

Data Visualization for Teaching and Learning Projects: Tips and Tricks



Creative Commons, attribution, non commercial, share alike.

Dec 13, 2022 1:30-3PM

Trish Varao-Sousa, Evaluation & Research Consultant

Natasha Pestonji-Dixon, Scholarship of Teaching and Learning (SoTL) Facilitator

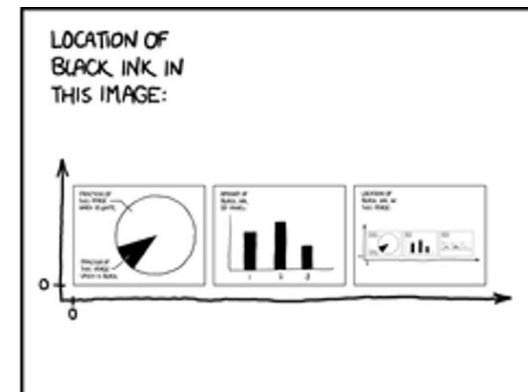
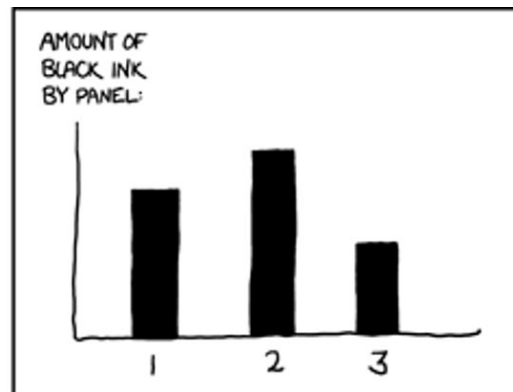
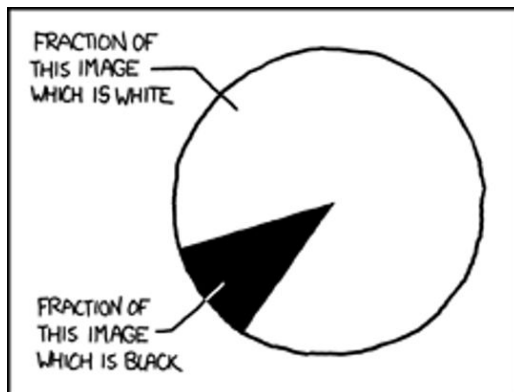


Creative Commons Attribution-Share Alike 2.0 Generic license
Source: https://commons.wikimedia.org/wiki/File:UBC_aerial_view.jpg

Workshop outline



- Principles of data visualization
- What is the purpose of your visualization?
- Basic do's and don'ts for data visualization
 - Bar/column charts
 - Line plots and scatterplots
 - Other less common visualizations
 - Visualizing qualitative data
- Best practices and accessibility considerations
- Q&A and resources



What brought you to today's session?



Thank you for taking time to fill out your responses in Jamboard!



Why are you visualizing this data?



Who is the audience?



Expert vs. novice

CC0 - Public Domain

What are they trying to learn from the figure/image?



Exploratory: To learn about the data
Explanatory: To tell a story

CC0 - Public Domain

Principles of data visualization



“Graphical excellence is that which gives the viewer the **greatest number of ideas in the shortest time** with the least ink in the smallest space.”

- Edward Tufte

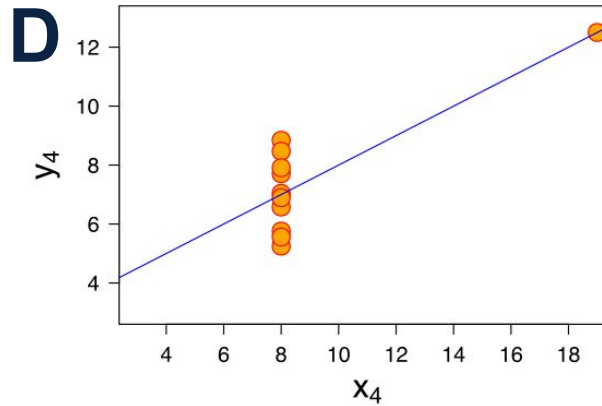
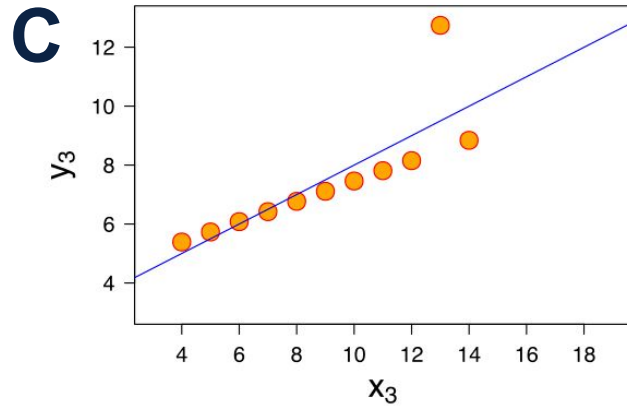
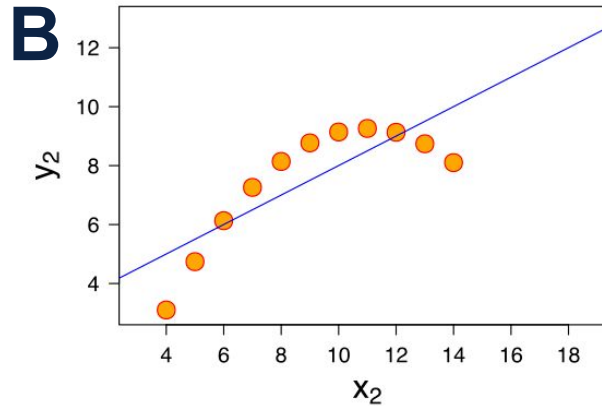
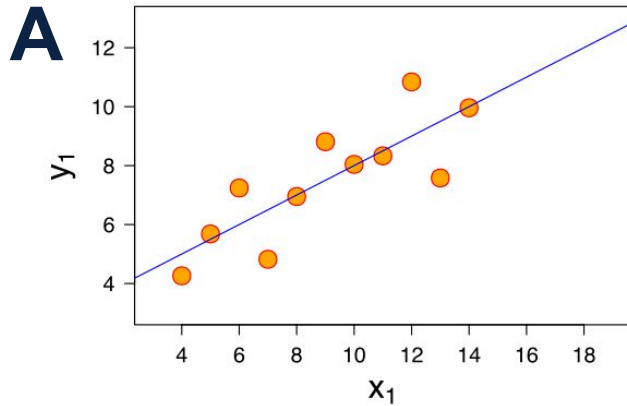
Principles of data visualization

A good visualization:

- 1) Represents quantities accurately
- 2) Clearly indicates how the values relate to one another
- 3) Makes obvious how people should use the information

(Few, 2013)

Which graph has the highest mean?



Anscombe's Quartet

All have the same:

- Mean
- Variance
- Correlation

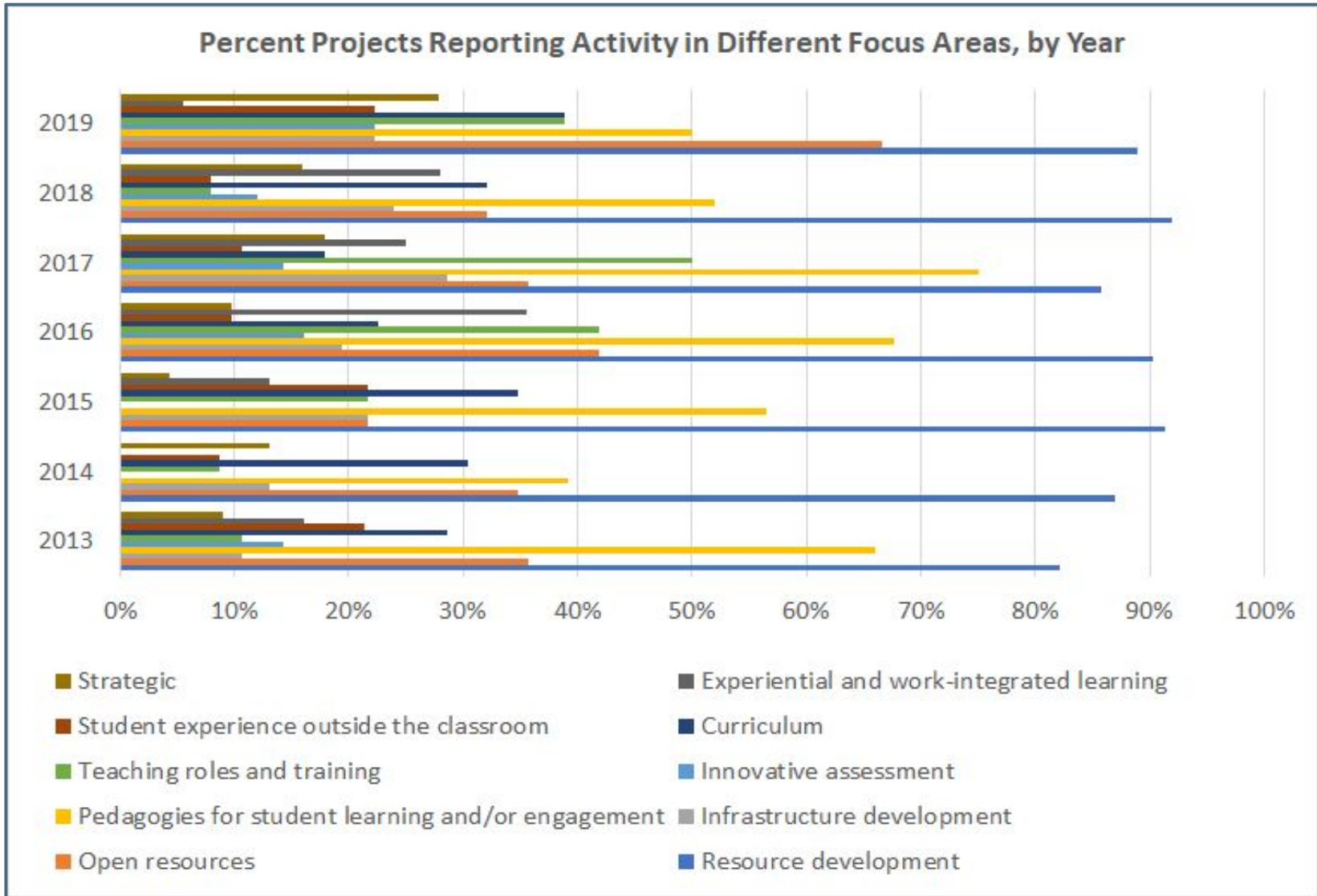
Common figure types: Bar/column charts



Useful for:

- comparing categories
- plotting means (e.g., grades, motivation scale means) or ranges of values (e.g., likert scale responses)

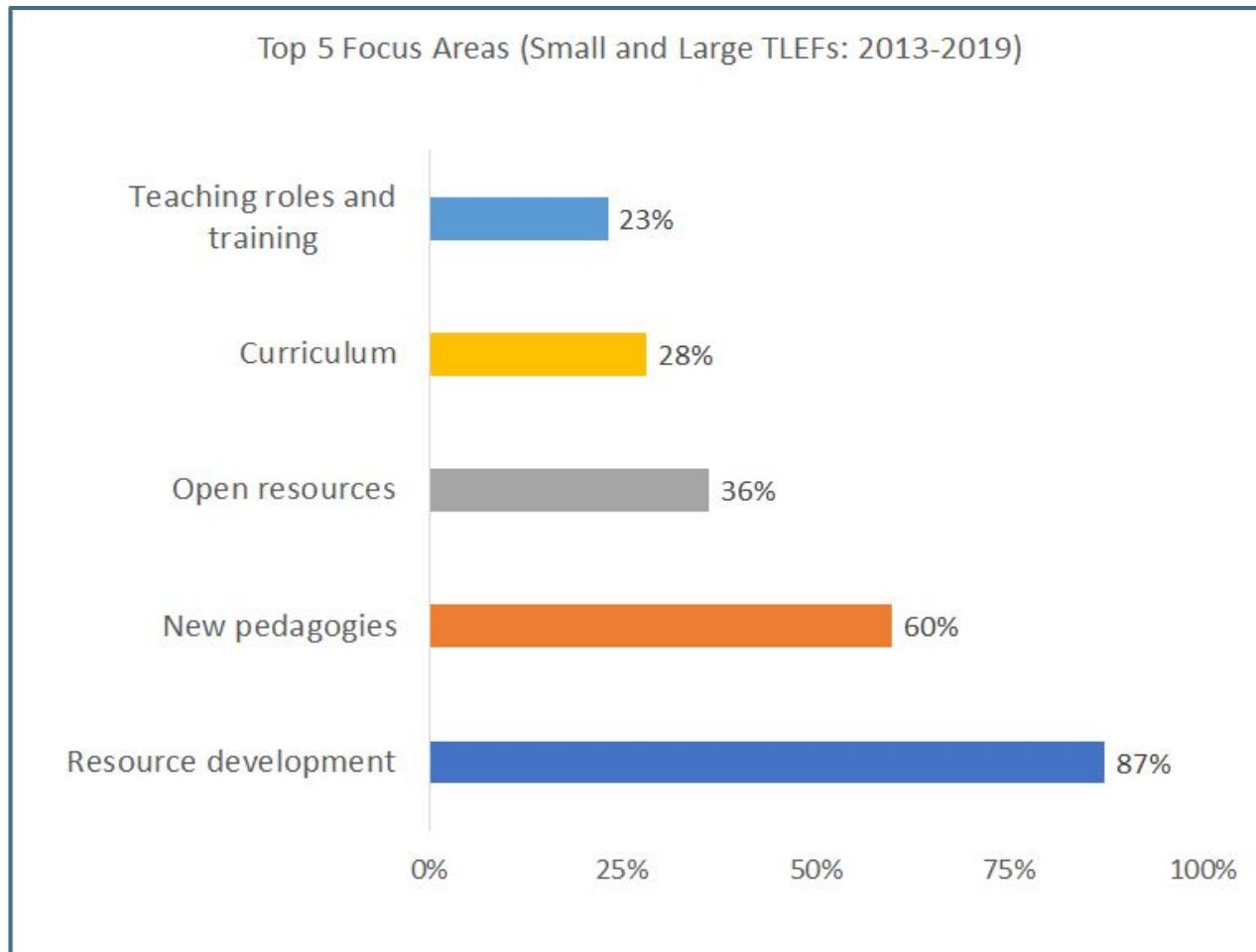
A problematic bar chart: What are the issues?



Easy ways to make your data clearer



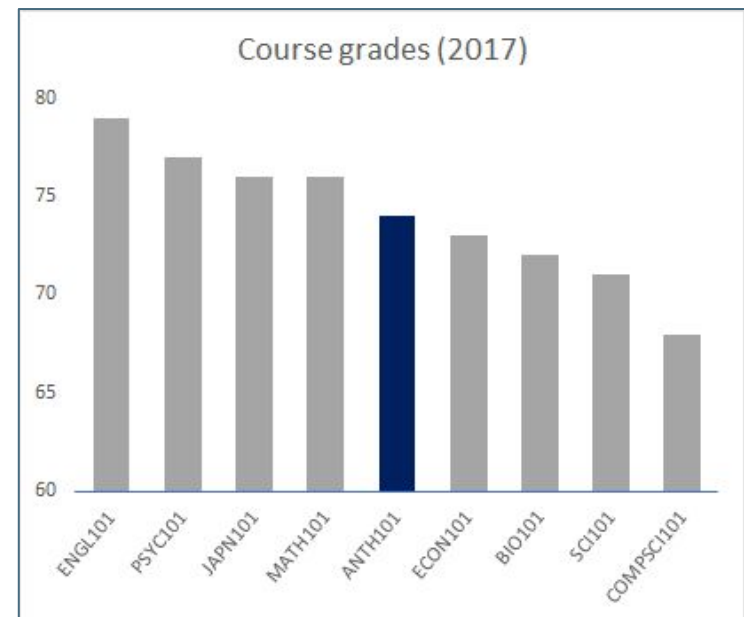
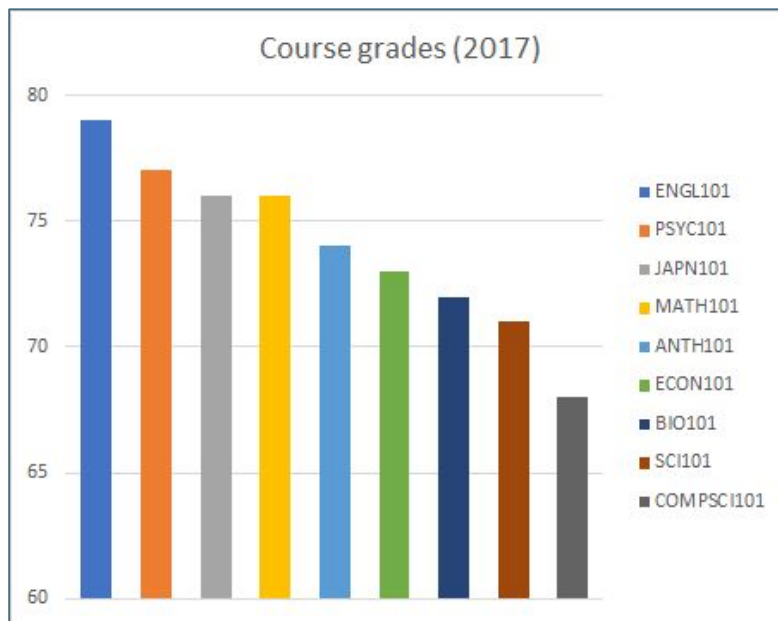
1) **Rank items** in a helpful way. With a small #, rank by count/percentage. With a large #, rank alphabetically and highlight key points.



Easy ways to make your data clearer



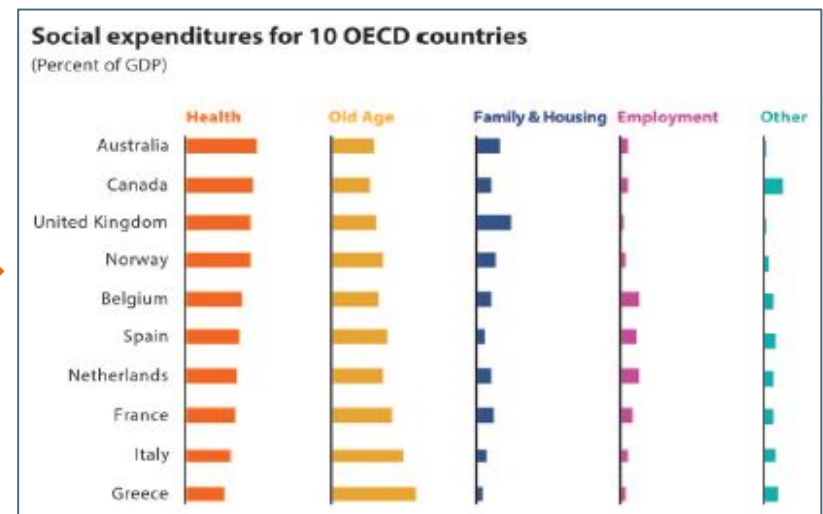
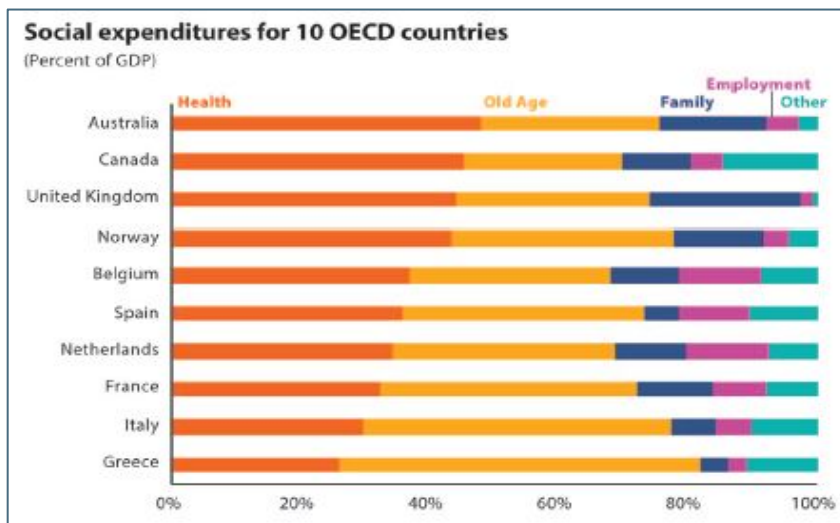
2) Use perceptual features (e.g., shape, colour, etc) to **highlight what is important** for the viewer to focus on.



Easy ways to make your data clearer



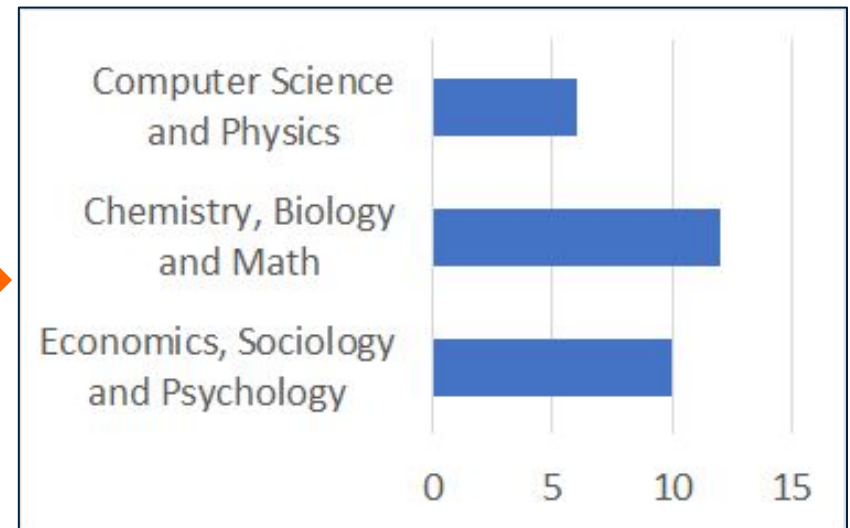
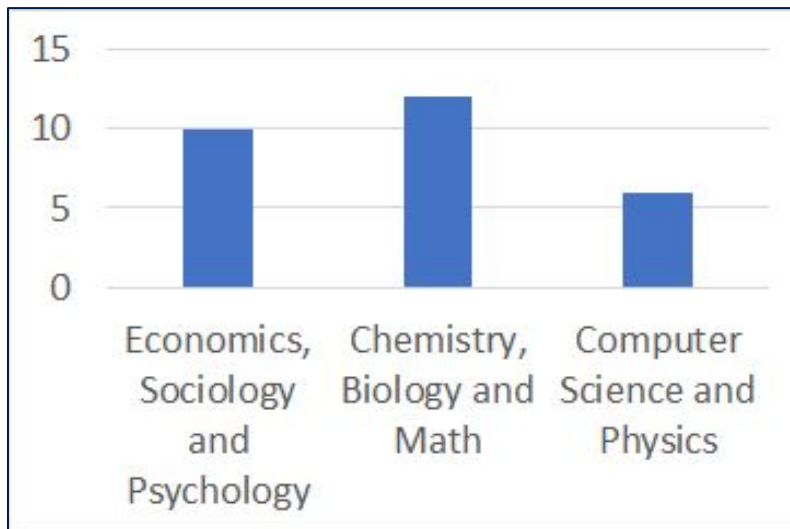
3) **Split the data** when you want to **compare values** within each category more easily (best when there are multiple factors being examined).



Easy ways to make your data clearer



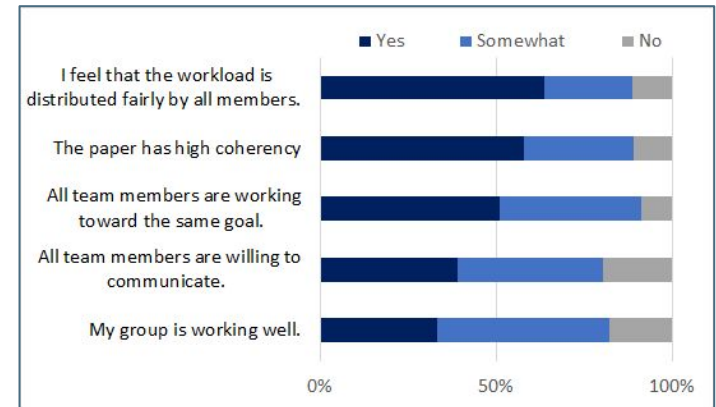
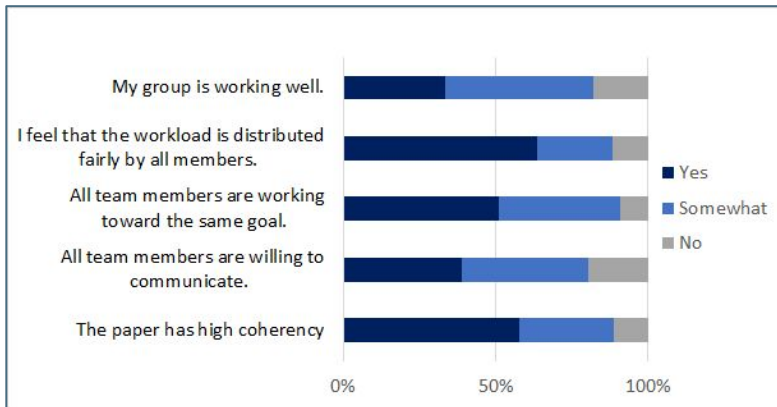
4) Use a column chart instead of bar chart if the labels are long.



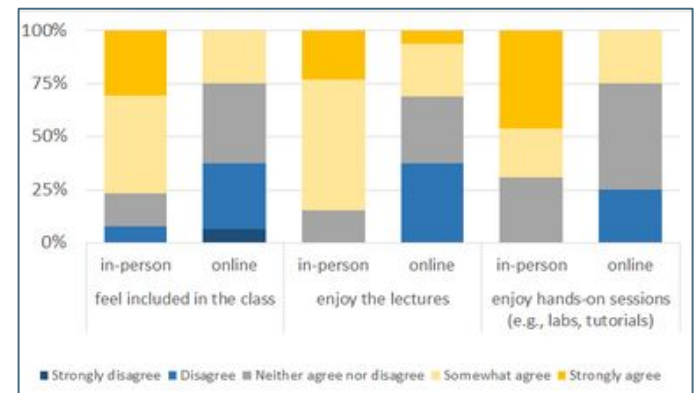
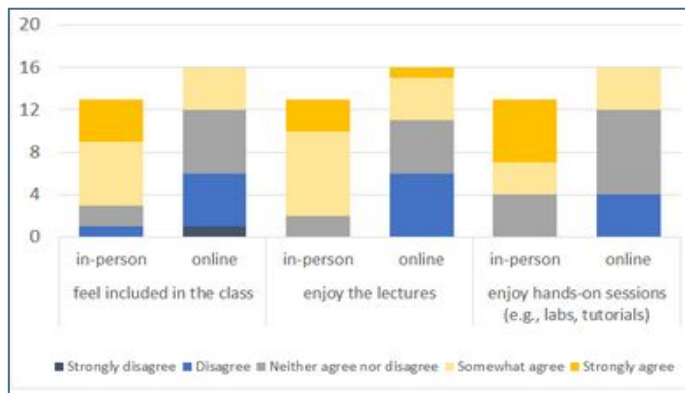
Tips for plotting Likert-type scale data



1) Sort “positive/strongly agree” data largest to smallest (or smallest to largest). Include the legend at the top.



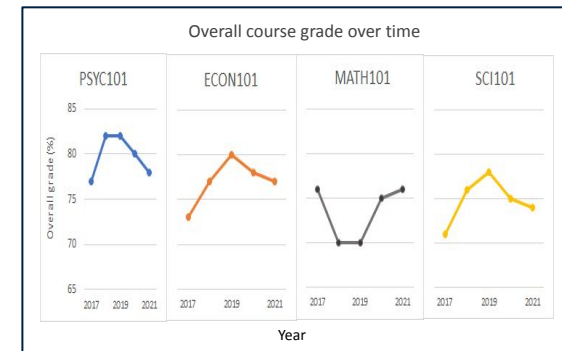
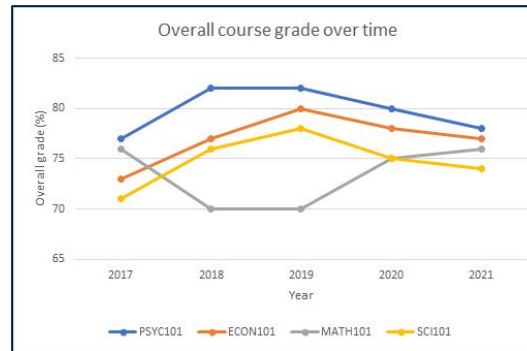
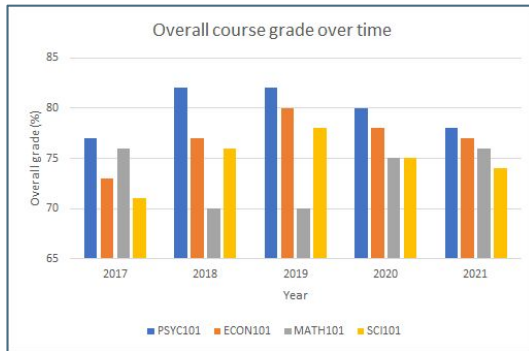
2) Use 100% stacked plots so that the data is on an equal “spread” and more easily comparable.



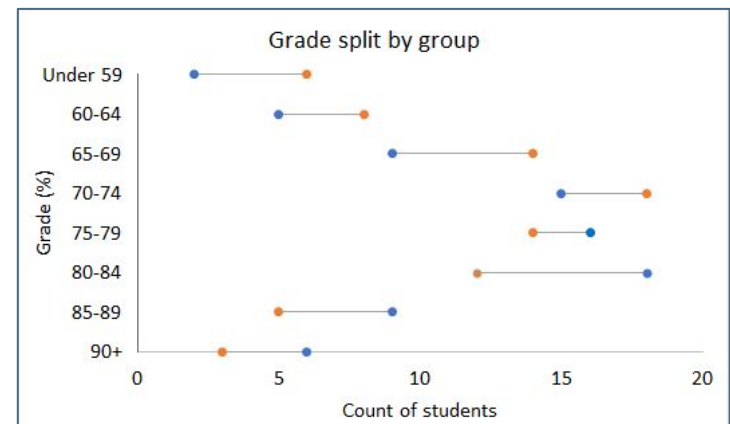
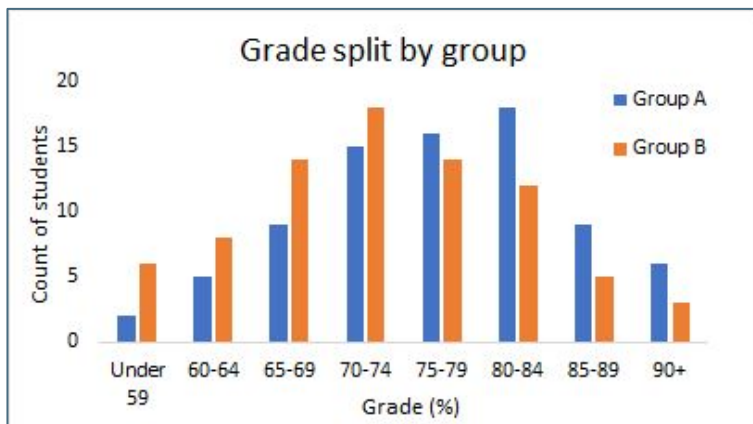
When a bar chart is too messy



A **slope chart** is typically better for showing changes over time compared to a bar chart.



A **lollipop chart (or dumbbell dot plot)** helps compare data more easily. Works well when you have only two categories of data that you are comparing.





CC0 1.0: Public Domain

Common figure types: Line & scatterplots



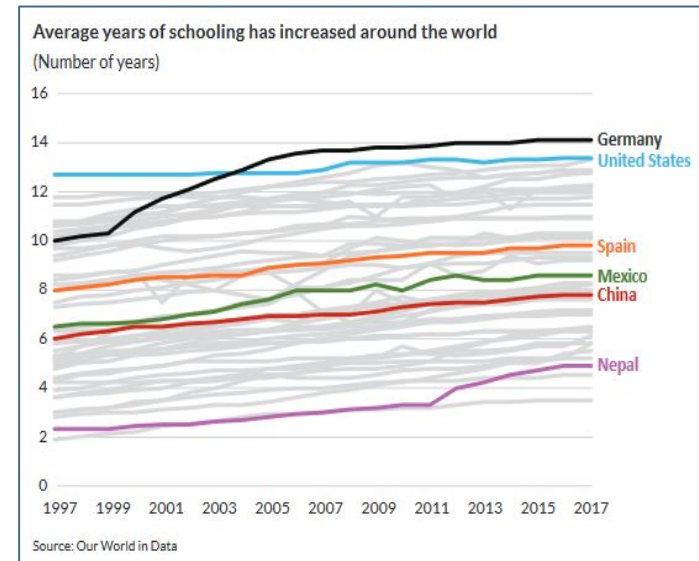
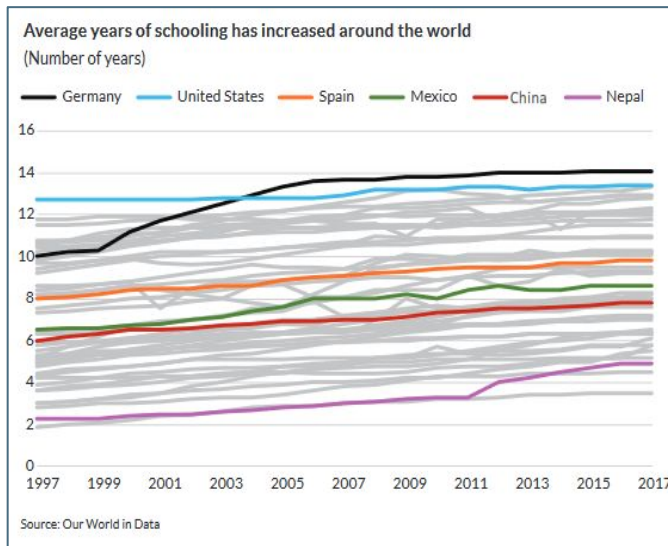
Useful for:

- Temporal data (e.g., changes over a term or years; pre/post data)
- Continuous data (e.g., student grades)

Easy ways to make your data clearer



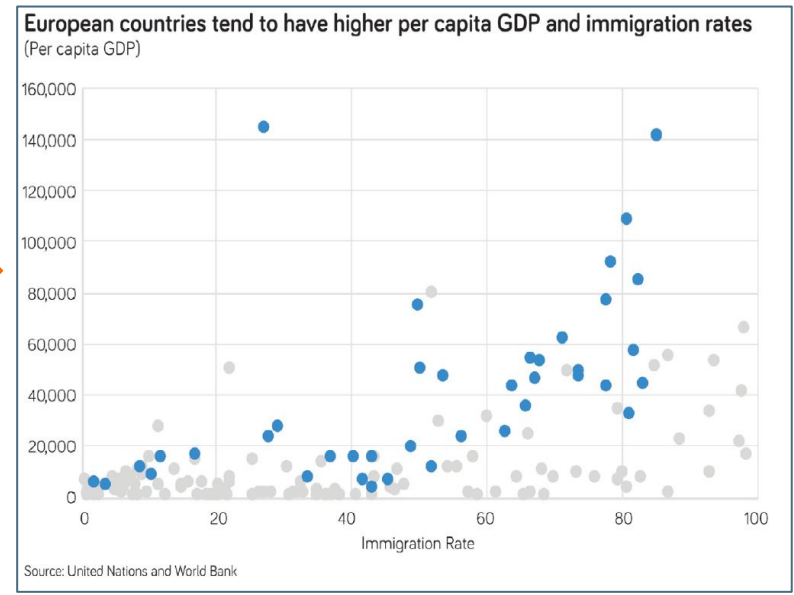
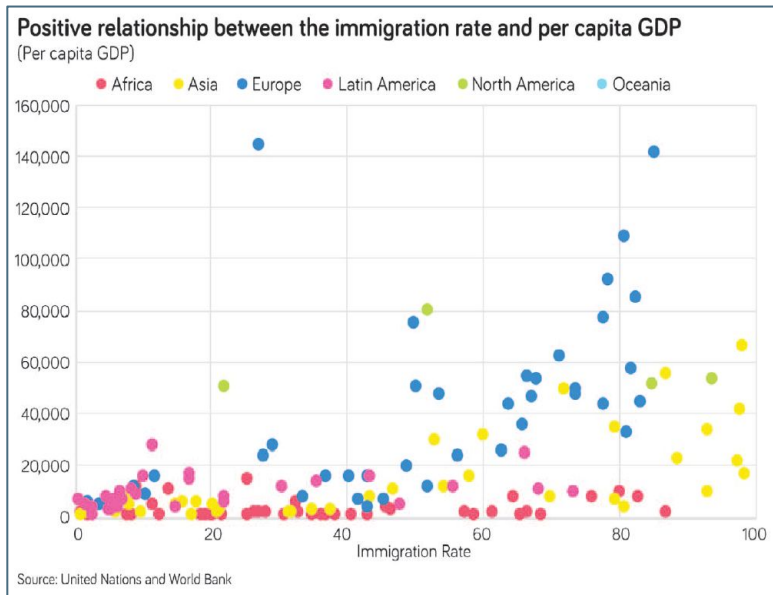
1) Reduce clutter (varying fonts, depth, axis lines). With multiple categories, use a feature to **highlight meaningful data** or **reduce plotting** to four lines per chart.



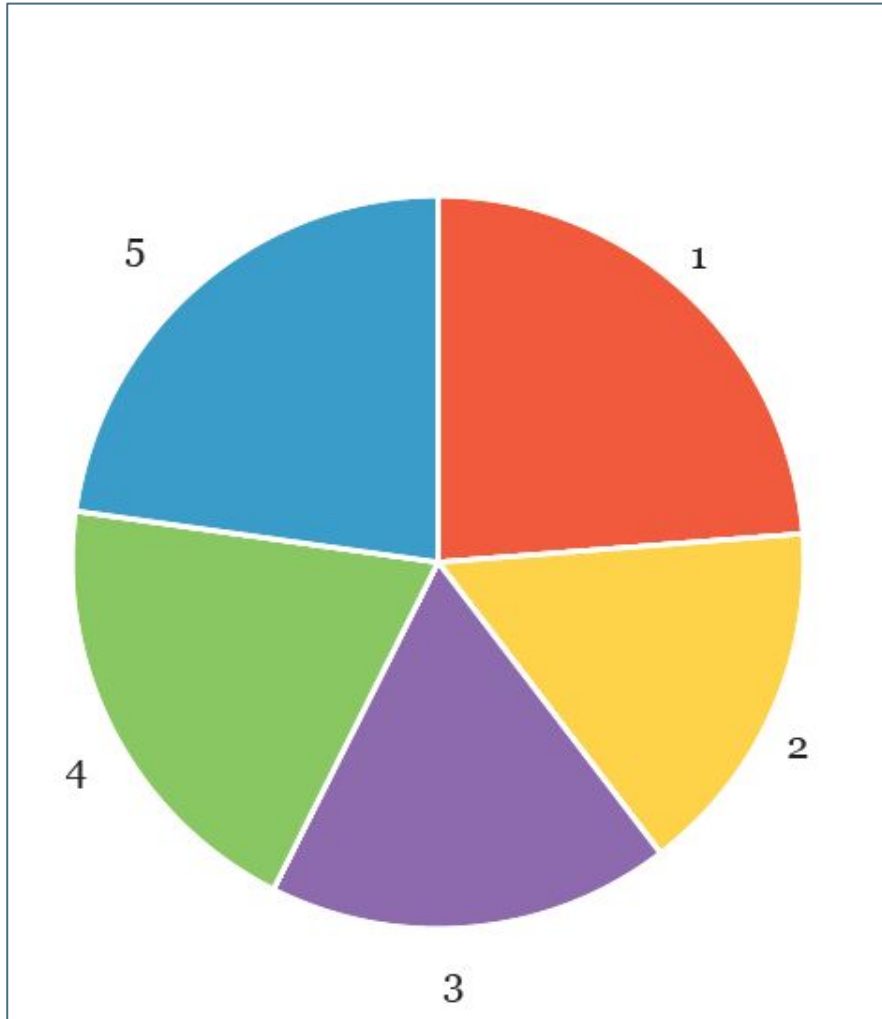
Easy ways to make your data clearer



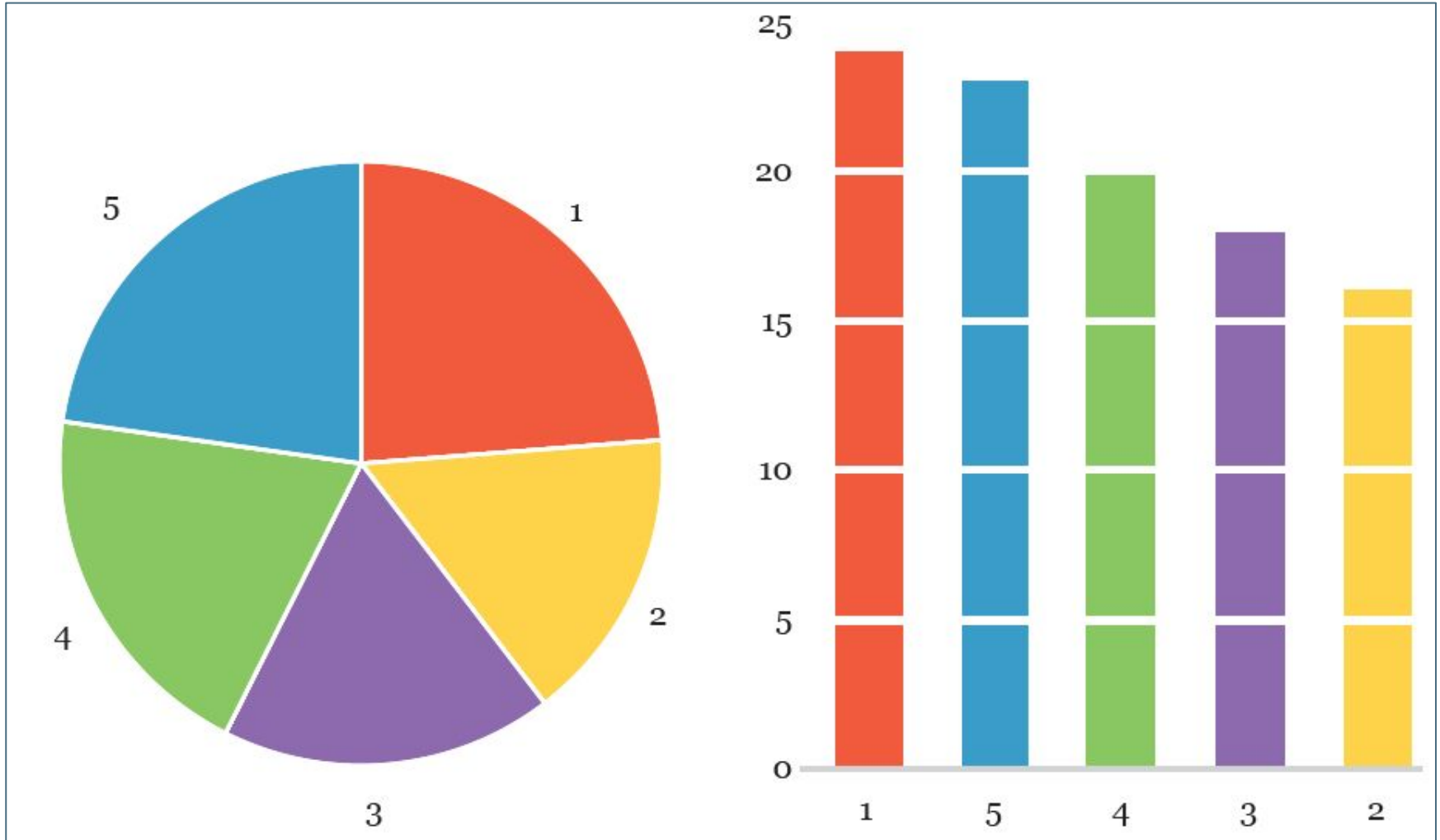
2) With a scatterplot, **highlight the focus** by adding color to a subset of data points, or **change the title** to emphasize a point.



Less common visualizations: Pie charts



Less common visualizations: Pie charts

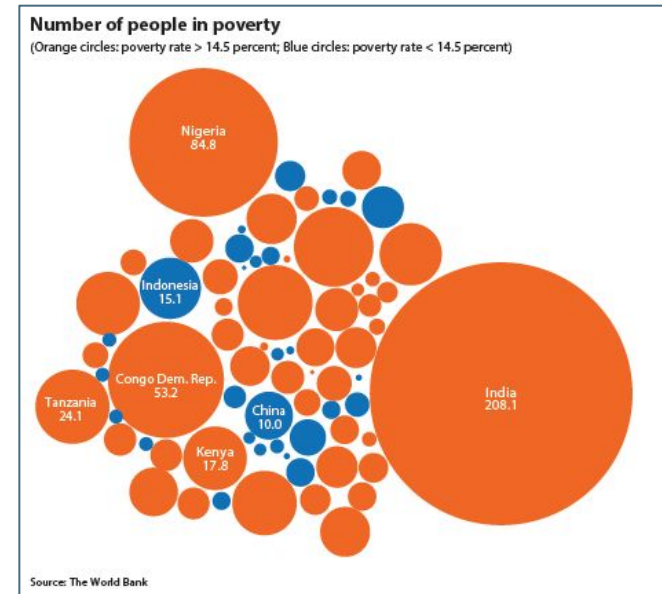


Less common visualizations: Pie & donut charts



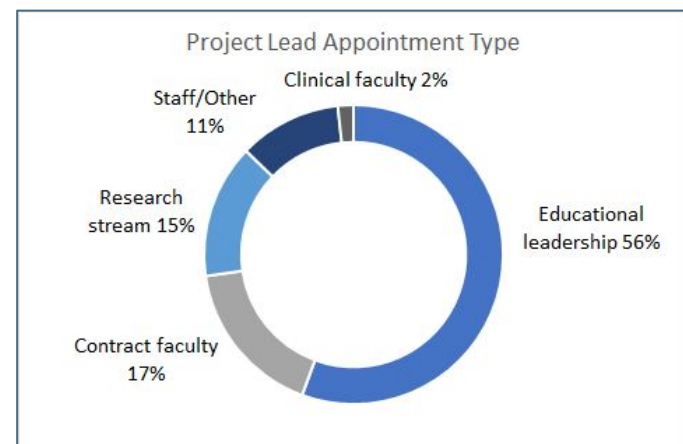
Pie charts and bubble plots:

- Very hard to compare areas
- Rarely the best option for presenting data



Donut charts:

- Help to reduce the issue of confusing area comparisons



Less common visualisations: Pictographs

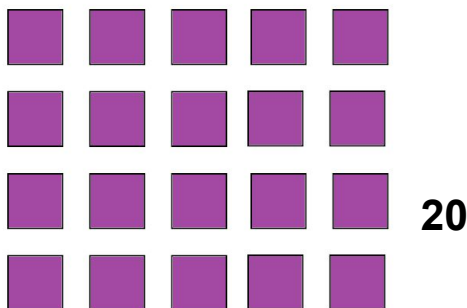


Best practices:

- **Don't** truncate, use full images
- **Don't** distort images to demonstrate increase/ decrease, instead use duplicates of the same image



- **Do** provide a numeric visual for # of items (don't make people count!)
- **Do** group in a way that is easy to tally



Visualizing qualitative data: Word clouds and word trees



Word clouds/trees:

- Can be helpful to get a quick snapshot of themes
- Difficult to compare specific frequencies
- Words may appear larger/more significant simply due to orientation or colour



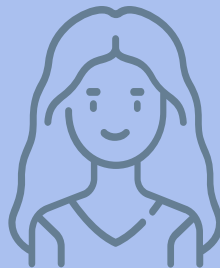
Visualizing qualitative data: Word clouds and word trees



Visualizing qualitative data: Using participant quotes

- Ensure you have student permission to share quotes
- Helpful to provide context
- Consider how icons may help contextualize or strengthen the visualization

“The worksheet allowed me to organize my ideas, and figure out what I was confused about. It helped me start writing my report much earlier.”



“I would have appreciated a clearer build up to the worksheet in class and more specific instructions. Also, feedback on the first draft would have been nice.”



Mixed data presentation



New Peer Feedback System

“I would like to see [system] used for giving feedback on other assignments in this course as well.”



“Because what I’m saying is anonymous, it makes me give more objective feedback and makes it more fair.”



“Anonymity makes people give lower quality feedback.”

“More efficient and convenient than other forms of giving feedback, like on Canvas. It was fast and easy to operate.”



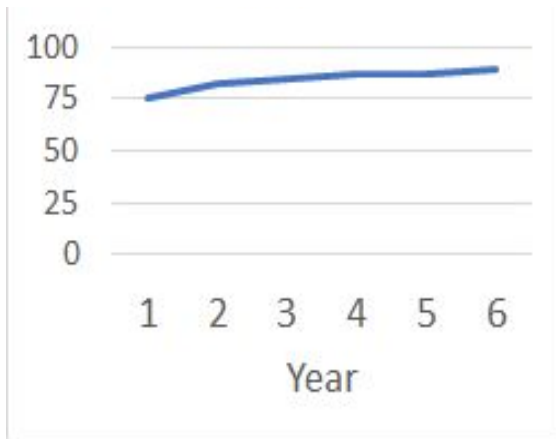
“The format of the text in the program was hard to read.”

Best practices: Scale considerations

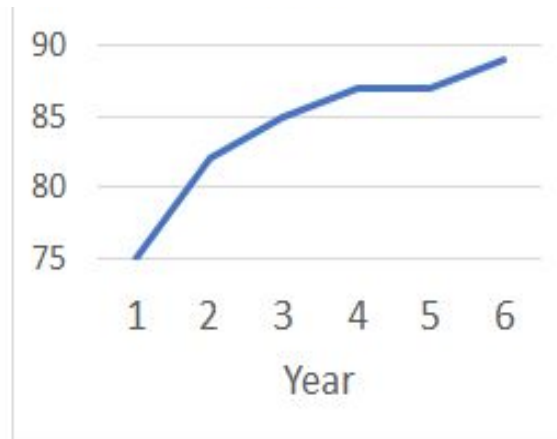


For the y-axis, use a scale that does not over/under exaggerate trends/differences.

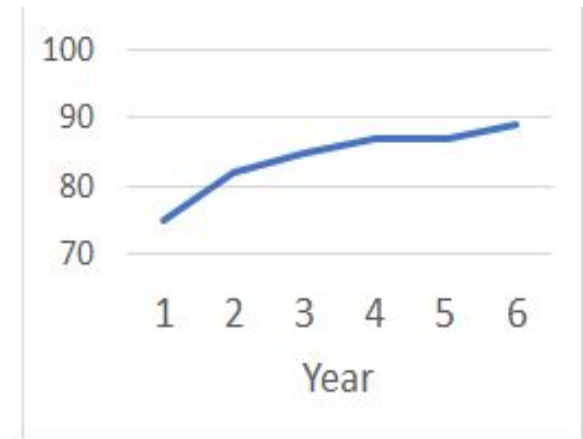
Grade Increase Over Time



Too flat,
minimizes difference



Too exaggerated,
maximizes difference



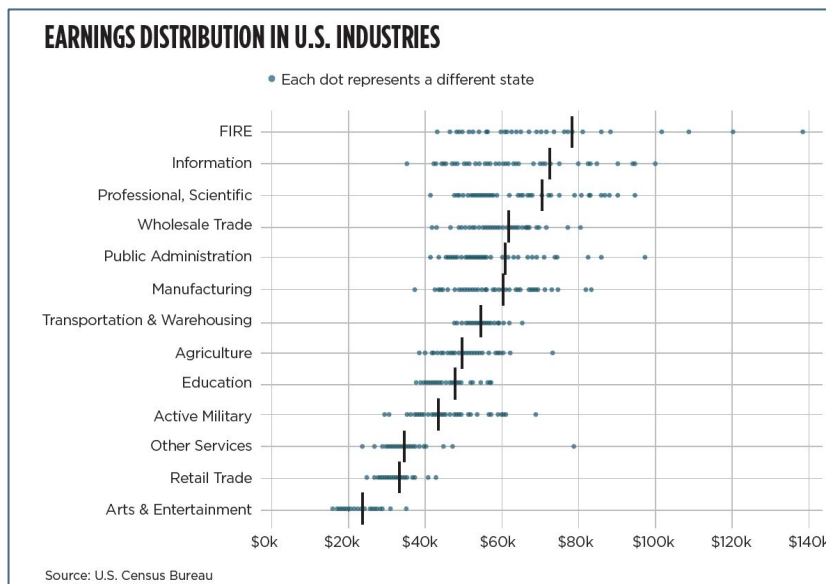
Best choice!

Best practices: Conveying uncertainty

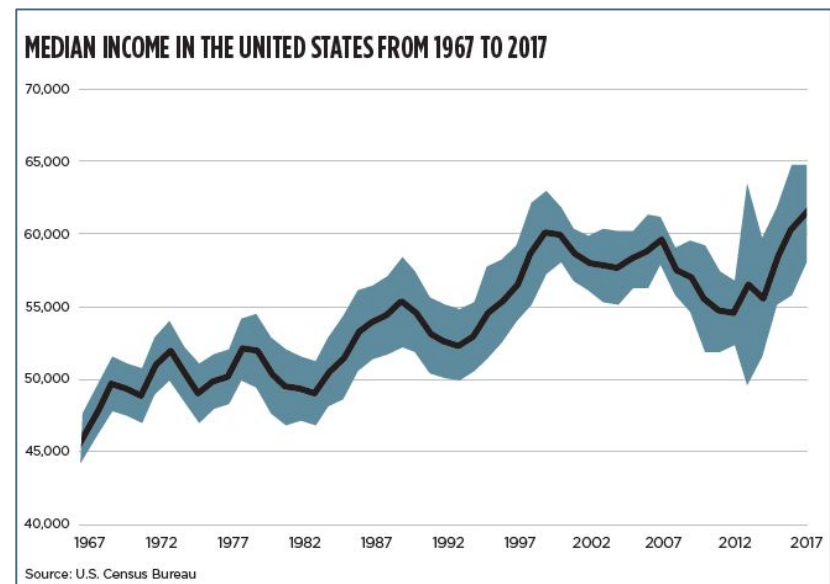


Data often contains (**unintentional**) uncertainty. Be clear about this uncertainty:

- Always note what percent of students completed the task/survey
- If possible, include error bars/confidence intervals to show variability
- If possible, display the entire distribution of the data for context



Distribution of data. Means indicated by bar.

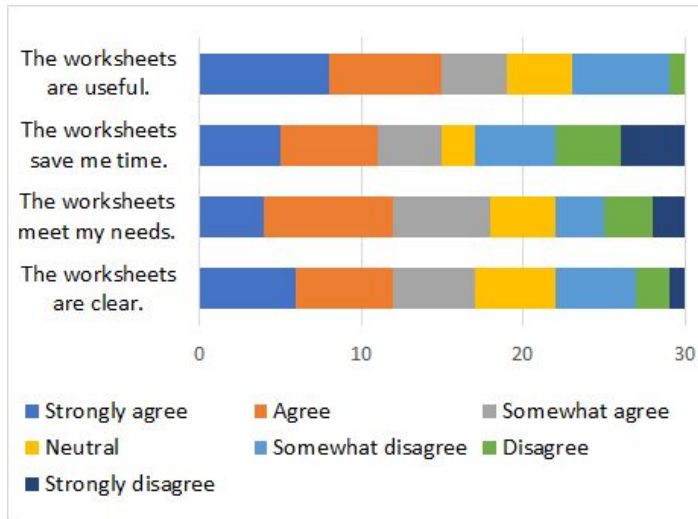


Linear line plot of data medians. Confidence intervals indicated by blue area.

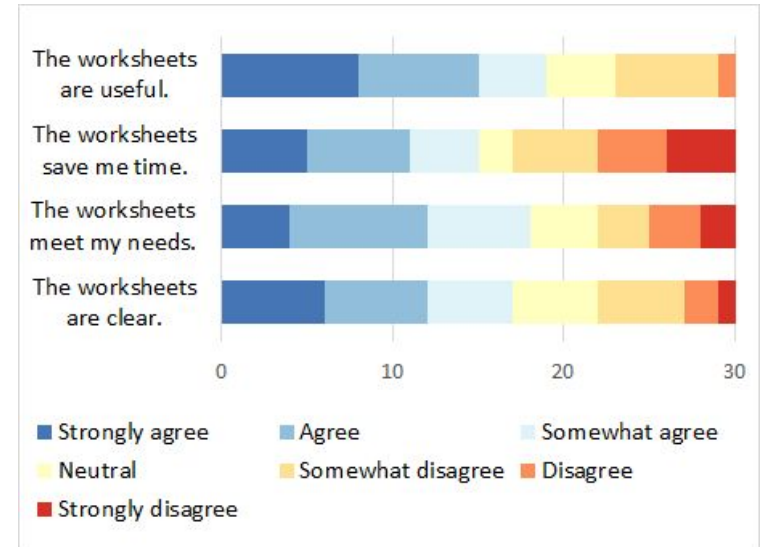
Best practices: Considering colour for scales



Instead of using mixed colours for an ordinal scale (e.g., Likert):



Use **divergent colours**:

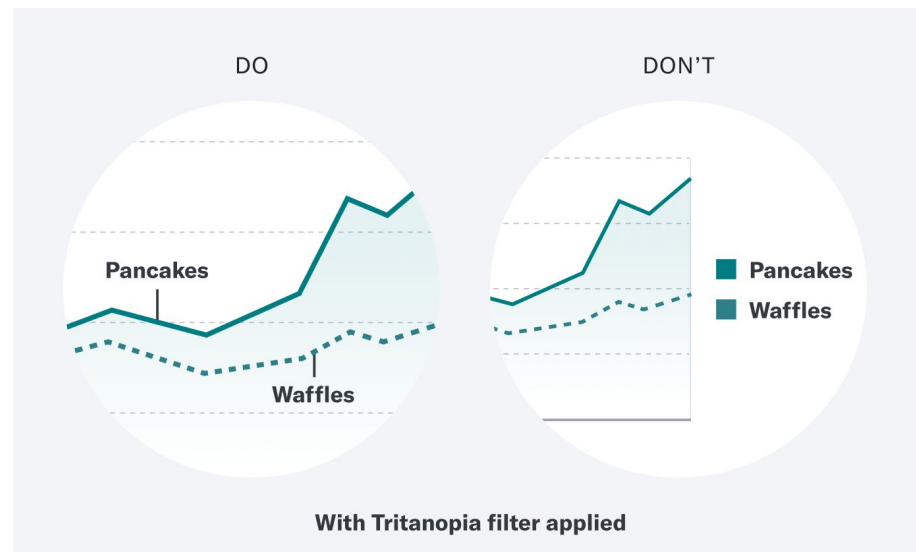


Use a **sequential colour** scale for continuous data, to show increase in numbers:



Use a **mix of colours** only for qualitative/categorical data (e.g., showcasing how different groups performed).

Best practices: Considering colour blindness



Adapted from:
https://www.hks.harvard.edu/sites/default/files/Academic%20Dean%27s%20Office/communications_program/workshop-materials/pp_jess_cohen-tanugi_design_principles_for_visualization_-_2-20-19.pdf

<https://www.betterment.com/design/accessible-data-visualization>

Resources



Content and examples were adapted from:

- Evergreen (2018) Effective data visualization
- Schwabish (2020) Better Data Visualizations
- Wong (2010) Guide to Information Graphics

Resources on making accessible visualizations:

- Various colour blind friendly palettes and colour codes can be found here:
[http://www.cookbook-r.com/Graphs/Colors_\(ggplot2\)/#a-colorblind-friendly-palette](http://www.cookbook-r.com/Graphs/Colors_(ggplot2)/#a-colorblind-friendly-palette)
- Trish's favorite website for choosing color schemes:
<https://colorbrewer2.org/#type=sequential&scheme=BuGn&n=5>
- “A Comprehensive Guide to Accessible Data Visualization”
<https://www.betterment.com/design/accessible-data-visualization>

Advanced qualitative visualizations:

<https://stephanieevergreen.com/wp-content/uploads/2021/07/Qualitative-Chooser.pdf>