

Data Visualization Field Guide for Education

Jared E. Knowles

March 2015

Tufte's Tips

- Have a properly chosen format and design
- Use words, numbers and drawing together
- Reflect balance, proportion, and a sense of relevant scale
- Display an accessible complexity of detail
- Have a narrative quality, tell a story about the data
- Draw in a professional manner, with the technical details done with care
- Avoid content-free decoration, including **chartjunk**

Planning

Data visualization is a tool for communicating a **specific feature** of a data set in an approachable and efficient manner.

- Who is the audience? What is their background? Their biases?
- What is the context? What would the user consider to be *good news*, what would be *bad news*? How different are these?
- What does the design style say?
- Are you designing this visualization for a specific data set, or to display different data in a dashboard environment?
- How long will this visualization last? Will it be updated next year? Does it need to be interactive?
- What visual elements best map onto the data I have – shapes, sizes, colors, or fill?

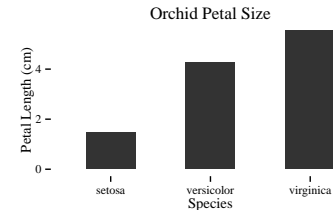


Figure 1: Sometimes the story is very simple.

Data visualization is done in service of the audience. If the goal is not to convey information, it is art!

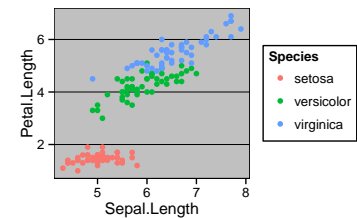


Figure 2: Theme says a lot. This says: 'I am in a hurry'.

How you turn dimensions in the into visual cues for your audience is everything.

Diamond Size and Price by Color and Cut

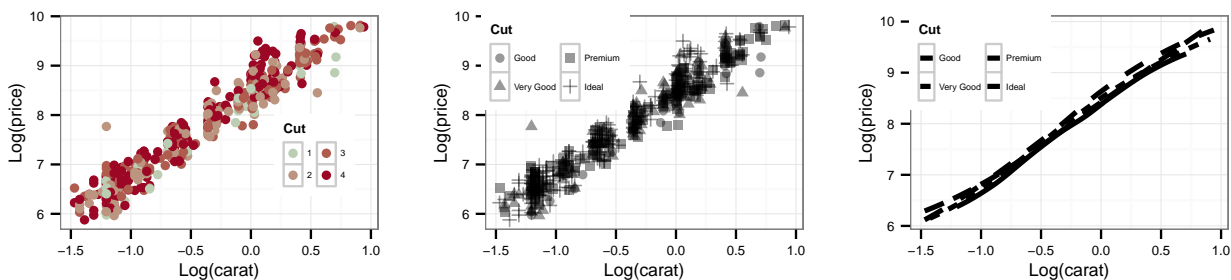


Figure 3: Shape and color work well for discrete data without too many levels, but one may work better than the other.

Dimensionality

If you got big data, I feel sorry for you son. you got 99ⁿ problems but data viz ain't one. ¹

¹ (adapted from Jay-Z, 2003)

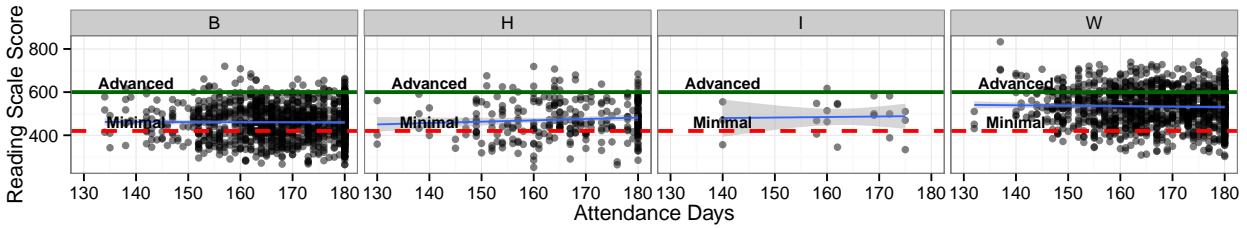


Figure 4: Add reference lines and annotate your charts with text. Guide the viewer, do not assume. Create visual context to invite and aid comparisons for the user.

Review

Before you submit, check that you have:

- **Axis labels** and a **title**: Is your purpose clear?
- A **legend**: Are all symbols explained?
- A **scale**: Is the relation of data points clear?
- **Annotations**: Is there enough context? Are key points highlighted?

Remember: If a picture is worth a thousand words, a good data visualization must always be *better than a table*.

Review this with *fresh eyes*. Can someone who has not seen this data before understand and interpret the main story here?

Technical Details

The final format and use case should inform the design:

- Will this be shared on the web, in a document, or in a presentation?
- Will the users want to interact with the output?
- Is the visualization intended to be shared and used widely, or is it for a short exploration?
- Do others need to be able to manipulate this image?

Raster files (.jpg, .png, and .gif) have a fixed scale, aspect ratio and size – not good for resizing or reshaping. Easy to share.

Vector files (.pdf and .svg) are zoomable and adjust to scale/aspect ratio changes. Viewable in browsers, but not able to be opened in PowerPoint, Word, etc.

Learn More

www.jaredknowles.com • <http://forums.datacope.org> • <http://eagereyes.org/about> • <http://datastori.es/>

Don't be afraid to model the data. A line of best fit or a smoother can provide a good summary to viewers to make comparisons across panels.

If it is not clear, consider encoding a key feature twice – e.g. using color and shape.

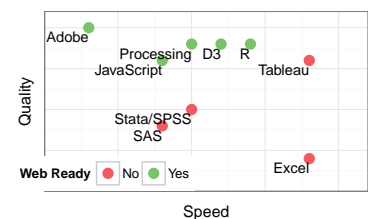


Figure 5: An opinionated ordering of the tools available to do data viz. Not comprehensive or complete and your mileage will vary depending on your comfort with certain technologies.

If sharing matters most, go raster, if quality matters most consider vector files.