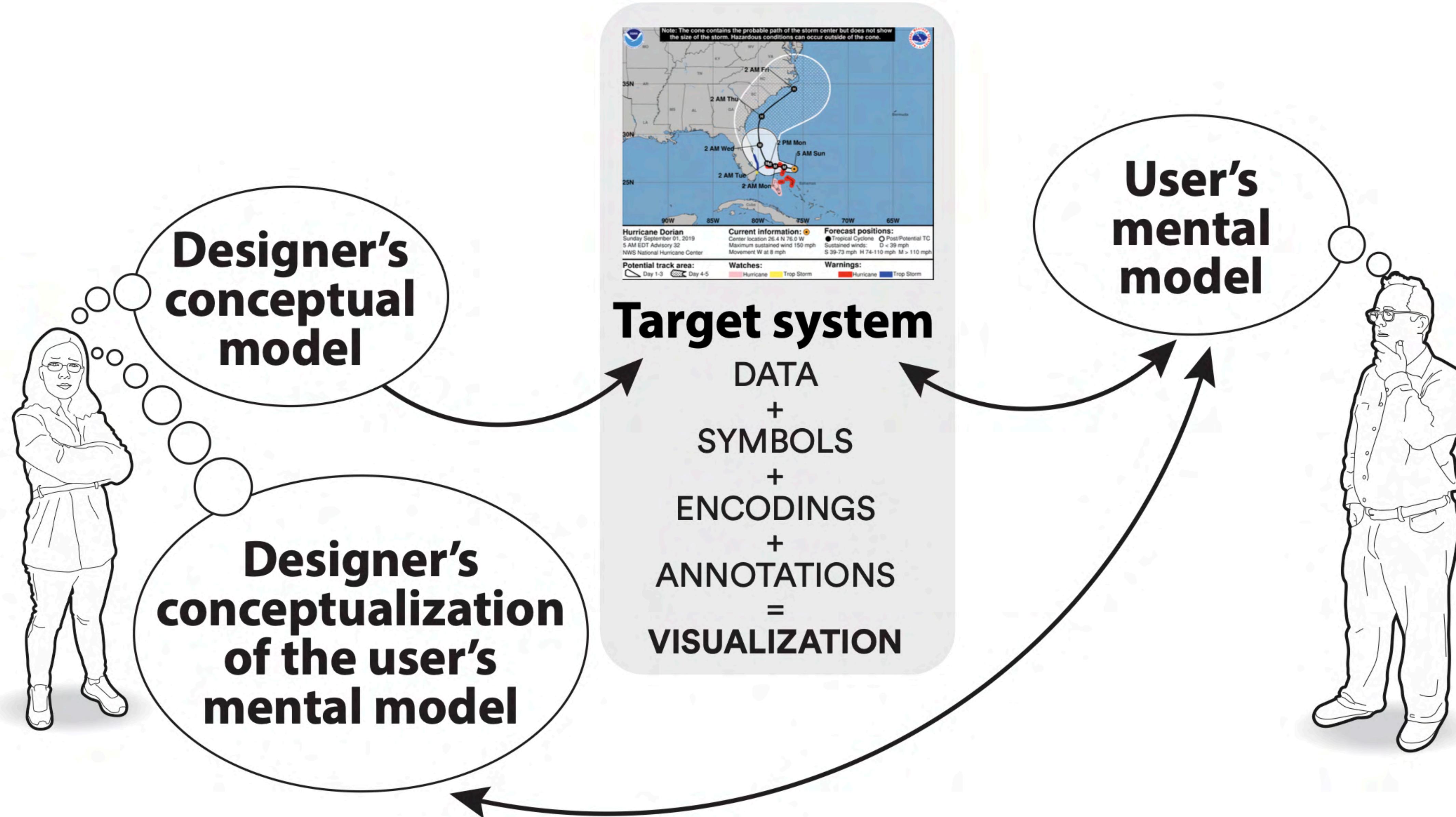
...but readers may interpret something different.



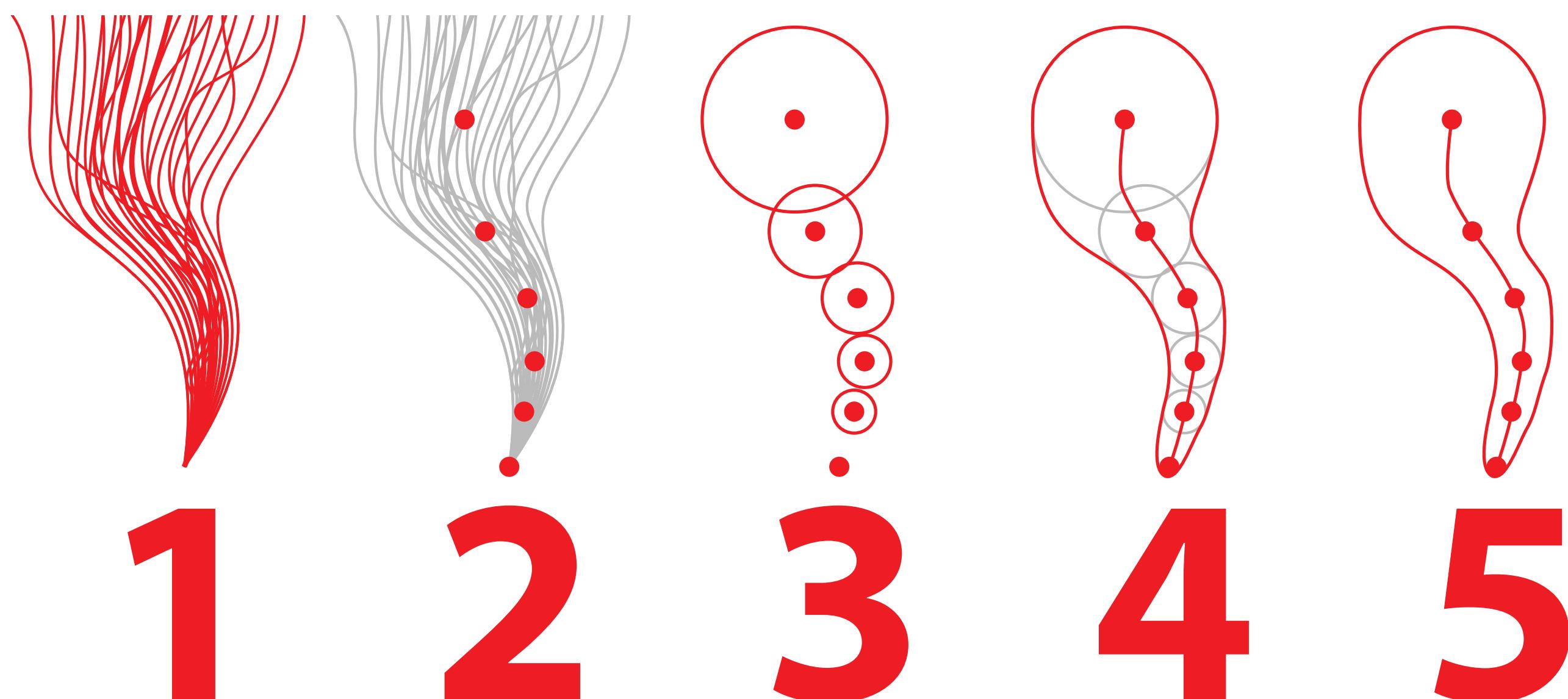
Traditional, unidirectional model of how graphical communication works



THE PSYCHOLOGY OF CHART COMPREHENSION, SIMPLIFIED

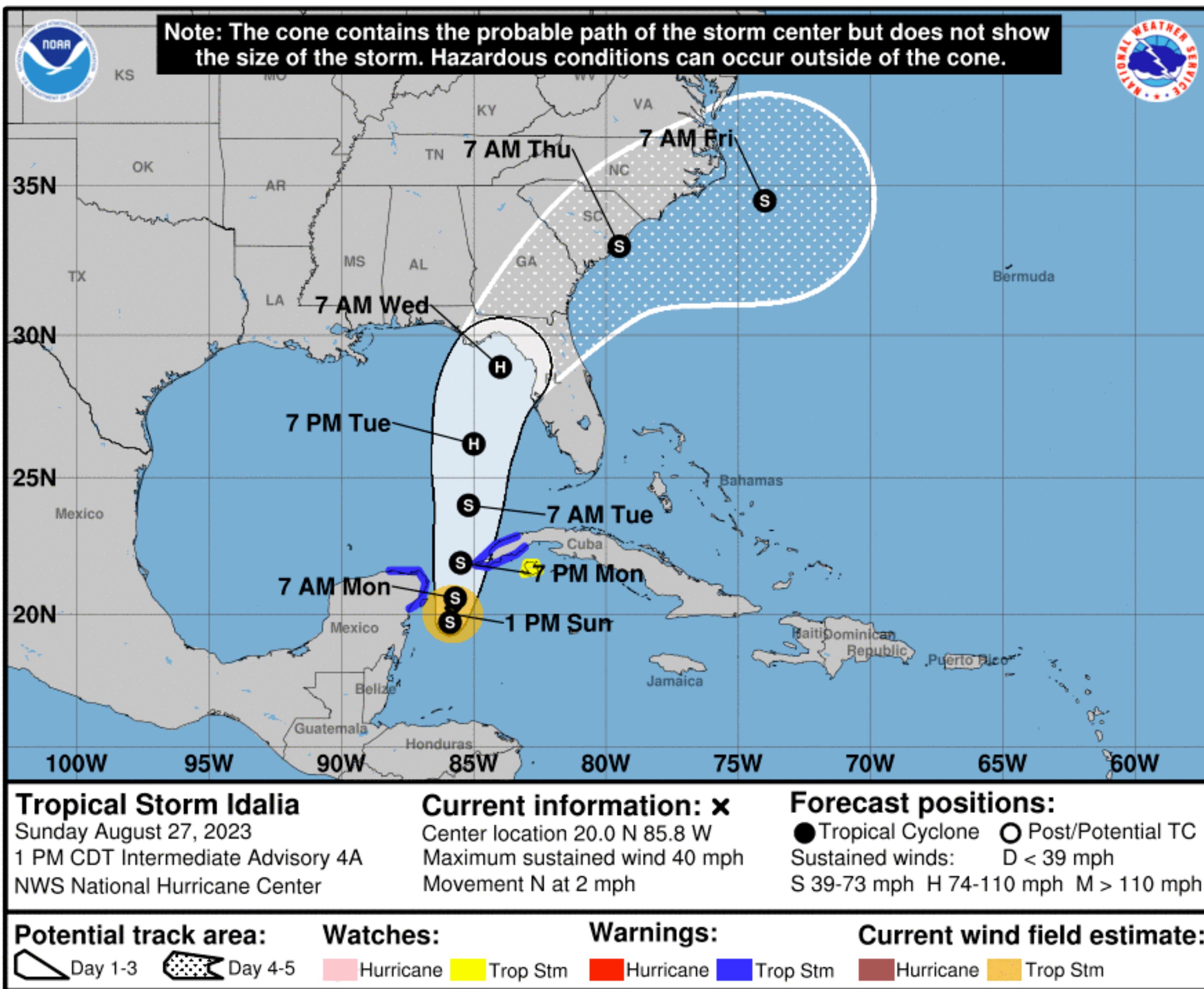


How to read the NHC cone of uncertainty



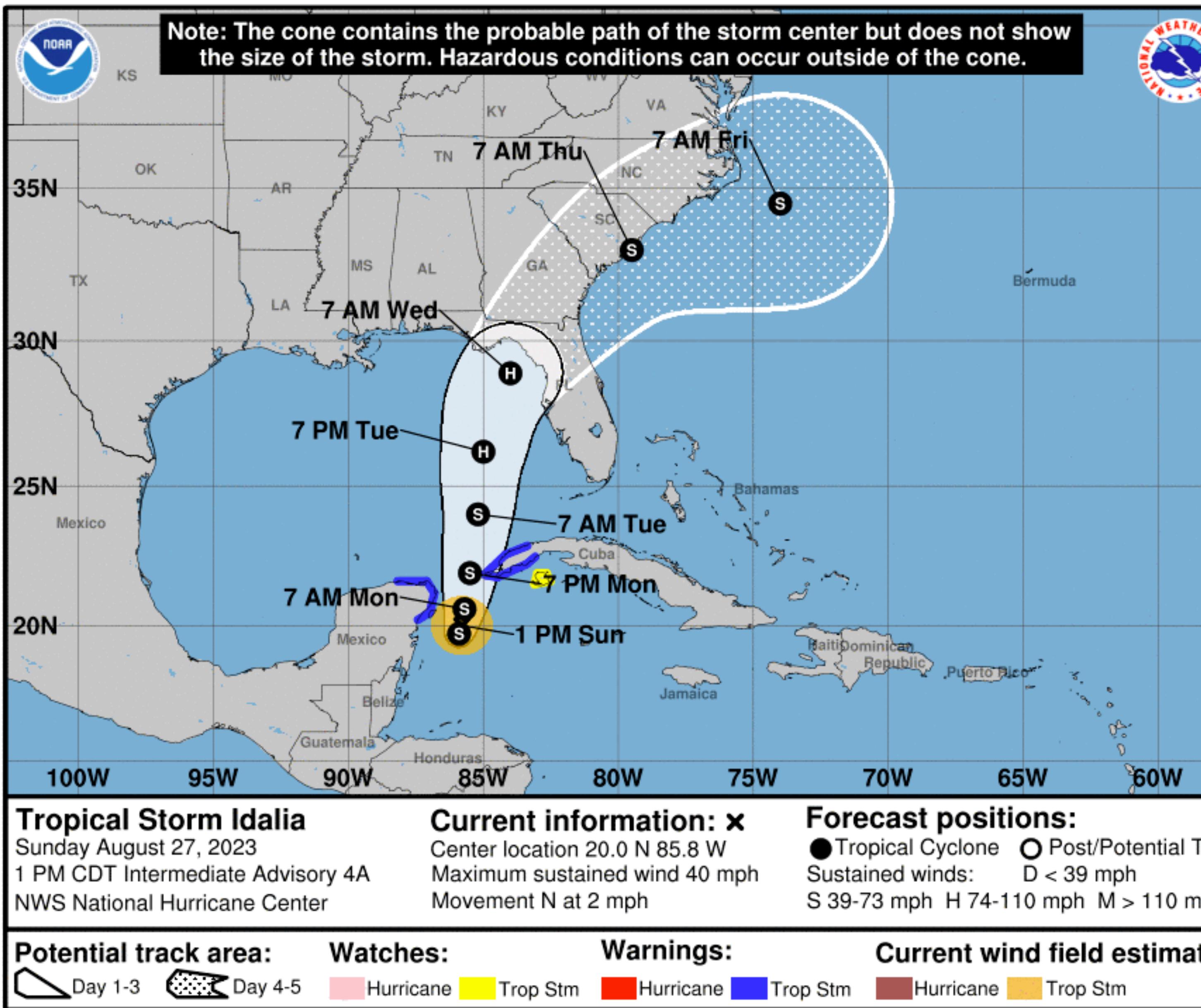
What do readers *really* need to know?

The possible path?

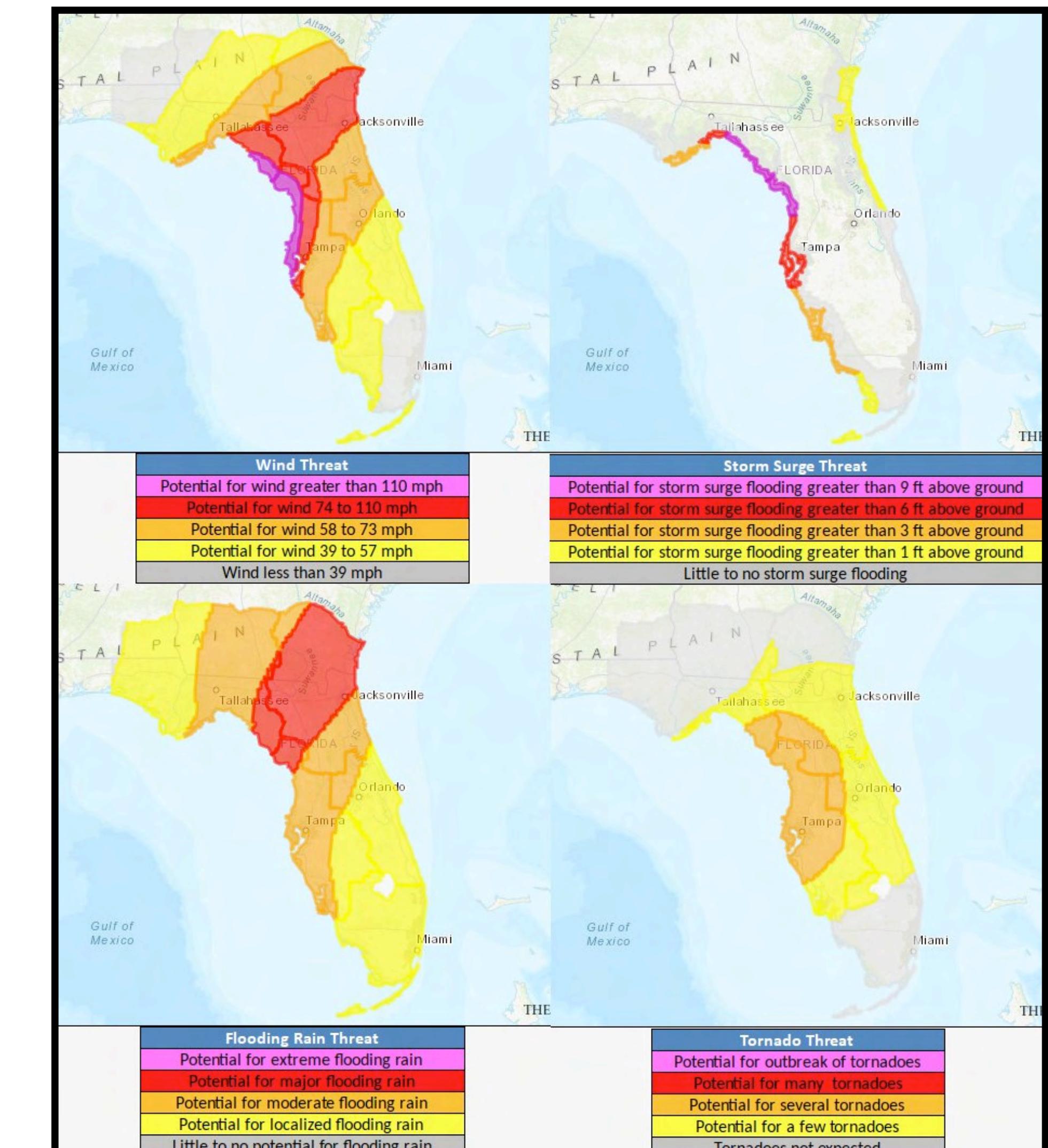


What do readers *really* need to know?

The possible path?



Or a hurricane's possible impacts (HTI)?



— Charts can mislead even if they are well designed —

Luís Melgar

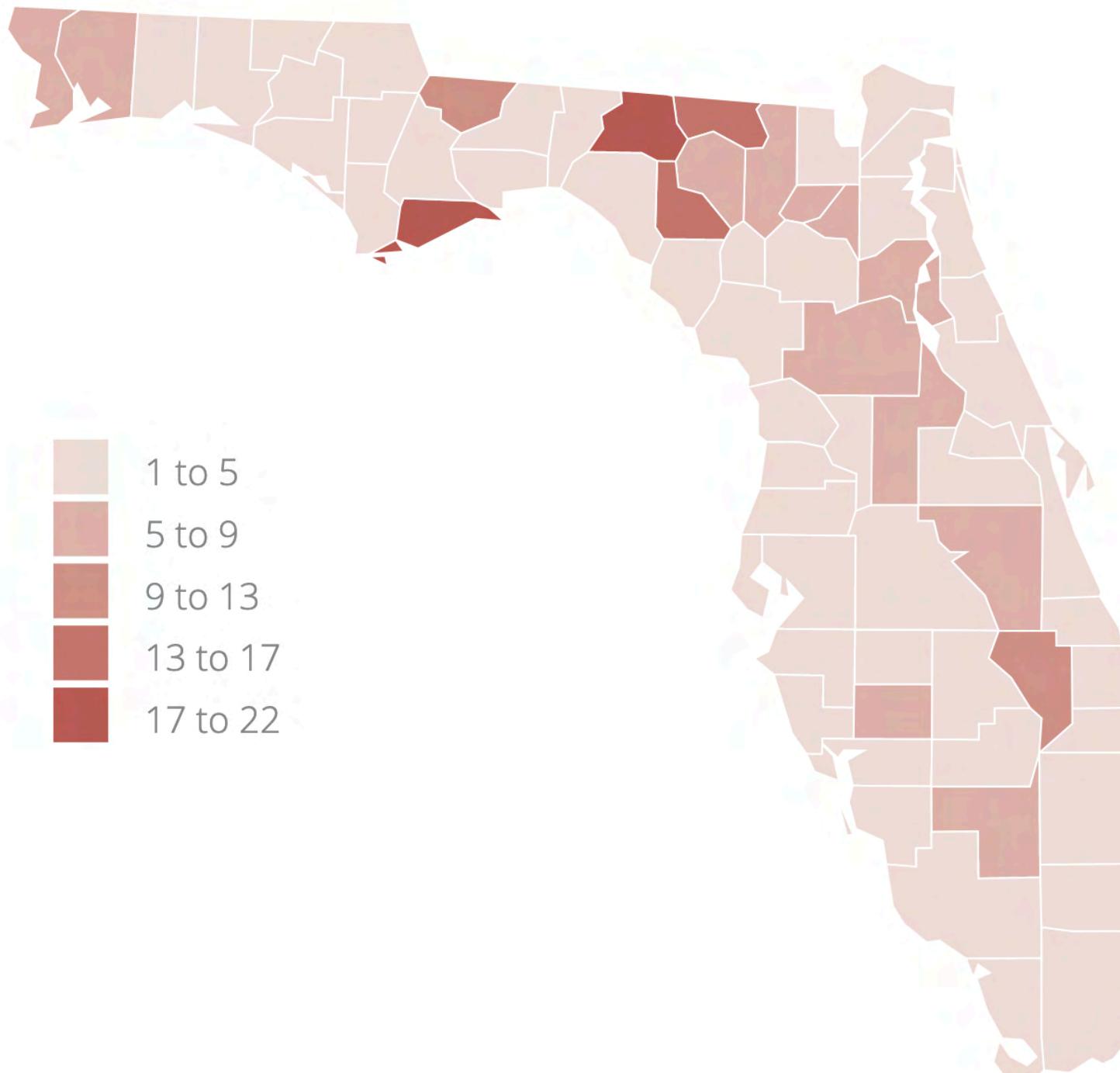




AT SCHOOL WITHOUT A ROOF

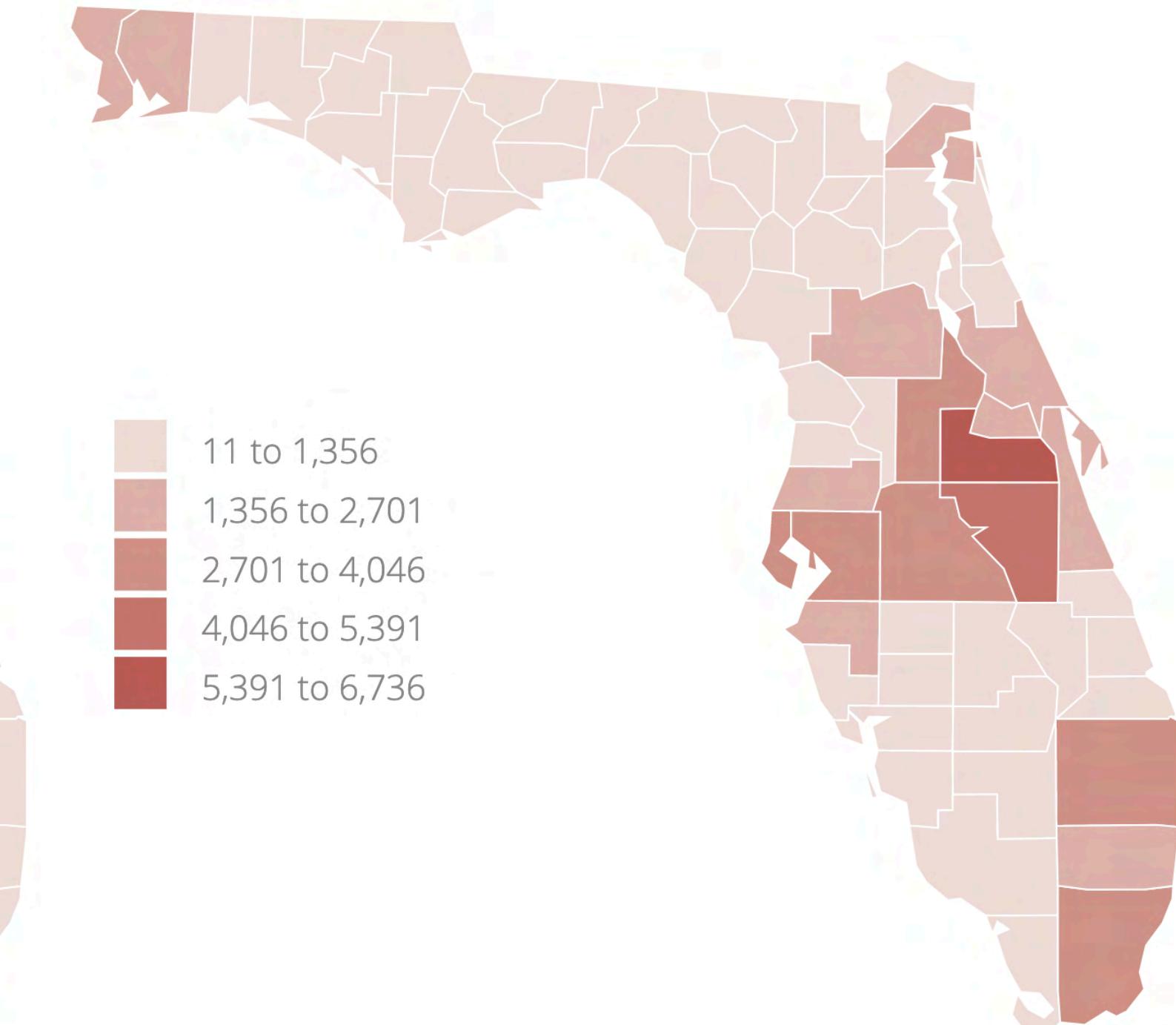
In Florida more than 71,000 students are homeless. During the last decade, this population rocketed as a result of the recession and how hard it has become for the poorest families to find affordable housing.

Percentage Total



1 to 5
5 to 9
9 to 13
13 to 17
17 to 22

Percentage Total



11 to 1,356
1,356 to 2,701
2,701 to 4,046
4,046 to 5,391
5,391 to 6,736

"A must-read for anyone who wants to stay informed."
—CATHY O'NEIL, best-selling author of *Weapons of Math Destruction*

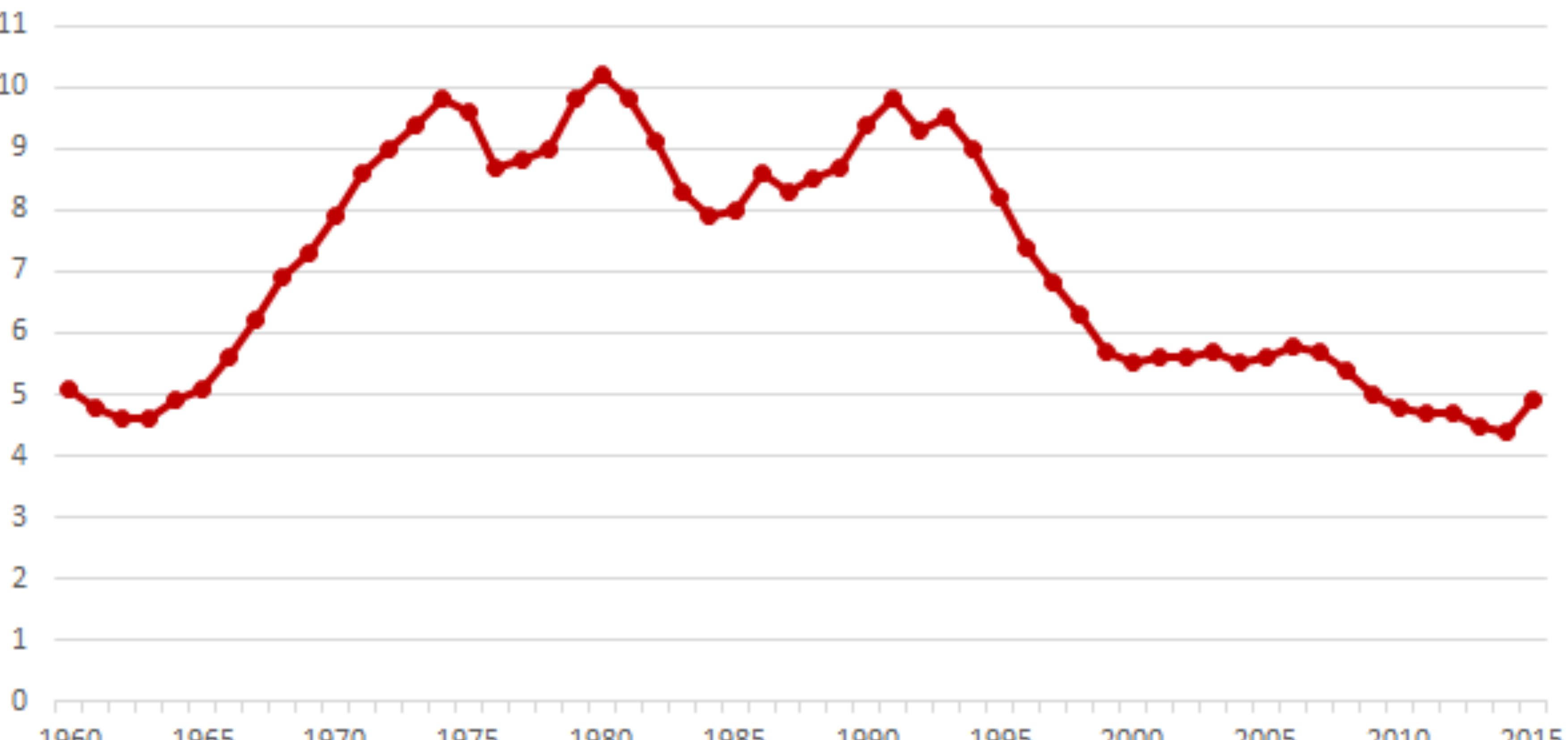
How Charts Lie



Getting Smarter about
Visual Information

Alberto Cairo

The U.S. Murder Rate (per 100,000 people)

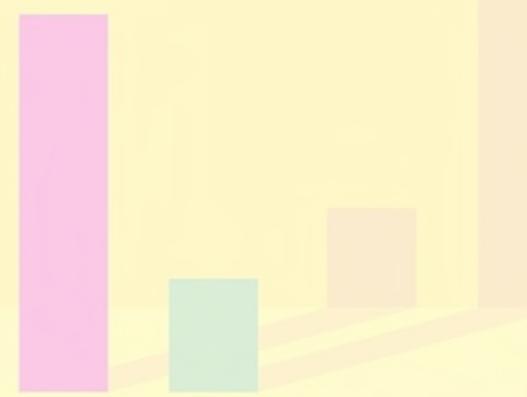


FactCheck.org chart based on FBI data

The danger of aggregating data too much,
and presenting just averages and other statistical summaries

"A must-read for anyone who wants to stay informed."
—CATHY O'NEIL, best-selling author of *Weapons of Math Destruction*

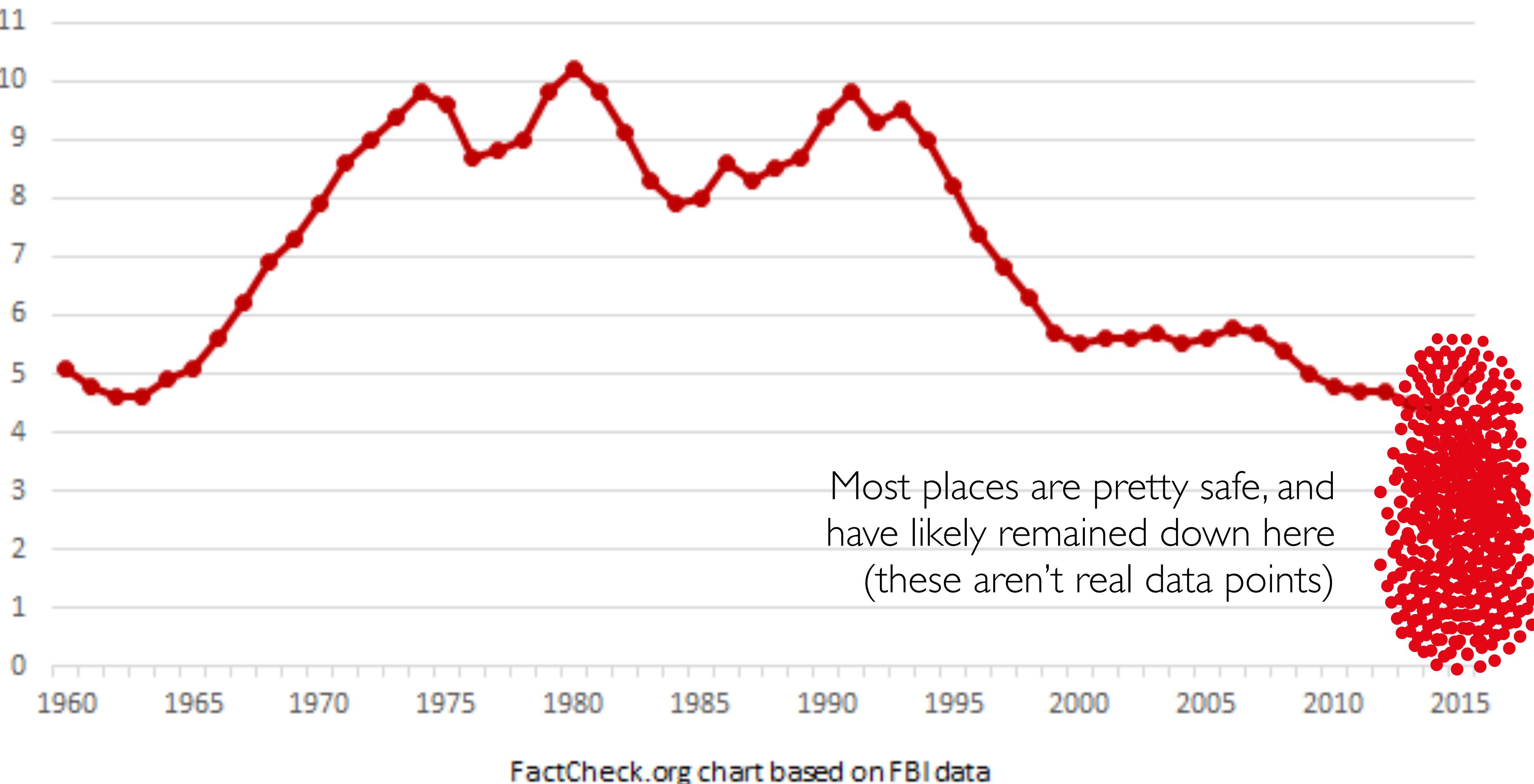
How Charts Lie



Getting Smarter about
Visual Information

Alberto Cairo

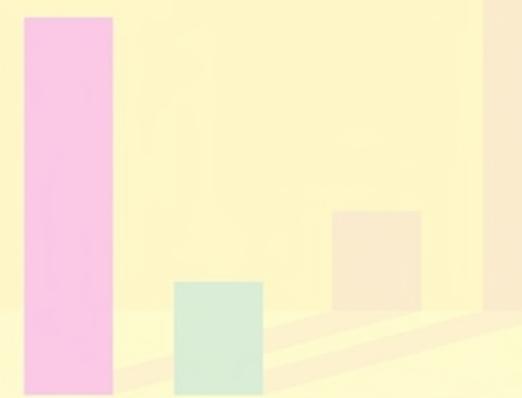
The U.S. Murder Rate (per 100,000 people)



The danger of aggregating data too much,
and presenting just averages and other statistical summaries

"A must-read for anyone who wants to stay informed."
—CATHY O'NEIL, best-selling author of *Weapons of Math Destruction*

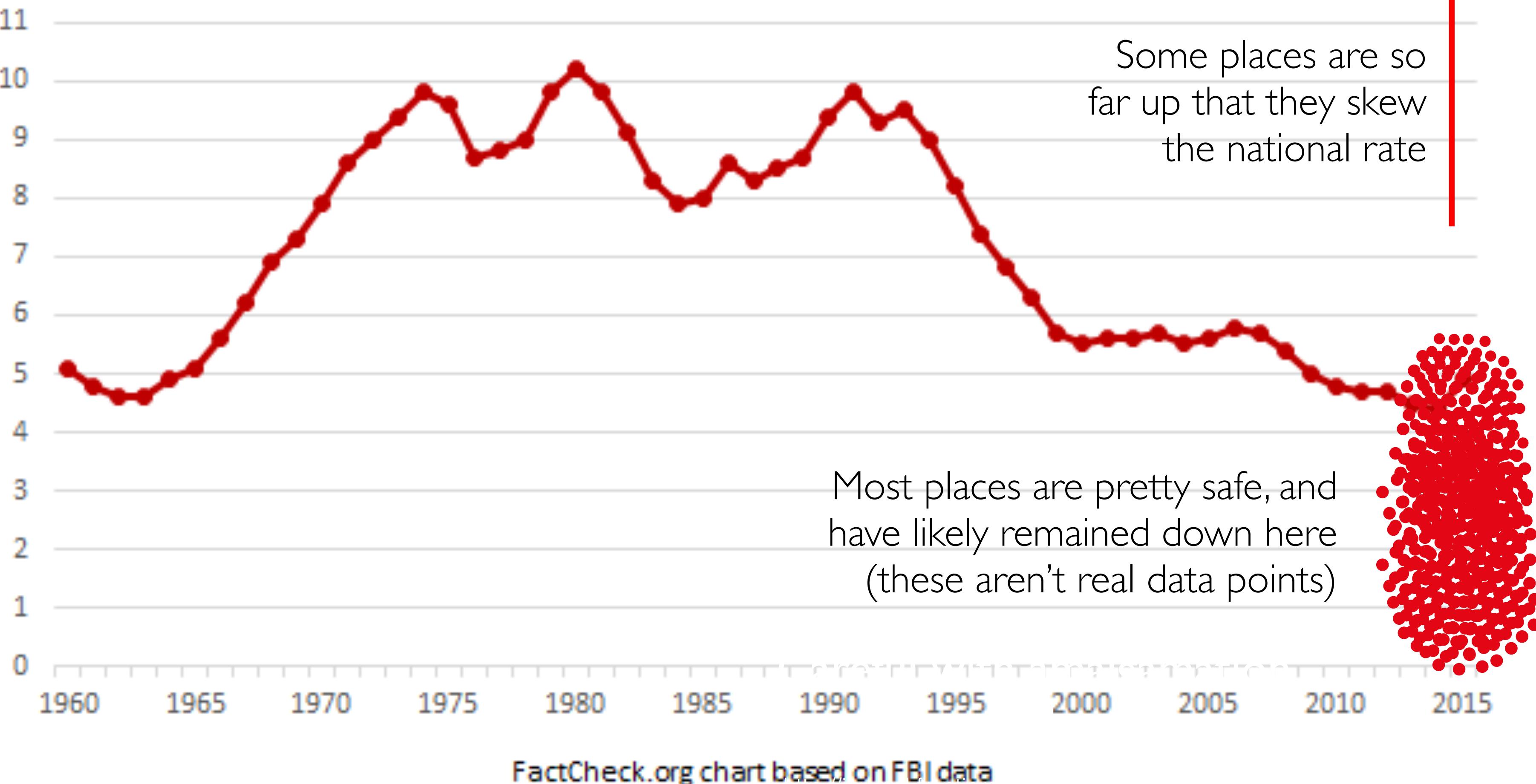
How Charts Lie

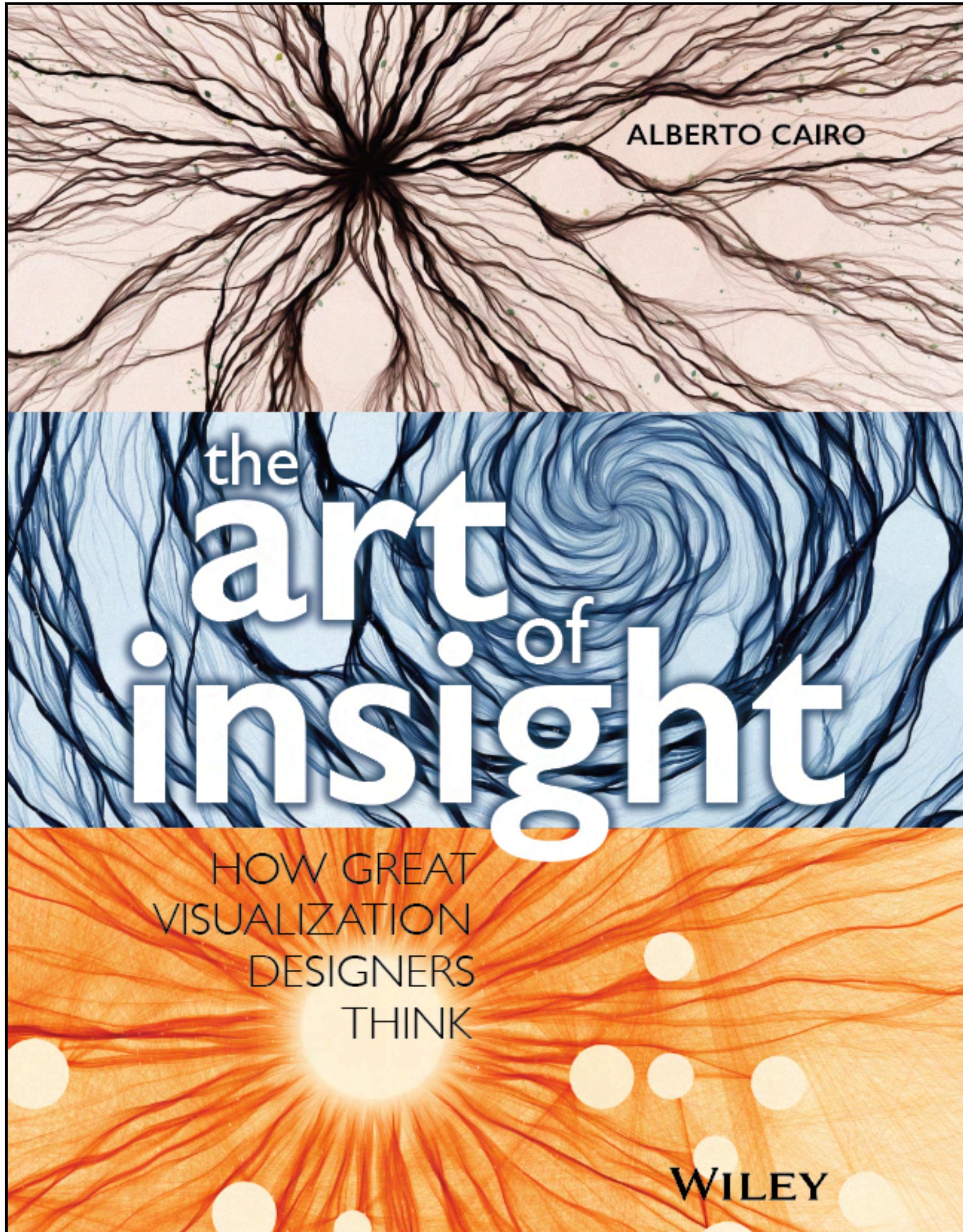


Getting Smarter about
Visual Information

Alberto Cairo

The U.S. Murder Rate (per 100,000 people)





Visualization design consists of **reasoning** about possible **choices** by considering the interplay between:

1. **Content:** The nature, origin, and limitations of the data.
2. **People:** Your audience.
3. **Intent:** The purpose(s) that we define.
4. **Constraints:** The limitations that we may face.
5. **Outcomes:** How the graphic is received.

Every design choice must be **deliberate**.

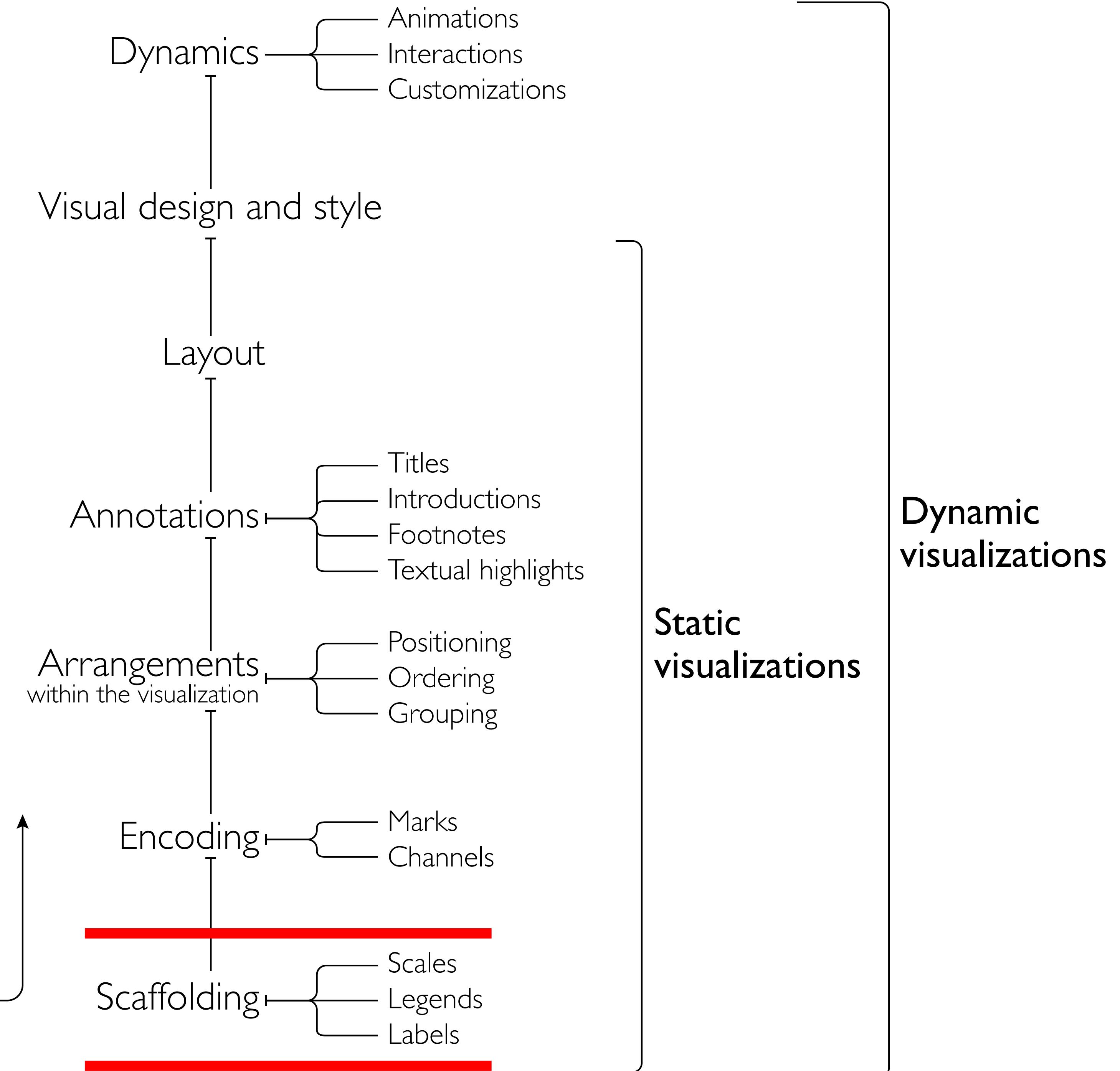
It is inevitably **subjective**, but it should never be **arbitrary**.

Reasoning about the components of a visualization

Visualization: Layers and elements to think about

**Considerations about
the information
to be visualized
—they influence
everything else.**

**Read from
the bottom-up**



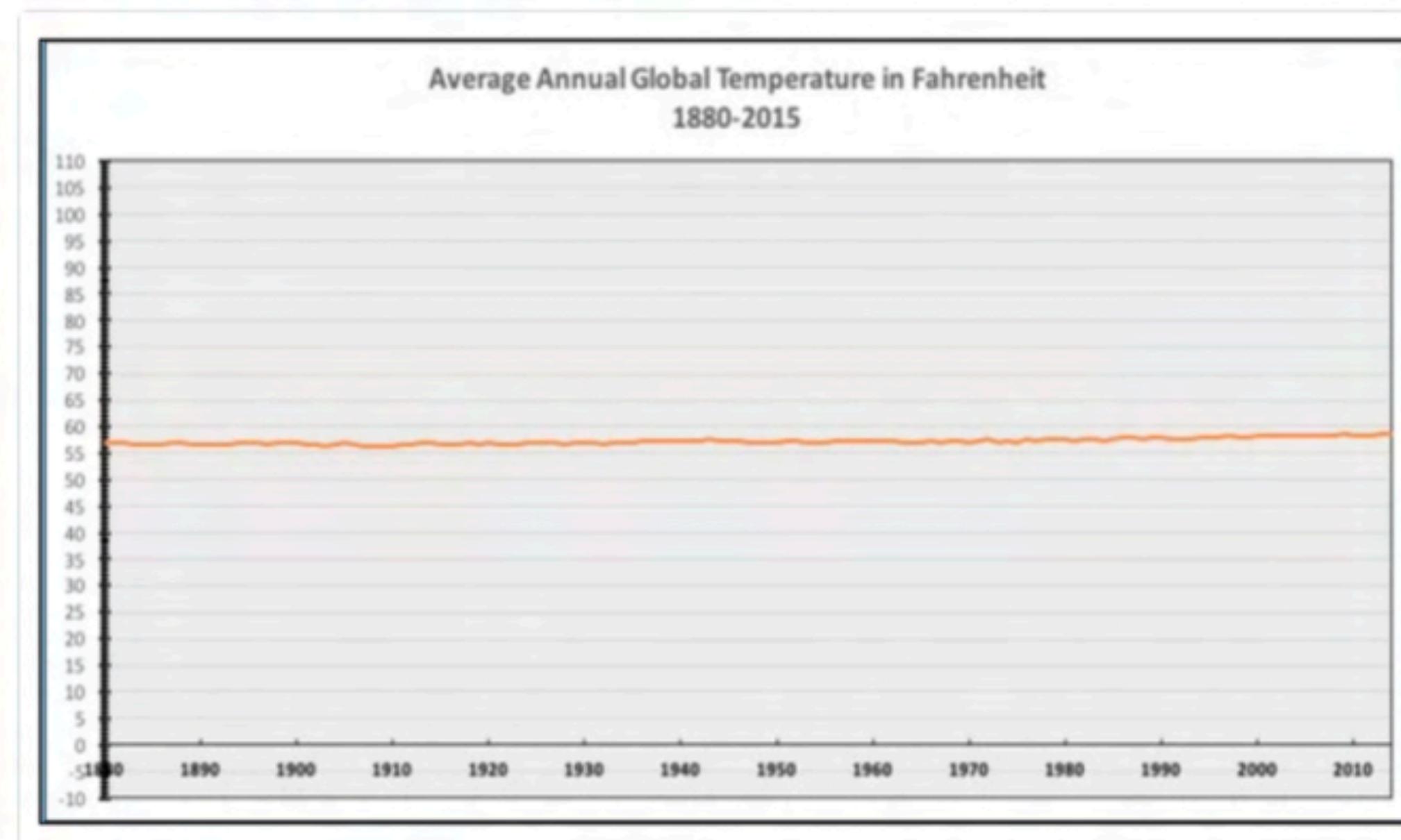
Decisions related to scales and legends can deeply affect our perception of the data



Follow

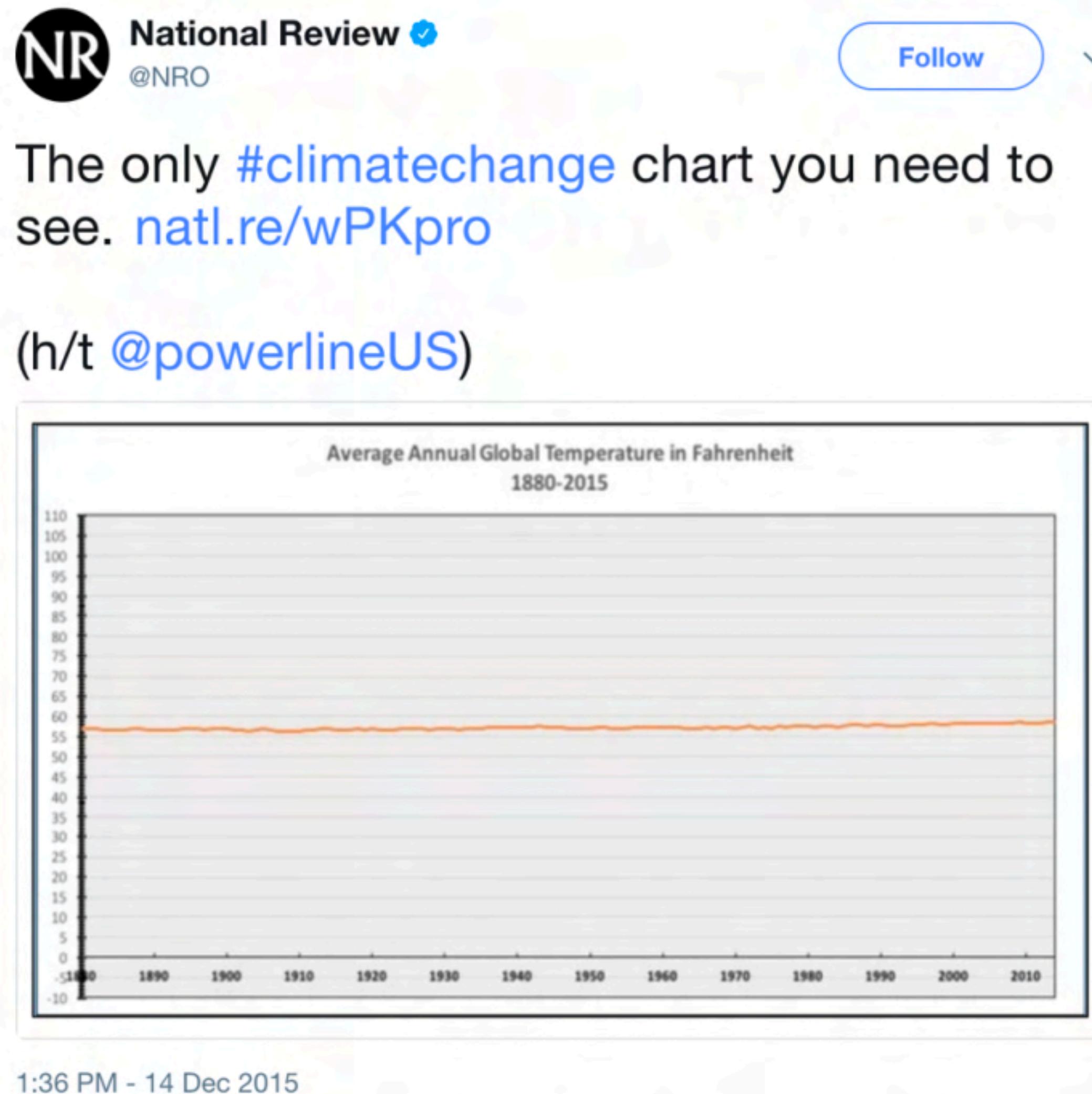
The only #climatechange chart you need to see. natl.re/wPKpro

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

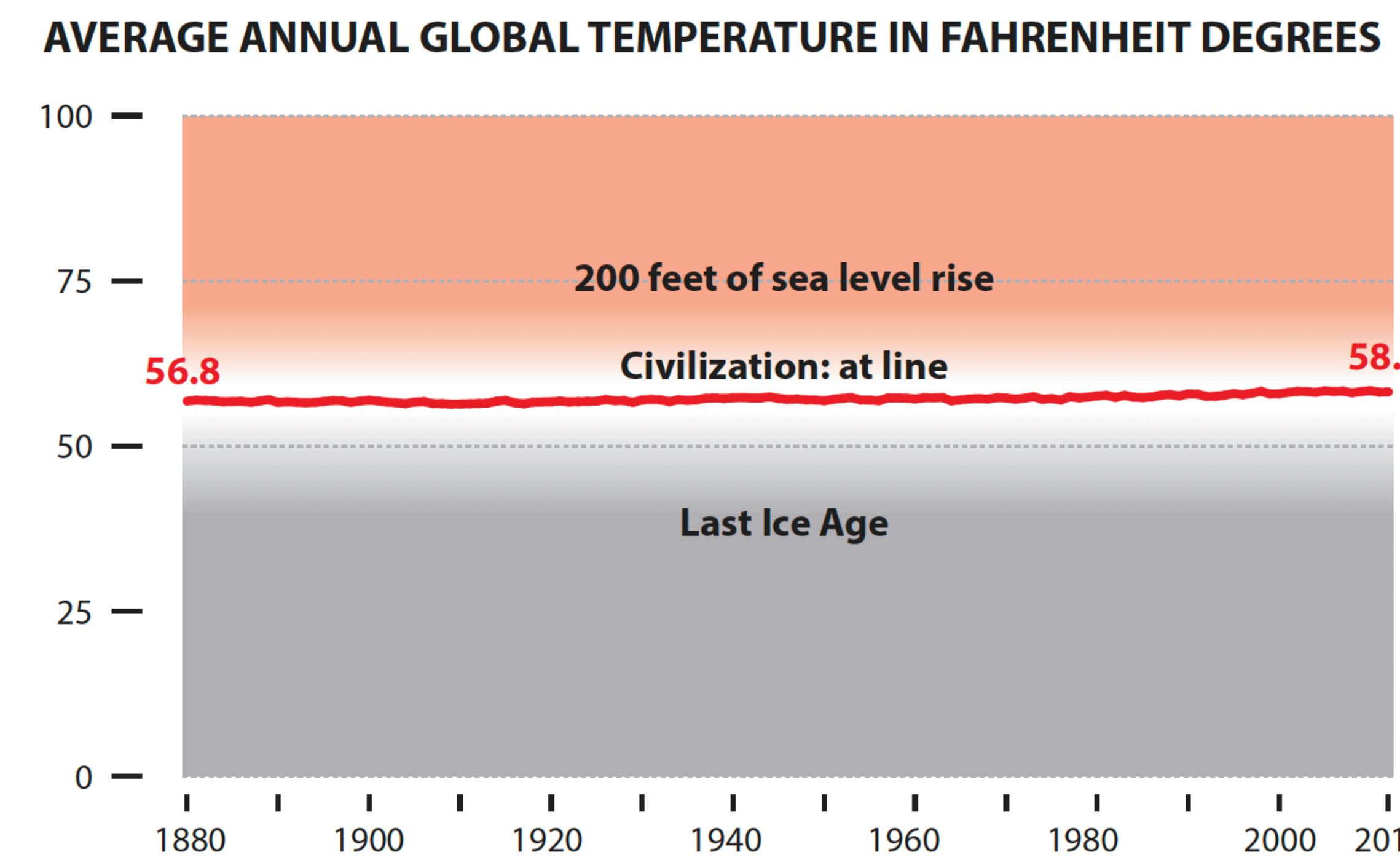


1:36 PM - 14 Dec 2015

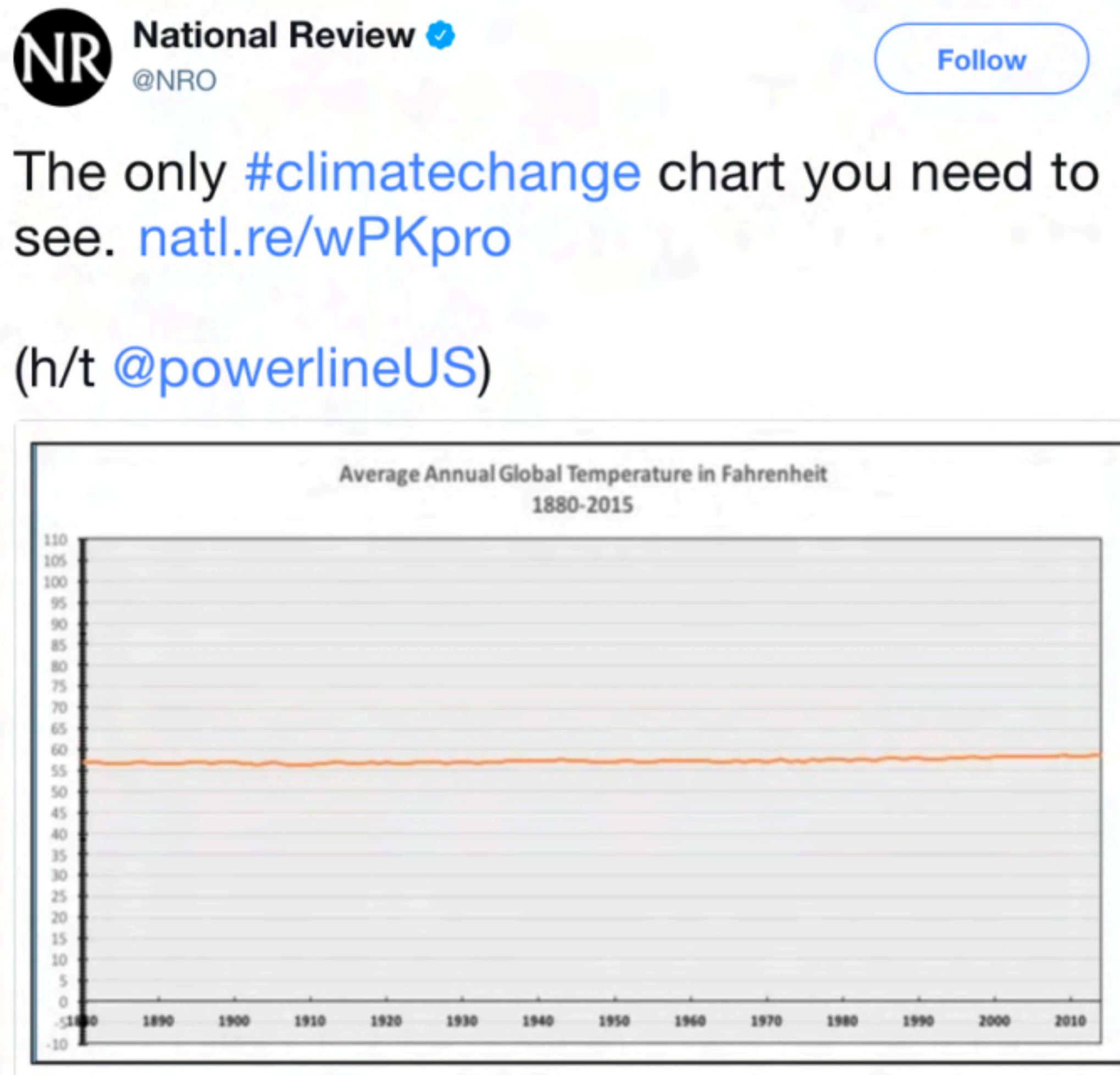
Decisions related to scales and legends can deeply affect our perception of the data



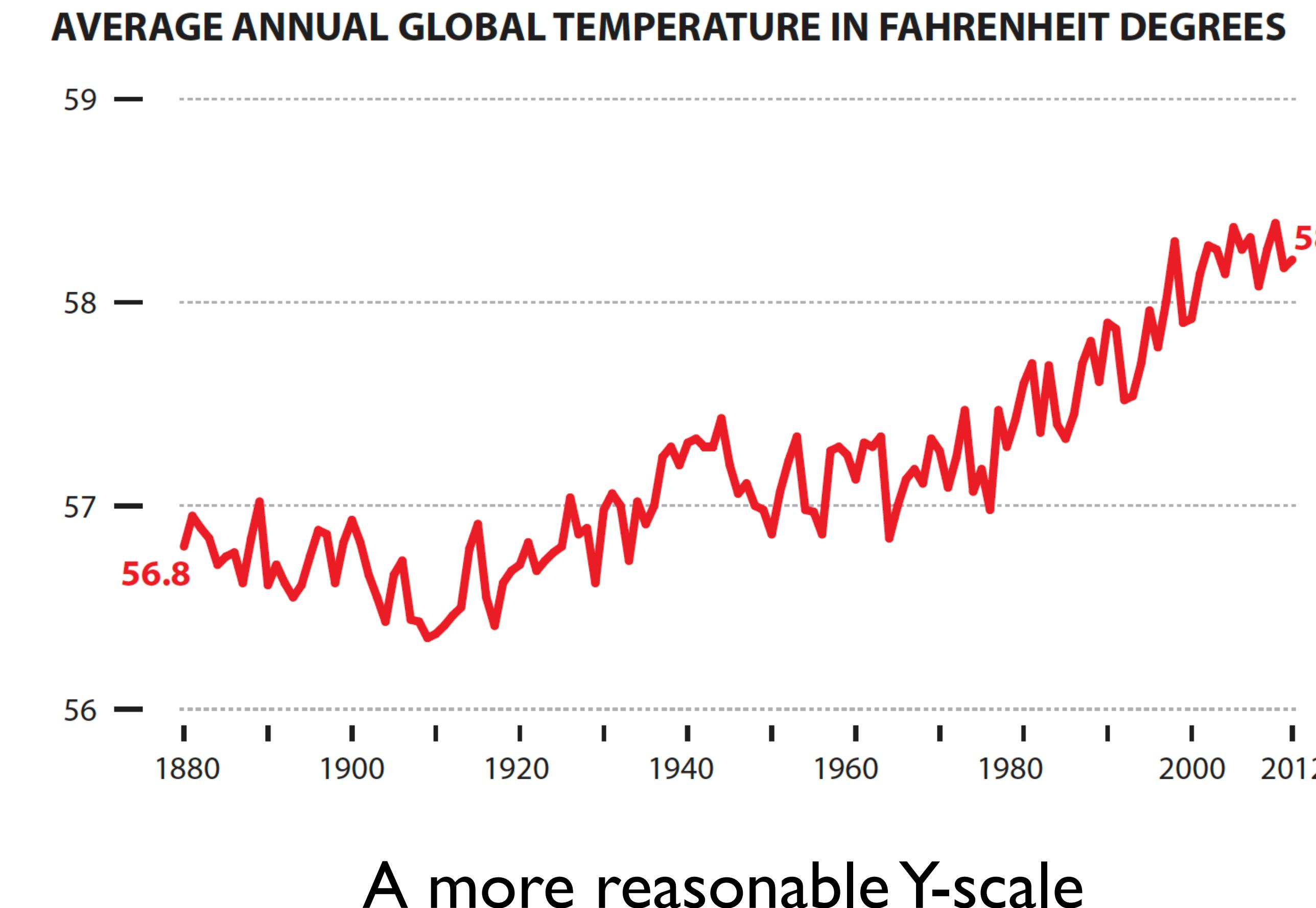
Richard Reiss, a research fellow at the City University of New York's Institute for Sustainable Cities added some tongue-in-cheek annotations to the original chart that reveal one of the many reasons why the choice of scales is so wrong:



Decisions related to scales and legends can deeply affect our perception of the data



1:36 PM - 14 Dec 2015



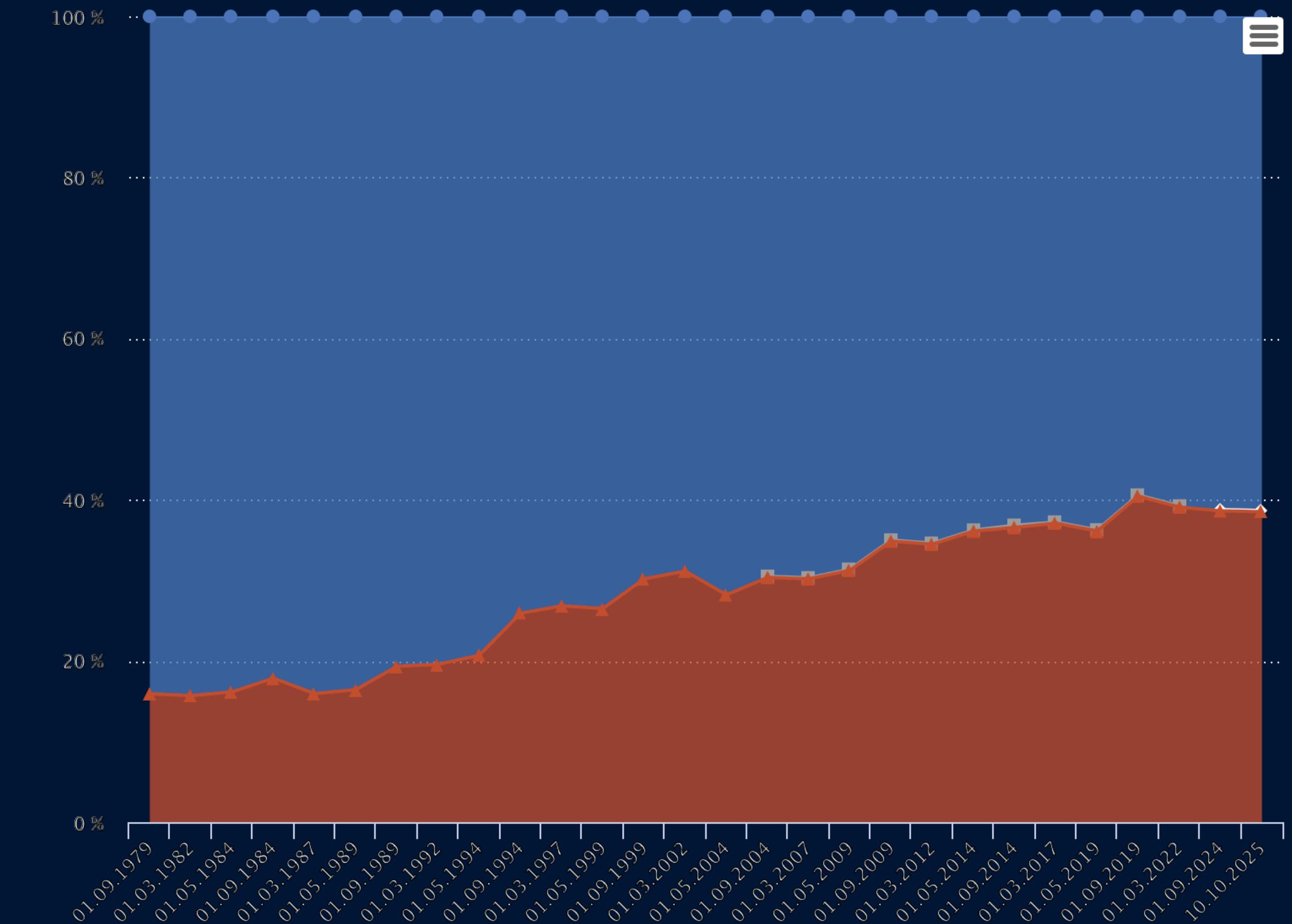
Proportion of female and male representation

This area chart shows the share of female and male representatives in the EP since 1979.

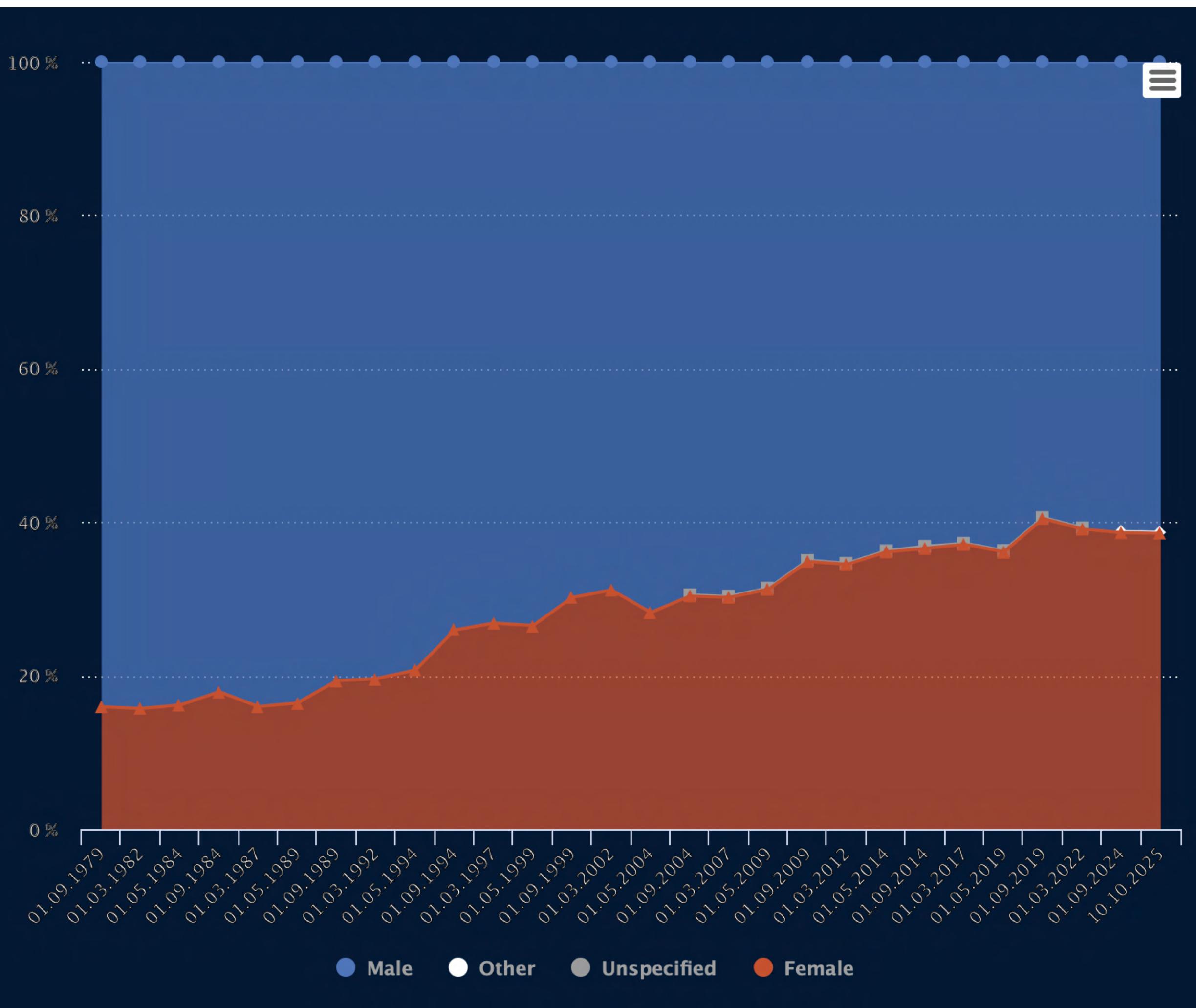
If you roll your mouse over the chart, you will see the number of female and male MEPs at the date shown.

Source: European Parliament

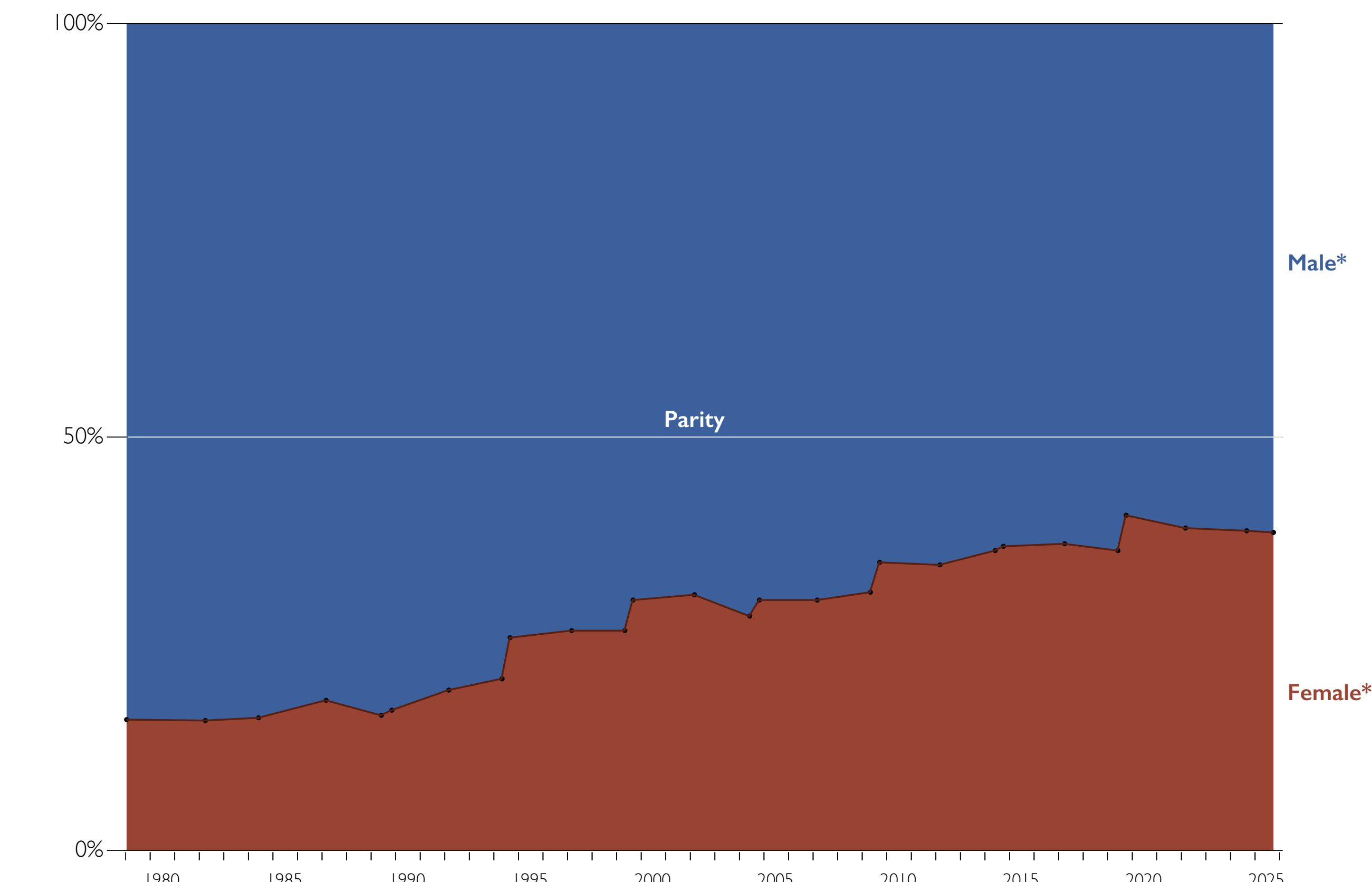
Notice that intervals on the X scale are even—but they shouldn't be, because time intervals aren't even!



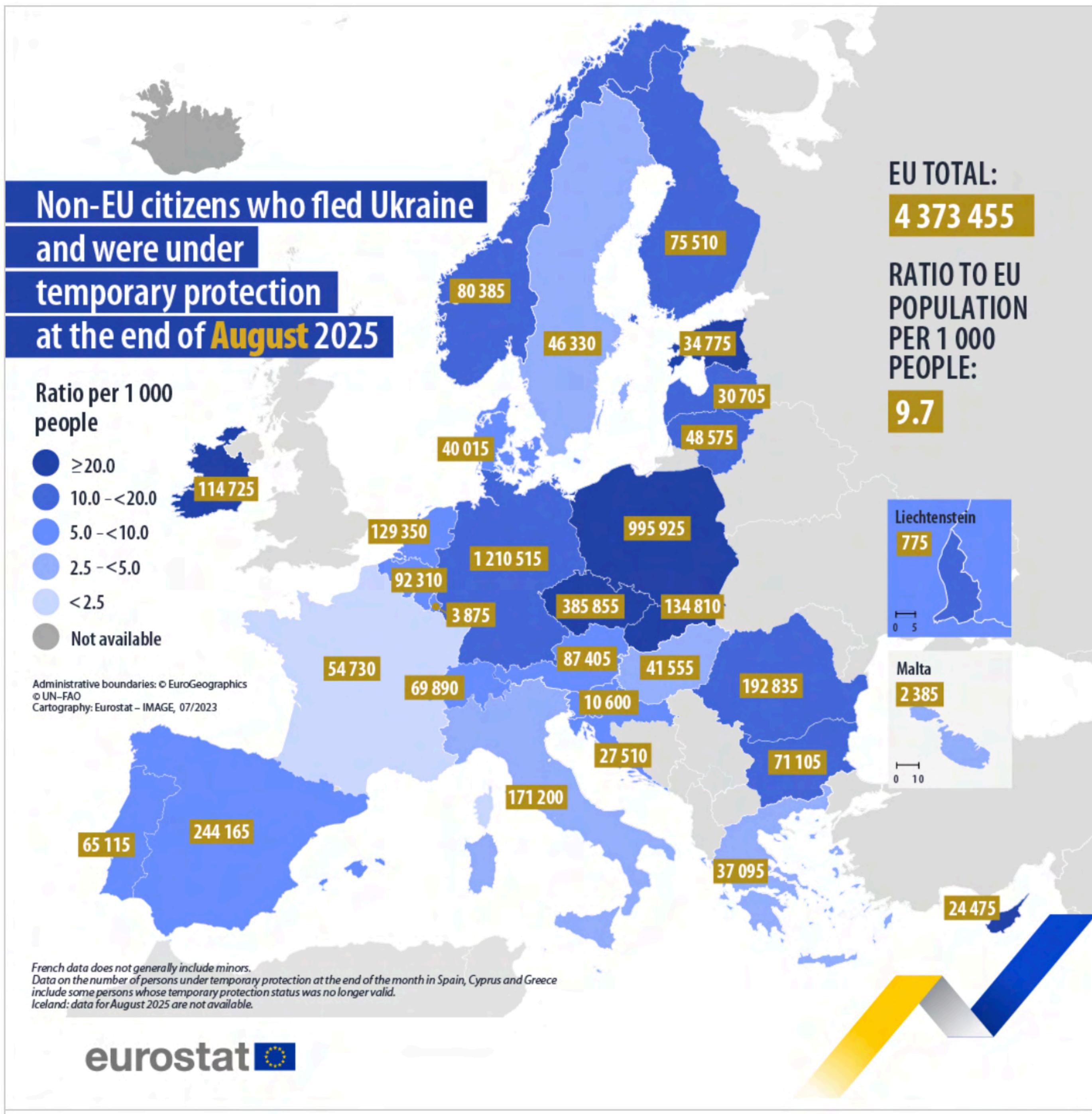
Original



Alternative (warning: this is just a quick draft)

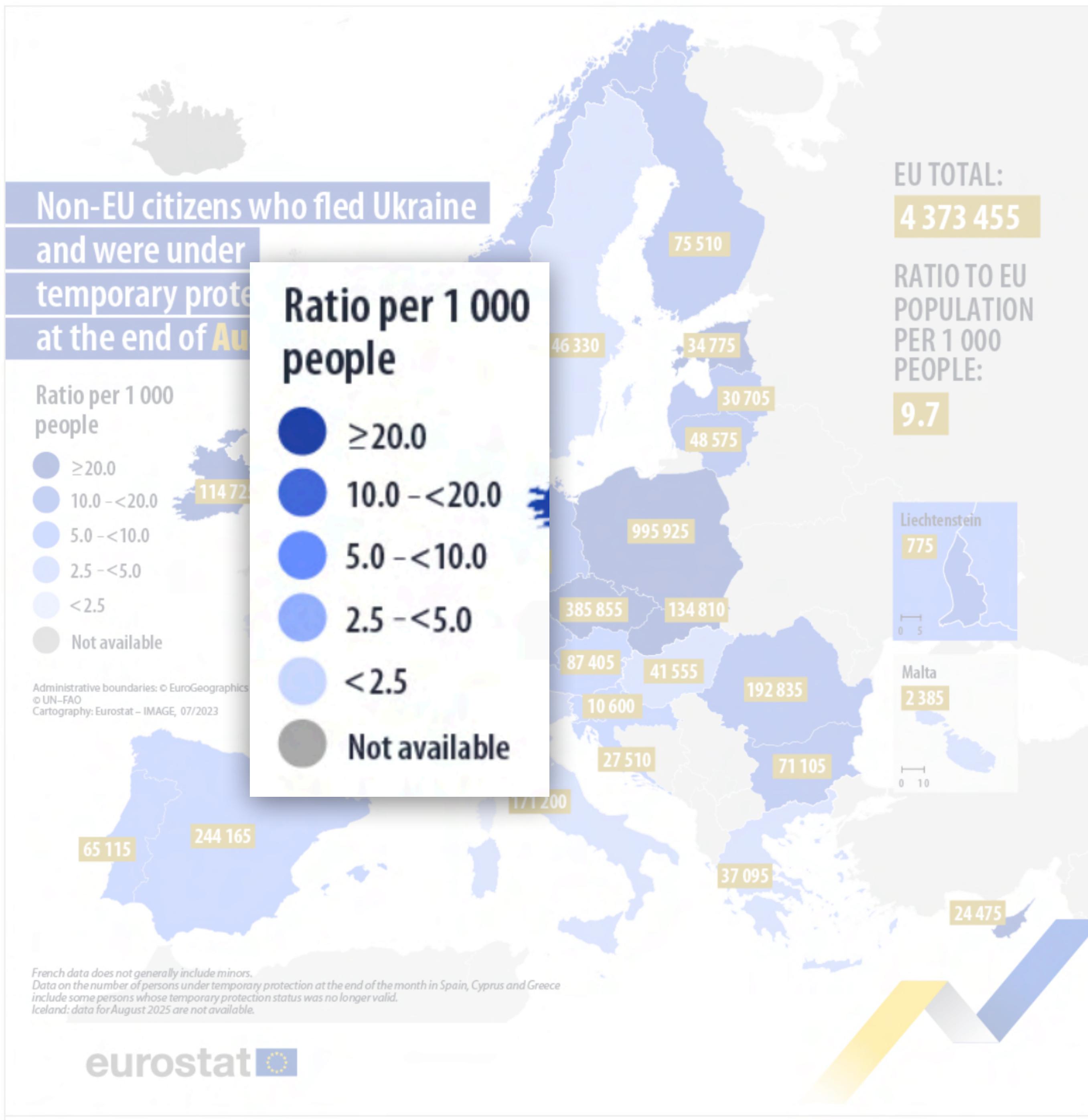


***Note about this quick-and-dirty chart makeover:** The original included the categories Other and Unspecified, corresponding to non-binary or transgender people. The percentages of these categories are too small to be represented accurately on a 0-100% scale. Therefore, an additional chart just for them would be needed.



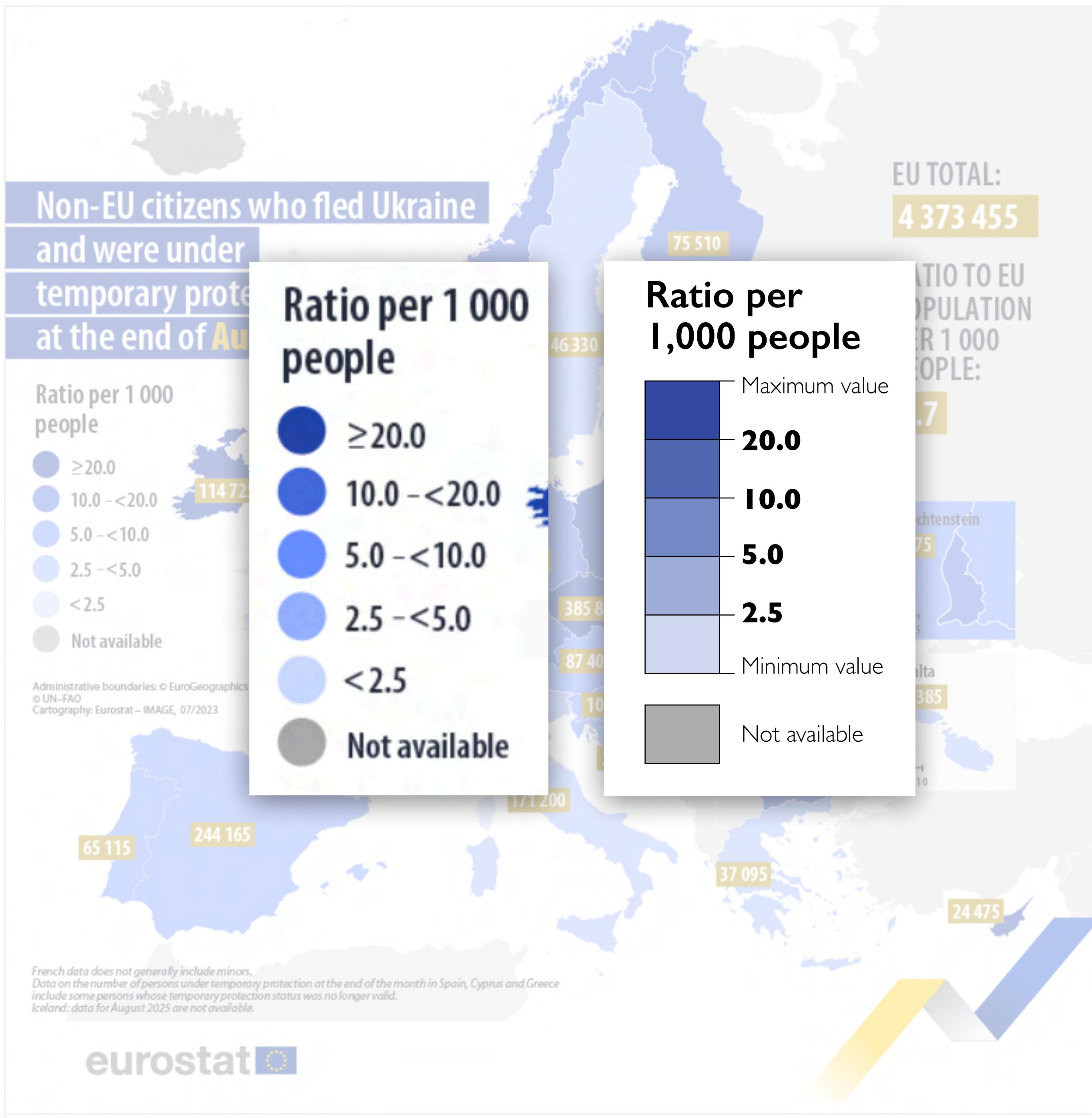
Map scales

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Temporary_protection_for_persons_fleeing_Ukraine_-_monthly_statistics



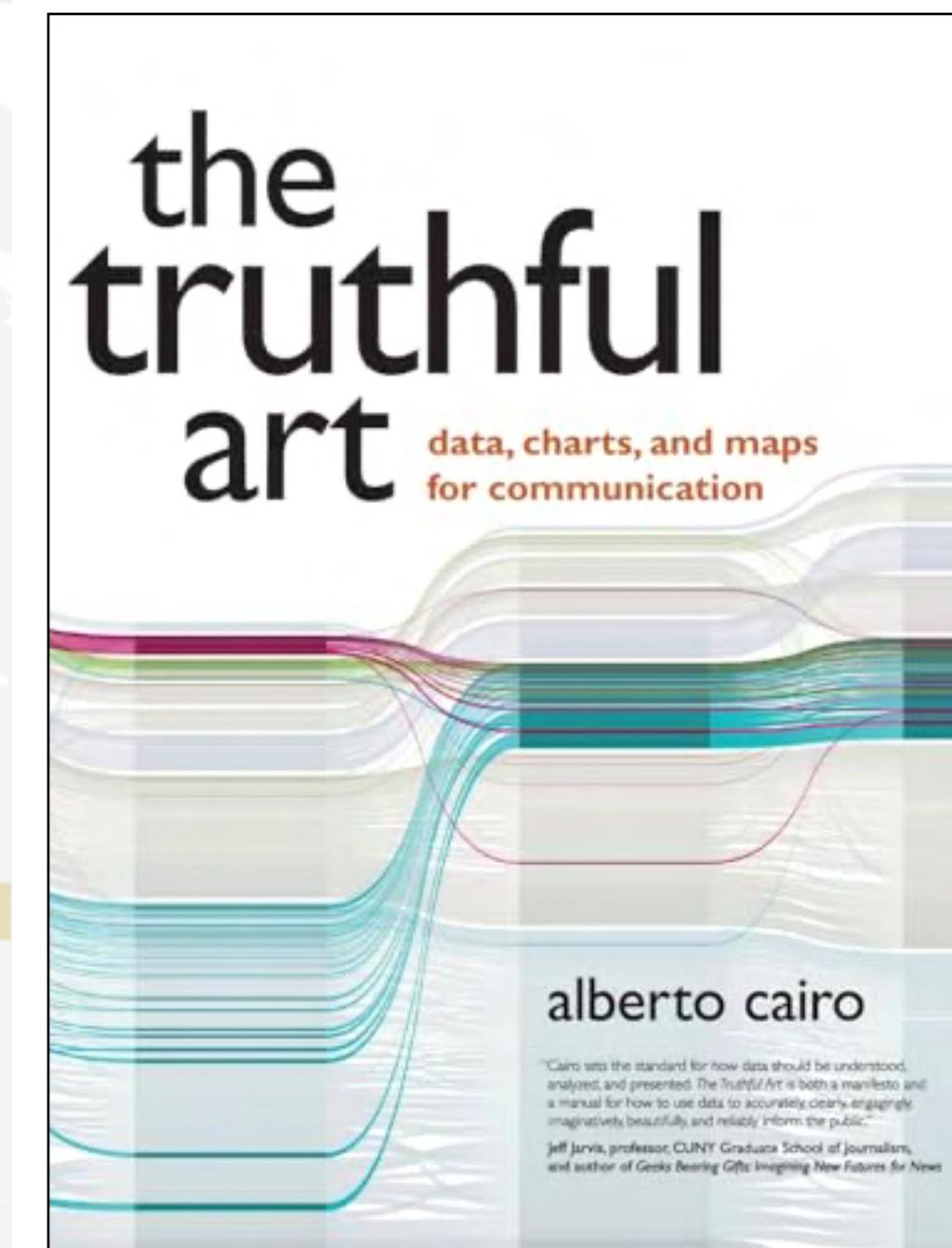
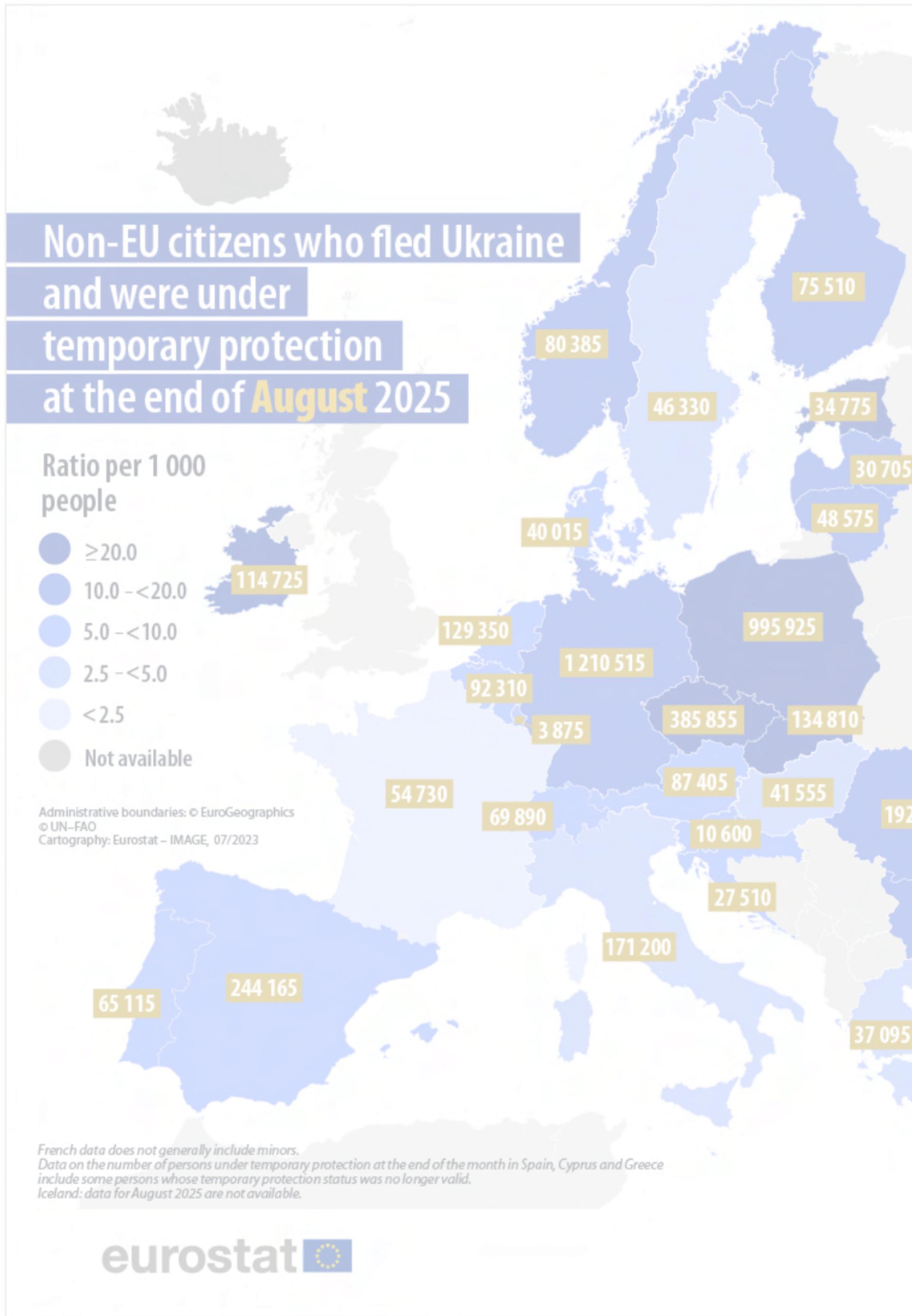
Map scales

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Temporary_protection_for_persons_fleeing_Ukraine_-_monthly_statistics



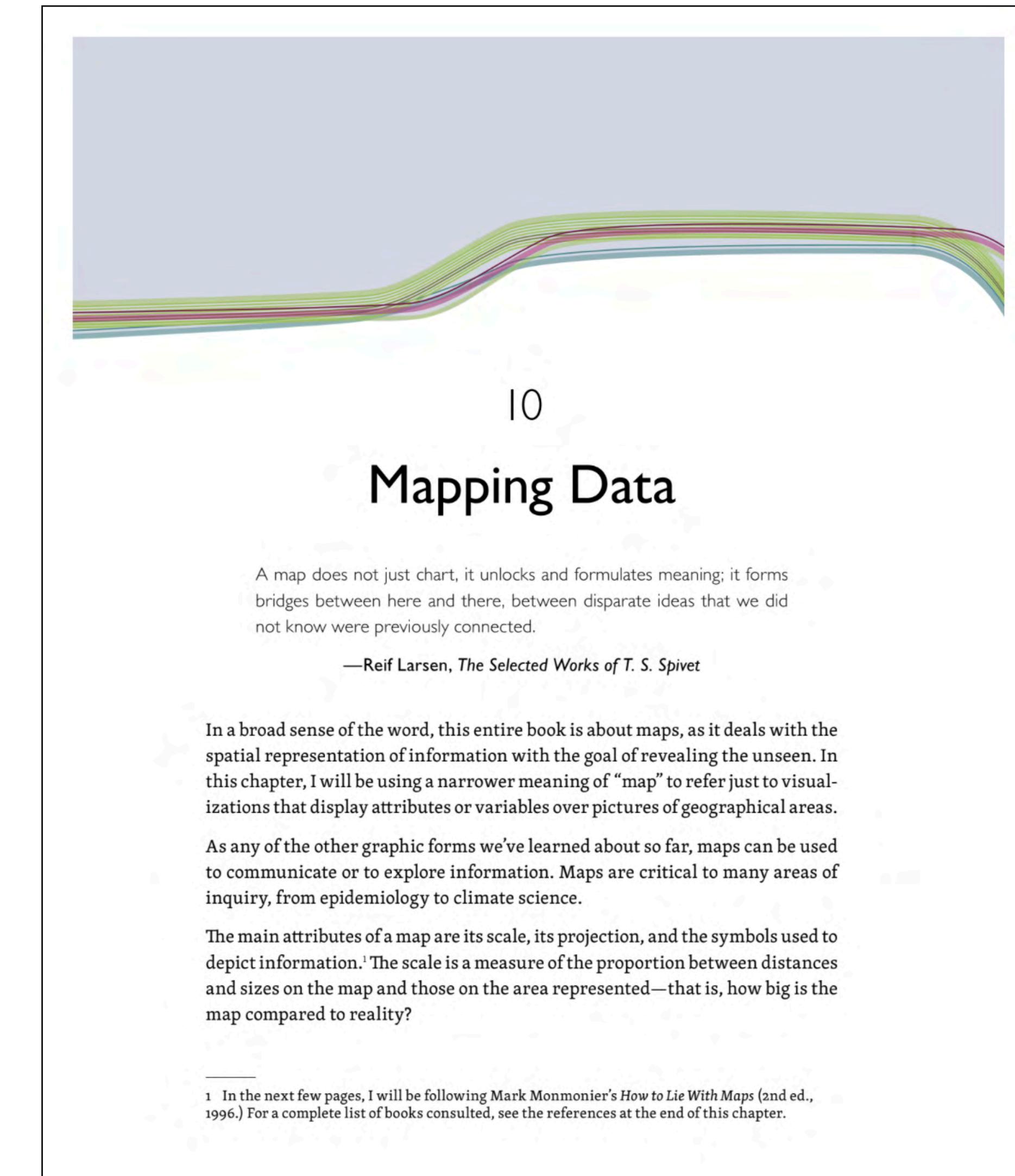
Map scales

https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Temporary_protection_for_persons_fleeing_Ukraine_-_monthly_statistics



Chapter 10 of The Truthful Art

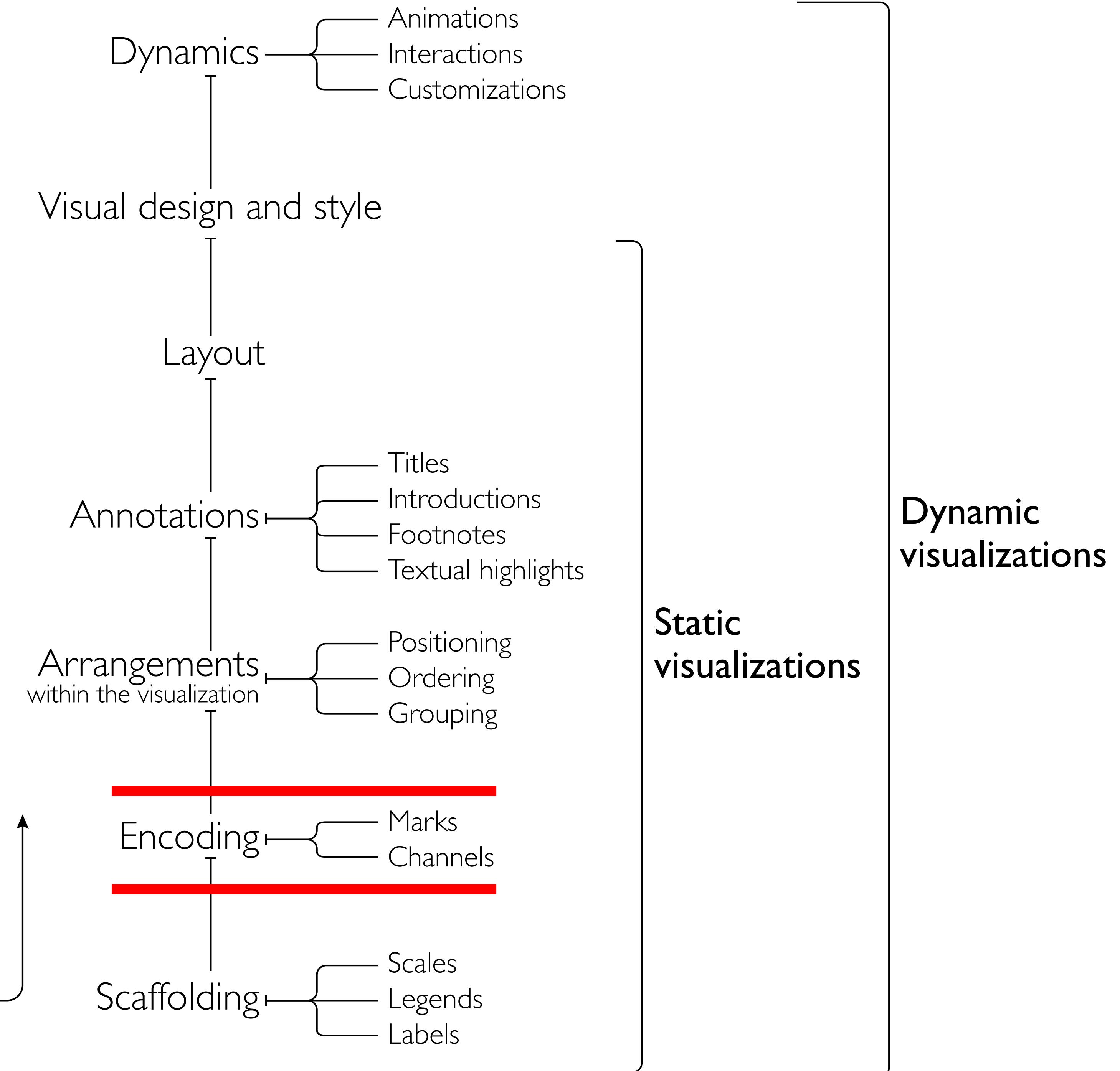
<https://drive.google.com/file/d/1Ap4ehrBDXBSDPcnPXCTeg9cVH1U9kOeX/>



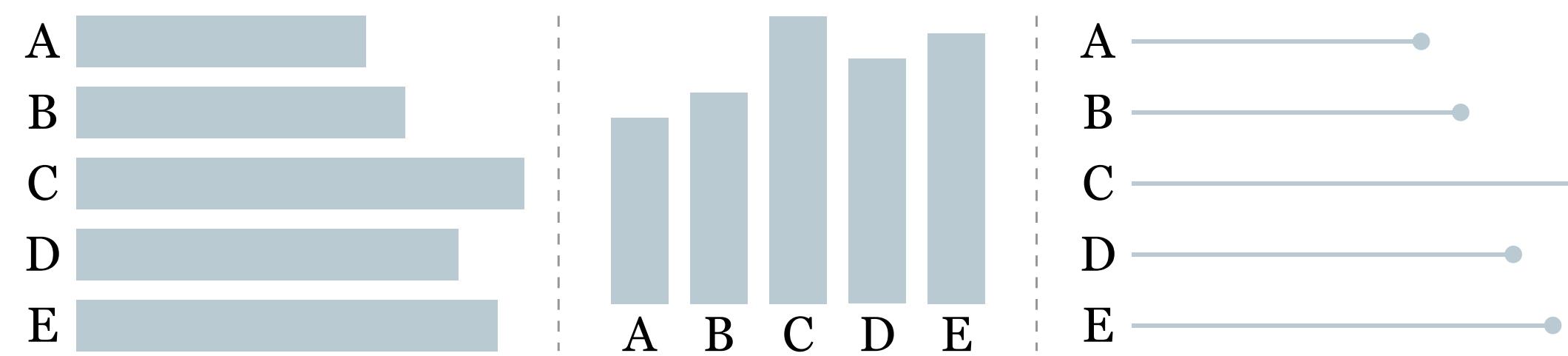
Visualization: Layers and elements to think about

**Considerations about
the information
to be visualized
—they influence
everything else.**

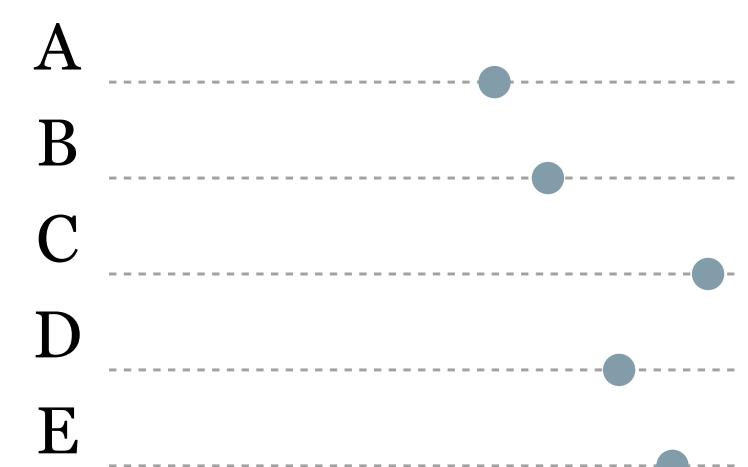
**Read from
the bottom-up**



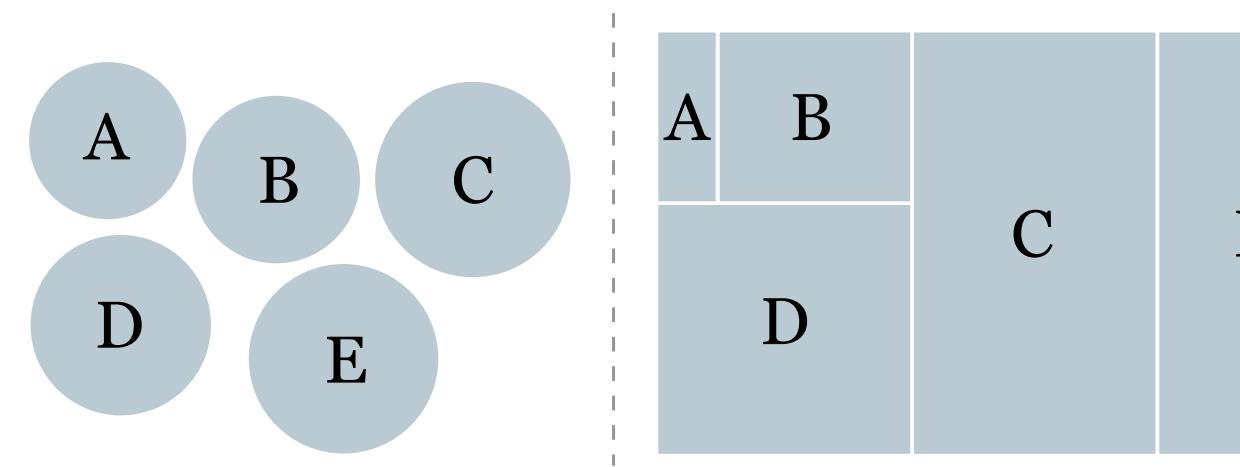
Length or height



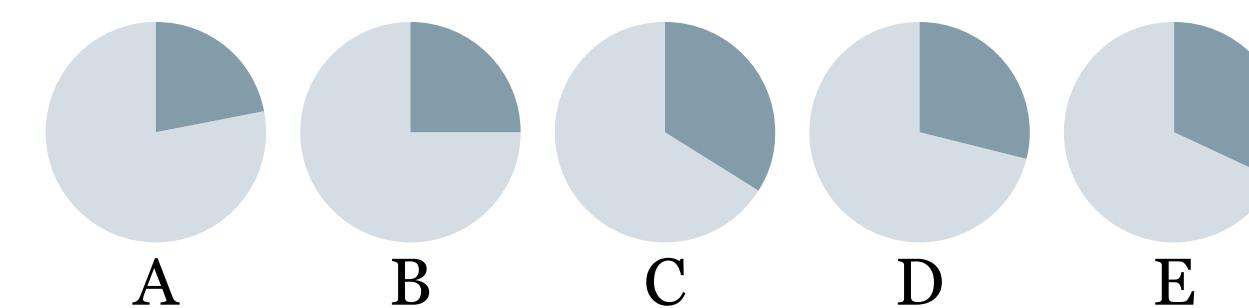
Position



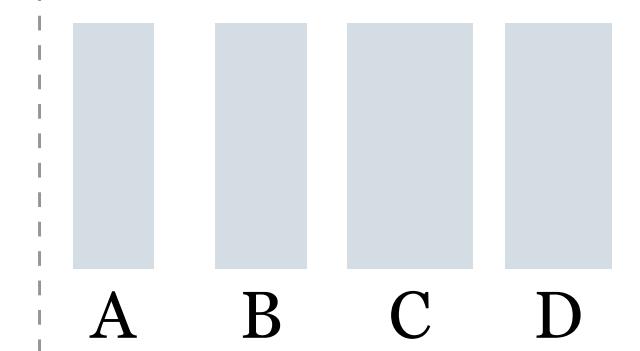
Area



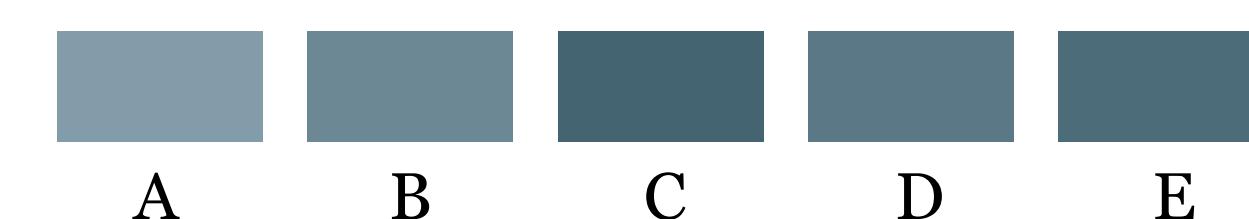
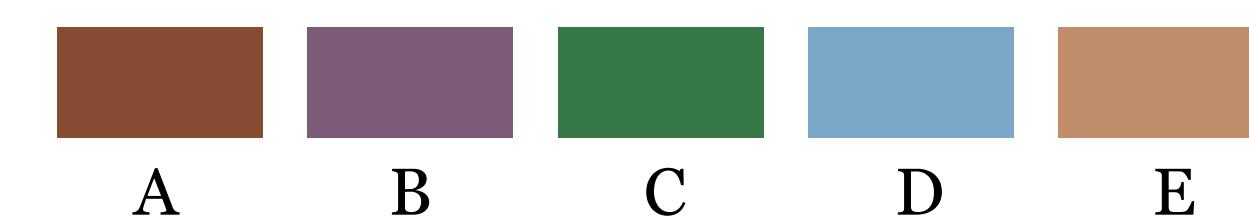
Angle/area



Line weight



Hue and shade



Figures represented
in all these graphics:
22%, 25%, 34%, 29%, 32%

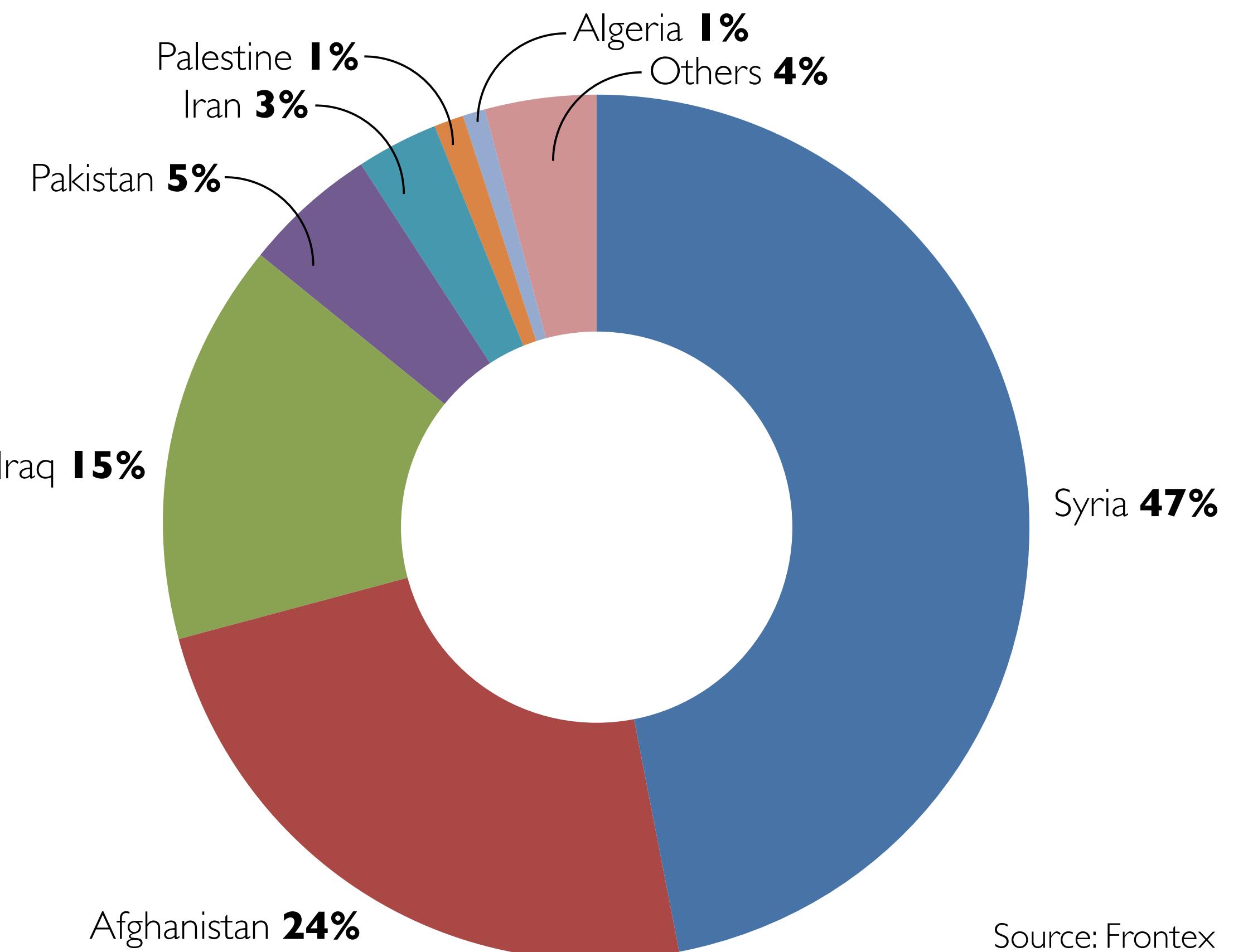
Encoding

Data visualization consists of mapping data onto attributes of objects—commonly abstract shapes, called **“marks”**.

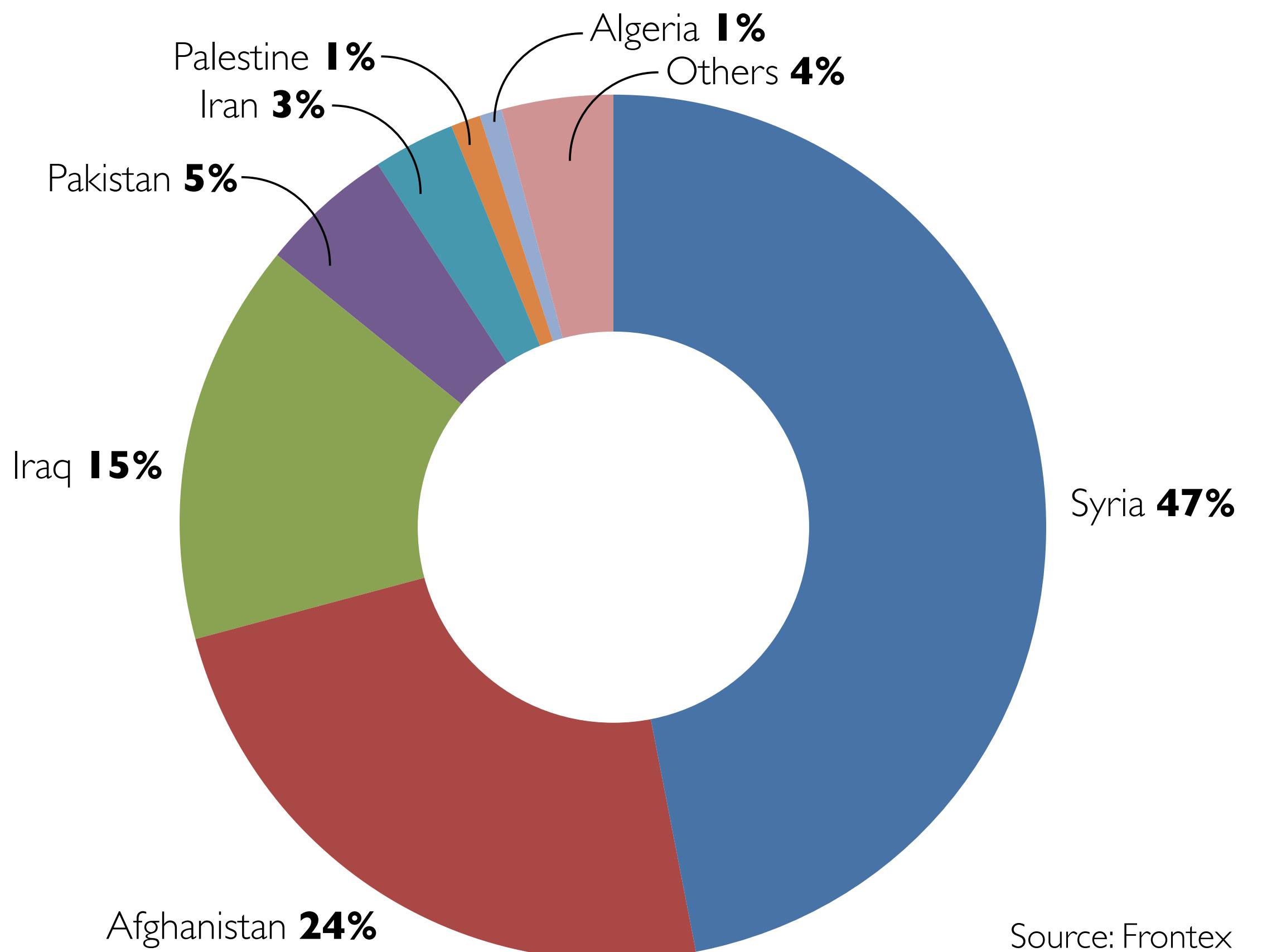
These attributes that vary with the data are called **“visual channels”**.

[https://web.cse.ohio-state.edu/~shen.94/
Melbourne/Slides/TamaraChp5.pdf](https://web.cse.ohio-state.edu/~shen.94/Melbourne/Slides/TamaraChp5.pdf)

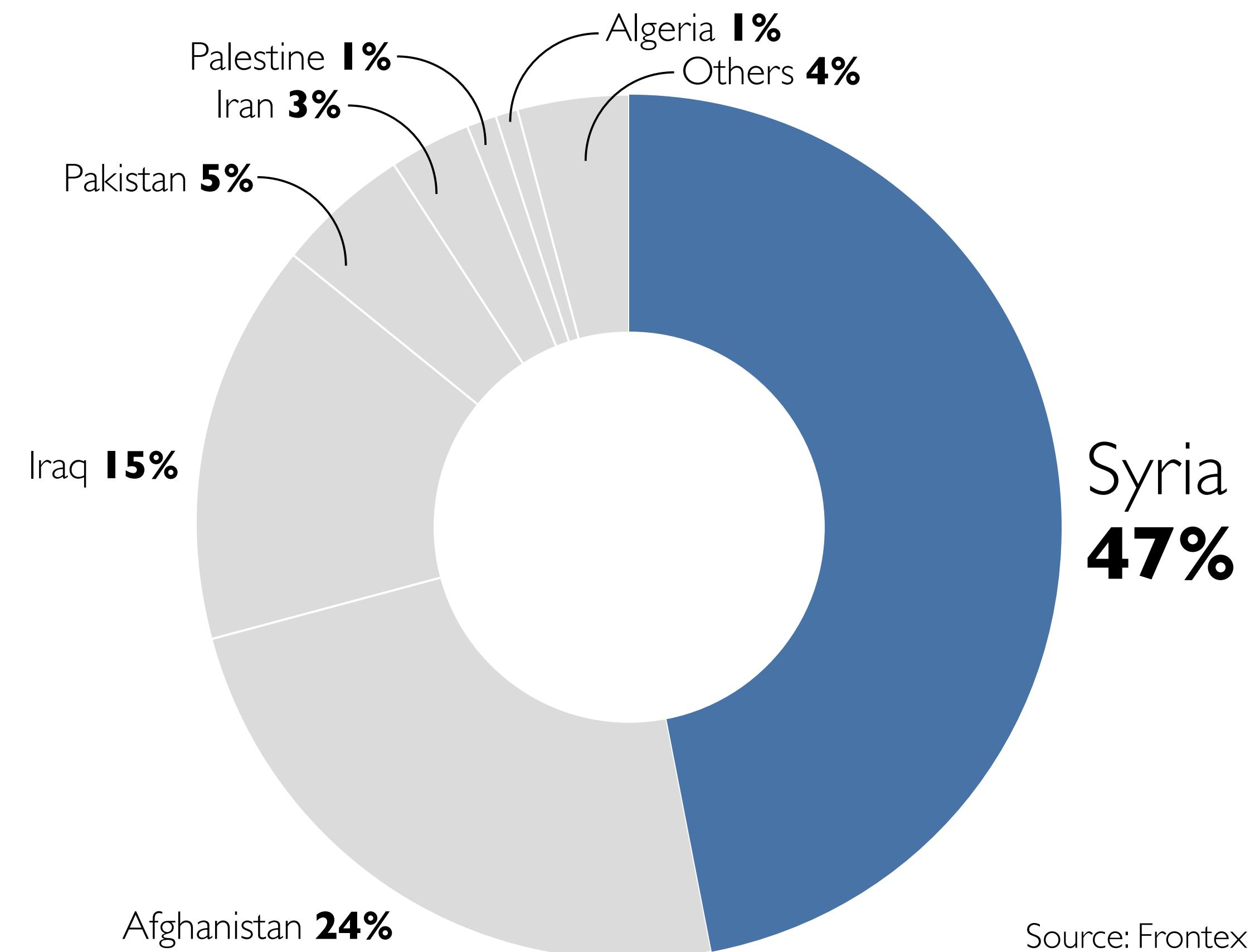
GREECE - Main nationalities of arriving migrants in 2016

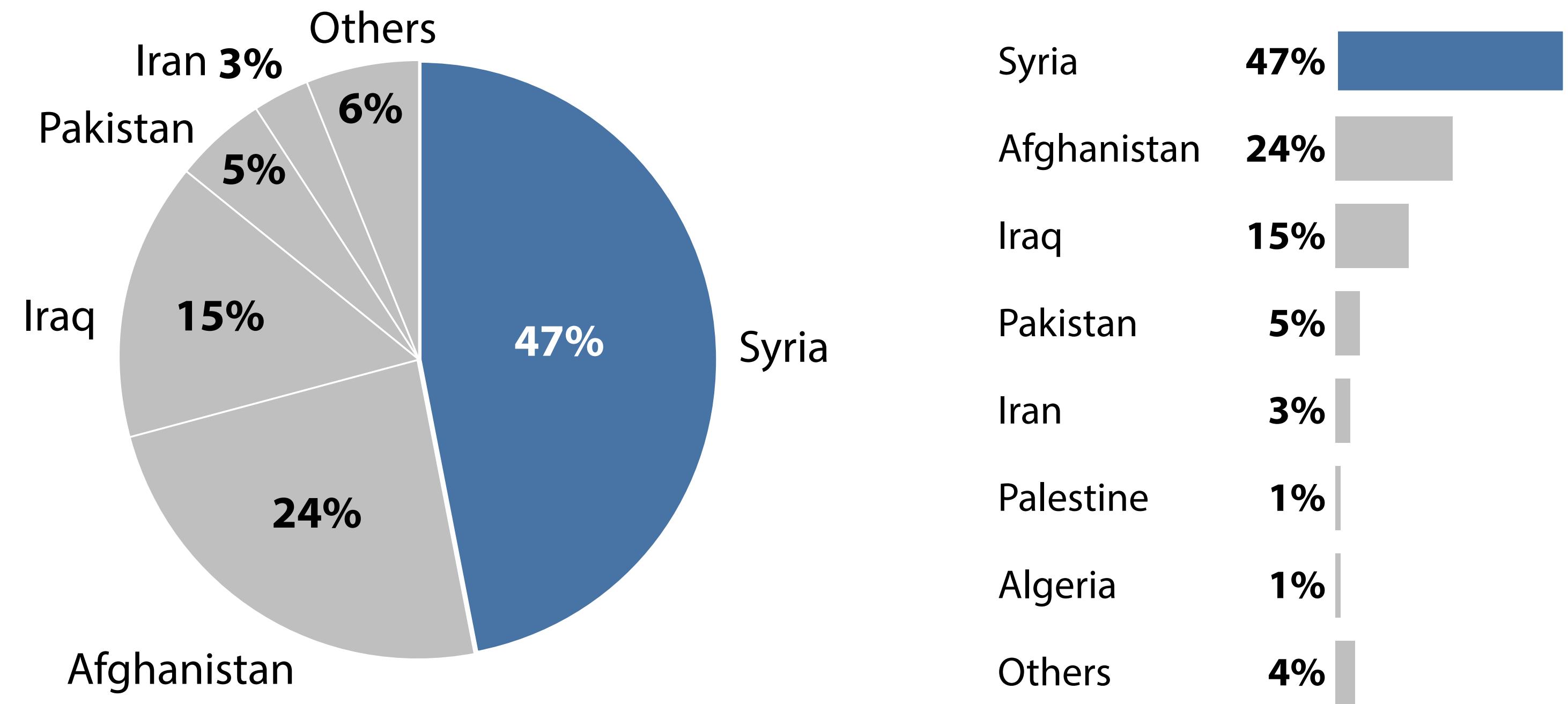


GREECE - Main nationalities of arriving migrants in 2016



GREECE - Main nationalities of arriving migrants in 2016

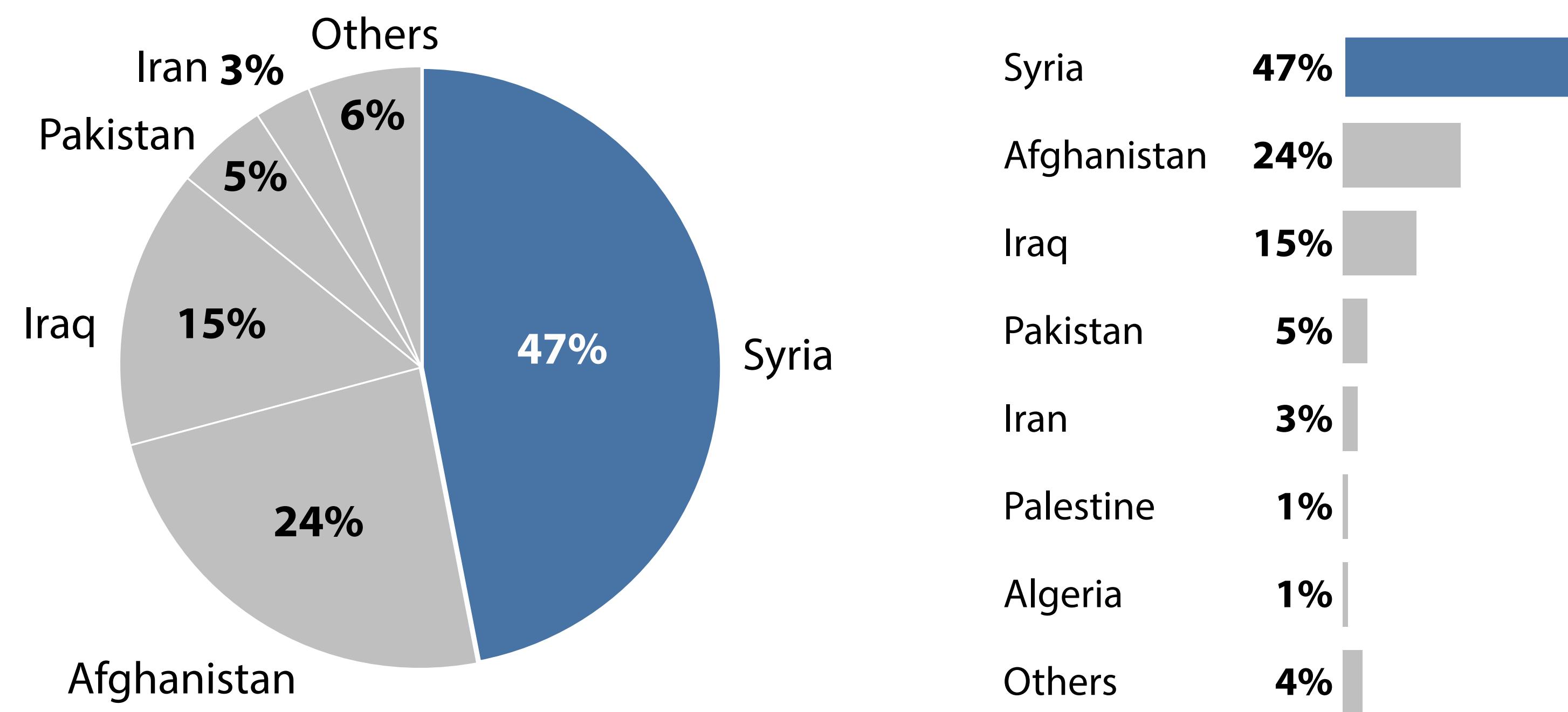
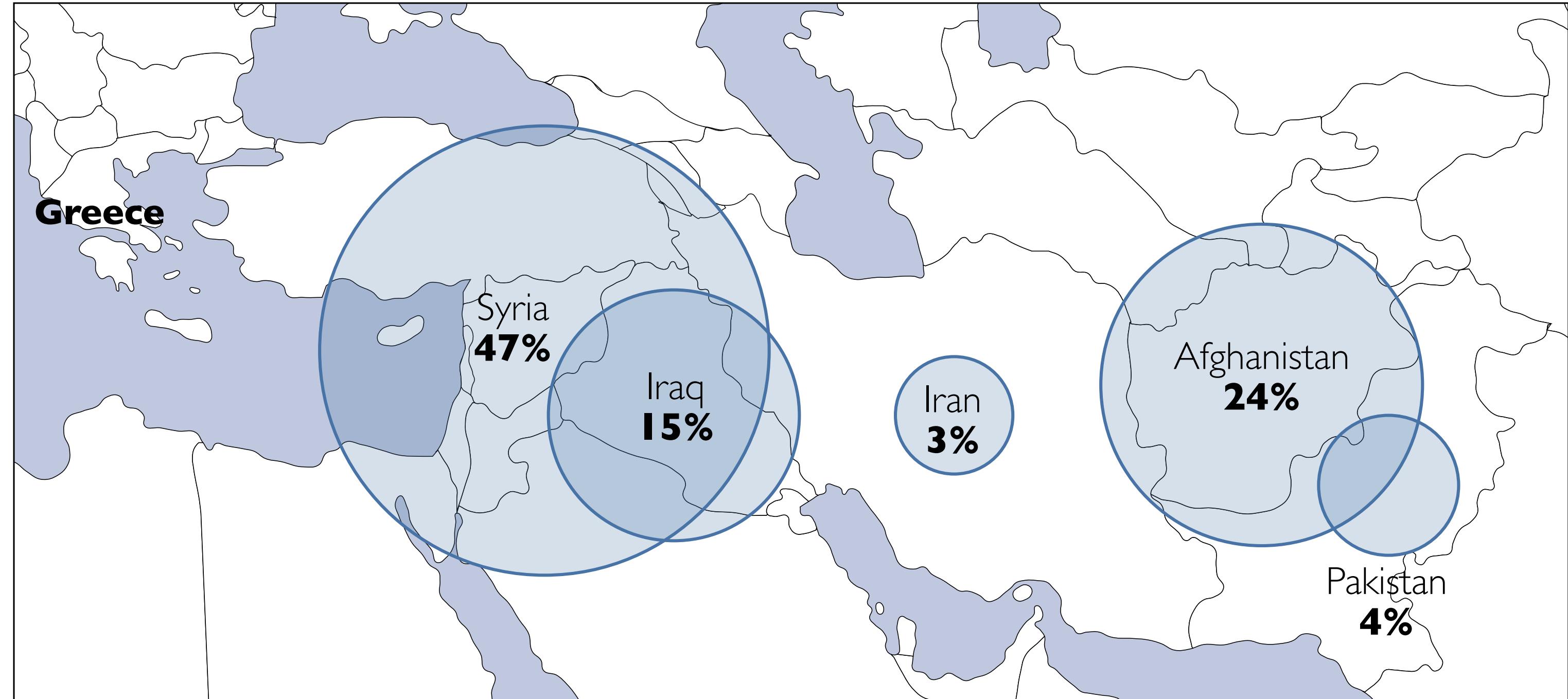




None of these charts is *right* or wrong, good or bad per se.

They are just **better** or **worse** depending on our **intent**:

What it is that we want to communicate, what we want our reader to **be able to see in the data**.



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; Lir Faunce; Graham Parish; Billy Ehrenberg; Paul McCalum; Martin Stabe
Inspired by the Graphic Continuum by Jon Schwabish and Severino Ribeiro

ft.com/vocabulary

FT

Deviation	Correlation	Ranking	Distribution	Change over Time	Part-to-whole	Magnitude	Spatial	Flow
Emphasise variations (+/-) from a fixed reference point. Typically zero but it can also be a target or a long-term average. Can focus on short-term sentiment (positive/negative).	Show the relationship between two or more variables. Be mindful that, unless you tell them otherwise, many readers will assume the relationships you show are causal (i.e. one causes the other).	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.	Show values in a dataset and how often they occur. The shape (or 'view') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.	Emphasis on changing trends. These can be short (intra-day) movements or extended series traversing decades or centuries. Choose the time period is important to provide suitable context for the reader.	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.	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 single entity (e.g. example, hard cash dollars or people) rather than a calculated rate or per cent.	Used only when precise locations or geographical patterns in data are more important to the reader than anything else.	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 Trade surplus/deficit, climate change	Example FT uses Inflation + unemployment, income & life expectancy	Example FT uses Wealth, deprivation, league tables, constituency election results	Example FT uses Income distribution, population, age/sex distribution	Example FT uses Share price movements, economic time series	Example FT uses Fiscal budgets, company structures, national election results	Example FT uses Commodity production, market capitalisation	Example FT uses Locator maps, population density, natural resource locations, natural disaster risk/impact, catchment areas, variation in election results	Example FT uses Movement of funds, trade, migrants, lawsuits, information: relationship graphs.
Diverging bar A simple standard bar chart that can handle both negative and positive magnitude values.	Scatterplot The standard way to show the relationship between two continuous variables, each of which has its own axis.	Ordered bar Standard bar charts display the ranks of values much more easily when sorted into order.	Histogram The standard way to show a statistical distribution - keep the gaps between columns small to highlight the shape of the data.	Line The standard way to show a changing time series. If data series are irregular, consider markers to represent data points.	Stacked column A simple way of showing part-to-whole relationships but can be difficult to read with more than a few components.	Column The standard way to compare the size of things. Must always start at 0 on the axis.	Basic choropleth (rate/ratio) The standard approach for putting data on a map - should always be rates rather than totals and in a sensible geography.	Sankey Shows changes in flows from one condition to at least one other; good for tracing the eventual outcome of a complex process.
Diverging stacked bar Perfect for presenting survey results which involve agreement (eg disagree/neutral/agree).	Line + Column A good way of showing the relationship between an amount (columns) and a rate (line).	Ordered column See above.	Boxplot Summarise multiple distributions by showing the median (centre) and range of the data.	Column + line Columns work well for showing change over time - but usually best with only one series of data at a time.	Proportional stacked bar A good way of showing the size and proportion of data at the same time - as long as the data are not too complicated.	Bar See above. Good when the data are not time series and labels have long category names.	Waterfall Designed to show the sequencing of data through a flow process, typically budgets. Can include +/- components.	
Spine chart Splits a single value into 2 contrasting components (eg Male/Female).	Connected scatterplot Usually used to show how the relationship between 2 variables has changed over time.	Ordered proportional symbol Use when there are big variations between values and/or seeing fine differences between data is not so important.	Violin plot Similar to a box plot but more effective with complex distributions (data that cannot be summarised with simple averages).	Line + column A good way of showing the relationship over time between an amount (columns) and a rate (line).	Pie A common way of showing part-to-whole values - but be aware that it's difficult to accurately compare the size of the segments.	Paired column As per standard column but allows for multiple series. Can become tricky to read with more than 2 series.	Flow map For showing unambiguous movement across a map.	
Surplus/deficit filled line The shaded area of these charts allows a balance to be shown relative to a baseline or between two series.	Bubble Like a scatterplot, but adds additional detail by sizing the circles according to a third variable.	Dot strip plot Data placed in order on a strip are a space-efficient method of laying out ranks across multiple categories.	Population pyramid A standard way for showing the age and sex breakdown of a population distribution; effectively back-to-back histograms.	Stack price Usually focused on day-to-day activity, these charts show opening/closing and high points of each day.	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).	Paired bar See above.	Contour map For showing areas of equal value on a map. Can use deviation colour schemes for showing +/- values.	
XY heatmap A good way of showing the patterns between 2 categories of data, less good at showing the differences in amounts.	Slope Perfect for showing how ranks have changed over time or very broken categories.	Dot strip plot Good for showing individual values in a distribution that be a proportion of the total when too many dots have the same value.	Dot plot A simple way of showing the change or range (min/max) of data across multiple categories.	Slope Good for showing changing data as long as the data can be simplified to 2 or 3 points without missing a key part of the story.	Treemap Use for hierarchical part-to-whole relationships; can be useful when there are many small segments.	Proportional stacked bar A good way of showing the size and proportion of data at the same time - as long as the data are not too complicated.	Equalised cartogram Converting each unit on a map to a regular and equally-sized shape - good for representing voting regions with equal value.	
Lollipop chart Lollipops draw more attention to the data value than standard bar/column and can also show rank and value effectively.	Barcode plot Like dot strip plots, good for displaying all the data in a table; they work best when highlighting individual values.	Cumulative curve A good way of showing changing data for two variables where there is a relatively clear pattern of progression.	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 where there is a relatively clear pattern of progression.	Sunburst Another way of visualising hierarchical part-to-whole relationships. Use sparingly (if at all) for obvious reasons.	Proportional symbol Use when there are big variations between values and/or seeing fine differences between data is not so important.	Scaled cartogram (value) Stretching and shrinking a map so that each area is sized according to a particular value.	
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).	Gridplot Good for showing % information; they work best when on whole numbers and work well in multiple layout form.	Venn Generally only used for schematic representation.	Radar chart A space-efficient way of showing values of multiple variables - but make sure they are organised in a way that makes sense to reader.	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 administrative unit.	
Seismogram Another alternative to the circle timeline for showing series where there are big variations in the data.	Waterfall Can be useful for showing part-to-whole relationships where some of the components are negative.			Parallel coordinates An alternative to radar charts - again, the arrangement of the variables is important. Usually benefits from highlighting values.				

Search by Function

View by List



Arc Diagram



Area Graph



Bar Chart



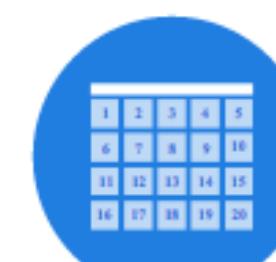
Box & Whisker Plot



Brainstorm



Bubble Chart



Calendar



Chord Diagram



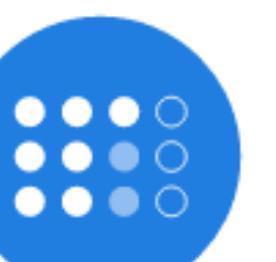
Choropleth Map



Circle Packing



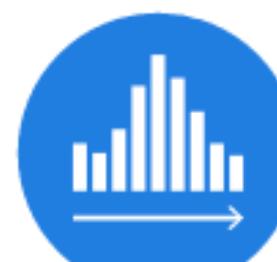
Donut Chart



Dot Matrix Chart



Flow Map



Histogram



Illustration Diagram



Line Graph



Marimekko Chart



Multi-set Bar Chart



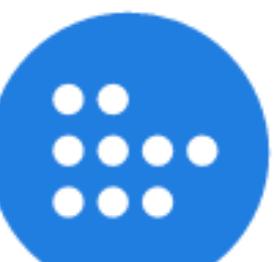
Nightingale Rose Chart



Non-ribbon Chord Diagram



Parallel Sets



Pictogram Chart



Pie Chart



Population Pyramid



Proportional Area Chart



Radar Chart



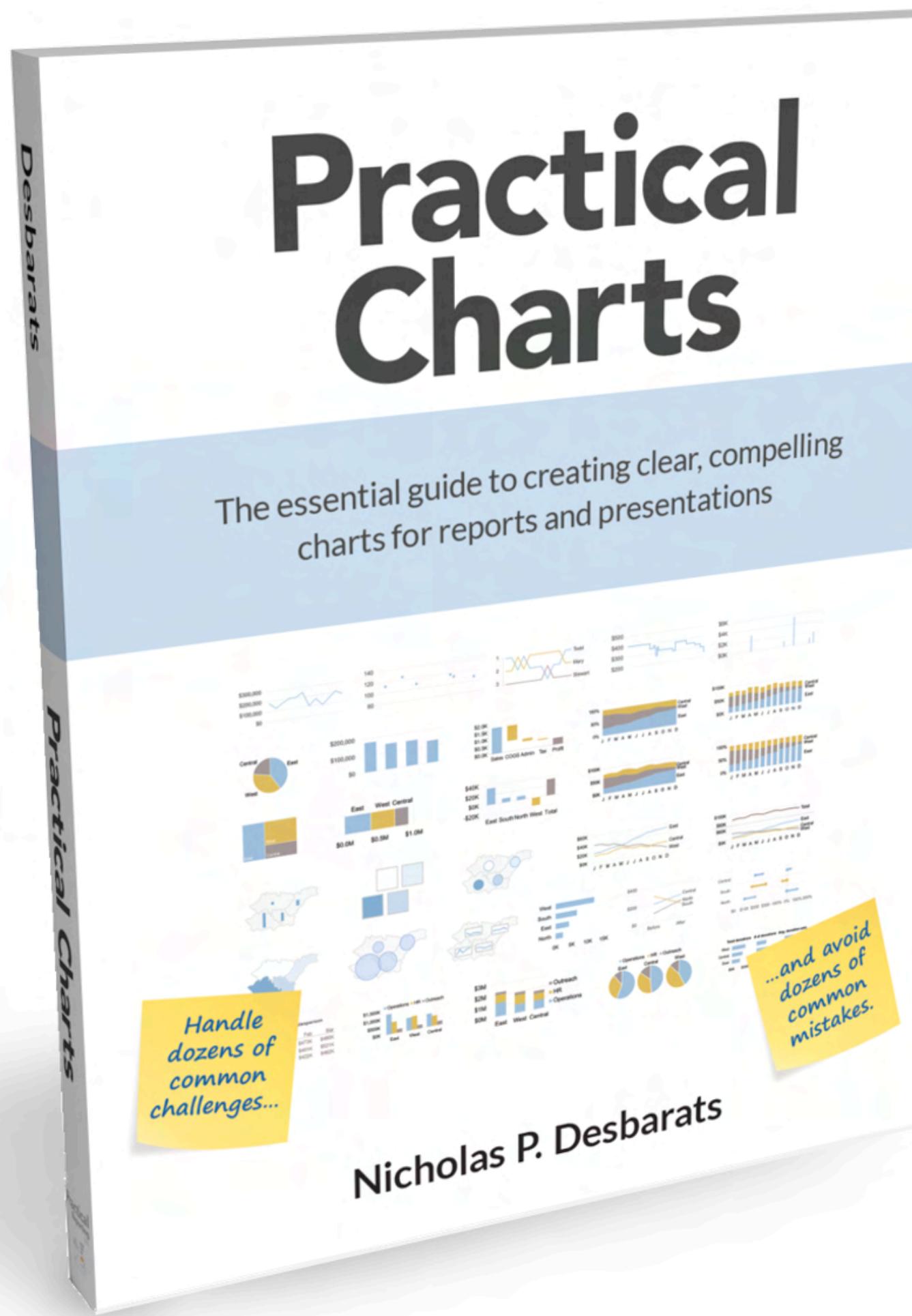
Radial Bar Chart



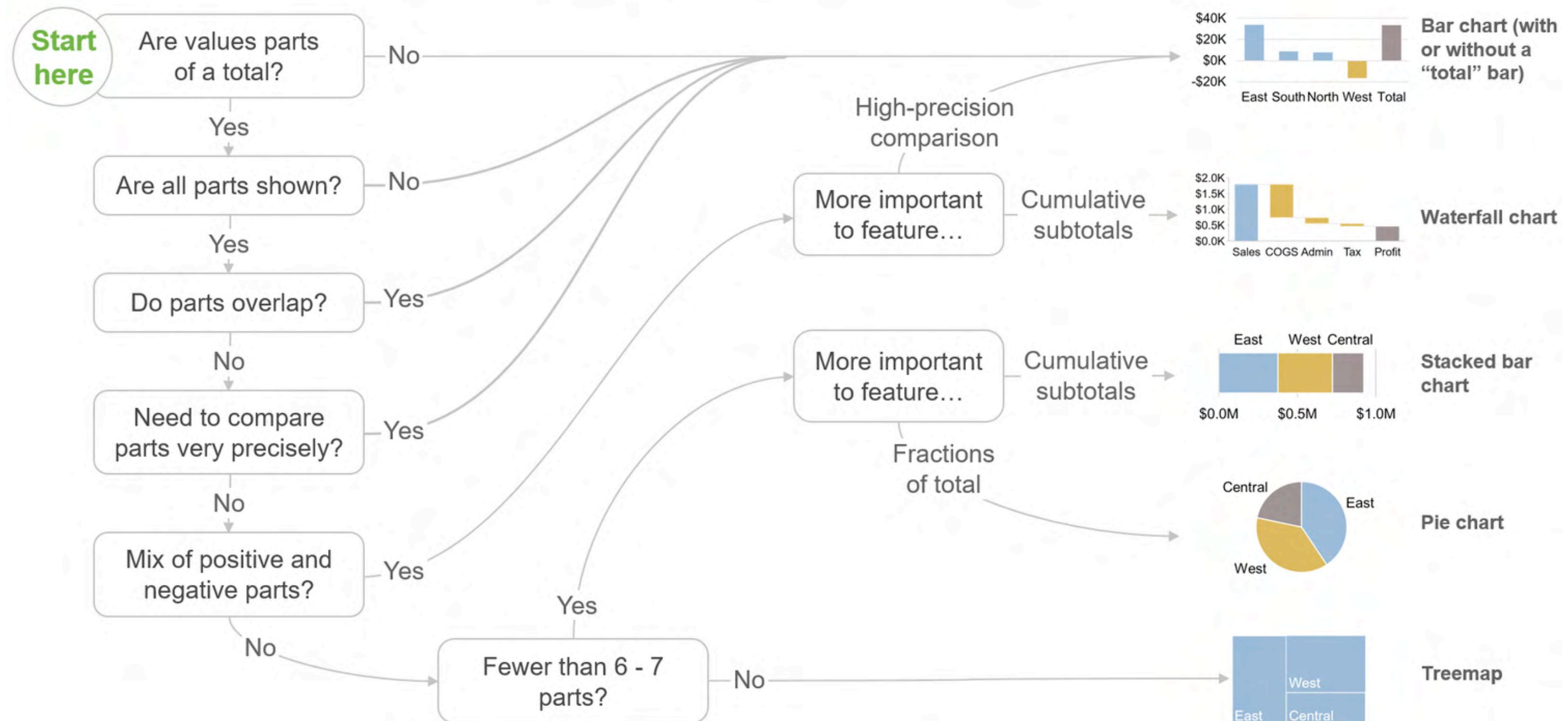
Sankey Diagram



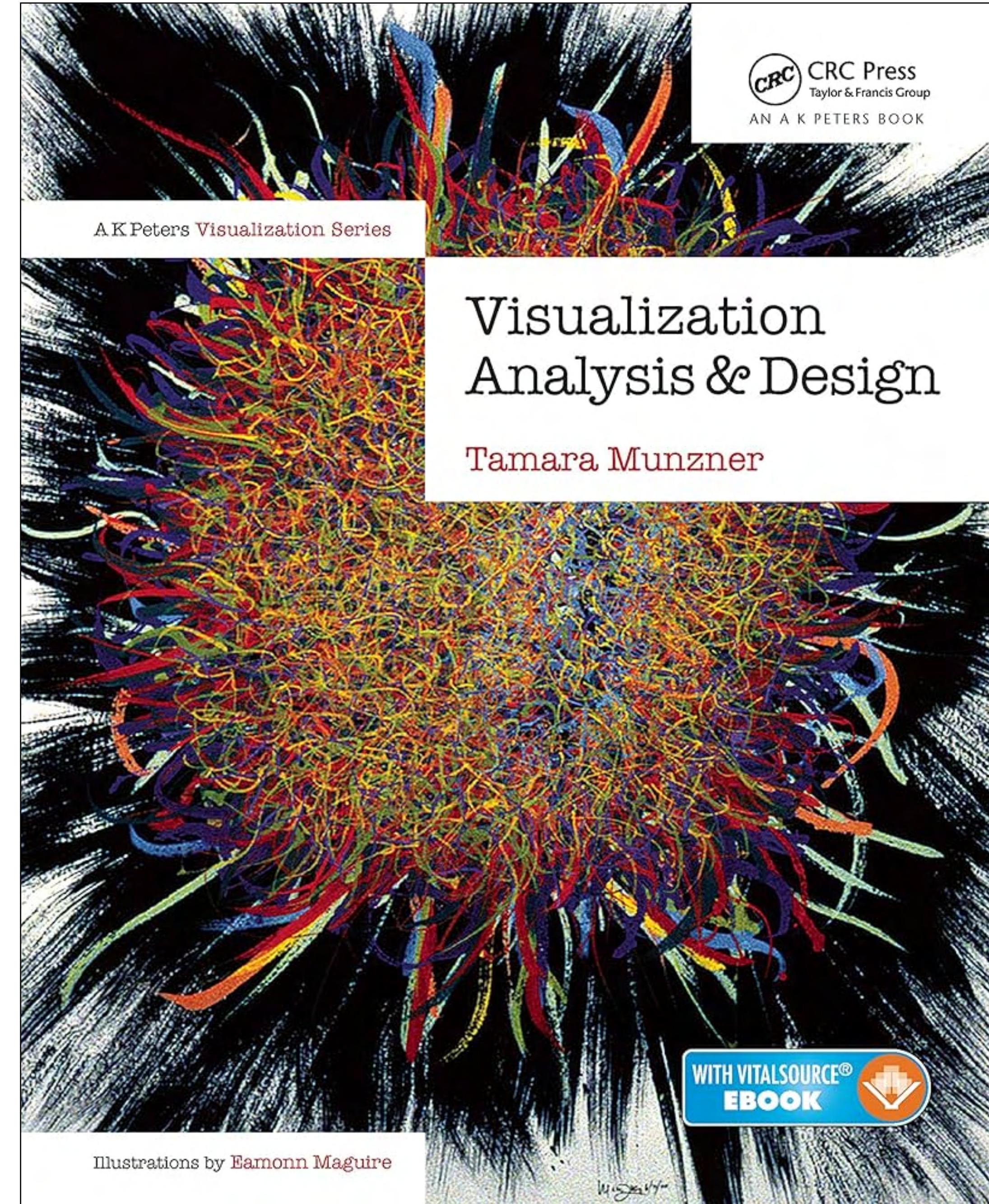
Scatterplot



Choosing a chart type to show the breakdown of a total



<https://www.practicalreporting.com/practical-charts-book>



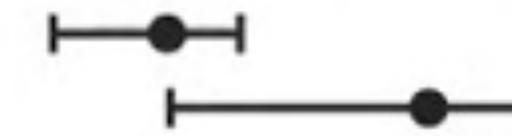
Encodings/channels hierarchy from Tamara Munzner's *Visualization Analysis and Design* (2014)

→ **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



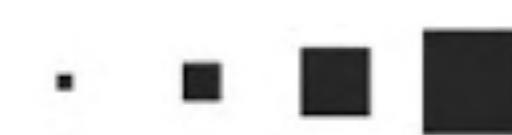
Length (1D size)



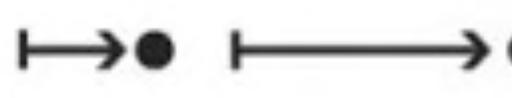
Tilt/angle



Area (2D size)



Depth (3D position)



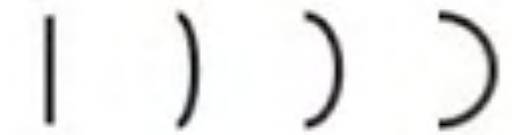
Color luminance



Color saturation



Curvature



Volume (3D size)



→ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion

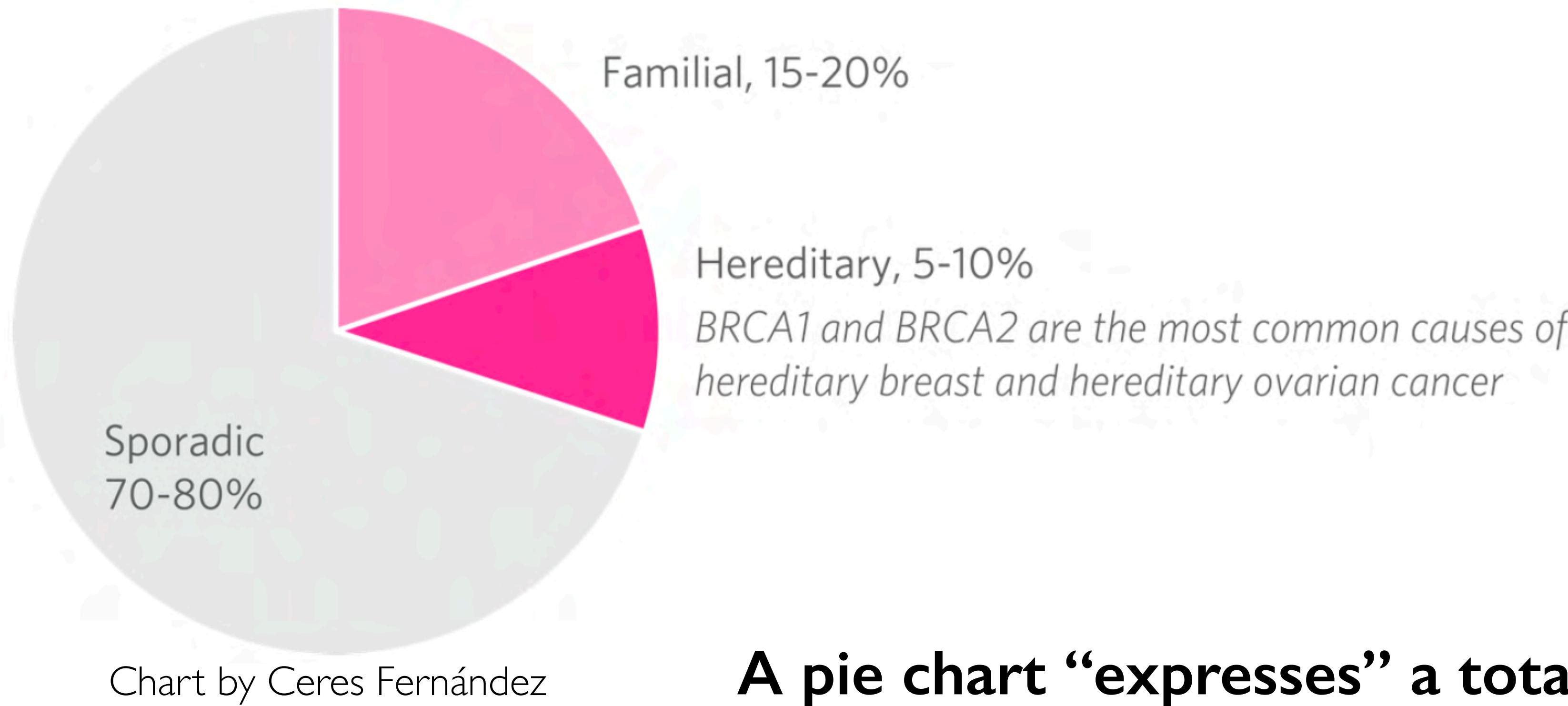


Shape

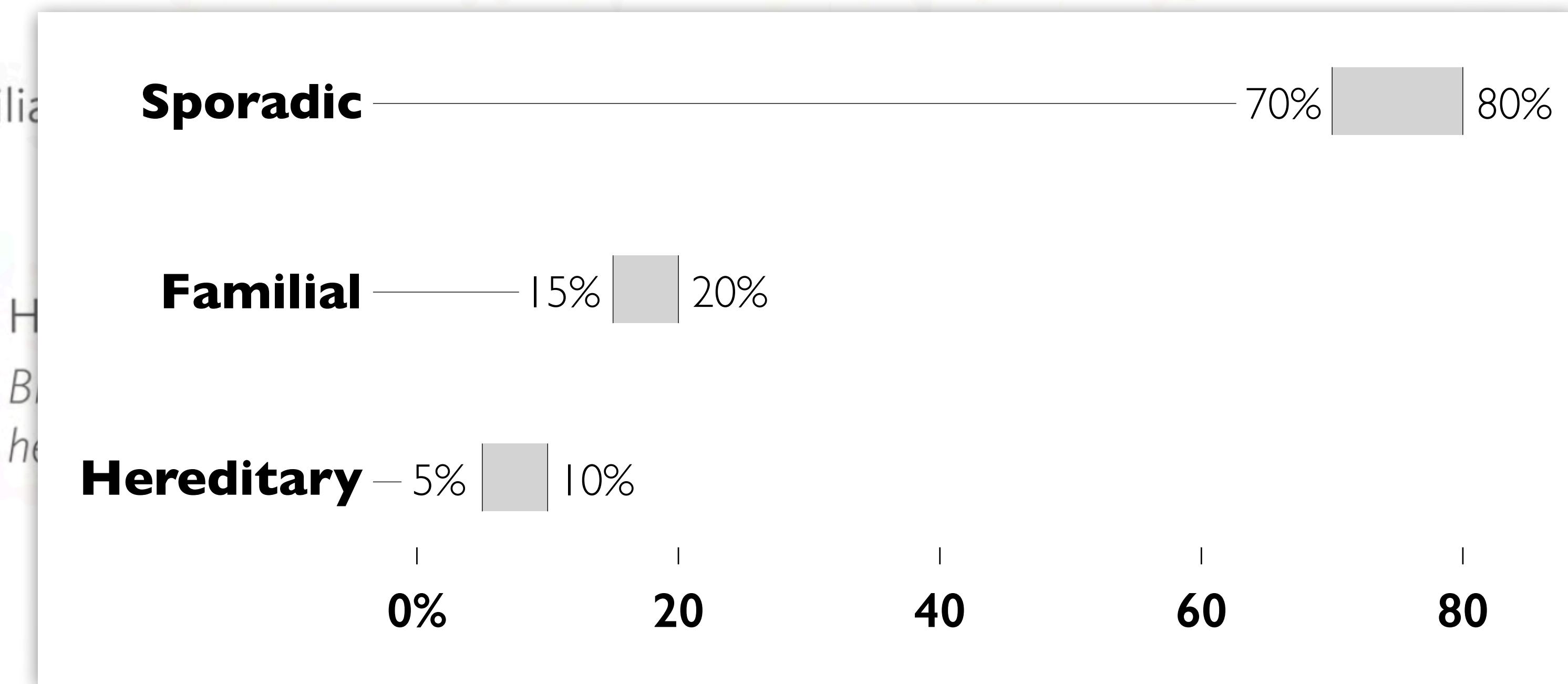
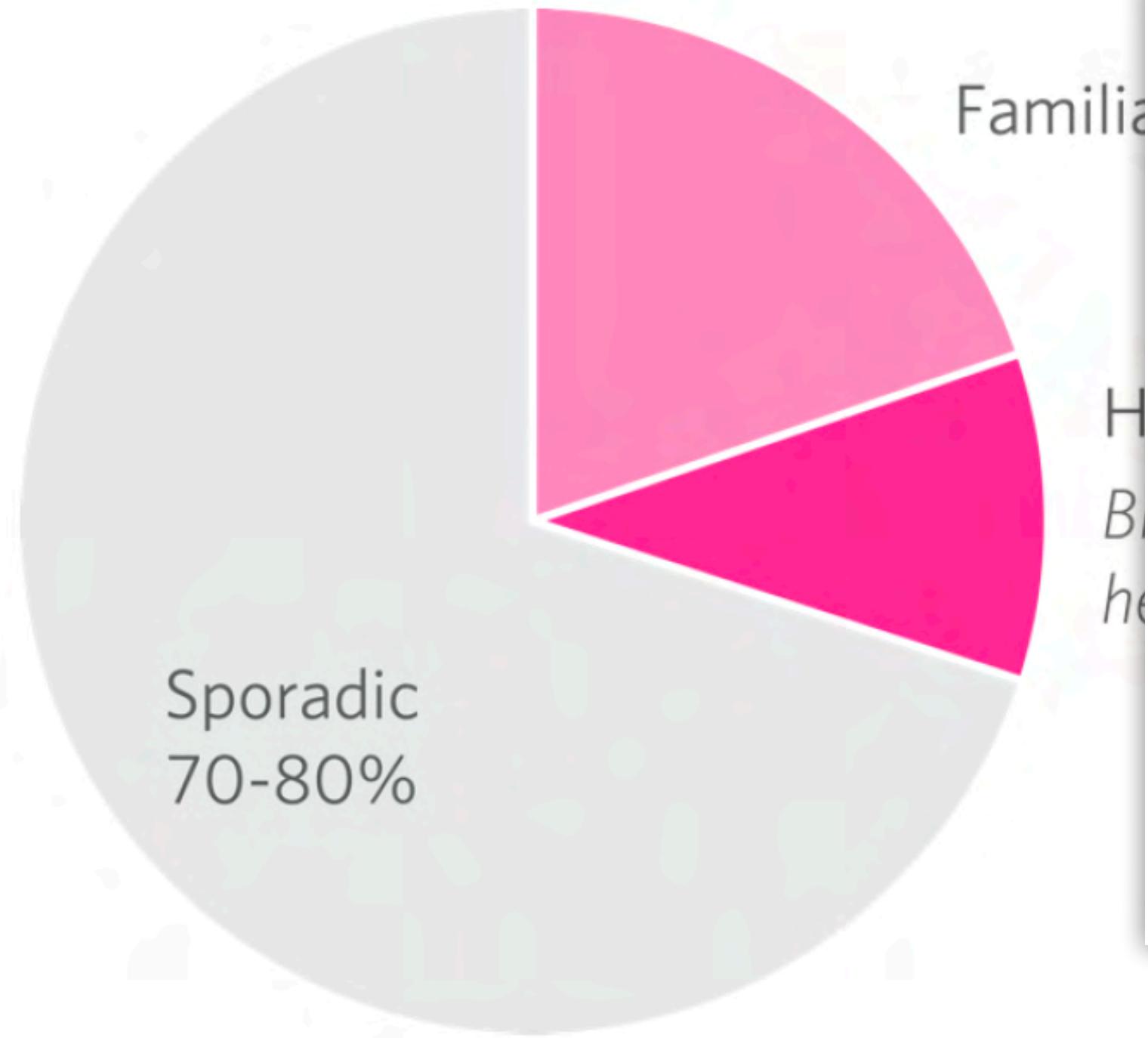


- **expressiveness** ←
 - match channel and data characteristics
- **effectiveness**
 - channels differ in accuracy of perception

- Hereditary cancer syndromes are the exception



- Hereditary cancer syndromes are the exception



This chart “expresses” ranges

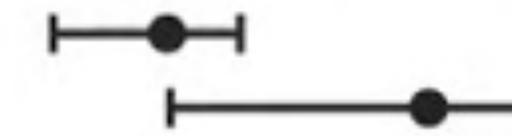
Encodings/channels hierarchy from Tamara Munzner's *Visualization Analysis and Design* (2014)

→ **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



Length (1D size)



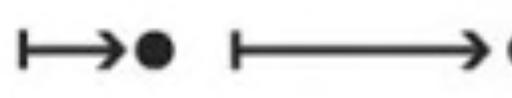
Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



→ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion



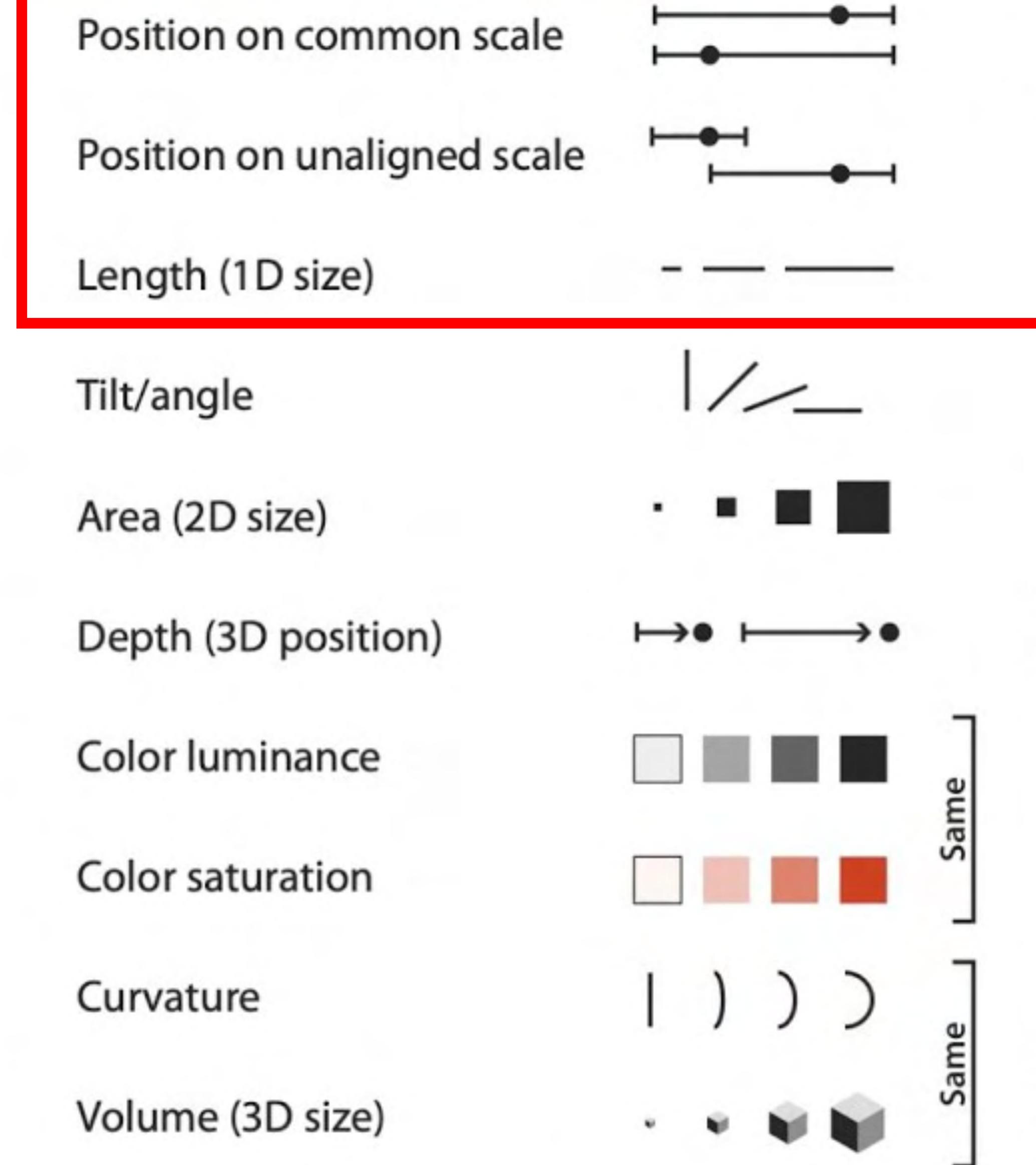
Shape



- expressiveness
 - match channel and data characteristics
- **effectiveness** ←
 - channels differ in accuracy of perception

Encodings/channels hierarchy from Tamara Munzner's *Visualization Analysis and Design* (2014)

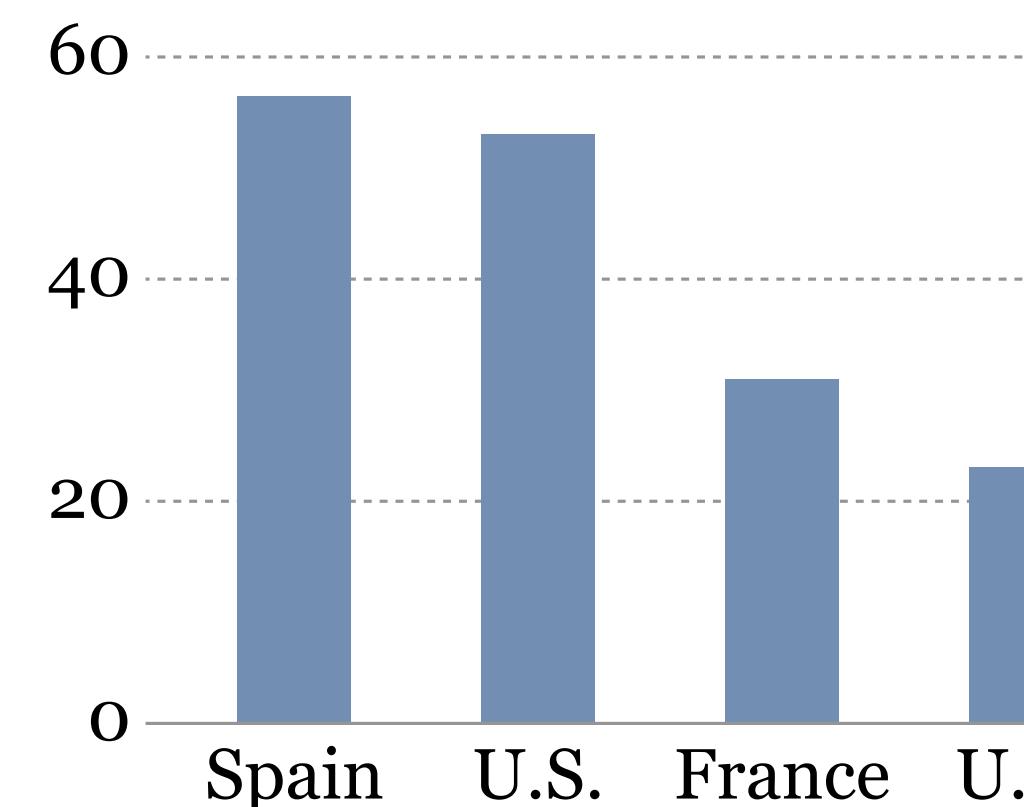
→ **Magnitude Channels: Ordered Attributes**



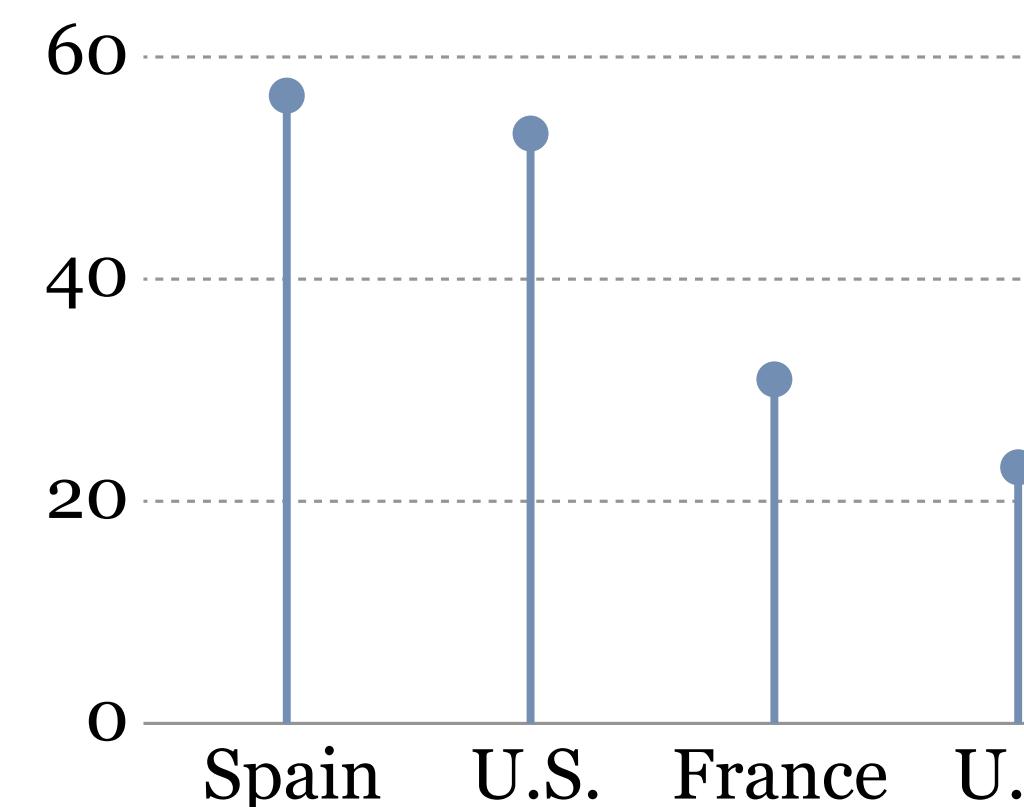
→ **Identity Channels: Categorical Attributes**



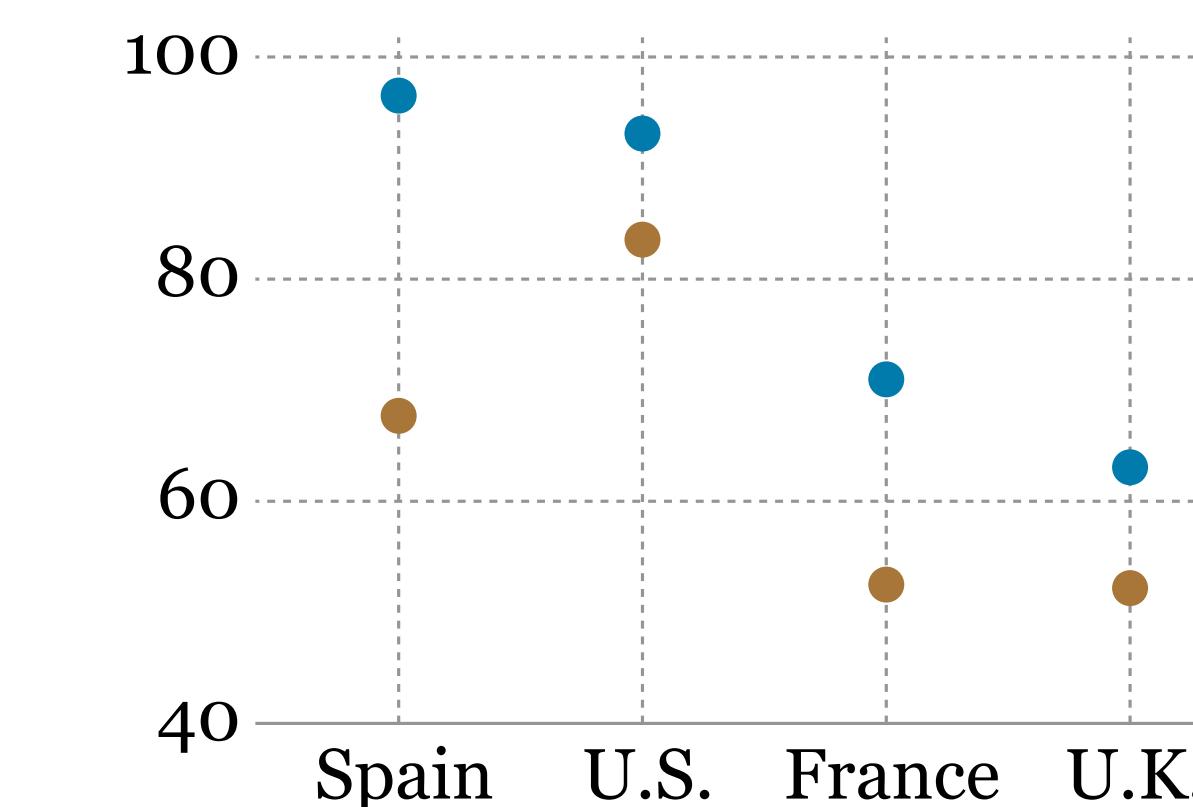
- expressiveness
 - match channel and data characteristics
- **effectiveness** ←
 - channels differ in accuracy of perception



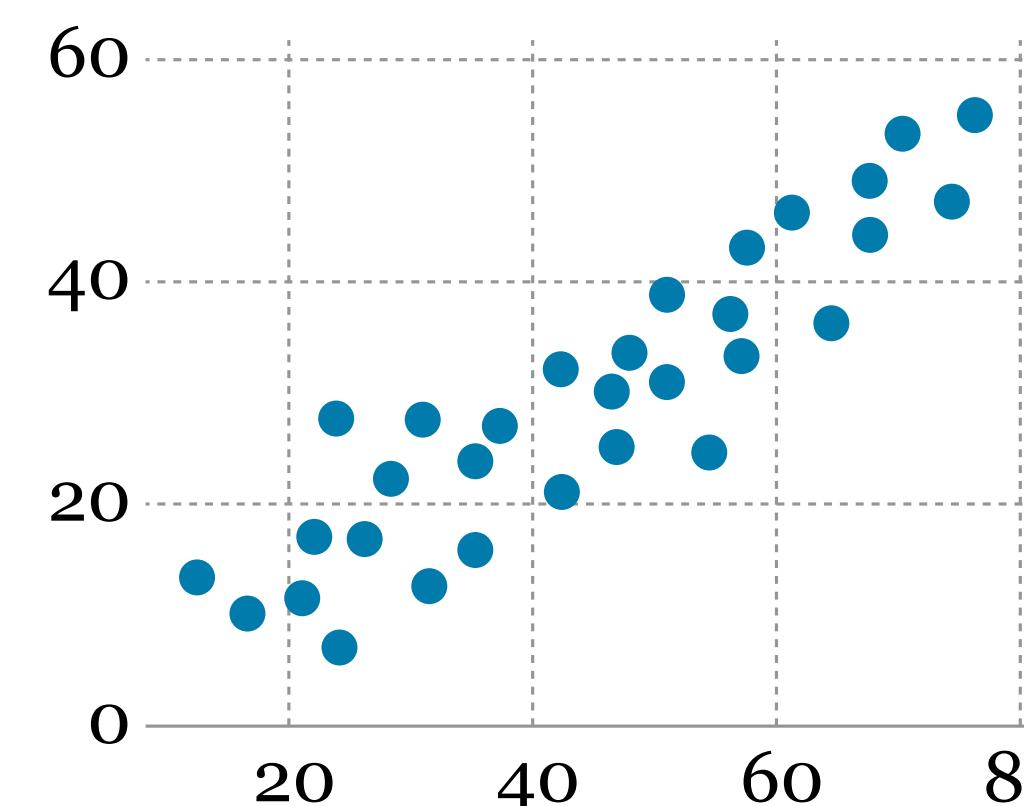
BAR CHART



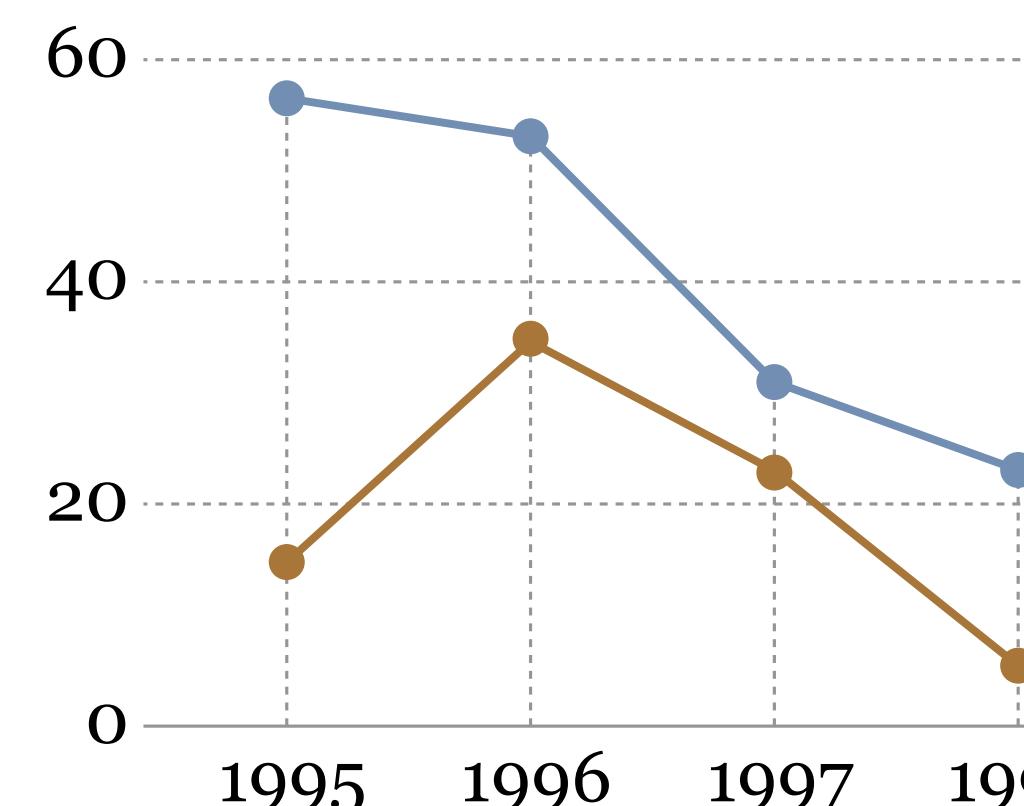
LOLLIPOP CHART



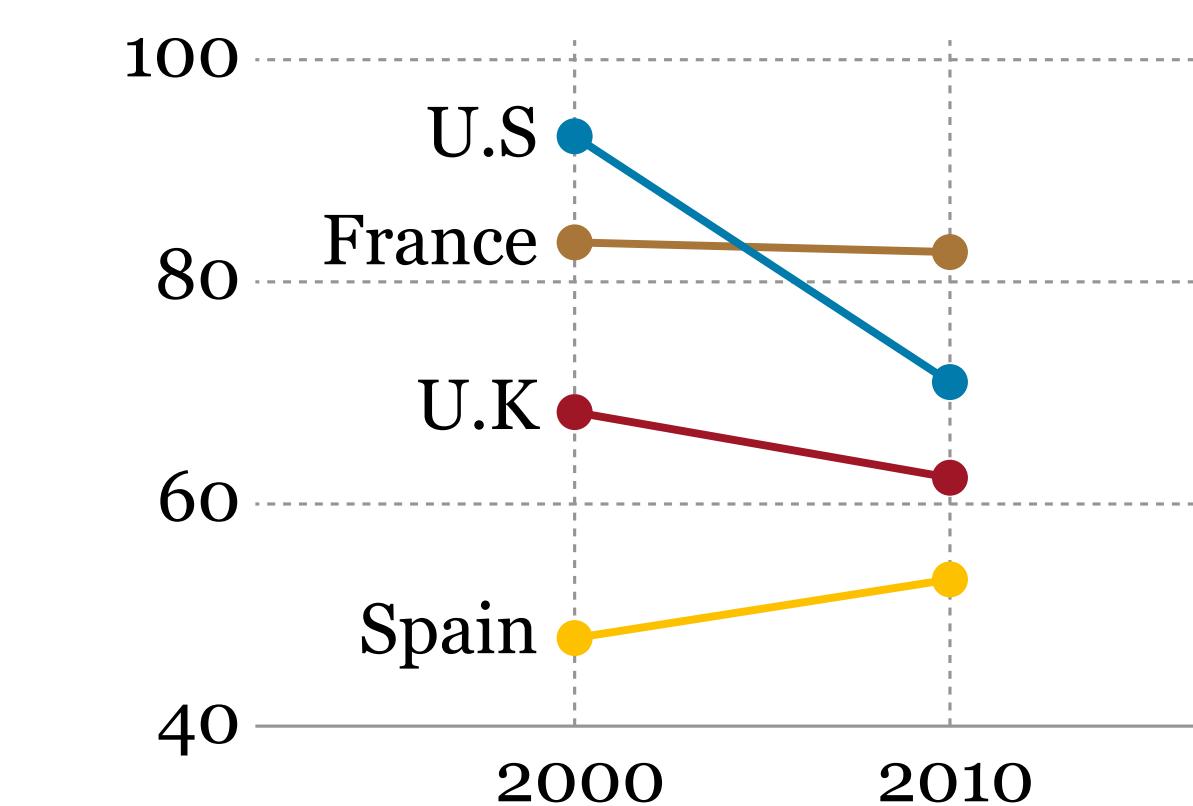
**DOT CHART
(or dot plot)**



**SCATTER CHART
(or scatter plot)**



**LINE CHART
(or time series chart)**



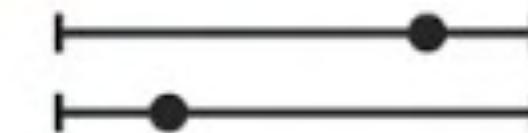
SLOPE CHART

Height, length, position are appropriate encodings
to help readers make accurate estimates

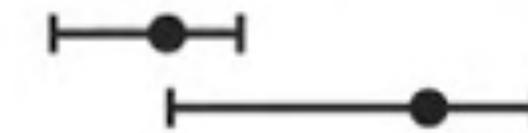
Encodings/channels hierarchy from Tamara Munzner's *Visualization Analysis and Design* (2014)

→ **Magnitude Channels: Ordered Attributes**

Position on common scale



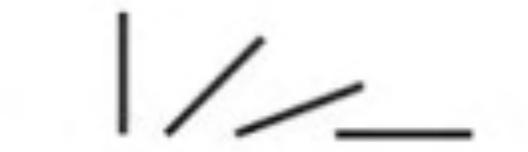
Position on unaligned scale



Length (1D size)



Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



→ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion



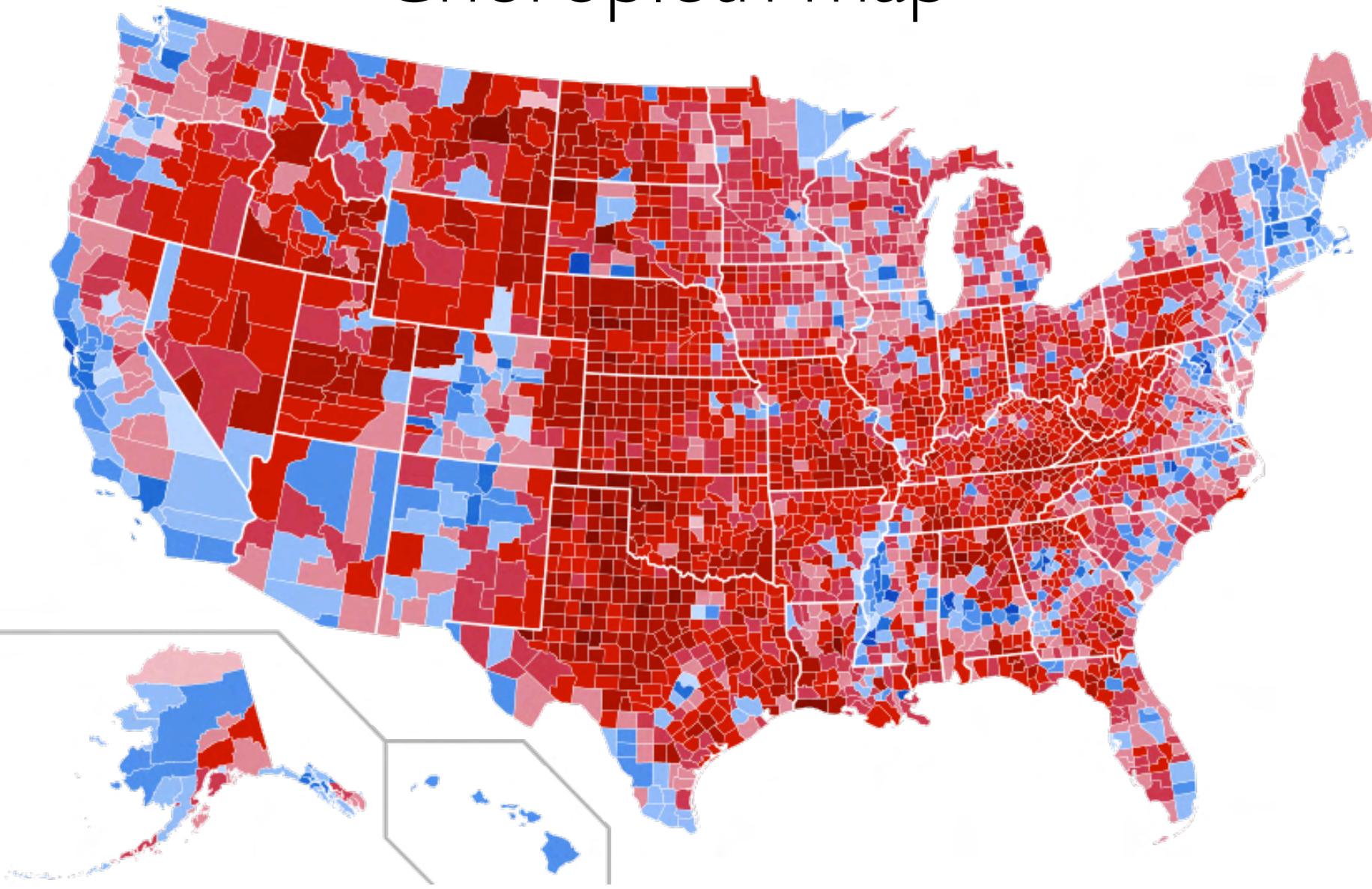
Shape



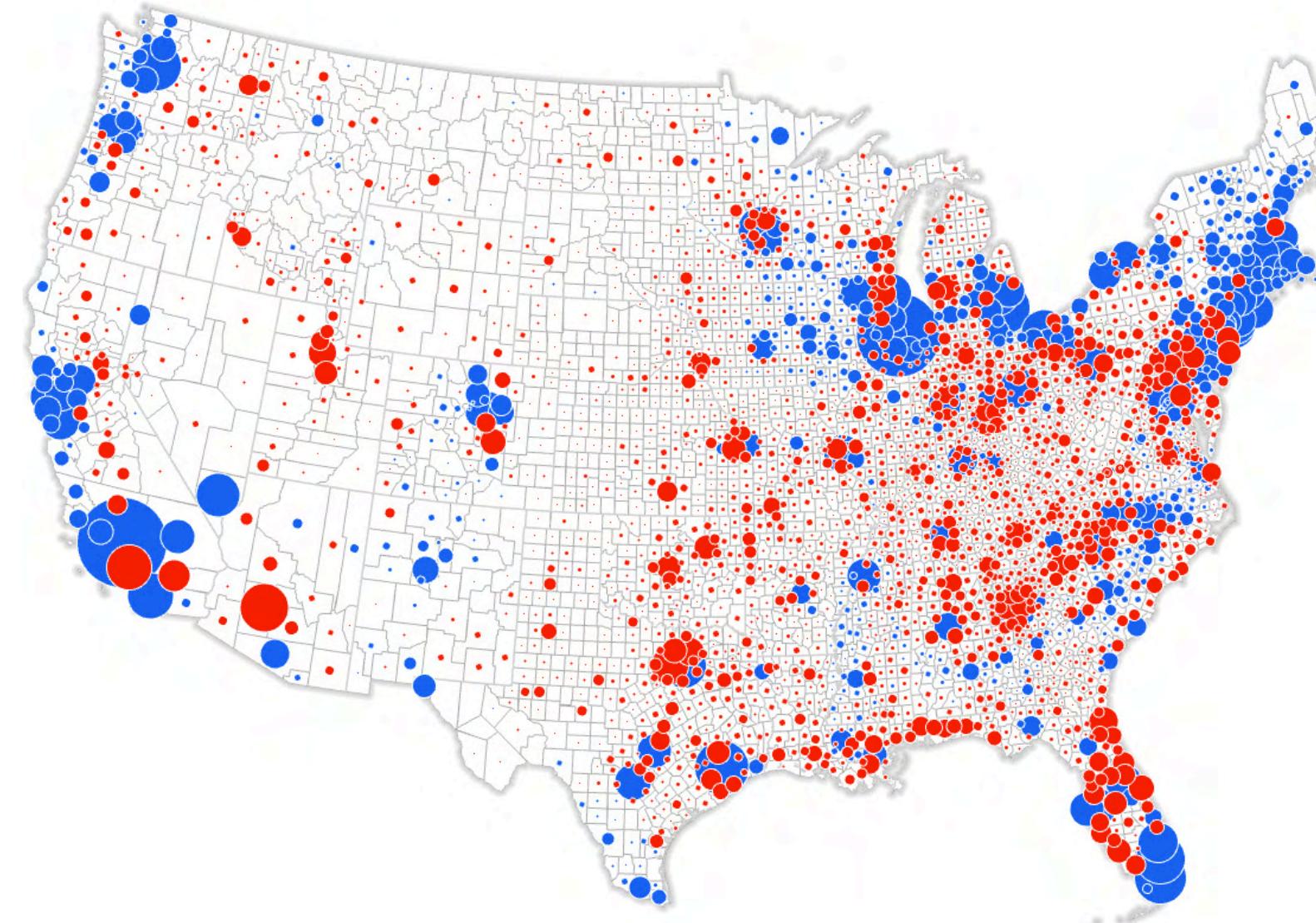
- **expressiveness**
 - match channel and data characteristics
- **effectiveness**
 - channels differ in accuracy of perception

A visual channel can be “effective” in the non-technical sense of the term because, in many cases, high accuracy of perception isn’t the goal.

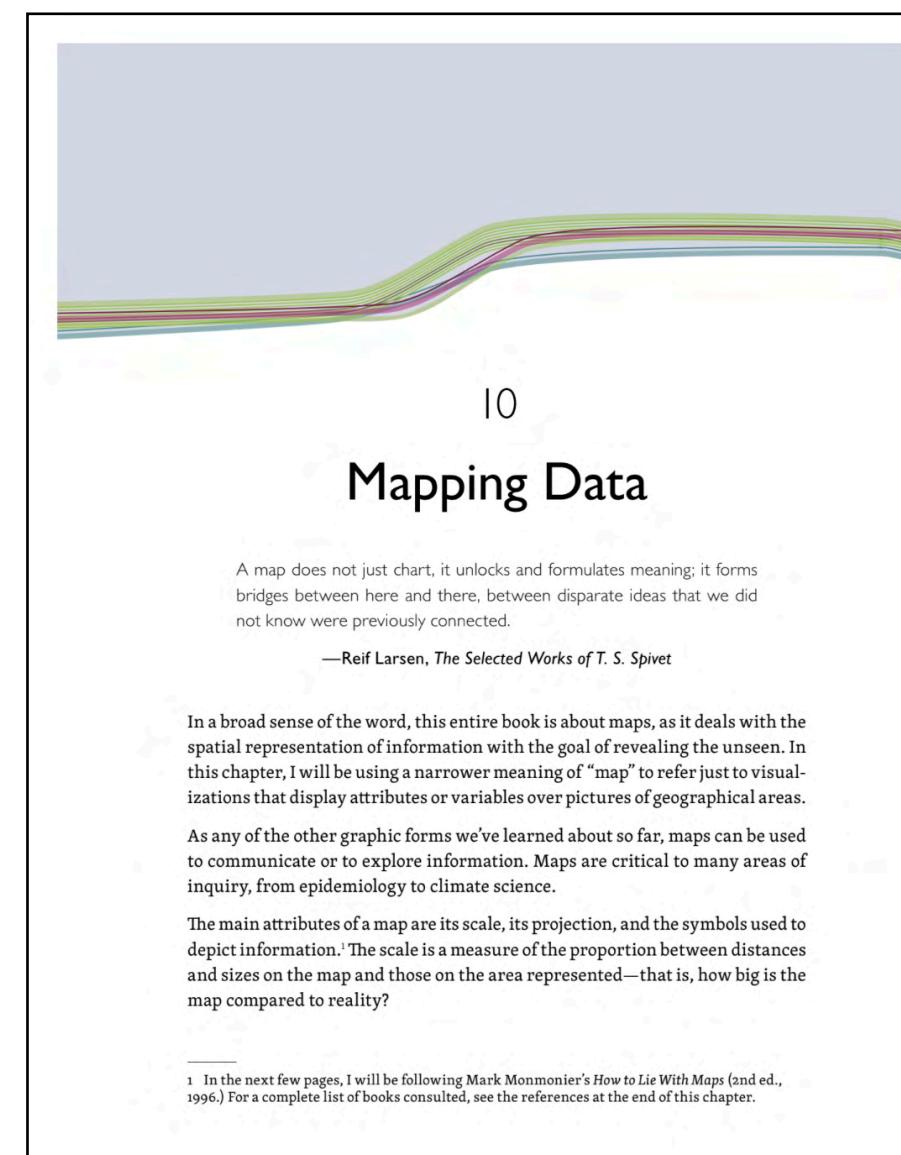
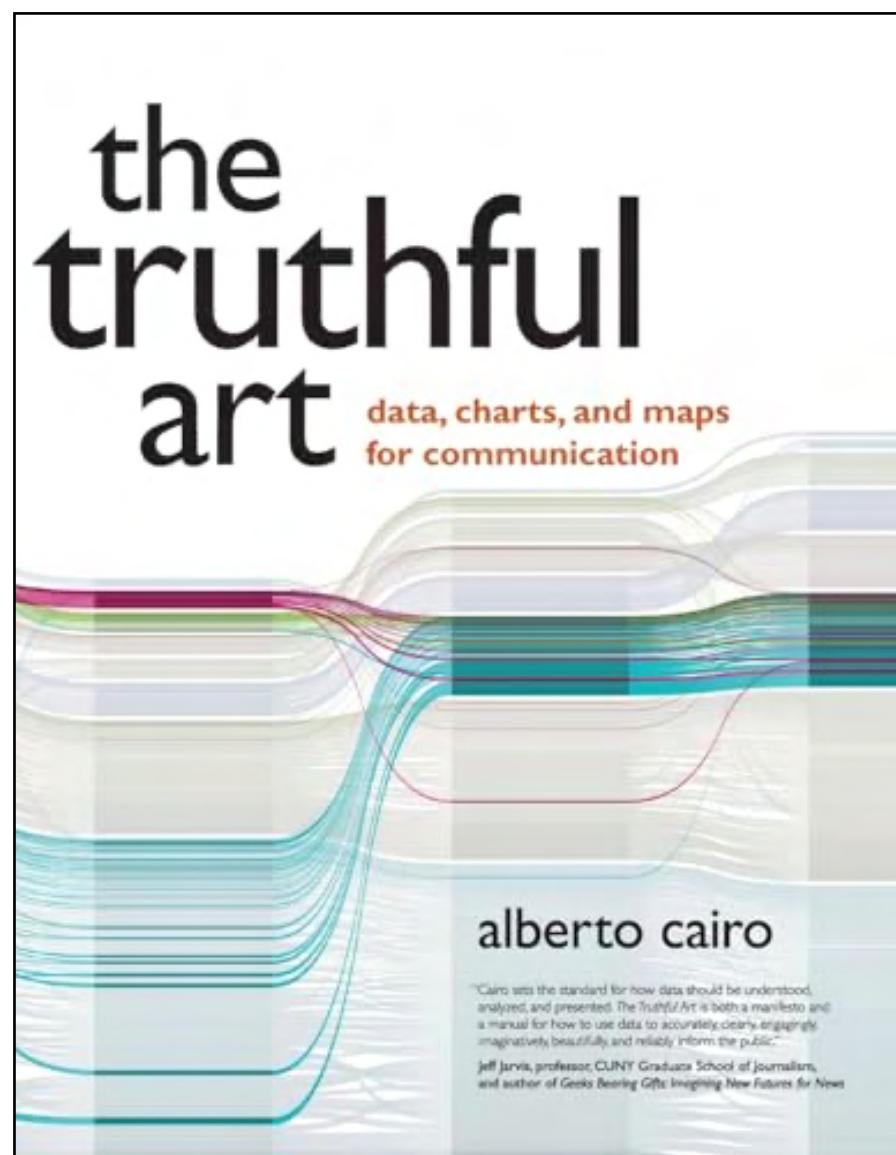
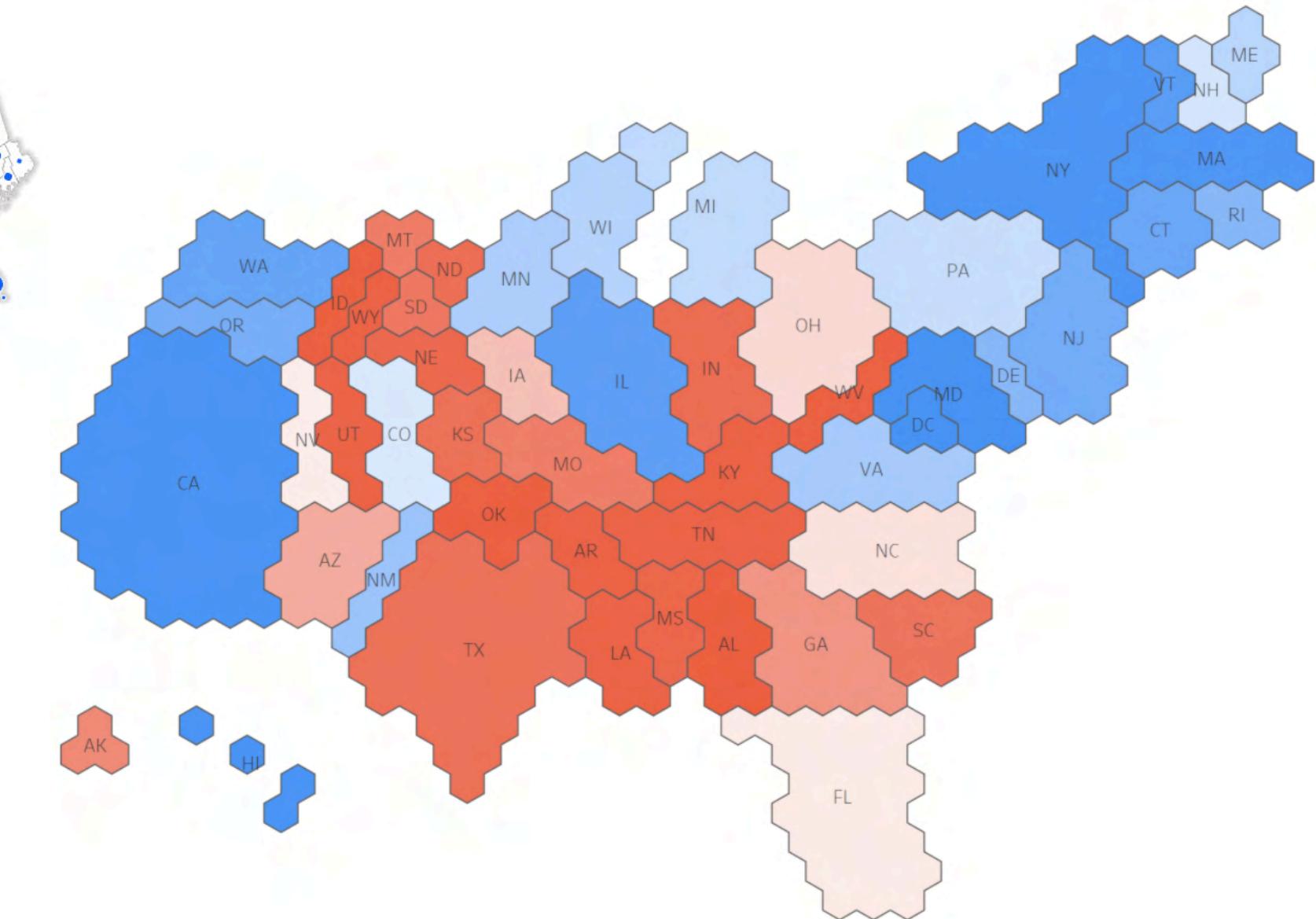
Choropleth map



Proportional symbol map



Cartogram



Chapter 10 of *The Truthful Art*

<https://drive.google.com/file/d/1Ap4ehrBDXBSDPcnPXCTeg9cVHIU9kOeX/>



European data

data.europa.eu The official portal for European data

| Home | Data | Academy | Open data maturity | Community | Publications |

Home > News & events > Events > Training series 'Visualising data with impact...

data.europa.eu event

Training series 'Visualising data with impact': foundations of effective data visualisation

14 October 2025

Online

Episode ONE - October 14

<https://data.europa.eu/en/news-events/events/training-data-visualisation-session-1-foundations-effective-data-visualisation>



European data

data.europa.eu The official portal for European data

| Home | Data | Academy | Open data maturity | Community | Publications |

Home > News & events > Events > Training series 'Visualising data with impact...

data.europa.eu event

Training series 'Visualising data with impact': designing with integrity

15 October 2025

Online

Episode TWO - October 15

<https://data.europa.eu/en/news-events/events/data-visualisation-training-session-2-designing-integrity>



European data

data.europa.eu The official portal for European data

| Home | Data | Academy | Open data maturity | Community | Publications |

Home > News & events > Events > Training series 'Visualising data with impact...

data.europa.eu event

Training series 'Visualising data with impact': data storytelling and creating your narrative

16 October 2025

Online

Episode THREE - October 16

Tomorrow's webinar

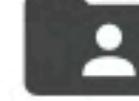
<https://data.europa.eu/en/news-events/events/data-visualisation-training-session-3-data-storytelling-and-creating-your>

 1_The_Basics

 2_History

 3_Chart_Taxonomies

 4_Perception_and_Accessibility

 5_Visual_Design

 6_Annotations_And_Storytelling

 7_Uncertainty

 8_Maps

Extra readings

<https://tinyurl.com/mr2st327>

Designing With Integrity

Thank you!

Alberto Cairo

OpenVisualizationAcademy.com