

Effective Data Visualisation

Data Exploration and Visualisation

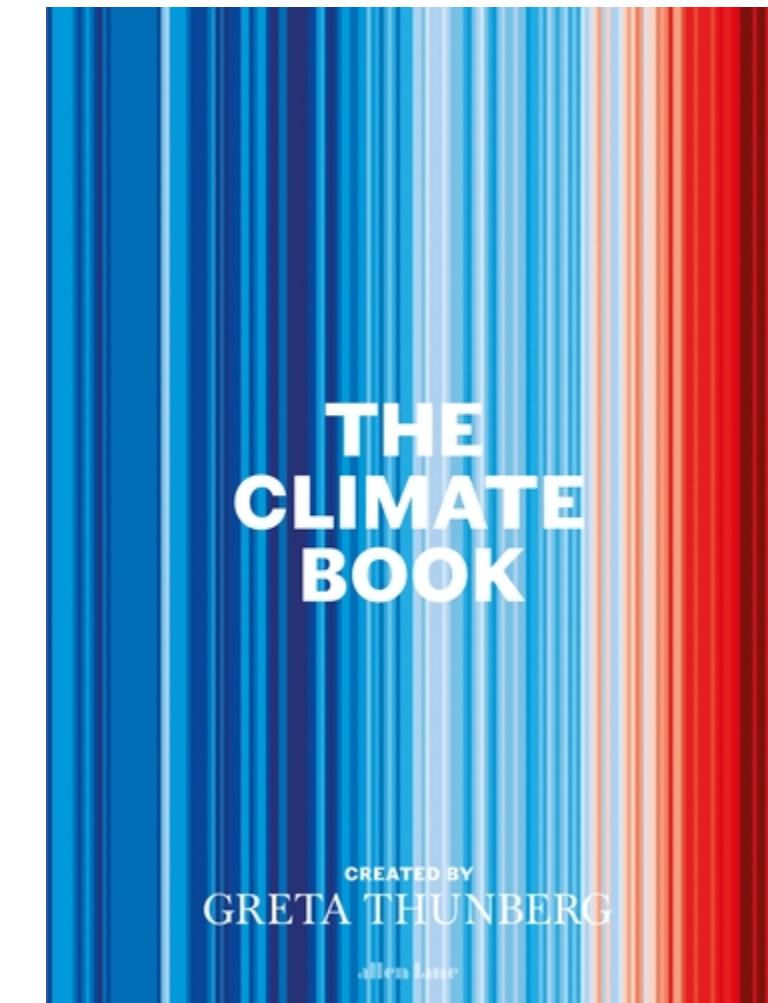
Dr Zak Varty

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.

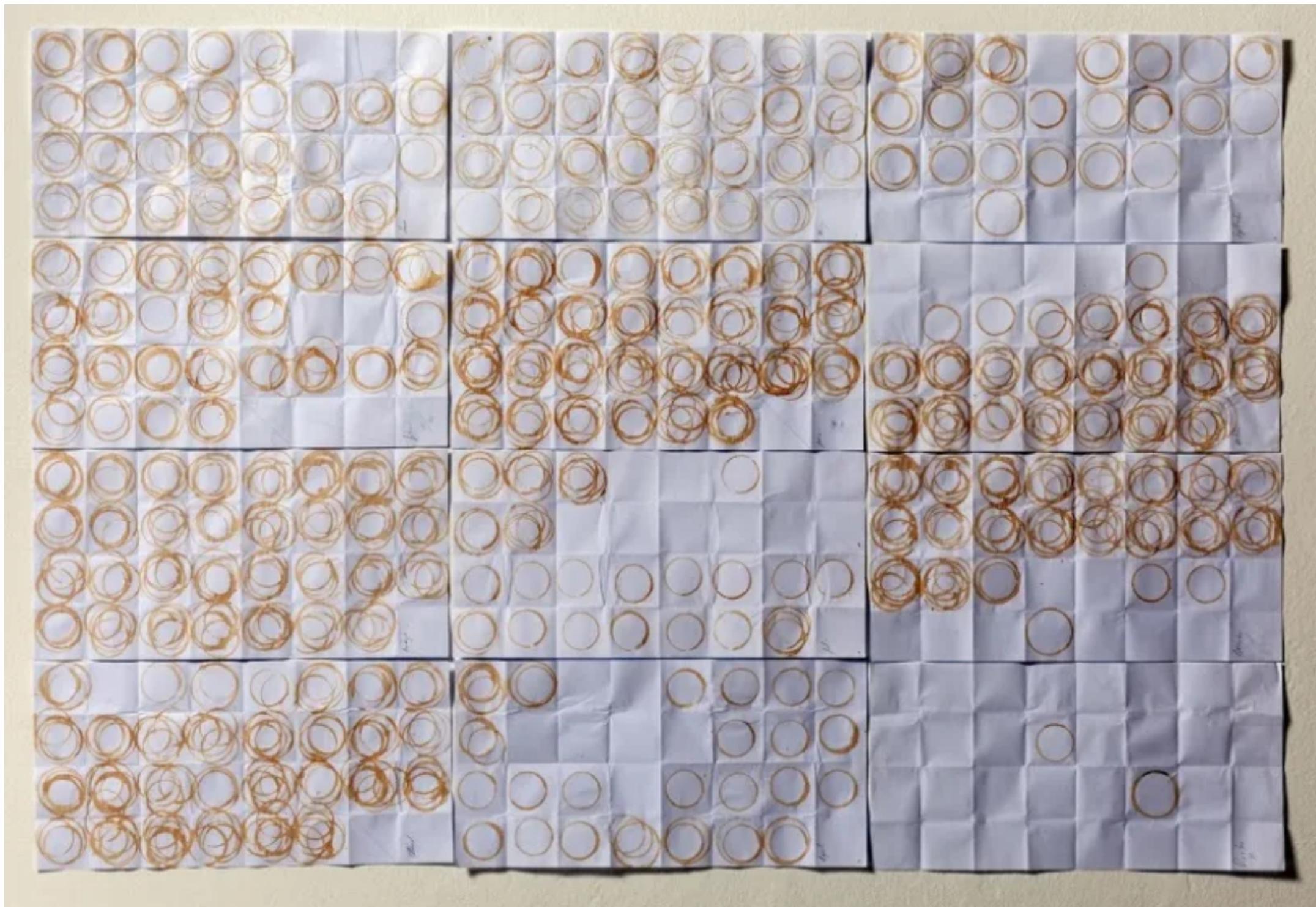
Reflective exercise, not a tutorial or rulebook.



Warming stripes graphic on the cover of “The Climate Book”

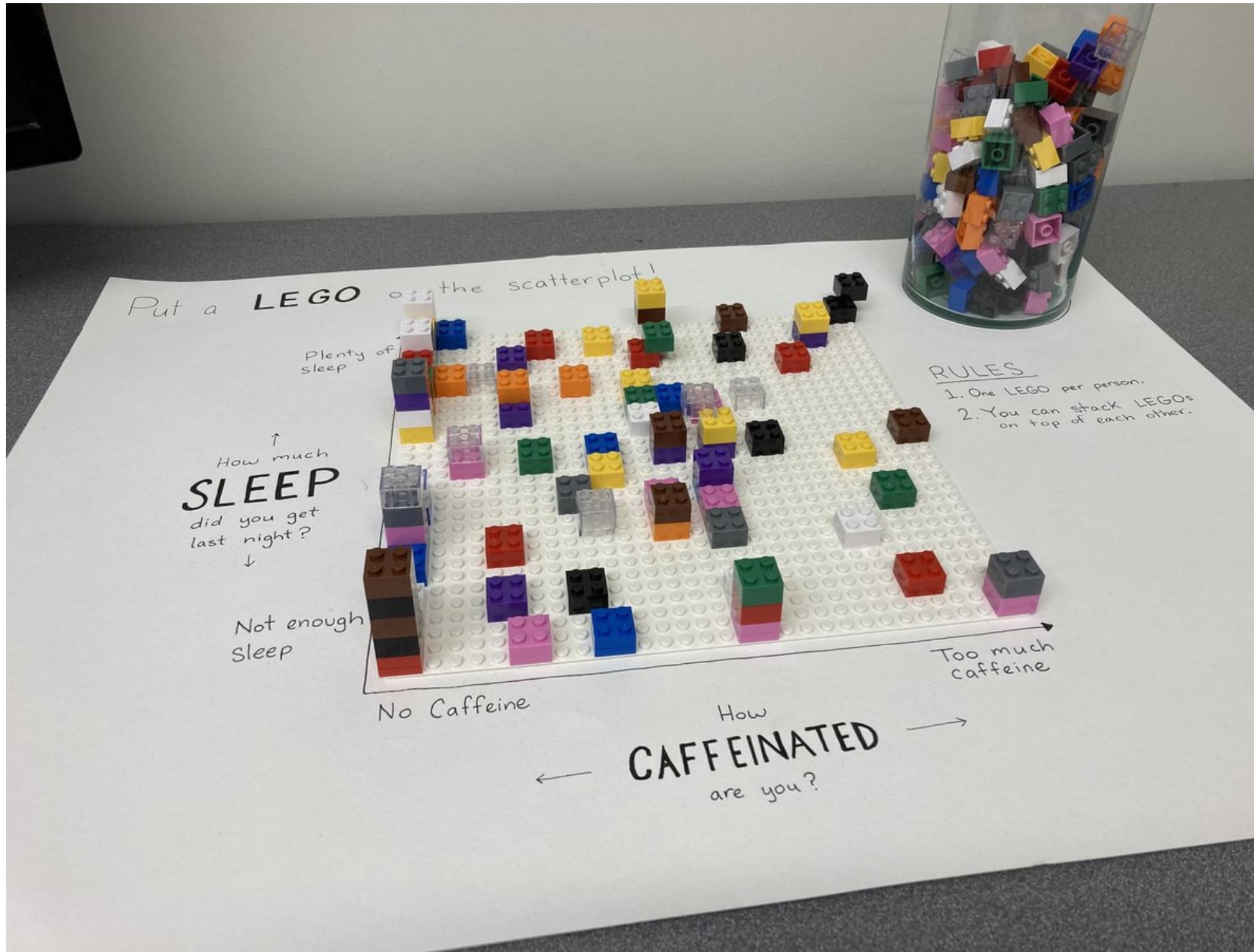
1: Think About Your Tools

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



Analysis



Journalism

Presentations



Reports

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

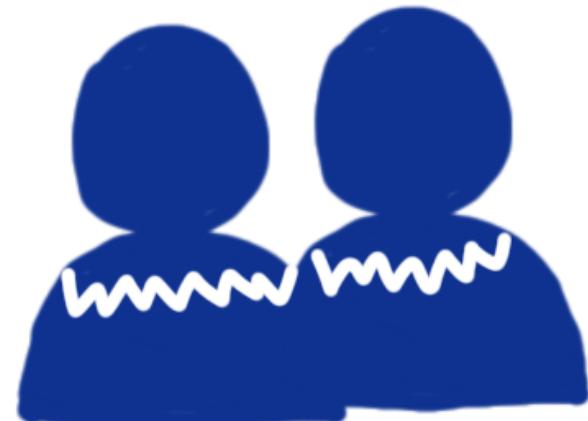
Know your audience



USER GROUP
← 1

USER GROUP

2 →



USER GROUP
← 3

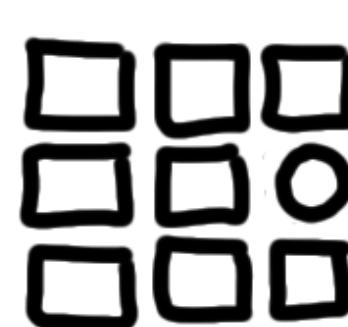
Who is the intended audience for your visualisation?

What knowledge do they bring with them?

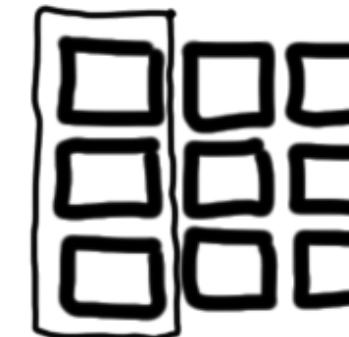
What assumptions and biases do they hold?

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

Preattentive Attributes



Shape



Enclosure



Weight



Saturation



Colour



Size



Marking



Orientation



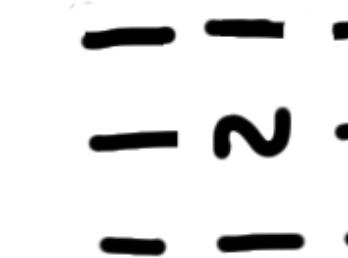
position



density

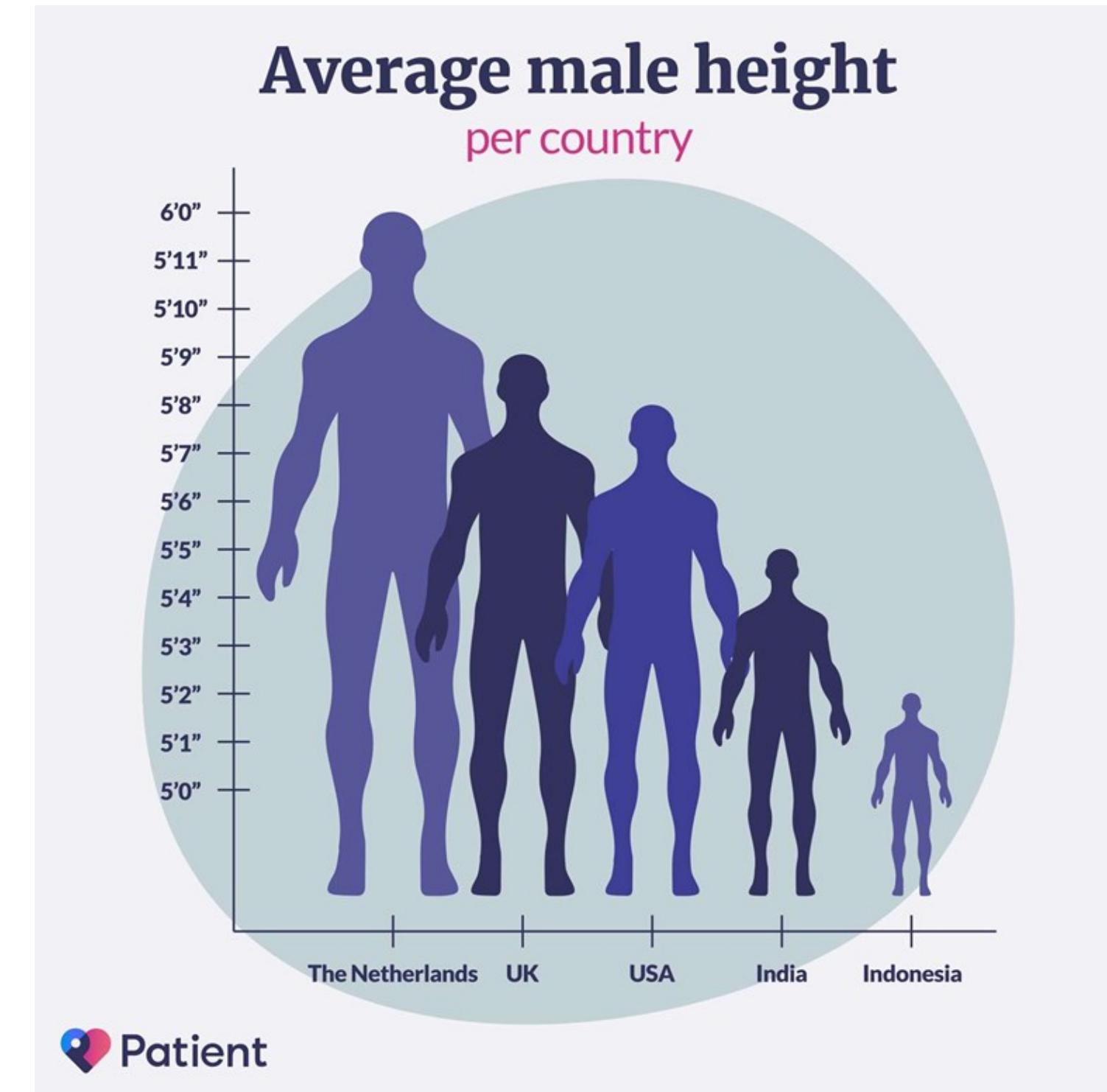


length



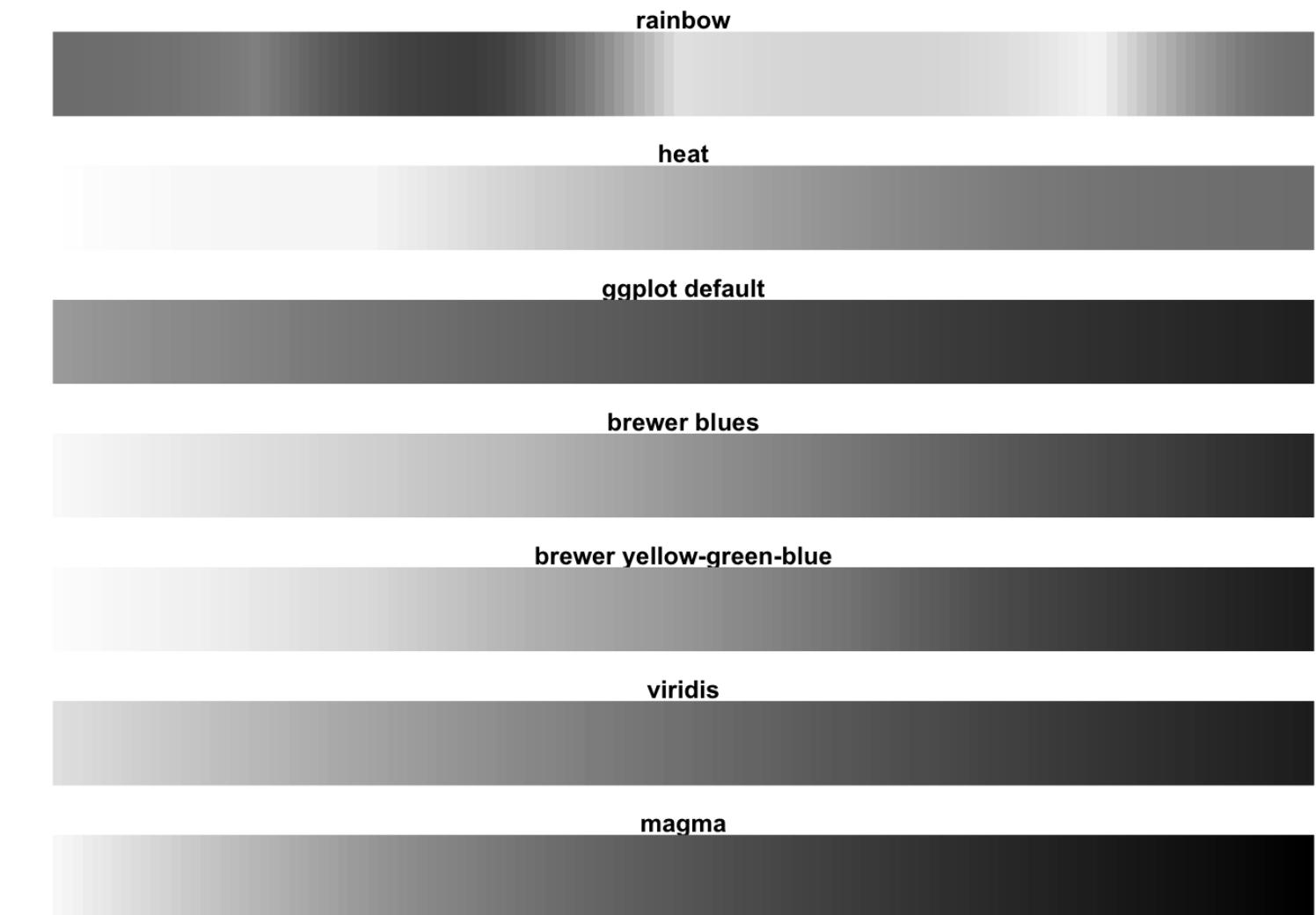
Curvature

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](#))

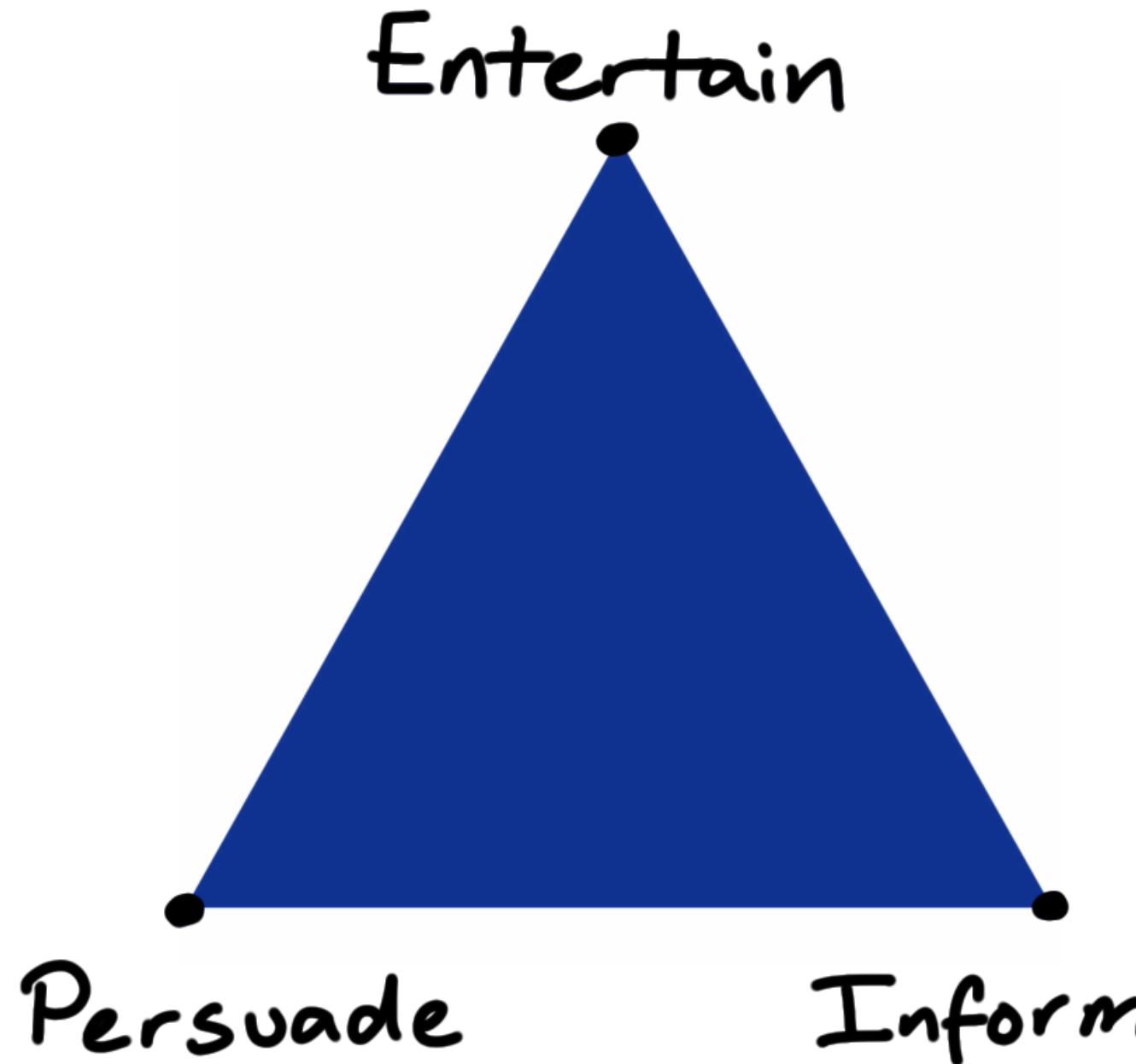
Titles

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

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

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)

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

