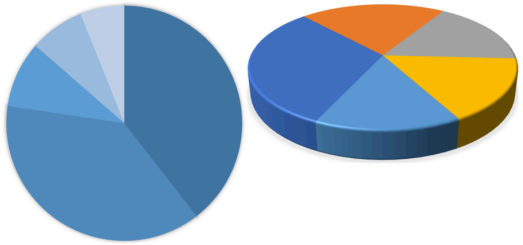
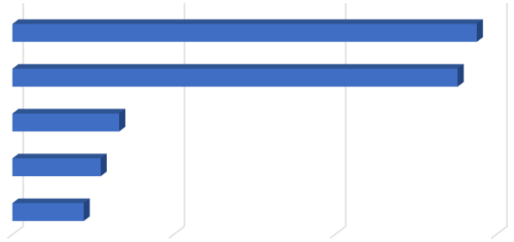
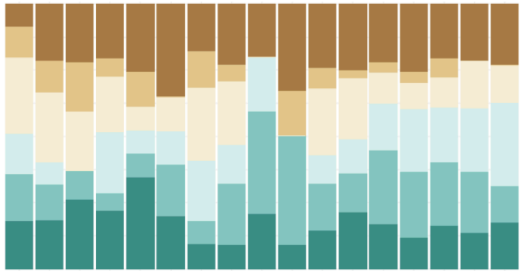






Creating more effective charts

Perception, reasoning, and credibility

2024 MIDFIELD Institute

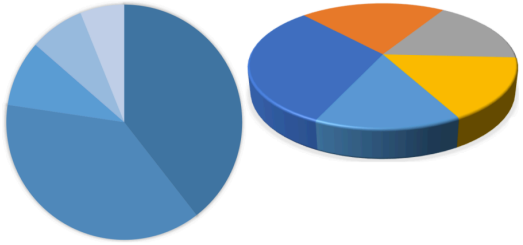
Richard Layton

https://www.graphdoctor.com

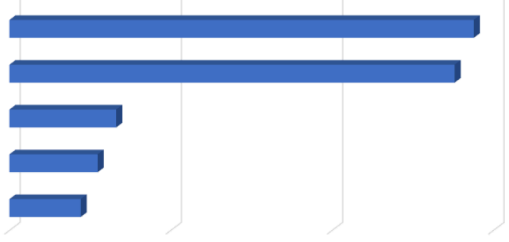
https://github.com/graphdr

graphdoctor@gmail.com

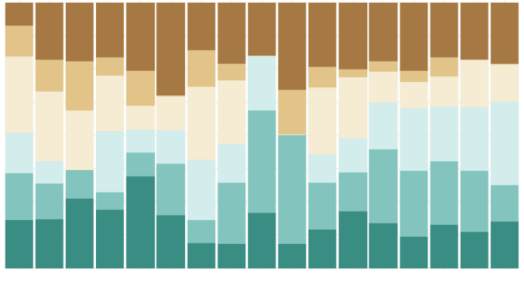
Title and exercises inspired by Naomi Robbins (2013) *Creating More Effective Graphs*, Chapter 2, Chart House.



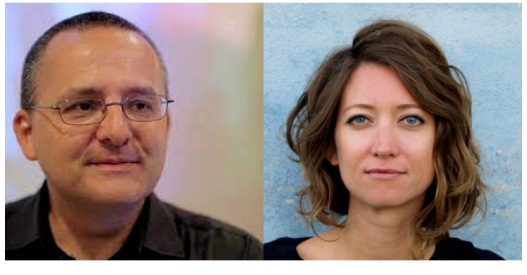
Effective alternatives to pie charts



Effective alternatives to bar charts



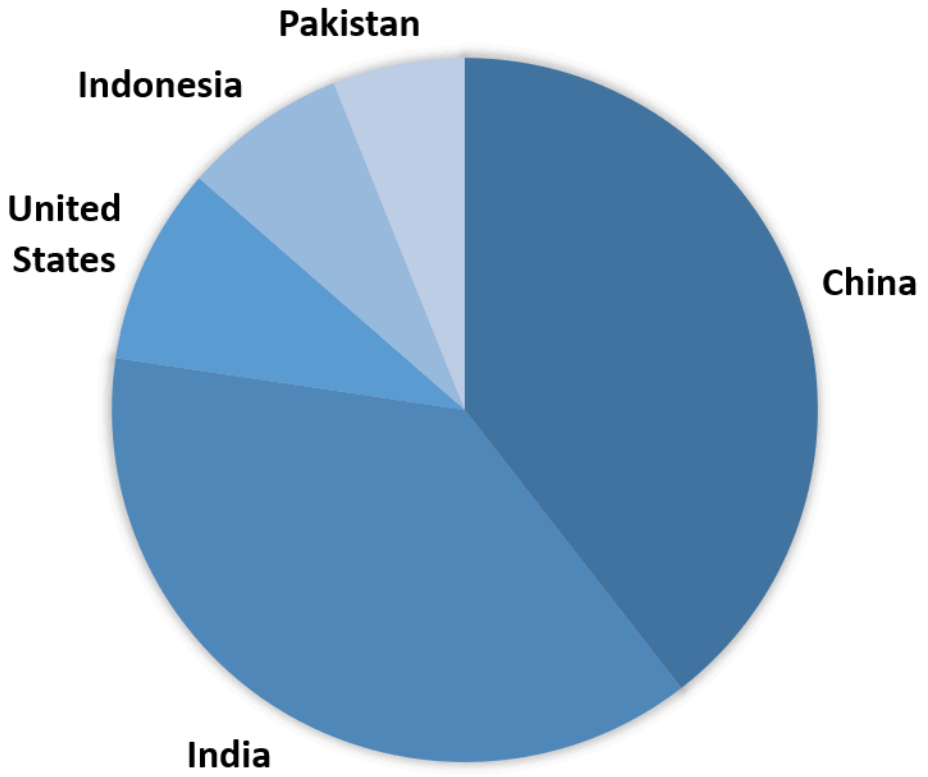
Aligning the design to the story



Advice from experts

Effective alternatives to pie charts

Judging pie slices is a low-accuracy task



- *Visually estimate* each country's percentage
- Fill-in the blanks in the table
- Total should be 100%

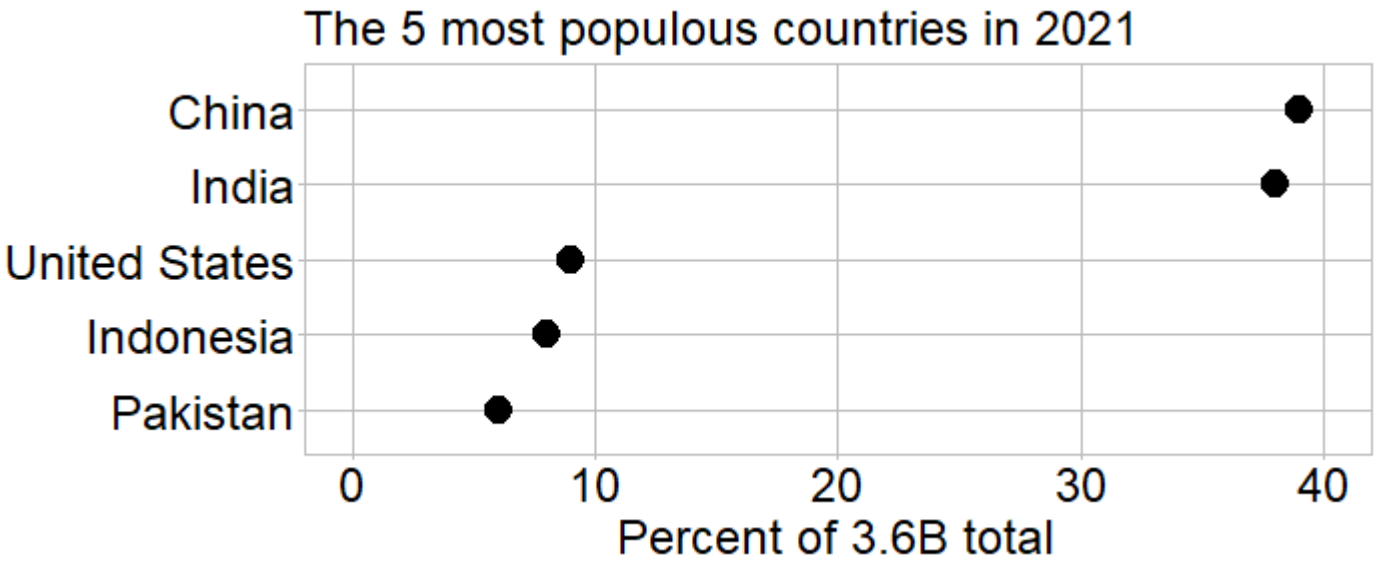
Country	Percentage
China	
India	
United States	
Indonesia	
Pakistan	

Data source: 2021 World population review (The pie represents 3.6B people total.)

Judging values along a common axis is a high-accuracy task

- The new chart displays the same data
- *Visually estimate* the percentages using the new chart
- Fill-in the blanks in the table

Country	Percentage
China	
India	
United States	
Indonesia	
Pakistan	

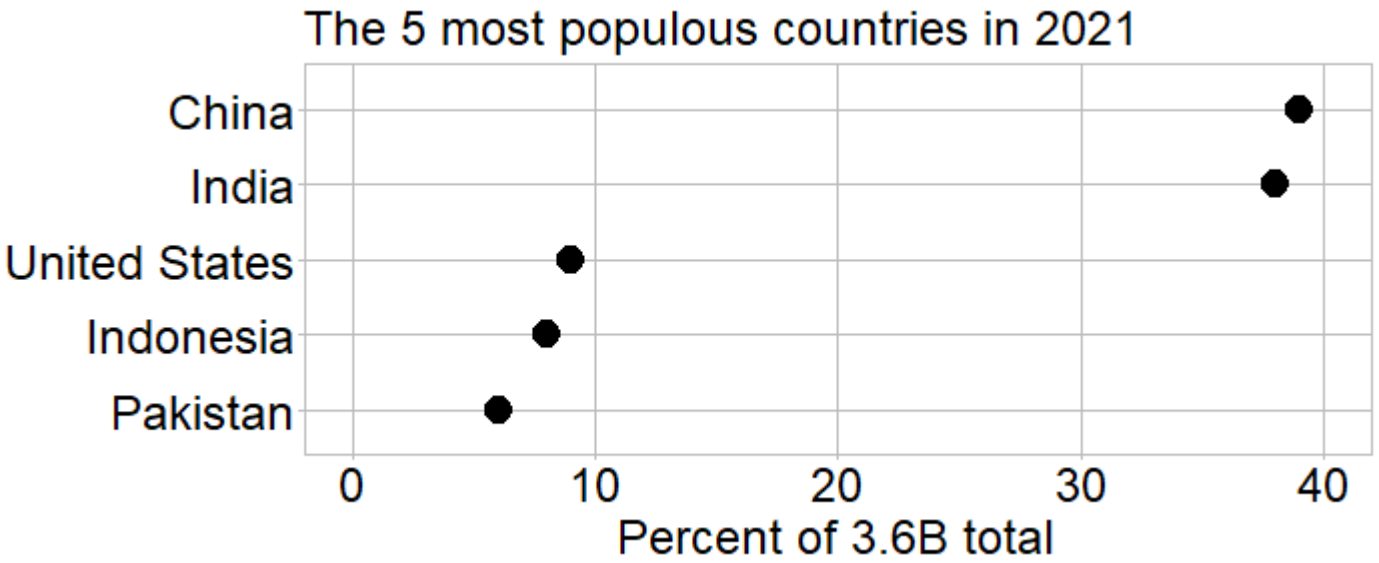


Data source: 2021 World population review

Judging values along a common axis is a high-accuracy task

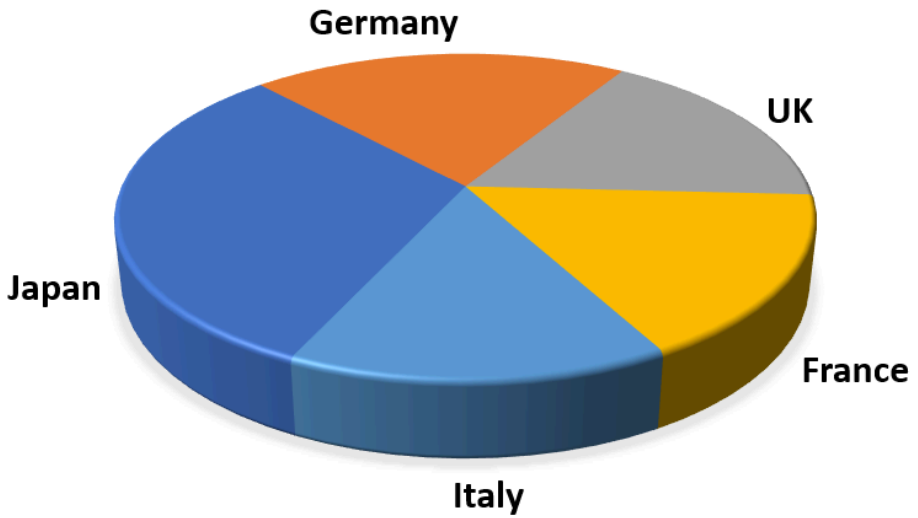
Compare your visual estimates to the data values.

Country	Percentage
China	39
India	38
United States	9
Indonesia	8
Pakistan	6



Data source: 2021 World population review

3D effects distort our judgment even further



- *Visually estimate* each country's percentage
- Fill-in the blanks in the table
- Total should be 100%

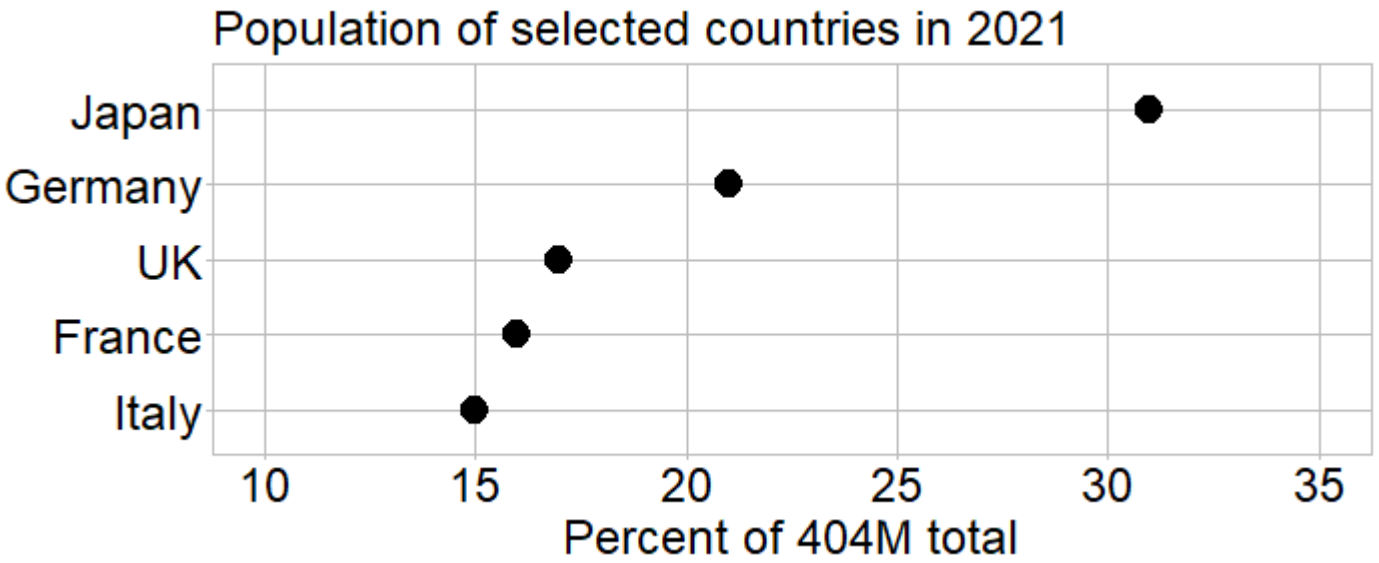
Country	Percentage
Japan	
Germany	
UK	
France	
Italy	

Data source: 2021 World population review (The pie represents 404M people.)

Again, a common scale improves our visual judgments

- The new chart displays the same data
- *Visually estimate* the percentages using the new chart
- Fill-in the blanks in the table

Country	Percentage
Japan	
Germany	
UK	
France	
Italy	

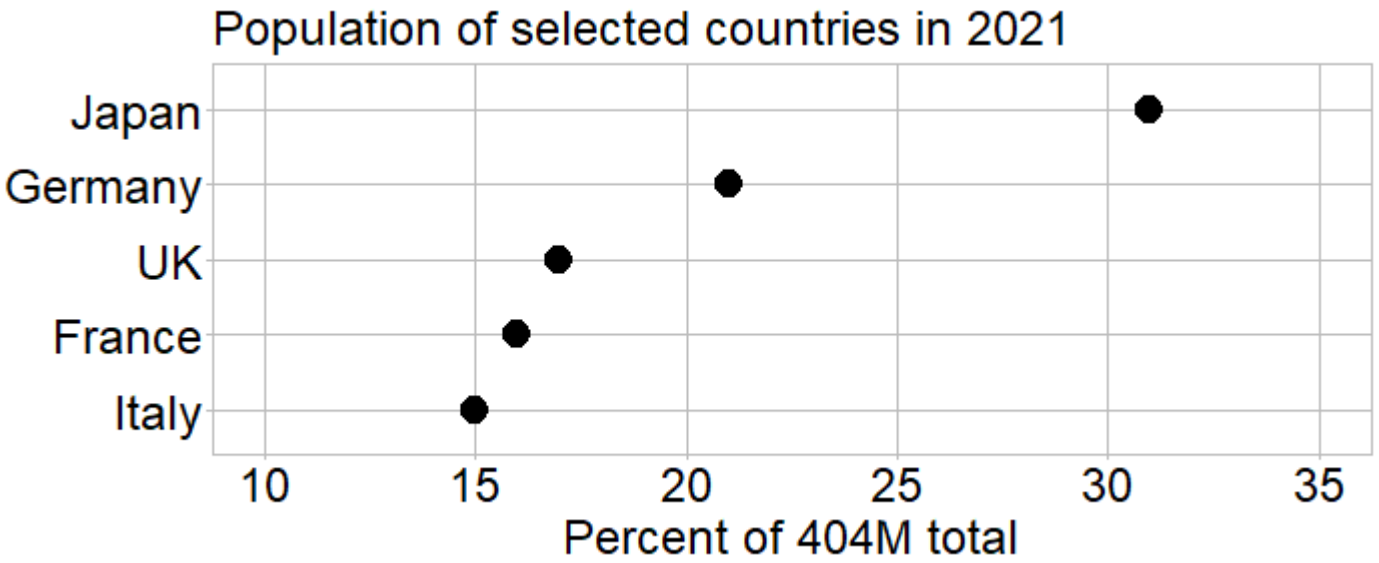


Data source: 2021 World population review

Again, a common scale improves our visual judgments

Compare your visual estimates to the data values.

Country	Percentage
Japan	31
Germany	21
UK	17
France	16
Italy	15



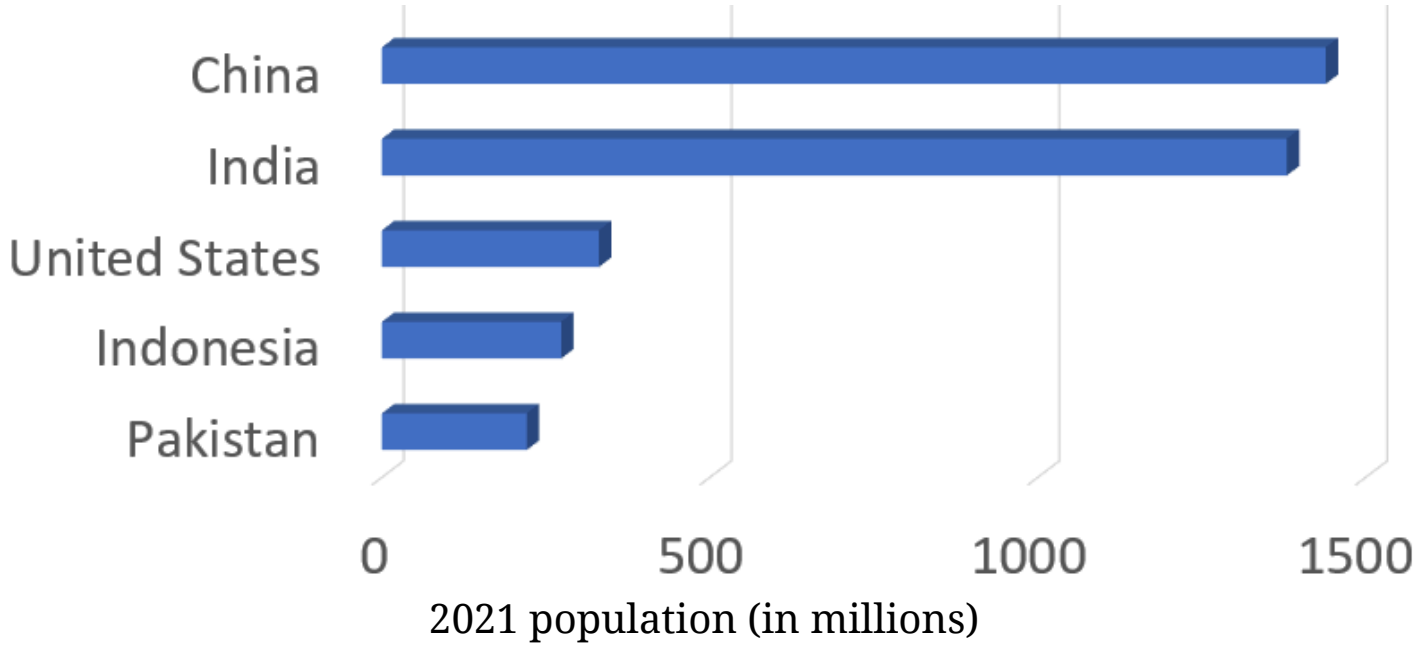
Data source: 2021 World population review

Effective alternatives to bar charts

3D effects always distort our judgment

- *Visually estimate* each country's population in millions
- Fill-in the blanks in the table

Country	Millions
China	
India	
United States	
Indonesia	
Pakistan	

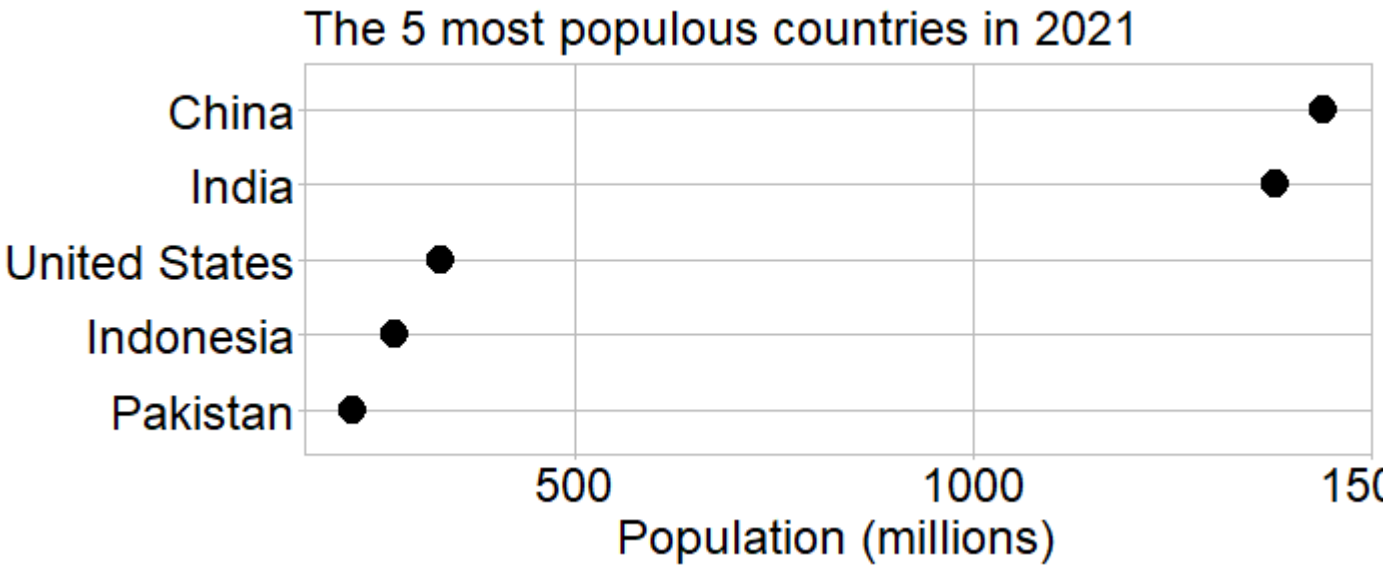


Data source: 2021 World population review

Same data—without 3D effects—along a common scale

- The new chart displays the same data
- *Visually estimate* the percentages using the new chart
- Fill-in the blanks in the table

Country	Millions
China	
India	
United States	
Indonesia	
Pakistan	

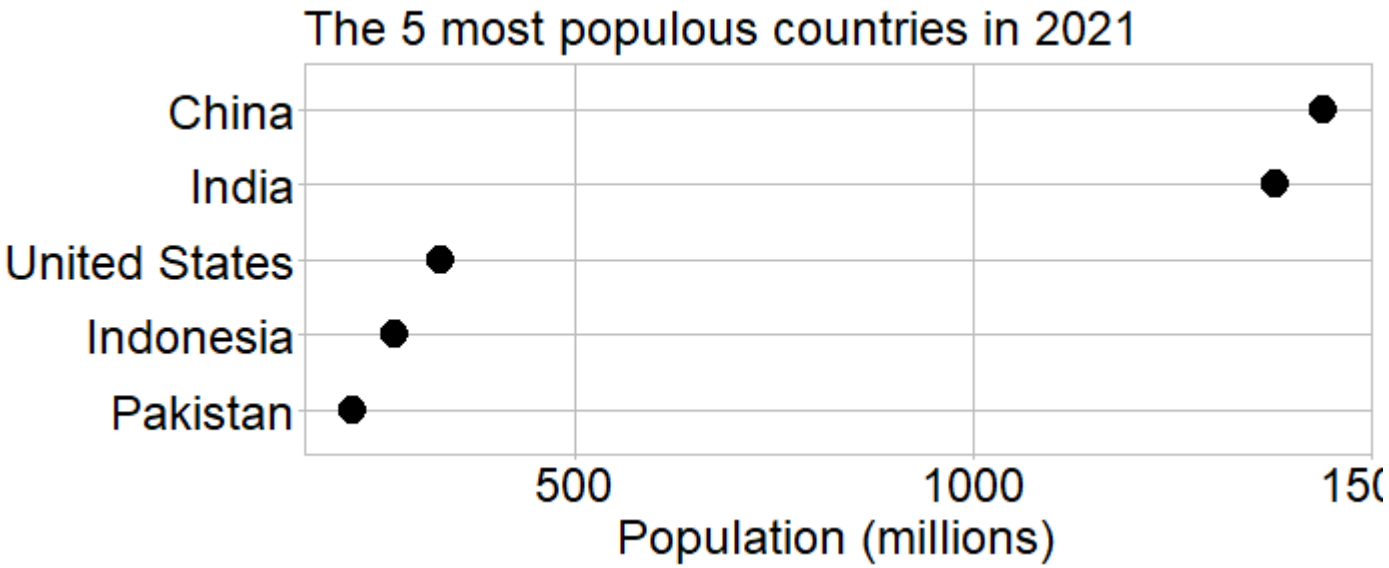


Data source: 2021 World population review

Same data—without 3D effects—along a common scale

Compare your visual estimates to the data values.

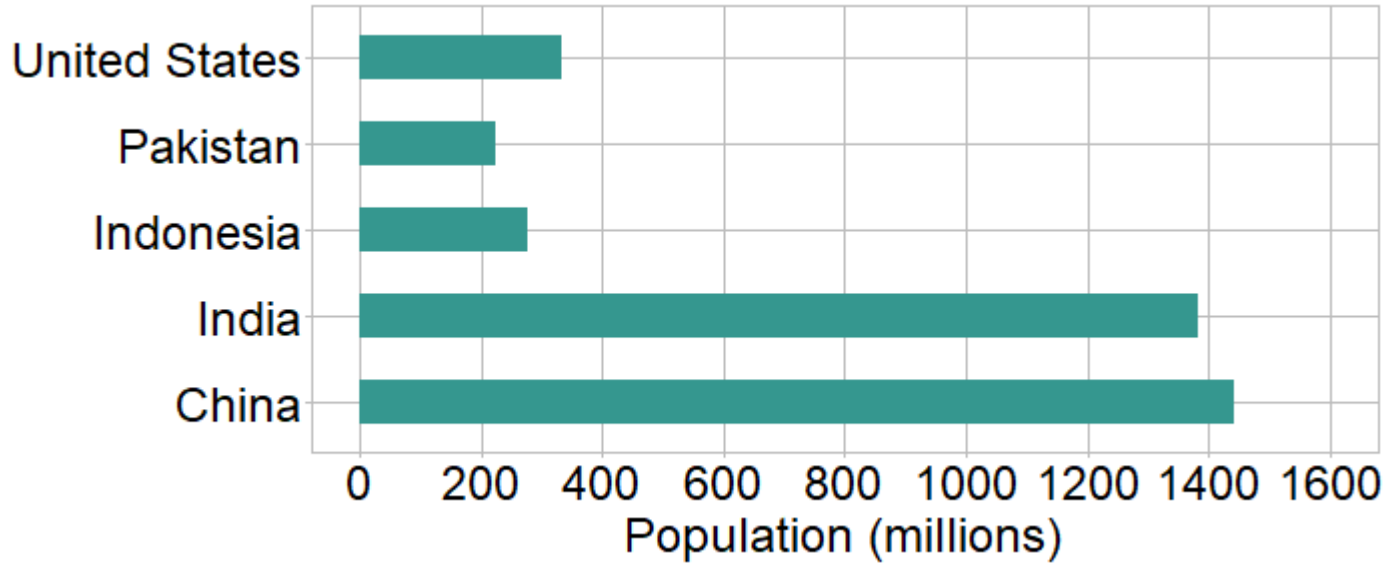
Country	Millions
China	1439
India	1380
United States	331
Indonesia	274
Pakistan	221



Data source: 2021 World population review

With a zero baseline and no 3D effects, bars are OK

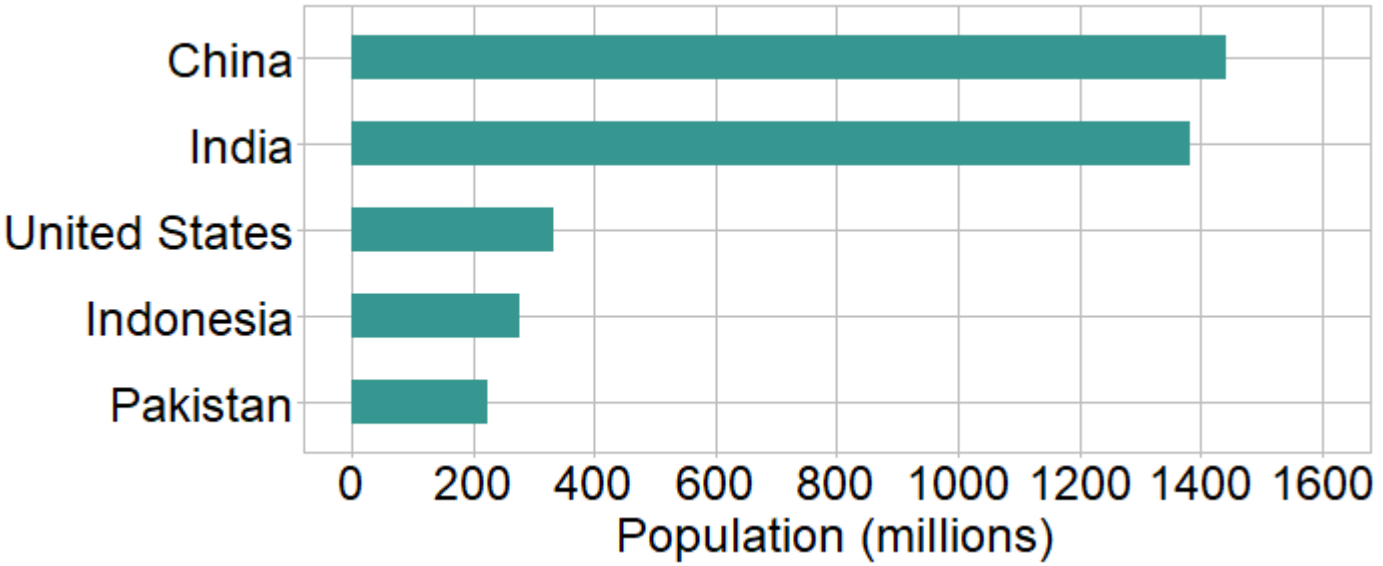
Bar charts must have a zero baseline to avoid deceiving your audience.



Data source: 2021 World population review

Row order affects the audience's ability to make comparisons

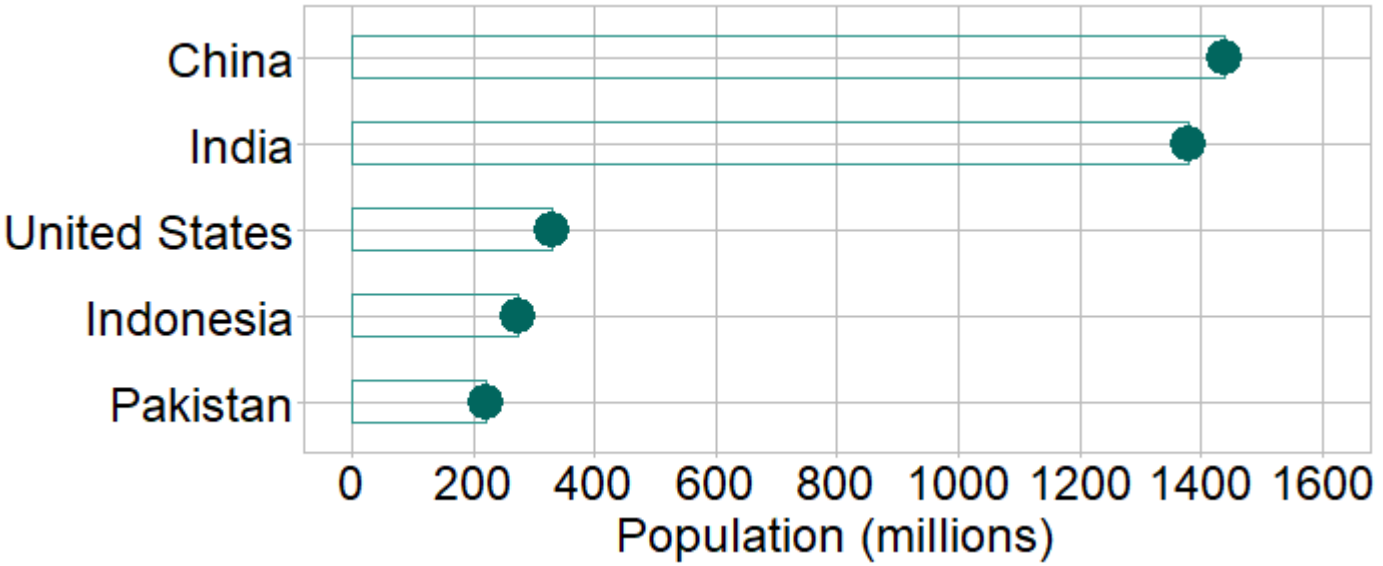
Ordering rows by the data values is usually superior to alphabetical order for visual comparisons.



Data source: 2021 World population review

Area is perceived differently than position

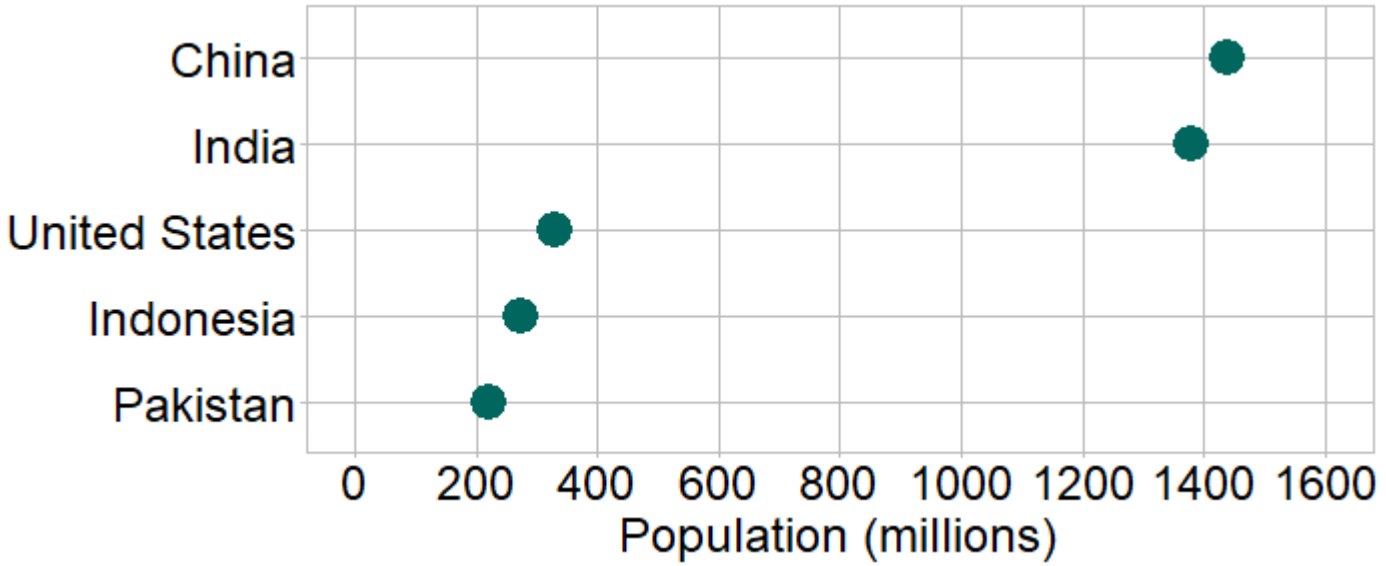
Area encodes no information. The only information in the bar is the position of its end point.



Data source: 2021 World population review

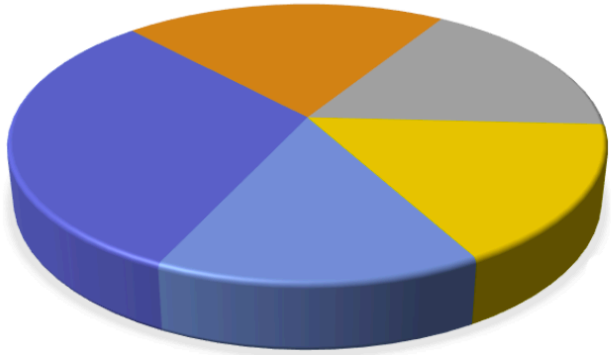
Dot charts allow direct visual comparison of quantities

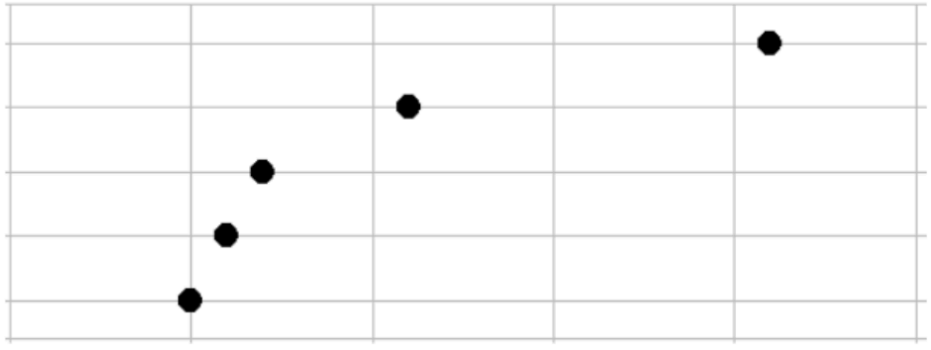
The result is the *dot chart* (Cleveland 1993).

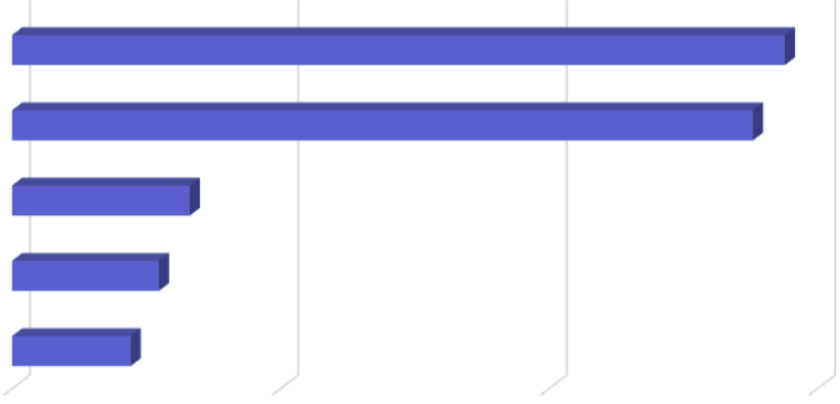


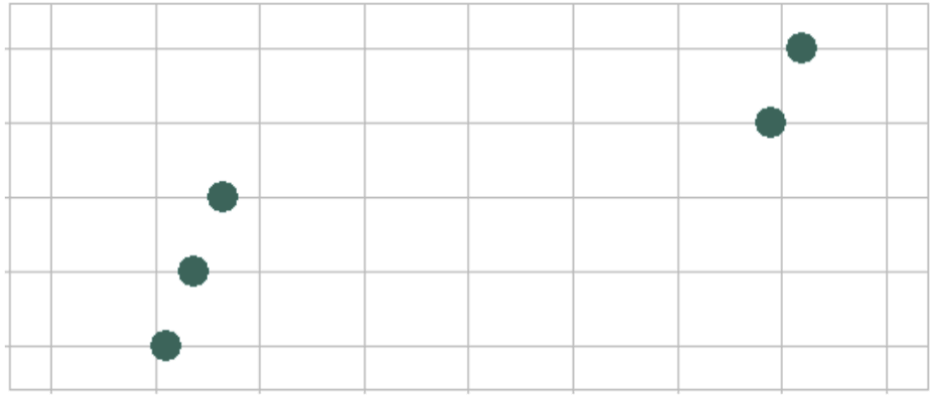
William S. Cleveland (1993) *Visualizing Data* Hobart Press, Summit, NJ.

Dot charts are effective replacements for pies and bars





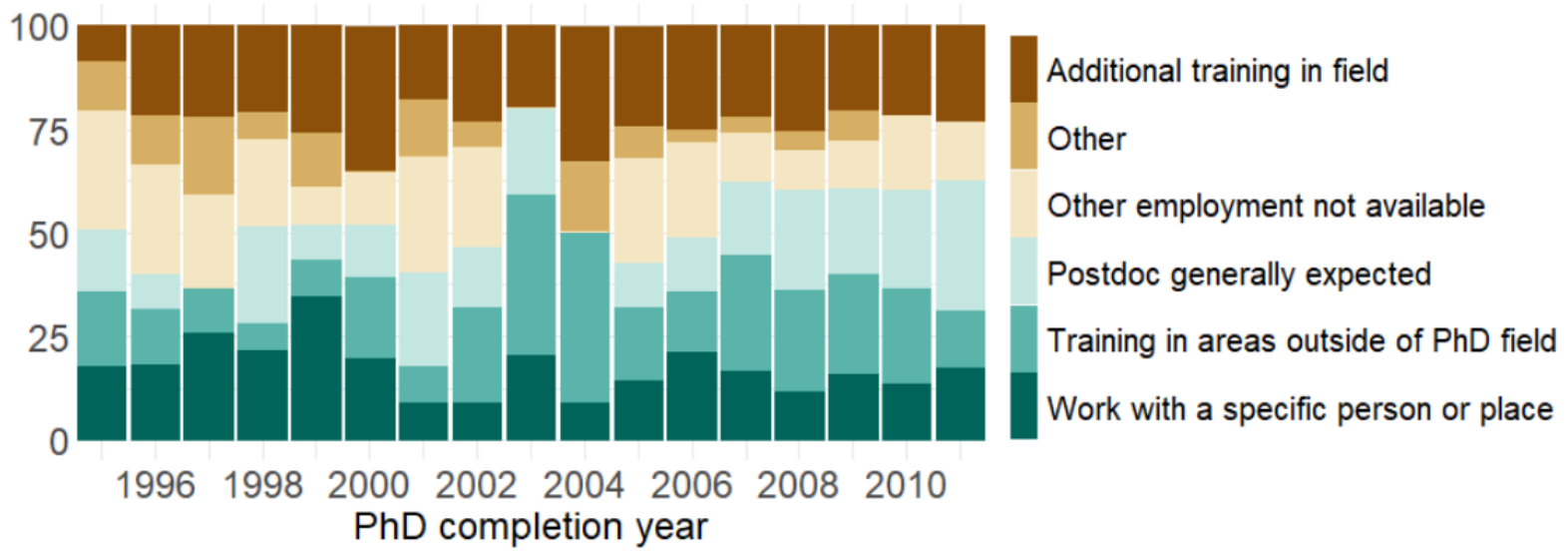




Aligning the design to the story

Visual grammar: charts encode information

Surveys: “What was your reason for taking this postdoc?”



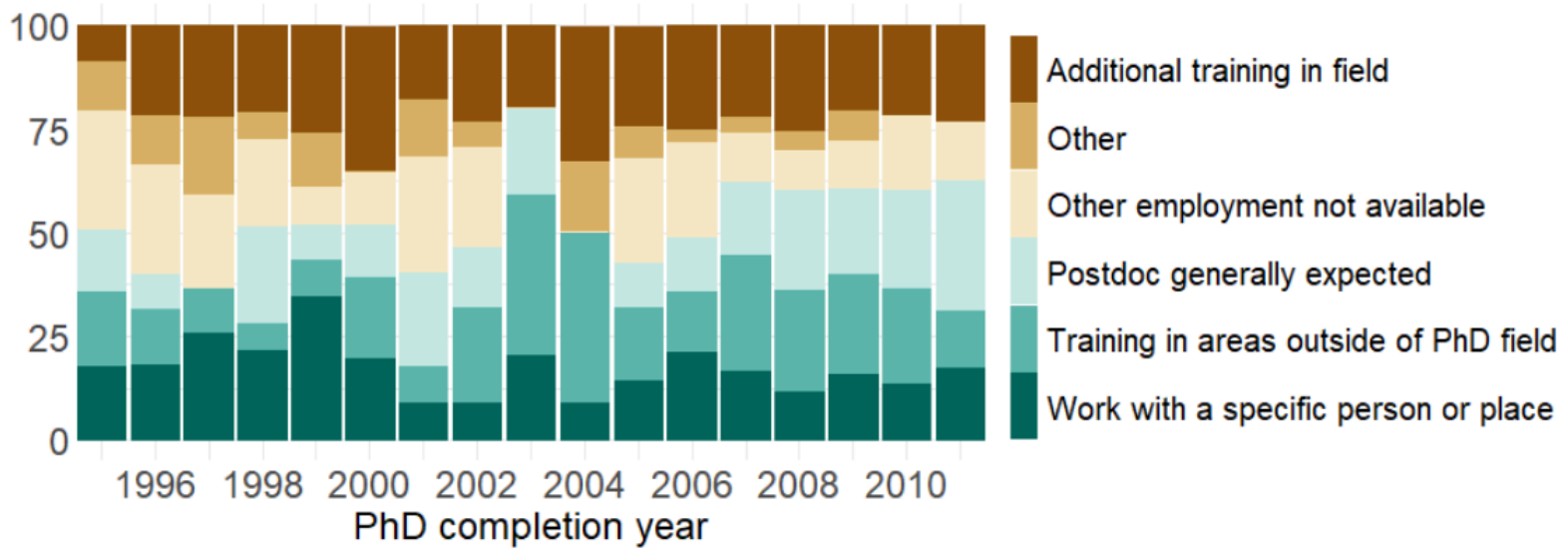
Decode the *information* before discussing *meaning*.

Select one color. What *information* does the color encode?

Data source: Main, Joyce B., Yanbing Wang, and Li Tan (2021) “The career outlook of engineering PhDs” *J of Engineering Education* 110 (4): 977–1002.

Visual rhetoric: charts convey meaning

Surveys: “What was your reason for taking this postdoc?”

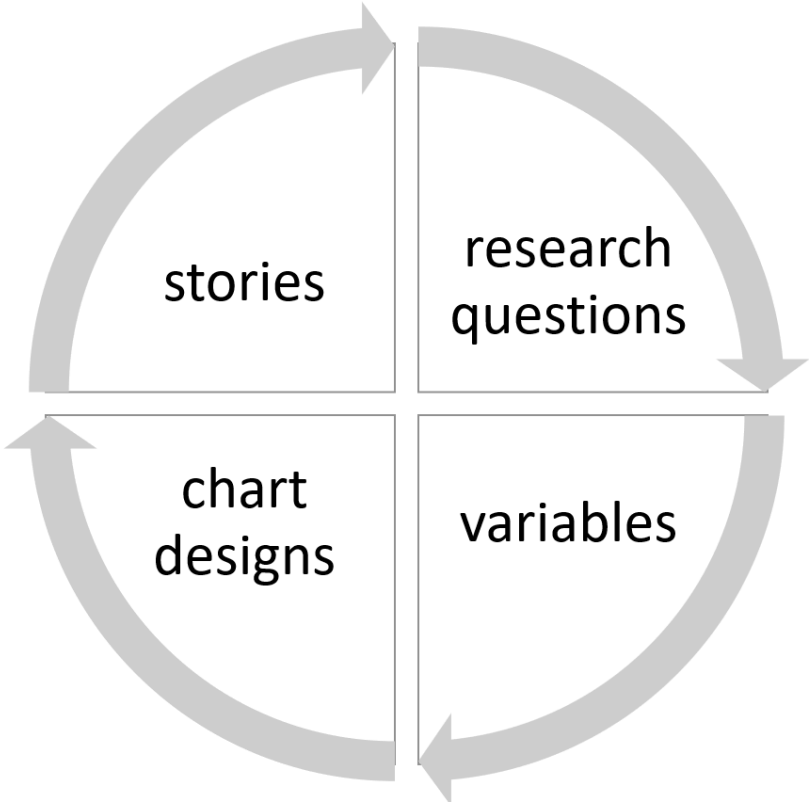


Agreeing on what the information *is*, we can consider what it *means*.

Meaning. Describe a story (if any) this chart *conveys to you.*

Data source: Main, Joyce B., Yanbing Wang, and Li Tan (2021) “The career outlook of engineering PhDs” *J of Engineering Education* 110 (4): 977–1002.

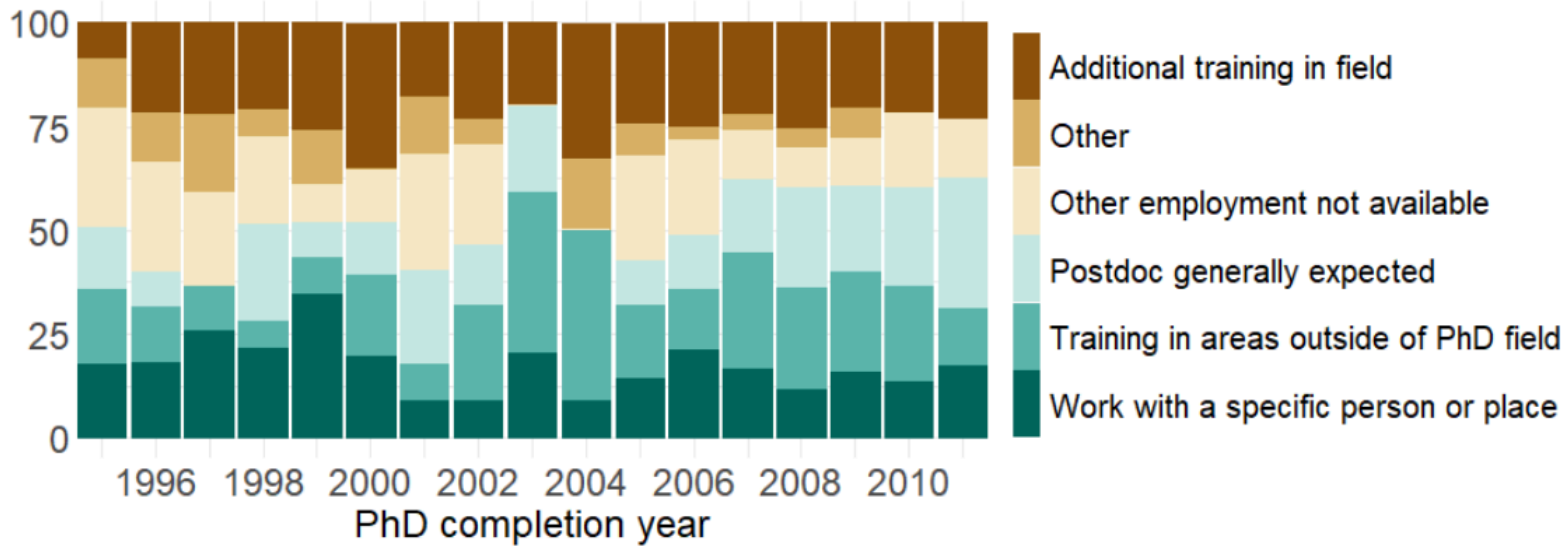
Visual grammar and visual rhetoric depend on the variables



- What is your question?
- What variables are measured?
- How are the variables classified?
- What chart designs suit these variables?
- What stories do the charts convey?
- How do the stories refine your questions?
- What new variables are needed?
- Repeat

What can we say about these variables?

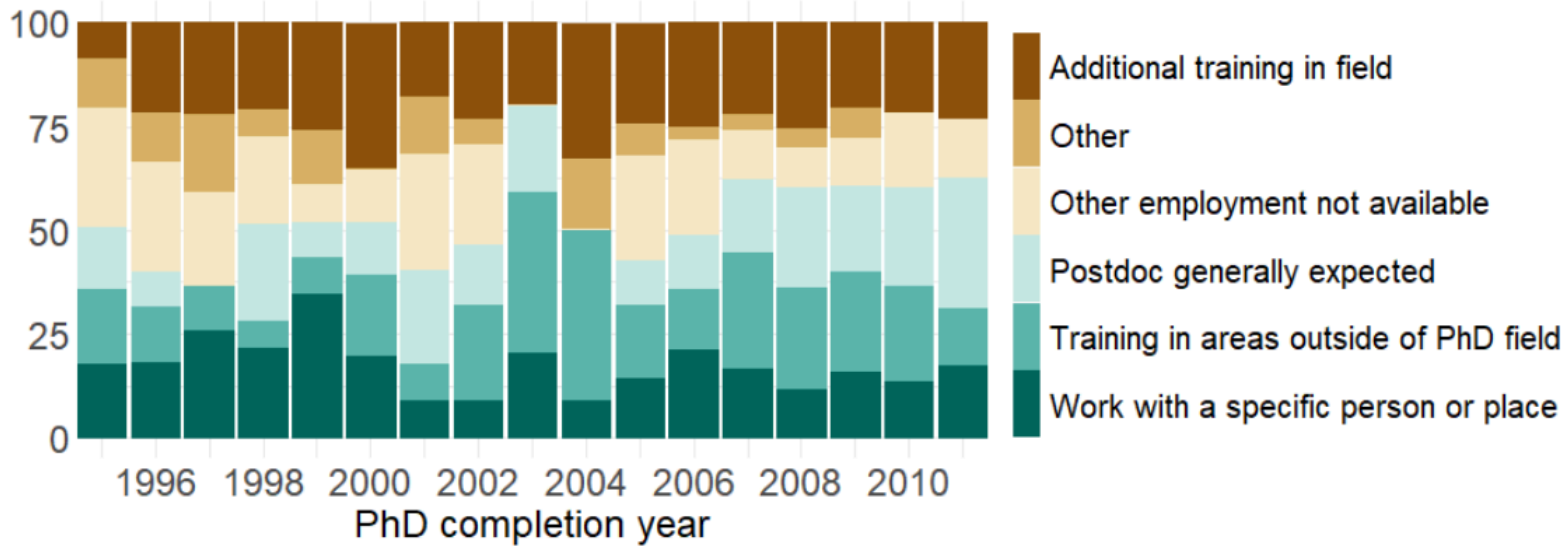
Surveys: “What was your reason for taking this postdoc?”



1. `_____PhD completion year_____` is a *categorical* variable.
2. `_____` is a *categorical* variable
3. `_____` is the *quantitative* variable
4. `_____` is the *independent* variable

What can we say about these variables?

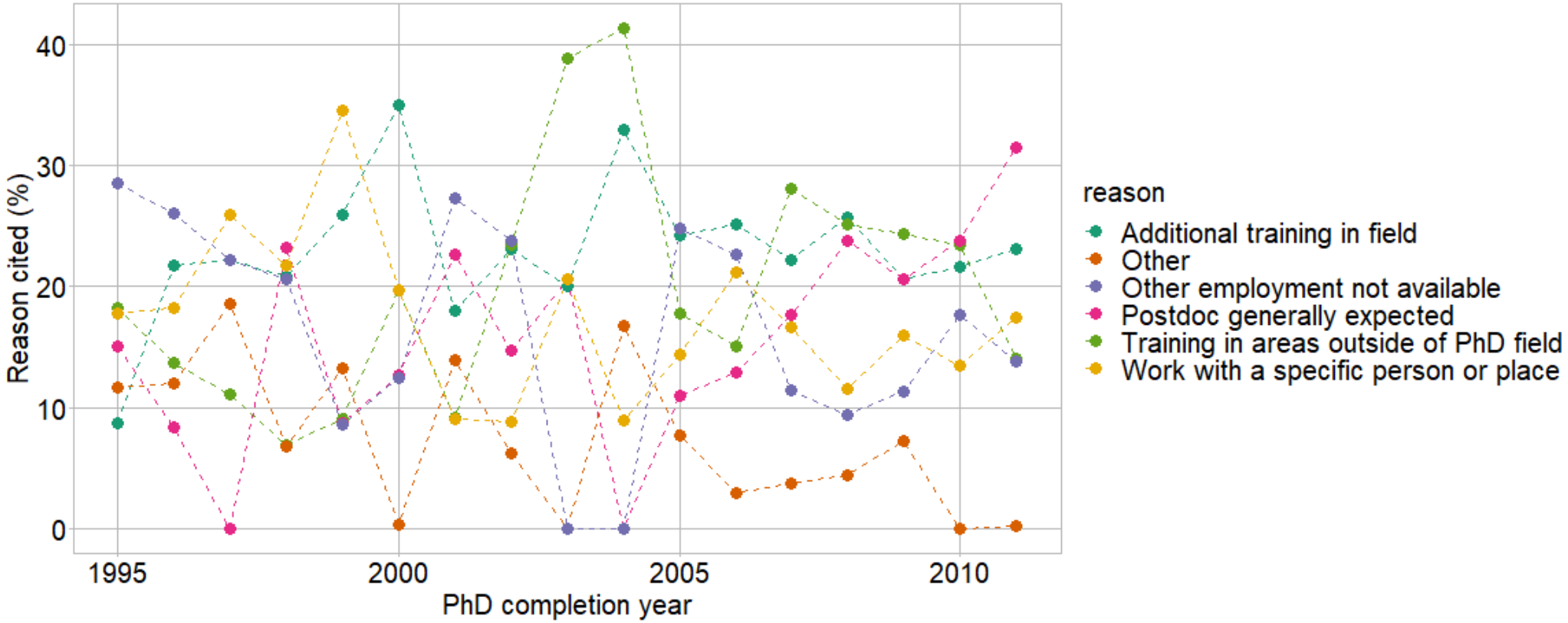
Surveys: “What was your reason for taking this postdoc?”



1. `_____PhD completion year_____` is a *categorical* variable.
2. `_____Postdoc reasons_____` is a *categorical* variable
3. `_____Percent citing a reason_____` is the *quantitative* variable (one observation per reason per year)
4. `_____PhD completion year_____` is the *independent* variable

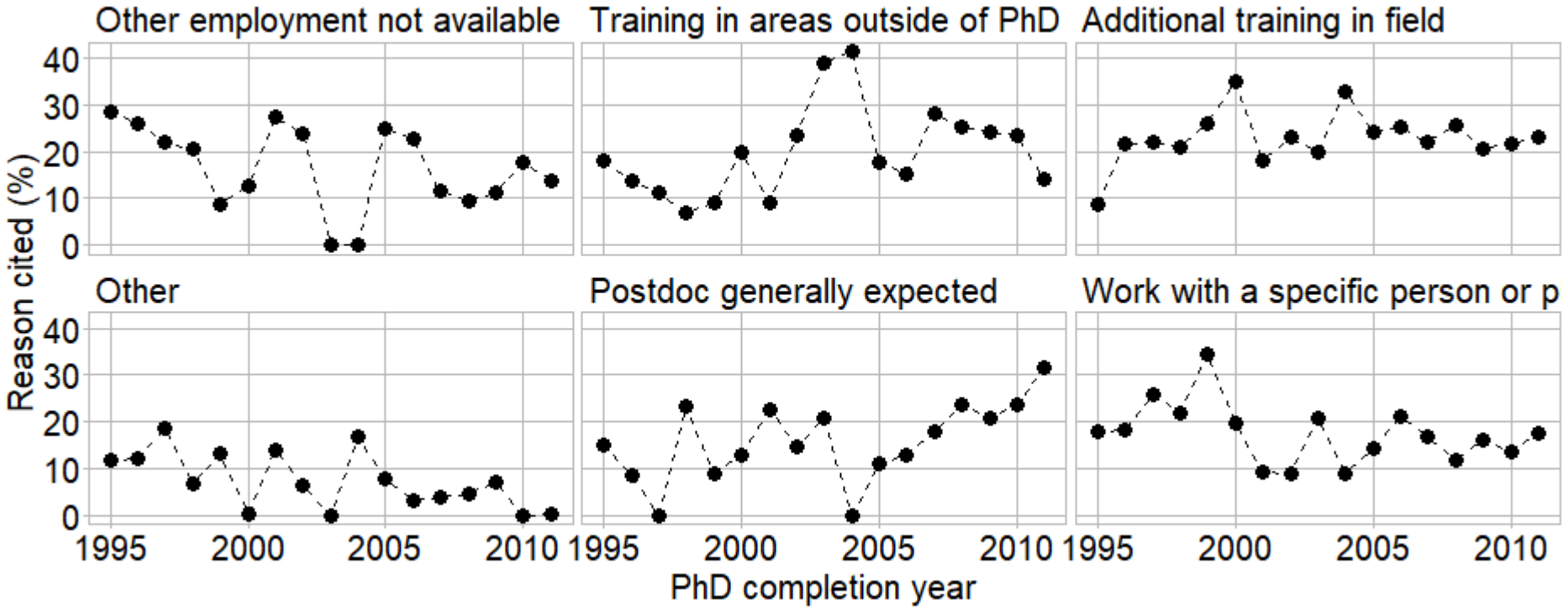
Time series? Use a line chart.

The original chart implies the data are a time series.



Time series? Use a line chart.

Un-clutter the display using one panel per reason



Meaning. Describe a story (if any) this chart *conveys to you.*

An unstated assumption underlies the visual muddle

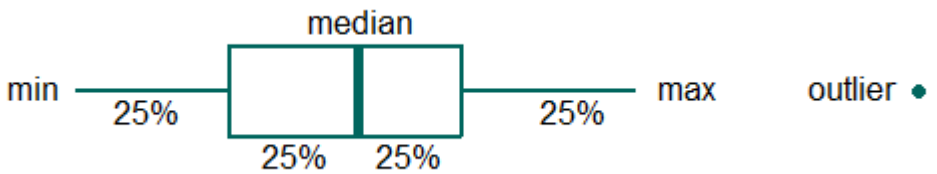
Emphasizing the trivial (Howard Wainer)

- Poor choice of independent variable
- Imposes a visual convention

Ignore time, the remaining variables are:

- percent citing a reason (quantitative) distributed
- reasons (categorical, 6 levels)

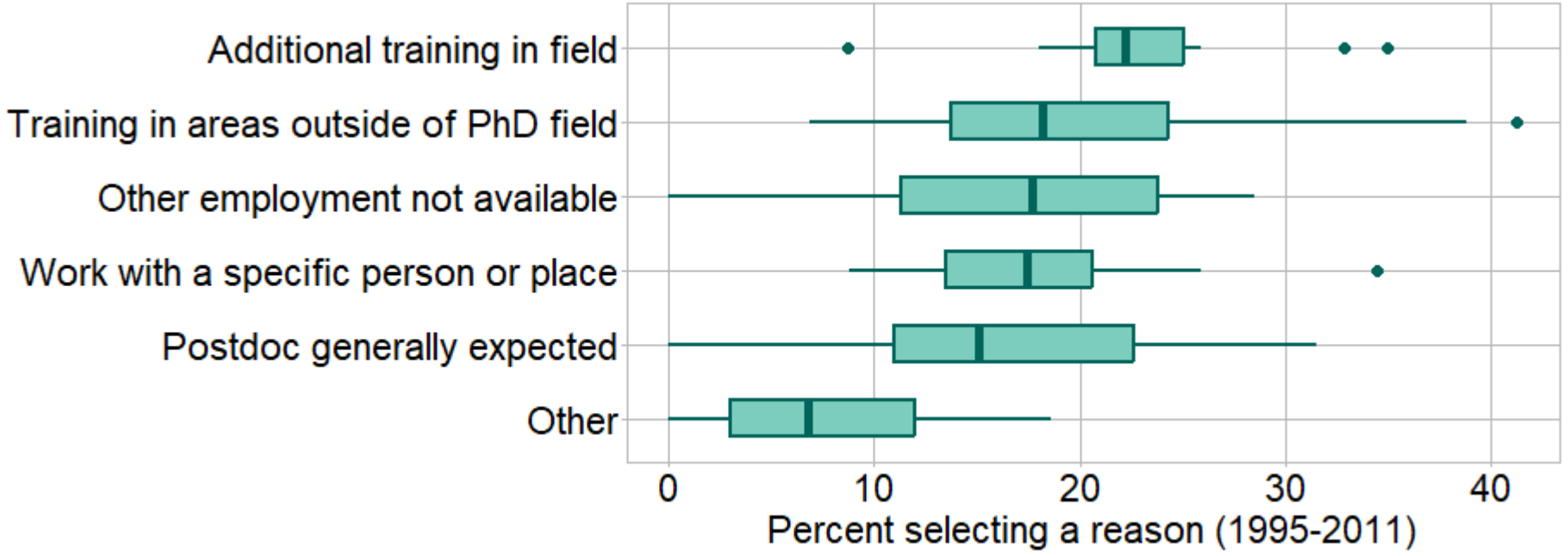
A distributed quantity is displayed in a box-and-whisker plot



Howard Wainer (1997) *Visual Revelations*, Lawrence Erlbaum Assoc., Mahwah, NJ, p. 30.

Distributions? Use a box-and-whisker plot.

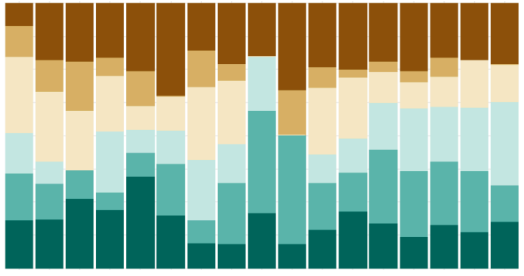
Surveys: “What was your reason for taking this postdoc?”



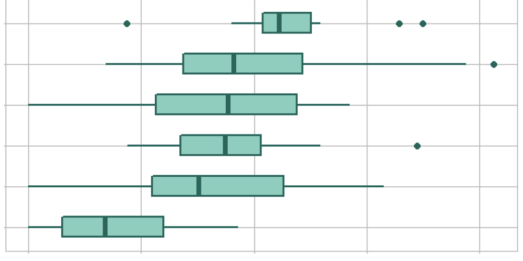
Meaning. Describe a story this chart *conveys to you.*

Reflect on perception, reasoning, and credibility

Compare the stacked bar



to the boxplot design.



Select any prompt. Outline your response:

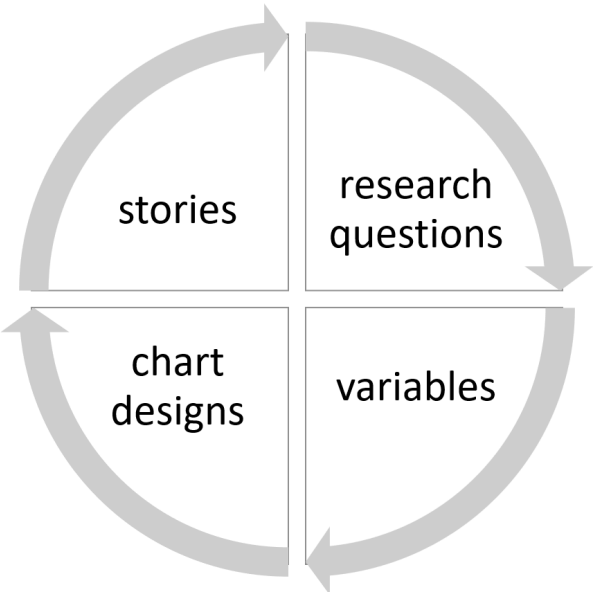
- Compare designs: Quantitative data are *perceived* accurately.
- Compare designs: *Reasoning* about the data is supported effectively.
- Compare designs: An argument is given *credible* visual support.

Advice from experts

Jean-Luc Doumont

The optimal graph design depends on two factors, primarily,

1. The message to be conveyed
2. The variables to be shown



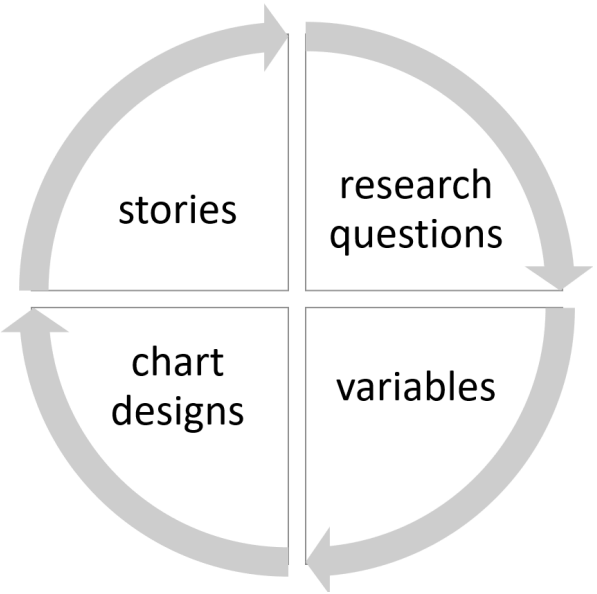
Jean-luc Doumont (2009) *Trees, Maps and Theorems*, Principiae.

Edward Tufte

The task of the designer is to give visual access

to the subtle and the difficult—that is,

the revelation of the complex.

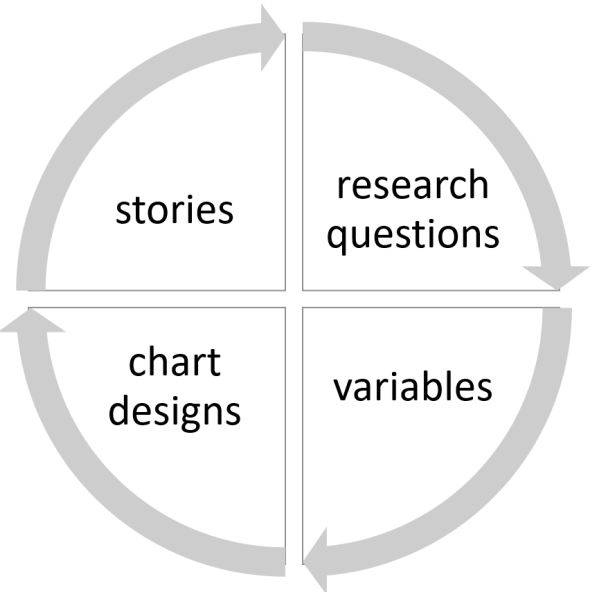


Edward Tufte (1983) *The Visual Display of Quantitative Information*, Graphics Press.

Stephanie Evergreen

What's your point?

Seriously, that's the most important question.

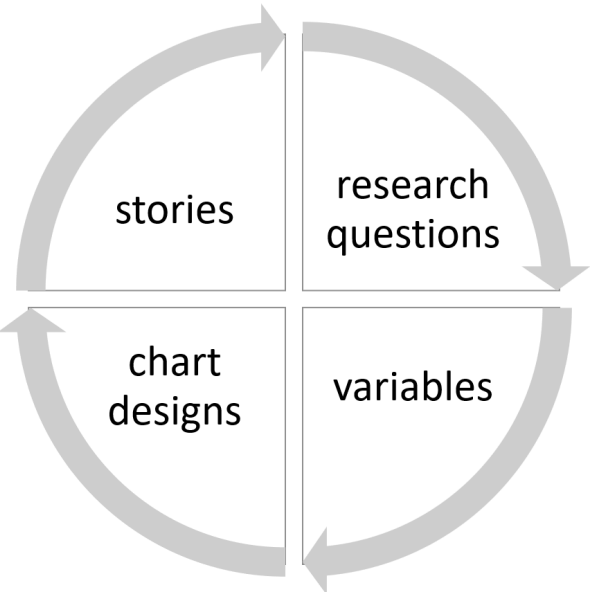


Stephanie D. H. Evergreen (2017) *Effective Data Visualization*, Sage.

Alberto Cairo

Charts can lie to to us because we read too much into them or see in them what we want to believe.

Don't lie to yourself (or others) with charts.



Alberto Cairo (2019) *How Charts Lie*, W.W. Norton

Ideas to consider

- **Characterize** the data structure and content
- **Explore** a story's context, causality, and complexity
- **Align** visual and verbal logic by revising iteratively
- **Edit** to suit the rhetorical goals for each audience
- **Control** every pixel—avoid thoughtless conformity
- **Question** are you seeing only what you want to believe?