Handout for 'How to Display Data Badly' by H. Wainer
Stat 696: Statistical Communication in Data Science

Andrew Bates, Josh Gloyd

Rules for Displaying Data Badly

The aim of good data graphics is to (A) Display Data, (B) Display Data Accurately, and (C) Display Data Clearly. Here are Wainer's twelve rules for displaying data badly along with the categories they fall under. Note: These rules tell us how to display data *badly* so these are rules that we should be breaking!

- **Showing Data**
    1. Show as Few Data as Possible (Minimize the Data Density)
    2. Hide What Data You Do Show (Minimize the Data-Ink Ratio)
- **Showing Data Accurately**
    3. Ignore the Visual Metaphor Altogether
    4. Only Order Matters
    5. Graph Data Out of Context
- **Showing Data Clearly**
    6. Change Scales Mid-Axis
    7. Emphasize the Trivial (Ignore the Important)
    8. Jiggle the Baseline
    9. Austria first!
    10. Label (a) Illegibly, (b) Incompletely, (c) Incorrectly, and (d) Ambiguously
    11. More Is Murkier: (a) More Decimal Places and (b) More Dimensions
    12. If It Has Been Done Well in the Past, Think of Another Way to Do It

Rule 9 – Austria First! may be hard to remember by name so perhaps a quote from the paper will help:

> Ordering graphs and tables alphabetically can obscure structure in the data that would have been obvious had the display been ordered by some aspect of the data.

So let's say you're making a bar chart with several variables. Then it makes sense to order them smallest-to-largest. But ordering them alphabetically might conceal information and is not really based on the data itself.

Wainer summarizes by saying:

> Examine the data carefully enough to know what they have to say, and then let them say it with a minimum of adornment.

You don't need to reinvent the wheel when preparing graphics for your writing. Know your audience and then use graphics to convey to them what your data is saying clearly and accurately.