

How Good People Make Bad Graphs

Unintentional Fibbing with Visualization

INTRODUCTION

Intentionally lying with visualization is easy.

Unintentionally lying with visualization is **easy**.

How do innocent motivations lead to misleading visualizations?

Everyone will and should make a bad visualization at some point.

(You just might not want to SHARE it widely.)

WARMEST YEARS

Climate Central, 2013

How 2012 STACKS UP



Potential Issues

- Range on y-axis from 53° to 55.5°
 - Bar charts should start at 0 on y-axis
 - Small ranges exaggerate small changes
- Sorting of values on x-axis by y-value
 - Unordered data can be sorted by y-value
 - Ordered data should be sorted by x-value

Potential Issues

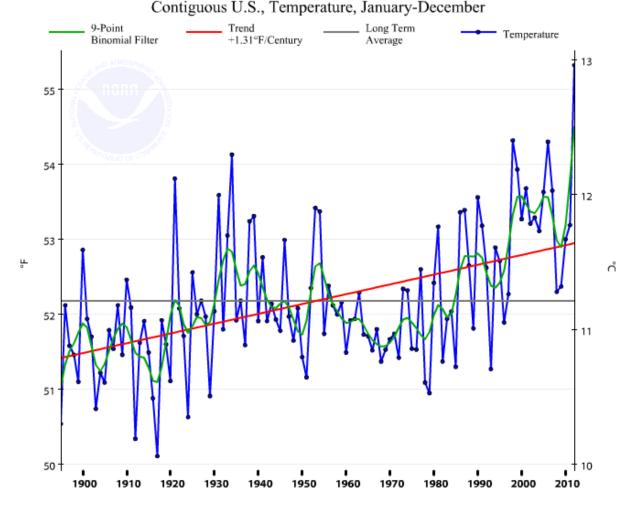
- Suggestive use of colors
 - Okay to use color to highlight data
 - Use of orange/red emphasizes warming
 - Intensity of foreground versus background similar
- Density could be improved slightly
 - Usually do not need data labels and axis lines
 - Accuracy to two decimal points on labels

Potential Issues

- Unnecessary use of 2.5D/3D hinting
 - Skewed axis forces brain to account for distance bars are away from viewer
 - Shadow and shiny color gradients suggest 3D bars which encourages comparison by volume
 - Axis lines sit "behind" the bars, requiring some adjustment to match to bar heights

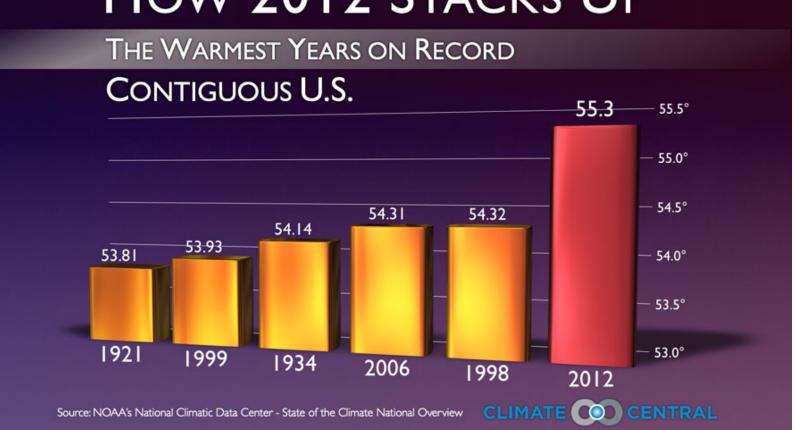
How did this happen?

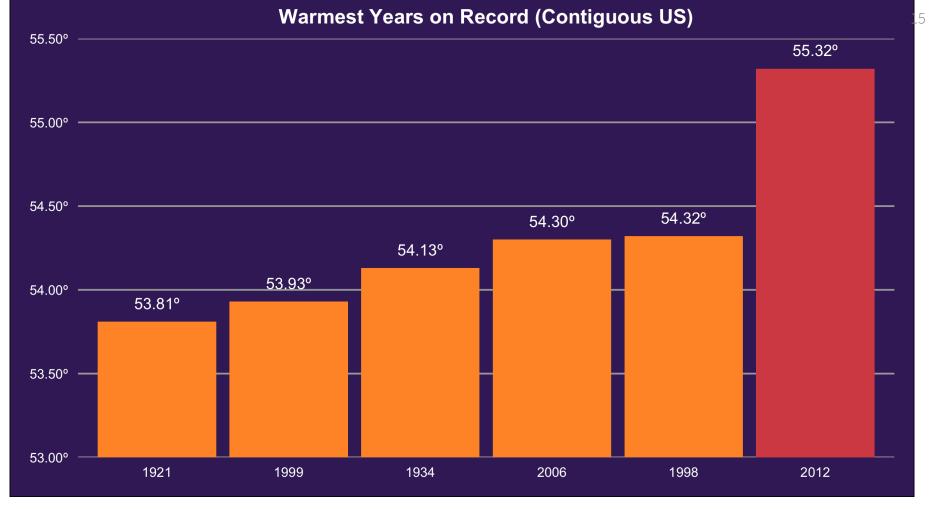
- Priority is to compare 2012 with other 5 hottest years
 - Removed context provided by other years and reordered x-axis to make comparison easier
- Wanted to emphasize change, even if small
 - Small changes in average temperature are considered a big deal
- Innocent design decisions to meet goals



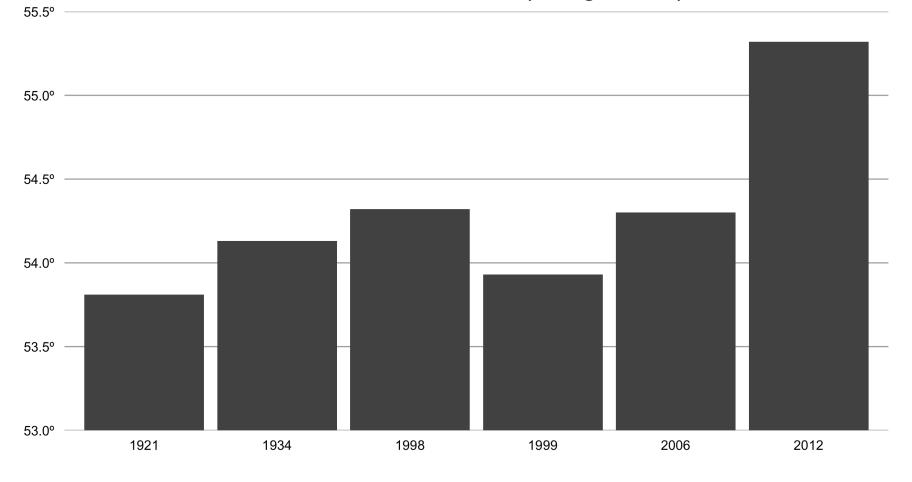
http://www.ncdc.noaa.gov/temp-and-precip/time-series/index.php?parameter=tmp&month=12&year=2012&filter=12&state=110&div=0

How 2012 STACKS UP

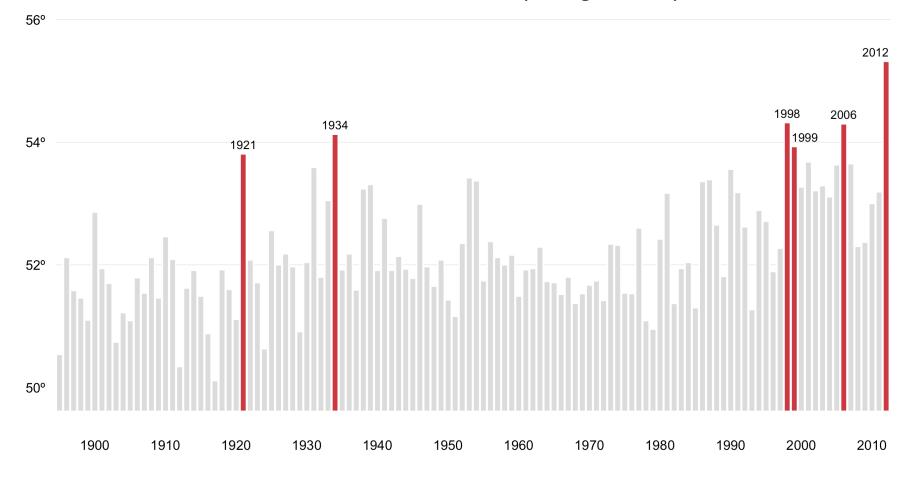




Using R and ggplot2

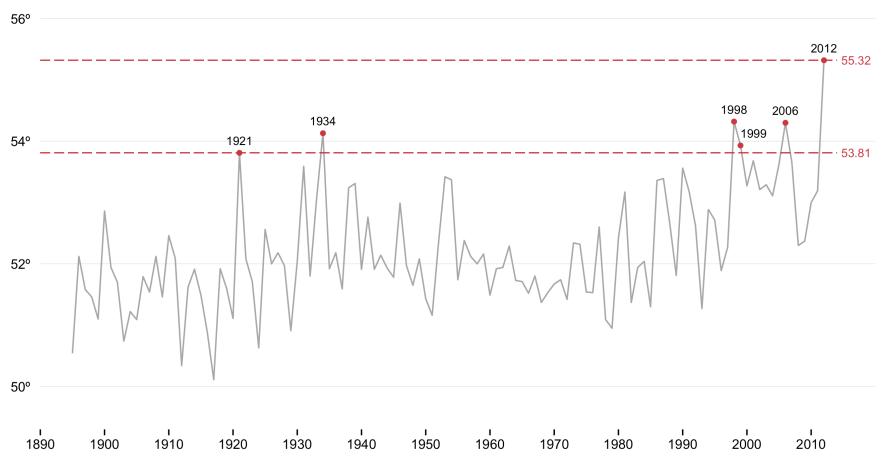


Removed sorting by temperature, unnecessary colors, and unnecessary text annotations.



Added in context of other years, while still focusing on warmest years. Still issues with bar plot not starting at 0.

Warmest Years on Record (Contiguous US)



Changed from bar to line, which is more appropriate when not starting at 0. Added focus using annotations.

QUESTIONS

http://sjengle.cs.usfca.edu/