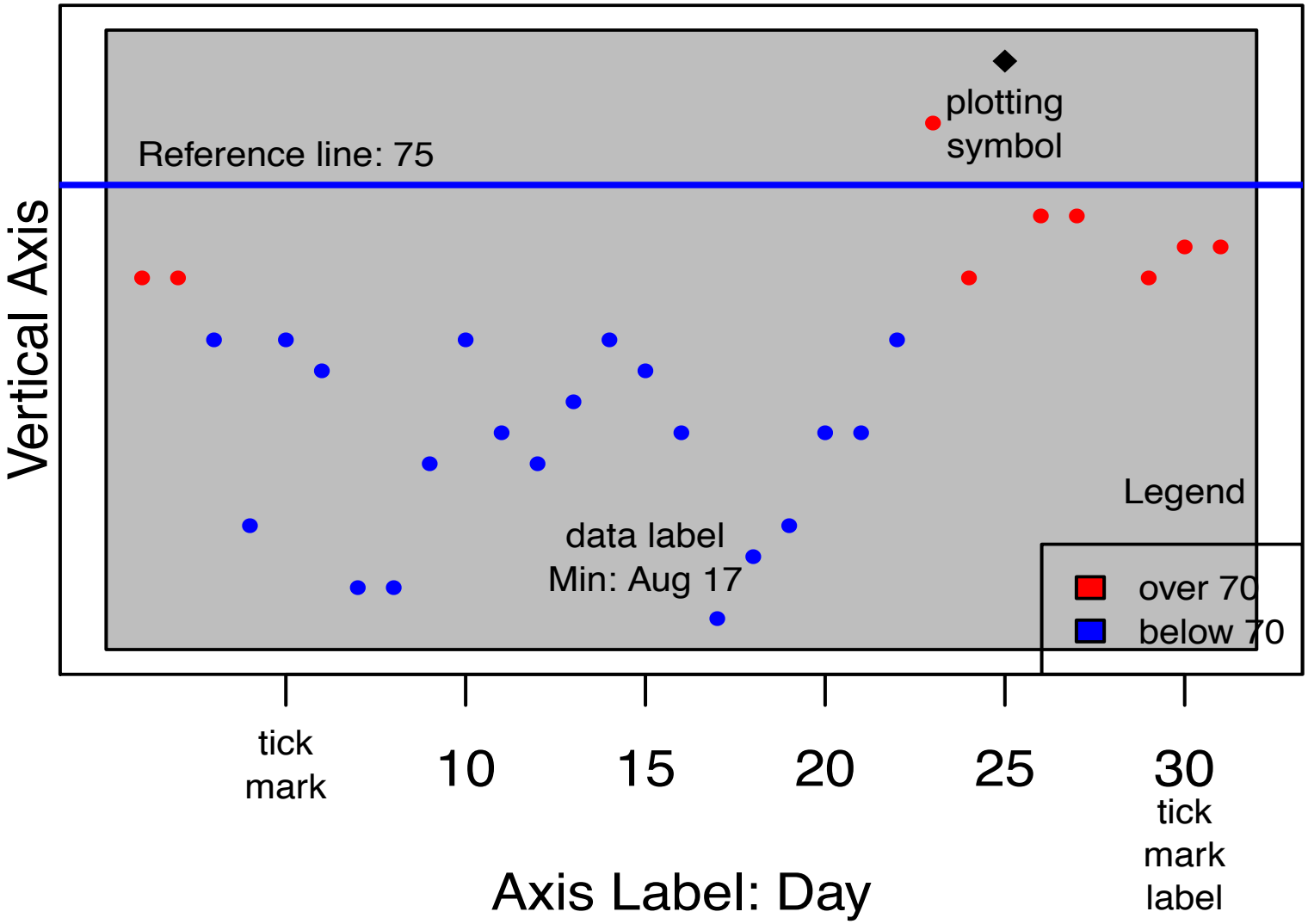
Graph Construction

Outline

- Vocabulary
- 3 Properties of good graph construction
  - Data stand out
  - Facilitate comparison
  - Information rich

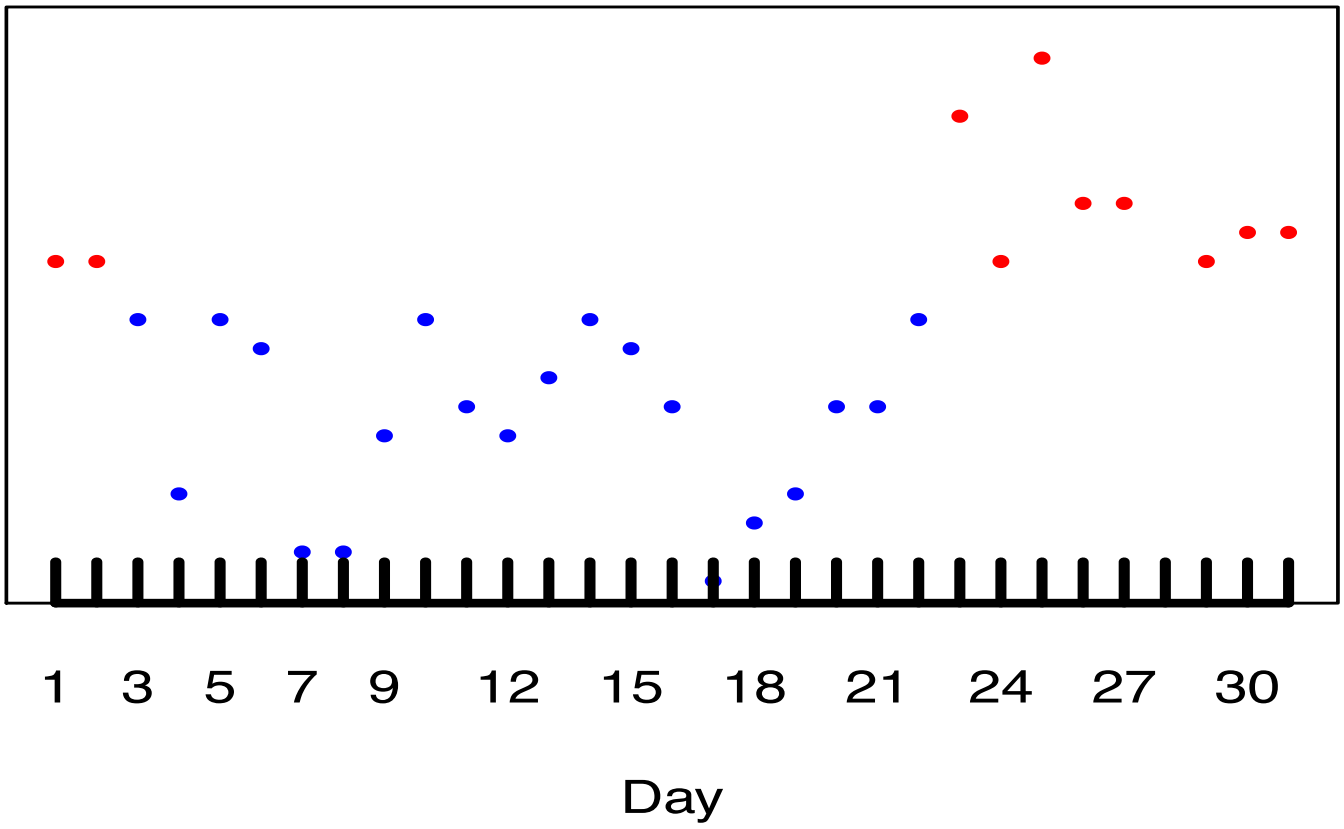
Vocabulary

Title: Temperature in August

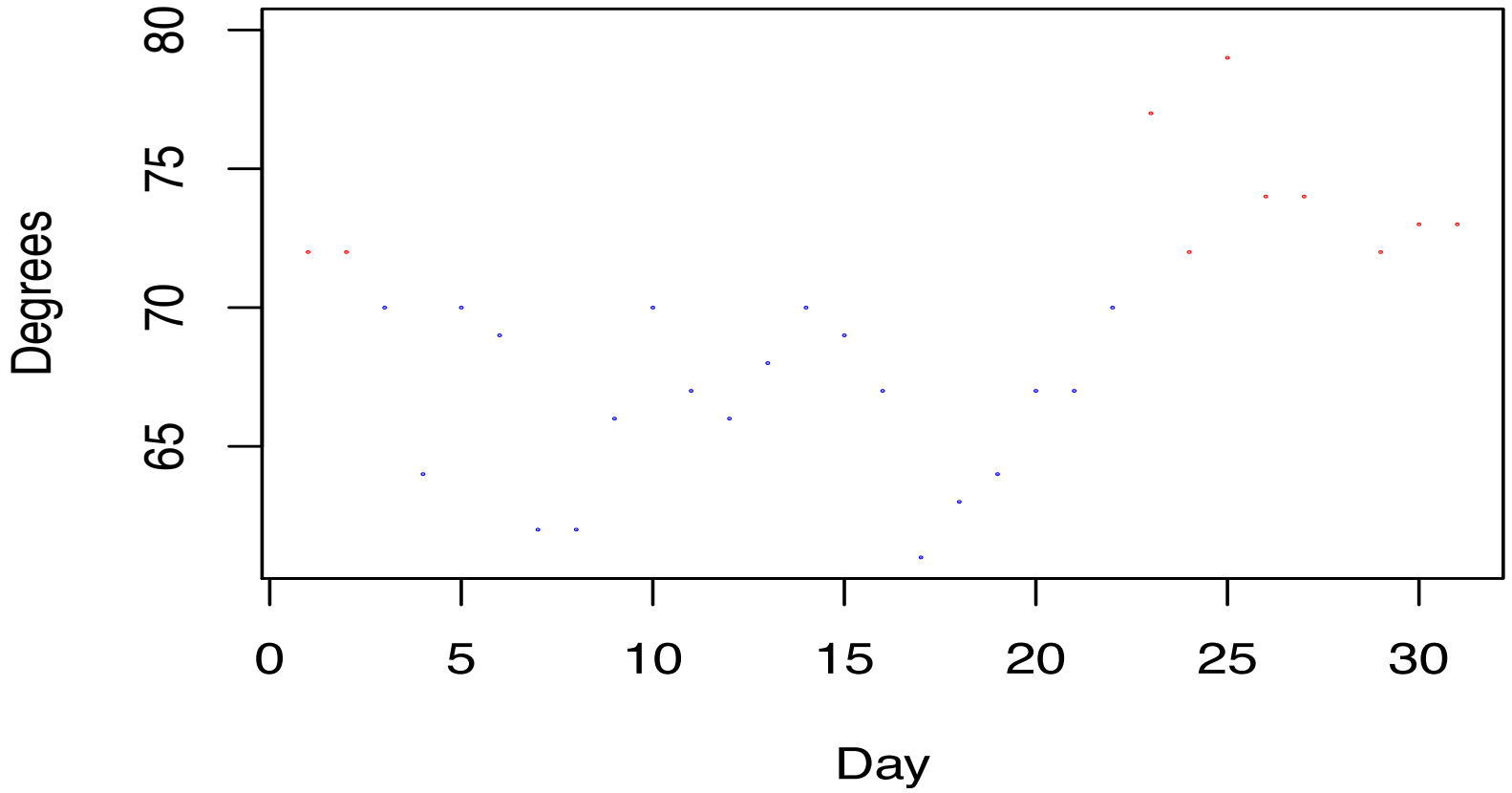


Data Stand Out

Avoid having other graph elements interfere with data

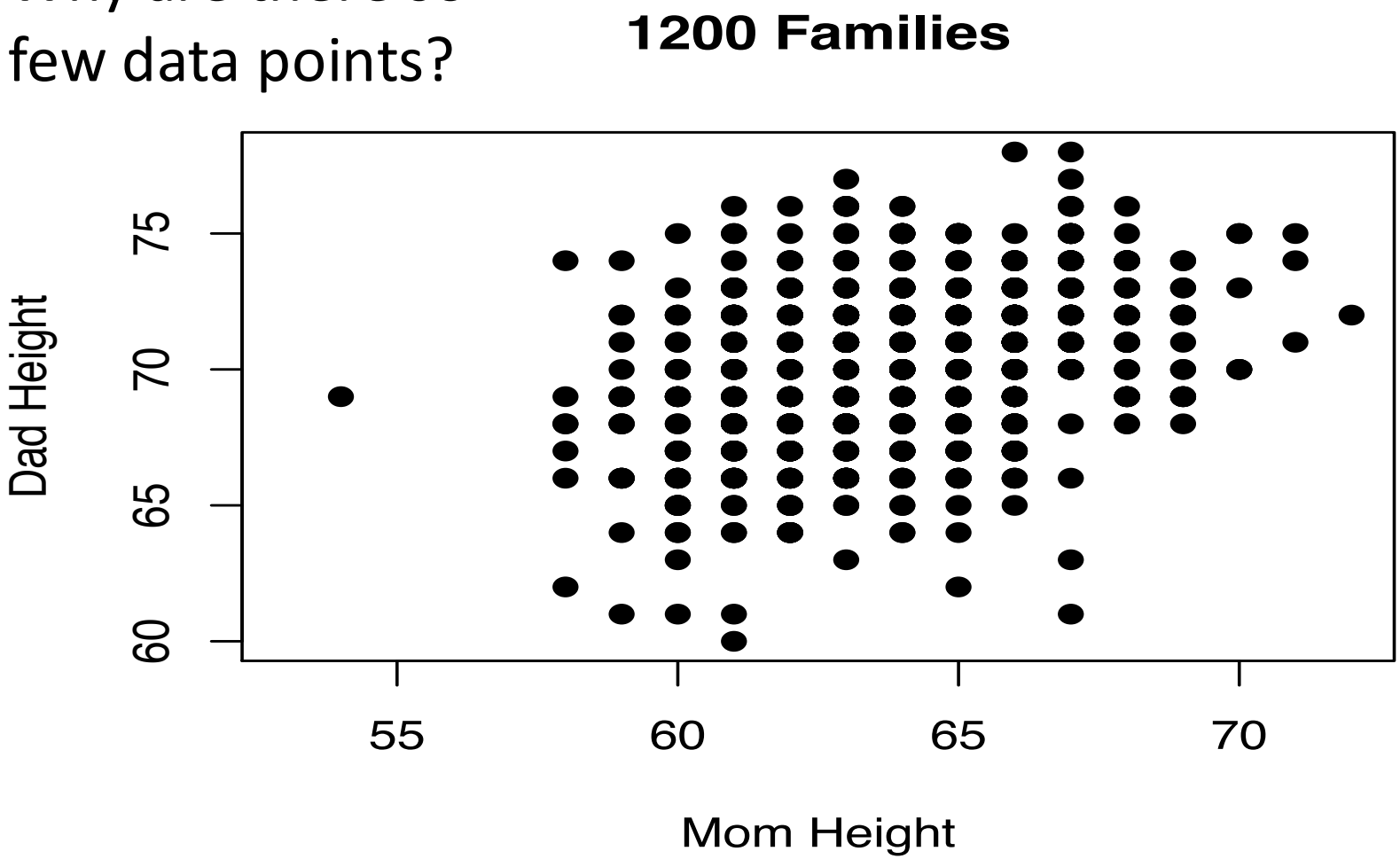


Use visually prominent symbols



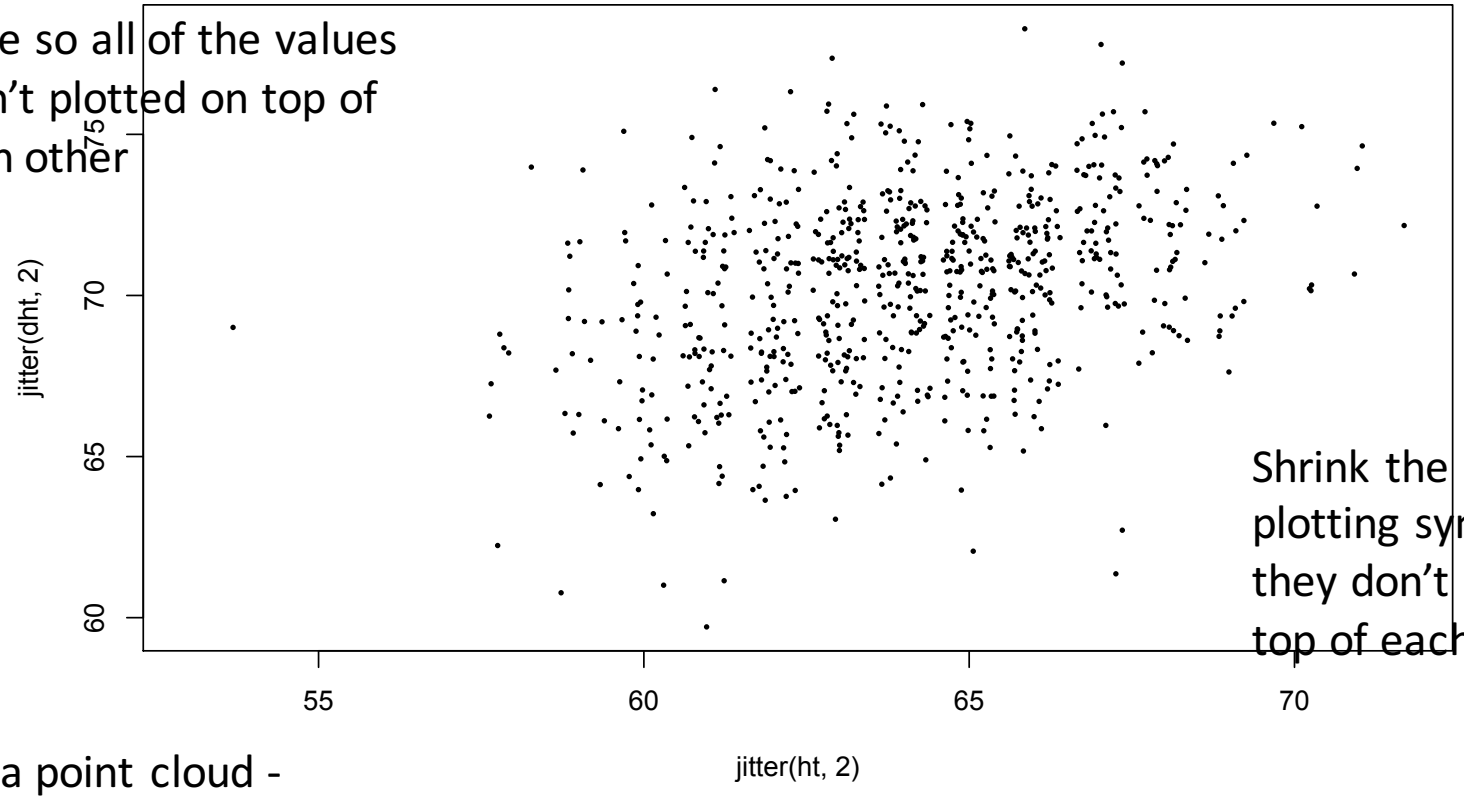
Avoid over-plotting

Why are there so few data points?



One way to avoid over plotting: Jitter the values

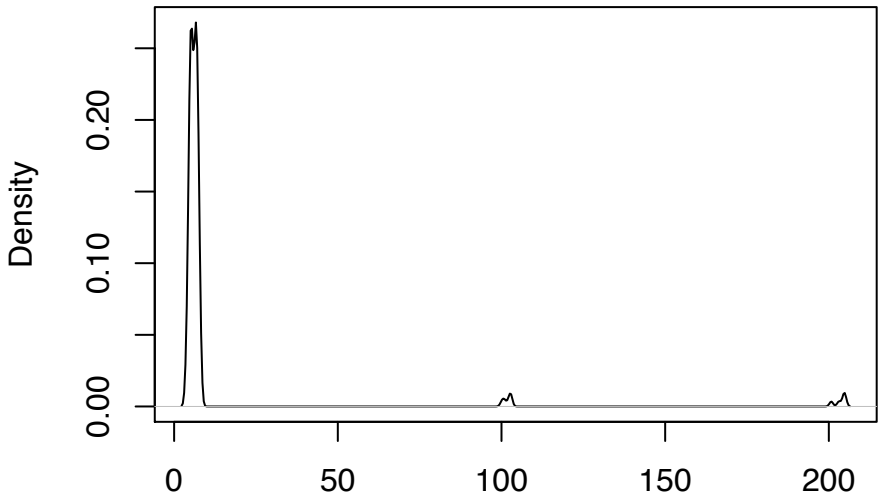
Add a little bit of random noise so all of the values aren't plotted on top of each other

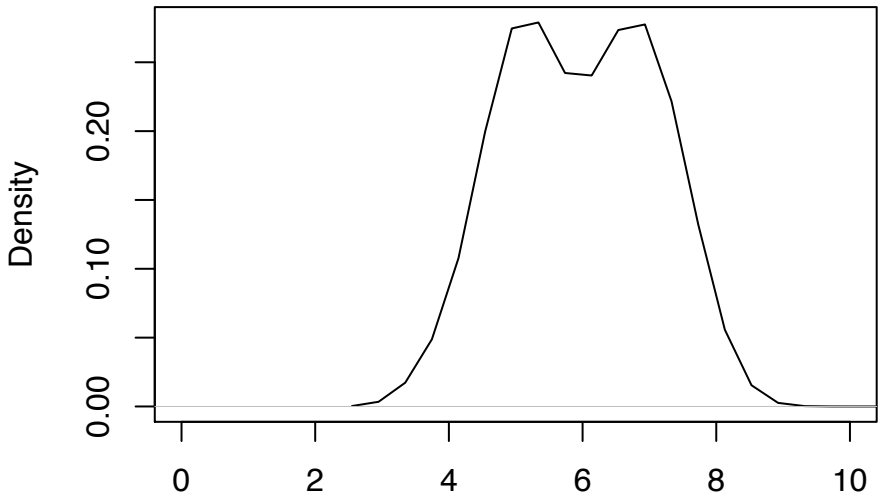


Shrink the plotting symbol so they don't plot on top of each other

See a point cloud -

Different values of data may obscure each other





Most of the data are in the 0 to 10 range.
The few large values obscure the bulk of the data.
Consider mentioning these large values in a caption, instead of showing them in the plot.

Choosing the Scale (limits) of the Axis

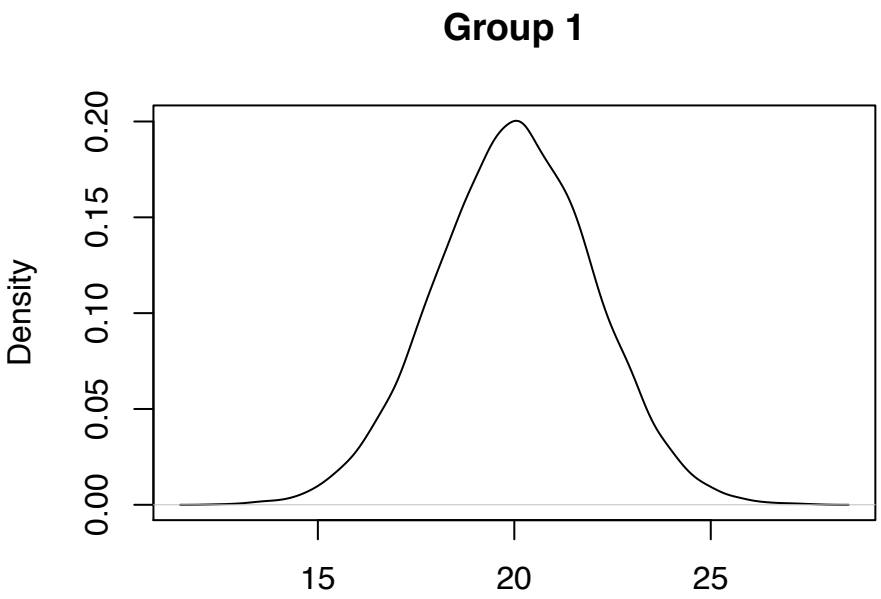
- Include all or nearly all of the data
- Fill data region
- Origin need not be on the axis
- Choose a transformation that improves resolution (to be continued)

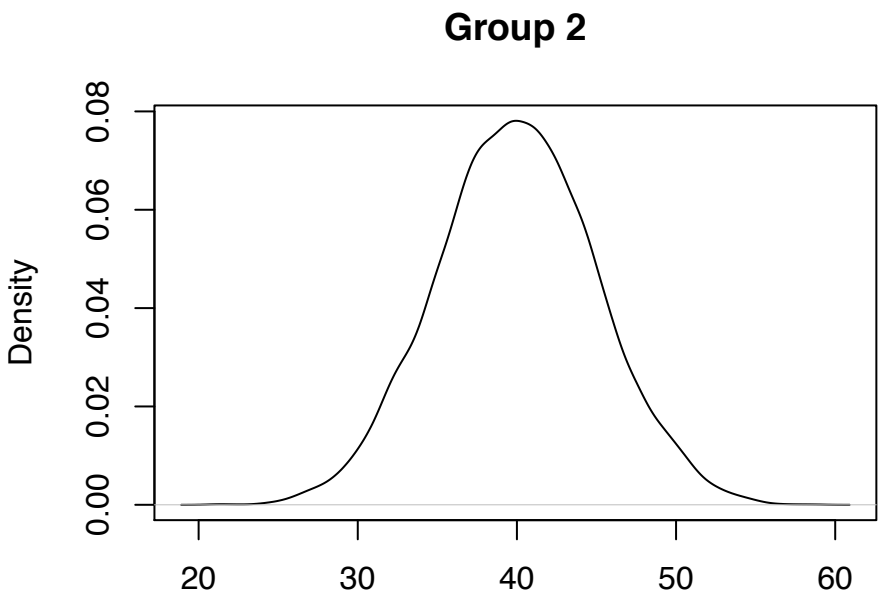
Eliminate superfluous material

- Chart junk – stuff that adds no meaning, e.g. butterflies on top of barplots, background images
- Extra tick marks and grid lines
- Unnecessary text and arrows
- Decimal places beyond the measurement error or the level of difference

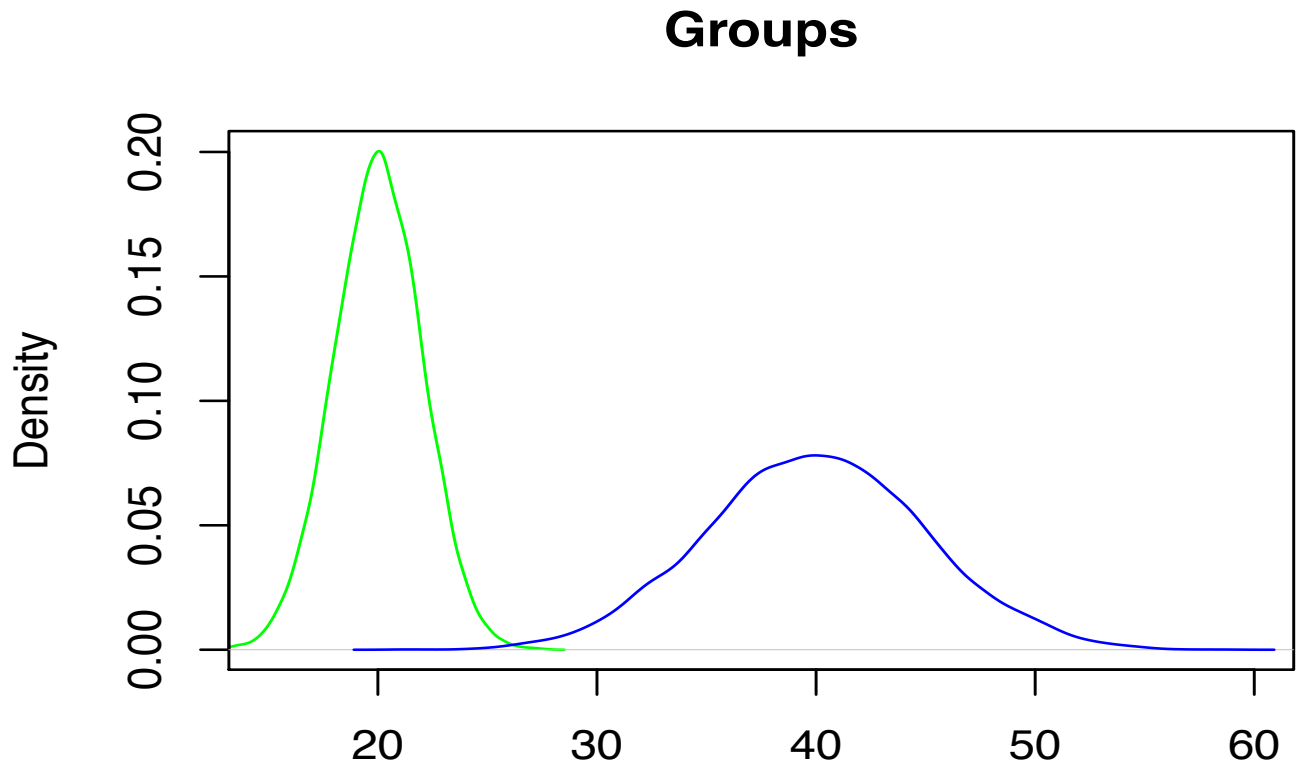
Facilitate Comparisons

Put Juxtaposed plots on same scale





Make it easy to distinguish elements of *superposed* plots (e.g. color)

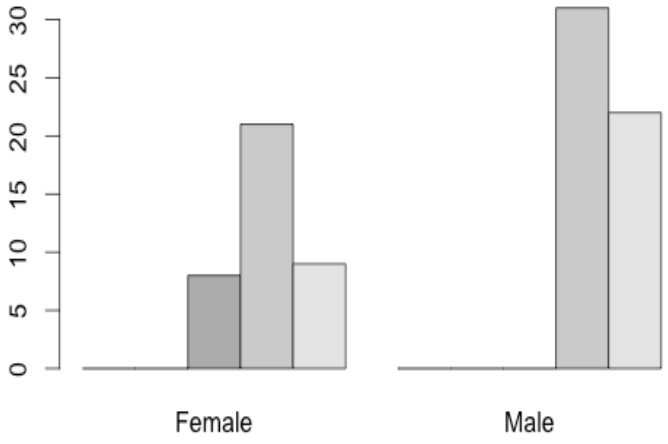


Choosing the Scale

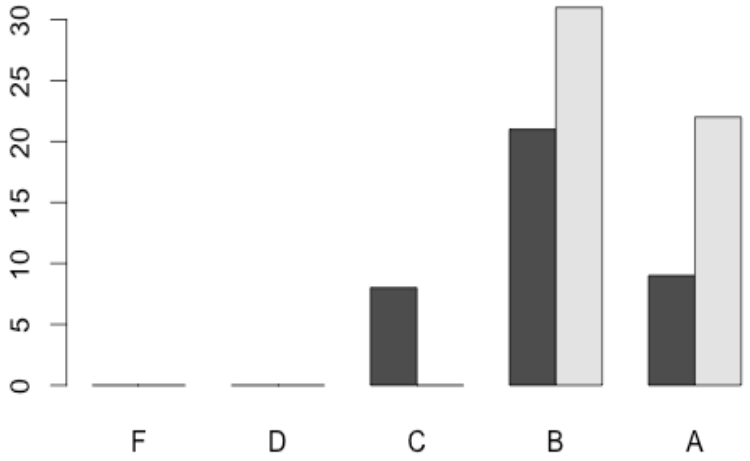
- Keep scales on x and y axes the same for both plots to facilitate the comparison
- Zoom in to focus on the region that contains the bulk of the data
- These two principles may go counter to one another
- Keep the scale the same throughout the plot (i.e., don’t change it mid-axis)

Emphasizes the important difference

A



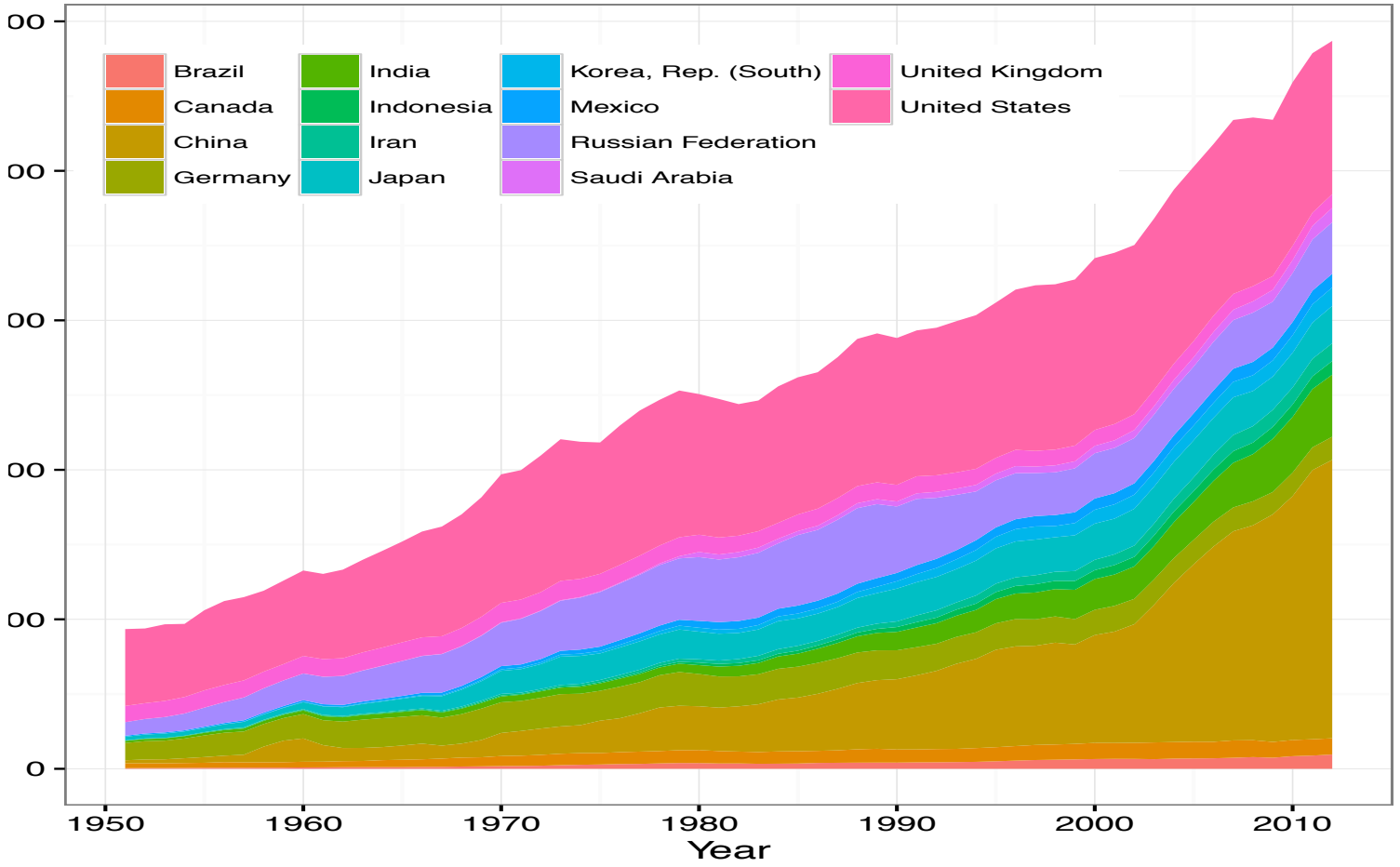
B



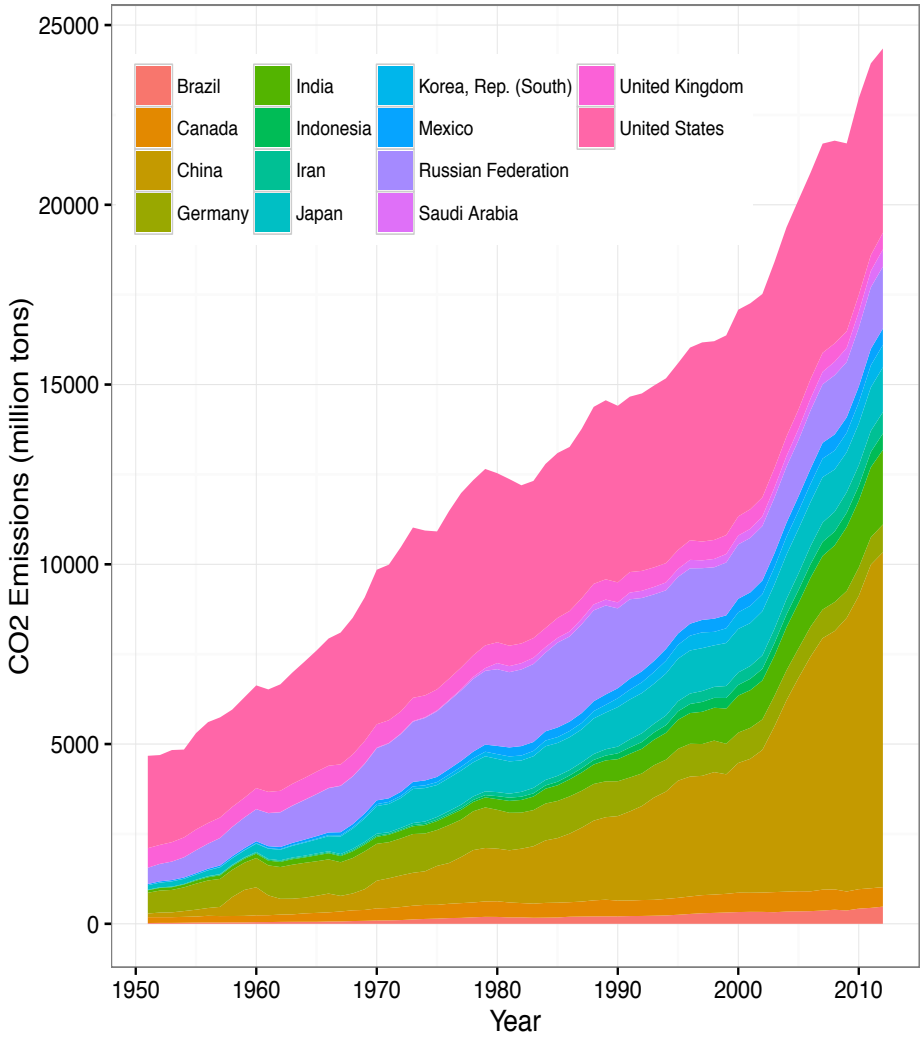
Which of these side-by-side bar plots emphasizes the important?

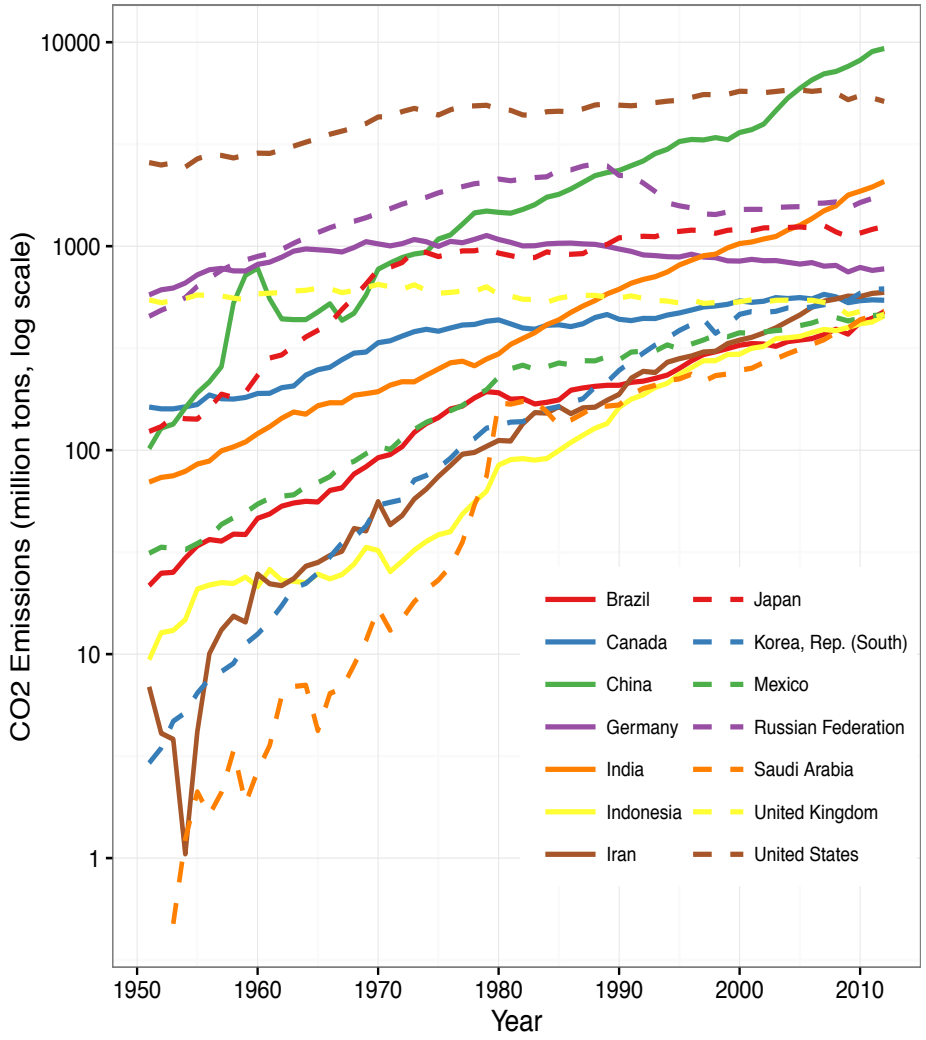
Avoid Jiggling the baseline

It is difficult to see how a country has changed over time because the bottom/base line moves



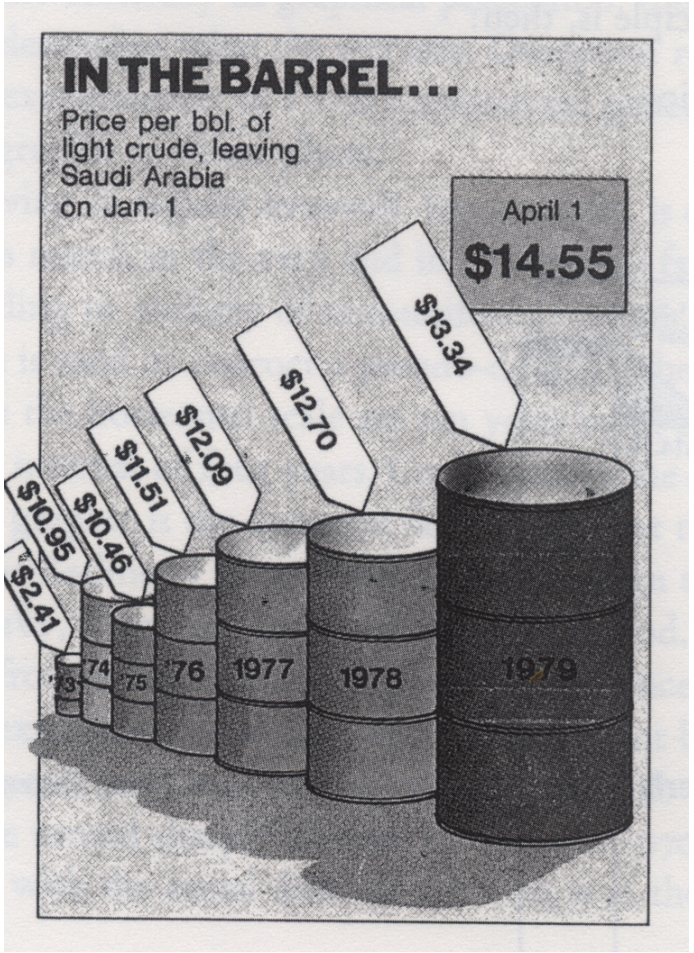
Avoid Jiggling the baseline



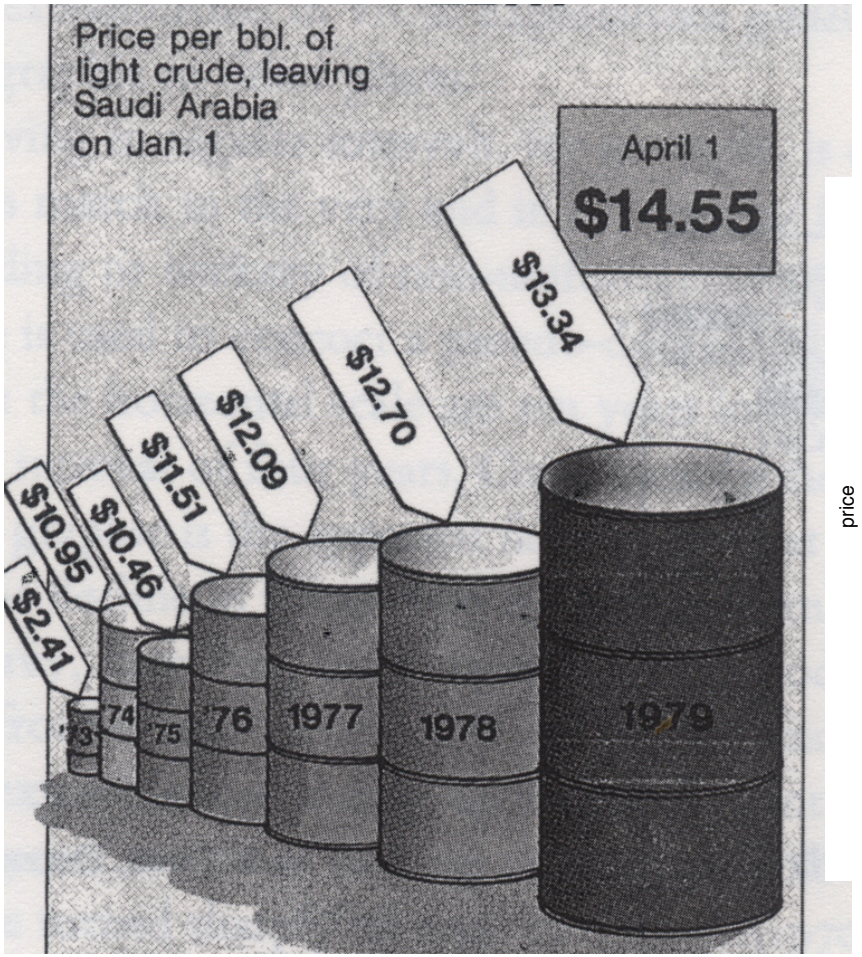


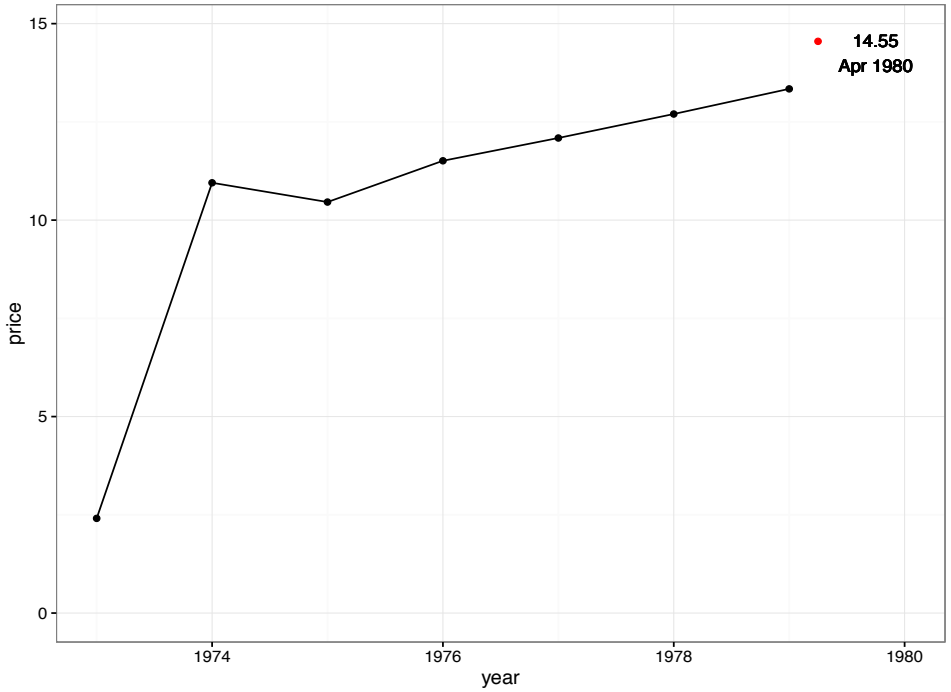
Comparison: volume, area, height

We naturally compare the volume of the barrels, but the change is really the height of the barrels



Comparison: volume, area, height





Information Rich

How to make a plot information rich

- Describe what you see in the **Caption**
- Add context with **Reference Markers** (lines and points) including text
- Add **Legends** and **Labels**
- Use color and plotting symbols to add more information
- Plot the same thing more than once in different ways/scales
- Reduce clutter

Captions

- Captions should be comprehensive
- Self-contained
- Captions should:
  - Describe what has been graphed
  - Draw attention to important features
  - Describe conclusions drawn from graph

Good Plot Making Practice

- Put major conclusions in graphical form
- Provide reference information
- Proof read for clarity and consistency
- Graphing is an iterative process
- Multiplicity is OK, i.e., two plots of the same variable may provide different messages
- Make plots data rich