Data Visualization & Design

Week 7

8 Different Ways to Visualize Changes Over Time

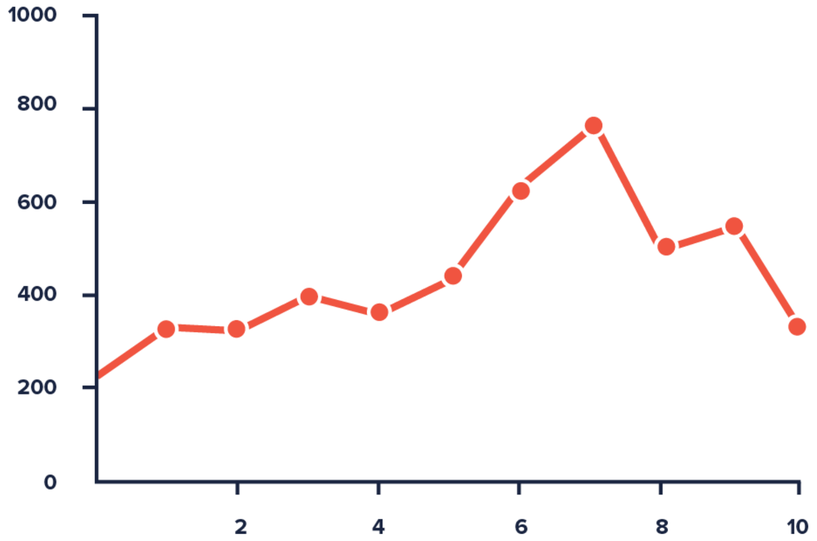
(see more at FlowingData)

Time-based data comes in many varieties: **increases**, **decreases**, **seasonal patterns**, **continuous**, **discrete**…

…**there is no one size fits all**, but the following approaches form good jumping-off points.

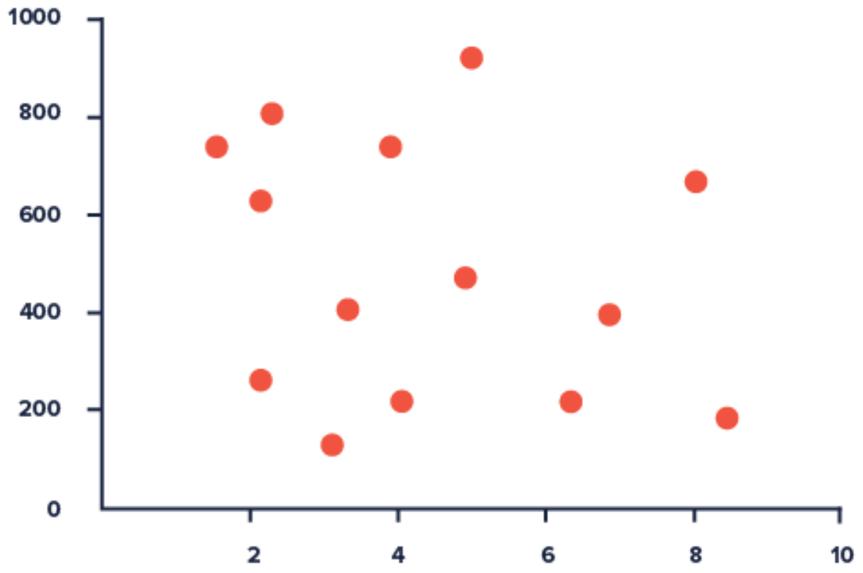
1. The **Line**

- The most basic option
- Works for most time series data
- Works with many points, or few
- Smoothing vs. adding points



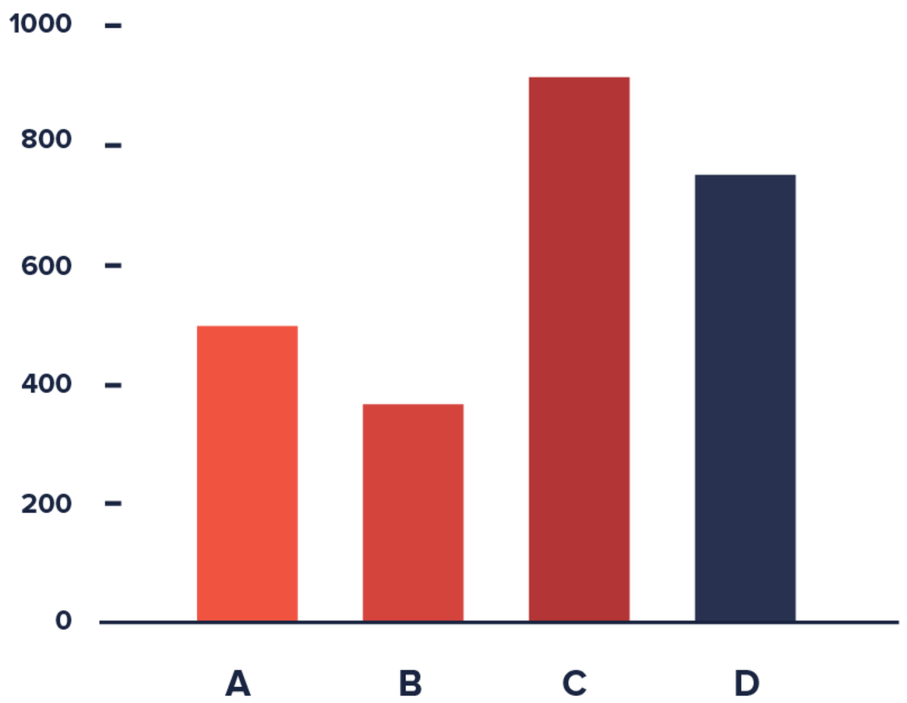
2. The **Scatter**

- Scatterplots work well with lots of data points
- Because dots are small, does not work well with few data points
- Work well when your measurements are not nicely structured (i.e. if increments are uneven, or if timestamps are missing)



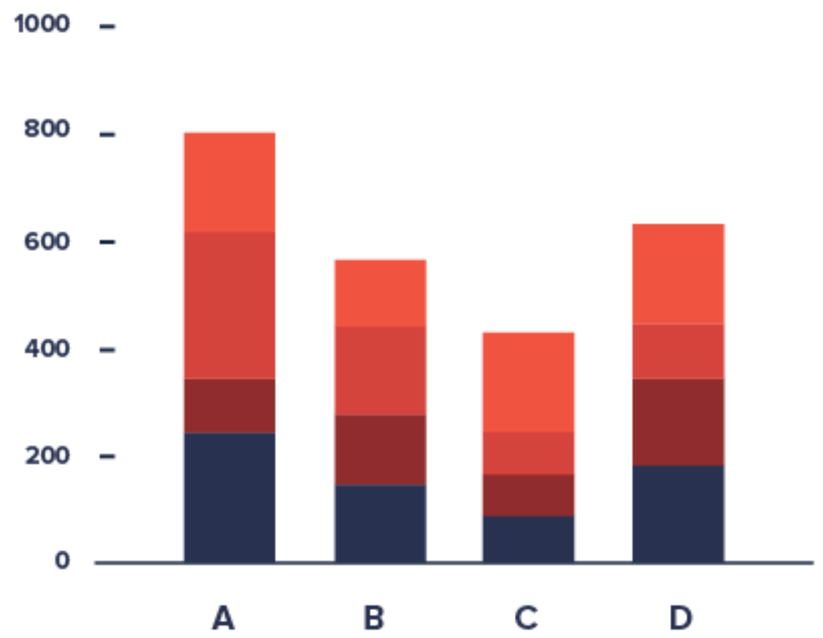
3. The **Bar**

- Work best for time series with distinct points in time, as opposed to more continuous data
- Tend to work better with data points evenly spaced through time



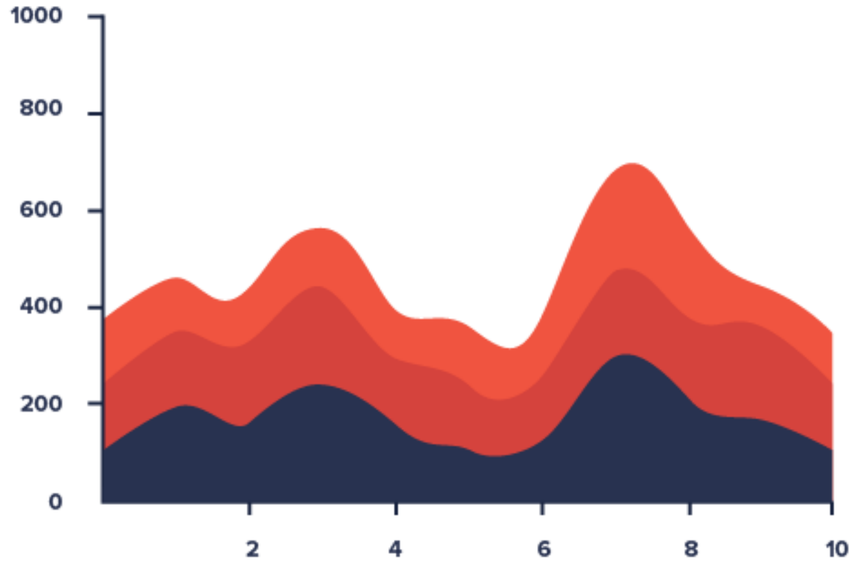
4. The **Stacked Bar**

- Use in the same way as a bar chart
- Accounts for multiple categories
- Stacks represent a significance in the sum of parts
- Keep the stacks ordered



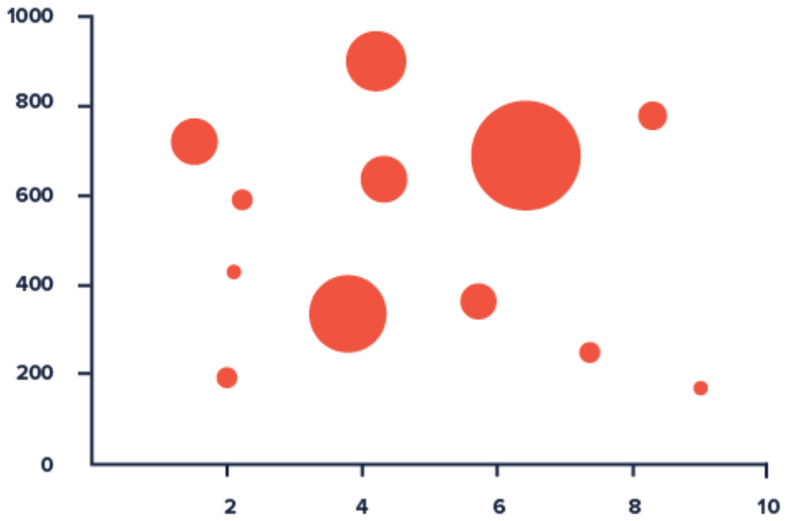
5. The **Stacked Area**

- Like a continuous stacked bar
- Use if you have a lot of data points in time, and not enough room for bars



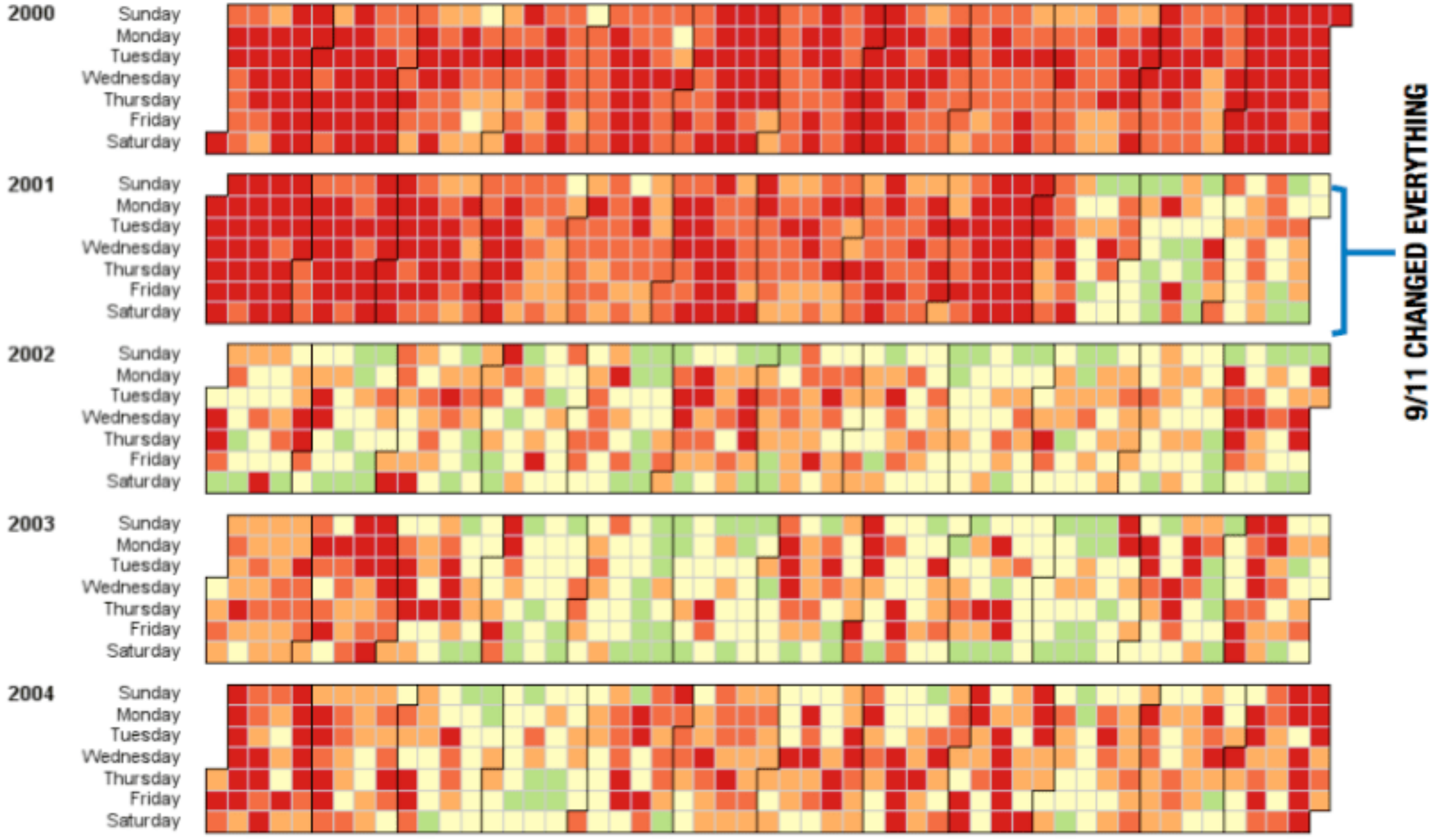
6. The **Bubble**

- Like a scatterplot
- Circles sized by some other metric
- (Popularized by Hans Rosling's TED talks)



7. The **Color Scale**

- Tends to be underutilized
- Easier to see differences in height than differences in gray
- Good if you are limited by space
- Choose a palette that is both perceptually distinct and accessible

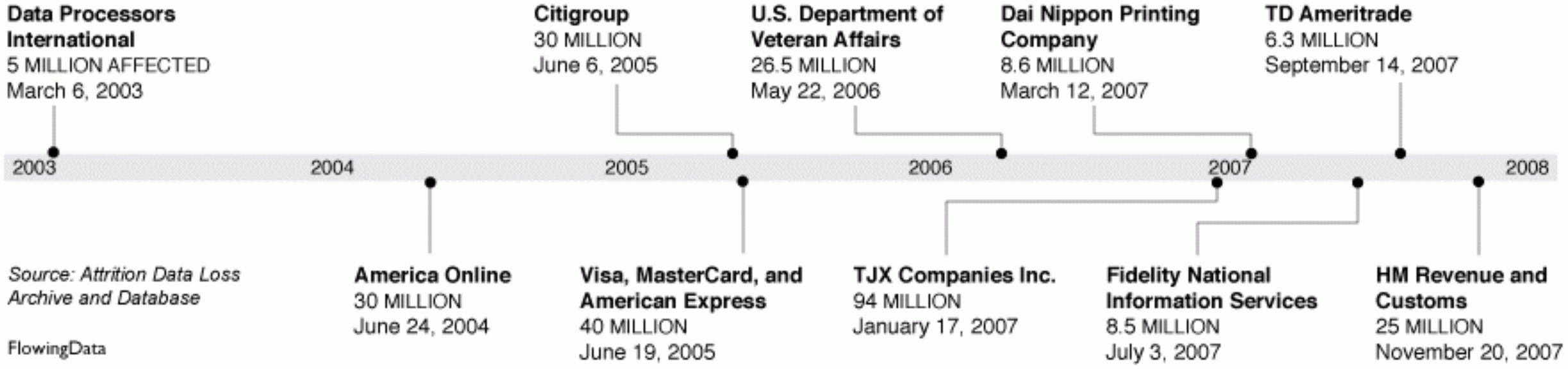

2000
2001
2002
2003
2004
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
9/11 CHANGED EVERYTHING

8. The **Timeline**

- Work best for events, where you're most interested in time of occurrence
- Don't work well for tons of data, but can be combined with others to good effect

10 Largest Data Breaches Since 2000

As more information goes digital, it becomes more important to protect against hackers.



Source: Attrition Data Loss Archive and Database

FlowingData

Honorable mentions

- Sankey Diagram
- Alluvial Diagram
- Animation

Introduction to R graphics with **ggplot2**

https://github.com/emilyfuhrman/datavis_design/blob/master/2017_Fall/Studios/02_Introduction_to_R_Graphics_with_ggplot2.md