Dec 16, 2024

Data Visualization

Week 10. The principles of plot design II

Reminder

There are three main principles of plot design:

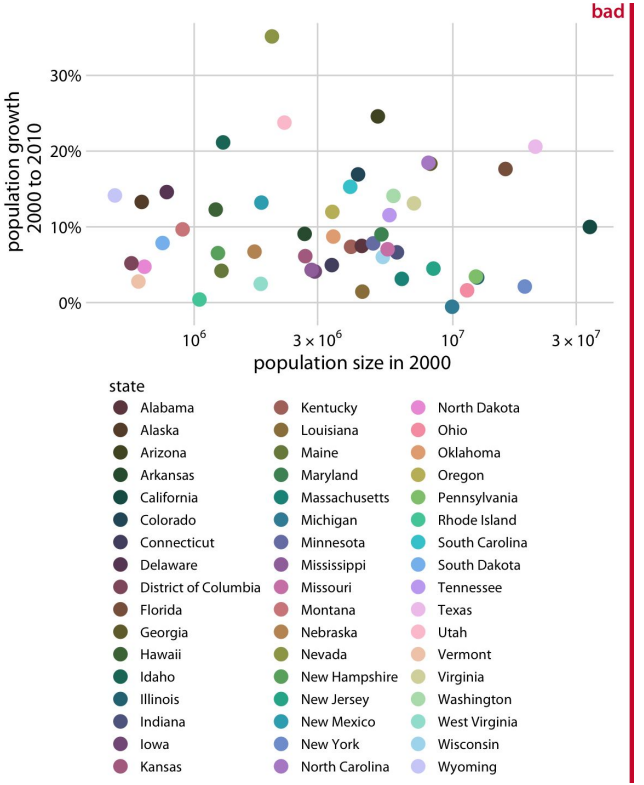
- proportionality
- overlapping points
- usage of color

Common Mistakes in Color Usage

- Color is the most effective aesthetic used for data visualization.
- Incorrect color choices can ruin even the best visualizations.
- To avoid such situations, color should be used only when it serves a purpose, it should be clear, and it should not distract from the data.

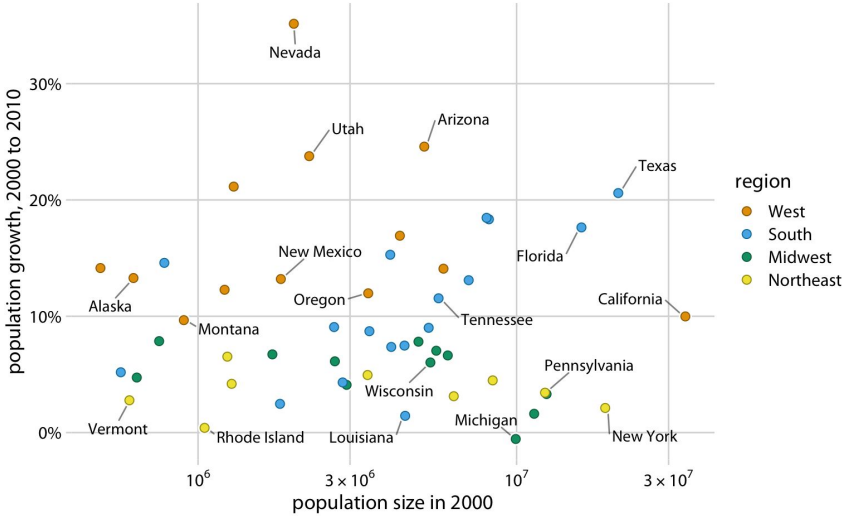
Common Mistakes in Color Usage

Overcoding information



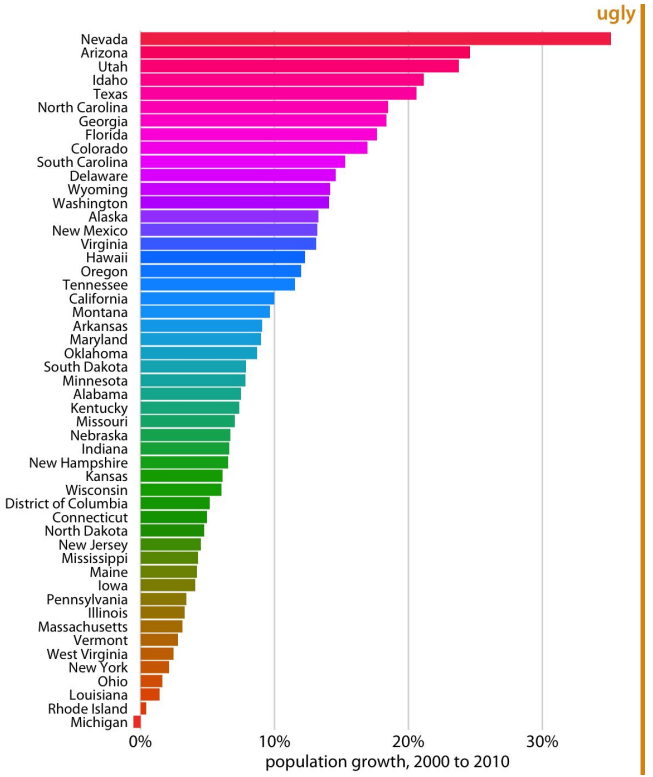
Common Mistakes in Color Usage

- As a general rule, categorical color scales work best when there are three to five different categories to color.
- When there are eight to ten or more categories, the task of matching colors to categories becomes challenging, even if the colors are distinct enough to be easily differentiated, making the visualization less useful.
- When a categorical variable has more than eight levels and differentiation between levels is necessary, direct labeling may be preferred over using colors.



Common Mistakes in Color Usage

Using Non-Monotonic Color Scales

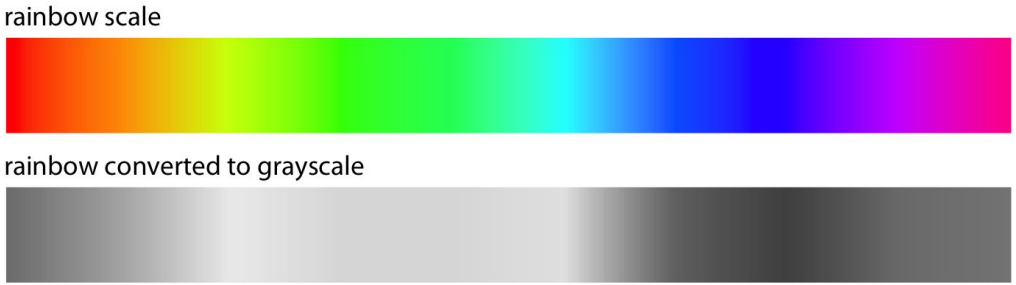


Common Mistakes in Color Usage

Two critical conditions for designing color scales:

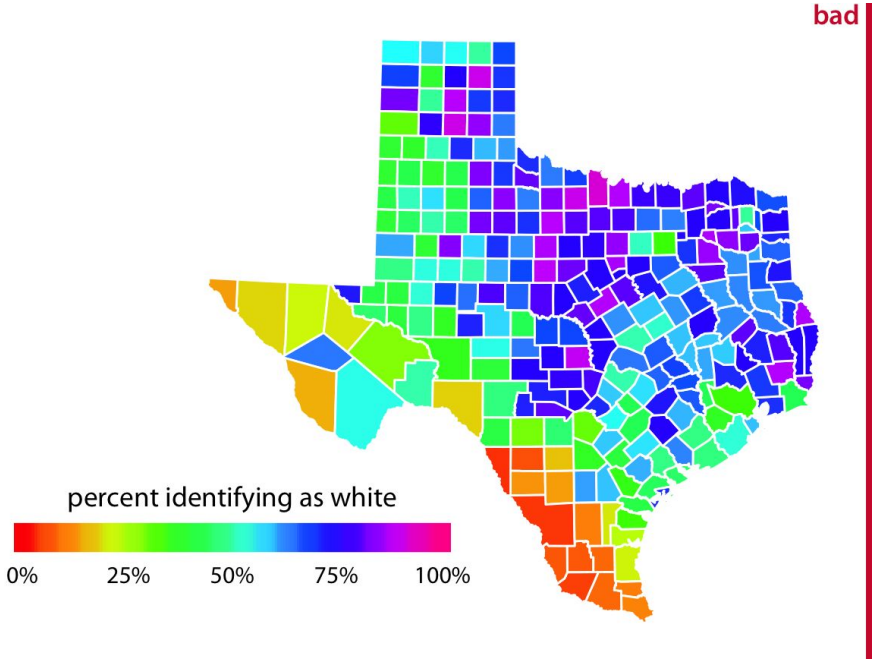
- Colors must clearly indicate which observation values are larger or smaller than others.
- The differences between colors should accurately represent the corresponding differences between values.

Many existing color scales may violate one or both of these conditions. The most well-known of these scales is the rainbow scale.



Common Mistakes in Color Usage

- In visualizing real data, the rainbow scale tends to hide data characteristics and/or emphasize arbitrary aspects of the data.
- Because the colors are highly saturated, looking at the graph for an extended period can be visually uncomfortable.



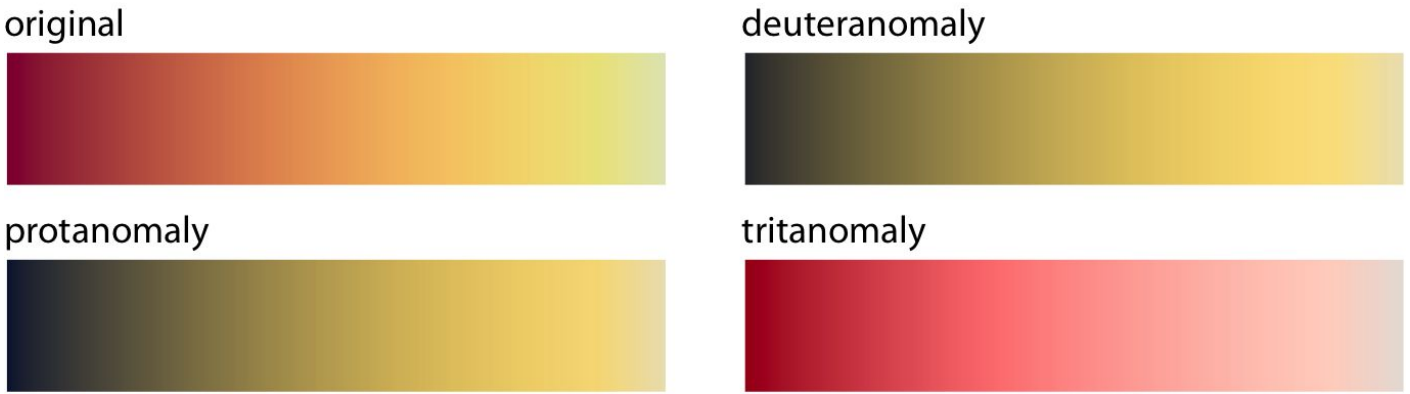
Common Mistakes in Color Usage

Color blindness

- Readers with color vision deficiency (color blindness) cannot distinguish colors that individuals without this condition can easily differentiate.
- There are three types of color vision deficiencies:
  - Red-green (deuteranomaly)
  - Blue-green (protanomaly)
  - Blue-yellow (tritanomaly)
- Approximately 8% of men and 0.5% of women experience some form of color vision deficiency. The most common issue is the inability to distinguish between red and green.

Common Mistakes in Color Usage

Red and green colors provide the strongest contrast for individuals with normal color vision, but for those with color vision deficiency, they become nearly indistinguishable.



Common Mistakes in Color Usage

Common Mistakes in Color Usage

Solutions for Color Blindness:

- Color scales specifically designed for individuals with color vision deficiency can be used.

- The colors used can be tested through a CVD (Color Vision Deficiency) simulator to ensure they are distinguishable by individuals with color vision deficiencies.

Reference

The notes and plots in the presentation are compiled from Claus O. Wilke's book, *Fundamentals of Data Visualization*.