

Session 4

Multidimensional Data



Online Course
**Data Visualization
for Professionals**



THE UNIVERSITY
of EDINBURGH

Benjamin Bach

June 2022

<http://benjbach.me>

<https://datavis-online.github.io>

-- Not for external use --

Outline

How to visualize data with dimensions (many attributes)?

- Low-dimensions (< 3)
- Higher dimensions (> 2)
- Dimensionality reduction (many!)

Outline

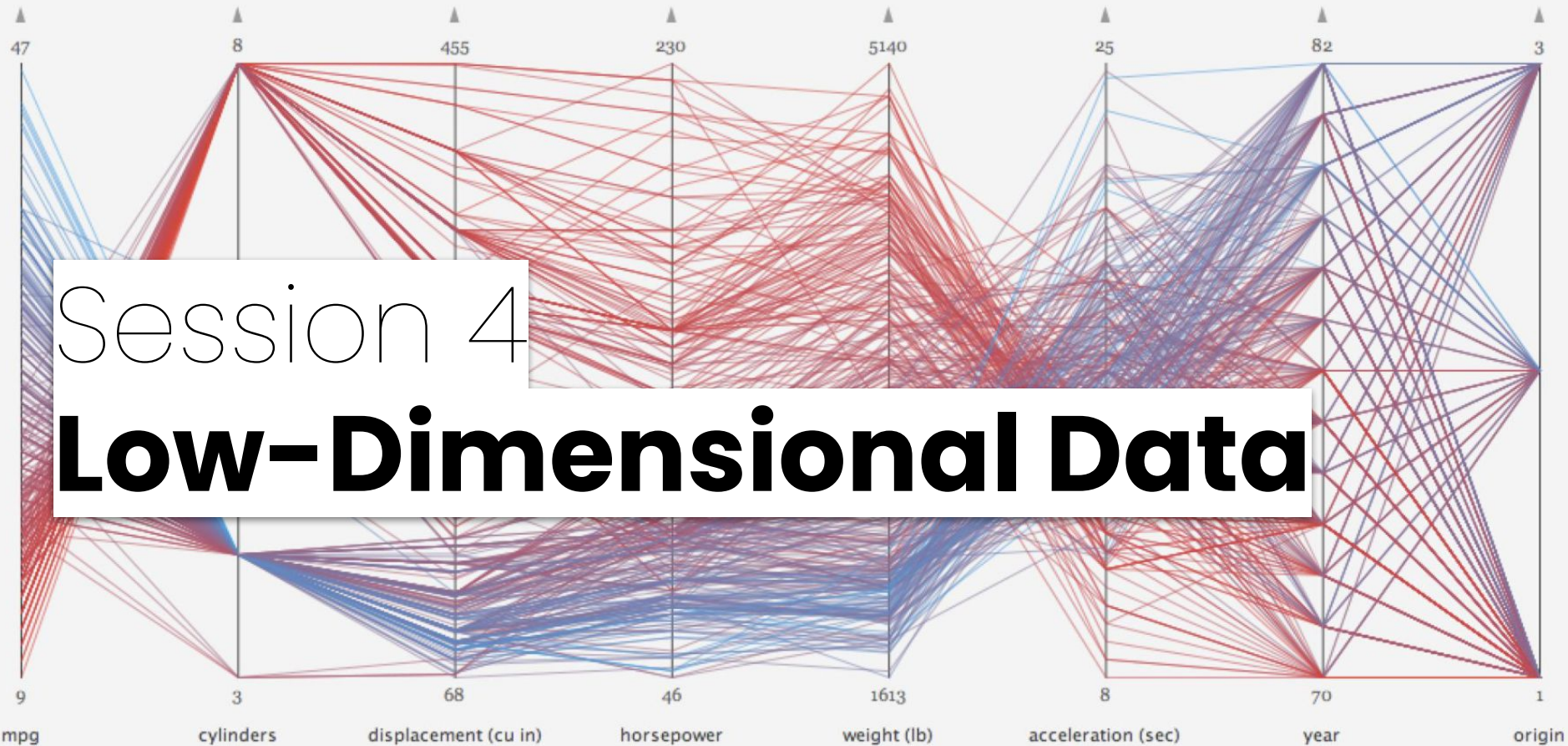
How to visualize data with dimensions (many attributes)?

- Low-dimensions (< 3)
- Higher dimensions (> 2)
- Dimensionality reduction (many!)

Outline

How to visualize data with dimensions (many attributes)?

- Low-dimensions (< 3)
- Higher dimensions (> 2)
- Dimensionality reduction (many!)



Online Course
**Data Visualization
for Professionals**



THE UNIVERSITY
of EDINBURGH

Benjamin Bach

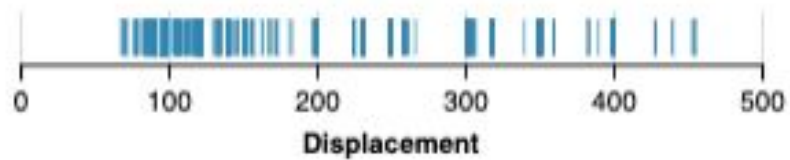
June 2022

<http://benjbach.me>

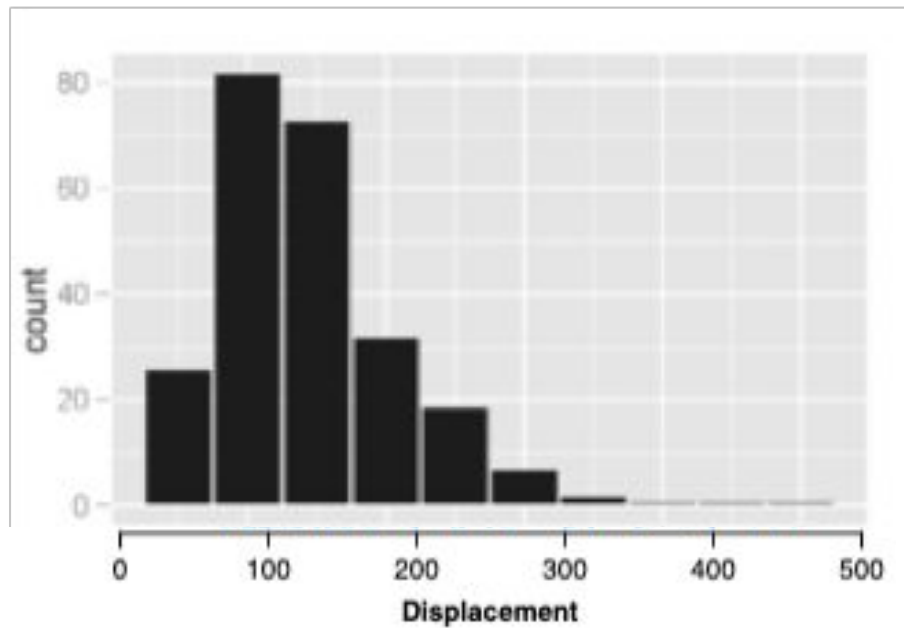
<https://datavis-online.github.io>

-- Not for external use --

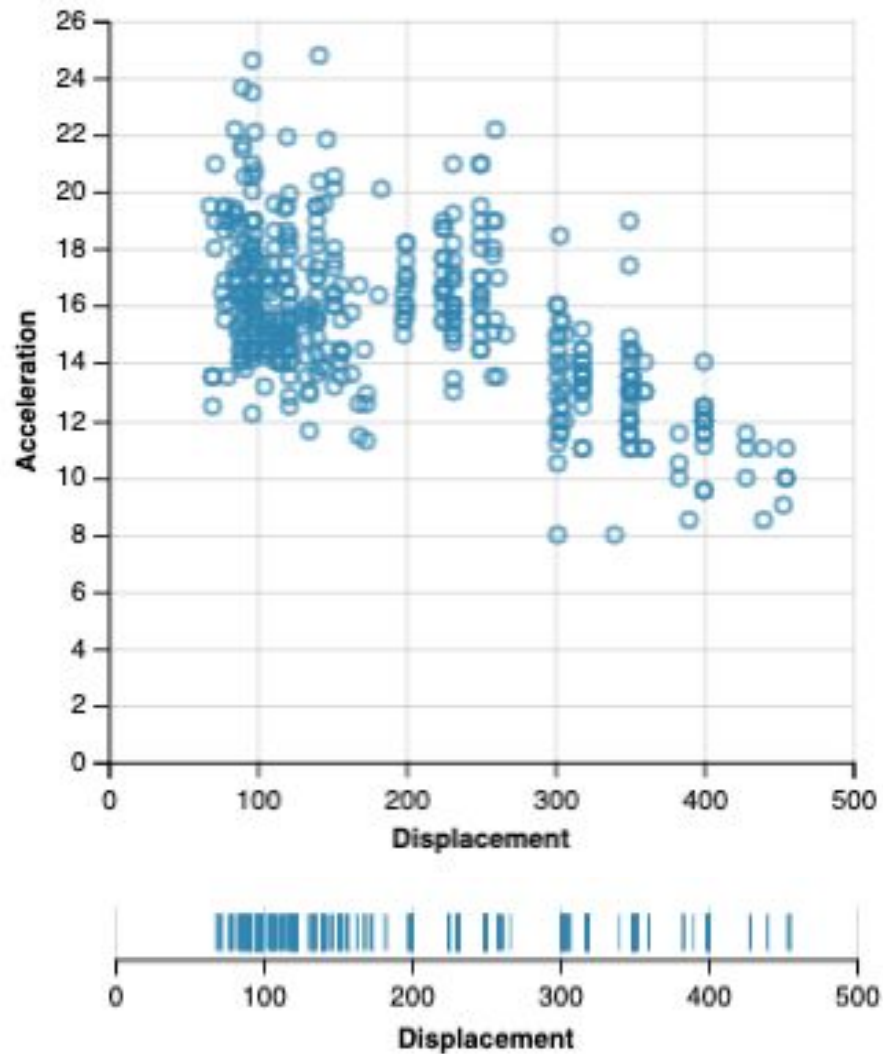
1 Dimension



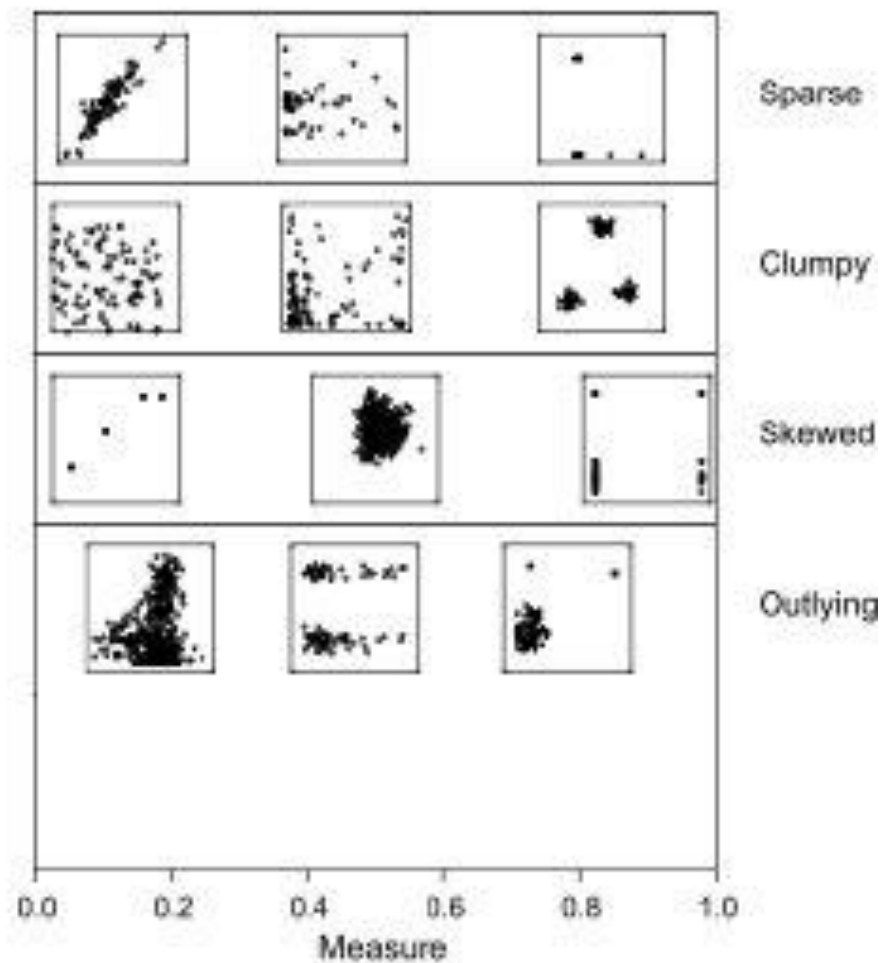
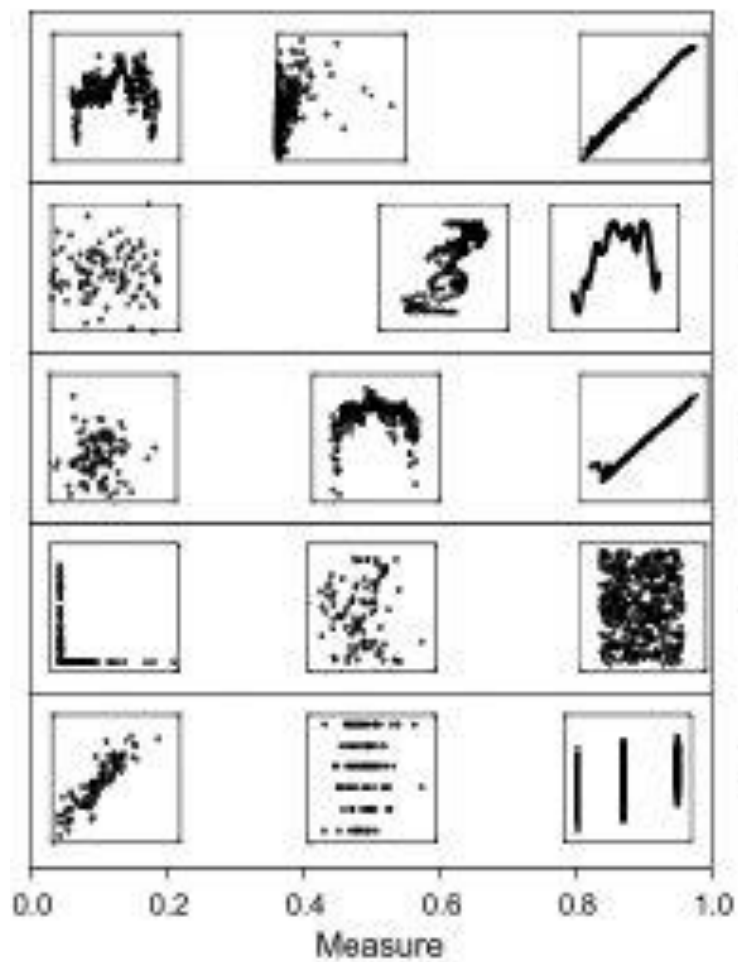
1 Dimension



2 Dimensions

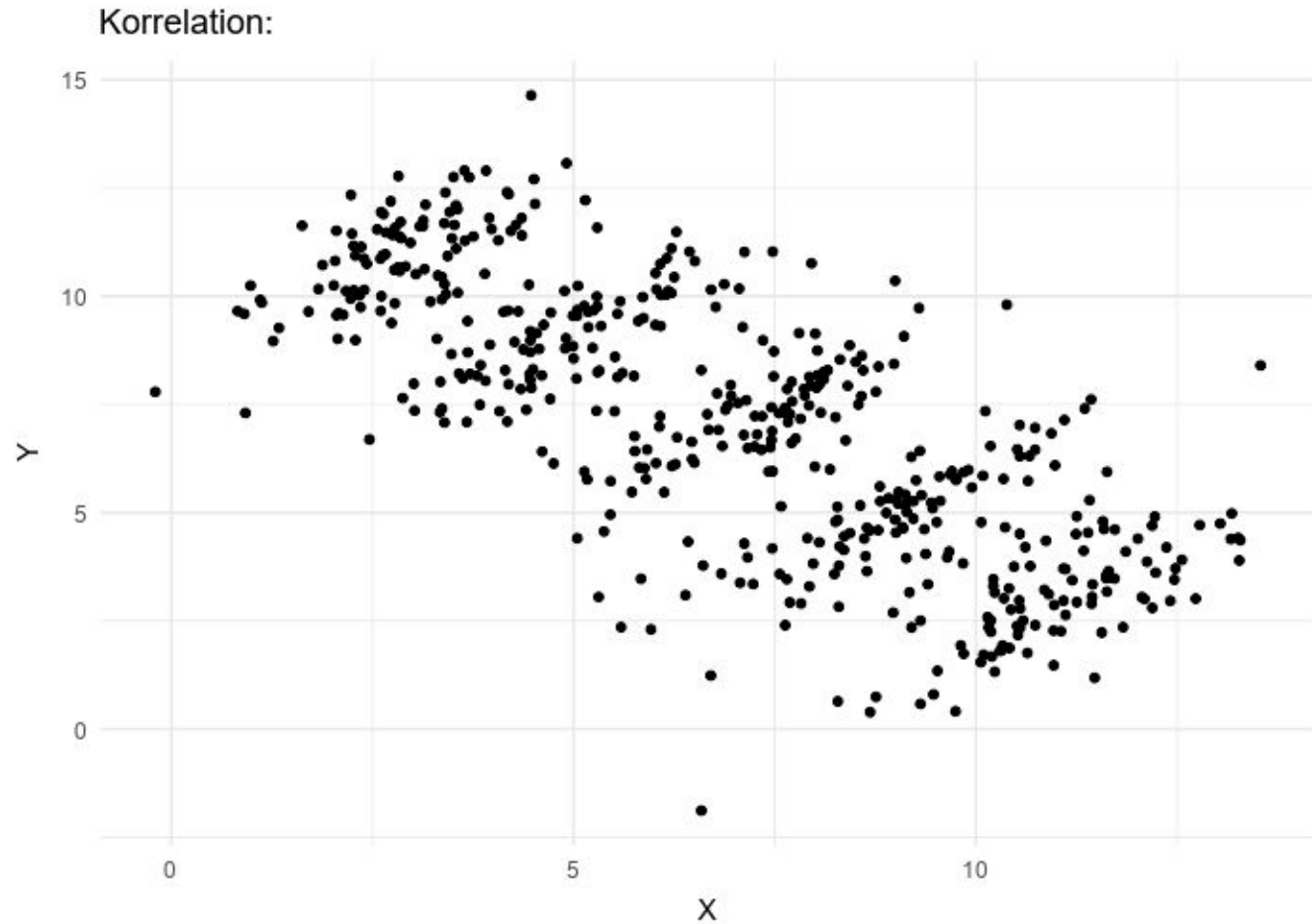


Quant + Quant: **Scatterplot**



Simpson's Paradox

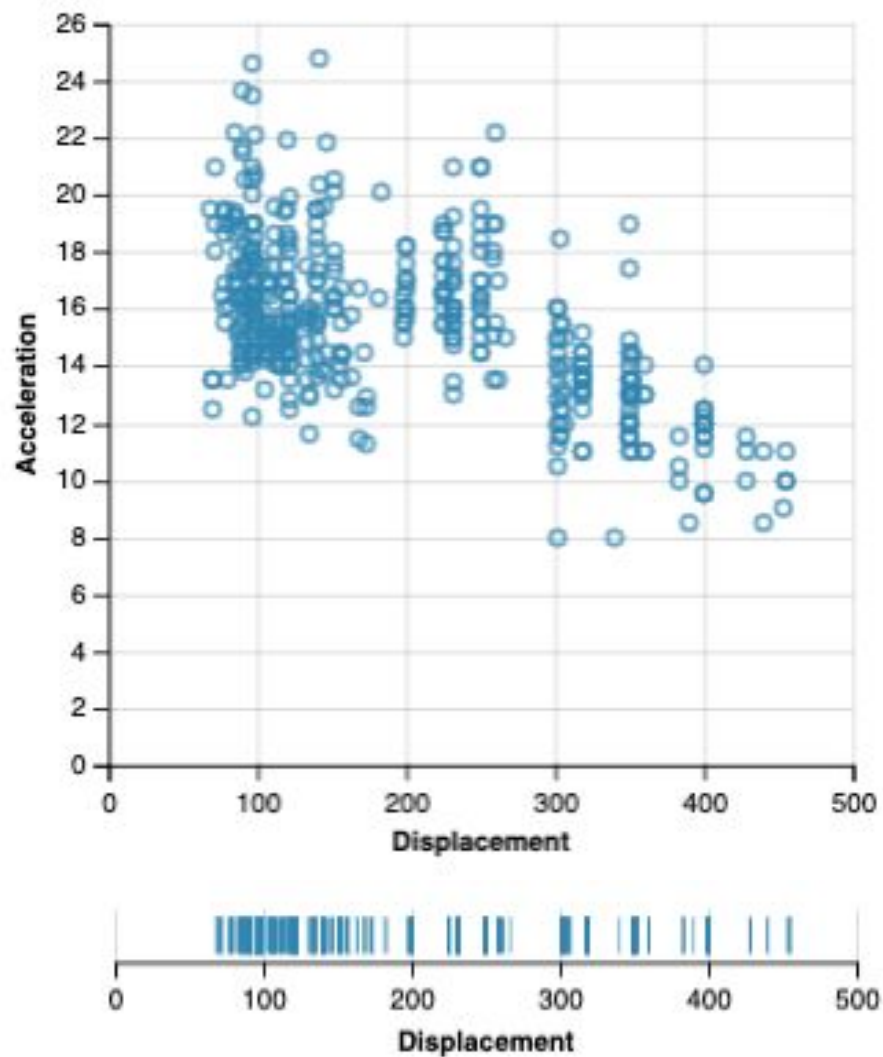
a trend appears in several different groups of data but disappears or reverses when these groups are combined.



Data types

	Quantitative	Ordered	Categorical
Quantitative			
Ordered			
Categorical			

Quant x Quant: **Scatterplot**



Quant x Quant: **Mekko Chart**

A world of difference

GDP per person at PPP* and share of global population, 2014 forecast

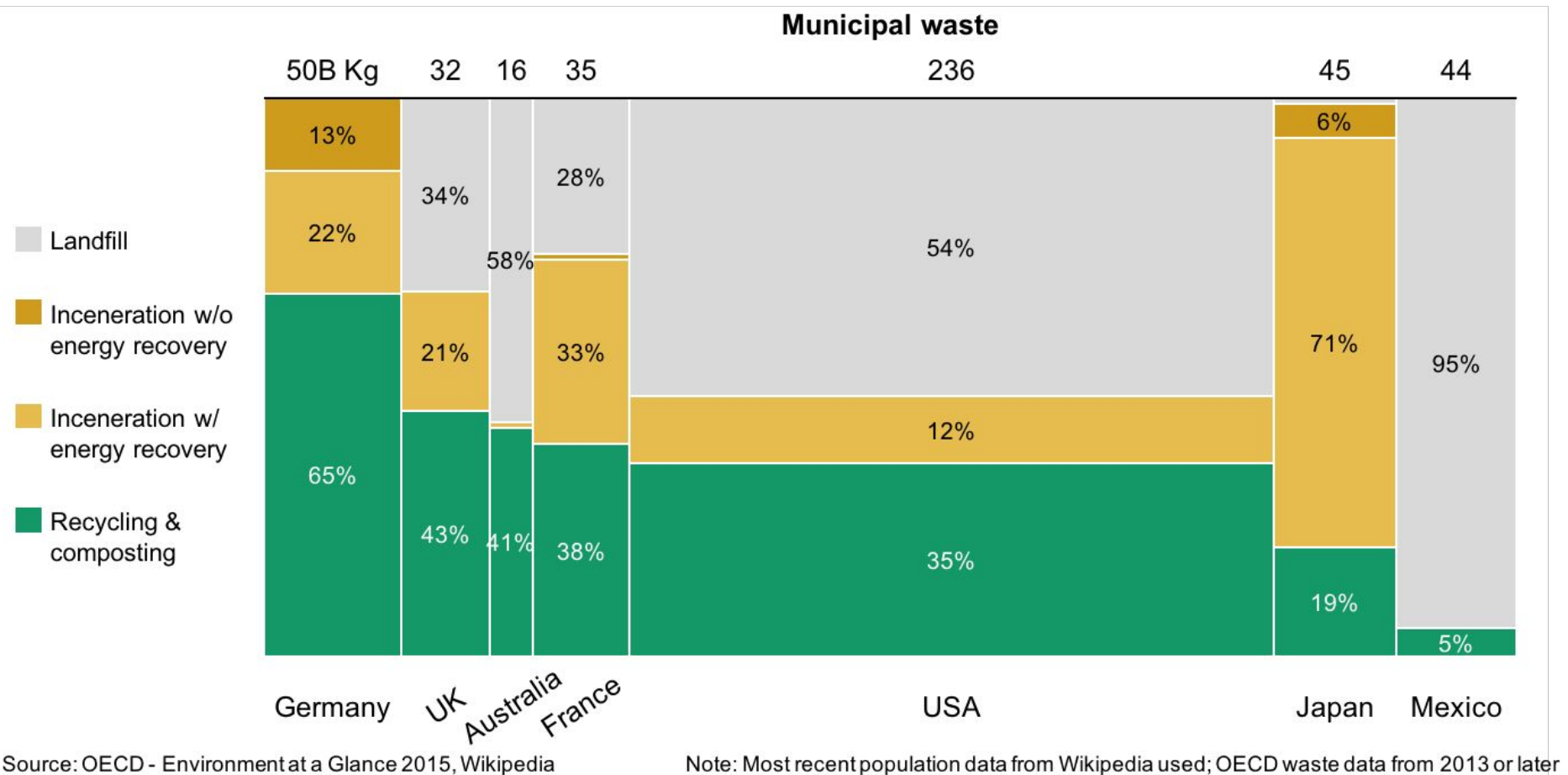


Sources: International Comparison Programme; IMF; *The Economist*

*Purchasing-power parity

Two quantities that if multiplied give a third quantity.

Quant x Quant: **Mekko Chart**

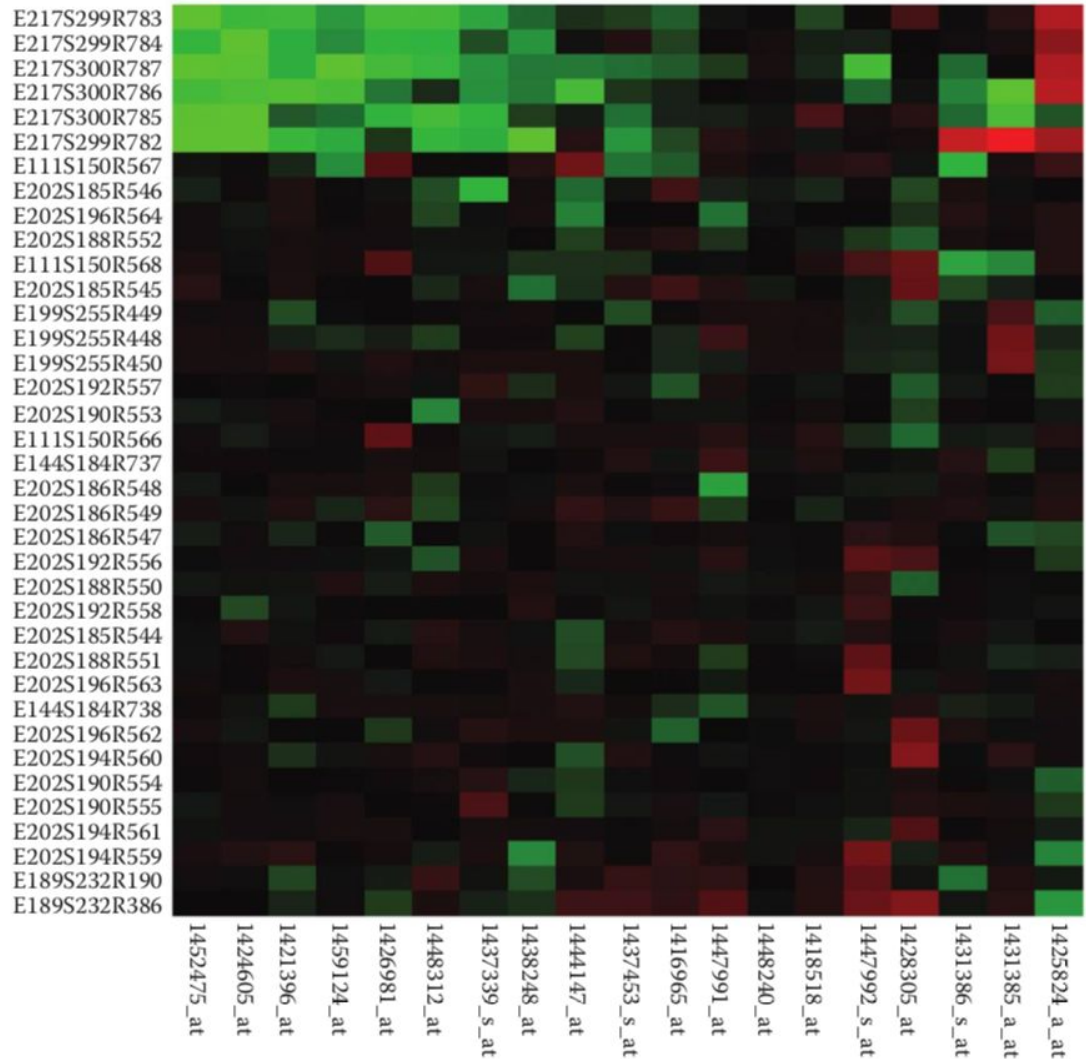


Two quantities that if multiplied give a third quantity.

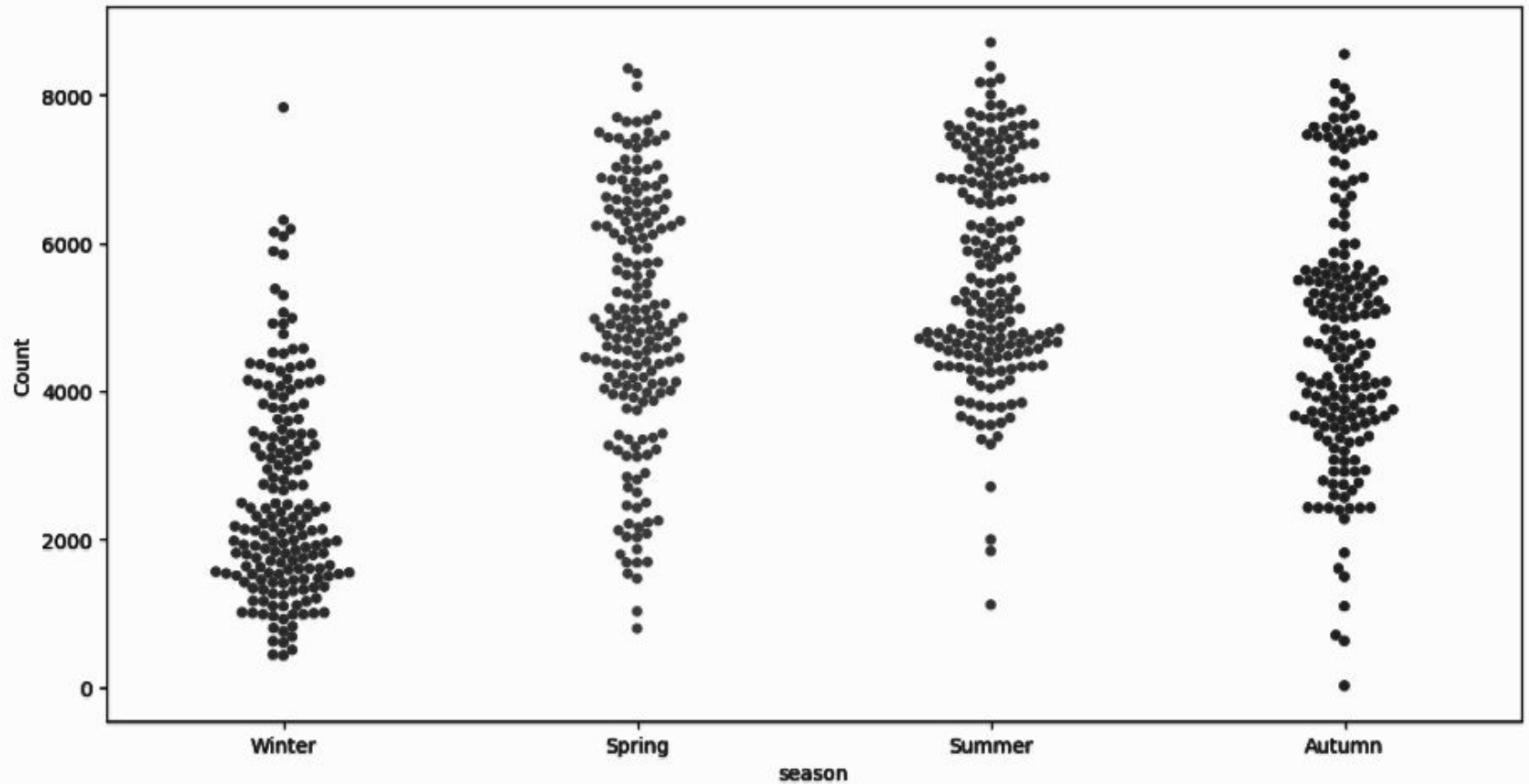
Q/C/O + O/C: Heatmaps

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

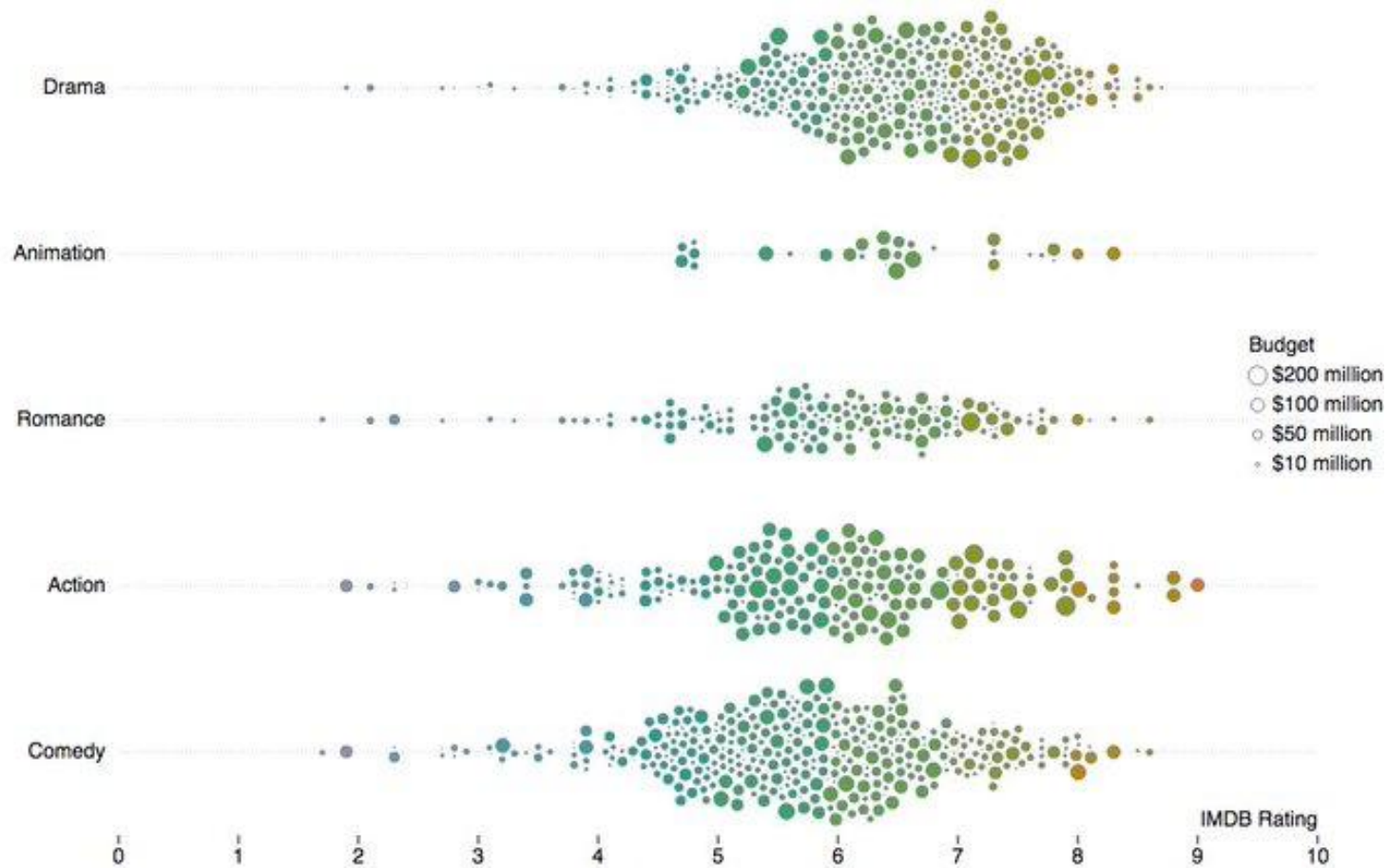
Requires complete field.



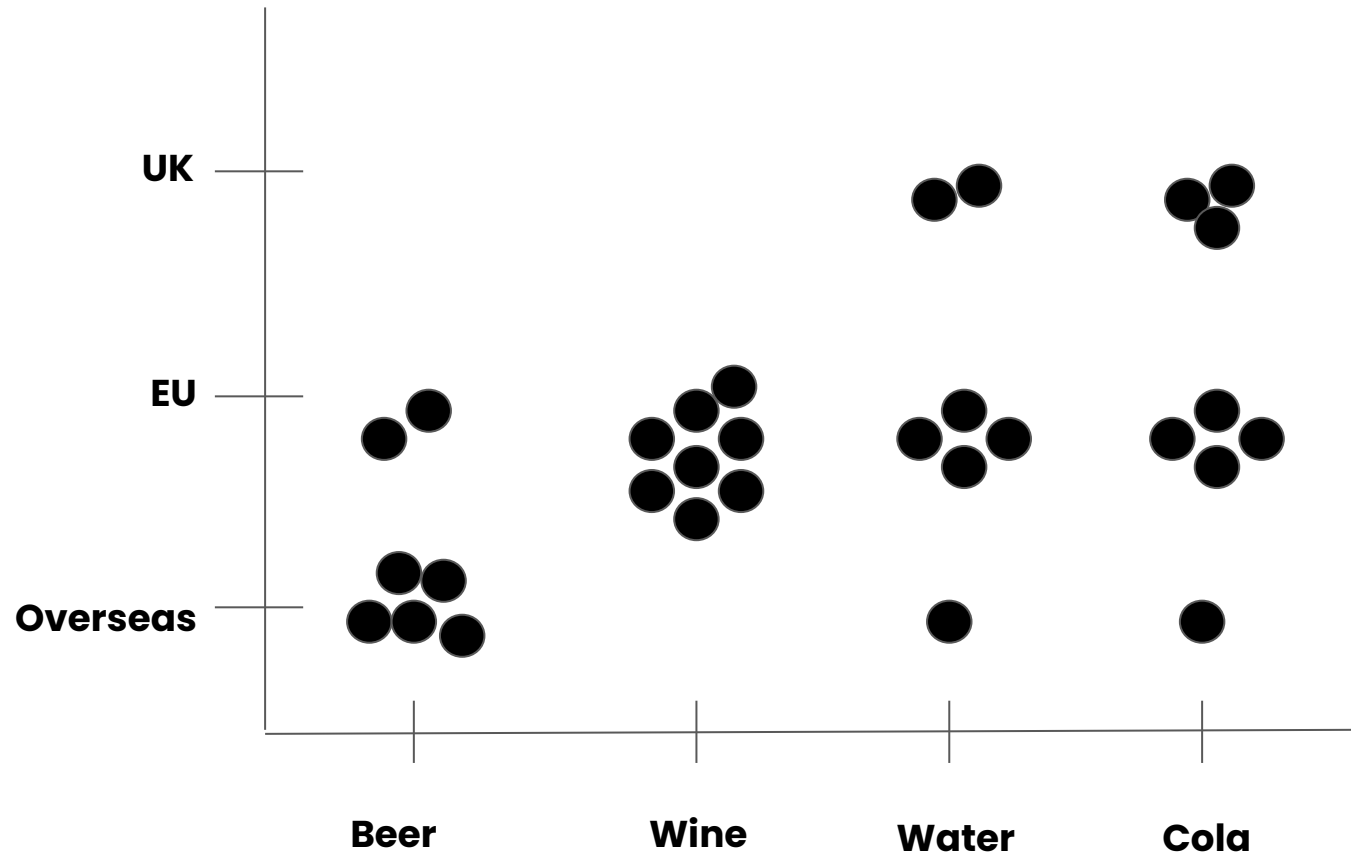
Quant + Ordered/Categorical: **Beeplots**



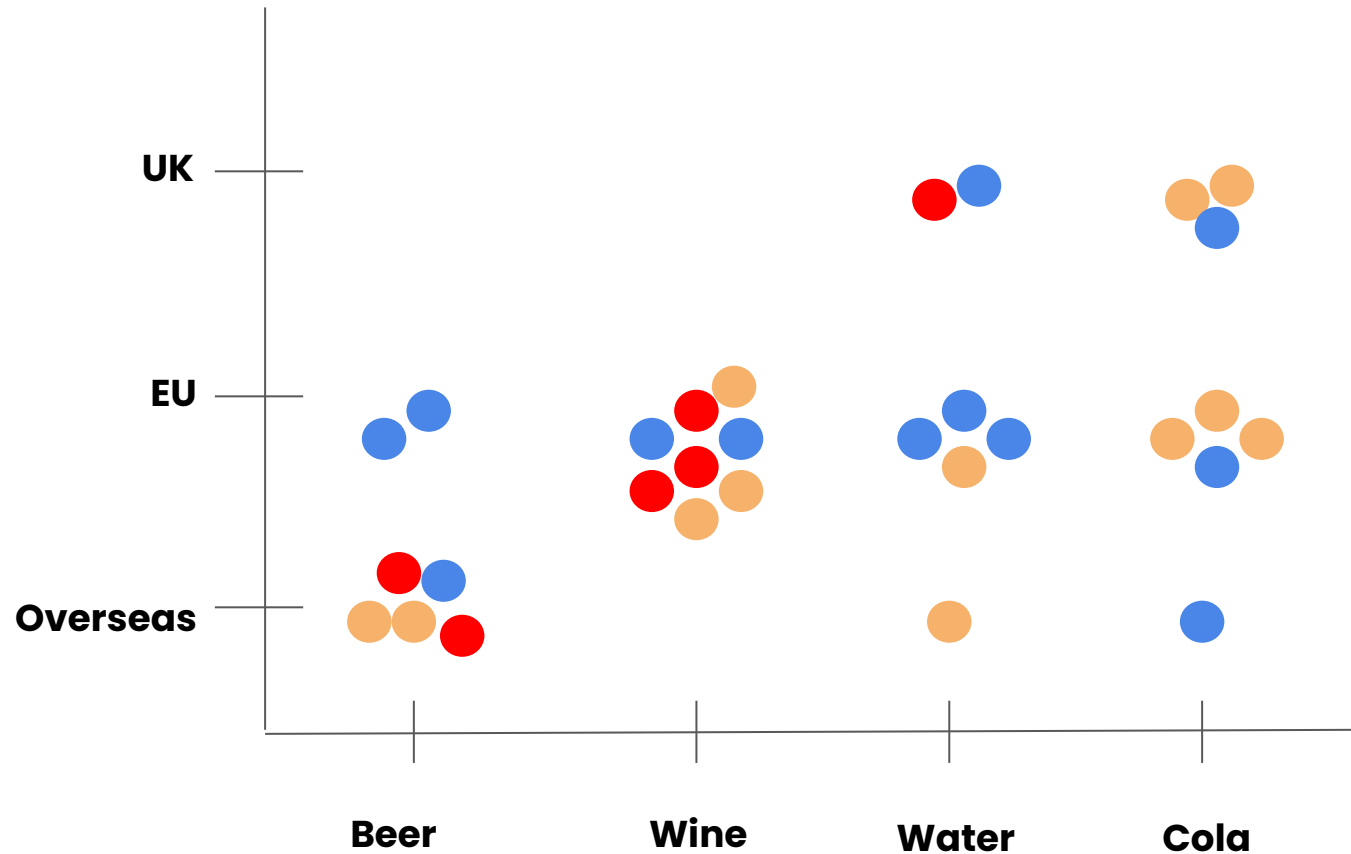
Quant + Ordered/Categorical: **Swarm plots**



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

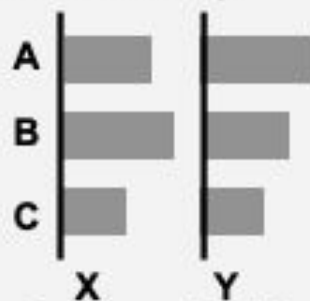


Categorical + Categorical: ***Swarm plots???***



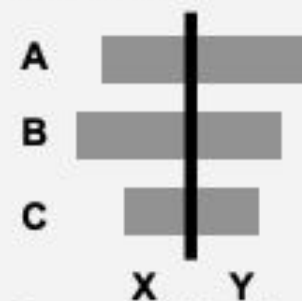
And again: **bar charts**

Bar Table X,Y



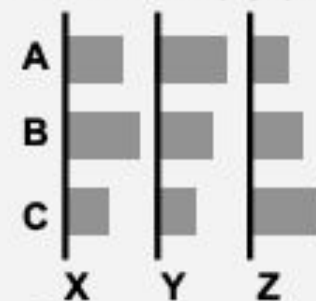
Compare X to Y,
'Small multiples'

Mirror Bar



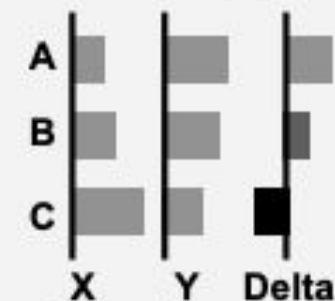
Compare X to Y

Bar Table X,Y,Z,...



Compare as many
as you like

Bar Table X,Y, Delta



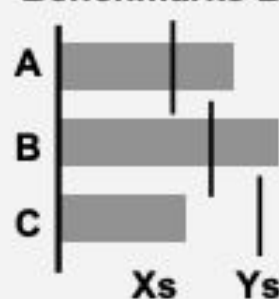
Comparisons are
slow. Plot critical
Deltas explicitly

Benchmark Bar



Compare X to a
benchmark

Benchmarks Bar



Compare X to Y.
Fancier version
called a 'Bullet graph'

Interleaved Bar



Compare X to Y
(not recommended)

Interleaves two
categories into
one spatial
dimension.
Typically better to
use Bar Table
(above) instead

... and many more!

Categories

Ordered Categories

Continuous Metrics

Metric, split by
1 category

Bar (Row)



Rows allow readable labels, while columns awkwardly turn text

Lollipop



More focus on the positions of tops. Fun factor +1

Dot Plot



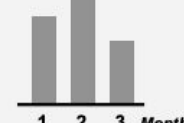
A non-zero y-axis base may be less misleading here

Bar (Column)



Histogram. Boxes help convey the underlying bins

Bar (Column)



Time moves horizontally. So use Column, not Row

Area



Adds continuity to x-axis

Line



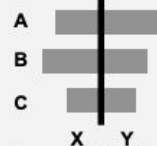
A non-zero y-axis base may be less misleading here

Bar Table X,Y



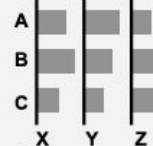
Compare X to Y, 'Small multiples'

Mirror Bar



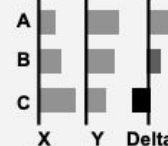
Compare X to Y

Bar Table X,Y,Z,...



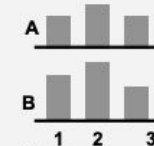
Compare as many as you like

Bar Table X,Y, Delta



Comparisons are slow. Plot critical Deltas explicitly

Bar Table



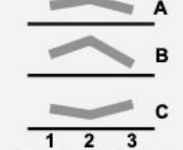
Compare a continuous metric across a category

Bar Line Table



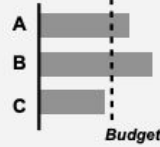
Compare two metrics

Line Table



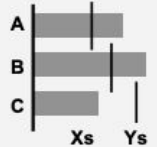
Trends visible, but use Lines (below) if precision is key

Benchmark Bar



Compare X to a benchmark

Benchmarks Bar



Compare X to Y. Fancier version called a 'Bullet graph'

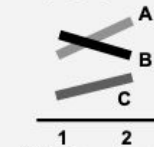
Interleaved Bar



Compare X to Y (not recommended) ←

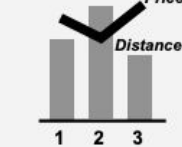
Interleaves two categories into one spatial dimension. Typically better to use Bar Table (above) instead

Slopegraph



With two values, slope encodes delta

Dual Axis



Use (above) instead. Crossings here are salient, but meaningless

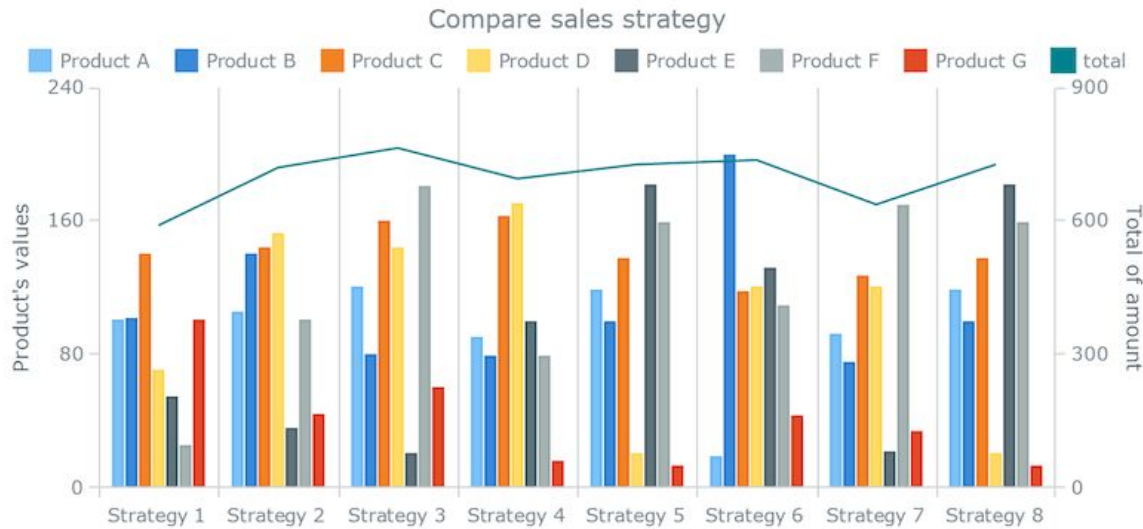
Lines



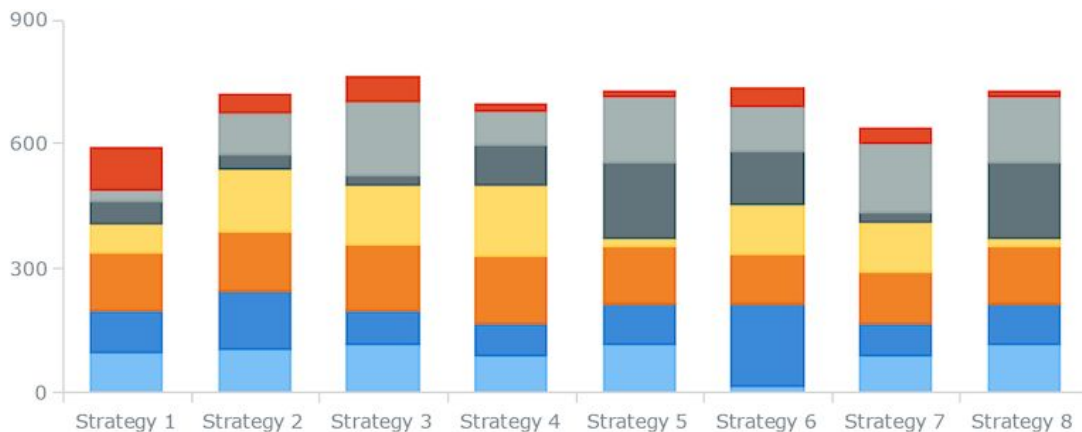
Compare many. Getting spaghetti? Use Line-Table (above)

... by
2+ categories

Stacked Bar Charts



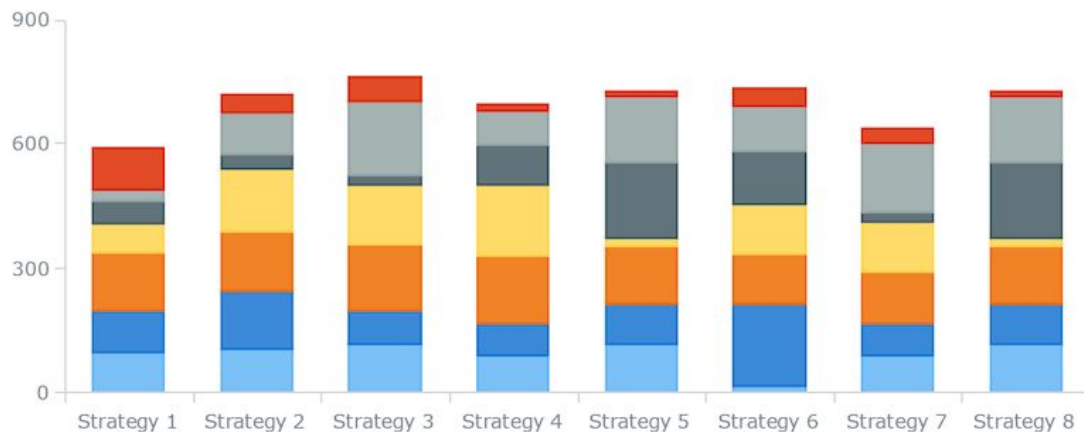
- Individual distributions
- Individual value comparison
- Details



Stacked Bar Charts

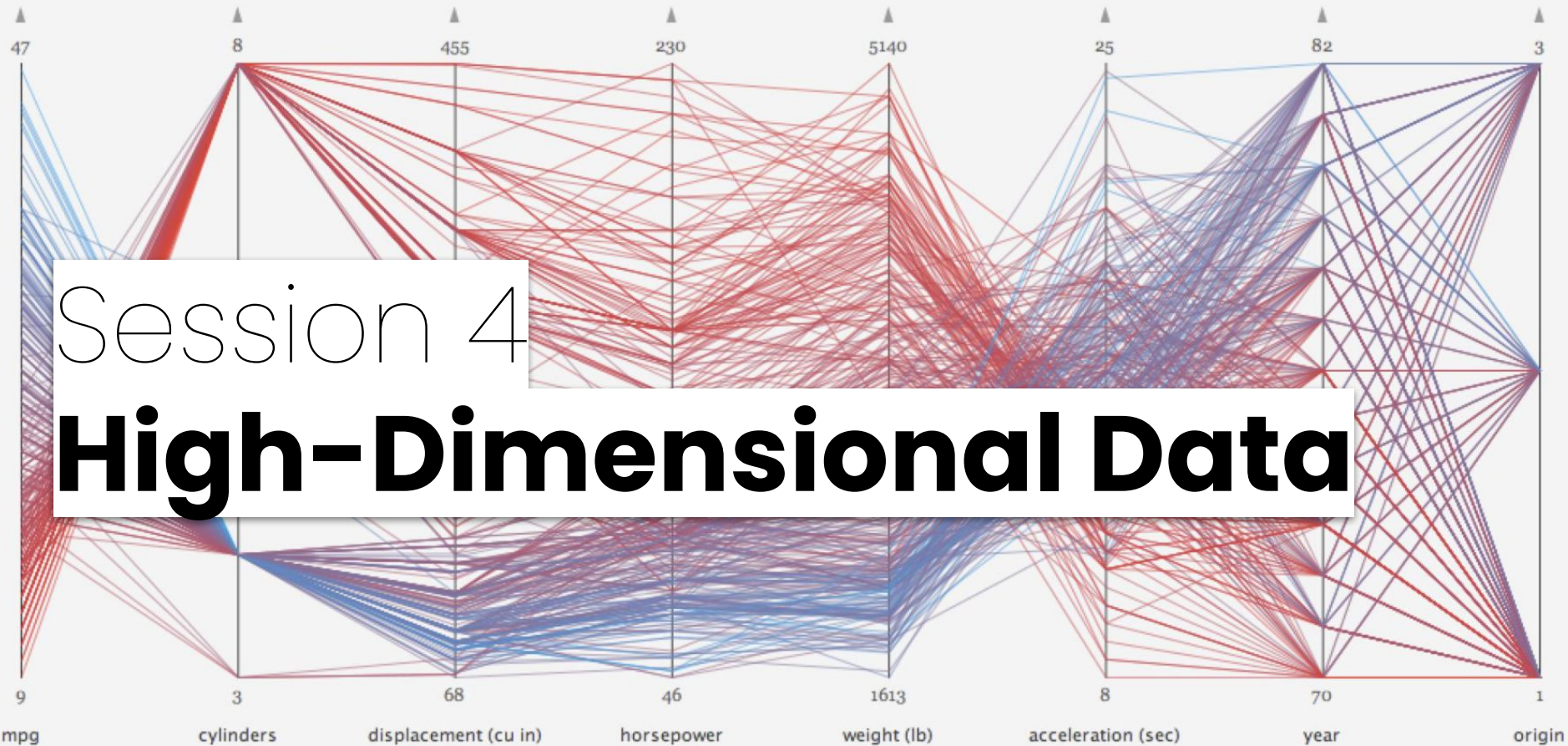


- Individual distributions
- Individual value comparison
- Details



- Sum comparison
- Overview

> requires good color palette!



Session 4

High-Dimensional Data



Online Course
**Data Visualization
for Professionals**



THE UNIVERSITY
of EDINBURGH

Benjamin Bach

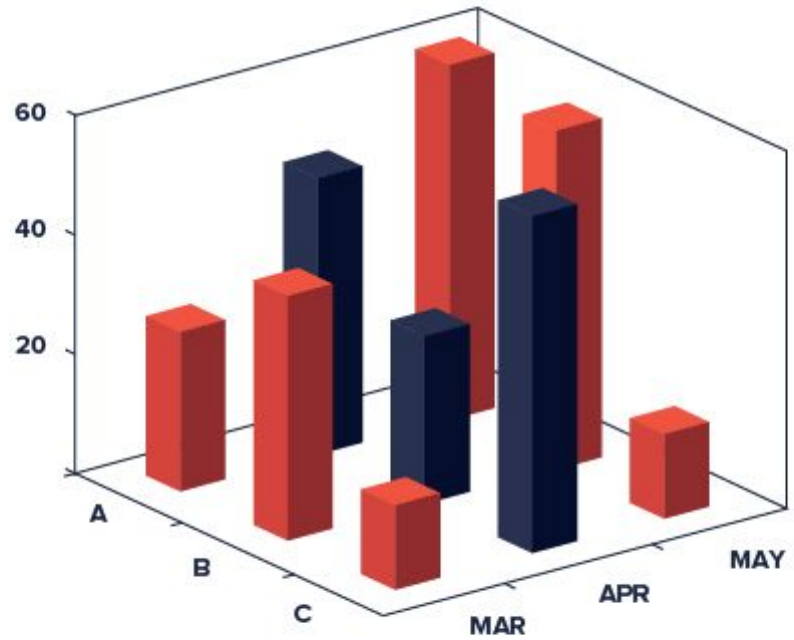
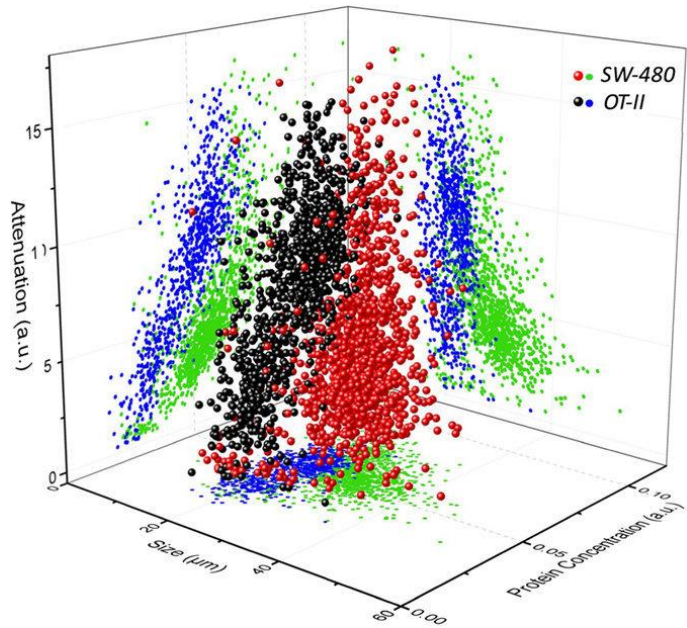
June 2022

<http://benjbach.me>

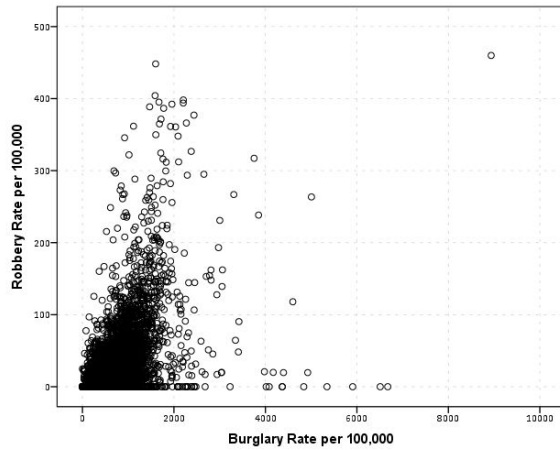
<https://datavis-online.github.io>

-- Not for external use --

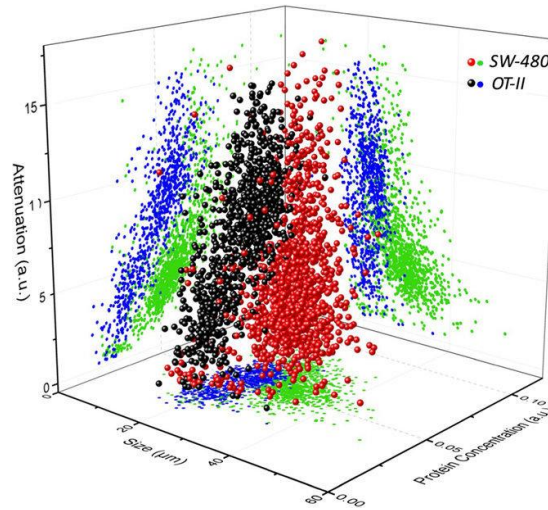
3 Dimensions...



...4 Dimensions?



2-dimensional

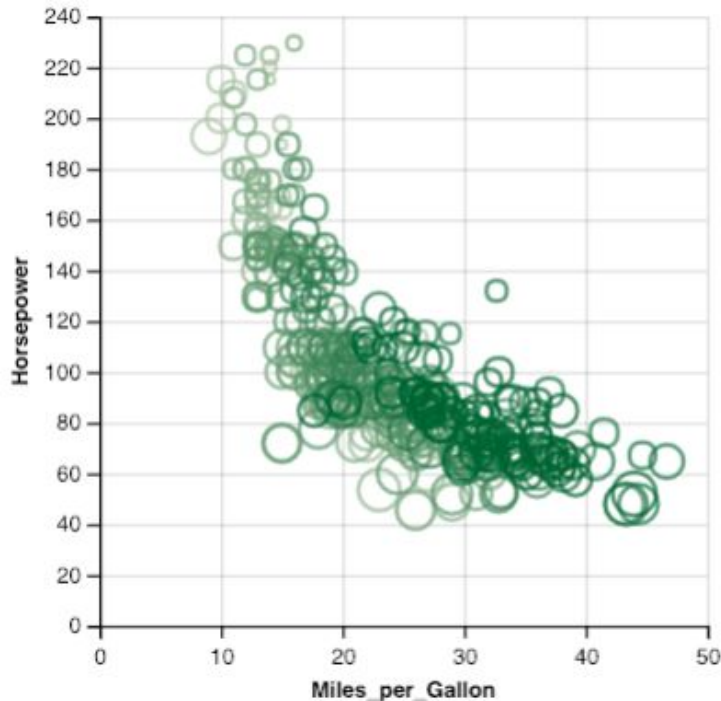


3-dimensional

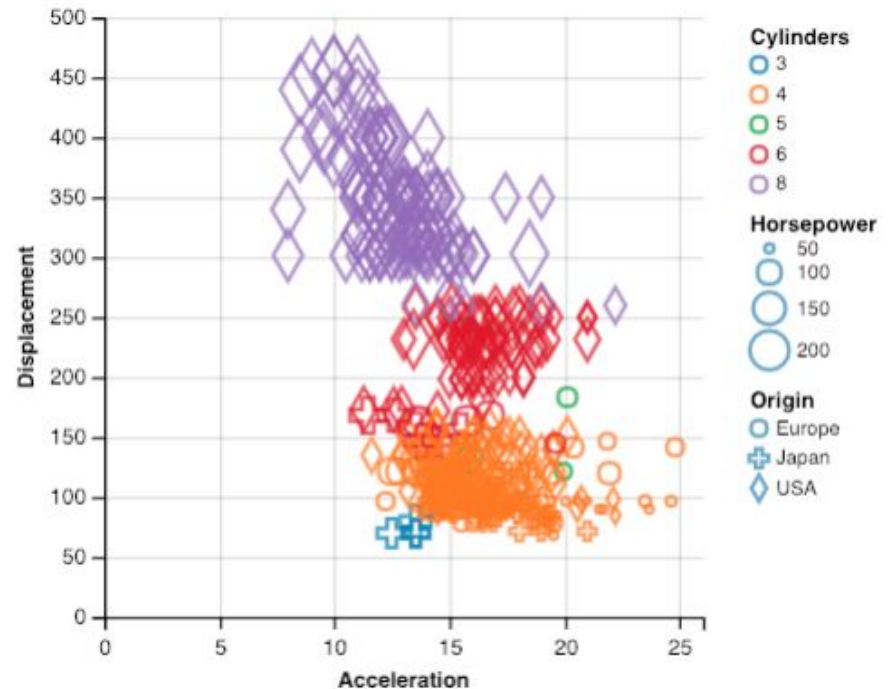
?

3-dimensional

1. Additional Visual Variable(s)

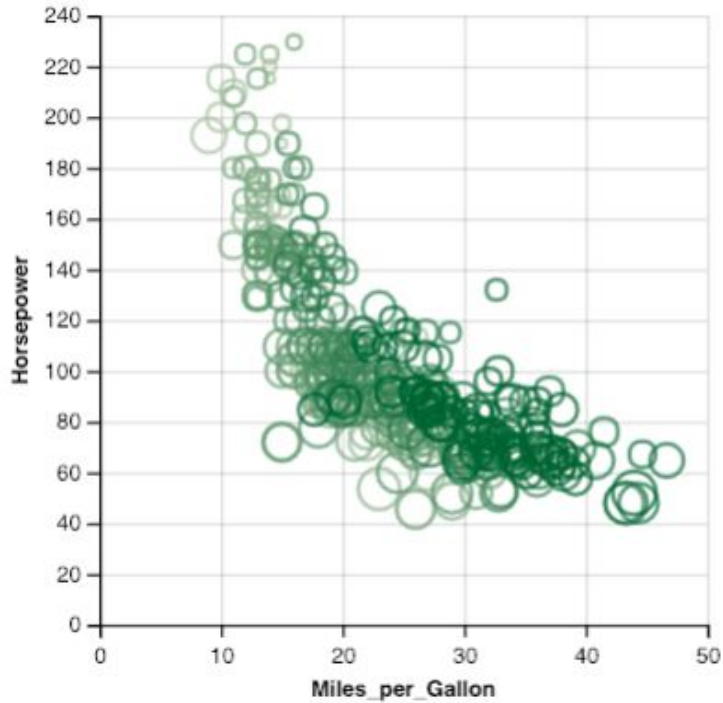


2-dimensional
+ 1 dimension (size)



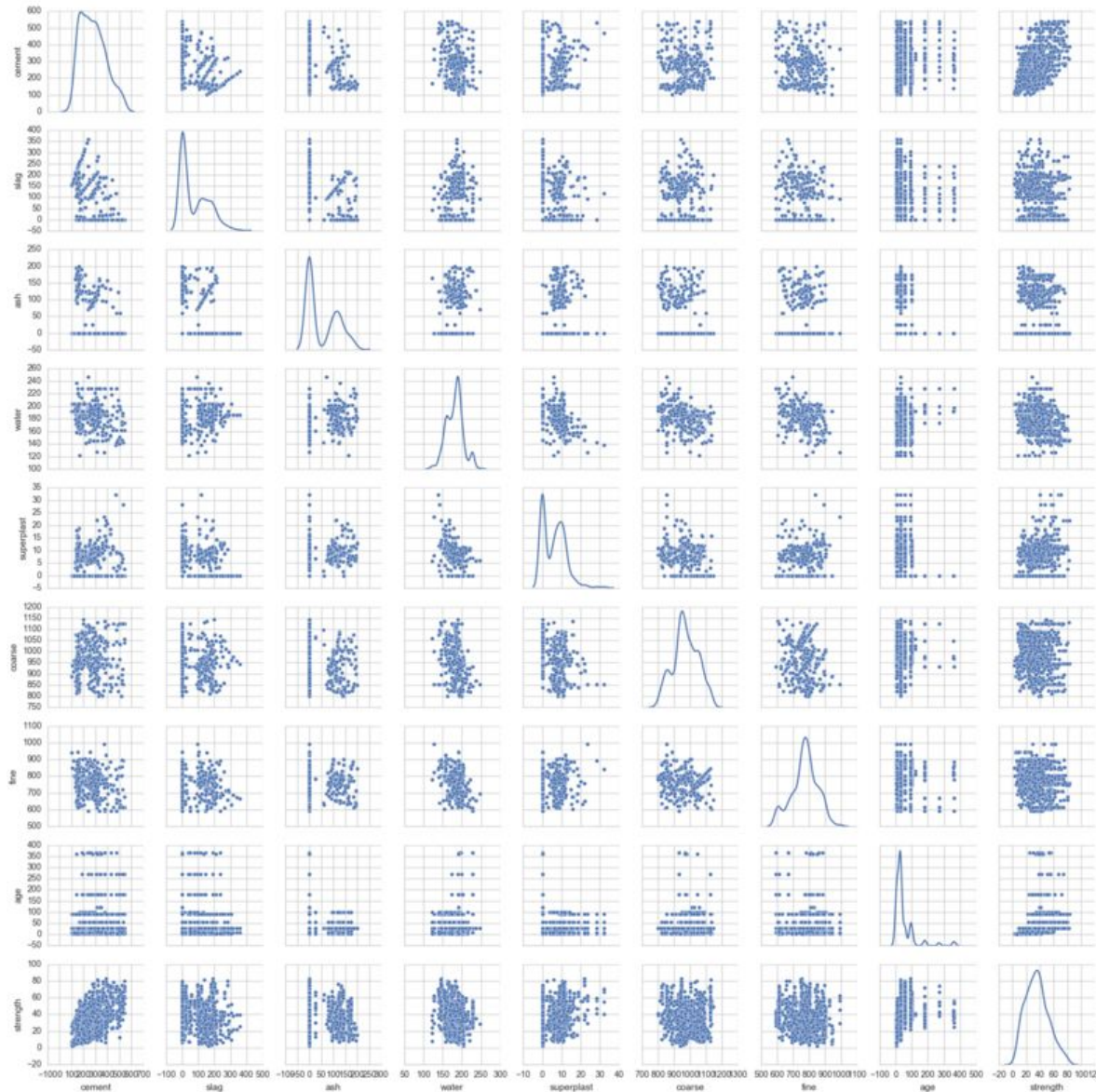
2-dimensional
+ 3 dimension (shape
+ color + size)

1. Additional Visual Variable(s)



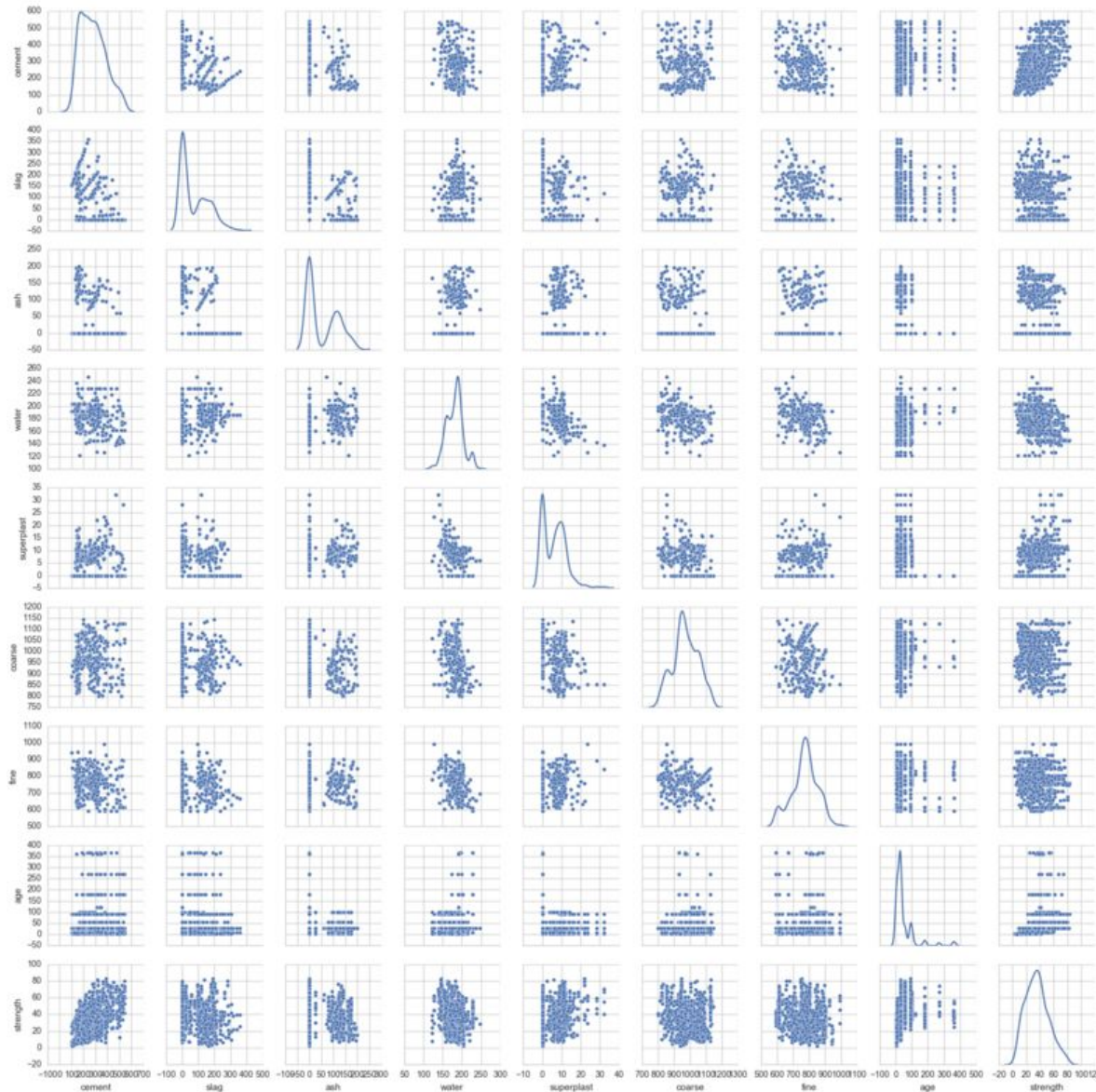
2-dimensional
+ 1 dimension (size)

2. Pairwise comparison: **Scatterplot Matrix**



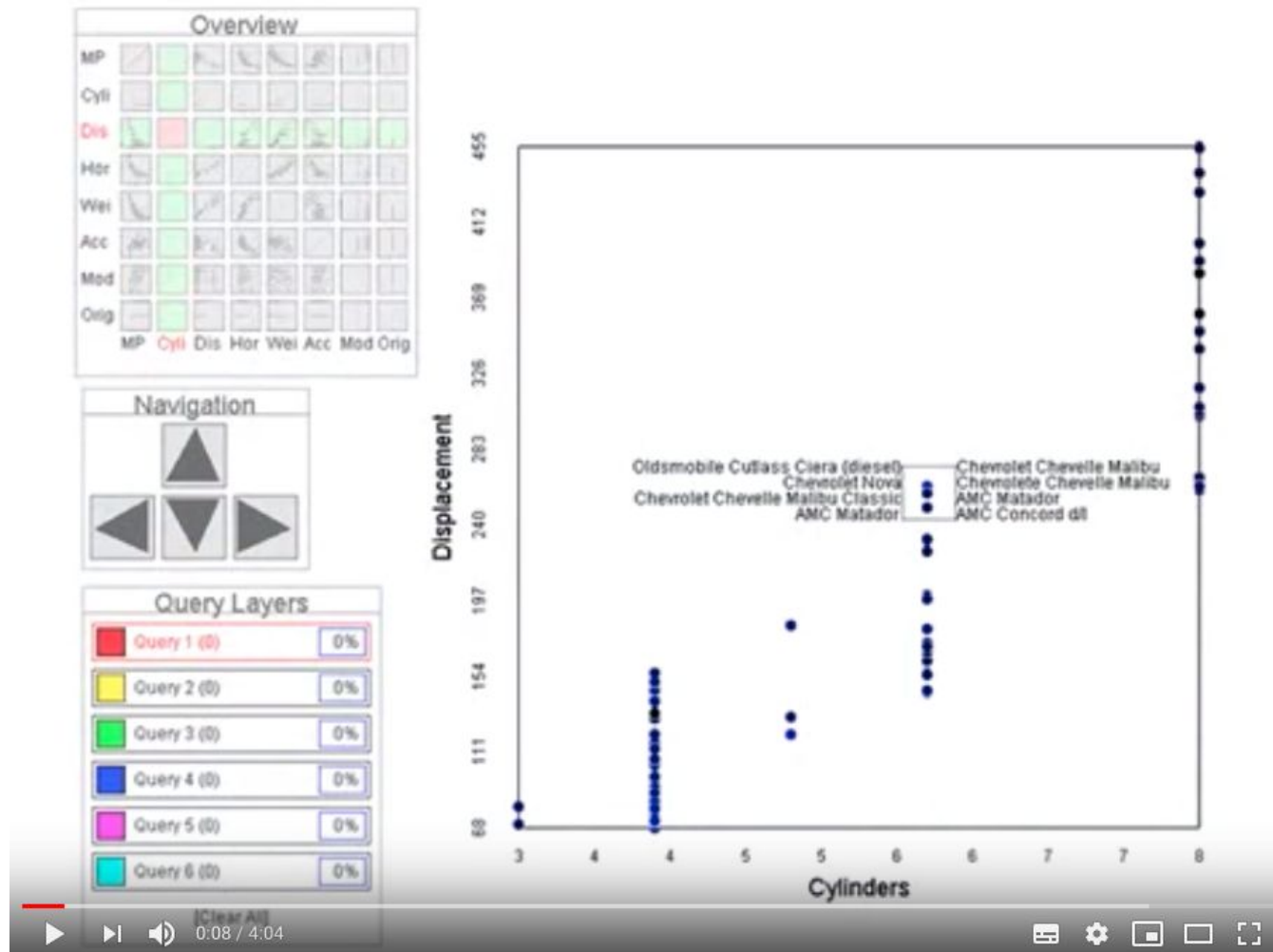
- + Scalable
- + Provides overview
- + Can use additional visualizations
- + Easy to understand and decode

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**

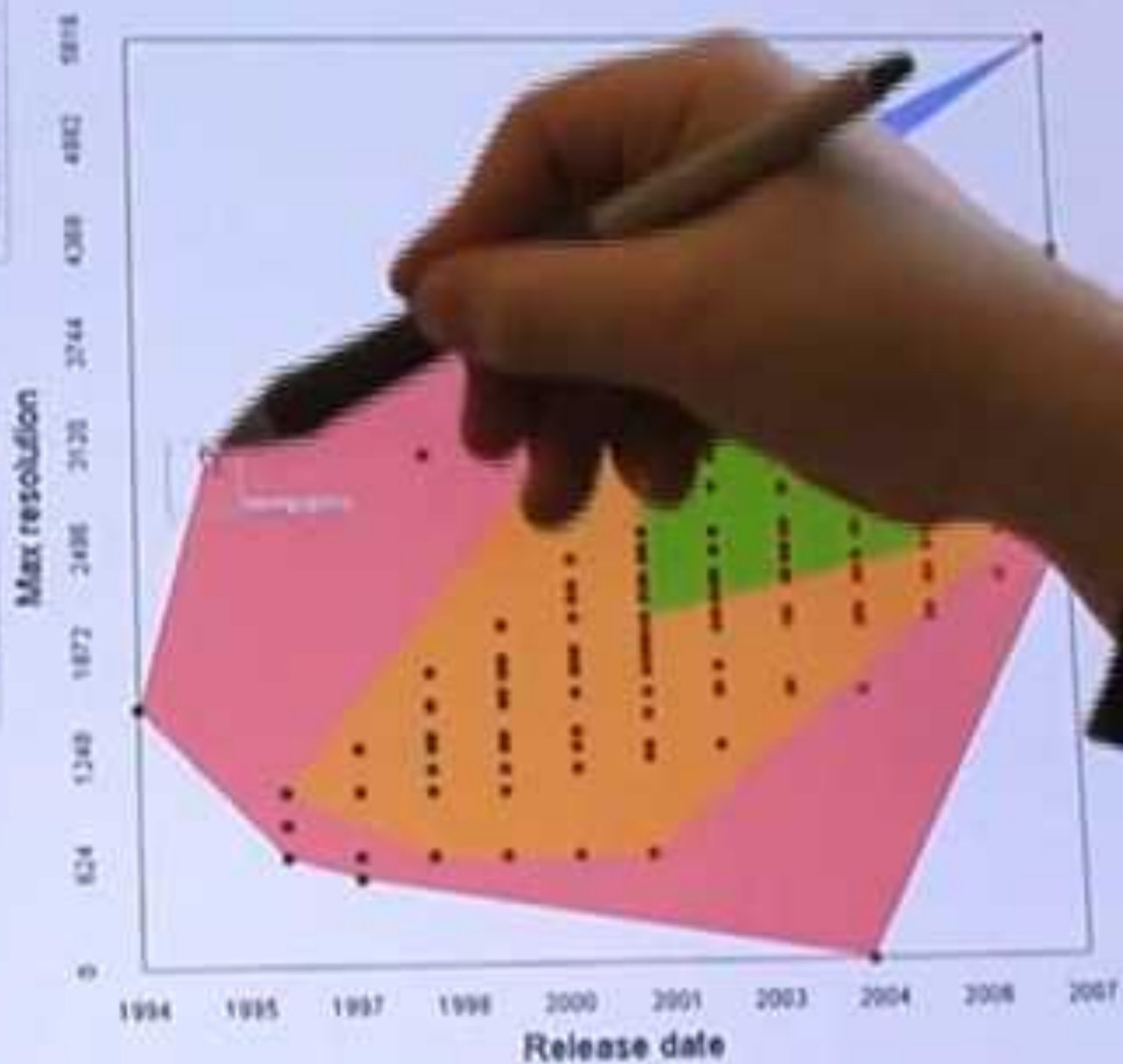


<https://www.youtube.com/watch?v=2bYIRcO-gwg>

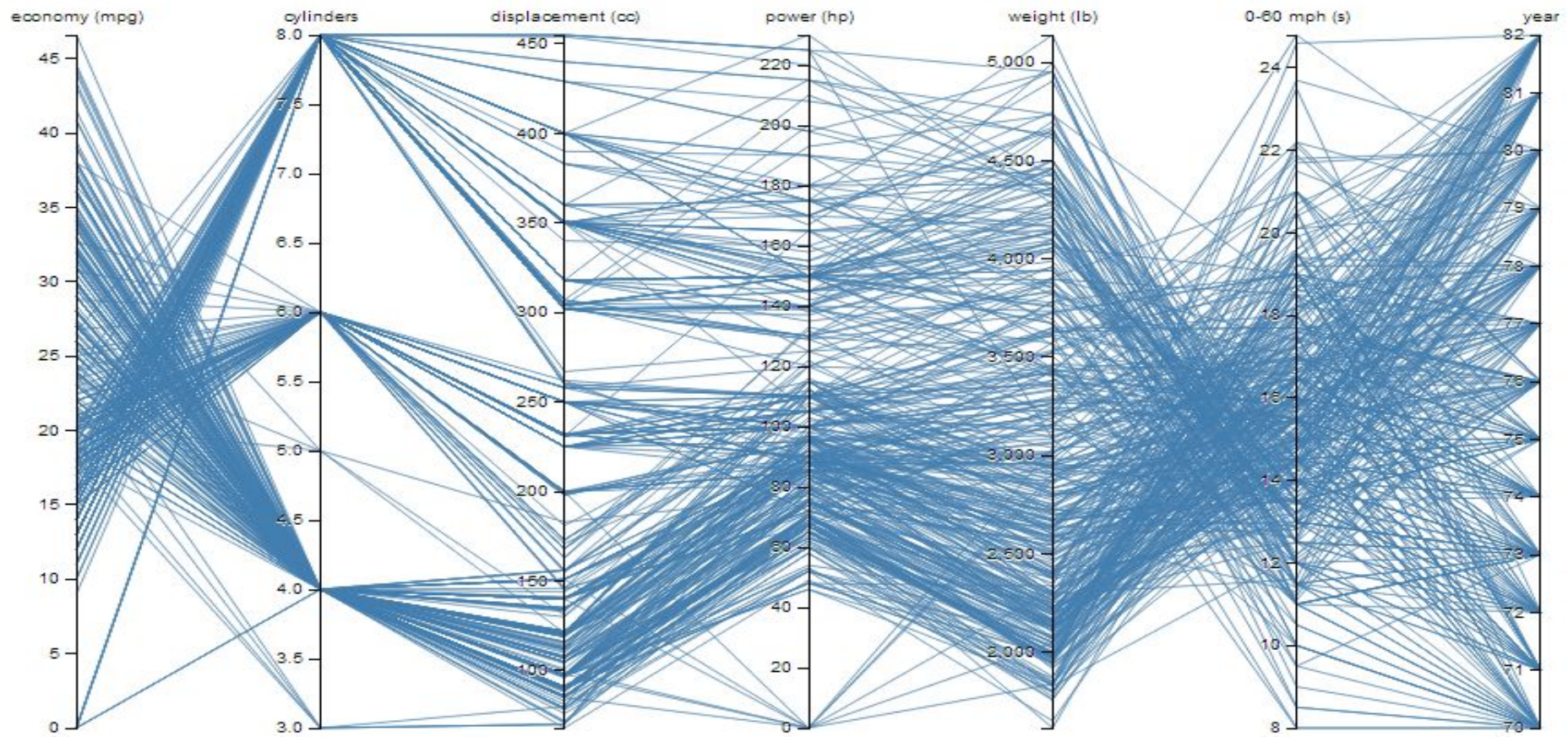
Elmqvist, Niklas, Pierre Dragicevic, and Jean-Daniel Fekete. "Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation." *IEEE transactions on Visualization and Computer Graphics* 14.6 (2008): 1539-1148.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

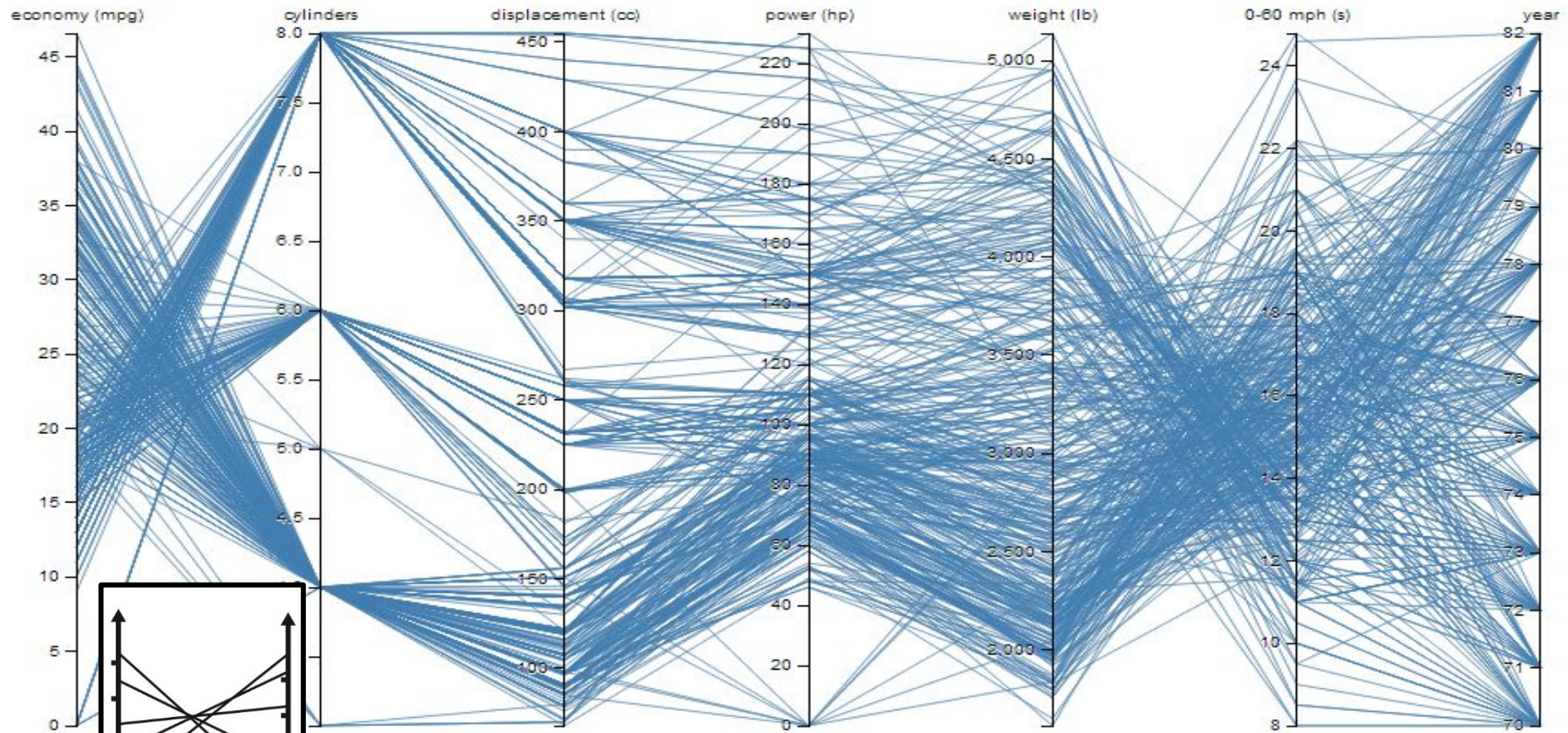
Query Layers		
Query 1 (2000)		1.0
Query 2 (2000)		1.0
Query 3 (2000)		1.0
Query 4 (2000)		1.0
Query 5 (2000)		1.0
Query 6 (2000)		1.0
© 2000		



4. Parallel Coordinates Plot (PCP)

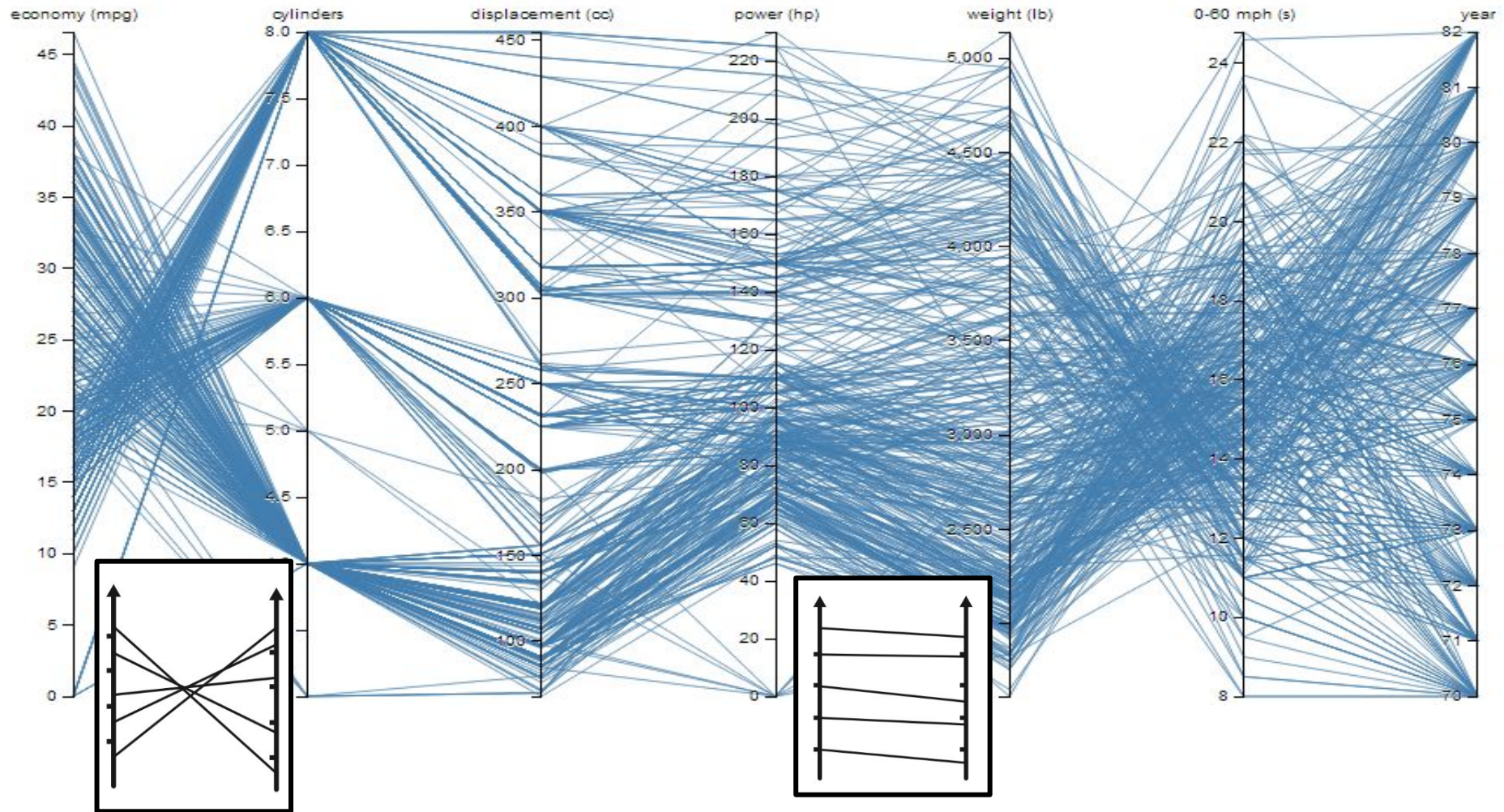


4. Parallel Coordinates Plot (PCP)



- + Scalable
- + Consise
- + Good overview

4. Parallel Coordinates Plot (PCP)

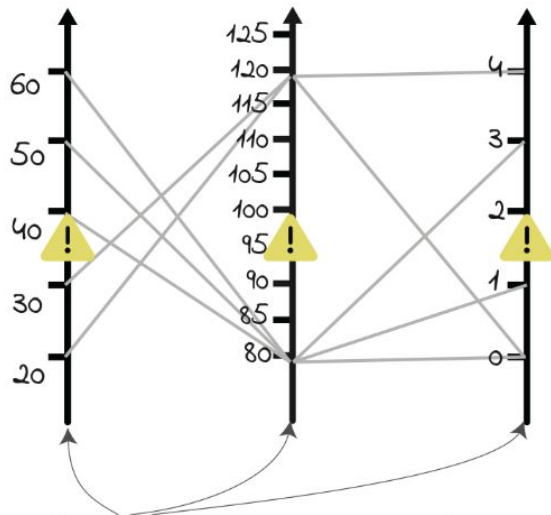


- + Scalable
- + Consise
- + Good overview

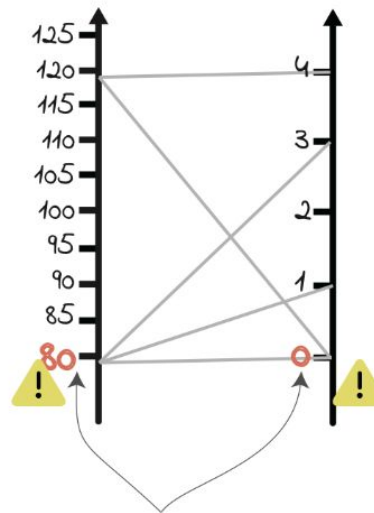
- Depending on ordering
- Can suffer from clutter
- Visual path following can be hard

PCPs Pitfalls

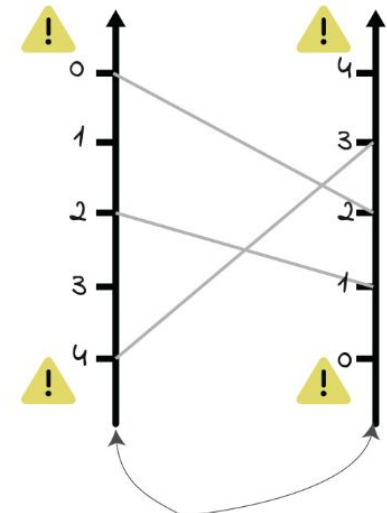
Axis scales



Truncated axes



Axes order



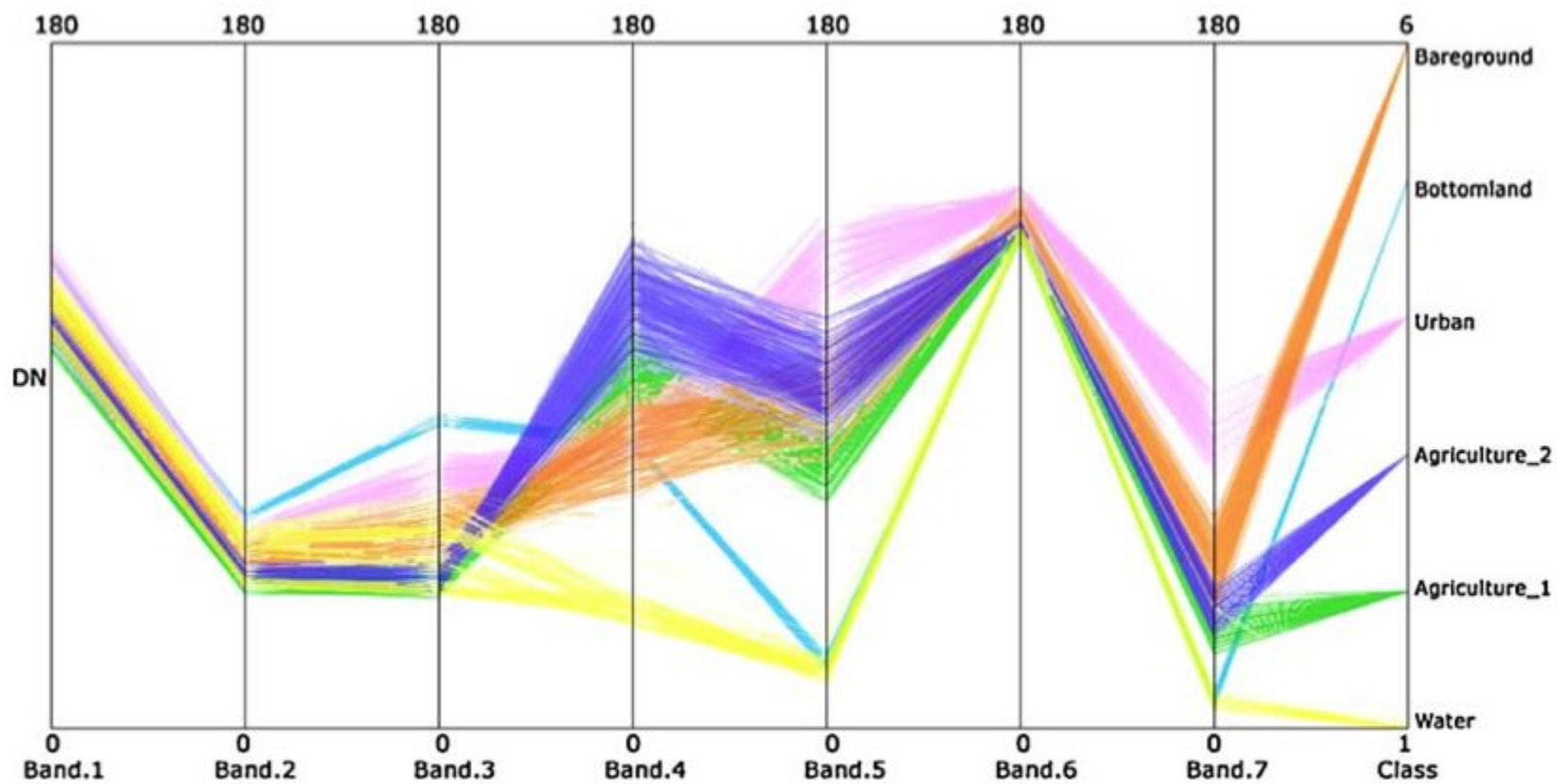
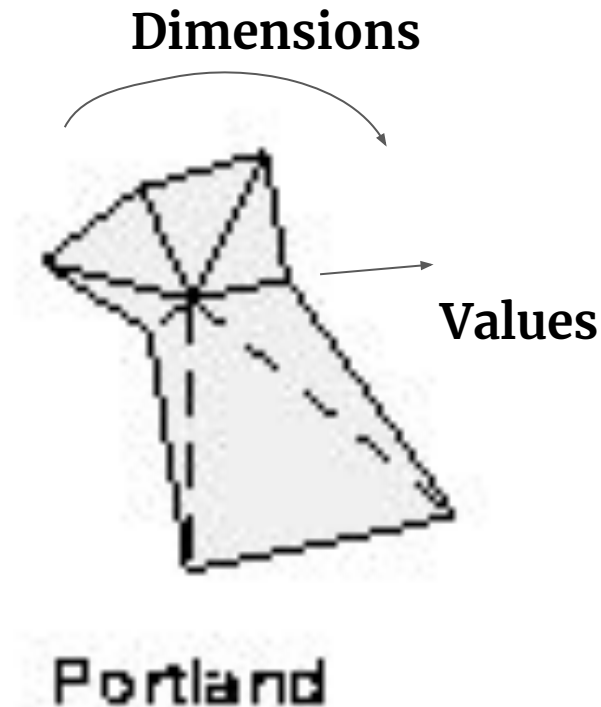


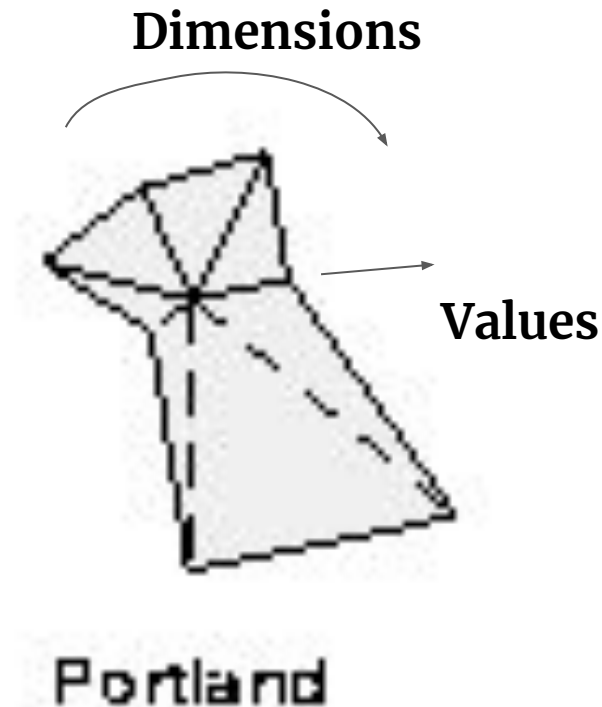
Fig. 3. PCP of sample data.

3. Glyphs: Star glyphs



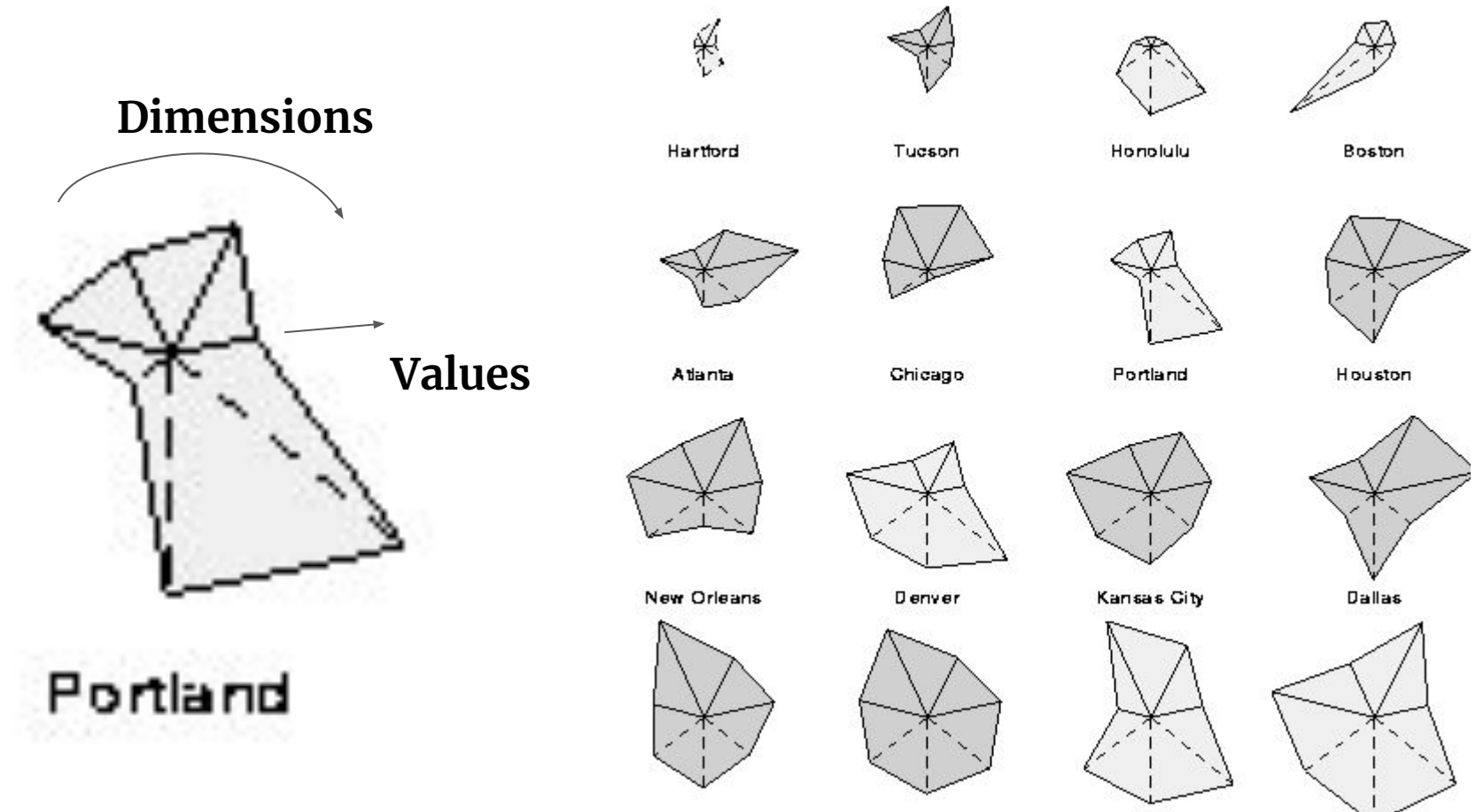
- 1) Data glyphs are [a] technique, in which single data points are encoded individually by assigning their dimensions to one or more marks and their visual variables

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

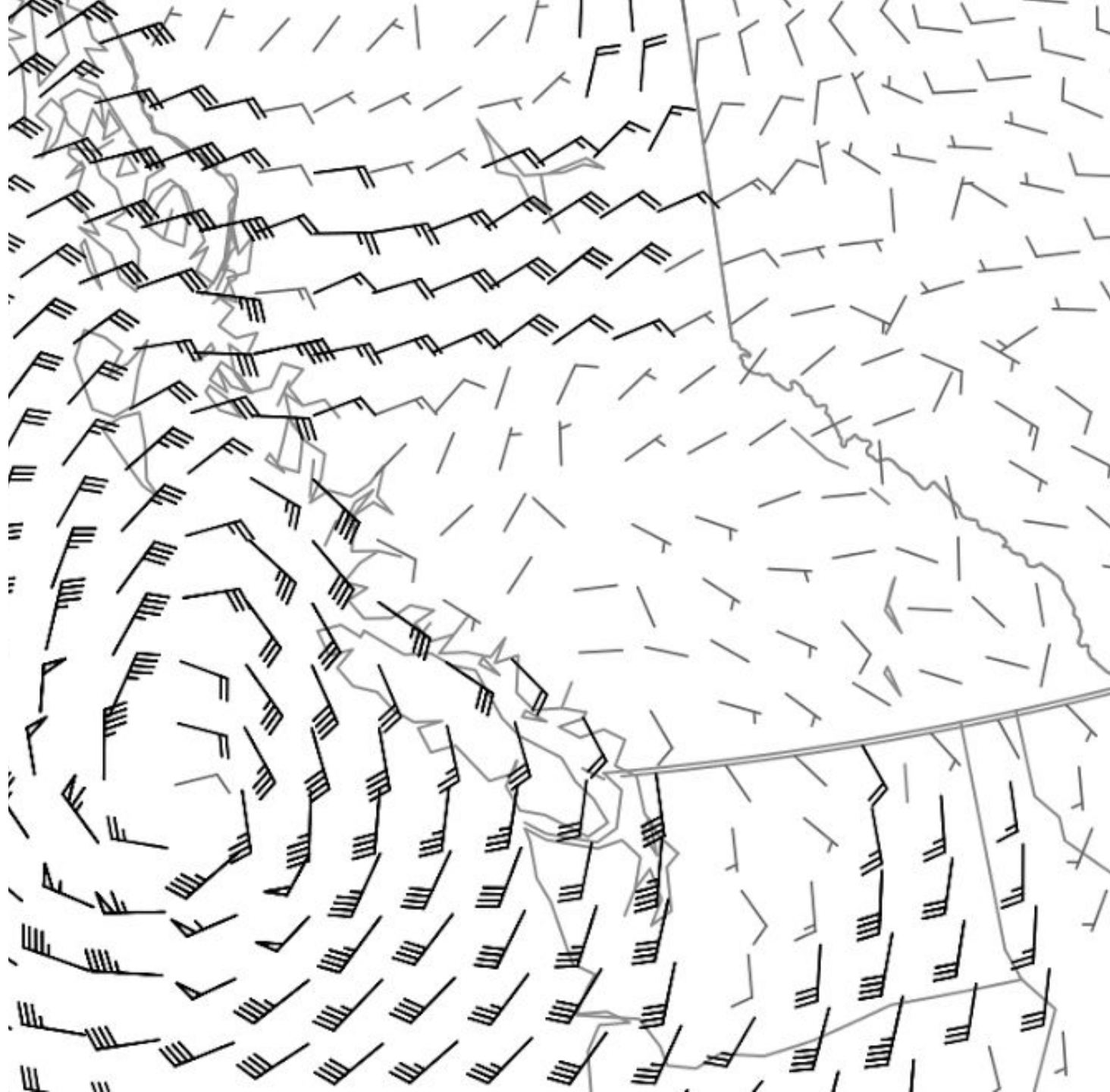
3. Glyphs: Star glyphs



- + Comparison
- + Outlier, Trends
- + Individual values

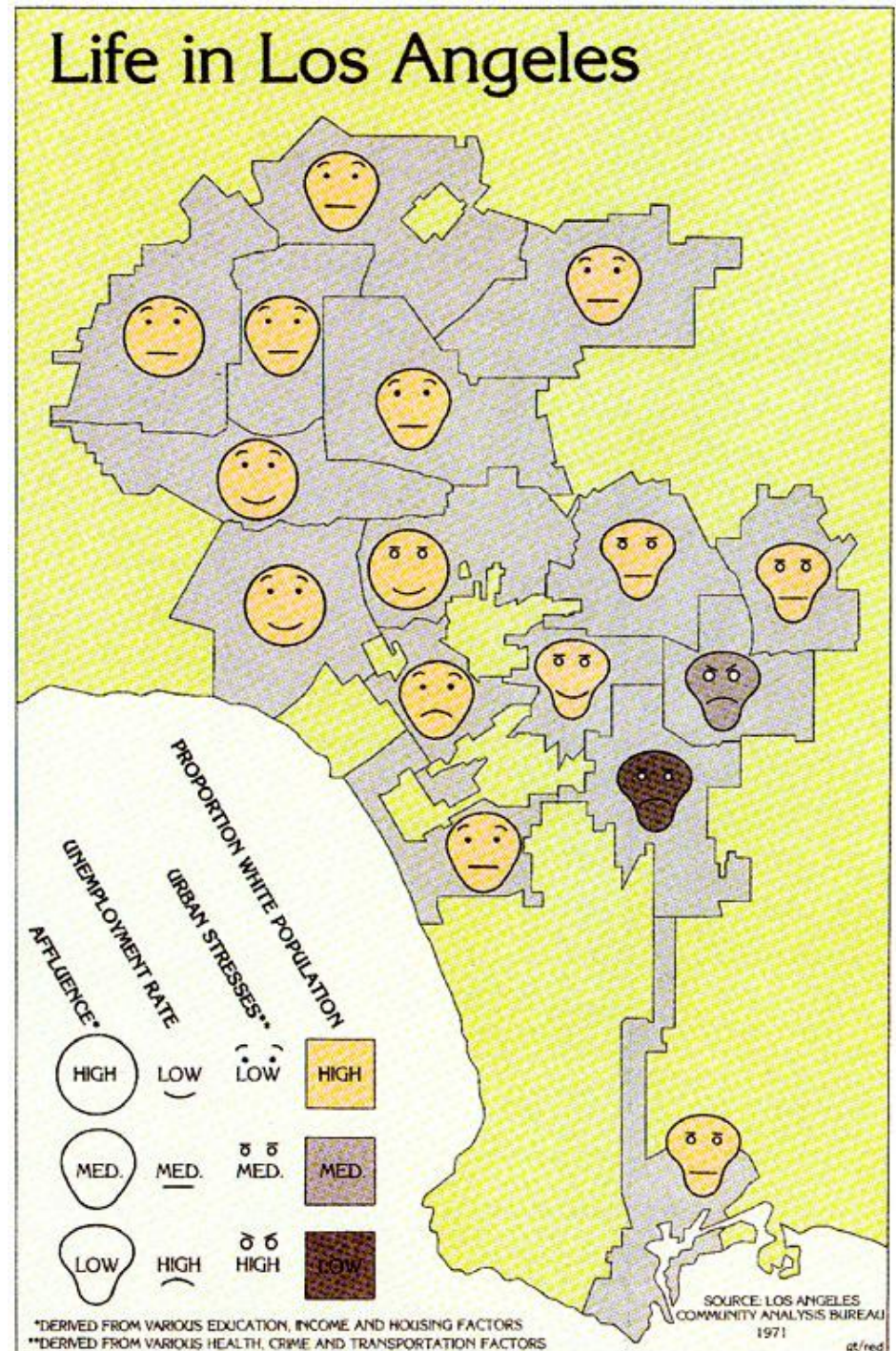
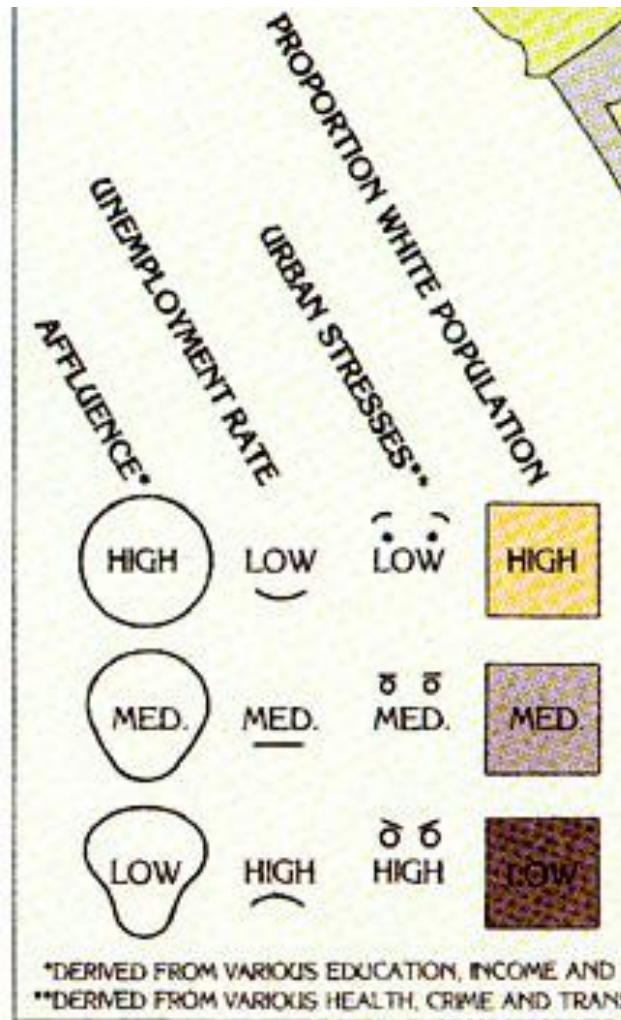
- Precise comparison
- Cluster dimensions

Weather glyphs



Chernoff Faces

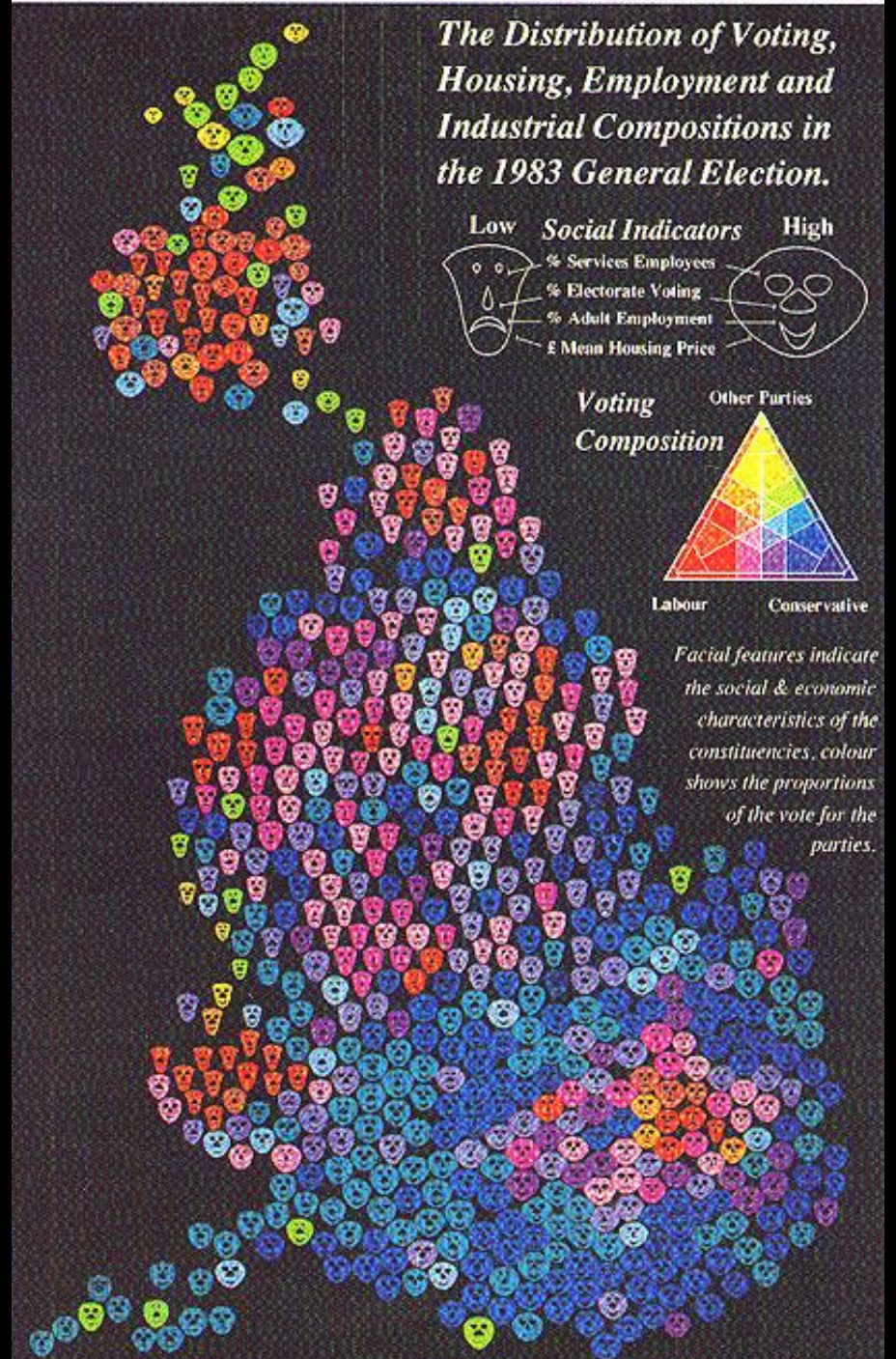
(Glyphs for geodata)



Chernoff Faces

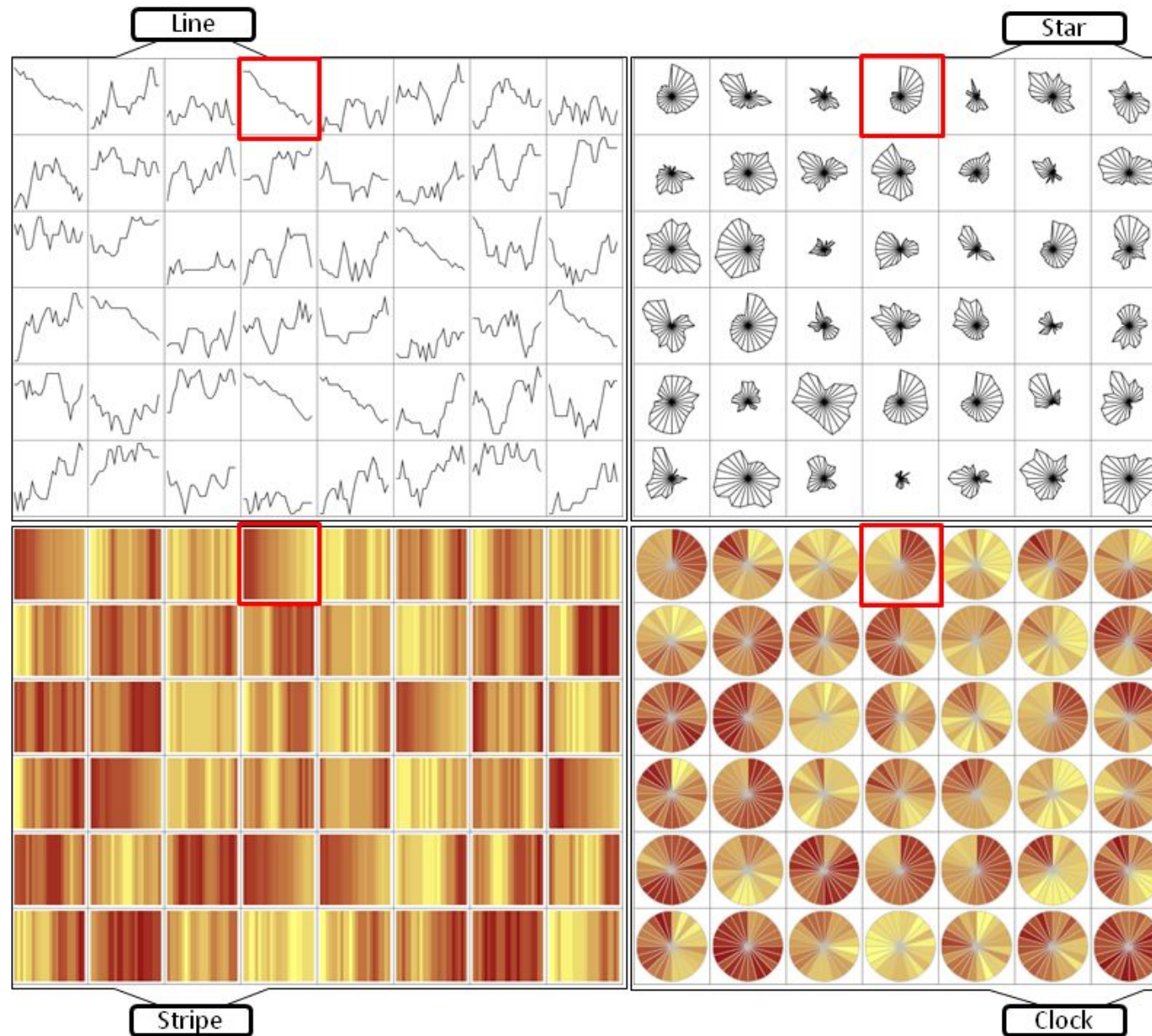
(Glyphs for geodata)

- + Individual values
- + Spatial correlation
- Some vis-variables are more prominent
- Some vis-variables are hard to perceive and estimate

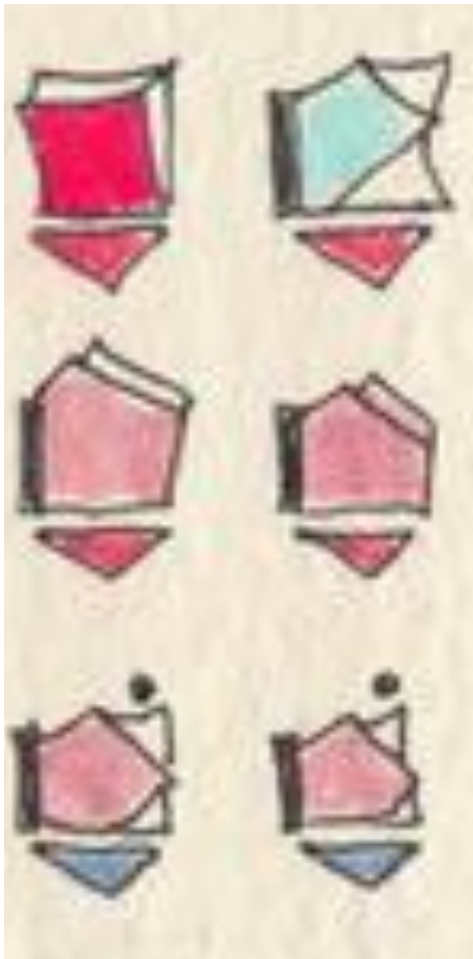


Glyphs for Temporal data

- Line glyphs for trends and peaks
- Radial for value look-up



Dear Data (fantastic glyphs)



66 Dear Data

WEEK 52 - ... Goodbye! ^{#Last week of Dear Data}

HOW TO READ IT:

This week I tracked all the "Goodbyes / ByeByes / goodnight" I said. Each Element is a goodbye I said, in chronological order.

SHAPE = "HOW"

□ = in real life

◻ = over the phone

◻ = skype / hangout

◻ = in public (public speech)

◻ = farewell to my old apartment!

* = missed goodnight to my boyfriend 'cause I fell asleep too early

→ physical contact : • kiss
• hug
• hand-shake
(no physical contact if dot is missing)

color = to whom?

color of the triangle = did I add something?

△ = good luck!

△ = have fun / enjoy / divert!

△ = talk soon / see you later / a dopo!

△ = thanks! / thanks for... / grazie

△ = Bacchus!

△ = have a nice day / evening

△ = love you!

■ mum

■ boyfriend

■ friend

■ you and Rebecca

■ coworker

■ acquaintance

■ client

■ grandma

■ stranger (sales man / waitress...)

from:
G. LUPU

BRUCKLYN
NY - USA



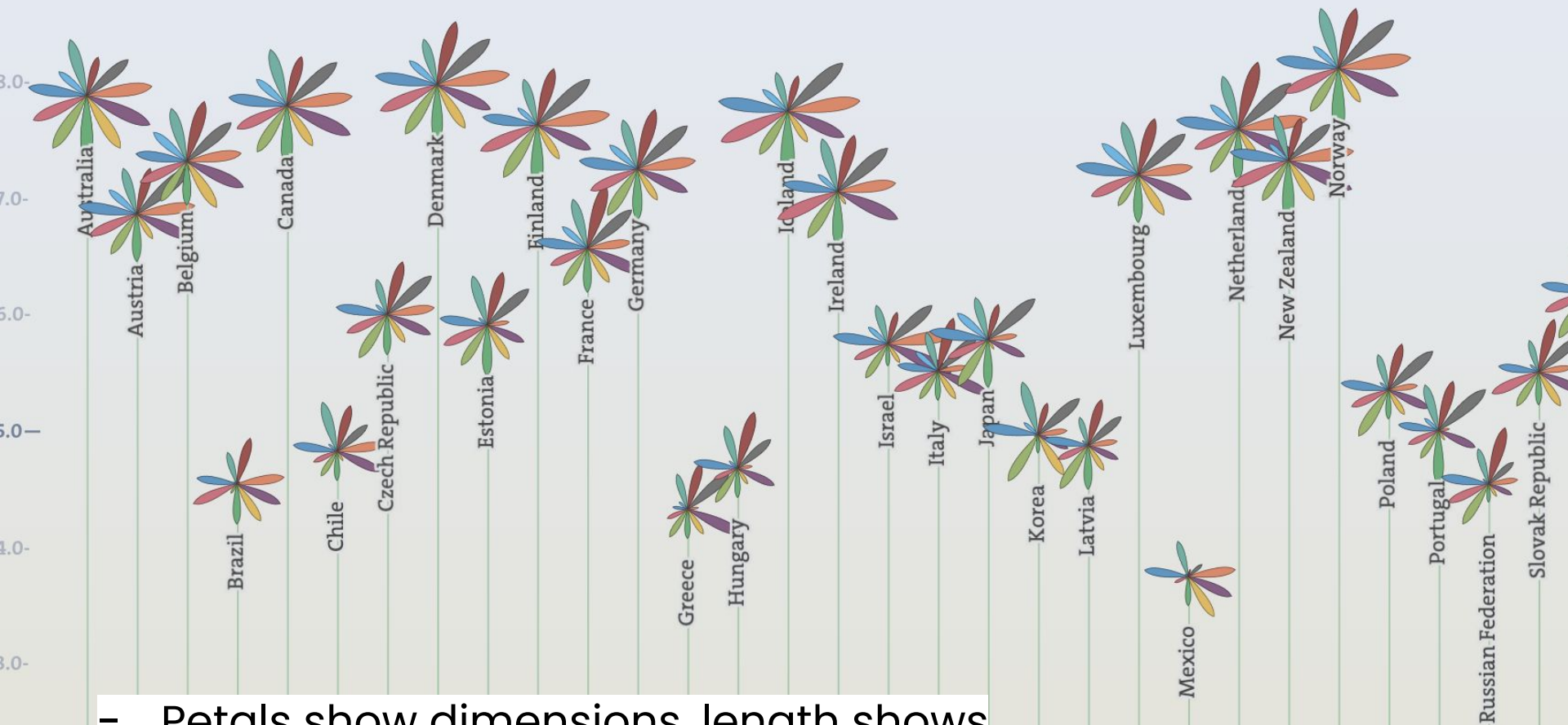
SEND TO:

STEFANIE POSAVEC

LONDON

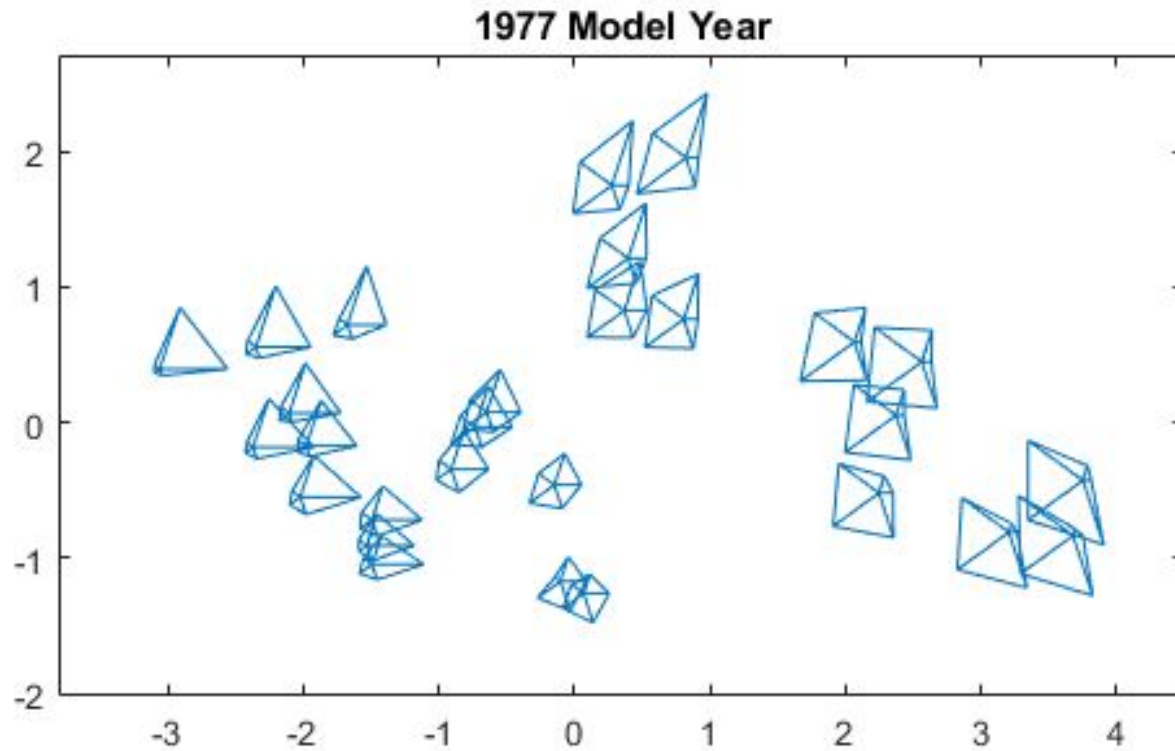
- UK -

England

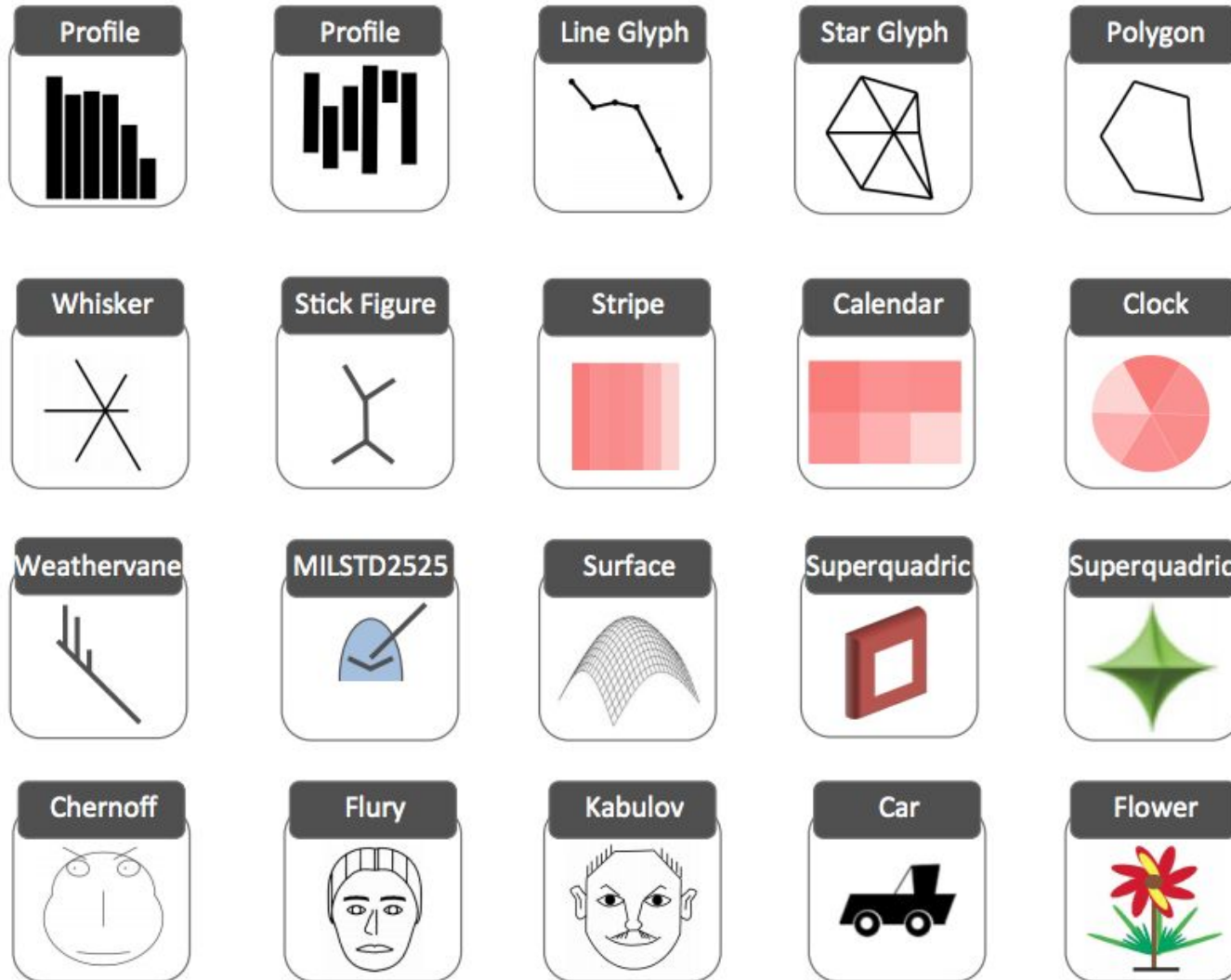


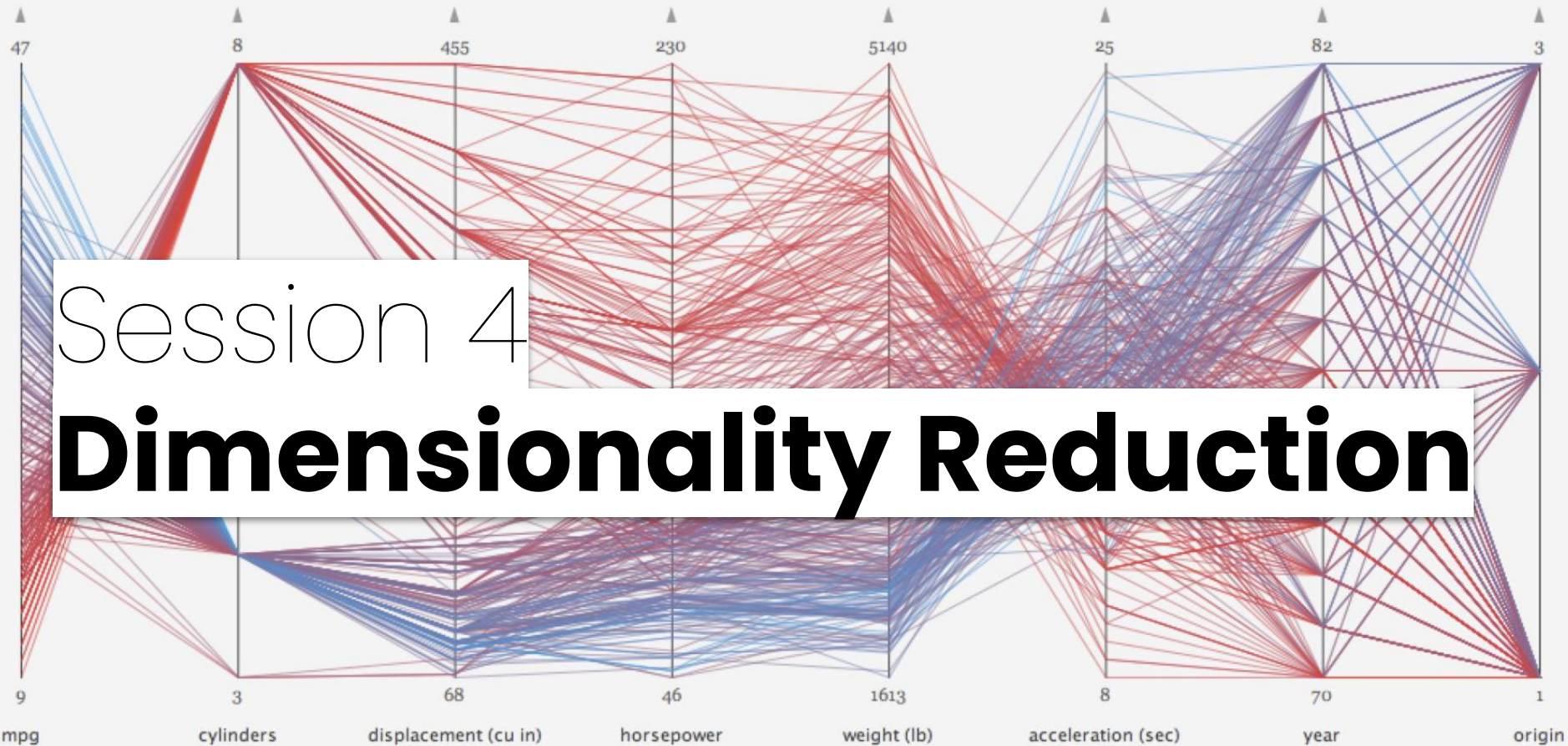
- Petals show dimensions, length shows attribute value
- Order country by values

Glyphs in Scatterplots



Glyph Design





Online Course
**Data Visualization
for Professionals**



THE UNIVERSITY
of EDINBURGH

Benjamin Bach

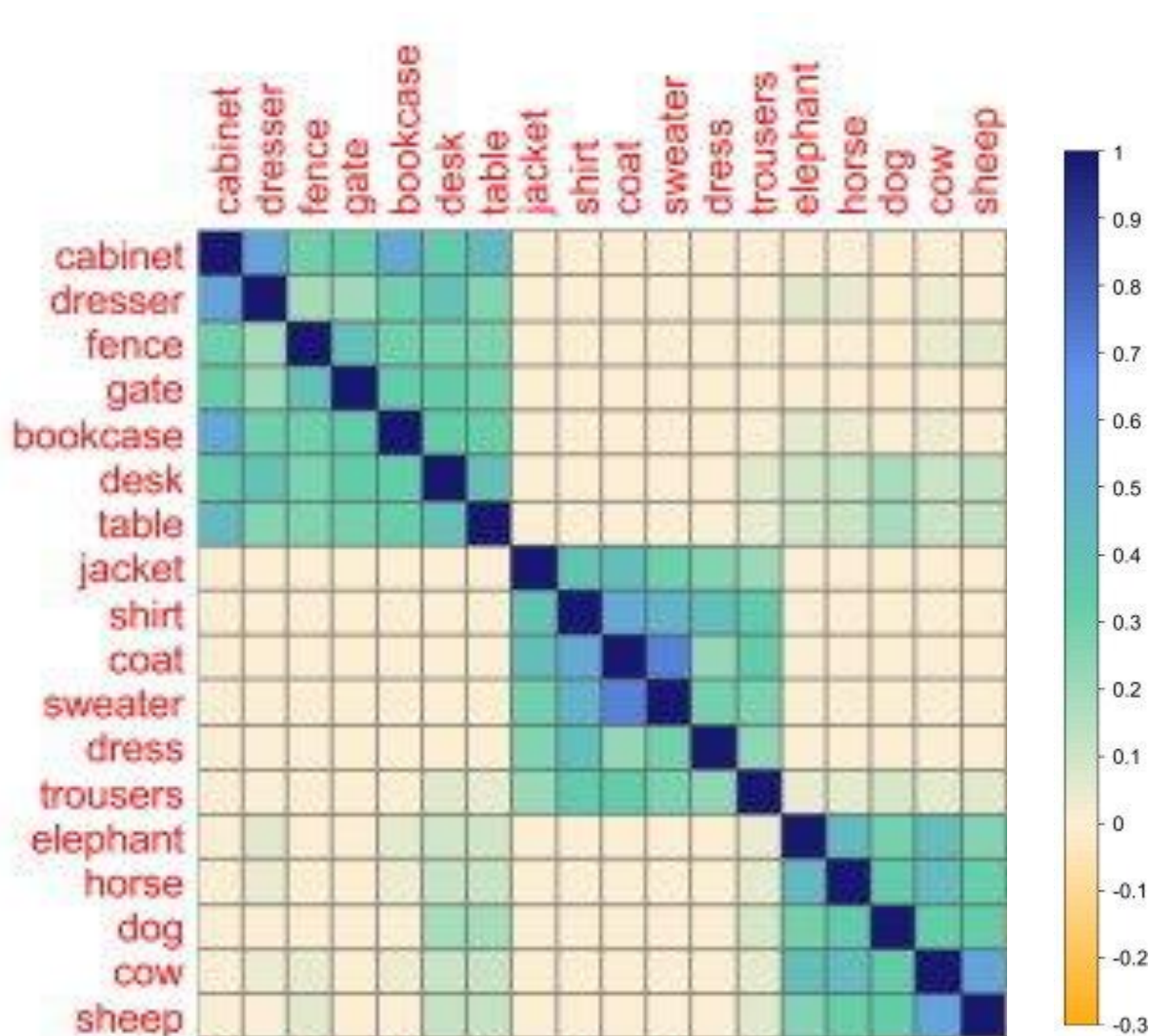
June 2022

<http://benjbach.me>

<https://datavis-online.github.io>

-- Not for external use --

5. Dimensionality Reduction: Similarity



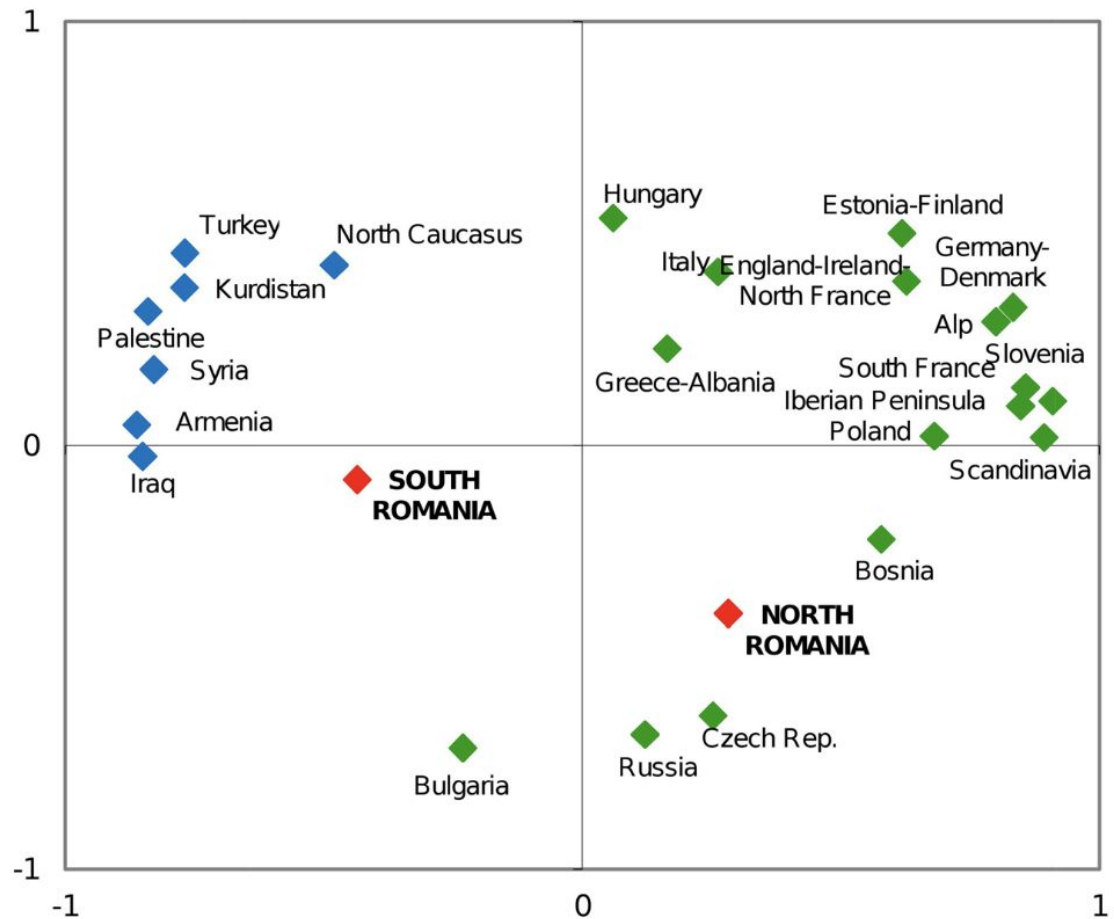
5. Dimensionality Reduction: MDS



Hervella Afonso, Montserrat, et al. "The Carpathian range represents a weak genetic barrier in South-East Europe." (2014).

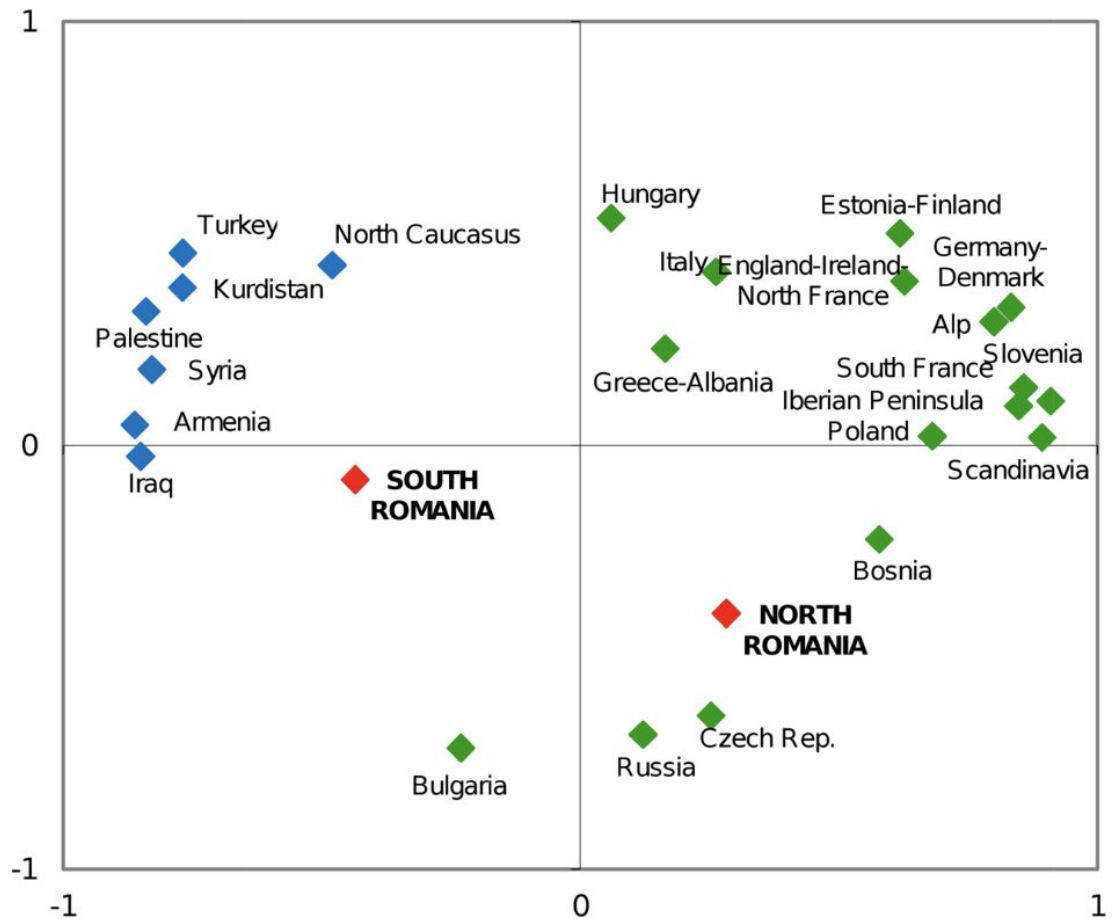
5. Dimensionality Reduction: MDS

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

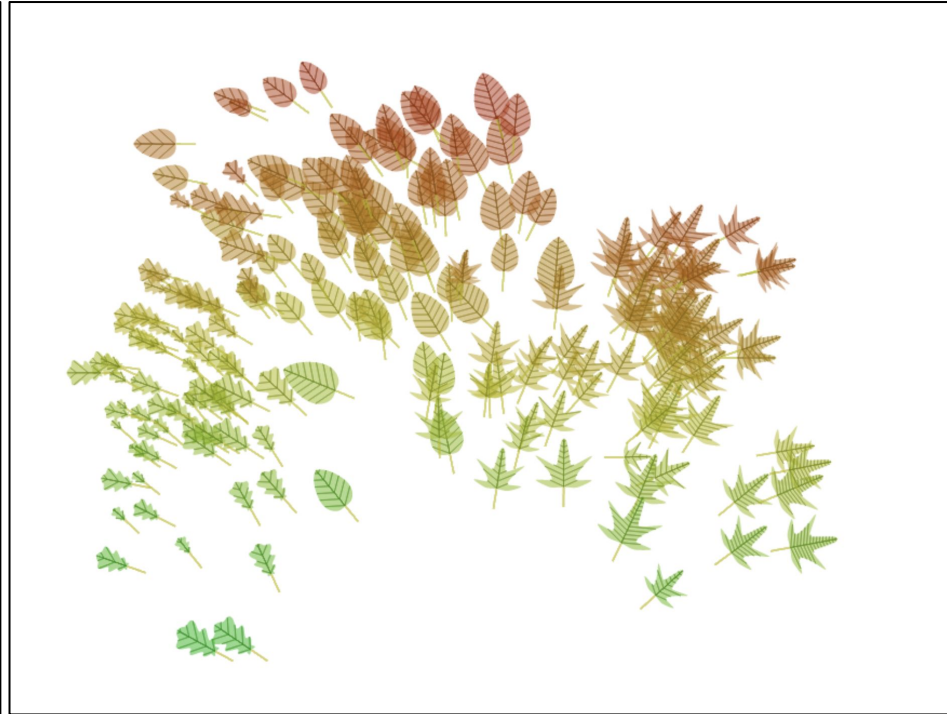
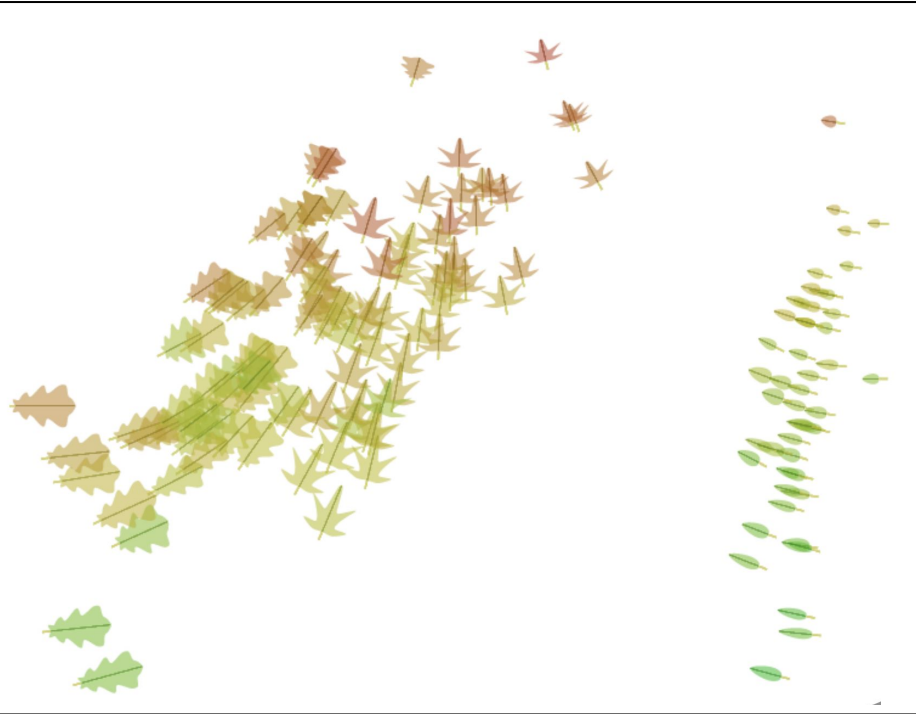


5. Dimensionality Reduction: MDS

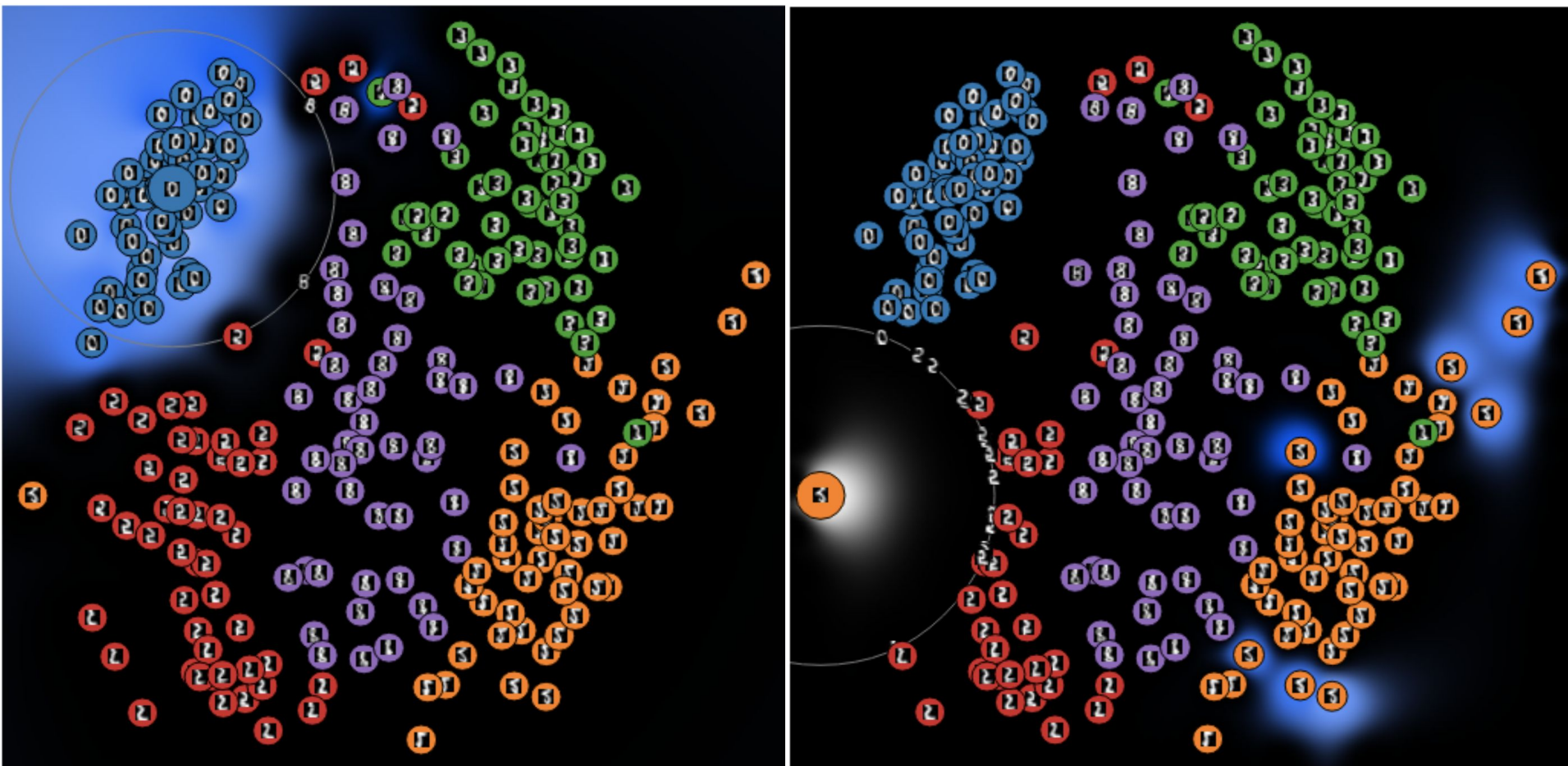
- + Dimension reduction
- + Can be 2D or 3D
- + Visual clustering
- Information loss
- Artifacts: **false neighbors** and **tears**



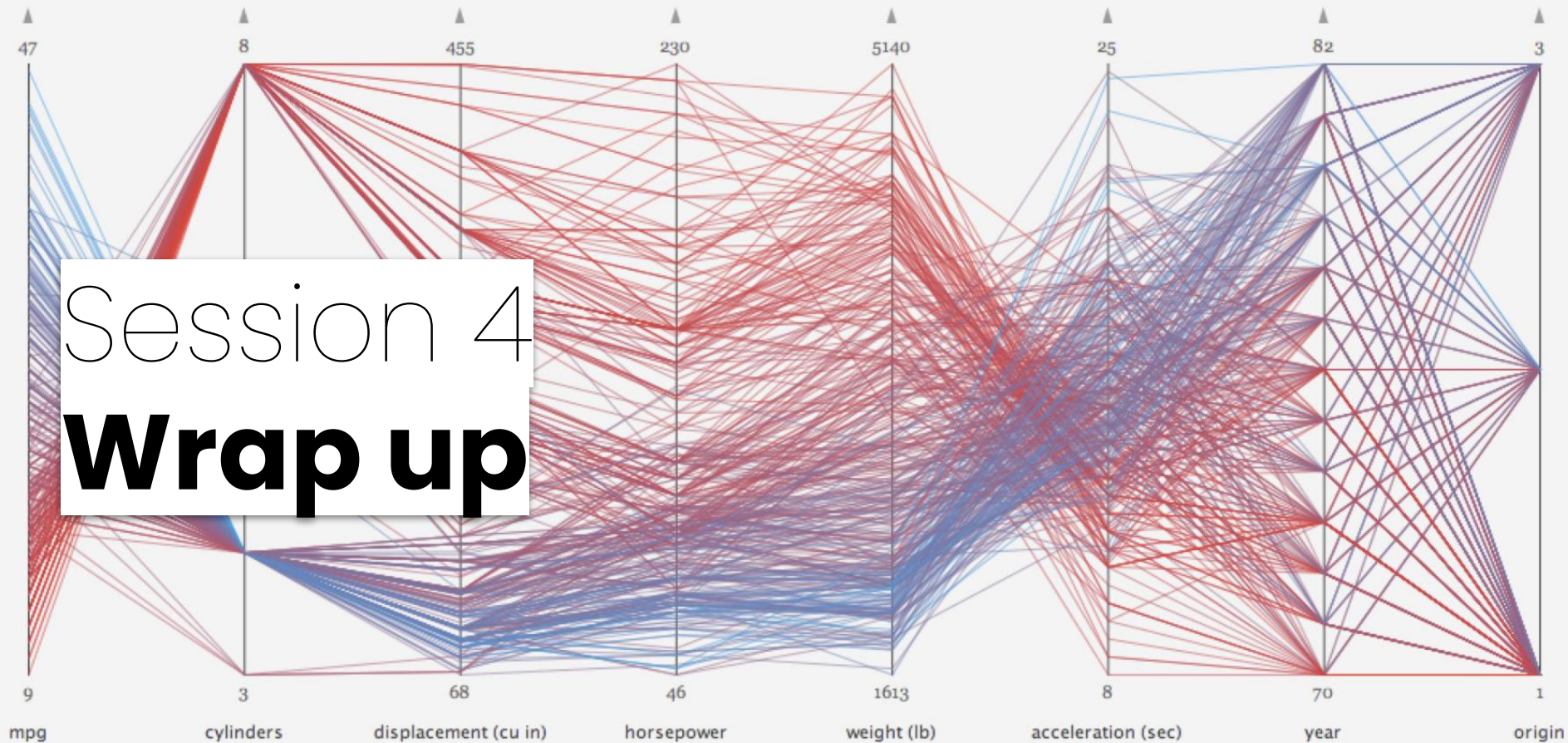
5. Dimensionality Reduction: Glyphs



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.



Online Course
**Data Visualization
for Professionals**



THE UNIVERSITY
of EDINBURGH

Benjamin Bach

June 2022

<http://benjbach.me>

<https://datavis-online.github.io>

-- Not for external use --

Outline

How to visualize data with dimensions (many attributes)?

- Low-dimensions
 - Scatterplot, Mekko chart, heatmap, Beeplots
- Higher dimensions (> 2)
 - Visual variables, scatterplot matrix, PCP, glyphs,
- Dimensionality reduction (many!)
 - + glyphs

Outline

How to visualize data with dimensions (many attributes)?

- Low-dimensions
 - Scatterplot, Mekko chart, heatmap, Beeplots
- Higher dimensions (> 2)
 - Visual variables, scatterplot matrix, PCP, glyphs,
- Dimensionality reduction (many!)
 - + glyphs

Outline

How to visualize data with dimensions (many attributes)?

- Low-dimensions
 - Scatterplot, Mekko chart, heatmap, Beeplots
- Higher dimensions (> 2)
 - Visual variables, scatterplot matrix, PCP, glyphs,
- Dimensionality reduction (many!)
 - + glyphs

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>