

Go Figure!

RSS Pre-Conference Workshop 2023

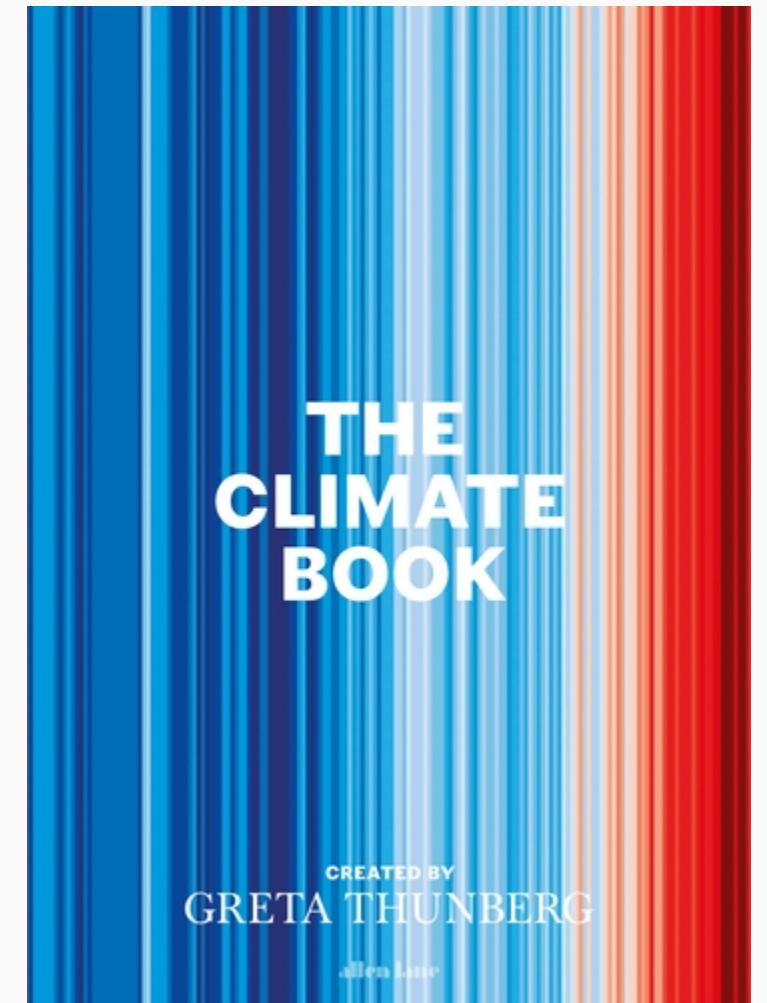
Zak Varty

More Than A Pretty Picture

Visualisation is a key skill in your data science tool kit:

- Rapidly explore data sets
- Model evaluation and diagnostics
- Sharing evidence
- Telling compelling stories.

Reflective exercise, not a tutorial or rulebook.

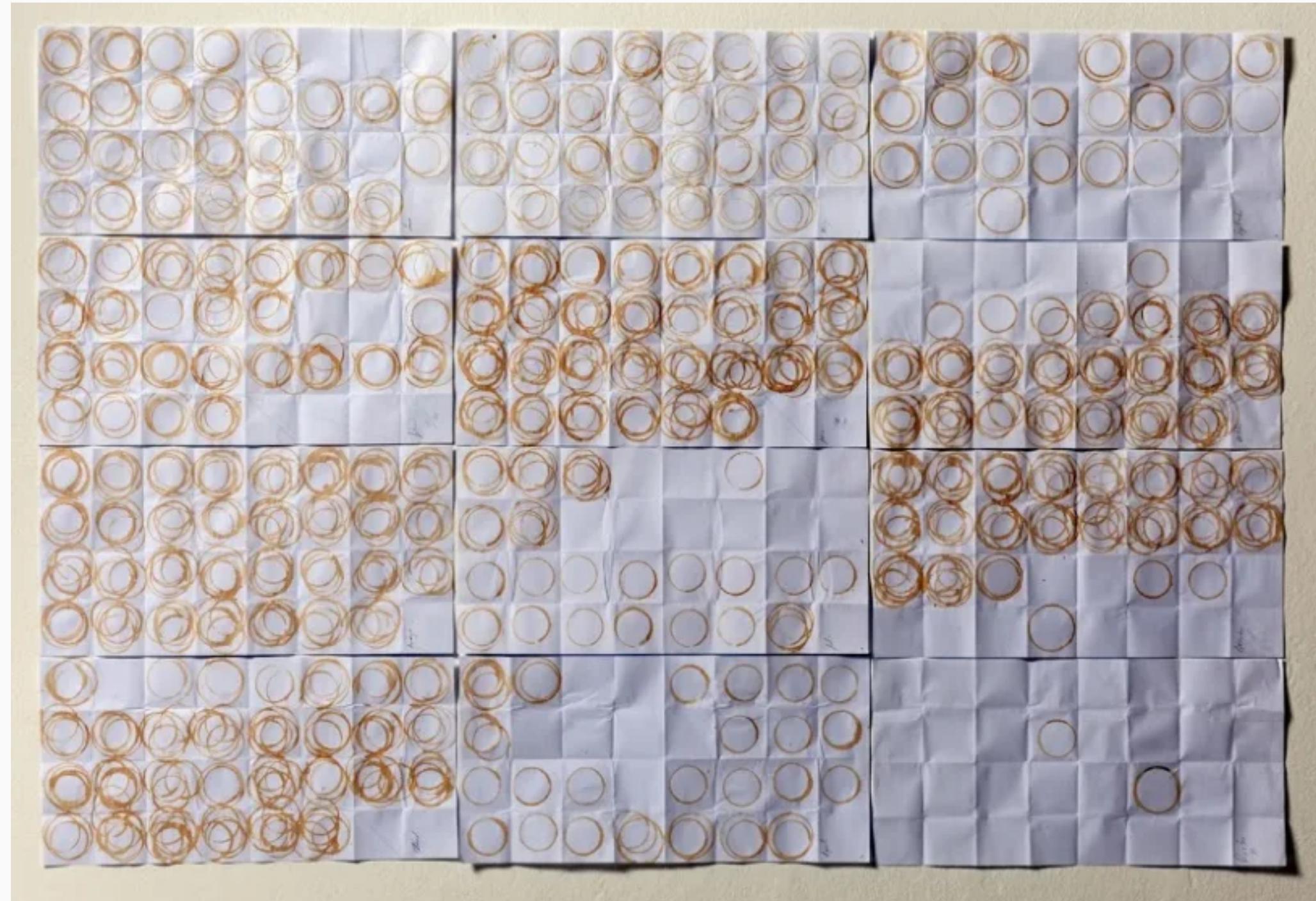


Warming stripes graphic on the cover
of "The Climate Book"

1: Think About Your Tools

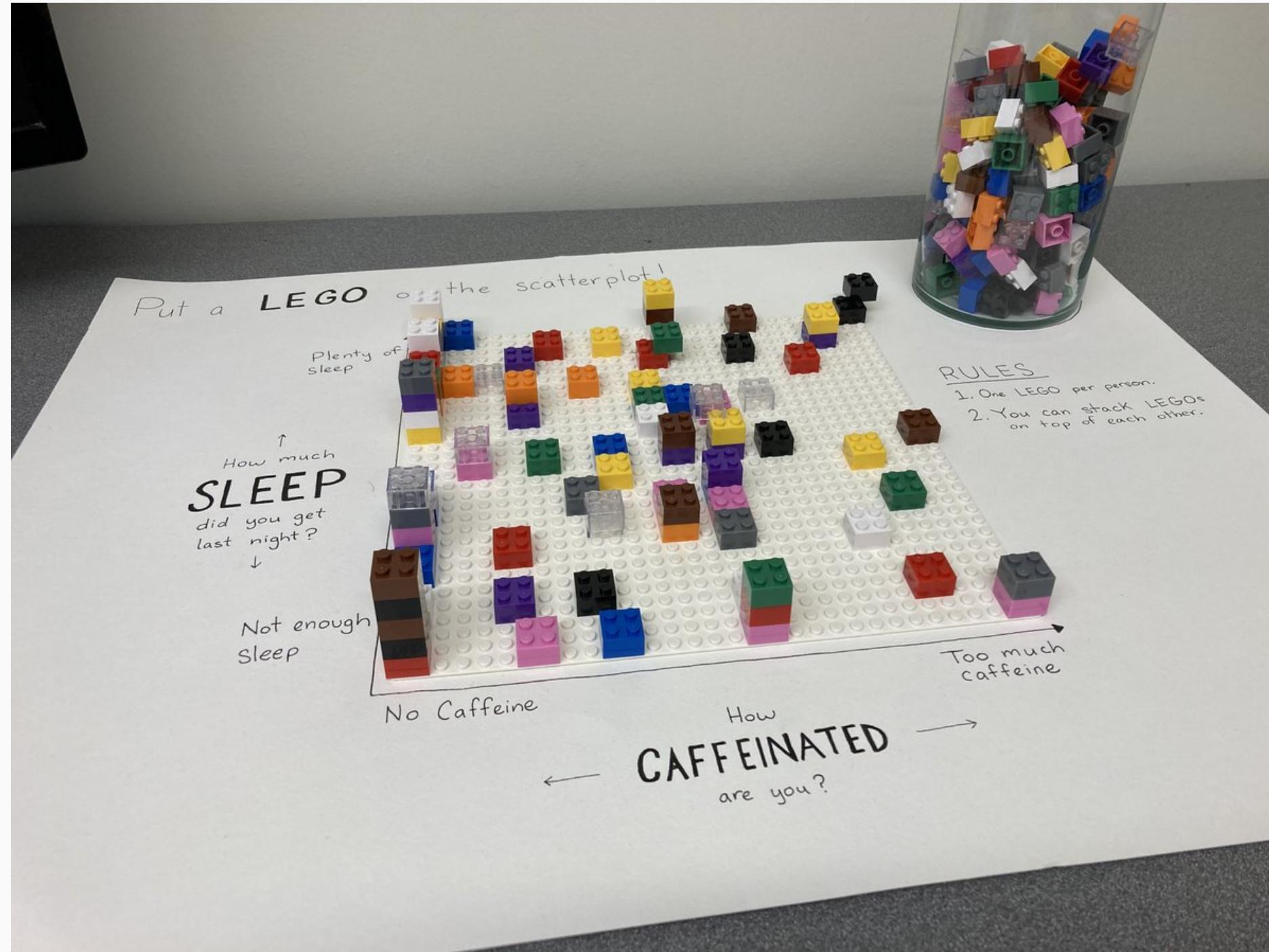


Data visualisation tools



Coffee consumption, visualised. Jaime Serra Palou.

Selecting your tools: Analogue or Digital



Caffeination vs sleep, shown in lego. Elsie Lee-Robbins

Staying in the tidyverse: {ggplot2}

Layered creation of graphics
from tidy data.

Learning {ggplot2}:

- Resources
- Tutorial
- Tidy Tuesday Github

A Layered Grammar of Graphics

Hadley WICKHAM

A grammar of graphics is a tool that enables us to concisely describe the components of a graphic. Such a grammar allows us to move beyond named graphics (e.g., the “scatterplot”) and gain insight into the deep structure that underlies statistical graphics. This article builds on Wilkinson, Anand, and Grossman (2005), describing extensions and refinements developed while building an open source implementation of the grammar of graphics for R, `ggplot2`.

The topics in this article include an introduction to the grammar by working through the process of creating a plot, and discussing the components that we need. The grammar is then presented formally and compared to Wilkinson’s grammar, highlighting the hierarchy of defaults, and the implications of embedding a graphical grammar into a programming language. The power of the grammar is illustrated with a selection of examples that explore different components and their interactions, in more detail. The article concludes by discussing some perceptual issues, and thinking about how we can build on the grammar to learn how to create graphical “poems.”

Supplemental materials are available online.

Key Words: Grammar of graphics; Statistical graphics.

2: Think About Your Medium

Where will your plot go?

Use cases: exploratory analysis, presentation, report / paper, data journalism.

Considerations:

- Time investment vs quality
- Image size / format
- Time spent with graphic



Presentations



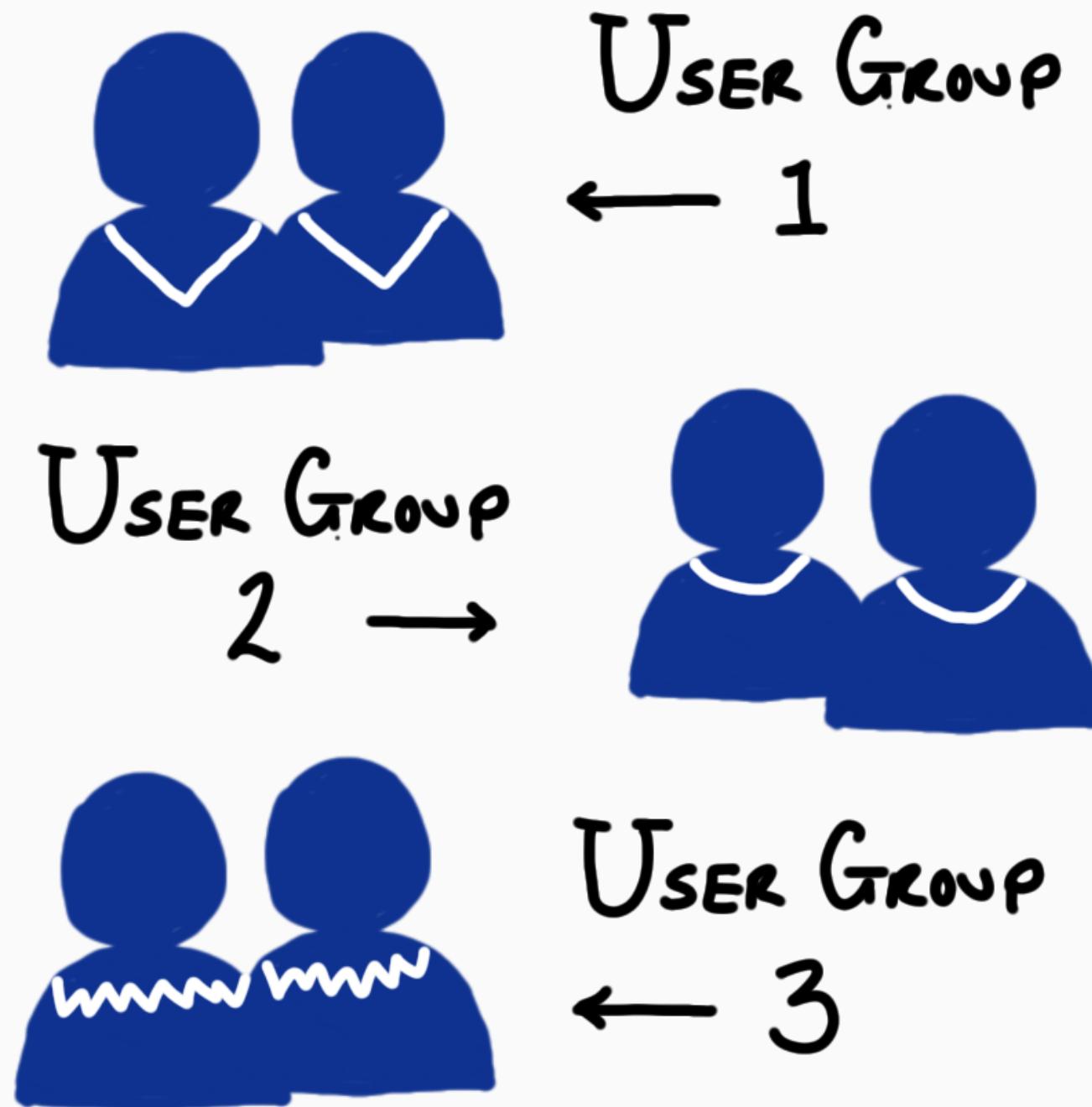
File types

- *Bitmap Graphics:* png, jpeg, gif
 - made of pixels
 - smaller file size
- *Vector Graphics:* pdf, eps, svg
 - made of vectors
 - larger file size



3: Think About Your Audience

Know your audience



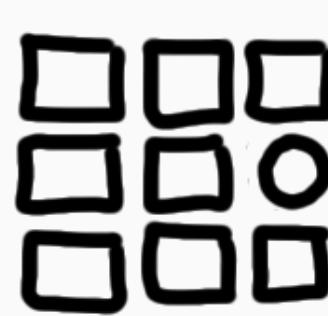
Who is the intended audience for your visualisation?

What knowledge do they bring with them?

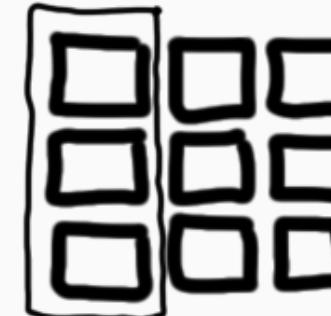
What assumptions and biases do they hold?

Creating *personas* for distinct user groups can be helpful.

Preattentive Attributes



Shape



Enclosure



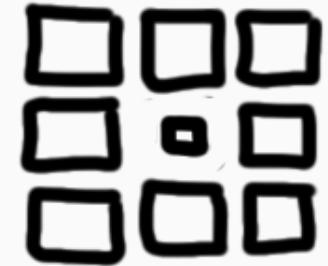
Weight



Saturation



Colour



Size



Marking



Orientation



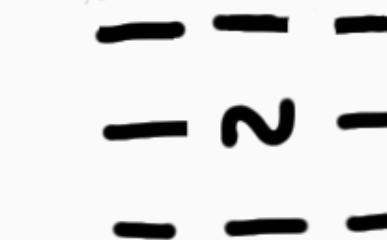
position



density

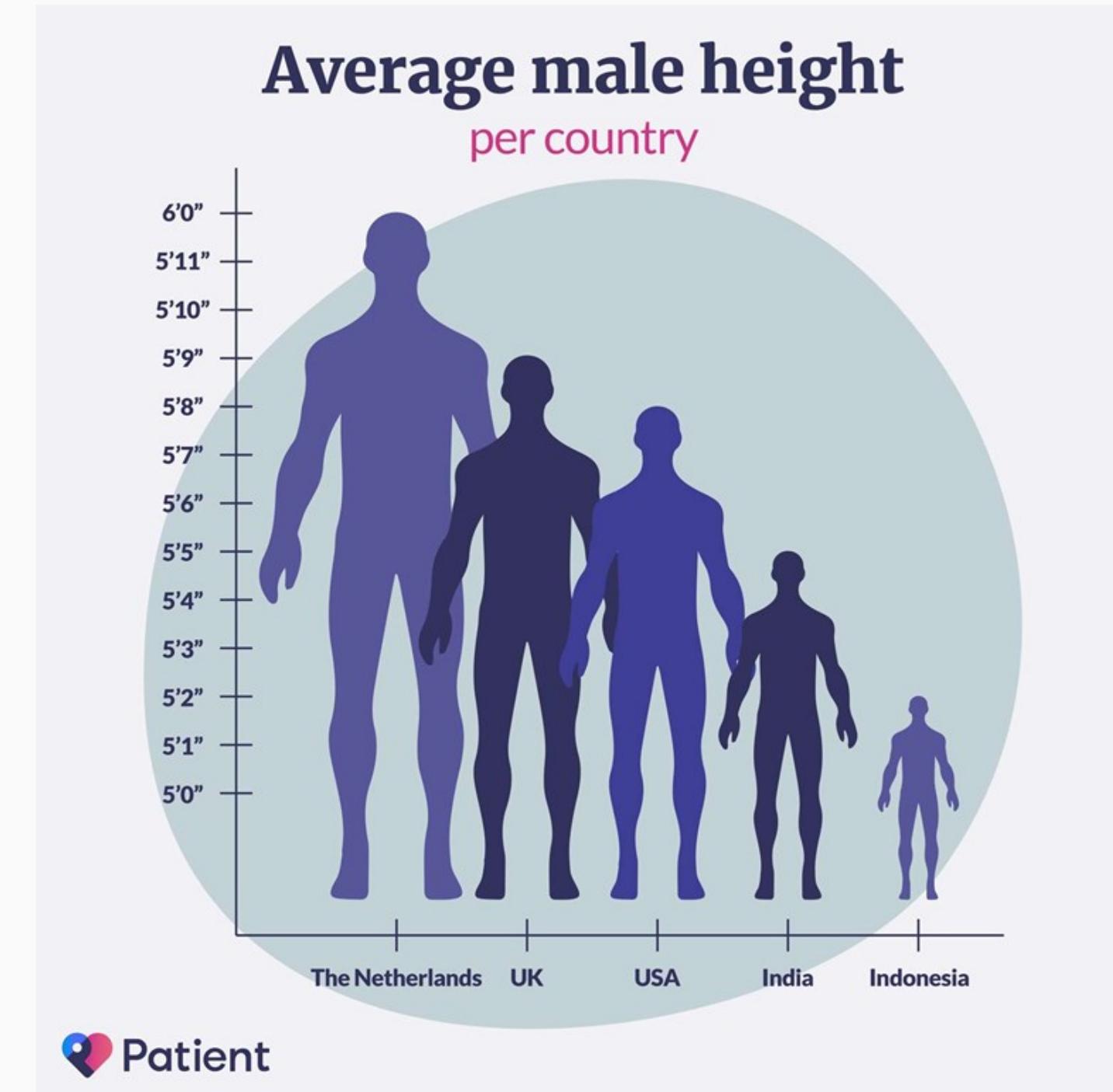


length



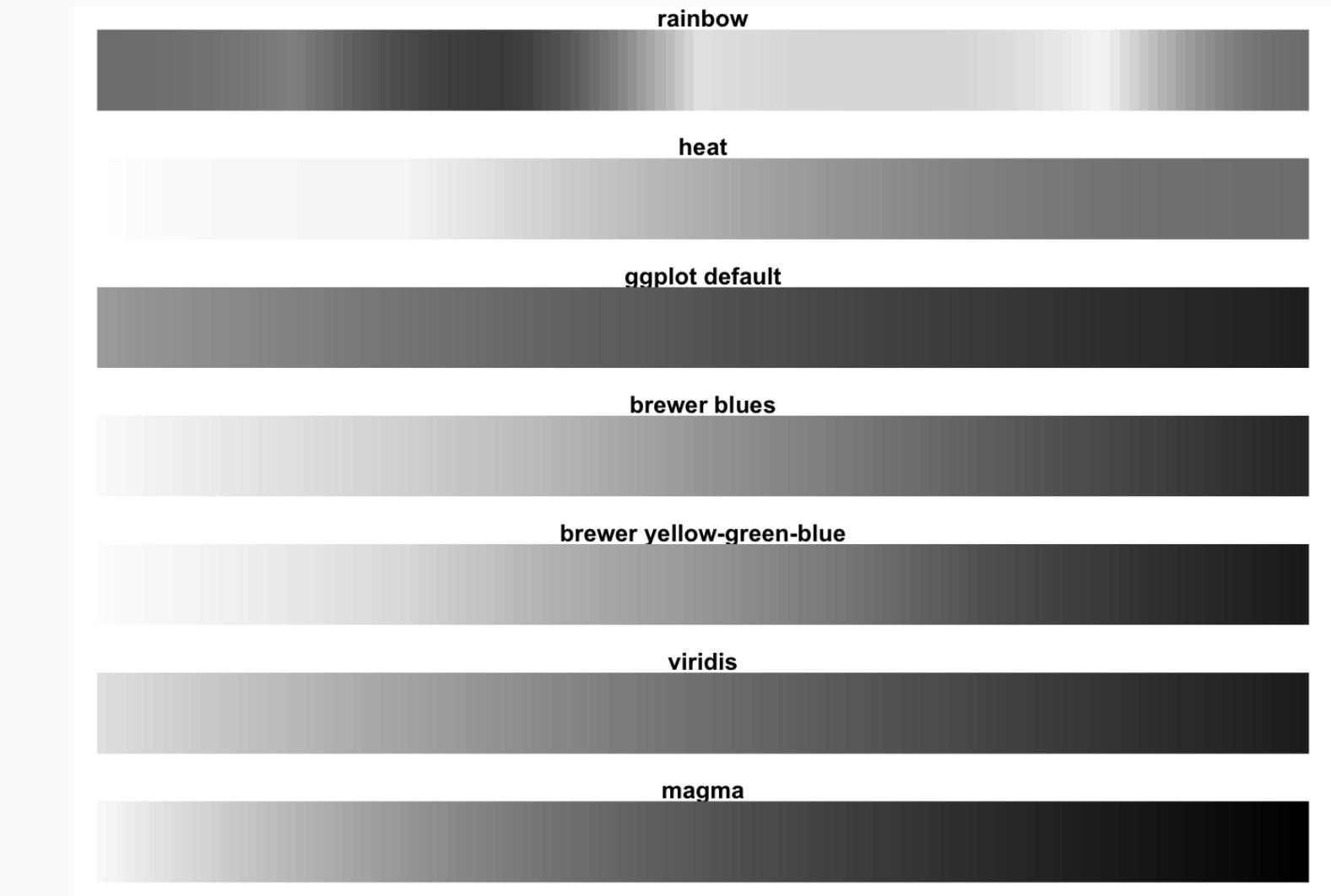
Curvature

First impressions count



Issues with scales, area and perspective

Visual perception



Alt text, titles and captions

Captions

Describes a figure or table so that it may be identified in a list of figures and (where appropriate).

Alternative text

Describes the content of an image for a person who cannot view it. ([Guide to writing alt-text](#))

Titles

Give additional context or identify key findings. Active titles are preferable.

~~Graph to show how X varies with Y~~

Adding Alt-Text to Images

HTML

```
1 
```

pdf (LaTeX)

```
1 \usepackage[tagged, highstructure]{accessibility}
2
3 \begin{figure}
4   \centering
5   \includegraphics[]{path/to/image.png}
6   \alt{Alt text goes here.}
7   \caption{Caption goes here.}
8   \label{fig:alt-text-example}
9 \end{figure}
```

Adding Alt-Text to Images

Markdown

Great for multi-output documents but many flavours.

Github / Jupyter:

```
1 ![Alt text goes here.](path/to/image.png)
```

Quarto:

```
1 ![Caption goes here](path/to/image.png){alt="Alt text goes here"}
```

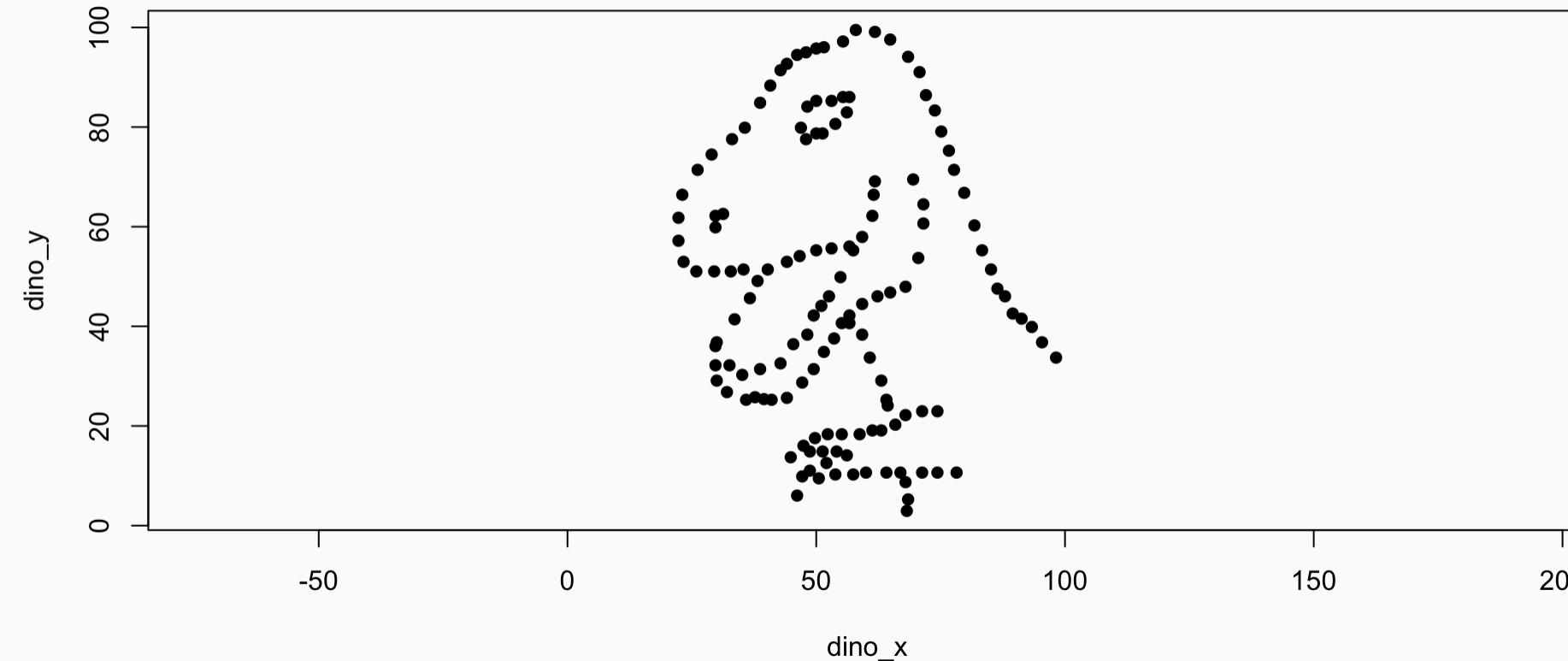
Adding Alt-Text to Plots

When using literate programming alt-text can be added as code block meta-data.

In Quarto:

```
1 ````{r make-dino-plot}
2 #| fig-cap: The Datasaurus from the Dino-Dozen.
3 #| fig-alt: Scatterplot in which the points form the outline of a T-Rex.
4 #| output-location: slide
5 par(bg = NA)
6 plot(x = dino_x, y = dino_y, pch = 16, asp = 1)
7 ````
```

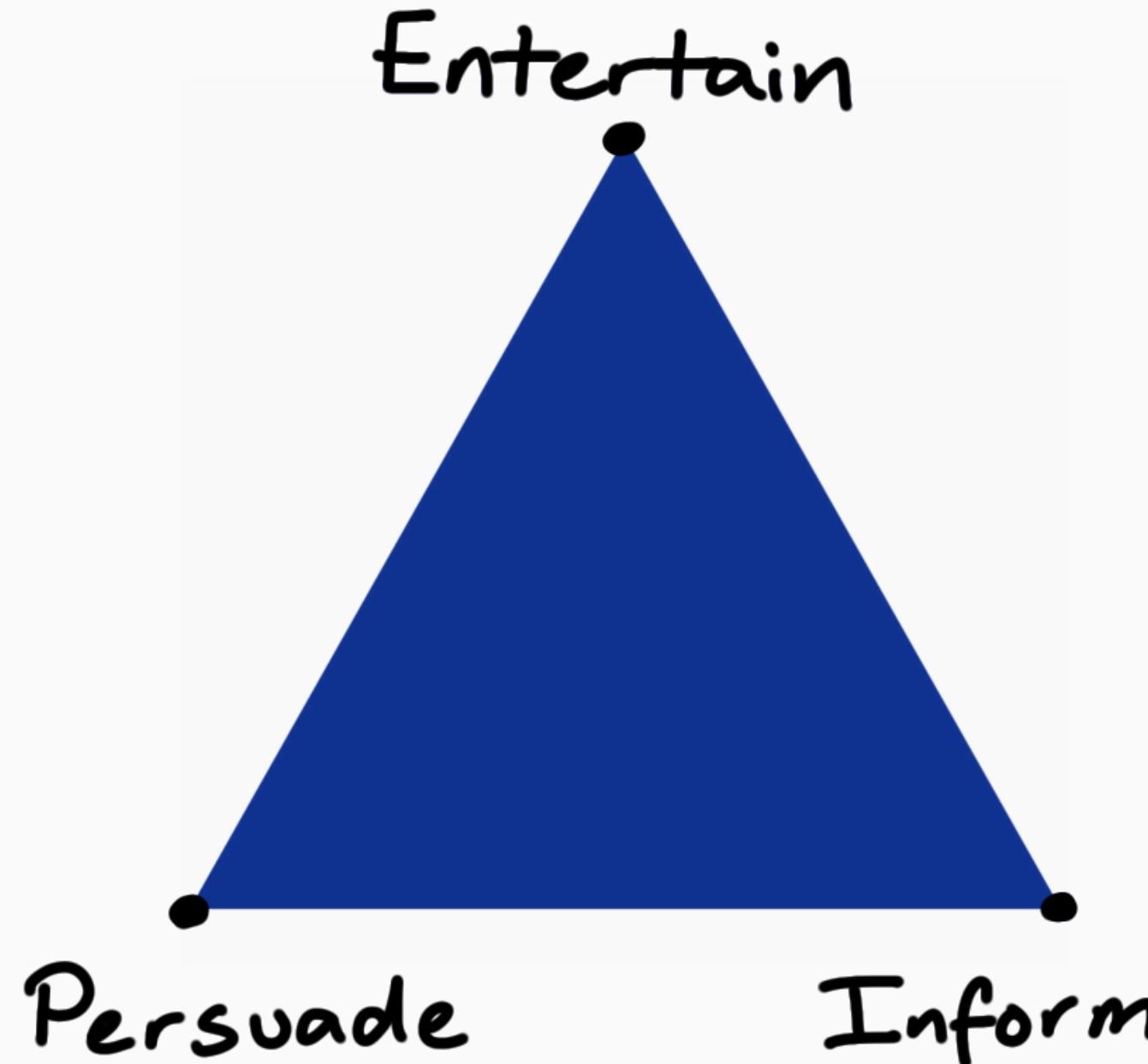
Adding Alt-Text to Plots



The Datasaurus from the Dino-Dozen.

4: Think About Your Story

Data visualisation as storytelling



- Where does your purpose fall on this triangle?
- No such thing as neutral presentation.
- Start with a hook.

5: Think About Your Guidelines

Standardise and document it

Decisions cost time, energy and money. (DRY)

Consider your design choices carefully and write down your decisions and reasoning.
(DRY)

This will form the basis of your own *style-guide for data visualisation*.

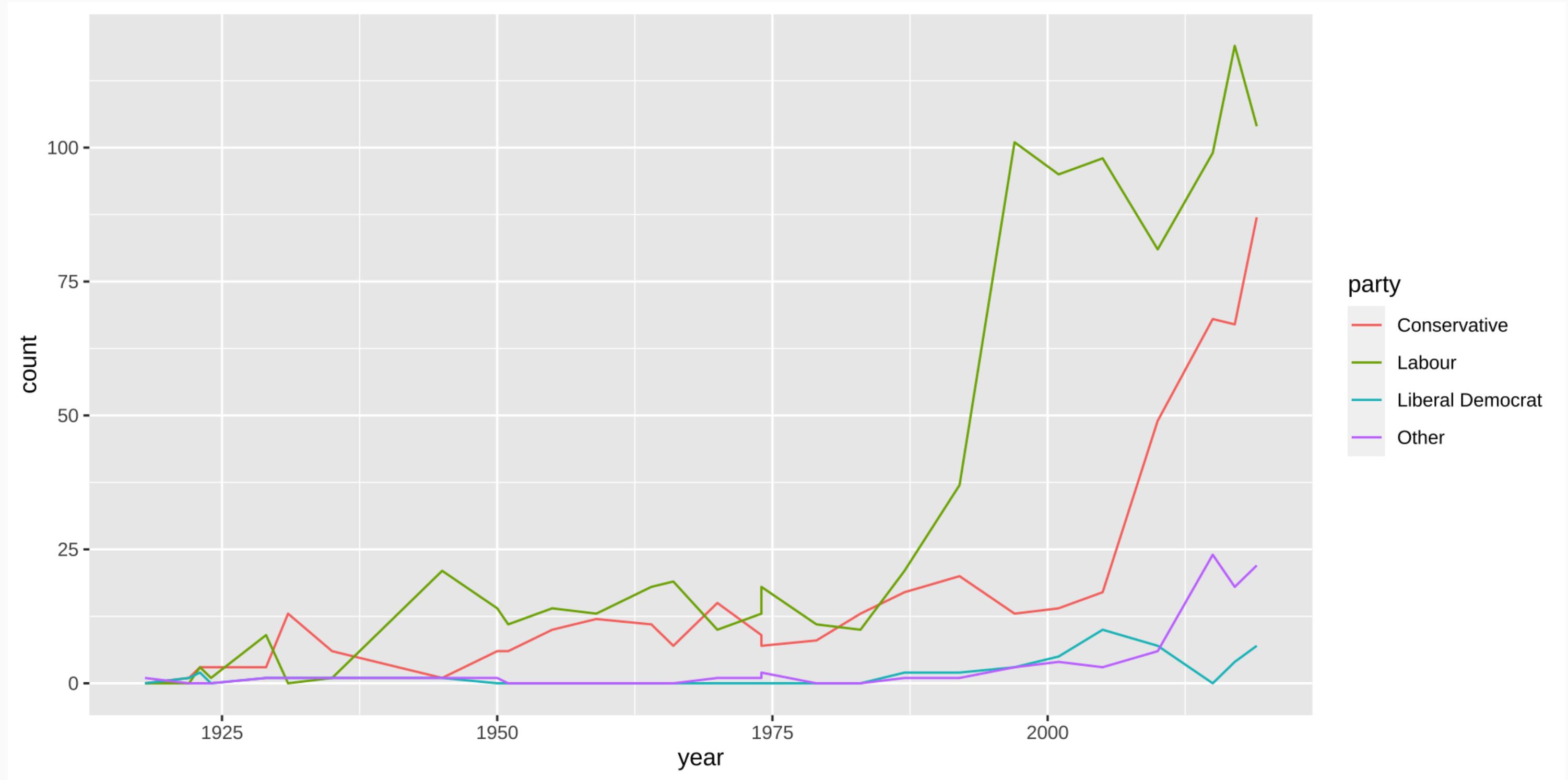
Style guides for data visualisation

- BBC
 - Infographics Guidelines
 - R Cookbook
 - {bbplot}
- The Economist
 - The Office for National Statistics
 - Eurostat
 - Urban Institute
 - The Pudding (learning resources)

A Worked example

- Quarto, R and `{ggplot}`.
- Blog post: women in politics.
- General audience, familiar with UK politics.
- Representation of women in parliament is improving over time.
- Style guidelines of blog and political parties.

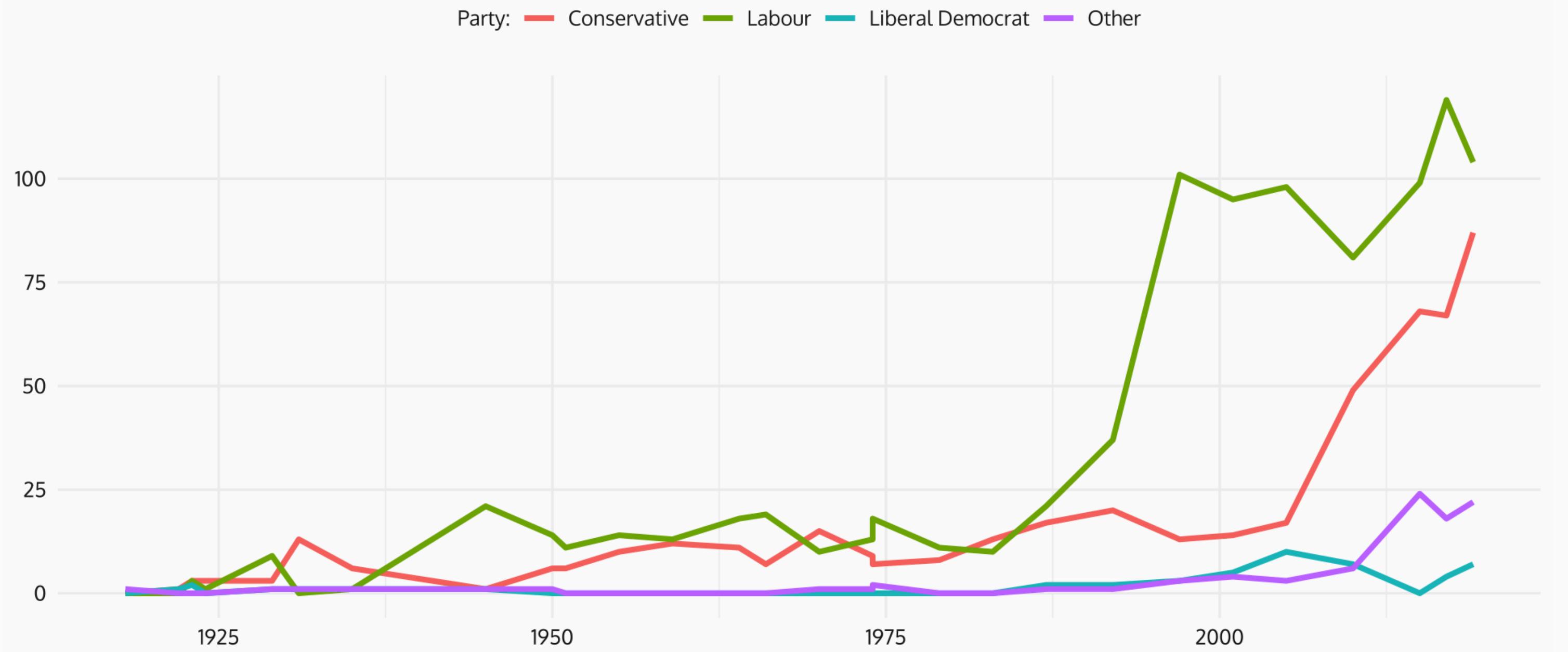
Out of the Box: A First Attempt



Applying Blog Styling

Number of Seats Won by Women in UK General Elections

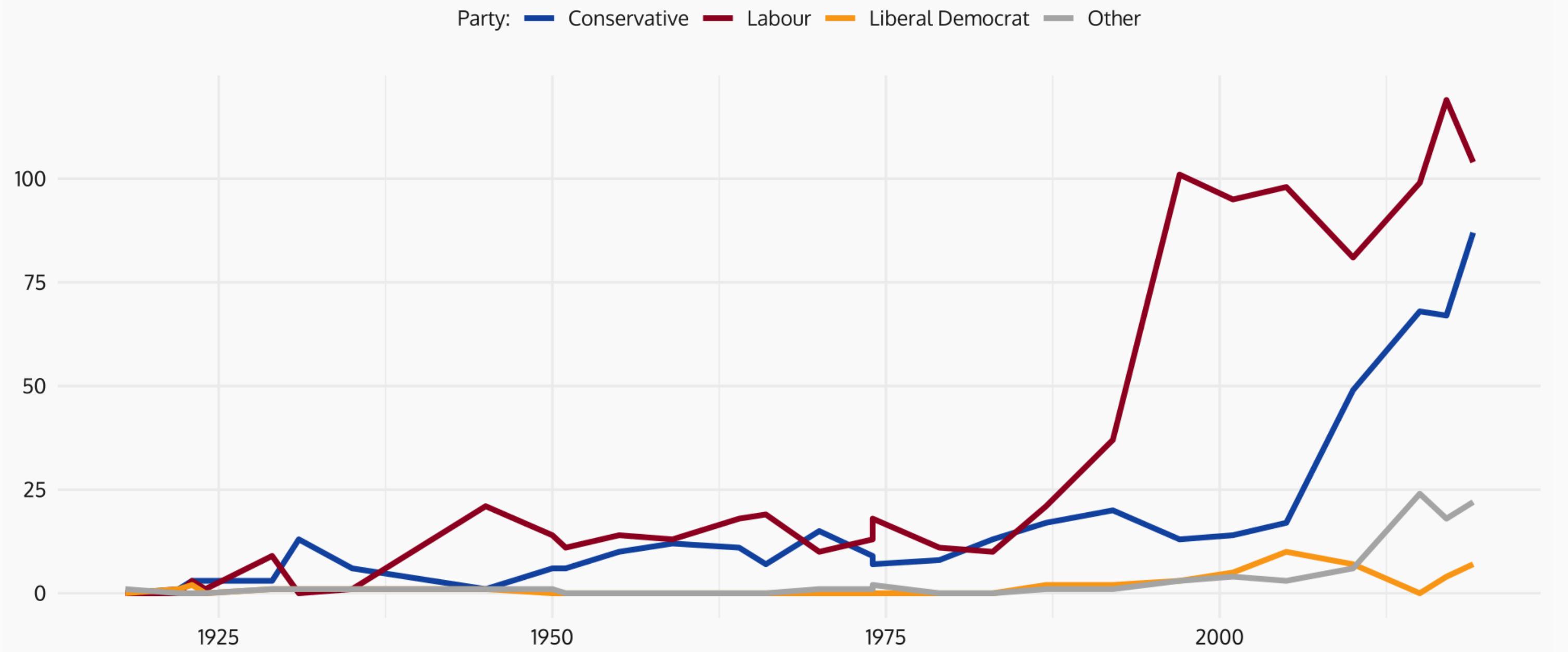
Data: House of Commons Library



Fix Confusing Colours

Number of Seats Won by Women in UK General Elections

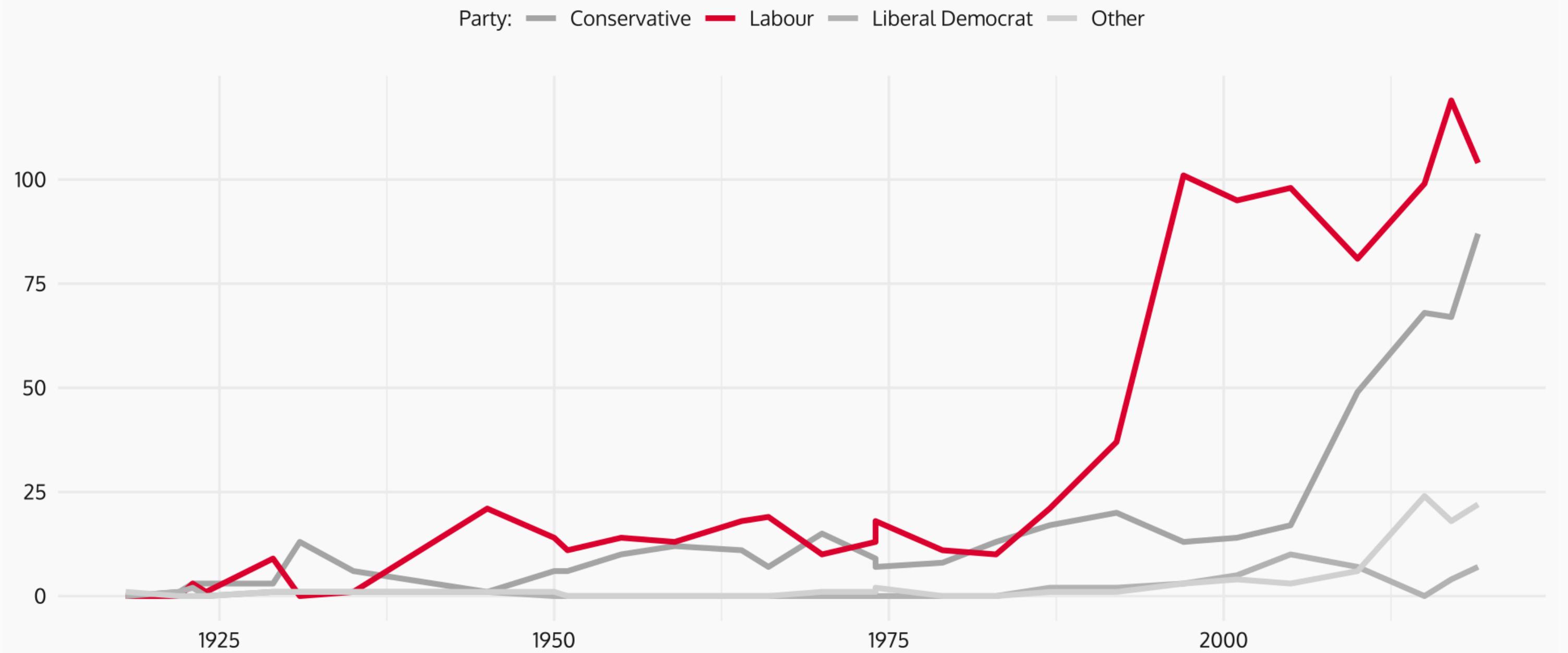
Data: House of Commons Library



Highlight important takeaway

Number of Seats Won by Women in UK General Elections

Data: House of Commons Library



Wrapping up

-  Think about your *tools*
-  Think about your *medium*
-  Think about your *audience*
-  Think about your *story*
-  Think about your *guidelines*

Image Credits

- The Climate Book - [Penguin](#)
- Coffee Cup - [Jaime Serra Palou](#)
- Lego coffee - [Elsie Lee-Robbins](#) via Twitter
- Pre-attentive attributes - Adapted from [Better Data Visualizations](#)
- Male Heights - [patient.info](#)
- Desaturated colour scales - [{viridis}](#) documentation

Further Resources

RSS Best Practices for Data Visualisation

How to make data outputs more readable, accessible, and impactful.

11:40-13:00 Tuesday, 5 September, 2023, Auditorium

- Reading list
- Jupyter Accessibility Workshop
- SuperCool blog post

Activity!

Exercise 1: Text to Plot

Split into 5 groups and spread around the room.

Try to draw the plot based on the alt-text provided.

Exercise 2: Plot to Text

Now for the inverse problem:

Write your own alt-text based on the plot provided.

Exercise 3: Text to plot

Put your alt-text to the test:

Pass your alt-text to another group. They have to try and recreate your plot!

Build Information

R version 4.2.2 (2022-10-31)

Platform: x86_64-apple-darwin17.0 (64-bit)

locale: en_US.UTF-8||en_US.UTF-8||en_US.UTF-8||C||en_US.UTF-8||en_US.UTF-8

attached base packages: *stats*, *graphics*, *grDevices*, *utils*, *datasets*, *methods* and *base*

other attached packages: *dplyr*(v.1.1.2), *ggplot2*(v.3.4.0) and *datasauRus*(v.0.1.6)

loaded via a namespace (and not attached): *Rcpp*(v.1.0.9), *cellranger*(v.1.1.0), *compiler*(v.4.2.2), *pillar*(v.1.9.0), *sysfonts*(v.0.8.8), *tools*(v.4.2.2), *digest*(v.0.6.31), *jsonlite*(v.1.8.4), *evaluate*(v.0.20), *lifecycle*(v.1.0.3), *tibble*(v.3.2.1), *gttable*(v.0.3.1), *pkgconfig*(v.2.0.3), *png*(v.0.1-7), *rlang*(v.1.1.0), *cli*(v.3.6.0), *rstudioapi*(v.0.14), *yaml*(v.2.3.6), *countdown*(v.0.4.0), *xfun*(v.0.36), *fastmap*(v.1.1.0), *showtextdb*(v.3.0), *withr*(v.2.5.0), *stringr*(v.1.5.0), *knitr*(v.1.41), *generics*(v.0.1.3), *vctrs*(v.0.6.2), *grid*(v.4.2.2), *tidyselect*(v.1.2.0), *glue*(v.1.6.2), *R6*(v.2.5.1), *fansi*(v.1.0.3), *readxl*(v.1.4.3), *rmarkdown*(v.2.19), *pander*(v.0.6.5), *whisker*(v.0.4.1), *farver*(v.2.1.1), *purrr*(v.1.0.1), *tidyr*(v.1.2.1), *magrittr*(v.2.0.3), *prismatic*(v.1.1.1), *ellipsis*(v.0.3.2), *scales*(v.1.2.1), *htmltools*(v.0.5.4), *showtext*(v.0.9-5), *zvplot*(v.0.0.0.9000), *colorspace*(v.2.1-0), *labeling*(v.0.4.2), *utf8*(v.1.2.2), *stringi*(v.1.7.12) and *munsell*(v.0.5.0)

