

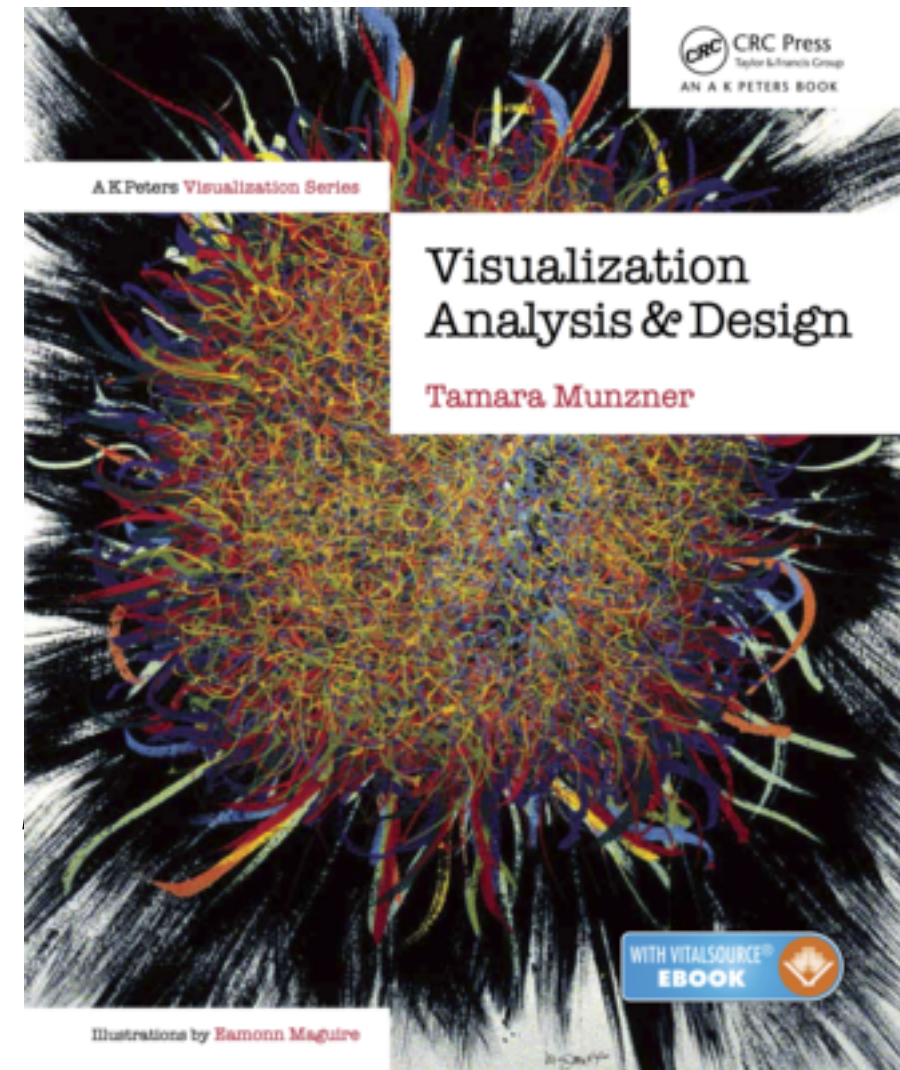
Visualization Analysis & Design, In More Depth

Tamara Munzner

Department of Computer Science

University of British Columbia

*Department of Industry, Innovation and Science, Economic and Analytical Services
June 23 2017, Canberra Australia*

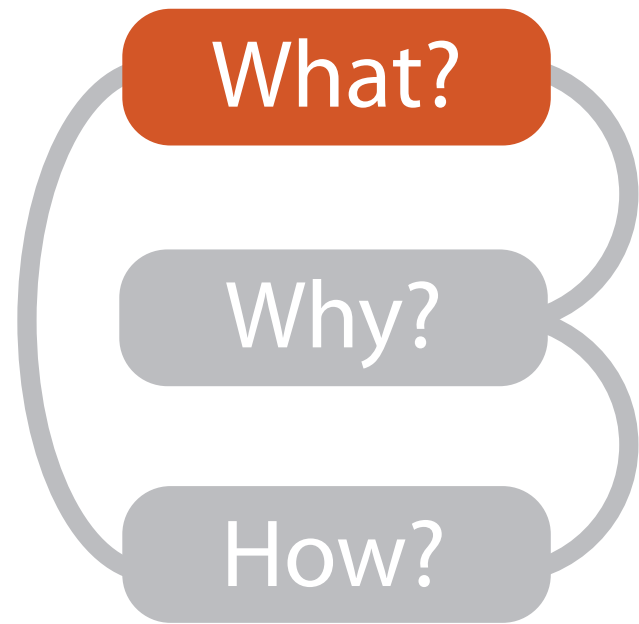


www.cs.ubc.ca/~tmm/talks.html#vad17can-aft

[@tamaramunzner](https://twitter.com/tamaramunzner)

Outline

- **Session 1** / 10-11:30am
Data Visualization Pitfalls to Avoid
 - Introduction
 - Color
 - Space: 2D vs 3D
- **Session 2** / 12:30-3pm
Visualization Analysis & Design, In More Depth
 - Marks and Channels, Perception
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate



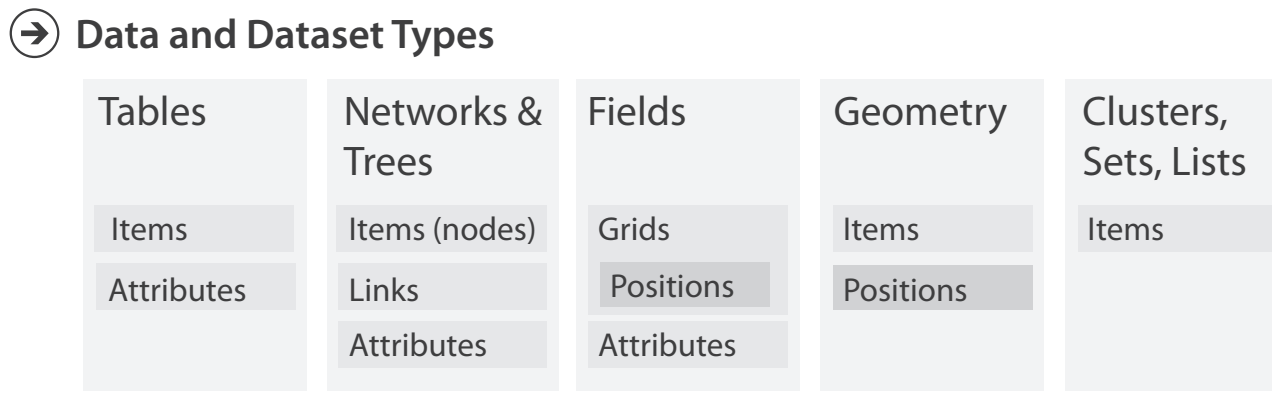
What?

Datasets

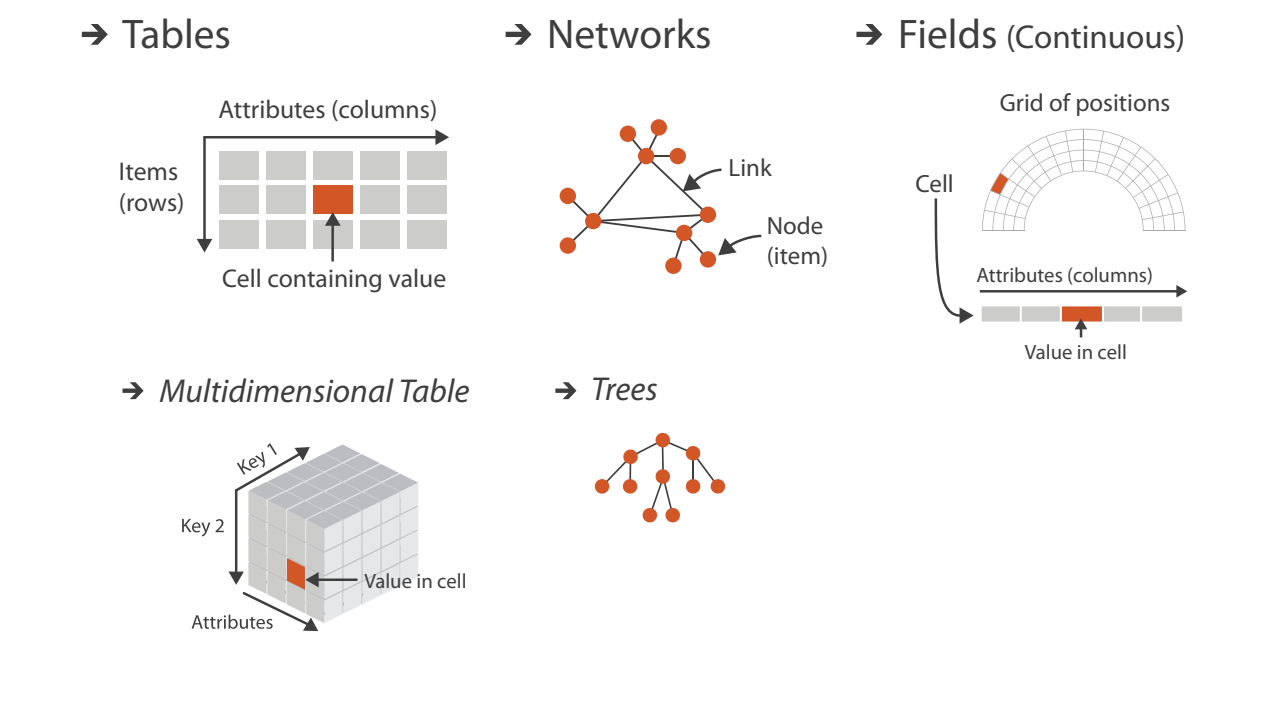
Attributes

- Data Types
 - Items
 - Attributes
 - Links
 - Positions
 - Grids

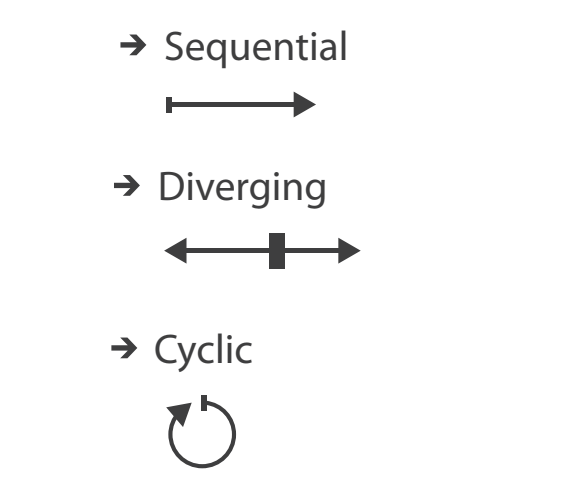
- Attribute Types
 - Categorical
 - + ● ■ ▲
 - Ordered
 - Ordinal
 - 👕 👕 👕
 - Quantitative
 - ┆ ┆ ┆



Dataset Types



Ordering Direction



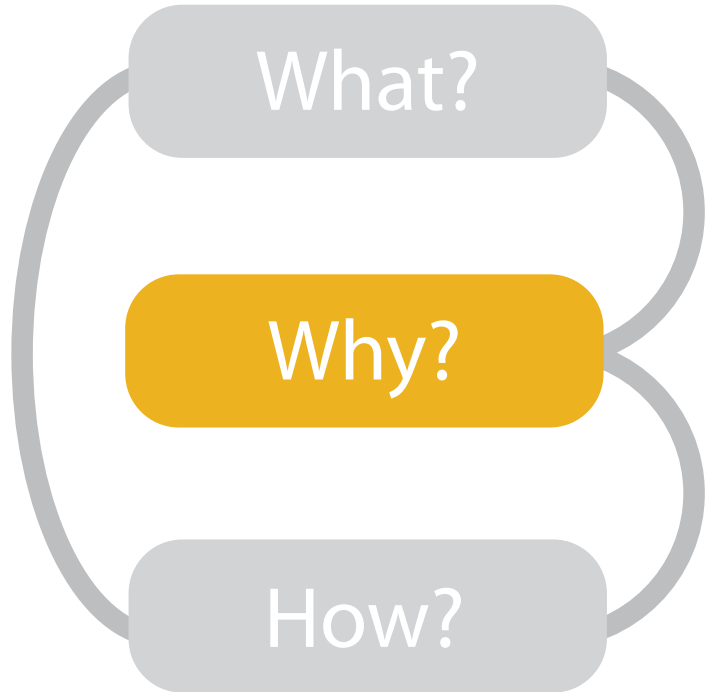
Geometry (Spatial)



Dataset Availability



Why?



👉 Actions

🎯 Targets

➔ **Analyze**

- ➔ Consume
 - ➔ Discover
 - ➔ Present
 - ➔ Enjoy
- ➔ Produce
 - ➔ Annotate
 - ➔ Record
 - ➔ Derive

➔ **Search**

	Target known	Target unknown
Location known	<i>Lookup</i>	<i>Browse</i>
Location unknown	<i>Locate</i>	<i>Explore</i>

➔ **Query**

- ➔ Identify
- ➔ Compare
- ➔ Summarize

➔ **All Data**

- ➔ Trends
- ➔ Outliers
- ➔ Features

➔ **Attributes**

- ➔ One
 - ➔ Distribution
 - ➔ Extremes
- ➔ Many
 - ➔ Dependency
 - ➔ Correlation
 - ➔ Similarity

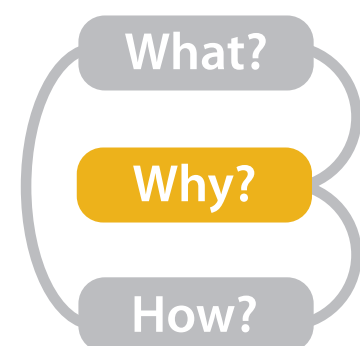
➔ **Network Data**

- ➔ Topology
- ➔ Paths

➔ **Spatial Data**

- ➔ Shape

- {action, target} pairs
 - discover distribution
 - compare trends
 - locate outliers
 - browse topology



How?

Encode

→ Arrange

→ Express



→ Separate



→ Order



→ Align



→ Use



→ Map

from **categorical** and **ordered** attributes

→ Color

→ Hue



→ Saturation



→ Luminance



→ Size, Angle, Curvature, ...



→ Shape



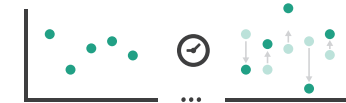
→ Motion

Direction, Rate, Frequency, ...

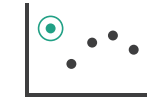


Manipulate

→ Change



→ Select



→ Navigate

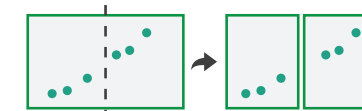


Facet

→ Juxtapose



→ Partition



→ Superimpose



Reduce

→ Filter



→ Aggregate



→ Embed



What?

Why?

How?

Channels: Rankings

➔ Magnitude Channels: Ordered Attributes



➔ Identity Channels: Categorical Attributes



Best
Effectiveness
Least

Same
Same

- **expressiveness principle**
 - match channel and data characteristics
- **effectiveness principle**
 - encode most important attributes with highest ranked channels

Channels: Expressiveness types and effectiveness rankings

➔ Magnitude Channels: Ordered Attributes

Position on common scale



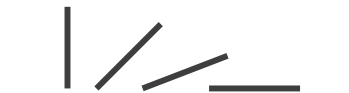
Position on unaligned scale



Length (1D size)



Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



Same

➔ Identity Channels: Categorical Attributes

Spatial region



Color hue



Motion



Shape



Best

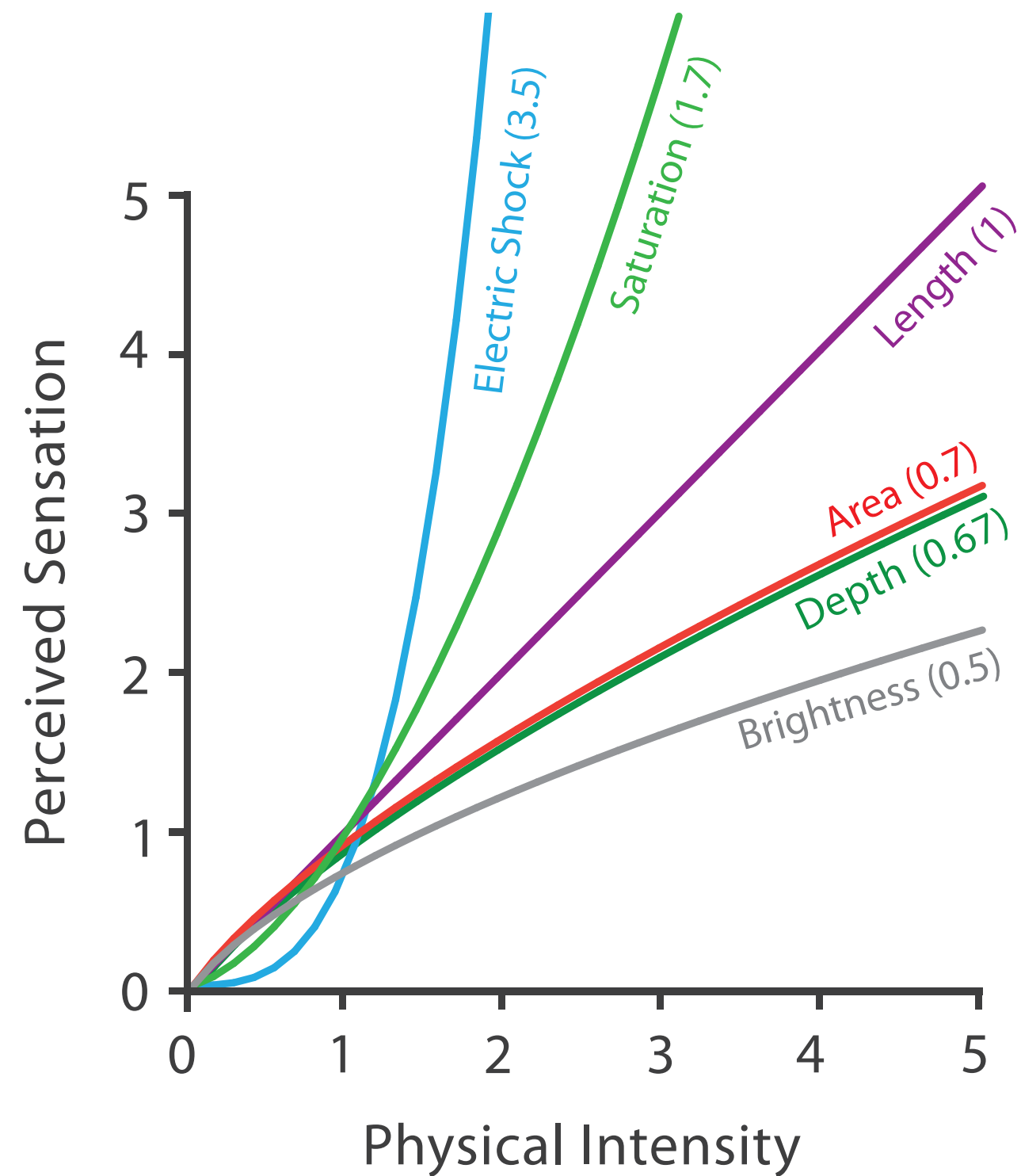
Effectiveness

Least

- **expressiveness principle**
 - match channel and data characteristics
- **effectiveness principle**
 - encode most important attributes with highest ranked channels
 - spatial position ranks high for both

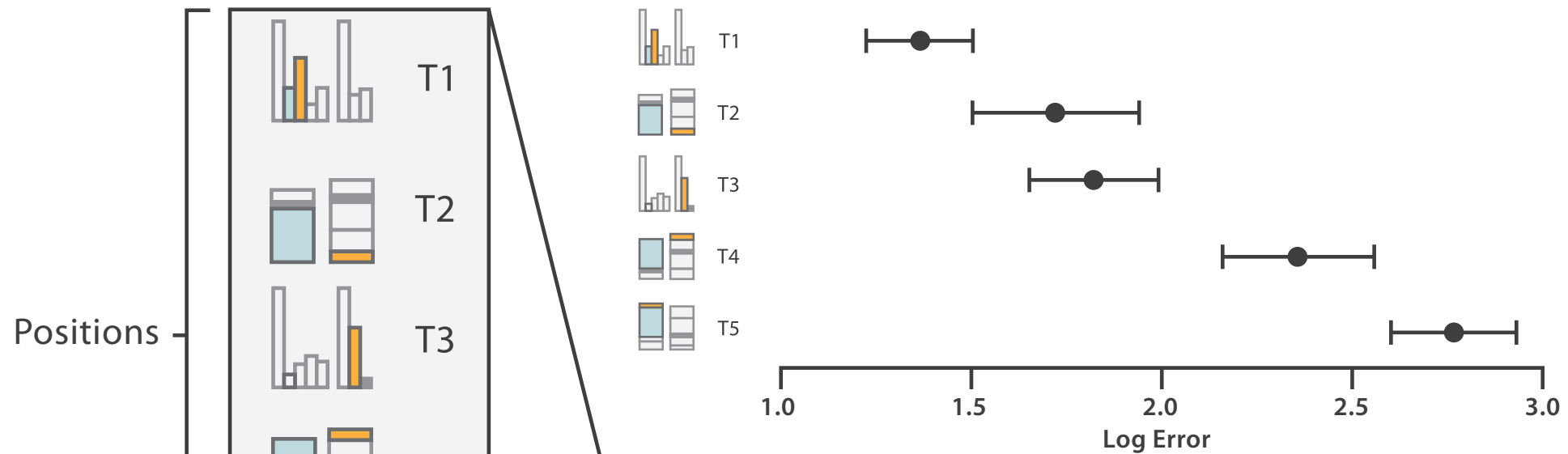
Accuracy: Fundamental Theory

Steven's Psychophysical Power Law: $S = I^N$

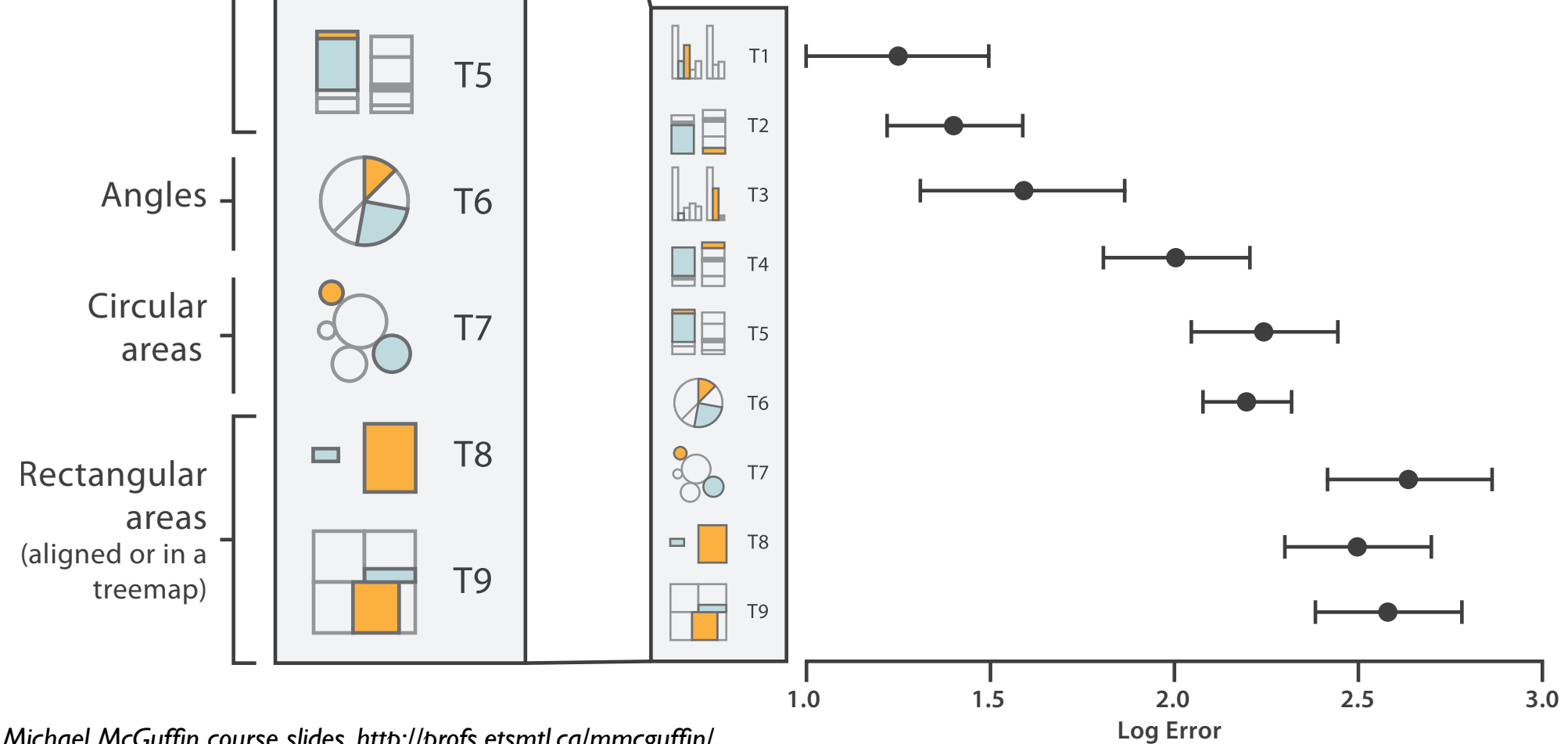


Accuracy: Vis experiments

Cleveland & McGill's Results



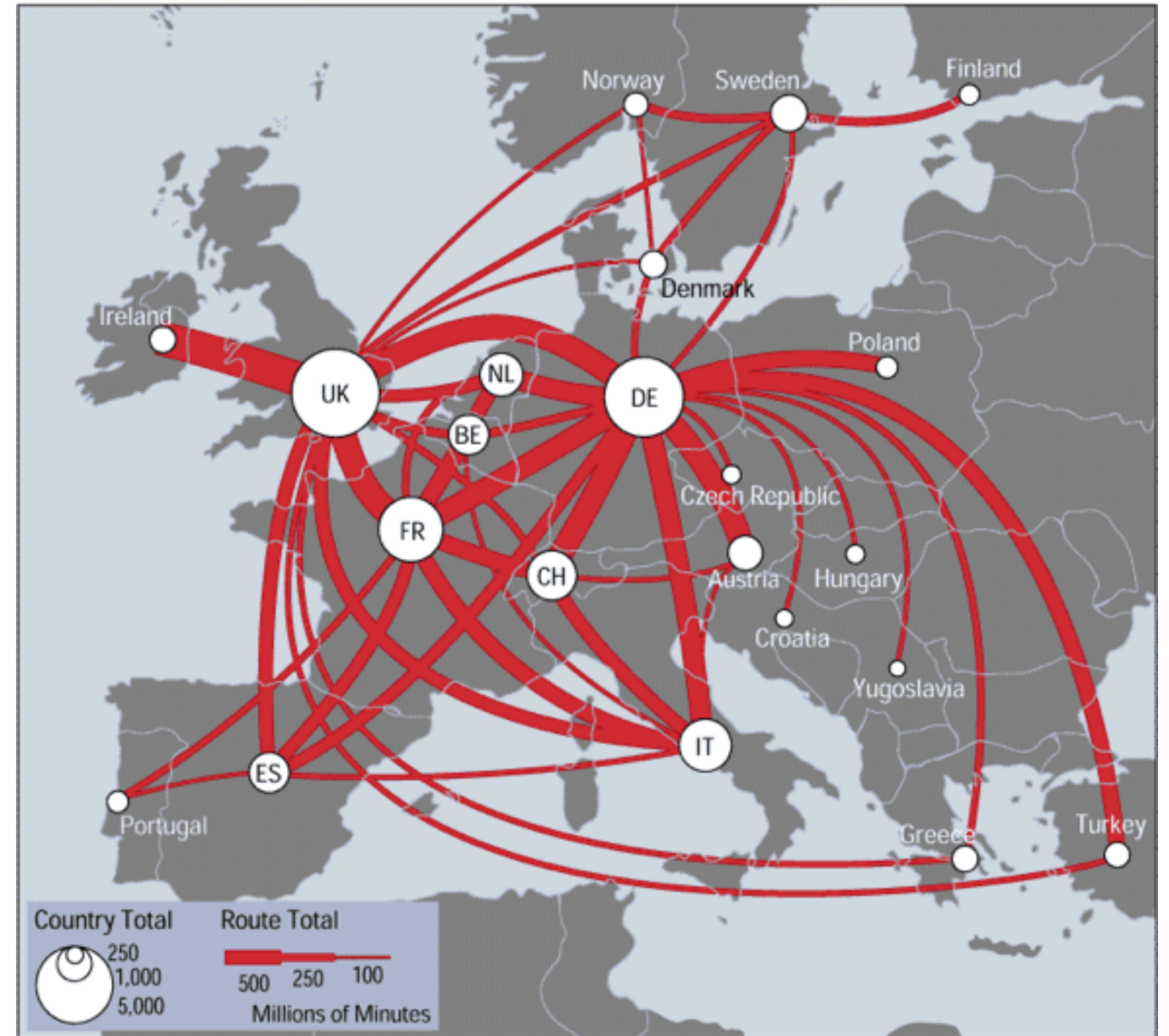
Crowdsourced Results



[Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. Heer and Bostock. Proc ACM Conf. Human Factors in Computing Systems (CHI) 2010, p. 203–212.]

Discriminability: How many usable steps?

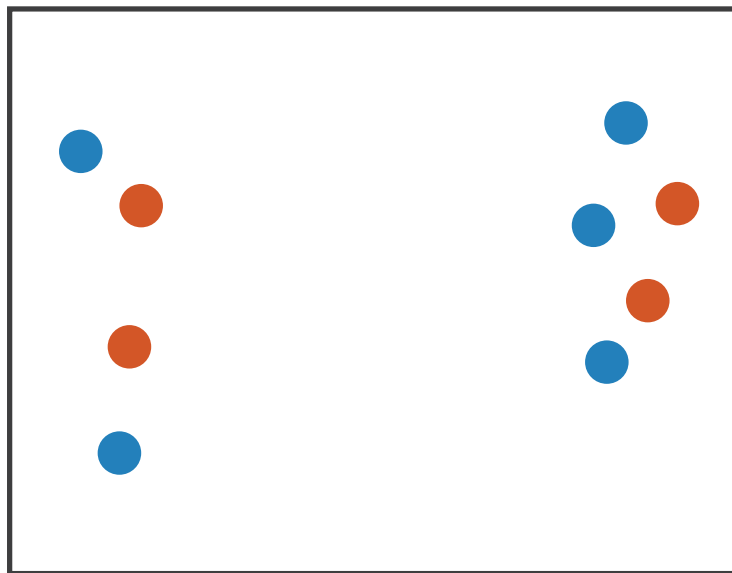
- must be sufficient for number of attribute levels to show
 - linewidth: few bins



[mappa.mundi.net/maps/maps_014/telegeography.html]

Separability vs. Integrality

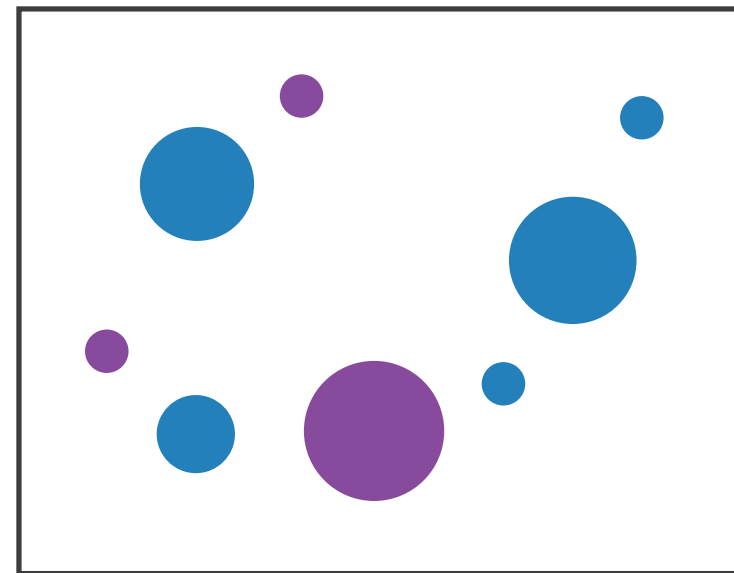
Position
+ Hue (Color)



Fully separable

2 groups each

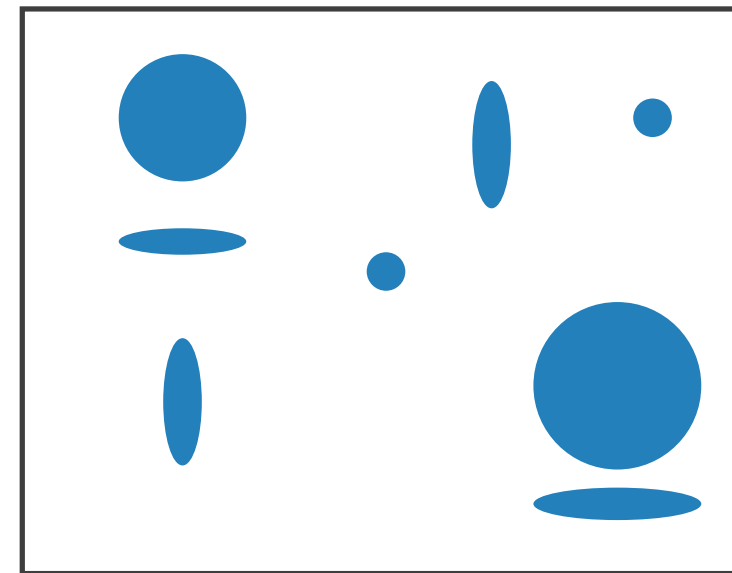
Size
+ Hue (Color)



Some interference

2 groups each

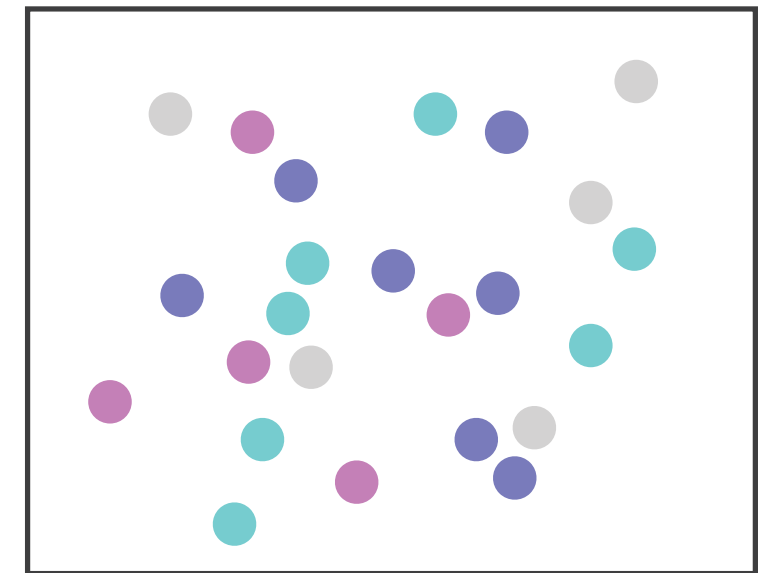
Width
+ Height



Some/significant
interference

3 groups total:
integral area

Red
+ Green

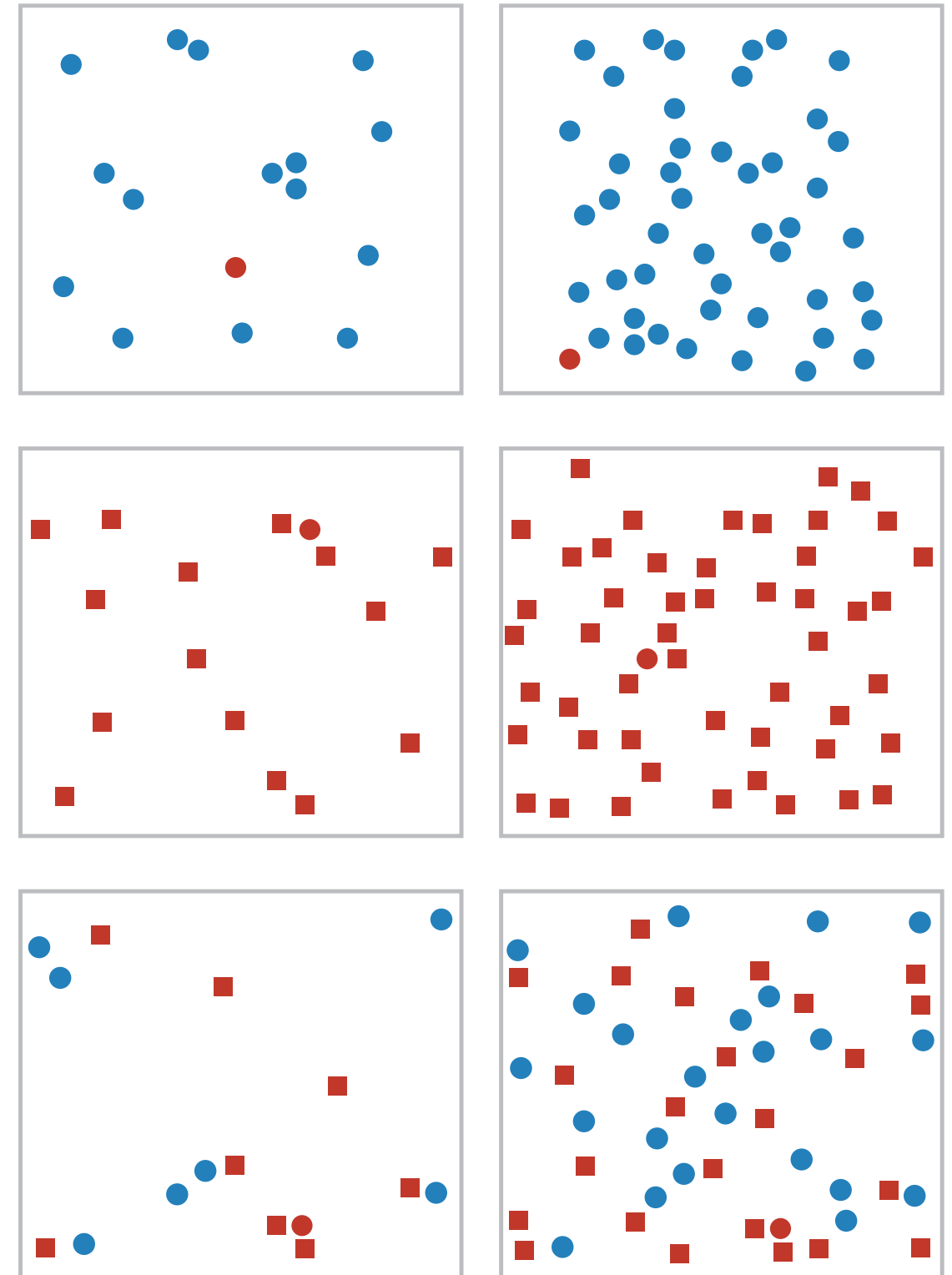


Major interference

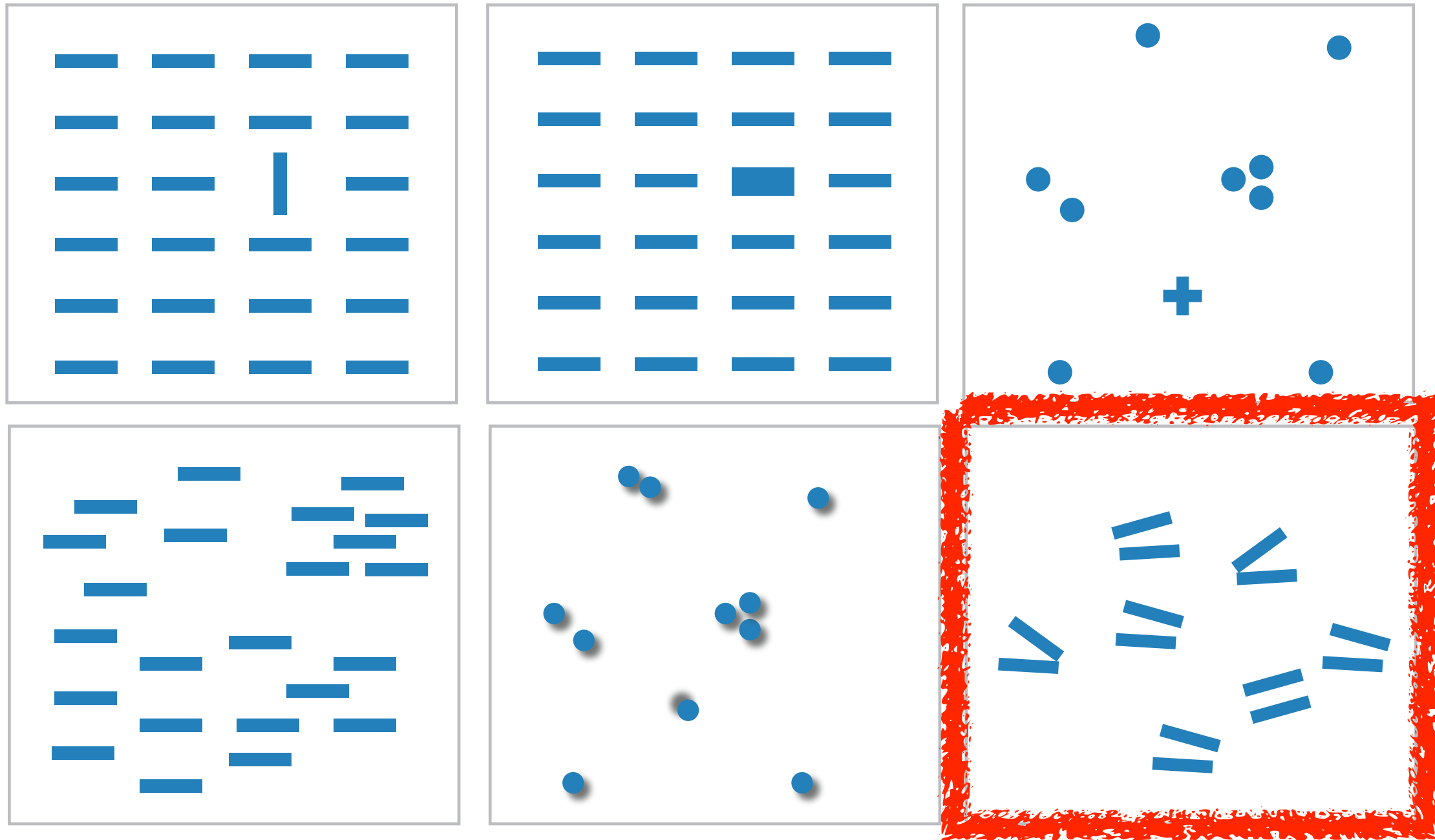
4 groups total:
integral hue

Popout

- find the red dot
 - how long does it take?
- parallel processing on many individual channels
 - speed independent of distractor count
 - speed depends on channel and amount of difference from distractors
- serial search for (almost all) combinations
 - speed depends on number of distractors



Popout



- many channels: tilt, size, shape, proximity, shadow direction, ...
- but not all! parallel line pairs do not pop out from tilted pairs

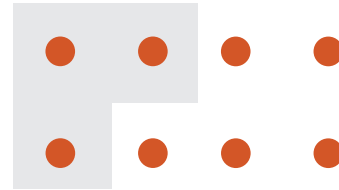
Grouping

- containment
- connection

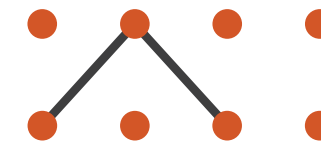
- proximity
 - same spatial region
- similarity
 - same values as other categorical channels

Marks as Links

➔ Containment



➔ Connection



➔ Identity Channels: Categorical Attributes

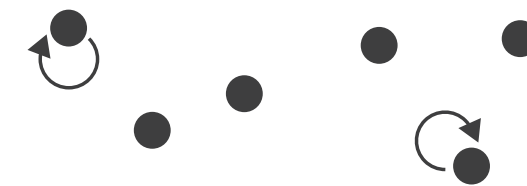
Spatial region



Color hue



Motion



Shape

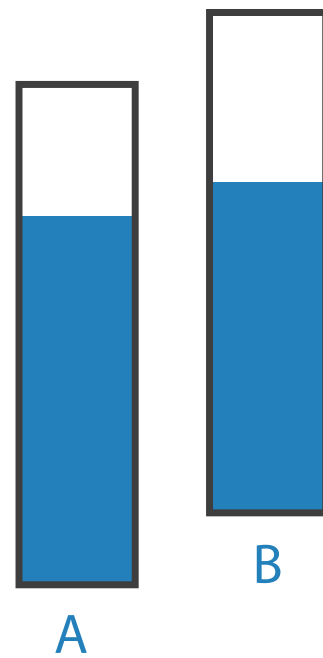


Relative vs. absolute judgements

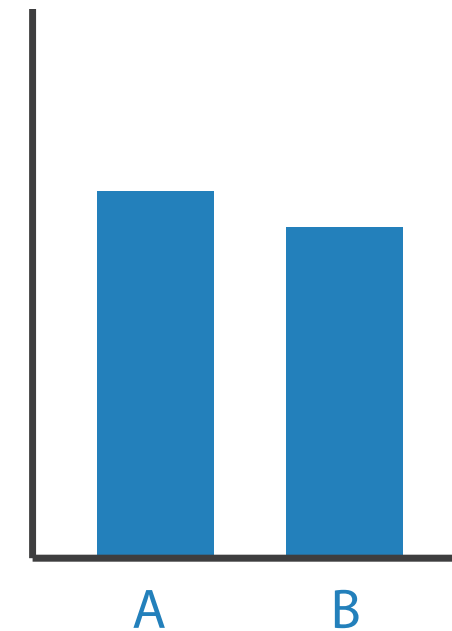
- perceptual system mostly operates with relative judgements, not absolute
 - that's why accuracy increases with common frame/scale and alignment
 - Weber's Law: ratio of increment to background is constant
 - filled rectangles differ in length by 1:9, difficult judgement
 - white rectangles differ in length by 1:2, easy judgement



length



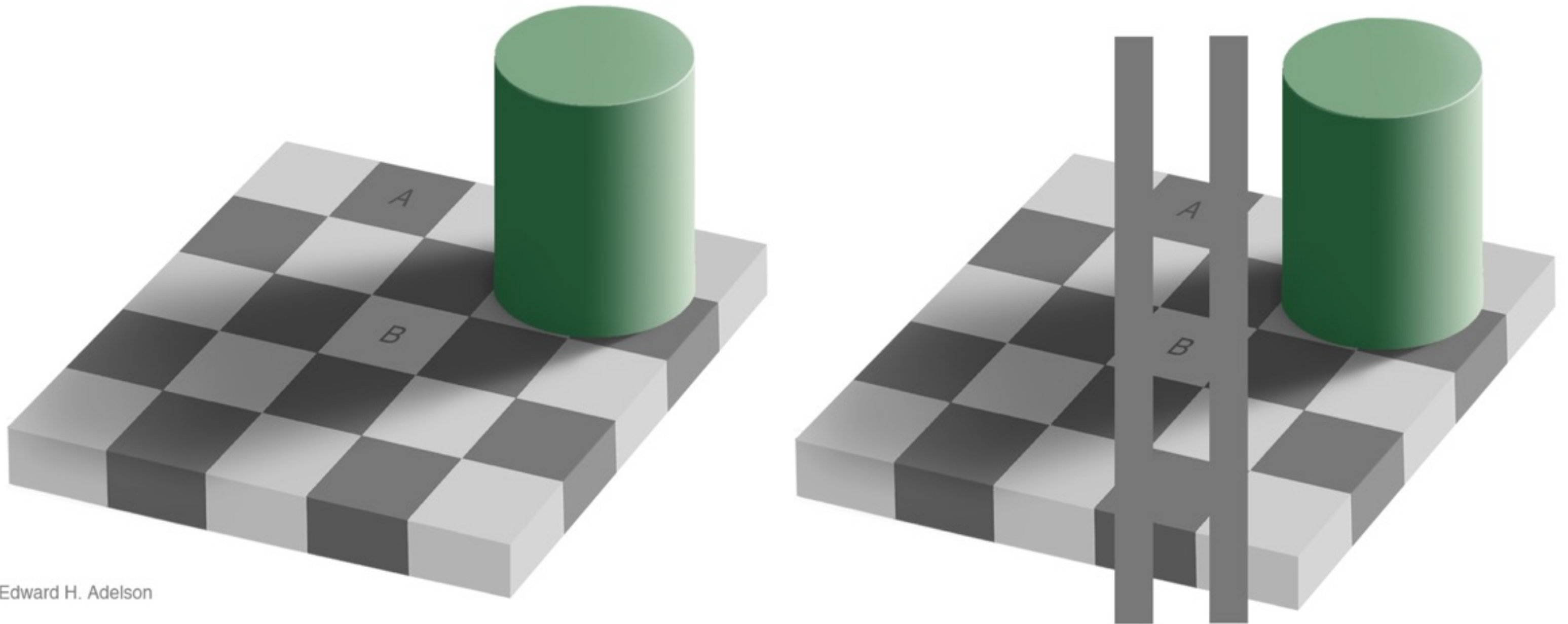
position along
unaligned
common scale



position along
aligned scale

Relative luminance judgements

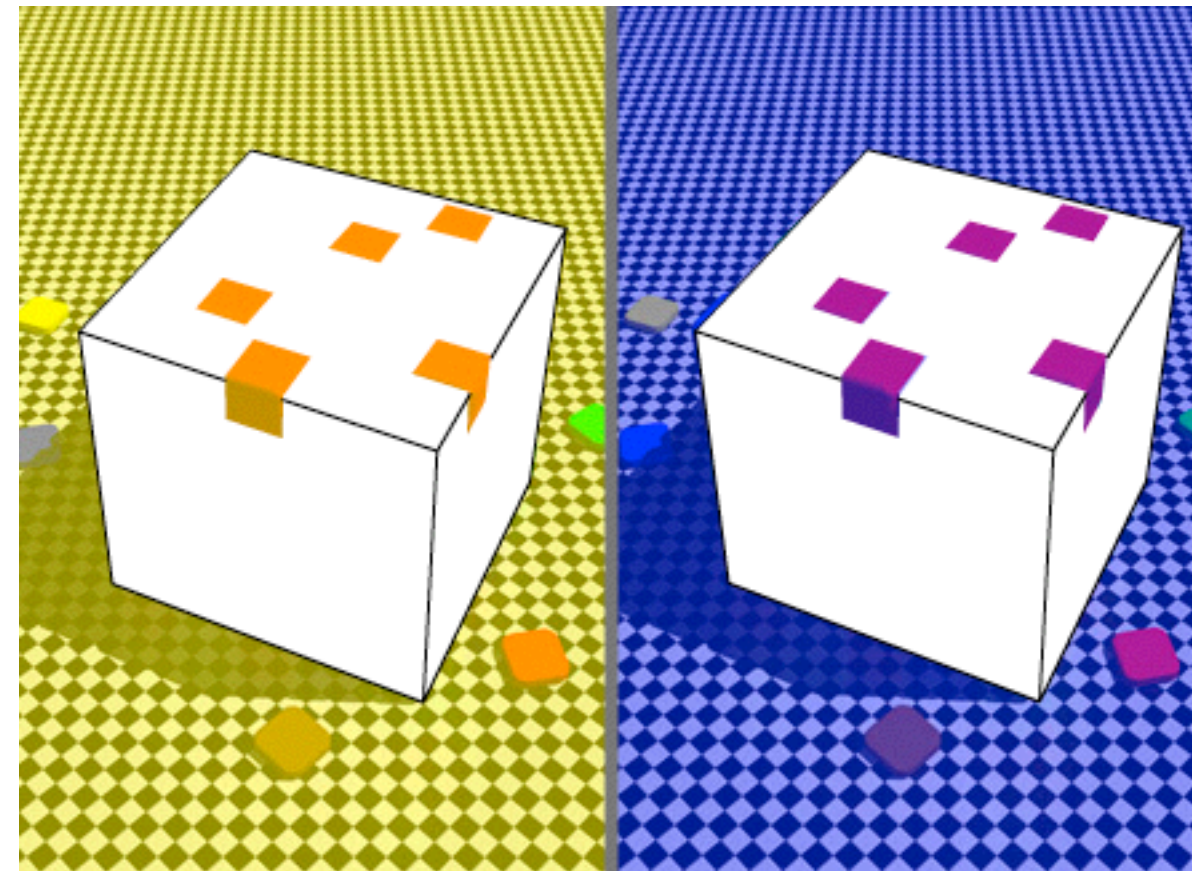
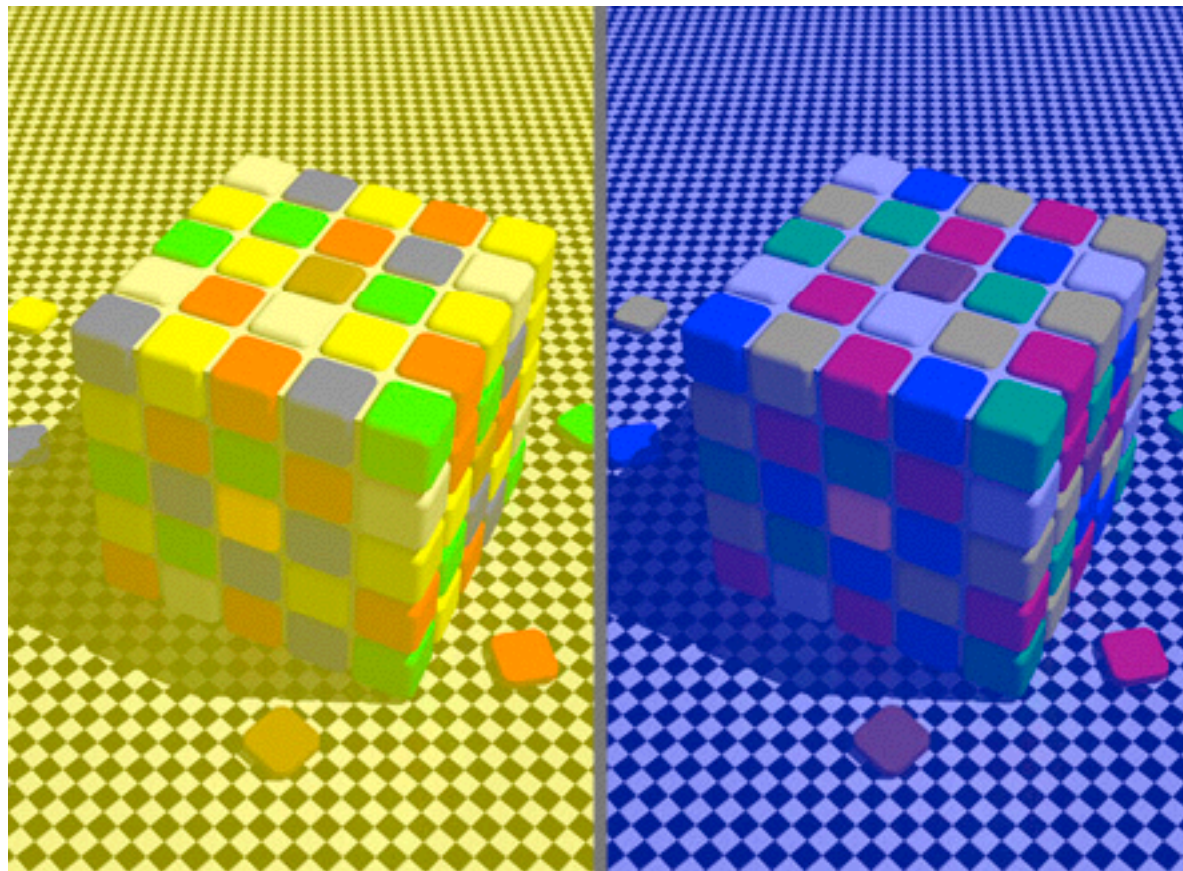
- perception of luminance is contextual based on contrast with surroundings



Edward H. Adelson

Relative color judgements

- color constancy across broad range of illumination conditions



Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
 - *Chap 5: Marks and Channels*
- *On the Theory of Scales of Measurement*. Stevens. Science 103:2684 (1946), 677–680.
- *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. Stevens. Wiley, 1975.
- *Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods*. Cleveland and McGill. Journ. American Statistical Association 79:387 (1984), 531–554.
- *Perception in Vision*. Healey. <http://www.csc.ncsu.edu/faculty/healey/PP>
- *Visual Thinking for Design*. Ware. Morgan Kaufmann, 2008.
- *Information Visualization: Perception for Design*, 3rd edition. Ware. Morgan Kaufmann /Academic Press, 2004.

Outline

- **Session 1** / 10-11:30am
Data Visualization Pitfalls to Avoid
 - Introduction
 - Color
 - Space: 2D vs 3D
- **Session 2** / 12:30-3pm
Visualization Analysis & Design, In More Depth
 - Marks and Channels, Perception
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

How?

Encode

→ Arrange

→ Express



→ Separate



→ Order



→ Align



→ Use



→ Map

from **categorical** and **ordered** attributes

→ Color

→ Hue



→ Saturation



→ Luminance



→ Size, Angle, Curvature, ...



→ Shape



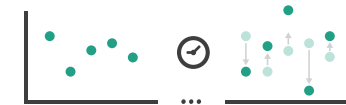
→ Motion

Direction, Rate, Frequency, ...

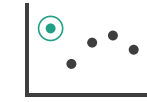


Manipulate

→ Change



→ Select



→ Navigate

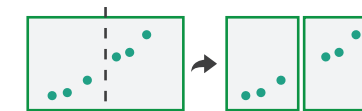


Facet

→ Juxtapose



→ Partition



→ Superimpose

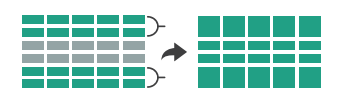


Reduce

→ Filter



→ Aggregate



→ Embed



What?

Why?

How?

Encode tables: Arrange space

Encode

➔ Arrange

➔ Express



➔ Separate



➔ Order



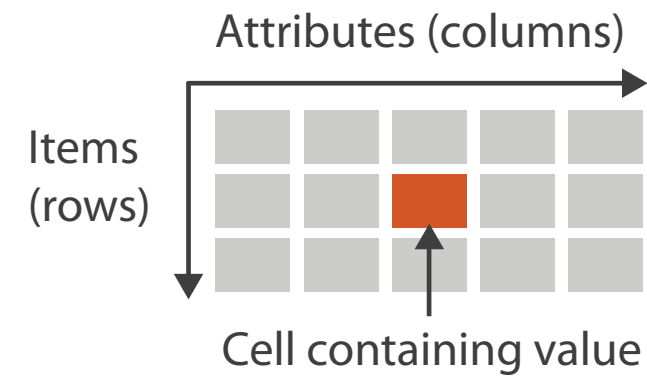
➔ Align



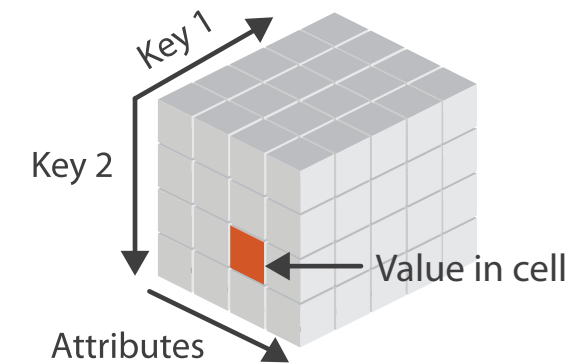
Keys and values

- 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 key count
 - 0, 1, 2, many...

→ Tables



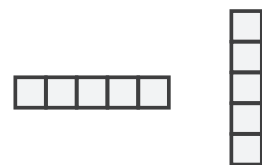
→ Multidimensional Table



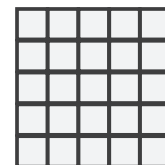
⊕ Express Values



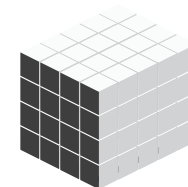
→ 1 Key
List



→ 2 Keys
Matrix



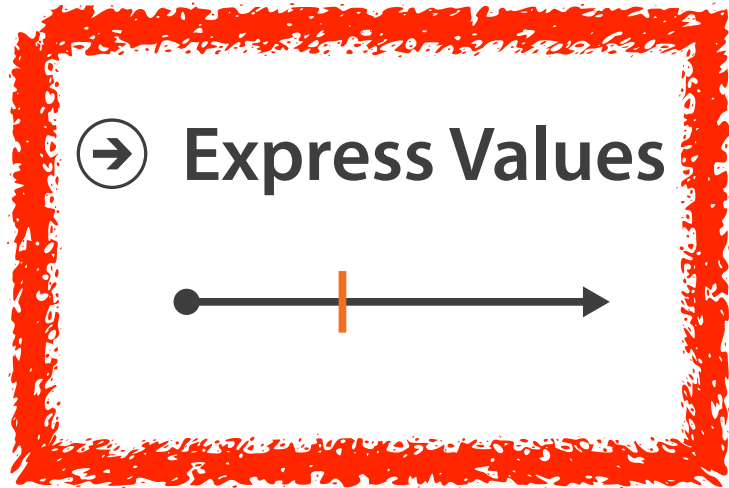
→ 3 Keys
Volume



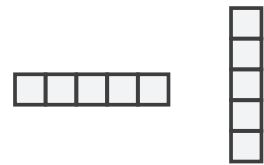
→ Many Keys
Recursive Subdivision



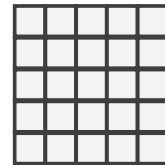
0 Keys



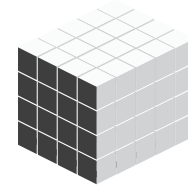
→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume



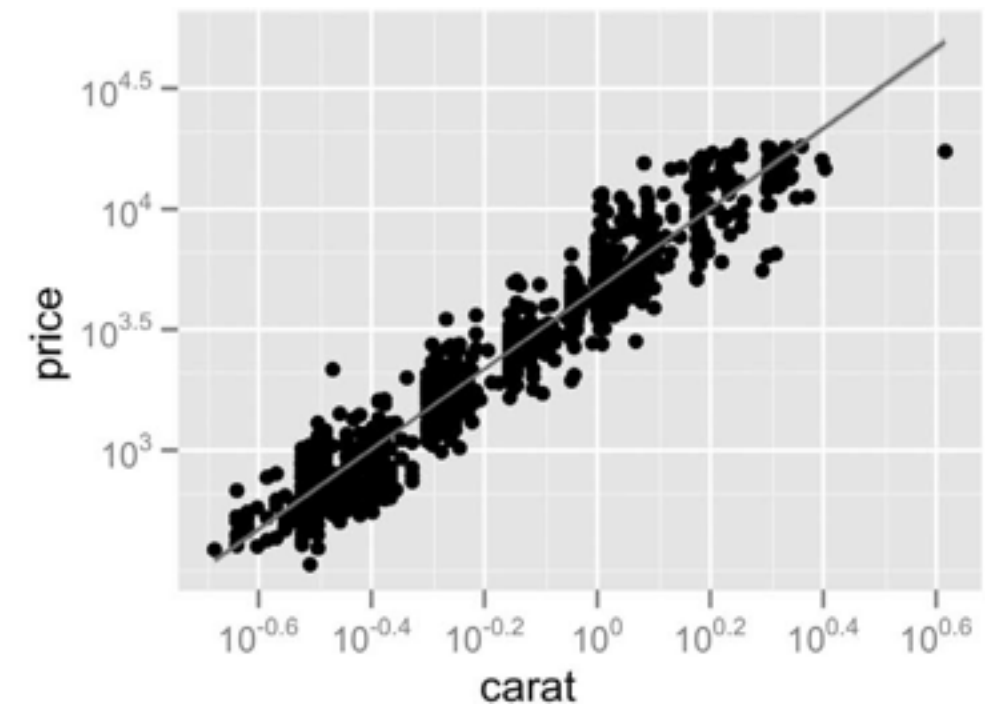
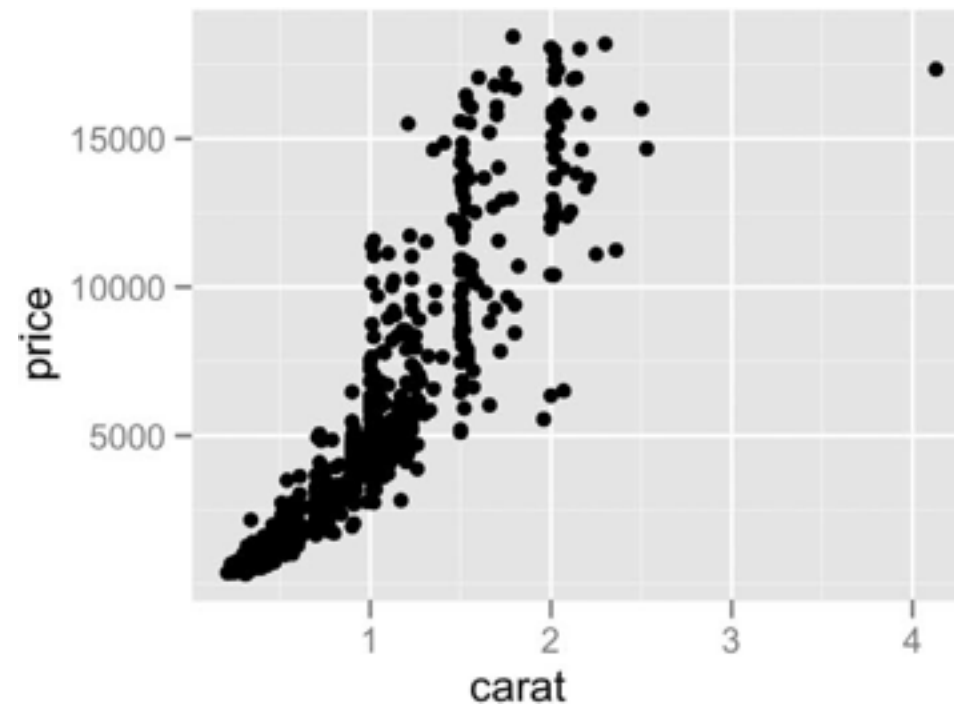
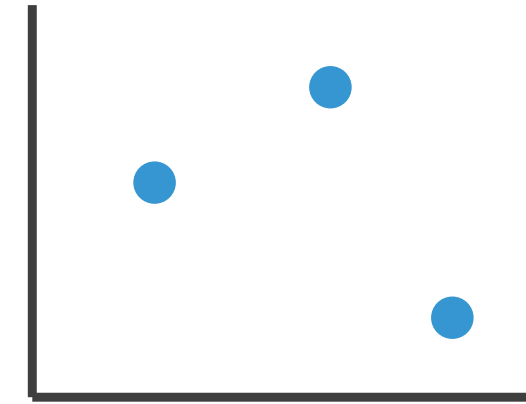
→ Many Keys
Recursive Subdivision



Idiom: scatterplot

- *express* values
 - 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

➔ Express Values

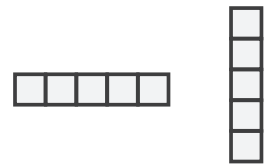


Some keys

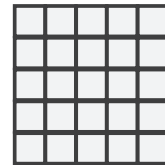
→ Express Values



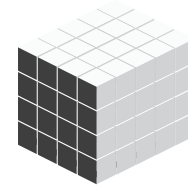
→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume



→ Many Keys
Recursive Subdivision

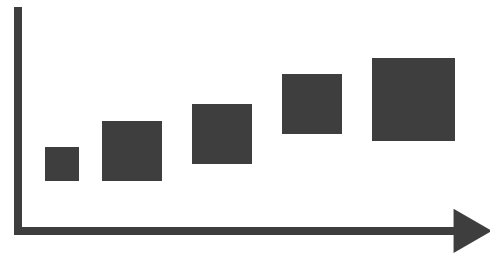


Some keys: Categorical regions

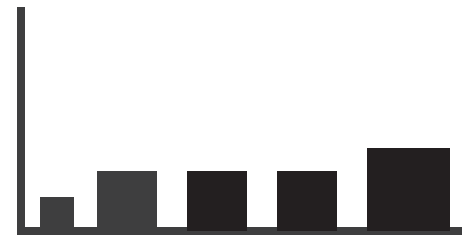
→ Separate



→ Order

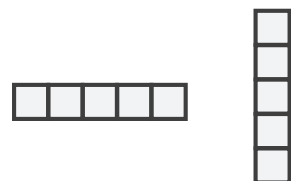


→ Align

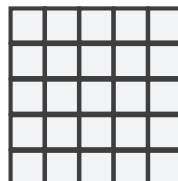


- **regions:** contiguous bounded areas distinct from each other
 - using space to *separate* (proximity)
 - following expressiveness principle for categorical attributes
- use ordered attribute to *order* and *align* regions

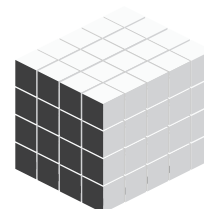
→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume



→ Many Keys
Recursive Subdivision



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

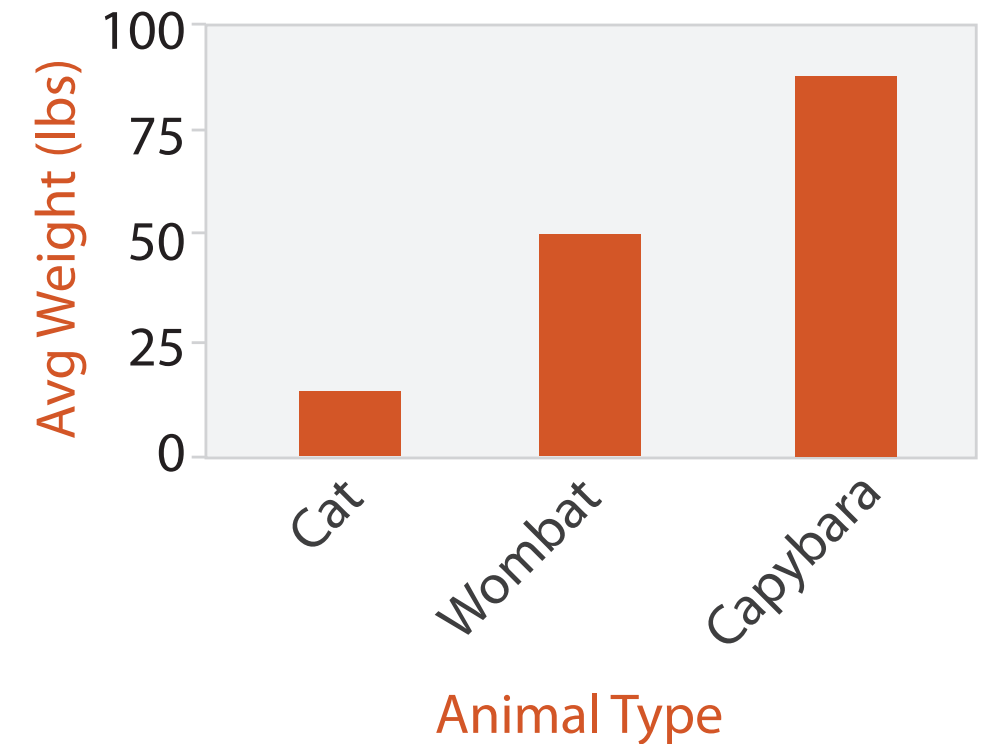
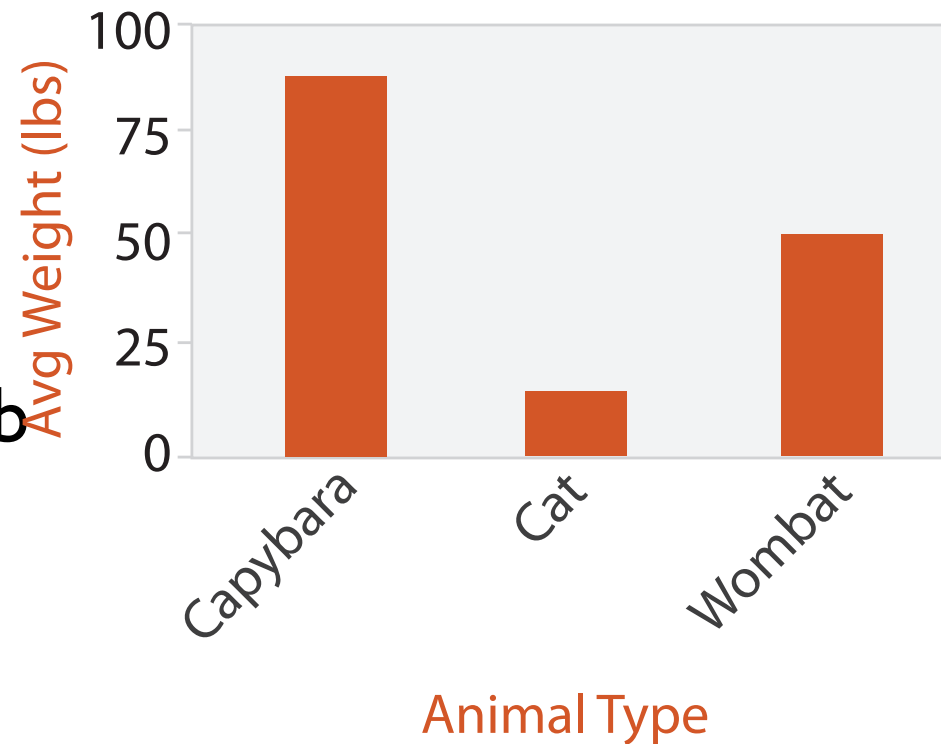
» by label (alphabetical), by length attrib (data-driven)

- task

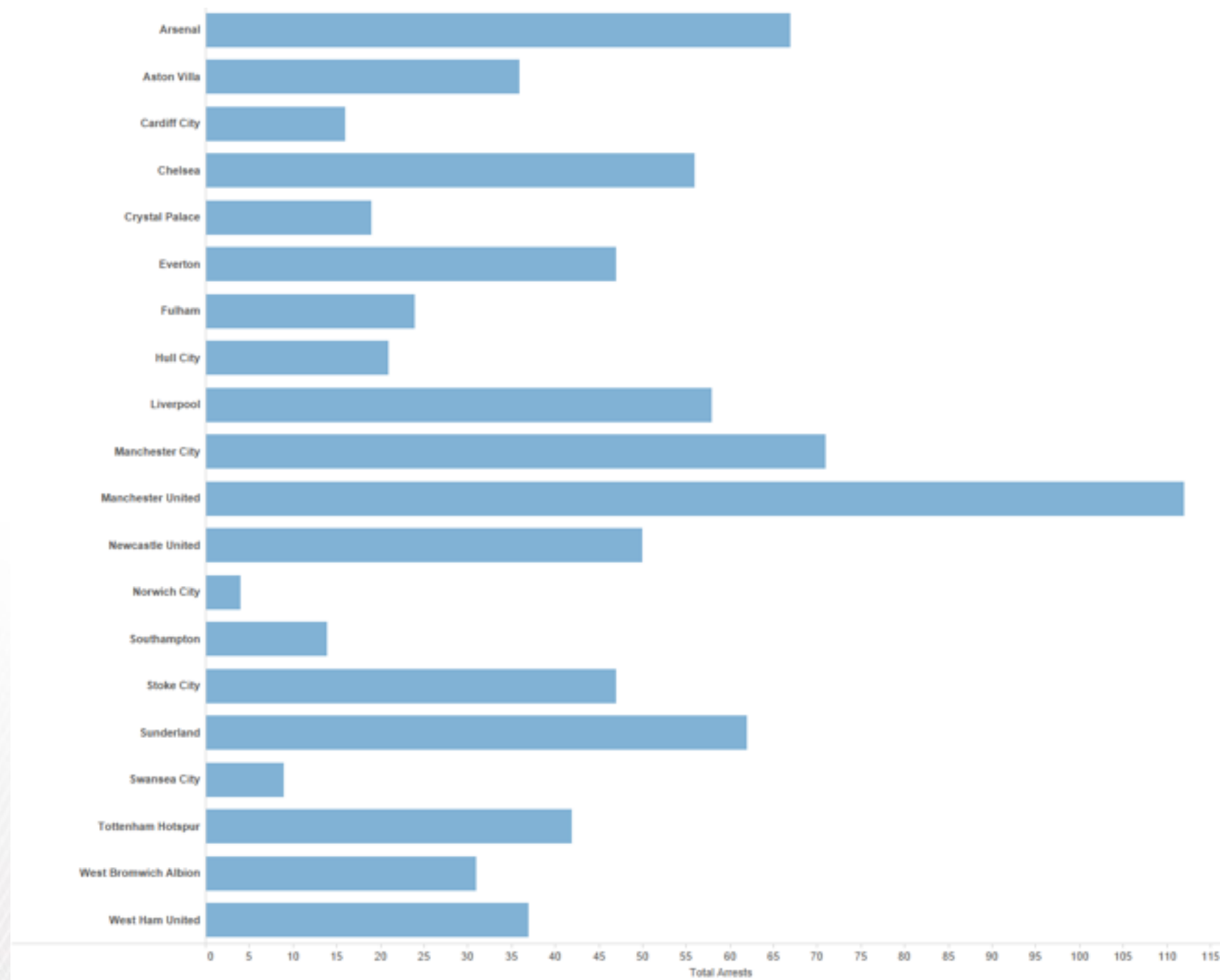
- compare, lookup values

- scalability

- dozens to hundreds of levels for key attrib



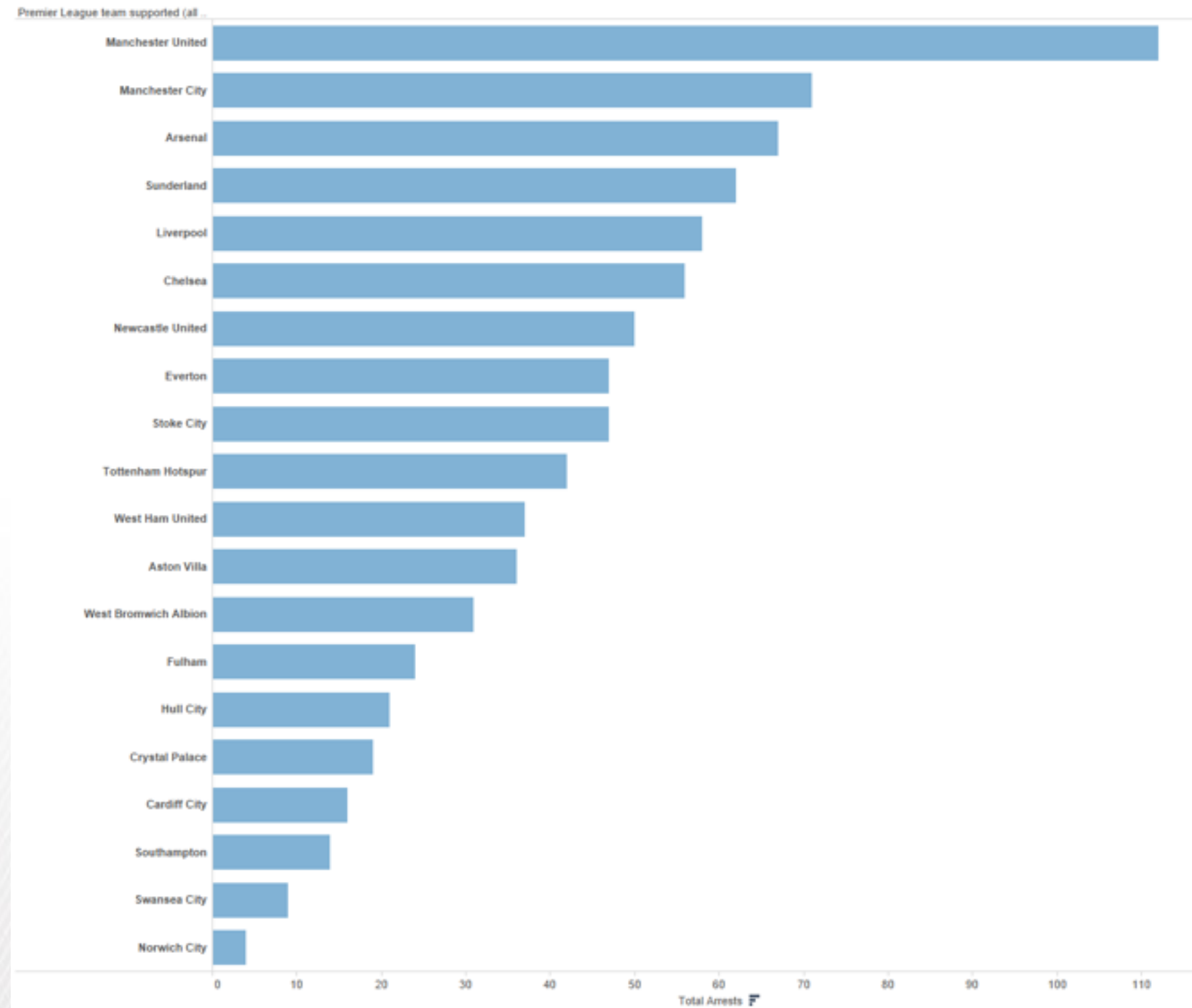
Separated and Aligned but not Ordered



LIMITATION: Hard to know rank. What's the 4th most? The 7th?

[Slide courtesy of Ben Jones]

Separated, Aligned and Ordered



[Slide courtesy of Ben Jones]

Separated but not Ordered or Aligned



LIMITATION: Hard to make comparisons

[Slide courtesy of Ben Jones]

Idiom: **line chart**

- one key, one value

- data

- 2 quant attribs

- mark: points

- line connection marks between them

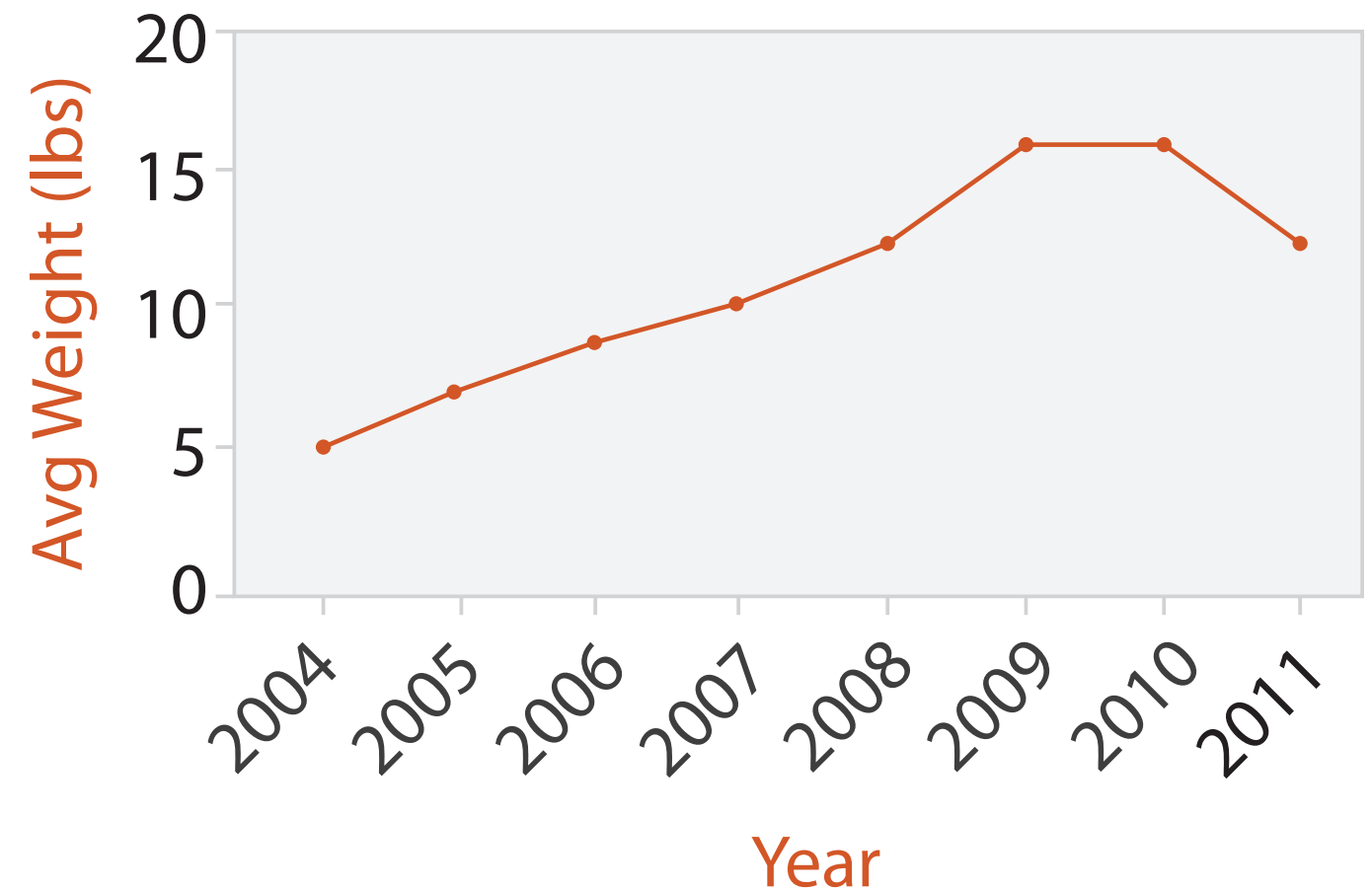
- channels

- 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



Idiom: **line chart / dot plot**

- one key, one value

- data

- 2 quant attribs

- mark: points

- line connection marks between them

- channels

- aligned lengths to express quant value
- separated and ordered by key attrib into horizontal regions

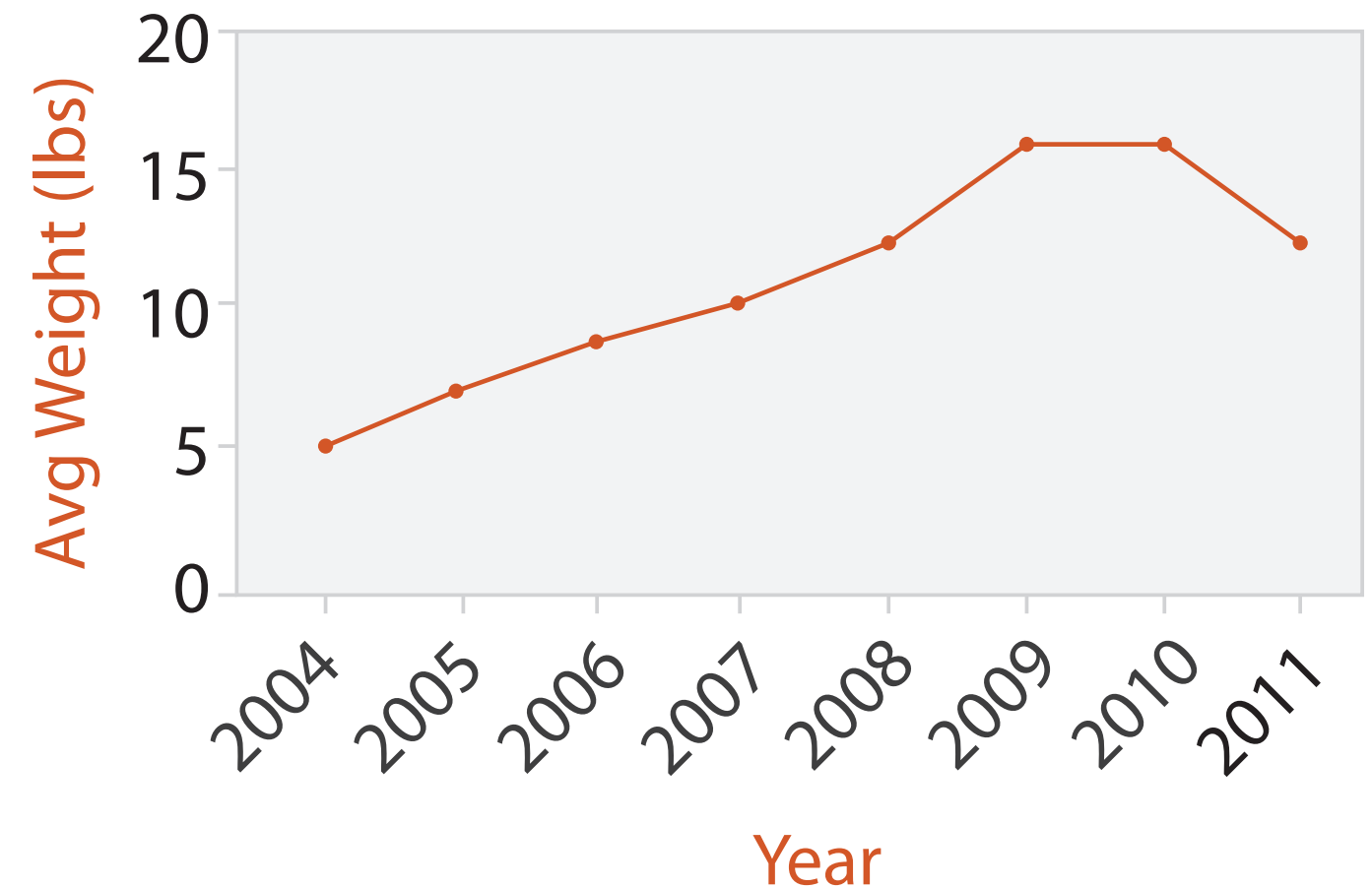
- task

- find trend

- connection marks emphasize ordering of items along key axis by explicitly showing relationship between one item and the next

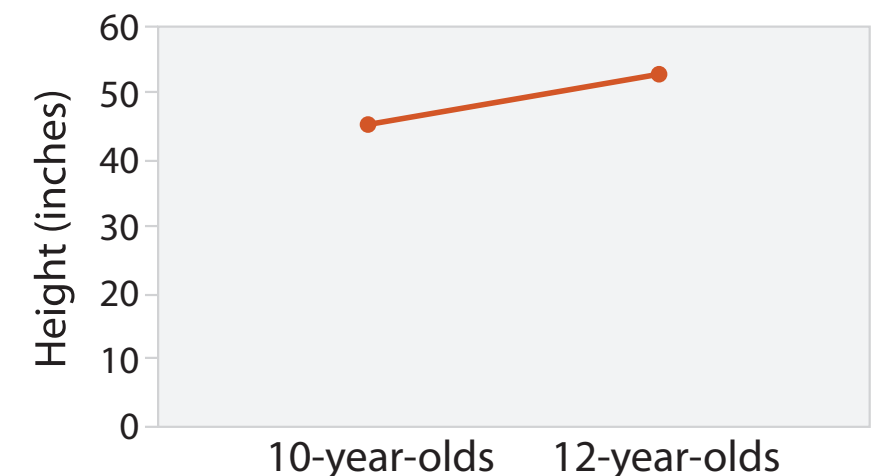
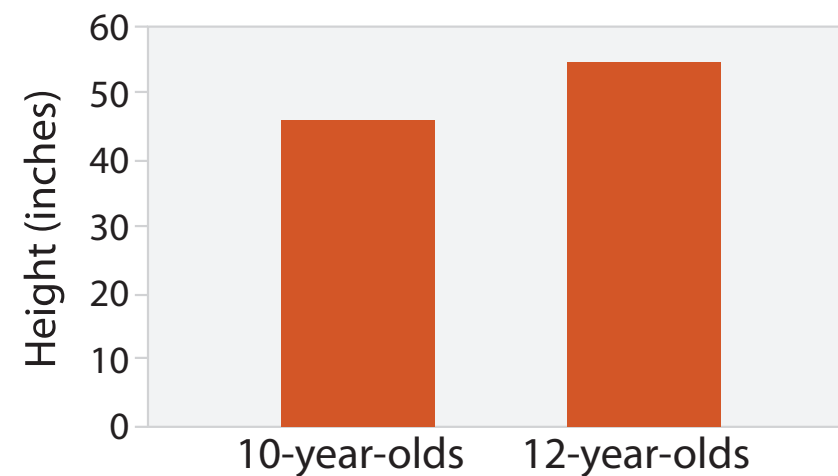
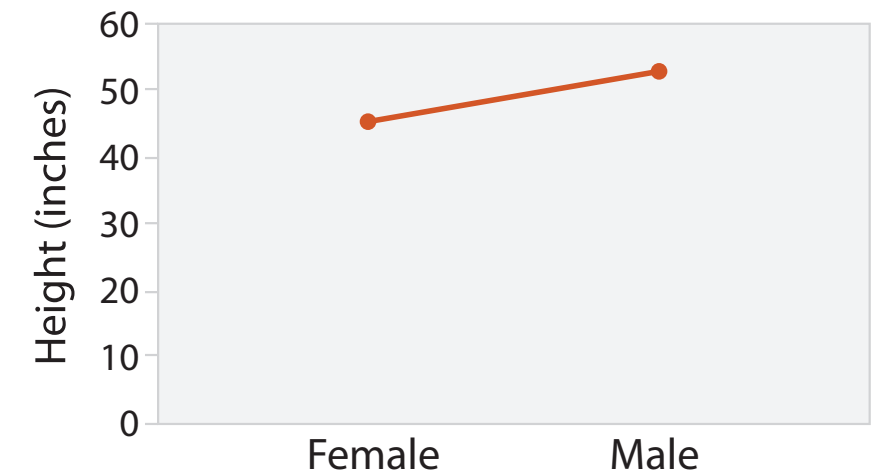
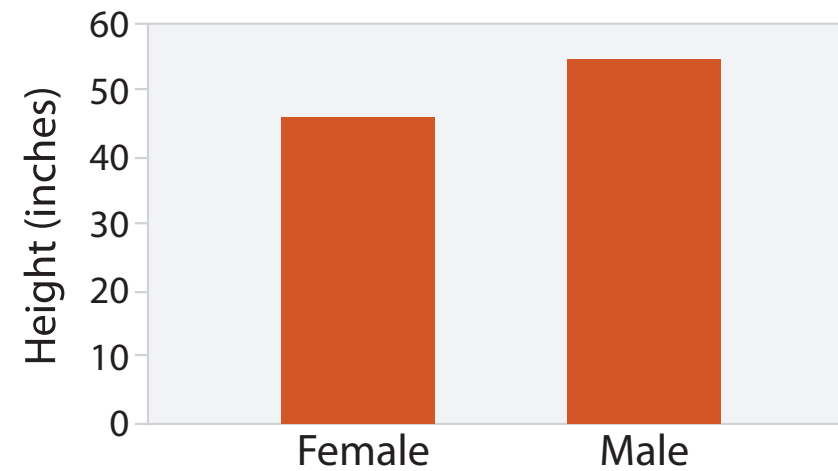
- scalability

- hundreds of key levels, hundreds of value levels



Choosing bar vs line charts

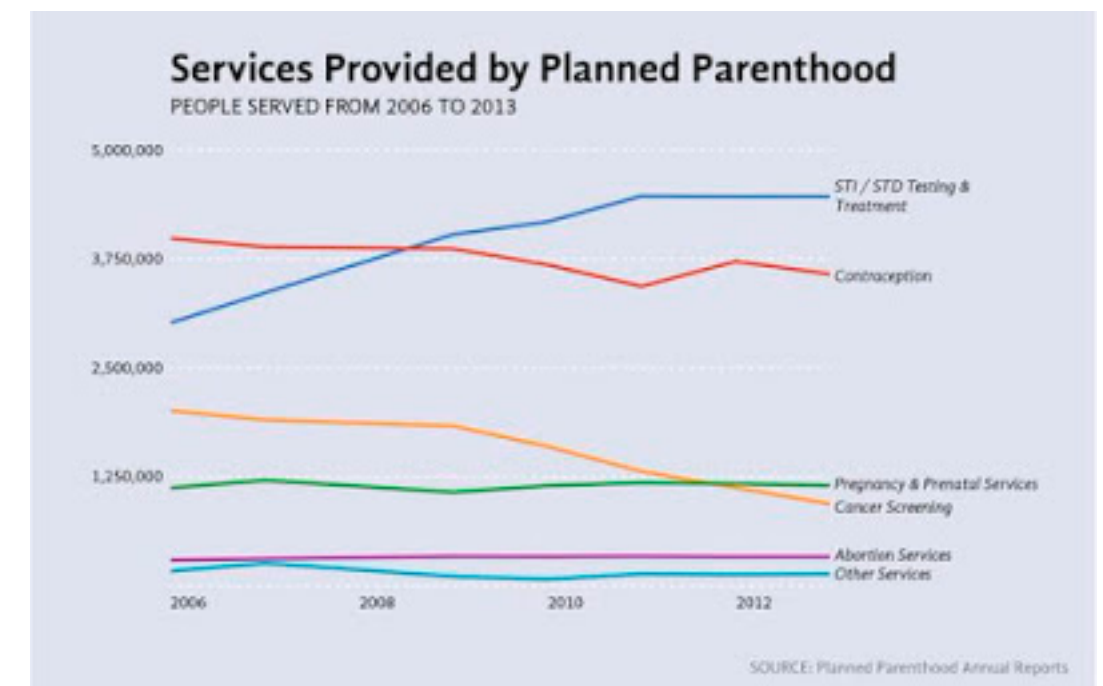
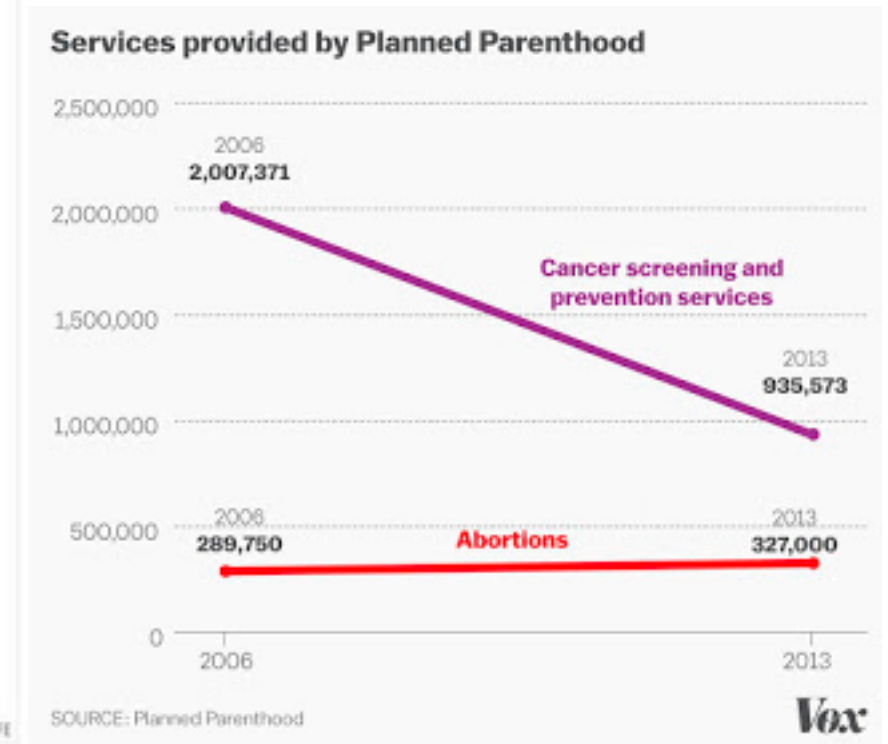
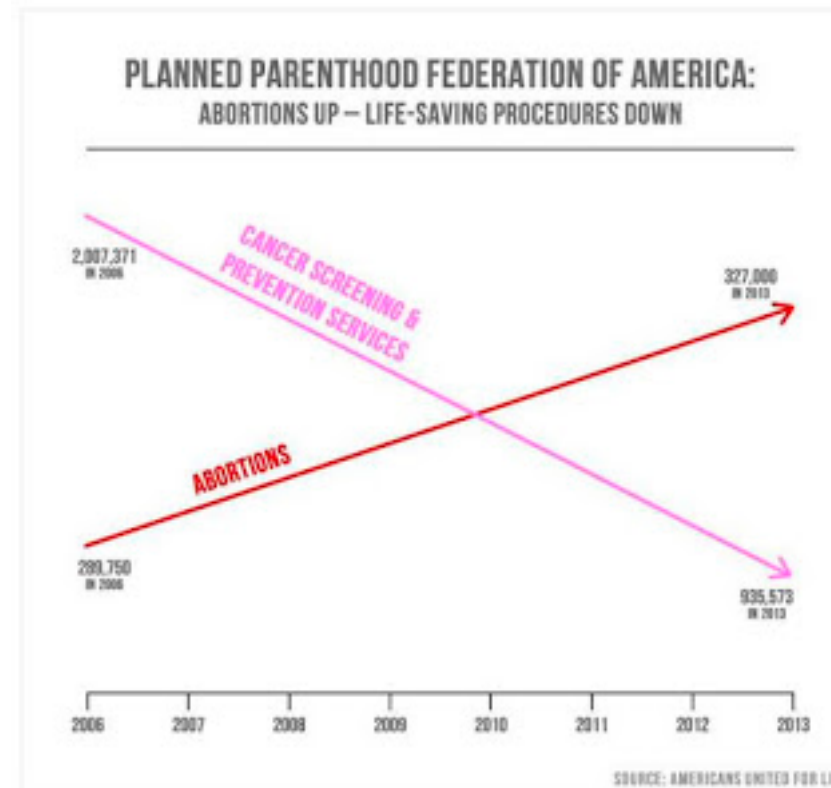
- depends on type of key attrib
 - bar charts if categorical
 - line charts if ordered
- do not use line charts for categorical key attribs
 - violates expressiveness principle
 - implication of trend so strong that it overrides semantics!
 - “The more male a person is, the taller he/she is”



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

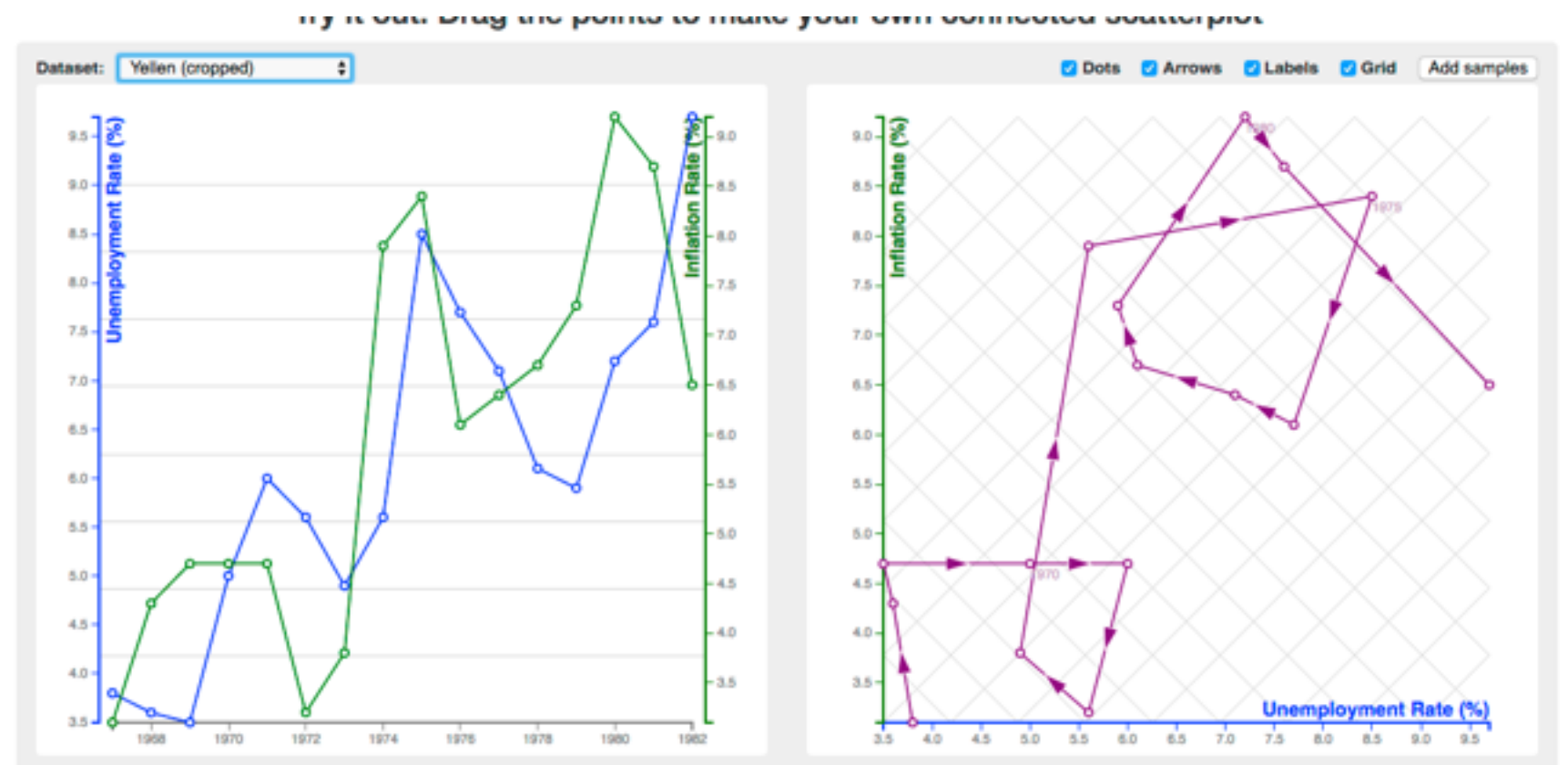
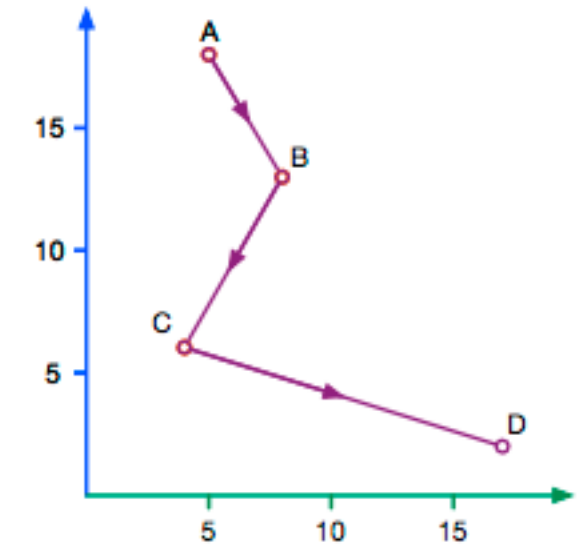
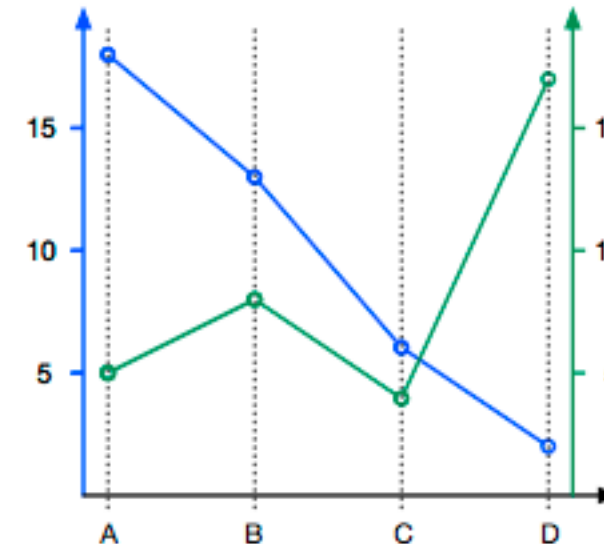
Chart axes

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

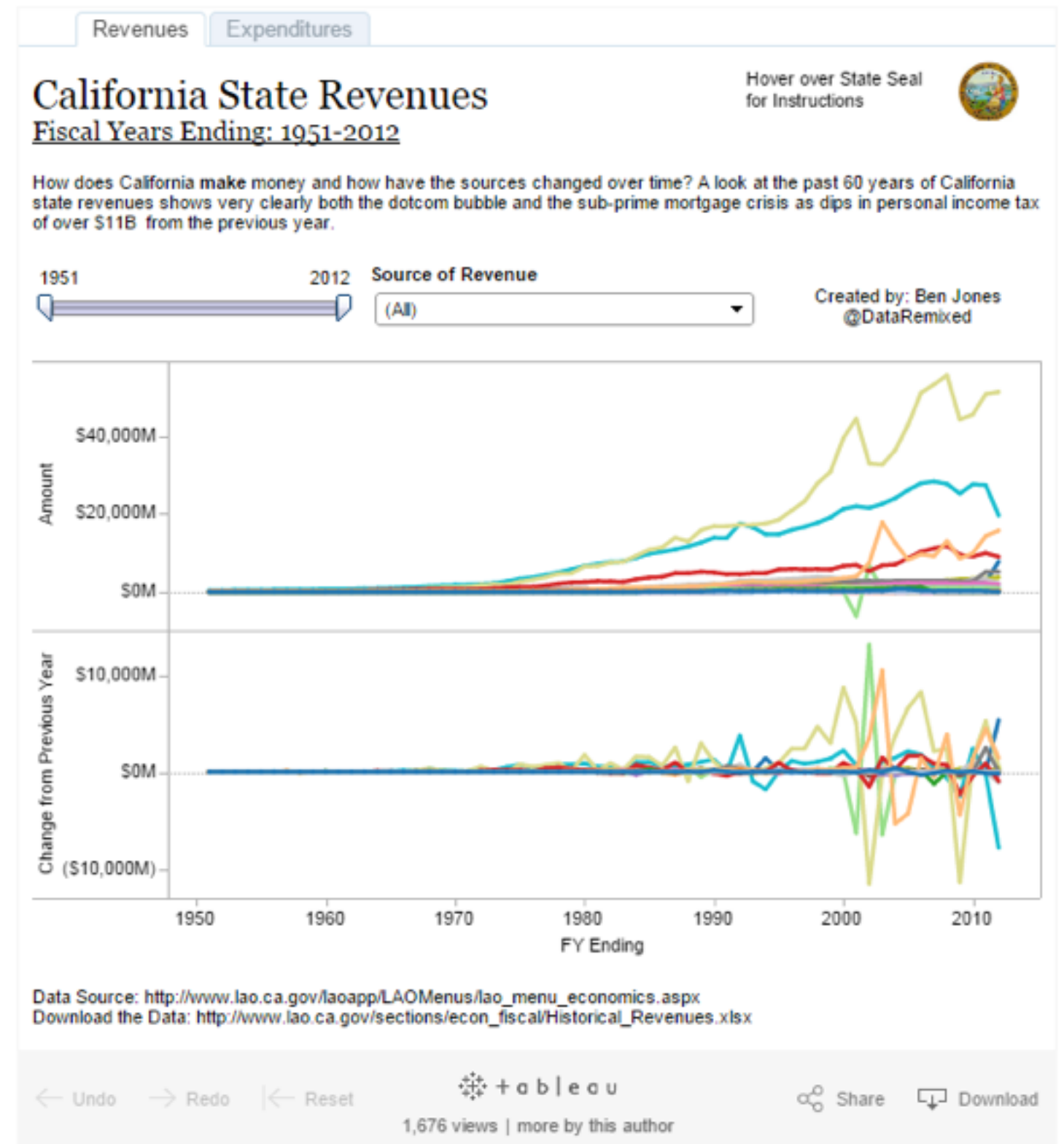


[The Connected Scatterplot for Presenting Paired Time Series.
Haroz, Kosara and Franconeri. IEEE TVCG 22(9):2174-86, 2016.]

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

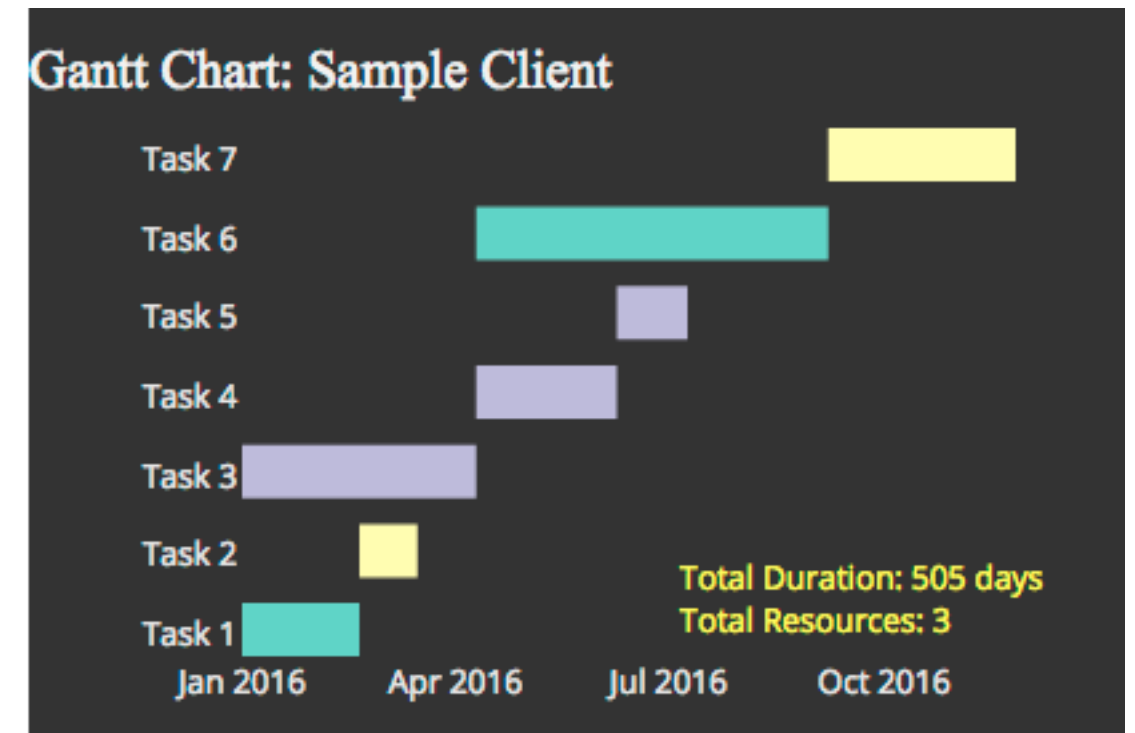
Idiom: Indexed line charts

- data: 2 quant attires
 - 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

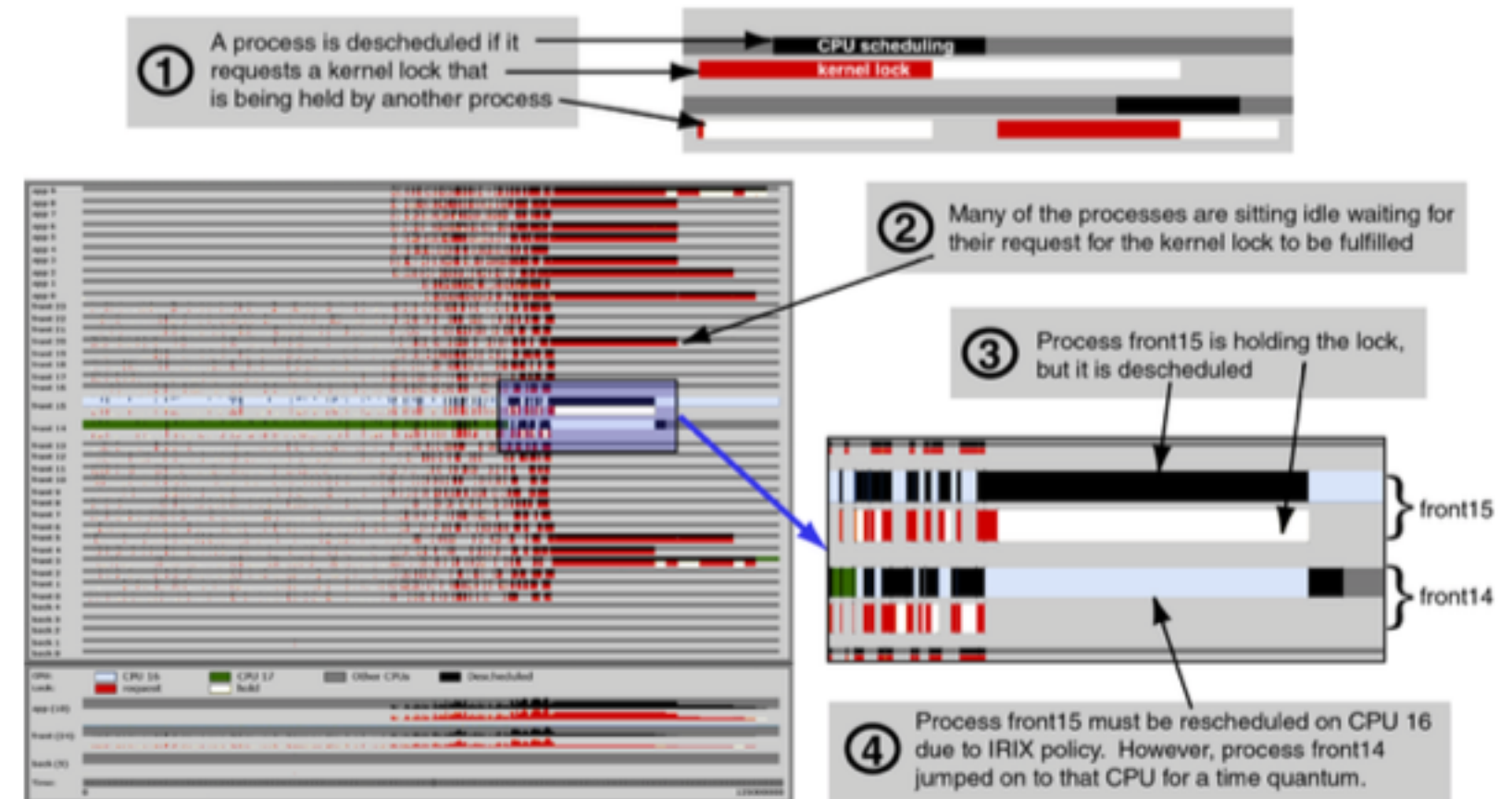


Idiom: Gantt charts

- 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
 - hundreds of value levels



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

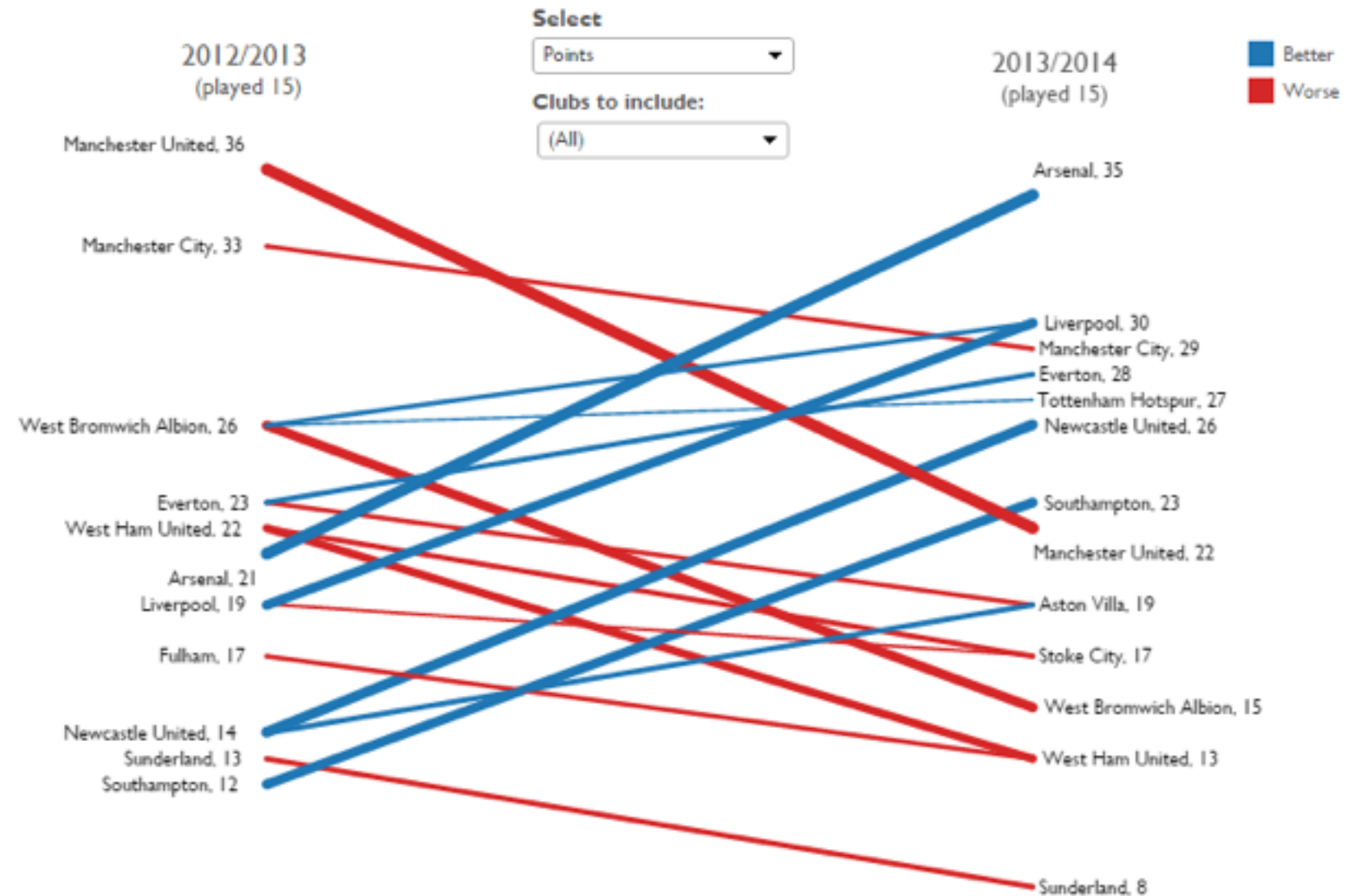


[Performance Analysis and Visualization of Parallel Systems Using SimOS and Rivet: A Case Study. Bosch, Stolte, Stoll, Rosenblum, and Hanrahan. Proc. HPCA 2000.]

Idiom: Slopegraphs

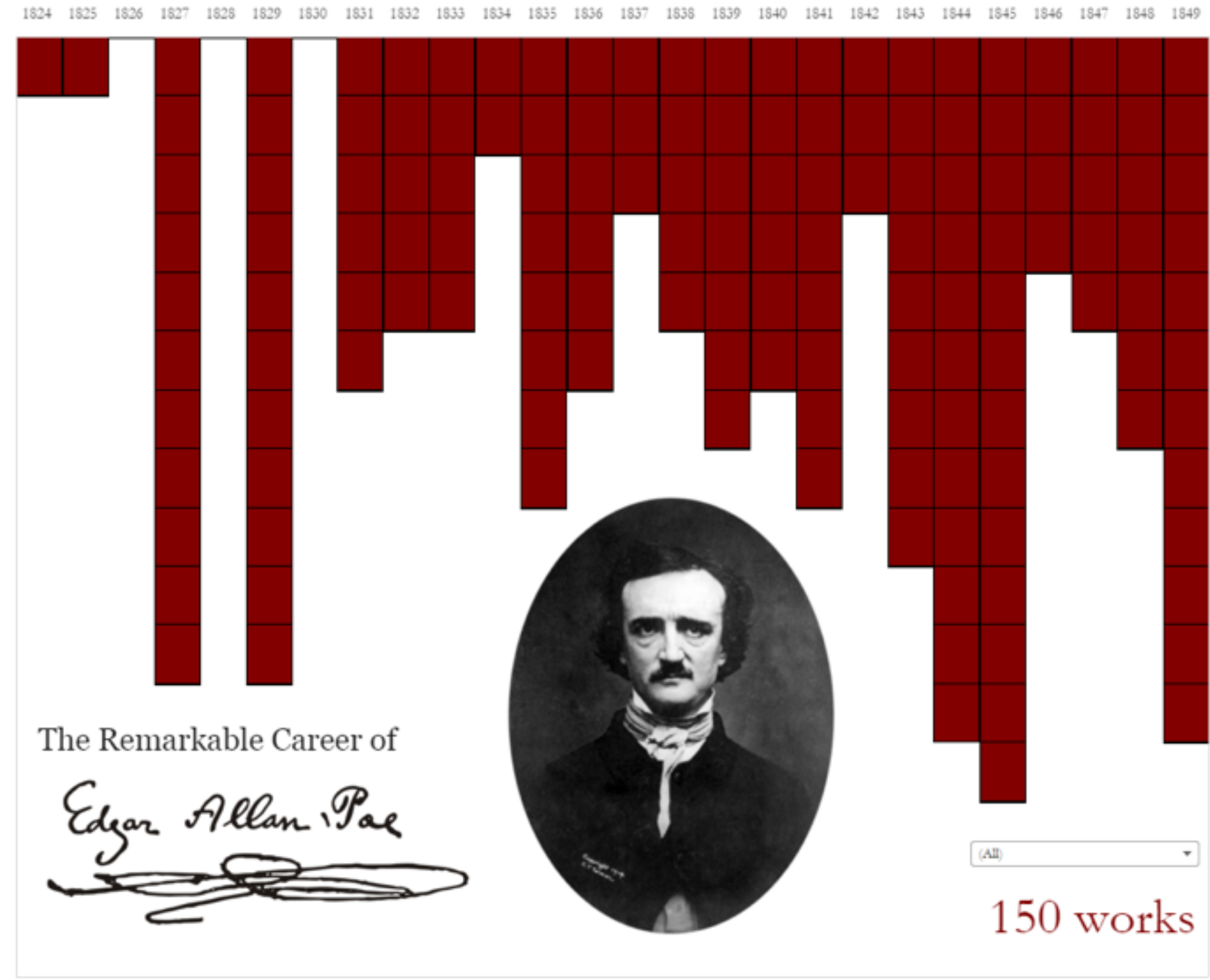
- two values
 - data
 - 2 quant value attribs
 - mark: point + line
 - line connecting mark between pts
 - channels
 - 2 vertical pos: express attrib value
 - task
 - emphasize changes in rank/value
 - scalability
 - hundreds of value levels

Barclay's Premier League Tables: Comparing 2012/2013 Starts to 2013/2014 Starts



Breaking conventions

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



Source: https://en.wikipedia.org/wiki/Edgar_Allan_Poe_bibliography

Ben Jones, 7 October 2015

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

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

2 Keys

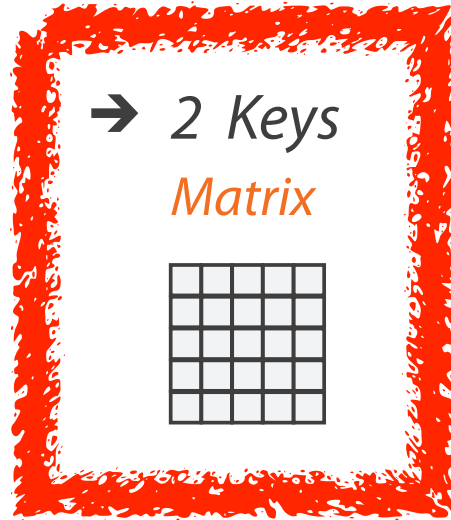
➔ Express Values



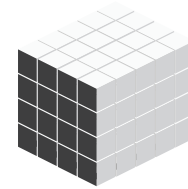
➔ 1 Key
List



➔ 2 Keys
Matrix



➔ 3 Keys
Volume



➔ Many Keys
Recursive Subdivision



Idiom: heatmap

- two keys, one value

- data

- 2 categ attribs (gene, experimental condition)
- 1 quant attrib (expression levels)

- marks: area

- separate and align in 2D matrix
 - indexed by 2 categorical attributes

- channels

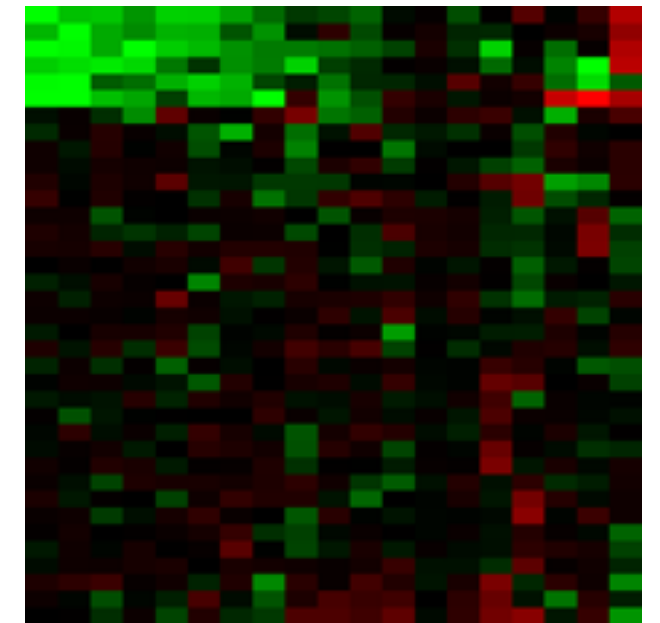
- color by quant attrib
 - (ordered diverging colormap)

- task

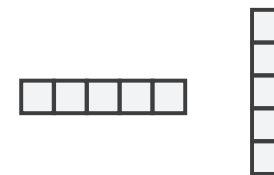
- find clusters, outliers

- scalability

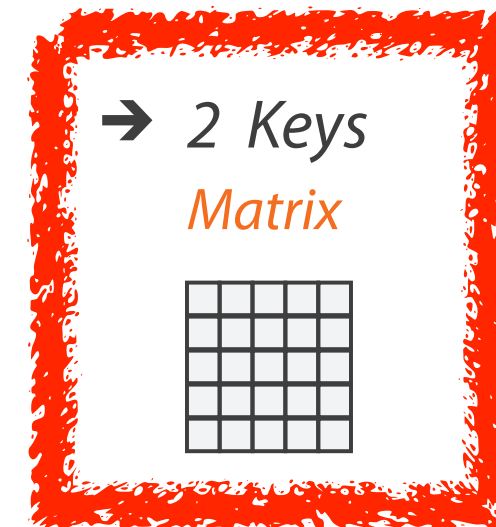
- 1M items, 100s of categ levels, ~10 quant attrib levels



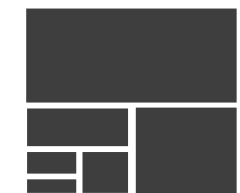
→ 1 Key
List



→ 2 Keys
Matrix

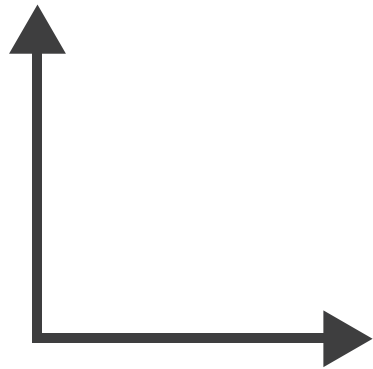


→ Many Keys
Recursive Subdivision

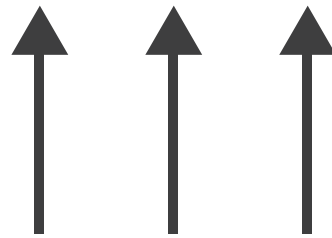


➔ Axis Orientation

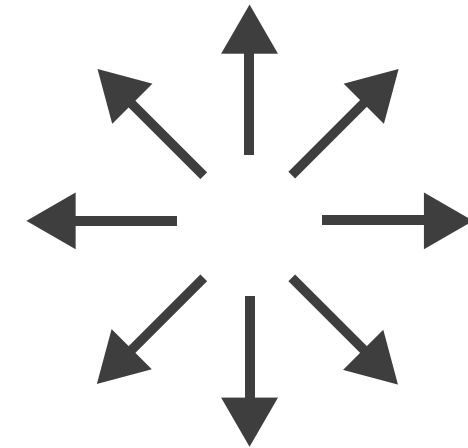
➔ Rectilinear



➔ Parallel



➔ Radial



Idioms: scatterplot matrix, parallel coordinates

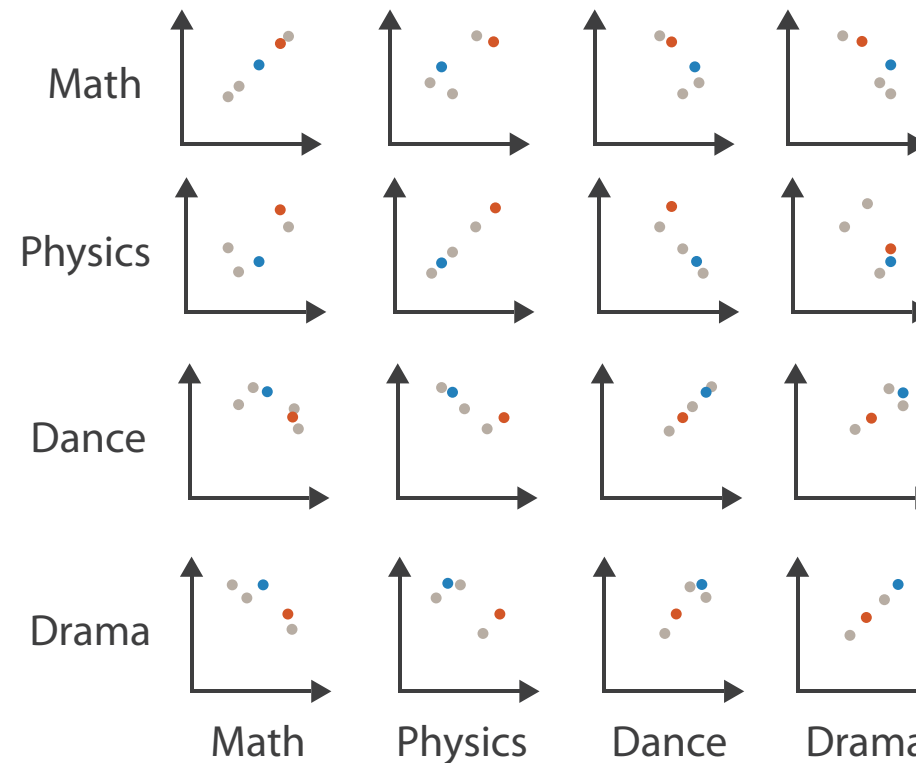
- scatterplot matrix (SPLOM)

- rectilinear axes, point mark
- all possible pairs of axes
- scalability
 - one dozen attribs
 - dozens to hundreds of items

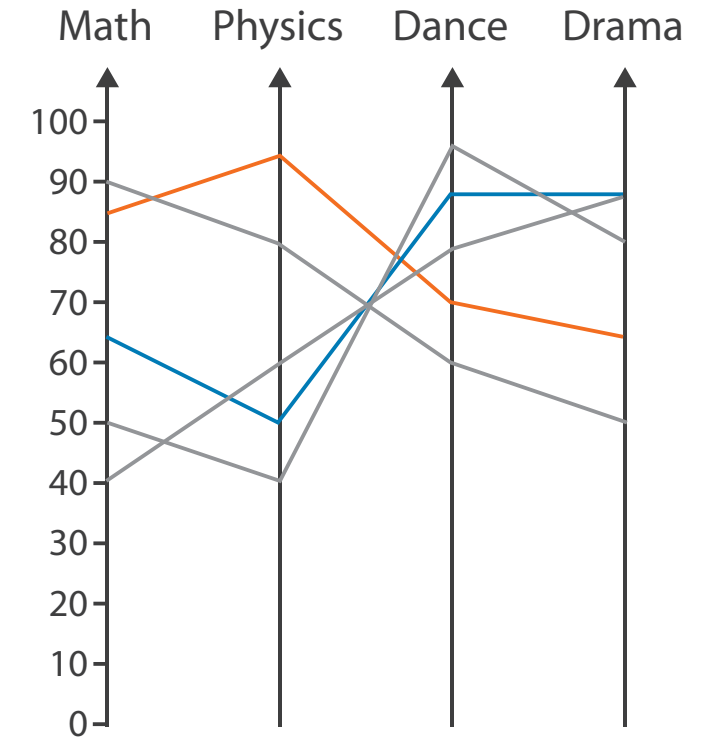
- parallel coordinates

- parallel axes, jagged line representing item
- rectilinear axes, item as point
 - axis ordering is major challenge
- scalability
 - dozens of attribs
 - hundreds of items

Scatterplot Matrix



Parallel Coordinates

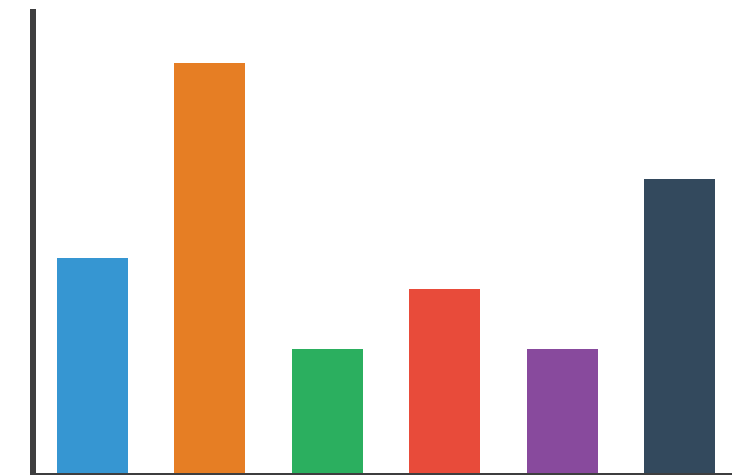


Table

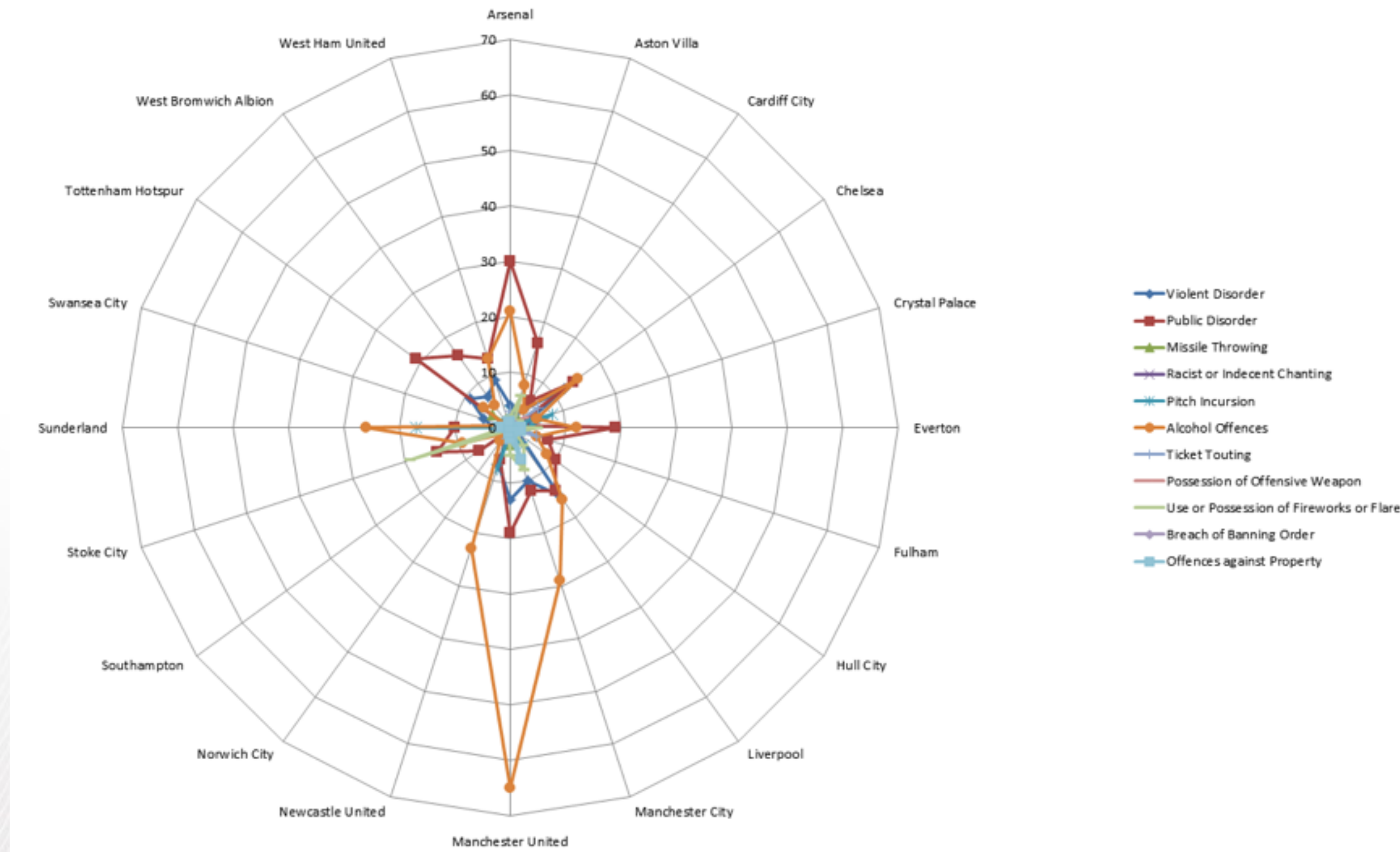
Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

Idioms: radial bar chart, star plot

- radial bar chart
 - radial axes meet at central ring, line mark
- star plot
 - radial axes, meet at central point, line mark
- bar chart
 - rectilinear axes, aligned vertically
 - less accurate than aligned with radial



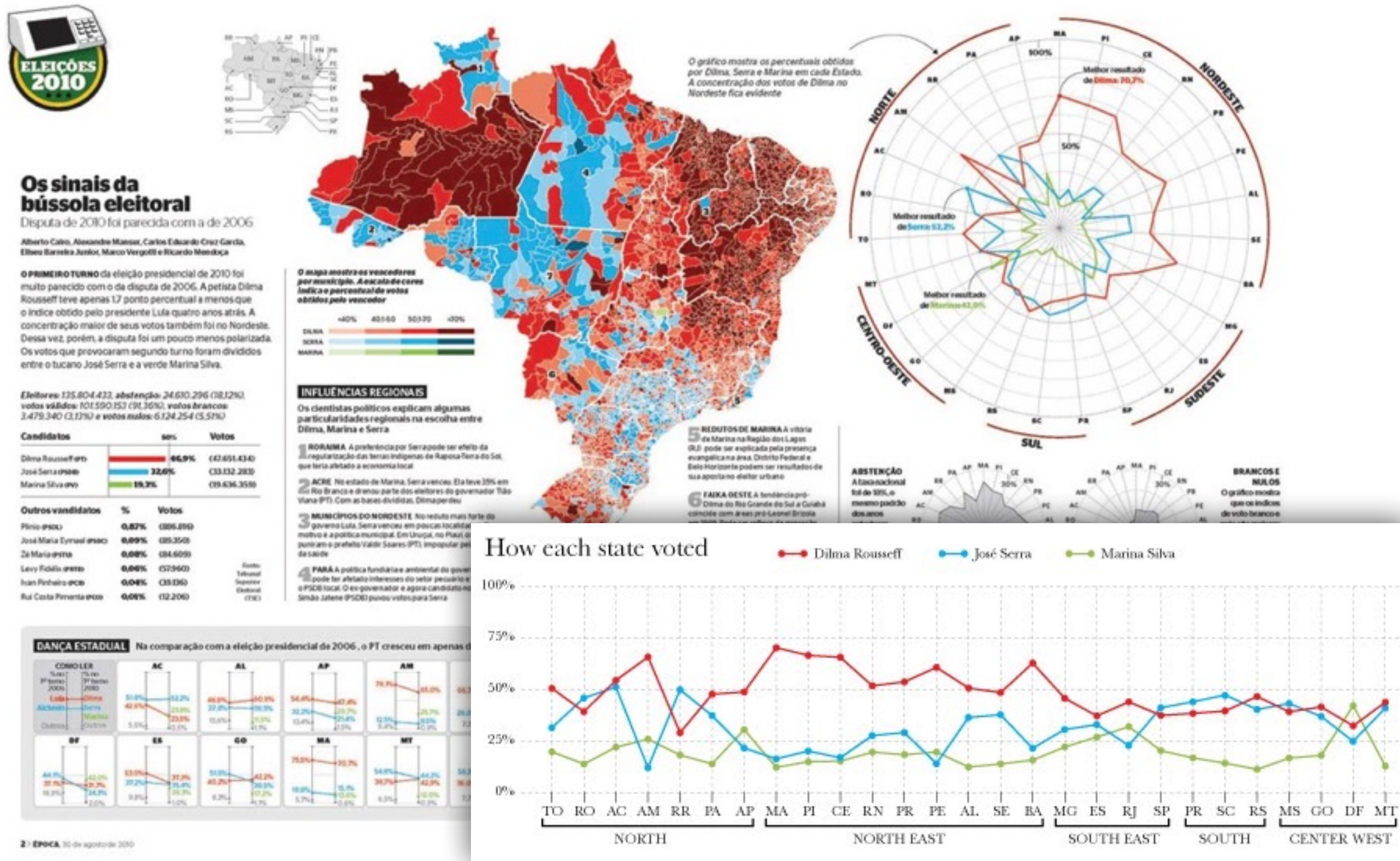
Radial Orientation: Radar Plots



LIMITATION: Not good when categories aren't cyclic

[Slide courtesy of Ben Jones]

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

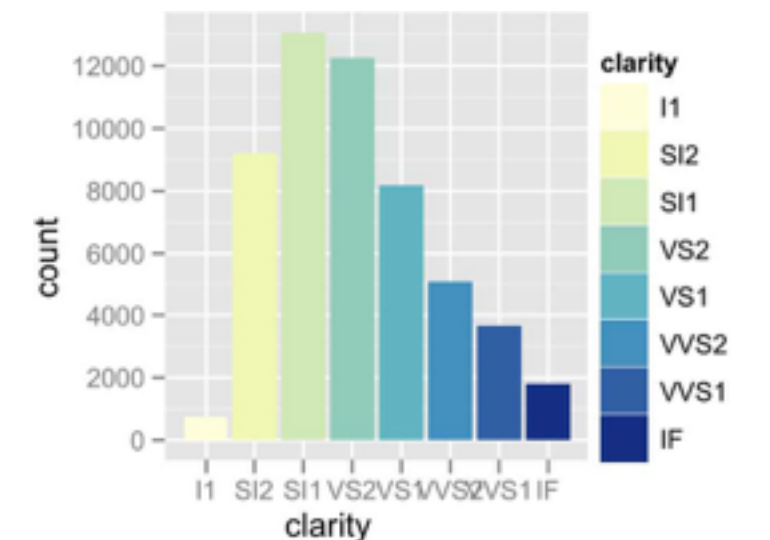
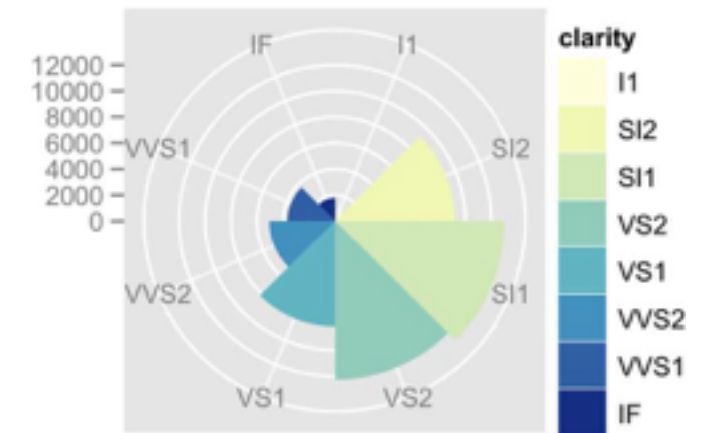
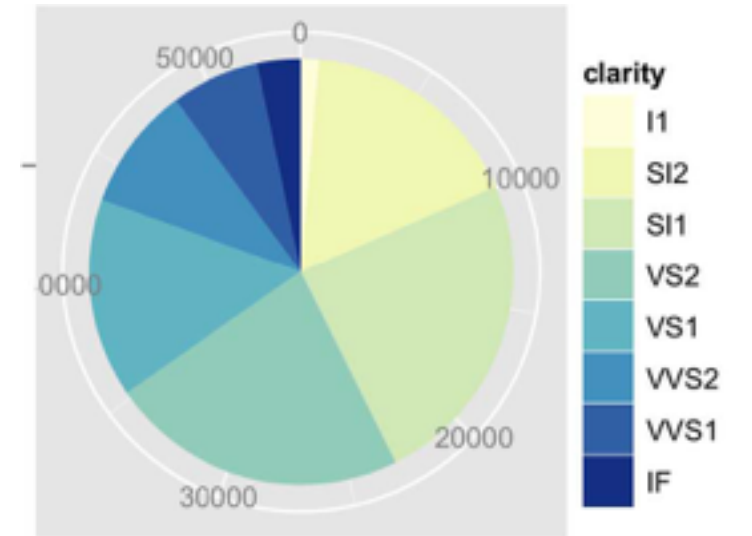


<http://www.thefunctionalart.com/2012/11/radar-graphs-avoid-them-999-of-time.html>

[Slide courtesy of Ben Jones]

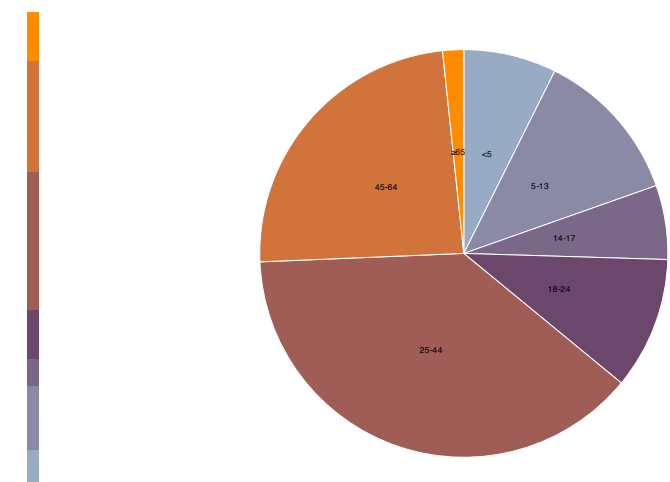
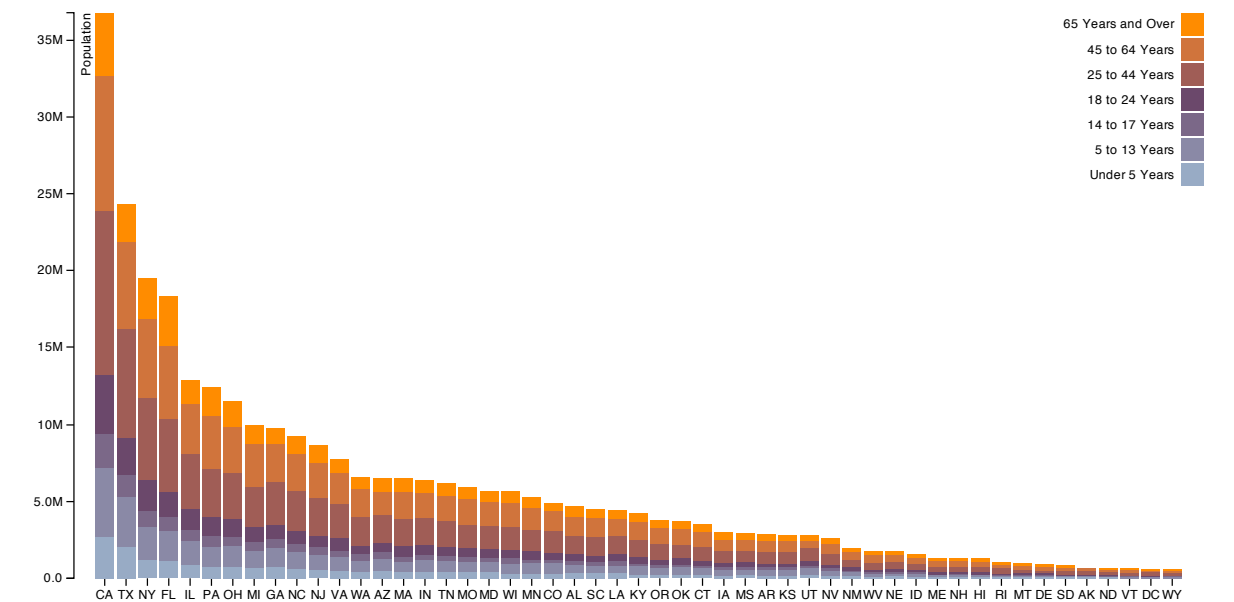
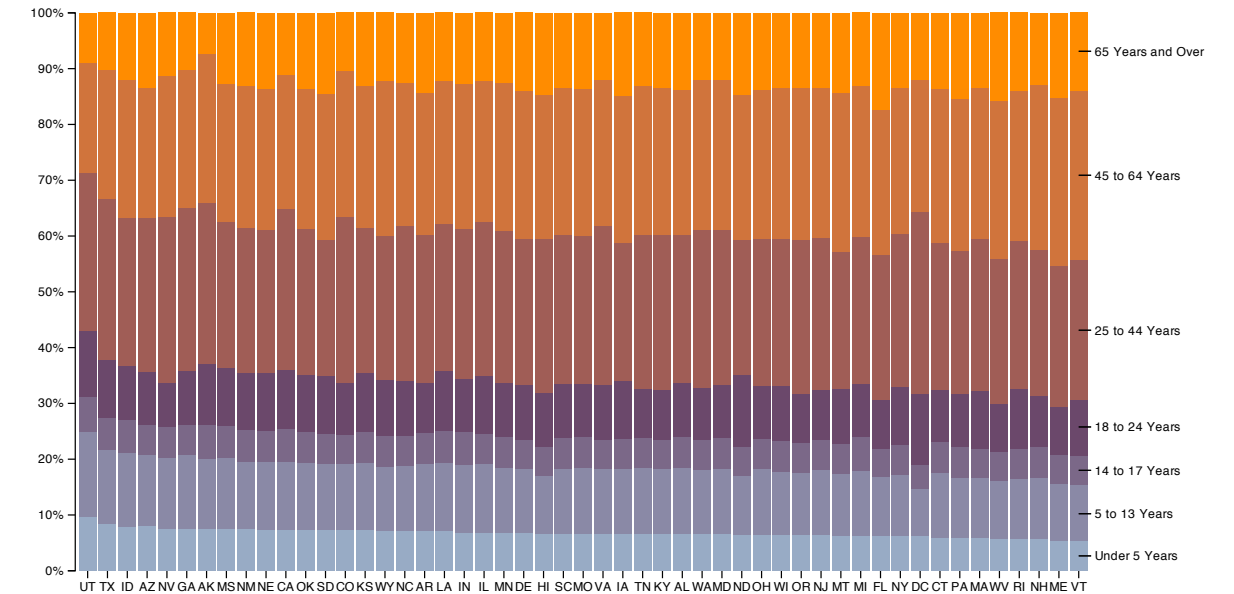
Idioms: pie chart, polar area chart

- pie chart
 - area marks with angle channel
 - accuracy: angle/area less accurate than line length
 - arclength also less accurate than line length
- polar area chart
 - area marks with length channel
 - more direct analog to bar charts
- data
 - I categ key attrib, I quant value attrib
- task
 - part-to-whole judgements



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/3887235>,

<http://bl.ocks.org/mbostock/3886208>,

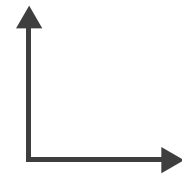
<http://bl.ocks.org/mbostock/3886394>.

Idiom: **glyphmaps**

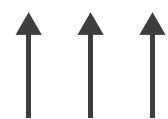
- rectilinear good for linear vs nonlinear trends
- radial good for cyclic patterns

➔ Axis Orientation

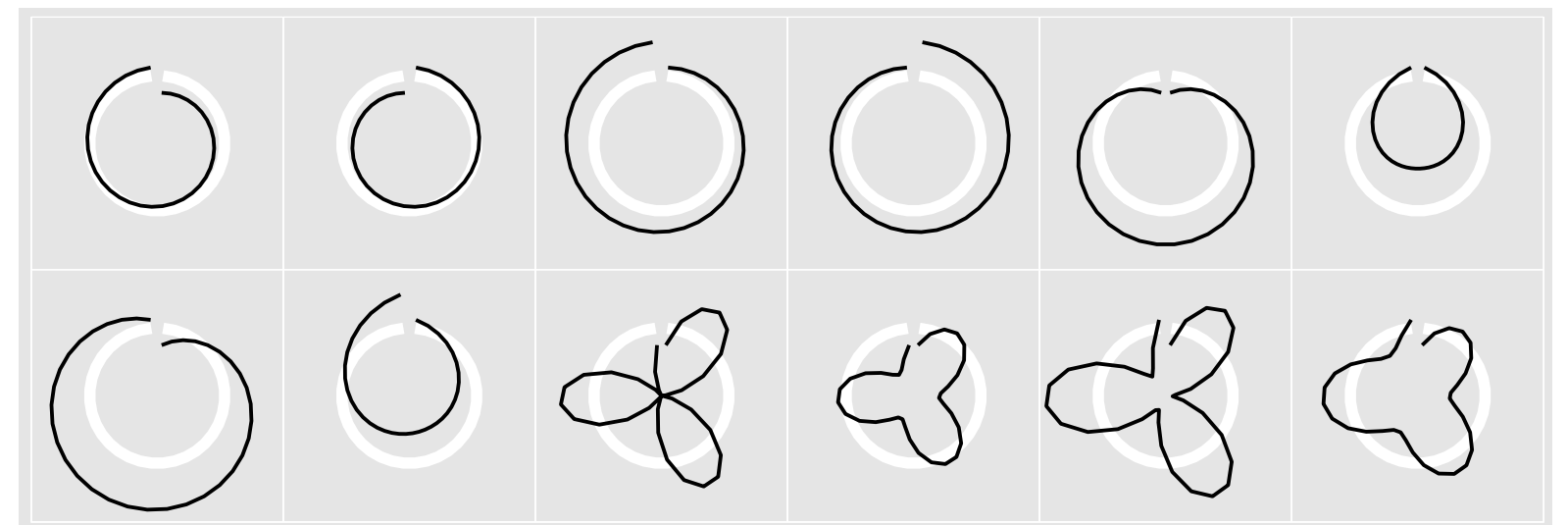
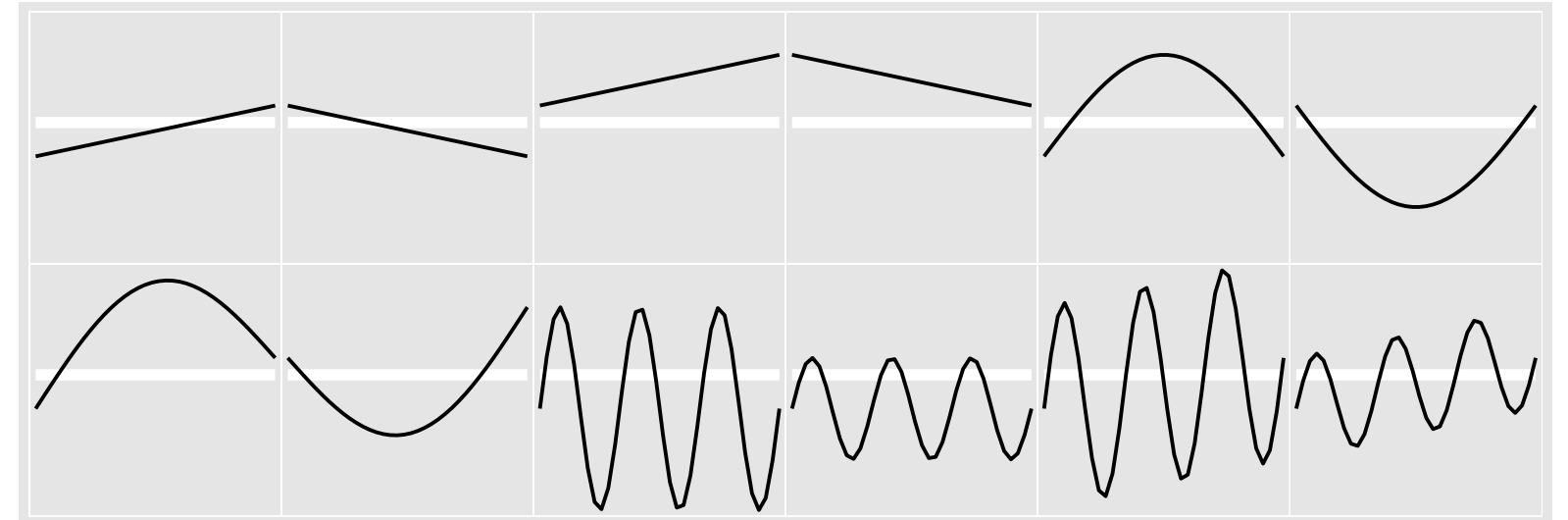
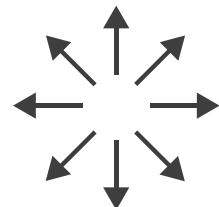
➔ Rectilinear



➔ Parallel



➔ Radial



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

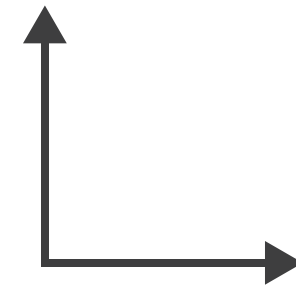
Orientation limitations

- **rectilinear: scalability wrt #axes**
 - 2 axes best
 - 3 problematic
 - more in afternoon
 - 4+ impossible
- **parallel: unfamiliarity, training time**
- **radial: perceptual limits**
 - angles lower precision than lengths
 - asymmetry between angle and length
 - can be exploited!

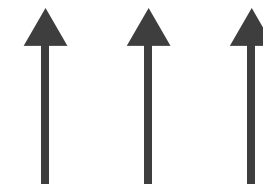
[Uncovering Strengths and Weaknesses of Radial Visualizations - an Empirical Approach. Diehl, Beck and Burch. IEEE TVCG (Proc. InfoVis) 16(6):935–942, 2010.]

➔ **Axis Orientation**

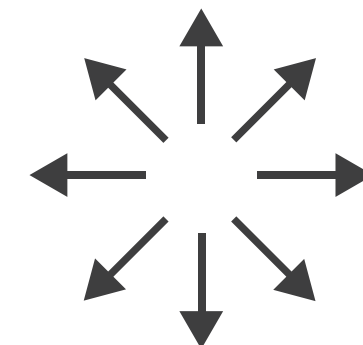
➔ Rectilinear



➔ Parallel

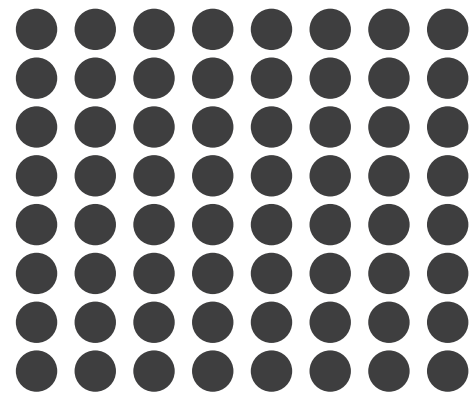


➔ Radial

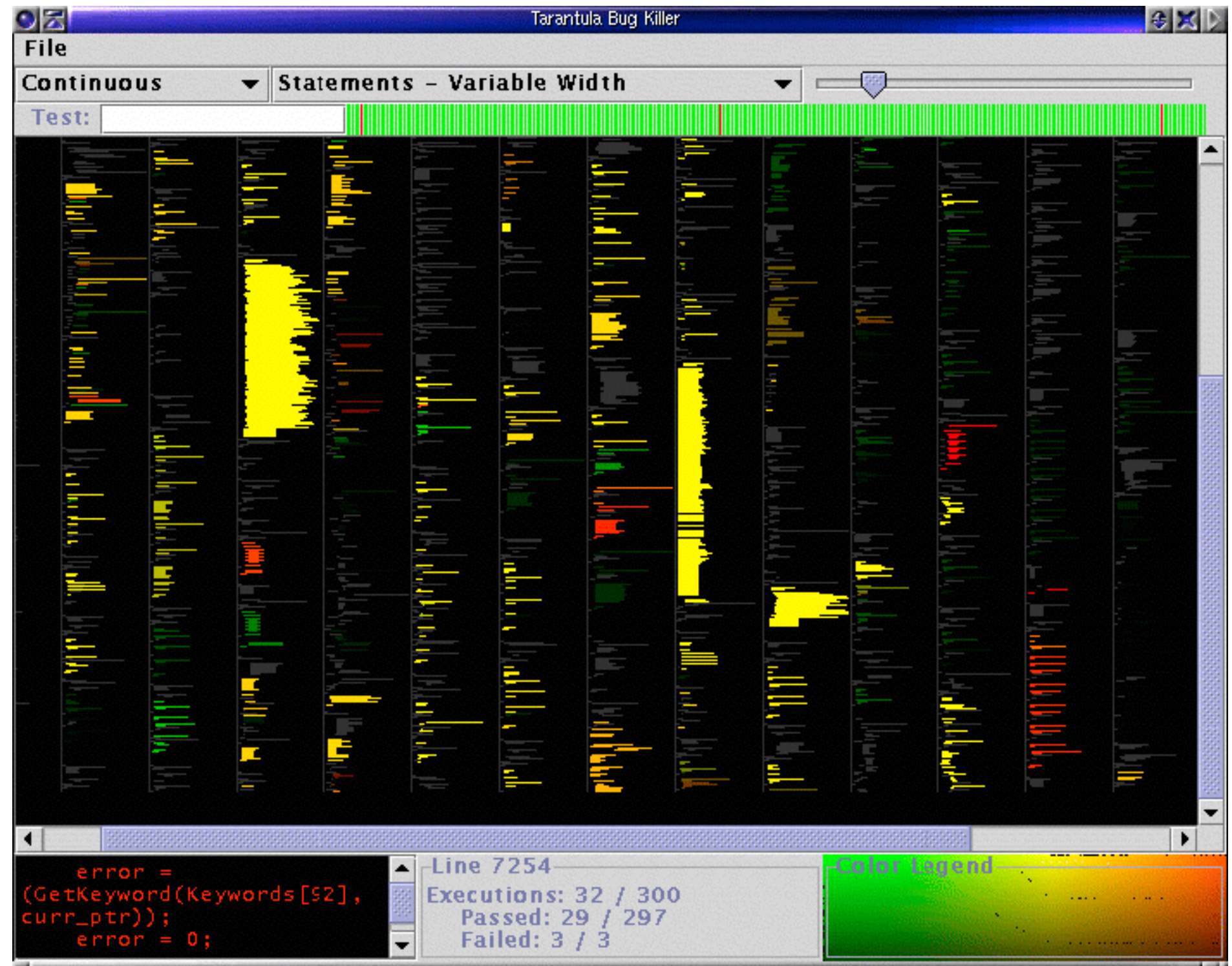


→ Layout Density

→ Dense



dense software overviews



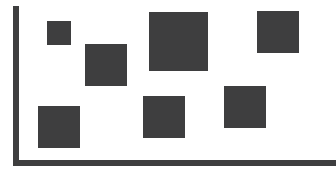
Arrange tables

→ Express Values

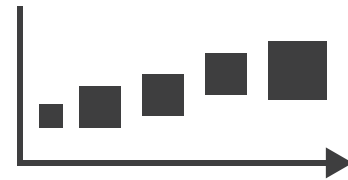


→ Separate, Order, Align Regions

→ Separate



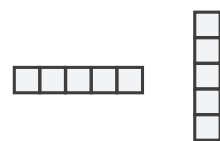
→ Order



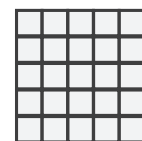
→ Align



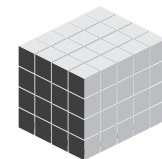
→ 1 Key *List*



→ 2 Keys *Matrix*



→ 3 Keys *Volume*

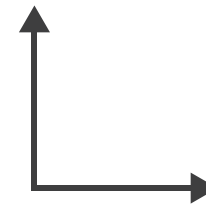


→ Many Keys *Recursive Subdivision*

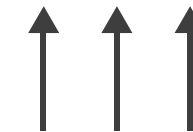


→ Axis Orientation

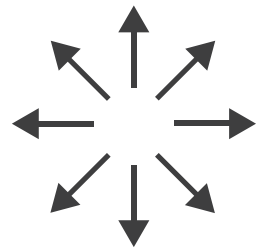
→ Rectilinear



→ Parallel

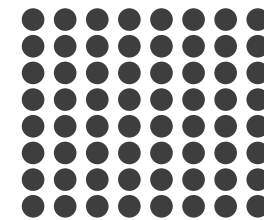


→ Radial



→ Layout Density

→ Dense



→ Space-Filling



Further reading

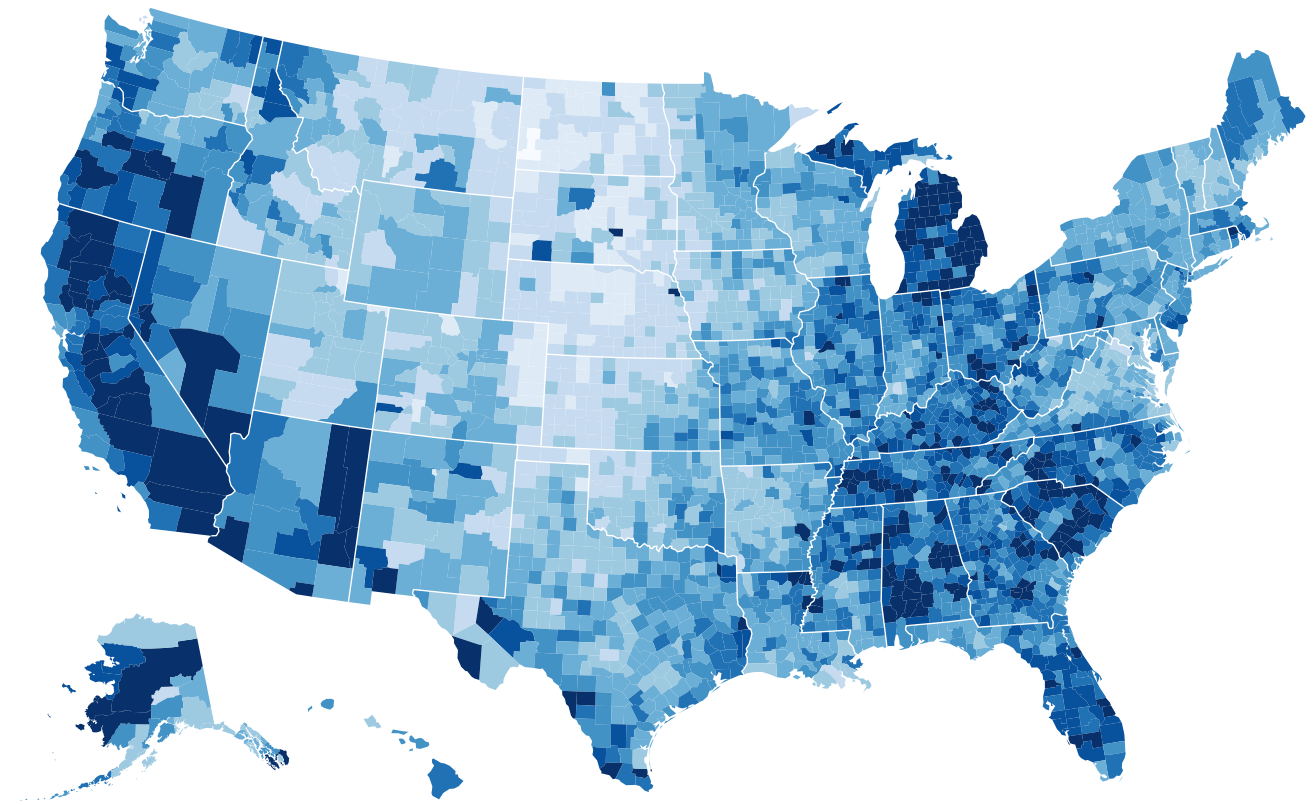
- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
–*Chap 7: Arrange Tables*
- Visualizing Data. Cleveland. Hobart Press, 1993.
- *A Brief History of Data Visualization*. Friendly. 2008.
<http://www.datavis.ca/milestones>

Outline

- **Session 1** / 10-11:30am
Data Visualization Pitfalls to Avoid
 - Introduction
 - Color
 - Space: 2D vs 3D
- **Session 2** / 12:30-3pm
Visualization Analysis & Design, In More Depth
 - Marks and Channels, Perception
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

Idiom: **choropleth map**

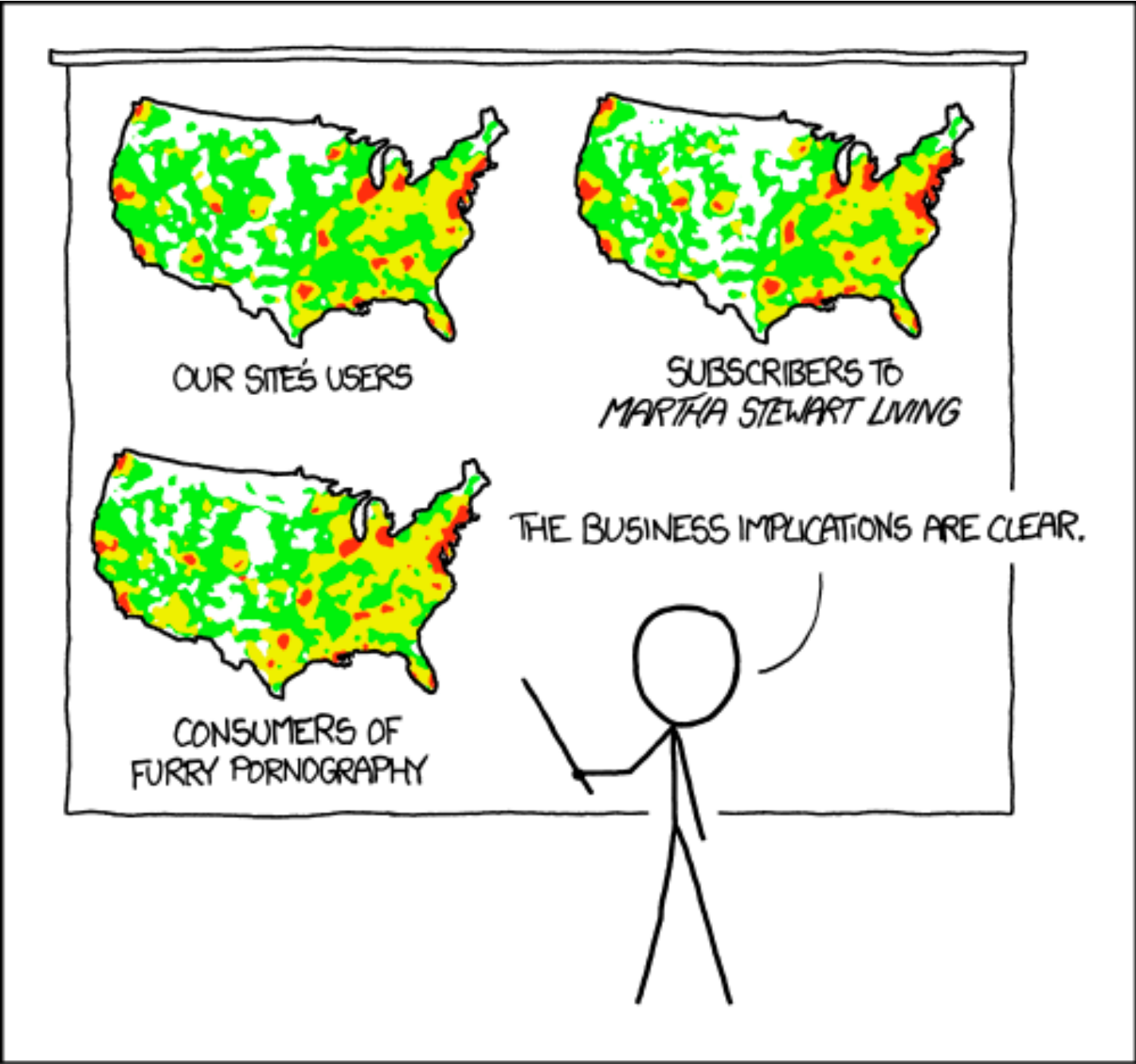
- *use* given spatial data
 - when central task is understanding spatial relationships
- data
 - geographic geometry
 - table with 1 quant attribute per region
- encoding
 - use given geometry for area mark boundaries
 - sequential segmented colormap *[more later]*



<http://bl.ocks.org/mbostock/4060606>

Beware: Population maps trickiness!

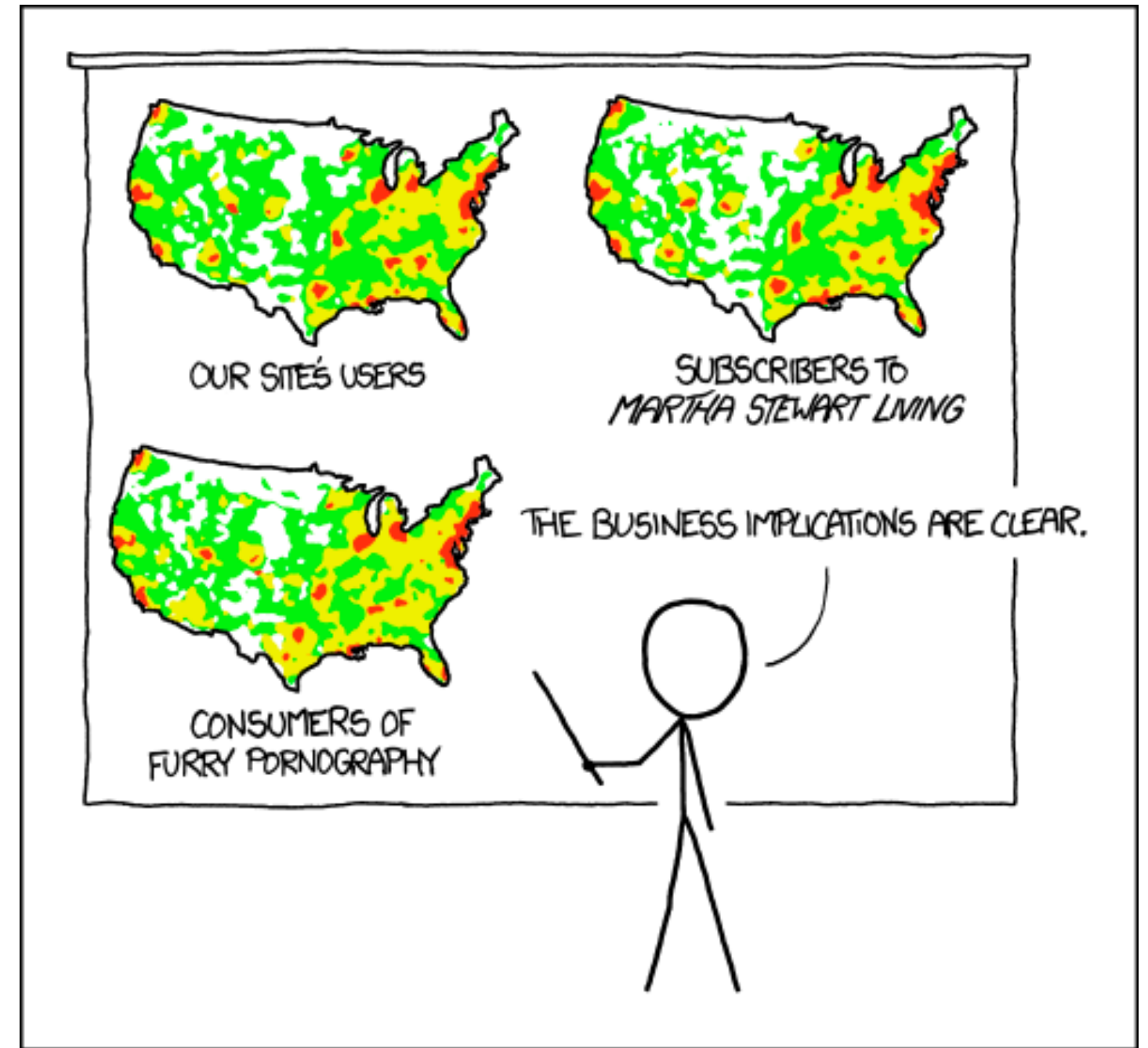
[<https://xkcd.com/1138>]



PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE
BASICALLY JUST POPULATION MAPS

Population maps trickiness

- beware!
 - absolute vs relative again
 - population density vs per capita
 - investigate with Ben Jones Tableau Public demo
 - <http://public.tableau.com/profile/ben.jones#!/vizhome/PopVsFin/PopVsFin>
- Are Maps of Financial Variables just Population Maps?*
- yes, unless you look at per capita (relative) numbers

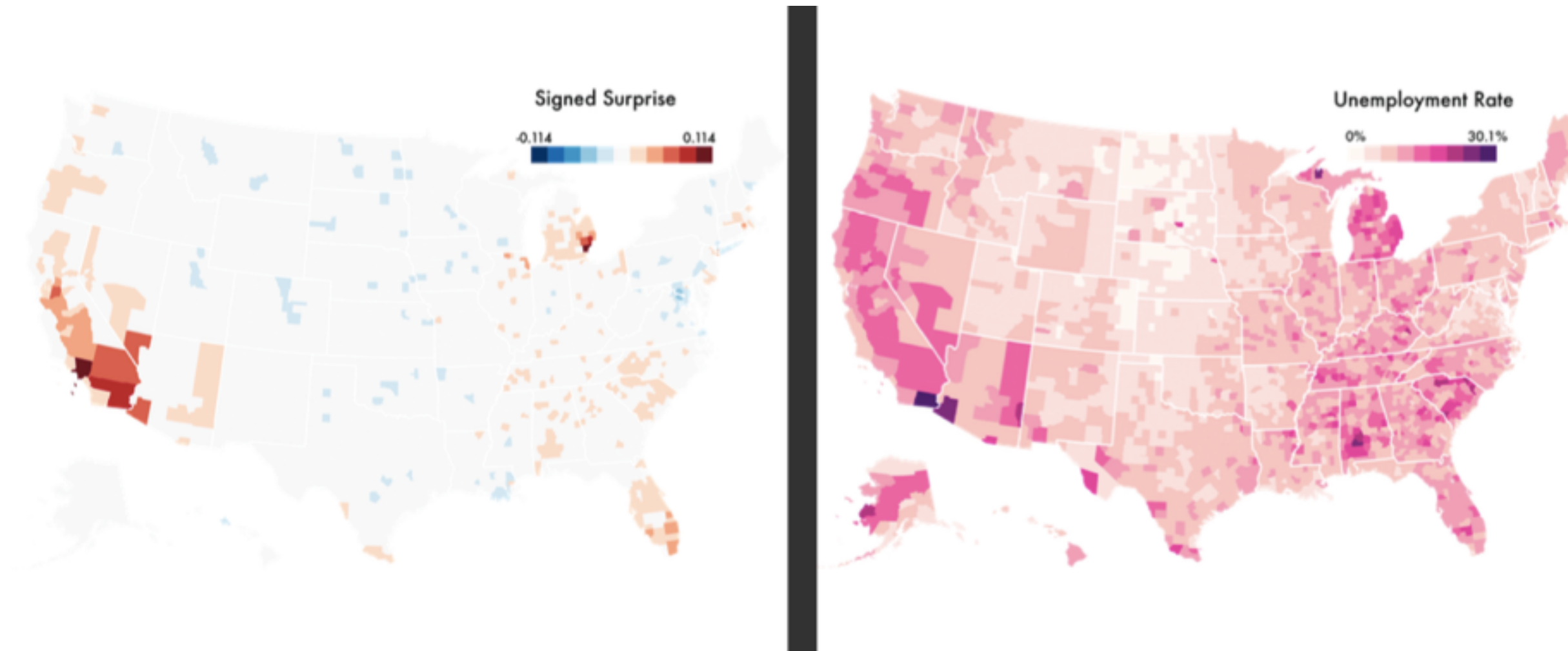


PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE
BASICALLY JUST POPULATION MAPS

[<https://xkcd.com/1138>]

Idiom: Bayesian surprise maps

- use models of expectations to highlight surprising values
- confounds (population) and variance (sparsity)



<https://medium.com/@uwdata/surprise-maps-showing-the-unexpected-e92b67398865>

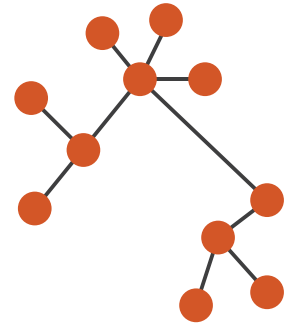
Outline

- **Session 1** / 10-11:30am
Data Visualization Pitfalls to Avoid
 - Introduction
 - Color
 - Space: 2D vs 3D
- **Session 2** / 12:30-3pm
Visualization Analysis & Design, In More Depth
 - Marks and Channels, Perception
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

Arrange networks and trees

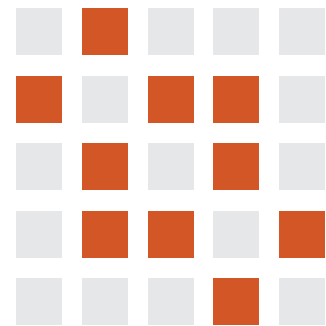
→ Node–Link Diagrams Connection Marks

✓ NETWORKS ✓ TREES



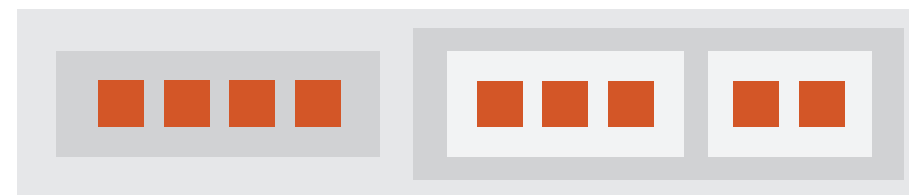
→ Adjacency Matrix Derived Table

✓ NETWORKS ✓ TREES



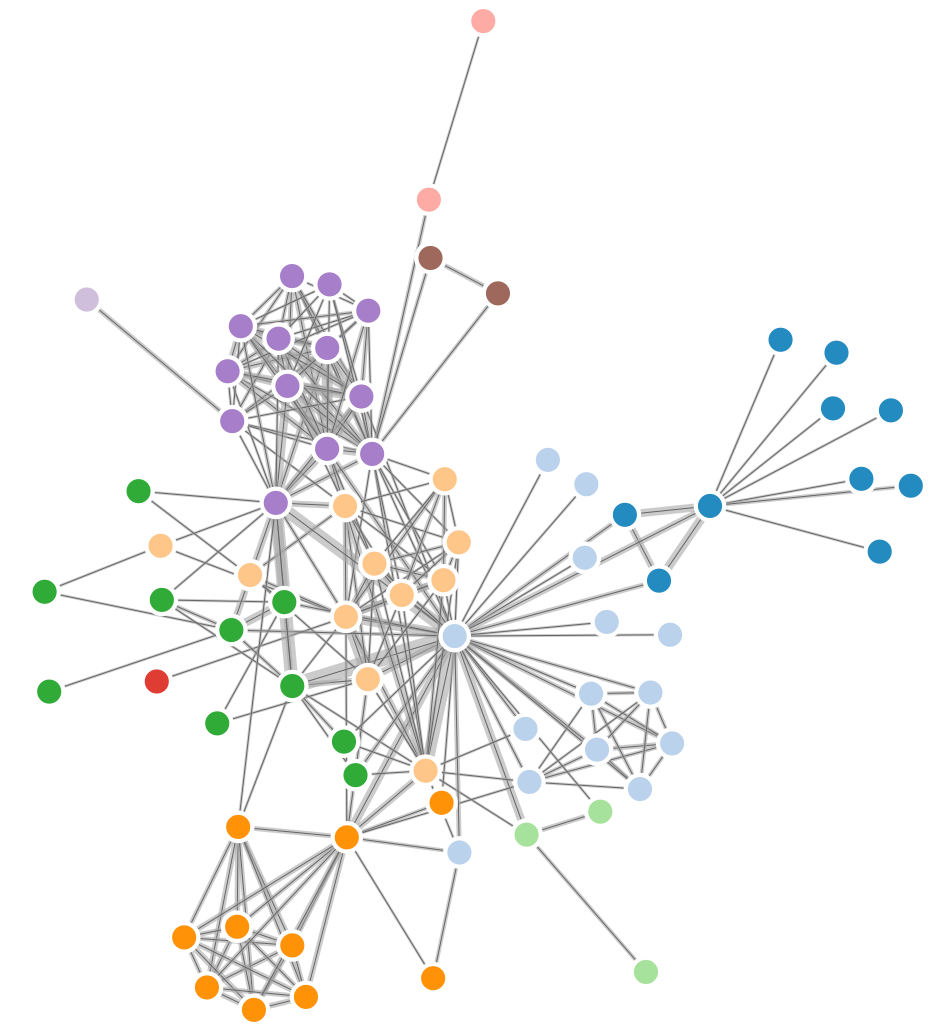
→ Enclosure Containment Marks

✗ NETWORKS ✓ TREES



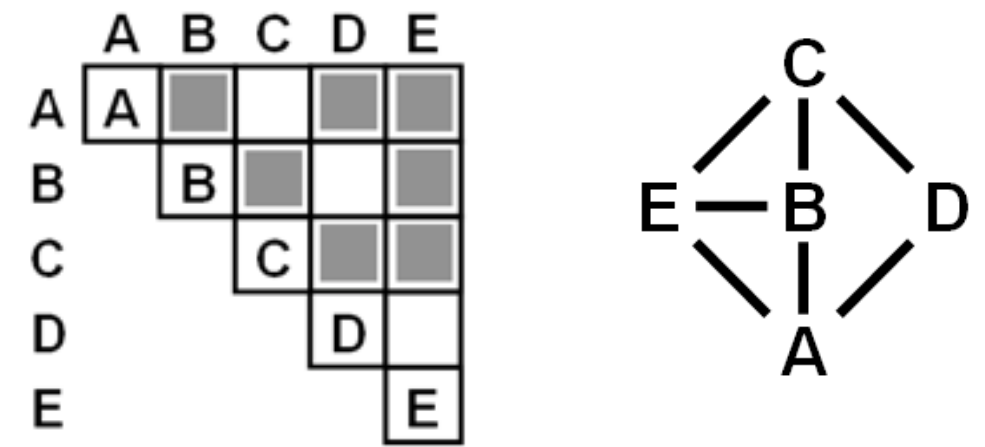
Idiom: **force-directed placement**

- visual encoding
 - link connection marks, node point marks
- considerations
 - spatial position: no meaning directly encoded
 - left free to minimize crossings
 - proximity semantics?
 - sometimes meaningful
 - sometimes arbitrary, artifact of layout algorithm
 - tension with length
 - long edges more visually salient than short
- tasks
 - explore topology; locate paths, clusters
- scalability
 - node/edge density $E < 4N$



Idiom: adjacency matrix view

- data: network
 - transform into same data/encoding as heatmap
- derived data: table from network
 - 1 quant attrib
 - weighted edge between nodes
 - 2 categ attribs: node list x 2
- visual encoding
 - cell shows presence/absence of edge
- scalability
 - 1K nodes, 1M edges



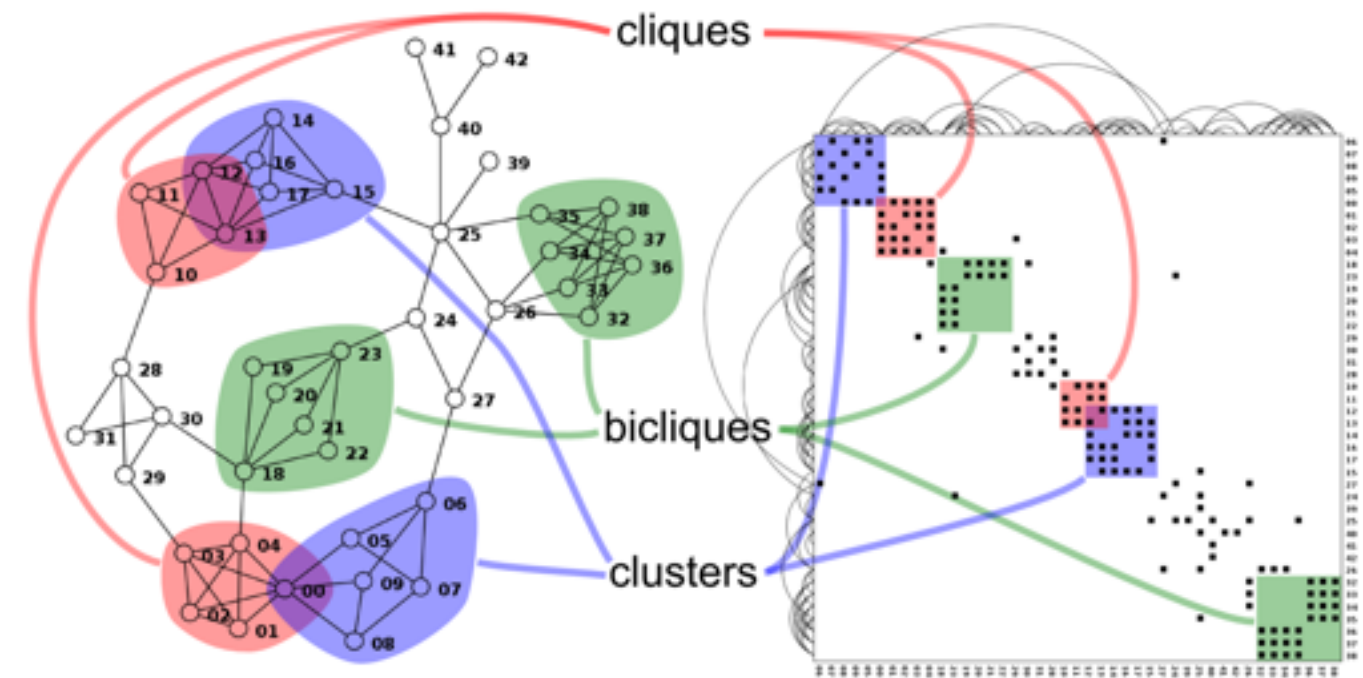
[NodeTrix: a Hybrid Visualization of Social Networks. Henry, Fekete, and McGuffin. IEEE TVCG (Proc. InfoVis) 13(6):1302-1309, 2007.]



[Points of view: Networks. Gehlenborg and Wong. Nature Methods 9:115.]

Connection vs. adjacency comparison

- adjacency matrix strengths
 - predictability, scalability, supports reordering
 - some topology tasks trainable
- node-link diagram strengths
 - topology understanding, path tracing
 - intuitive, no training needed
- empirical study
 - node-link best for small networks
 - matrix best for large networks
 - if tasks don't involve topological structure!

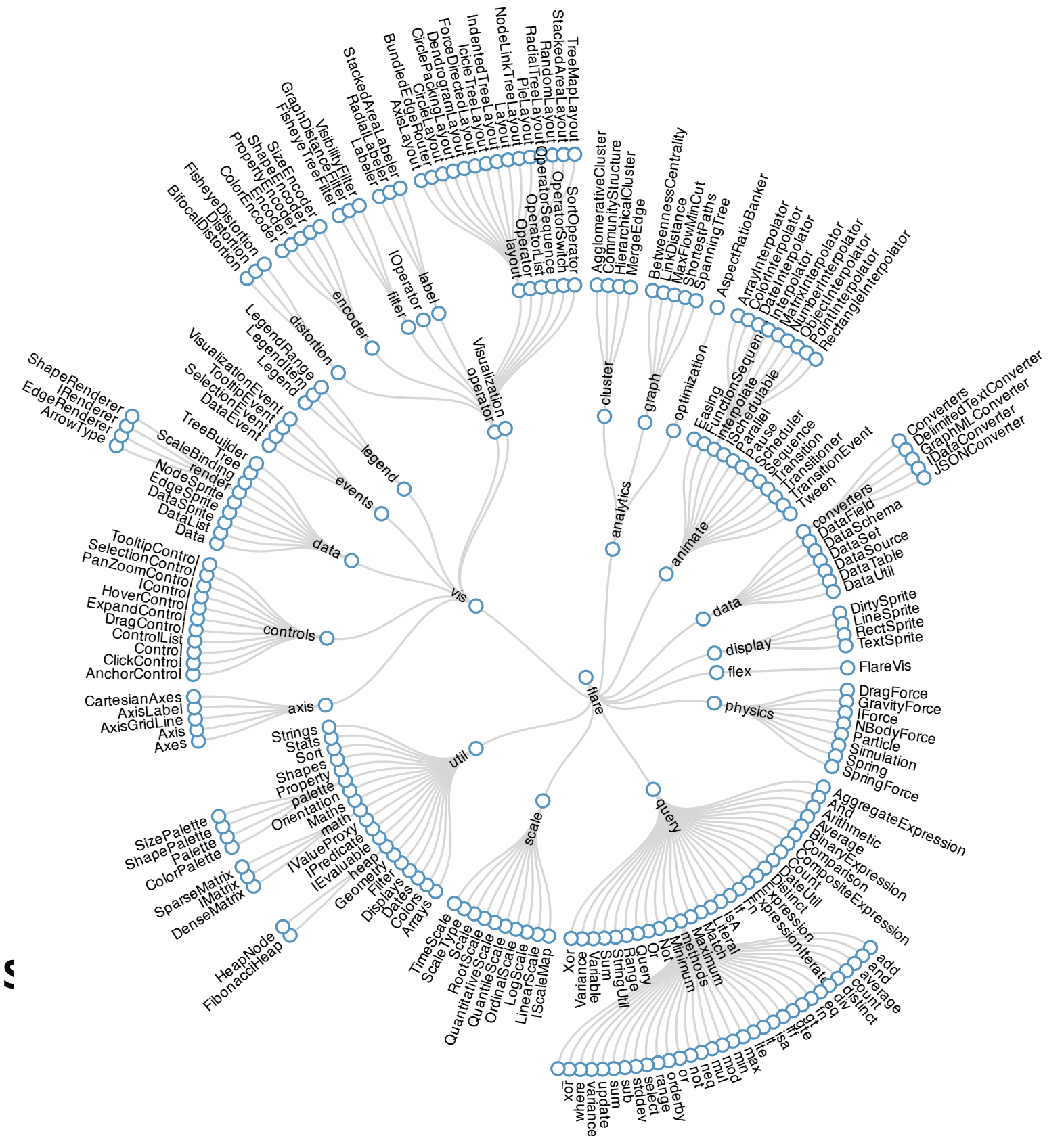


<http://www.michaelmcguffin.com/courses/vis/patternsInAdjacencyMatrix.png>

[On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. Ghoniem, Fekete, and Castagliola. Information Visualization 4:2 (2005), 114–135.]

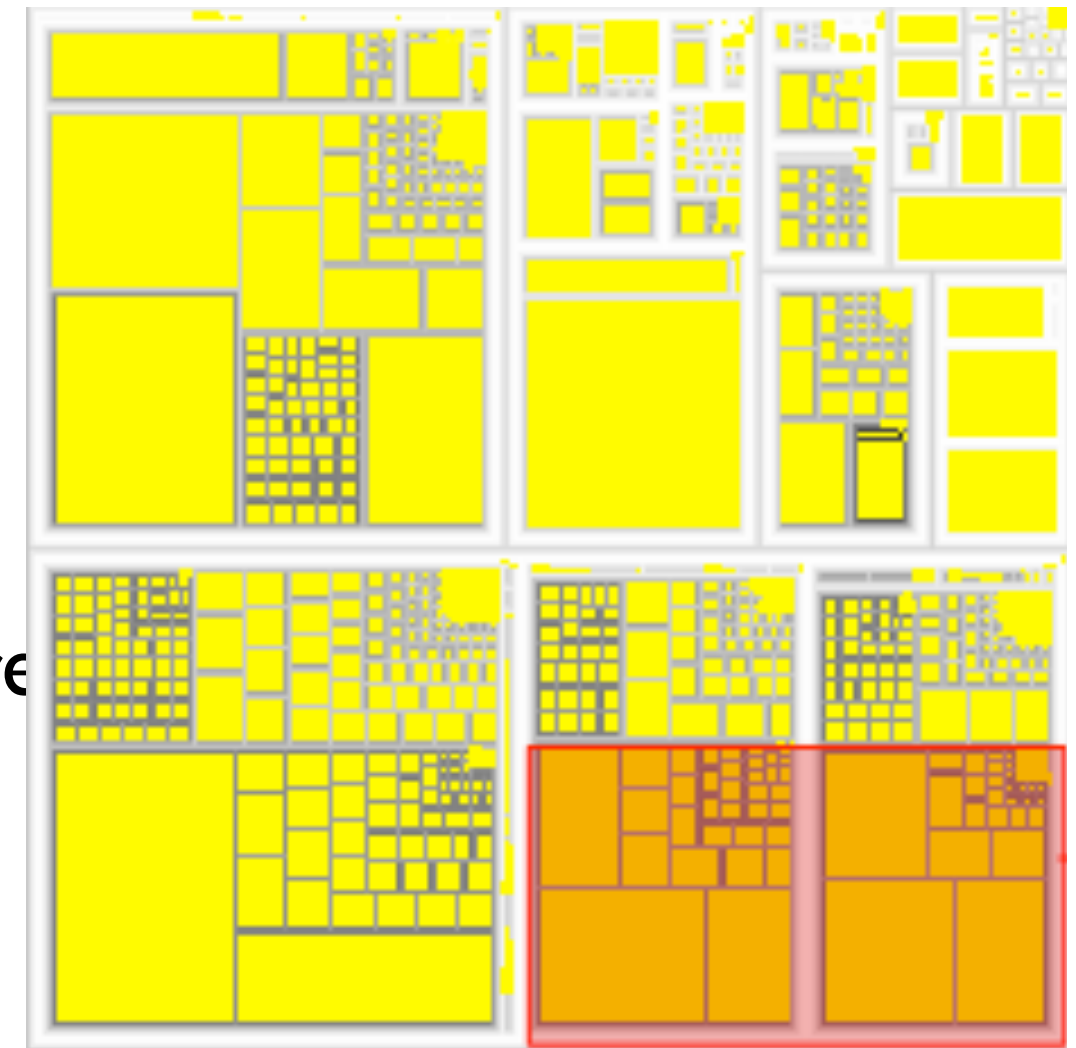
Idiom: radial node-link tree

- data
 - tree
- encoding
 - link connection marks
 - point node marks
 - radial axis orientation
 - angular proximity: siblings
 - distance from center: depth in tree
- tasks
 - understanding topology, following paths
- scalability
 - 1K - 10K nodes



Idiom: **treemap**

- data
 - tree
 - 1 quant attrib at leaf nodes
- encoding
 - area containment marks for hierarchical structure
 - rectilinear orientation
 - size encodes quant attrib
- tasks
 - query attribute at leaf nodes
- scalability
 - 1M leaf nodes

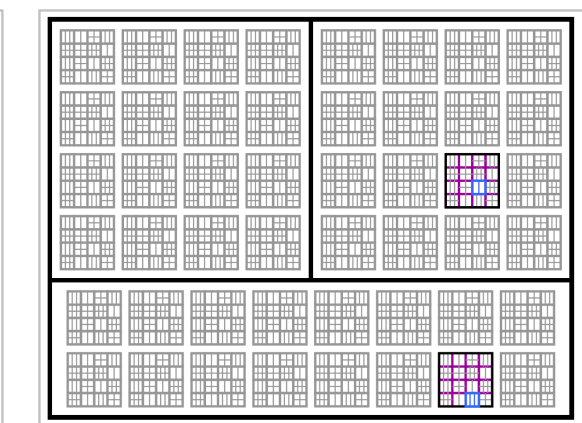
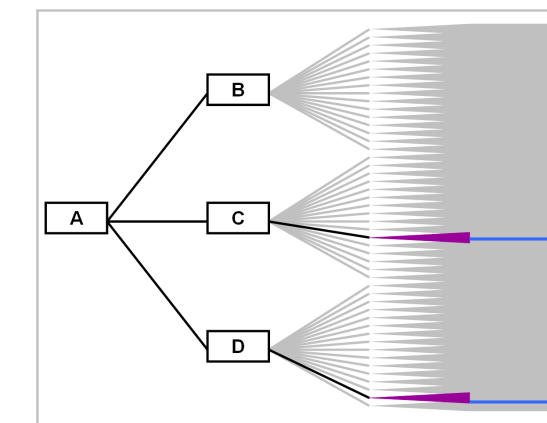
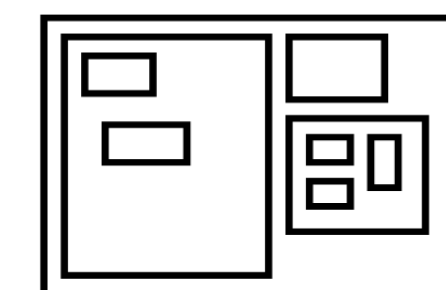
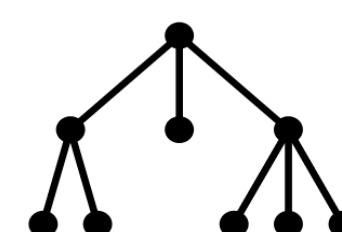
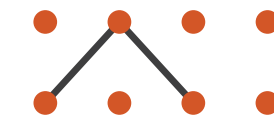
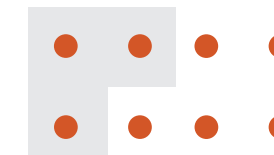


http://tulip.labri.fr/Documentation/3_7/userHandbook/html/ch06.html

Link marks: Connection and containment

- marks as links (vs. nodes)
 - common case in network drawing
 - 1D case: connection
 - ex: all node-link diagrams
 - emphasizes topology, path tracing
 - networks and trees
 - 2D case: containment
 - ex: all treemap variants
 - emphasizes attribute values at leaves (size coding)
 - only trees

➔ Containment ➔ Connection



Node-Link Diagram

Treemap

[Elastic Hierarchies: Combining Treemaps and Node-Link Diagrams. Dong, McGuffin, and Chignell. Proc. InfoVis 2005, p. 57-64.]

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
–*Chap 9: Arrange Networks and Trees*
- Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges. von Landesberger et al. Computer Graphics Forum 30:6 (2011), 1719–1749.
- Simple Algorithms for Network Visualization: A Tutorial. McGuffin. Tsinghua Science and Technology (Special Issue on Visualization and Computer Graphics) 17:4 (2012), 383–398.
- Drawing on Physical Analogies. Brandes. In Drawing Graphs: Methods and Models, LNCS Tutorial, 2025, edited by M. Kaufmann and D. Wagner, LNCS Tutorial, 2025, pp. 71–86. Springer-Verlag, 2001.
- <http://www.treevis.net> Treevis.net: A Tree Visualization Reference. Schulz. IEEE Computer Graphics and Applications 31:6 (2011), 11–15.
- Perceptual Guidelines for Creating Rectangular Treemaps. Kong, Heer, and Agrawala. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis) 16:6 (2010), 990–998.

Outline

- **Session 1** / 10-11:30am
Data Visualization Pitfalls to Avoid
 - Introduction
 - Color
 - Space: 2D vs 3D
- **Session 2** / 12:30-3pm
Visualization Analysis & Design, In More Depth
 - Marks and Channels, Perception
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

How?

Encode

→ Arrange

→ Express



→ Separate



→ Order



→ Align



→ Use



→ Map

from **categorical** and **ordered** attributes

→ Color

→ Hue



→ Saturation



→ Luminance



→ Size, Angle, Curvature, ...



→ Shape



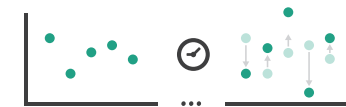
→ Motion

Direction, Rate, Frequency, ...

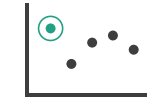


Manipulate

→ Change



→ Select



→ Navigate

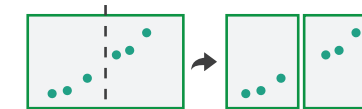


Facet

→ Juxtapose



→ Partition



→ Superimpose



Reduce

→ Filter



→ Aggregate



→ Embed



What?

Why?

How?

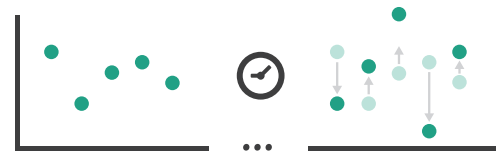
How to handle complexity: 1 previous strategy + 3 more

→ *Derive*

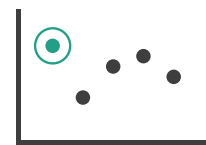


Manipulate

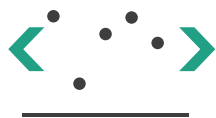
→ Change



→ Select

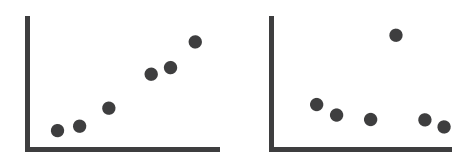


→ Navigate

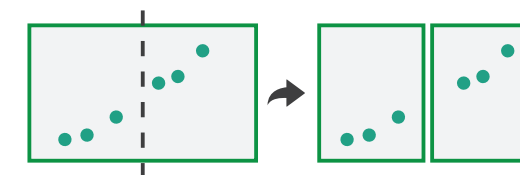


Facet

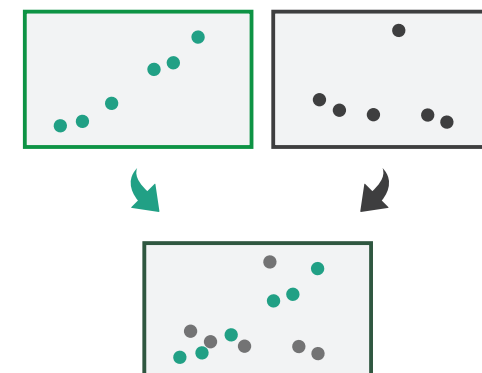
→ Juxtapose



→ Partition

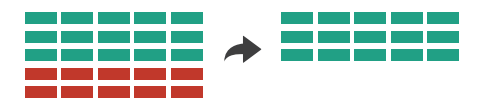


→ Superimpose



Reduce

→ Filter



→ Aggregate



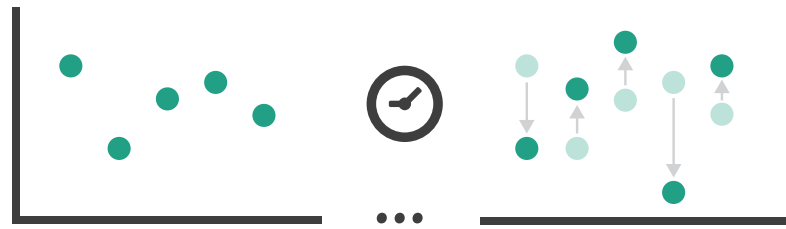
→ Embed



- derive new data to show within view
- change view over time
- facet across multiple views
- reduce items/attributes within single view

Manipulate

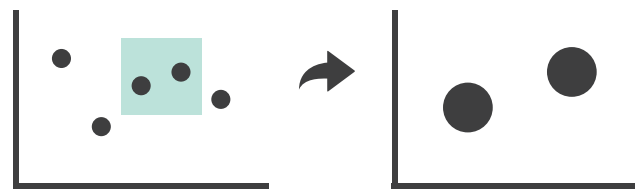
→ Change over Time



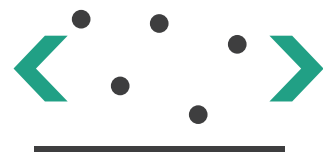
→ Navigate

→ Item Reduction

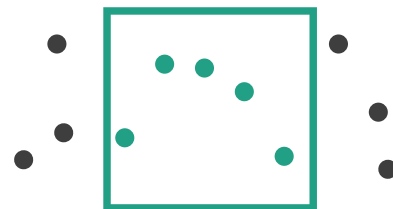
→ Zoom
Geometric or *Semantic*



→ Pan/Translate

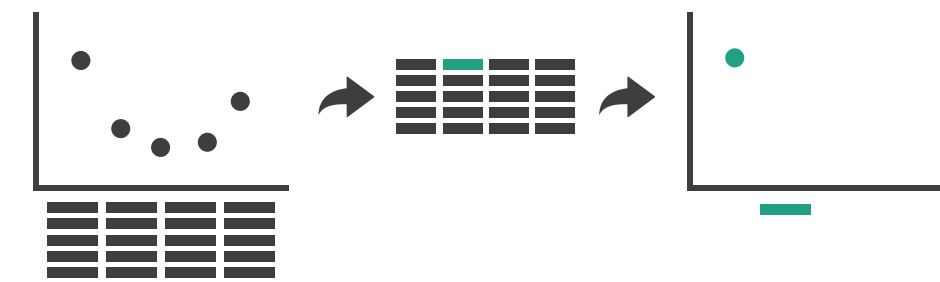


→ Constrained

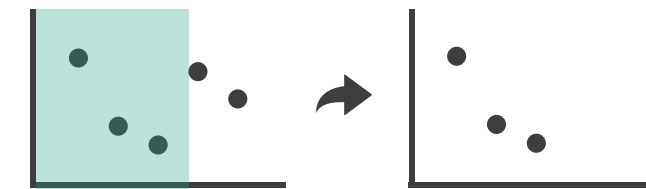


→ Attribute Reduction

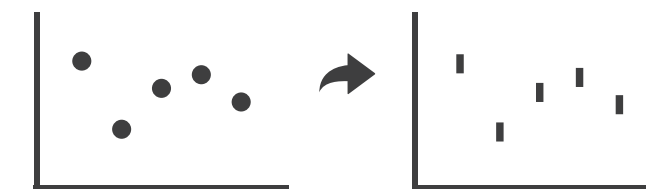
→ Slice



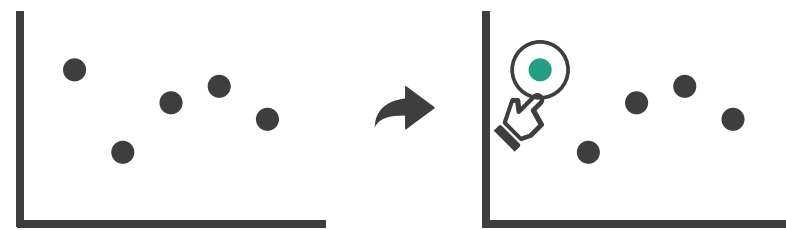
→ Cut



→ Project



→ Select



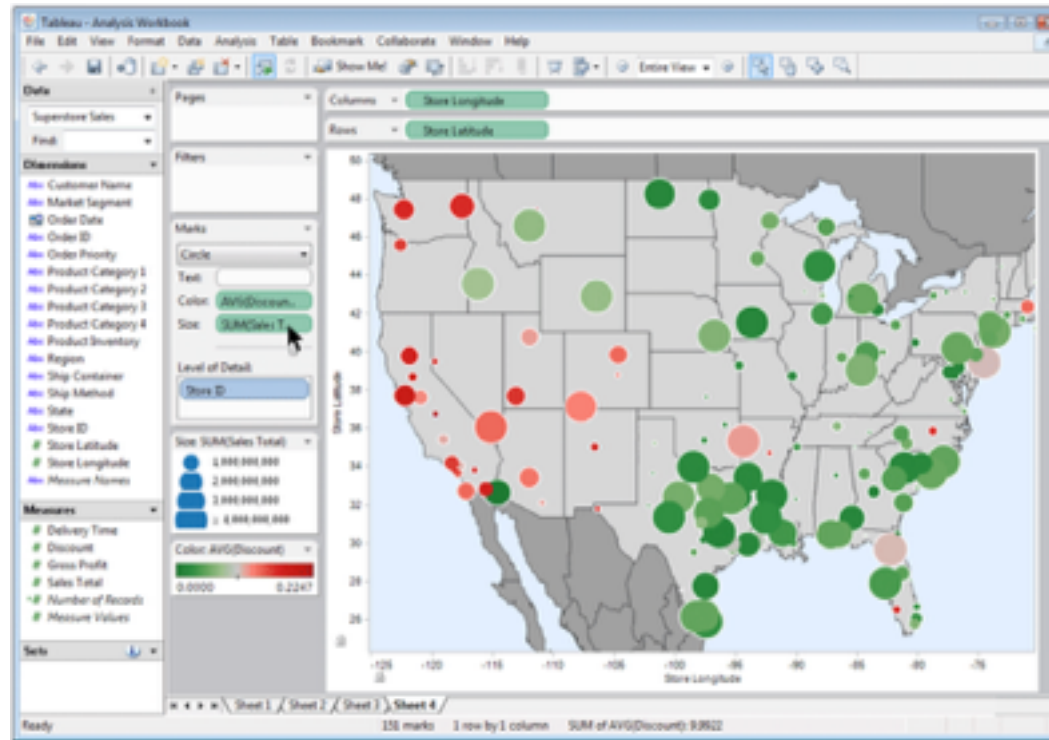
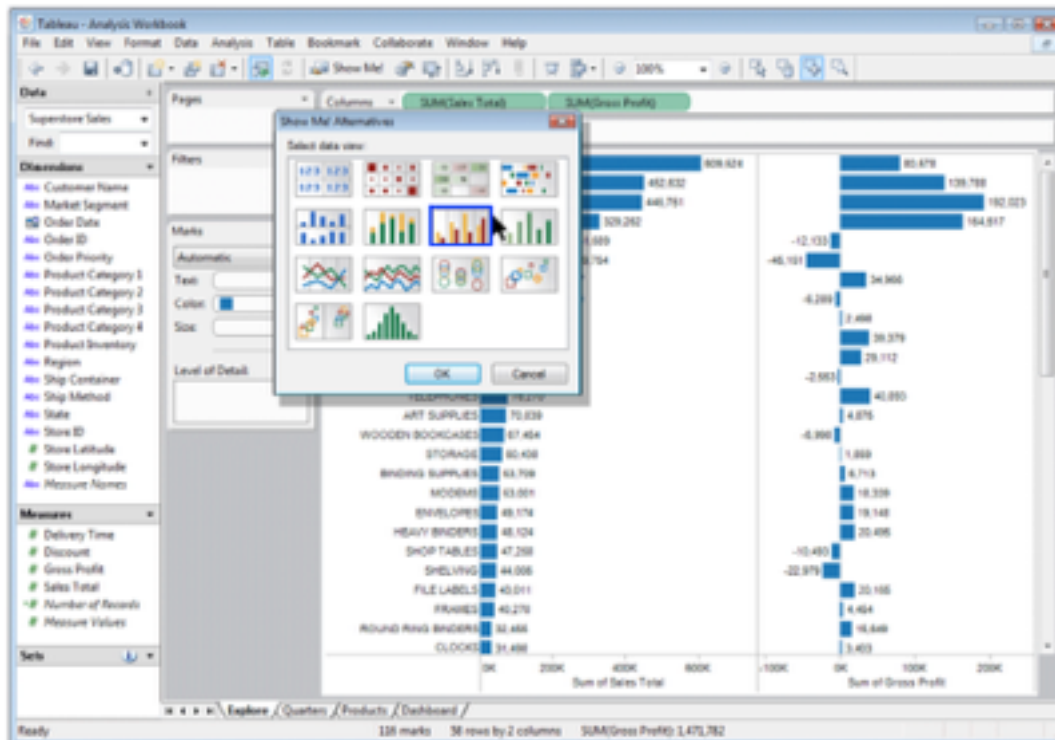
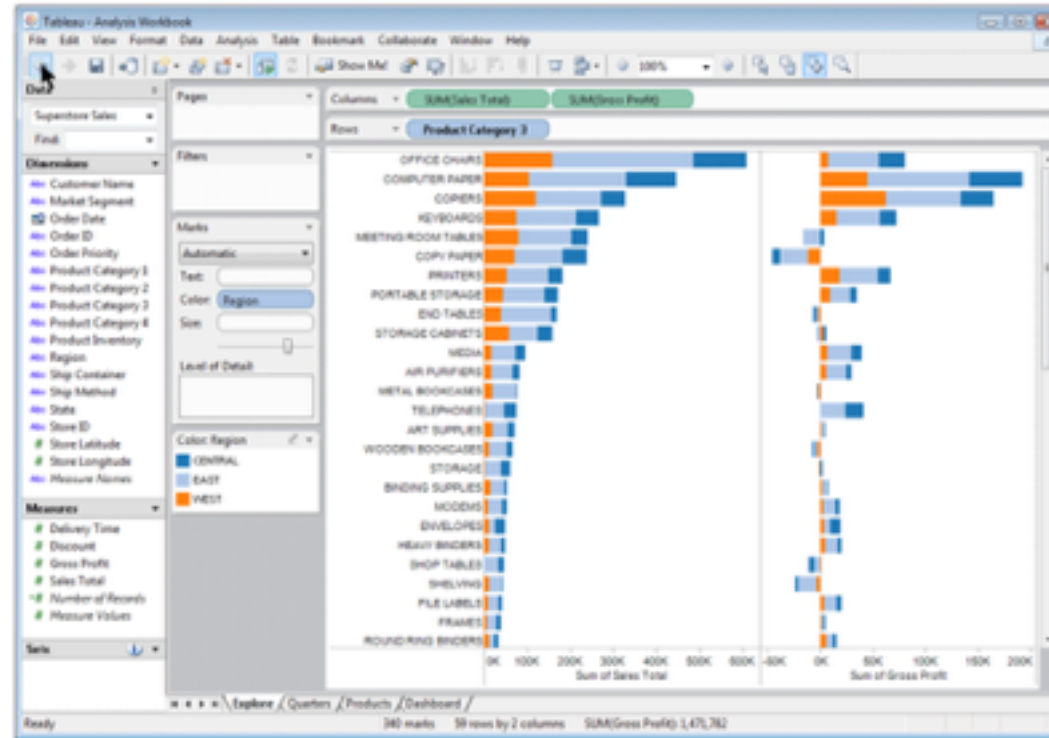
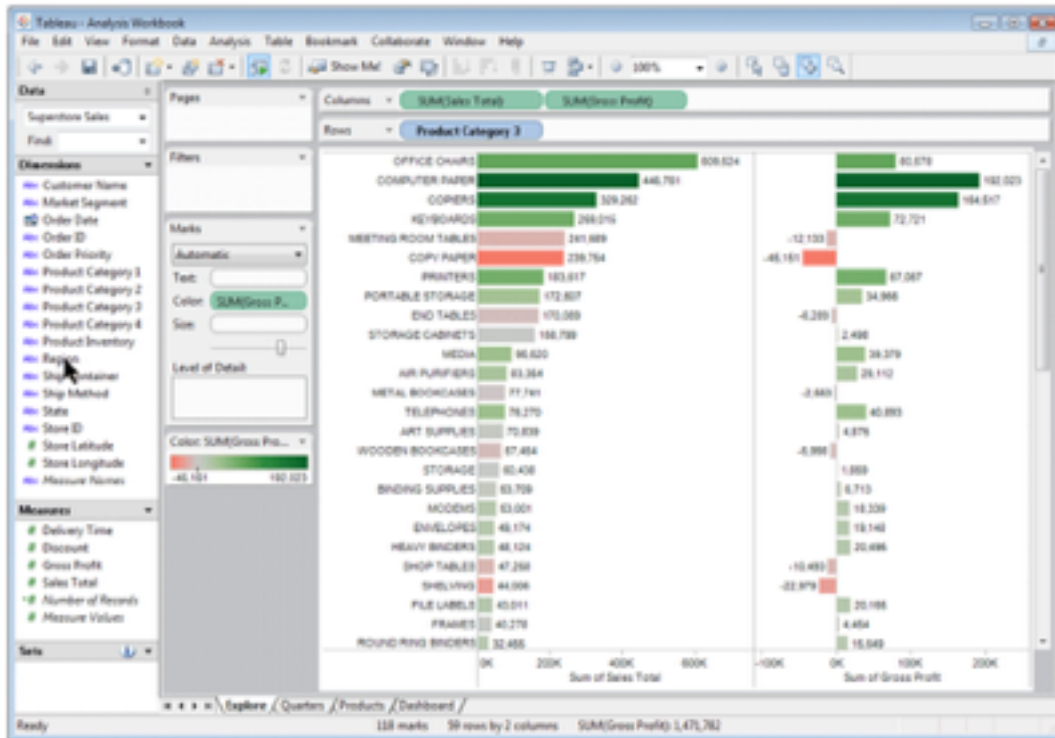
Change over time

- change any of the other choices
 - encoding itself
 - parameters
 - arrange: rearrange, reorder
 - aggregation level, what is filtered...

 - interaction entails change

Idiom: Re-encode

System: Tableau

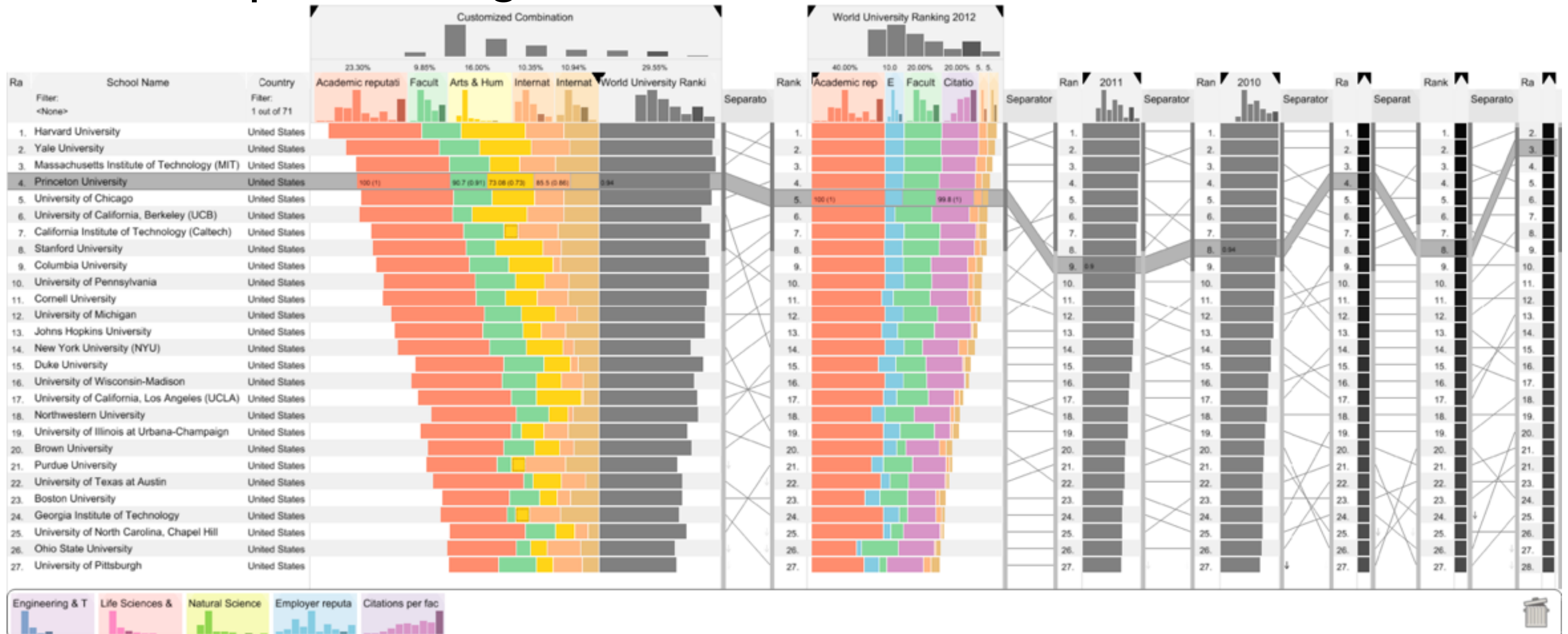


made using Tableau, <http://tableausoftware.com>

Idiom: Reorder

System: LineUp

- data: tables with many attributes
- task: compare rankings

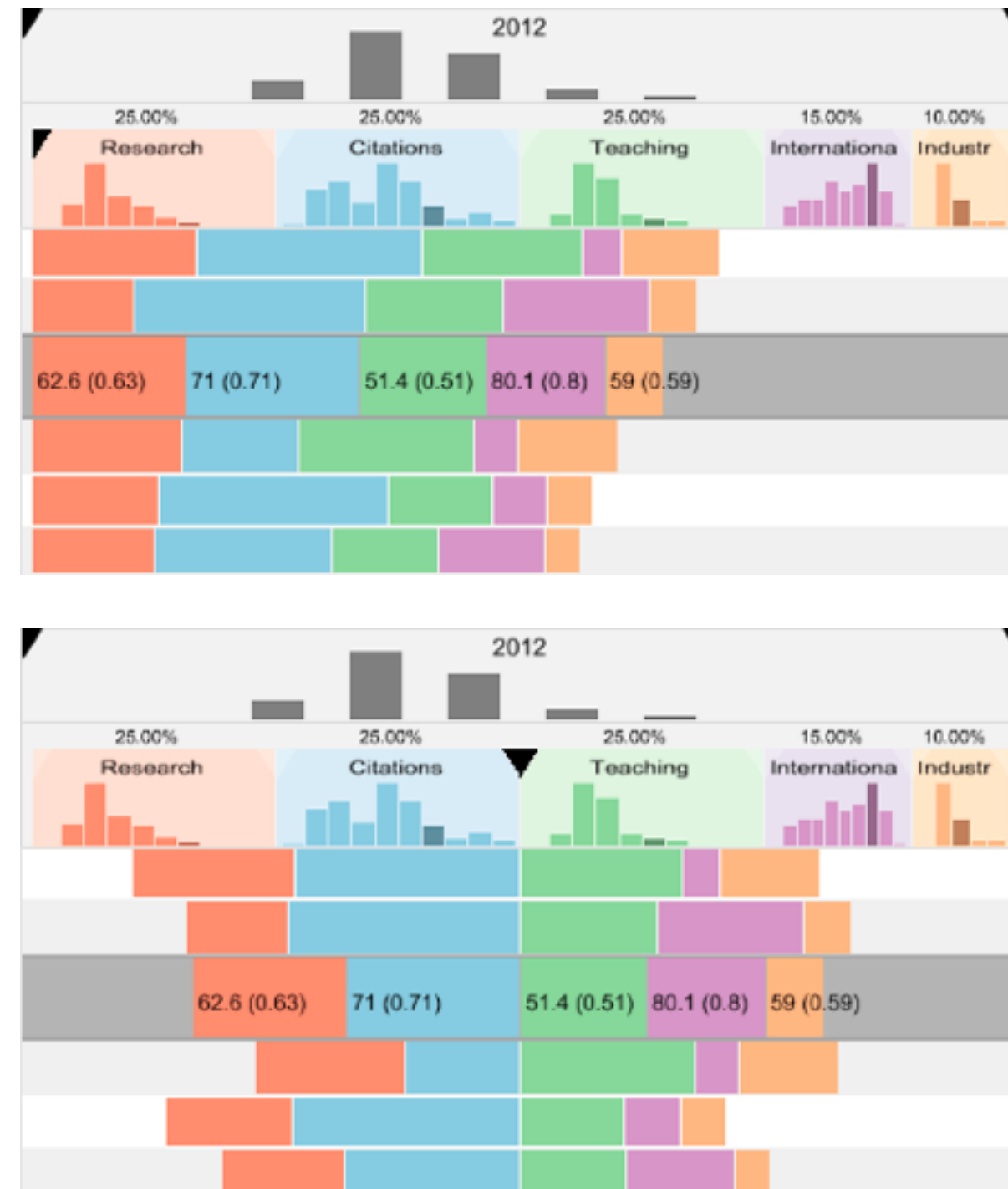


[LineUp: Visual Analysis of Multi-Attribute Rankings. Gratzl, Lex, Gehlenborg, Pfister, and Streit. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2013) 19:12 (2013), 2277–2286.]

Idiom: **Realign**

- stacked bars
 - easy to compare
 - first segment
 - total bar
- align to different segment
 - supports flexible comparison

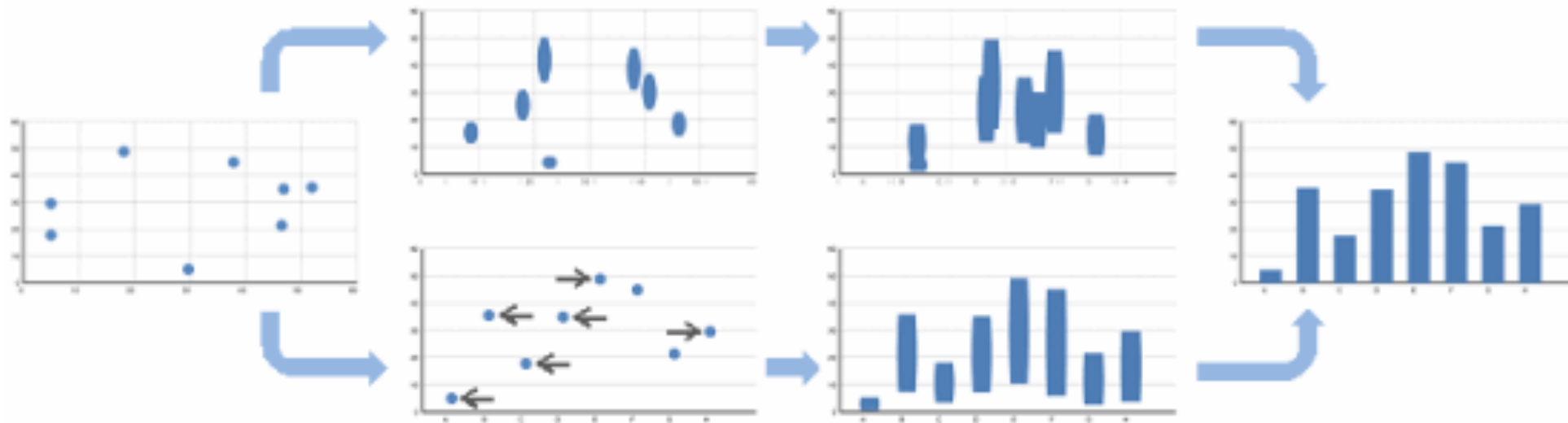
System: **LineUp**



[LineUp: Visual Analysis of Multi-Attribute Rankings. Gratzl, Lex, Gehlenborg, Pfister, and Streit. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2013) 19:12 (2013), 2277–2286.]

Idiom: **Animated transitions**

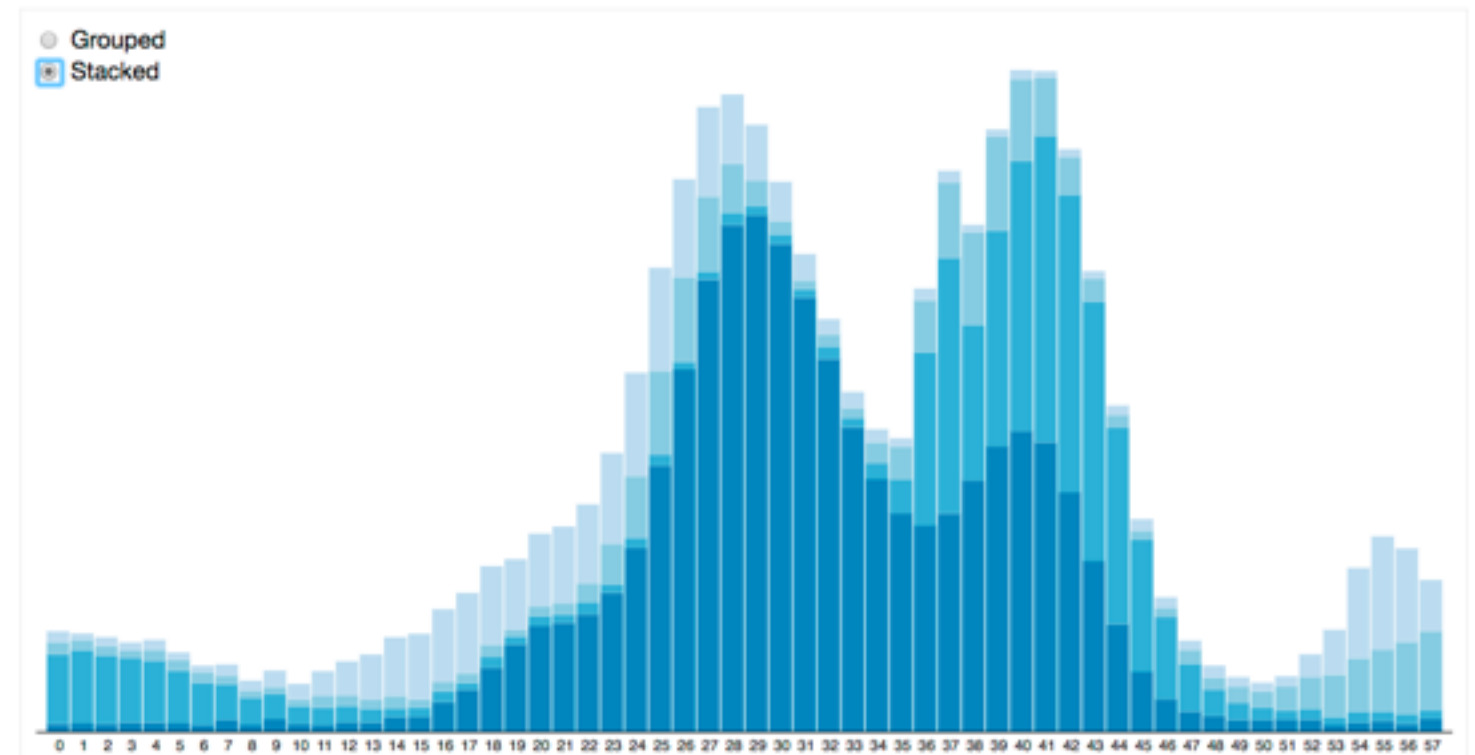
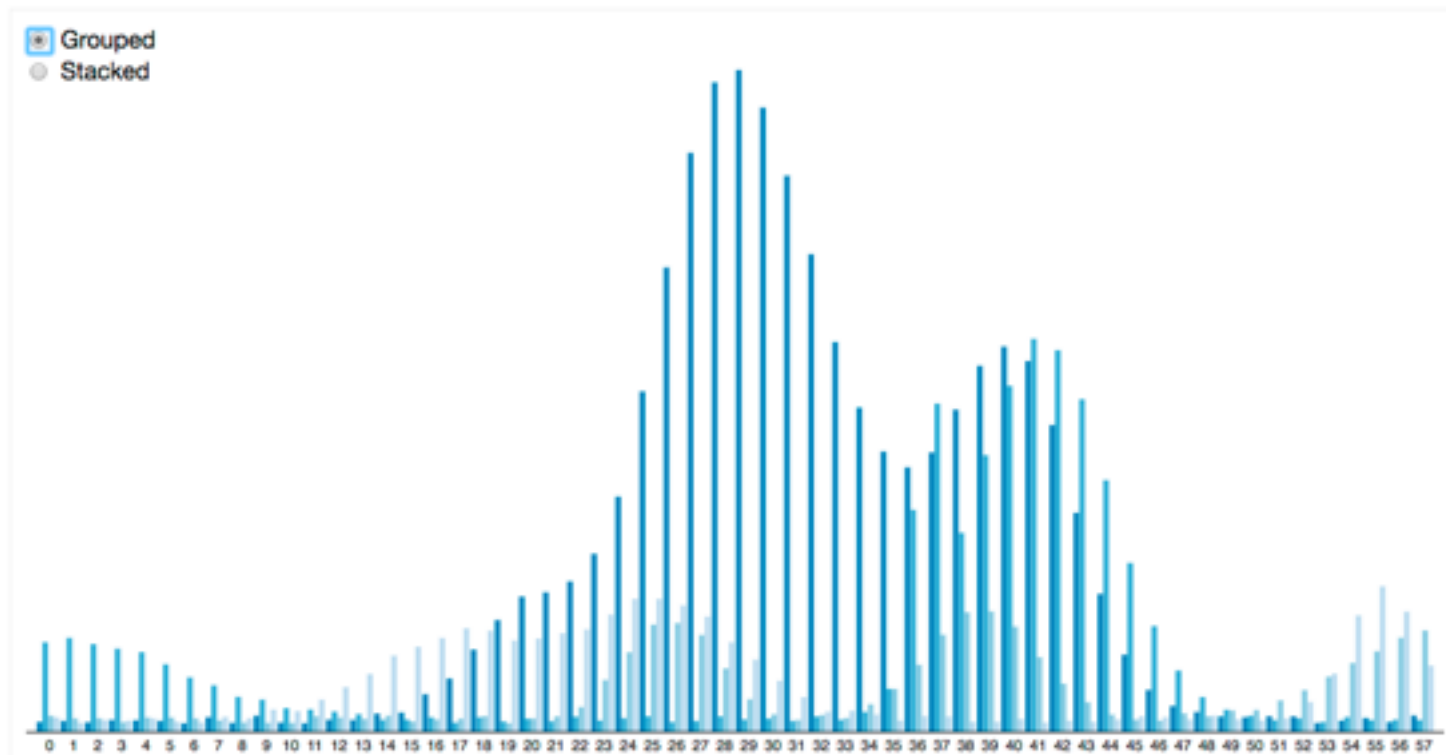
- smooth interpolation from one state to another
 - alternative to jump cuts, supports item tracking
 - best case for animation
 - staging to reduce cognitive load
- example: animated transitions in statistical data graphics



video: vimeo.com/19278444

Idiom: **Animated transitions - visual encoding change**

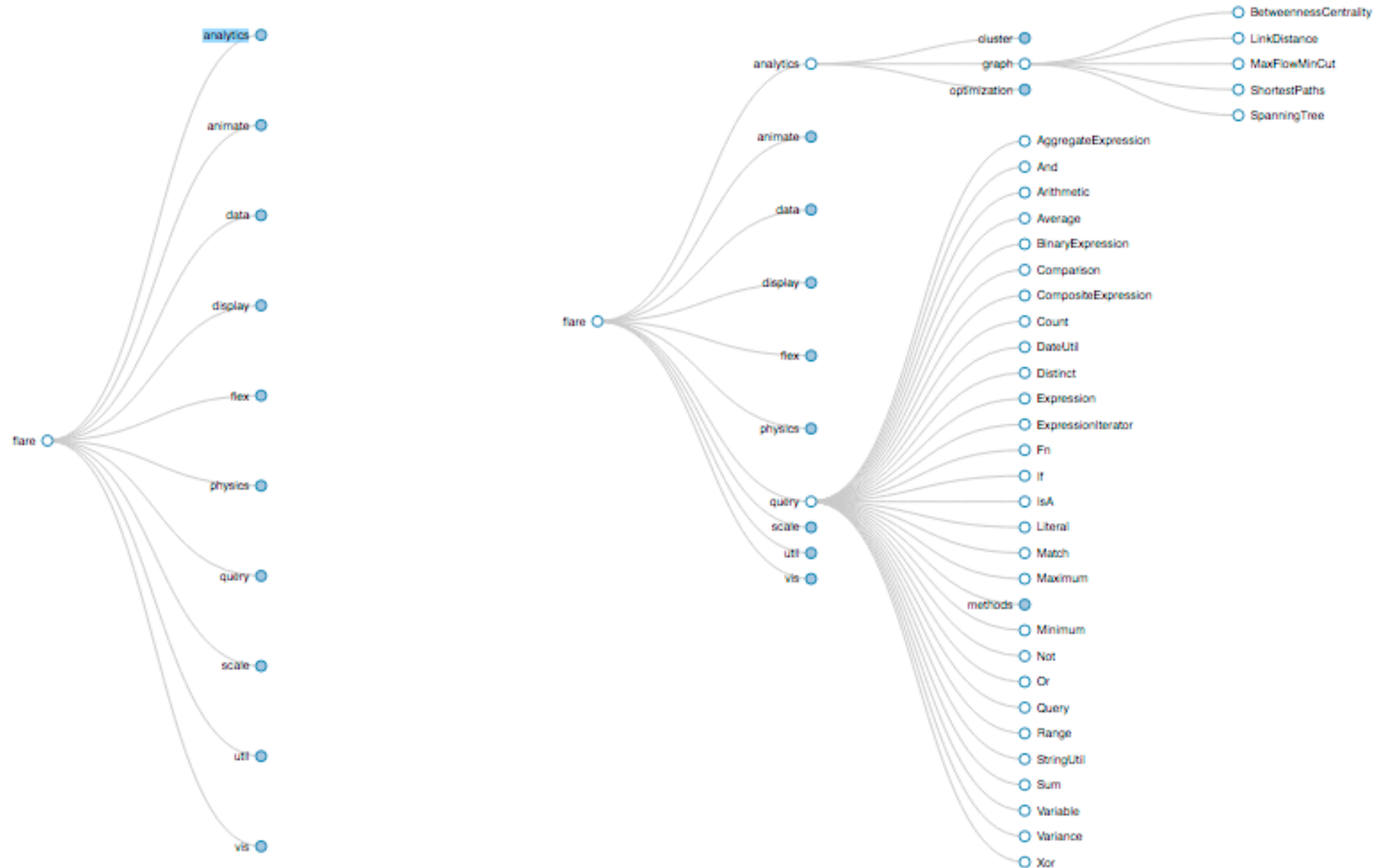
- smooth transition from one state to another
 - alternative to jump cuts, supports item tracking
 - best case for animation
 - staging to reduce cognitive load



[Stacked to Grouped Bars](<http://bl.ocks.org/mbostock/3943967>)

Idiom: **Animated transition** - tree detail

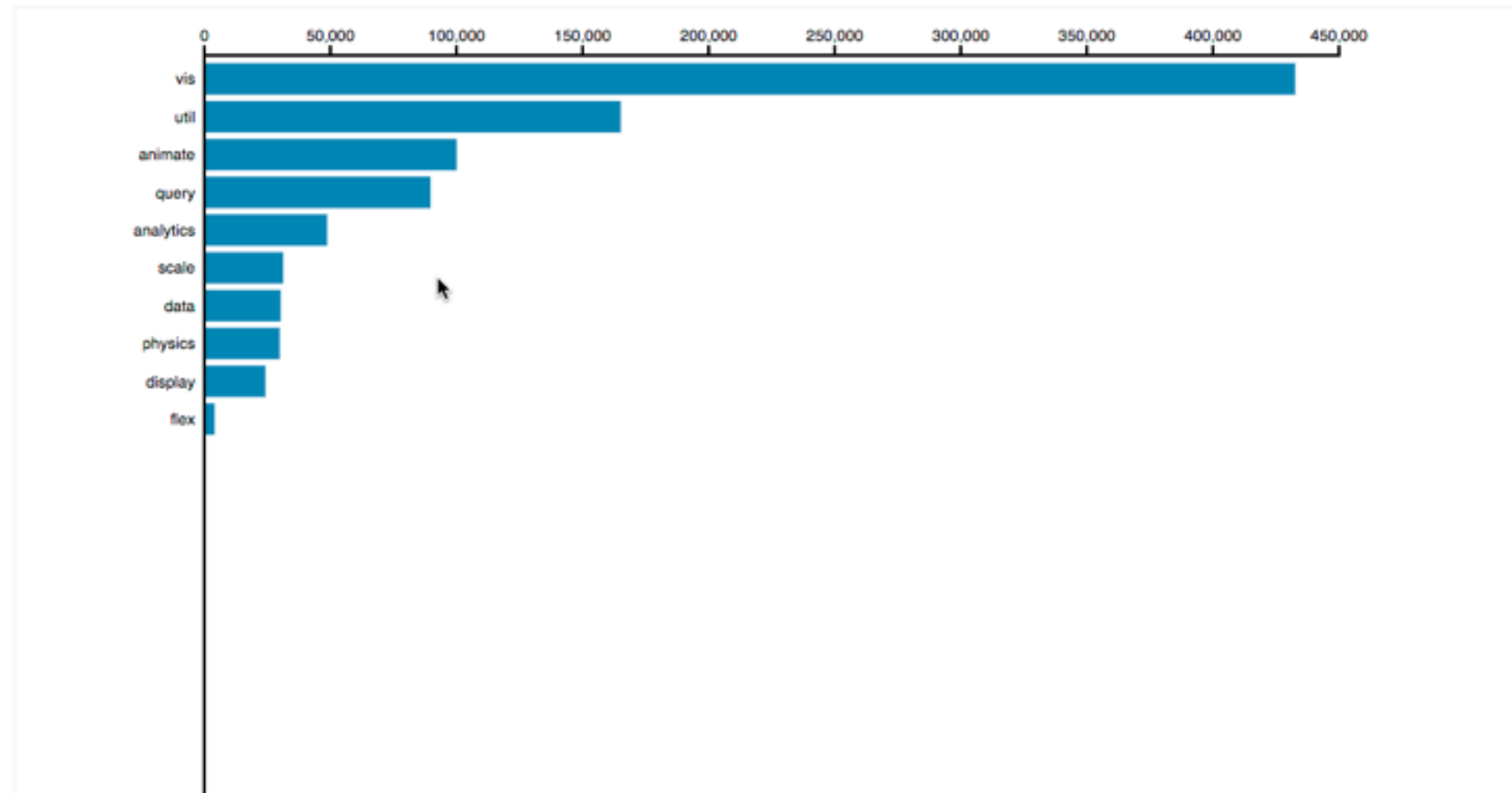
- animated transition
 - network drilldown/rollup



[Collapsible Tree](<https://blocks.org/mbostock/4339083>)

Idiom: **Animated transition - bar detail**

- example: hierarchical bar chart
 - add detail during transition to new level of detail



[Hierarchical Bar Chart](<https://blocks.org/mbostock/1283663>)

Navigate: Changing item visibility

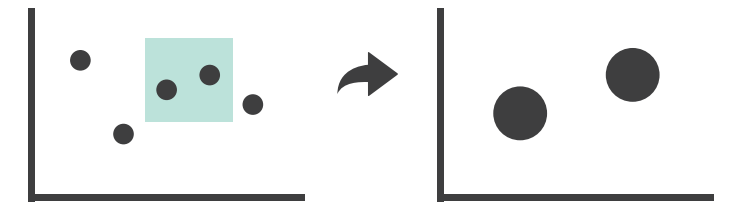
- change viewpoint
 - changes which items are visible within view
 - camera metaphor
 - zoom
 - geometric zoom: familiar semantics
 - semantic zoom: adapt object representation based on available pixels
 - » dramatic change, or more subtle one
 - pan/translate
 - rotate
 - especially in 3D
 - constrained navigation
 - often with animated transitions
 - often based on selection set

➔ Navigate

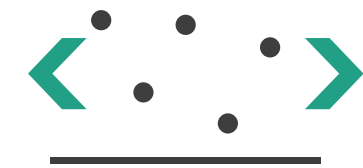
➔ Item Reduction

➔ Zoom

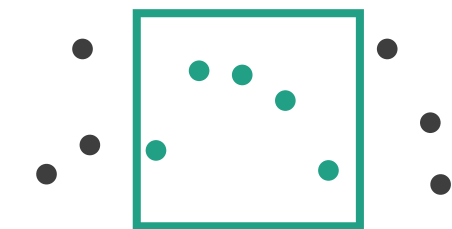
Geometric or *Semantic*



➔ Pan/Translate



➔ Constrained



Further reading

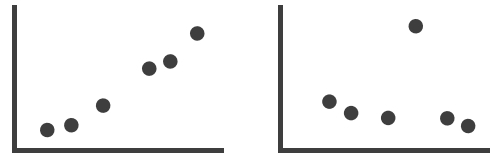
- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
 - Chap 11: Manipulate View*
- *Animated Transitions in Statistical Data Graphics*. Heer and Robertson. IEEE Trans. on Visualization and Computer Graphics (Proc. InfoVis07) 13:6 (2007), 1240–1247.
- *Selection: 524,288 Ways to Say “This is Interesting”*. Wills. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 54–61, 1996.
- *Smooth and efficient zooming and panning*. van Wijk and Nuij. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 15–22, 2003.
- *Starting Simple - adding value to static visualisation through simple interaction*. Dix and Ellis. Proc. Advanced Visual Interfaces (AVI), pp. 124–134, 1998.

Outline

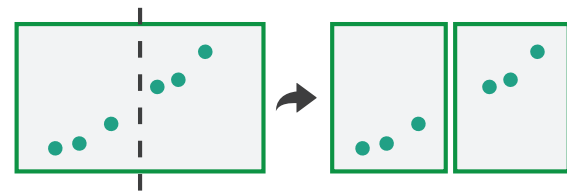
- **Session 1** / 10-11:30am
Data Visualization Pitfalls to Avoid
 - Introduction
 - Color
 - Space: 2D vs 3D
- **Session 2** / 12:30-3pm
Visualization Analysis & Design, In More Depth
 - Marks and Channels, Perception
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

Facet

→ Juxtapose



→ Partition



→ Superimpose



Juxtapose and coordinate views

→ Share Encoding: Same/Different

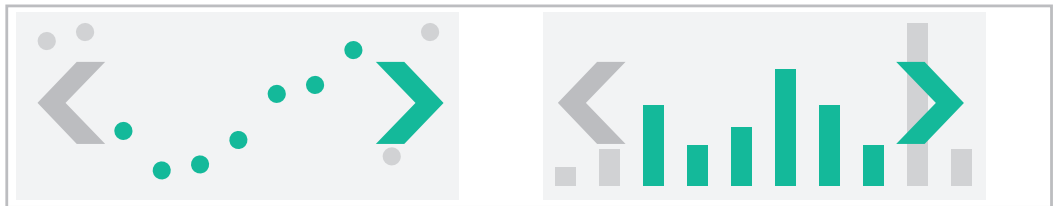
→ *Linked Highlighting*



→ Share Data: All/Subset/None



→ Share Navigation



Idiom: **Linked highlighting**

System: **EDV**

- see how regions contiguous in one view are distributed within another
 - powerful and pervasive interaction idiom
- encoding: different
 - *multiform*
- data: all shared



[Visual Exploration of Large Structured Datasets. Wills. Proc. New Techniques and Trends in Statistics (NTTS), pp. 237–246. IOS Press, 1995.]

Idiom: **bird's-eye maps**

System: **Google Maps**

- encoding: same
- data: subset shared
- navigation: shared
 - bidirectional linking

- differences
 - viewpoint
 - (size)

- ***overview-detail***



[A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. Cockburn, Karlson, and Bederson. *ACM Computing Surveys* 41:1 (2008), 1–31.]

Idiom: **Small multiples**

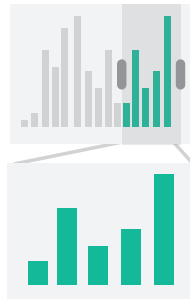
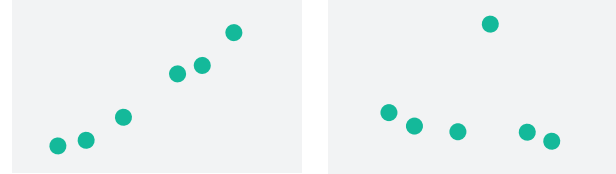


System: **Cerebral**

- encoding: same
- data: none shared
 - different attributes for node colors
 - (same network layout)
- navigation: shared



[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Kincaid. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2008)* 14:6 (2008), 1253–1260.]

Coordinate views: Design choice interaction

		Data		
		All	Subset	None
Encoding	Same	Redundant	 Overview/ Detail	 Small Multiples
	Different	 Multiform	 Multiform, Overview/ Detail	No Linkage

- why juxtapose views?

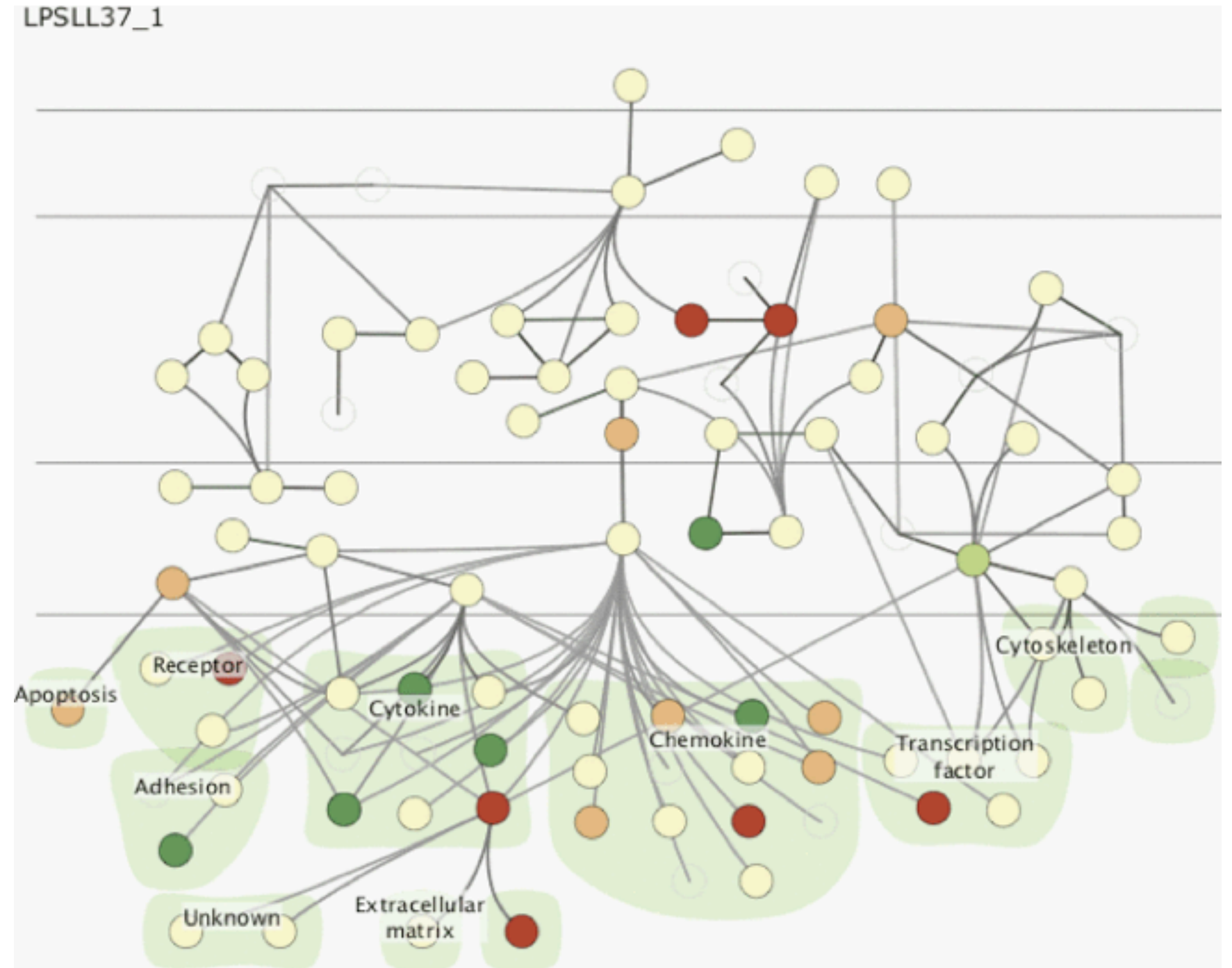
- benefits: eyes vs memory

- lower cognitive load to move eyes between 2 views than remembering previous state with single changing view

- costs: display area, 2 views side by side each have only half the area of one view

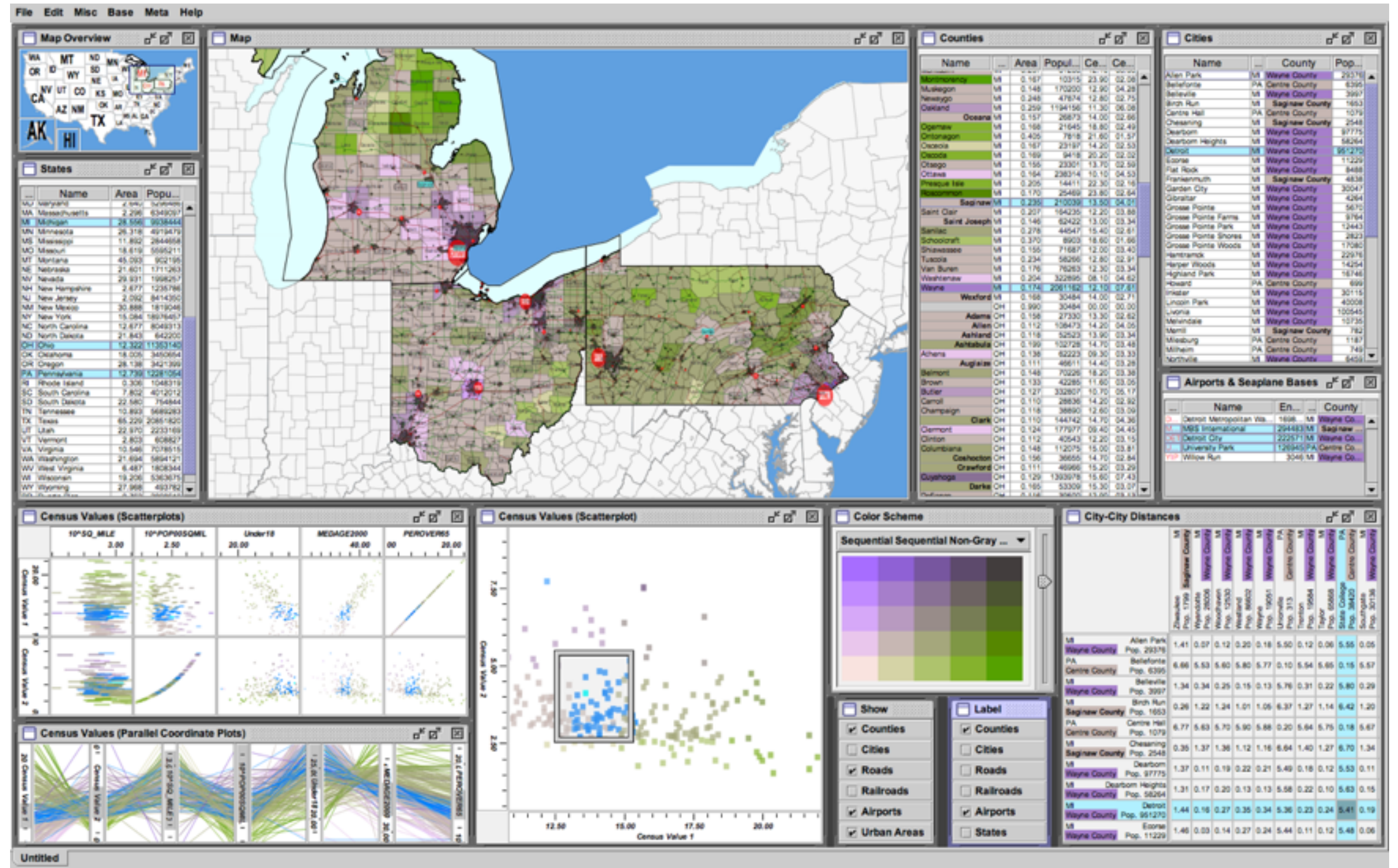
Why not animation?

- disparate frames and regions: comparison difficult
 - vs contiguous frames
 - vs small region
 - vs coherent motion of group
- safe special case
 - animated transitions



System: **Improvise**

- investigate power of multiple views
 - pushing limits on view count, interaction complexity
 - how many is ok?
 - open research question
 - reorderable lists
 - easy lookup
 - useful when linked to other encodings

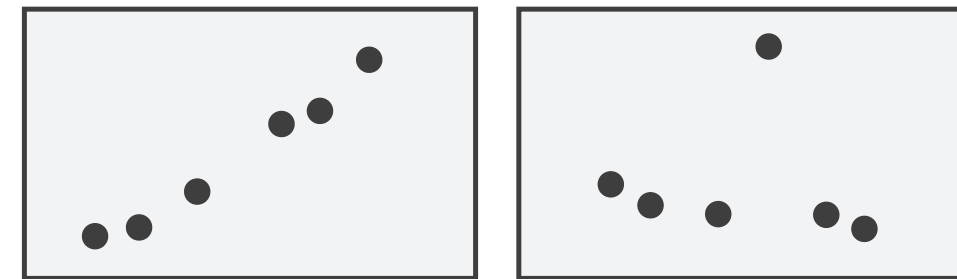


[Building Highly-Coordinated Visualizations In Improvise. Weaver. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 159–166, 2004.]

Partition into views

- how to divide data between views → Partition into Side-by-Side Views

- split into regions by attributes
- encodes association between items using spatial proximity
- order of splits has major implications for what patterns are visible



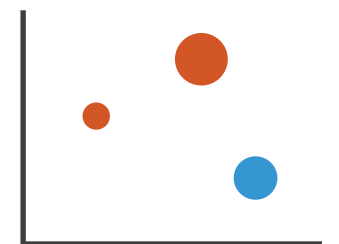
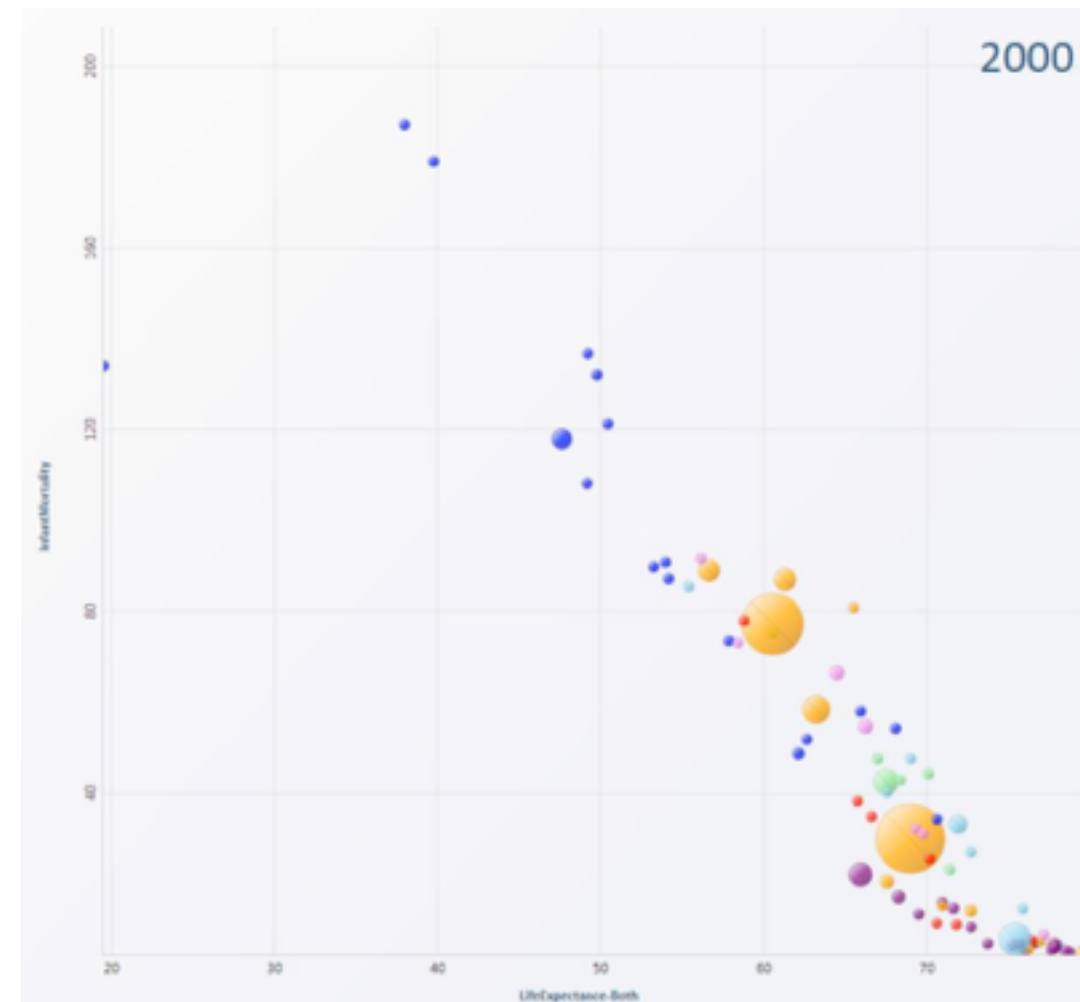
- no strict dividing line

– *view: big/detailed*

- contiguous region in which visually encoded data is shown on the display

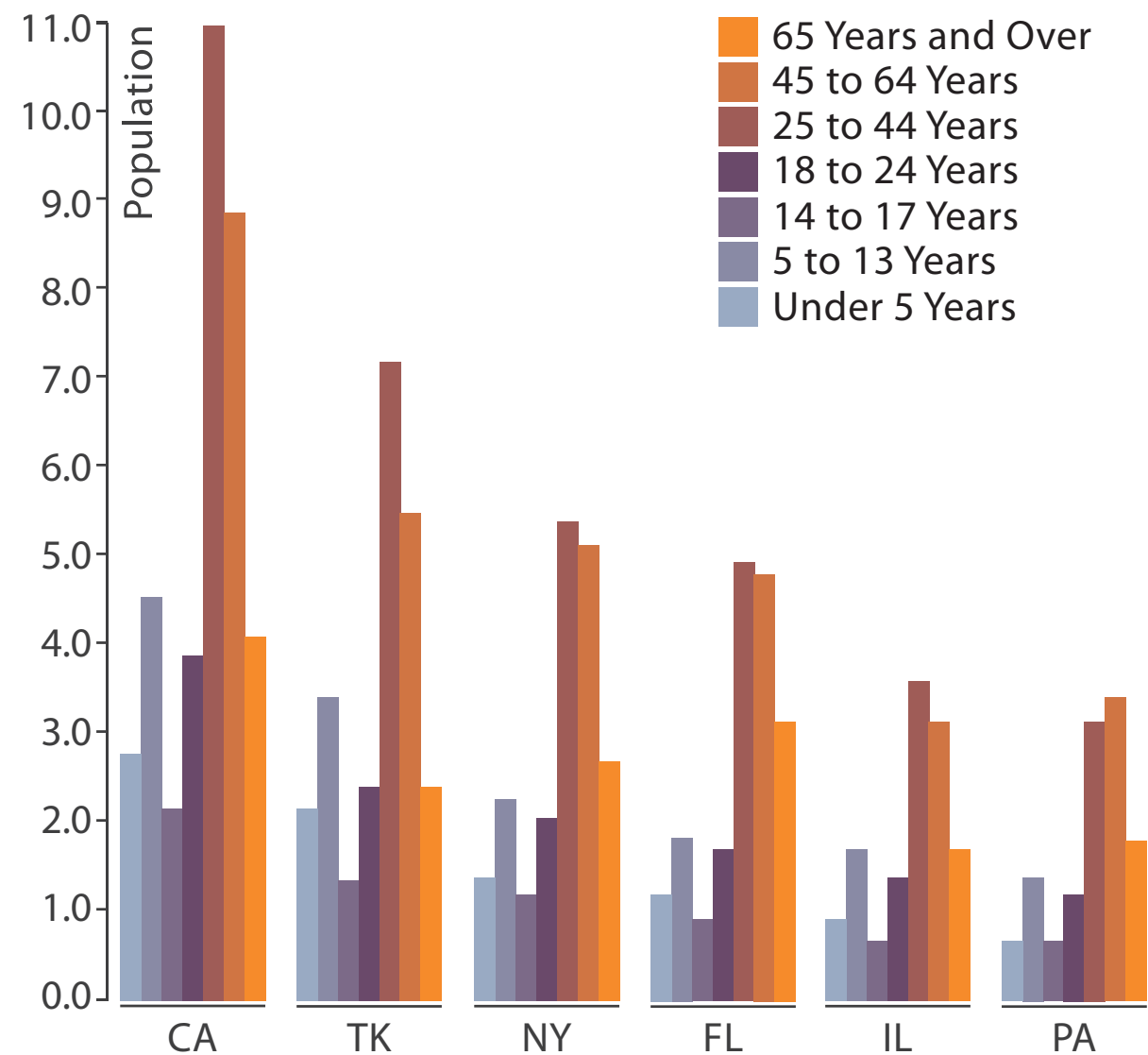
– *glyph: small/iconic*

- object with internal structure that arises from multiple marks

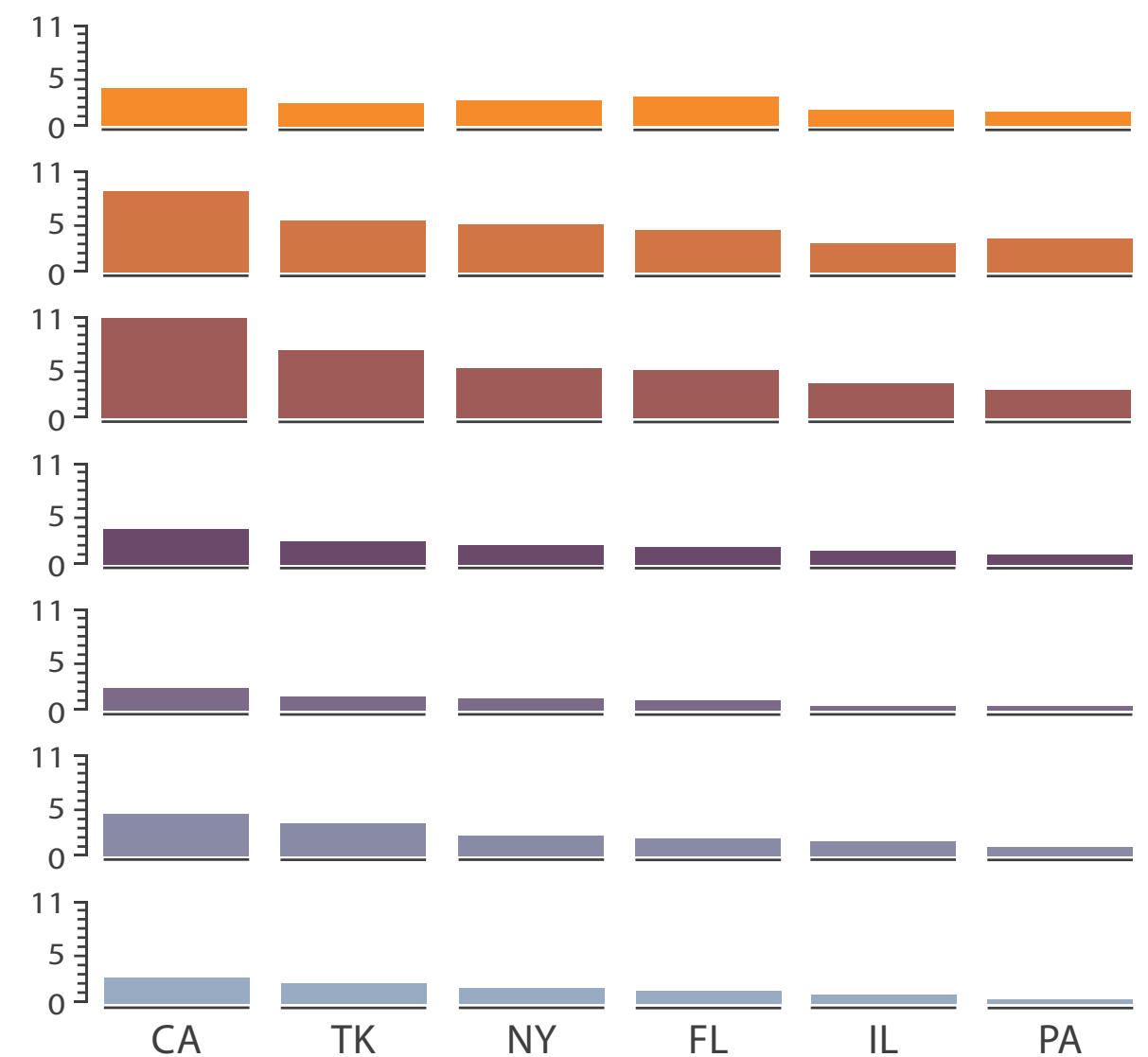


Partitioning: List alignment

- single bar chart with grouped bars
 - split by state into regions
 - complex glyph within each region showing all ages
 - compare: easy within state, hard across ages



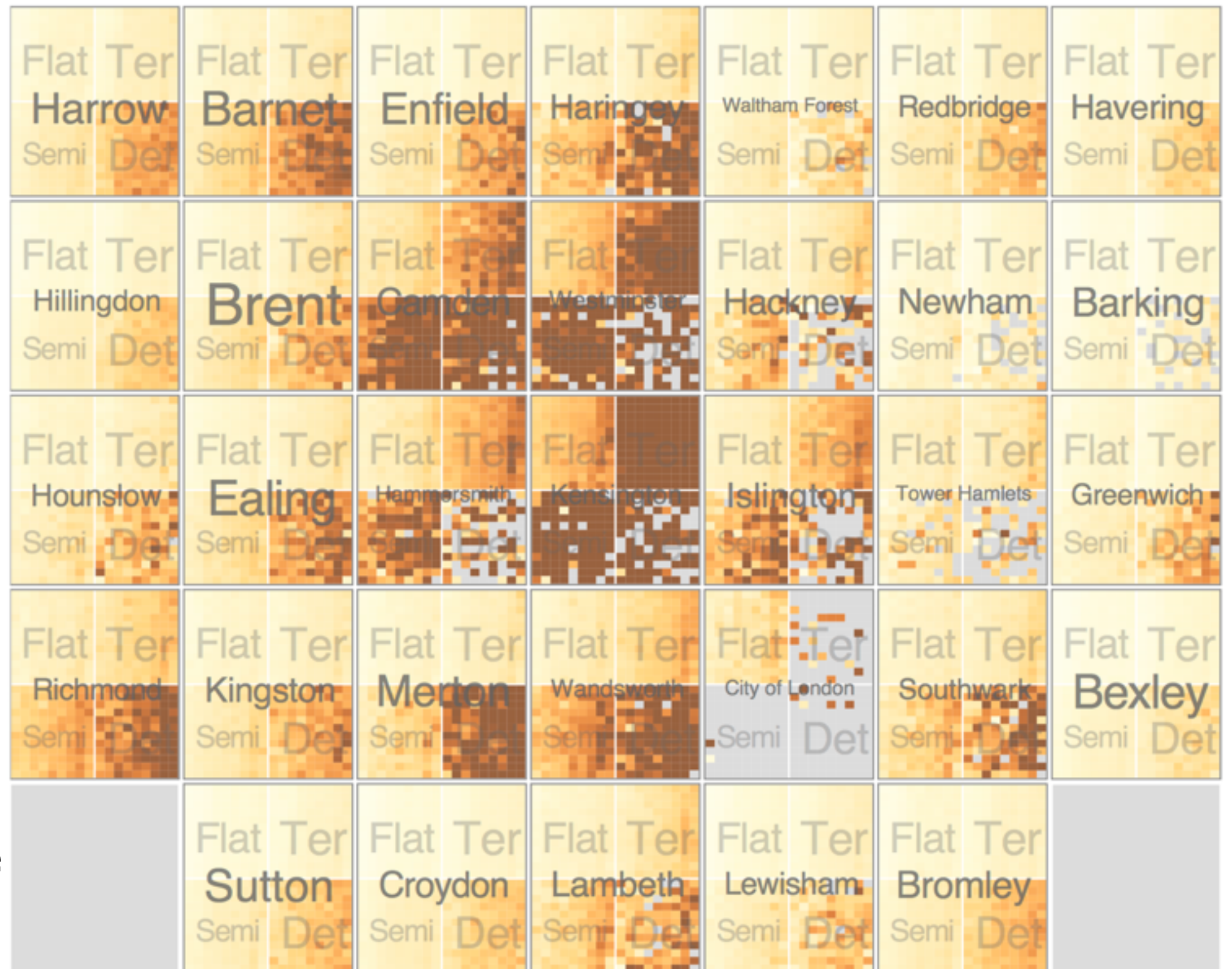
- small-multiple bar charts
 - split by age into regions
 - one chart per region
 - compare: easy within age, harder across states



Partitioning: Recursive subdivision

System: **HIVE**

- split by neighborhood
- then by type
- then time
 - years as rows
 - months as columns
- color by price
- neighborhood patterns
 - where it's expensive
 - where you pay much more for detached type



Partitioning: Recursive subdivision

System: **HIVE**

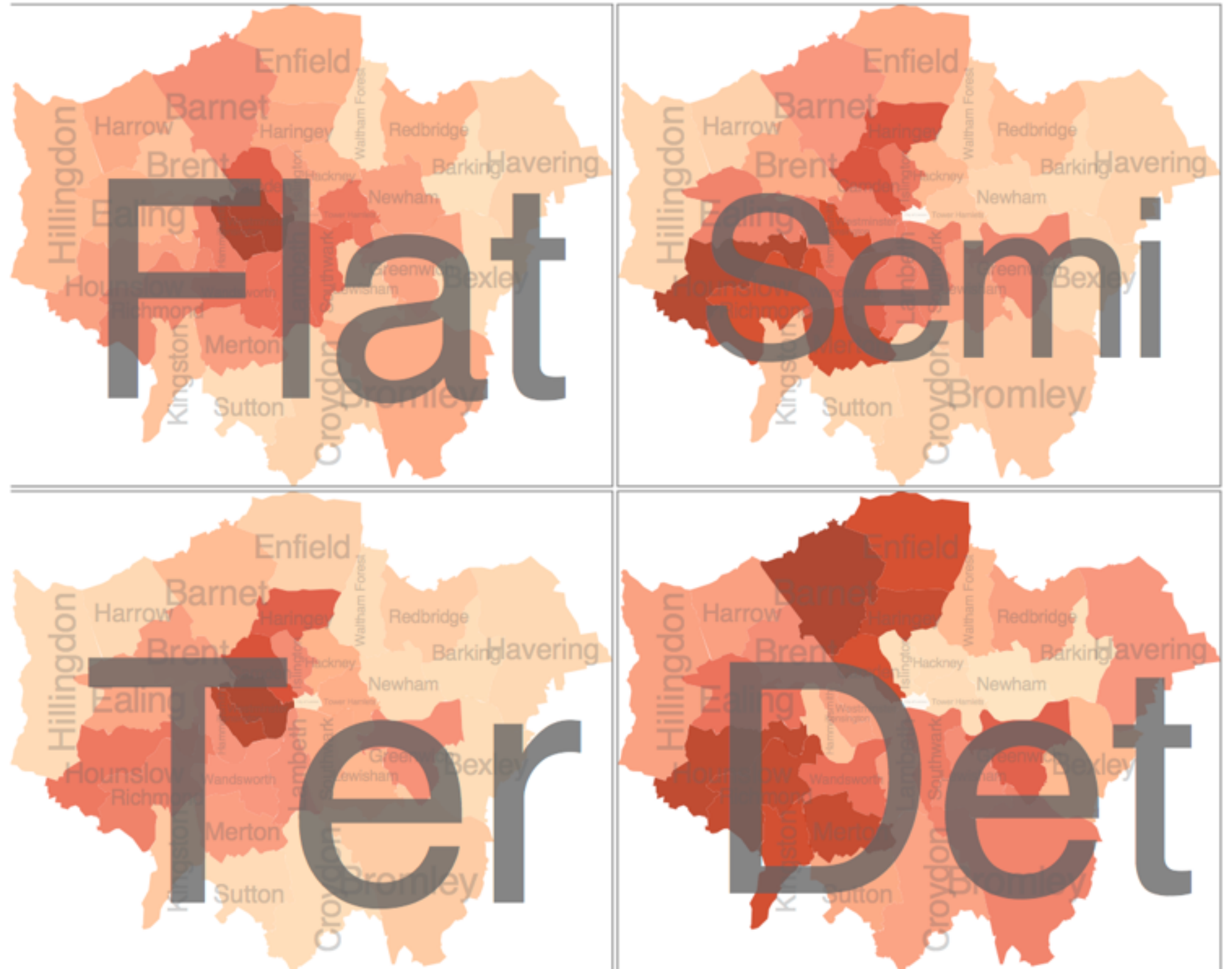
- switch order of splits
 - type then neighborhood
- switch color
 - by price variation
- type patterns
 - within specific type, which neighborhoods inconsistent



Partitioning: Recursive subdivision

System: **HIVE**

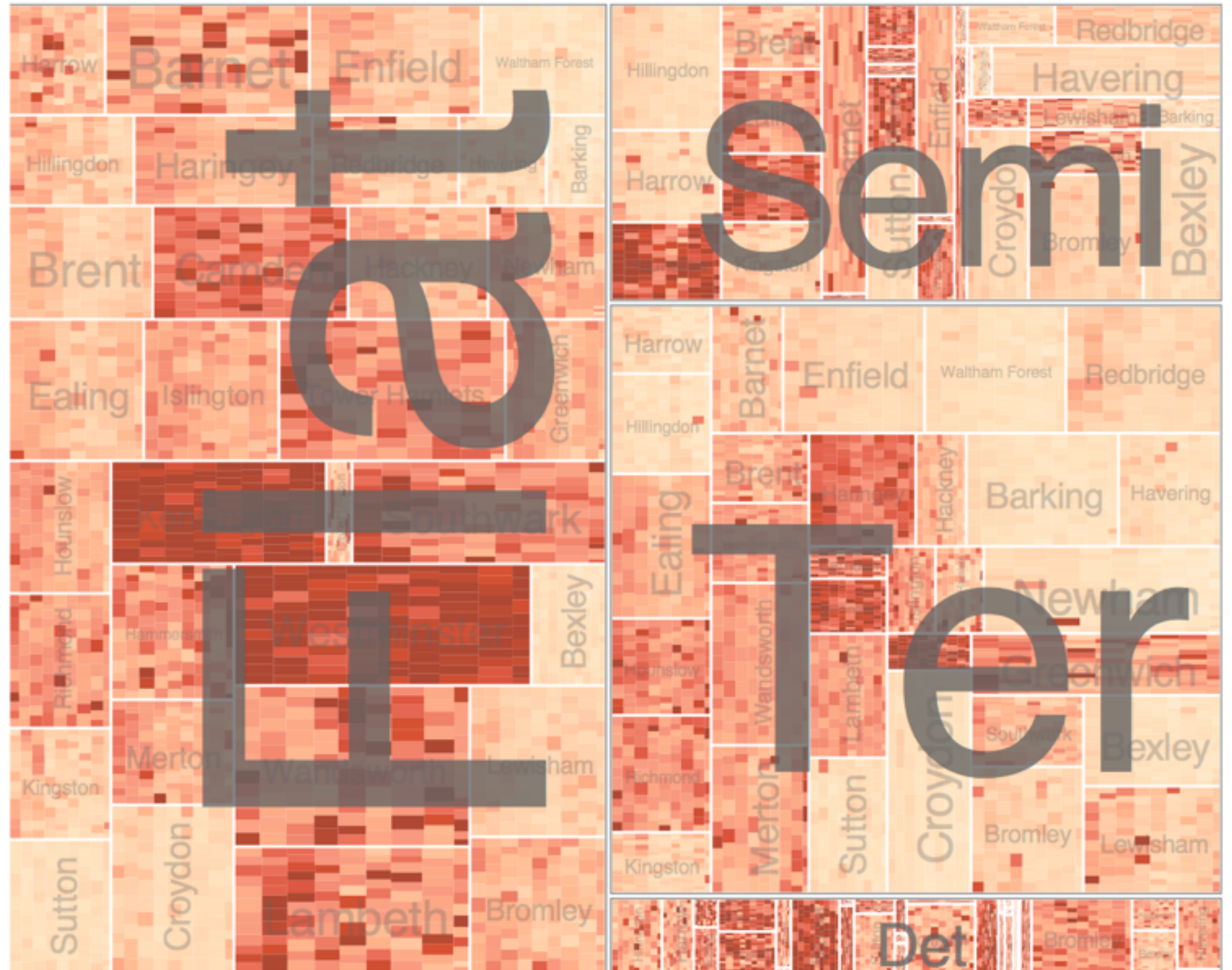
- different encoding for second-level regions
– choropleth maps



Partitioning: Recursive subdivision

System: **HIVE**

- size regions by sale counts
 - not uniformly
- result: treemap



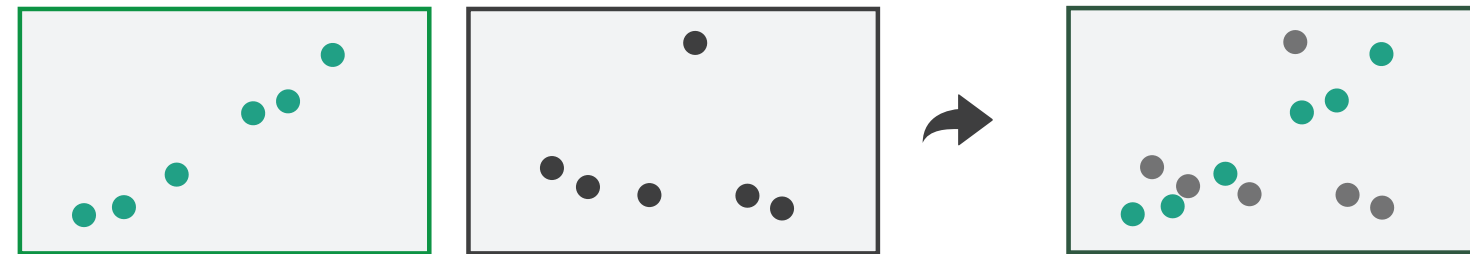
Superimpose layers

- *layer*: set of objects spread out over region
 - each set is visually distinguishable group
 - extent: whole view

➔ Superimpose Layers

- design choices

- how many layers, how to distinguish?
 - encode with different, nonoverlapping channels
 - two layers achievable, three with careful design
- small static set, or dynamic from many possible?



Static visual layering

- foreground layer: roads
 - hue, size distinguishing main from minor
 - high luminance contrast from background
- background layer: regions
 - desaturated colors for water, parks, land areas
- user can selectively focus attention
- “get it right in black and white”
 - check luminance contrast with greyscale view

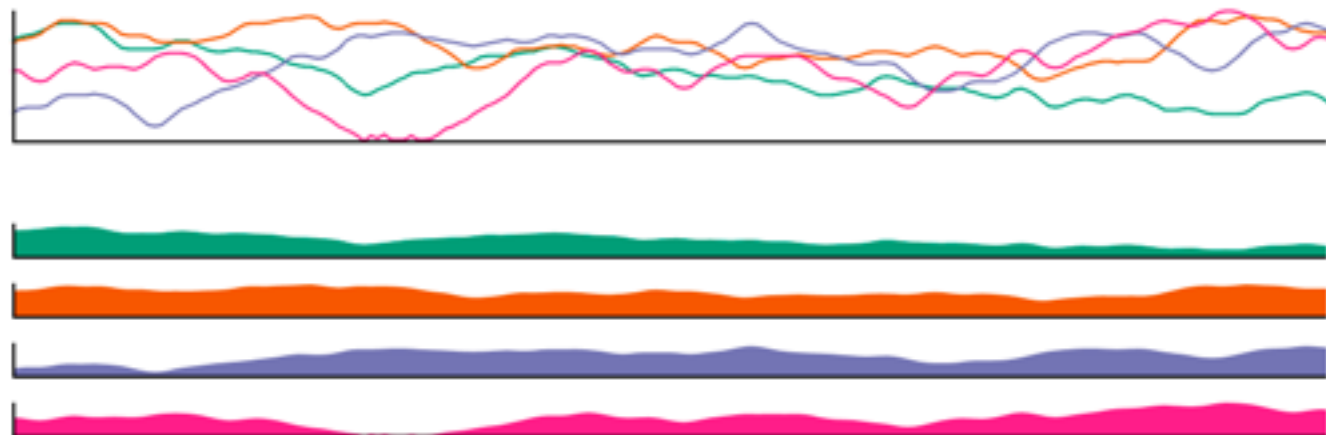


[Get it right in black and white. Stone. 2010.

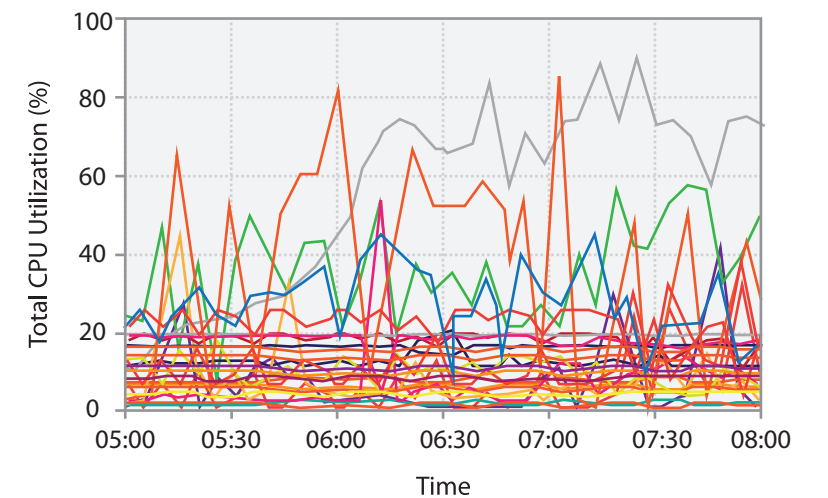
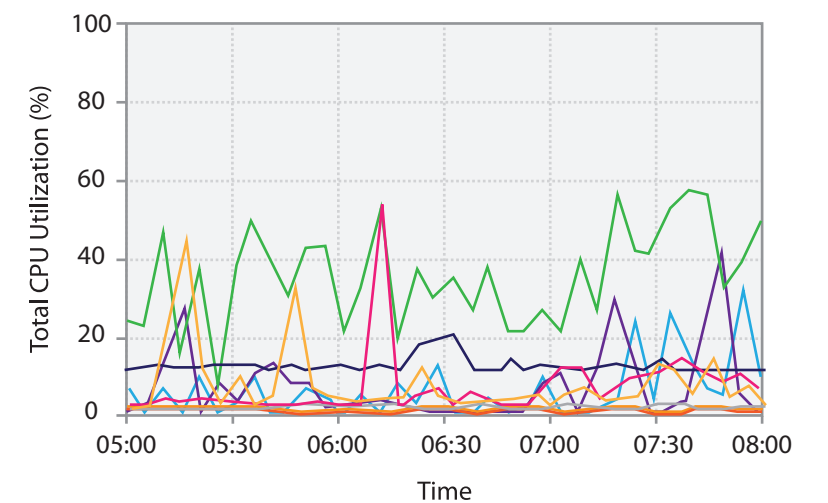
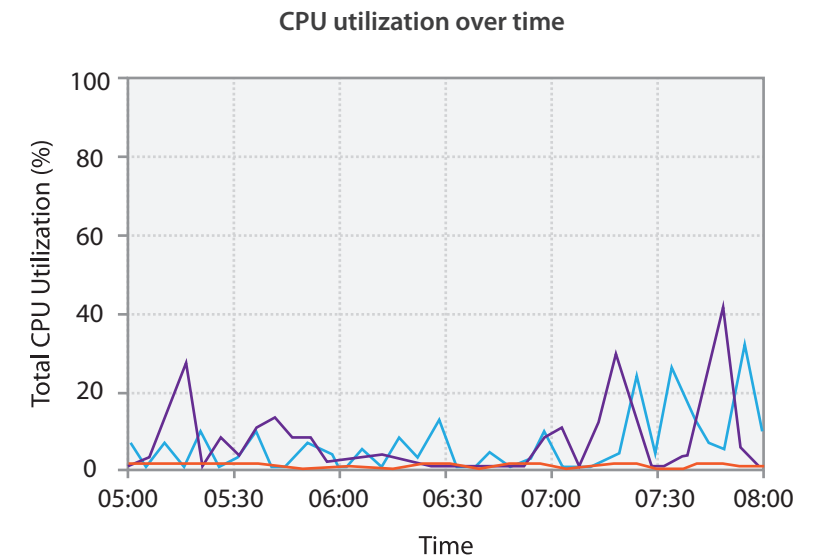
<http://www.stonesc.com/wordpress/2010/03/get-it-right-in-black-and-white>]

Superimposing limits

- few layers, but many lines
 - up to a few dozen
 - but not hundreds
- superimpose vs juxtapose: empirical study
 - superimposed for local, multiple for global
 - tasks
 - local: maximum, global: slope, discrimination
 - same screen space for all multiples vs single superimposed

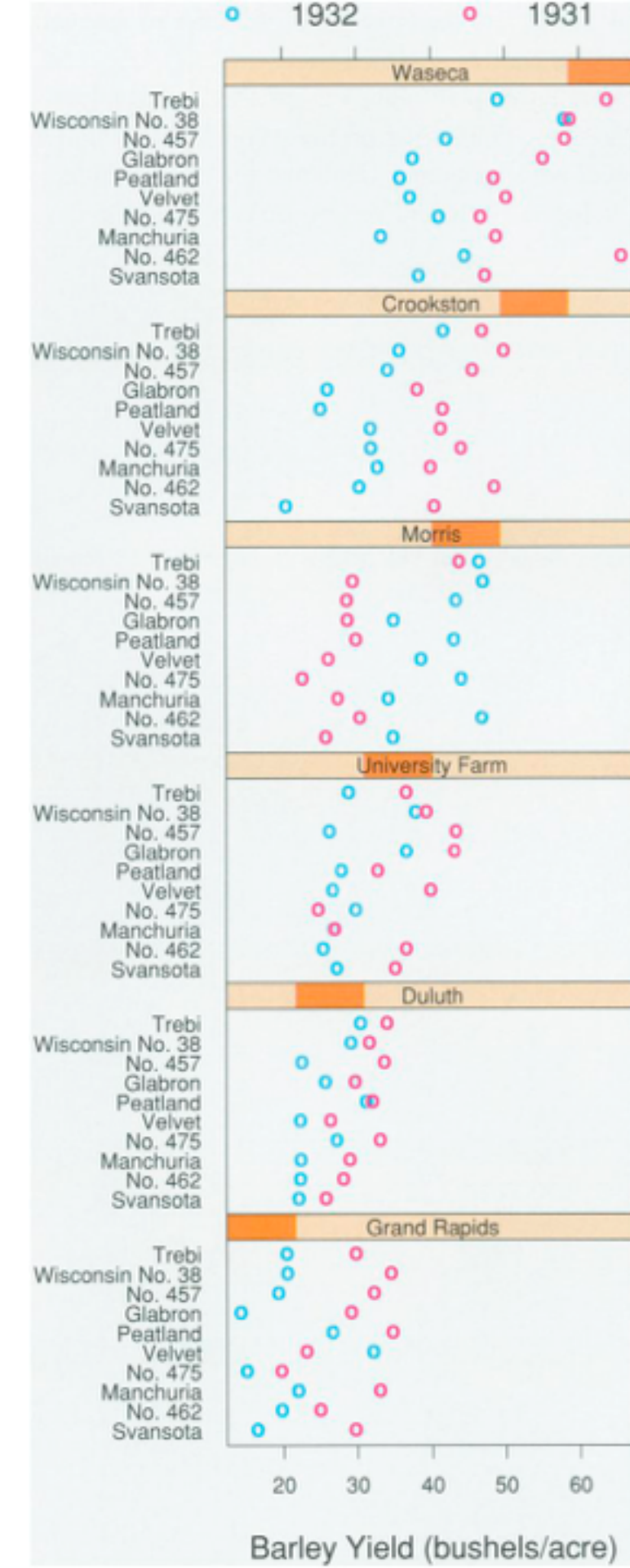


[Graphical Perception of Multiple Time Series. Javed, McDonnel, and Elmqvist. IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE InfoVis 2010) 16:6 (2010), 927–934.]



Idiom: Trellis plots

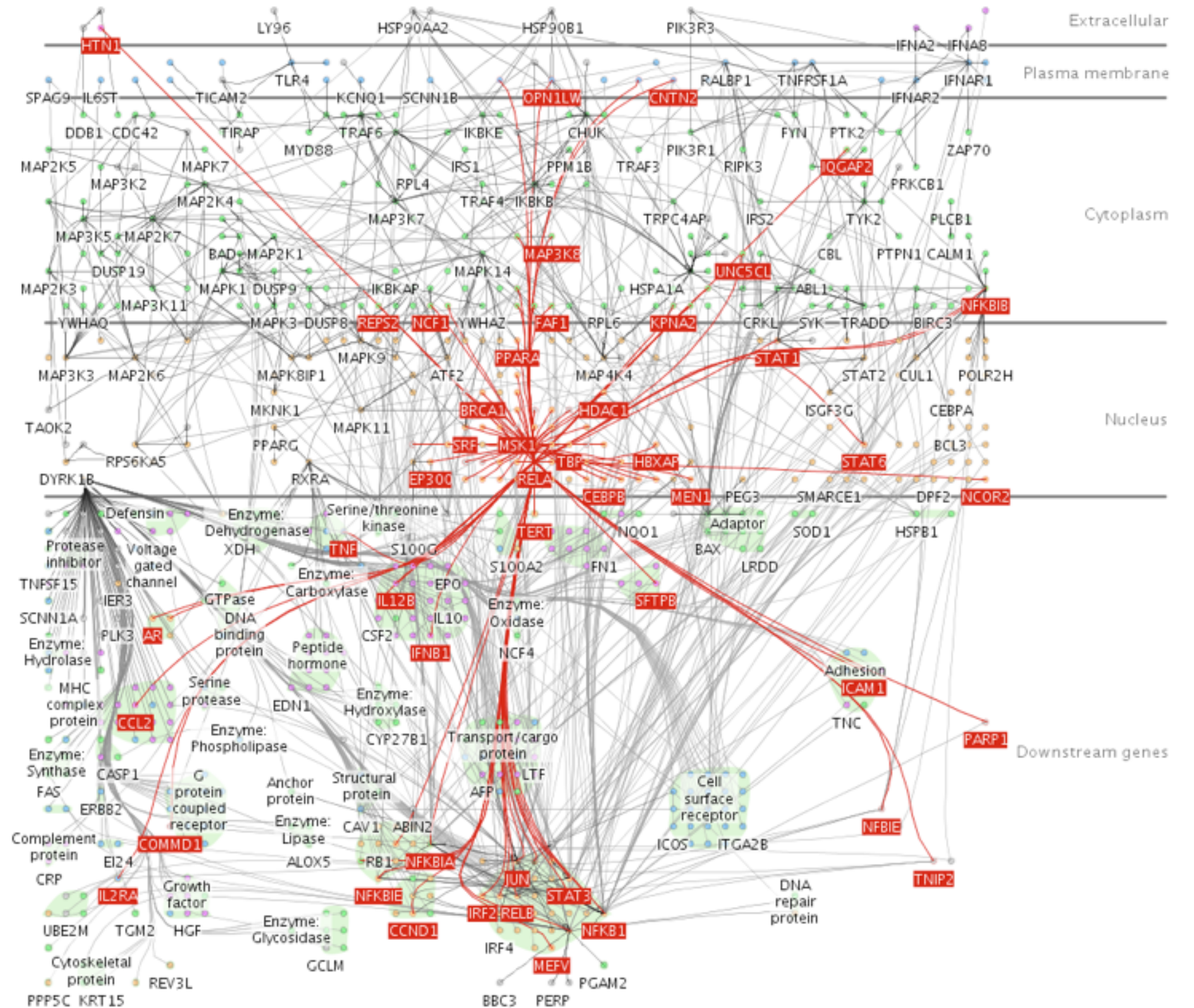
- superimpose within same frame
 - color code by year



Dynamic visual layering

- interactive, from selection
 - lightweight: click
 - very lightweight: hover
- ex: 1-hop neighbors

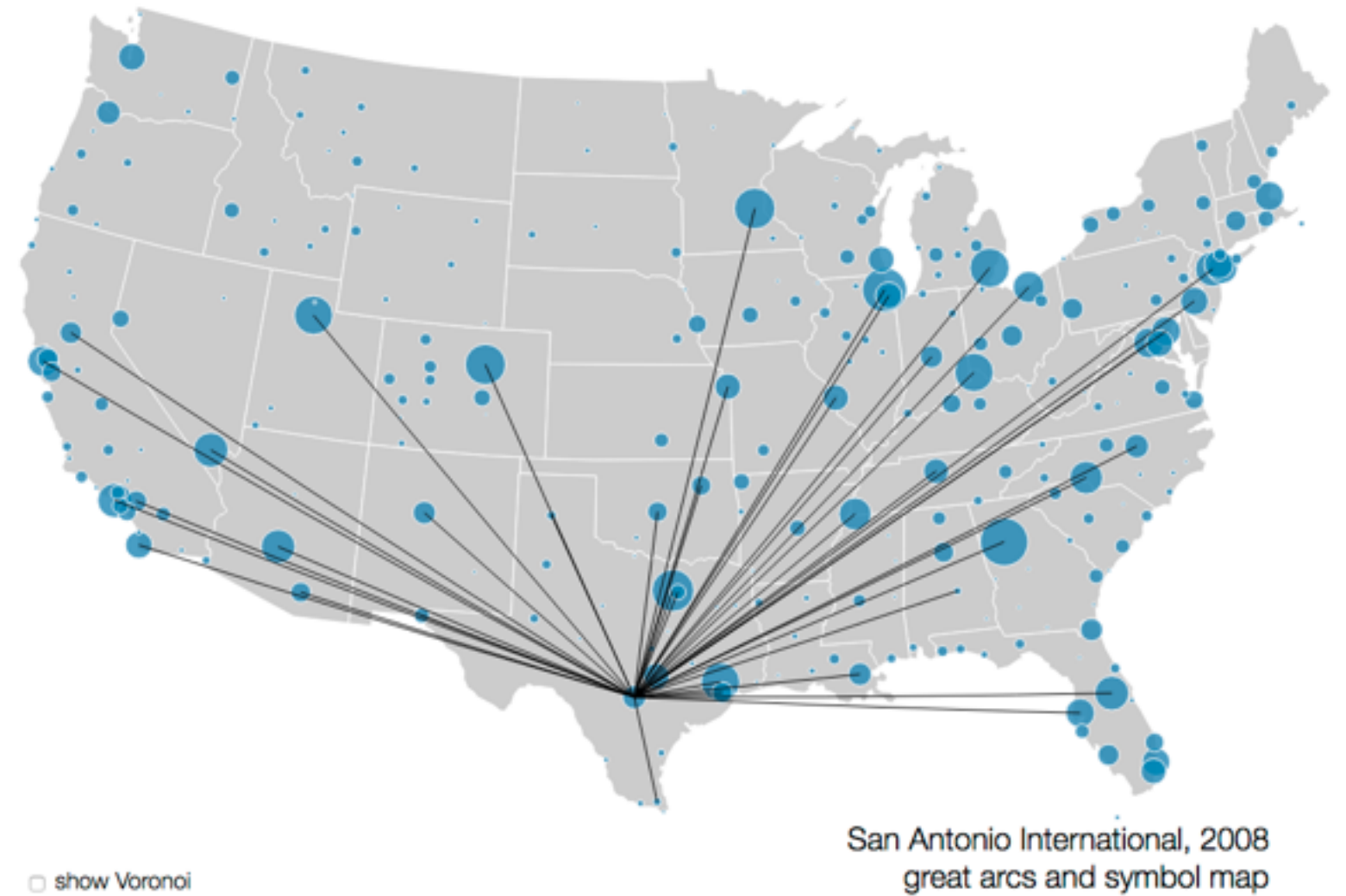
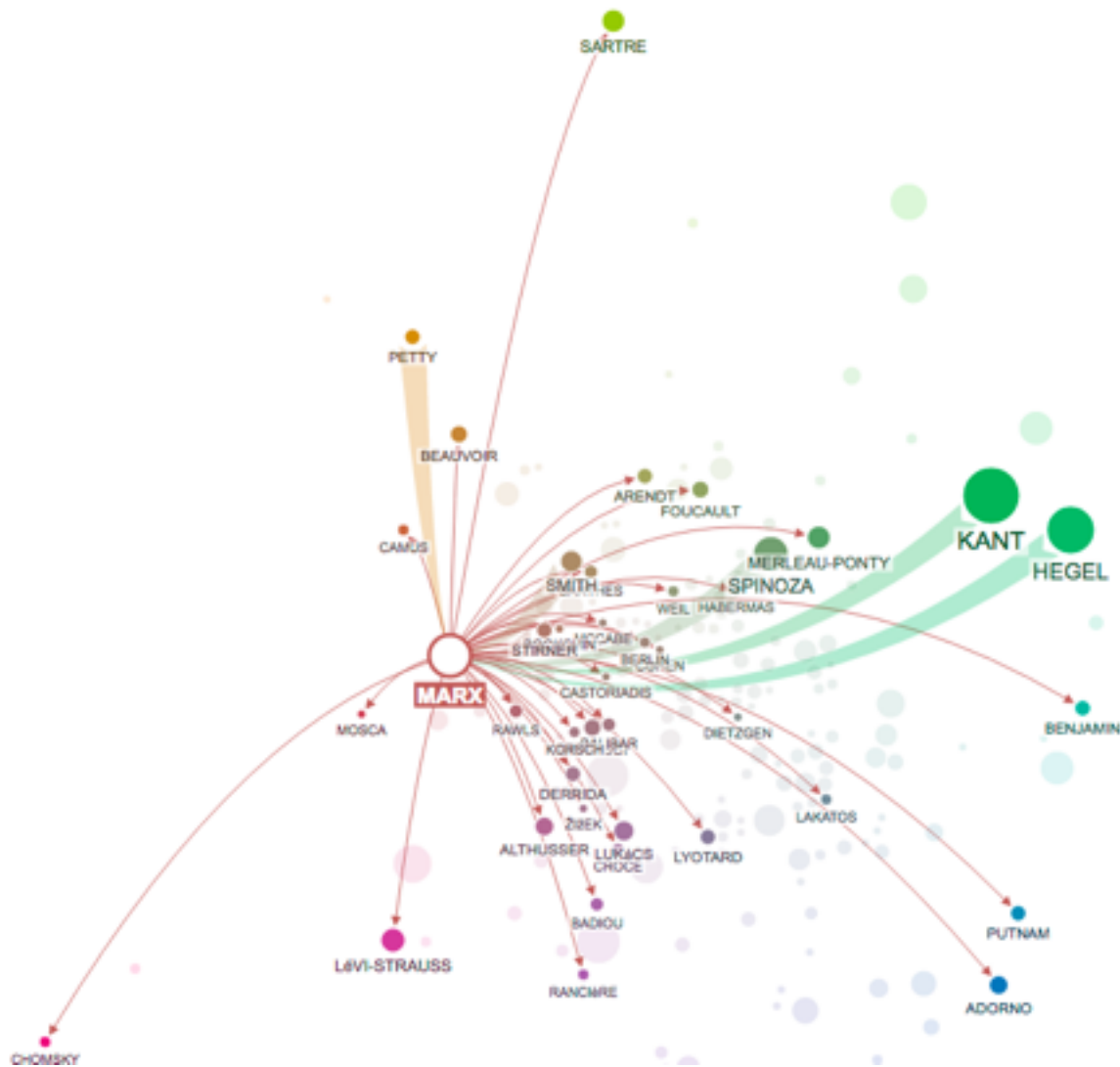
System: Cerebral



[Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. Barsky, Gardy, Hancock, and Munzner. *Bioinformatics* 23:8 (2007), 1040–1042.]

Dynamic visual layering

- one-hop neighbour highlighting demos: click vs hover



<http://mariandoerk.de/edgemaps/demo/>

<http://mbostock.github.io/d3/talk/20111116/airports.html>

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
–*Chap 12: Facet Into Multiple Views*
- *A Review of Overview+Detail, Zooming, and Focus+Context Interfaces*. Cockburn, Karlson, and Bederson. ACM Computing Surveys 41:1 (2008), 1–31.
- *A Guide to Visual Multi-Level Interface Design From Synthesis of Empirical Study Evidence*. Lam and Munzner. Synthesis Lectures on Visualization Series, Morgan Claypool, 2010.
- *Zooming versus multiple window interfaces: Cognitive costs of visual comparisons*. Plumlee and Ware. ACM Trans. on Computer-Human Interaction (ToCHI) 13:2 (2006), 179–209.
- *Exploring the Design Space of Composite Visualization*. Javed and Elmqvist. Proc. Pacific Visualization Symp. (PacificVis), pp. 1–9, 2012.
- *Visual Comparison for Information Visualization*. Gleicher, Albers, Walker, Jusufi, Hansen, and Roberts. Information Visualization 10:4 (2011), 289–309.
- *Guidelines for Using Multiple Views in Information Visualizations*. Baldonado, Woodruff, and Kuchinsky. In Proc. ACM Advanced Visual Interfaces (AVI), pp. 110–119, 2000.
- *Cross-Filtered Views for Multidimensional Visual Analysis*. Weaver. IEEE Trans. Visualization and Computer Graphics 16:2 (Proc. InfoVis 2010), 192–204, 2010.
- *Linked Data Views*. Wills. In Handbook of Data Visualization, Computational Statistics, edited by Unwin, Chen, and Härdle, pp. 216–241. Springer-Verlag, 2008.
- *Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications*. Borgo, Kehrer, Chung, Maguire, Laramee, Hauser, Ward, and Chen. In Eurographics State of the Art Reports, pp. 39–63, 2013.

Outline

- **Session 1** / 10-11:30am
Data Visualization Pitfalls to Avoid
 - Introduction
 - Color
 - Space: 2D vs 3D
- **Session 2** / 12:30-3pm
Visualization Analysis & Design, In More Depth
 - Marks and Channels, Perception
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

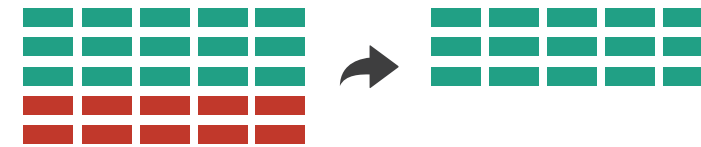
Reduce items and attributes

- reduce/increase: inverses
- filter
 - pro: straightforward and intuitive
 - to understand and compute
 - con: out of sight, out of mind
- aggregation
 - pro: inform about whole set
 - con: difficult to avoid losing signal
- not mutually exclusive
 - combine filter, aggregate
 - combine reduce, change, facet

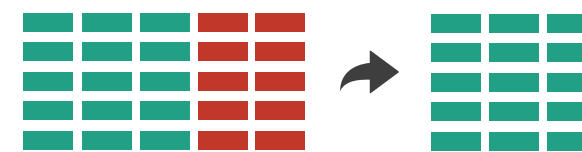
Reducing Items and Attributes

① Filter

→ Items

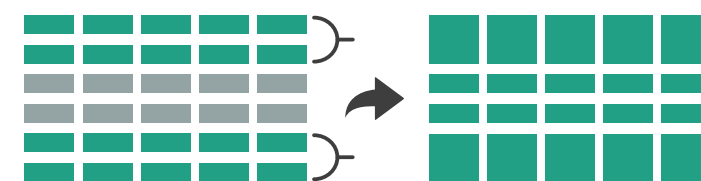


→ Attributes

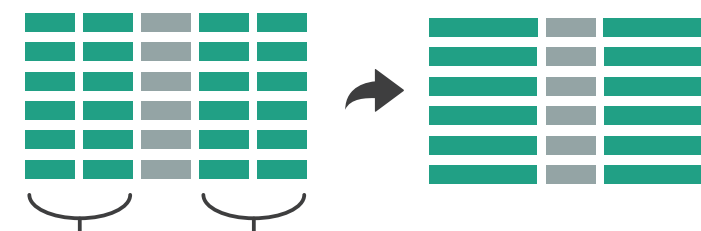


② Aggregate

→ Items



→ Attributes

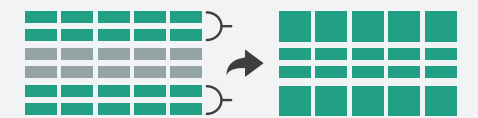


Reduce

① Filter



② Aggregate



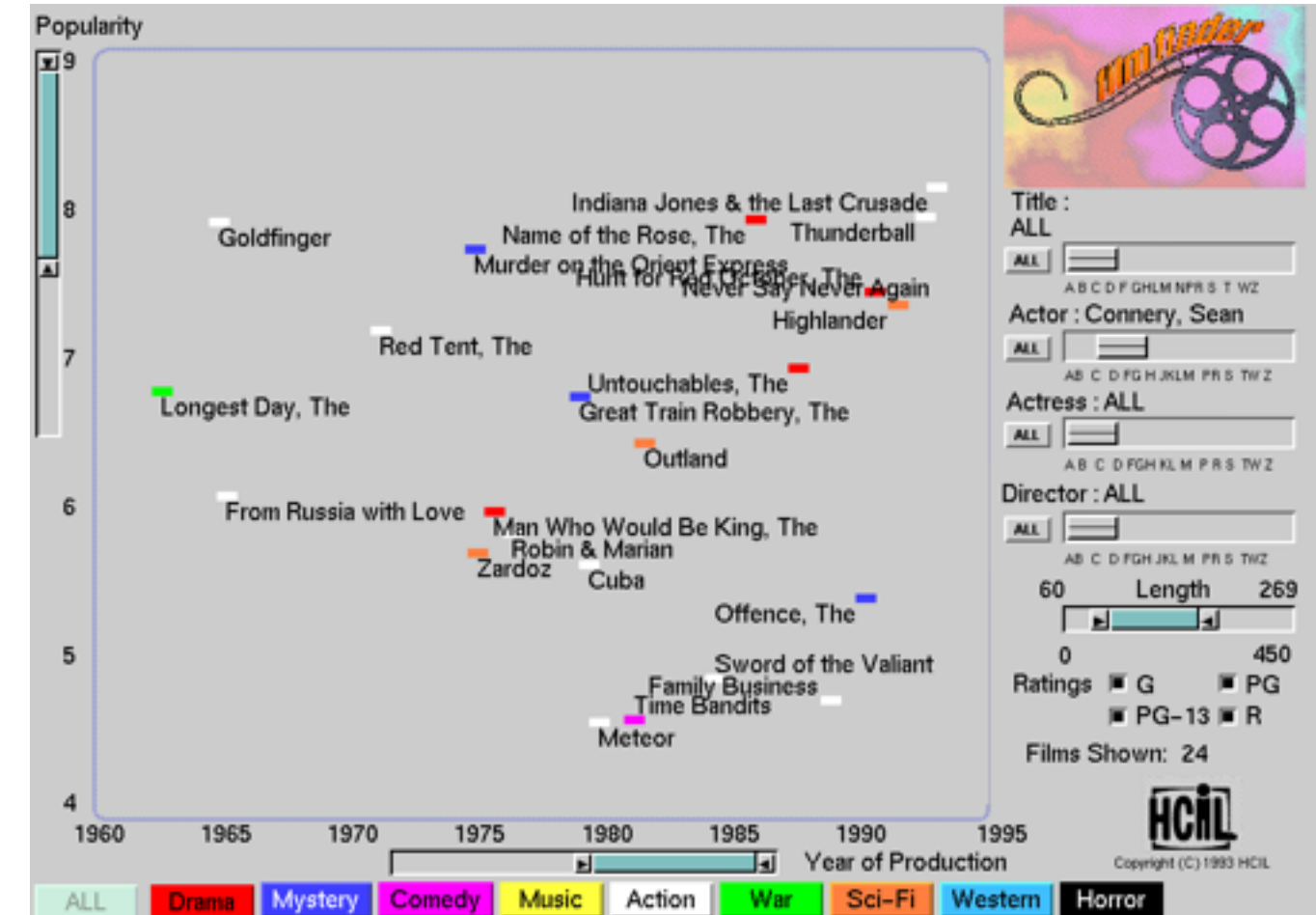
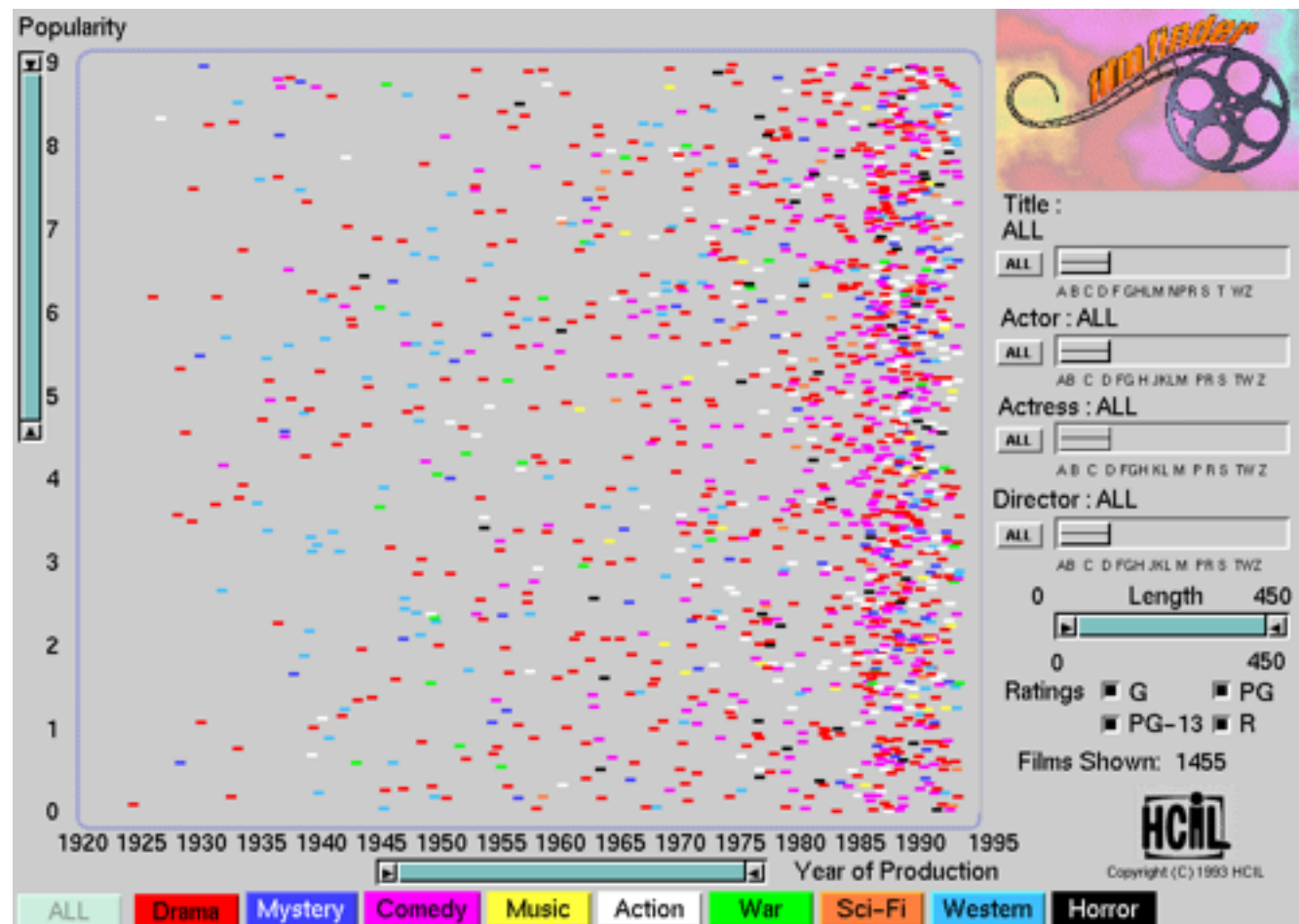
③ Embed



Idiom: **dynamic filtering**

System: **FilmFinder**

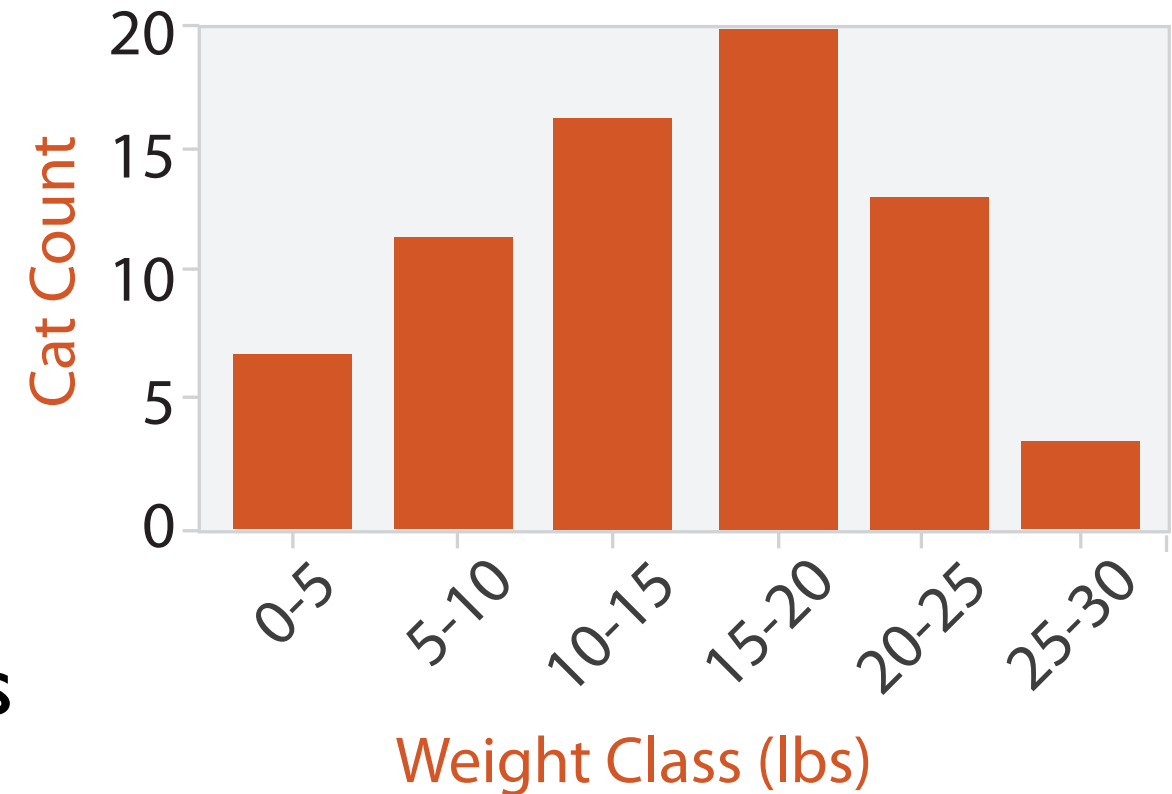
- item filtering
- browse through tightly coupled interaction
 - alternative to queries that might return far too many or too few



[Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Ahlberg and Shneiderman. Proc. ACM Conf. on Human Factors in Computing Systems (CHI), pp. 313–317, 1994.]

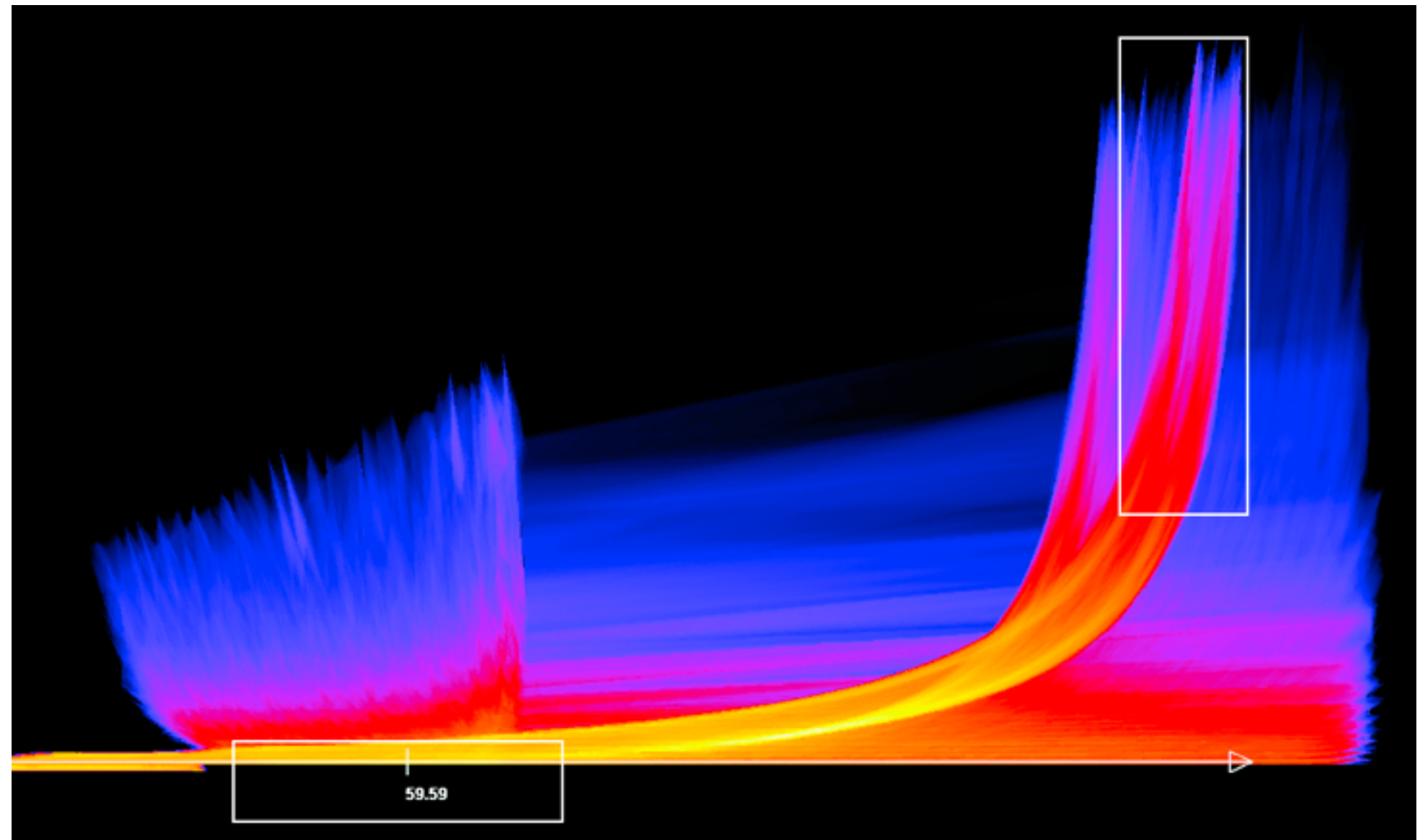
Idiom: **histogram**

- static item aggregation
- task: find distribution
- data: table
- derived data
 - new table: keys are bins, values are counts
- bin size crucial
 - pattern can change dramatically depending on discretization
 - opportunity for interaction: control bin size on the fly



Continuous scatterplot

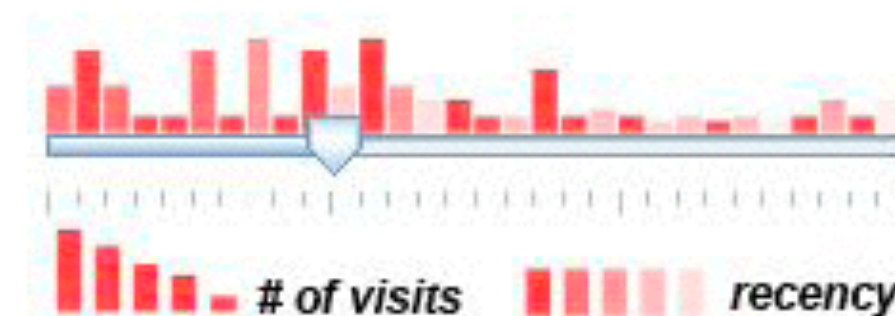
- static item aggregation
- data: table
- derived data: table
 - key attribs x,y for pixels
 - quant attrib: overplot density
- dense space-filling 2D matrix
- color: sequential categorical hue + ordered luminance



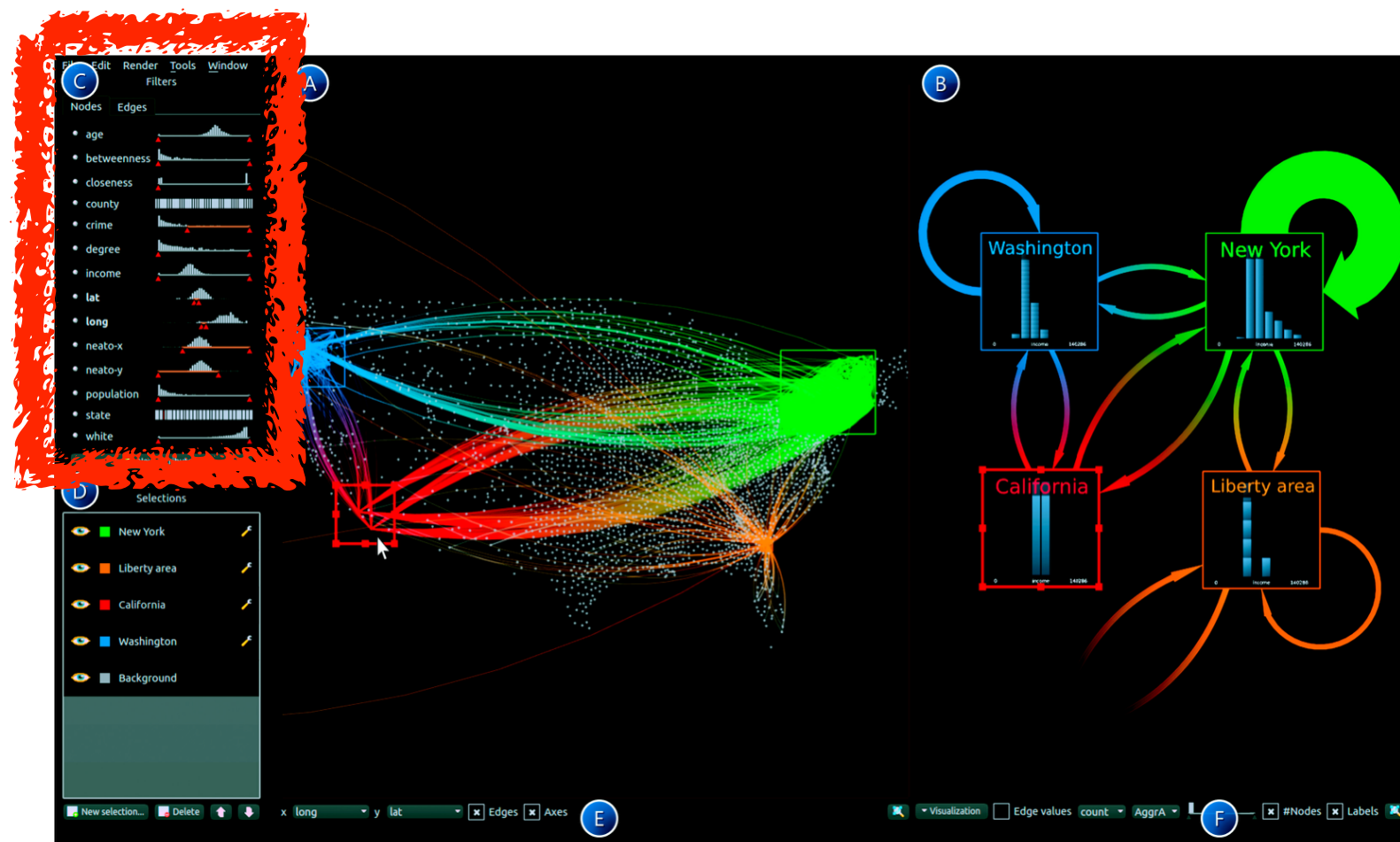
colormap
[Continuous Scatterplots. Bachthaler and Weiskopf. IEEE TVCG (Proc.Vis 08) 14:6 (2008), 1428–1435. 2008.]

Idiom: scented widgets

- augmented widgets show information scent
 - cues to show whether value in drilling down further vs looking elsewhere
- concise use of space: histogram on slider

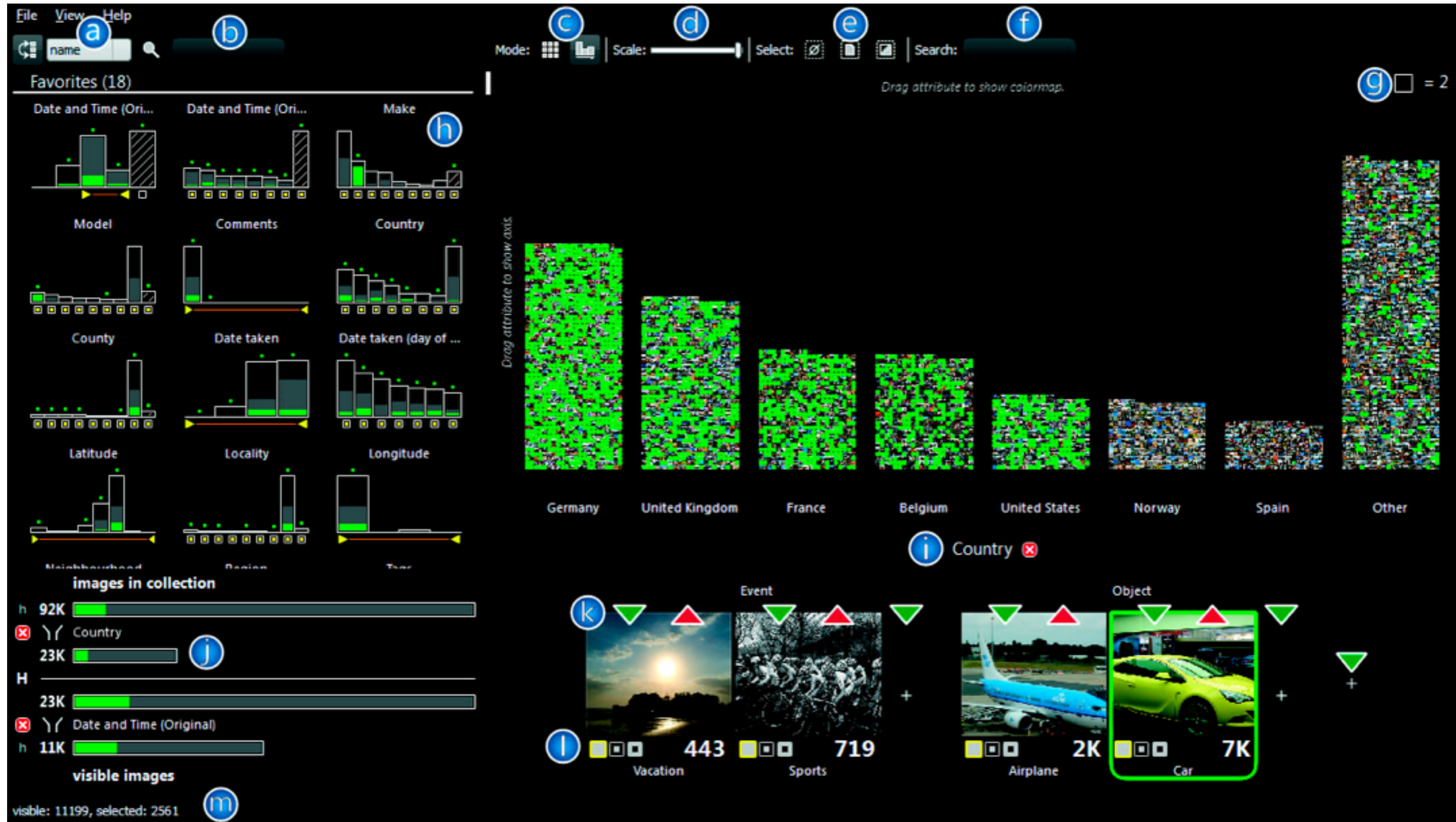


[Scented Widgets: Improving Navigation Cues with Embedded Visualizations. Willett, Heer, and Agrawala. IEEE TVCG (Proc. InfoVis 2007) 13:6 (2007), 1129–1136.]



[Multivariate Network Exploration and Presentation: From Detail to Overview via Selections and Aggregations. van den Elzen, van Wijk, IEEE TVCG 20(12): 2014 (Proc. InfoVis 2014).]

Scented histogram bisliders: detailed

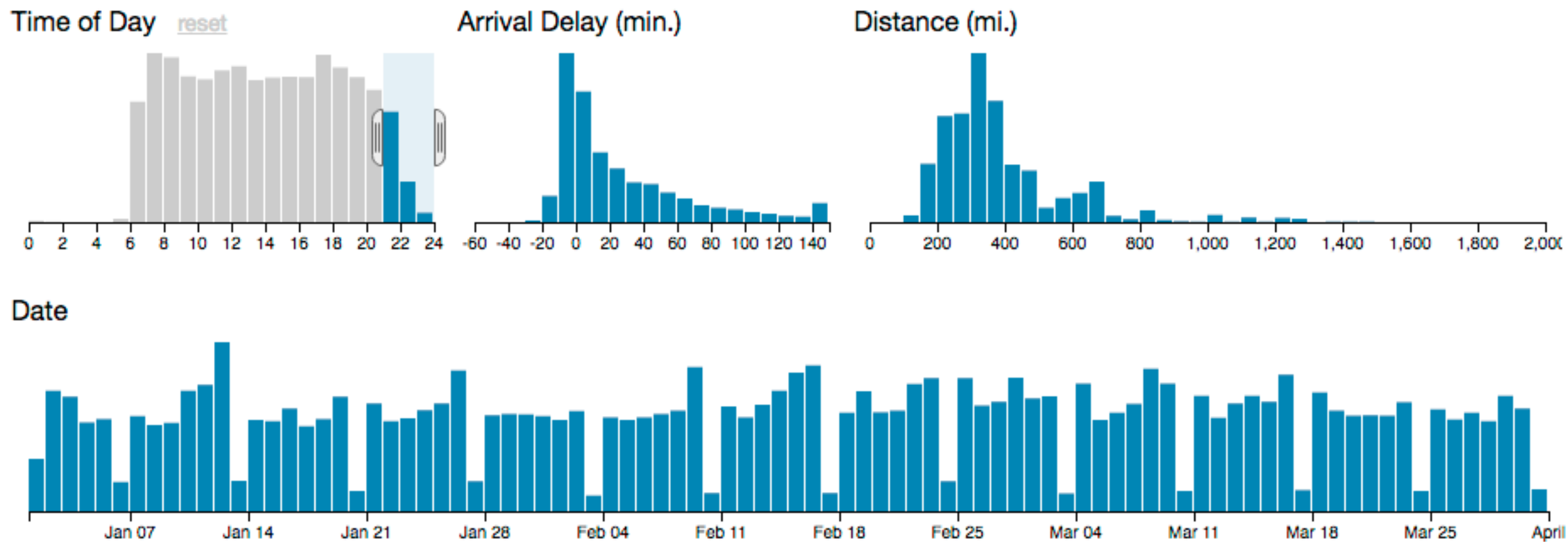


[ICLIC: Interactive categorization of large image collections. van der Corput and van Wijk. Proc. PacificVis 2016.]

Idiom: **cross filtering**

System: **Crossfilter**

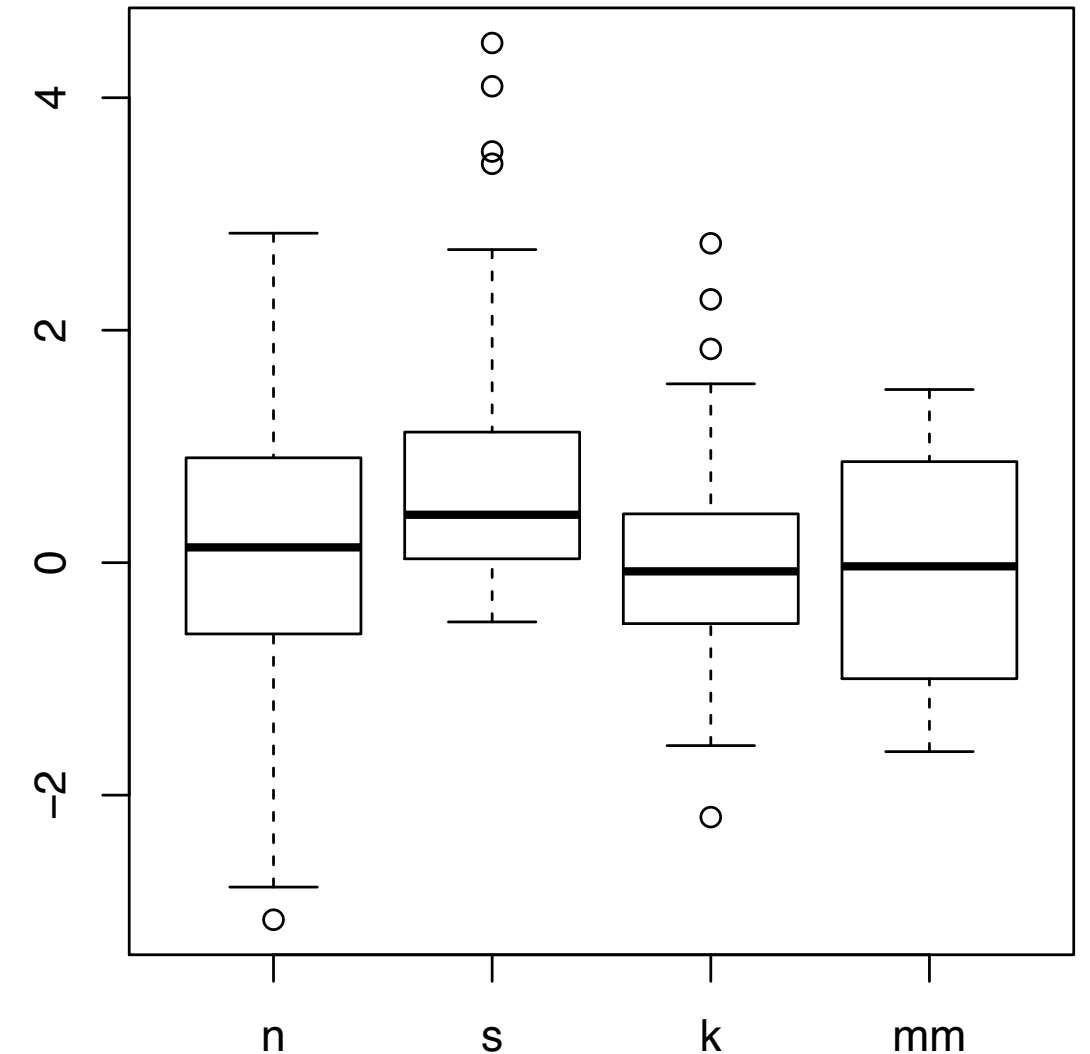
- item filtering
- coordinated views/controls combined
 - all scented histogram bisliders update when any ranges change



[\[http://square.github.io/crossfilter/\]](http://square.github.io/crossfilter/)

Idiom: **boxplot**

- static item aggregation
- task: find distribution
- data: table
- derived data
 - 5 quant attribs
 - median: central line
 - lower and upper quartile: boxes
 - lower upper fences: whiskers
 - values beyond which items are outliers
 - outliers beyond fence cutoffs explicitly shown



[40 years of boxplots. Wickham and Stryjewski. 2012. had.co.nz]

Spatial aggregation

- MAUP: Modifiable Areal Unit Problem

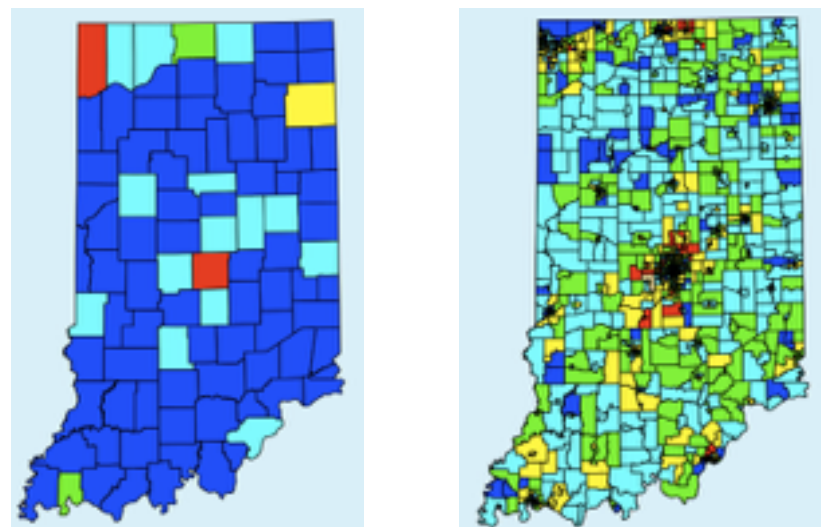
- gerrymandering (manipulating voting district boundaries) is only one example!

- zone effects



[http://www.e-education.psu.edu/geog486/l4_p7.html, Fig 4.cg.6]

- scale effects



<https://blog.cartographica.com/blog/2011/5/19/the-modifiable-areal-unit-problem-in-gis.html>

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
 - Chap 13: Reduce Items and Attributes*
- *Hierarchical Aggregation for Information Visualization: Overview, Techniques and Design Guidelines*. Elmqvist and Fekete. IEEE Transactions on Visualization and Computer Graphics 16:3 (2010), 439–454.
- *A Review of Overview+Detail, Zooming, and Focus+Context Interfaces*. Cockburn, Karlson, and Bederson. ACM Computing Surveys 41:1 (2008), 1–31.
- *A Guide to Visual Multi-Level Interface Design From Synthesis of Empirical Study Evidence*. Lam and Munzner. Synthesis Lectures on Visualization Series, Morgan Claypool, 2010.

More Information

[@tamaramunzner](https://twitter.com/tamaramunzner)

- this talk

<http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft>

- book page (including tutorial lecture slides)

<http://www.cs.ubc.ca/~tmm/vadbook>

– 20% promo code for book+ebook combo:
HVN17

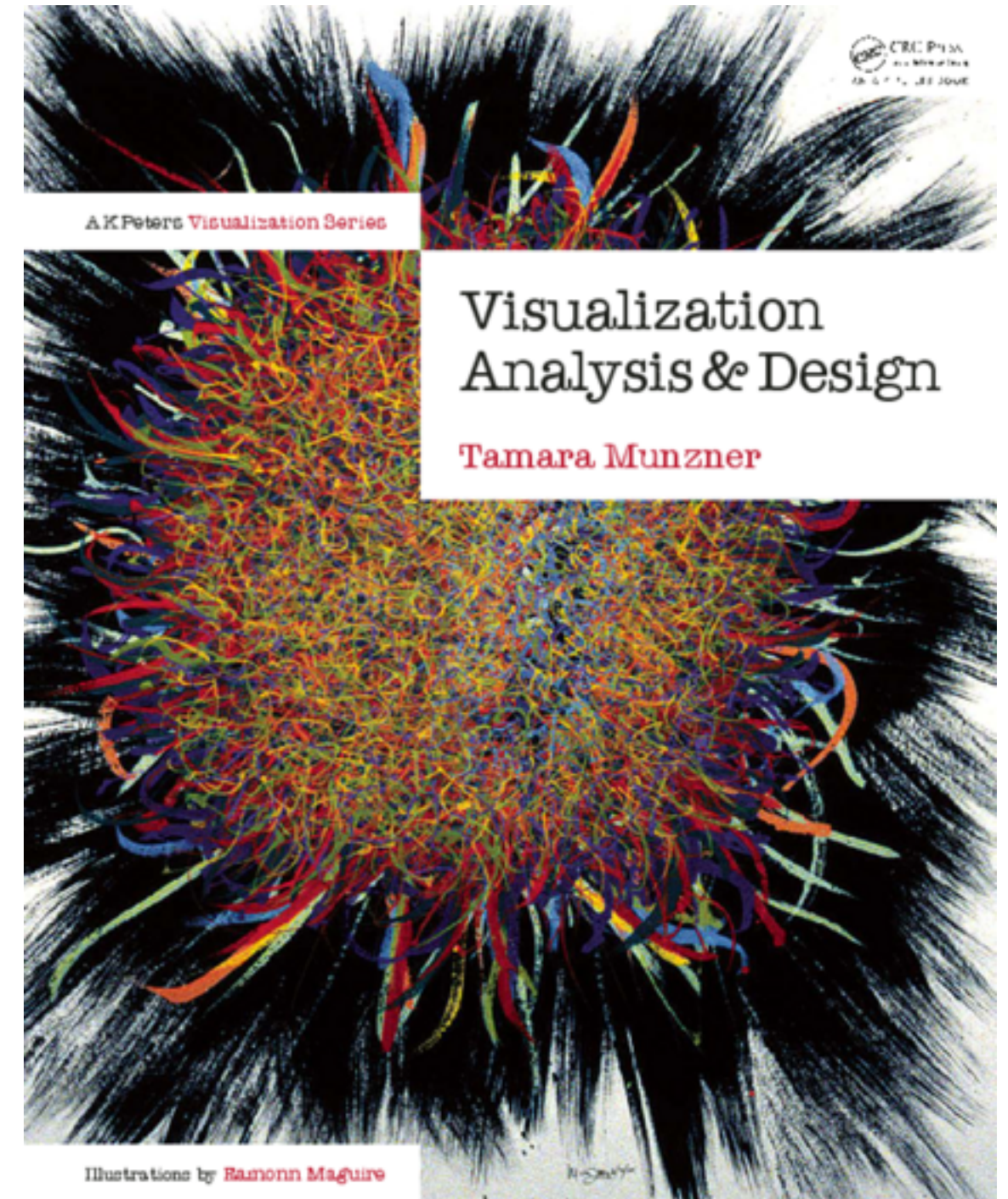
– <http://www.crcpress.com/product/isbn/9781466508910>

– illustrations: Eamonn Maguire

- papers, videos, software, talks, courses

<http://www.cs.ubc.ca/group/infovis>

<http://www.cs.ubc.ca/~tmm>



Visualization Analysis and Design.
Munzner. A K Peters Visualization Series, CRC Press, Visualization Series, 2014.