

# EGM101 – Skills Toolbox

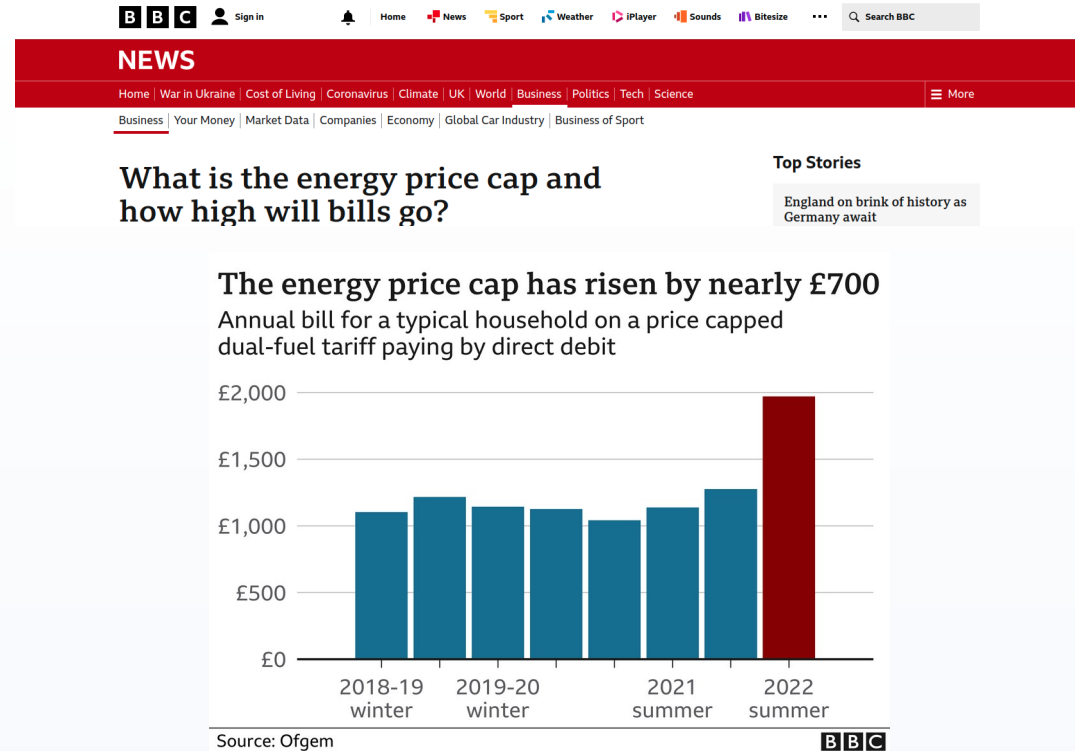
Week 5, Part 2: Presenting Data

# Why do we present data?

- Science/research does not happen in a vacuum
- Need to communicate!
  - To whom?
  - Why should they care?
- Need to explain research:
  - What is it?
  - What did you find?
  - Why does it matter?
- Presenting data is part of communicating *effectively* to an audience



- Reports
  - Arguably most common form
  - Press articles
  - Peer-reviewed research
  - “gray” literature
- Tables
- Charts/Figures/Graphs
  - [r/datavisualization](https://r-datavisualization.github.io/)



<https://www.bbc.co.uk/news/business-58090533>

# Tables: general rules

- Explanation:
  - Title (stand-alone)
  - Caption (report/article)
- Headings
  - Rows
  - Columns
  - Units!
- Remember significant figures!

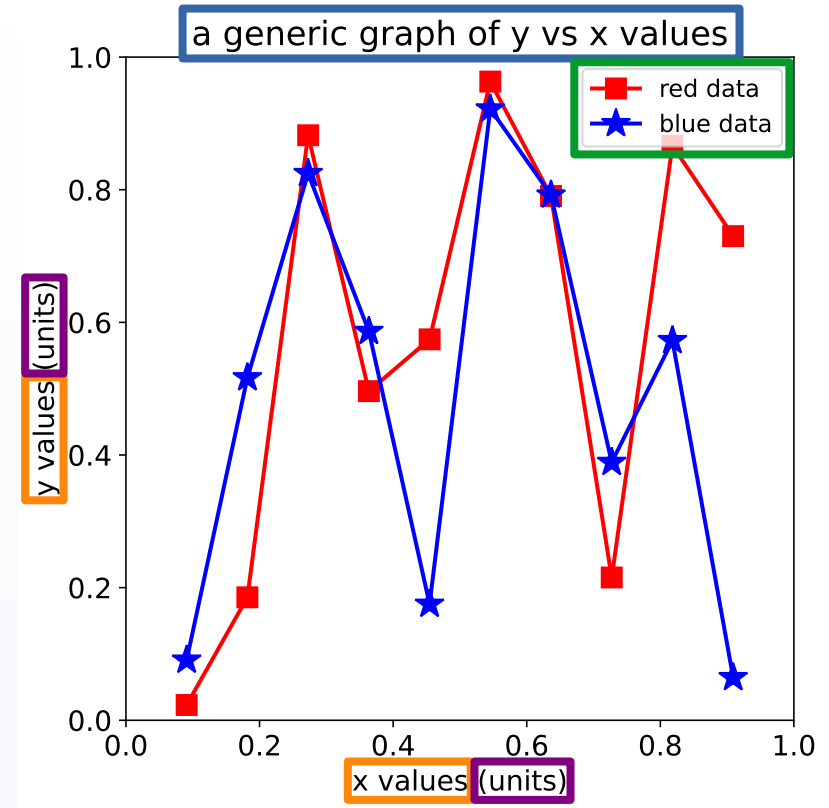
Ocean Area

| <b>Name</b> | <b>Area</b><br>1000 km <sup>2</sup> | <b>Area</b><br>% |
|-------------|-------------------------------------|------------------|
| Pacific     | 155.56                              | 43.07            |
| Atlantic    | 76.76                               | 21.26            |
| Indian      | 68.56                               | 18.98            |
| Southern    | 20.33                               | 5.63             |
| Arctic      | 14.06                               | 3.89             |
| Other       | 25.87                               | 7.16             |

**Table 1:** Area (in 1000s of km<sup>2</sup> and %) of the world's oceans

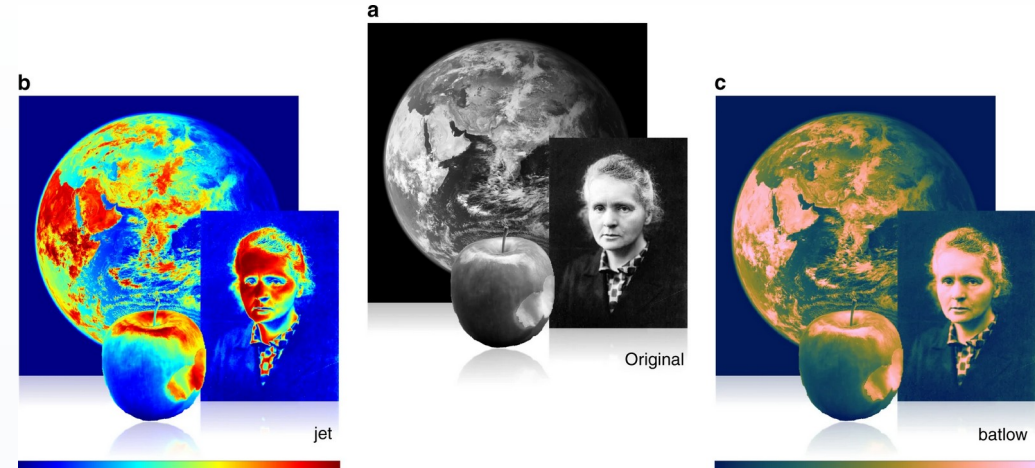
# Charts/figures/graphs: general rules

- Needs an explanation:
  - Title (stand-alone)
  - Caption (report/article)
- Clearly-labeled axes
  - Units!
- Key/Legend:
  - Explains colors/symbols



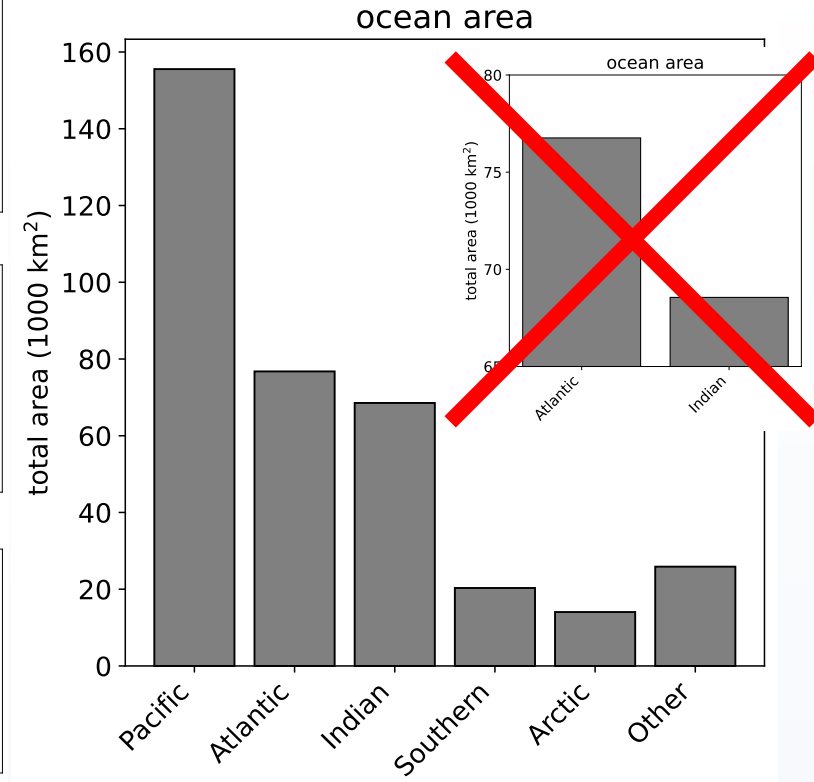
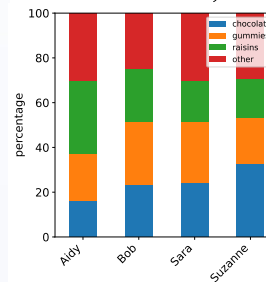
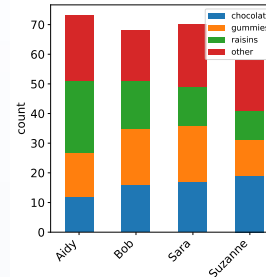
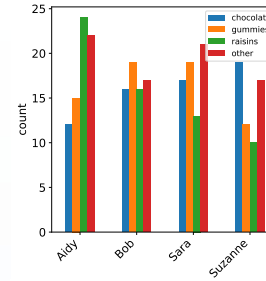
# A word on color

- Color is an incredibly effective way to communicate in charts/figures
  - But, it can be abused/misused
  - Not everyone sees color in the same way
- Avoid:
  - Color maps that distort
  - Inaccessible color combinations
- Think about the data + what you are trying to show!



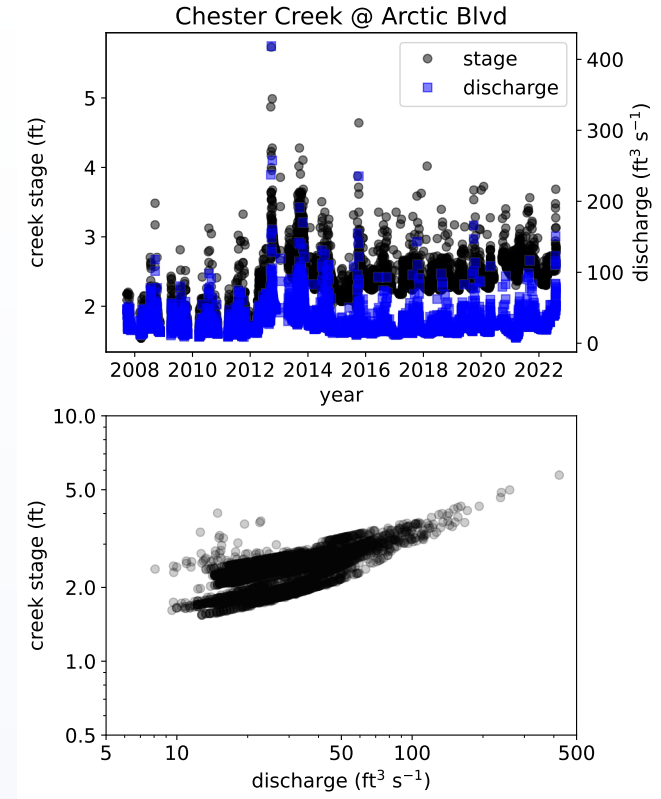
Crameri et al., 2020

- Height/length of bars is proportional to represented values
- Some variations:
  - Grouped
  - Component/stacked
  - Percent component/stacked
- Tell a story about magnitudes
  - Axis should *always* include the full bar!



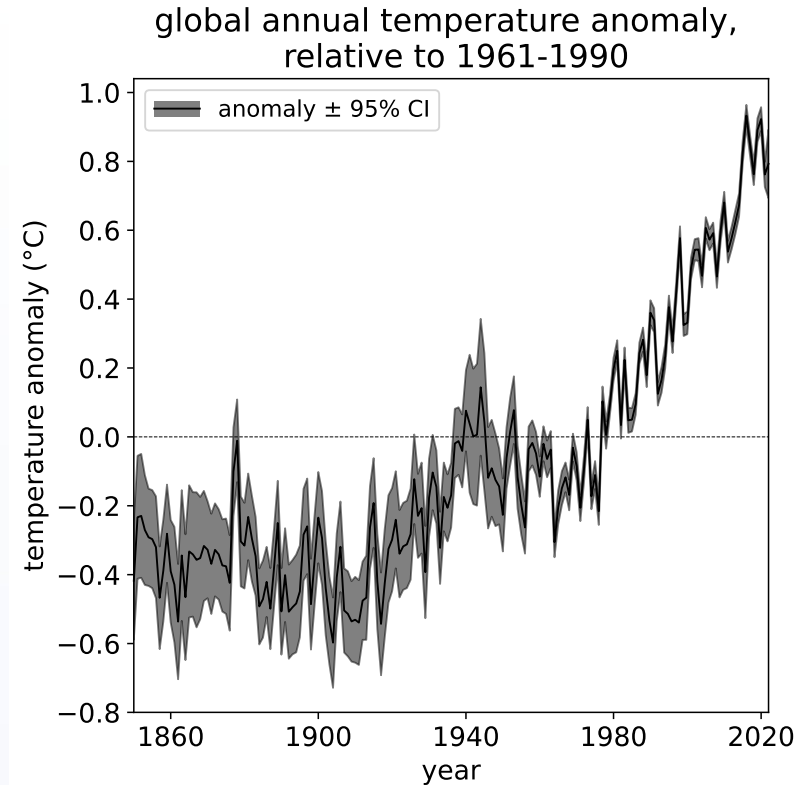
- Display paired values of two (or more) variables
- Helps to show relationships\* between variables
- Logarithmic scale:
  - Helps show variation with large differences in values
  - Typically plots  $\log_{10}(\text{data})$
  - Ticks aren't evenly spaced

\*if they exist! (more on this next week)





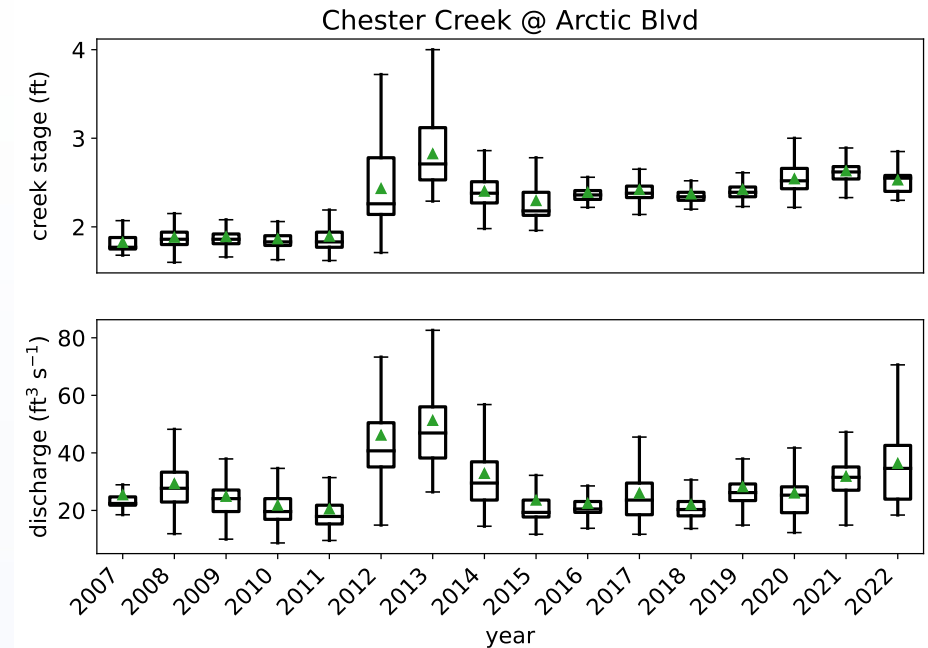
- Tell us about changes
  - Between two variables
  - Over time
- If there isn't a reason to connect the dots, don't!



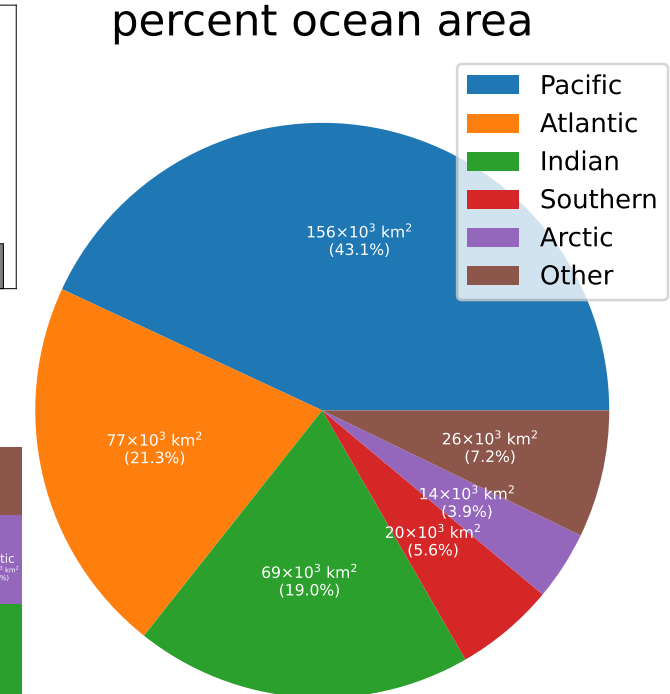
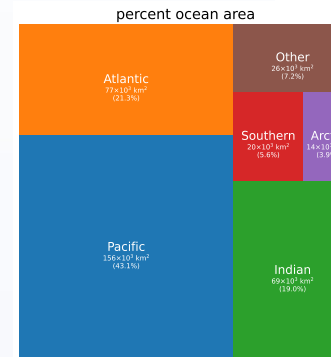
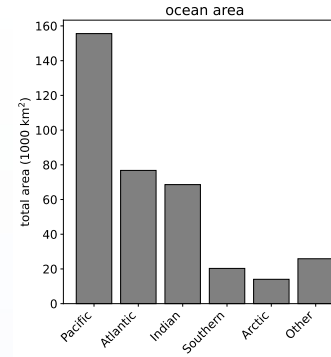
Data Source: <https://www.metoffice.gov.uk/hadobs/hadcrut5/data/current/download.html>

# Box (& whisker) plots

- Show concentration of data
- Box: middle 50% of the data (**Interquartile Range**)
  - Line in middle: **median**
- Whiskers are usually:
  - Maximum/minimum values
  - $1.5 * \text{IQR}$  (showing outliers)



- Shows proportion of a whole
  - If your data doesn't add to a whole, use something else
- Problem: humans are not great at “eyeballing” angles
  - Alternatives:
    - Bar chart
    - Treemap
- 3D pie charts are a great way to lie, though.



- Part of research is communicating the results
- Several ways to present data effectively
- For charts/figures:
  - Choose the right tool for the job
  - Remember to label!
  - Try to use color effectively

- Illowsky and Dean, Chapters 2.1, 2.4
- Caswell, Chapter 5
- Weiss, Chapters 2.2 – 2.5
- Huff, “The Gee-Whiz Graph” (Chapter 5)
- Bergstrom and West, “Data Visualization” (Chapter 7)
- Effective Communication, Better Science [[Scientific American](#)]
- Crameri et al., 2020 [[Nature Communications](#)]