Data Visualization

A (Very) Brief Overview

What's your graphic IQ?

- http://www.perceptualedge.com/files/GraphDesignIQ.html

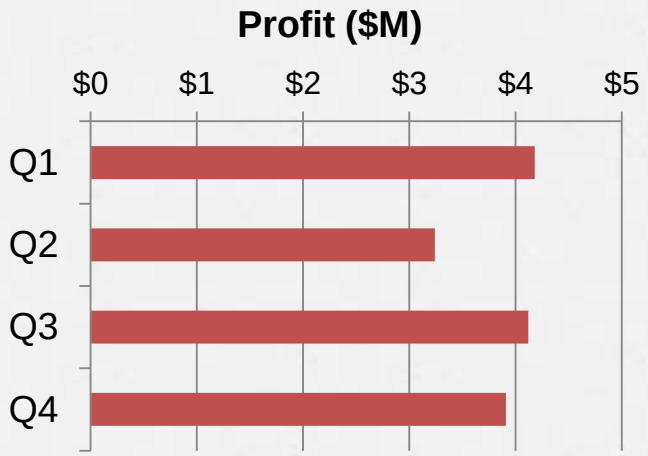
What is data viz?

- Data visualization helps us translate numbers into a picture that we can interpret more easily.
- Two of the biggest reasons to use data visualization are to…
  - Make sense of data (i.e., help *you* see patterns)
  - Communicate data to others (i.e., executive presentations)
- Good visualizations allow our eyes to do some of the heavy lifting in data processing

What is data viz?

- The low profit in Q2 is more apparent in the graph than in the table.

	Profit ($M)
Q 1	$4.18
Q 2	$3.24
Q 3	$4.12
Q 4	$3.91



Audience matters!

- Visualizations can be useful, even if you're the only one to see them.
- For presentations, consider if everyone will have a printed copy, or if they need to be able to see fine details on a screen.

	Q1	Q2	Q3
Group 1	4.19	4.25	4.50
Group 2	4.00	4.32	4.31
Group 3	4.22	4.18	4.54
Group 4	4.35	4.41	4.62

	Q1	Q2	Q3
Group 1	4.19	4.25	4.50
Group 2	4.00	4.32	4.31
Group 3	4.22	4.18	4.54
Group 4	4.35	4.41	4.62

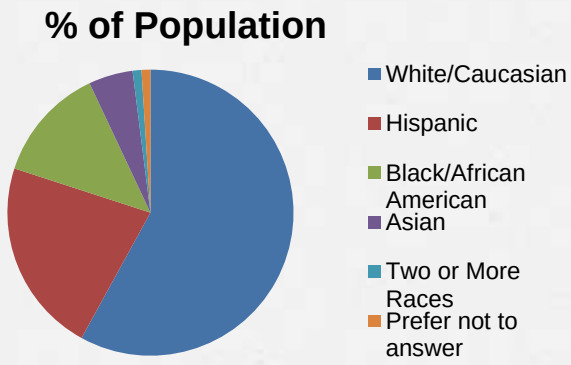
Purpose matters!

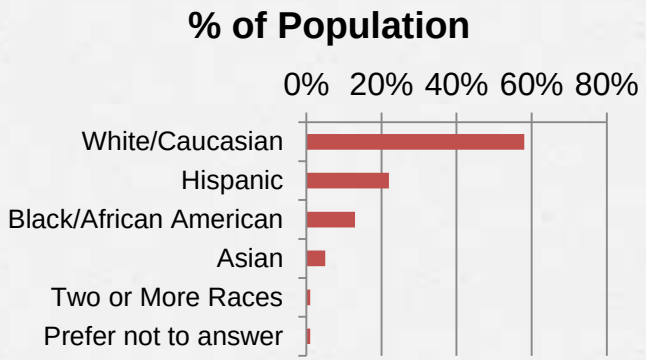
- Presentations are often expected to function as both the visual aid during the meeting, and a reference for those unable to attend.

- Alternatives:
  - 2 documents with difference purposes
  - Use “speaker notes”

Content matters!

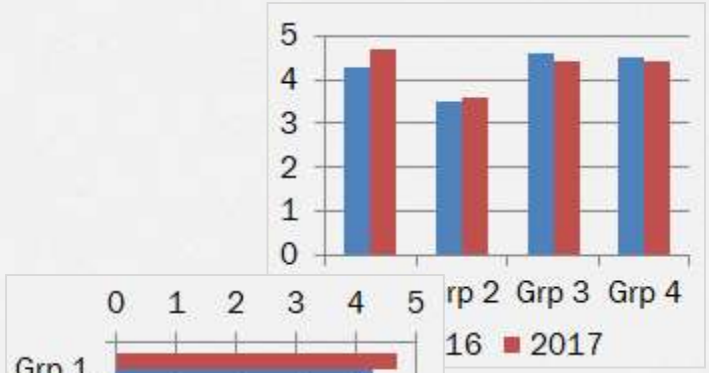
- Different chart types have different strengths and weaknesses.
- Of the two options below, which do you think is easier to interpret? Why?

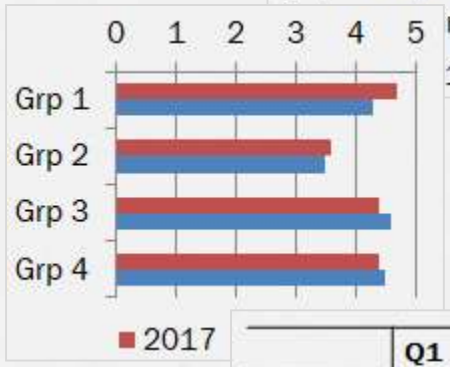




Types of Data: Comparisons

- **Column charts** are likely the best option if you have few groups (< 8), but title length can be a problem
- **Bar charts** work better if you have more groups or longer titles
- **Heat maps** are good if you're comparing more than one group on more than one metric



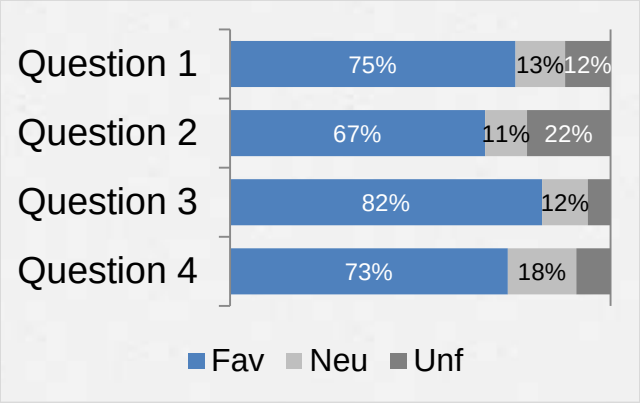


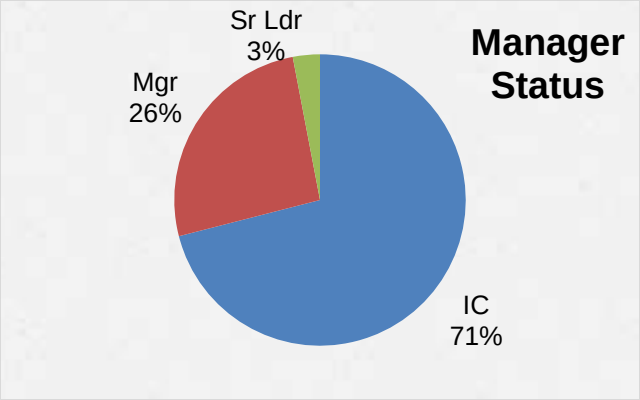
	Q1	Q2	Q3
Group 1	4.19	4.25	4.50
Group 2	4.00	4.32	4.31
Group 3	4.22	4.18	4.54
Group 4	4.35	4.41	4.62

Types of Data: Comparisons, continued

- **Bullet graphs** have become popular in dashboards, since they're a compact option to show a value against a specific metric as well as a showing what's considered good, fair, and bad.

Types of Data: Composition

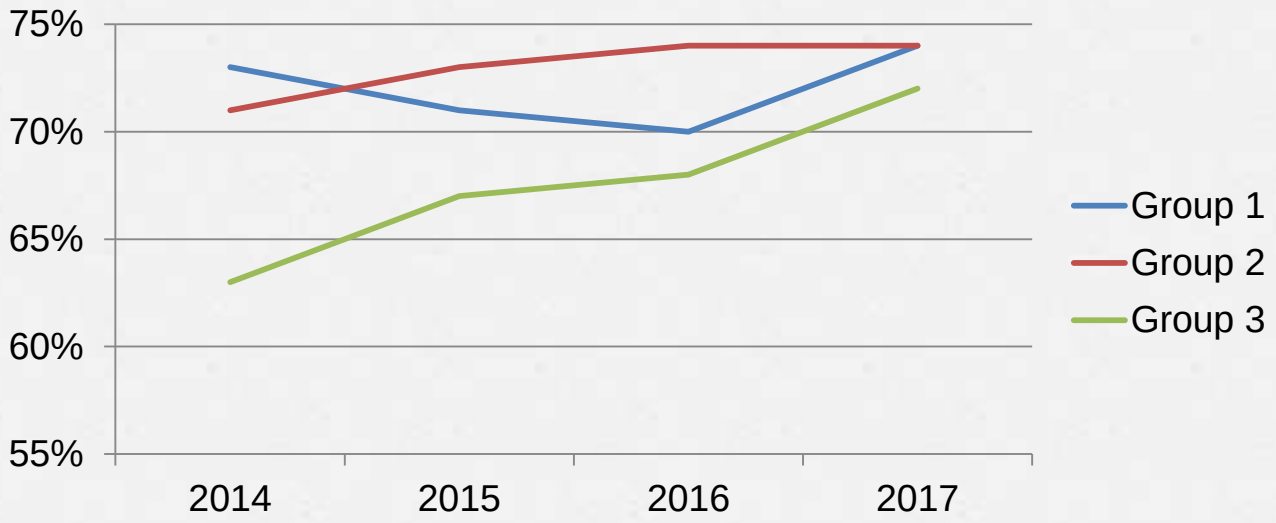




- **Stacked bar** or **column charts** are frequently used for survey results
- **Pie charts** can be used, but be careful to avoid common issues with pies

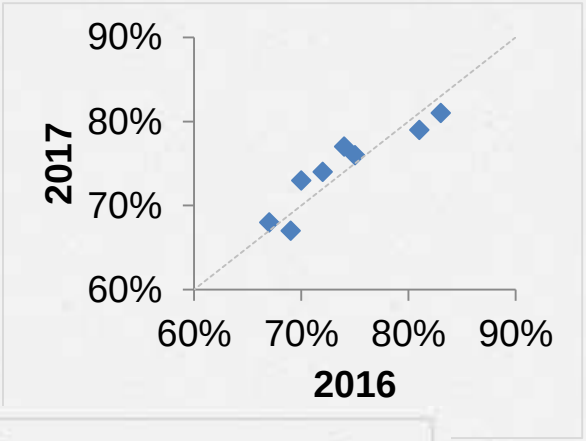
Types of Data: Trend

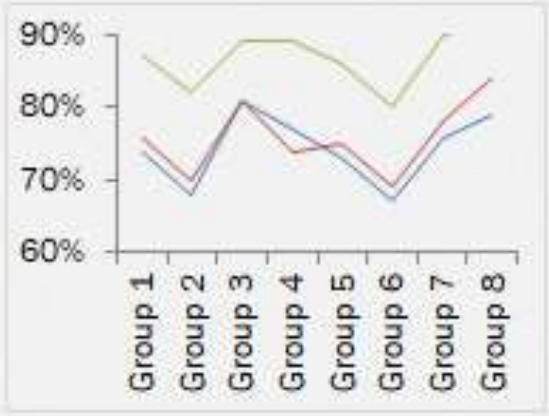
- **Line charts** are commonly used to display long-term trend data.



Types of Data: Relationship

- **Scatterplots** are the most common graphic to show relationship data, but bubble charts can be used if you need to look at a third variable
- You can also use **line charts** to see relationships by comparing the shape of the graphs
- **Network graphs** can be used to show relationships between individuals





Tips for Useful Charts & Graphs

- Use the full axis, particularly on column and bar charts
- Highlight the most important information
- Consider if legends and labels detract from your point
  - *This isn't to say you shouldn't label your data, but sometimes it's redundant or could be placed in a better way*
- Consider the "data-to-ink" ratio
- Pass the "squint" test
- Think about sort order
- Ask for a second opinion

Common Problems in Presentations

- Too much information
- Poor color choices
- Variety for the sake of variety
- Inconsistent axes from one slide to the next
- Neglecting to label charts
- Not following common conventions

Making Bad Charts Better

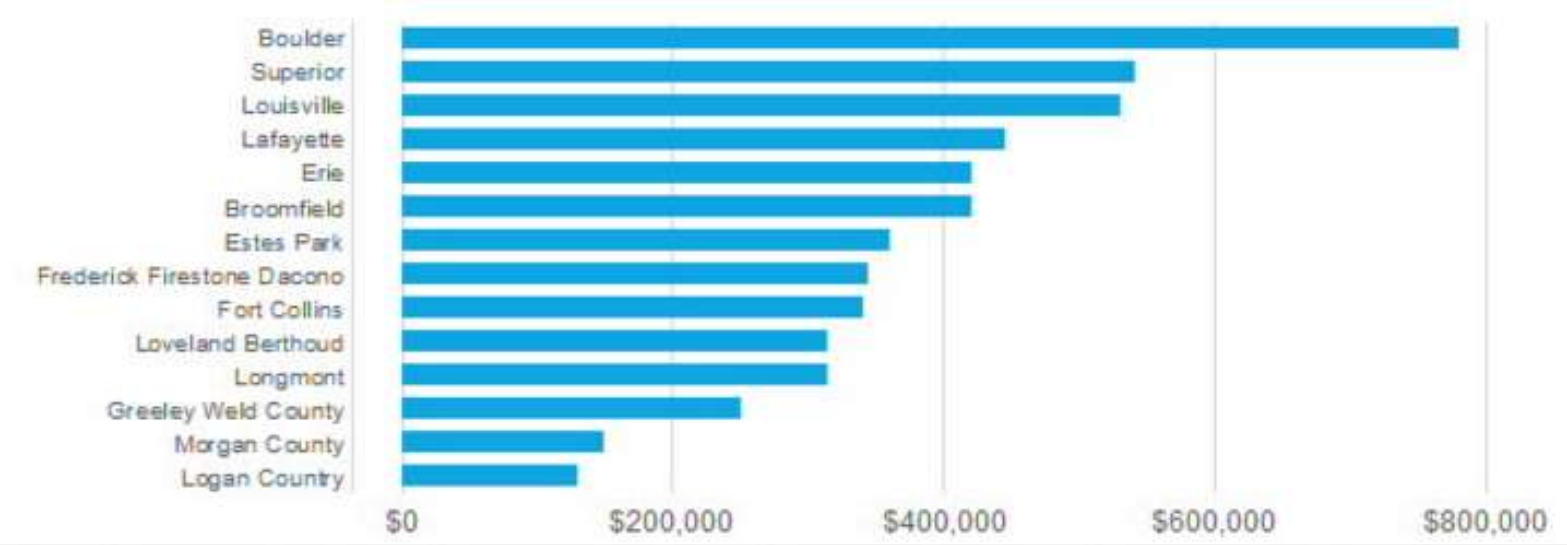
Pretty vs. Useful

- While this chart is pretty and at least labeled, it's hard to read quickly. Using an alphabetic scale on the x-axis doesn't do anything to enhance interpretation.

Example from: http://viz.wtf/post/147196565281/alphabet-on-the-x-axis-for-reals#notes

Pretty vs. Useful, continued

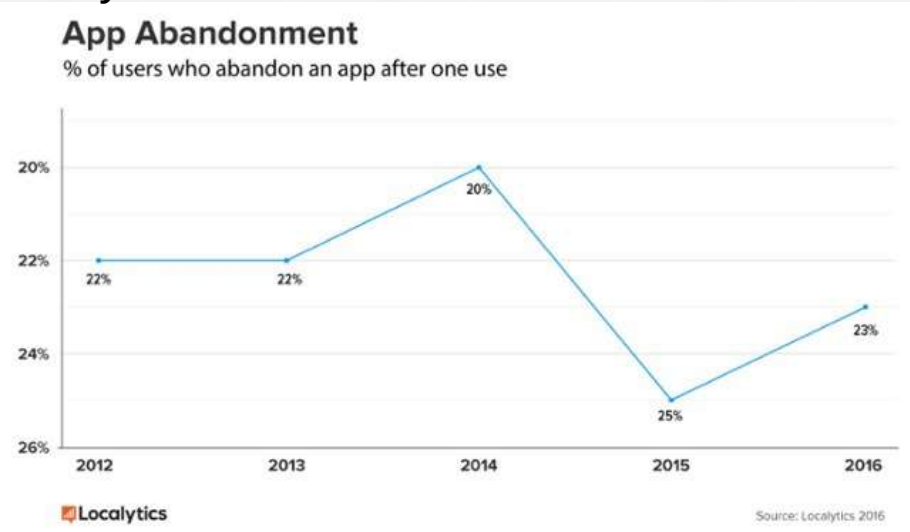
- This bar chart is a better option for conveying the same data.

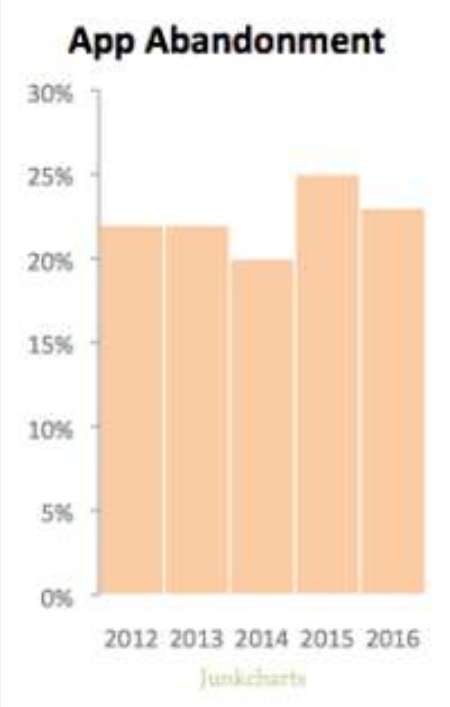


Example from: http://viz.wtf/post/147196565281/alphabet-on-the-x-axis-for-reals#notes

Poor Axis Choice

The graph below uses a line to show trend (a common convention), but the upside-down, truncated y-axis makes it hard to read. A bar chart version would be easier to follow, but you could just fix the y-axis.

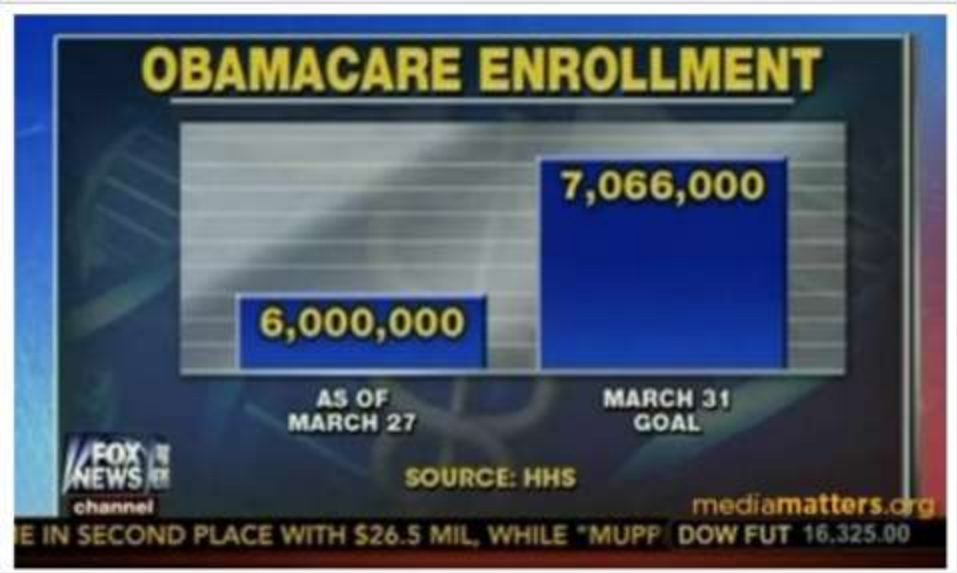


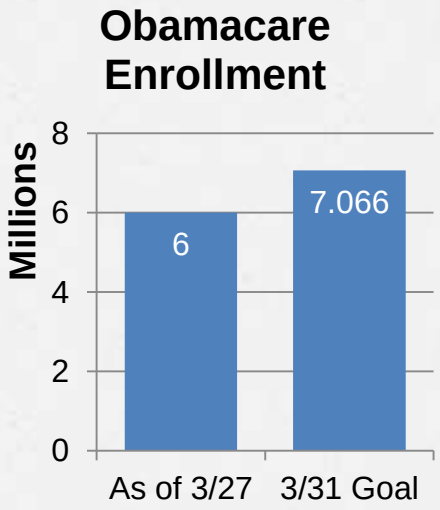


Example from: http://junkcharts.typepad.com/junk_charts/2016/06/what-doesnt-help-readers-on-the-chart-and-what-does-help-off-the-chart.html

Republicans are bad at graphs…

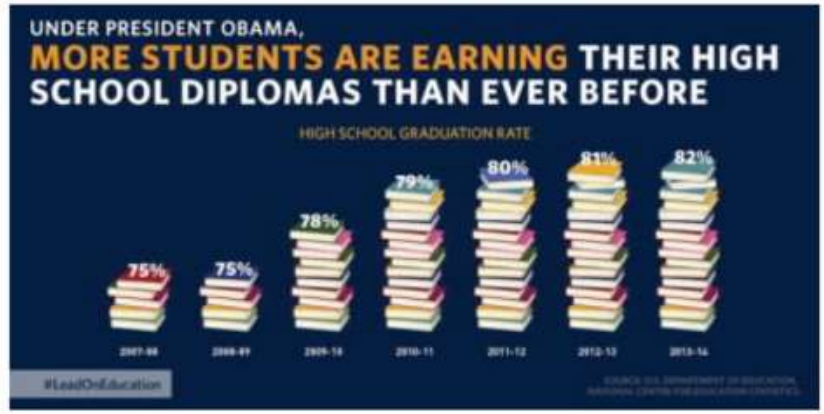
- Media sources with a clear bias are notorious for using a truncated axis to their advantage.

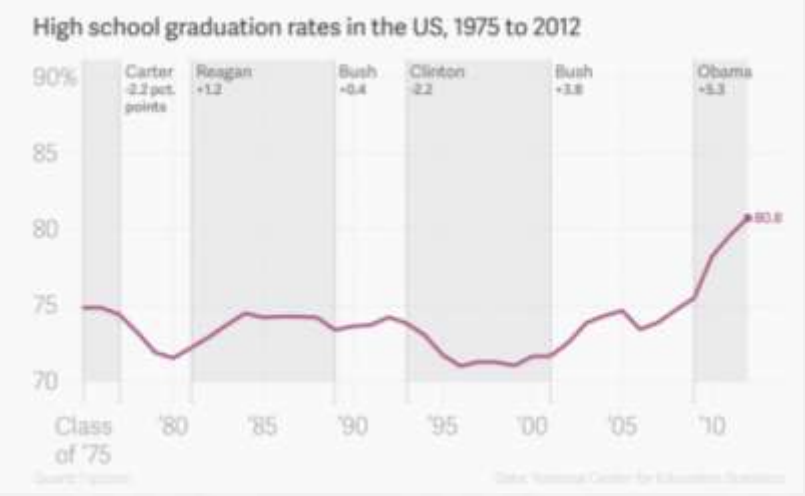




… but so are Democrats

- In addition to truncating the axis, this one also fails to show that the upward trend started before Obama took office.

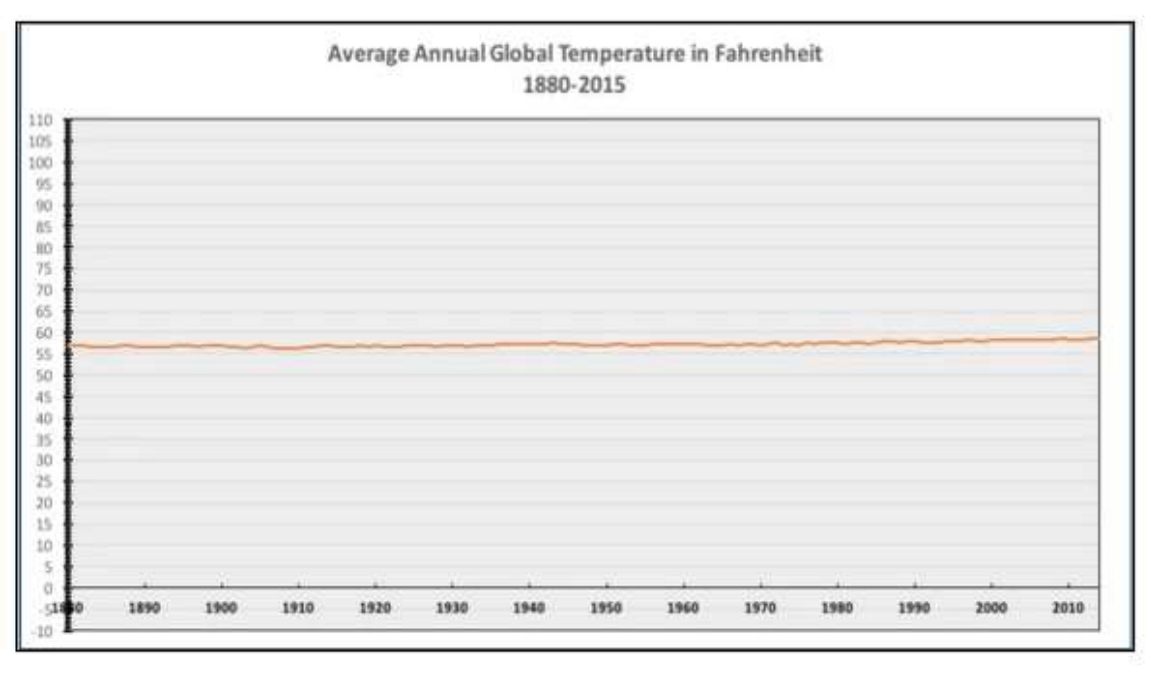




Example from: https://qz.com/580859/the-most-misleading-charts-of-2015-fixed/

Yet Another Poor Axis Choice

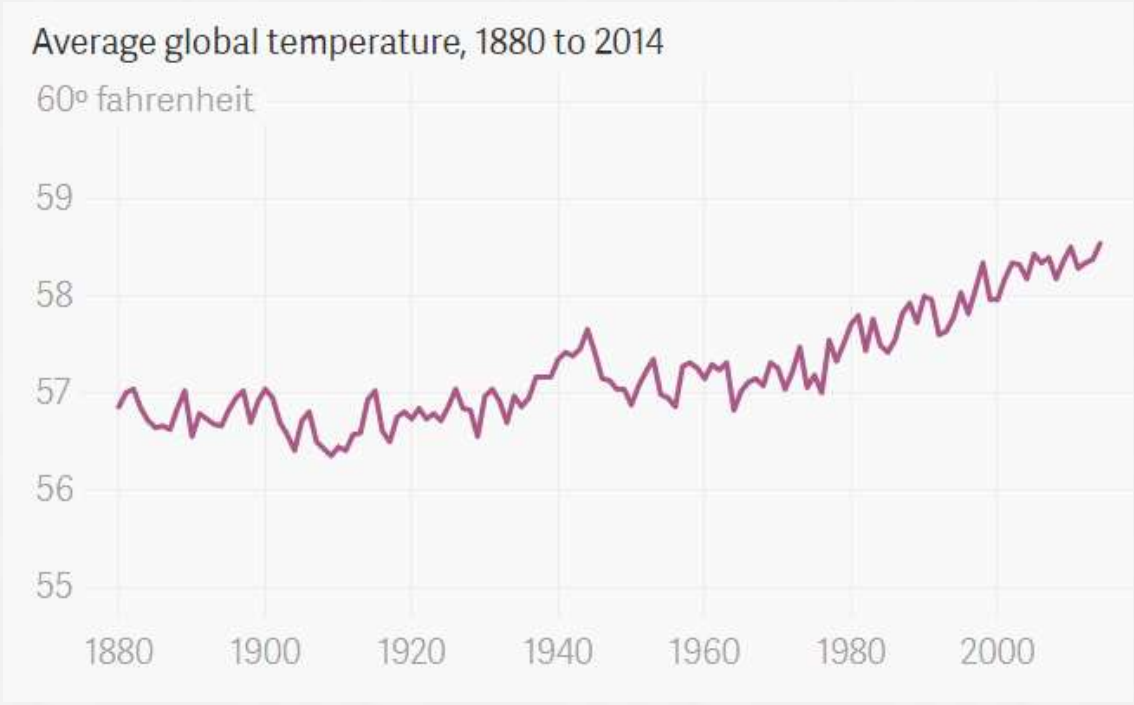
- This is one of those times that it ***doesn't*** make sense to start the y-axis at 0.



Example from: https://qz.com/580859/the-most-misleading-charts-of-2015-fixed/

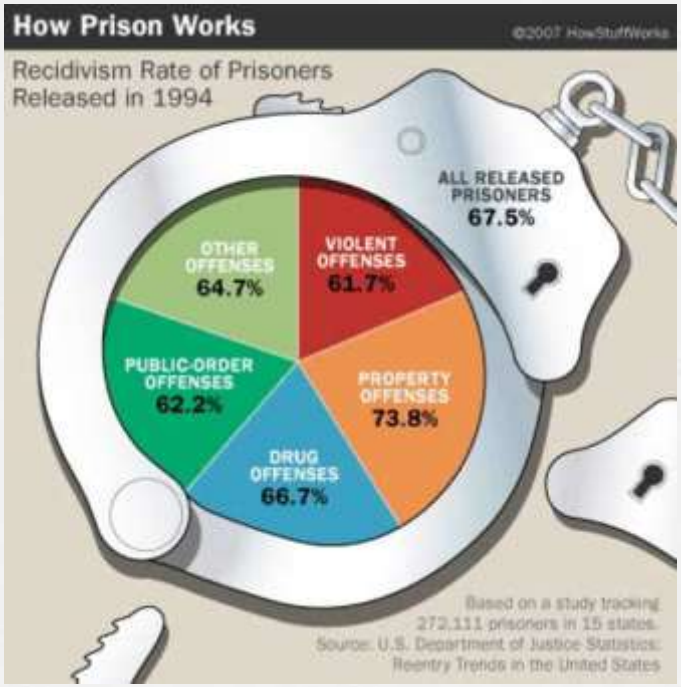
Yet Another Poor Axis Choice, continued

- The trend is more apparent when we use a more realistic axis.

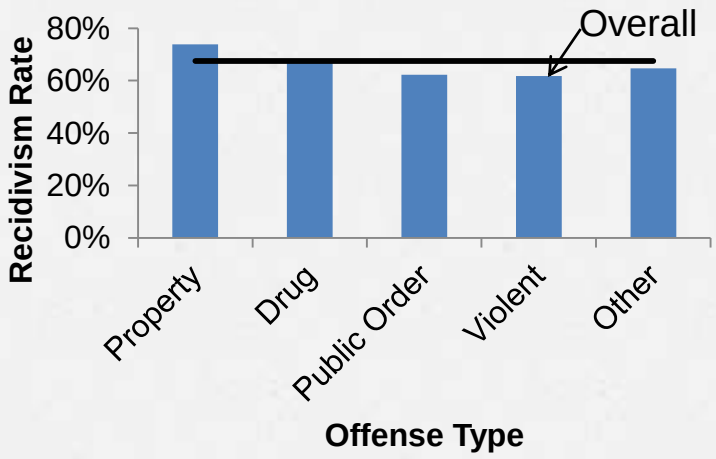


Example from: https://qz.com/580859/the-most-misleading-charts-of-2015-fixed/

How Not to Make a Pie Chart

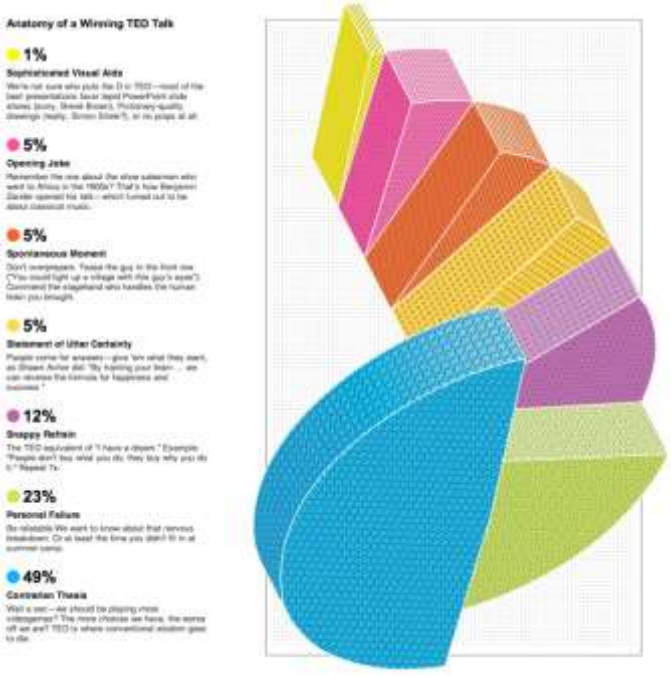


- Pie charts should ***never*** be used to show values on a multiple-select item. Use a bar or column chart instead.

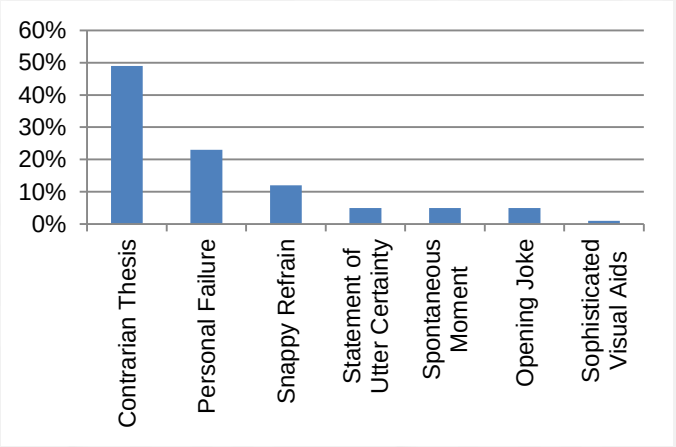


Example from: http://viz.wtf/post/162169270900/what-is-the-recidivism-rate-for-pie-chart

How Not to Make a Pie Chart, continued



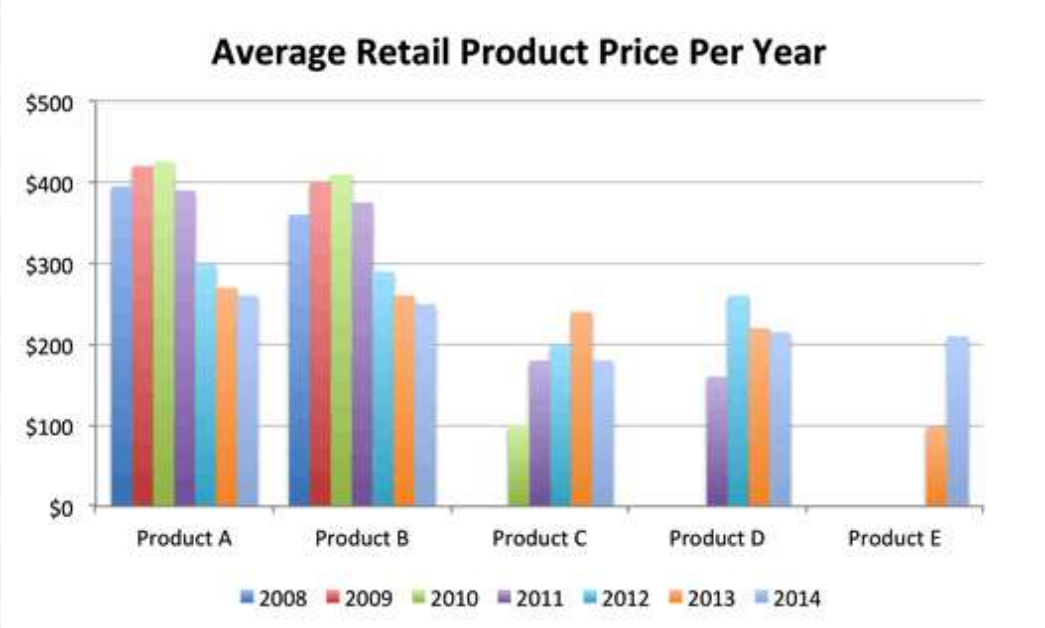
- While this pie does add up to 100%, the design of the graphic makes it nearly impossible to read.



Example from: http://viz.wtf/post/60203066686/the-spiral-staircase-courtesy-of-janwillemtulp

Telling Your Story

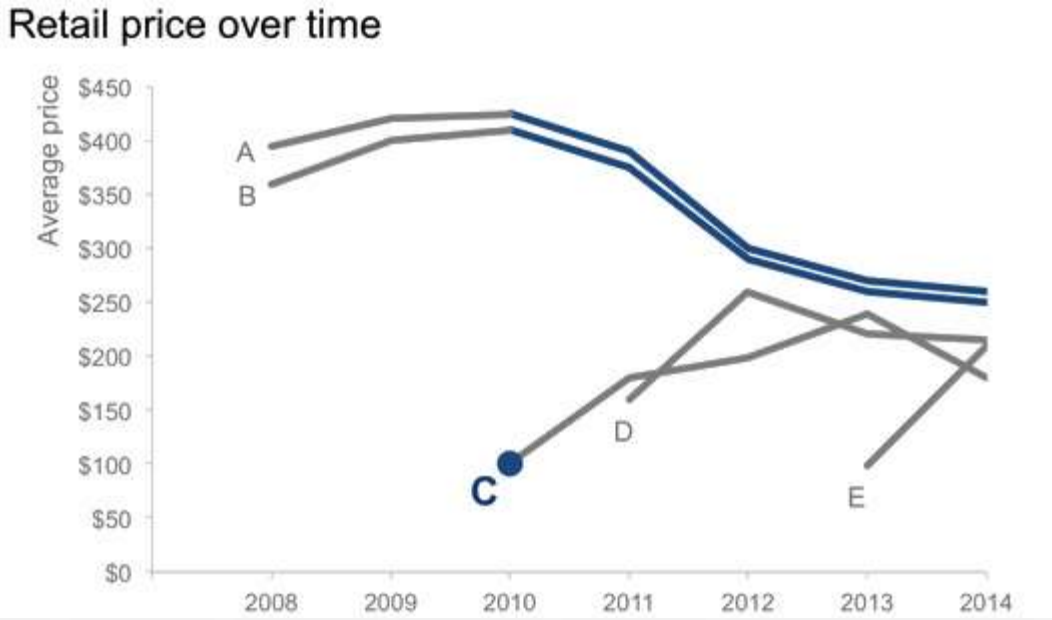
- The headline to go with this chart was "*Price has declined for all products on the market since the launch of Product C in 2010.*"



Example from: http://www.storytellingwithdata.com/blog/2014/05/the-story-you-want-to-telland-one-your

Telling Your Story, continued

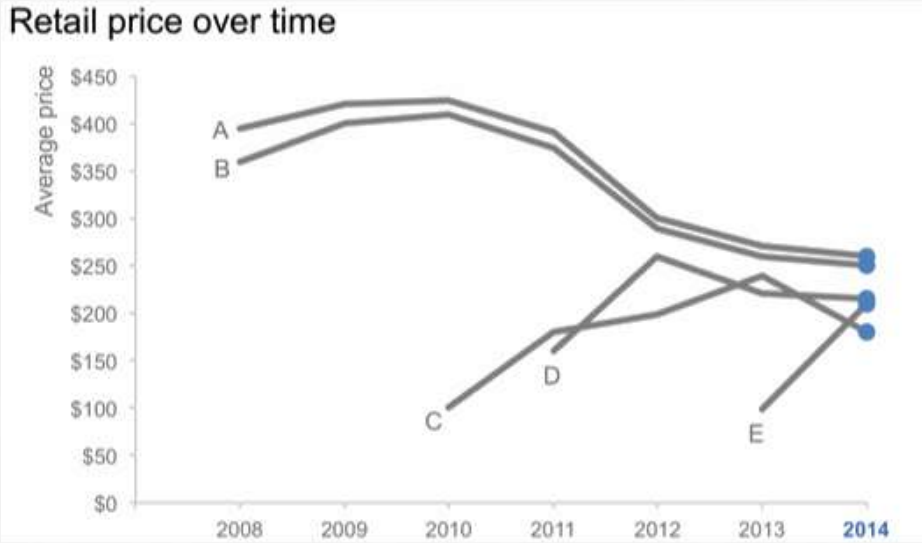
- Since we're looking at trend data, a line chart would make it easier to see where the points are for each year/product.



Example from: http://www.storytellingwithdata.com/blog/2014/05/the-story-you-want-to-telland-one-your

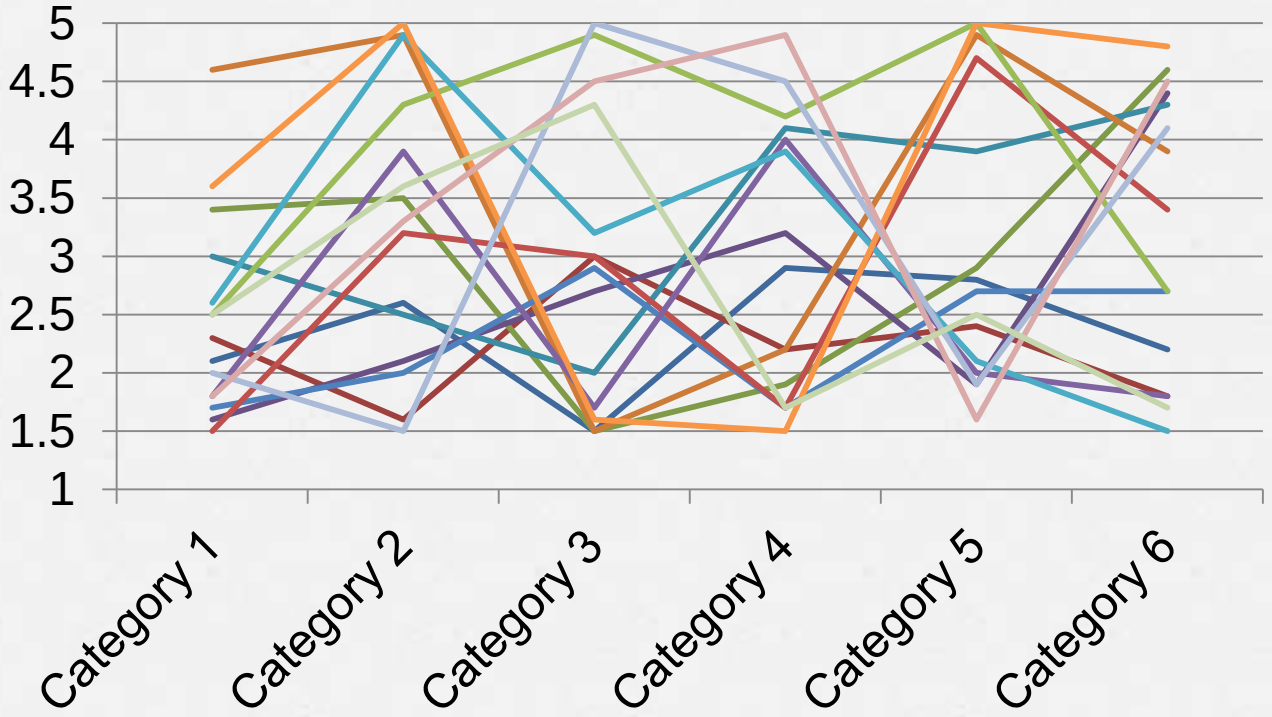
Telling Your Story, continued

- Alternate headline: *"As of 2014, retail prices have converged across products, with an average retail price of $223, ranging from a low of $180 (Product C) to a high of $260 (Product A).*

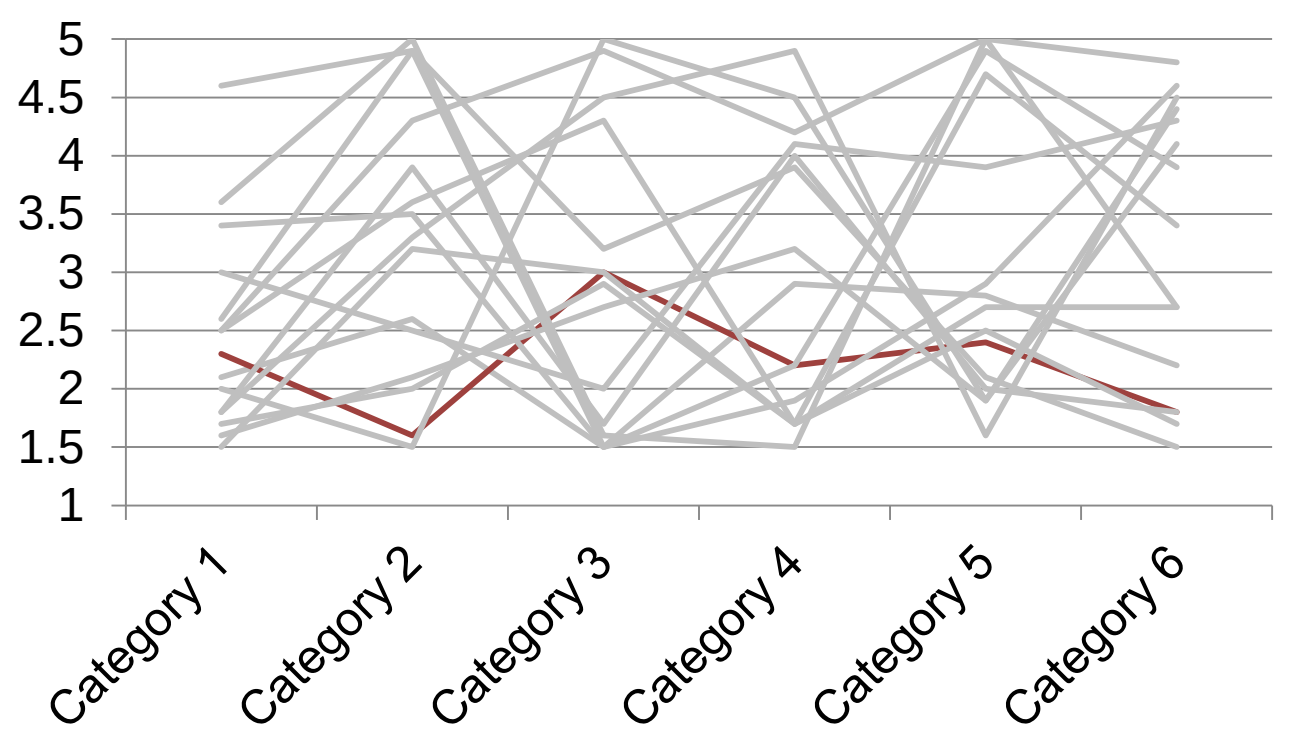


Example from: http://www.storytellingwithdata.com/blog/2014/05/the-story-you-want-to-telland-one-your

Use of Color & Number of Groups



Use of Color & Number of Groups, continued



Appendix

Chart Chooser

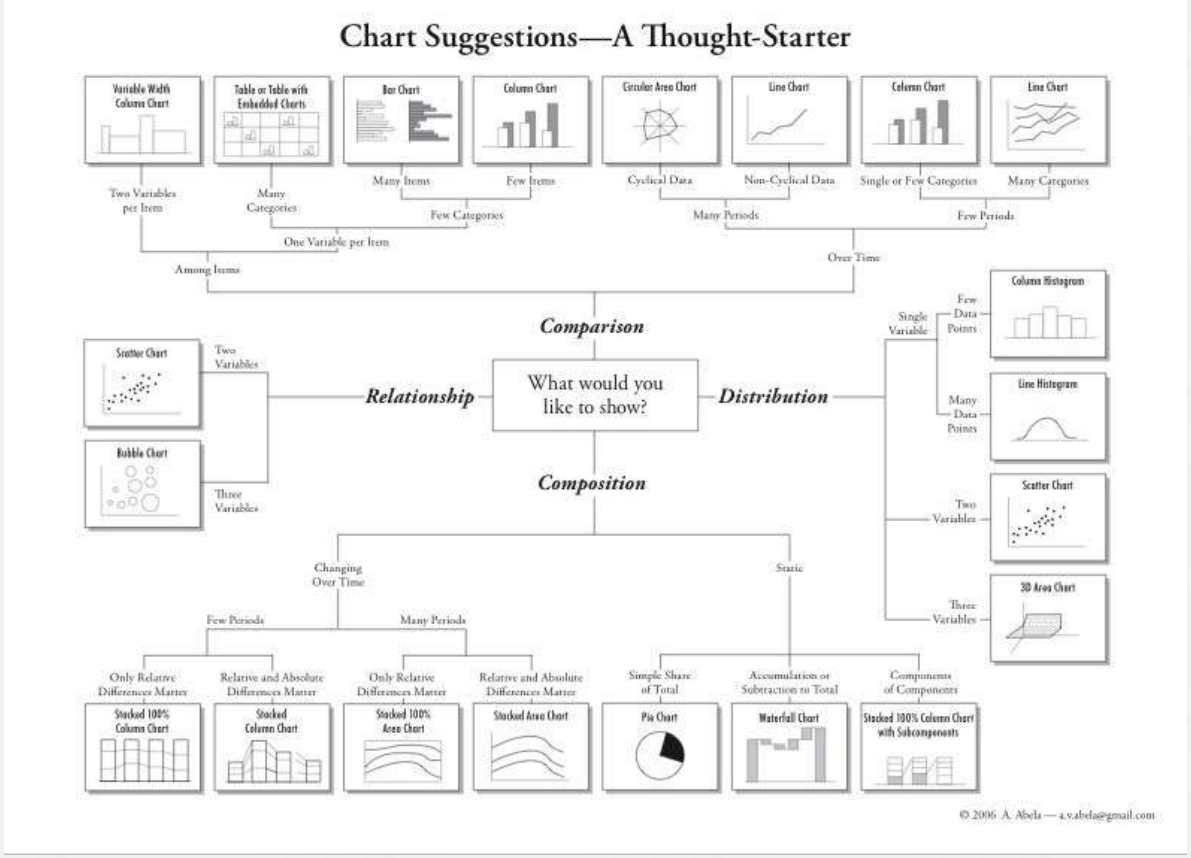
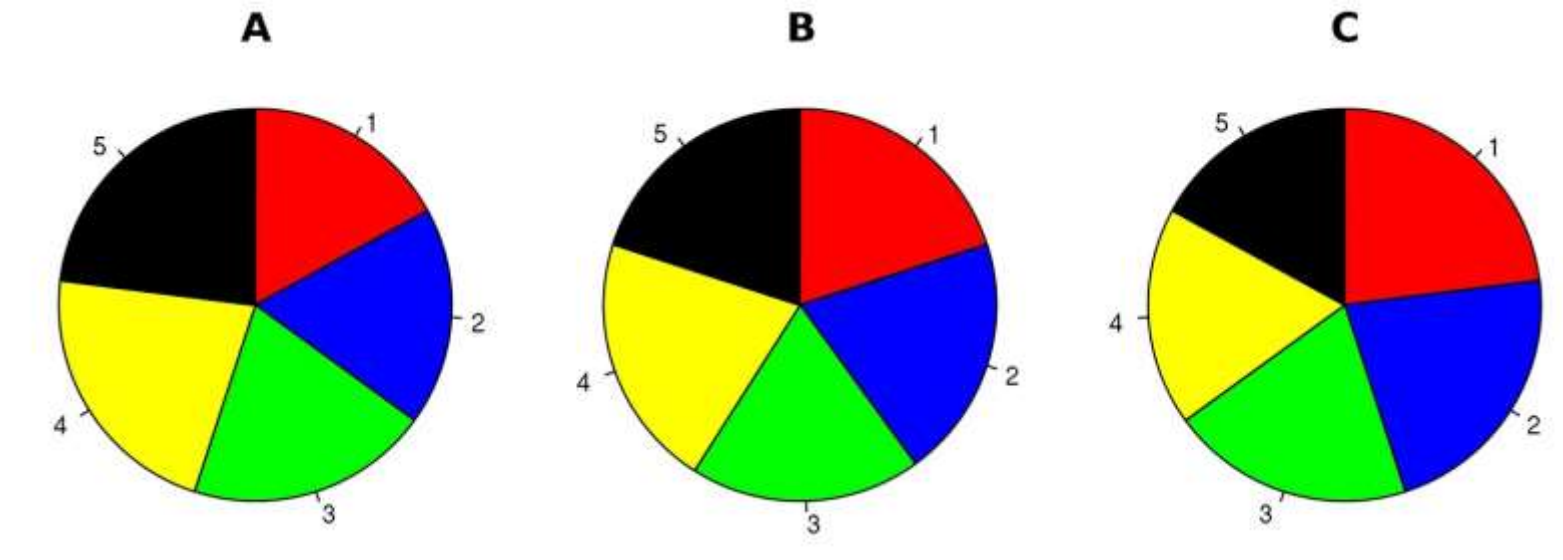


Chart Chooser from: http://extremepresentation.typepad.com/blog/2006/09/choosing_a_good.html

Why Are Pie Charts Disliked?

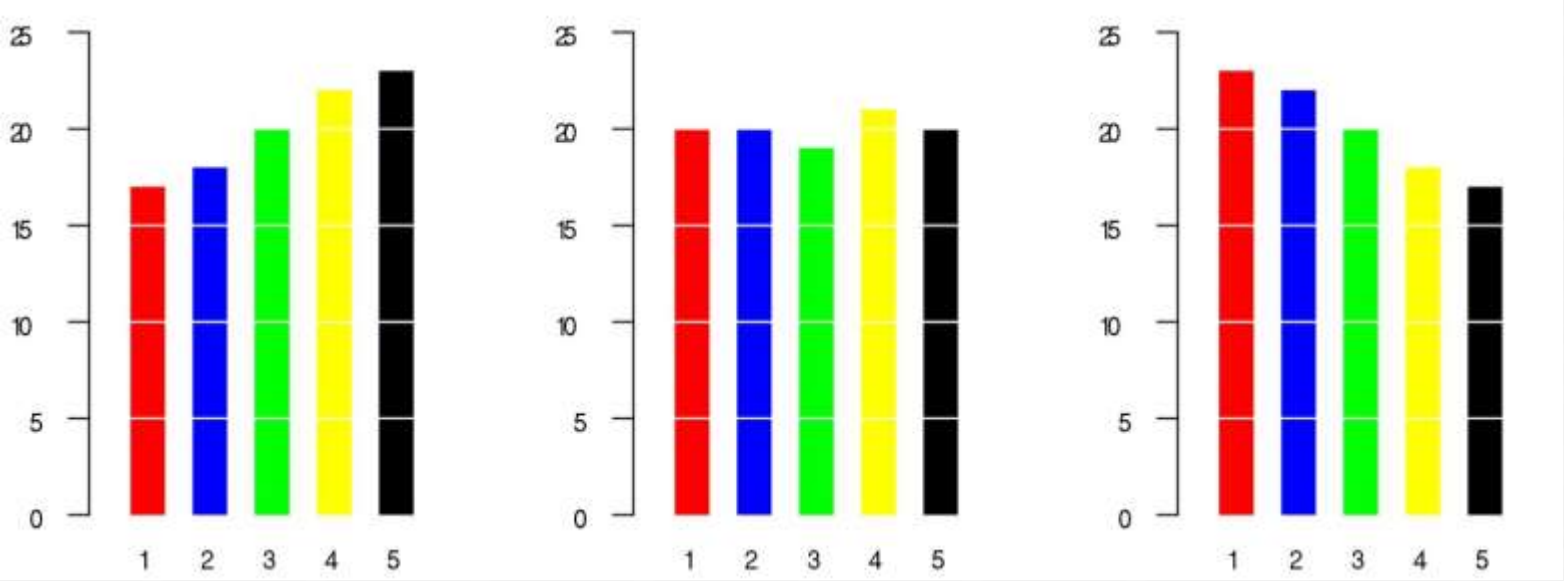
- In each of these charts, can you identify the largest slice? How does it compare to the second largest?



Example from: http://prsync.com/oracle/pie-charts-just-dont-work-when-comparing-data---number--of-top--reasons-to-never-ever-use-a-pie-chart--23294/

Why Are Pie Charts Disliked?

- Here's the same data as column charts- it's much easier to see differences between groups.



Example from: http://prsync.com/oracle/pie-charts-just-dont-work-when-comparing-data---number--of-top--reasons-to-never-ever-use-a-pie-chart--23294/

If You *Really* Need to Use a Pie Chart…

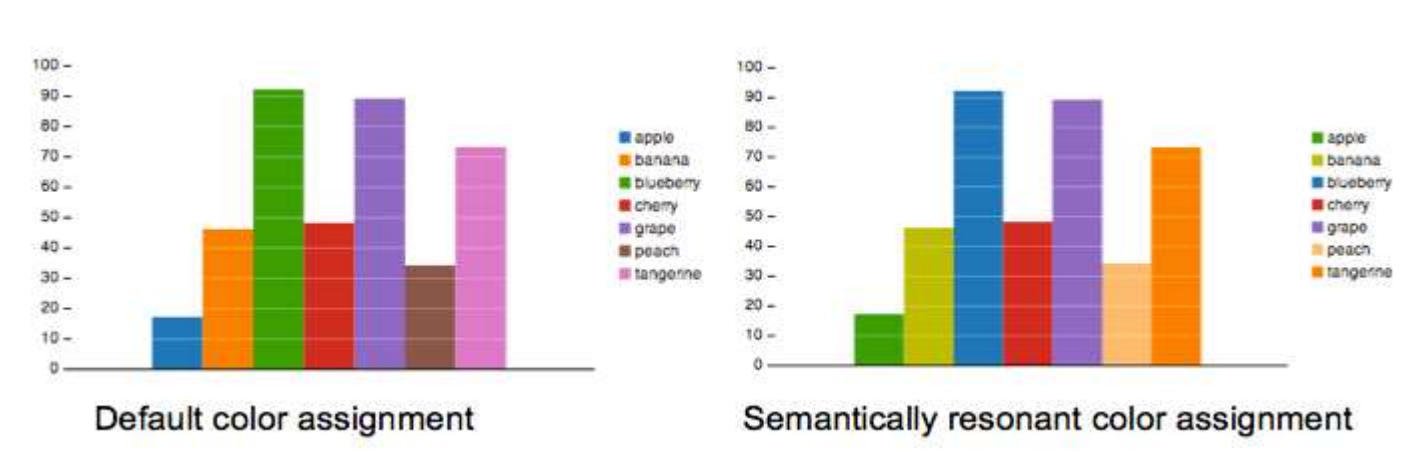
- There are some do's and don't's that are specific to pies if you *really* feel you need to use them.
  - Arrange the slices in a way that makes sense.
  - Don't use them for more than 2-3 categories.
  - Don't use 3D. Ever.
  - Add numeric values as labels so that the end user doesn't have to guess. It's also usually helpful to put the category in the label instead of using a legend.
  - Don't "explode" your pies.
  - Don't use pies for questions that allow more than one response.

Excel Tricks

- Trying to get Excel to do something it's not really designed to do? Check the Peltier Tech site.
- https://peltiertech.com/Excel/Charts/ChartIndex.html

More on Color Choices

- Semantically resonant colors
  - https://hbr.org/2014/04/the-right-colors-make-data-easier-to-read



More on Color Choices

- If you know someone in the audience is color-blind or will be printing the presentation, there are specific palettes that are “friendly”.
  - http://www.vischeck.com has examples of what various images look like to individuals with color-blindness
  - http://colorbrewer2.org gives sample “friendly” palettes

Other Resources

There are several great data viz practitioners with excellent books and websites. Some to look for:

- Stephen Few
- Edward Tufte
- Nathan Yau
- Albert Cairo
- Cole Nussbaumer Knaflic
- Junk Charts
  http://junkcharts.typepad.com/
- WTF Visualizations
  http://viz.wtf/