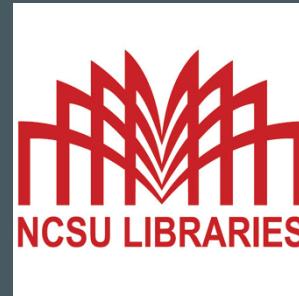


[go.ncsu.edu/startviz2](http://go.ncsu.edu/startviz2)

# Getting Started with Data Visualization: Elements of Design!

...

Alison Blaine, Jennifer Garrett, & Hannah Rainey



# Goals

After this session you should leave with:

- an understanding of the process for making good visualizations
- an introduction to core design concepts
- exposure to various chart types
- practice doing a visualization “make over”

Most importantly, we hope you learn that you can always get more help!

# Agenda

1. Overview
2. Basic Design Considerations
3. Characteristics of Bad & Good Visualizations
4. Best practices for visualizing data
5. Resources for choosing the right chart type
6. Hands-on activity

Workshop materials folder

[go.ncsu.edu/startviz2](http://go.ncsu.edu/startviz2)

[go.ncsu.edu/startviz2](http://go.ncsu.edu/startviz2)

# What is data visualization?

---

## Why visualize?

# Grammar of Graphics

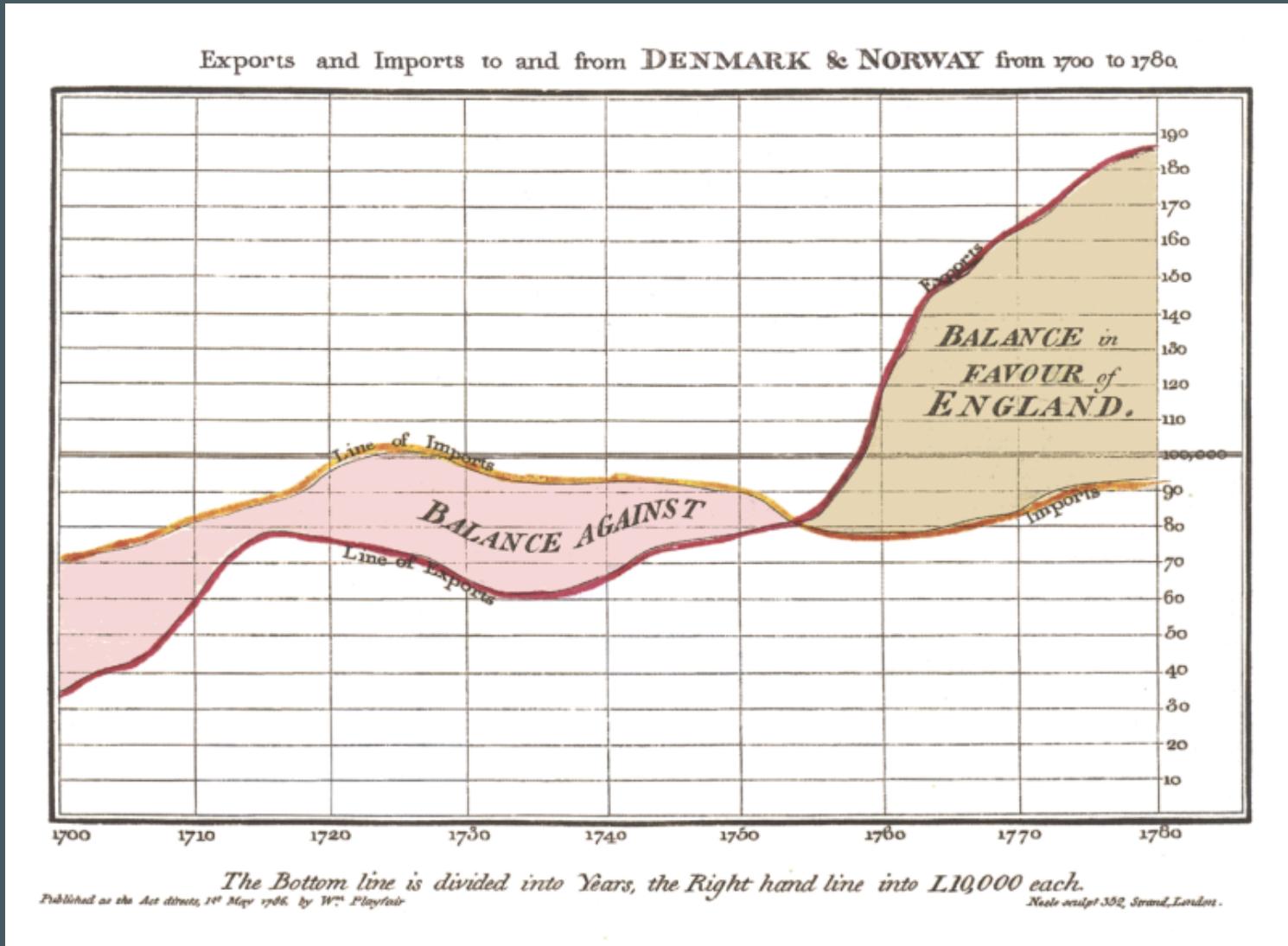
Originated by [Leland Wilkinson](#), simplified by [Hadley Wickham](#) and others. A “grammar” for creating graphics gives us tools to create new types of visualizations. Data visualization is taking data and mapping it onto visual aesthetic, (points, lines, bars, etc).

- Data
- Geometry - visual encoding or symbol onto which the data is mapped. Each geom. has aesthetics (length, position, etc)
- Grouping - subsets the data
- Statistics - transforms or summarizes data
- Mapping - maps the data onto the geometry

[Source](#)

# Data Visualization is an old endeavor.

William Playfair, *Commercial and Political Atlas*, 1786



# New Visualization Challenges



# Enduring Challenges



By Diego Delso, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=35691051>

# Interoperability Between Open Source Tools



# Proliferation of Visualization Software

## Commercial

AutoCAD

GEOPAK

ArcGIS

SAS Visual Analytics

Tableau Desktop

Find software at NCSU

Libraries:

[www.lib.ncsu.edu/software](http://www.lib.ncsu.edu/software)

## Open Source

R (ggplot2, ggvis, shiny)

Python (Bokeh, Seaborn,  
Altair)

Plotly

Tableau Public

Vega-Lite

VTK

ParaView

QGIS

Carto

D3 & other .js libraries

# Data Visualization Process

getting data

getting to know the data

setting goals

determining if a visualization is needed

deciding what to visualize

cleaning

**visualizing**

**designing**

exporting/embedding for presentation

# Basic Design Considerations

Color

Type

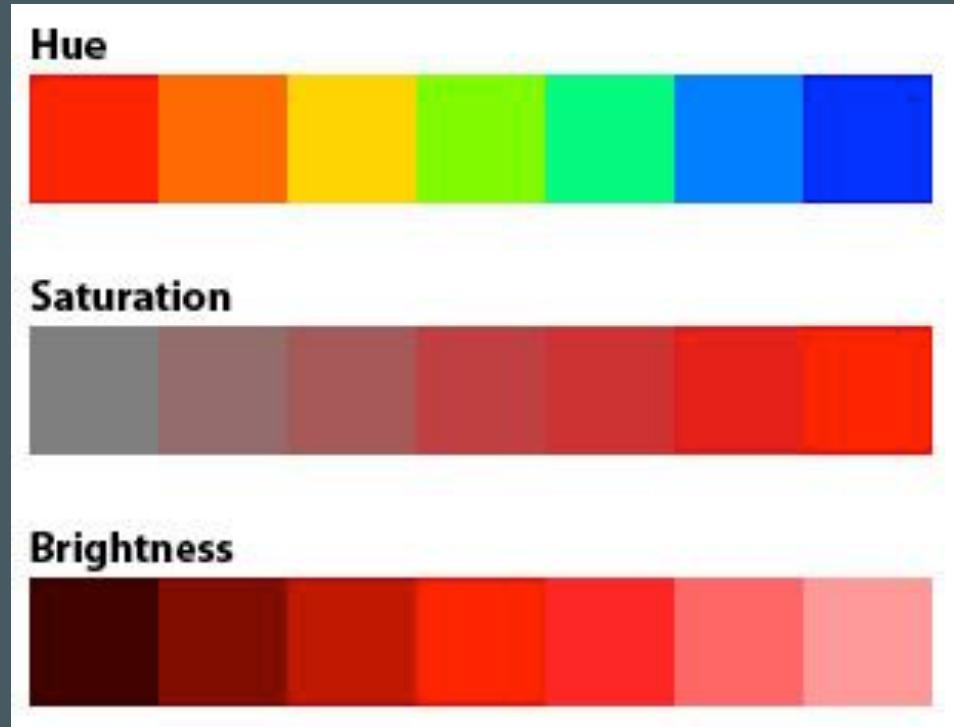
S P A C E

# Color Characteristics

Hue = color name

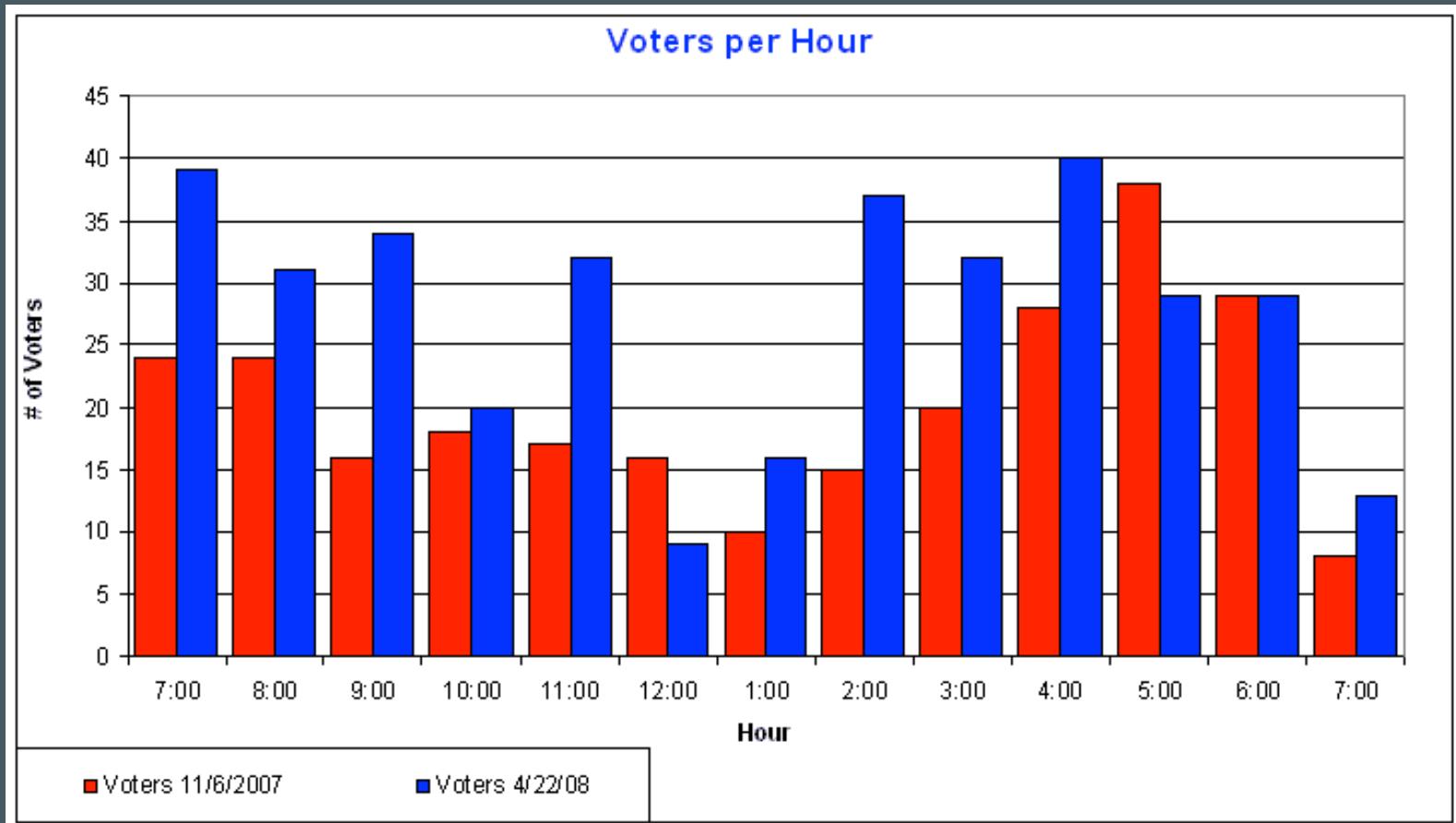
Chroma = saturation

Value = brightness



- Choose different hues when comparing categories that are not ordered. Consider grayer or more pastel shades over bold hues.
- Keep same hue but vary chroma or brightness to show an ordered relationship

# Color has cultural significance



Source: [http://ridgerunnerconsulting.com/2008/04/more\\_voting\\_statistics/](http://ridgerunnerconsulting.com/2008/04/more_voting_statistics/)

# Tools for Picking Color Wisely

Color Brewer 2

Adobe Color Wheel

Number of data classes: 3

Nature of your data:  
 sequential  diverging  qualitative

Pick a color scheme:  
Multi-hue:   
Single hue:

Only show:  
 colorblind safe  
 print friendly  
 photocopy safe

Context:  
 roads  
 cities  
 borders

Background:  
 solid color  terrain

color transparency

how to use | updates | downloads | credits

COLORBREWER 2.0  
color advice for cartography

3-class BuPu

EXPORT

HEX

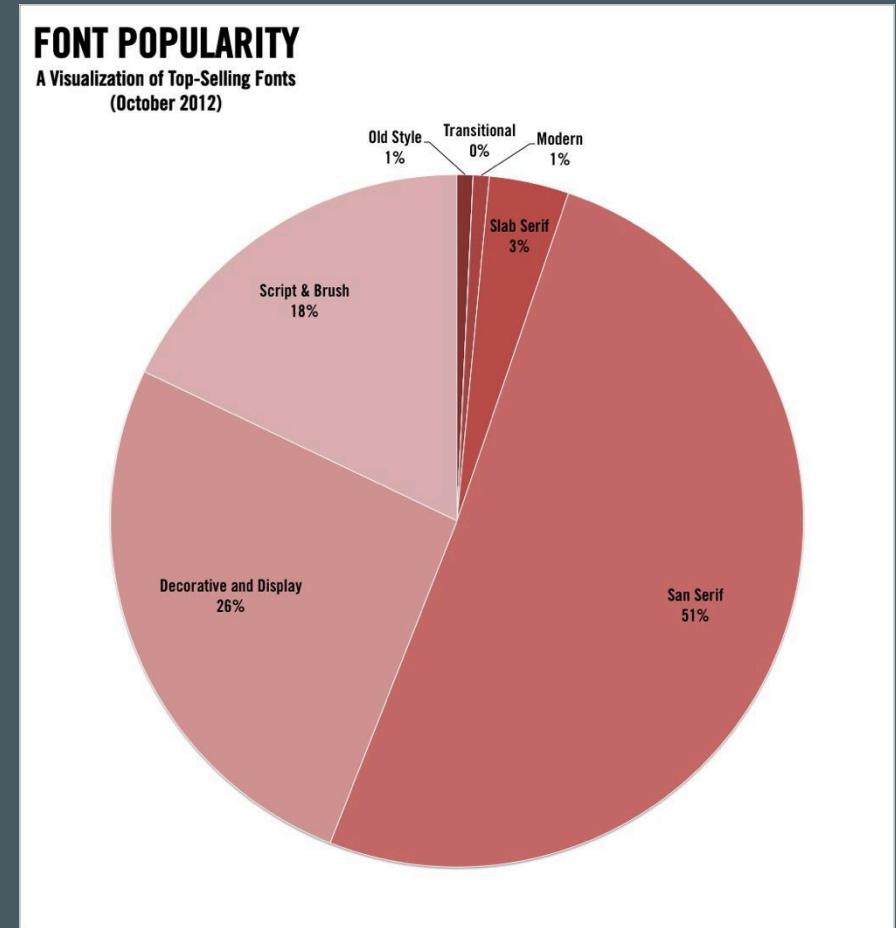
#e0ecf4  
#9ebcda  
#8856a7

# Type



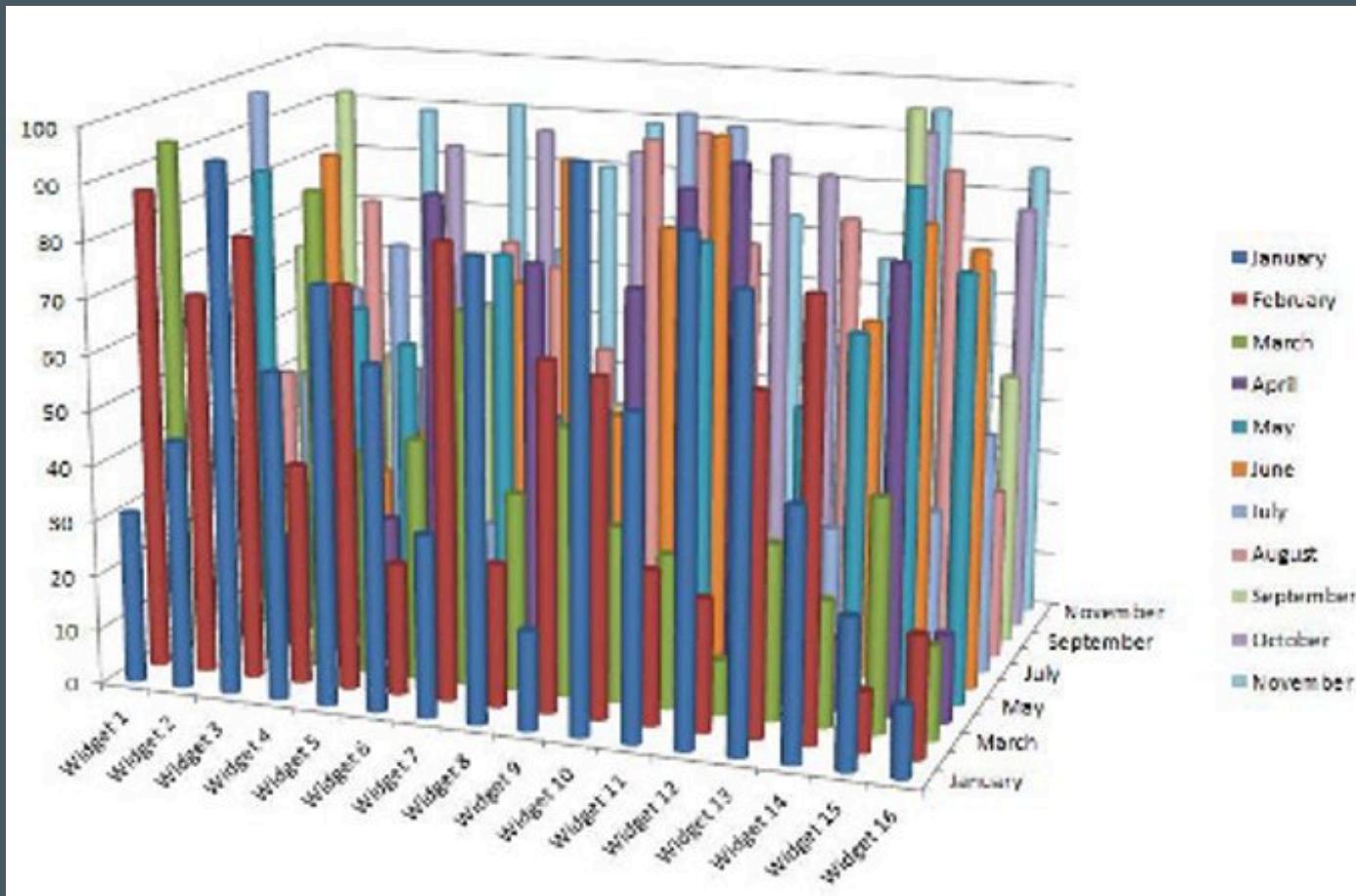
Source

Sans Serif fonts are better for digital displays and smaller captions  
Limit the use of different font types to two!



Source

# Space

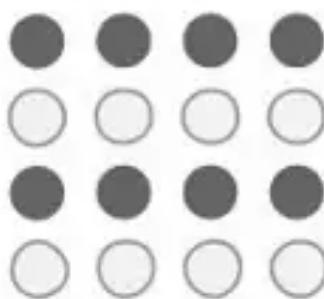


# Gestalt principles

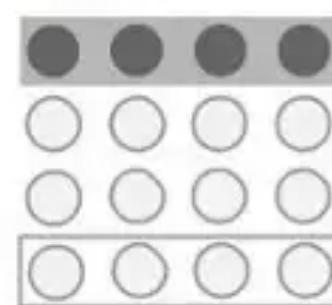
Proximity



Similarity



Enclosure



Symmetry



Closure



Continuity



Connection

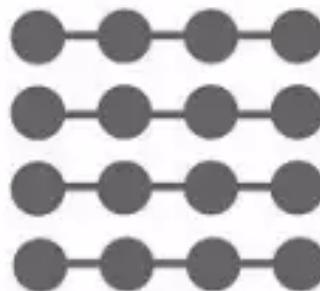


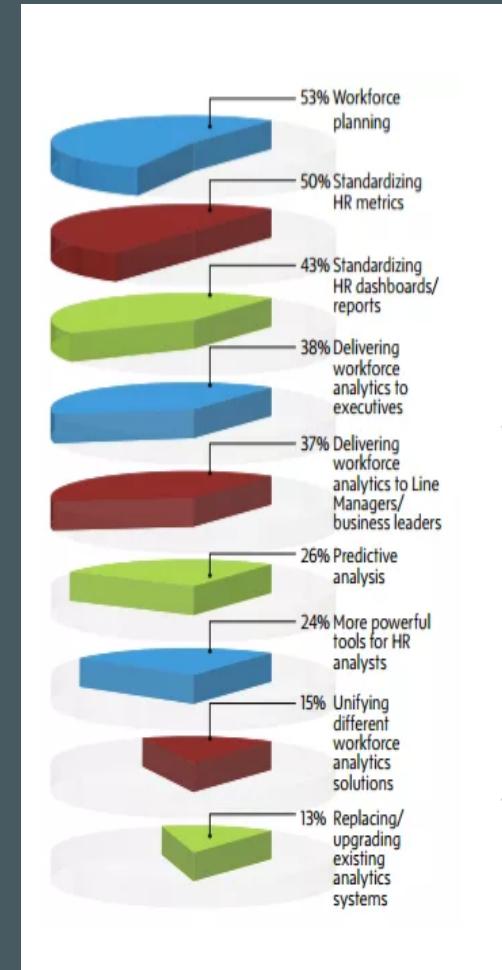
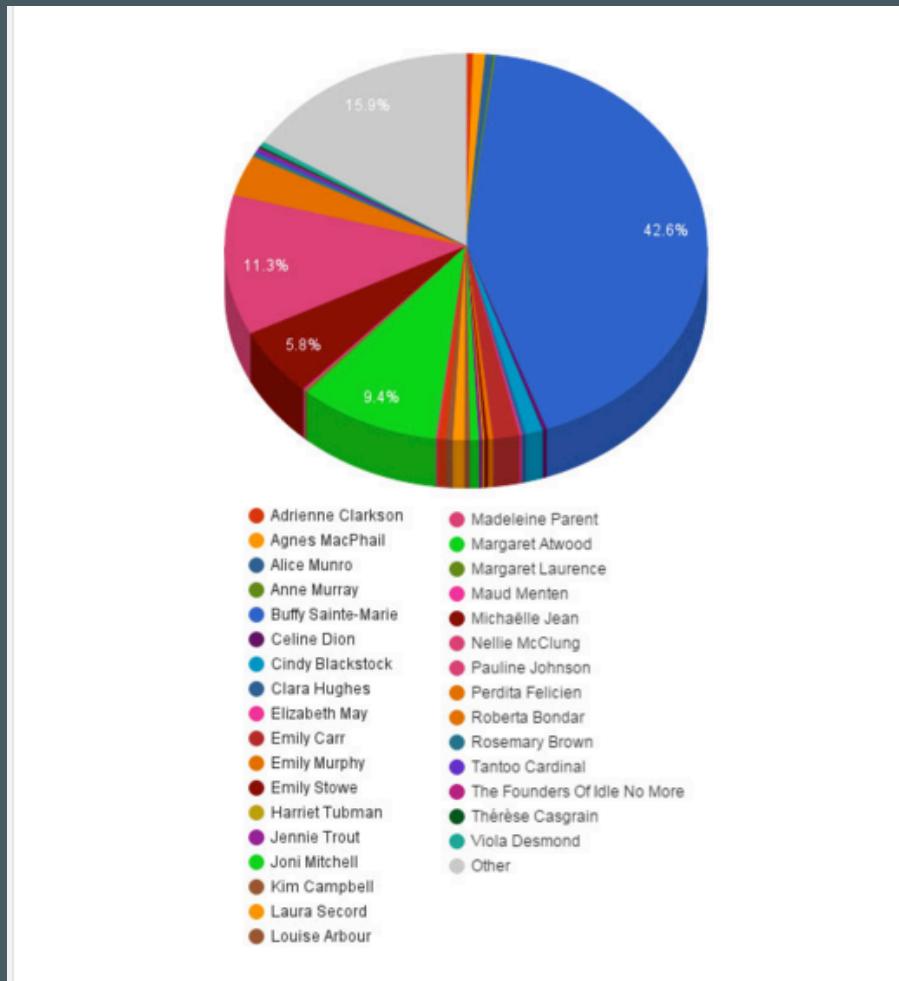
Figure & ground



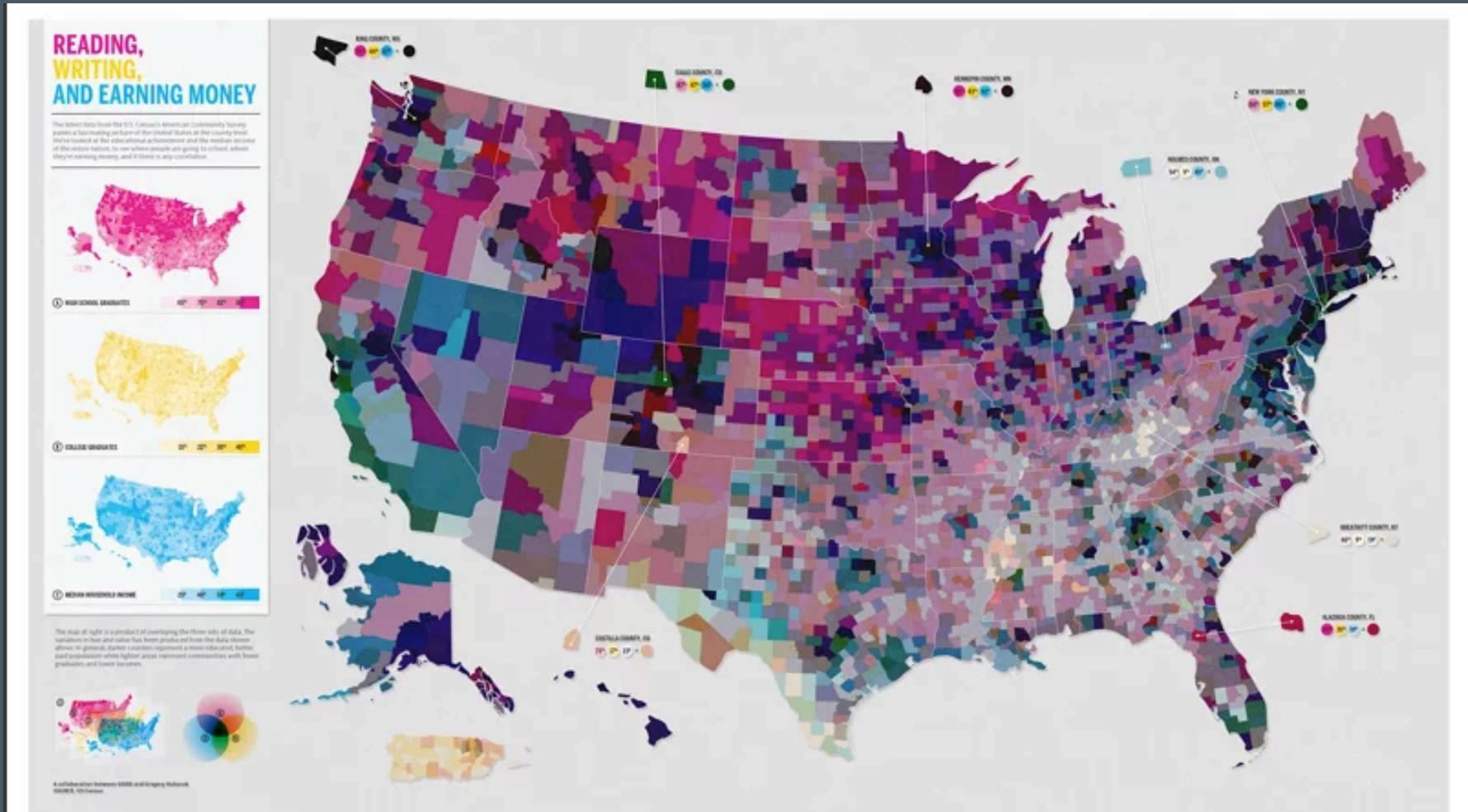
[Source](#)

# Good vs. “Bad” Visualizations

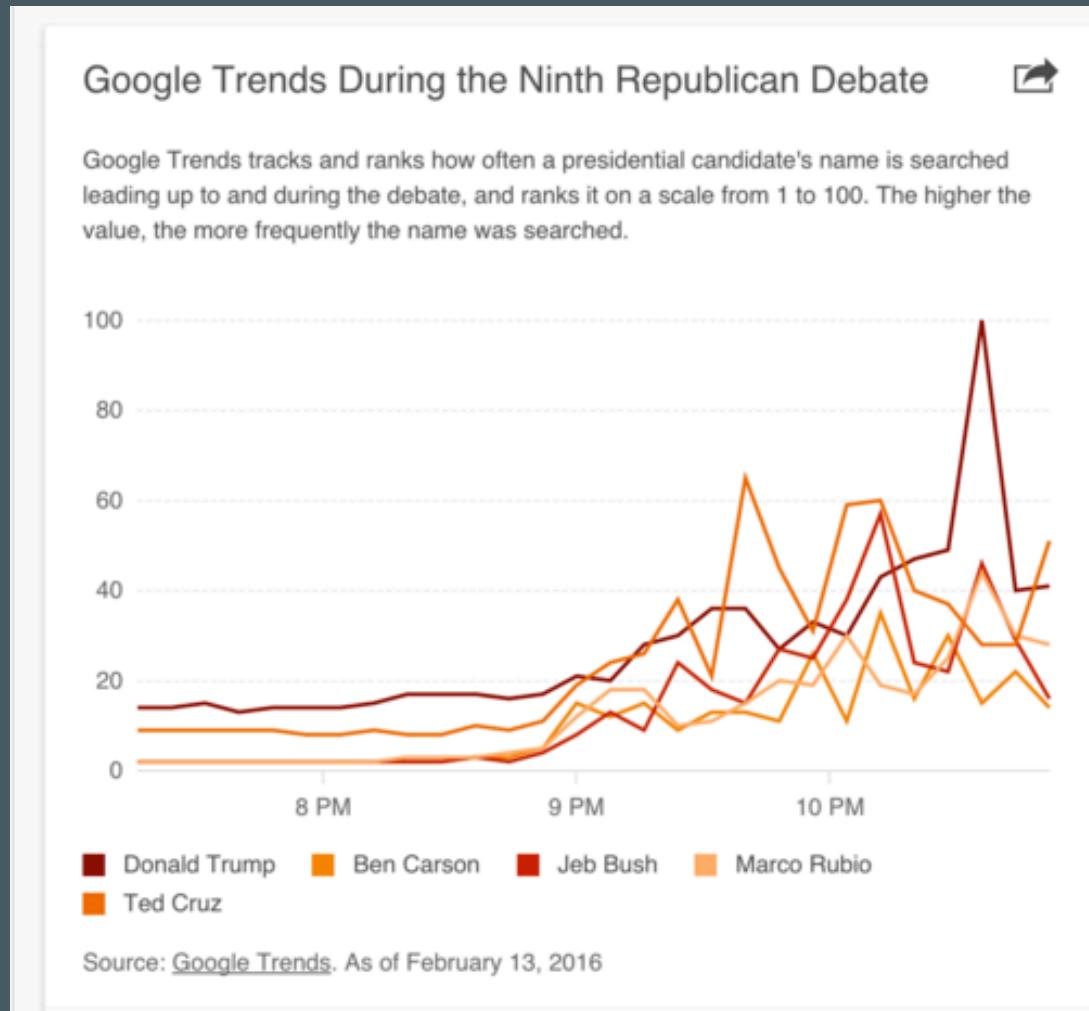
# Characteristics of bad visualizations



# Bad visualizations confuse rather than reveal



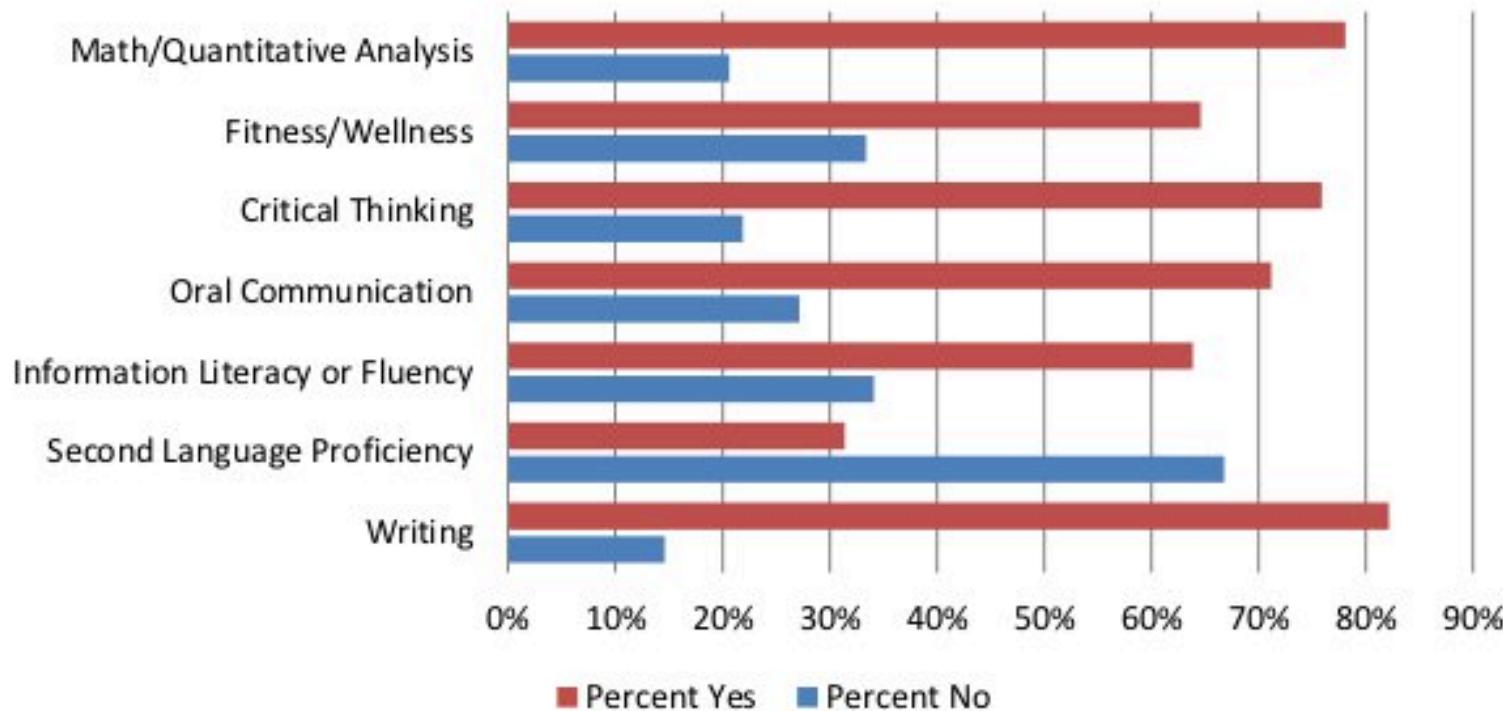
# Bad visualizations are disorienting



# Mediocre visualizations aren't good

## Should this Skill Be Included?

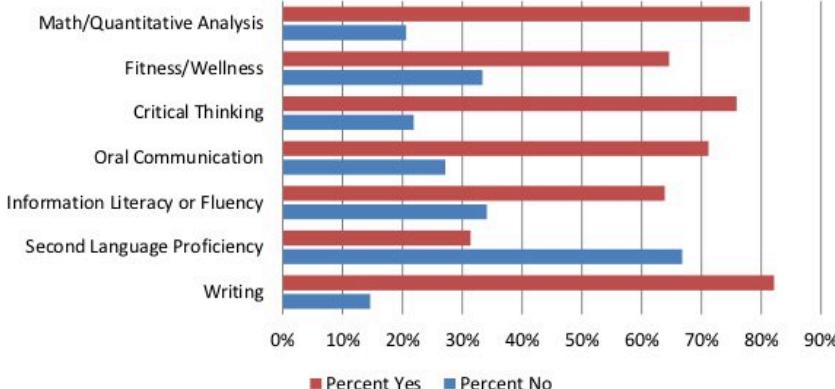
N=548



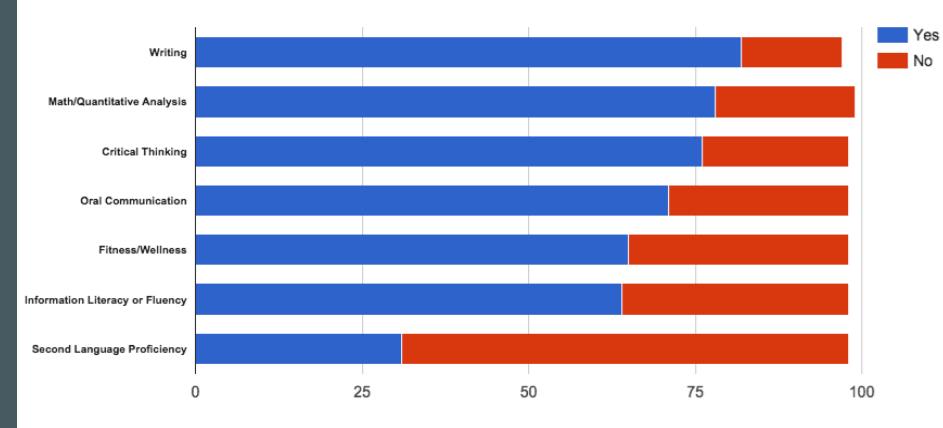
# Visualization Makeover: Make it easier to interpret

Should this Skill Be Included?

N=548



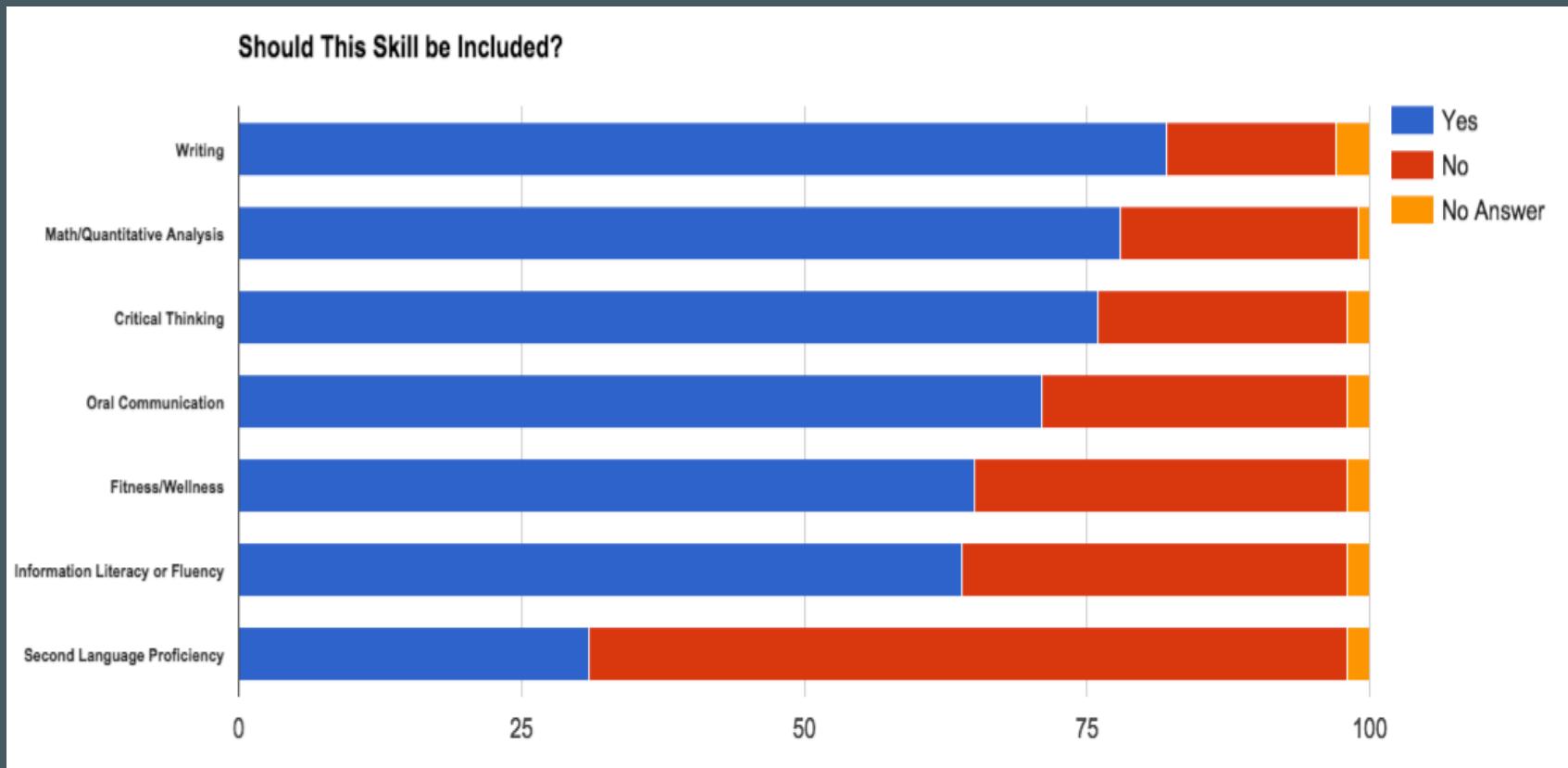
Should this Skill Be Included?



Regular bar chart, unordered

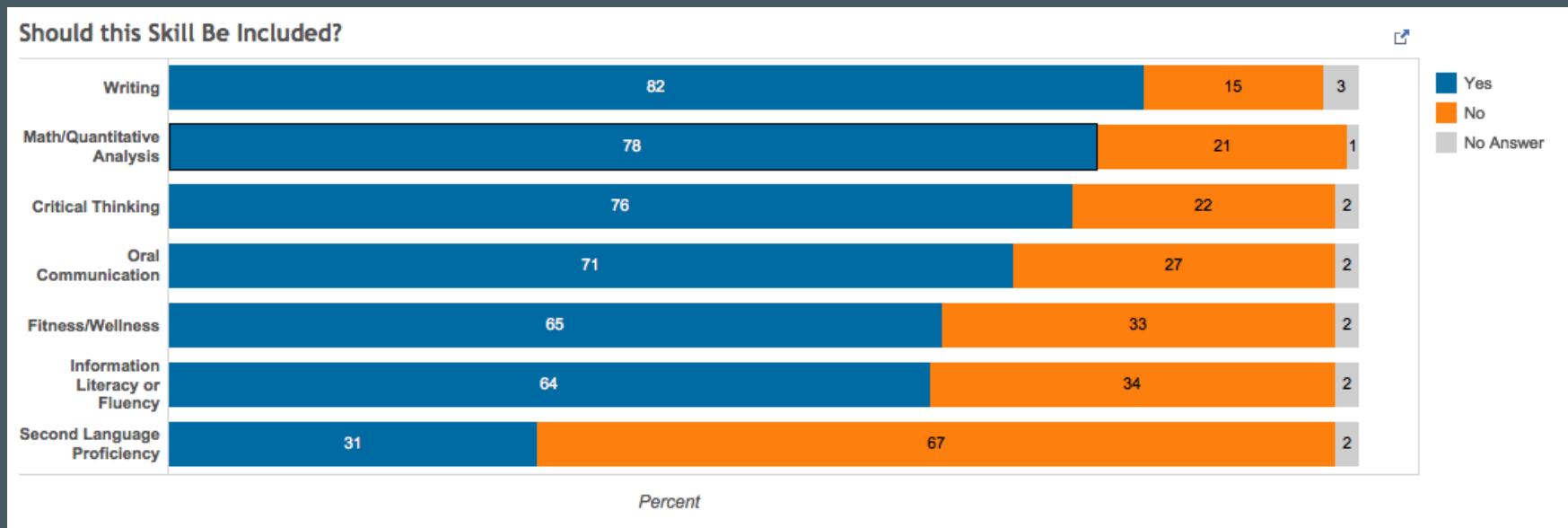
Stacked bar chart, ordered by % Yes

# Correct misleading visuals



Tool used: Google sheets

# Taking it a step farther: Labels and color



Tool used: Tableau Public ([www.tableau.com](http://www.tableau.com))

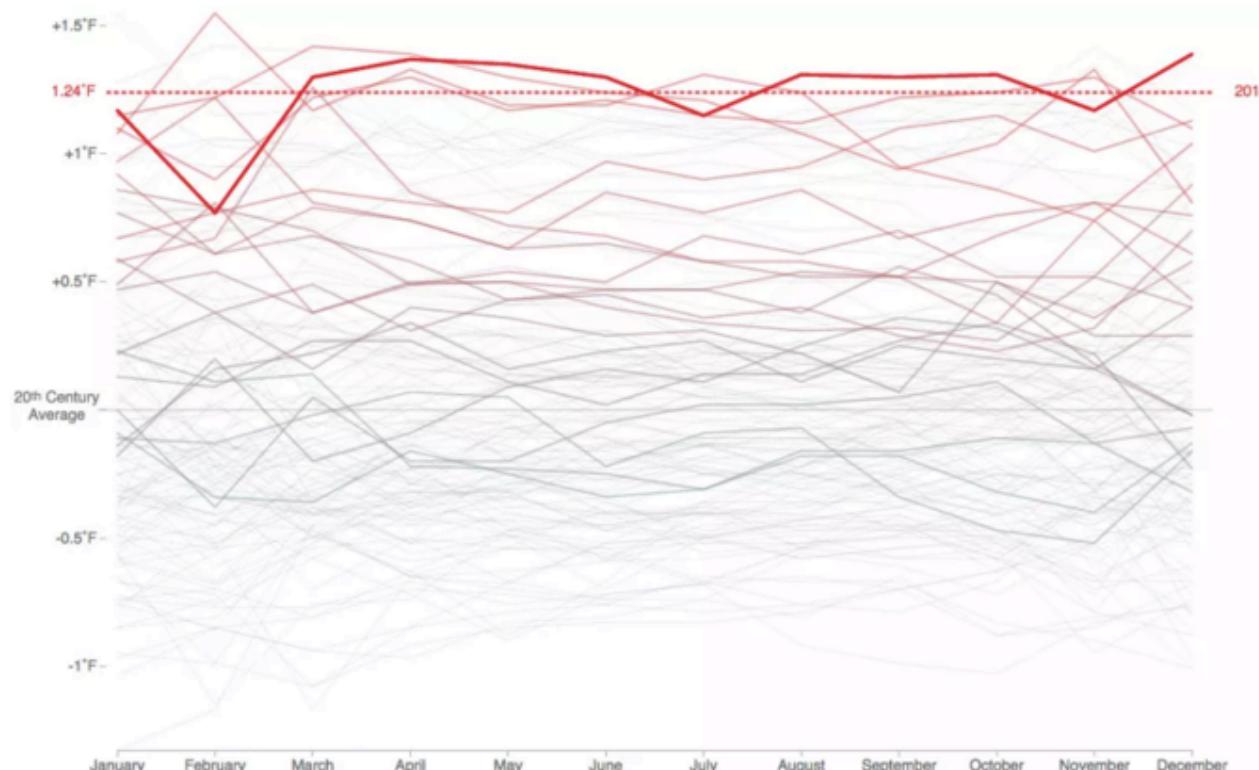
# Best practices for making good visualizations

1. Make them easy to interpret (8-second rule).
2. Tell a story with annotation (titles, subtitles, and notations on the graph near the data)
3. Do not use hue to convey order. Blue does not come before yellow.
4. Utilize the length, size and/or position of visual marks to convey quantitative information.
5. Provide multiple levels of detail.
6. Consider color blindness.
7. Familiarize yourself with Gestalt Principles.

# Good visualizations point out what's important. Bold colors used to highlight only.

## 8. 2014 Was the Hottest Year on Record

The obvious way to visualize the data was as a single line chart. A split-up view got the point across much better.

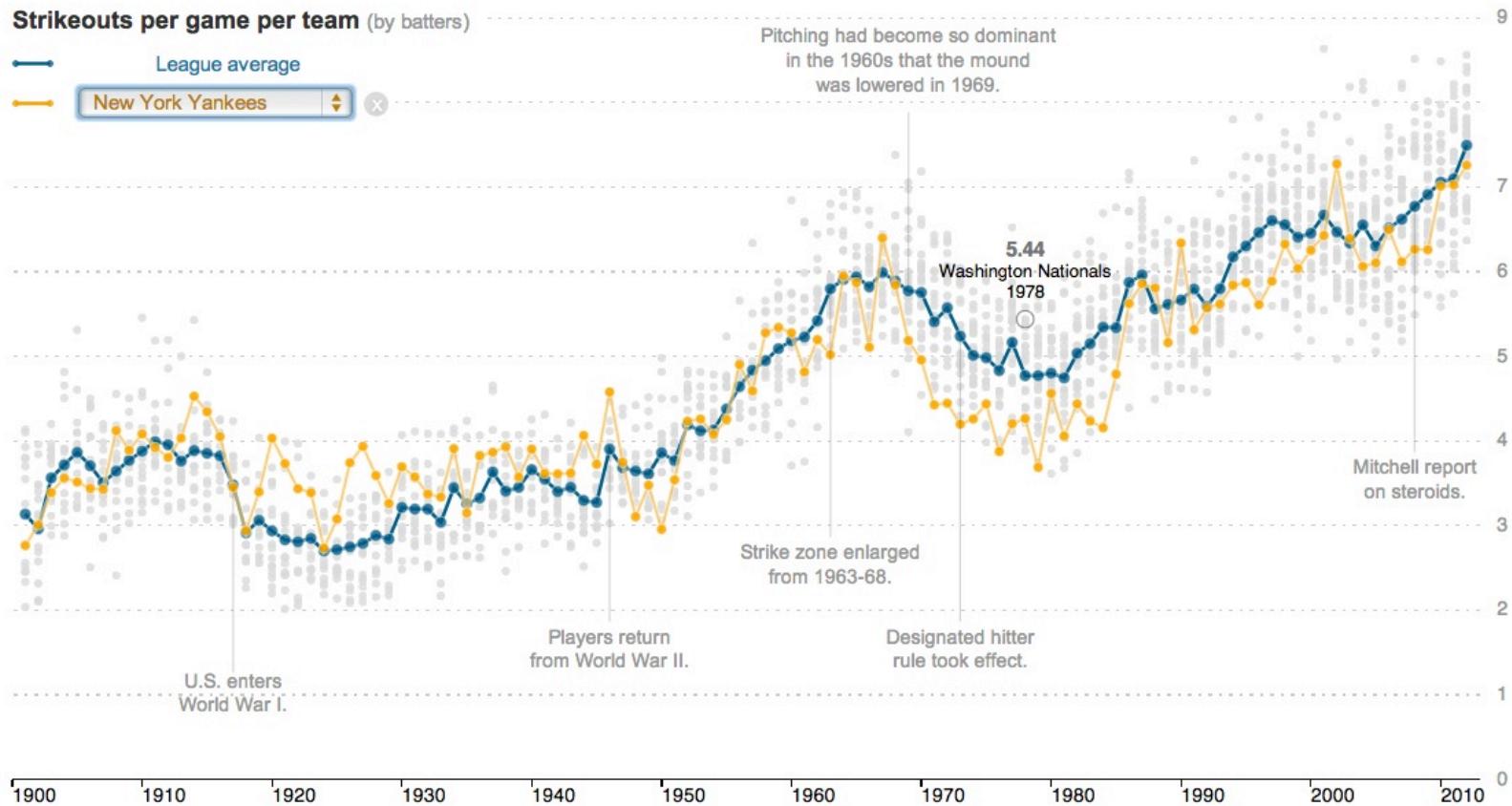


Source:  
Flowingdata.com

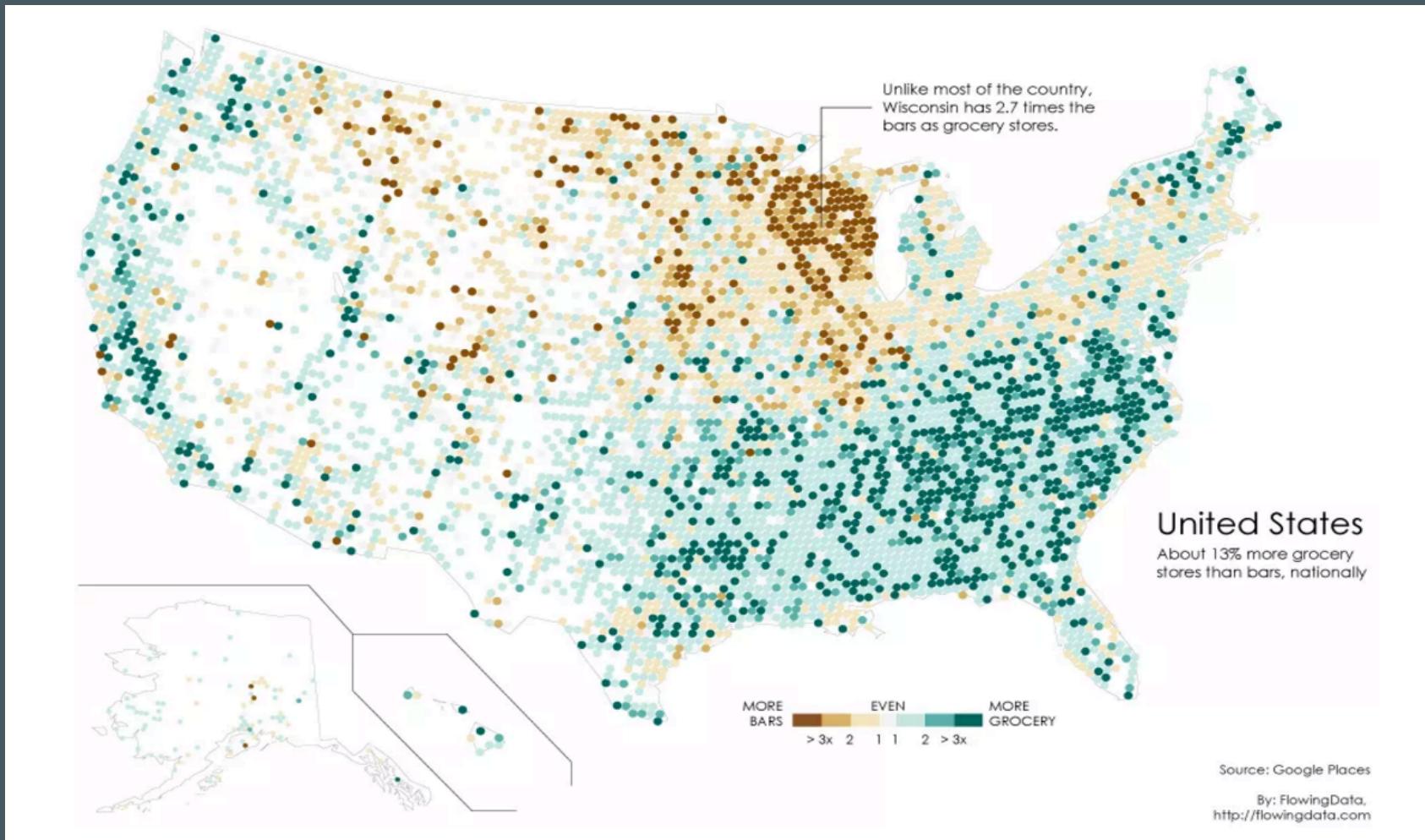
# Good visualizations tell a story

## Strikeouts on the Rise

There were more strikeouts in 2012 than at any other time in major league history.



# Good visualizations use color and notation to aid interpretation



# Choosing the Right Chart Type

(or rather, geometries and aesthetics)

# Research findings

From Robert Kosara

The Science of What We Do (And Don't) Know About Data

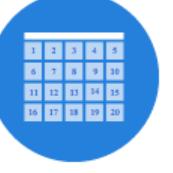
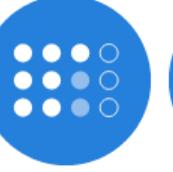
Visualization:

“...lines imply transitions whereas bars imply individual values.”

“We remember points not where they are in the plot, but shift them towards clusters in our memory, and let them drift slightly downwards.”

# Data Viz Catalogue - [www.datavizcatalogue.com](http://www.datavizcatalogue.com)

Search by Function      View by List

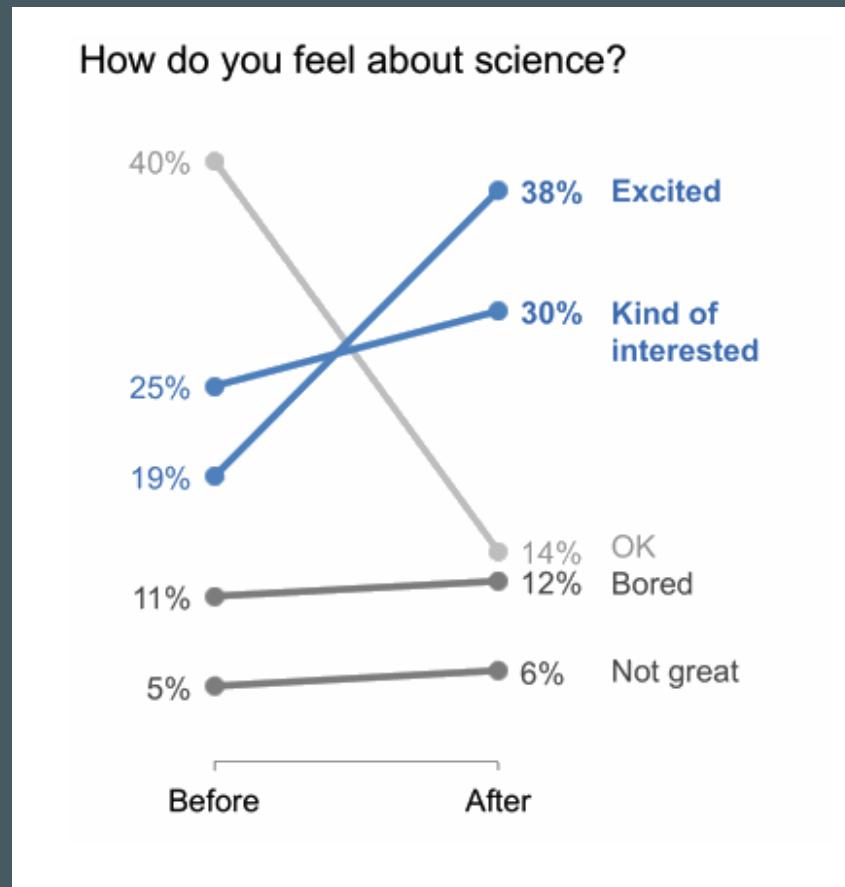
					
Arc Diagram	Area Graph	Bar Chart	Box & Whisker Plot	Brainstorm	Bubble Chart
					
Bubble Map	Bullet Graph	Calendar	Chord Diagram	Choropleth Map	Circle Packing
					
Connection Map	Density Plot	Donut Chart	Dot Map	Dot Matrix Chart	Flow Map
					
Histogram	Illustration Diagram	Line Graph	Marimekko Chart	Multi-set Bar Chart	Nightingale Rose Chart

# Alternatives to the status quo

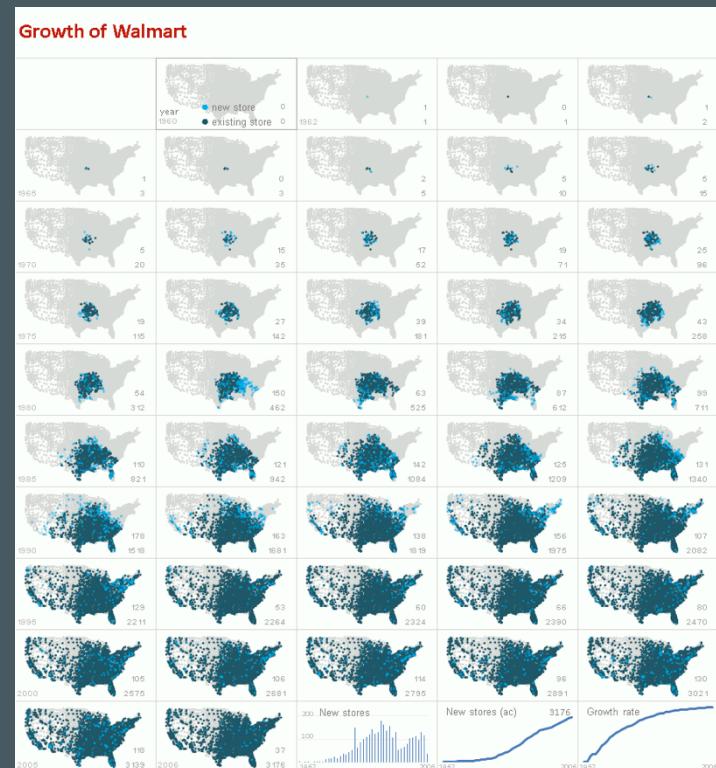
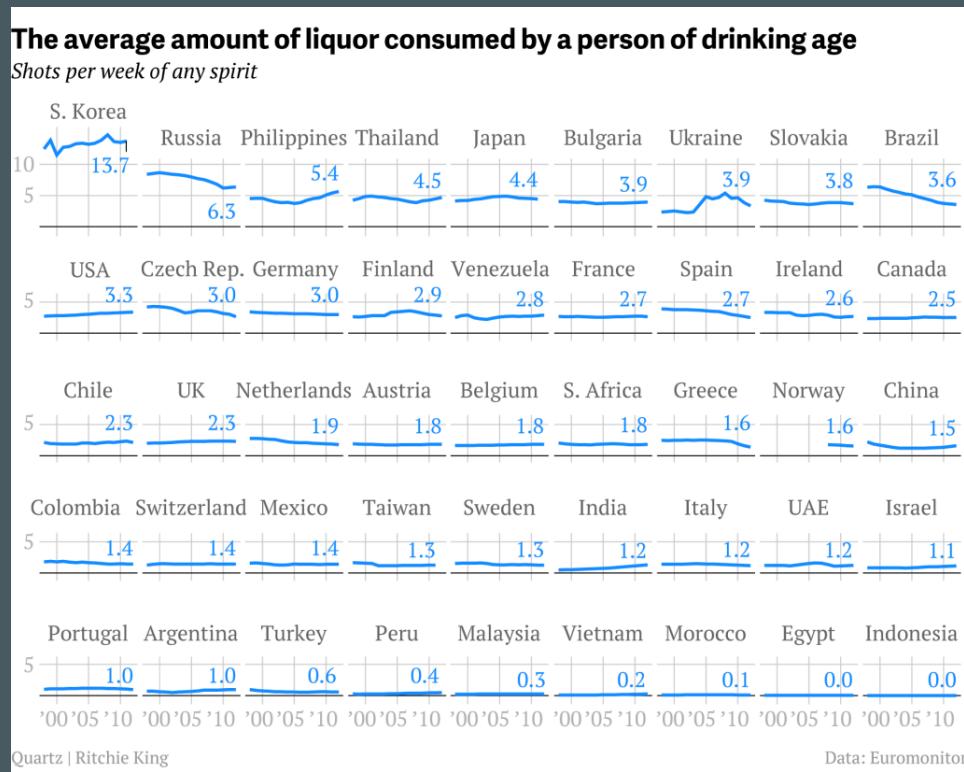
Common chart types may be fine, but there are others out there:

- “Six alternatives to the clustered bar chart” - Ann K. Emery  
[annkemery.com/clustered-bar-chart](http://annkemery.com/clustered-bar-chart)
- “Alternatives to pies” - Cole Nussbaumer Knaflic  
[storytellingwithdata.com/blog/2014/06/alternatives-to-pies](http://storytellingwithdata.com/blog/2014/06/alternatives-to-pies)

# Slope graphs: showing change over time



# Small multiples charts and maps: good for showing lots of data without overcrowding



# Further Resources

“Color Harmonies: Basic Techniques for Combining Colors” (Tiger Color website)

“Choosing Colors for Data Visualization” (Maureen Stone)

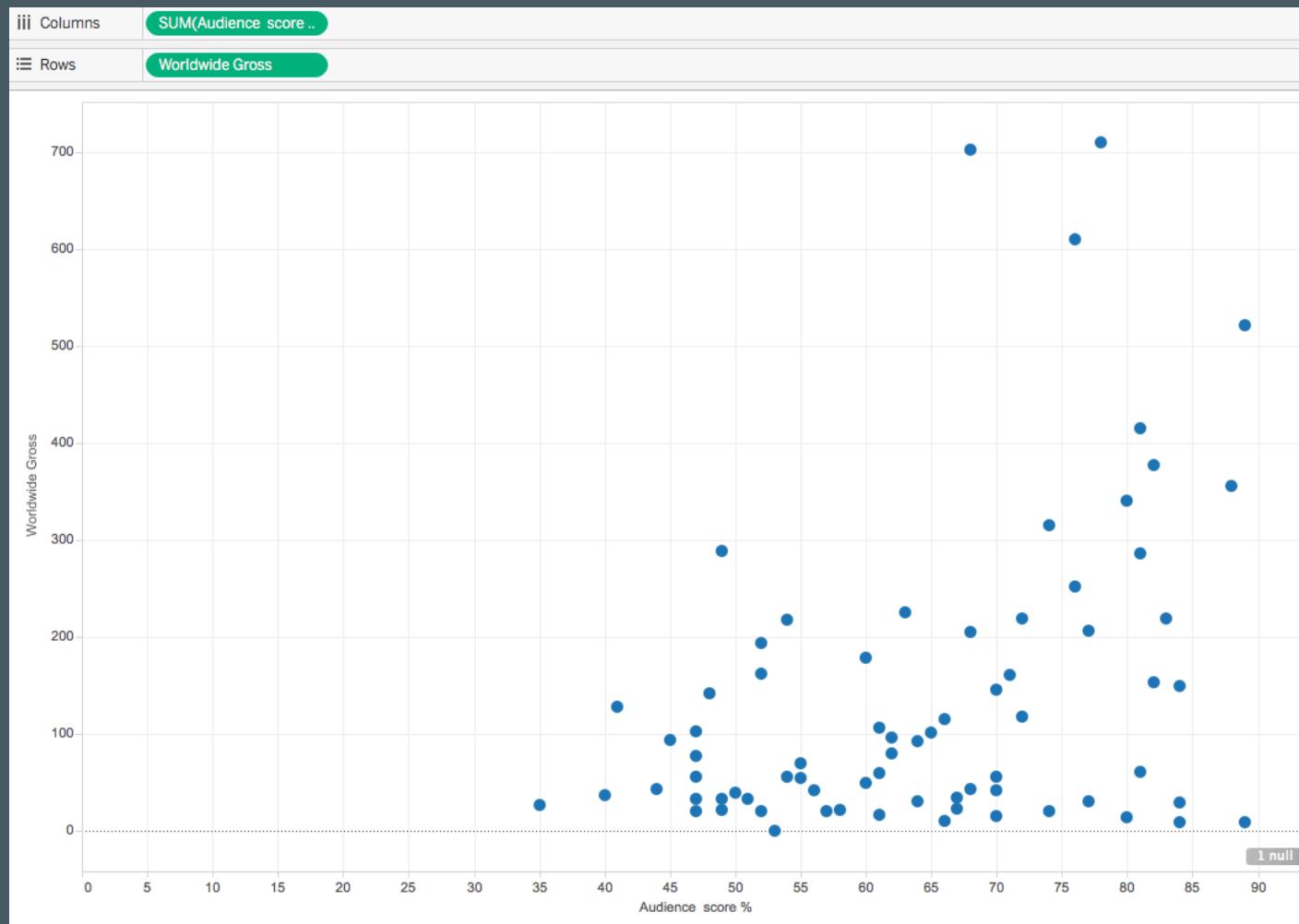
Data visualization [electronic resource] : a successful design process (Andy Kirk)

The functional art : an introduction to information graphics and visualization (Alberto Cairo)

[go.ncsu.edu/startviz2](http://go.ncsu.edu/startviz2)

Discuss examples on your handout  
in small groups

How could they be better?  
What would you do differently?



## Example 1: Scatterplot

This graph shows audience score (X-axis) and Worldwide Gross revenue in millions of dollars (Y-axis). The purpose of this graph is to see if there is a correlation between box office gross revenue and a film's audience score.

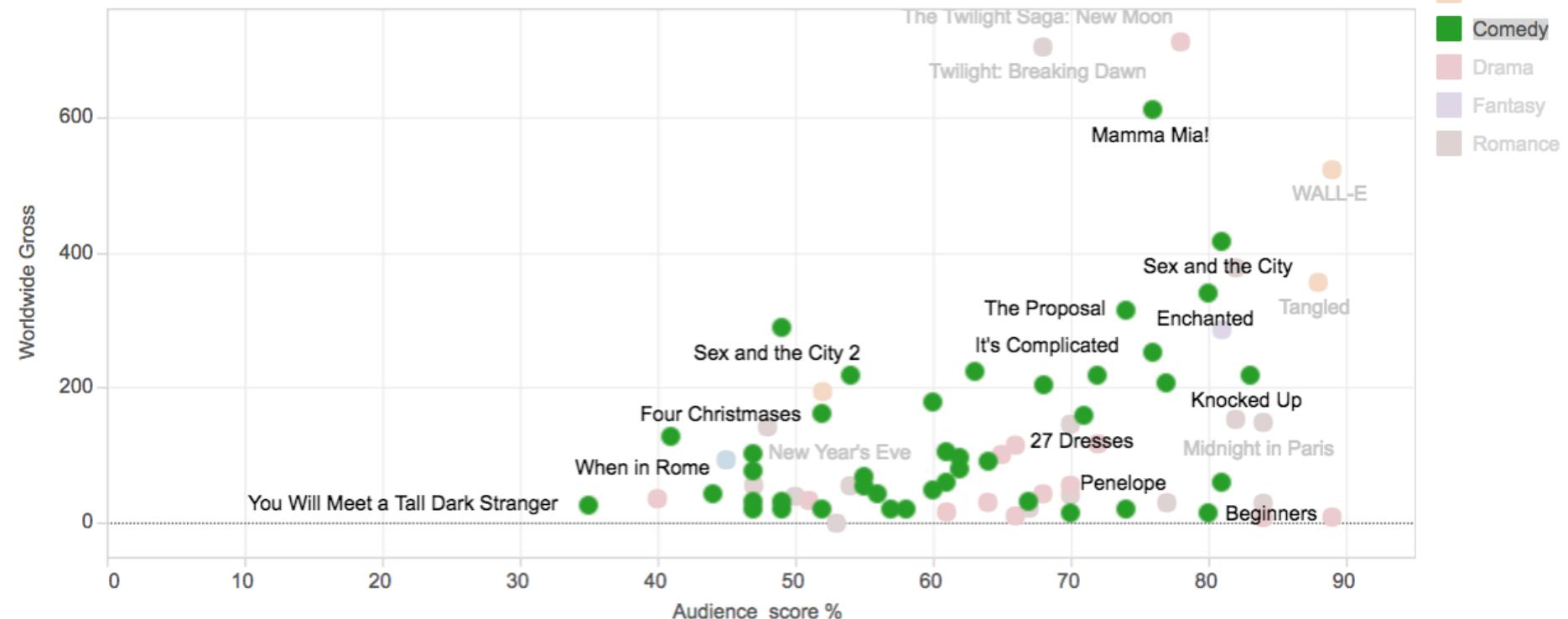
# Improved example

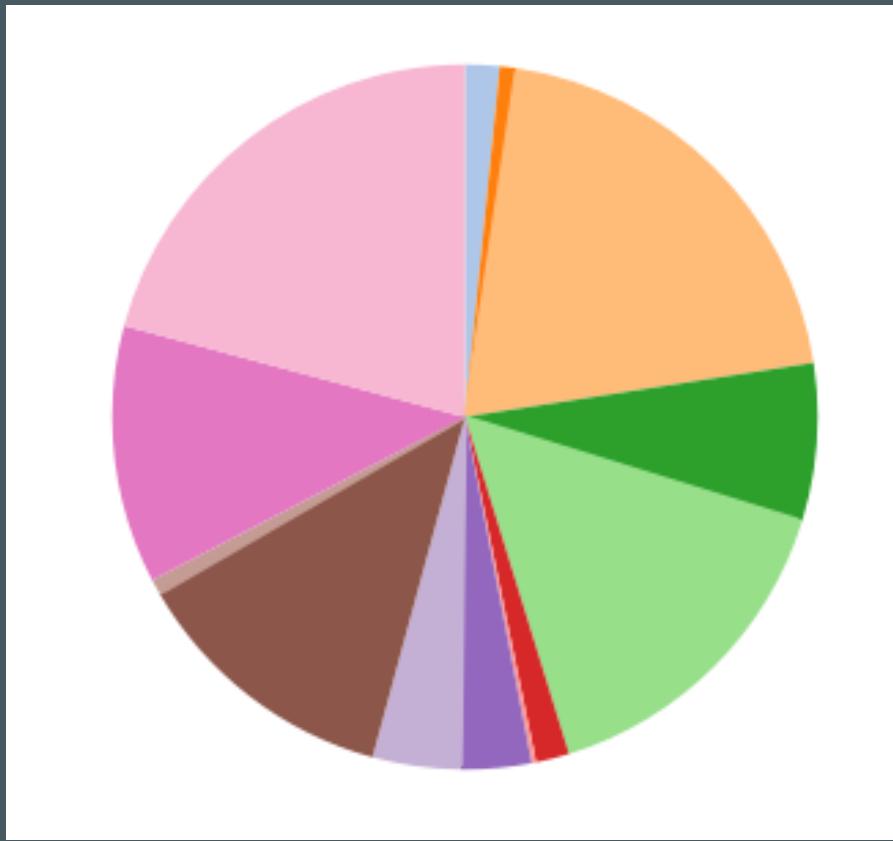
## Comedy Wins Best Genre In Dollars Earned

Among Hollywood's highest grossing films of all time, 55% are comedy; yet comedies fell just short of getting the highest audience scores.

Genre

- Action
- Animation
- Comedy
- Drama
- Fantasy
- Romance

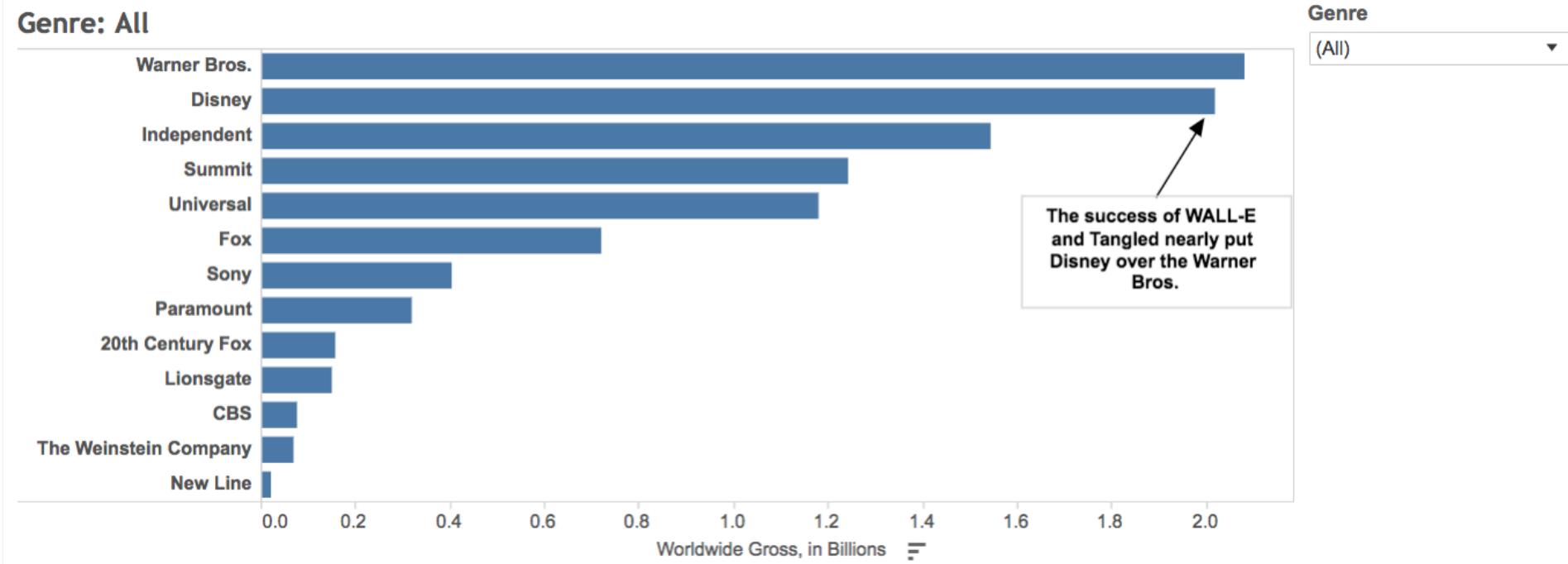


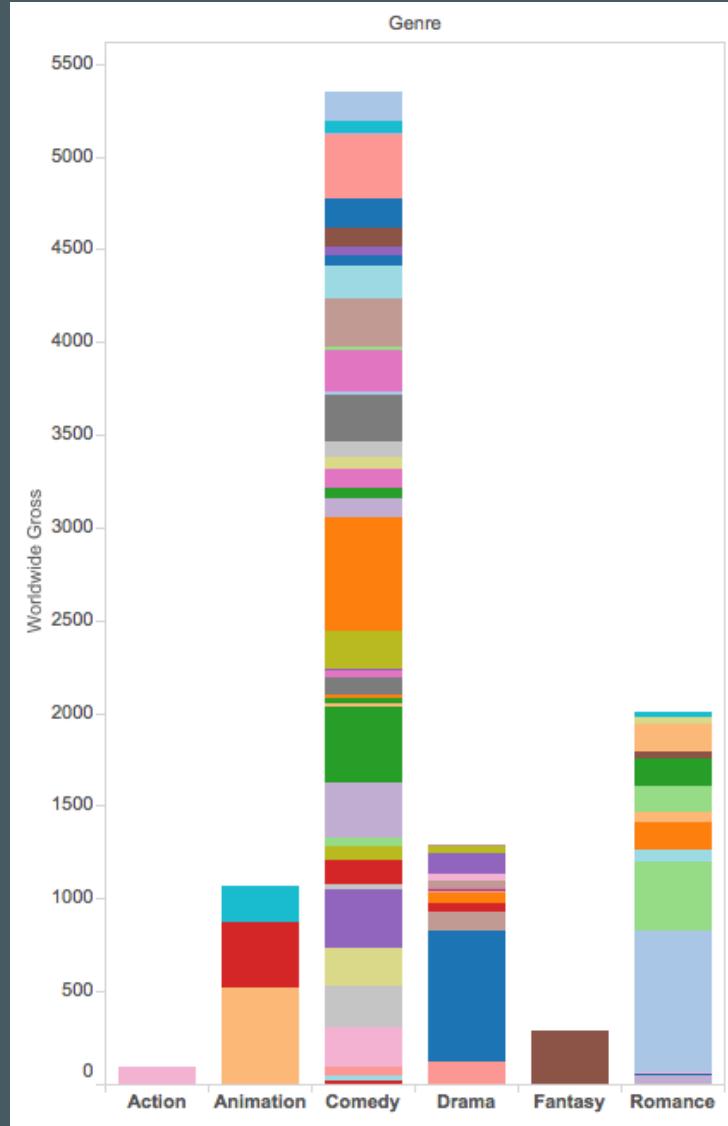


## Example 2: Pie Chart

This pie chart attempts to show how much film studios compare in terms of gross revenue.

# Improved example





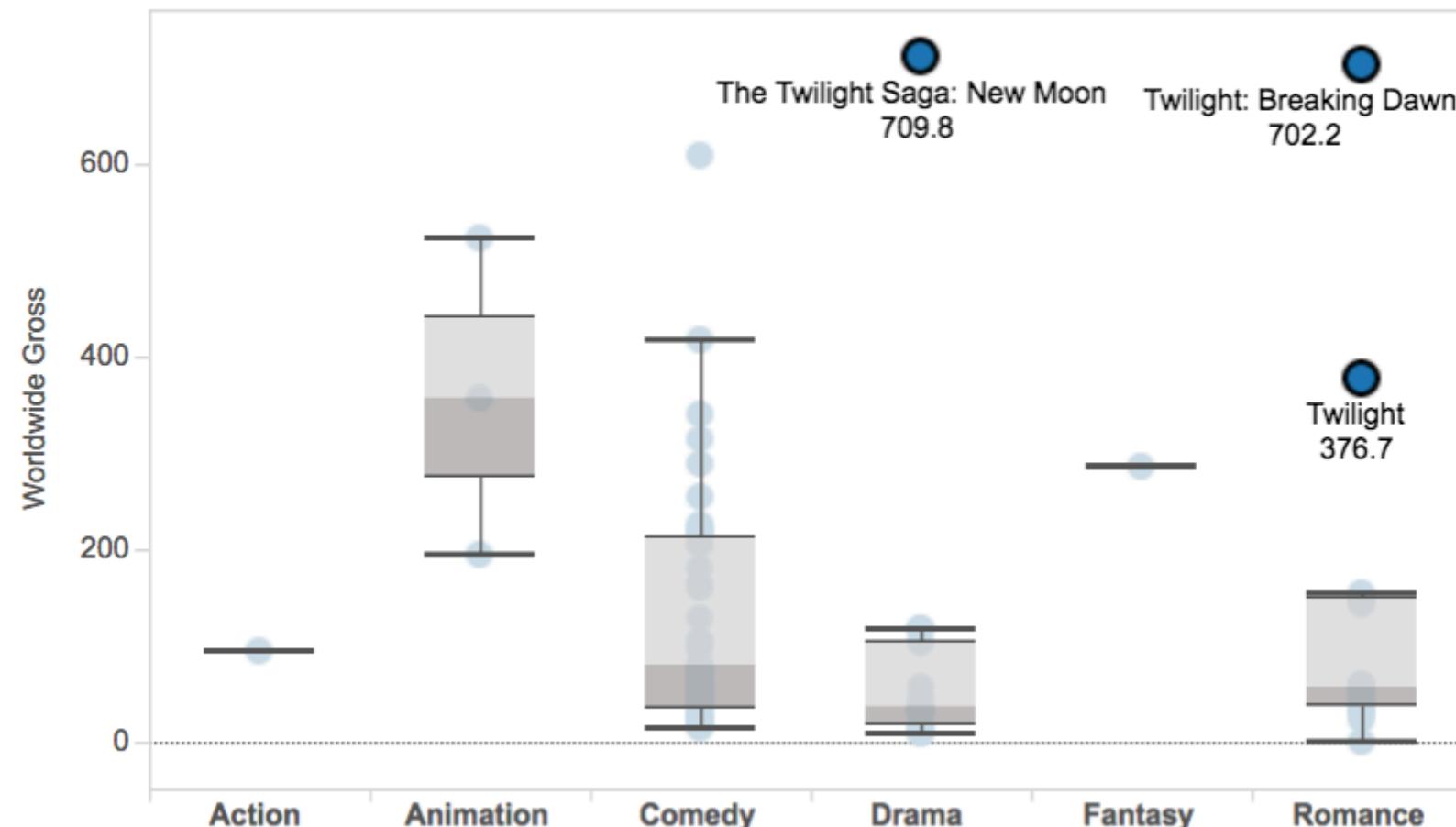
### Example 3: Bar chart

This bar chart shows worldwide gross by film and genre. It's too busy and needs help!

# Improved example

## Twilight Transcends

Among Hollywood's top grossing films, the Twilight movies are in the stratosphere in terms of dollars earned.



# Improved examples

Here's a link to a workbook with the examples from the activity:

[https://public.tableau.com/profile/al.blaine#/vizhome/ImprovedExamples/  
AudienceScorevsWorldwideGross](https://public.tableau.com/profile/al.blaine#/vizhome/ImprovedExamples/AudienceScorevsWorldwideGross)