

# More on Multivariate Data & Representations



CS 4460 - Information Visualization  
Jim Foley

Original PPTs John Stasko, augmented by J. Foley. Last update September 2014

# Quick Review – Common MV Rep'ns



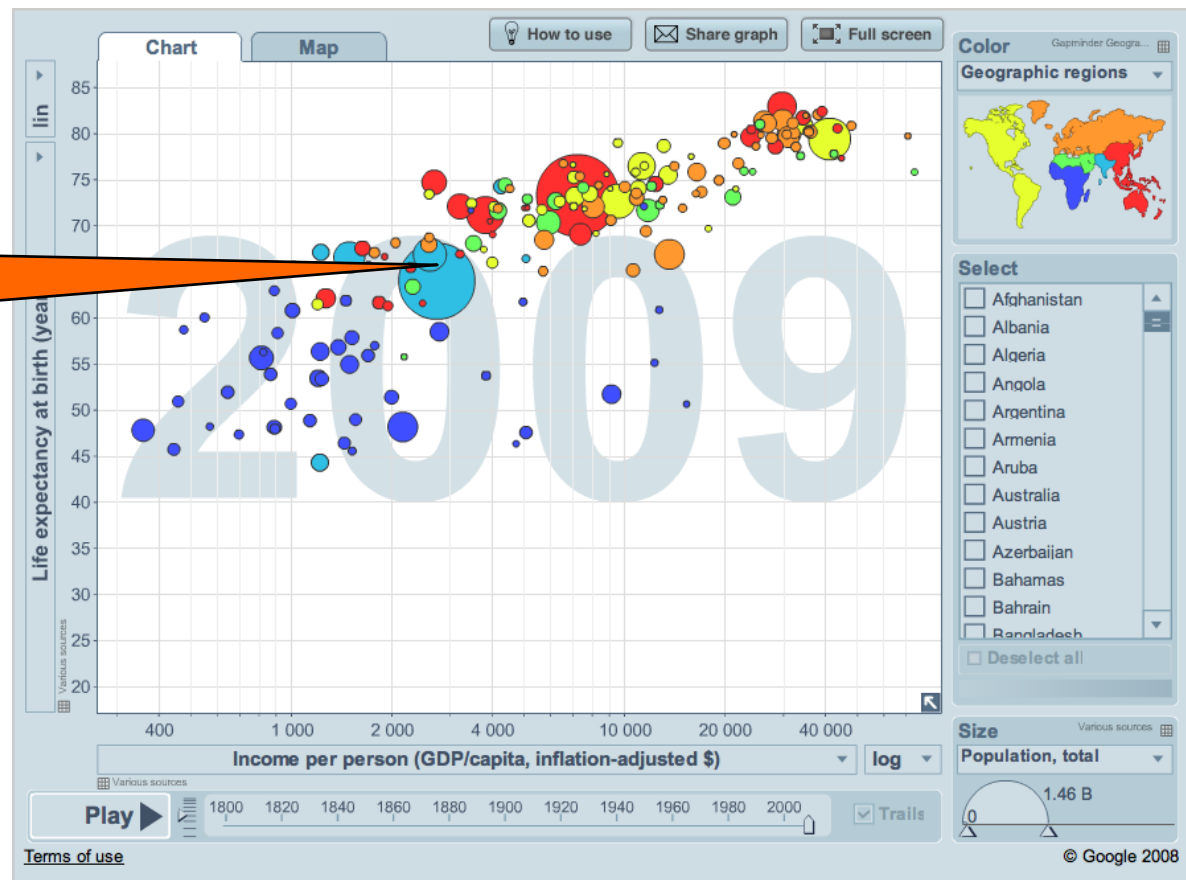
- Scatter plot
- Enhanced scatter plot (as in Gapminder)
- Bar charts
  - Simple, stacked, clustered
- Pie chart
- Chernoff faces ☹️
- Small multiples of any of the above
  - Scatter plot matrix etc.
- Parallel coordinates

# Enhanced Scatter Plot - Gapminder

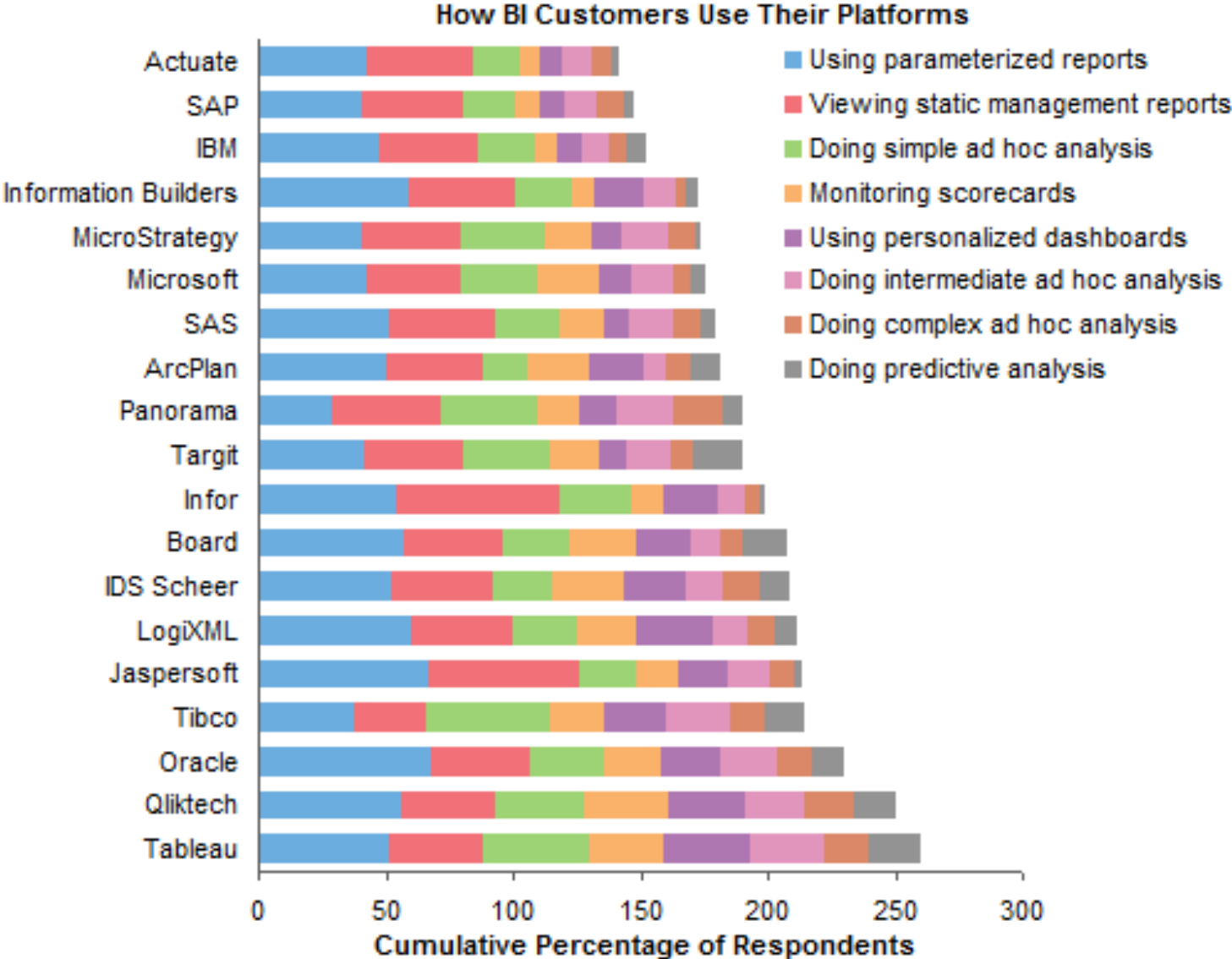


Remember the idea of Marks?

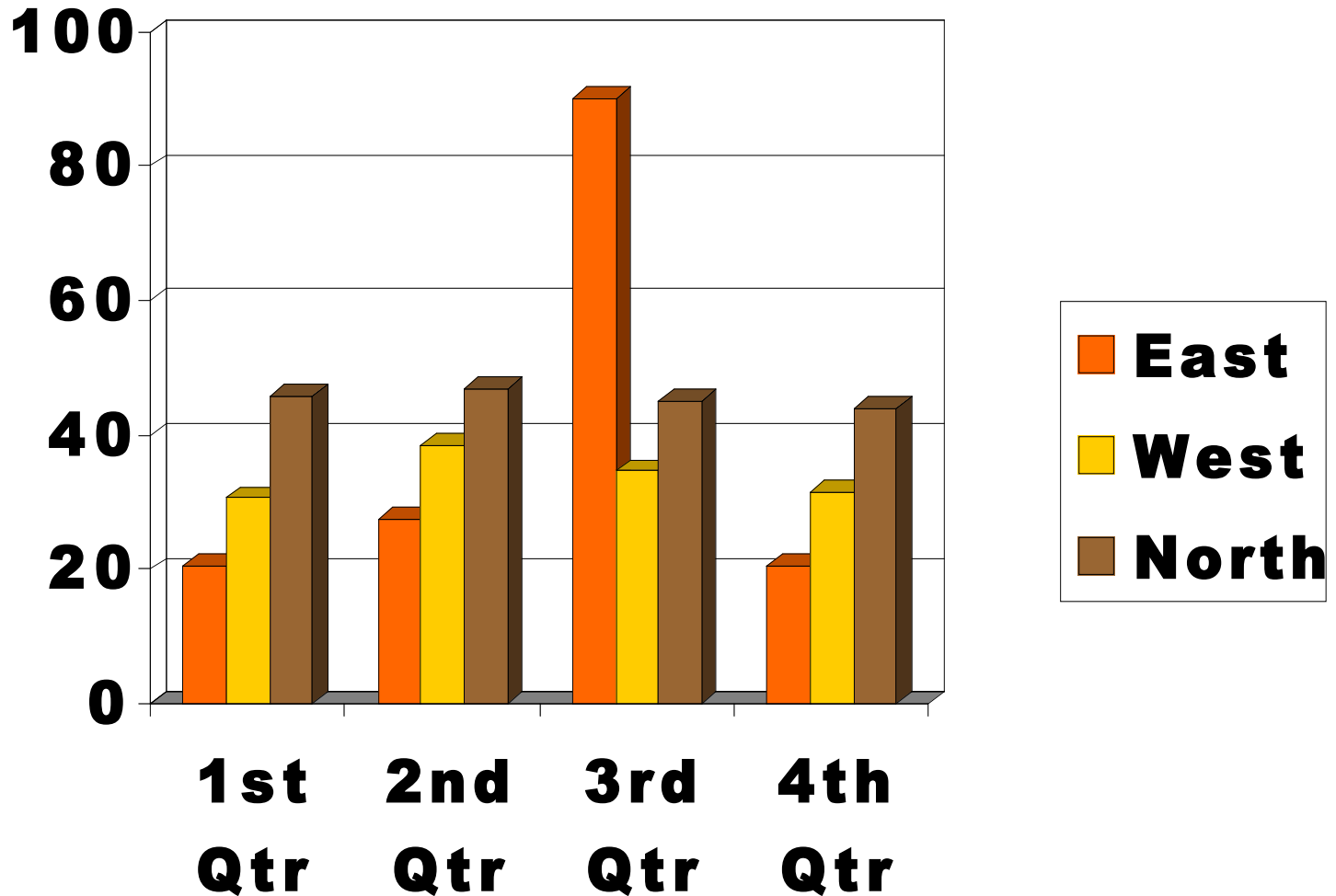
These are Marks!



# Stacked Bar Chart



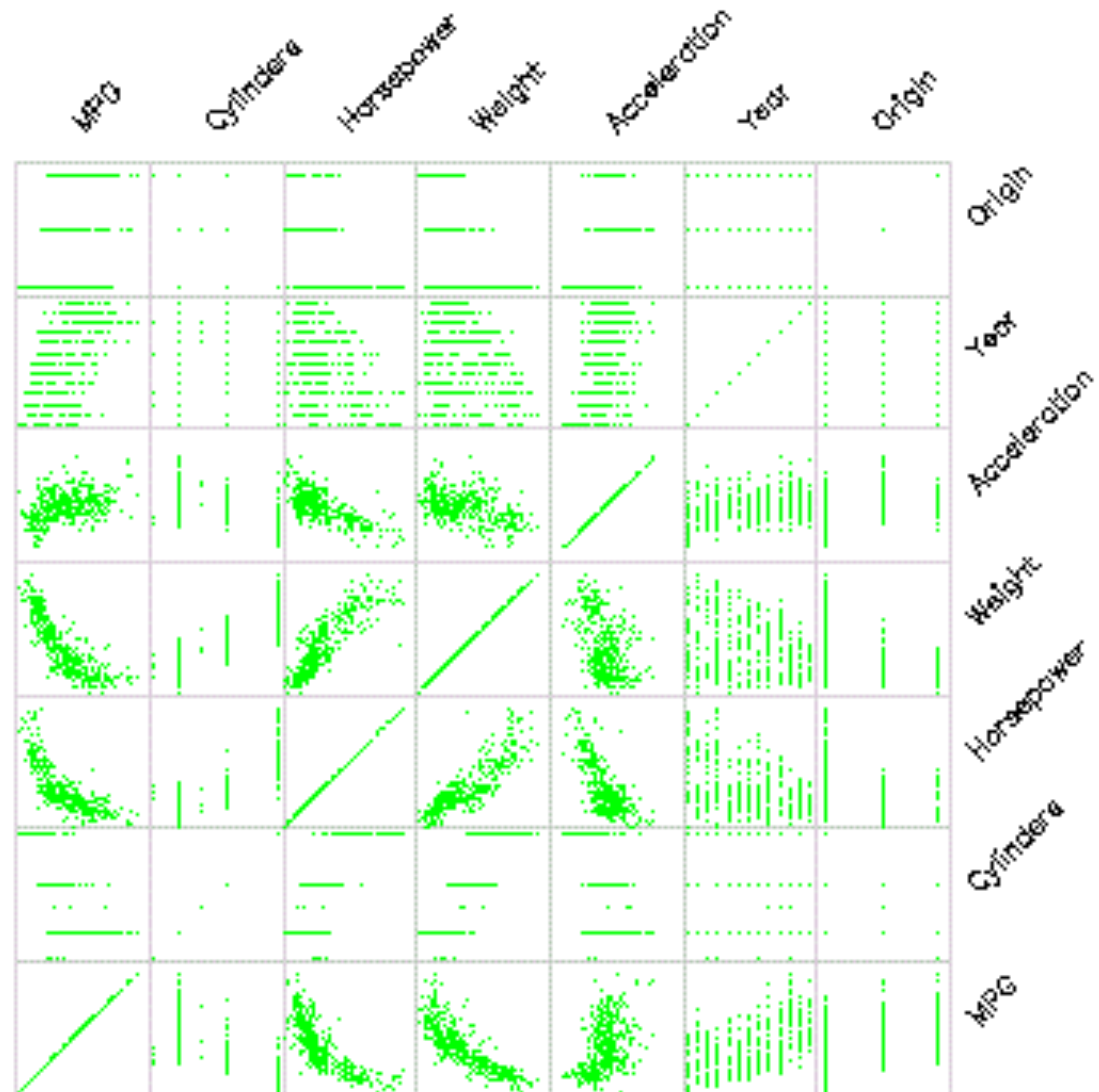
# Clustered Bar Chart



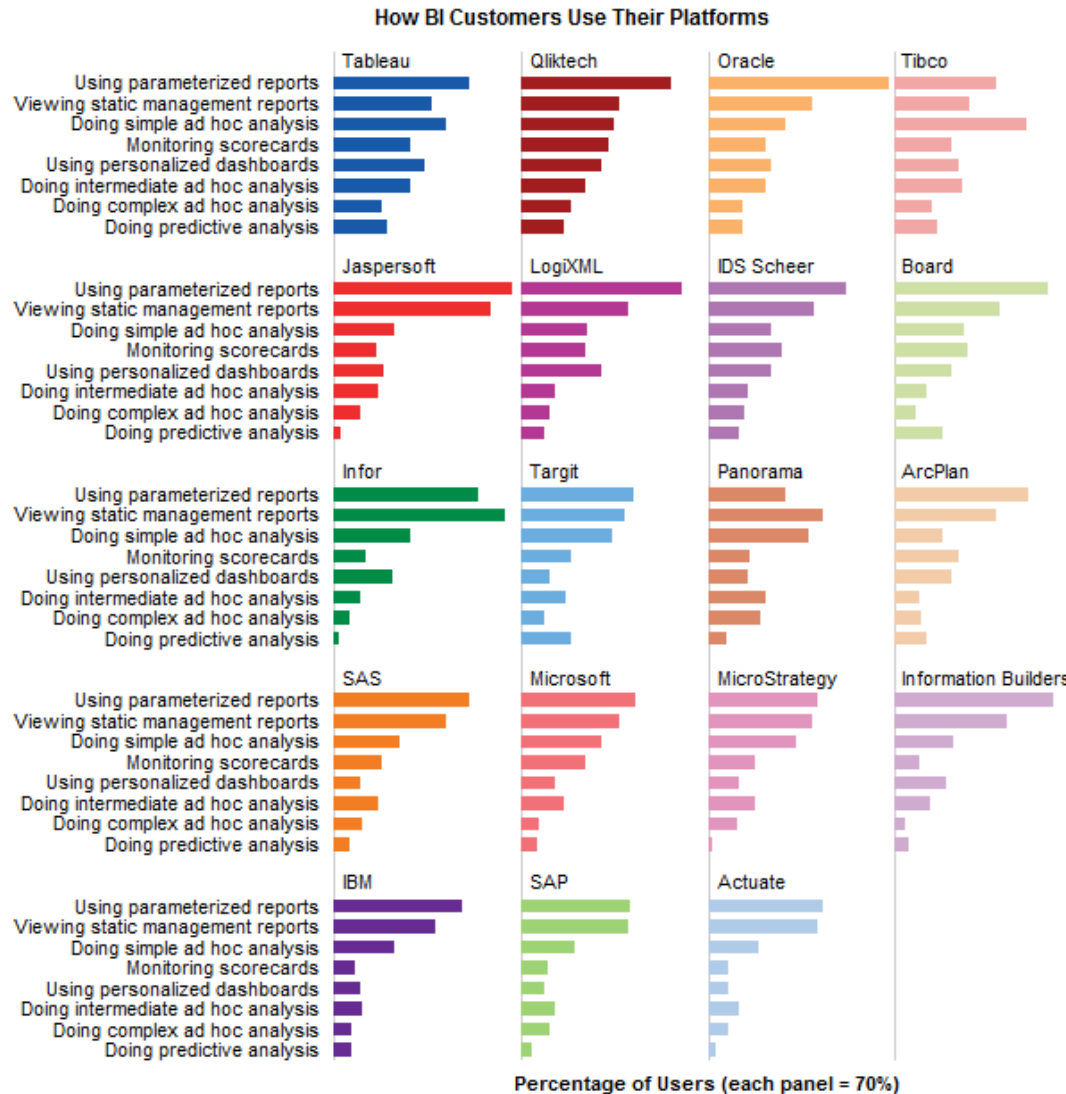
# Small Multiples: Scatterplot Matrix



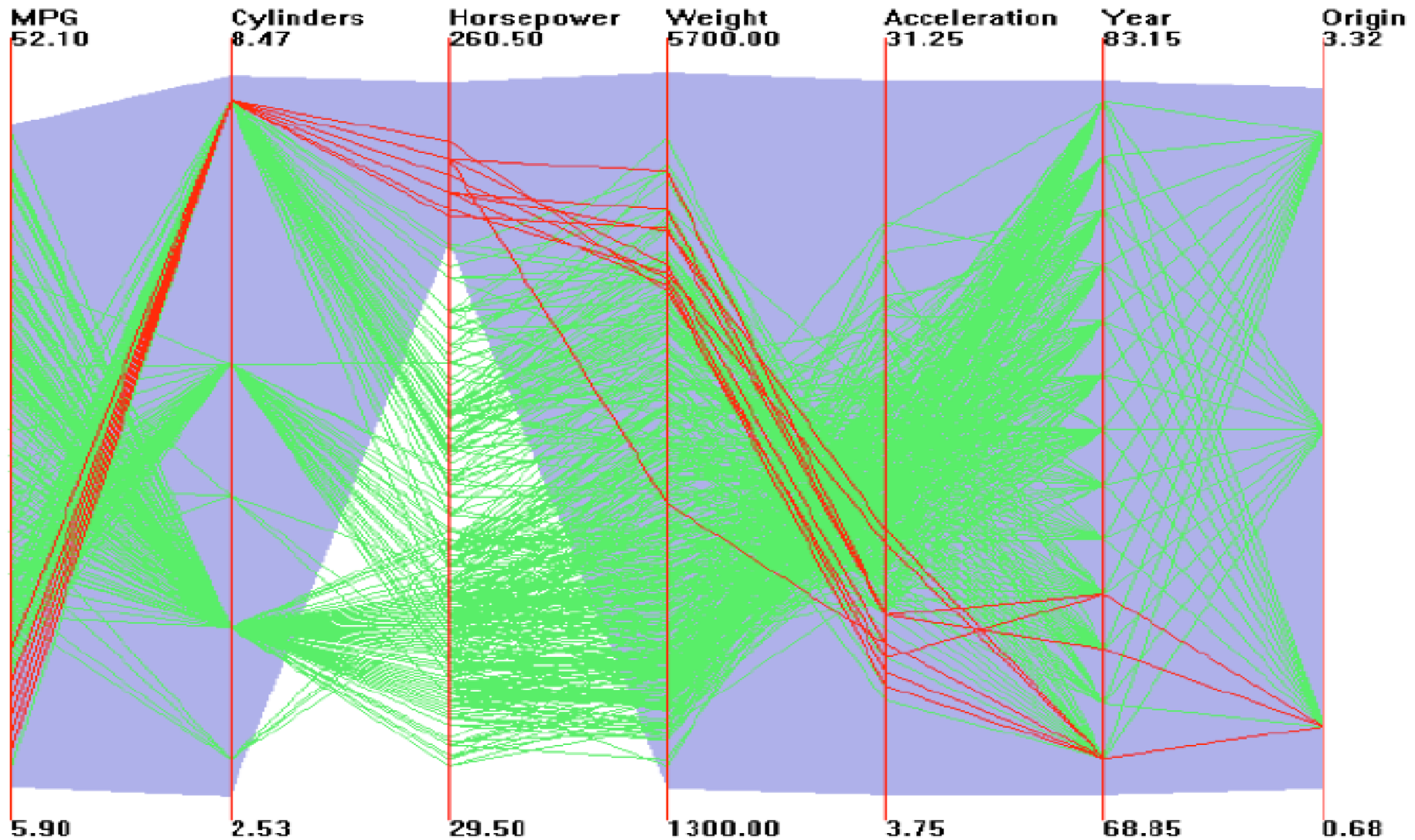
Represent each pair of variables in its own 2-D scatterplot



# Small Multiples: Bar Charts

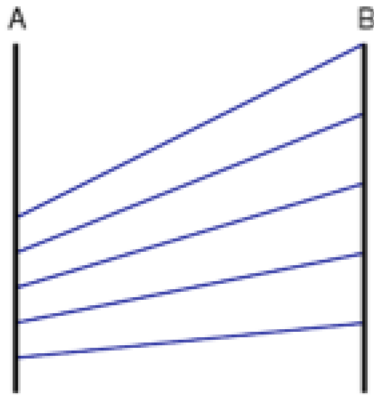


# Automobile Data in Parallel Coords

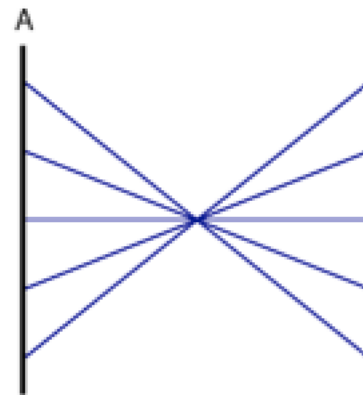
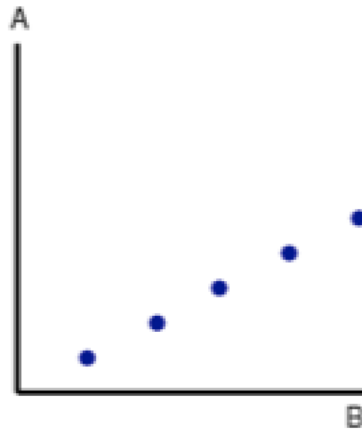




# Parallel Coords $\Leftrightarrow$ Scatter Plots



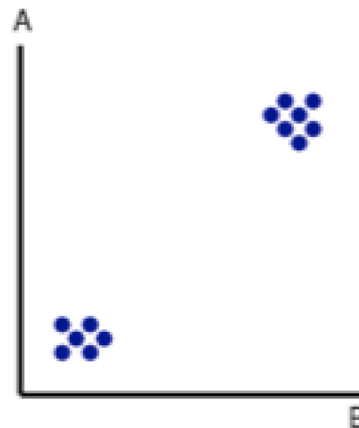
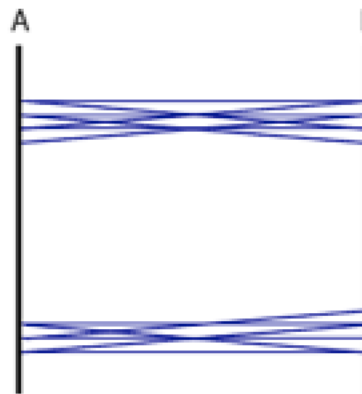
(a) A positive linear relationship.



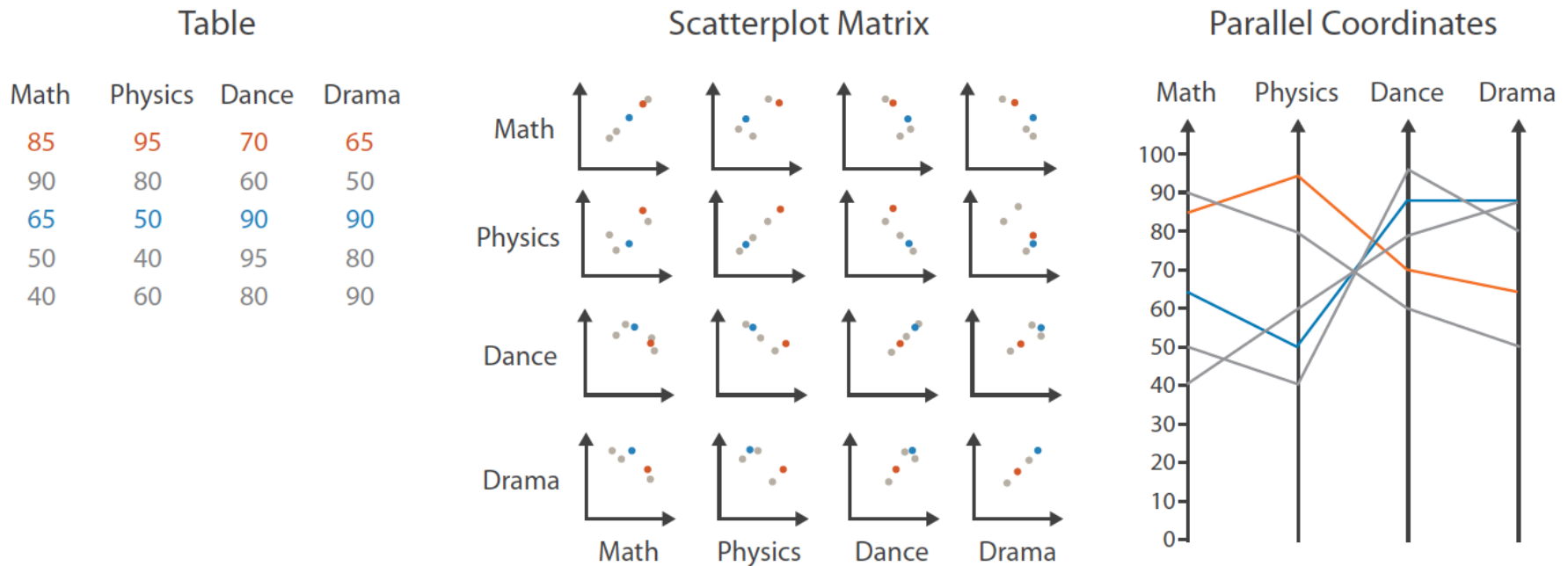
(a) A positive linear relationship.



(b) A negative linear relationship.



# Table, Scatterplot, Parallel Coords



**Figure 7.12.** Comparison of scatterplot matrix and parallel coordinate idioms for a small data table. After [McGuffin 14].

From Munzer

# Parallel Coordinate Example



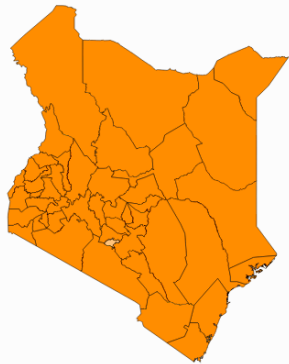
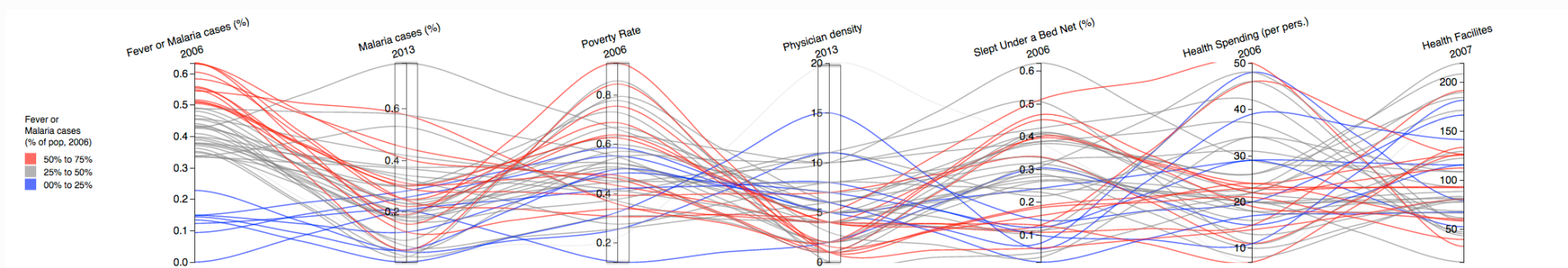
Fork me on GitHub

## Malaria Data Exploration

### Kenya counties in 2006 - 2013

Each line represents a county. They are colored by percentage of people that had fever or malaria in 2006. To filter data, click and drag along an axis. To reorder the dimensions, drag the axes titles around.

Next step, explore the data! For example, you can drag the Poverty Rate axis next to the Fever or Malaria axis and then filter on the Fever or Malaria axis to check for correlations. While you filter and move the filter along the axis, you can also observe the selected counties on the map.



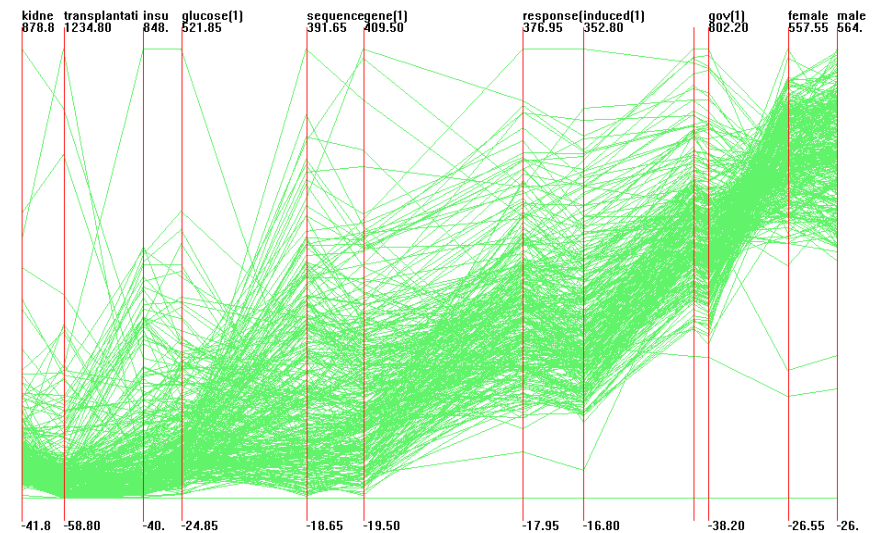
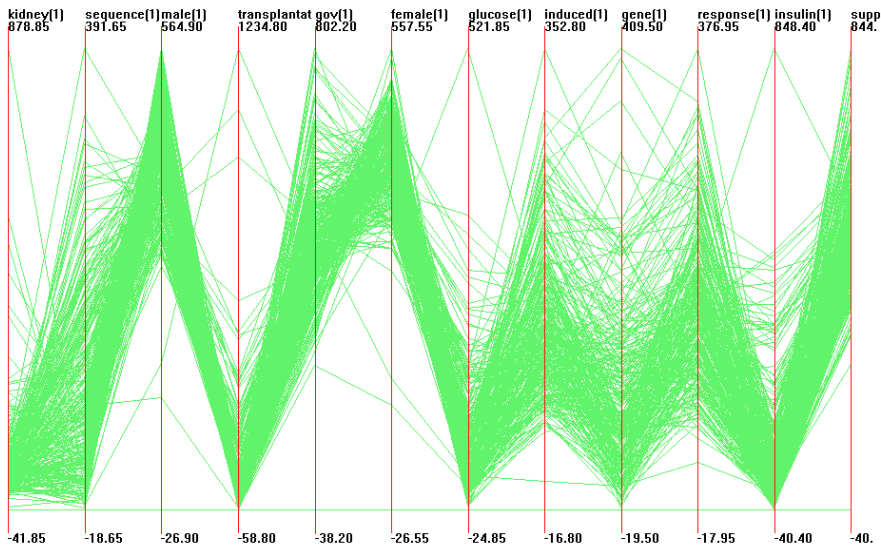
County	Fever or Malaria cases (...)	Malaria cases (%) // 201...	Poverty Rate // 2006	Physician density // 2013	Slept Under a Bed Net (...)	Health Spending (per per...	Health Facilities // 2007
BARINGO	0.149	0.28051	0.585	5	0.242	29	115
BOMET	0.395	0.19962	0.507	2	0.263	20	77
BUNGOMA	0.376	0.24506	0.522	4	0.278	20	88
BUSIA	0.428	0.7751	0.66	6	0.407	20	81
ELGEYO MARAKWET	0.094	0.25528	0.552	5	0.075	48	79
EMBU	0.398	0.34122	0.408	11	0.218	34	108
GARISSA	0.466	0.04769	0.545	10	0.384	37	61
HOMA BAY	0.583	0.44668	0.431	4	0.396	24	125
ISILO	0.337	0.2985	0.631	10	0.426	42	43
KAJIADO	0.123	0.20442	0.121	2	0.304	15	115
KAKAMEGA	0.379	0.37295	0.521	5	0.267	23	219
KERICHO	0.516	0.29122	0.392	7	0.338	11	126

<https://realimpactanalytics.github.io/d4g-hackathon-malaria-viz/>

# Dimensional Reordering



Which dimensions are most like each other?



Same dimensions ordered according to similarity. Shows overall trends more clearly.

Yang et al, InfoVis '03

# Enhanced Parallel Coordinates

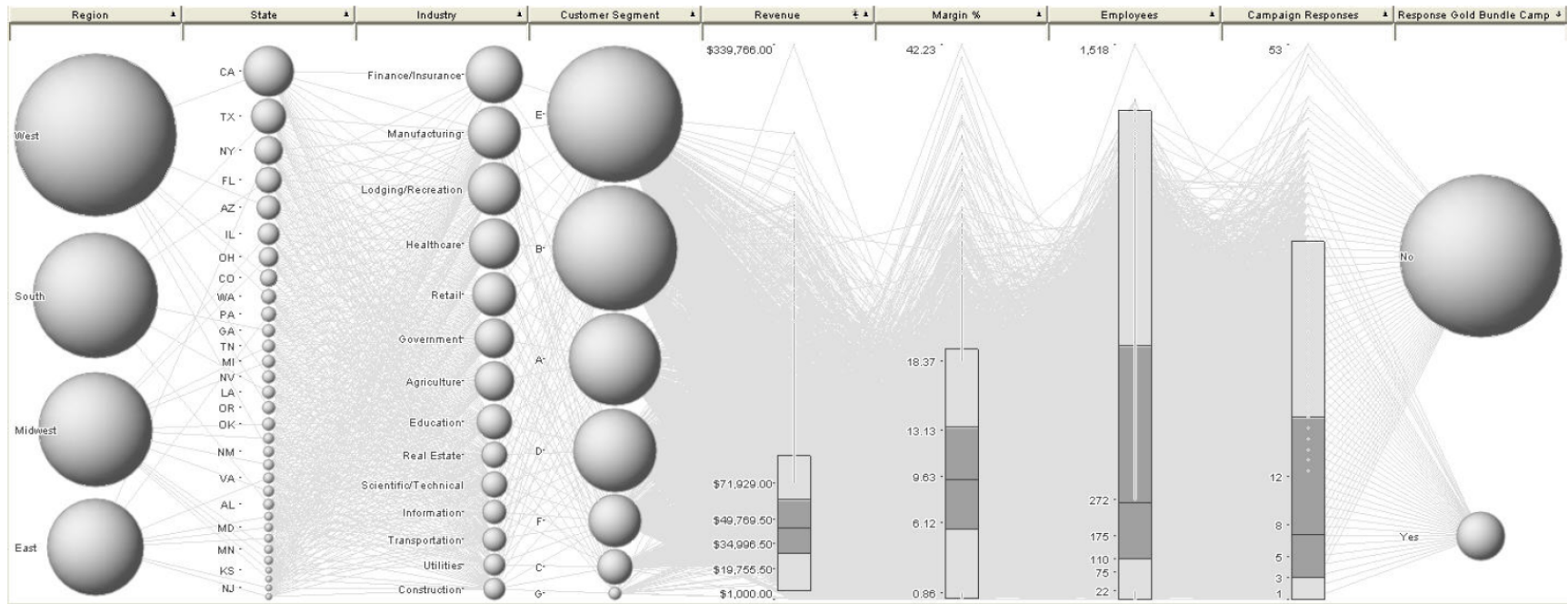
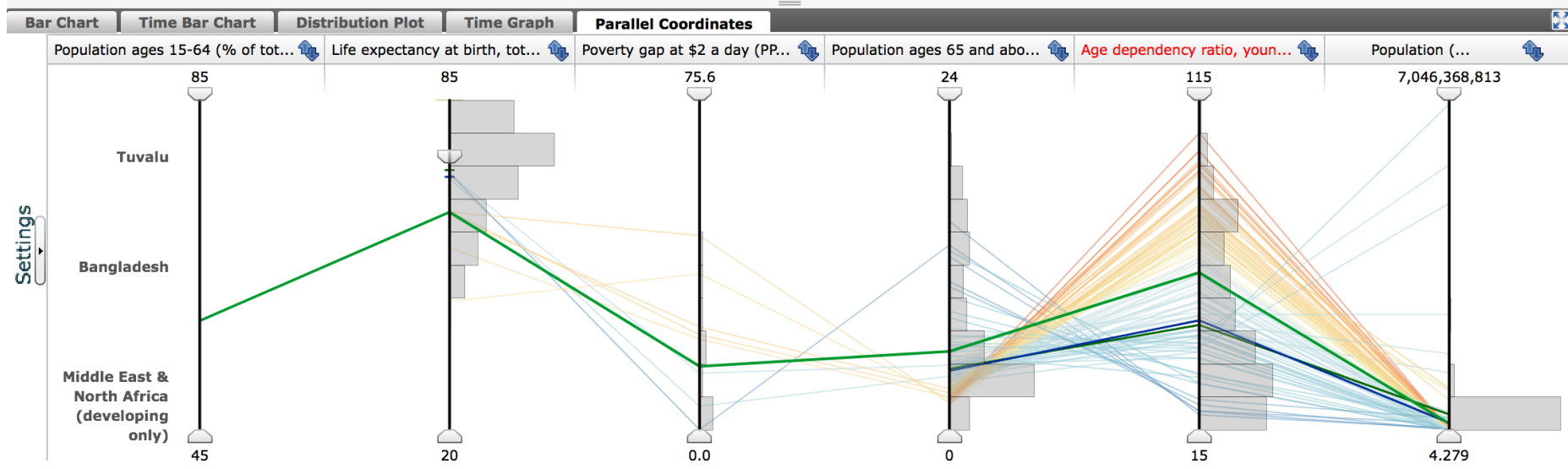
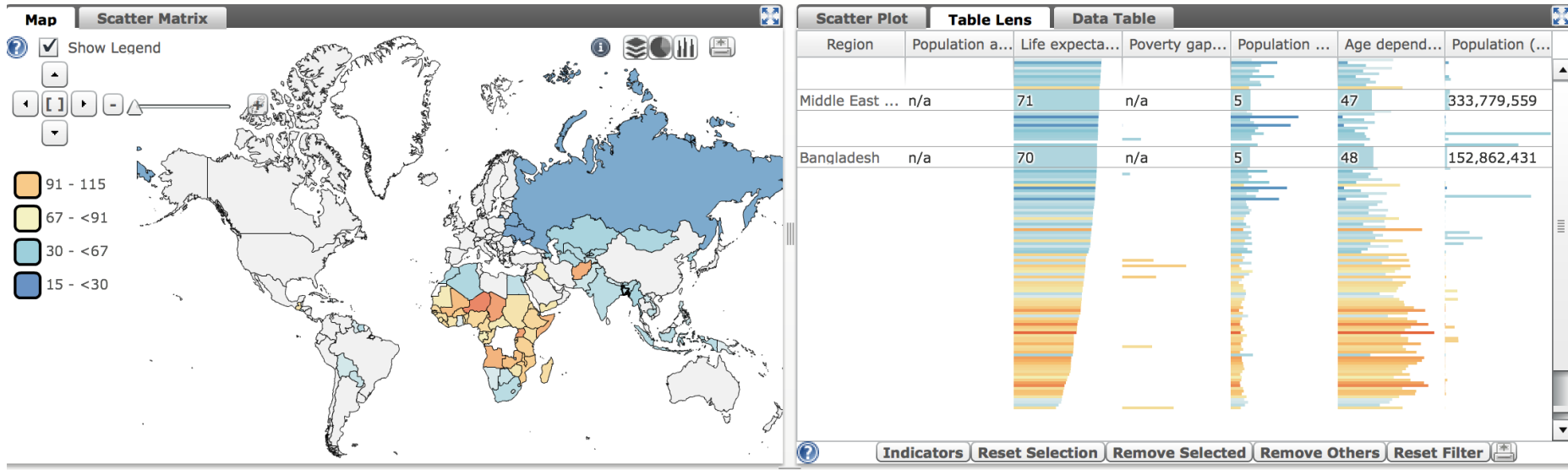


Figure 8: A parabox (another name for parallel coordinates) graph from Advizor Solutions.

Ordinal variables double coded; quantitative show distribution stats; how might nominal be encoded?

# Enhanced PCs – World Statistics Explorer



# Design Exercise

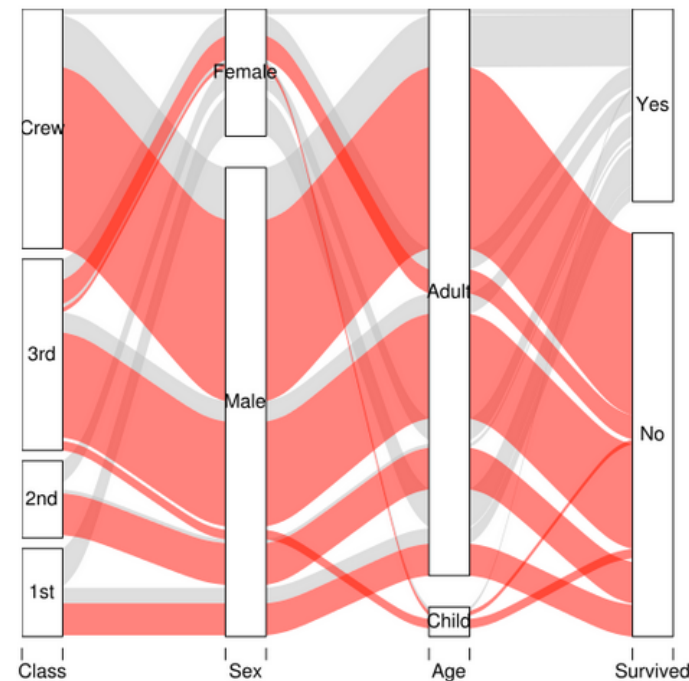


- On handout
- Count off by project team to work on first goal or second goal

# Parallel Sets



- Cousin of parallel coordinates
- Nominal & Ordinal variables
  - Or quantitative variables put into 'bins'  
Ordinal bins: 0-5, 5-10, 10-15, etc
- Does not show each case
- Instead shows size of sets



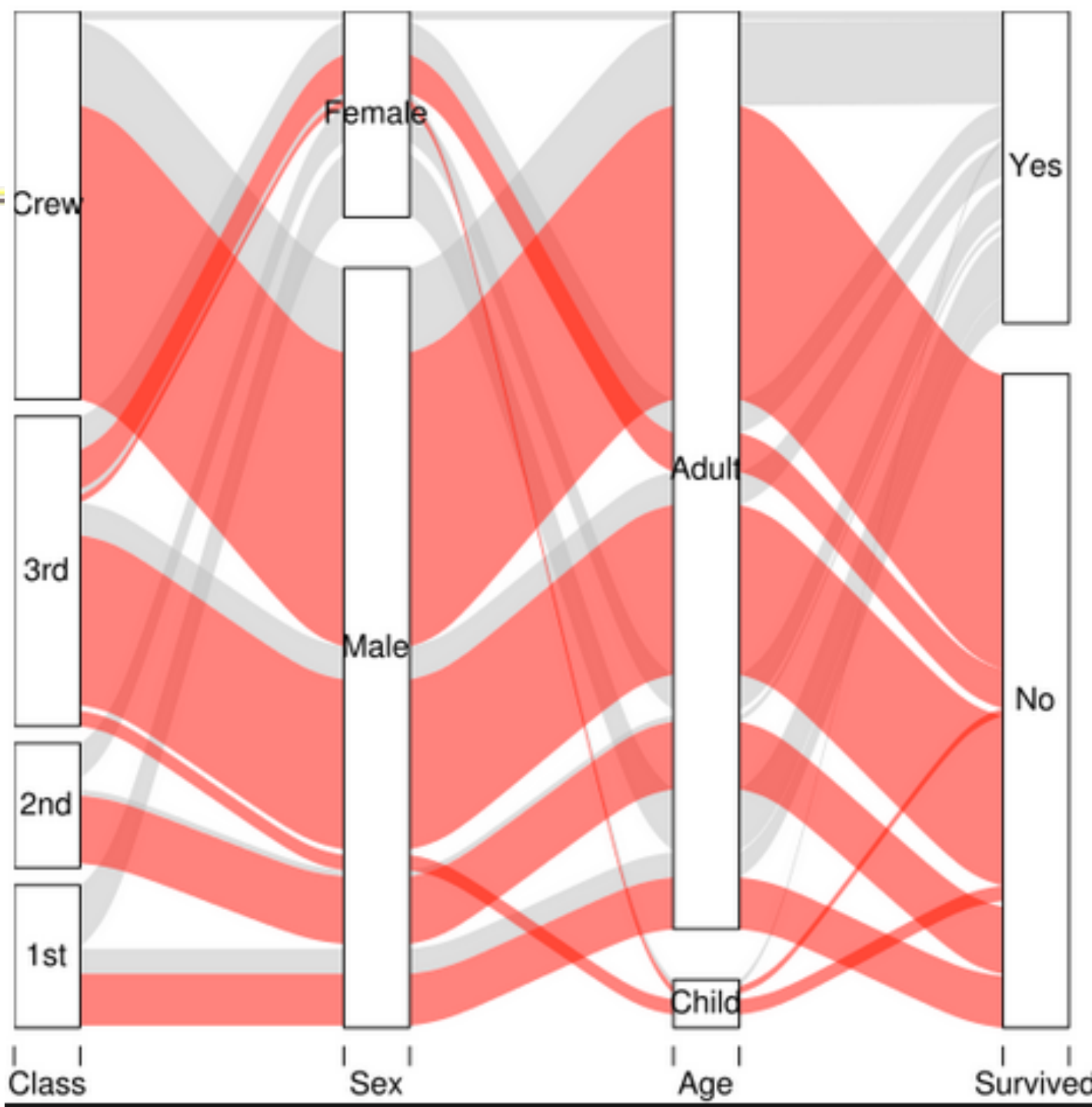


Titanic Data

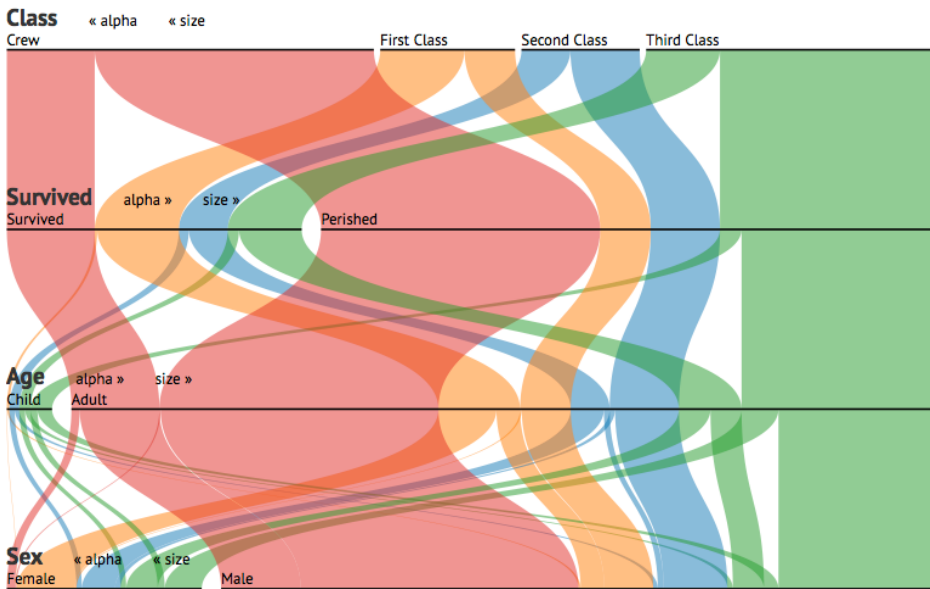
Red did not survive ☹️

Note similarity to parallel coords

What would parallel coord vis be?



# Variations



Horizontal “parallel sets”

Curved edges

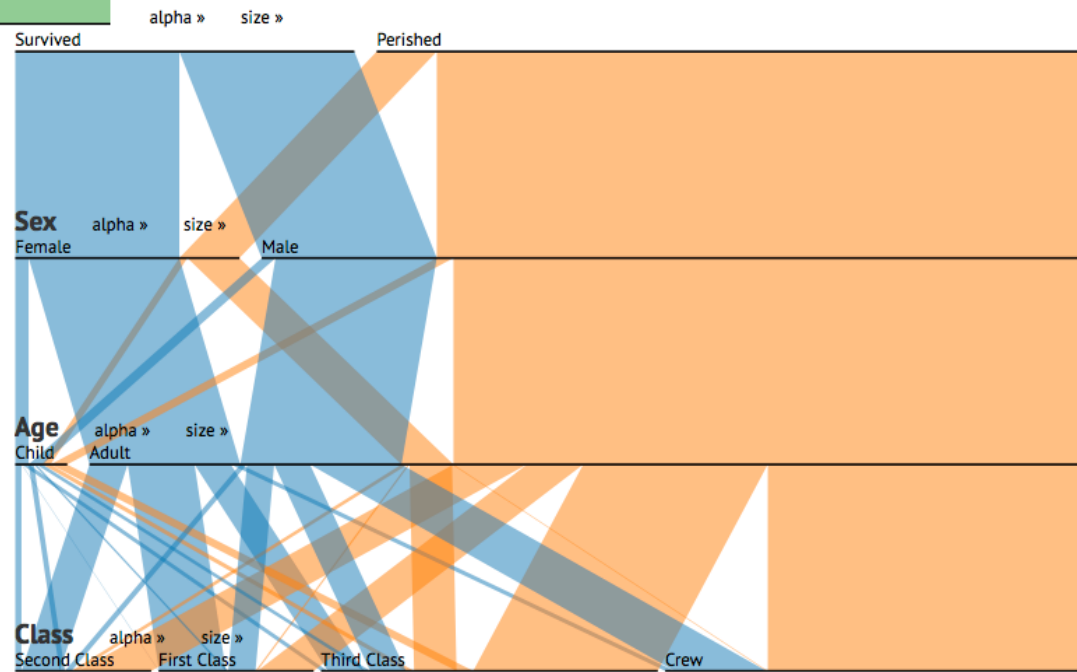
Reorder variables

Browser program:

<https://www.jasondavies.com/parallel-sets/>

Downloadable program:

<https://code.google.com/p/parsets/downloads/list>

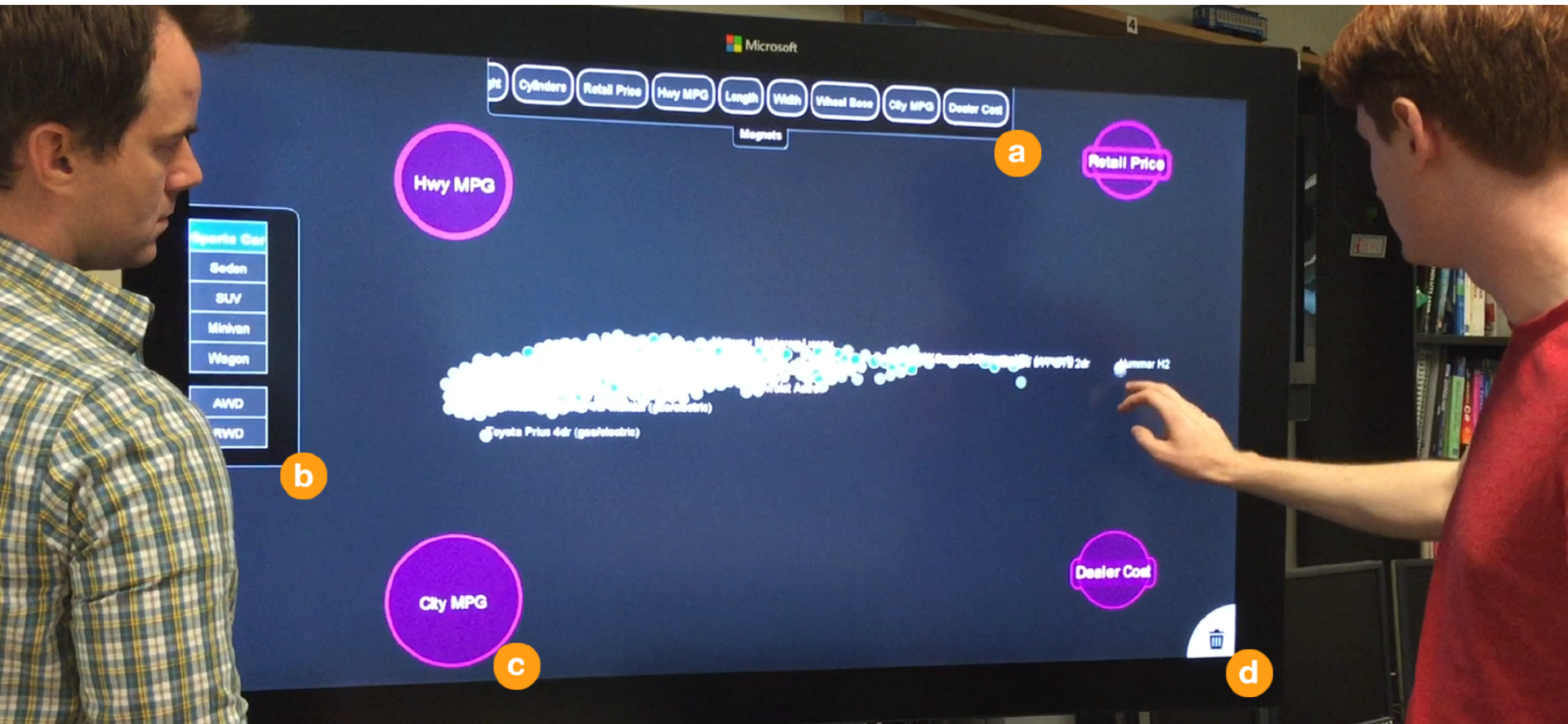


# Design Exercise – Titanic Data



- Other ways to visualize the Titanic Data
- Work with project team
- Sketch idea(s)
  - Aka crowdsourcing 😊

# Dust & Magnets



# Dust & Magnets Video



# The End







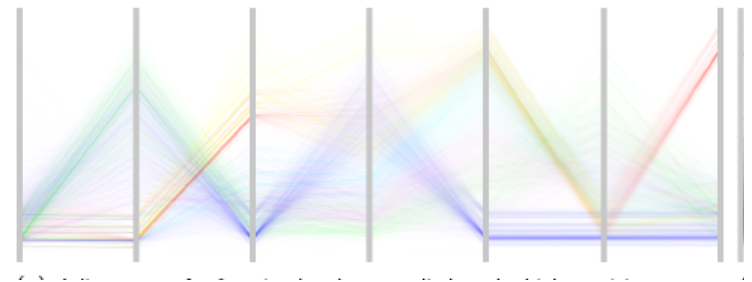
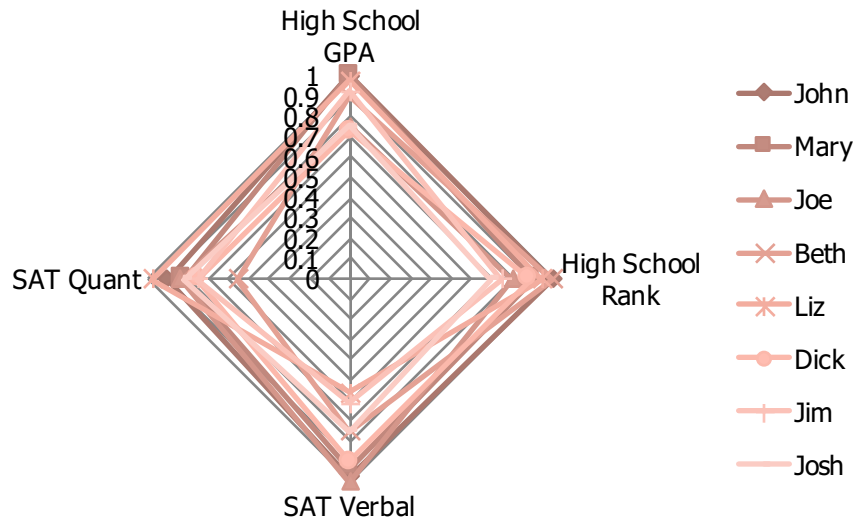




# In-class Discussion



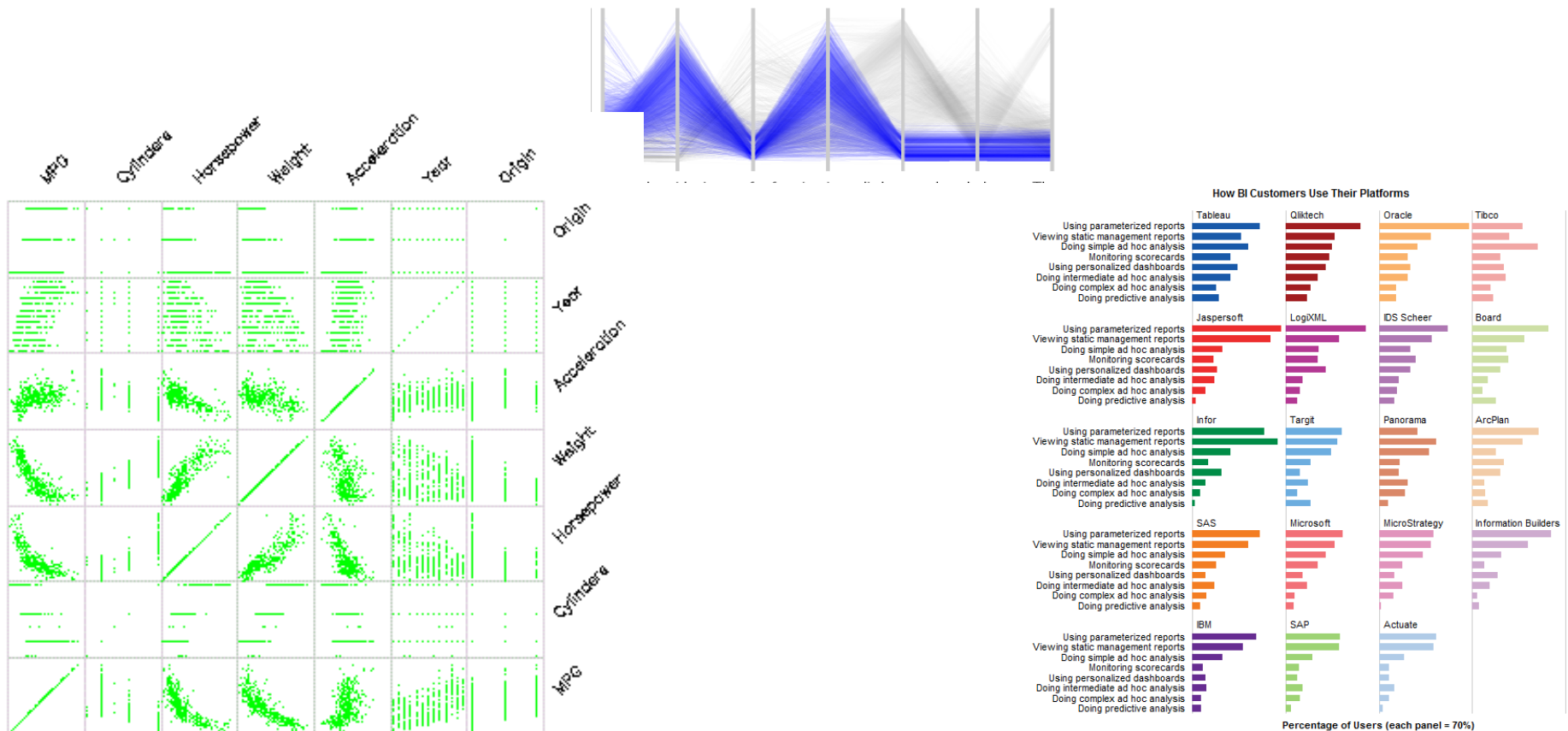
- Pros/cons of star plots and parallel coordinates; after all, they show the same information



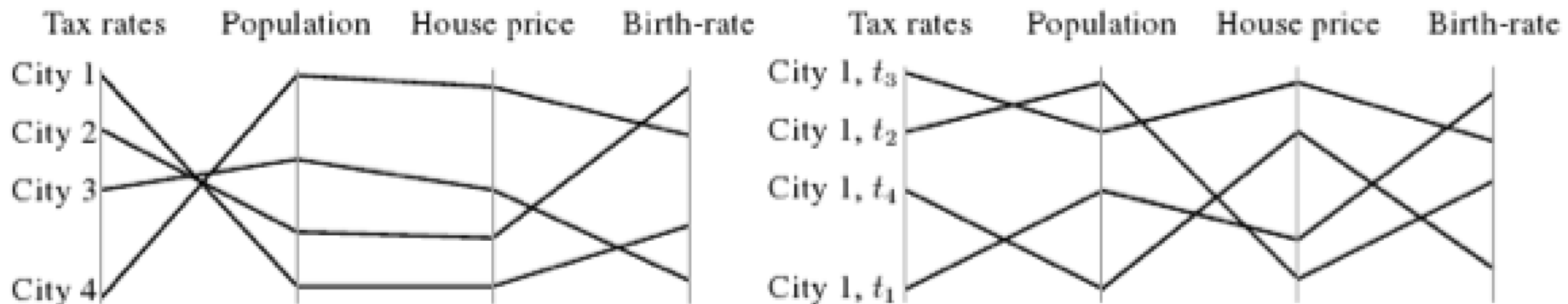
# In-class discussion



- Pros/cons of  $n \times n$  grid of scatter plots vs.  $n$  parallel coordinates vs. lots of bar charts



# Time in Parallel Coords



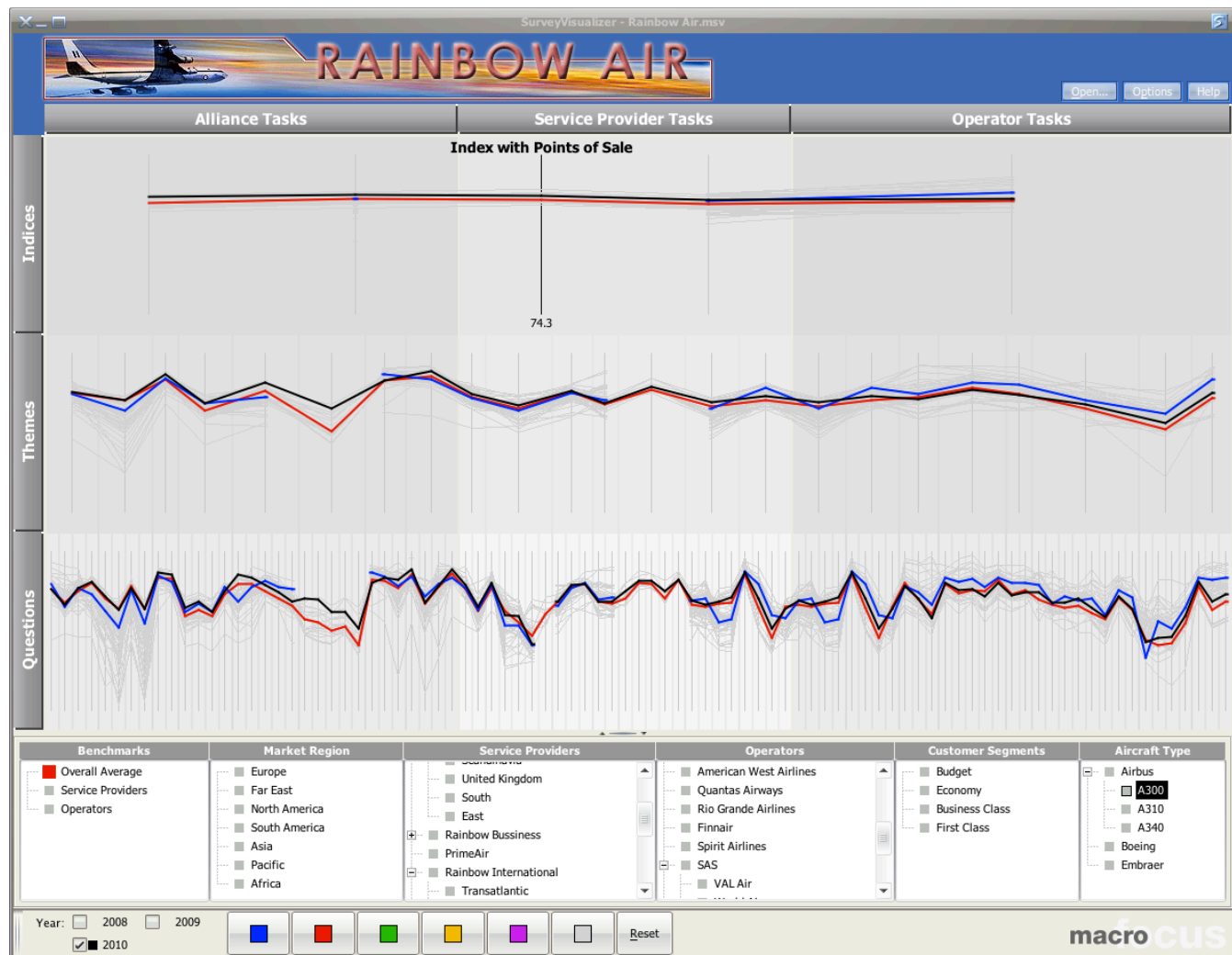
For how many  
values of time will  
this work?

# Survey Visualizer



Free download from <http://www.macrofocus.com/public/products/surveyvisualizer/>

Live Demo

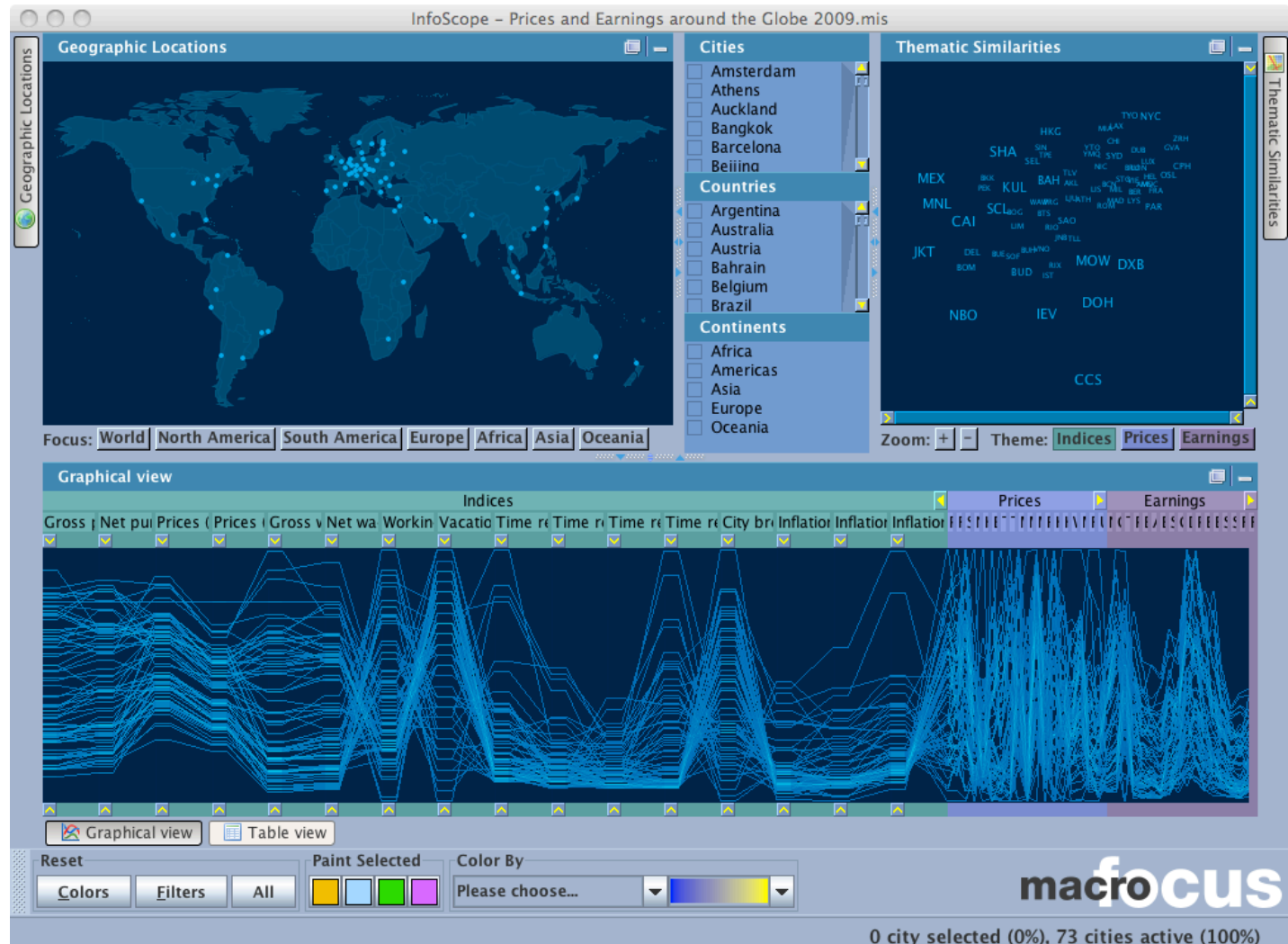


# InfoScope

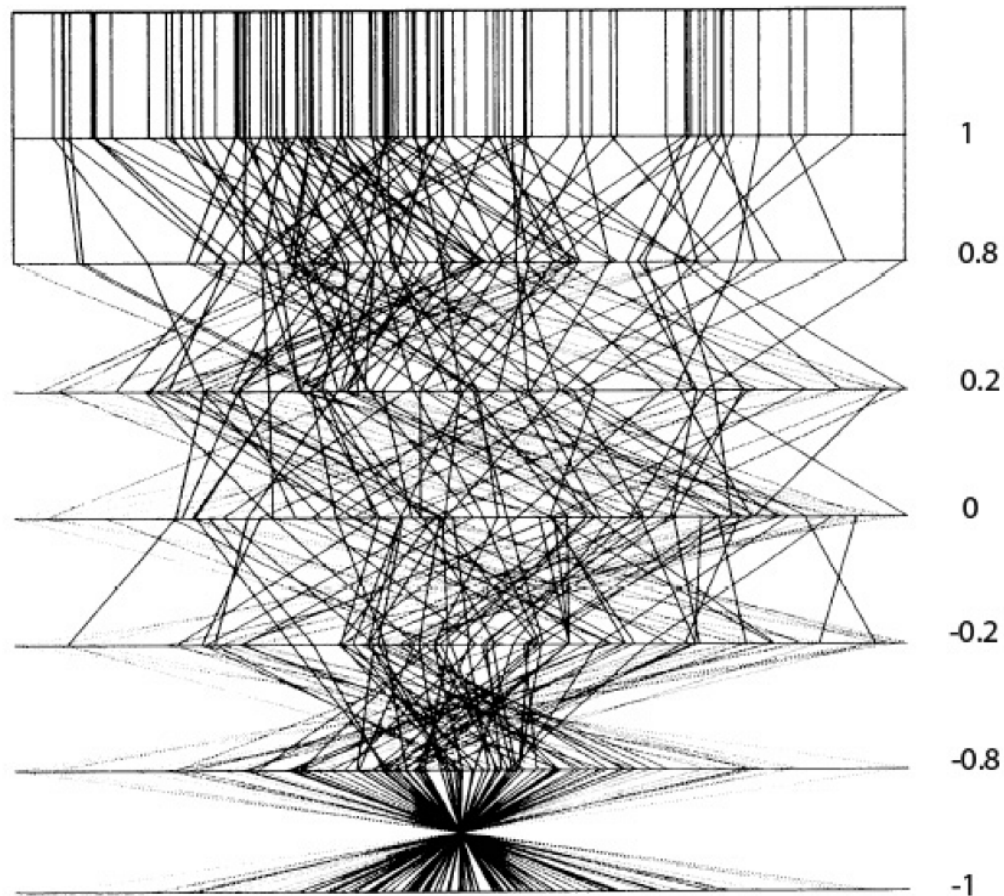


Free download from <http://www.macrofocus.com/public/products/infoscope/download/>

Live Demo



# Making Sense of Par. Coords.



From  
Munzer

CS **Figure 7.13.** Parallel coordinates were designed to show correlation between neighboring axes. At the top, parallel lines show perfect positive correlation. At the bottom, all of the lines cross over each other at a single spot in between the two axes, showing perfect negative correlation. In the middle, the mix of crossings shows uncorrelated data. From [Wegman 90, Figure 3].

# In-Class Discussion



- You have a data set about a large number of people containing
  - Number of years of education
  - Age
  - Height
  - Weight
  - Weight index (1 means ideal weight given height, greater than 1 means overweight, less than 1 means underweight)
  - Gender
- What are some interesting questions?
  - How would you visualize this data to answer?



# In-class Discussion

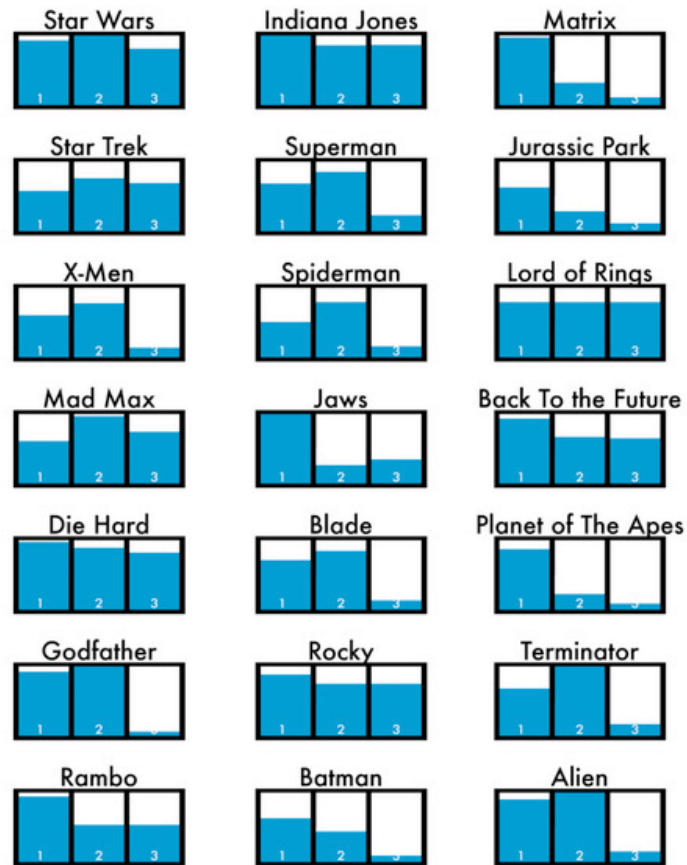


- Data set about car ownership
  - Car brand
  - Car year
  - Car weight
  - Owner age
  - Owner gender
- What are some interesting questions?
  - How would you visualize this data to answer?
  - How might you encode all this information into a single InfoViz?

# Small Multiples: Bar Chart



## THE TRILOGY METER



#1 In A Series of Pop Cultural Charts

DANMETH.COM

<http://danmeth.com/post/77471620/my-trilogy-meter-1-in-a-series-of-pop-cultural>