

# VISUALIZING MULTIVARIATE DATA

1. INTRODUCTION

2. VISUAL ENCODING

3. BASIC CHART TYPES

Basics

4. INTERACTION

5. VISUALIZATION DESIGN

6. DATA PREPROCESSING

7. RECAP 1<sup>st</sup> Half

8. MULTIVARIATE DATAVIS

9. TEMPORAL DATAVIS

10. GEOSPATIAL DATAVIS

11. GRAPH DATAVIS

12. TEXT DATAVIS

13. VISUAL ANALYTICS

14. RECAP 2<sup>nd</sup> Half

Visualization  
Building Blocks  
& Processes

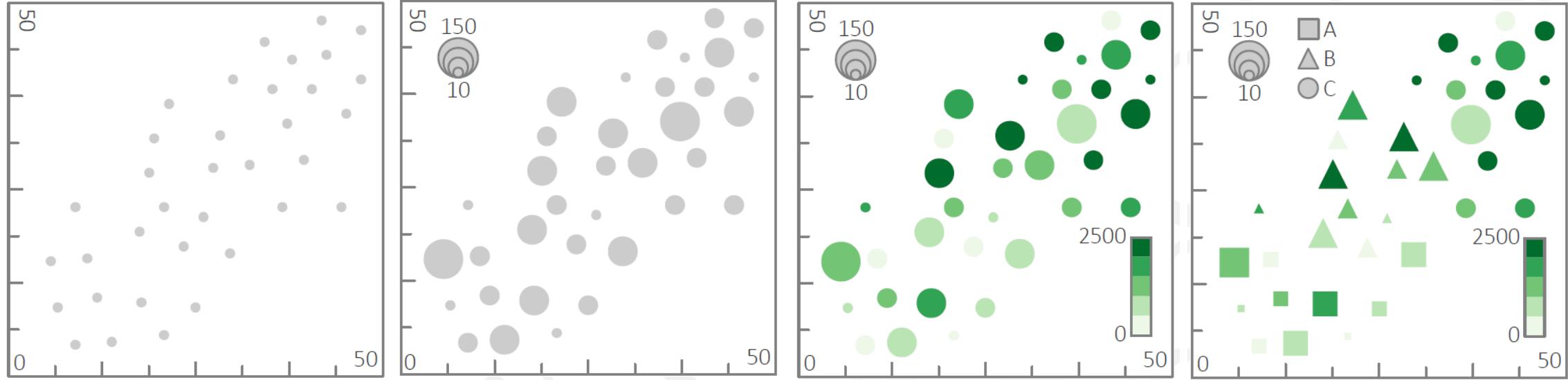
Visualization  
Techniques

Visualization  
in Practice



# WHY TALK ABOUT MULTIVARIATE VIS?

---



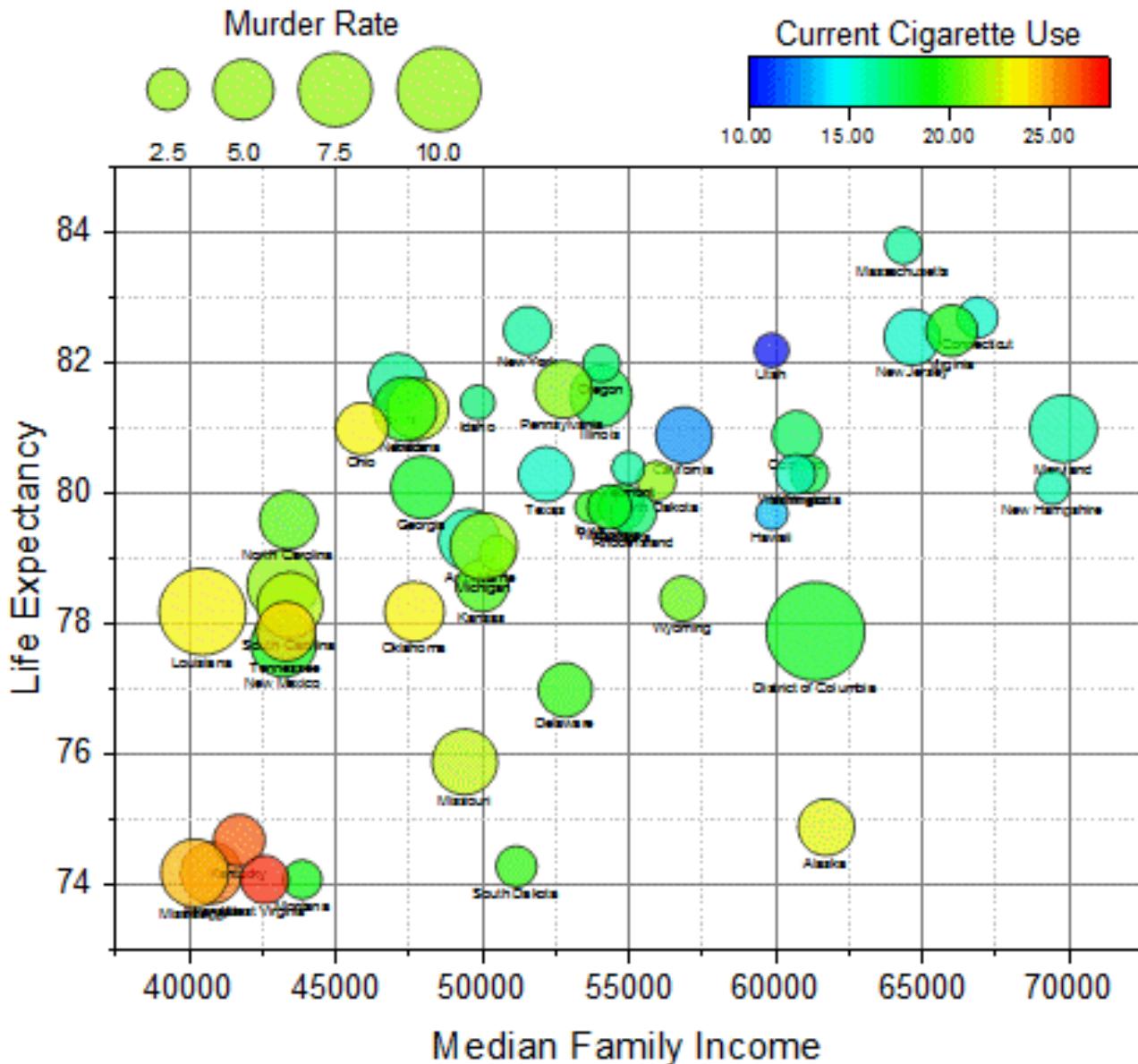
Mapping two attributes  
(Position)

Mapping three attributes  
(Position, Area)

Mapping four Attributes  
(Position, Area, Color)

Mapping five Attributes  
(Position, Area, Color, Shape)

This does not scale!



[OriginLab Documentation]

<https://www.originlab.com/www/products/GraphGallery.aspx?GID=416>



# OVERVIEW

---

## Using Points

- Ternary Plots
- Scatterplot Matrices (SPLOMs)
- RadViz

## Using (Poly)Lines

- Parallel Coordinates (PCs)
- Parallel Sets
- Radar Charts

## Using Nesting / Hierarchies

- Mosaic Plots
- Trellis Displays

## Others

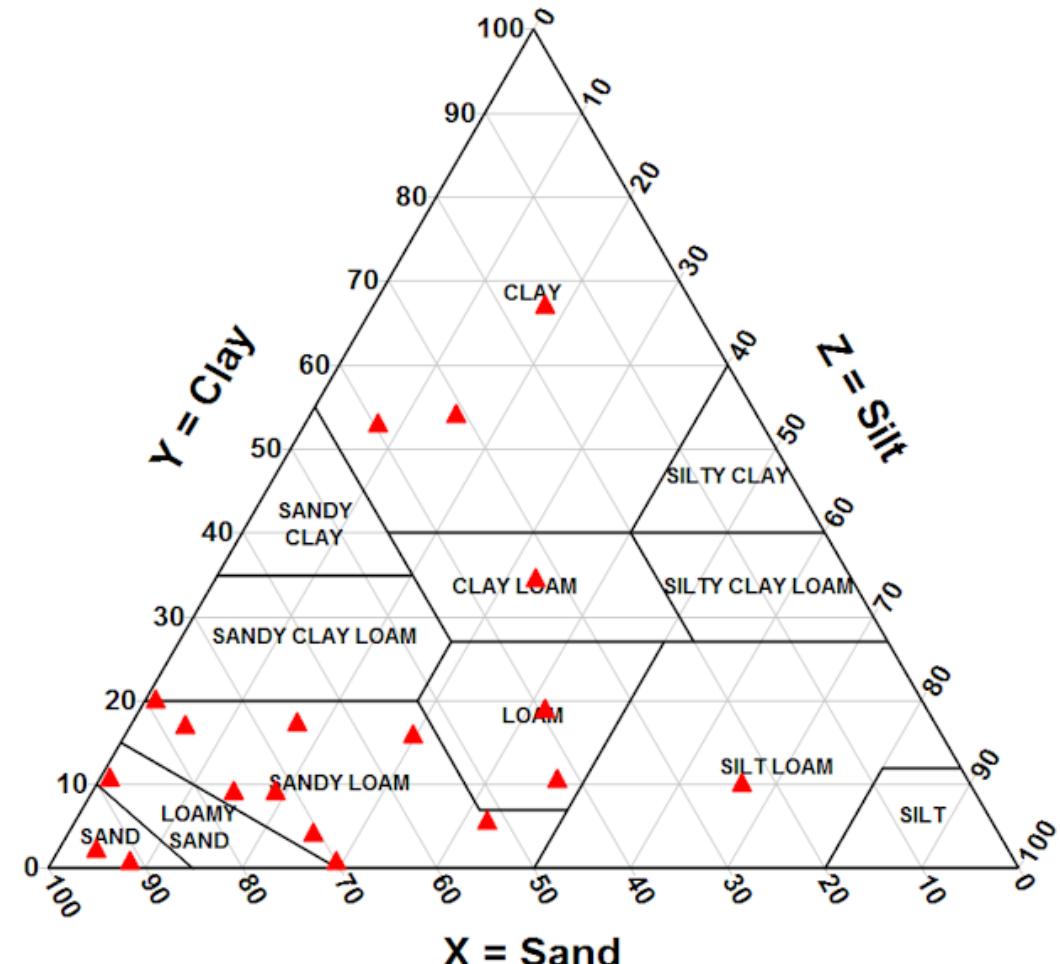
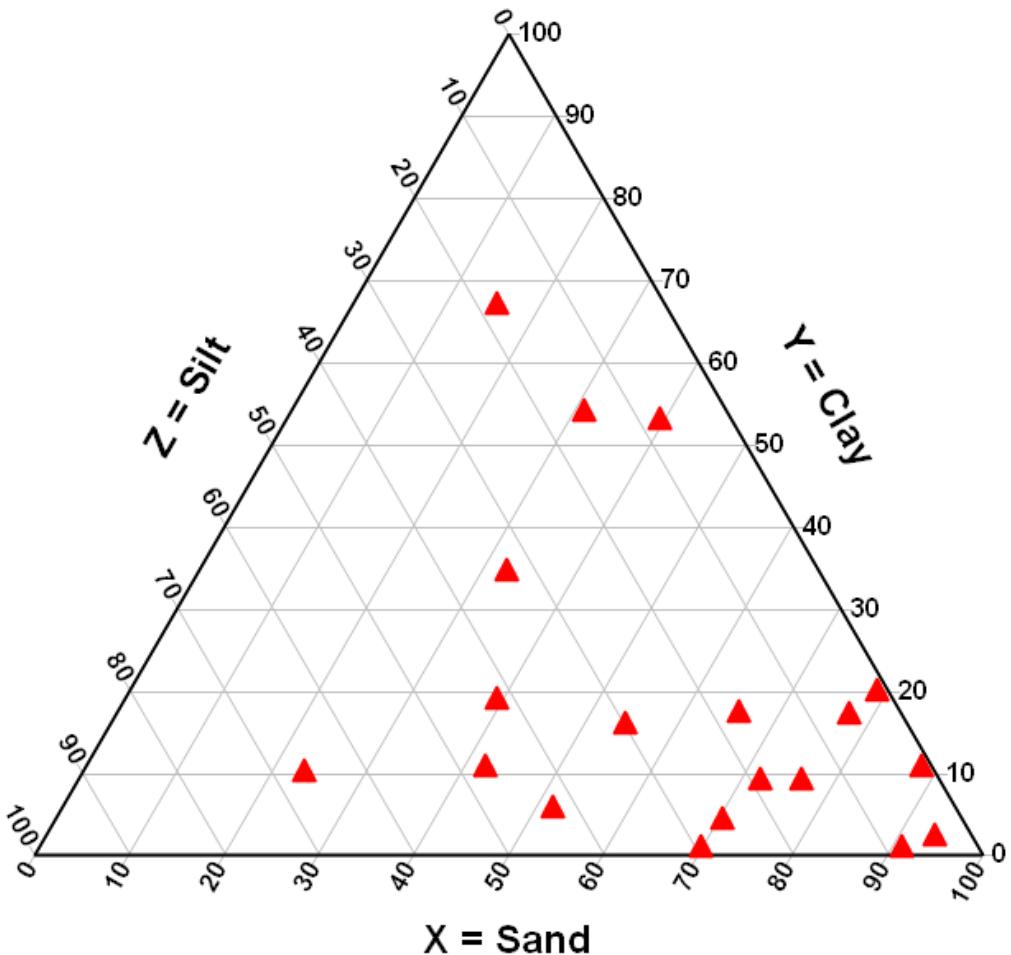
- Glyphs
- Pixel-based Visualizations

# USING POINTS

# TERNARY PLOTS

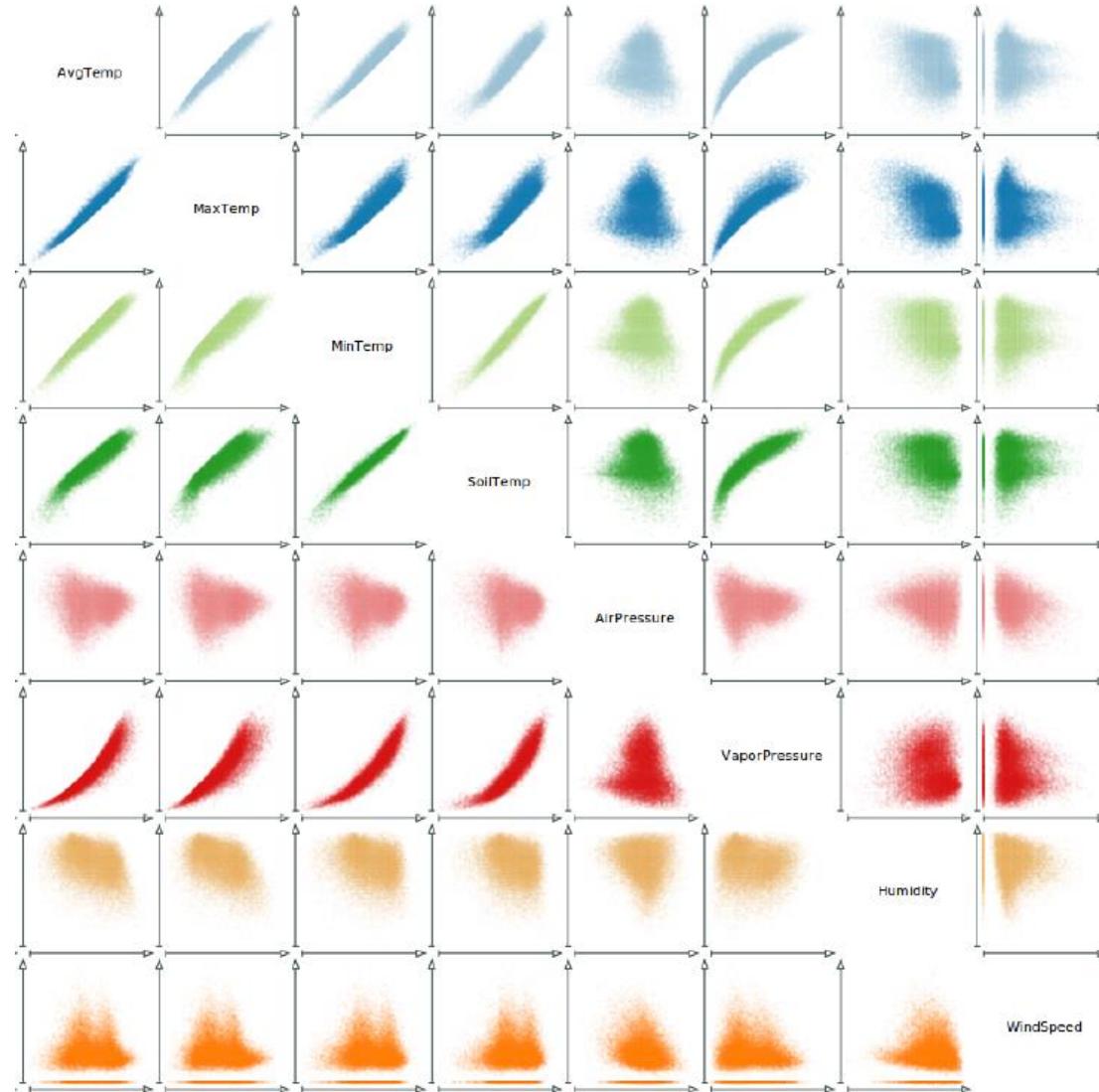
[ [https://en.wikipedia.org/wiki/Ternary\\_plot](https://en.wikipedia.org/wiki/Ternary_plot) ]

= Scatterplots for 3 variables that add up to a given total (e.g., 1 or 100%)



# SCATTERPLOT MATRIX

---

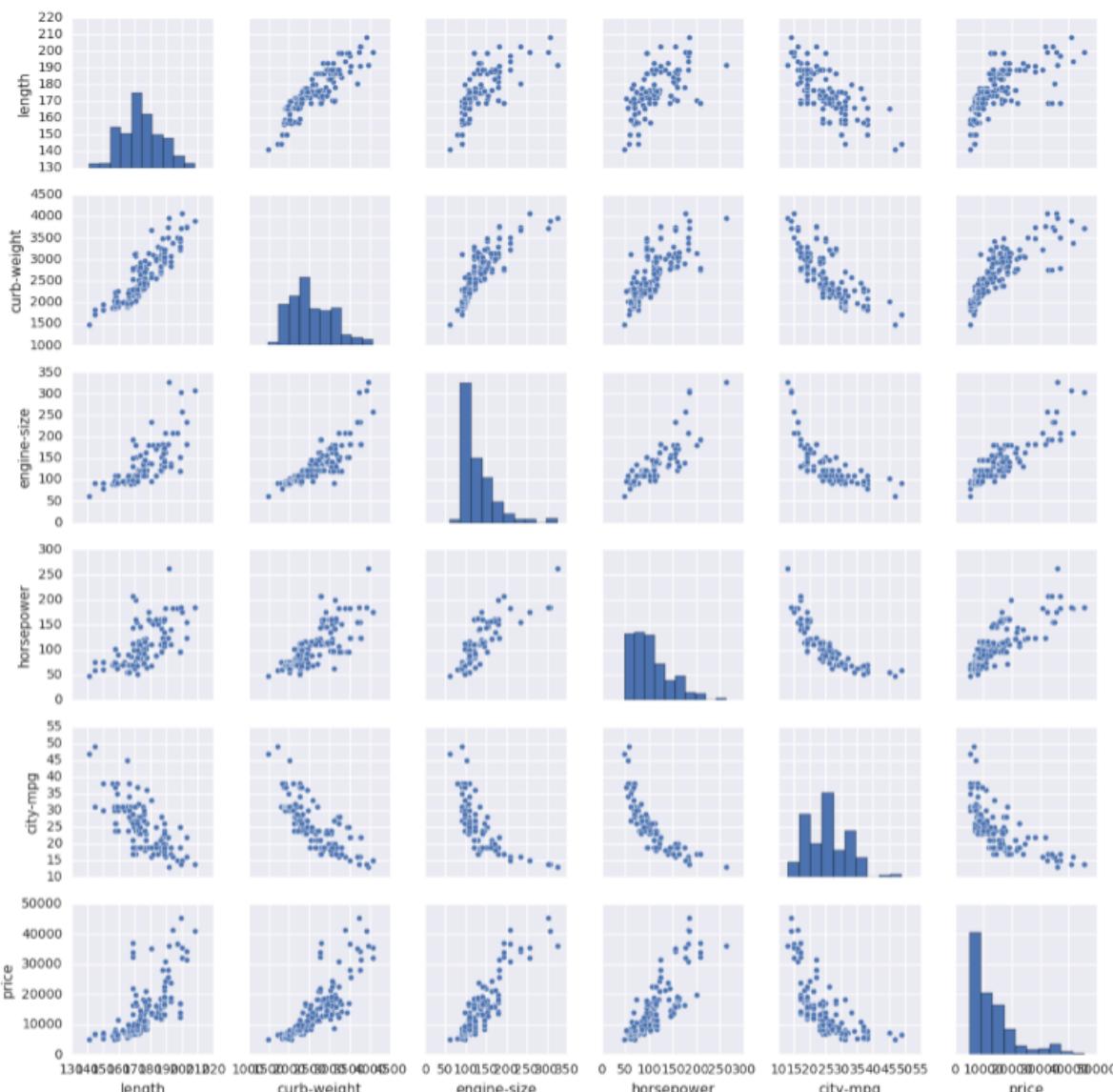


**Scatterplot Matrix:**  
One Scatterplot per variable  
combination in a matrix layout



# SCATTERPLOT MATRIX W/ HISTOGRAMS

---

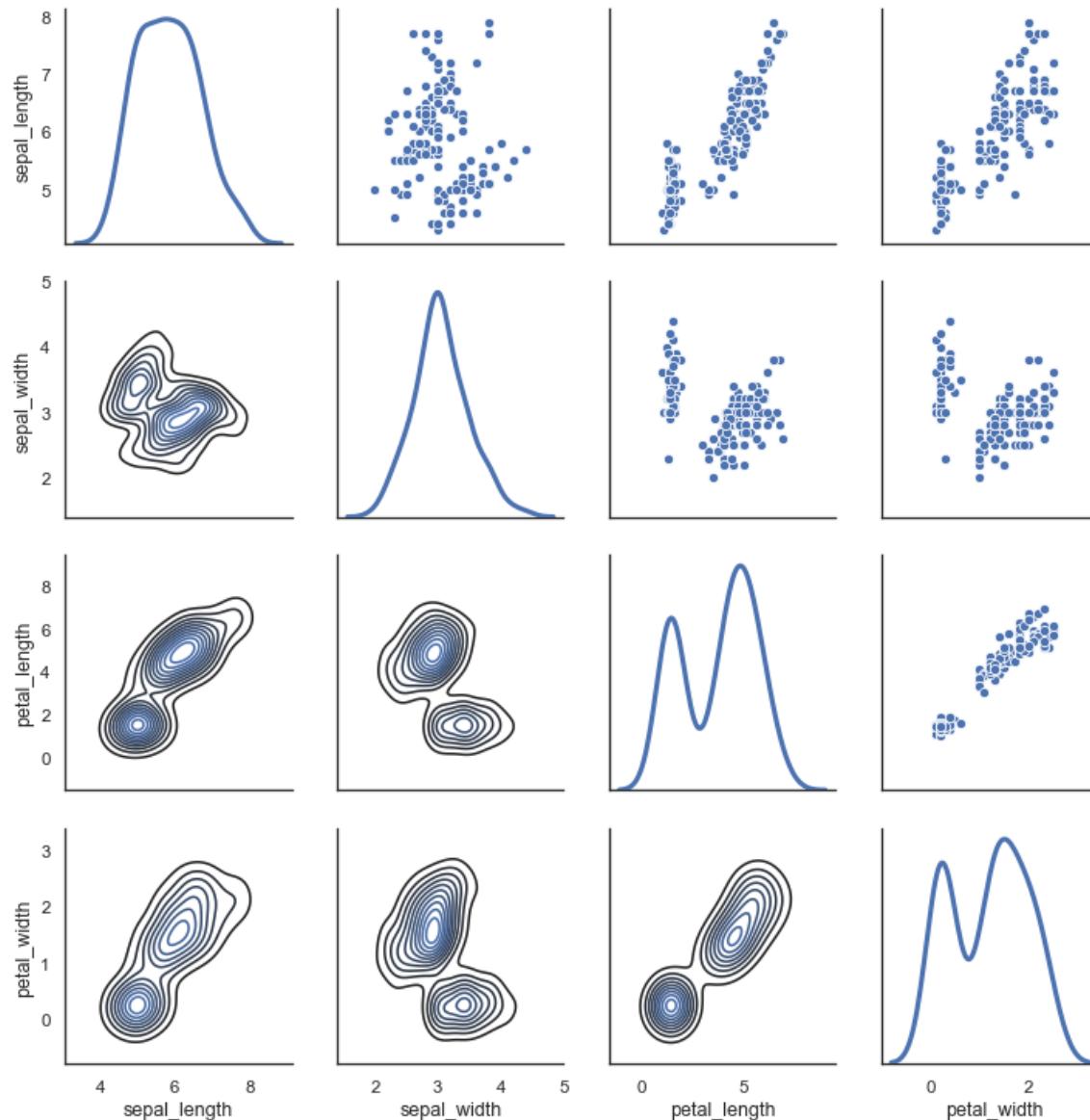


**Scatterplot Matrix /w Histograms:**  
Embed Histograms showing the value distributions for each attribute in the otherwise unused main diagonal of the scatterplot matrix



# PAIRED SPLOM / DENSITY CONTOURS

---



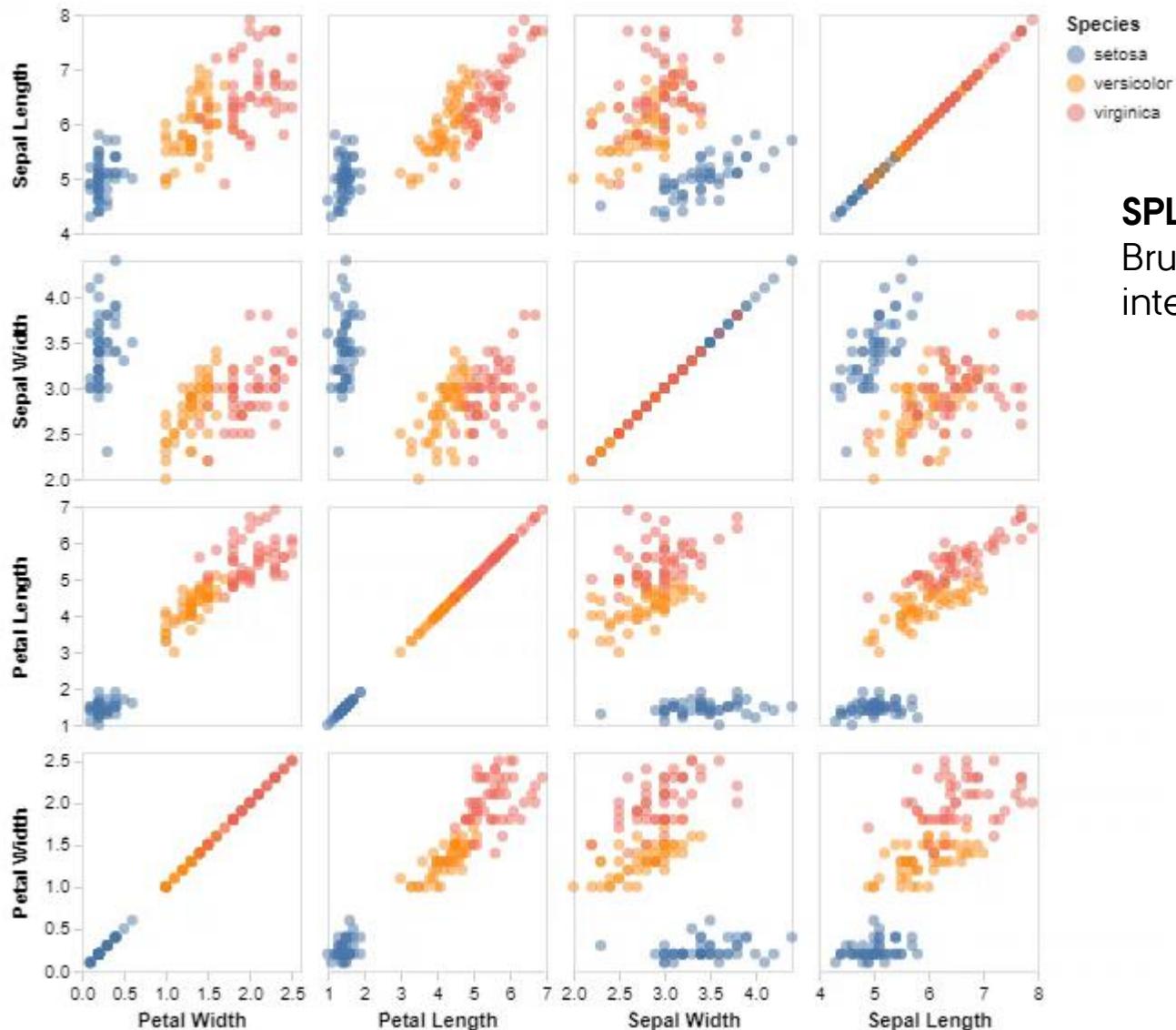
## SPLOM w/ Density Contours:

Use mirrored half of the SPLOM to embed other plots – e.g., density contours to highlight areas of potential overplotting in the SPLOM



# INTERACTION: LINKING & BRUSHING

---



**SPLOM w/ Linking & Brushing:**  
Brushing & Linking are essential  
interactions in a SPLOM



# Rolling the Dice

Multidimensional Visual Exploration  
using Scatterplot Matrix Navigation

Niklas Elmquist  
Pierre Dragicevic  
Jean-Daniel Fekete

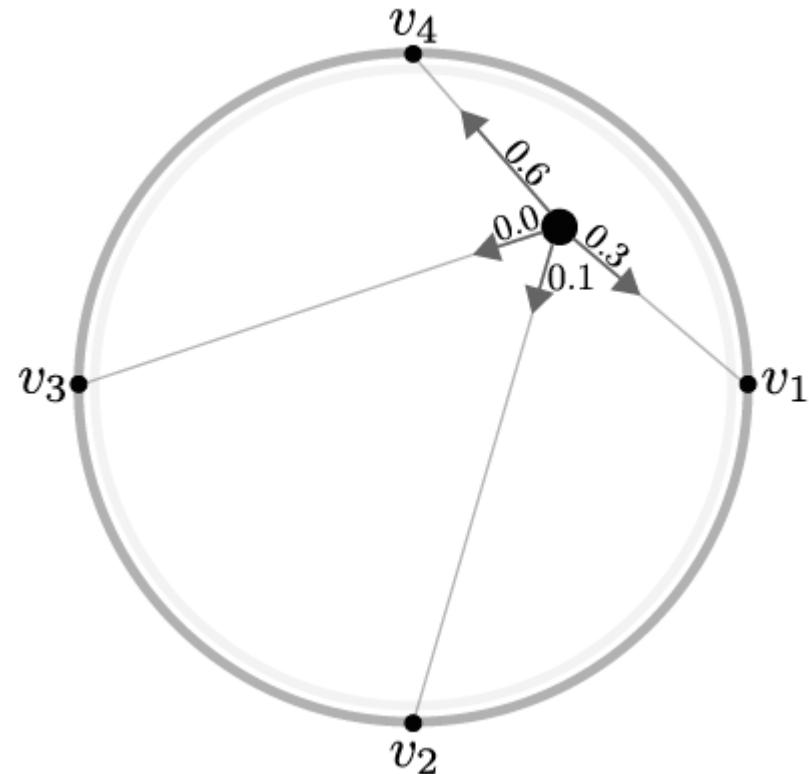
INRIA

# RADVIZ

[Hoffman et al. 1997]

In RadViz, each attribute is assigned an anchor on the circumference of a circular drawing area.

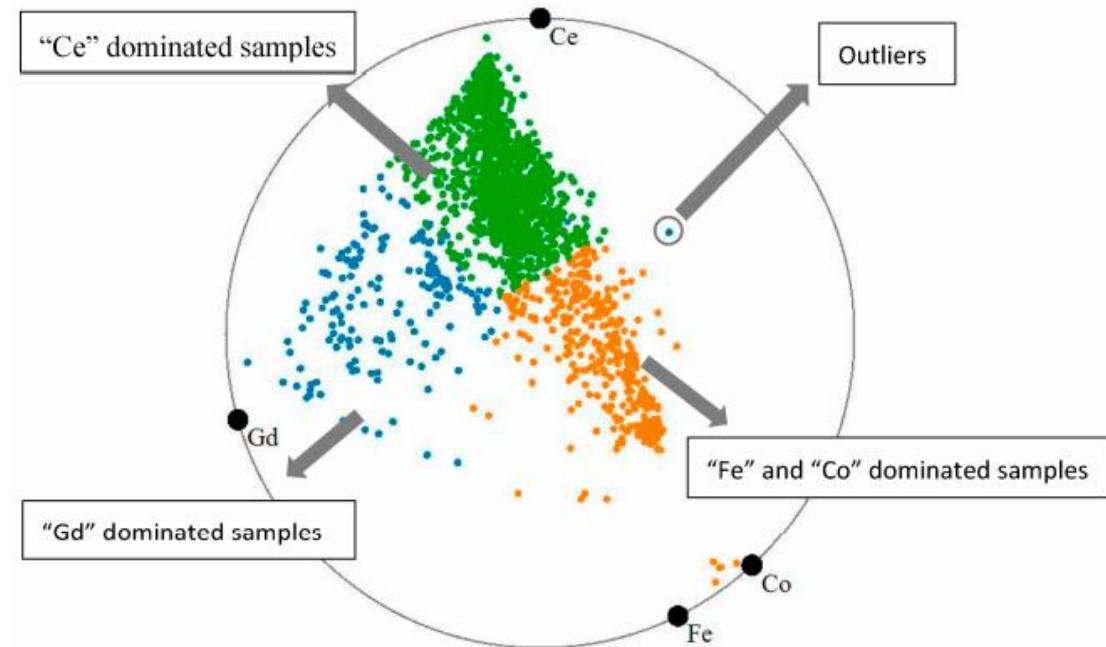
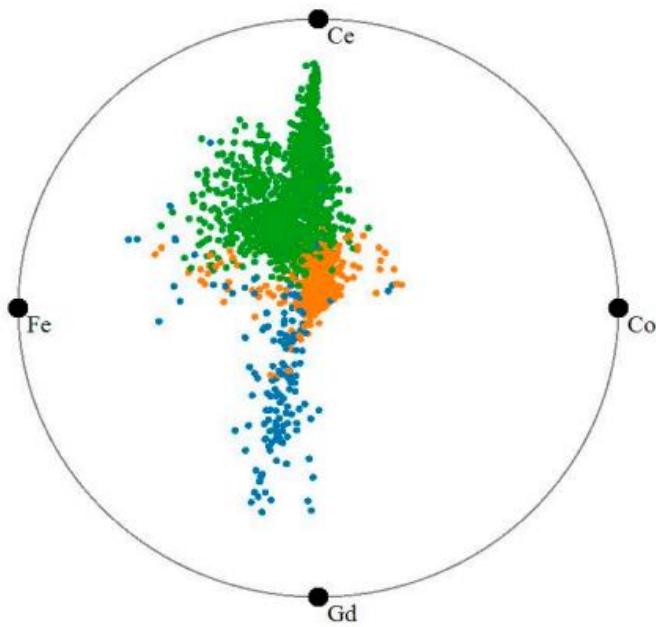
Each data point is “pulled towards” the different attribute anchors with a force proportional to the respective value.  
(Barycentric Coordinates)



# RADVIZ DELUXE

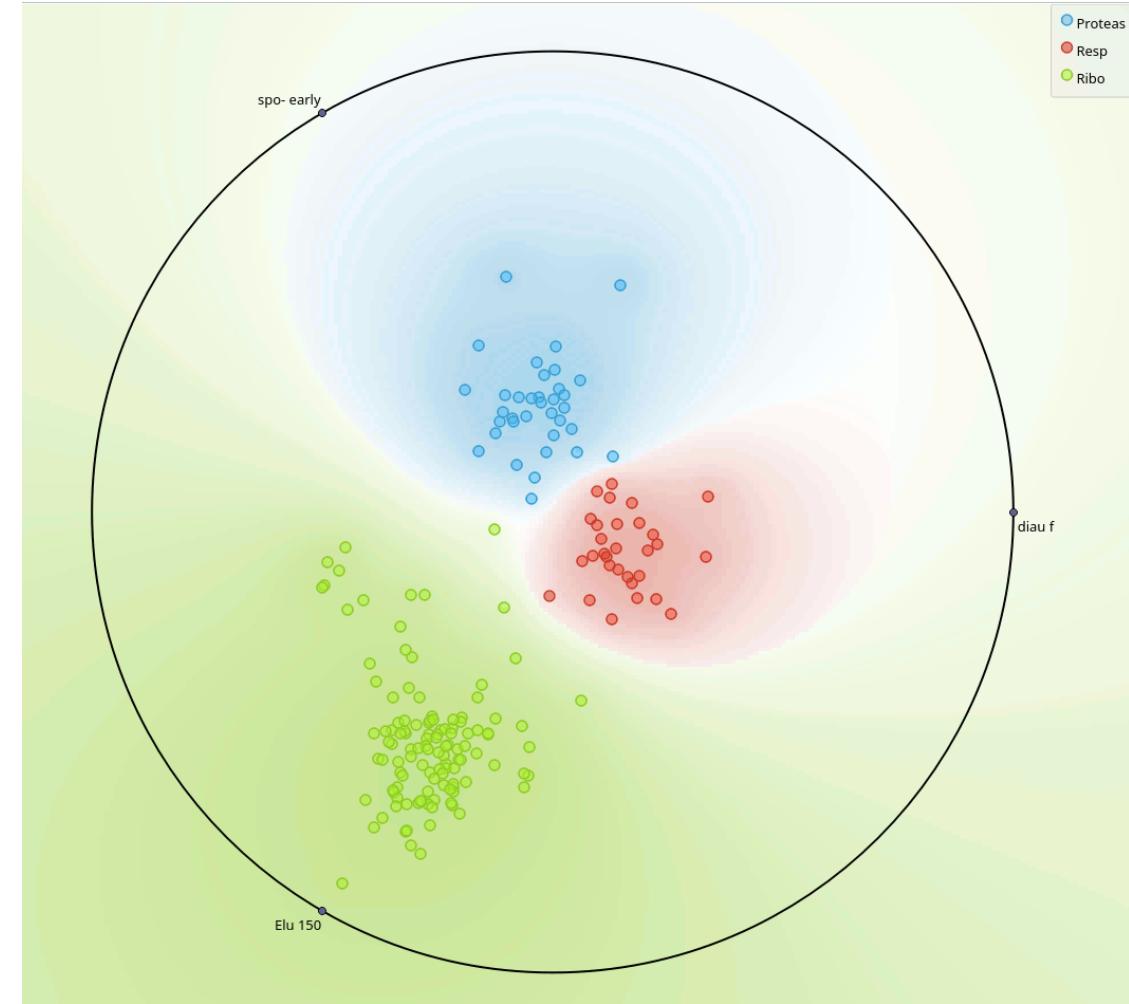
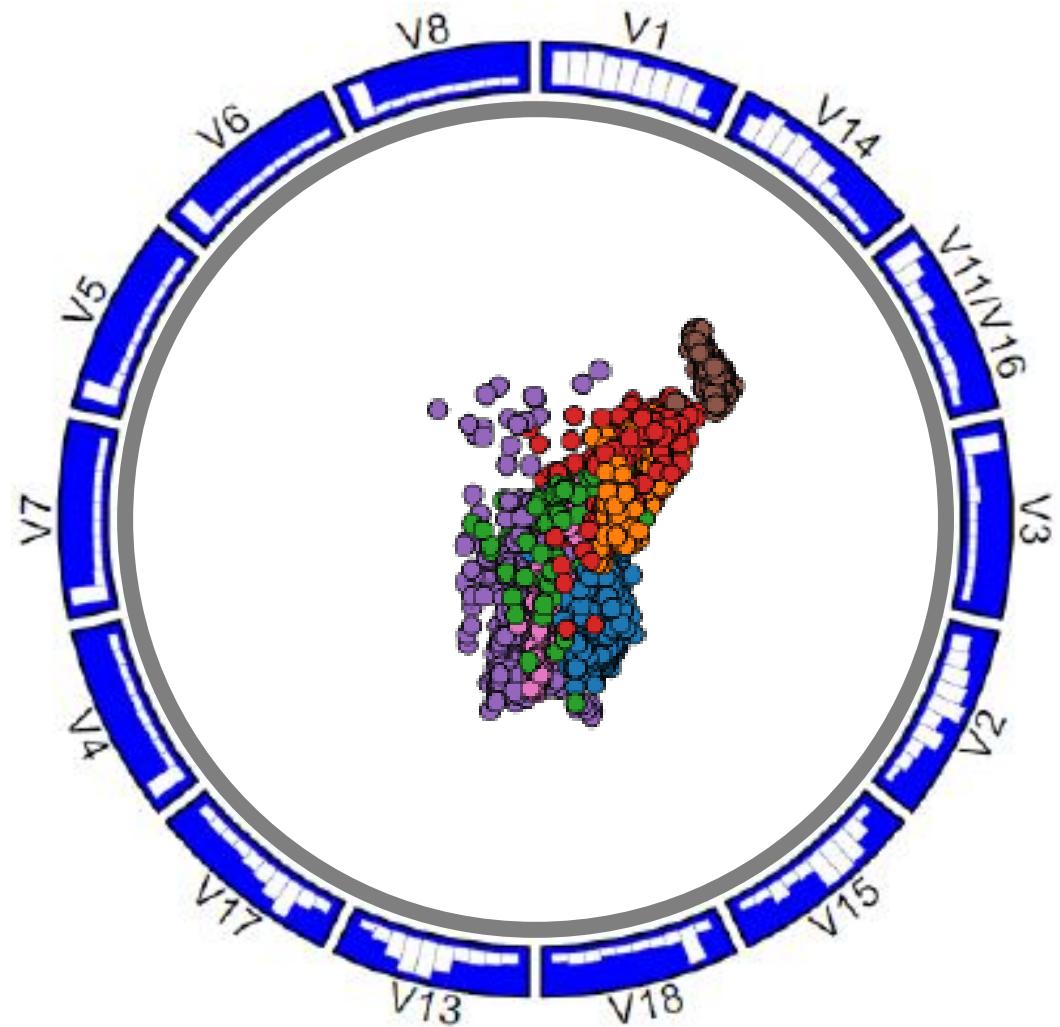
[Cheng et al. 2017]

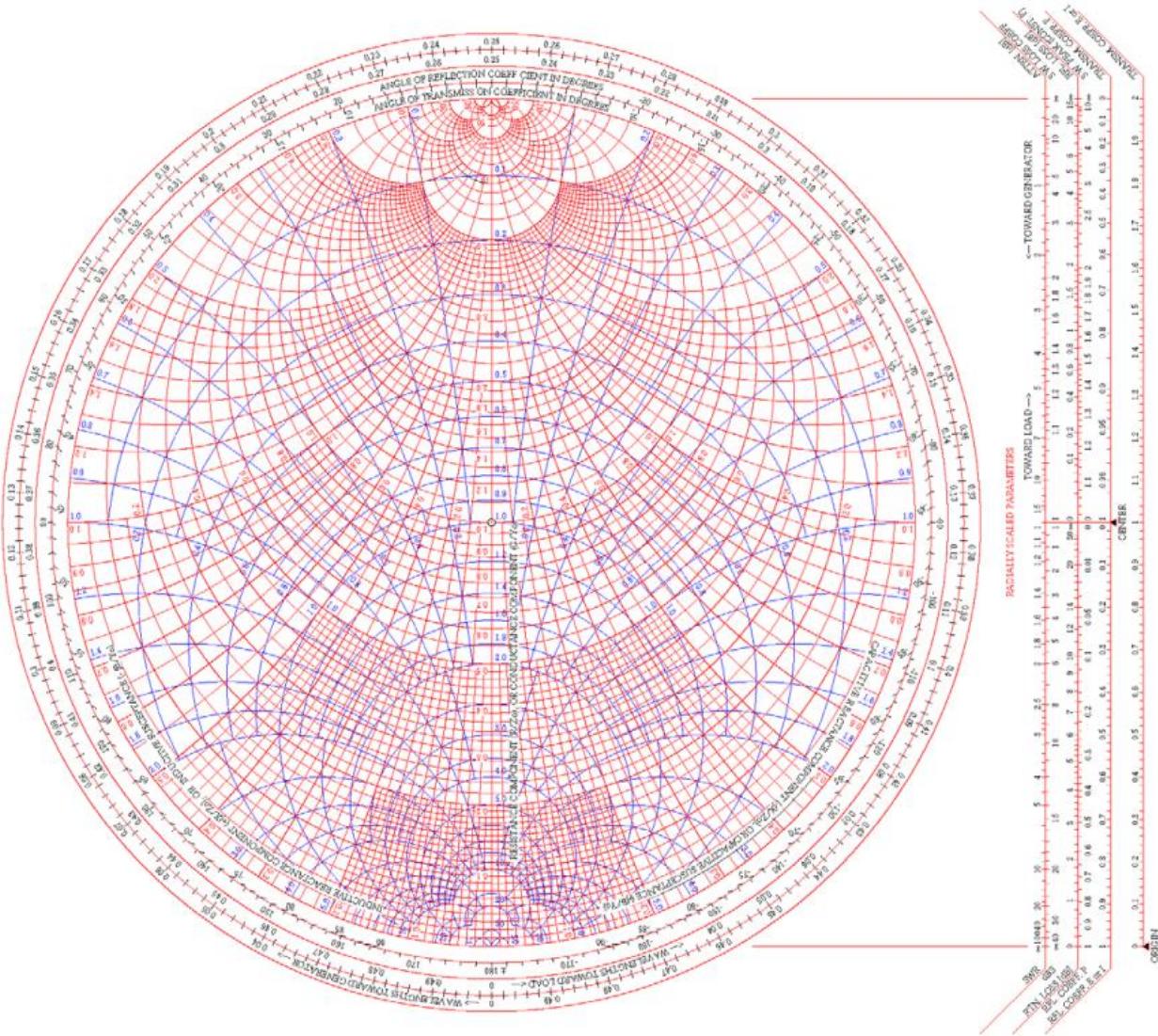
RadViz Deluxe corrects for some of the errors in RadViz – e.g., through a better placement of the anchors that reflects their pairwise correlations:



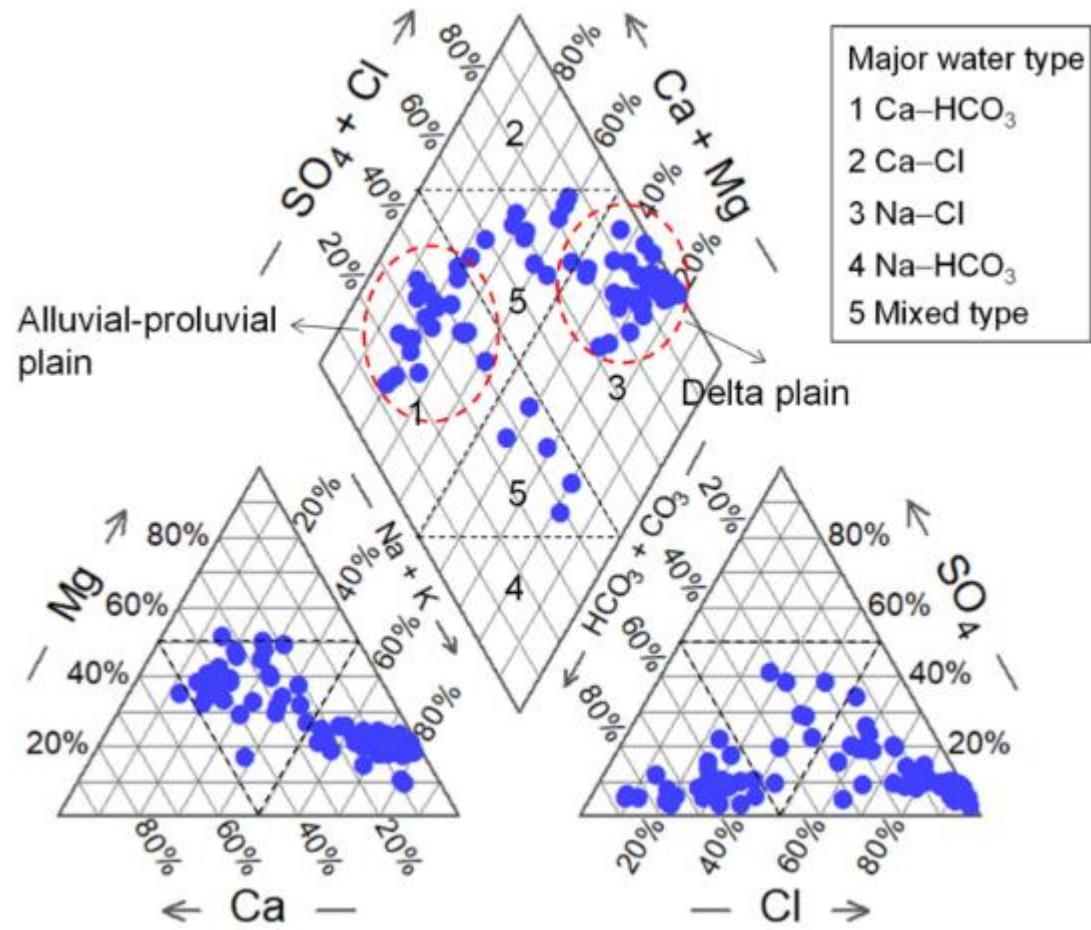
# RADVIZ W/ HISTOGRAMS OR DENSITIES

---





[ [https://en.wikipedia.org/wiki/Smith\\_chart](https://en.wikipedia.org/wiki/Smith_chart) ]



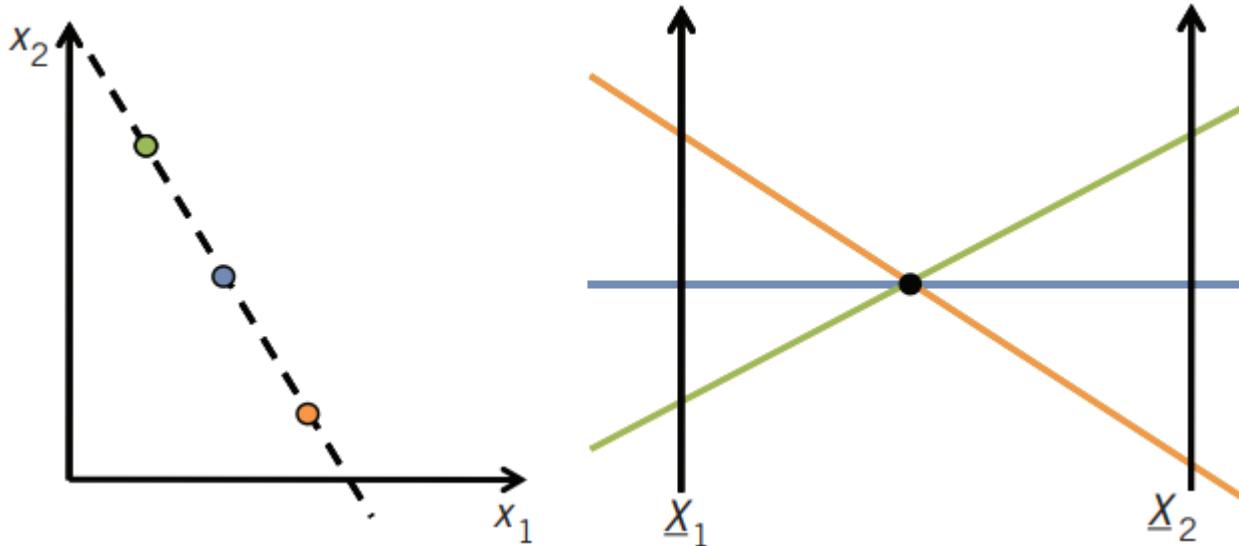
[ [https://en.wikipedia.org/wiki/Piper\\_diagram](https://en.wikipedia.org/wiki/Piper_diagram) ]

# USING LINES

# FROM CARTESIAN COORDINATES TO PCS

---

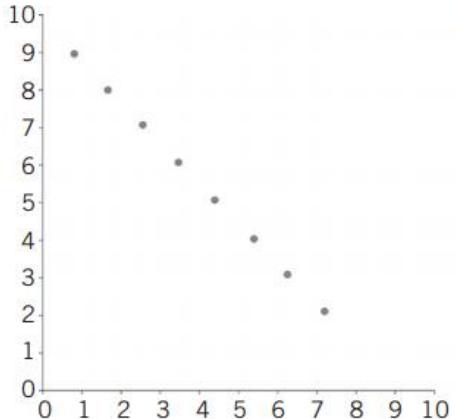
[Inselberg 2009]



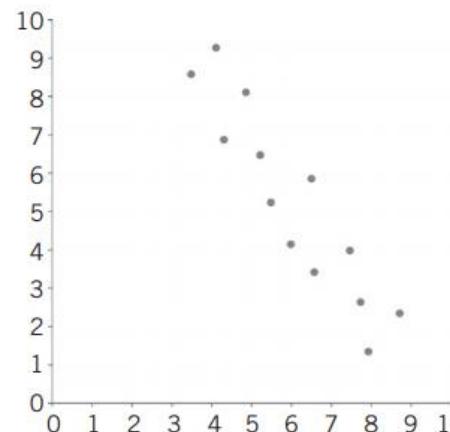
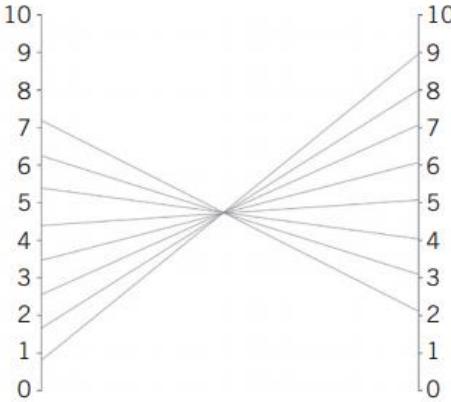
Each point in Cartesian Coordinates becomes a line in Parallel Coordinates

# PARALLEL COORDINATE PATTERNS

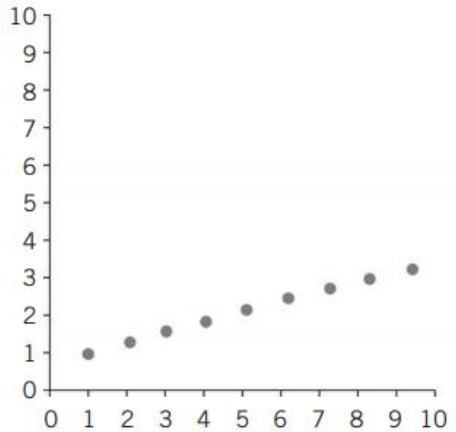
---



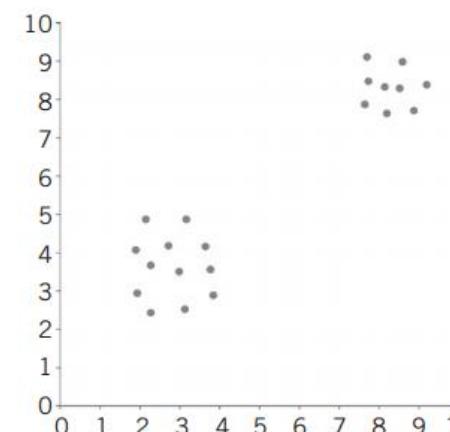
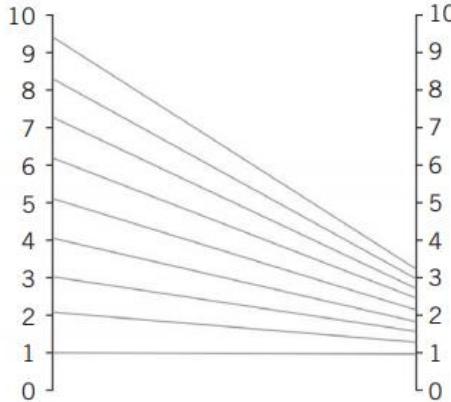
Inversely Correlated



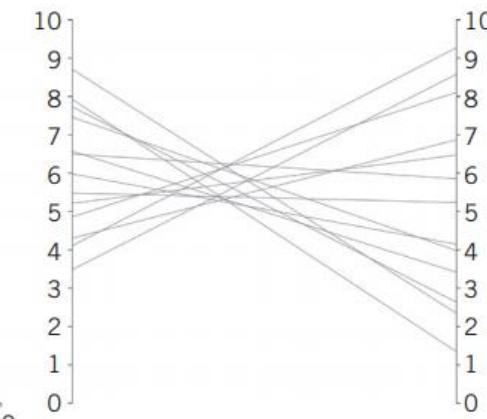
Inversely Correlated



Correlated

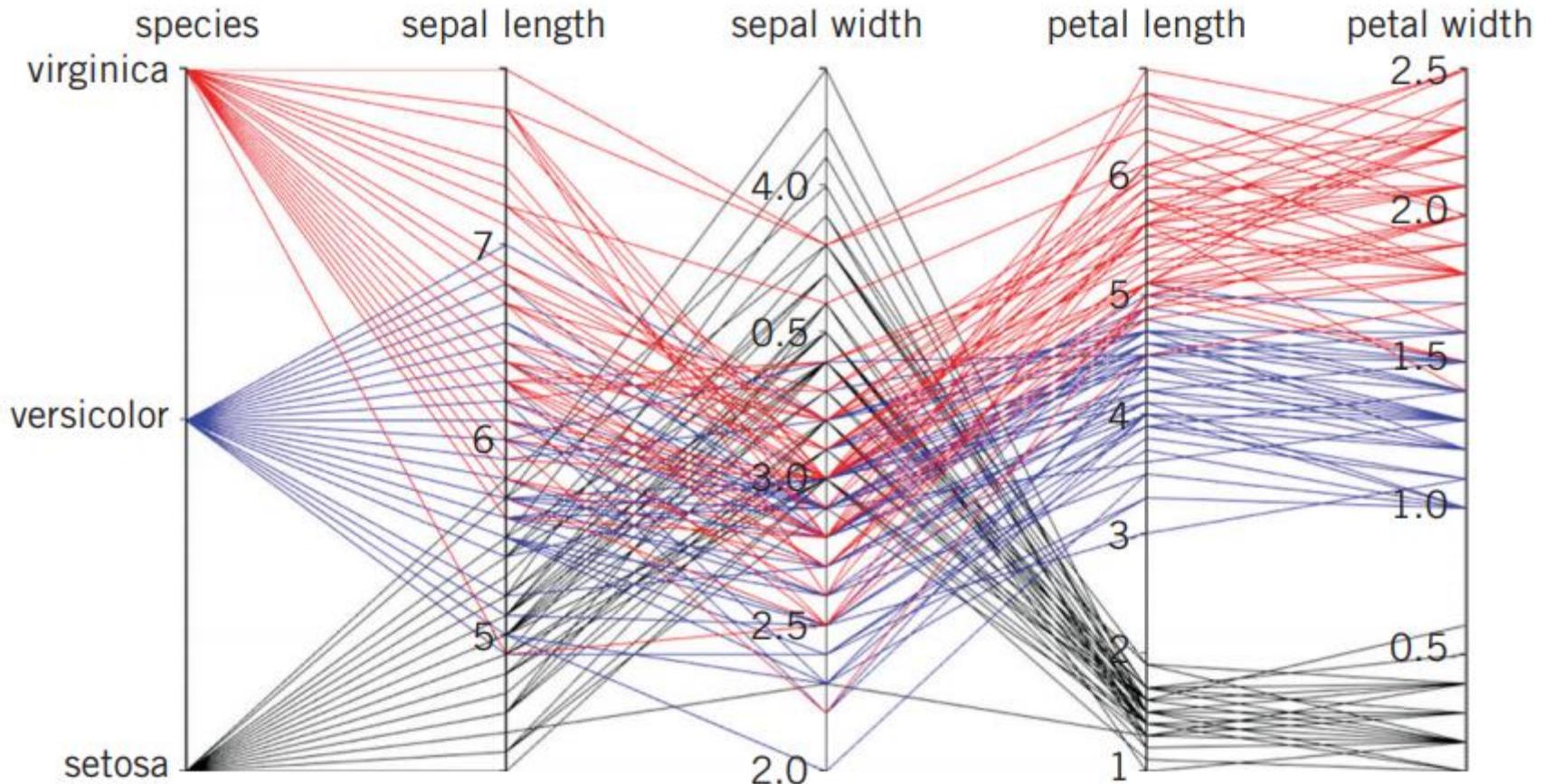


Clustered



# FROM LINES TO POLYLINES

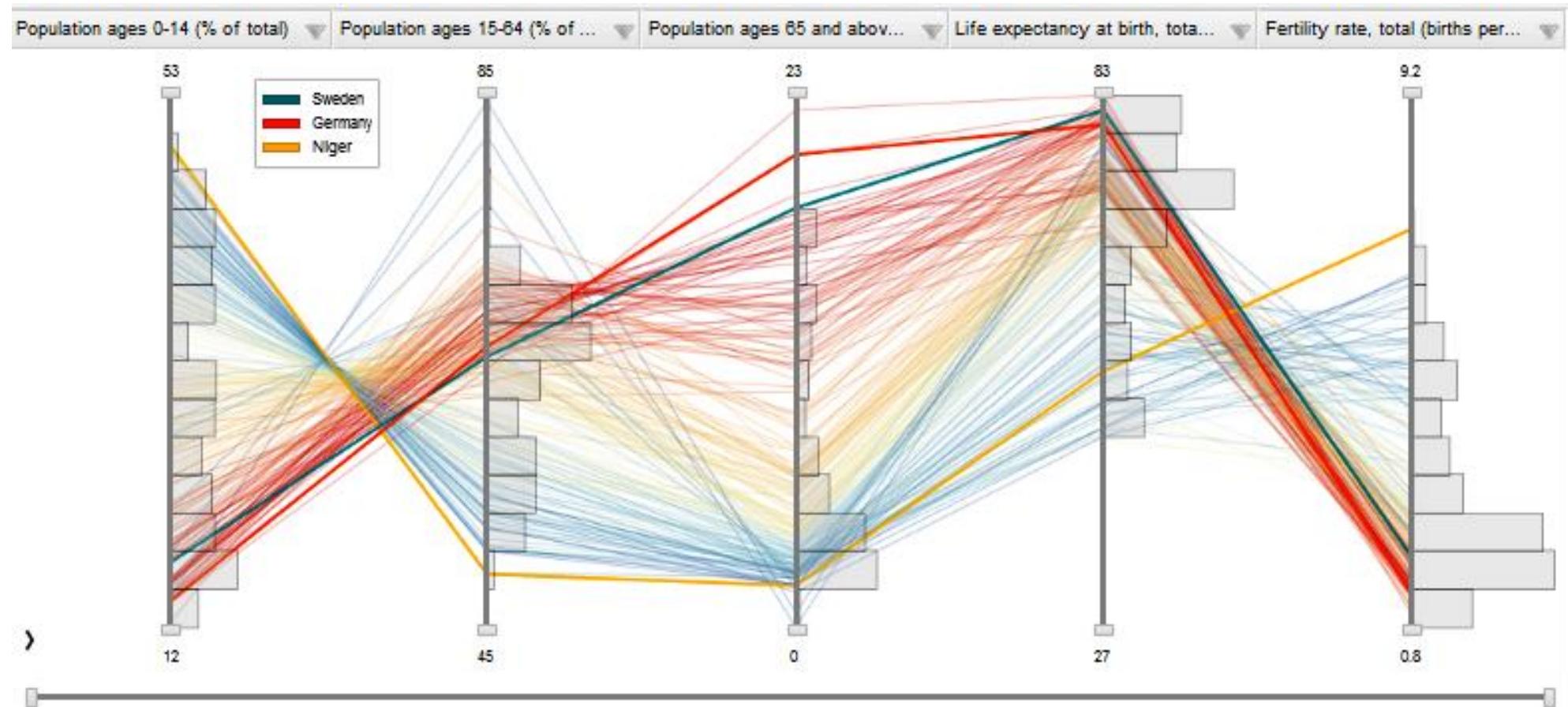
---



# PCS /W HISTOGRAMS

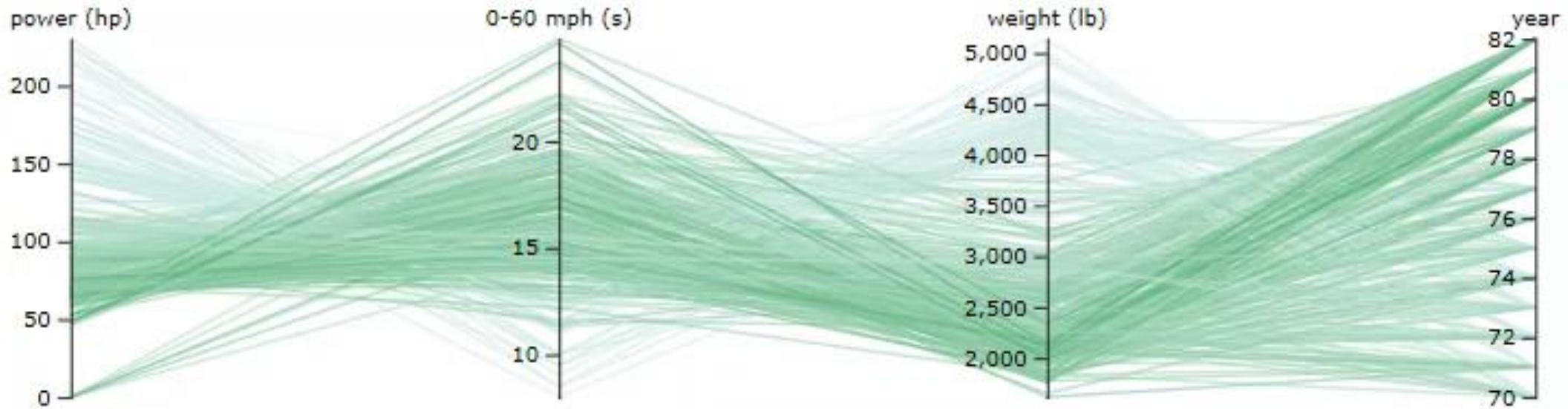
---

**Parallel Coordinates /w Histograms:**  
Embed Histograms showing the value distributions for each attribute along the different axes



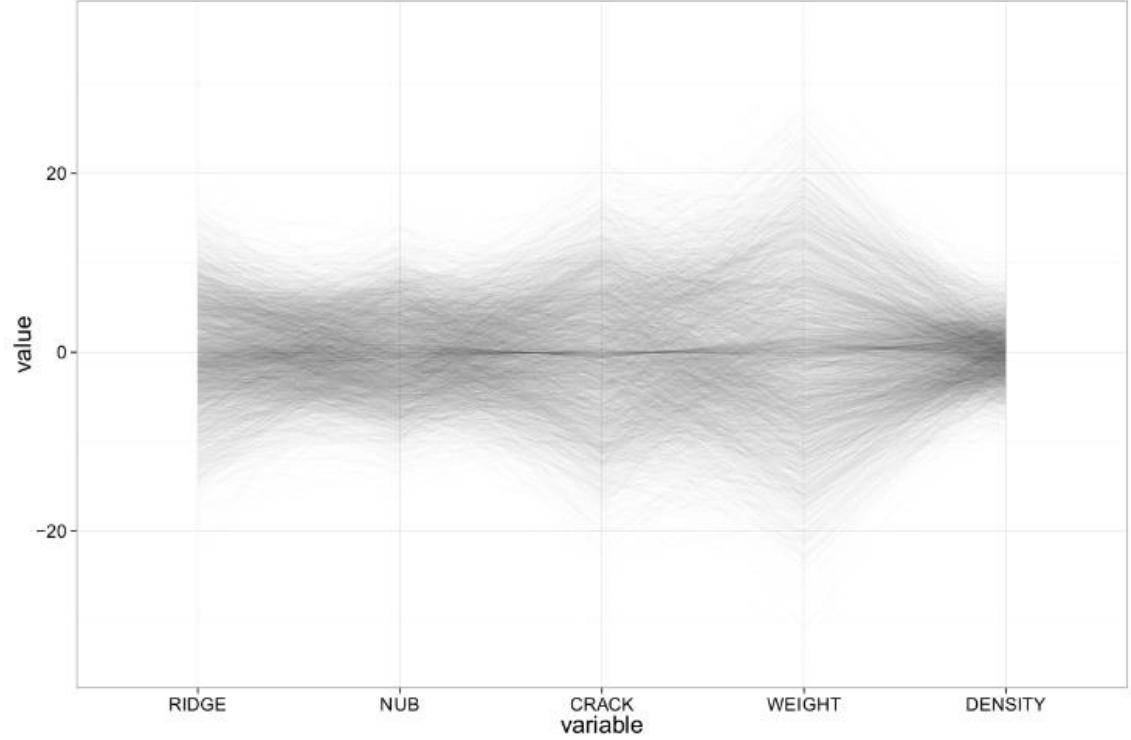
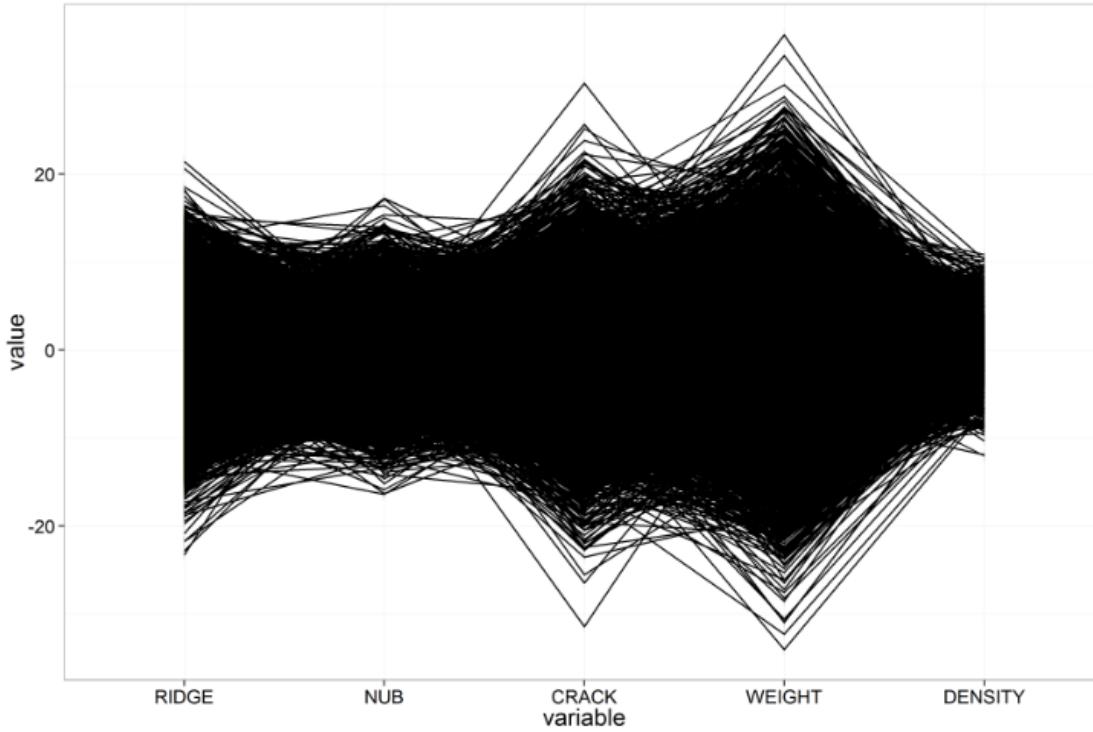
# INTERACTION: REORDERING, BRUSHING

---



# REDUCING THE VISUAL CLUTTER

Alpha Blending

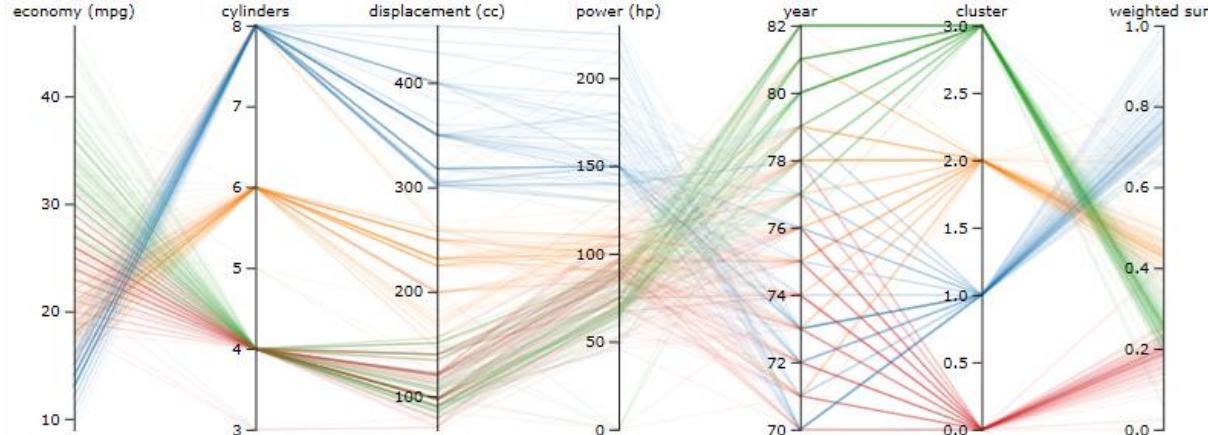


<https://markusthill.github.io/visualizing-high-dimensional-data-with-parallel-coordinates/>

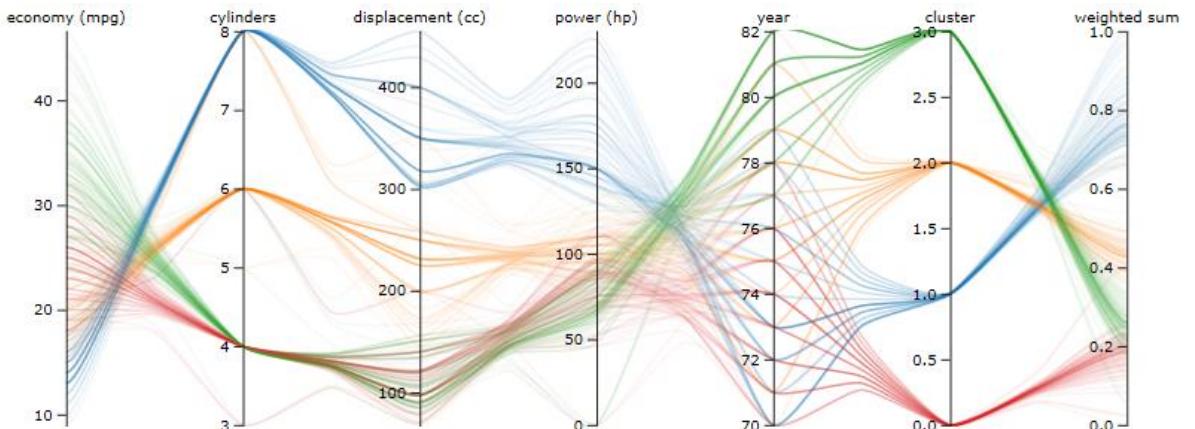
# REDUCING THE VISUAL CLUTTER

---

[Raseman et al. 2019]  
<https://parasoljs.github.io>



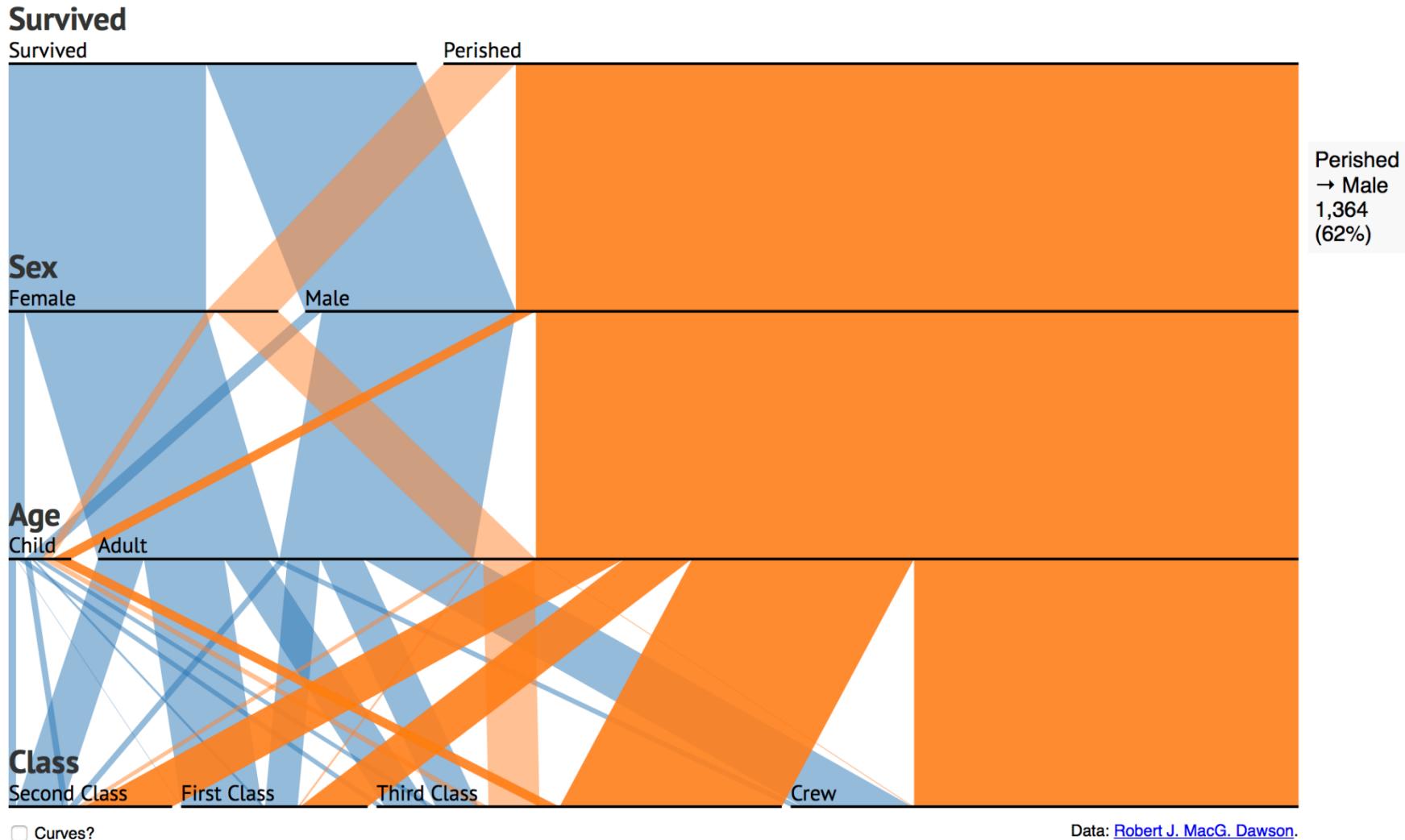
Bundling



# PARALLEL SETS

[Bendix et al. 2005]

## Titanic Survivors



AARHUS  
UNIVERSITY

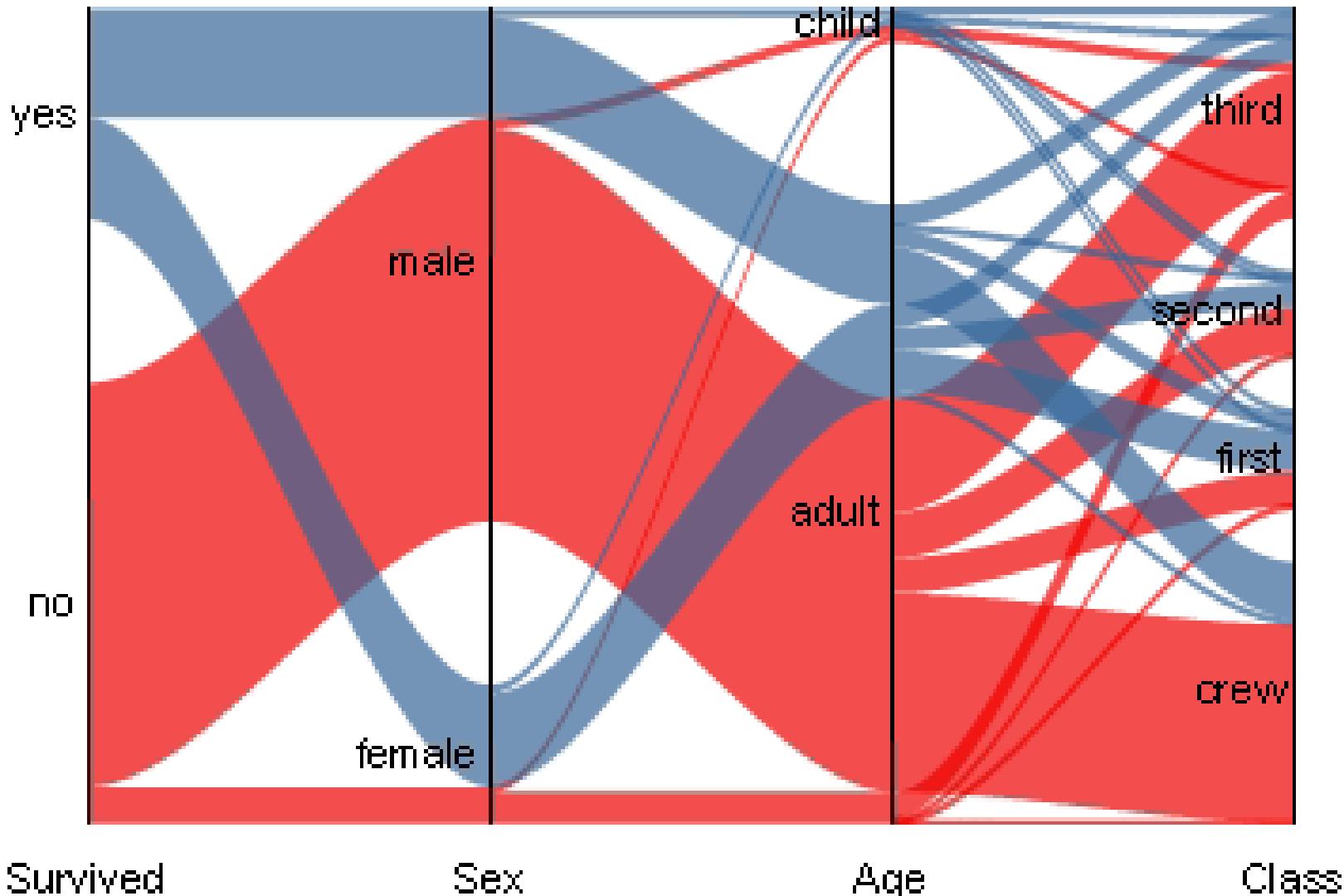
DEPARTMENT OF COMPUTER SCIENCE



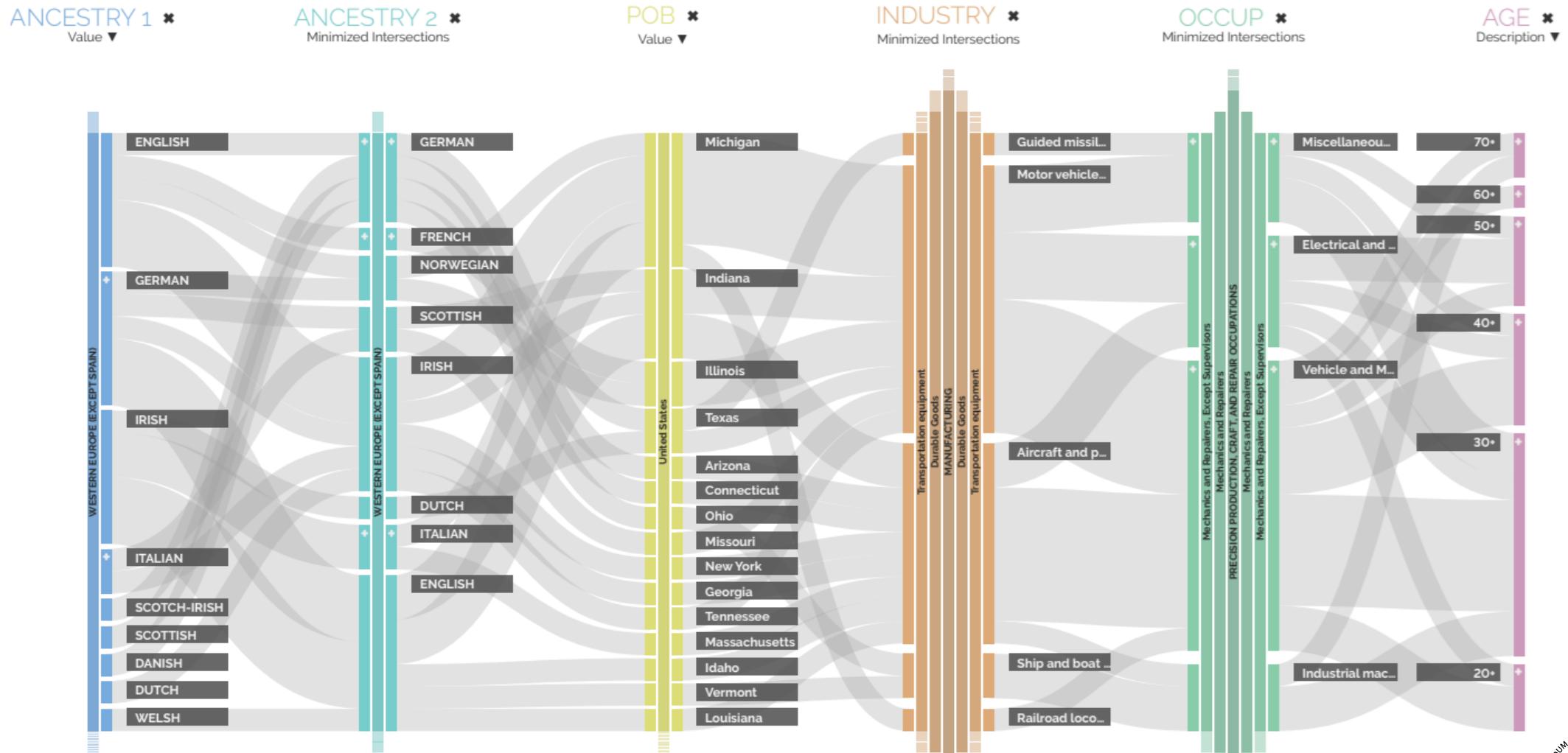
Data: [Robert J. MacG. Dawson](#).

# PARALLEL SETS

---



# PARALLEL HIERARCHIES [VOSOUGH ET AL. 2018]





## ANCESTRY 1 ×

Description ▾

- + ITALIAN
- + IRISH
- + SCOTTISH
- + SWEDISH
- + SCOTCH-IRISH
- + NORWEGIAN
- + WELSH
- + GERMAN
- + WESTERN EUROPE (EXCEPT SPAIN)
- + FRENCH
- + ENGLISH
- + DUTCH

## ANCESTRY 2 ×

Description ▾

- + IRISH
- + NORWEGIAN
- + ITALIAN
- + SCOTTISH
- + SWEDISH
- + SCOTCH-IRISH
- + WELSH
- + GERMAN
- + WESTERN EUROPE (EXCEPT SPAIN)
- + FRENCH
- + ENGLISH
- + DUTCH

## INDUSTRY ×

Description ▾

- + Furniture and fl...
- + Professional an...
- + Not specified ...
- + Transportation ...
- + MANUFACTURING
- + Durable Goods
- + Non-Durable Goods
- + Stone, clay, gla...
- + Metal Industries
- + Machinery and ...
- + Lumber and w...
- + Electrical mac...

## OCCUP ×

Description ▾

- + Professional S...
- + MANAGERIAL AND PROFESSIONAL SPECIALTY OCCUPATIONS
- + Executive, Ad...

## SEX ×

Description ▾

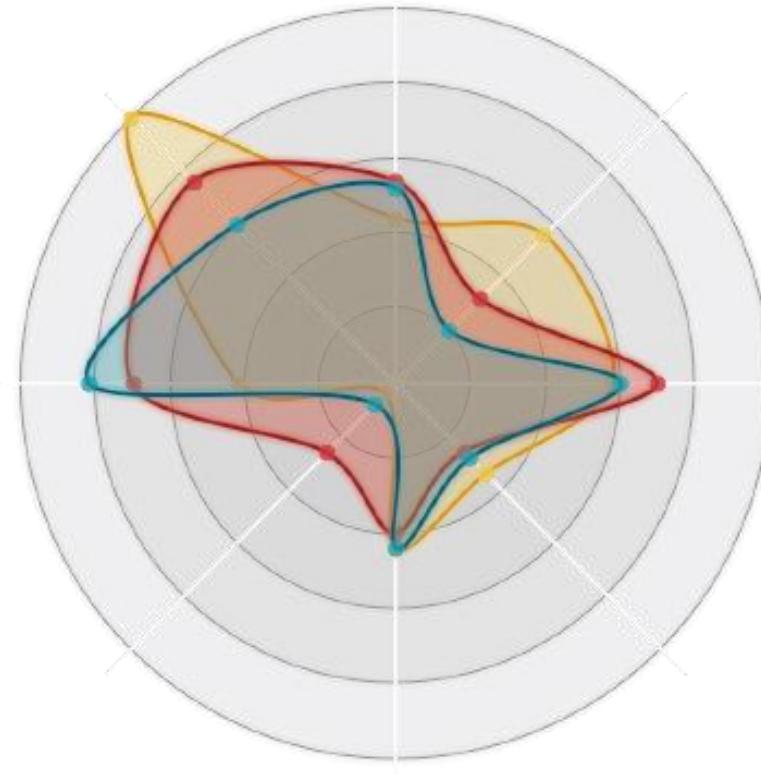
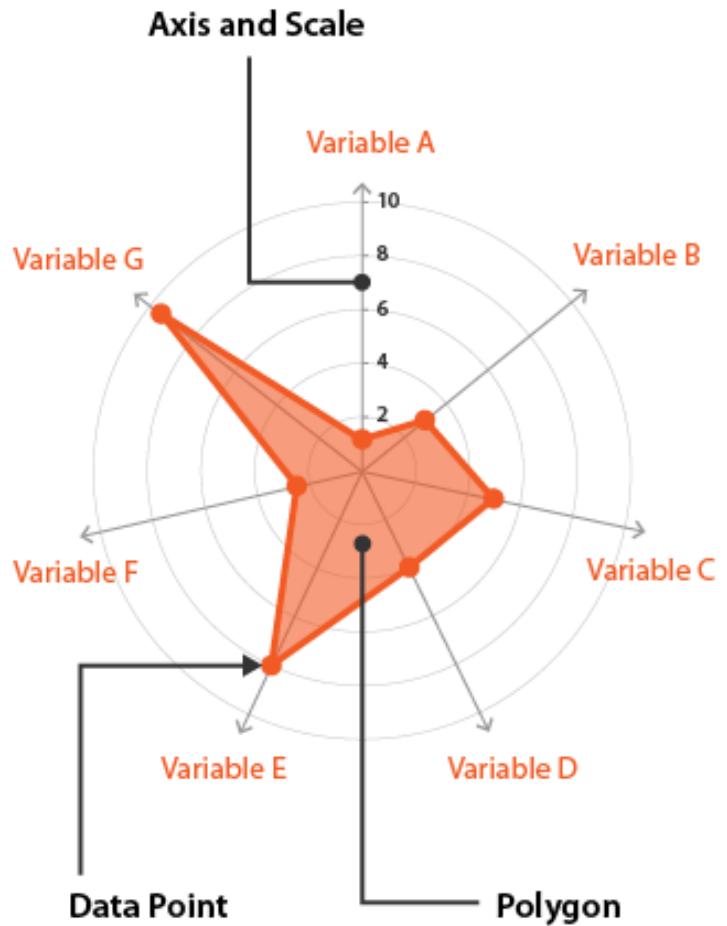
Male

Female



# RADAR CHART

---



# WHICH SPORT IS THE TOUGHEST?

ESPN enlisted a panel of expert sports scientists, academicians, journalists, and a two-sport star athlete to score 60 sports across 10 different categories. This resulted in boxing being named the most demanding sport, and fishing the least. Choose a sport below to find out more.



Select a Sport...

All

All

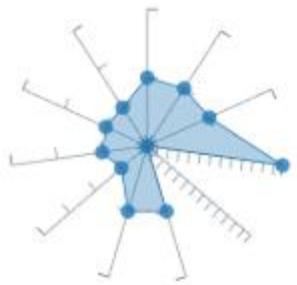
RANK: All

AVERAGE SCORE: 1.5 to 7.2

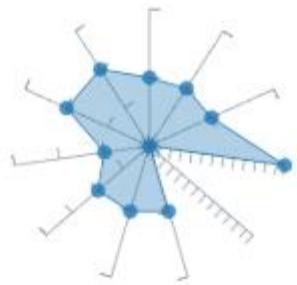
Agility	4.8
Analytical Aptitude	4.8
Durability	4.6
Endurance	5.1
Flexibility	4.7
Hand-Eye Coordination	5.0
Nerve	4.6
Power	5.5
Speed	4.8
Strength	5.2

# RADAR CHART SMALL MULTIPLES

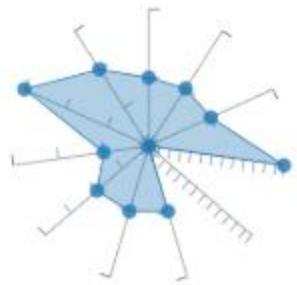
---



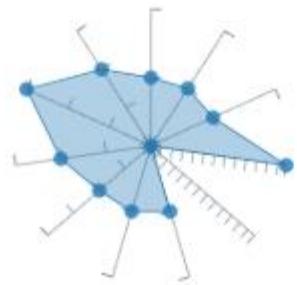
TRIE NODE



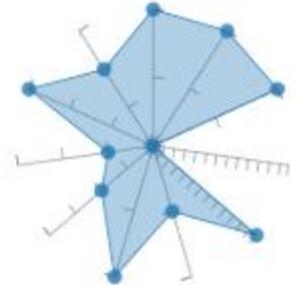
DATAPAGE (SORTED)



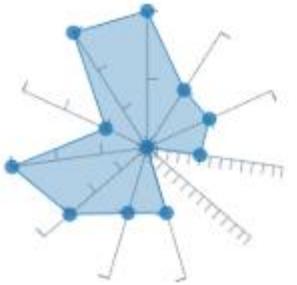
DATAPAGE



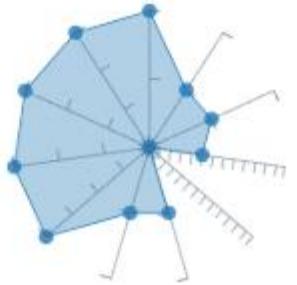
DATAPAGE (COMPR.)



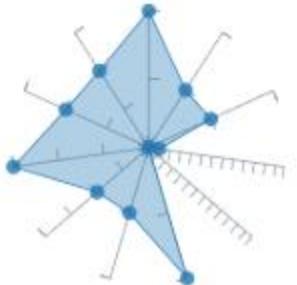
LINKED LIST



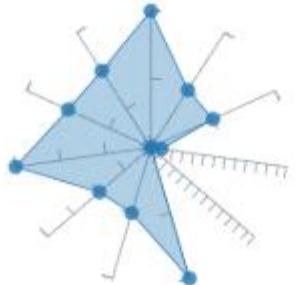
RANGE PARTITIONING



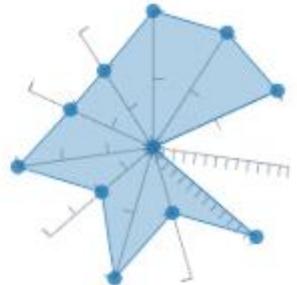
HASH PARTITIONING



B+TREE NODE



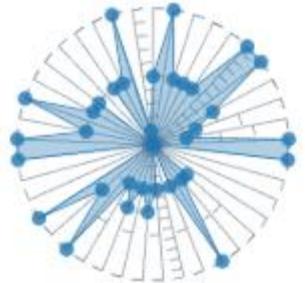
B-TREE NODE



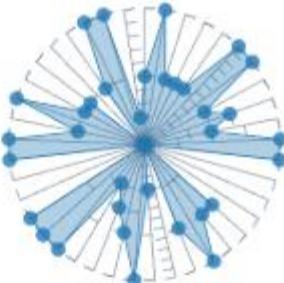
SKIP LIST

# RADAR CHART SMALL MULTIPLES

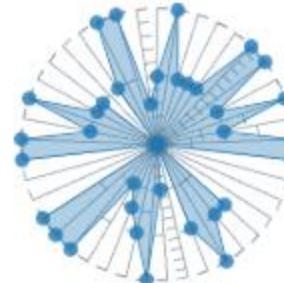
---



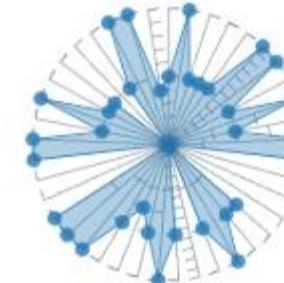
TRIE NODE



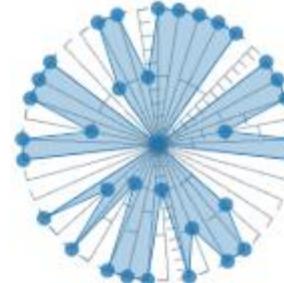
DATAPAGE (SORTED)



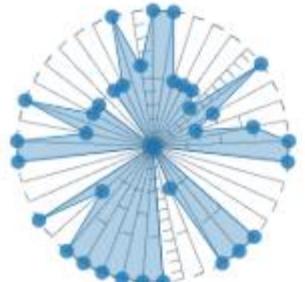
DATAPAGE



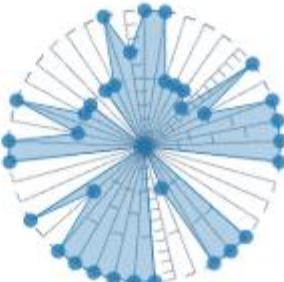
DATAPAGE (COMPR.)



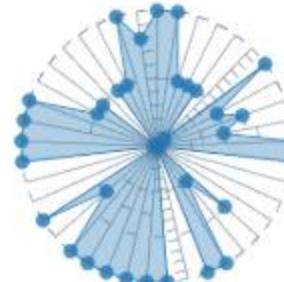
LINKED LIST



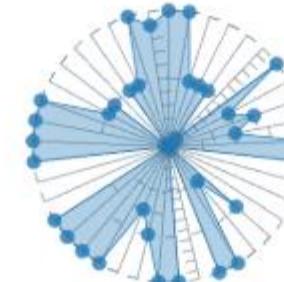
RANGE PARTITIONING



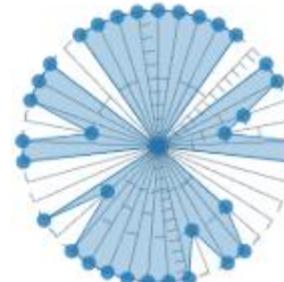
HASH PARTITIONING



B+TREE NODE



B-TREE NODE



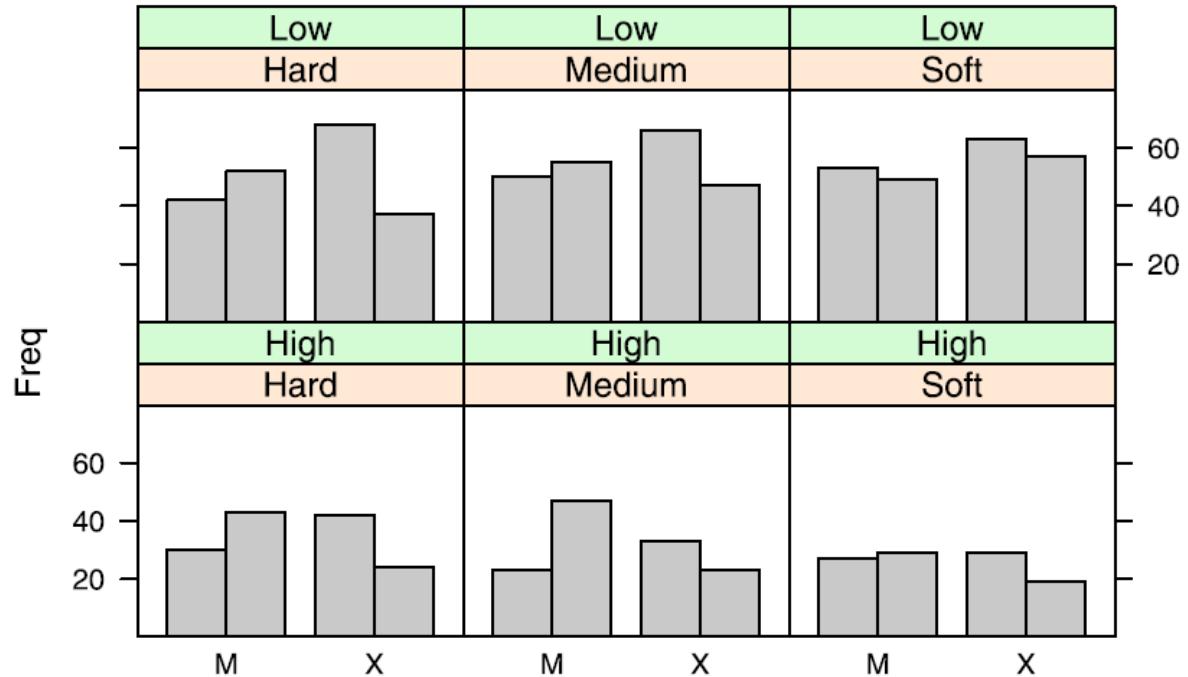
SKIP LIST

# USING NESTING

# TRELLIS DISPLAYS

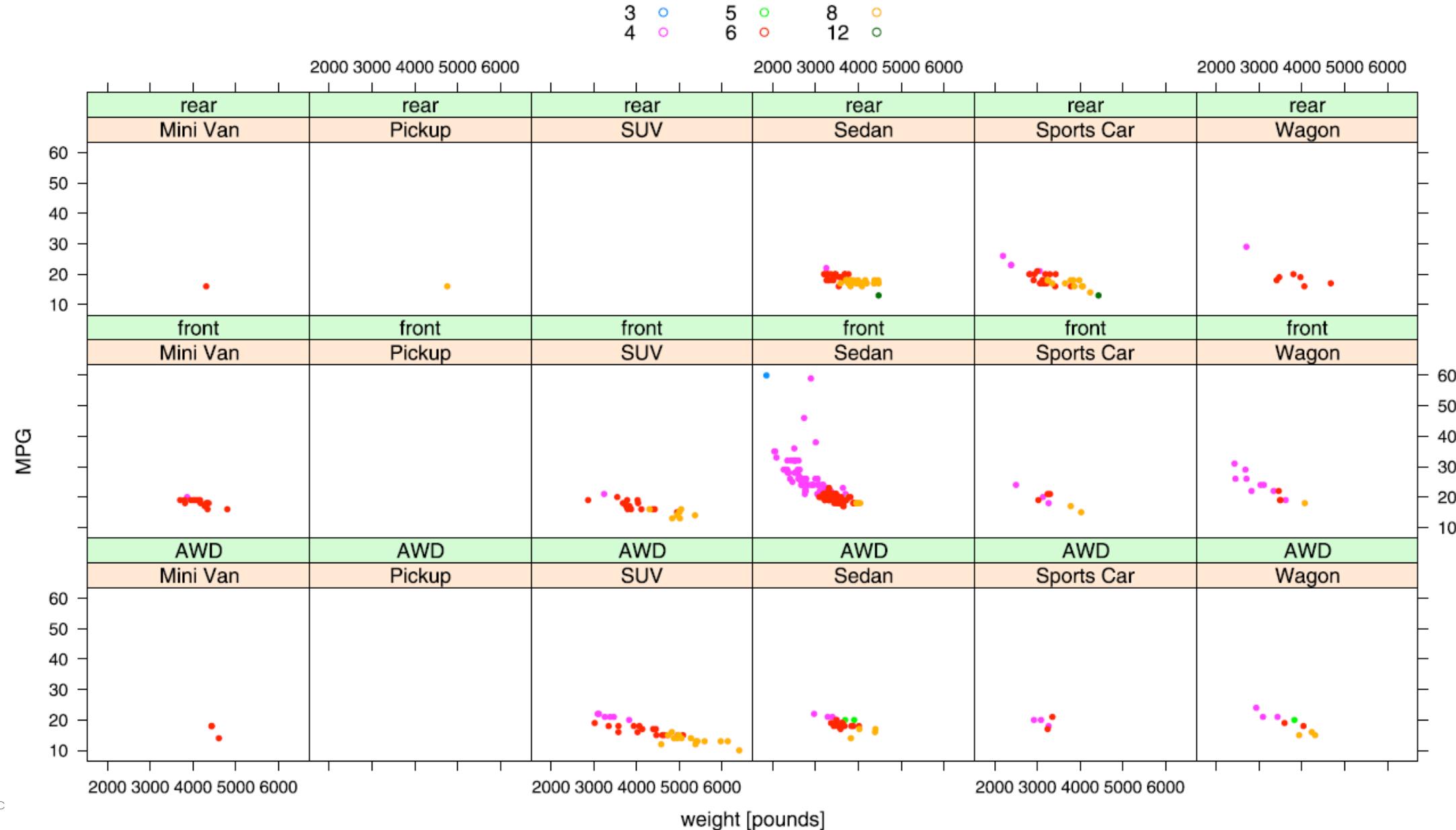
---

Trellis = grid-like structure for systematic display of complex multivariate data relations



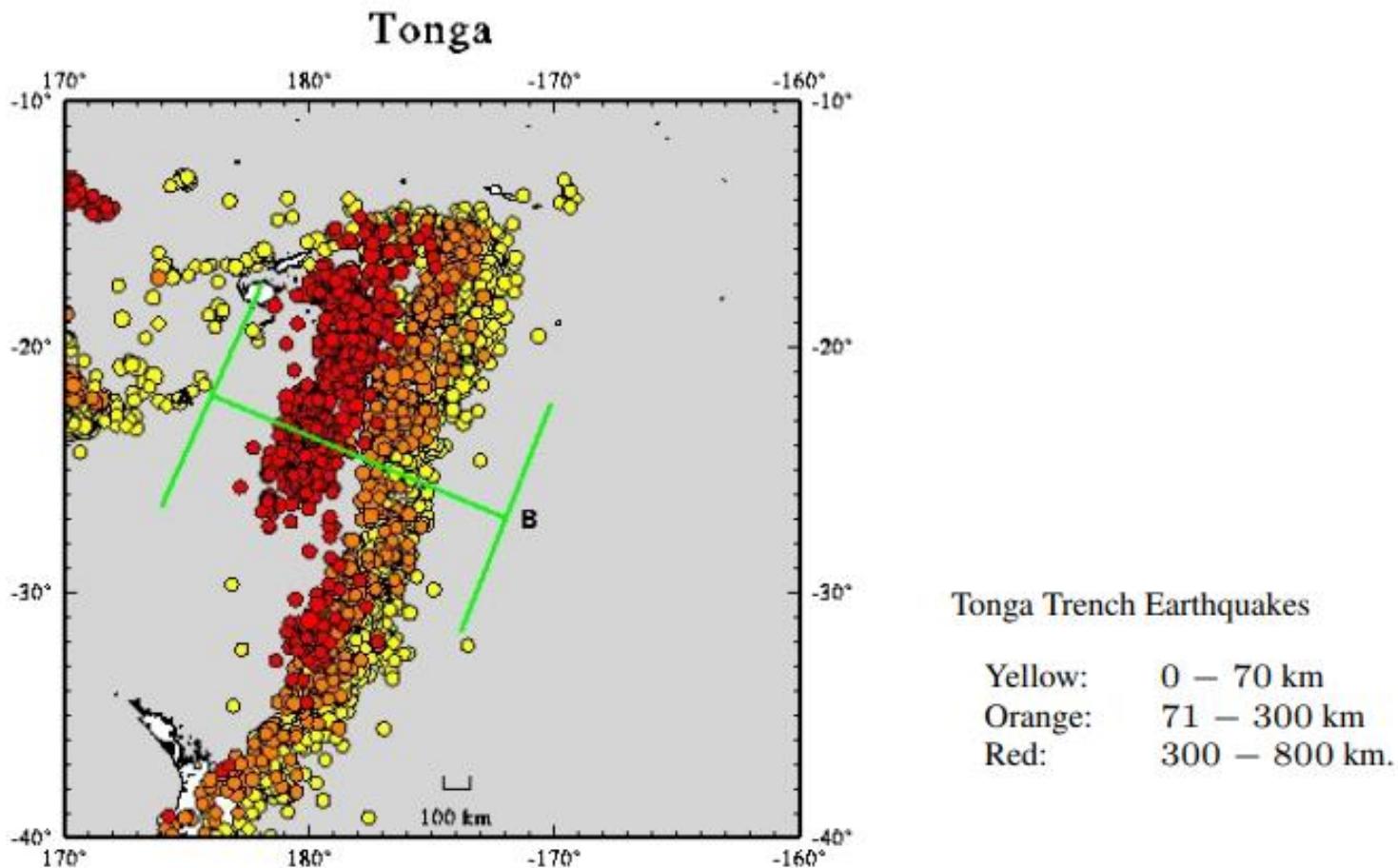
# TRELLIS DISPLAYS

---



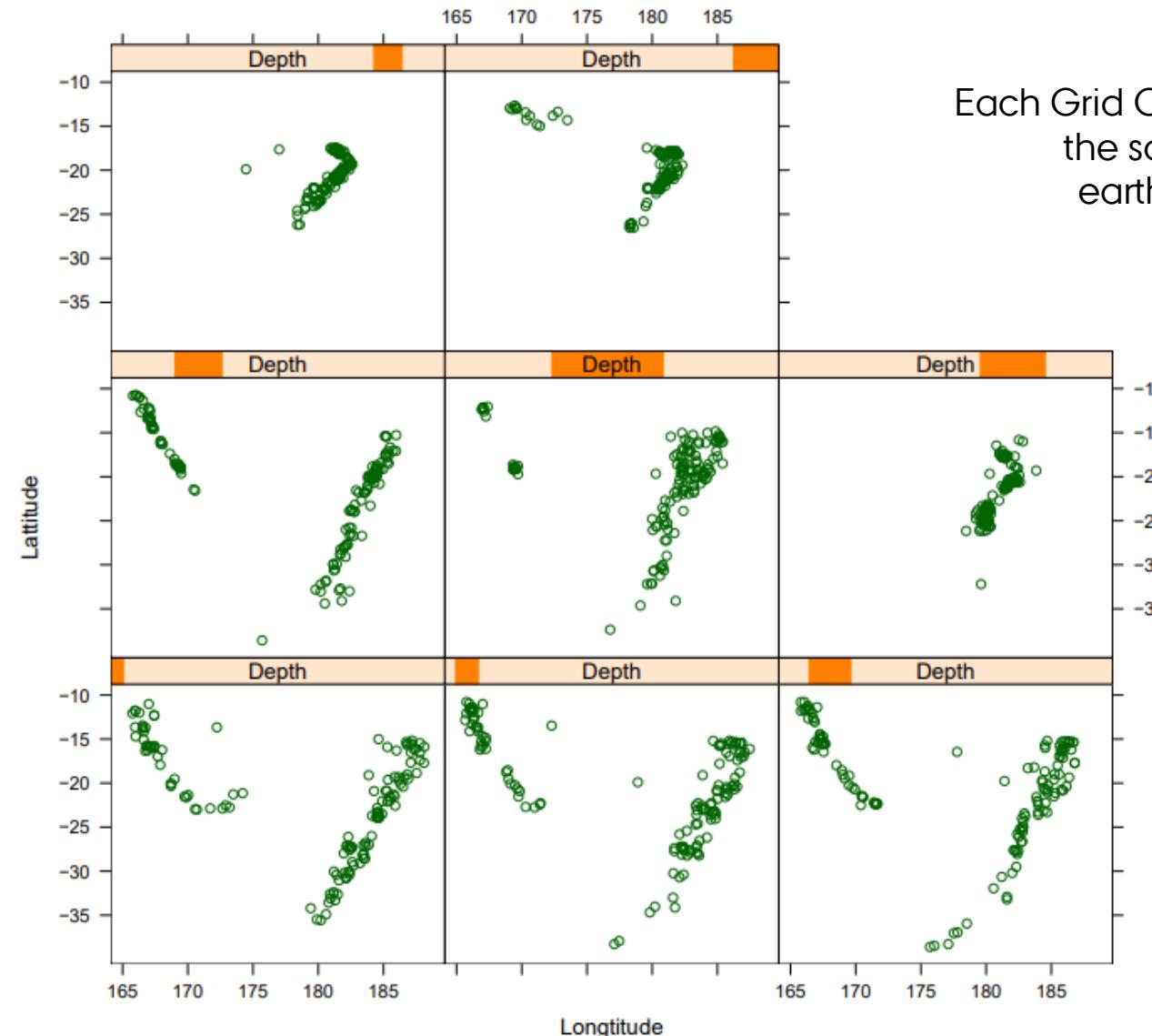
# TRELLIS DISPLAY FOR CONTINUOUS VALUES

---



# TRELLIS DISPLAY FOR CONTINUOUS VALUES

Deep Earthquakes ->

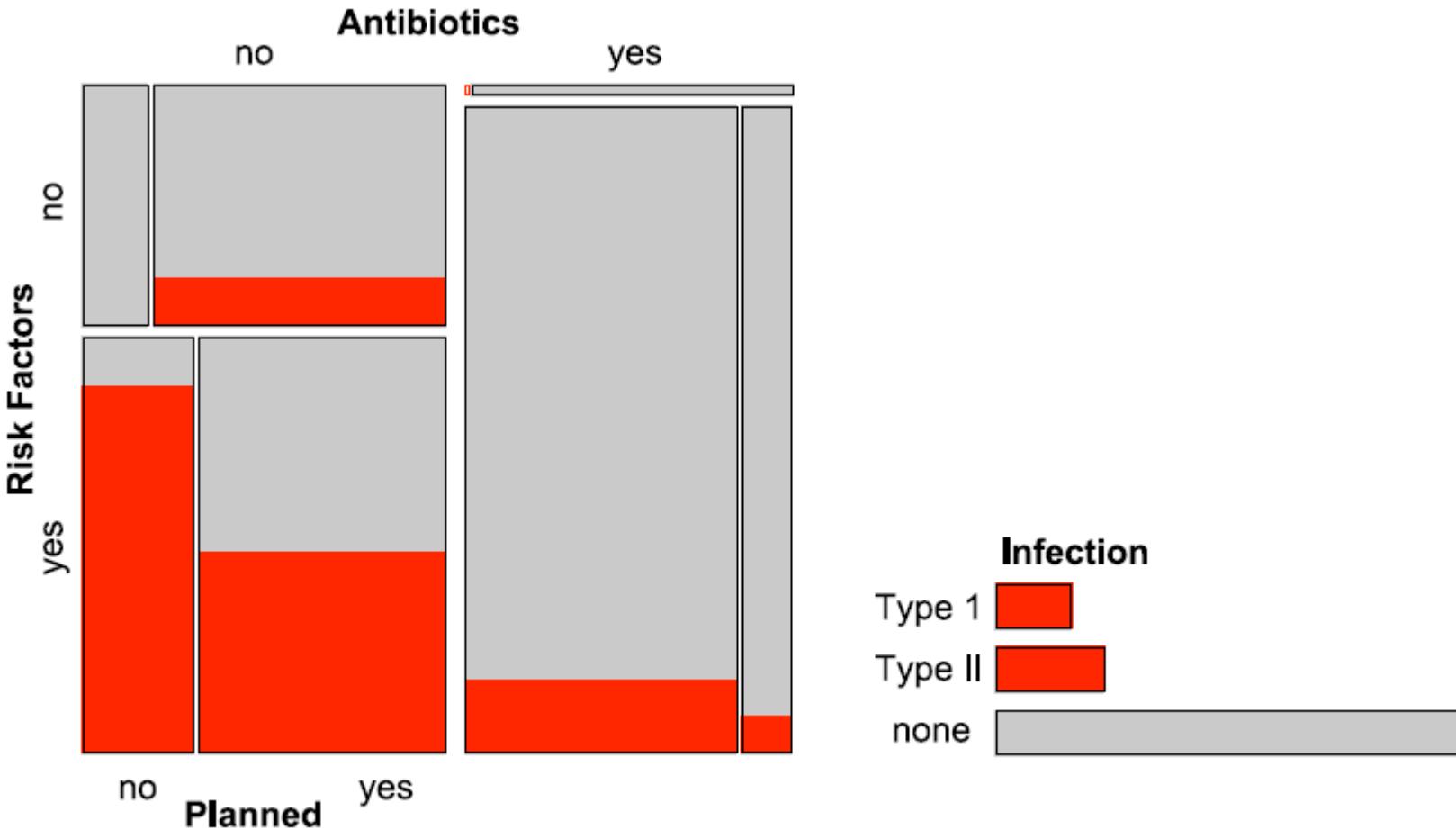


Each Grid Cell has approx.  
the same number of  
earthquakes (1/8<sup>th</sup>)

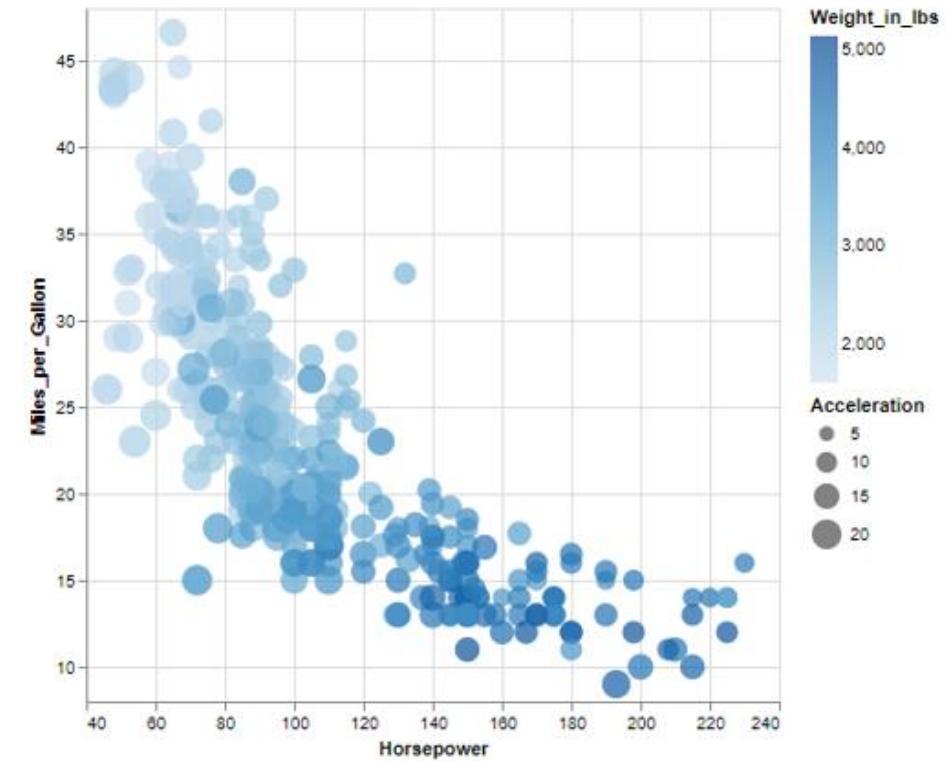
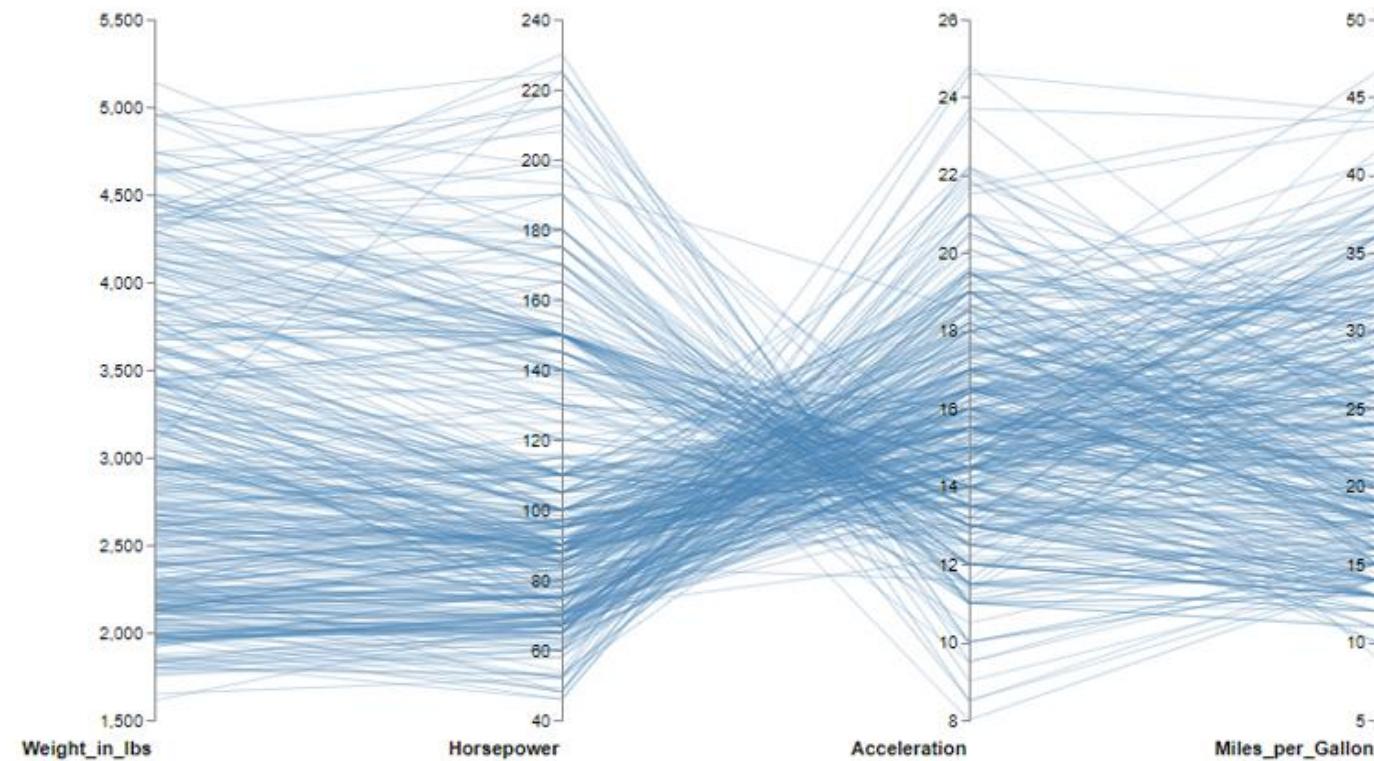
Shallow Earthquakes ->

# MOSAIC PLOTS

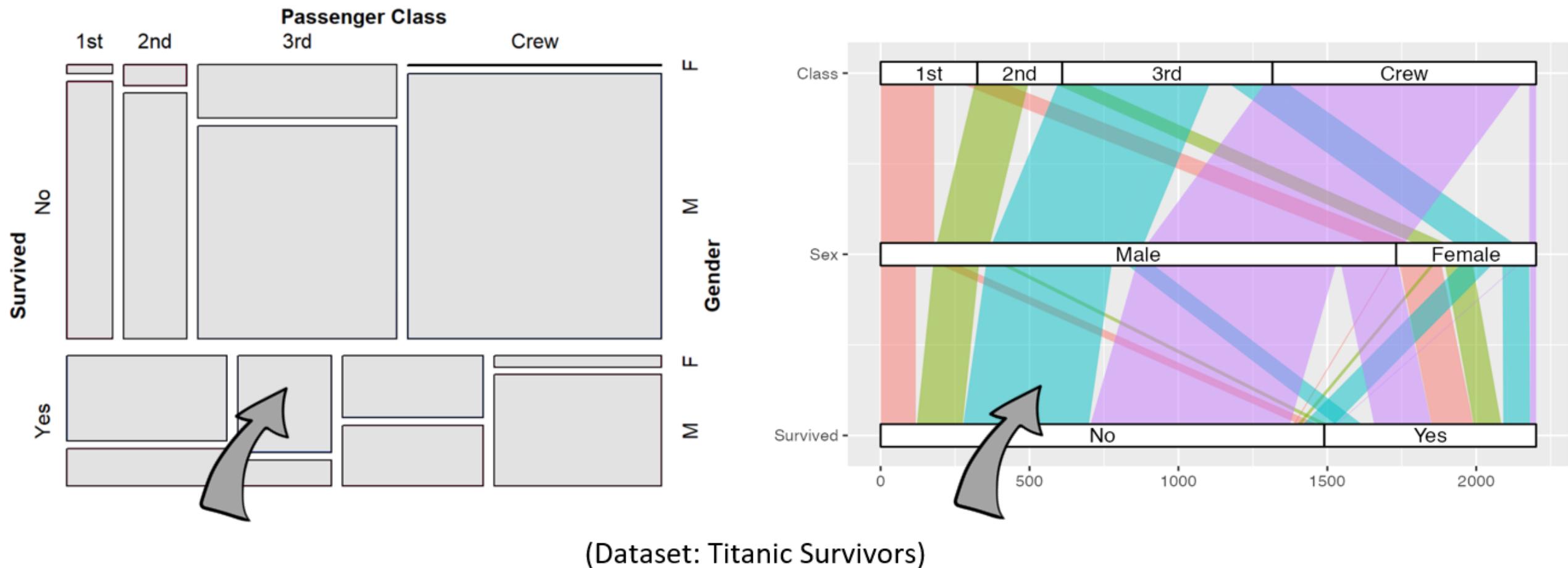
---



# CHART PRACTICE



(Dataset: Cars)

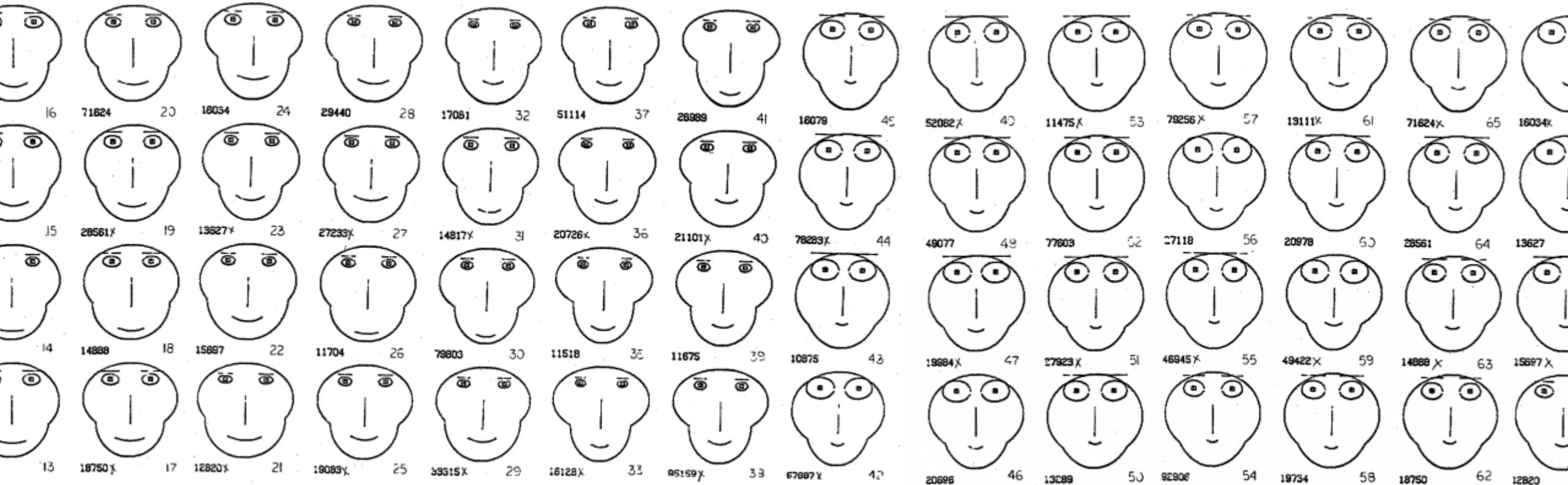


# OTHER TECHNIQUES

# GLYPHS – CHERNOFF FACES

---

Morris, Ebert, Rheingans: Experimental analysis of the effectiveness of features in Chernoff faces, 2000



Data attributes encoded in the different visual features of these faces –  
e.g., length of eyebrows, size of eyes, size of mouth,...

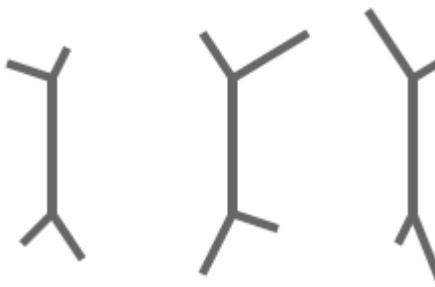
# GLYPH VARIANTS

---

Star Glyphs



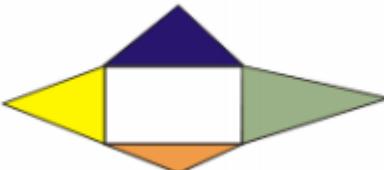
Stick Figures



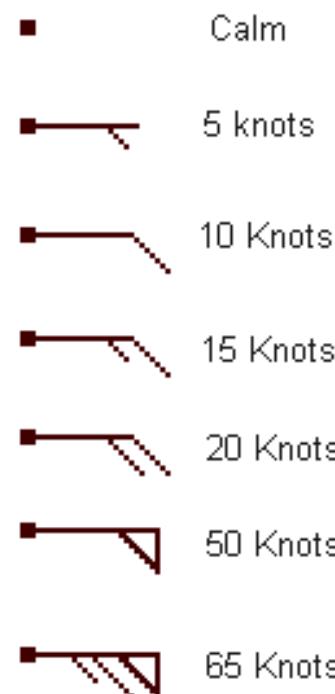
Anderson's Metroglyph



Data Jacks

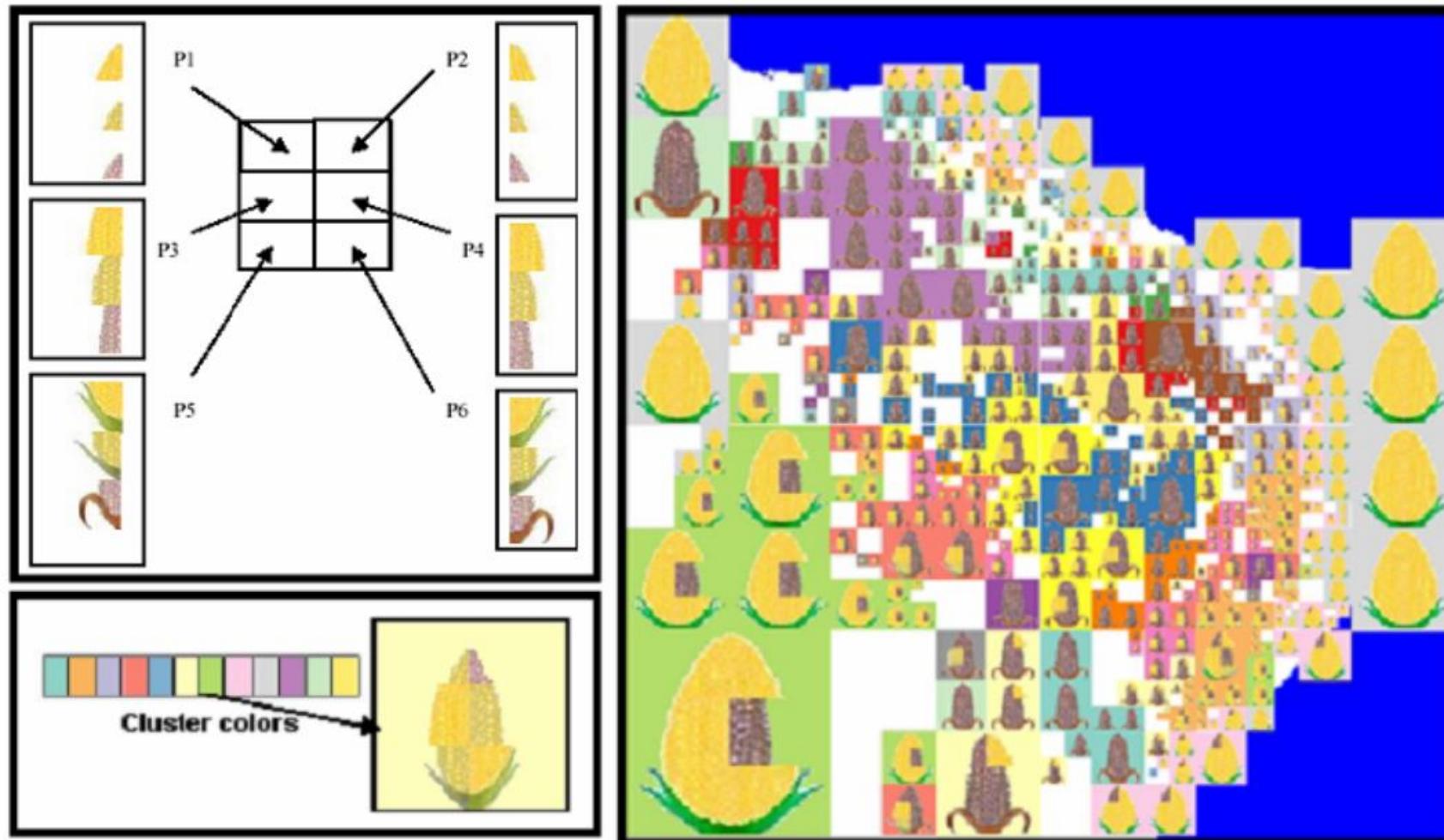


Wind Bars

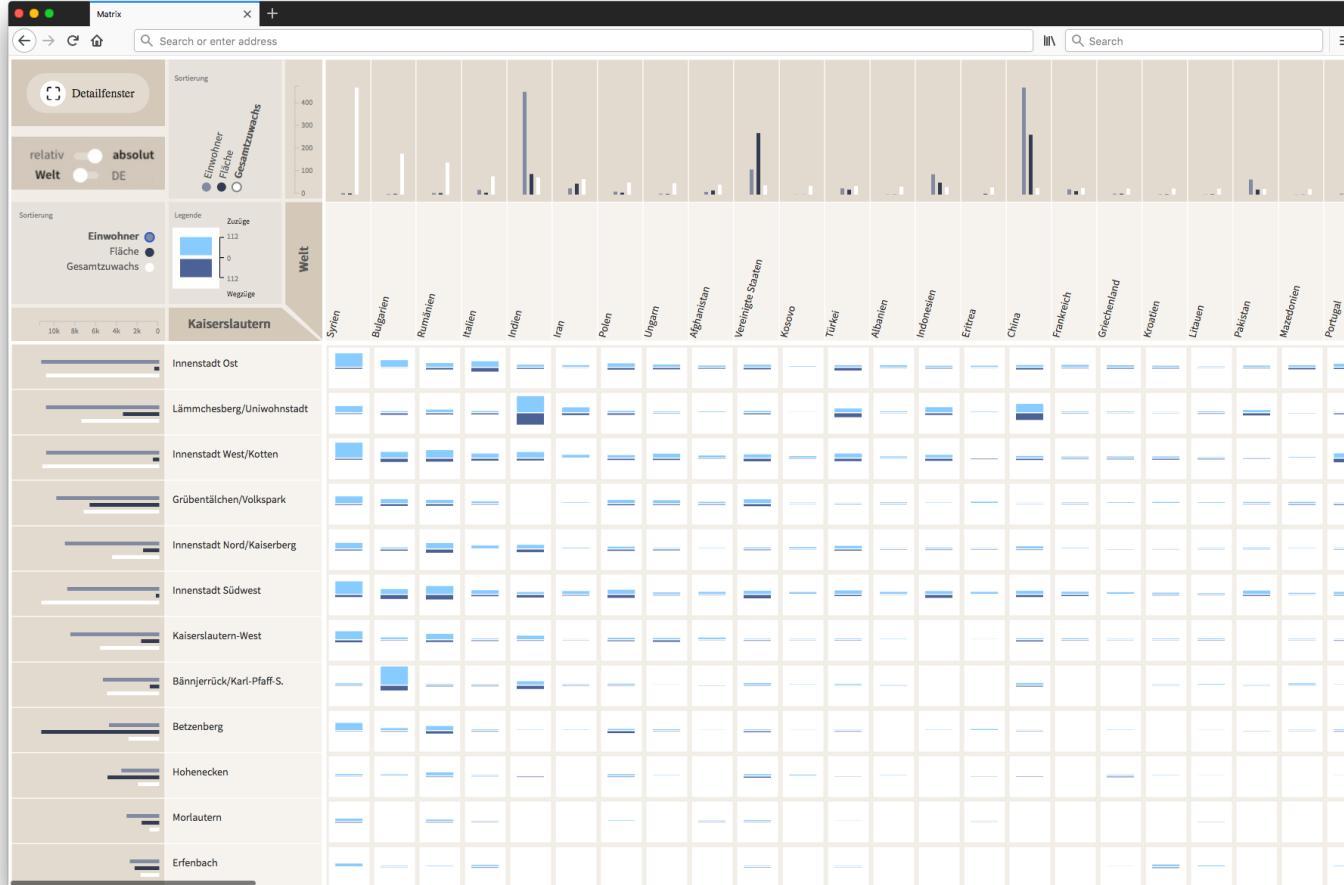


# ALMOST NO LIMITS: CORN GLYPHS

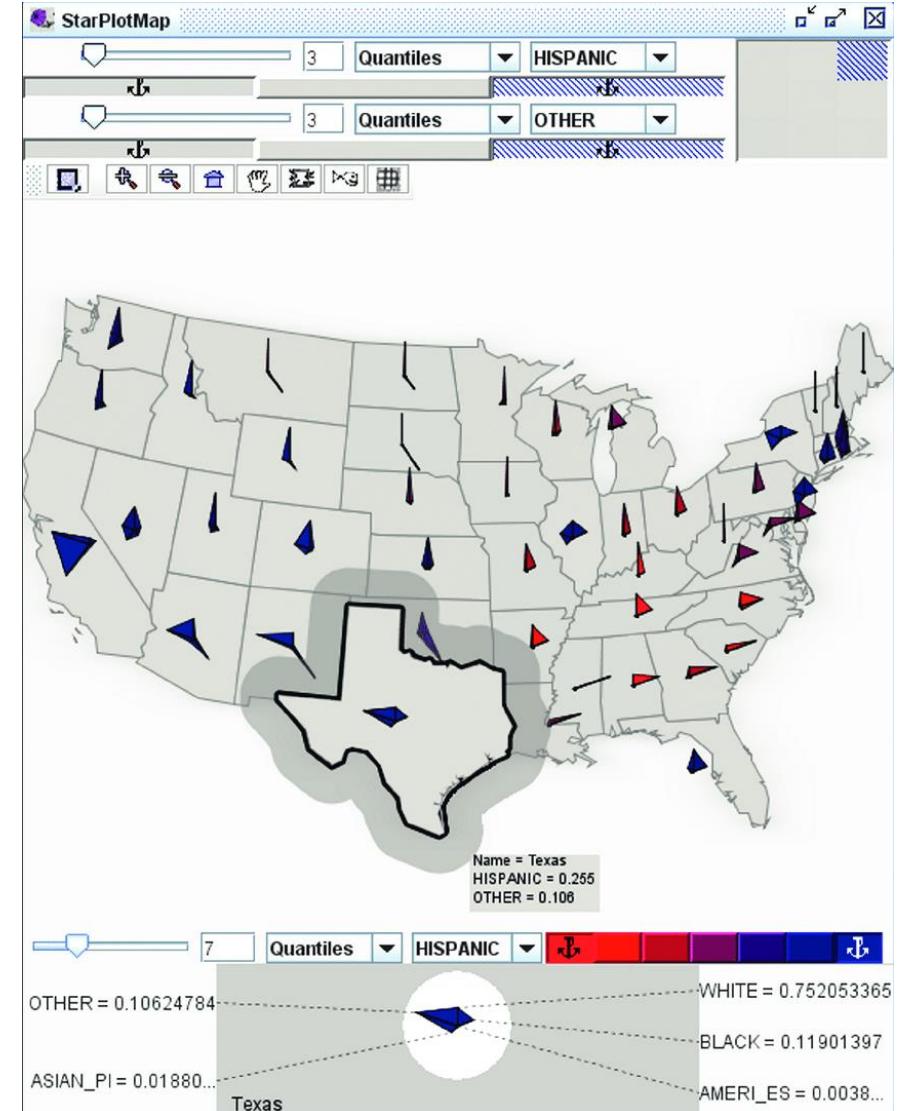
[Nocke et al. 2005]



# USED FOR EMBEDDING



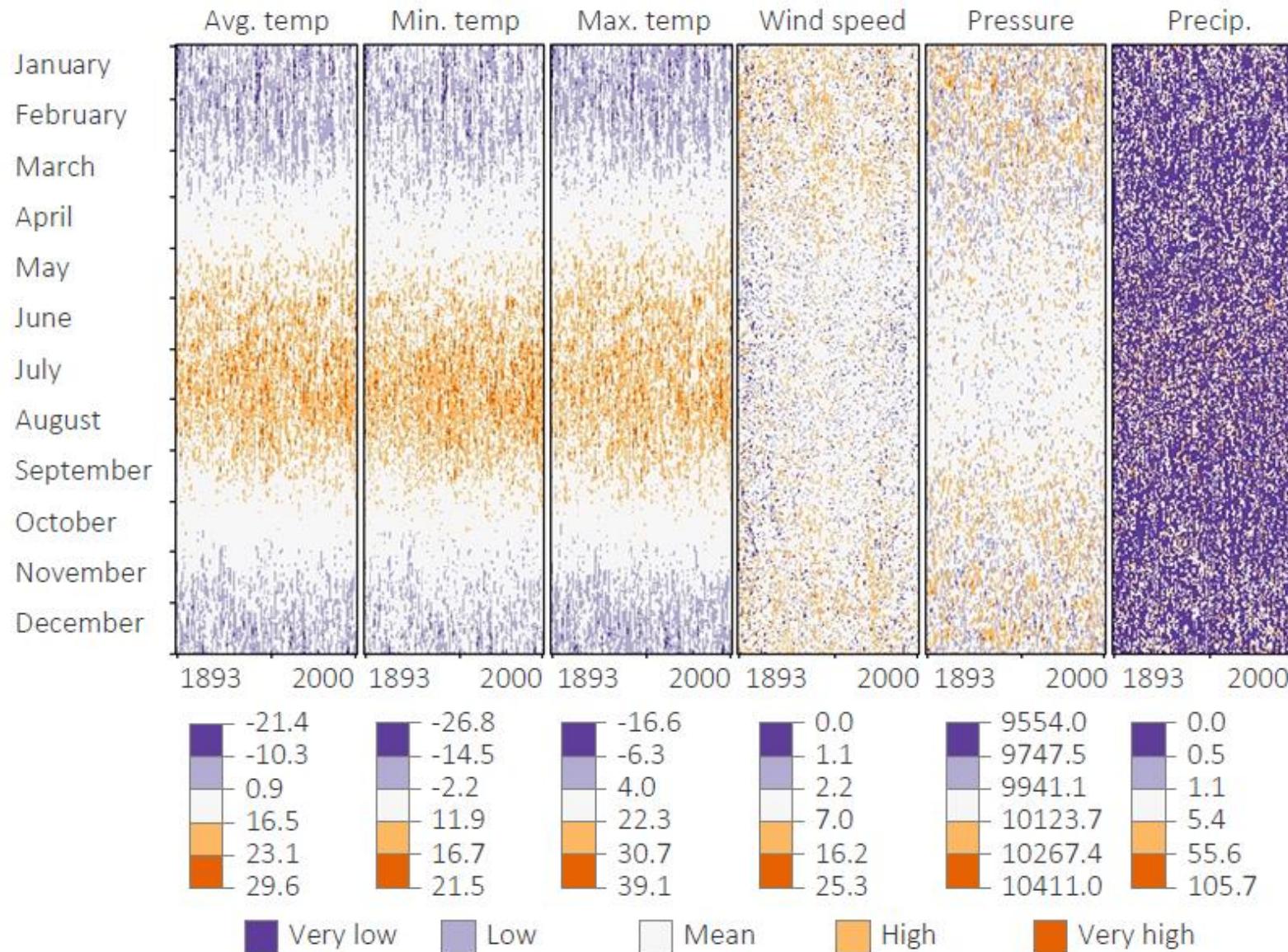
[ <https://infovis-mannheim.de/wandervis/> ]



[Klipfel et al. 2009]

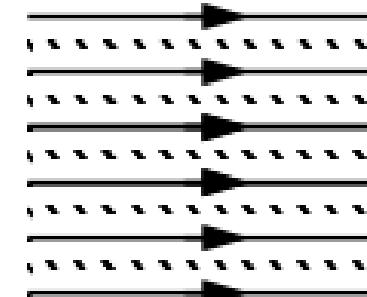
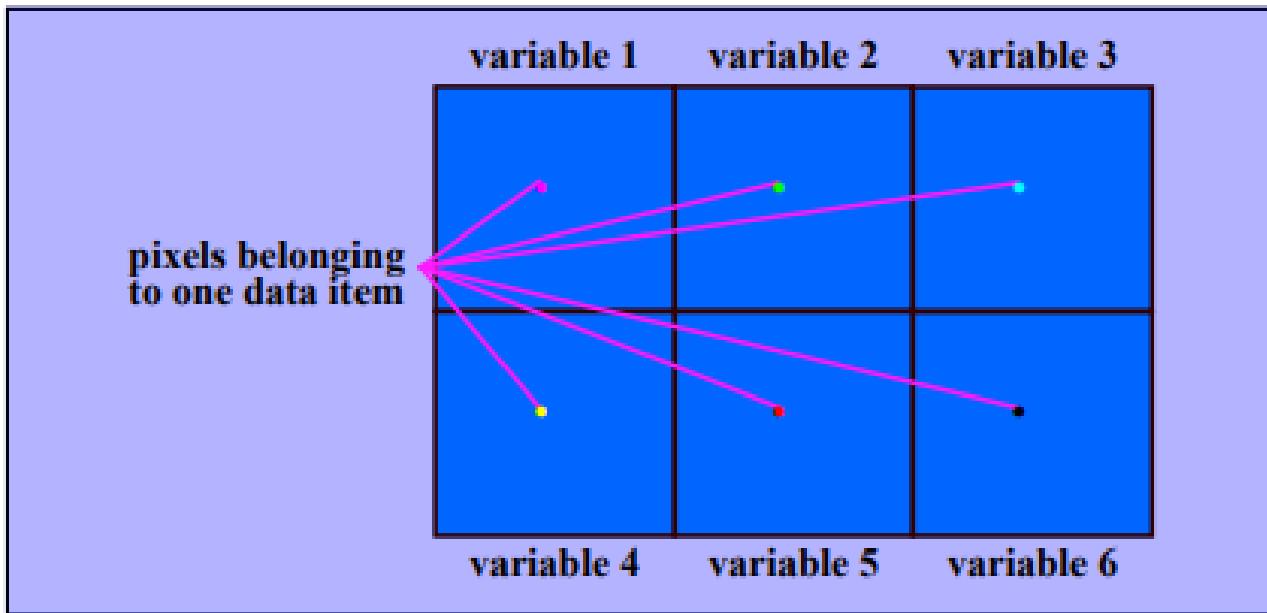
# PIXEL-BASED TECHNIQUES

[Keim 2000]

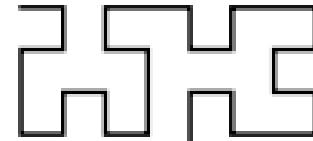


# PIXEL-BASED TECHNIQUES

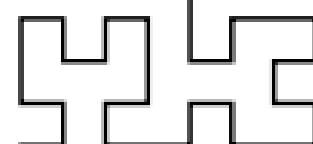
[Keim 2000]



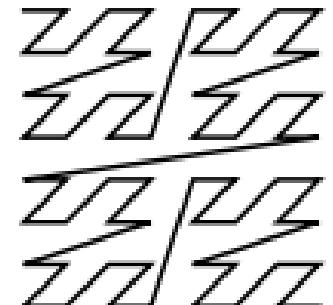
a. line-by-line



b. column-by-column



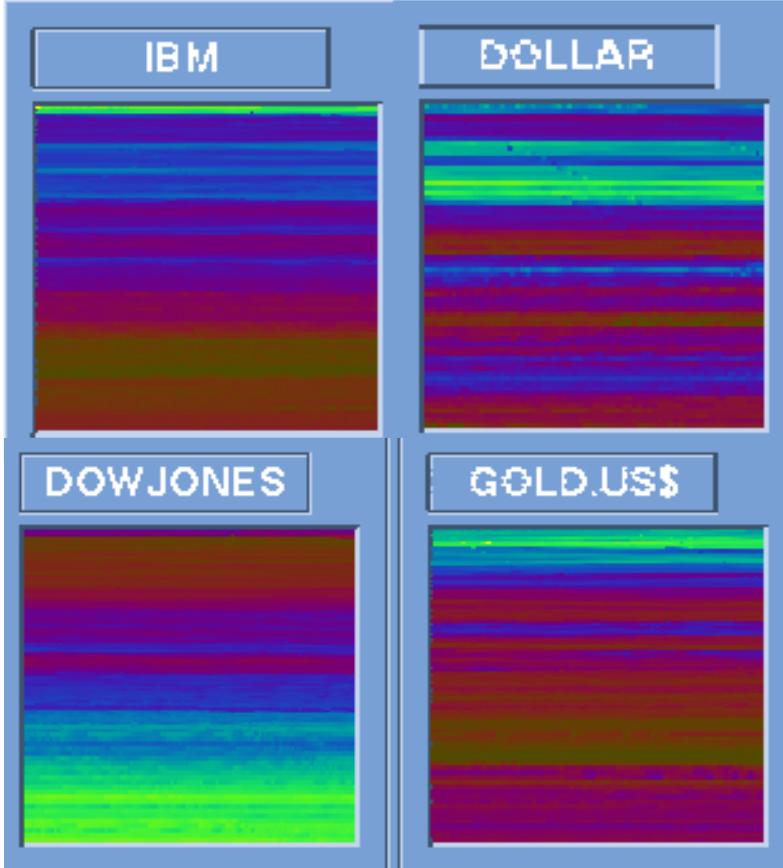
c. Peano-Hilbert



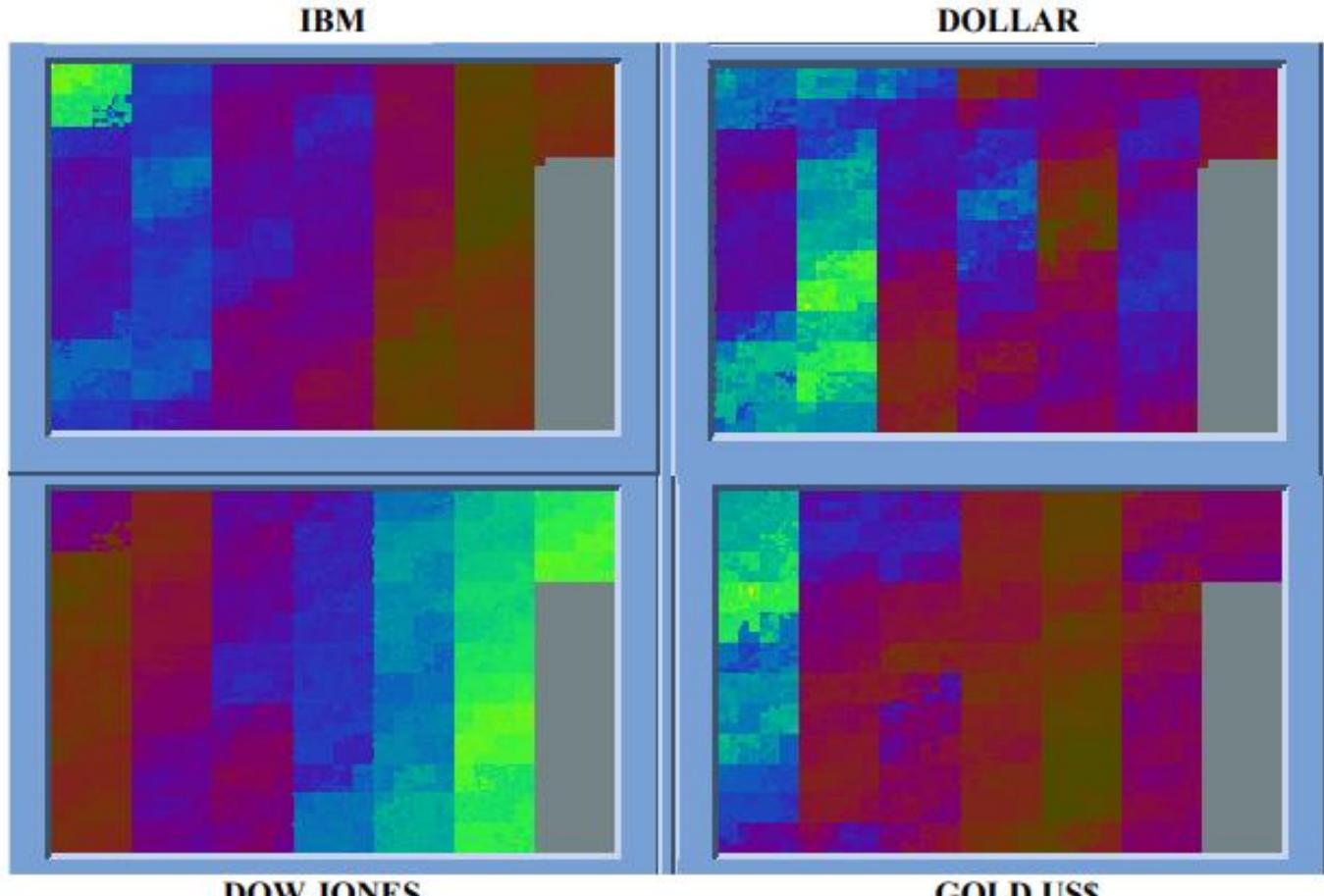
d. Morton

# PIXEL-BASED TECHNIQUES

[Keim 2000]



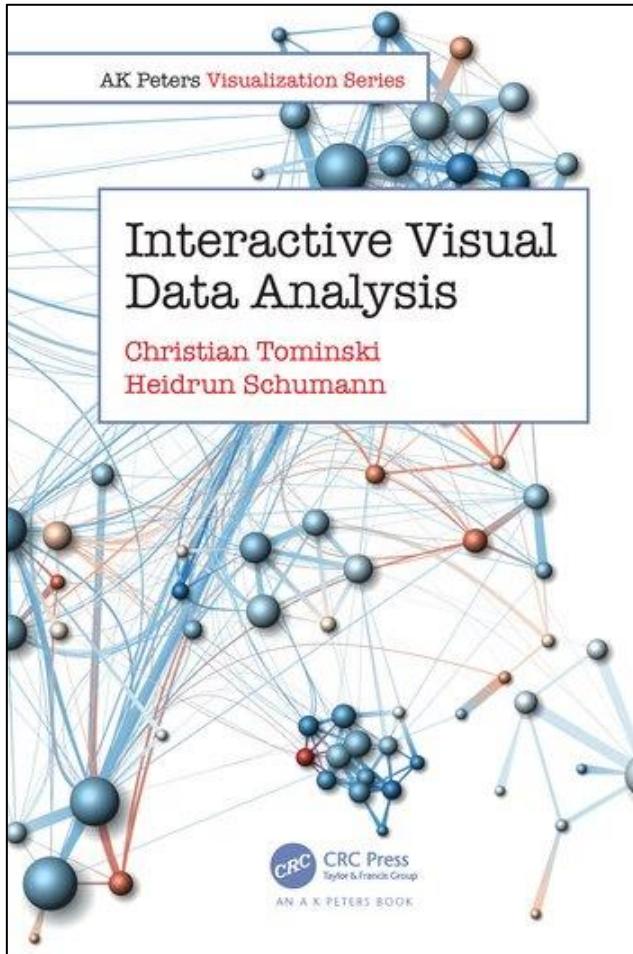
Line-by-line Back & Forth  
Arrangement



Calendar Arrangement

# WEEKLY READING

---



## 82 ■ Interactive Visual Data Analysis

### 3.3 VISUALIZATION OF TEMPORAL DATA

Time is an exceptional dimension. Virtually everything around us is governed by the steady progress of time. So it comes as no surprise that much of the data that people seek to understand are connected to time. In this section, we add the dimension of time  $T$  to our considerations. In other words, we are interested in methods and techniques for visualizing time and temporal data, where the latter primarily means communicating the dependency of data attributes on time  $T \rightarrow A$ . As we will see in a moment, time and data that depend on time are quite special and require dedicated techniques for their visualization.

#### 3.3.1 Time and Temporal Data

Time is not just another data attribute. Time has several properties that need to be taken into account when visualizing data that are connected to time. In the pages to come, we will briefly characterize time and temporal data.

##### *Characterizing Time*

Philosophers have pondered the concept of time for ages. Here we want to concentrate on aspects of time that are relevant for the visual analysis of temporal data.

Before going into the details of time, there is a more general aspect to be mentioned: Analog to the different data domains discussed in Section 2.2.1, the time domain can be characterized as ordinal, discrete, or continuous. For *ordinal* time, only equality and partial order relations are defined. For *discrete* time, a mapping exists from the time domain to the set of integers, which makes it possible to measure distances in time. *Continuous* time conceptually maps to the set of real numbers. As such, continuous time is dense, which, in general, complicates the data handling and visualization.

Let us next look at the specifics of time in more detail. In particular, we will deal with four aspects: primitives, arrangement, granularity, and structure

# LIST OF LITERATURE SOURCES

---

- Elmqvist et al. 2008 (Rolling the Dice): <https://doi.org/10.1109/TVCG.2008.153>
- Hoffman et al. 1997 (Original RadViz): <https://doi.org/10.1109/VISUAL.1997.663916>
- Cheng et al. 2017 (Radviz Deluxe): <https://doi.org/10.3390/pr5040075>
- Pagliosa & Telea 2019 (Radviz++): <https://doi.org/10.3390/informatics6020016>
- Inselberg 2009 (PCS): <https://doi.org/10.1007/978-0-387-68628-8>
- Raseman et al. 2019: <https://doi.org/10.1016/j.envsoft.2019.03.005>
- Bendix et al. 2005: <https://doi.org/10.1109/INFVIS.2005.1532139>
- Vosough et al. 2018: <https://doi.org/10.1016/j.cag.2018.07.009>
- Morris et al. 2000: <https://doi.org/10.1111/12.384865>
- Fuchs 2015 (Glyph Design for Temporal and Multi-Dimensional Data): <https://d-nb.info/1112944869/34>
- Nocke et al. 2005: <https://doi.org/10.1109/IV.2005.58>
- Klippel et al. 2009: <https://doi.org/10.1559/152304009788188808>
- Tong et al. 2016: <https://doi.org/10.2312/eurovisshort.20161154>
- Fuchs et al. 2015: [https://doi.org/10.1007/978-3-319-29971-6\\_7](https://doi.org/10.1007/978-3-319-29971-6_7)
- Chung et al. 2015: <https://doi.org/10.1177/1473871613511959>
- Keim 2000: <https://doi.org/10.1109/2945.841121>