

Tutorial: Scientific Data Visualization: Do's and Don'ts

Prof. Max Ruzz

Introduction

Effective data visualization is not about making data look fancy; it is about telling the truth, clearly, efficiently, and memorably. Below you will find five pairs of visualizations,—each showing a **good** and a **bad** example—together with brief comments.

1. Histogram + Density Curve

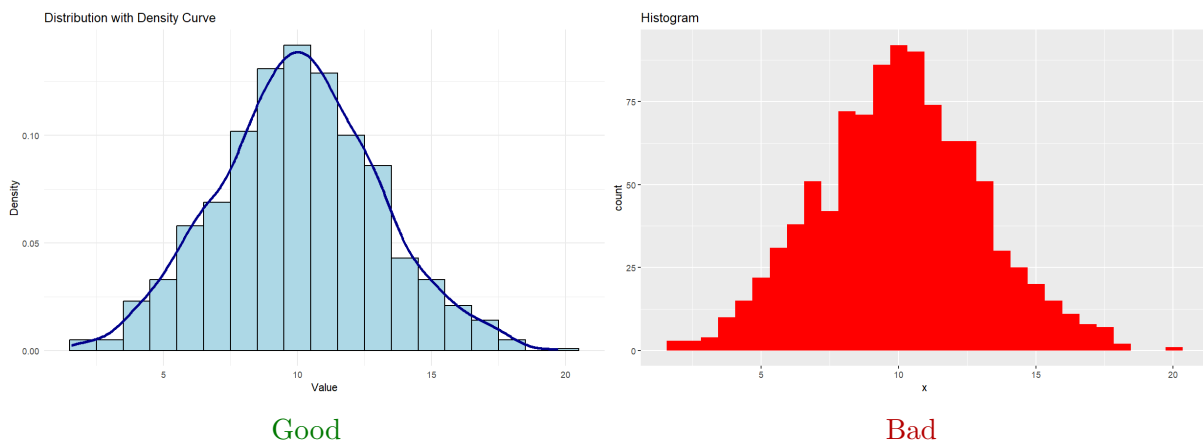


Figure 1: Overlaying a density curve (left) clarifies the distribution, whereas the right plot omits context and uses an aggressive colour.

2. Volcano Plot

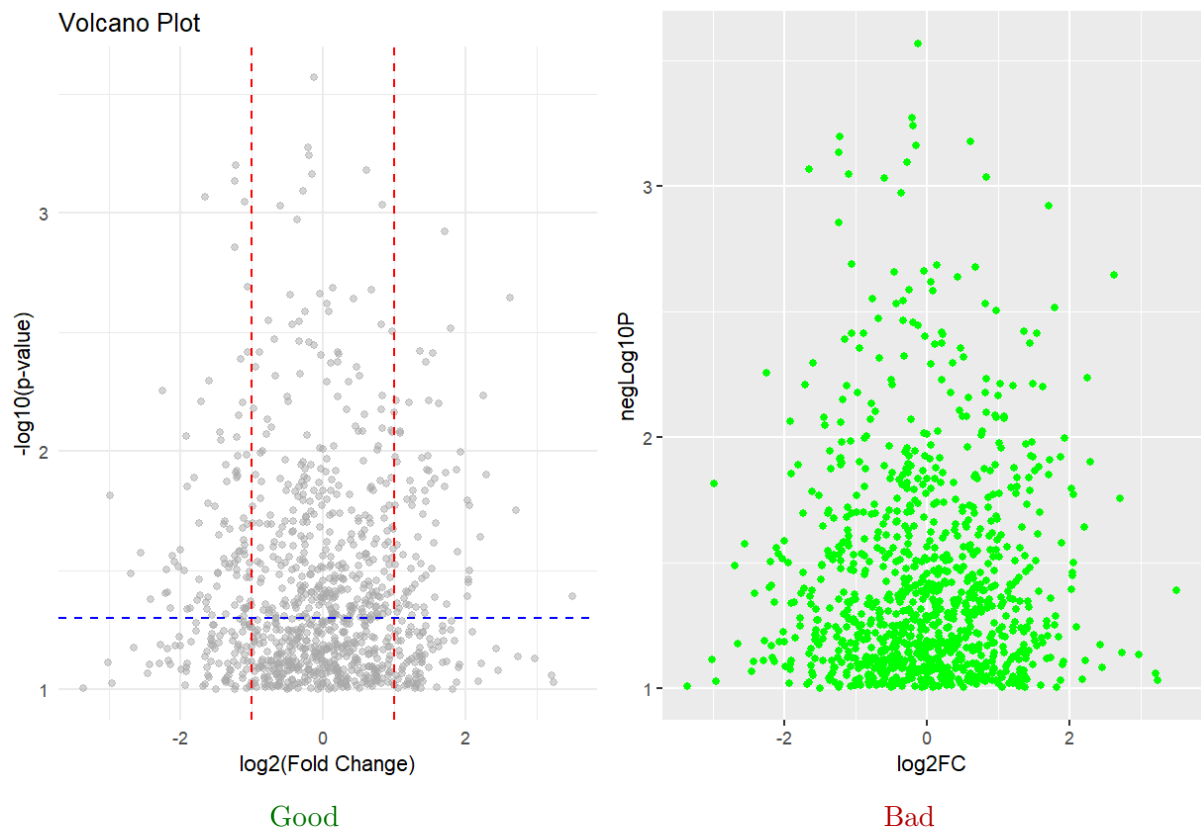


Figure 2: The good volcano plot displays significance thresholds and clear colouring; the bad example lacks context and misleads.

3. Gene Expression Heatmap

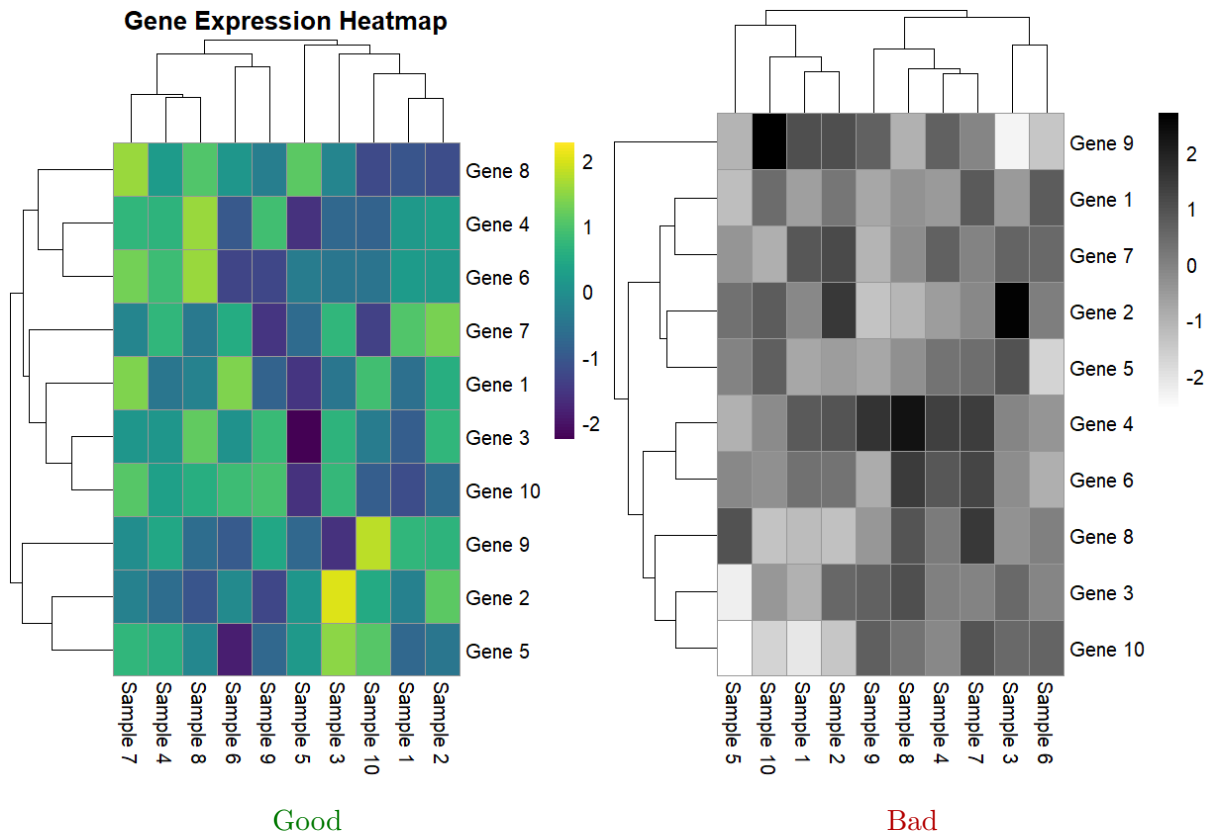


Figure 3: Left: normalised rows, perceptually uniform palette. Right: raw values and black–white gradient obscure patterns.

4. Boxplot with Raw Points

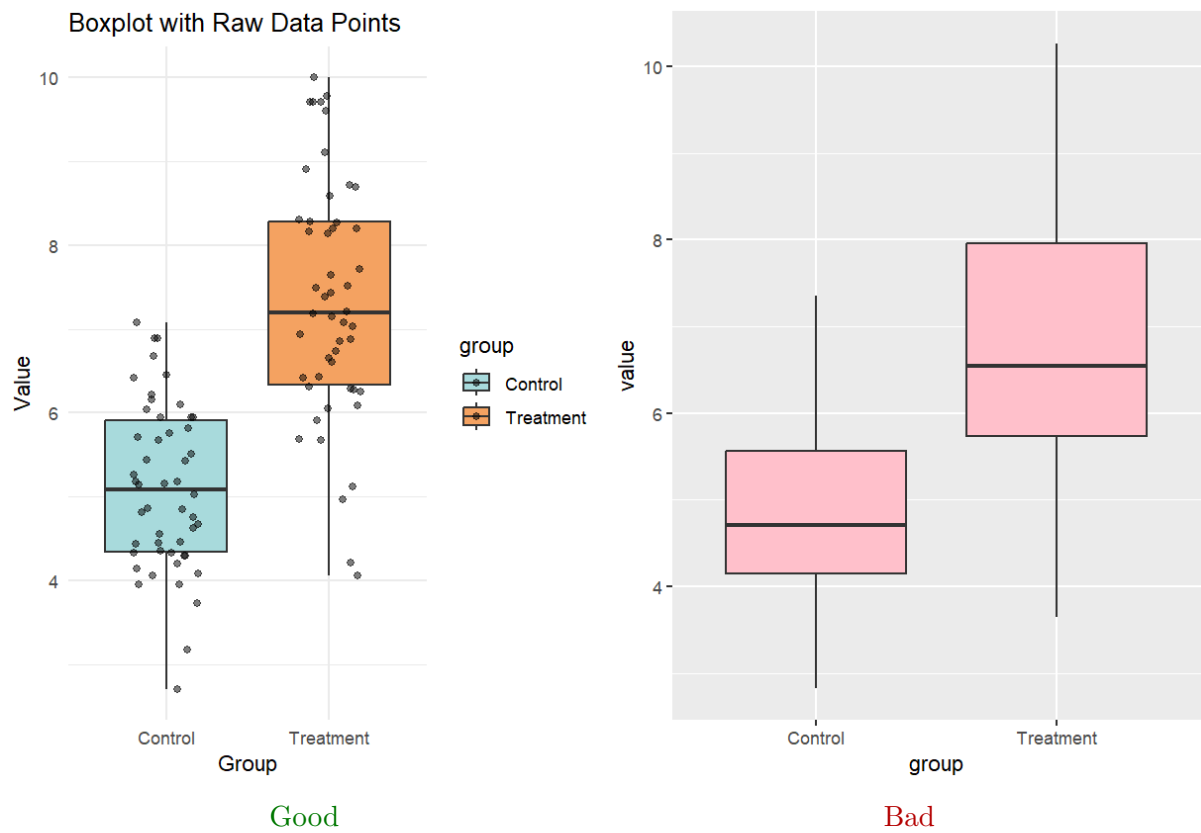


Figure 4: Combining a boxplot with jittered points (left) reveals distribution and avoids hiding outliers. Right plot hides raw data and uses distracting colour.

5. Barplot with Standard Error

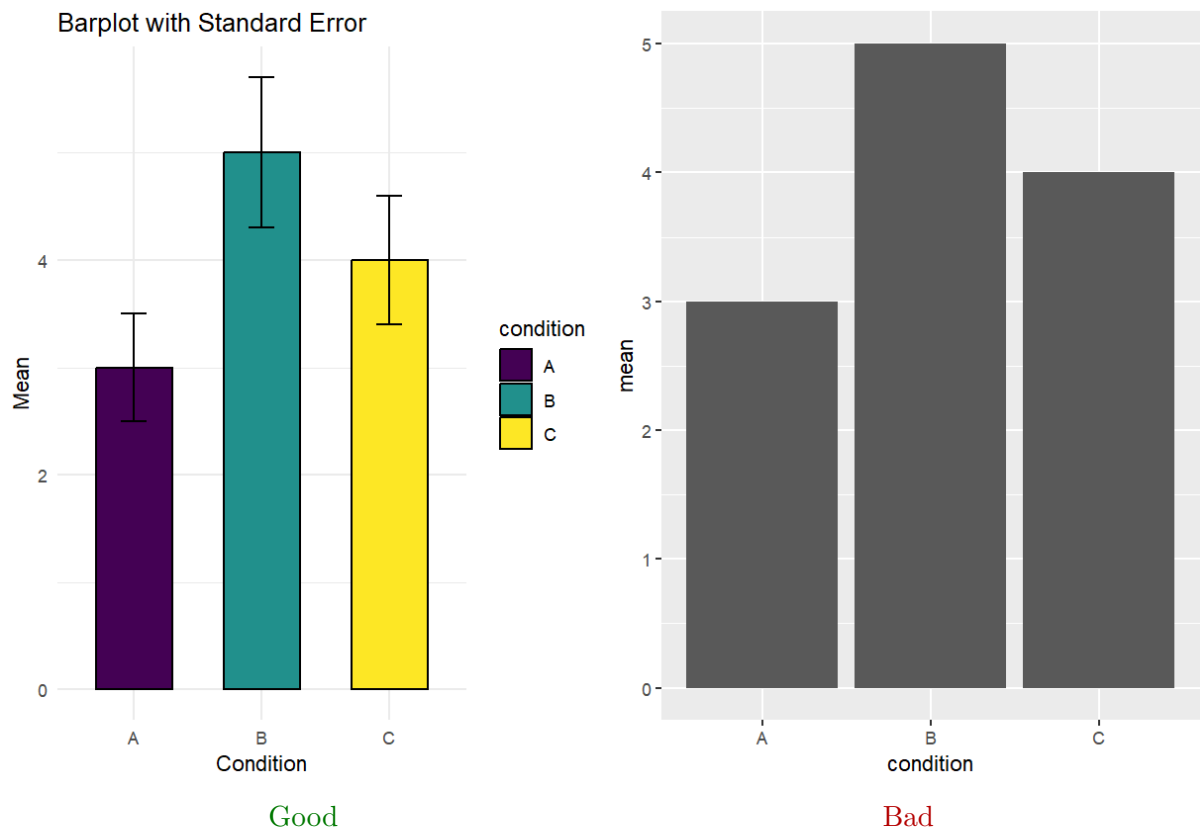


Figure 5: Standard error bars (left) communicate uncertainty; the right plot omits them, encouraging over-interpretation.

Summary:

- Always label axes and include units.
- Use perceptually uniform or publication-appropriate palettes.
- Show variability (density curves, raw points, error bars) whenever possible.
- Avoid visual clutter and misleading axis limits.
- Start with the data, not the graphic.

Happy plotting ;)