

Tips for good visualization

Adapted from Geocenter (Laura Hughes) (source: <https://bit.ly/2XXDAUy>)

Know your audience, understand what you're trying to convey, and how they'll interact with the information (Online? In a report? On a poster?):

- Do you want to explore the data (or let others explore)?
- Do you want to foster dialogue between two parties?
- Do you want to tell a compelling story?

Simplify, simplify. Eliminate anything you don't need, like:

- Data that aren't pertinent to your story (Note: this does not mean you get to cherry-pick which data looks best, but if certain variables are ancillary to the story, don't include them. Edit but don't skew, and be transparent about what you've excluded!)
- Make sure colors have a clear meaning and purpose. Don't go overboard.
- Tick marks and axis labels. Use as many as you need to make it clear, but don't go overboard.
- Excess numerical digits (your axis does not need to be labeled \$1.894327849 B; \$1.9 B will work just fine).
- Remove unnecessary grid lines, boxes, and tick marks.
- 3D charts or embellishments: they almost always make things more confusing.

Fancy software is not needed to make good visualizations (but it can help).

People should be able to understand the whole story by looking at the abstract, figures, and figure captions. All of the following should be good, descriptive, and BIG (it must be readable):

- Titles
- Captions
- Annotations
- Axes (with units!)
- Legends (if you have to have one)
- Scale bars (like a temperature gauge)
- Sources for your data, if it's from someone else

Label things directly wherever possible. Don't make the reader work to figure out what the orange star means.

Avoid uncommon acronyms wherever you can. If you have room to spell it out, and it's not something obvious like USAID, do it! This prevents the reader from having to do unnecessary work to search within the text to find what the acronym means.

Be consistent wherever you can:

- Axis limits: if you compare two plots, make sure the x-axis and y-axis have the same range.
- Have colors mean the same thing everywhere.
- Start axes at zero unless you have a good reason not to.

Order your variables in a meaningful way.

You can also group and transform data. Convert to millions; create a new ratio; calculate a percentage.

Don't mislead people with your representation.

- Don't cherry-pick data. If you exclude data, have a justifiable reason that is more than "It makes the figure look better."
- Cite where you got your information.
- Be clear if info is zero or if it's missing.
- Highlight any limitations of the data or uncertainty in their measurement, such as using error bars.

Sketch and try out different options. Every chart type, color, symbol, and placement has its own merits; see what works for your data and your story.

Most importantly: Good data visualizations will not save you from bad data.