Basic Spatial Arrangements

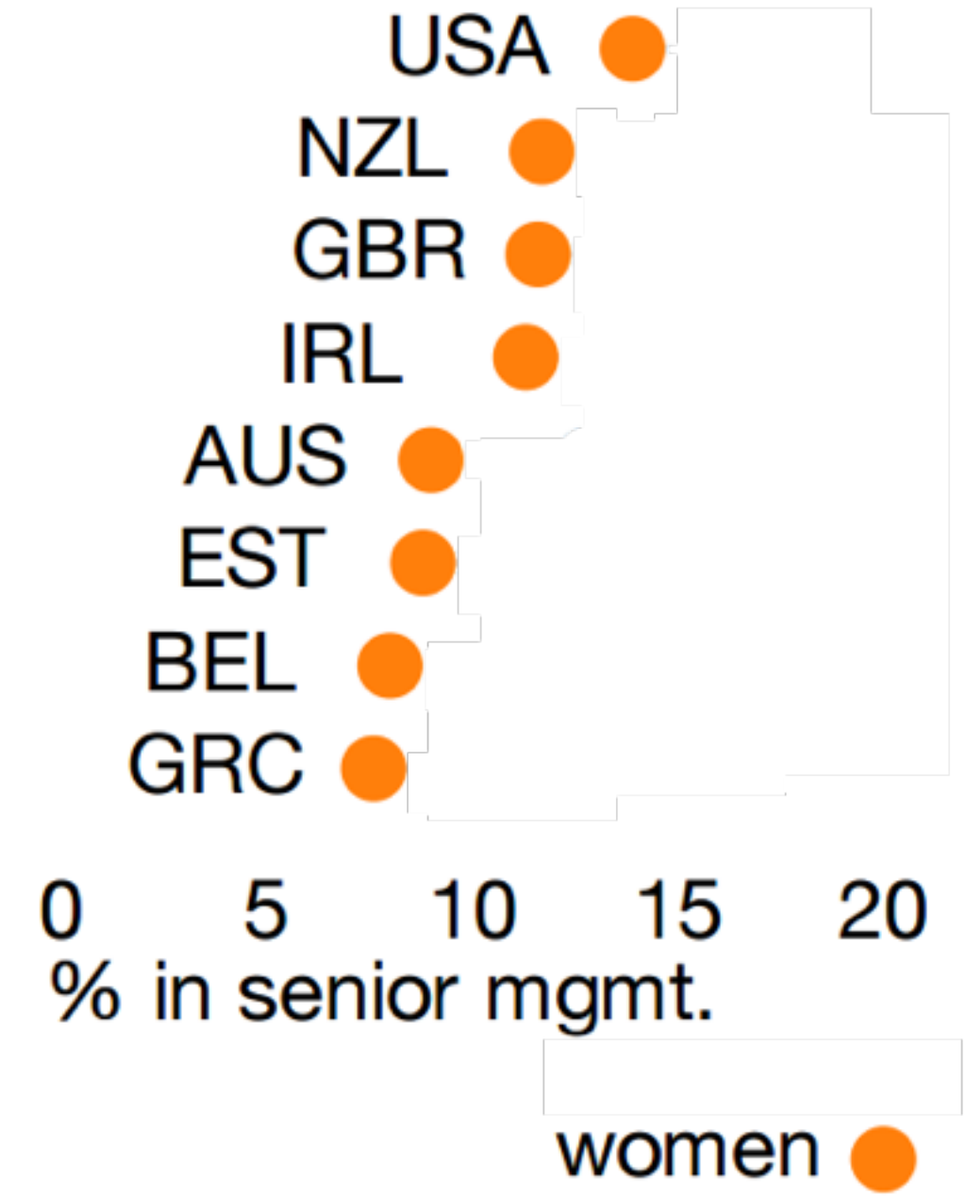
CSC544
Techniques, Lecture 3

Suggested reading:
Munzner's Visualization Analysis and Design, **Chapter 7**

Few dimensions

Dot Plots

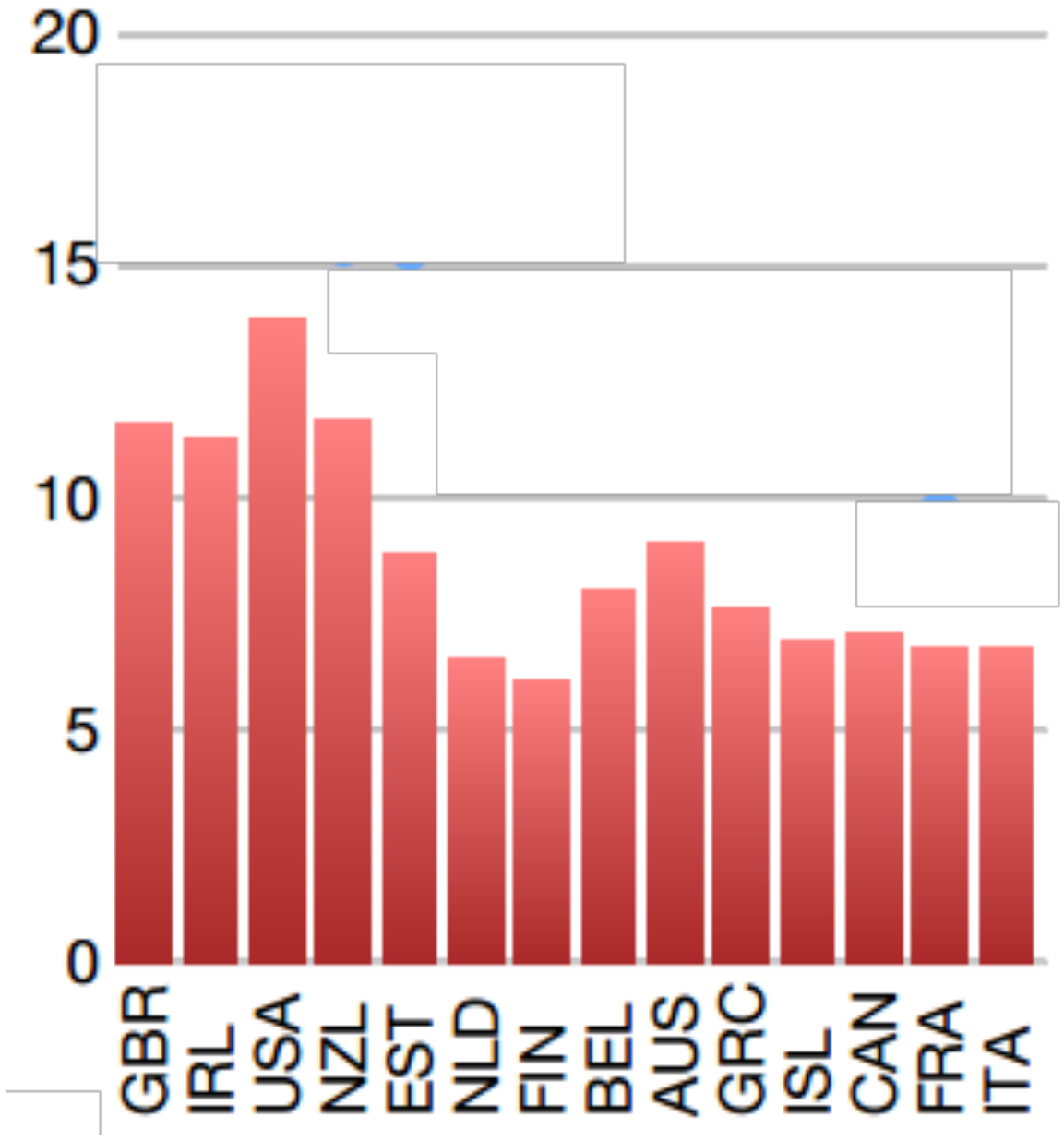
- One categorical attribute (usually a “key”)
- One quantitative attribute (usually a “value”)



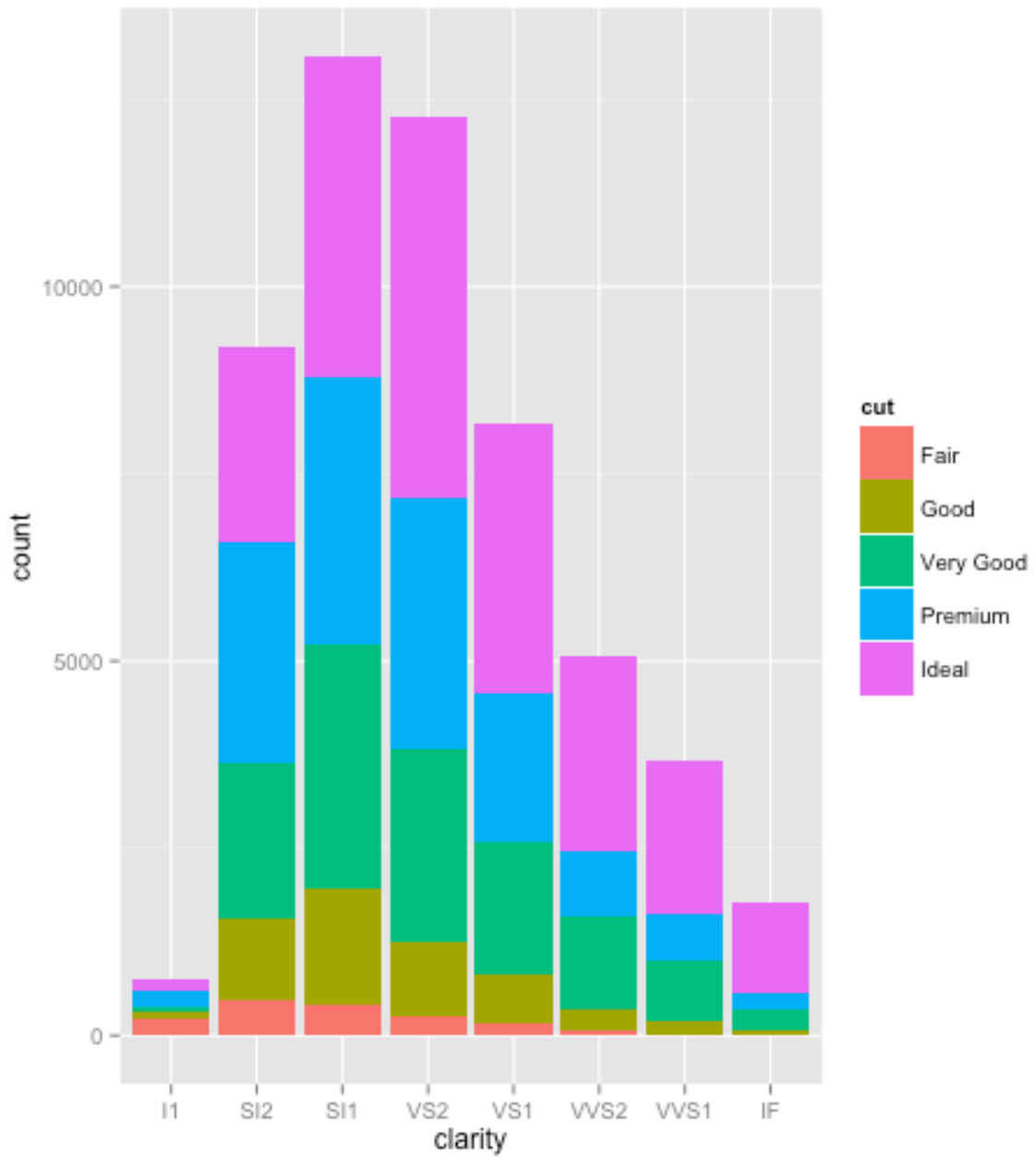
Redesign Dot Plots

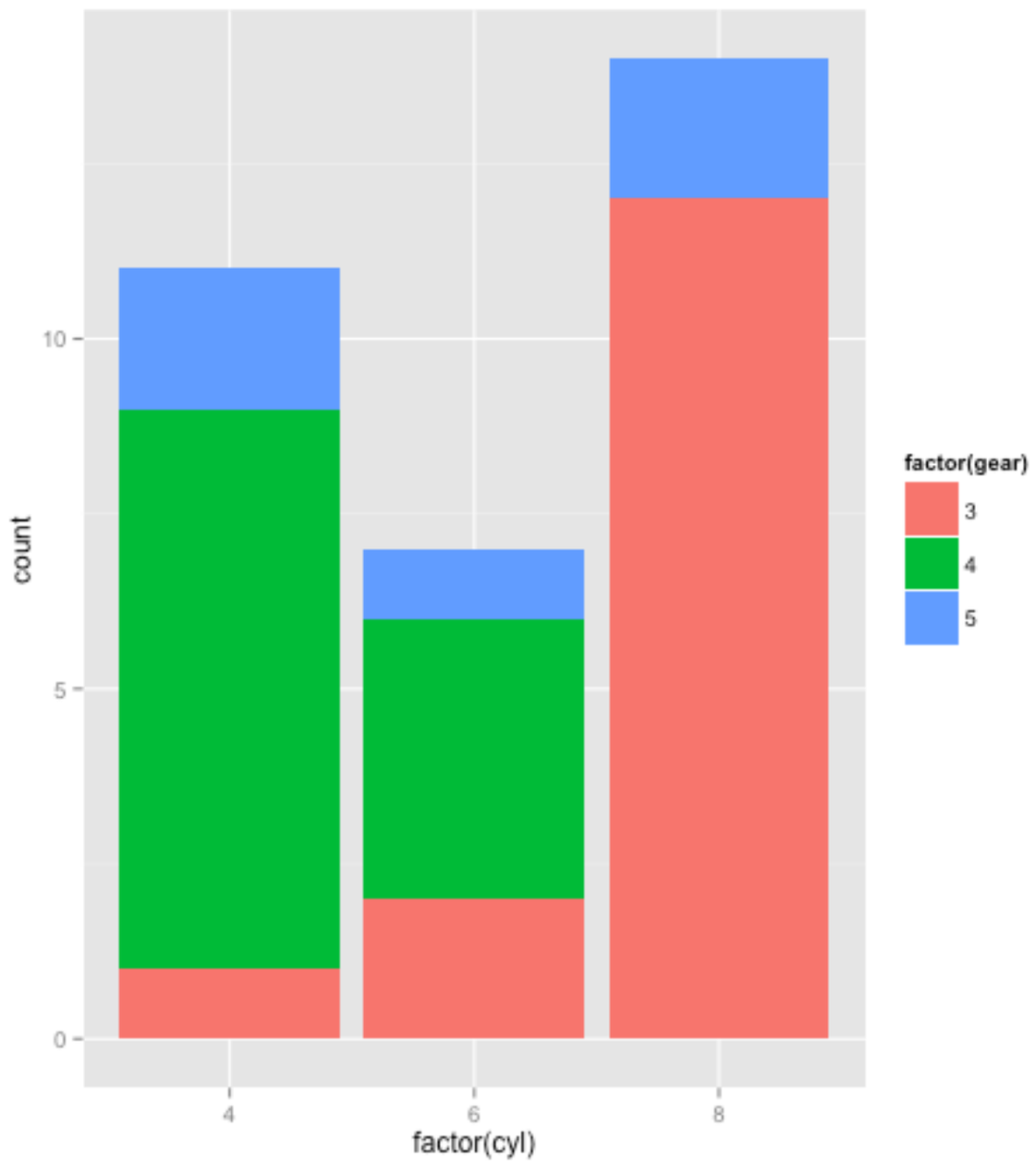
Bar Charts

- One categorical attribute (usually a “key”)
- One quantitative attribute (usually a “value”)
- But usually not the best option, given Cleveland and McGill

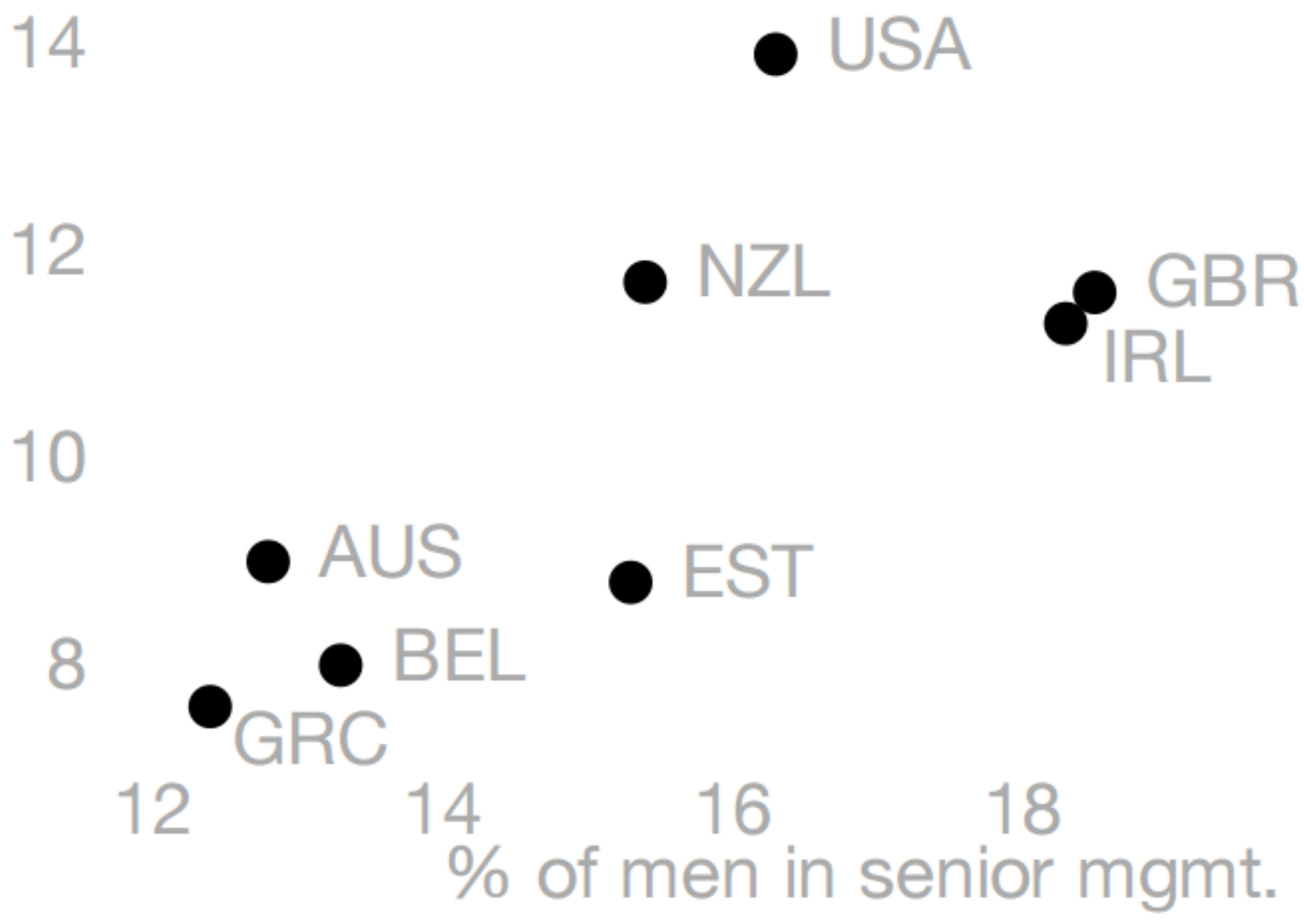


Stacked Bar Charts

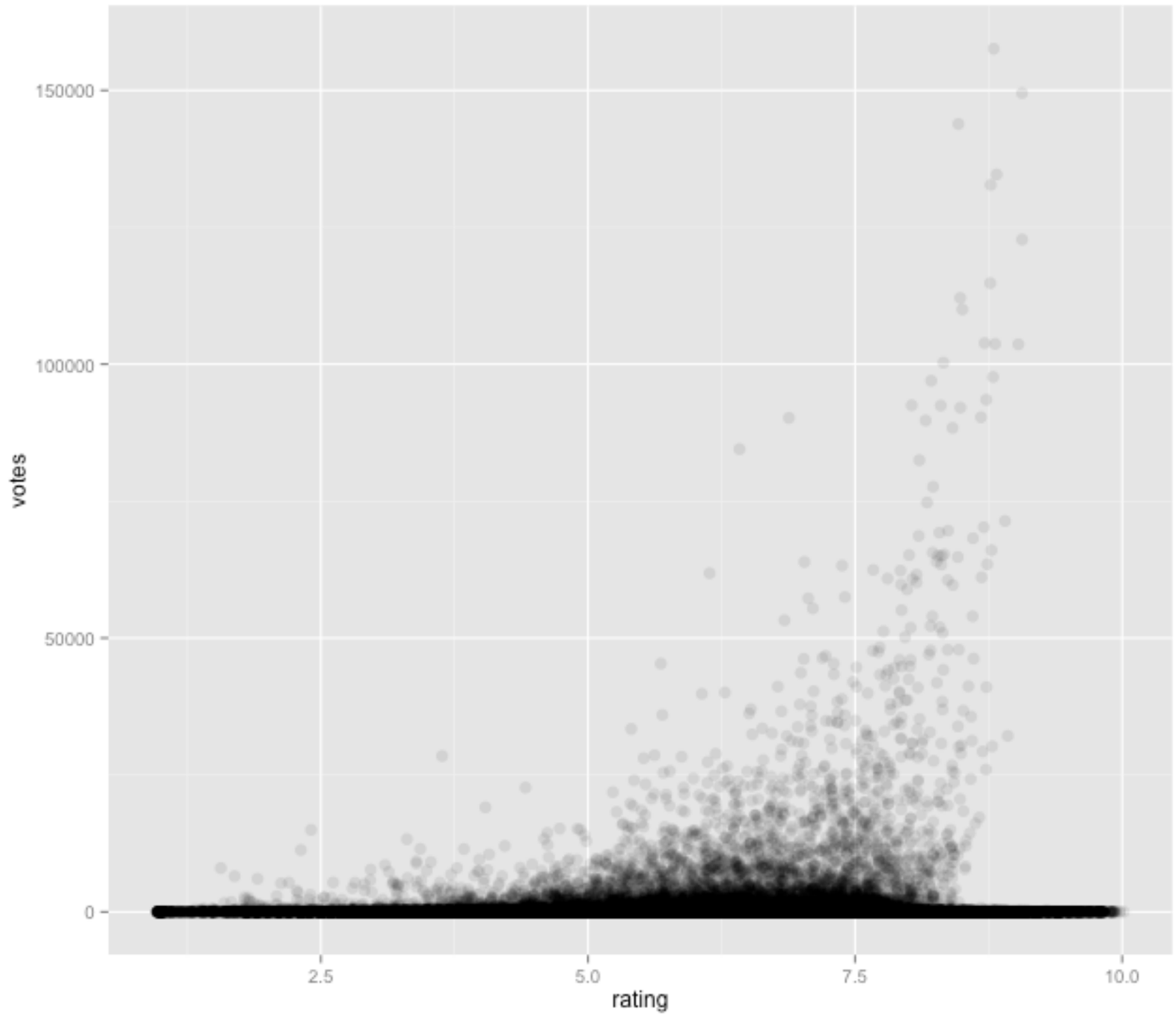




Scatterplots

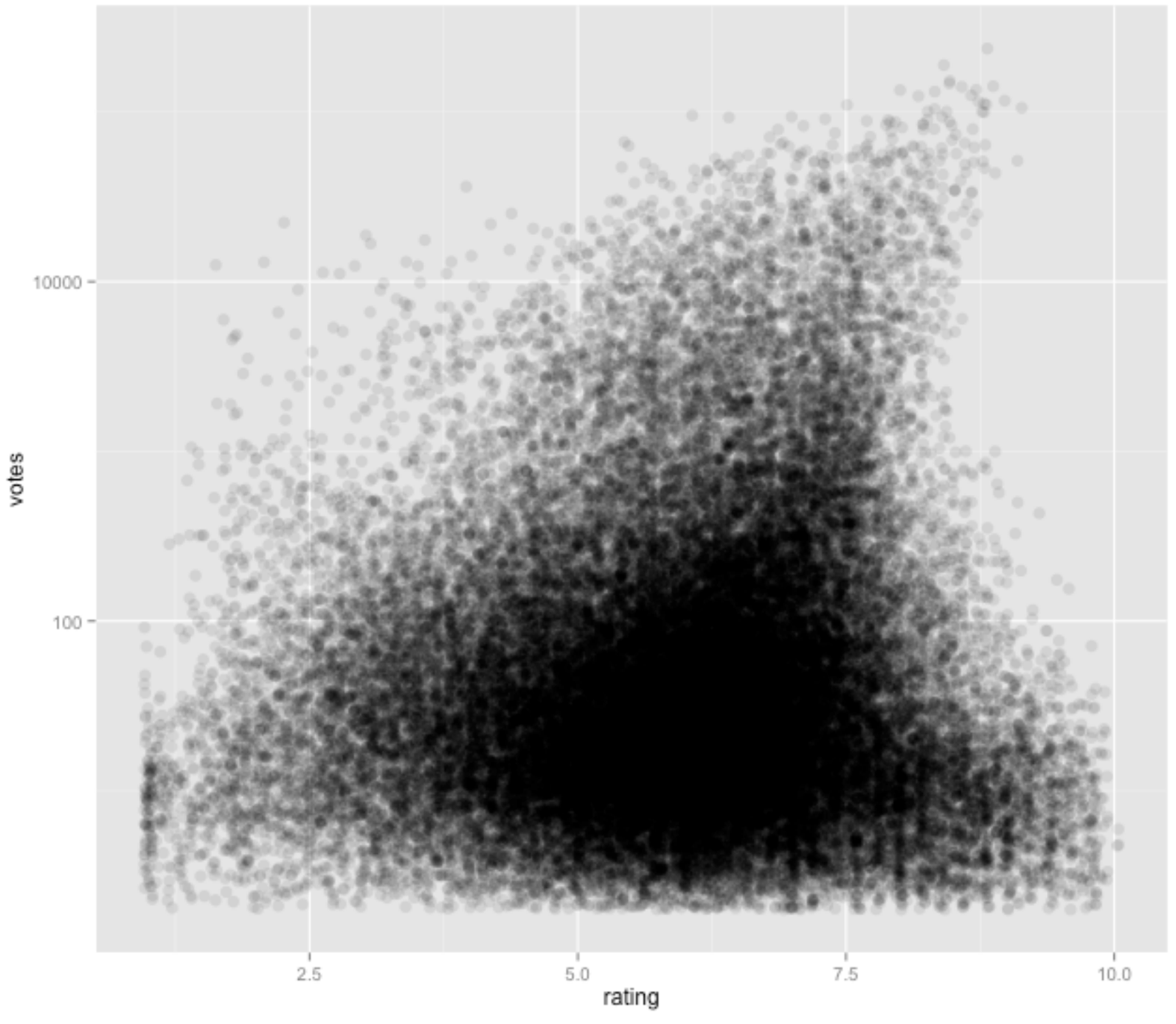


Transformations



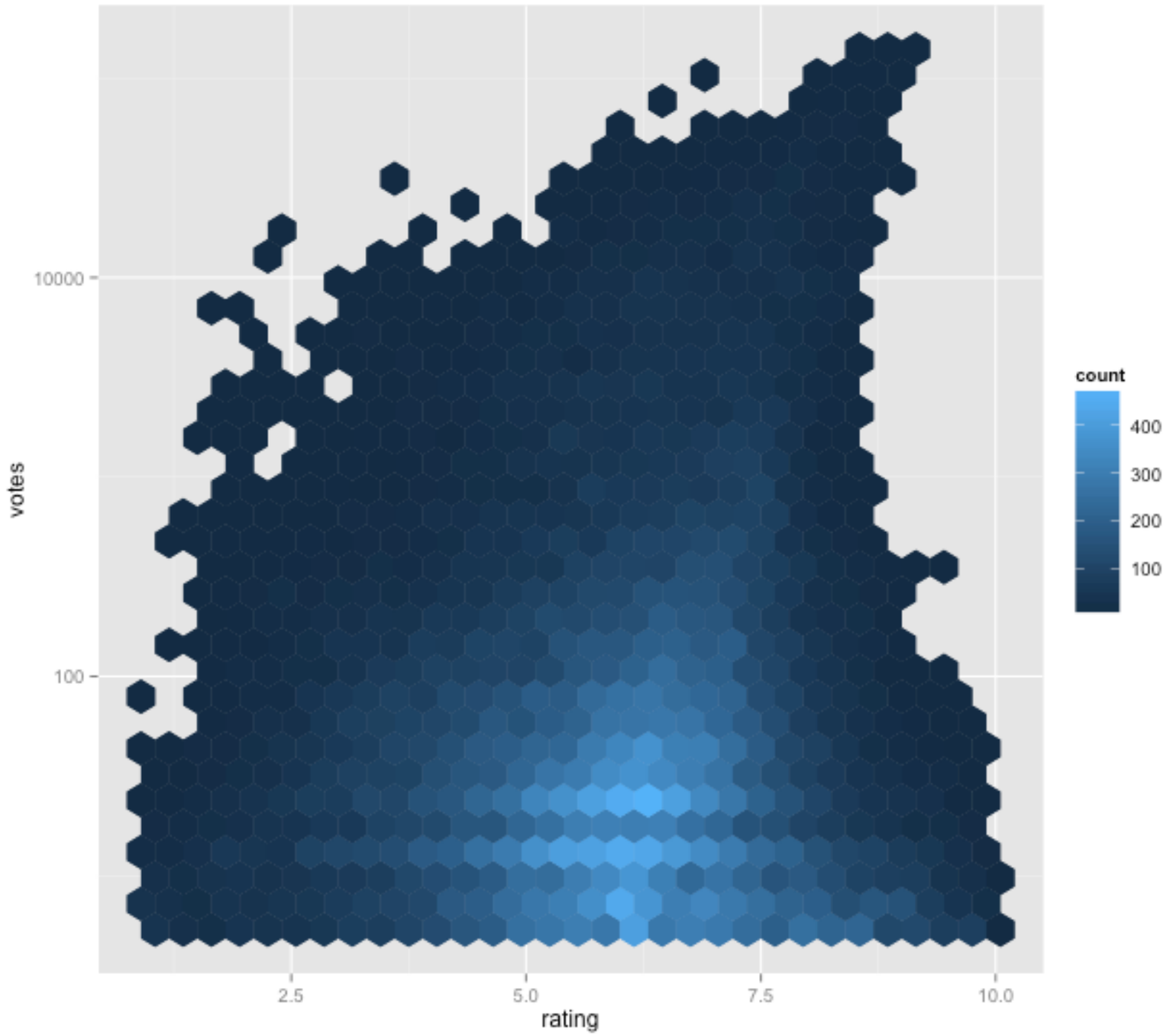
IMDB
Avg Rating
vs. #votes

Transformations



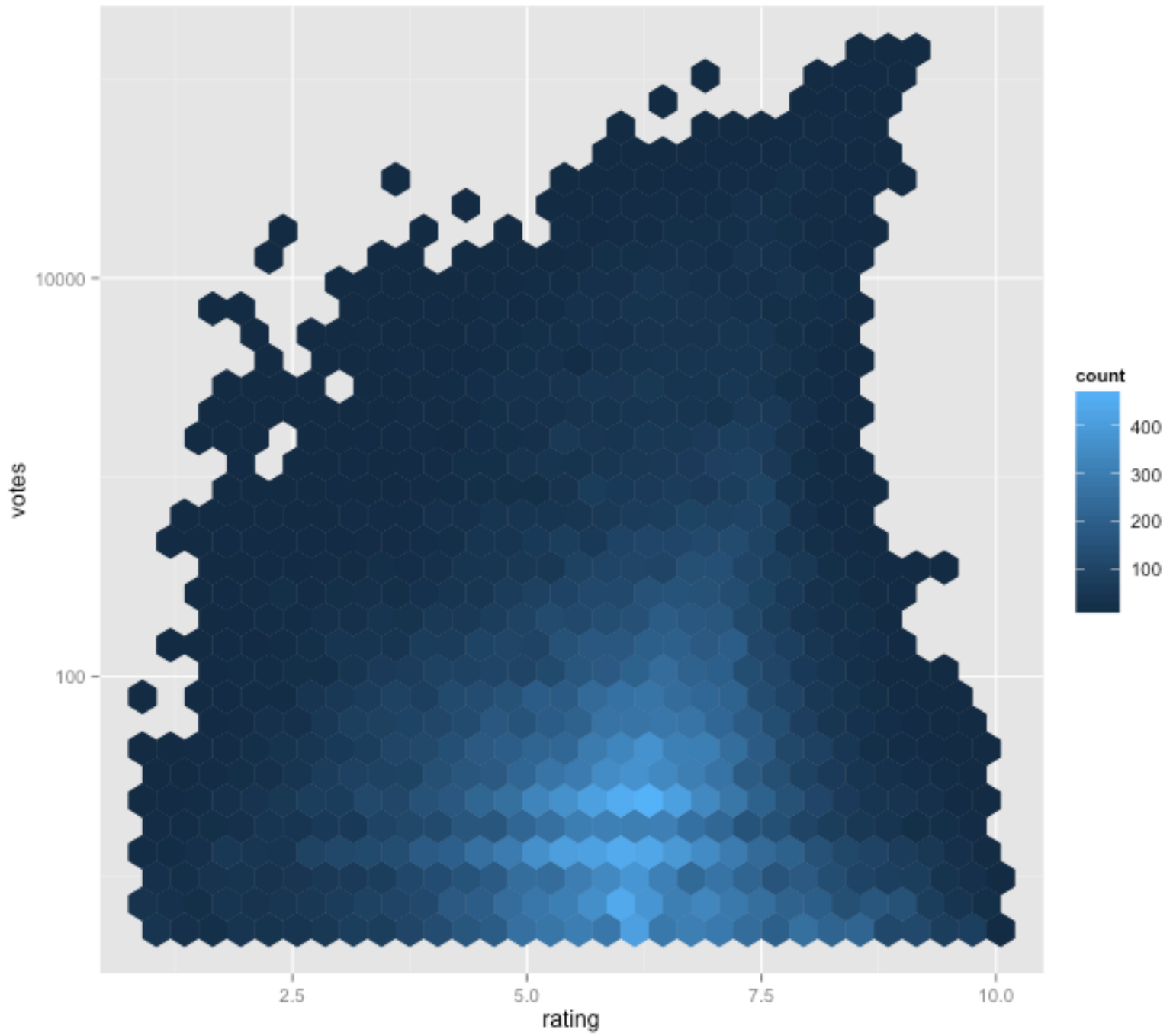
IMDB
Avg Rating
vs. #votes

Transformations



IMDB
Avg Rating
vs. #votes

Aside: how not to (sort) by average rating



IMDB
Avg Rating
vs. #votes

Aside: how not to (sort) by average rating

PROBLEM: You are a web programmer. You have users. Your users rate stuff on your site. You want to put the highest-rated stuff at the top and lowest-rated at the bottom. You need some sort of "score" to sort by.

WRONG SOLUTION #1: Score = (Positive ratings) – (Negative ratings)

WRONG SOLUTION #2: Score = Average rating = (Positive ratings) / (Total ratings)

https://www.evanmiller.org/how-not-to-sort-by-average-rating.html

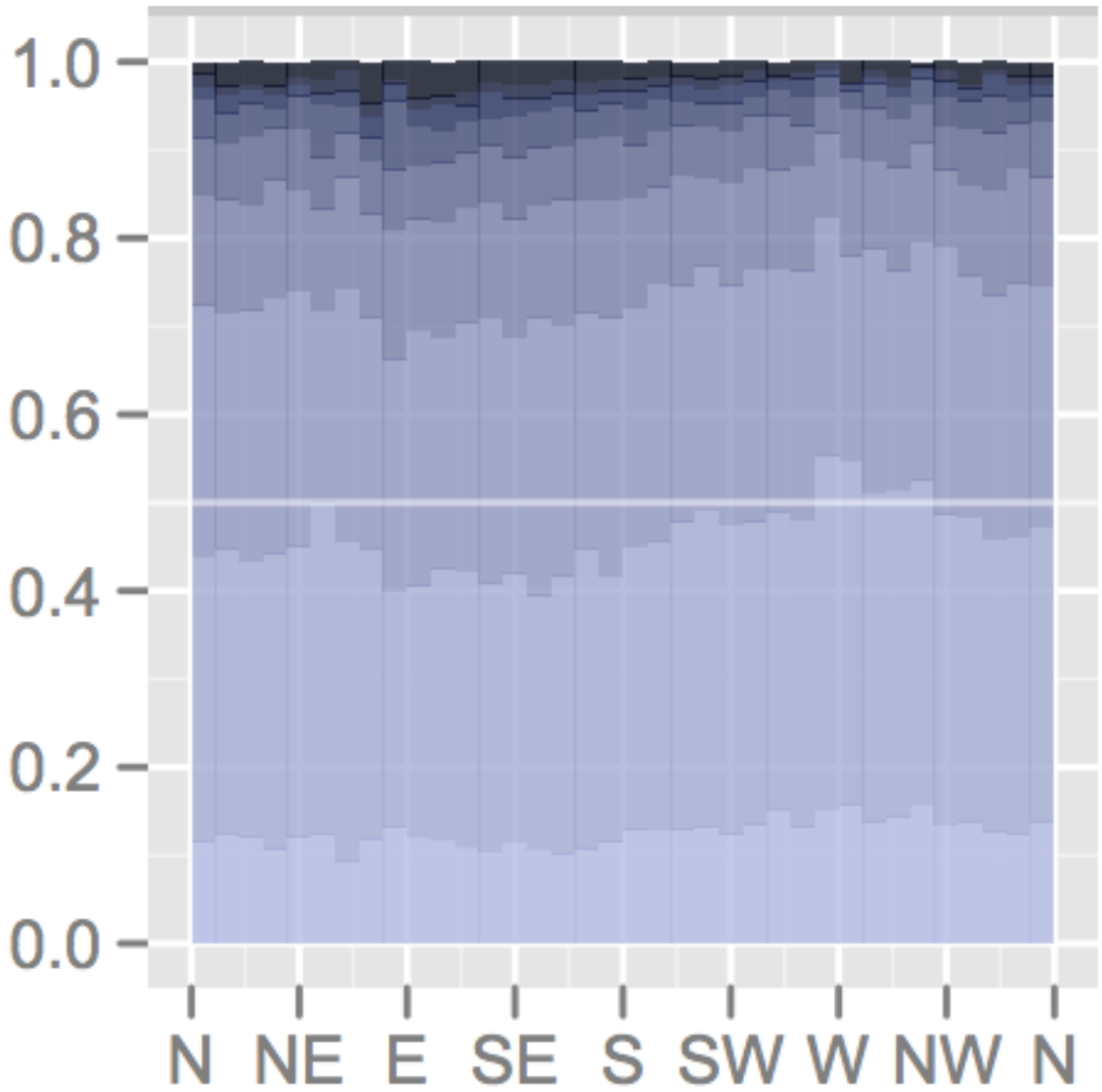
CORRECT SOLUTION: Score = Lower bound of Wilson score confidence interval for a Bernoulli parameter

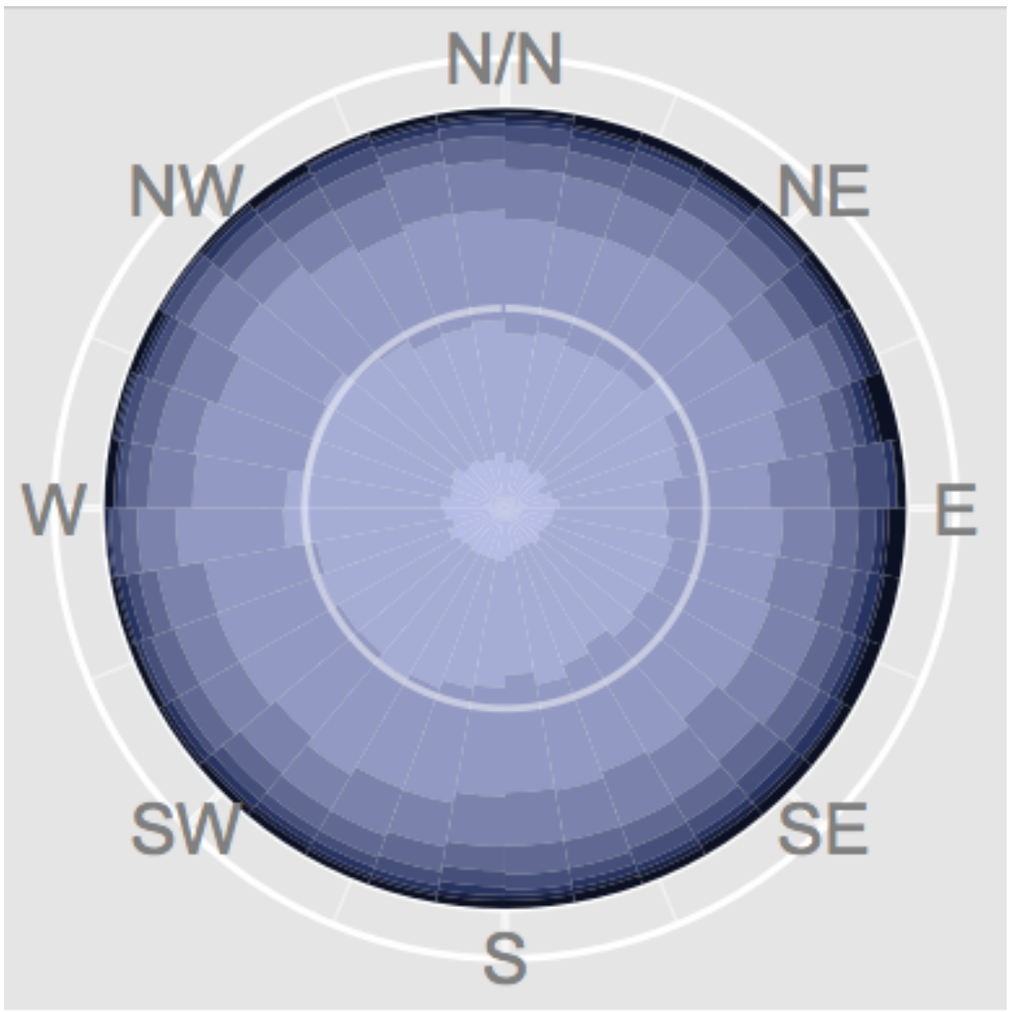
Say what: We need to balance the proportion of positive ratings with the uncertainty of a small number of observations. Fortunately, the math for this was worked out in 1927 by Edwin B. Wilson. What we want to ask is: *Given the ratings I have, there is a 95% chance that the "real" fraction of positive ratings is at least what?* Wilson gives the answer. Considering only positive and negative ratings (i.e. not a 5-star scale), the lower bound on the proportion of positive ratings is given by:

$$\left(\hat{p} + \frac{z^2_{\alpha/2}}{2n} \pm z_{\alpha/2}\sqrt{[\hat{p}(1-\hat{p}) + z^2_{\alpha/2}/4n]/n}\right)/(1 + z^2_{\alpha/2}/n).$$

https://www.evanmiller.org/how-not-to-sort-by-average-rating.html

Transformations

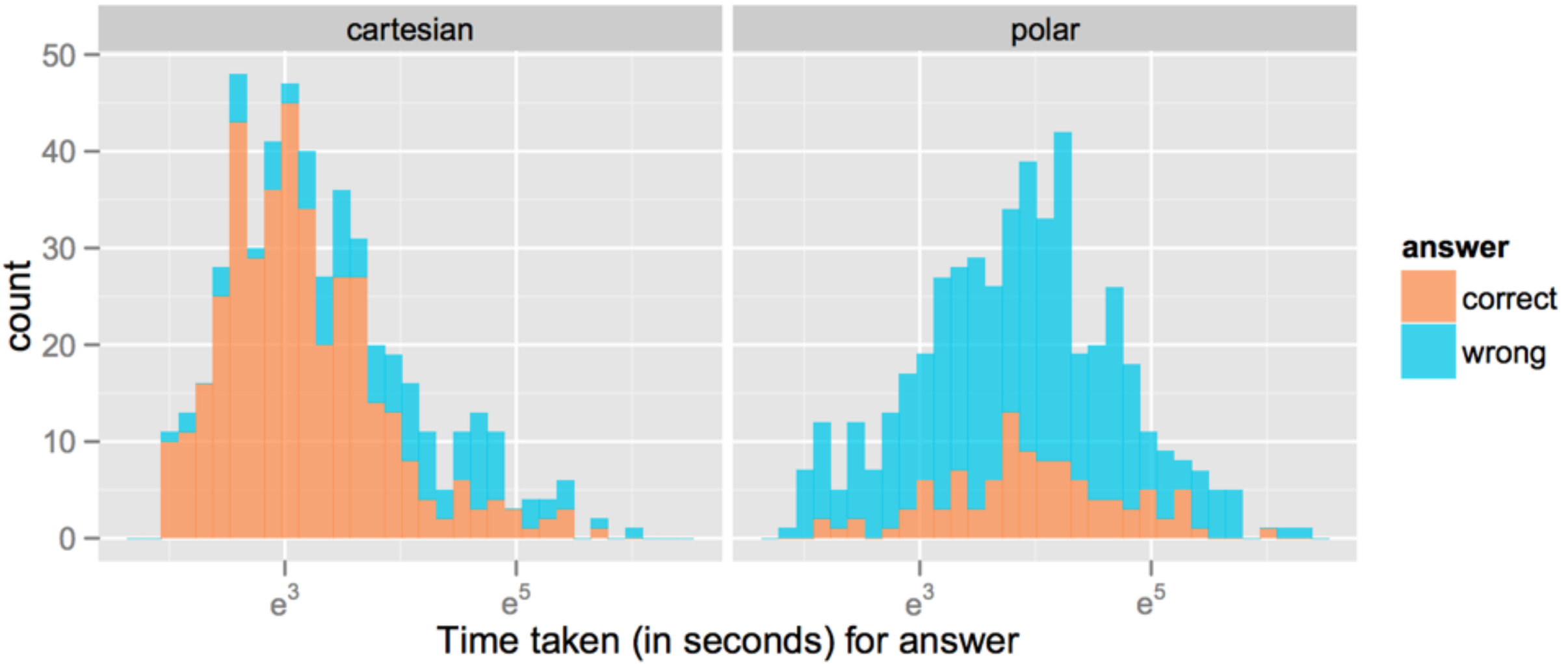




Hoffman et al., TVCG 2012

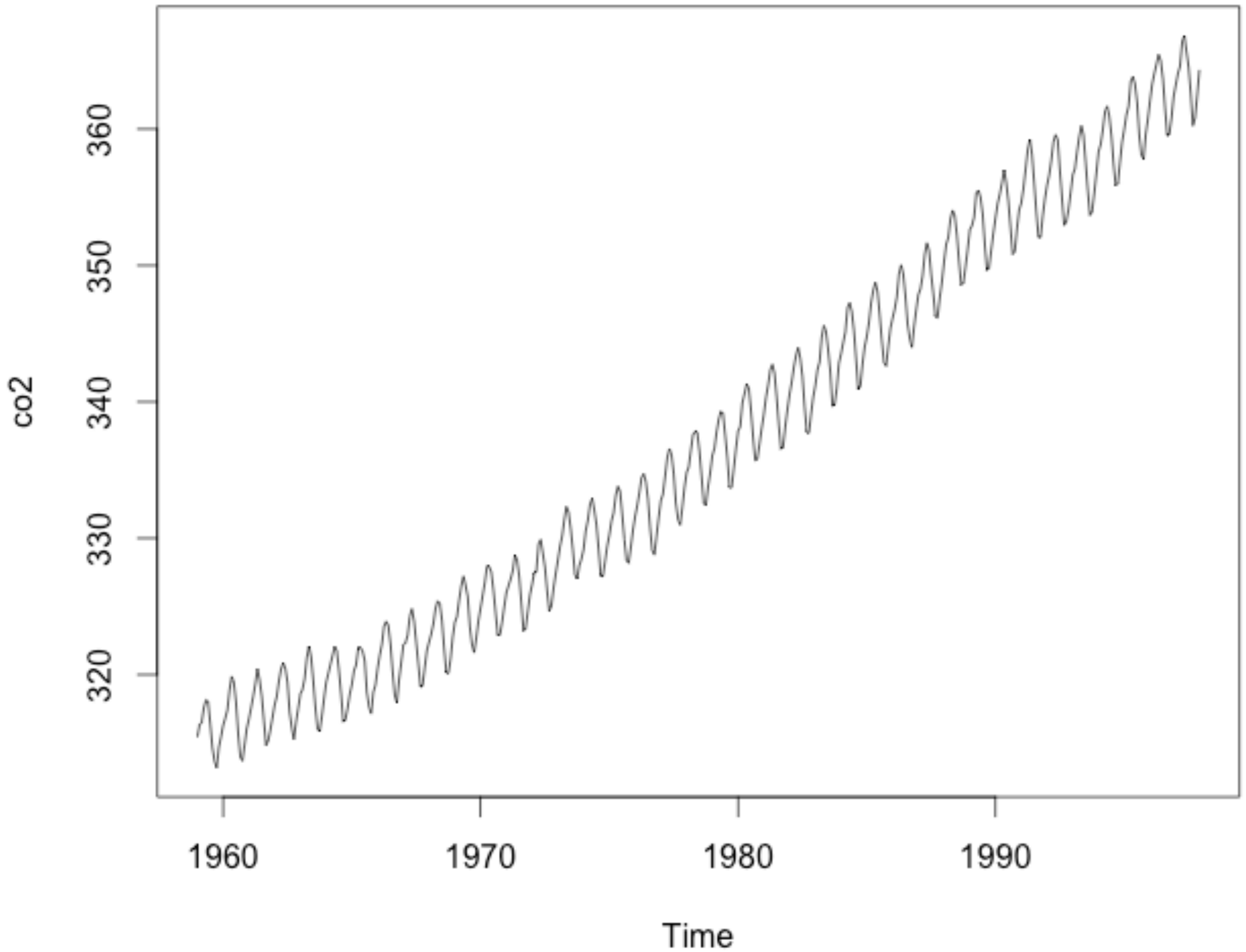
BUT

Transformations

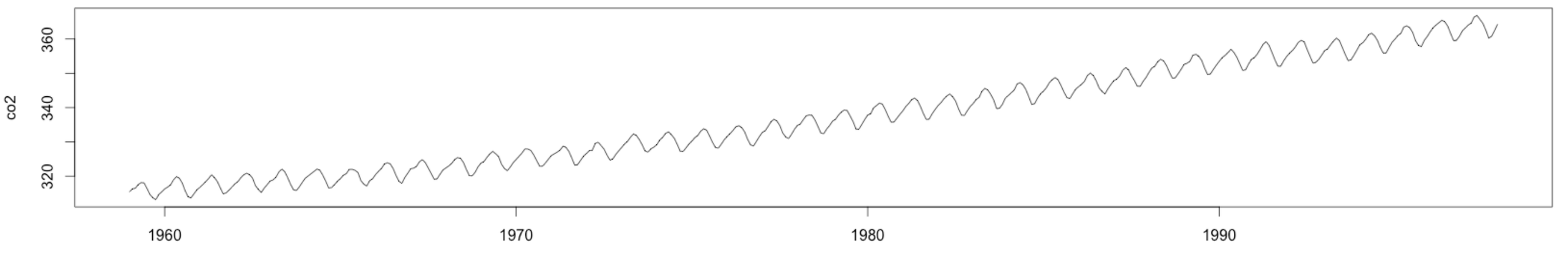


Hoffman et al., TVCG 2012

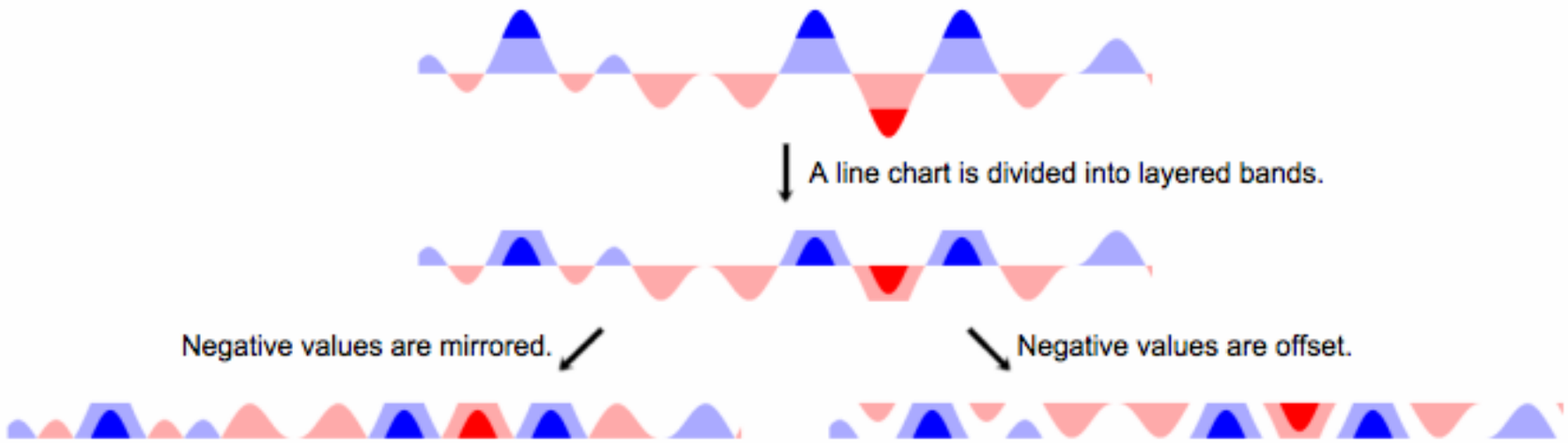
Line Charts

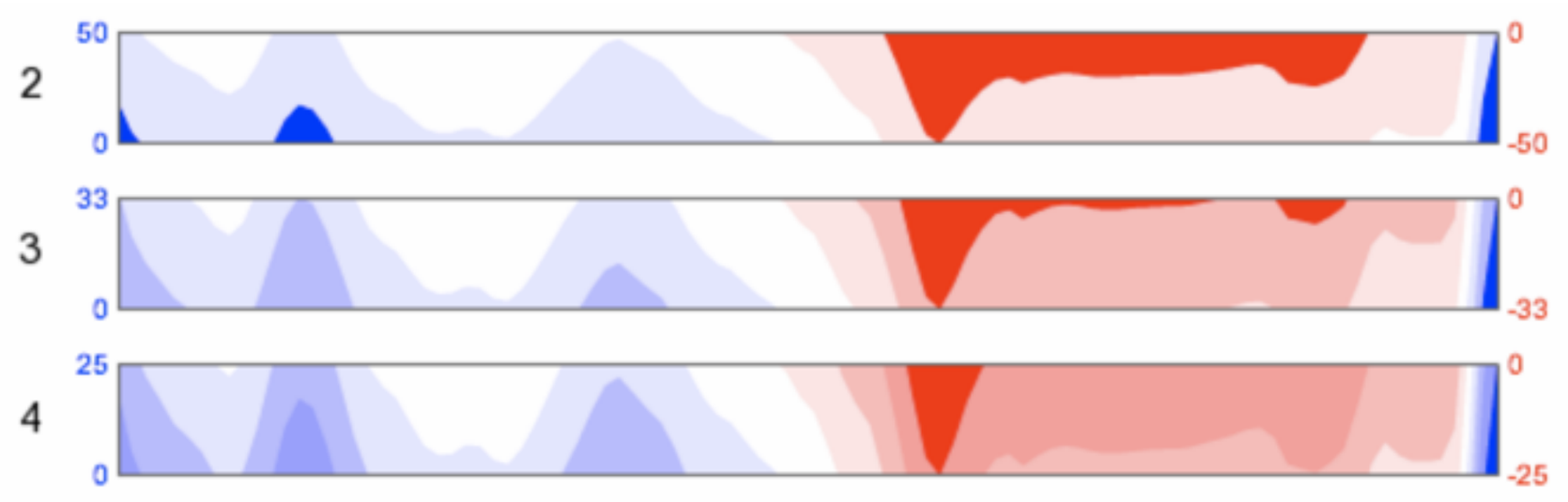


Bank to 45 degrees



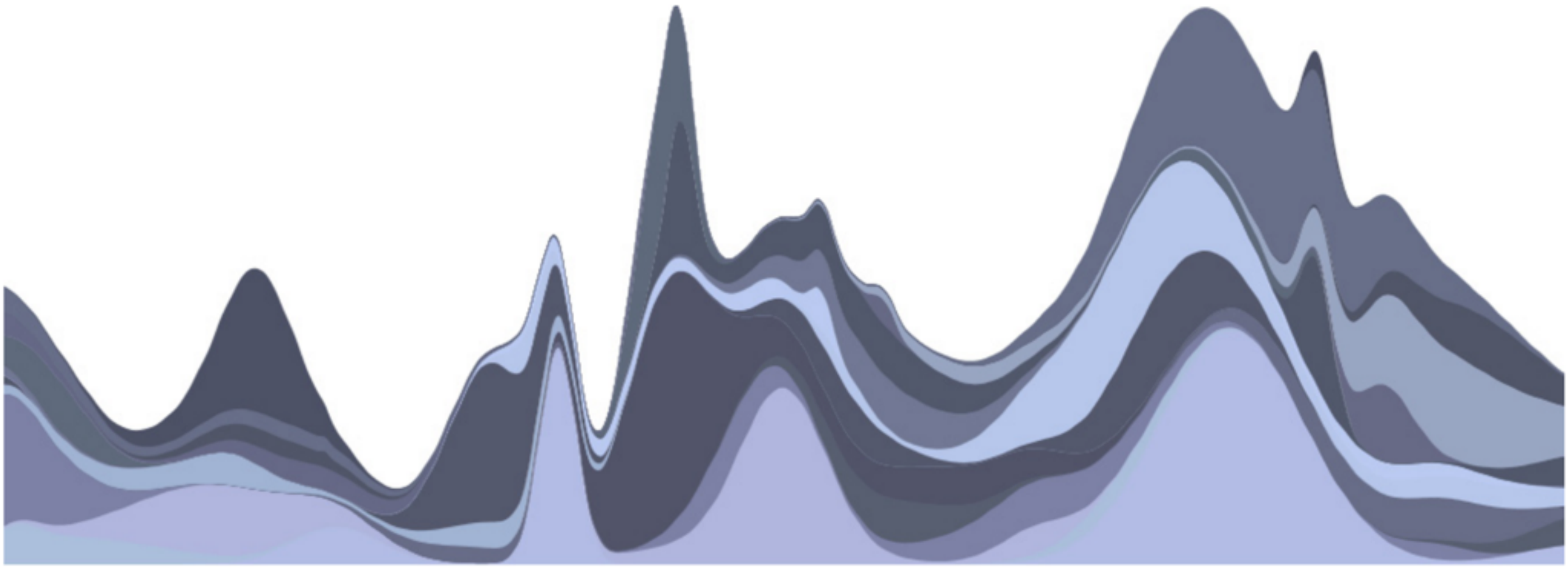
Horizon Charts





http://bl.ocks.org/mbostock/1483226

Streamgraphs



http://www.leebyron.com/else/streamgraph/

Streamgraphs

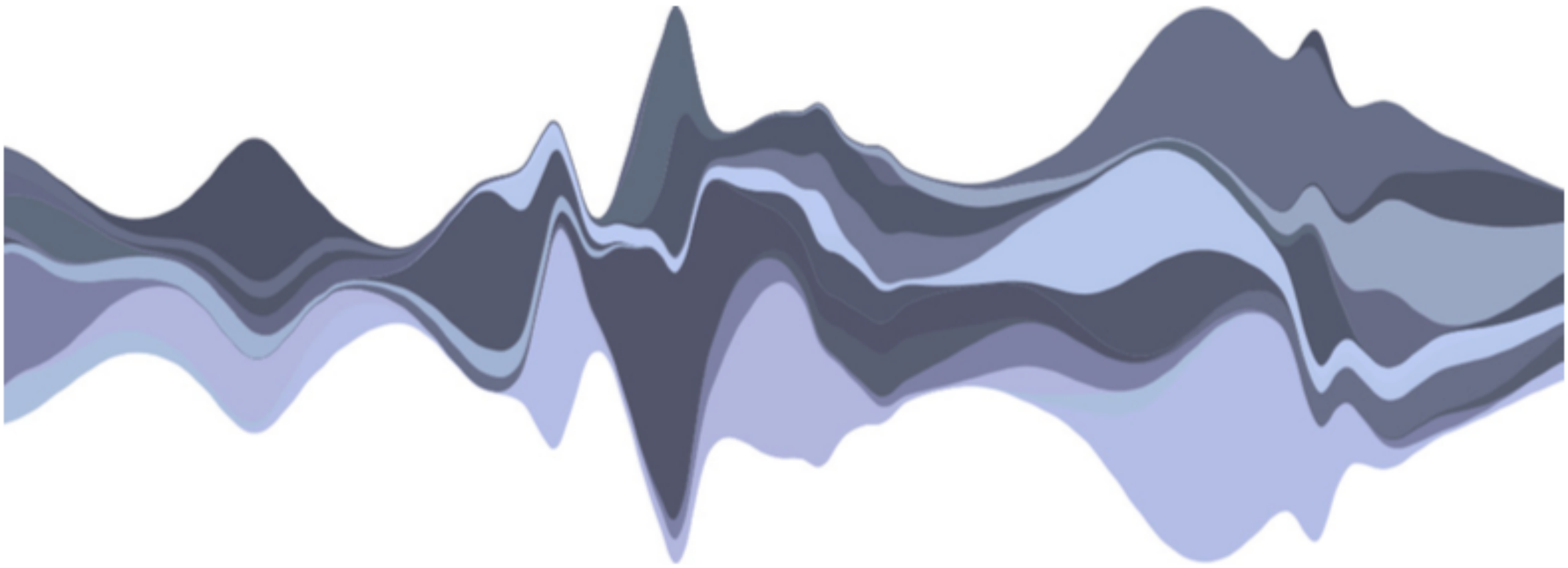


fig 6 – the same data set using the ThemeRiver layout algorithm

http://www.leebyron.com/else/streamgraph/

Streamgraphs

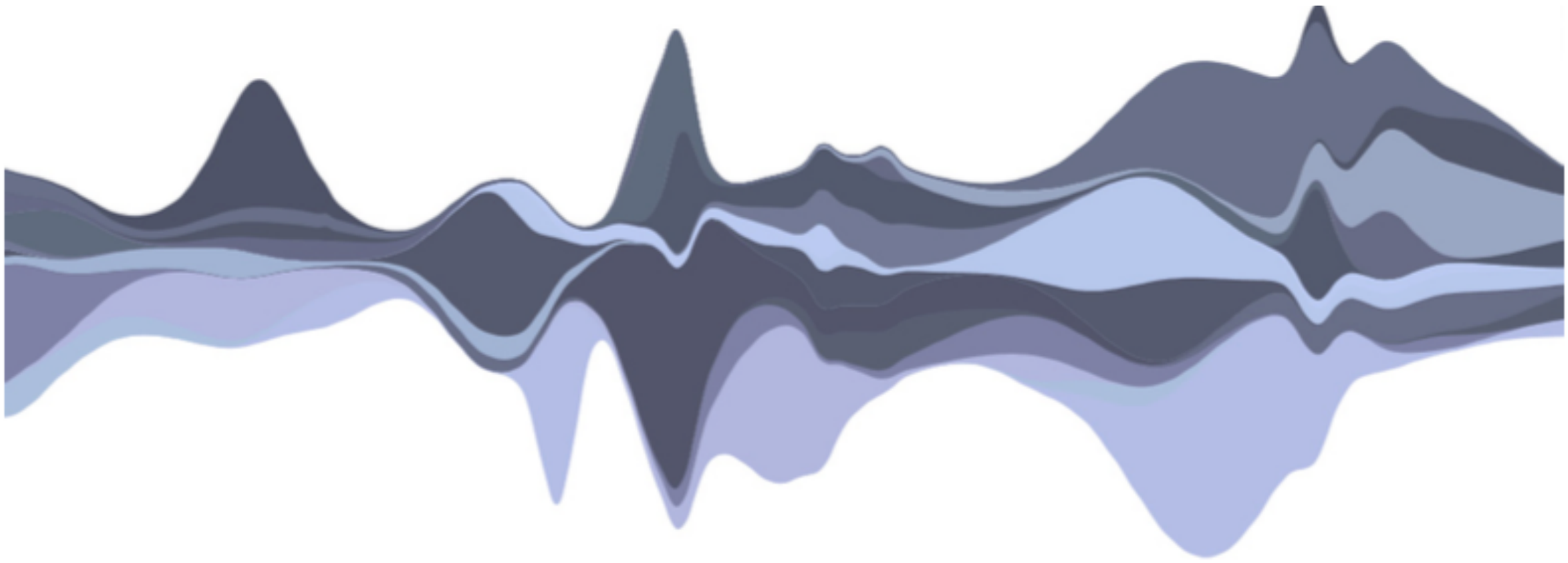


fig 7 – the same data set optimized to reduce the ”wiggle” function, or overall variation in slope

http://www.leebyron.com/else/streamgraph/

Streamgraphs

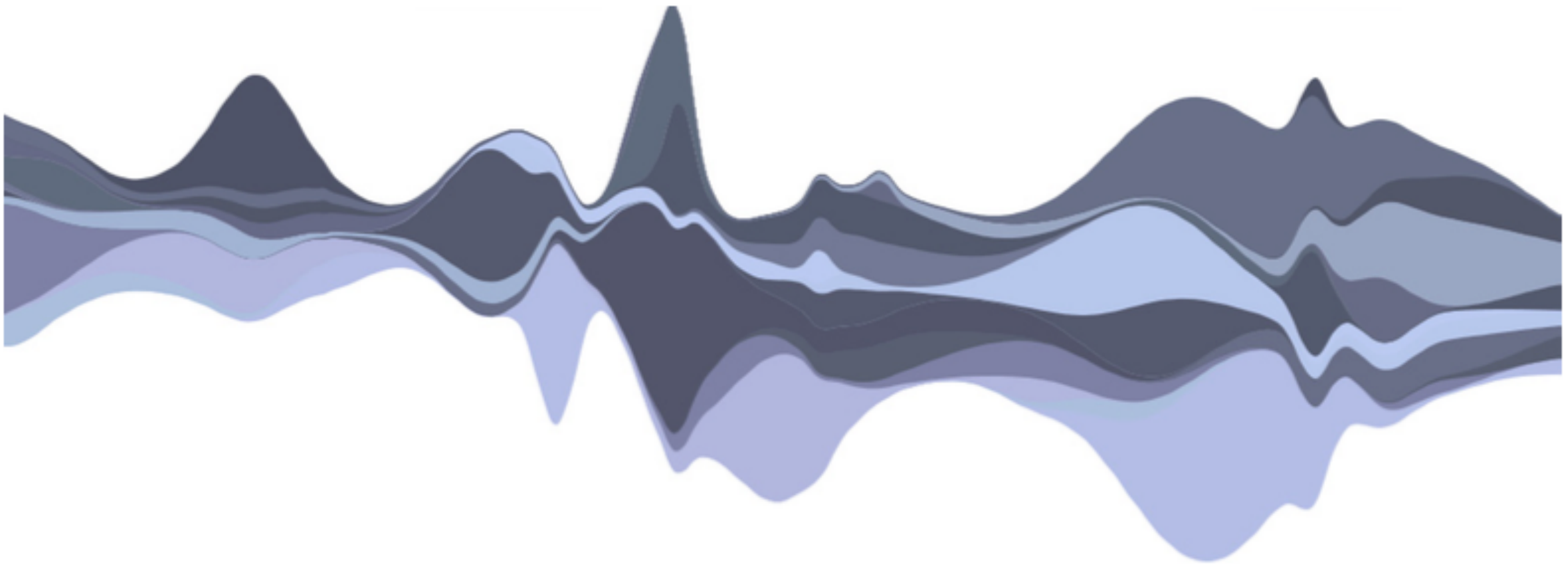
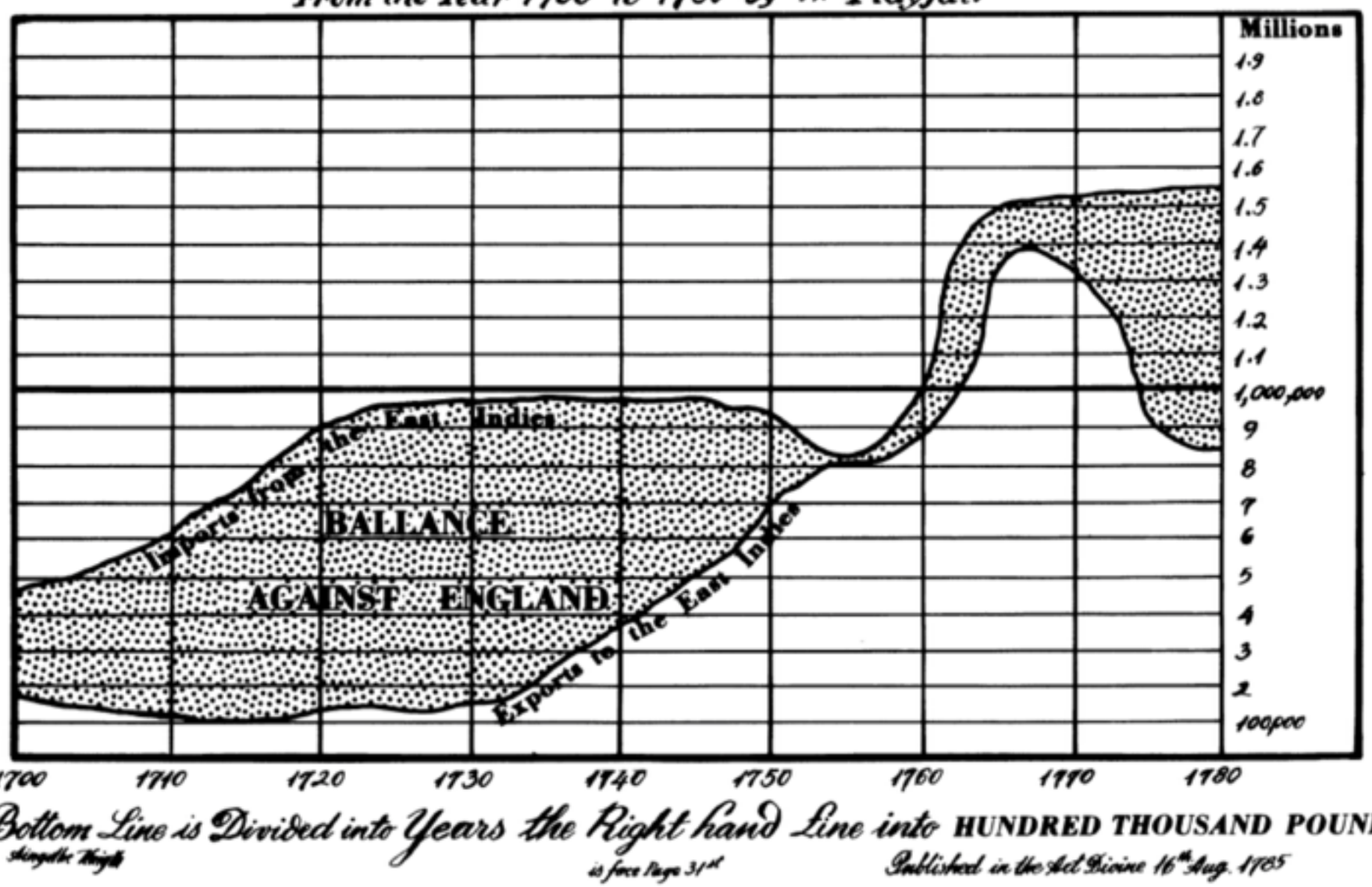


fig 8 – the same data set optimized to reduce the "weighted_wiggle," the algorithm used in Streamgraph

http://www.leebyron.com/else/streamgraph/

Streamgraphs



Streamgraphs

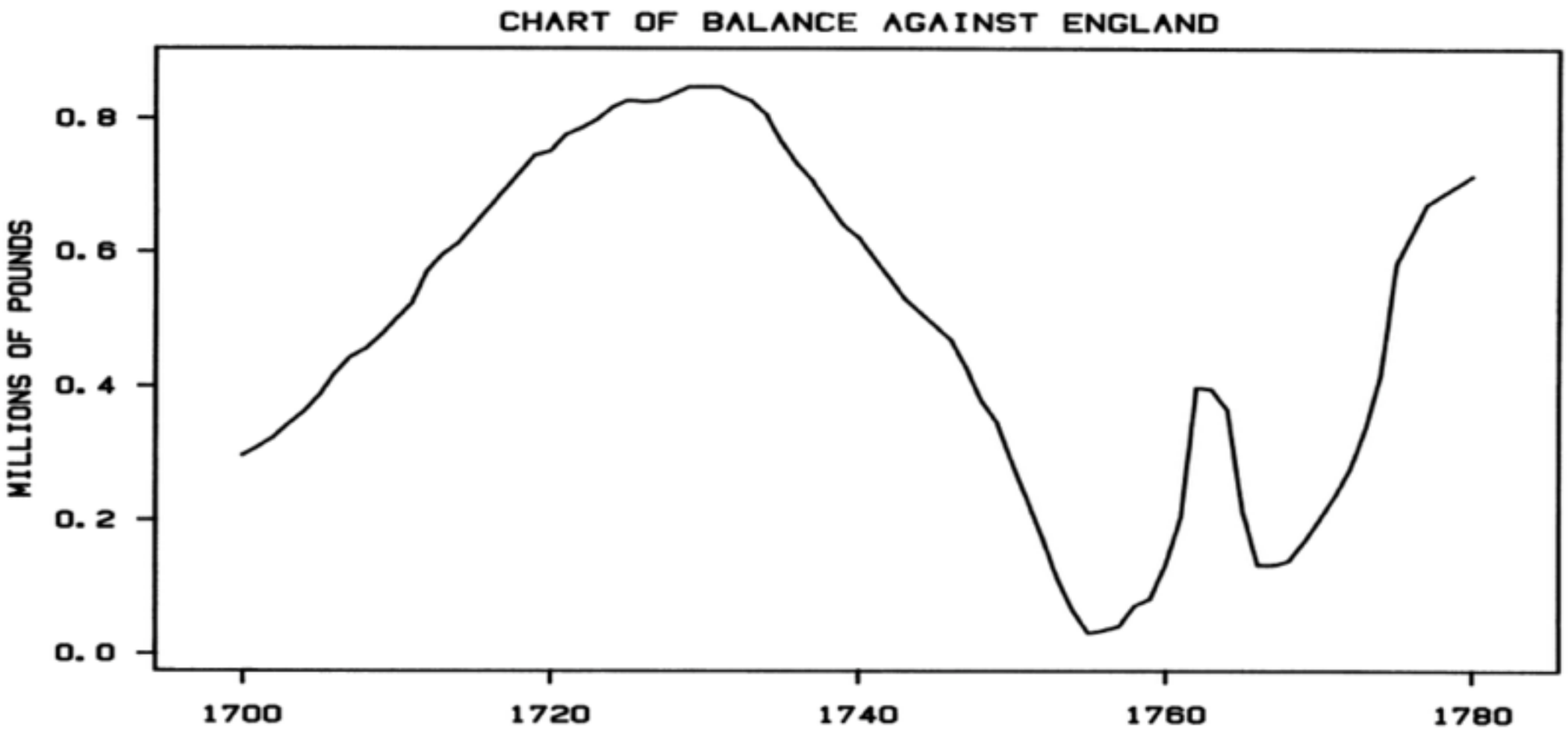


Figure 28. Playfair data.

Hoffman and Vendettuoli

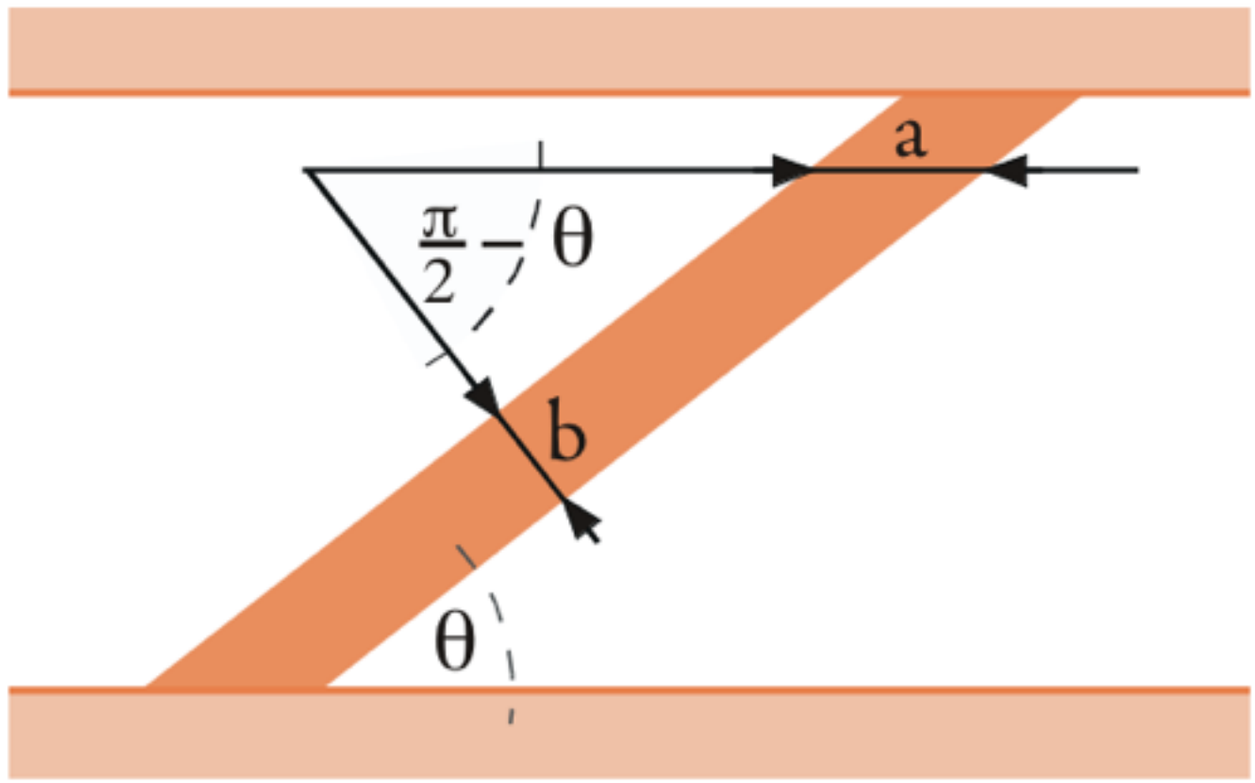
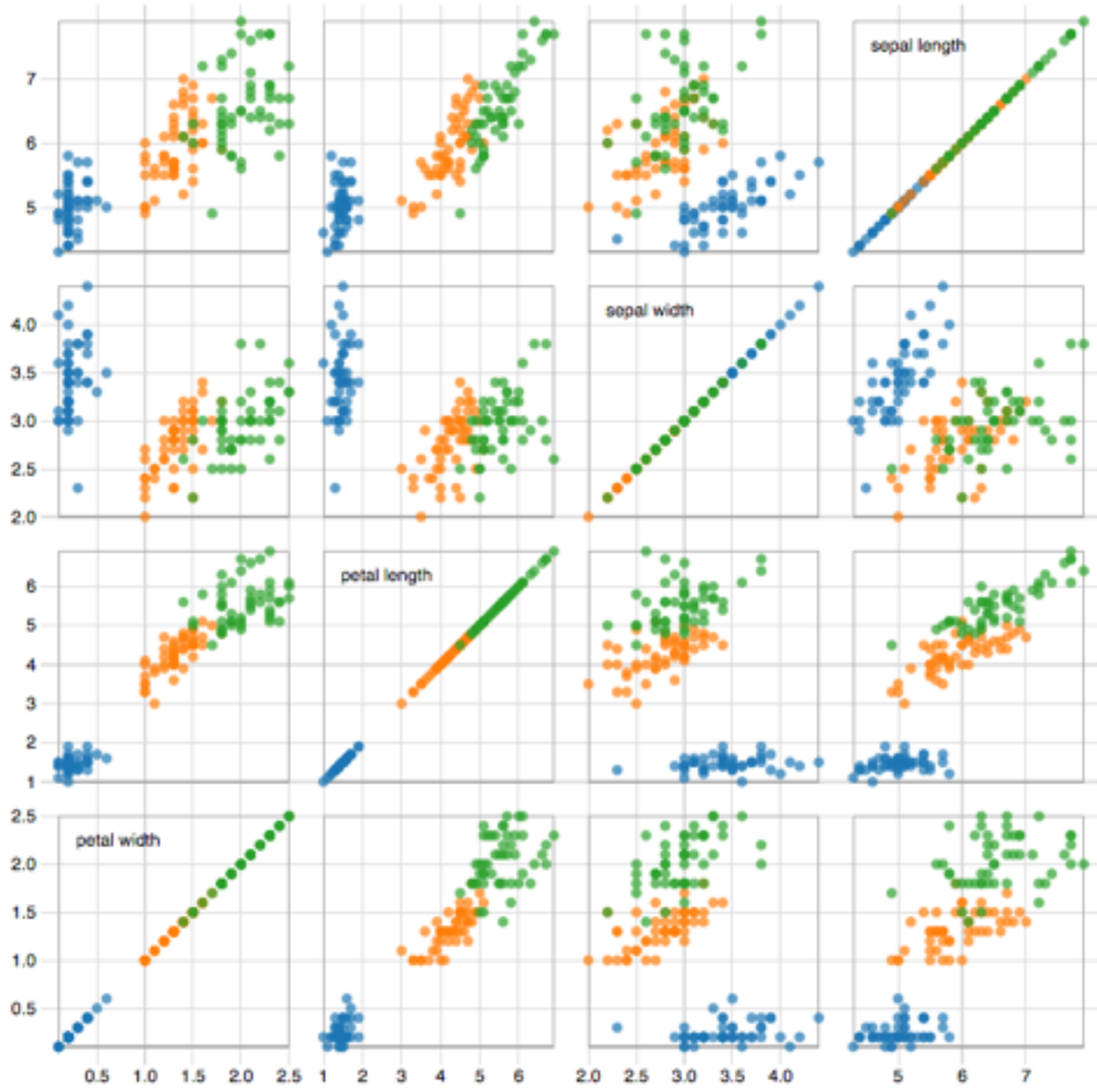


Fig. 5. Sketch of line width assessments: (a) is showing horizontal width, (b) shows width orthogonal to the slope. Survey results in section 4.2 indicate that observers associate line width more with orthogonal width w_o (b) than horizontal width w_h (a).

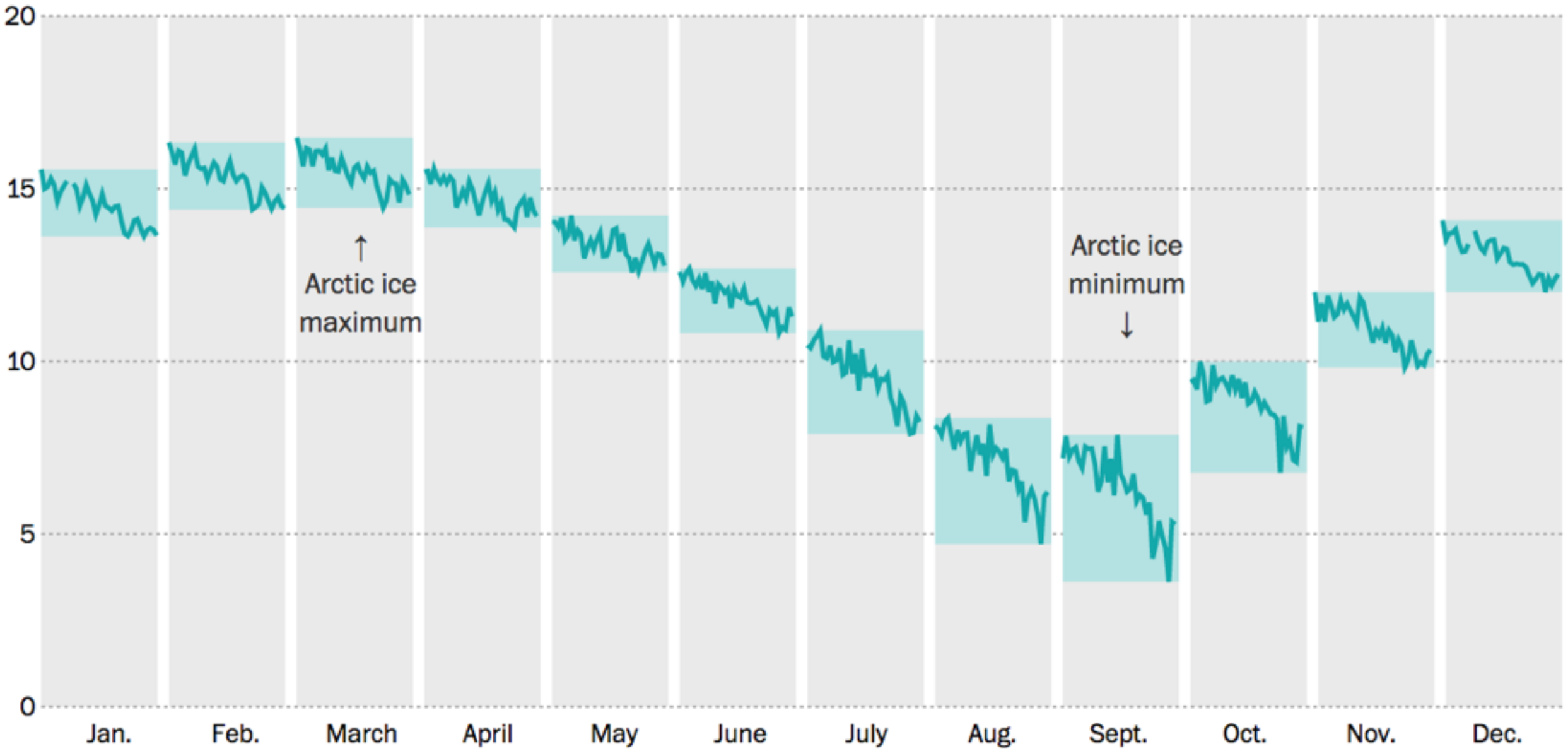
Many dimensions

Small Multiples



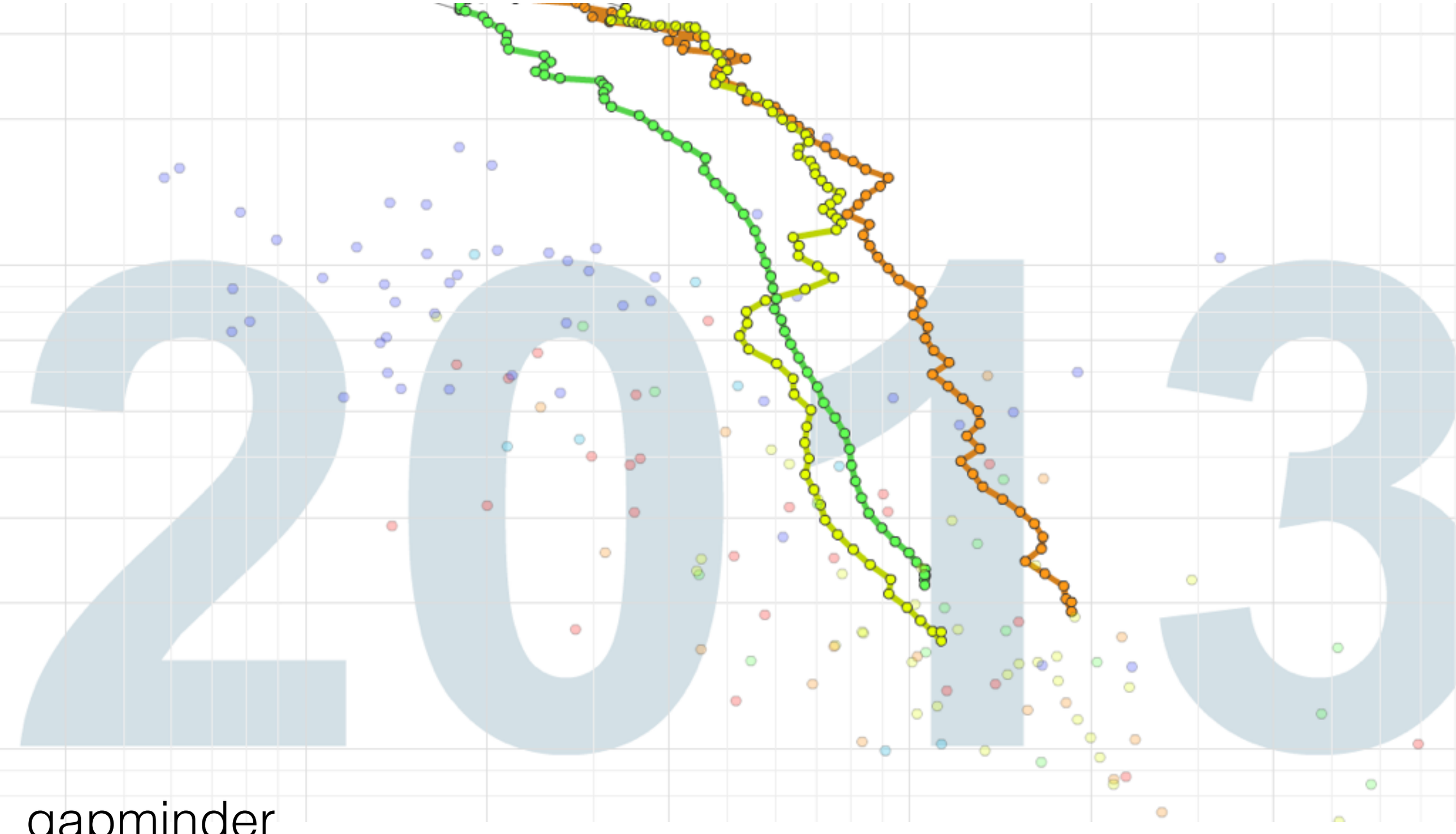
http://bl.ocks.org/mbostock/4063663

Small Multiples



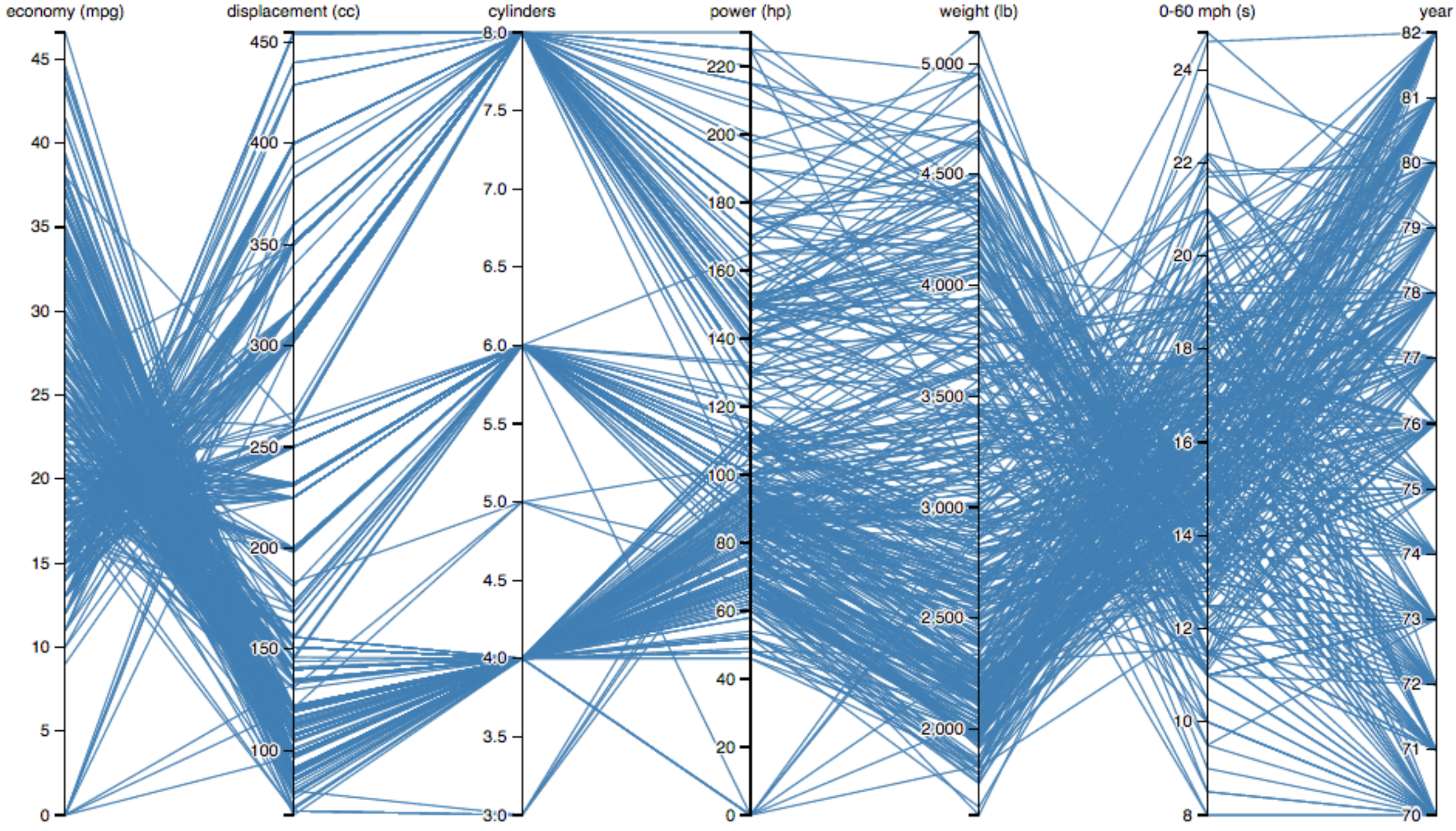
http://www.washingtonpost.com/graphics/national/arctic-ice-2015/

Multivariate Time Series



gapminder

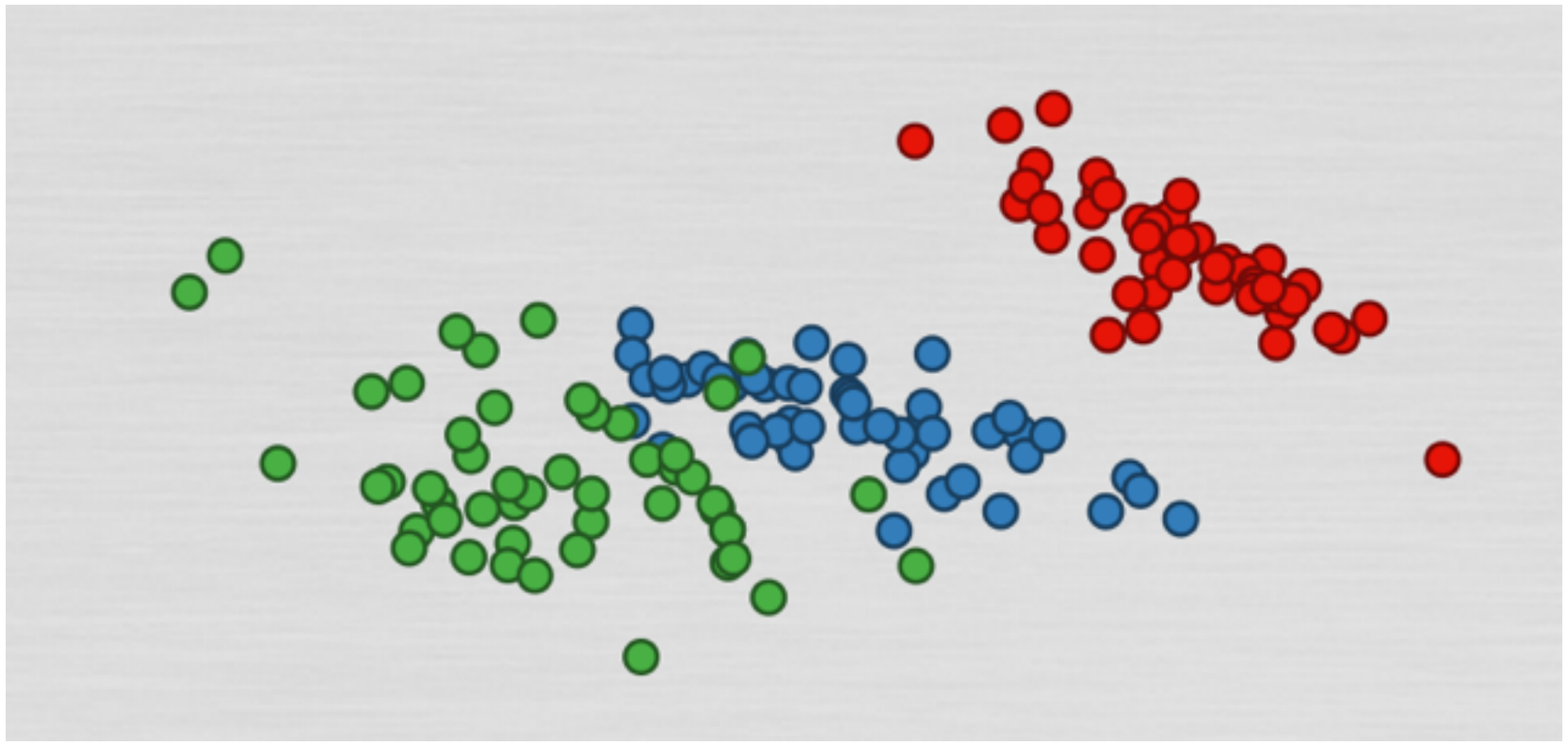
Parallel Coordinates



http://bl.ocks.org/jasondavies/1341281

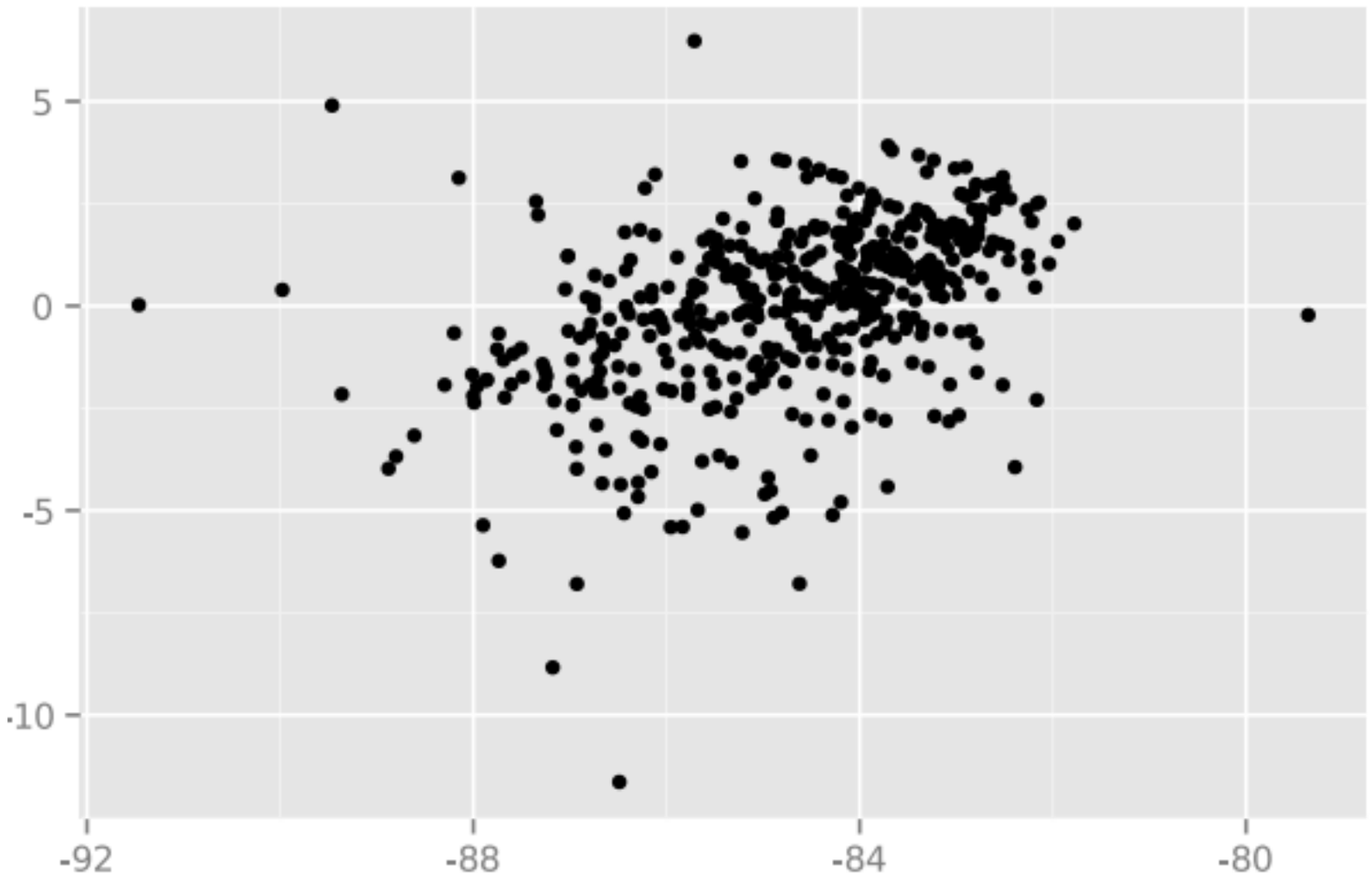
Later in the course: Dimensionality Reduction

Later in the course: Dimensionality Reduction

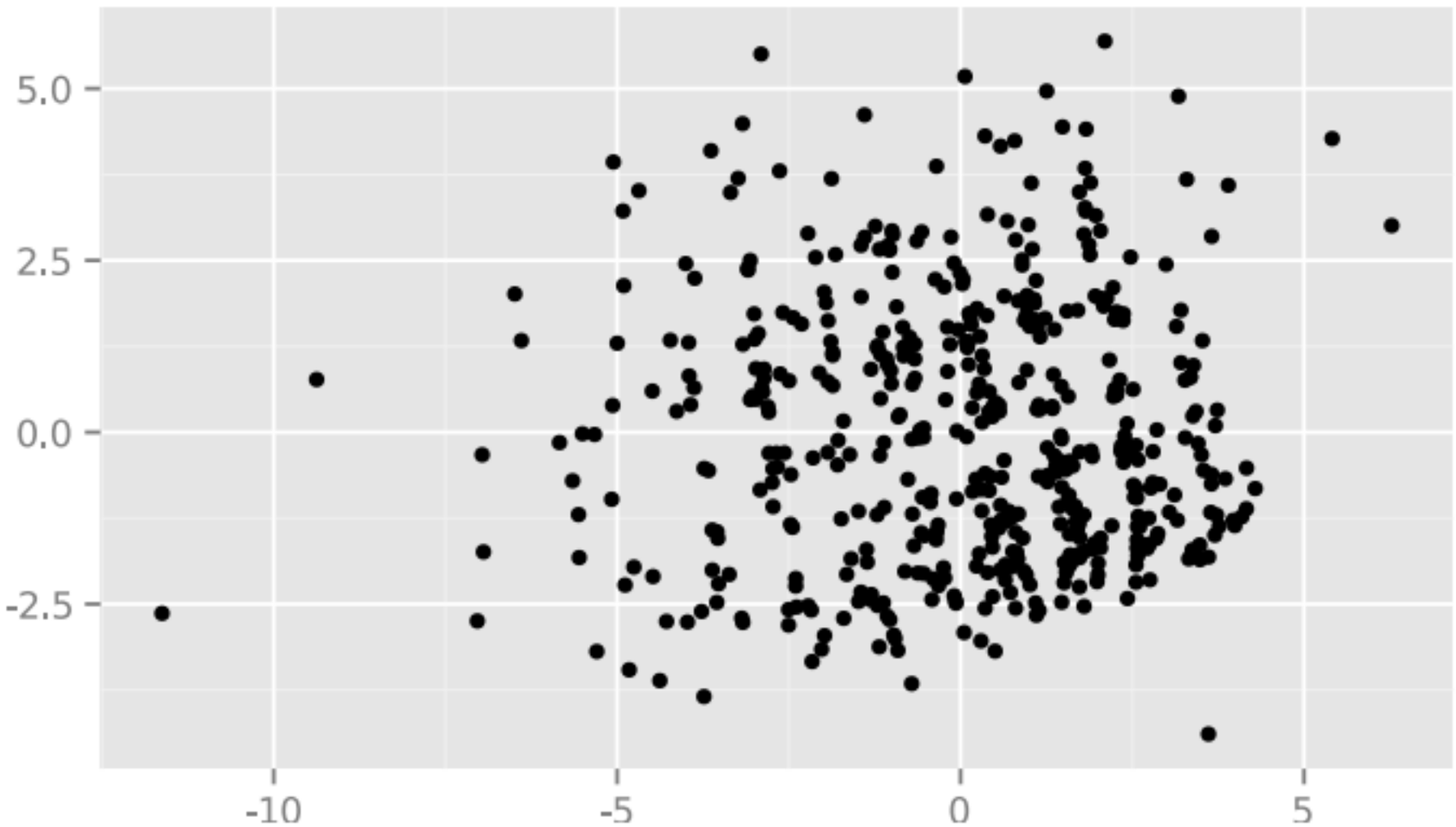


http://cscheid.github.io/lux/demos/tour/tour.html

Later in the course: Dimensionality Reduction



Later in the course: Dimensionality Reduction



Later in the course: Dimensionality Reduction