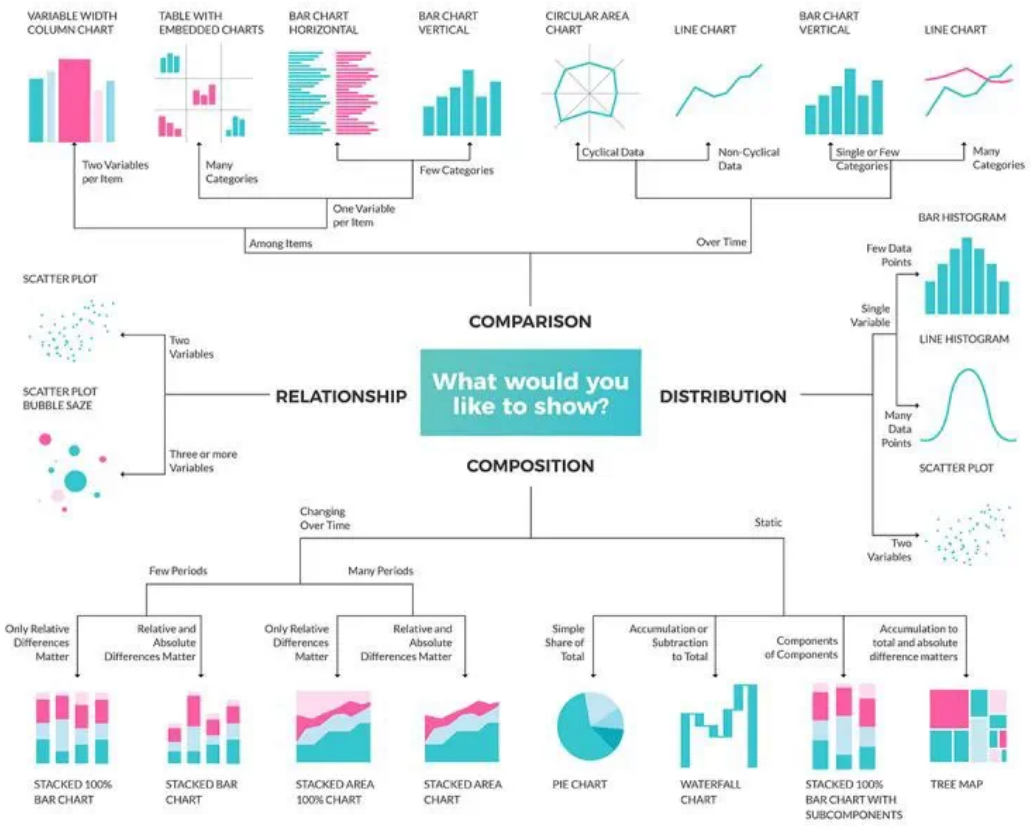
How to build data stories with data?

Data Story = Data + Narrative + Visualization

1. **Create a narrative to the target audience with an intent**
    a. Characters: Research and empathize with your audience.
    b. Settings: Set the scene by explaining your problem.
    c. Conflict: Identify root issue.
    d. Resolution: Propose your solution based on your data analysis.
2. **Finding insights**
    a. Categorize them by asking: Is this insight…
        i. BIG - Is it numerically and statically significant?
        ii. USEFUL - Is it actionable?
        iii. SURPRISING - Is it something that stakeholders didn't know?
3. **Craft a story**
    a. Draft an outline (problem, approach, solution)
    b. Start with takeaways - Add one sentence summary (message, moral)
    c. Find the analysis to support your story - Find data, charts and colors.
    d. Convert analysis to messages - Do your data analysis
    e. Structure the message - Pyramid structure (top takeaways - bottom support)
    f. Re-order your message
        i. Emotional contrast, start with bad news and then move to good ones.
4. **Design data stories**
    a. Sketch on a paper
    b. Charts
    c. Colors palettes - Enhanced comprehension, pattern recognition, emphasis, aesthetics, branding
    d. Libraries (ggplot, d3.js, matplotlib, …)

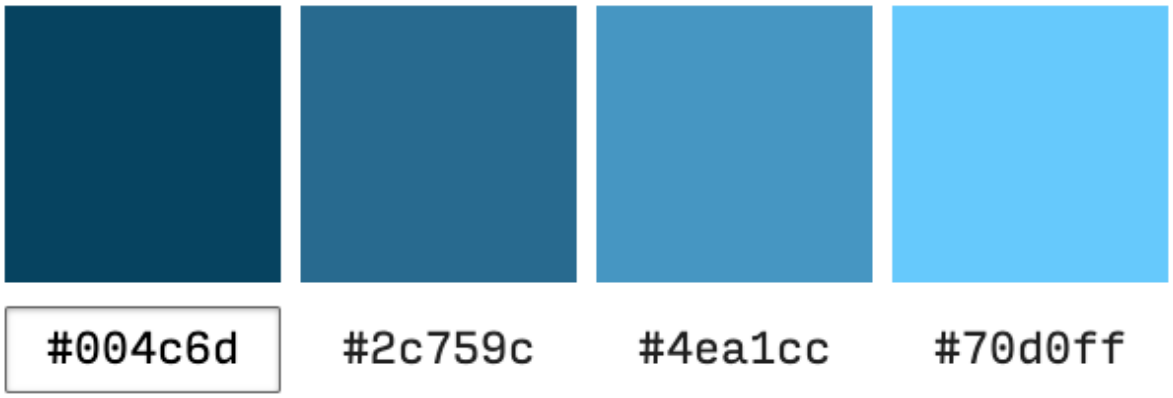
Charts: What would you like to show?



Colors: What would you like to emphasize?

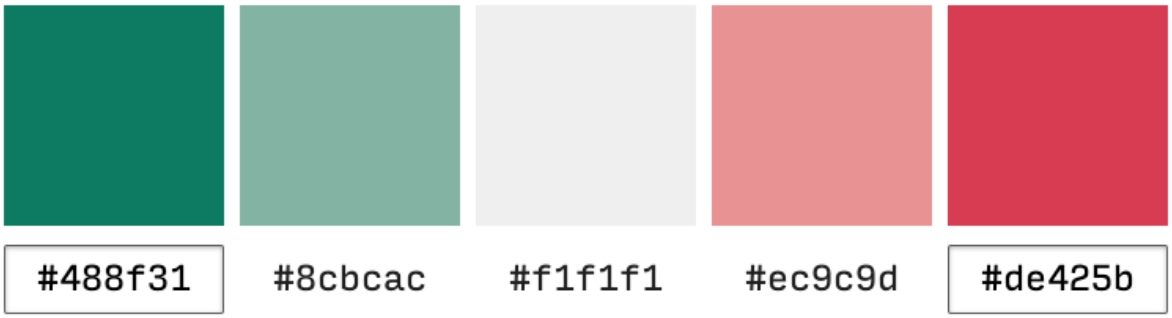
Palette

Hue



The Single Hue Scales are most useful for visualizations where you're showing the value of a single variable. Typically, the darker variation will represent a higher value, and a neutral color (even white) will represent a value closer to zero.

Divergent



The Divergent Color Scales are most useful for visualizations where you're showing a transition from (a) one extreme, through a (b) neutral middle, and finally to a (c) opposite extreme.

Libraries

https://d3js.org/ for javascript

https://ggplot2.tidyverse.org/ for R

https://matplotlib.org/ for python