**Good vs. Bad Data Displays**

Marlee Salisbury

PSYC6135

A graph with lines and text

Description automatically generated with medium confidenceA graph with lines and numbers

Description automatically generated with medium confidenceAs we learned in the first (thought-provoking) class, good visualizations are clear and precise in their communication of data. They emphasize comparisons meaningfully, avoid clutter or unnecessary information, and eliminate the need to “scratch your head” by communicating data efficiently, effectively, and—ideally—parsimoniously without compromising the data integrity. As I explored some literature in my research area, I came across the following “bad” graphs that even a lay person may see as rudimentary (perhaps even lazy).

Figure 1 above is a line graph comparing boys and girls externalizing trajectories over time. While the intent of the graph is clear—to show differences between boys and girls—its execution is overly simplistic. The graph provides no details about the statistical significance of the trends, what the values on the Y-axis mean with respect to levels of externalizing behaviours, and what each time point represents (1 year? 5 years? 10 years?), leaving me with an incomplete understanding of the data. It does not convey any additional insights beyond what could be summarized in a sentence in the manuscript. Additionally, the plain design, including the inefficient use of space (e.g., significant white space around the data), fails to engage the viewer and ultimately reduces its communicative value. The use of a legend also seems unnecessary given there are only two lines; labels beside the lines or the use of colour would have been more effective.

Figure 2 is similarly problematic. This graph is intended to communicate how different attachment styles relate to externalizing behaviours over time. In addition to the points made about Figure 1, its use of dashed lines without distinct thicknesses or colour makes it difficult to differentiate between the categories—again, likely solved by using line labels rather than a legend (which, by the way, who approved the lines merging into the legend??). There are no annotations or explanations, so it is unclear exactly what the trends mean (are any of the slopes significant? Differences between the slopes? And, if so, what does this mean about each attachment style?).

A graph showing the number of patients with covid-19

Description automatically generatedThankfully, I came across some much more effective data displays. One great example is the correlation matrix in Figure 1 below. The purpose of this matrix was to show relationships between the various variables in the study. It effectively used a colour gradient—green for positive correlations and orange for negative ones—which provided an intuitive visual cue for interpreting the strength and direction of the correlations, even at a glance. As well, each cell includes exact correlation values which makes the matrix precise. There is effective use of a legend, annotations (e.g., about significance), and additional information about the data (e.g., sample size) to add context without adding unnecessary “chart junk” that might detract from the data. Overall, it was a great example of how to use various visual elements to effectively and efficiently represent the data and make comparisons easy on the viewer.

A diagram of a graph

Description automatically generated with medium confidenceAnother example of a good visualization is a regression coefficient plot summarizing the results of a multiple linear regression analysis in Figure 5. Specifically, it highlights the strength and direction of relationships between the predictors and the outcome variable (loneliness during the pandemic). What made it effective was the inclusion of regression and significance values directly on the plot, eliminating the need to cross-reference these details in the text or tables. Additionally, confidence intervals are displayed as horizontal bars, which allows the viewer to visually understand the uncertainty around each coefficient. The design was clean and minimalist—no heavy grid lines or unnecessary details—which helped to keep the focus on the data and ensures the data can be easily interpreted by the viewer. Overall, this visualization seemed intentional, utilizing clarity, precision, and aesthetic appeal to achieve a visually effective and engaging display.

**References**

Doom, J. R., Seok, D., Narayan, A. J., & Fox, K. R. (2021). Adverse and benevolent childhood

experiences predict mental health during the COVID-19 pandemic. *Adversity and resilience science*, *2*, 193-204.

Warmuth, K. A., Cummings, E. M., & Davies, P. T. (2022). A prospective longitudinal study of

mother–child attachment and externalizing trajectories in boys and girls. *Child Psychiatry & Human Development*, 1-12.