

This article argues that using well-designed graphs is superior to using tables to display numerical information when making comparisons, which is a primary motivation in statistical research. The authors first review previous work on data representations, and then discuss the key elements in designing a well-displayed graph to present data comparisons. Furthermore, a few examples from JASA publications are provided to demonstrate the advantages of using graphs over tables for a better display of numerical data. In the end, the authors emphasized that a well-designed graph is more informative and straightforward to read, though it may involve more effort to be constructed. They also list key steps to construct graphs: identification of key comparisons; application of comparison lines or axes on small plots; and establishment of control over the display.

I think it is common sense that graphs are more readable than tables in general. In my weekly progress report, I can tell that the PIs are happier to see well-explained graphs rather than listing everything in a table. The authors provide solid examples to demonstrate the advantage and necessity of using graphs to display numerical comparisons in statistical research. Though I agree that graphs are preferred to display numerical data in general, I would like to see scenarios when graphs do not have advantages to visualize comparisons. One example I can think of is a confusion matrix: though we may use heatmaps to show the magnitude of the values, it could be challenging to merge/stack multiple confusion matrices together. In this case, I guess it should be OK to just show a concatenated confusion matrix in its tabular form and only plot the F1 scores when multiple confusion matrices need to be shown. In addition, I have a couple of comments/questions on the examples: 1) Is it acceptable to combine subplots of Figure 5b using scatter plots when the same model B is the baseline? 2) In Figure 6b, is it better to show mean values to the plots?