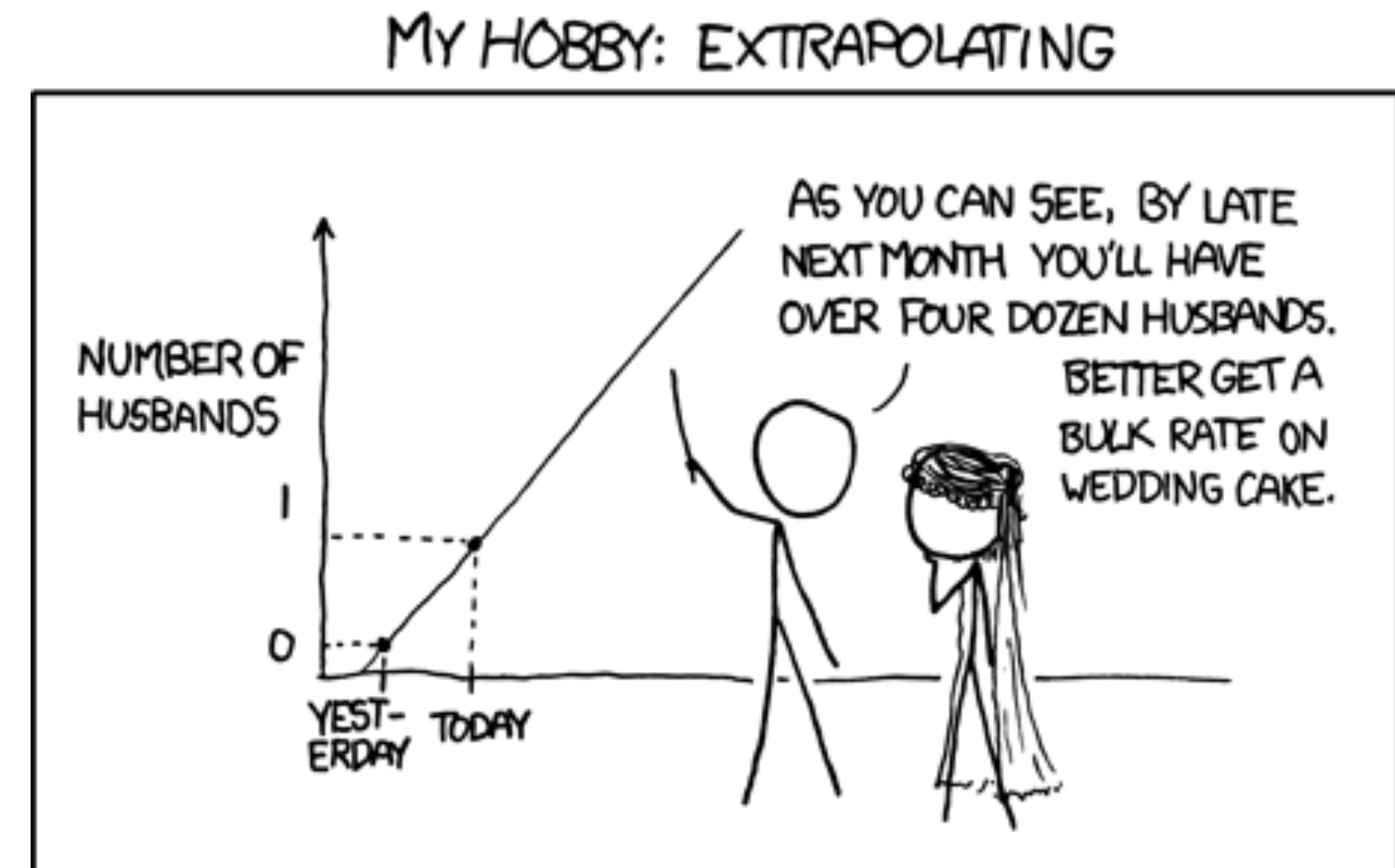


CS-5630 / CS-6630 Visualization

Tables

Alexander Lex
alex@sci.utah.edu

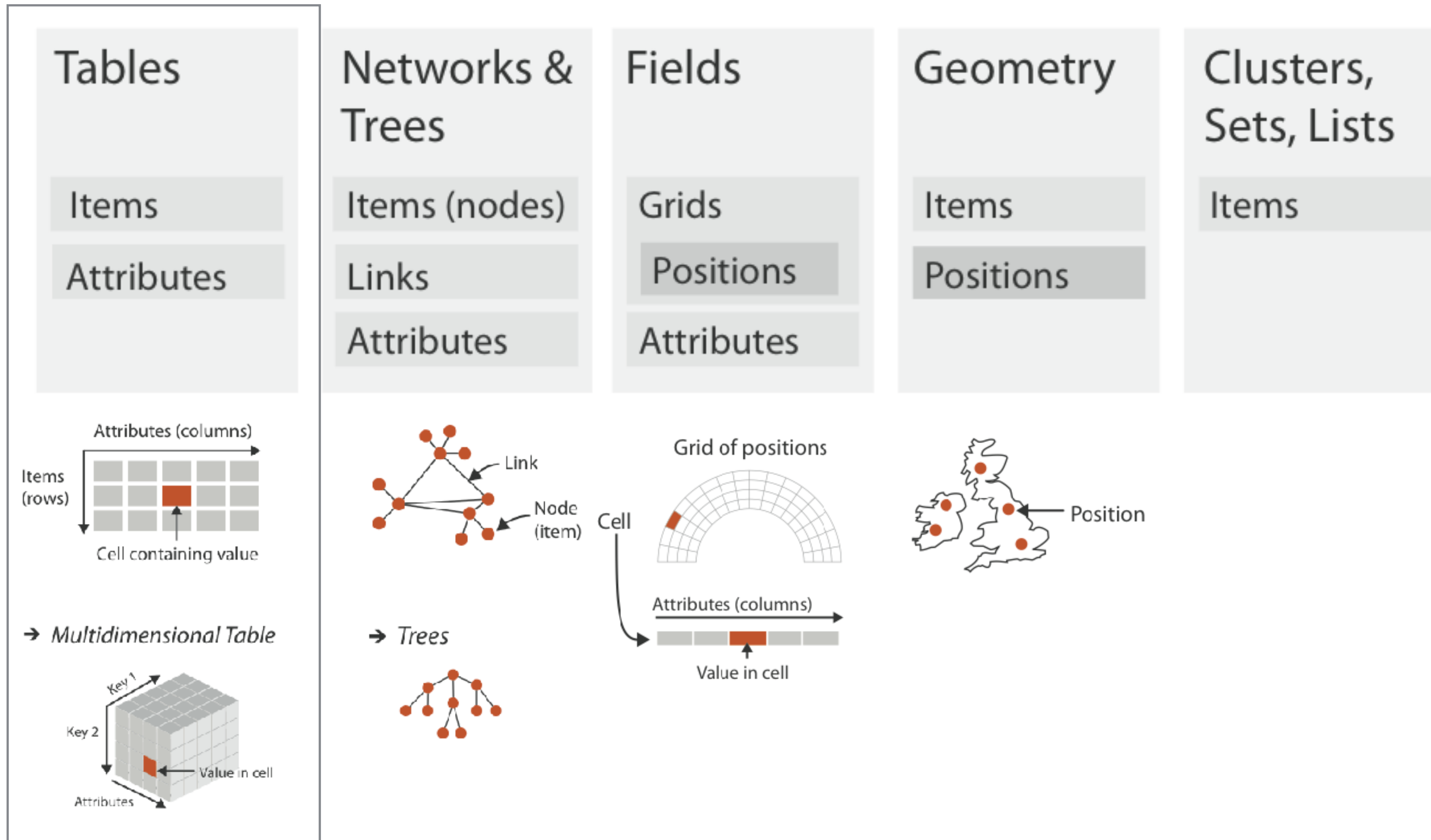


Organizational

Contacted by TA this week for feedback on project

No more standing office hours - arrange meetings

dataset types



Arrange Tables

① Express Values



② Separate, Order, Align Regions

→ Separate



→ Order



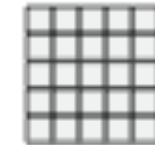
→ Align



→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume



→ Many Keys
Recursive Subdivision



③ Axis Orientation

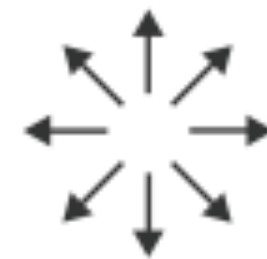
→ Rectilinear



→ Parallel

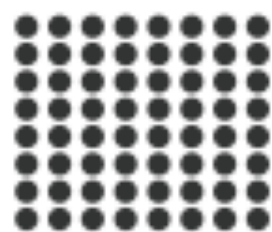


→ Radial



④ Layout Density

→ Dense

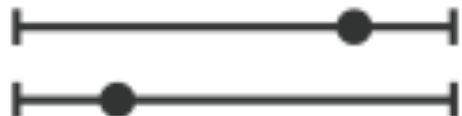


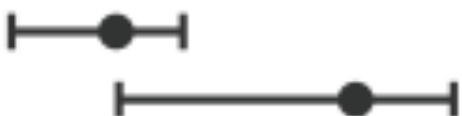
→ Space-Filling




spatial channels are the most effective for all attribute types


➔ **Magnitude Channels: Ordered Attributes**

Position on common scale 

Position on unaligned scale 

Length (1D size) 


➔ **Identity Channels: Categorical Attributes**

Spatial region 

Tilt/angle 

Area (2D size) 

Depth (3D position) 

Color luminance 

Color saturation 

Curvature 

Most ▲

Effectiveness

Same

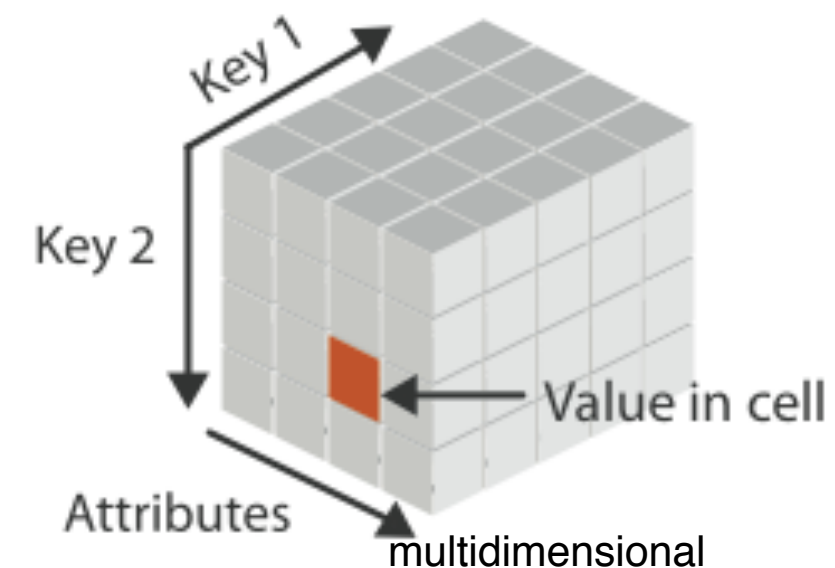
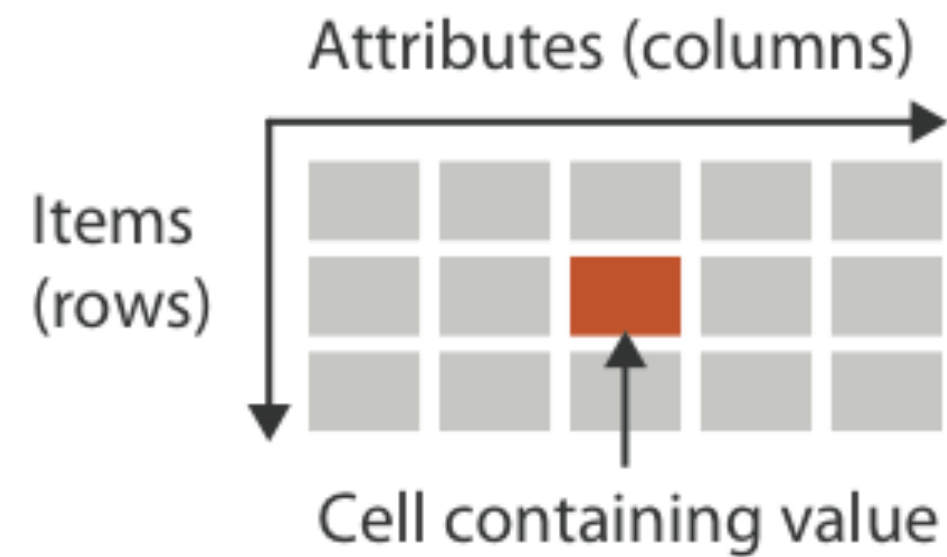
Color hue 

Motion 

Shape 

recall: attribute semantics

when we arrange tabular data, attributes are chosen to be keys and values



Scale of Tables

Need different approaches for “normal” and “high-dimensional” tables.

How many dimensions?

~50 – tractable with “just” vis

~1000 – need analytical methods

How many records?

~ 1000 – “just” vis is fine

>> 10,000 – need analytical methods

Homogeneity

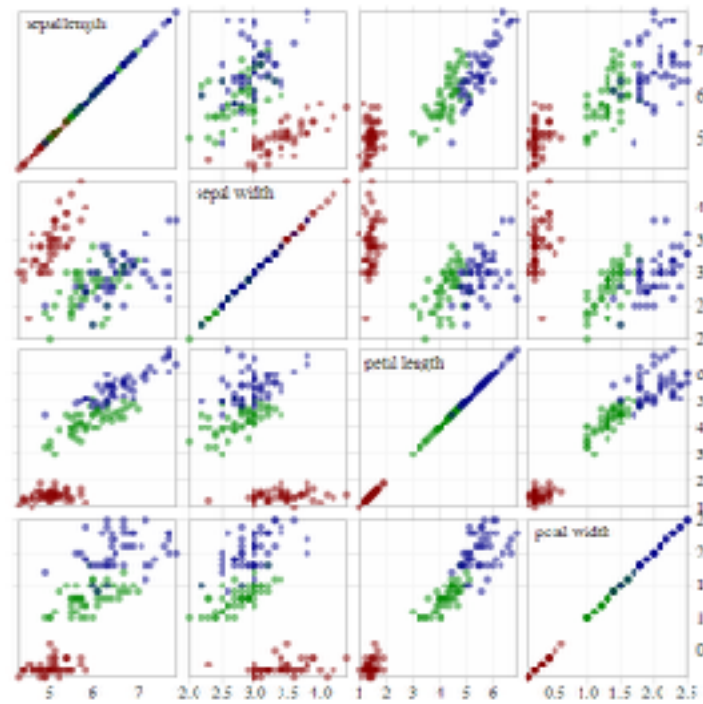
Same data type?

Same scales?

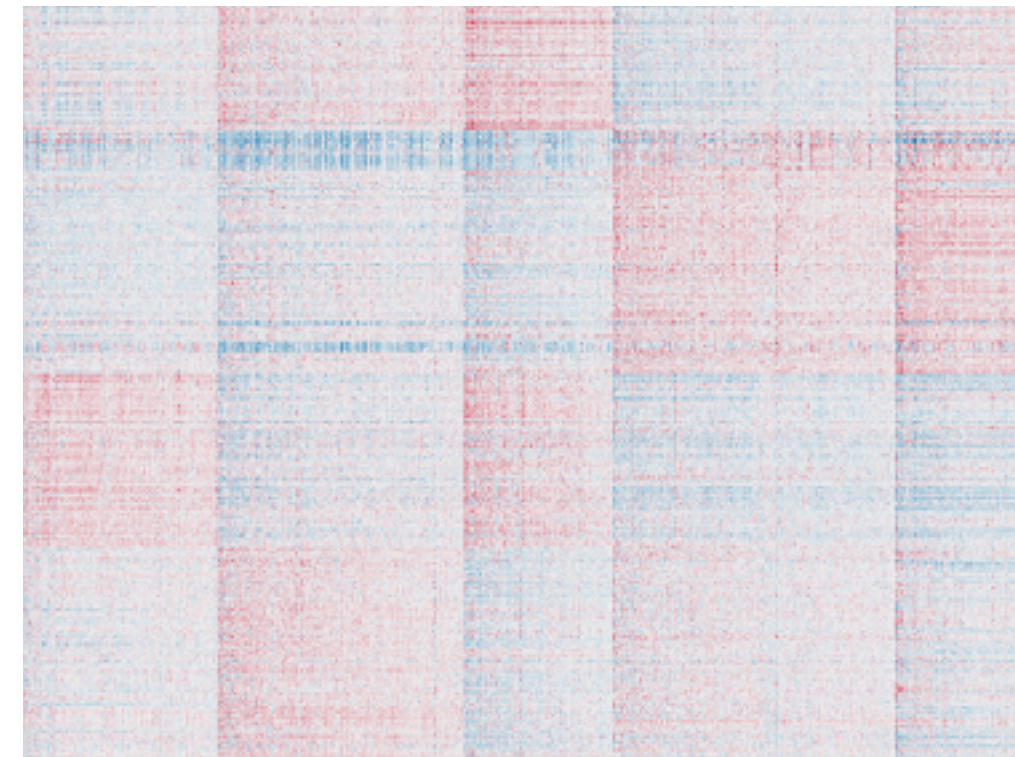
	Age	Gender	Height
Bob	25	M	181
Alice	22	F	185
Chris	19	M	175

	BPM 1	BPM 2	BPM 3
Bob	65	120	145
Alice	80	135	185
Chris	45	115	135

Analytic Component



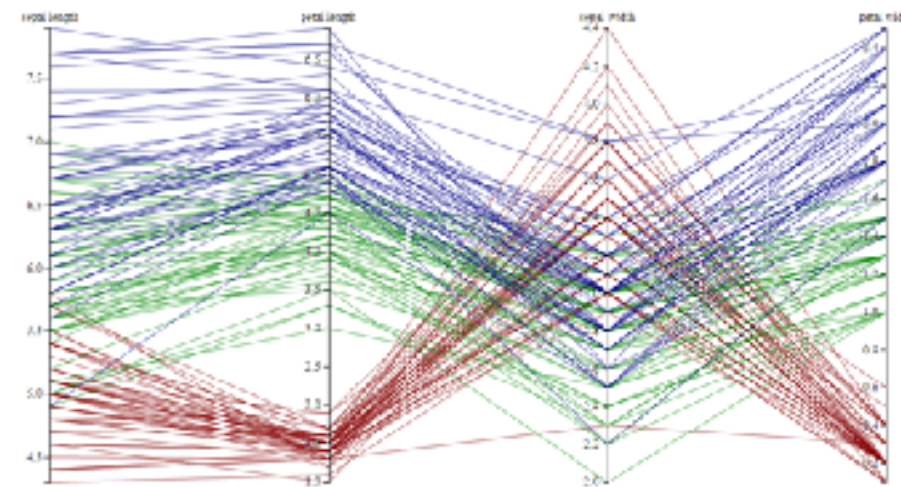
Scatterplot Matrices
[Bostock]



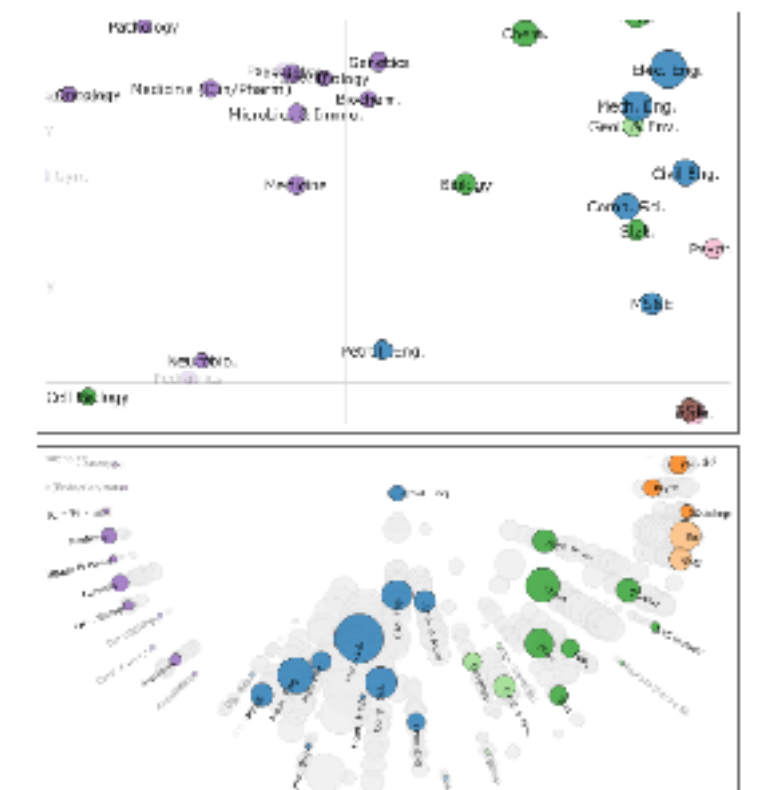
Pixel-based visualizations /
heat maps



Multidimensional Scaling
[Doerk 2011]



Parallel Coordinates
[Bostock]



[Chuang 2012]

no / little analytics

strong analytics
component

Express Values

No Keys

encode using zero keys: scatterplots

Arrange Tables

① Express Values



② Separate, Order, Align Regions

→ Separate



→ Order



→ Align



③ Axis Orientation

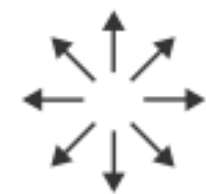
→ Rectilinear



→ Parallel

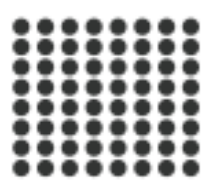


→ Radial



④ Layout Density

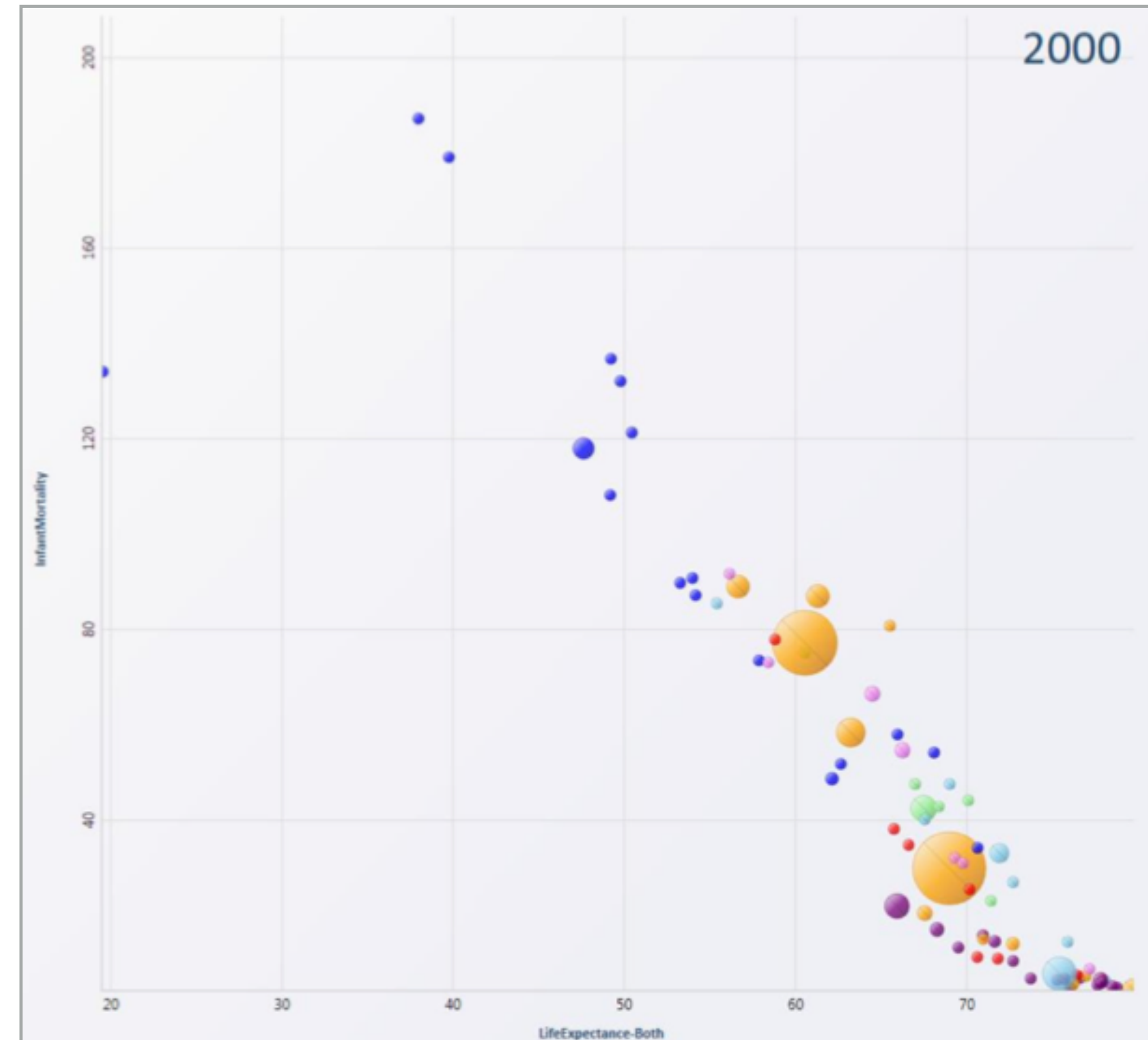
→ Dense



→ Space-Filling



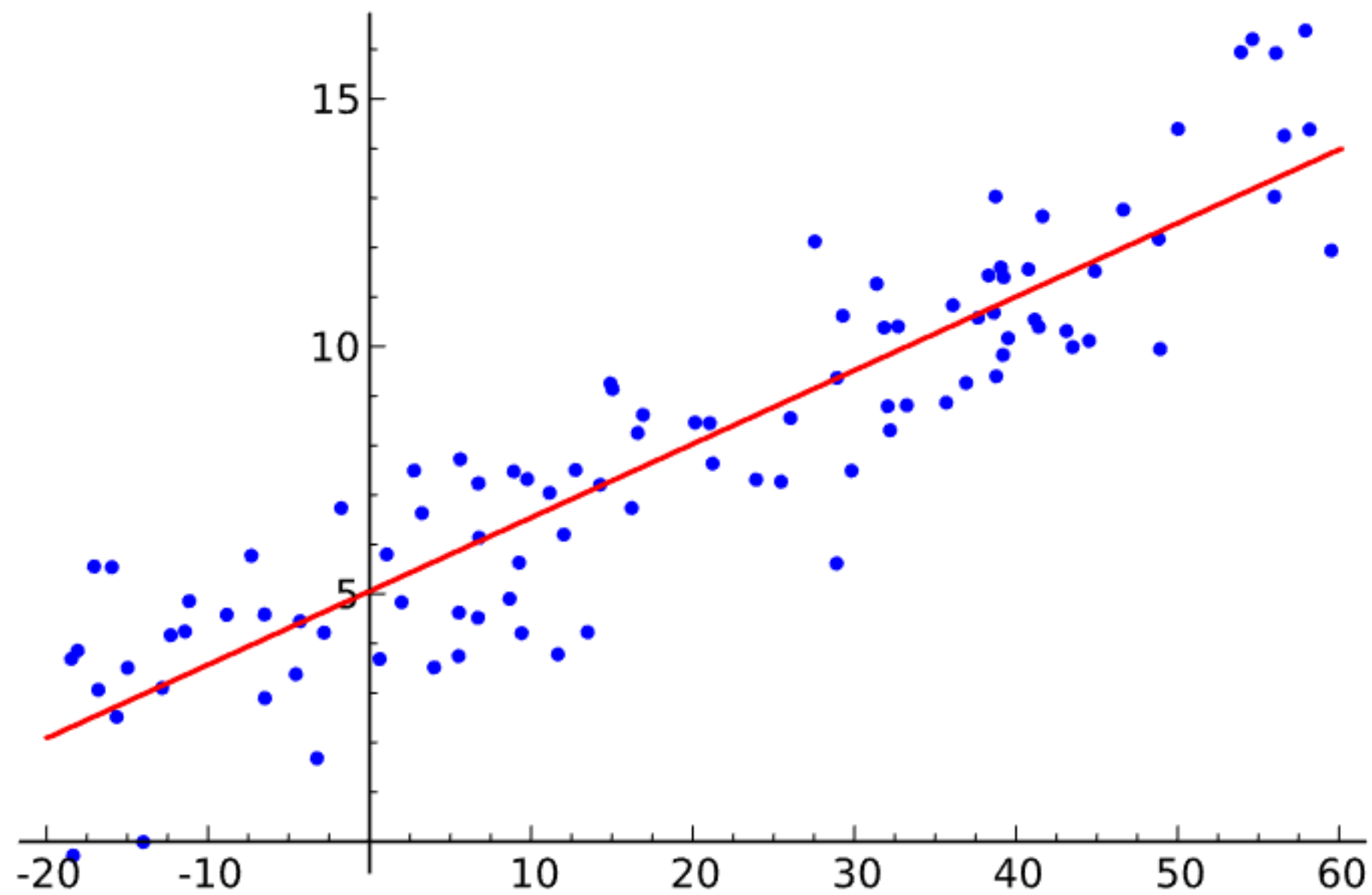
Infant Mortality



Life Expectance

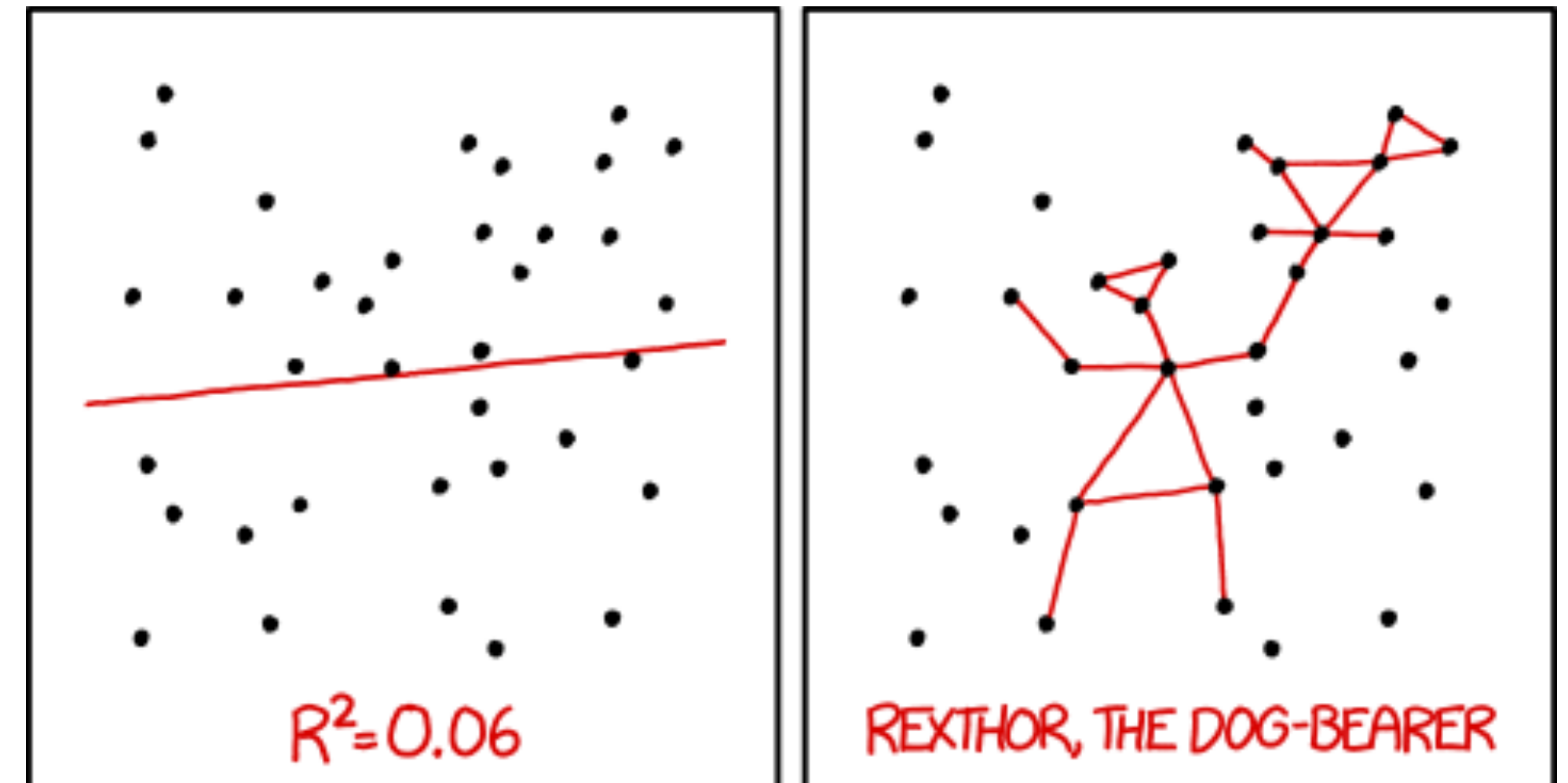
Regression Lines

$$y \sim \beta_0 + \beta_1 x$$



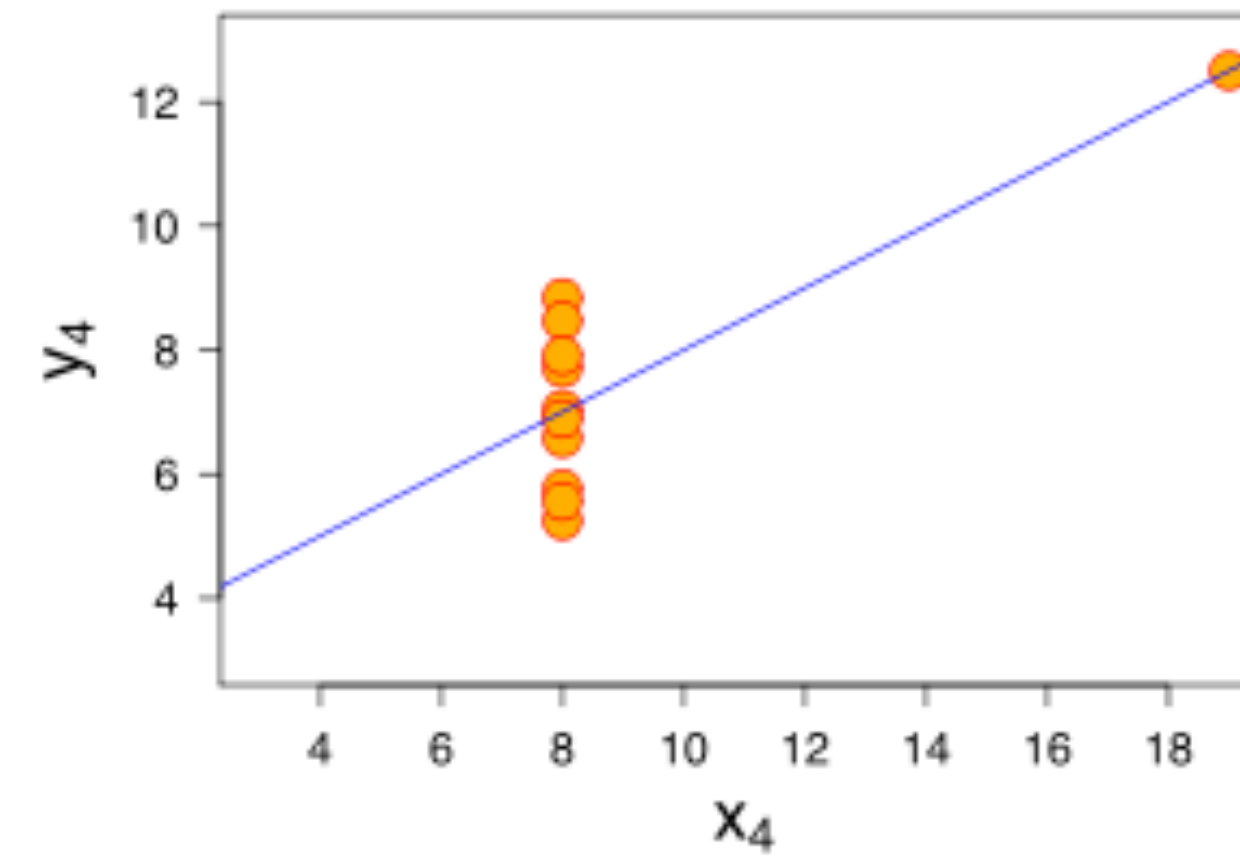
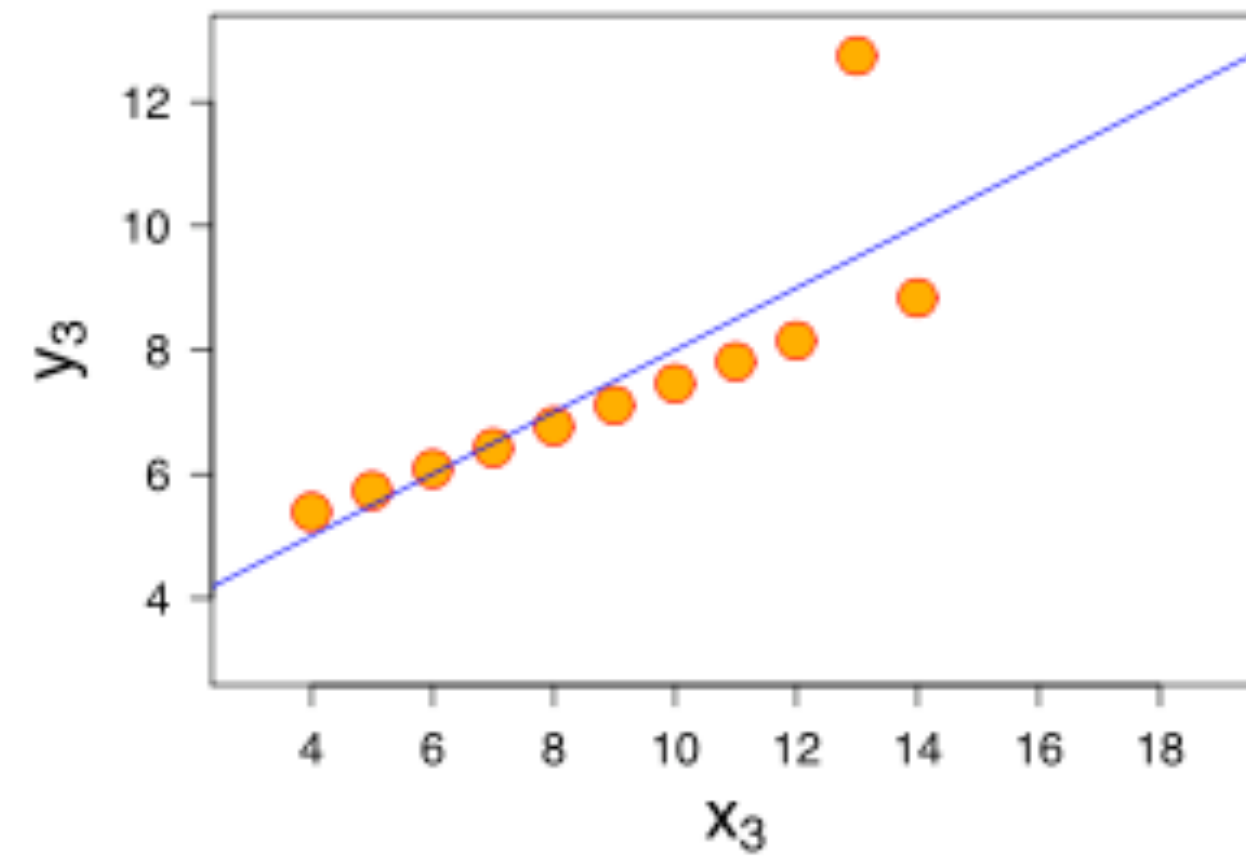
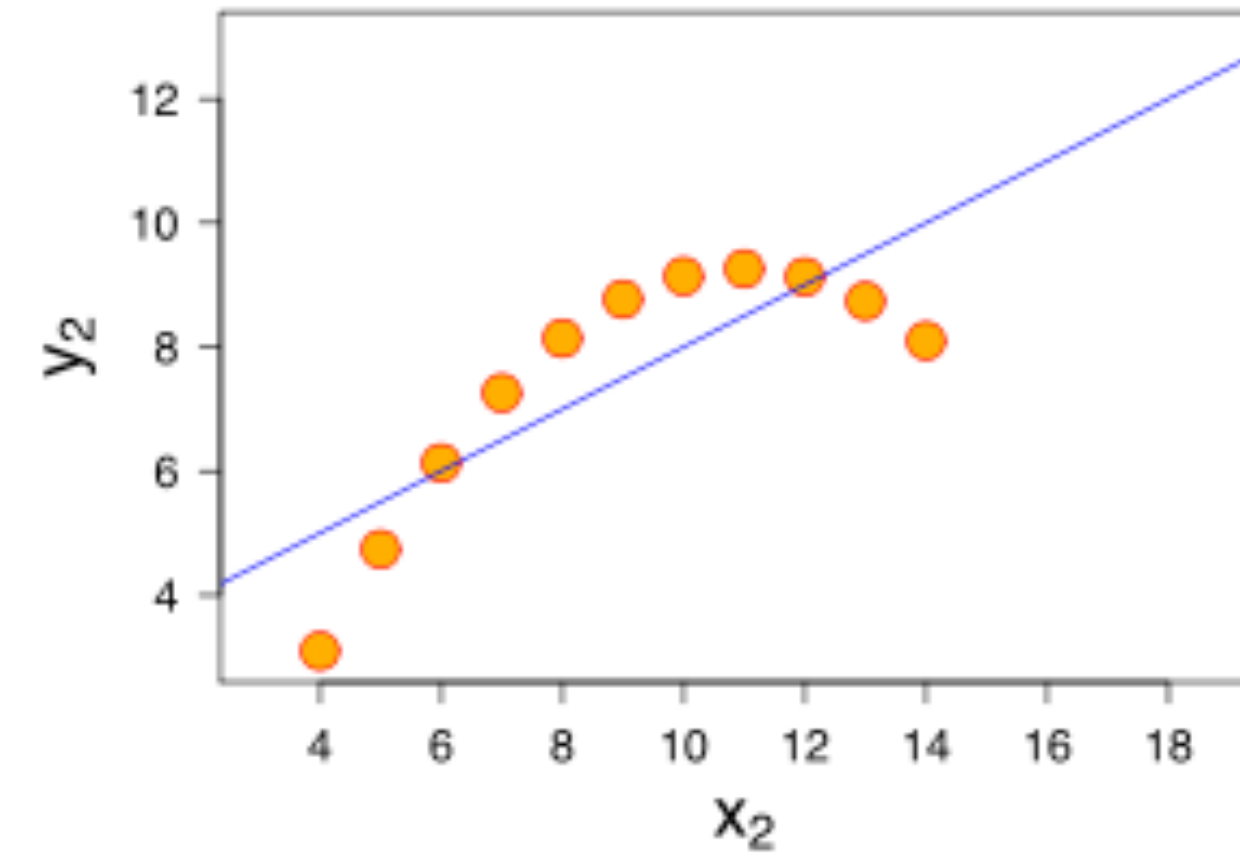
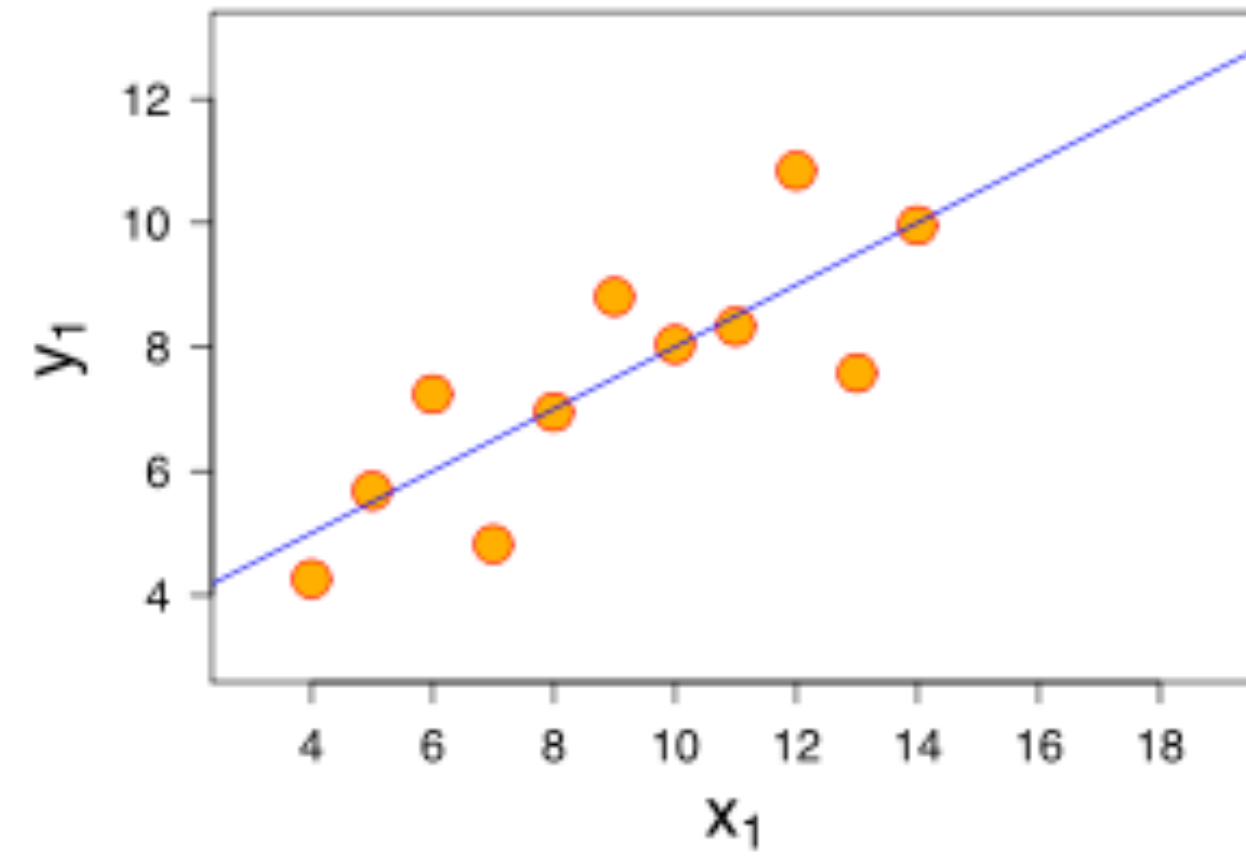
Goal: Find the best values of β_0 and β_1 , denoted $\hat{\beta}_0$ and $\hat{\beta}_1$, so that the prediction $y = \hat{\beta}_0 + \hat{\beta}_1 x$ “best fits” the data.

Approach: use least squares to minimize the sum of the squares of the errors



I DON'T TRUST LINEAR REGRESSIONS WHEN IT'S HARDER TO GUESS THE DIRECTION OF THE CORRELATION FROM THE SCATTER PLOT THAN TO FIND NEW CONSTELLATIONS ON IT.

Anscombe's Quartet



Encode one Key
Attribute

encode one key attribute: bar, dot, & line charts

Arrange Tables

Express Values



Separate, Order, Align Regions

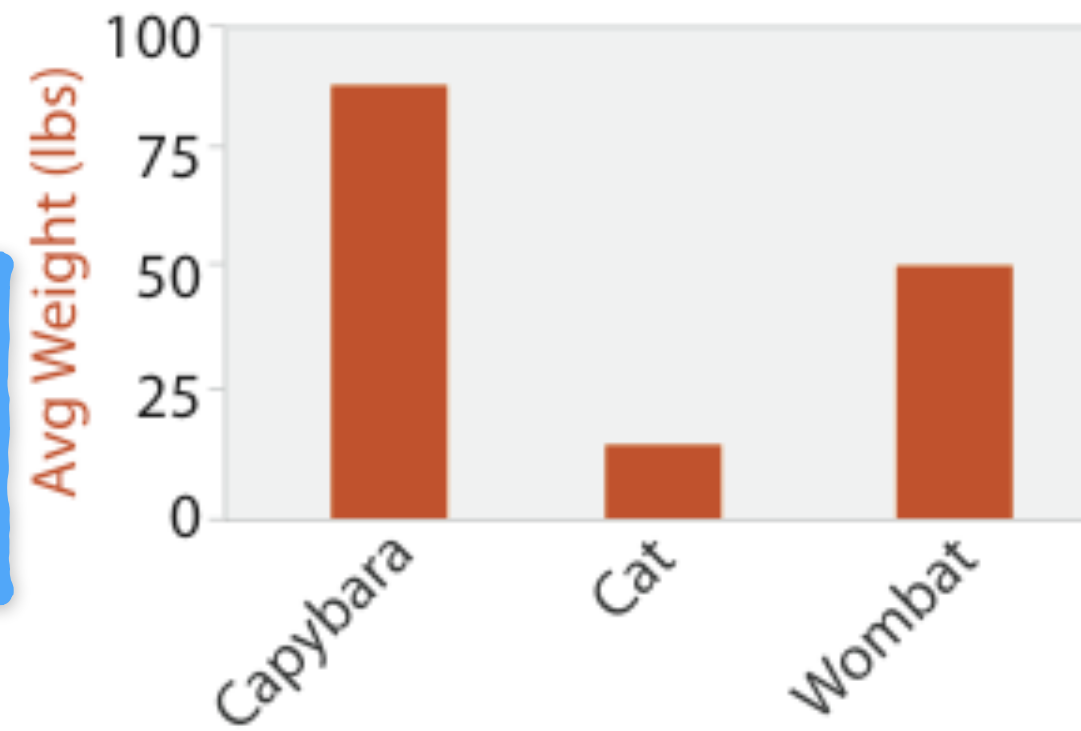
Separate



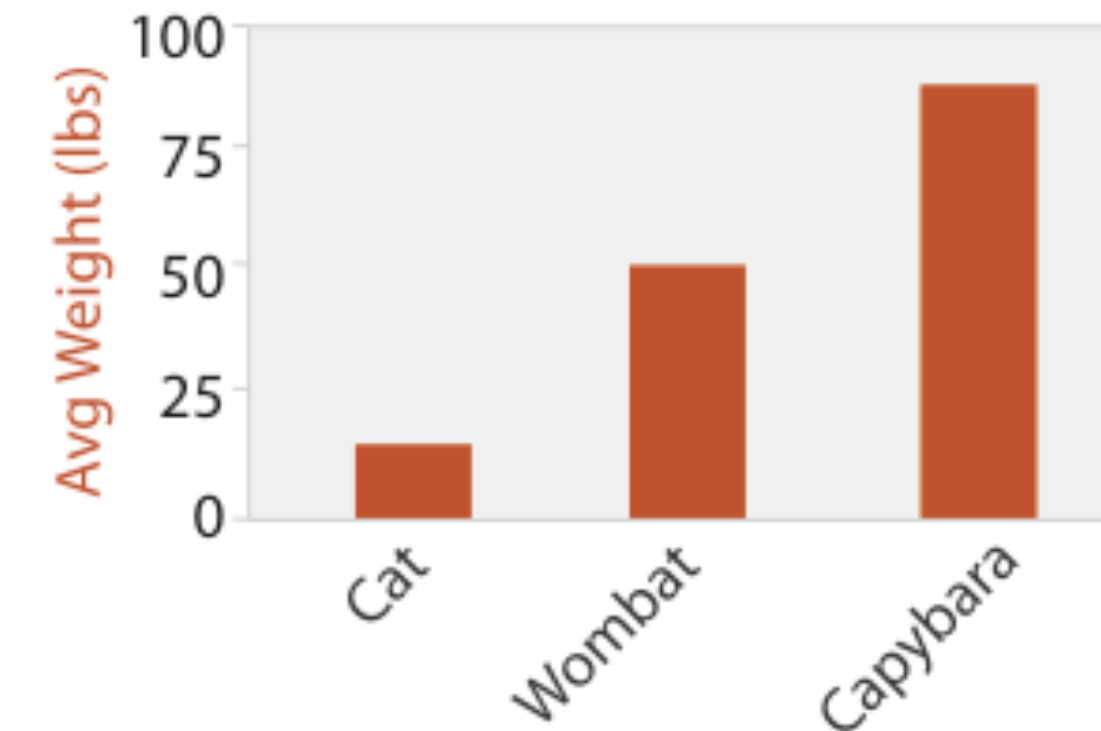
Order



Align



Animal Type



Animal Type

Axis Orientation

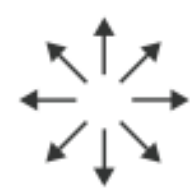
Rectilinear



Parallel



Radial

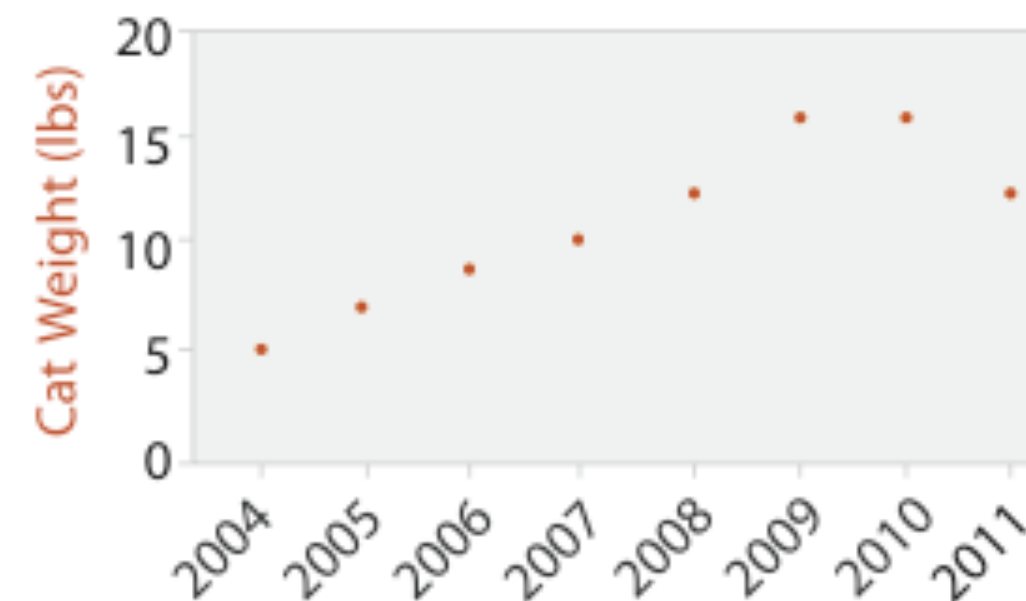


Layout Density

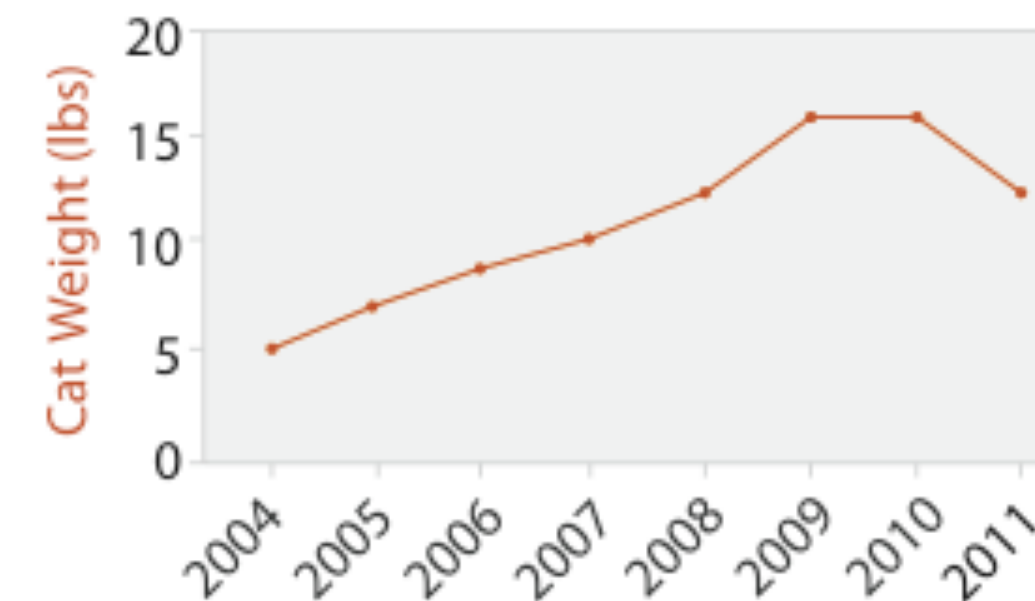
Dense



Space-Filling



Year



Year

Encode Multiple Key Attributes

Arrange Tables

① Express Values



② Separate, Order, Align Regions

→ Separate



→ Order



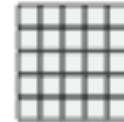
→ Align



→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume



→ Many Keys
Recursive Subdivision



③ Axis Orientation

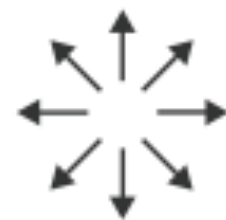
→ Rectilinear



→ Parallel

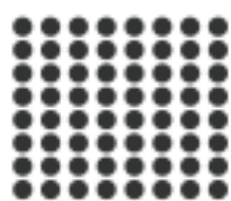


→ Radial



④ Layout Density

→ Dense



→ Space-Filling



Stacked Bar Chart

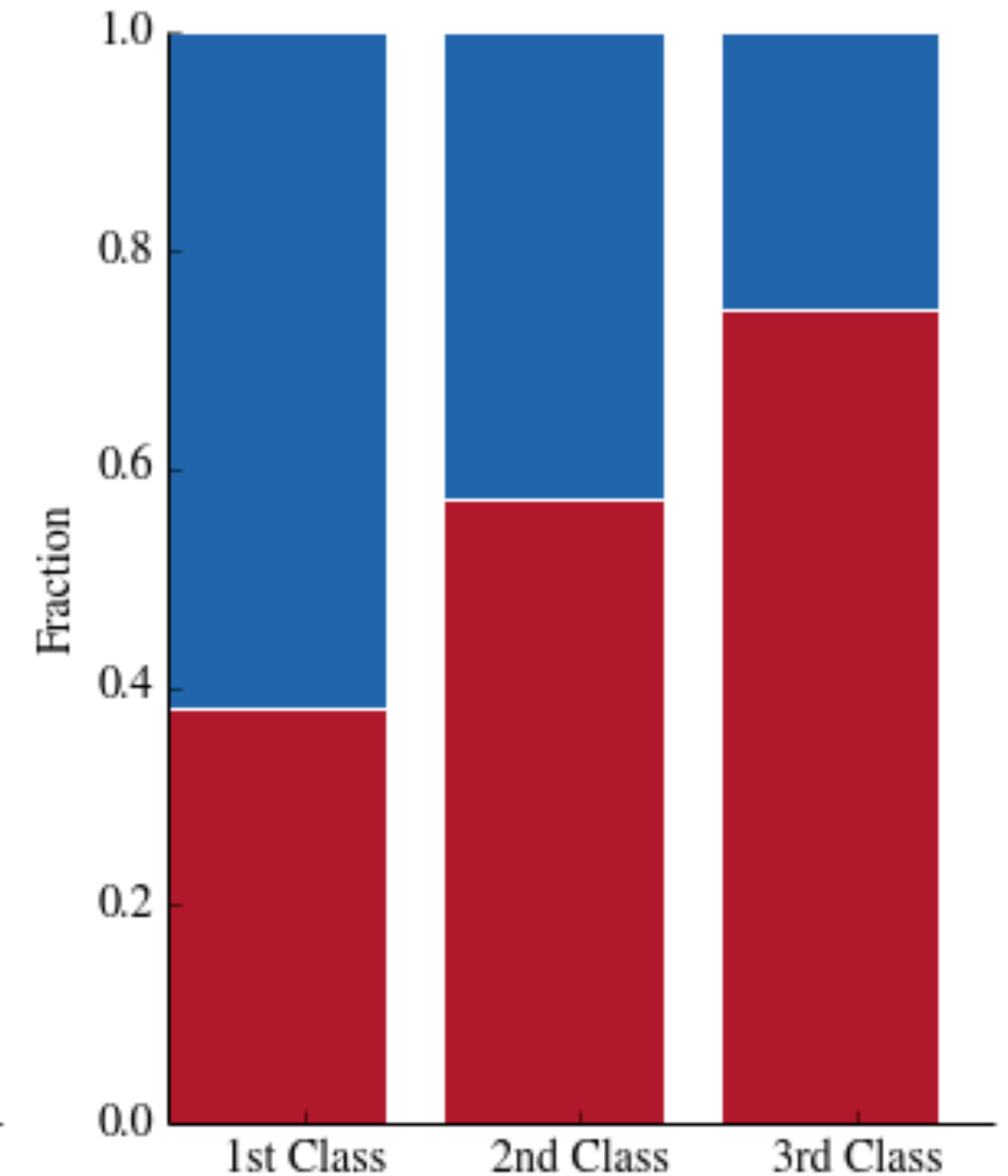
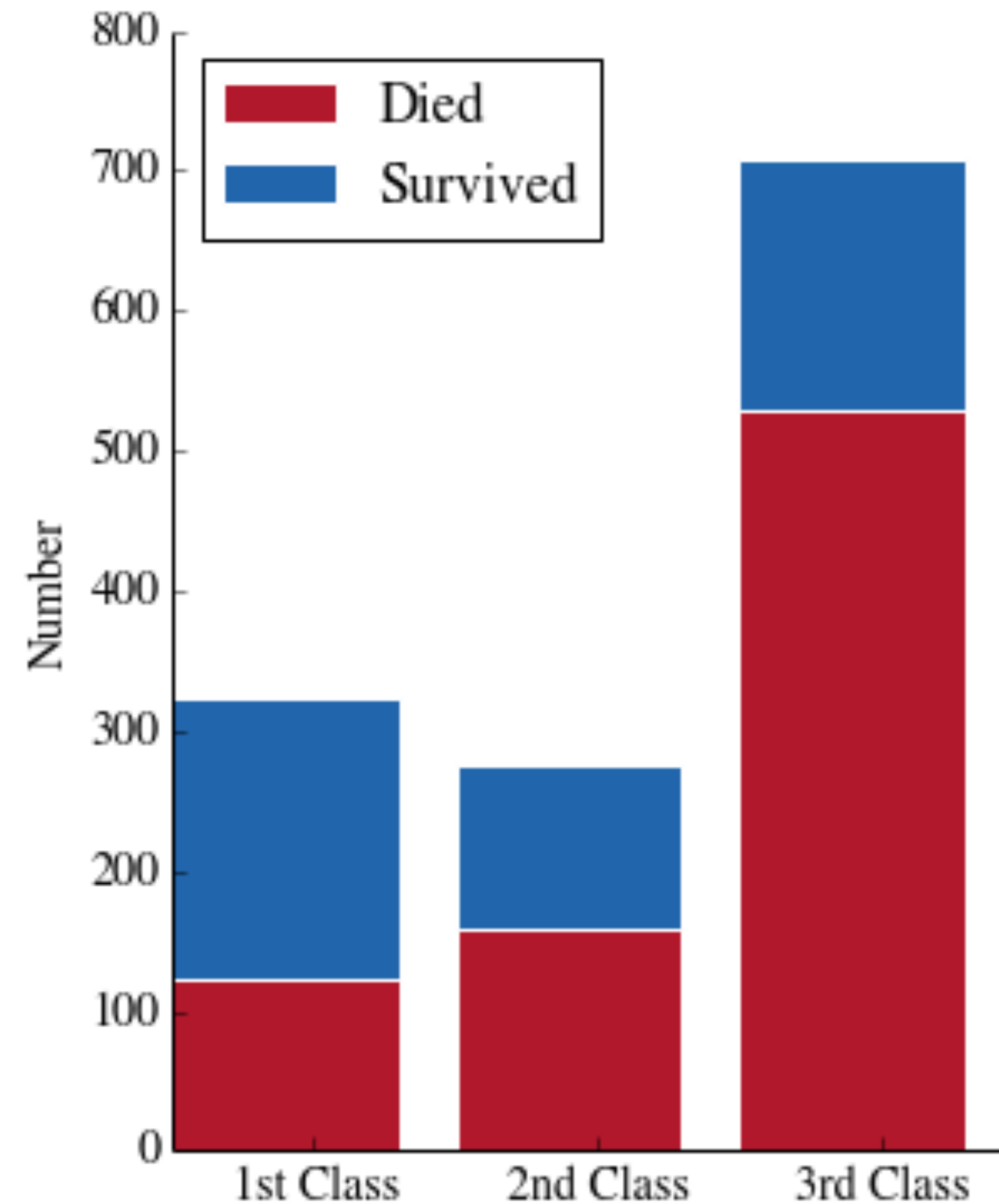
Keys: Class, Survival

Class is spatial

Survival is color

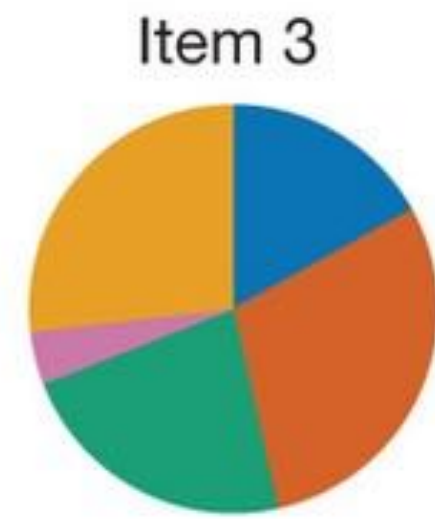
Left: absolute values

Right: proportional values

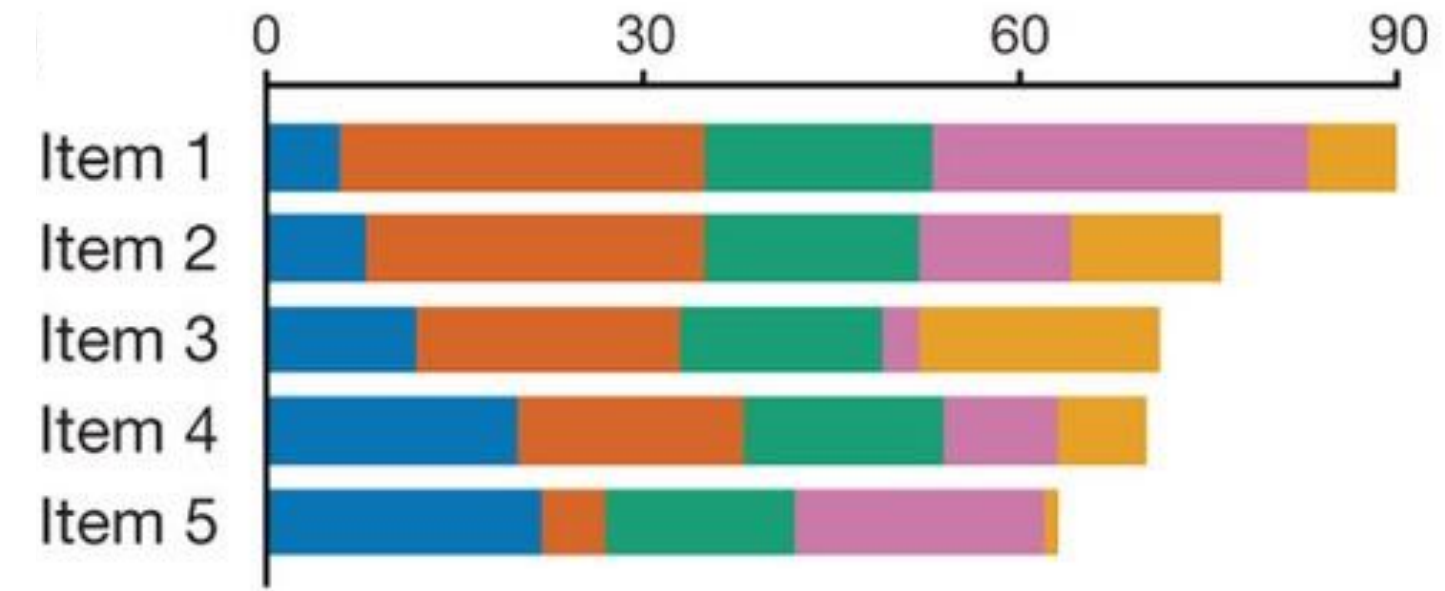


Comparison of bar chart types

- Category 1 ●
- Category 2 ●
- Category 3 ●
- Category 4 ●
- Category 5 ●

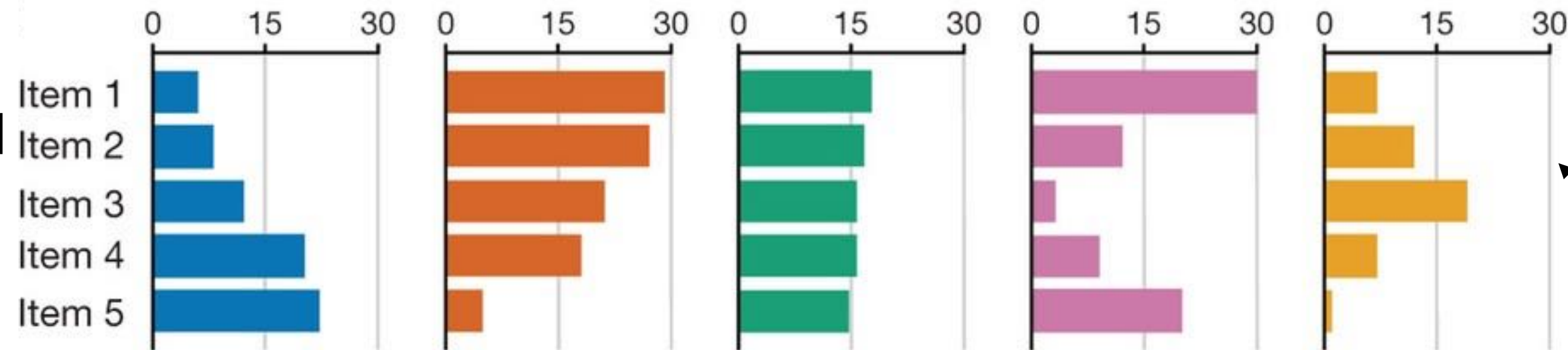


Pie Chart

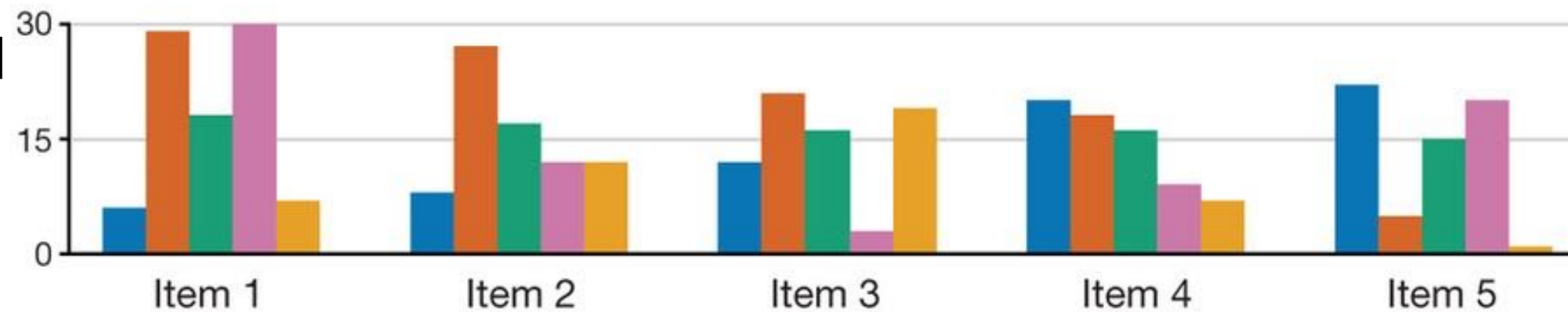


Stacked bar chart

Layered Bar Chart

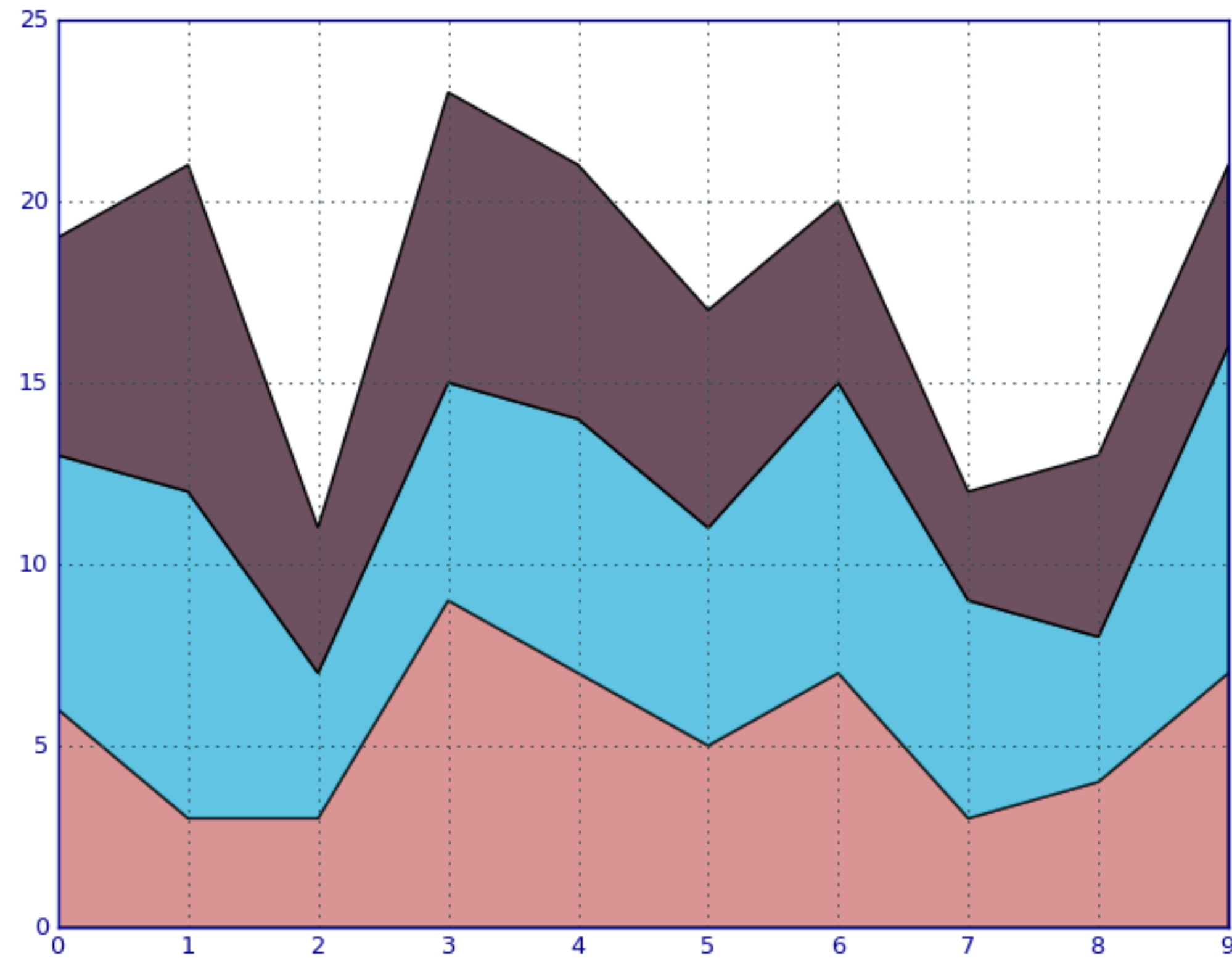


Grouped Bar Chart



Small Multiples

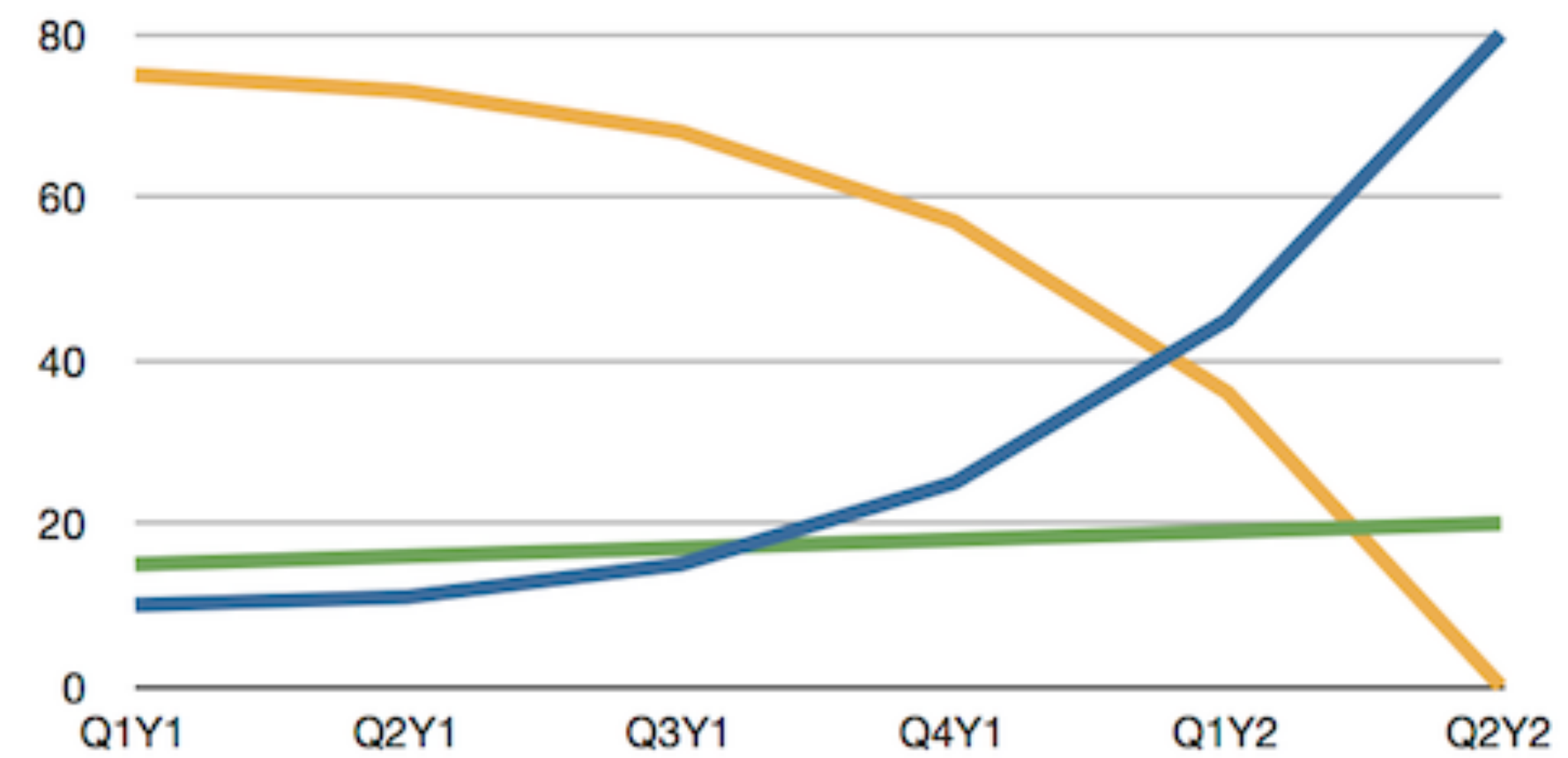
Stacked Area Chart



100% Stacked Area Chart



Stacked Area vs. Line Graphs



Can you spot the trends?

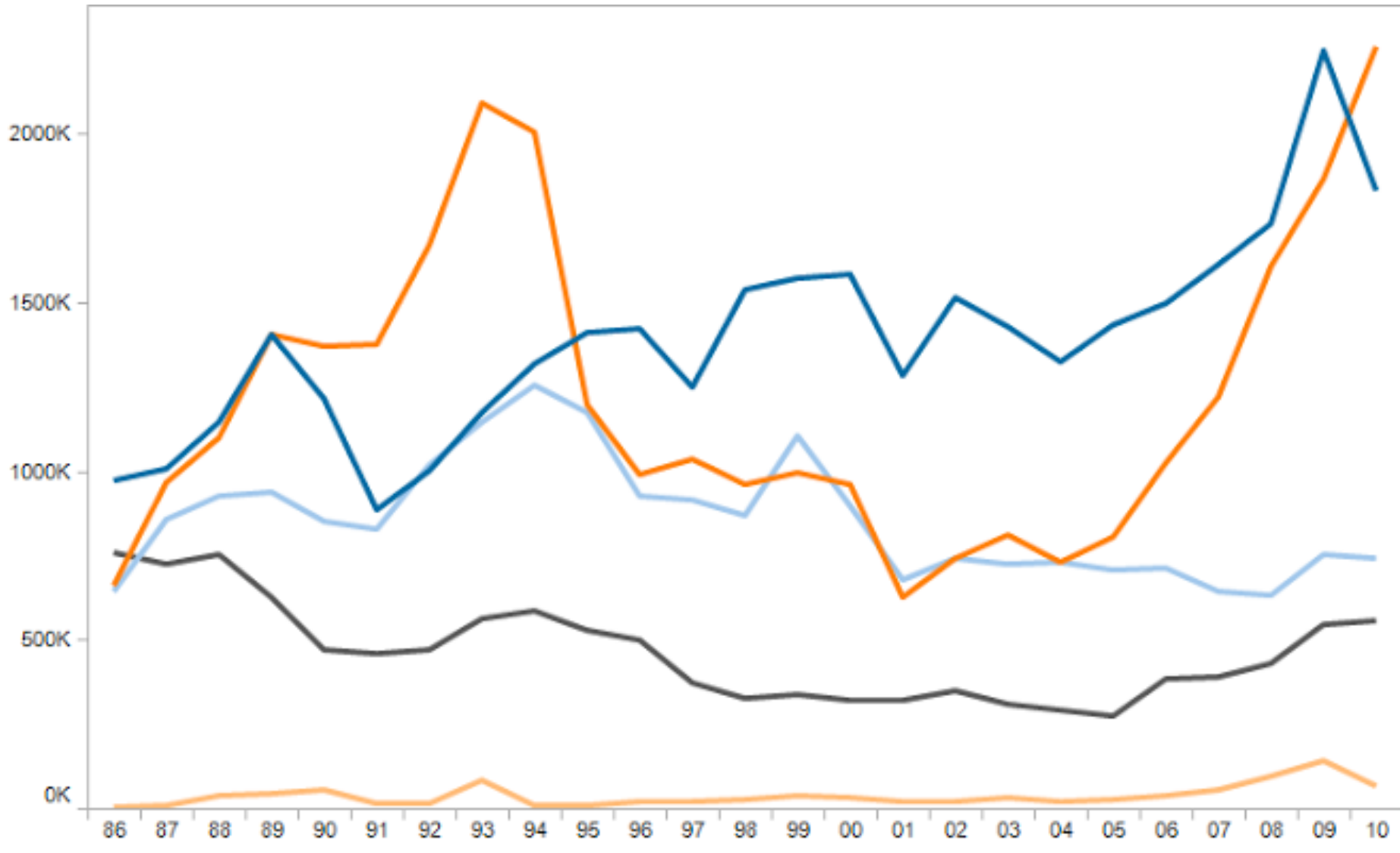
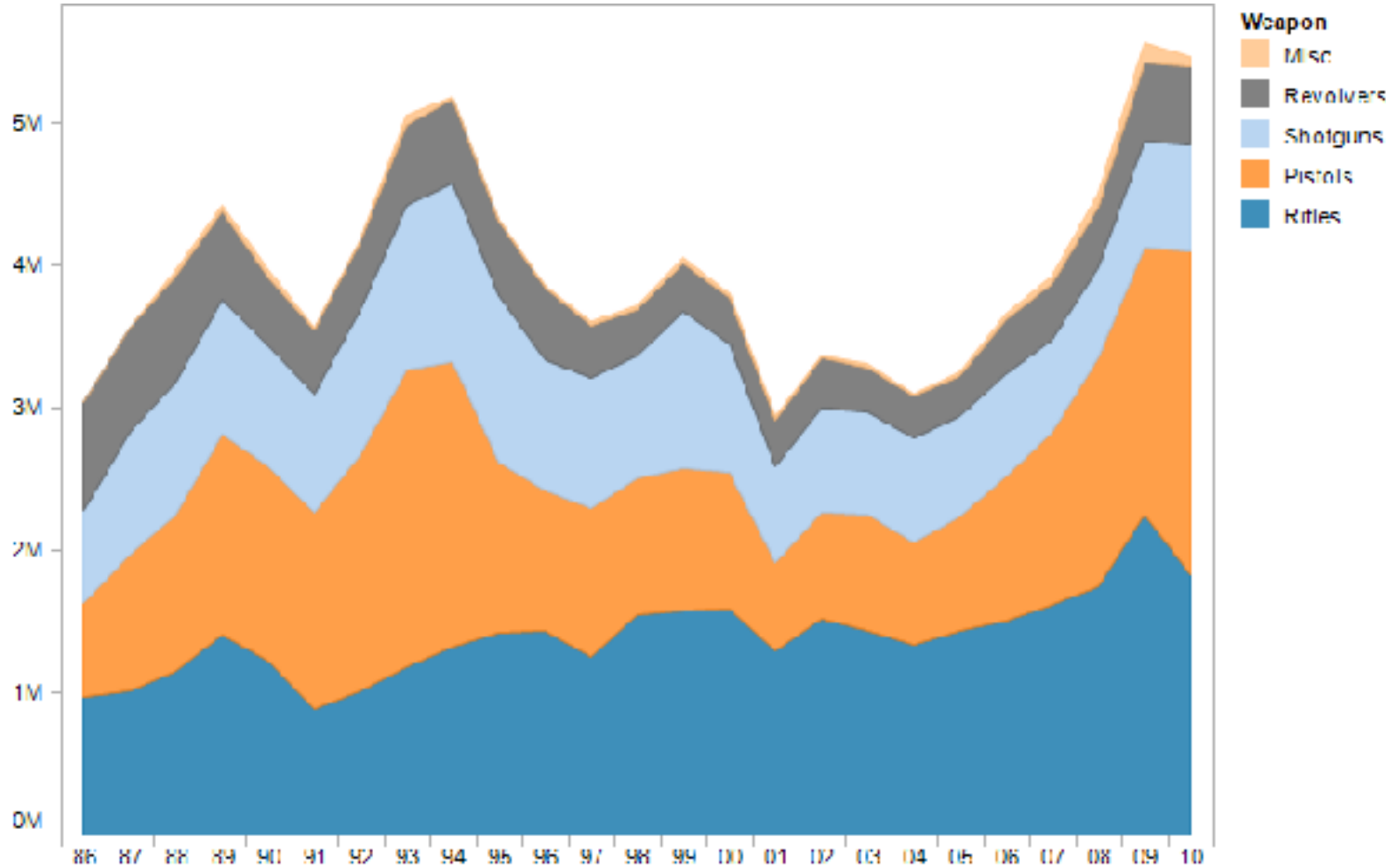
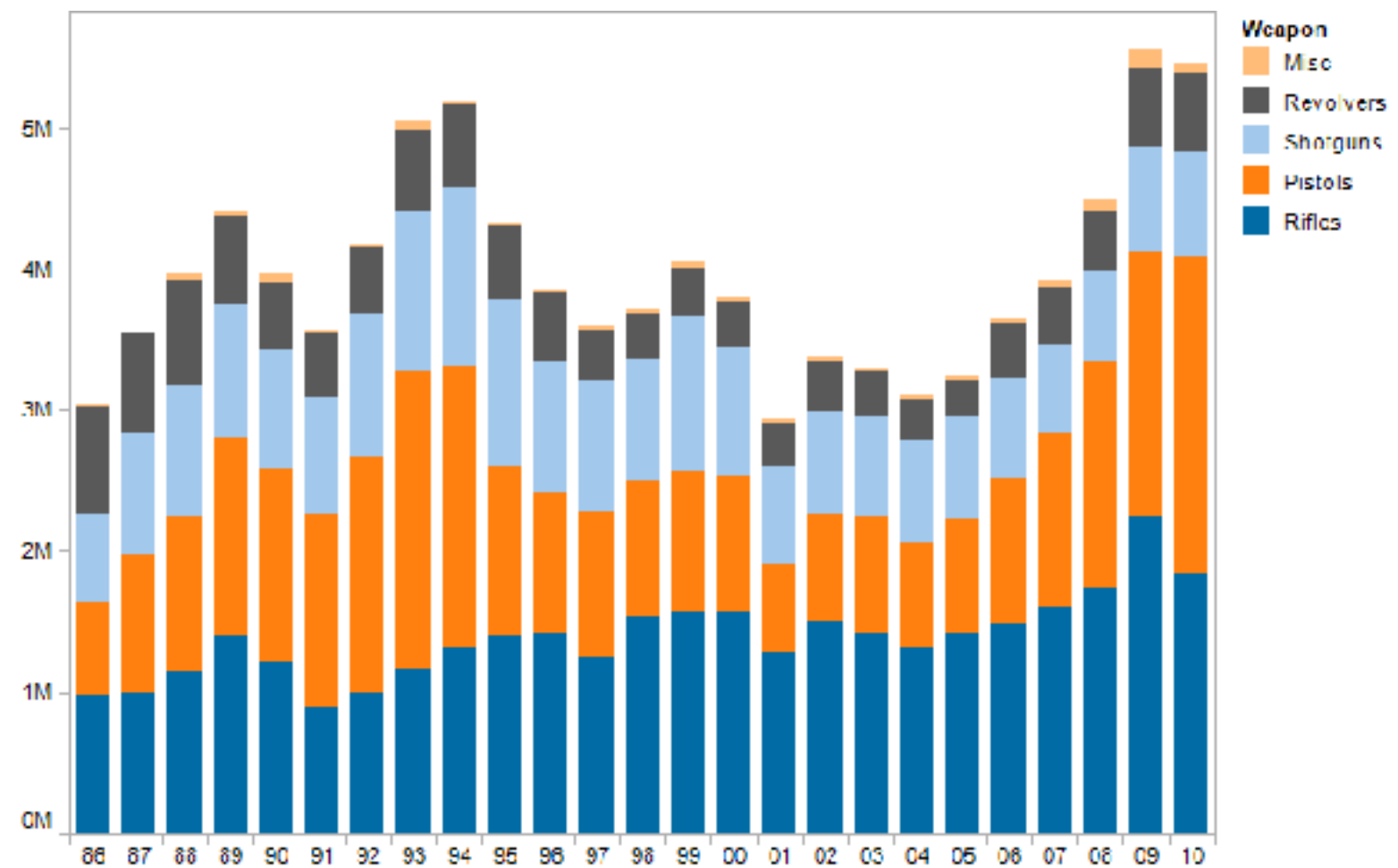
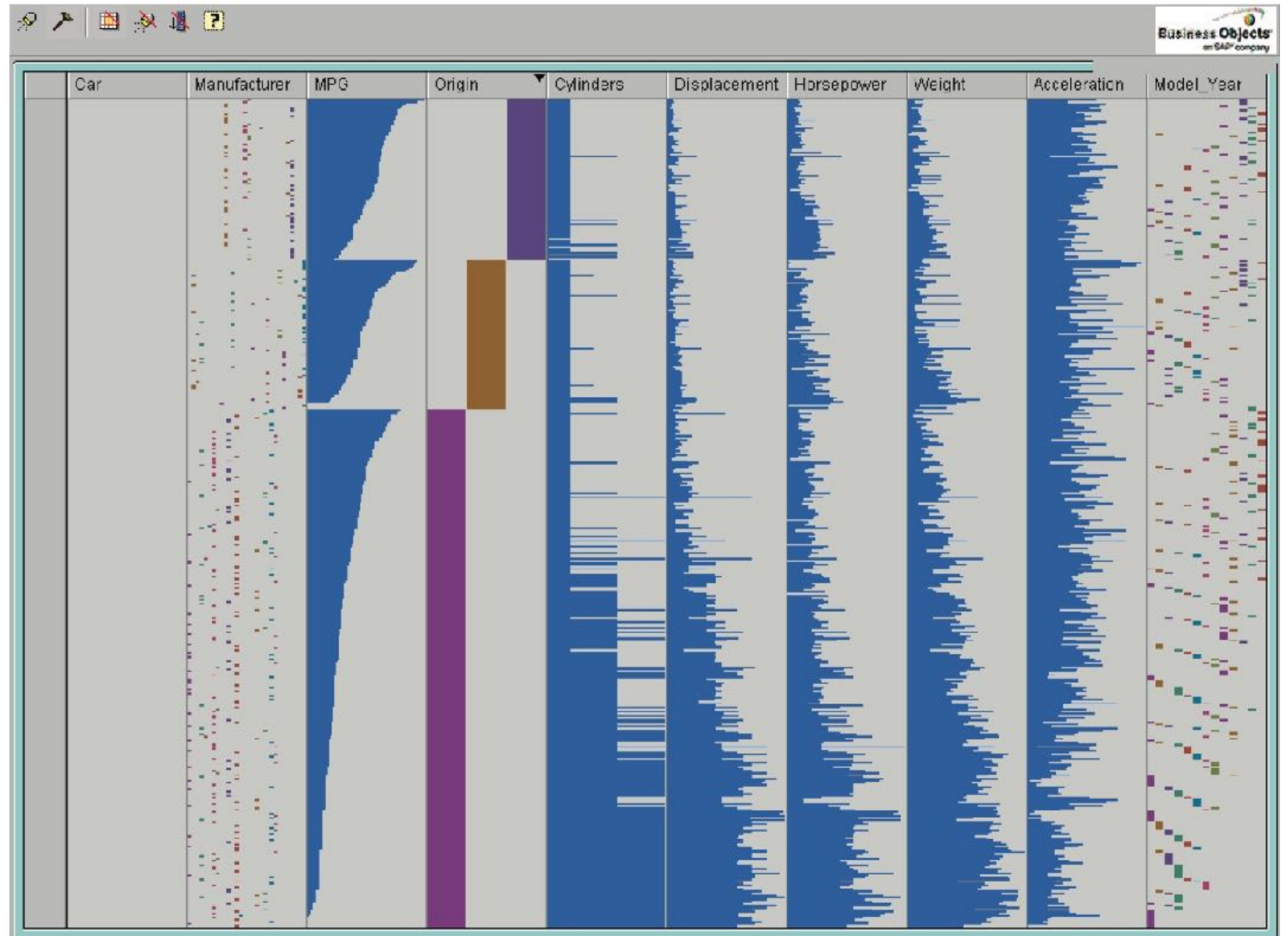
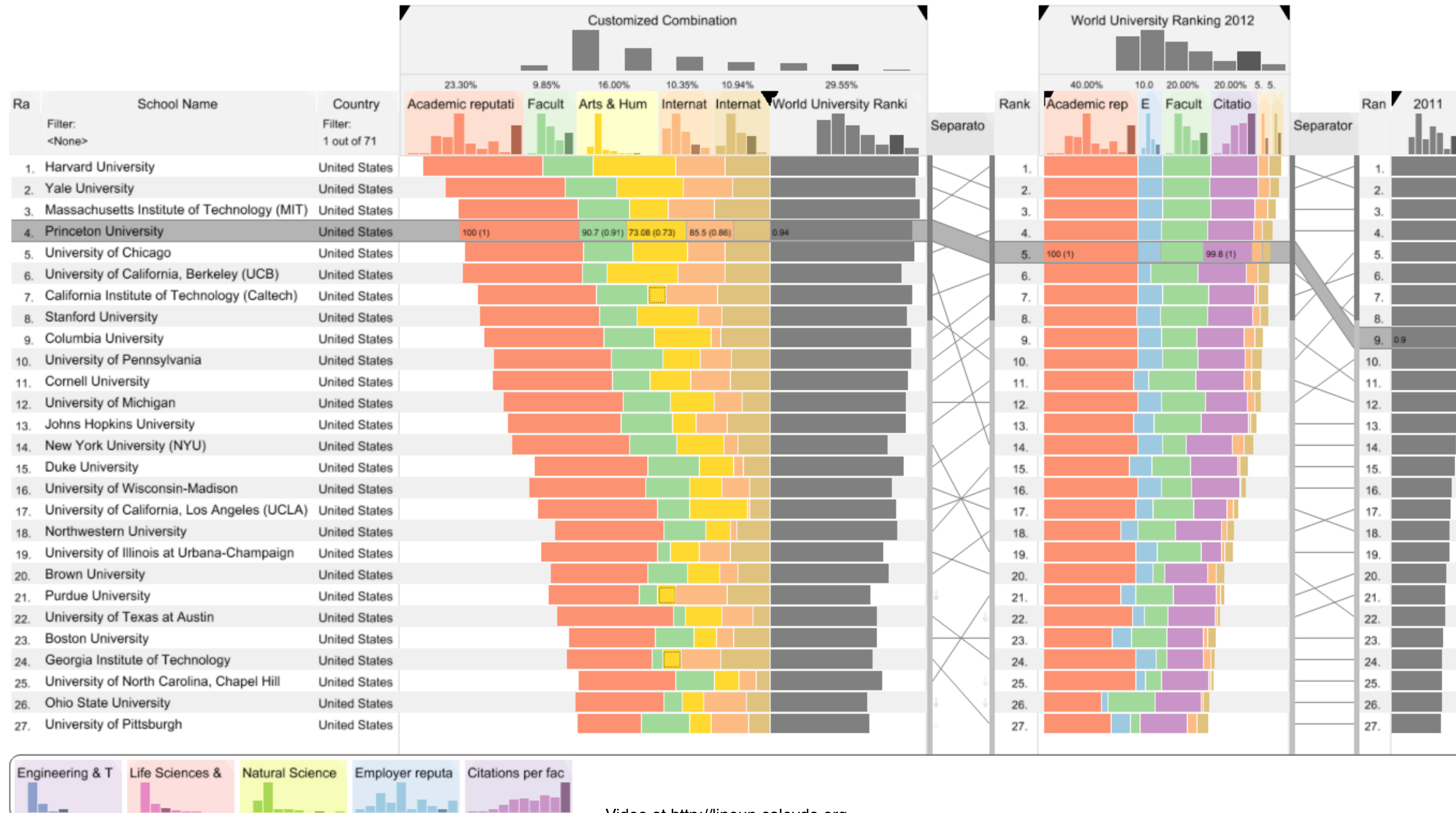


Table Lens

Interactive table-based representation



LineUp





iTunes Store Top 10 Song

Rank	Song	Artist
1	Roar	Katy Perry
2	Wake Me Up	Avicii
3	Royals	Lorde
4	Applause	Lady Gaga
5	Wrecking Ball	Miley Cyrus
6	Hold On, We're Going Home (feat. Majid Jordan)	Drake
7	Blurred Lines (feat. T.I. & Pharrell)	Robin Thicke
8	That's My Kind of Night	Ke\$ha
9	Can't Stop	Miley Cyrus
10	Proactive	Line Dragons

UNIVERSITY RANKINGS - 2012

Rank	University	Country	QS Stars Rating	Academic Reputation	Employer Reputation	Faculty Student Ratio	International Faculty	International Students
1	Massachusetts Institute of Technology	United States	5	100.00	100	98.9	98.0	98.0
2	Harvard University	United States	5	100.00	100	98.9	98.0	98.0
3	Stanford University	United States	5	100.00	100	98.9	98.0	98.0
4	Yale University	United States	5	100.00	100	98.9	98.0	98.0
5	University of California Berkeley	United States	5	100.00	100	98.9	98.0	98.0

US NEWS BEST CARS

Best Affordable Small Cars

Rank	Model	Price	Performance	Interior	Safety	Reliability	MSRP
1	Volkswagen Jetta	16,999	8.0	8.3	9.0	8.5	16,999
2	Volkswagen Jetta	17,999	8.7	8.3	9.0	8.5	17,999
3	Volkswagen Jetta	18,999	7.5	8.0	9.0	8.5	18,999

SJR Journal Rankings

Rank	Journal Title	Index	Total Docs. (2012)	Total Docs. (Years)	Total Refs.	Total Cites (Years)	Cites / Doc. (Years)	Country
1	IEEE Transactions on Pattern Analysis and Machine Intelligence	8.094	200	195	8,249	3,724	54.3	USA
2	ACM Computing Surveys	6.751	81	30	4,902	2,024	7.11	USA
3	Foundations and Trends in Information Retrieval	6.530	12	10	661	283	22.130	USA
4	Foundations and Trends in Computer Vision	6.167	171	3	1,746	10	0.72	USA
5	Journal of the ACM	5.997	11	114	3	4,911	44.81	USA
6	ACM Transactions on Internet Technology	5.949	81	7	1,111	91	13.46	USA
7	Foundations and Trends in Machine Learning	5.570	6	27	3	15.30	16.710	USA

Ranking.com

Rank	Domain	Page Rank	Category	Language	Advanced
1	google.com	10	Search	English	Yes
2	youtube.com	10	Video	English	Yes
3	facebook.com	10	Social	English	Yes
4	yahoo.com	10	Search	English	Yes
5	wikipedia.org	10	Reference	English	Yes
6	amazon.com	10	Shopping	English	Yes
7	apple.com	10	Technology	English	Yes
8	microsoft.com	10	Technology	English	Yes
9	ask.com	10	Search	English	Yes
10	wordpress.com	10	Web	English	Yes

The World's Billionaires

Rank	Name	Net Worth	Industry	Country
1	Larry Ellison	\$43 B	Software	United States
2	Charles Koch	\$34 B	Diversified	United States
3	David Koch	\$34 B	Diversified	United States
4	Michael Bloomberg	\$33 B	Media	United States
5	Mark Zuckerberg	\$31 B	Technology	United States
6	Li Ka-shing	\$31 B	Diversified	Hong Kong

Things to do in Atlanta

Attractions travelers recommend

Rank	Attraction	Reviews
1	Fox Theatre	576 reviews
2	Atlanta Botanical Garden	697 reviews



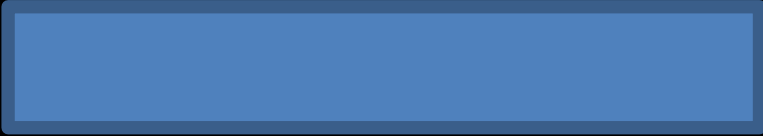


World University Rankings 2011-2012

Rank	Institution	Country	Score
1	California Institute of Technology	United States	93.7
2	University of Oxford	United Kingdom	93.6
3	Stanford University	United States	93.1
4	Harvard University	United States	92.7
5	Massachusetts Institute of Technology	United States	92.6
6	Princeton University	United States	90.6
7	University of Cambridge	United Kingdom	90.5
8	Imperial College London	United Kingdom	90.4

U.S. and Canada Box Office



















Rank	Title	Distributor	Weekend Gross	Cumulative Gross	Wks Out	# of Theaters
1	EST MODUS IN REBUS	The Weinstein Company	\$20,201,300	\$79,466,400	3	3330
2	\$18,472,900	\$18,472,900	1	2735

Rankings are popular

Rank	University	Score
1.	MIT, USA	
2.	Harvard, U	
3.	Princeton,	
4.	Cambridge	
5.	Oxford, UK	 4.0

Support Multiple Attributes

$$\text{Score} = f(A, B, C)$$

Rank	University	A	B	C
1.	MIT, USA			
2.	Harvard, USA		 	
3.	Princeton, US		 	
4.	Cambridge, U		 	
5.	Oxford, UK			

Combiner functions: $f(A,B,C)$

(Weighted) sum

$$\text{Score} = w_a A + w_b B + w_c C$$

→ Serial

Maximum

$$\text{Score} = \max(A, B, C)$$

→ Parallel

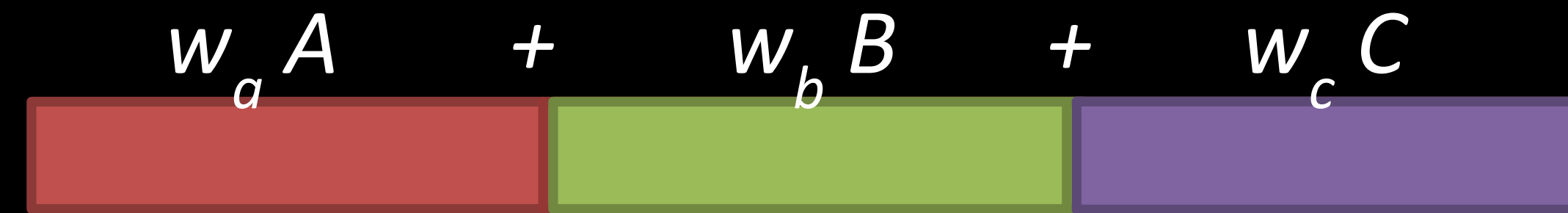
Product

Nesting

...

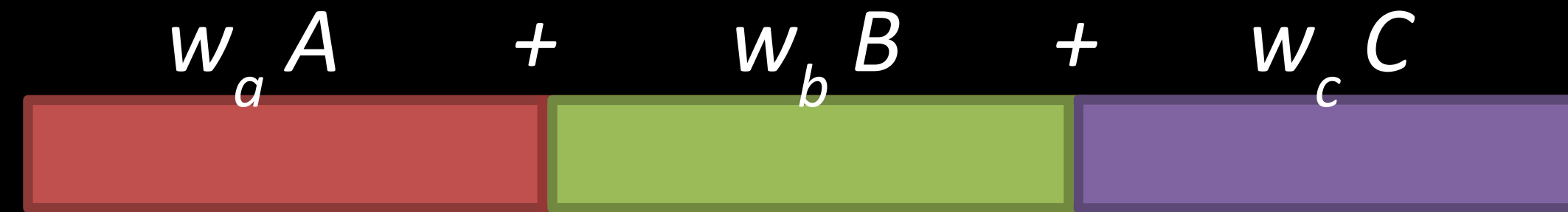
→ Complex
Combiners

Serial Combiner (as Stacked Bar)



Rank	University	A	B	C
1.	MIT, USA			
2.	Harvard, USA			
3.	Princeton, USA			
4.	Cambridge, UK			
5.	Oxford, UK			

Serial Combiner (as Stacked Bar)



Rank	University	A	B	C
1.	MIT, USA	Large	Large	Large
2.	Harvard, USA	Large	Large	Large
3.	Princeton, USA	Medium	Large	Large
4.	Cambridge, UK	Large	Large	Small
5.	Oxford, UK	Large	Small	Small

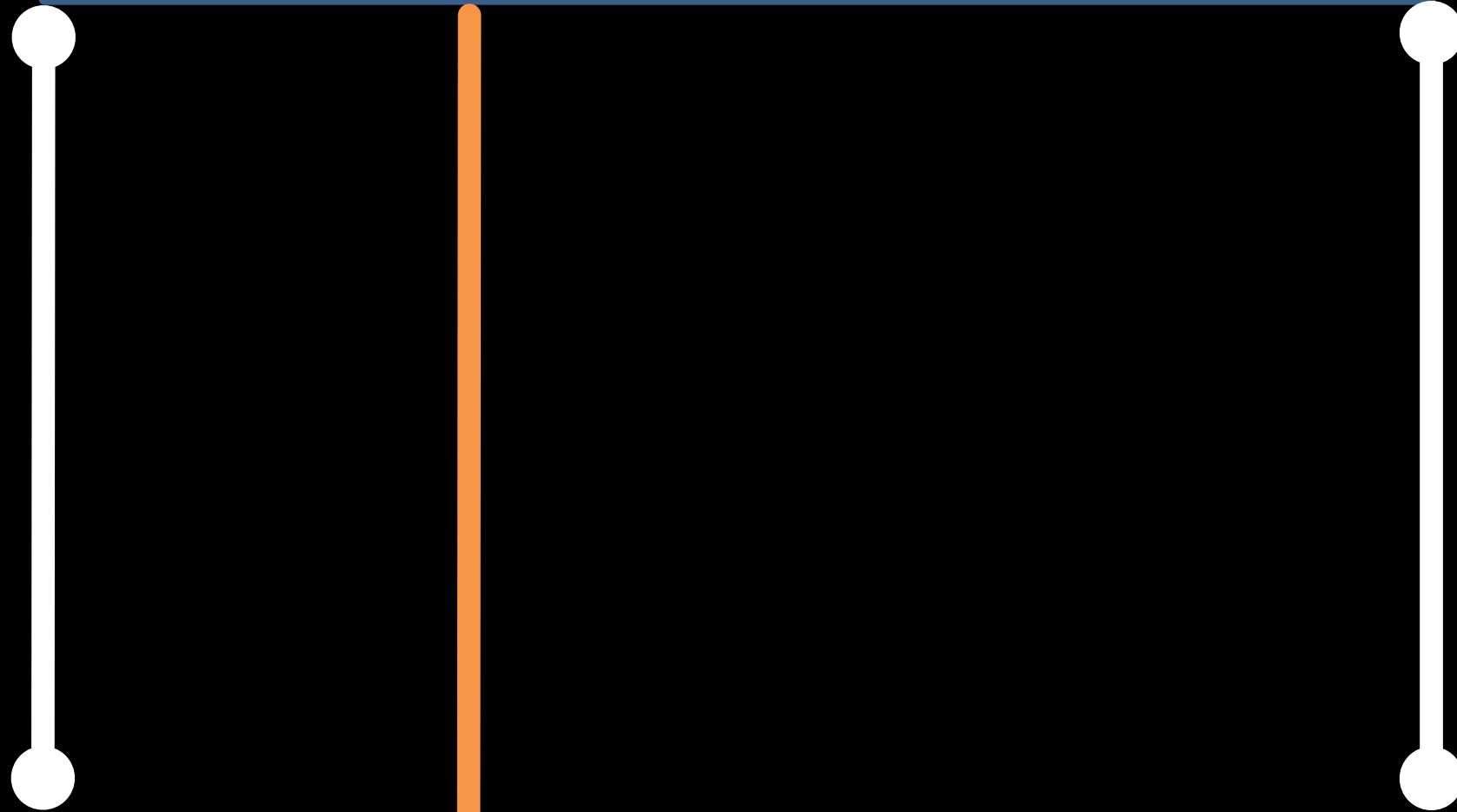
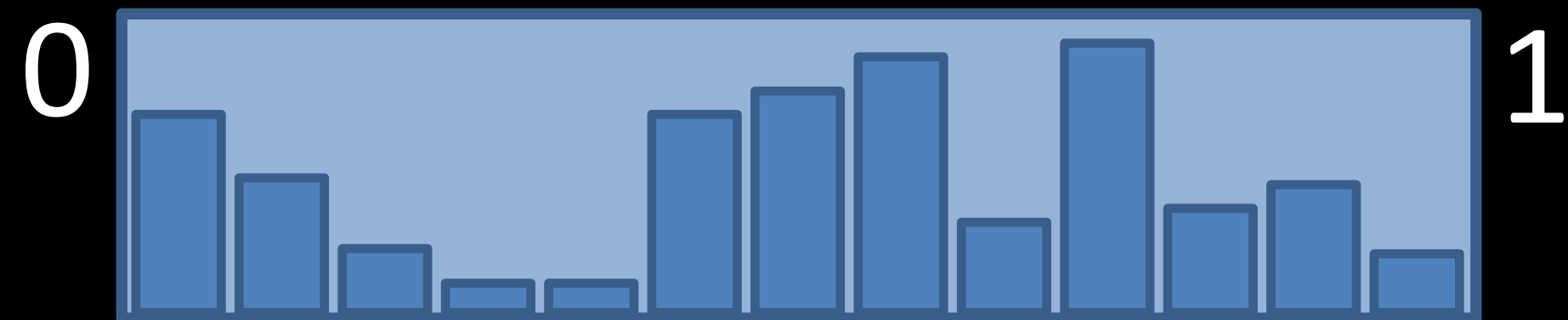
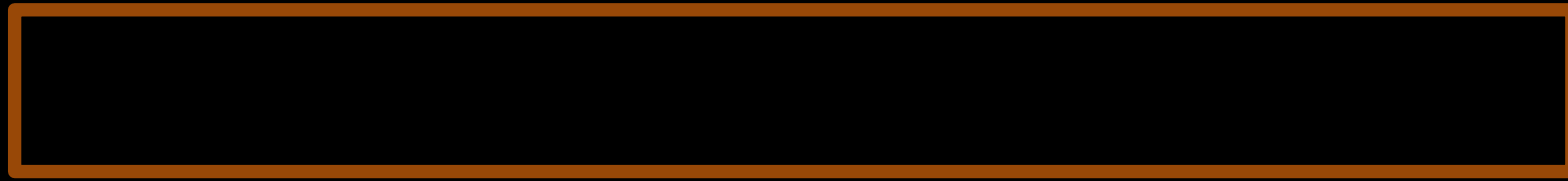
Serial Combiner (as Stacked Bar)



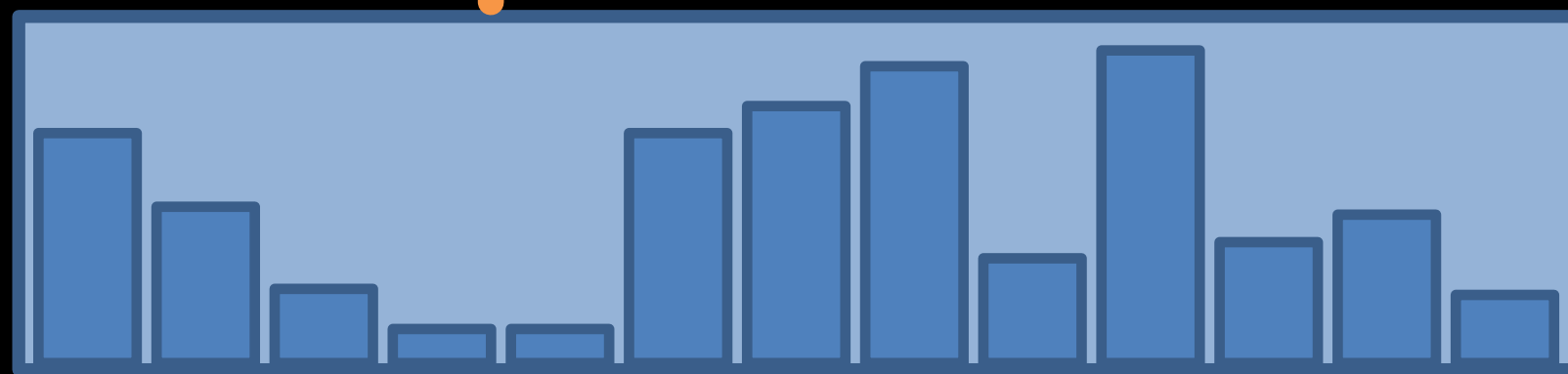
Rank	University	A	B	C
1.	MIT, USA	Large red segment	Medium green segment	Medium purple segment
2.	Harvard, USA	Large red segment	Medium green segment	Medium purple segment
3.	Princeton, USA	Medium red segment	Medium green segment	Medium purple segment
4.	Cambridge, UK	Large red segment	Small green segment	Small purple segment
5.	Oxford, UK	Large red segment	Very small green segment	Small purple segment

Rank	School Name	Country	Faculty/student ratio	Employer reputation	Citations per faculty
	Filter: <None>	Filter: 2 out of 72			
1.	American University	United States			
2.	Arizona State University	United States			
3.	Aston University	United Kingdom			
4.	Birkbeck College, University of L	United Kingdom			
5.	Boston College	United States			
6.	Boston University	United States			
7.	Brandeis University	United States			
8.	Brown University	United States			
9.	Brunel University	United Kingdom			
10.	California Institute of Technology	United States			
11.	Cardiff University	United Kingdom			
12.	Case Western Reserve University	United States			
13.	City University London	United Kingdom			
14.	College of William & Mary	United States			
15.	Colorado State University	United States			
16.	Columbia University	United States			
17.	Cornell University	United States			
18.	Cranfield University	United Kingdom			
19.	Dartmouth College	United States			
20.	Drexel University	United States			
21.	Duke University	United States			
22.	Durham University	United Kingdom			

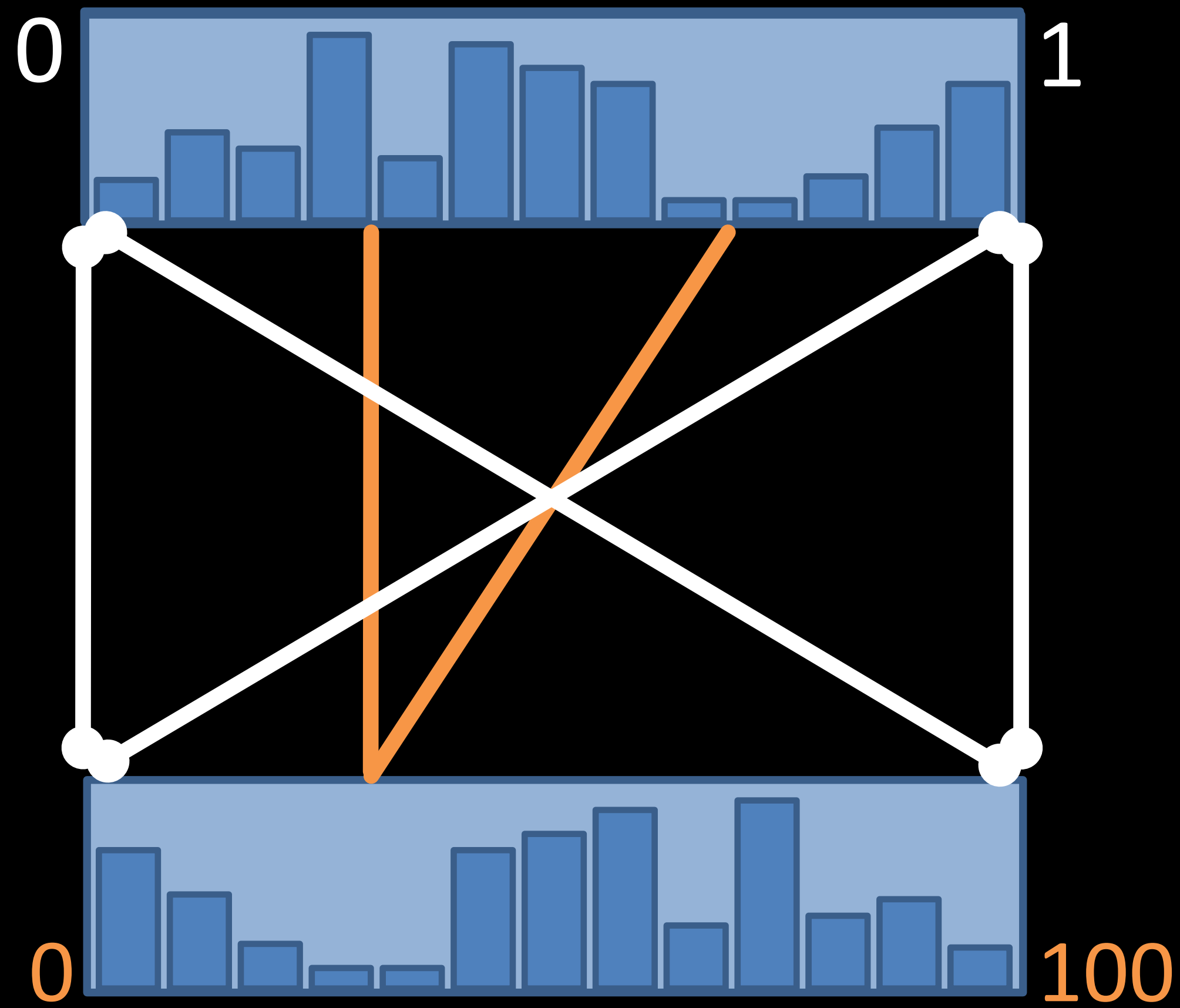
Flexible Mapping of Attributes to Scores

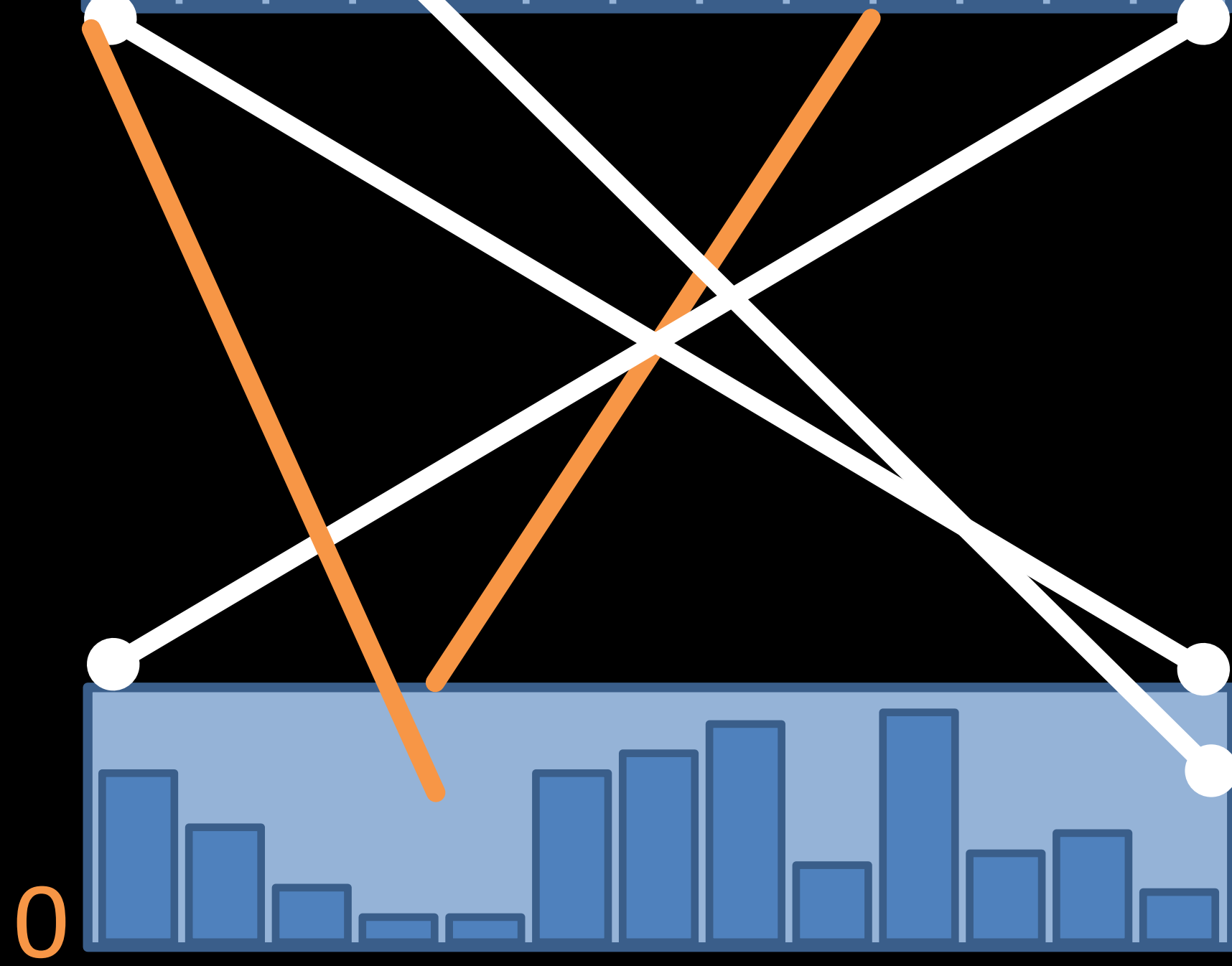
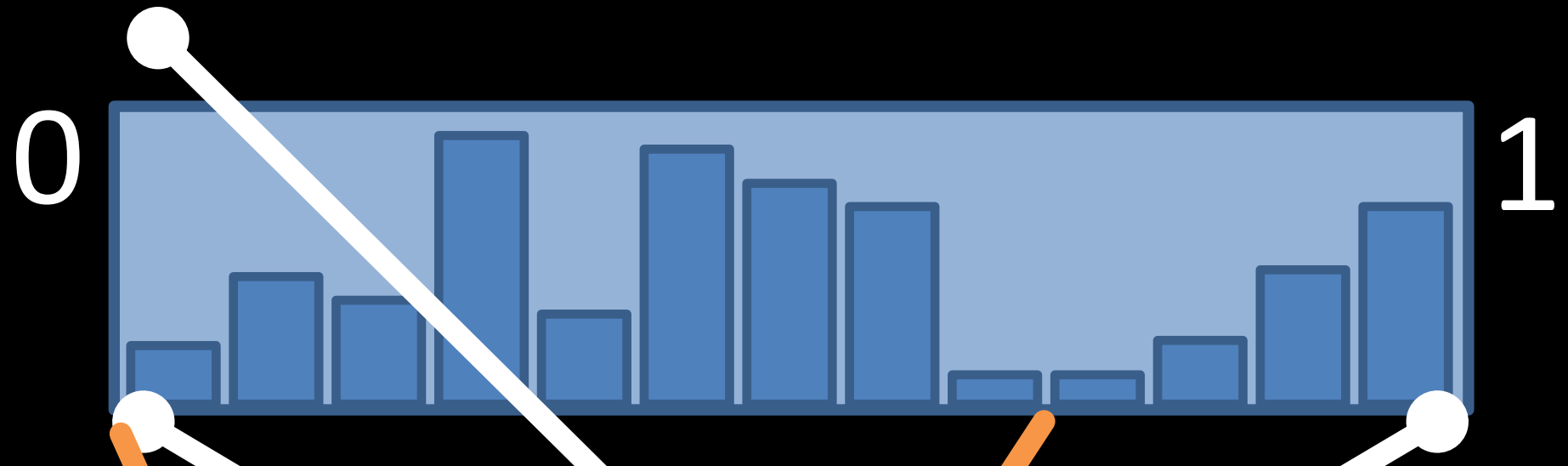


Mi0



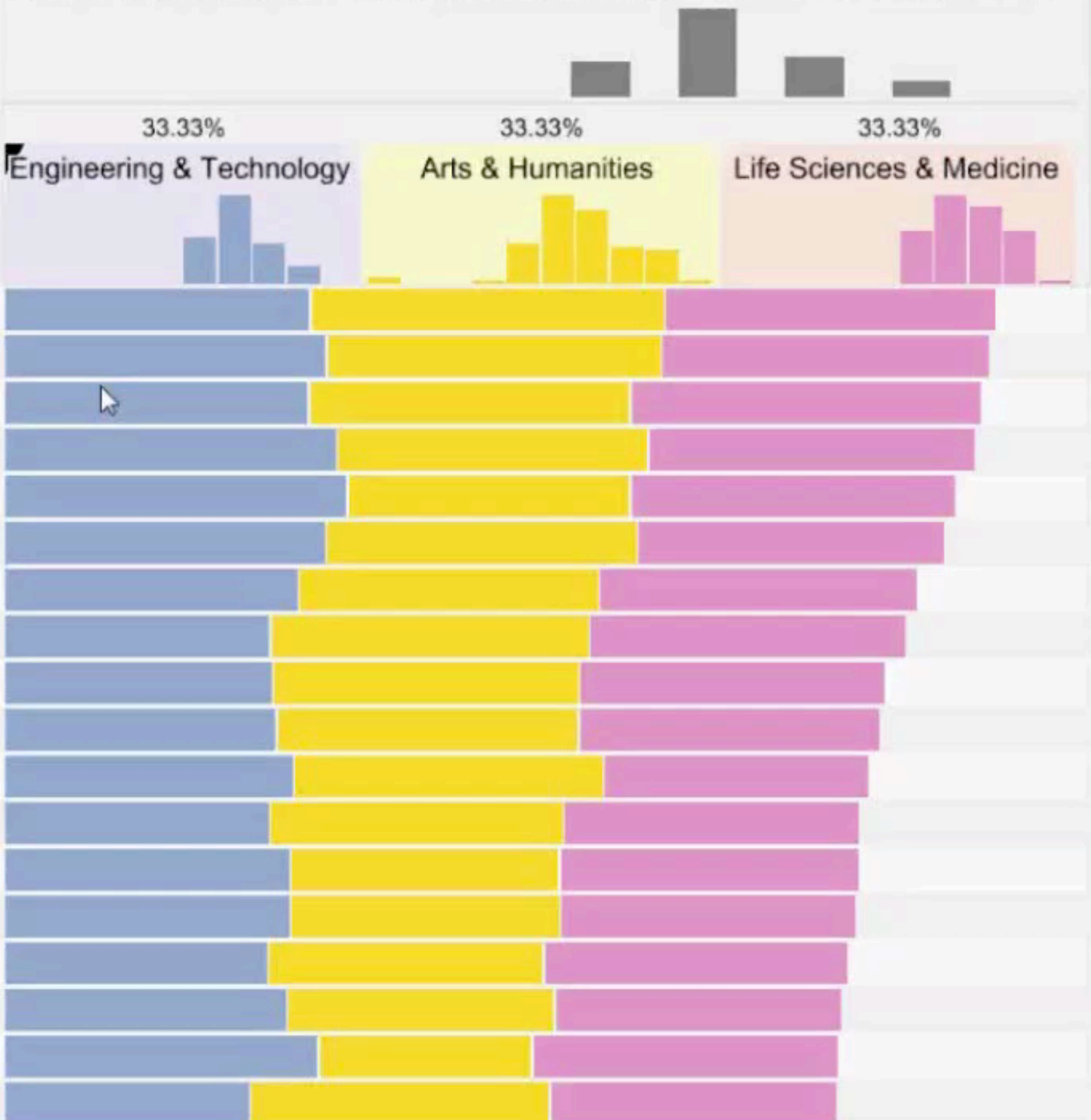
Max





SUM (Engineering & Technology, Arts & Humanities, Life Sciences & Medicine)

Ran	School Name	Country
	Filter: <None>	Filter: 2 out of 43
1.	University of Oxford	United Kingdom
2.	University of Cambridge	United Kingdom
3.	Harvard University	United States
4.	Stanford University	United States
5.	Massachusetts Institute of Technology (MIT)	United States
6.	University of California, Berkeley (UCB)	United States
7.	University of California, Los Angeles (UCL)	United States
8.	Yale University	United States
9.	UCL (University College London)	United Kingdom
10.	Columbia University	United States
11.	Princeton University	United States
12.	University of Edinburgh	United Kingdom
13.	University of Michigan	United States
14.	Cornell University	United States
15.	University of Pennsylvania	United States
16.	The University of Manchester	United Kingdom
17.	Imperial College London	United Kingdom
18.	University of Chicago	United States

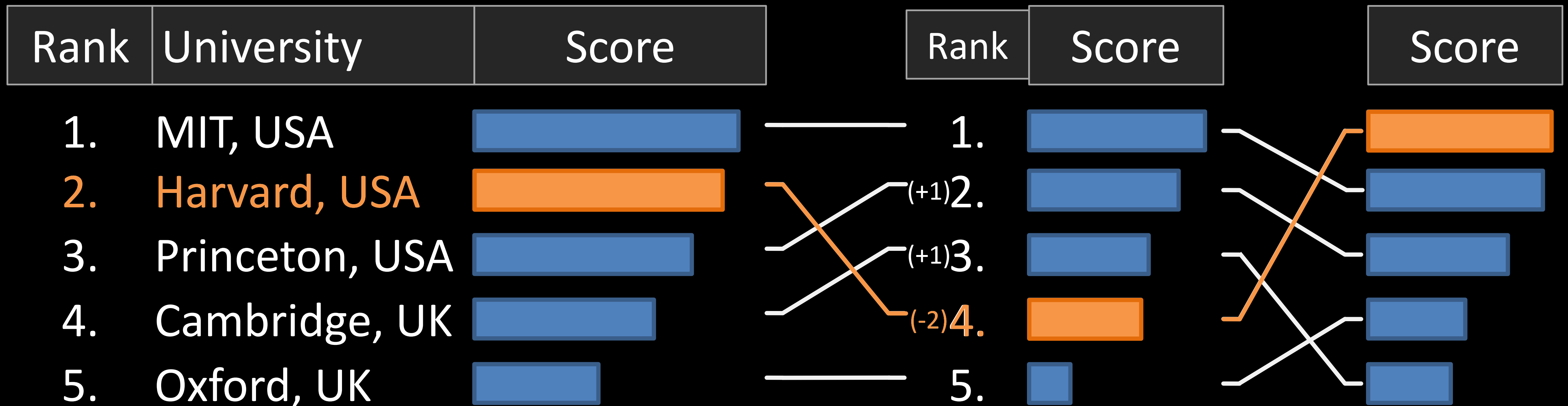


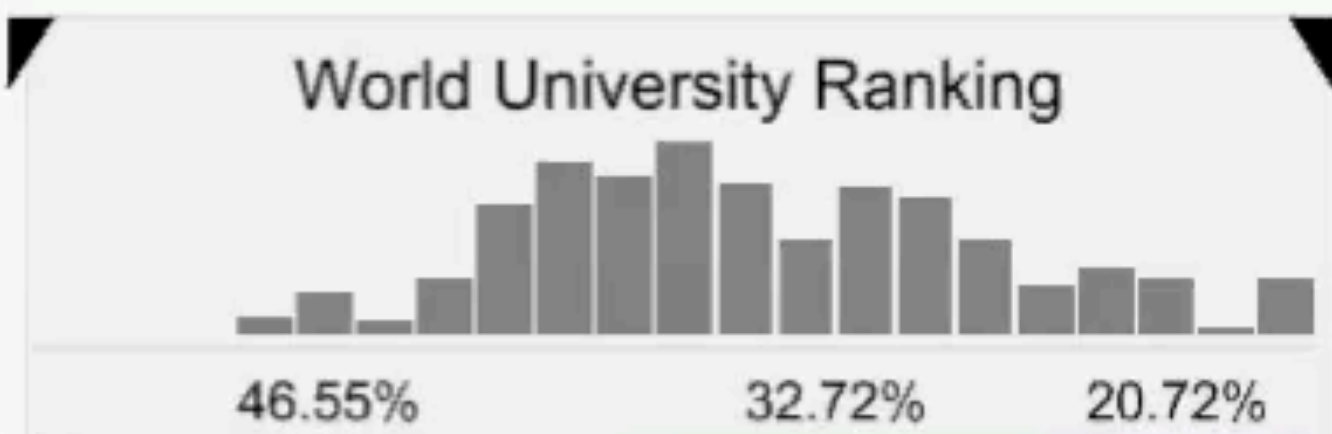
Compare Rankings

Bump Charts

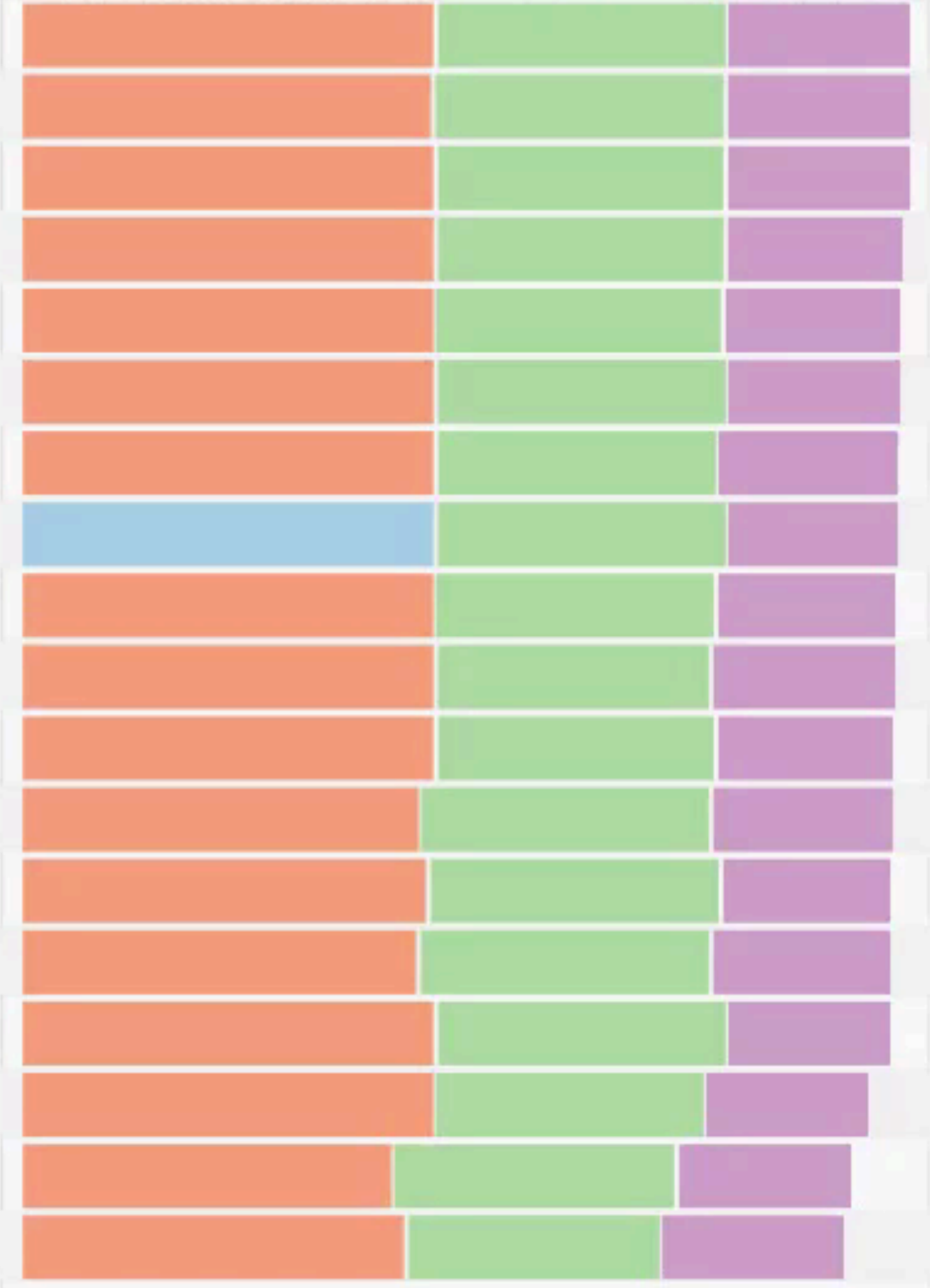


Bump Charts





Rank	School Name
Filter:	<None>
1.	Massachusetts Institute of Te
2.	California Institute of Technol
3.	Harvard University
4.	University of Cambridge
5.	UCL (University College Lond
6.	University of Oxford
7.	Princeton University
8.	Imperial College London
9.	University of Chicago
10.	Stanford University
11.	Columbia University
12.	Duke University
13.	University of Pennsylvania
14.	Johns Hopkins University
15.	Yale University
16.	University of Michigan
17.	Ecole normale supérieure, Pa
18.	Northwestern University



<http://lineup.caleydo.org>

Pixel Based Displays

Each cell is a “pixel”, value encoded in color / value

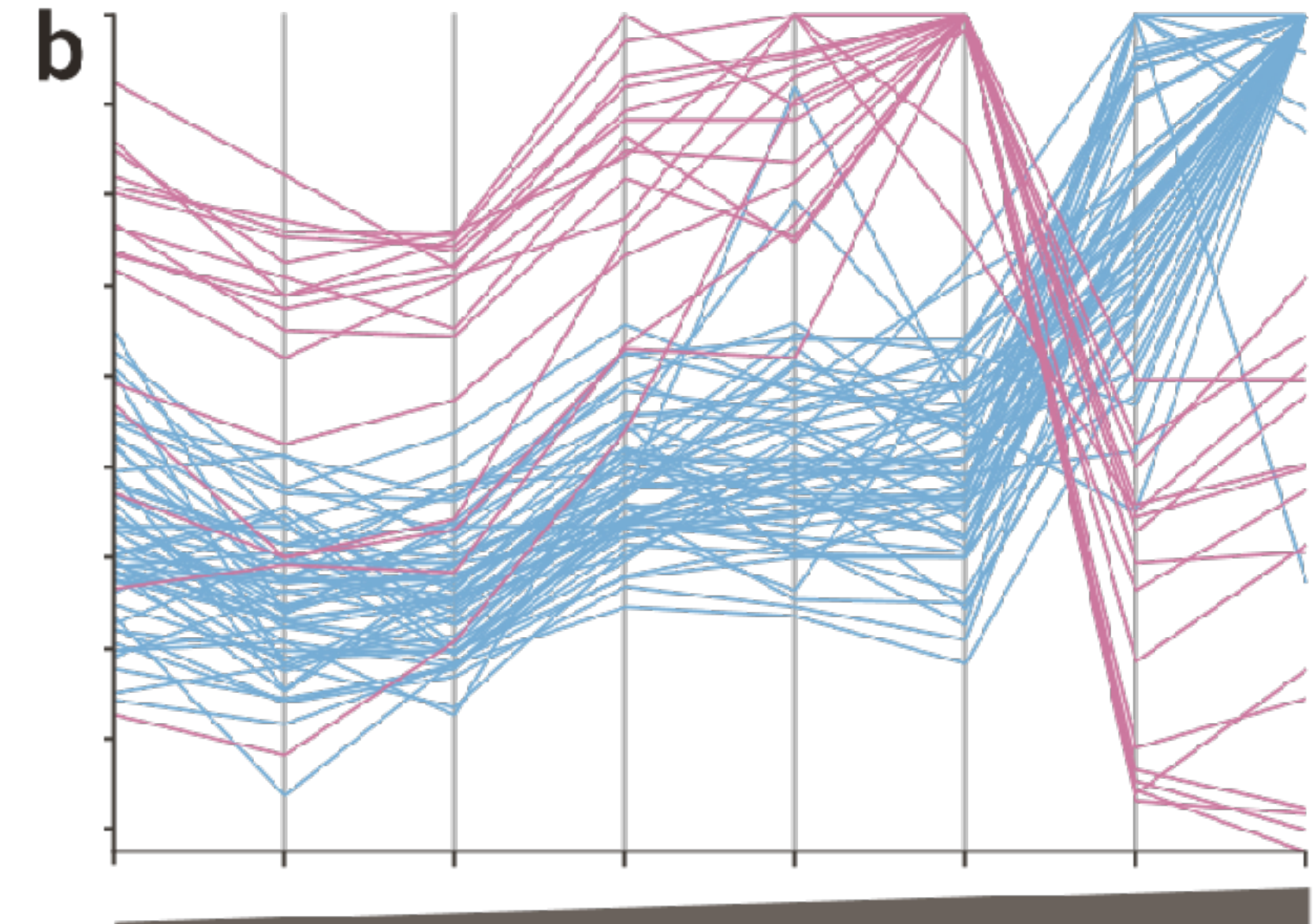
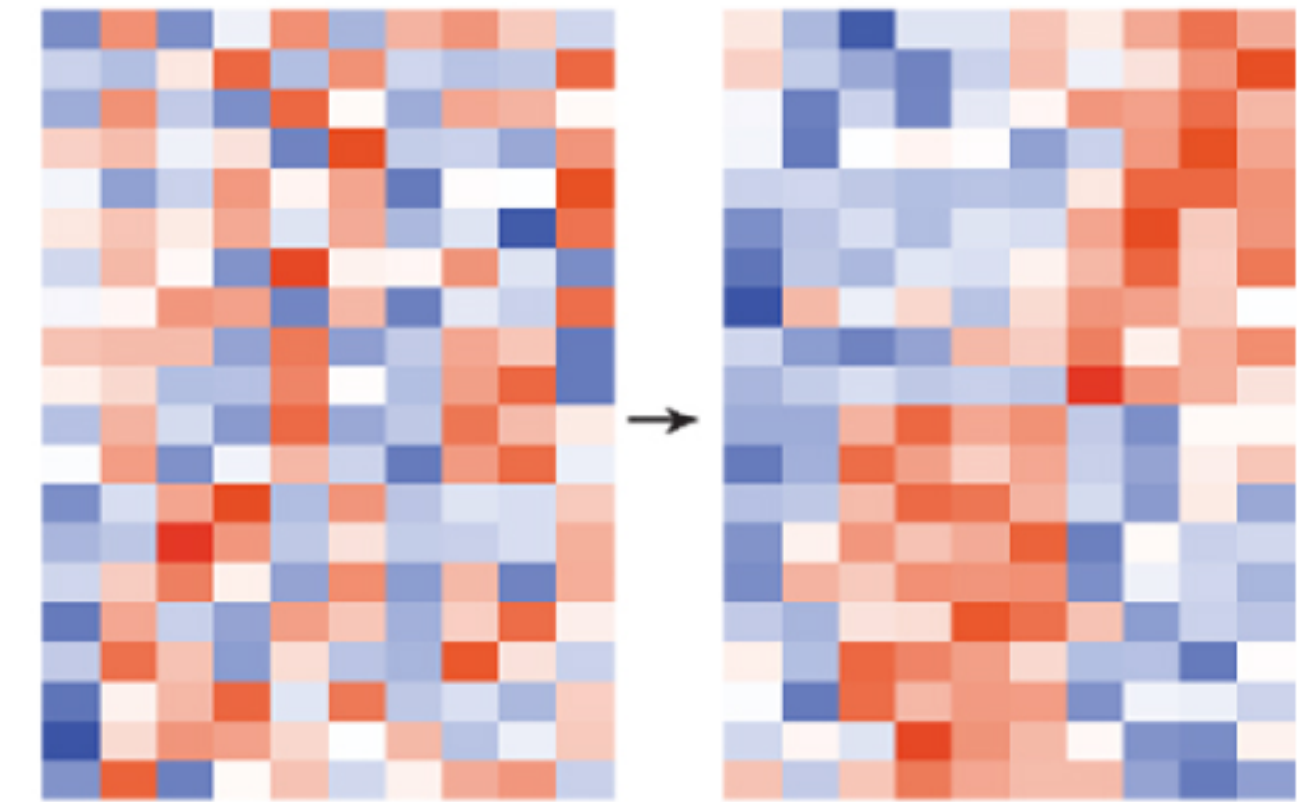
Ordering critical for interpretation

If no ordering inherent, clustering is used

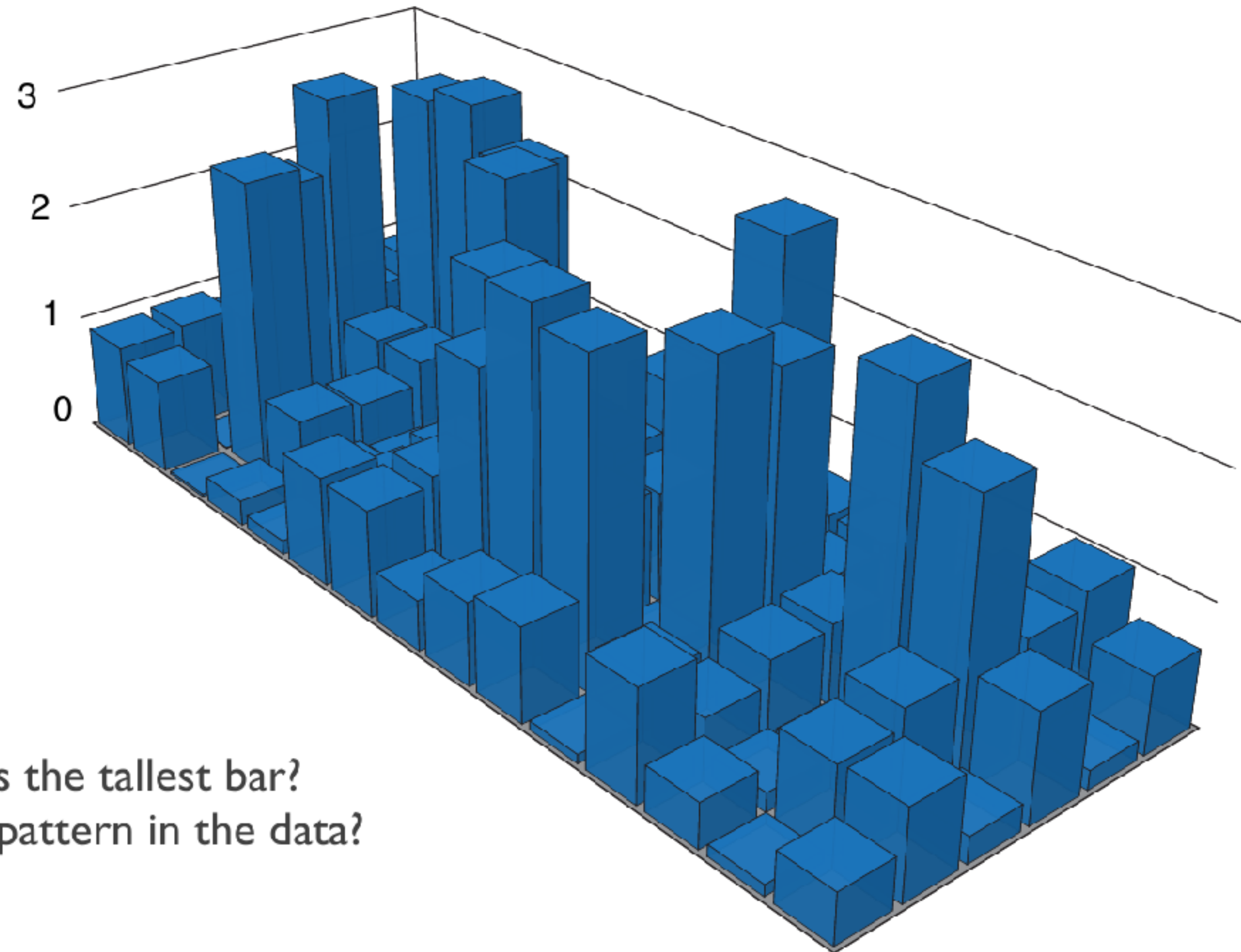
Scalable – 1 px per item

Good for homogeneous data

same scale & type

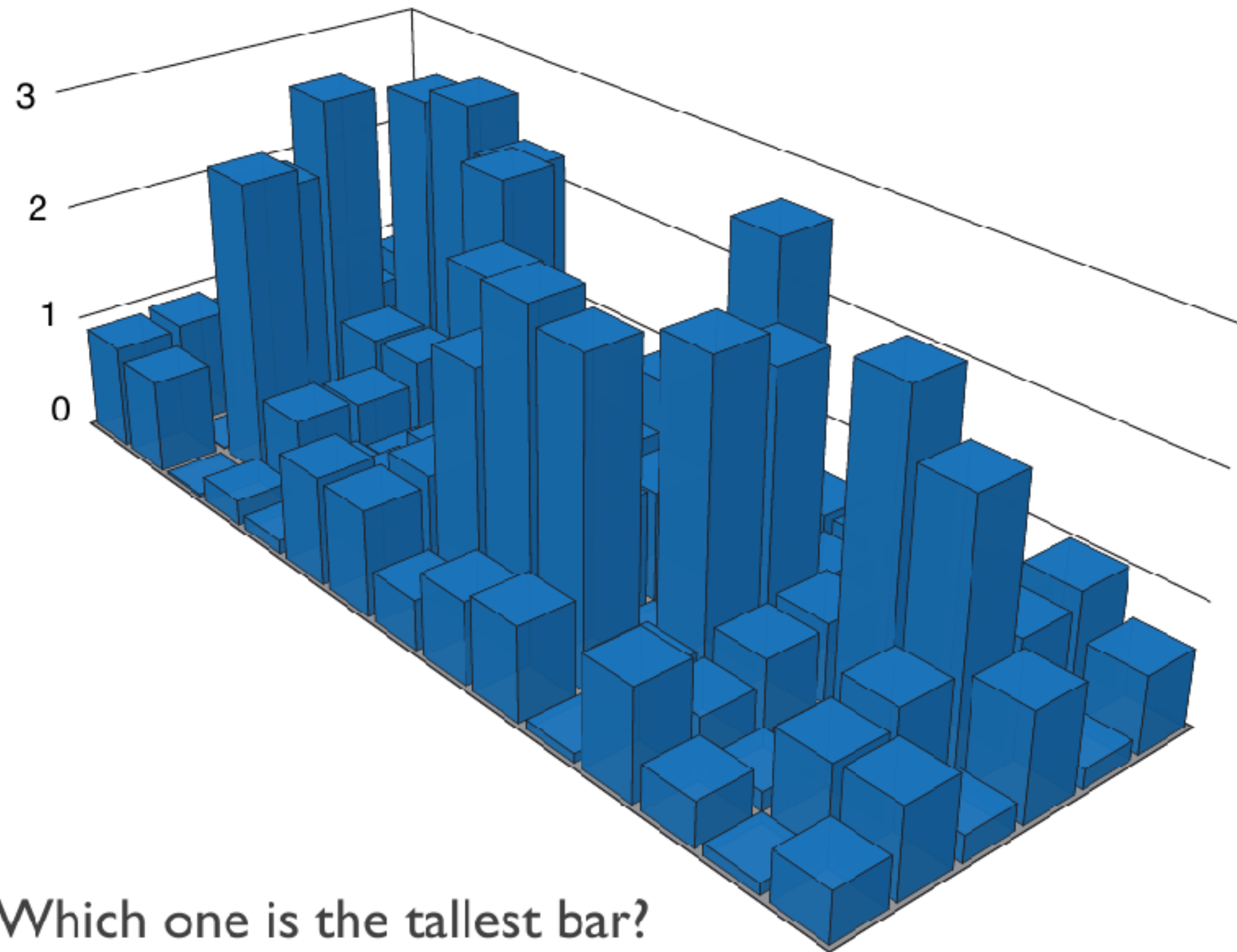


3D Pitfall: Occlusion & Perspective

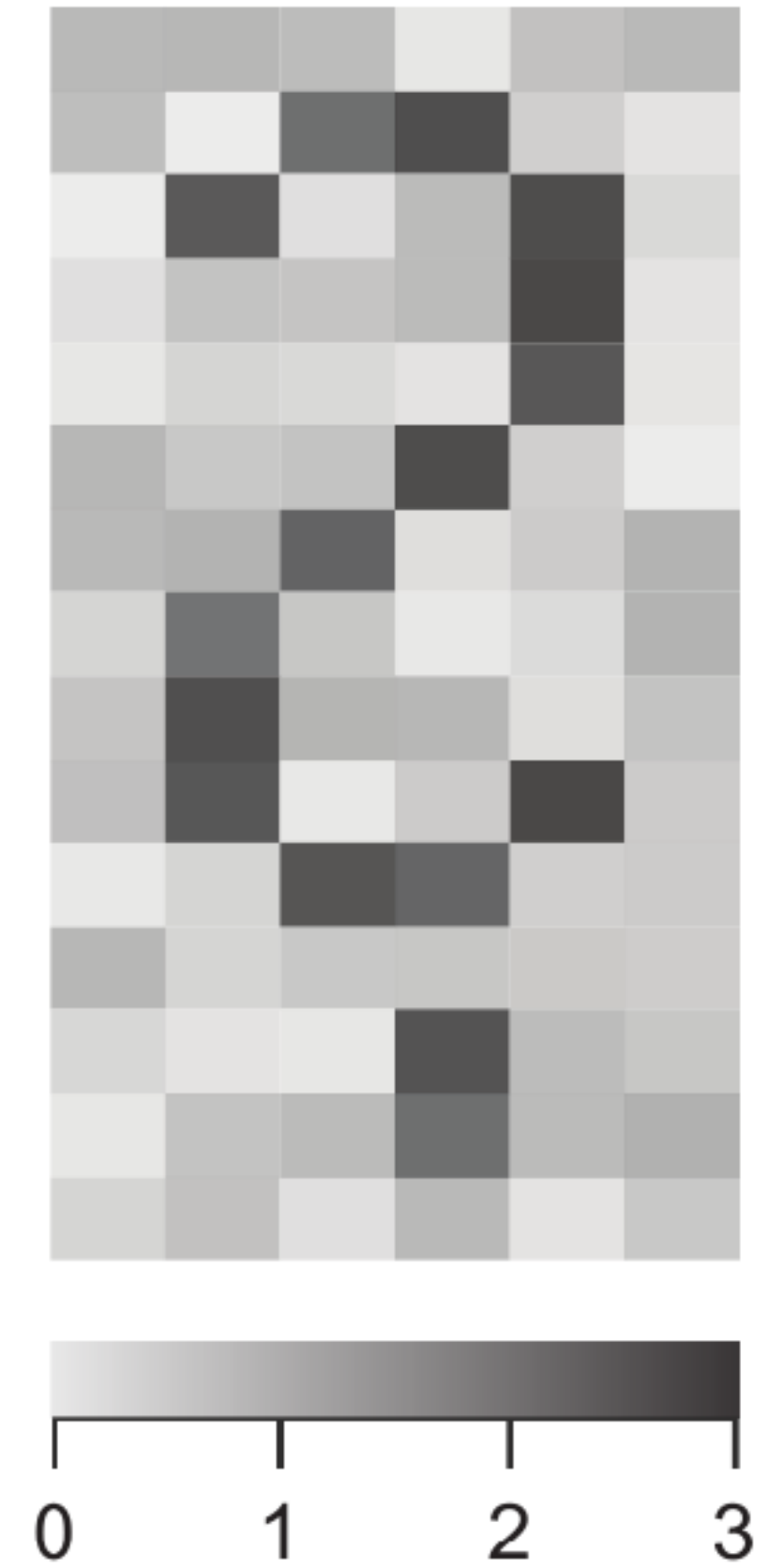


Which one is the tallest bar?
What is the pattern in the data?

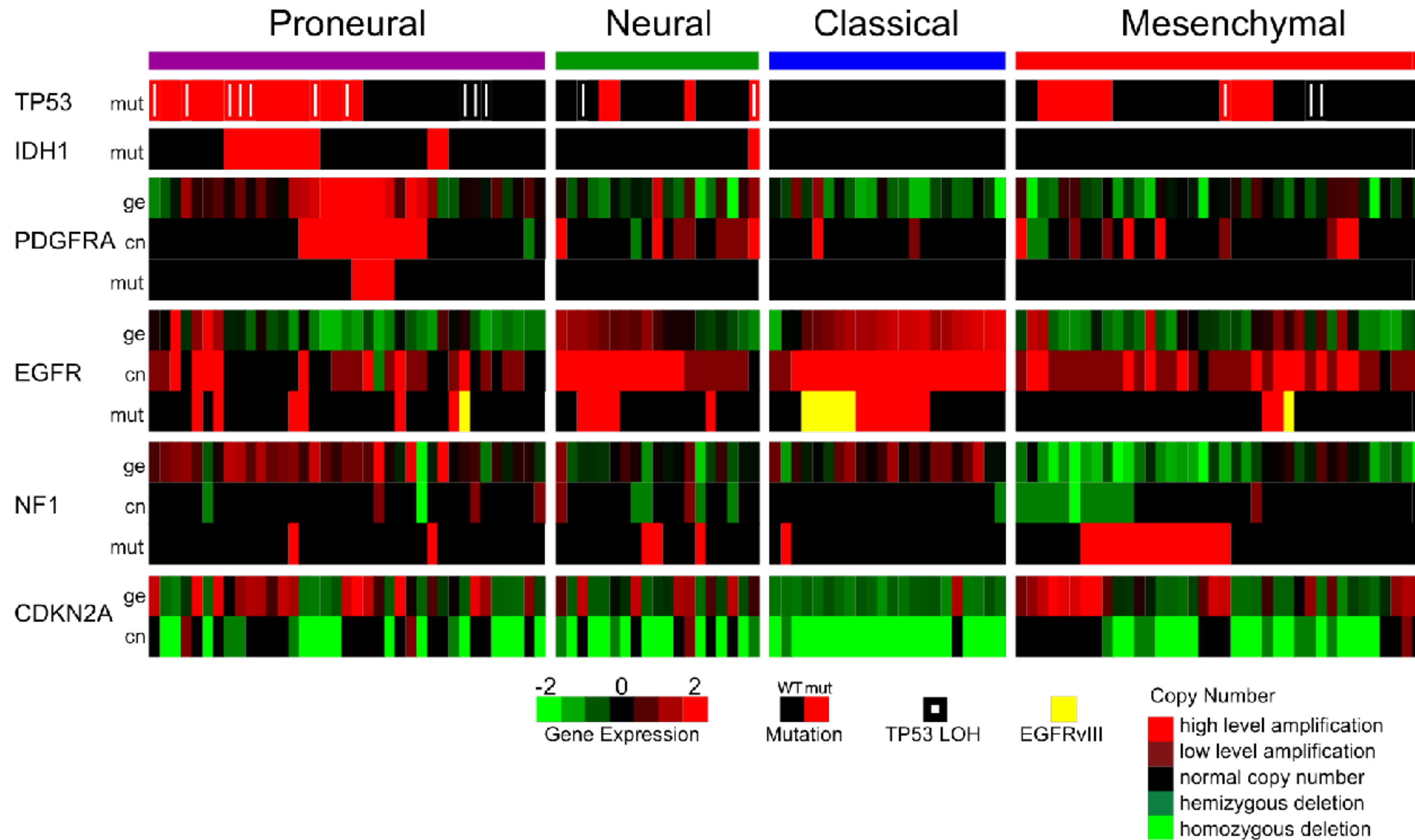
3D Pitfall: Occlusion & Perspective



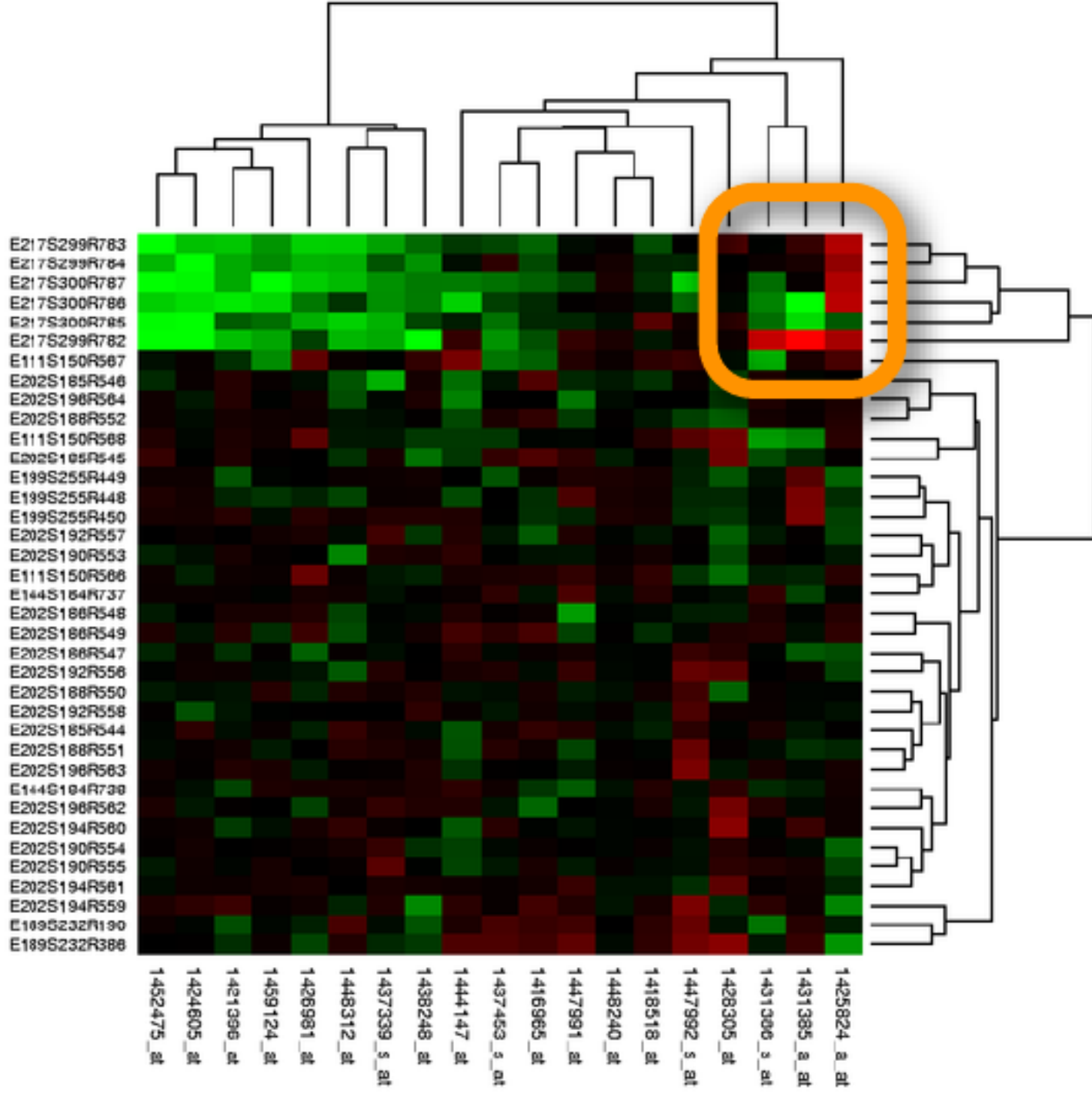
Which one is the tallest bar?
What is the pattern in the data?



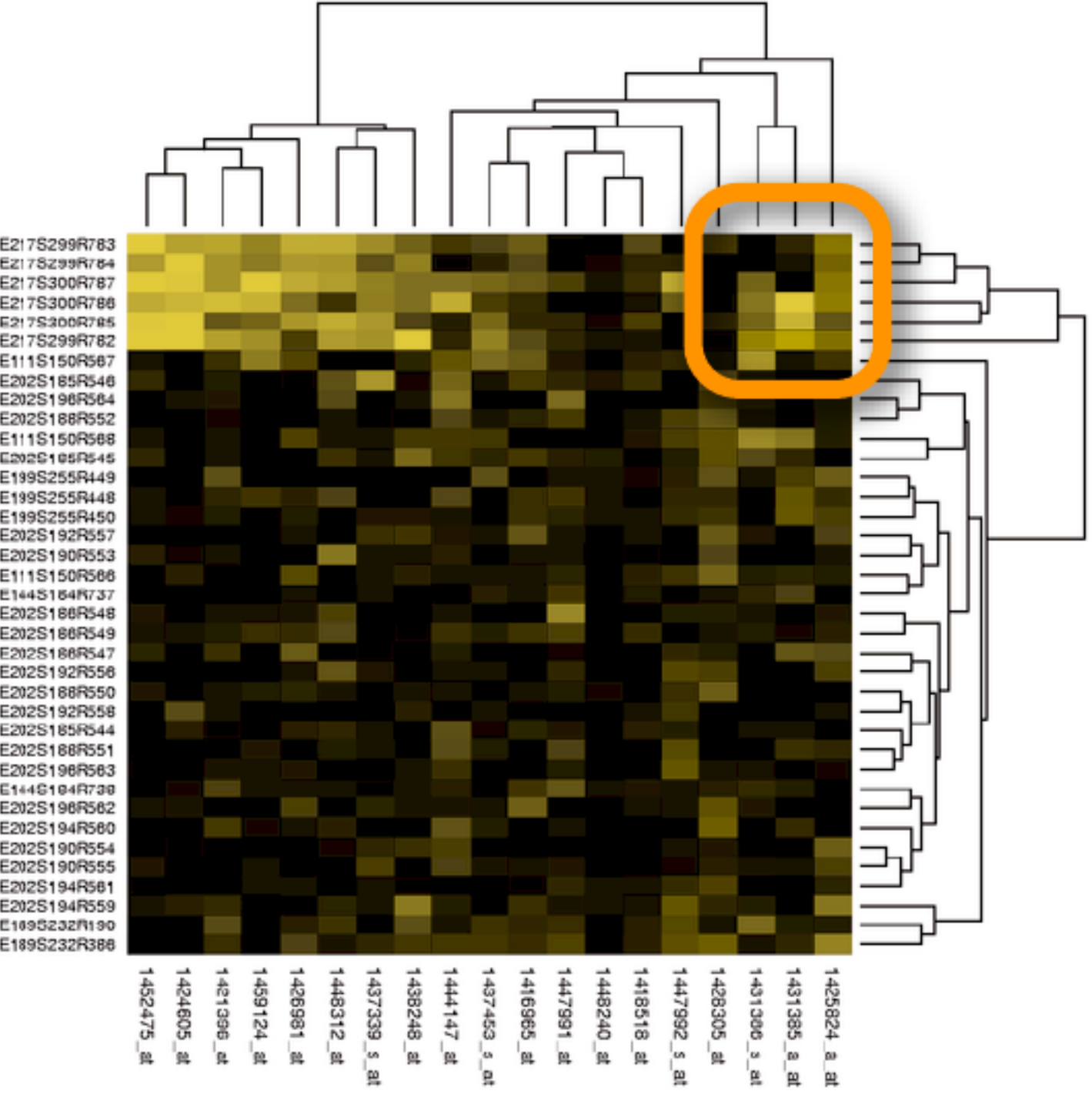
Heterogeneous Data?



Bad Color Mapping

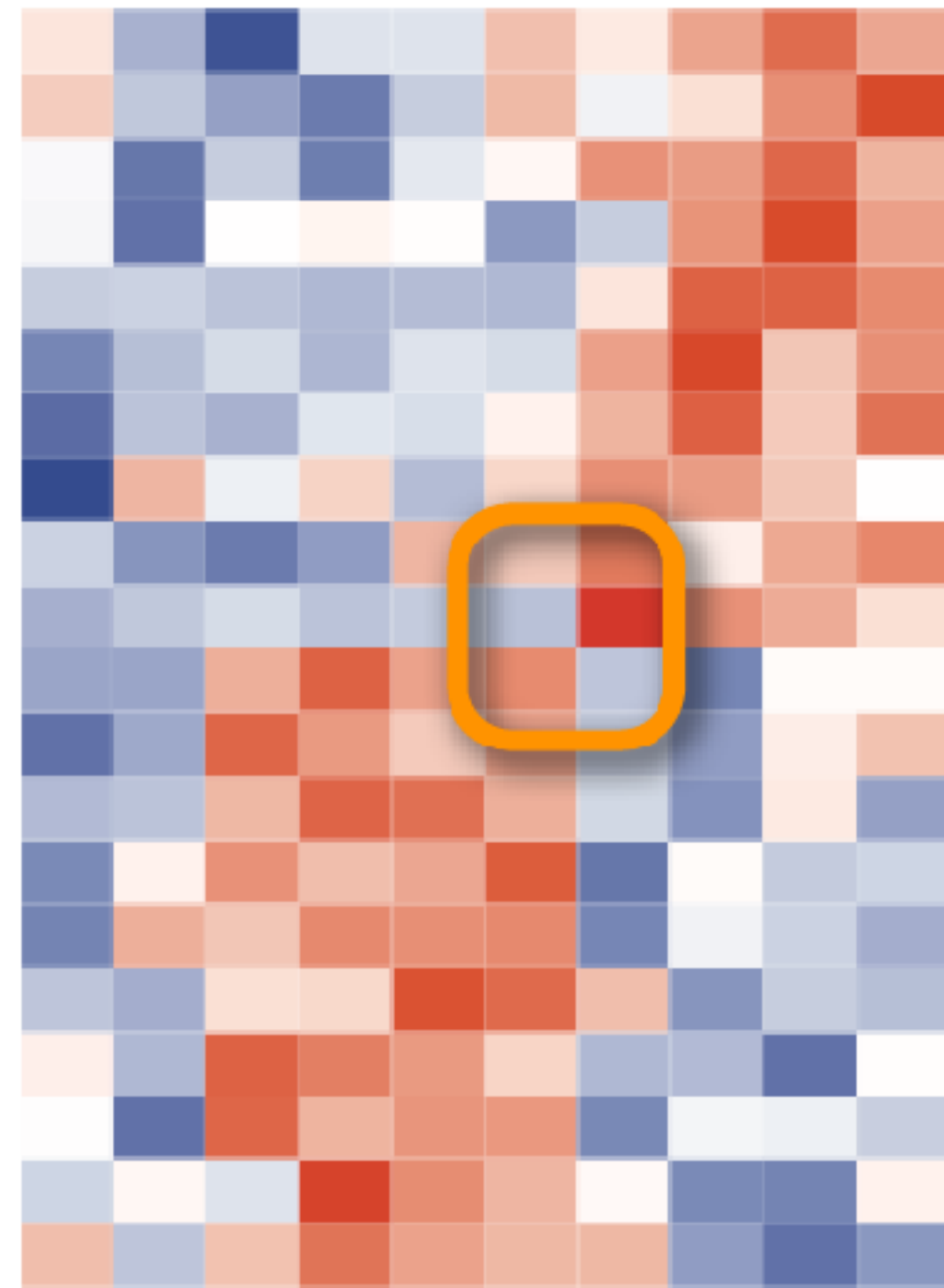


Normal Vision

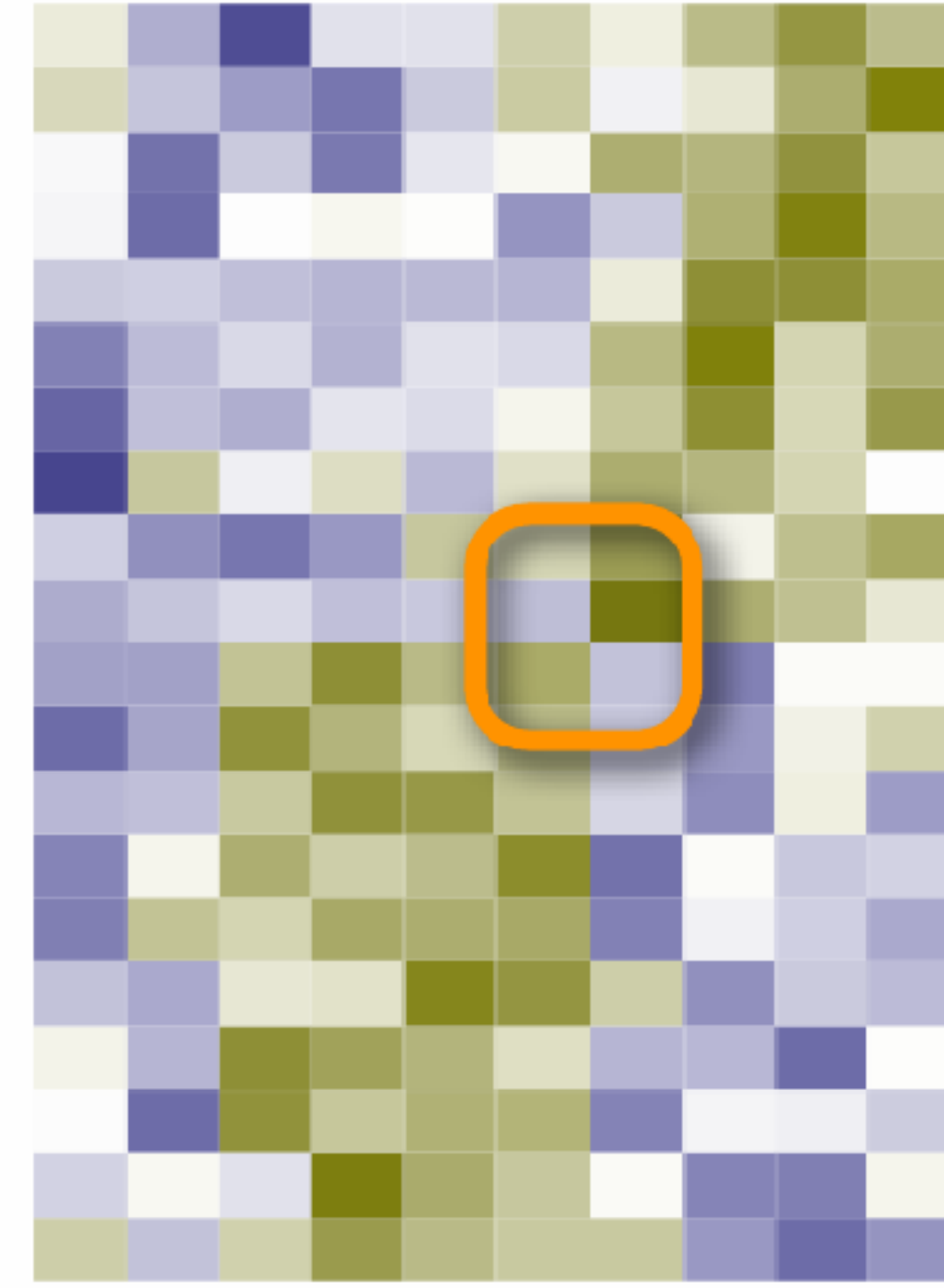


Deuteranope Vision
("Red-Green Blindness")

Good Color Mapping

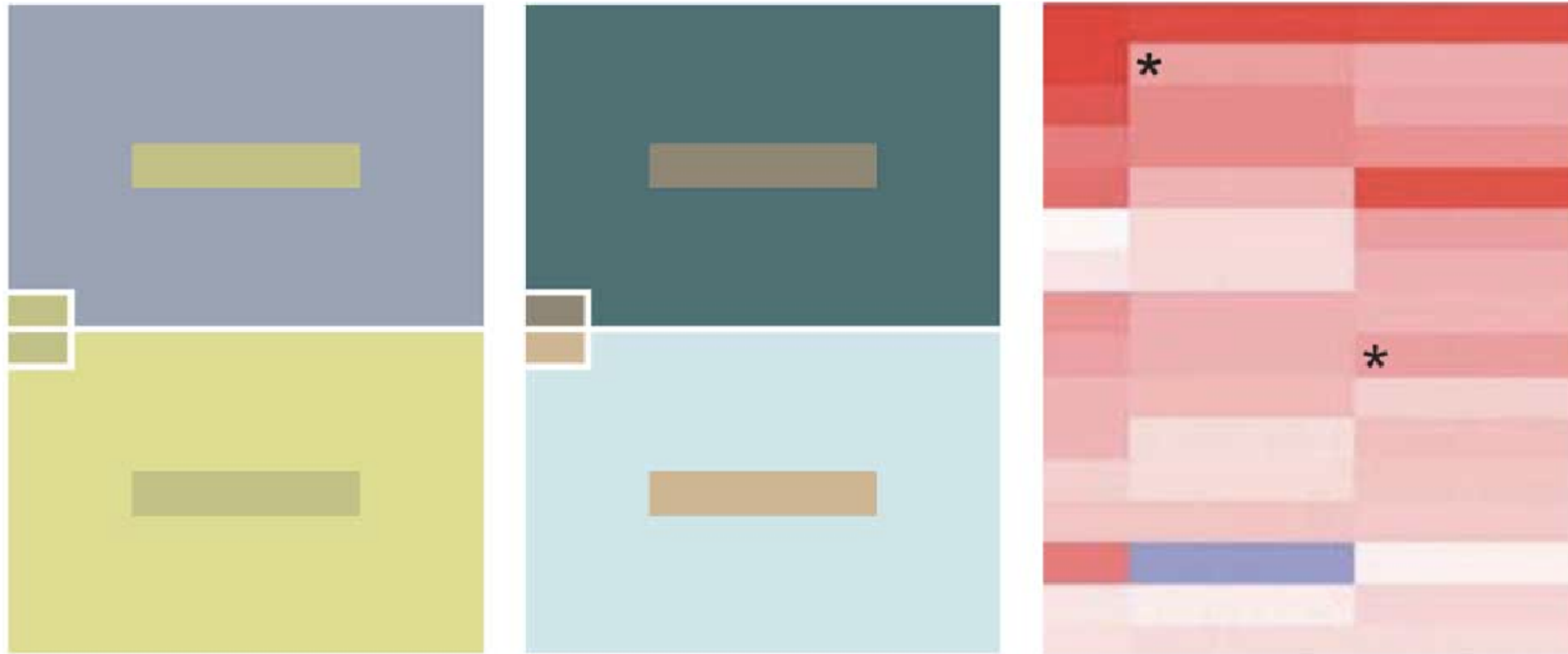


Normal Vision

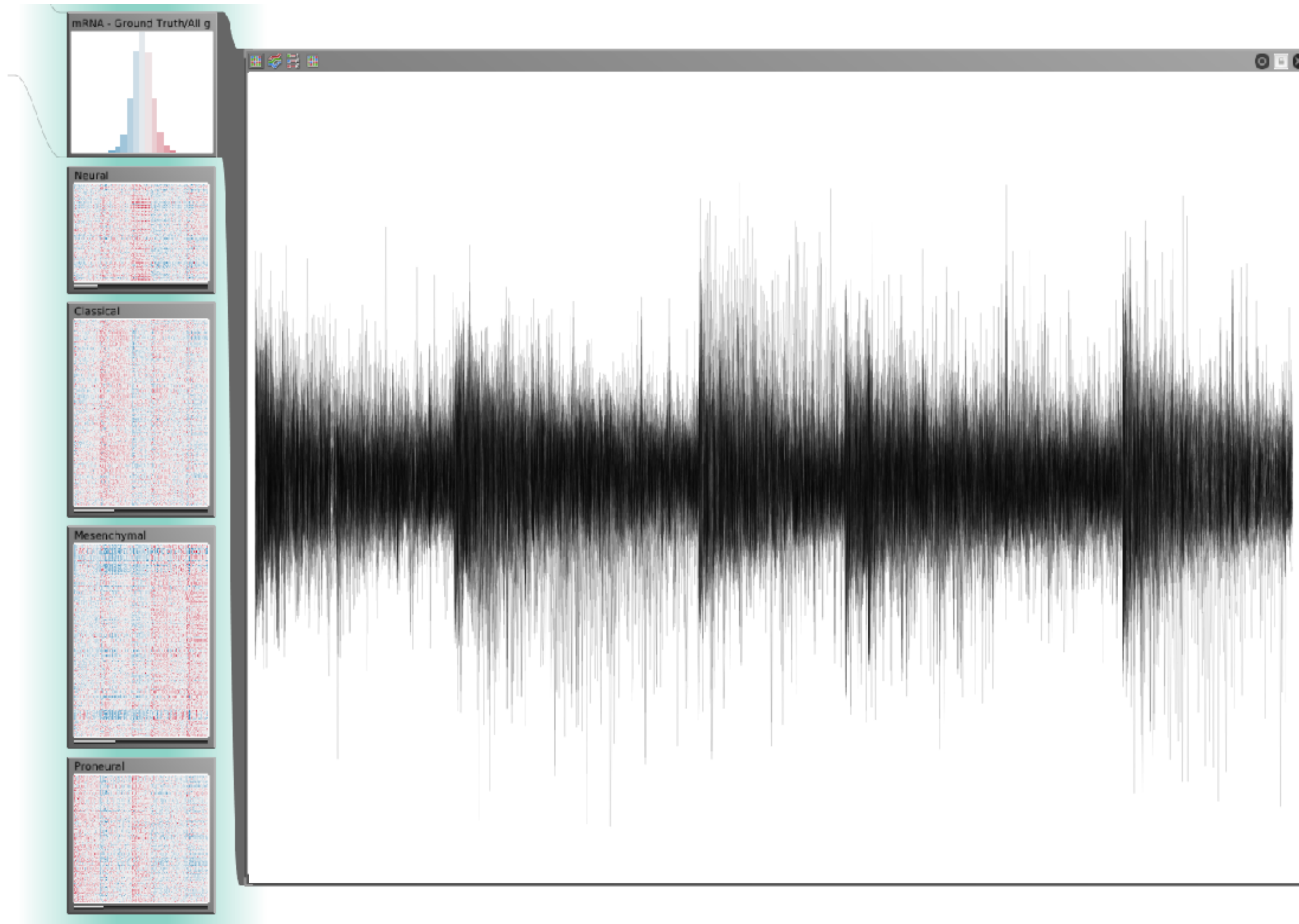


Deuteranope Vision
("Red-Green Blindness")

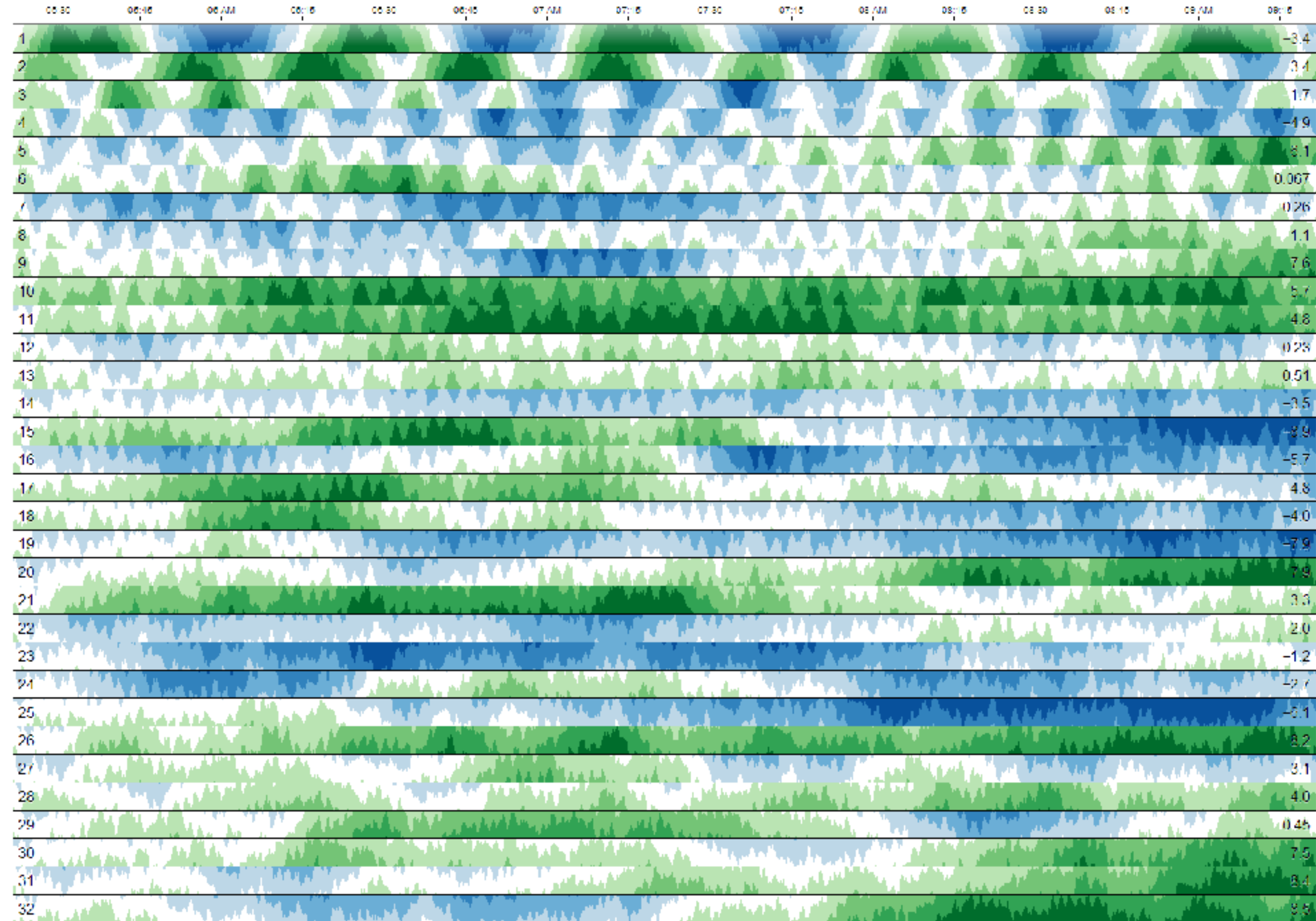
Color is relative!



Clustered Heat Map

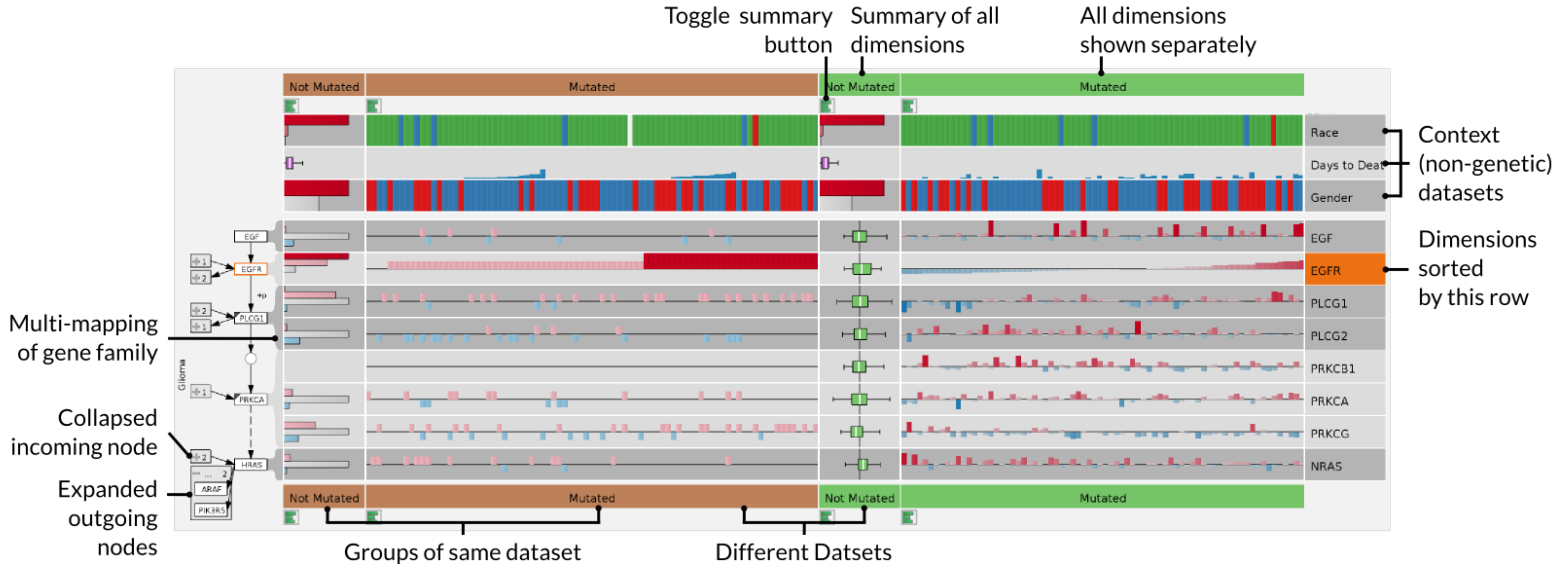


Multiple Line Charts

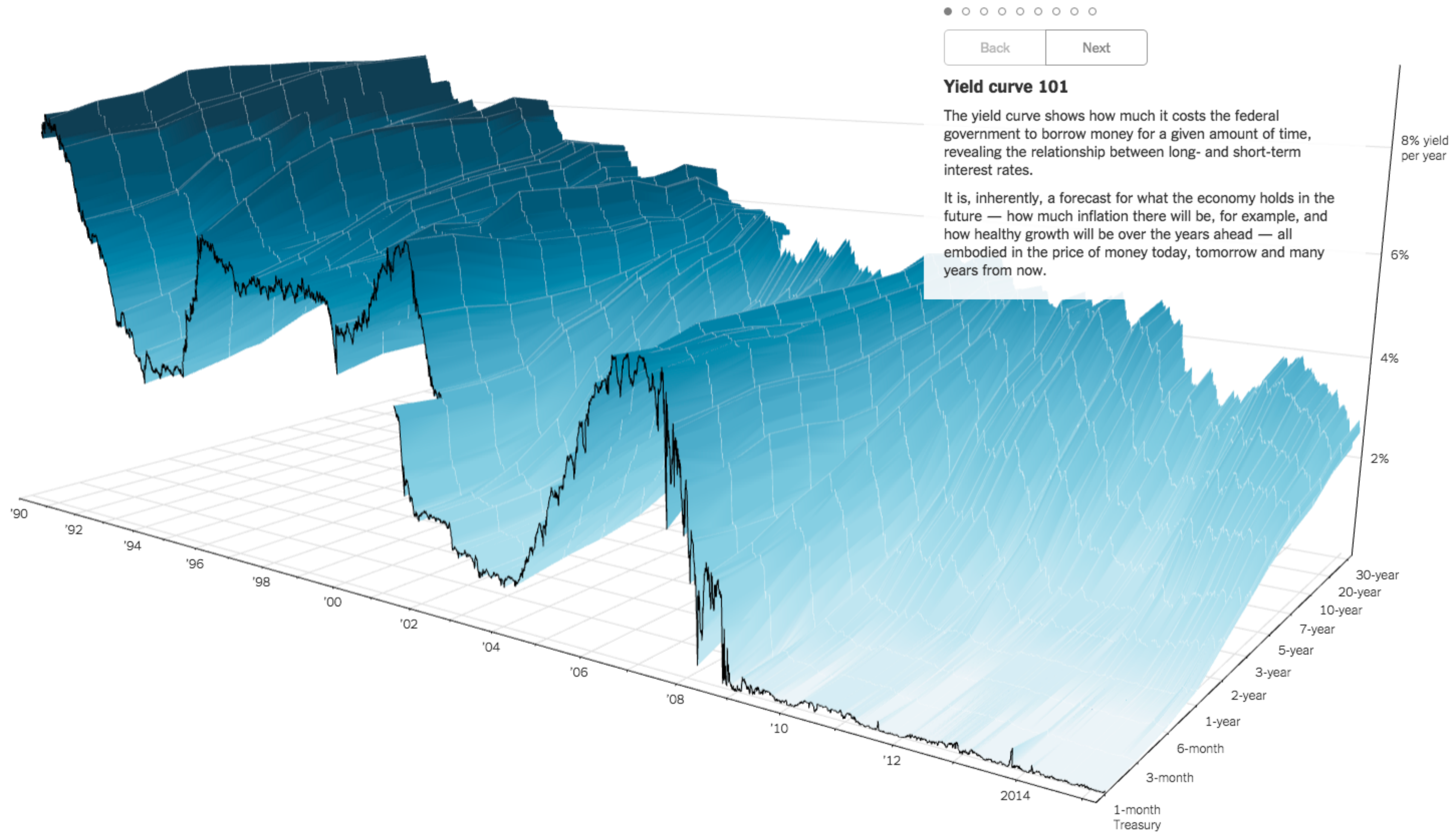


<http://square.github.io/cubism/>

Combining Various Charts



Design Critique



Document: <https://goo.gl/W6w0il>
 Website: <http://goo.gl/D3mlsy>

Spatial Axis Orientation

S

Arrange Tables

⌚ Express Values



⌚ Separate, Order, Align Regions

→ Separate



→ Order



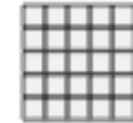
→ Align



→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume



→ Many Keys
Recursive Subdivision



n

⌚ Axis Orientation

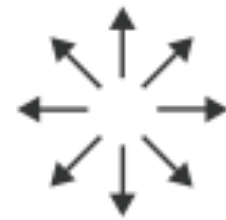
→ Rectilinear



→ Parallel

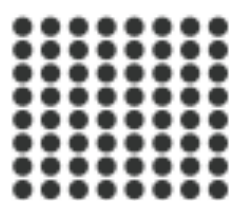


→ Radial



⌚ Layout Density

→ Dense



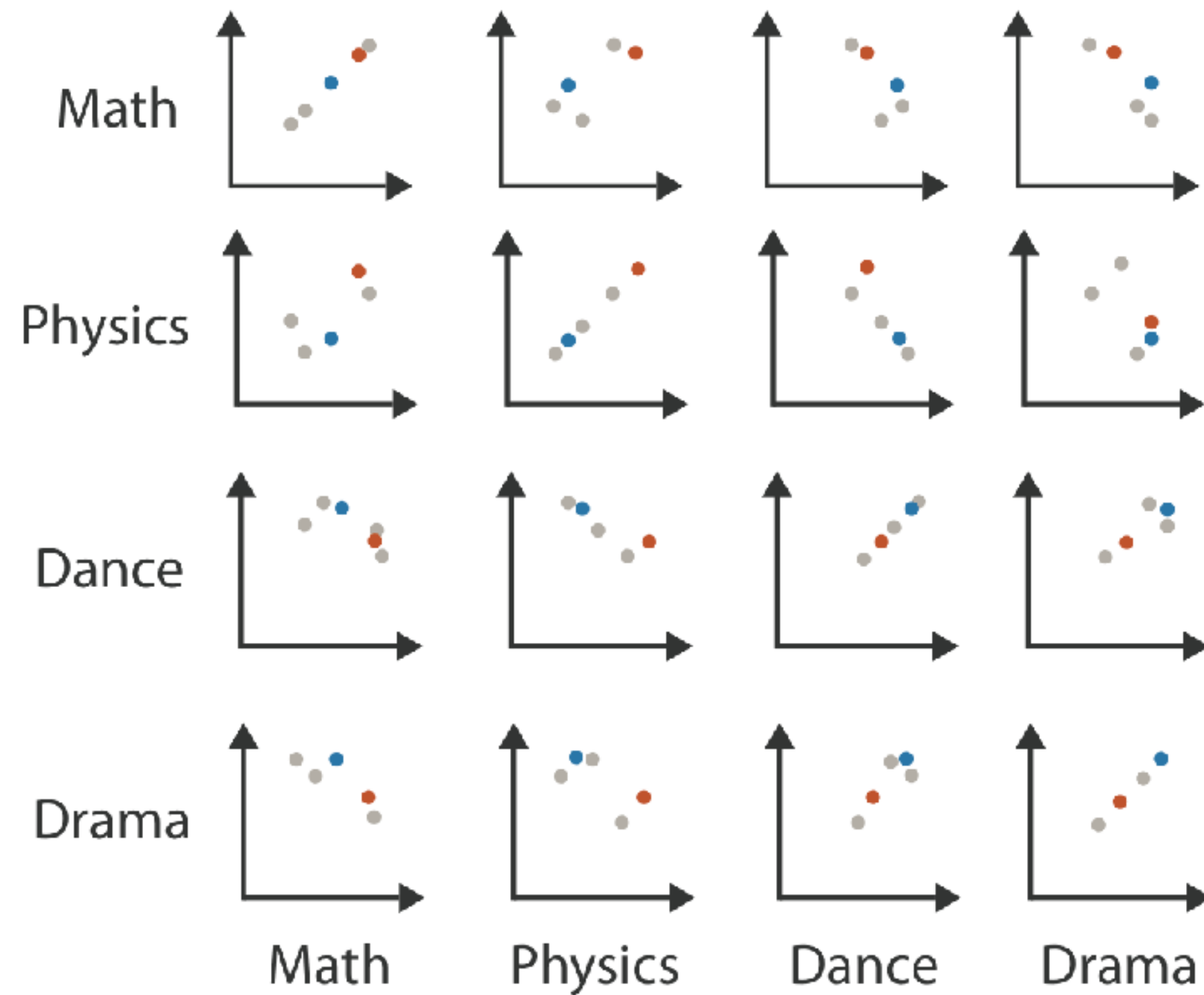
→ Space-Filling



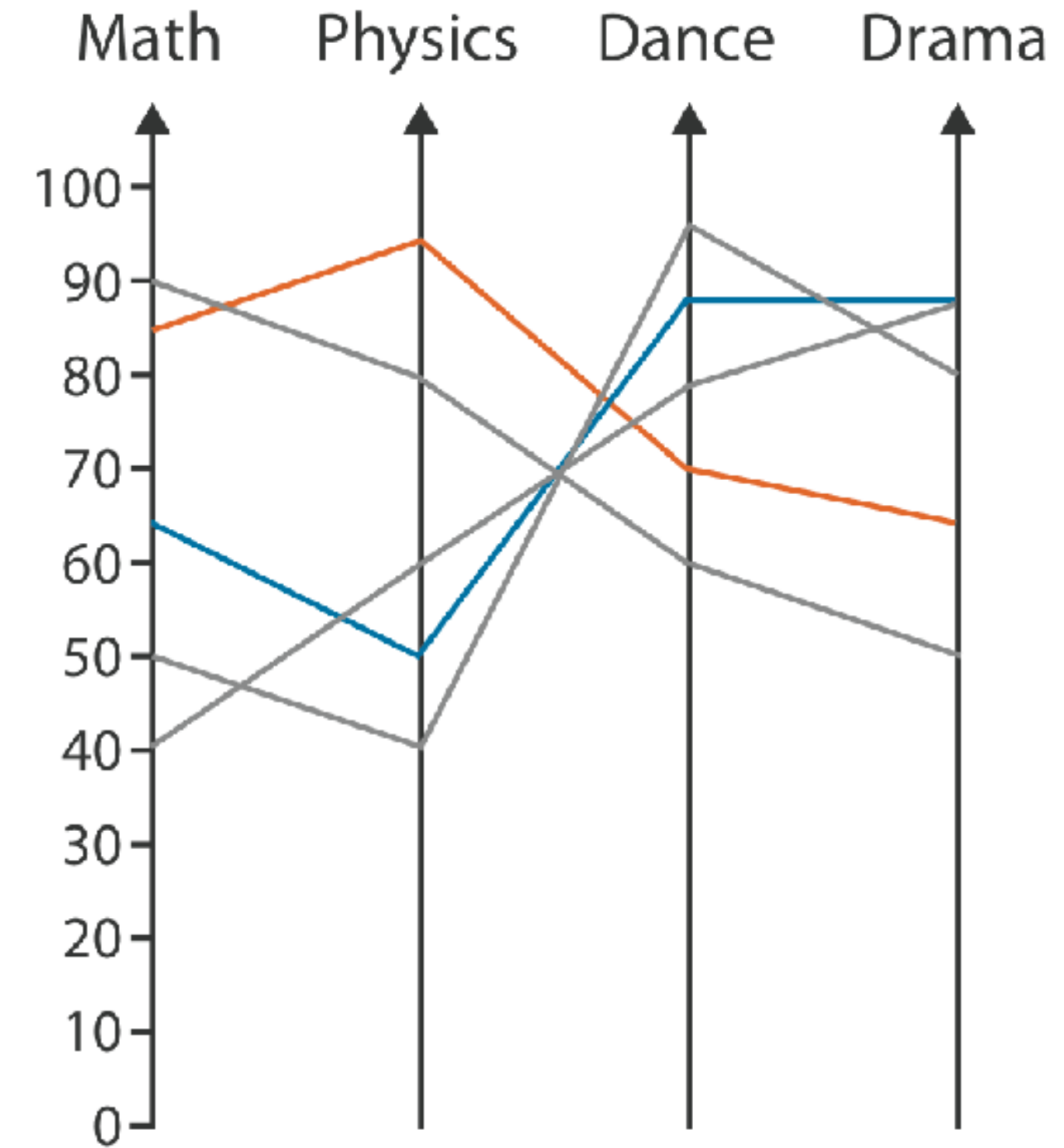
Table

	Math	Physics	Dance	Drama
	85	95	70	65
	90	80	60	50
	65	50	90	90
	50	40	95	80
	40	60	80	90

Scatterplot Matrix



Parallel Coordinates



Spatial Axis Orientation

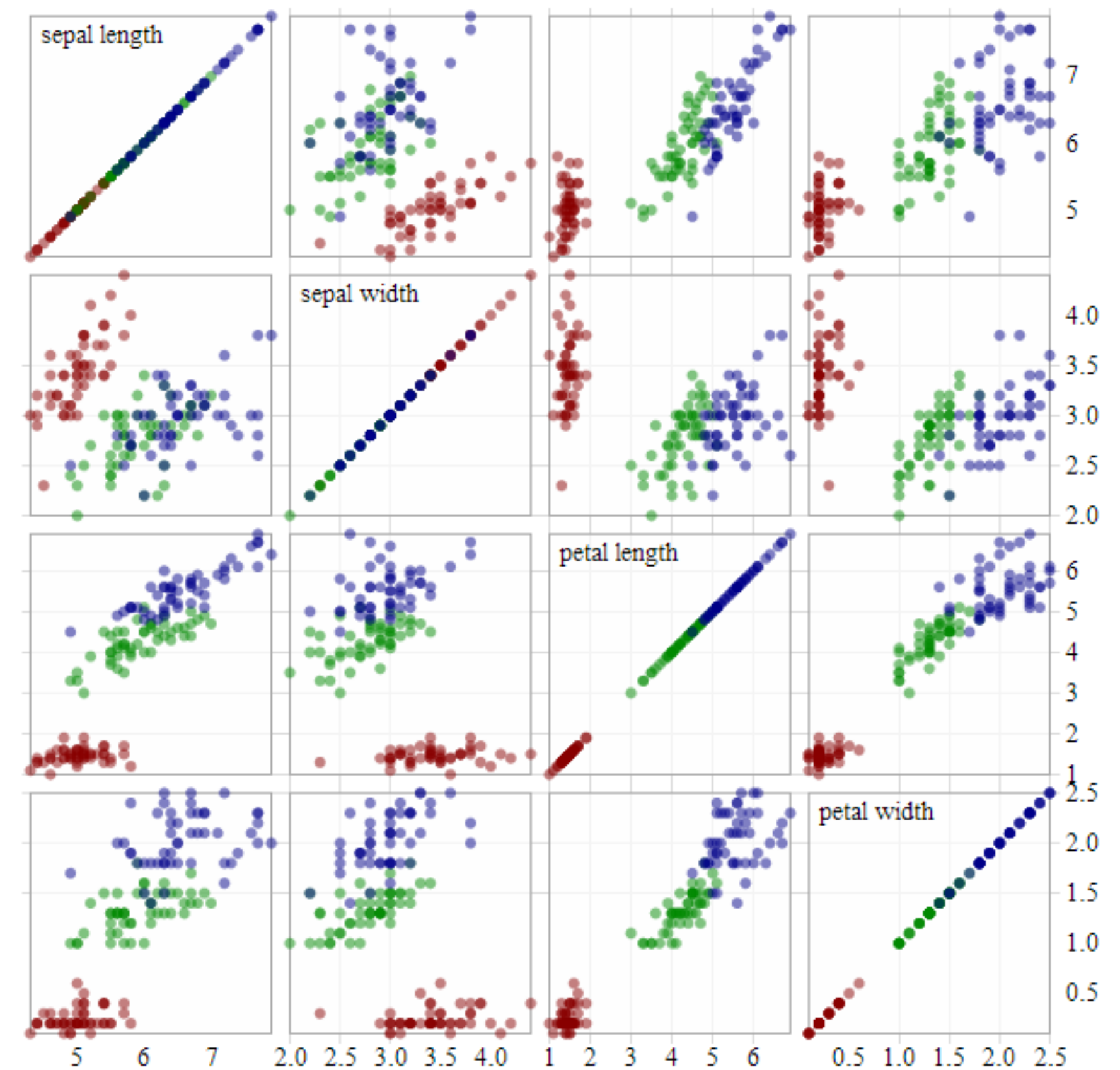
Scatterplot Matrix

Scatterplot Matrices (SPLOM)

Matrix of size $d \times d$

Each row/column is one dimension

Each cell plots a scatterplot of two dimensions



Scatterplot Matrices

Limited scalability (~20 dimensions, ~500-1k records)

Brushing is important

Often combined with “Focus Scatterplot” as F+C technique

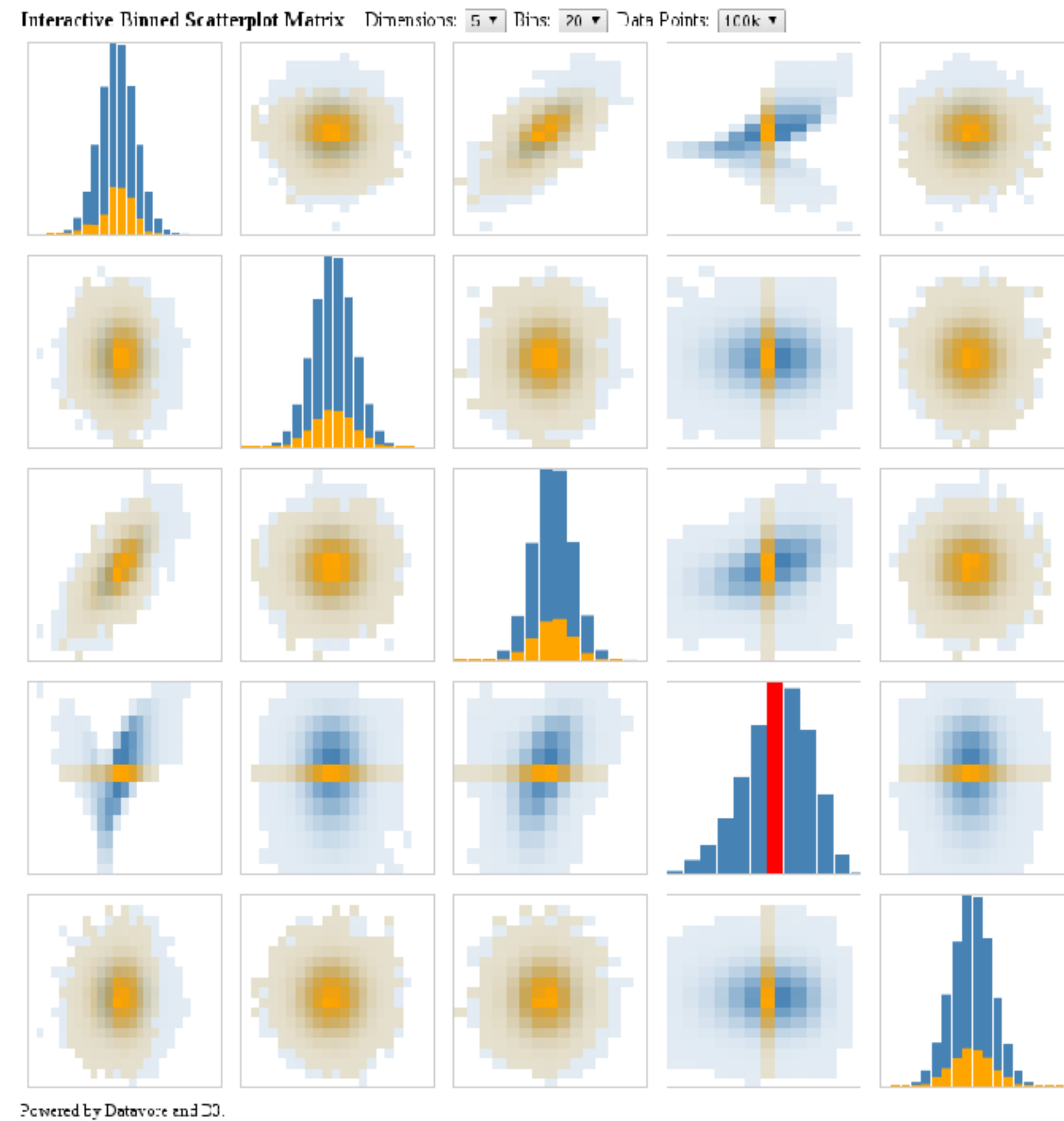
Algorithmic approaches:

Clustering & aggregating records

Choosing dimensions

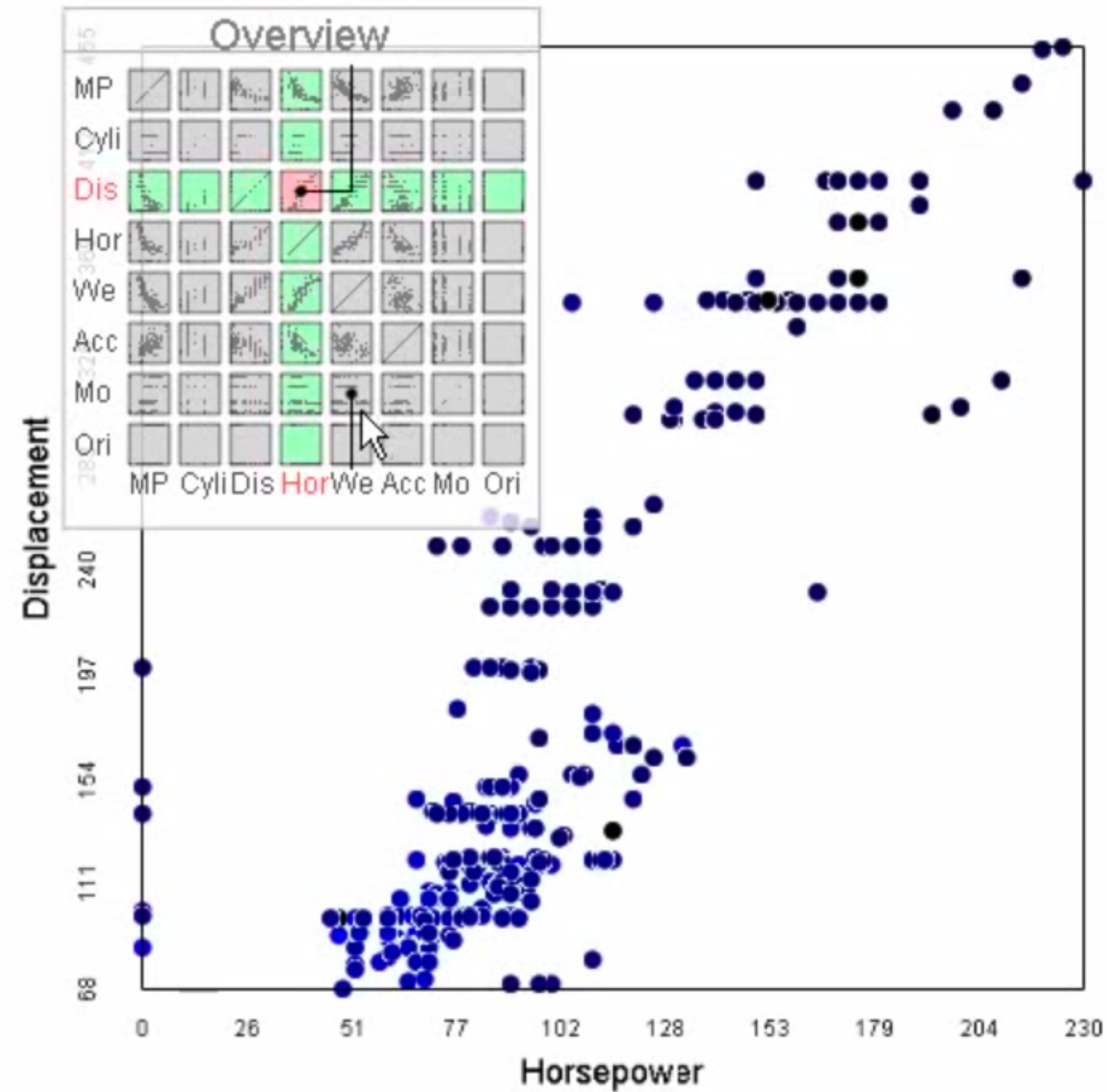
Choosing order

SPLOM Aggregation - Heat Map



Datavore: <http://vis.stanford.edu/projects/datavore/splom/>

SPLOM F+C, Navigation



[Elmqvist]

Spatial Axis Orientation

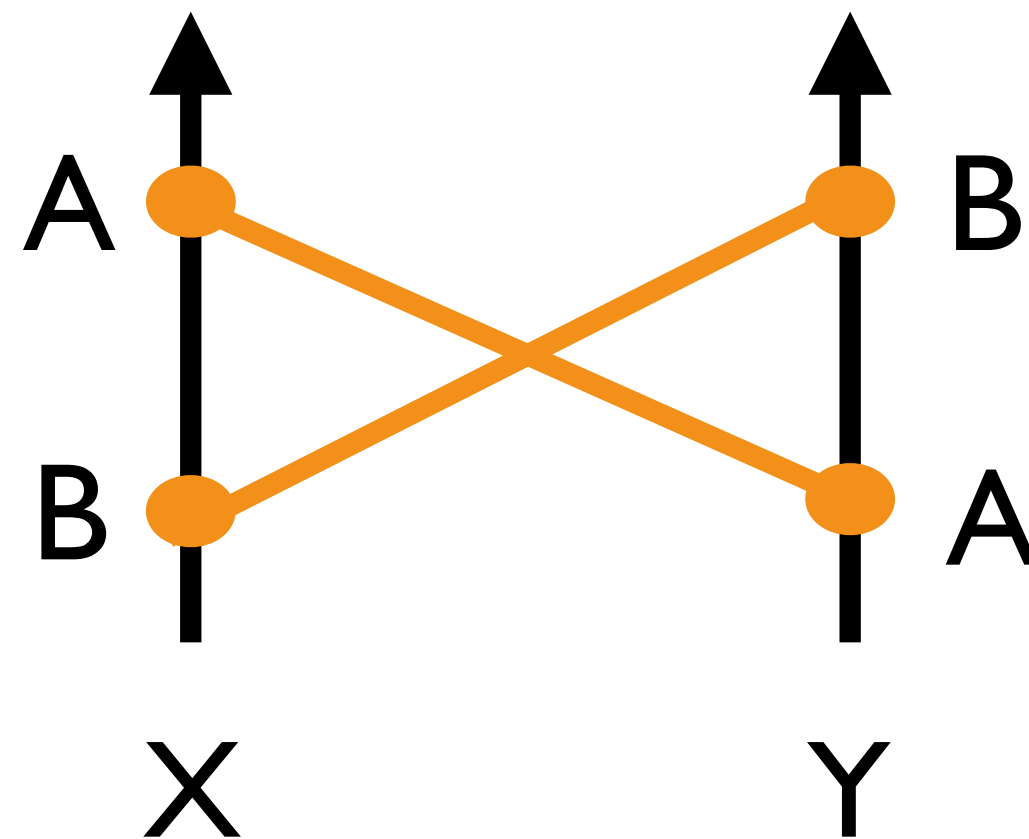
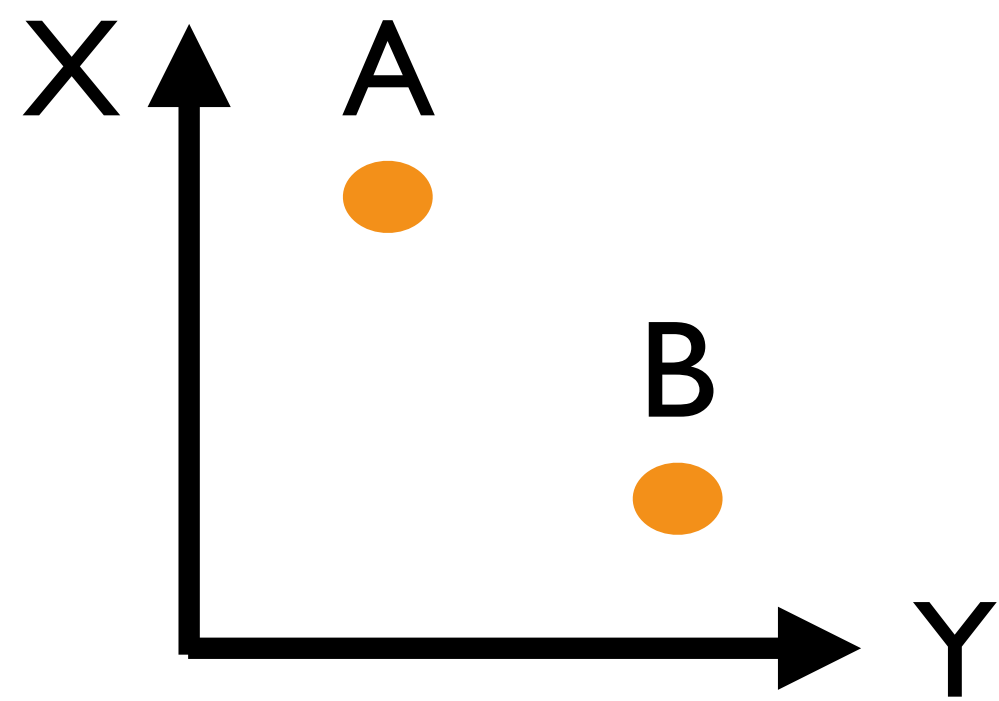
Parallel Coordinates

Parallel Coordinates (PC)

Inselberg 1985

Axes represent attributes

Lines connecting axes represent items



Parallel Coordinates

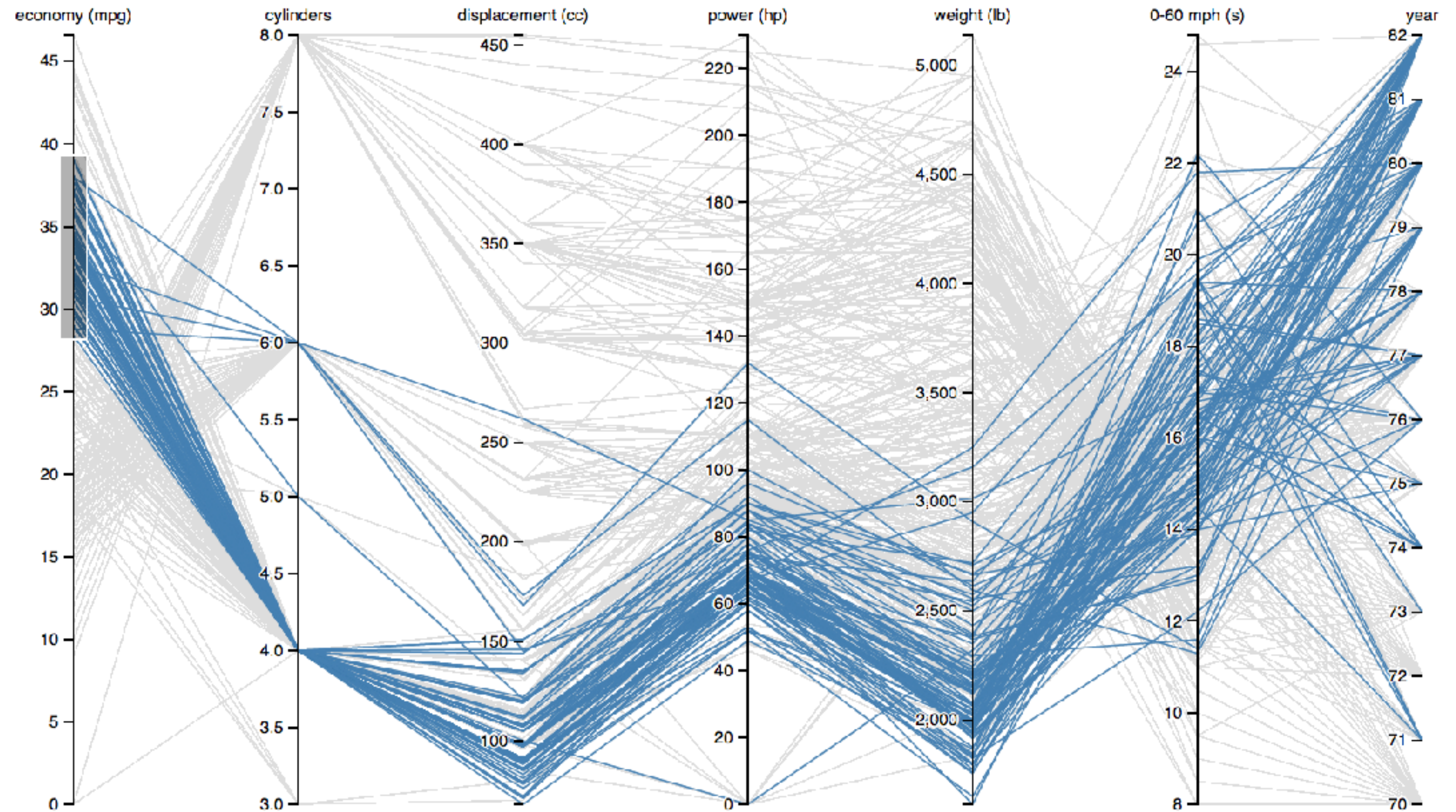
Each axis represents dimension

Lines connecting axis represent records

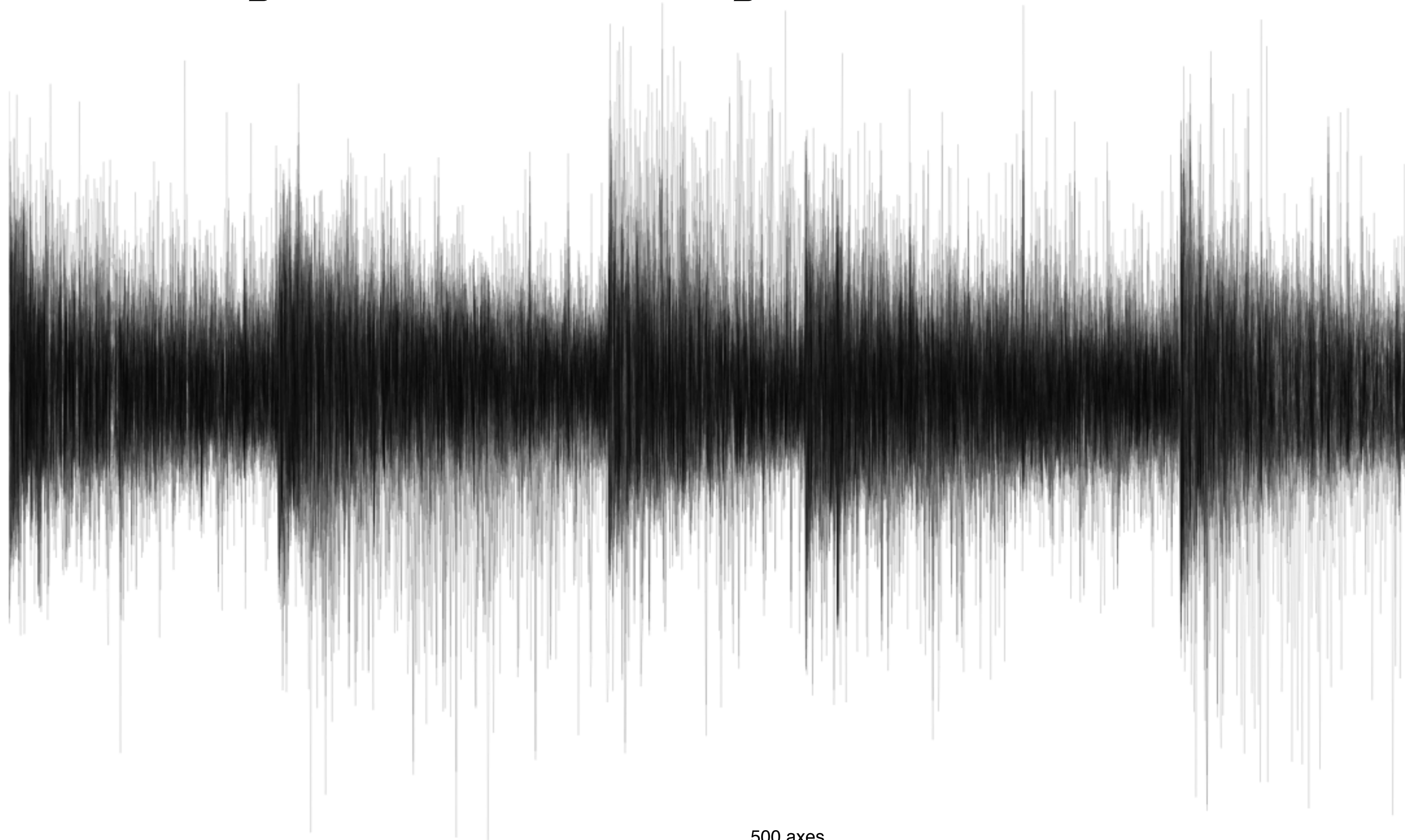
Suitable for

all tabular data types

heterogeneous data



PC Limitation: Scalability to Many Dimensions



500 axes

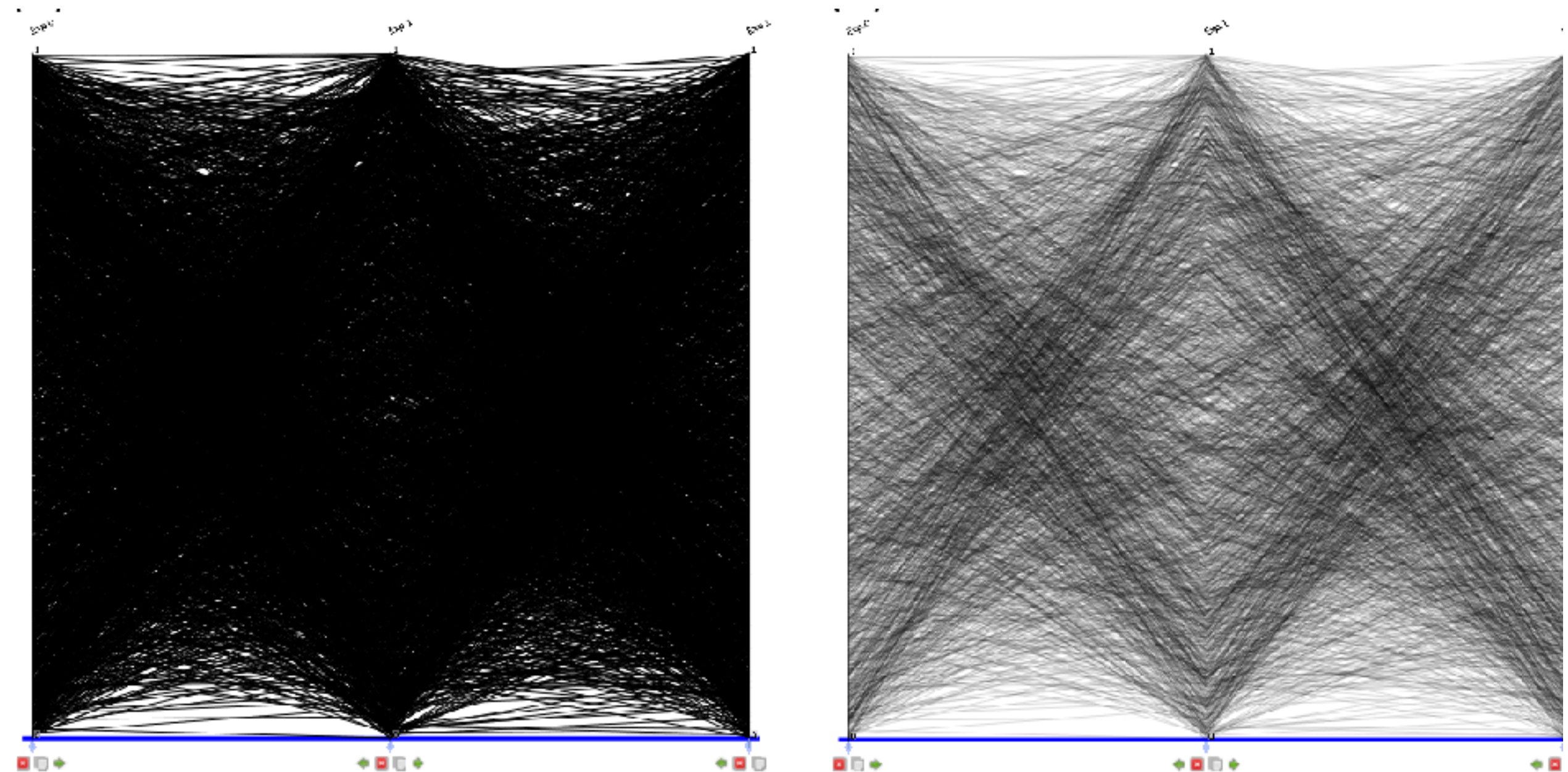
PC Limitation: Scalability to Many Items

Solutions:

Transparency

Bundling, Clustering

Sampling



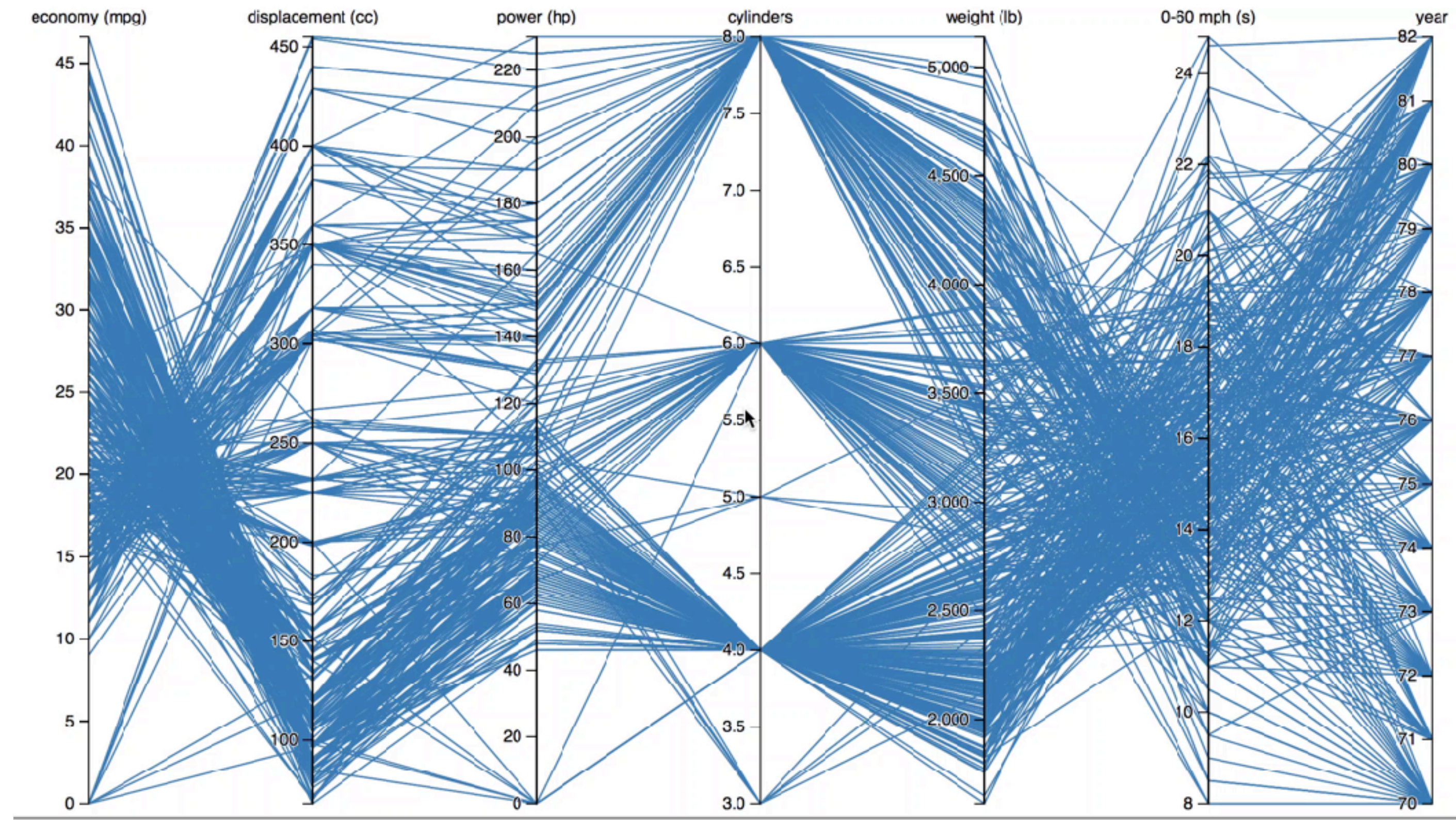
PC Limitations

Correlations only between adjacent axes

Solution: Interaction

Brushing

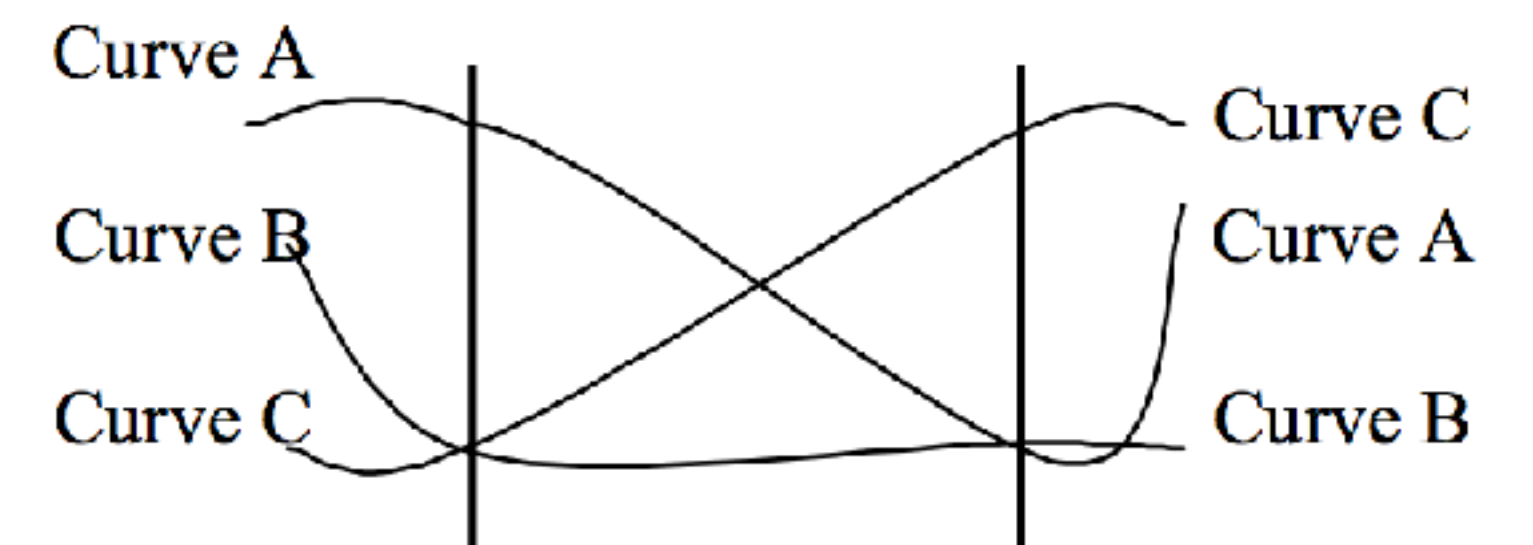
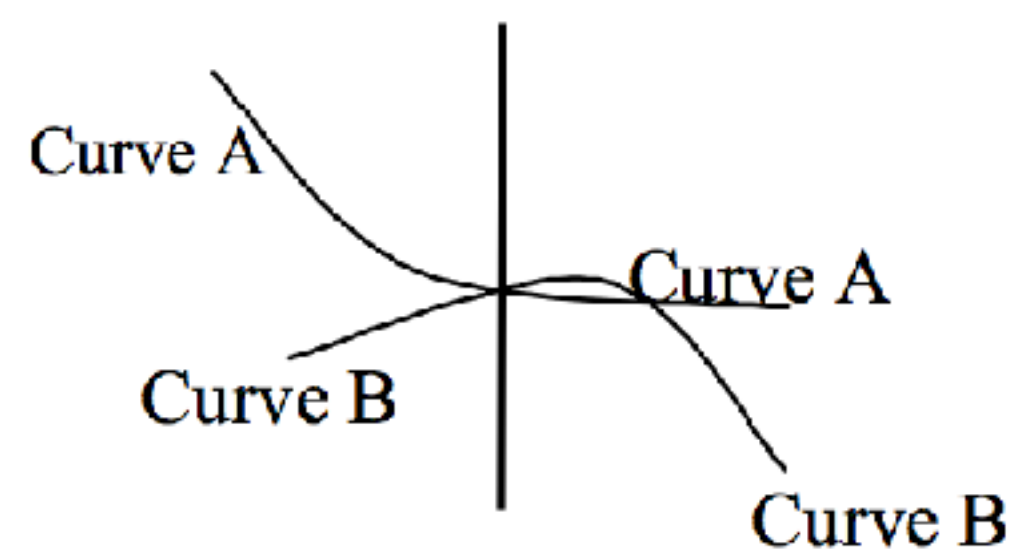
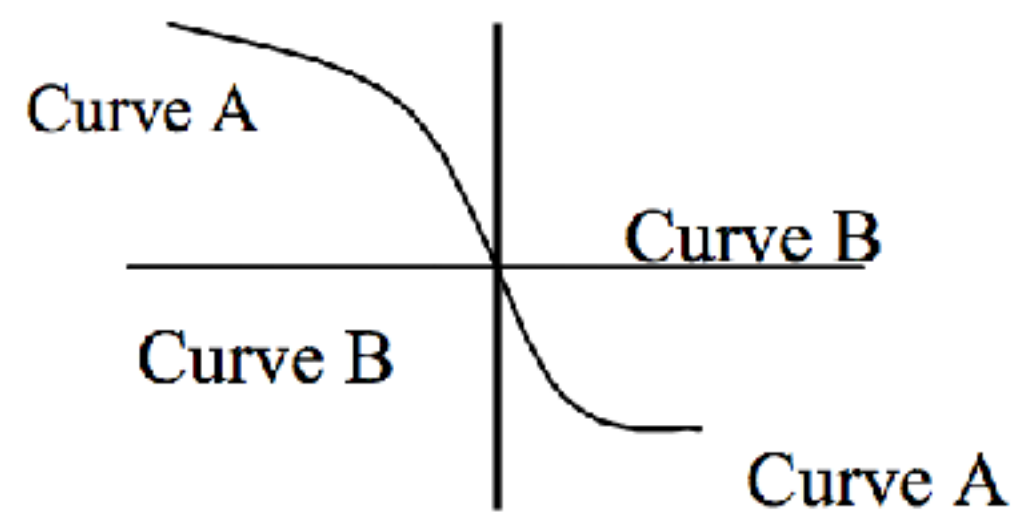
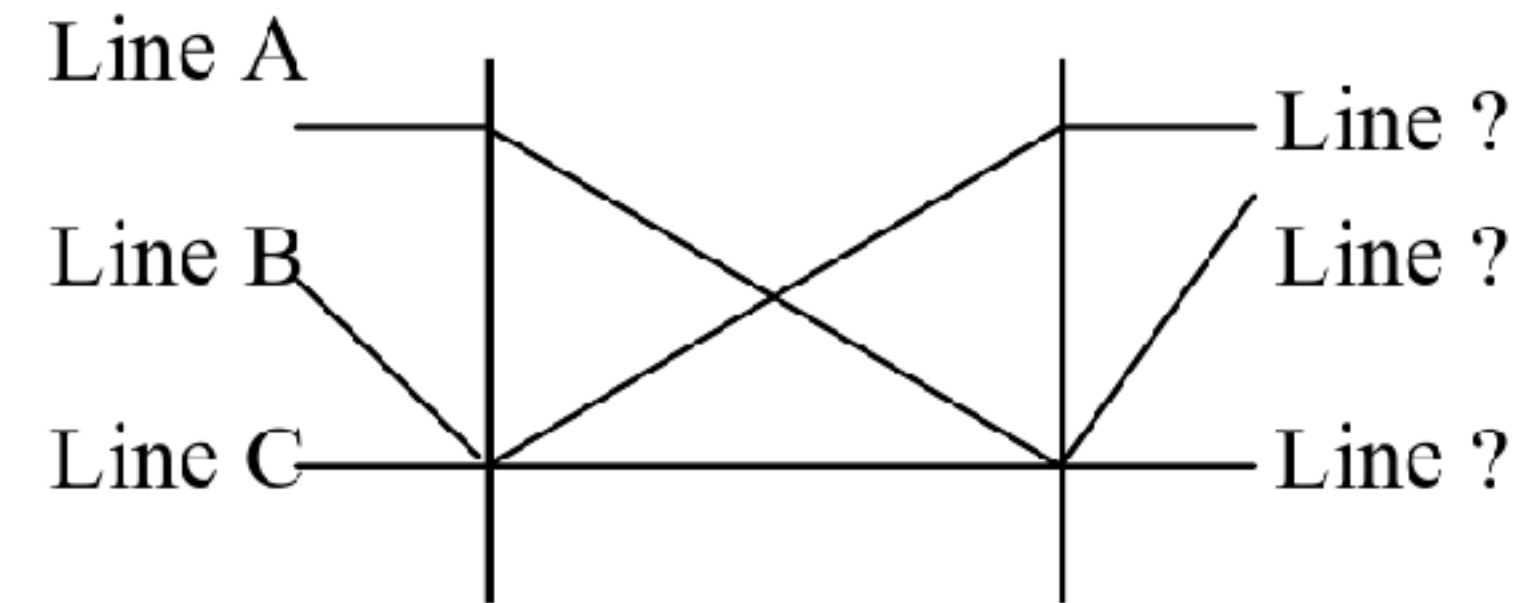
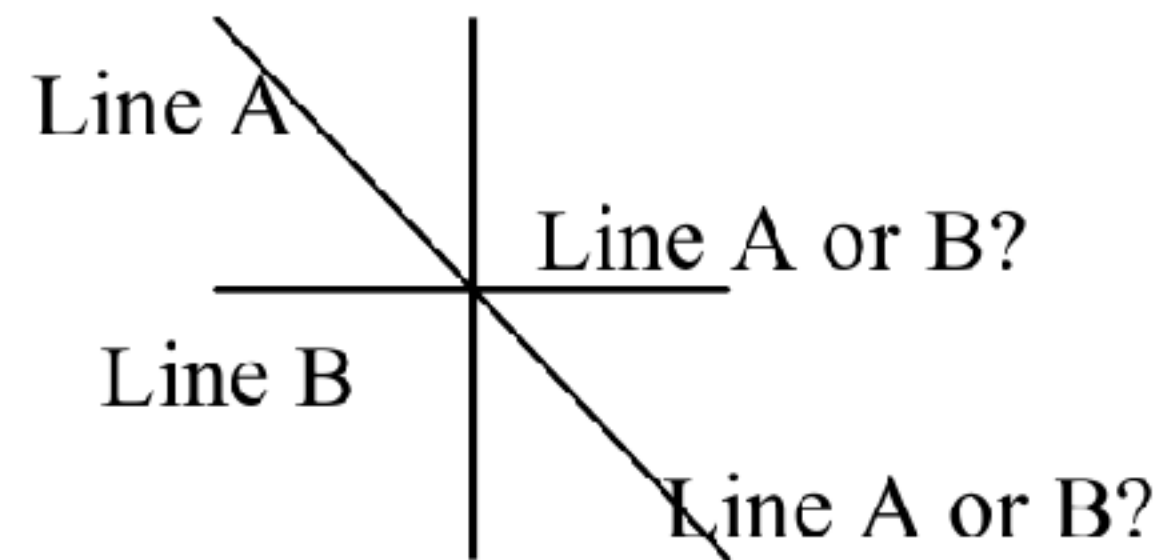
Let user change order



PC Limitation: Ambiguity

Solutions:

Brushing
Curves



Parallel Coordinates

Shows primarily relationships between adjacent axis

Limited scalability (~50 dimensions, ~1-5k records)

Transparency of lines

Interaction is crucial

Axis reordering

Brushing

Filtering

Algorithmic support:

Choosing dimensions

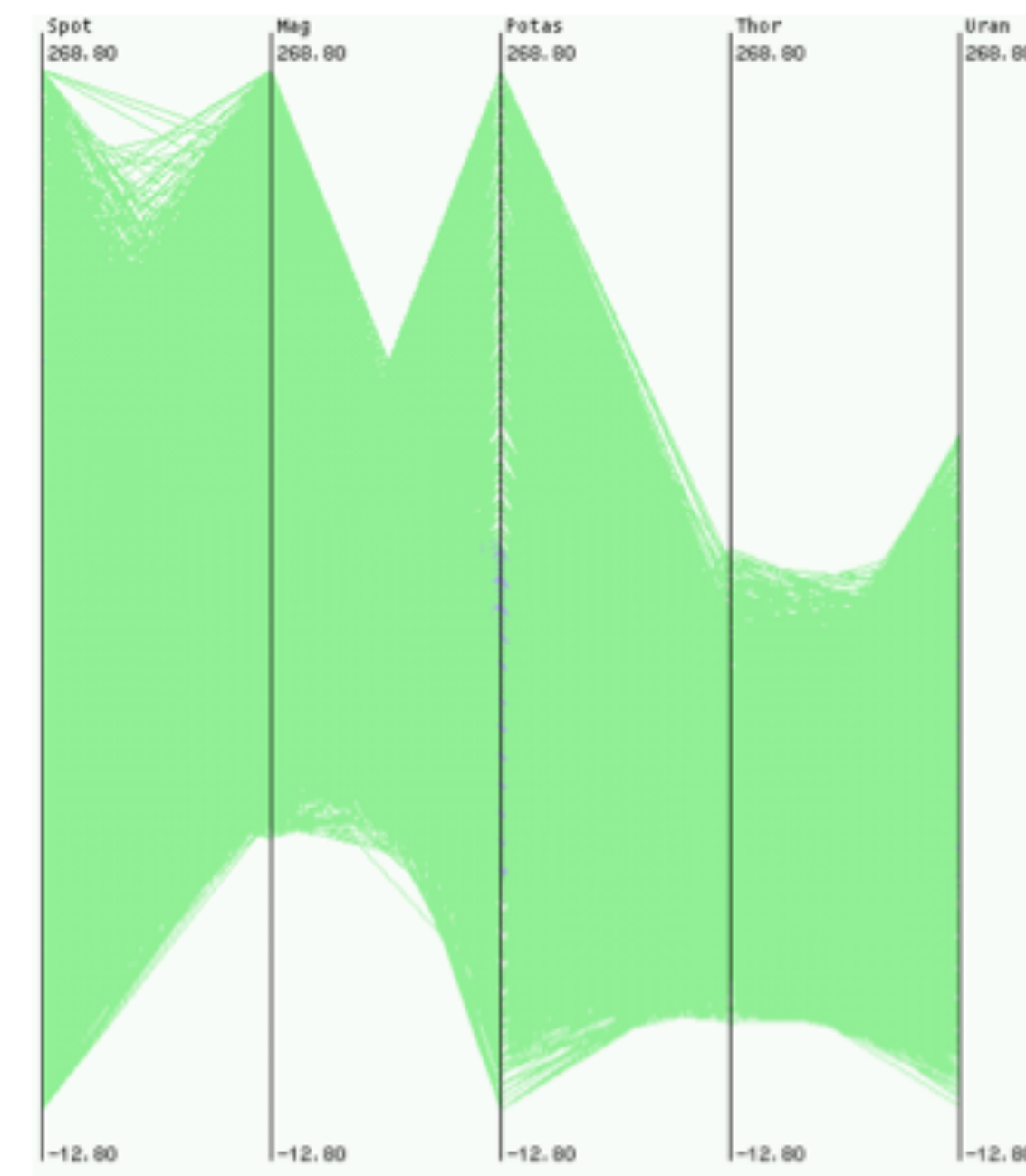
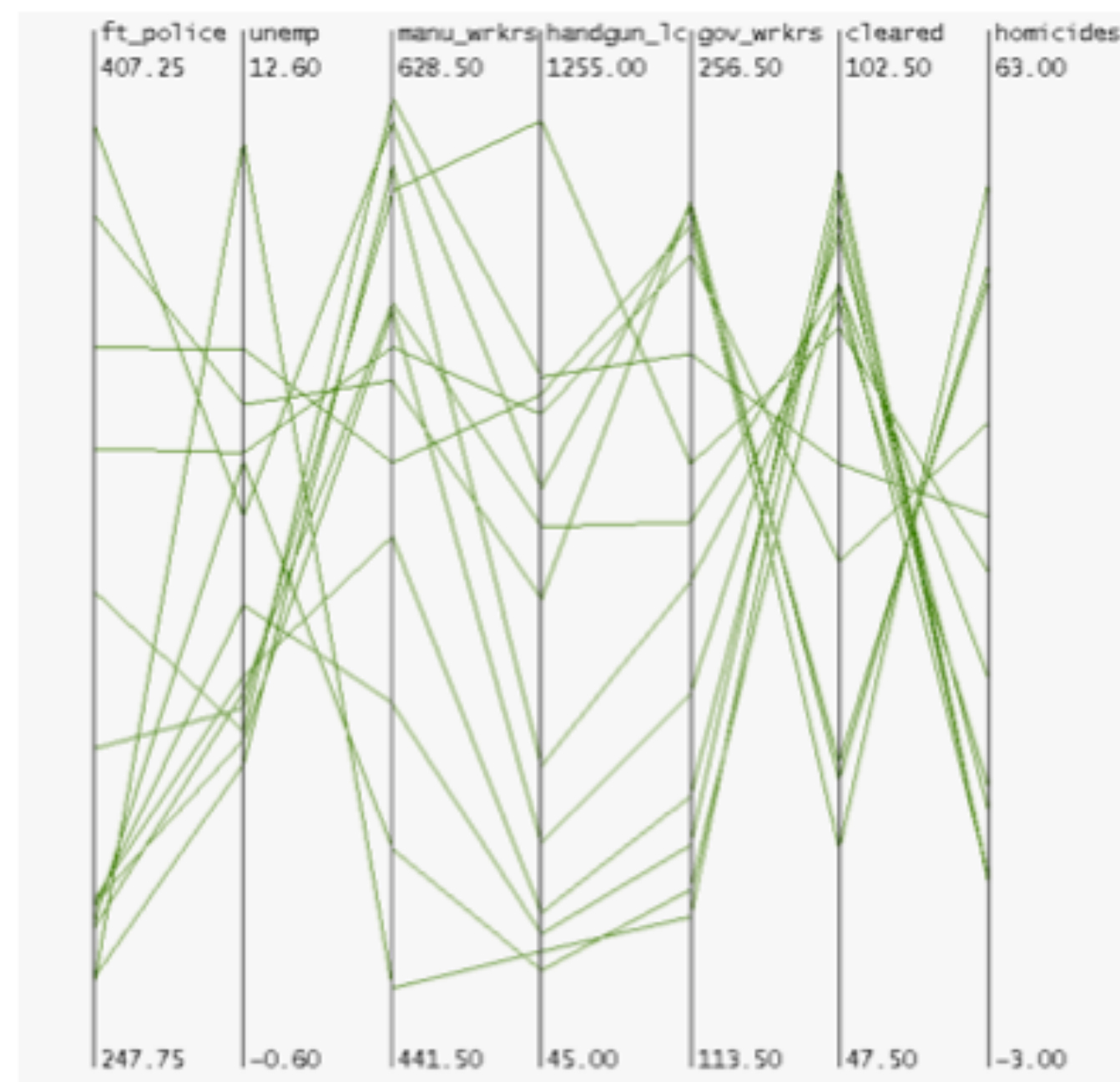
Choosing order

Clustering & aggregating records

HIERARCHICAL PARALLEL COORDINATES

goal: scale up parallel coordinates to large datasets

challenge: overplotting/occlusion



HPC: ENCODING DERIVED DATA

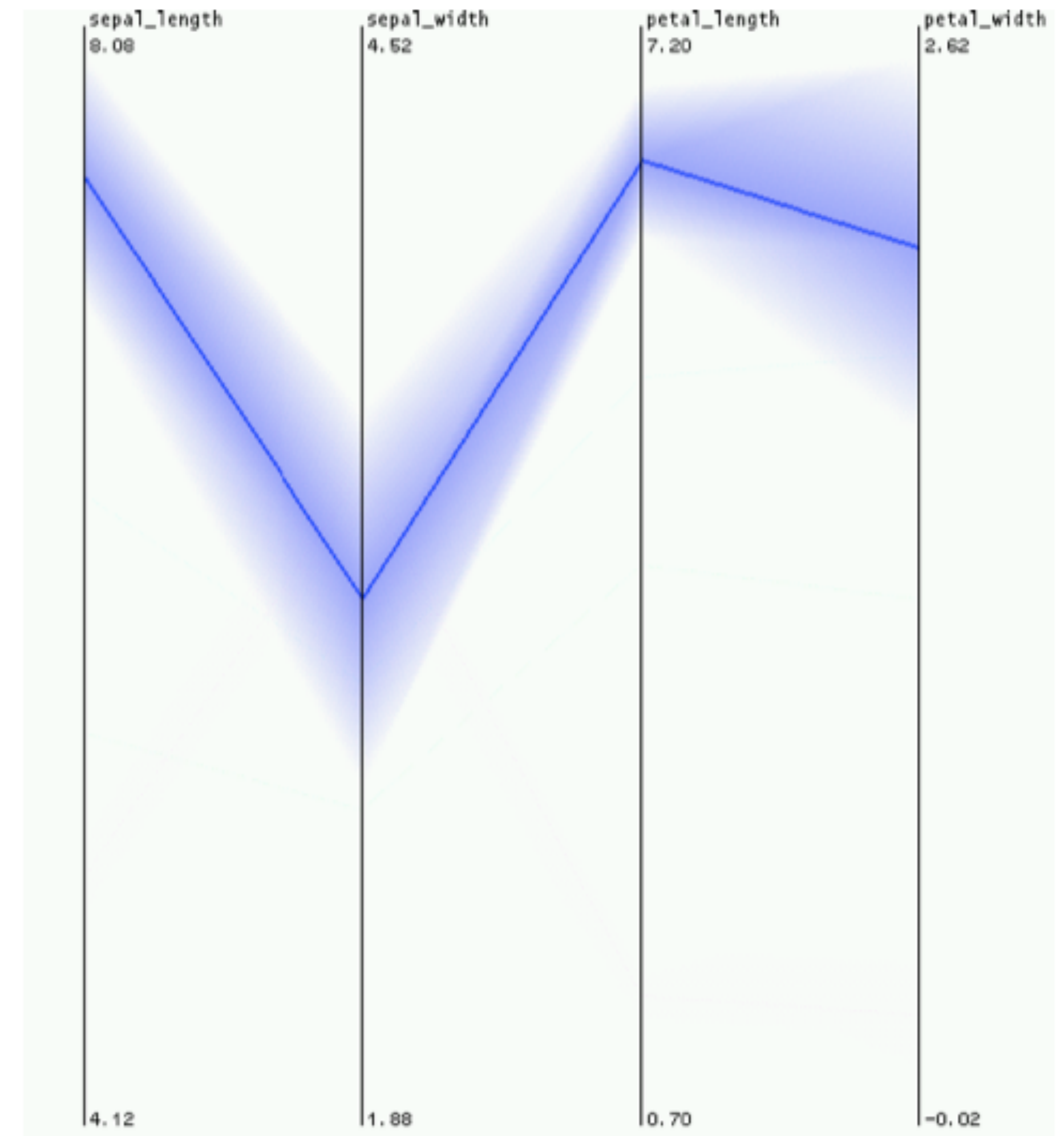
visual representation: variable-width opacity bands

show whole cluster, not just single item

min / max: spatial position

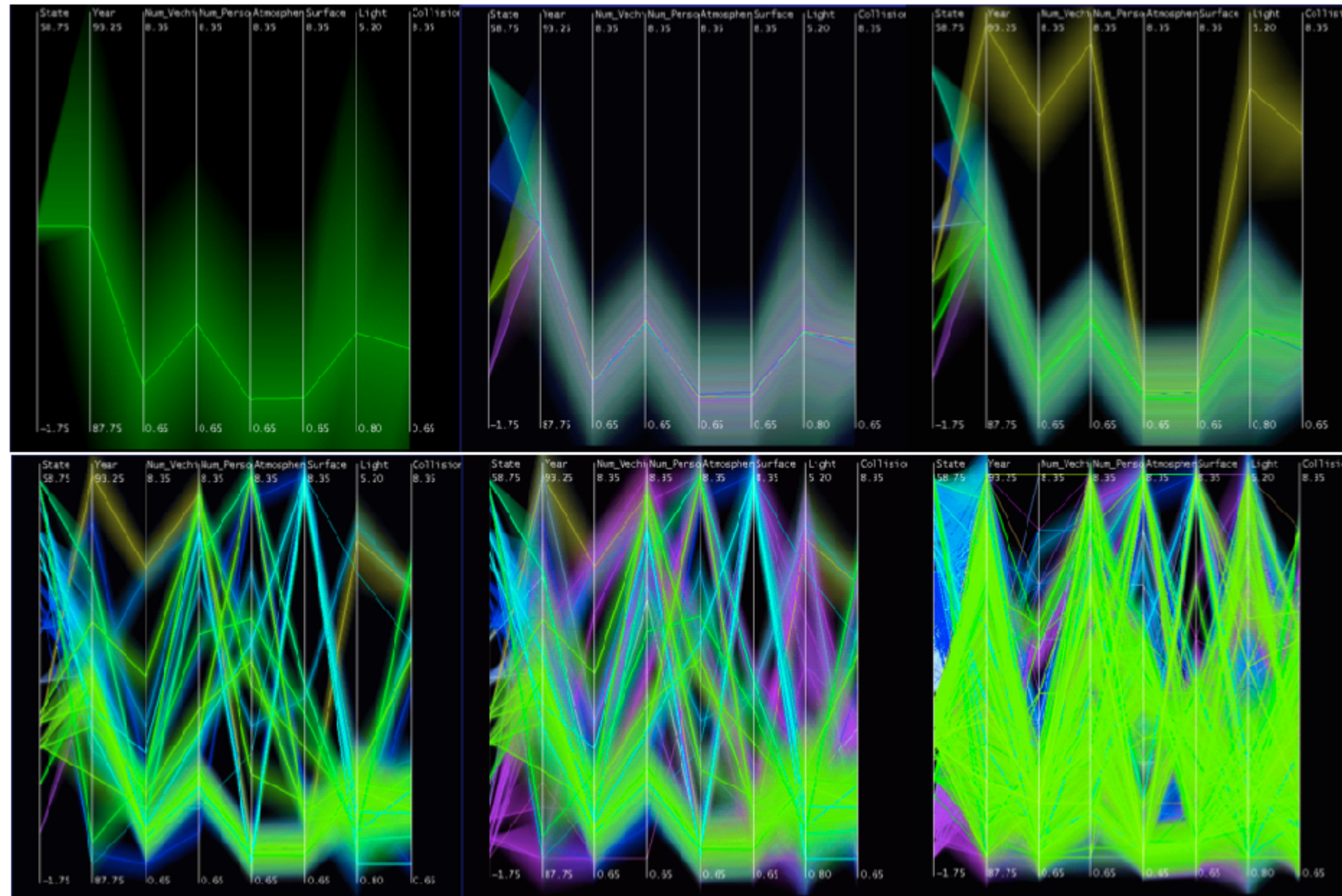
cluster density: transparency

mean: opaque



HPC: INTERACTING WITH DERIVED DATA

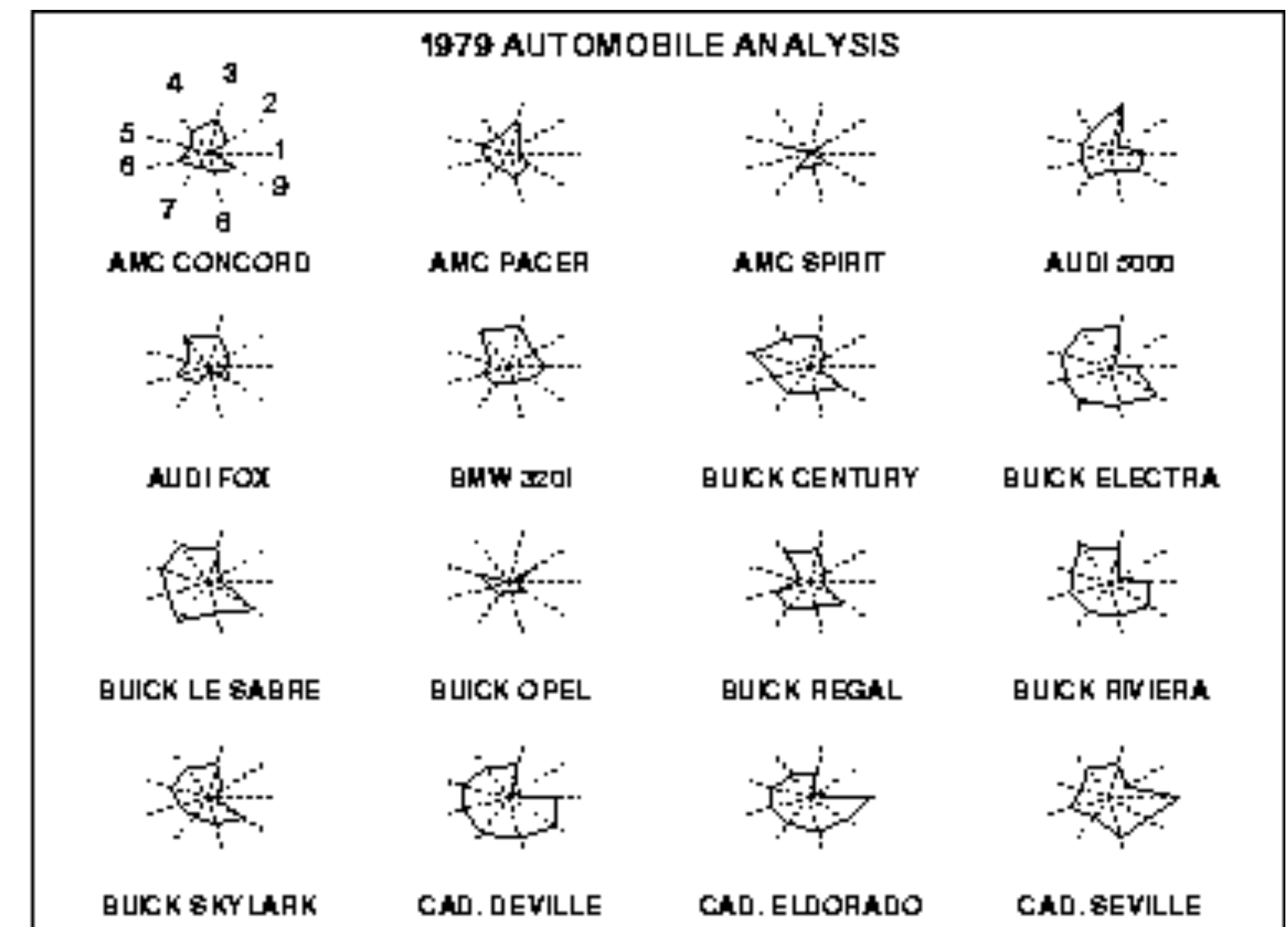
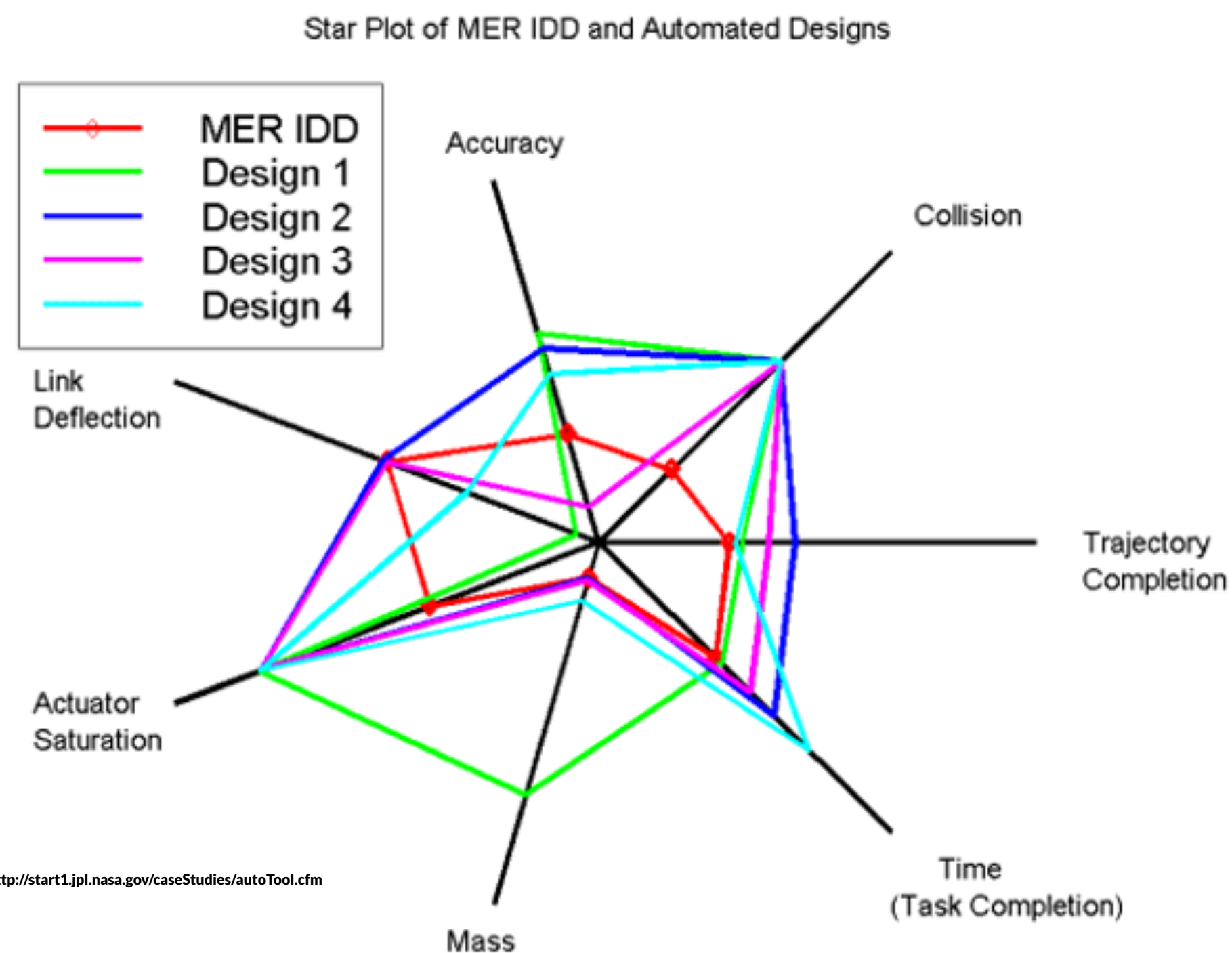
interactively change level of detail to navigate cluster hierarchy



Star Plot

[Coekin1969]

Similar to parallel coordinates
Radiate from a common origin



<http://blocks.org/kevinschau/raw/8833989/>

Data Reduction

Sampling

Don't show every element, show a (random) subset

Efficient for large dataset

Apply only for display purposes

Outlier-preserving approaches

Filtering

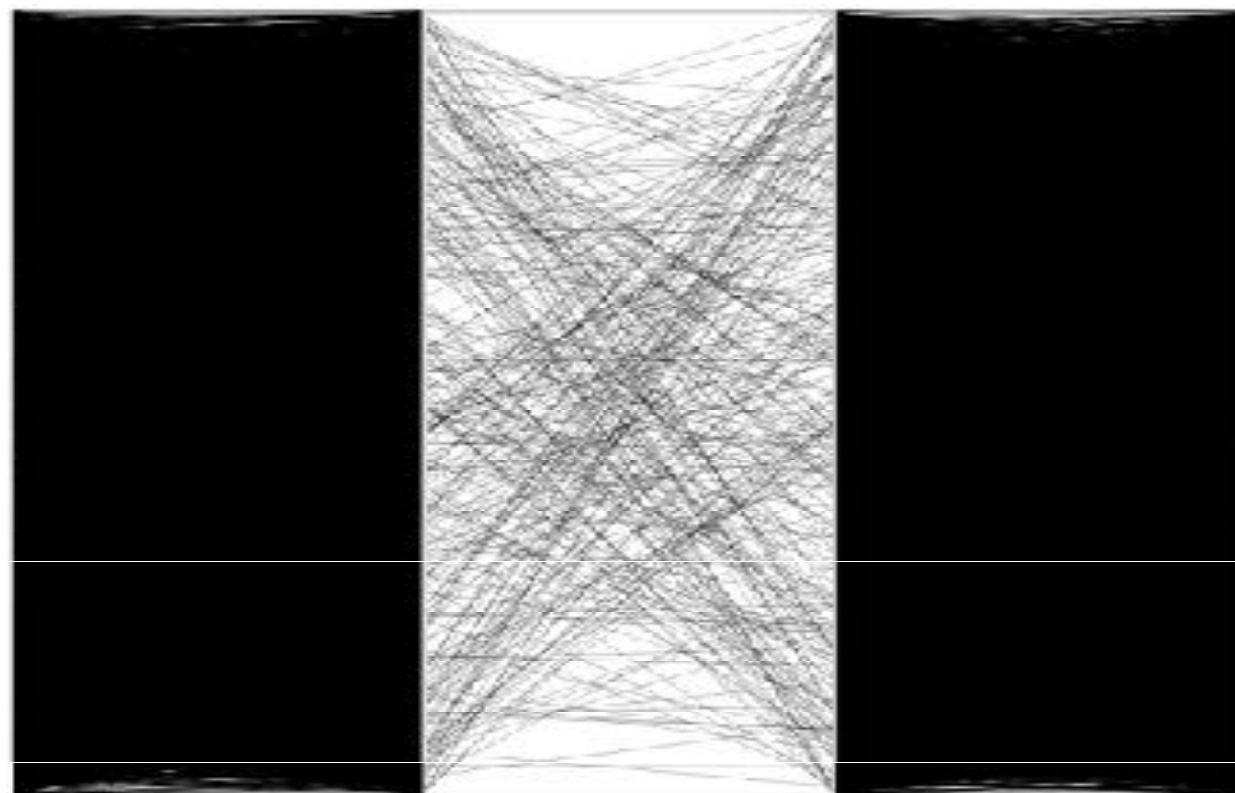
Define criteria to remove data, e.g.,

minimum variability

> / < / = specific value for one dimension

consistency in replicates, ...

Can be interactive, combined with sampling

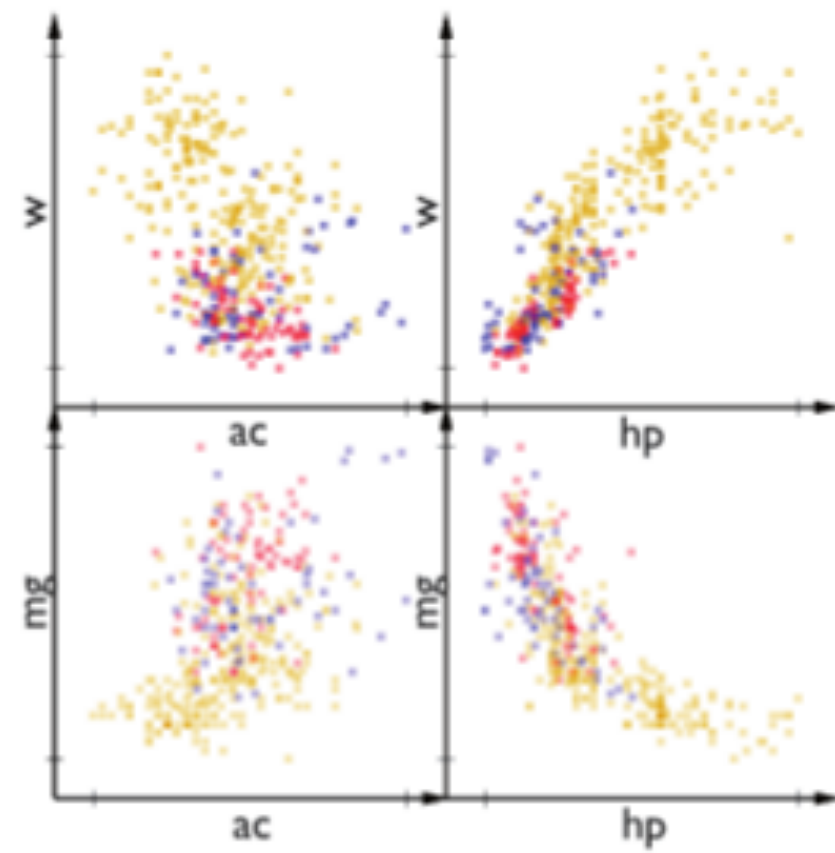


[Ellis & Dix, 2006]

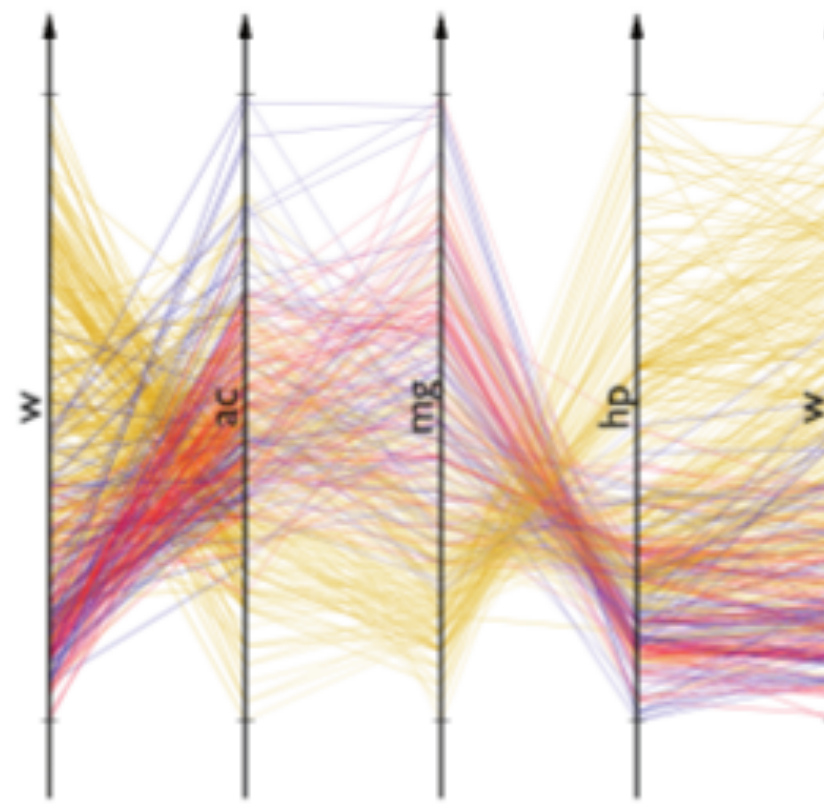
Spatial Axis Orientation

Hybrids

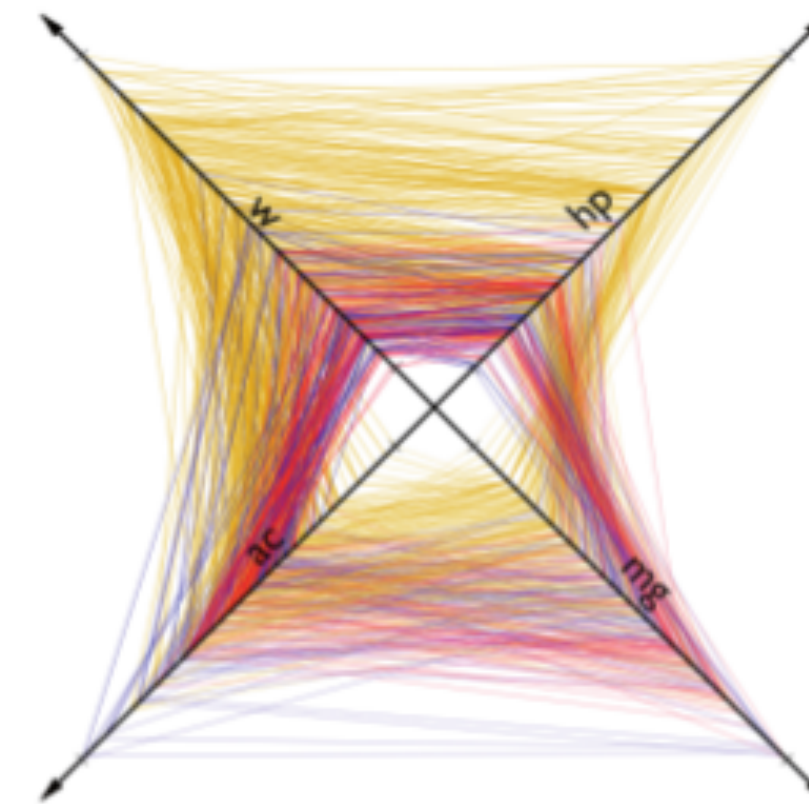
Flexible Linked Axes (FLINA)



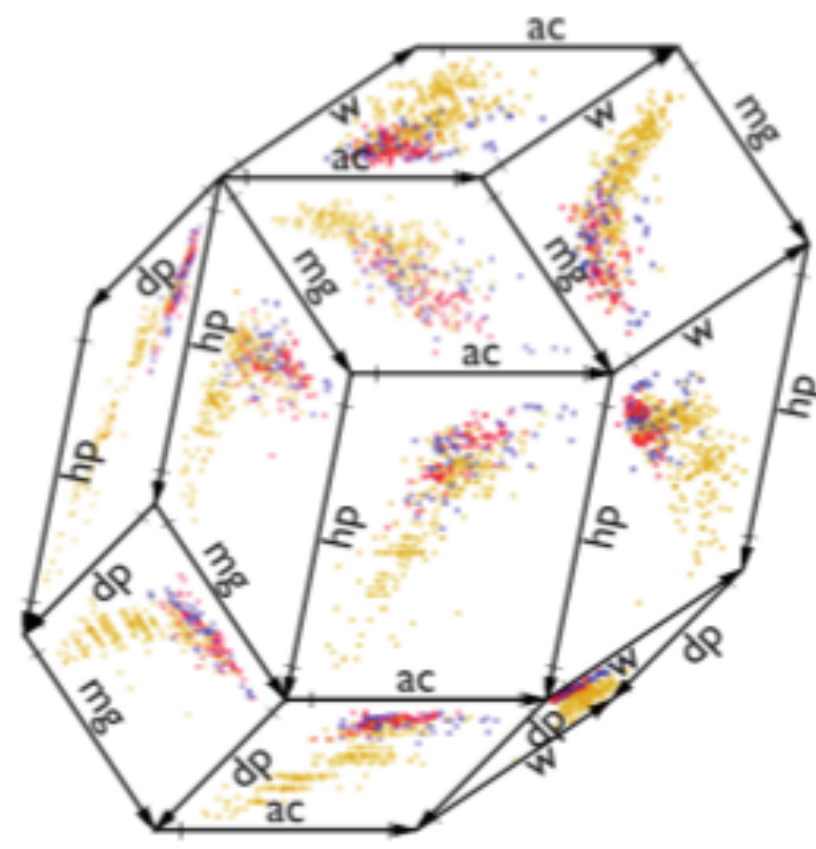
(a) scatterplots



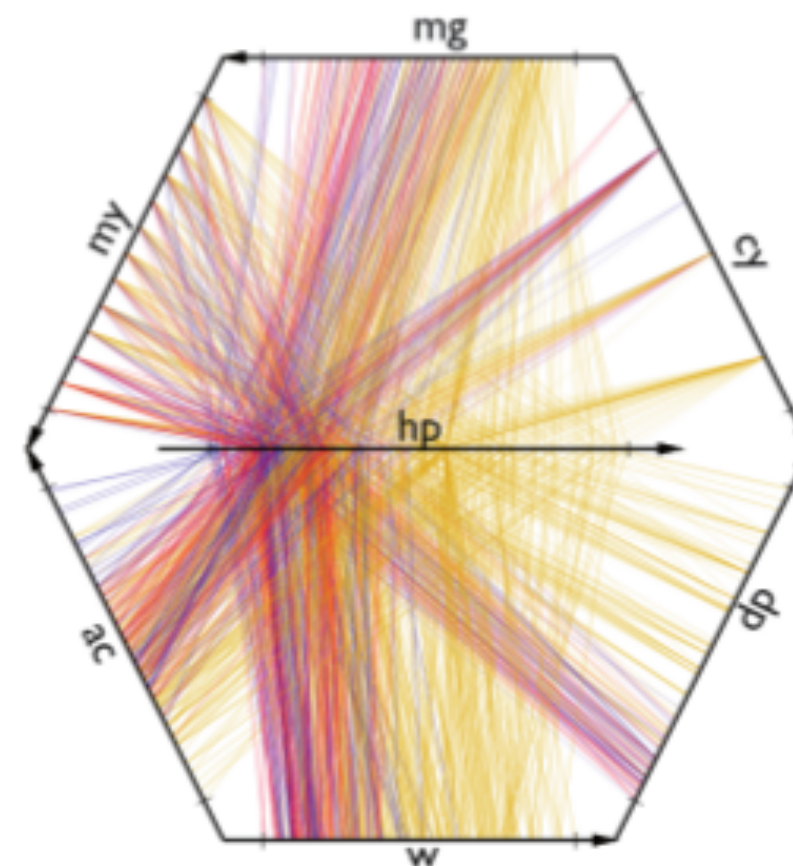
(b) Parallel Coordinates Plot



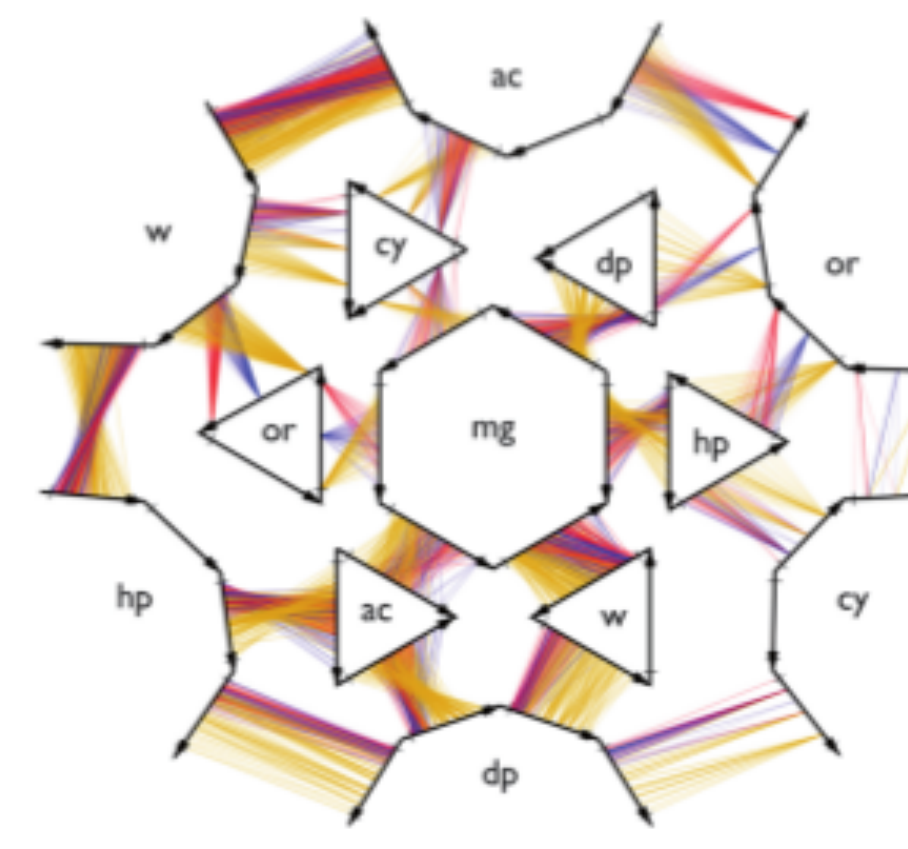
(c) radar chart



(d) Hyperbox

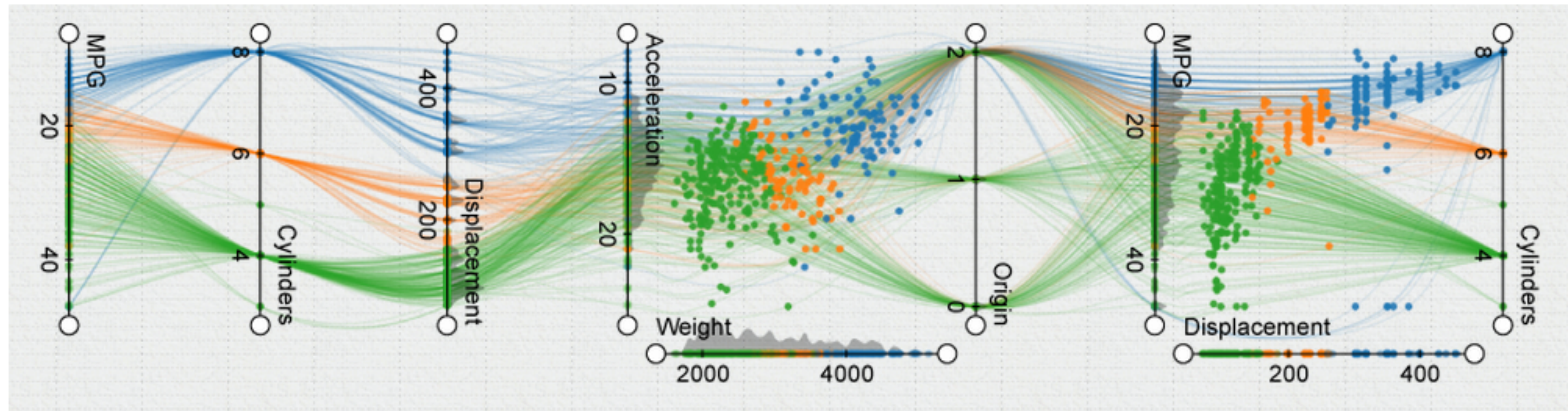


(e) Time Wheel



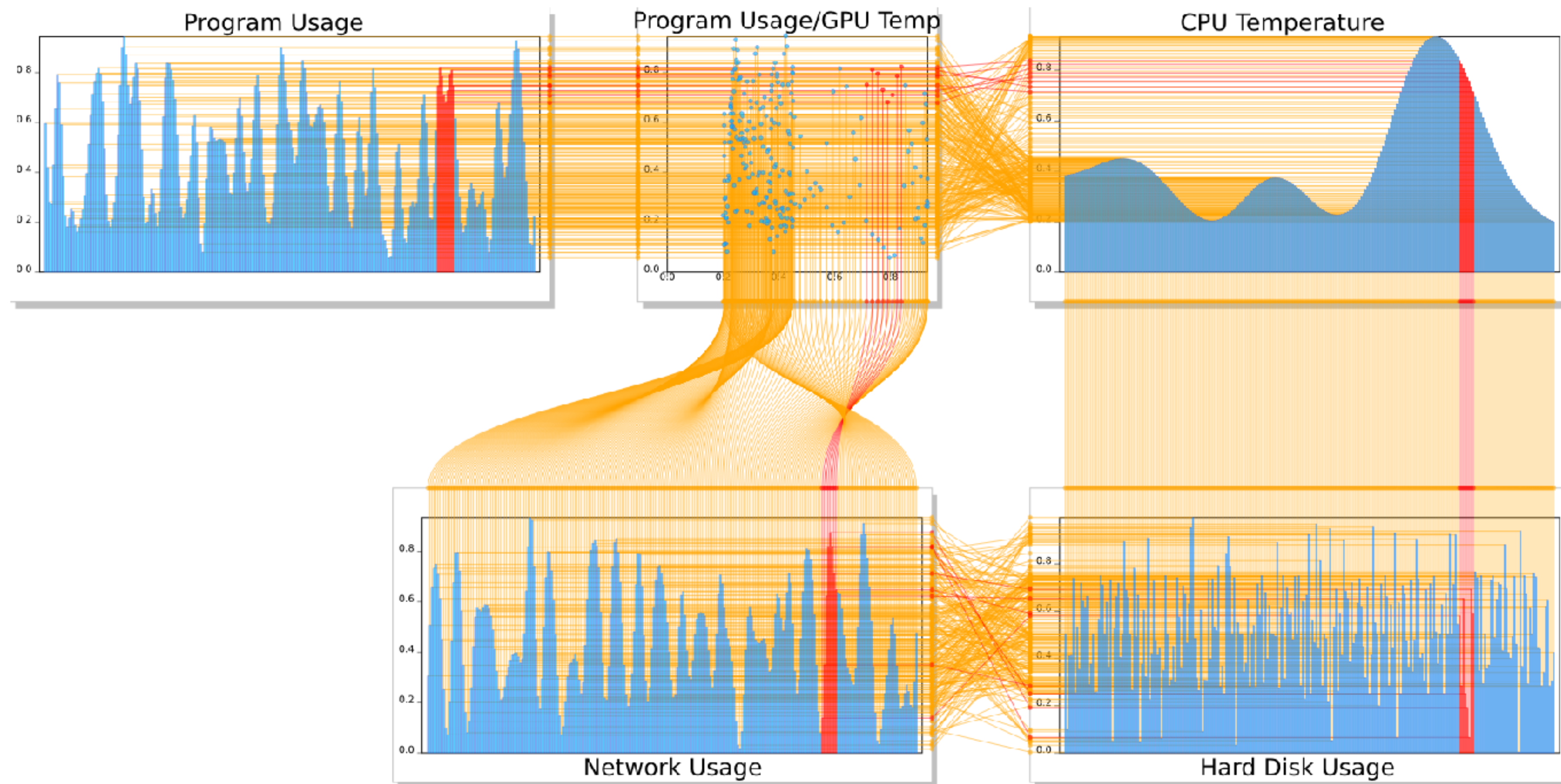
(f) Many-to-many PCP

Web-based implementation of FLINA concept

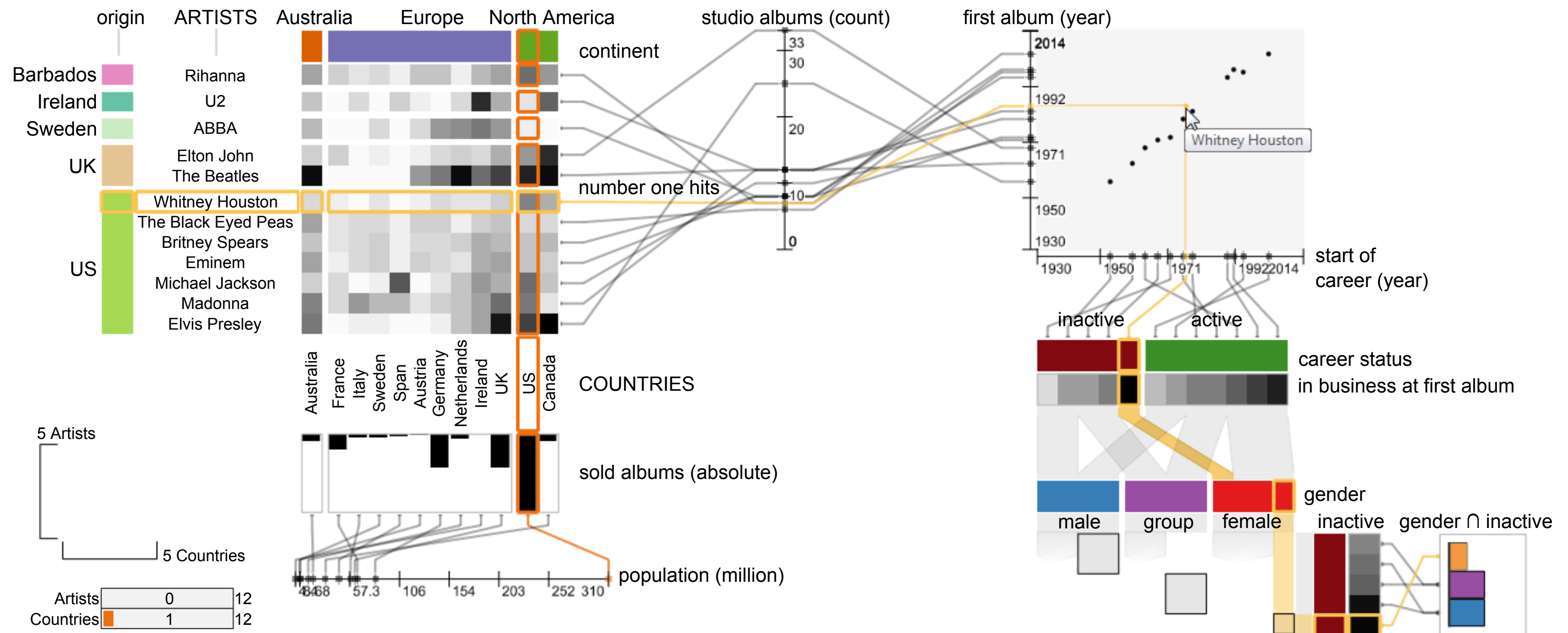


<http://vis.pku.edu.cn/mddv/val/>

Connected Charts



Domino



Spatial Axis Orientation

Parallel Sets

Parallel Sets

builds on PC to better handle categorical data

discrete

small number of values

no implied ordering between attributes

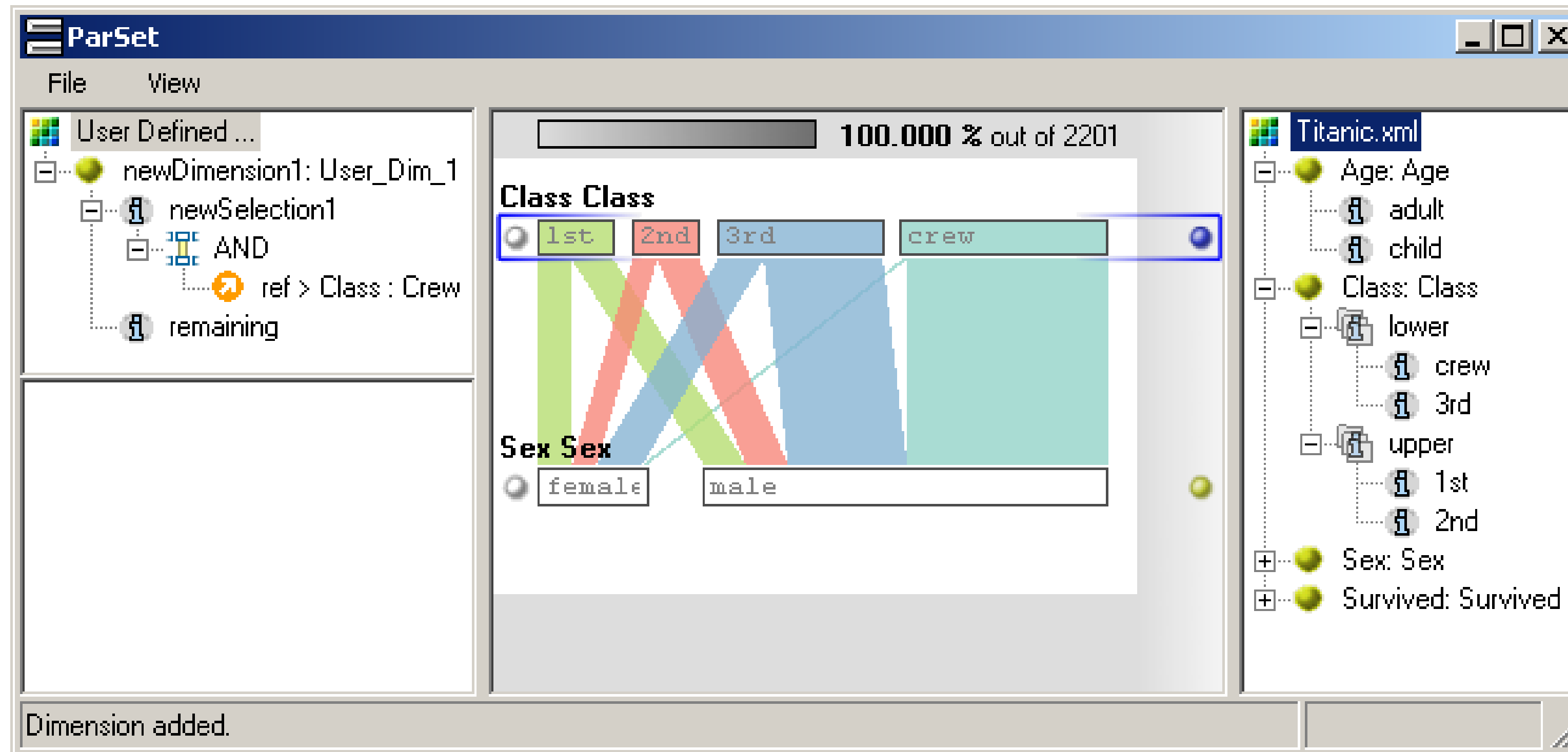
task: find relationship between attributes

interaction driven technique

Visual Encoding

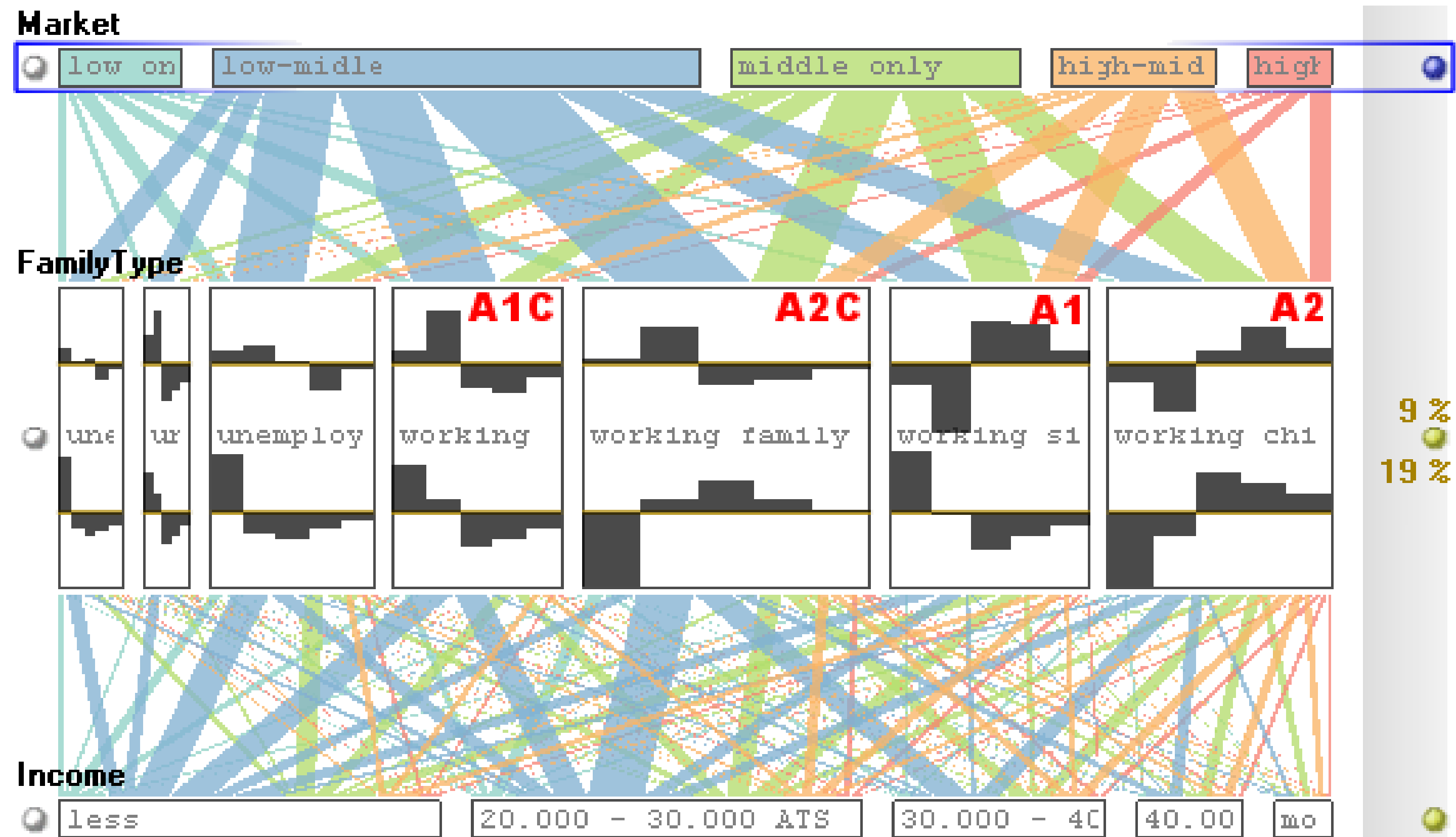
boxes scaled by frequency

color coded by values for current active dimension

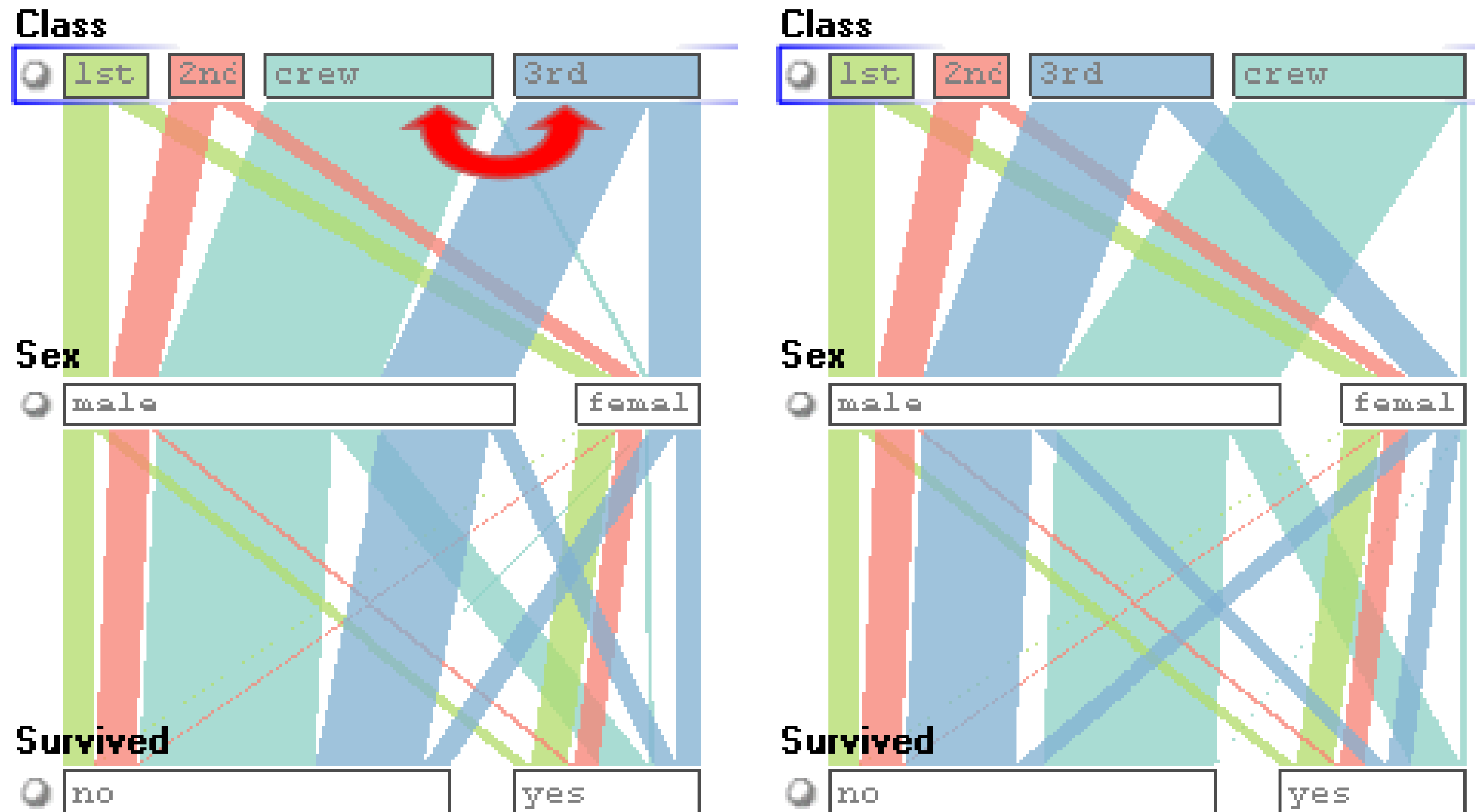


Visual Encoding

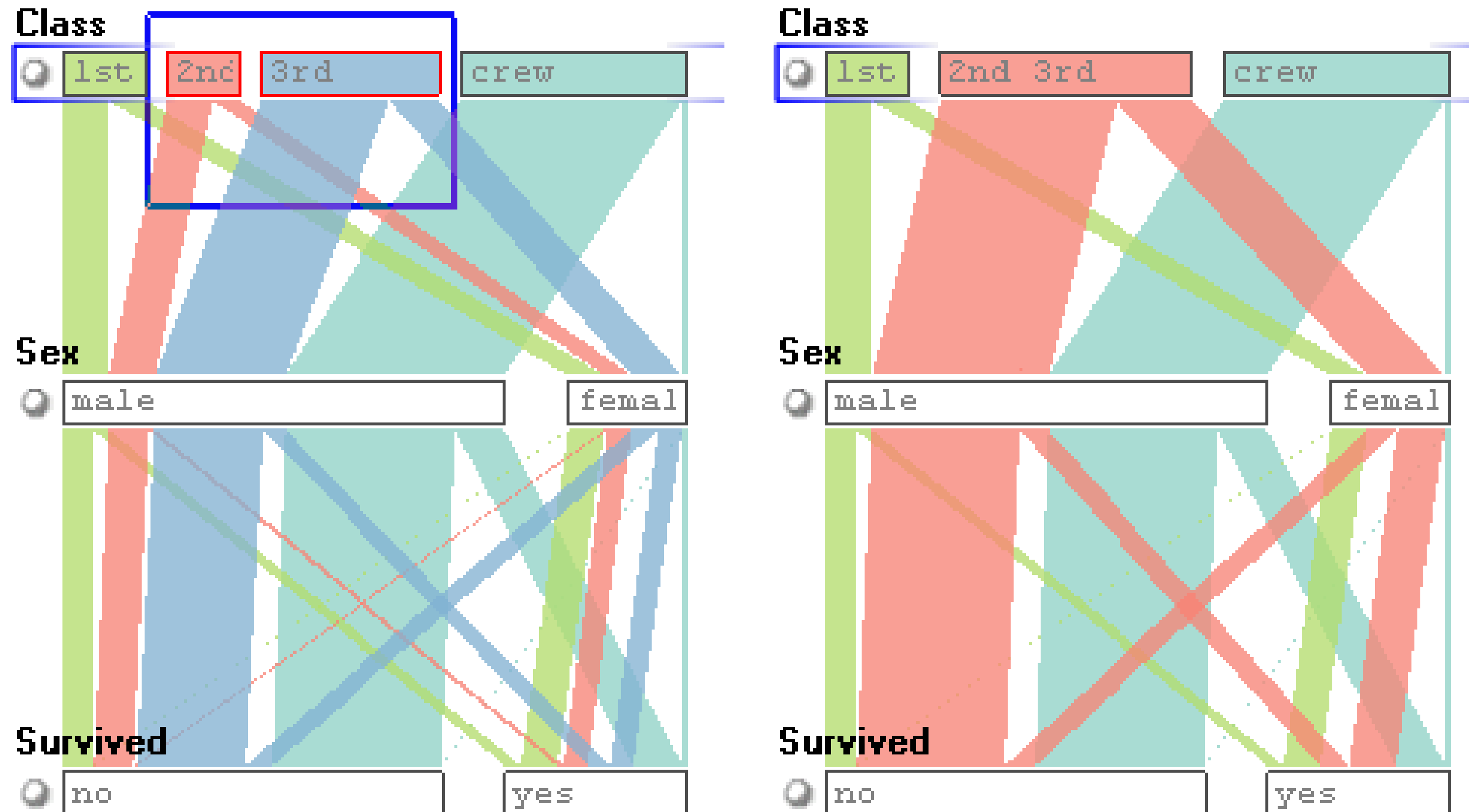
- boxes expand to show histogram



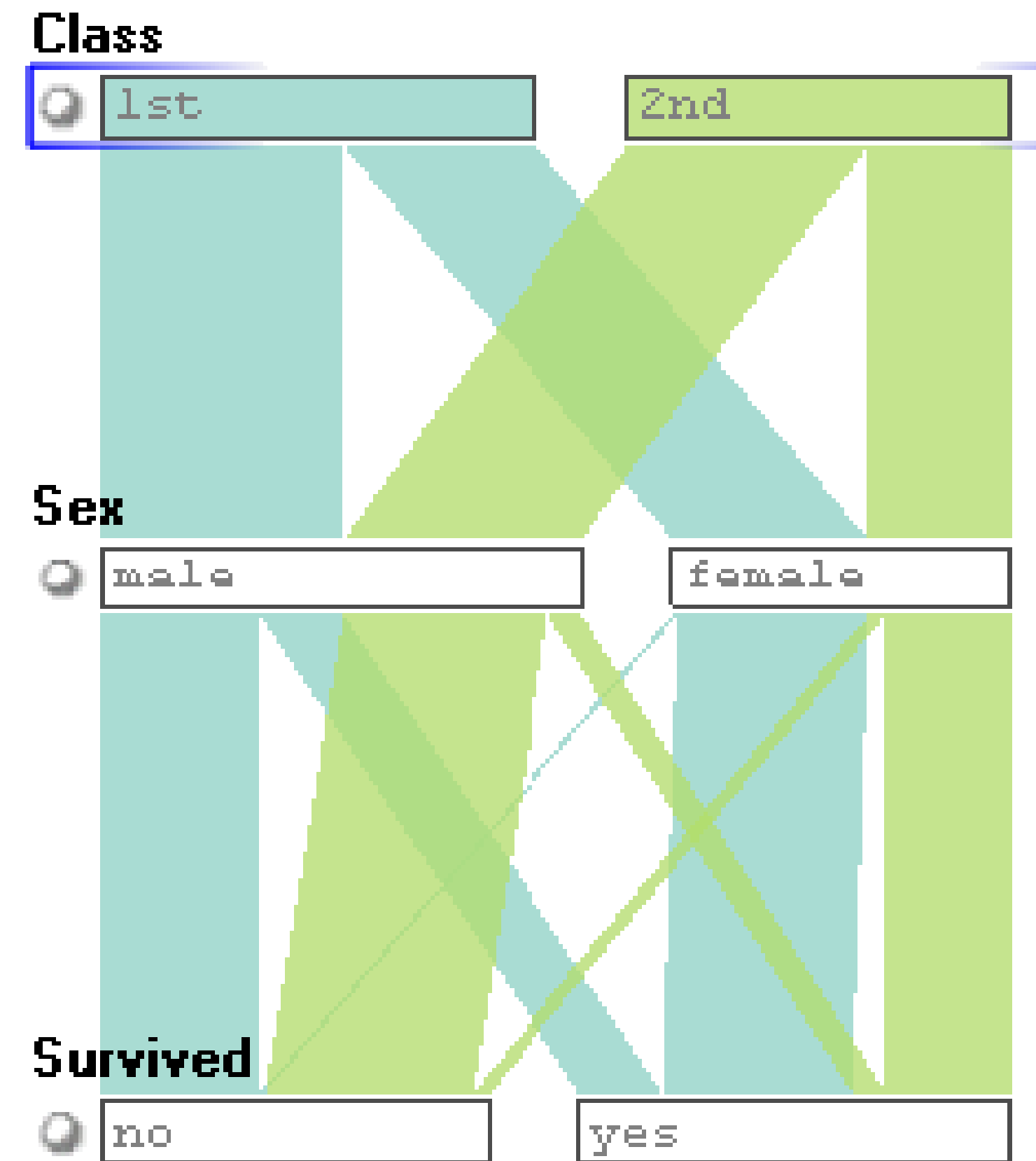
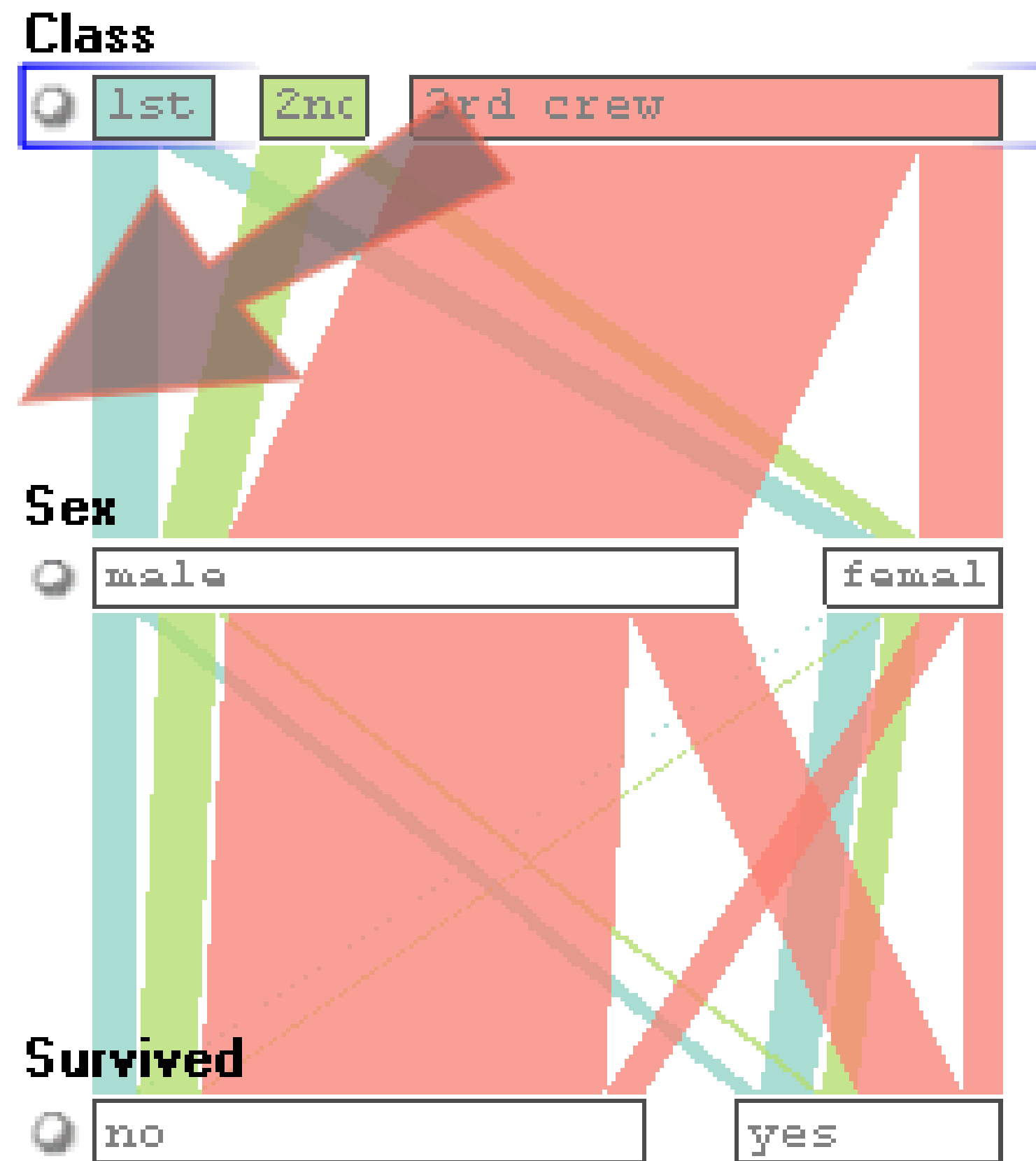
Interaction: Reorder



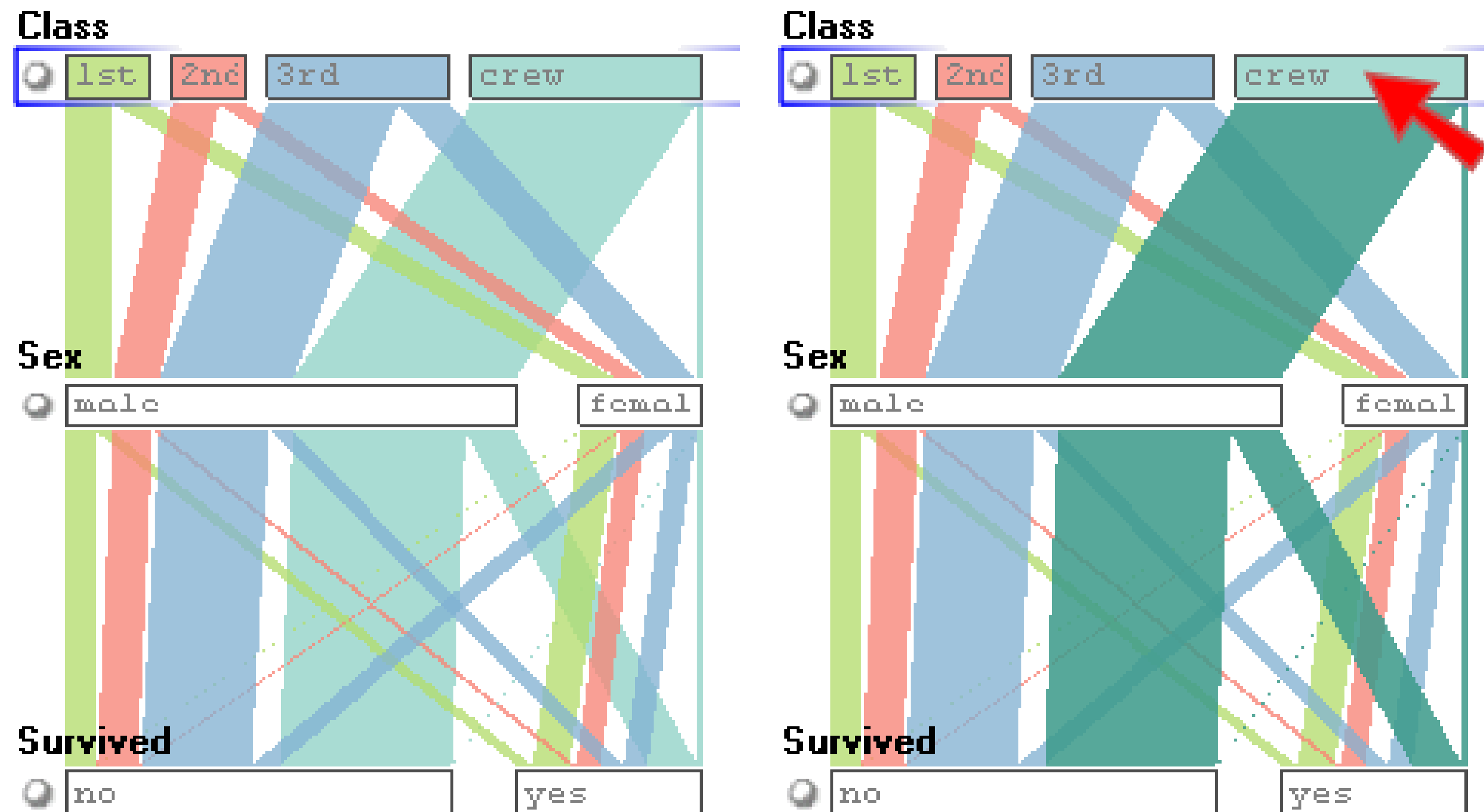
Interaction: Aggregate



Interaction: Filter



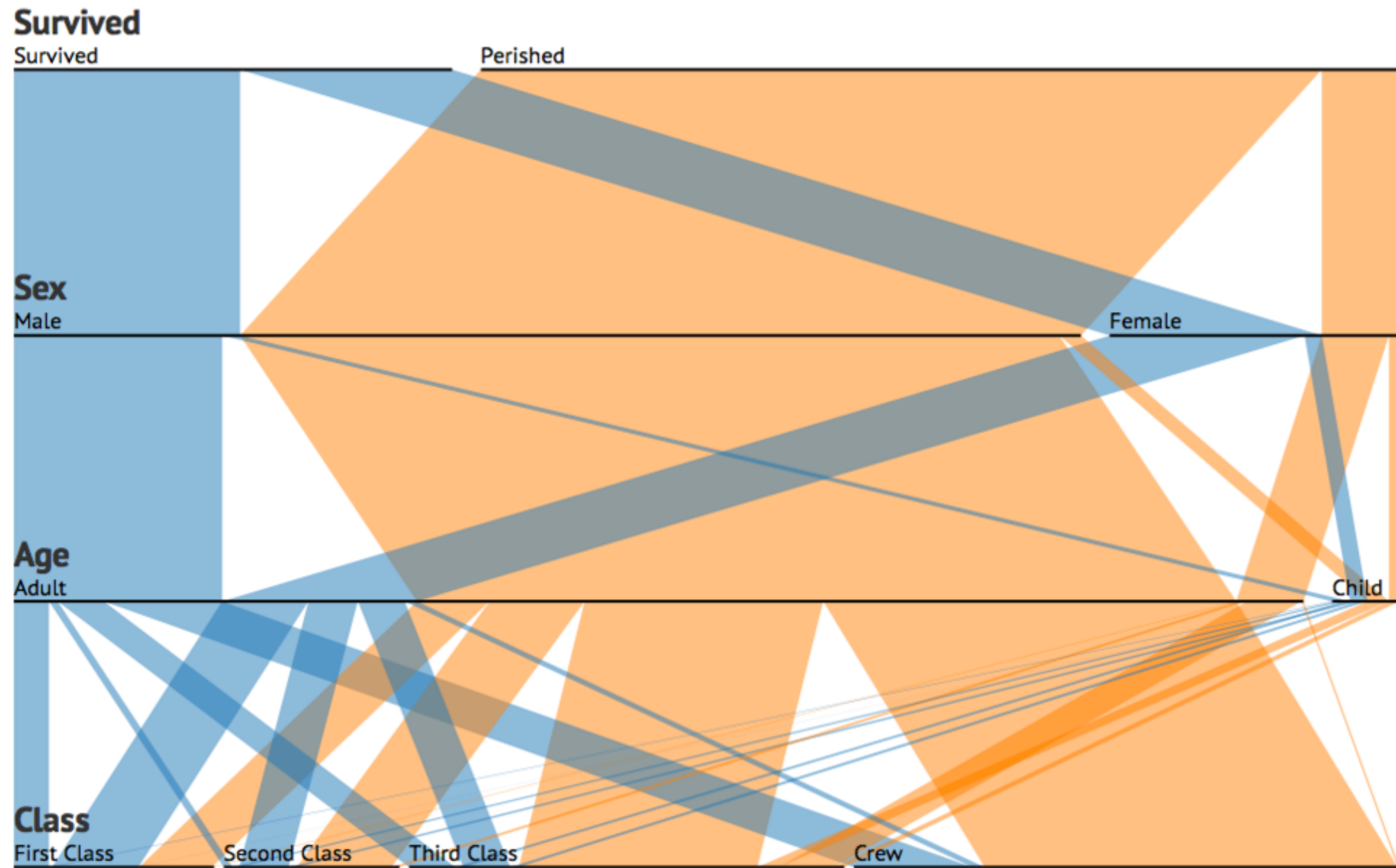
Interaction: Highlight



Parallel Sets

A visualisation technique for multidimensional categorical data.

Titanic Survivors



Curves?

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

Filling Space

Arrange Tables

① Express Values



② Separate, Order, Align Regions

→ Separate



→ Order



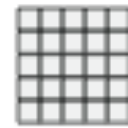
→ Align



→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume



→ Many Keys
Recursive Subdivision



③ Axis Orientation

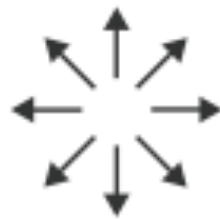
→ Rectilinear



→ Parallel

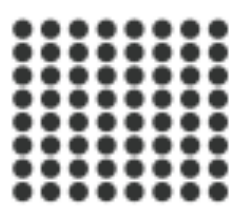


→ Radial



④ Layout Density

→ Dense



→ Space-Filling



Dense pixel display: VisDB

represent each data item, or each attribute in an item as a single pixel

can fit as many items on the screen as there are pixels, on the order of millions

relies heavily on color coding

challenge: what's the layout?

The data...

large database where each item has multiple attributes (on the order of 10)

goal: visualize the relevance of set of items which satisfy a query

plot out data items in a spiral pattern, ordered by relevance

