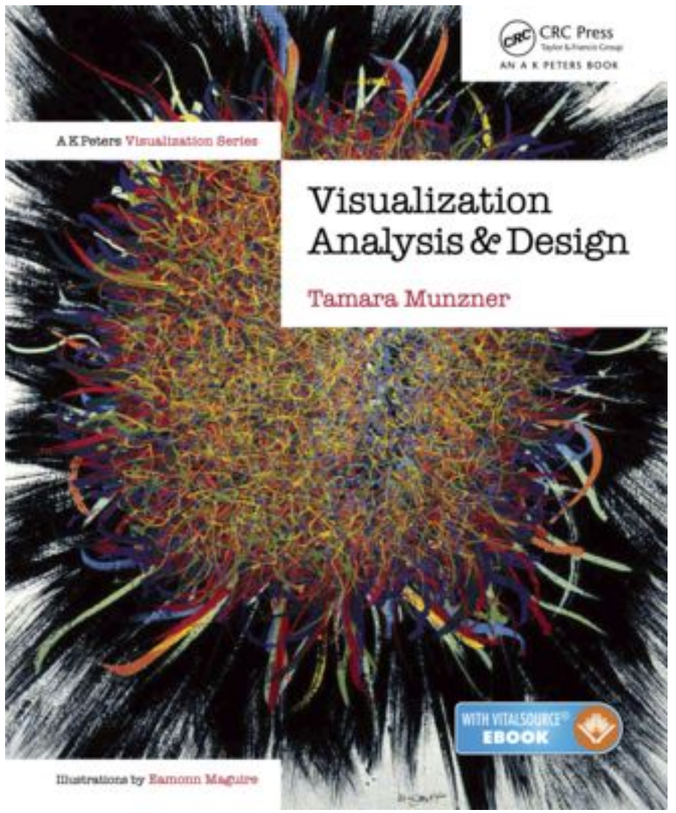
Visualization Analysis & Design

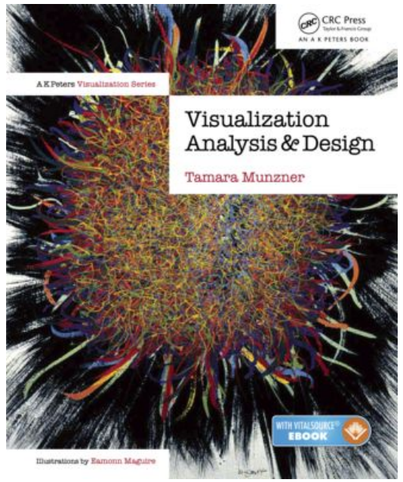
Teaching Slides

Tamara Munzner & Jean-Daniel Fekete



@tamaramunzner

Visualization Analysis & Design



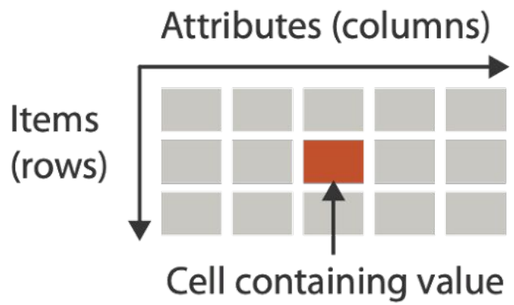
Arrange Tables (Ch 7) I

Tamara Munzner
Department of Computer Science
University of British Columbia
@tamaramunzner

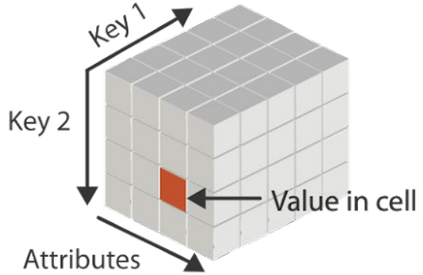
Focus on Tables

Dataset Types

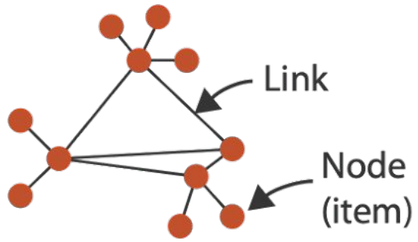
Tables



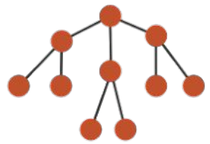
Multidimensional Table



Networks

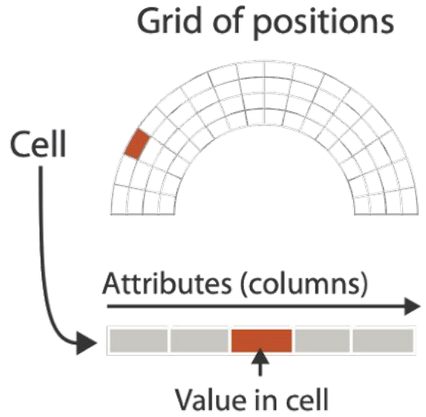


Trees



Spatial

Fields (Continuous)

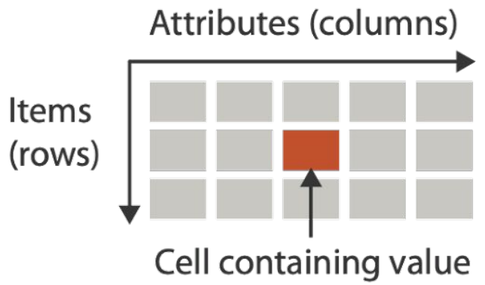


Geometry (Spatial)

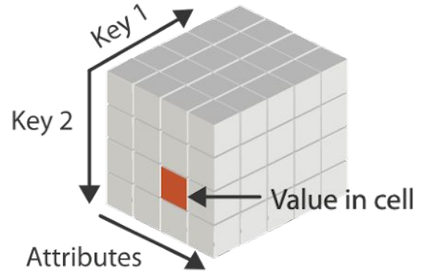
Keys and values

- key
  - independent attribute
  - used as unique index to look up items
  - simple tables: 1 key
  - multidimensional tables: multiple keys
- value
  - dependent attribute, value of cell

➔ Tables



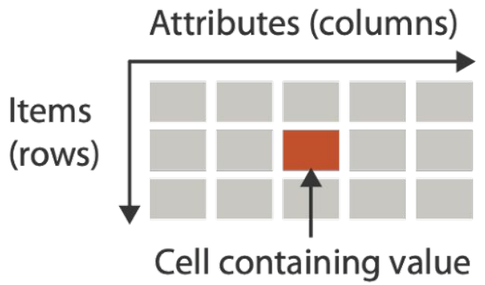
➔ *Multidimensional Table*



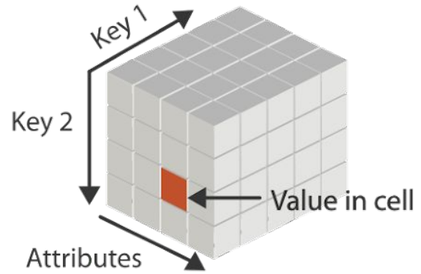
Keys and values

- key
  - independent attribute
  - used as unique index to look up items
  - simple tables: 1 key
  - multidimensional tables: multiple keys
- value
  - dependent attribute, value of cell
- classify arrangements by keys used
  - 0, 1, 2, ...

➔ Tables

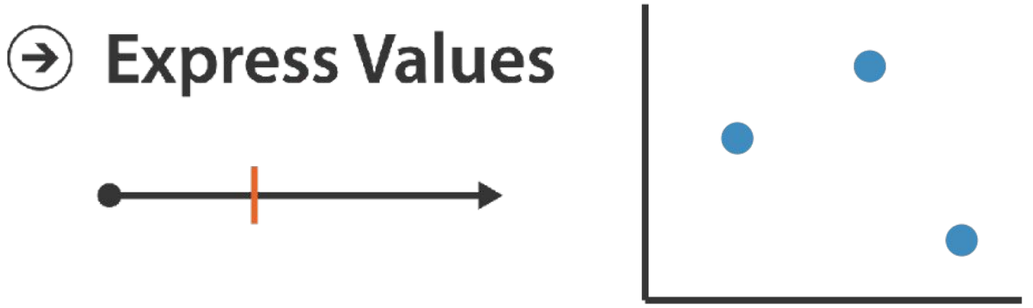


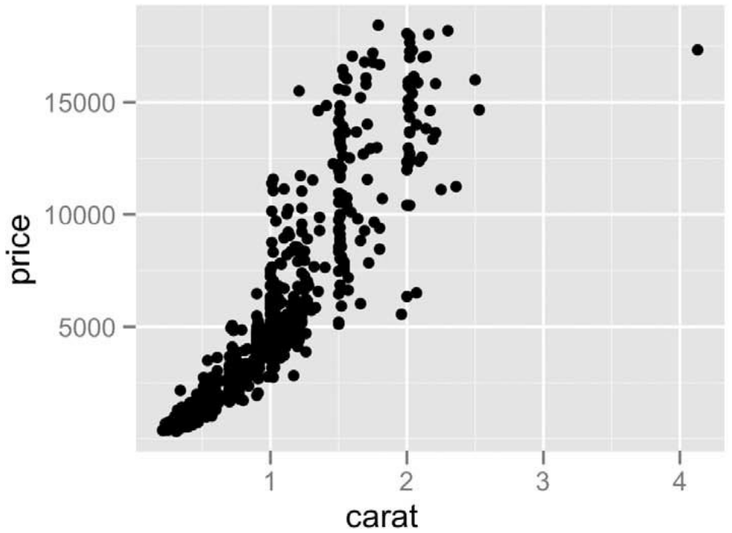
➔ *Multidimensional Table*

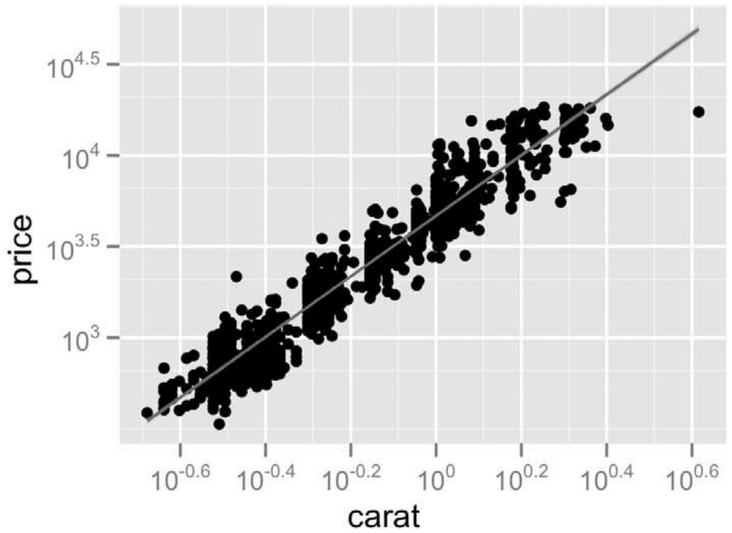


Idiom: **scatterplot**

- **express** values (magnitudes)
  - quantitative attributes
- no keys, only values



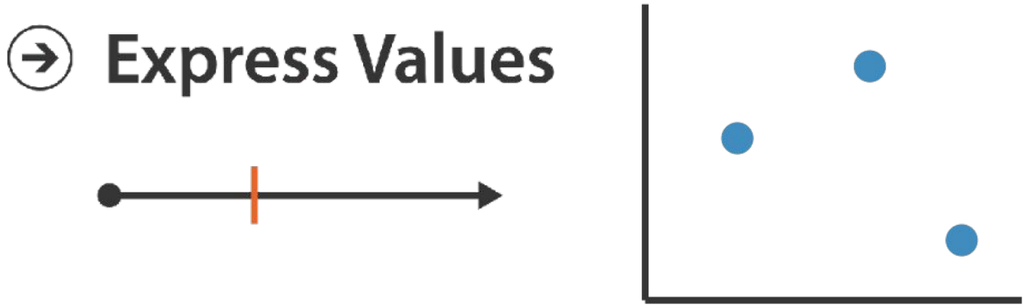


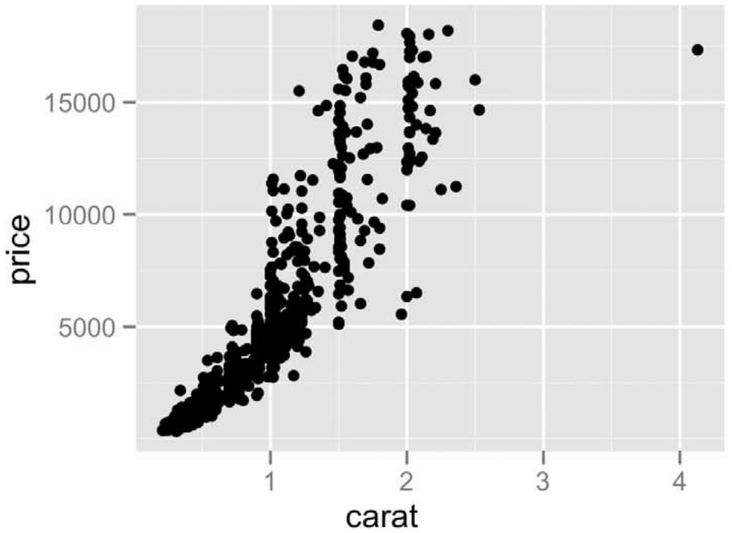


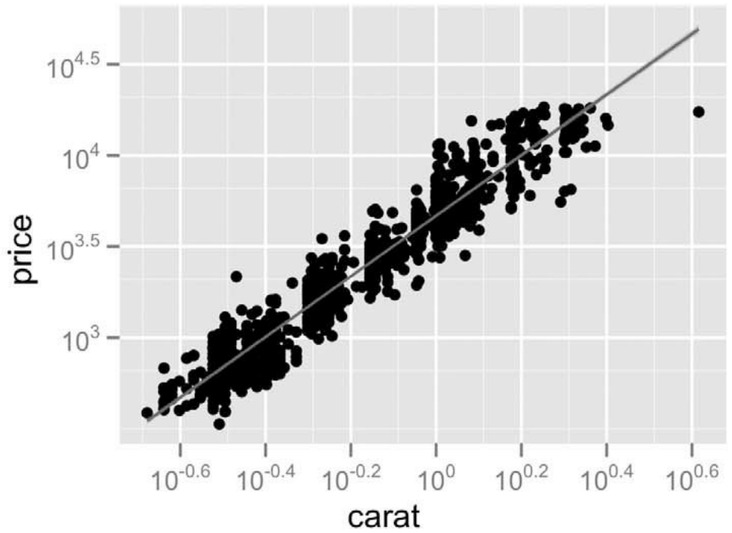
[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

Idiom: **scatterplot**

- **express** values (magnitudes)
  - quantitative attributes
- no keys, only values
  - data
    - 2 quant attribs
  - mark: points
  - channels
    - horiz + vert position

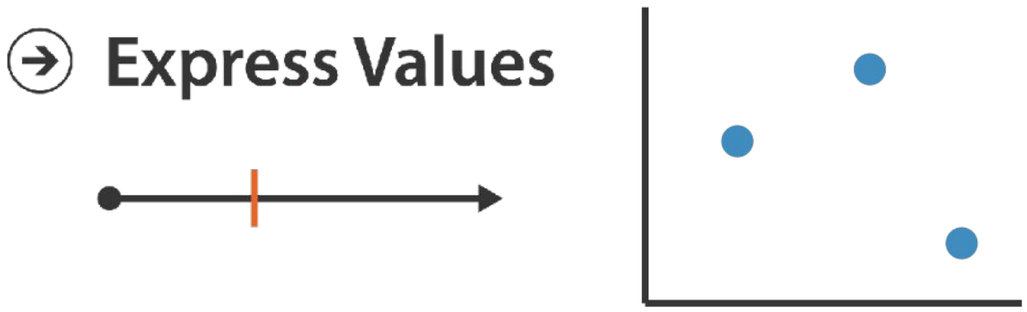




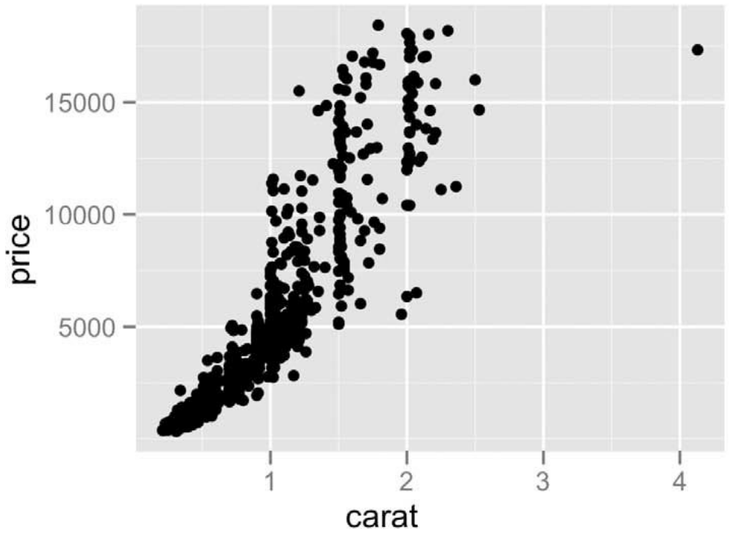


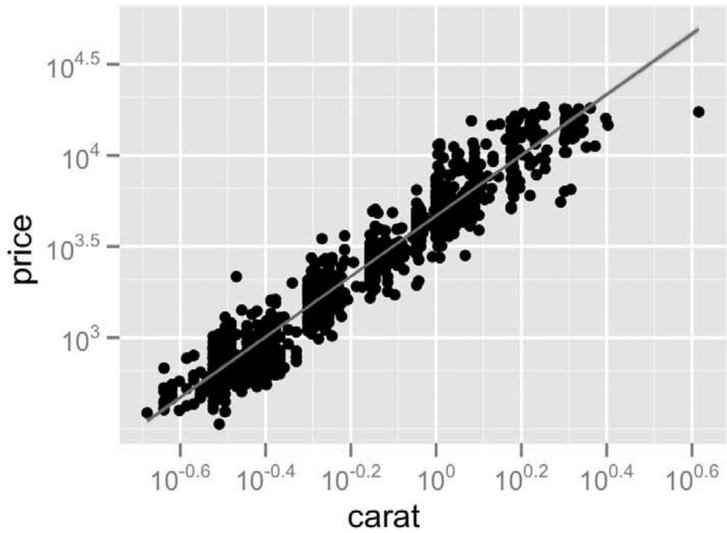
[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

Idiom: **scatterplot**



- **express** values (magnitudes)
  - quantitative attributes
- no keys, only values
  - data
    - 2 quant attribs
  - mark: points
  - channels
    - horiz + vert position
  - tasks
    - find trends, outliers, distribution, correlation, clusters
  - scalability
    - hundreds of items

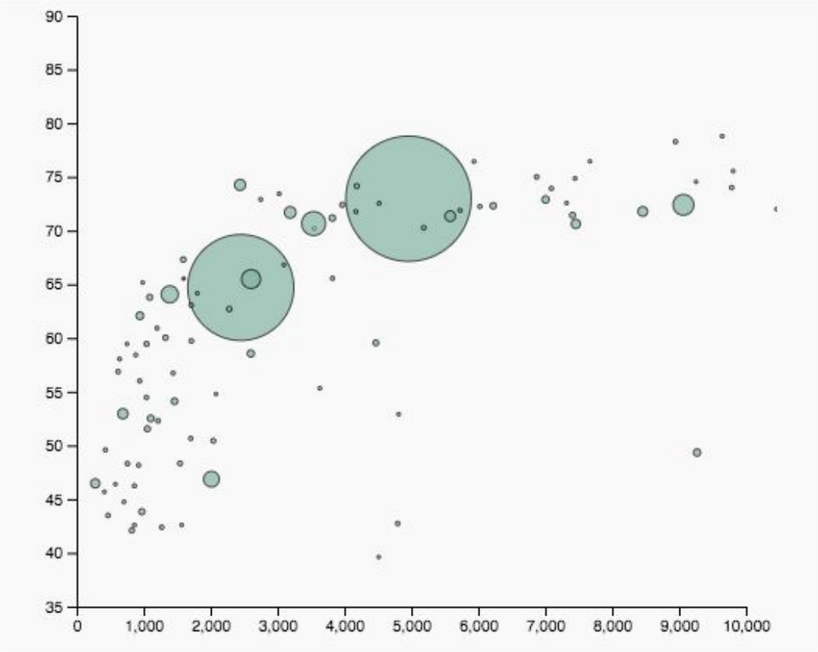


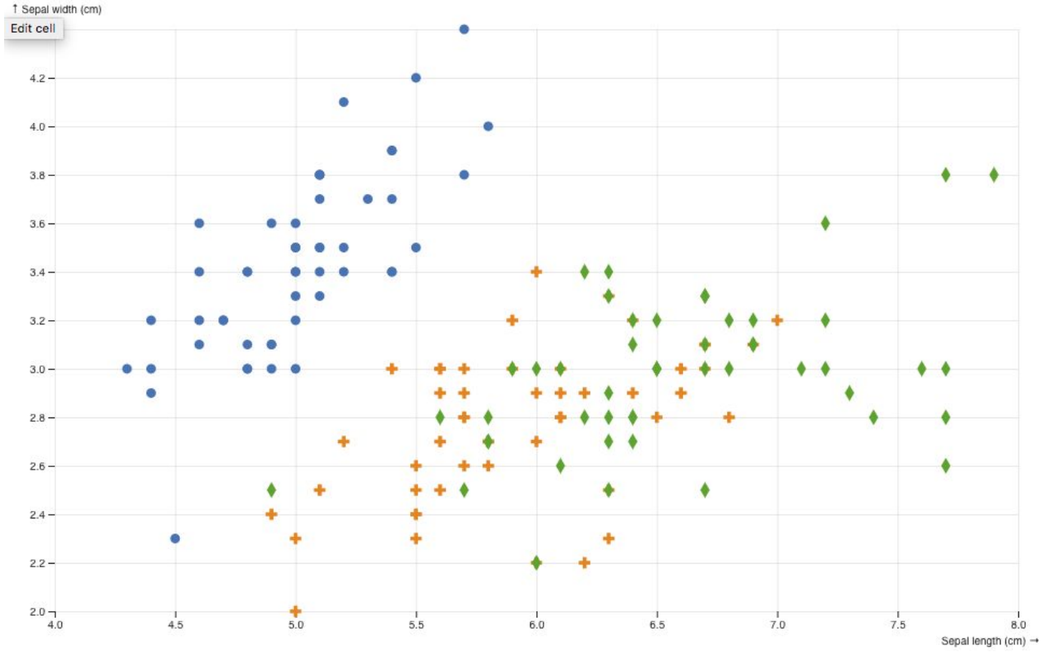


[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

Scatterplots: Encoding more channels

- additional channels viable since using point marks
  - color
  - size (1 quant attribute, used to control 2D area)
    - note radius would mislead, take square root since area grows quadratically
  - shape





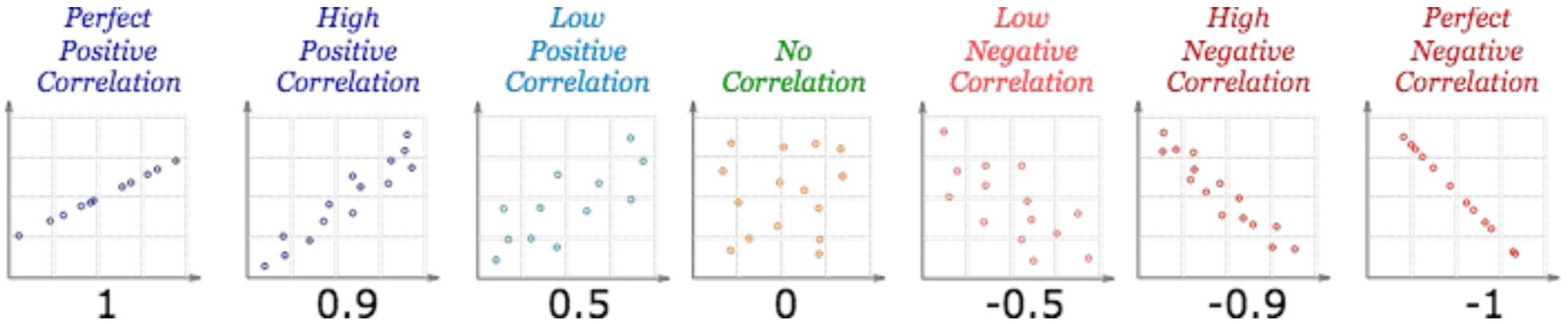
https://www.d3-graph-gallery.com/graph/bubble_basic.html

https://observablehq.com/@d3/scatterplot-with-shapes

Scatterplot tasks

Scatterplot tasks

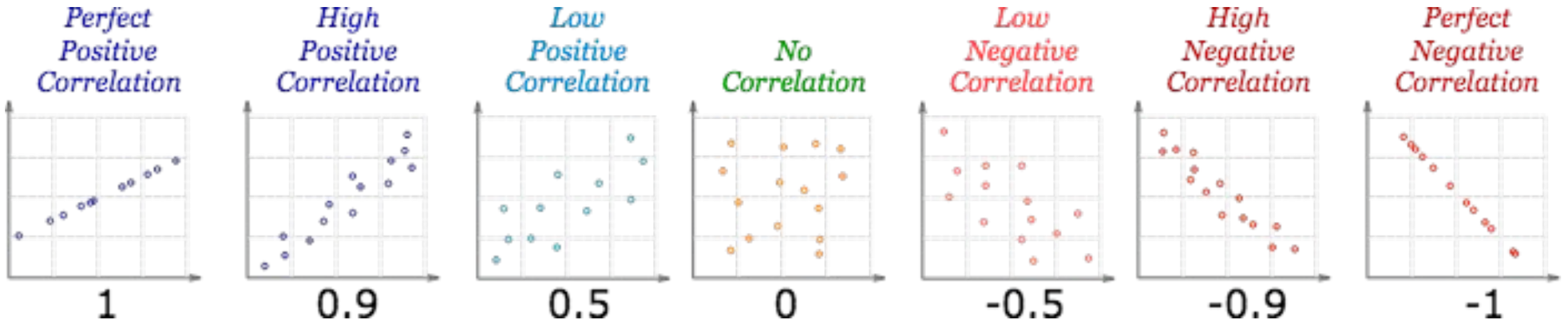
- correlation



https://www.mathsisfun.com/data/scatter-xy-plots.html

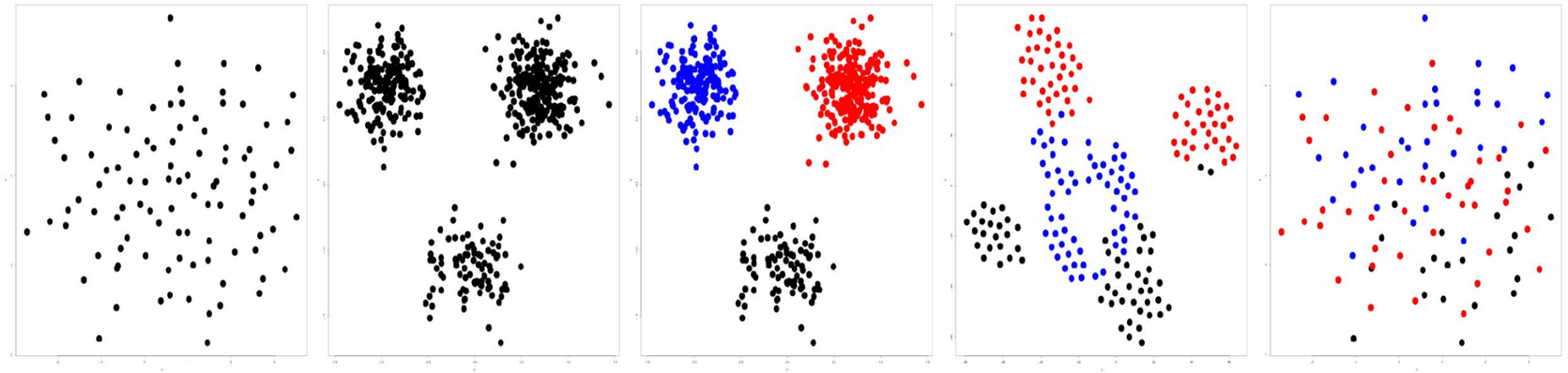
Scatterplot tasks

- correlation



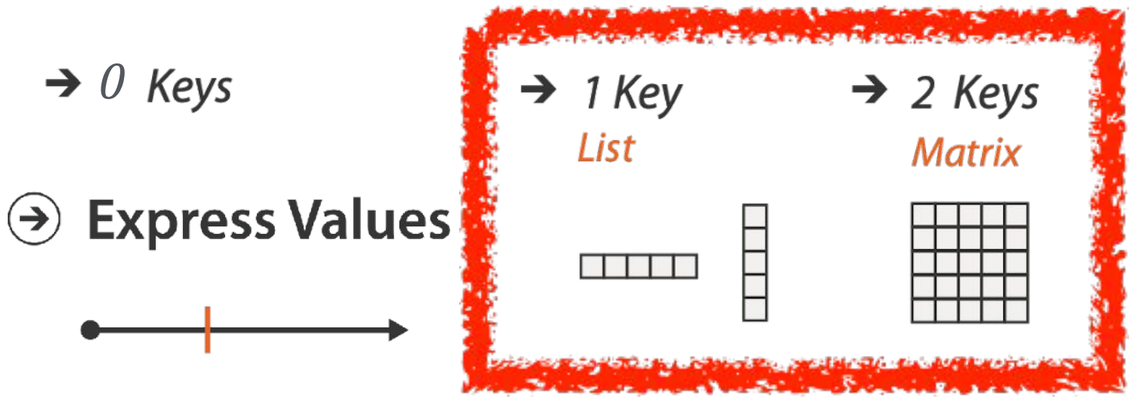
https://www.mathsisfun.com/data/scatter-xy-plots.html

- clusters/groups, and clusters vs classes



https://www.cs.ubc.ca/labs/imager/tr/2014/DRVisTasks/

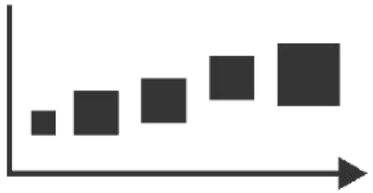
Some keys



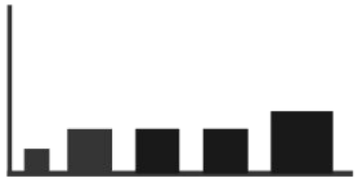
Some keys: Categorical regions

- Separate

- Order



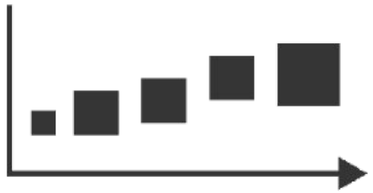
- Align



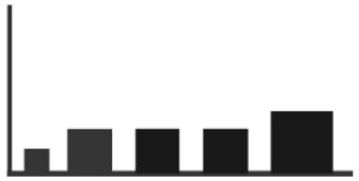
Regions: Separate, order, align

→ Separate

→ Order



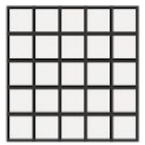
→ Align



- regions: contiguous bounded areas distinct from each other
  - separate into spatial regions: one mark per region (for now)
- use categorical or ordered attribute to separate into regions
  - no conflict with expressiveness principle for categorical attributes
- use ordered attribute to order and align regions

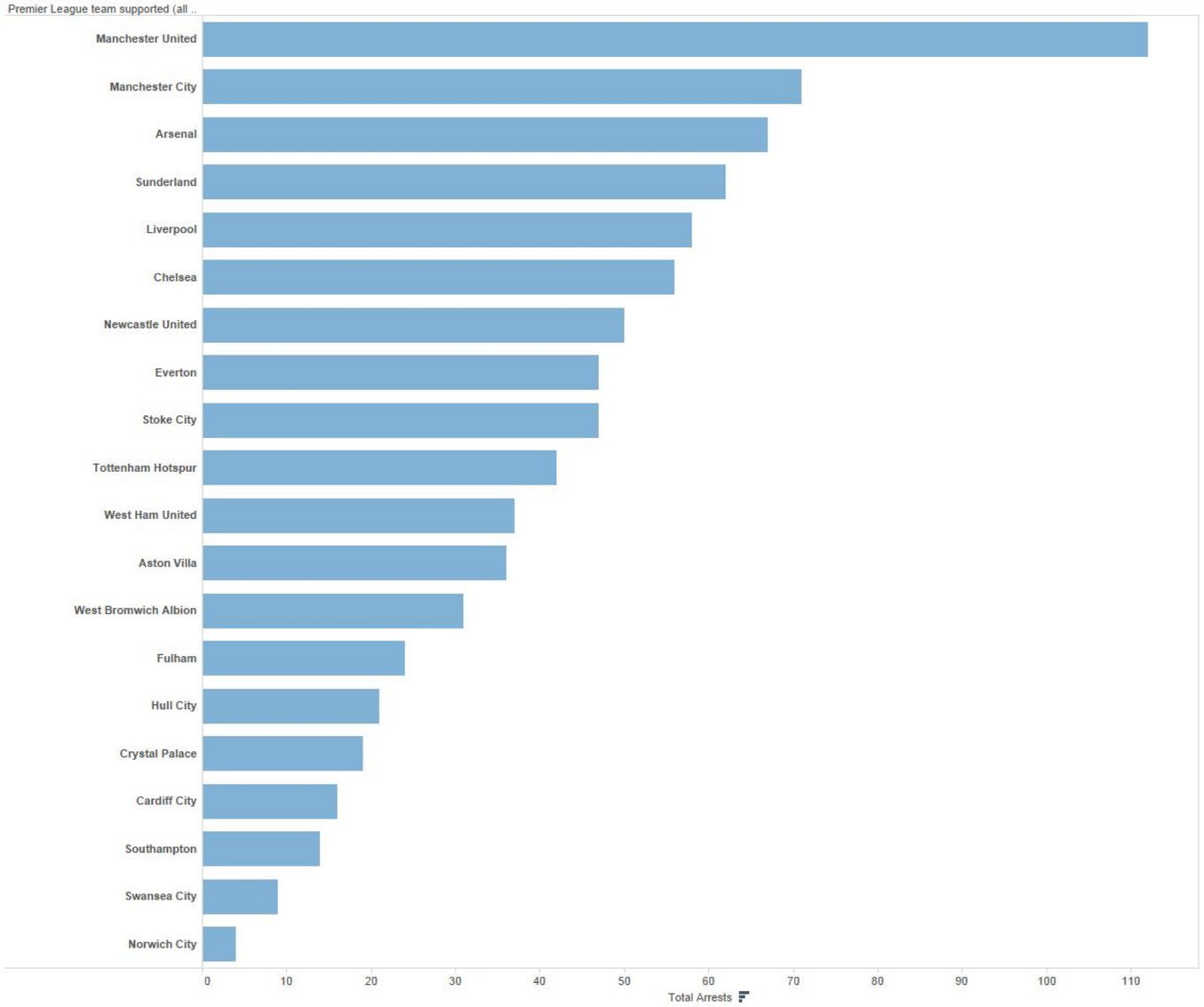
→ *1 Key*
List

→ *2 Keys*
Matrix



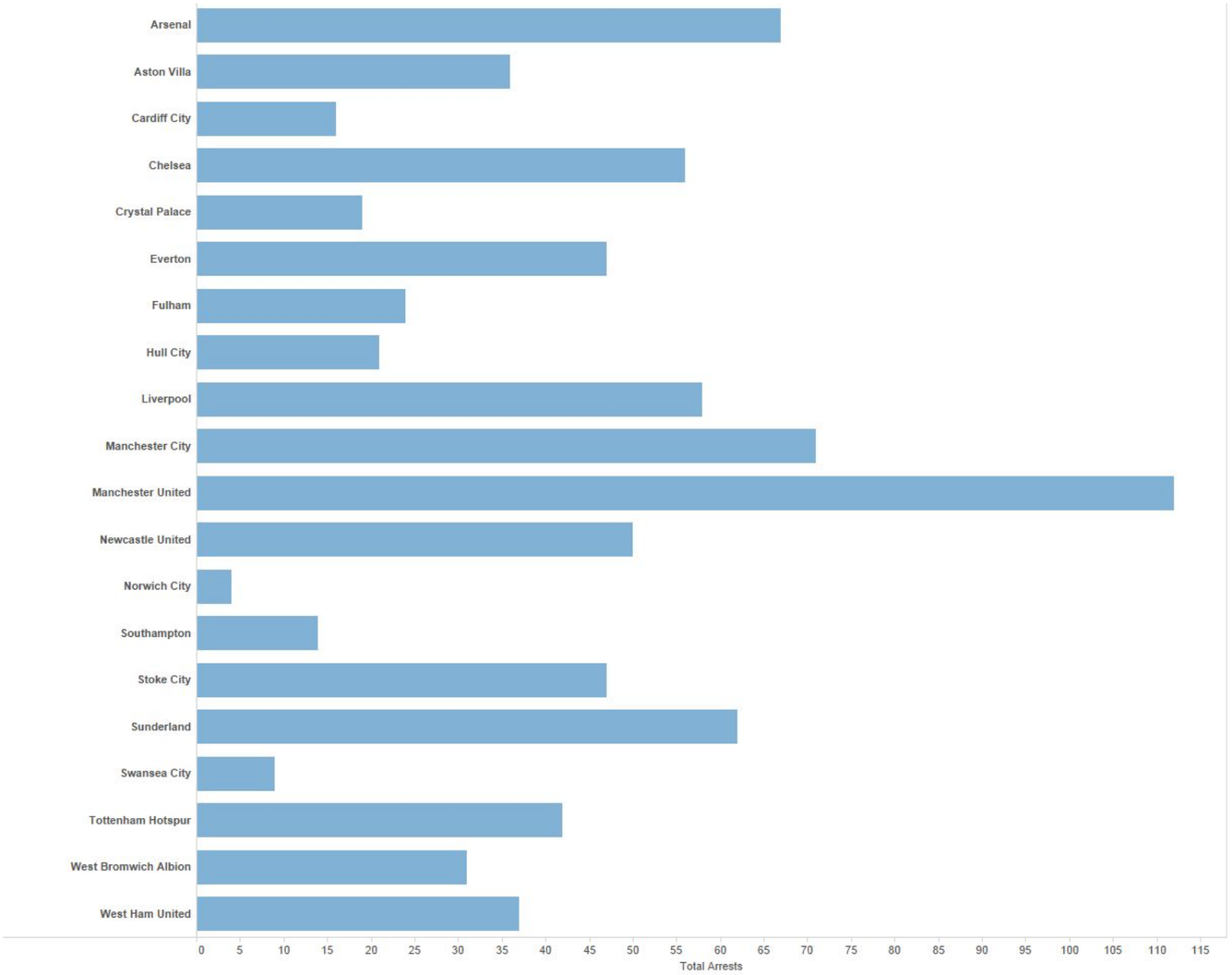
Separated and aligned and ordered

- best case



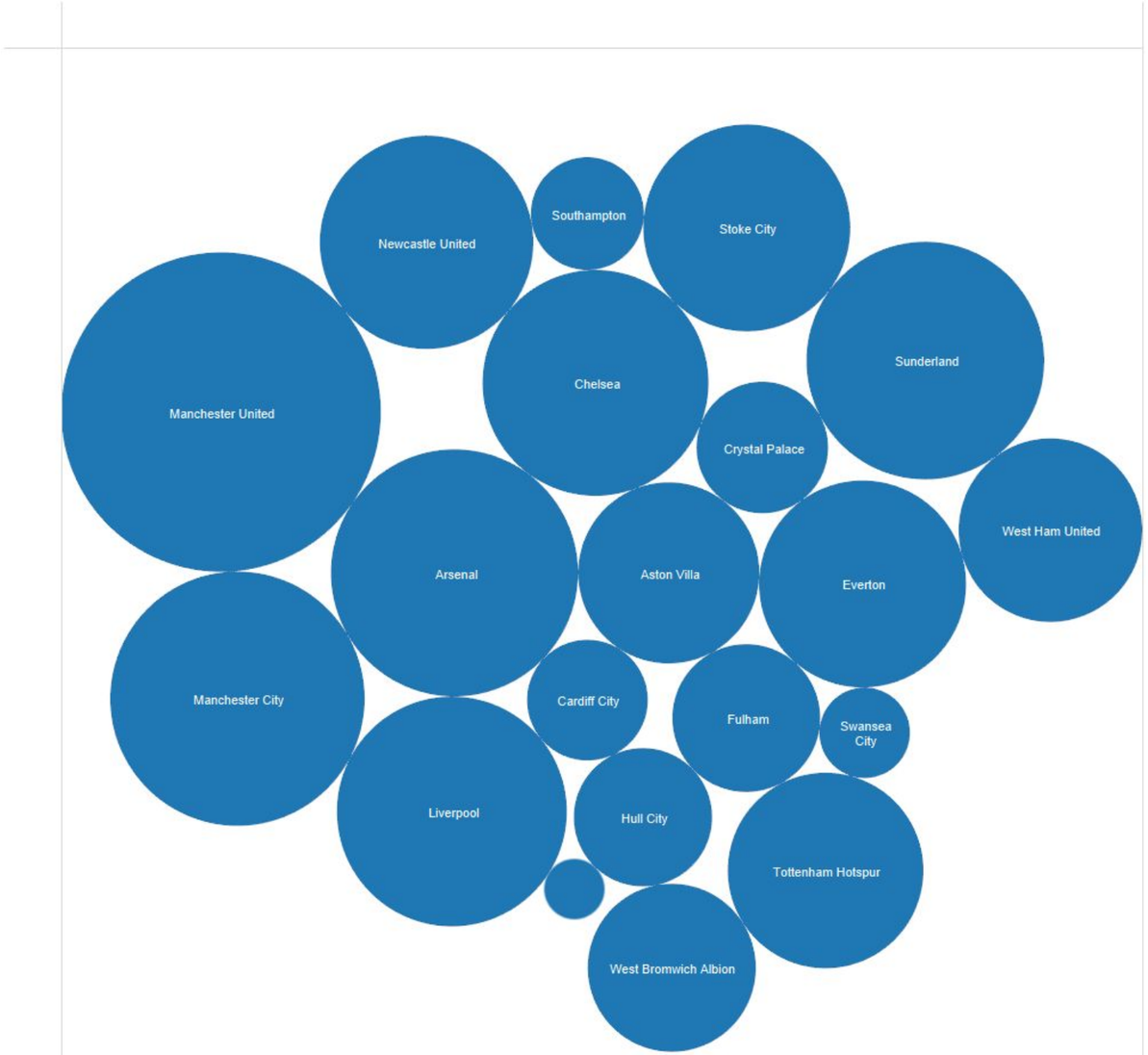
Separated and aligned but not ordered

- limitation: hard to know rank. what's 4th? what's 7th?



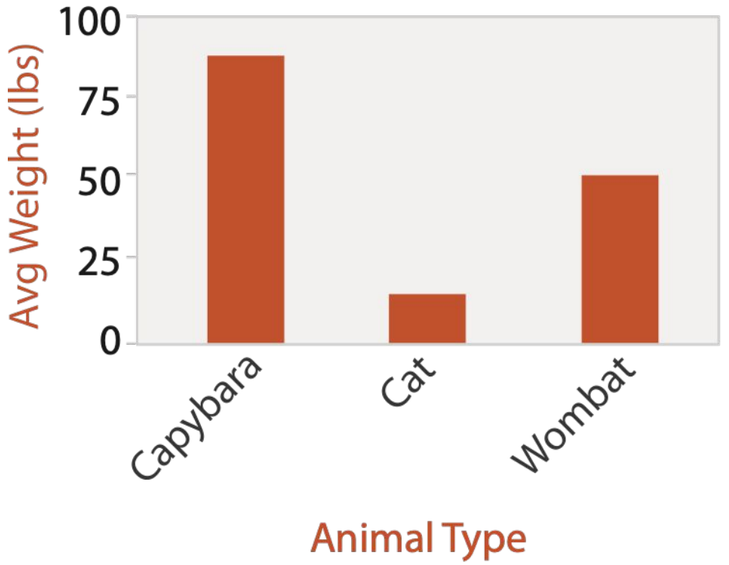
Separated but not aligned or ordered

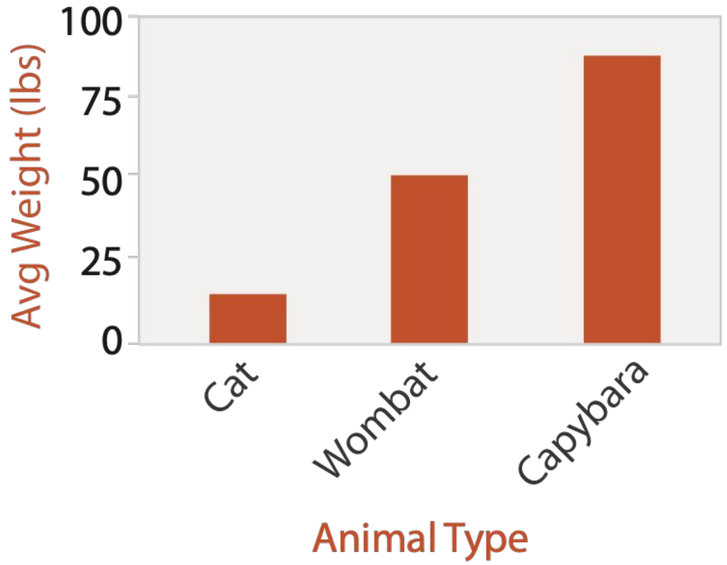
- limitation: hard to make comparisons with size (vs aligned position)



Idiom: **bar chart**

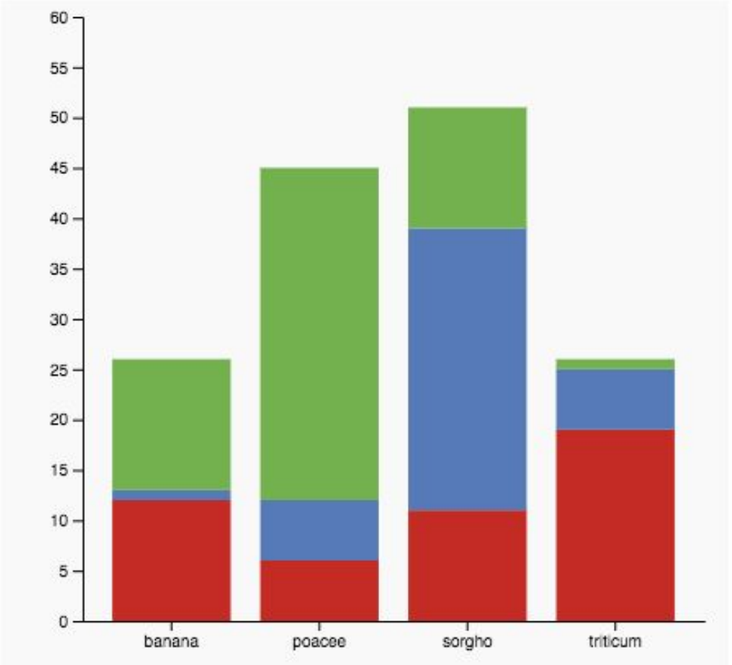
- one key, one value
  - data
    - 1 categ attrib, 1 quant attrib
  - mark: lines
  - channels
    - length to express quant value
    - spatial regions: one per mark
      - separated horizontally, aligned vertically
      - ordered by quant attrib
        - » by label (alphabetical), by length attrib (data-driven)
  - task
    - compare, lookup values
  - scalability
    - dozens to hundreds of levels for key attrib [bars], hundreds for values





Idiom: **stacked bar chart**

- one more key
  - data
    - 2 categ attrib, 1 quant attrib
  - mark: vertical stack of line marks
    - **glyph**: composite object, internal structure from multiple marks
  - channels
    - length and color hue
    - spatial regions: one per glyph
      - aligned: full glyph, lowest bar component
      - unaligned: other bar components
  - task
    - part-to-whole relationship
  - scalability: asymmetric
    - for *stacked* key attrib, 10-12 levels [segments]
    - for *main* key attrib, dozens to hundreds of levels [bars]

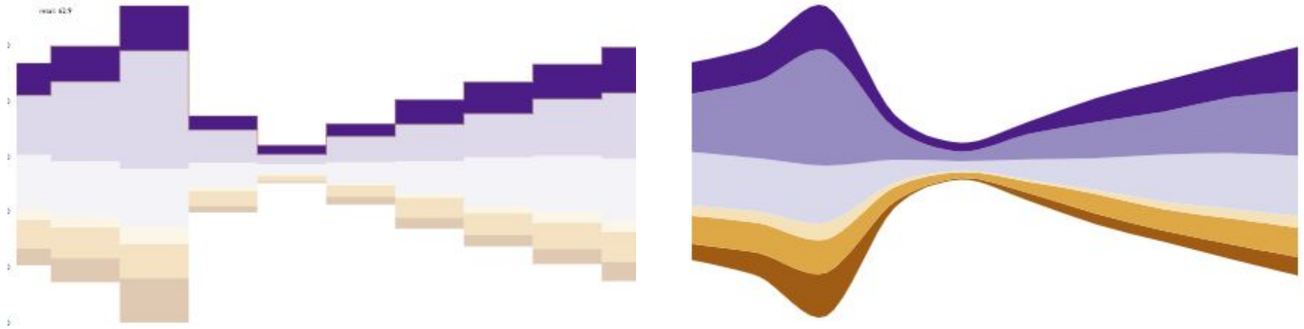


https://www.d3-graph-gallery.com/graph/b

ot_stacked_basicWide.html

Idiom: **streamgraph**

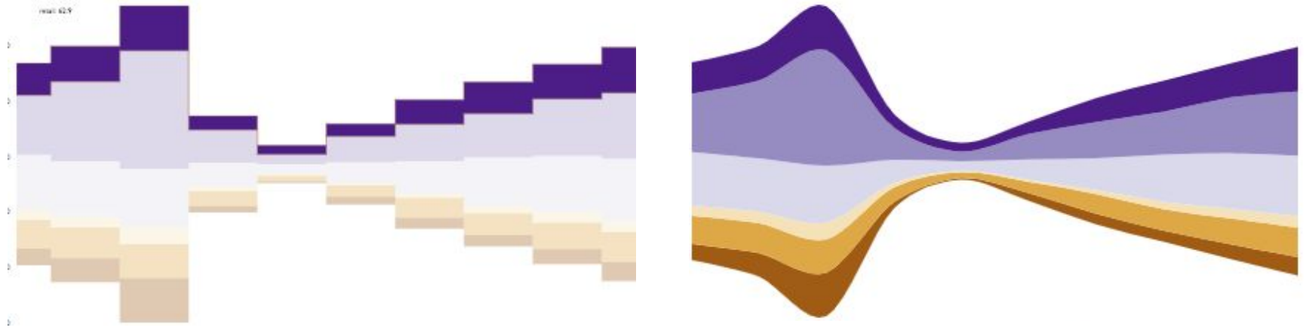
- generalized stacked graph
  - emphasizing horizontal continuity
    - vs vertical items
  - data
    - 1 categ key attrib (movies)
    - 1 ordered key attrib (time)
    - 1 quant value attrib (counts)
  - derived data
    - geometry: layers, where height encodes counts
    - 1 quant attrib (layer ordering)



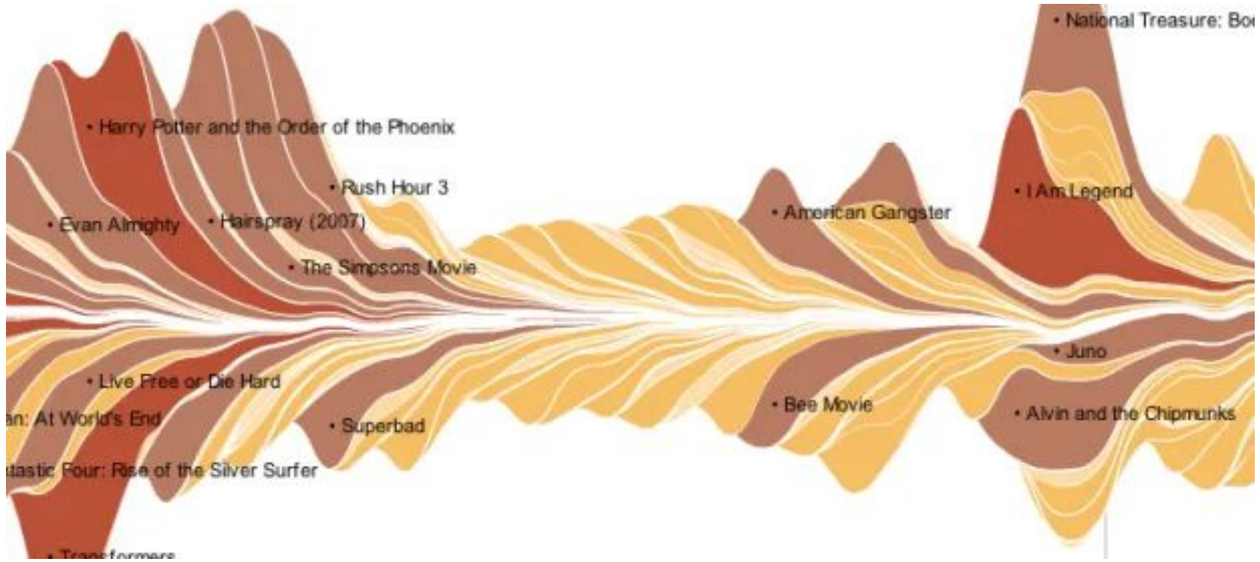
[Stacked Graphs Geometry & Aesthetics. Byron and Wattenberg. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2008) 14(6): 1245–1252, (2008).]

Idiom: **streamgraph**

- generalized stacked graph
  - emphasizing horizontal continuity
    - vs vertical items
  - data
    - 1 categ key attrib (movies)
    - 1 ordered key attrib (time)
    - 1 quant value attrib (counts)
  - derived data
    - geometry: layers, where height encodes counts
    - 1 quant attrib (layer ordering)
  - scalability
    - hundreds of time keys
    - dozens to hundreds of movies keys
      - more than stacked bars: most layers don't extend across whole chart



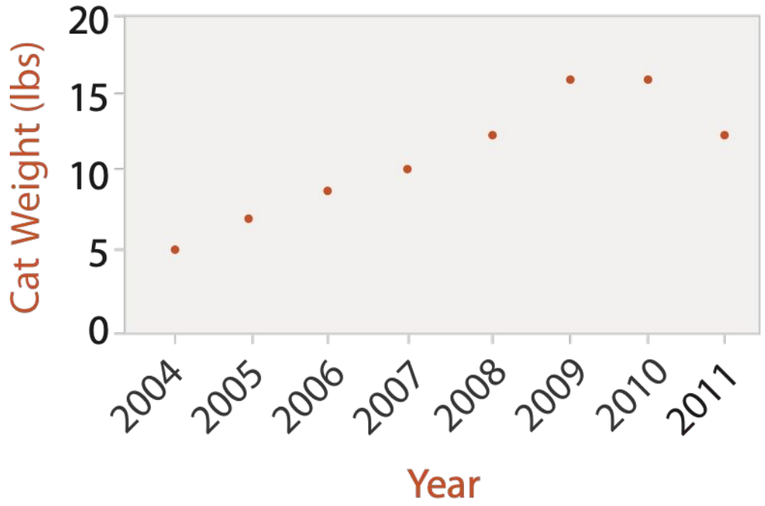
[Stacked Graphs Geometry & Aesthetics. Byron and Wattenberg. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2008) 14(6): 1245–1252, (2008).]

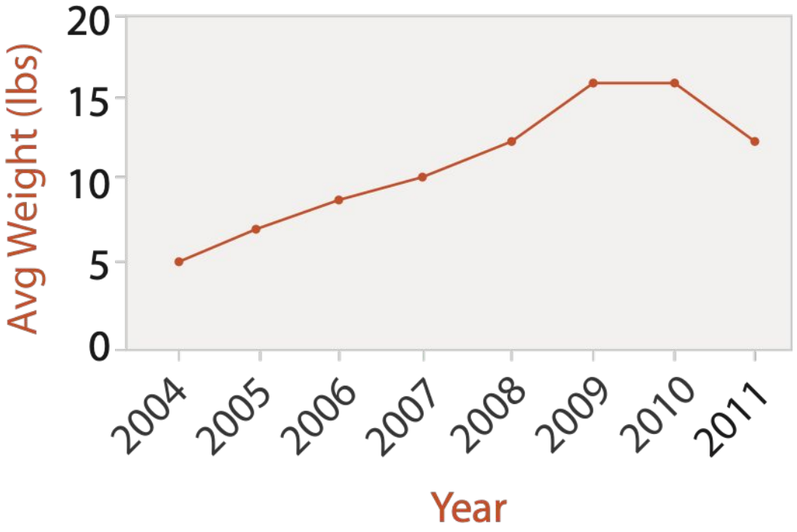


https://flowingdata.com/2008/02/25/ebb-and-flow-of-box-office-receipts-over-past-20-years/

Idiom: **dot / line chart**

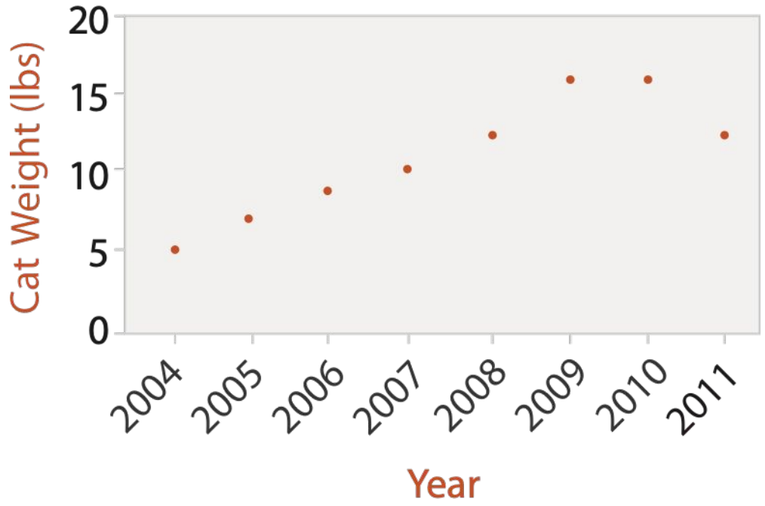
- one key, one value
  - data
    - 2 quant attribs
  - mark: points
    AND line connection marks between them
  - channels
    - aligned lengths to express quant value
    - separated and ordered by key attrib into horizontal regions

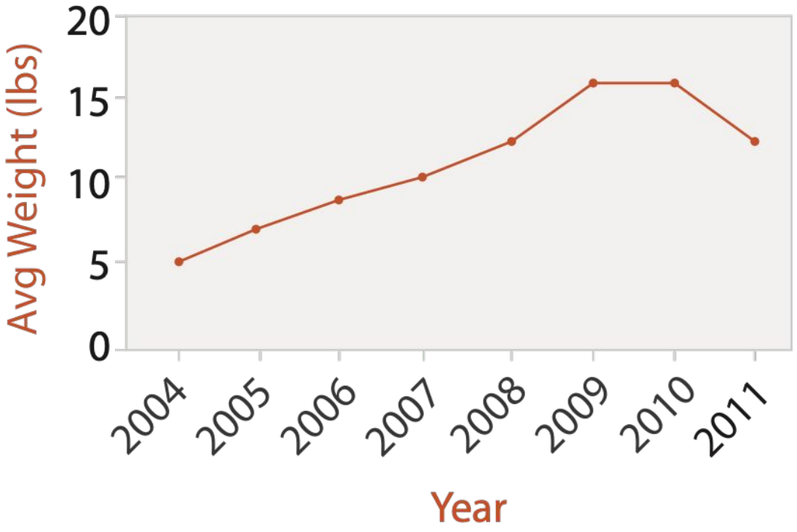




Idiom: **dot / line chart**

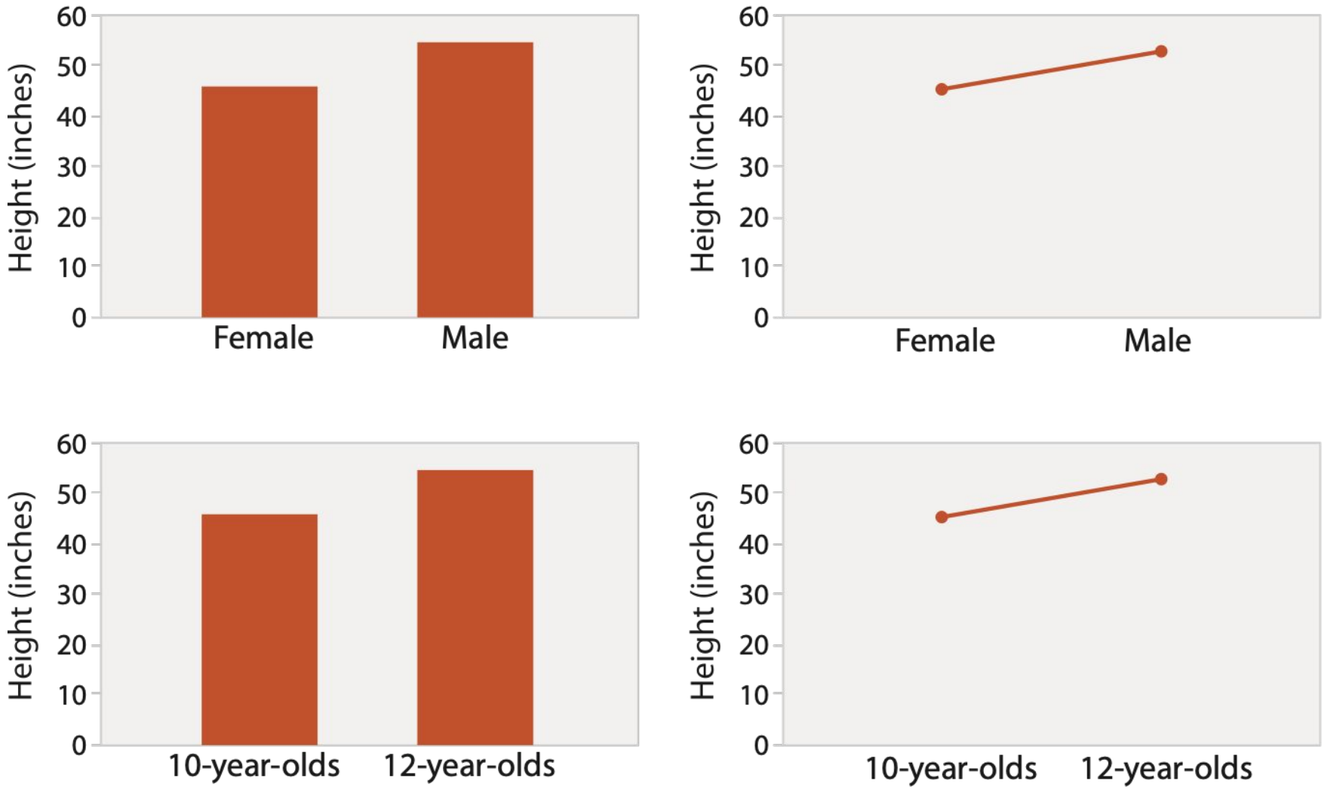
- one key, one value
  - data
    - 2 quant attribs
  - mark: points
    AND line connection marks between them
  - channels
    - aligned lengths to express quant value
    - separated and ordered by key attrib into horizontal regions
  - task
    - find trend
      - connection marks emphasize ordering of items along key axis by explicitly showing relationship between one item and the next
  - scalability
    - hundreds of key levels, hundreds of value levels





Choosing bar vs line charts

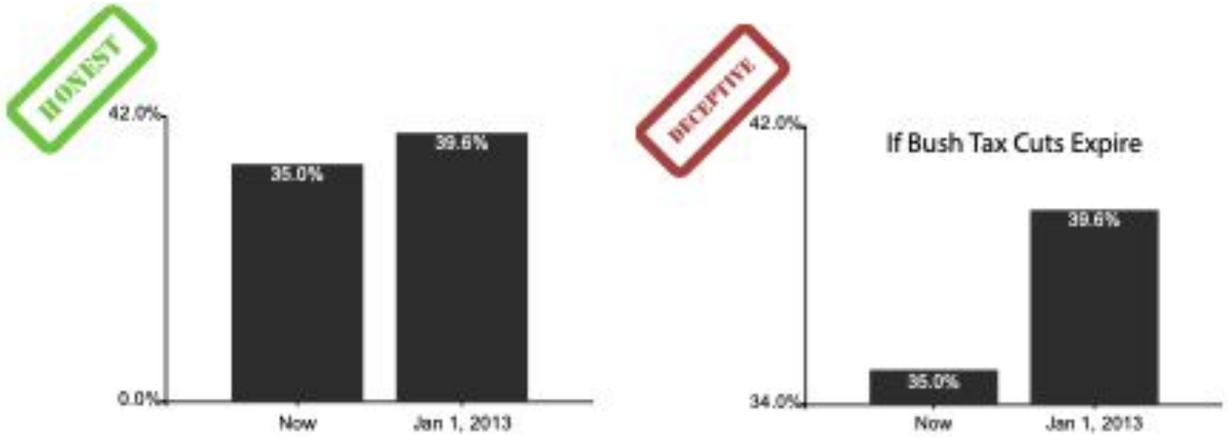
- depends on type of key attrib
  - bar charts if categorical
  - line charts if ordered
- do not use line charts for categorical key attribs
  - violates expressiveness principle
    - implication of trend so strong that it overrides semantics!
      - "The more male a person is, the taller he/she is"



after [Bars and Lines: A Study of Graphic Communication. Zacks and Tversky. Memory and Cognition 27:6 (1999), 1073–1079.]

Chart axes: avoid cropping y axis

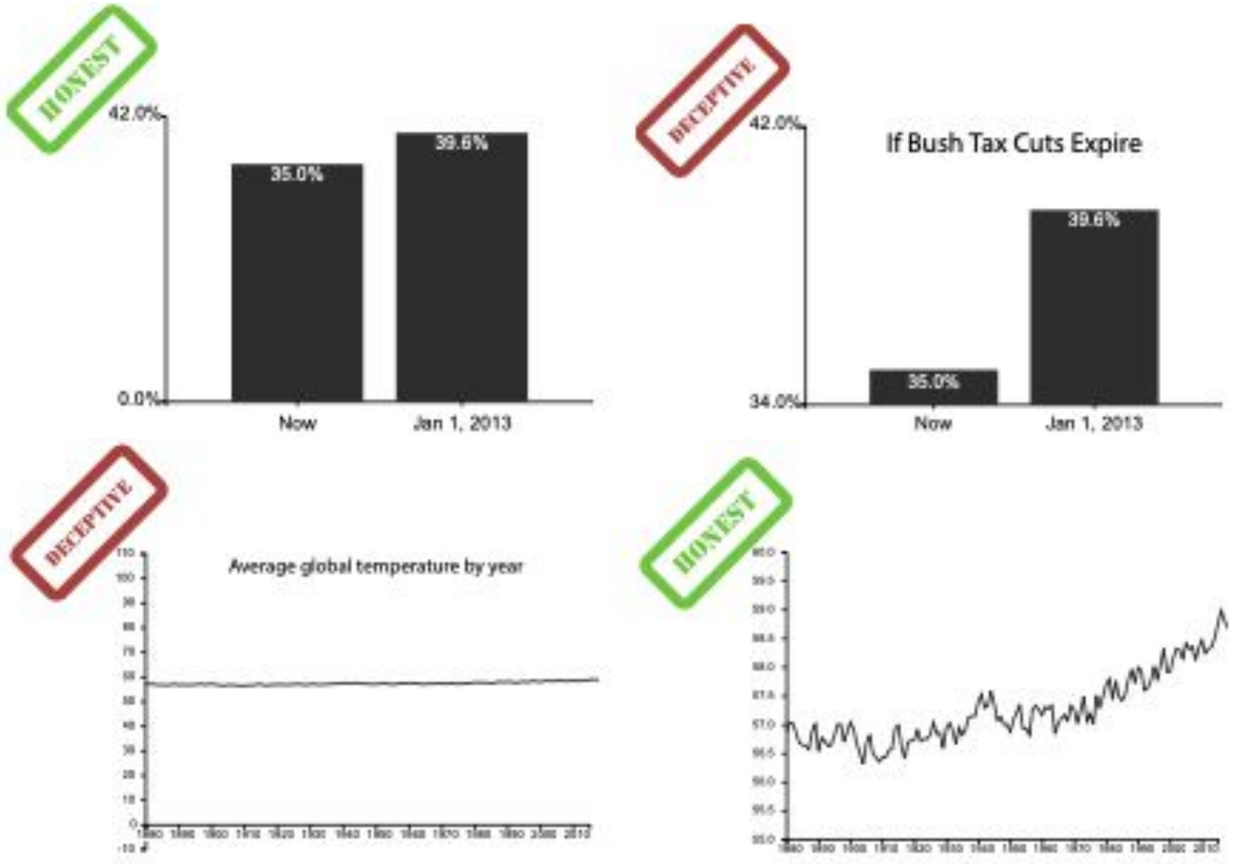
- include 0 at bottom left or slope misleads



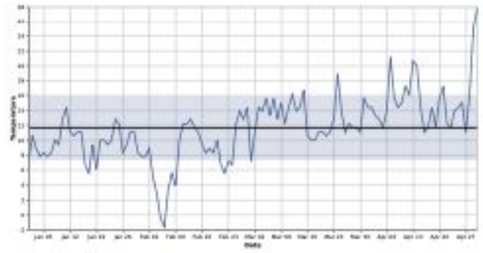
[Truncating the Y-Axis: Threat or Menace?
Correll, Bertini, & Franconeri, CHI 2020.]

Chart axes: avoid cropping y axis

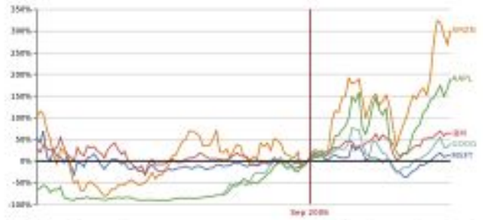
- include 0 at bottom left or slope misleads
  - some exceptions (arbitrary 0, small change matters)



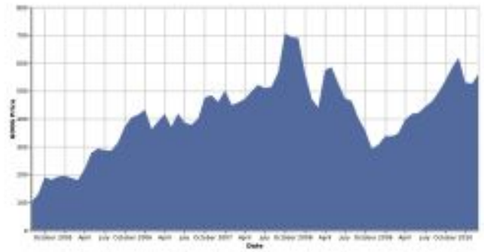
[Truncating the Y-Axis: Threat or Menace? Correll, Bertini, & Franconeri, CHI 2020.]



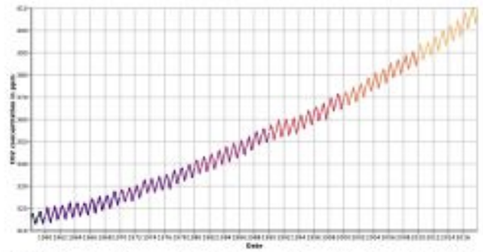
(a) Statistical process charts rely on comparison to an expected value, and so deviations from that value, not from zero, are important



(b) Index charts compare to an indexed value rather than zero.



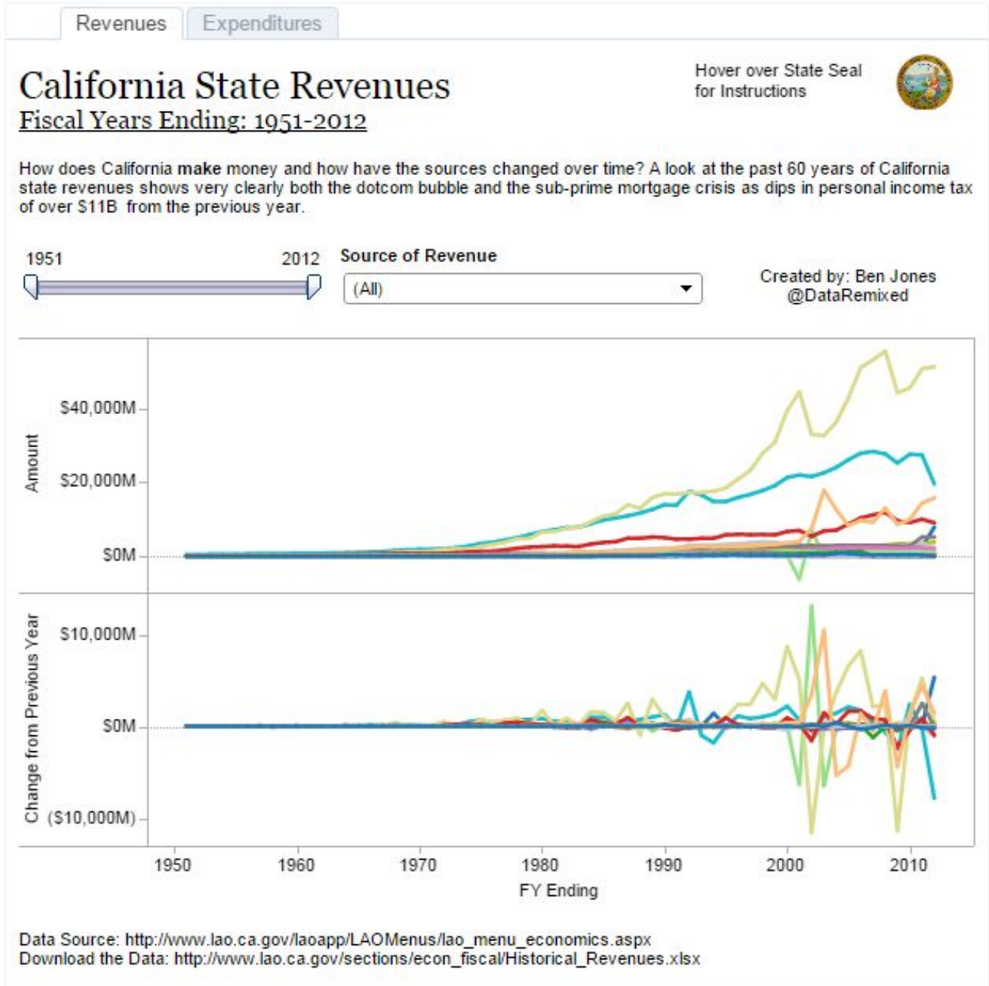
(c) Stock charts must show small differences in stock value, as these can translate to enormous monetary gains or losses.



(d) Climate Anomaly charts rely on both highlighting deviation from a non-zero expected value but also emphasize the potentially disastrous impact of even minute changes in climate.

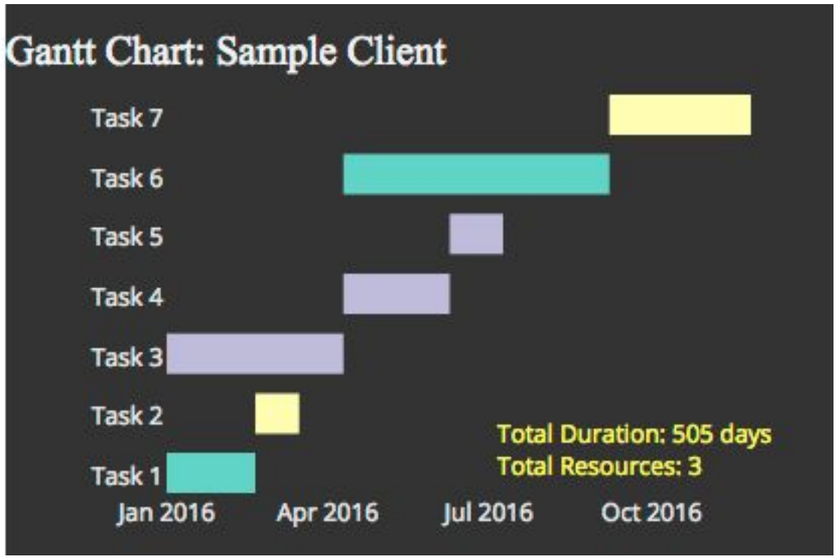
Idiom: **Indexed line charts**

- data: 2 quant attribs
  - 1 key + 1 value
- derived data: new quant value attrib
  - index
  - plot instead of original value
- task: show change over time
  - principle: normalized, not absolute
- scalability
  - same as standard line chart



ublic.tableau.com/profile/ben.jones#!/vizhome/CAStateRevenues/Revenues

Idiom: **Gantt charts**

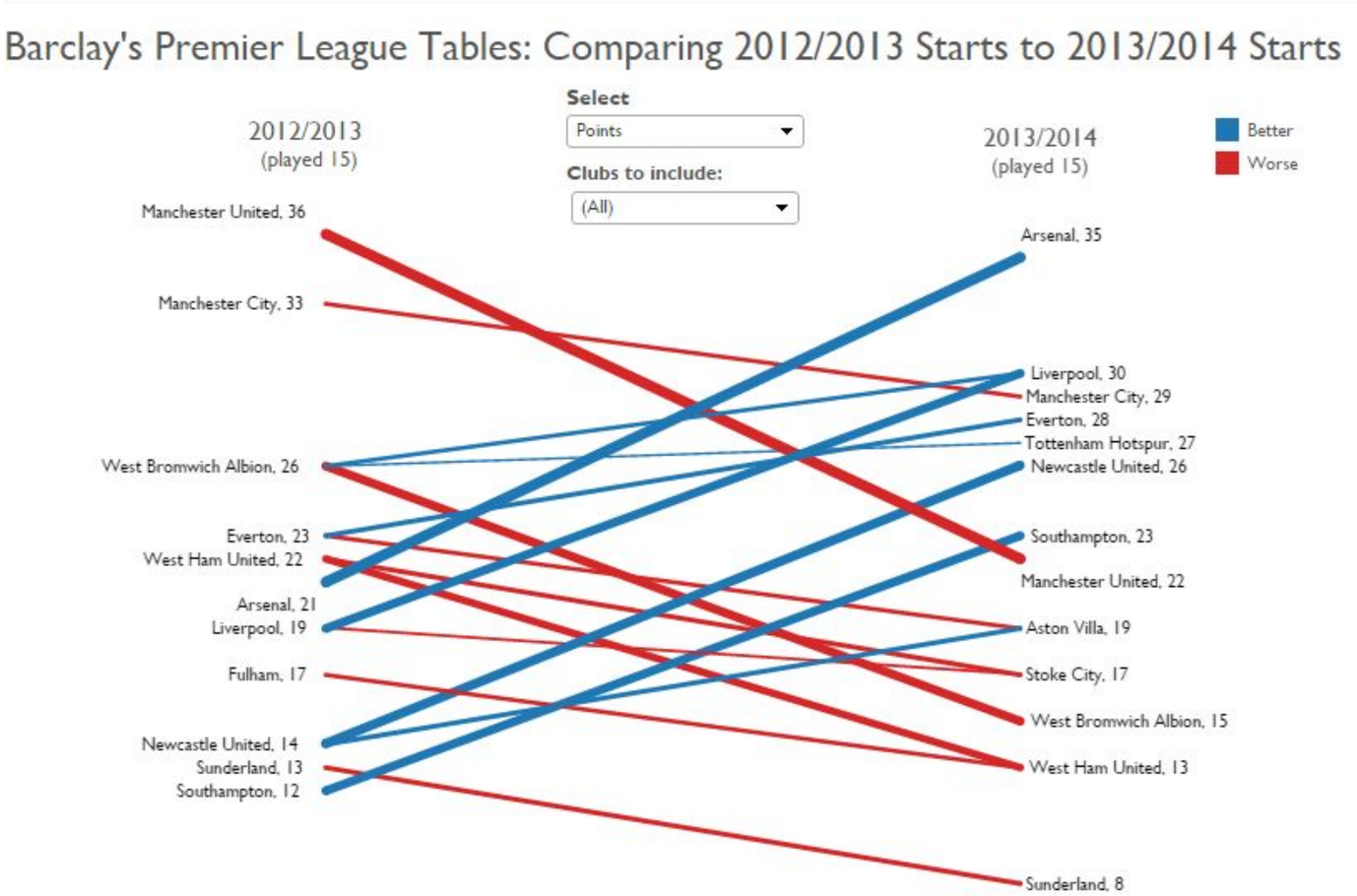


https://www.r-bloggers.com/gantt-charts-in-r-using-plotly/

- one key, two (related) values
  - data
    - 1 categ attrib, 2 quant attribs
  - mark: line
    - length: duration
  - channels
    - horiz position: start time (+end from duration)
  - task
    - emphasize temporal overlaps & start/end dependencies between items
  - scalability
    - dozens of key levels [bars]
    - hundreds of value levels [durations]

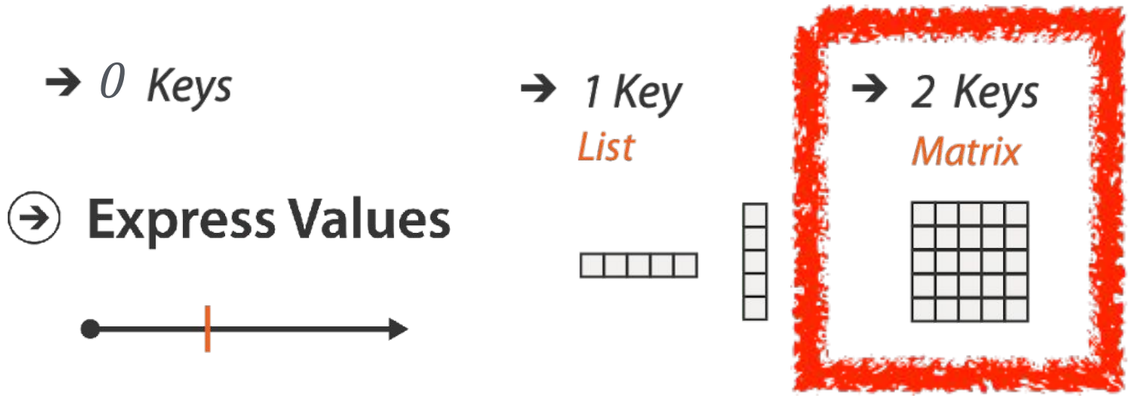
Idiom: **Slopegraphs**

- two values
  - data
    - 2 quant value attribs
    - (1 derived attrib: change magnitude)
  - mark: point + line
    - line connecting mark between pts
  - channels
    - 2 vertical pos: express attrib value
    - (linewidth/size, color)
  - task
    - emphasize changes in rank/value
  - scalability
    - hundreds of value levels
    - dozens of items



https://public.tableau.com/profile/ben.jones#!/vizhome/Slopegraphs/Slopegraphs

2 Keys

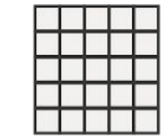


Idiom: **heatmap**

- two keys, one value
  - data
    - 2 categ attribs (gene, experimental condition)
    - 1 quant attrib (expression levels)
  - marks: point
    - separate and align in 2D matrix
      - indexed by 2 categorical attributes
  - channels
    - color by quant attrib
      - (ordered diverging colormap)
  - task
    - find clusters, outliers
  - scalability
    - 1M items, 100s of categ levels, ~10 quant attrib levels

→ *2 Keys*

Matrix



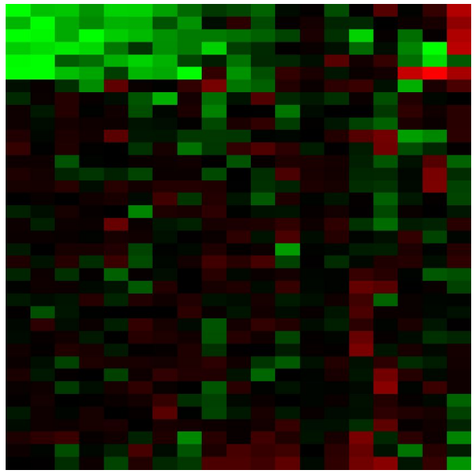
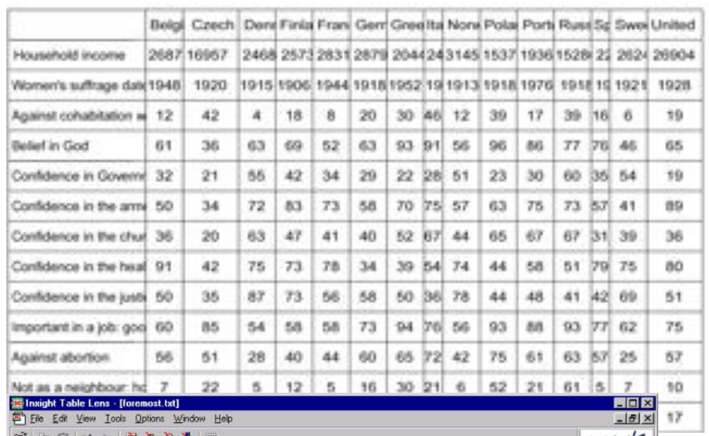
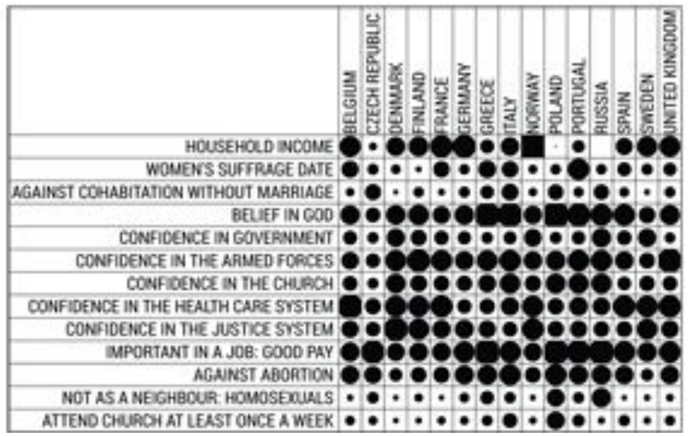


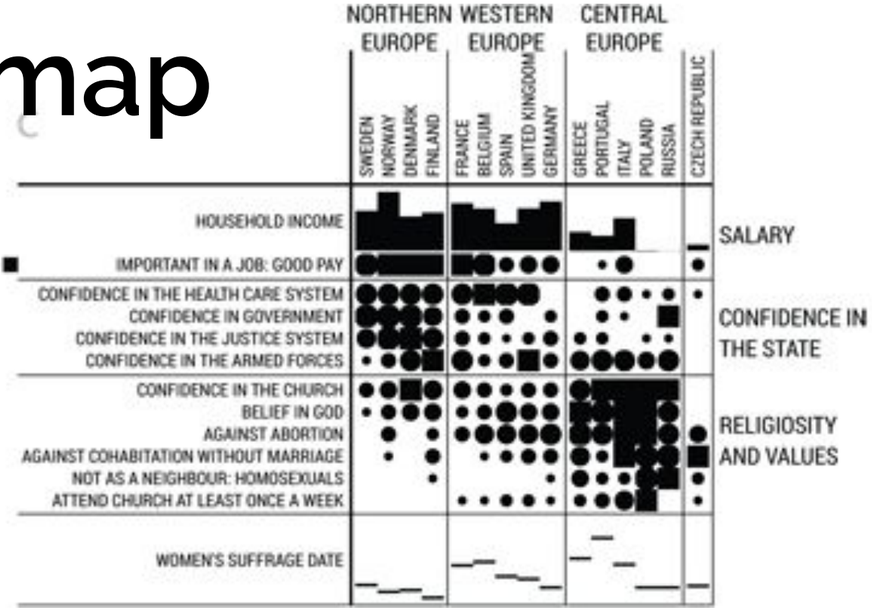
Table heatmap

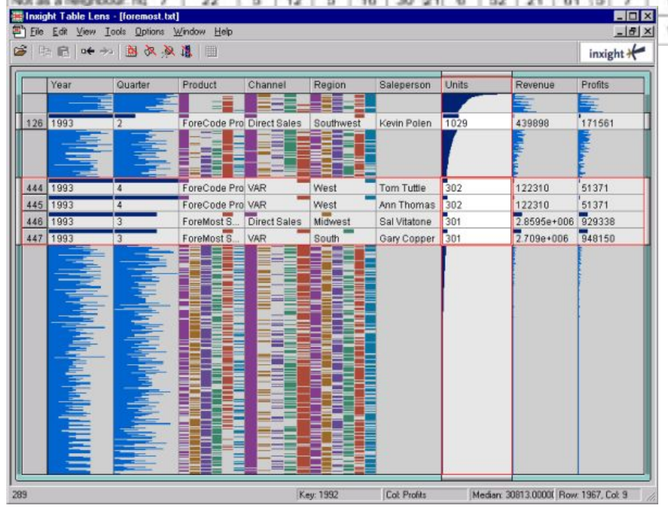
Reorderable table + Heatmap

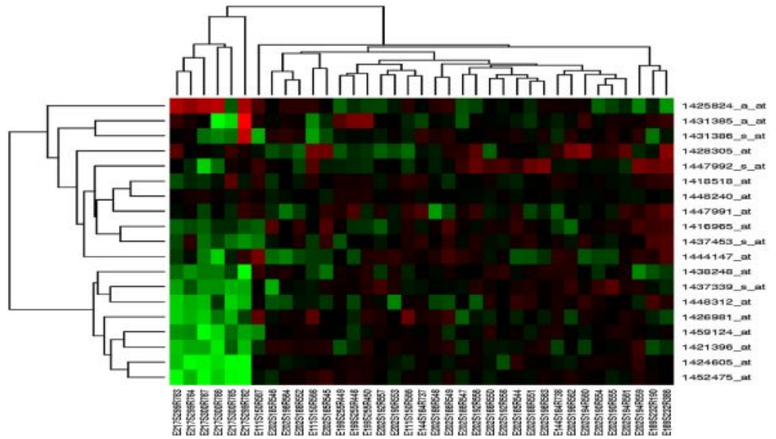


	Belgi	Czech	Denr	Finla	Fran	Gerr	Gree	Ita	Norv	Pola	Porti	Russ	Sp	Swe	United
Household income	2687	16957	2468	2573	2831	2879	2044	24	3145	1537	1936	1528	22	2824	26904
Women's suffrage date	1948	1920	1915	1906	1944	1918	1952	19	1913	1918	1976	1918	19	1921	1928
Against cohabitation w	12	42	4	18	8	20	30	46	12	39	17	39	16	6	19
Belief in God	61	36	63	69	52	63	93	91	56	96	86	77	76	46	65
Confidence in Governm	32	21	55	42	34	29	22	28	51	23	30	60	35	54	19
Confidence in the arm	50	34	72	83	73	58	70	75	57	63	75	73	57	41	89
Confidence in the chur	36	20	63	47	41	40	52	67	44	65	67	67	31	39	36
Confidence in the heal	91	42	75	73	78	34	39	54	74	44	58	51	79	75	80
Confidence in the justi	50	35	87	73	56	58	50	36	78	44	48	41	42	69	51
Important in a job: goo	60	85	54	58	58	73	94	76	56	93	88	93	77	62	75
Against abortion	56	51	28	40	44	60	65	72	42	75	61	63	57	25	57
Not as a neighbour: ho	7	22	5	12	5	16	30	21	6	52	21	61	5	7	10
Attend church at least	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17

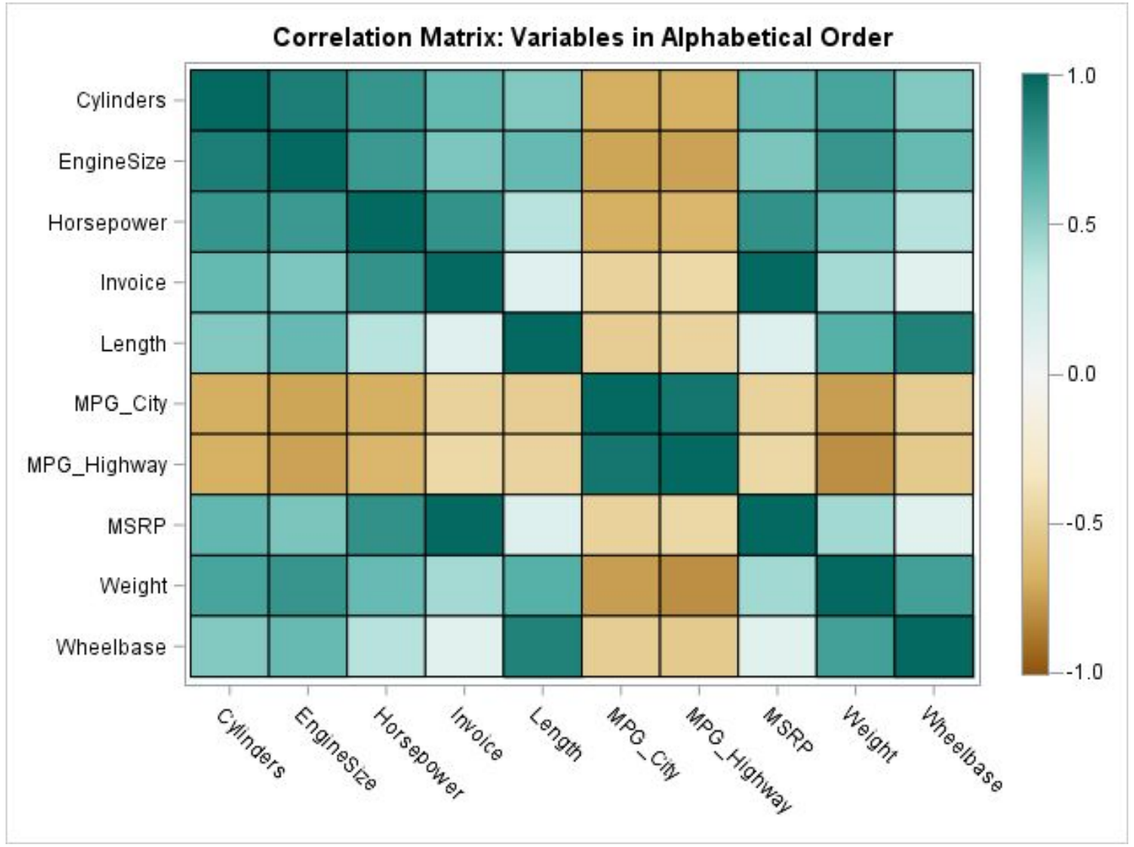


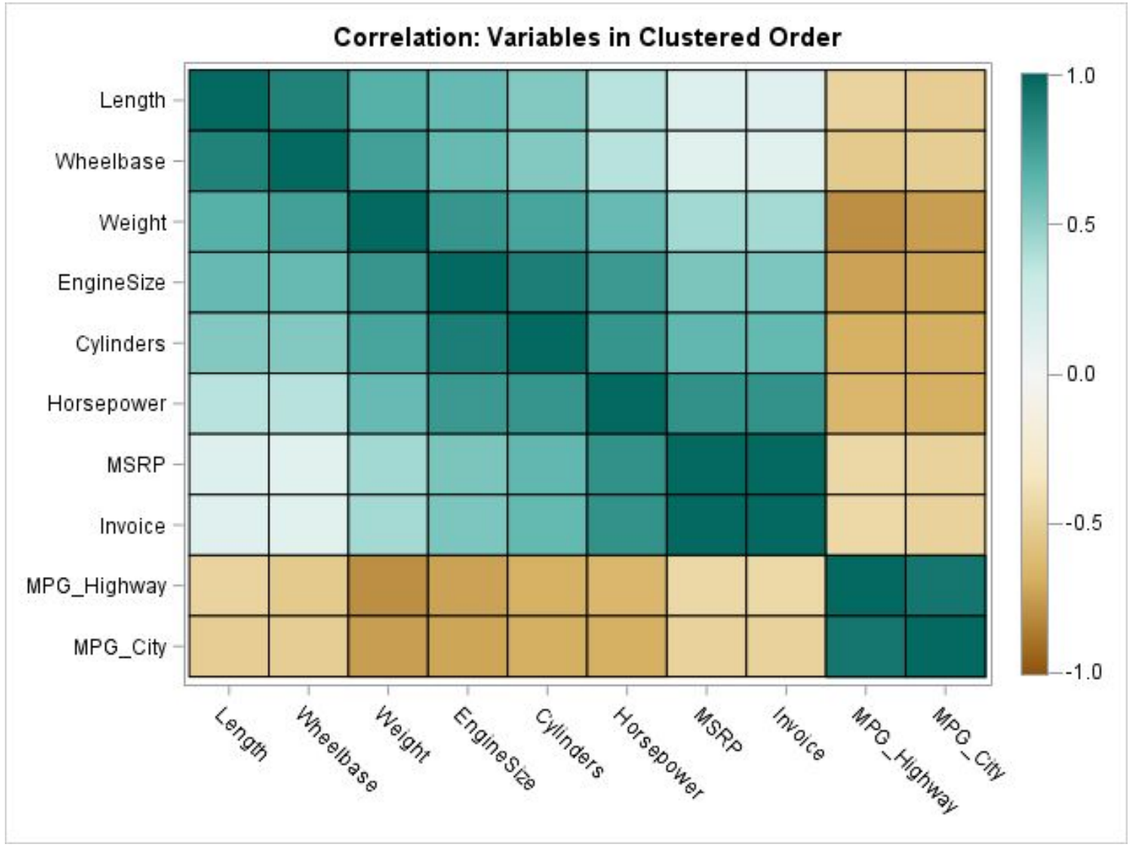






Heatmap reordering





https://blogs.sas.com/content/iml/2018/05/02/reorder-variables-correlation-heat-map.html

Idiom: **cluster heatmap**

- in addition
  - derived data
    - 2 cluster hierarchies
  - dendrogram
    - parent-child relationships in tree with connection line marks
    - leaves aligned so interior branch heights easy to compare
  - heatmap
    - marks (re-)ordered by cluster hierarchy traversal
    - task: assess quality of clusters found by automatic methods

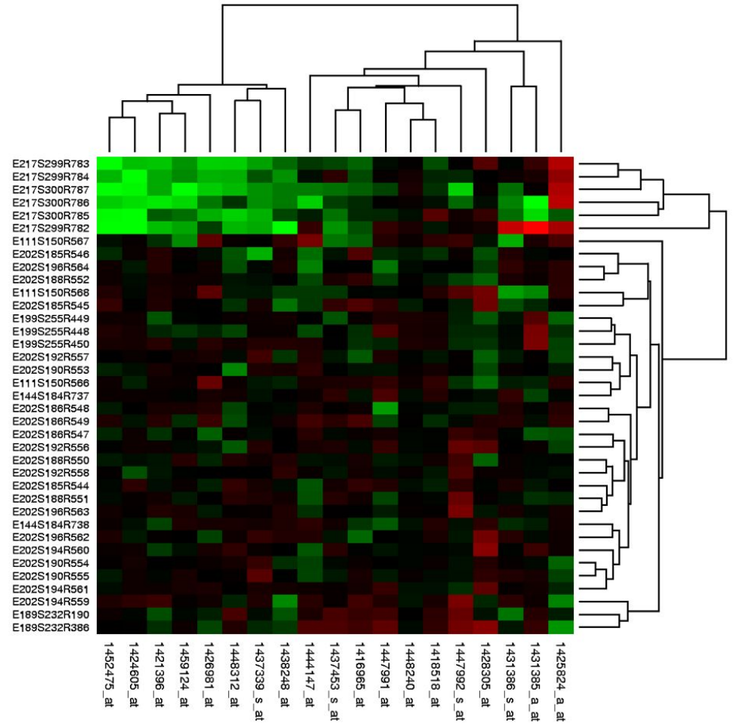
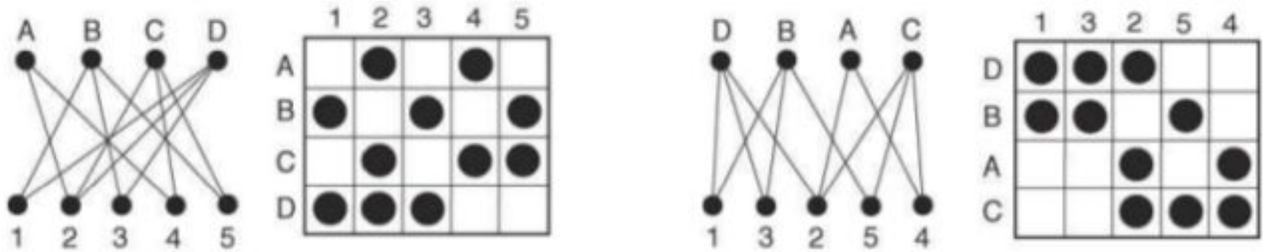
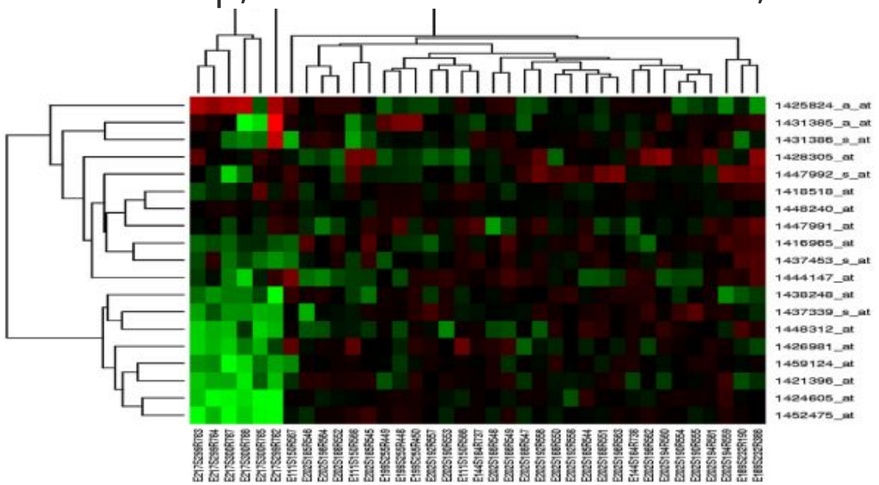


Table + Heatmap

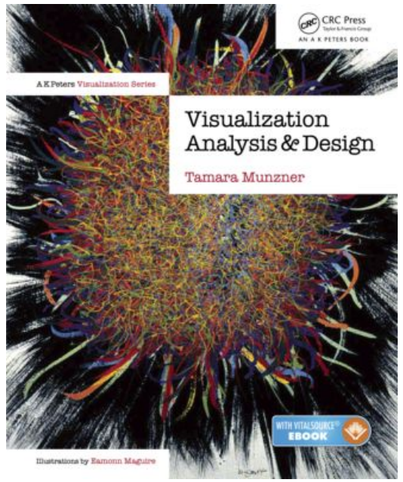
The table order is essential

- Michael Behrisch, Benjamin Bach, Nathalie Henry Riche, Tobias Schreck, Jean-Daniel Fekete. Matrix Reordering Methods for Table and Network Visualization. *Computer Graphics Forum*, Wiley, 2016, 35, pp.24. ⟨10.1111/cgf.12935⟩. ⟨hal-01326759⟩
- **Nathan van Beusekom, Wouter Meulemans, Bettina Speckmann, Simultaneous Matrix Orderings for Graph Collections, IEEE Transactions on Visualization and Computer Graphics, to appear, 2022**
- Leland Wilkinson & Michael Friendly (2009) The History of the Cluster Heat Map, The American Statistician, 63:2, 179-184, DOI: 10.1198/tas.2009.0033





Visualization Analysis & Design

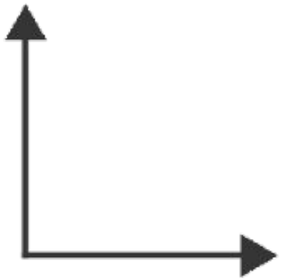


Tables (Ch 7) II

Tamara Munzner
Department of Computer Science
University of British Columbia
@tamaramunzner

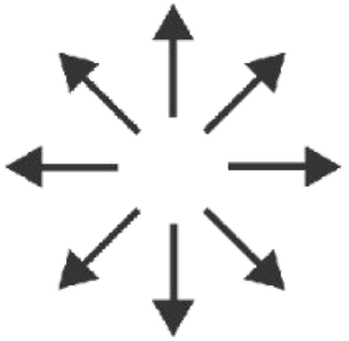
Axis Orientation

- Rectilinear



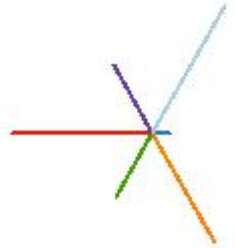
- Parallel

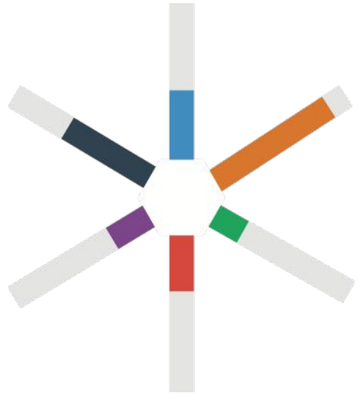
- Radial

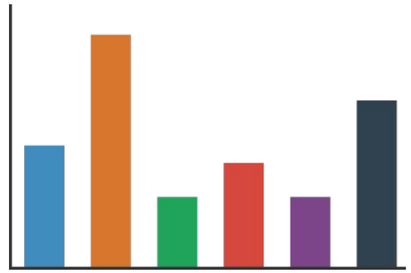


Idioms: **radial bar chart, star plot**

- star plot
  - line mark, radial axes meet at central point
- radial bar chart
  - line mark, radial axes meet at central ring
  - channels: length, angle/orientation
- bar chart
  - rectilinear axes, aligned vertically







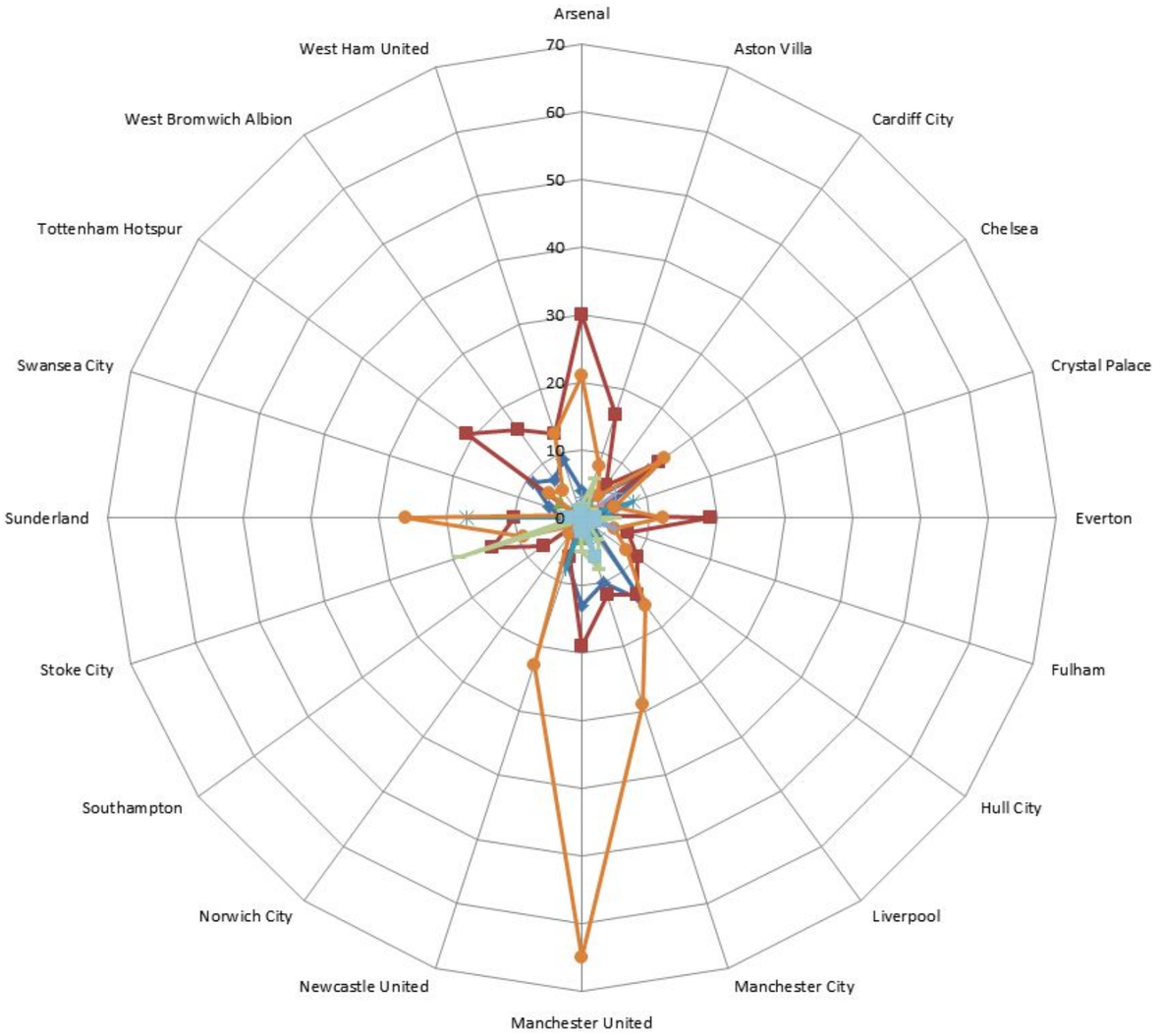
- accuracy
  - length not aligned with radial layouts
    - less accurately perceived than rectilinear aligned

[Vismon: Facilitating Risk Assessment and Decision Making In Fisheries Management. Booshehrian, Möller, Peterman, and Munzner. Technical Report TR 2011-04, Simon Fraser University, School of Computing Science, 2011.]

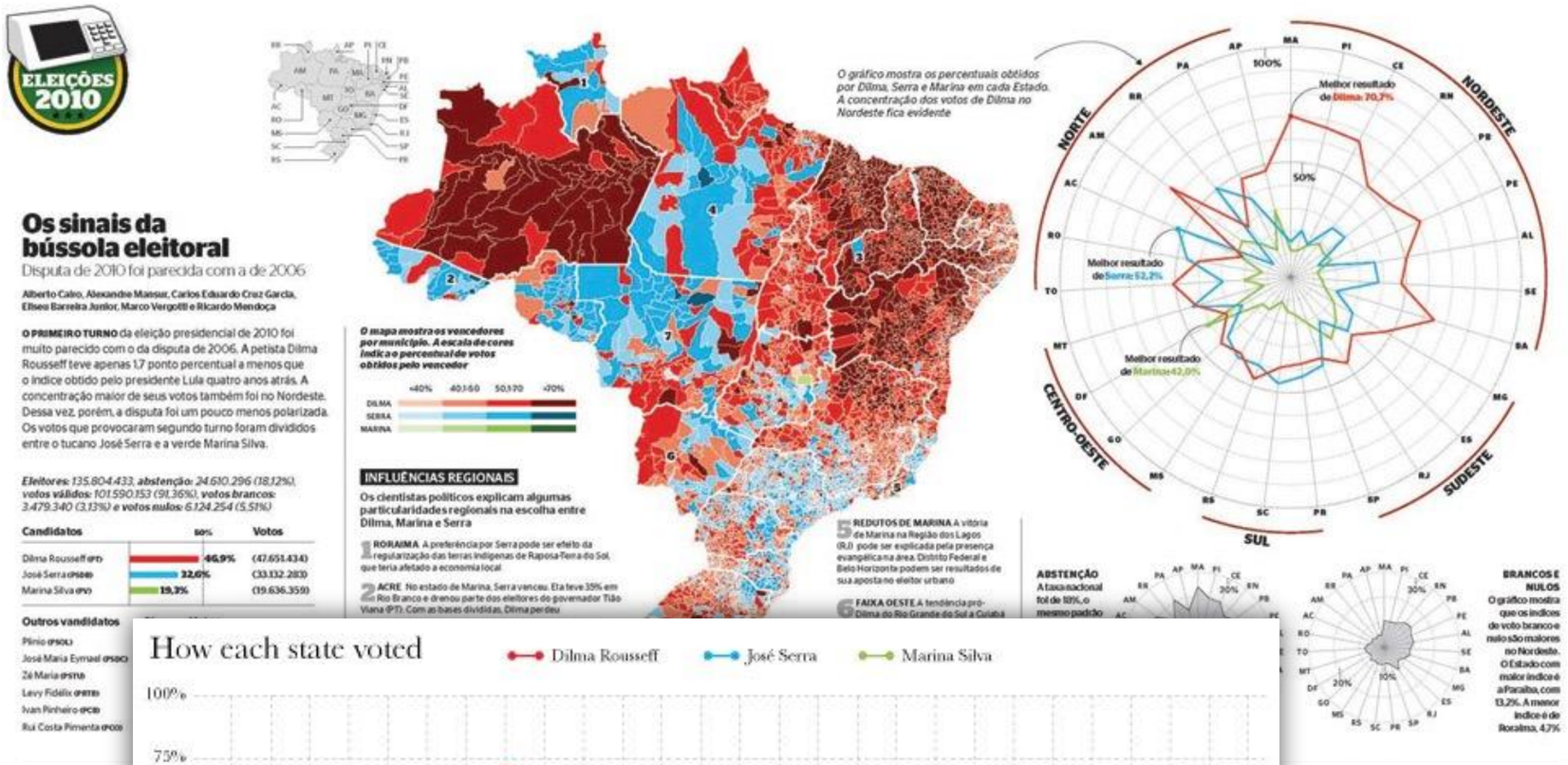
Idiom: **radar plot**

- radial line chart
  - point marks, radial layout
  - connecting line marks

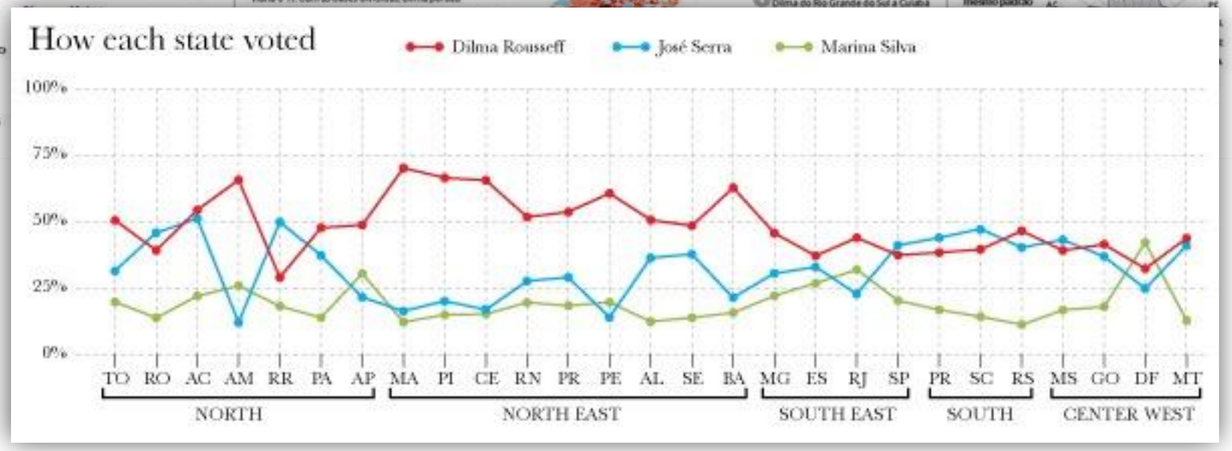
- avoid unless data is cyclic



“Radar graphs: Avoid them (99.9% of the time)”



original
difficult to interpret

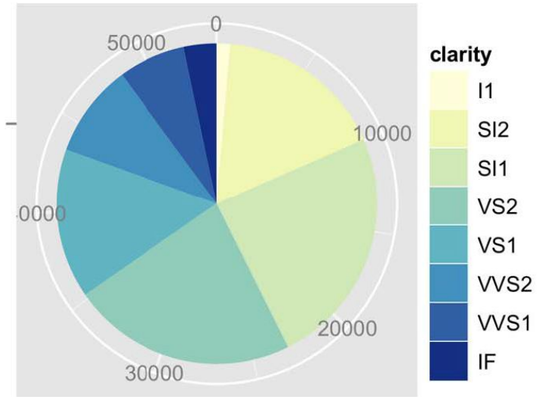


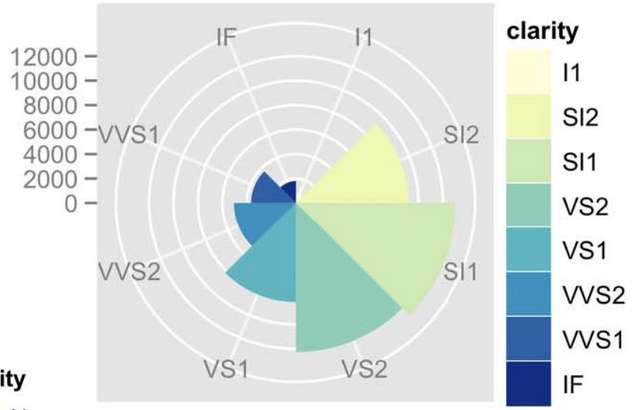
redesign for
rectilinear

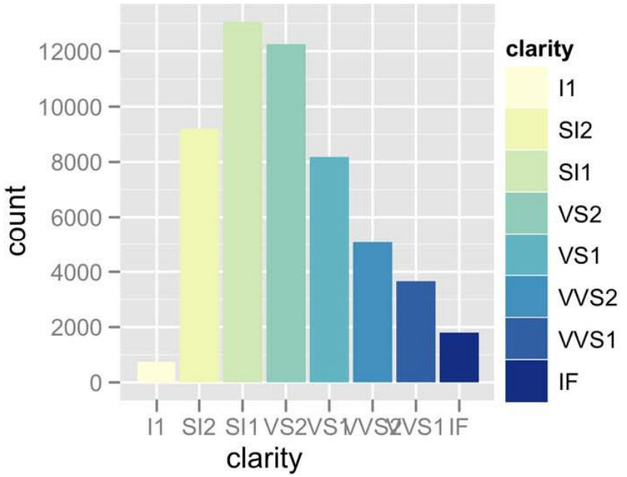
www.thefunctionalart.com/2012/11/radar-graphs-avoid-them-999-of-time.html

Idioms: **pie chart, coxcomb chart**

- pie chart
  - **interlocking area** marks with angle channel: **2D area varies**
    - separated & ordered radially, uniform height
  - accuracy: area less accurate than rectilinear aligned line length
  - **task: part-to-whole judgements**
- coxcomb chart
  - line marks with length channel: **1D length varies**
    - separated & ordered radially, uniform width
  - direct analog to radial bar charts
- data
  - 1 categ key attrib, 1 quant value attrib



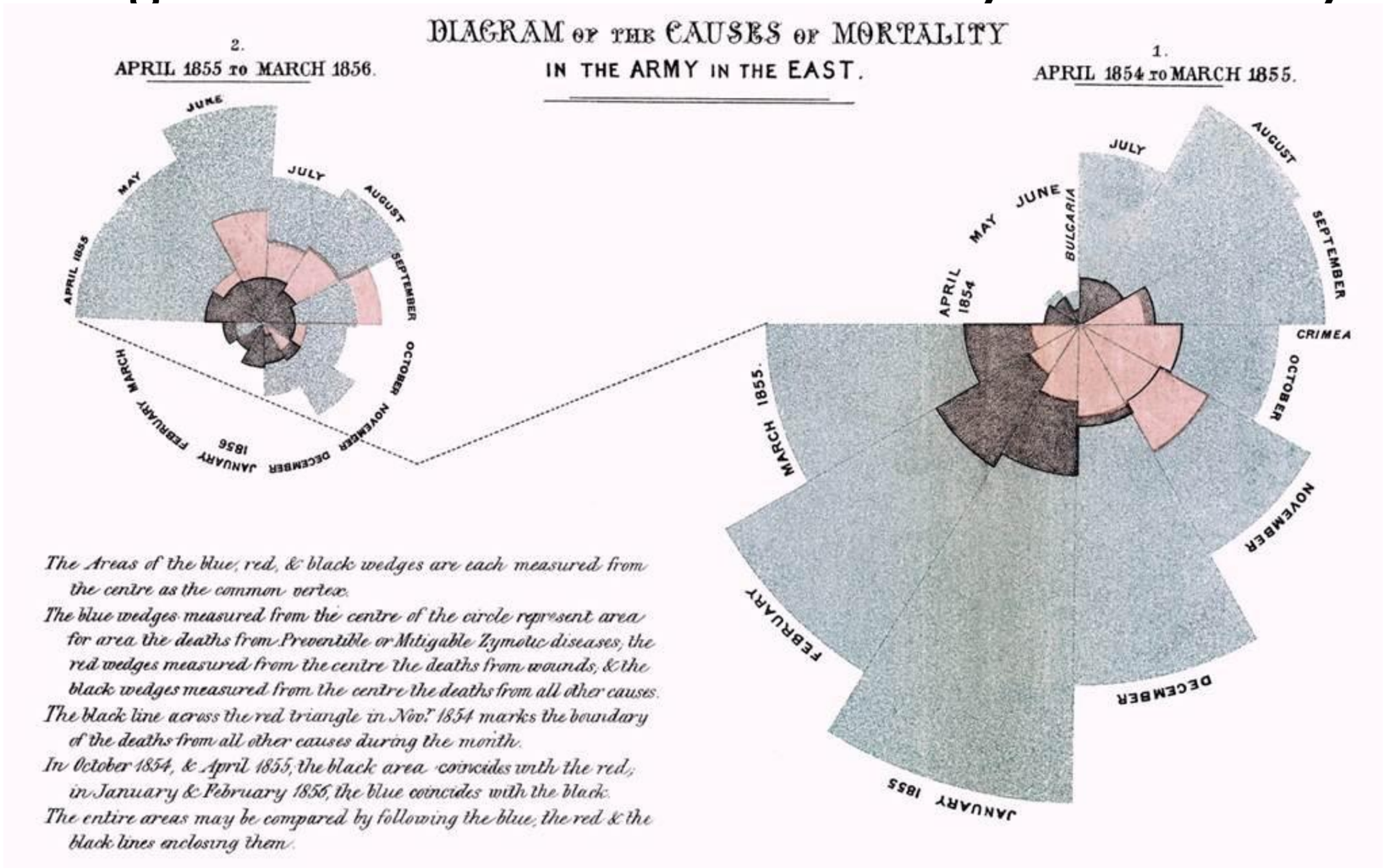




[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

Coxcomb / nightingale rose / polar area chart

- invented by Florence Nightingale:
  Diagram of the Causes of Mortality in the Army in the East

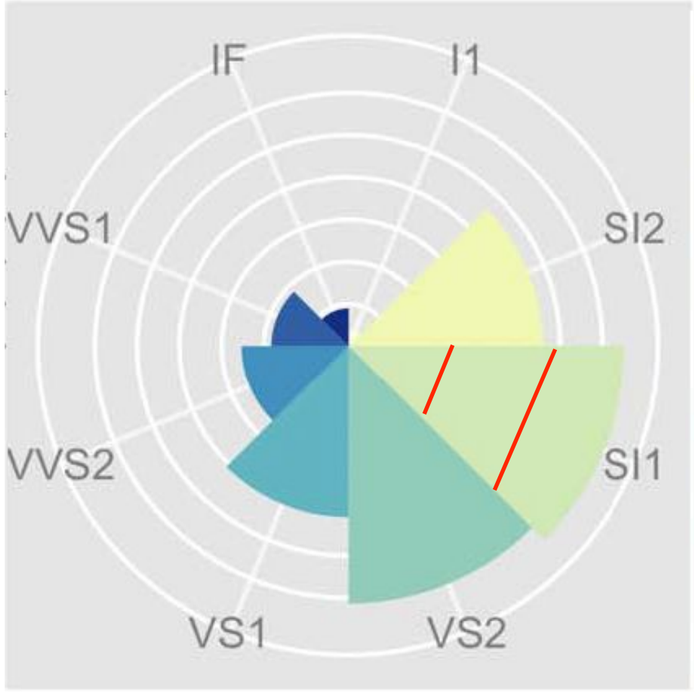


Coxcomb: perception

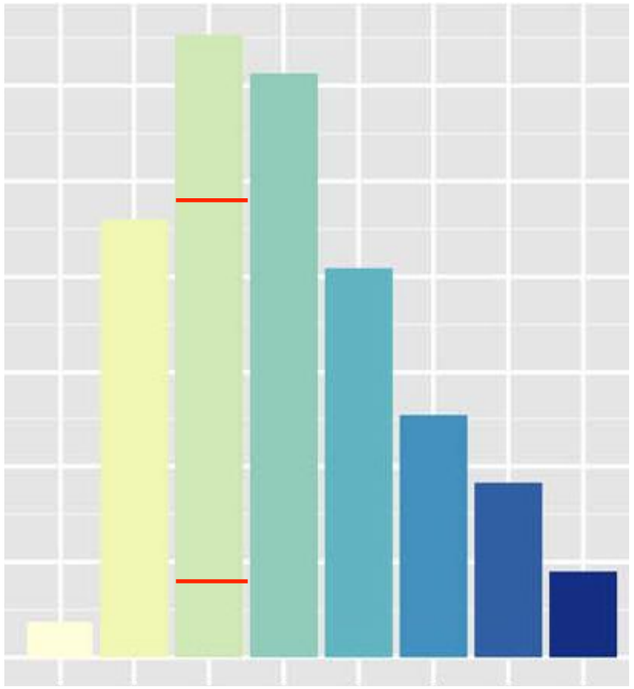
- encode: 1D length
- decode/perceive: 2D area

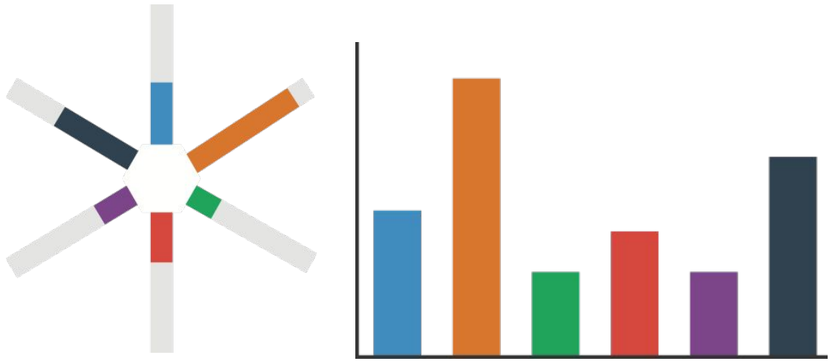
- nonuniform line/sector width as length increases
  - so area variation is nonlinear wrt line mark length!

- bar chart safer: uniform width, so area is linear with line mark length
  - both radial & rectilinear cases



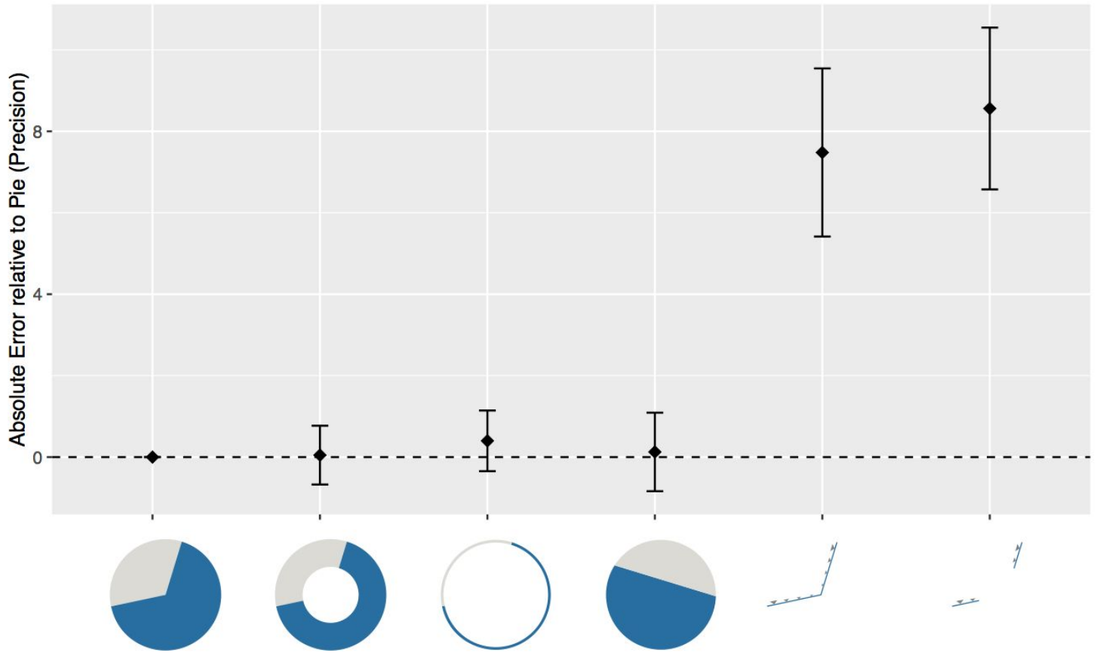
nonuniform width as length increases


uniform width as length increases


radial & rectilinear bars: uniform width as length increases

Pie charts: perception

- some empirical evidence that people respond to arc length
  - decode/perceive: not angles
  - maybe also areas?…
- donut charts no worse than pie charts



I Data Encodings in Pie and Donut Charts. Skau and Kosara. Proc. EuroVis 2016.]

gereyes.org/blog/2016/an-illustrated-tour-of-the-pie-chart-study-results

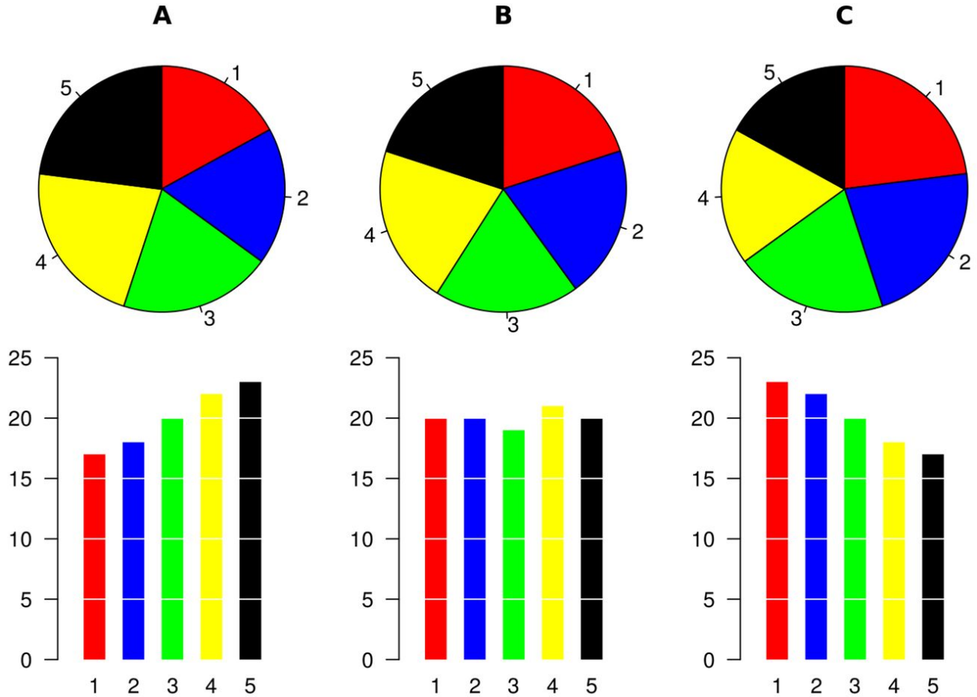
Pie charts: best practices

- not so bad for two (or few) levels, for part-to-whole task

https://eagereyes.org/pie-charts

Pie charts: best practices

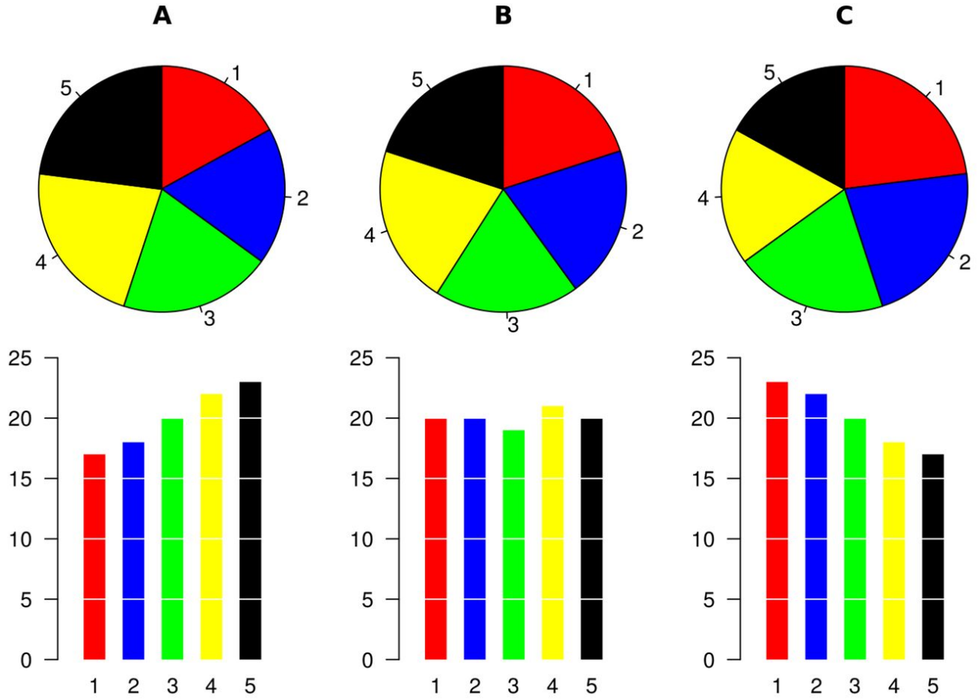
- not so bad for two (or few) levels, for part-to-whole task
- dubious for several levels if details matter

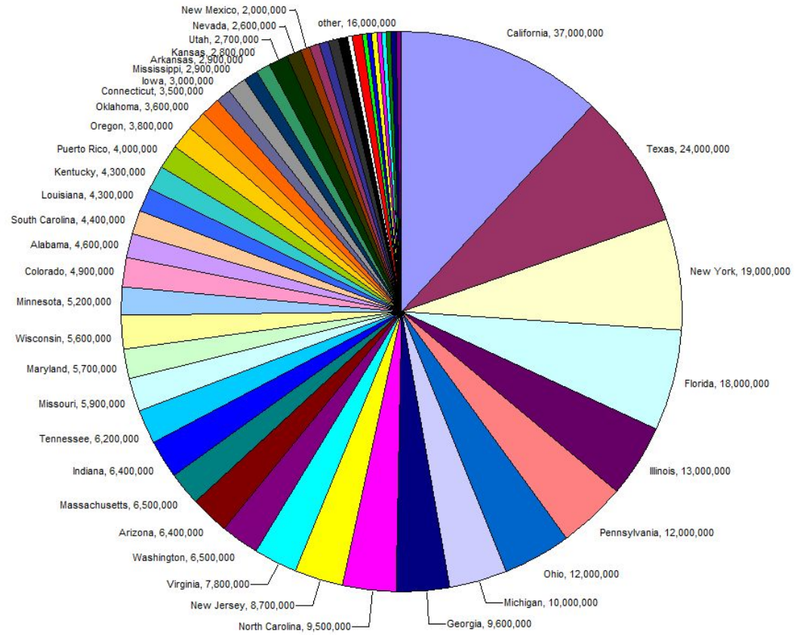


https://eagereyes.org/pie-charts

Pie charts: best practices

- not so bad for two (or few) levels, for part-to-whole task
- dubious for several levels if details matter
- terrible for many levels

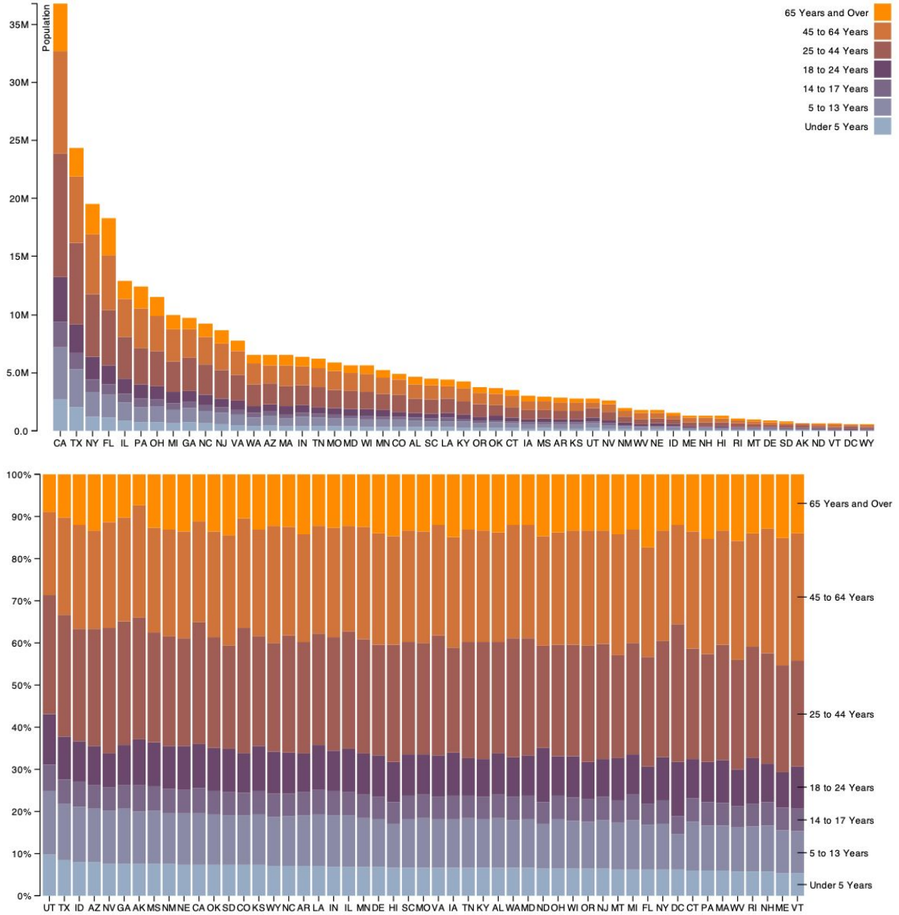


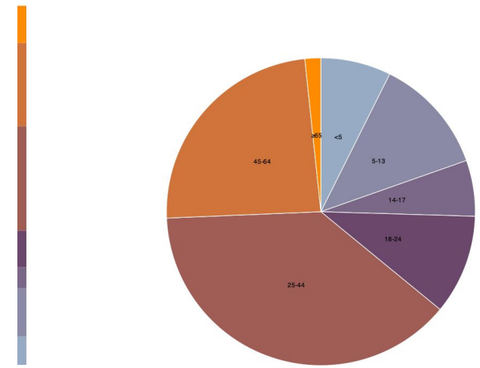


https://eagereyes.org/pie-charts

Idioms: **normalized stacked bar chart**

- task
  - part-to-whole judgements
- normalized stacked bar chart
  - stacked bar chart, normalized to full vert height
  - single stacked bar equivalent to full pie
    - high information density: requires narrow rectangle
- pie chart
  - information density: requires large circle





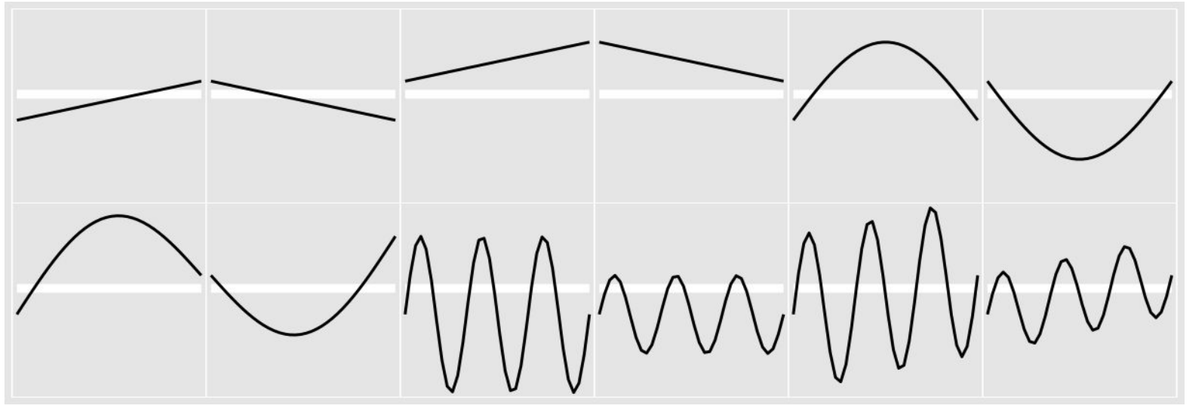
http://bl.ocks.org/mbostock/3886208

,

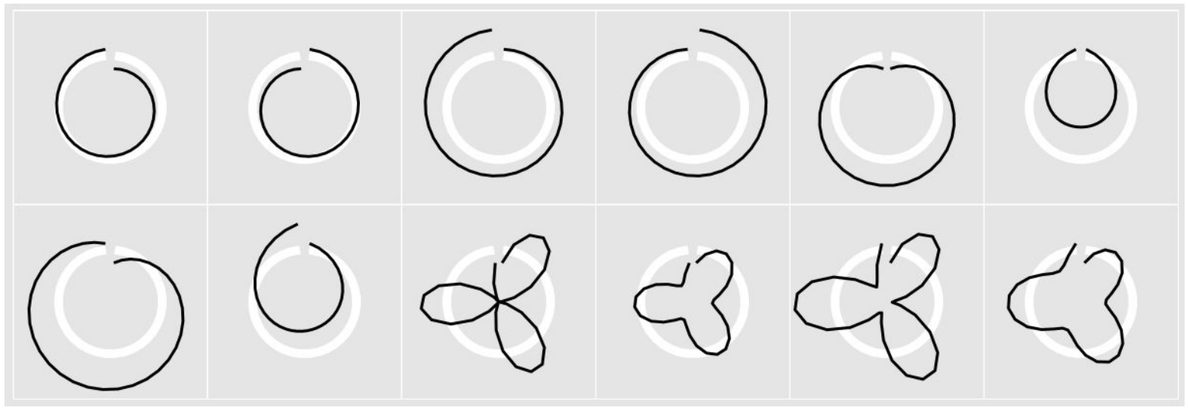
http://bl.ocks.org/mbostock/3887235

Idiom: **glyphmaps**

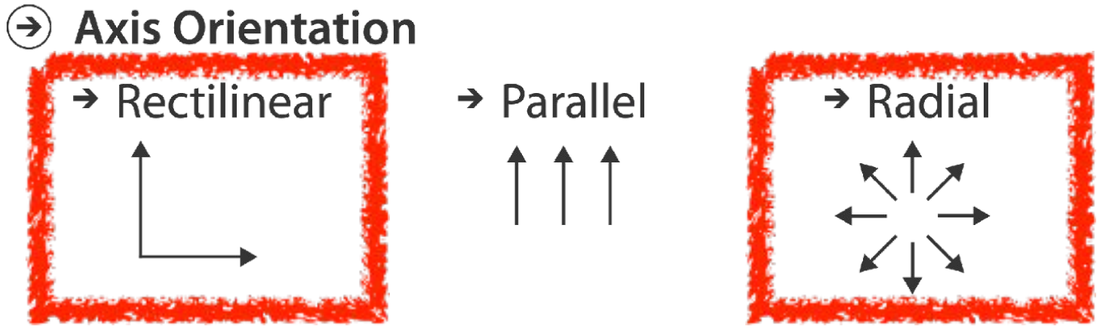
- rectilinear good for linear vs nonlinear trends



- radial good for cyclic patterns
  - evaluating periodicity

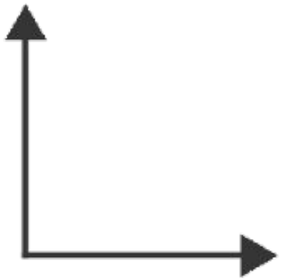


[Glyph-maps for Visually Exploring Temporal Patterns in Climate Data and Models. Wickham, Hofmann, Wickham, and Cook. Environmetrics 23:5 (2012), 382–393.]



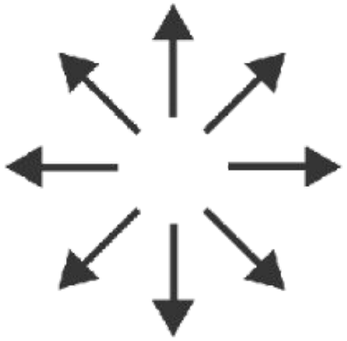
Axis Orientation

- Rectilinear



- Parallel

- Radial



Idiom: **SPLOM**

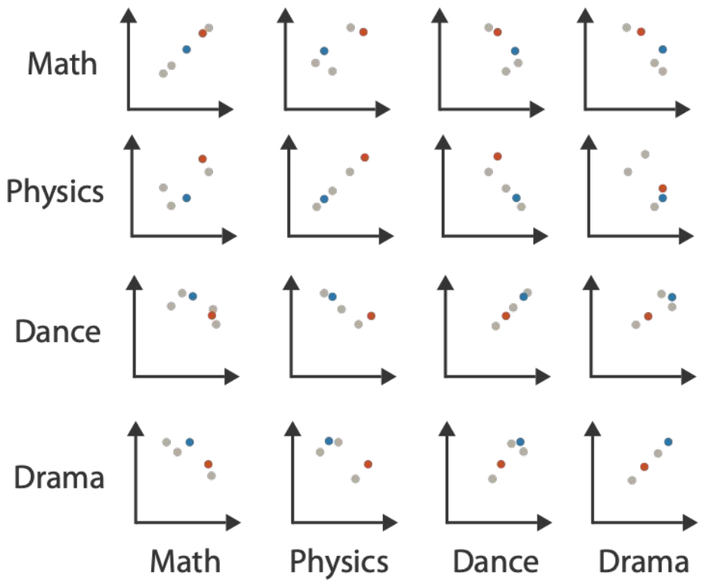
- scatterplot matrix (SPLOM)
  - rectilinear axes, point mark
  - all possible pairs of axes
  - scalability
    - one dozen attribs
    - dozens to hundreds of items

Wilkinson et al., 2005

Idioms: **parallel coordinates**

- scatterplot limitation
  - visual representation with orthogonal axes
  - can show only two attributes with spatial position channel

Scatterplot Matrix



Table

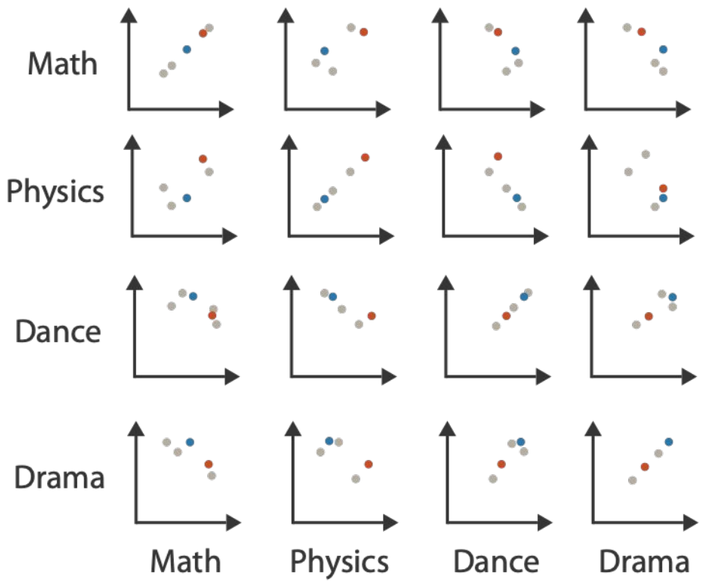
Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

after [Visualization Course Figures. McGuffin, 2014. http://www.michaelmcguffin.com/courses/vis/]

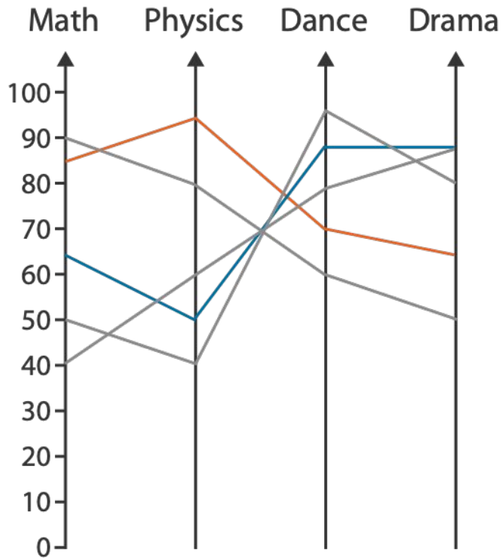
Idioms: **parallel coordinates**

- scatterplot limitation
  - visual representation with orthogonal axes
  - can show only two attributes with spatial position channel
- alternative: line up axes in parallel to show many attributes with position
  - item encoded with a line with n segments
  - n is the number of attributes shown
- parallel coordinates
  - parallel axes, jagged line for item
  - rectilinear axes, item as point
    - axis ordering is major challenge
  - scalability
    - dozens of attribs
    - hundreds of items

Scatterplot Matrix



Parallel Coordinates



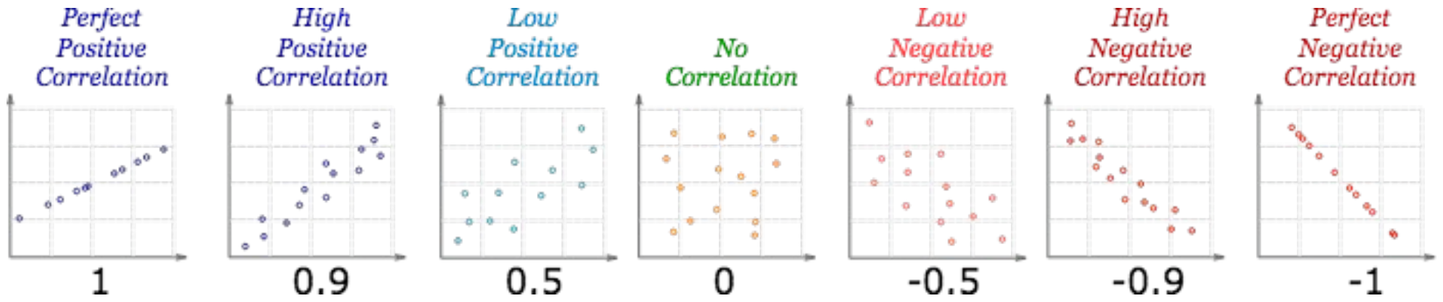
Table

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

after [Visualization Course Figures. McGuffin, 2014. http://www.michaelmcguffin.com/courses/vis/]

Task: Correlation

- scatterplot matrix
  - positive correlation
    - diagonal low-to-high
  - negative correlation
    - diagonal high-to-low
  - uncorrelated: spread out
- parallel coordinates
  - positive correlation
    - parallel line segments
  - negative correlation
    - all segments cross at halfway point
  - uncorrelated
    - scattered crossings



https://www.mathsisfun.com/data/scatter-xy-plots.html

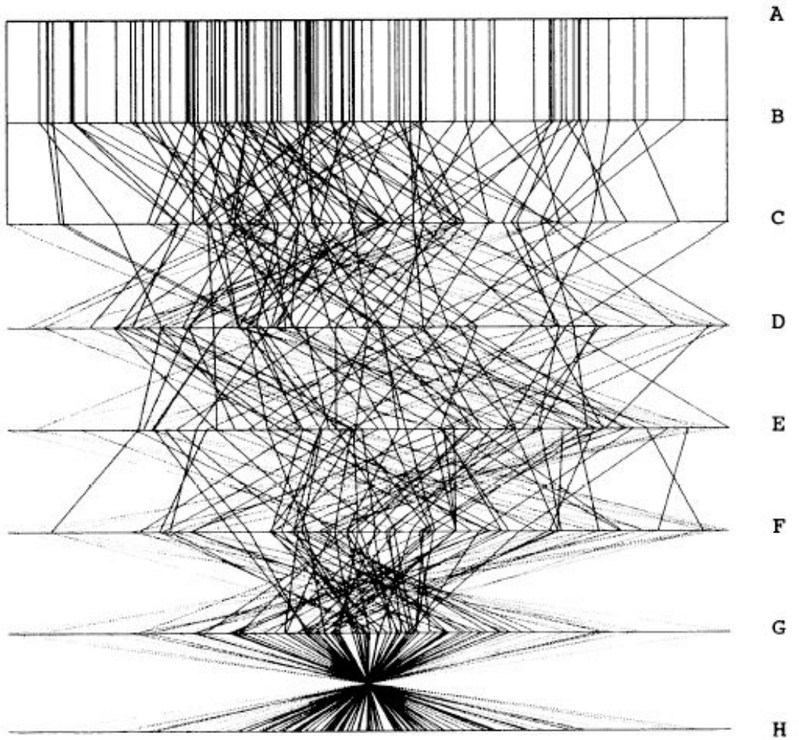
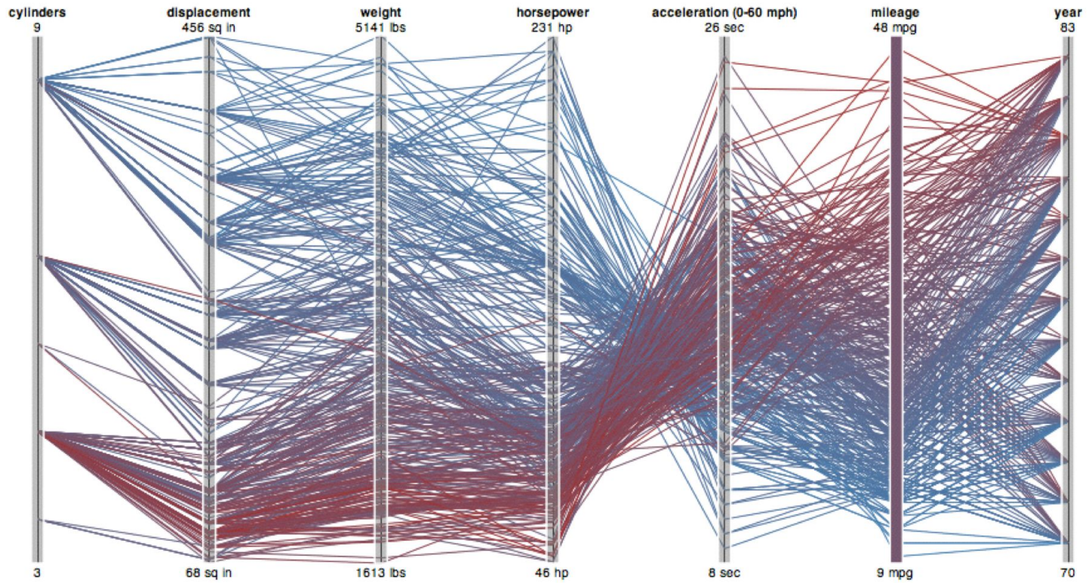


Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of $\rho = 1, .8, .2, 0, -.2, -.8,$ and -1.

[Hyperdimensional Data Analysis Using Parallel Coordinates. Wegman. Journ. American Statistical Association 85:411 (1990), 664–675.]

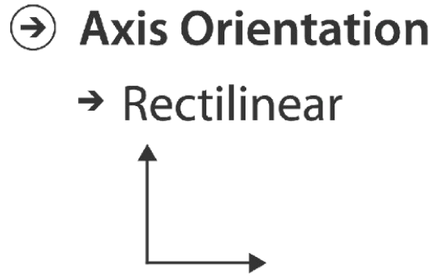
Parallel coordinates, limitations

- visible patterns only between neighboring axis pairs
- how to pick axis order?
  - usual solution: reorderable axes, interactive exploration
  - same weakness as many other techniques
    - downside of interaction: human-powered search
  - some algorithms proposed, none fully solve



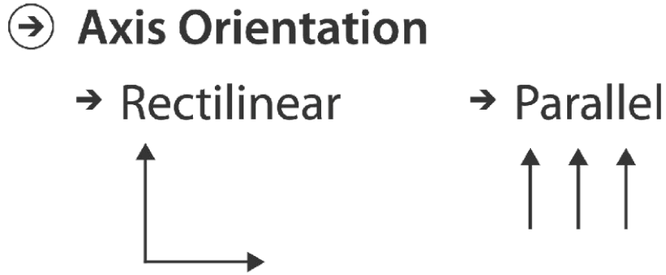
Orientation limitations

- rectilinear: scalability wrt #axes
  - 2 axes best, 3 problematic, 4+ impossible



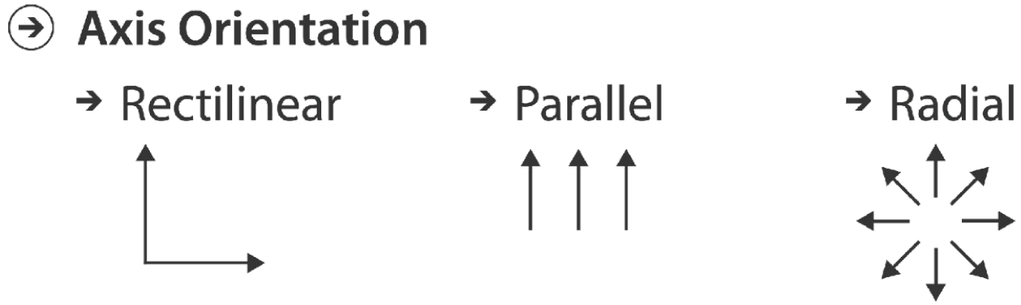
Orientation limitations

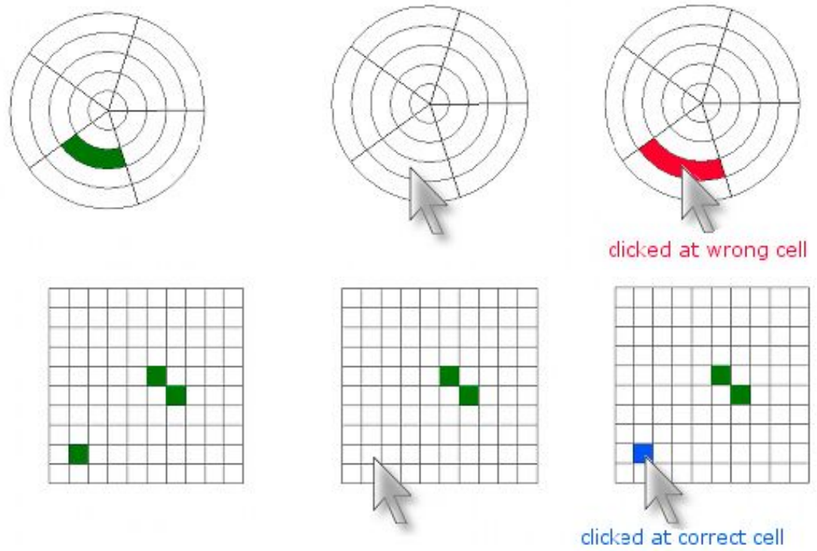
- rectilinear: scalability wrt #axes
  - 2 axes best, 3 problematic, 4+ impossible
- parallel: unfamiliarity, training time



Orientation limitations

- rectilinear: scalability wrt #axes
  - 2 axes best, 3 problematic, 4+ impossible
- parallel: unfamiliarity, training time
- radial: perceptual limits
  - polar coordinate asymmetry
    - angles lower precision than length
    - nonuniform sector width/size depending on radial distance
  - frequently problematic
    - but sometimes can be deliberately exploited!
      - for 2 attribs of very unequal importance



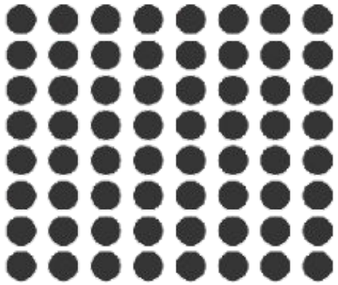


[Uncovering Strengths and Weaknesses of Radial Visualizations - an Empirical Approach. Diehl, Beck and Burch. IEEE TVCG (Proc. InfoVis) 16(6):935--942, 2010.]

Layout density

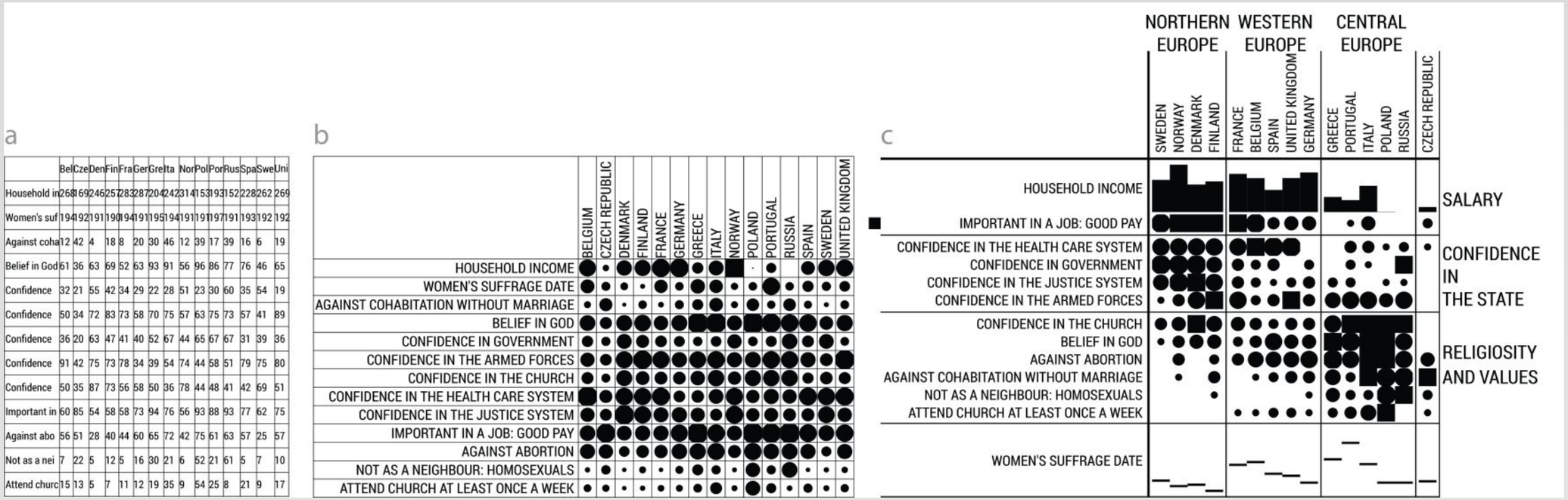
Layout Density

- Dense



- Space-Filling

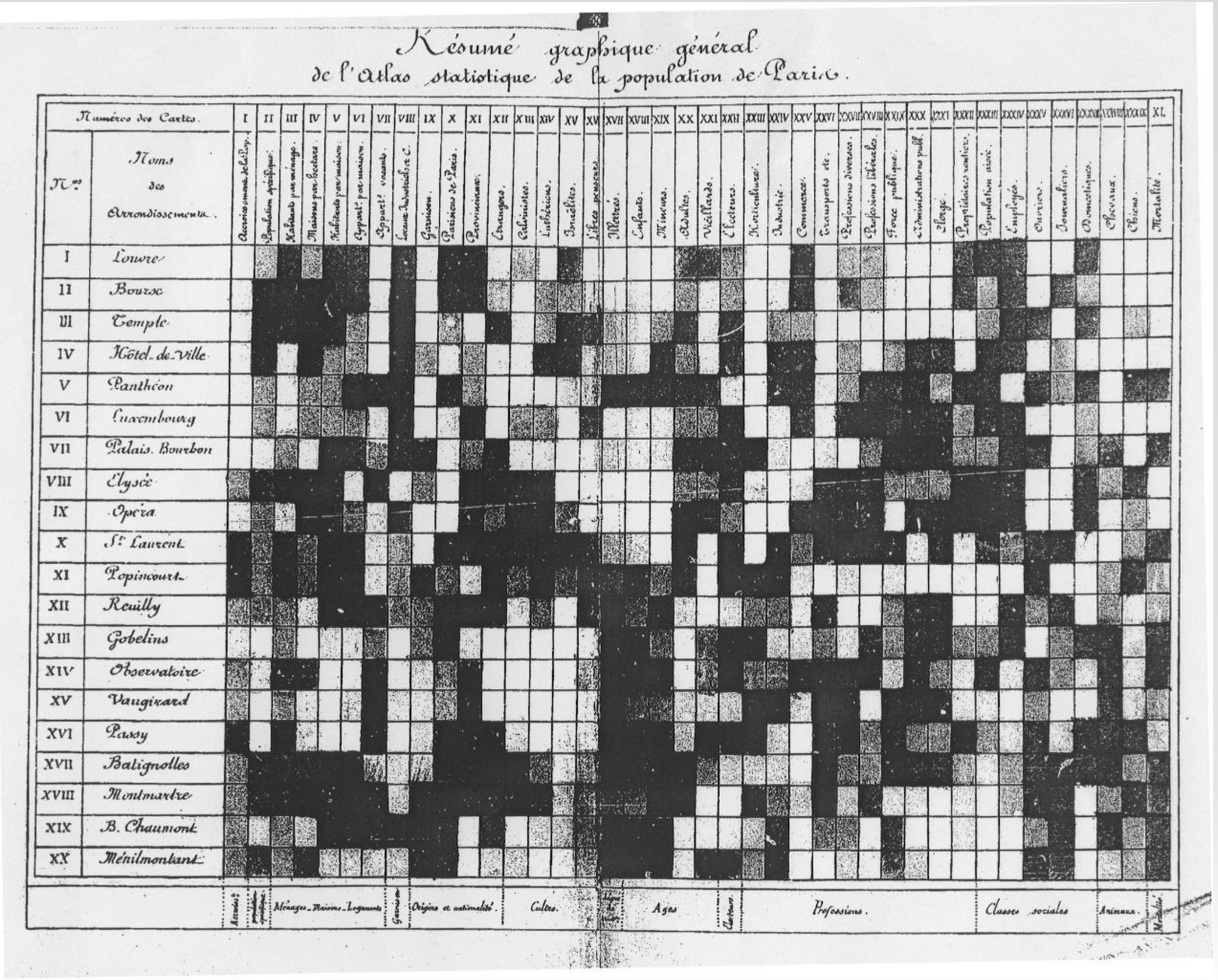
Revisiting Bertin Matrices: New Interactions for Crafting Tabular Visualizations



Charles Perin, Pierre Dragicevic, Jean-Daniel Fekete

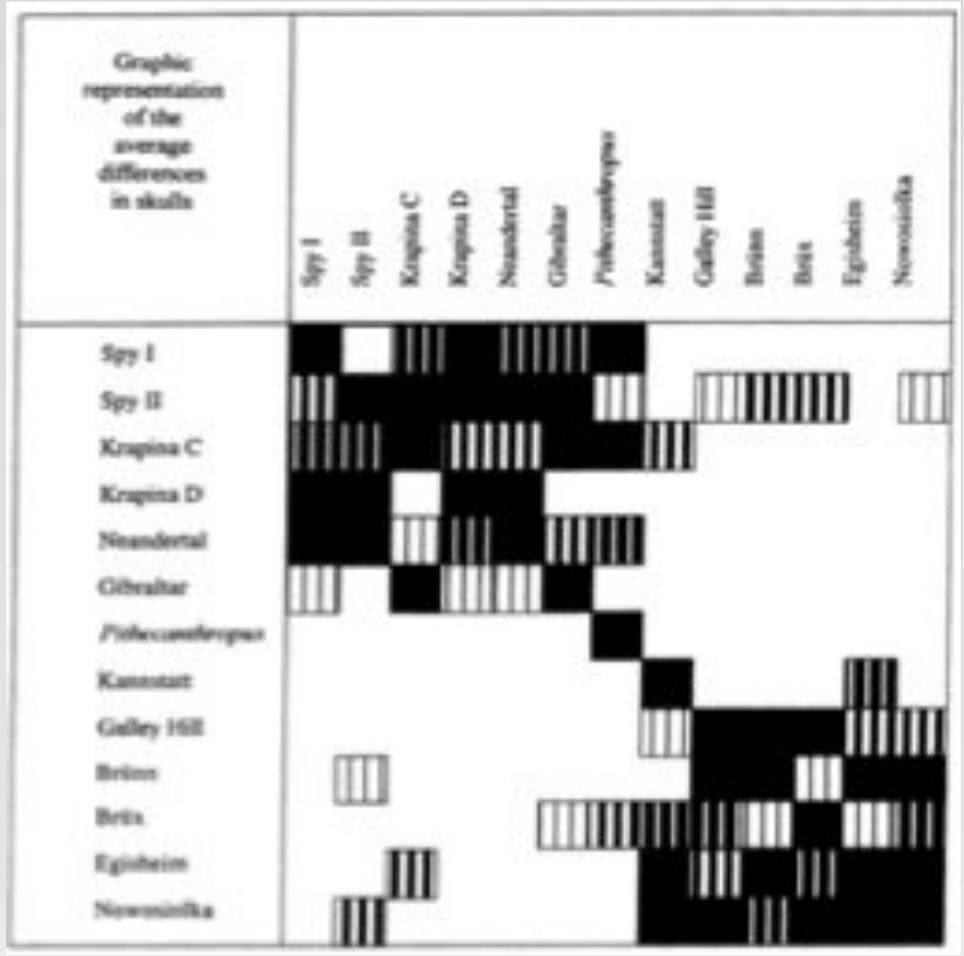
www.aviz.fr/bertifier

informatics mathematics
Aviz
Visual Analytics Project
UNIVERSITÉ PARIS SUD
CNRS LIMSI



More examples here:
www.aviz.fr/bertifier

J. Czekanowski 1909



5 b&w patterns

Manual reordering

J. Bertin 70s-80s

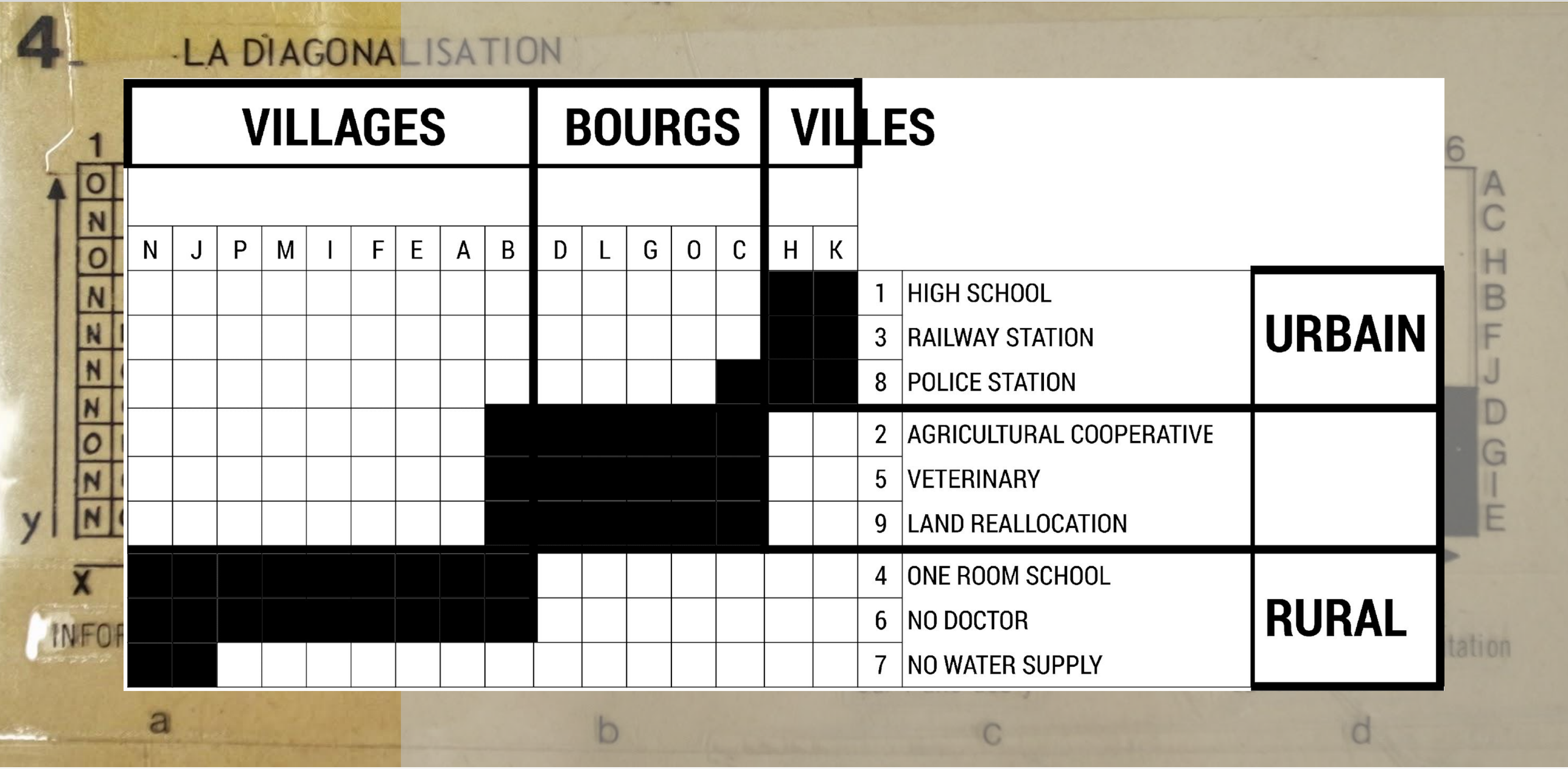
Manual reordering

Dedicated device

J. Bertin 1975

Empirical method

"[Manipulation] is fundamental. It is the internal mobility of the image that characterize the modern graphic"


4
LA DIAGONALISATION
VILLAGES
BOURGS
VILLES
N J P M I F E A B D L G O C H K
1 HIGH SCHOOL
3 RAILWAY STATION
8 POLICE STATION
URBAIN
2 AGRICULTURAL COOPERATIVE
5 VETERINARY
9 LAND REALLOCATION
4 ONE ROOM SCHOOL
6 NO DOCTOR
7 NO WATER SUPPLY
RURAL
a
b
c
d

For small n (< 10-50)

A few visualization techniques exist:

- Scatterplot matrix (SPLOM)
- Parallel coordinates
- Matrix/Heatmap

For more:

- projection methods reduce n to 2 (or 3), trying to maintain distances

MNIST Benchmark

database of handwritten digits for training ML systems

28x28 = 784 pixel of

255 greylevels

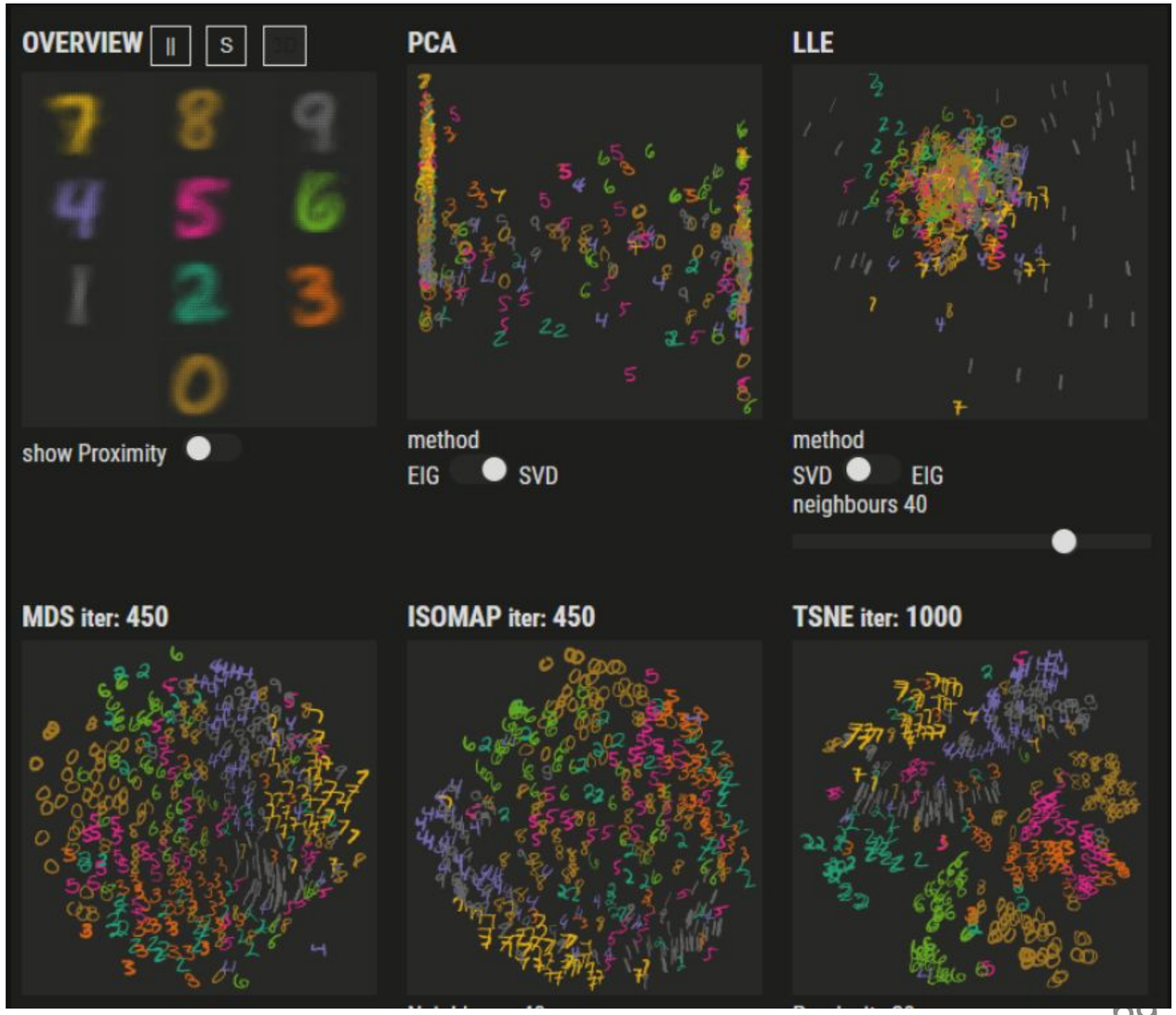
60,000 training

10,000 testing

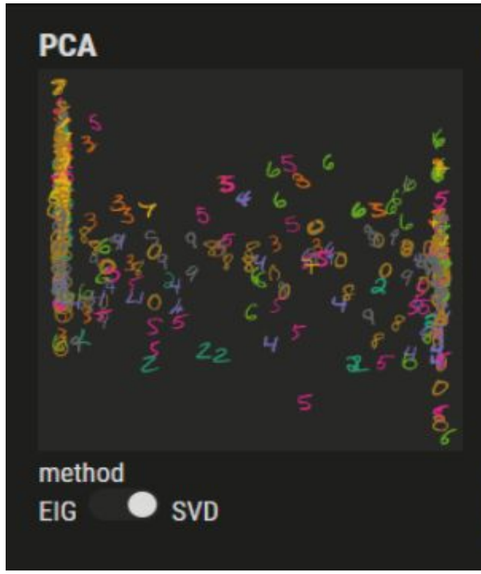
Projections

Transform nD data into 2D data

- PCA
- MDS
- t-SNE
- UMAP



PCA



Classical linear dimension reduction

- Very fast now, see e.g., scikit learn
  - stochastic algorithms
- out-of-core with IncrementalPCA
- Provide a stable global structure
  - not good at preserving local structure

Manifold Algorithms use KNN

Full distance matrix impossible for large HD data

- approximate k-nearest-neighbors
- M. Aumüller, E. Bernhardsson, A. Faithfull: ANN-Benchmarks: A Benchmarking Tool for Approximate Nearest Neighbor Algorithms. Information Systems 2019. DOI: 10.1016/j.is.2019.02.006
- Huge differences, improving every year
- Use PyNNDescent for now
  - https://pynndescent.readthedocs.io/en/latest/

MDS

Distances in HD = distances in 2D (metric)

Order in HD = order in 2D (non metric)

- Slow (not scalable)
- Limited for HD due to intrinsic limits
- Don't use it for HD data
- See http://colah.github.io/posts/2014-10-Visualizing-MNIST/

John A. Lee, Michel Verleysen, Shift-invariant similarities circumvent distance concentration in stochastic neighbor embedding and variants, Procedia Computer Science, Volume 4, 2011, Pages 538-547, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2011.04.056 .

t-SNE

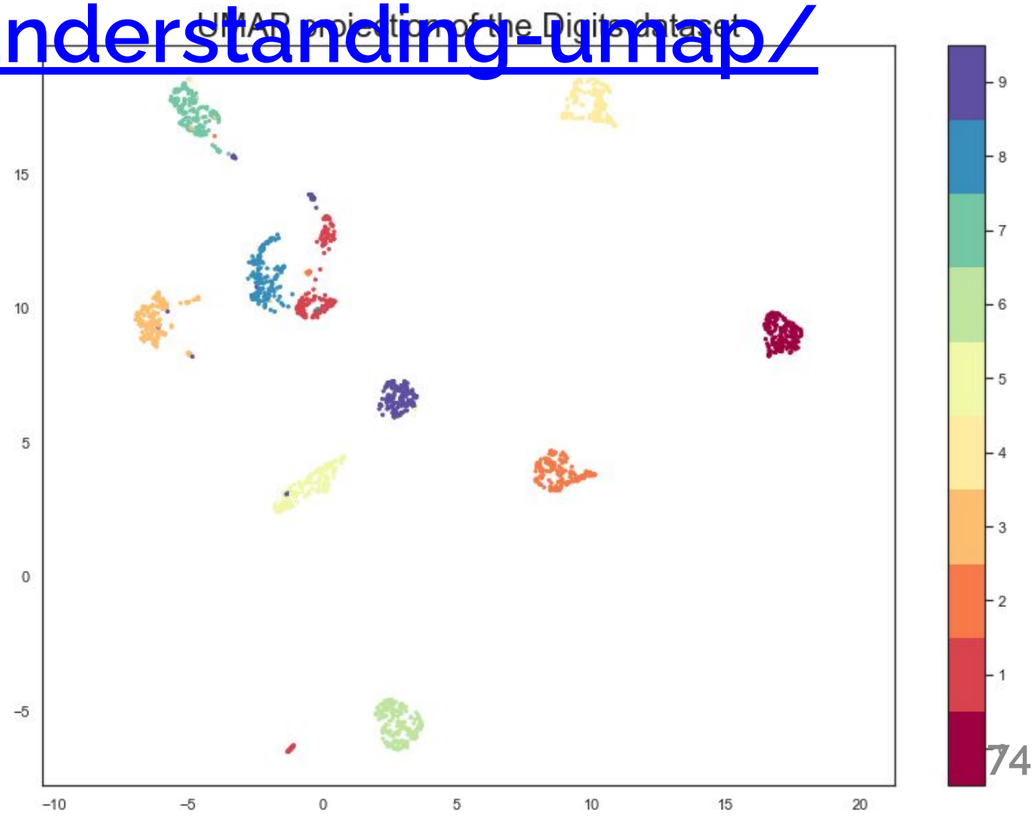
Manifold method using KNN

- Popular, many implementations
- Scale with GPU (1m points 1k dims)
  - https://github.com/CannyLab/tsne-cuda
- Hyperparameters are important
  - https://distill.pub/2016/misread-tsne/

UMAP

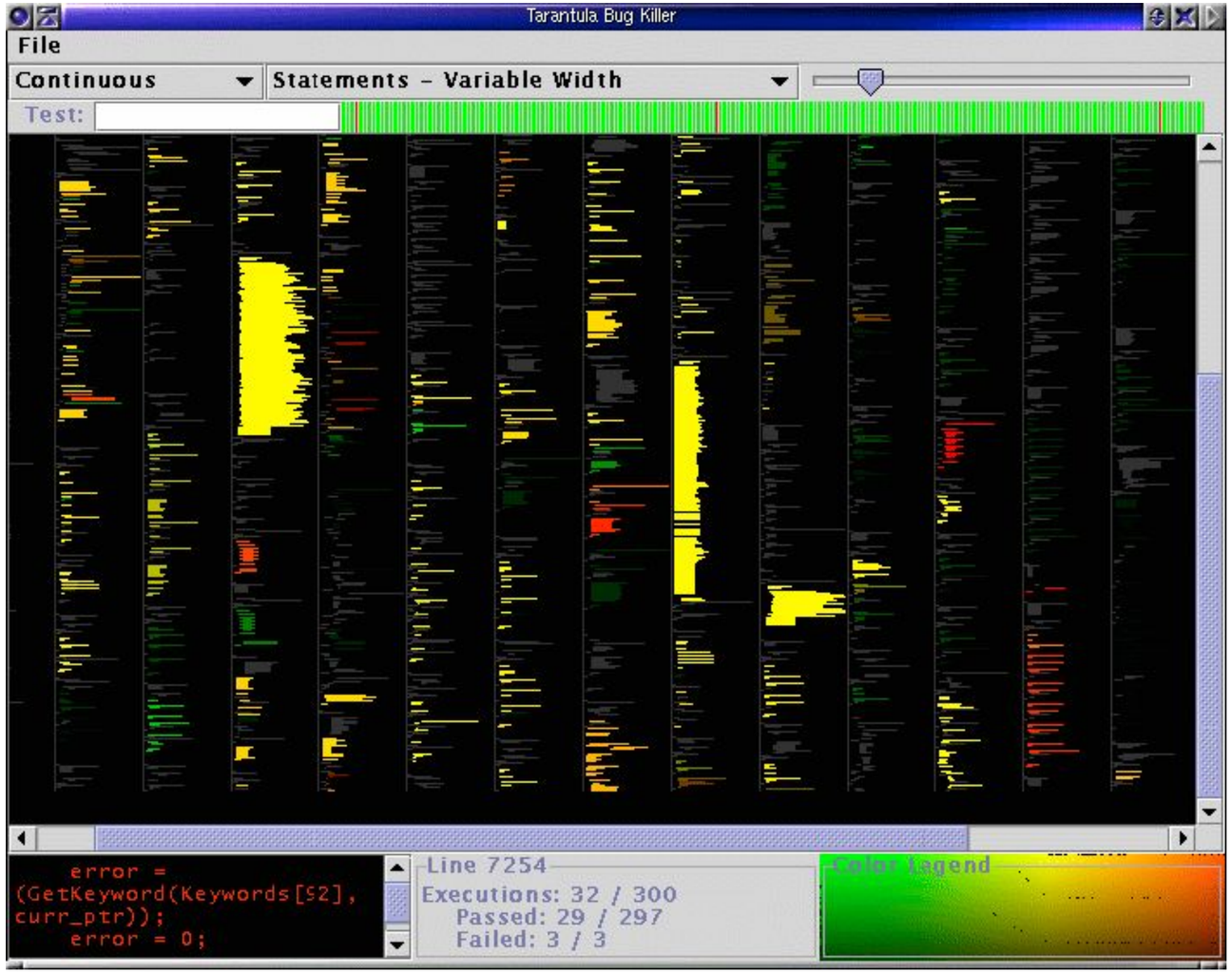
Like t-SNE but faster and more flexible

- https://umap-learn.readthedocs.io/en/latest/
- https://pair-code.github.io/understanding-umap/
- Can fit and transform
- Global + local = scale
  - learn of 20%
  - transform 80%



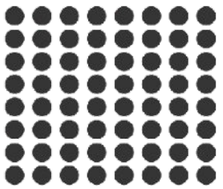
Idiom: Dense software overviews

- data: text
  - text + 1 quant attrib per line
- derived data:
  - one pixel high line
  - length according to original
- color line by attrib
- scalability
  - 10K+ lines



Layout Density

→ Dense



[Visualization of test information to assist fault localization. Jones, Harrold, Stasko. Proc. ICSE 2002, p 467-477.]

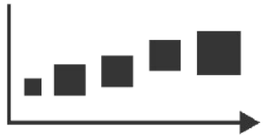
Arrange tables

Express Values

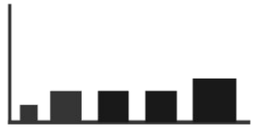
Separate, Order, Align Regions

Separate

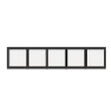
Order



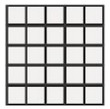
Align



1 Key List

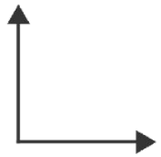


2 Keys Matrix



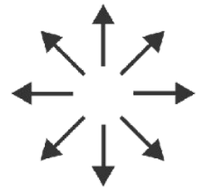
Axis Orientation

Rectilinear



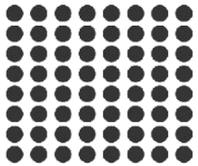
Parallel

Radial



Layout Density

Dense



How?

Encode

Arrange

- Express
- Separate
- Order
- Align
- Use

Map

from categorical and ordered attributes

- Color
  - Hue
  - Saturation
  - Luminance
- Size, Angle, Curvature, ...
- Shape
- Motion

Direction, Rate, Frequency, ...

Manipulate

Change

Select

Navigate

Facet

Juxtapose

Partition

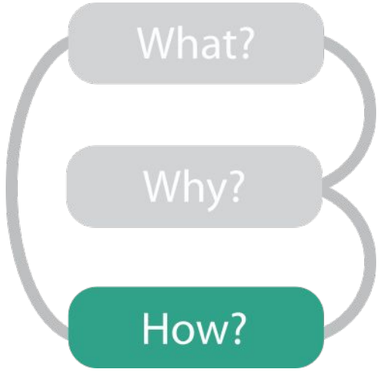
Superimpose

Reduce

Filter

Aggregate

Embed



How?

Encode

Arrange

- Express
- Separate
- Order
- Align
- Use

Map

from categorical and ordered attributes

- Color
  - *Hue*
  - *Saturation*
  - *Luminance*
- Size, Angle, Curvature, ...
- Shape
- Motion

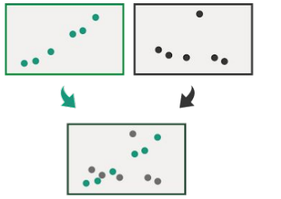
Direction, Rate, Frequency, ...

Manipulate

- Change
- Select
- Navigate

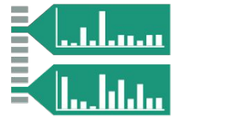
Facet

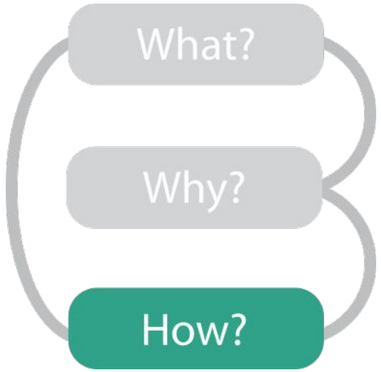
- Juxtapose
- Partition
- Superimpose



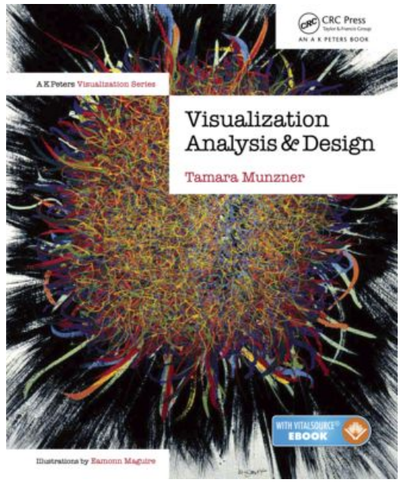
Reduce

- Filter
- Aggregate
- Embed





Visualization Analysis & Design

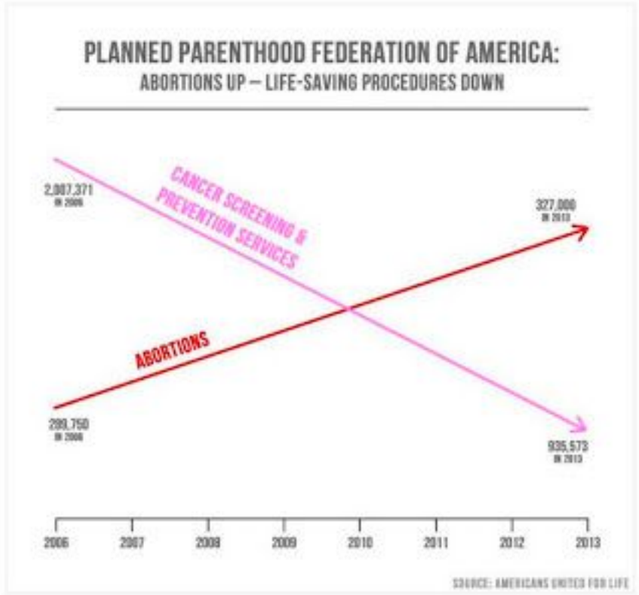


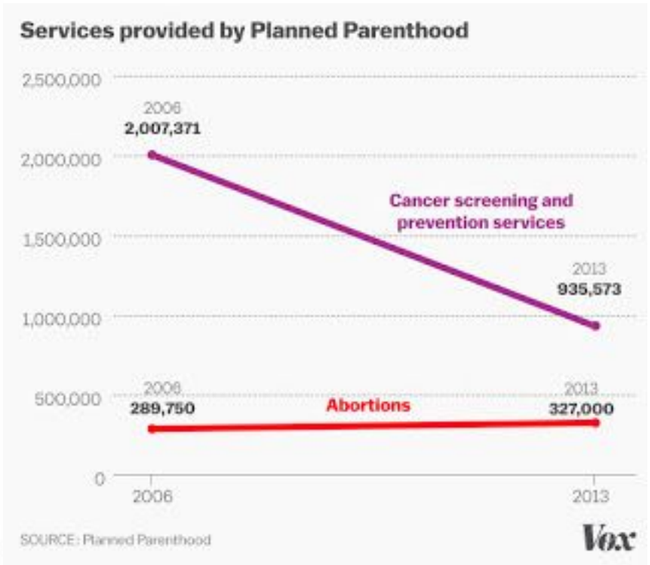
Tables (Ch 7) III

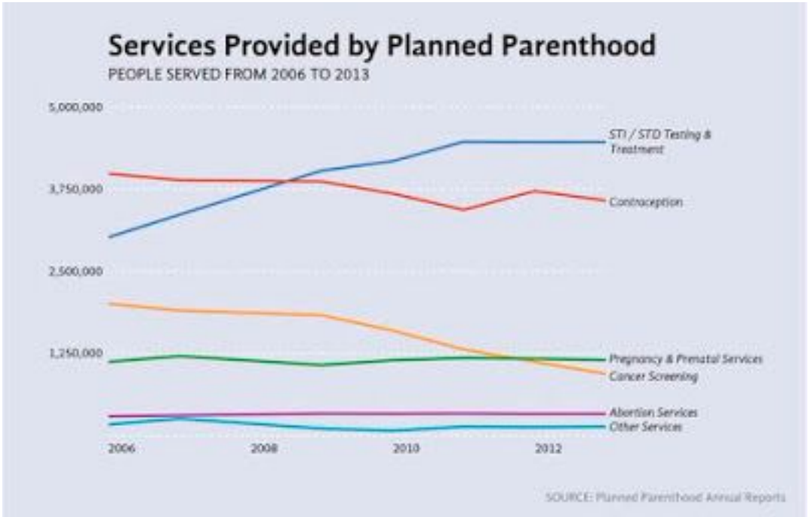
Tamara Munzner
Department of Computer Science
University of British Columbia
@tamaramunzner

Chart axes

- labelled axis is critical
- avoid cropping y-axis
  - include 0 at bottom left
  - or slope misleads



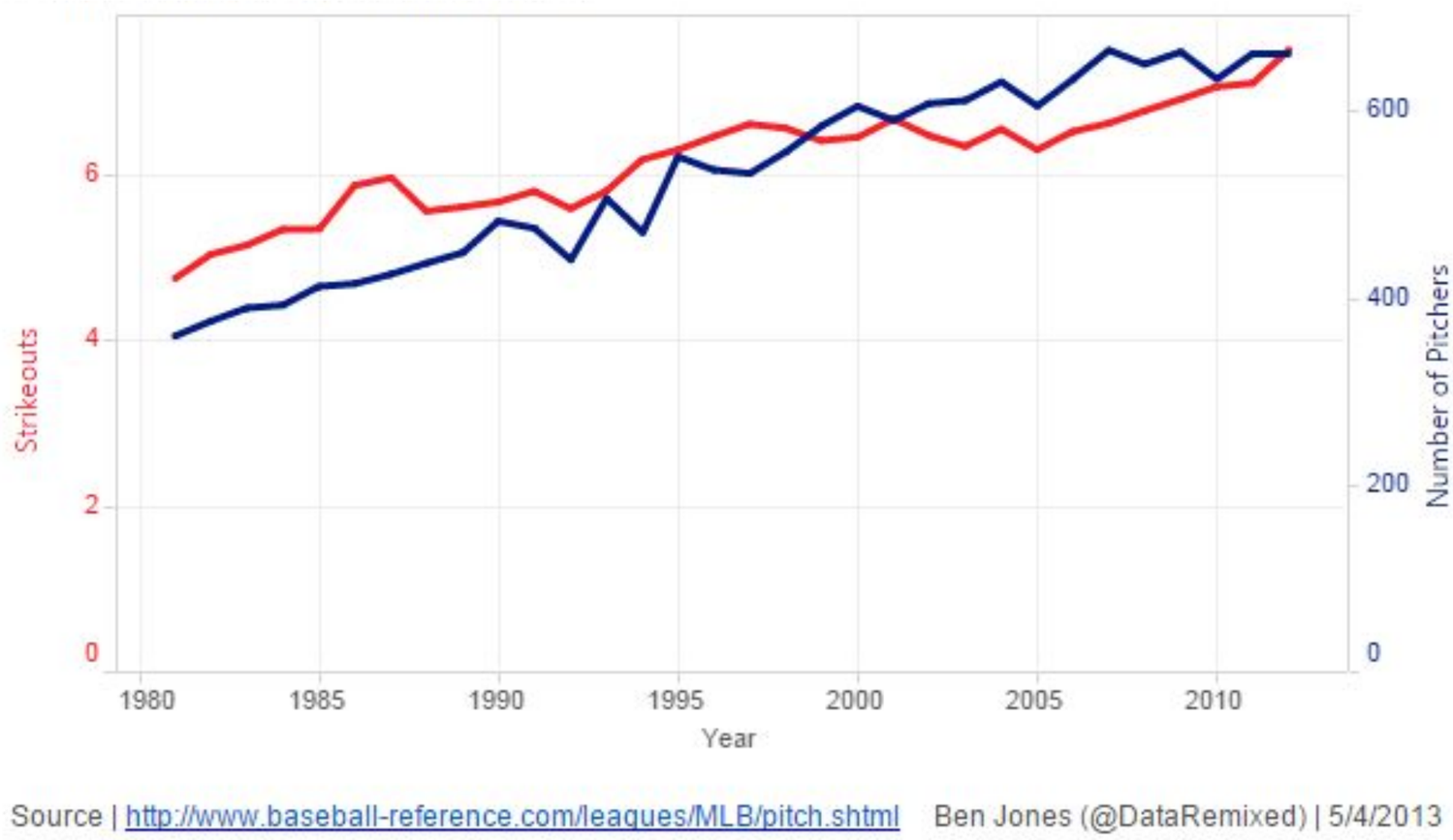




http://www.thefunctionalart.com/2015/10/if-you-see-bullshit-say-bullshit.html

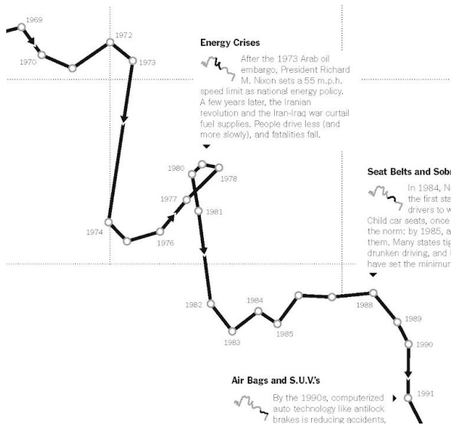
Idiom: **dual-axis line charts**

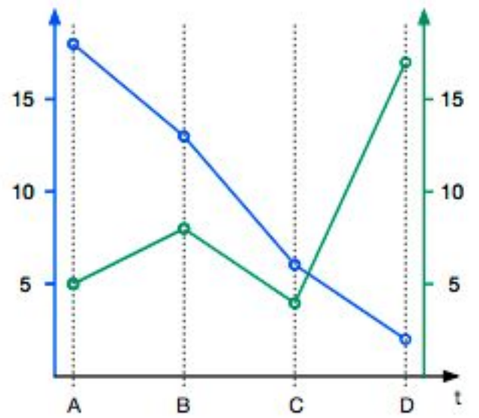
- controversial
  - acceptable if commensurate
  - beware, very easy to mislead!

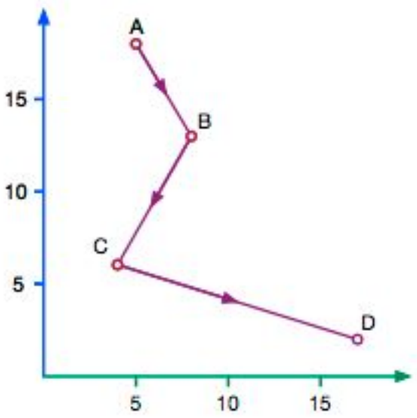


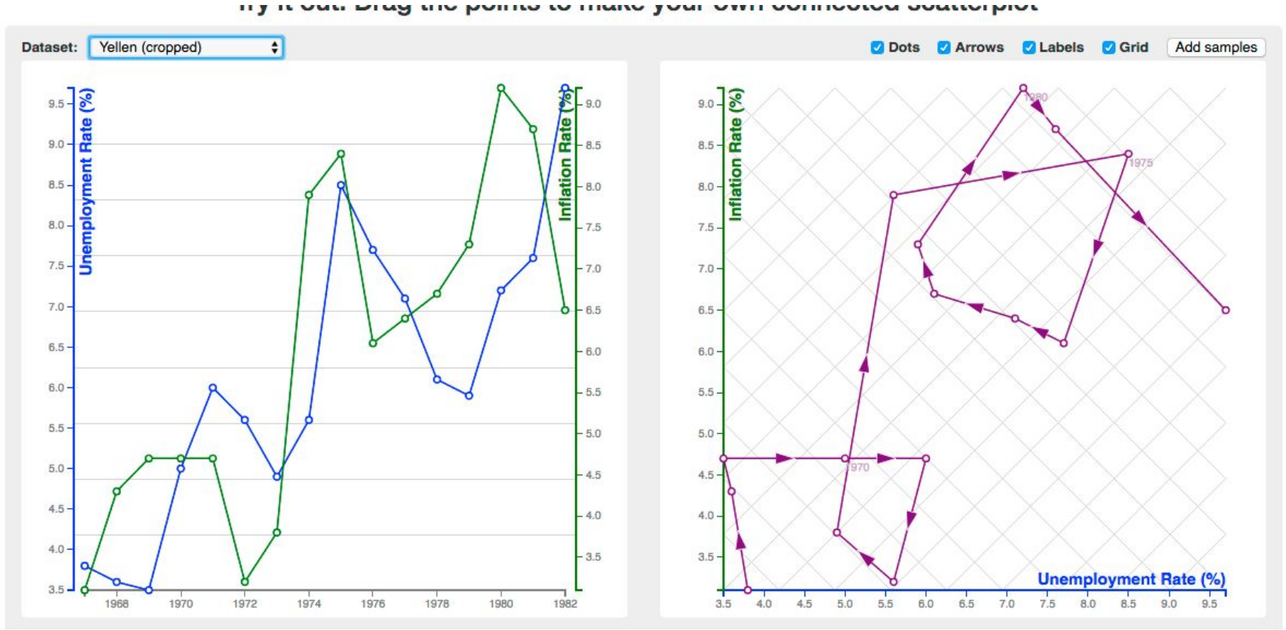
Idiom: **connected scatterplots**

- scatterplot with line connection marks
  - popular in journalism
  - horiz + vert axes: value attribs
  - line connection marks: temporal order
  - alternative to dual-axis charts
    - horiz: time
    - vert: two value attribs
- empirical study
  - engaging, but correlation unclear









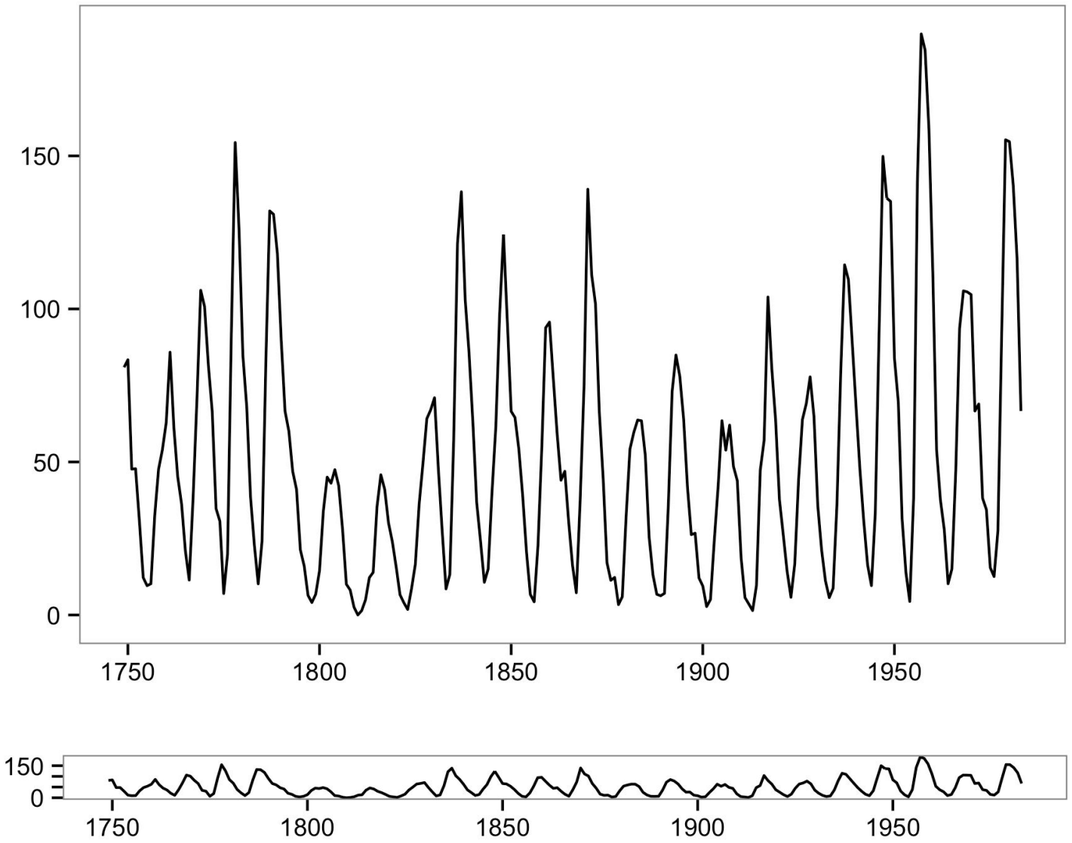
[The Connected Scatterplot for Presenting Paired Time Series. Haroz, Kosara and Franconeri. IEEE TVCG 22(9):2174-86, 2016.]

http://steveharoz.com/research/connected_scatterplot/

Choosing line chart aspect ratios

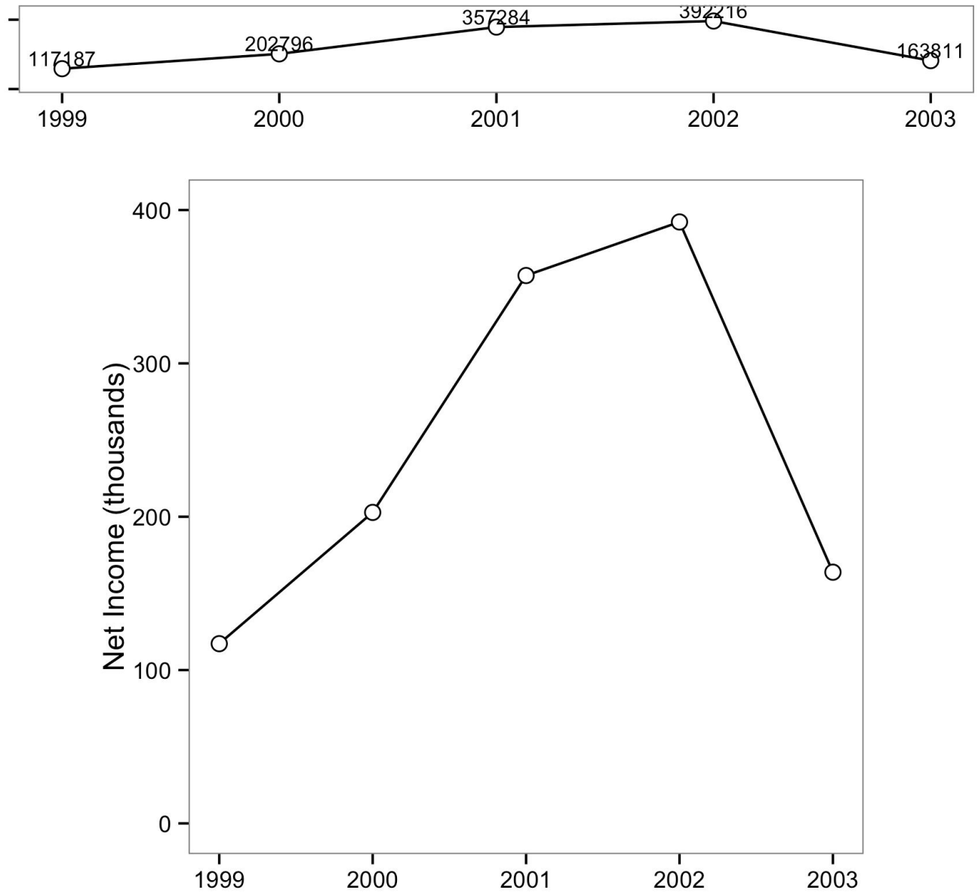
- 1: banking to 45 (1980s)
  - Cleveland perceptual argument: most accurate angle judgement at 45

Fig 7.1 Sunspot Data: Aspect Ratio 1



https://github.com/jennybc/r-graph-catalog/tree/master/figures/fig07-01_sunspot-data-aspect-ratio-1

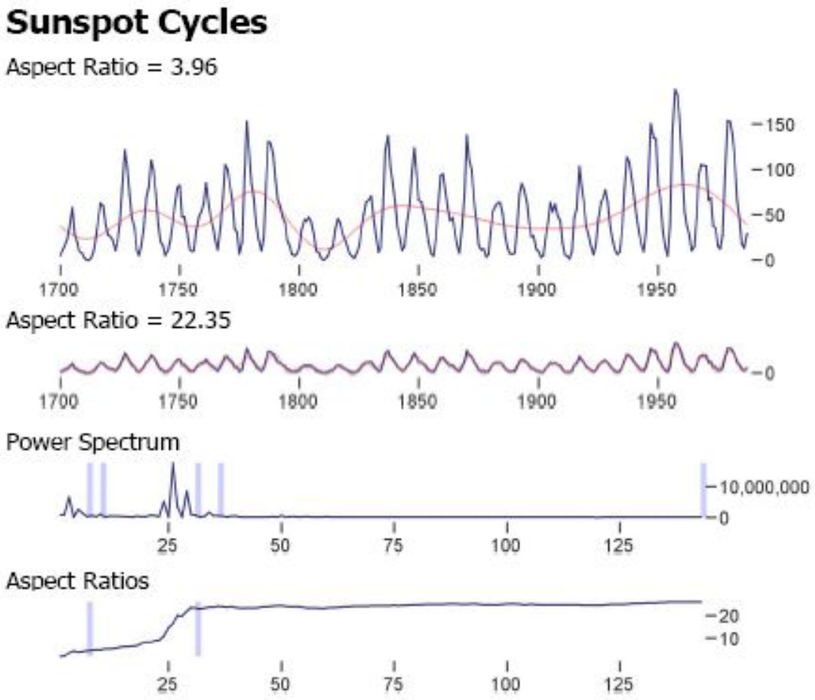
Fig 7.2 Annual Report: Aspect Ratio 2



https://github.com/jennybc/r-graph-catalog/tree/master/figures/fig07-02_annual-report-aspect-ratio-2

Choosing line chart aspect ratios

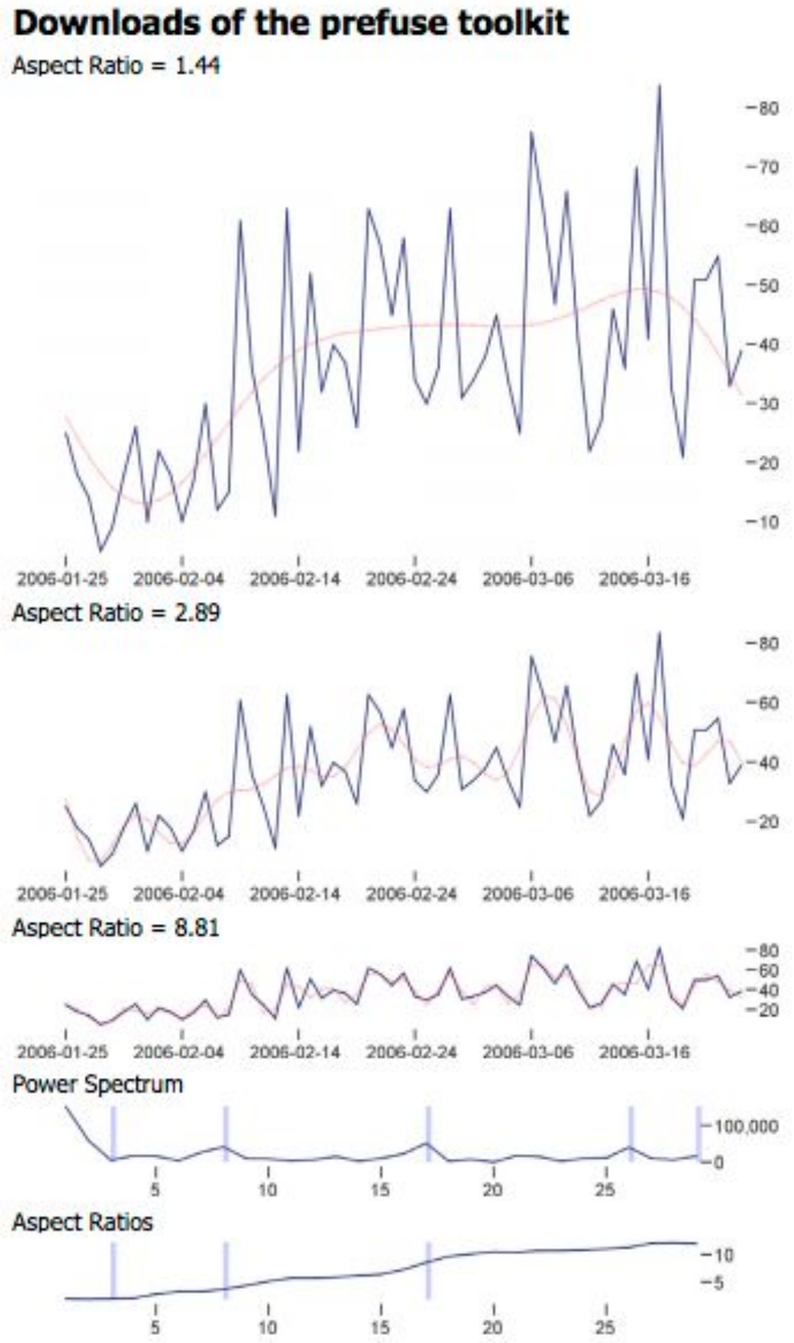
- 2: multi scale banking to 45 (2006)
  - frequency domain analysis to find ratios
    - FFT the data, convolve with Gaussian to smooth
  - find interesting spikes/ranges in power spectrum
    - cull nearby regions if similar, ensure overview
  - create trend curves (red) for each aspect ratio



overall

weekly

daily



[Multi-Scale Banking to 45 Degrees.

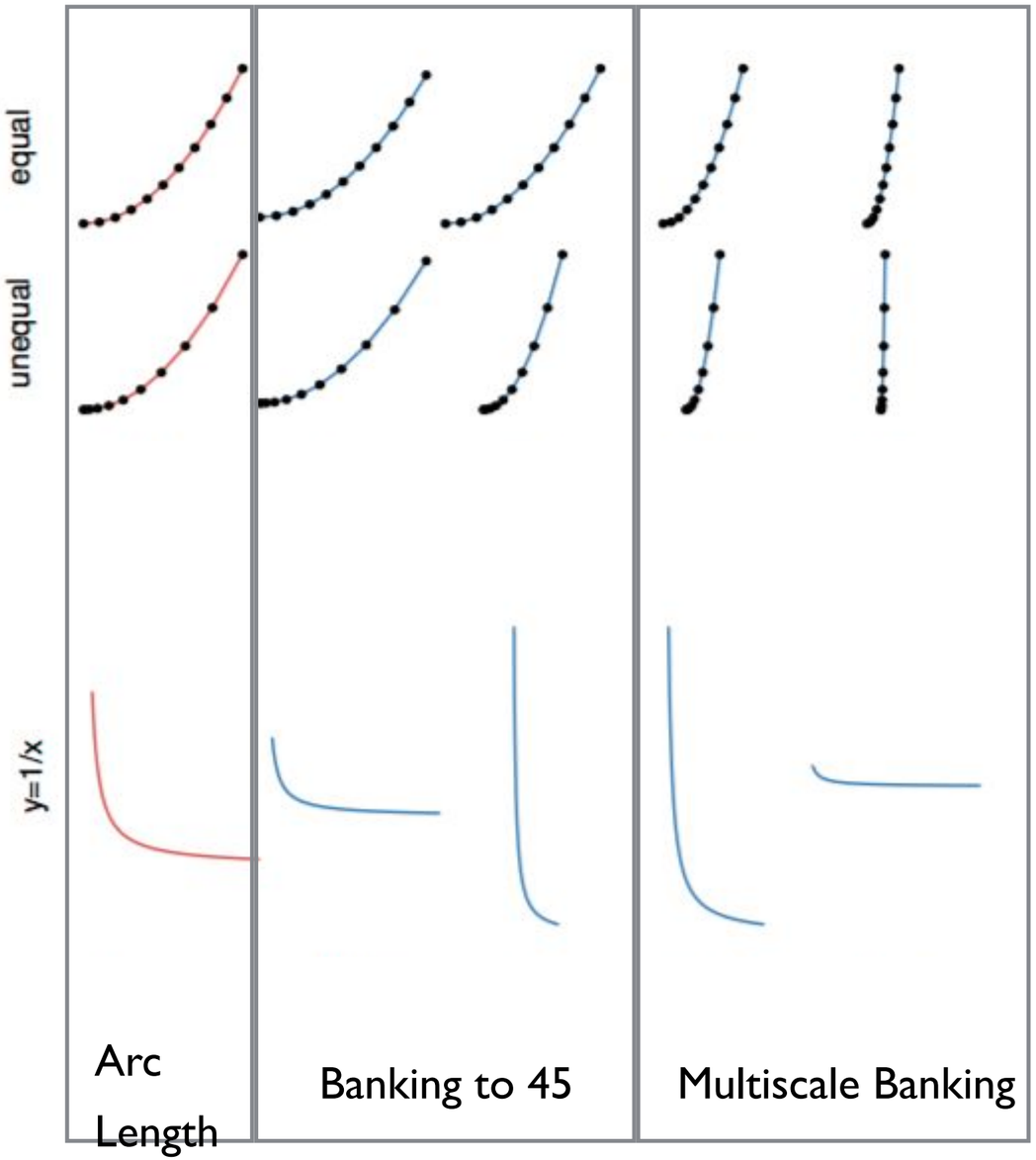
Heer and Agrawala, Proc InfoVis

2006]

Choosing line chart aspect ratios

- 3: arc length based aspect ratio (2011)
  - minimize the arc length of curve while keeping the area of the plot constant
  - parametrization and scale invariant
  - symmetry preserving
  - robust & fast to compute

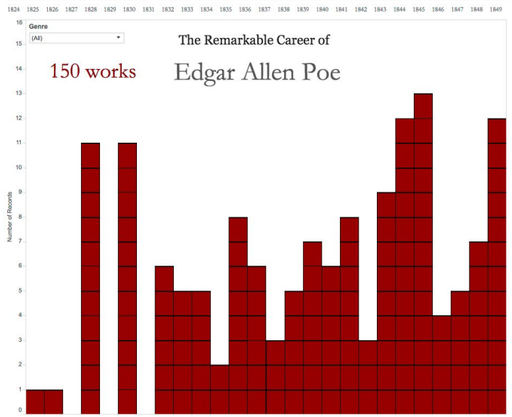
- meta-points from this progression
  - young field; prescriptive advice changes rapidly
  - reasonable defaults required deep dive into perception meets math

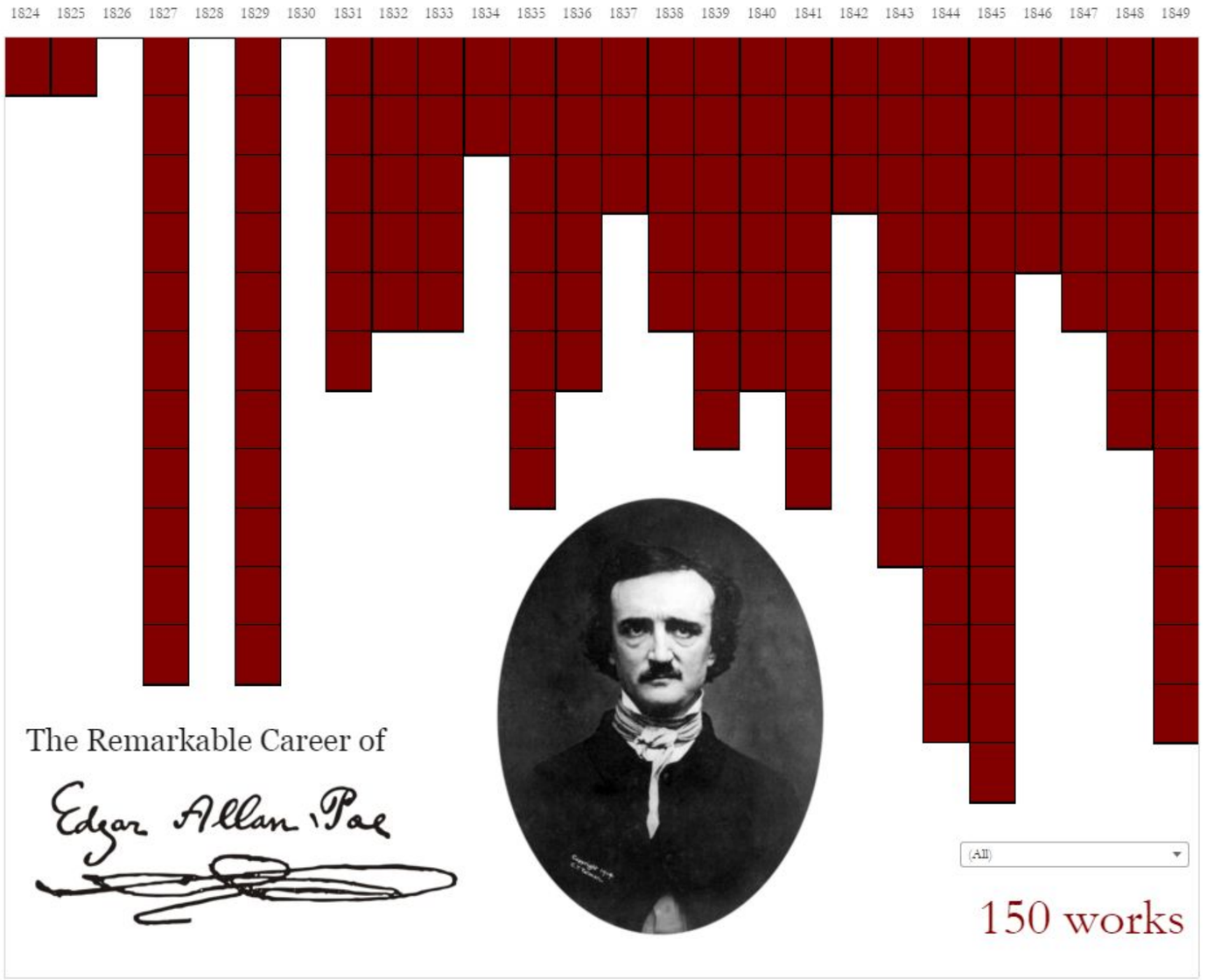


ngth-Based Aspect Ratio Selection. Talbot, Gerth, and Hanrahan. Proc InfoVis 2011]

Breaking conventions

- presentation vs exploration
  - engaging/evocative
  - inverted y axis
    - blood drips down on Poe





https://public.tableau.com/profile/ben.jones#!/vizhome/EdgarAllanPoeBoring/EdgarAllenPoeBoring

https://public.tableau.com/profile/ben.jones#!/vizhome/EdgarAllanPoeViz/EdgarAllanPoeViz

[Slide inspired by Ben Jones]