

Data Visualization with ggplot2

Session 4: Making Good Visualizations

Nathan Barron

✉ nathanbarron@ou.edu

4 Principles of Good Data Visualization

Know your data

- You can't visualize what you don't understand
- A deep knowledge of the data, its structure, and the insights to be drawn are essential to data visualization
- When preparing your visualization, **form follows function**

Data for price of groceries since 1970

year	item	unit	price	type
1970	milk	gallon	0.49	namebrand
1970	eggs	dozen	0.60	namebrand
1970	butter	lb	0.87	namebrand
1970	flour	5lbs	0.59	namebrand
1970	sugar	5lbs	0.67	namebrand
1971	milk	gallon	0.50	namebrand
1971	eggs	dozen	0.53	namebrand
1971	butter	lb	0.88	namebrand
1971	flour	5lbs	0.60	namebrand
1971	sugar	5lbs	0.79	namebrand

Sketch

The **creative** and **> programming** aspects of data visualization require different parts of your brain.

1. Start with *What do I want?*
2. Draw it
3. Follow up with *How do I get that?*
4. Code it!

Visualize the price of groceries since 1970

Simple Visuals

Don't try to include too much or irrelevant information.

Do provide detailed captions that offer additional context and information to the graphic

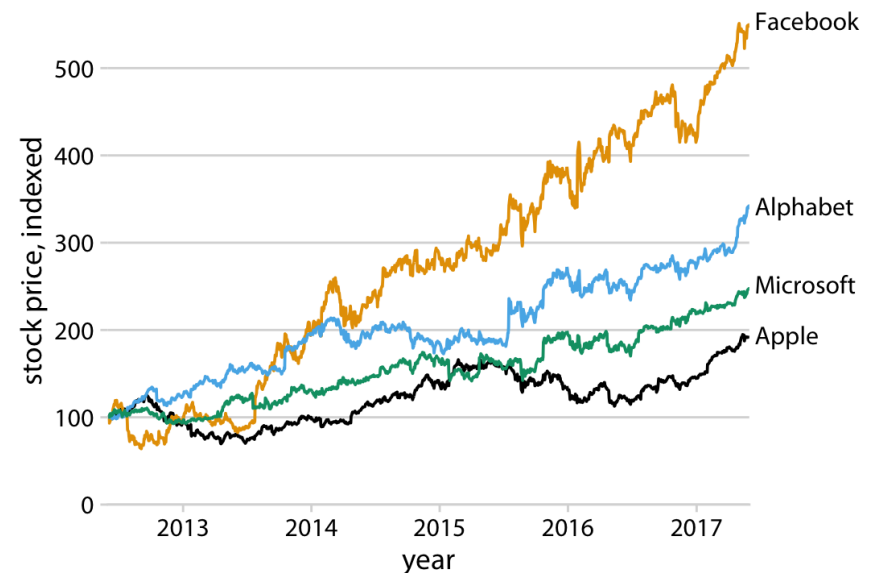
Use Color Effectively

- Only use colors when:
 - There are fewer than 6 categories; or
 - Visualizing continuous data with color gradient
- Use monotonic scales
- Use accessible colors: [Color Blindness Simulator](#)
- Use colors and shading to highlight important information

[Color Palette Generator](#)

Legends

- If there is a clear visual ordering in your data, make sure to match it in the legend.
- Whenever possible, design your figures so they don't need a legend.



Tell a story well

Don't try to tell too much in a single visualization

Do know *what* you want to convey

Do design for the 'aha' moment