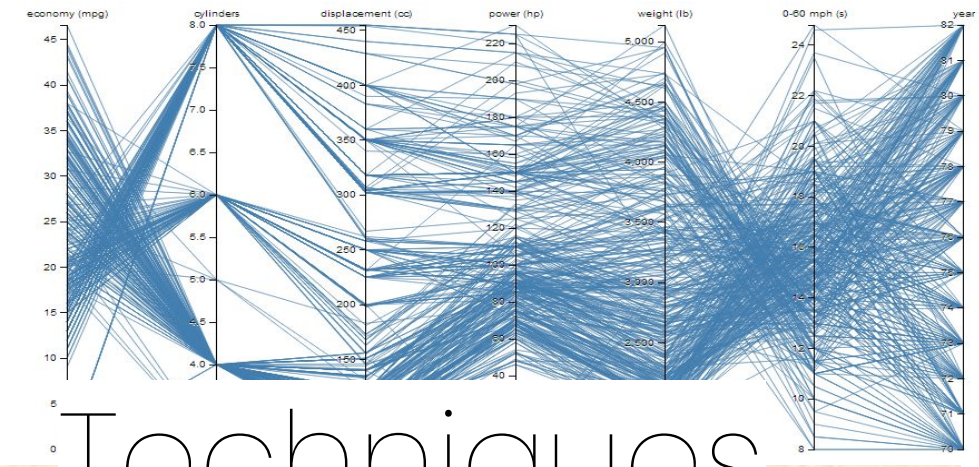
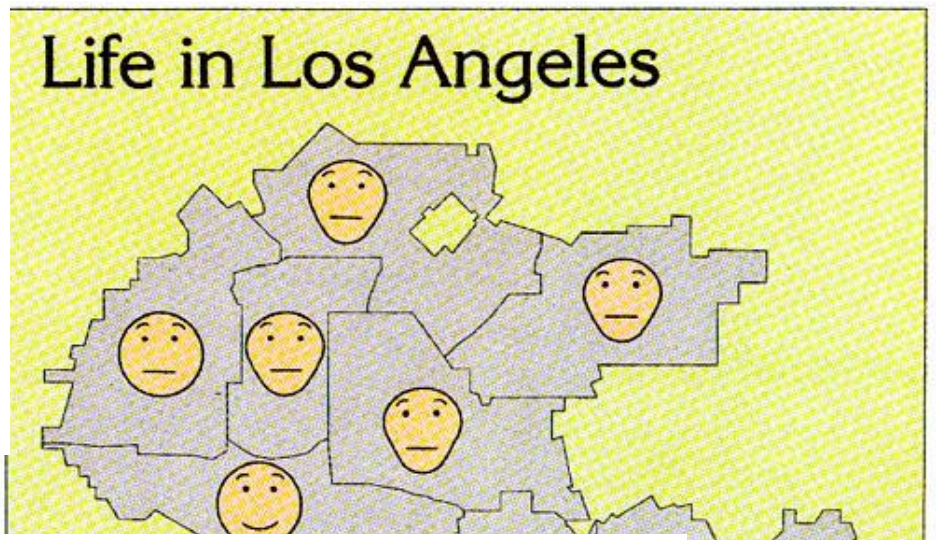
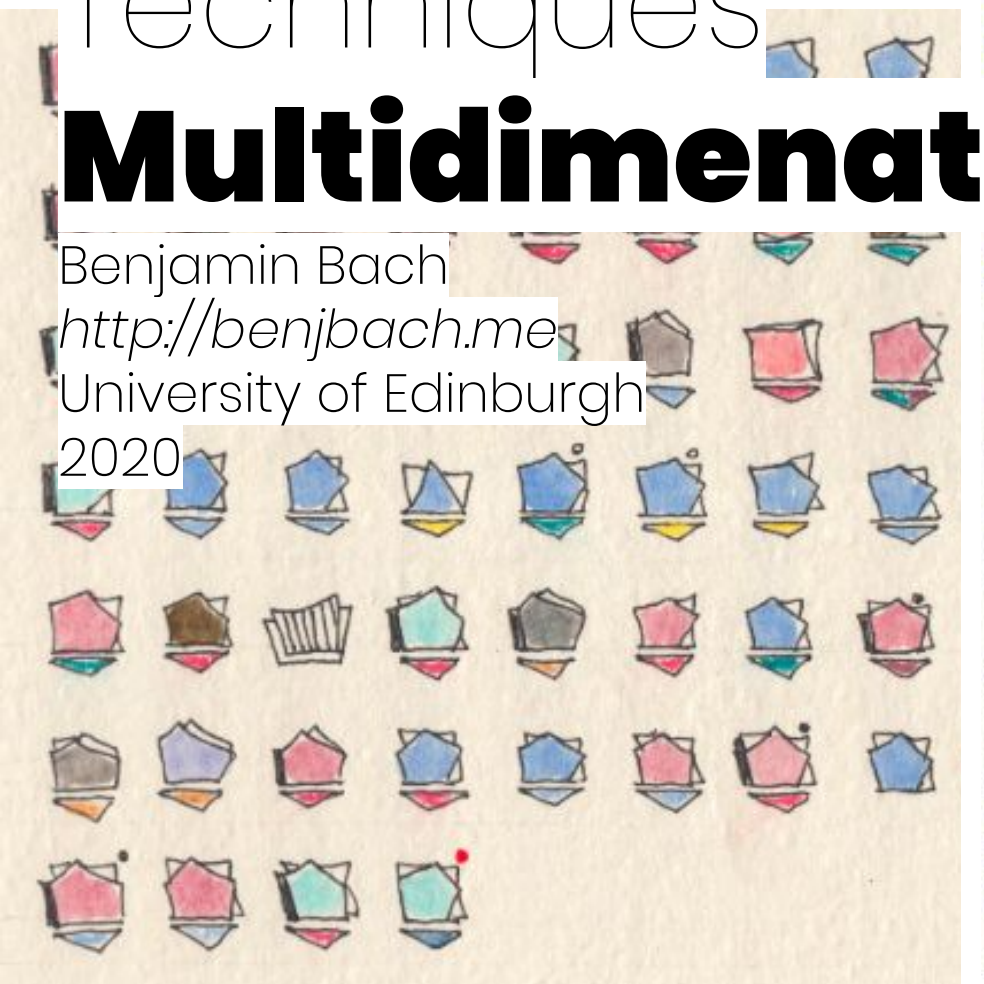


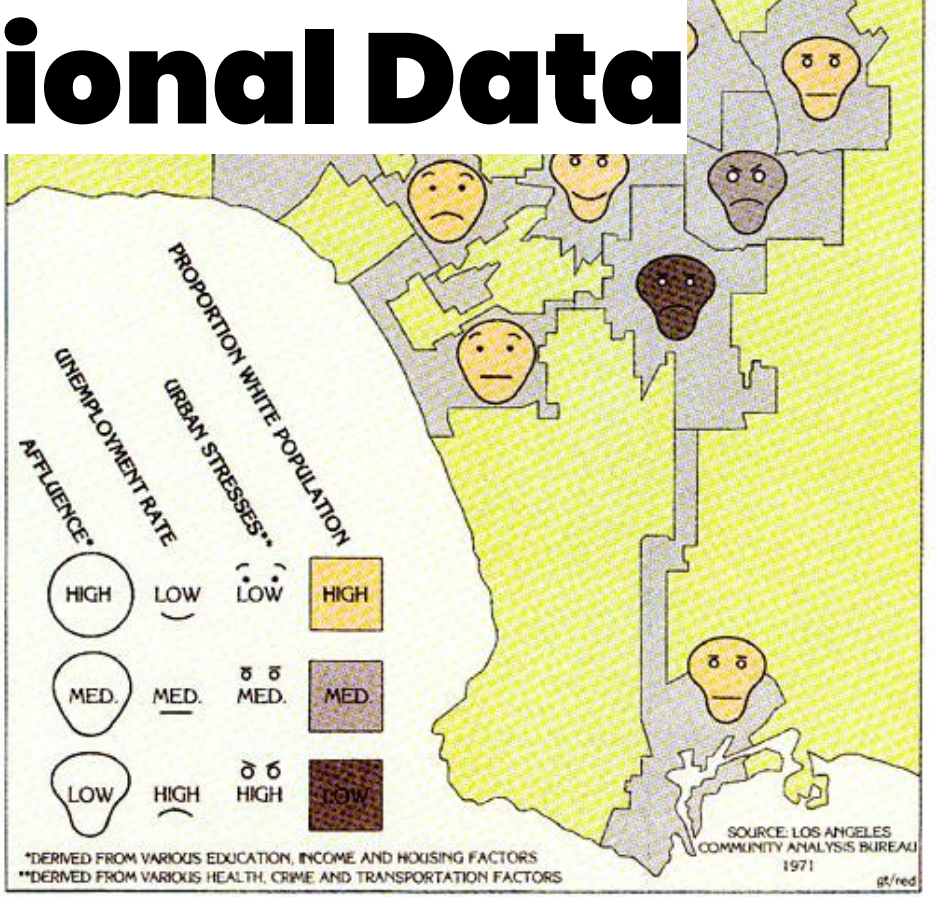


Techniques

Multidimenational Data

Benjamin Bach
http://benjbach.me
University of Edinburgh
2020

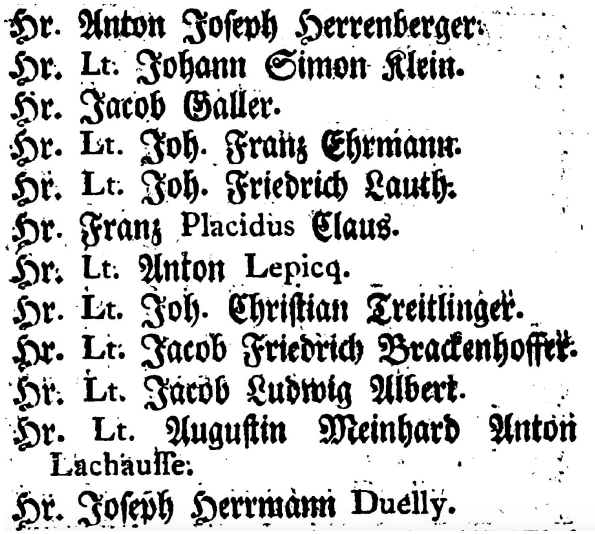




Data Types: Ben Shneiderman

- **1-dimensional**
- 2-dimensional
- 3-dimensional
- Multidimensional
- Temporal
- Trees
- Networks

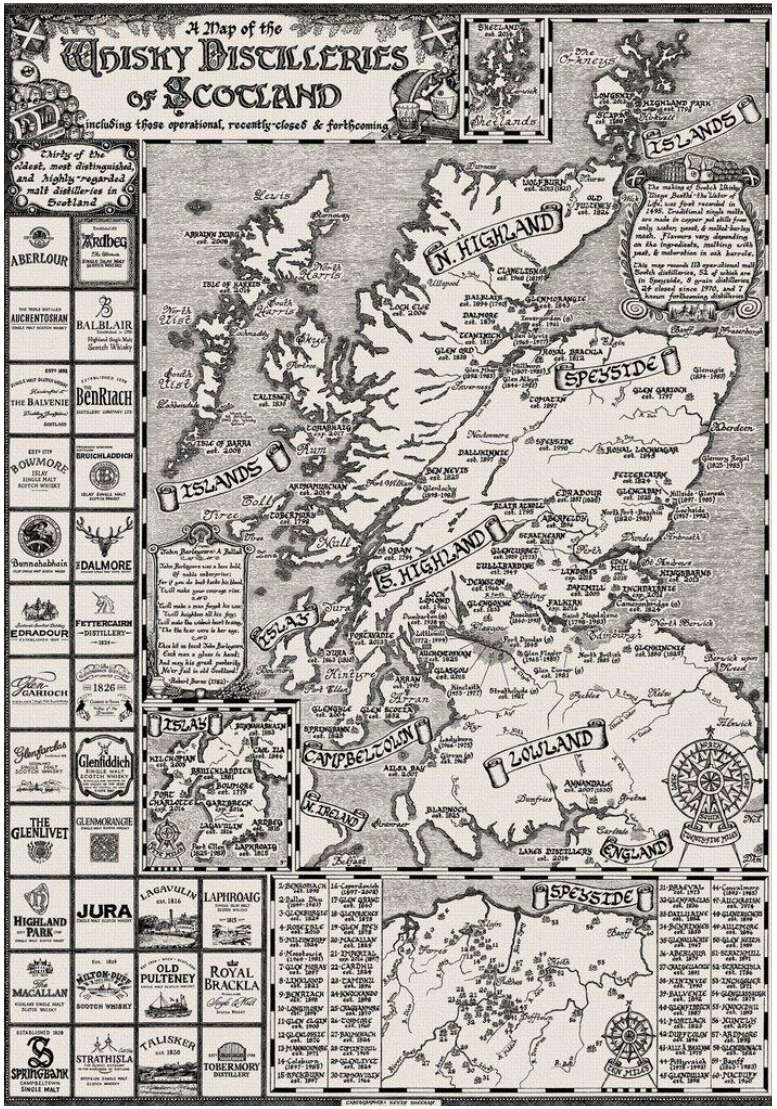
total_bill
16.99
10.34
21.01
23.68
24.59
25.29
8.77
26.88
15.04
14.78
10.27
35.26


Hr. Anton Joſeph Herrenberger.
Hr. Lt. Johann Simon Klein.
Hr. Jacob Galler.
Hr. Lt. Joh. Franz Ehrmann.
Hr. Lt. Joh. Friedrich Lauth.
Hr. Franz Placidus Claus.
Hr. Lt. Anton Lepicq.
Hr. Lt. Joh. Chriſtian Treitlinger.
Hr. Lt. Jacob Friedrich Brackenhoffer.
Hr. Lt. Jacob Ludwig Albert.
Hr. Lt. Auguſtin Meinhard Anton Lachauſſe.
Hr. Joſeph Herrmann Duelly.

Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

Data Types: Ben Shneiderman

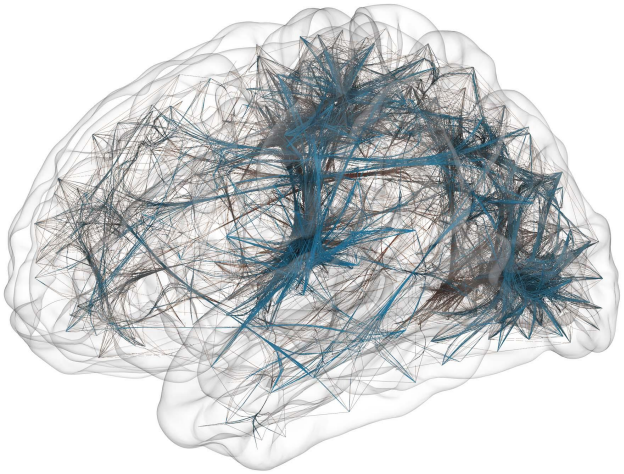
- 1-dimensional
- **2-dimensional**
- 3-dimensional
- Multidimensional
- Temporal
- Trees
- Networks

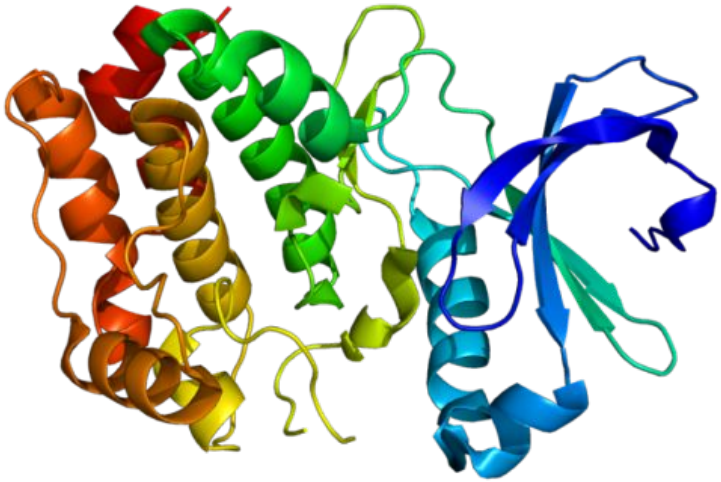


Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

Data Types: Ben Shneiderman

- 1-dimensional
- 2-dimensional
- **3-dimensional**
- Multidimensional
- Temporal
- Trees
- Networks





Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

Data Types: Ben Shneiderman

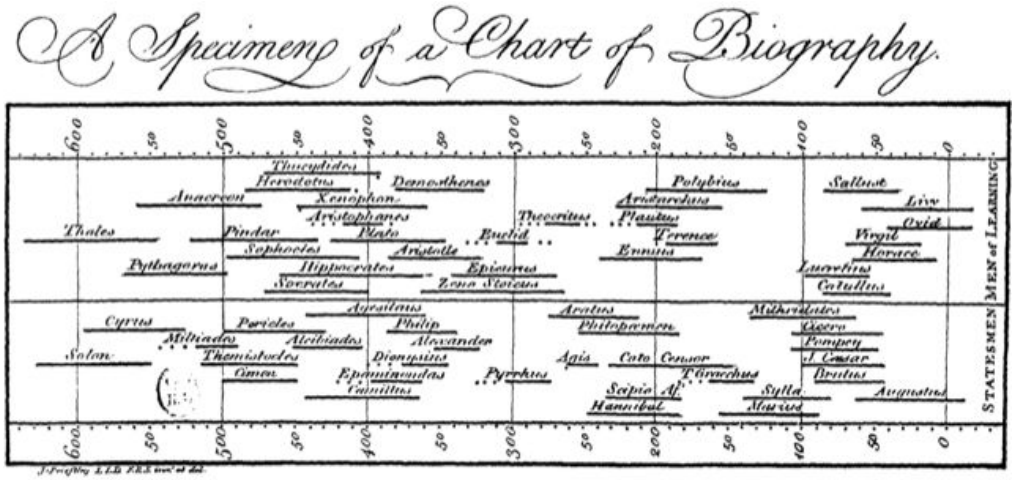
- 1-dimensional
- 2-dimensional
- 3-dimensional
- **Multidimensional**
- Temporal
- Trees
- Networks

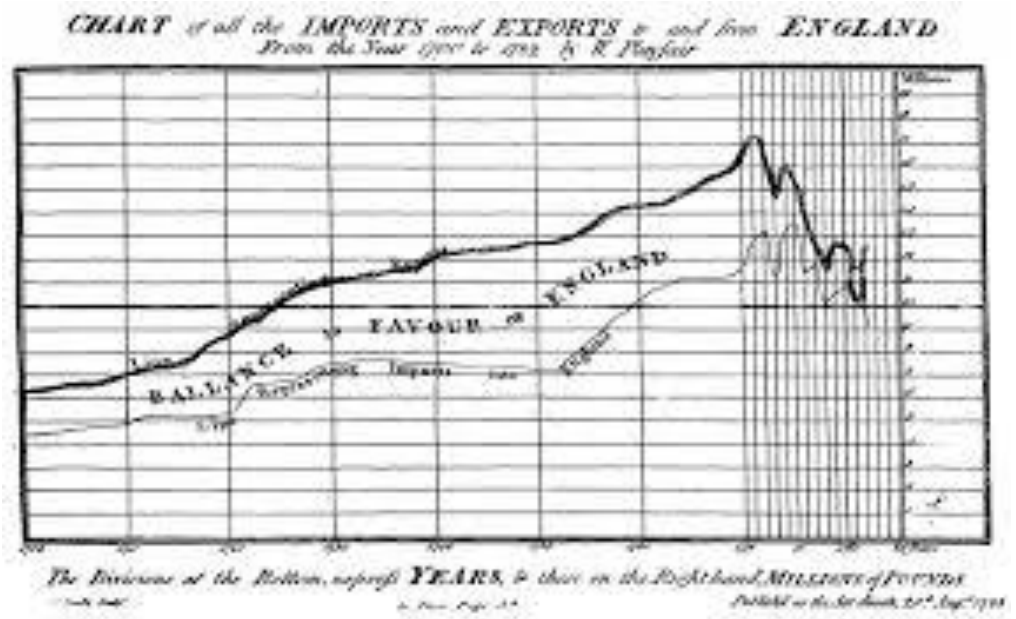
total_bill	tip	sex	smoker	day
16.99	1.01	Female	No	Sun
10.34	1.66	Male	No	Sun
21.01	3.5	Male	No	Sun
23.68	3.31	Male	No	Sun
24.59	3.61	Female	Yes	Sun
25.29	4.71	Male	No	Sun
8.77	2	Male	No	Sun
26.88	3.12	Male	No	Sun
15.04	1.96	Male	Yes	Sun
14.78	3.23	Male	Yes	Sun
10.27	1.71	Male	Yes	Sun
35.26	5	Female	No	Sun

Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

Data Types: Ben Shneiderman

- 1-dimensional
- 2-dimensional
- 3-dimensional
- Multidimensional
- **Temporal**
- Trees
- Networks

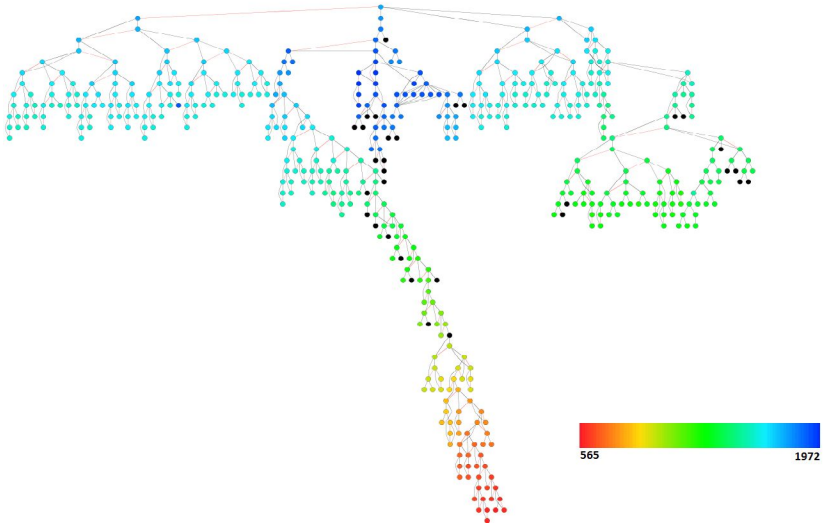


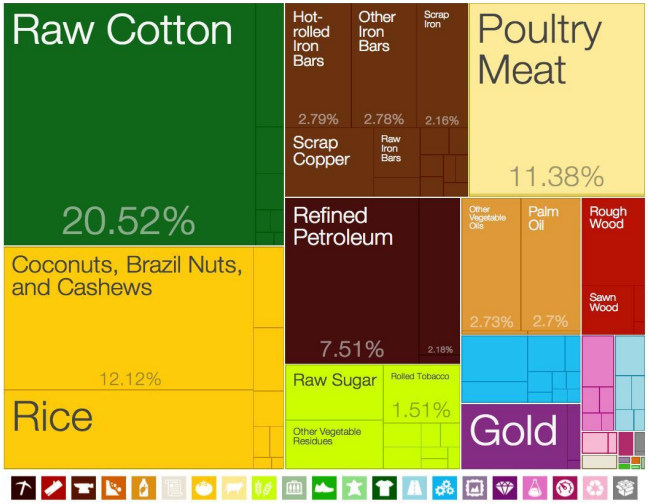


Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

Data Types: Ben Shneiderman

- 1-dimensional
- 2-dimensional
- 3-dimensional
- Multidimensional
- Temporal
- **Trees**
- Networks





Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

Data Types: Ben Shneiderman

- 1-dimensional
- 2-dimensional
- 3-dimensional
- Multidimensional
- Temporal
- Trees
- **Networks**

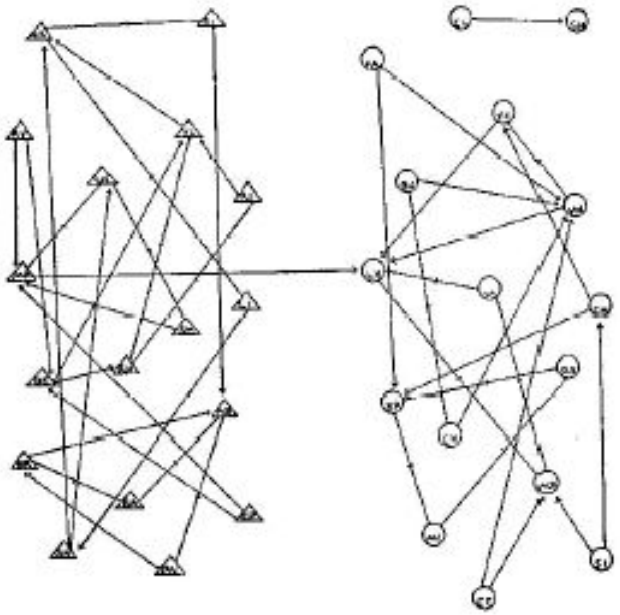
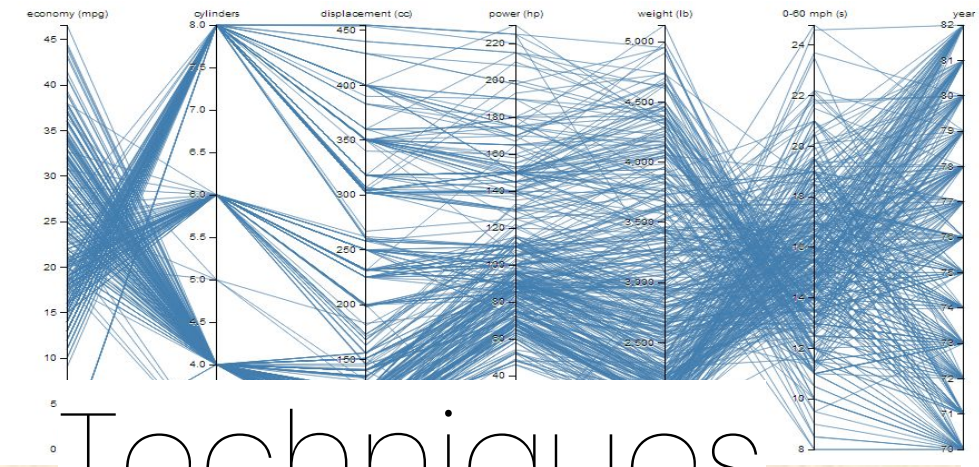
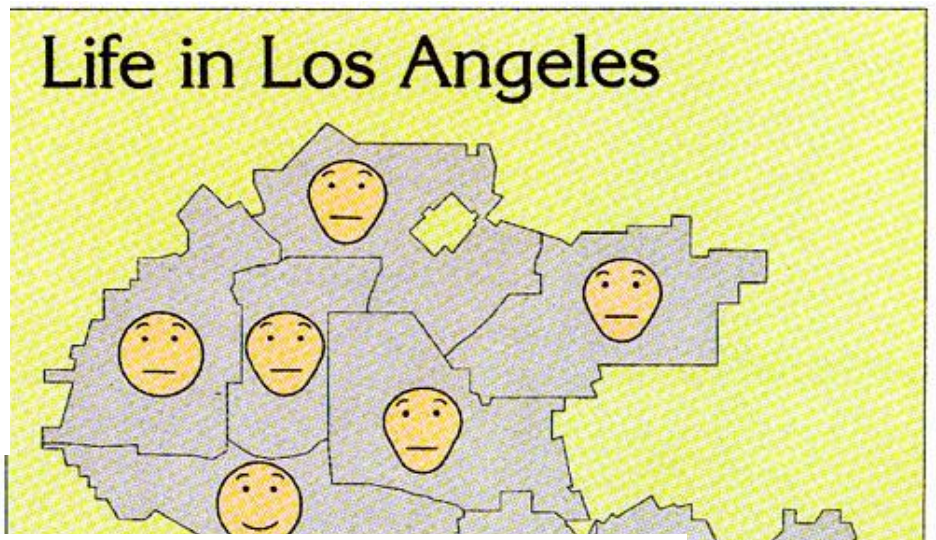


Figure 1. An Attraction Network in a Fourth Grade Class (from Moreno [19], p. 38).

Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

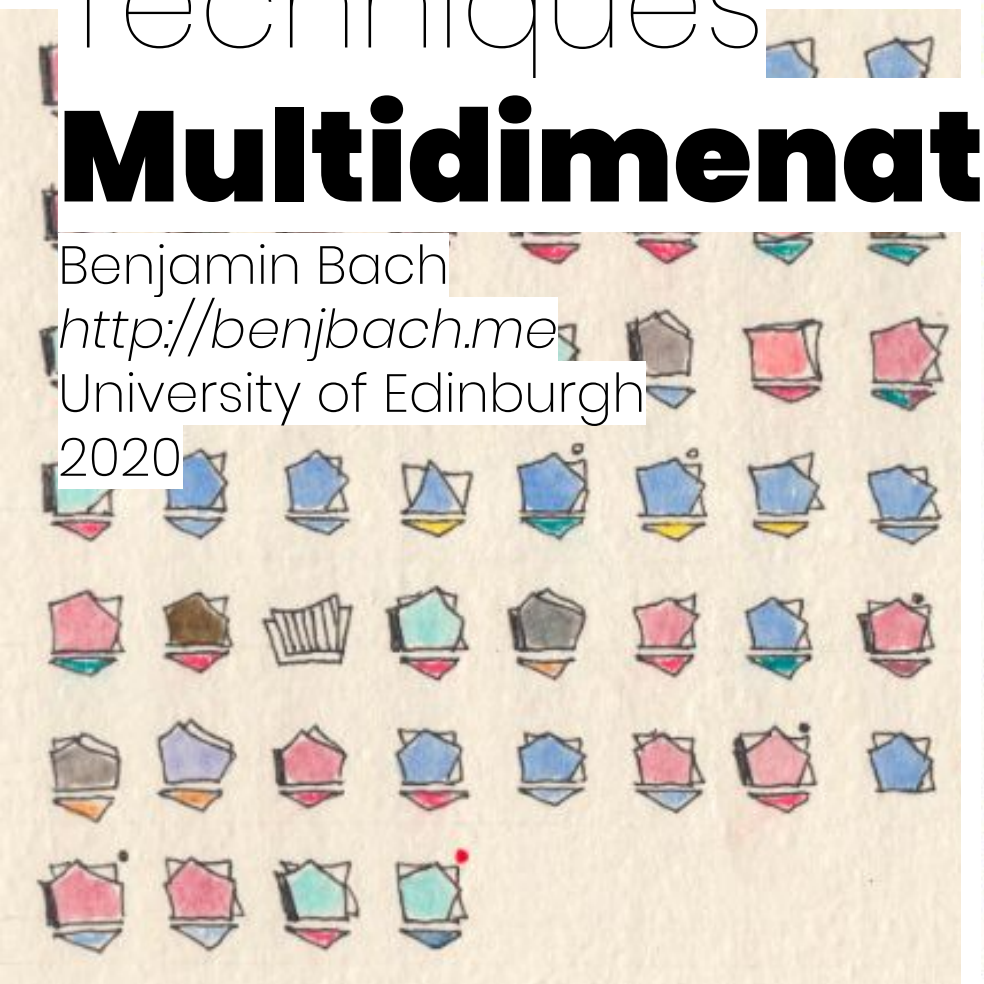


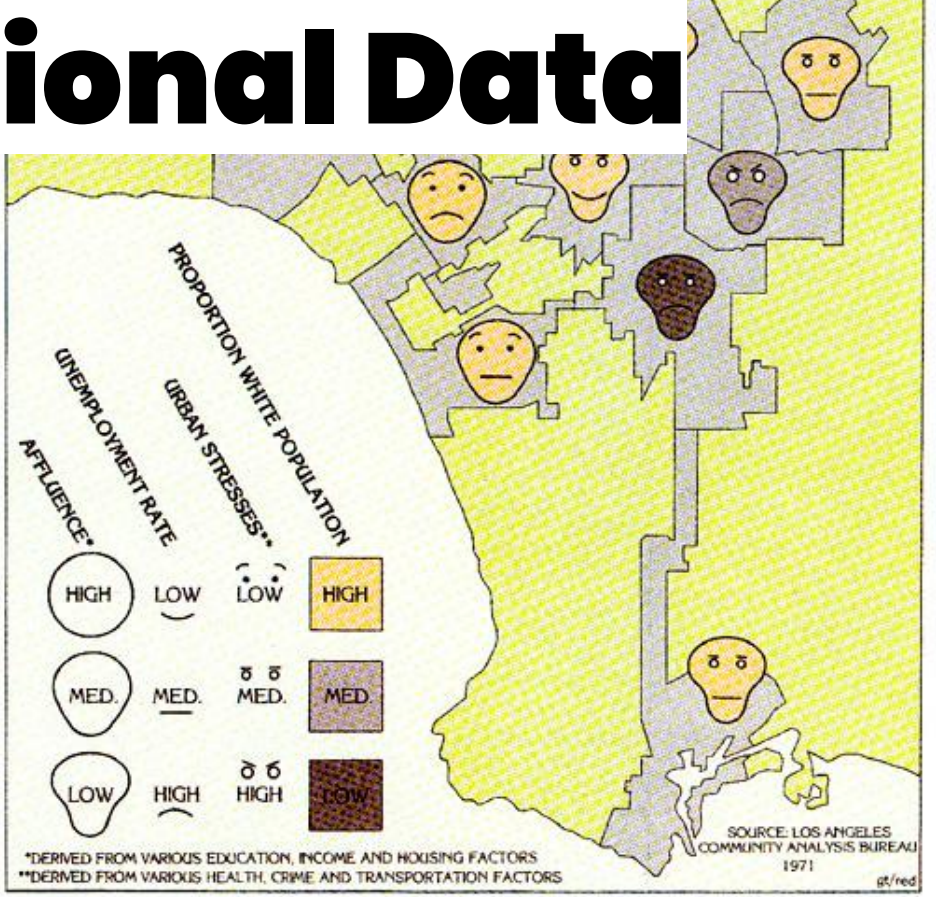


Techniques

Multidimenational Data

Benjamin Bach
http://benjbach.me
University of Edinburgh
2020



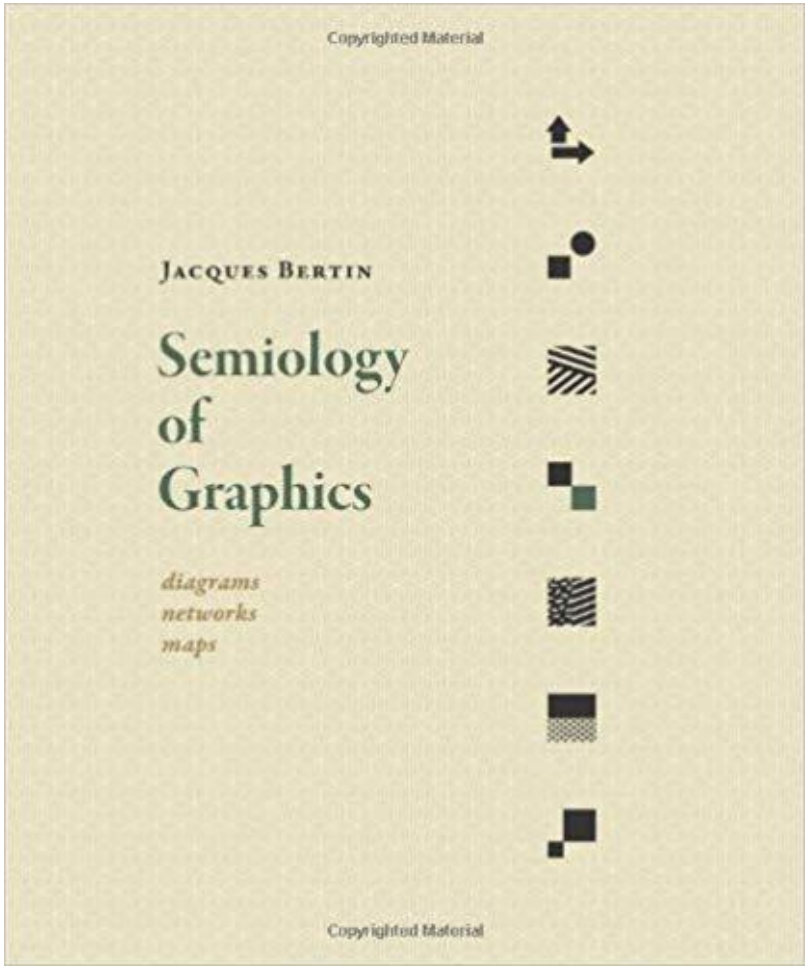


Multidimensional data

total_bill	tip	sex	smoker	day	time	size
16.99	1.01	Female	No	Sun	Dinner	2
10.34	1.66	Male	No	Sun	Dinner	3
21.01	3.5	Male	No	Sun	Dinner	3
23.68	3.31	Male	No	Sun	Dinner	2
24.59	3.61	Female	Yes	Sun	Lunch	4
25.29	4.71	Male	No	Sun	Lunch	4
8.77	2	Male	No	Sun	Dinner	2
26.88	3.12	Male	No	Sun	Dinner	4
15.04	1.96	Male	Yes	Sun	Dinner	2
14.78	3.23	Male	Yes	Sun	Dinner	2
10.27	1.71	Male	Yes	Sun	Dinner	2
35.26	5	Female	No	Sun	Lunch	4

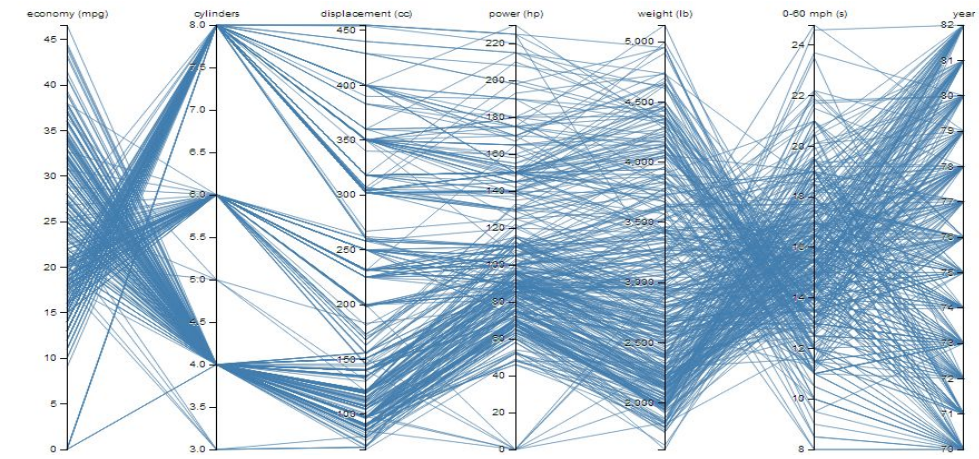
Data types: **Jaques Bertin**

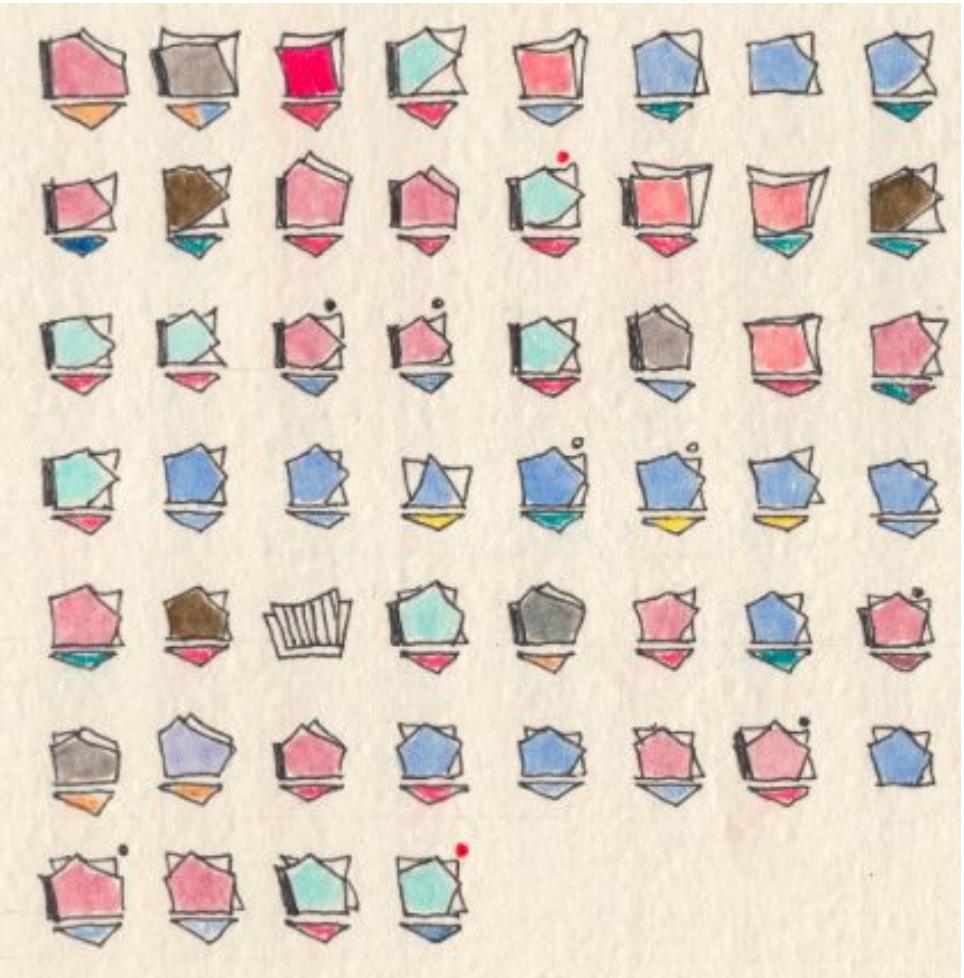
- Quantitative data **Q**
- Ordinal **O**
- Categorical **C**

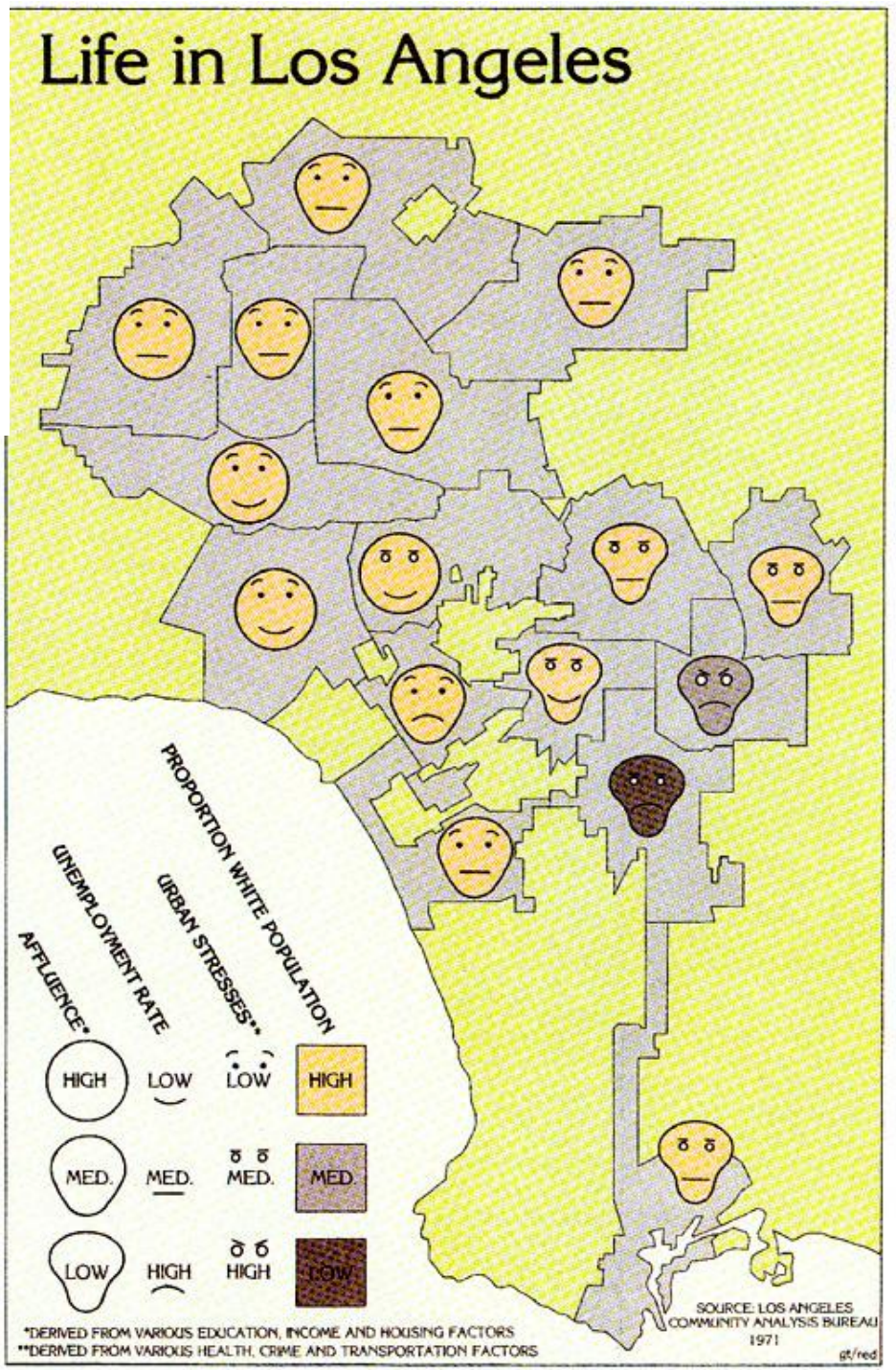


Multivariate data

Q	Q	C	C	Q	O	Q
total_bill	tip	sex	smoker	day	time	size
16.99	1.01	Female	No	Sun	Dinner	2
10.34	1.66	Male	No	Sun	Dinner	3
21.01	3.5	Male	No	Sun	Dinner	3
23.68	3.31	Male	No	Sun	Dinner	2
24.59	3.61	Female	Yes	Sun	Lunch	4
25.29	4.71	Male	No	Sun	Lunch	4
8.77	2	Male	No	Sun	Dinner	2
26.88	3.12	Male	No	Sun	Dinner	4
15.04	1.96	Male	Yes	Sun	Dinner	2
14.78	3.23	Male	Yes	Sun	Dinner	2
10.27	1.71	Male	Yes	Sun	Dinner	2
35.26	5	Female	No	Sun	Lunch	4

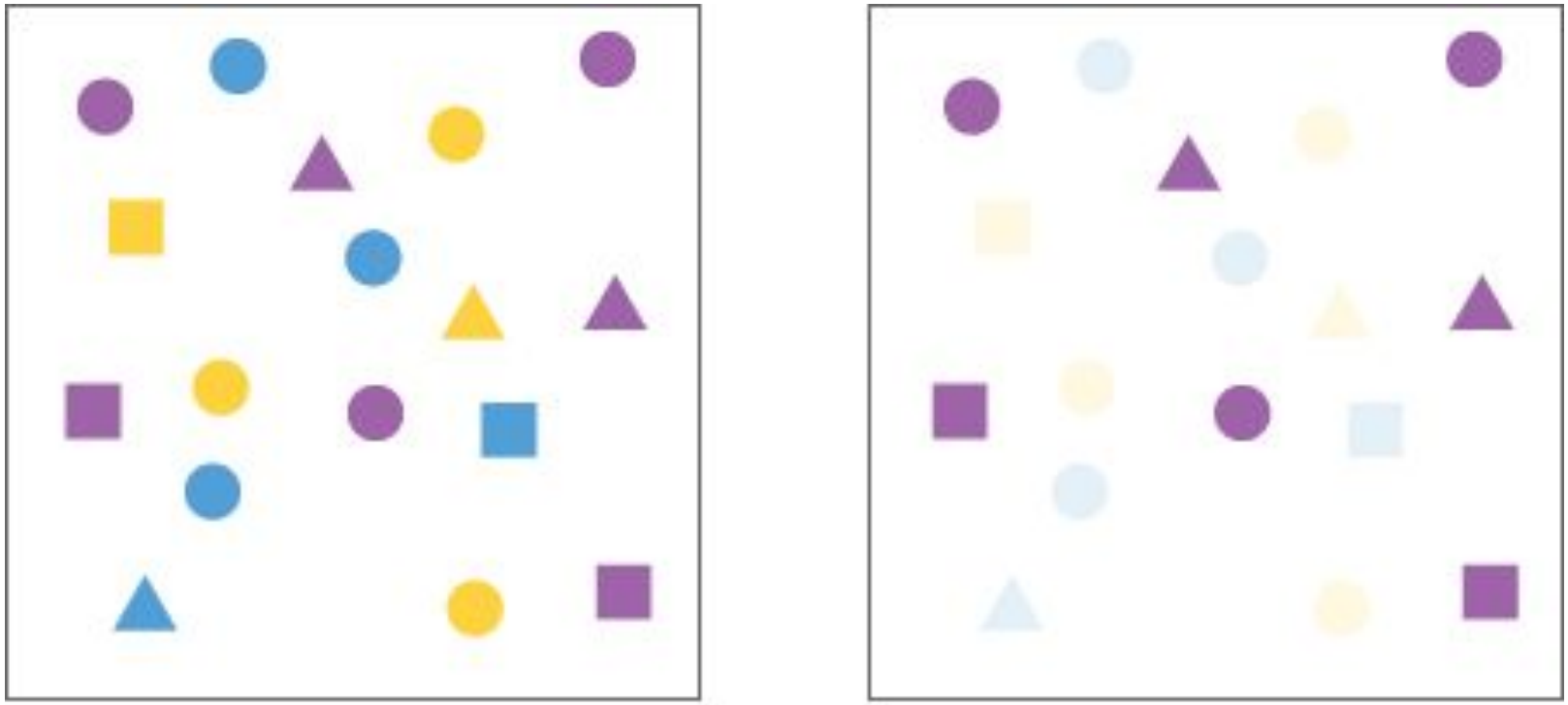

economy (mpg)
cylinders
displacement (cc)
power (hp)
weight (lb)
0-60 mph (s)
year




Life in Los Angeles
PROPORTION WHITE POPULATION
URBAN STRESSES**
UNEMPLOYMENT RATE
AFFLUENCE*
HIGH
LOW
LOW
HIGH
MED.
MED.
MED.
MED.
LOW
HIGH
HIGH
LOW
*DERIVED FROM VARIOUS EDUCATION, INCOME AND HOUSING FACTORS
**DERIVED FROM VARIOUS HEALTH, CRIME AND TRANSPORTATION FACTORS
SOURCE: LOS ANGELES COMMUNITY ANALYSIS BUREAU 1971
gt/red

Outline

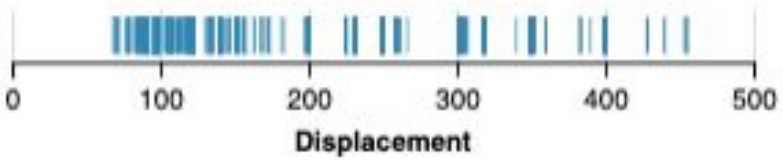
- Low-dimensions
- Higher dimensions
- Dimensionality reduction



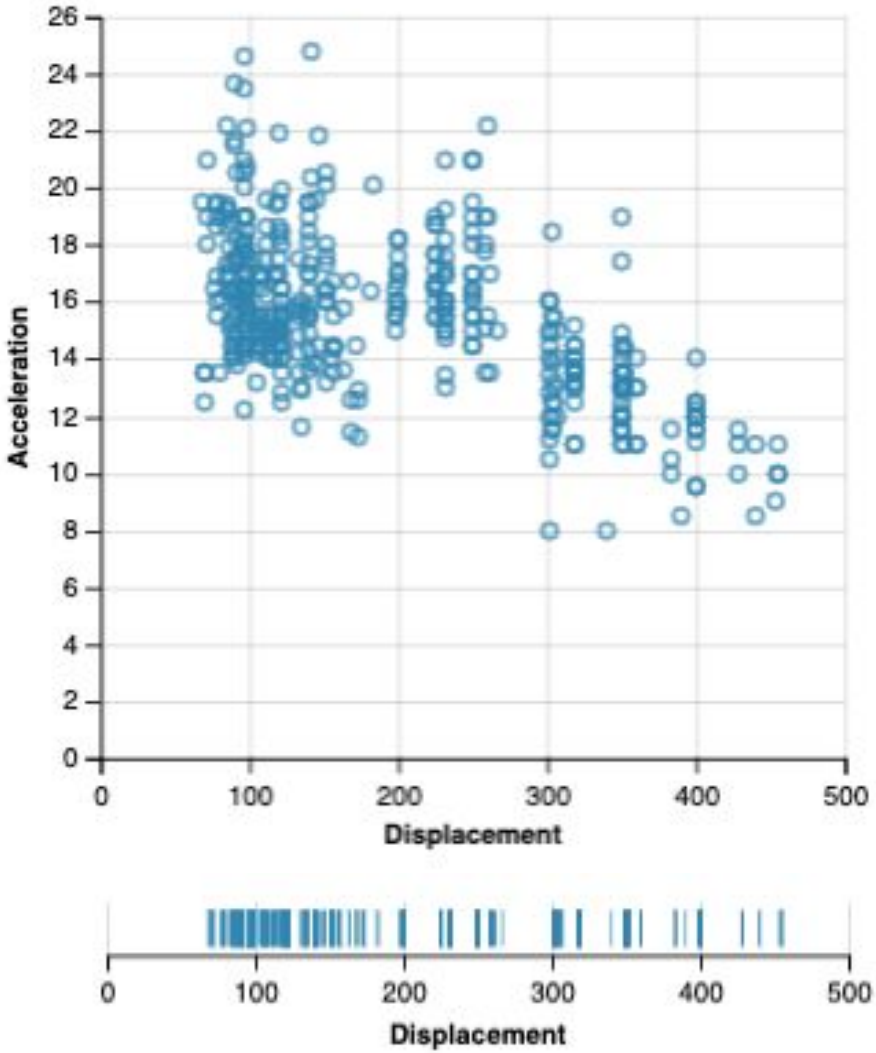
Symbols can be perceived as a group despite variation in shape, e.g., these similarly sized purple symbols appear as a group. Shape is therefore *associative*.

https://www.axismaps.com/guide/general/visual-variables/

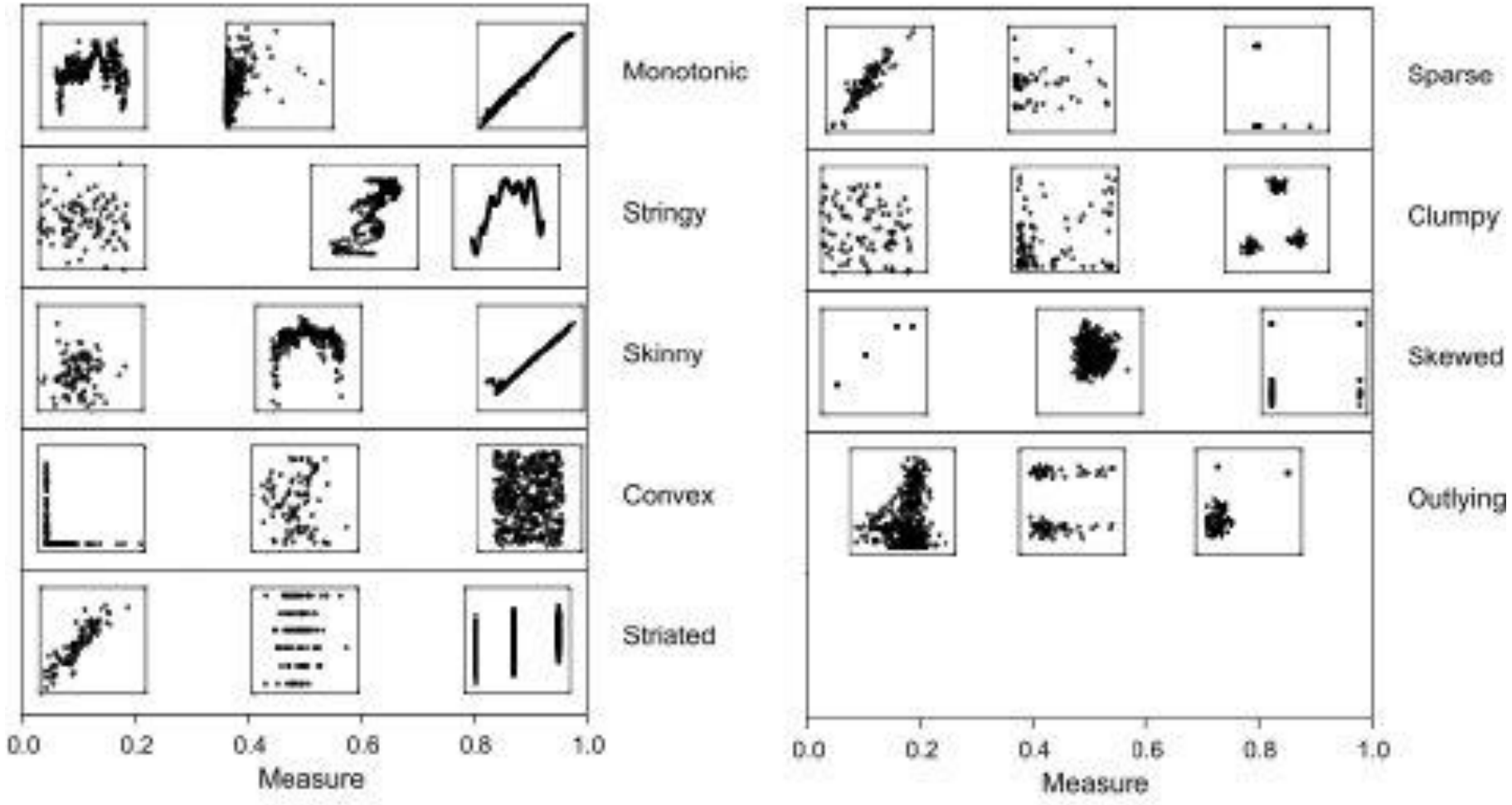
1 Dimension



2 Dimensions

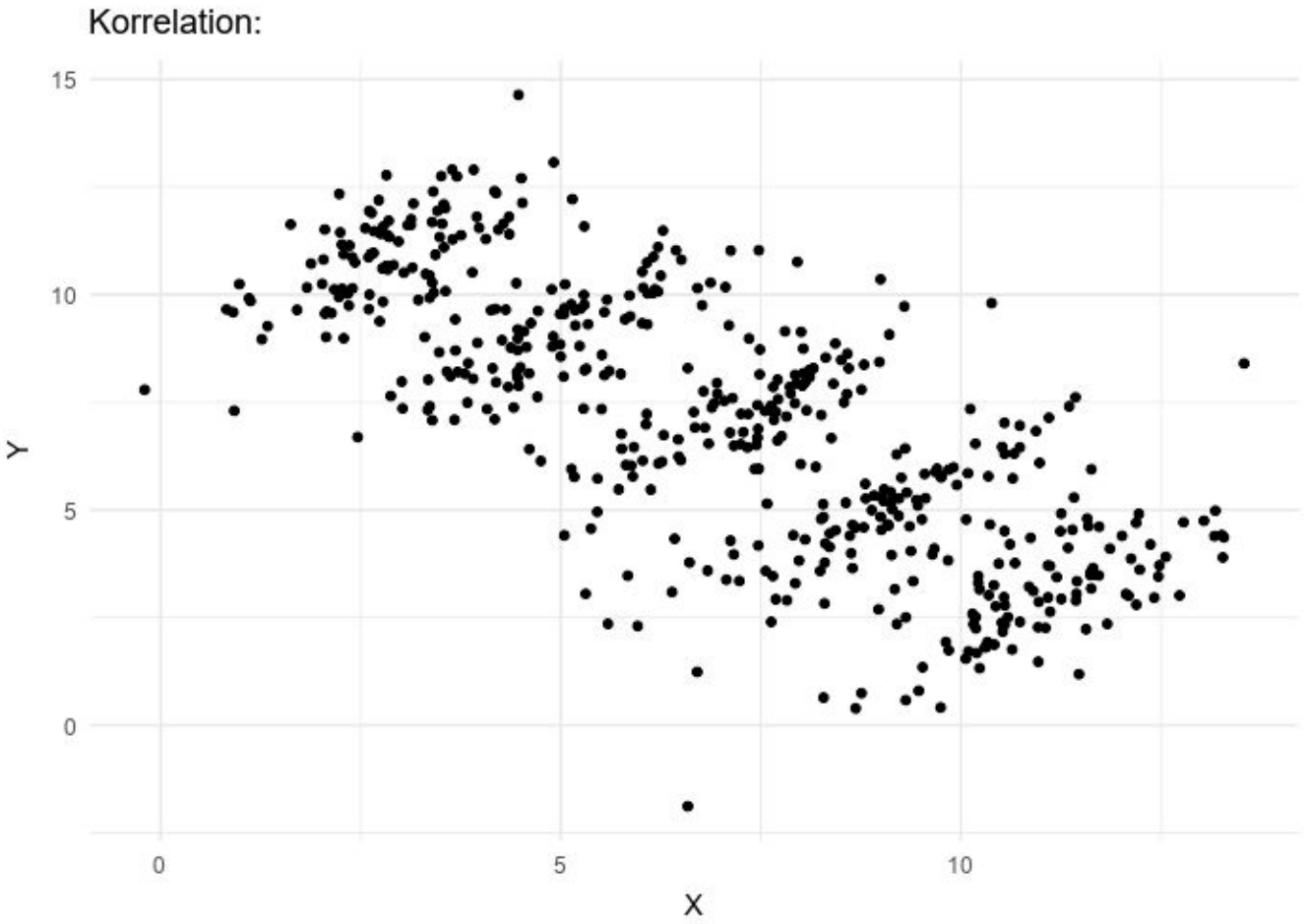


Quant + Quant: **Scatterplot**

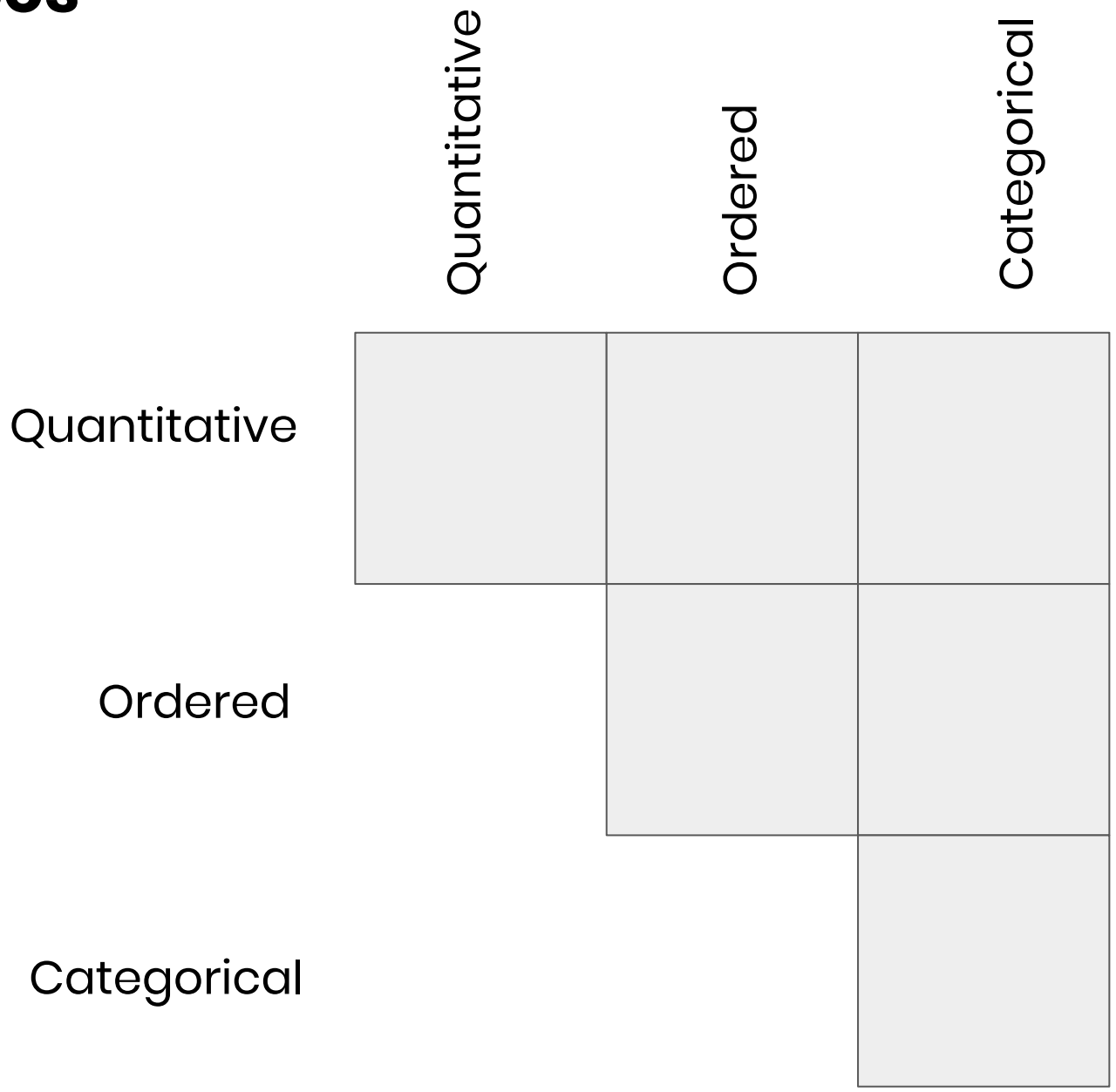


Simpson's Paradox

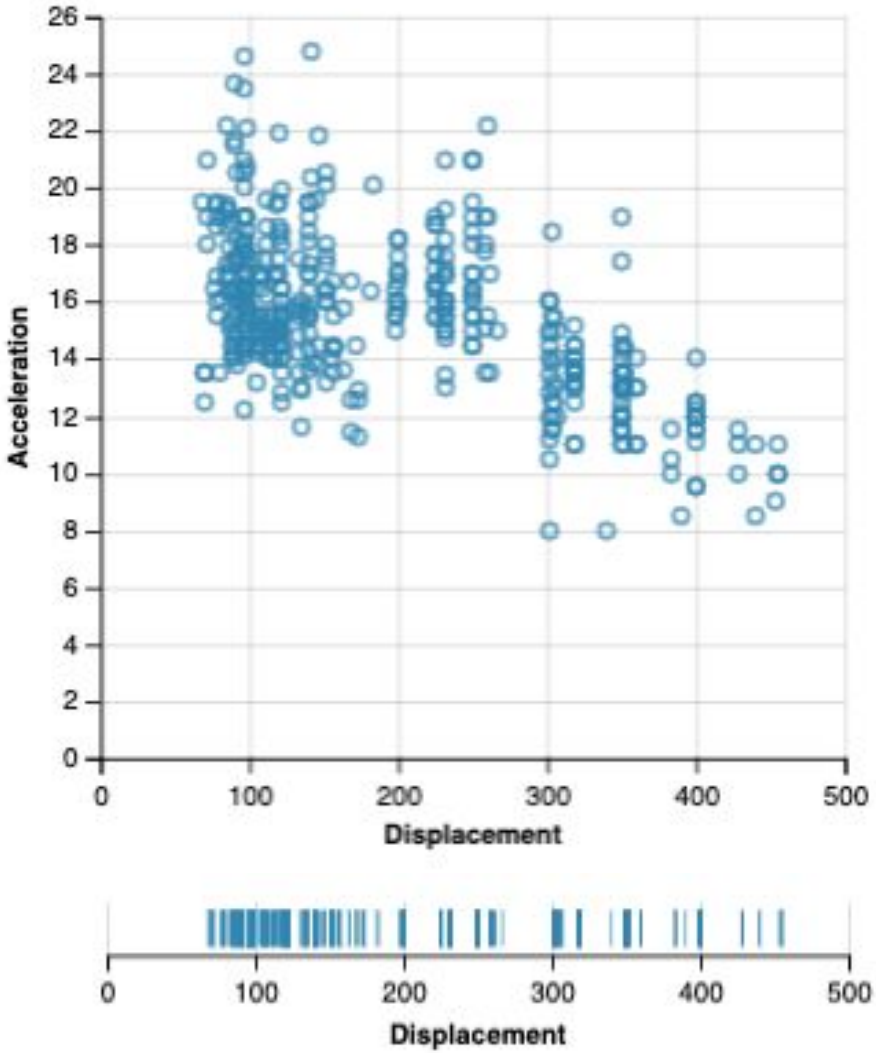
a trend appears in several different groups of data but disappears or reverses when these groups are combined.



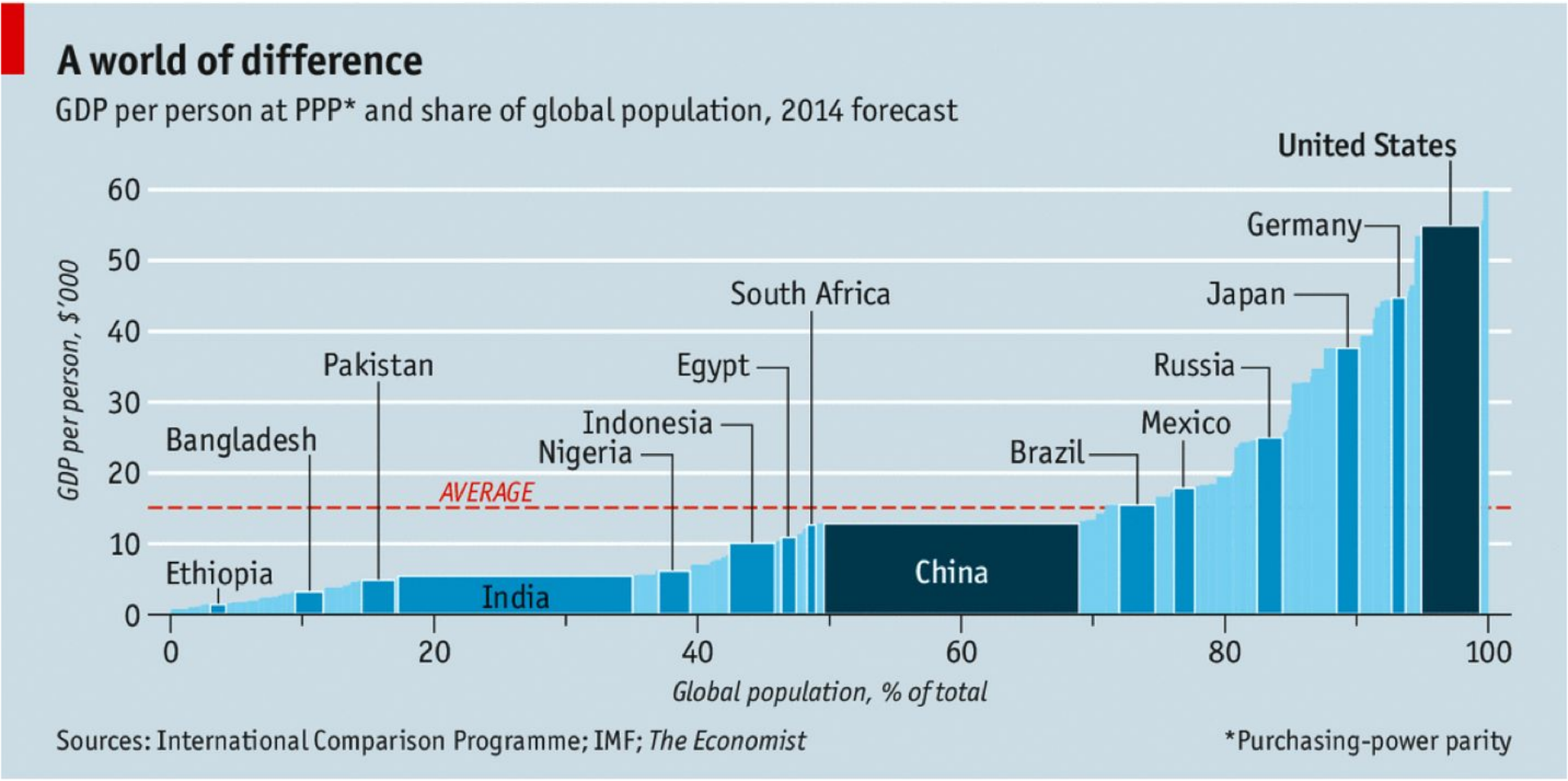
Data types


Quantitative
Ordered
Categorical
Quantitative
Ordered
Categorical

Quant x Quant: **Scatterplot**

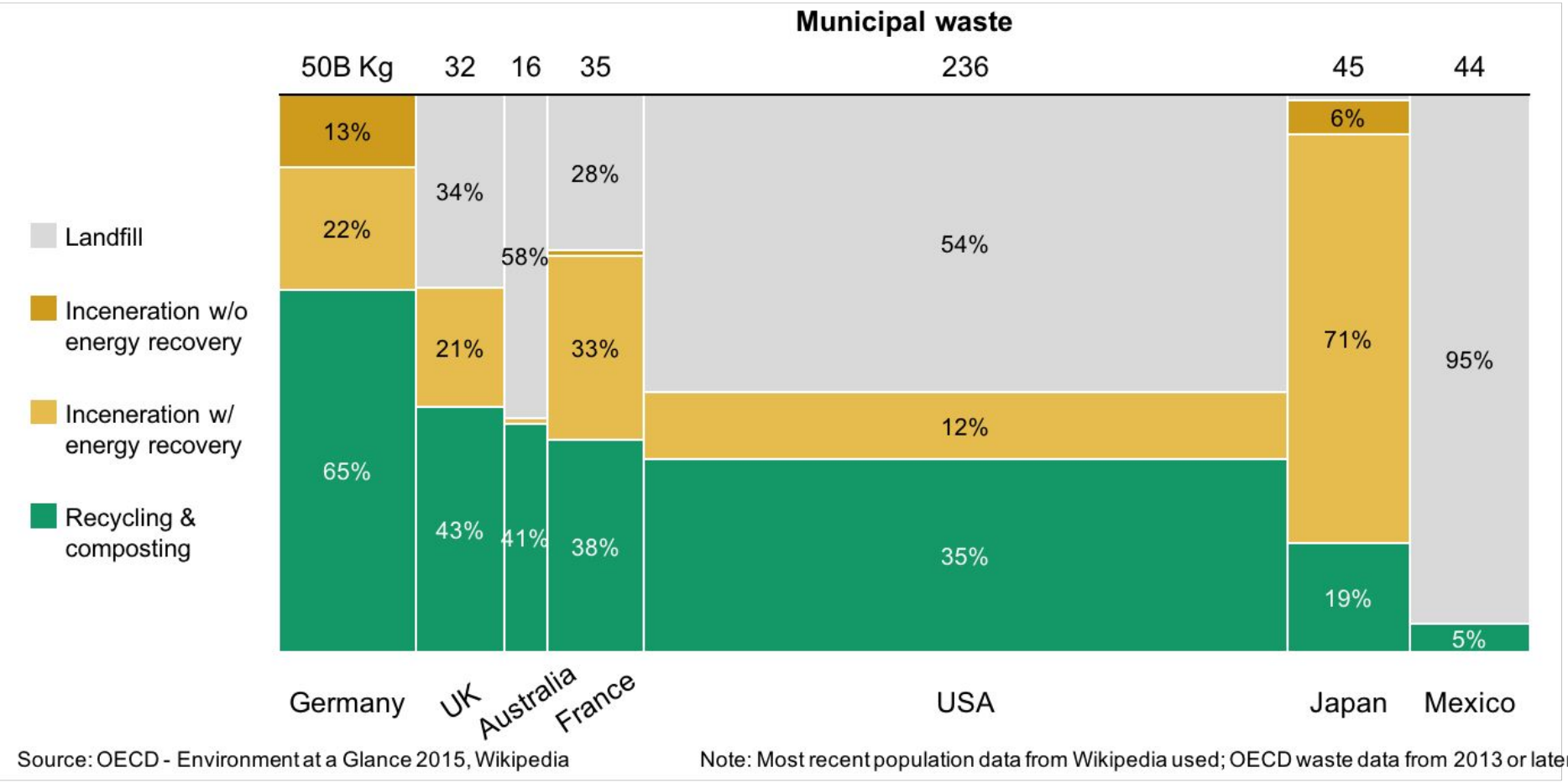


Quant x Quant: **Mekko Chart**



Two quantities that if multipiled give a third quantity.

Quant x Quant: **Mekko Chart**

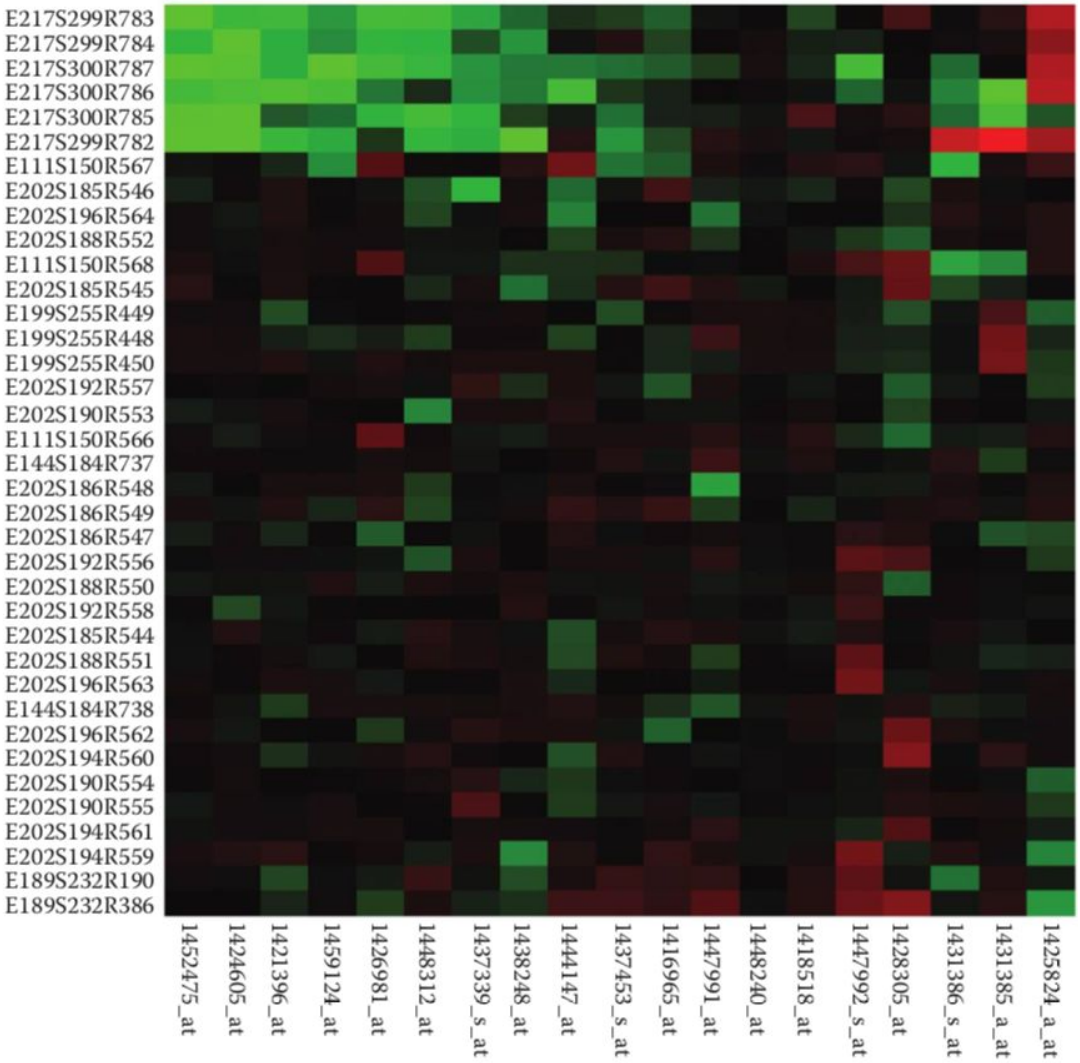


Two quantities that if multiplied give a third quantity.

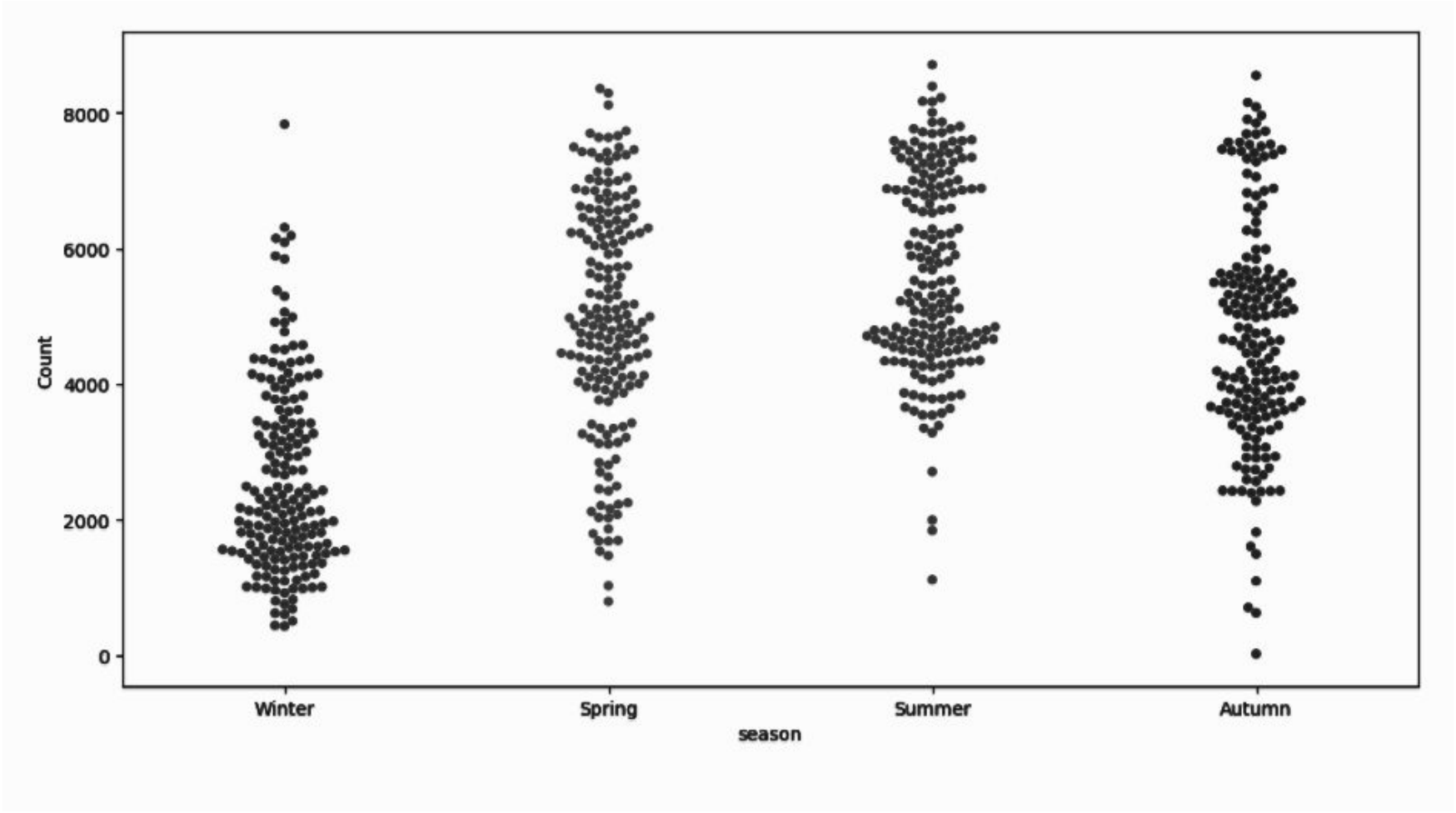
Quant + Ordered/Categorical: **Heatmaps**

A quantity (cells), depending on two factors (ordered or categorical)

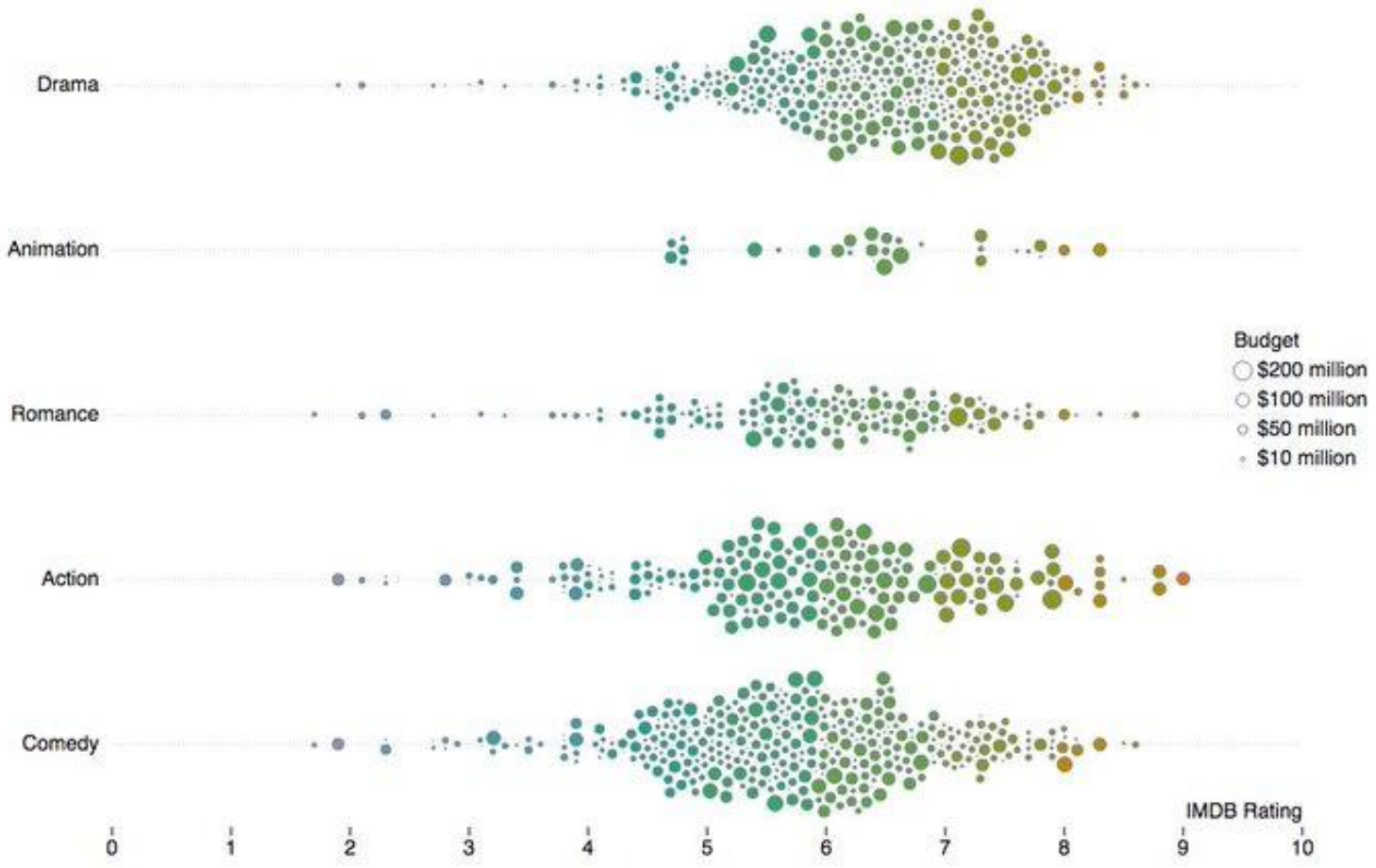
Requires complete field.



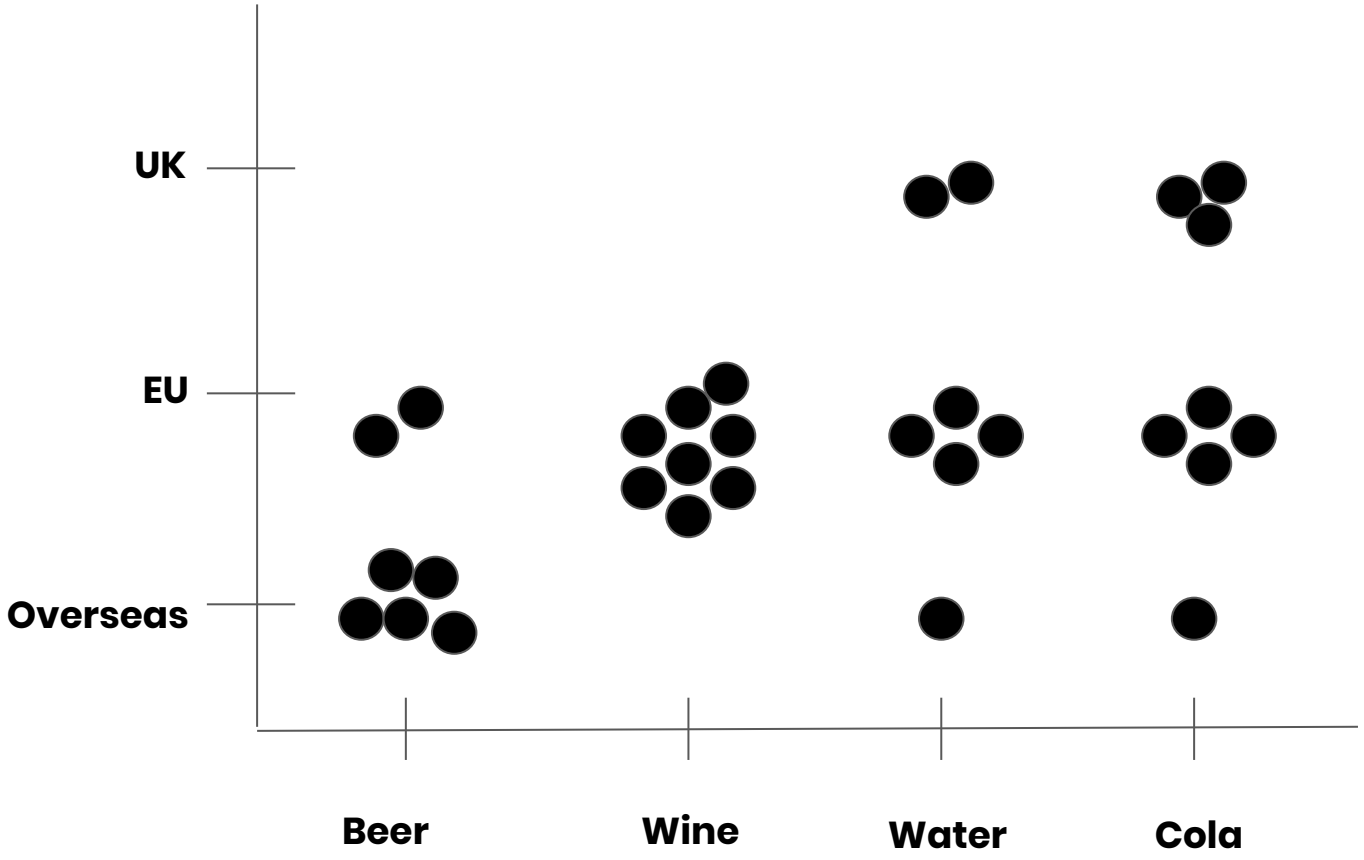
Quant + Ordered/Categorical: **Beeplots**



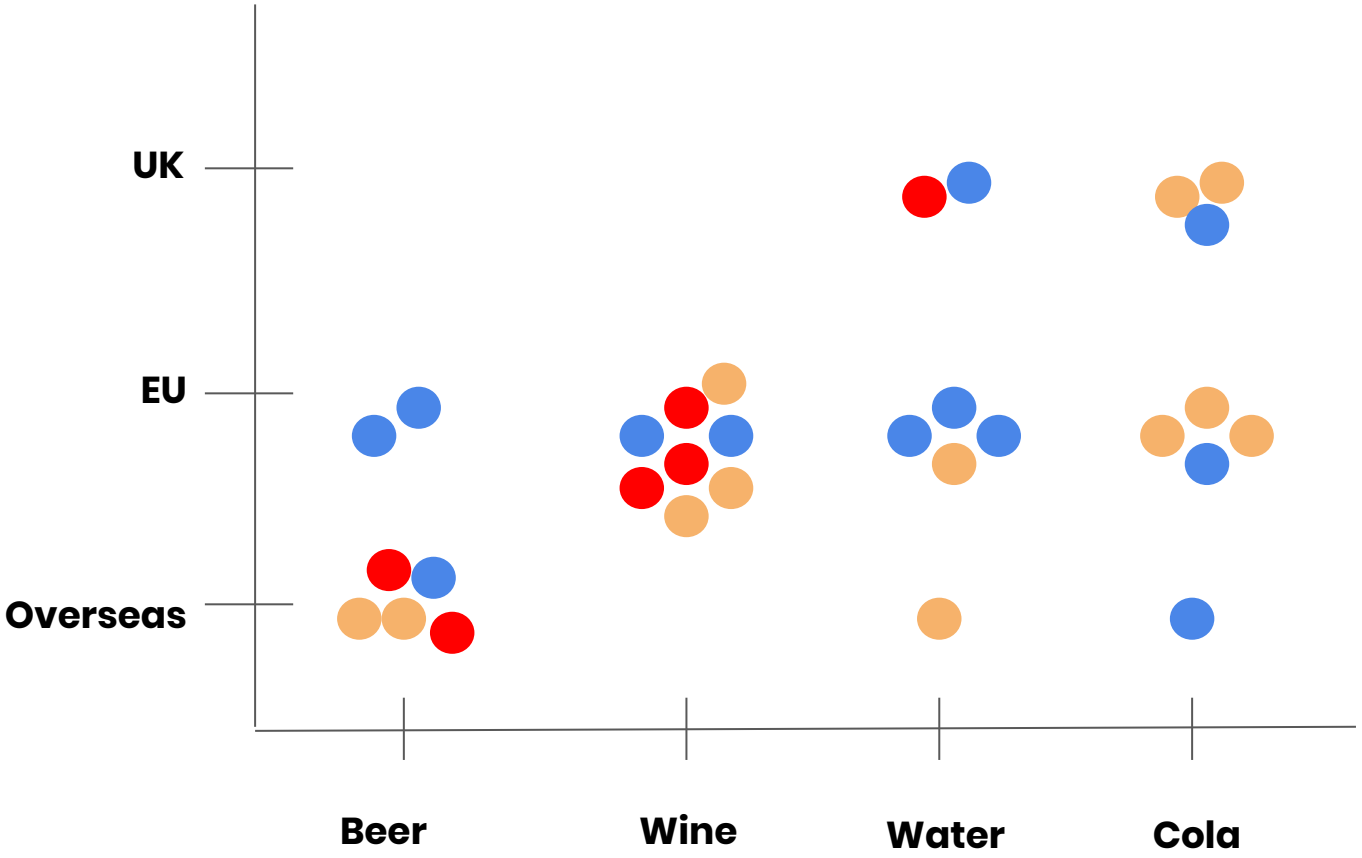
Quant + Categorical: **Swarm plots**



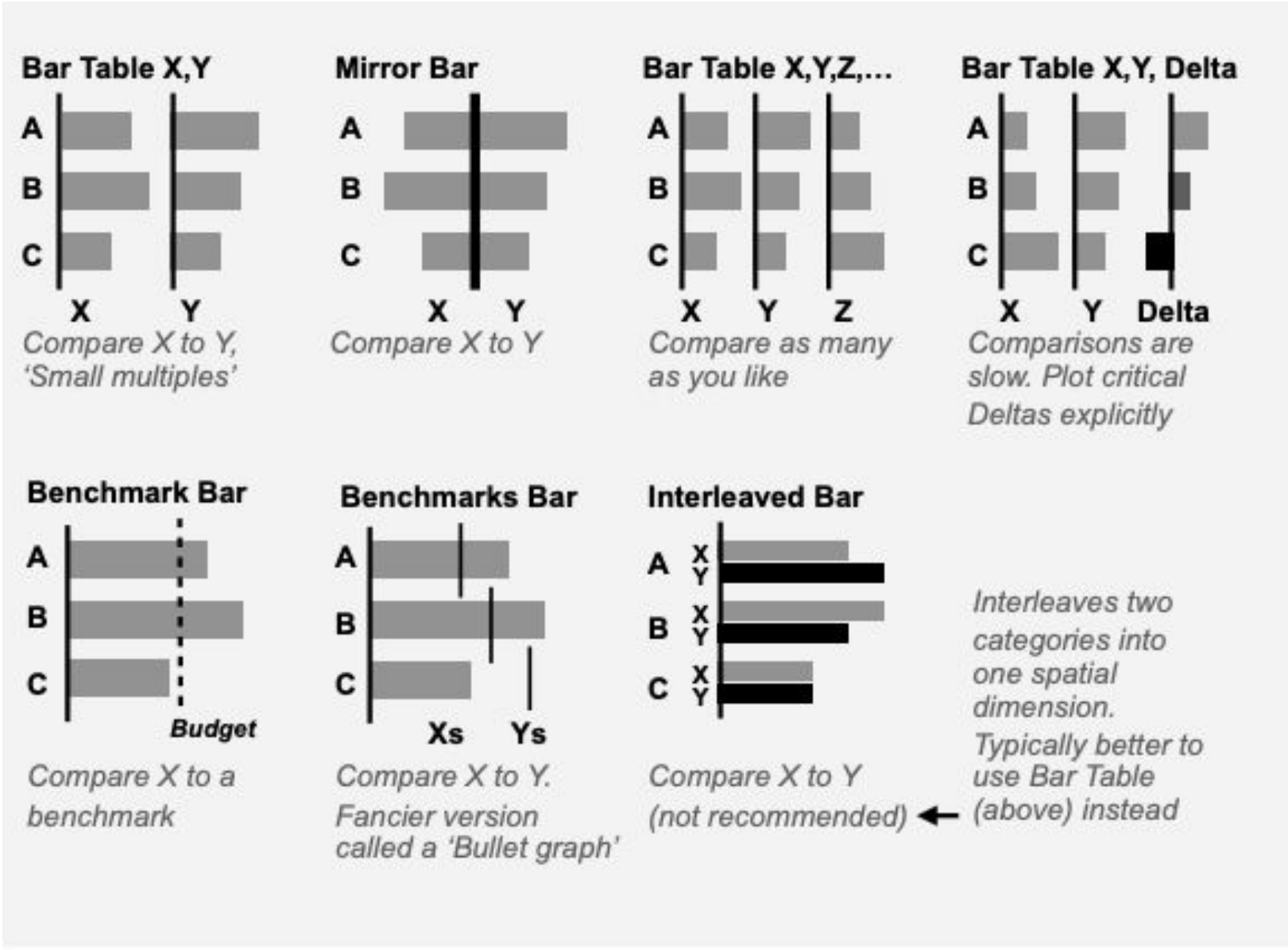
Categorical + Categorical: ***Swarm plots???***



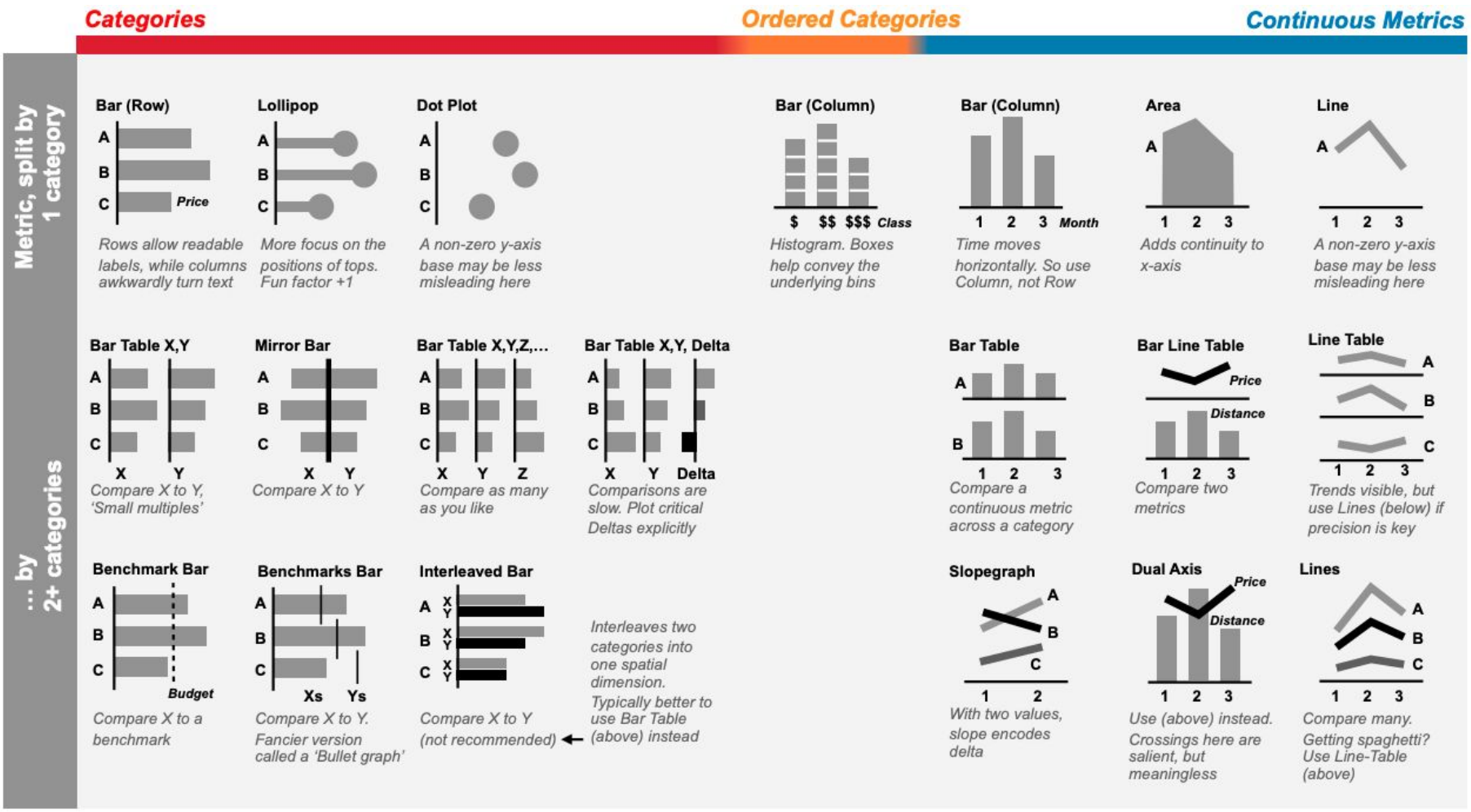
Categorical + Categorical: ***Swarm plots???***



And again: **bar charts**



… and many more!



http://experception.net/

… and much more!!!

… by Part-to-Whole, split by 2+ categories

Pie

Screams 'Percentages!'

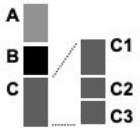
Stacked Bar (Row)

More precise and flex, but less screaming

Stacked Bar (Col)

Now I'm standing

Breakout Bar



Let's zoom in here. User different colors. Global starts left (or top)

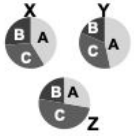
Waterfall

Waterfalls are vertical stacked bars that narrate financial values in a (typically) artificially imposed ordering

These lines are identical, with equal vertical separation at each X

Beware of an illusion for these: seeing differences (lines), or category values (stacked area) can be difficult and even misleading

MultiPie

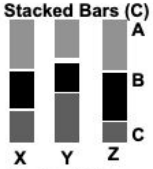


Please don't… (not recommended)

Stacked Bars (R)

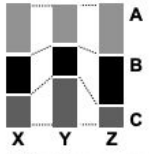
…Use this instead

Stacked Bars (C)



Horizontal flow implies an ordering

… with lines



Added lines suggest continuity, help depict changes

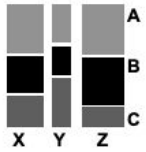
No legends. Instead, directly label actual values

Stacked area

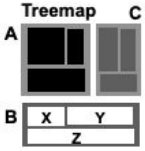
Now it's definitely continuous

Mekko stacks in X and Y simultaneously. Here XYZ might be absolute values of a market, and ABC are company market shares as %'s

Mekko



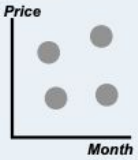
Treemap



Hierarchy, ~3 levels max. Size & color code different metrics. Often misused (info)

Metrics: relationships to other metrics

Scatter



A most powerful chart

Parallel Coordinates

Instead of Perpendicular Cartesian axes, Parallel format allows more axes

Look at this number *Just look at it.*

Dot Array

Dot Array %

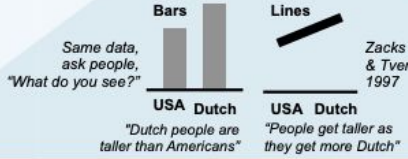
Icon Array (ISOTYPE, info)

Icon Array %

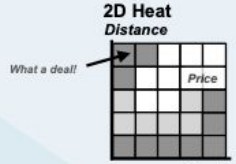
43%

Huge Number

Connecting Lines imply continuous data

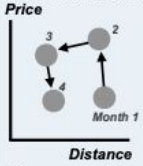


2D Heat



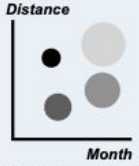
2D Histogram

Connected Scatter



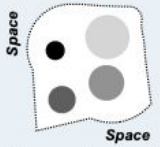
The power of a scatterplot, with points connected by time (info)

Rosling Scatter



Watch Hans Rosling's famous TED talk. He adds two more metrics, and moves it over time

Map



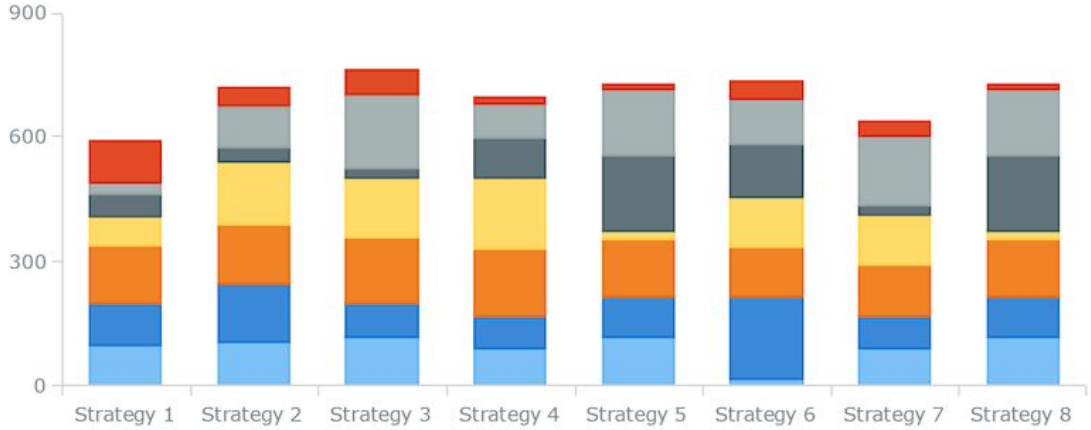
Maps and scatterplots share the same DNA
Color = Flight Time
Size = Price

v1.2 Download the most recent version at ExperCeption.net

@SteveFranconeri

Stacked Bar Charts

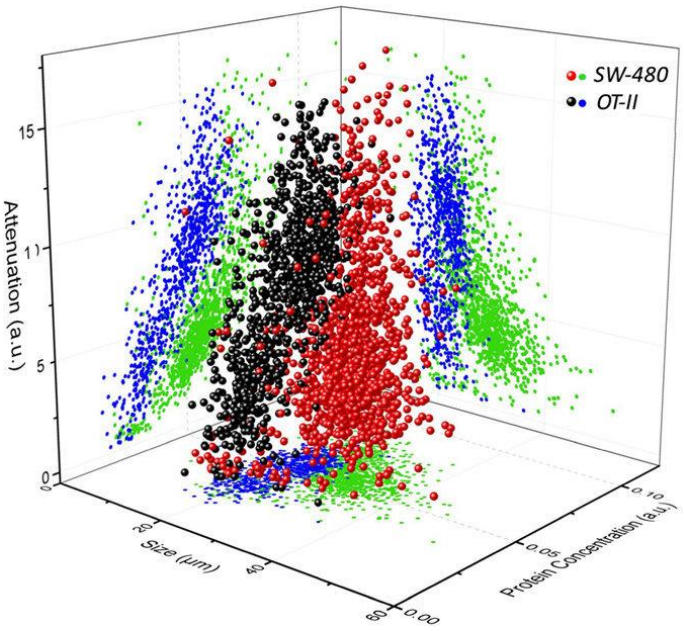
- Individual distributions
- Individual value comparison
- Details

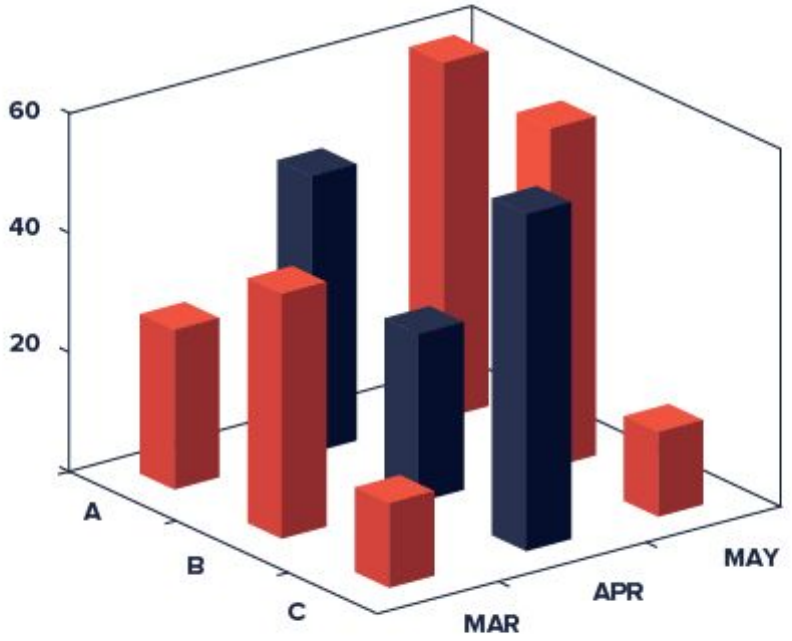


- Sum comparison
- Overview

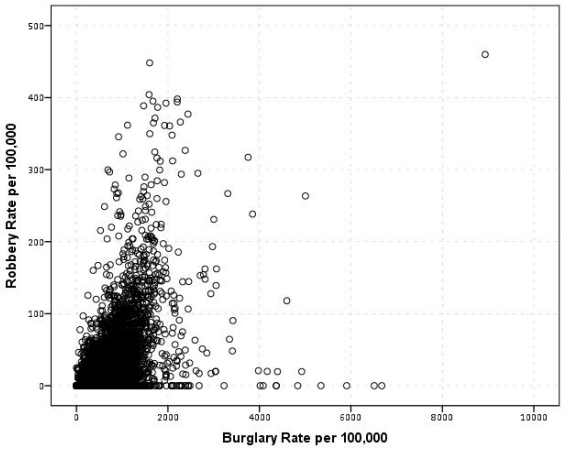
> requires good color pallette!

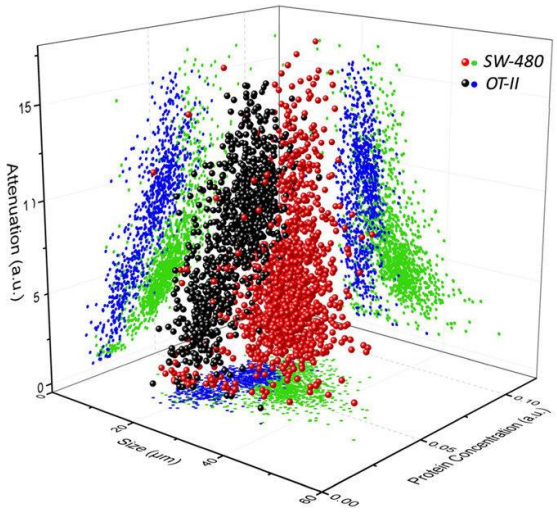
3 Dimensions...





...4 Dimensions?

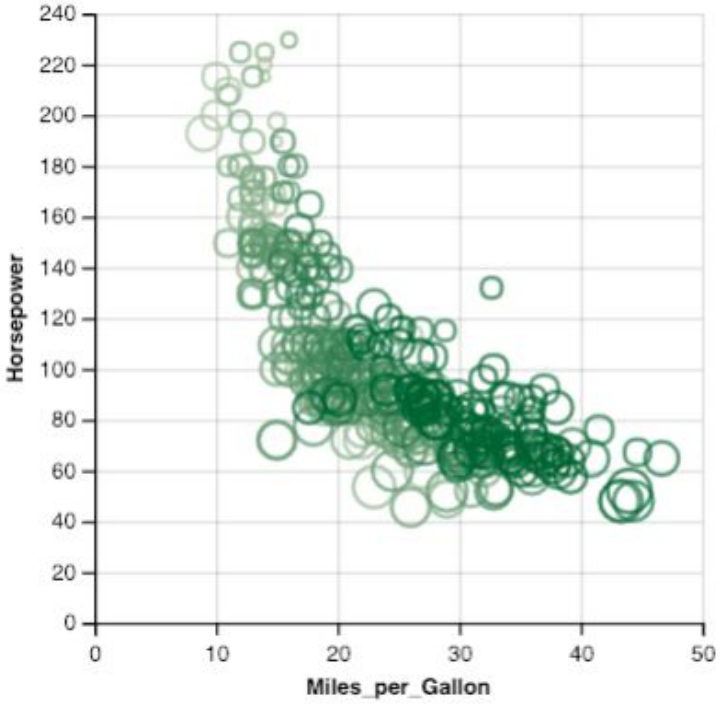




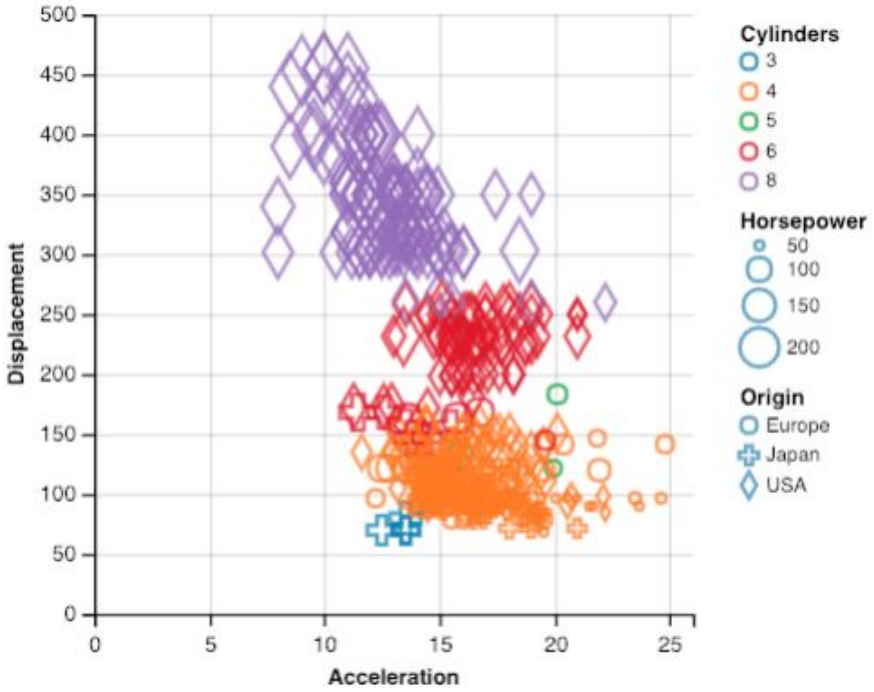
?

2-dimensional **3-dimensional** **3-dimensional**

1. Additional Visual Variable(s)

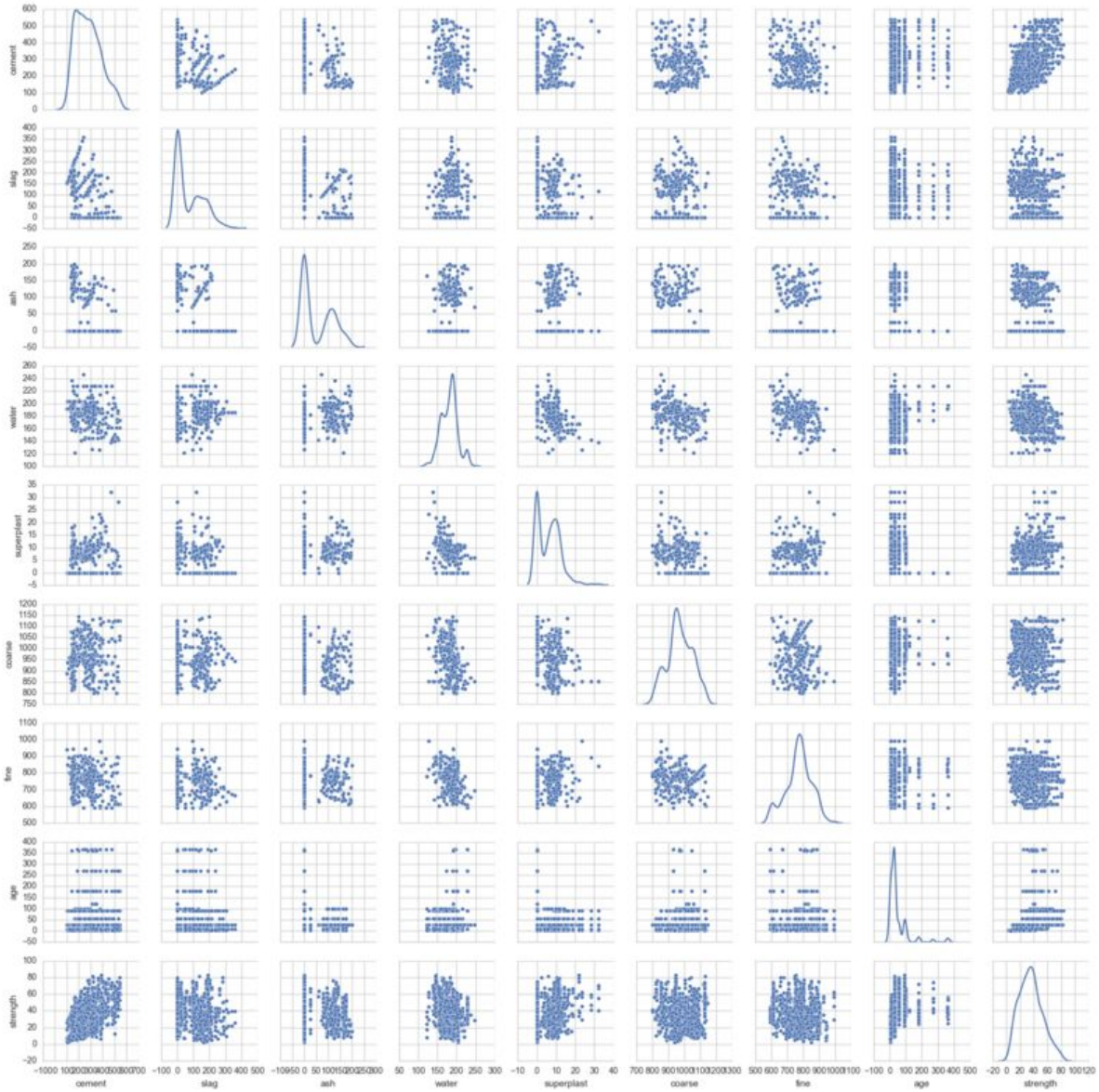


2-dimensional
+ 1 dimension (size)



2-dimensional
+ 2 dimension
(shape + color)

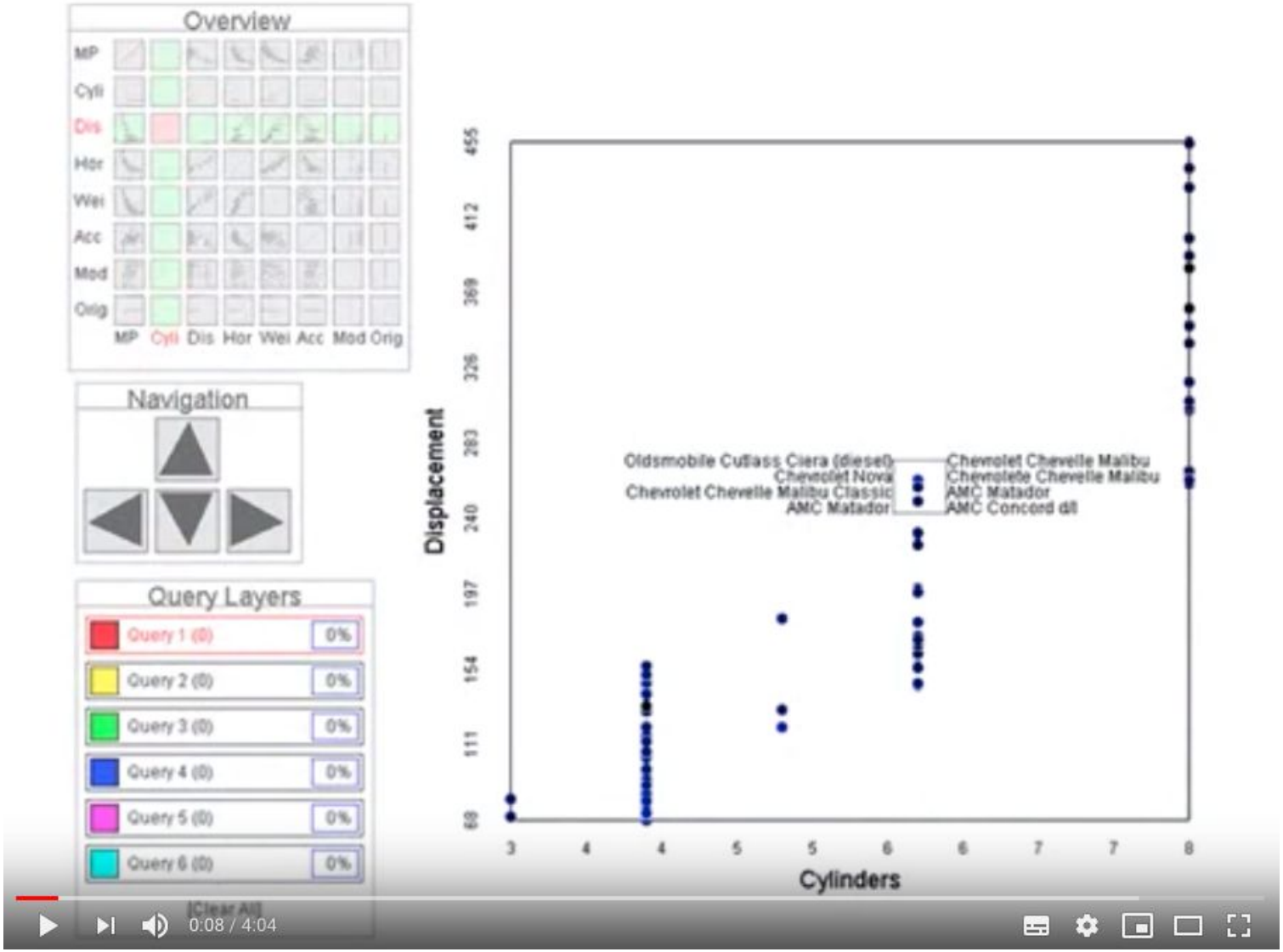
2. Pairwise comparison: **Scatterplot Matrix**



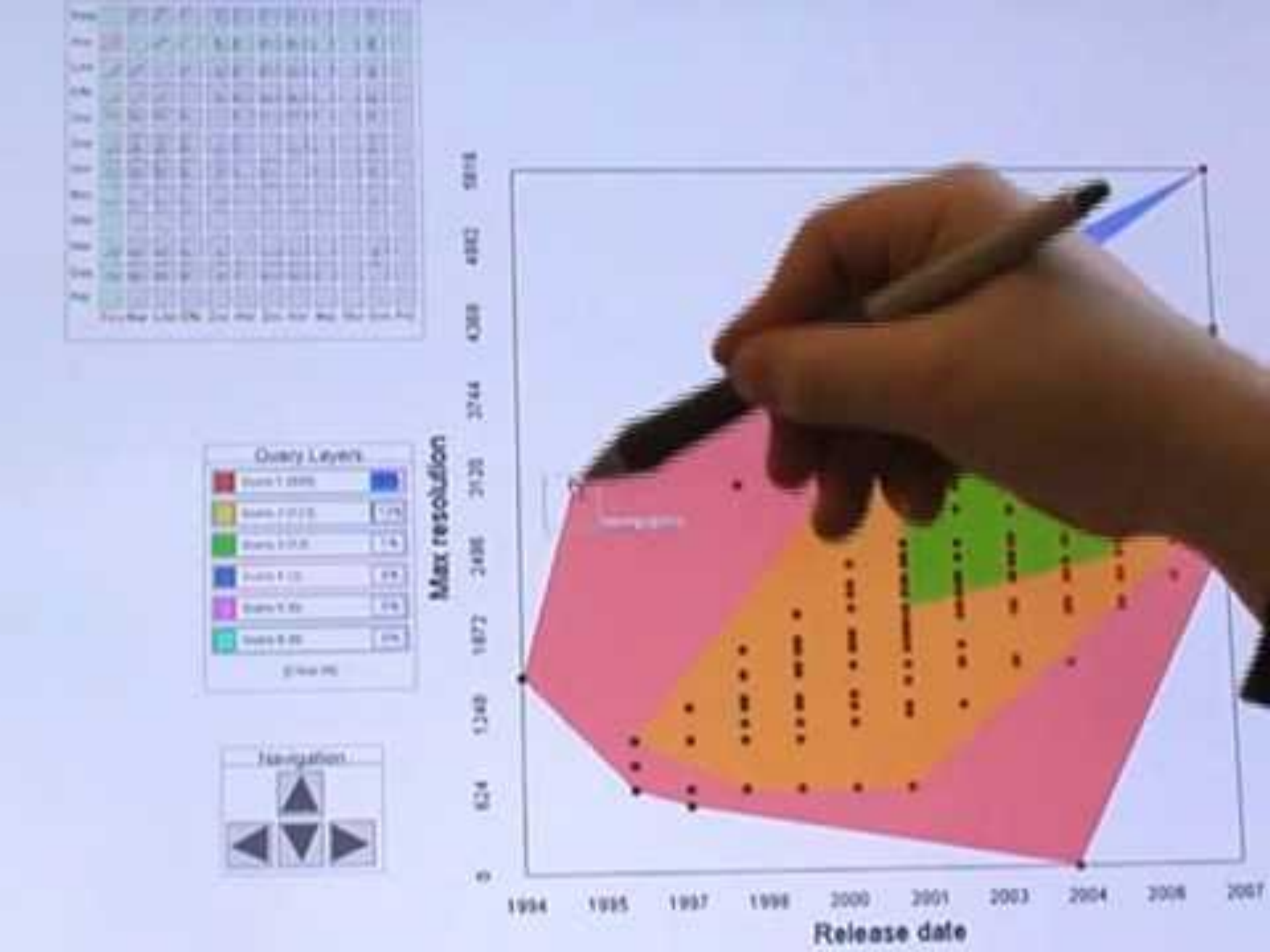
+ Scalable
+ Provides overview
+ Can use additional visualizations
+ Easy to understand and decode

- Many dimensions require pan+zoom

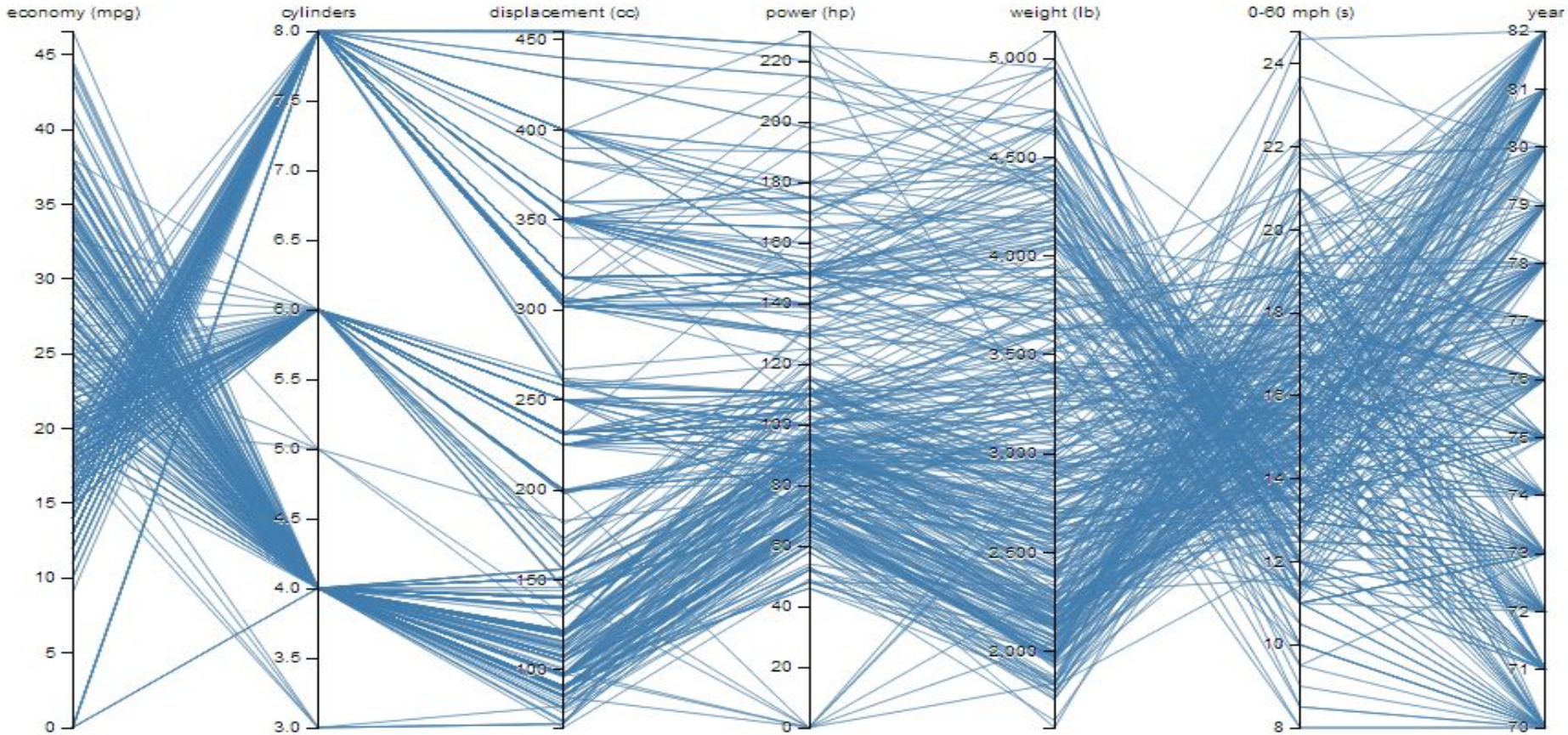
Scatterplot Navigation: **Scatterdice**



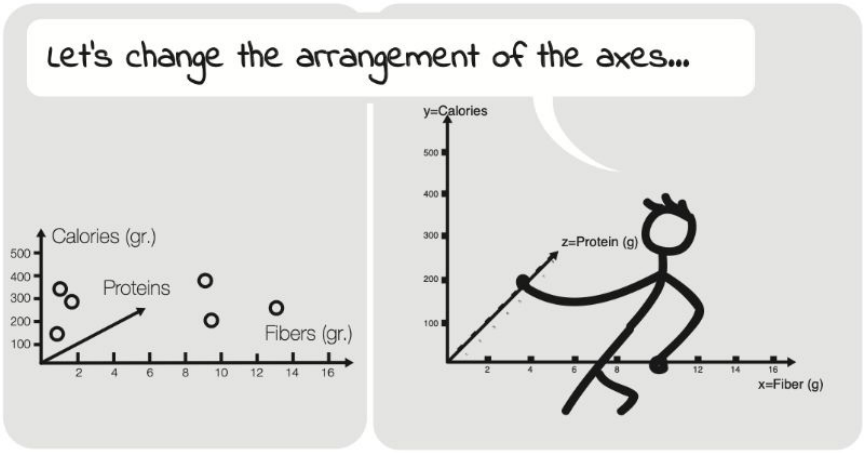
Elmqvist, Niklas, Pierre Dragicevic, and Jean-Daniel Fekete. "Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation." *IEEE transactions on Visualization and Computer Graphics* 14.6 (2008): 1539-1148.


Query Layers
Navigation
Max resolution
0
624
1248
1872
2496
3120
3744
4368
4992
5616
1994
1995
1997
1998
2000
2001
2003
2004
2006
2007
Release date

4. Parallel Coordinates Plot (PCP)

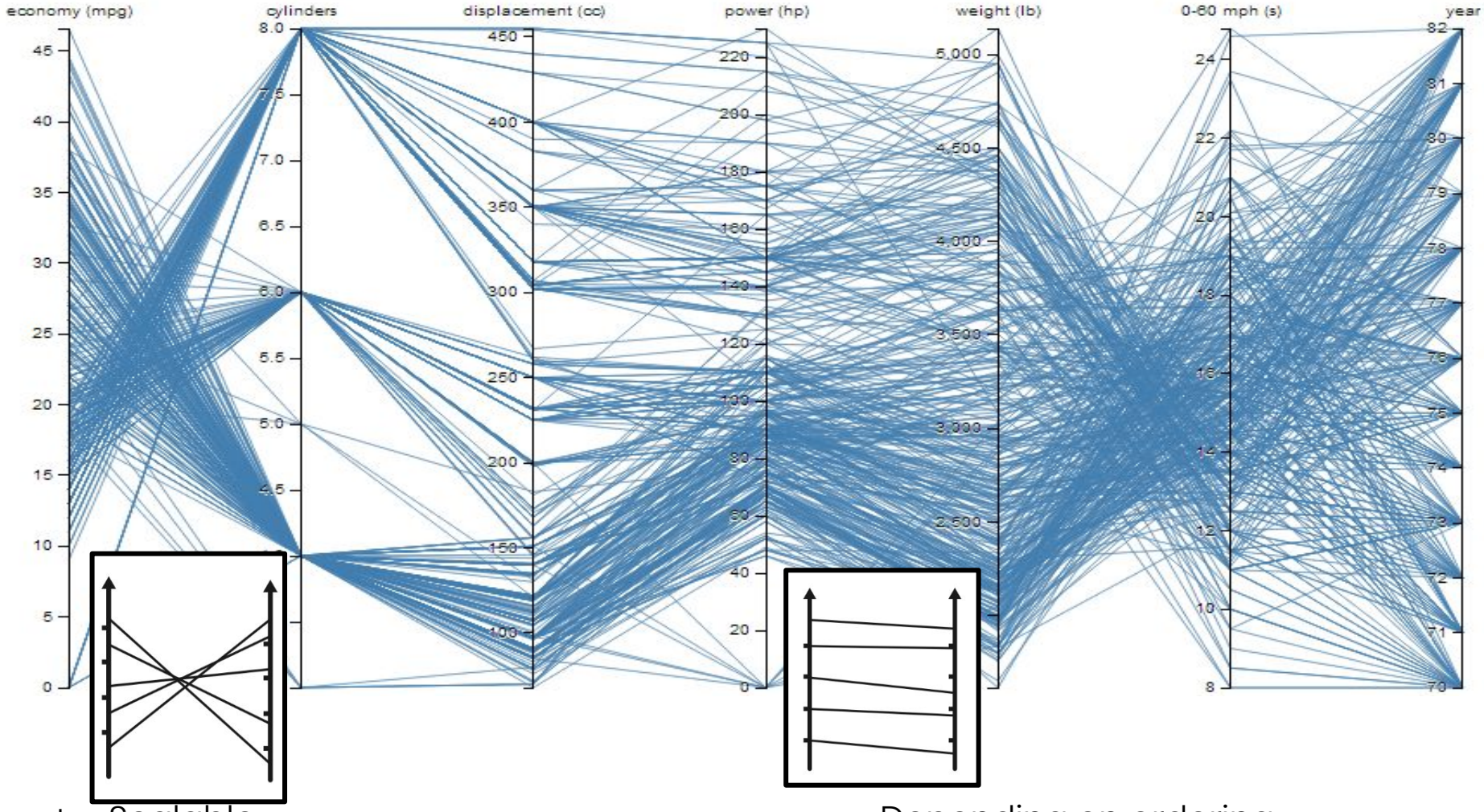


4. Parallel Coordinates Plot (PCP)



https://visualizationcheatsheets.github.io/pcp.html

4. Parallel Coordinates Plot (PCP)

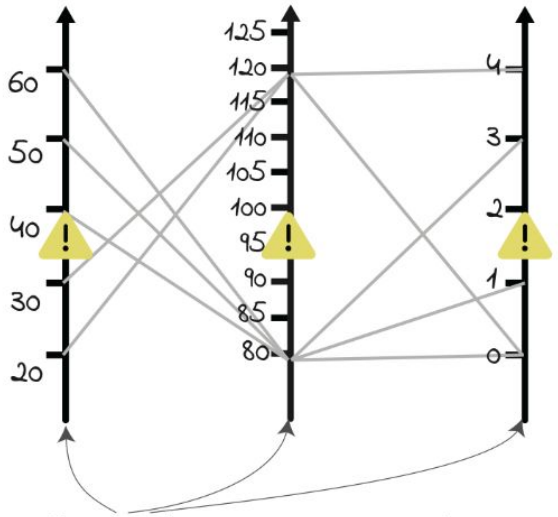


+ Scalable
+ Consise
+ Good overview

- Depending on ordering
- Can suffer from clutter
- Visual path following can be hard

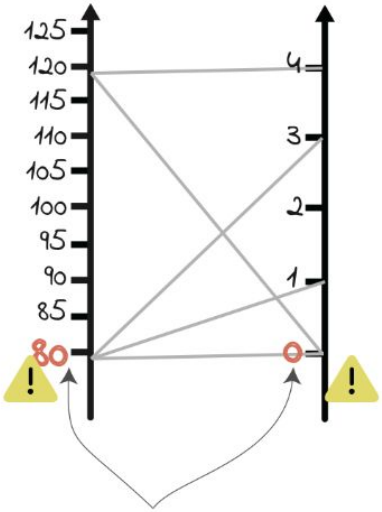
PCPs Pitfalls

Axis scales



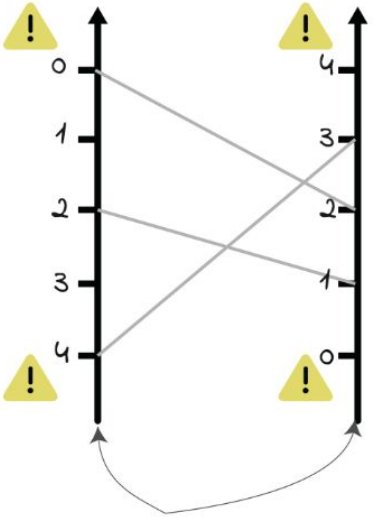
Different dimensions usually have different scales and units.

Truncated axes



Values on axes can start form values other than '0'.

Axes order



Values on axes can be either decending or acending.

https://visualizationcheatsheets.github.io/pcp.html

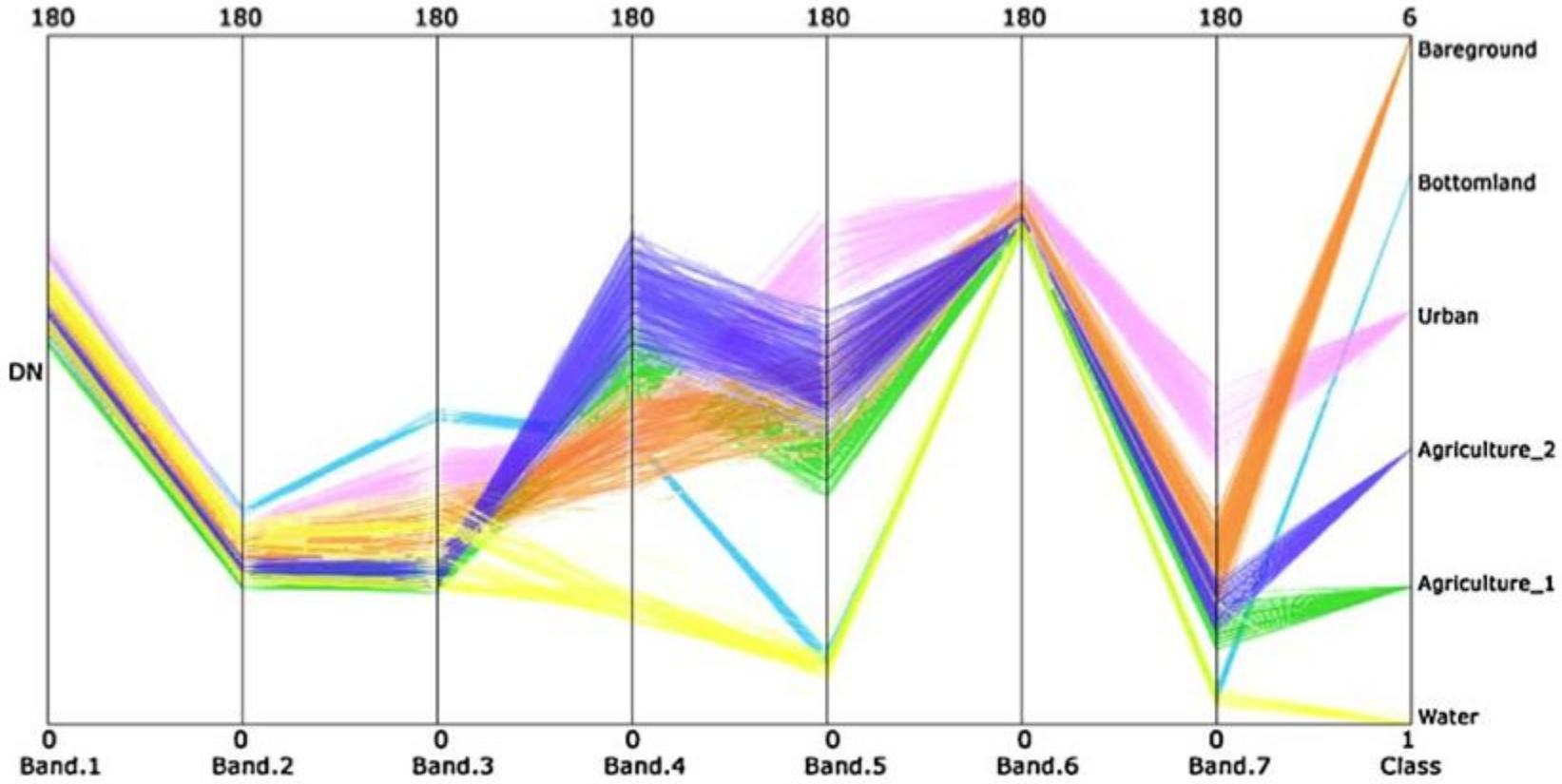
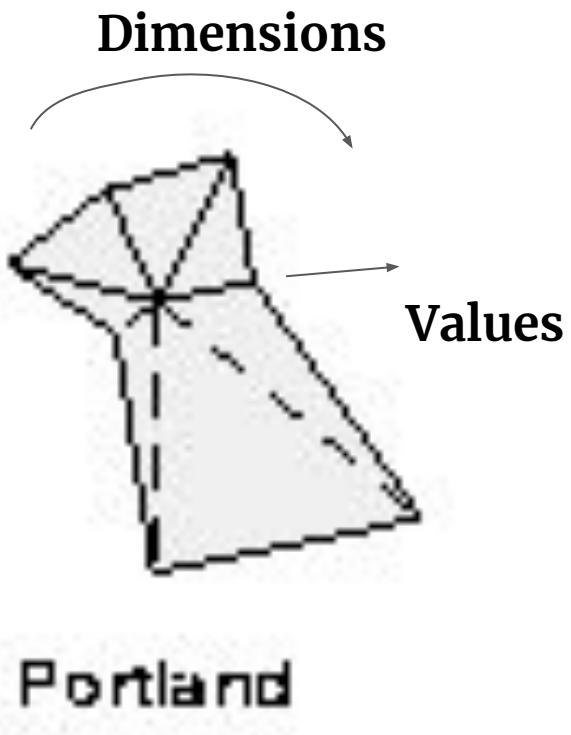


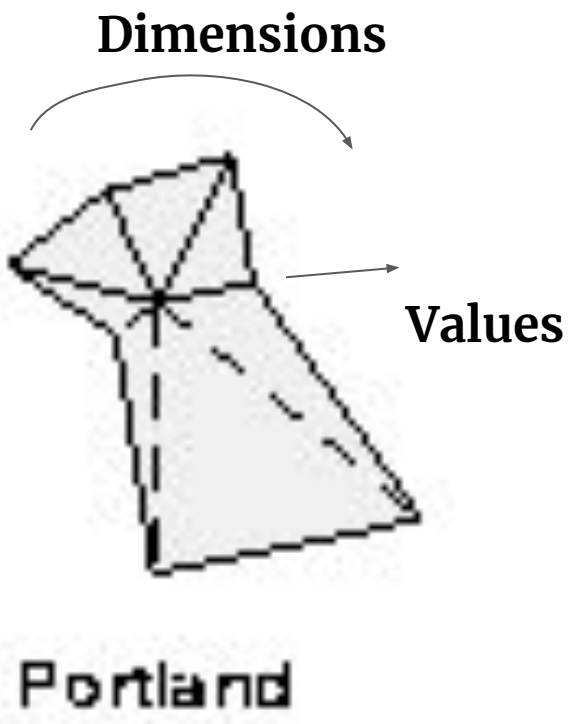
Fig. 3. PCP of sample data.

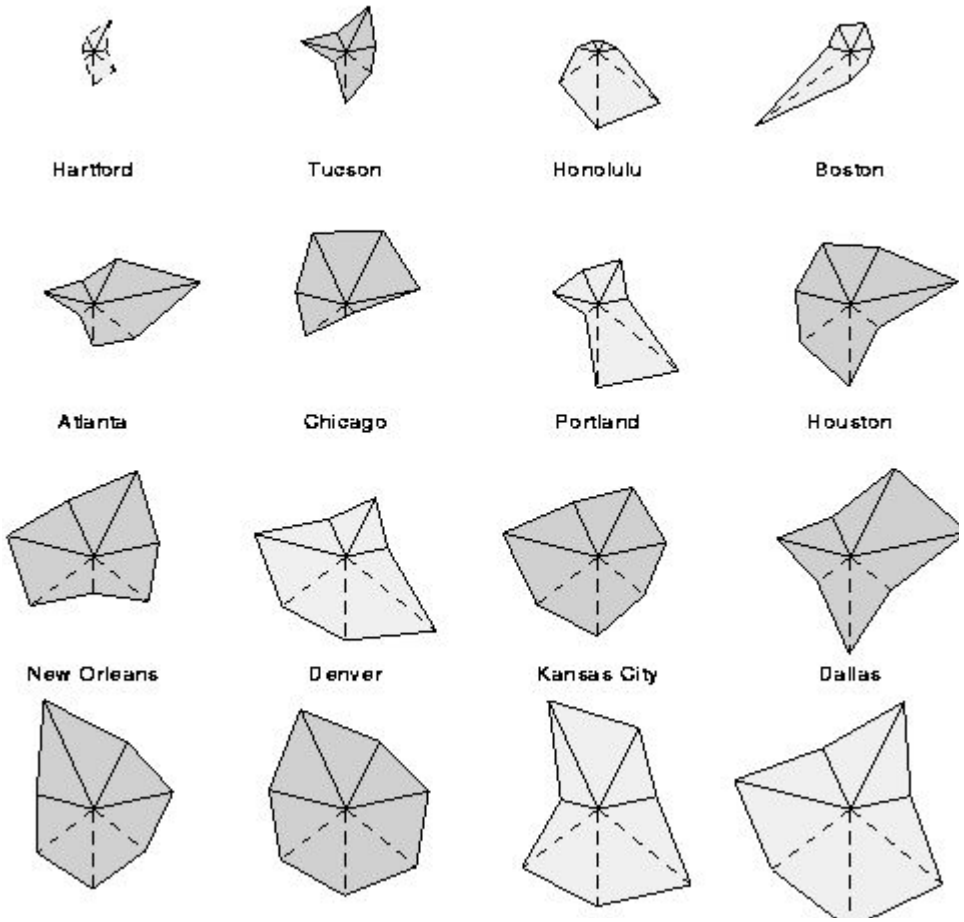
3. Glyphs: Star glyphs



1) Data glyphs are [a] technique, in which single data points are encoded individually by assigning their dimensions to one or more marks and their visual variables
2) Each glyph can be placed spatially independently from others

3. Glyphs: Star glyphs

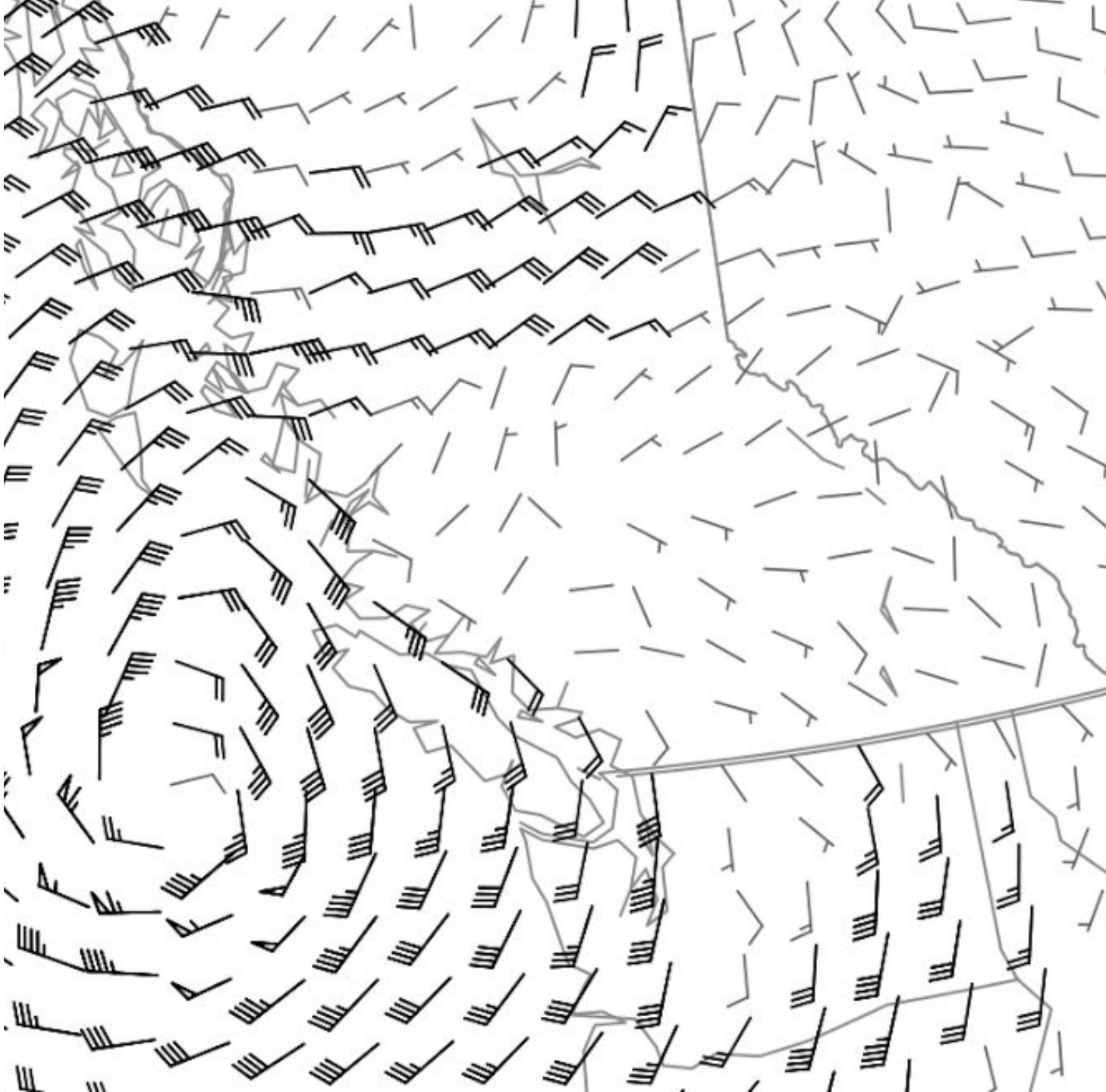




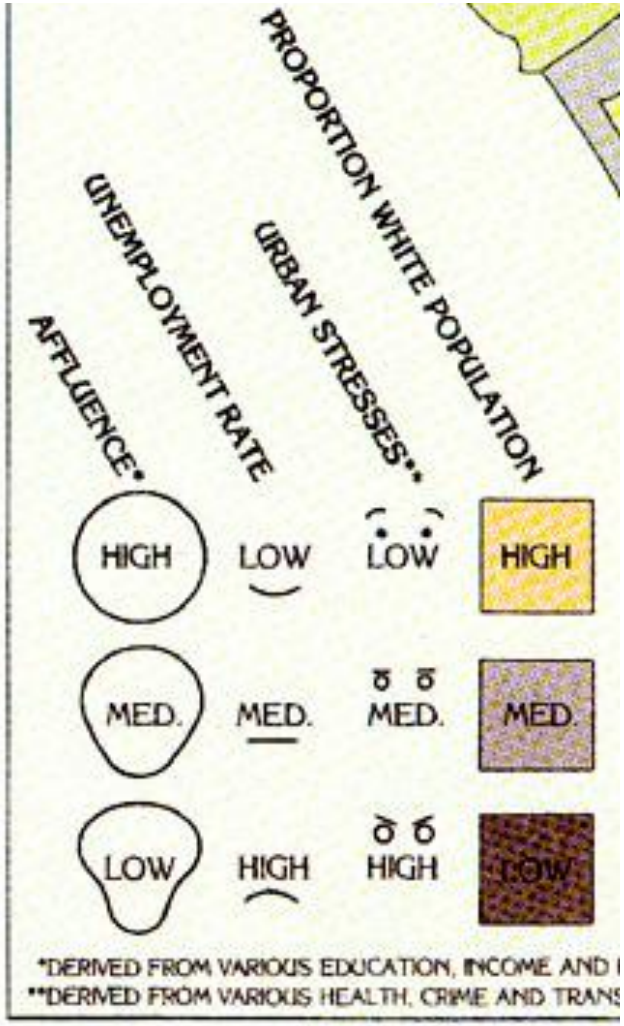
+ Comparison
+ Outlier, Trends
+ Individual values

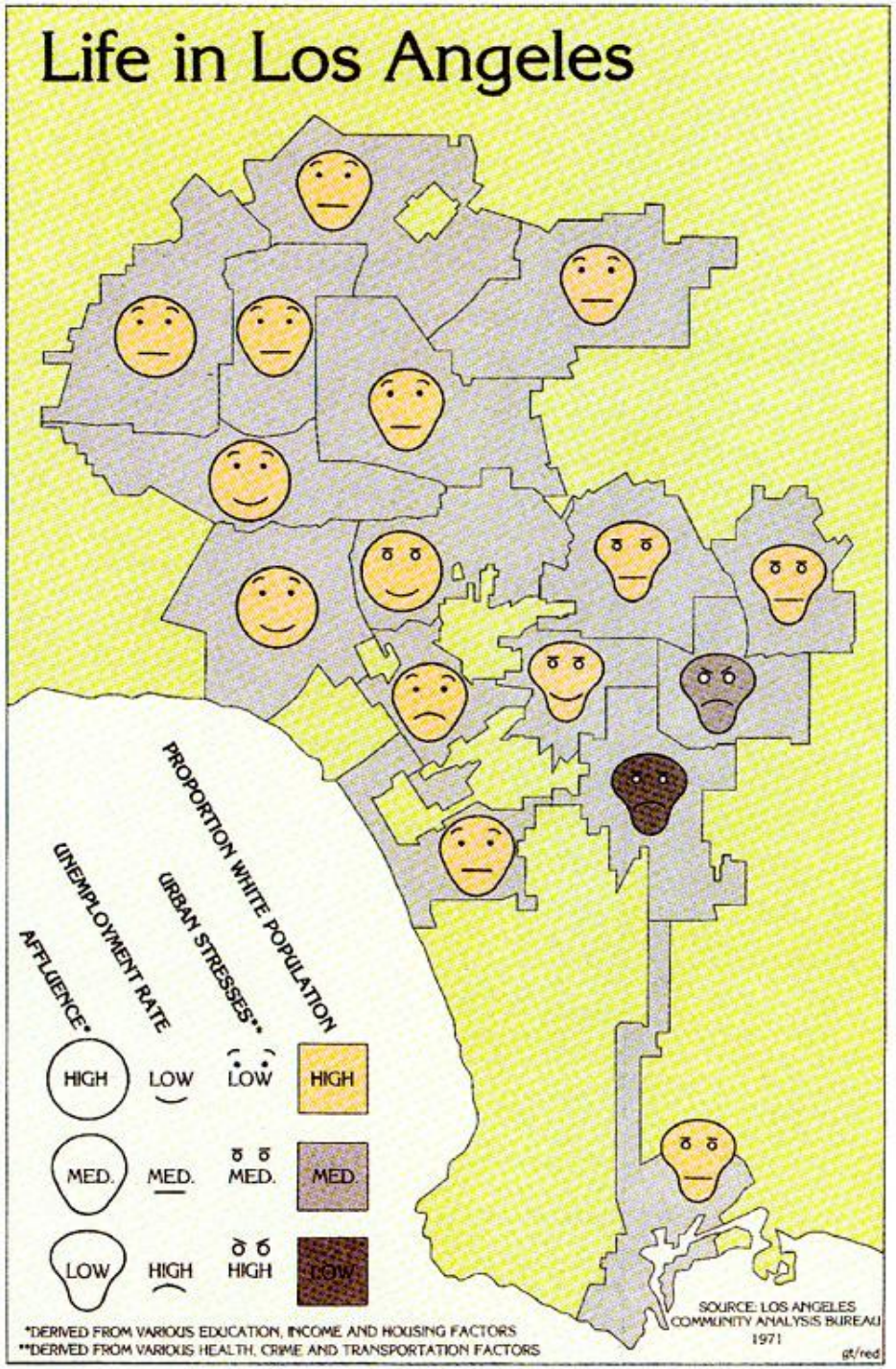
- Precise comparison
- Cluster dimensions

Weather glyphs

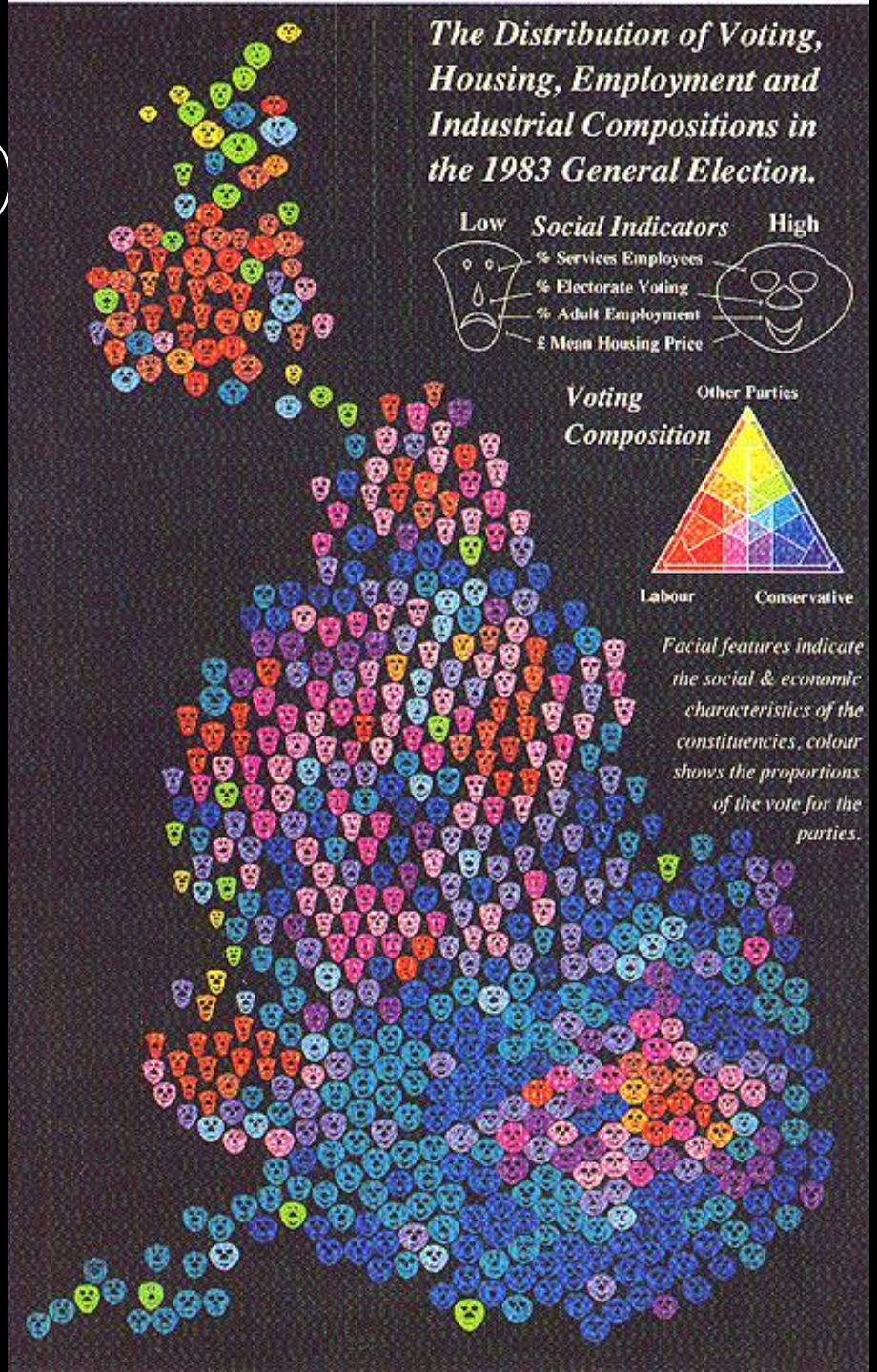


Chernoff Faces
(Glyphs for geodata)





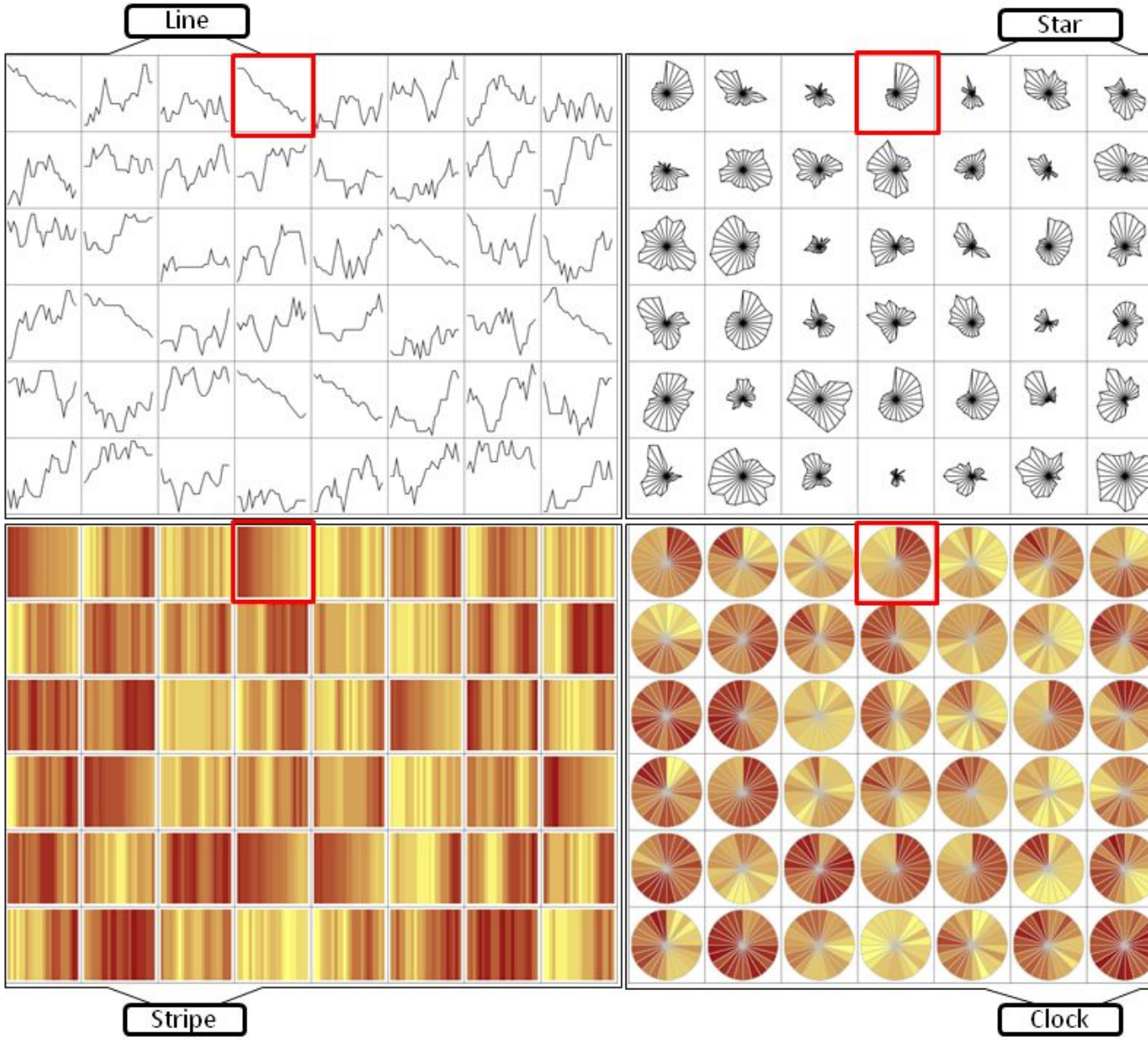
Chernoff Faces
(Glyphs for geodata)



+ Individual values
+ Spatial correlation

- Some vis-variables are more prominent
- Some vis-variables are hard to perceive and estimate

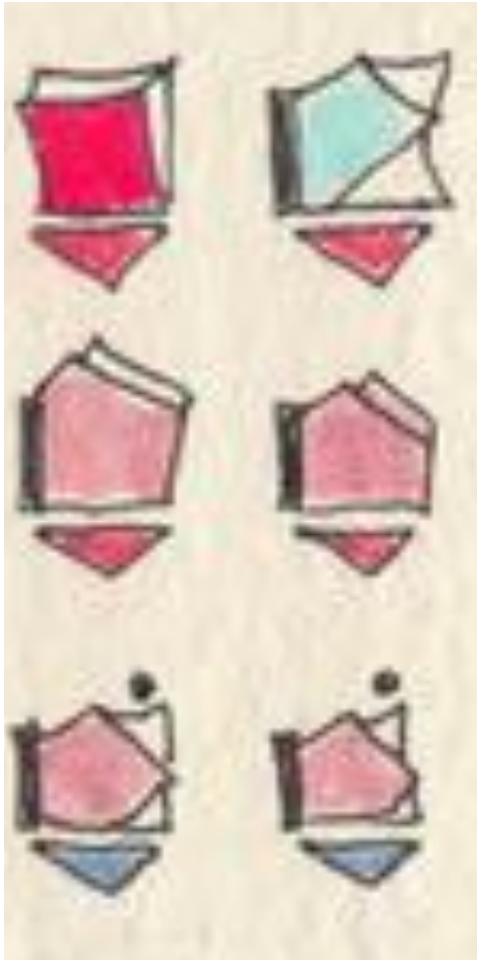
Glyphs for Temporal data

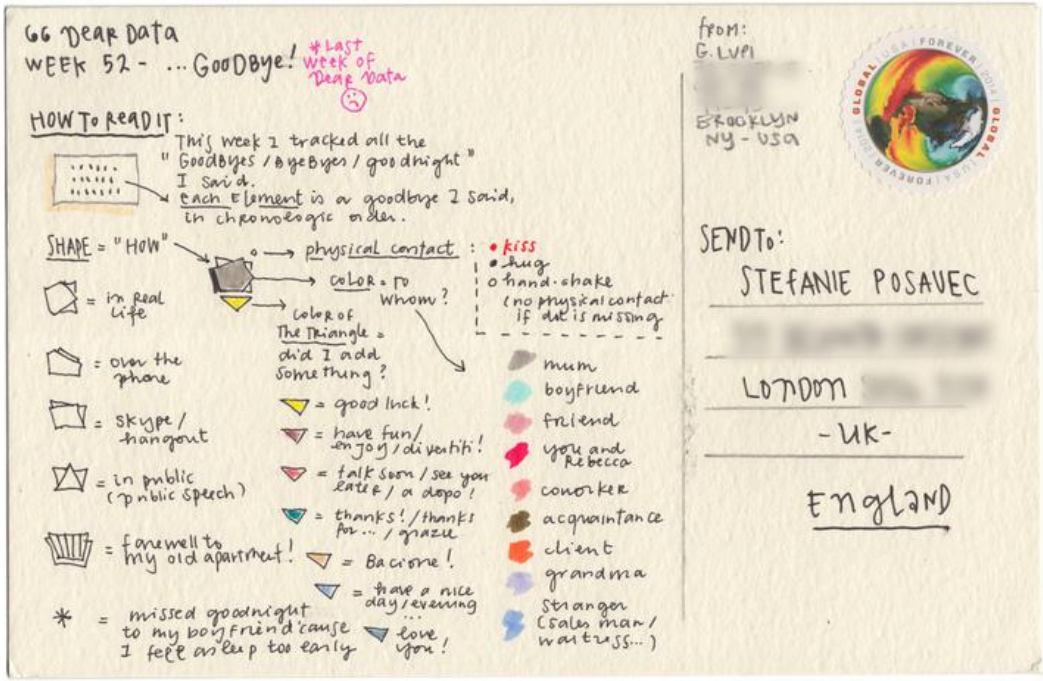


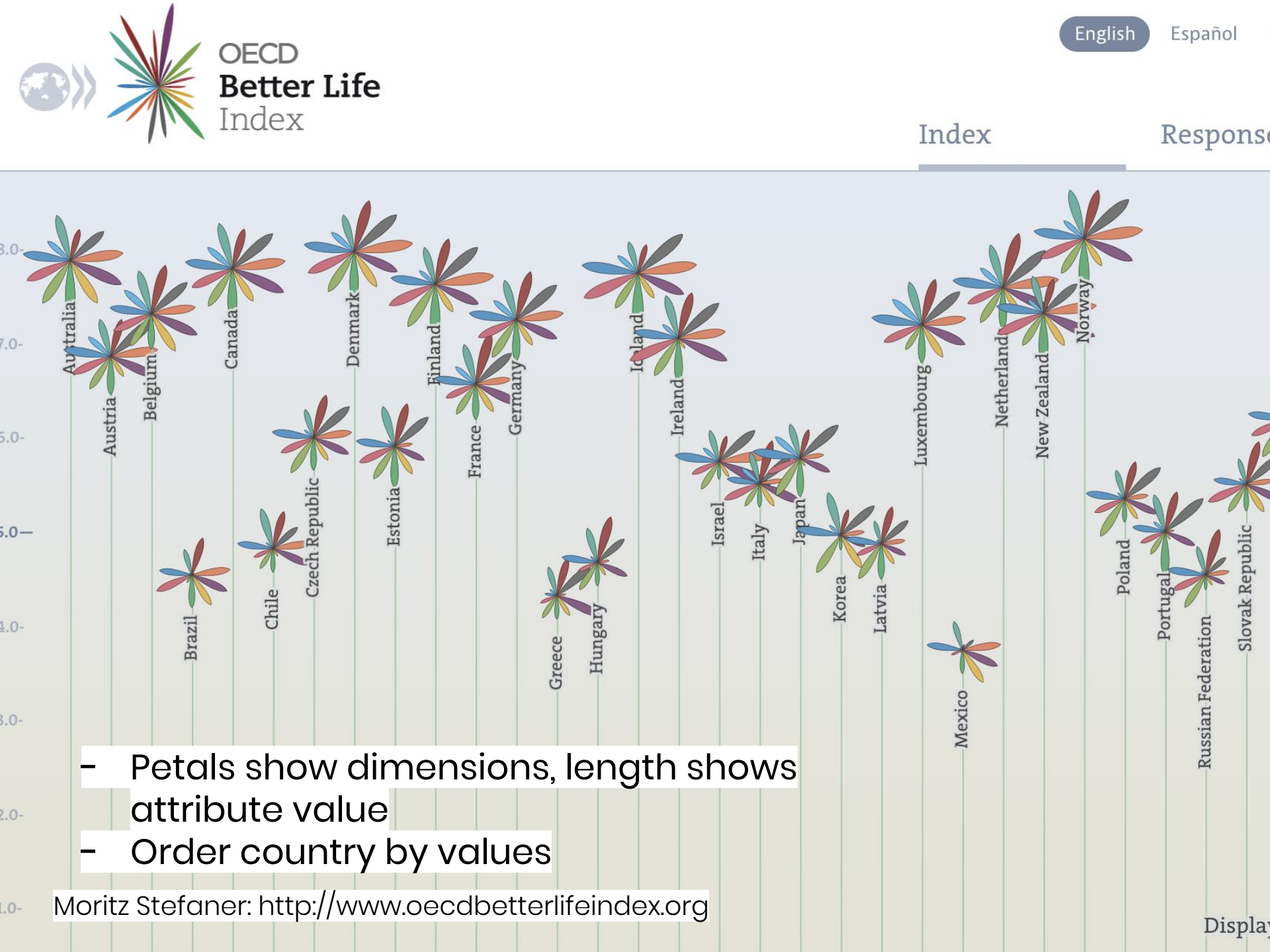
- Lineglyhs for trends and peaks
- Radial for value look-up

Fuchs, Johannes, et al. "Evaluation of alternative glyph designs for time series data in a small multiple setting." *Proceedings of the SIGCHI conference on human factors in computing systems*. 2013.

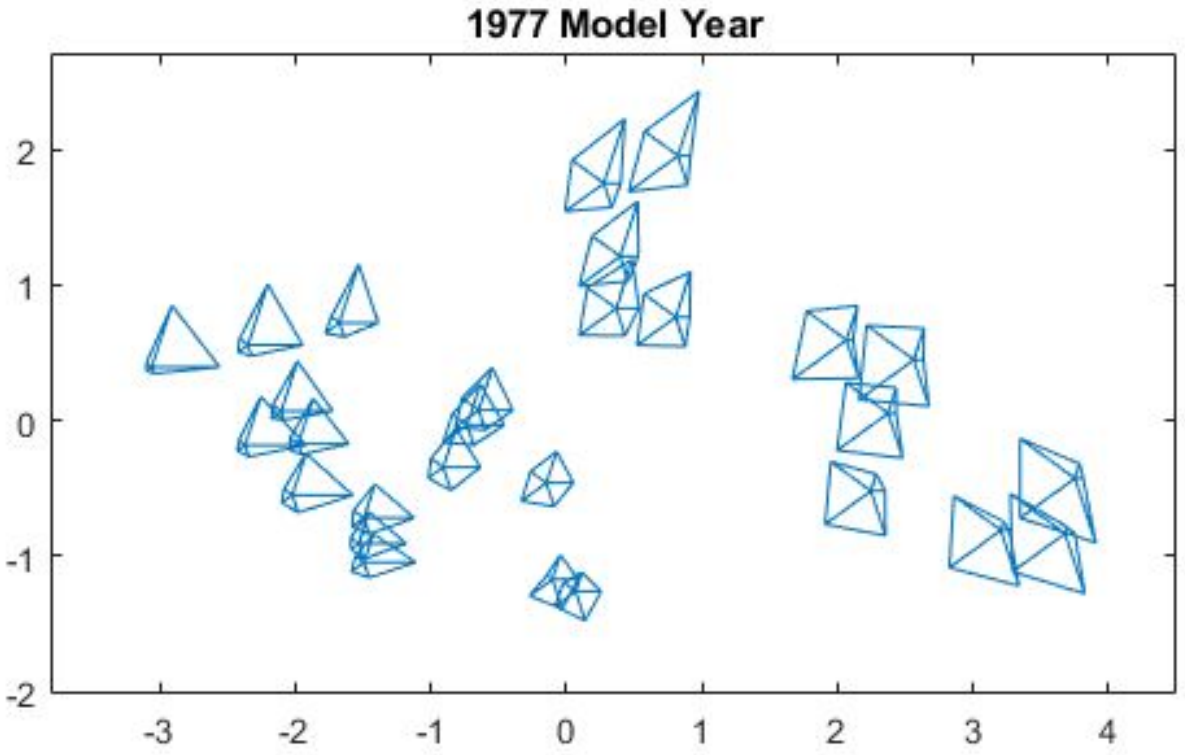
Dear Data
(fantastic glyphs)



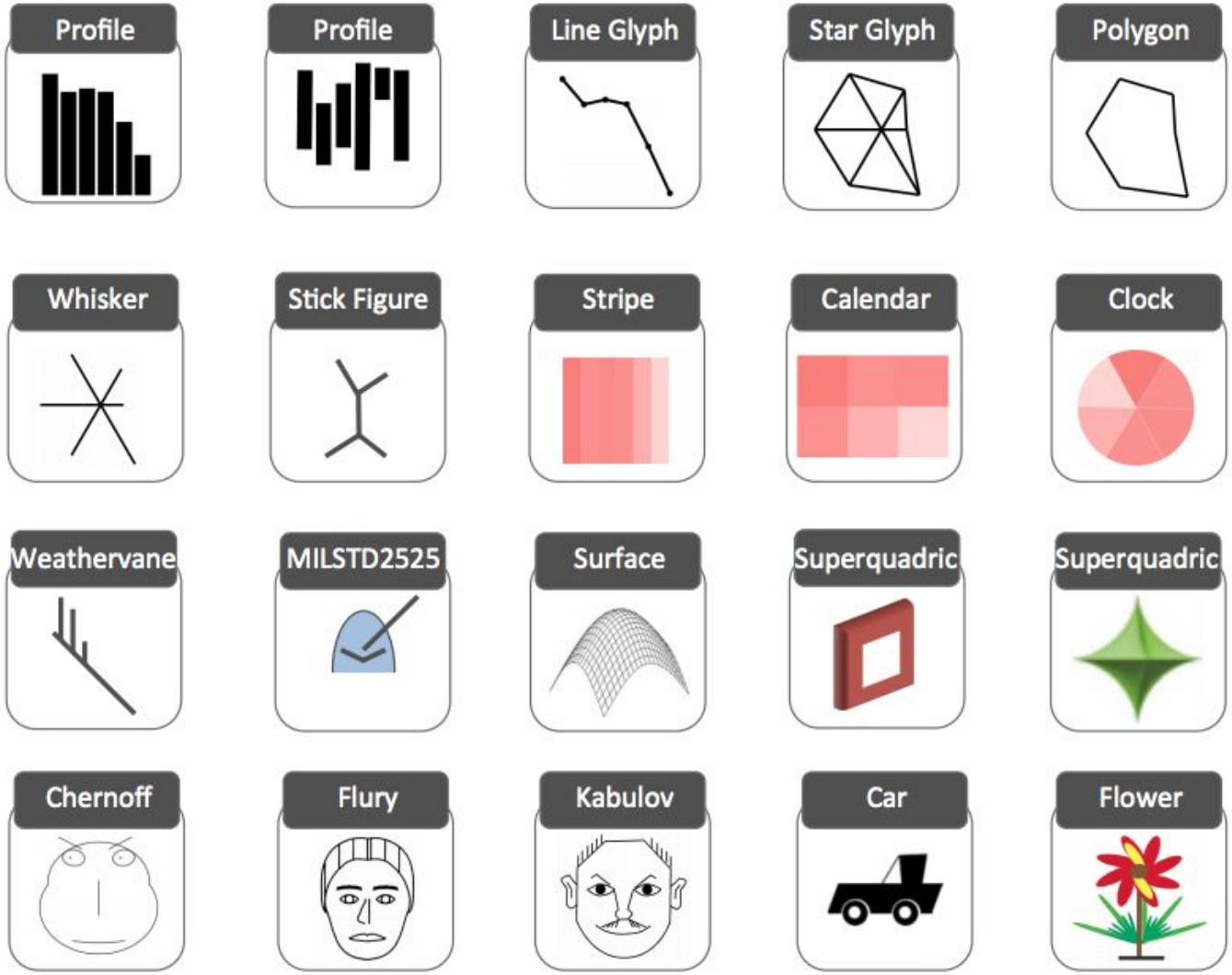



OECD
Better Life
Index
English
Español
Index
Respons
Australia
Austria
Belgium
Canada
Denmark
Finland
France
Germany
Iceland
Ireland
Israel
Italy
Japan
Luxembourg
Netherland
New Zealand
Norway
Brazil
Chile
Czech Republic
Estonia
Greece
Hungary
Korea
Latvia
Mexico
Poland
Portugal
Russian Federation
Slovak Republic
- Petals show dimensions, length shows attribute value
- Order country by values
Moritz Stefaner: http://www.oecdbetterlifeindex.org

Glyphs in Scatterplots

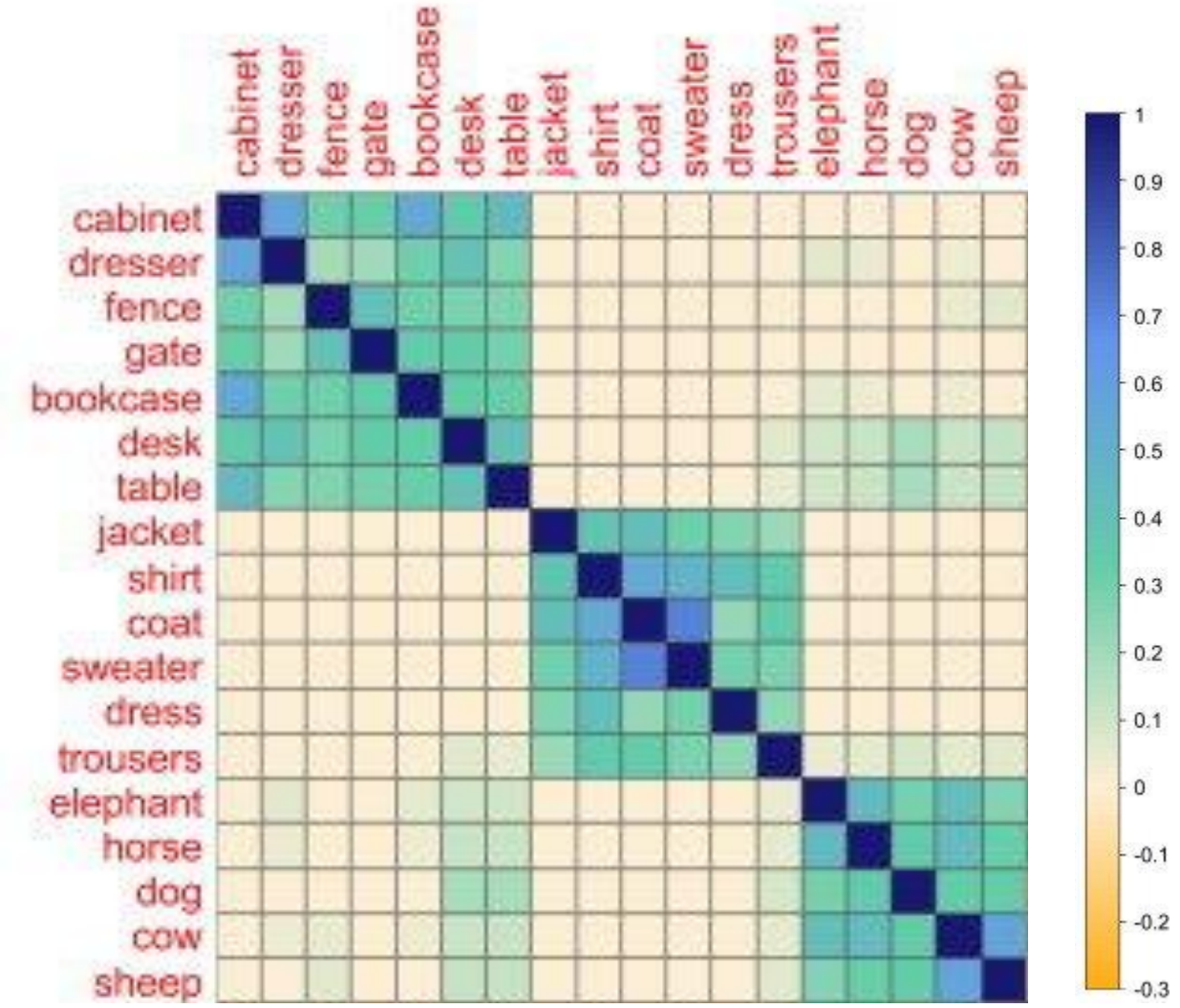


Glyph Design



Fuchs, Johannes, et al. "A systematic review of experimental studies on data glyphs." *IEEE transactions on visualization and computer graphics* 23.7 (2016): 1863-1879.

5. Dimensionality Reduction: Similarity

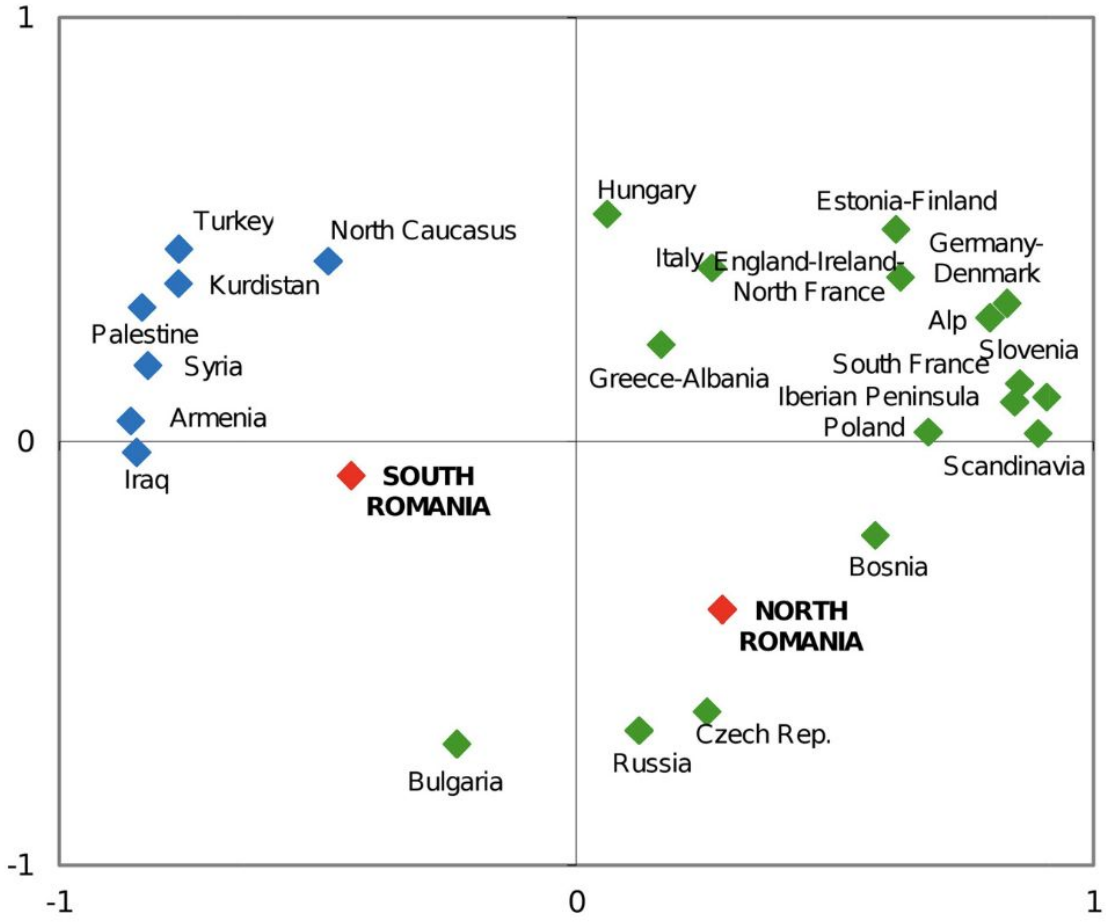


5. Dimensionality Reduction: MDS

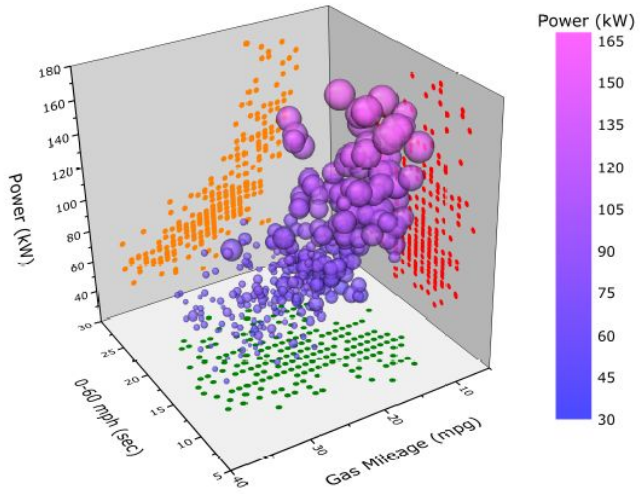
5. Dimensionality Reduction: MDS

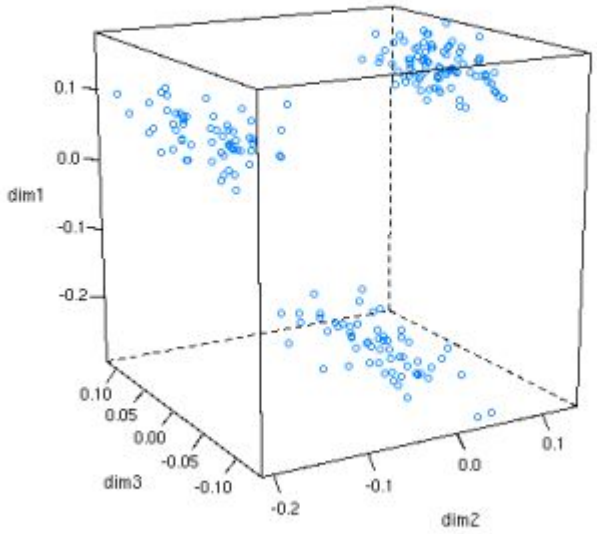
+ Dimension reduction
+ Can be 2D or 3D
+ Visual clustering

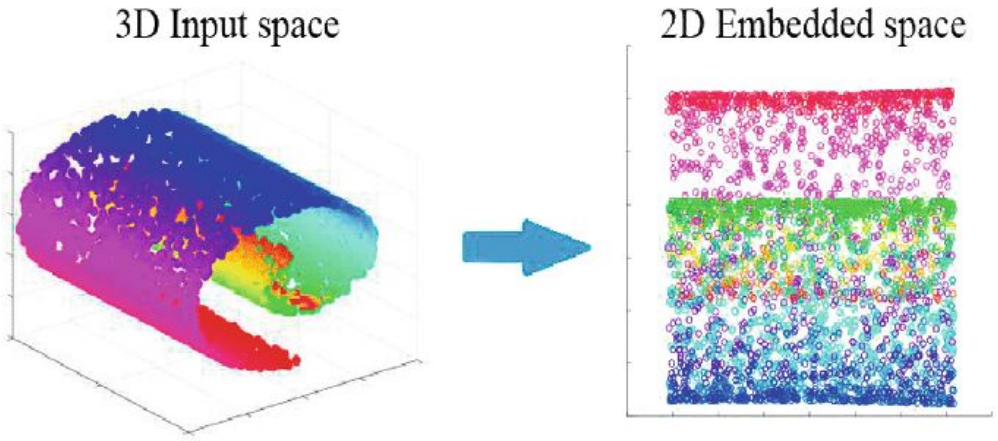
- Information loss
- Artifacts: **false neighbors** and **tears**



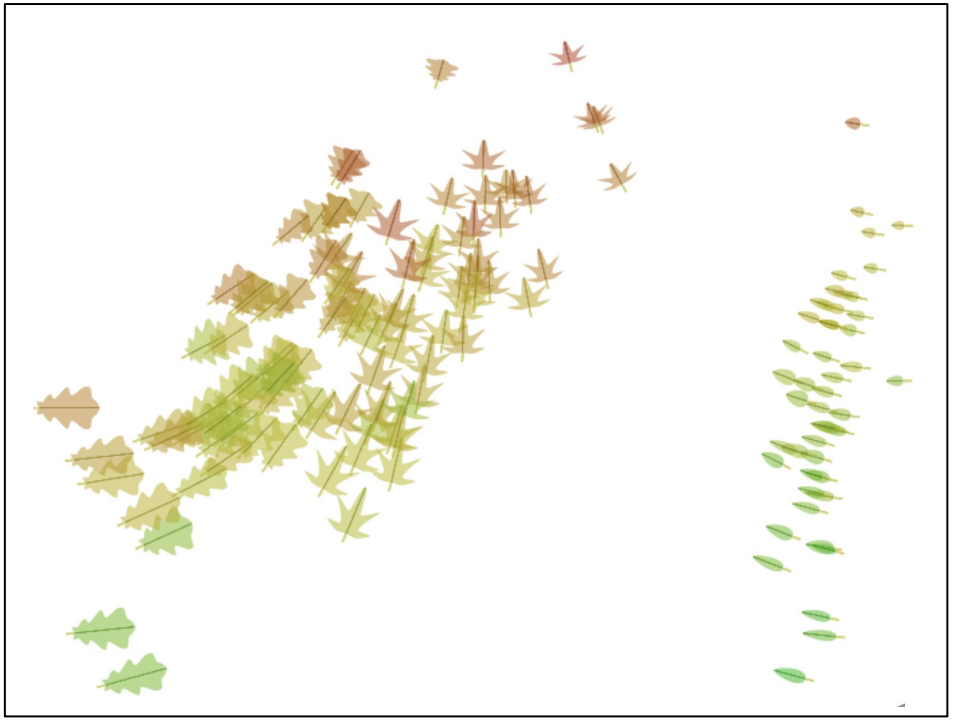
5. Dimensionality Reduction: 3D

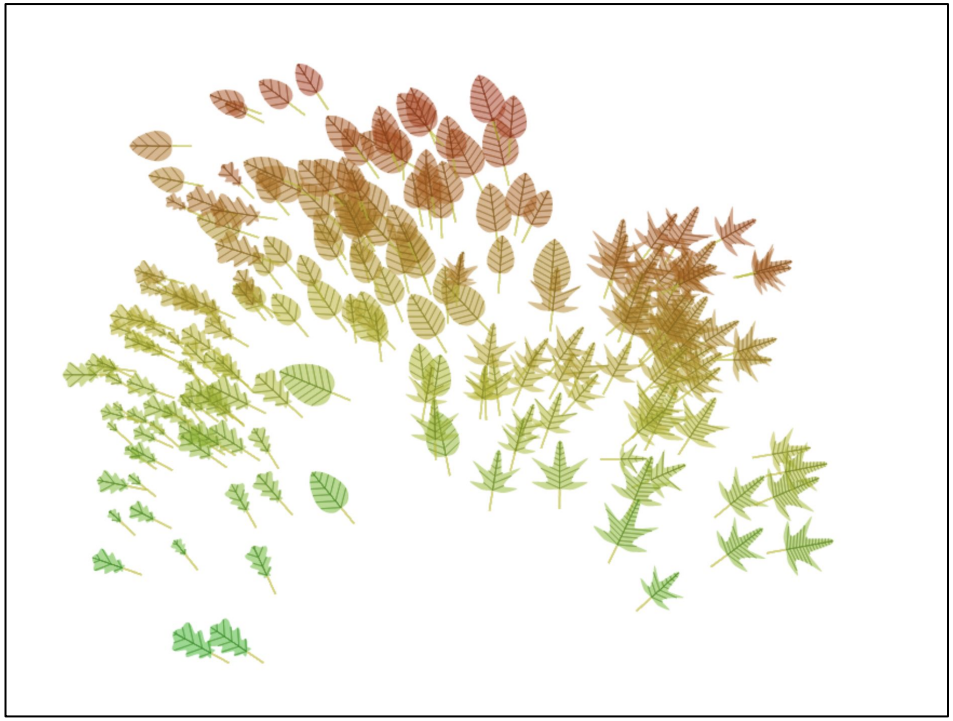




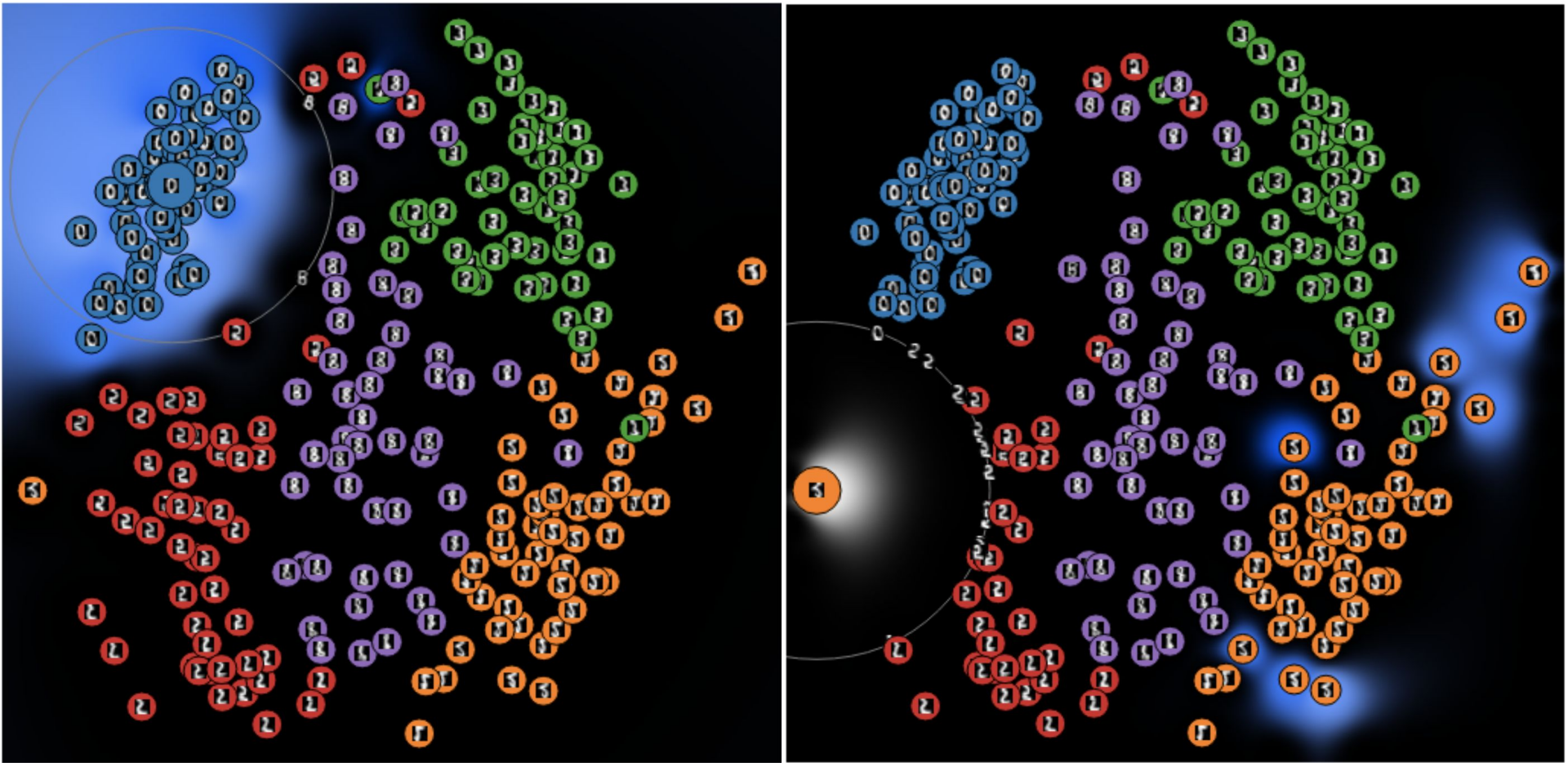


5. Dimensionality Reduction: Glyphs





5. Dimensionality reduction: Interaction



Heulot, Nicolas, Michael Aupetit, and Jean-Daniel Fekete. "Proxilens: Interactive exploration of high-dimensional data using projections." *VAMP: EuroVis Workshop on Visual Analytics using Multidimensional Projections*. The Eurographics Association, 2013.

Further Readings

- Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.
- Fuchs, Johannes, et al. "A systematic review of experimental studies on data glyphs." *IEEE transactions on visualization and computer graphics* 23.7 (2016): 1863-1879.
- Heinrich, Julian, and Daniel Weiskopf. "State of the Art of Parallel Coordinates." *Eurographics (STARs)*. 2013.
- https://visualizationcheatsheets.github.io/pcp.html