More effective data visualisation

Visualisation for communication with others

Dr Nicola Rennie

Introduction

Data visualisation specialist, mainly working with R, Python, and D3.

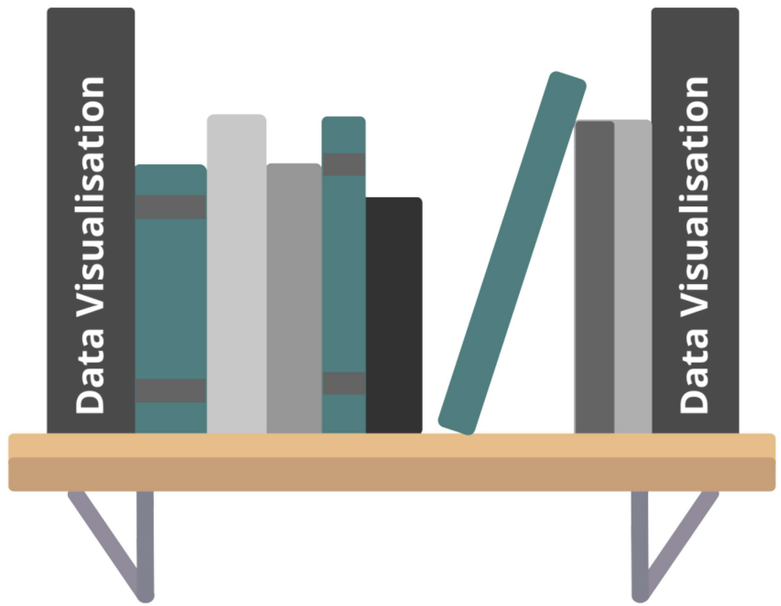
Background in statistics, operational research, and data science consultancy.

Co-author of Royal Statistical Society's *Best Practices for Data Visualisation* guidance.

Why visualise data?

Data visualisation has two main purposes:

- Exploratory data analysis and identifying data issues
- Communicating insights and results



> See nrennie.rbind.io/training-data-visualisation/slides1.html for exploratory data analysis discussion.

Outline

This session will include:

- the role of visualisation in communication;
- how to design more effective visualisations;
- examples of good and bad charts!

Communicating insights with data visualisation

Grab attention

Visualisations stand out. If a reader is short on time or uncertain about whether a document is of interest, an attention-grabbing visualisation may entice them to start reading.

Improve access to information

Textual descriptions can be lengthy and hard to read, and are frequently less precise than a visual depiction showing data points and axes.

Summarise content

Visual displays allow for summarising complex textual content, aiding the reader in memorising key points.

Designing better charts

What are you trying to communicate?

Who are you trying to communicate it to?

Visualisations should be intuitive to the group of people you are communicating with.

Think about your audience

- Dissertation presentation?
- Dissertation report?
- Journal article?
- Magazine article

- Conference presentation?
- Stakeholder meeting?
  - NHS data analysts?
  - Clinicians?
  - Policy makers?

> These are different audiences. They might each require a different type of visualisation for the same data.

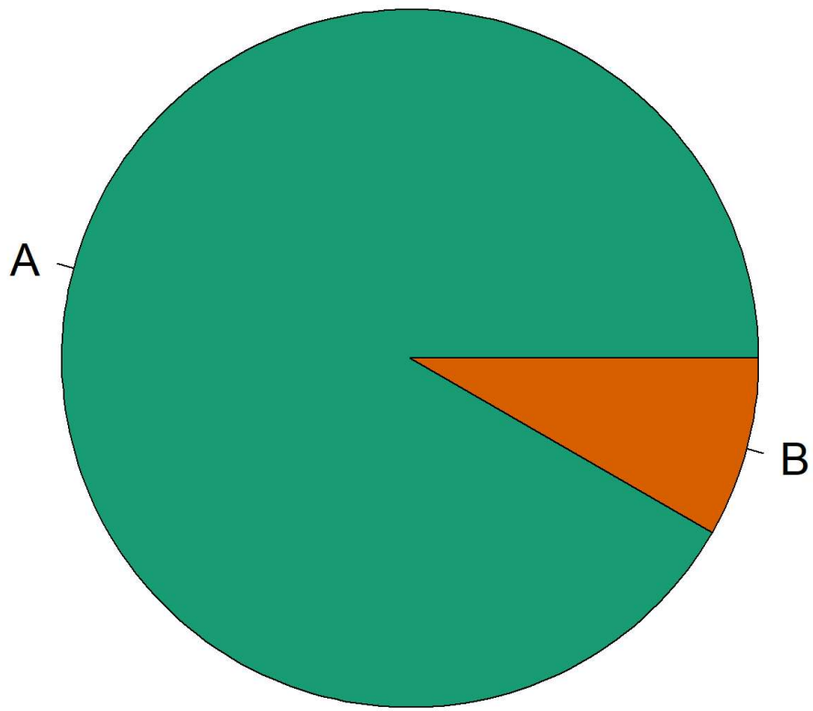
What makes charts less intuitive?

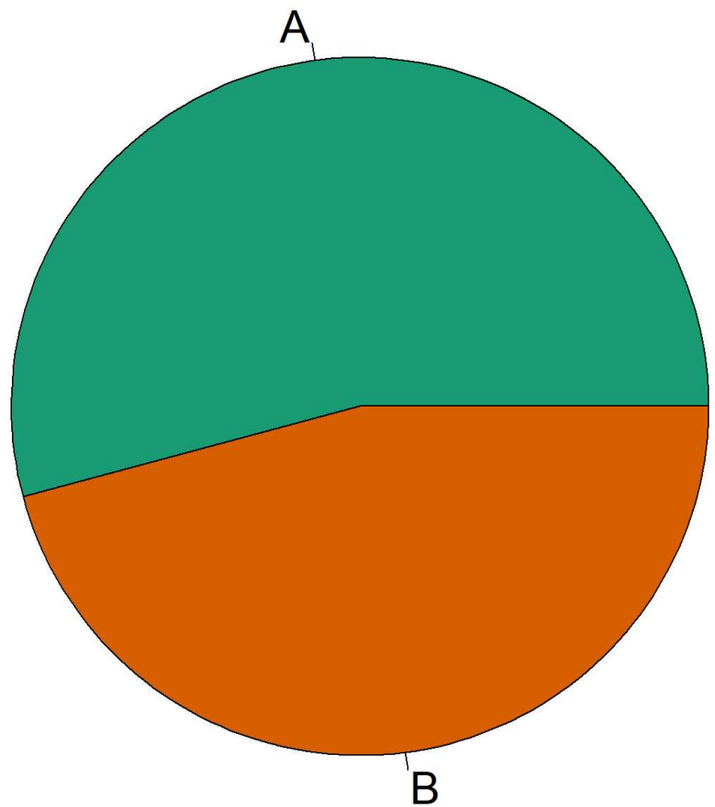
- Type of chart doesn’t match the type of data
- Inconsistent colour schemes
- Poor choices of axes
- Overly complex chart design
- Not enough explanation

Some types of charts are less intuitive than others.

Why do people hate pie charts?

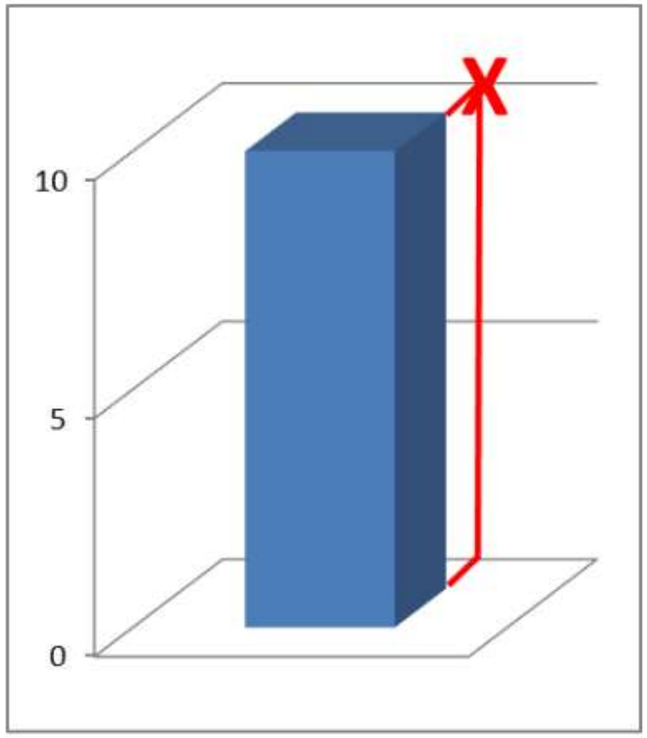
Good data visualisation isn't just based on opinions...

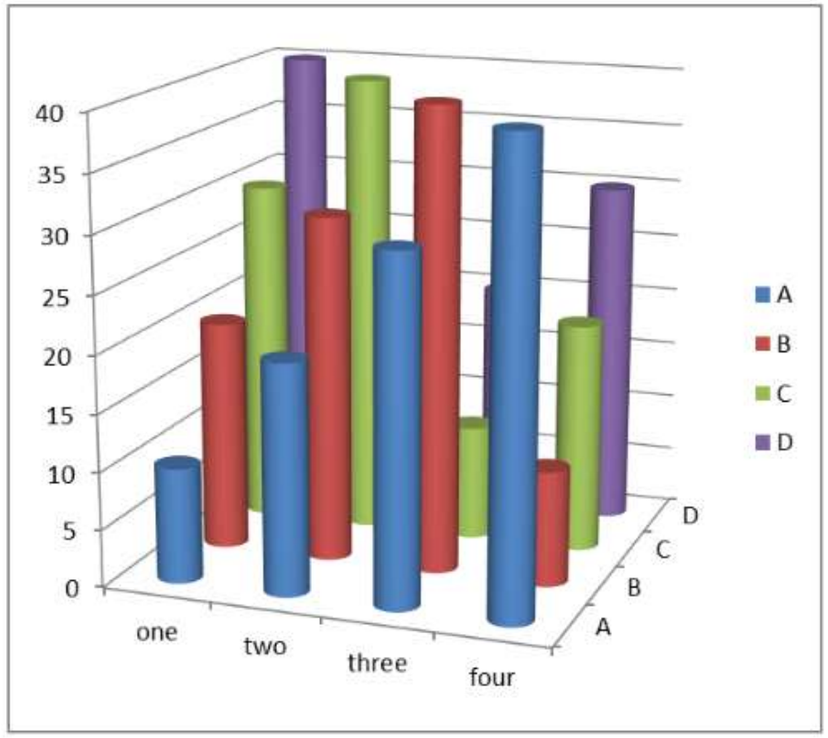




Why do people hate 3D charts?

On the plot on the left, how tall is the bar?





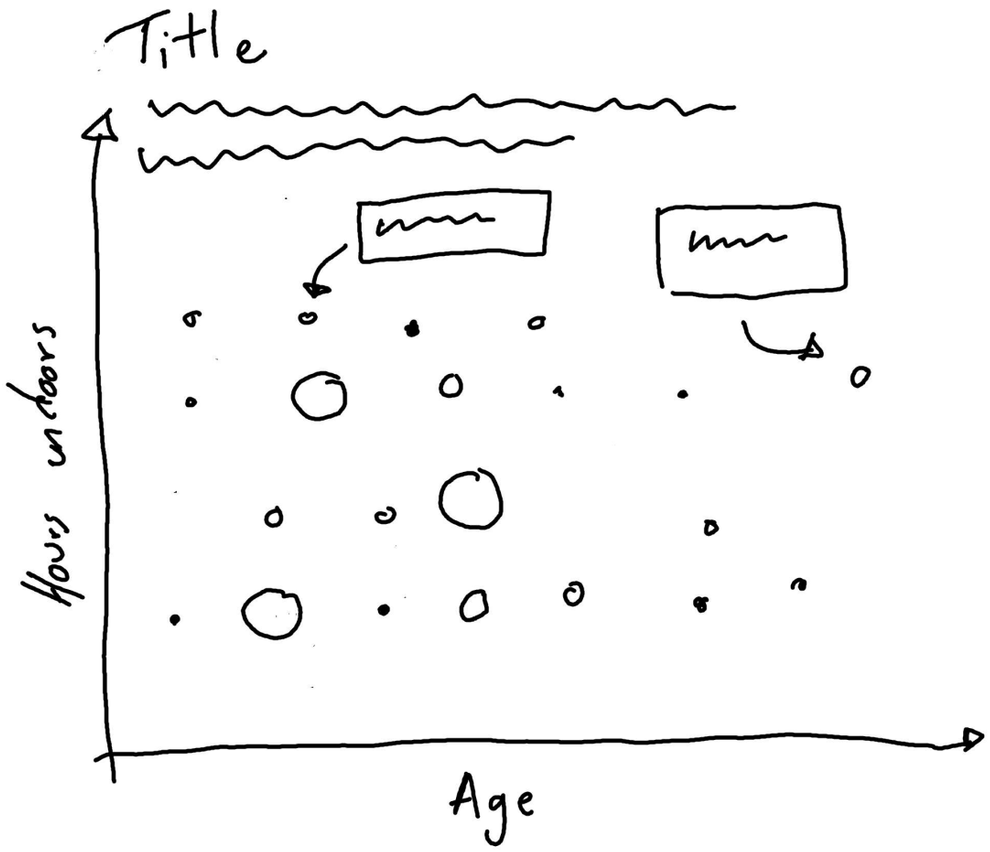
The design process

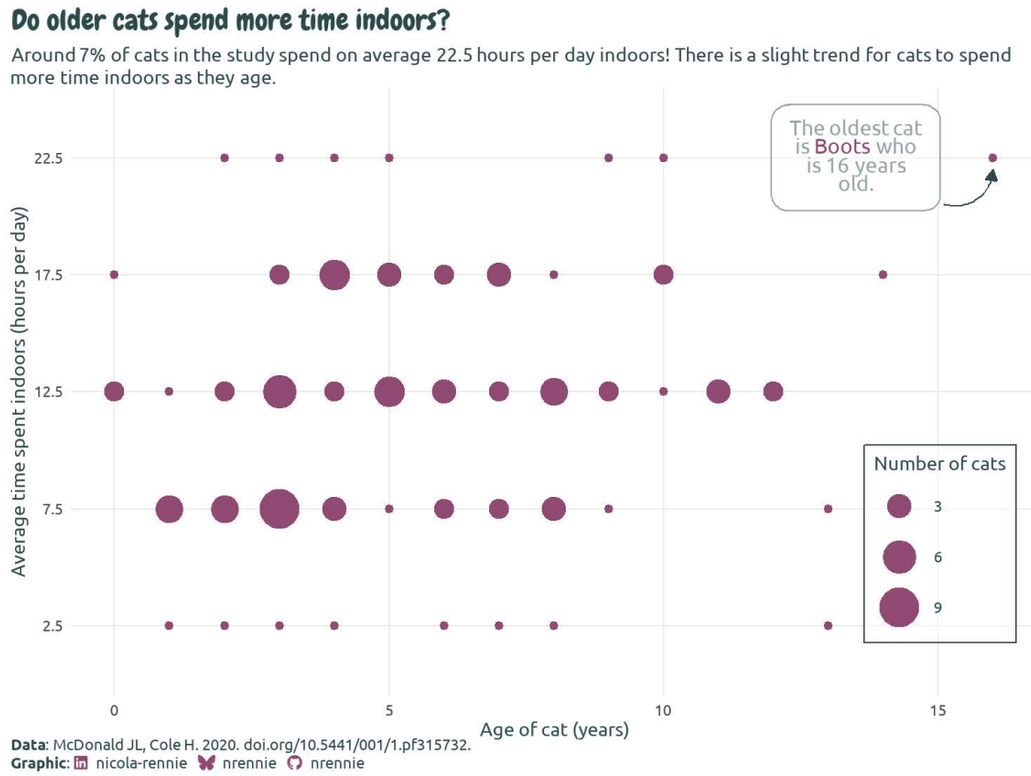
Questions to answer before you build

- What type of chart?
- How will you design that chart?
  - What orientation is the plot?
  - Are there multiple plots?
  - Should you use colour?
  - What does the title say?

> Even after you decide to make a bar chart, for example, there are still a lot of choices to be made about how that bar chart will look.

Start with sketching

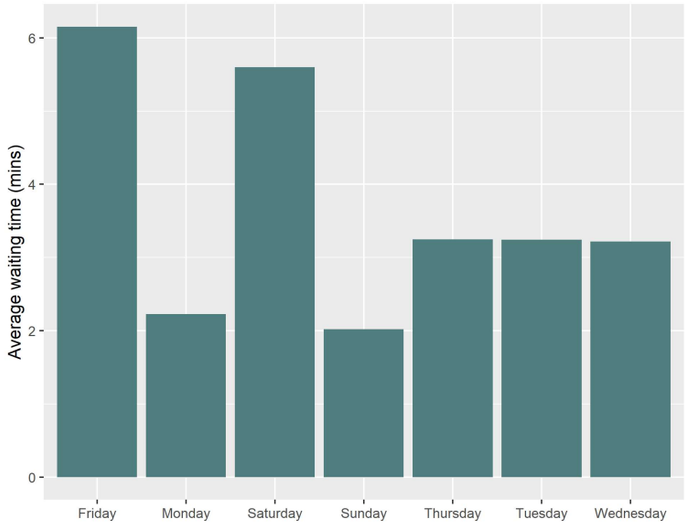




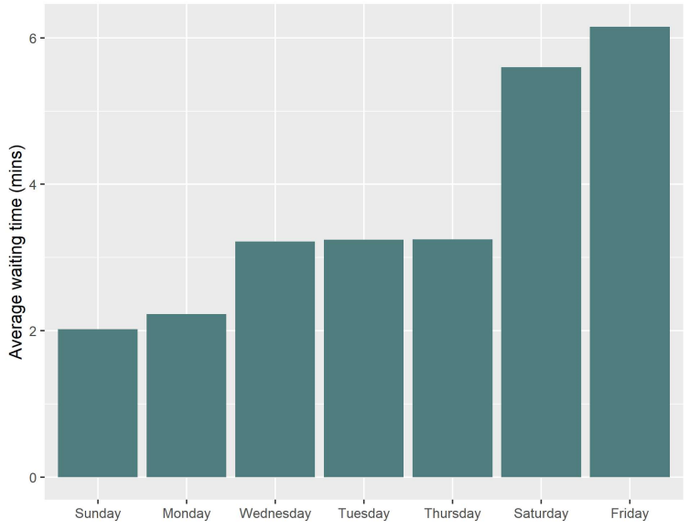
Choosing appropriate axes

Ordering discrete variables

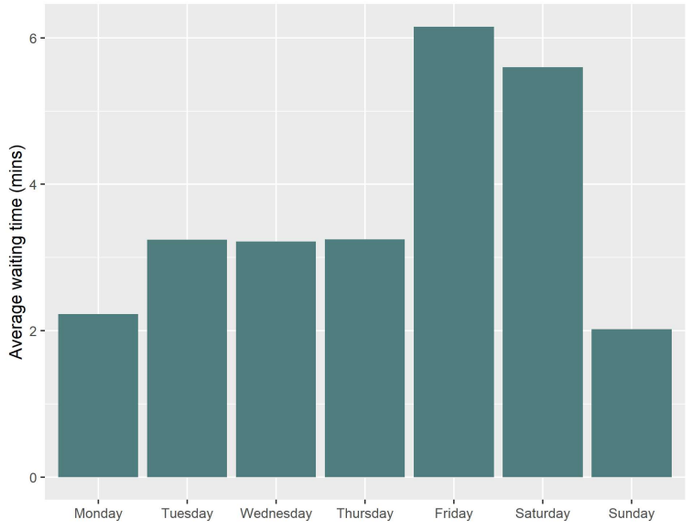
Alphabetical order:



Magnitude ordered:

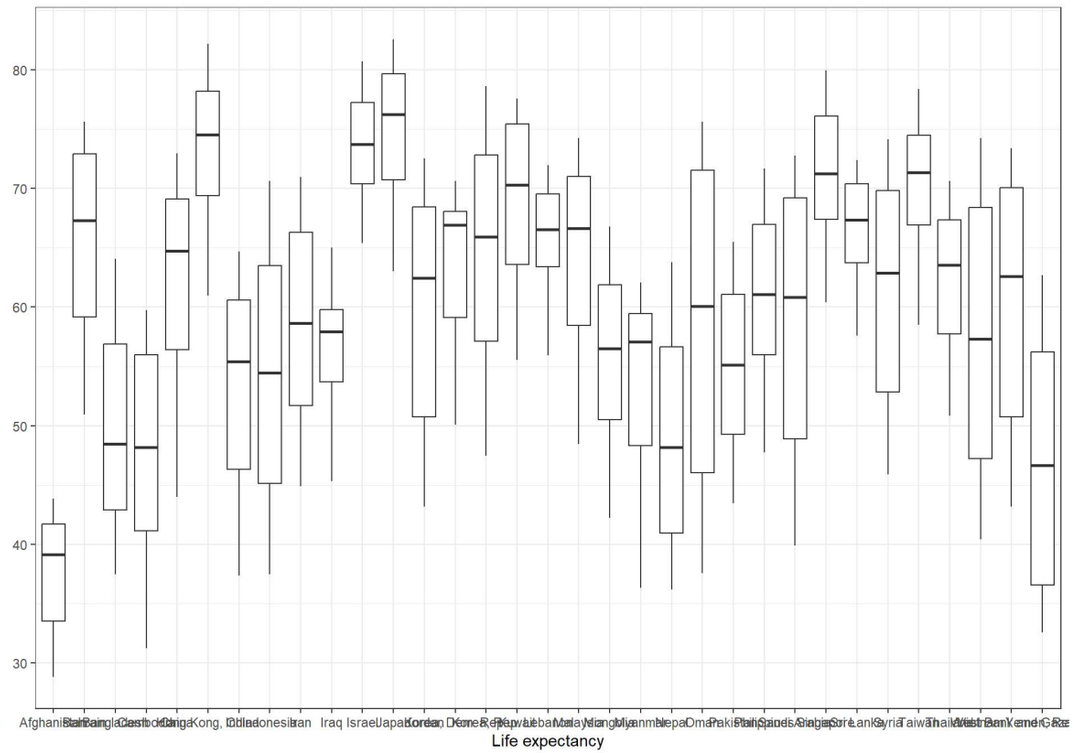


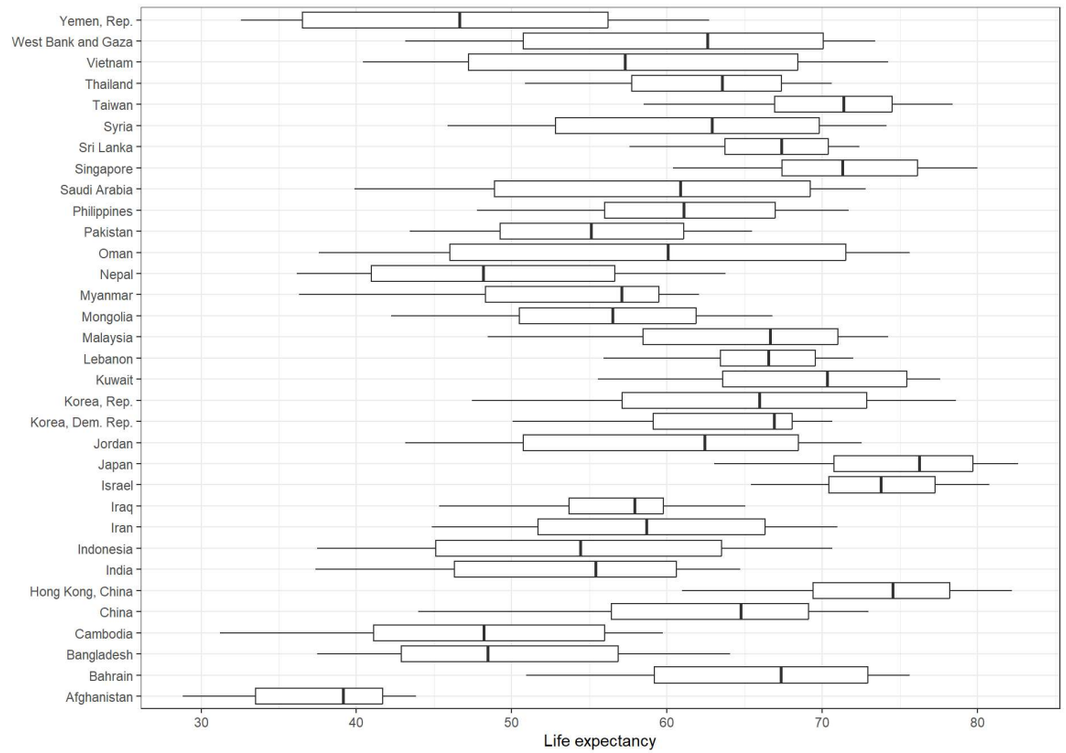
Naturally ordered:



Don’t make it more difficult than it needs to be

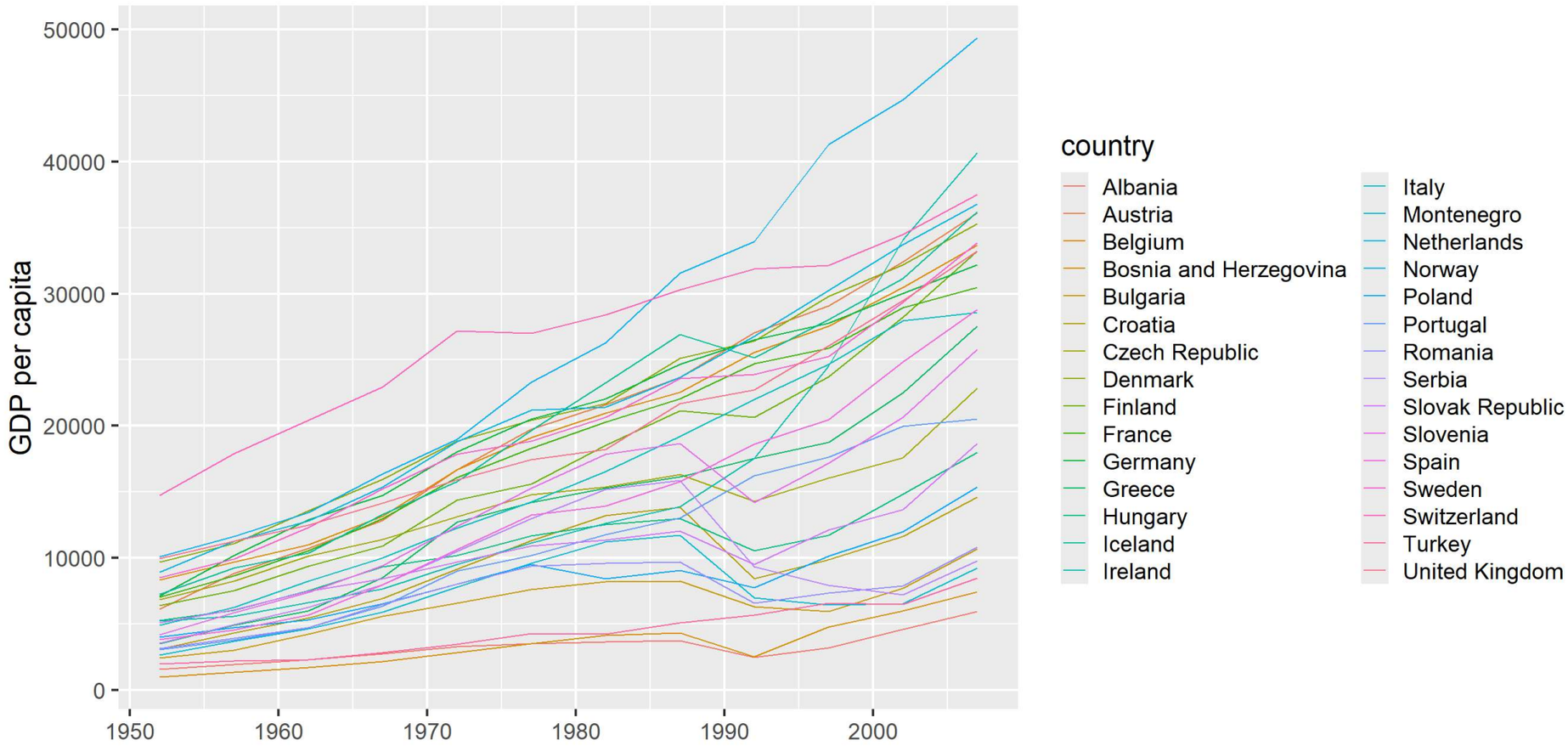
Longer labels are best on the y-axis, horizontally.





Plotting multiple values

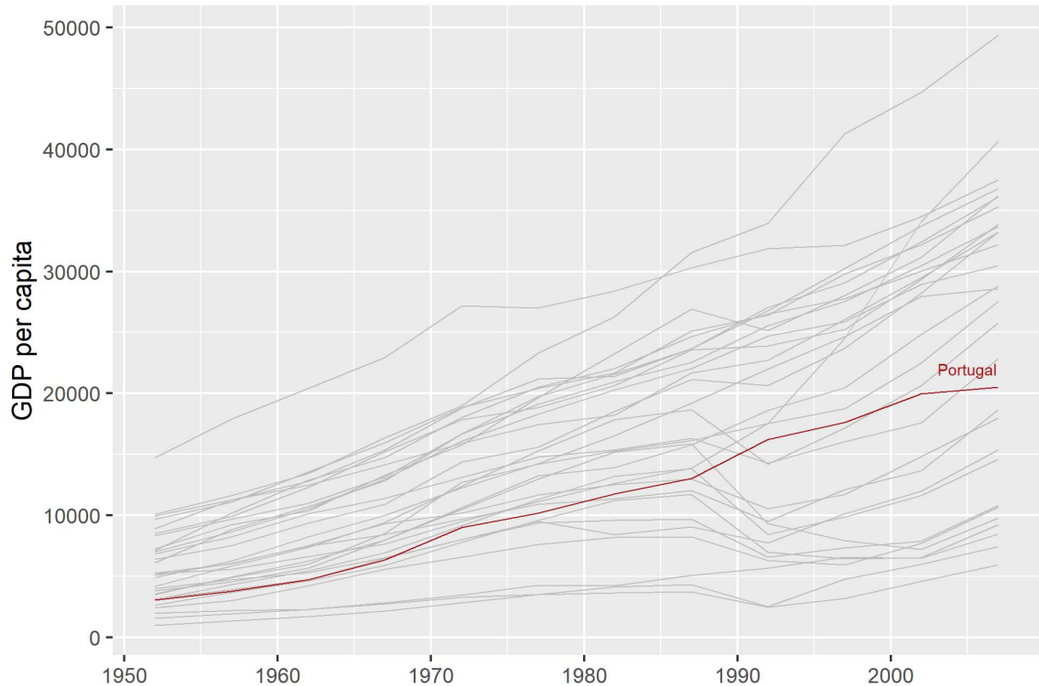
Avoid *spaghetti* plots!



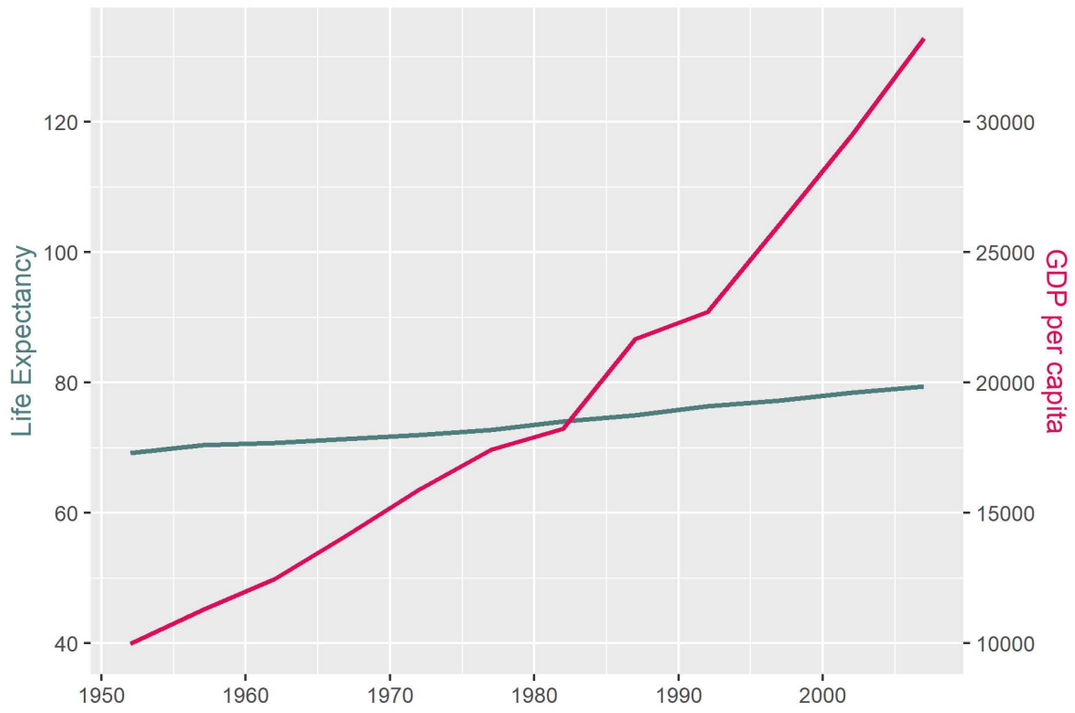
Effectively plotting multiple values

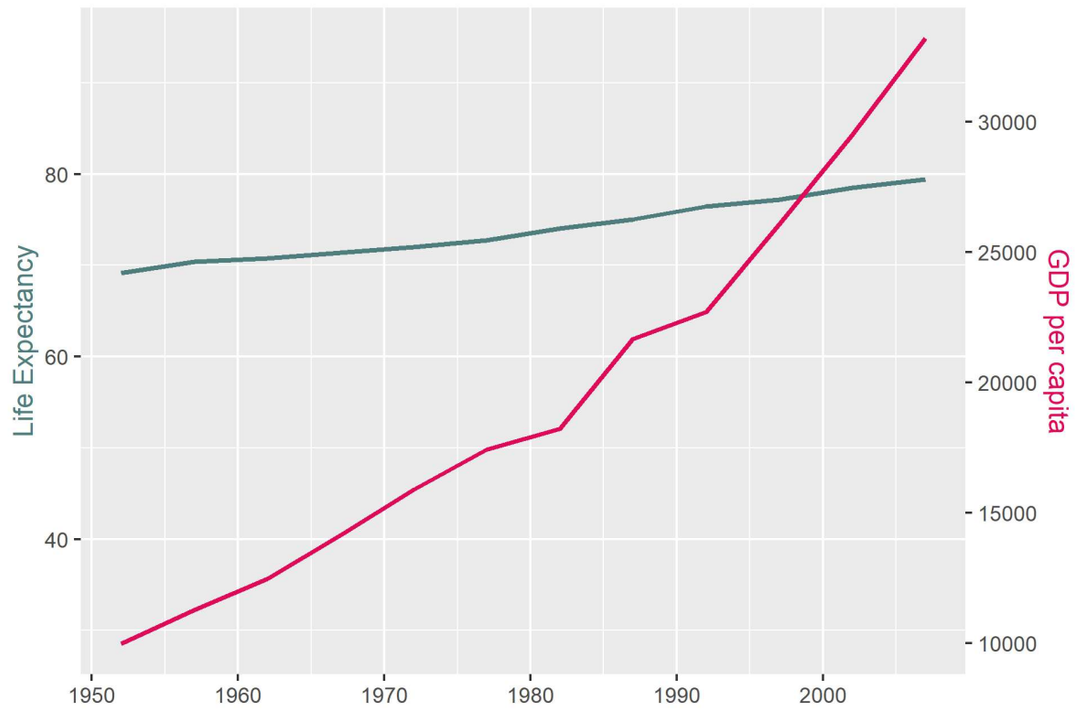
Alternatives to *spaghetti*:

- Show a smaller number of lines
  (e.g. compare a few countries to average)
- Use facets (AKA small multiples)
- Use colour only to highlight lines



Plotting multiple variables

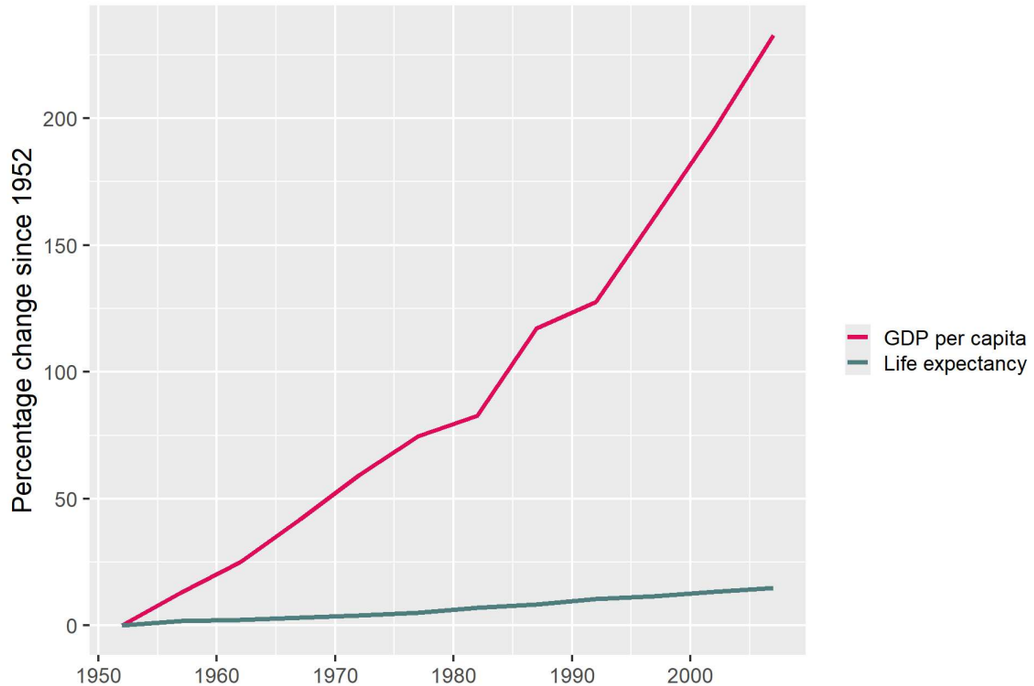




Effectively plotting multiple variables

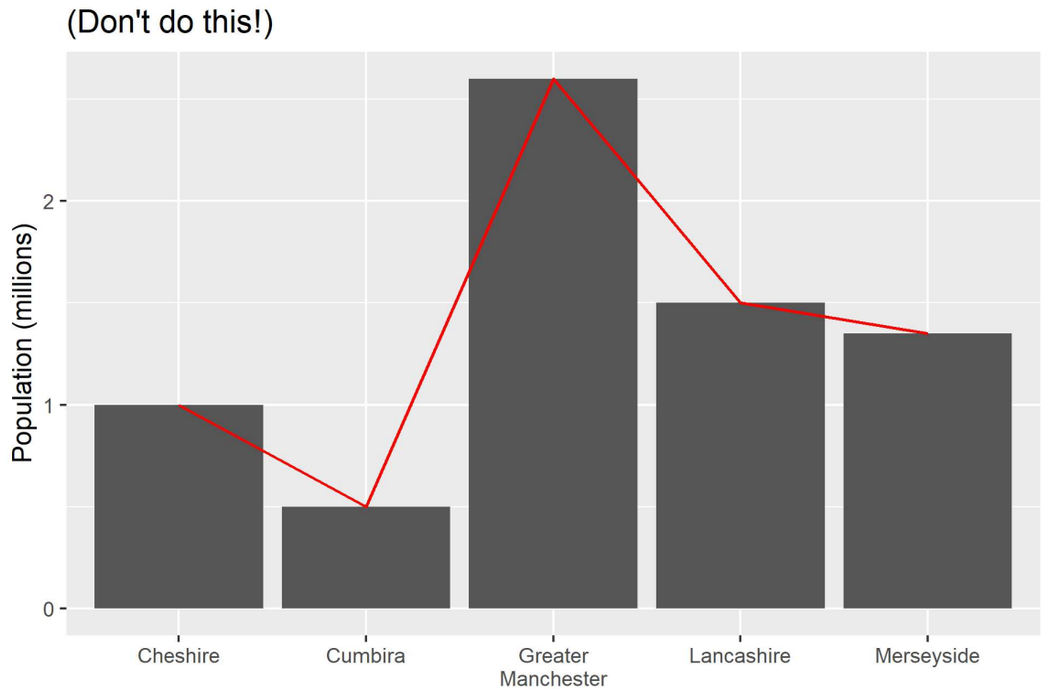
Some alternatives:

- Separate plots, each with their own axis, and place the plots side-by-side.
- Plot different variables on the x- and y-axis.
- Rescale the variables, rather than the axis.

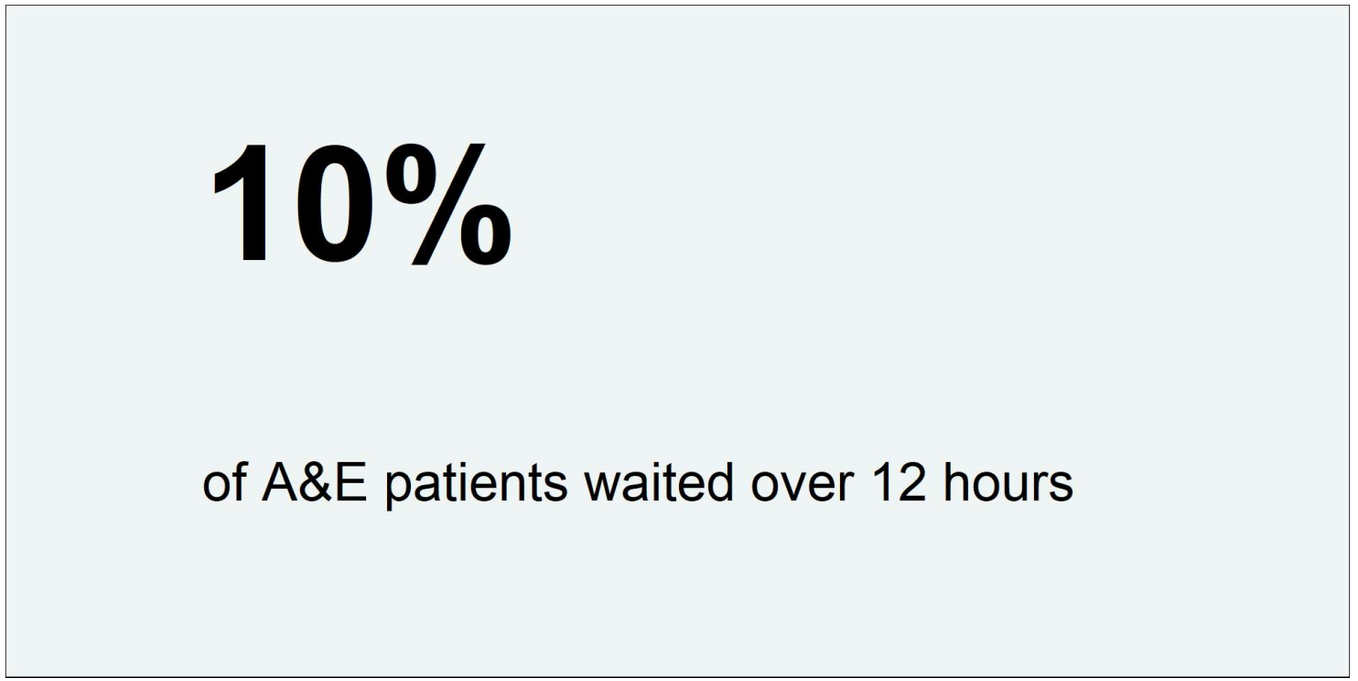


Know when lines are appropriate

- Suggest an order
- Suggest continuity



Sometimes just a number is enough...



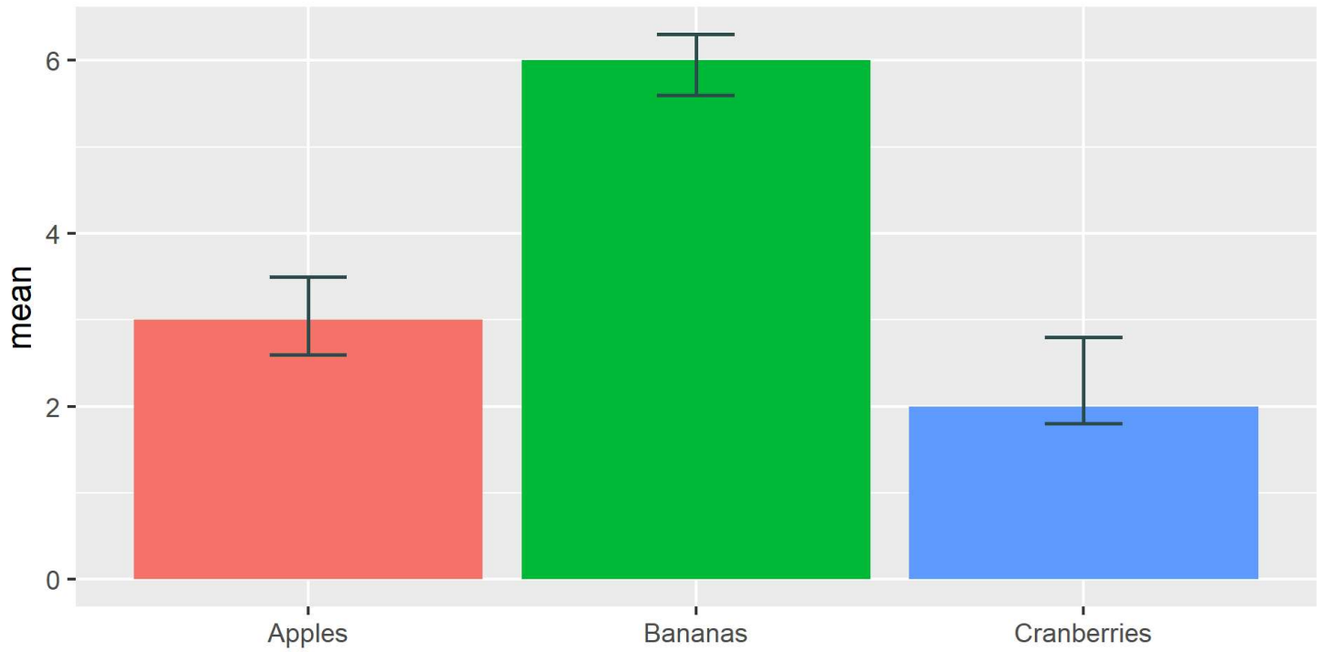
... but sometimes it's not!

> Keep it simple, but don't over simplify.

Discussion 1

03:00

- What are the problems with this type of chart?
- What could you do instead?



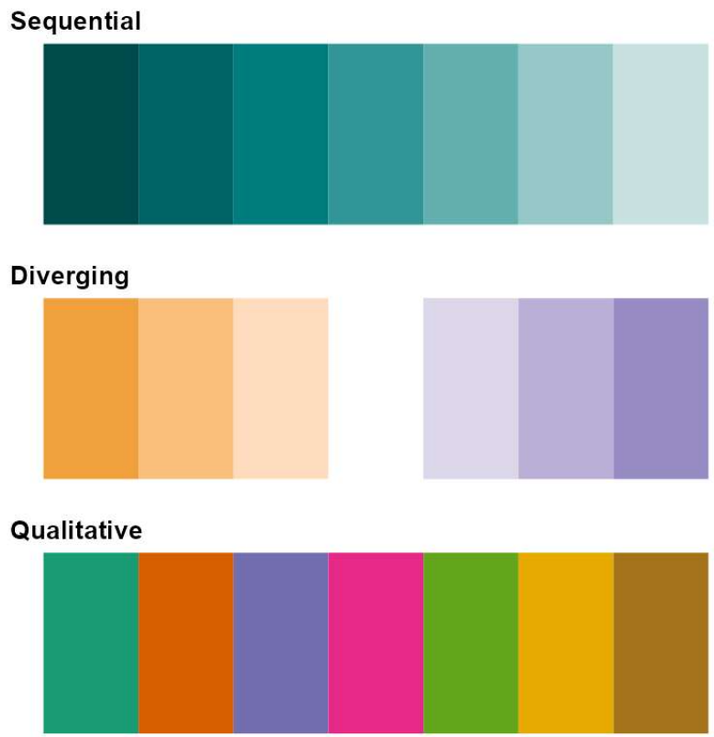
The use of colour

Colours

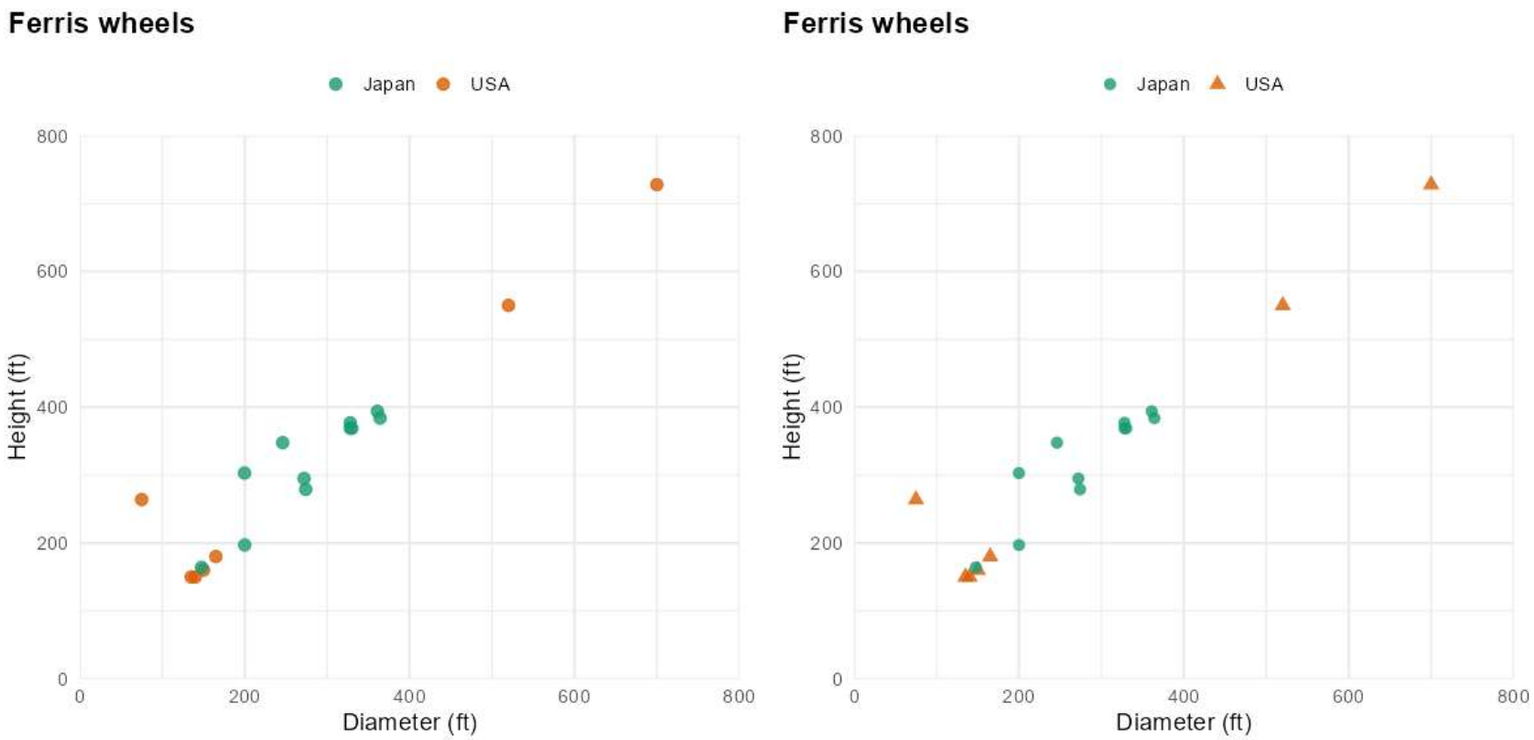
Why use colours in data visualisation?

- Colours should serve a purpose,
  e.g. discerning groups of data
- Colours can highlight or emphasise parts of your data.
- Not always the most effective for,
  e.g. communicating differences between variables.

Different types of colour palettes for different types of data.



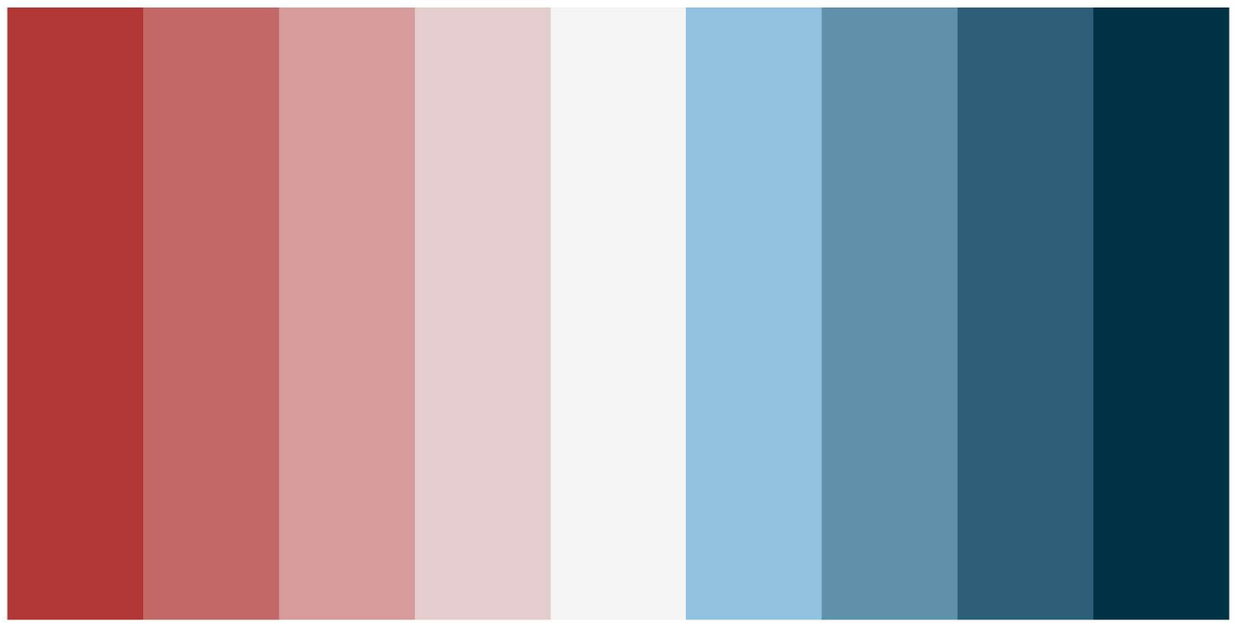
Differentiating categories with colour



Tip: In a series of charts, assign the same colour to the same variable in each chart.

Are *intuitive* colours always best?

Example: red and blue used to show hot and cold



> Tip: **never** switch to the opposite meaning!

Are *intuitive* colours always best?

Example: red and green used to show bad and good

> Tip: check the accessibility of colour palettes.

Accessible colours

Check for colourblind friendly plots with `colorblindr::cvd_grid(g)`.

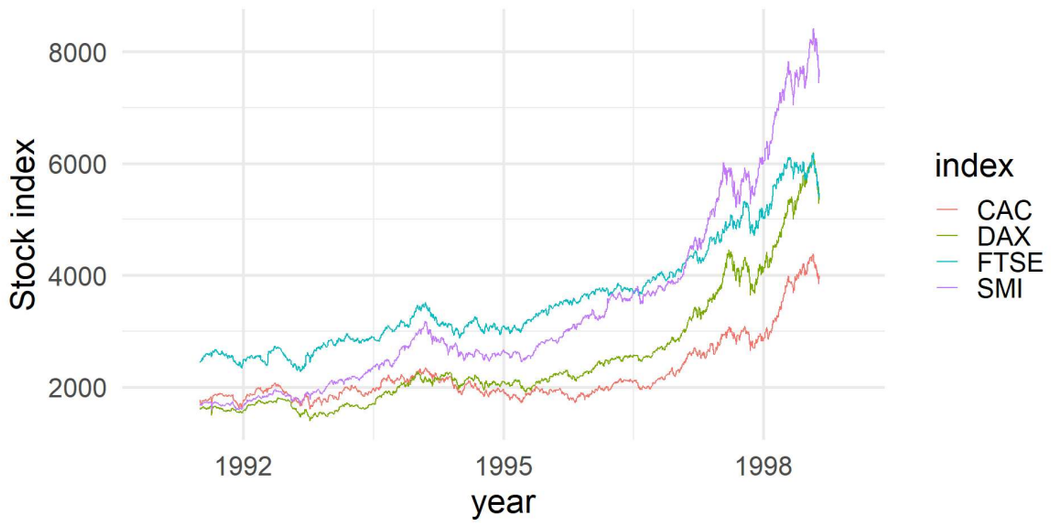
Are *intuitive* colours always best?

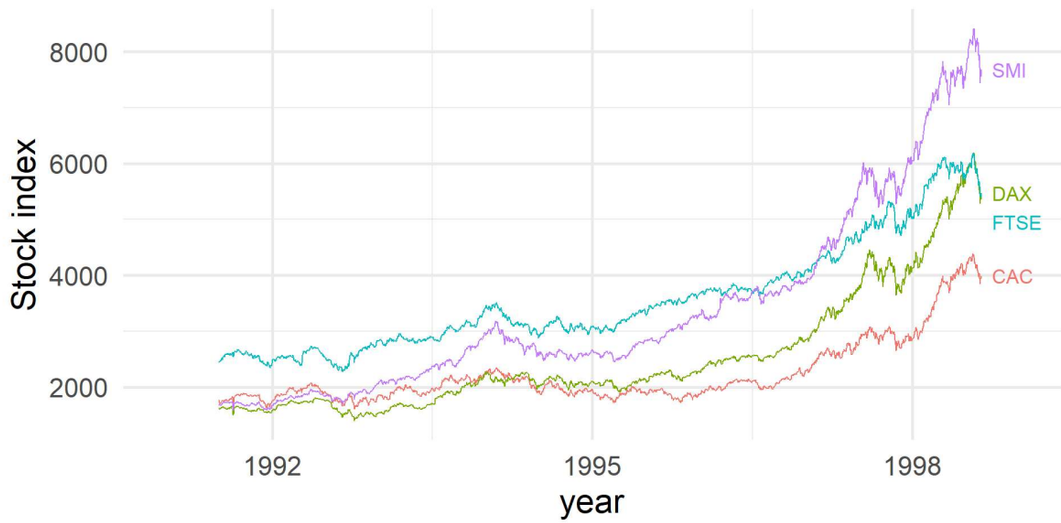
Example: pink and blue used to show women and men

> Tip: think about colour associations.

Legends

- Order and orientation of the legend matches the order and orientation of the data
- Use direct labelling.





The use of text

Why use text?

Just because you're making a picture doesn't mean you can't also use words!

- Improve clarity
- Explain how to interpret a complex chart
- Highlight interesting data points
- Summarise a message

Textual elements

Title: A concise description of the main message you want the visualisation to show.

Subtitle and/or Caption: A more specific description of the data used, and how the visualisation should be interpreted. Attribution of the data source. Any caveats about the data or analysis.

Annotations: Highlight specific points, or time periods.

Alt text: An accessible text description of the visualisation, including chart type and the main message.

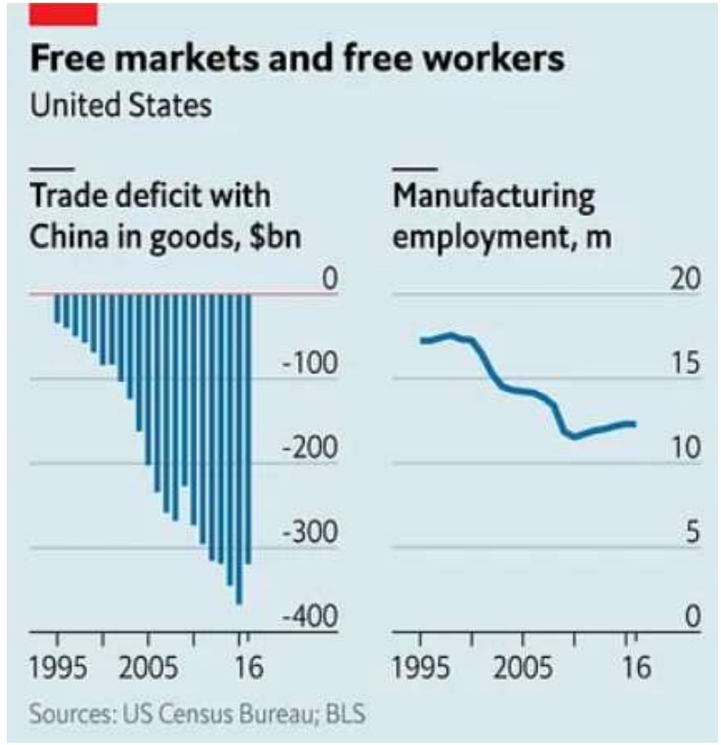
Choosing a font

- Font size: larger fonts are (usually) better
- Font colour: ensure sufficient contrast
- Font face: highlight text using bold font, avoid italics
- Font family: choose a clear font with distinguishable features. Does a `1` look different from a lowercase `l`, and from an uppercase `I`?

Discussion 2

03:00

In groups, discuss the following chart. What is good and bad about it?



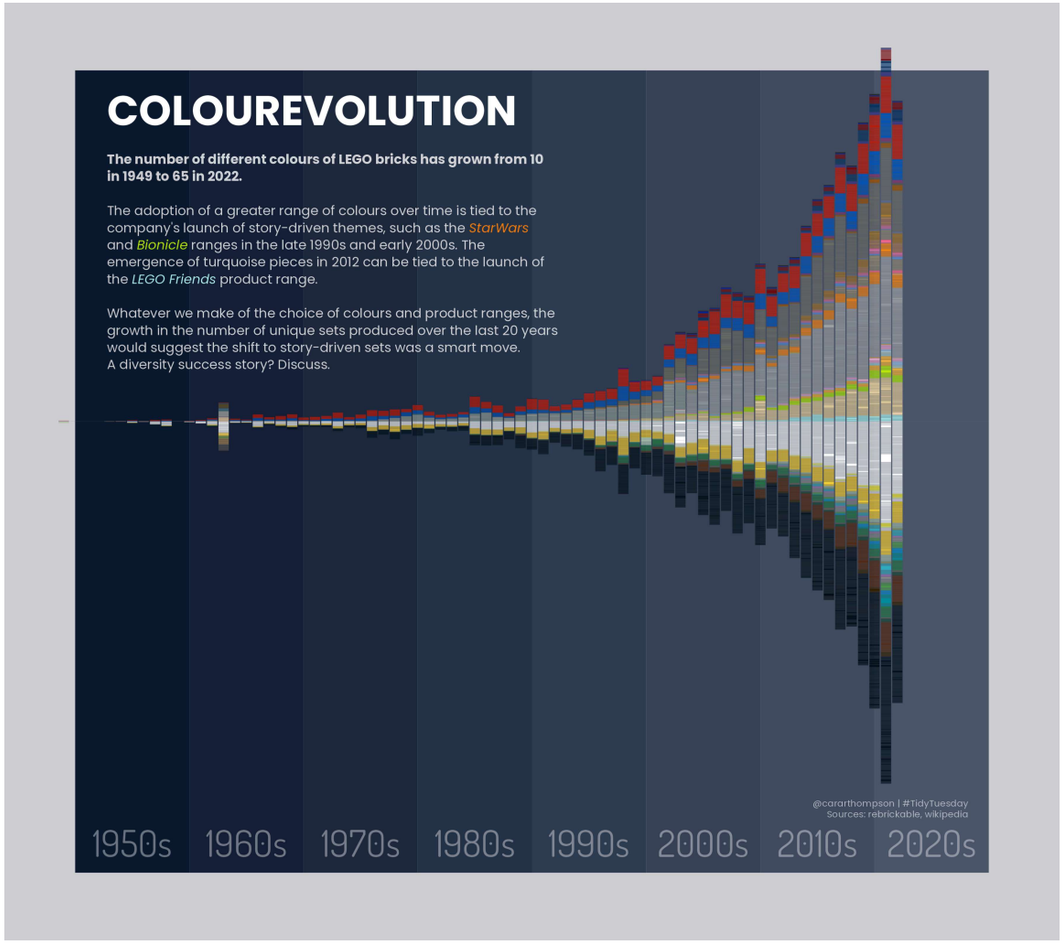
Source: medium.economist.com/mistakes-weve-drawn-a-few-8cdd8a42d368

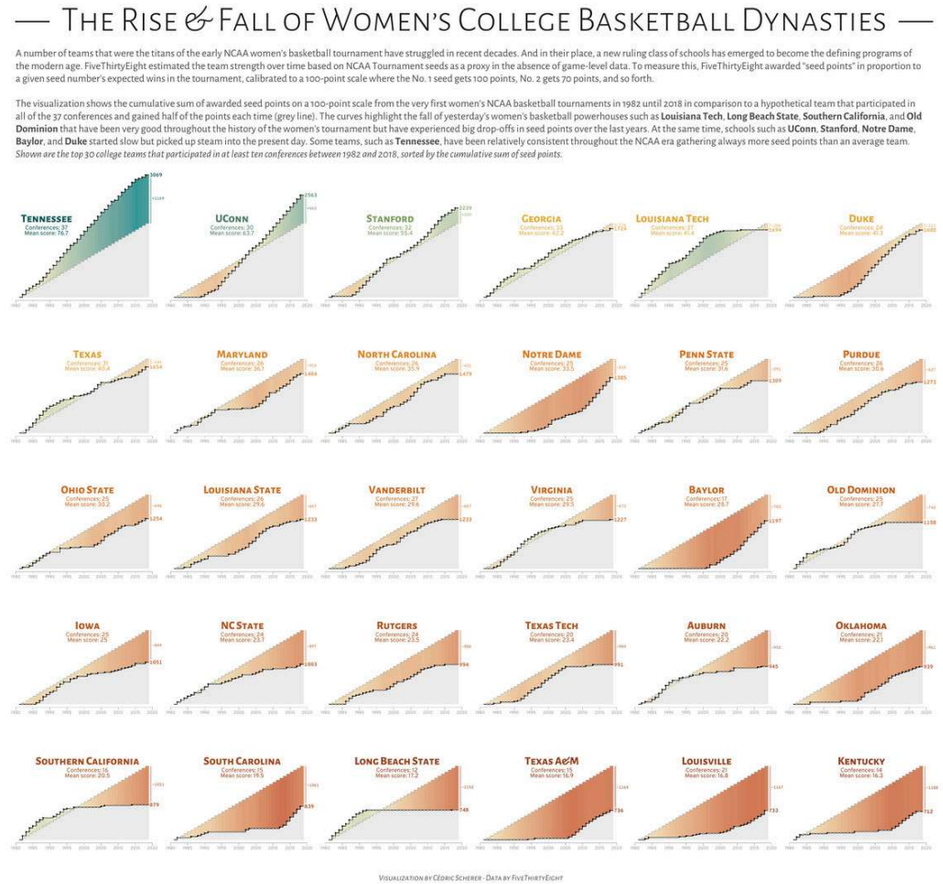
Data visualisation in practice

Key points

- Think about your audience
- Actively design visualisations with that audience in mind
- Not all *popular* chart types are the best choice

Good charts don’t have to be boring!



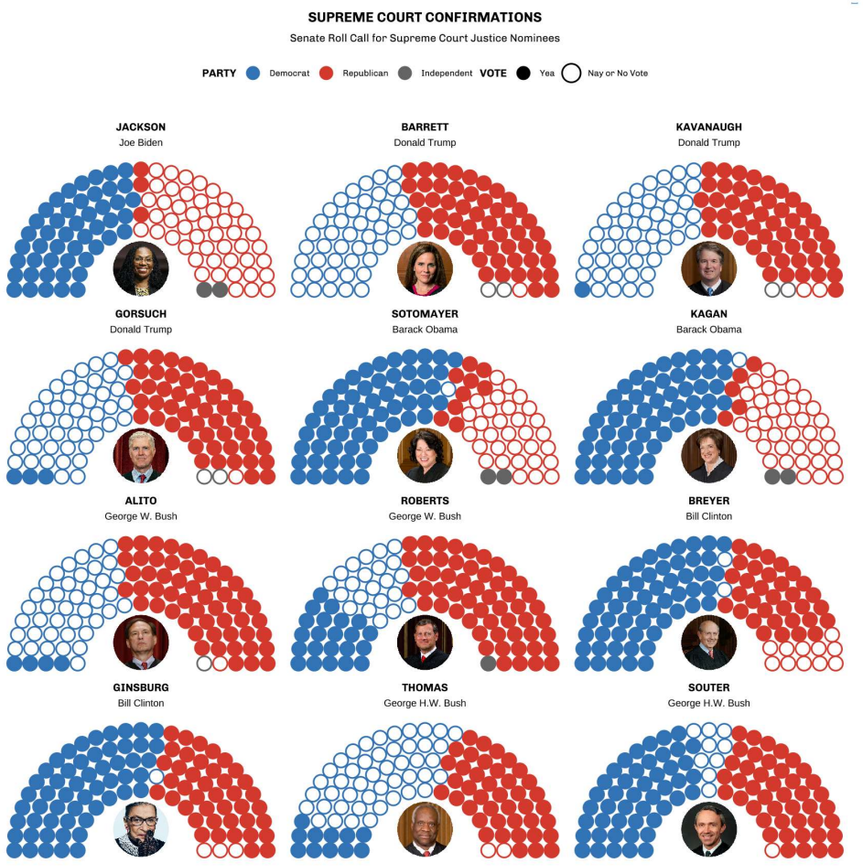


Cara Thompson (cararthompson.com)

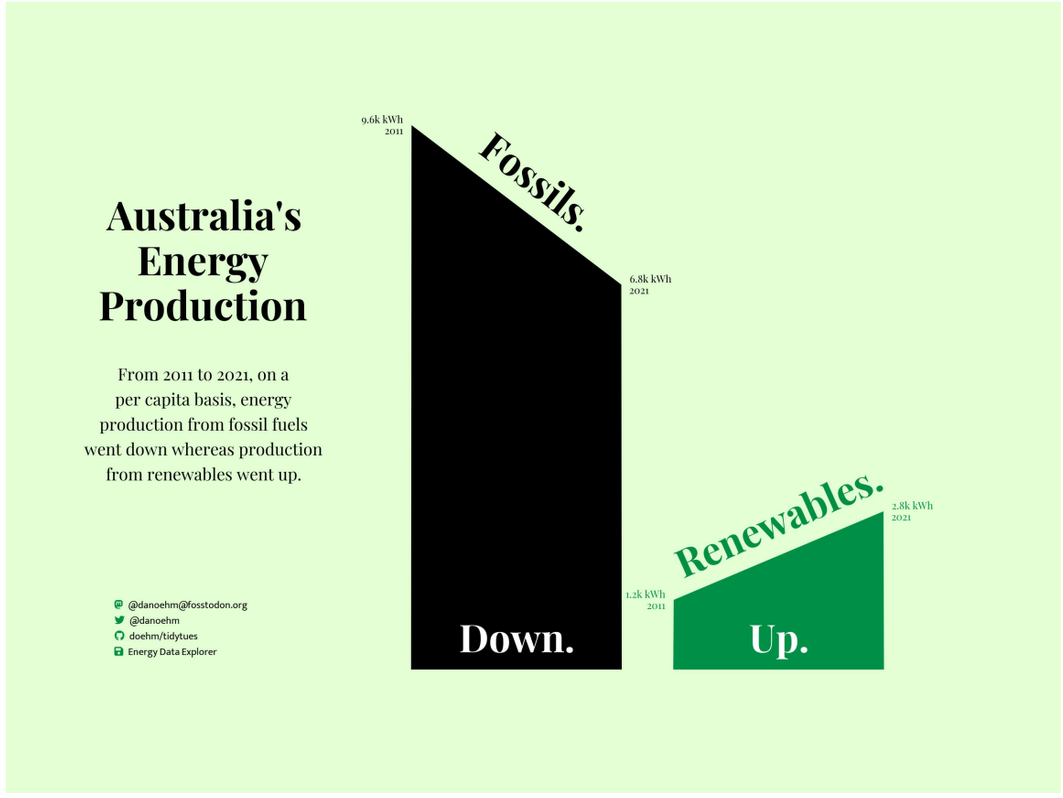
Cedric Scherer (cedricscherer.com)

nrennie.rbind.io/training-data-visualisation

Good charts don’t have to be boring!



Tanya Shapiro (tanyaviz.com)

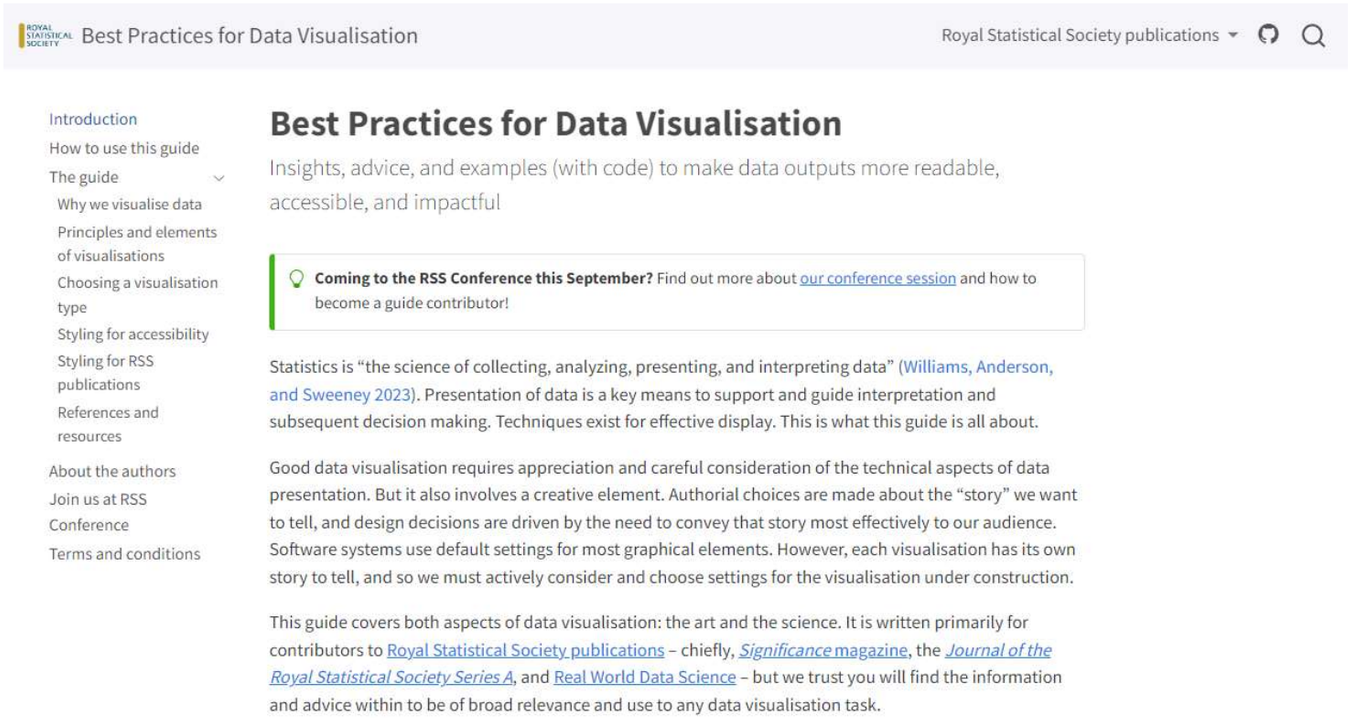


Dan Oehm (gradientdescending.com)

Resources

RSS Data Visualisation Guide

royal-statistical-society.github.io/datavisguide



Data visualisation resources

Fundamentals of Data Visualization: clauswilke.com/dataviz

Civil Service Guidance: analysisfunction.civilservice.gov.uk/policy-store/data-visualisation-charts/

DataWrapper Blog: blog.datawrapper.de

Financial Times Visual Vocabulary: ft-interactive.github.io/visual-vocabulary

nrennie.rbind.io

nicola-rennie

nrennie