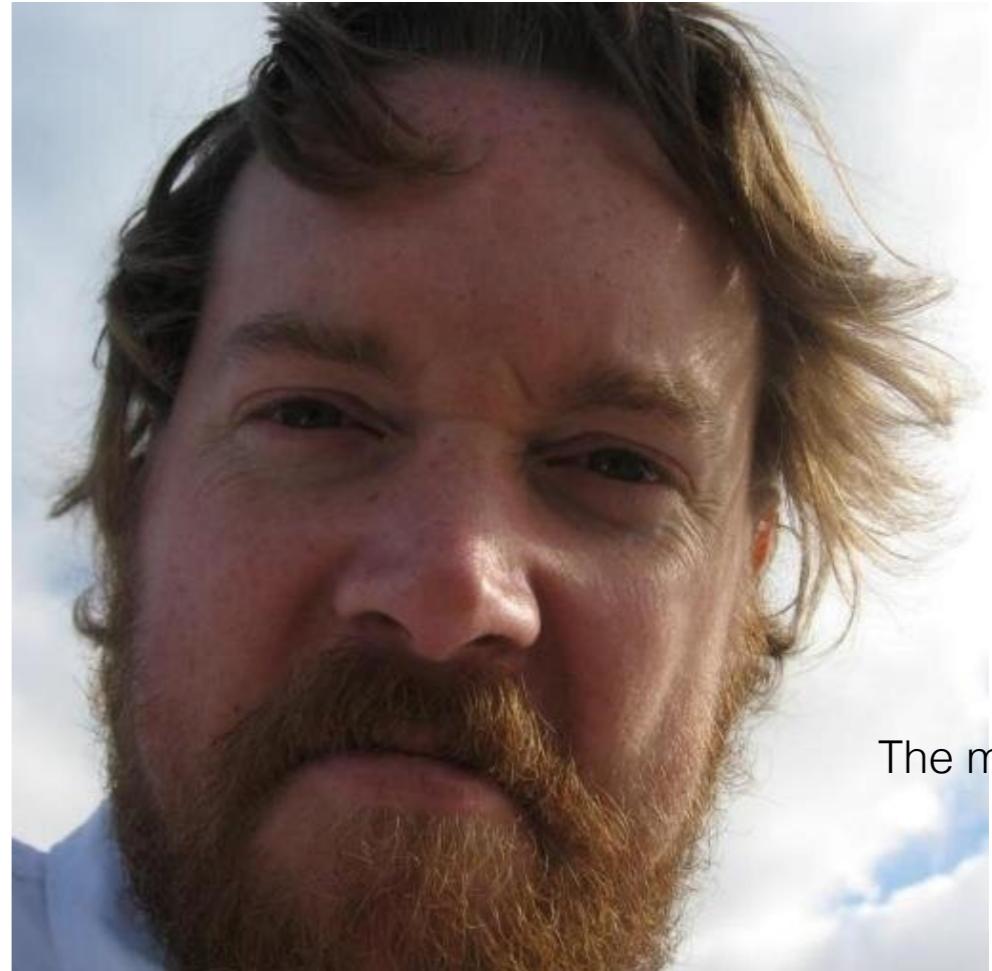


Data Driven Documents

A use case of d3 in a data science workflow

What is this d3 nonsense?

- Created by [Mike Bostock](#).
- Formally PhD Student under Jeffery Heer at University of Washington.
- Formally NY Times, now independent.



The messiah

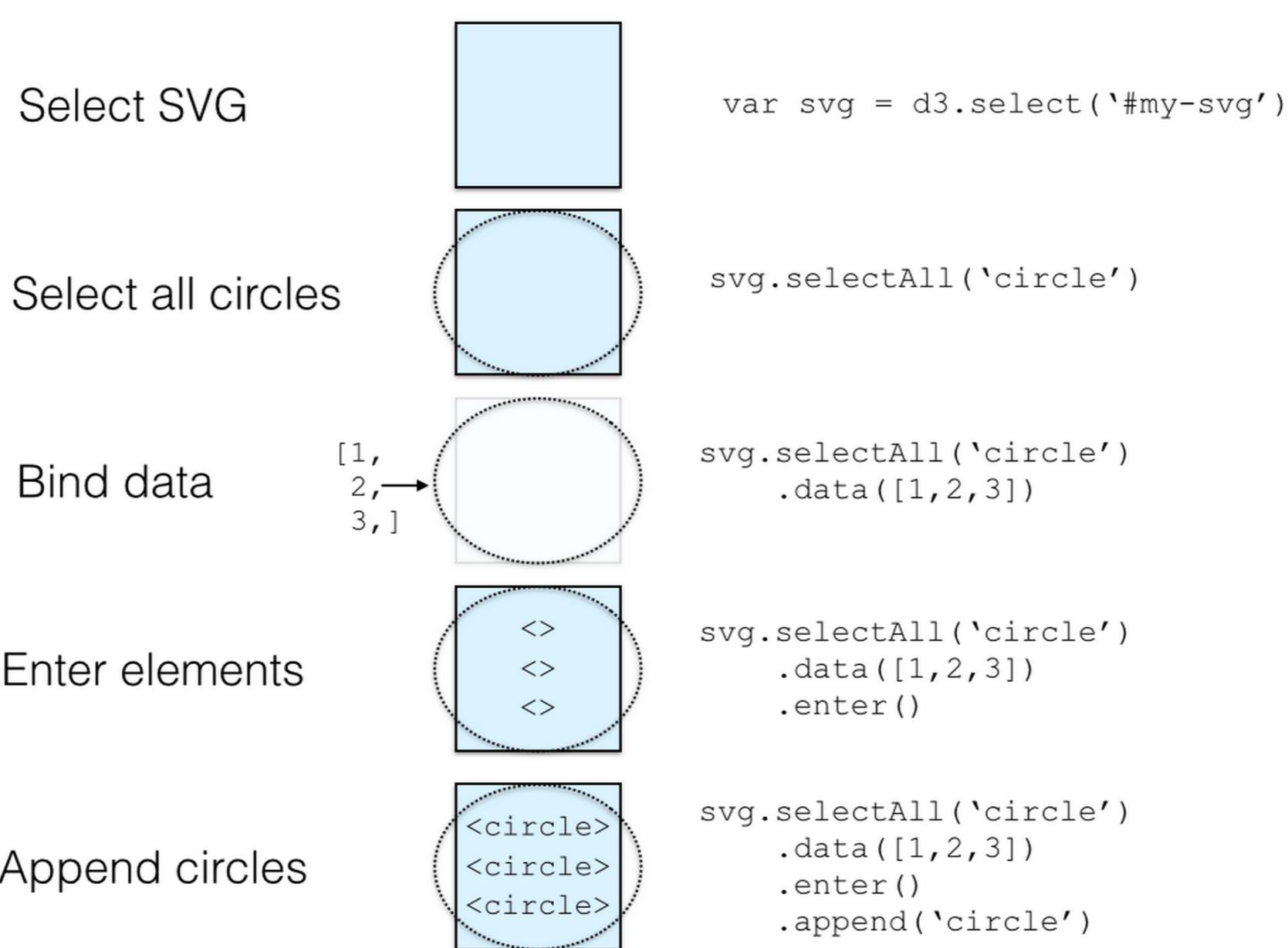
What does it do?

```
d3.select(this).attr("what") = "magic"
```

- It is (one of) a (million) javascript framework(s).
- At its core it is a method of binding Data and Driving the Dom (elements of a webpage).
- Also has fantastic interpolation abilities for transitions and drawing shapes.

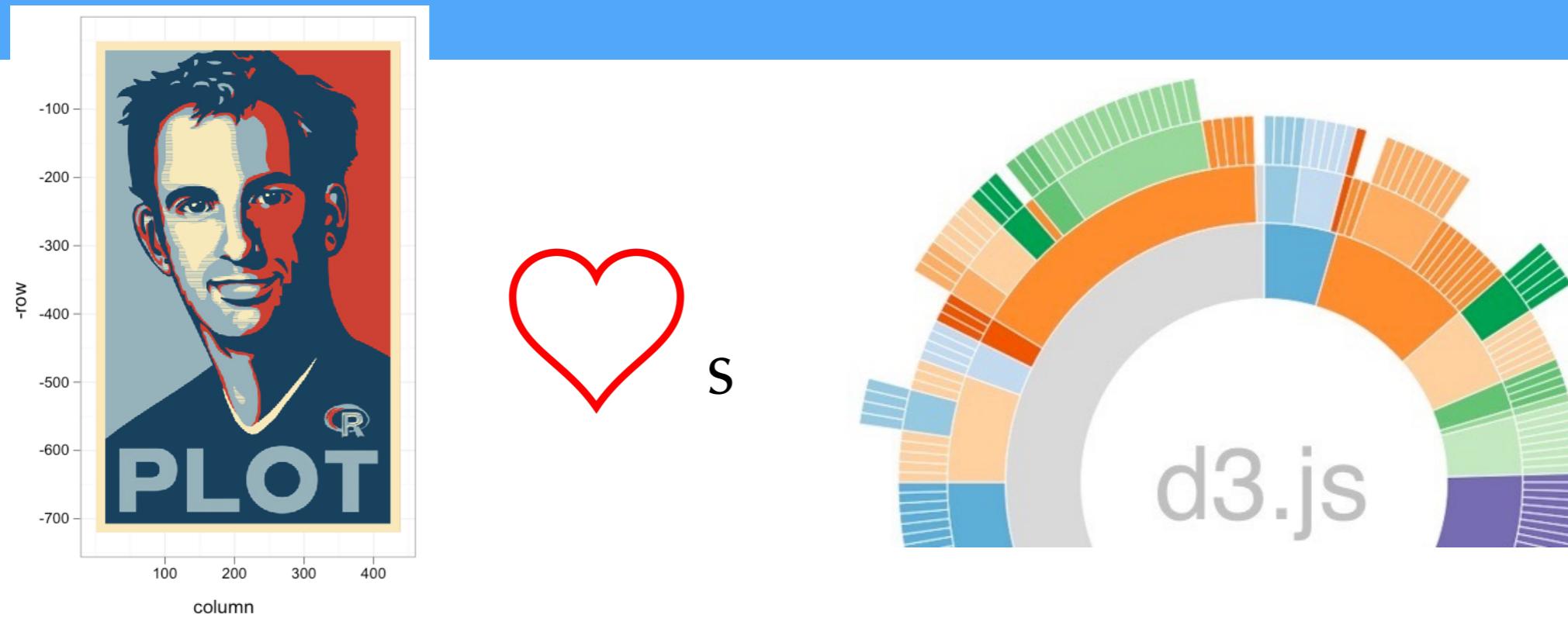
A quick overview

- How d3 operates (at a high level)*.



*Shamelessly stolen from Michael Freeman's "[Scrollytelling](#)"

Similarities to R



- Ggplot2 is in many ways a spiritual brother to d3.
- Follows the grammar of graphics:
(Data → Geometry → Mapping)
- I think they compliment each other rather than compete. Ggplot for exploration static, d3 for explanatory/ dynamic.

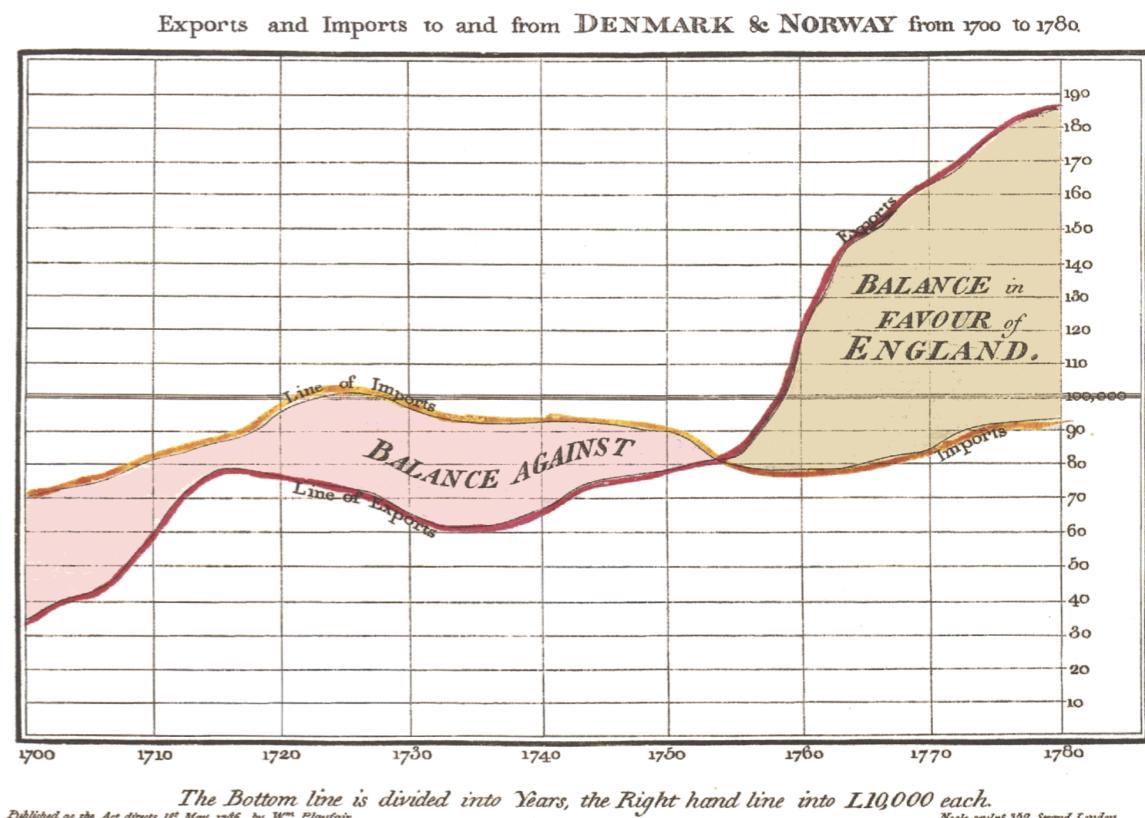
Why is data visualization important?

- (I think) data visualization is a form of empathy.
- It enables the dissemination of information/knowledge across disciplinary boundaries.
- A language of science.

	2003	2002	2001	2000	1999	1998
USA	10,882	10,383	10,020	9,762	9,213	8,720
EU	10,970	9,040	8,303	8,234	8,901	8,889
UK	1,765	1,564	1,430	1,438	1,460	1,423
China	1,575	1,434	1,345	1,252	1,158	1,148
India	599	510	479	457	447	414

GDP, based on exchange rates, over time. Values in billion USDs.

VS.



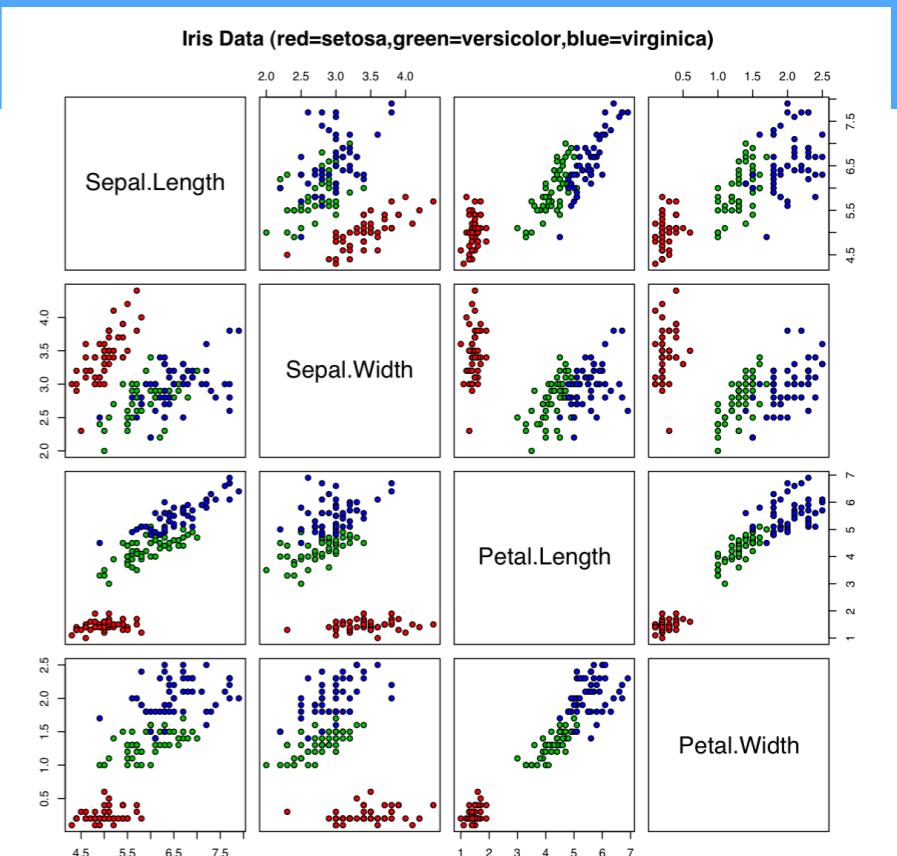
Sources: [table](#), [Playfair](#)

The two types of Data Viz

“Lies, damn lies, and statistics...”

- Exploratory (ggplot/ matplotlib ...)

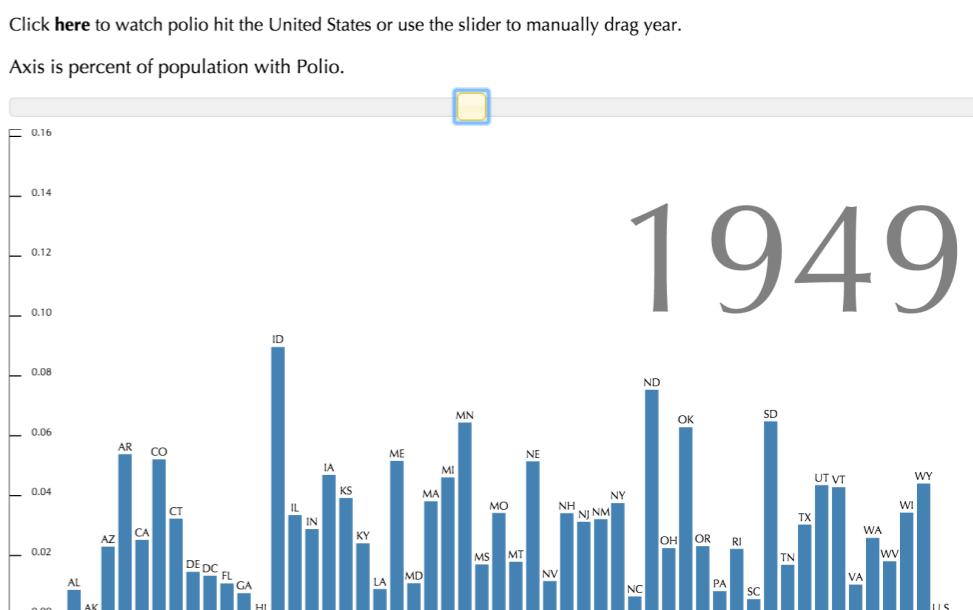
- What is going on in the data?
- Must be careful to do without bias.



- Explanatory (d3, ggviz, tableau?,...)

Visualizing the impact of Polio:

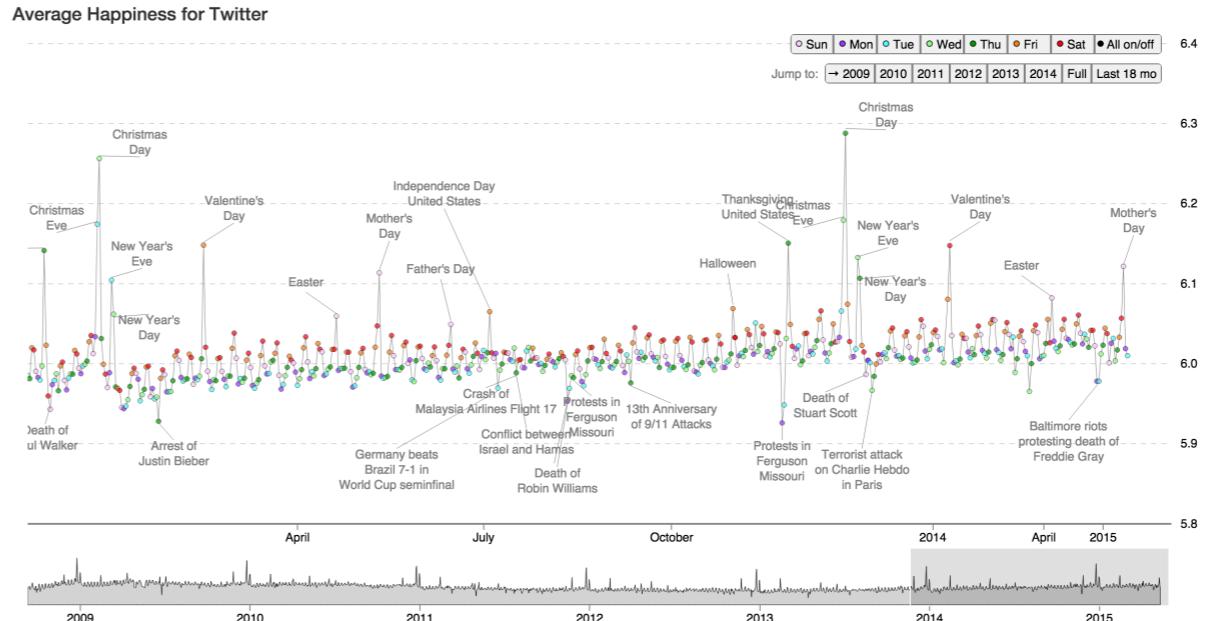
- Let others know what is going on in the data.
- Tell a story. (**Danger!** you are biased.)



Source: [iris data](#)

Now the cool stuff:

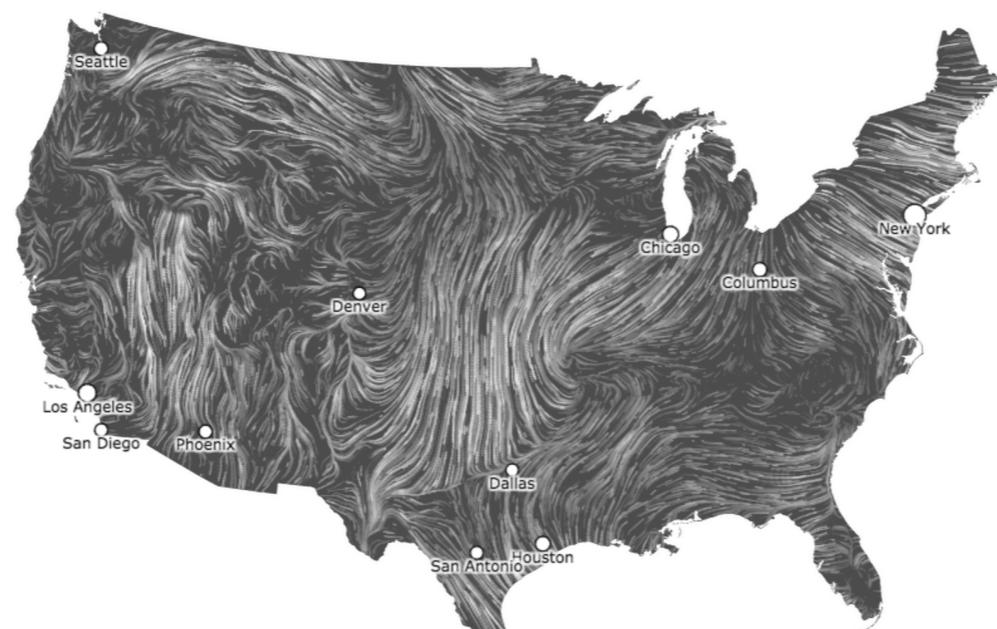
- UVM's Hedonometer



Gestalt Principles for Data Visualization

Similarity, Proximity & Enclosure

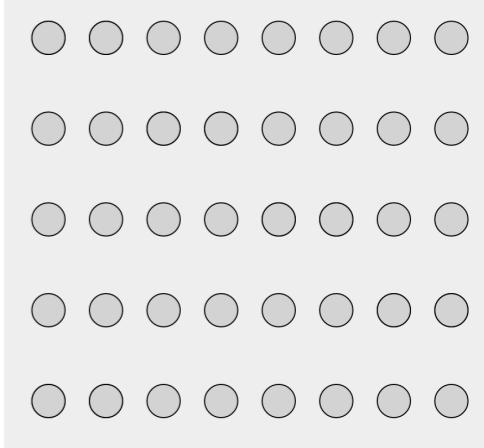
- Elijah Meek's Gestalt



Introduction

At a recent talk I challenged the audience to define several gestalt principles based solely on representative figures. This "academic" approach to data visualization seems in opposition to a "pragmatic" approach that focuses on best practices and prior art demonstrated in the growing library of data visualization books and 2-day seminars.

But let me suggest that gestalt is very much a pragmatic aspect of creating data visualization, in fact a necessary aspect if you plan to do more than simple bar and line charts (and perhaps even for those simple charts). This exploration of three of the most simple gestalt principles focuses on how they operate and how they might act in tandem with and in opposition to each other. I also include some gestures toward how the gestalt may already be influencing what we think of not as cognitive qualities but as design and style in data visualization.

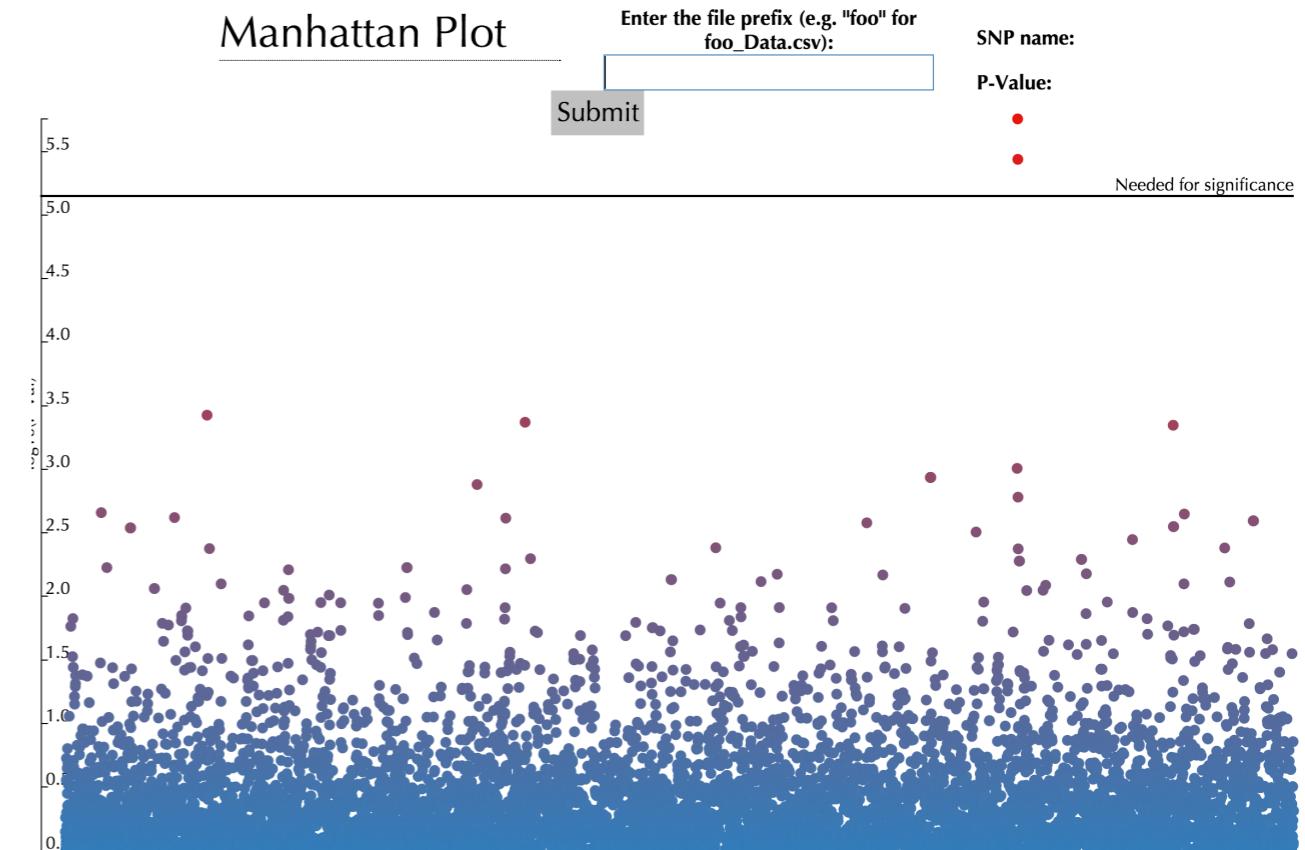


- Wind Map

nickstrayer.me/#projects

- Exploratory

- Help investigate genetic links to traits
- Fed by a simple r-script



- Explanatory

- Data was cleaned (in python)
- Hosted to help show trends

California power plants

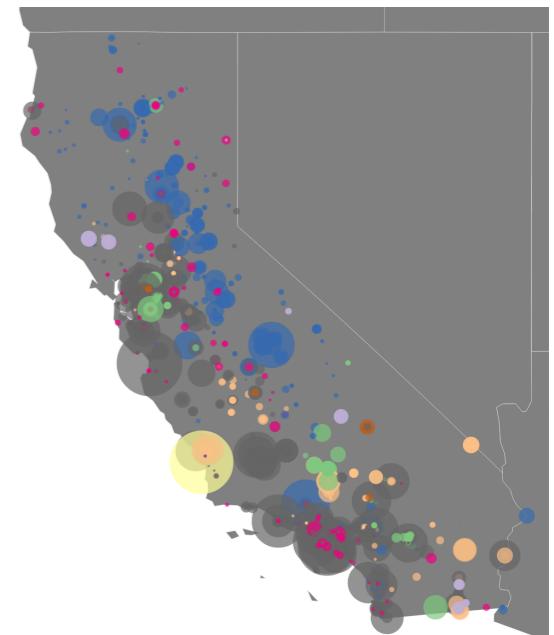


Plant Name:

Type:

Capacity:

Address:



Want to Learn it?

- Interactive Data Visualization for the Web
 - Scott Murray, USF.
 - What I started out with.
 - Great introduction from the very basics up to a semi-decent proficiency.
 - Free and has interactive code chunks to play with what you just read about.
- Easy as Pi
 - Scott Murray again.
 - Cool interactive story-based learning experience.
 - Scrolls, which is awesome.
- d3js.org
 - Hundreds upon hundreds of examples, most with code to go along.
 - Probably better for someone who knows javascript a bit, or has read the aforementioned book.

Contact me

- Feel free to send me any questions you may have about d3 or data viz. I will do my very best to answer them.
- @NicholasStrayer
- nick.strayer@gmail.com
- github.com/nstrayer
- This talk is available on github [here](#).