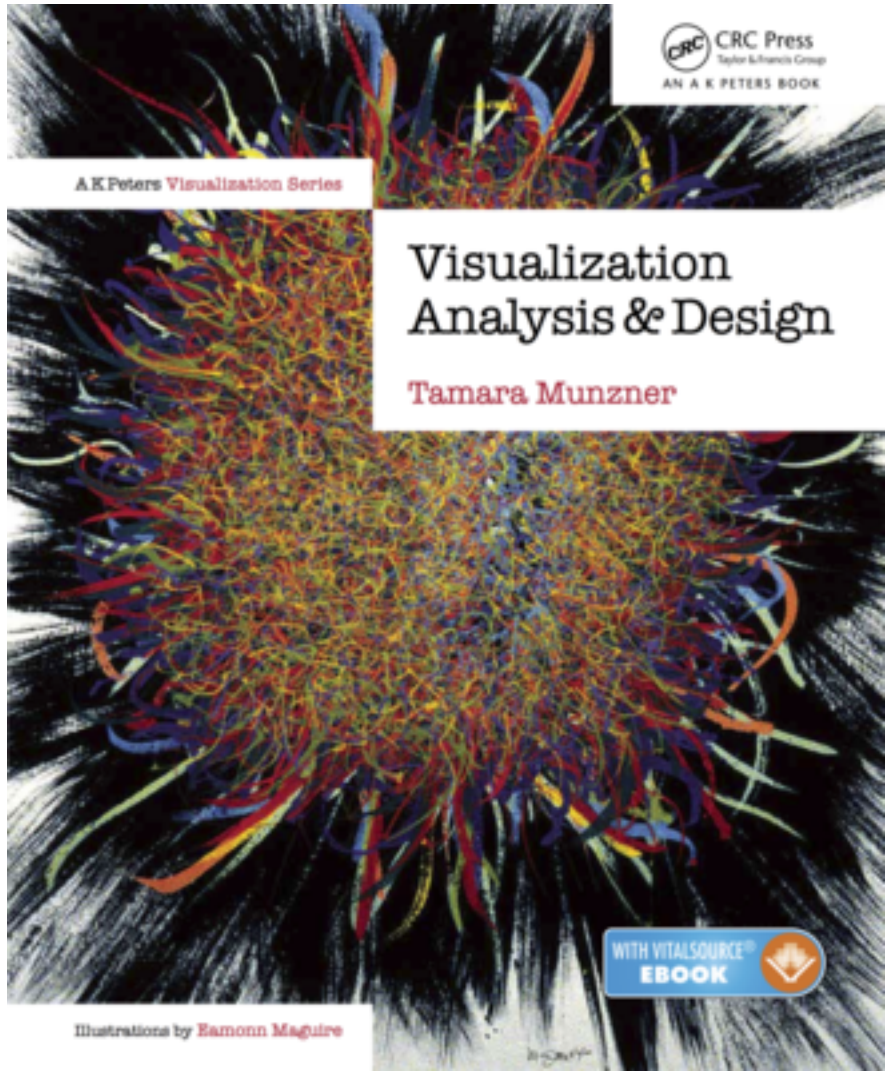
Visualization Analysis & Design
All Book/Teaching Slides



Tamara Munzner
Department of Computer Science
University of British Columbia

All Book/Teaching Slides
Last change: 30 Jan 2019

www.cs.ubc.ca/~tmm/talks.html#vadallslides

@tamaramunzner

Contents

Ch 1. What's Vis, and Why Do It?

Visualization (vis) defined & motivated

short version: alternate to next 3 slides

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- human in the loop needs the details
  - doesn't know exactly what questions to ask in advance
  - longterm exploratory analysis
    - ***speed up*** through human-in-the-loop visual data analysis
  - presentation of known results
  - stepping stone towards automation: refining, trustbuilding
- intended task, measurable definitions of effectiveness

Defining visualization (vis)

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

Why?...

Visualization (vis) defined & motivated

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

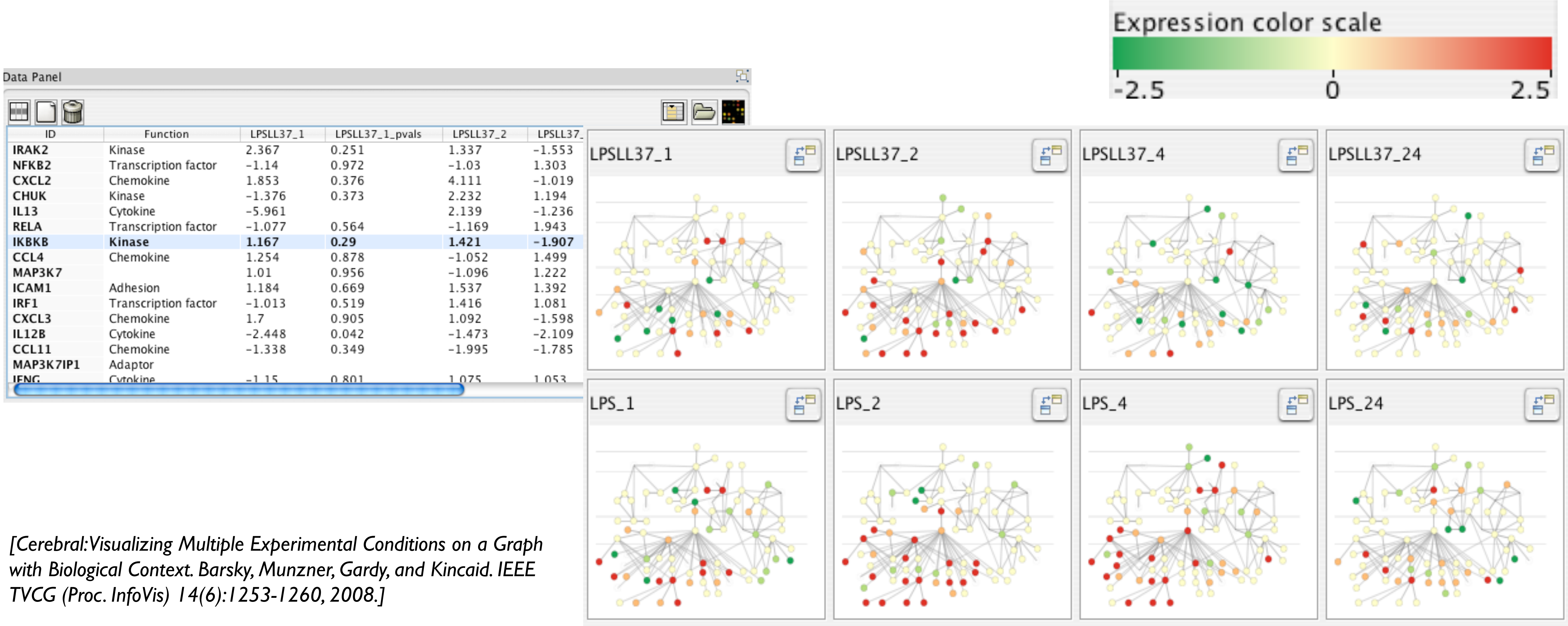
Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- human in the loop needs the details & no trusted automatic solution exists
  - doesn't know exactly what questions to ask in advance
  - exploratory data analysis
    - ***speed up*** through human-in-the-loop visual data analysis
  - present known results to others
  - stepping stone towards automation
    - before model creation to provide understanding
    - during algorithm creation to refine, debug, set parameters
    - before or during deployment to build trust and monitor

Why use an external representation?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- external representation: replace cognition with perception



[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Kincaid. IEEE TVCG (Proc. InfoVis) 14(6):1253-1260, 2008.]

Why depend on vision?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- human visual system is high-bandwidth channel to brain
  - overview possible due to background processing
    - subjective experience of seeing everything simultaneously
    - significant processing occurs in parallel and pre-attentively
- sound: lower bandwidth and different semantics
  - overview not supported
    - subjective experience of sequential stream
- touch/haptics: impoverished record/replay capacity
  - only very low-bandwidth communication thus far
- taste, smell: no viable record/replay devices

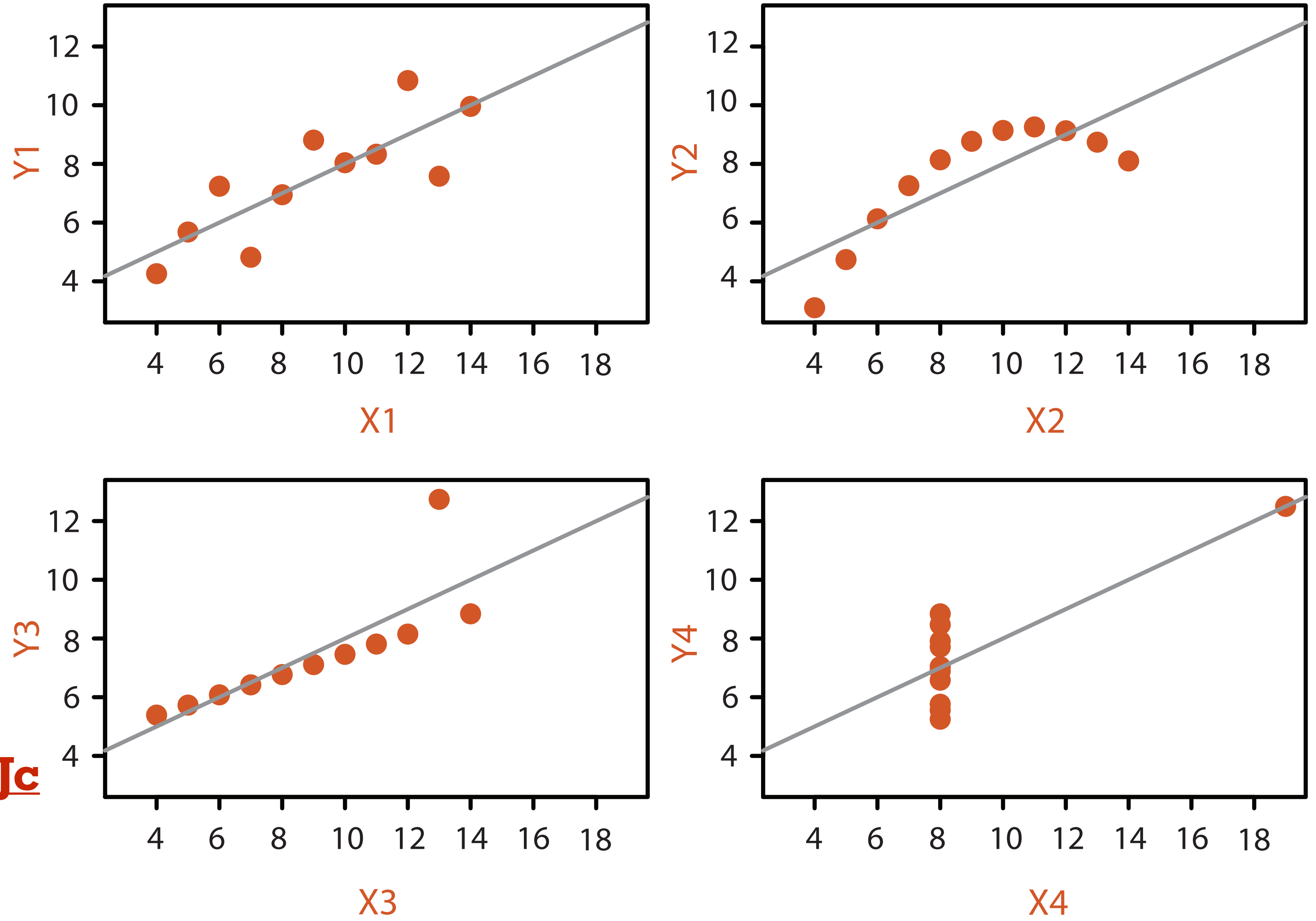
Why represent all the data?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- summaries lose information, details matter
  - confirm expected and find unexpected patterns
  - assess validity of statistical model

Anscombe's Quartet

Identical statistics	
x mean	9
x variance	10
y mean	7.5
y variance	3.75
x/y correlation	0.816



https://www.youtube.com/watch?v=DbJyPELmhJc

Same Stats, Different Graphs

Why focus on tasks and effectiveness?

Computer-based visualization systems provide visual representations of $\boxed{\text{datasets}}$ designed to help people carry out $\boxed{\text{tasks}}$ more $\boxed{\text{effectively}}$

- effectiveness requires match between data/task and representation
  - set of representations is huge
  - many are ineffective mismatch for specific data/task combo
  - increases chance of finding good solutions if you understand full space of possibilities
- what counts as effective?
  - novel: enable entirely new kinds of analysis
  - faster: speed up existing workflows
- how to validate effectiveness
  - many methods, must pick appropriate one for your context

What resource limitations are we faced with?

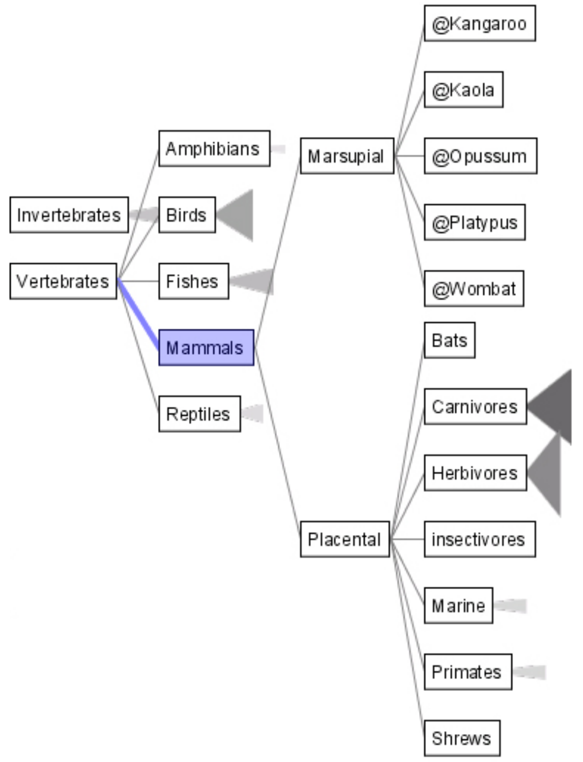
Vis designers must take into account three very different kinds of resource limitations: those of computers, of humans, and of displays.

- computational limits
  - processing time
  - system memory
- human limits
  - human attention and memory
- display limits
  - pixels are precious resource, the most constrained resource
  - **information density**: ratio of space used to encode info vs unused whitespace
    - tradeoff between clutter and wasting space, find sweet spot between dense and sparse

Why analyze?

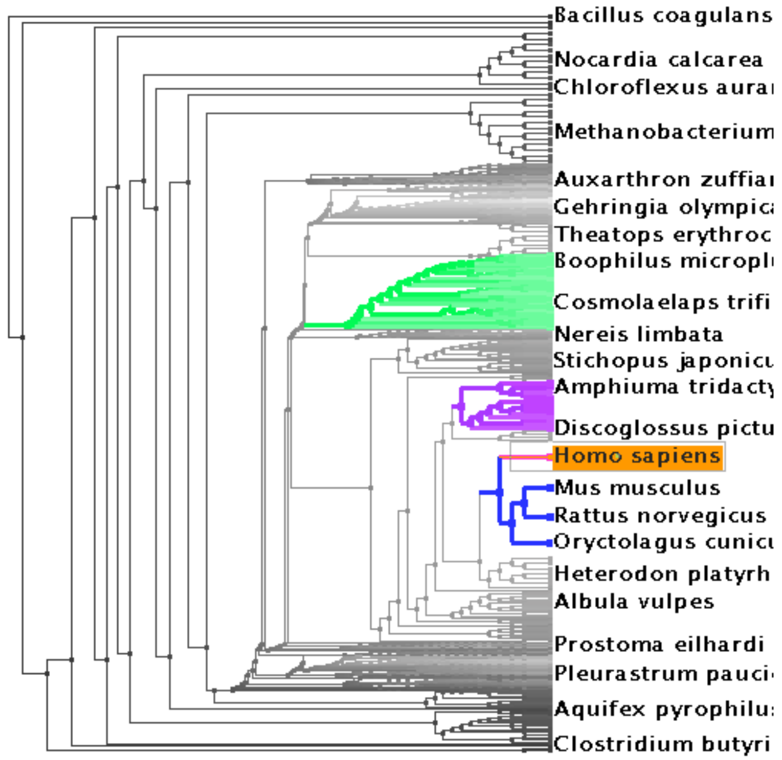
- imposes structure on huge design space
  - –scaffold to help you think systematically about choices
  - –analyzing existing as stepping stone to designing new
  - –most possibilities ineffective for particular task/data combination

SpaceTree

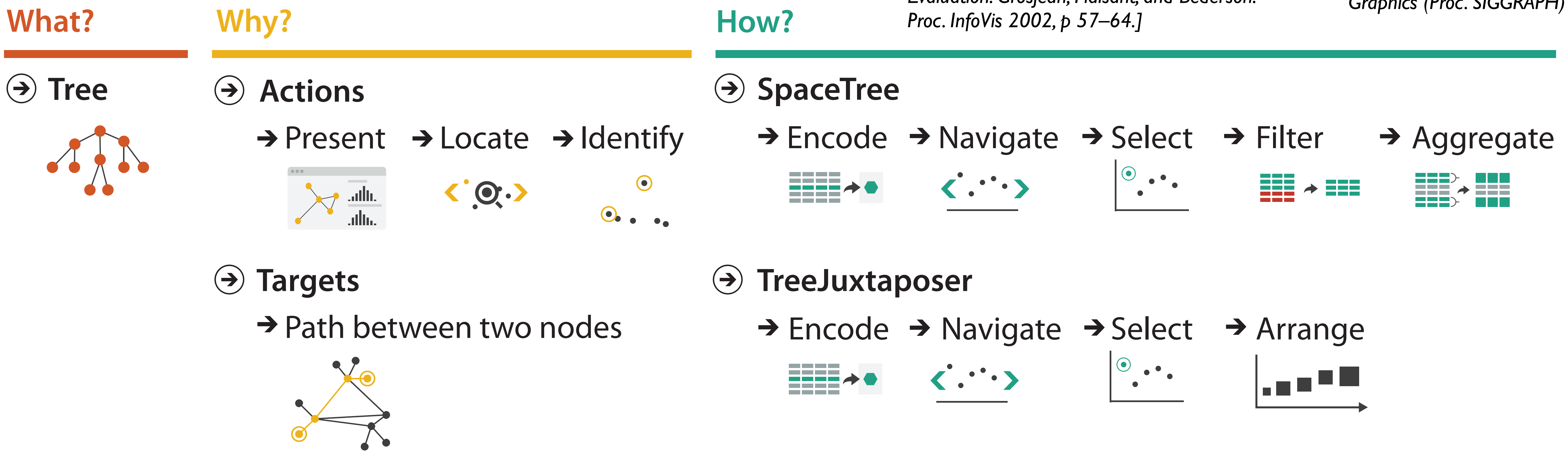


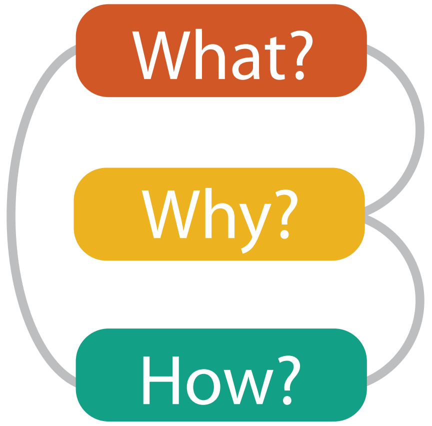
[SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation. Grosjean, Plaisant, and Bederson. Proc. InfoVis 2002, p 57–64.]

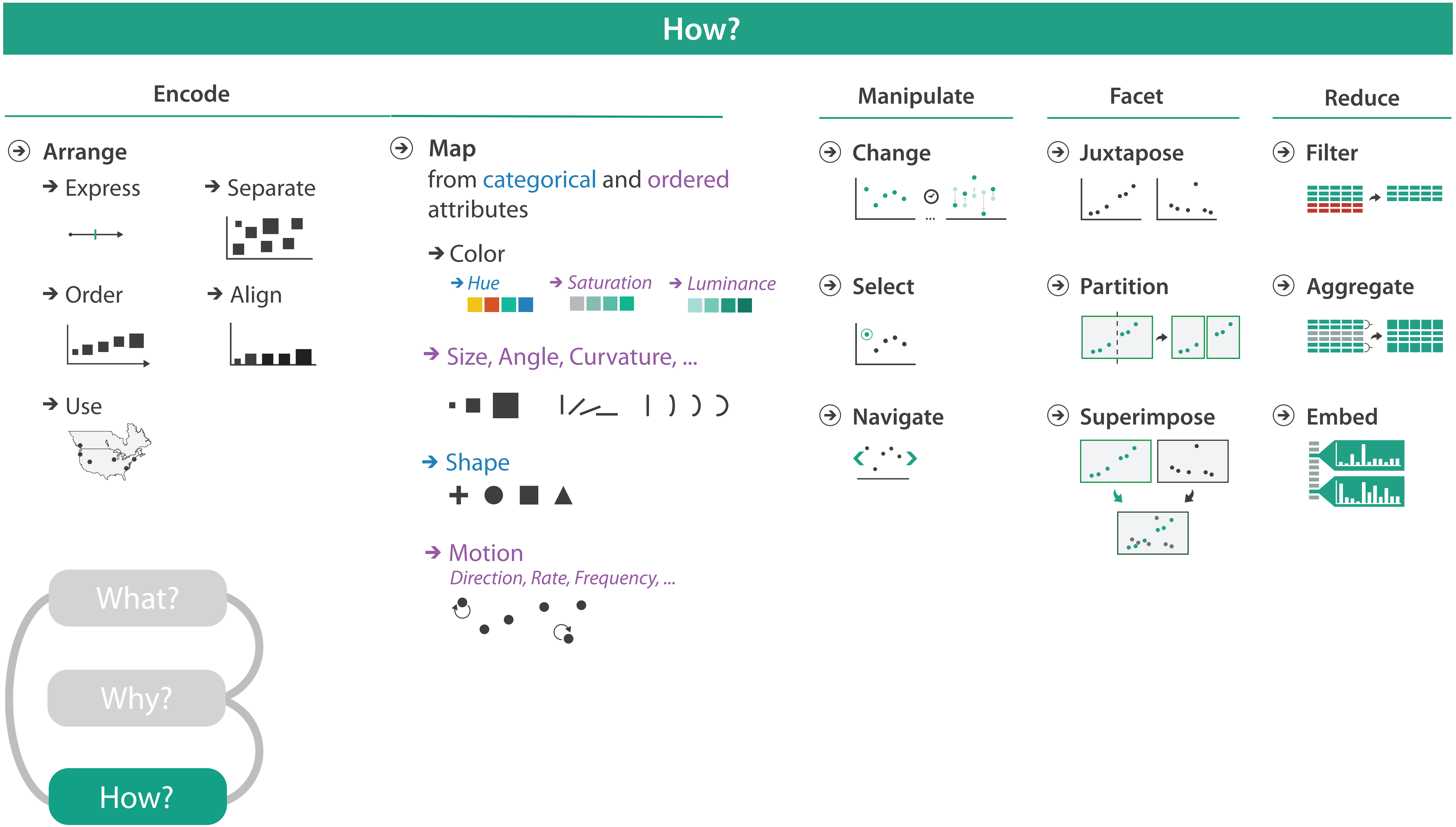
TreeJuxtaposer



[TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context With Guaranteed Visibility. ACM Trans. on Graphics (Proc. SIGGRAPH) 22:453– 462, 2003.]



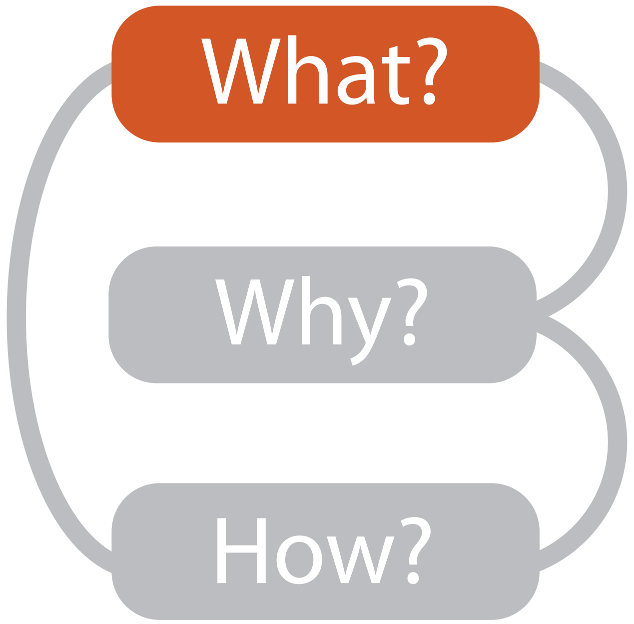



How?
Encode
Manipulate
Facet
Reduce
Arrange
Express
Separate
Order
Align
Use
Map
from categorical and ordered attributes
Color
Hue
Saturation
Luminance
Size, Angle, Curvature, ...
Shape
Motion
Direction, Rate, Frequency, ...
Change
Select
Navigate
Juxtapose
Partition
Superimpose
Filter
Aggregate
Embed
What?
Why?
How?

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
  - Chap 1: What's Vis, and Why Do It?
- The Nature of External Representations in Problem Solving. Jiajie Zhang. Cognitive Science 21:2 (1997), 179-217.
- A Representational Analysis of Numeration Systems. Jiajie Zhang and Donald A. Norman. Cognition 57 (1995), 271-295.
- Why a Diagram Is (Sometimes) Worth Ten Thousand Words.. Jill H. Larkin and Herbert A. Simon. Cognitive Science 11:1 (1987), 65-99.
- Graphs in Statistical Analysis.F.J. Anscombe. American Statistician 27 (1973), 17-21.
- Design Study Methodology: Reflections from the Trenches and the Stacks. Michael Sedlmair, Miriah Meyer, and Tamara Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2012), 18(12):2431-2440, 2012.
- Information Visualization: Perception for Design, 3rd edition, Colin Ware, Morgan Kaufmann, 2013.
- Current approaches to change blindness Daniel J. Simons. Visual Cognition 7, 1/2/3 (2000), 1-15.
- Semiology of Graphics, Jacques Bertin, Gauthier-Villars 1967, EHESS 1998
- The Visual Display of Quantitative Information. Edward R. Tufte. Graphics Press, 1983.

Ch 2. What: Data Abstraction



What?

Datasets

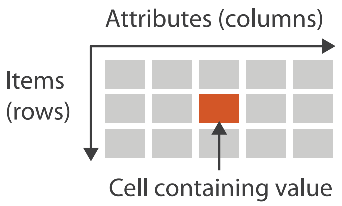
Data Types
- → Items
- → Attributes
- → Links
- → Positions
- → Grids

Data and Dataset Types

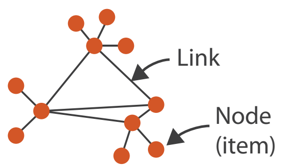
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

Dataset Types

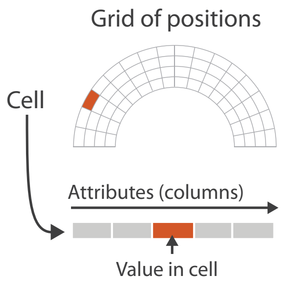
→ Tables



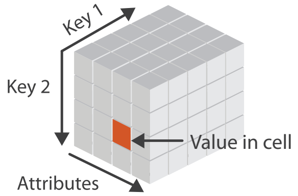
→ Networks



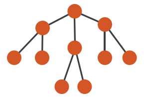
→ Fields (Continuous)



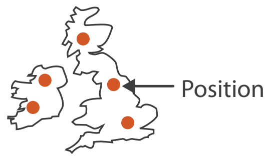
→ *Multidimensional Table*



→ *Trees*



→ Geometry (Spatial)



Dataset Availability

→ Static

→ Dynamic

Attributes

Attribute Types

→ Categorical

→ Ordered

→ *Ordinal*

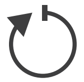
→ *Quantitative*

Ordering Direction

→ Sequential

→ Diverging

→ Cyclic

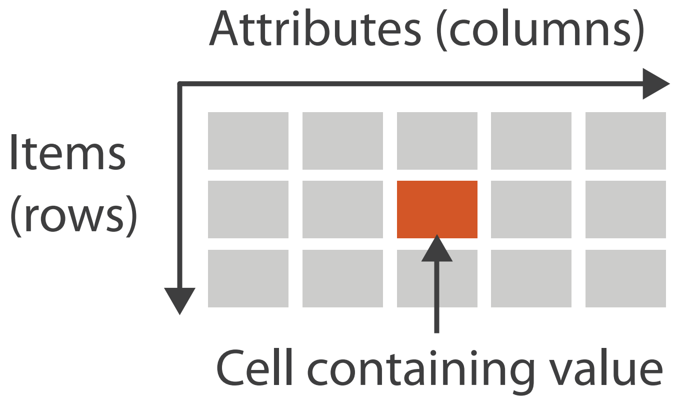


Types: Datasets and data

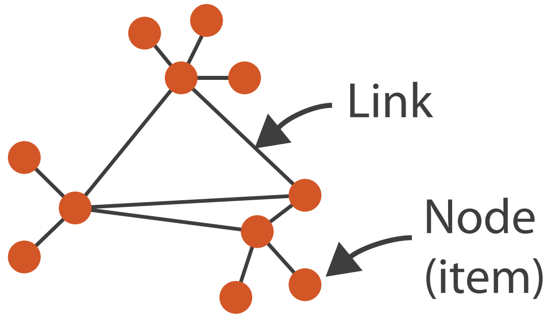
short version: alternate to next 3 slides

Dataset Types

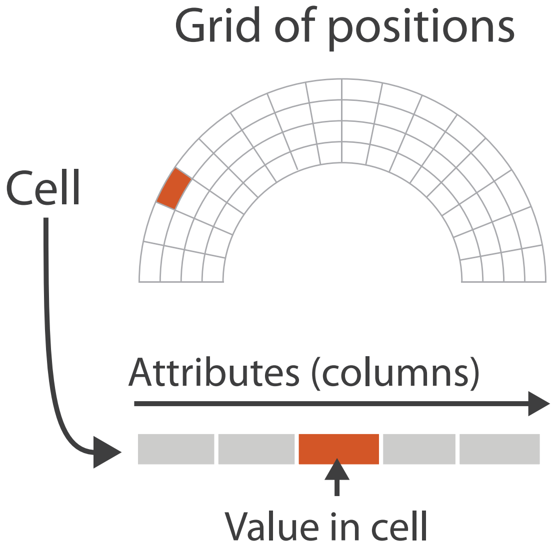
- Tables



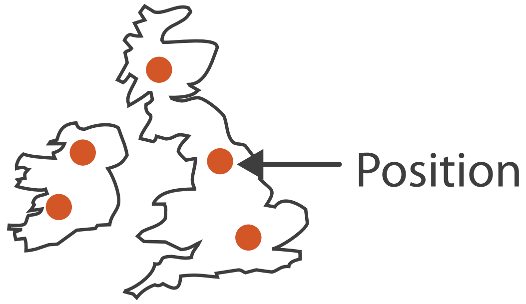
- Networks



- Spatial
  - Fields (Continuous)



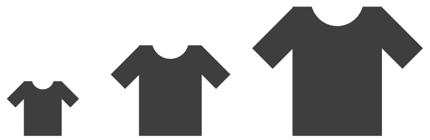
  - Geometry (Spatial)



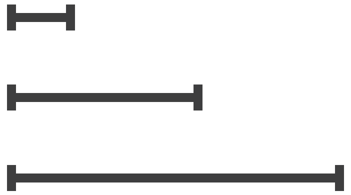
Attribute Types

- Categorical

- Ordered
  - *Ordinal*



  - *Quantitative*

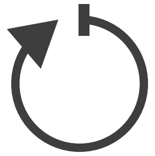


Ordering Direction

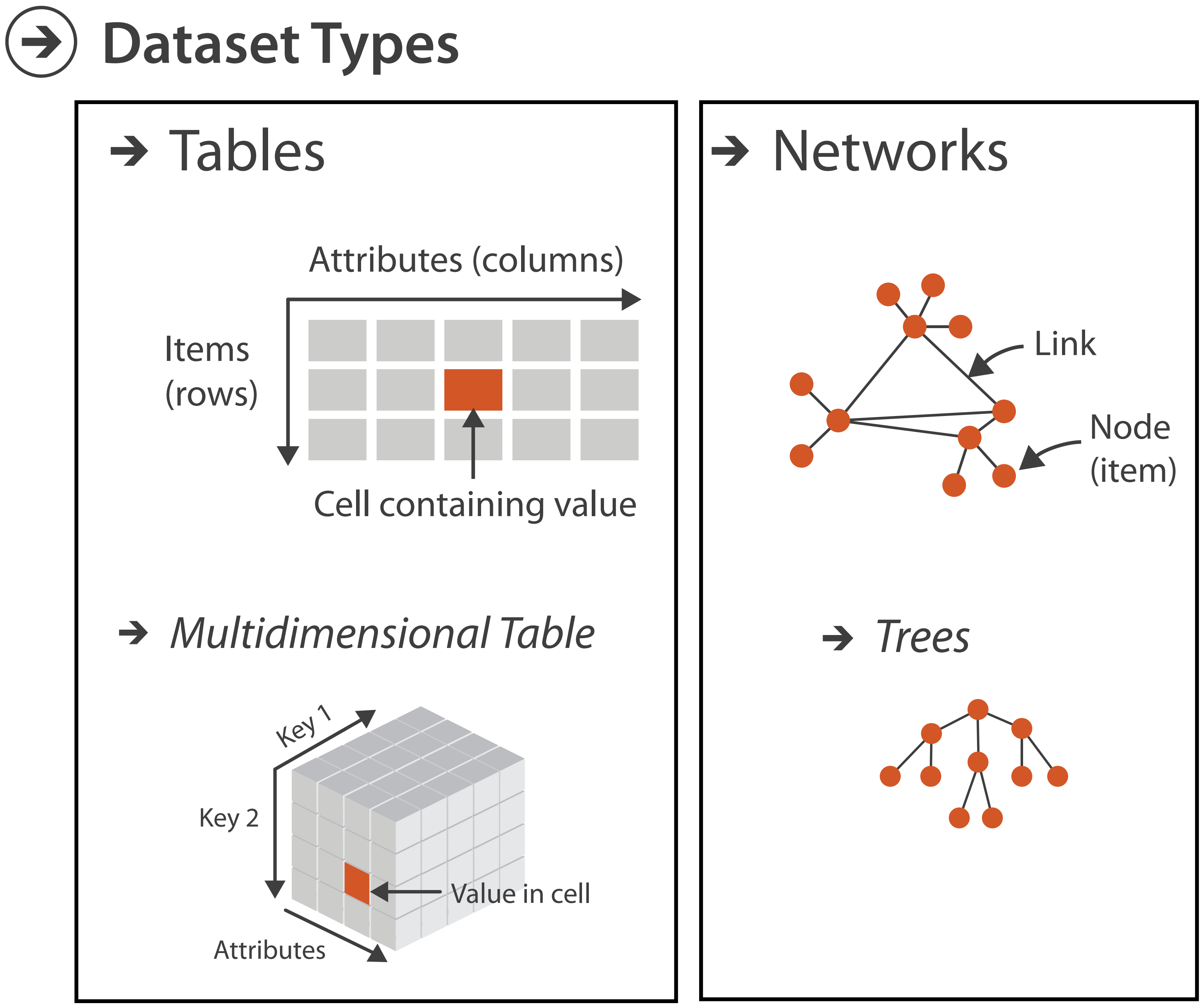
- Sequential

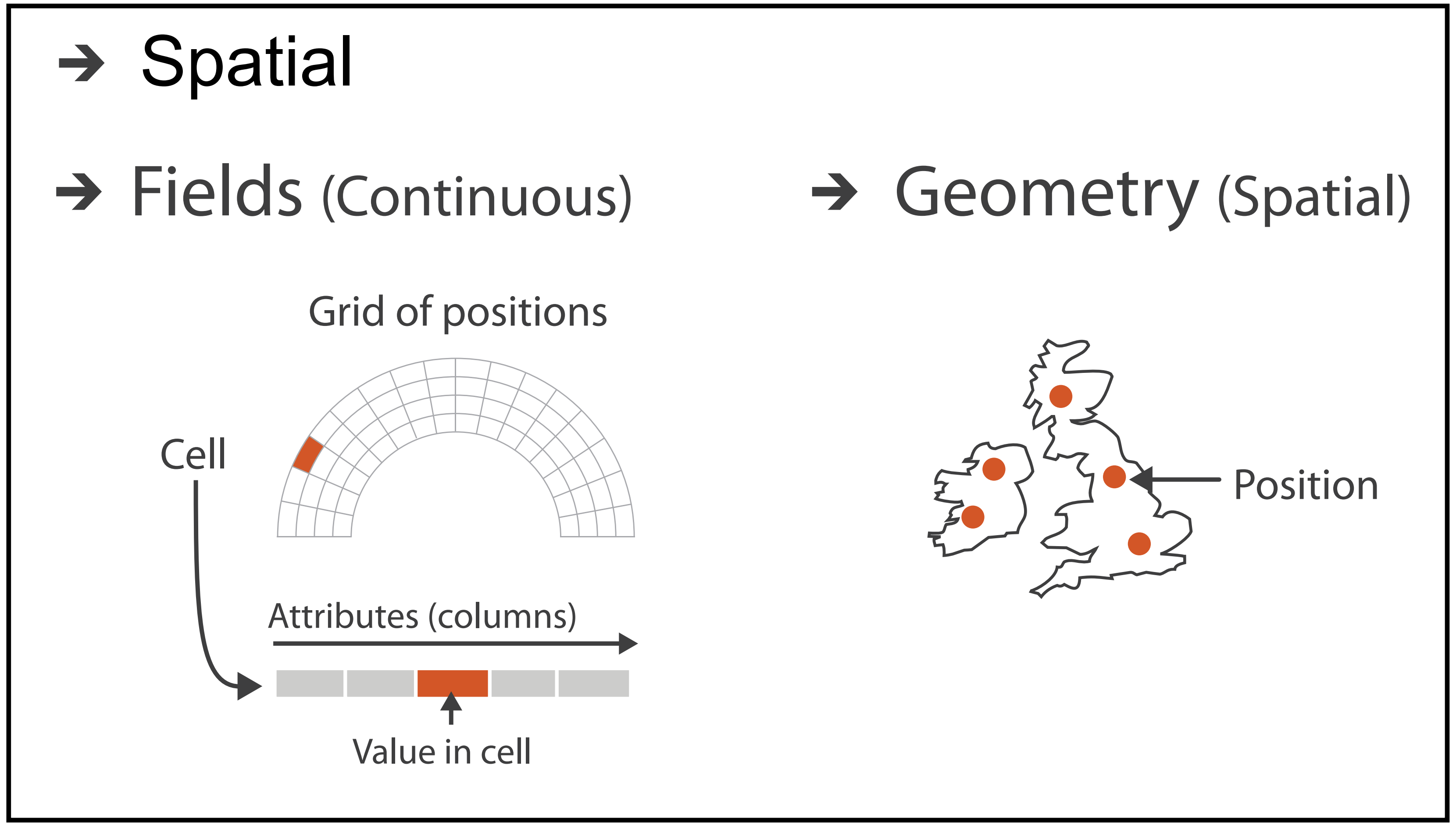
- Diverging

- Cyclic



Three major datatypes





- visualization vs computer graphics
  – geometry is design decision

Dataset and data types

➔ Data and Dataset Types

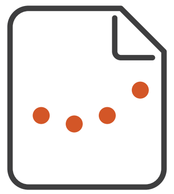
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

➔ Data Types

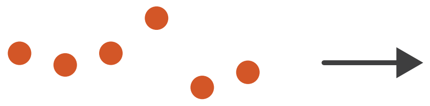
➔ Items ➔ Attributes ➔ Links ➔ Positions ➔ Grids

➔ Dataset Availability

➔ Static



➔ Dynamic

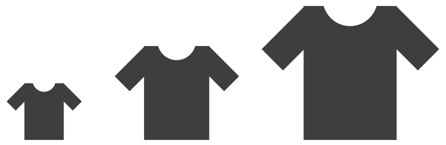


Attribute types

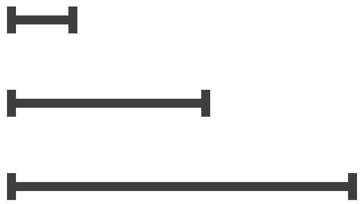
Attribute Types

- Categorical

- Ordered
  - *Ordinal*



  - *Quantitative*

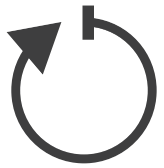


Ordering Direction

- Sequential

- Diverging

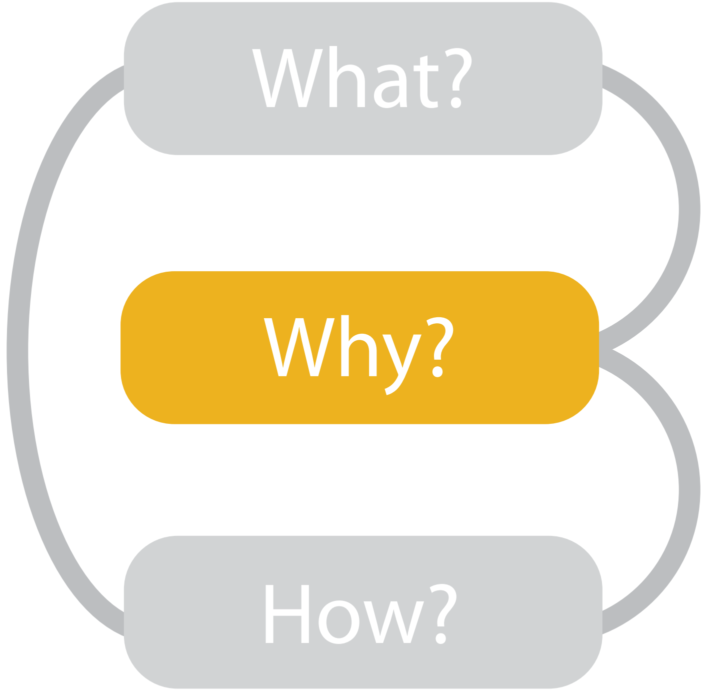
- Cyclic



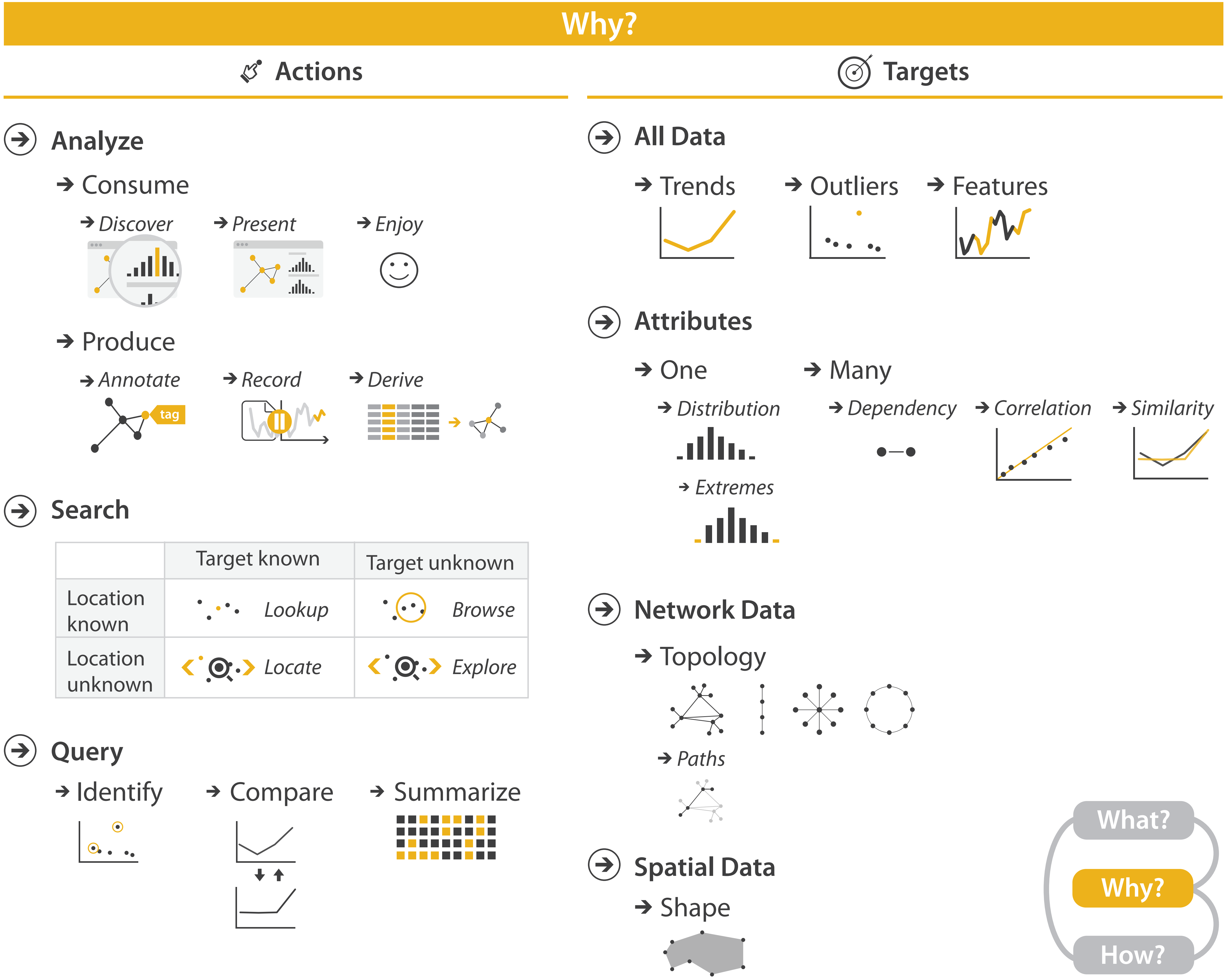
Further reading, full Ch 2

- Readings in Information Visualization: Using Vision To Think, Chapter 1. Stuart K. Card, Jock Mackinlay, and Ben Shneiderman. Morgan Kaufmann, 1999.
- Rethinking Visualization: A High-Level Taxonomy. InfoVis 2004, p 151-158, 2004.
- The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations Ben Shneiderman, Proc. 1996 IEEE Visual Languages
- Data Visualization: Principles and Practice, 2nd ed. Alexandru Telea, CRC Press, 2014.
- Interactive Data Visualization: Foundations, Techniques, and Applications, 2nd ed. Matthew O. Ward, Georges Grinstein, Daniel Keim. CRC Press, 2015.
- The Visualization Handbook. Charles Hansen and Chris Johnson, eds. Academic Press, 2004.
- Visualization Toolkit: An Object-Oriented Approach to 3D Graphics, 4th ed. Will Schroeder, Ken Martin, and Bill Lorensen. Kitware 2006.
- The Structure of the Information Visualization Design Space. Stuart Card and Jock Mackinlay, Proc. InfoVis 97.
- Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases (extended paper) Chris Stolte, Diane Tang and Pat Hanrahan. IEEE TVCG 8(1):52-65 2002.
- Visualization of Time-Oriented Data. Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, Chris Tominski. Springer 2011.

Ch 3. Why: Task Abstraction



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



Actions: Analyze, Query

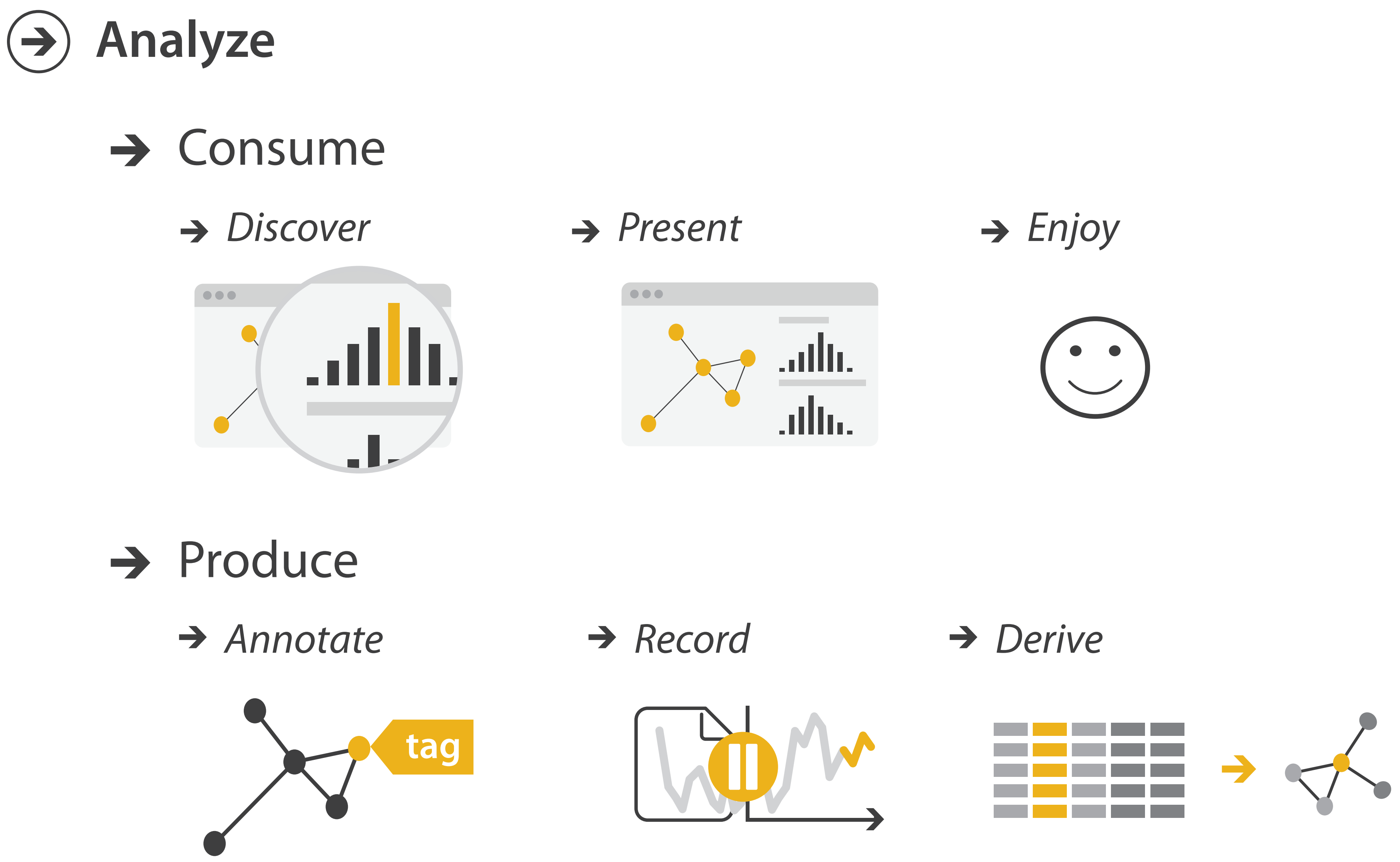
short version: alternate to next 4 slides

- analyze
  - consume
    - discover vs present
      - aka explore vs explain
    - enjoy
      - aka casual, social
  - produce
    - annotate, record, derive
- query
  - how much data matters?
    - one, some, all
- independent choices
  - analyze, query, (search)

Actions: Analyze

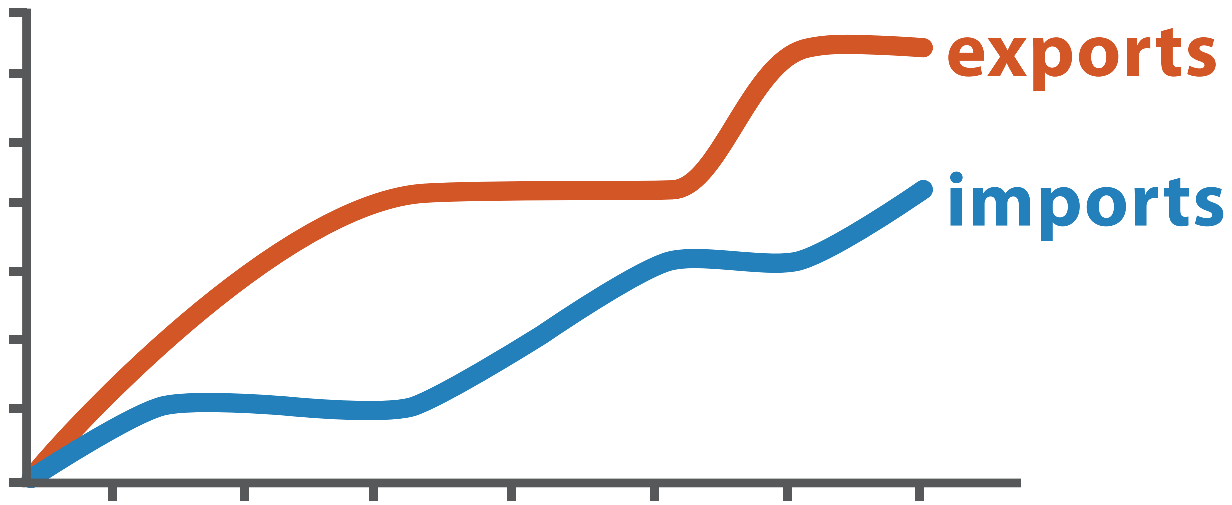
- consume
  - discover vs present
    - classic split
    - aka explore vs explain
  - enjoy
    - newcomer
    - aka casual, social

- produce
  - annotate, record
  - derive
    - crucial design choice

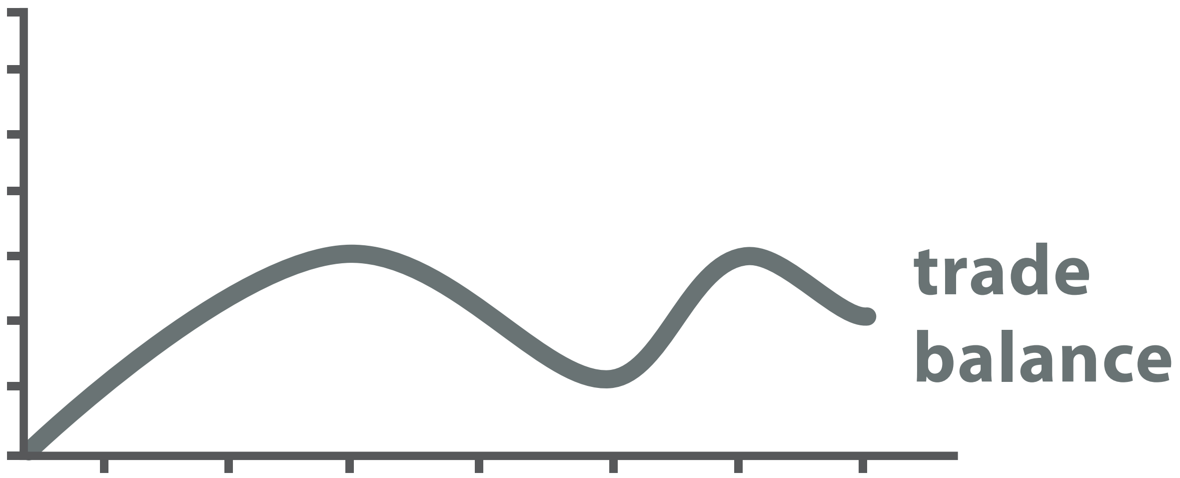


Derive

- don't just draw what you're given!
  - decide what the right thing to show is
  - create it with a series of transformations from the original dataset
  - draw that
- one of the four major strategies for handling complexity



Original Data



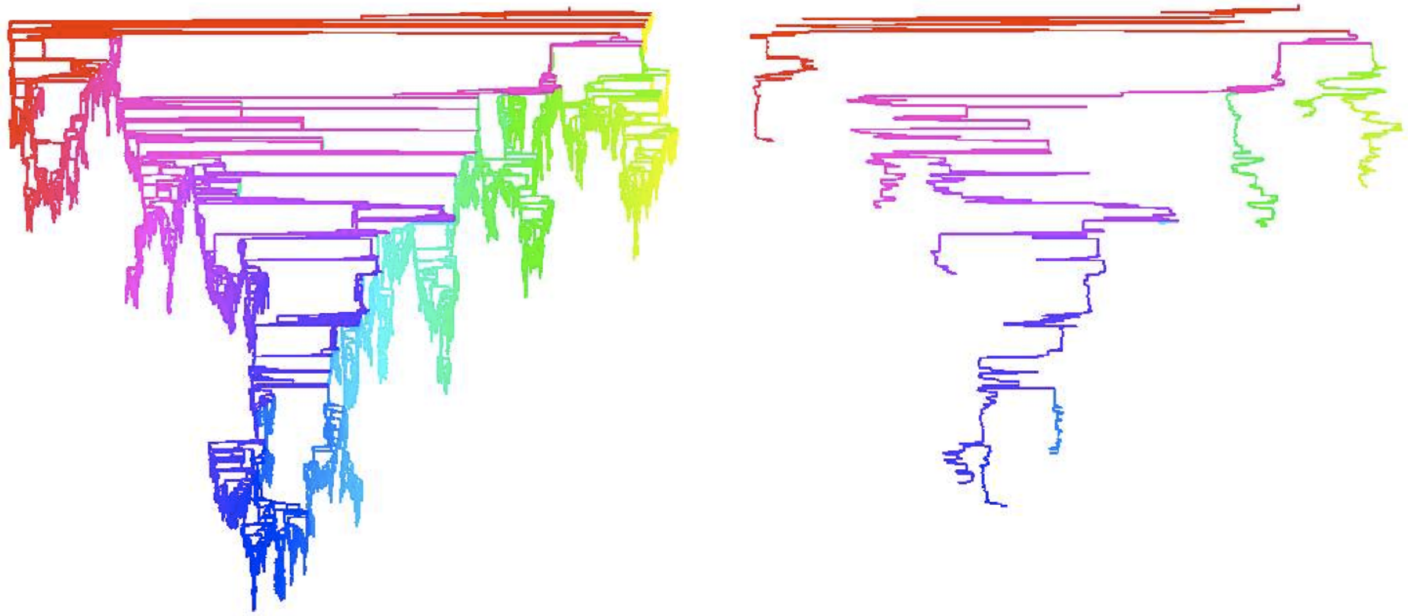
trade balance = exports − imports

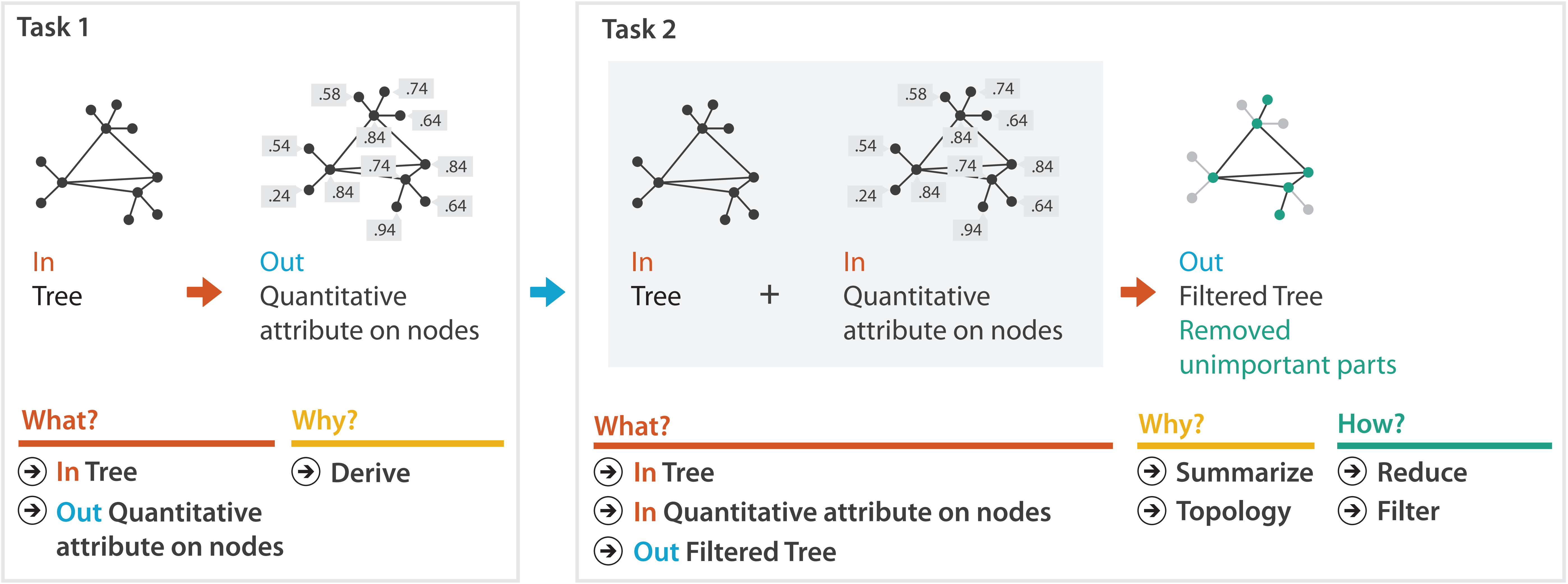
Derived Data

Analysis example: Derive one attribute

- Strahler number
  - centrality metric for trees/networks
  - derived quantitative attribute
  - draw top 5K of 500K for good skeleton

[Using Strahler numbers for real time visual exploration of huge graphs. Auber. Proc. Intl. Conf. Computer Vision and Graphics, pp. 56–69, 2002.]





Actions: Search, query

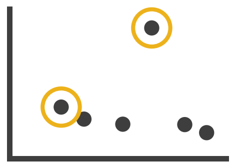
- what does user know?
  - target, location
- how much of the data matters?
  - one, some, all
- independent choices for each of these three levels
  - analyze, search, query
  - mix and match

Search

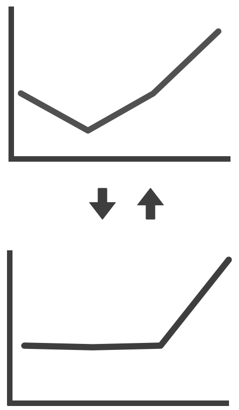
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Query

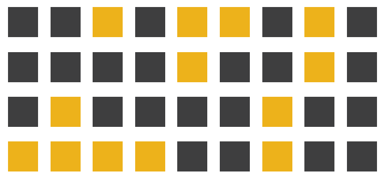
→ Identify



→ Compare



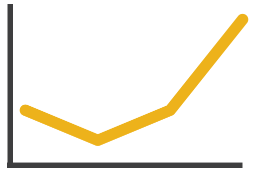
→ Summarize



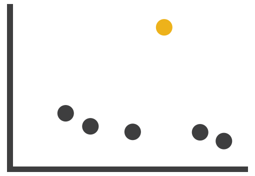
Why: Targets

All Data

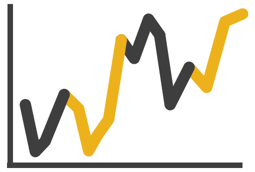
- Trends



- Outliers

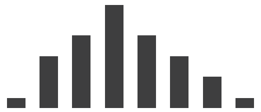


- Features

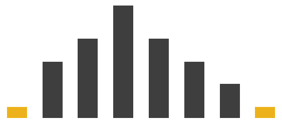


Attributes

- One
  - *Distribution*

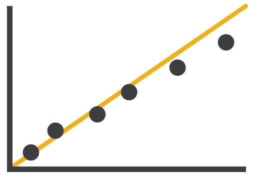


  - *Extremes*

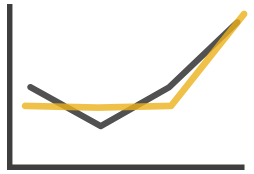


- Many
  - *Dependency*

  - *Correlation*

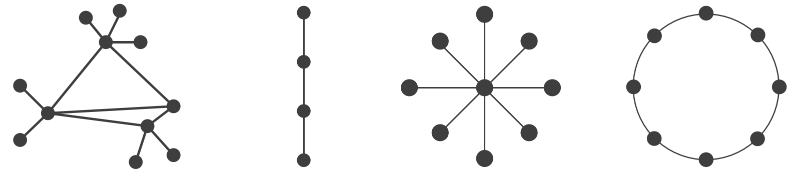


  - *Similarity*

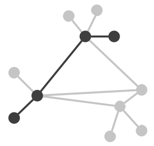


Network Data

- Topology

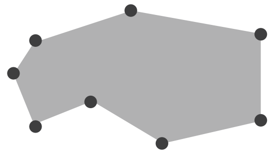


  - *Paths*



Spatial Data

- Shape



Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
  - *Chap 2: What: Data Abstraction*
  - *Chap 3: Why: Task Abstraction*
- *A Multi-Level Typology of Abstract Visualization Tasks*. Brehmer and Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis) 19:12 (2013), 2376–2385.
- *Low-Level Components of Analytic Activity in Information Visualization.* Amar, Eagan, and Stasko. Proc. IEEE InfoVis 2005, p 111–117.
- *A taxonomy of tools that support the fluent and flexible use of visualizations*. Heer and Shneiderman. Communications of the ACM 55:4 (2012), 45–54.
- *Rethinking Visualization: A High-Level Taxonomy*. Tory and Möller. Proc. IEEE InfoVis 2004, p 151–158.
- Visualization of Time-Oriented Data. Aigner, Miksch, Schumann, and Tominski. Springer, 2011.

Further reading, full Ch 3

- A Multi-Level Typology of Abstract Visualization Tasks.. Matthew Brehmer and Tamara Munzner. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 13) 19:12 (2013), 2376-2385.
- A characterization of the scientific data analysis process. Rebecca R. Springmeyer, Meera M. Blattner, and Nelson M. Max. Proc. Vis 1992, p 235-252.
- Low-Level Components of Analytic Activity in Information Visualization. Robert Amar, James Eagan, and John Stasko. Proc. InfoVis 05, pp. 111-117.
- Task taxonomy for graph visualization. Bongshin Lee, Catherine Plaisant, Cynthia Sims Parr, Jean-Daniel Fekete, and Nathalie Henry. Proc. BELIV 2006.
- Interactive Dynamics for Visual Analysis. Jeffrey Heer and Ben Shneiderman. Communications of the ACM, 55(4), pp. 45-54, 2012.
- What does the user want to see?: what do the data want to be? A. Johannes Pretorius and Jarke J. van Wijk. Information Visualization 8(3): 153-166, 2009.
- Chapter 1, Readings in Information Visualization: Using Vision to Think. Stuart Card, Jock Mackinlay, and Ben Shneiderman, Morgan Kaufmann 1999.
- An Operator Interaction Framework for Visualization Systems. Ed H. Chi and John T. Riedl. Proc. InfoVis 1998, p 63-70.
- Nominal, Ordinal, Interval, and Ratio Typologies are Misleading. Paul F. Velleman and Leland Wilkinson. The American Statistician 47(1):65-72, 1993.
- Rethinking Visualization: A High-Level Taxonomy. Melanie Tory and Torsten Möller, Proc. InfoVis 2004, pp. 151-158.
- SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation. Catherine Plaisant, Jesse Grosjean, and Ben B. Bederson. Proc. InfoVis 2002.
- TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility Tamara Munzner, Francois Guimbretiere, Serdar Tasiran, Li Zhang, and Yunhong Zhou. SIGGRAPH 2003.
- Feature detection in linked derived spaces. Chris Henze. Proc. Visualization (Vis) 1998, p 87-94.
- Using Strahler numbers for real time visual exploration of huge graphs. David Auber. Intl Conf. Computer Vision and Graphics, 2002, p 56-69.

Ch 4. Analysis: Four Levels for Validation

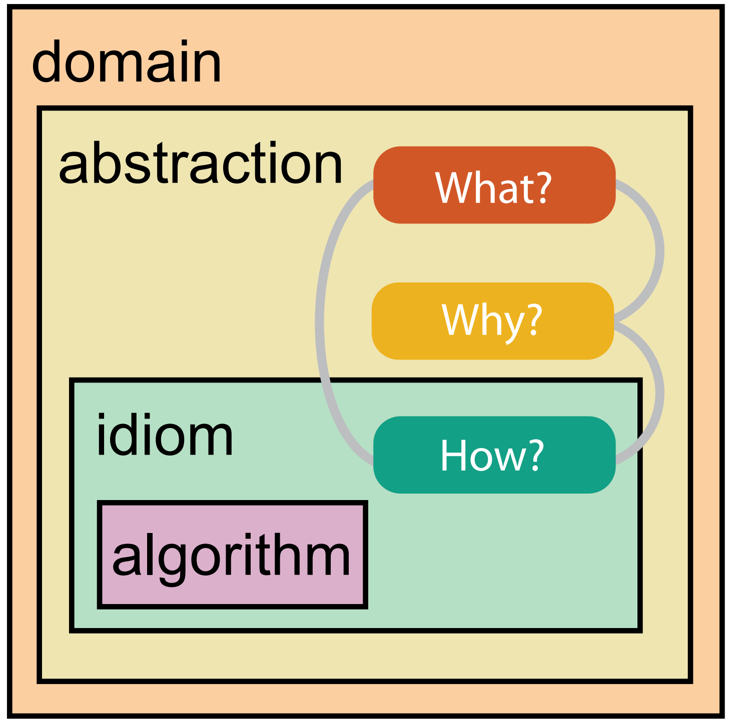
How to evaluate a visualization: So many methods, how to pick?

- Computational benchmarks?
  - quant: system performance, memory
- User study in lab setting?
  - quant: (human) time and error rates, preferences
  - qual: behavior/strategy observations
- Field study of deployed system?
  - quant: usage logs
  - qual: interviews with users, case studies, observations
- Analysis of results?
  - quant: metrics computed on result images
  - qual: consider what structure is visible in result images
- Justification of choices?
  - qual: perceptual principles, best practices

Nested model: Four levels of visualization design

[A Nested Model of Visualization Design and Validation.
Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

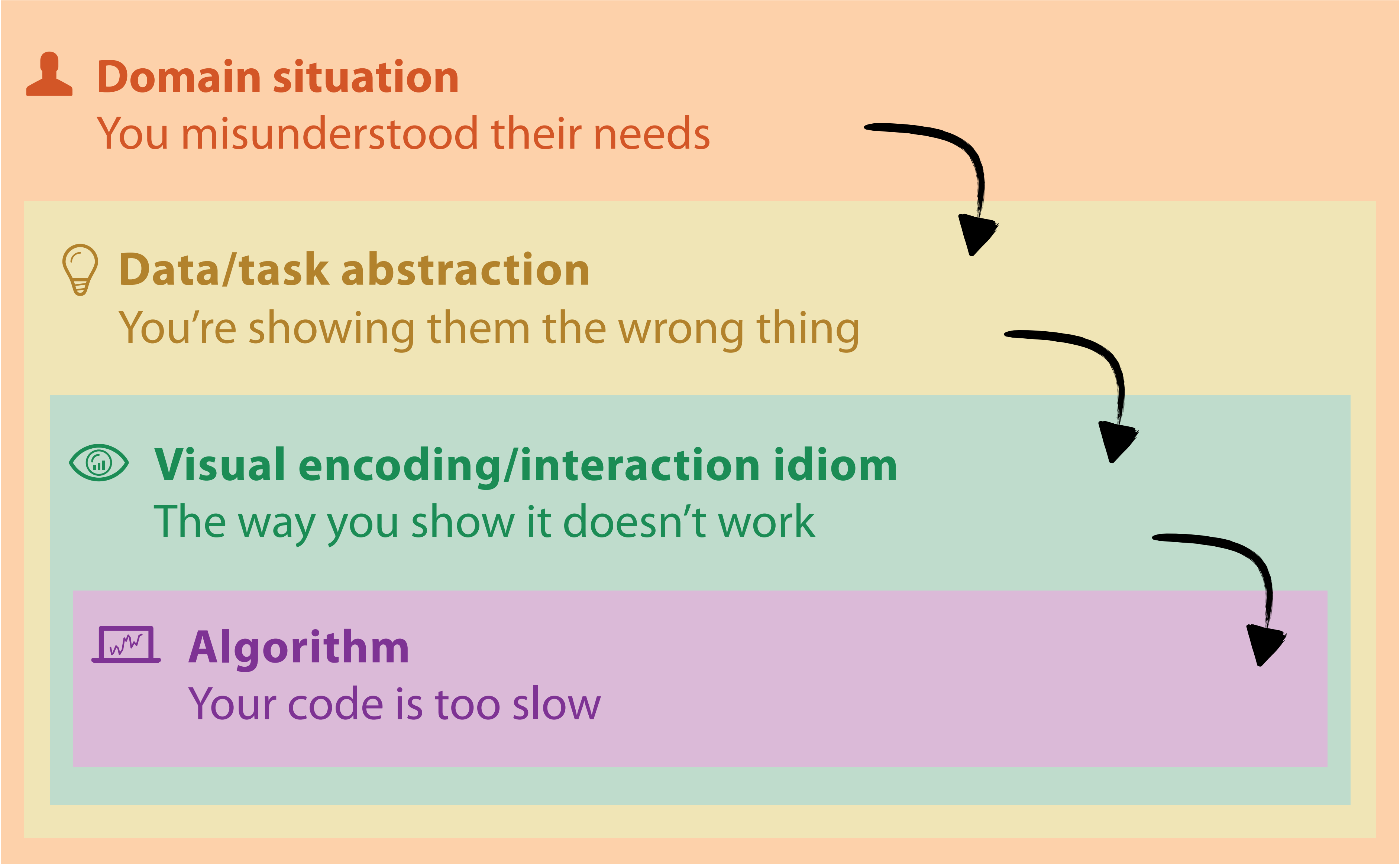
- *domain situation*
  - who are the target users?
- *abstraction*
  - translate from specifics of domain to vocabulary of visualization
    - **what** is shown? **data** abstraction
    - **why** is the user looking at it? **task** abstraction
- *idiom*
  - **how** is it shown?
    - **visual encoding** idiom: how to draw
    - **interaction** idiom: how to manipulate
- *algorithm*
  - efficient computation



[A Multi-Level Typology of Abstract Visualization Tasks
Brehmer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

Different threats to validity at each level

- cascading effects downstream



Interdisciplinary: need methods from different fields at each level

- mix of qual and quant approaches (typically)

anthropology/ ethnography

Domain situation
Observe target users using existing tools — qual

Data/task abstraction

design

Visual encoding/interaction idiom
Justify design with respect to alternatives — qual

computer science

Algorithm
Measure system time/memory
Analyze computational complexity — **quant**

psychology

Analyze results qualitatively — qual
Measure human time with lab experiment (*lab study*) — **quant**

anthropology/ ethnography

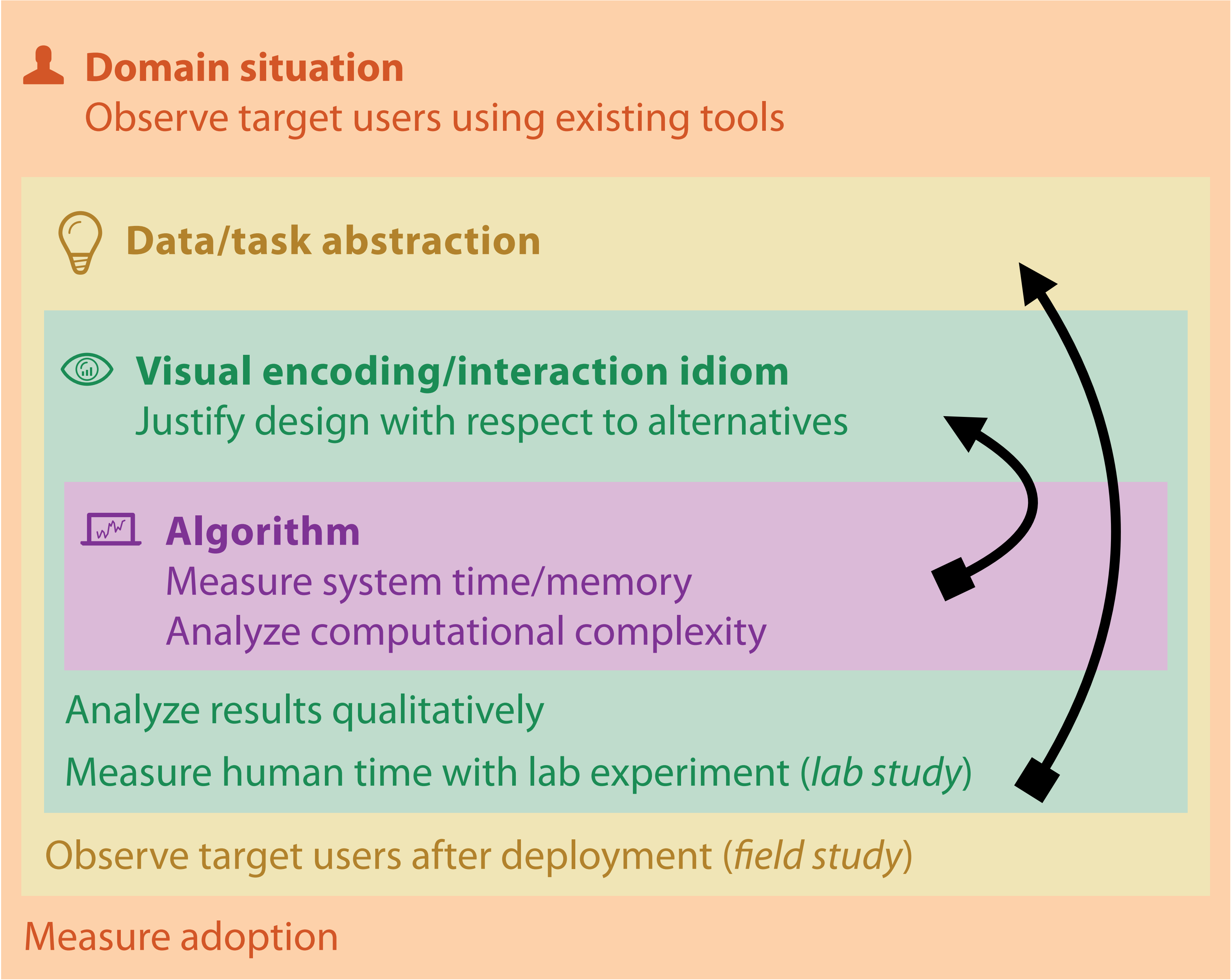
Observe target users after deployment (*field study*) — qual
Measure adoption — **quant**

problem-driven work

technique-driven work

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Mismatches: Common problem



benchmarks can't confirm design

lab studies can't confirm task abstraction

[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Analysis examples: Single paper includes only subset of methods

MatrixExplorer. Henry and Fekete. InfoVis 2006.

- observe and interview target users
- justify encoding/interaction design
- measure system time/memory
- qualitative result image analysis

LiveRAC. McLachlan, Munzner, Koutsofios, and North. CHI 2008.

- observe and interview target users
- justify encoding/interaction design
- qualitative result image analysis
- field study, document deployed usage

An energy model for visual graph clustering. (LinLog) Noack. Graph Drawing 2003

- qualitative/quantitative image analysis

Effectiveness of animation in trend visualization. Robertson et al. InfoVis 2008.

- lab study, measure time/errors for operation

Interactive visualization of genealogical graphs. McGuffin and Balakrishnan. InfoVis 2005.

- justify encoding/interaction design
- qualitative result image analysis
- test on target users, get utility anecdotes

Flow map layout. Phan et al. InfoVis 2005.

- justify encoding/interaction design
- computational complexity analysis
- measure system time/memory
- qualitative result image analysis

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
  - *Chap 4: Analysis: Four Levels for Validation*
- Storks Deliver Babies (p= 0.008). Robert Matthews. Teaching Statistics 22(2):36-38, 2000.
- The Earth is spherical ($p < 0.05$): alternative methods of statistical inference. Kim J. Vicente and Gerard L. Torenvliet. Theoretical Issues in Ergonomics Science, 1(3):248-271, 2000.
- The Prospects for Psychological Science in Human-Computer Interaction. Allen Newell and Stuart K. Card. Journal Human-Computer Interaction 1(3):209-242, 1985.
- How to do good research, get it published in SIGKDD and get it cited!, Eamonn Keogh, SIGKDD Tutorial 2009.
- False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. Joseph P. Simmons, Leif D. Nelson and Uri Simonsohn. Psychological Science 22(11):1359-1366, 2011.
- Externalisation - how writing changes thinking.. Alan Dix. Interfaces, Autumn 2008.

Guerilla/Discount Usability

- grab a few people and watch them use your interface
  - even 3-5 gives substantial coverage of major usability problems
  - agile/lean qualitative, vs formal quantitative user studies
    - goal is not statistical significance!
- think-aloud protocol
  - contextual inquiry (conversations back and forth) vs fly on the wall (you're silent)

Further reading, usability

- 7 Step Guide to Guerrilla Usability Testing, Markus Piper
  - https://userbrain.net/blog/7-step-guide-guerrilla-usability-testing-diy-usability-testing-method
- The Art of Guerrilla Usability Testing, David Peter Simon
  - http://www.uxbooth.com/articles/the-art-of-guerrilla-usability-testing/
- Discount Usability: 20 Years, Jakob Nielsen
  - https://www.nngroup.com/articles/discount-usability-20-years/
- Interaction Design: Beyond Human-Computer Interaction
  - Preece, Sharp, Rogers. Wiley, 4th edition, 2015.
- About Face: The Essentials of Interaction Design
  - Cooper, Reimann, Cronin, Noessel. Wiley, 4th edition, 2014.
- Task-Centered User Interface Design. Lewis & Rieman, 1994
  - http://hcibib.org/tcuid/
- Designing with the Mind in Mind. Jeff Johnson. Morgan Kaufmann, 2nd, 2014.

Ch 5. Marks and Channels

How to encode: Arrange space, map channels

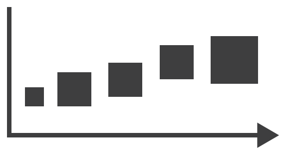
Encode

Arrange

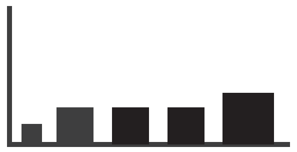
- Express

- Separate

- Order



- Align



- Use

Map

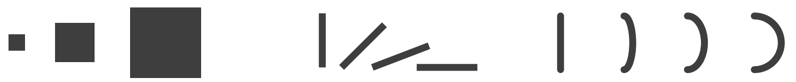
from categorical and ordered attributes

- Color
  - *Hue*

  - *Saturation*

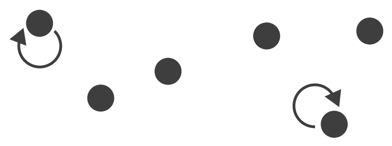
  - *Luminance*

- Size, Angle, Curvature, ...



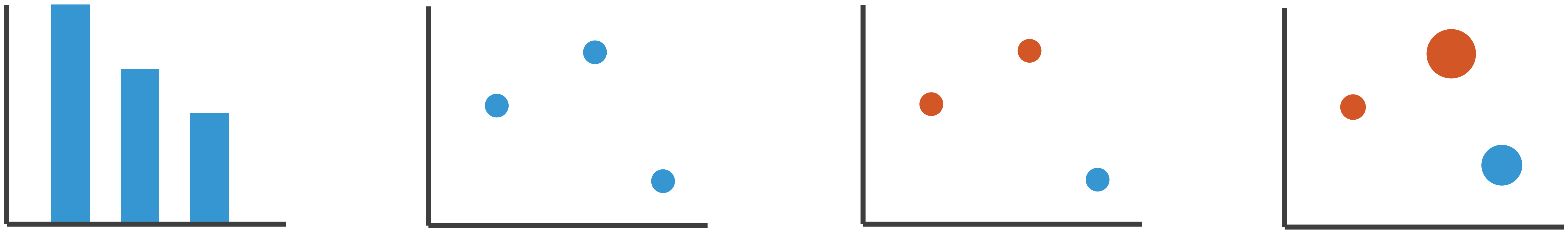
- Shape

- Motion
  Direction, Rate, Frequency, ...



Visual encoding

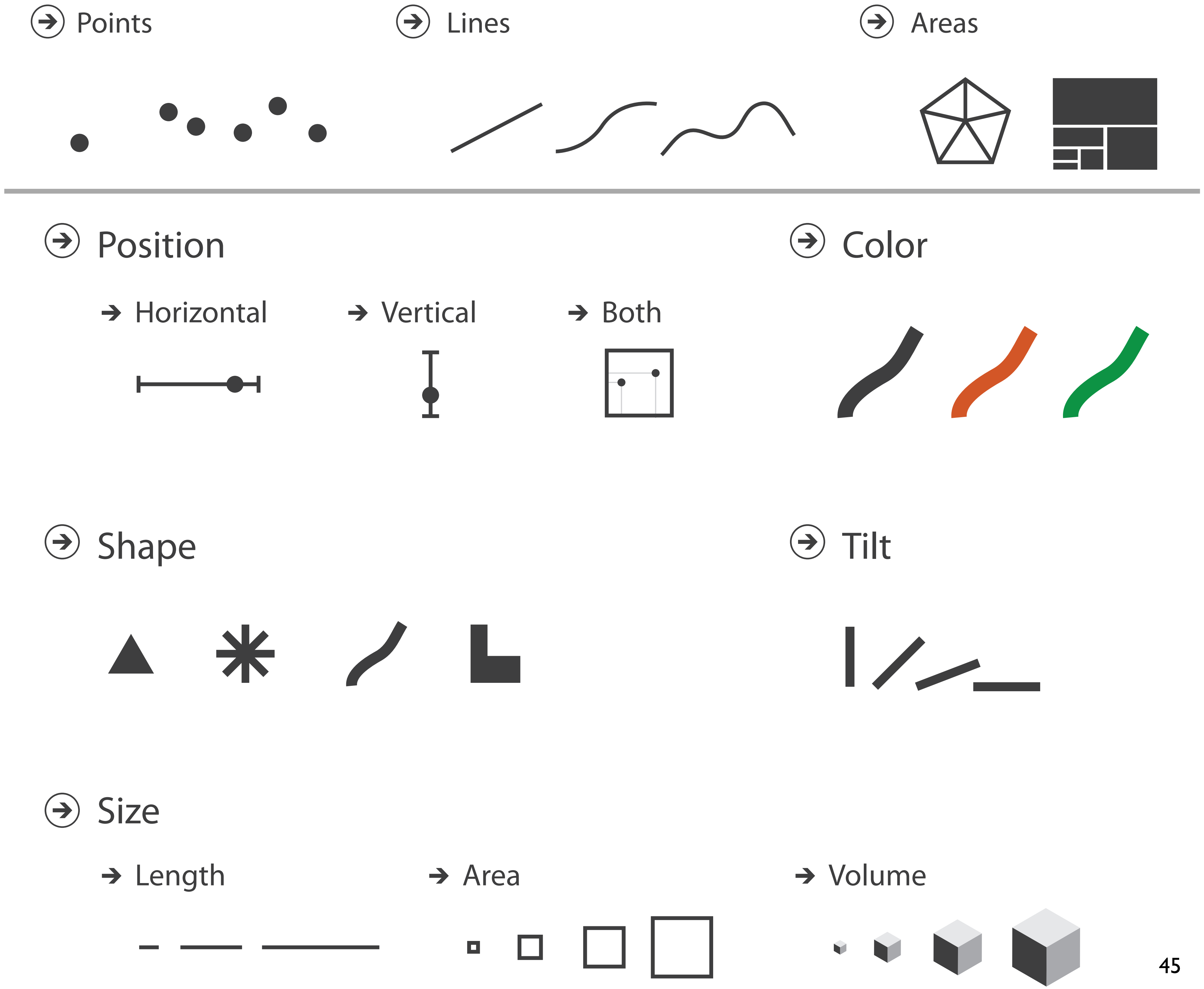
- analyze idiom structure



Definitions: Marks and channels

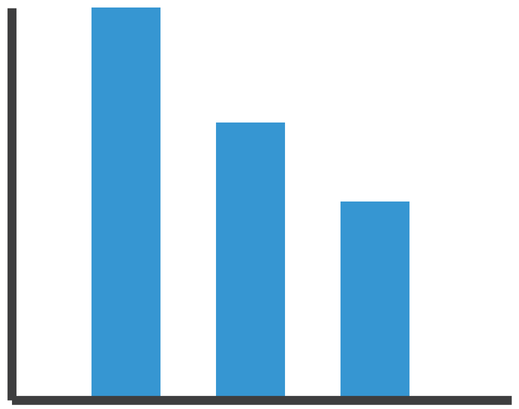
- marks
  - geometric primitives

- channels
  - control appearance of marks
  - can redundantly code with multiple channels



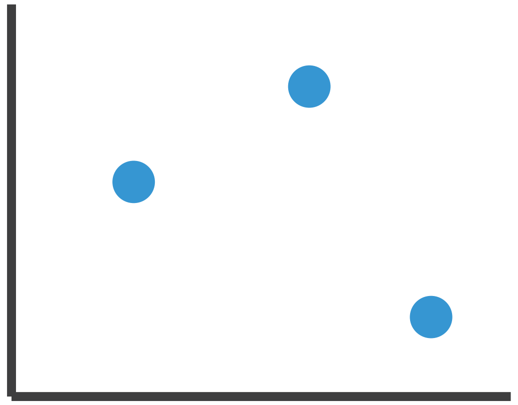
Visual encoding

- analyze idiom structure
  - as combination of marks and channels



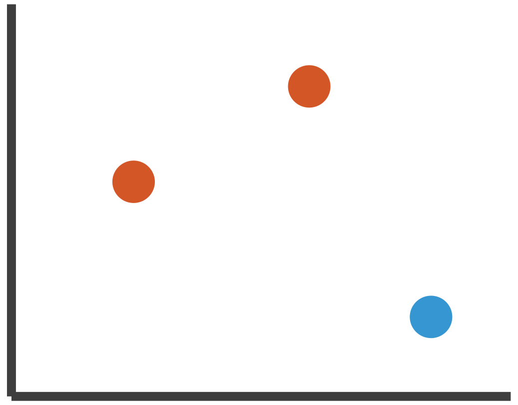
1:
vertical position

mark: line



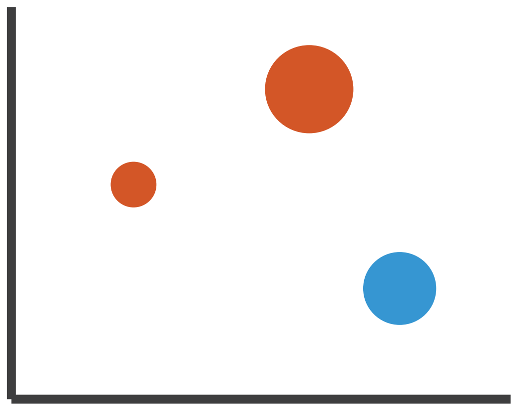
2:
vertical position
horizontal position

mark: point



3:
vertical position
horizontal position
color hue

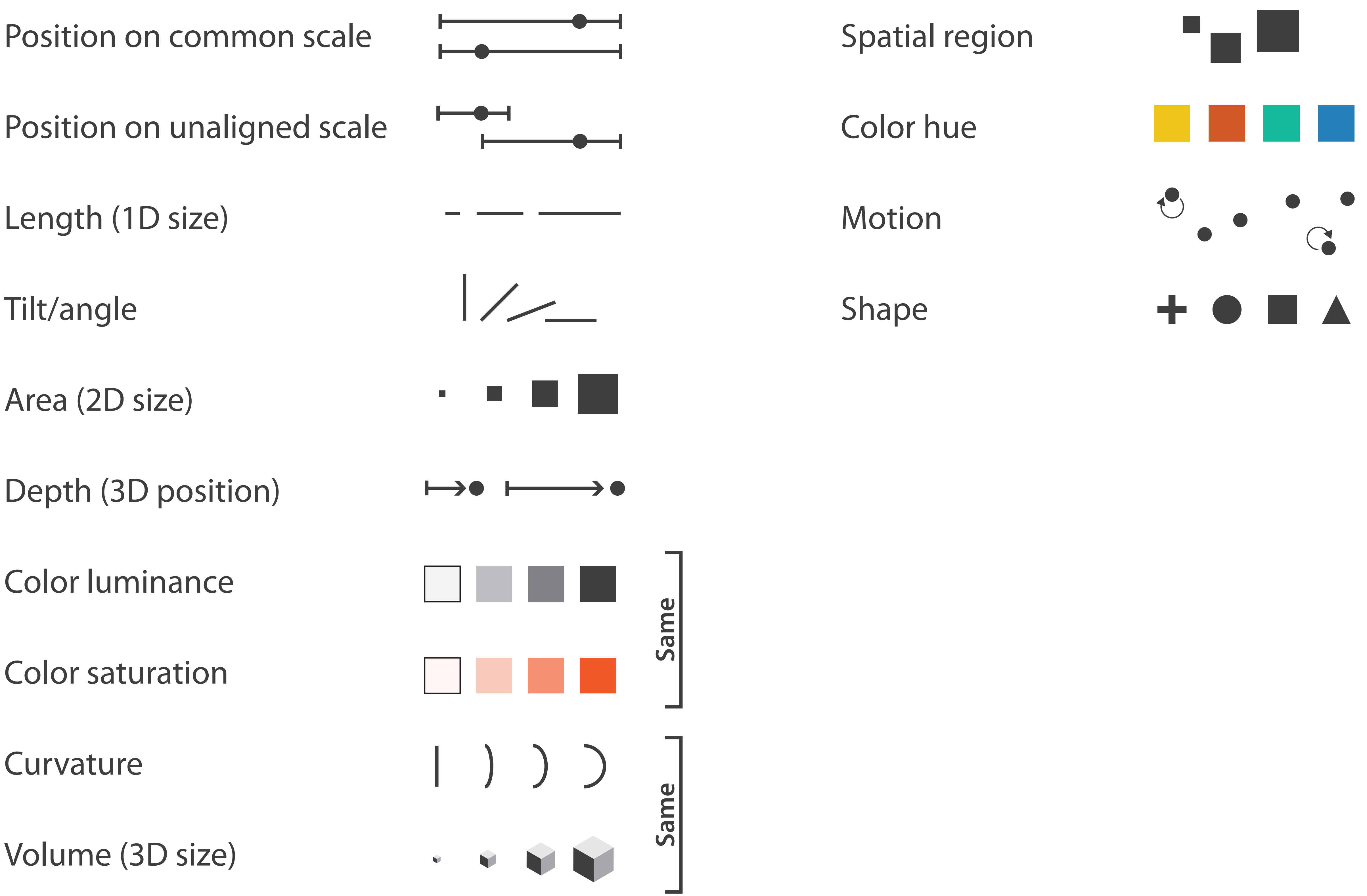
mark: point



4:
vertical position
horizontal position
color hue
size (area)

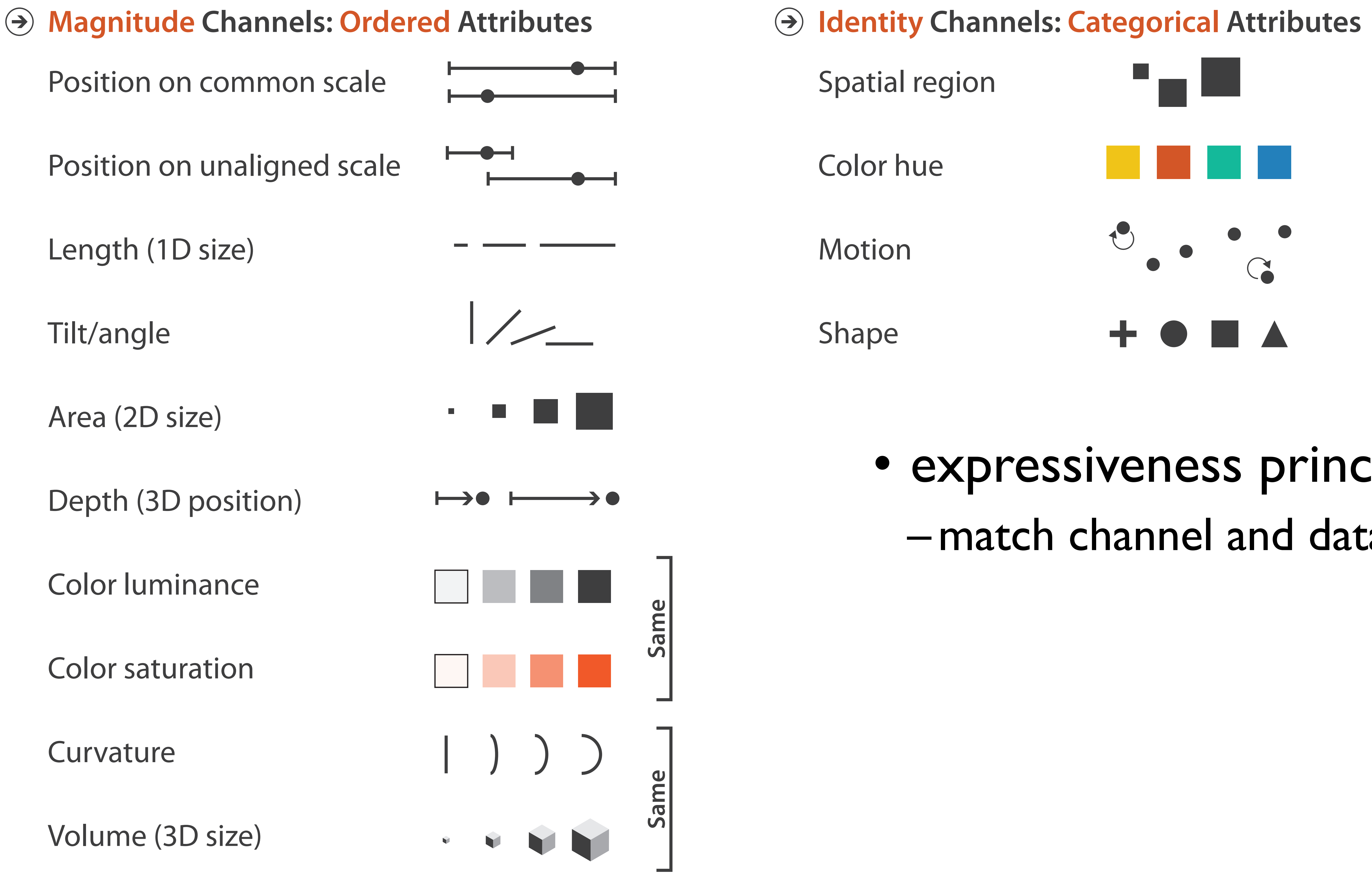
mark: point

Channels



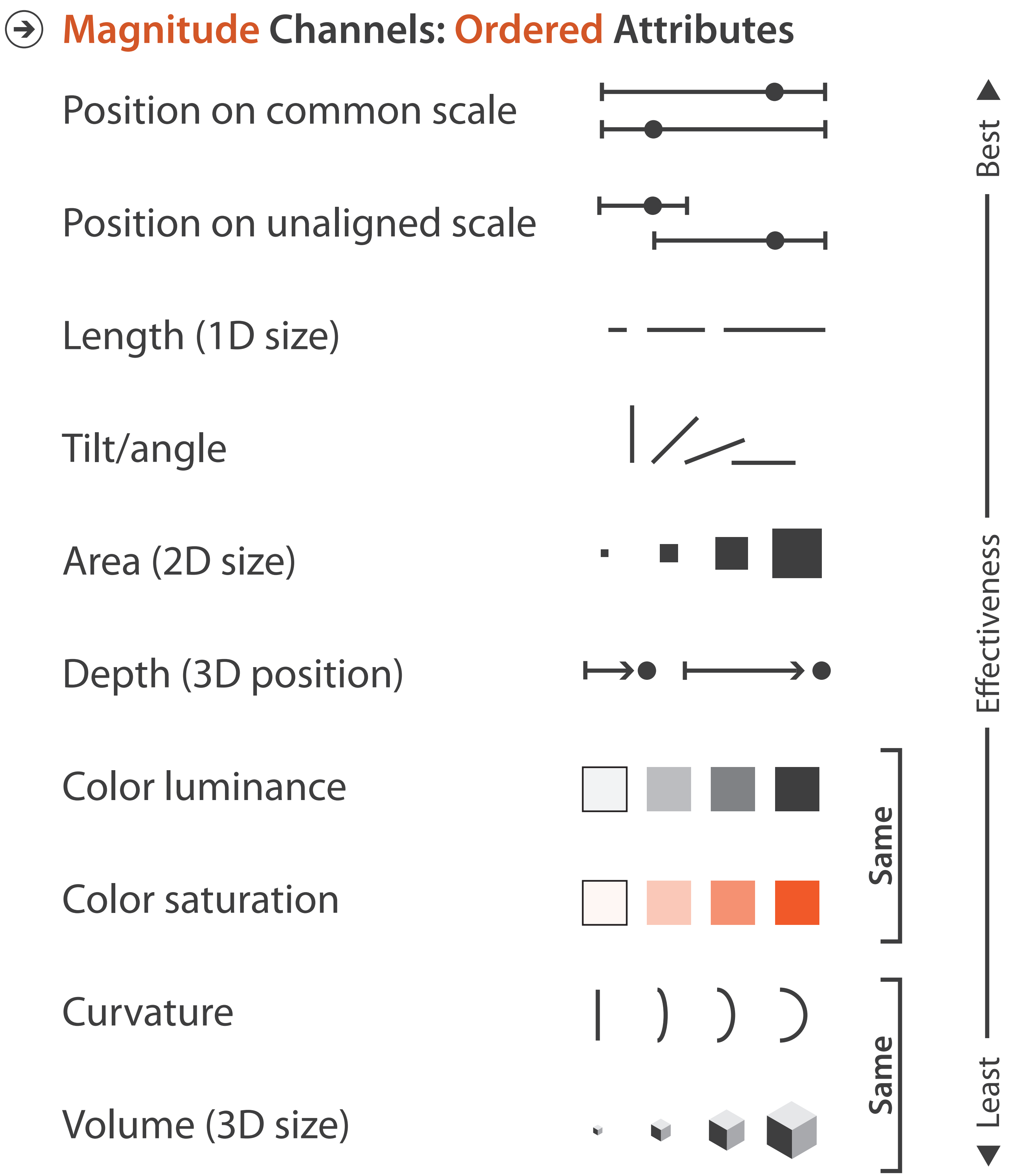
Position on common scale
Position on unaligned scale
Length (1D size)
Tilt/angle
Area (2D size)
Depth (3D position)
Color luminance
Color saturation
Curvature
Volume (3D size)
Same
Same
Spatial region
Color hue
Motion
Shape

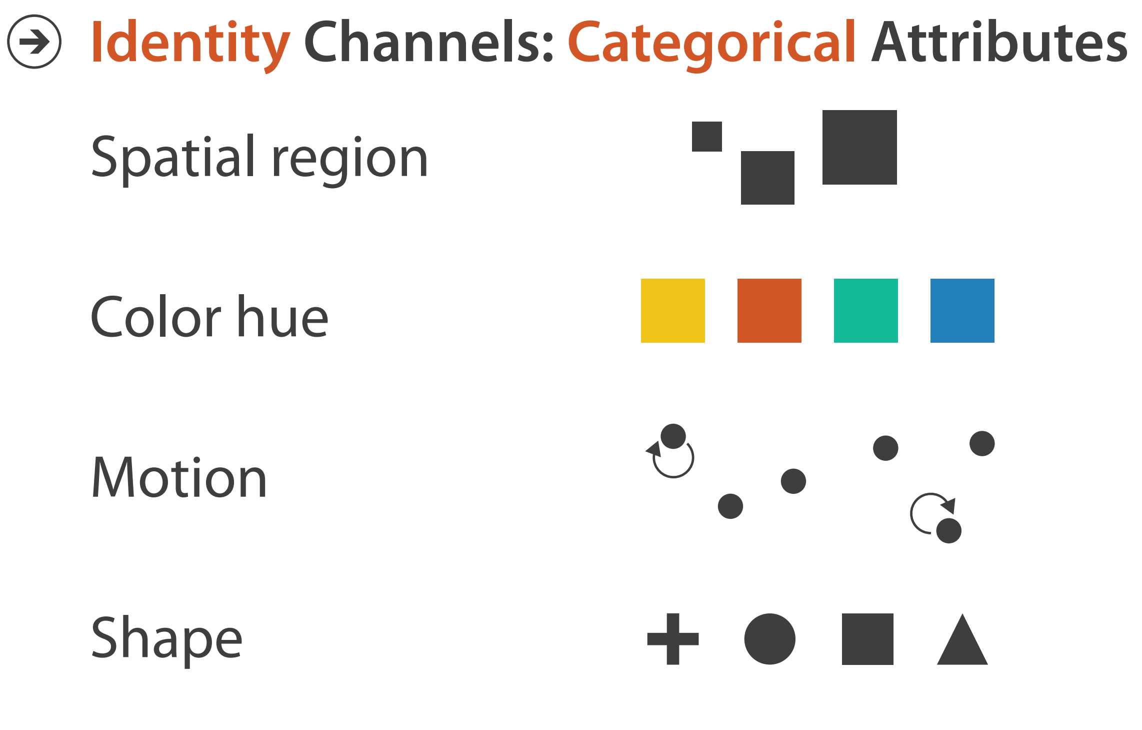
Channels: Matching Types



- expressiveness principle
  - match channel and data characteristics

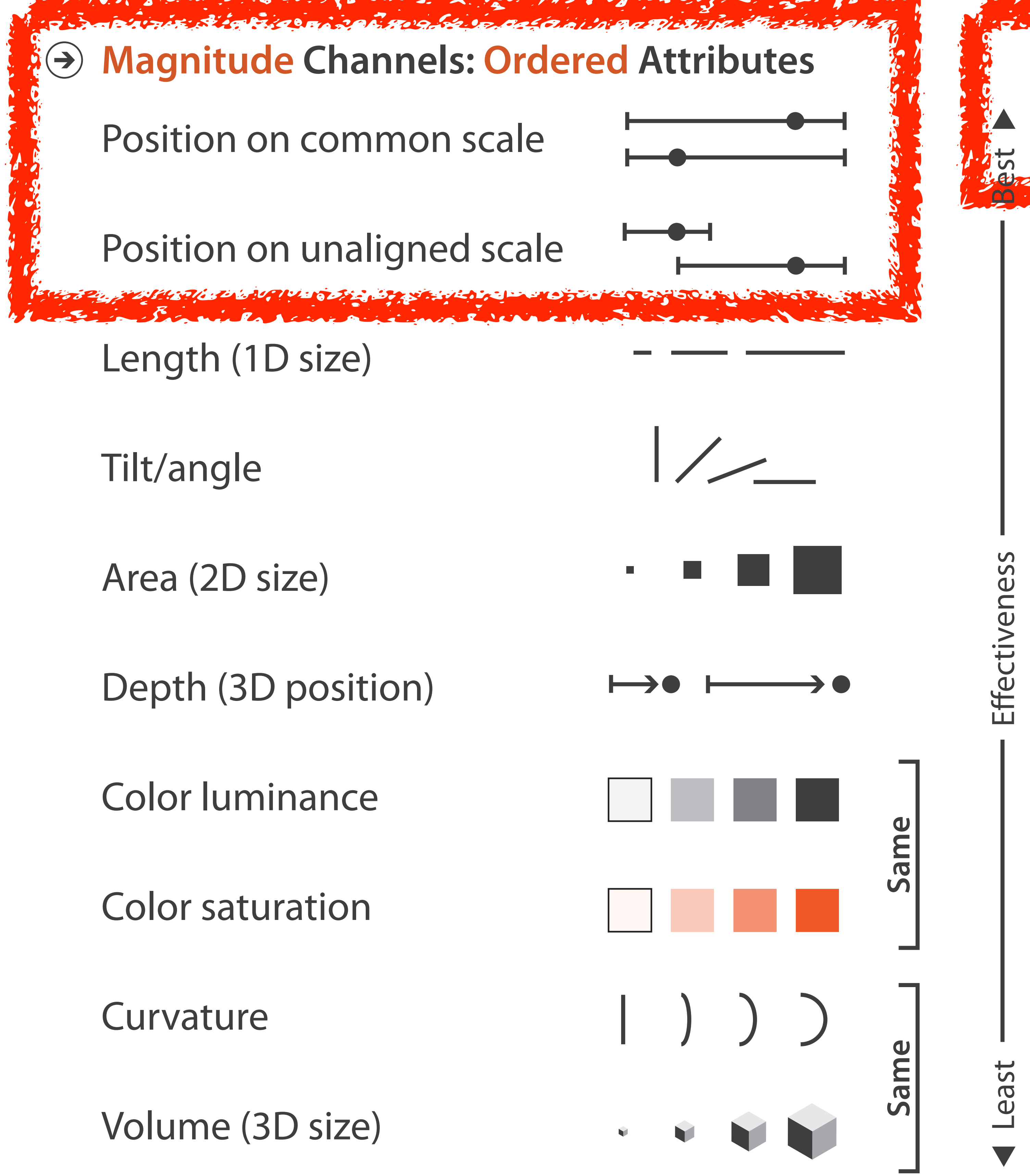
Channels: Rankings

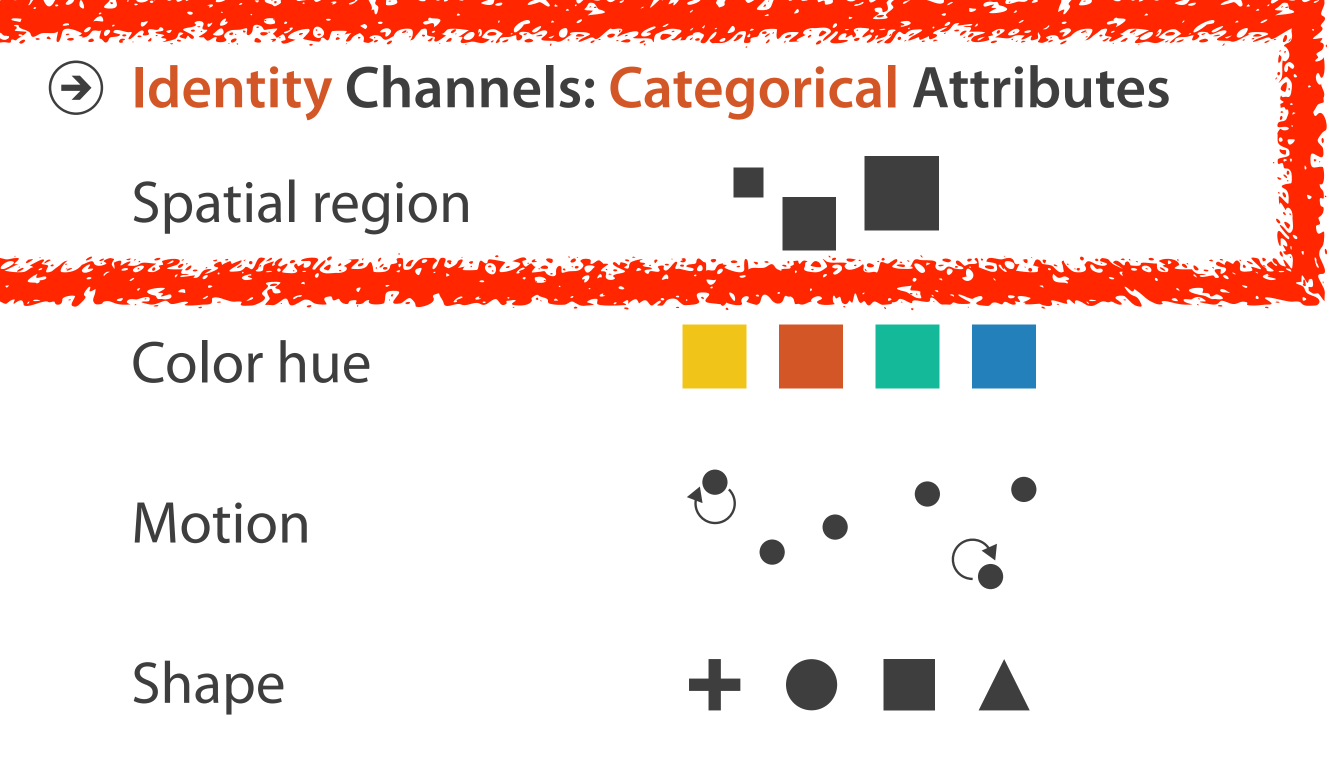




- expressiveness principle
  - match channel and data characteristics
- effectiveness principle
  - encode most important attributes with highest ranked channels

Channels: Expressiveness types and effectiveness rankings

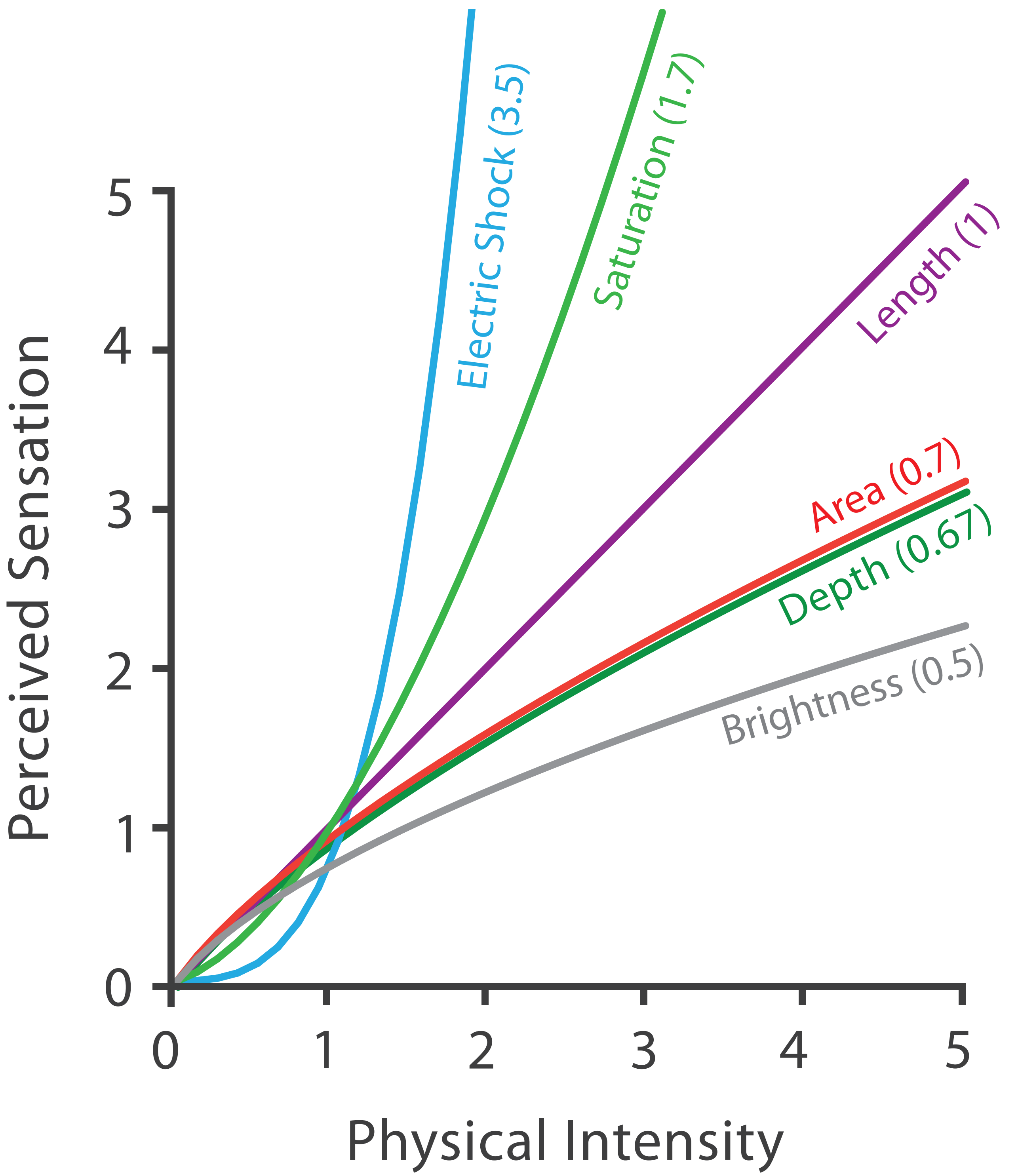




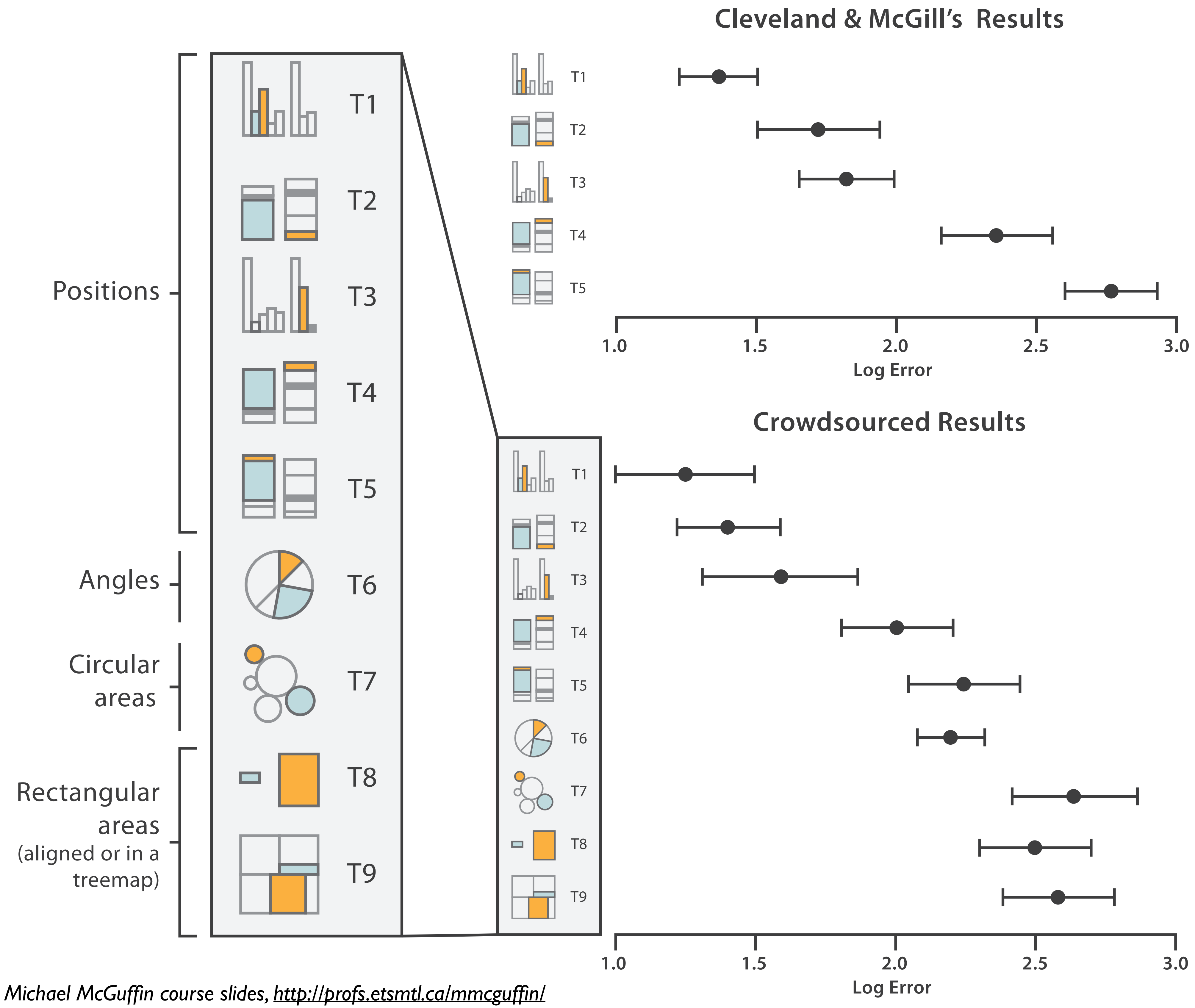
- expressiveness principle
  - match channel and data characteristics
- effectiveness principle
  - encode most important attributes with highest ranked channels
  - spatial position ranks high for both

Accuracy: Fundamental Theory

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



Accuracy: Vis experiments

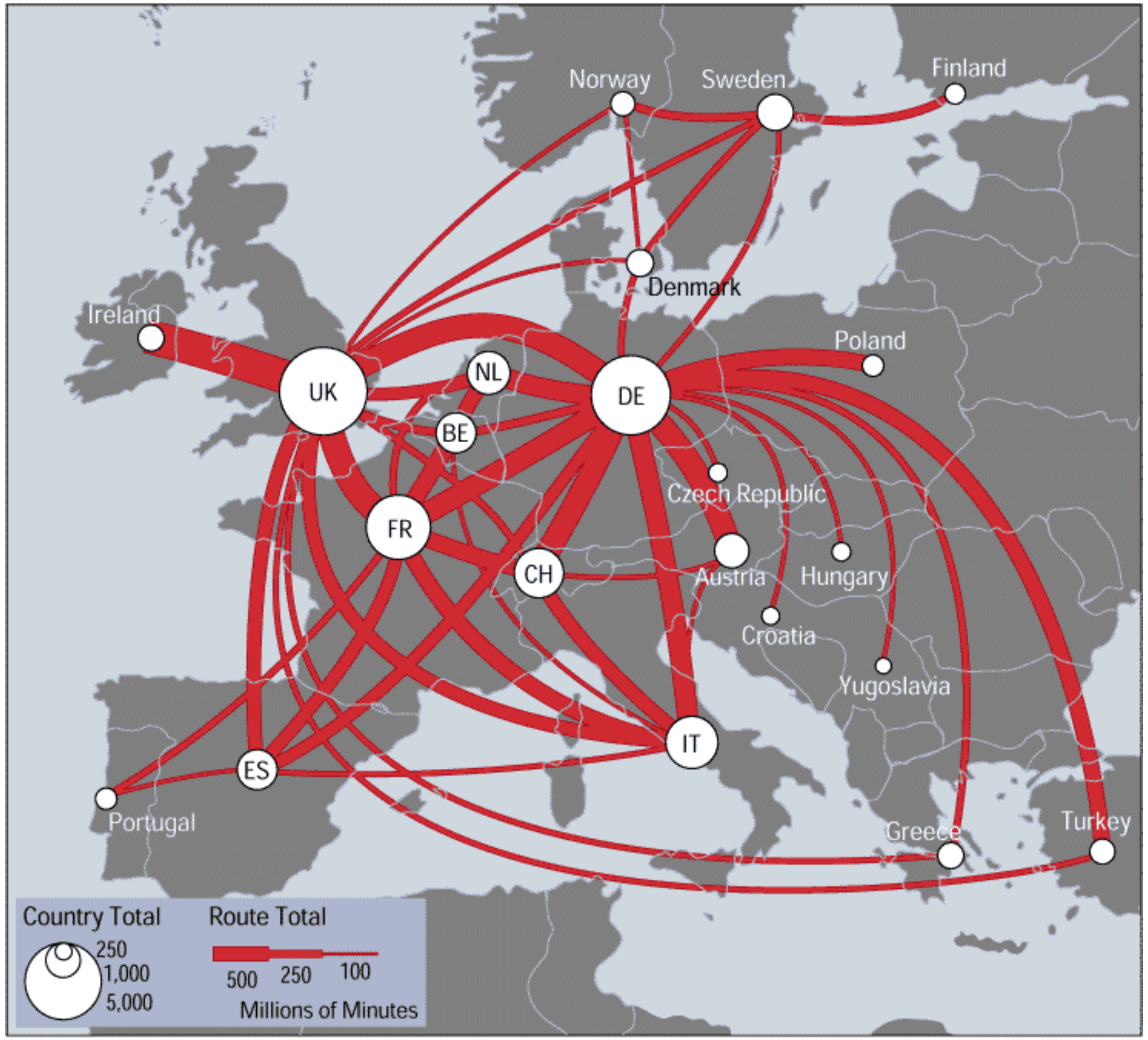


[Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. Heer and Bostock. Proc ACM Conf. Human Factors in Computing Systems (CHI) 2010, p. 203–212.]

after Michael McGuffin course slides, http://profs.etsmtl.ca/mmcguffin/

Discriminability: How many usable steps?

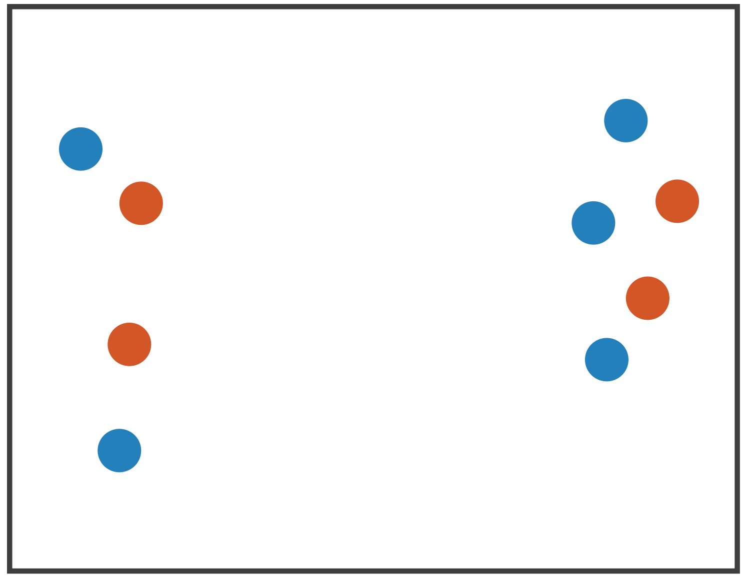
- must be sufficient for number of attribute levels to show
  - linewidth: few bins but salient



[mappa.mundi.net/maps/maps 014/telegeography.html]

Separability vs. Integrality

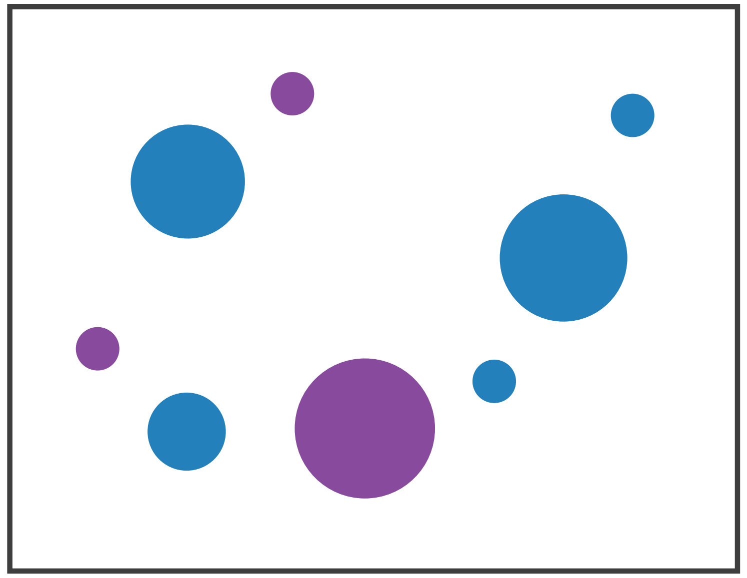
Position
+ Hue (Color)



Fully separable

2 groups each

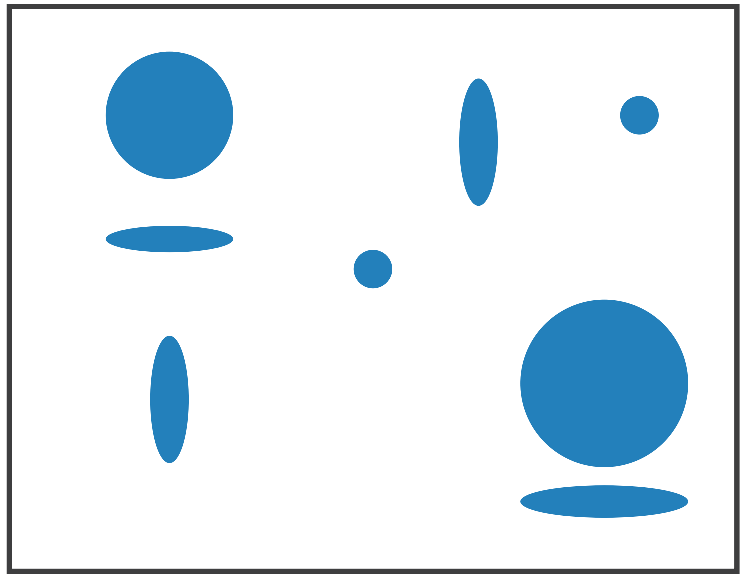
Size
+ Hue (Color)



Some interference

2 groups each

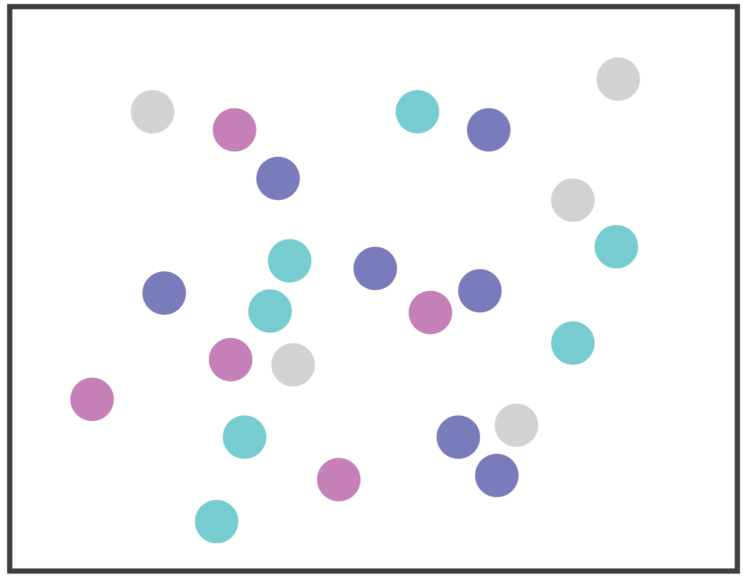
Width
+ Height



Some/significant interference

3 groups total: integral area

Red
+ Green

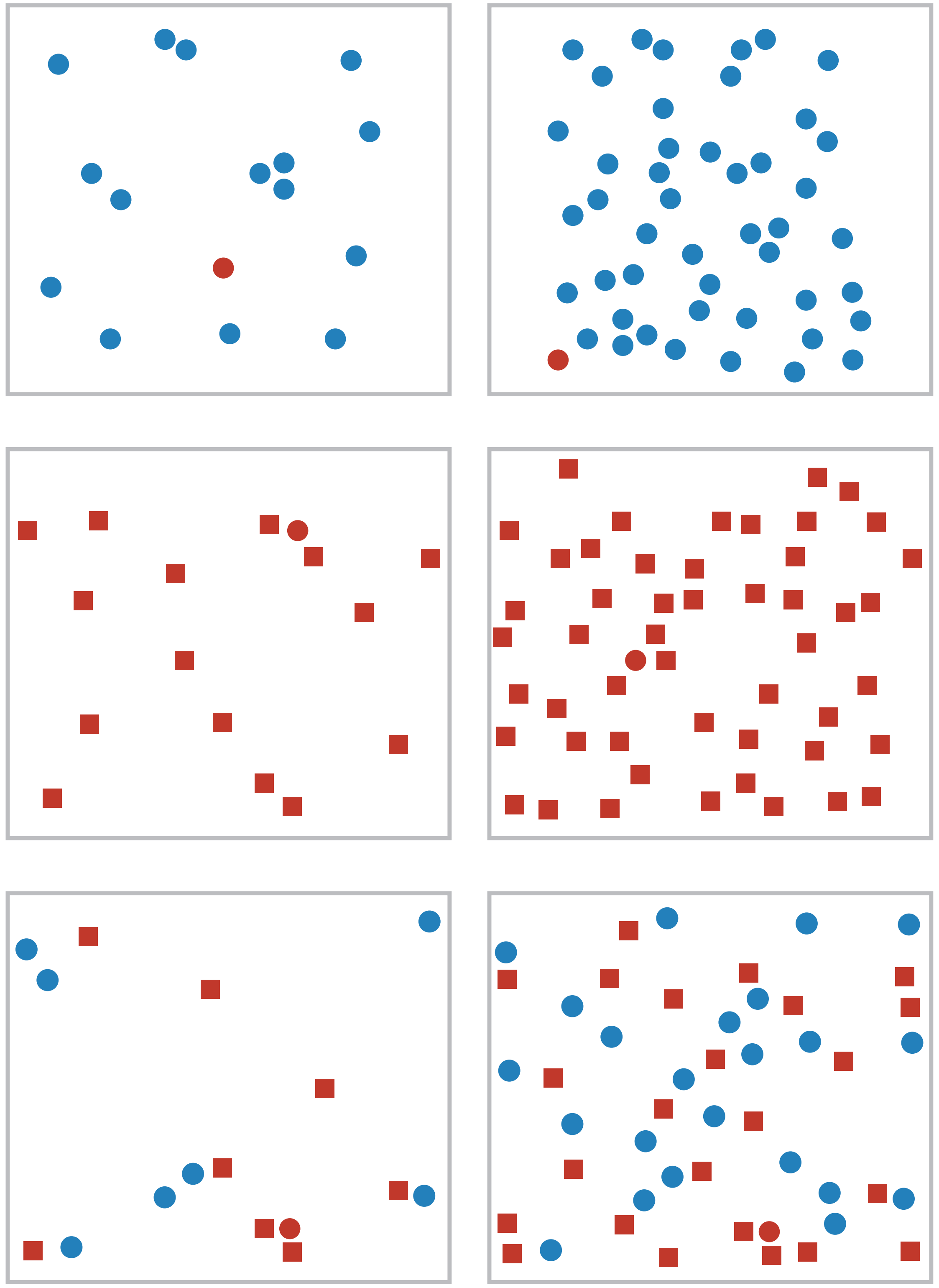


Major interference

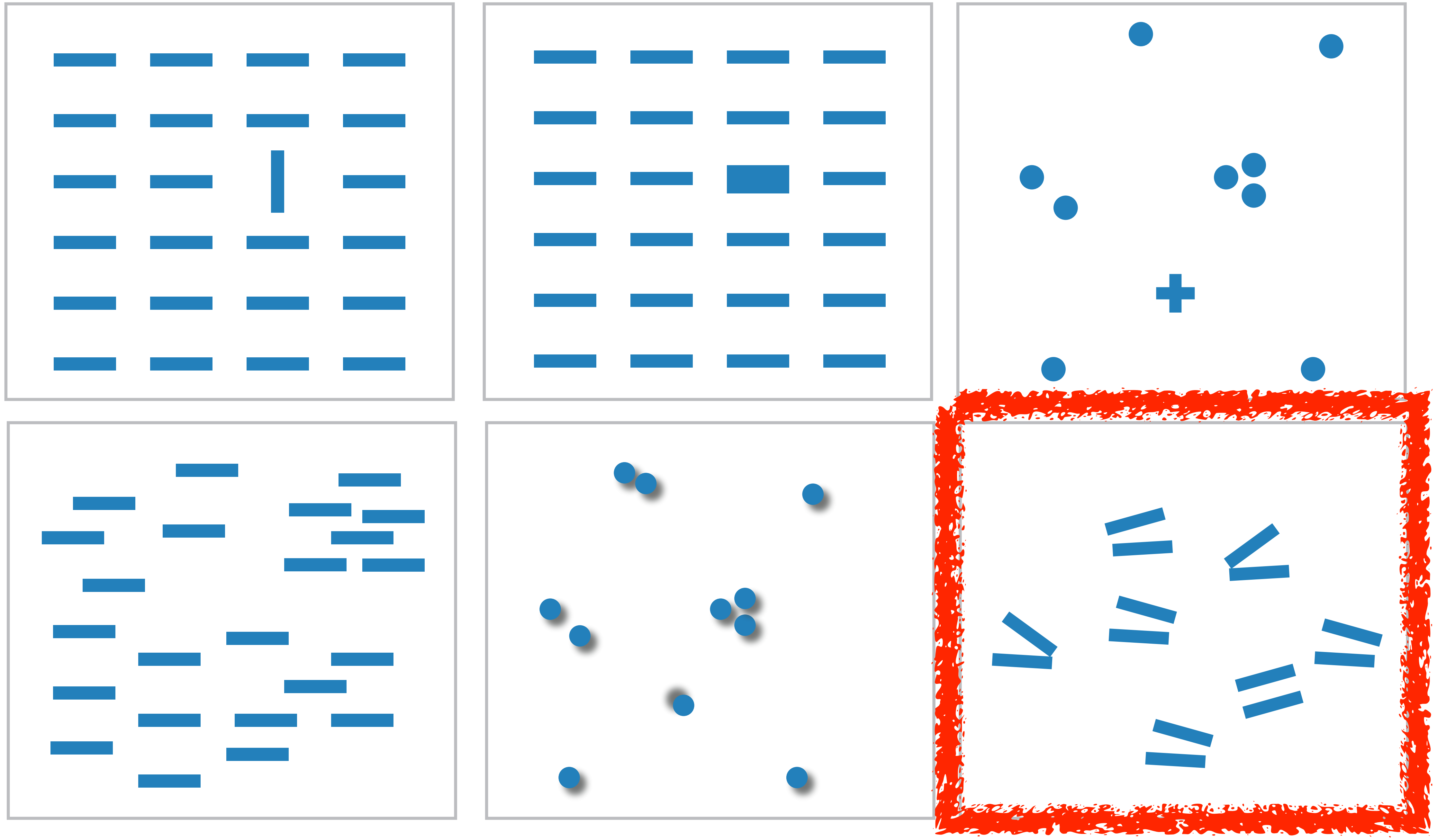
4 groups total: integral hue

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

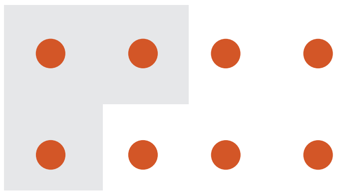
Grouping

- containment
- connection

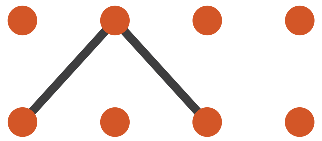
- proximity
  - same spatial region
- similarity
  - same values as other categorical channels

Marks as Links

Containment



Connection

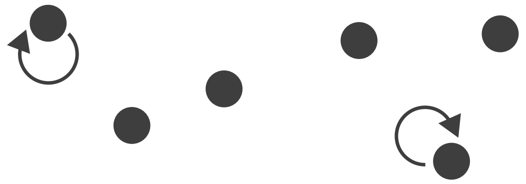


Identity Channels: Categorical Attributes

Spatial region

Color hue

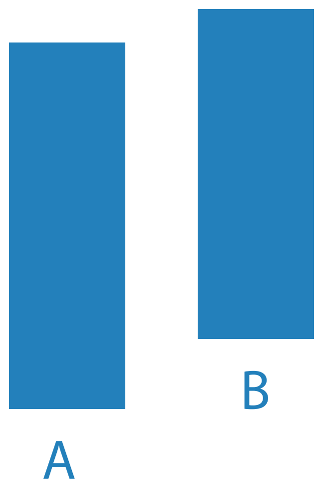
Motion



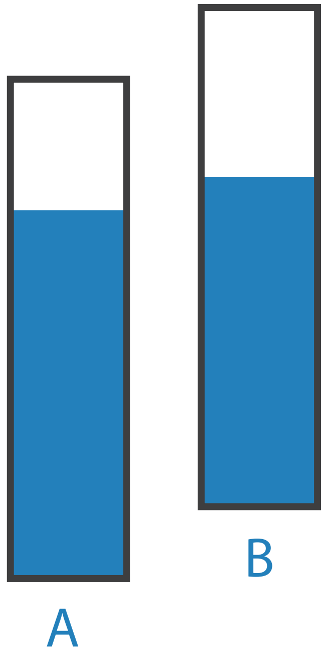
Shape

Relative vs. absolute judgements

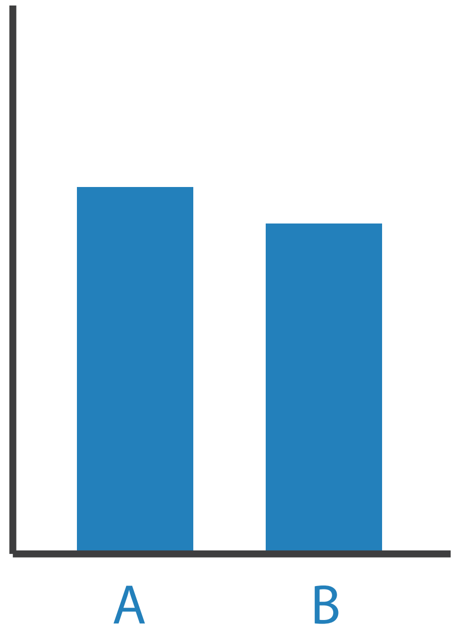
- perceptual system mostly operates with relative judgements, not absolute
  - that's why accuracy increases with common frame/scale and alignment
  - Weber's Law: ratio of increment to background is constant
    - filled rectangles differ in length by 1:9, difficult judgement
    - white rectangles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



length



position along unaligned common scale

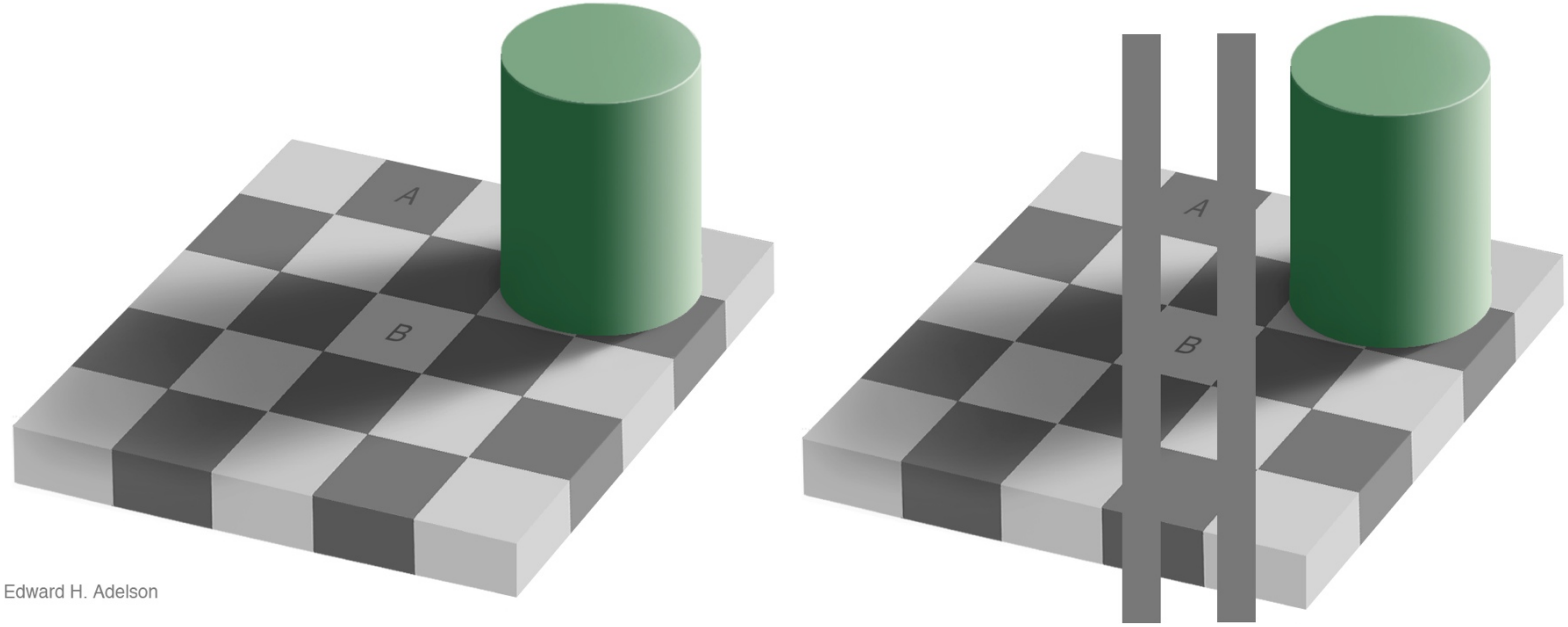


position along aligned scale

after [Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. Cleveland and McGill. Journ. American Statistical Association 79:387 (1984), 531–554.]

Relative luminance judgements

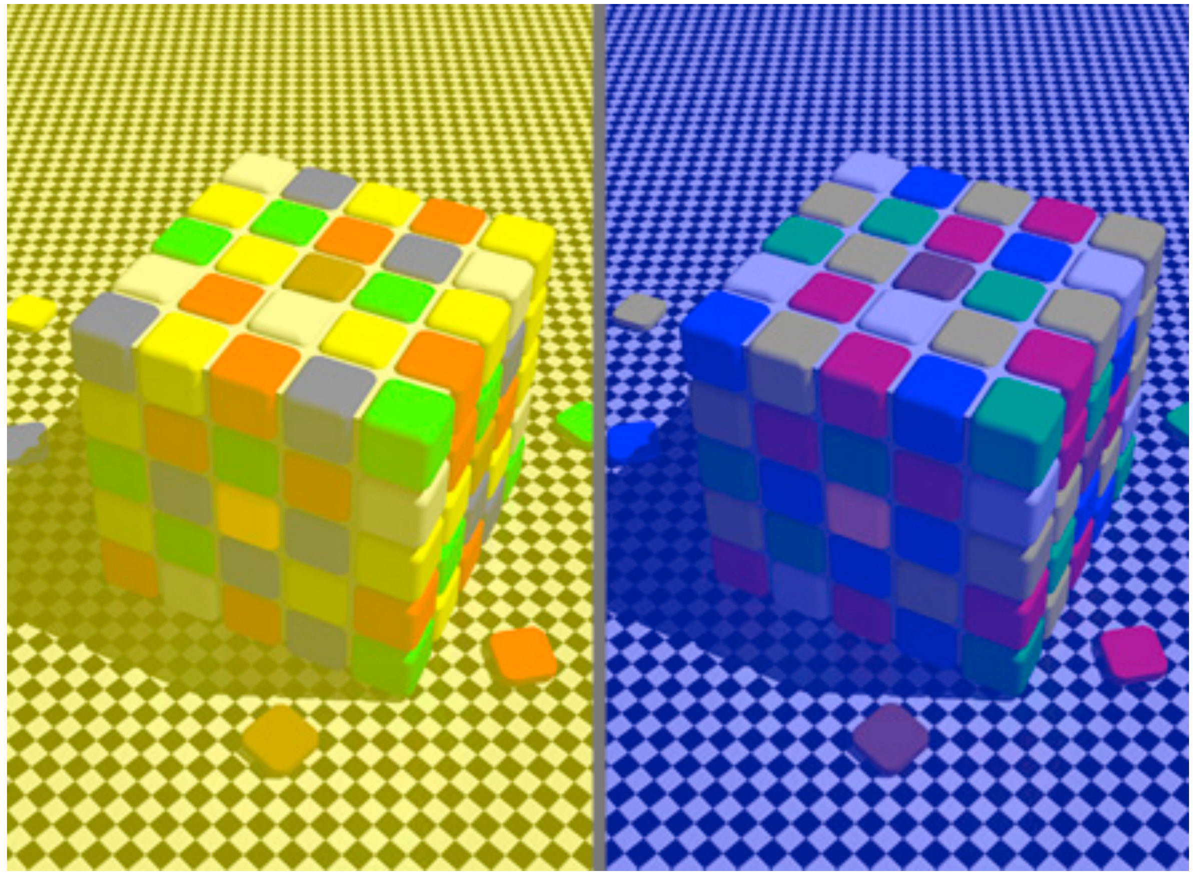
- perception of luminance is contextual based on contrast with surroundings

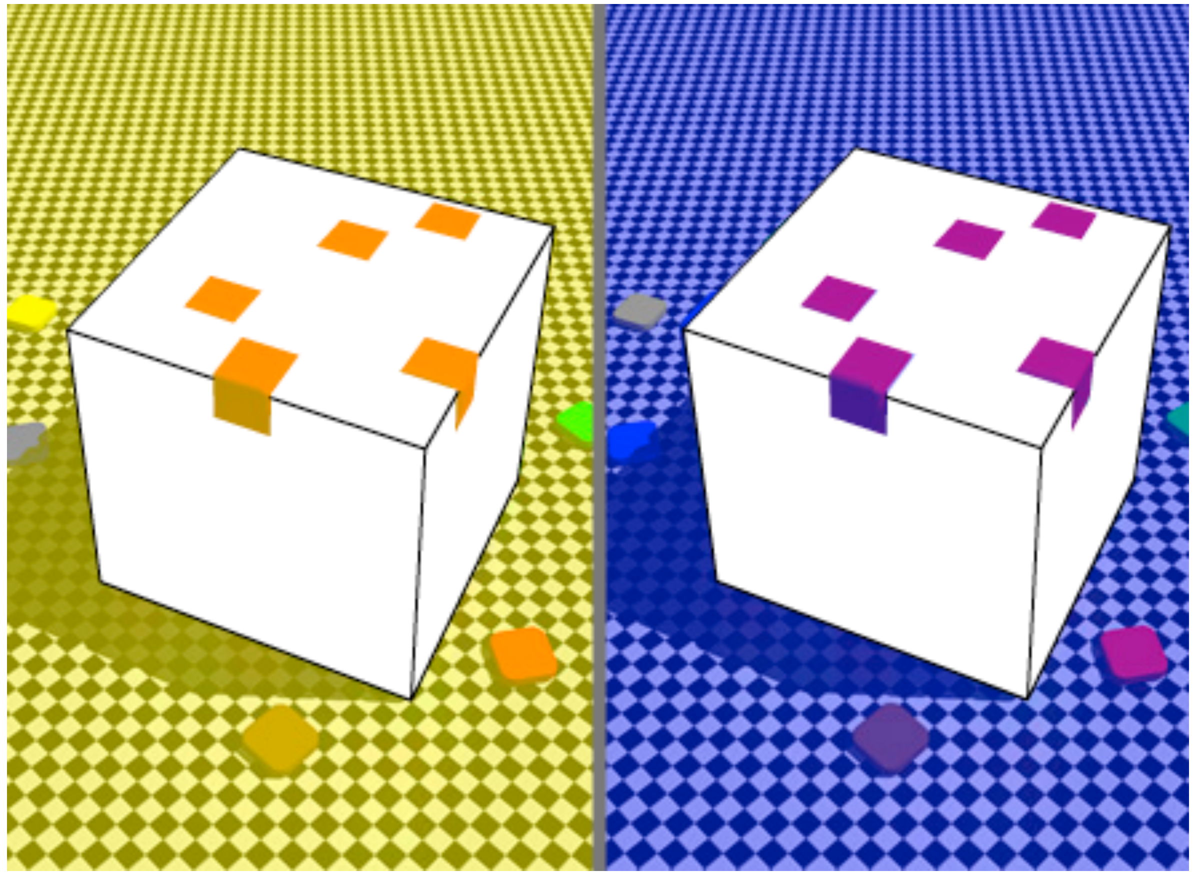


http://persci.mit.edu/gallery/checkershadow

Relative color judgements

- color constancy across broad range of illumination conditions





http://www.purveslab.net/seeforyourself/

Further Reading

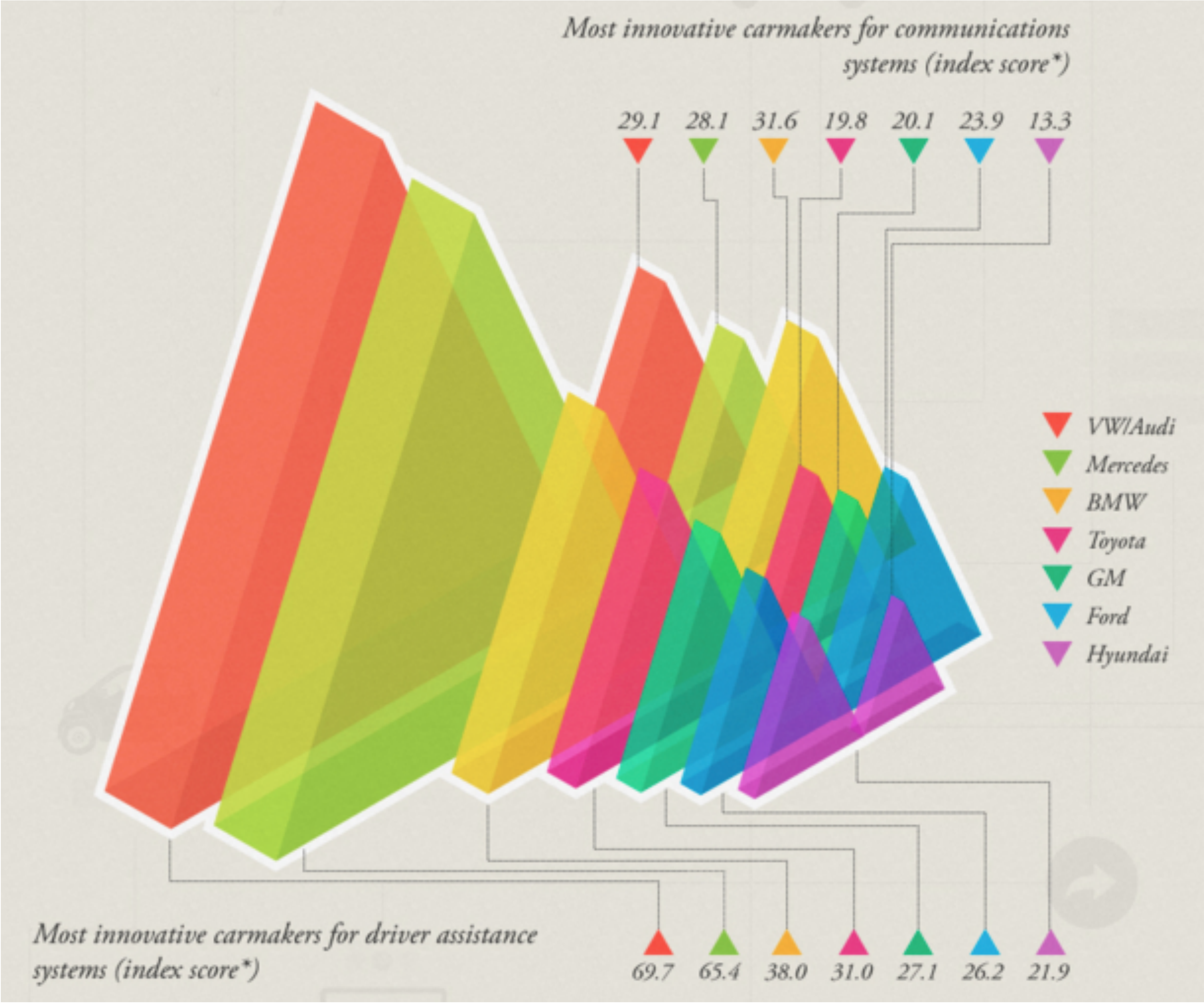
- Visualization Analysis and Design. Munzner. AK Peters Visualization Series. CRC Press. 2014
  - Chap 5: Marks and Channels
- Semiology of Graphics, Jacques Bertin, Gauthier-Villars 1967, EHESS 1998.
- Graphical Perception: Theory, Experimentation and the Application to the Development of Graphical Models. William S. Cleveland, Robert McGill, J. Am. Stat. Assoc. 79:387, pp. 531-554, 1984.
- A Model for Studying Display Methods of Statistical Graphics (with Discussion). William S. Cleveland. Journal of Computational and Statistical Graphics 2(4):323-364 1993.
- Automating the Design of Graphical Presentations of Relational Information. Jock Mackinlay, ACM Transaction on Graphics, vol. 5, no. 2, April 1986, pp. 110-141.
- Information Visualization: Perception for Design, 3rd edition, Colin Ware, Morgan Kaufmann, 2013.
- How Maps Work: Representation, Visualization, and Design. Alan M. MacEachren. Guilford Press, 1995.
- Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. Jeffrey Heer and Michael Bostock. Proc. CHI 2010.
- Taxonomy-Based Glyph Design---With a Case Study on Visualizing Workflows of Biological Experiments. Eamonn Maguire, Philippe Rocca-Serra, Susanna-Assunta Sansone, Jim Davies, and Min Chen. IEEE TVCG (Proc. InfoVis 12) 18(12):2603-2612 2012.
- Glyph-Based Visualization: Foundations, Design Guidelines, Techniques and Applications. Rita Borgo, Johannes Kehrer, David H.S. Chung, Eamonn Maguire, Robert S. Laramee, Helwig Hauser, Matthew Ward, and Min Chen. Eurographics State of the Art Reports (STAR):39-63 2013.
- On the Theory of Scales of Measurement. S. S. Stevens. Science 103(2684):677-680, 1946.
- Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects, S. S. Stevens, Wiley 1975.
- Visual Thinking for Design, Colin Ware, Morgan Kaufmann 2008.
- Perception in Vision web page with demos, Christopher Healey. (see also Attention and Visual Memory in Visualization and Computer Graphics, Christopher G. Healey and James T. Enns, IEEE TVCG 18(7):1170-1188 2012.)
- Feature Analysis in Early Vision: Evidence from Search Asymmetries. Treisman and Gormican. Psychological Review 95(1):15-48, 1988.

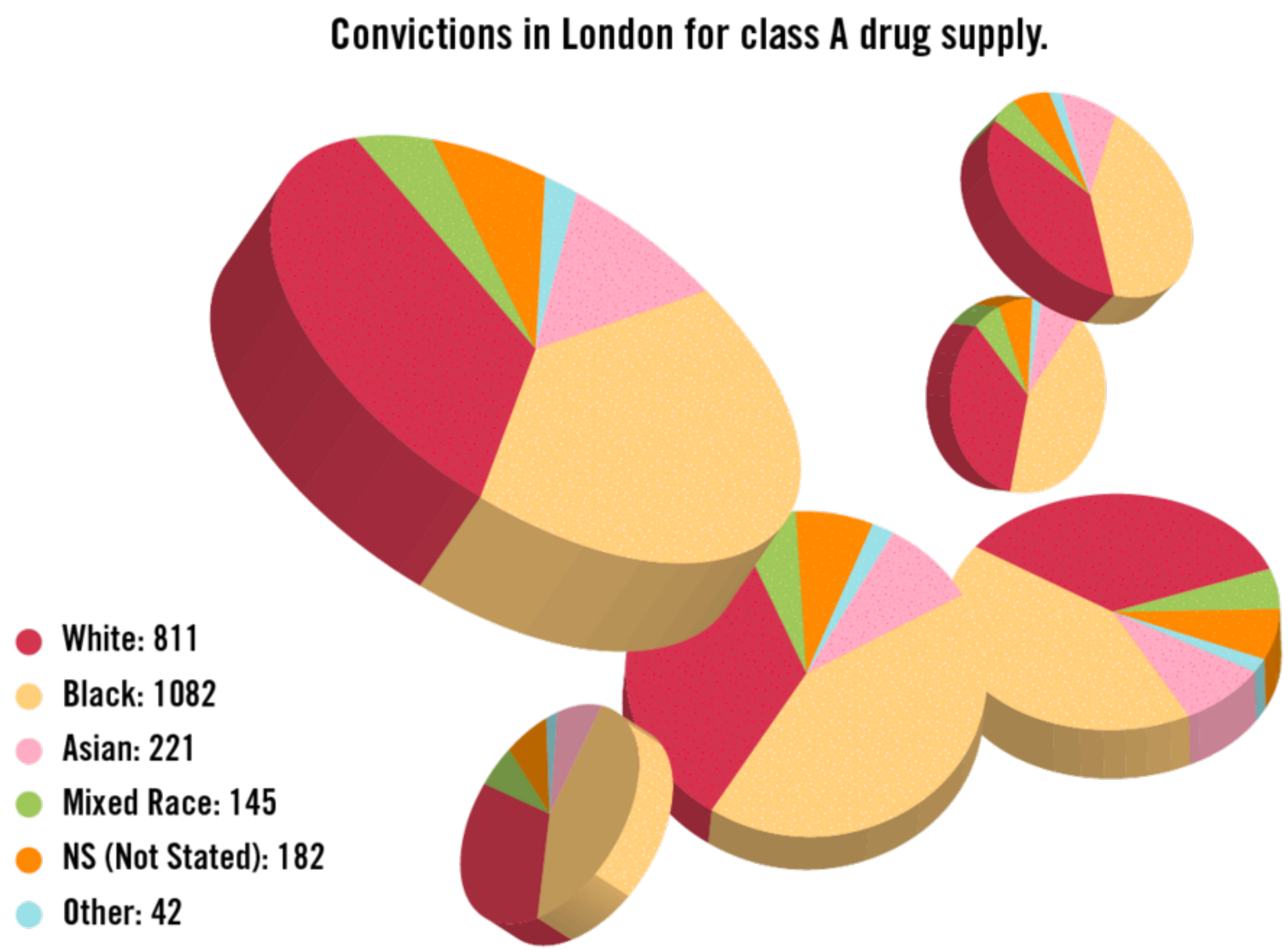
Ch 6. Rules of Thumb

Rules of Thumb

- No unjustified 3D
  - –Power of the plane
  - –Disparity of depth
  - –Occlusion hides information
  - –Perspective distortion dangers
  - –Tilted text isn't legible
- No unjustified 2D
- Eyes beat memory
- Resolution over immersion
- Overview first, zoom and filter, details on demand
- Responsiveness is required
- Function first, form next

Unjustified 3D all too common, in the news and elsewhere



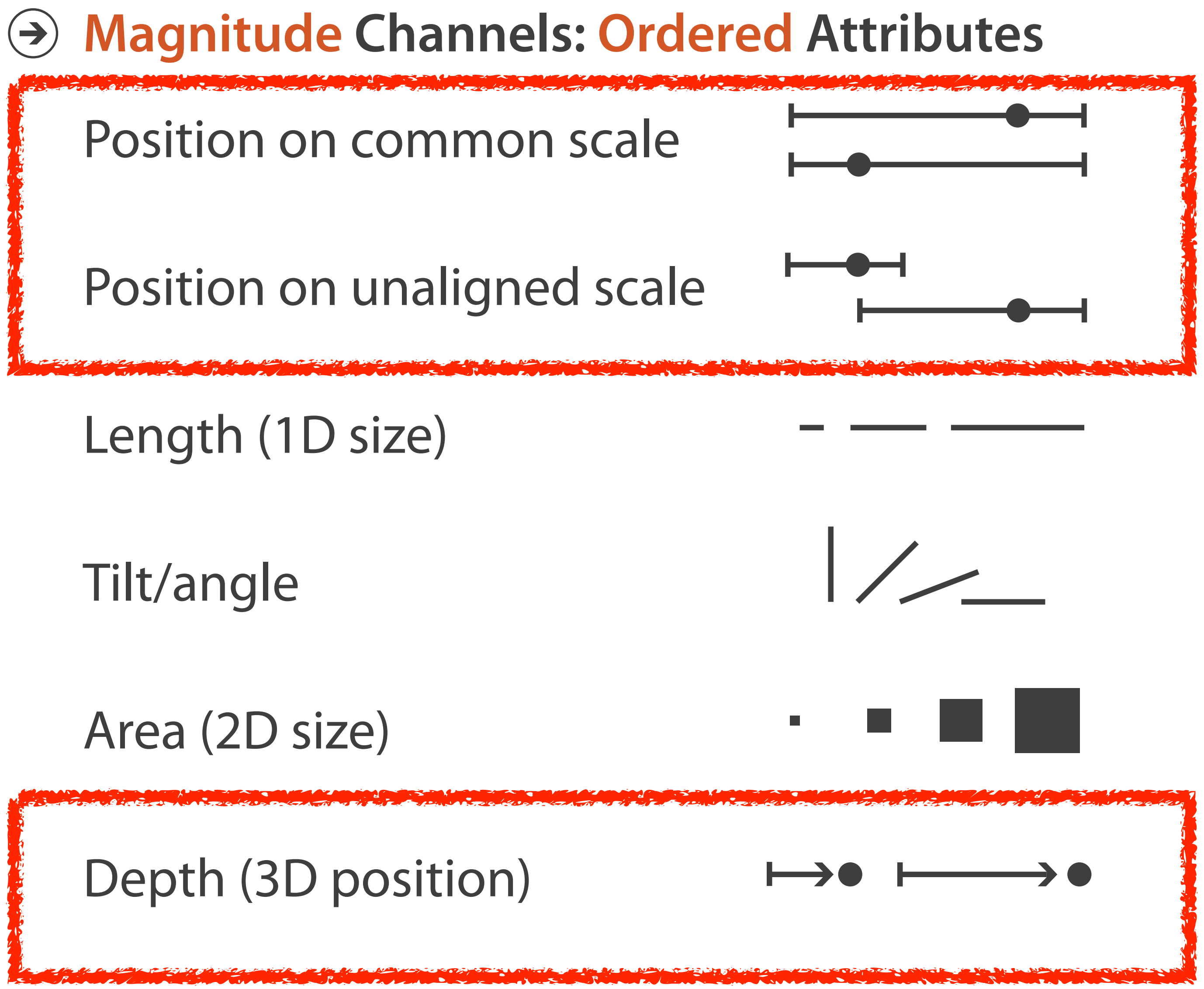


http://viz.wtf/post/137826497077/eye-popping-3d-triangles

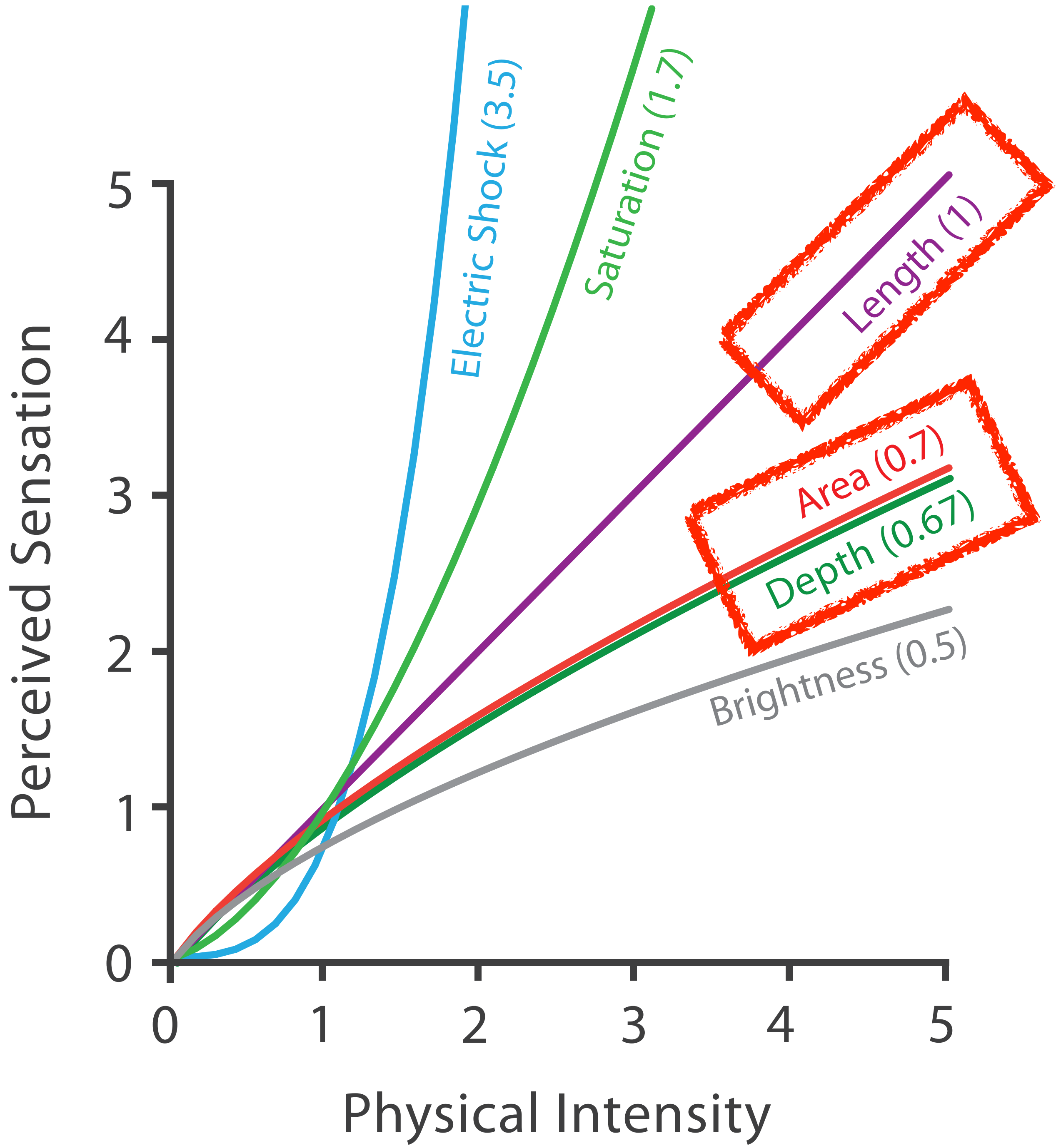
http://viz.wtf/post/139002022202/designer-drugs-ht-ducqn

Depth vs power of the plane

- high-ranked spatial position channels: **planar** spatial position
  - not depth!

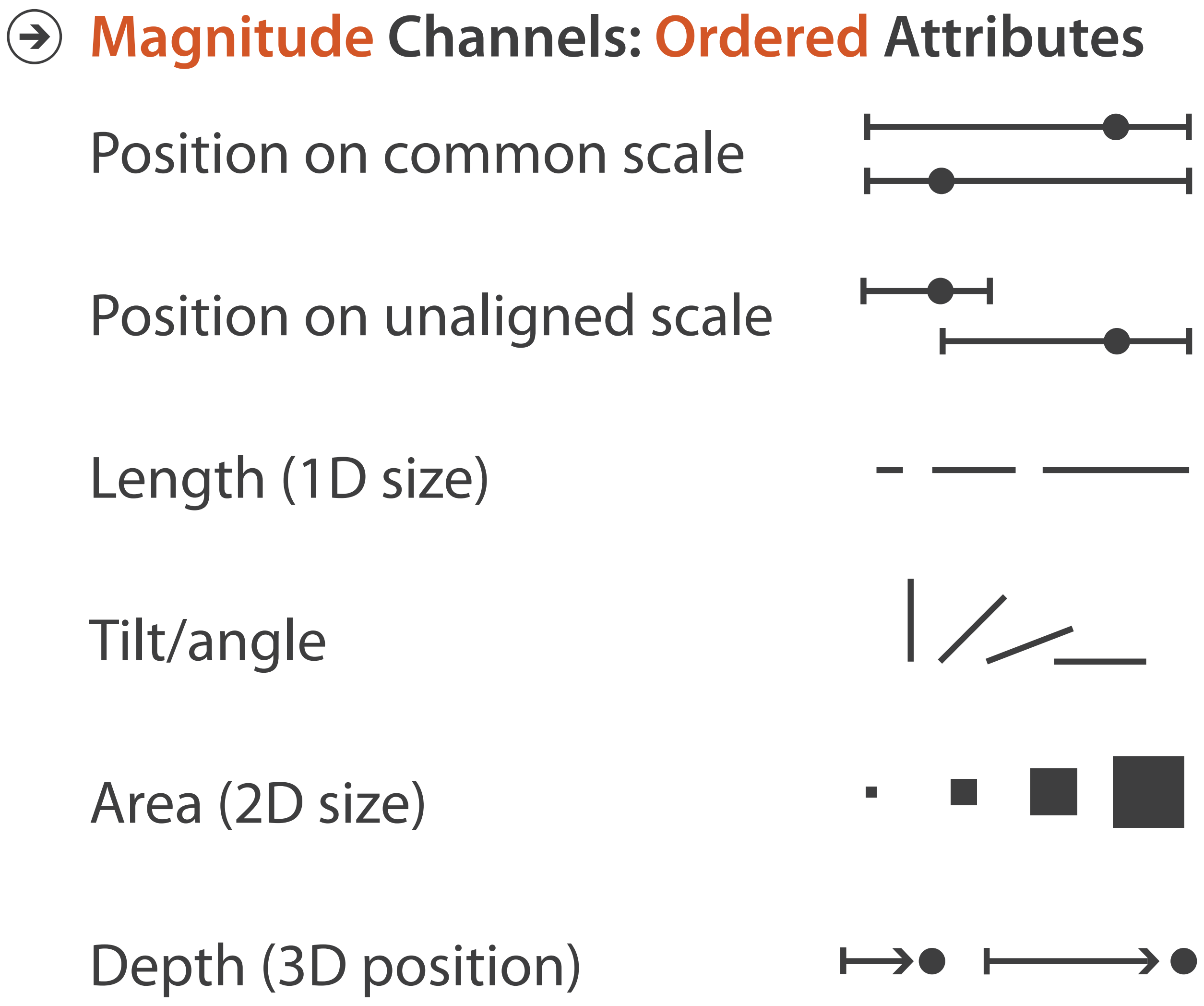


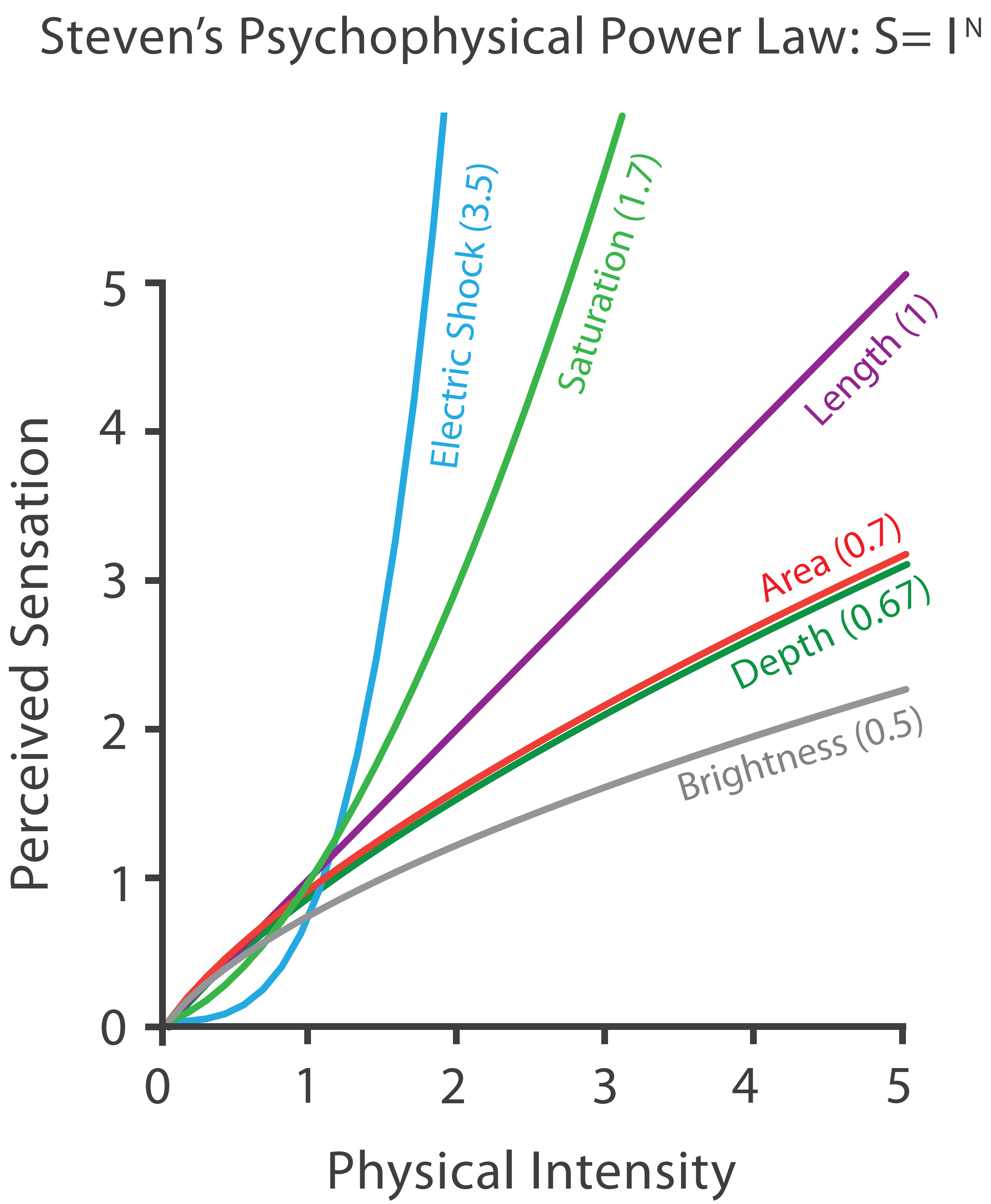
Steven's Psychophysical Power Law: $S = I^N$



No unjustified 3D: Power of the plane

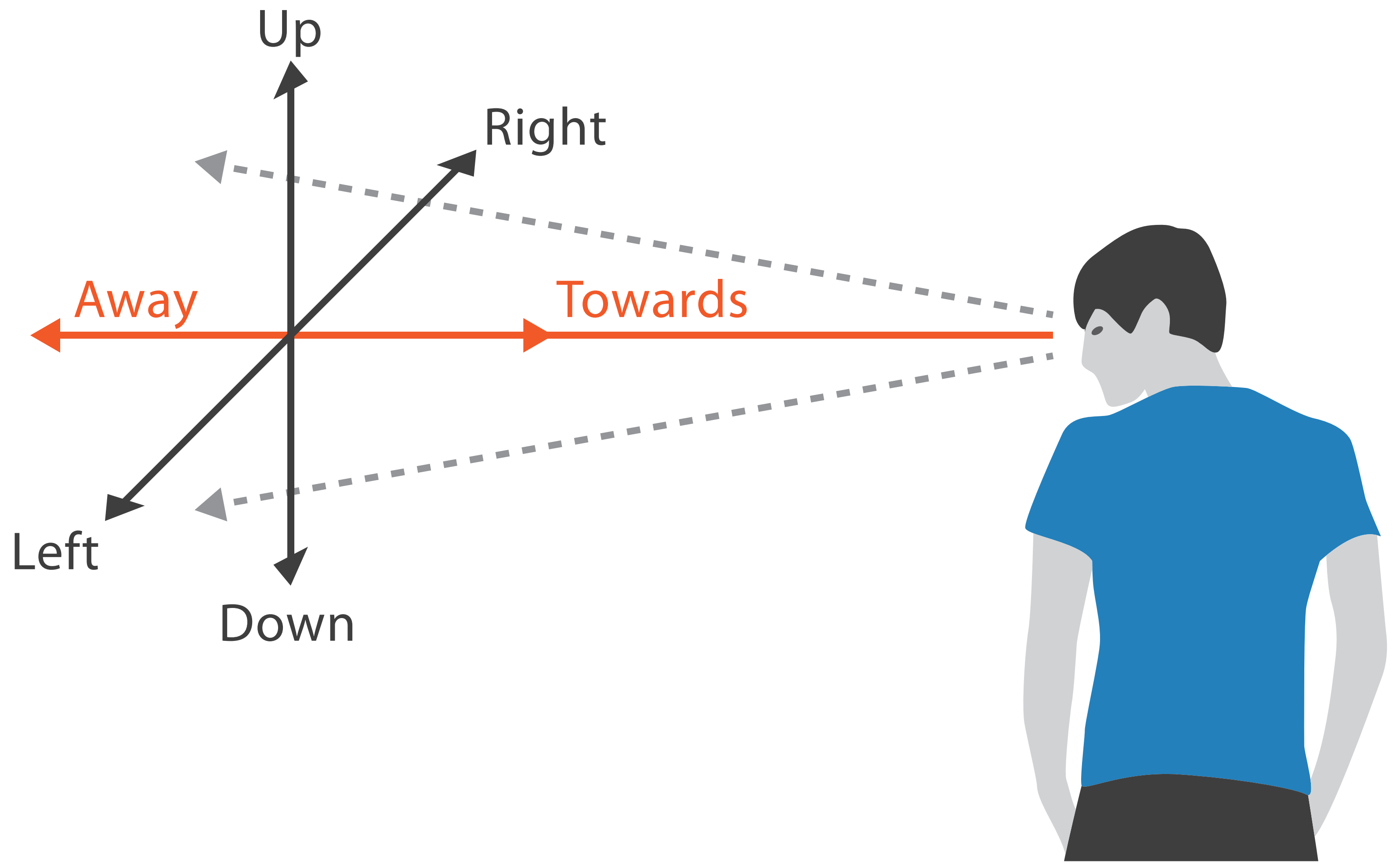
- high-ranked spatial position channels: **planar** spatial position
  - not depth!

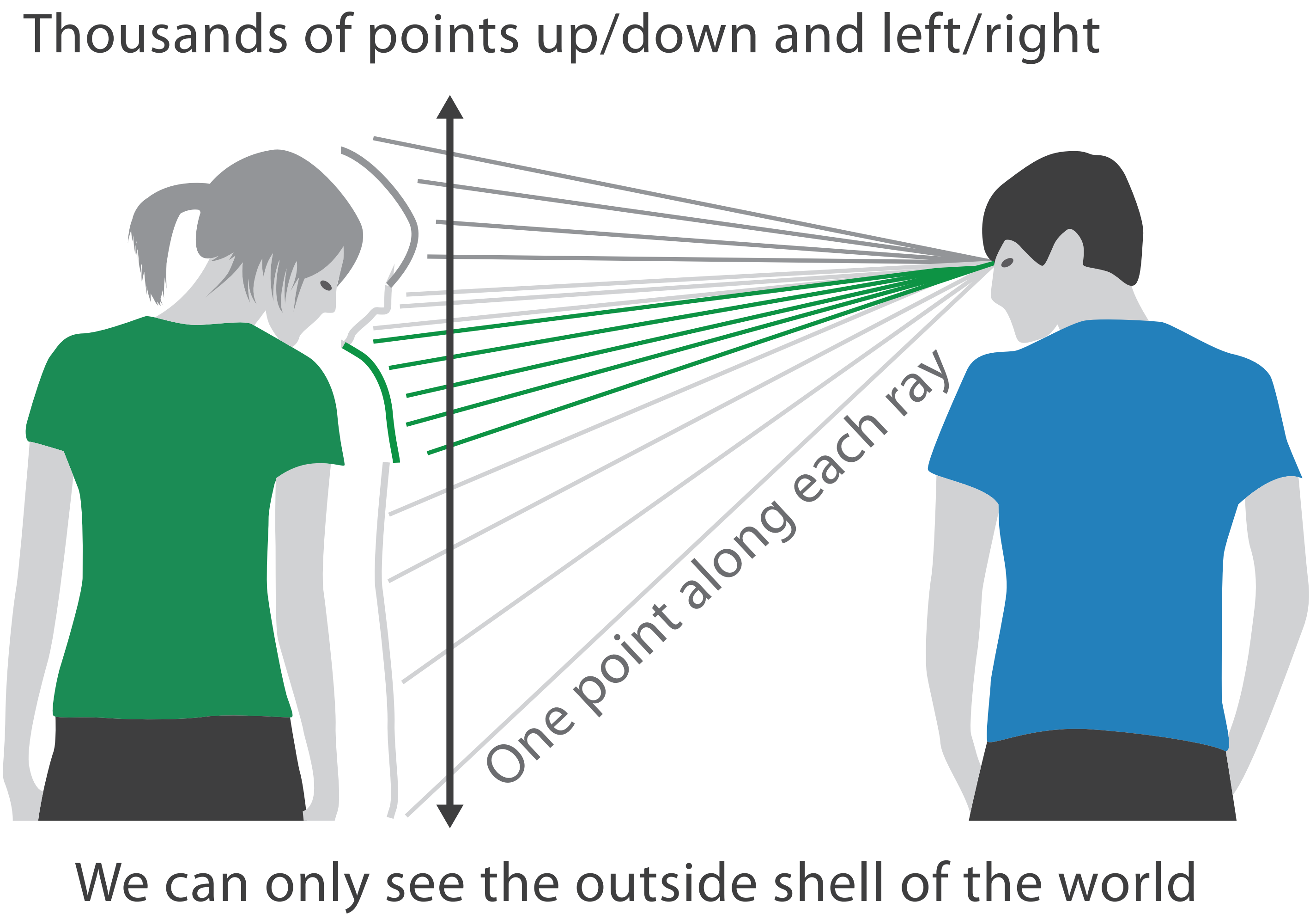




No unjustified 3D: Danger of depth

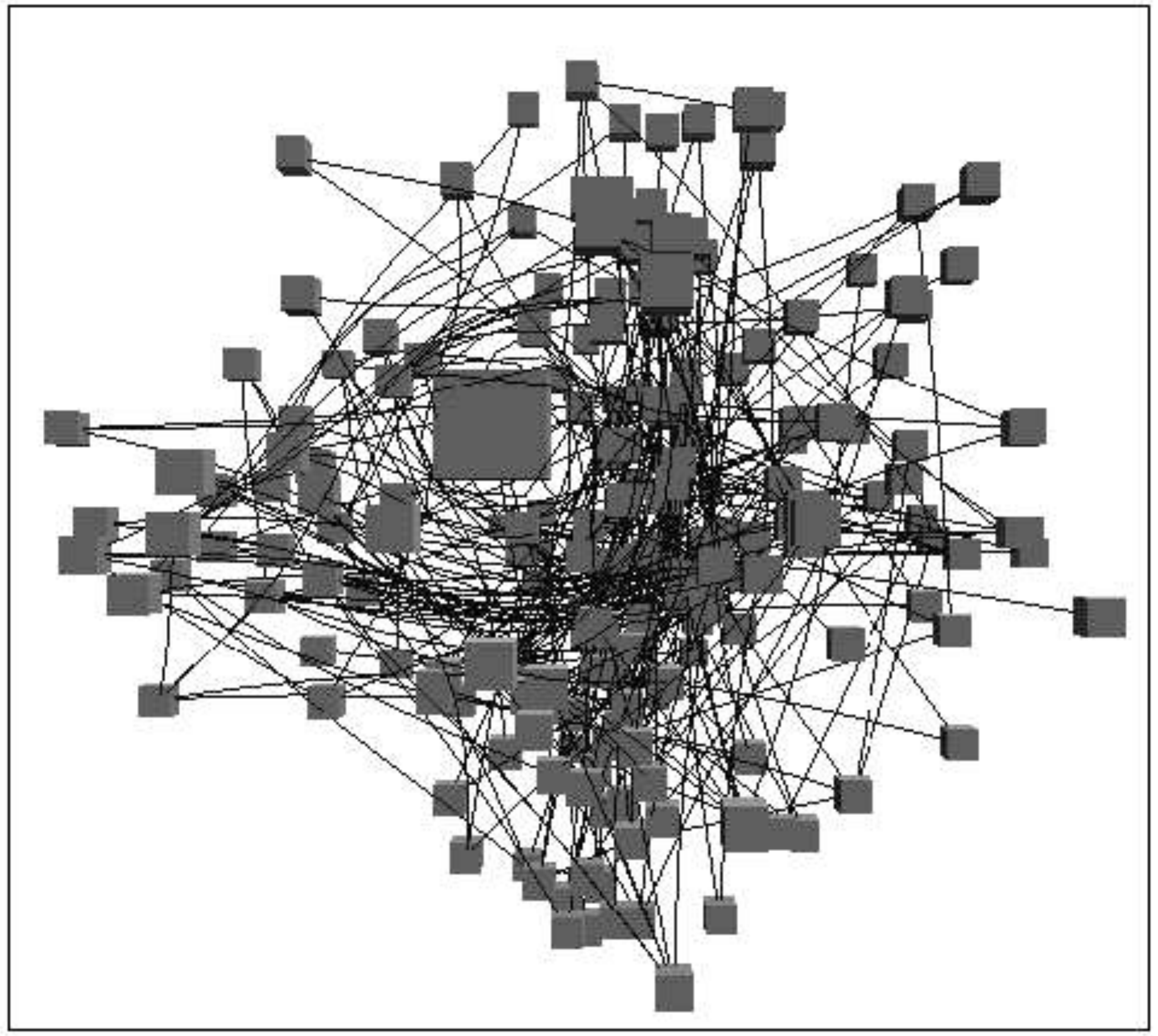
- we don't really live in 3D: we **see** in 2.05D
  - acquire more info on image plane quickly from eye movements
  - acquire more info for depth slower, from head/body motion





Occlusion hides information

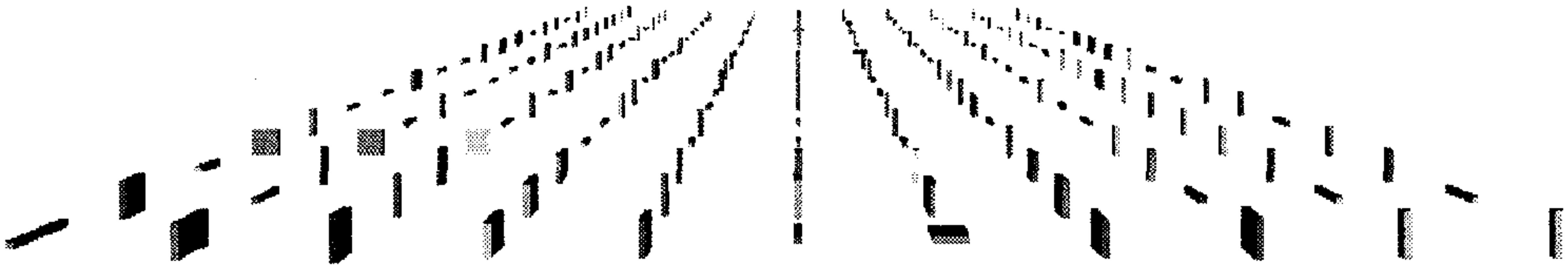
- occlusion
- interaction can resolve, but at cost of time and cognitive load



[Distortion Viewing Techniques for 3D Data. Carpendale et al. InfoVis 1996.]

Perspective distortion loses information

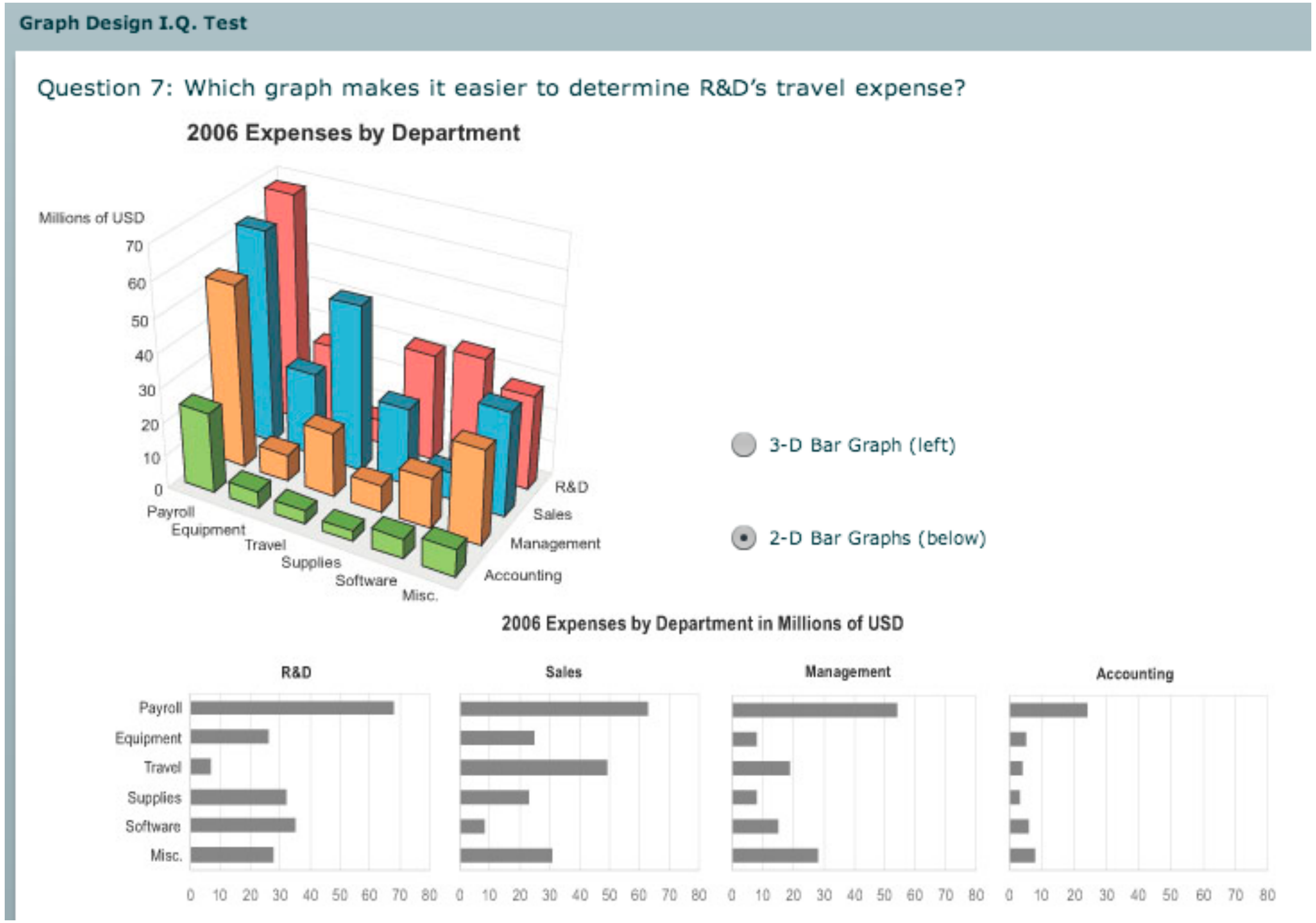
- perspective distortion
  - interferes with all size channel encodings
  - power of the plane is lost!



[Visualizing the Results of Multimedia Web Search Engines. Mukherjea, Hirata, and Hara. InfoVis 96]

3D vs 2D bar charts

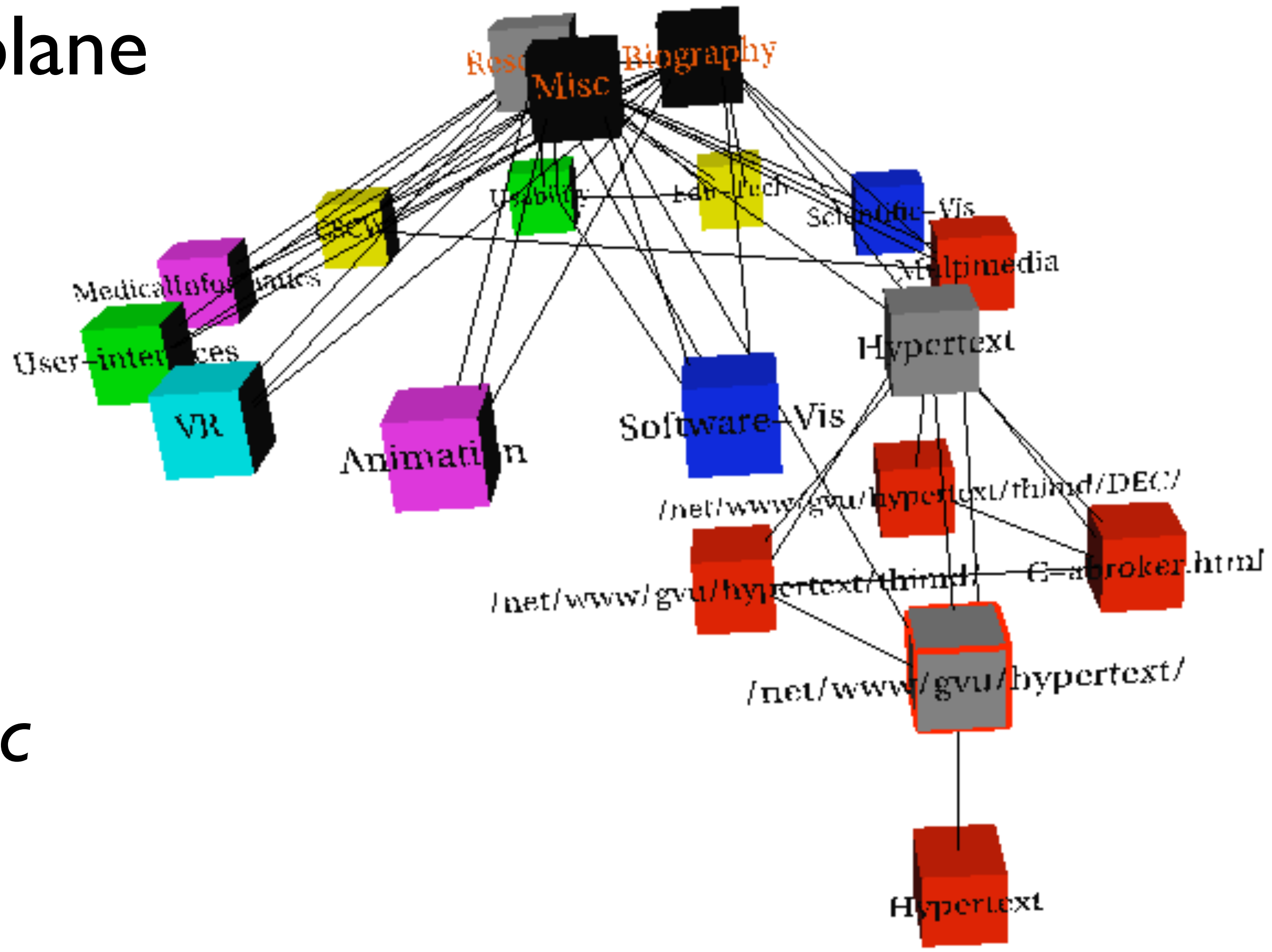
- 3D bars very difficult to justify!
  - perspective distortion
  - occlusion
- faceting into 2D almost always better choice



[http://perceptualedge.com/files/GraphDesignIQ.html]

Tilted text isn't legible

- text legibility
  - –far worse when tilted from image plane



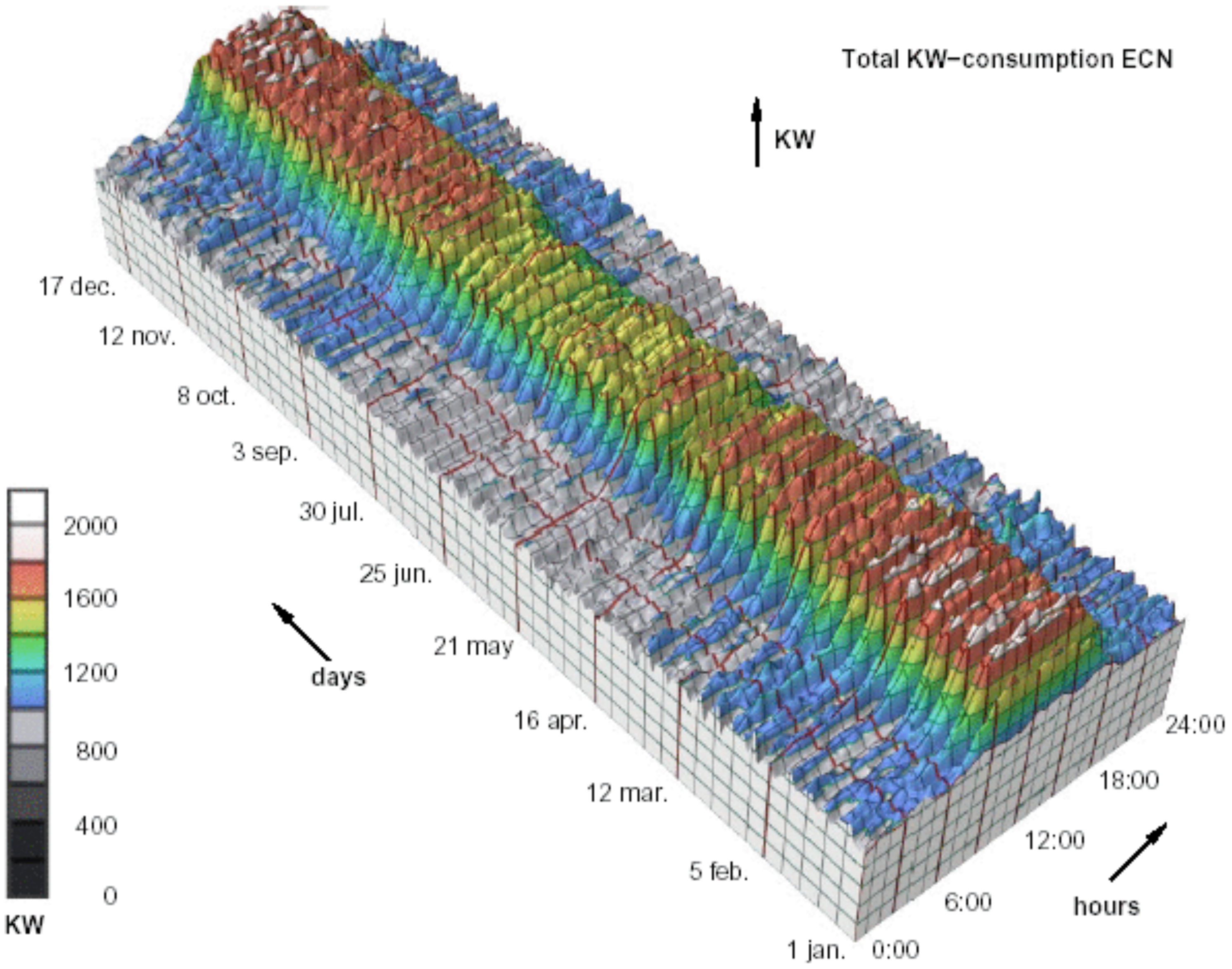
- further reading

[Exploring and Reducing the Effects of Orientation on Text Readability in Volumetric Displays. Grossman et al. CHI 2007]

[Visualizing the World-Wide Web with the Navigational View Builder. Mukherjea and Foley. Computer Networks and ISDN Systems, 1995.]

No unjustified 3D example: Time-series data

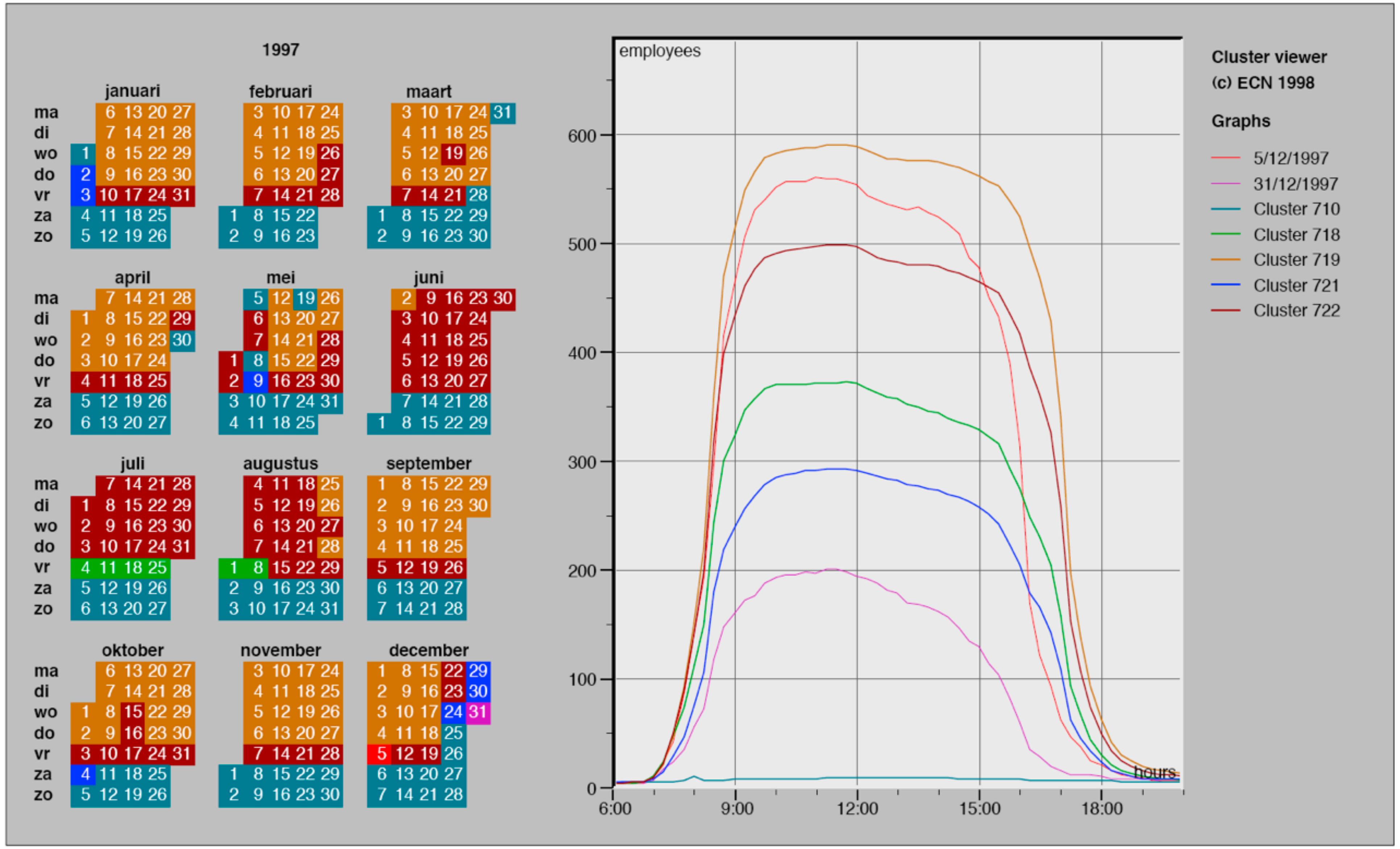
- extruded curves: detailed comparisons impossible



[Cluster and Calendar based Visualization of Time Series Data. van Wijk and van Selow, Proc. InfoVis 99.]

No unjustified 3D example: Transform for new data abstraction

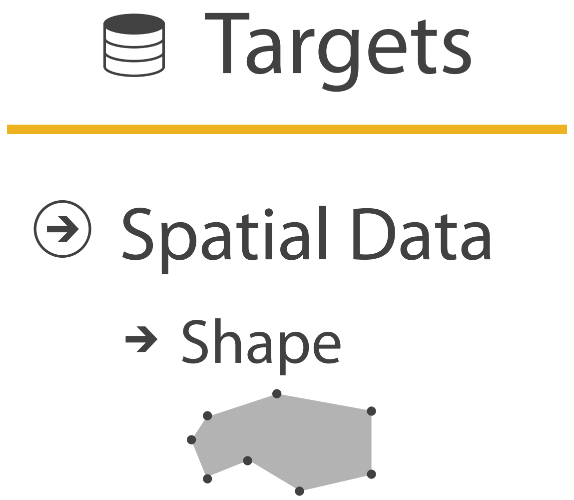
- derived data: cluster hierarchy
- juxtapose multiple views: calendar, superimposed 2D curves

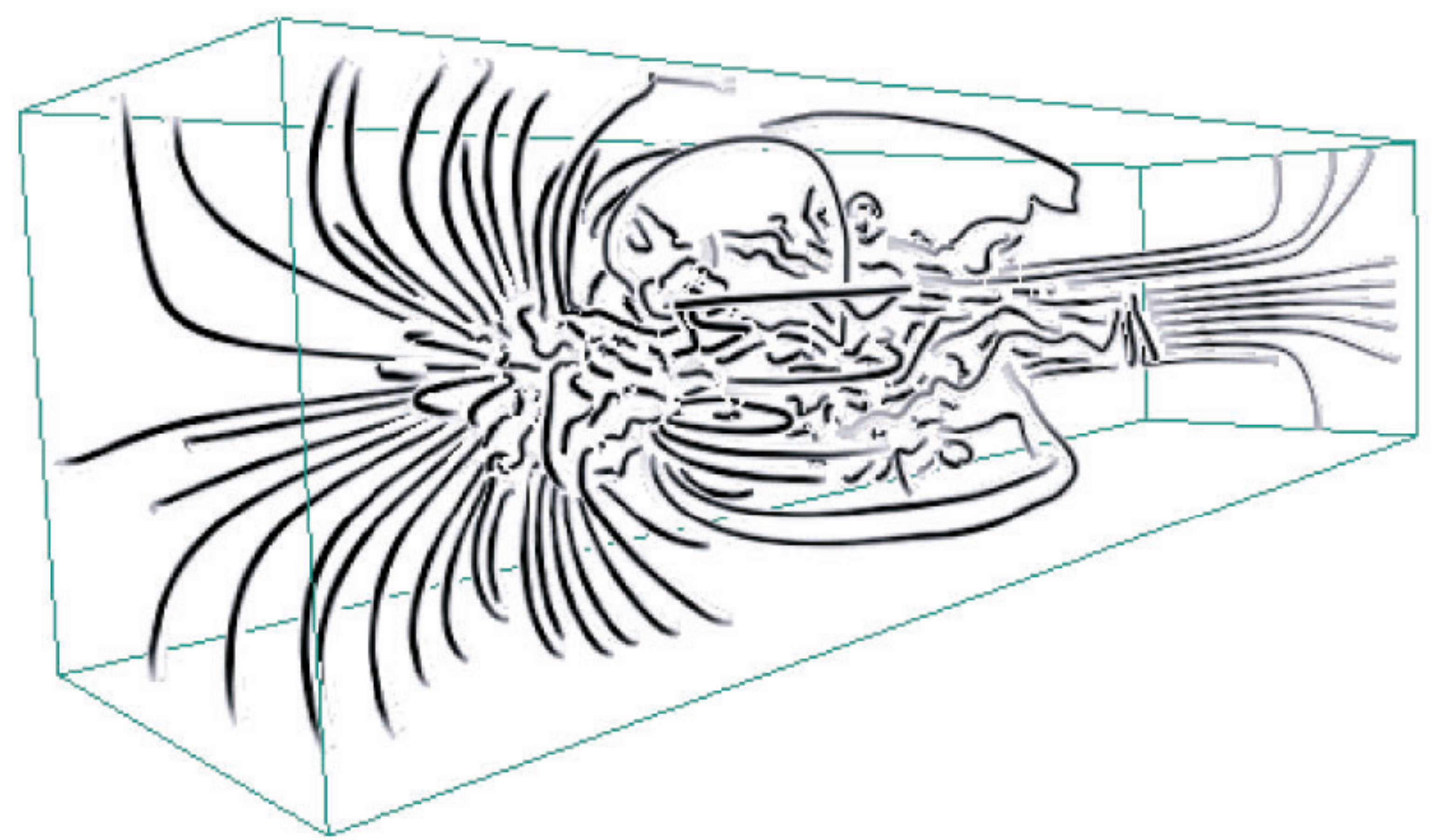


[Cluster and Calendar based Visualization of Time Series Data. van Wijk and van Selow, Proc. InfoVis 99.]

Justified 3D: shape perception

- benefits outweigh costs when task is shape perception for 3D spatial data
  - interactive navigation supports synthesis across many viewpoints

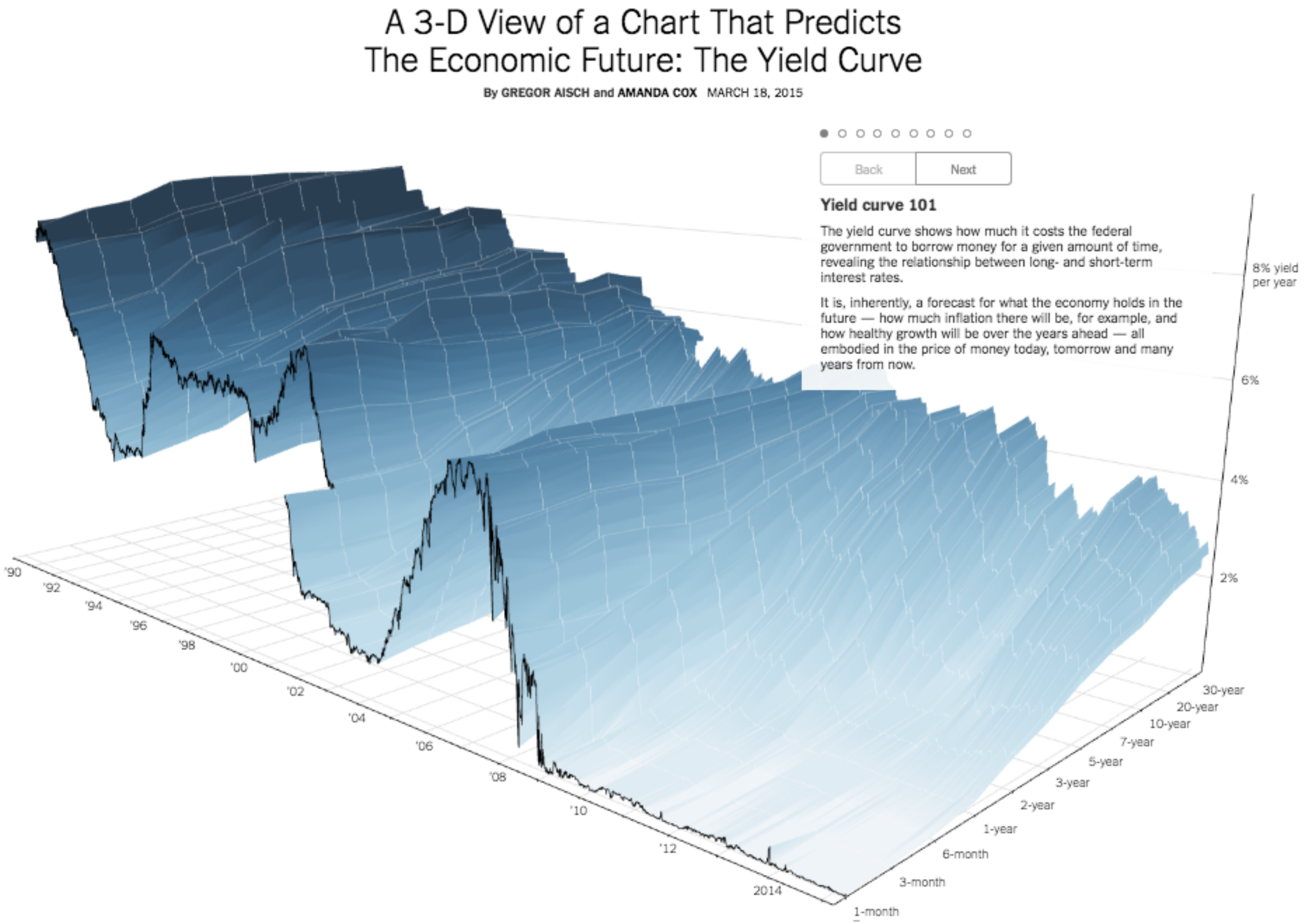




[Image-Based Streamline Generation and Rendering. Li and Shen. IEEE Trans. Visualization and Computer Graphics (TVCG) 13:3 (2007), 630–640.]

Justified 3D: Economic growth curve

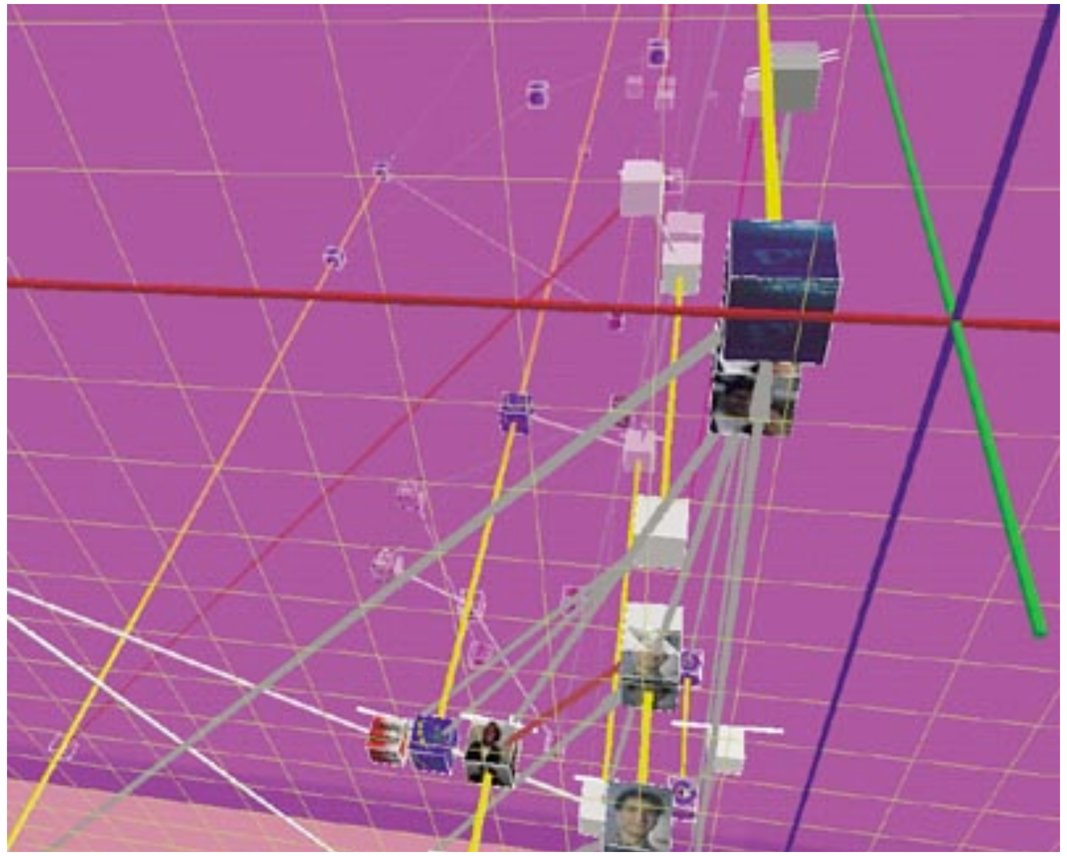
- constrained navigation steps through carefully designed viewpoints



http://www.nytimes.com/interactive/2015/03/19/upshot/3d-yield-curve-economic-growth.html

No unjustified 3D

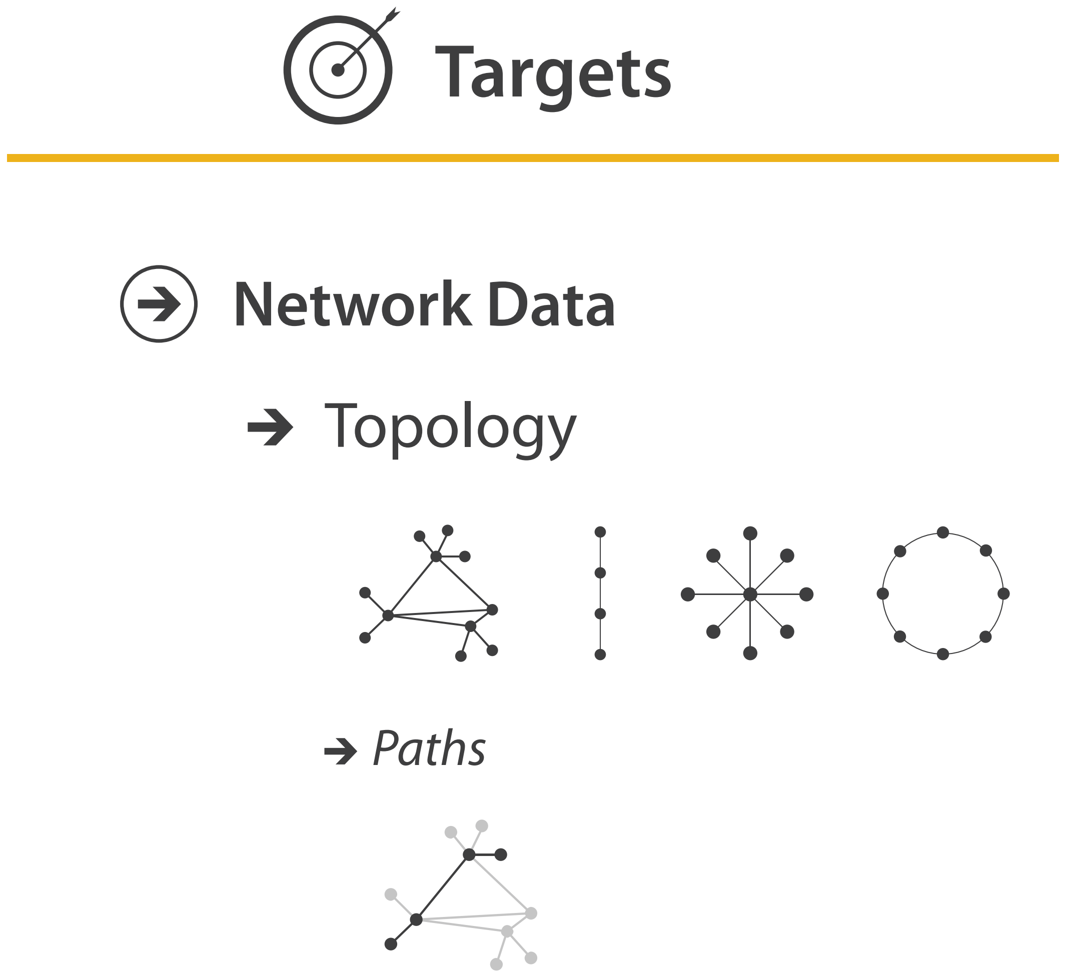
- 3D legitimate for true 3D spatial data
- 3D needs very careful justification for abstract data
  - enthusiasm in 1990s, but now skepticism
  - be especially careful with 3D for point clouds or networks



[WEBPATH-a three dimensional Web history. Frecon and Smith. Proc. InfoVis 1999]

No unjustified 2D

- consider whether network data requires 2D spatial layout
  - especially if reading text is central to task!
  - arranging as network means lower information density and harder label lookup compared to text lists
- benefits outweigh costs when topological structure/context important for task
  - be especially careful for search results, document collections, ontologies



Eyes beat memory

- principle: external cognition vs. internal memory
  - easy to compare by moving eyes between side-by-side views
  - harder to compare visible item to memory of what you saw
- implications for animation
  - great for choreographed storytelling
  - great for transitions between two states
  - poor for many states with changes everywhere
    - consider small multiples instead

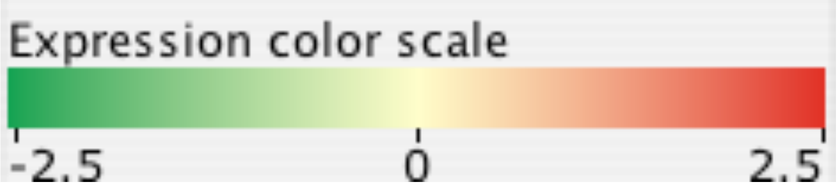
literal — abstract

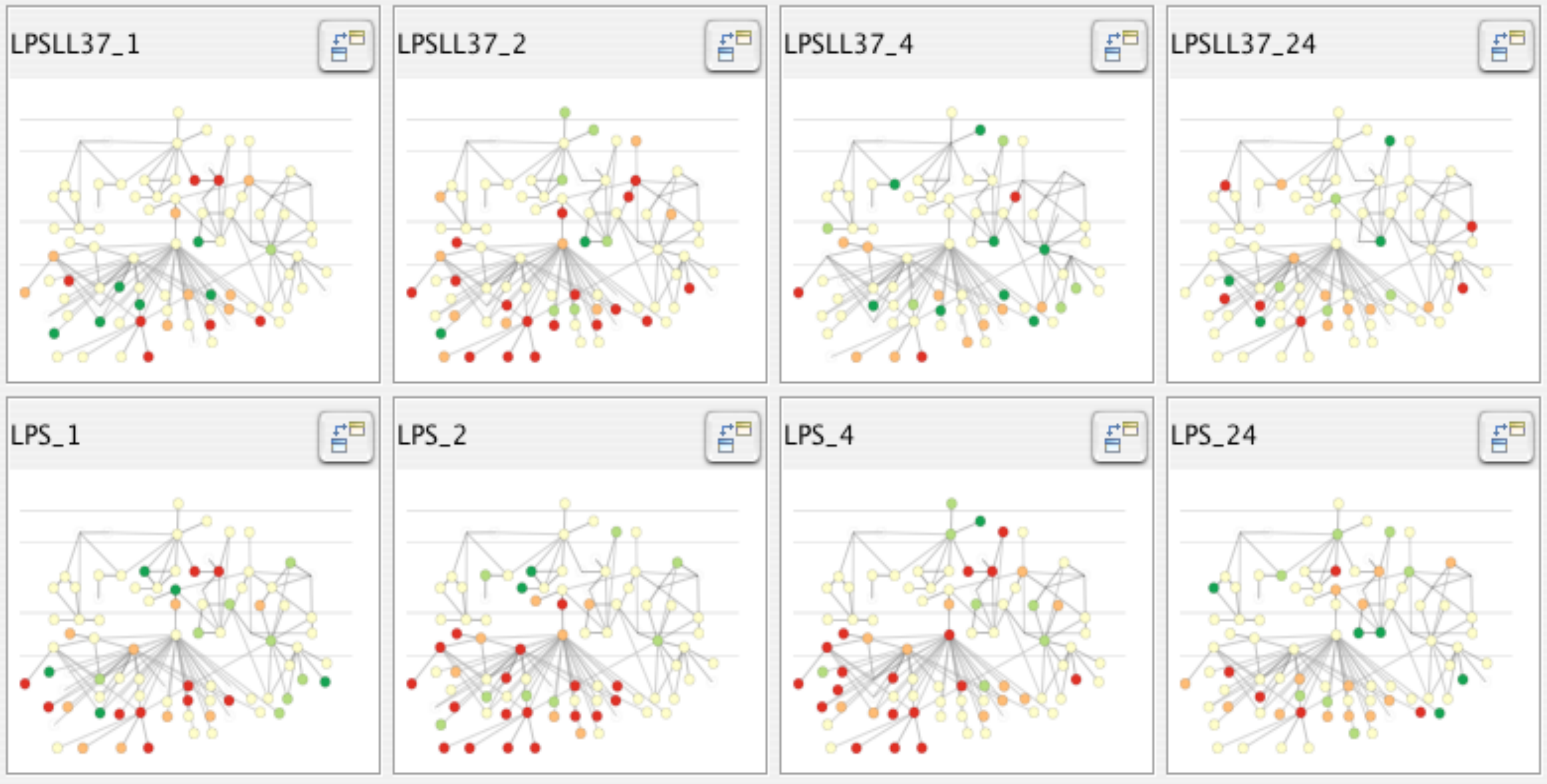
animation — small multiples

show time with time ⟷ show time with space

Eyes beat memory example: Cerebral

- small multiples: one graph instance per experimental condition
  - same spatial layout
  - color differently, by condition



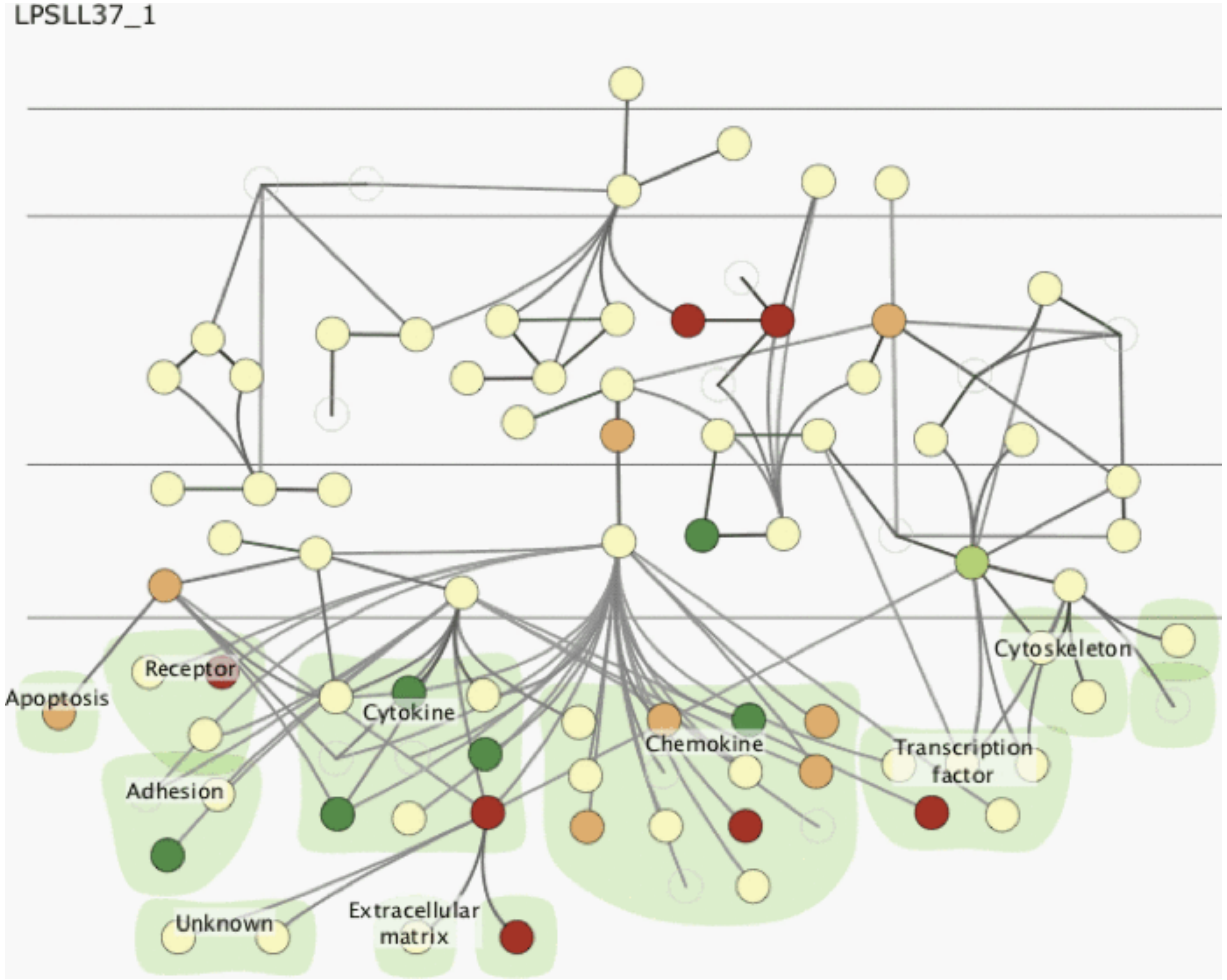


[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Kincaid. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2008) 14:6 (2008), 1253–1260.]

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

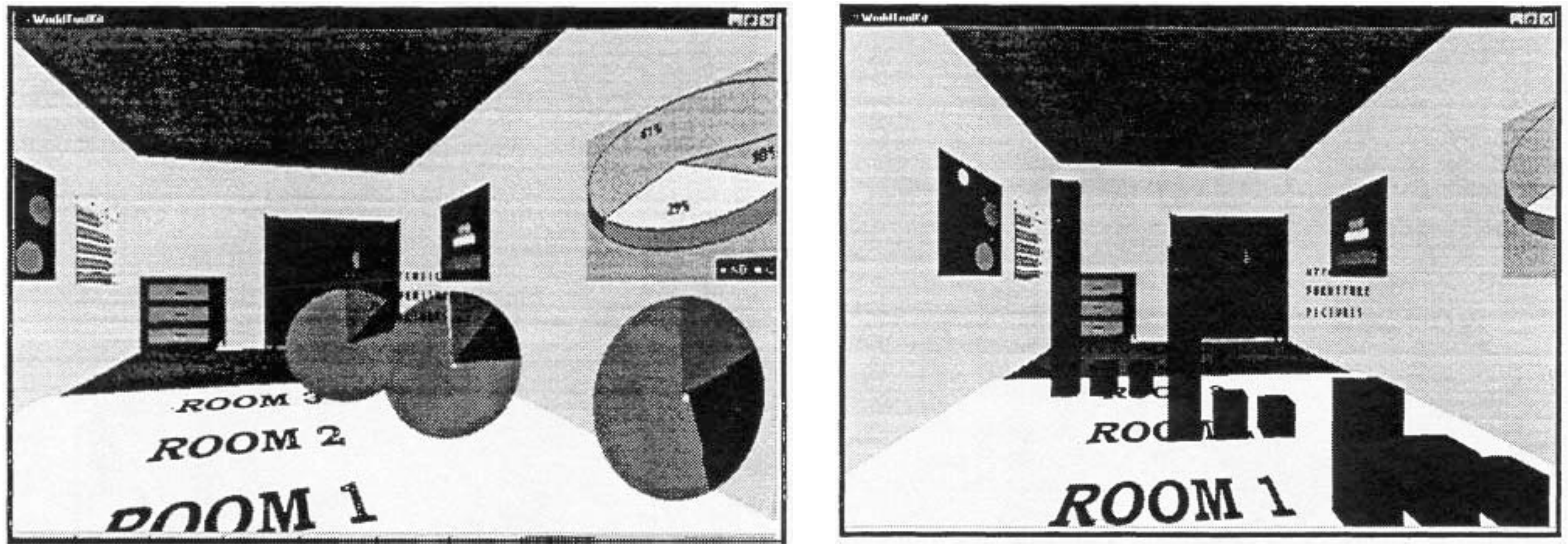


Change blindness

- if attention is directed elsewhere, even drastic changes not noticeable
  - door experiment
- change blindness demos
  - mask in between images

Resolution beats immersion

- immersion typically not helpful for abstract data
  - do not need sense of presence or stereoscopic 3D
  - desktop also better for workflow integration
- resolution much more important: pixels are the scarcest resource
- virtual reality for abstract data difficult to justify thus far
  - but stay tuned with second wave



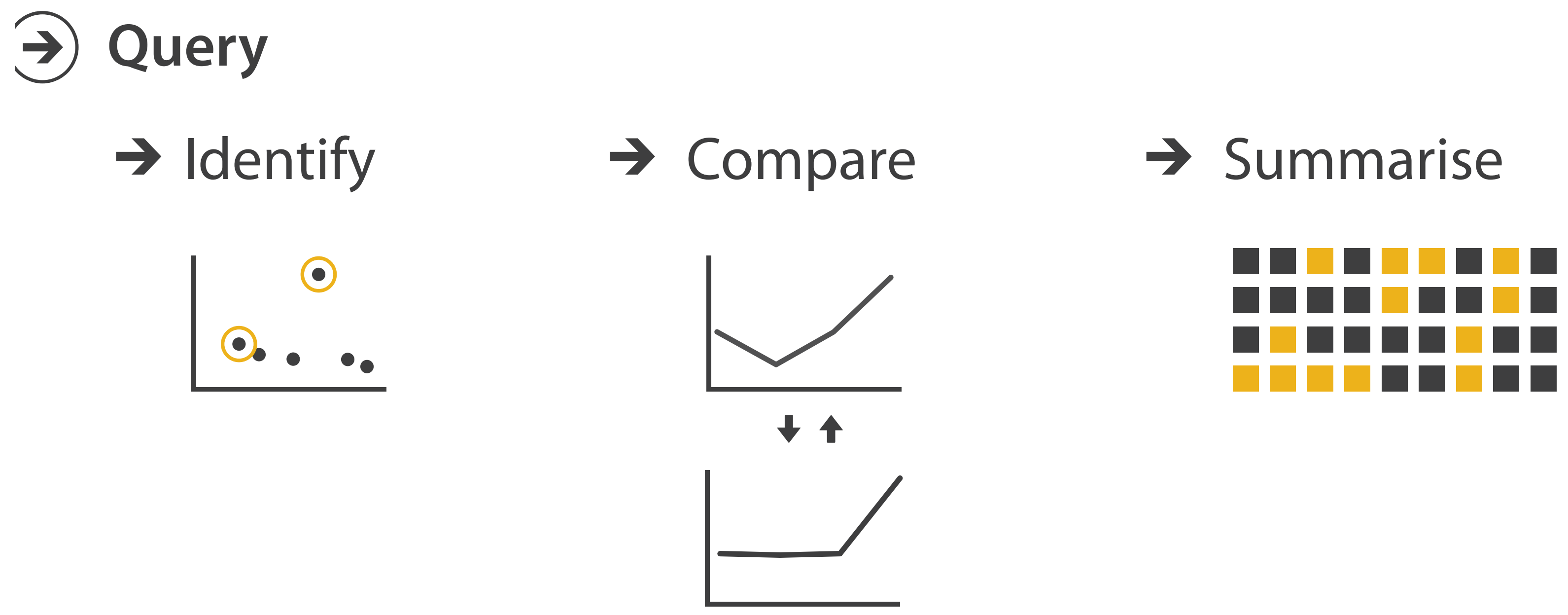
[Development of an information visualization tool using virtual reality. Kirner and Martins. Proc. Symp. Applied Computing 2000]

Overview first, zoom and filter, details on demand

- influential mantra from Shneiderman

[The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Shneiderman. Proc. IEEE Visual Languages, pp. 336–343, 1996.]

- overview = summary
  - –microcosm of full vis design problem



Rule of thumb: **Responsiveness is required**

- *visual feedback: three rough categories*
  - *0.1 seconds: perceptual processing*
    - subsecond response for mouseover highlighting - ballistic motion
  - *1 second: immediate response*
    - fast response after mouseclick, button press - Fitts' Law limits on motor control
  - *10 seconds: brief tasks*
    - bounded response after dialog box - mental model of heavyweight operation (file load)
- scalability considerations
  - highlight selection without complete redraw of view (graphics frontbuffer)
  - show hourglass for multi-second operations (check for cancel/undo)
  - show progress bar for long operations (process in background thread)
  - rendering speed when item count is large (guaranteed frame rate)

Function first, form next

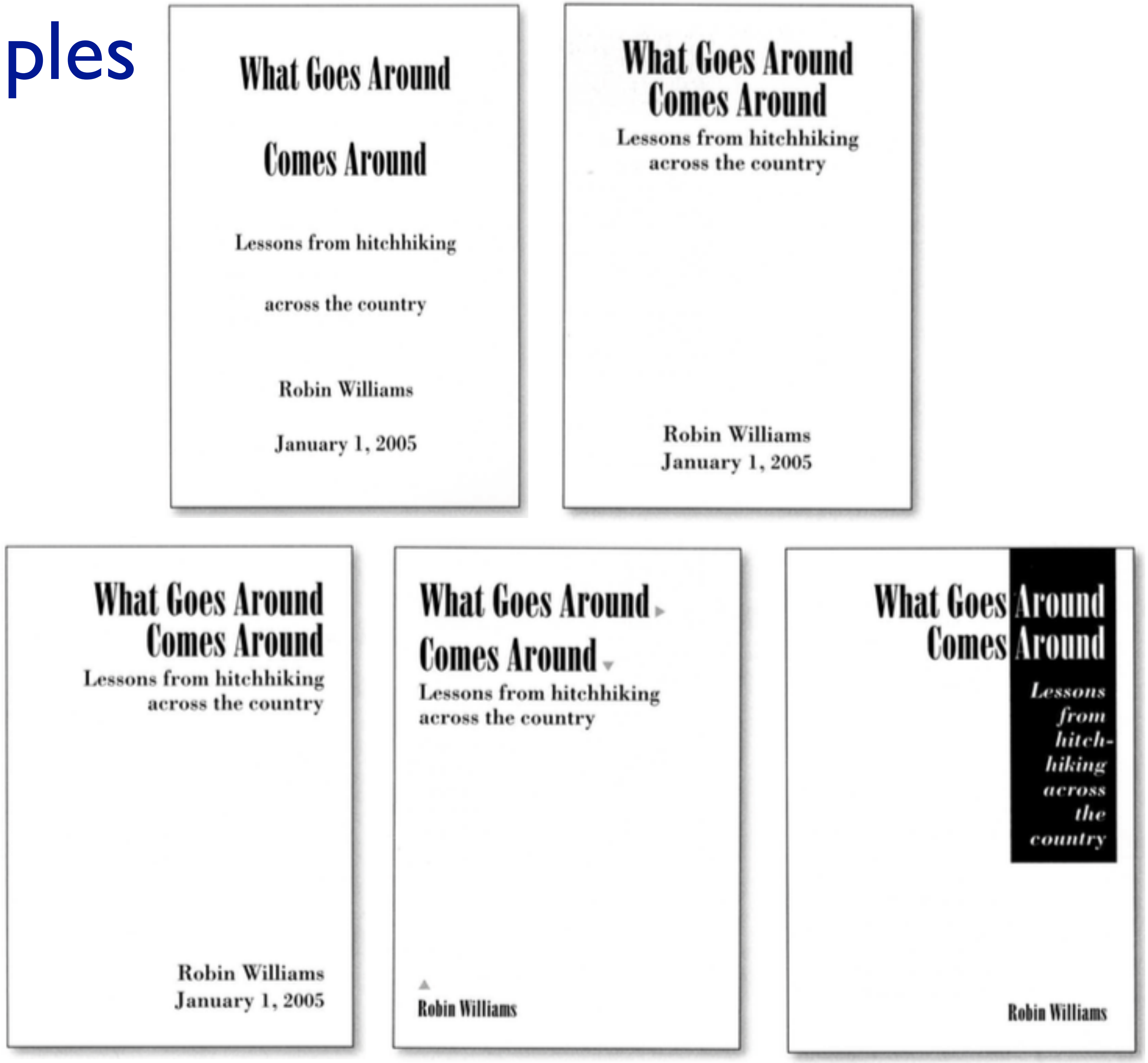
- start with focus on functionality
  - possible to improve aesthetics later on, as refinement
  - if no expertise in-house, find good graphic designer to work with
  - aesthetics do matter: another level of function
    - visual hierarchy, alignment, flow
    - Gestalt principles in action

- dangerous to start with aesthetics
  - usually impossible to add function retroactively

[The Non-Designer's Design Book. Robin Williams. 3rd edition. Peachpit Press, 2008.]

Form: Basic graphic design principles

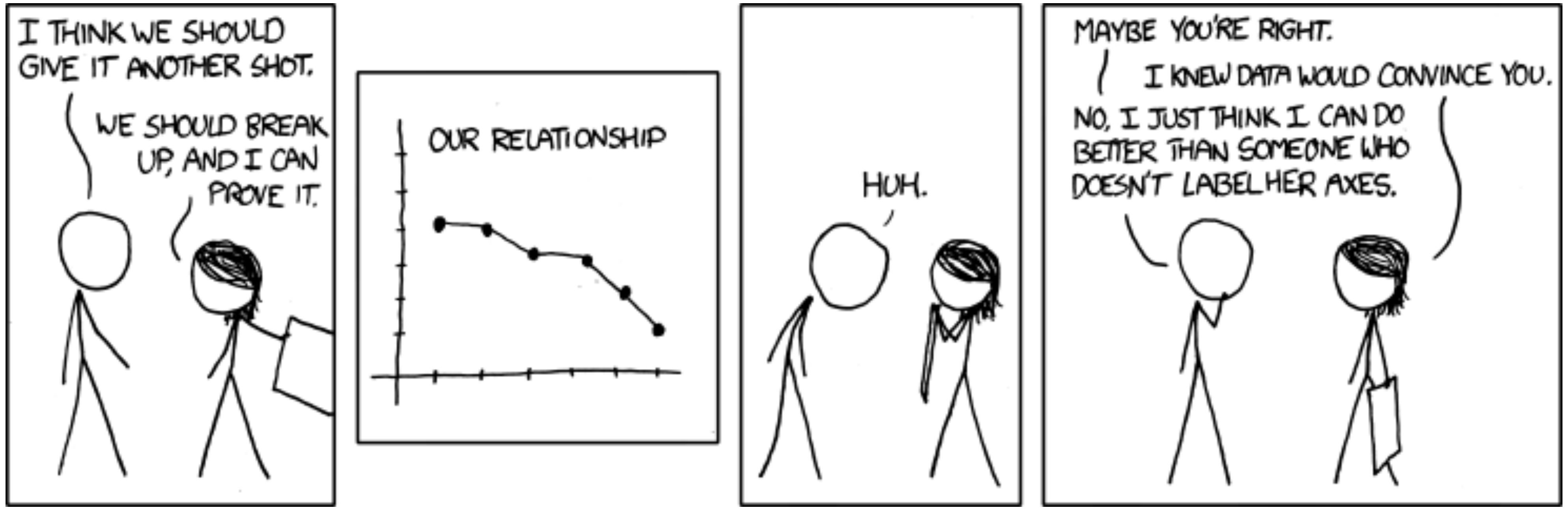
- proximity
  - do group related items together
  - avoid equal whitespace between unrelated
- alignment
  - do find/make strong line, stick to it
  - avoid automatic centering
- repetition
  - do unify by pushing existing consistencies
- contrast
  - if not identical, then very different
  - avoid similar



- buy now and read cover to cover - very practical, worth your time, fast read!
  The Non-Designer's Design Book, 4th ed. Robin Williams, Peachpit Press, 2015.

Best practices: Labelling

- make visualizations as self-documenting as possible
  - meaningful & useful title, labels, legends
    - axes and panes/subwindows should have labels
      - and axes should have good mix/max boundary tick marks
    - everything that's plotted should have a legend
      - and own header/labels if not redundant with main title
    - use reasonable numerical format
      - avoid scientific notation in most cases



[https://xkcd.com/833/]

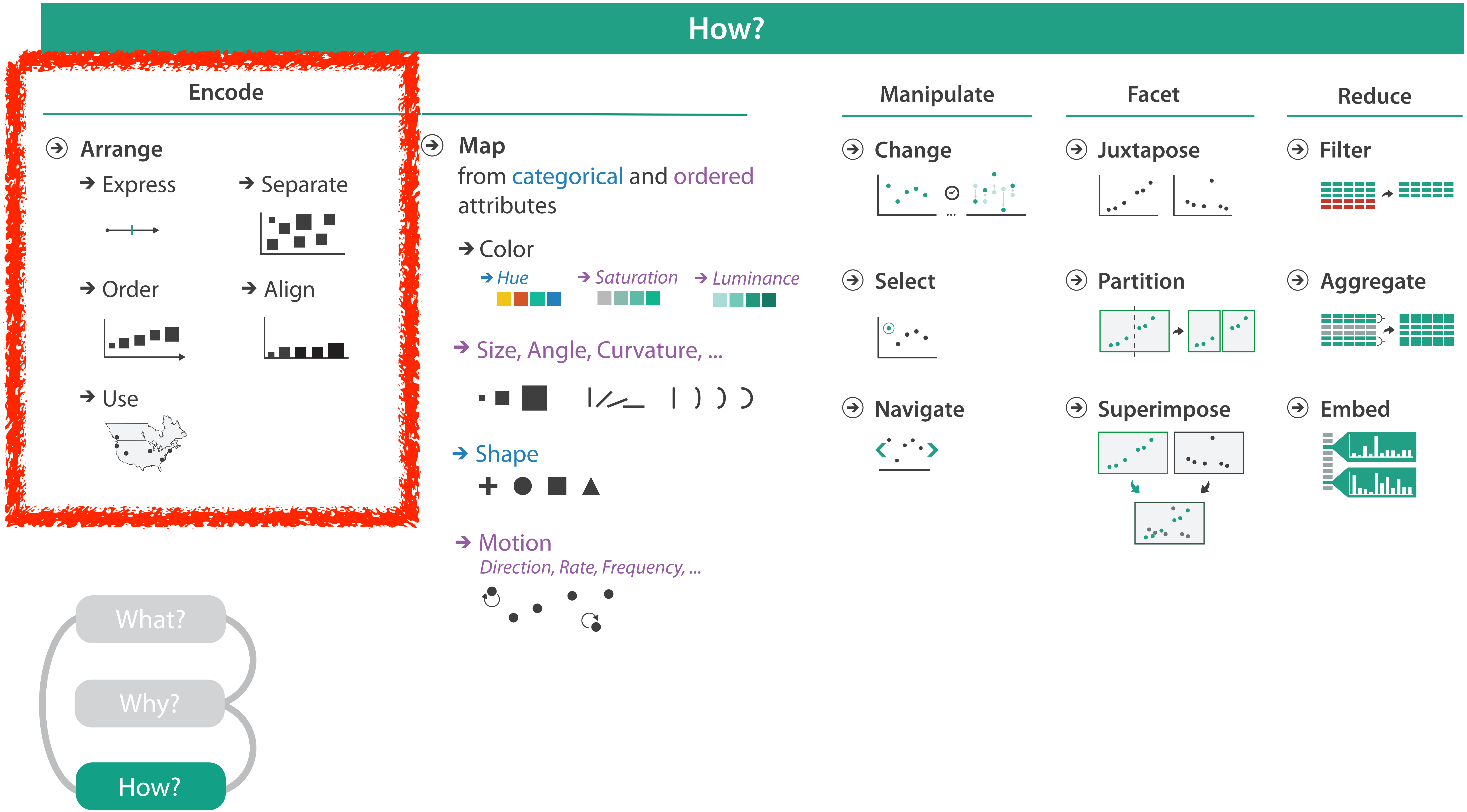
Further reading

- Visualization Analysis and Design. Tamara Munzner. CRC Press, 2014.
  - *Chap 6: Rules of Thumb*
- Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules. Jeff Johnson. Morgan Kaufmann, 2010.
  - *Chap 12: We Have Time Requirements*
- The Non-Designer's Design Book. 3rd edition. Robin Williams. Peachpit Press, 2008.

Further reading, full

- Visual Thinking for Design, Colin Ware, Morgan Kaufmann 2008.
- Information Visualization: Perception for Design, 3rd edition, Colin Ware, Morgan Kaufmann, 2013.
- The Use of 2-D and 3-D Displays for Shape Understanding versus Relative Position Tasks. Mark St. John, Michael B. Cowen, Harvey S. Smallman, and Heather M. Oonk. Human Factors 43:1 (2001), 79-98.
- An Evaluation of Cone Trees. Andy Cockburn and Bruce McKenzie. In People and Computers XIV: Usability or Else. British Computer Society Conference on Human Computer Interaction, pp. 425-436. Springer, 2000.
- 3D or Not 3D? Evaluating the Effect of the Third Dimension in a Document Management System. Andy Cockburn and Bruce McKenzie. Proc. CHI 2003, p 434-441.
- Evaluating Spatial Memory in Two and Three Dimensions. Andy Cockburn and Bruce McKenzie. International Journal of Human-Computer Studies. 61(30):359-373.
- Supporting and Exploiting Spatial Memory in User Interfaces. Joey Scarr, Andy Cockburn, and Carl Gutwin. Foundations and Trends in Human-Computer Interaction. 2013. 6:1 1-84.
- Principles of Traditional Animation Applied to Computer Animation John Lasseter, Proceedings of SIGGRAPH 87, Computer Graphics, 21(4), pp. 35-44, July 1987.
- Animation: Can It Facilitate? Barbara Tversky, Julie Morrison, Mireille Betrancourt. International Journal of Human Computer Studies 57:4, pp 247-262, 2002.
- Structuring information interfaces for procedural learning. Jeffrey M. Zacks and Barbara Tversky. Journal of Experimental Psychology: Applied, Vol 9(2), Jun 2003, 88-100.
- Effectiveness of Animation in Trend Visualization. George Robertson and Roland Fernandez and Danyel Fisher and Bongshin Lee and John Stasko. IEEE Trans. on Visualization and Computer Graphics 14(6):1325-1332, 2008 (Proc. InfoVis08).
- Current Approaches to Change Blindness. Daniel J. Simons. Visual Cognition 7:1/2/3 (2000), 1-15.
- The eyes have it: A task by data type taxonomy for information visualizations. Ben Shneiderman. Proc. Conf. Visual Languages 1996, p 336-343.
- The Notion of Overview in Information Visualization. Kaspar Hornbaek and Morten Hertzum. International Journal of Human-Computer Studies 69:7-8 (2011), 509-525.
- The Information Visualizer, an Information Workspace. Stuart Card, George Robertson, and Jock Mackinlay. Proc. CHI 1991, p 181-186.
- Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules. Jeff Johnson. Morgan Kaufmann, 2010.
- A Framework of Interaction Costs in Information Visualization. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 08) 14:6 (2008), 1149-1156.
- Toward a Deeper Understanding of the Role of Interaction in Information Visualization. Ji Soo Yi, Youn Ah Kang, John T. Stasko, and Julie A. Jacko. TVCG (Proc. InfoVis 07) 13:6 (2007), 1224-1231.
- Get It Right in Black and White. Maureen Stone. Functional Color, 2010.
- The Non-Designer's Design Book. Robin Williams. Peachpit Press, 2008.

Ch 7. Arrange Tables


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?

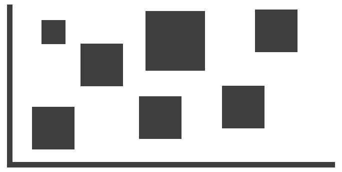
Encode tables: Arrange space

Encode

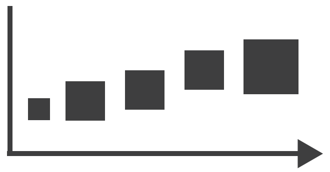
➔ Arrange

- ➔ Express

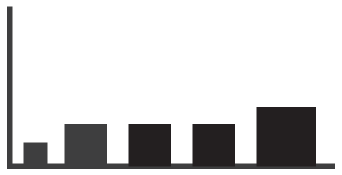
- ➔ Separate



- ➔ Order



- ➔ Align

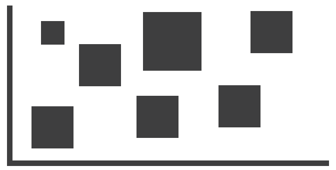


Arrange tables

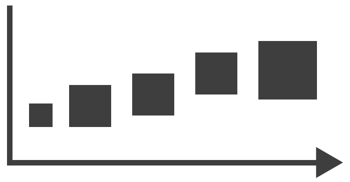
Express Values

Separate, Order, Align Regions

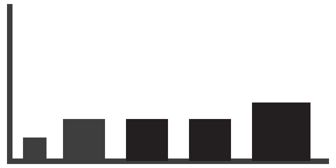
- Separate



- Order

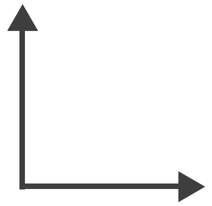


- Align

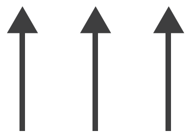


Axis Orientation

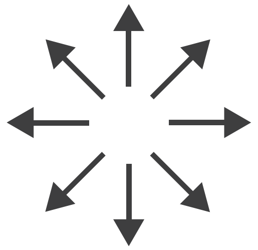
- Rectilinear



- Parallel

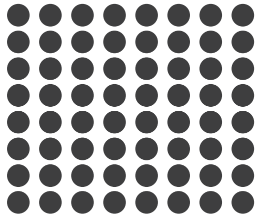


- Radial



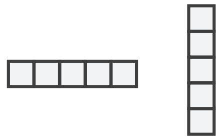
Layout Density

- Dense

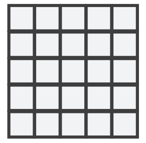


- Space-Filling

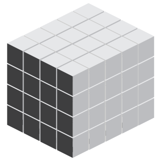
- *1 Key*
  List



- *2 Keys*
  Matrix



- *3 Keys*
  Volume

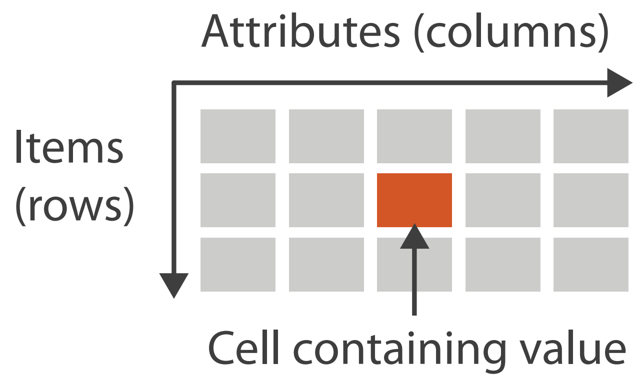


- *Many Keys*
  Recursive Subdivision

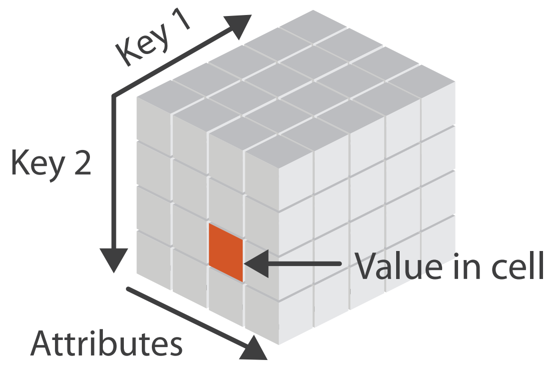
Keys and values

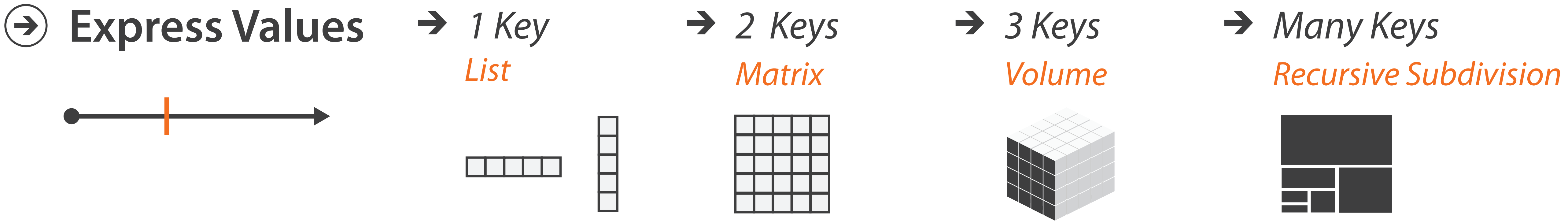
- key
  - independent attribute
  - used as unique index to look up items
  - simple tables: 1 key
  - multidimensional tables: multiple keys
- value
  - dependent attribute, value of cell
- classify arrangements by key count
  - 0, 1, 2, many...

→ Tables

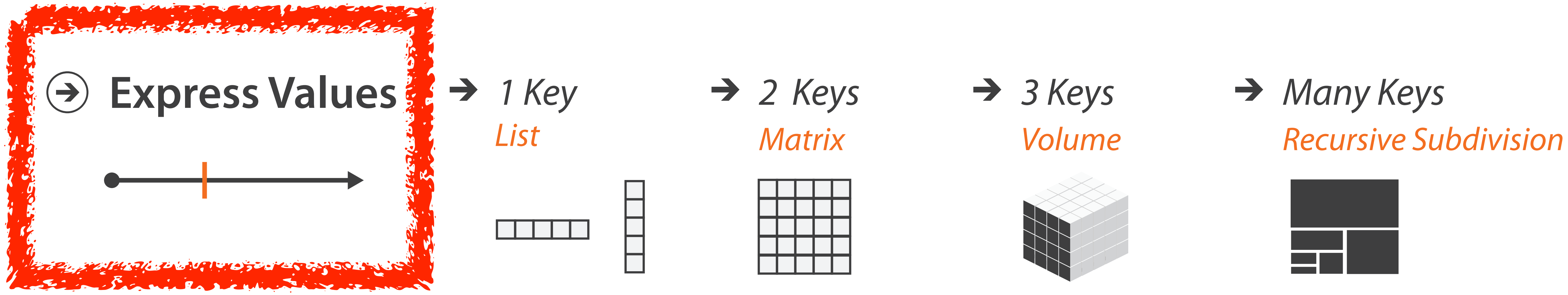


→ *Multidimensional Table*

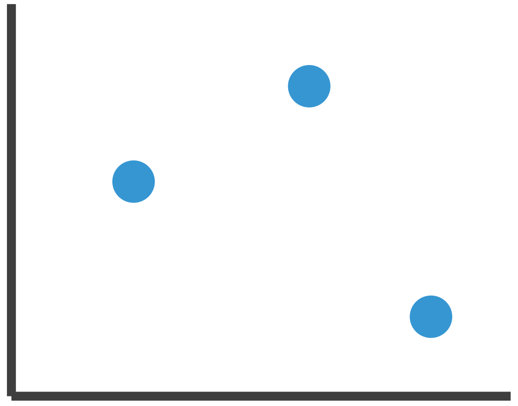




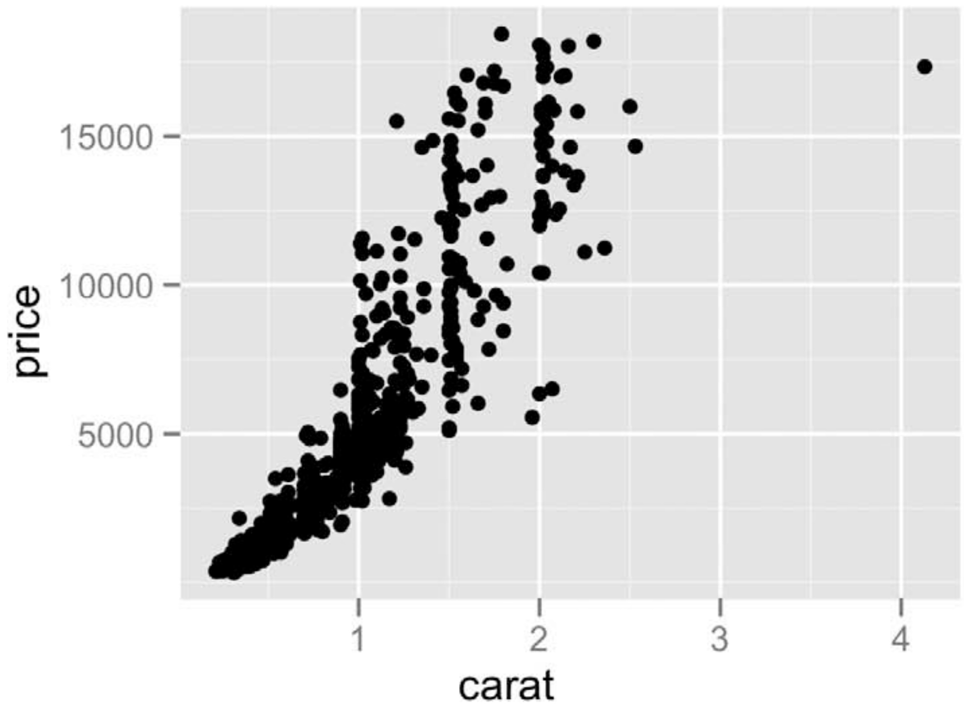
0 Keys

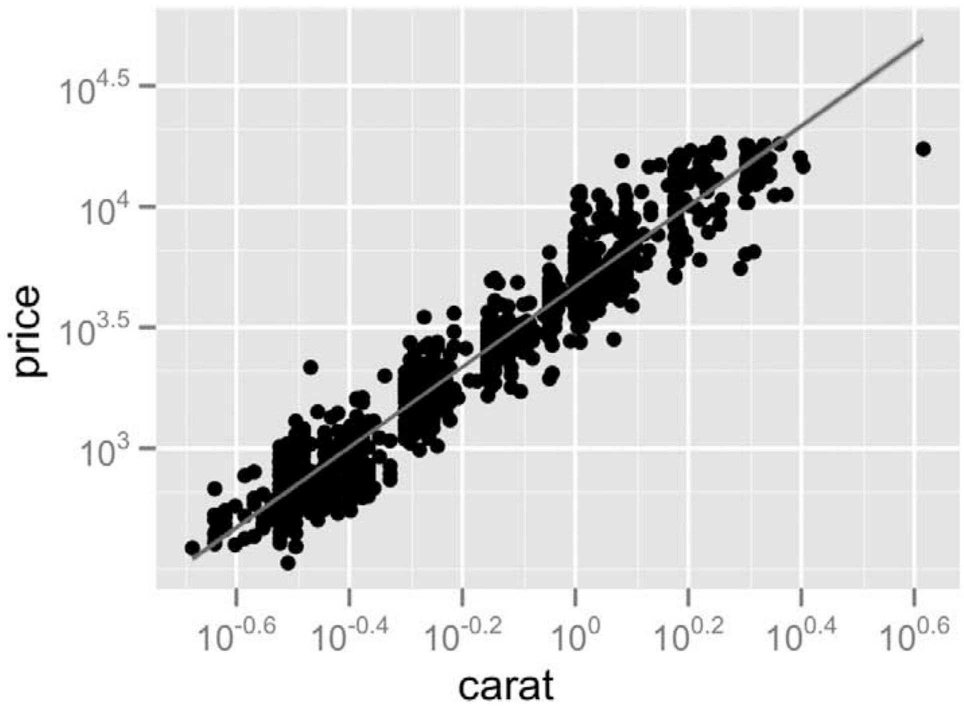

Express Values
1 Key
List
2 Keys
Matrix
3 Keys
Volume
Many Keys
Recursive Subdivision

Idiom: **scatterplot**



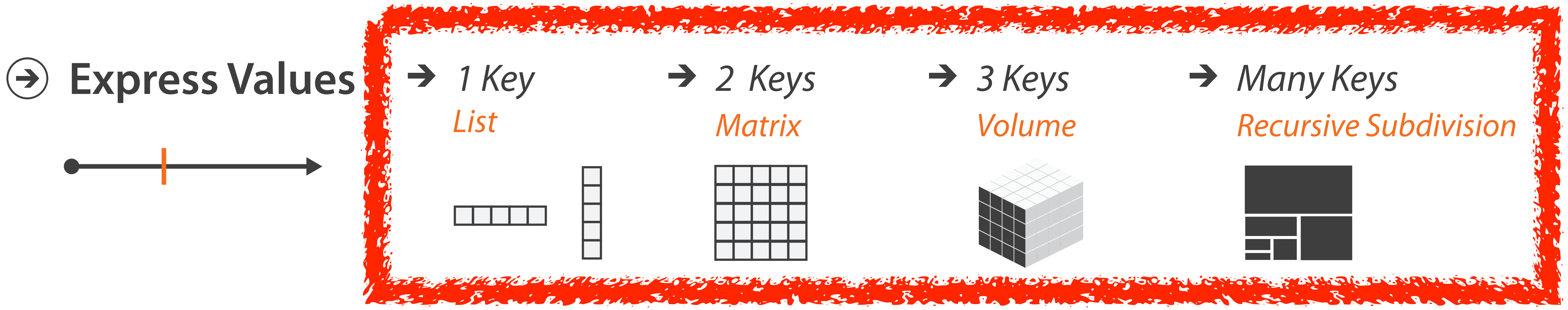
- ***express*** values
  - quantitative attributes
- no keys, only values
  - data
    - 2 quant attribs
  - mark: points
  - channels
    - horiz + vert position
  - tasks
    - find trends, outliers, distribution, correlation, clusters
  - scalability
    - hundreds of items





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

Some keys

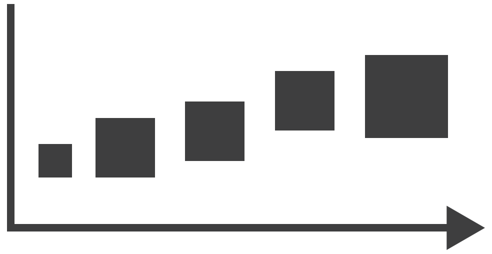


Express Values
1 Key
List
2 Keys
Matrix
3 Keys
Volume
Many Keys
Recursive Subdivision

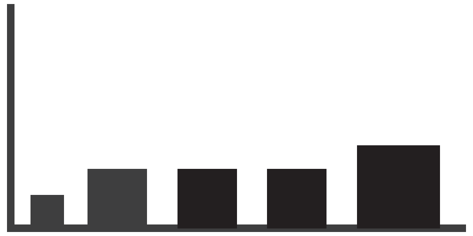
Some keys: Categorical regions

➔ Separate

➔ Order

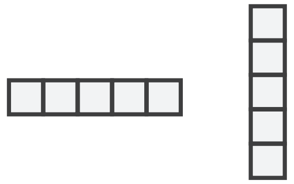


➔ Align

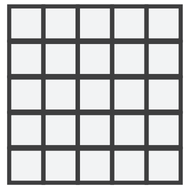


- **regions**: contiguous bounded areas distinct from each other
  - using space to ***separate*** (proximity)
  - following expressiveness principle for categorical attributes
- use ordered attribute to ***order*** and ***align*** regions

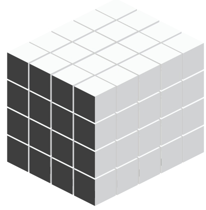
➔ *1 Key*
List



➔ *2 Keys*
Matrix



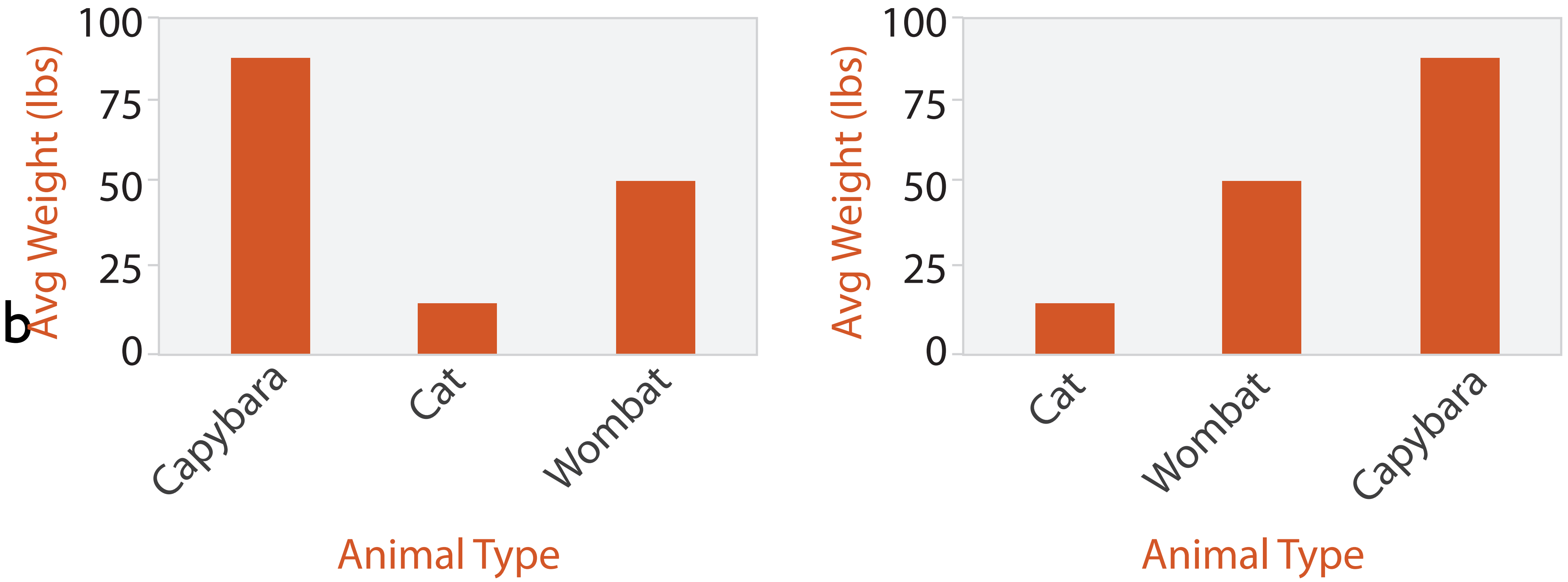
➔ *3 Keys*
Volume



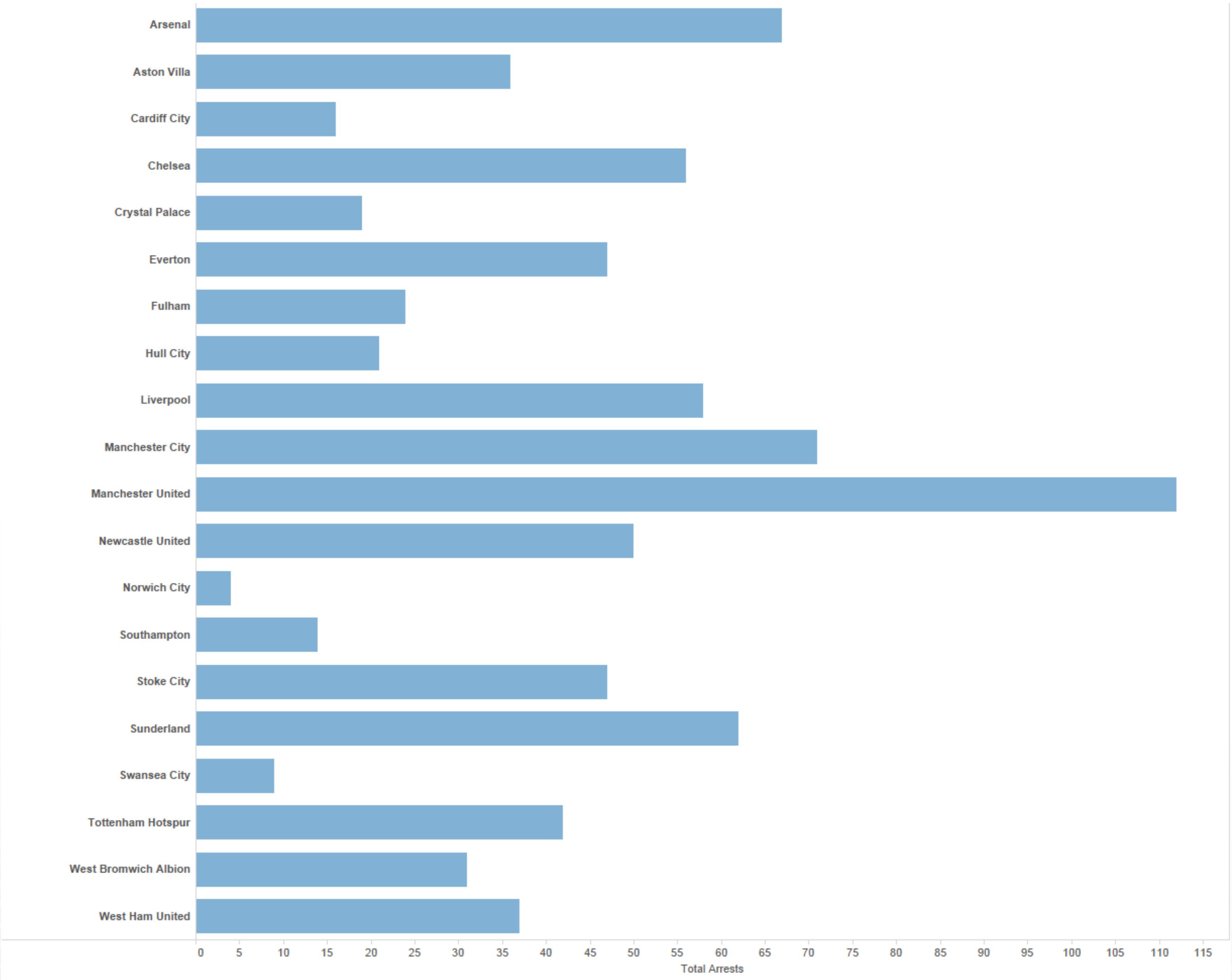
➔ *Many Keys*
Recursive Subdivision

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



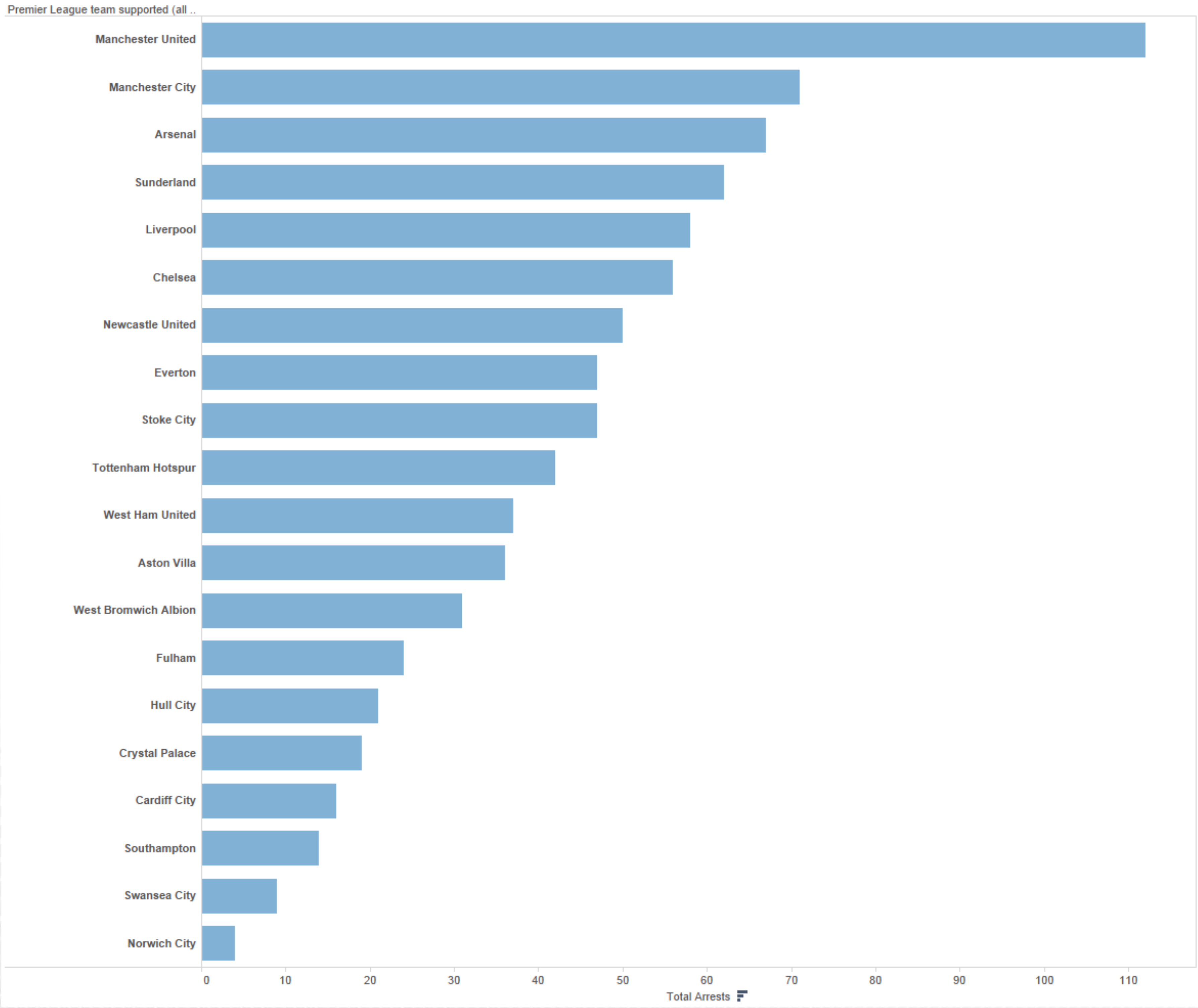
Separated and Aligned but not Ordered



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

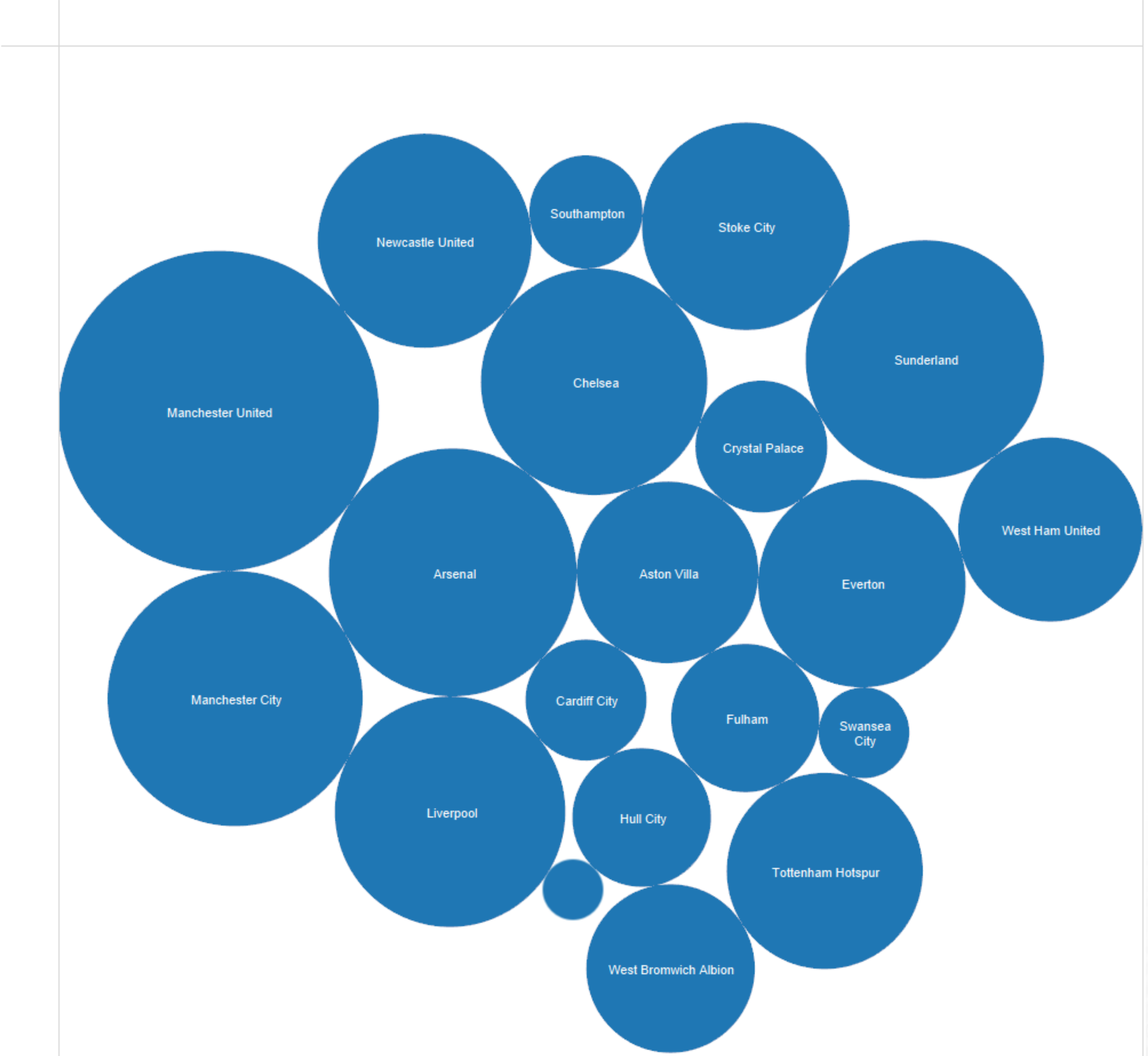
[Slide courtesy of Ben Jones]

Separated, Aligned and Ordered



[Slide courtesy of Ben Jones]

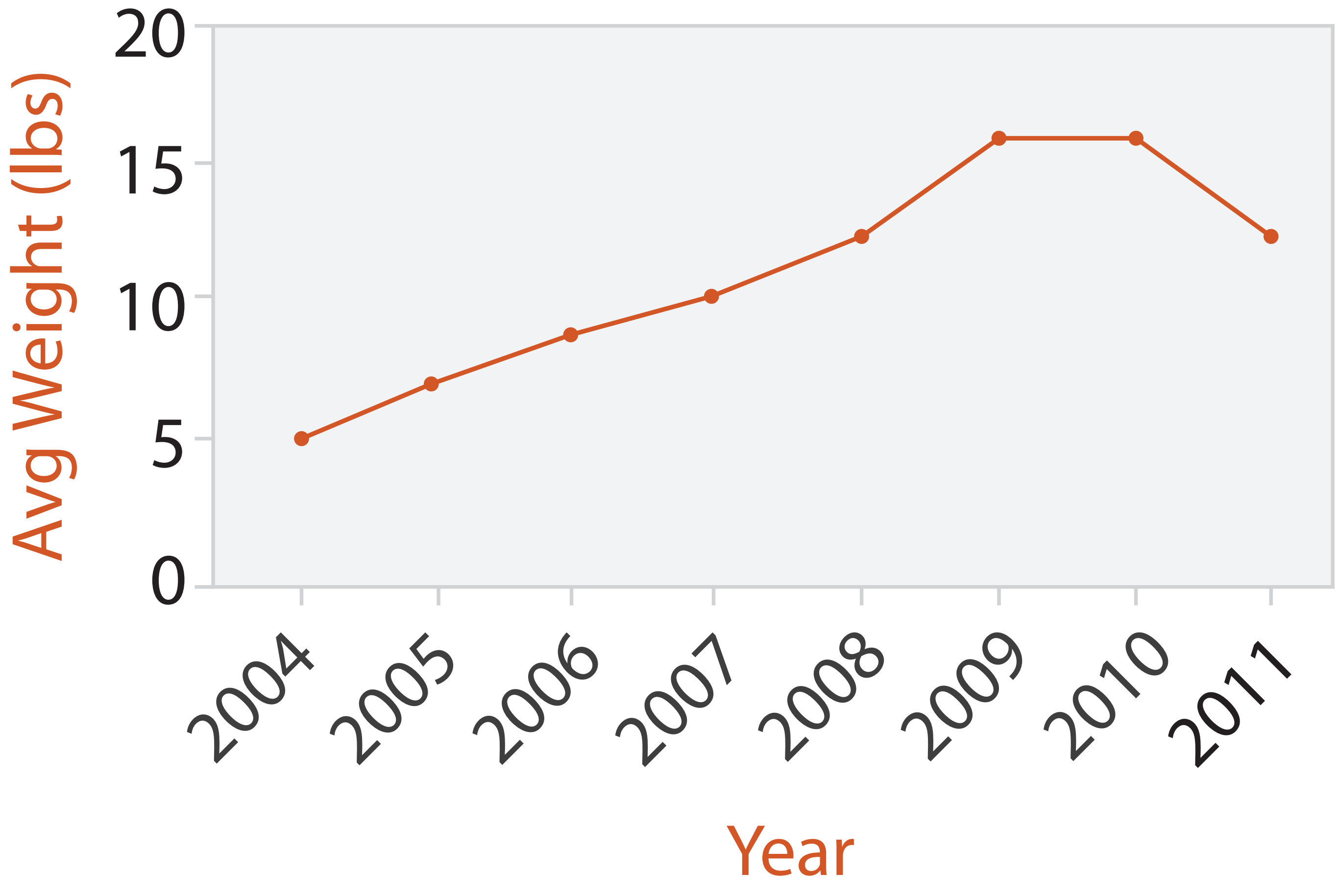
Separated but not Ordered or Aligned



LIMITATION: Hard to make comparisons

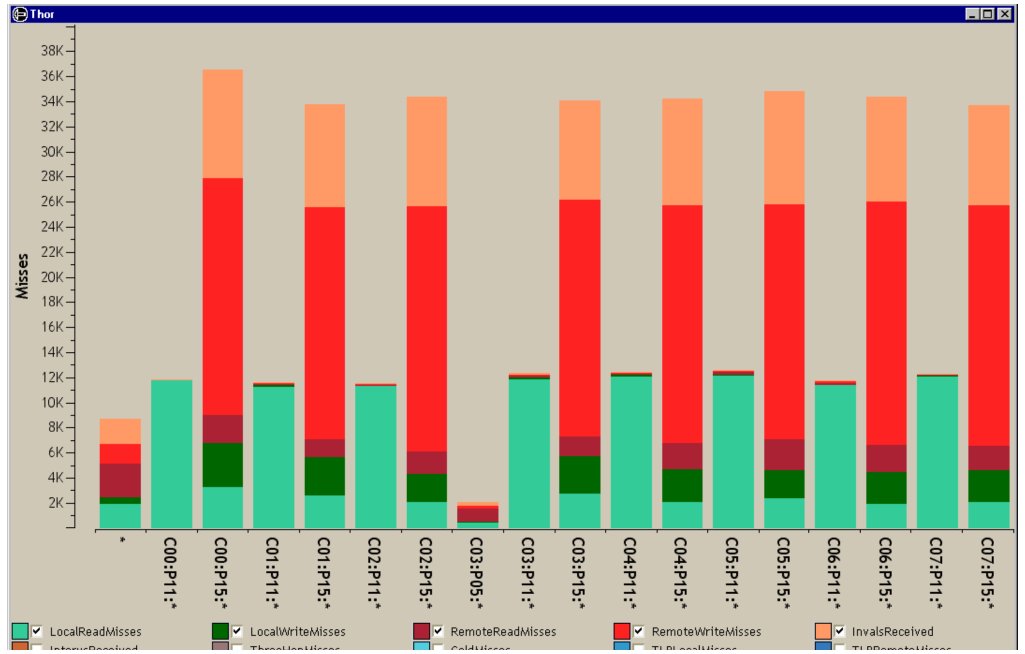
[Slide courtesy of Ben Jones]

Idiom: **line chart / 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

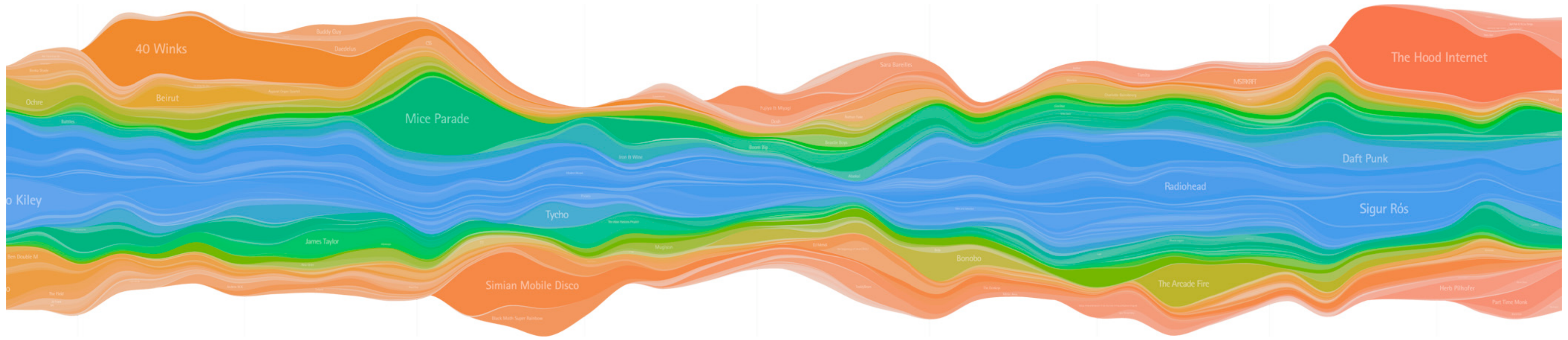
Idiom: **stacked bar chart**



- 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
    - several to one dozen levels for stacked attrib

[Using Visualization to Understand the Behavior of Computer Systems. Bosch. Ph.D. thesis, Stanford Computer Science, 2001.]

Idiom: **streamgraph**

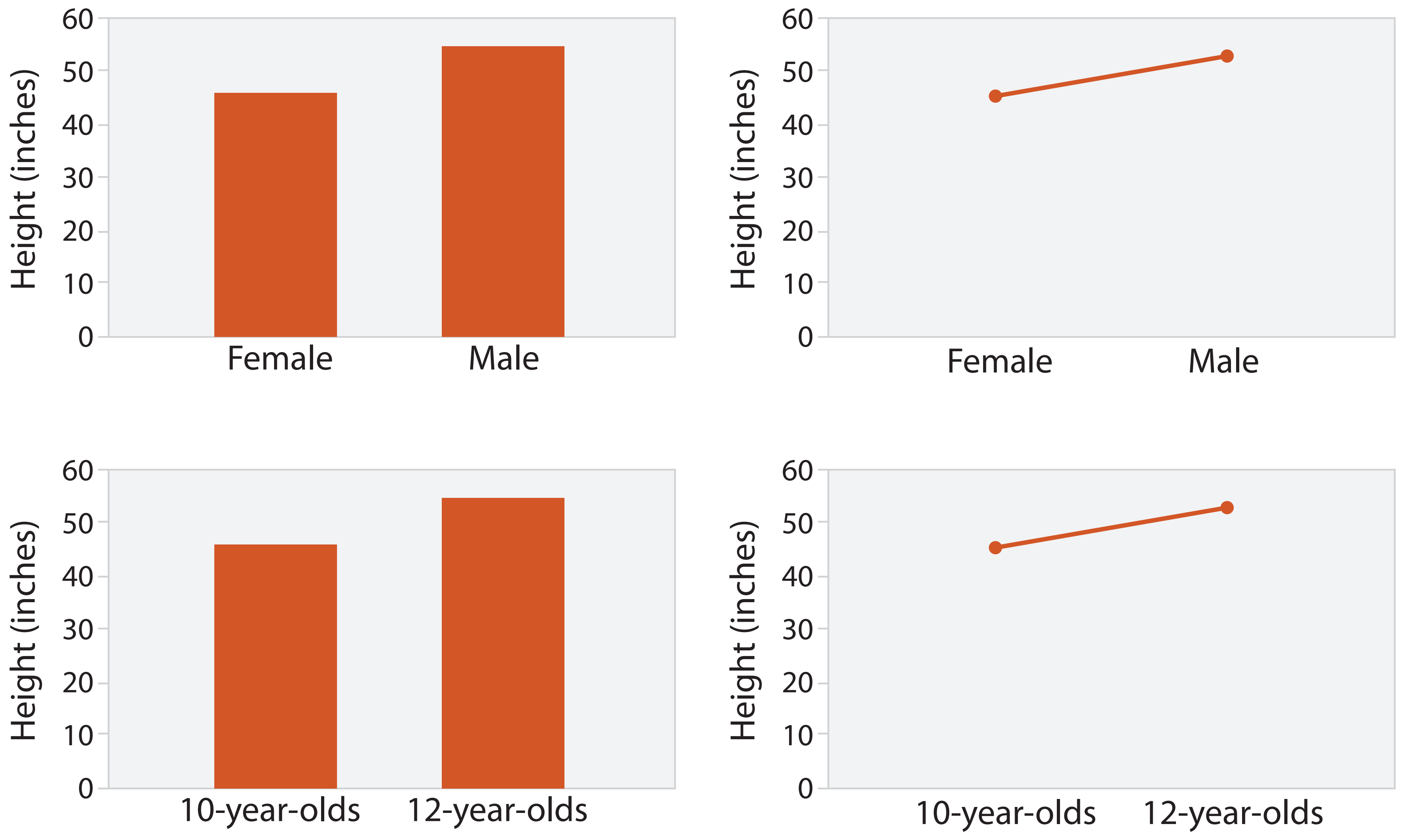


[Stacked Graphs Geometry & Aesthetics. Byron and Wattenberg. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2008) 14(6): 1245–1252, (2008).]

- generalized stacked graph
  - emphasizing horizontal continuity
    - vs vertical items
  - data
    - 1 categ key attrib (artist)
    - 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 artist keys
      - more than stacked bars, since most layers don't extend across whole chart

Choosing bar vs line charts

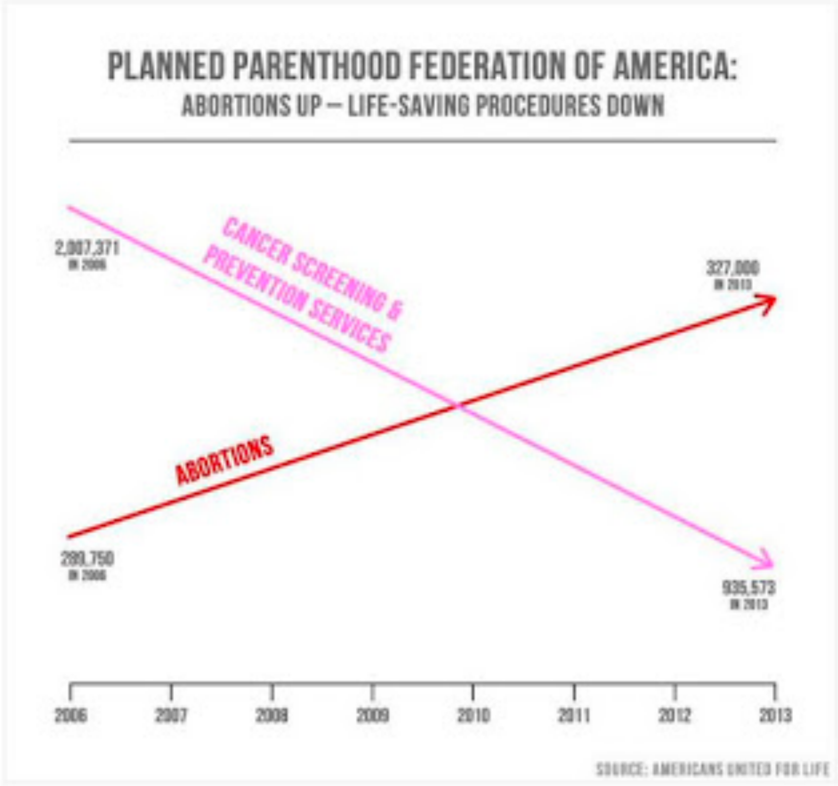
- depends on type of key attrib
  - bar charts if categorical
  - line charts if ordered
- do not use line charts for categorical key attribs
  - violates expressiveness principle
    - implication of trend so strong that it overrides semantics!
      - “The more male a person is, the taller he/she is”

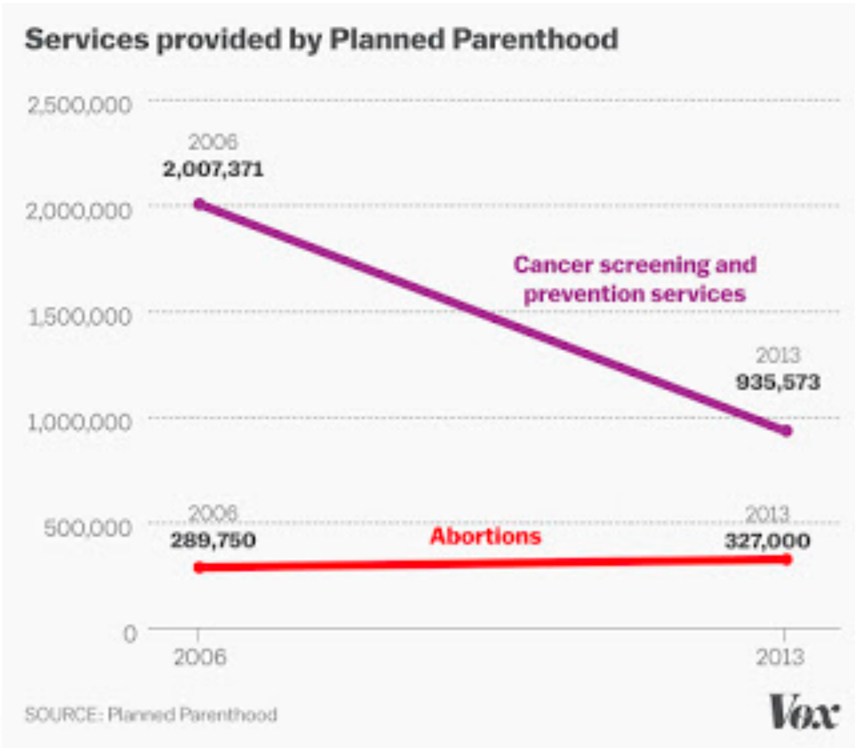


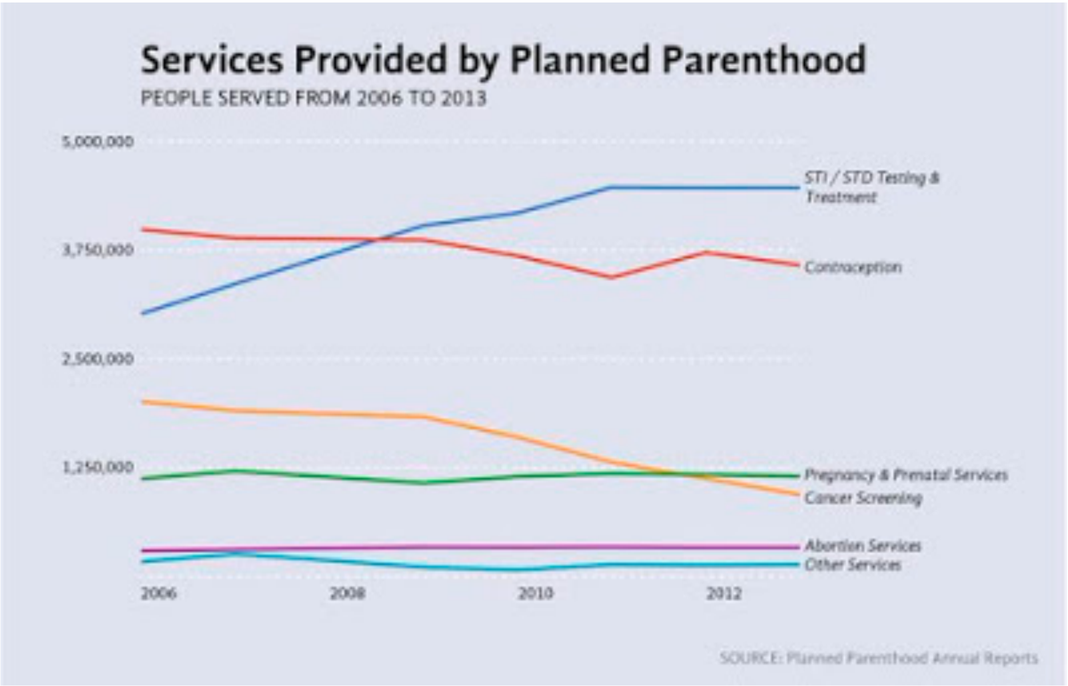
after [Bars and Lines: A Study of Graphic Communication. Zacks and Tversky. Memory and Cognition 27:6 (1999), 1073–1079.]

Chart axes

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



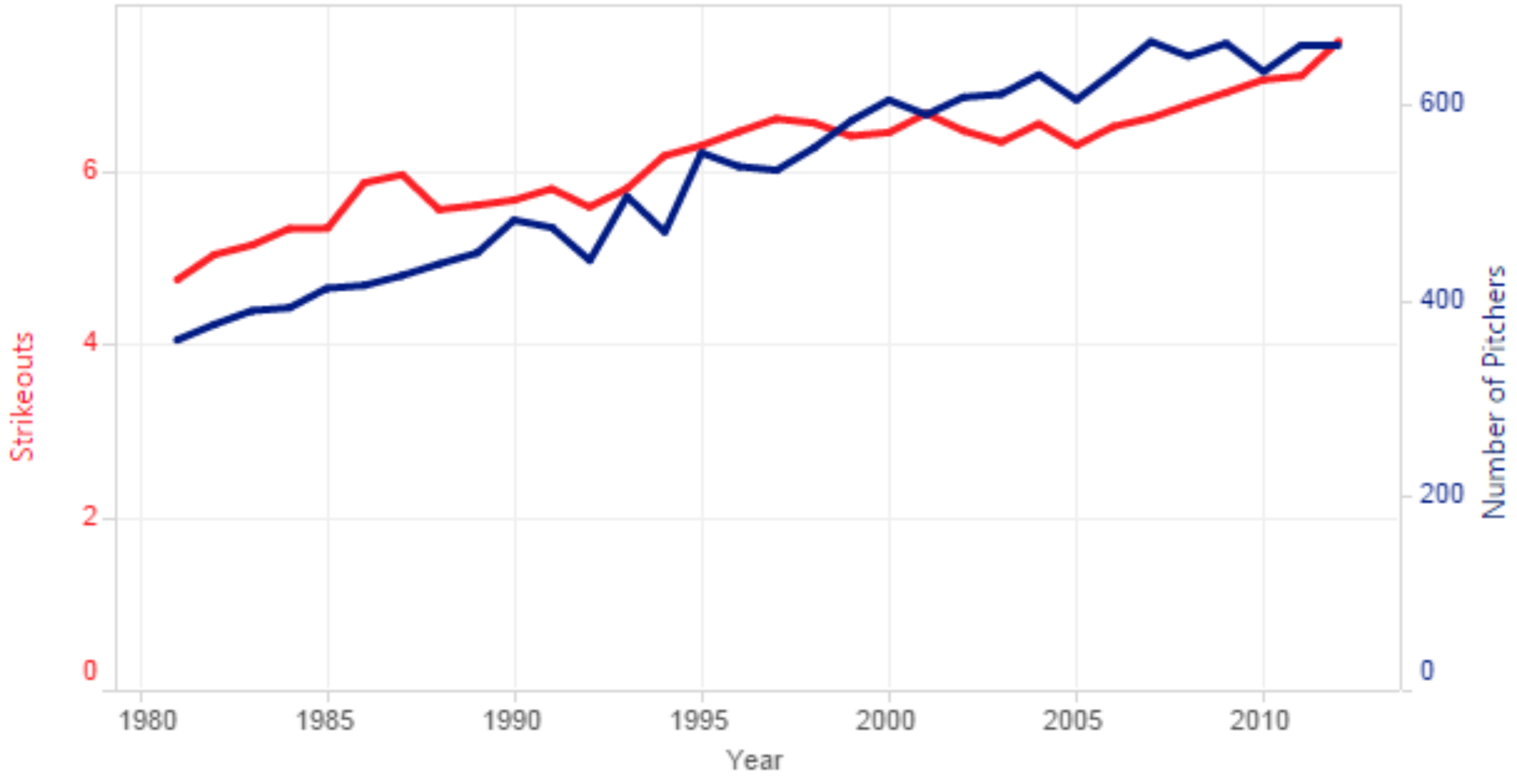




http://www.thefunctionalart.com/2015/10/if-you-see-bullshit-say-bullshit.html

Idiom: **dual-axis line charts**

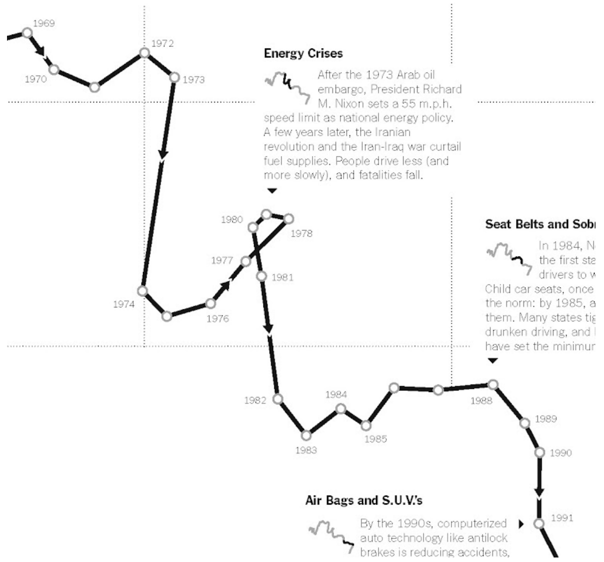
- controversial
  - acceptable if commensurate
  - beware, very easy to mislead!

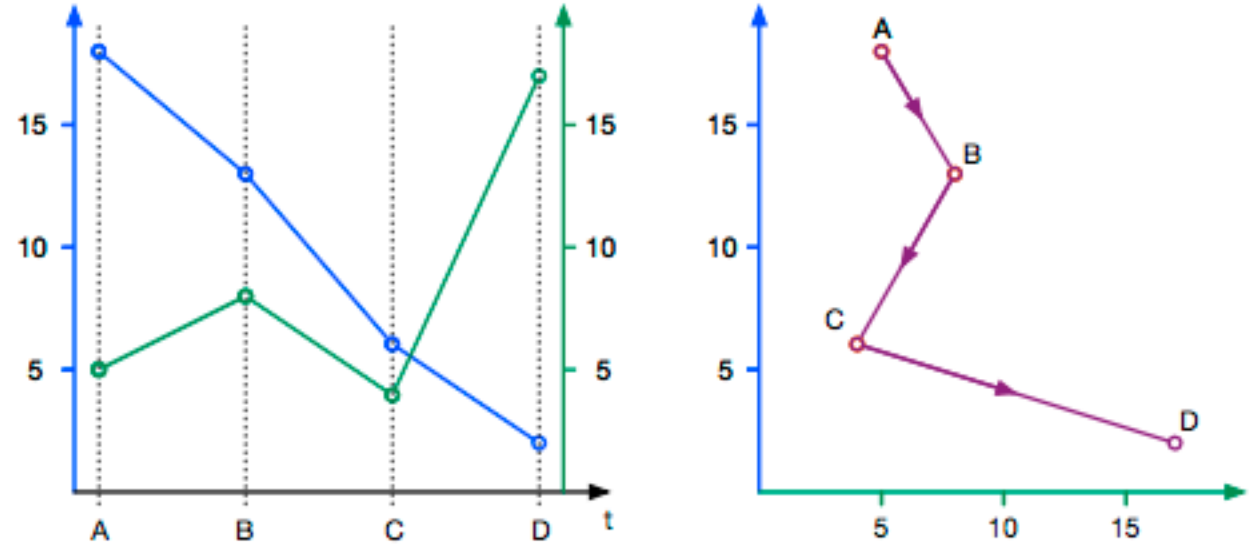


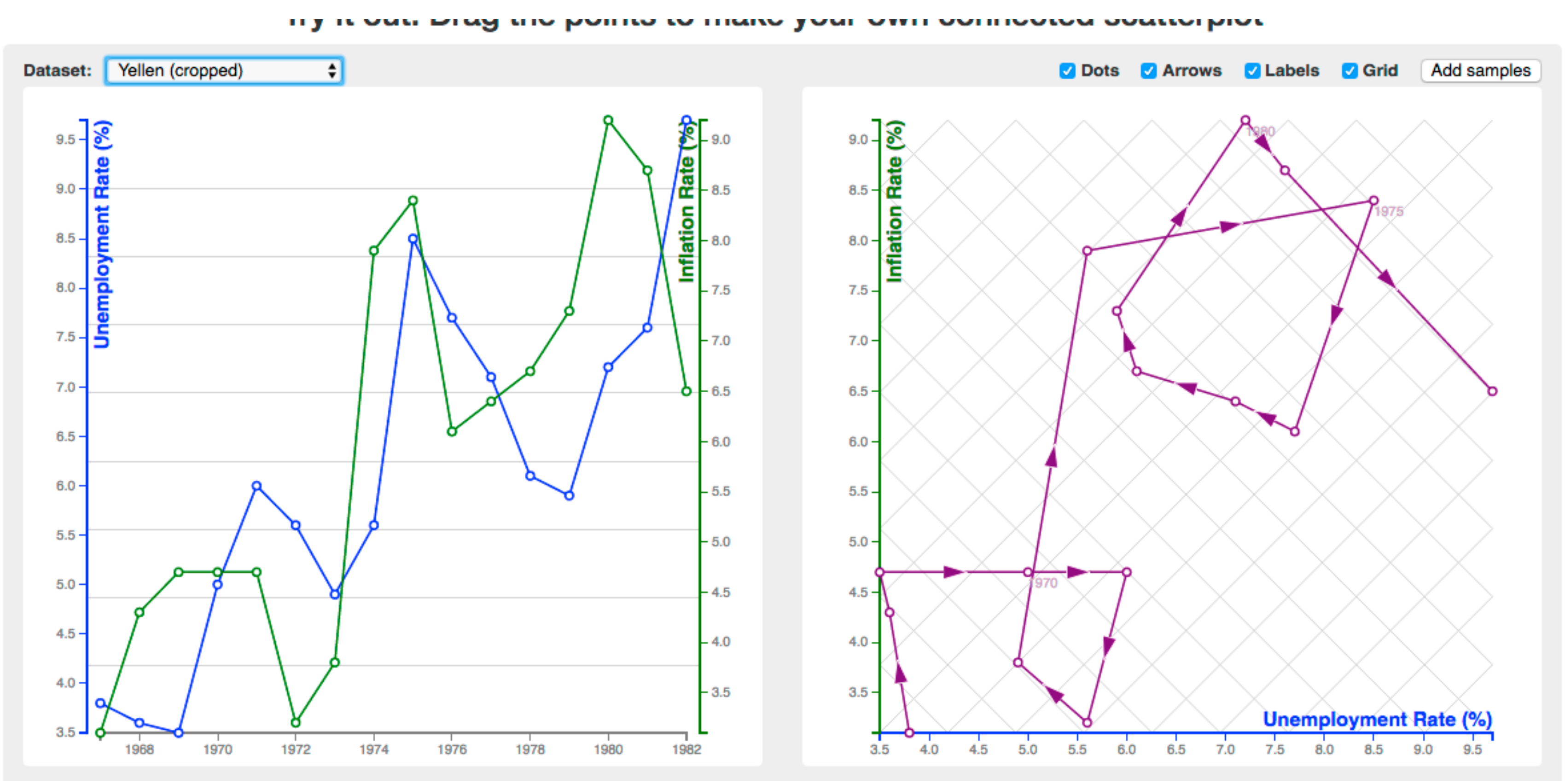
Source | http://www.baseball-reference.com/leagues/MLB/pitch.shtml Ben Jones (@DataRemixed) | 5/4/2013

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





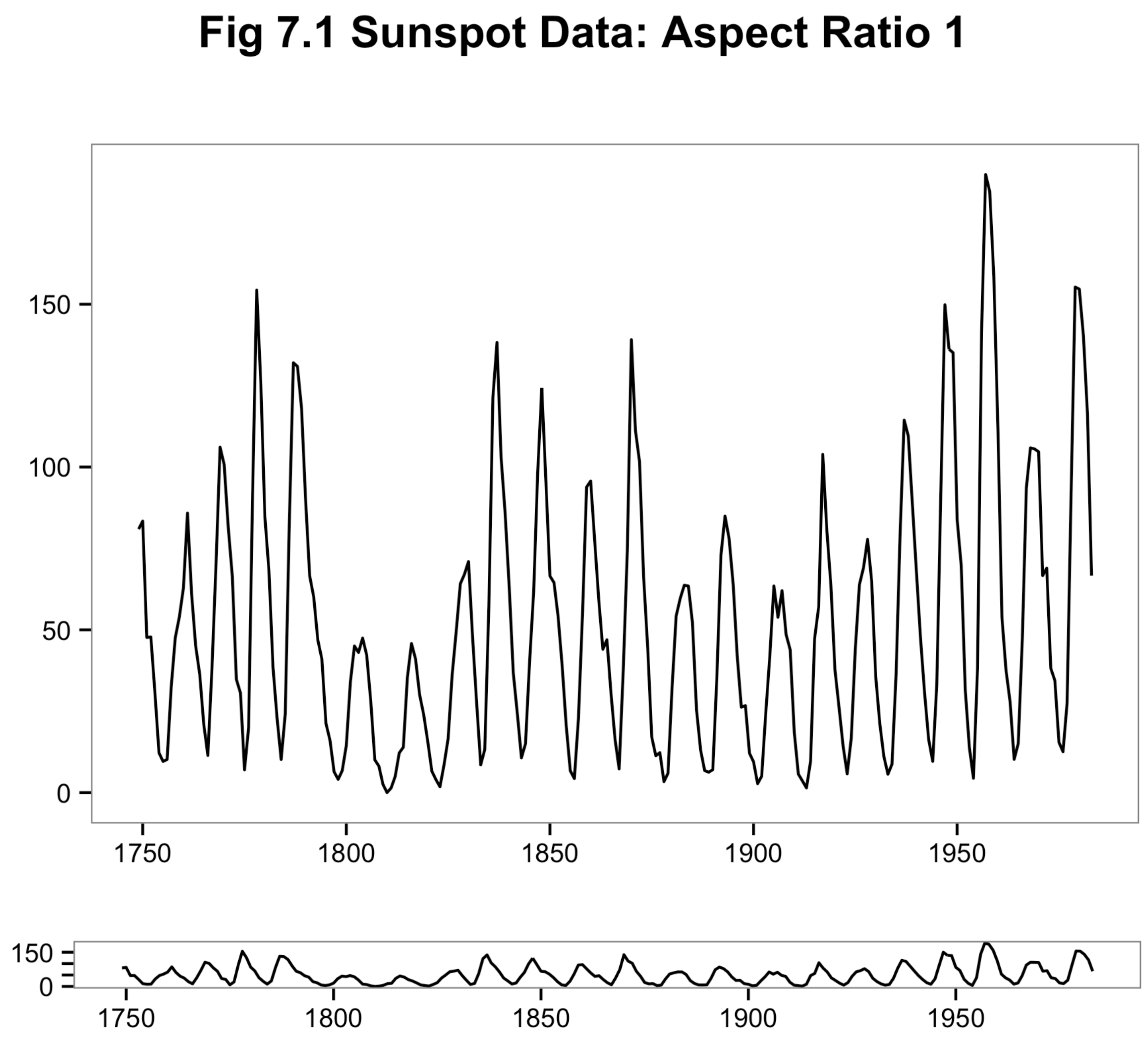


[The Connected Scatterplot for Presenting Paired Time Series. Haroz, Kosara and Franconeri. IEEE TVCG 22(9):2174-86, 2016.]

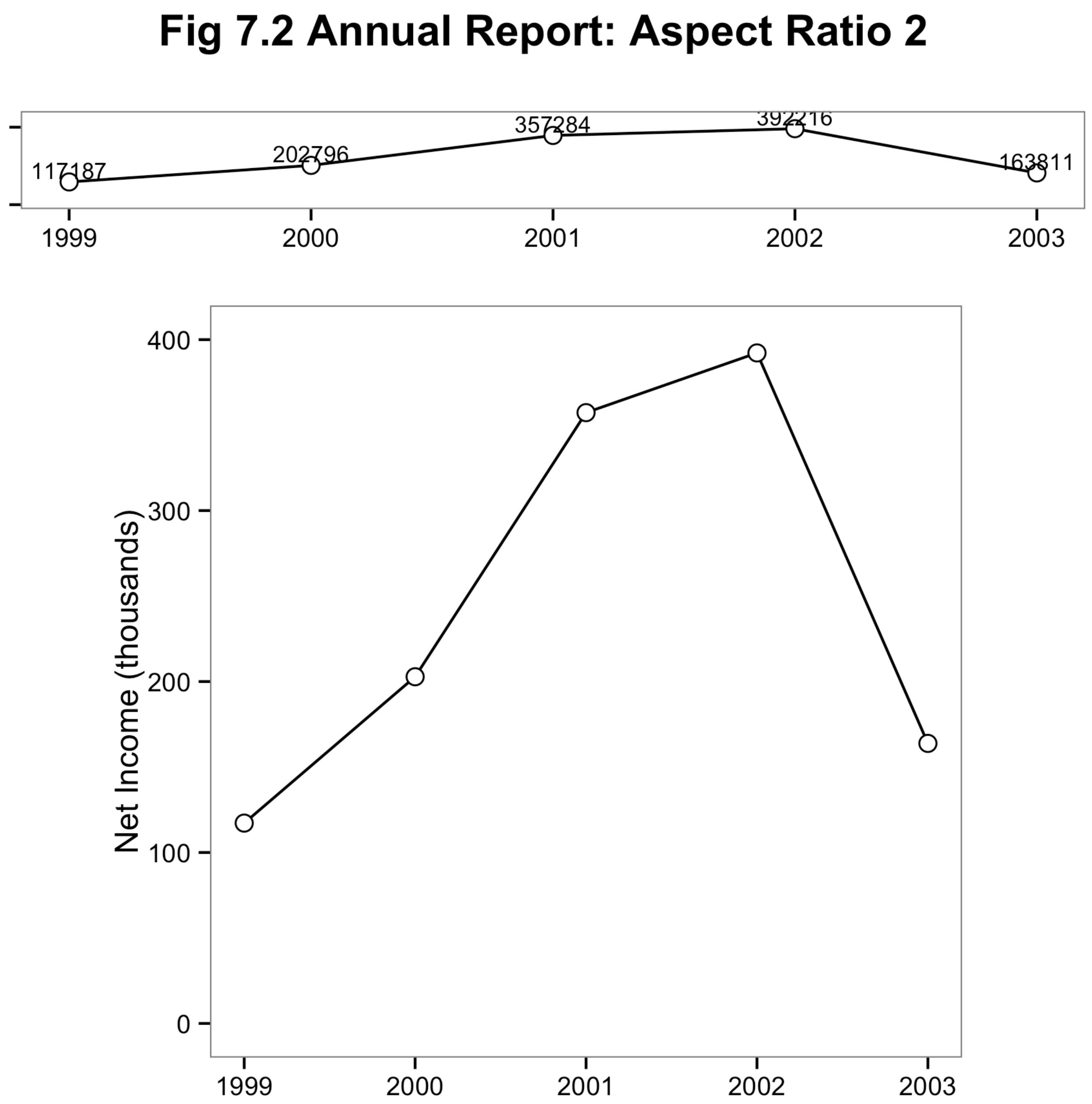
http://steveharoz.com/research/connected_scatterplot/

Choosing line chart aspect ratios

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



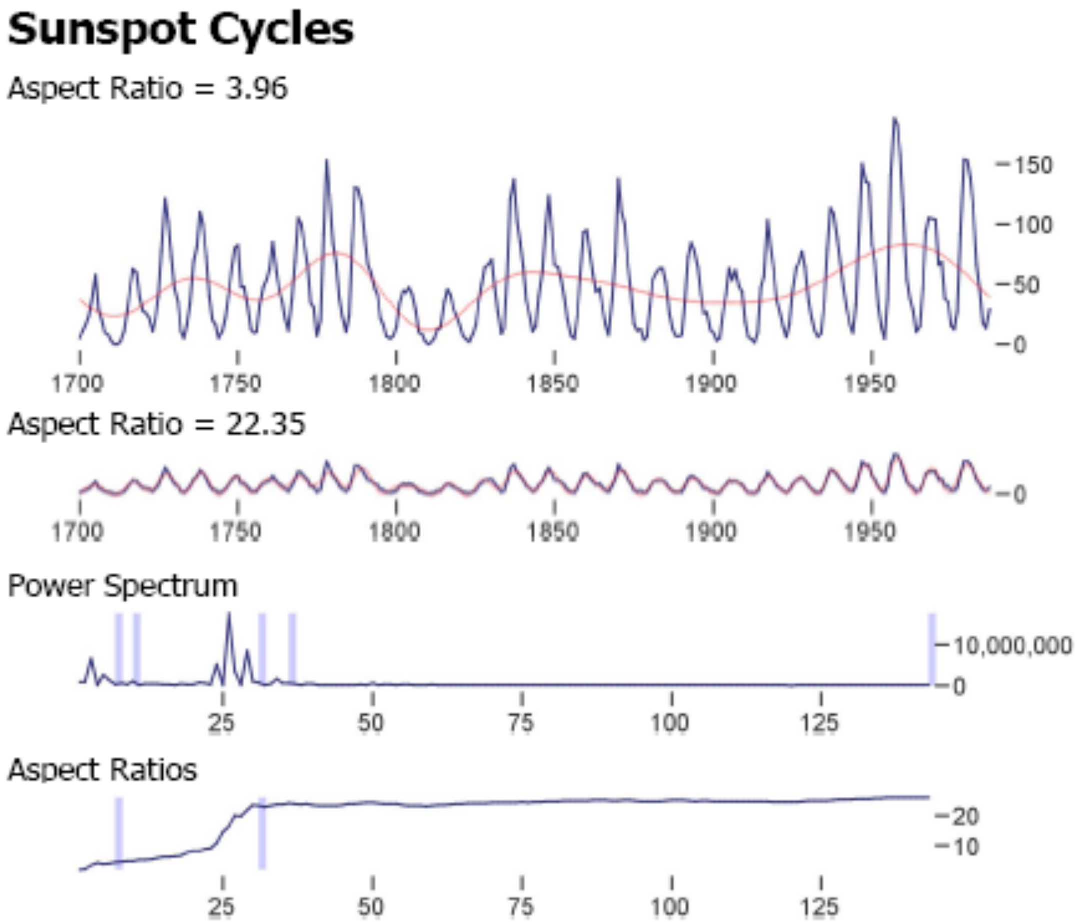
https://github.com/jennybc/r-graph-catalog/tree/master/figures/fig07-01_sunspot-data-aspect-ratio-1



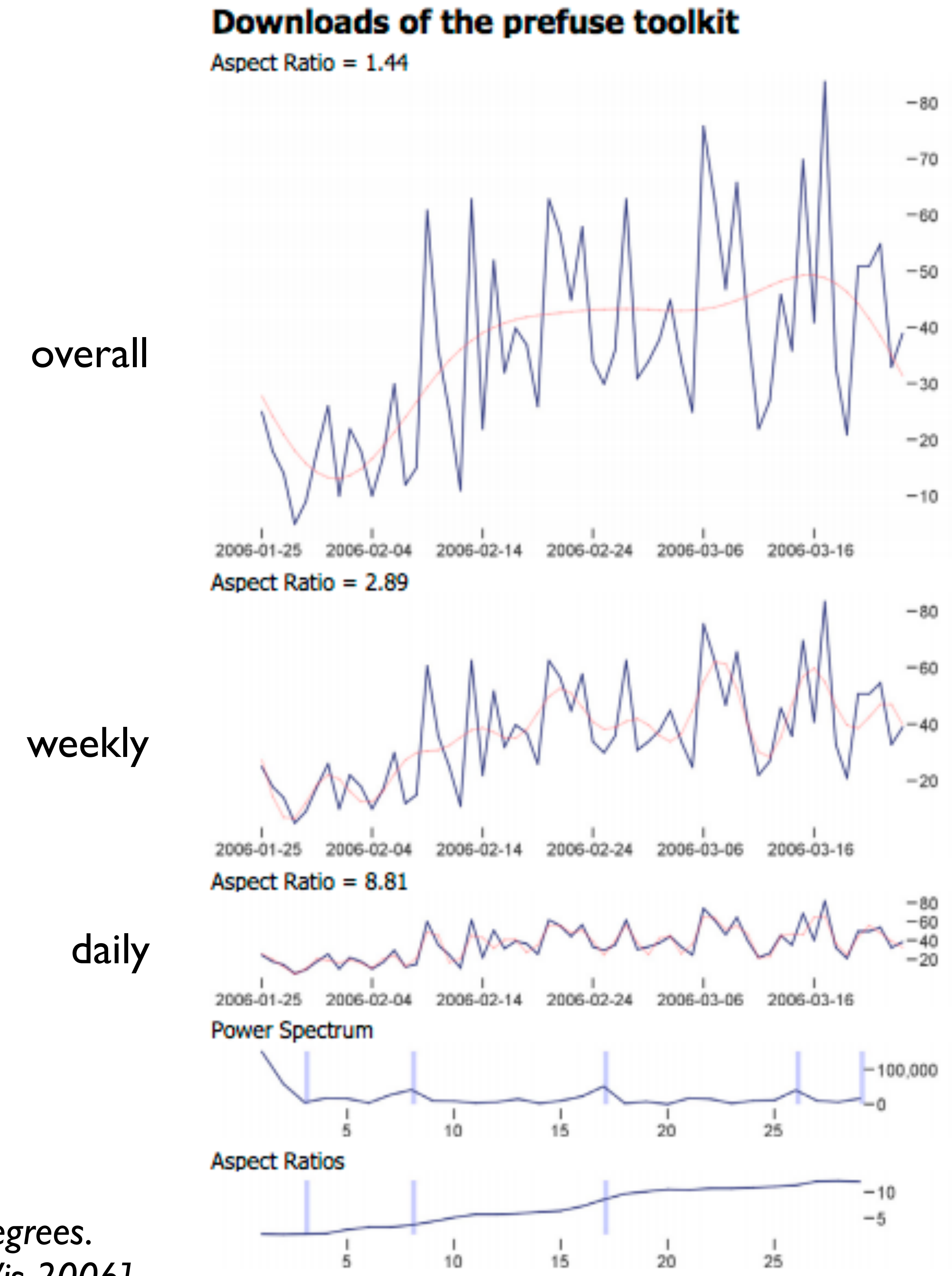
https://github.com/jennybc/r-graph-catalog/tree/master/figures/fig07-02_annual-report-aspect-ratio-2

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



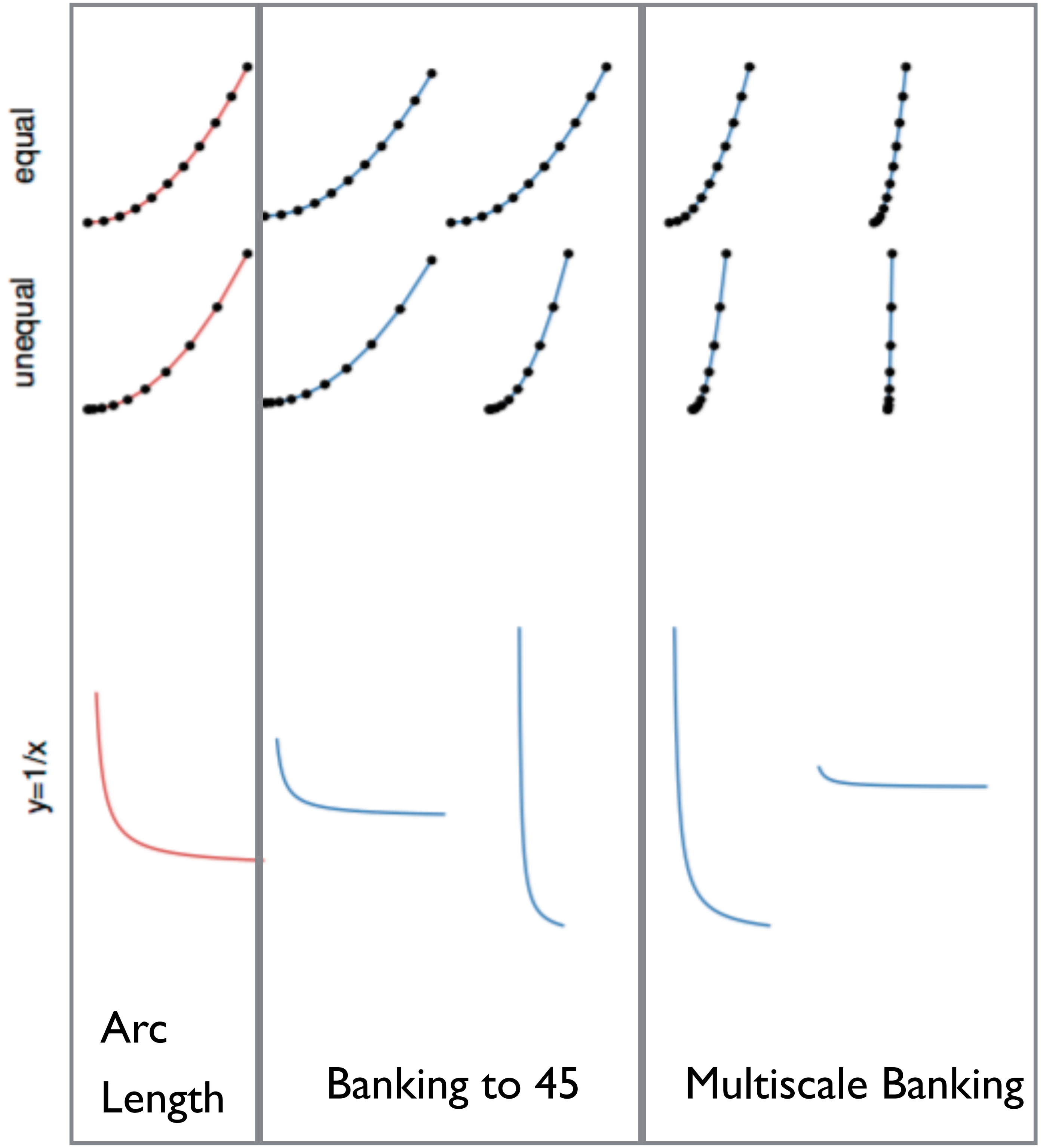
[Multi-Scale Banking to 45 Degrees. Heer and Agrawala, Proc InfoVis 2006]



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

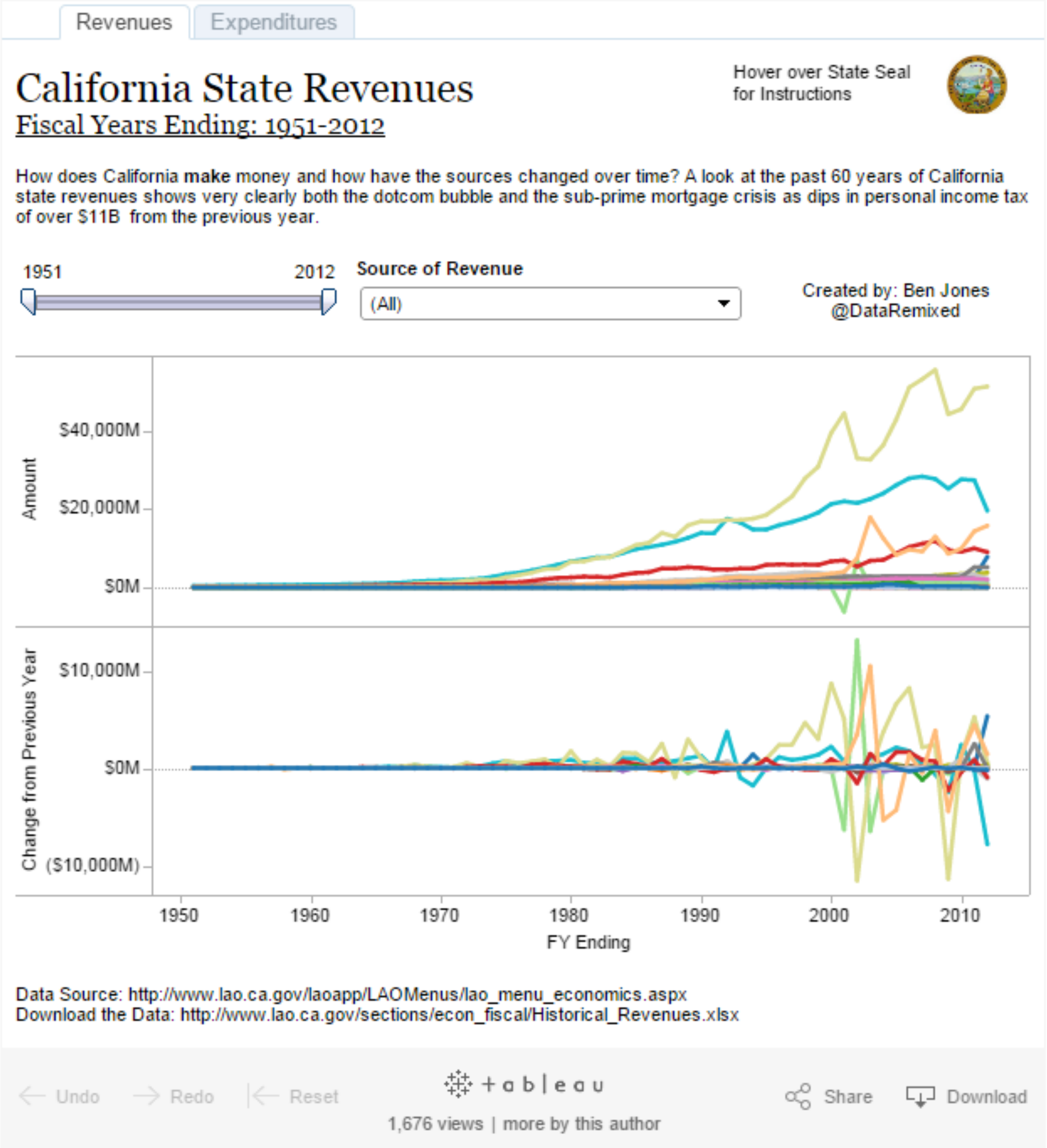
- meta-points from this progression
  - young field; prescriptive advice changes rapidly
  - reasonable defaults required deep dive into perception meets math



[Arc Length-Based Aspect Ratio Selection. Talbot, Gerth, and Hanrahan. Proc InfoVis 2011]

Idiom: **Indexed line charts**

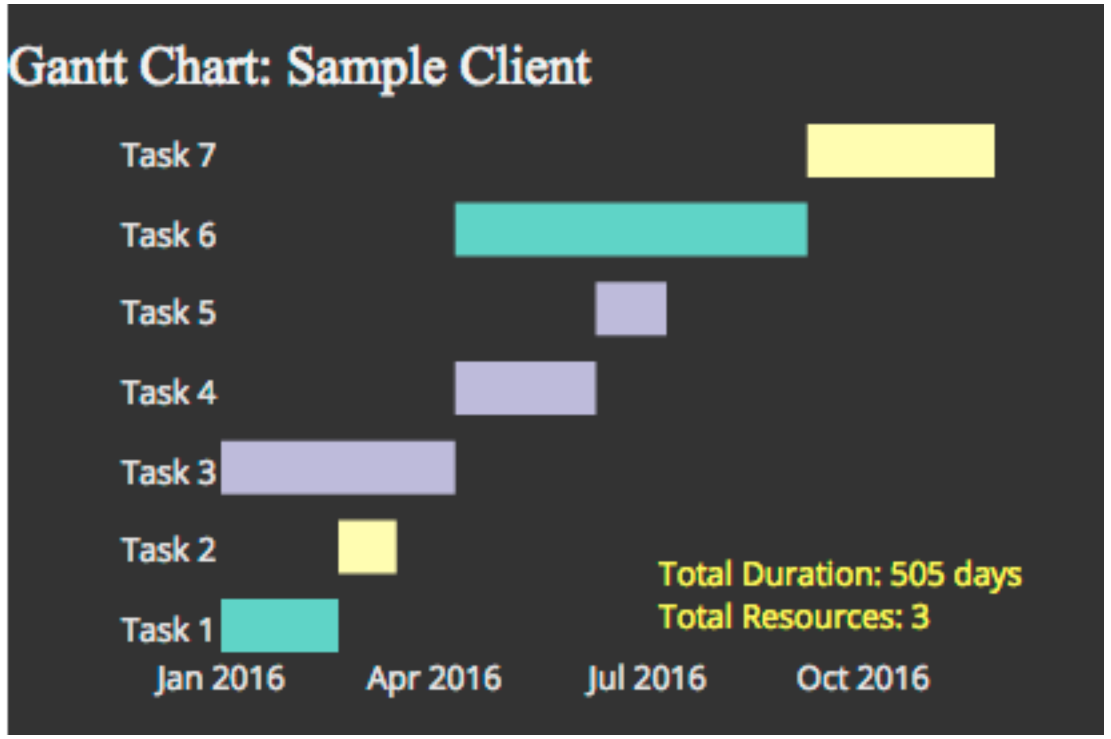
- data: 2 quant attires
  - 1 key + 1 value
- derived data: new quant value attrib
  - index
  - plot instead of original value
- task: show change over time
  - principle: normalized, not absolute
- scalability
  - same as standard line chart



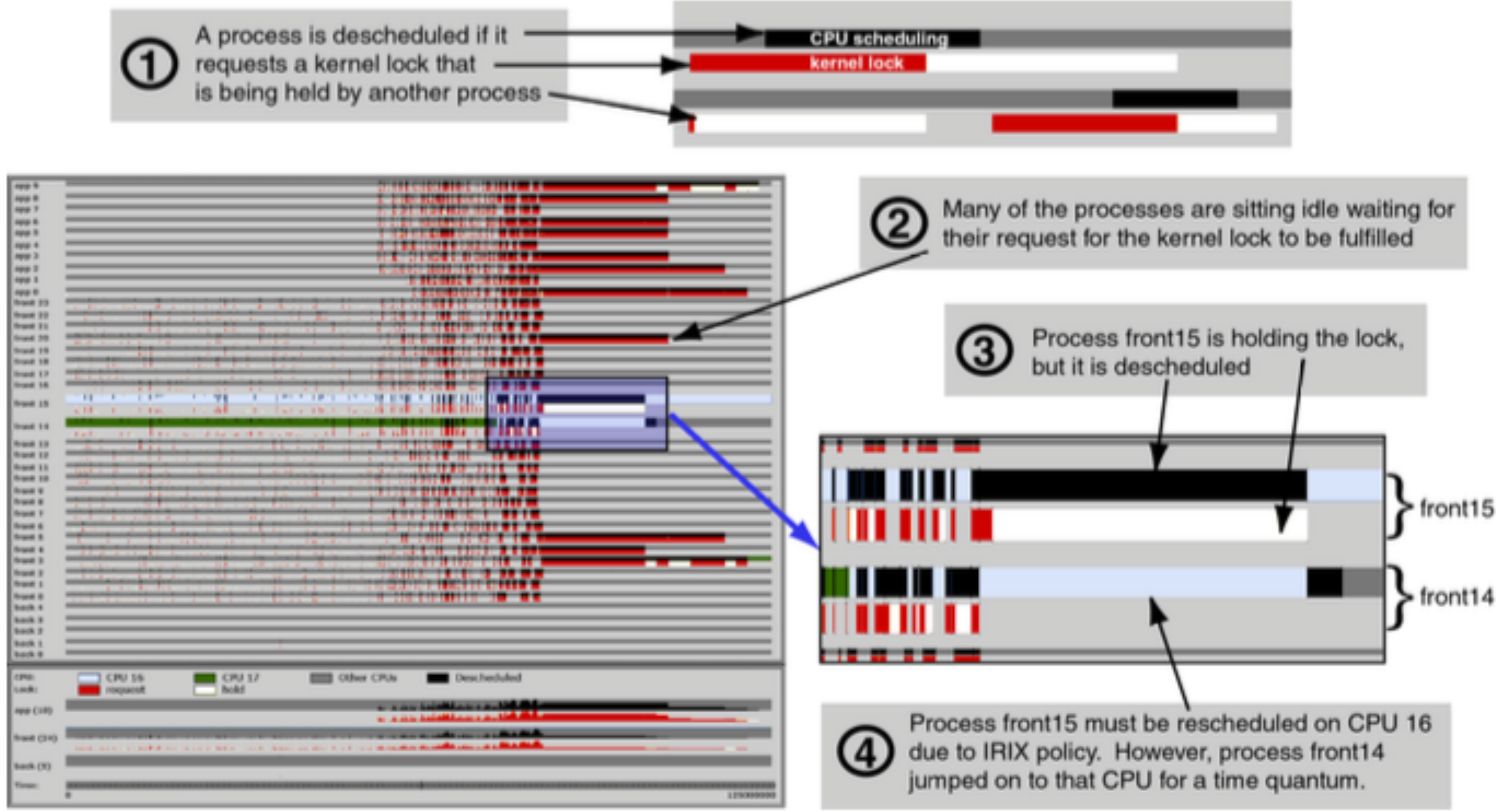
https://public.tableau.com/profile/ben.jones#!/vizhome/CAStateRevenues/Revenues

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



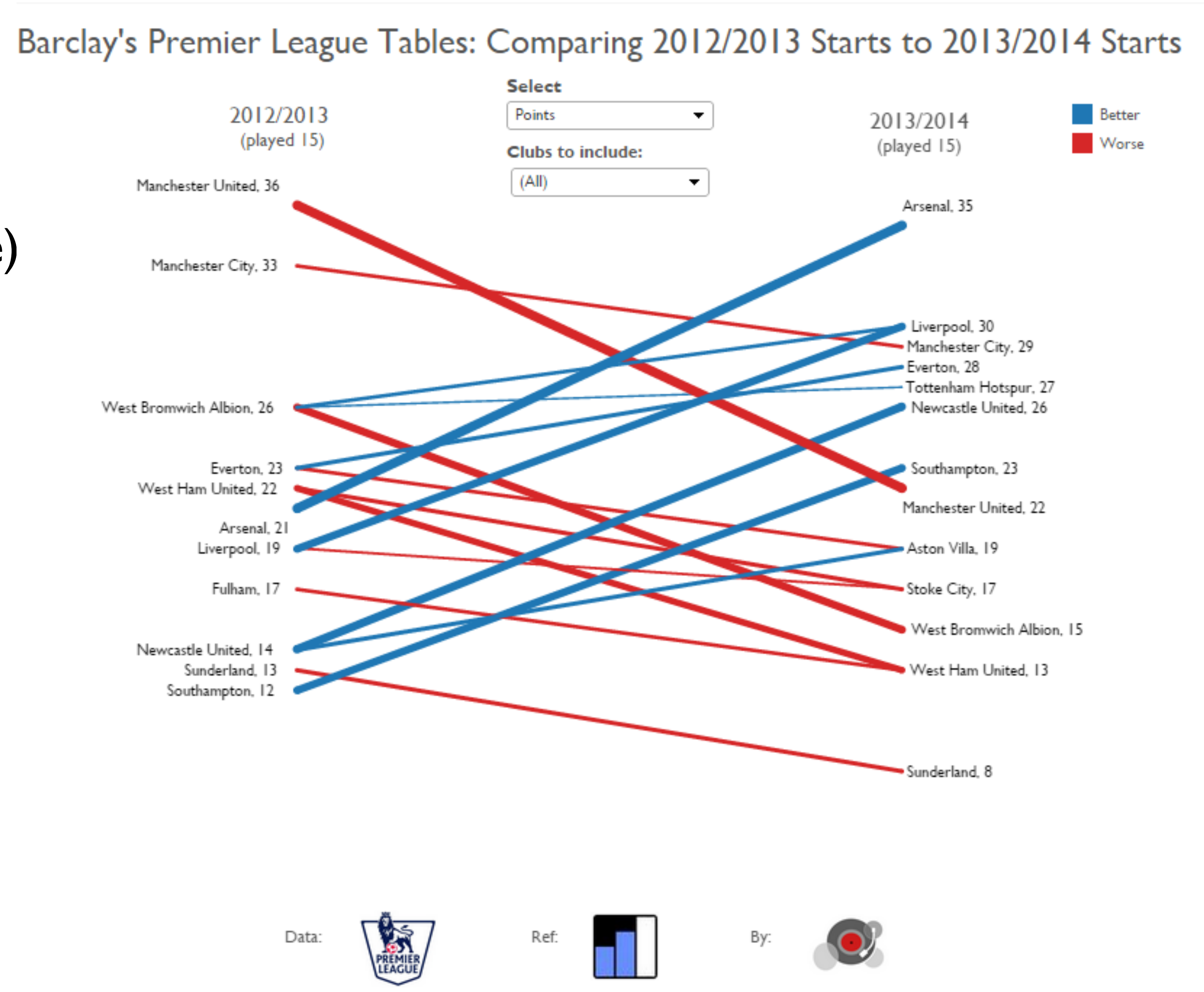
https://www.r-bloggers.com/gantt-charts-in-r-using-plotly/



[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**

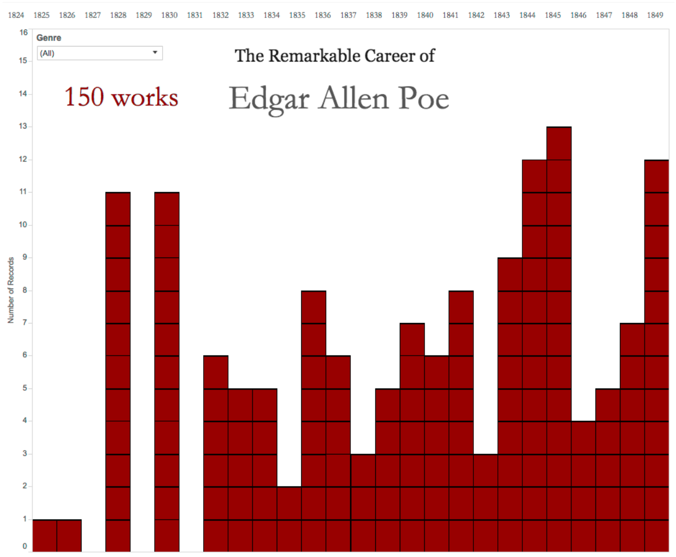
- two values
  - data
    - 2 quant value attribs
    - (1 derived attrib: change magnitude)
  - mark: point + line
    - line connecting mark between pts
  - channels
    - 2 vertical pos: express attrib value
    - (linewidth/size, color)
  - task
    - emphasize changes in rank/value
  - scalability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



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

Breaking conventions

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

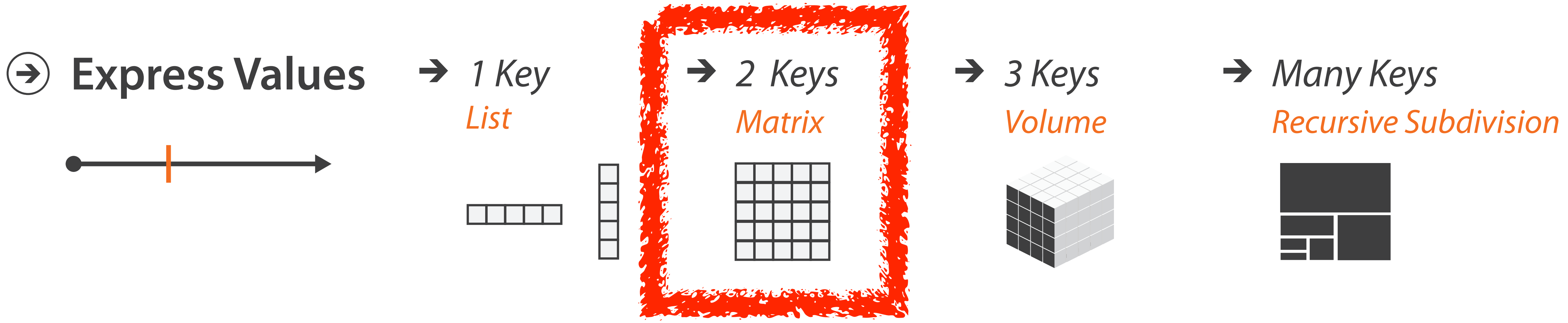


https://public.tableau.com/profile/ben.jones#!/vizhome/EdgarAllanPoeBoring/EdgarAllenPoeBoring

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

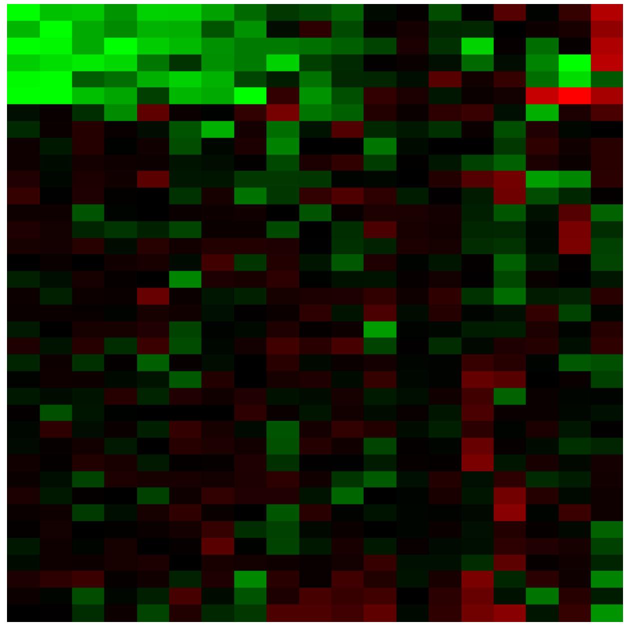
[Slide inspired by Ben Jones]

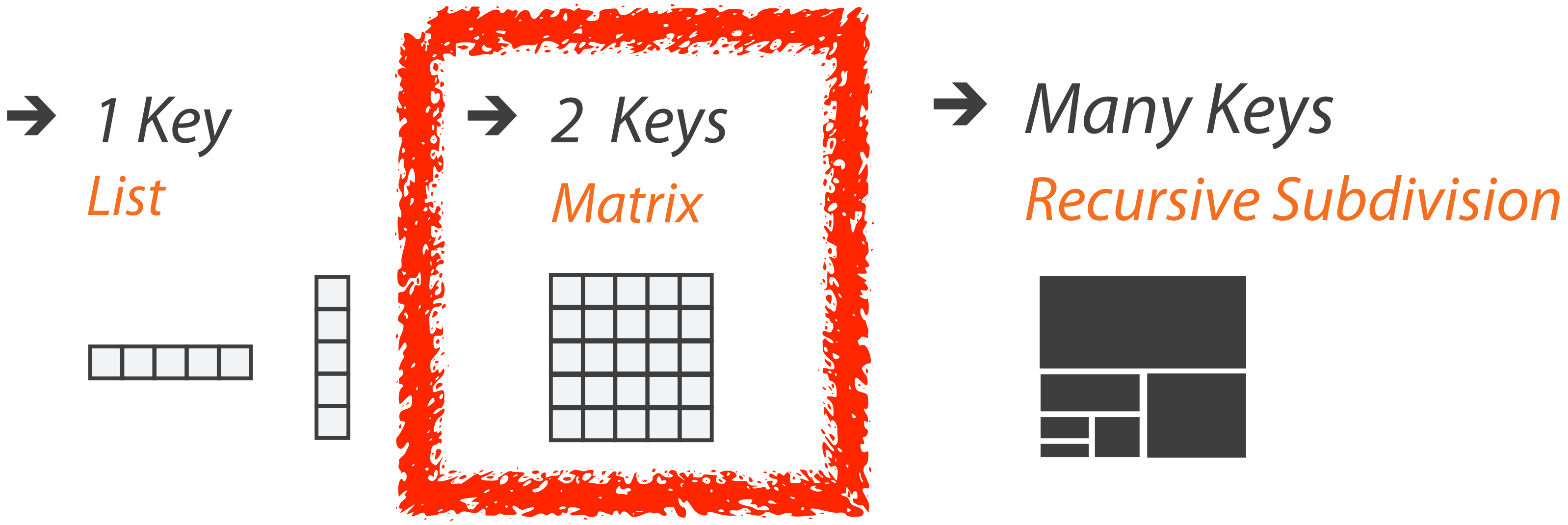
2 Keys


Express Values
1 Key
List
2 Keys
Matrix
3 Keys
Volume
Many Keys
Recursive Subdivision

Idiom: **heatmap**

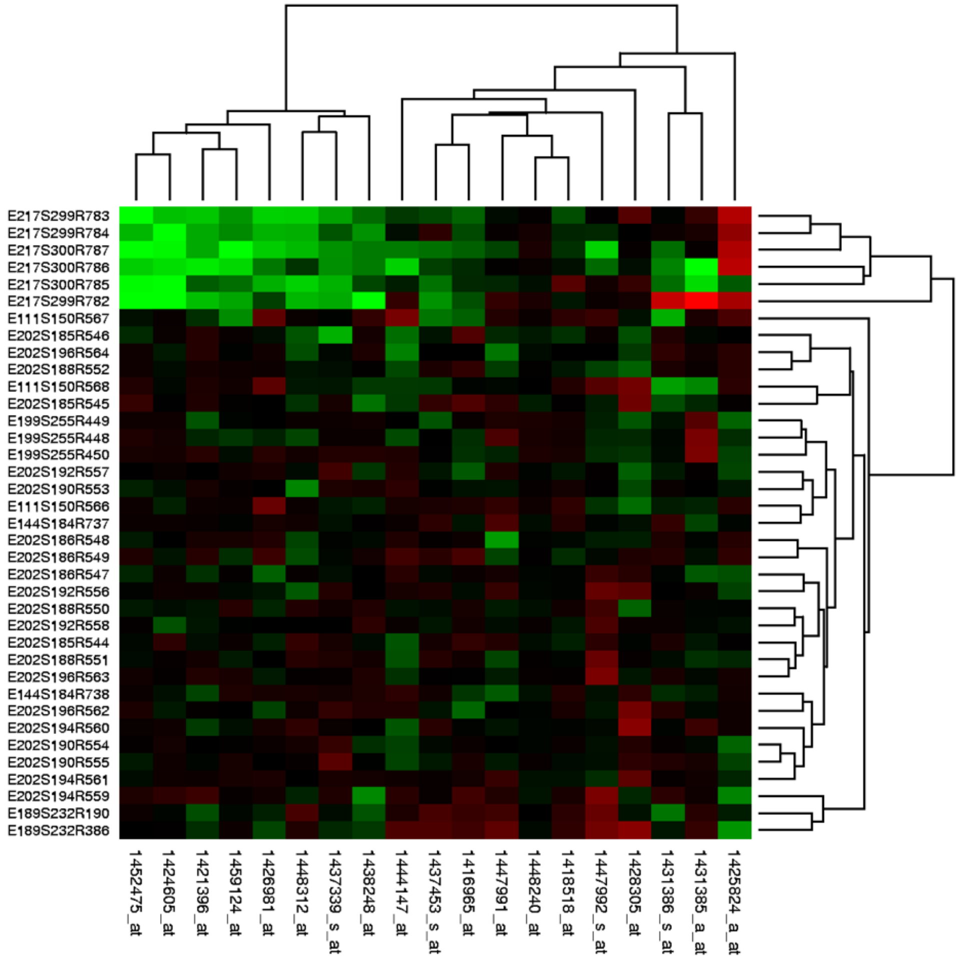
- two keys, one value
  - data
    - 2 categ attribs (gene, experimental condition)
    - 1 quant attrib (expression levels)
  - marks: area
    - separate and align in 2D matrix
      - indexed by 2 categorical attributes
  - channels
    - color by quant attrib
      - (ordered diverging colormap)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





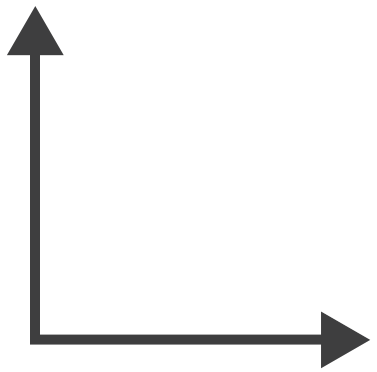
Idiom: **cluster heatmap**

- 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

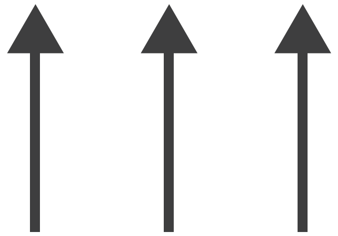


Axis Orientation

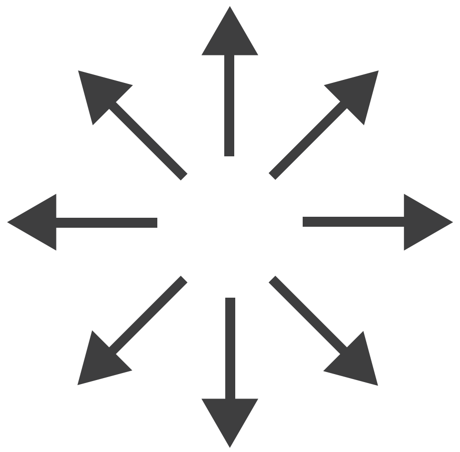
- Rectilinear



- Parallel

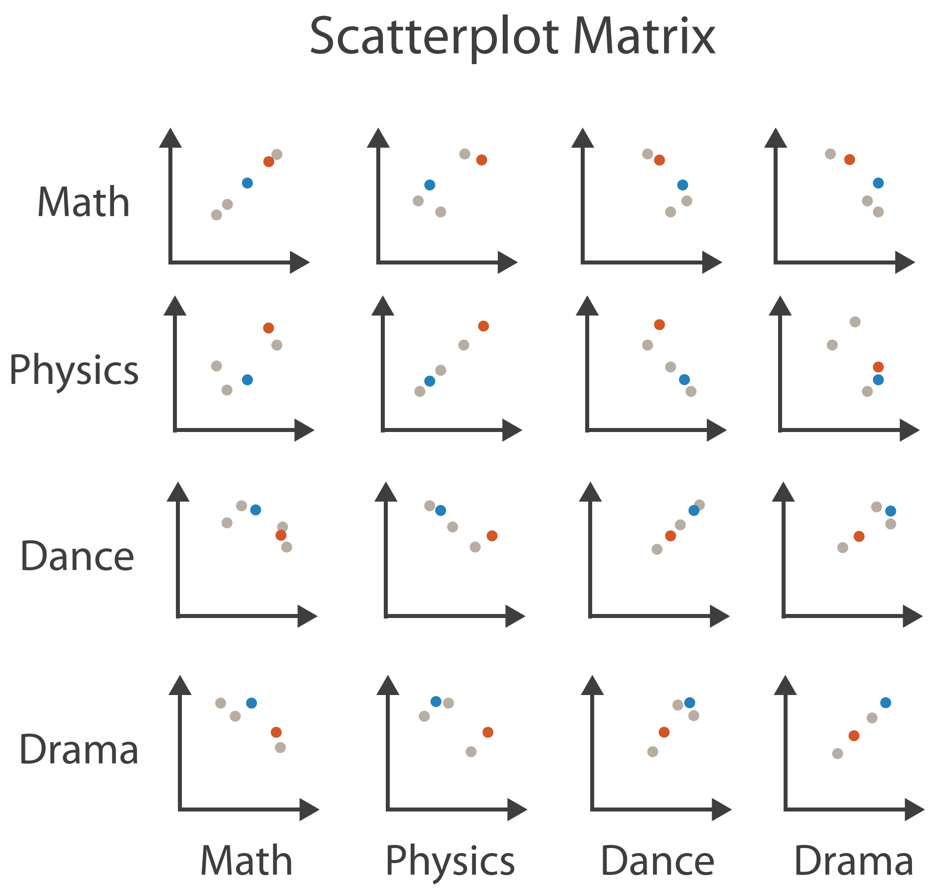


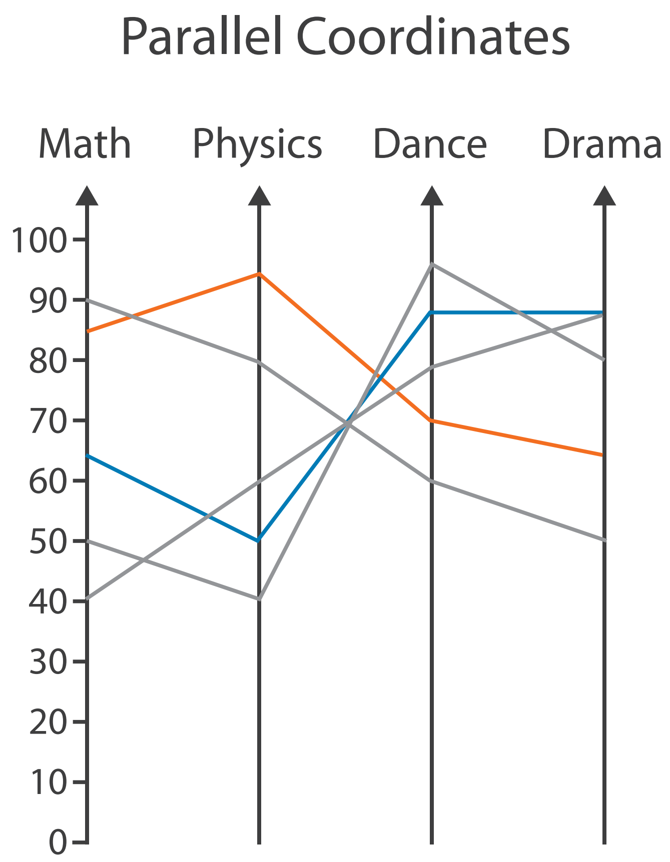
- Radial



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





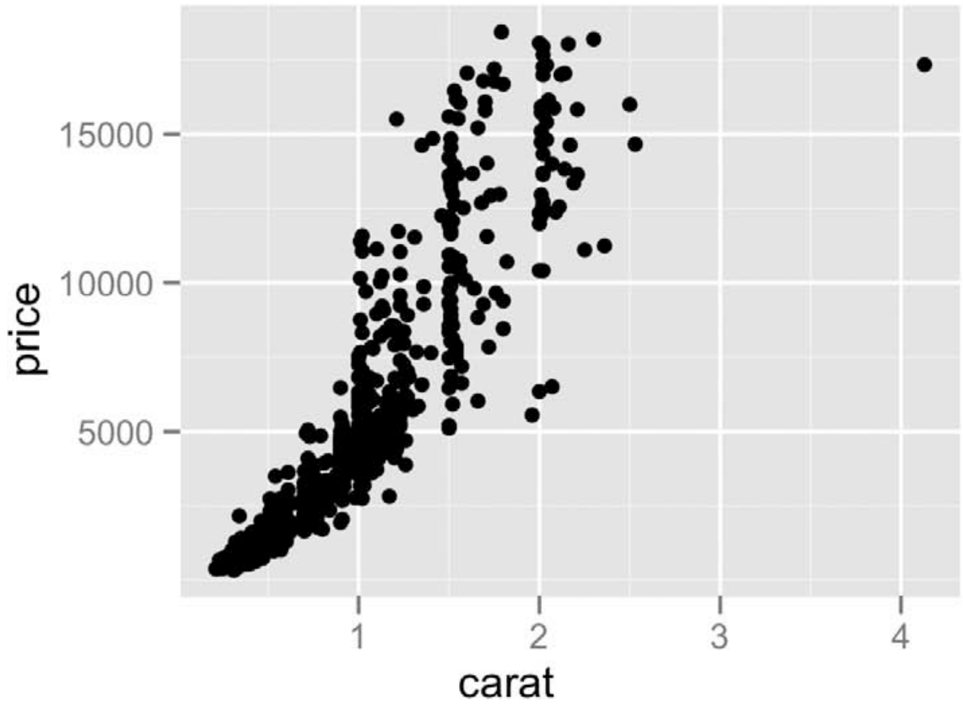
Table

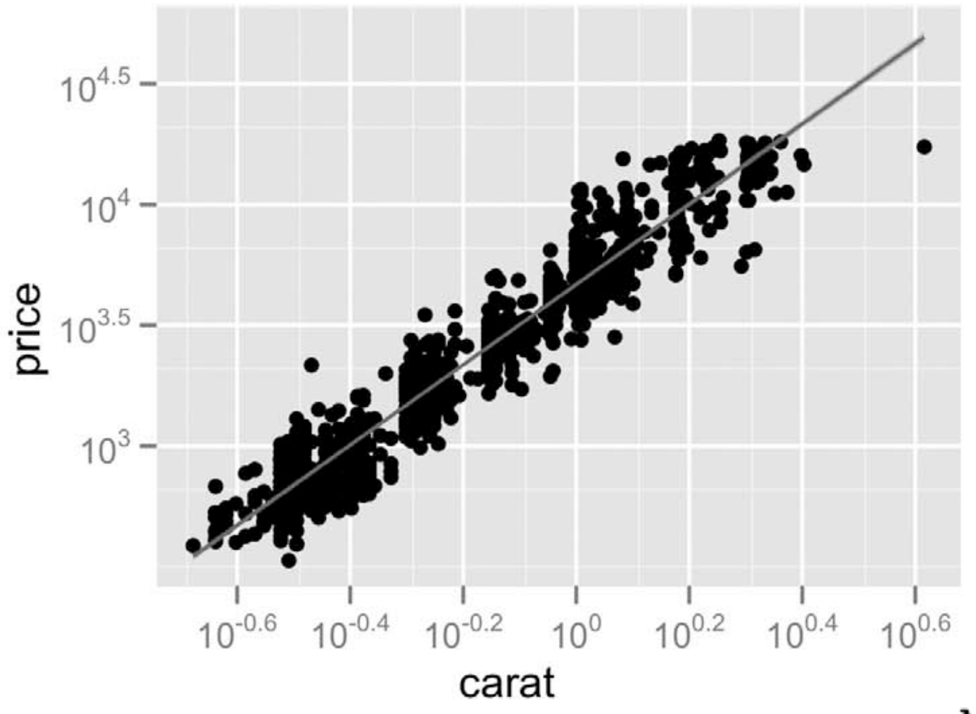
Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

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





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

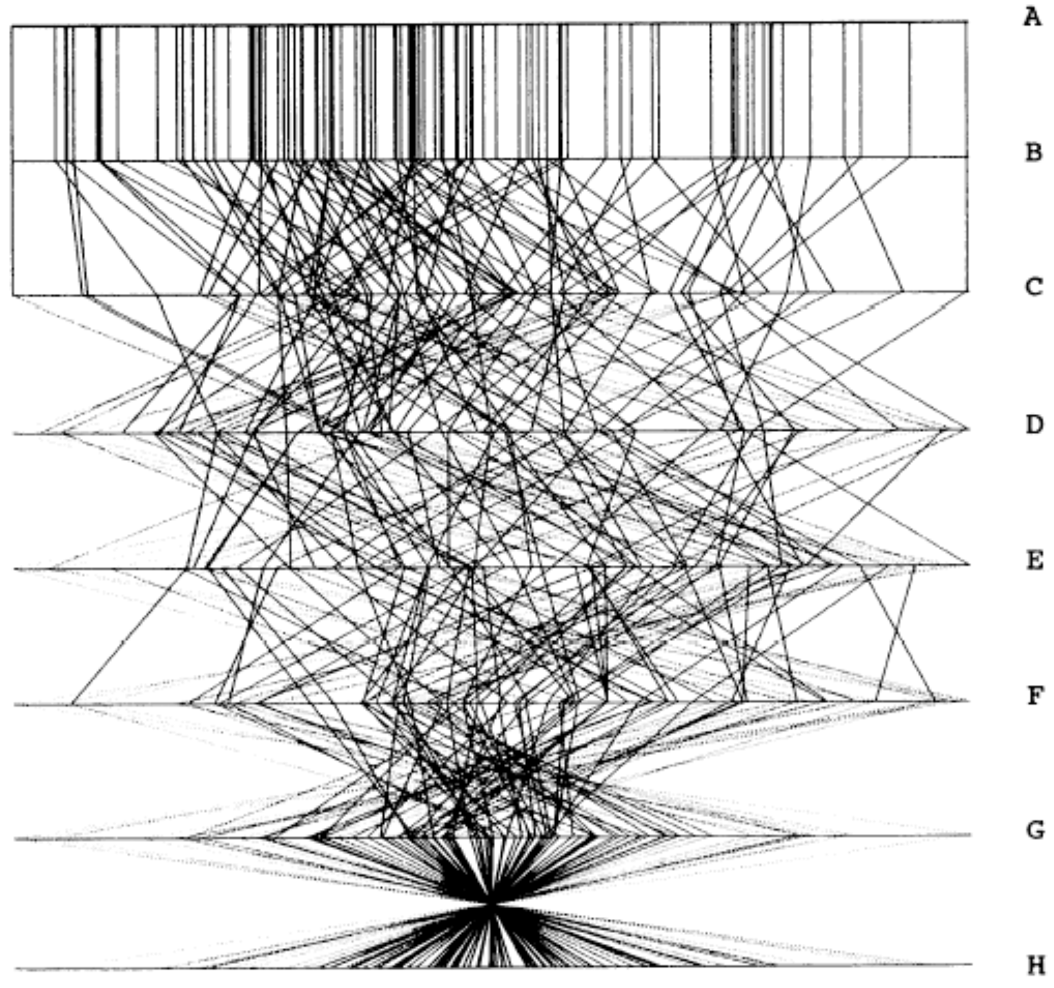
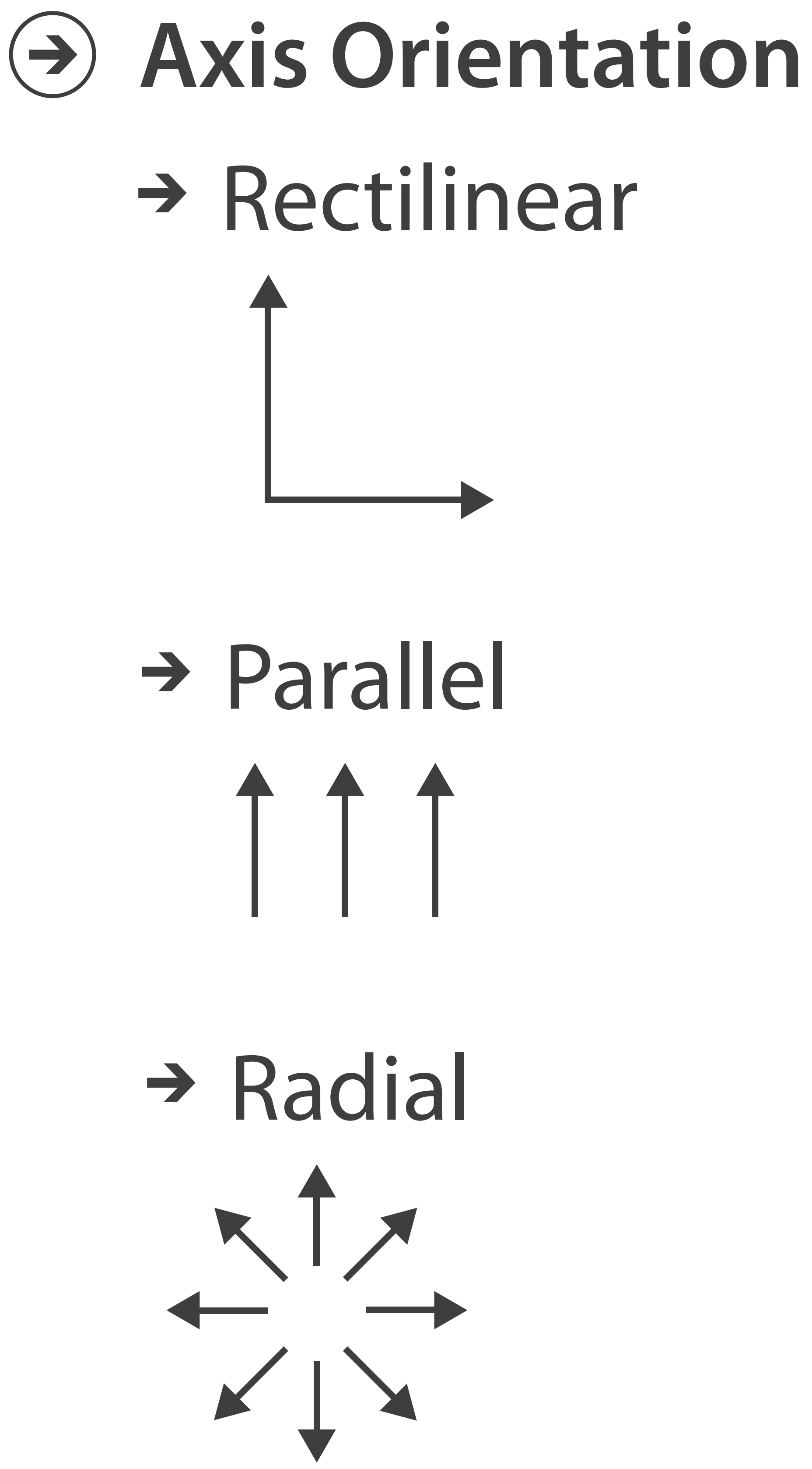


Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of ρ = 1, .8, .2, 0, −.2, −.8, and −1.

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

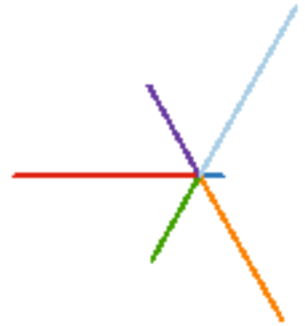
Orientation limitations

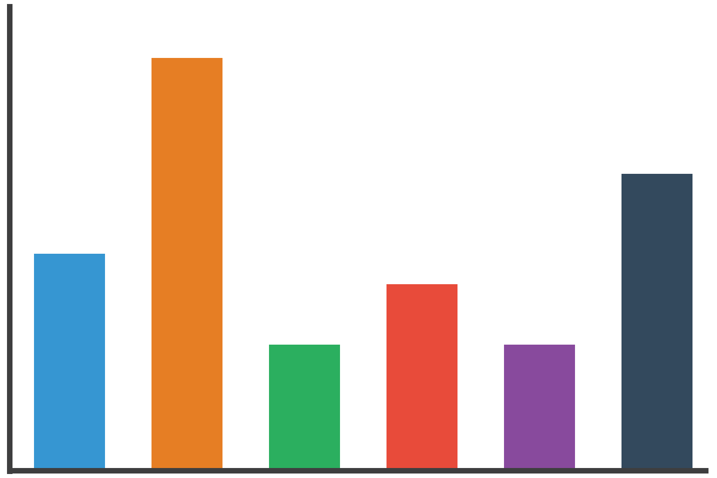
- rectilinear: scalability wrt #axes
    - 2 axes best
    - 3 problematic
    - 4+ impossible
- parallel: unfamiliarity, training time



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

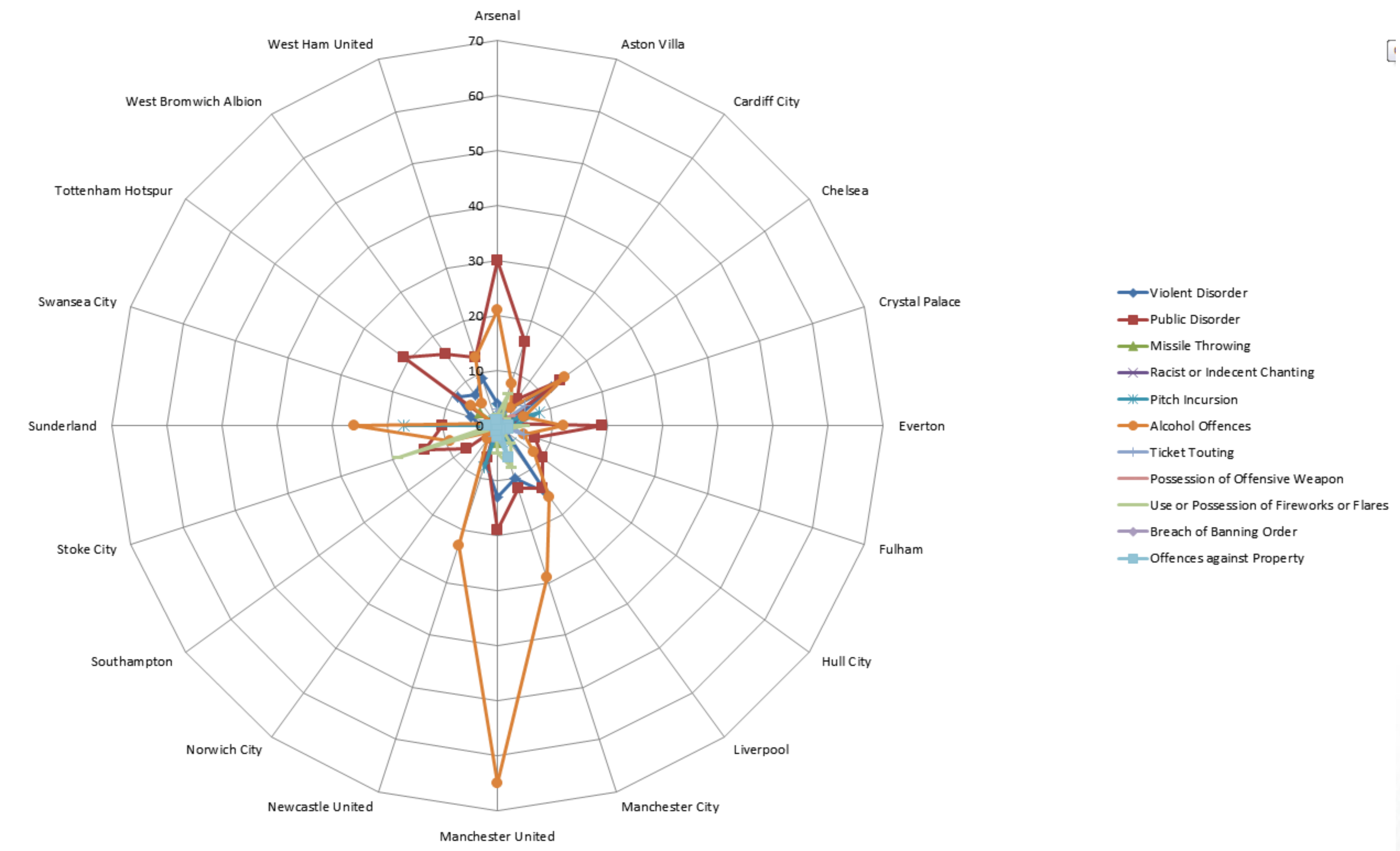




- accuracy
  - length unaligned with radial
    - less accurate than aligned with rectilinear

[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.]

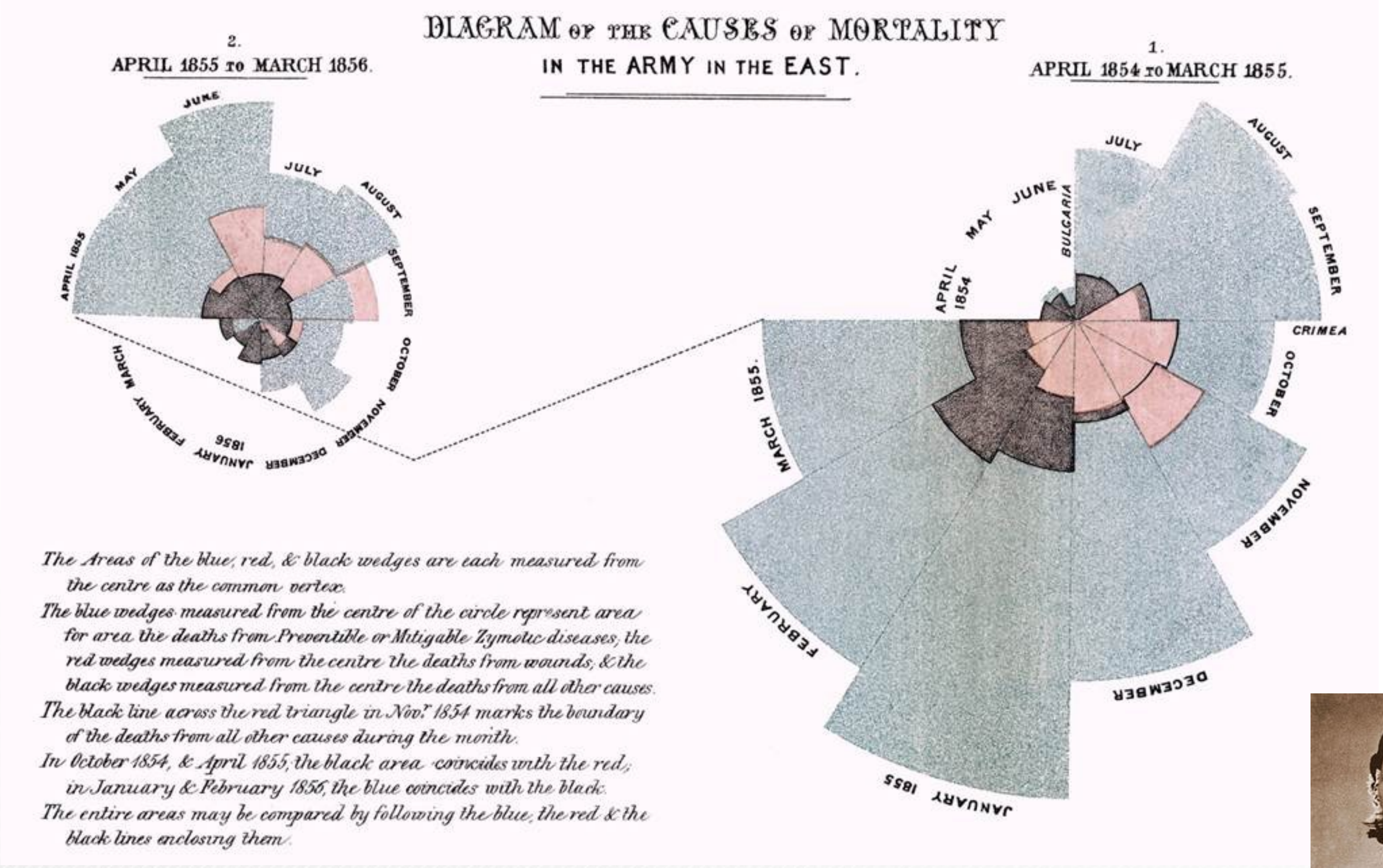
Radial Orientation: Radar Plots



LIMITATION: Not good when categories aren't cyclic

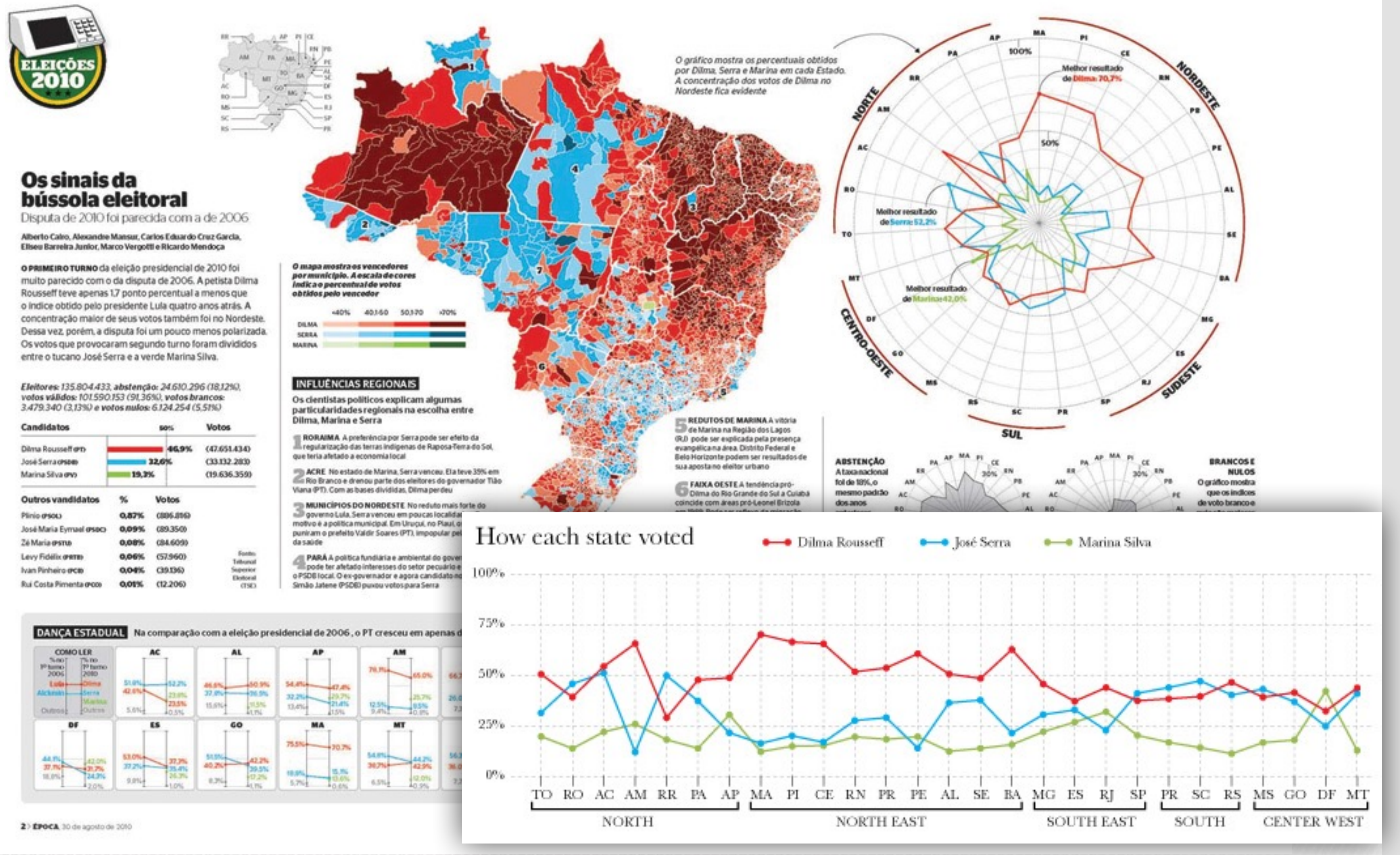
[Slide courtesy of Ben Jones]

"Diagram of the causes of mortality in the army in the East" (1858)



[Slide courtesy of Ben Jones]

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

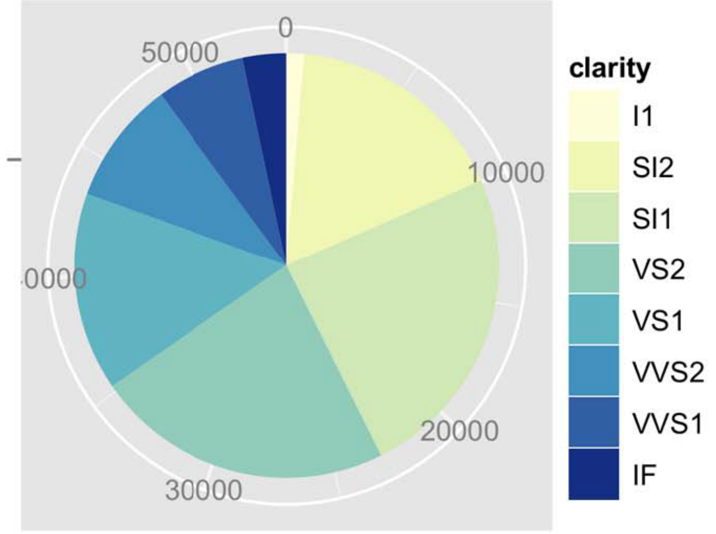


http://www.thefunctionalart.com/2012/11/radar-graphs-avoid-them-999-of-time.html

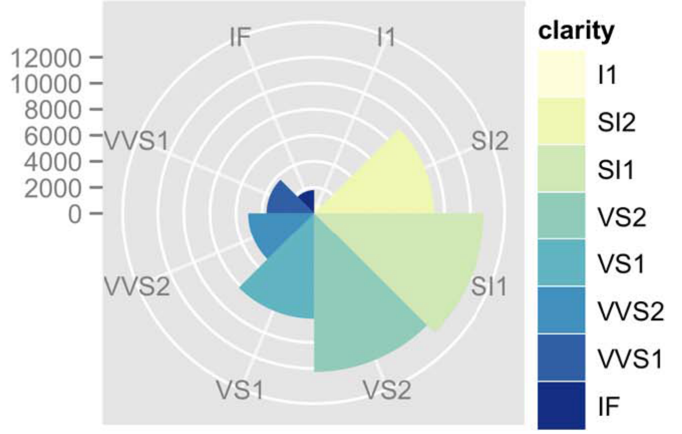
[Slide courtesy of Ben Jones]

Idioms: **pie chart, polar area chart**

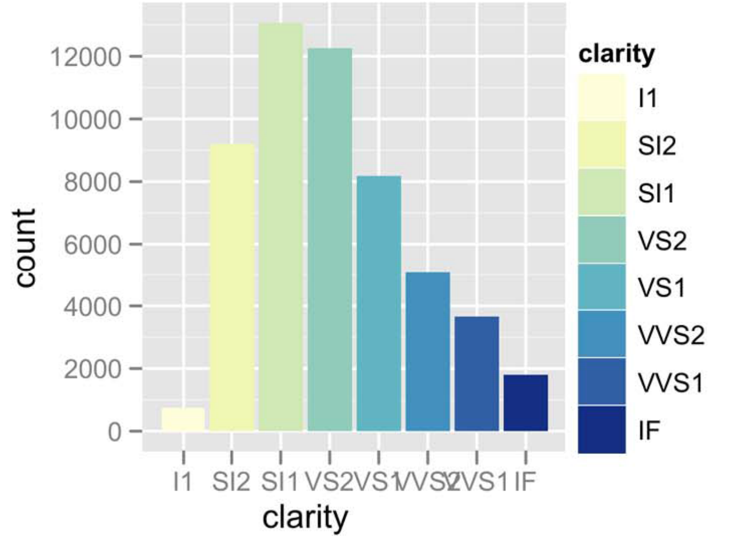
- pie chart
  - area marks with angle channel
  - accuracy: angle/area much less accurate than line length



- polar area chart
  - area marks with length channel
  - more direct analog to bar charts



- data
  - 1 categ key attrib, 1 quant value attrib
- task
  - part-to-whole judgements

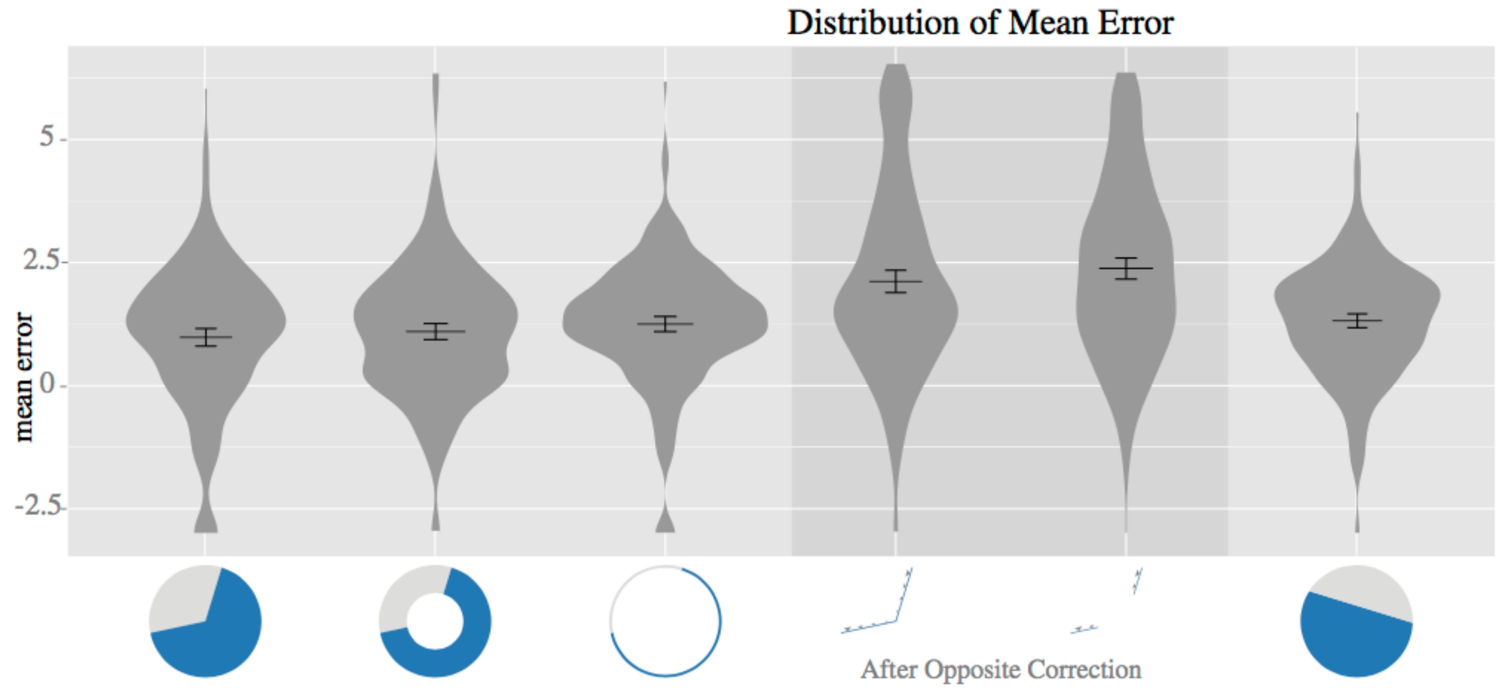


[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

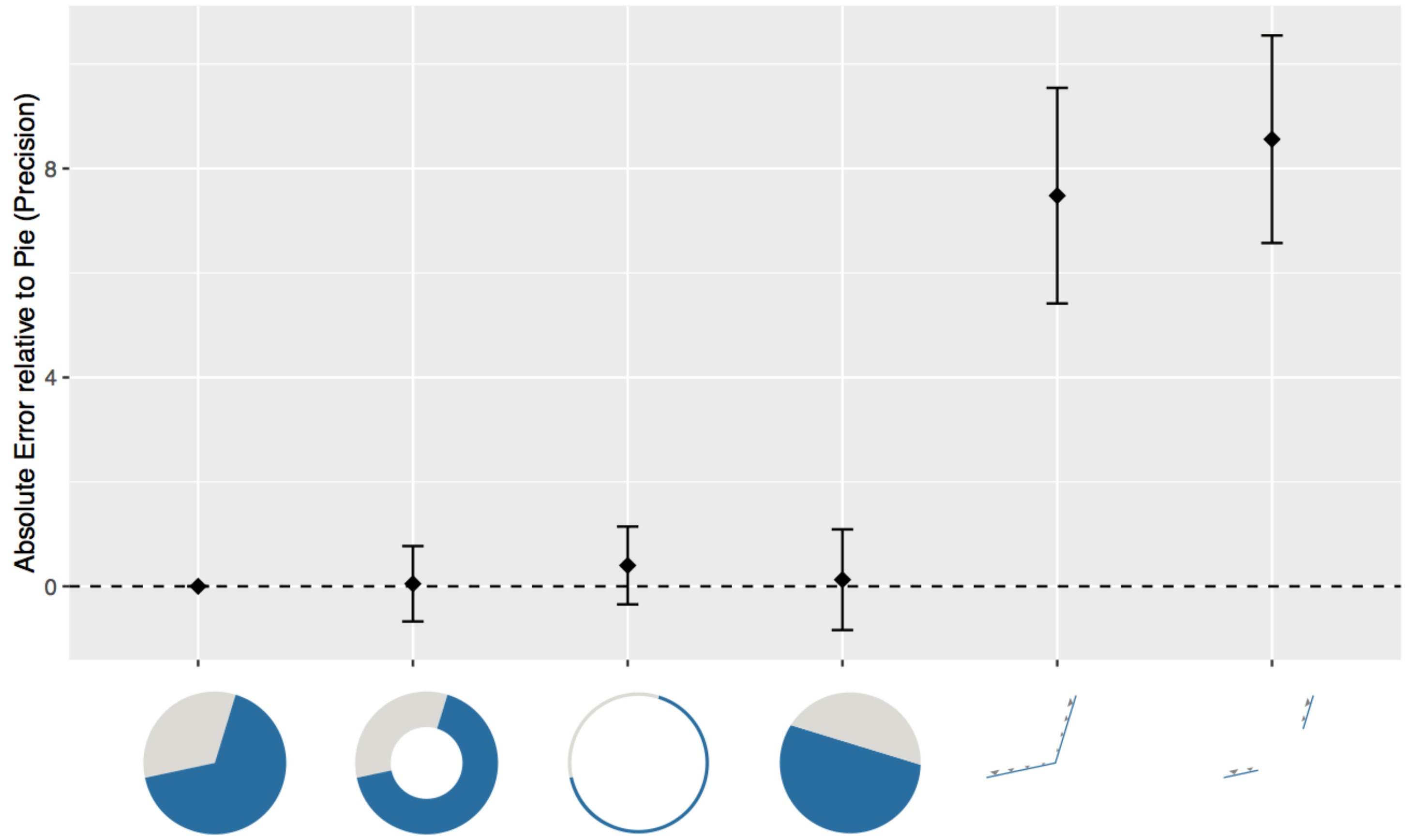
Pie chart perception

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

[Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts. Skau and Kosara. Proc. EuroVis 2016.]



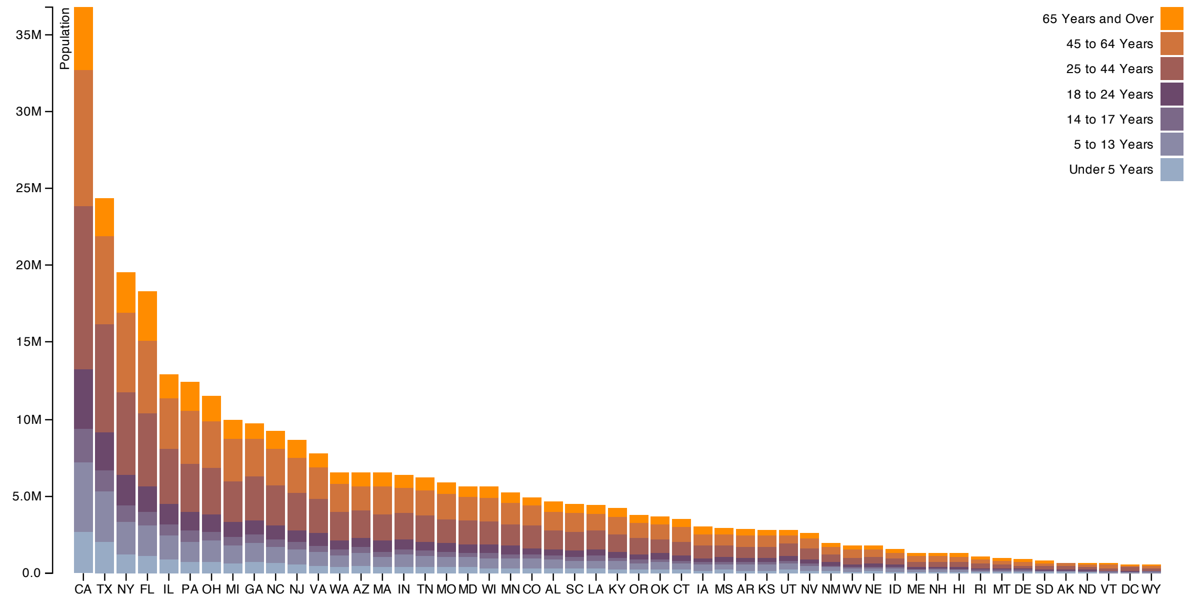
- meta-points
  - redesign of paper figures in later blog post
    - violin plots good for analysis but too detailed for presentation
  - my advice: still dubious for pie/donut charts
    - sometimes ok if just 2 attribs

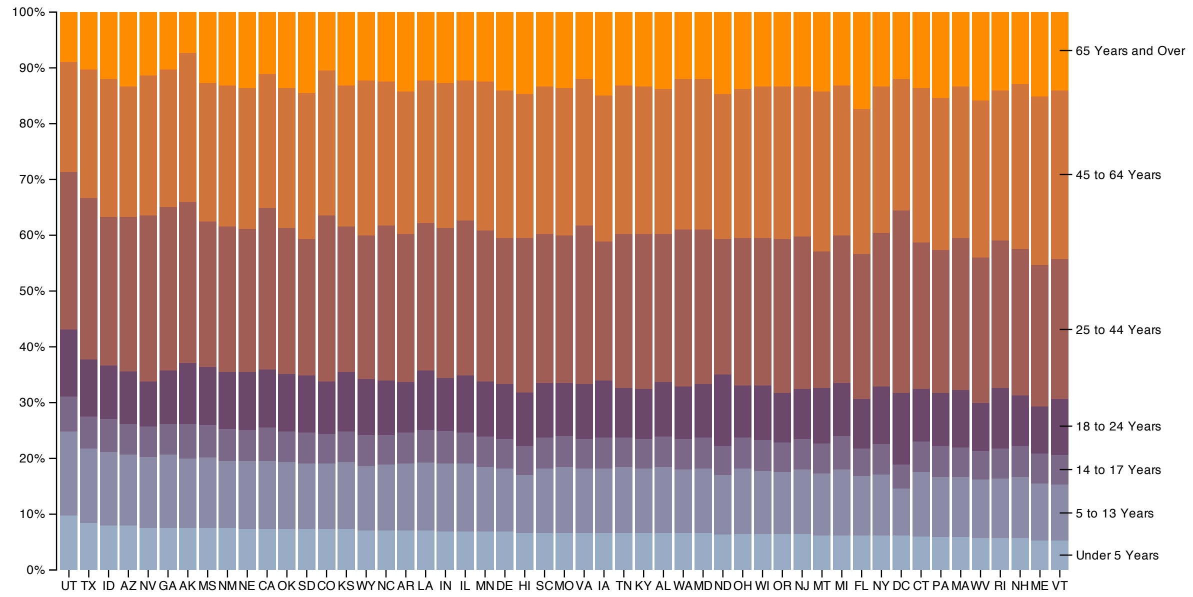


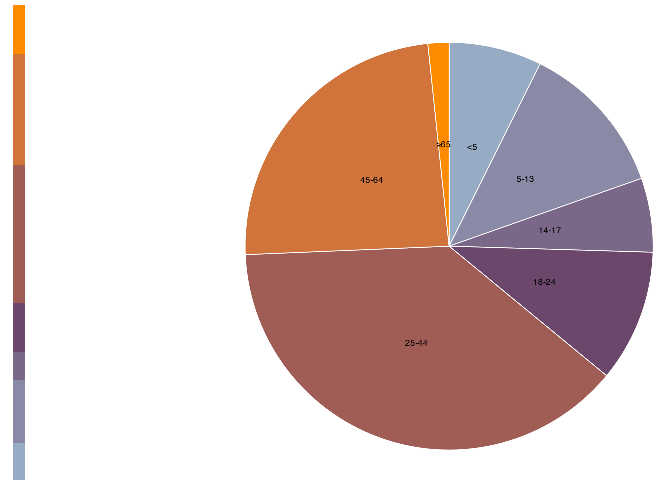
https://eagereyes.org/blog/2016/an-illustrated-tour-of-the-pie-chart-study-results

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







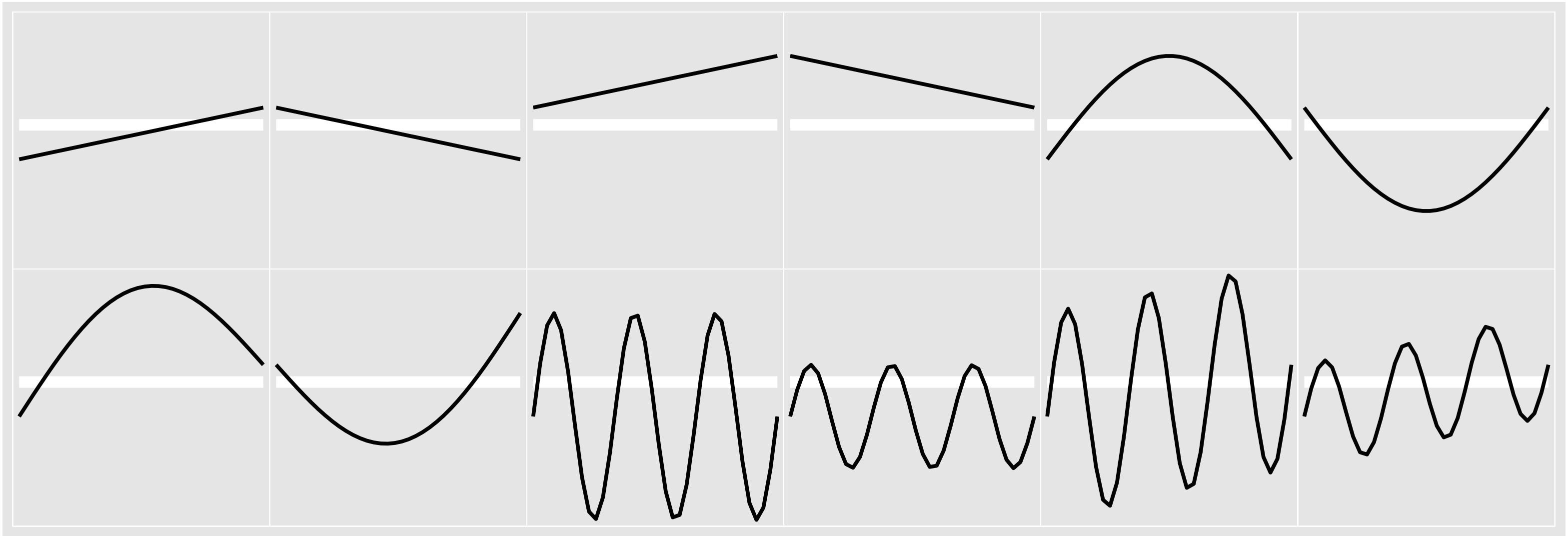
http://bl.ocks.org/mbostock/3886208,

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

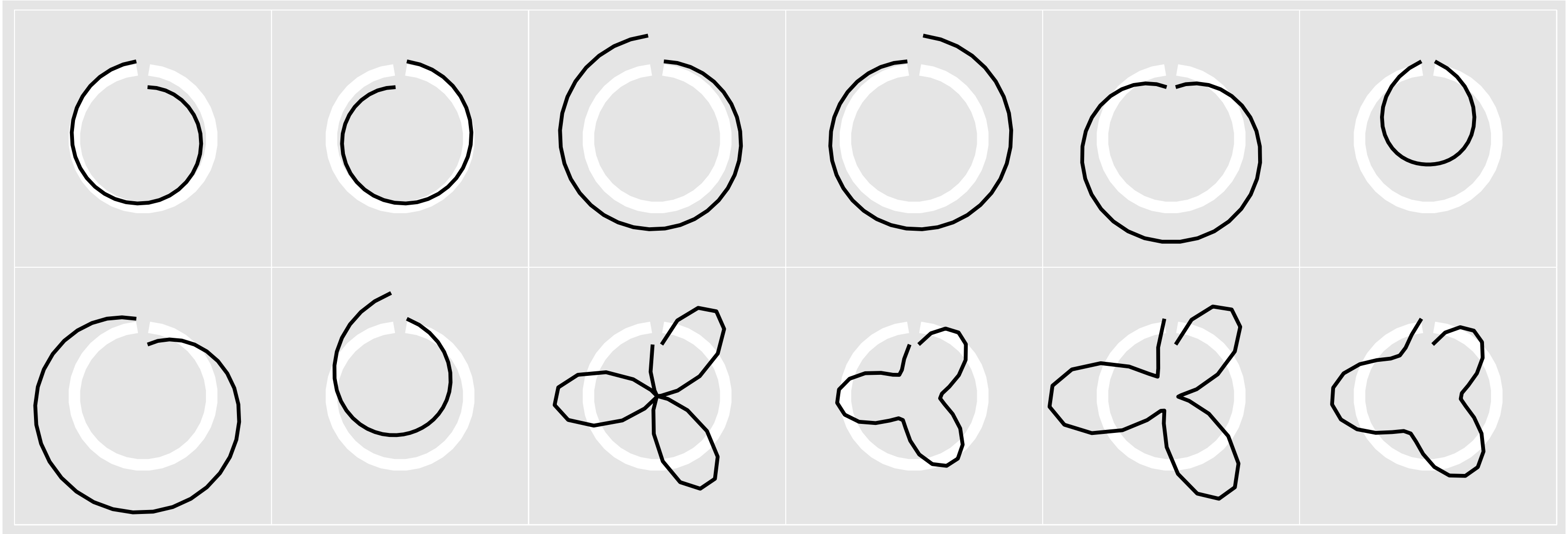
http://bl.ocks.org/mbostock/3886394.

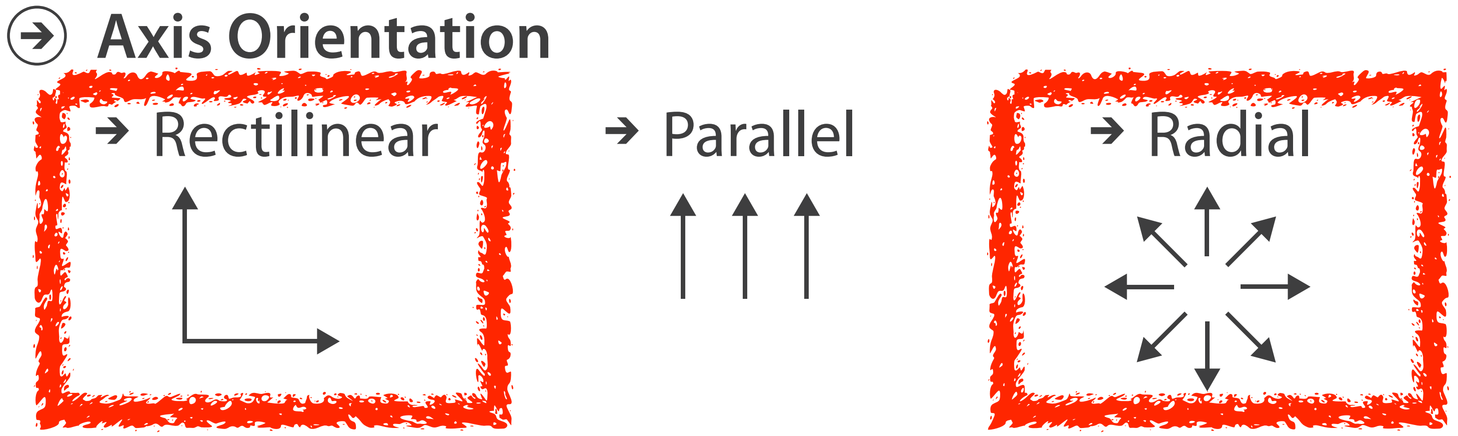
Idiom: **glyphmaps**

- rectilinear good for linear vs nonlinear trends



- radial good for cyclic patterns

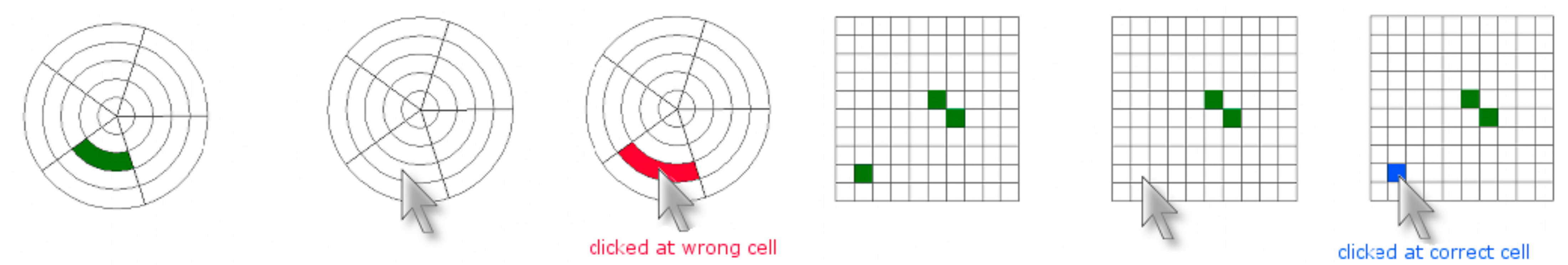




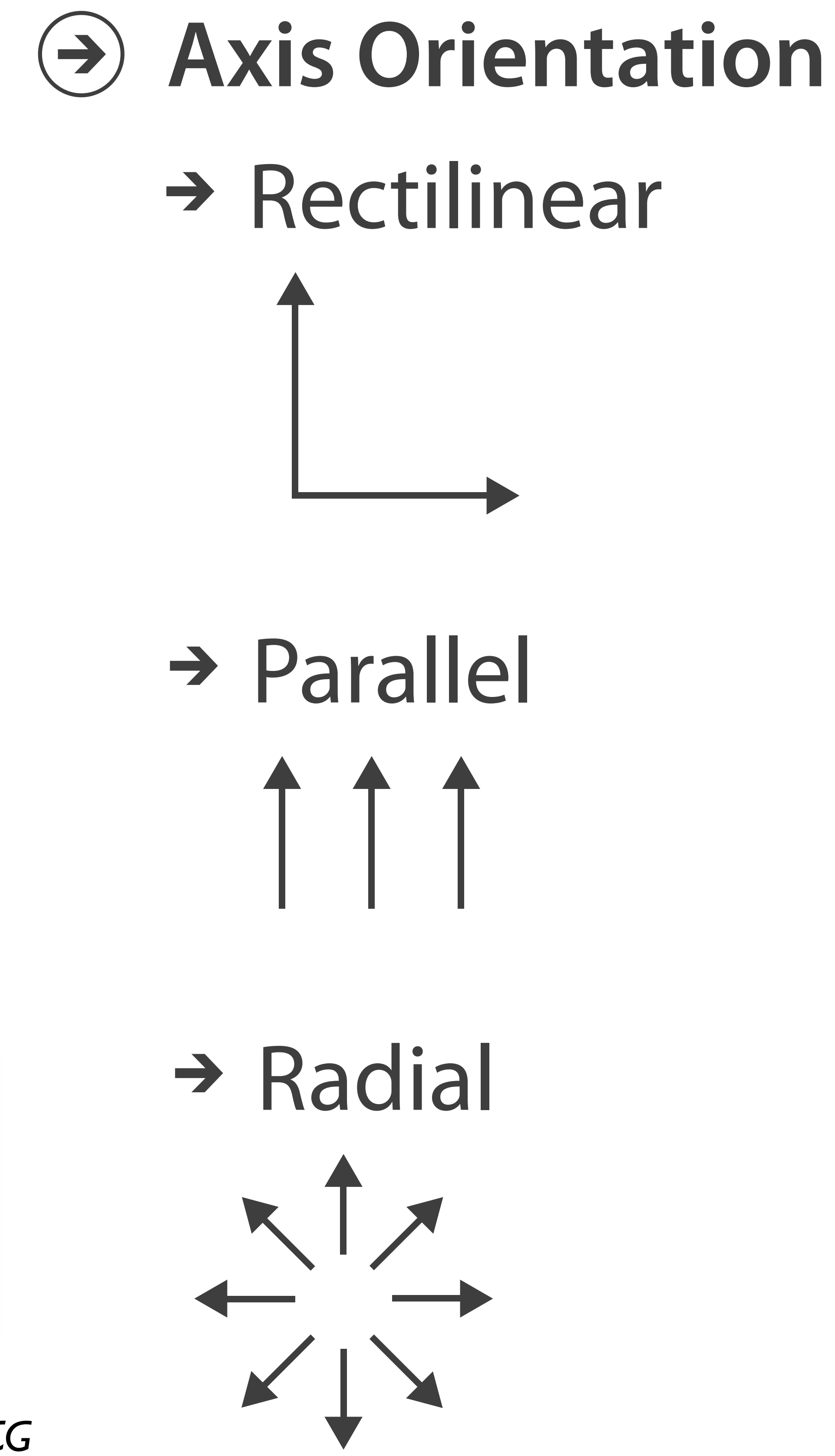
[Glyph-maps for Visually Exploring Temporal Patterns in Climate Data and Models. Wickham, Hofmann, Wickham, and Cook. Environmetrics 23:5 (2012), 382–393.]

Radial orientation

- perceptual limits
  - –polar coordinate asymmetry
    - angles lower precision than lengths
    - frequently problematic
    - sometimes can be deliberately exploited!
      - for 2 attribs of very unequal importance



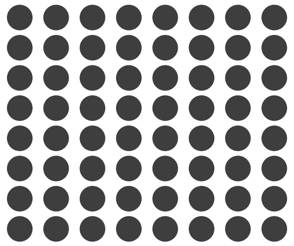
[*Uncovering Strengths and Weaknesses of Radial Visualizations - an Empirical Approach.* Diehl, Beck and Burch. IEEE TVCG (Proc. InfoVis) 16(6):935–942, 2010.]



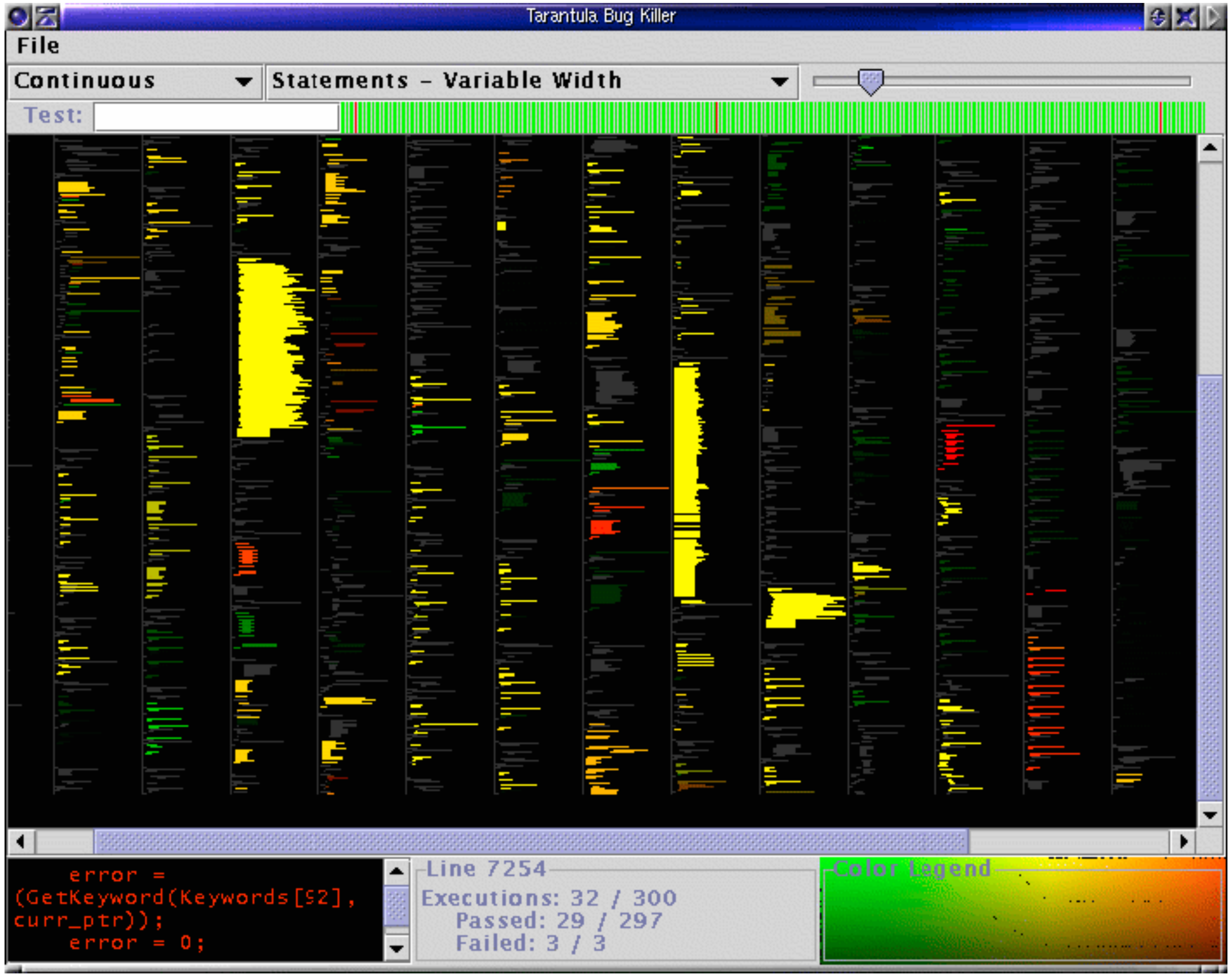
Idiom: Dense software overviews

Layout Density

→ Dense



- 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



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

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

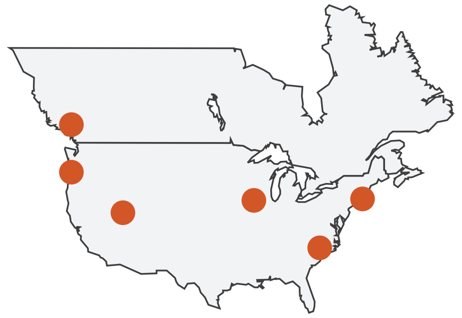
Further reading, full

- The Structure of the Information Visualization Design Space. Stuart K. Card and Jock D. Mackinlay. Proc InfoVis 1999, pp. 92-99.
- Visual Techniques for Exploring Databases.. Daniel A. Keim. KDD 1997 Tutorial Notes, 1997.
- The Grammar of Graphics, 2nd edition. Leland Wilkinson. Springer, 2005.
- Visualization Course Figures. Michael McGuffin, 2014
- The Generalized Pairs Plot. John W. Emerson, Walton A. Green, Barret Schloerke, Dianne Cook, Heike Hofmann, and Hadley Wickham. Journal of Computational and Graphical Statistics 22:1 (2012), 79-91.
- Interactive Dynamics for Visual Analysis: A Taxonomy of Tools That Support the Fluent and Flexible Use of Visualizations. Jeffrey Heer and Ben Shneiderman. Queue 10:2 (2012), 30-55.
- Sémiologie Graphique: Les diagrammes - Les réseaux - Les cartes. Jacques Bertin. Gauthier-Villard, 1967. Reissued by Editions de l'Ecole des Hautes Etudes en Sciences in 1999.
- web site: A Brief History of Data Visualization.. Michael Friendly. article: A Brief History of Data Visualization. Michael Friendly. In Handbook of Data Visualization, Computational Statistics, edited by Antony Unwin, Chun-houh Chen, and Wolfgang K. HÃ¤rdle, pp. 15-56. Springer, 2008.
- Visualizing Data. William S. Cleveland. Hobart Press, 1993.
- ThemeRiver: Visualizing Theme Changes over Time. Susan Havre, Beth Hetzler, and Lucy Nowell. Proc. InfoVis 2000, pp. 9-20.
- Stacked Graphs Geometry & Aesthetics.. Lee Byron and Martin Wattenberg. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 08) 14:6 (2008), 1245-1252.
- Bars and Lines: A Study of Graphic Communication.. Jeff Zacks and Barbara Tversky. Memory and Cognition 27:6 (1999), 1073-1079.
- The Shape Parameter of a Two-Variable Graph. William S. Cleveland, Marylyn E. McGill, and Robert McGill. Journal of the American Statistical Association 83:402 (1988), 289-300.
- Multi-Scale Banking to 45 Degrees.. Jeffrey Heer and Maneesh Agrawala. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 06) 12:5 (2006), 701-708.
- The History of the Cluster Heat Map.. Leland Wilkinson and Michael Friendly. The American Statistician 63:2 (2009), 179-184.
- Seriation and Matrix Reordering Methods: An Historical Overview.. Innar Liiv. Journal of Statistical Analysis and Data Mining 3:2 (2010), 70-91.
- Parallel Coordinates: A Tool for Visualizing Multi-Dimensional Geometry. Alfred Inselberg and Bernard Dimsdale. Proc. IEEE Conf. Visualization (Vis), 1990.
- Parallel Coordinates: Visual Multidimensional Geometry and Its Applications.. Alfred Inselberg. Springer, 2009.
- Uncovering Strengths and Weaknesses of Radial Visualizations - An Empirical Approach. Stephan Diehl, Fabian Beck, and Michael Burch. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 10) 16:6 (2010), 935-942.
- A Survey of Radial Methods for Information Visualization.. Geoffrey M. Draper, Yarden Livnat, and Richard F. Riesenfeld. IEEE Transactions on Visualization and Computer Graphics 15:5 (2009), 759-776.
- Designing Pixel-Oriented Visualization Techniques: Theory and Applications. Daniel A. Keim. IEEE Transactions on Visualization and Computer Graphics 6:1 (2000), 59-78.
- Seesoft - A Tool for Visualizing Line Oriented Software Statistics. Stephen G. Eick, Joseph L. Steffen, and Eric E Sumner, Jr. IEEE Transactions on Software Engineering 18:11 (1992), 957-968.
- Visualization of Test Information to Assist Fault Localization.. James A. Jones, Mary Jean Harrold, and John Stasko. Proc. International Conference on Software Engineering (ICSE), pp. 467-477. ACM, 2002.

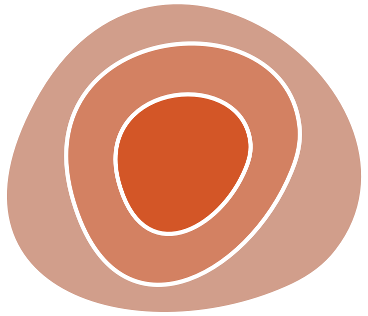
Ch 8. Arrange Spatial Data

Arrange spatial data

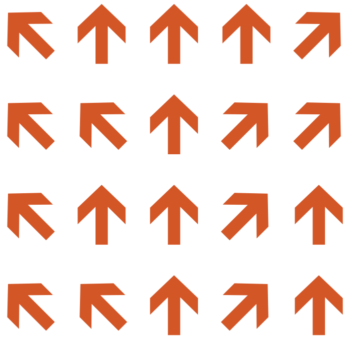
- Use Given
  - Geometry
    - *Geographic*
    - *Other Derived*



  - Spatial Fields
    - *Scalar Fields (one value per cell)*
      - *Isocontours*
      - *Direct Volume Rendering*

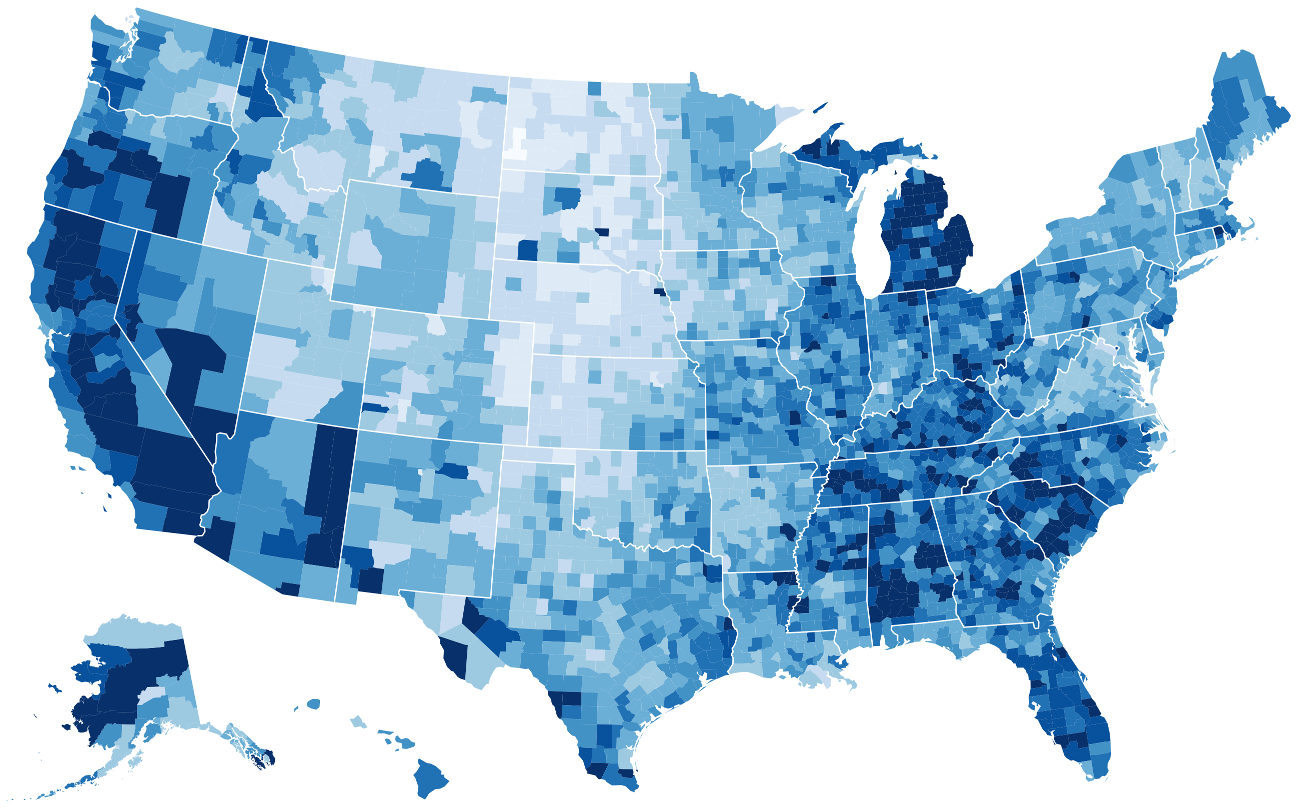


    - *Vector and Tensor Fields (many values per cell)*
      - *Flow Glyphs (local)*
      - *Geometric (sparse seeds)*
      - *Textures (dense seeds)*
      - *Features (globally derived)*



Idiom: **choropleth map**

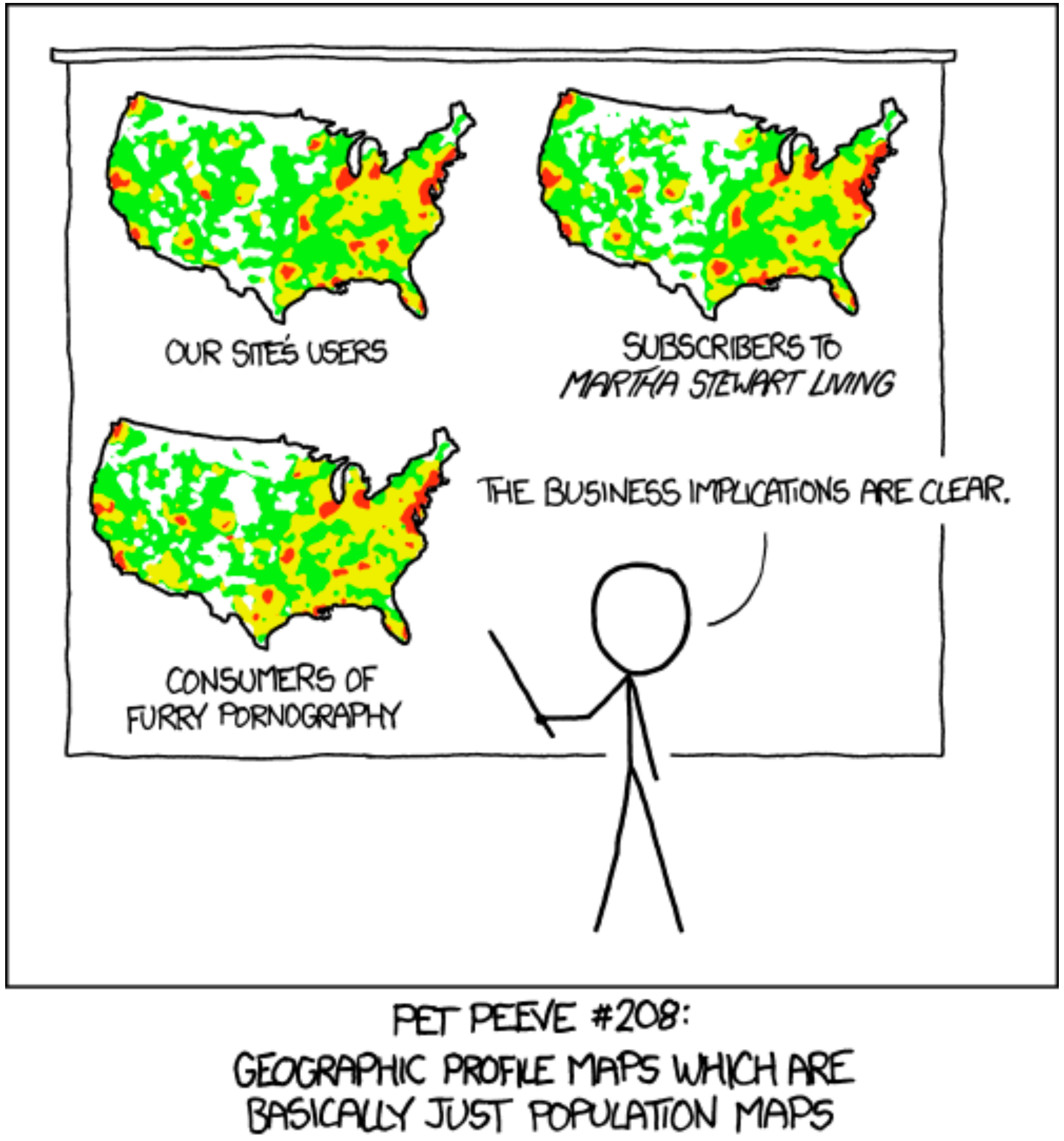
- ***use*** given spatial data
  - –when central task is understanding spatial relationships
- data
  - –geographic geometry
  - –table with 1 quant attribute per region
- encoding
  - –use given geometry for area mark boundaries
  - –sequential segmented colormap *[more later]*
  - –(geographic heat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



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

Population maps trickiness

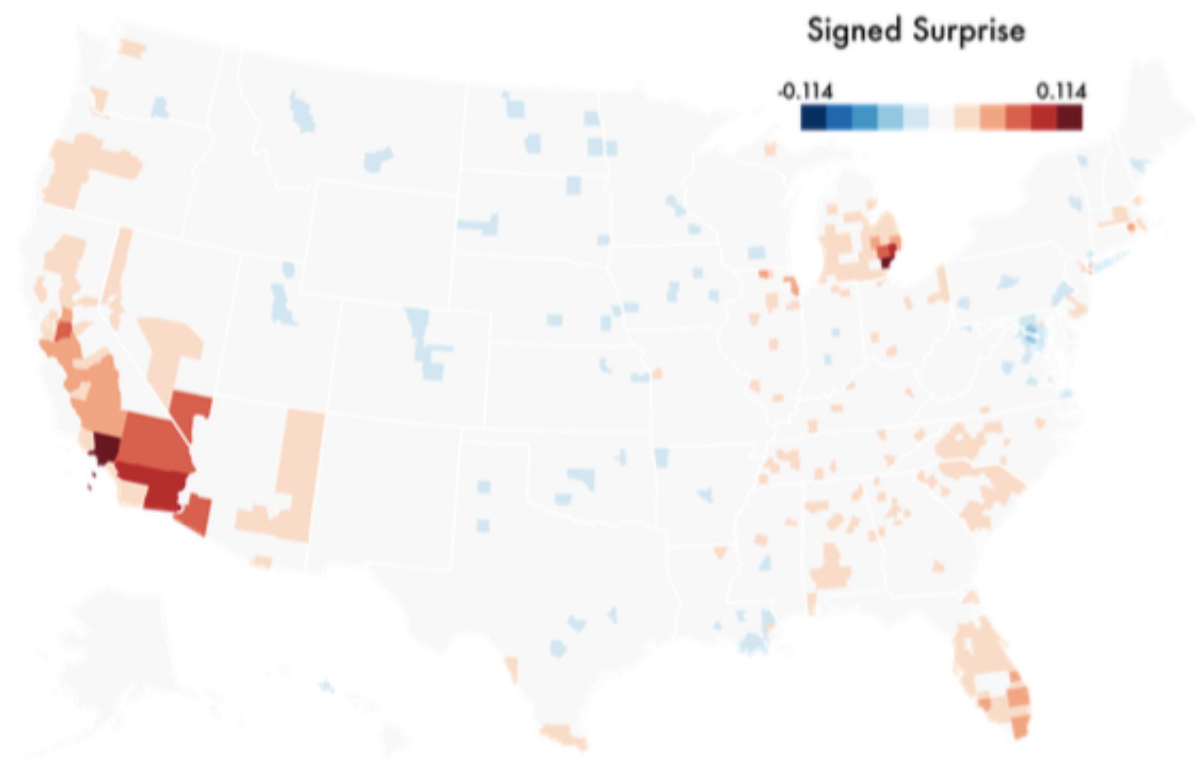
- beware!
- absolute vs relative again
  - population density vs per capita
- investigate with Ben Jones Tableau Public demo
  - http://public.tableau.com/profile/ben.jones#!/vizhome/PopVsFin/PopVsFin *Are Maps of Financial Variables just Population Maps?*
    - yes, unless you look at per capita (relative)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

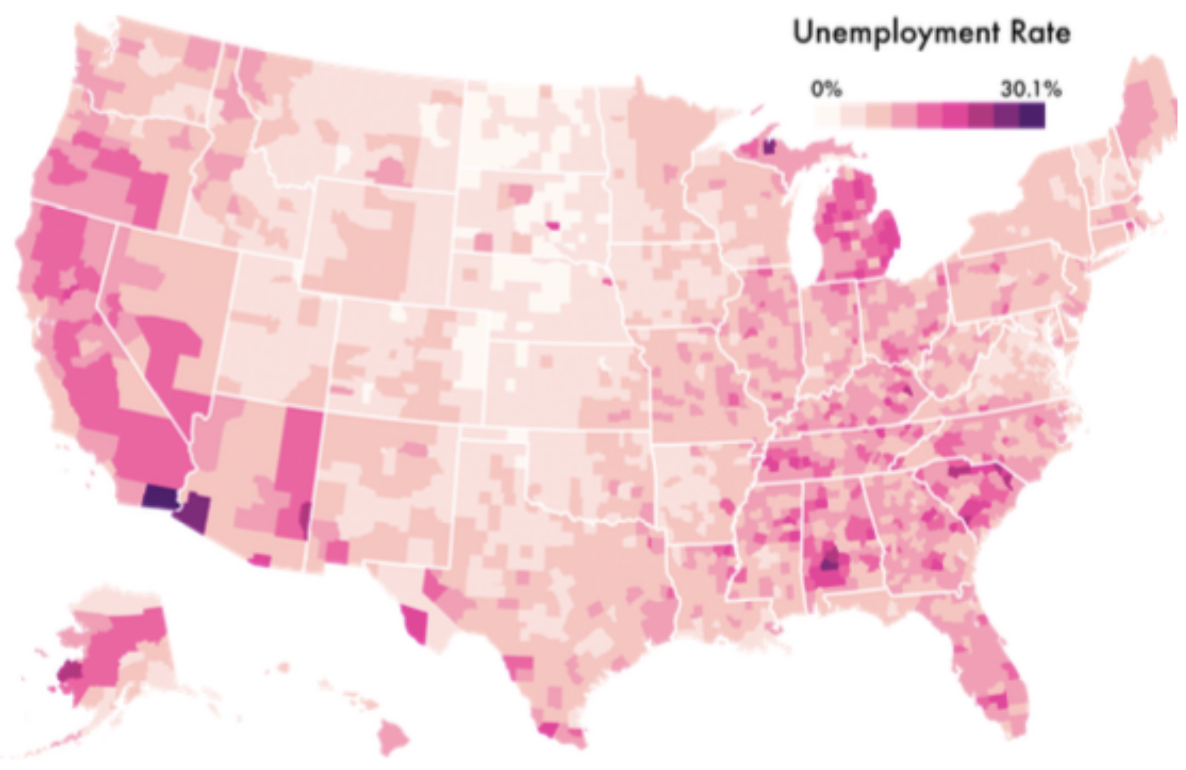


[https://xkcd.com/1138]

Idiom: Bayesian surprise maps

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



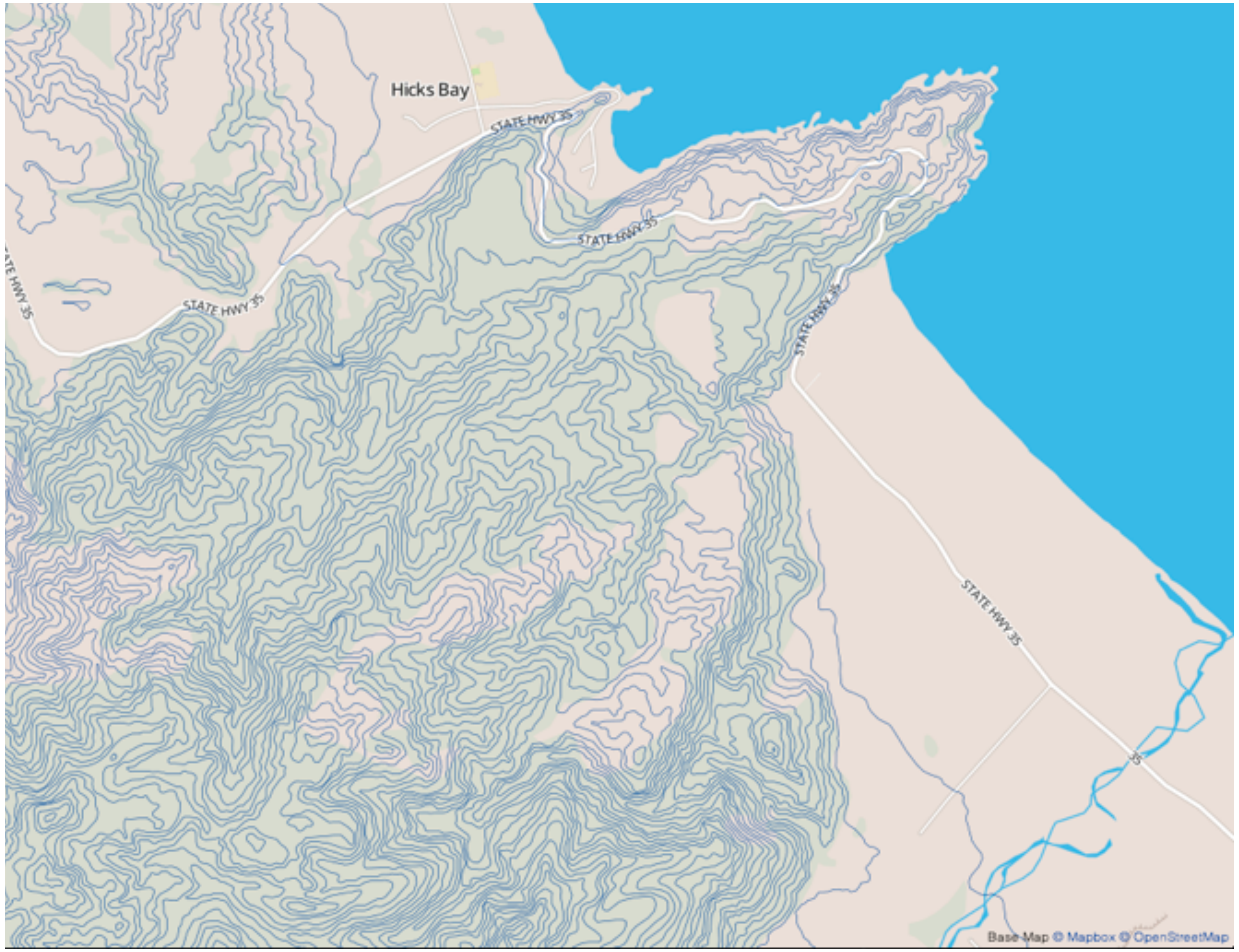


[Surprise! Bayesian Weighting for De-Biasing Thematic Maps. Correll and Heer. Proc InfoVis 2016]

https://medium.com/@uwdata/surprise-maps-showing-the-unexpected-e92b67398865 https://idl.cs.washington.edu/papers/surprise-maps/

Idiom: **topographic map**

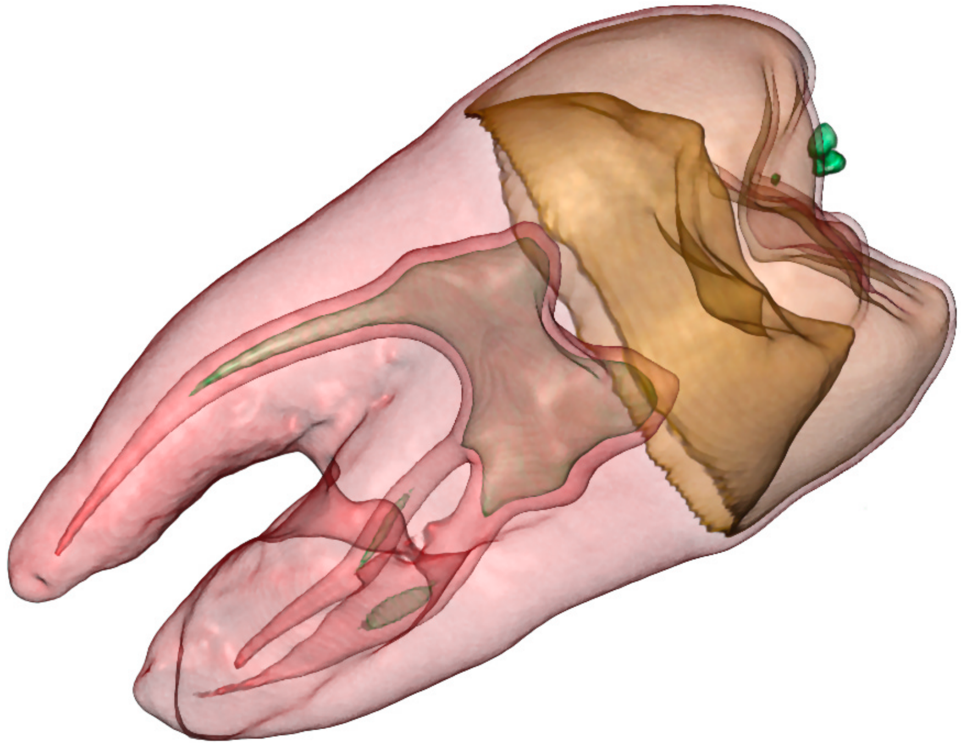
- data
  - –geographic geometry
  - –scalar spatial field
    - 1 quant attribute per grid cell
- derived data
  - –isoline geometry
    - isocontours computed for specific levels of scalar values



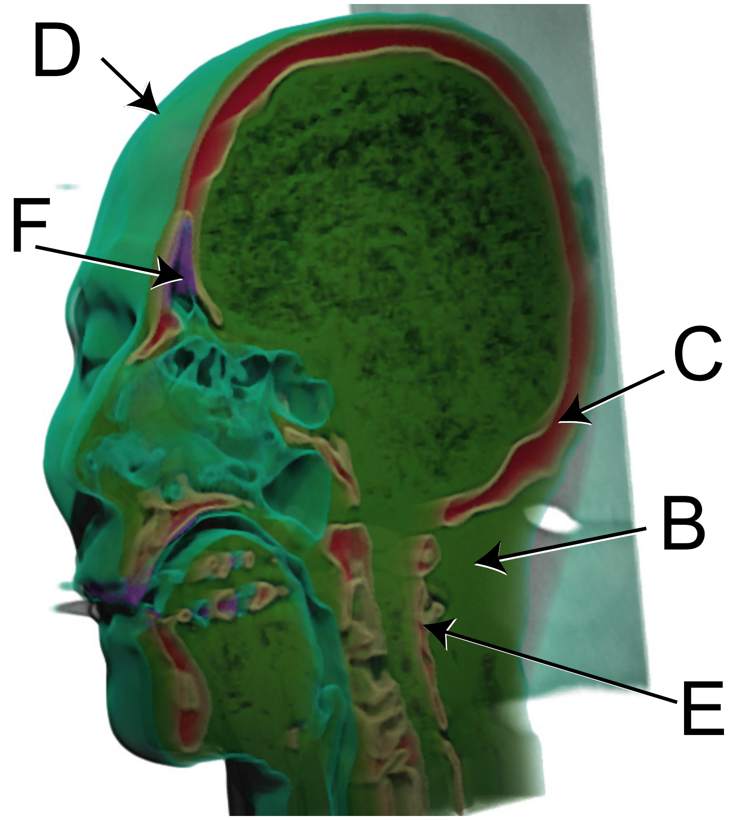
Land Information New Zealand Data Service

Idioms: **isosurfaces, direct volume rendering**

- data
  - scalar spatial field
    - 1 quant attribute per grid cell
- task
  - shape understanding, spatial relationships
- isosurface
  - derived data: isocontours computed for specific levels of scalar values
- direct volume rendering
  - transfer function maps scalar values to color, opacity



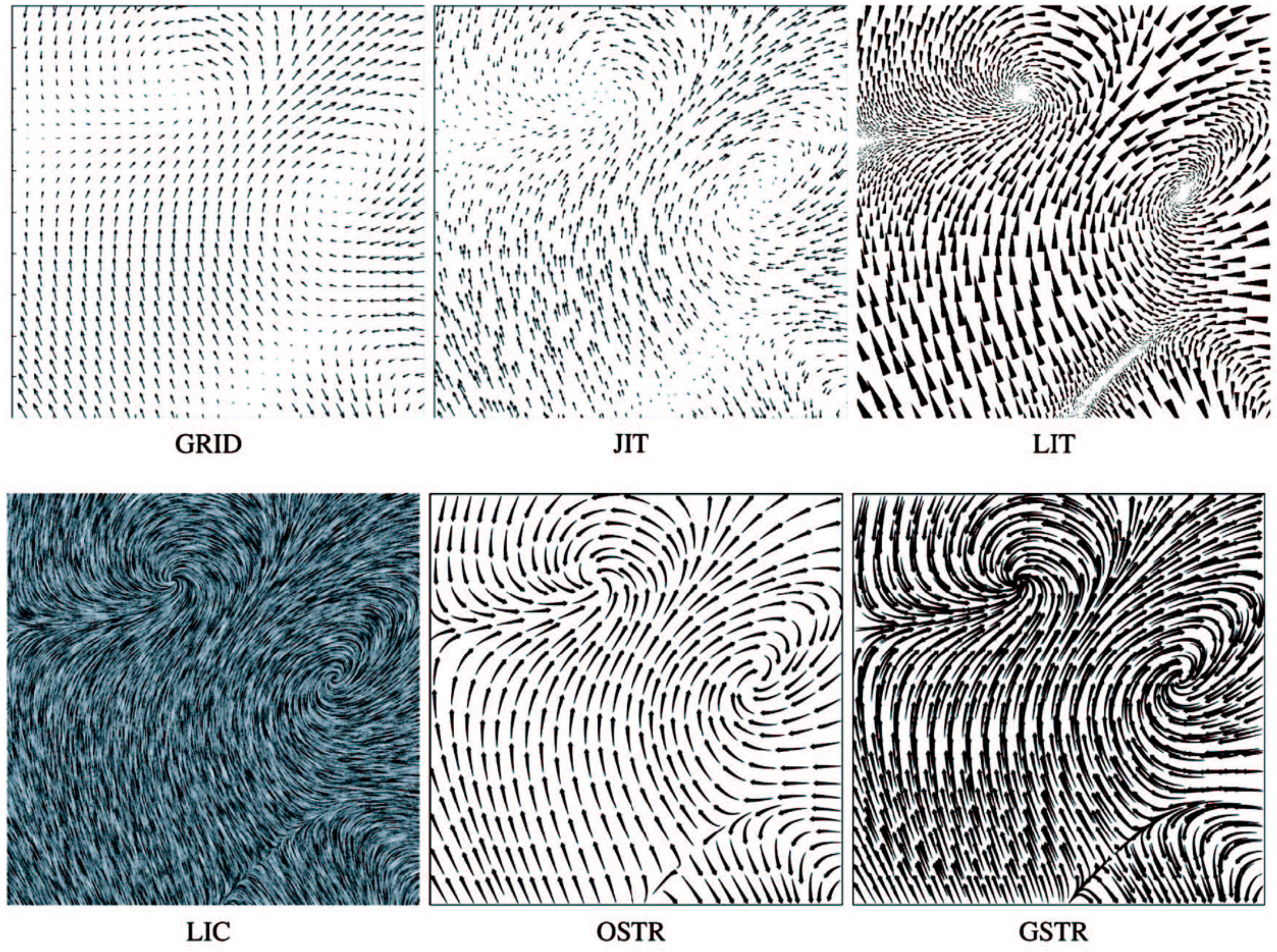
[Interactive Volume Rendering Techniques. Kniss. Master's thesis, University of Utah Computer Science, 2002.]



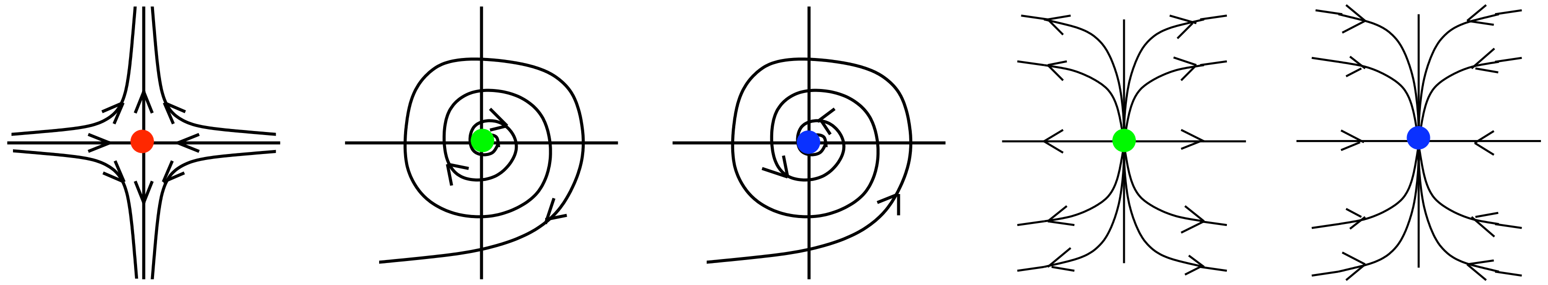
[Multidimensional Transfer Functions for Volume Rendering. Kniss, Kindlmann, and Hansen. In The Visualization Handbook, edited by Charles Hansen and Christopher Johnson, pp. 189–210. Elsevier, 2005.]

Vector and tensor fields

- data
  - many attribs per cell
- idiom families
  - flow glyphs
    - purely local
  - geometric flow
    - derived data from tracing particle trajectories
    - sparse set of seed points
  - texture flow
    - derived data, dense seeds
  - feature flow
    - global computation to detect features
      - encoded with one of methods above



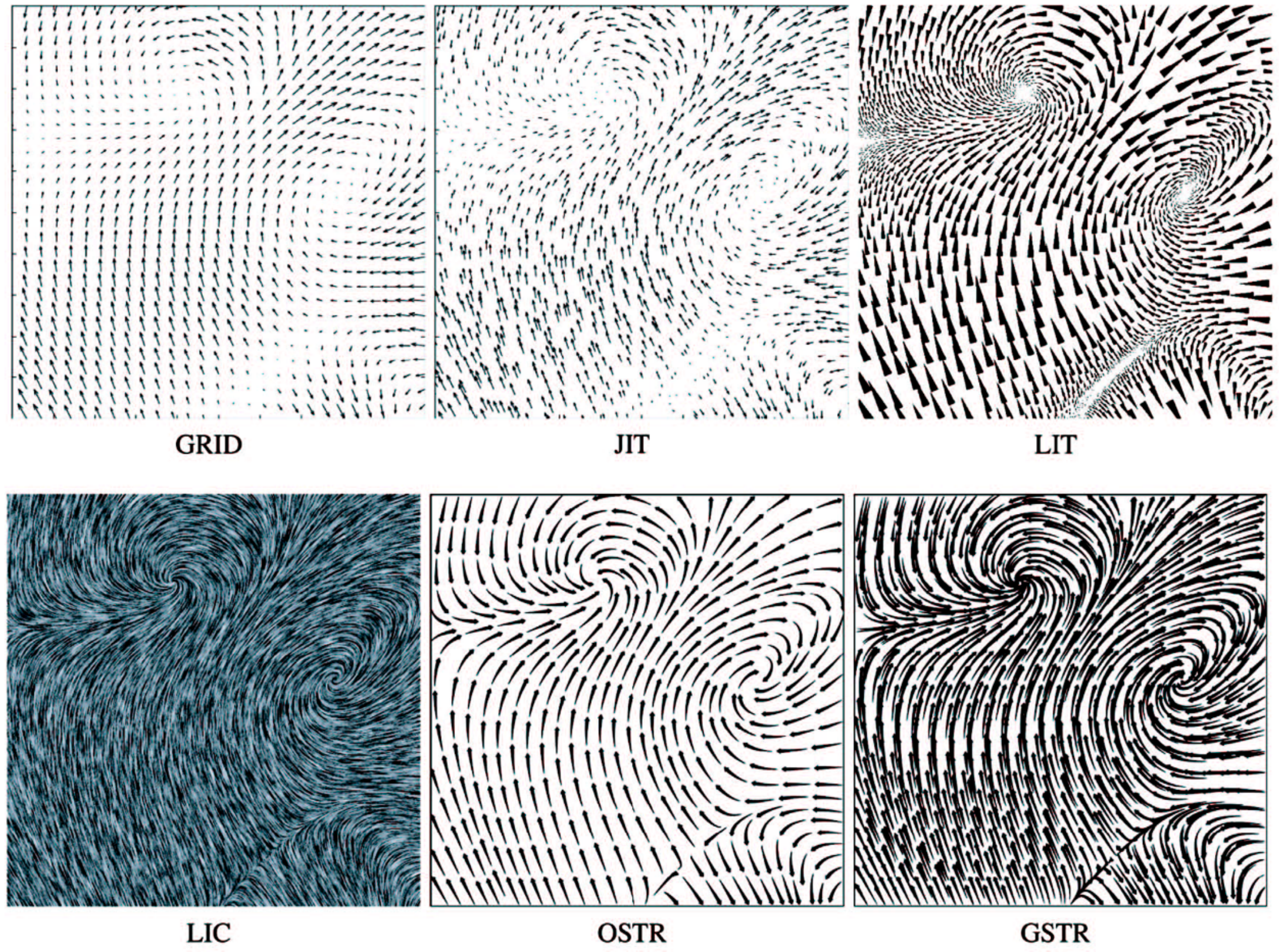
[Comparing 2D vector field visualization methods: A user study. Laidlaw et al. IEEE Trans. Visualization and Computer Graphics (TVCG) 11:1 (2005), 59–70.]



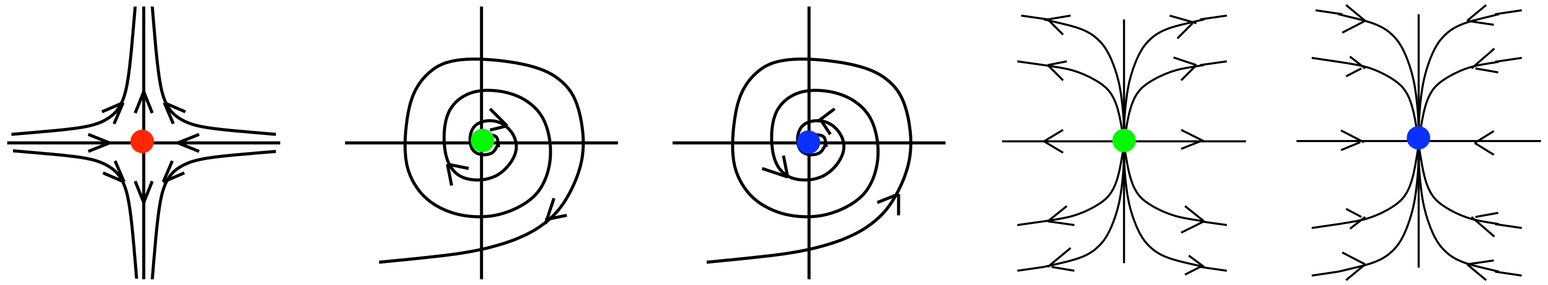
[Topology tracking for the visualization of time-dependent two-dimensional flows. Tricoche, Wischgoll, Scheuermann, and Hagen. Computers & Graphics 26:2 (2002), 249–257.]

Vector fields

- empirical study tasks
  - finding critical points, identifying their types
  - identifying what type of critical point is at a specific location
  - predicting where a particle starting at a specified point will end up (advection)



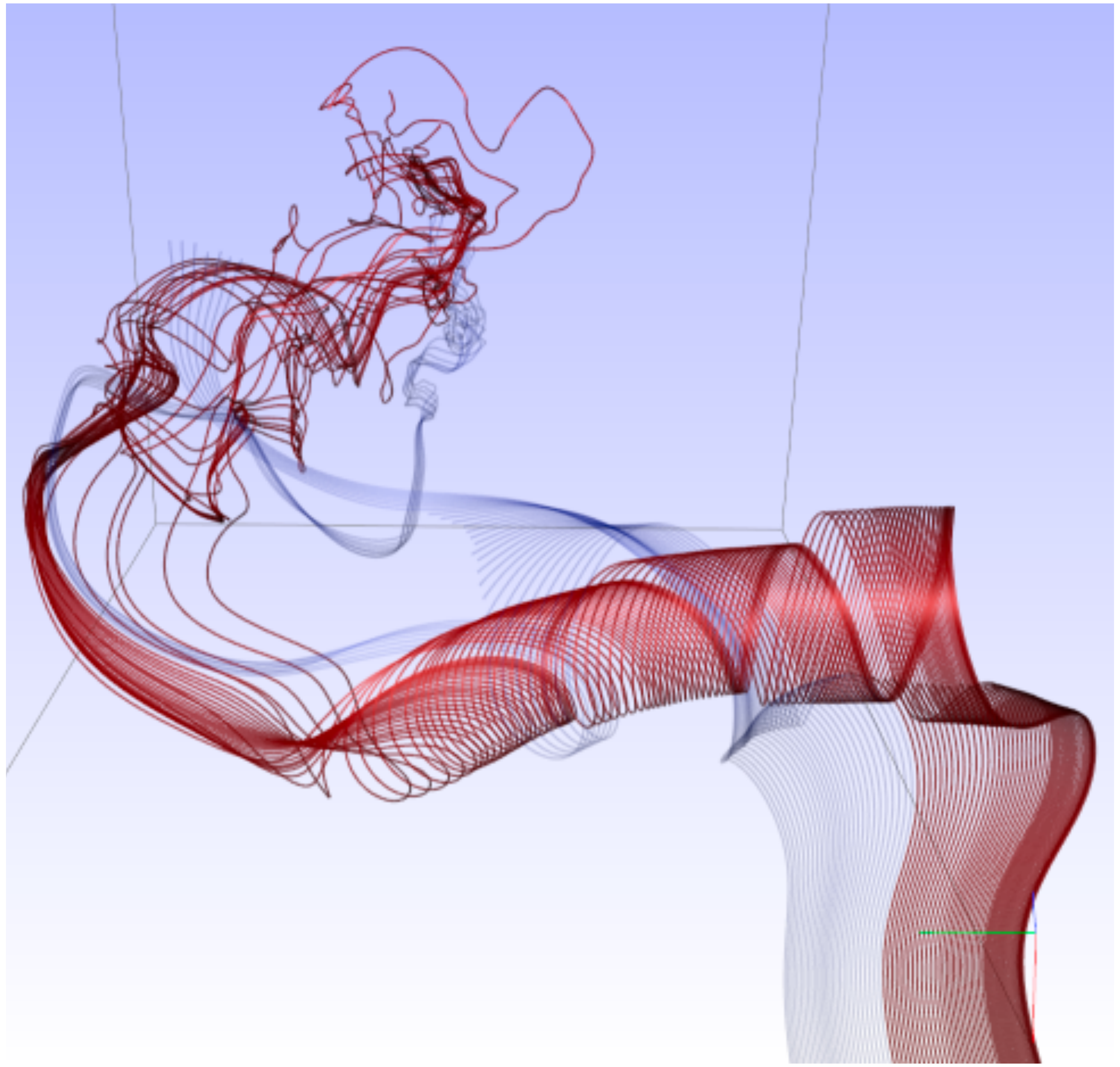
[Comparing 2D vector field visualization methods: A user study. Laidlaw et al. IEEE Trans. Visualization and Computer Graphics (TVCG) 11:1 (2005), 59–70.]



[Topology tracking for the visualization of time-dependent two-dimensional flows. Tricoche, Wischgoll, Scheuermann, and Hagen. Computers & Graphics 26:2 (2002), 249–257.]

Idiom: **similarity-clustered streamlines**

- data
  - 3D vector field
- derived data (from field)
  - streamlines: trajectory particle will follow
- derived data (per streamline)
  - curvature, torsion, tortuosity
  - signature: complex weighted combination
  - compute cluster hierarchy across all signatures
  - encode: color and opacity by cluster
- tasks
  - find features, query shape
- scalability
  - millions of samples, hundreds of streamlines



[Similarity Measures for Enhancing Interactive Streamline Seeding. McLoughlin,. Jones, Laramee, Malki, Masters, and. Hansen. IEEE Trans. Visualization and Computer Graphics 19:8 (2013), 1342–1353.]

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014.
  - *Chap 8: Arrange Spatial Data*
- How Maps Work: Representation, Visualization, and Design. MacEachren. Guilford Press, 1995.
- Overview of visualization. Schroeder and. Martin. In The Visualization Handbook, edited by Charles Hansen and Christopher Johnson, pp. 3–39. Elsevier, 2005.
- Real-Time Volume Graphics. Engel, Hadwiger, Kniss, Reza-Salama, and Weiskopf. AK Peters, 2006.
- Overview of flow visualization. Weiskopf and Erlebacher. In The Visualization Handbook, edited by Charles Hansen and Christopher Johnson, pp. 261–278. Elsevier, 2005.

Further reading, full

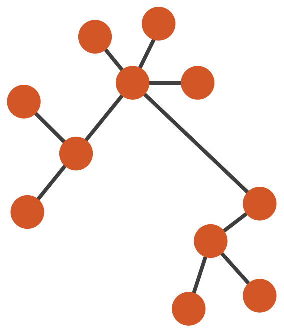
- web site: A Brief History of Data Visualization, http://www.datavis.ca/milestones, Michael Friendly. article: A Brief History of Data Visualization. Michael Friendly. In Handbook of Data Visualization, Computational Statistics, edited by Antony Unwin, Chun-houh Chen, and Wolfgang K. Härdle, pp. 15-56. Springer, 2008.
- The Evolution of Thematic Cartography/A Research Methodology and Historical Review.. Alan M. MacEachren. The Canadian Cartographer 16:1 (1979), 17-33.
- Thematic Cartography and Geovisualization, Third edition. Terry A. Slocum, Robert B. McMaster, Fritz C. Kessler, and Hugh H. Howard. Prentice Hall, 2008.
- Overview of Visualization. William J. Schroeder and Kenneth M. Martin. In The Visualization Handbook, edited by Charles Hansen and Christopher Johnson, pp. 3-39. Elsevier, 2005.
- Overview of Volume Rendering.. Arie Kaufman and Klaus Mueller. In The Visualization Handbook, edited by Charles C. Hansen and Christopher R. Johnson, pp. 127-174. Elsevier, 2005.
- Marching Cubes: A High Resolution 3D Surface Construction Algorithm.. William E. Lorensen and Harvey E. Cline. Computer Graphics (Proc. SIGGRAPH 87) 21:4 (1987), 163-169.
- A Survey of the Marching Cubes Algorithm. Timothy S. Newman and Hong Yi. Computers & Graphics 30:5 (2006), 854-879.
- Simplifying Flexible Isosurfaces Using Local Geometric Measures. Hamish Carr, Jack Snoeyink, and Michiel van de Panne. Proc. IEEE Conf. Visualization (Vis) 2004, pp. 497-504.
- Real-Time Volume Graphics. Klaus Engel, Markus Hadwiger, Joe Kniss, Christof Reza-Salama, and Daniel Weiskopf. A K Peters, 2006.
- Volume Rendering.. Robert A. Drebin, Loren C. Carpenter, and Pat Hanrahan. Computer Graphics (Proc. SIGGRAPH 88) 22:4 (1988), 65-74.
- Display of Surfaces from Volume Data. Marc Levoy. IEEE Computer Graphics and Applications 8:3 (1988), 29-37.
- Interactive Volume Rendering Techniques. Joe Kniss. Master's thesis, University of Utah, Department of Computer Science, 2002. a
- Multidimensional Transfer Functions for Volume Rendering. Joe Kniss, Gordon Kindlmann, and Charles Hansen. In The Visualization Handbook, edited by Charles Hansen and Christopher Johnson, pp. 189-210. Elsevier, 2005.
- Overview of Flow Visualization. Daniel Weiskopf and Gordon Erlebacher. In The Visualization Handbook, edited by Charles Hansen and Christopher Johnson, pp. 261-278. Elsevier, 2005.
- Over Two Decades of Integration-Based Geometric Flow Visualization. Tony McLouglin, Robert S. Laramee, Ronald Peikert, Frits H. Post, and Min Chen. Computer Graphics Forum (Proc. Eurographics 09, State of the Art Reports) 6:29 (2010), 1807-1829.
- The State of the Art in Flow Visualization: Dense and Texture-Based Techniques. Robert S. Laramee, Helwig Hauser, Helmut Doleisch, Benjamin Vrolijk, Frits H. Post, and Daniel Weiskopf. Computer Graphics Forum (Proc. Eurographics 04) 23:2 (2004), 203-221.
- The State of the Art in Flow Visualisation: Feature Extraction and Tracking. Frits H. Post, Benjamin Vrolijka, Helwig Hauser, Robert S. Laramee, and Helmut Doleisch. Computer Graphics Forum (Proc. Eurographics 03) 22:4 (2003), 1-17.
- Imaging Vector Fields Using Line Integral Convolution. Brian Cabral and Leith Casey Leedom. SIGGRAPH 1993, pp. 263-270.
- Visualization and Processing of Tensor Fields. Joachim Weickert and Hans Hagen, editors. Springer, 2006.
- An Introduction to Visualization of Diffusion Tensor Imaging and Its Applications. A. Vilanova, S. Zhang, G. Kindlmann, and D. Laidlaw. In Visualization and Processing of Tensor Fields, pp. 121-153. Springer, 2006.
- Superquadric Tensor Glyphs. Gordon Kindlmann. Proc. Eurographics/IEEE Conference on Visualization (VisSym), pp. 147-154. Eurographics, 2004.

Ch 9. Arrange Networks and Trees

Arrange networks and trees

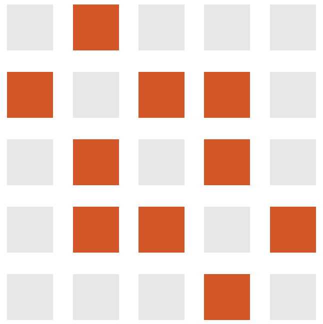
Node–Link Diagrams
Connection Marks

✔ NETWORKS ✔ TREES



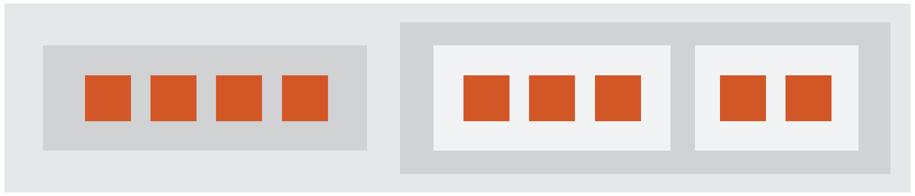
Adjacency Matrix
Derived Table

✔ NETWORKS ✔ TREES



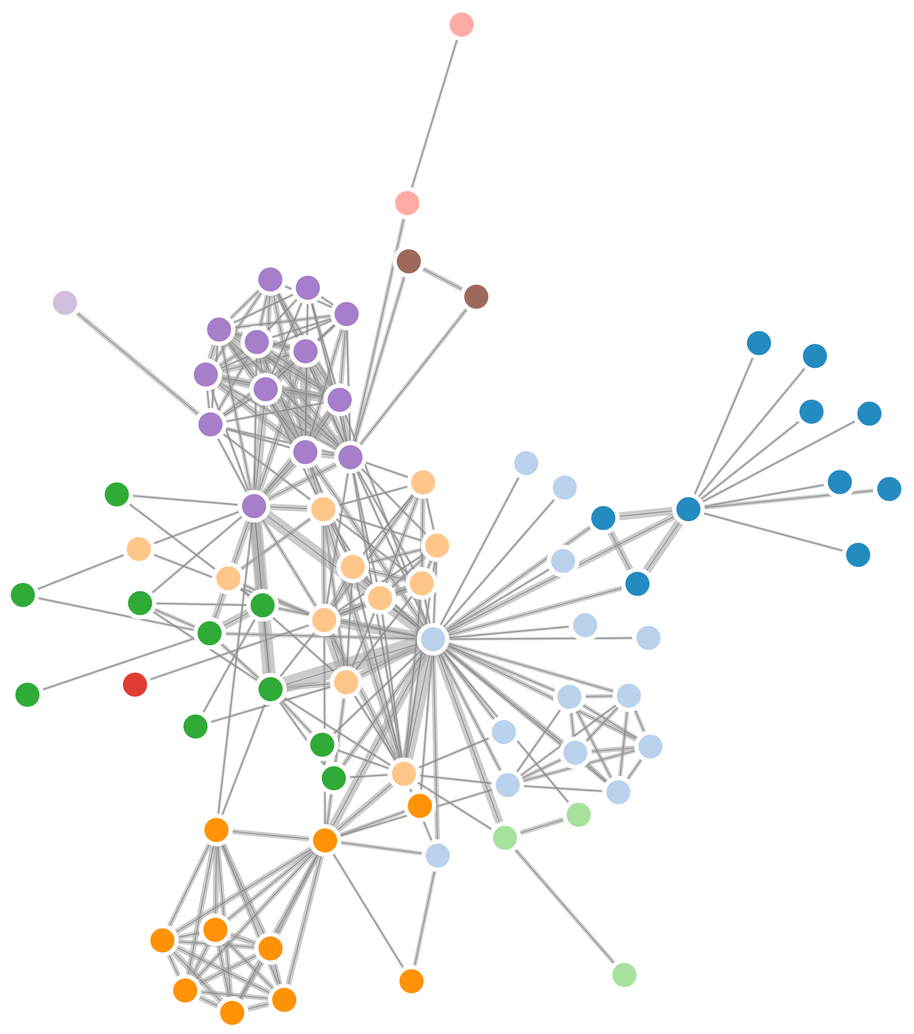
Enclosure
Containment Marks

✕ NETWORKS ✔ TREES



Idiom: **force-directed placement**

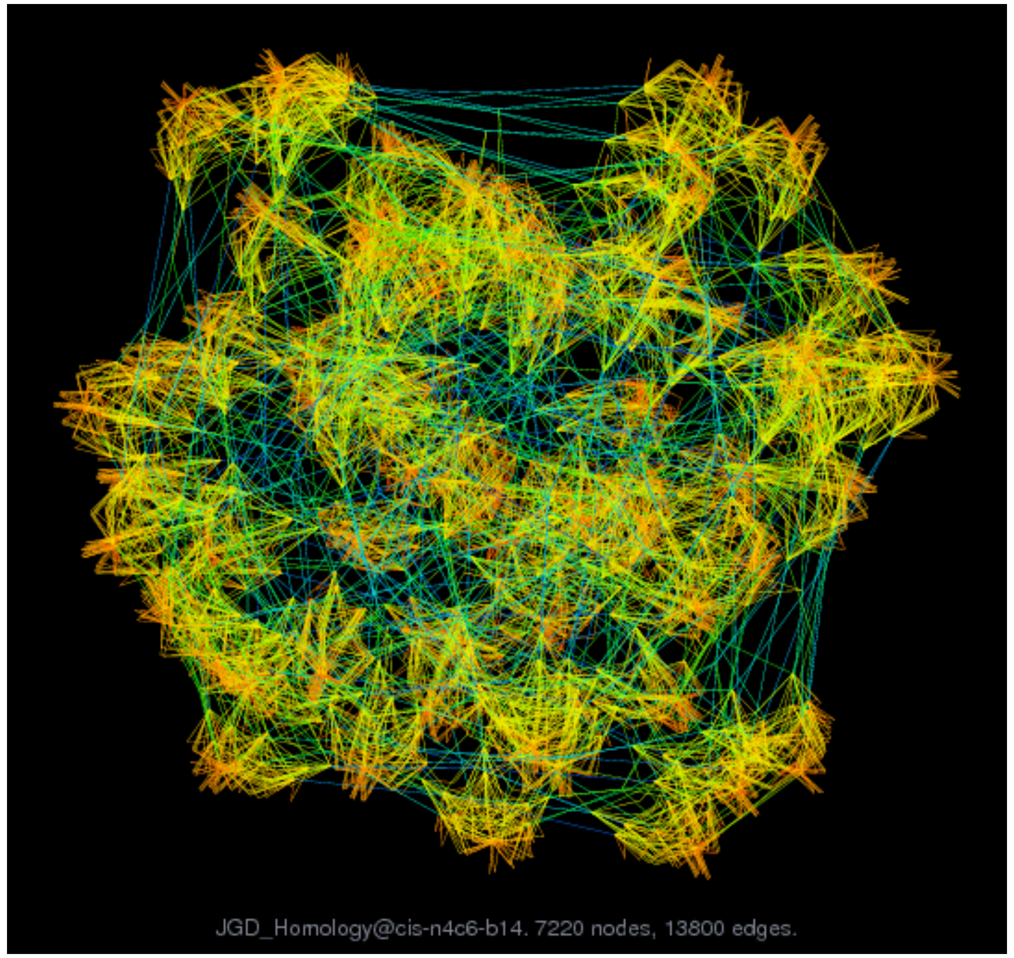
- visual encoding
  - –link connection marks, node point marks
- considerations
  - –spatial position: no meaning directly encoded
    - left free to minimize crossings
  - –proximity semantics?
    - sometimes meaningful
    - sometimes arbitrary, artifact of layout algorithm
    - tension with length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



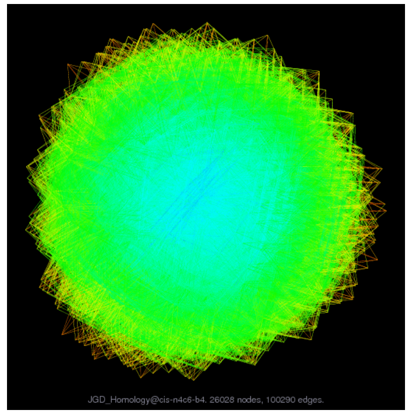
http://mbostock.github.com/d3/ex/force.html

Idiom: **sfdp** (multi-level force-directed placement)

- data
  - original: network
  - derived: cluster hierarchy atop it
- considerations
  - better algorithm for same encoding technique
    - same: fundamental use of space
    - hierarchy used for algorithm speed/quality but not shown explicitly
    - (more on algorithm vs encoding in afternoon)
- scalability
  - nodes, edges: 1K-10K
  - hairball problem eventually hits



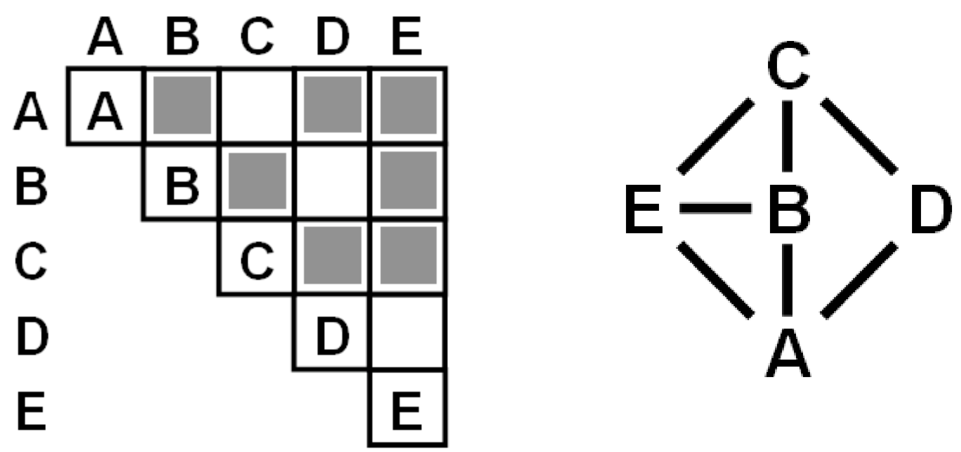
[Efficient and high quality force-directed graph drawing. Hu. The Mathematica Journal 10:37–71, 2005.]



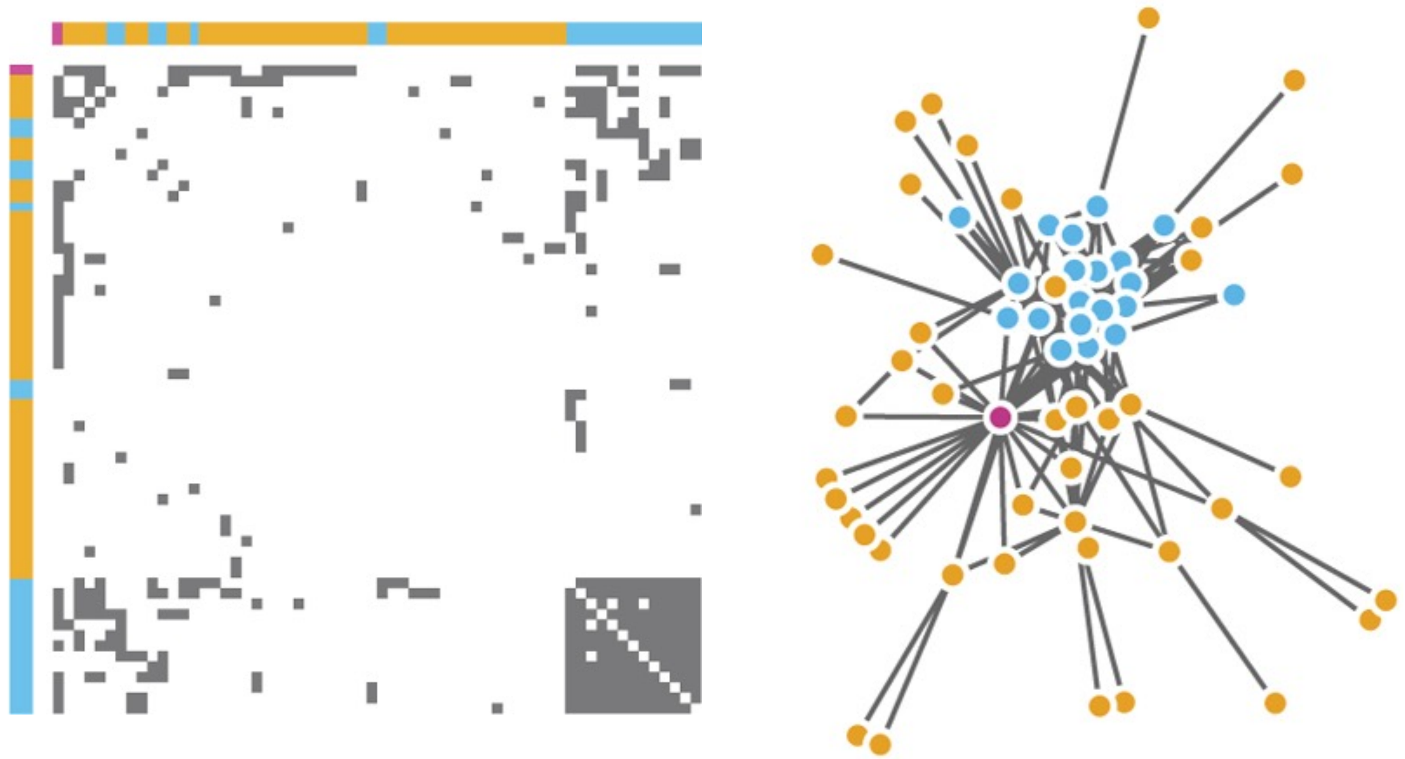
http://www.research.att.com/yifanhu/GALLERY/GRAPHS/index1.html

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



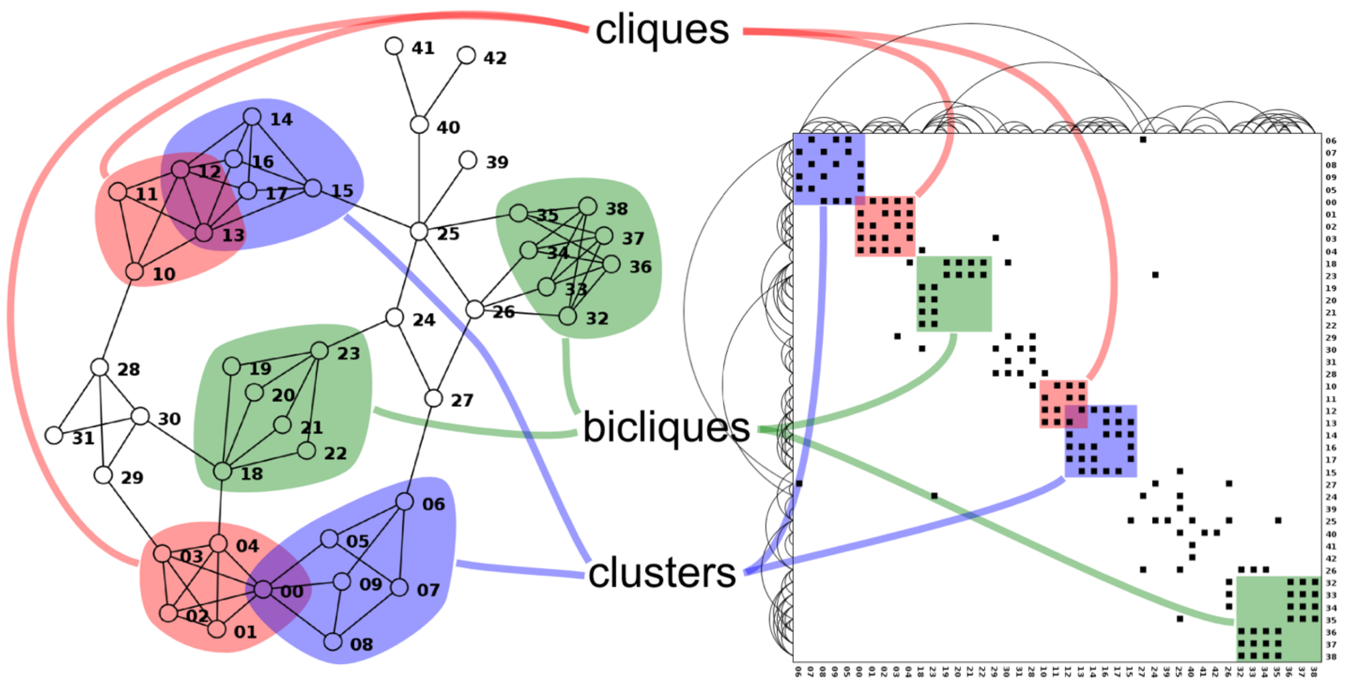
[NodeTrix: a Hybrid Visualization of Social Networks. Henry, Fekete, and McGuffin. IEEE TVCG (Proc. InfoVis) 13(6):1302-1309, 2007.]



[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!

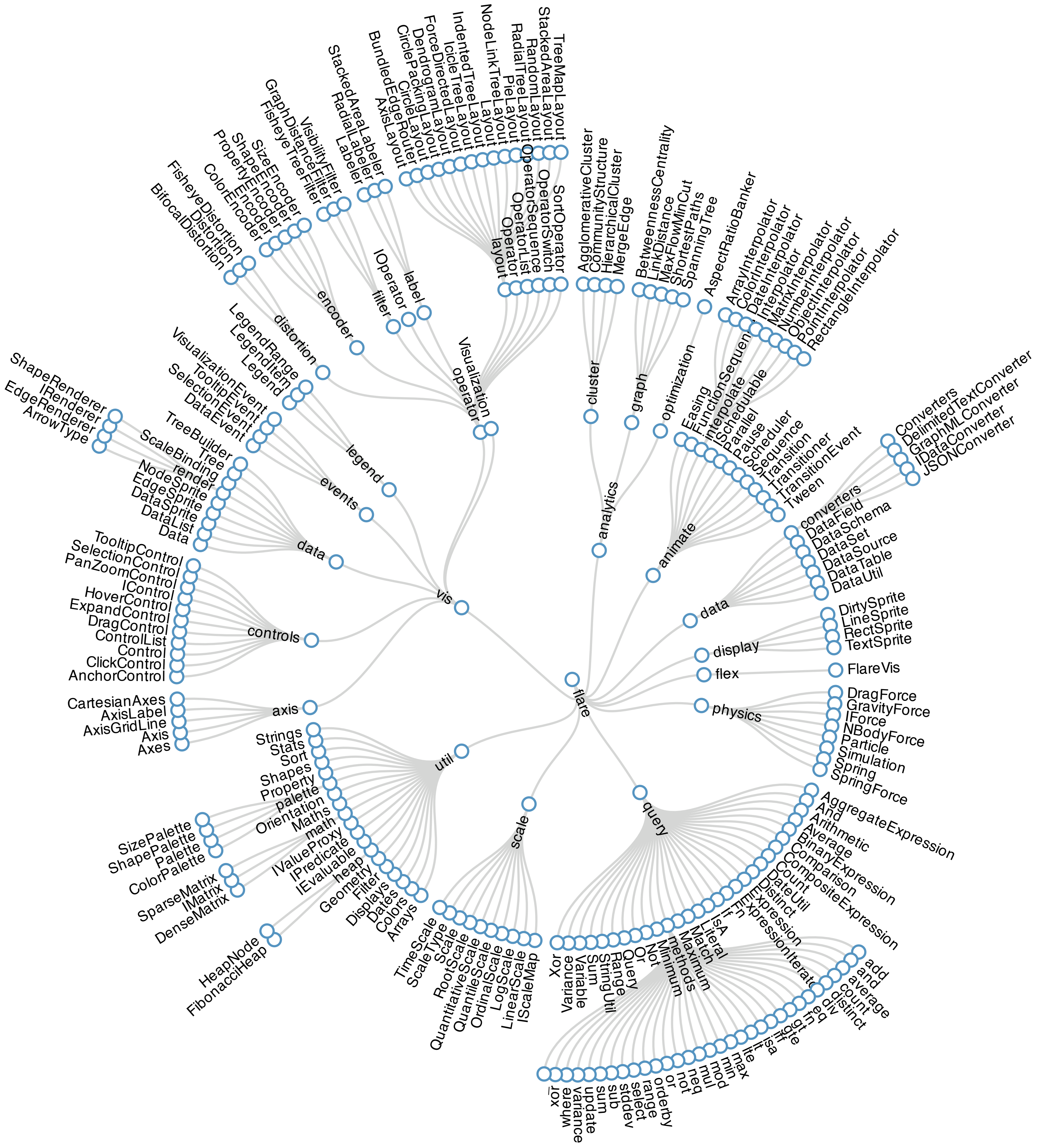


http://www.michaelmcguffin.com/courses/vis/patternsInAdjacencyMatrix.png

[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**

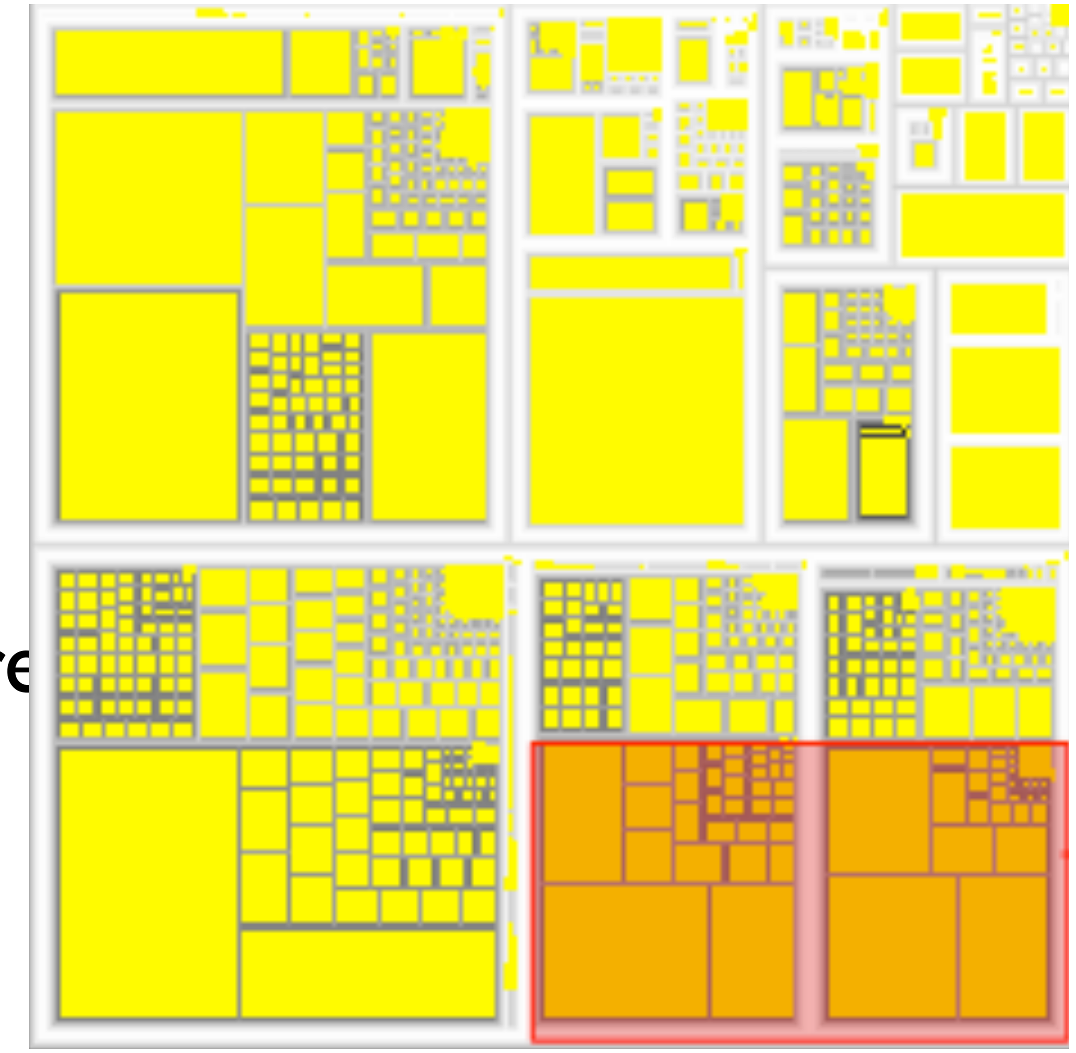
- data
  - tree
- encoding
  - link connection marks
  - point node marks
  - radial axis orientation
    - angular proximity: siblings
    - distance from center: depth in tree
- tasks
  - understanding topology, following paths
- scalability
  - 1K - 10K nodes



http://mbostock.github.com/d3/ex/tree.html

Idiom: **treemap**

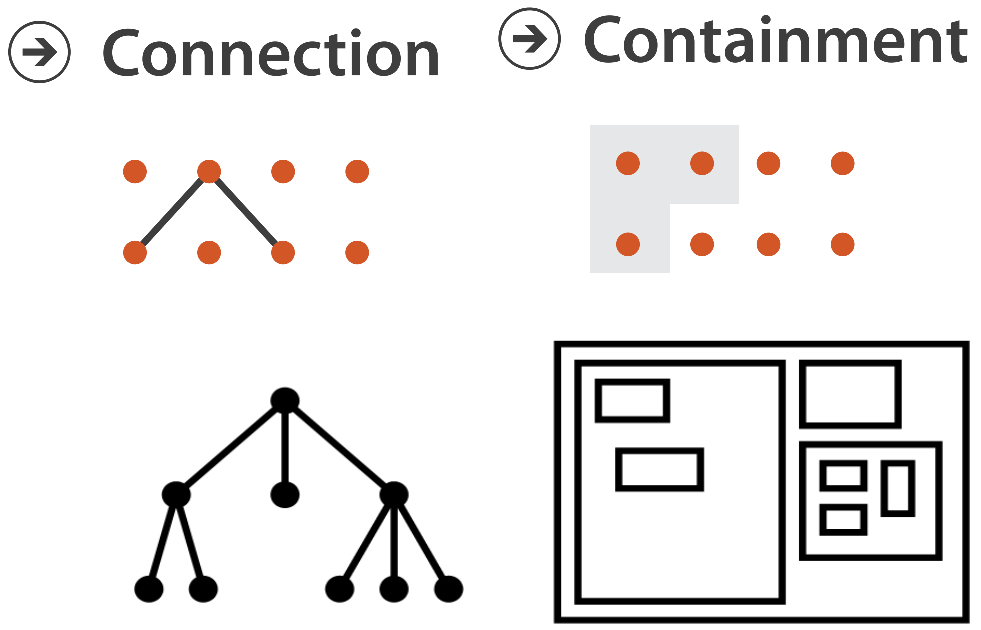
- data
  - tree
  - 1 quant attrib at leaf nodes
- encoding
  - area containment marks for hierarchical structure
  - rectilinear orientation
  - size encodes quant attrib
- tasks
  - query attribute at leaf nodes
- scalability
  - 1M leaf nodes

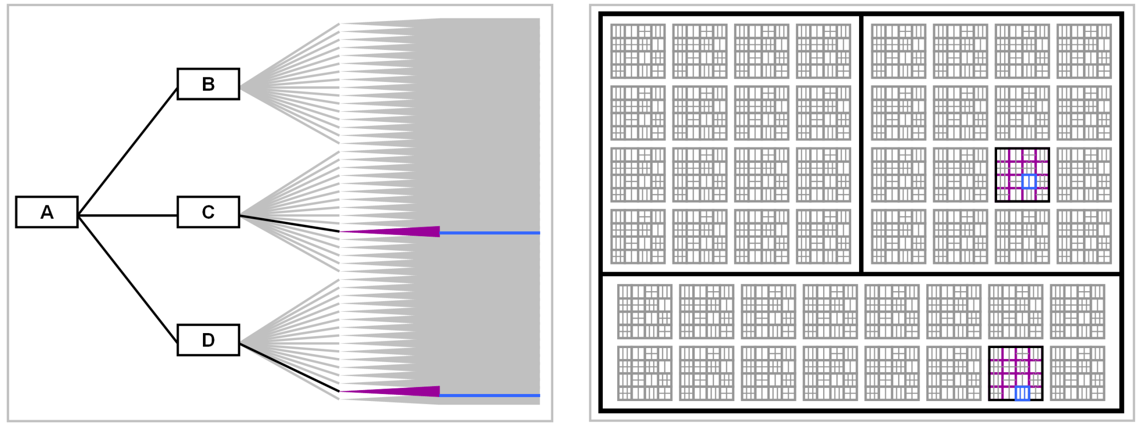


http://tulip.labri.fr/Documentation/3_7/userHandbook/html/ch06.html

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



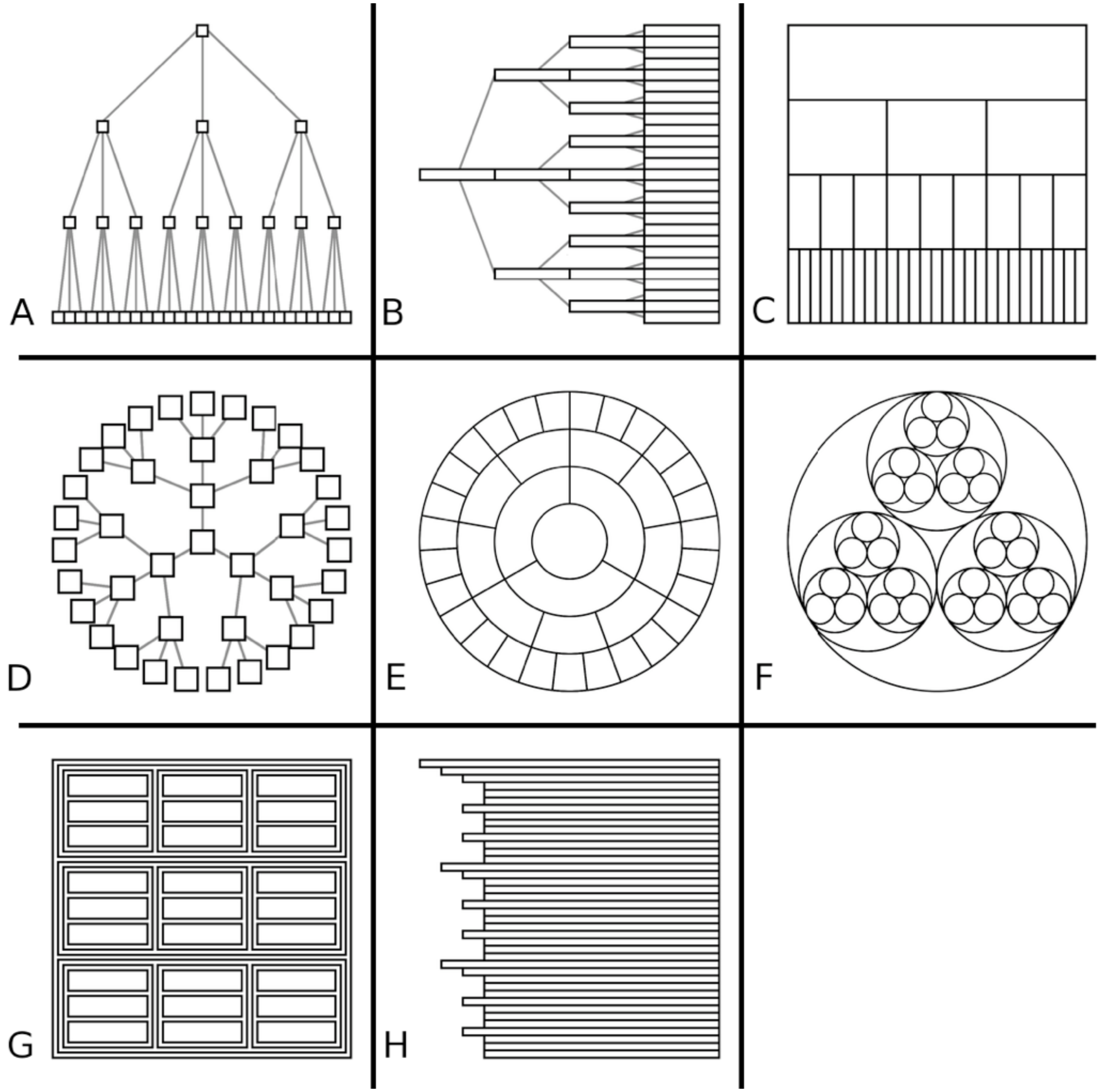


Node–Link Diagram **Treemap**

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

Tree drawing idioms comparison

- data shown
  - link relationships
  - tree depth
  - sibling order
- design choices
  - connection vs containment link marks
  - rectilinear vs radial layout
  - spatial position channels
- considerations
  - redundant? arbitrary?
  - information density?
    - avoid wasting space



[Quantifying the Space-Efficiency of 2D Graphical Representations of Trees. McGuffin and Robert. Information Visualization 9:2 (2010), 115–140.]

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.

Further reading, full

- Graph Visualisation in Information Visualisation: A Survey. Ivan Herman, Guy Melançon, and M. Scott Marshall. IEEE Transactions on Visualization and Computer Graphics (TVCG) 6:1 (2000), 24-44.
- 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.
- Just How Dense Are Dense Graphs in the Real World?: A Methodological Note. Guy Melançon. Proc AVI Workshop BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV). ACM, 2006.
- Drawing on Physical Analogies. Ulrik 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.
- A Fast Adaptive Layout Algorithm for Undirected Graphs. A. Frick, A. Ludwig, and H. Mehldau. Proc. International Symposium on Graph Drawing (GD 94), Lecture Notes in Computer Science, 894, pp. 388-403. Springer, 1995.
- Efficient and High Quality Force-Directed Graph Drawing. Yifan Hu. The Mathematica Journal 10 (2005), 37-71.
- Drawing Large Graphs with a Potential-Field-Based Multilevel Algorithm. S. Hachul and M. Jünger. Proc International Symposium on Graph Drawing (GD 04), Lecture Notes in Computer Science, 3383, pp. 285-295. Springer, 2004.
- TopoLayout: Multilevel Graph Layout by Topological Features. Daniel Archambault, Tamara Munzner, and David Auber. IEEE Transactions on Visualization and Computer Graphics 13:2 (2007), 305-317.
- MatrixExplorer: A Dual-Representation System to Explore Social Networks. Nathalie Henry and Jean-Daniel Fekete. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 06) 12:5 (2006), 677-684.
- NodeTrix: A Hybrid Visualization of Social Networks. Nathalie Henry, Jean-Daniel Fekete, and Michael McGuffin. IEEE Transactions on Computer Graphics and Visualization (Proc. InfoVis 07) 13:6 (2007), 1302-1309.
- On the Readability of Graphs Using Node-Link and Matrix-Based Representations: A Controlled Experiment and Statistical Analysis. Mohammad Ghoniem, Jean-Daniel Fekete, and Philippe Castagliola. Information Visualization 4:2 (2005), 114-135.
- Quantifying the Space-Efficiency of 2D Graphical Representations of Trees. Michael J. McGuffin and Jean-Marc Robert. Information Visualization 9:2 (2010), 115-140.
- web site: treevis.net. article: A Tree Visualization Reference. Hans-Jö rg Schulz. IEEE Computer Graphics and Applications 31:6 (2011), 11-15.
- The Design Space of Implicit Hierarchy Visualization: A Survey. Hans-J&oouml;rg Schulz, Steffen Hadlak, and Heidrun Schumann. IEEE Transactions on Visualization and Computer Graphics 17:4(2011), 393-411.
- Treemaps: A Space-Filling Approach to the Visualization of Hierarchical Information. Brian Johnson and Ben Shneiderman. Proc IEEE Conference on Visualization (Vis) 1991, pp. 284-291.
- Perceptual Guidelines for Creating Rectangular Treemaps. Kong, Heer, and Agrawala. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis) 16:6 (2010), 990-998.

Ch 10. Map Color and Other Channels

Color

- Color Encoding
  - Hue
  - Saturation
  - Luminance
- Color Map
  - Categorical
  - Ordered
    - *Sequential*
    - *Diverging*
  - Bivariate

Size, Angle, Curvature, ...

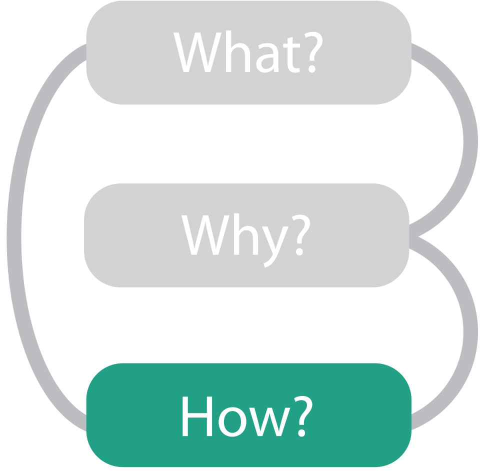
- Length
- Angle
- Area
- Curvature
- Volume

Shape

Motion

- Motion
  Direction, Rate, Frequency, ...

Idiom design choices: Encode



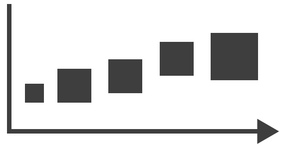
Encode

Arrange

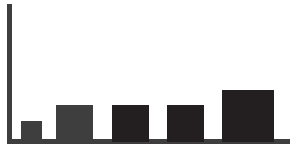
→ Express

→ Separate

→ Order



→ Align



→ Use

Map
from categorical and ordered attributes

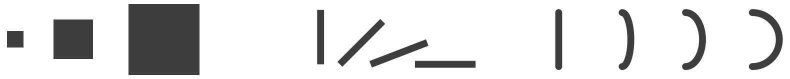
→ Color

→ *Hue*

→ *Saturation*

→ *Luminance*

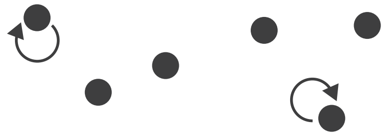
→ Size, Angle, Curvature, ...



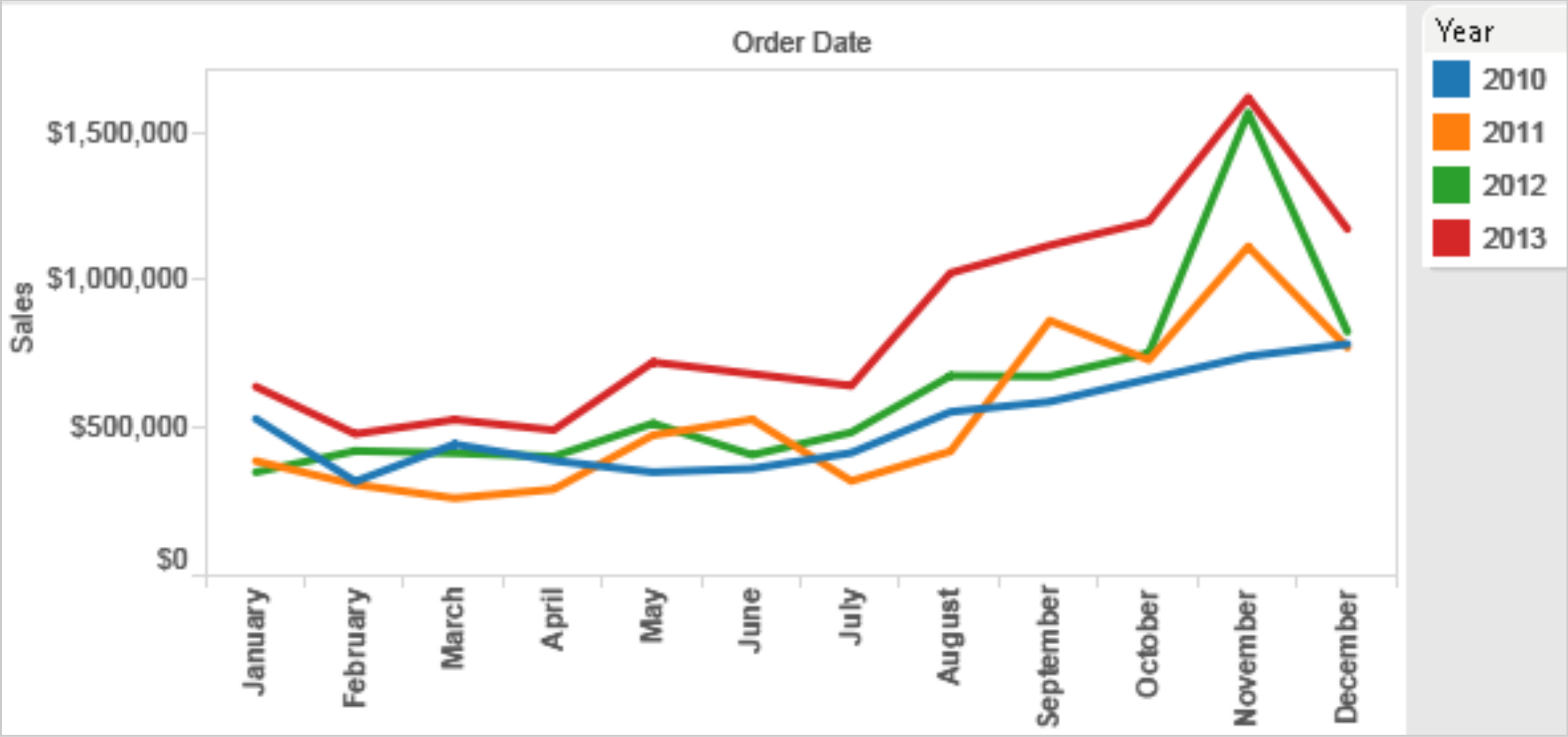
→ Shape

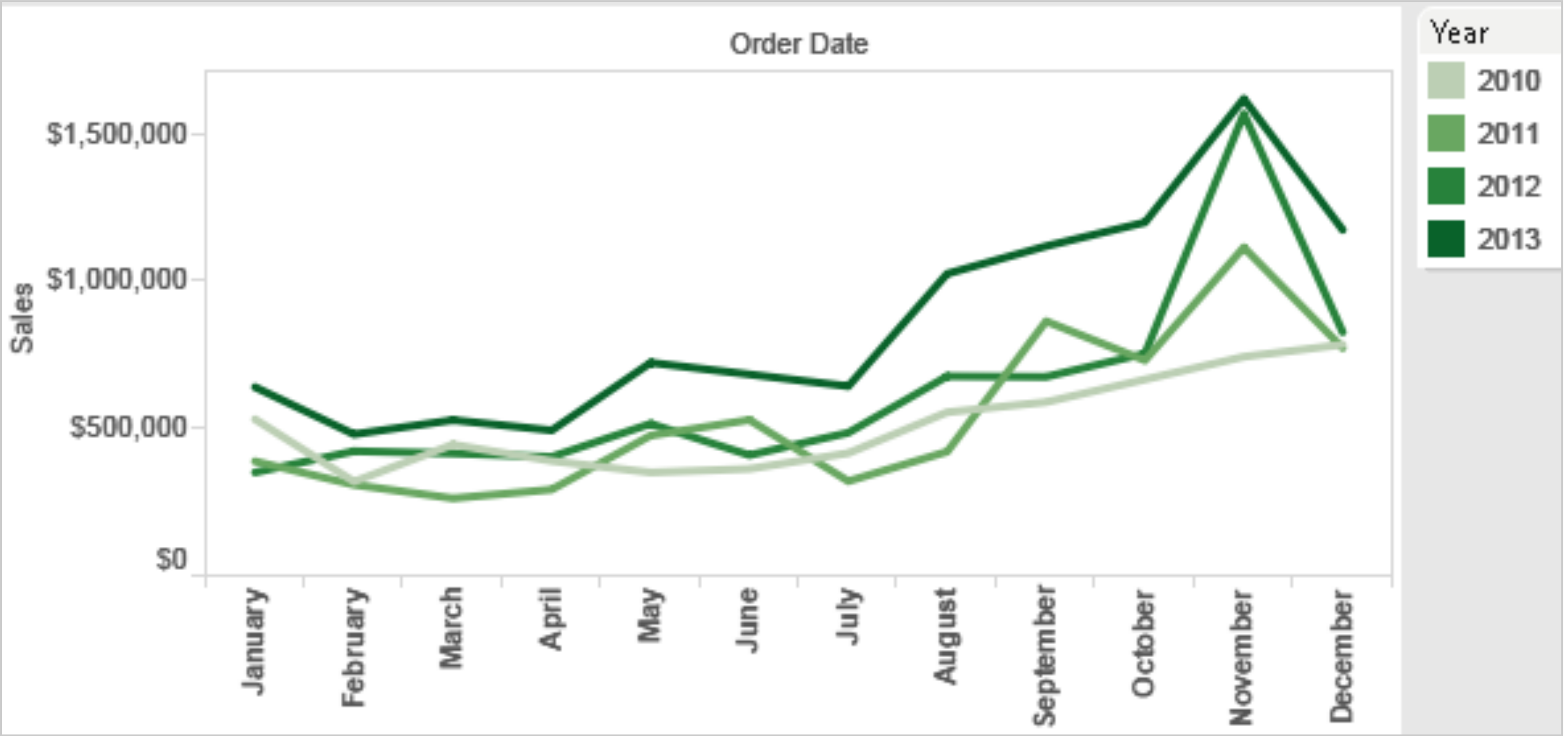
→ Motion

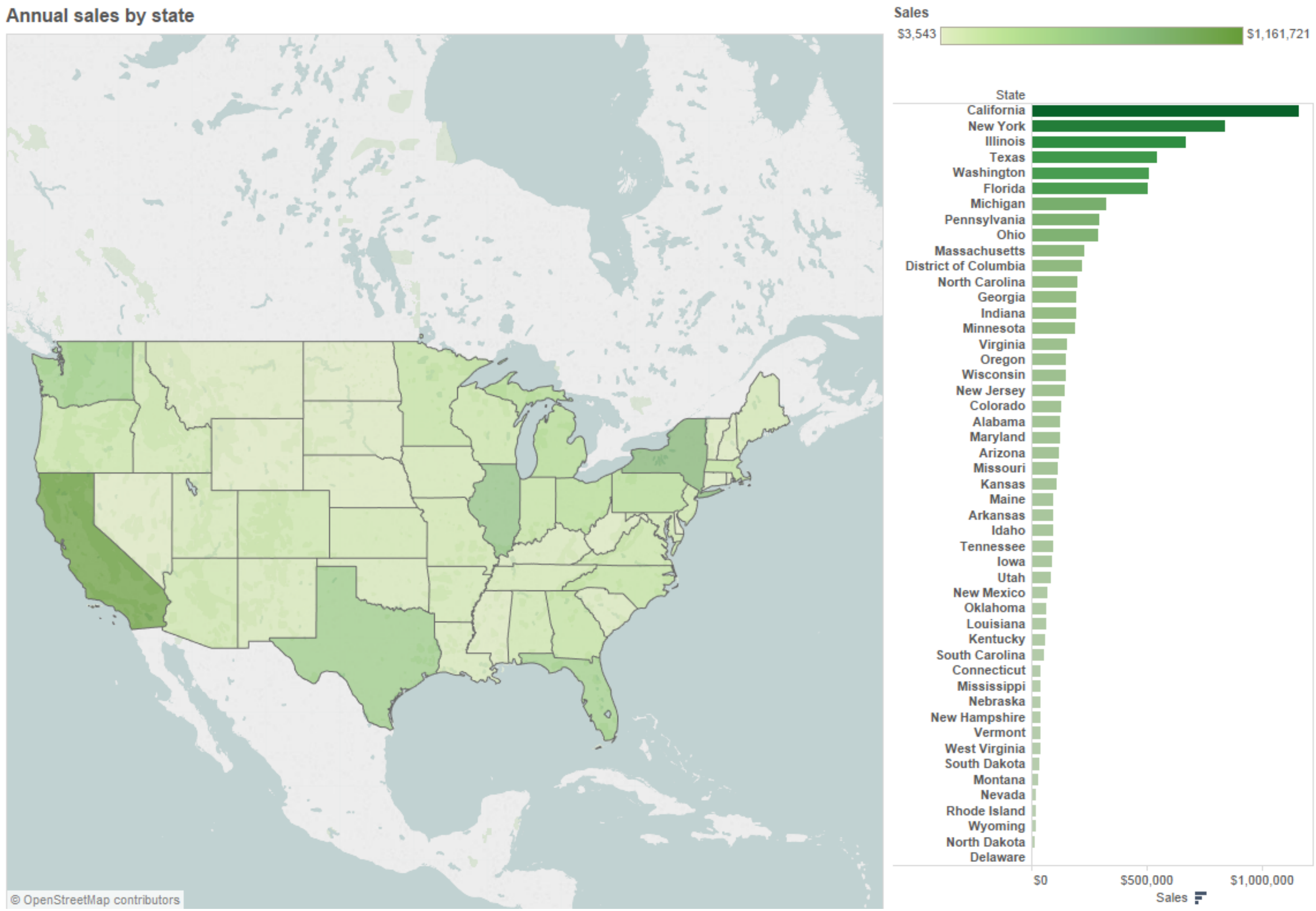
Direction, Rate, Frequency, ...



Categorical vs ordered color





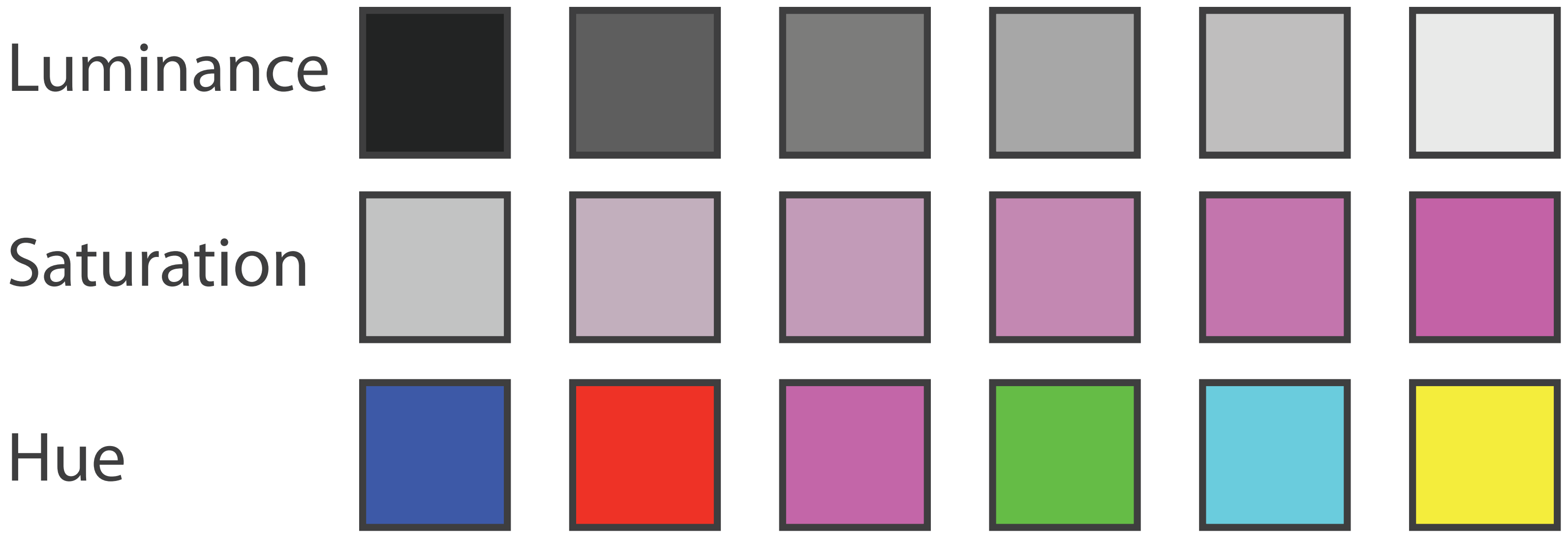


[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

Decomposing color

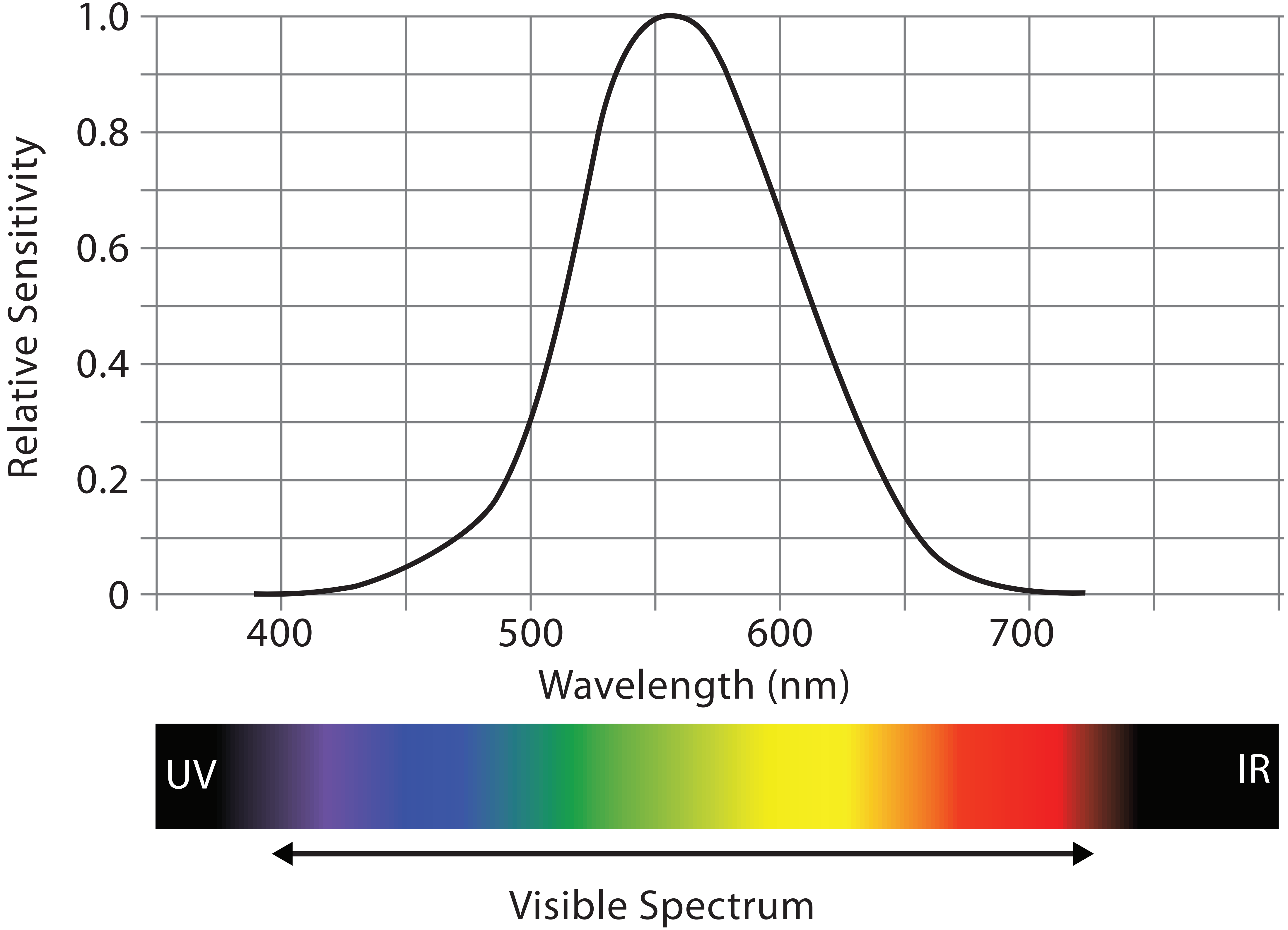
- first rule of color: do not talk about color!
  - color is confusing if treated as monolithic

- decompose into three channels
  - ordered can show magnitude
    - luminance: how bright
    - saturation: how colorful
  - categorical can show identity
    - hue: what color



- channels have different properties
  - what they convey directly to perceptual system
  - how much they can convey: how many discriminable bins can we use?

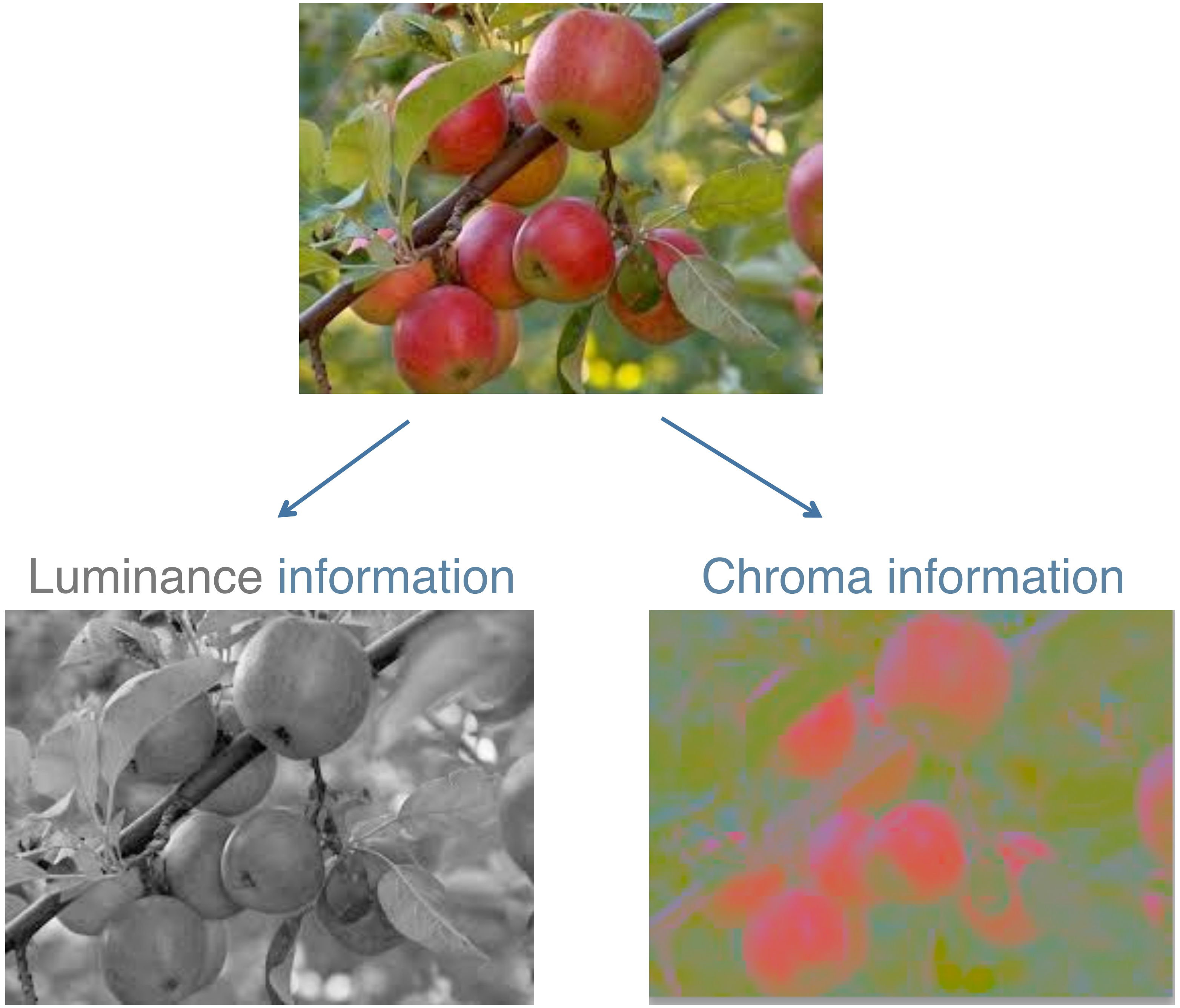
Spectral sensitivity



Luminance

- need luminance for edge detection
  - fine-grained detail only visible through luminance contrast
  - legible text requires luminance contrast!

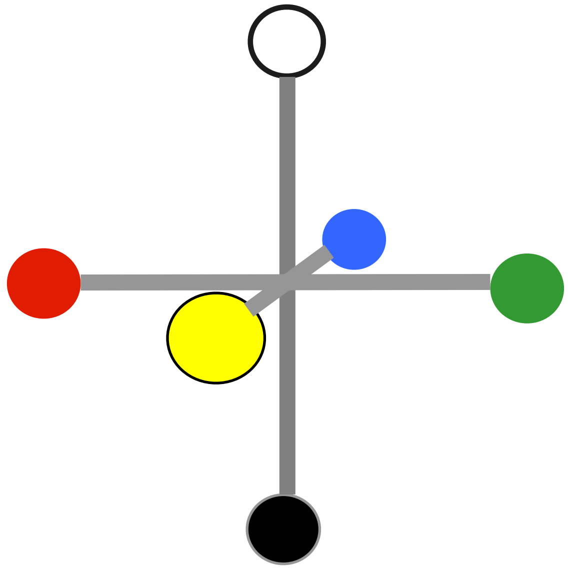
- intrinsic perceptual ordering

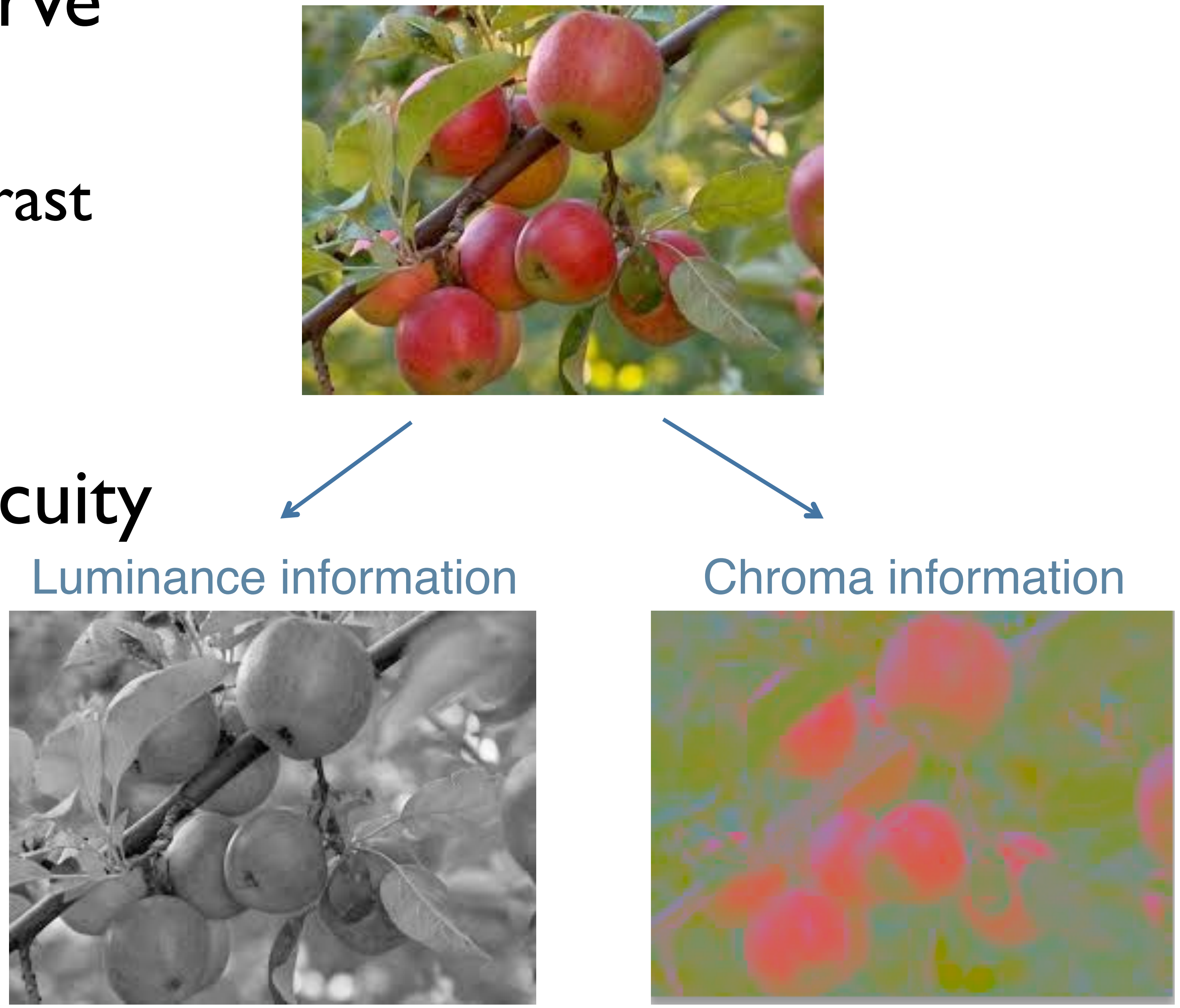


[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

Opponent color and color deficiency

- perceptual processing before optic nerve
  - one achromatic luminance channel (L*)
    - edge detection through luminance contrast
  - 2 chroma channels
    - red-green (a*) & yellow-blue axis (b*)
- "color blind": one axis has degraded acuity
  - 8% of men are red/green color deficient
  - blue/yellow is rare





[Seriously Colorful: Advanced Color Principles & Practices. Stone. Tableau Customer Conference 2014.]

Color spaces

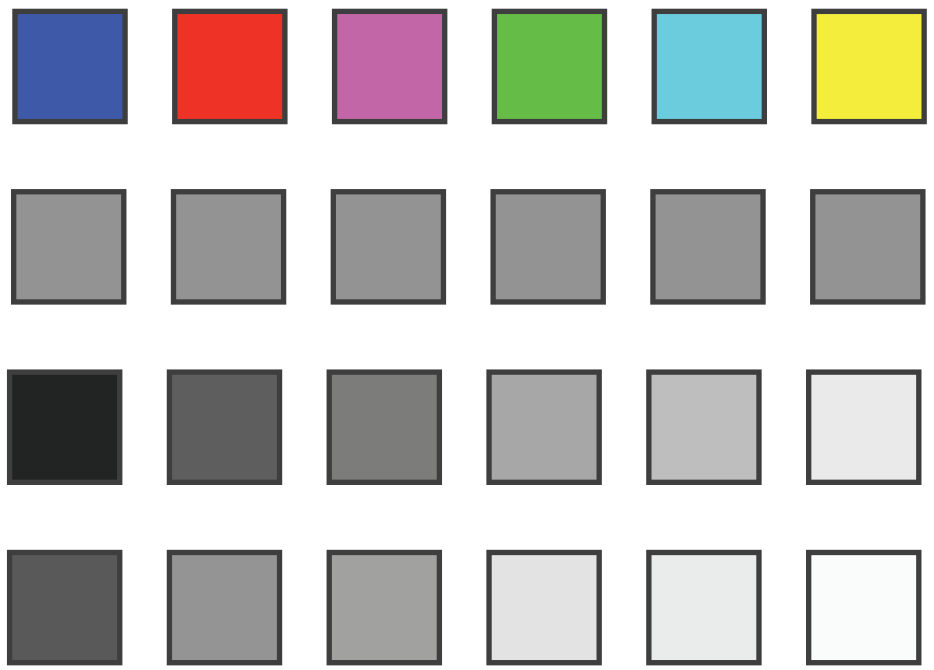
- CIE L*a*b*: good for computation
  – L* intuitive: perceptually linear luminance
  – a*b* axes: perceptually linear but nonintuitive
- RGB: good for display hardware
  – poor for encoding
- HSL/HSV: somewhat better for encoding
  – hue/saturation wheel intuitive
  – beware: only pseudo-perceptual!
  – lightness (L) or value (V) ≠ luminance or L*
- Luminance, hue, saturation
  – good for encoding
  – but not standard graphics/tools colorspace

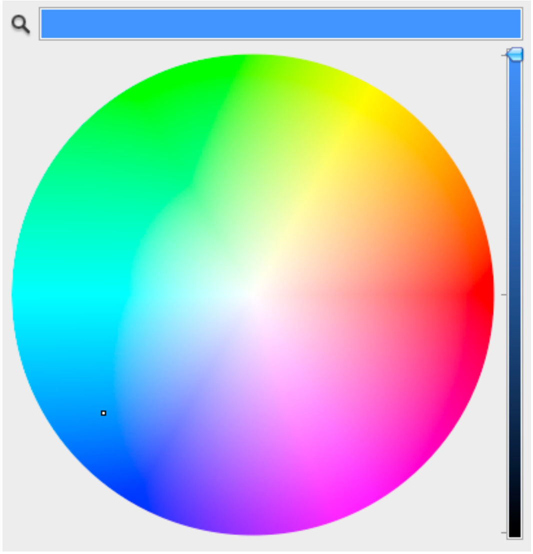
Corners of the RGB color cube

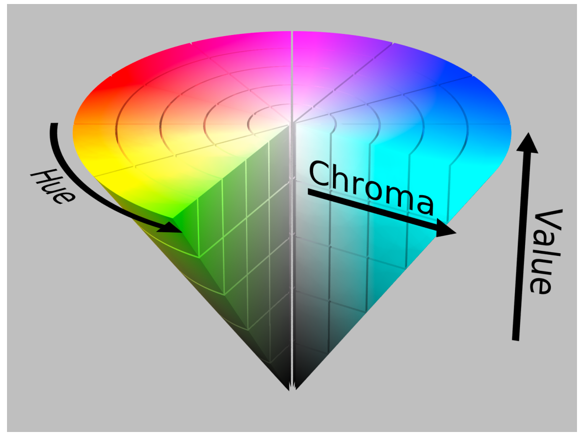
L from HLS
All the same

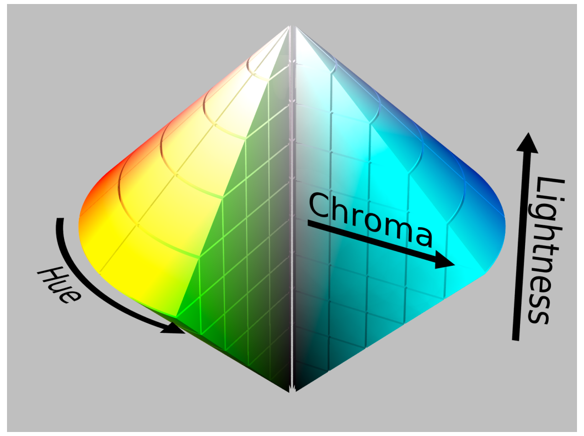
Luminance values

L* values



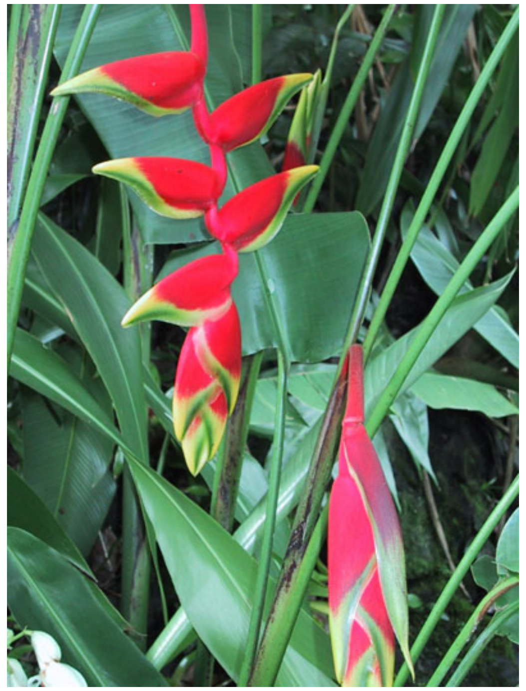




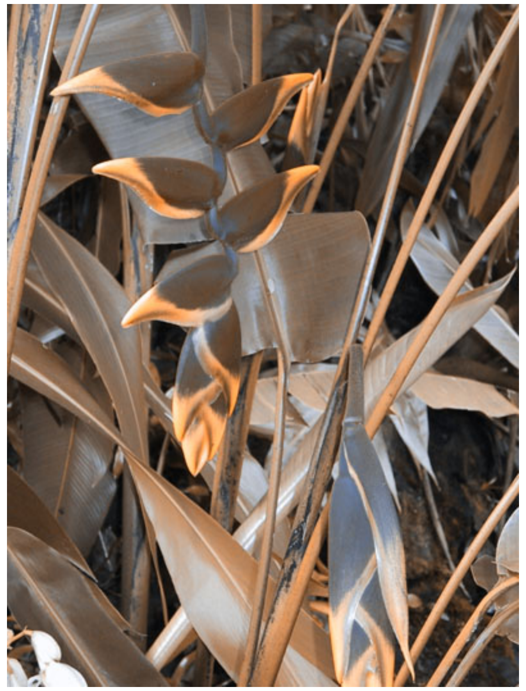


[https://en.wikipedia.org/wiki/HSL_and_HSV]

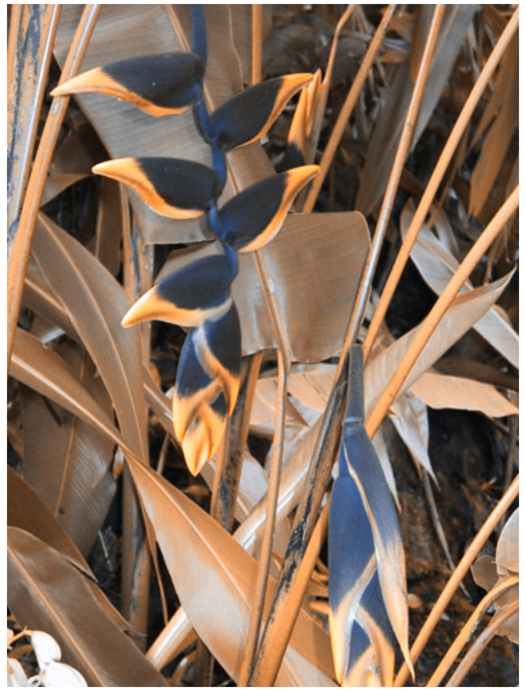
Designing for color deficiency: Check with simulator



Normal vision



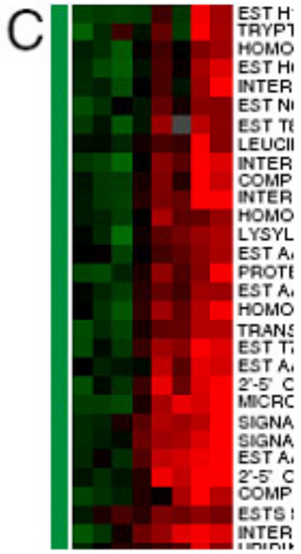
Deuteranope



Protanope

Tritanope

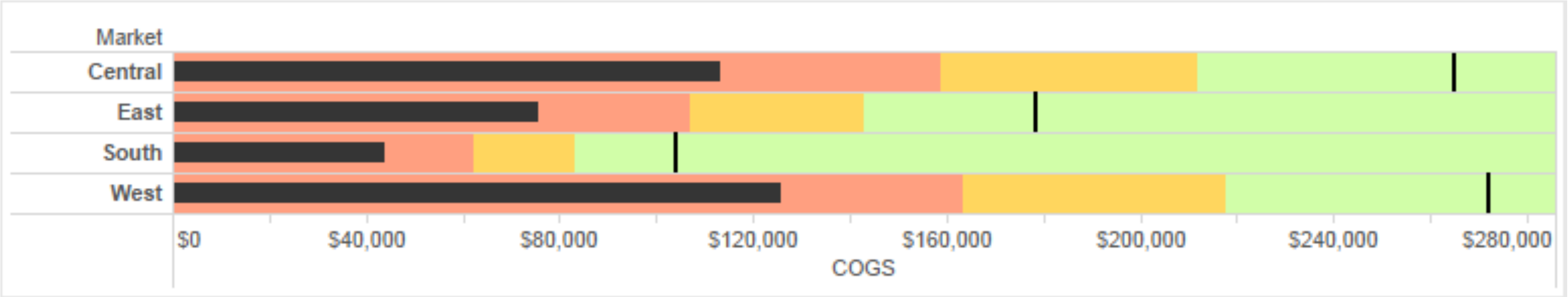
http://rehue.net

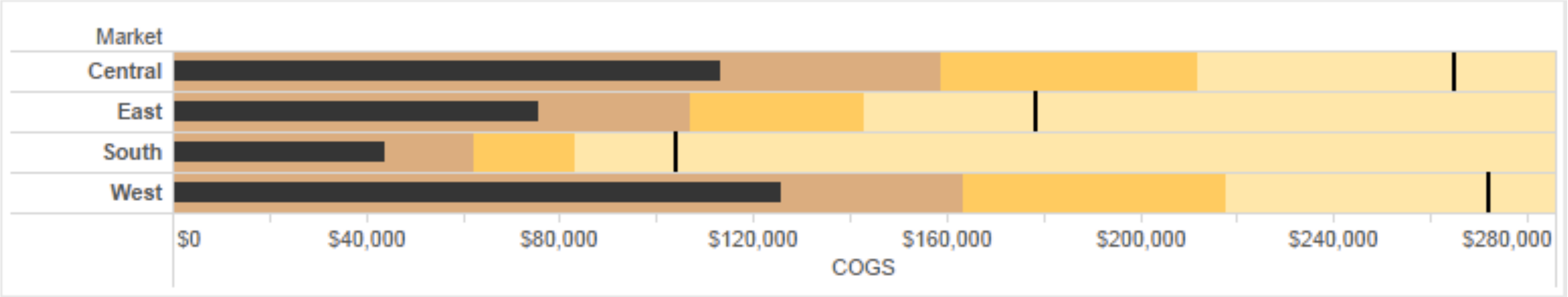


[Seriously Colorful: Advanced Color Principles & Practices. Stone. Tableau Customer Conference 2014.]

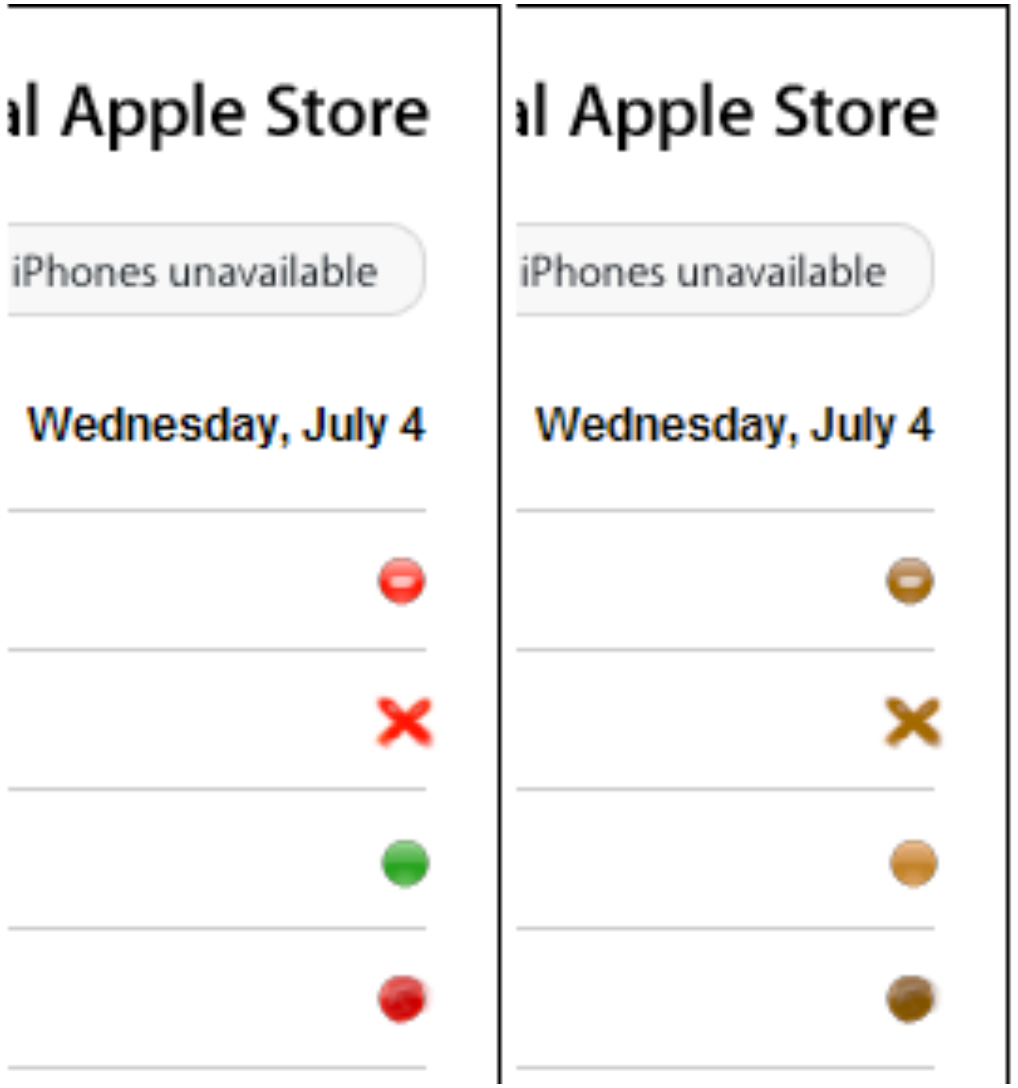
Designing for color deficiency: Avoid encoding by hue alone

- redundantly encode
  - vary luminance
  - change shape





Deuteranope simulation

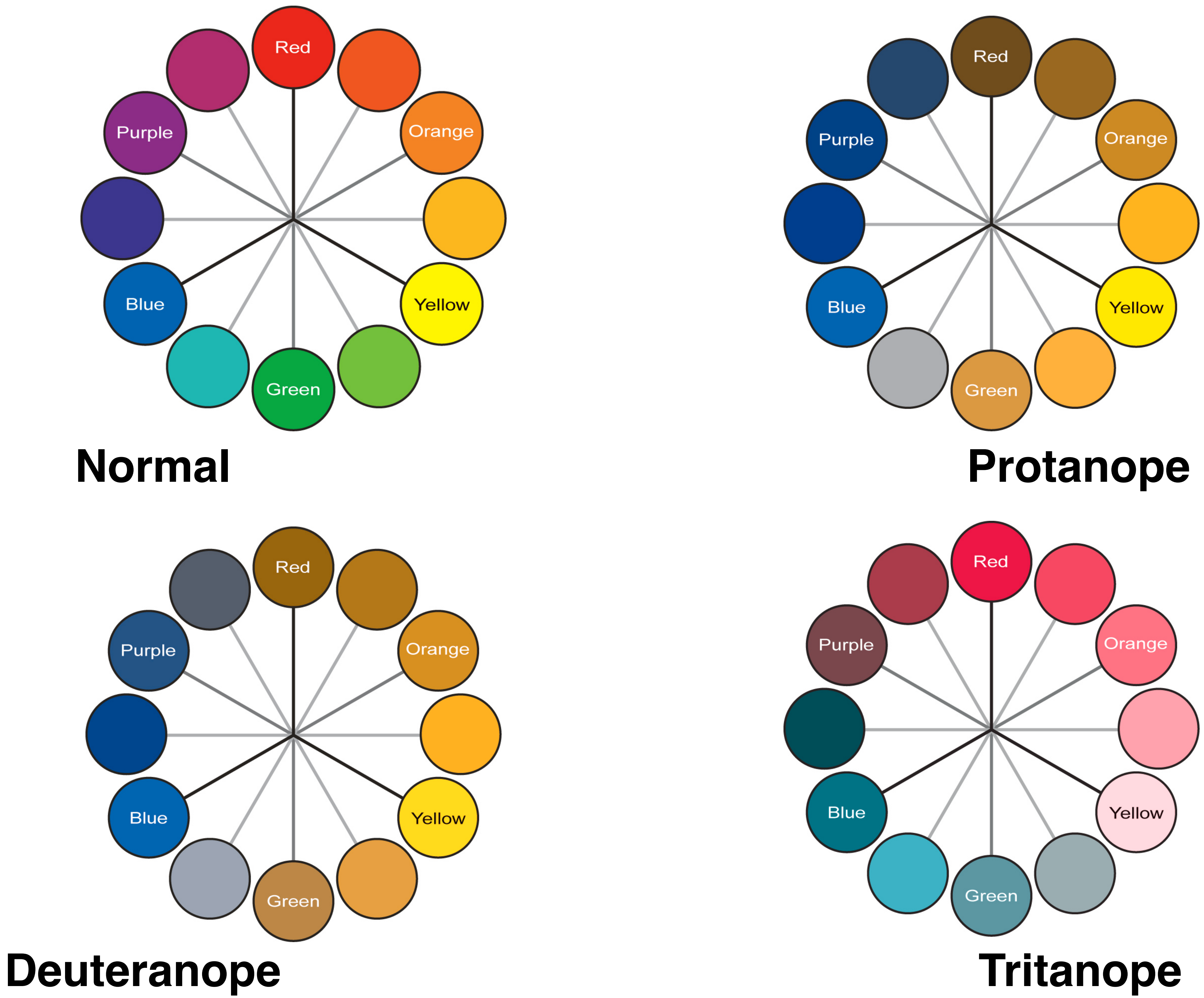


Change the shape

Vary luminance

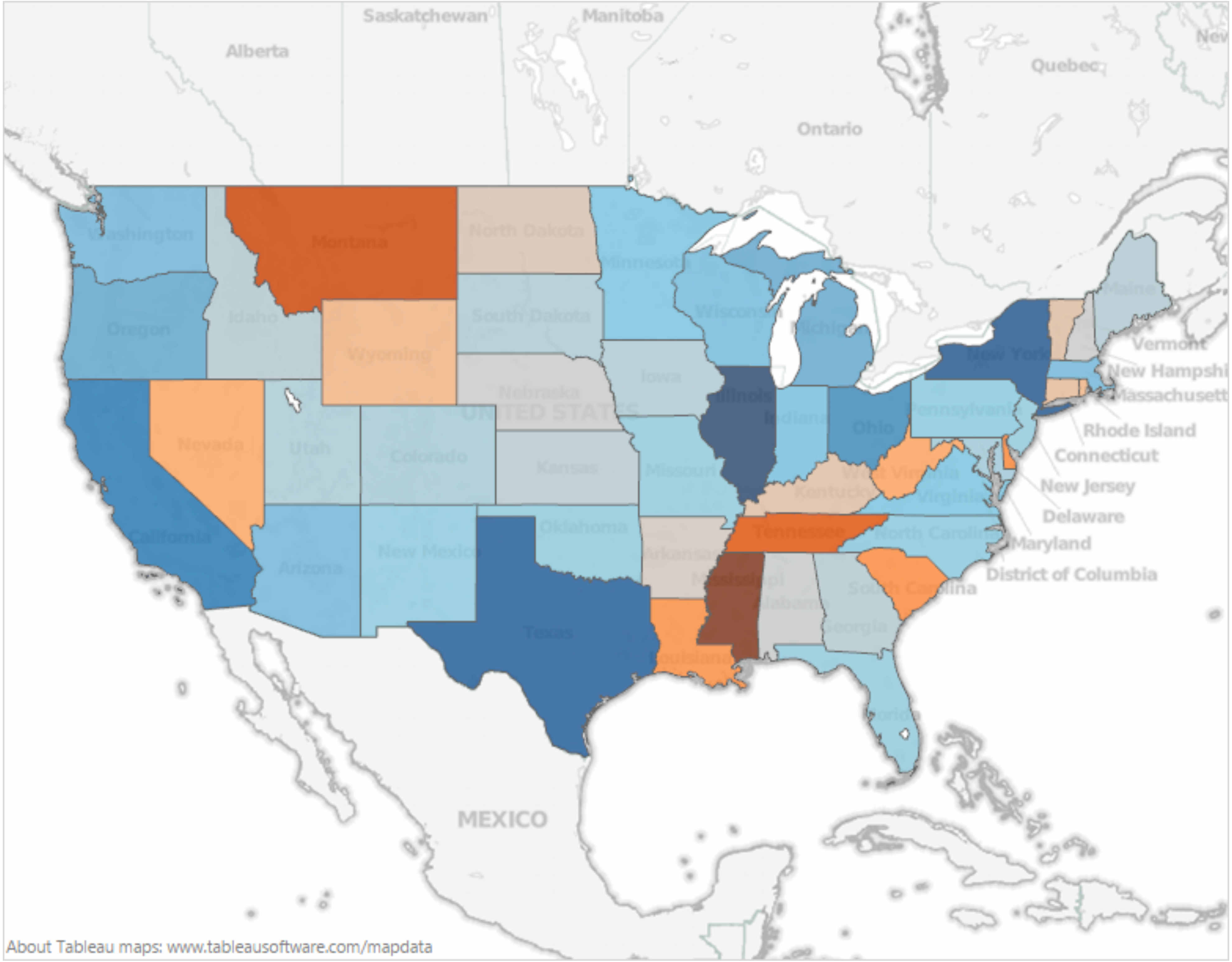
[Seriously Colorful: Advanced Color Principles & Practices. Stone. Tableau Customer Conference 2014.]

Color deficiency: Reduces color to 2 dimensions



[Seriously Colorful: Advanced Color Principles & Practices. Stone. Tableau Customer Conference 2014.]

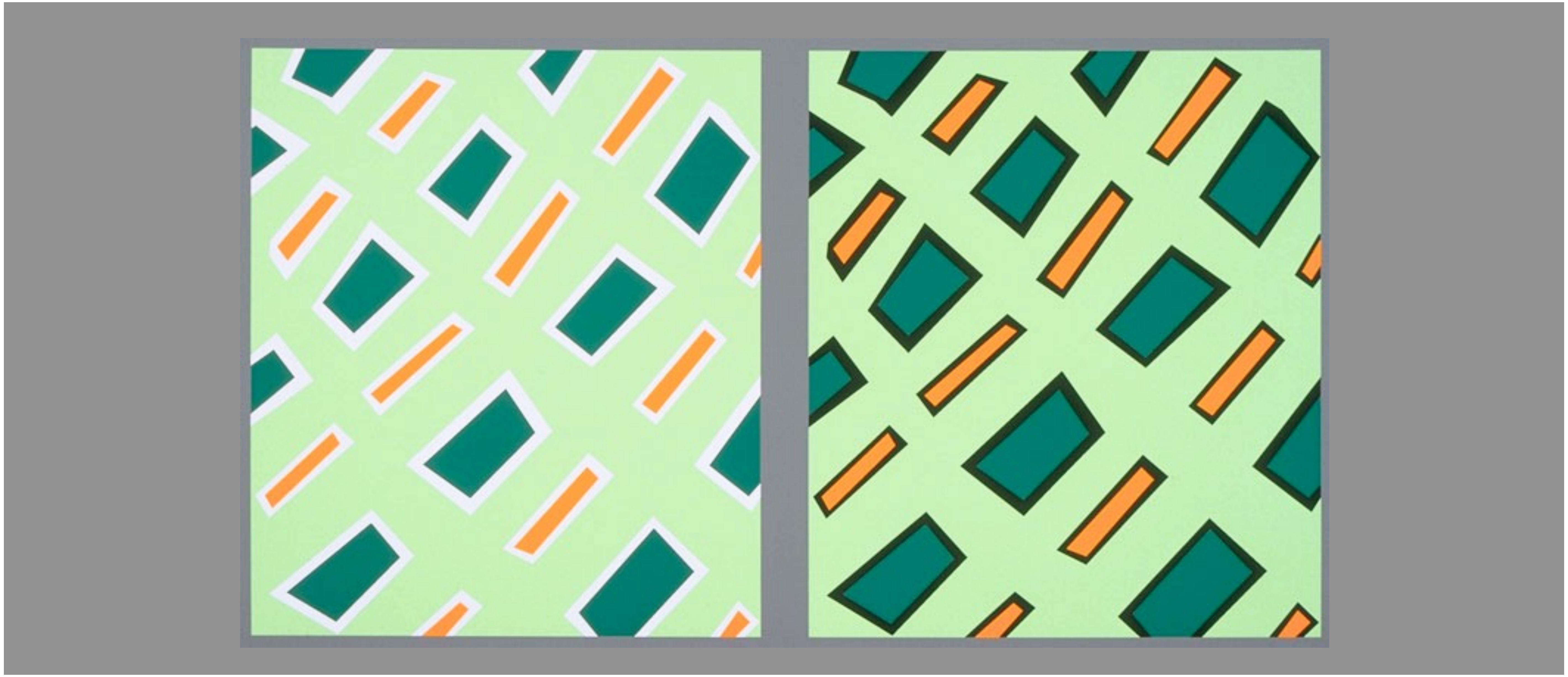
Designing for color deficiency: Blue-Orange is safe



[Seriously Colorful: Advanced Color Principles & Practices. Stone. Tableau Customer Conference 2014.]

Bezold Effect: Outlines matter

- color constancy: simultaneous contrast effect



[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

Relative judgements: Color & illumination

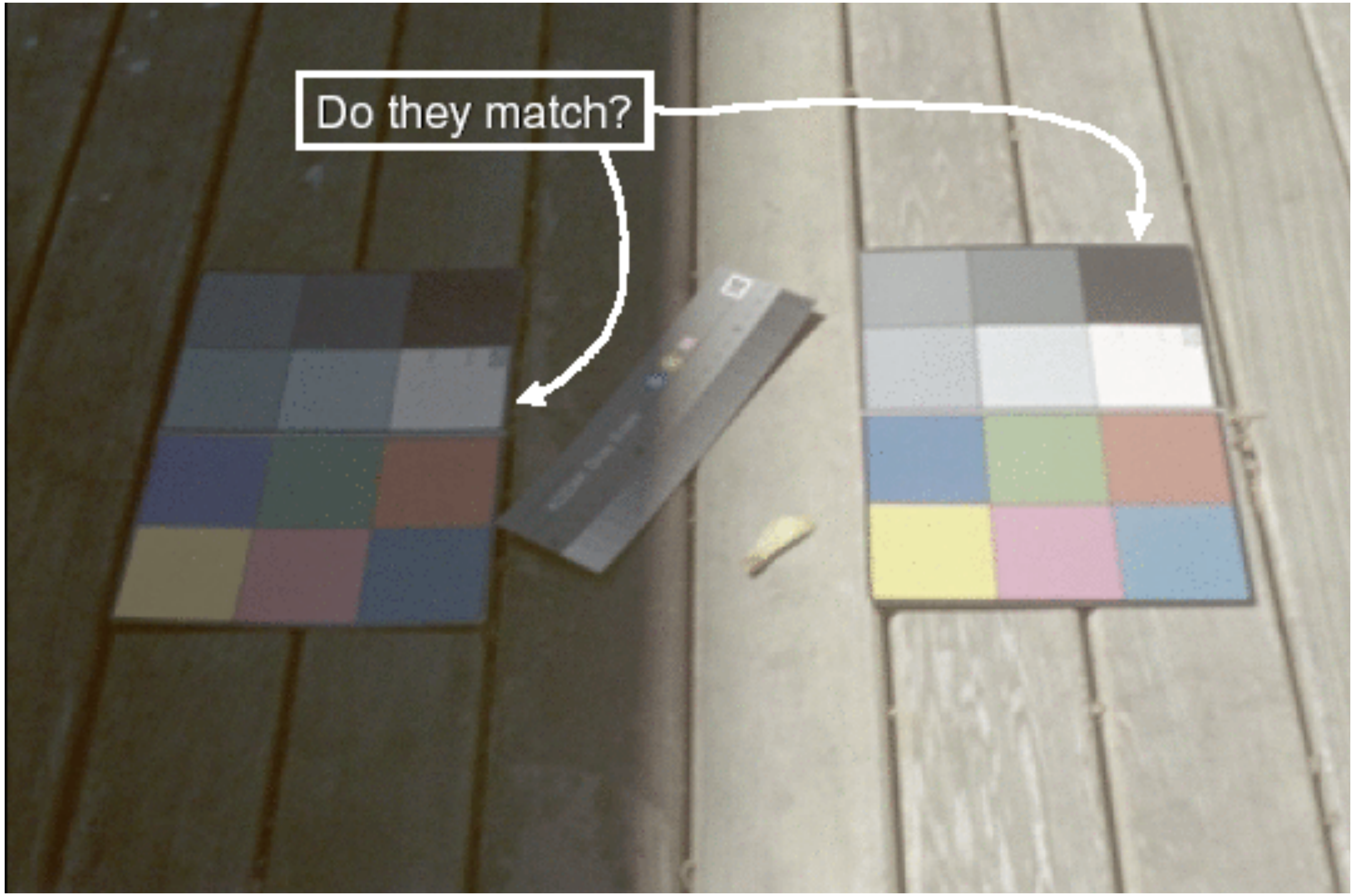


Image courtesy of John McCann

Relative judgements: Color & illumination

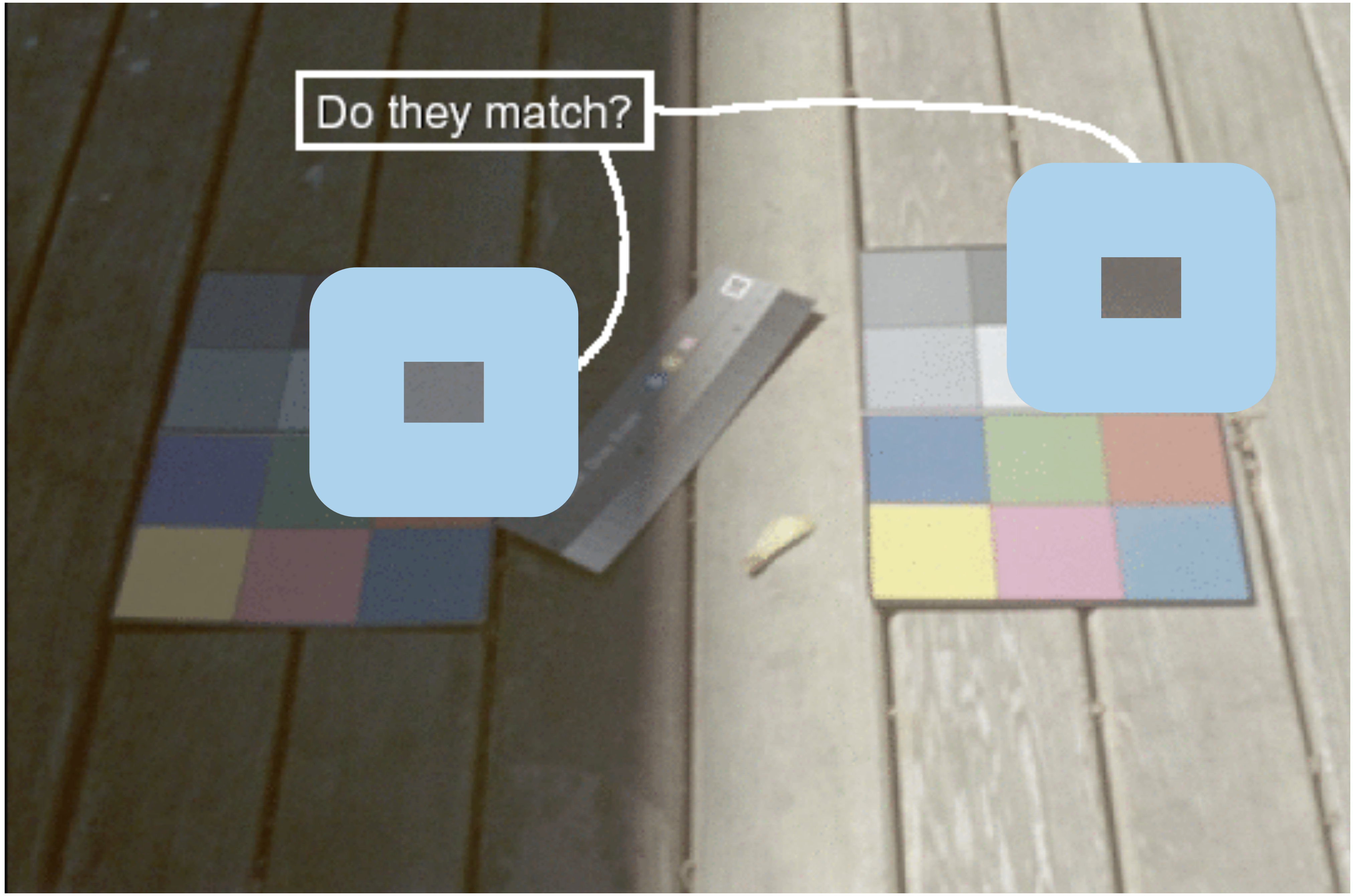
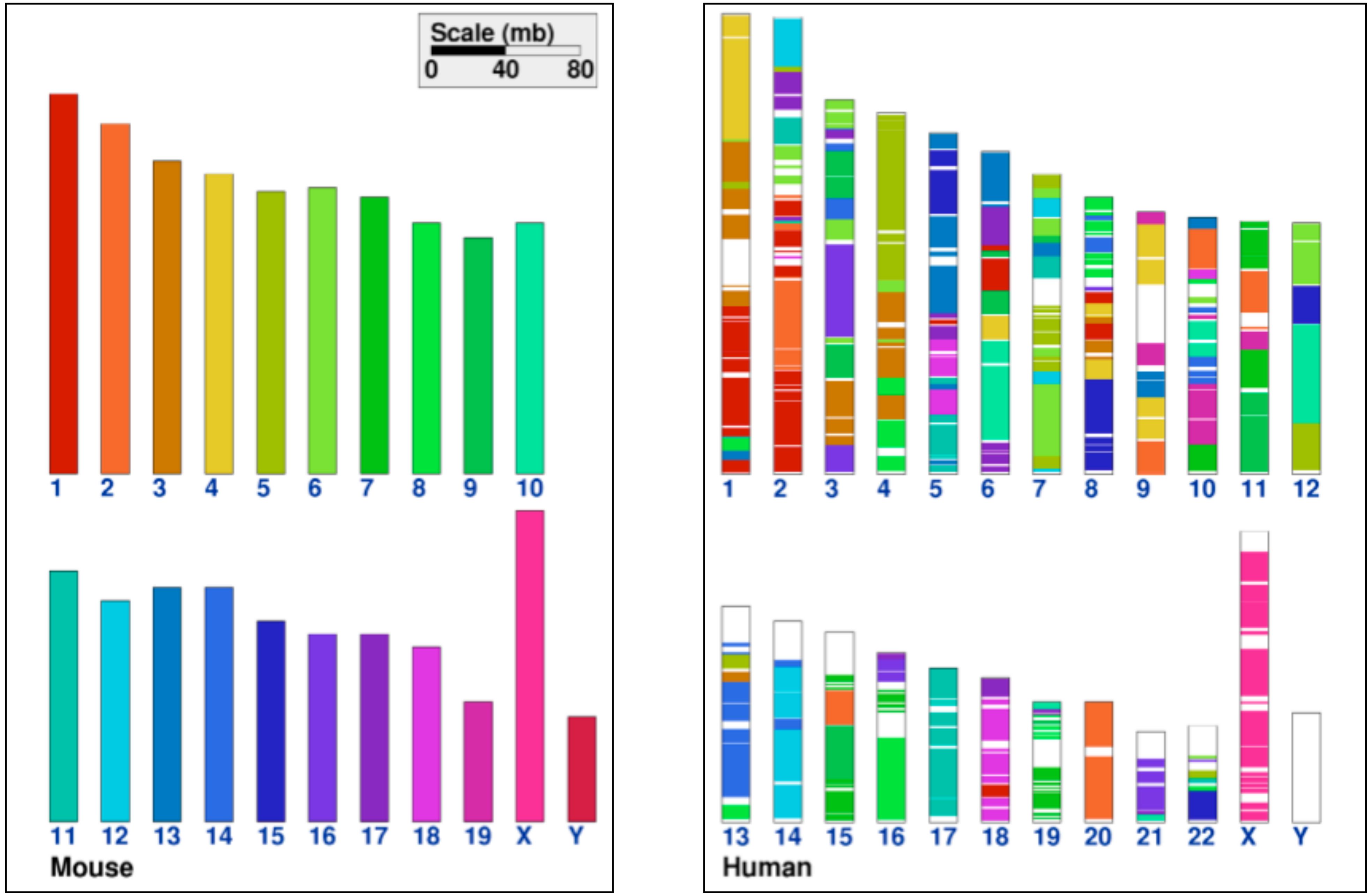


Image courtesy of John McCann via Maureen Stone

Categorical color: limited number of discriminable bins

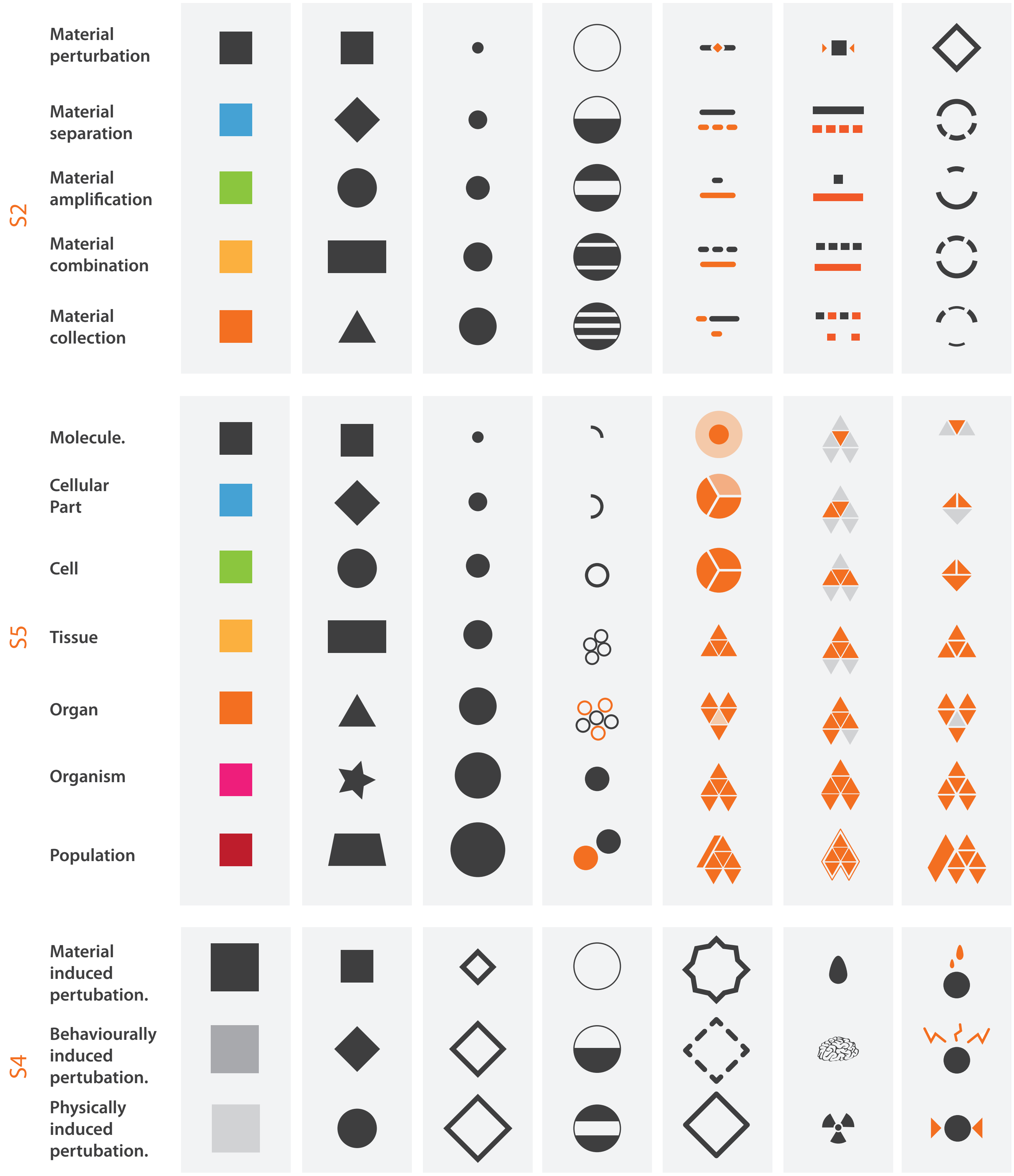
- human perception built on relative comparisons
  - great if color contiguous
  - surprisingly bad for absolute comparisons
- noncontiguous small regions of color
  - fewer bins than you want
  - rule of thumb: 6-12 bins, including background and highlights



[Cinteny: flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. Sinha and Meller. BMC Bioinformatics, 8:82, 2007.]

Glyphs

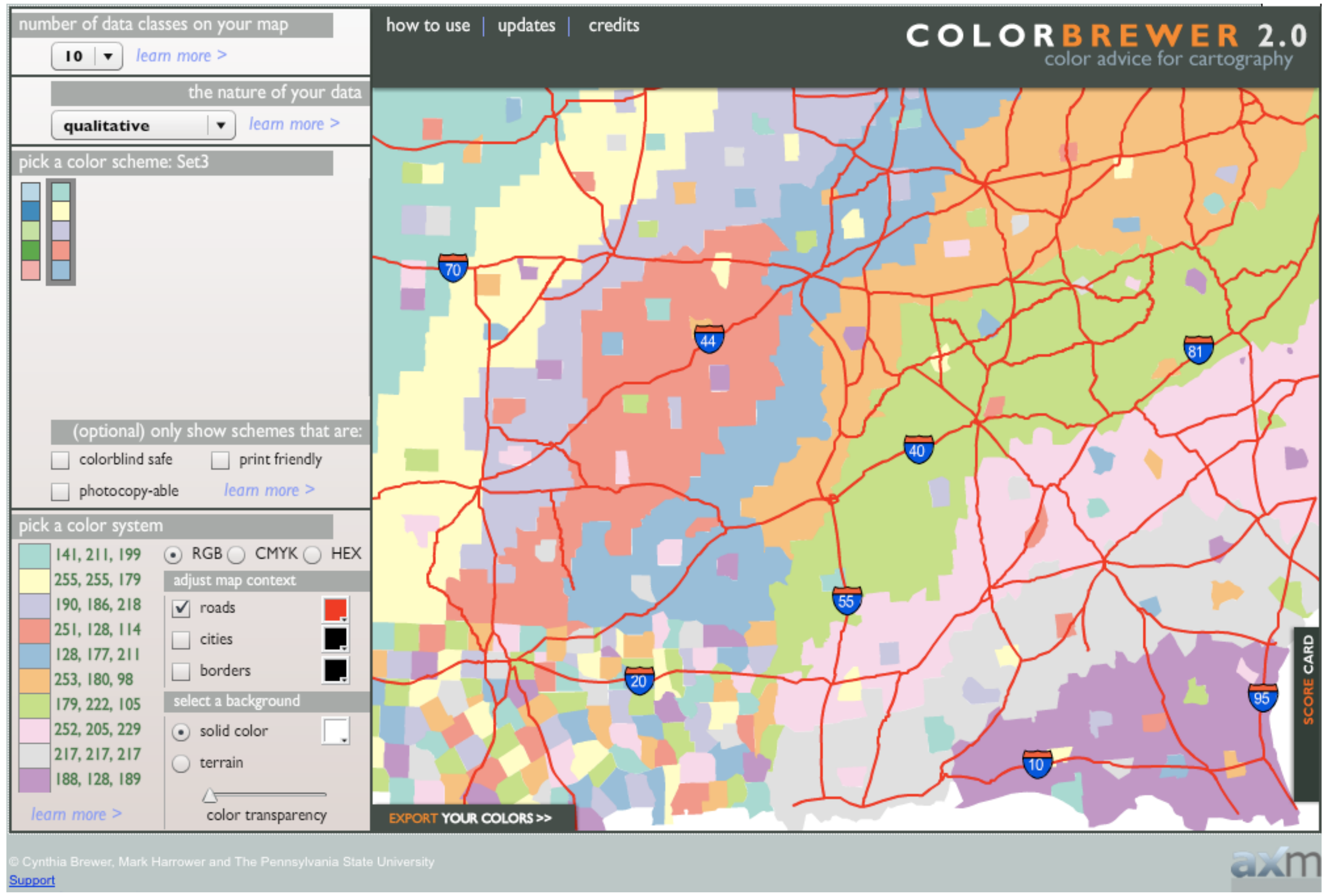
- glyphs: composite objects
  - internal structure with multiple marks
- alternative to color coding
  - or coding with any single channel

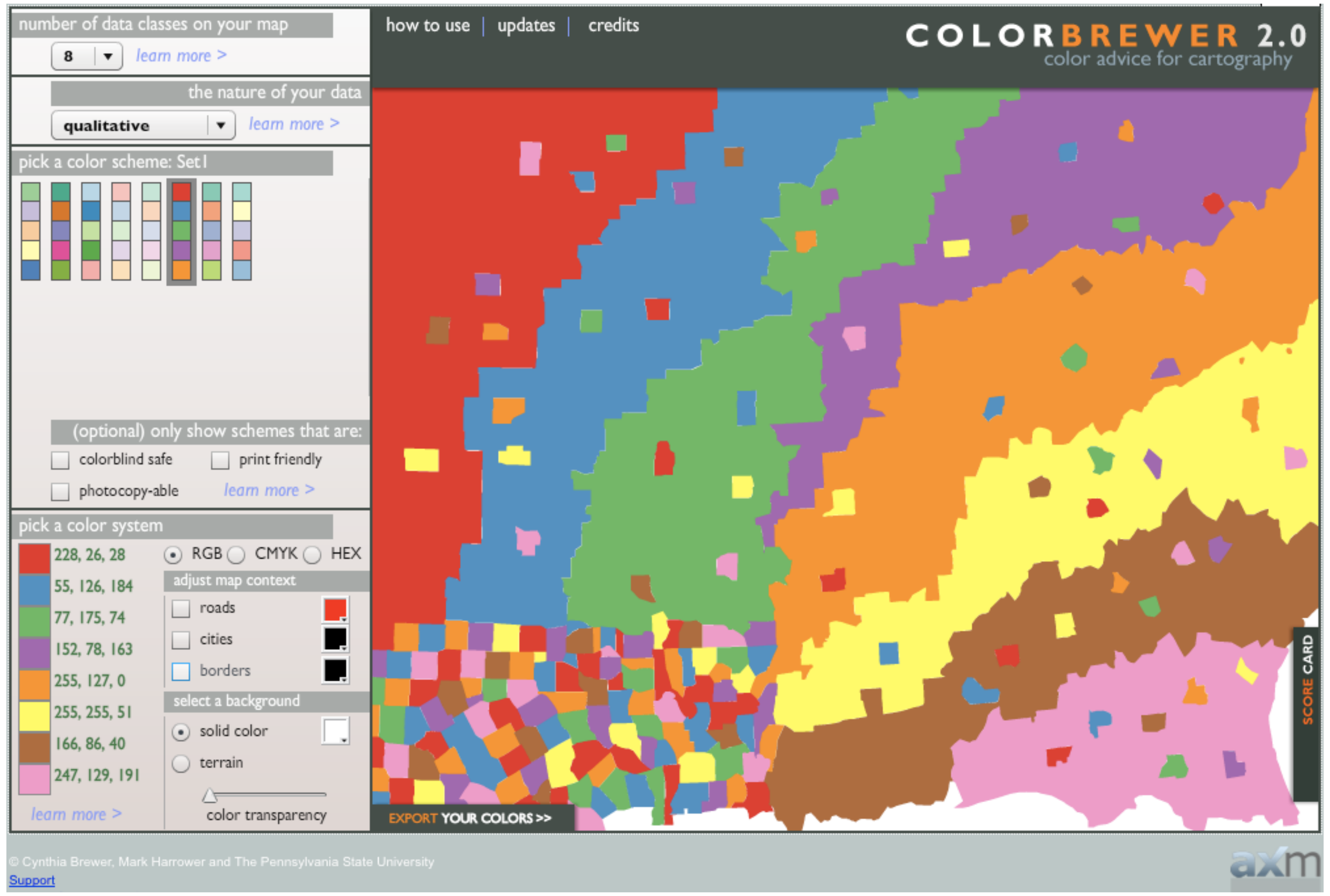


[Fig 5. Taxonomy-Based Glyph Design - with a Case Study on Visualizing Workflows of Biological Experiments. Maguire, Rocca-Serra, Sansone, Davies, and Chen. IEEE Trans. Visualization and Computer Graphics 18:12:2603-2612 (Proc. InfoVis 12).]

ColorBrewer

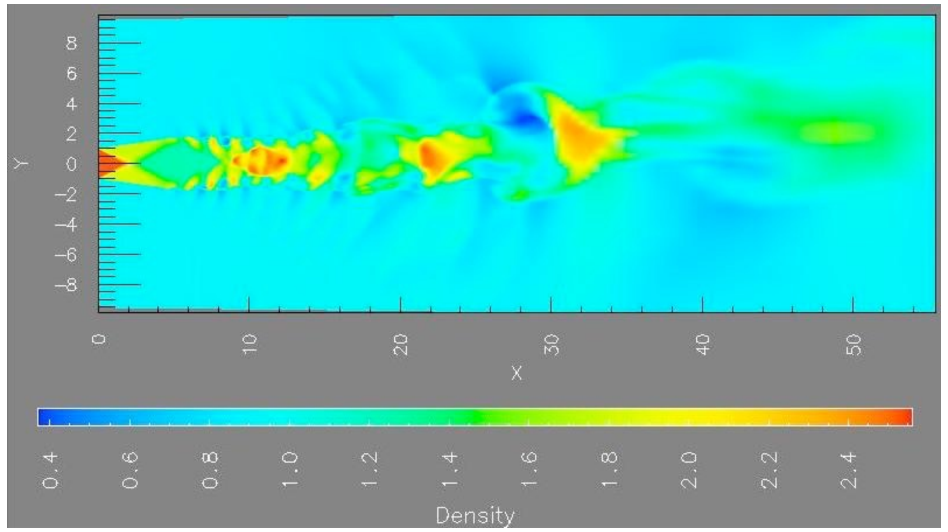
- http://www.colorbrewer2.org
- saturation and area example: size affects salience!



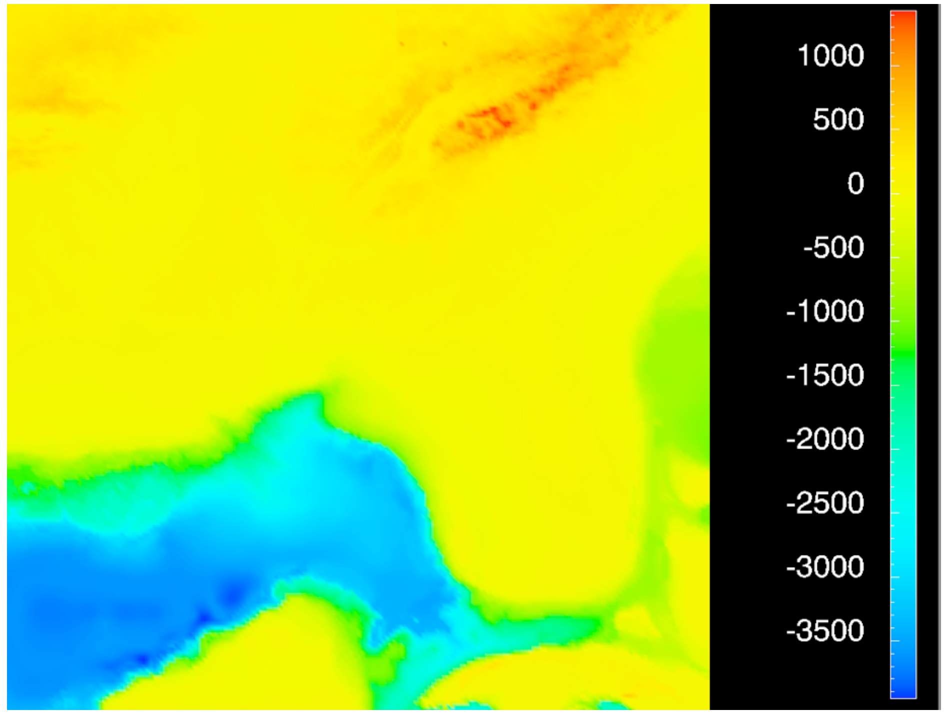


Ordered color: Rainbow is poor default

- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable



[A Rule-based Tool for Assisting Colormap Selection. Bergman,. Rogowitz, and. Treinish. Proc. IEEE Visualization (Vis), pp. 118–125, 1995.]

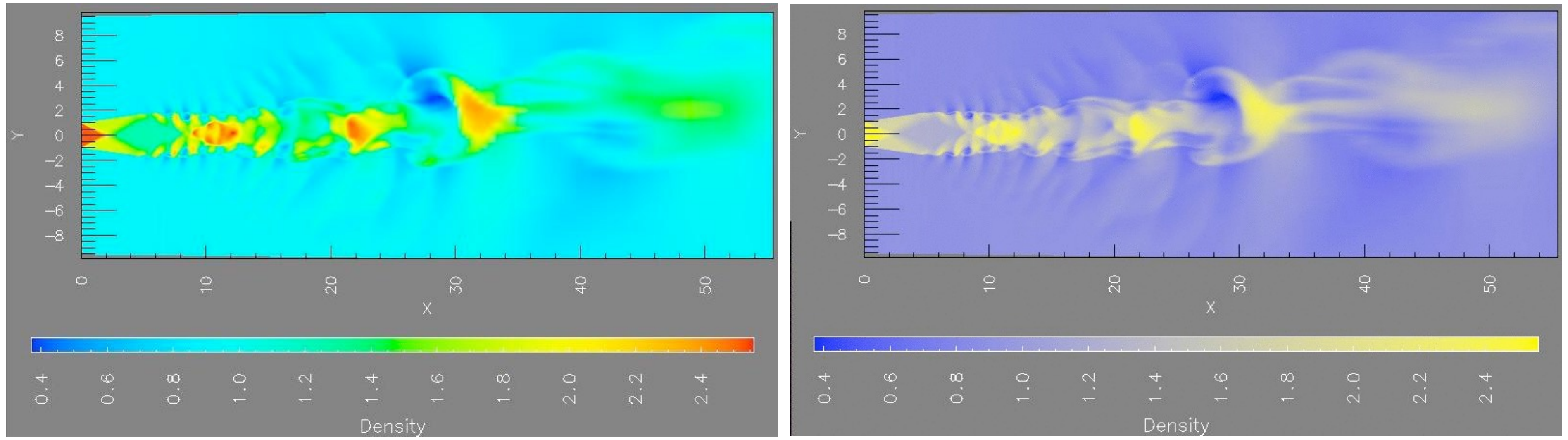


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

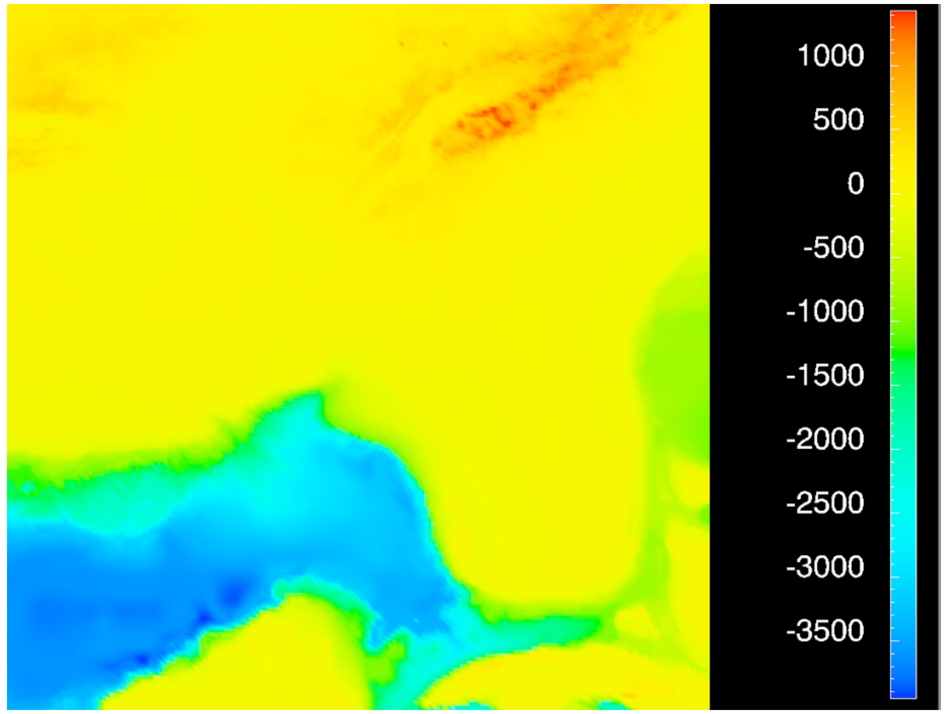
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

Ordered color: Rainbow is poor default

- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable
- alternatives
  - large-scale structure: fewer hues



[A Rule-based Tool for Assisting Colormap Selection. Bergman,. Rogowitz, and. Treinish. Proc. IEEE Visualization (Vis), pp. 118–125, 1995.]

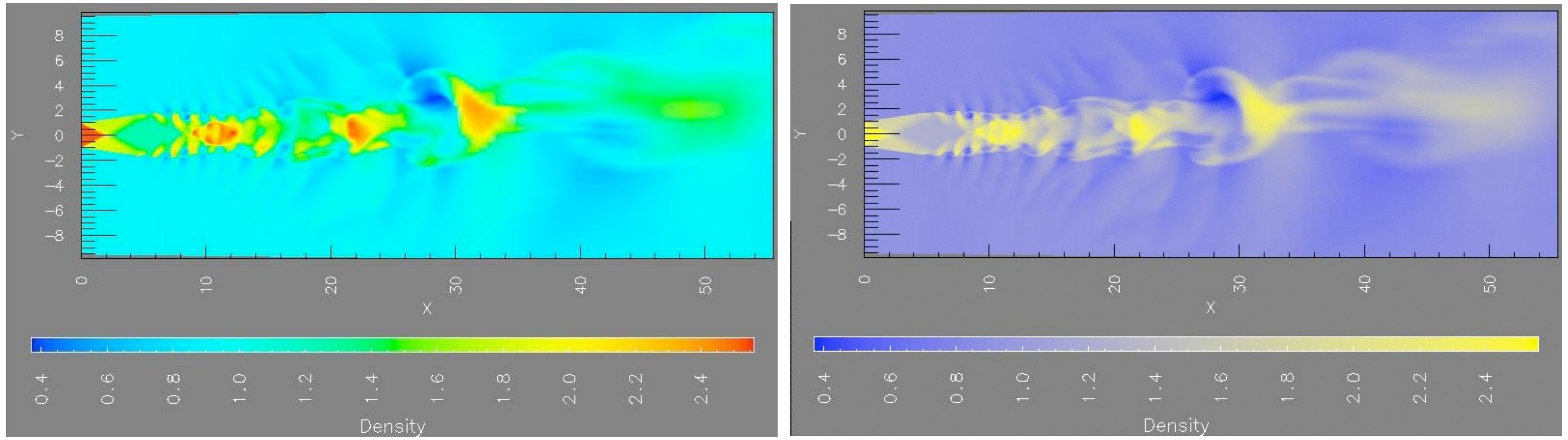


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

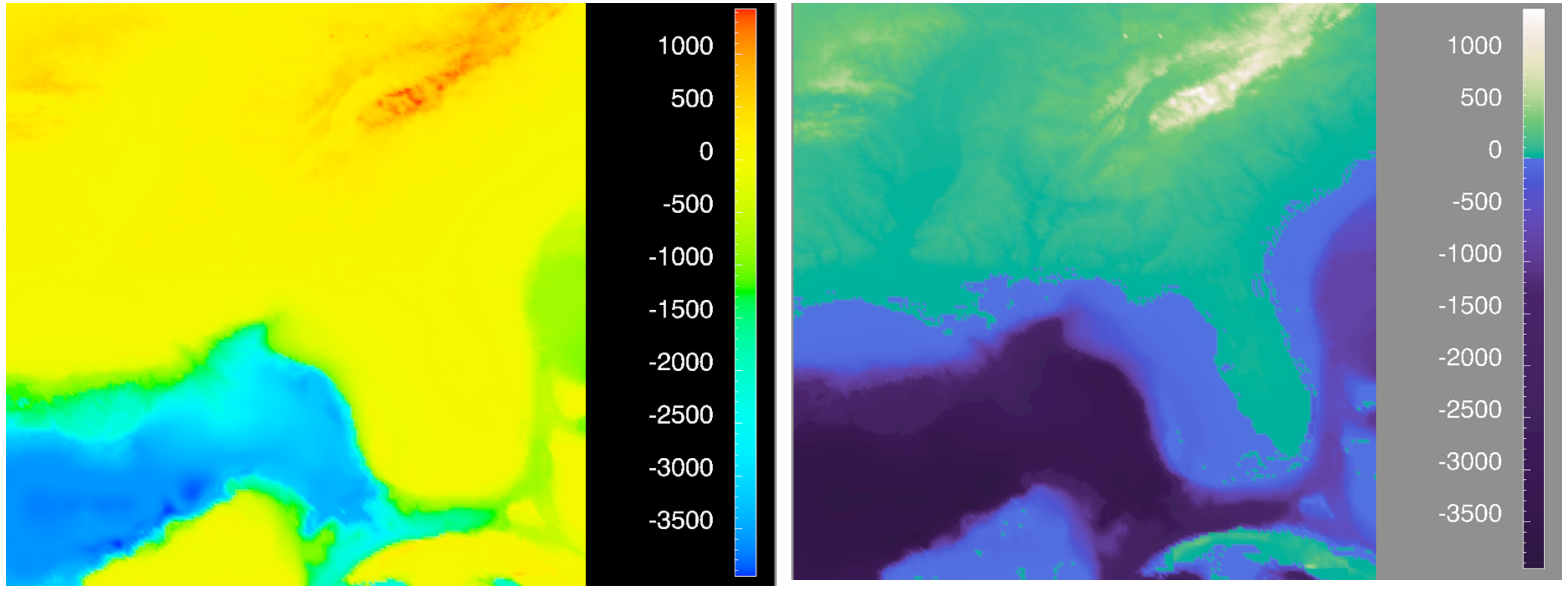
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

Ordered color: Rainbow is poor default

- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable
- alternatives
  - large-scale structure: fewer hues
  - fine structure: multiple hues with monotonically increasing luminance [eg viridis R/python]



[A Rule-based Tool for Assisting Colormap Selection. Bergman,. Rogowitz, and. Treinish. Proc. IEEE Visualization (Vis), pp. 118–125, 1995.]

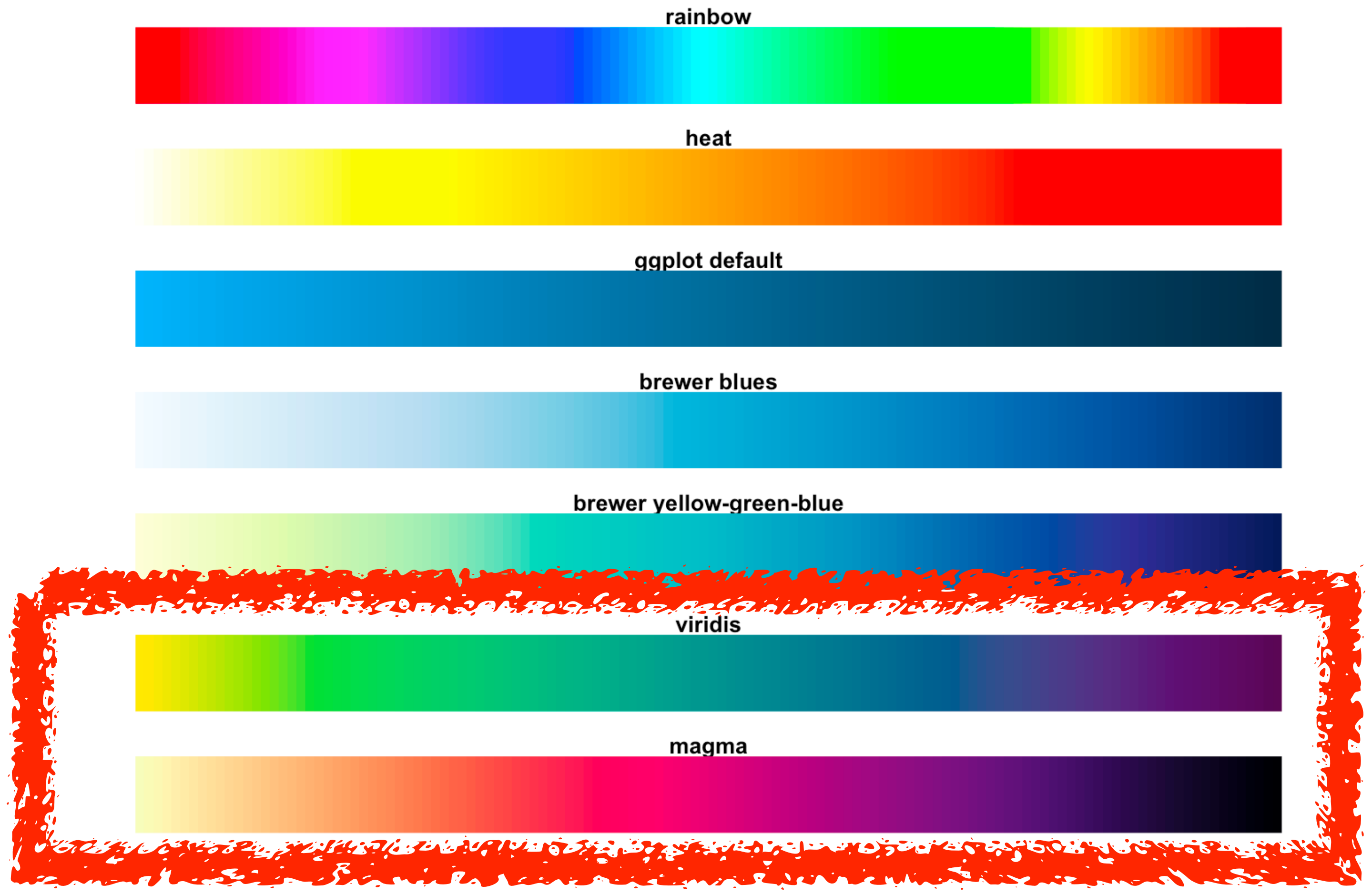


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

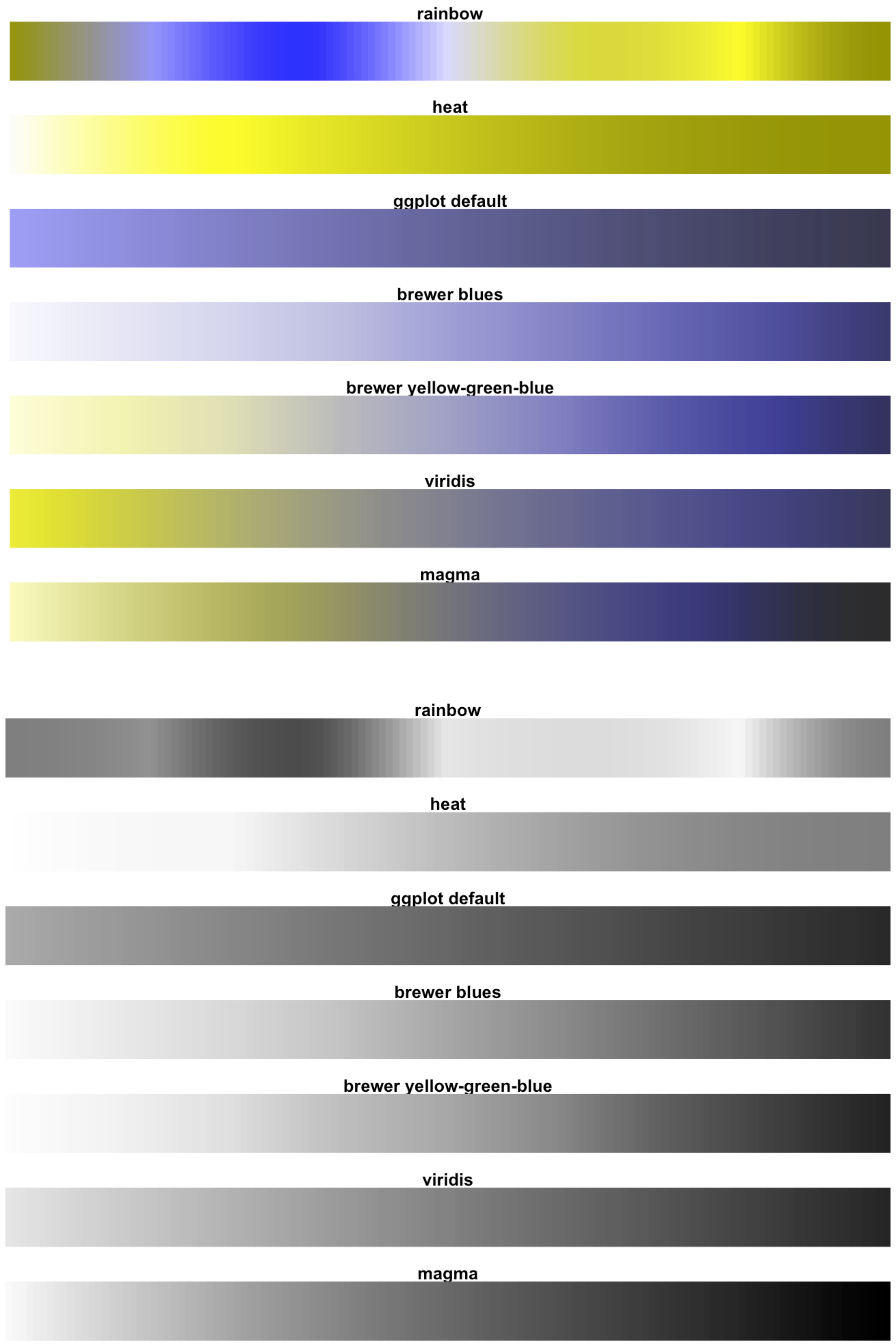
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

Viridis

- colorful, perceptually uniform, colorblind-safe, monotonically increasing luminance

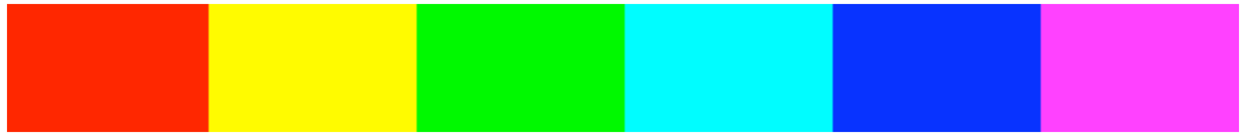


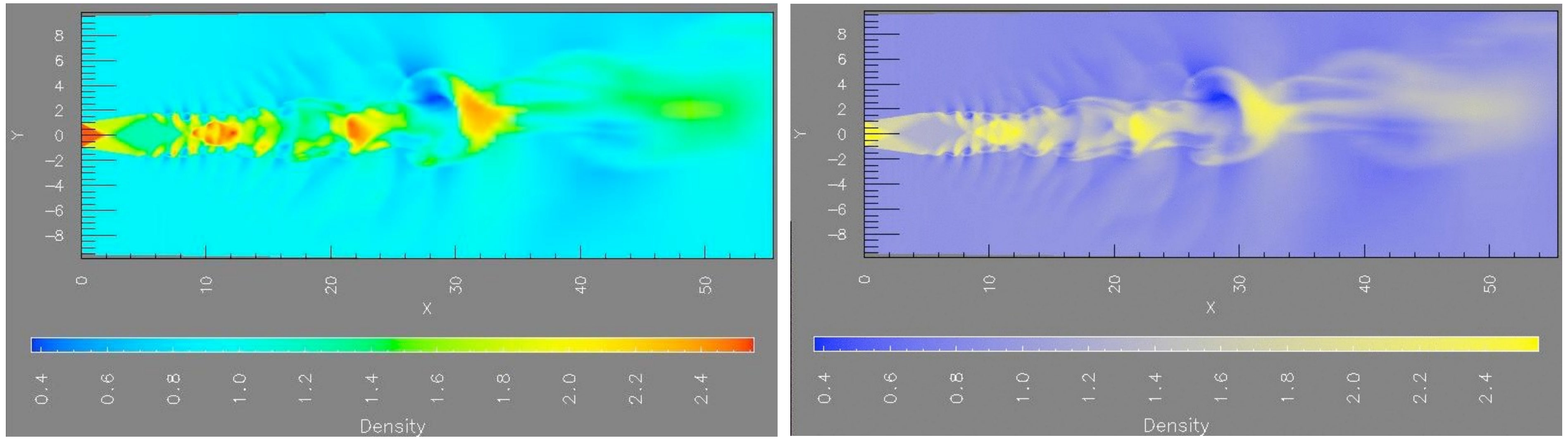
https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html



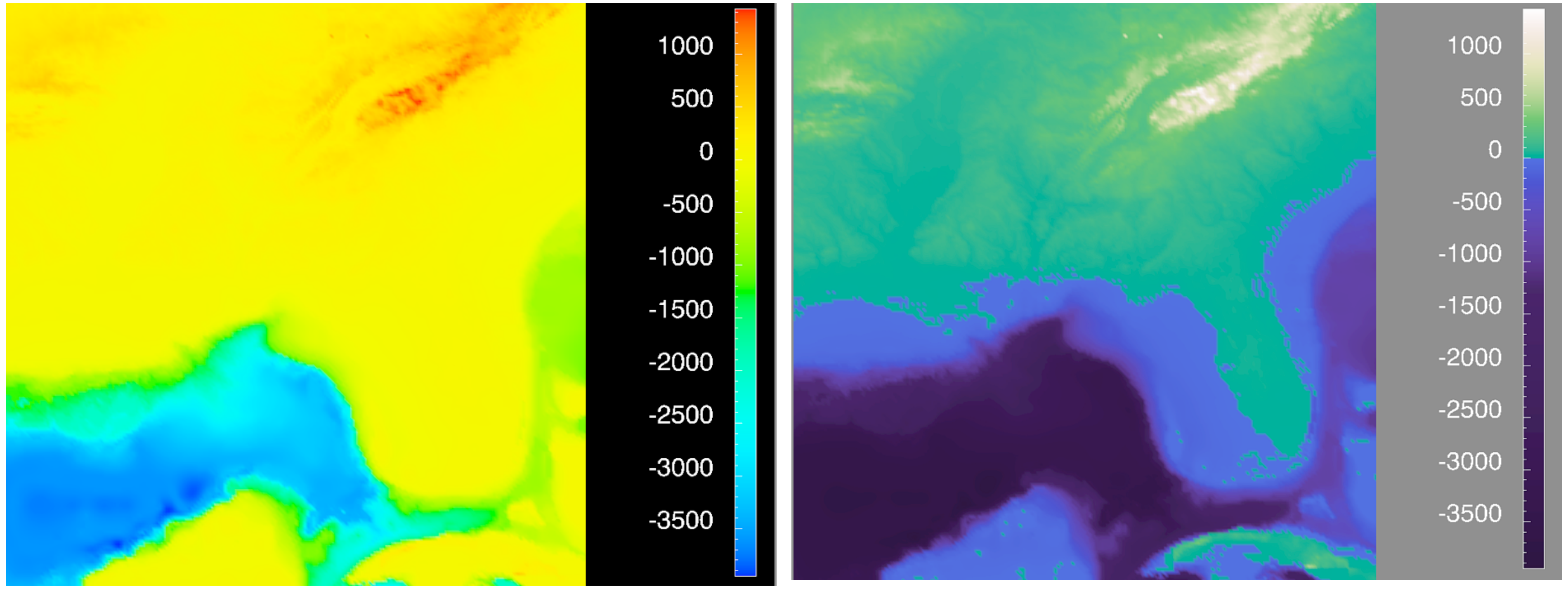
Ordered color: Rainbow is poor default

- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable
- alternatives
  - large-scale structure: fewer hues
  - fine structure: multiple hues with monotonically increasing luminance [eg viridis R/python]
  - segmented rainbows for binned or categorical





[A Rule-based Tool for Assisting Colormap Selection. Bergman,. Rogowitz, and. Treinish. Proc. IEEE Visualization (Vis), pp. 118–125, 1995.]



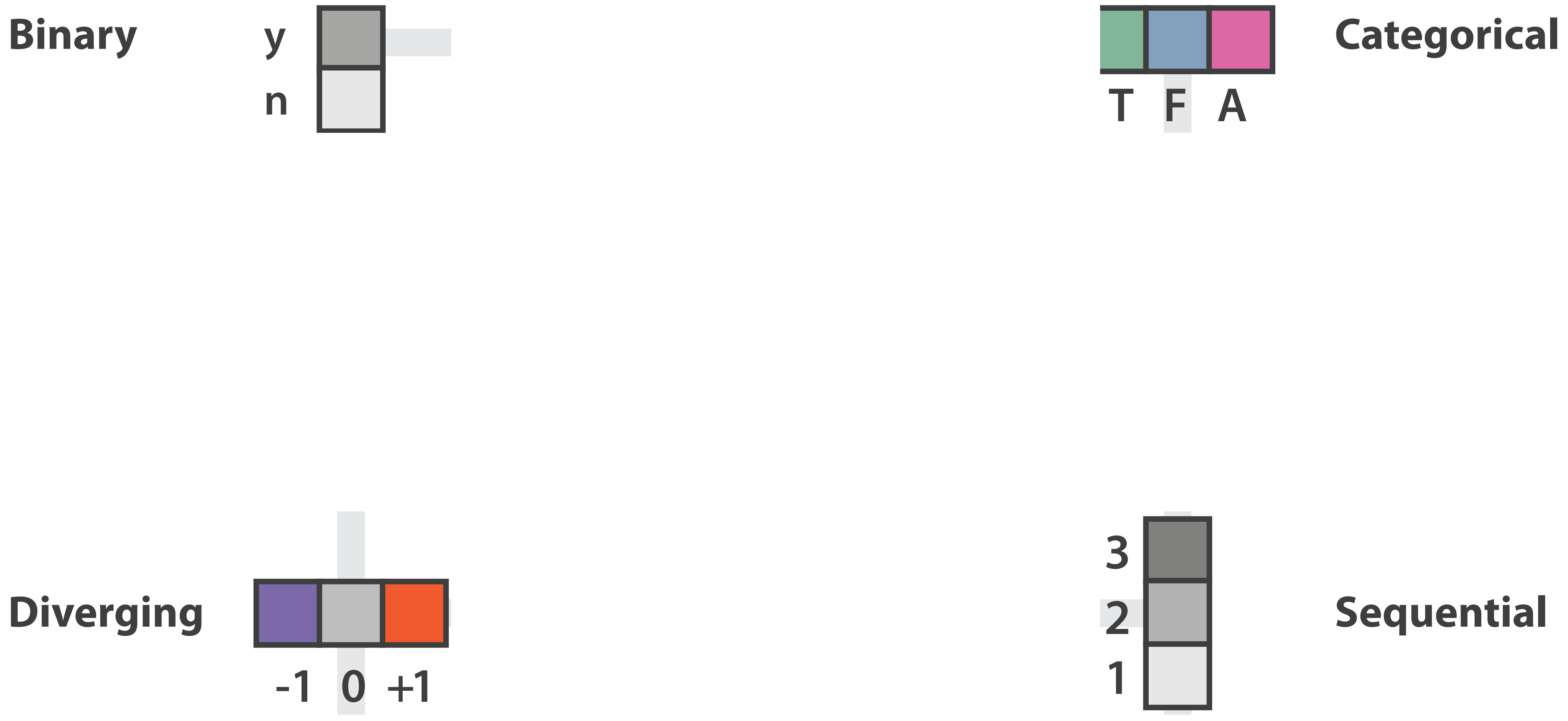
[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

Colormaps

- ➔ Categorical

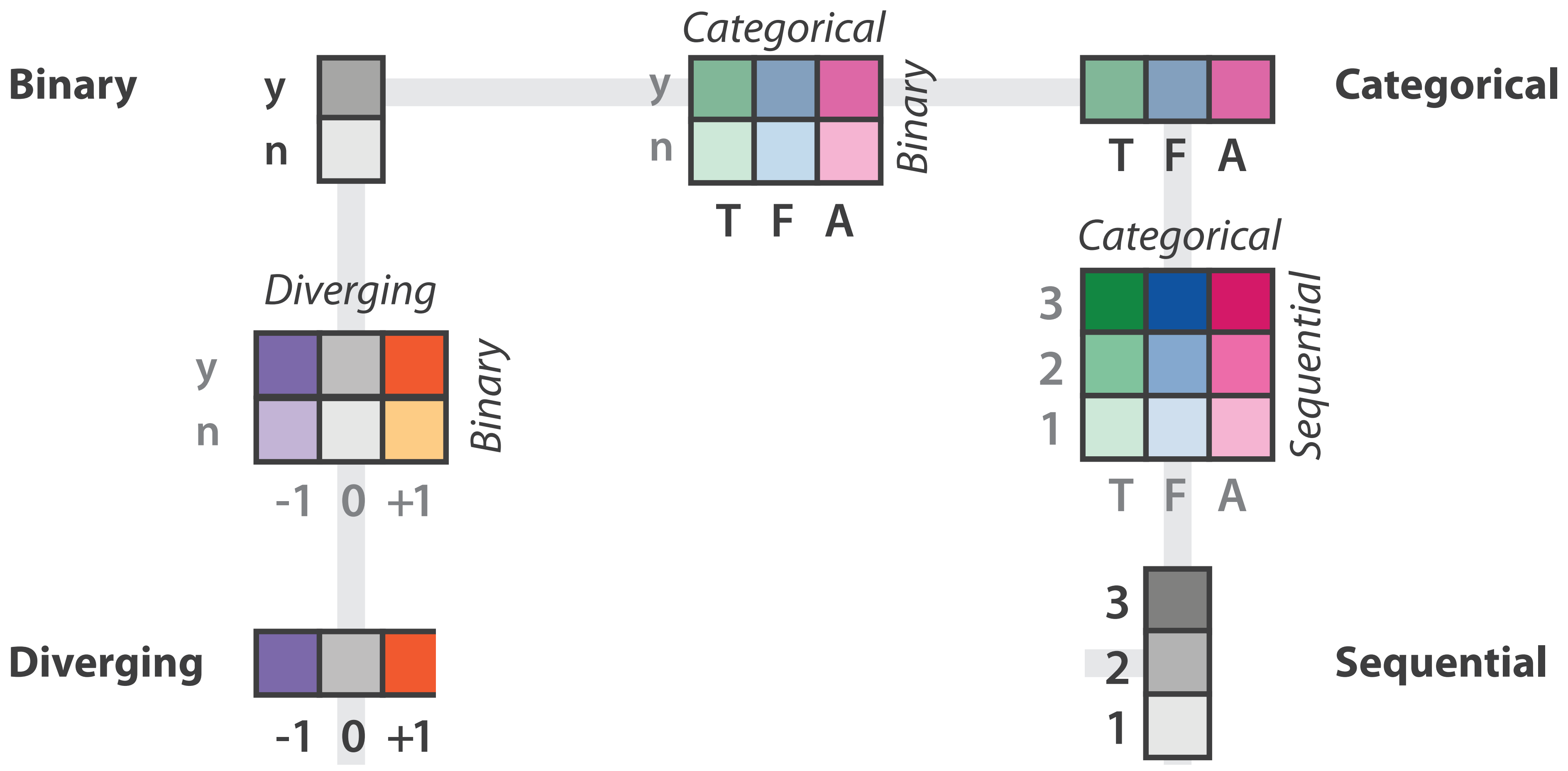
- ➔ Ordered
  - ➔ *Sequential*
  - ➔ *Diverging*



after [Color Use Guidelines for Mapping and Visualization. Brewer, 1994. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html]

Colormaps

- ➔ Categorical
- ➔ Ordered
  - ➔ *Sequential*
  - ➔ *Diverging*
- ➔ Bivariate

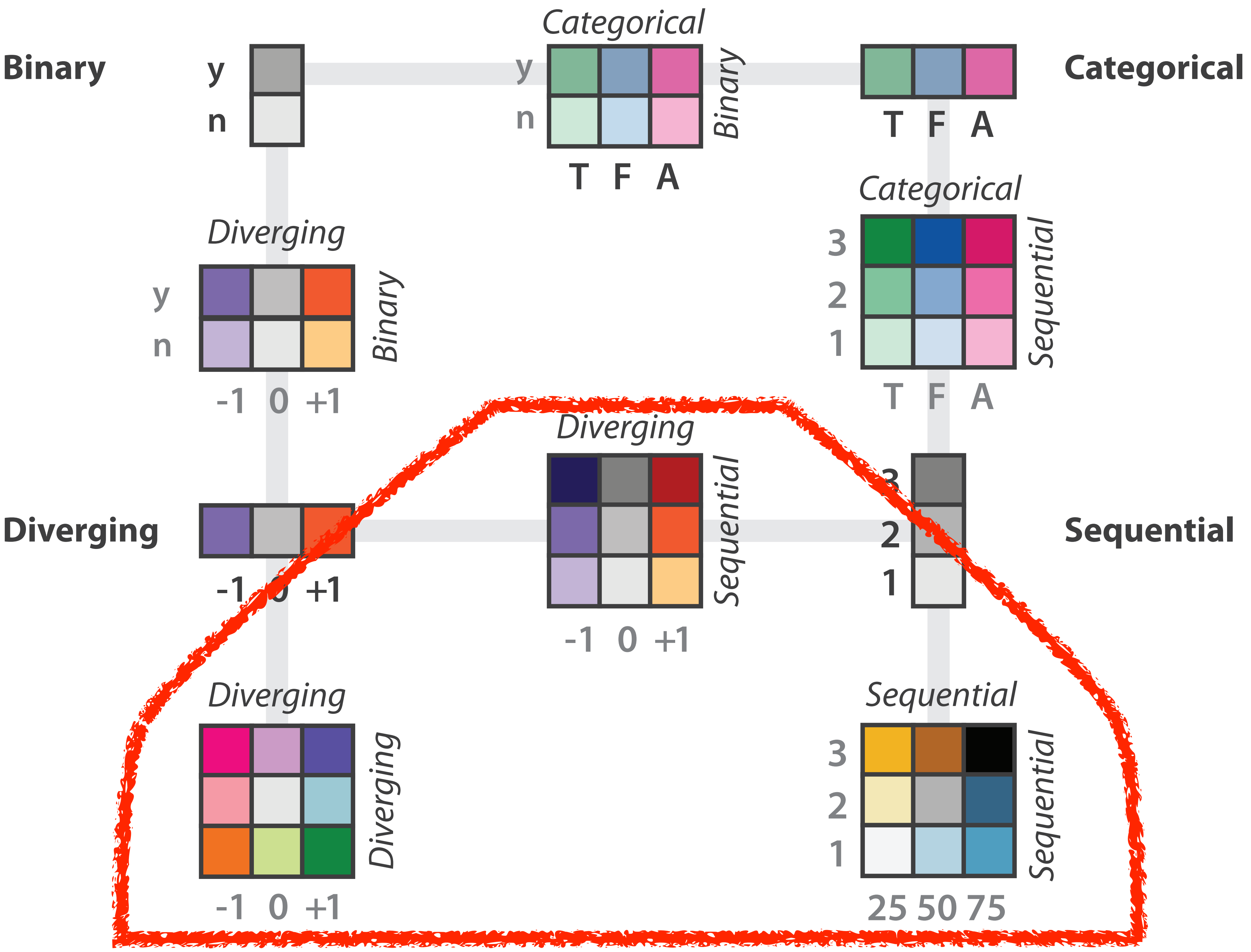


after [Color Use Guidelines for Mapping and Visualization. Brewer, 1994. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html]

Colormaps

- Categorical
- Ordered
  - *Sequential*
  - *Diverging*
- Bivariate

use with care!



after [Color Use Guidelines for Mapping and Visualization. Brewer, 1994. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html]

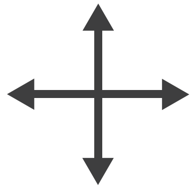
Colormaps

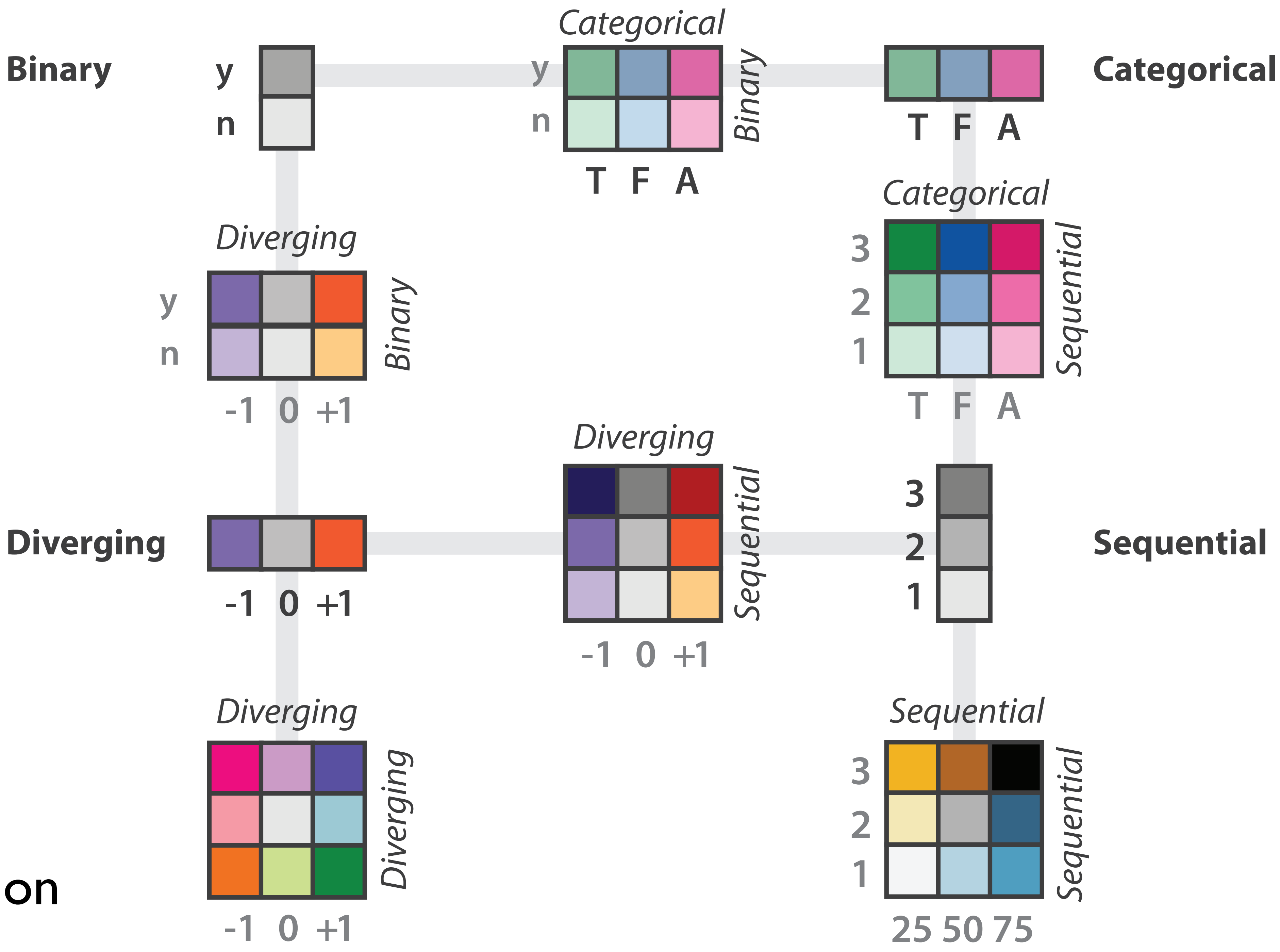
- ➔ Categorical

- ➔ Ordered
  - ➔ *Sequential*

  - ➔ *Diverging*

- ➔ Bivariate



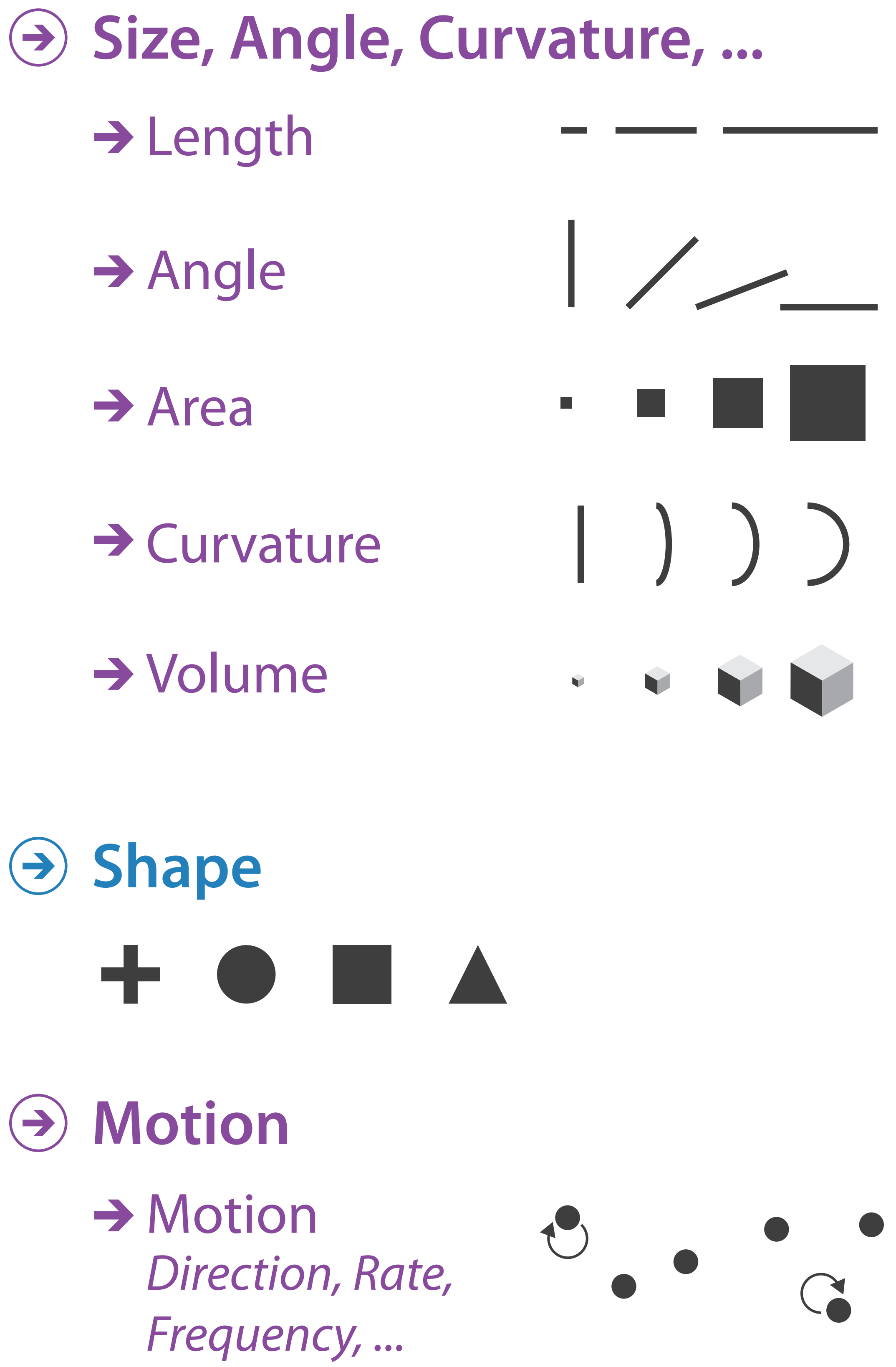


- color channel interactions
  - size heavily affects salience
    - small regions need high saturation
    - large need low saturation
  - saturation & luminance: 3-4 bins max
    - also not separable from transparency

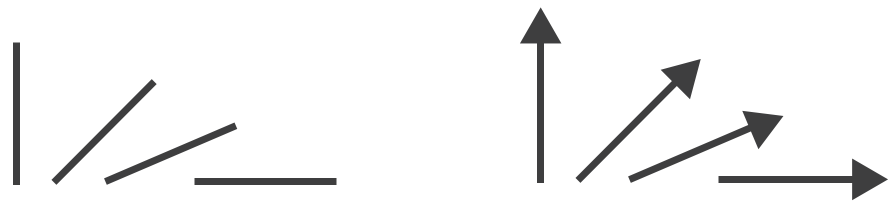
after [Color Use Guidelines for Mapping and Visualization. Brewer, 1994. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html]

Map other channels

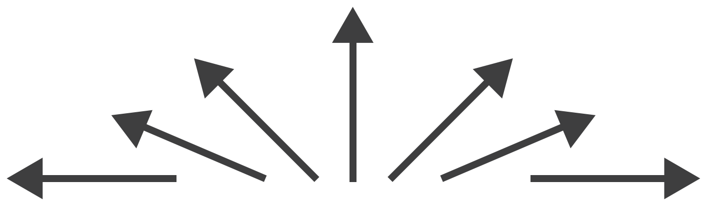
- size
  - length accurate, 2D area ok, 3D volume poor
- angle
  - nonlinear accuracy
    - horizontal, vertical, exact diagonal
- shape
  - complex combination of lower-level primitives
  - many bins
- motion
  - highly separable against static
    - binary: great for highlighting
  - use with care to avoid irritation



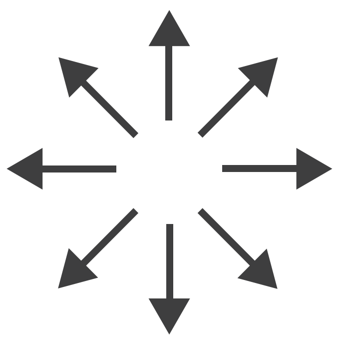
Angle



Sequential ordered
line mark or arrow glyph



Diverging ordered
arrow glyph



Cyclic ordered
arrow glyph

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, 2014
  - *Chap 10: Map Color and Other Channels*
- ColorBrewer, Brewer.
  - http://www.colorbrewer2.org
- *Color In Information Display*. Stone. IEEE Vis Course Notes, 2006.
  - http://www.stonesc.com/Vis06
- A Field Guide to Digital Color. Stone. AK Peters, 2003.
- *Rainbow Color Map (Still) Considered Harmful.* Borland and Taylor. IEEE Computer Graphics and Applications 27:2 (2007), 14–17.
- Visual Thinking for Design. Ware. Morgan Kaufmann, 2008.
- Information Visualization: Perception for Design, 3rd edition. Ware. Morgan Kaufmann / Academic Press, 2004.
- https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html

Further reading, full

- Information Visualization: Perception for Design, 3rd edition, Colin Ware, Morgan Kaufmann, 2013.
- A Field Guide To Digital Color, Maureen Stone, AK Peters 2003.
- Representing Colors as Three Numbers, Maureen Stone, IEEE Computer Graphics and Applications, 25(4), July 2005, pp. 78-85.
- Color use guidelines for data representation C. Brewer, 1999.
- How Not to Lie with Visualization, Bernice E. Rogowitz and Lloyd A. Treinish, Computers In Physics 10(3) May/June 1996, pp 268-273.
- Rainbow Color Map (Still) Considered Harmful. David Borland and Russell M. Taylor, III. IEEE Computer Graphics and Applications 27:2 (2007), 14-17.

- Information Visualization: Perception for Design, 3rd edition, Colin Ware, Morgan Kaufmann, 2013.

- A Rule-Based Tool for Assisting Colormap Selection.. Lawrence D. Bergman, Bernice E. Rogowitz, and Lloyd A. Treinish. Proc. IEEE Visualization (Vis) 1995, 118-125.
- An Empirical Inquiry Concerning Human Understanding of Two-Variable Color Maps.. Howard Wainer and Carl M. Francolini. The American Statistician 34:2 (1980), 81-93.
- Motion to Support Rapid Interactive Queries on Node-Link Diagrams.. Colin Ware and Robert Bobrow. Transactions on Applied Perception (TAP) 1:1 (2004), 3-18.

Ch 11. Manipulate View

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

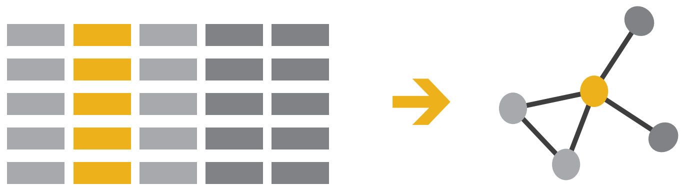
What?

Why?

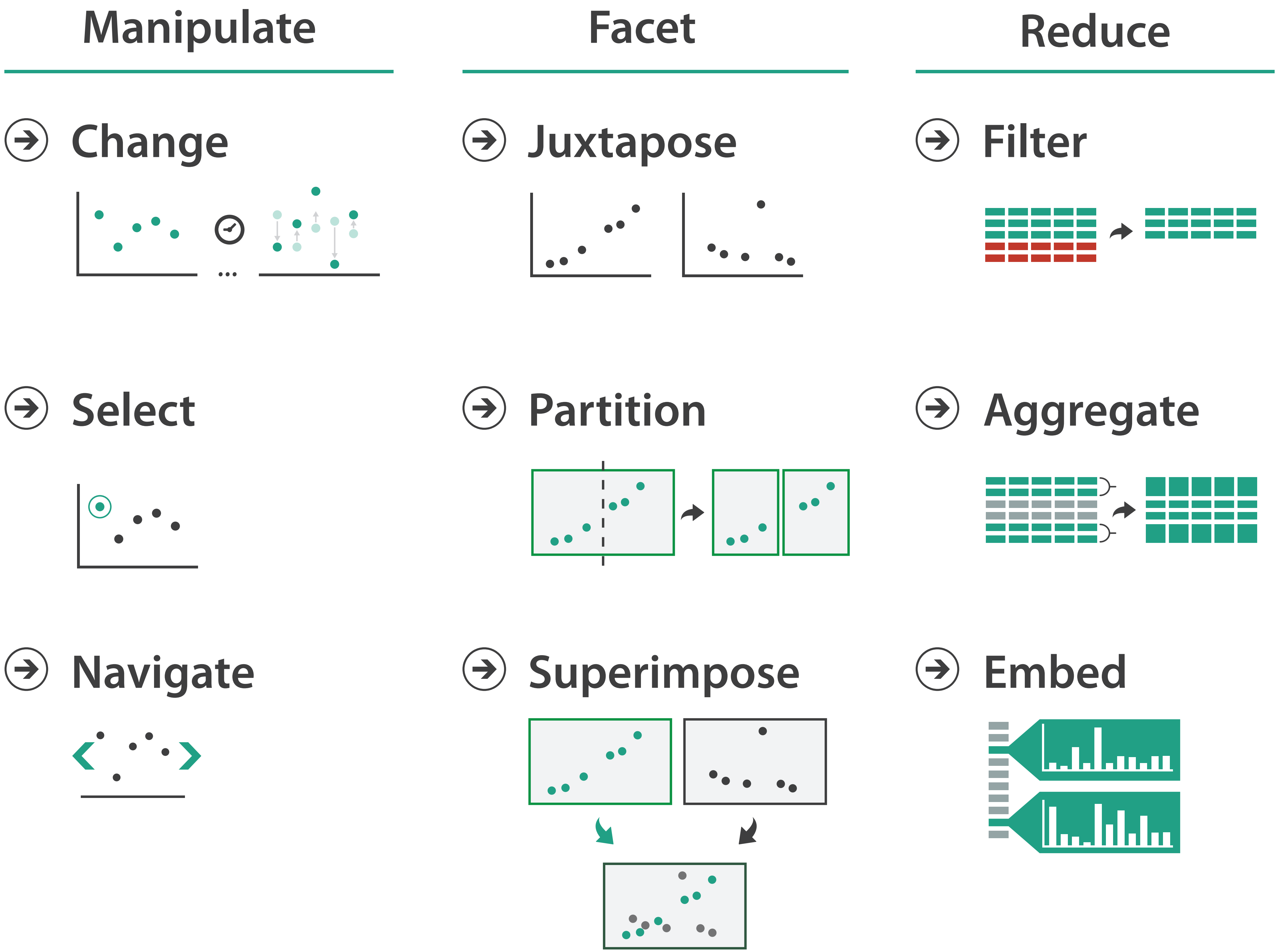
How?

How to handle complexity: 1 previous strategy + 3 more

→ *Derive*

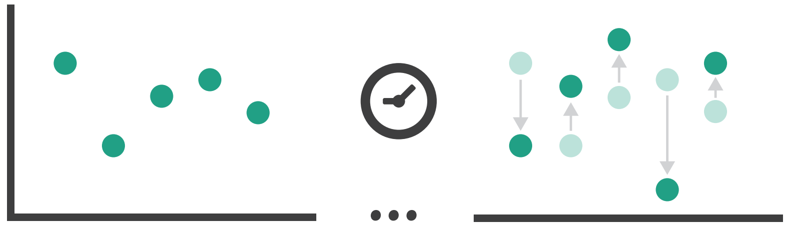


- derive new data to show within view
- change view over time
- facet across multiple views
- reduce items/attributes within single view

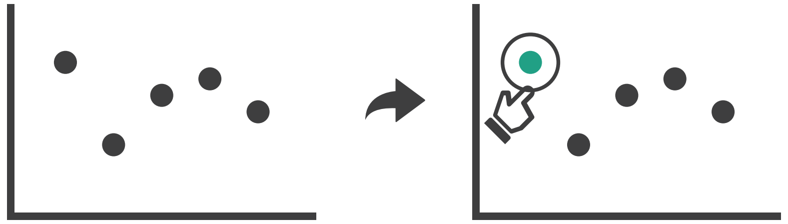


Manipulate

Change over Time



Select

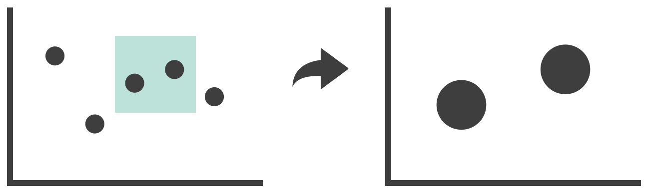


Navigate

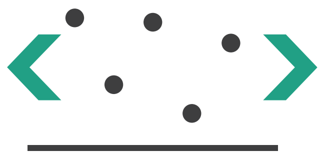
Item Reduction

Zoom

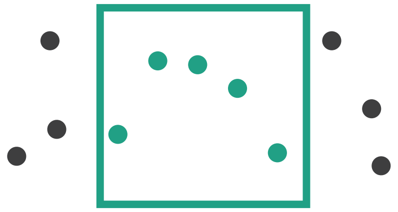
Geometric or *Semantic*



Pan/Translate



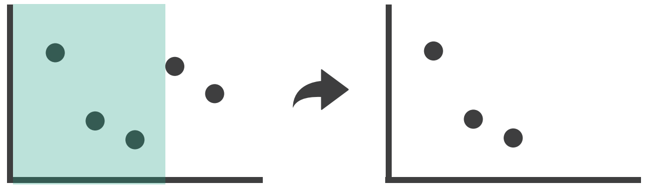
Constrained



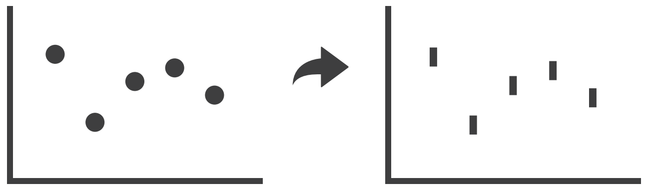
Attribute Reduction

Slice

Cut



Project

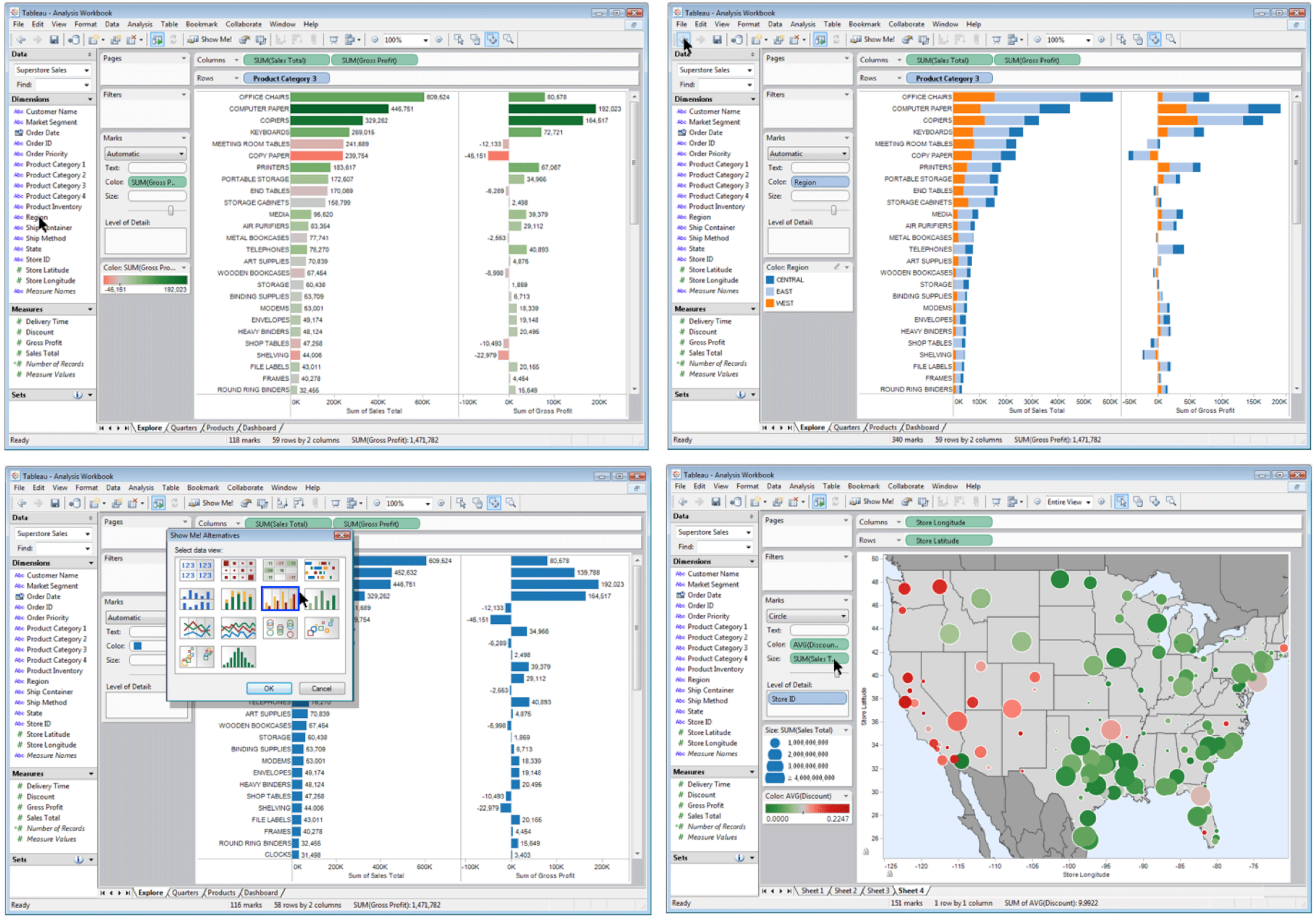


Change over time

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

  - interaction entails change

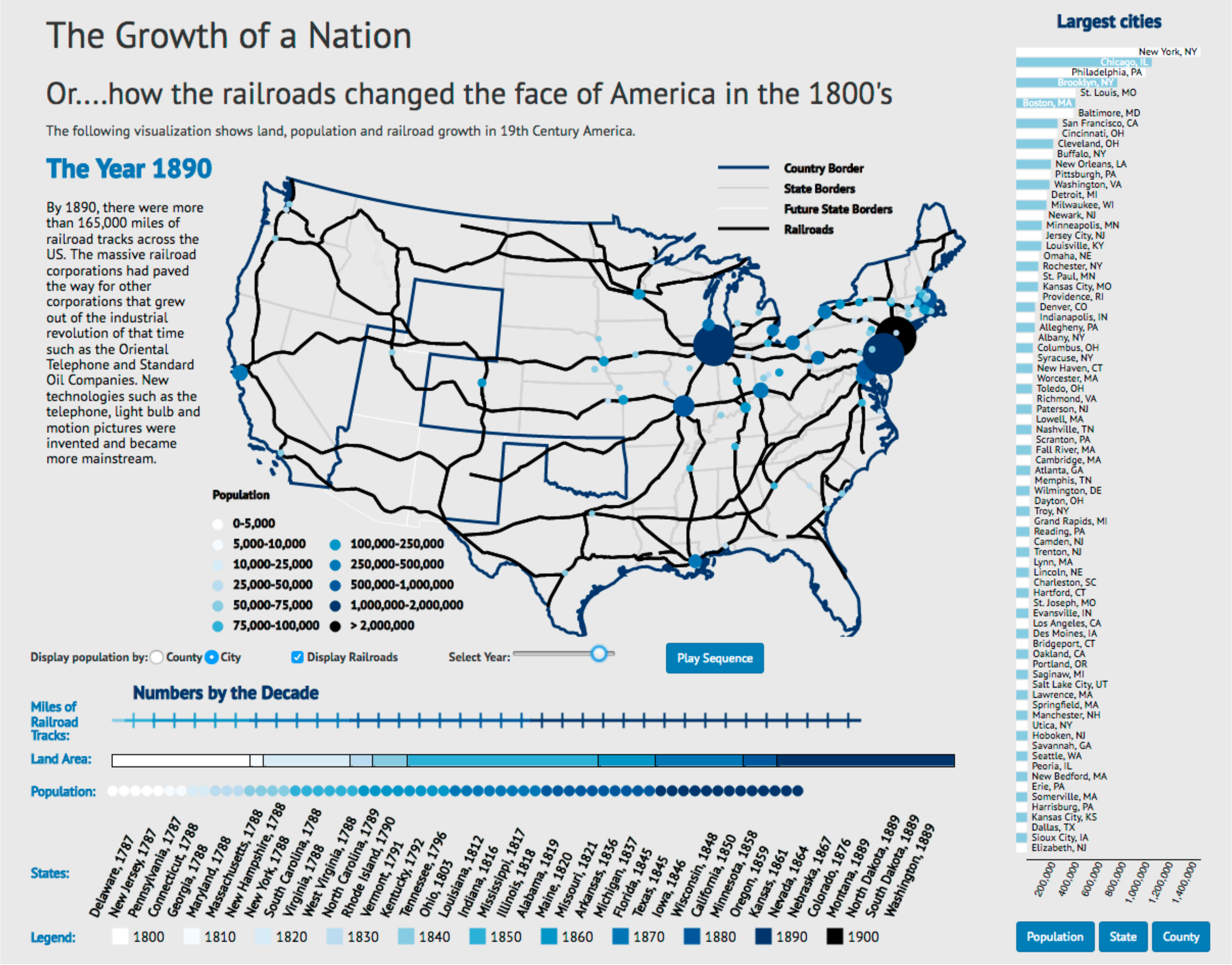
Idiom: **Re-encode** System: **Tableau**



made using Tableau, http://tableausoftware.com

Idiom: **Change parameters**

- widgets and controls
  - sliders, buttons, radio buttons, checkboxes, dropdowns/comboboxes
- pros
  - clear affordances, self-documenting (with labels)
- cons
  - uses screen space
- design choices
  - separated vs interleaved
    - controls & canvas

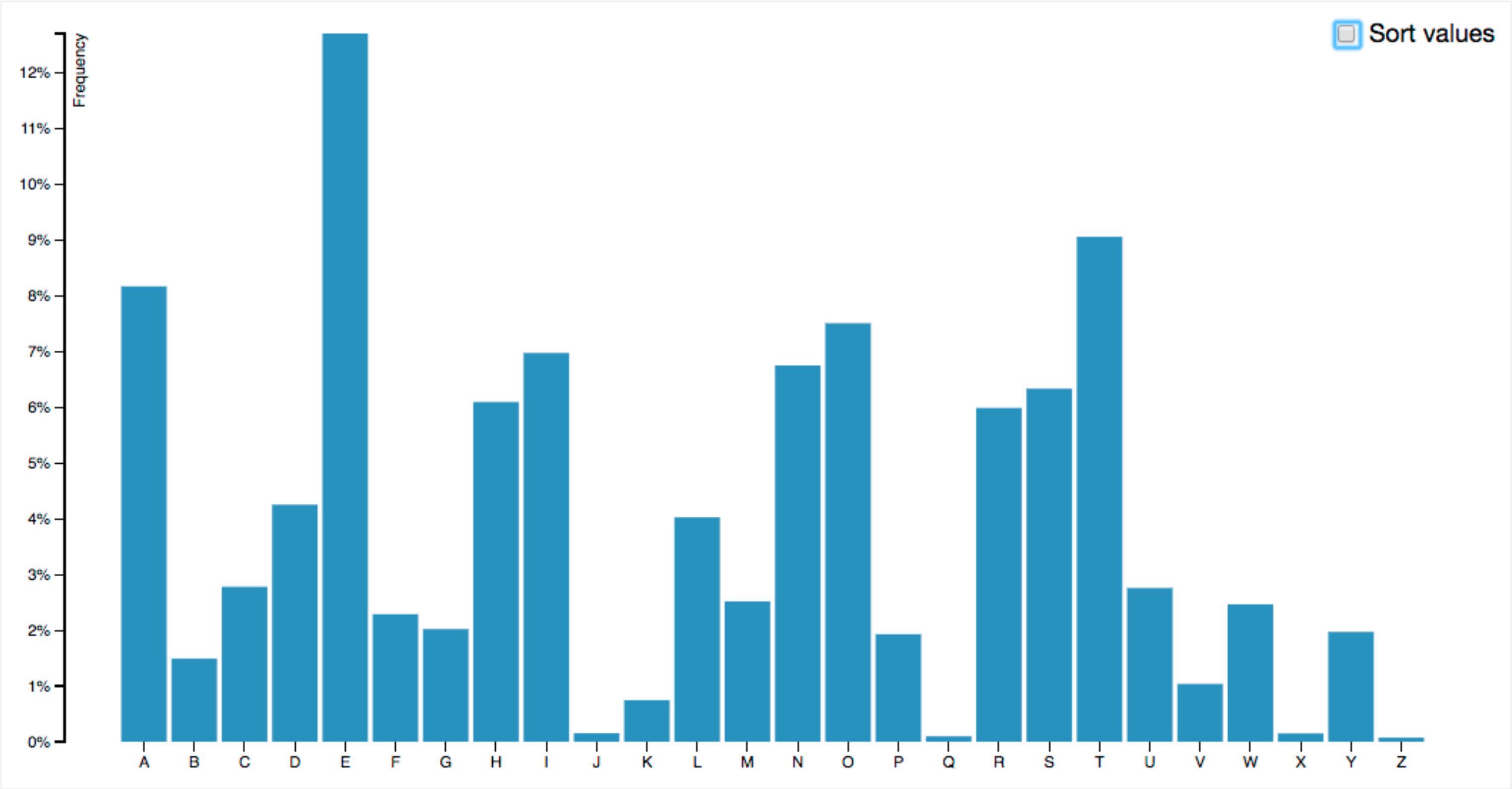


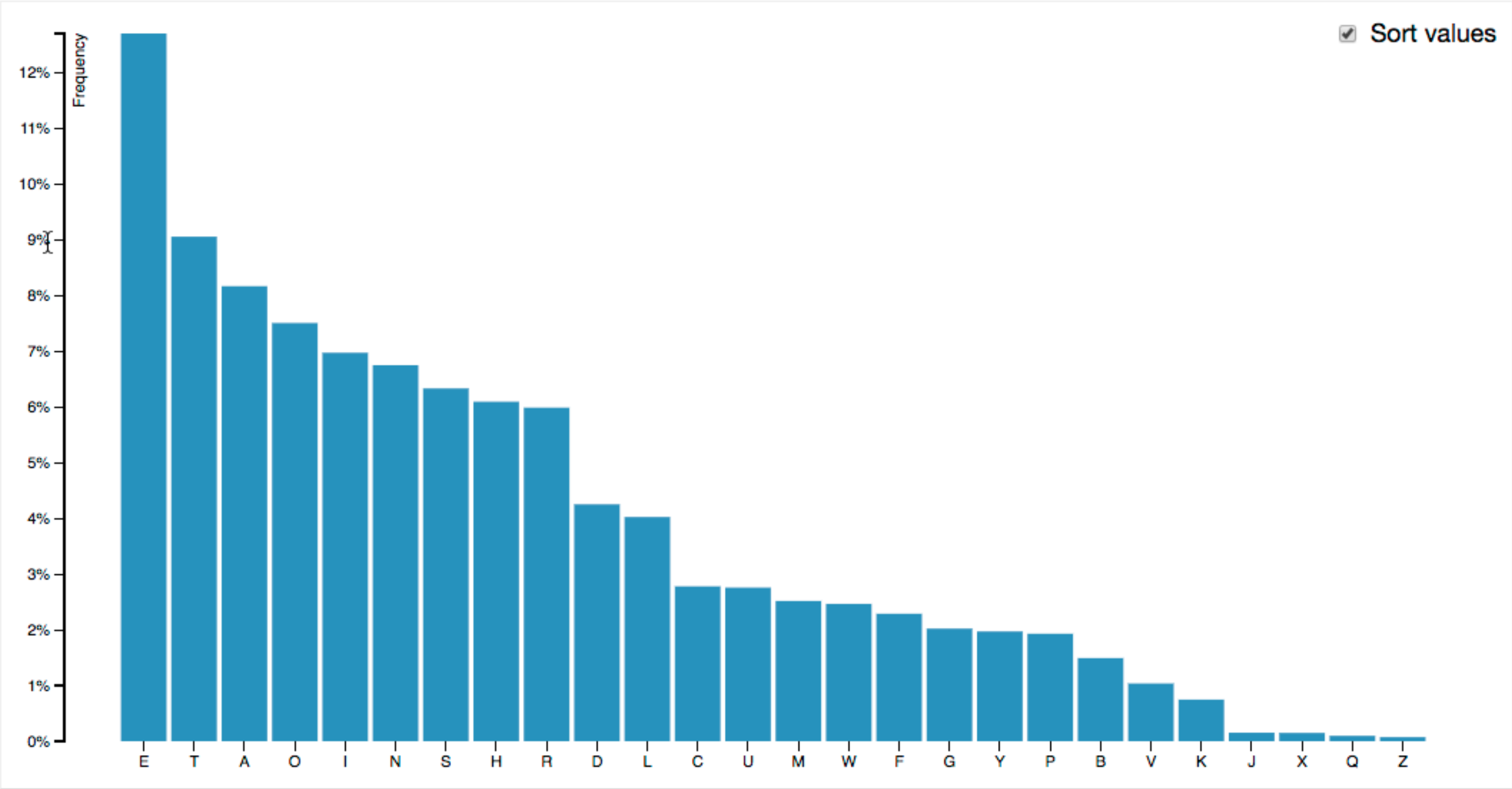
[Growth of a Nation](http://laurenwood.github.io/)

slide inspired by: Alexander Lex, Utah

Idiom: **Change order/arrangement**

- what: simple table
- how: data-driven reordering
- why: find extreme values, trends



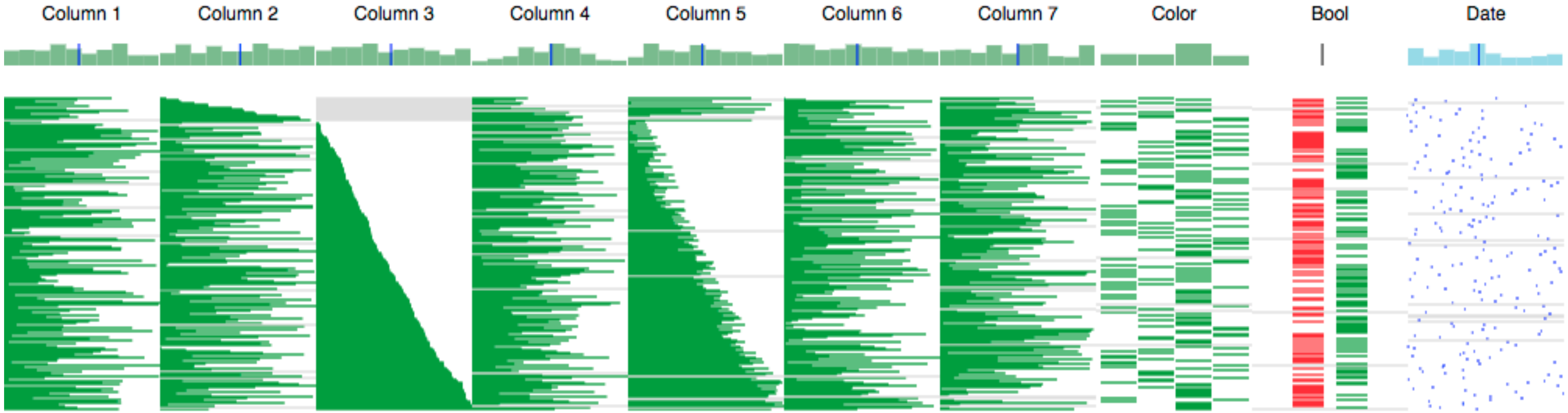


[Sortable Bar Chart](https://bl.ocks.org/mbostock/3885705)

Idiom: **Reorder**

System: **DataStripes**

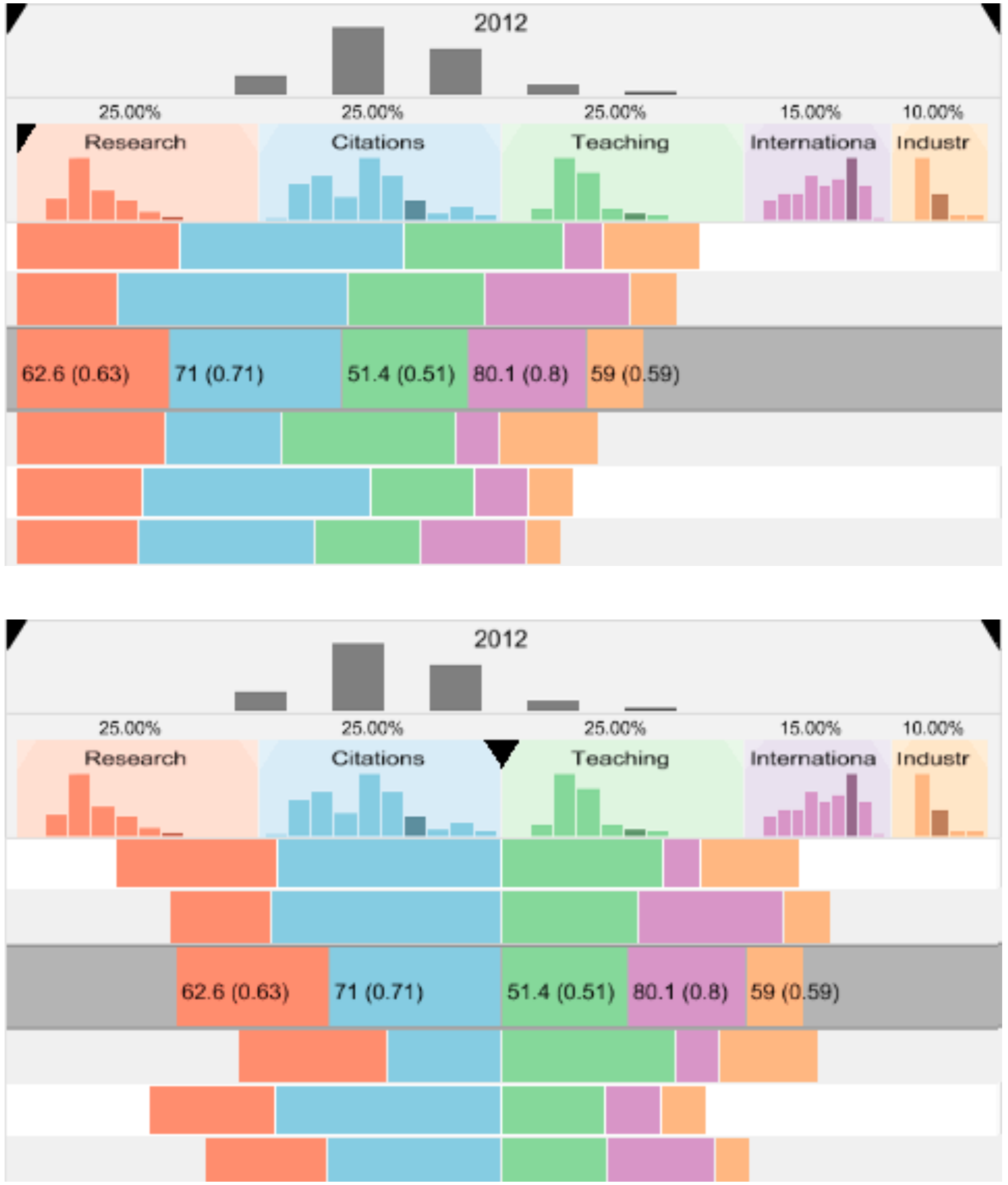
- what: table with many attributes
- how: data-driven reordering by selecting column
- why: find correlations between attributes



[*http://carlmanaster.github.io/datastripes/*]

Idiom: **Change alignment** System: **LineUp**

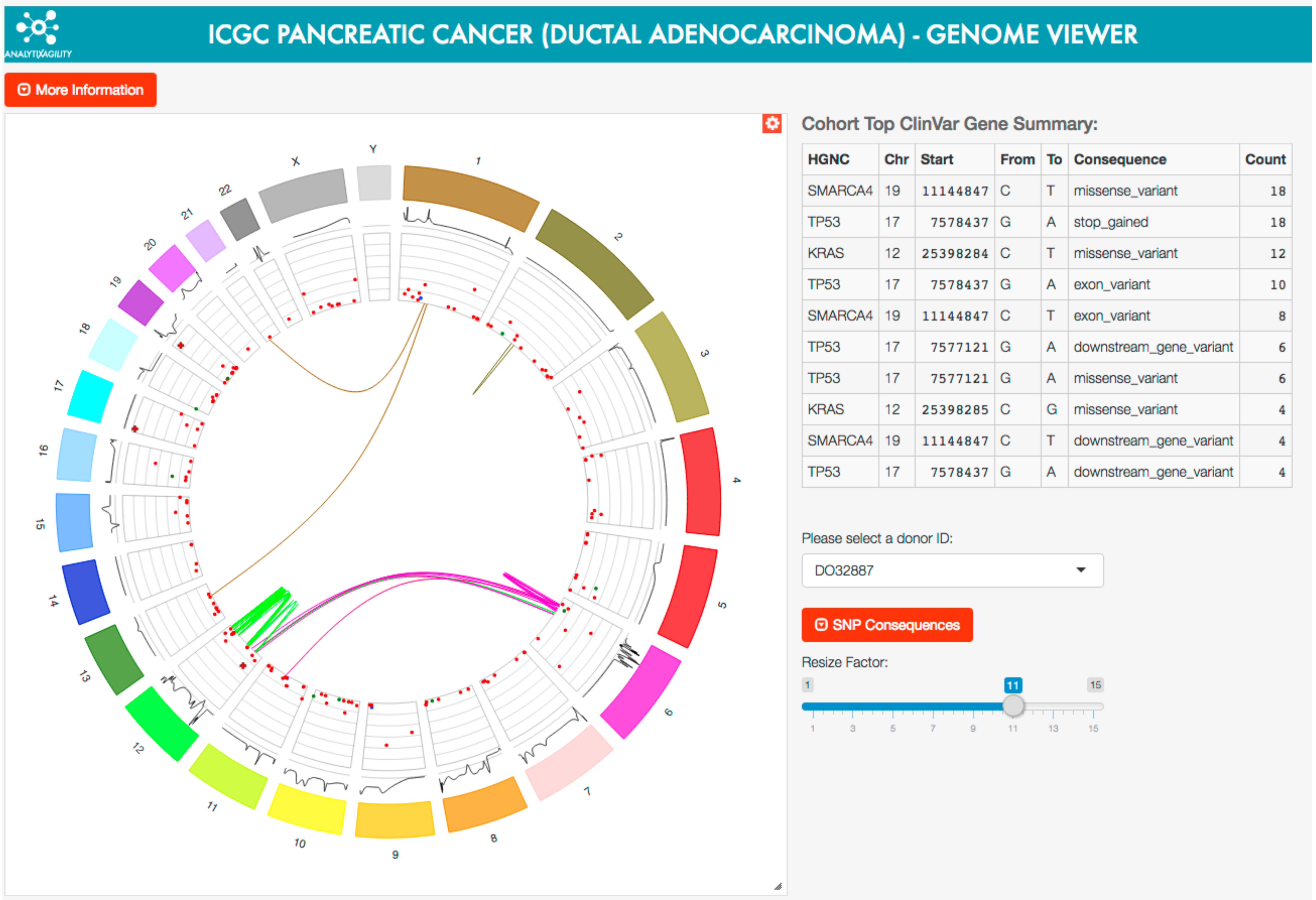
- stacked bars
  - easy to compare
    - first segment
    - total bar
- align to different segment
  - supports flexible comparison



[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.]

Shiny example

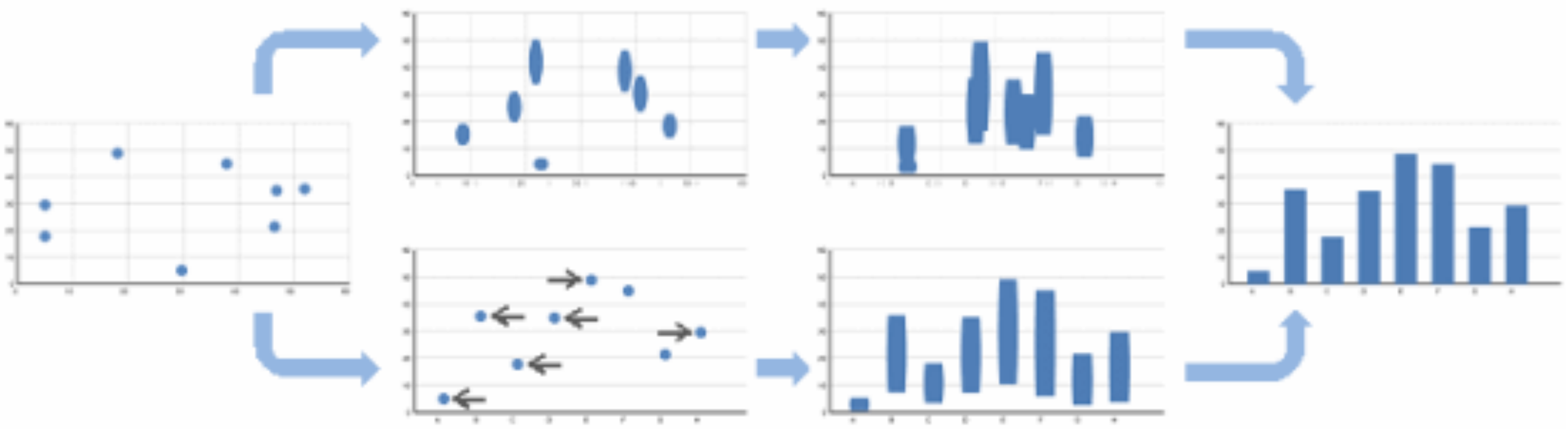
- APGI genome browser
  - tooling: R/Shiny
  - interactivity
    - tooltip detail on demand on hover
    - expand/contract chromosomes
    - expand/contract control panes



https://gallery.shinyapps.io/genome_browser/

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

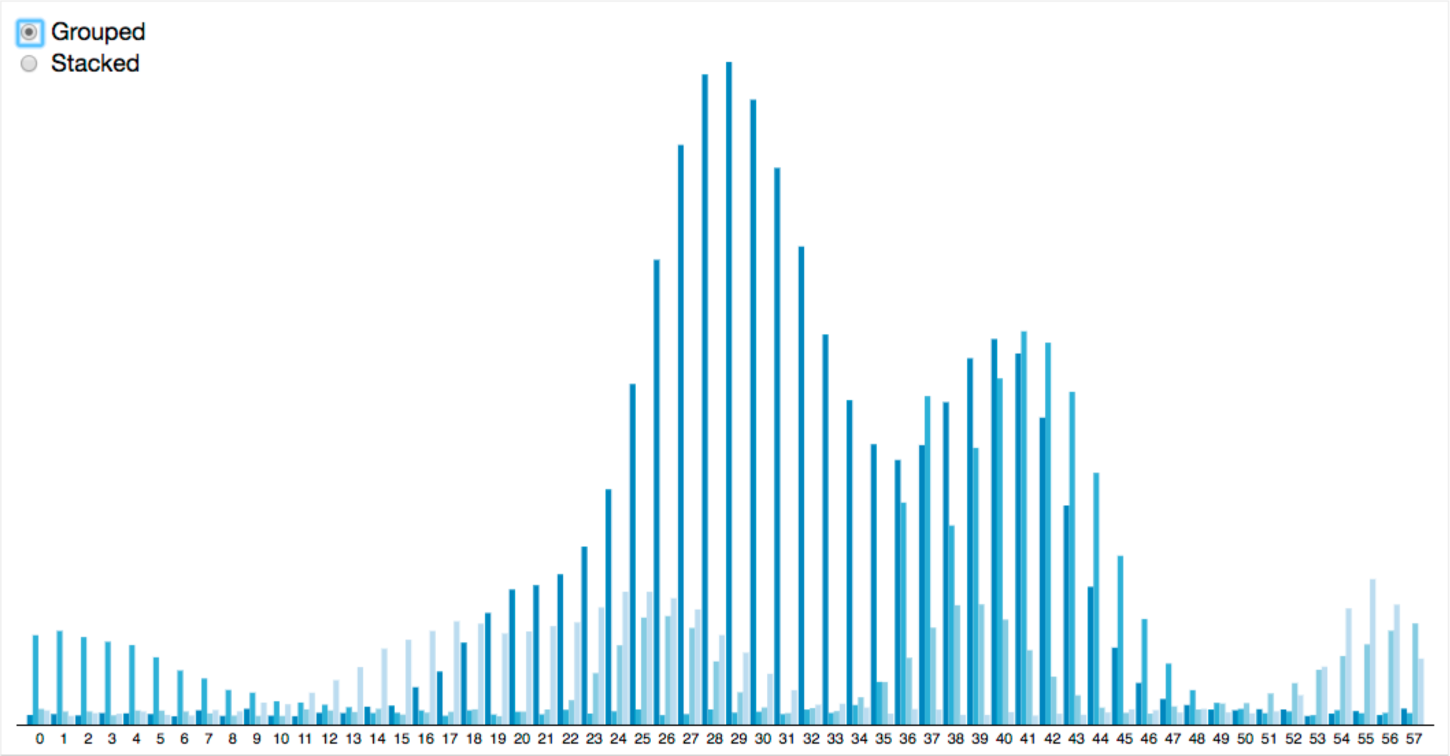


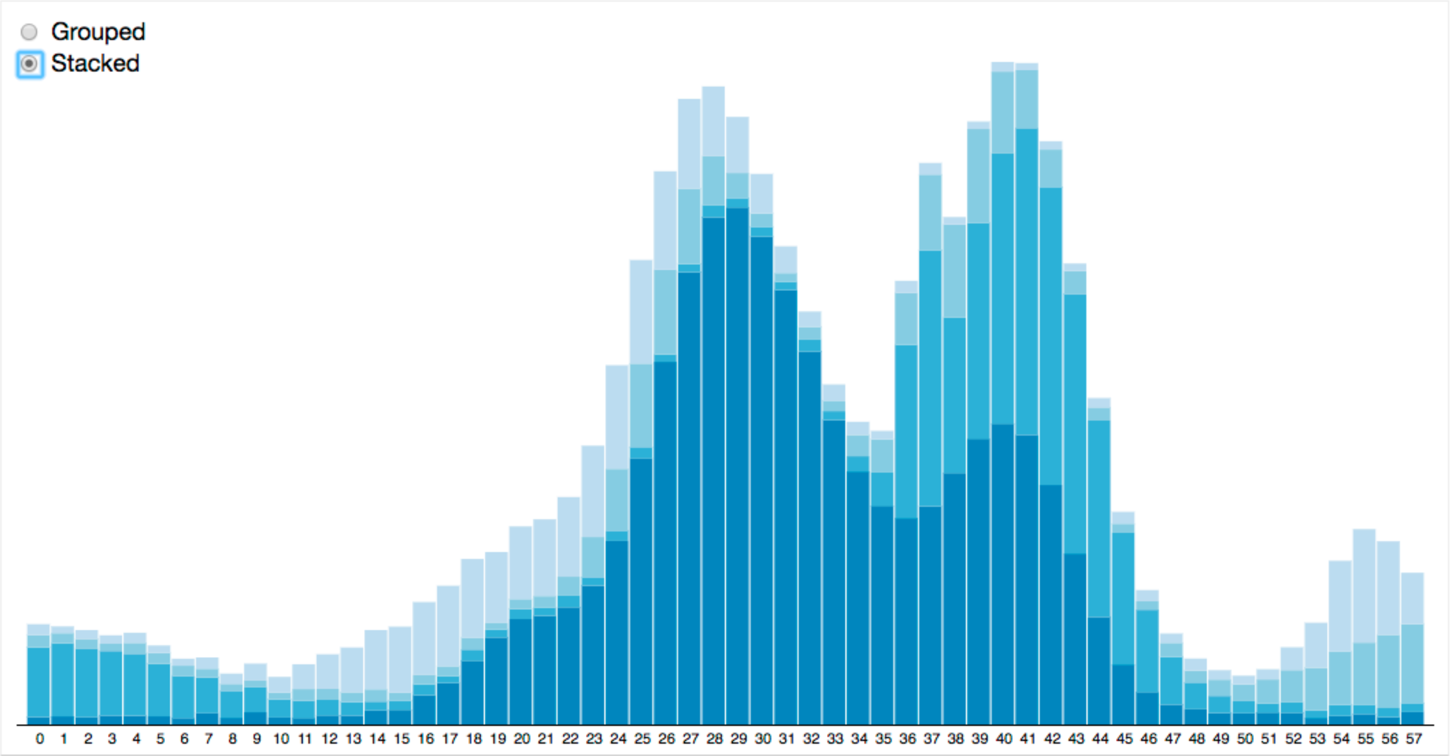
video: vimeo.com/19278444

[Animated Transitions in Statistical Data Graphics. Heer and Robertson. IEEE TVCG (Proc InfoVis 2007) 13(6):1240-1247, 2007]

Idiom: **Animated transitions - visual encoding**

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

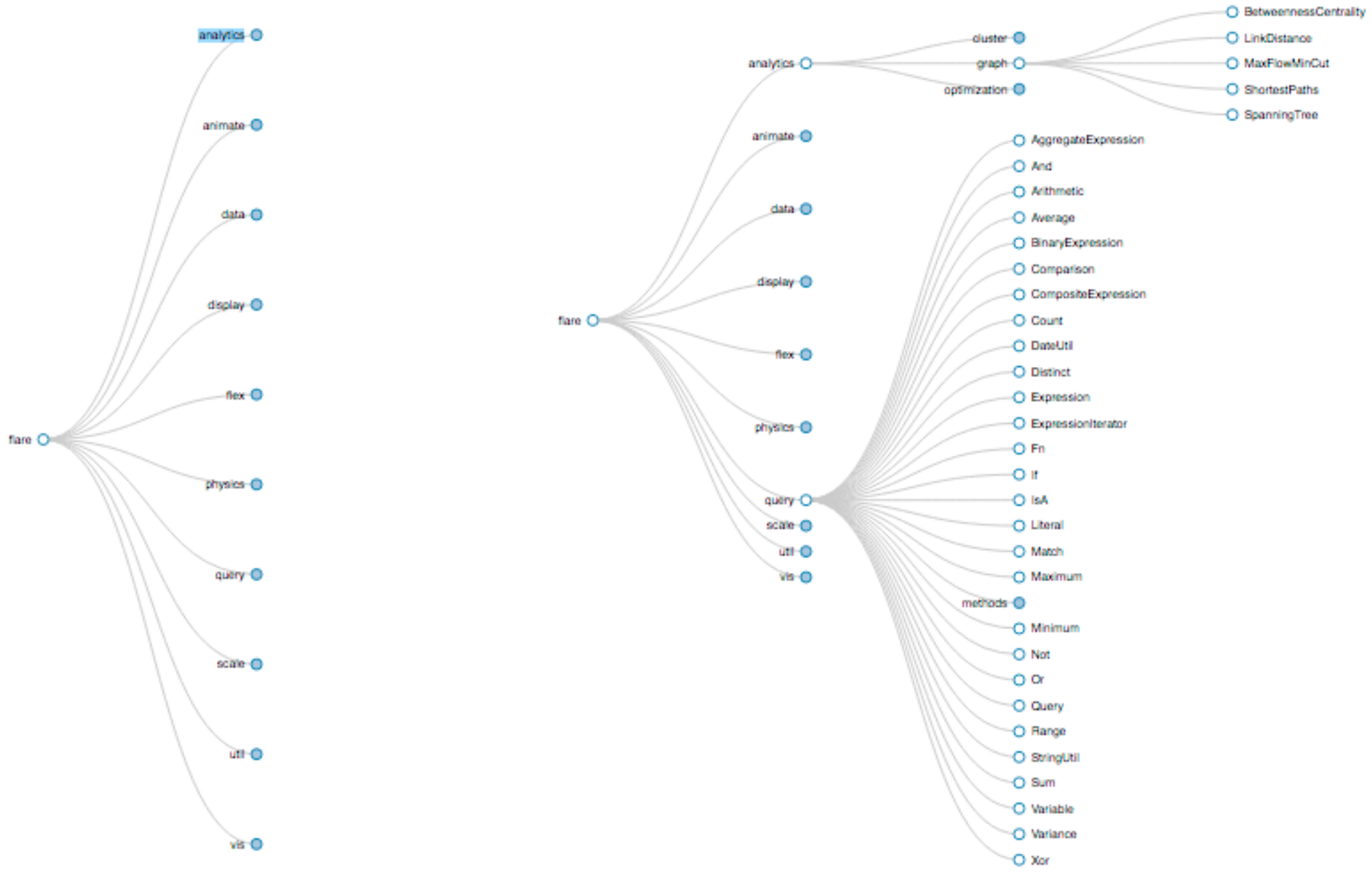




[Stacked to Grouped Bars](http://bl.ocks.org/mbostock/3943967)

Idiom: **Animated transition - tree detail**

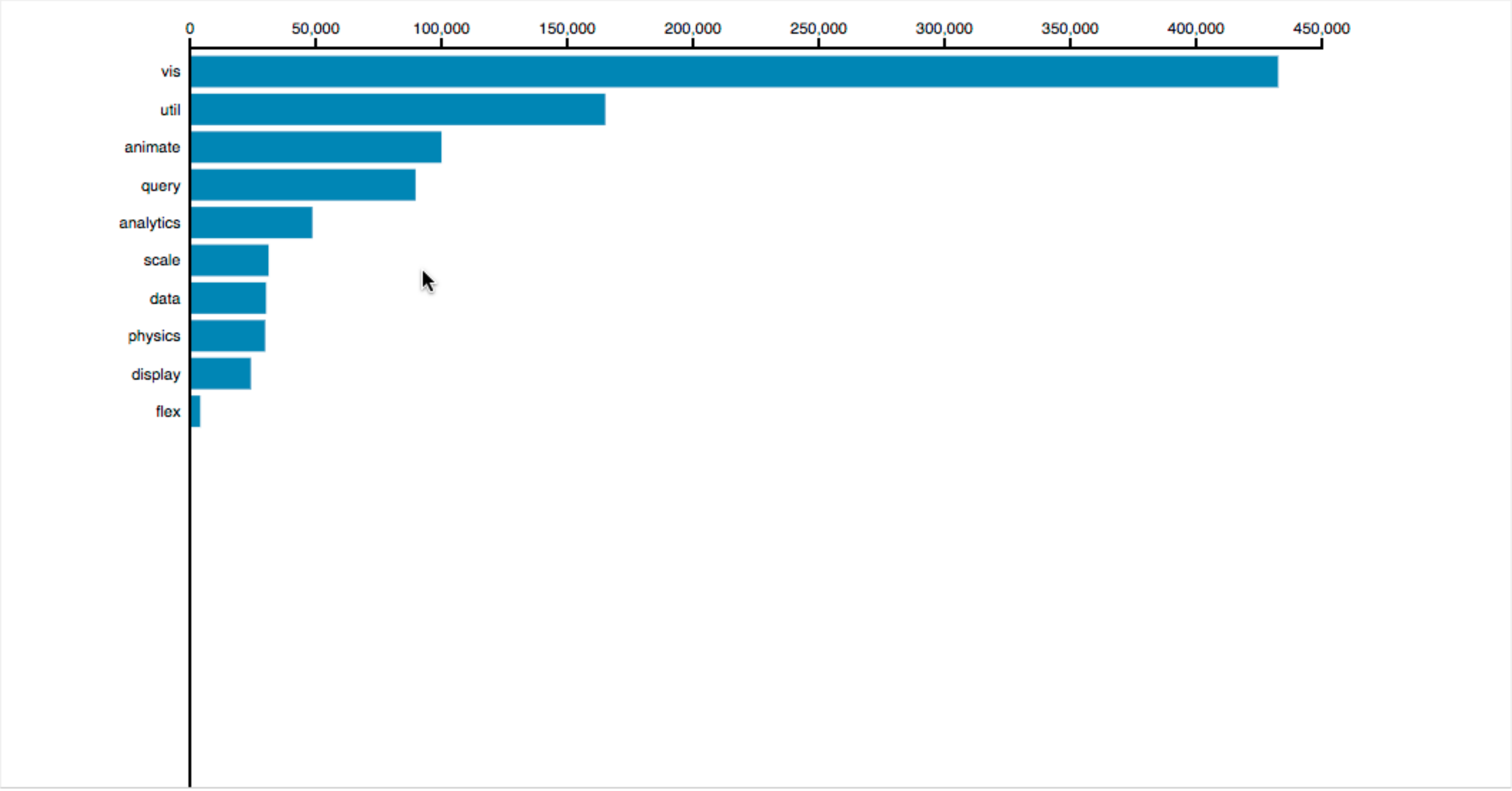
- animated transition
  - network drilldown/rollup



[Collapsible Tree](https://bl.ocks.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://bl.ocks.org/mbostock/1283663)

Interaction technology

- what do you design for?
  - mouse & keyboard on desktop?
    - large screens, hover, multiple clicks
  - touch interaction on mobile?
    - small screens, no hover, just tap
  - gestures from video / sensors?
    - ergonomic reality vs movie bombast
  - eye tracking?

Data visualization and the news - Gregor Aisch (37 min)

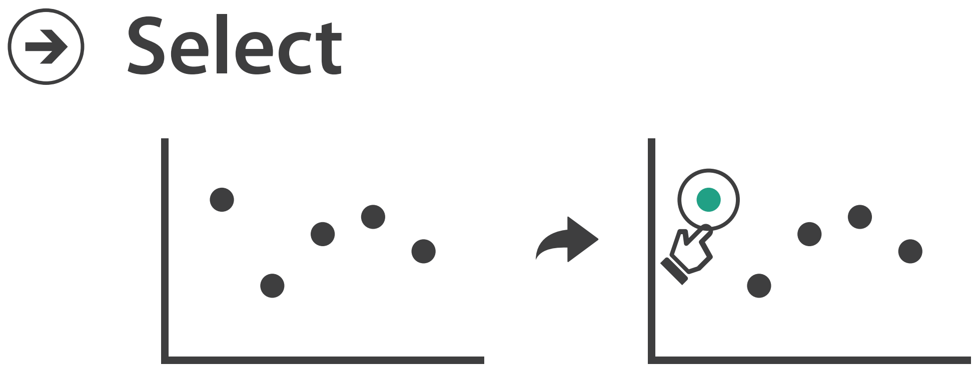
vimeo.com/182590214

I Hate Tom Cruise - Alex Kauffmann (5 min)

www.youtube.com/watch?v=QXLfT9sFcbc

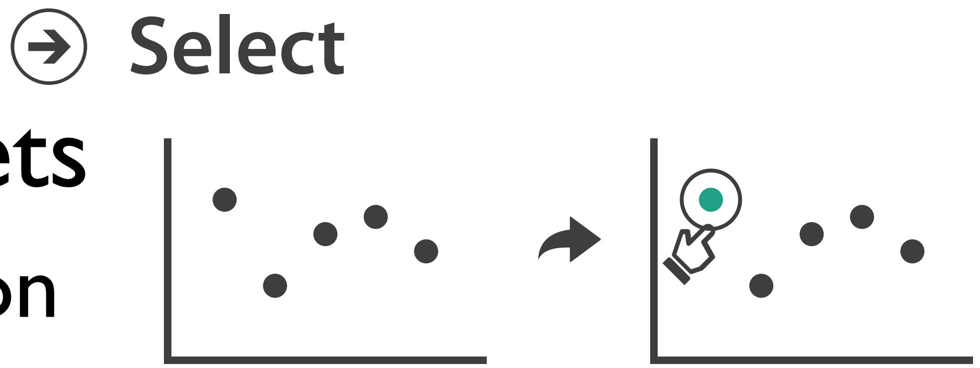
slide inspired by: Alexander Lex, Utah

Selection



- selection: basic operation for most interaction
- design choices
  - how many selection types?
    - interaction modalities
      - click/tap (heavyweight) vs hover (lightweight but not available on most touchscreens)
      - multiple click types (shift-click, option-click, …)
      - proximity beyond click/hover (touching vs nearby vs distant)
    - application semantics
      - adding to selection set vs replacing selection
      - can selection be null?
        - ex: toggle so nothing selected if click on background
      - primary vs secondary (ex: source/target nodes in network)
      - group membership (add/delete items, name group, …)

Highlighting



- highlight: change visual encoding for selection targets
  - visual feedback closely tied to but separable from selection (interaction)
- design choices: typical visual channels
  - change item color
    - but hides existing color coding
  - add outline mark
  - change size (ex: increase outline mark linewidth)
  - change shape (ex: from solid to dashed line for link mark)
- unusual channels: motion
  - motion: usually avoid for single view
    - with multiple views, could justify to draw attention to other views

Tooltips

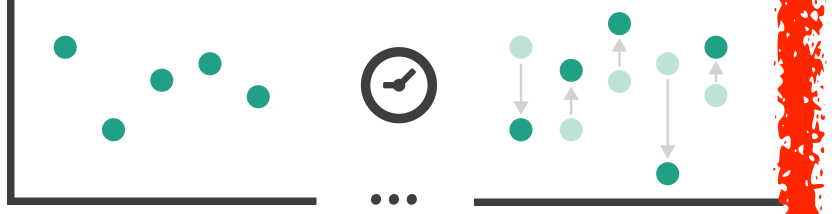
- popup information for selection
  - hover or click
  - can provide useful additional detail on demand
  - beware: does not support overview!
    - always consider if there’s a way to visually encode directly to provide overview
    - “If you make a rollover or tooltip, assume nobody will see it. If it's important, make it explicit.“
      - Gregor Aisch, NYTimes

Rule of thumb: **Responsiveness is required**

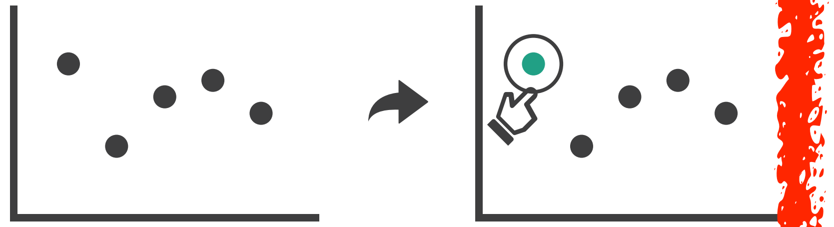
- *visual feedback: three rough categories*
  - *0.1 seconds: perceptual processing*
    - subsecond response for mouseover highlighting - ballistic motion
  - *1 second: immediate response*
    - fast response after mouseclick, button press - Fitts' Law limits on motor control
  - *10 seconds: brief tasks*
    - bounded response after dialog box - mental model of heavyweight operation (file load)
- scalability considerations
  - highlight selection without complete redraw of view (graphics frontbuffer)
  - show hourglass for multi-second operations (check for cancel/undo)
  - show progress bar for long operations (process in background thread)
  - rendering speed when item count is large (guaranteed frame rate)

Manipulate

➔ Change over Time



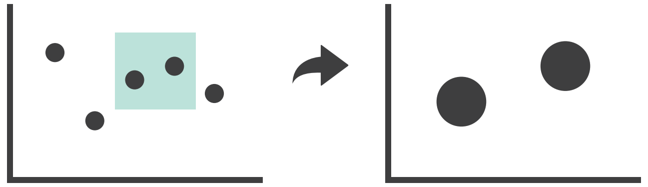
➔ Select



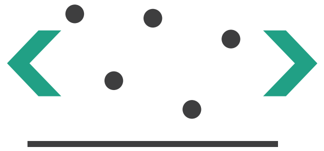
➔ Navigate

➔ Item Reduction

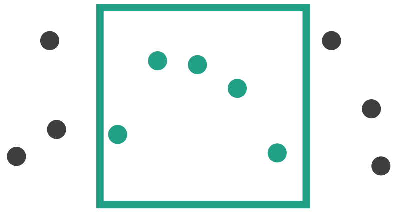
➔ *Zoom*
Geometric or *Semantic*



➔ *Pan/Translate*



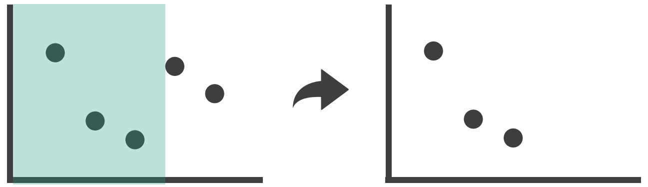
➔ *Constrained*



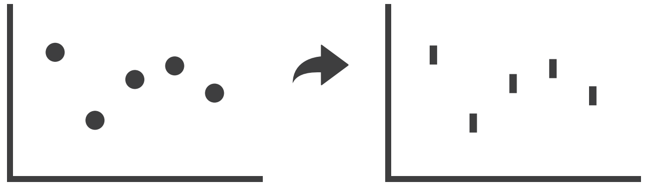
➔ Attribute Reduction

➔ *Slice*

➔ *Cut*



➔ *Project*



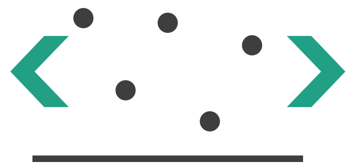
Navigate: Changing viewpoint/visibility

- change viewpoint
  - changes which items are visible within view
- camera metaphor
  - pan/translate/scroll
    - move up/down/sideways

➔ **Navigate**

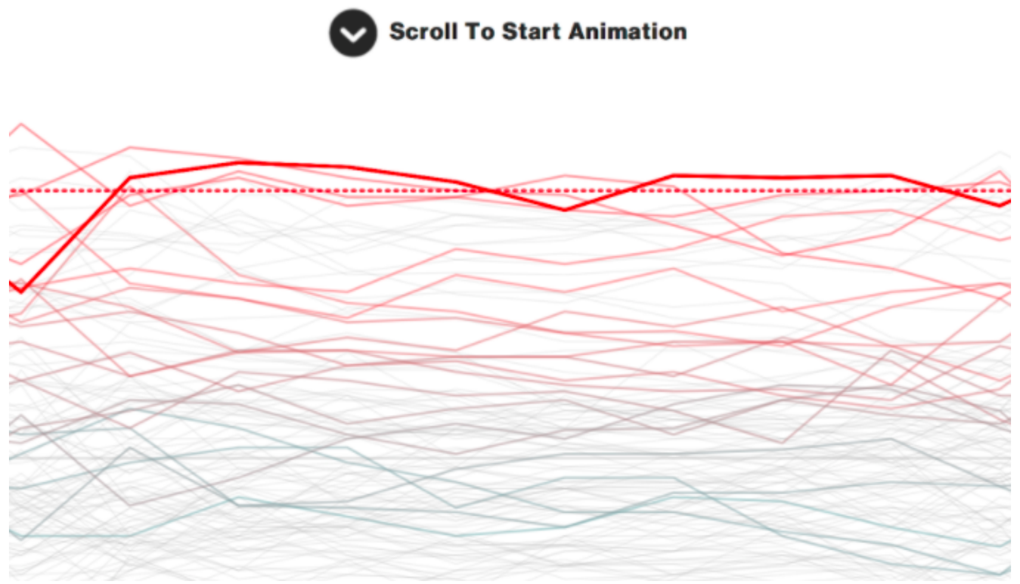
➔ Item Reduction

➔ *Pan/Translate*



Idiom: **Scrollytelling**

- how: navigate page by scrolling (panning down)
- pros:
  - familiar & intuitive, from standard web browsing
  - linear (only up & down) vs possible overload of click-based interface choices
- cons:
  - full-screen mode may lack affordances
  - scrolljacking, no direct access
  - unexpected behaviour
  - continuous control for discrete steps

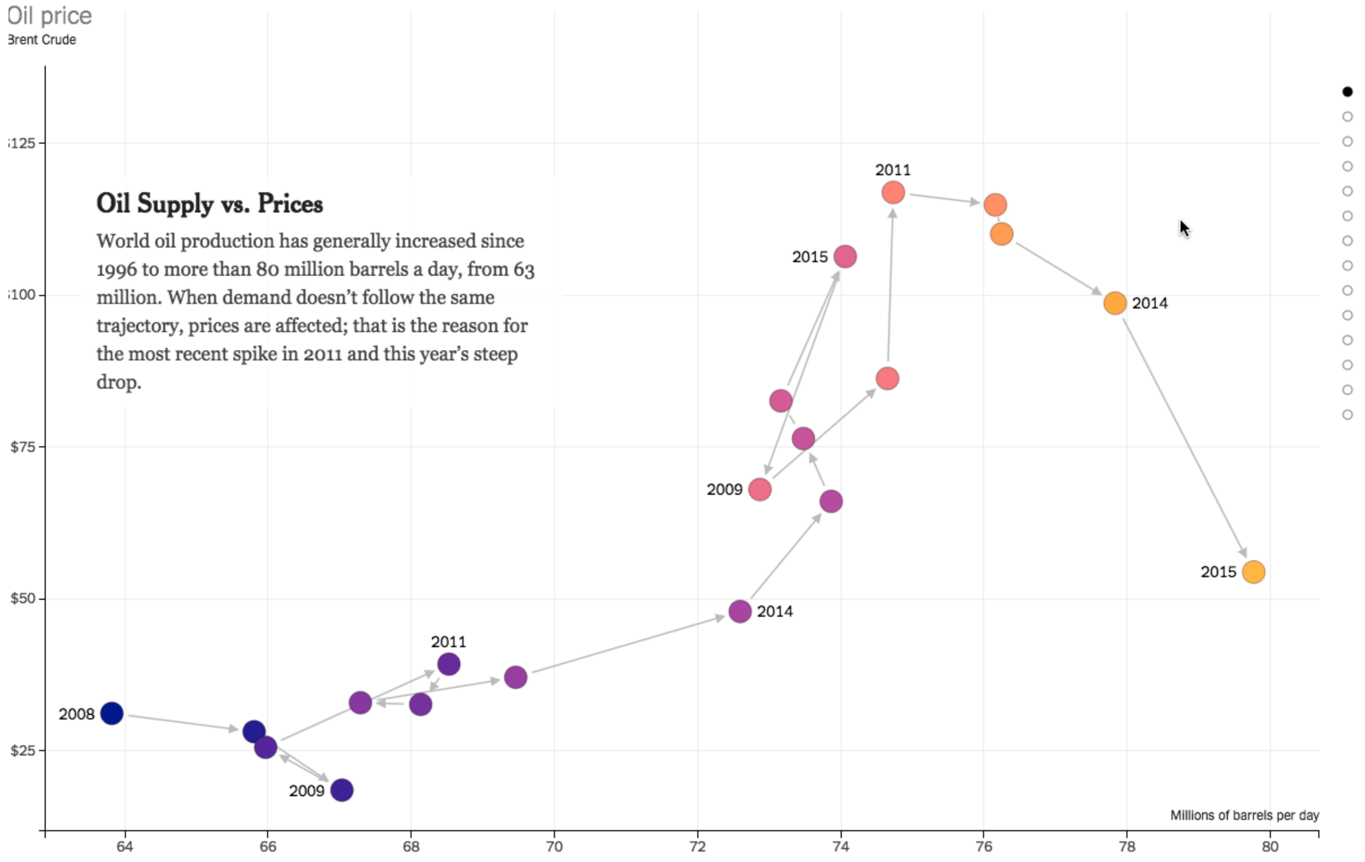


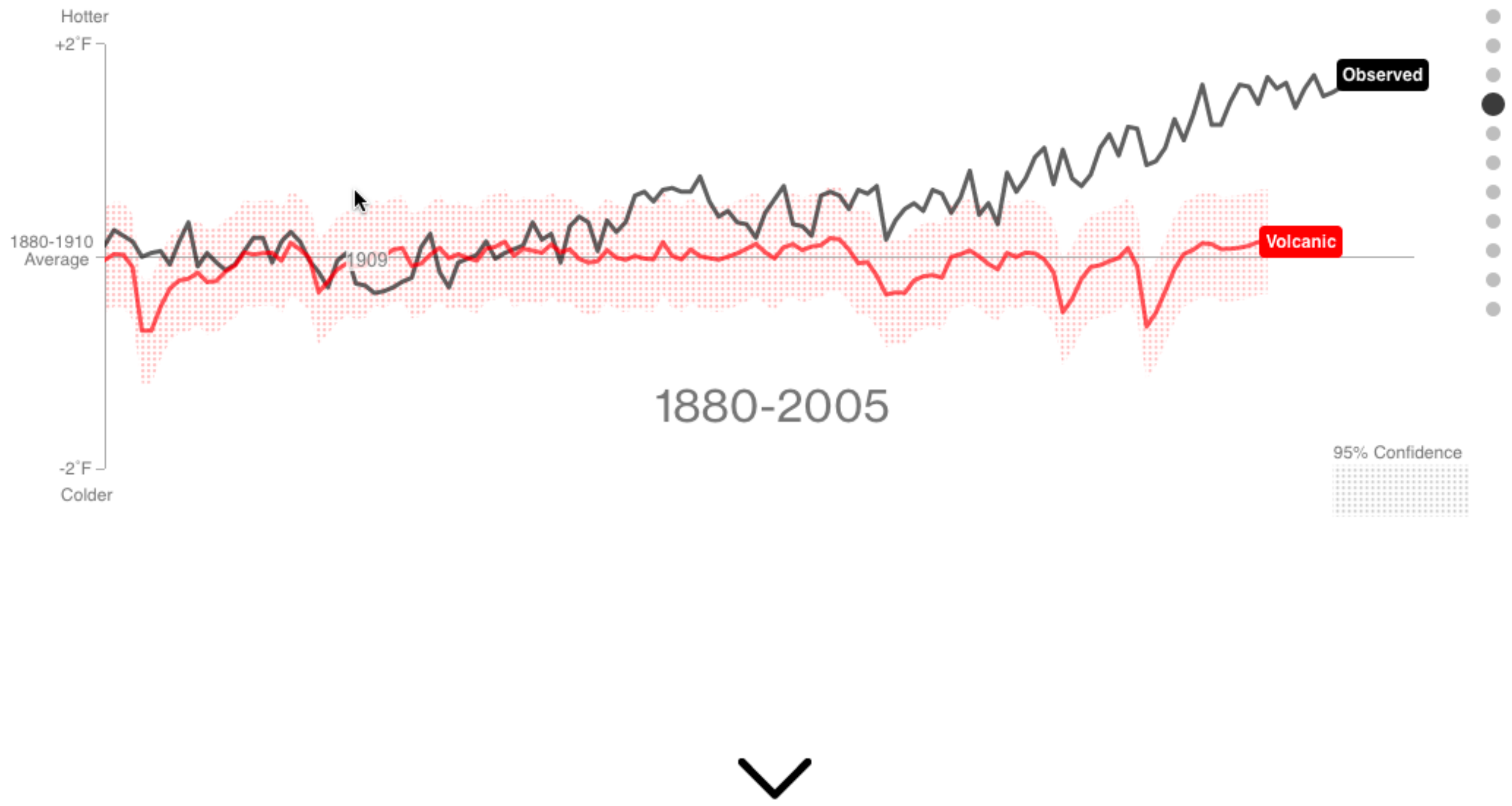
https://eagereyes.org/blog/2016/the-scrollytelling-scourge

[How to Scroll, Bostock](https://bost.ocks.org/mike/scroll/)

slide inspired by: Alexander Lex, Utah

Scrollytelling examples





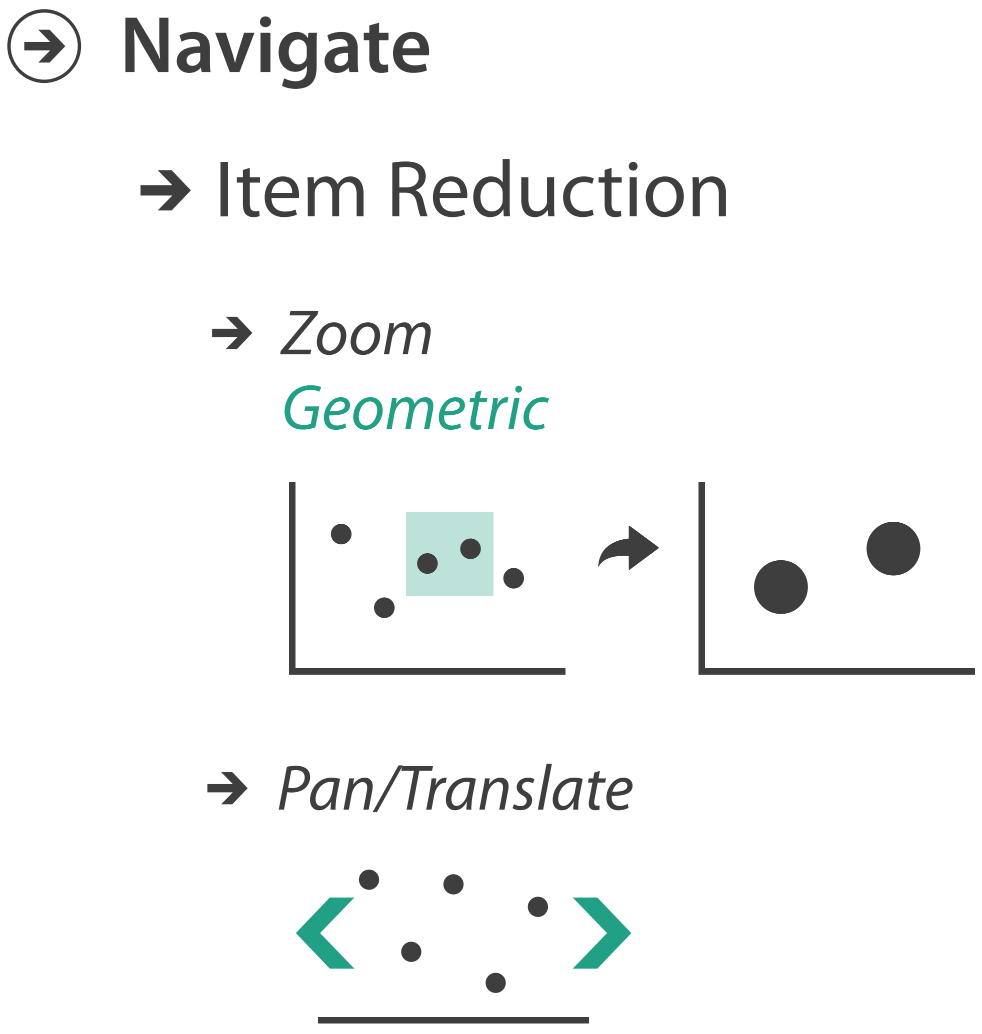
https://www.nytimes.com/interactive/2015/09/30/business/how-the-us-and-opec-drive-oil-prices.html?_r=1

https://www.bloomberg.com/graphics/2015-whats-warming-the-world/

slide inspired by: Alexander Lex, Utah

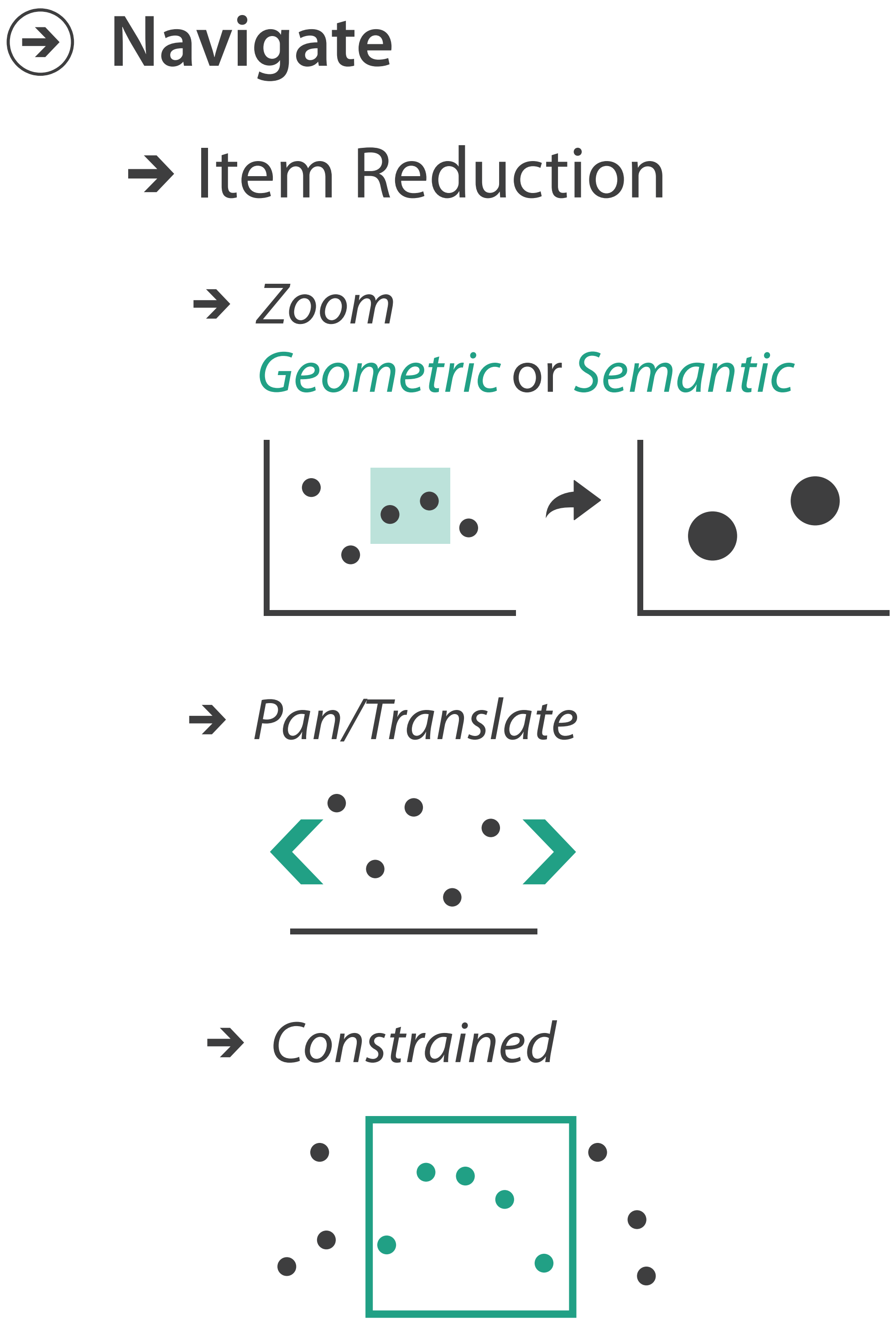
Navigate: Changing viewpoint/visibility

- change viewpoint
  - changes which items are visible within view
- camera metaphor
  - pan/translate/scroll
    - move up/down/sideways
  - rotate/spin
    - typically in 3D
  - zoom in/out
    - enlarge/shrink world == move camera closer/further
    - geometric zoom: standard, like moving physical object



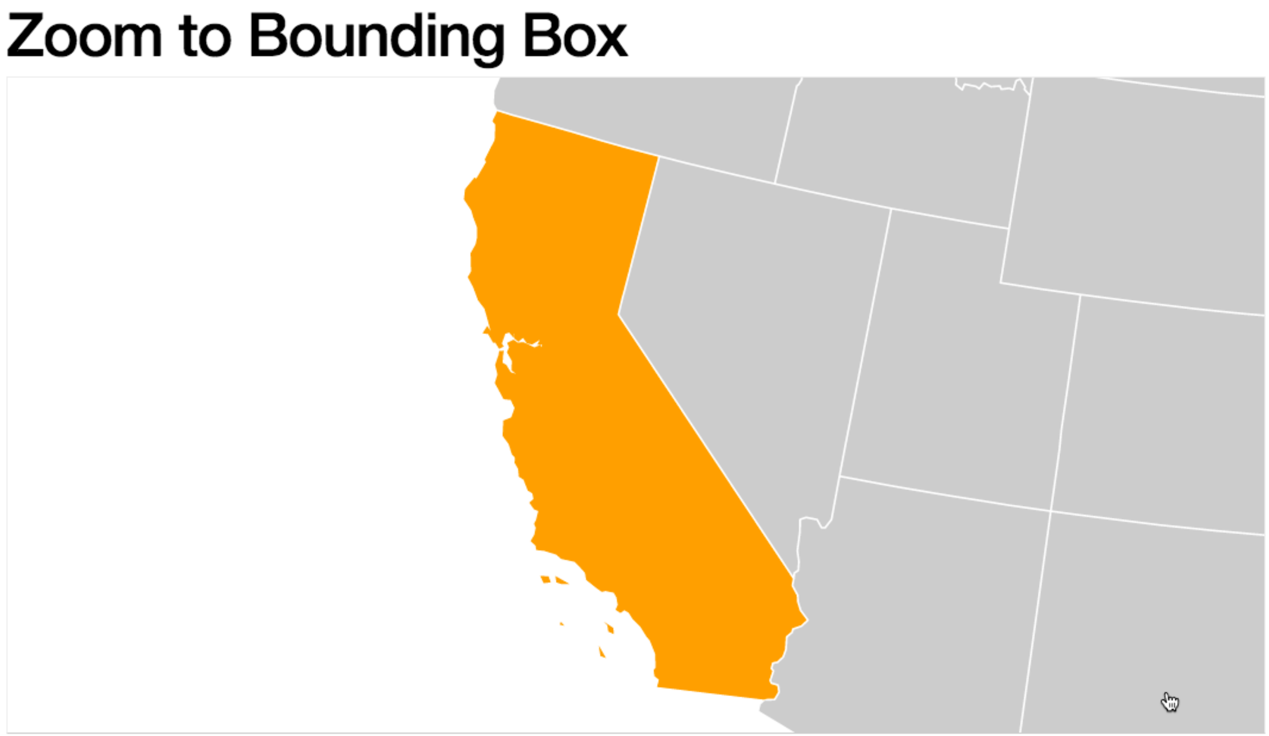
Navigate: Unconstrained vs constrained

- unconstrained navigation
  - easy to implement for designer
  - hard to control for user
    - easy to overshoot/undershoot
- constrained navigation
  - typically uses animated transitions
  - trajectory automatically computed based on selection
    - just click; selection ends up framed nicely in final viewport



Idiom: **Animated transition + constrained navigation**

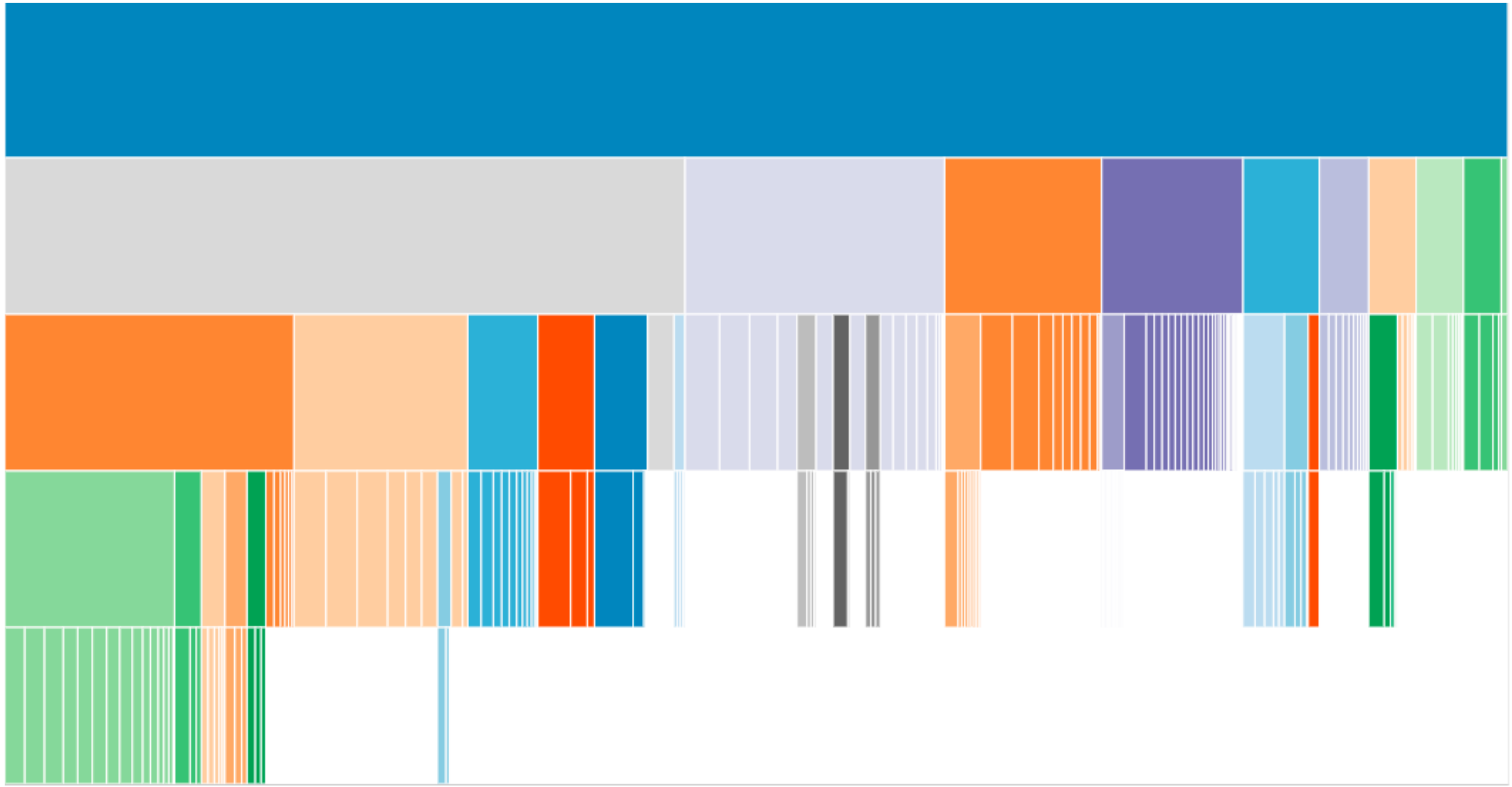
- example: geographic map
  - simple zoom, only viewport changes, shapes preserved



[Zoom to Bounding Box](https://bl.ocks.org/mbostock/4699541)

Idiom: **Animated transition + constrained navigation**

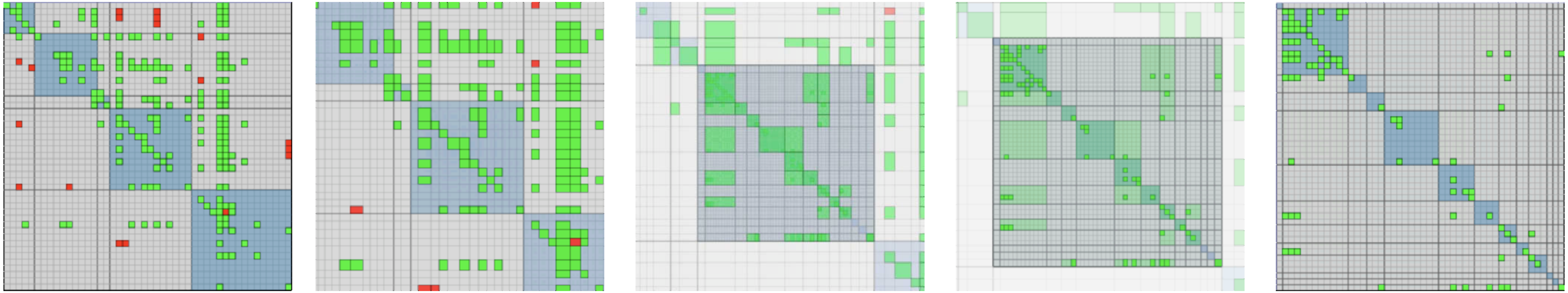
- example: icicle plot
  - transition into containing mark causes aspect ratio (shape) change



[Zoomable Icicle](https://bl.ocks.org/mbostock/1005873)

Idiom: **Animated transition + constrained navigation**

- example: multilevel matrix views
  - add detail during transition
  - movie: http://www.win.tue.nl/vis1/home/fvham/matrix/Zoomin.avi
  - movie: http://www.win.tue.nl/vis1/home/fvham/matrix/Zoomout.avi
  - movie: http://www.win.tue.nl/vis1/home/fvham/matrix/Pan.avi

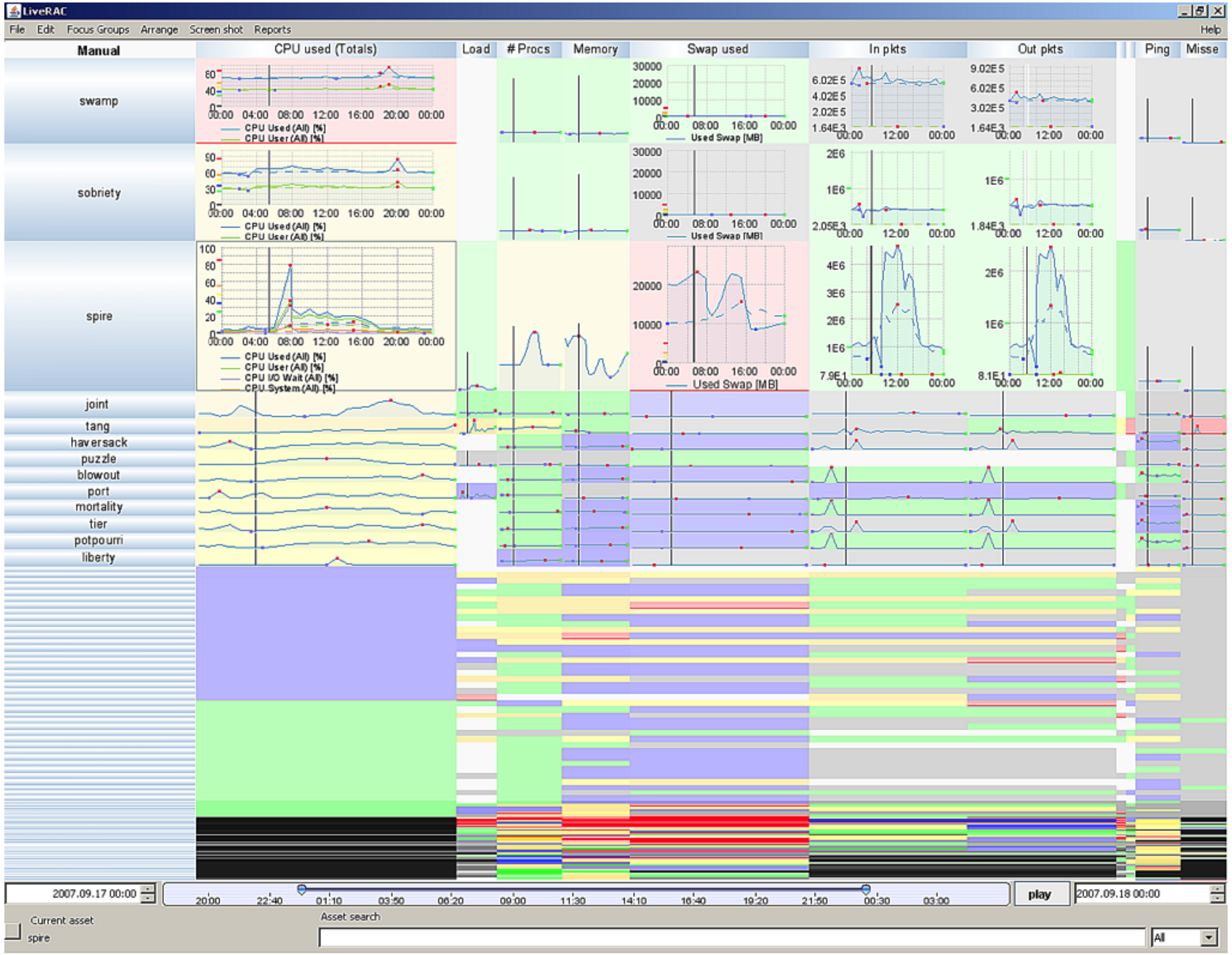


[Using Multilevel Call Matrices in Large Software Projects. van Ham. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 227–232, 2003.]

Idiom: **Semantic zooming**

System: **LiveRAC**

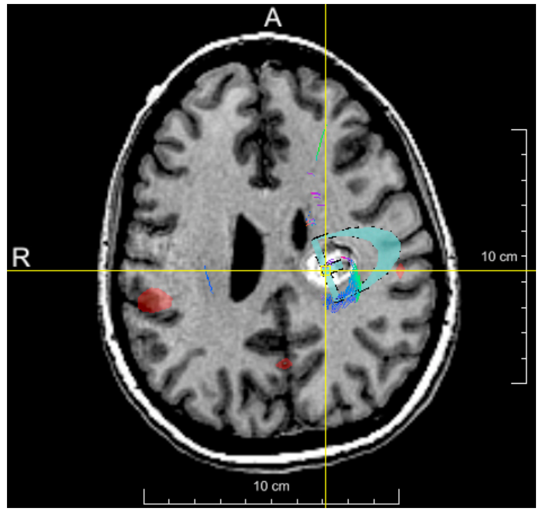
- semantic zoom
  - alternative to geometric zoom
  - resolution-aware layout adapts to available space
  - goal: legible at multiple scales
  - dramatic or subtle effects
- visual encoding change
  - colored box
  - sparkline
  - simple line chart
  - full chart: axes and tickmarks

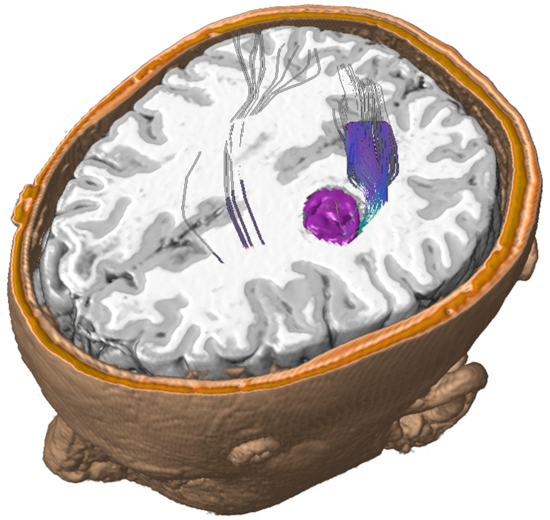


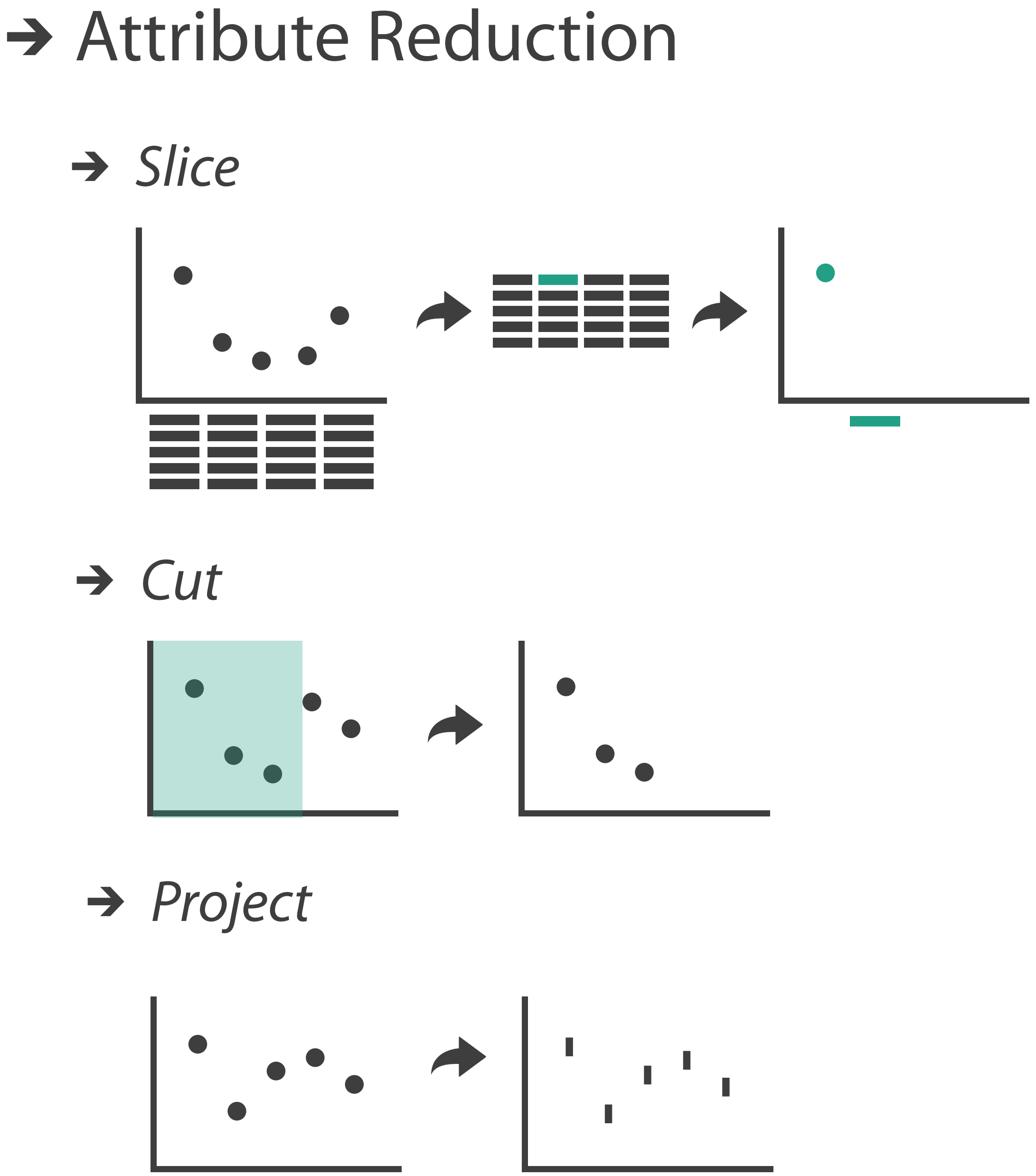
[LiveRAC - Interactive Visual Exploration of System Management Time-Series Data. McLachlan, Munzner, Koutsofios, and North. Proc. ACM Conf. Human Factors in Computing Systems (CHI), pp. 1483–1492, 2008.]

Navigate: Reducing attributes

- continuation of camera metaphor
  - slice
    - show only items matching specific value for given attribute: slicing plane
    - axis aligned, or arbitrary alignment
  - cut
    - show only items on far slide of plane from camera
  - project
    - change mathematics of image creation
      - orthographic (eliminate 3rd dimension)
      - perspective (foreshortening captures limited 3D information)





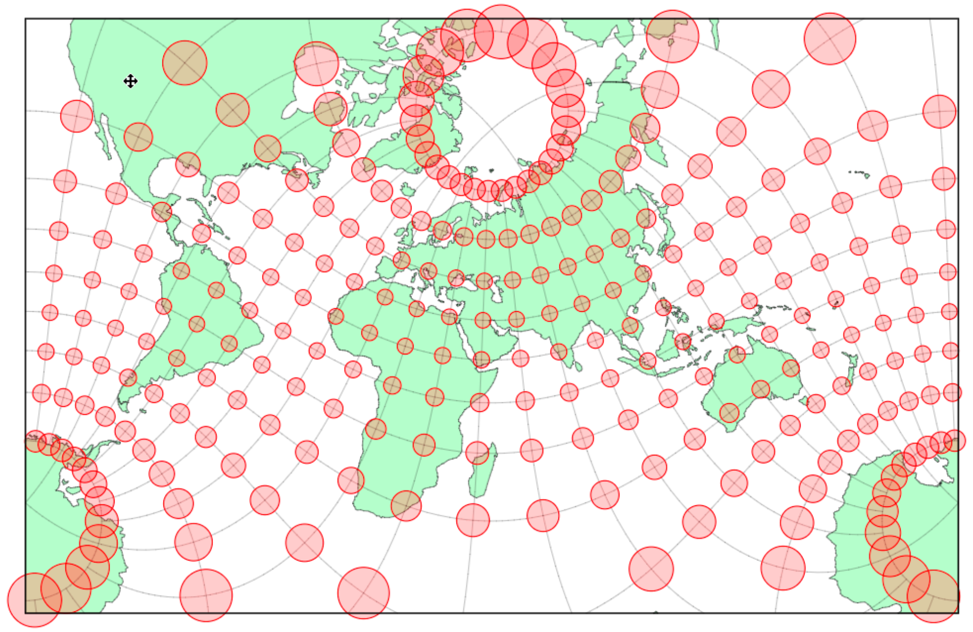


[Interactive Visualization of Multimodal Volume Data for Neurosurgical Tumor Treatment. Rieder, Ritter, Raspe, and Peitgen. Computer Graphics Forum (Proc. EuroVis 2008) 27:3 (2008), 1055–1062.]

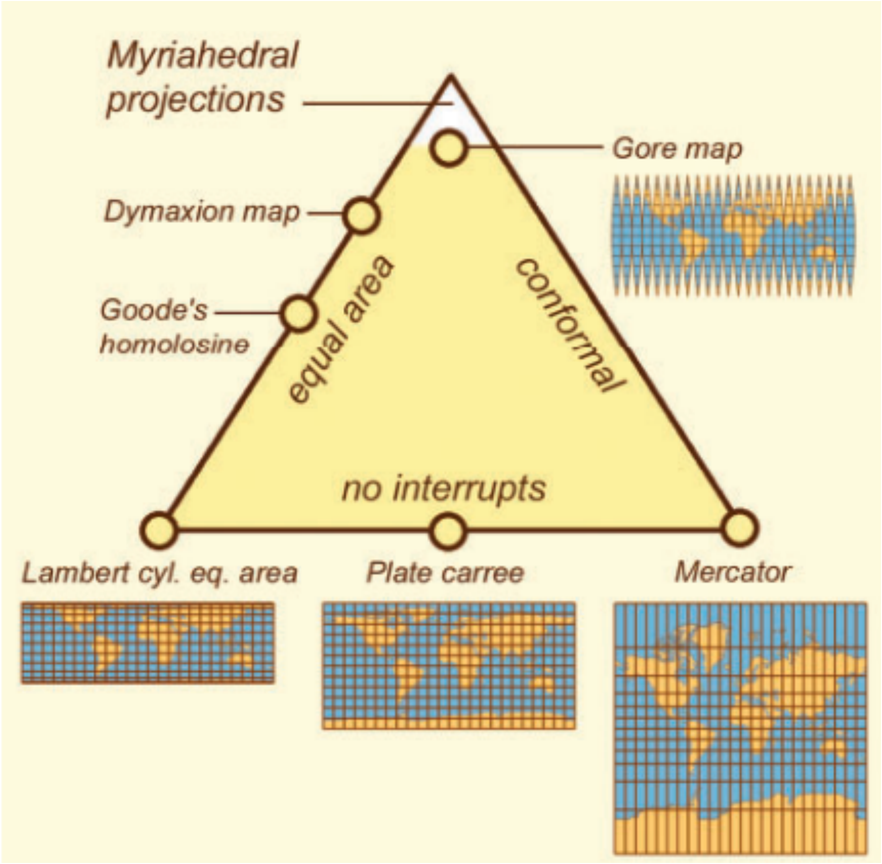
Navigate: Cartographic projections

- project from 2D sphere surface to 2D plane
  - can only fully preserve 2 out of 3
    - angles: conformal
    - area: equal area
    - contiguity: no interruptions

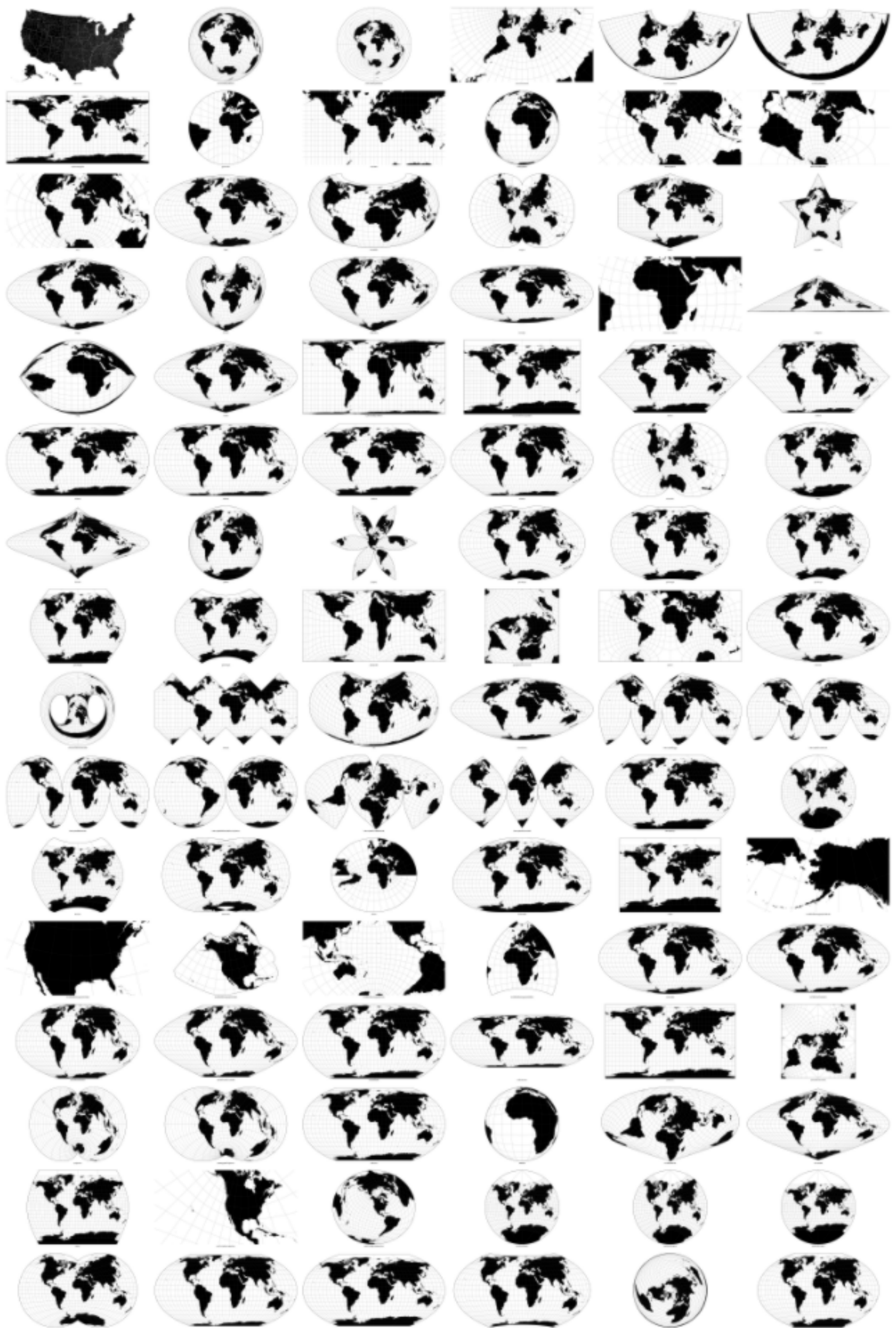
Tissot's Indicatrix



https://www.jasondavies.com/maps/tissot/



https://www.win.tue.nl/~vanwijk/myriahedral/



[Every Map Projection](https://bl.ocks.org/mbostock/29cddc0006f8b98eff12e60dd08f59a7)

Interaction benefits

- interaction pros
  - major advantage of computer-based vs paper-based visualization
  - flexible, powerful, intuitive
    - exploratory data analysis: change as you go during analysis process
    - fluid task switching: different visual encodings support different tasks
  - animated transitions provide excellent support
    - empirical evidence that animated transitions help people stay oriented

Interaction limitations

- interaction has a time cost
  - sometimes minor, sometimes significant
  - degenerates to human-powered search in worst case
- remembering previous state imposes cognitive load
  - *rule of thumb: eyes over memory*
    - *hard to compare visible item to memory of what you saw*
    - ex: maintaining context/orientation when navigating
    - ex: tracking complex changes during animation
- controls may take screen real estate
  - or invisible functionality may be difficult to discover (lack of affordances)
- users may not interact as planned by designer
  - NYTimes logs show ~90% don't interact beyond scrollytelling - Aisch, 2016

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.

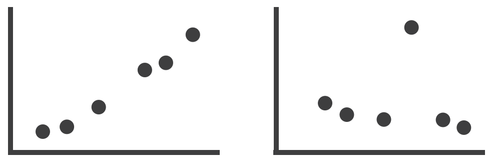
Further reading, full

- Starting Simple - Adding Value to Static Visualisation Through Simple Interaction.. A. Dix and G. Ellis. Proc. Advanced Visual Interfaces (AVI) 1998, 124-134.
- Animated Transitions in Statistical Data Graphics Jeffrey Heer and George G. Robertson. IEEE TVCG (Proc. InfoVis 2007) 13(6): 1240-1247, 2007.[Archived version]
- Selection: 524,288 Ways To Say 'This Is Interesting'. Graham J. Wills. Proc. InfoVis 1996, p 54-61.
- Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics Ben Bederson, and James D Hollan, Proc UIST 94.
- LiveRAC - Interactive Visual Exploration of System Management Time-Series Data. Peter McLachlan, Tamara Munzner, Eleftherios Koutsofios, Stephen North. Proc. Conf. on Human Factors in Computing Systems (CHI) 2008, 1483-1492.
- Rapid Controlled Movement Through a Virtual 3D Workspace Jock Mackinlay, Stuart Card, and George Robertson. Proc SIGGRAPH '90, pp 171-176.
- Smooth and Efficient Zooming and Panning. Jack J. van Wijk and Wim A.A. Nuij, Proc. InfoVis 2003, p. 15-22.

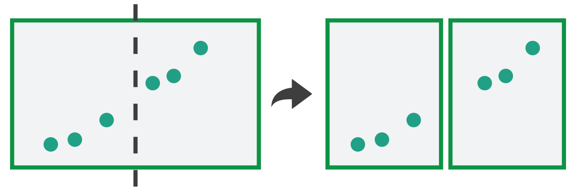
Ch 12: Facet into Multiple Views

Facet

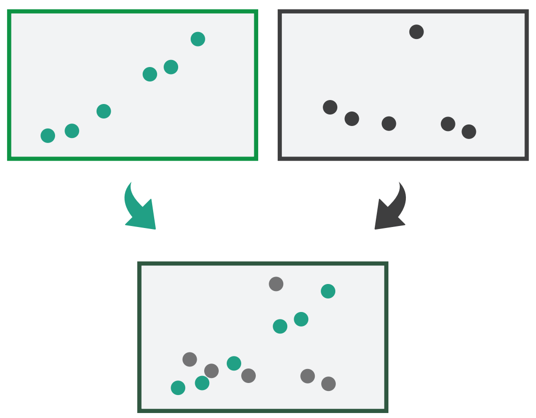
Juxtapose



Partition

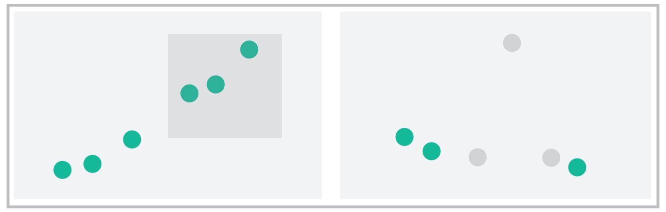


Superimpose

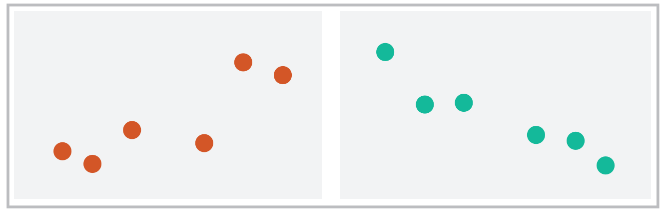


Juxtapose and coordinate views

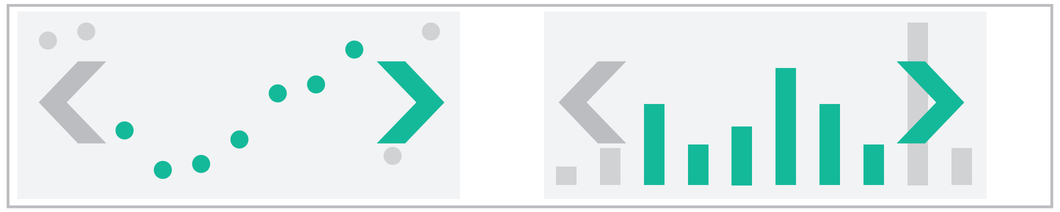
- Share Encoding: Same/Different
  - *Linked Highlighting*



- Share Data: All/Subset/None

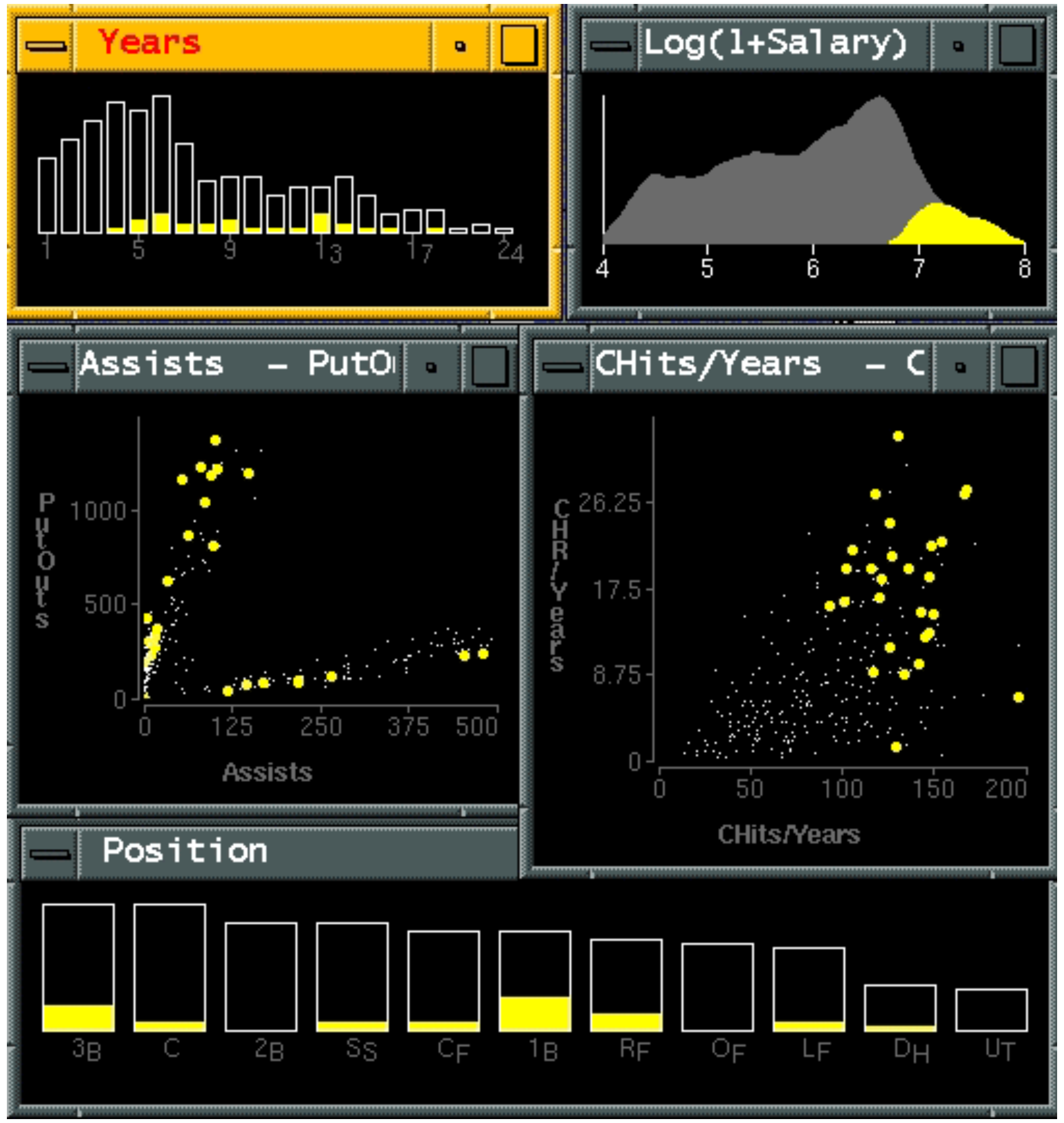


- Share Navigation



Idiom: **Linked highlighting** System: **EDV**

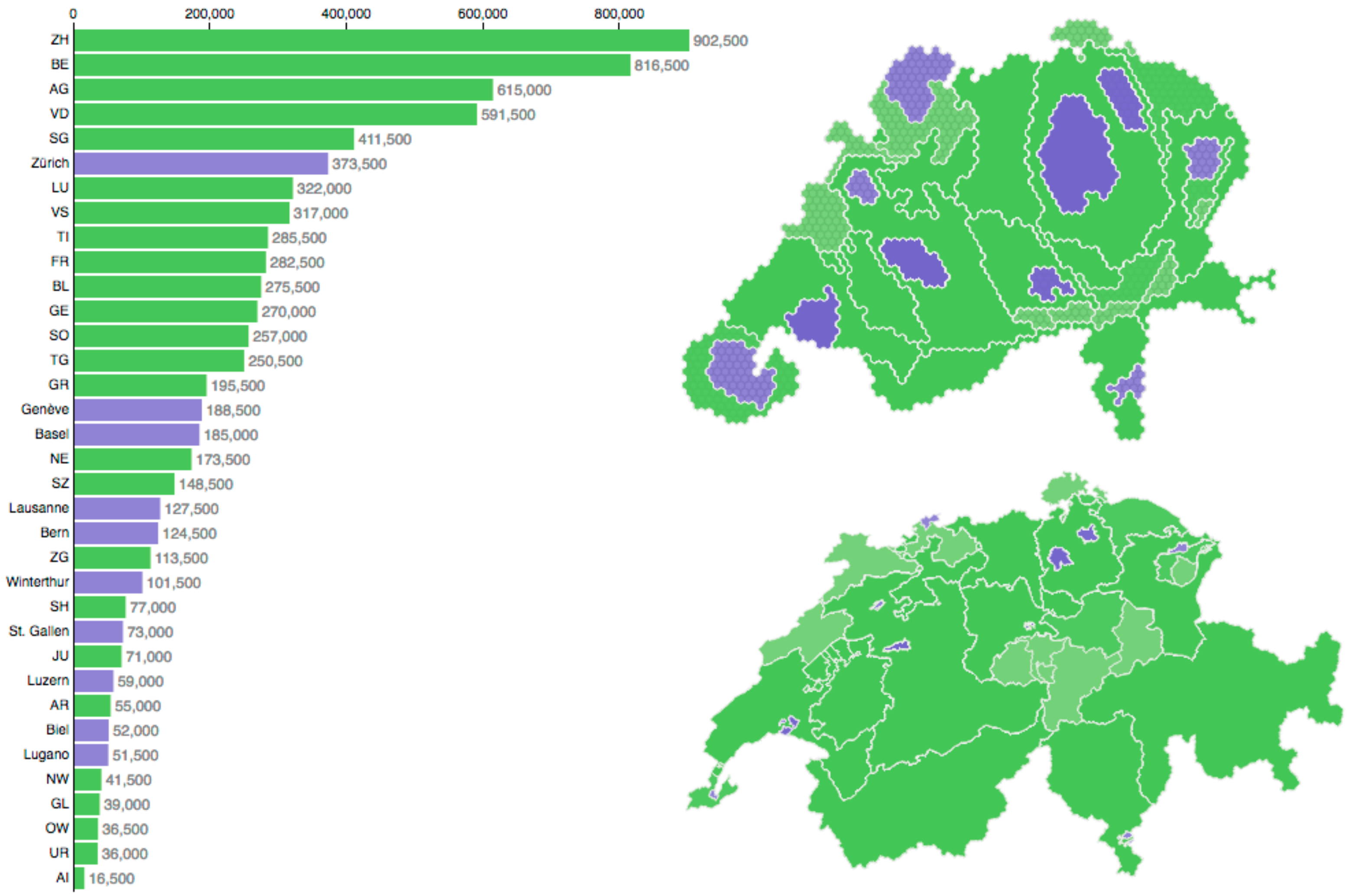
- see how regions contiguous in one view are distributed within another
  - powerful and pervasive interaction idiom
- encoding: different
  - ***multiform***
- data: all shared
- aka: brushing and linking

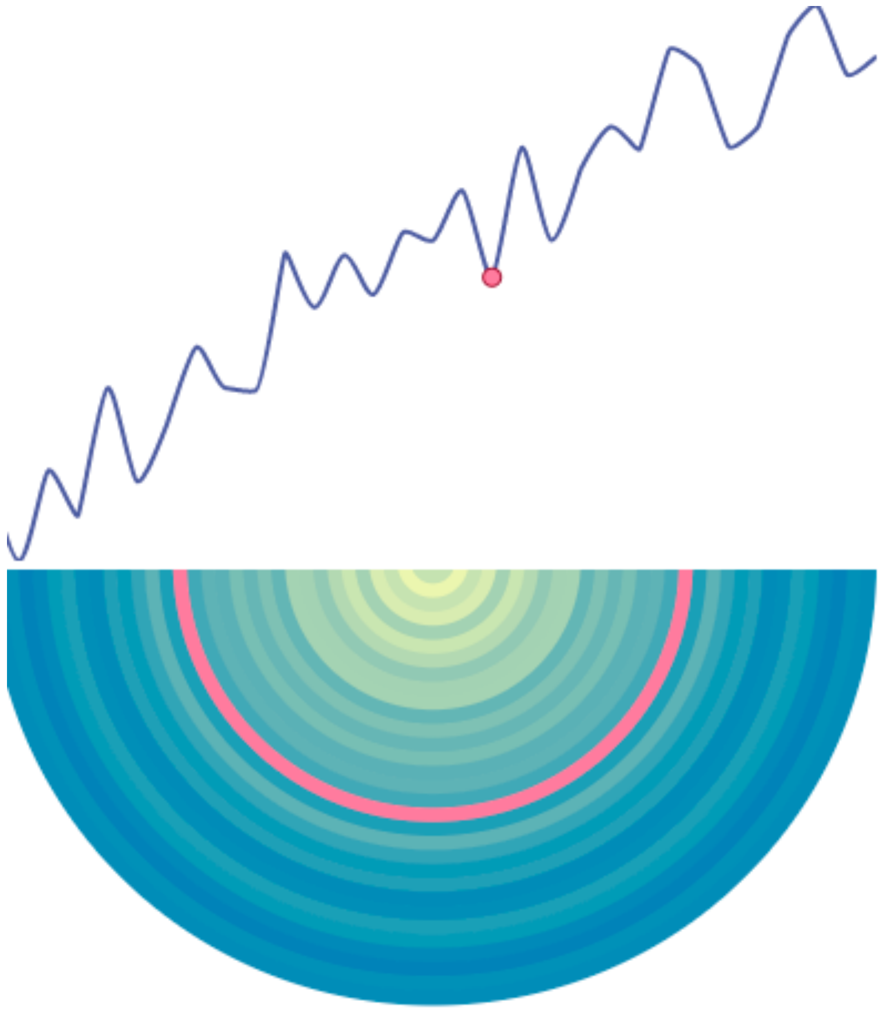


[Visual Exploration of Large Structured Datasets. Wills. Proc. New Techniques and Trends in Statistics (NTTS), pp. 237–246. IOS Press, 1995.]

Linked views

- unidirectional vs bidirectional linking



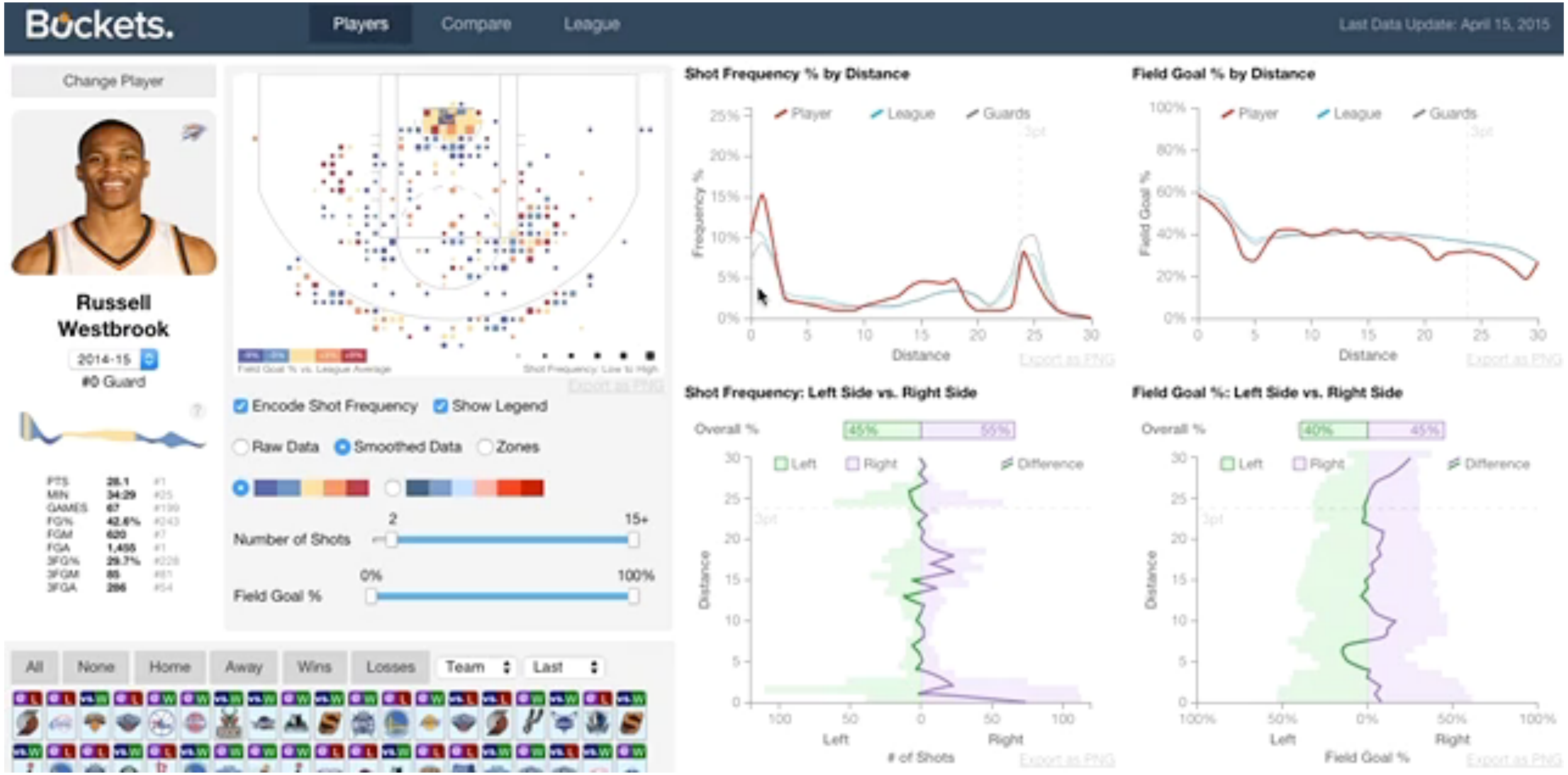


http://www.ralphstraumann.ch/projects/swiss-population-cartogram/

http://peterbeshai.com/linked-highlighting-react-d3-reflux/

Linked views: Multidirectional linking

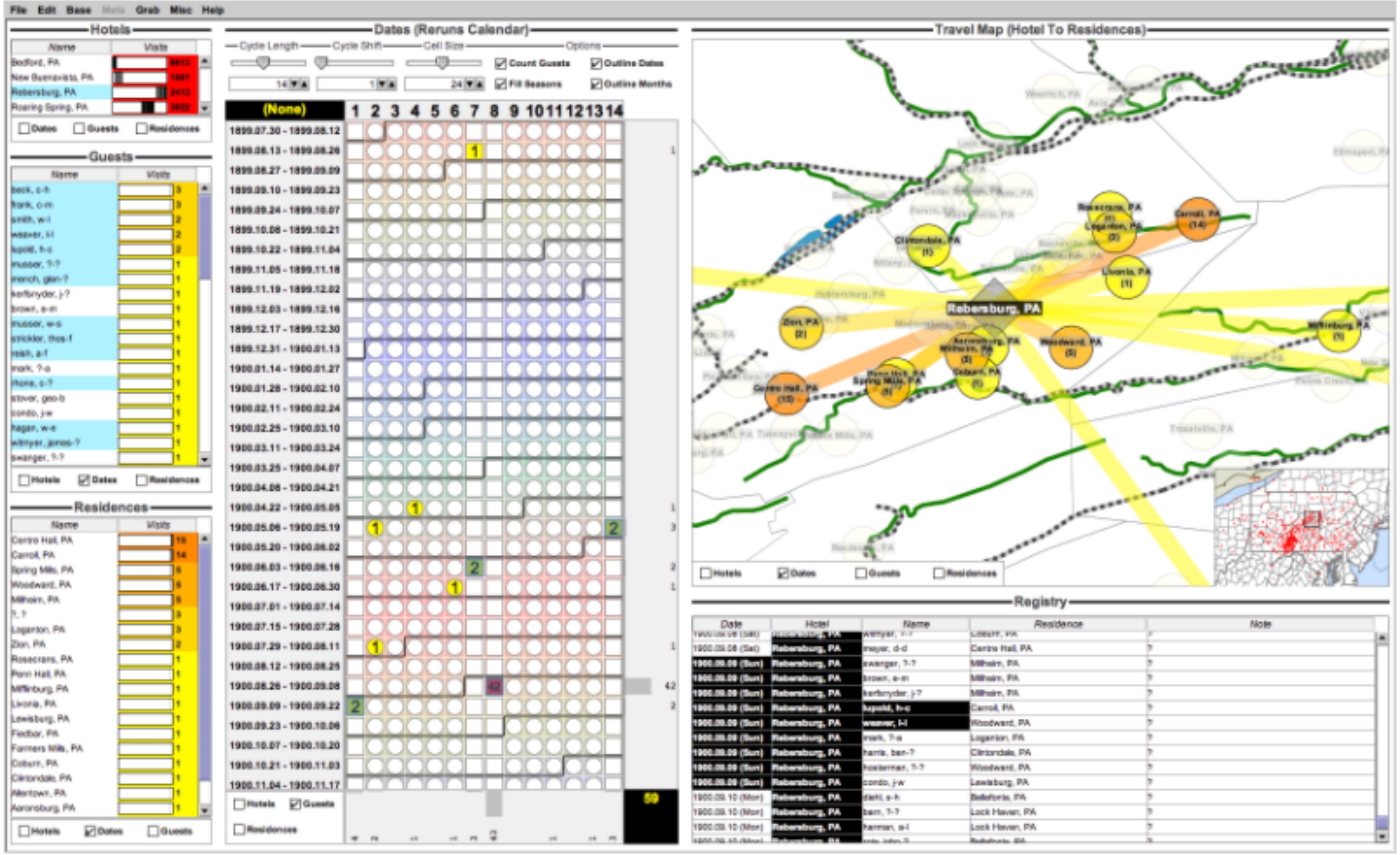
System: **Buckets**



http://buckets.peterbeshai.com/

https://medium.com/@pbesh/linked-highlighting-with-react-d3-js-and-reflux-16e9c0b2210b

Video: Visual Analysis of Historical Hotel Visitation Patterns

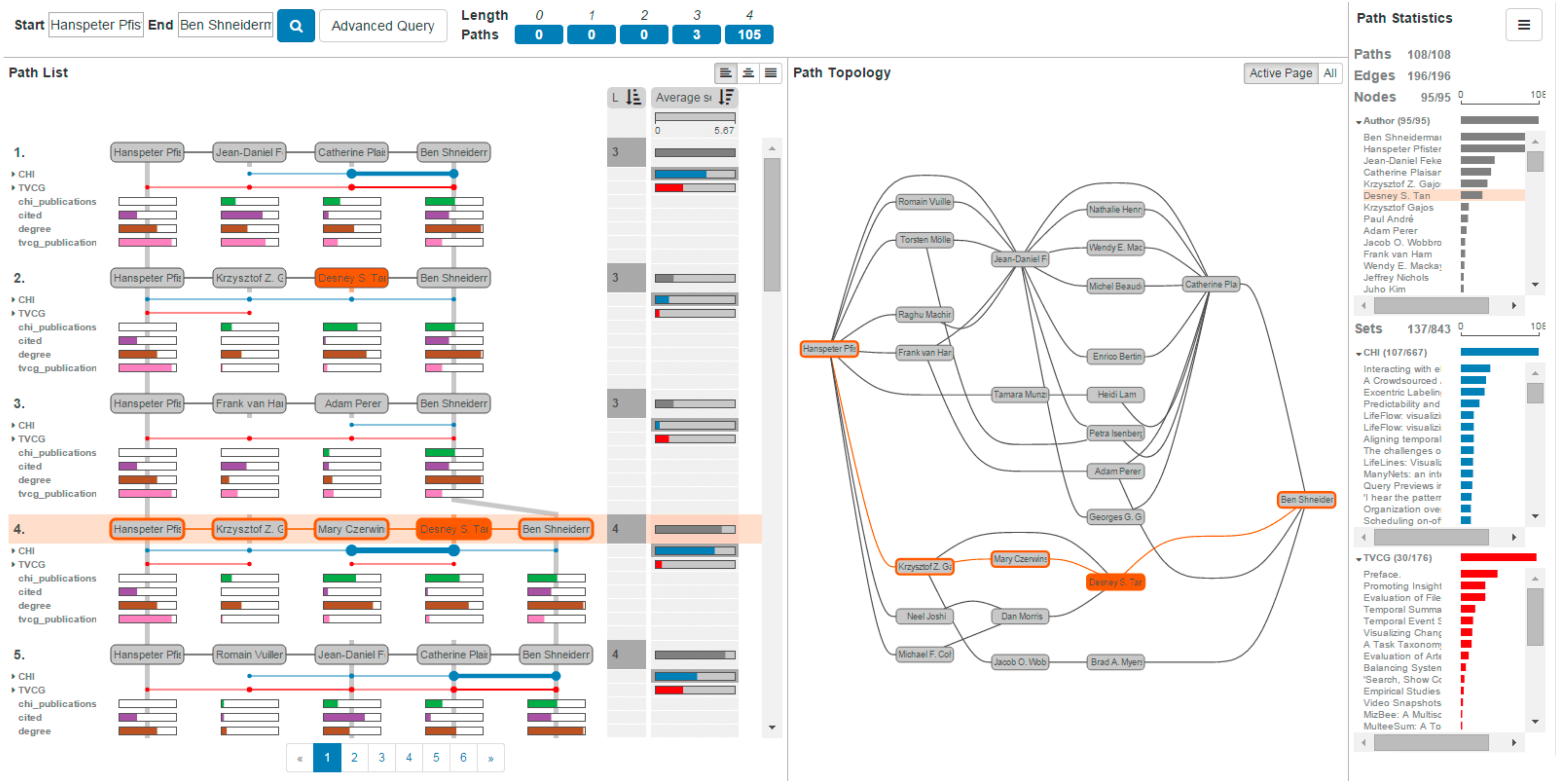


https://www.youtube.com/watch?v=Tzsv6wkZoiQ

http://www.cs.ou.edu/~weaver/improvise/examples/hotels/

Complex linked multiform views

System: **Pathfinder**



https://www.youtube.com/watch?v=aZF7AC8aNXo

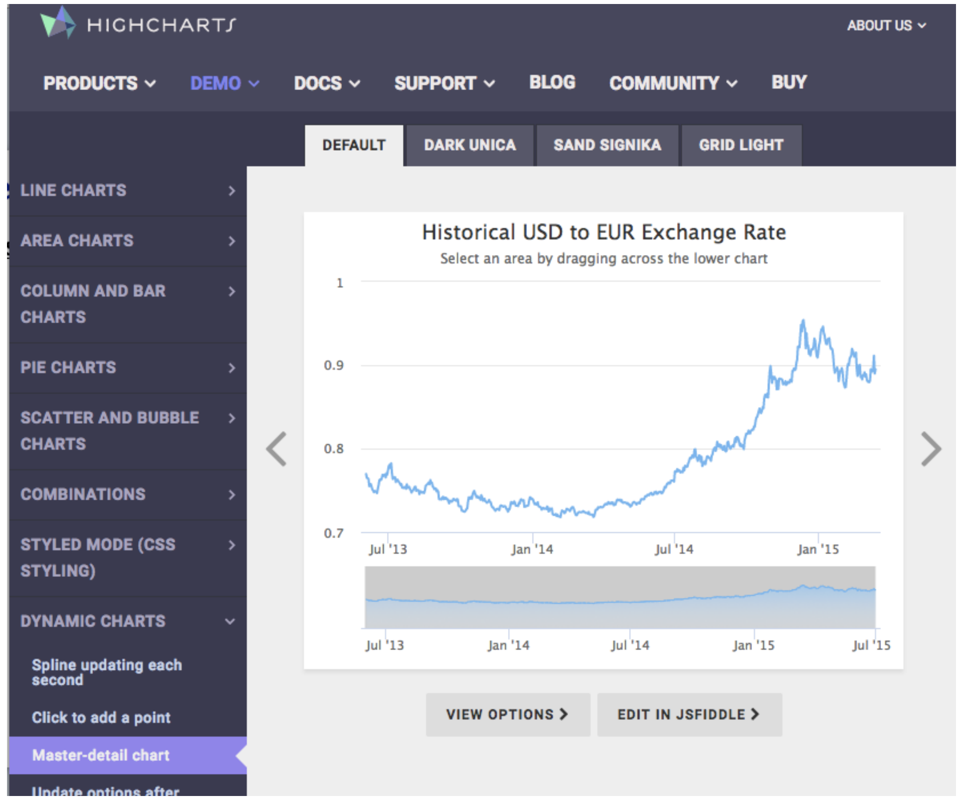
Idiom: **Overview-detail views** System: **Google Maps**

- encoding: same
- data: subset shared
- navigation: shared
  - bidirectional linking
- differences
  - viewpoint
  - (size)
- special case: ***birds-eye map***

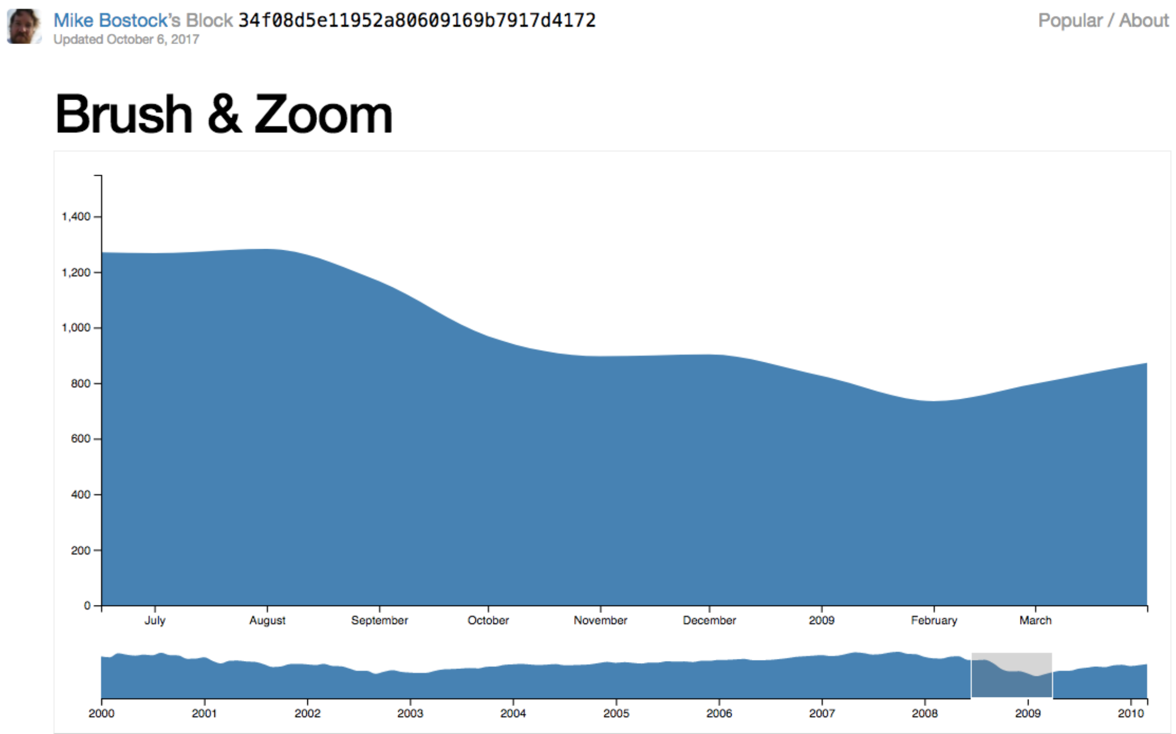
[A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. Cockburn, Karlson, and Bederson. ACM Computing Surveys 41:1 (2008), 1–31.]

Idiom: **Overview-detail navigation**

- encoding: same
- data: subset shared
- navigation: shared
  - unidirectional linking
  - select in small overview
  - change extent in large detail view



https://www.highcharts.com/demo/dynamic-master-detail

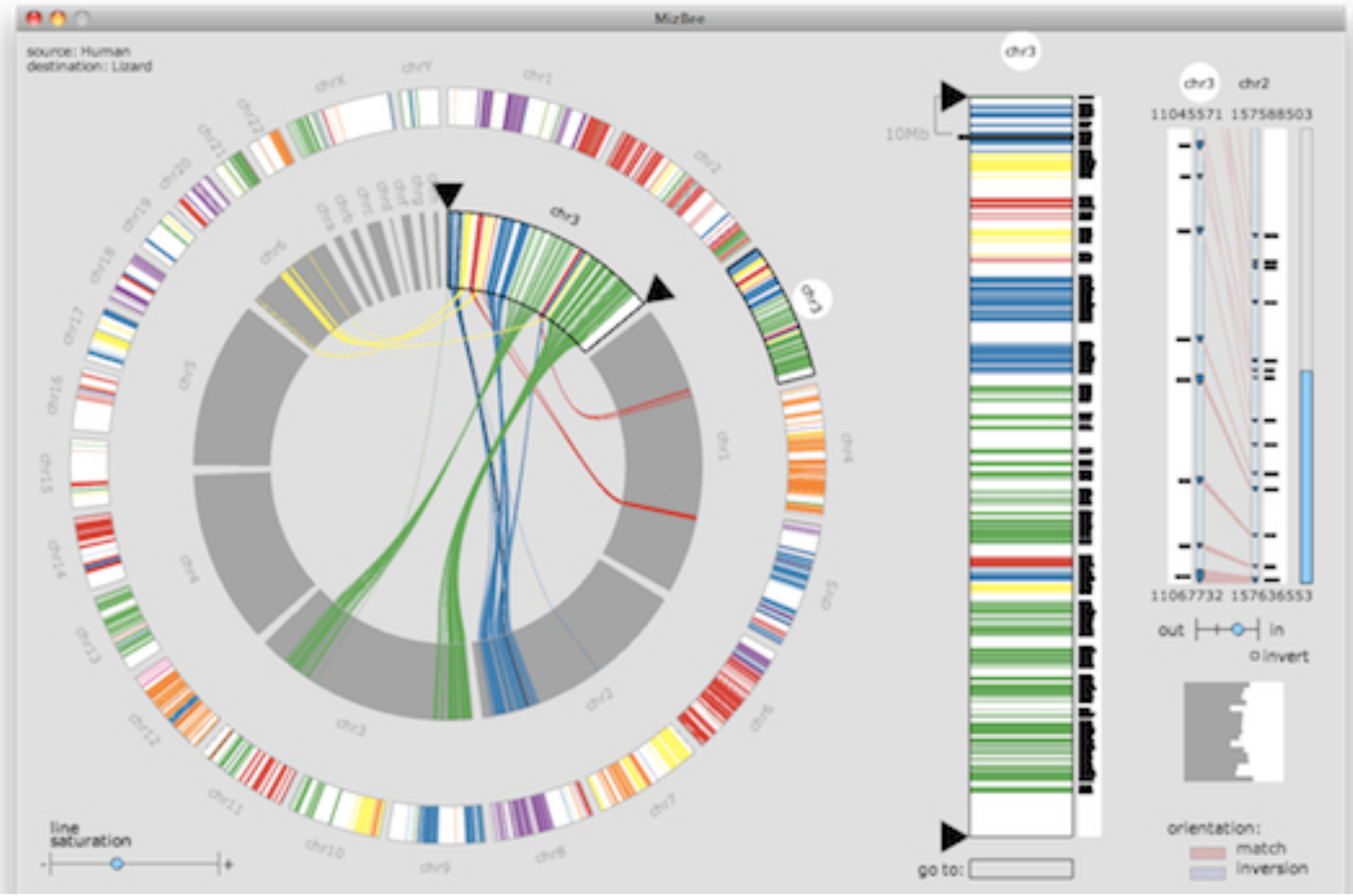


[*https://bl.ocks.org/mbostock/34f08d5e11952a80609169b7917d4172*](https://bl.ocks.org/mbostock/34f08d5e11952a80609169b7917d4172)

Overview-detail

System: **MizBee**

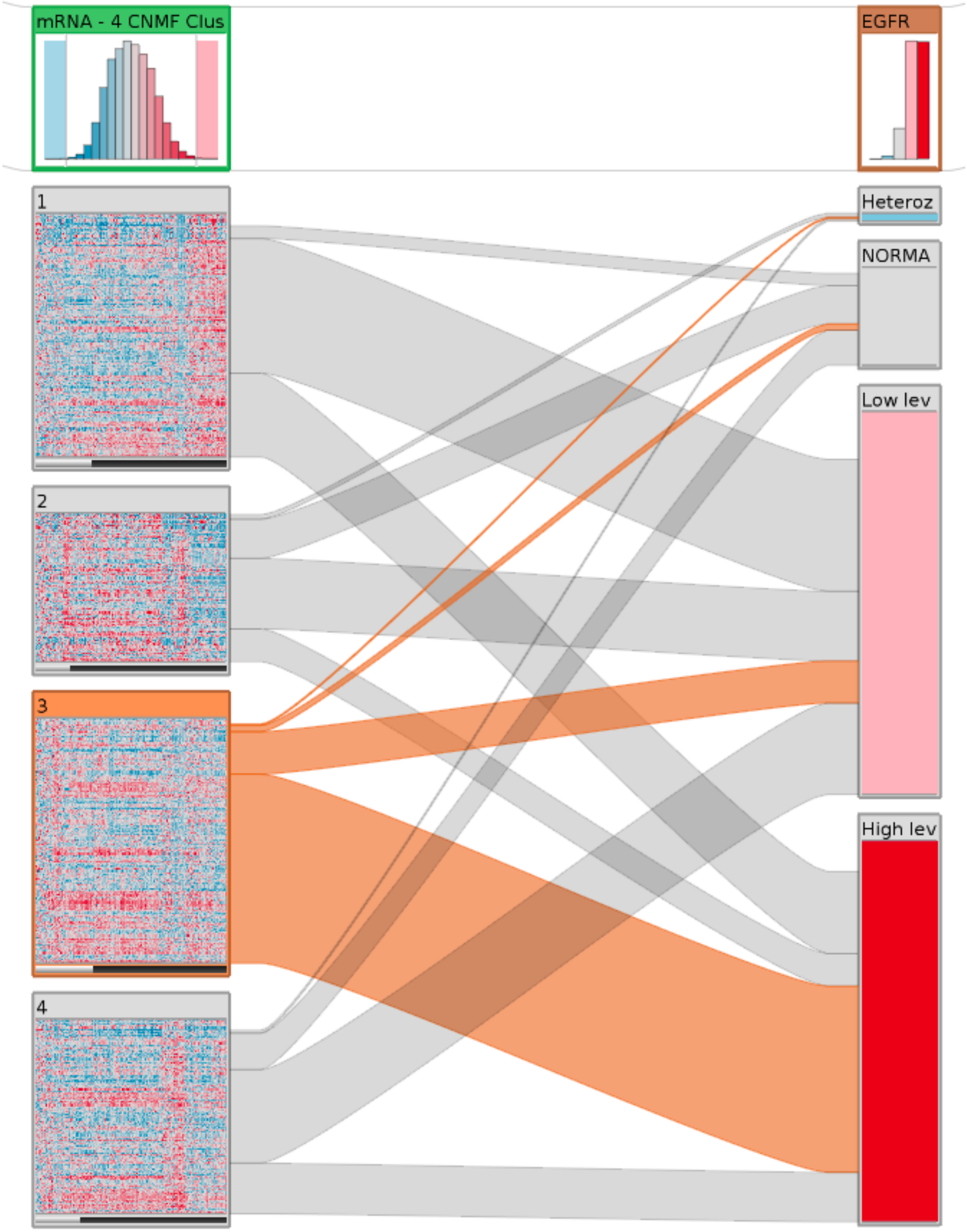
- multiscale: three viewing levels
  - linked views
  - dynamic filtering
  - tooling: processing
    (modern version: p5js.org)



https://www.youtube.com/watch?v=86p7brwuz2g

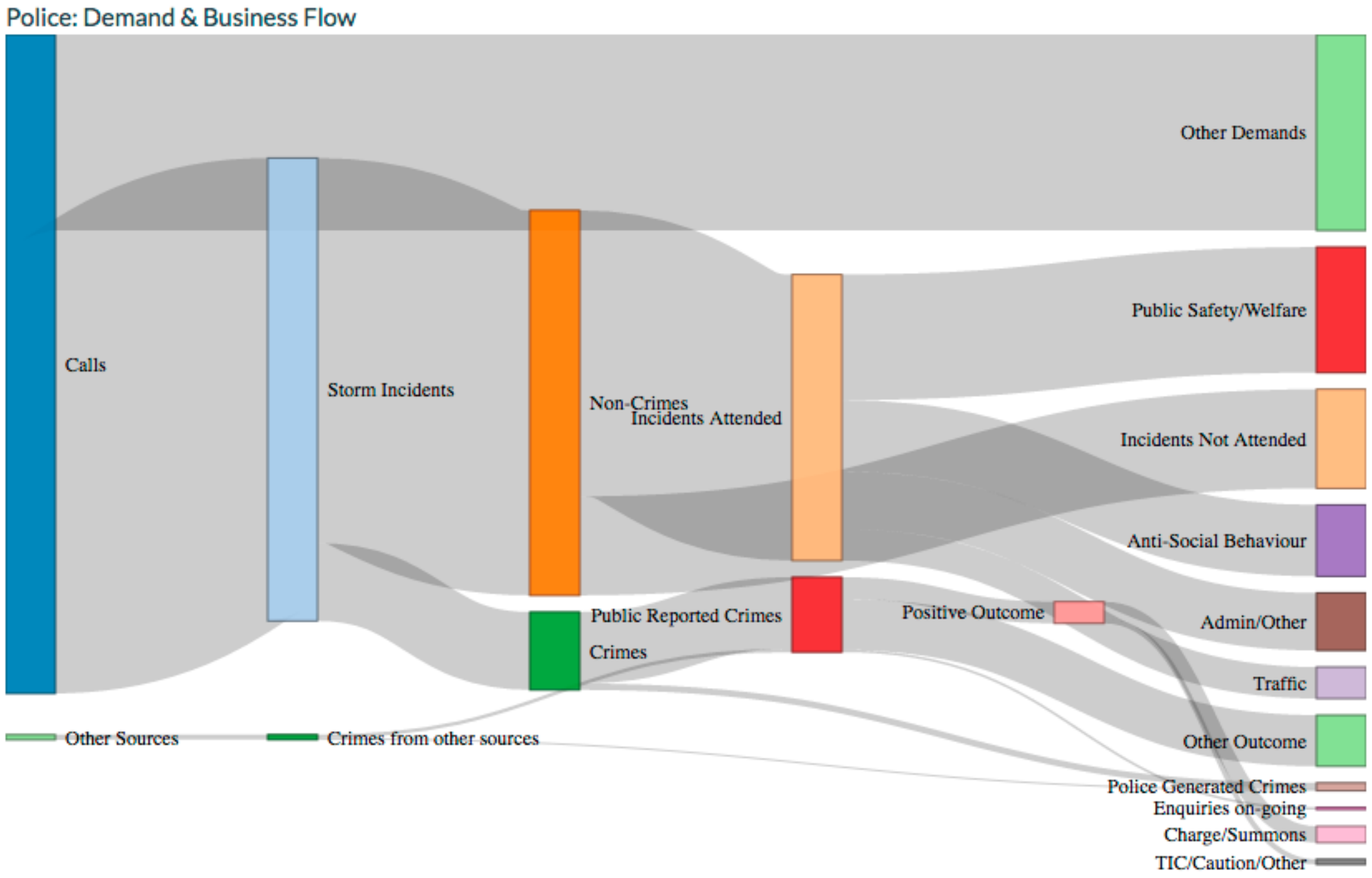
Overview-detail

System: **StratomeX**



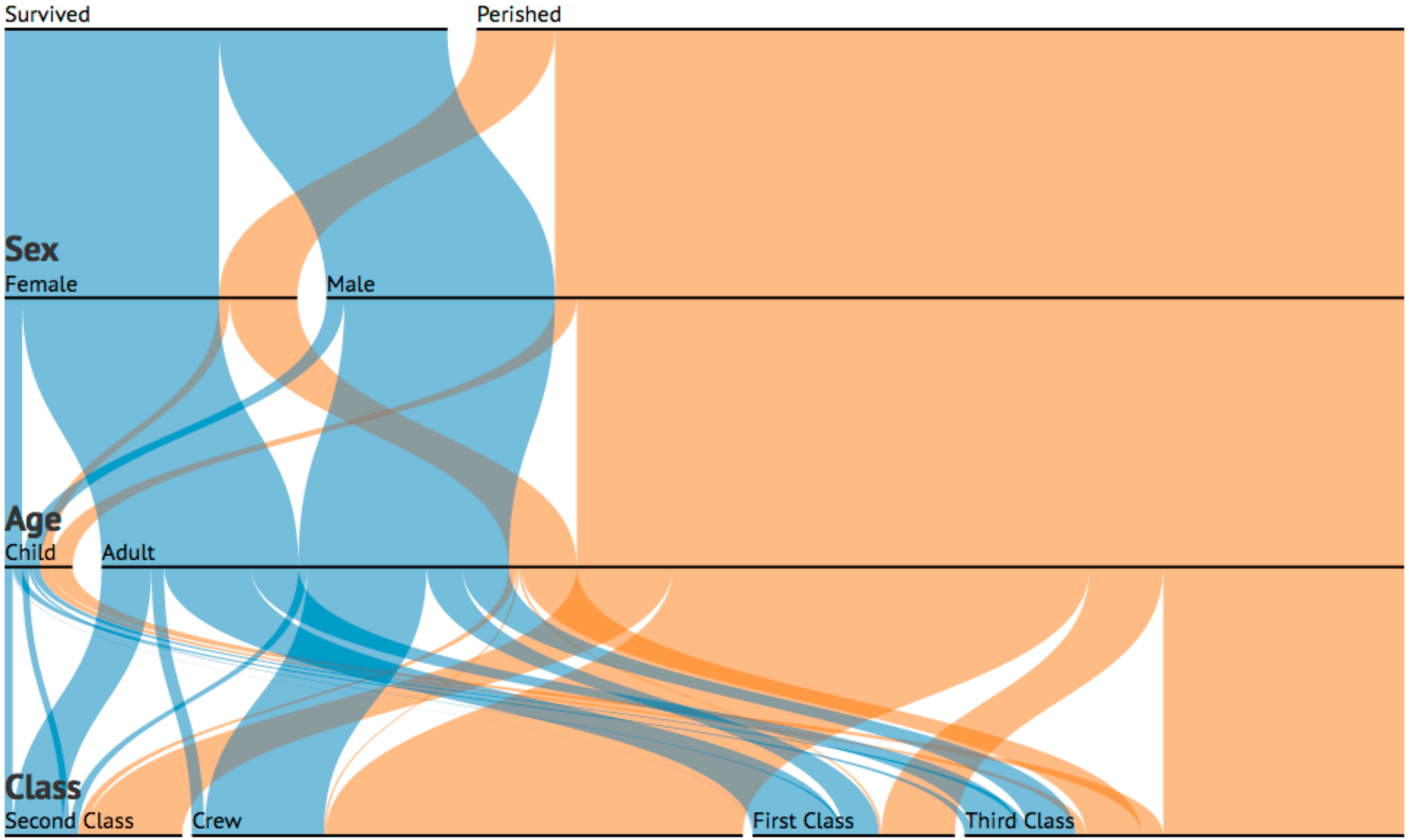
https://www.youtube.com/watch?v=UcKDbGqHsdE

Flows: R/Shiny



https://gallery.shinyapps.io/TSupplyDemand/

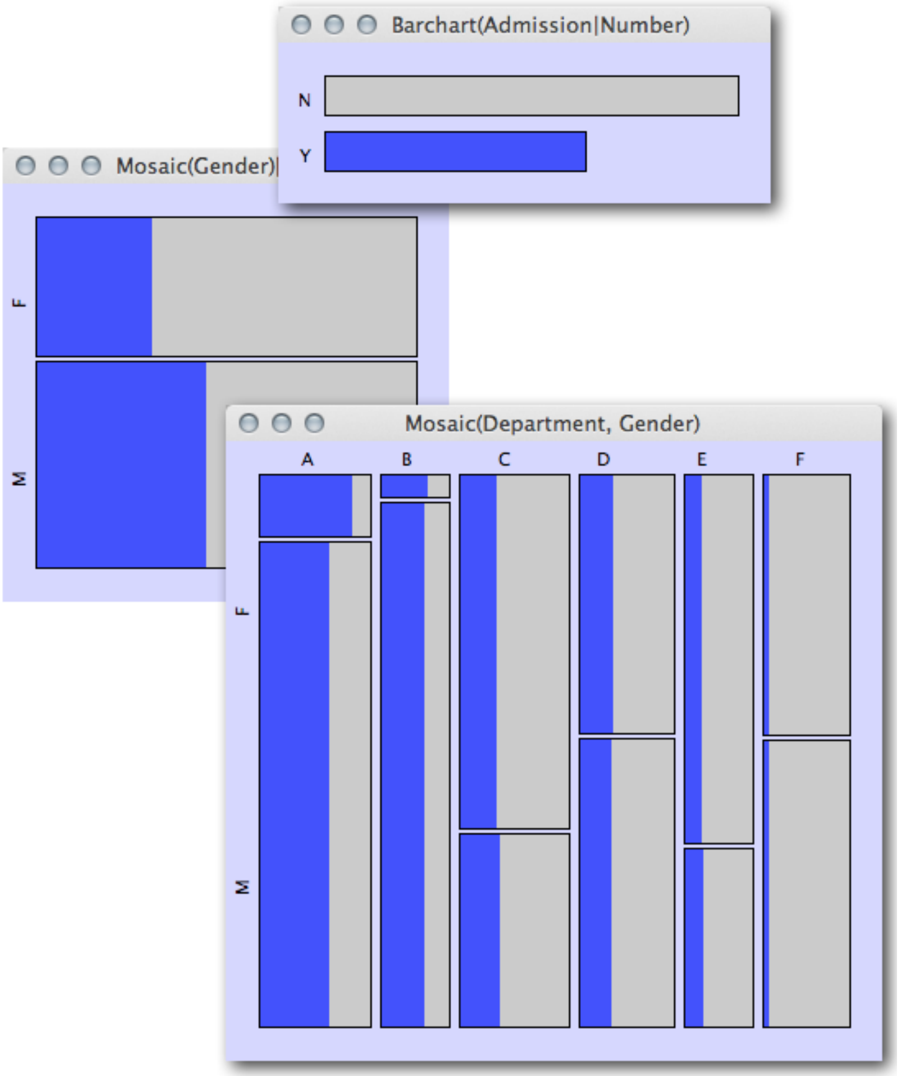
Idiom: **Parallel sets**



https://www.jasondavies.com/parallel-sets/

https://eagereyes.org/parallel-sets

Idiom: **Mosaic plots** System: **Mondrian**



http://www.theusrus.de/blog/understanding-mosaic-plots/

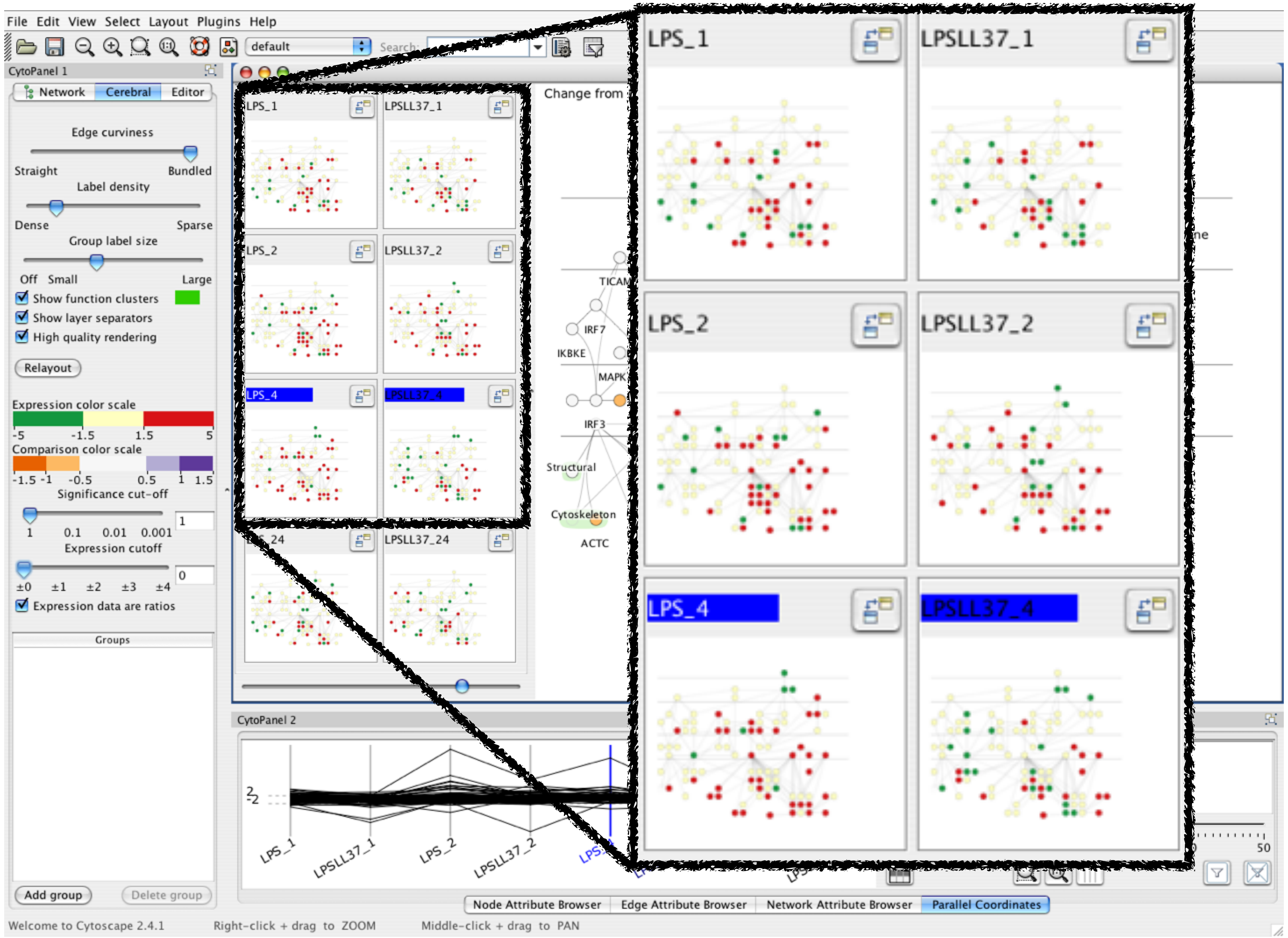
http://www.theusrus.de/Mondrian/

http://www.theusrus.de/blog/making-movies/

Idiom: **Small multiples**

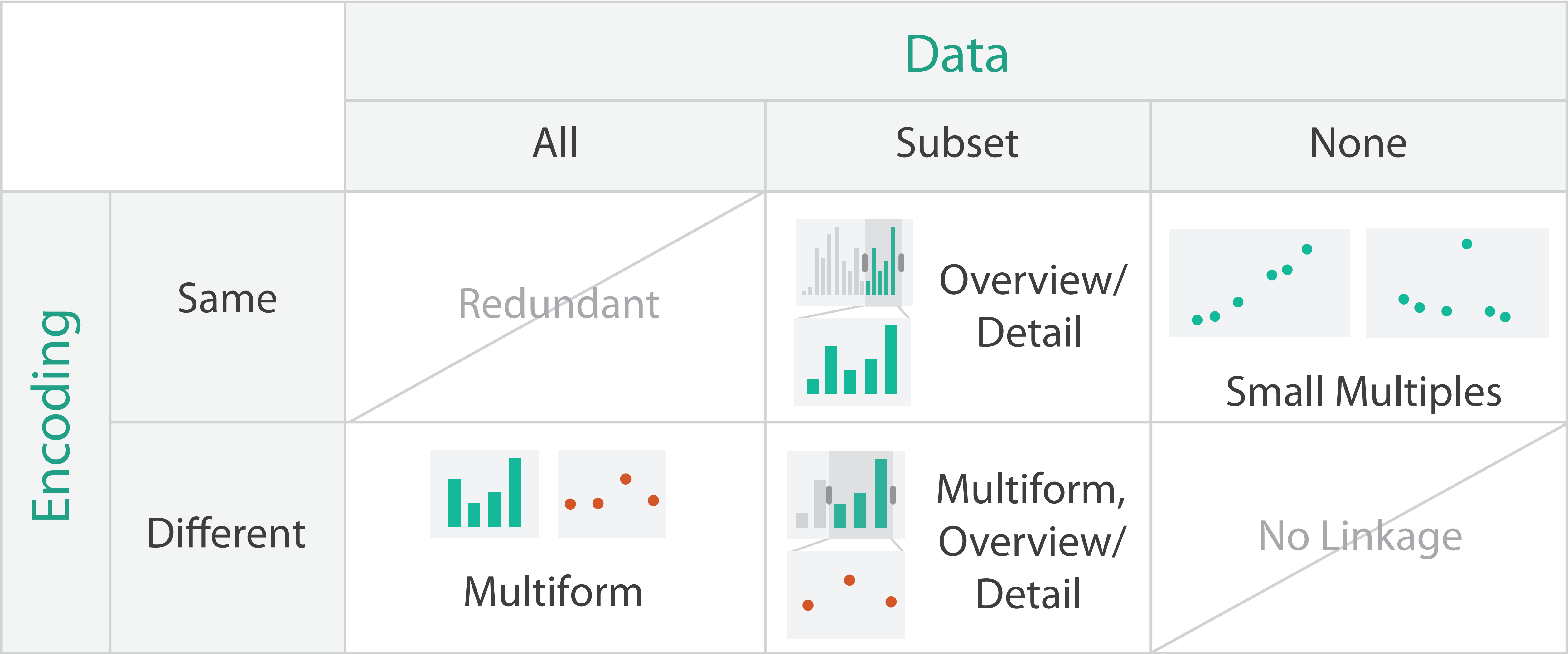
System: **Cerebral**

- 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

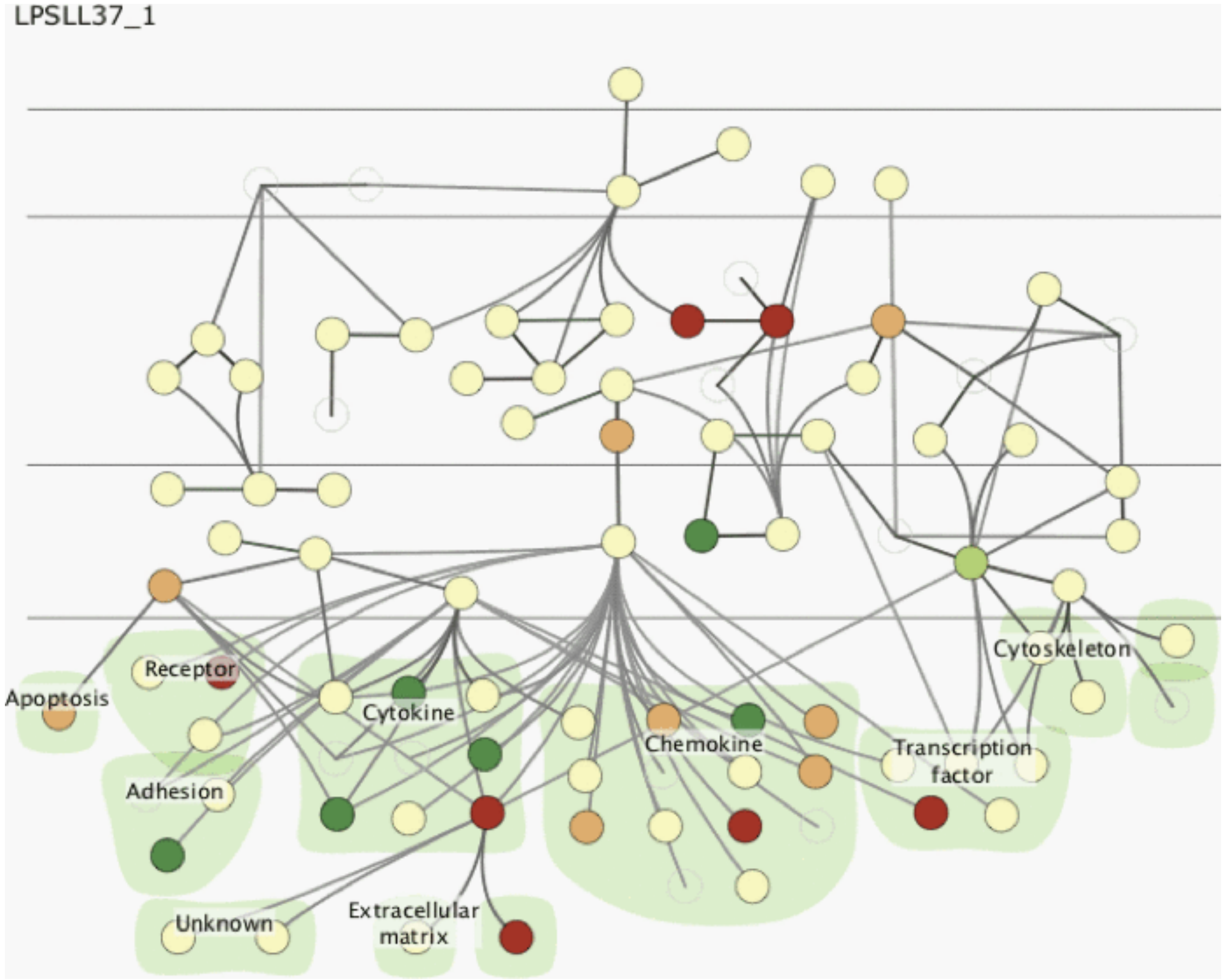


- why juxtapose views?
  - benefits: eyes vs memory
    - lower cognitive load to move eyes between 2 views than remembering previous state with single changing view
  - costs: display area, 2 views side by side each have only half the area of one view

Why not animation?

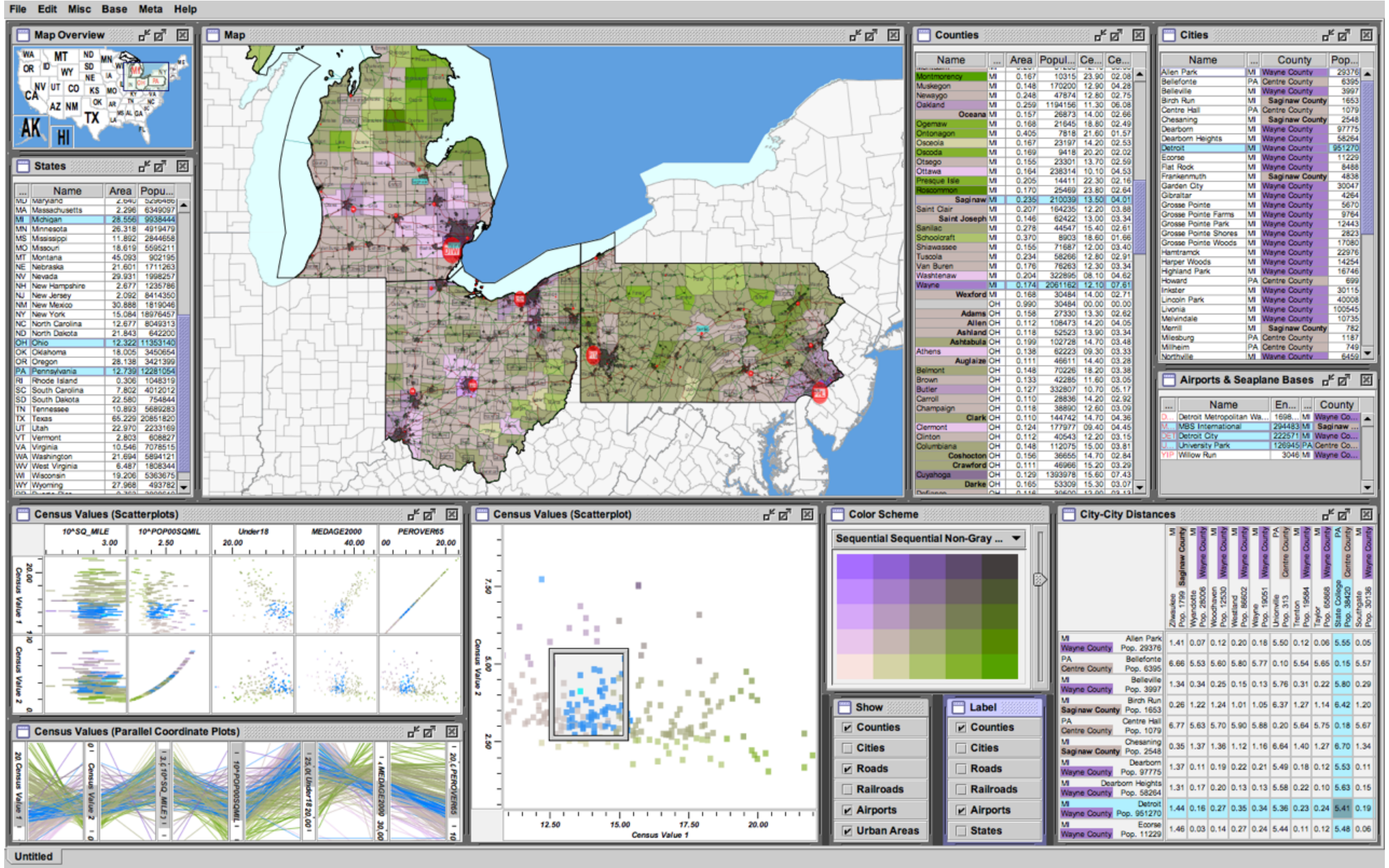
- disparate frames and regions: comparison difficult
  - –vs contiguous frames
  - –vs small region
  - –vs coherent motion of group

- 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

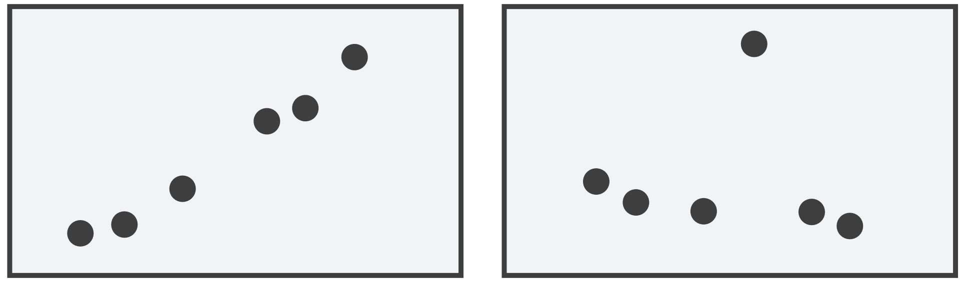


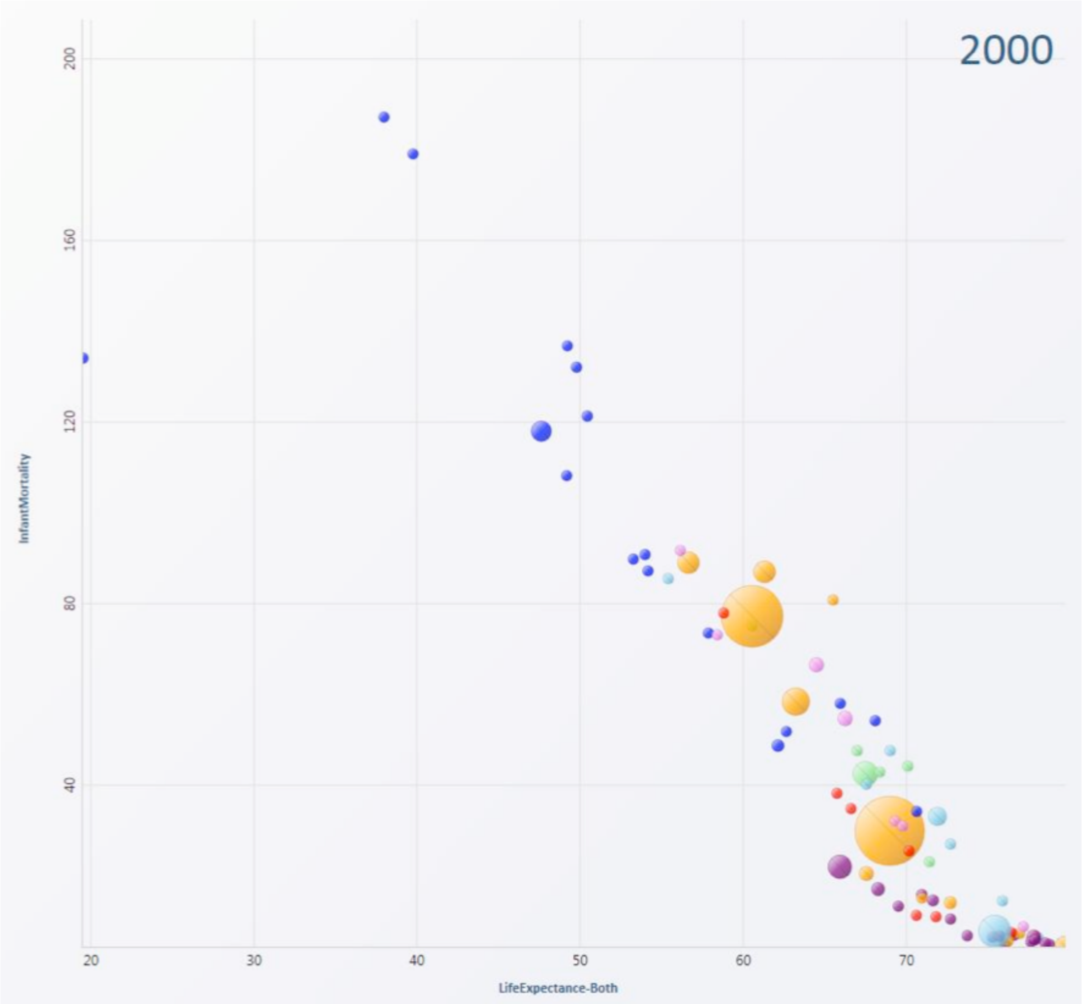
[Building Highly-Coordinated Visualizations In Improvise. Weaver. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 159–166, 2004.]

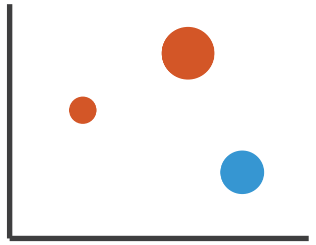
Partition into views

- how to divide data between views
  - 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

➔ Partition into Side-by-Side Views

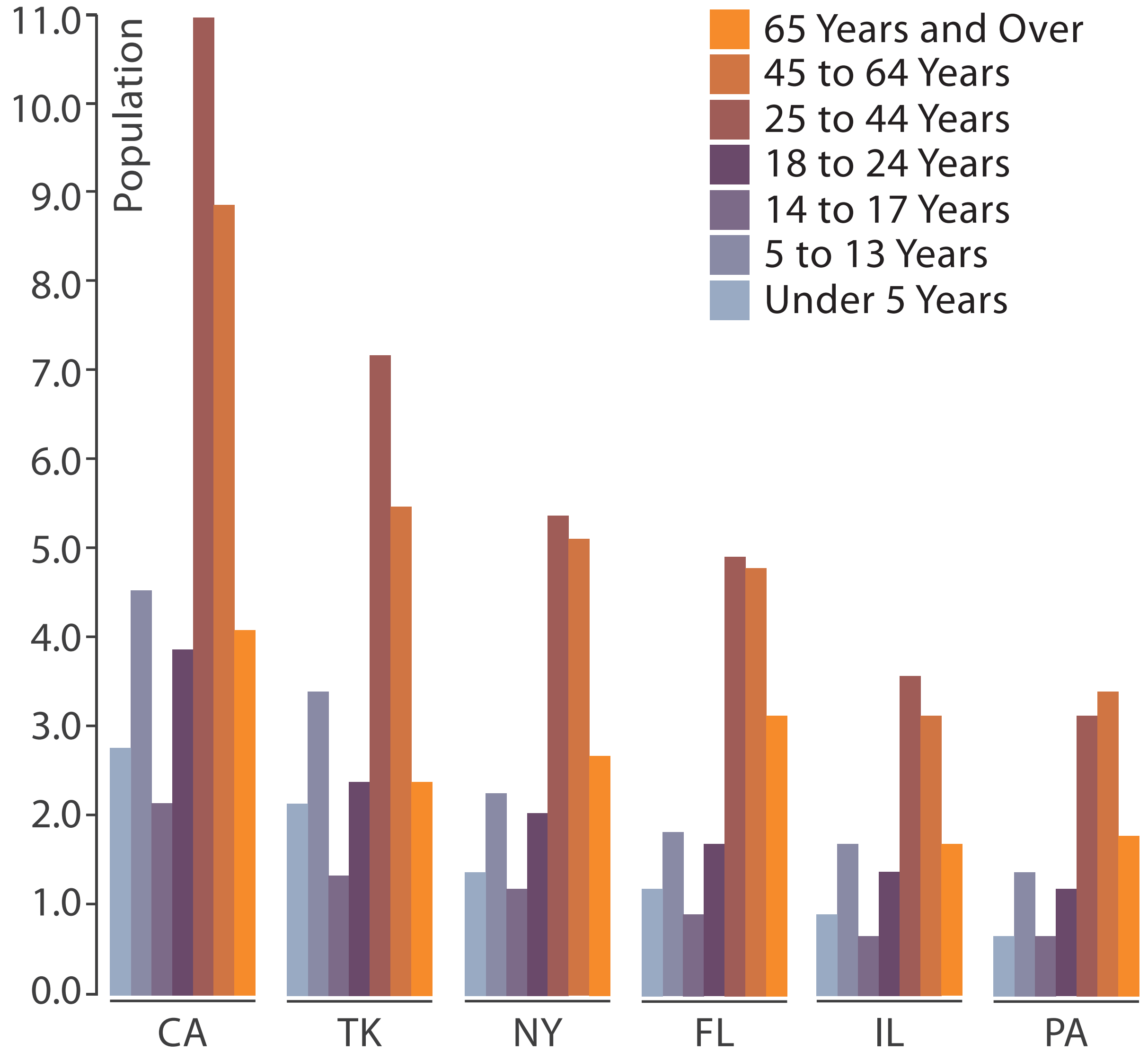




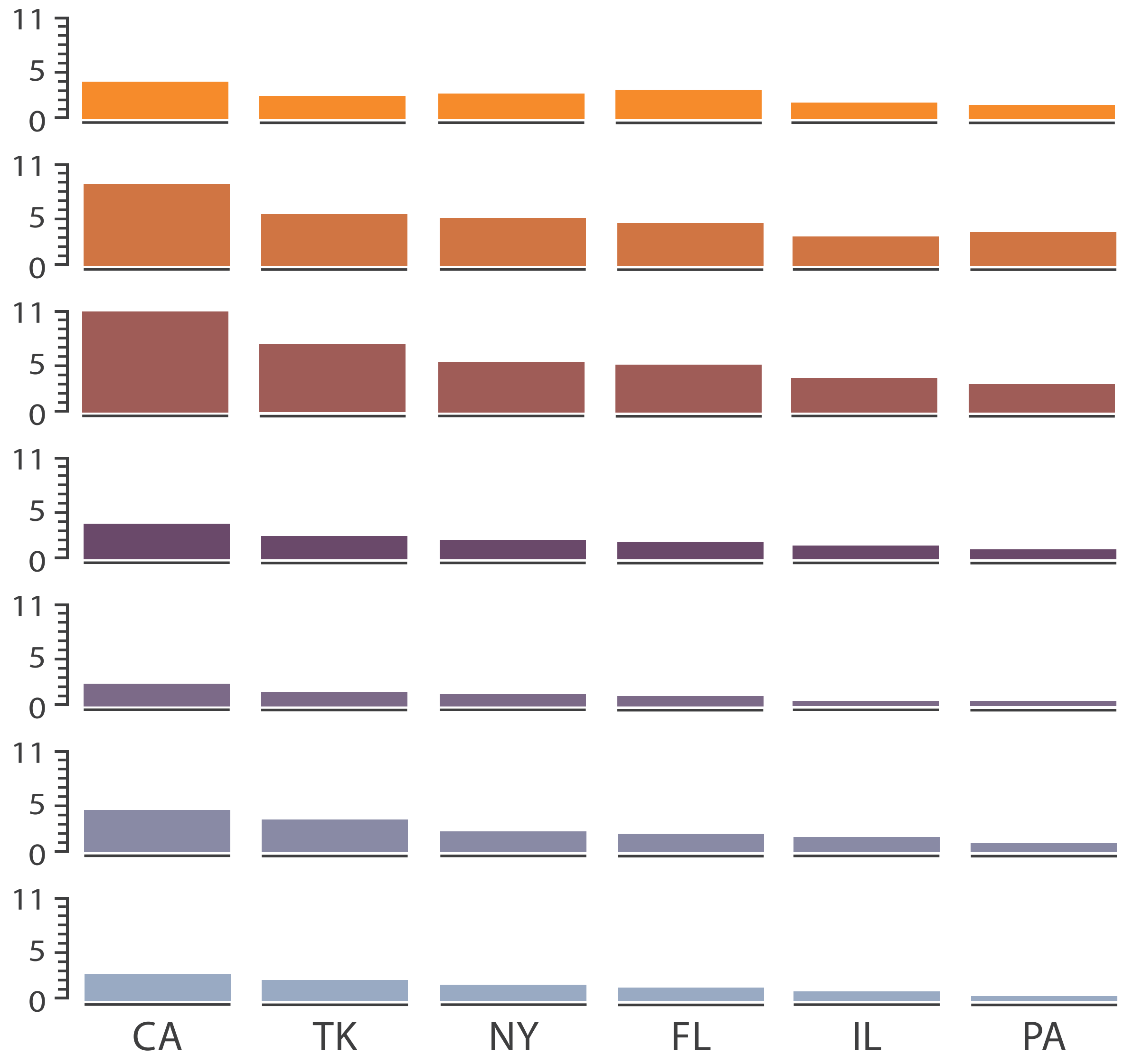


Partitioning: List alignment

- 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

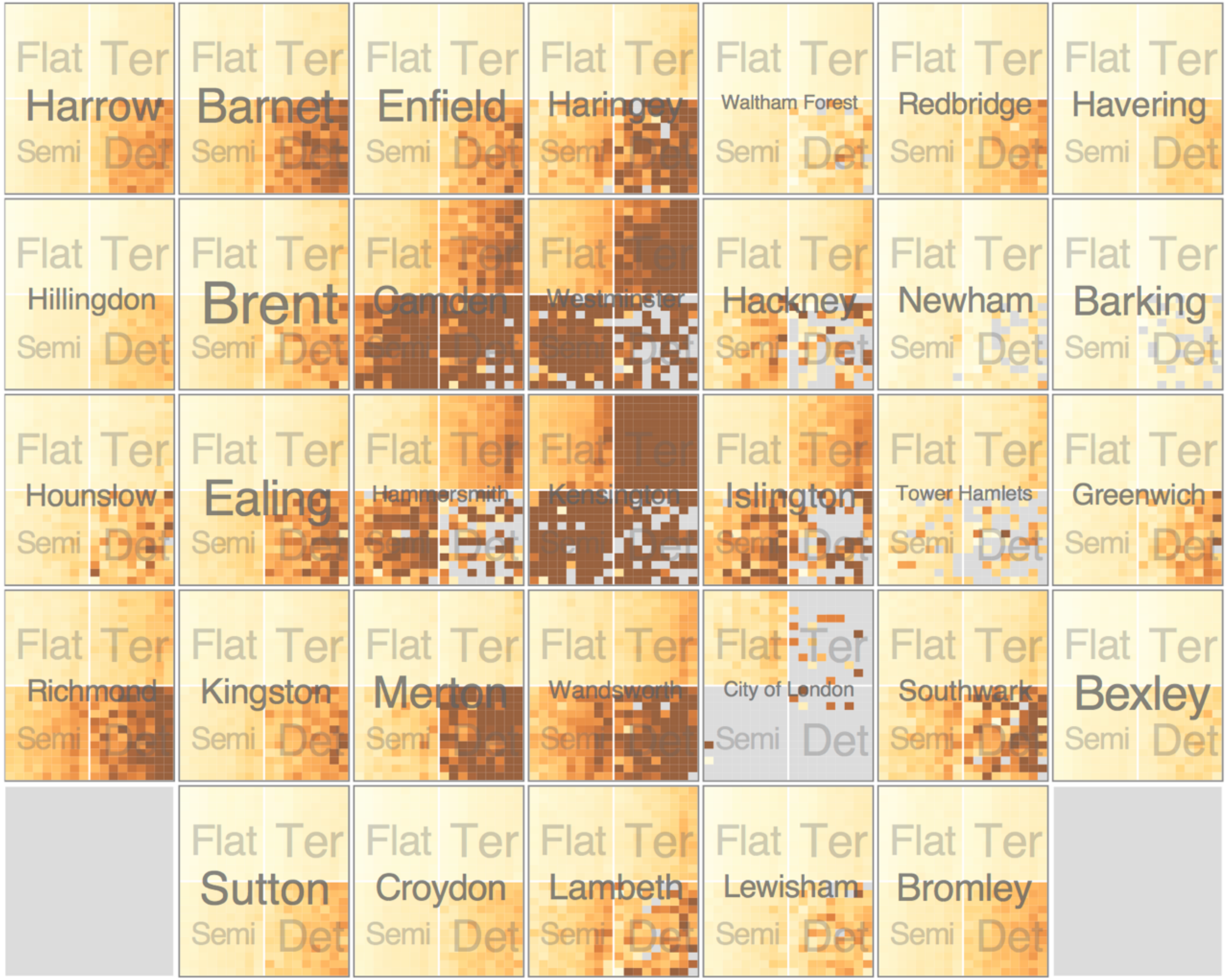


Partitioning: Recursive subdivision

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



[Configuring Hierarchical Layouts to Address Research Questions. Slingsby, Dykes, and Wood. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2009) 15:6 (2009), 977–984.]

Partitioning: Recursive subdivision

System: **HIVE**

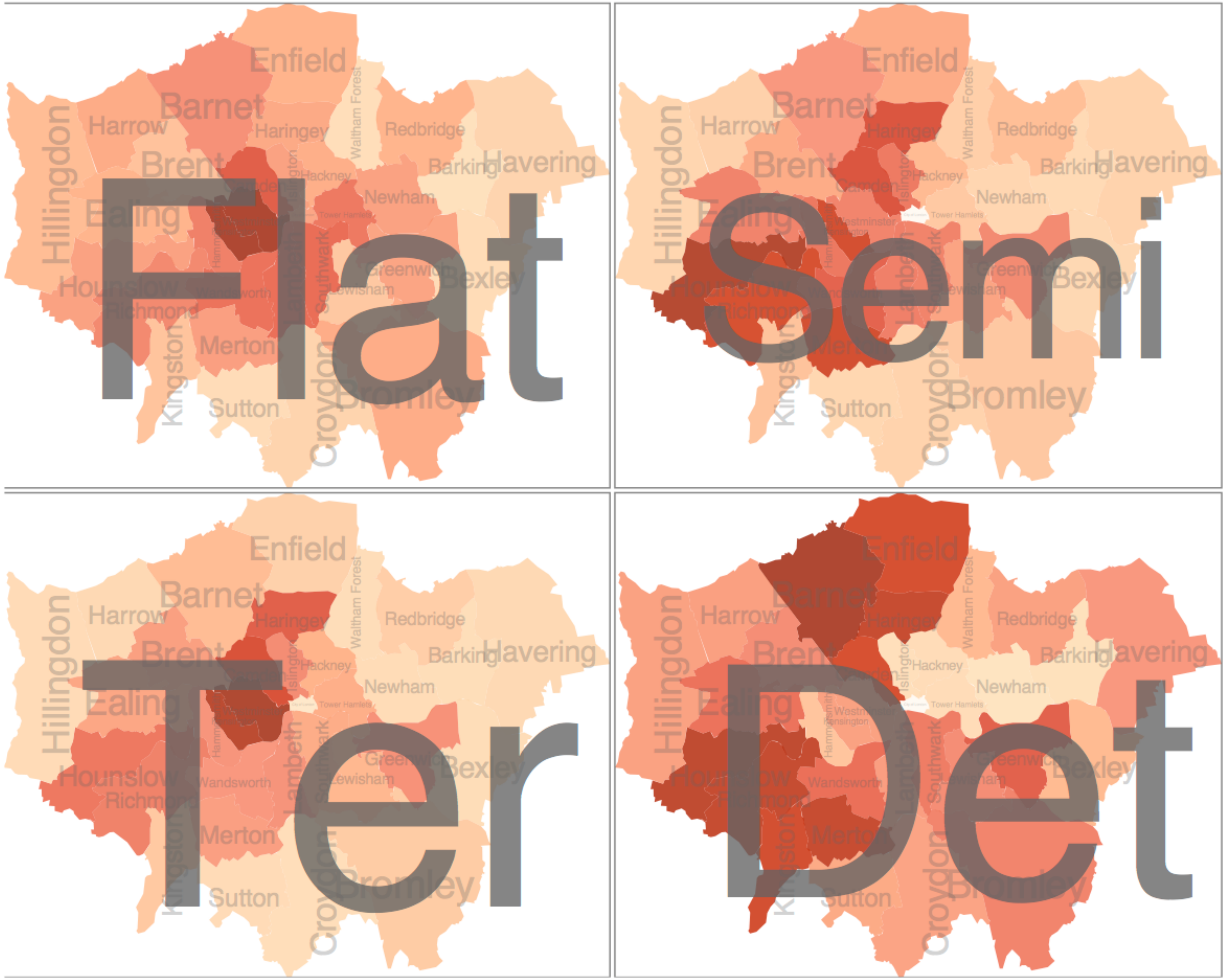
- switch order of splits
  - type then neighborhood
- switch color
  - by price variation

- type patterns
  - within specific type, which neighborhoods inconsistent

[Configuring Hierarchical Layouts to Address Research Questions. Slingsby, Dykes, and Wood. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2009) 15:6 (2009), 977–984.]

Partitioning: Recursive subdivision

- different encoding for second-level regions
  - –choropleth maps

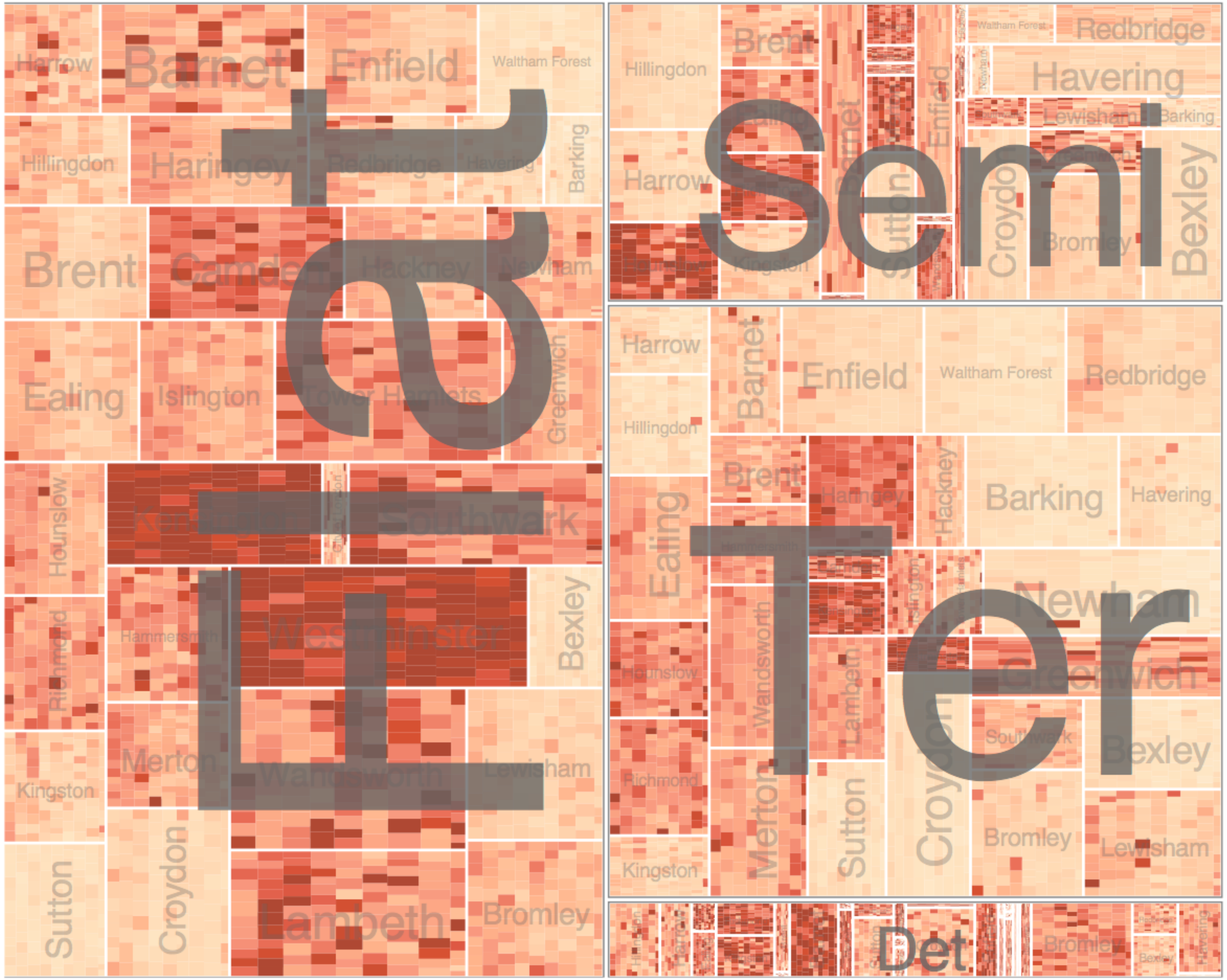


[Configuring Hierarchical Layouts to Address Research Questions. Slingsby, Dykes, and Wood. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2009) 15:6 (2009), 977–984.]

Partitioning: Recursive subdivision

System: **HIVE**

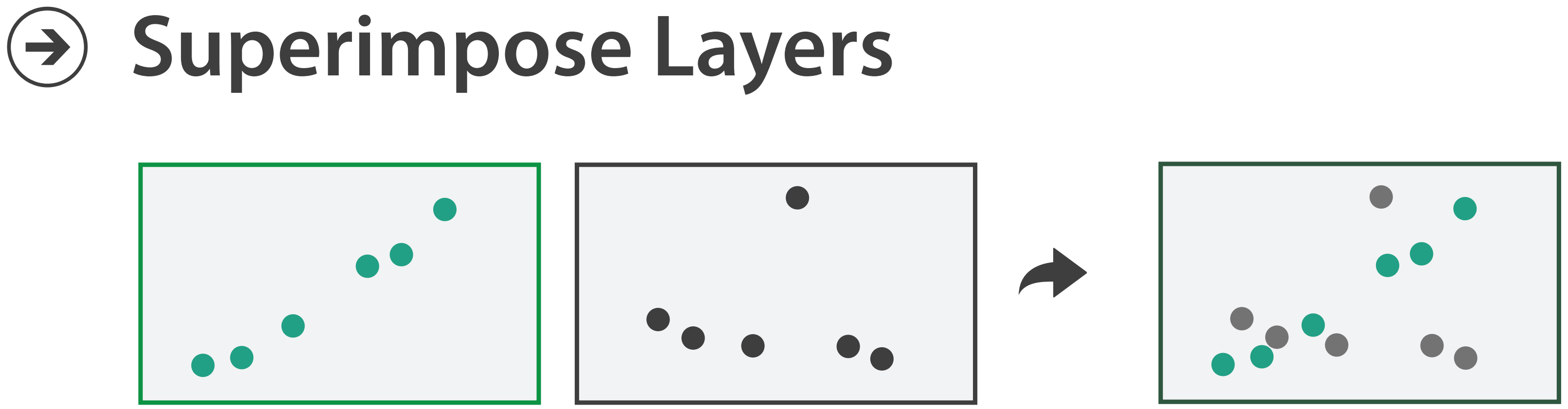
- size regions by sale counts
  - not uniformly
- result: treemap



[Configuring Hierarchical Layouts to Address Research Questions. Slingsby, Dykes, and Wood. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2009) 15:6 (2009), 977–984.]

Superimpose layers

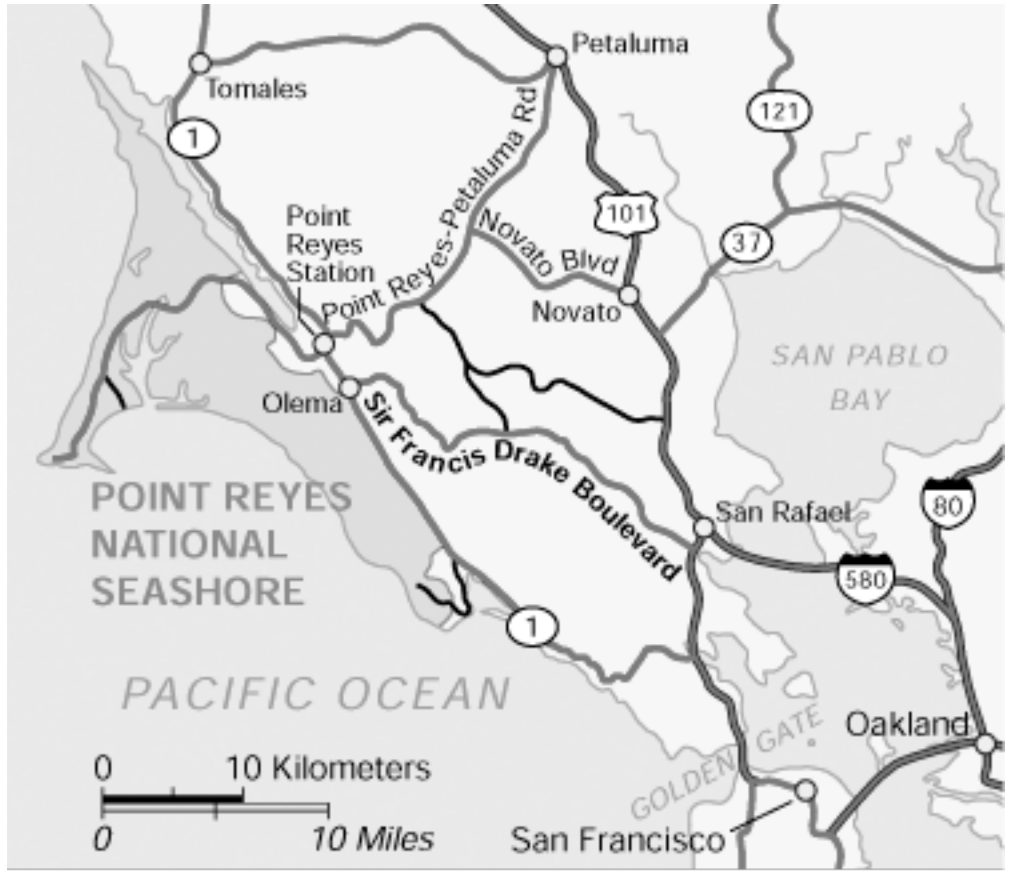
- ***layer***: set of objects spread out over region
  - –each set is visually distinguishable group
  - –extent: whole view
- design choices
  - –how many layers, how to distinguish?
    - encode with different, nonoverlapping channels
    - two layers achieveable, three with careful design
  - –small static set, or dynamic from many possible?



Static visual layering

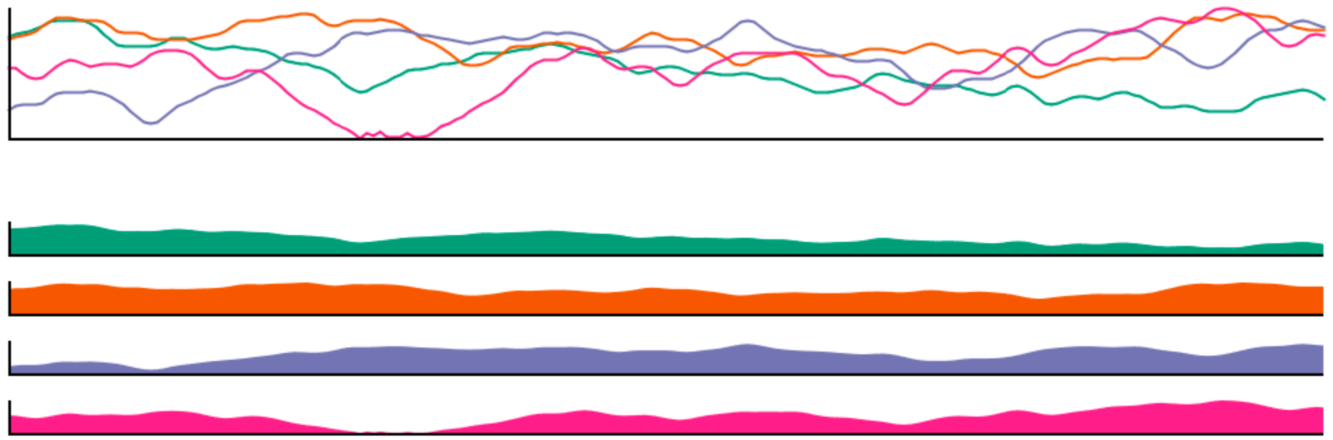
- foreground layer: roads
  - hue, size distinguishing main from minor
  - high luminance contrast from background
- background layer: regions
  - desaturated colors for water, parks, land areas
- user can selectively focus attention
- "get it right in black and white"
  - check luminance contrast with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

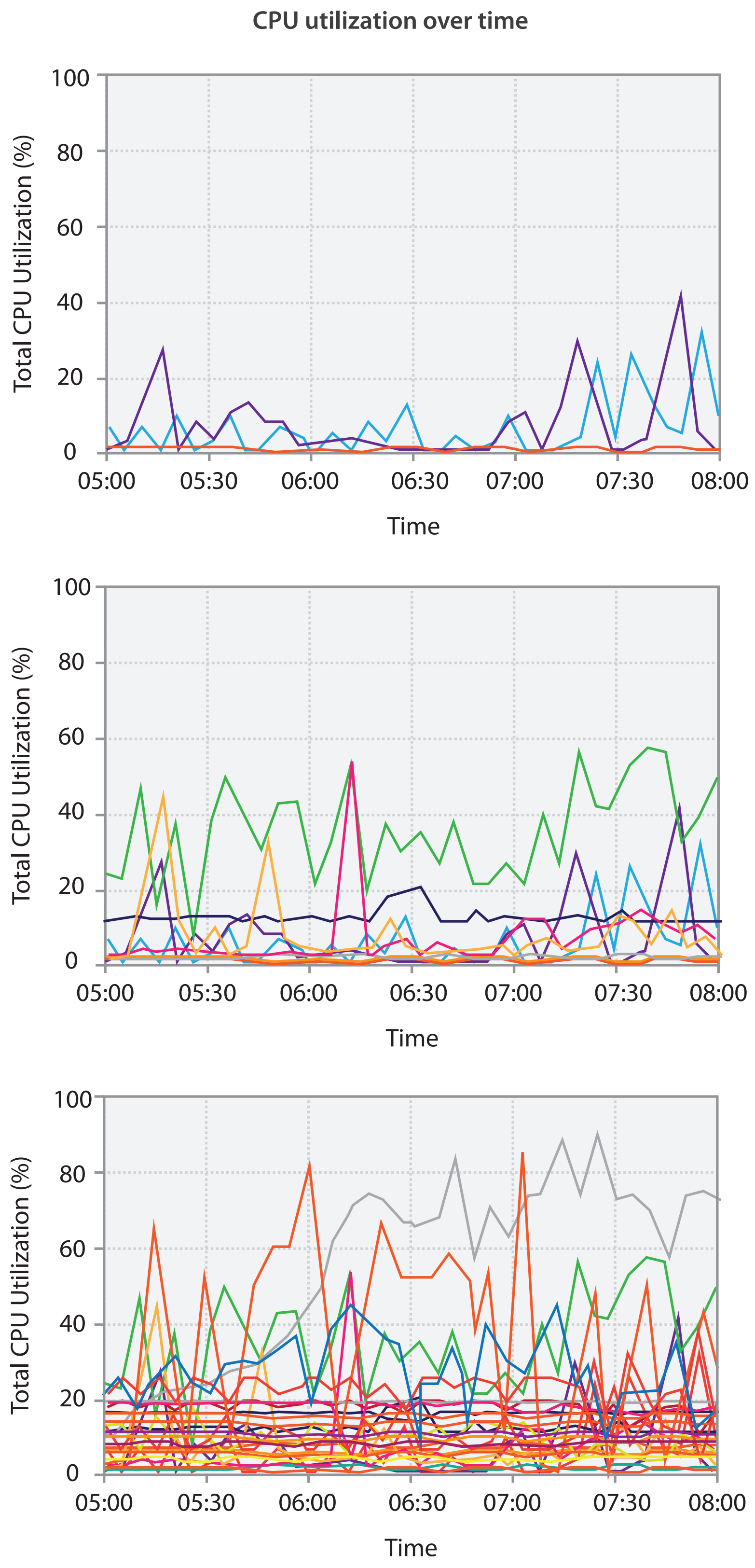


Superimposing limits

- 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

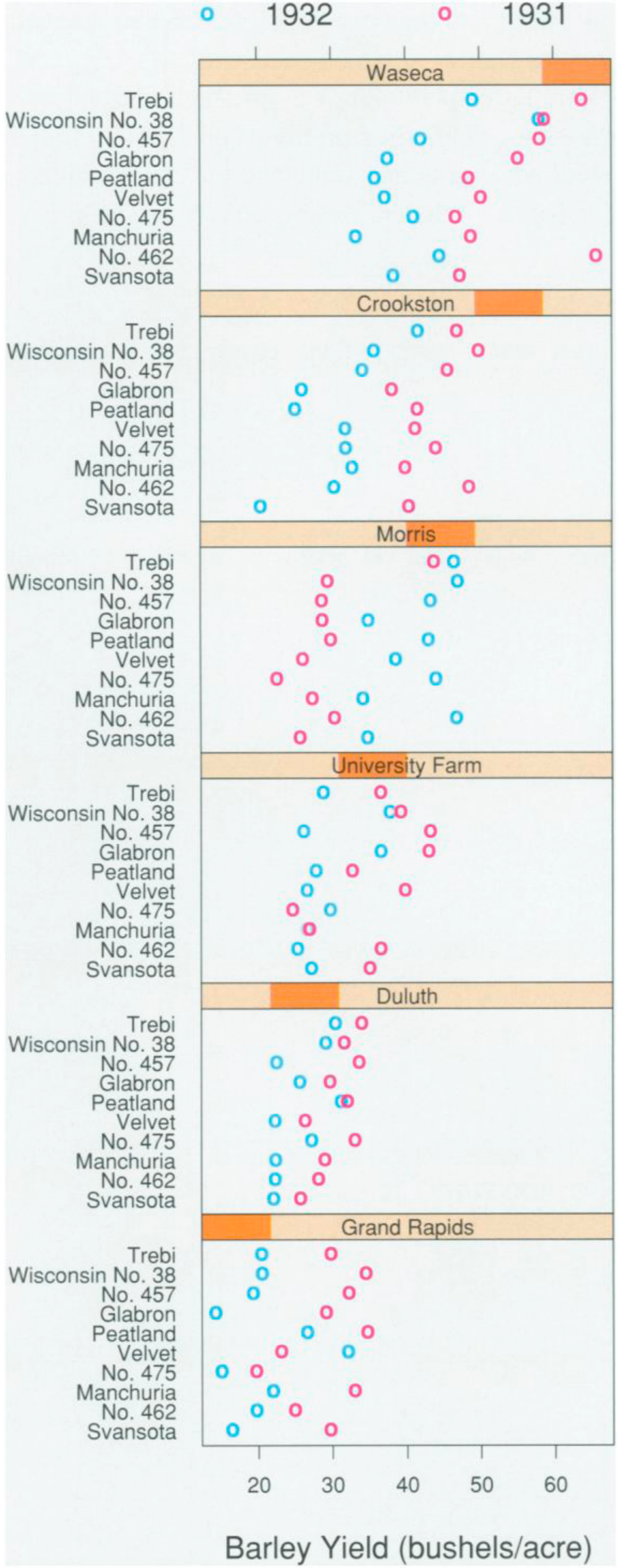


[Graphical Perception of Multiple Time Series. Javed, McDonnel, and Elmqvist. IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE InfoVis 2010) 16:6 (2010), 927–934.]



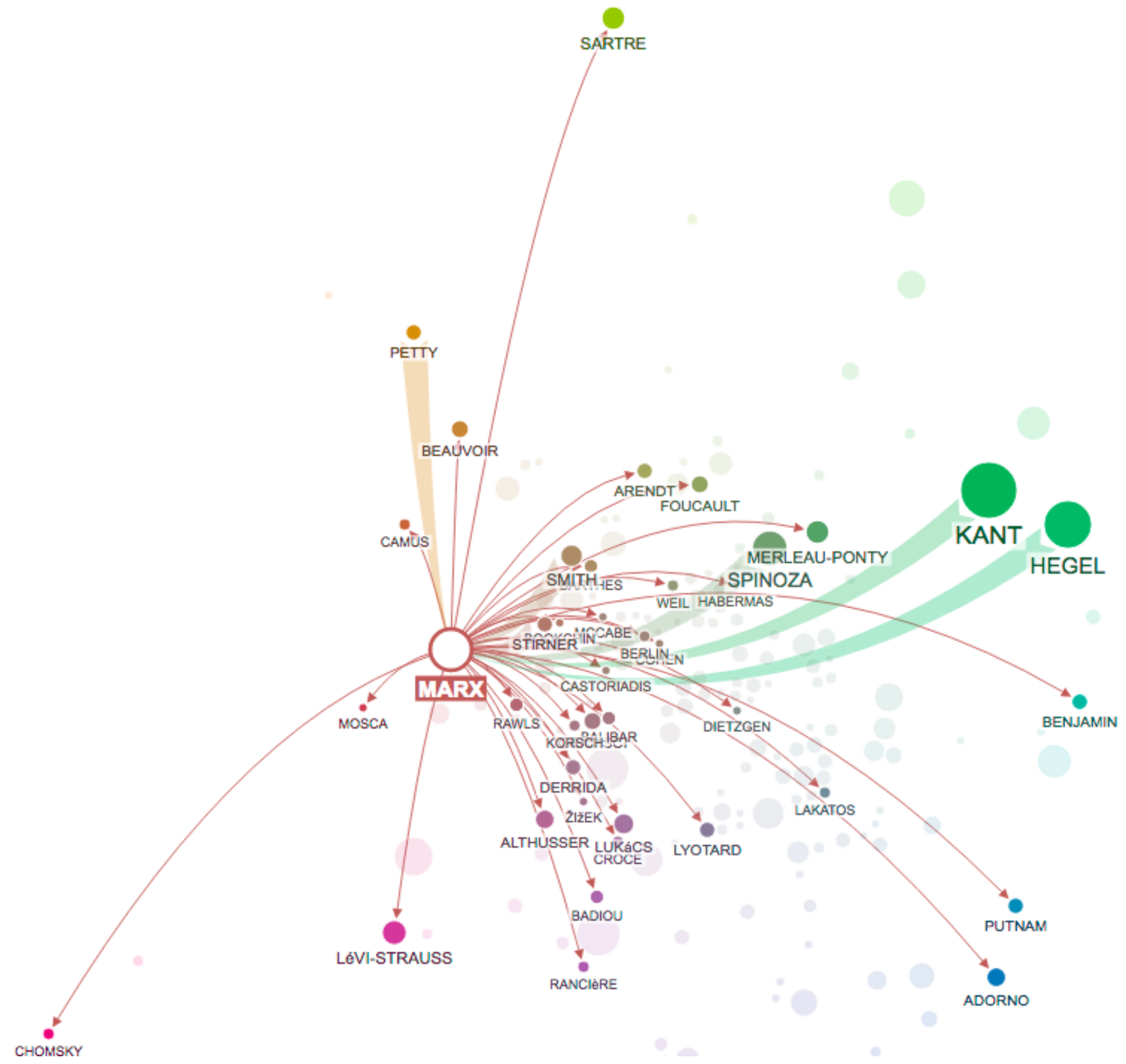
Idiom: **Trellis plots**

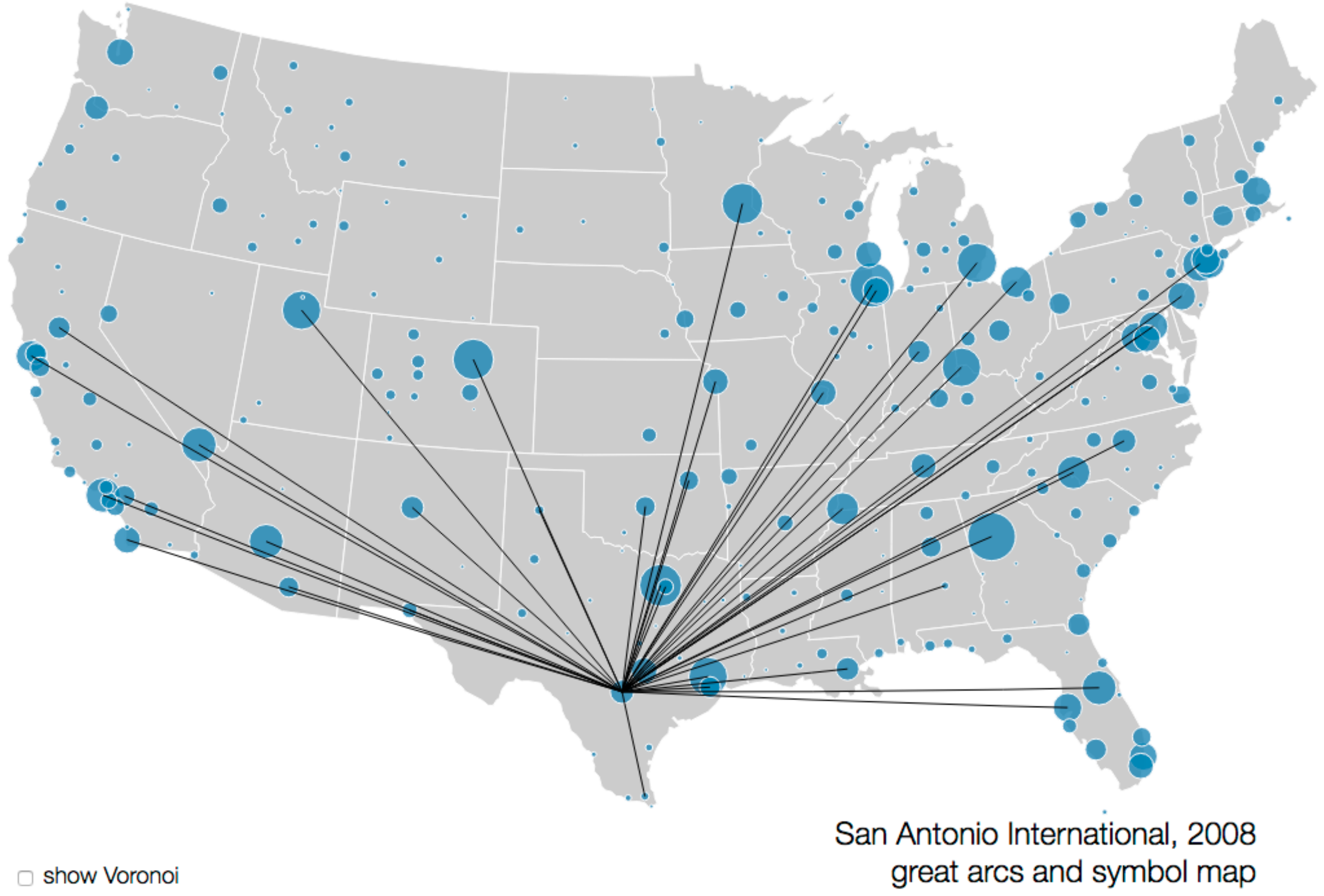
- superimpose within same frame
  - color code by year
- partitioning
  - split by site, rows are wheat varieties
- main-effects ordering
  - derive value of median for group, use to order
  - order rows within view by variety median
  - order views themselves by site median



Dynamic visual layering

- interactive based on selection
- one-hop neighbour highlighting demos: click vs hover (lightweight)





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.

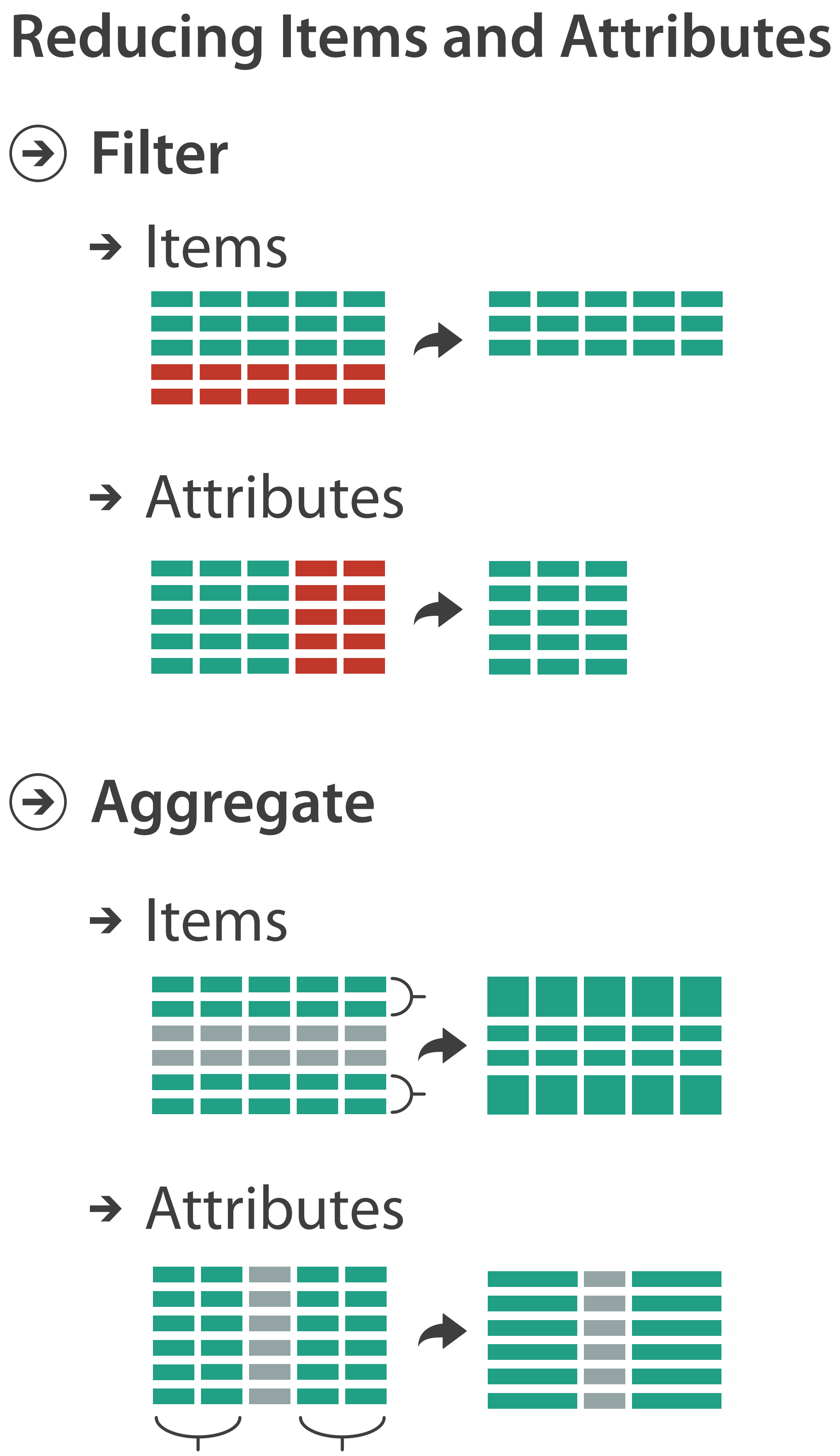
Further reading, full

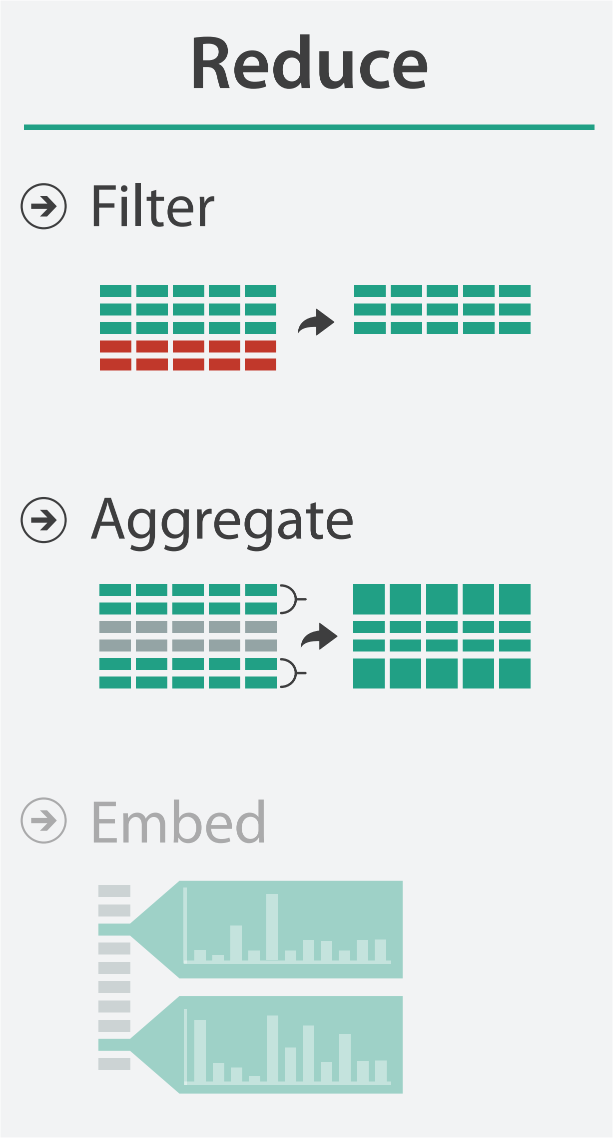
- A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. Computing Surveys 41:1 (2008), pp. 1-31.
- A Guide to Visual Multi-Level Interface Design from Synthesis of Empirical Study Evidence. Heidi Lam and Tamara Munzner. Synthesis Lectures on Visualization Series, Morgan Claypool, 2010.
- Zooming versus Multiple Window Interfaces: Cognitive Costs of Visual Comparisons. M. Plumlee and C. Ware. Transactions on Computer-Human Interaction (ToCHI) 13:2 (2006), pp. 179-209.
- Exploring the Design Space of Composite Visualization. Waqas Javed and Niklas Elmqvist. Proc. IEEE Symposium on Pacific Visualization (PacificVis), pp. 1-9. IEEE Computer Society, 2012.
- Visual Comparison for Information Visualization. Michael Gleicher, Danielle Albers, Rick Walker, Ilir Jusufi, Charles D. Hansen, and Jonathan C. Roberts. Information Visualization 10:4 (2011), pp. 289-309.
- Guidelines for Using Multiple Views in Information Visualizations. Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. Proc. International Working Conference on Advanced Visual Interfaces (AVI), pp. 110-119. ACM, 2000.
- State of the Art: Coordinated & Multiple Views in Exploratory Visualization. Jonathan C. Roberts. Proc. Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV), pp. 61-71. IEEE Computer Society, 2007.
- Building Highly-Coordinated Visualizations in Improvise. Chris Weaver. Proc. IEEE Symposium on Information Visualization (InfoVis), pp. 159-166. IEEE Computer Society, 2004.
- Cross-Filtered Views for Multidimensional Visual Analysis. Chris Weaver. IEEE TVCG (2010), pp. 192-204.
- Configuring Hierarchical Layouts to Address Research Questions. Adrian Slingsby, Jason Dykes, and Jo Wood. IEEE TVCG (Proc. InfoVis 2009), pp. 977-984.
- Spatially Ordered Treemaps. Jo Wood and Jason Dykes. IEEE TVCG (Proc. InfoVis 2008), pp. 1348-1355.
- Glyph-Based Visualization: Foundations, Design Guidelines, Techniques and Applications. Rita Borgo, Johannes Kehrer, David H.S. Chung, Eamonn Maguire, Robert S. Laramee, Helwig Hauser, Matthew Ward, and Min Chen. Eurographics State of the Art Reports, pp. 39-63. Eurographics, 2013.
- Multivariate Data Glyphs: Principles and Practice. Matthew O. Ward. In Handbook of Data Visualization, Computational Statistics, edited by Antony Unwin, Chun-houh Chen, and Wolfgang K. HÃ¤rdle, pp. 179-198.
- Information Visualization: Perception for Design Colin Ware. Third edition. Morgan Kaufmann, 2013.
- A Taxonomy of Glyph Placement Strategies for Multidimensional Data Visualization. Matthew O. Ward. Information Visualization 1:3-4 (2002), pp. 194-210.
- Taxonomy-Based Glyph Designâ€"With a Case Study on Visualizing Workflows of Biological Experiments. Eamonn Maguire, Philippe Rocca-Serra, Susanna-Assunta Sansone, Jim Davies, and Min Chen. IEEE TVCG (Proc. InfoVis 2012), pp. 2603-2612.
- Color Enhanced Star Plot Glyphsâ€"Can Salient Shape Characteristics Be Overcome? Alexander Klippel, Frank Hardisty, Rui Li, and Chris Weaver. Cartographica 44:3 (2009), 217-231.
- Brushing Scatterplots Richard A. Becker and William S. Cleveland. Technometrics 29 (1987), 127-142.
- Linked Data Views Graham J. Wills. Technometrics 29 (1987), 127-142.
- Get It Right in Black and White. Maureen Stone. Functional Color, 2010.
- Dot Plots: A Useful Alternative to Bar Charts. Naomi B. Robbins. 2006.

Ch 13: Reduce Items and Attributes

Reduce items and attributes

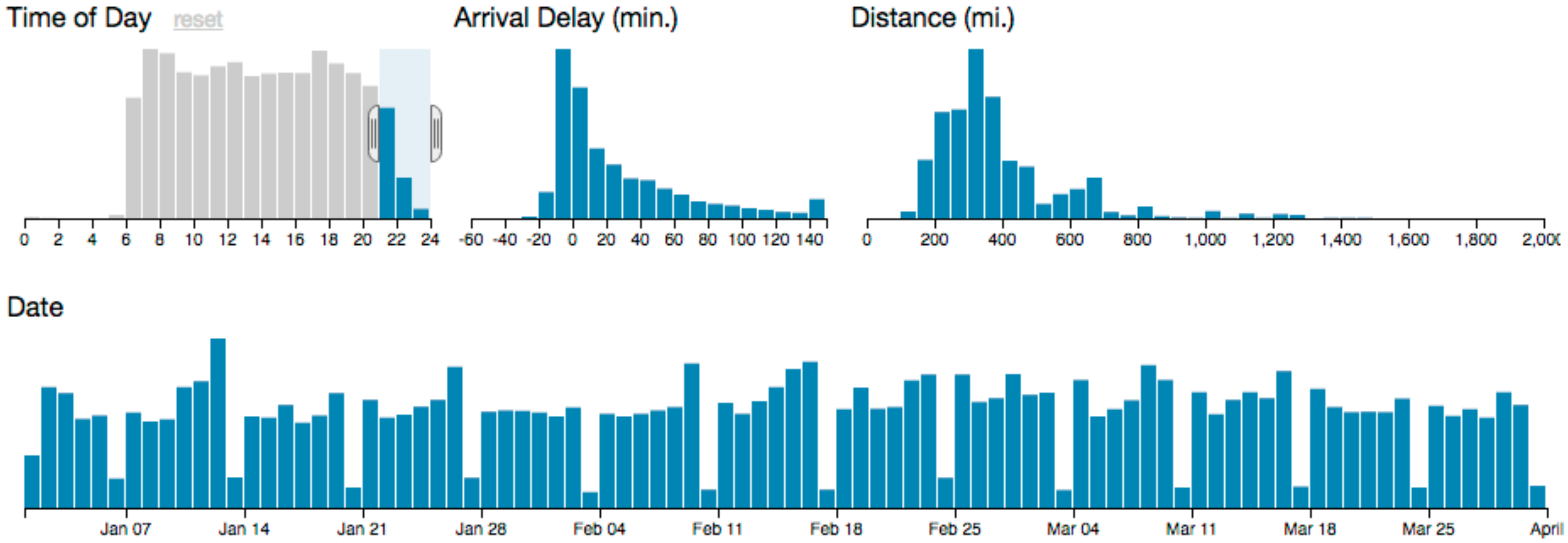
- reduce/increase: inverses
- filter
  - pro: straightforward and intuitive
    - to understand and compute
  - con: out of sight, out of mind
- aggregation
  - pro: inform about whole set
  - con: difficult to avoid losing signal
- not mutually exclusive
  - combine filter, aggregate
  - combine reduce,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





Idiom: **cross filtering** System: **Crossfilter**

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



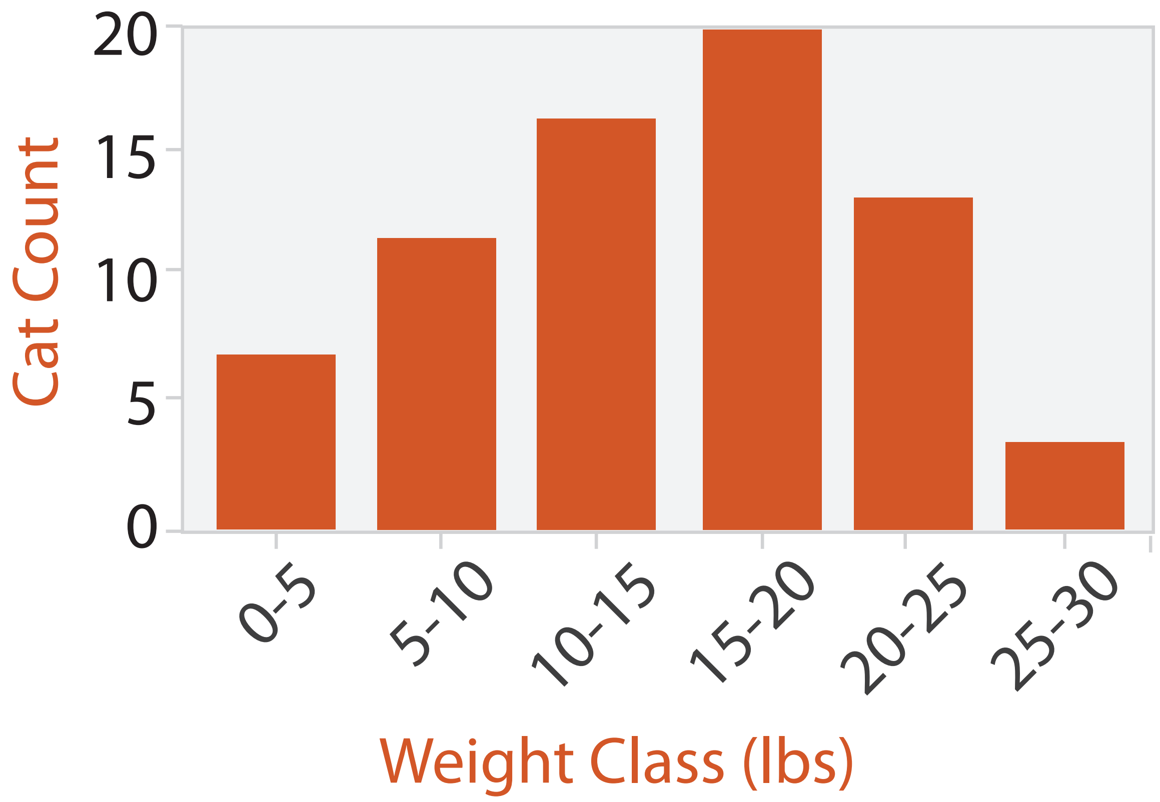
[http://square.github.io/crossfilter/]

Idiom: **cross filtering**

[https://www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html?_r=0]

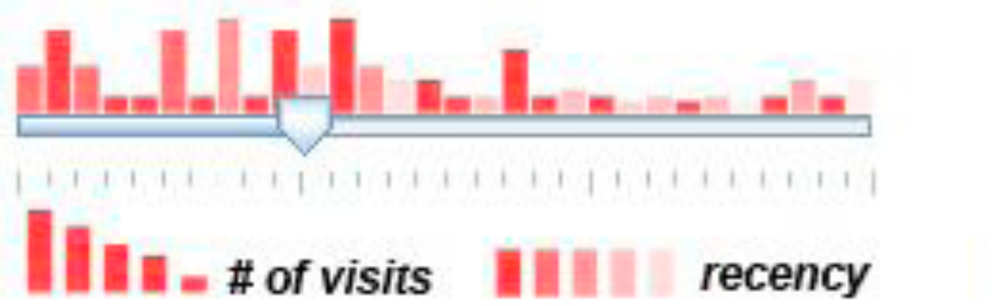
Idiom: **histogram**

- 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

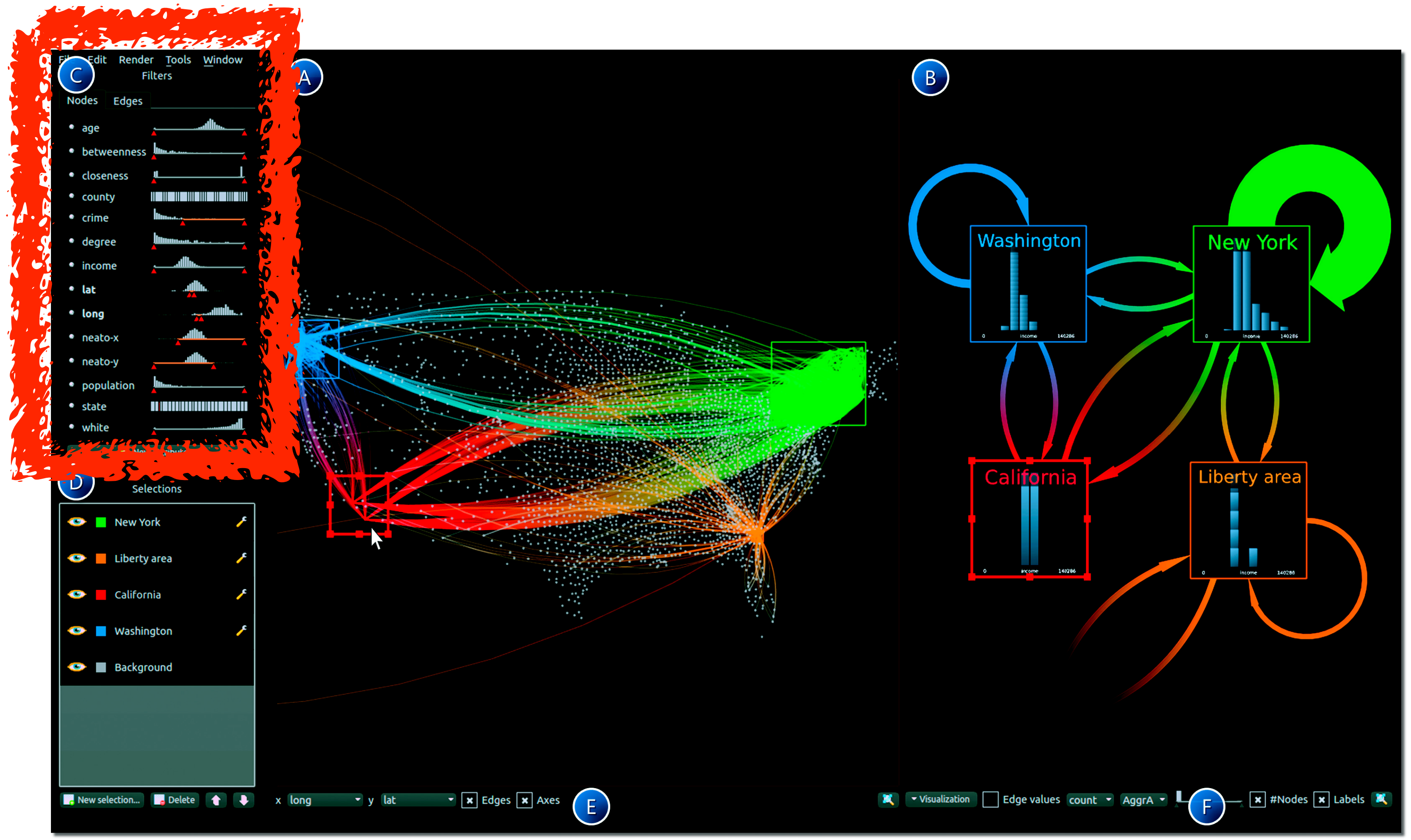


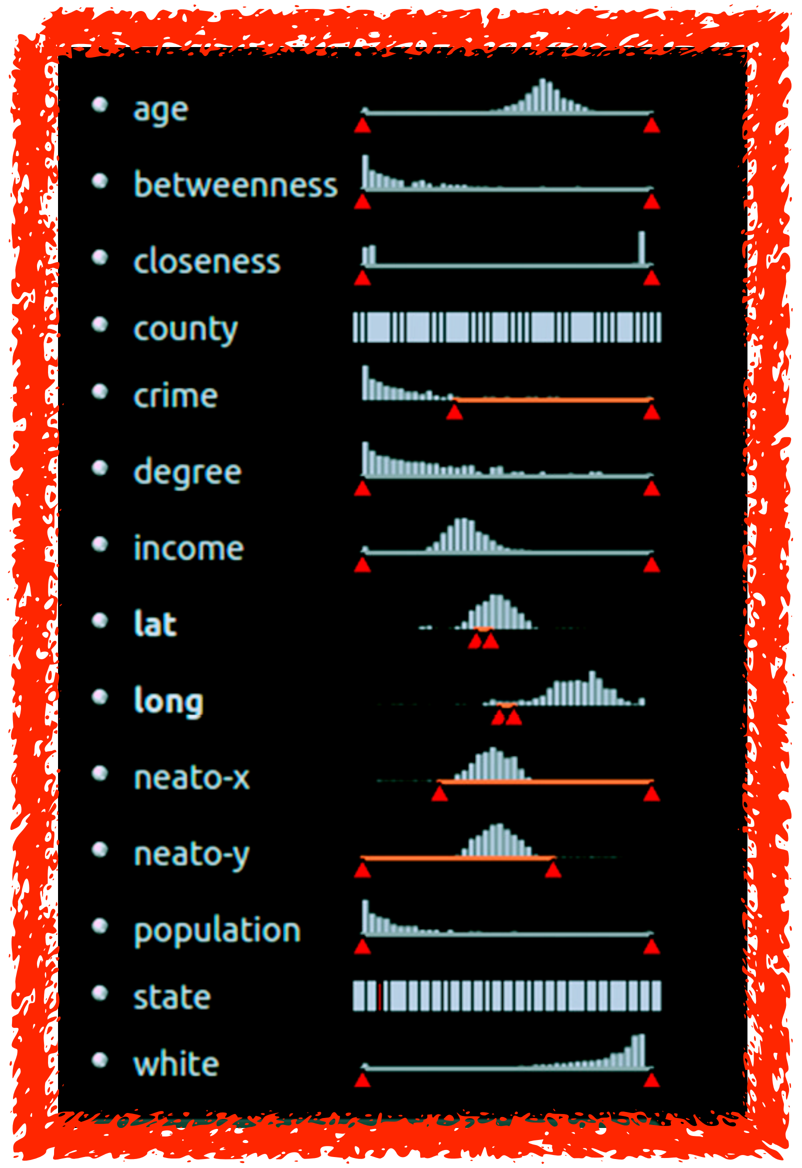
Idiom: **scented widgets**

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



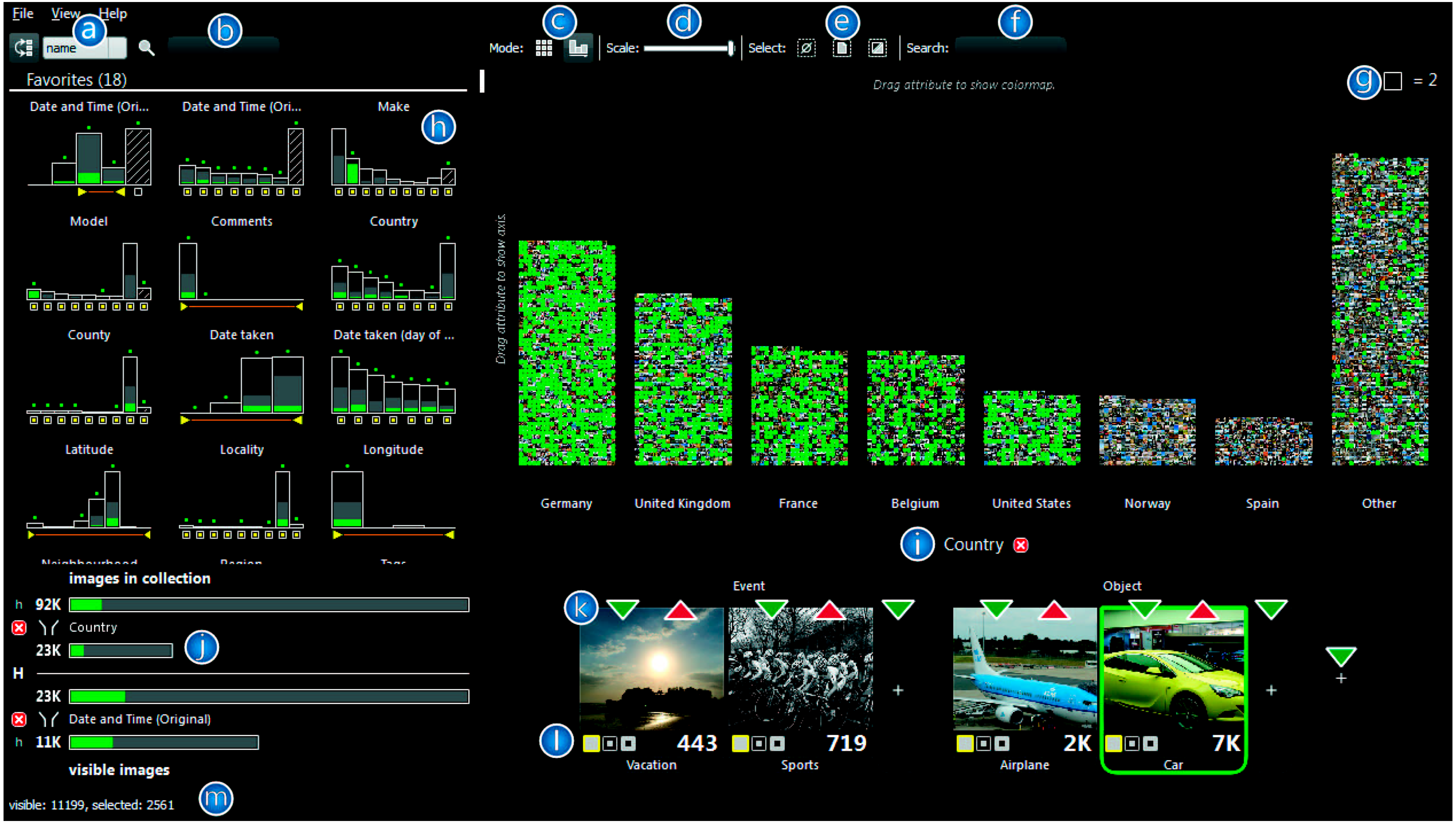
[Scented Widgets: Improving Navigation Cues with Embedded Visualizations. Willett, Heer, and Agrawala. IEEE TVCG (Proc. InfoVis 2007) 13:6 (2007), 1129–1136.]





[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).]

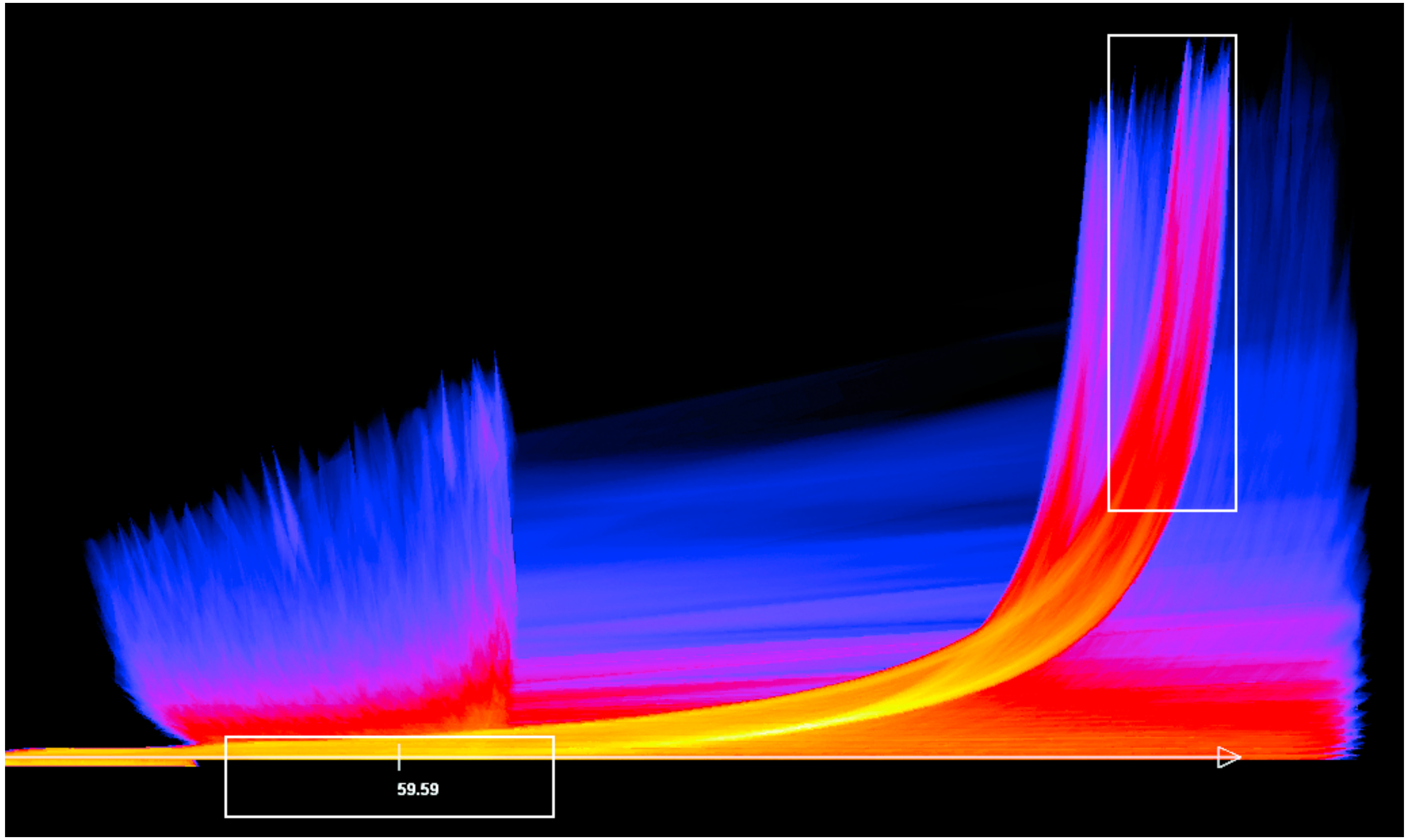
Scented histogram bisliders: detailed



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

Idiom: **Continuous scatterplot**

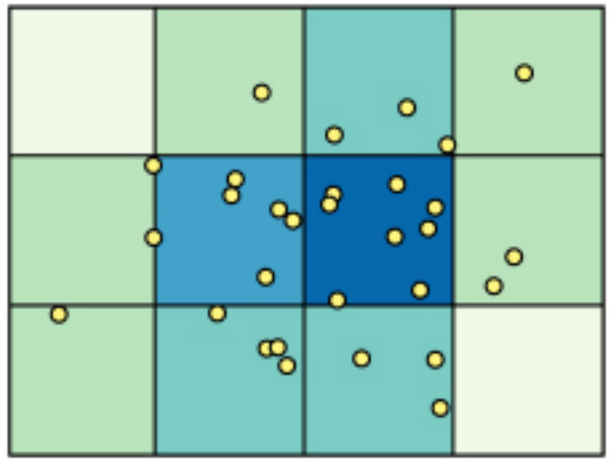
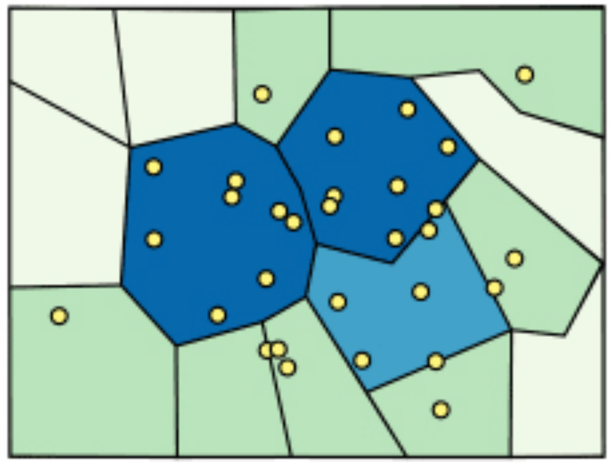
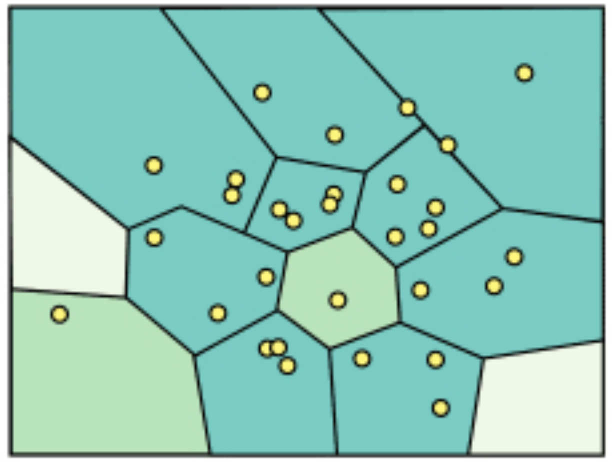
- static item aggregation
- data: table
- derived data: table
  - key attribs x,y for pixels
  - quant attrib: overplot density
- dense space-filling 2D matrix
- color: sequential categorical hue + ordered luminance colormap



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

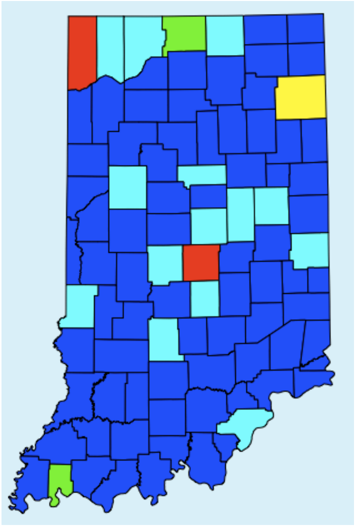
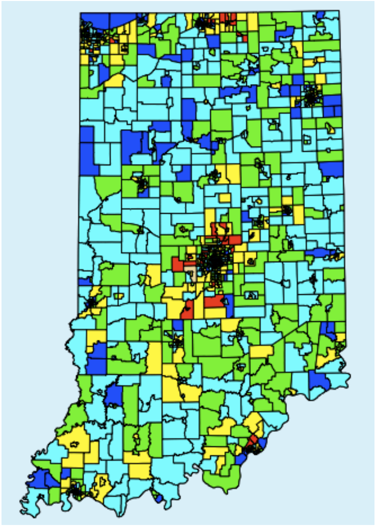
Spatial aggregation

- 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]

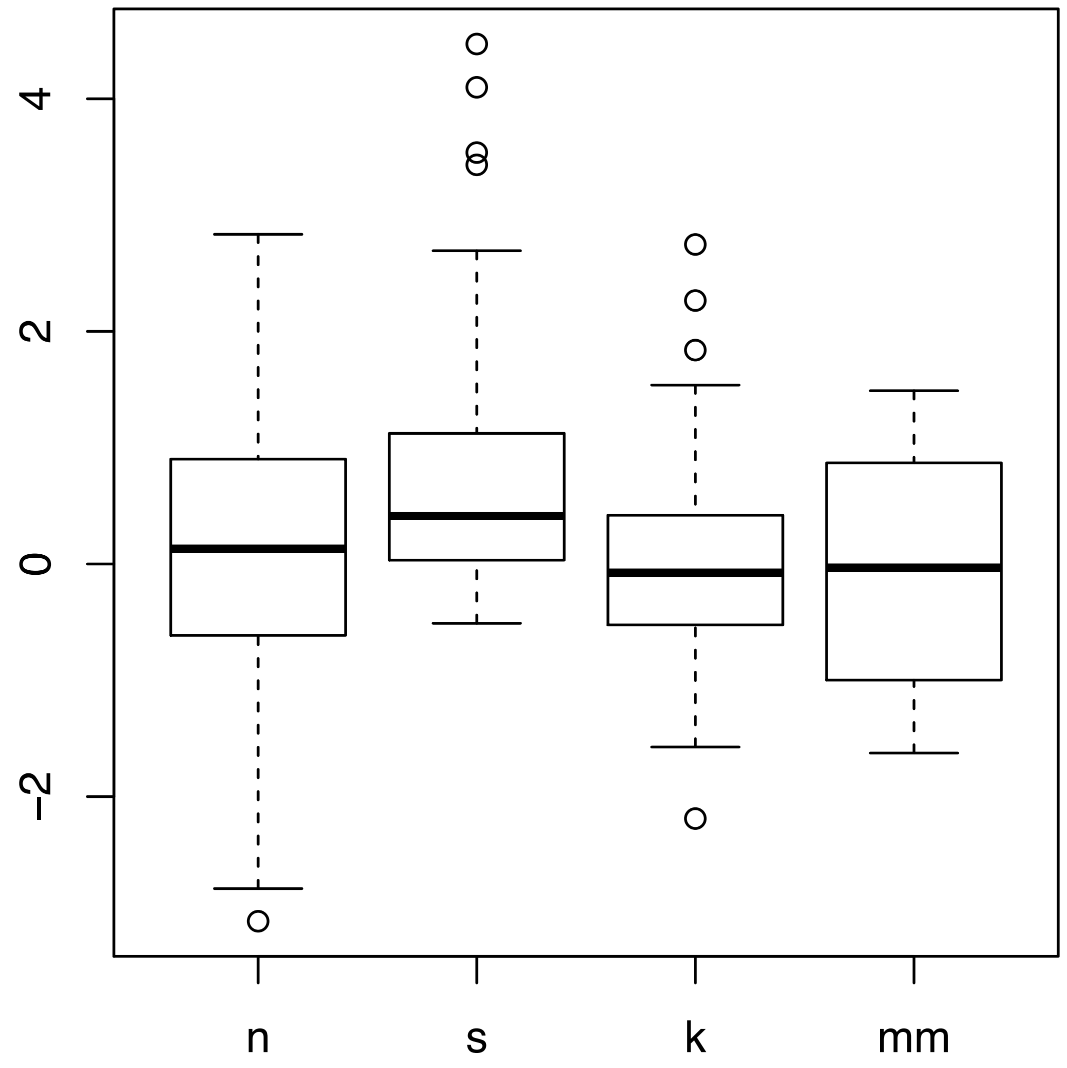
  - scale effects

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

Idiom: **boxplot**

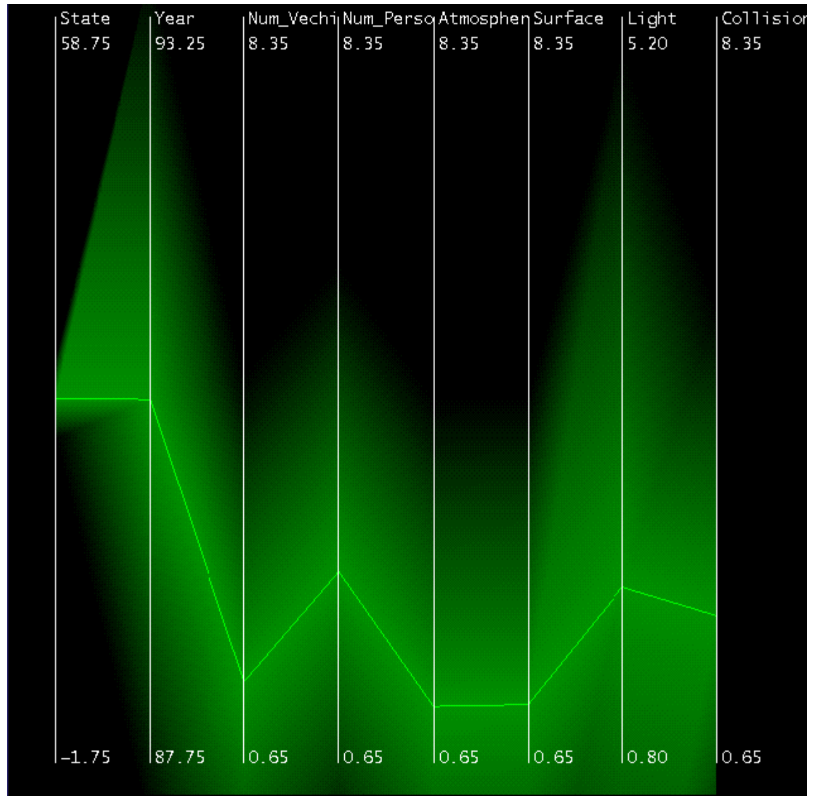
- static item aggregation
- task: find distribution
- data: table
- derived data
  - 5 quant attribs
    - median: central line
    - lower and upper quartile: boxes
    - lower upper fences: whiskers
      - values beyond which items are outliers
  - outliers beyond fence cutoffs explicitly shown

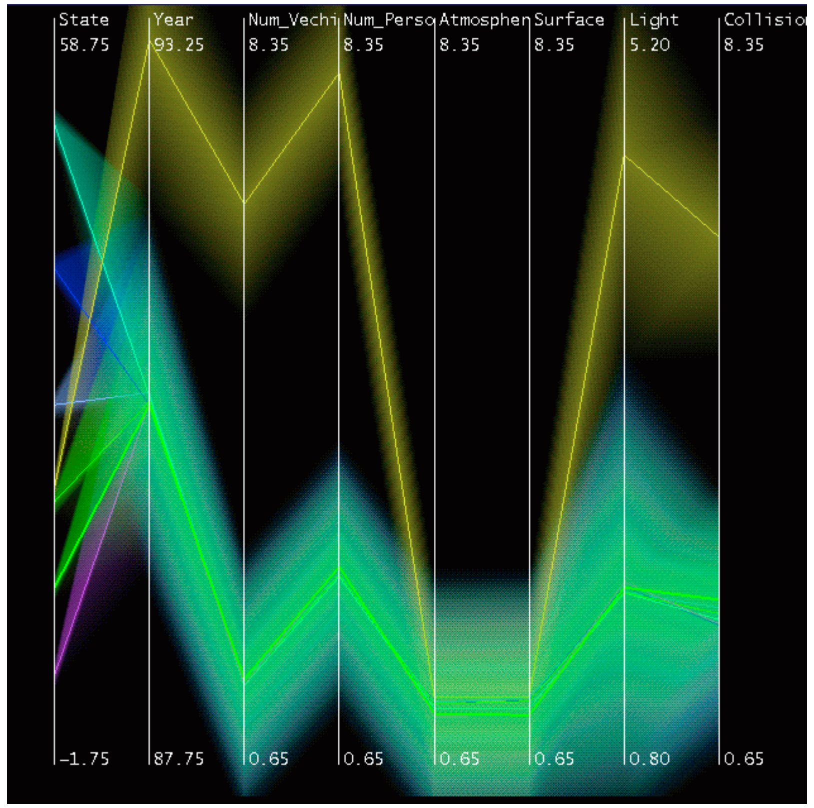


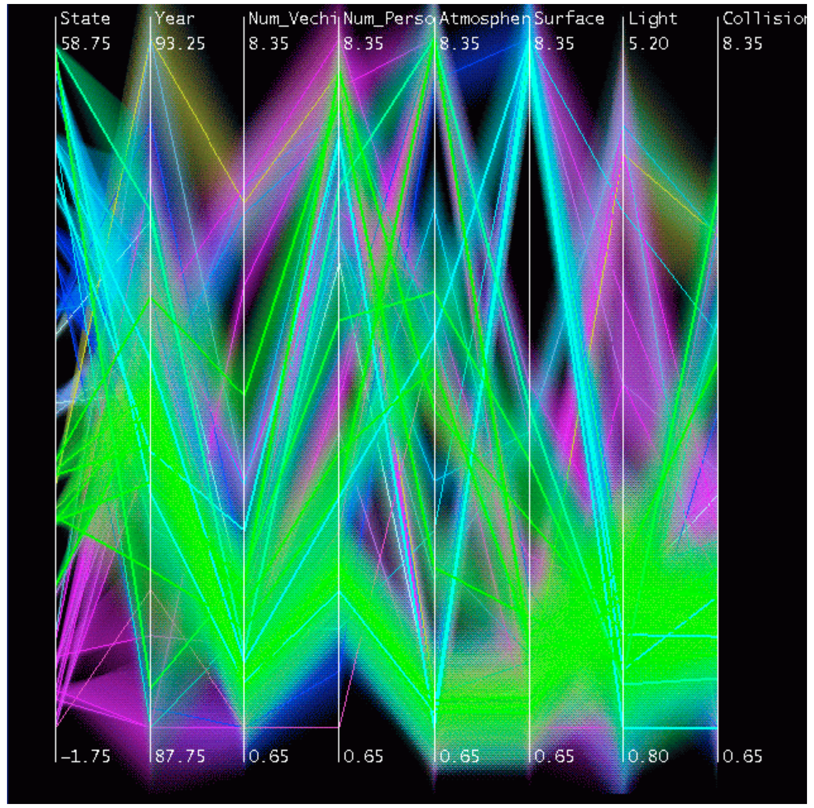
[40 years of boxplots. Wickham and Stryjewski. 2012. had.co.nz]

Idiom: **Hierarchical parallel coordinates**

- dynamic item aggregation
- derived data: ***hierarchical clustering***
- encoding:
  - –cluster band with variable transparency, line at mean, width by min/max values
  - –color by proximity in hierarchy



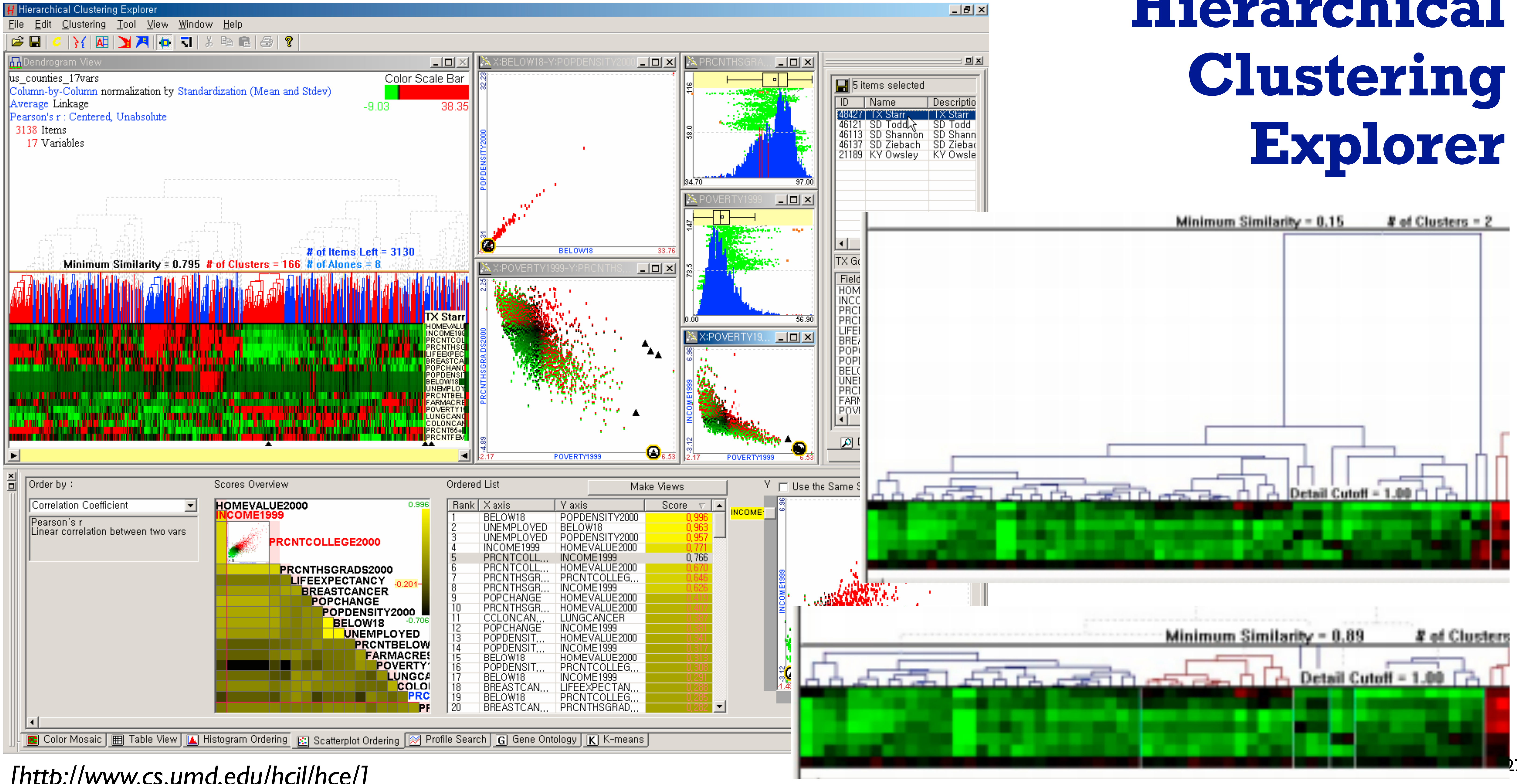




[Hierarchical Parallel Coordinates for Exploration of Large Datasets. Fua, Ward, and Rundensteiner. Proc. IEEE Visualization Conference (Vis '99), pp. 43– 50, 1999.]

Idiom: aggregation via hierarchical clustering (visible)

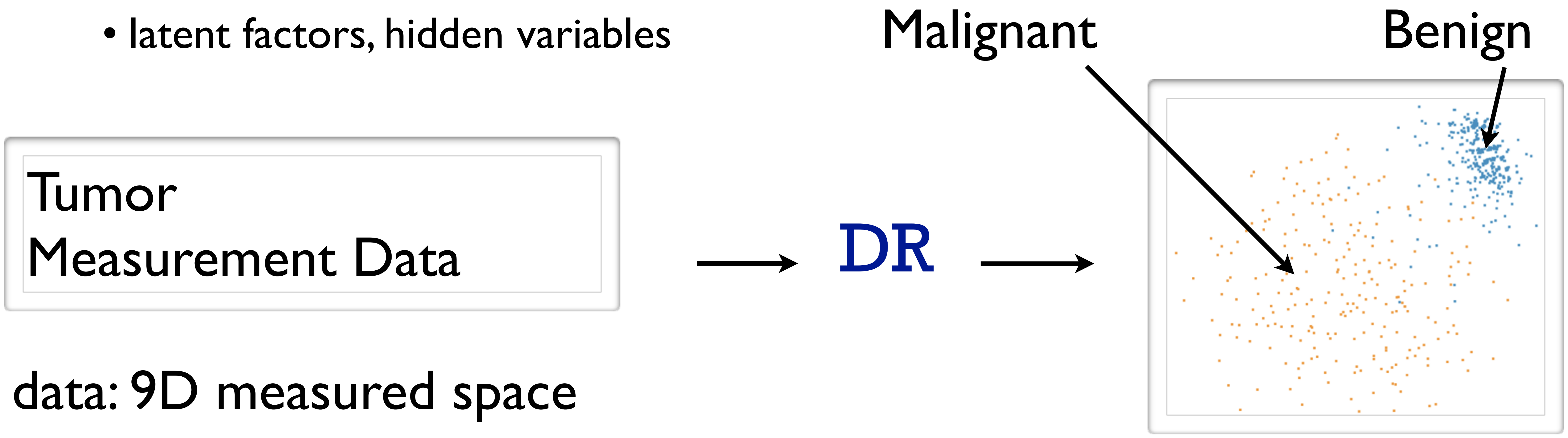
System: **Hierarchical Clustering Explorer**



[http://www.cs.umd.edu/hcil/hce/]

Dimensionality reduction

- attribute aggregation
  - derive low-dimensional target space from high-dimensional measured space
    - capture most of variance with minimal error
  - use when you can't directly measure what you care about
    - true dimensionality of dataset conjectured to be smaller than dimensionality of measurements
    - latent factors, hidden variables



data: 9D measured space

derived data: 2D target space

Dimensionality vs attribute reduction

- vocab use in field not consistent
  - dimension/attribute
- attribute reduction: reduce set with filtering
  - includes orthographic projection
- dimensionality reduction: create smaller set of new dims/attribs
  - typically implies dimensional aggregation, not just filtering
  - vocab: projection/mapping

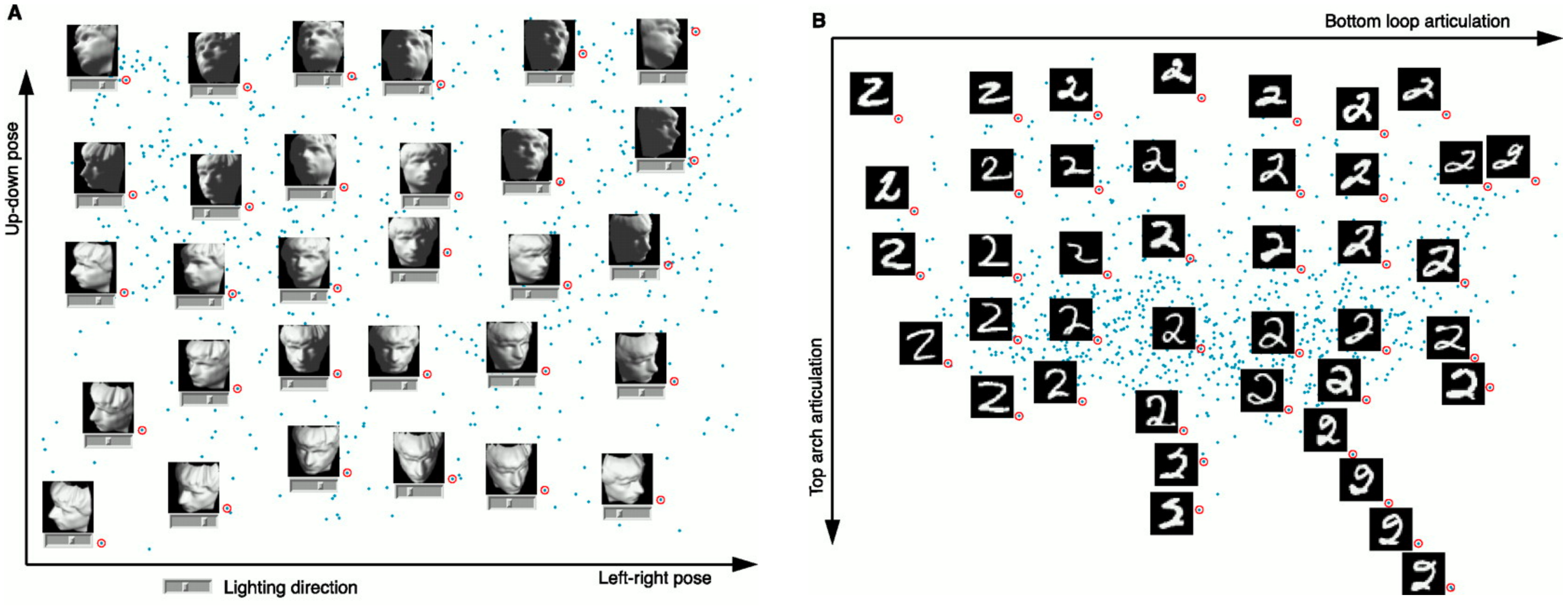
Dimensionality reduction & visualization

- why do people do DR?
  - improve performance of downstream algorithm
    - avoid curse of dimensionality
  - data analysis
    - if look at the output: visual data analysis
- abstract tasks when visualizing DR data
  - dimension-oriented tasks
    - naming synthesized dims, mapping synthesized dims to original dims
  - cluster-oriented tasks
    - verifying clusters, naming clusters, matching clusters and classes

[Visualizing Dimensionally-Reduced Data: Interviews with Analysts and a Characterization of Task Sequences. Brehmer, Sedlmair, Ingram, and Munzner. Proc. BELIV 2014.]

Dimension-oriented tasks

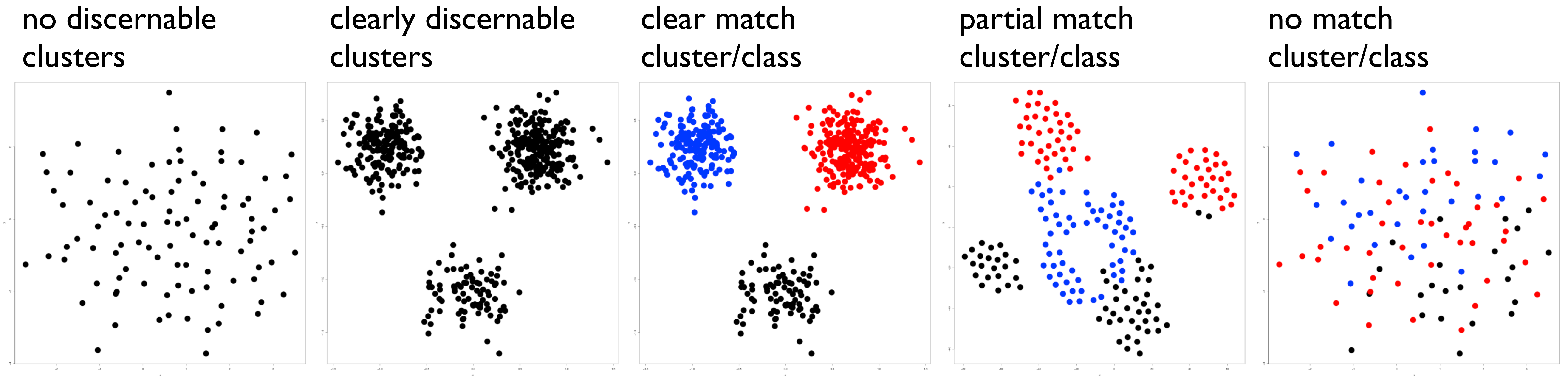
- naming synthesized dims: inspect data represented by lowD points



[A global geometric framework for nonlinear dimensionality reduction. Tenenbaum, de Silva, and Langford. Science, 290(5500):2319–2323, 2000.]

Cluster-oriented tasks

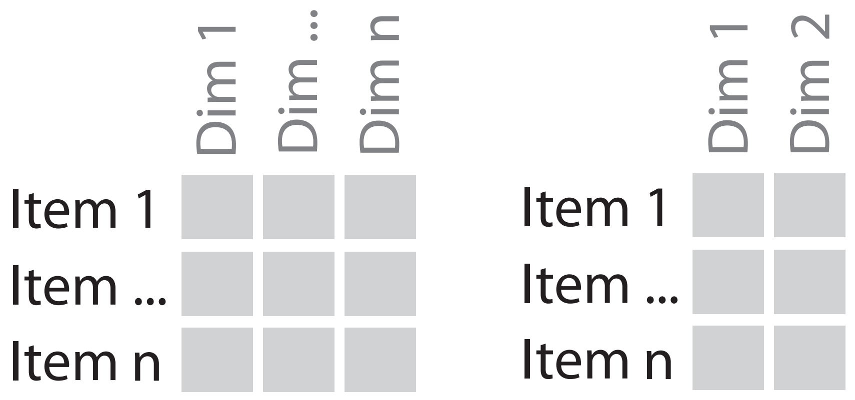
- verifying, naming, matching to classes



[Visualizing Dimensionally-Reduced Data: Interviews with Analysts and a Characterization of Task Sequences. Brehmer, Sedlmair, Ingram, and Munzner. Proc. BELIV 2014.]

Idiom: **Dimensionality reduction for documents**

Task 1



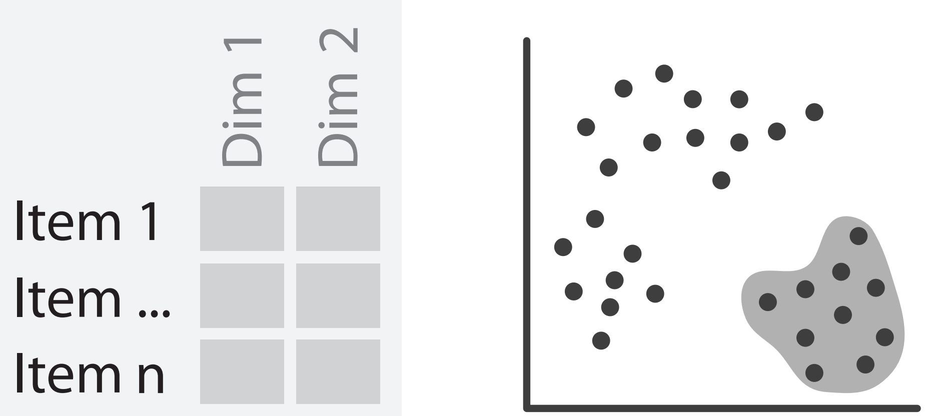
In
HD data

Out
2D data

What?
- In High-dimensional data
- Out 2D data

Why?
- Produce
- Derive

Task 2



In
2D data

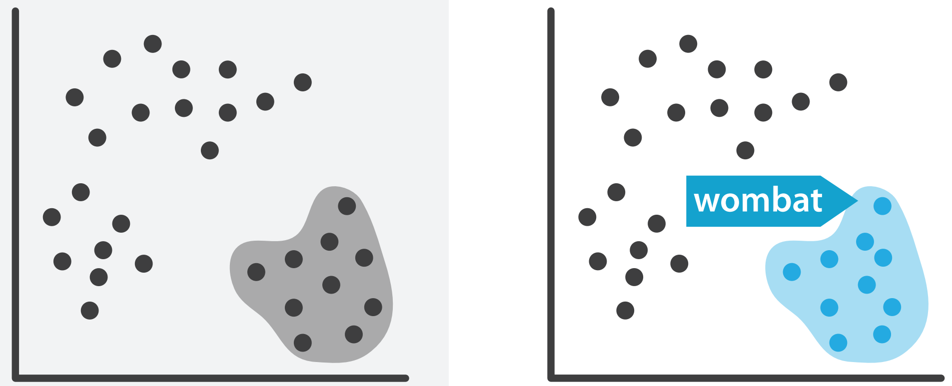
Out
Scatterplot
Clusters & points

What?
- In 2D data
- Out Scatterplot
- Out Clusters & points

Why?
- Discover
- Explore
- Identify

How?
- Encode
- Navigate
- Select

Task 3



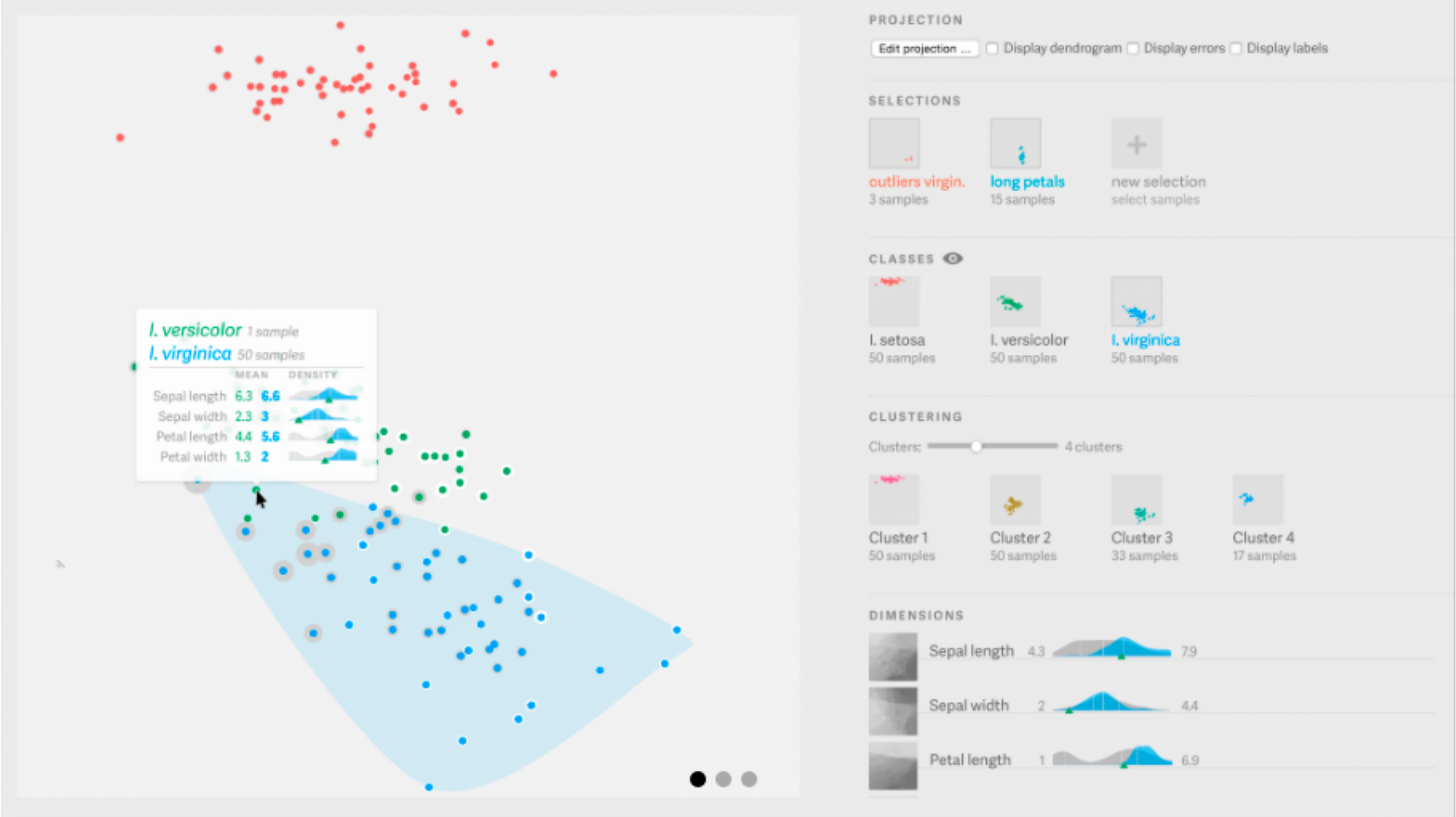
In
Scatterplot
Clusters & points

Out
Labels for clusters

What?
- In Scatterplot
- In Clusters & points
- Out Labels for clusters

Why?
- Produce
- Annotate

Interacting with dimensionally reduced data

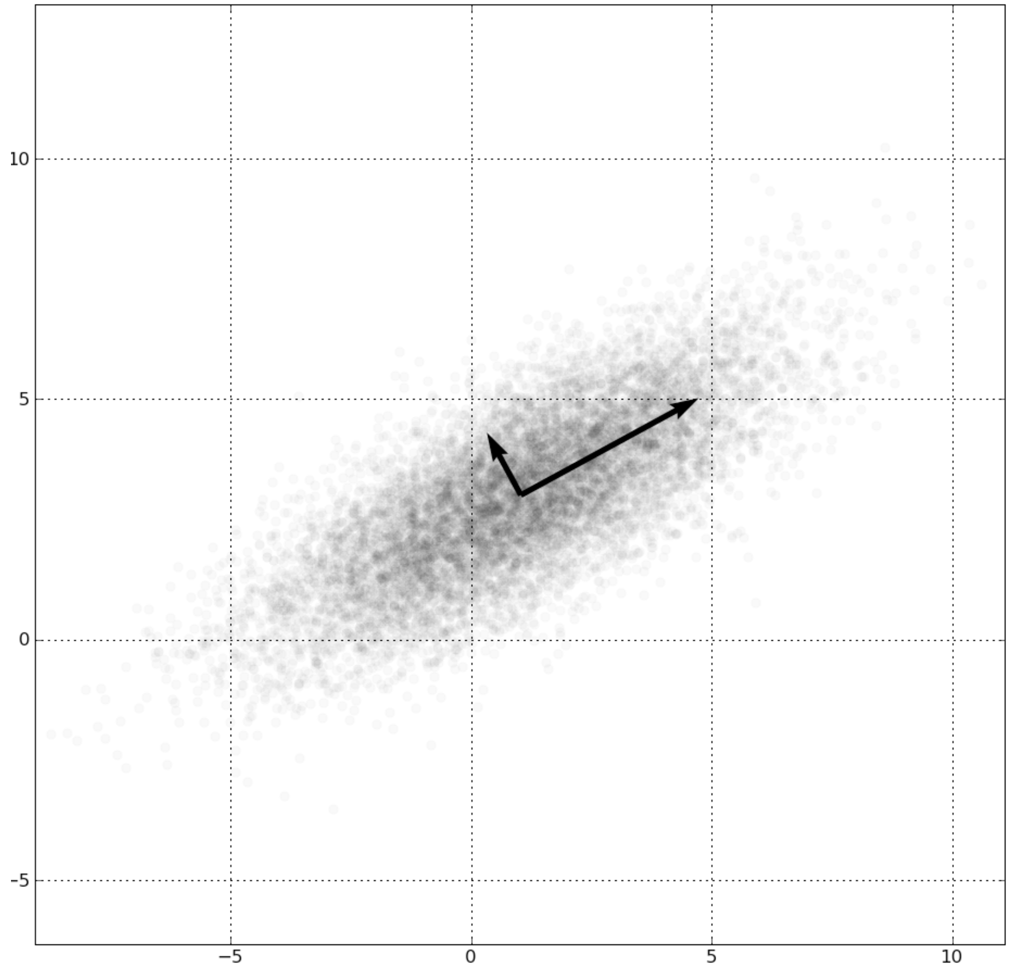


[https://uclab.fh-potsdam.de/projects/probing-projections/]

[Probing Projections: Interaction Techniques for Interpreting Arrangements and Errors of Dimensionality Reductions. Stahnke, Dörk, Müller, and Thom. IEEE TVCG (Proc. InfoVis 2015) 22(1):629-38 2016.]

Linear dimensionality reduction

- principal components analysis (PCA)
  - finding axes: first with most variance, second with next most, …
  - describe location of each point as linear combination of weights for each axis
    - mapping synthesized dims to original dims



[http://en.wikipedia.org/wiki/File:GaussianScatterPCA.png]

Nonlinear dimensionality reduction

- pro: can handle curved rather than linear structure
- cons: lose all ties to original dims/attribs
  - new dimensions often cannot be easily related to originals
    - mapping synthesized dims to original dims task is difficult
- many techniques proposed
  - many literatures: visualization, machine learning, optimization, psychology, ...
  - techniques: t-SNE, MDS (multidimensional scaling), charting, isomap, LLE,…
    - t-SNE: excellent for clusters
      - but some trickiness remains: http://distill.pub/2016/misread-tsne/
    - MDS: confusingly, entire family of techniques, both linear and nonlinear
      - minimize stress or strain metrics
      - early formulations equivalent to PCA

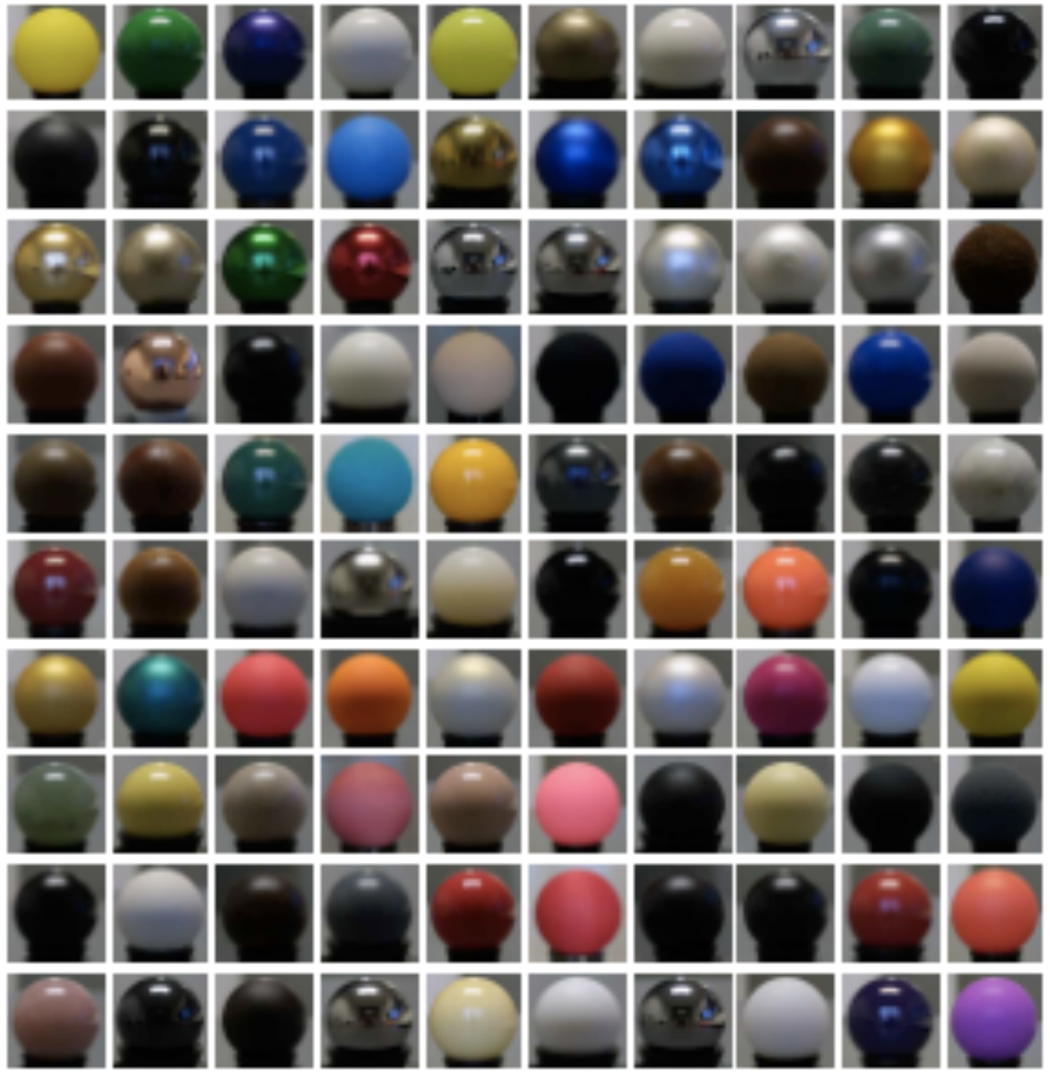
VDA with DR example: nonlinear vs linear

- DR for computer graphics reflectance model
  - goal: simulate how light bounces off materials to make realistic pictures
    - computer graphics: BRDF (reflectance)
  - idea: measure what light does with real materials

[Fig 2. Matusik, Pfister, Brand, and McMillan. A Data-Driven Reflectance Model. SIGGRAPH 2003]

Capturing & using material reflectance

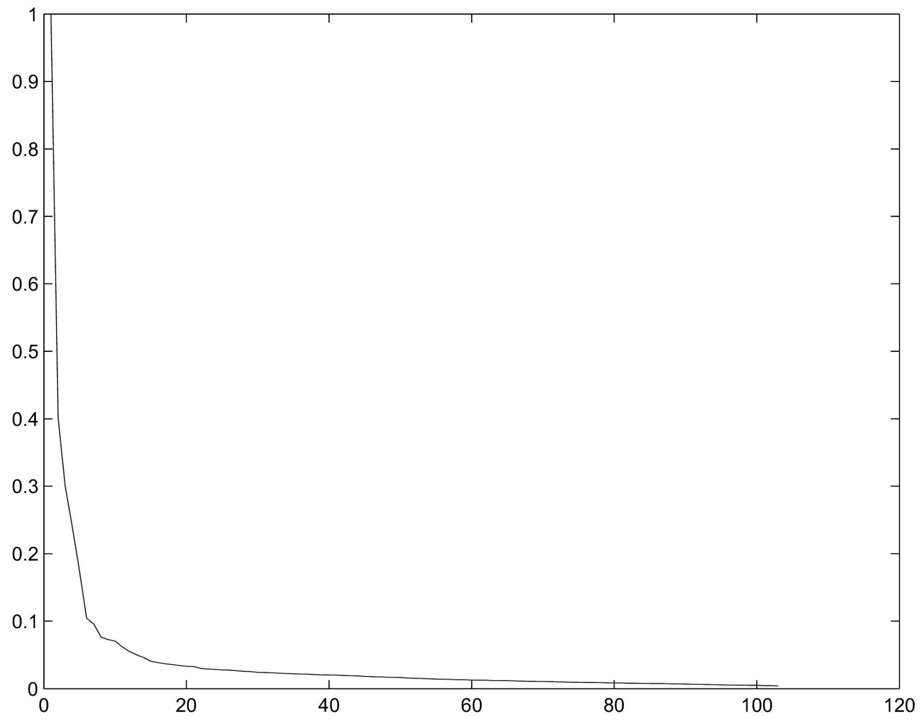
- reflectance measurement: interaction of light with real materials (spheres)
- result: 104 high-res images of material
  - each image 4M pixels
- goal: image synthesis
  - simulate completely new materials
- need for more concise model
  - 104 materials * 4M pixels = 400M dims
  - want concise model with meaningful knobs
    - how shiny/greasy/metallic
    - DR to the rescue!

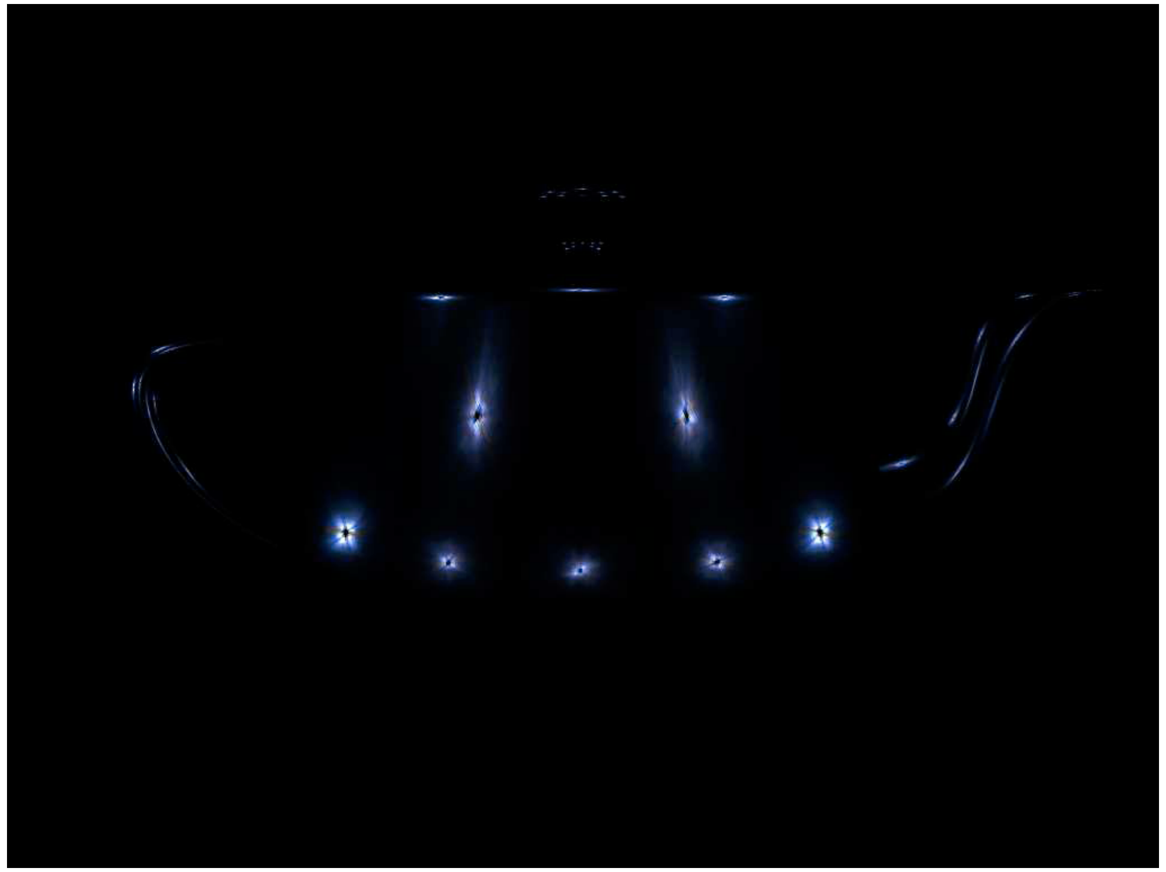


[Figs 5/6. Matusik et al. A Data-Driven Reflectance Model. SIGGRAPH 2003]

Linear DR

- first try: PCA (linear)
- result: error falls off sharply after ~45 dimensions
  - scree plots: error vs number of dimensions in lowD projection
- problem: physically impossible intermediate points when simulating new materials
  - specular highlights cannot have holes!

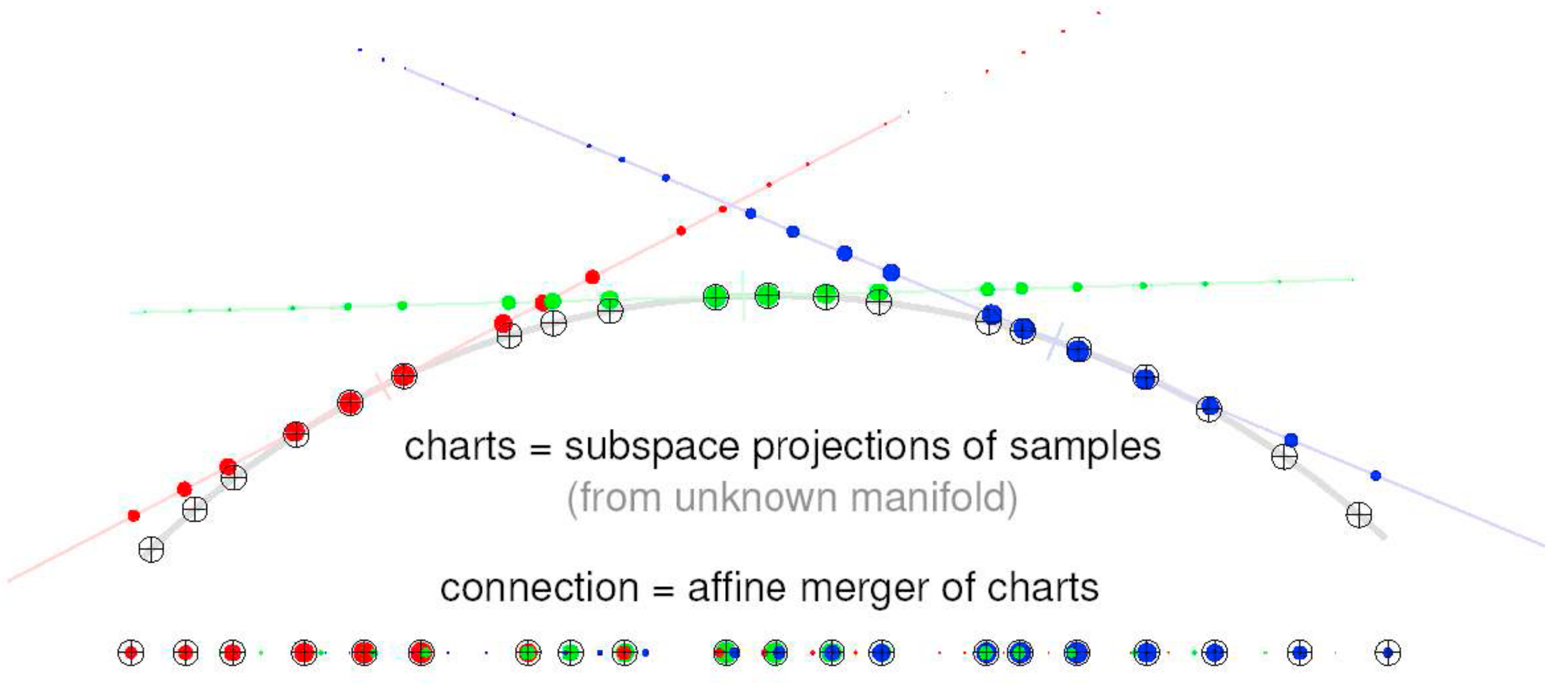




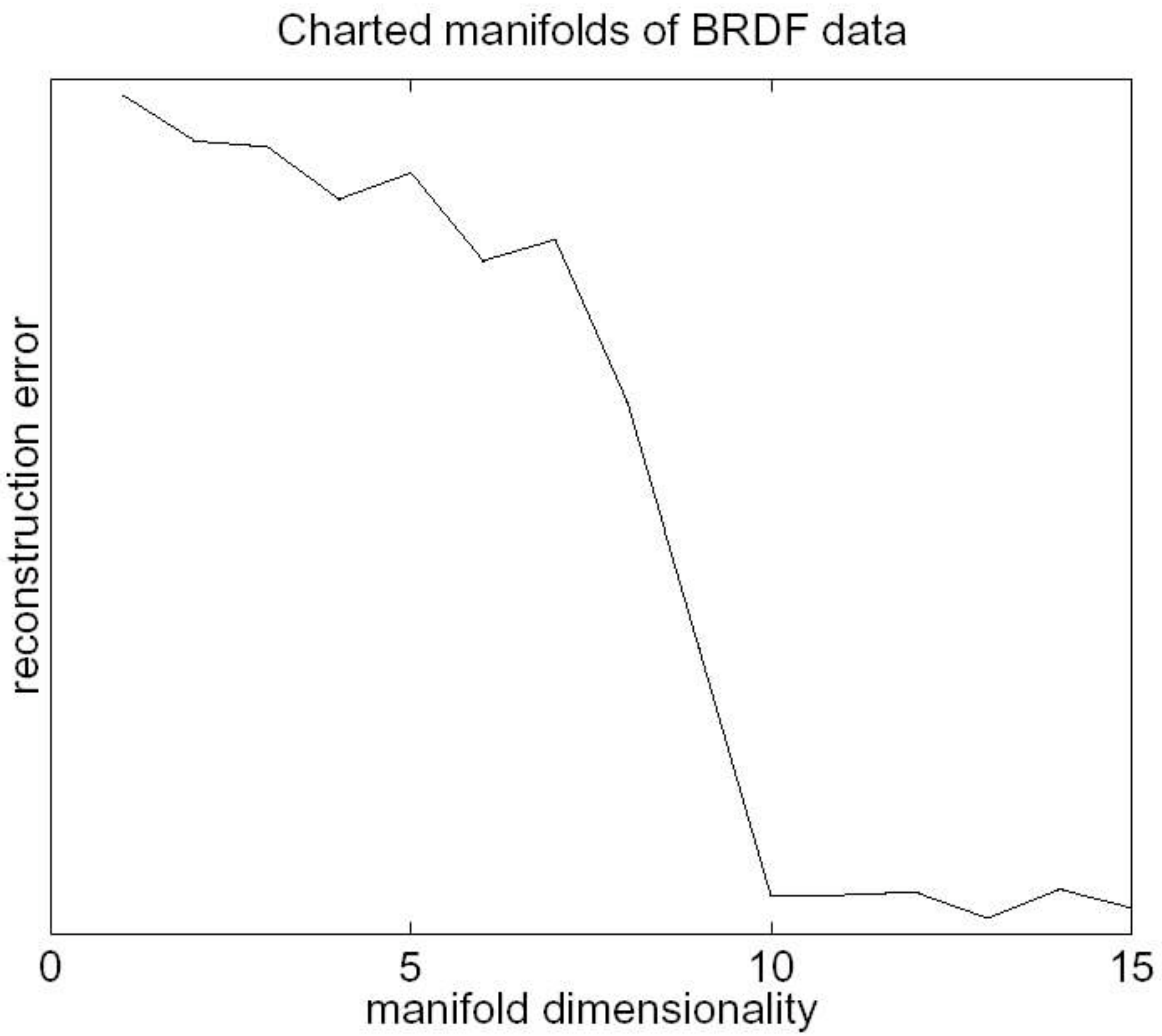
[Figs 6/7. Matusik et al. A Data-Driven Reflectance Model. SIGGRAPH 2003]

Nonlinear DR

- second try: charting (nonlinear DR technique)
  - scree plot suggests 10-15 dims
  - note: dim estimate depends on technique used!

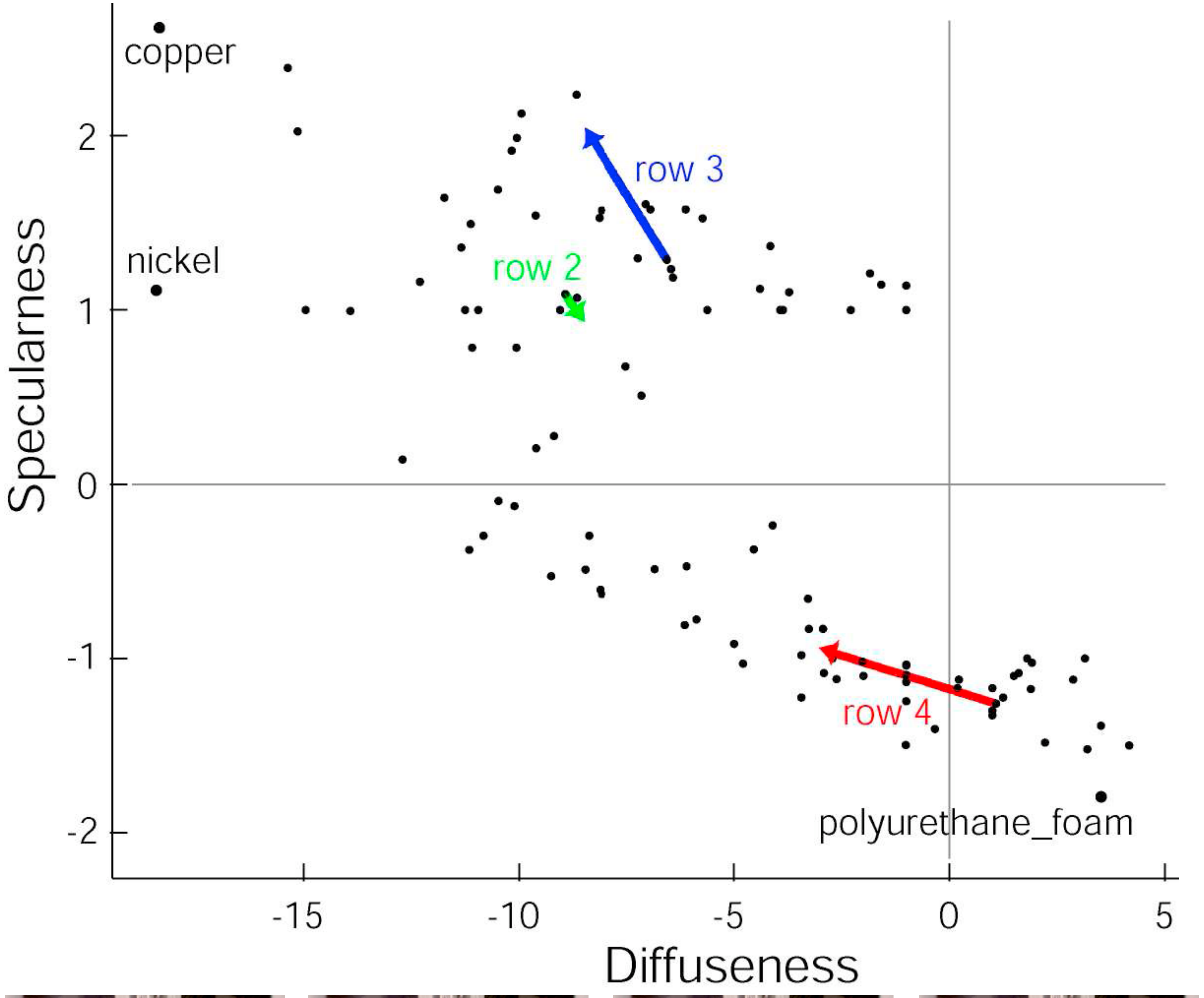


[Fig 10/11. Matusik et al. A Data-Driven Reflectance Model. SIGGRAPH 2003]



Finding semantics for synthetic dimensions

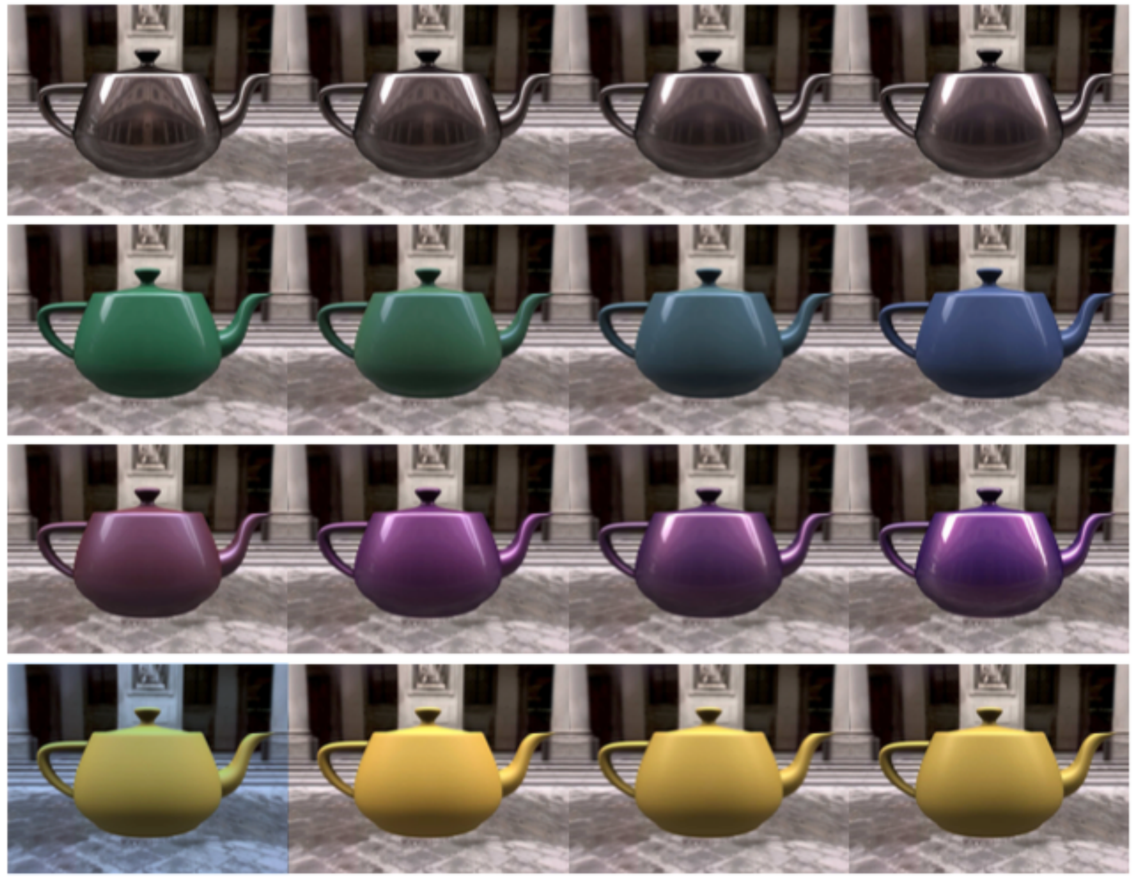
- look for meaning in scatterplots
  - synthetic dims created by algorithm but named by human analysts
  - points represent real-world images (spheres)
  - people inspect images corresponding to points to decide if axis could have meaningful name
- cross-check meaning
  - arrows show simulated images (teapots) made from model
  - check if those match dimension semantics



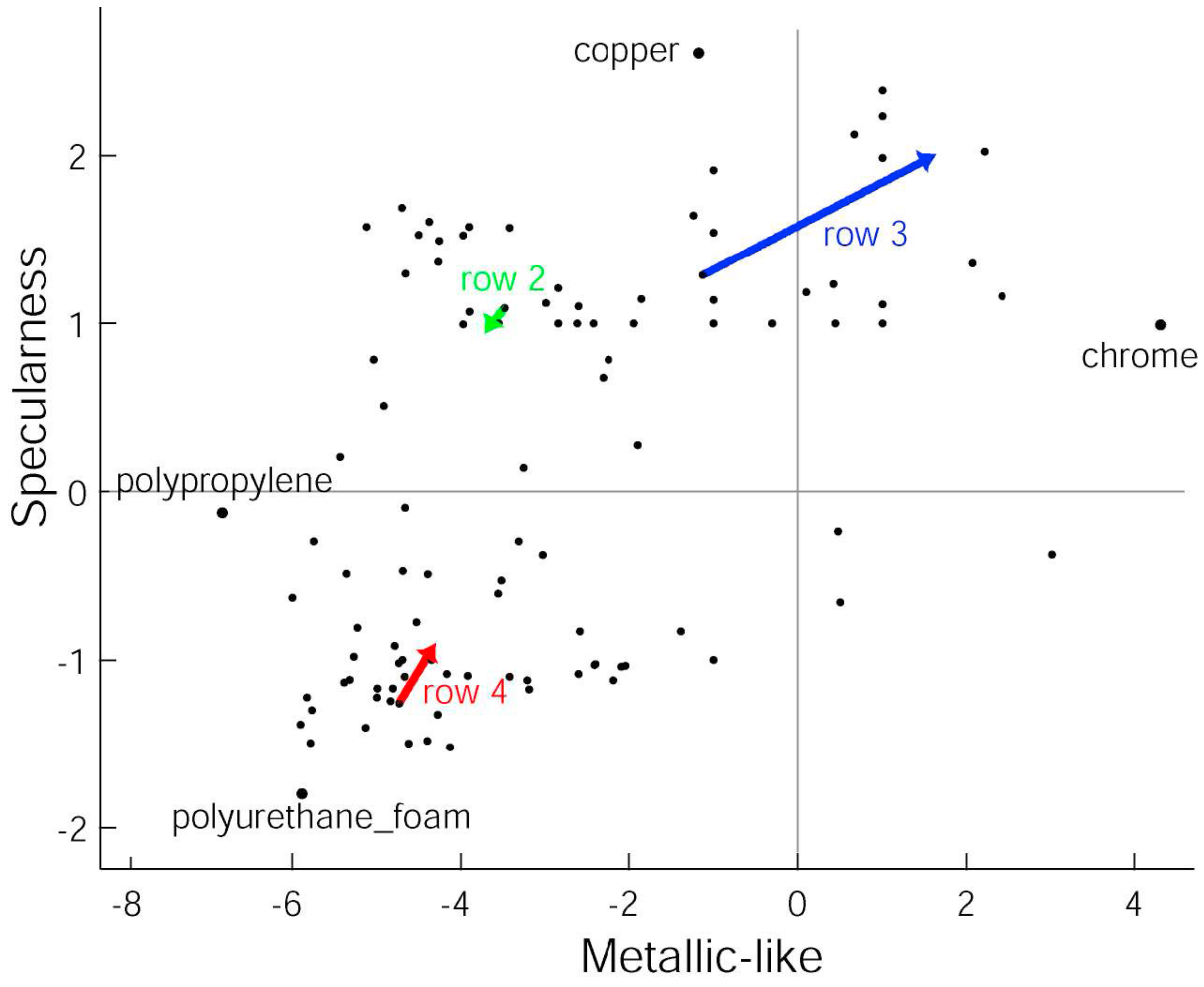
row 4

[Fig 12/16. Matusik et al. A Data-Driven Reflectance Model. SIGGRAPH 2003]

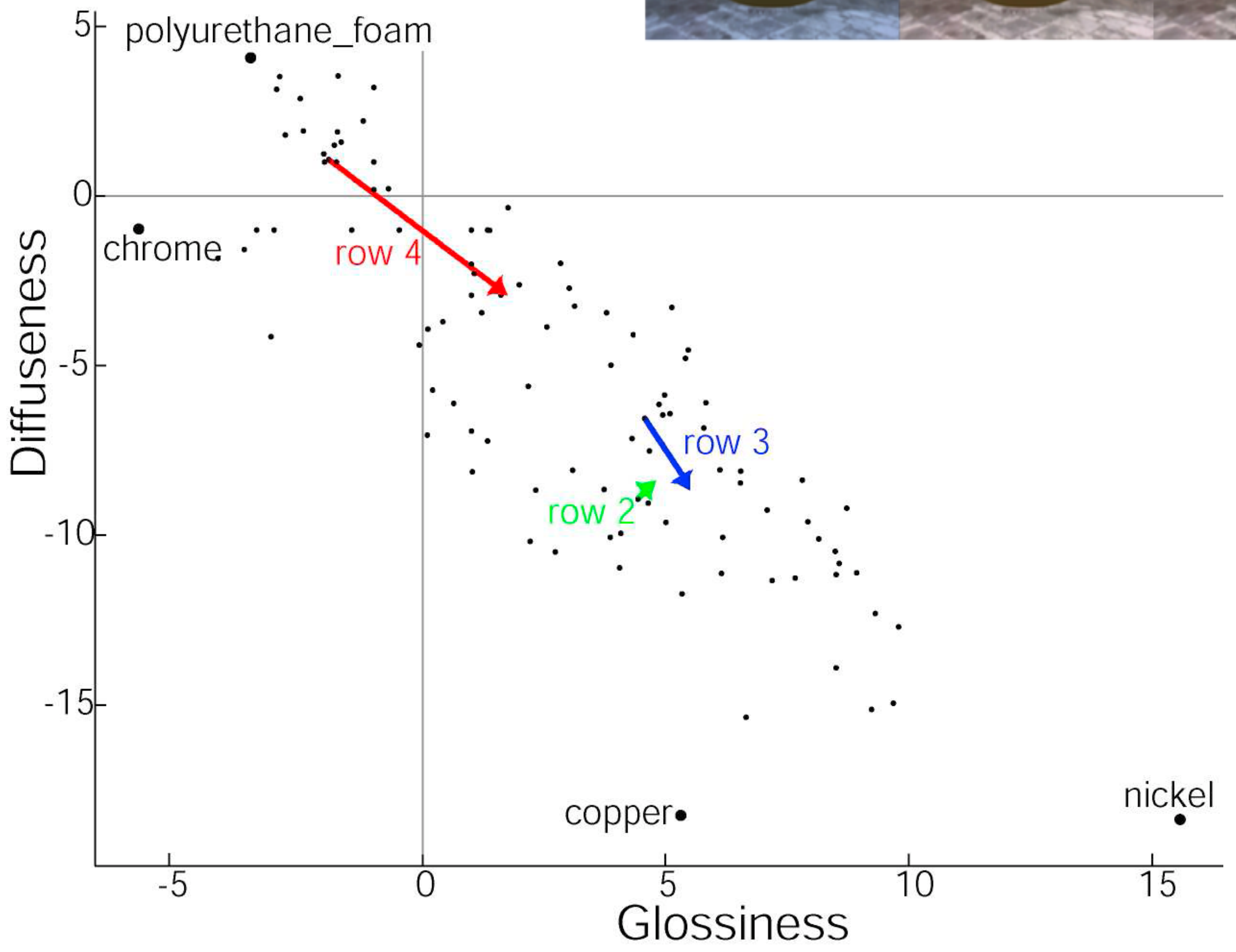
Understanding synthetic dimensions



Specular-Metallic



Diffuseness-Glossiness



[Fig 13/14/16. Matusik et al. A Data-Driven Reflectance Model. SIGGRAPH 2003]

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.

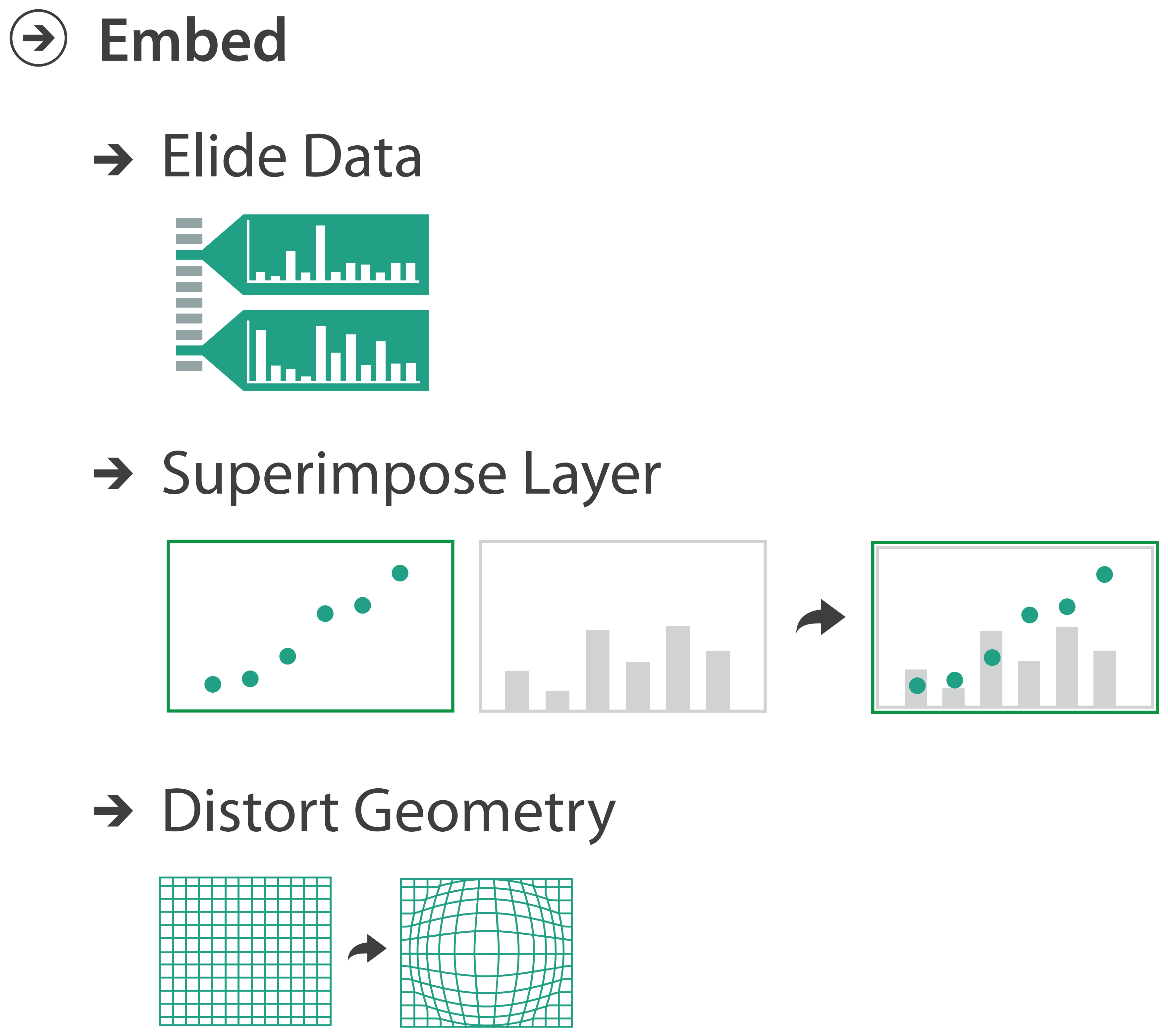
Further reading, full

- Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. Chris Ahlberg and Ben Shneiderman. Proc. CHI, pp. 313-317. ACM, 1994.
- Scented Widgets: Improving Navigation Cues with Embedded Visualizations. Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. IEEE TVCG (Proc. InfoVis 2007), pp. 1129-1136.
- Information Foraging Theory: Adaptive Interaction with Information. Peter Pirolli. Oxford University Press, 2007.
- Exploratory Data Analysis. John W. Tukey. Addison-Wesley, 1977.
- 40 Years of Boxplots Hadley Wickham and Lisa Stryjewski. Technical report, had.co.nz, 2012.
- Hierarchical Aggregation for Information Visualization: Overview, Techniques and Design Guidelines. Niklas Elmqvist and Jean-Daniel Fekete. IEEE TVCG (2010), pp. 439-454.
- Hierarchical Parallel Coordinates for Exploration of Large Datasets. Ying-Huey Fua, Matthew O. Ward, and Elke A. Rundensteiner. Proc. IEEE Conference on Visualization (Vis), pp. 43-50. IEEE Computer Society, 1999.
- The Modifiable Areal Unit Problem (MAUP) David Wong. In The SAGE Handbook of Spatial Analysis, edited by A. Stewart Fotheringham and Peter A. Rogerson, pp. 105-123. Sage, 2009.
- The Modifiable Areal Unit Problem. Stan Openshaw. Number 38 in Concepts and Techniques in Modern Geography, Geo Books, 1984.
- Geographically Weighted Visualization: Interactive Graphics for Scale-Varying Exploratory Analysis. Jason Dykes and Chris Brunsdon. IEEE TVCG (Proc. InfoVis 2007), pp. 1161-1168.
- Interactive Hierarchical Dimension Ordering, Spacing and Filtering for Exploration of High Dimensional Datasets. Jing Yang, Wei Peng, Matthew O. Ward, and Elke A. Rundensteiner. Proc. InfoVis, pp. 105-112. IEEE Computer Society, 2003.
- Clutter Reduction in Multi-Dimensional Data Visualization Using Dimension Reordering. Wei Peng, Matthew O. Ward, and Elke A. Rundensteiner. Proc. InfoVis, pp. 89-96. IEEE Computer Society, 2004.
- Value and Relation Display for Interactive Exploration of High Dimensional Datasets. Jing Yang, Anilkumar Patro, Shiping Huang, Nishant Mehta, Matthew O. Ward, and Elke A. Rundensteiner. Proc. InfoVis, pp. 73-80. IEEE Computer Society, 2004.
- Visual Hierarchical Dimension Reduction for Exploration of High Dimensional Datasets. Jing Yang, Matthew O. Ward, Elke A. Rundensteiner, and Shiping Huang. Proc. Eurographics/IEEE Symposium on Visualization (VisSym), pp. 19-28. Eurographics, 2003.
- DimStiller: Workflows for Dimensional Analysis and Reduction. Stephen Ingram, Tamara Munzner, Veronika Irvine, Melanie Tory, Steven Bergner, and Torsten Moeller. Proc. VAST, pp. 3-10. IEEE Computer Society, 2010.
- Similarity Clustering of Dimensions for an Enhanced Visualization of Multidimensional Data. Michael Ankerst, Stefan Berchtold, and Daniel A. Keim. Proc. InfoVis, pp. 52-60. IEEE Computer Society, 1998.
- Discussion of a Set of Points in Terms of Their Mutual Distances. G. Young and A. S. Householder. Psychometrika 3:1.
- Multidimensional Scaling: I. Theory and Method. W. S. Torgerson. Psychometrika 17 (1952), pp. 401-419.
- A Linear Iteration Time Layout Algorithm for Visualising High Dimensional Data. M. Chalmers. Proc. IEEE VIS, pp. 127-132. IEEE Computer Society, 1996.
- Glimmer: Multilevel MDS on the GPU. Stephen Ingram, Tamara Munzner, and Marc Olano. IEEE TVCG (2009), pp. 249-261.
- Empirical Guidance on Scatterplot and Dimension Reduction Technique Choices. Michael Sedlmair, Tamara Munzner, and Melanie Tory. IEEE TVCG (Proc. InfoVis 2013), pp. 2634-2643.

Ch 14. Embed: Focus+Context

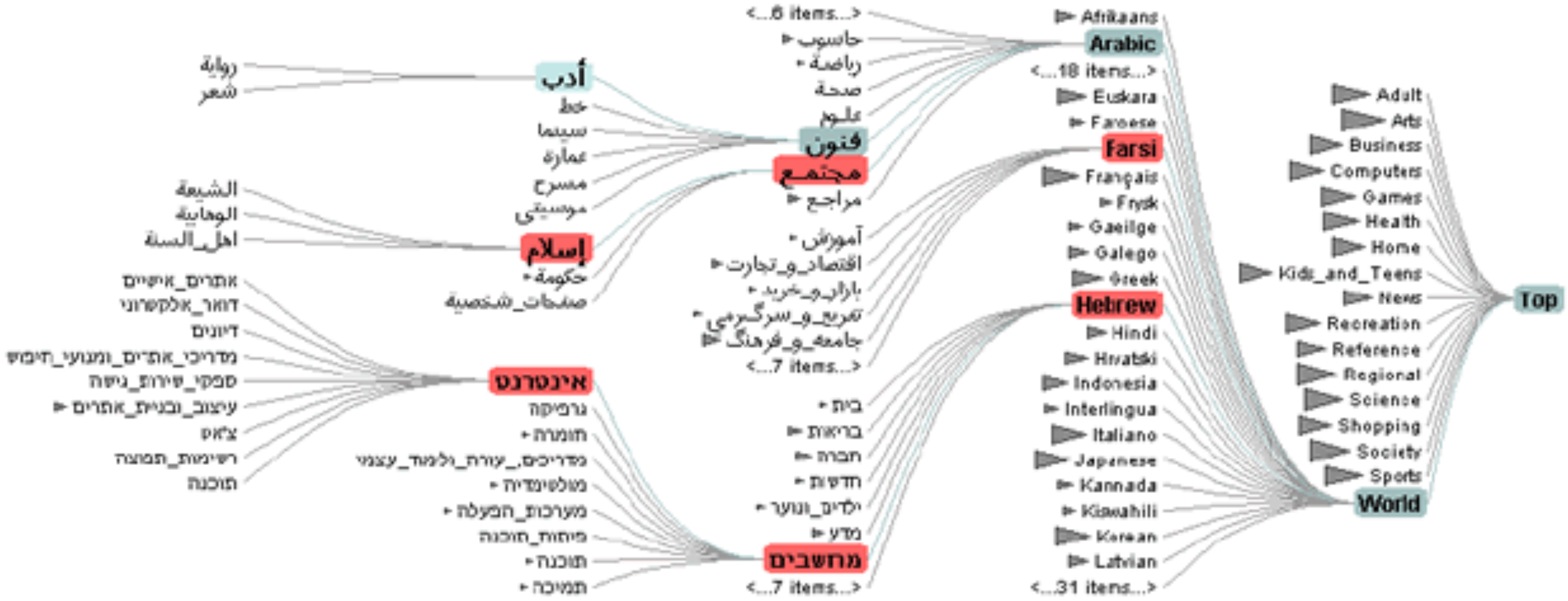
Embed: Focus+Context

- combine information within single view
- elide
  - –selectively filter and aggregate
- superimpose layer
  - –local lens
- distortion design choices
  - –region shape: radial, rectilinear, complex
  - –how many regions: one, many
  - –region extent: local, global
  - –interaction metaphor



Idiom: **DOITrees Revisited**

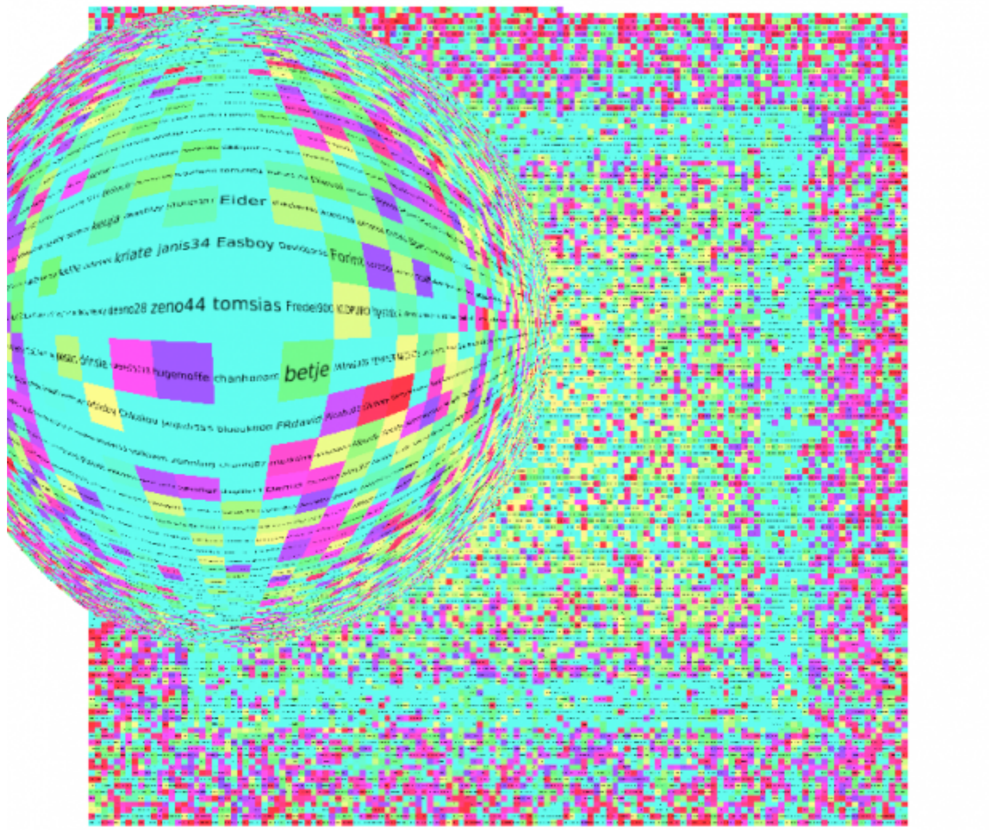
- elide
  - some items dynamically filtered out
  - some items dynamically aggregated together
  - some items shown in detail

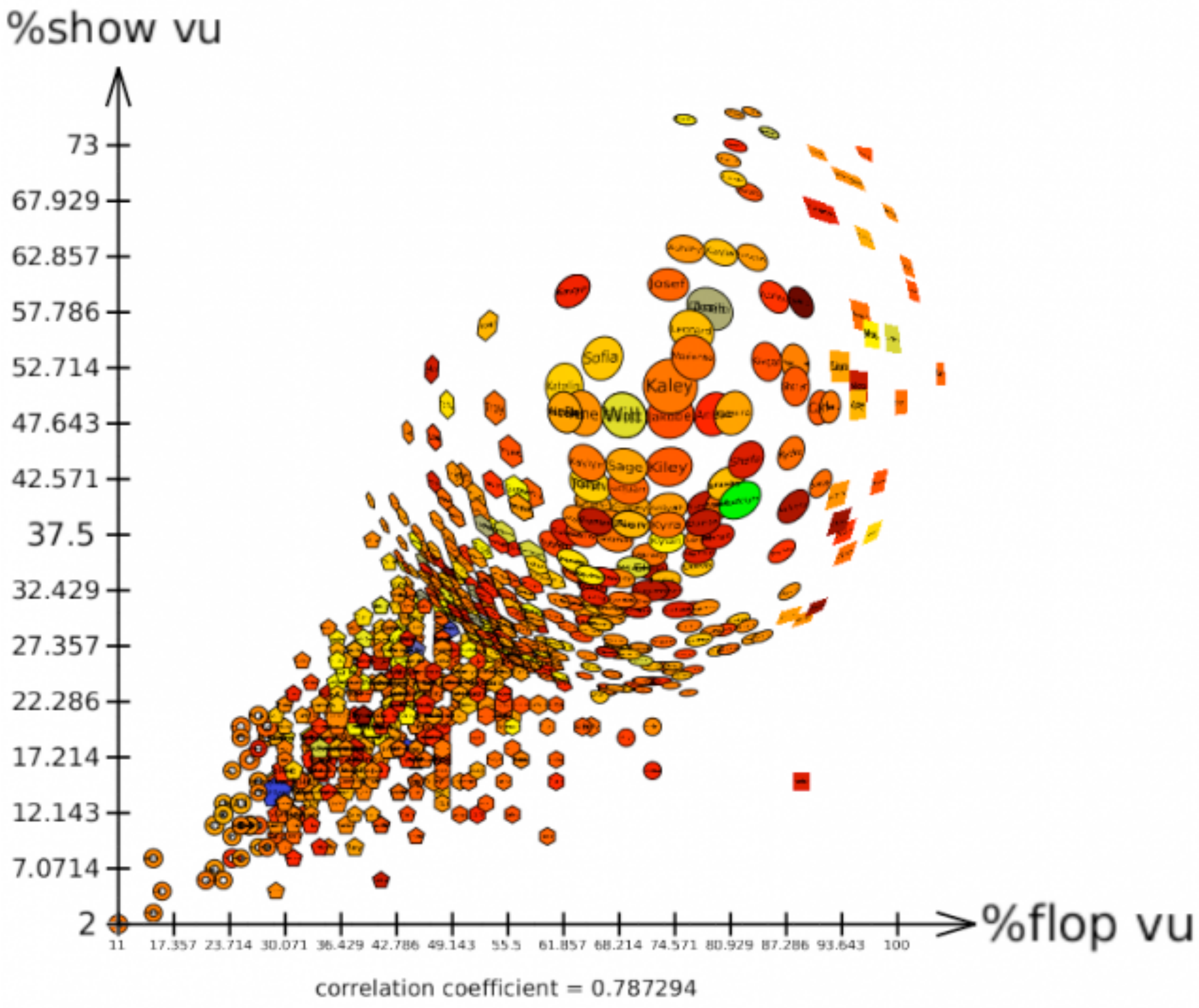


[DOITrees Revisited: Scalable, Space-Constrained Visualization of Hierarchical Data. Heer and Card. Proc. Advanced Visual Interfaces (AVI), pp. 421–424, 2004.]

Idiom: **Fisheye Lens**

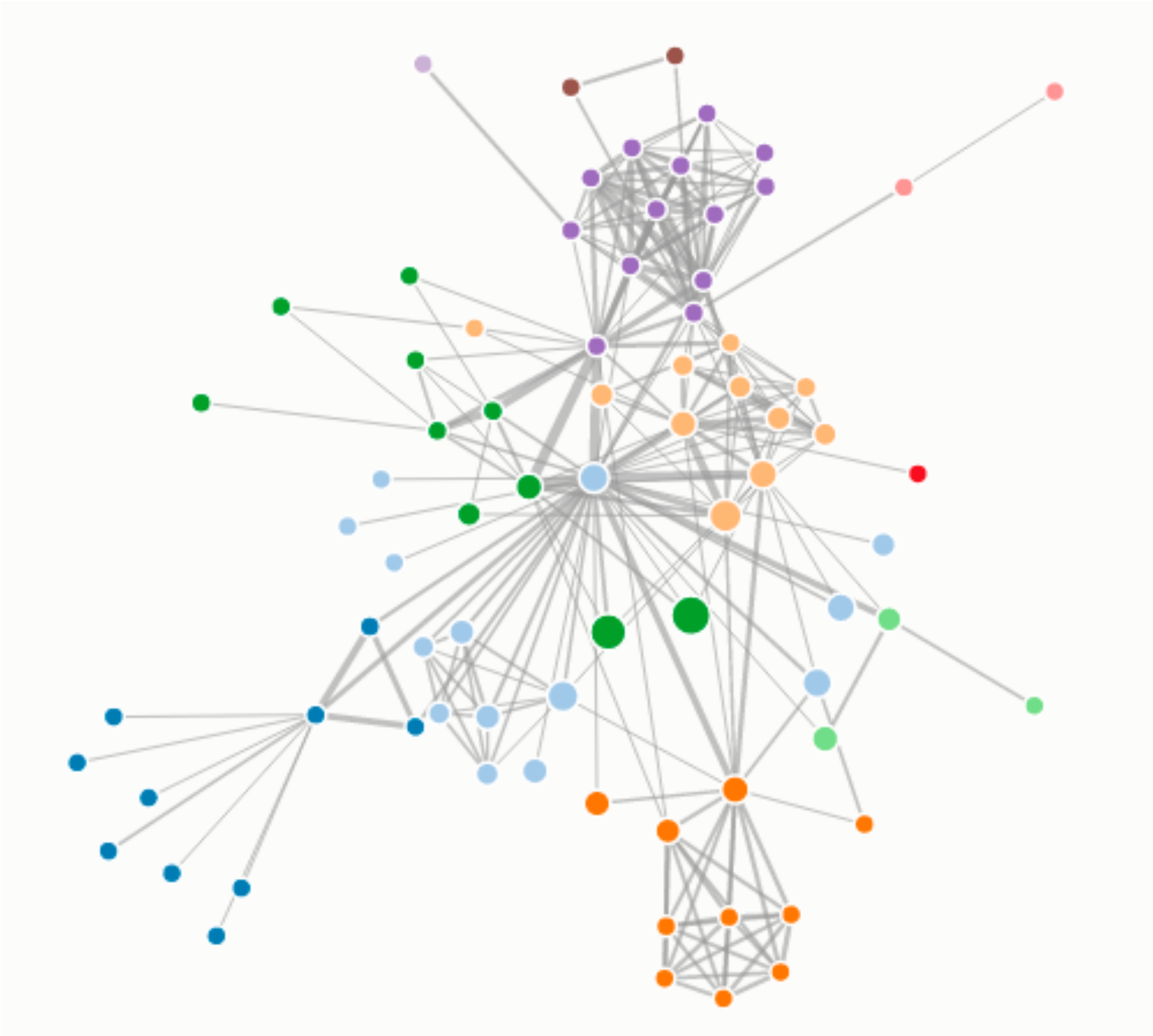
- distort geometry
  - shape: radial
  - focus: single extent
  - extent: local
  - metaphor: draggable lens





http://tulip.labri.fr/TulipDrupal/?q=node/351
http://tulip.labri.fr/TulipDrupal/?q=node/371

Idiom: **Fisheye Lens** System: **D3**

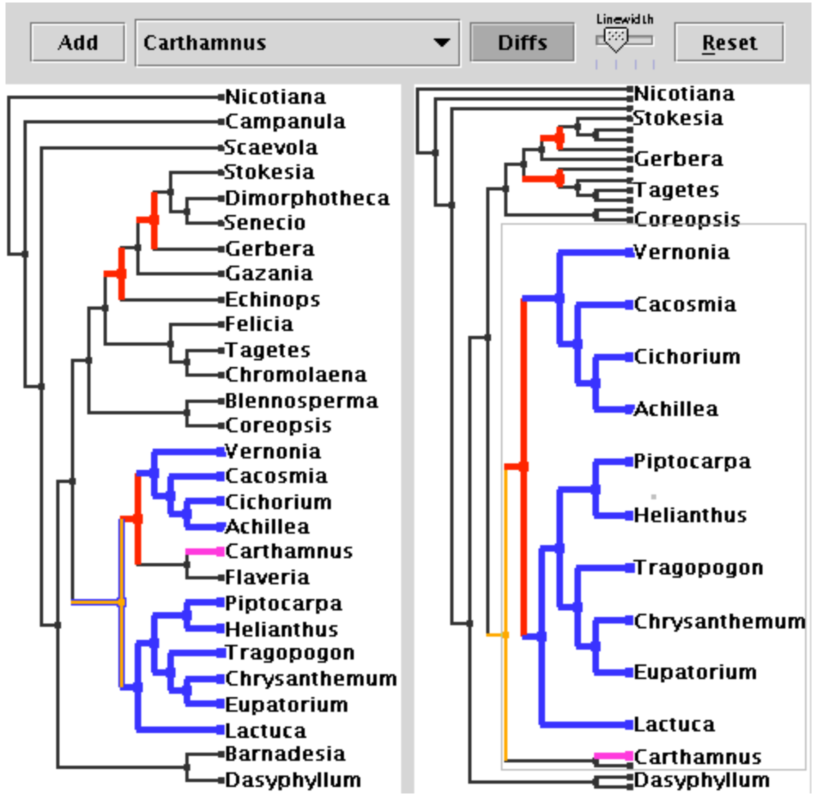


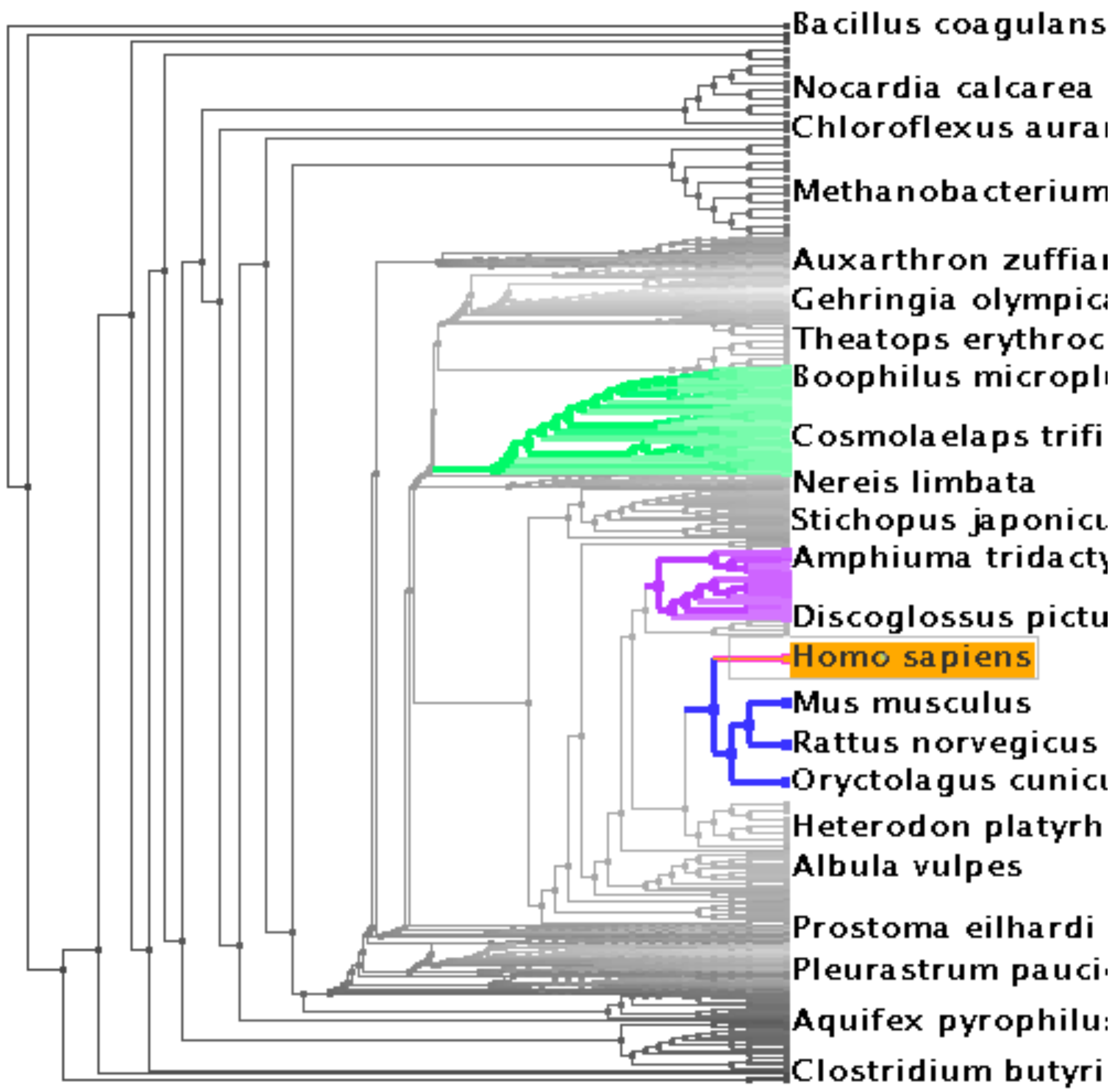
[D3 Fisheye Lens](https://bost.ocks.org/mike/fisheye/)

Idiom: **Stretch and Squish Navigation**

System: **TreeJuxtaposer**

- distort geometry
  - –shape: rectilinear
  - –foci: multiple
  - –impact: global
  - –metaphor: stretch and squish, borders fixed





[https://youtu.be/GdaPj8a9QEo]

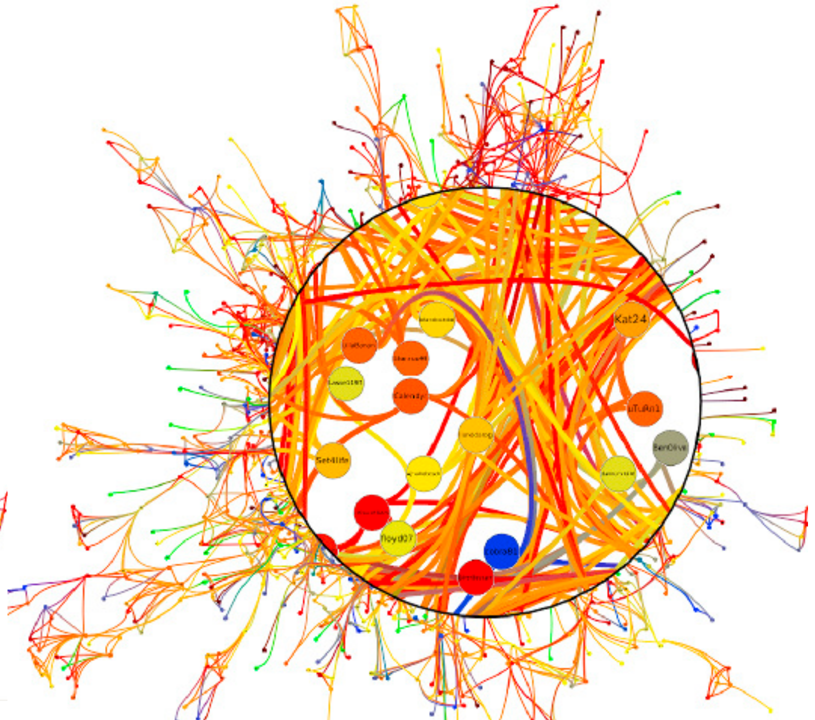
[TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context With Guaranteed Visibility. Munzner, Guimbretiere, Tasiran, Zhang, and Zhou. ACM Transactions on Graphics (Proc. SIGGRAPH) 22:3 (2003), 453– 462.]

Distortion costs and benefits

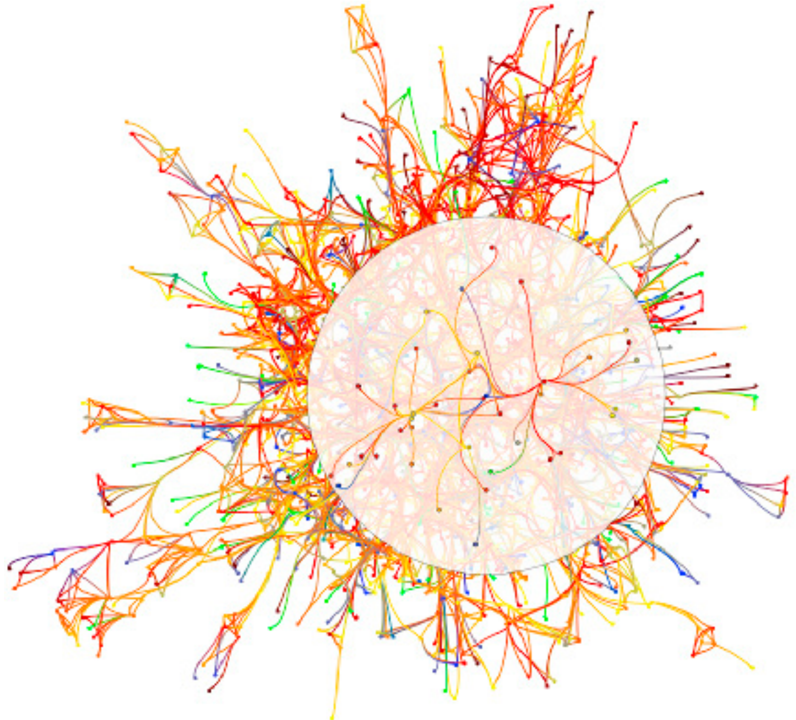
- benefits
  - combine focus and context information in single view
- costs
  - length comparisons impaired
    - network/tree topology comparisons unaffected: connection, containment
  - effects of distortion unclear if original structure unfamiliar
  - object constancy/tracking maybe impaired

fisheye lens

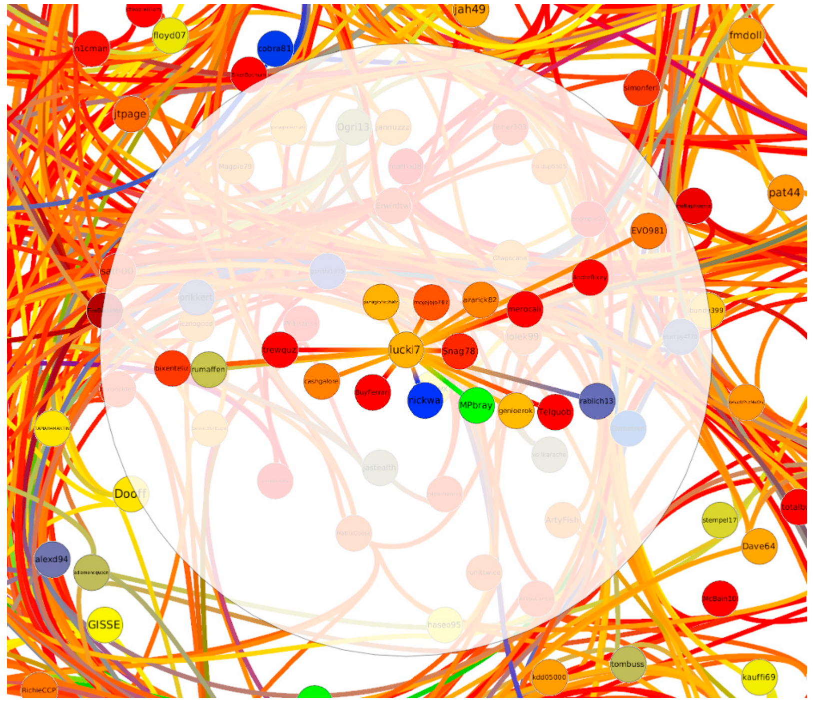
magnifying lens



neighborhood layering



Bring and Go



[https://www.youtube.com/watch?v=hm2oFBqVM9o]

[Living Flows: Enhanced Exploration of Edge-Bundled Graphs Based on GPU-Intensive Edge Rendering. Lambert, Auber, and Melançon. Proc. Intl. Conf. Information Visualisation (IV), pp. 523–530, 2010.]

Further reading

- Visualization Analysis and Design. Munzner. AK Peters / CRC Press, Oct 2014.
  – *Chap 14: Embed: Focus+Context*
- *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.
- *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 Fisheye Follow-up: Further Reflection on Focus + Context.* Furnas. Proc. ACM Conf. Human Factors in Computing Systems (CHI), pp. 999–1008, 2006.

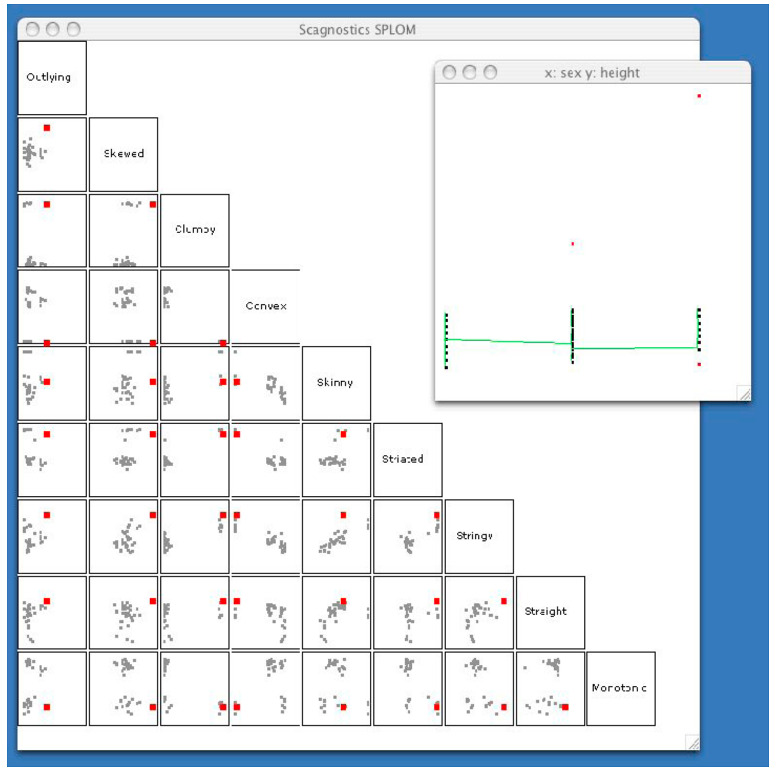
Further reading, full

- A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. Computing Surveys 41:1 (2008), pp. 1-31.
- A Guide to Visual Multi-Level Interface Design from Synthesis of Empirical Study Evidence. Heidi Lam and Tamara Munzner. Synthesis Lectures on Visualization Series, Morgan Claypool, 2010.
- Data Base Navigation: An Office Environment for the Professional. Robert Spence and Mark Apperley. Behaviour and Information Technology 1:1 (1982), pp. 43-54.
- A Polyfocal Projection for Statistical Surfaces. Naftali Kadmon and Eli Shlomi. The Cartographic Journal 15:1 (1978), pp. 36-41.
- A Review and Taxonomy of Distortion-Oriented Presentation Techniques. Ying K. Leung and Mark Apperley. Transactions on Computer-Human Interaction (ToCHI) 1:2 (1994), pp. 126-160.
- The FISHEYE View: A New Look at Structured Files. George W. Furnas. Technical report, Bell Laboratories Technical Memorandum No. 82-11221-22, 1982.
- Generalized Fisheye Views. George W. Furnas. Proc. CHI 1986, pp. 16-23.
- A Fisheye Follow-up: Further Reflection on Focus + Context. George W. Furnas. Proc. CHI 2006, pp. 999-1008.
- The Perspective Wall: Detail and Context Smoothly Integrated. Jock D. Mackinlay, George G. Robertson, and Stuart K. Card. Proc. CHI 1991, pp. 173-179.
- Cone Trees: Animated 3D Visualizations of Hierarchical Information. George Robertson, Jock Mackinlay, and Stuart Card. Proc. CHI 1991, pp. 189-194.
- Three-Dimensional Pliable Surfaces: For Effective Presentation of Visual Information. M. Sheelagh T. Carpendale, David J. Cowperthwaite, and F. David Fracchia. Proc. of the Symposium on User Interface Software and Technology (UIST), pp. 217-226. ACM, 1995.
- Distortion Viewing Techniques for 3D Data. M. Sheelagh T. Carpendale, David J. Cowperthwaite, and F. David Fracchia. Proc. InfoVis, pp. 46-53. IEEE Computer Society, 1996.
- Nonlinear Magnification Fields. T. Alan Keahey and Edward L. Robertson. Proc. InfoVis, pp. 51-58. IEEE Computer Society, 1997.
- A Focus+Content Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. John Lamping, Ramana Rao, and Peter Pirolli. Proc. CHI, pp. 401-408. ACM, 1995.
- Exploring Large Graphs in 3D Hyperbolic Space. Tamara Munzner. IEEE Computer Graphics and Applications 18:4 (1998), pp. 18-23.
- TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context with Guaranteed Visibility. Tamara Munzner, Francois Guimbretiere, Serdar Tasiran, Li Zhang, and Yunhong Zhou. Proc. SIGGRAPH (2003), pp. 453-462.
- PRISAD: Partitioned Rendering Infrastructure for Scalable Accordion Drawing (Extended Version). James Slack, Kristian Hildebrand, and Tamara Munzner. Information Visualization 5:2 (2006), pp. 137-151.

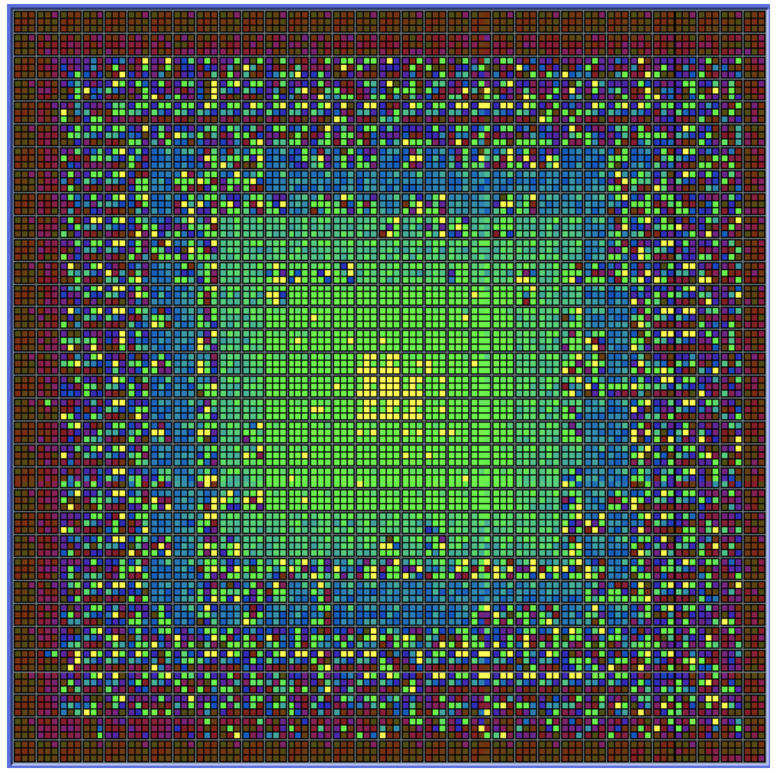
Ch 15: Analysis Case Studies

Analysis Case Studies

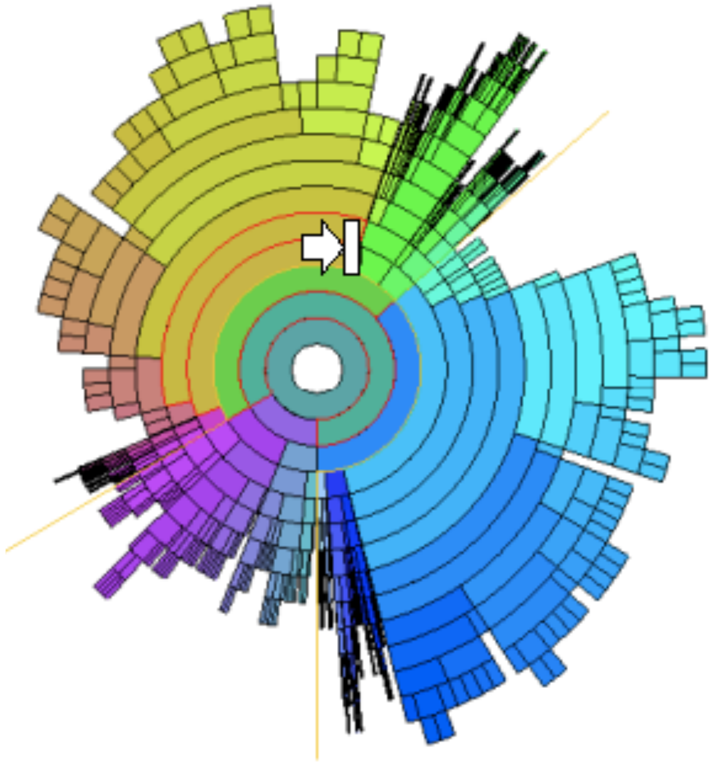
Scagnostics



VisDB

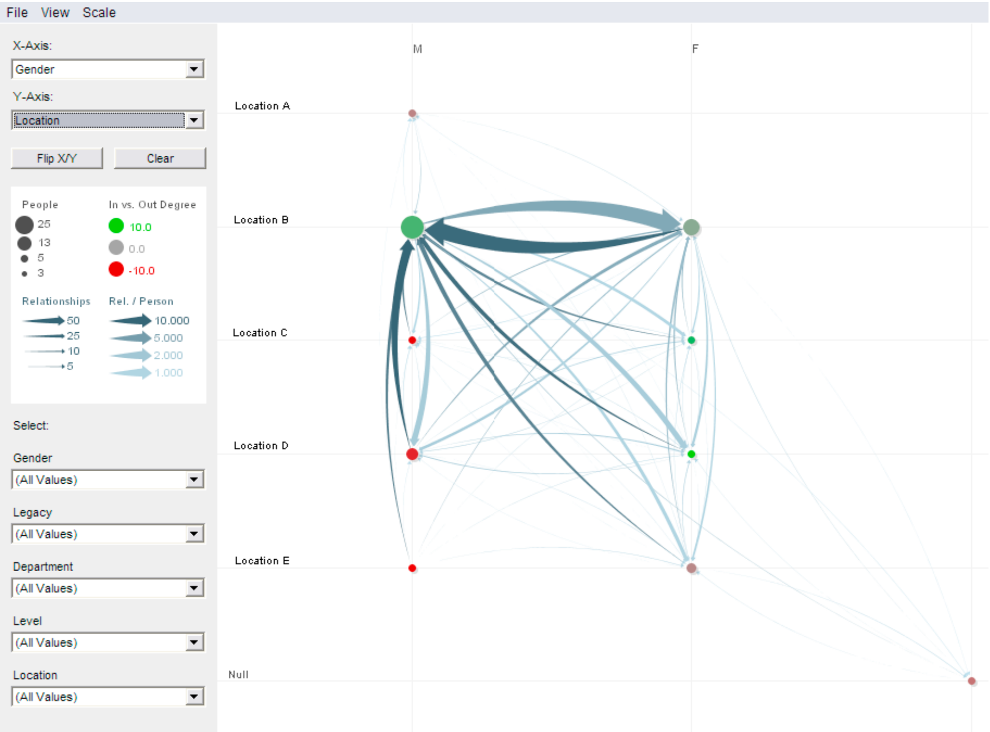


InterRing

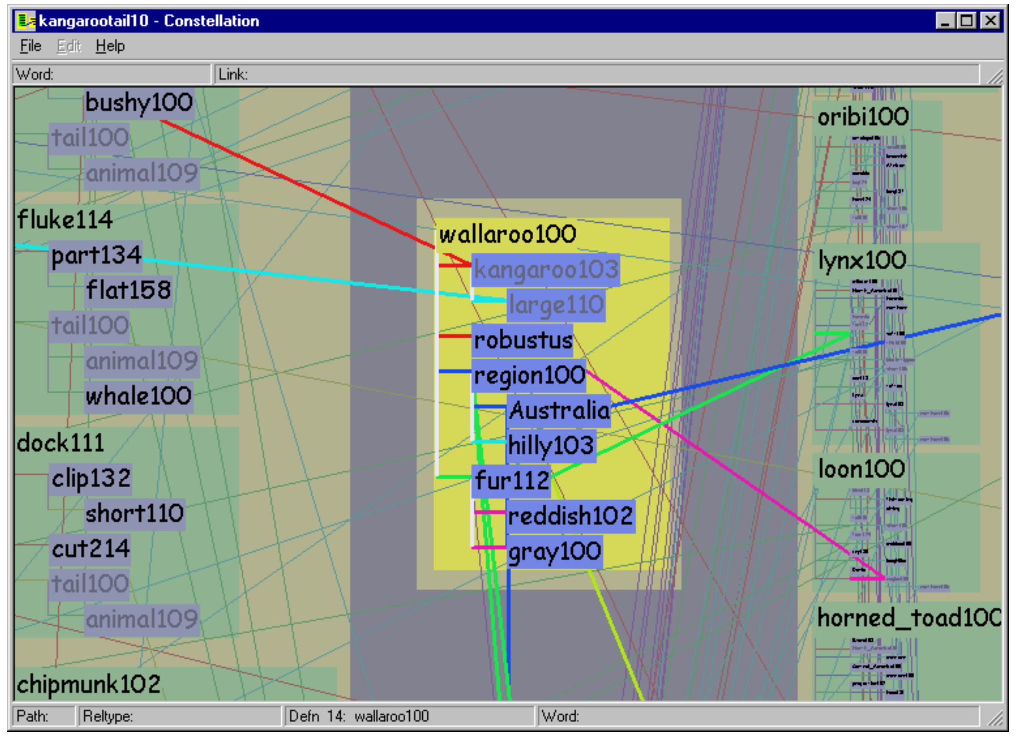


HCE

PivotGraph

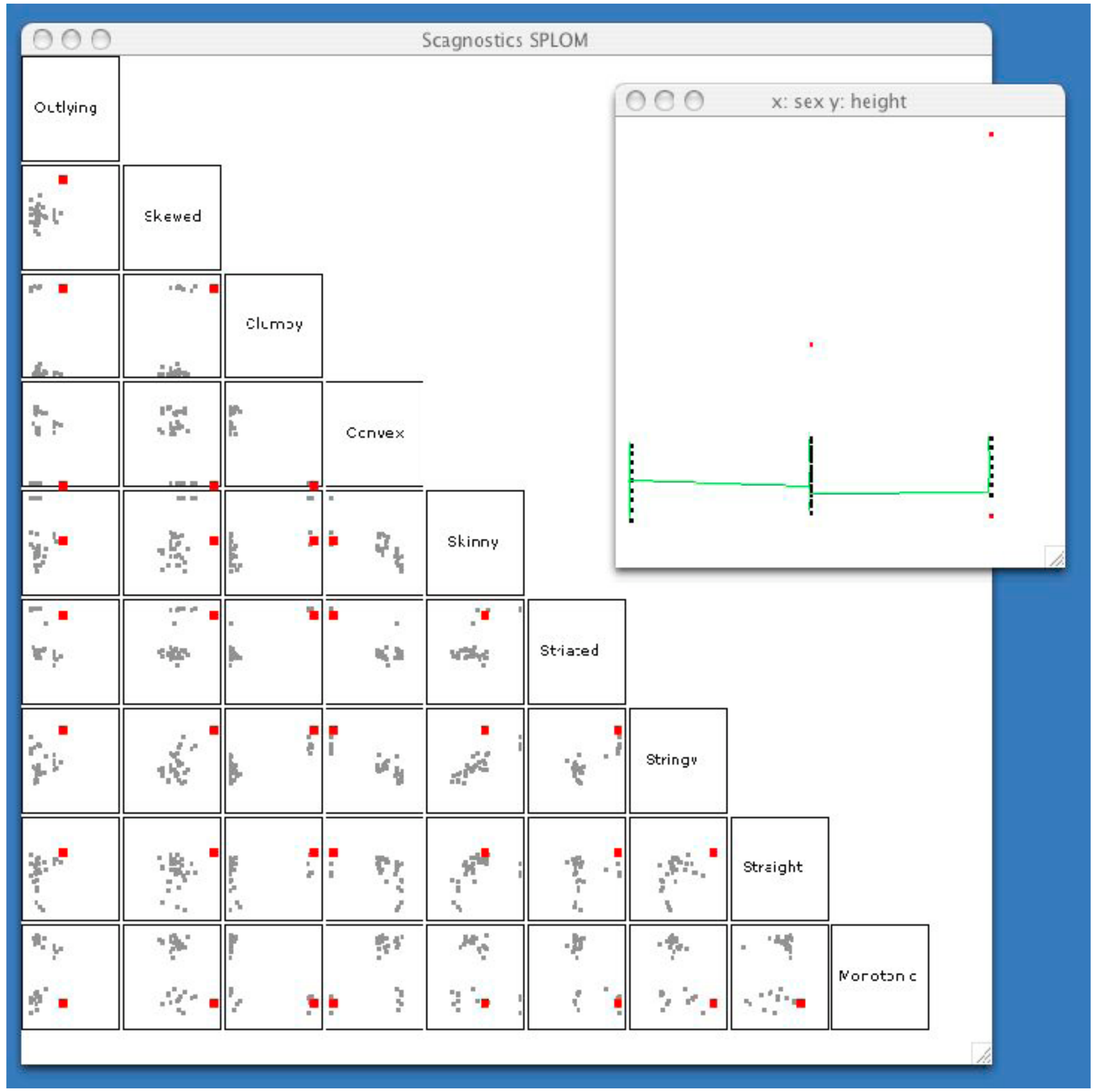


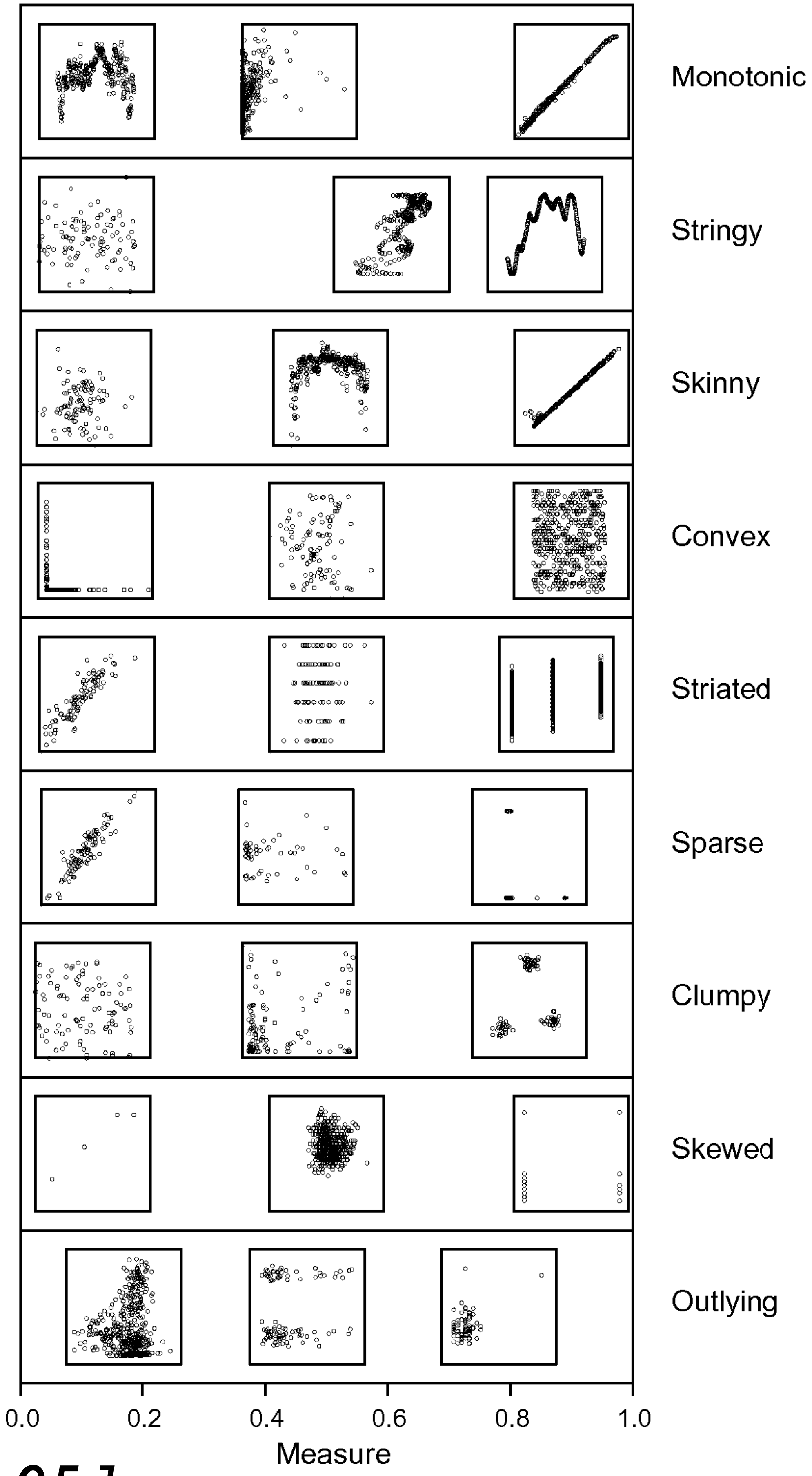
Constellation



Graph-Theoretic Scagnostics

- scatterplot diagnostics
  - scagnostics SPLOM: each point is one original scatterplot





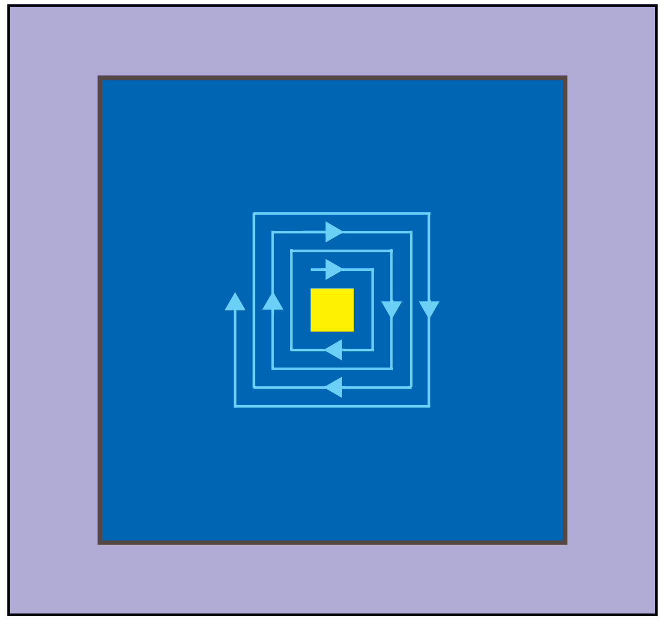
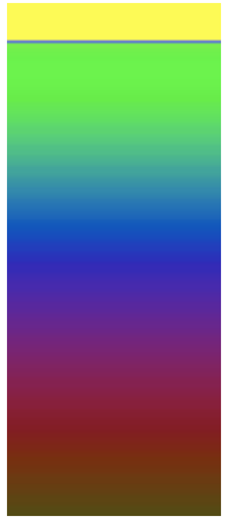
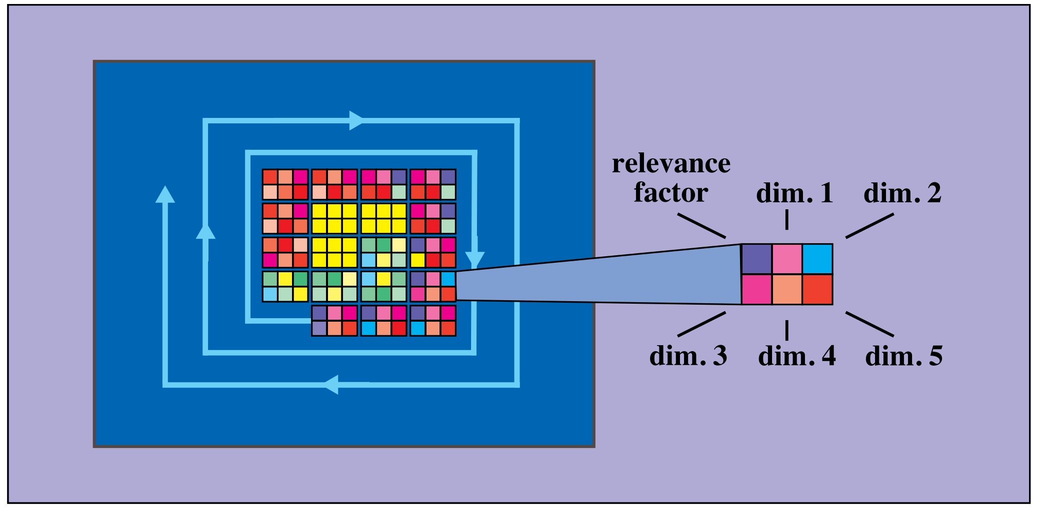
[Graph-Theoretic Scagnostics Wilkinson, Anand, and Grossman. Proc InfoVis 05.]

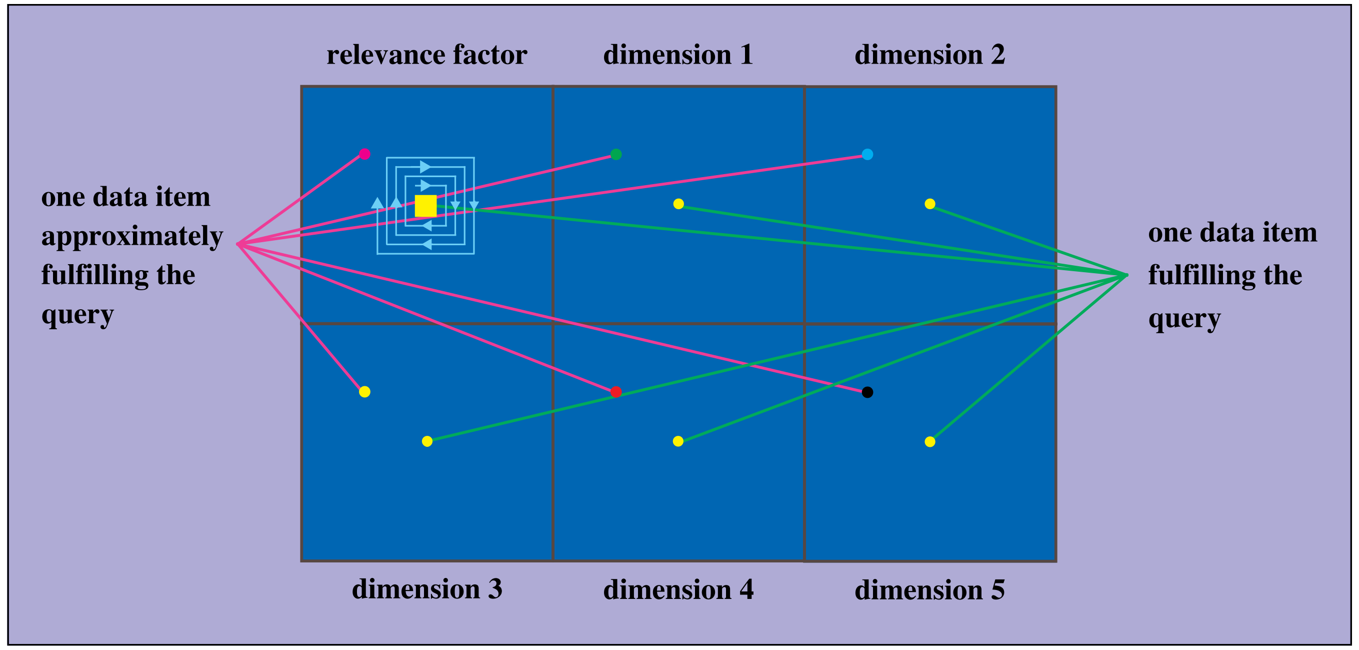
Scagnostics analysis

System	Scagnostics
What: Data	Table.
What: Derived	Nine quantitative attributes per scatterplot (pairwise combination of original attributes).
Why: Tasks	Identify, compare, and summarize; distributions and correlation.
How: Encode	Scatterplot, scatterplot matrix.
How: Manipulate	Select.
How: Facet	Juxtaposed small-multiple views coordinated with linked highlighting, popup detail view.
Scale	Original attributes: dozens.

VisDB

- table: draw pixels sorted, colored by relevance
- group by attribute or partition by attribute into multiple views

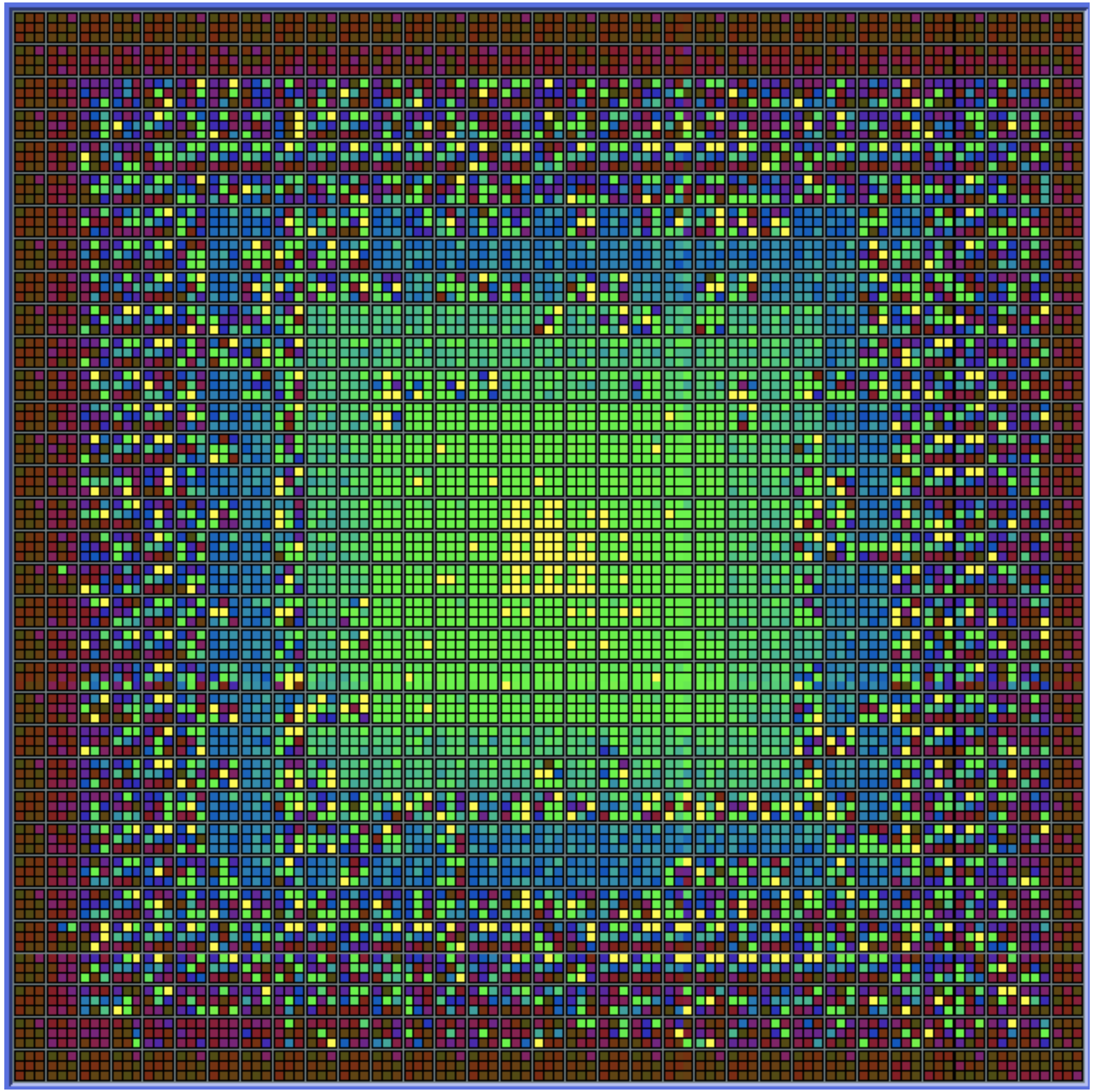






[VisDB: Database Exploration using Multidimensional Visualization, Keim and Kriegel, IEEE CG&A, 1994]

VisDB Results

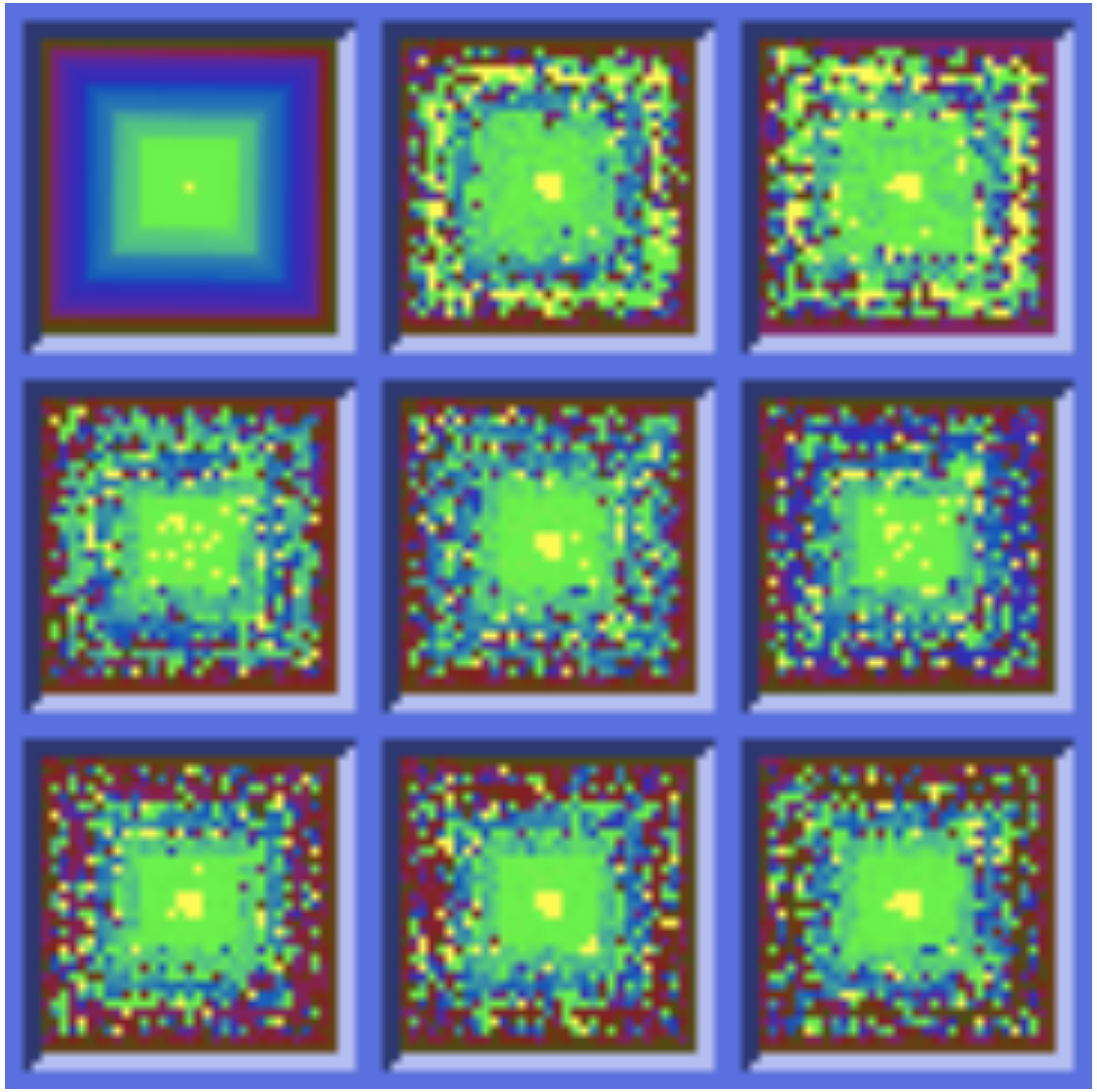
- partition into many small regions: dimensions grouped together



[VisDB: Database Exploration using Multidimensional Visualization, Keim and Kriegel, IEEE CG&A, 1994]

VisDB Results

- partition into small number of views
  - –inspect each attribute



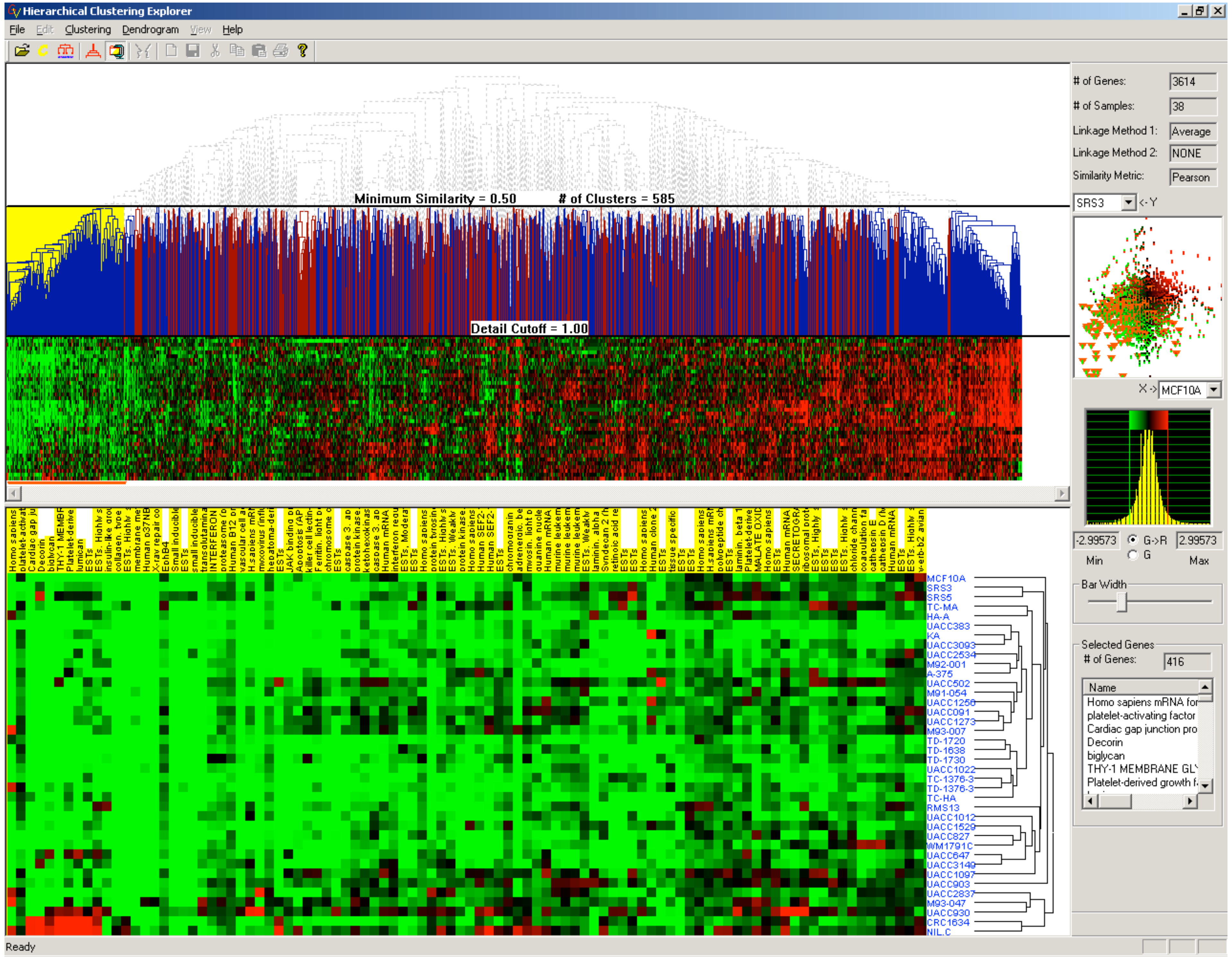
[*VisDB: Database Exploration using Multidimensional Visualization, Keim and Kriegel, IEEE CG&A, 1994*]

VisDB Analysis

System	VisDB
What: Data	Table (database) with k attributes; query returning table subset (database query).
What: Derived	$k + 1$ quantitative attributes per original item: query relevance for the k original attributes plus overall relevance.
Why: Tasks	Characterize distribution within attribute, find groups of similar values within attribute, find outliers within attribute, find correlation between attributes, find similar items.
How: Encode	Dense, space-filling; area marks in spiral layout; colormap: categorical hues and ordered luminance.
How: Facet	Layout 1: partition by attribute into per-attribute views, small multiples. Layout 2: partition by items into per-item glyphs.
How: Reduce	Filtering
Scale	Attributes: one dozen. Total items: several million. Visible items (using multiple views, in total): one million. Visible items (using glyphs): 100,000

Hierarchical Clustering Explorer

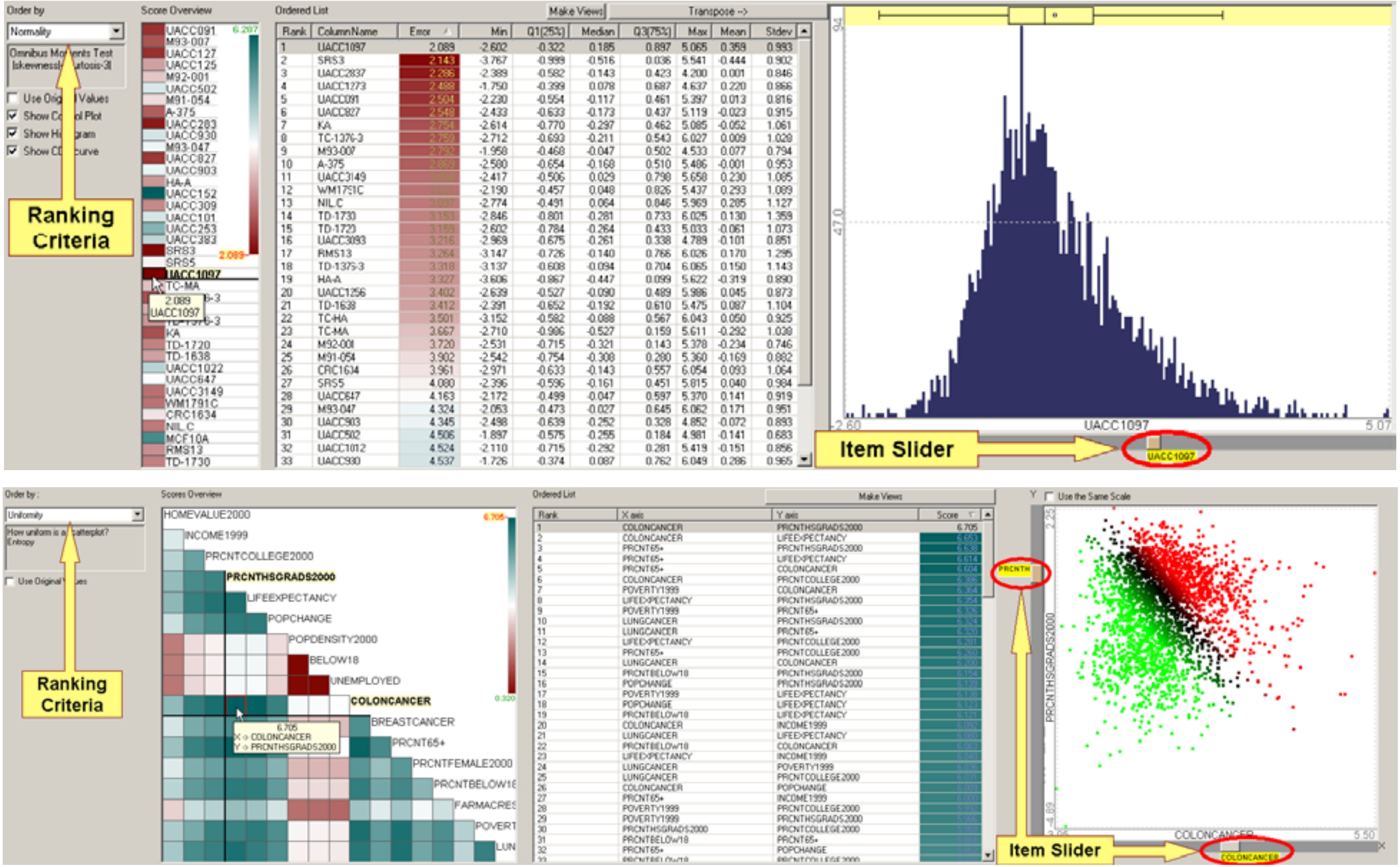
- heatmap, dendrogram
- multiple views



[Interactively Exploring Hierarchical Clustering Results. Seo and Shneiderman, <u>IEEE Computer 35</u>(7): 80-86 (2002)]

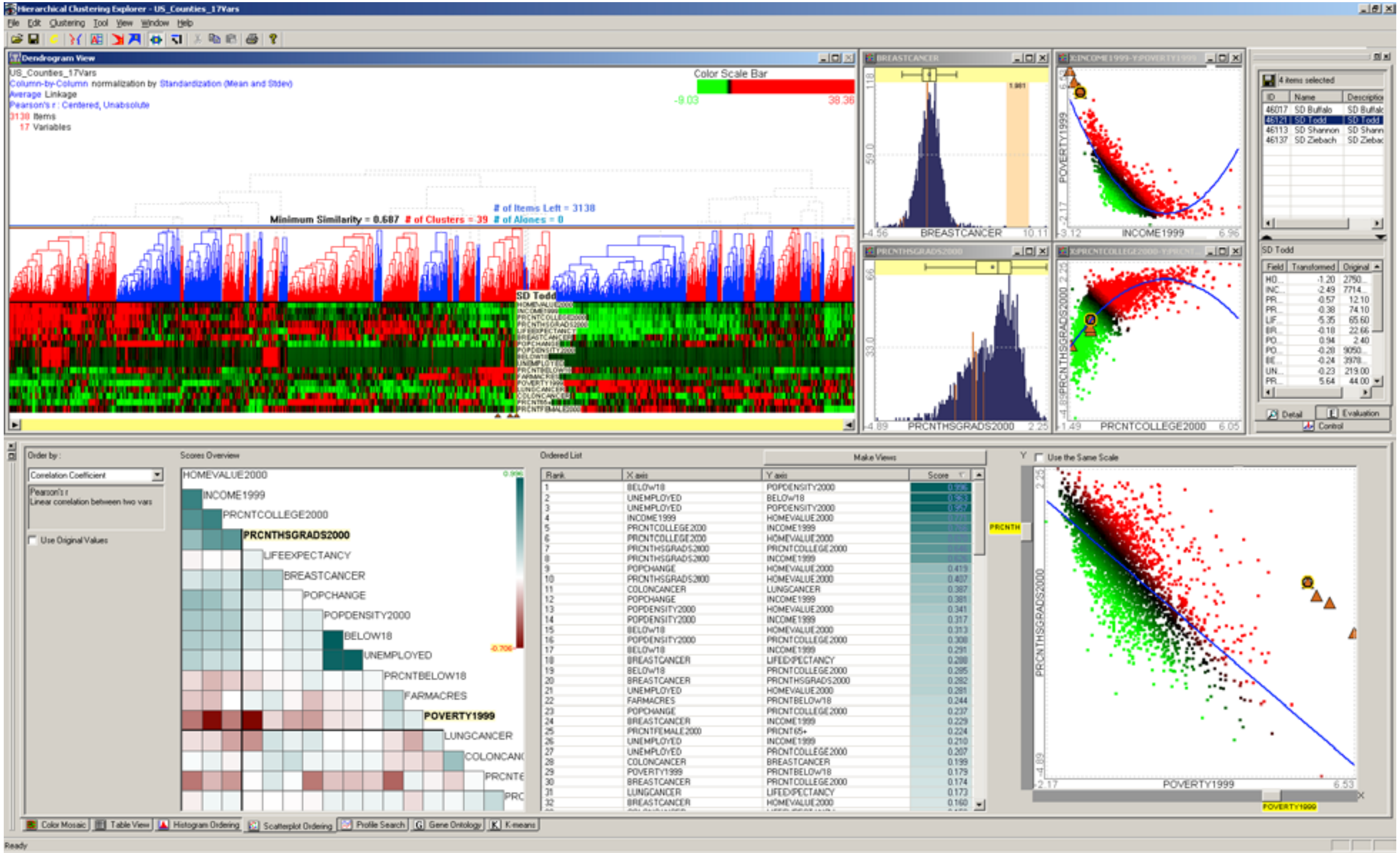
HCE

- rank by feature idiom
  - 1D list
  - 2D matrix



A rank-by-feature framework for interactive exploration of multidimensional data. Seo and Shneiderman. Information Visualization 4(2): 96-113 (2005)

HCE

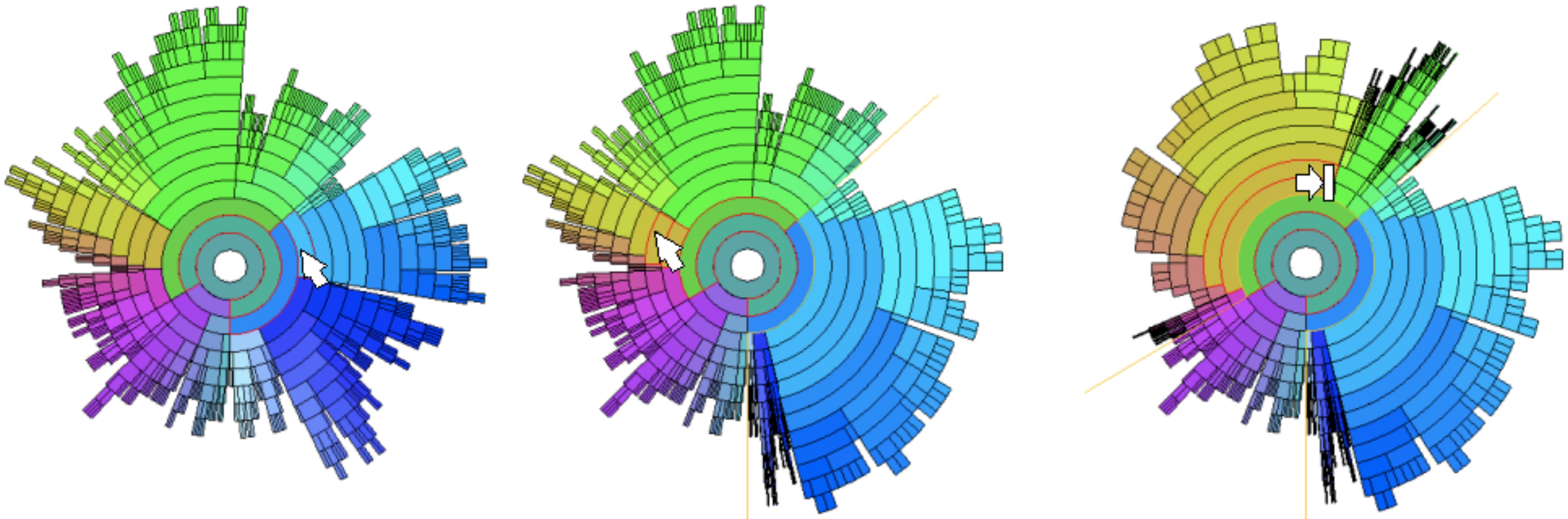


A rank-by-feature framework for interactive exploration of multidimensional data. Seo and Shneiderman. Information Visualization 4(2): 96-113 (2005)

HCE Analysis

System	Hierarchical Clustering Explorer (HCE)
What: Data	Multidimensional table: two categorical key attributes (genes, conditions); one quantitative value attribute (gene activity level in condition).
What: Derived	Hierarchical clustering of table rows and columns (for cluster heatmap); quantitative derived attributes for each attribute and pairwise attribute combination; quantitative derived attribute for each ranking criterion and original attribute combination.
Why: Tasks	Find correlation between attributes; find clusters, gaps, outliers, trends within items.
How: Encode	Cluster heatmap, scatterplots, histograms, boxplots. Rank-by-feature overviews: continuous diverging colormaps on area marks in reorderable 2D matrix or 1D list alignment.
How: Reduce	Dynamic filtering; dynamic aggregation.
How: Manipulate	Navigate with pan/scroll.
How: Facet	Multiform with linked highlighting and shared spatial position; overview–detail with selection in overview populating detail view.
Scale	Genes (key attribute): 20,000. Conditions (key attribute): 80. Gene activity in condition (quantitative value attribute): $20{,}000 \times 80 = 1{,}600{,}000$.

InterRing



original hierarchy | blue subtree expanded | tan subtree expanded

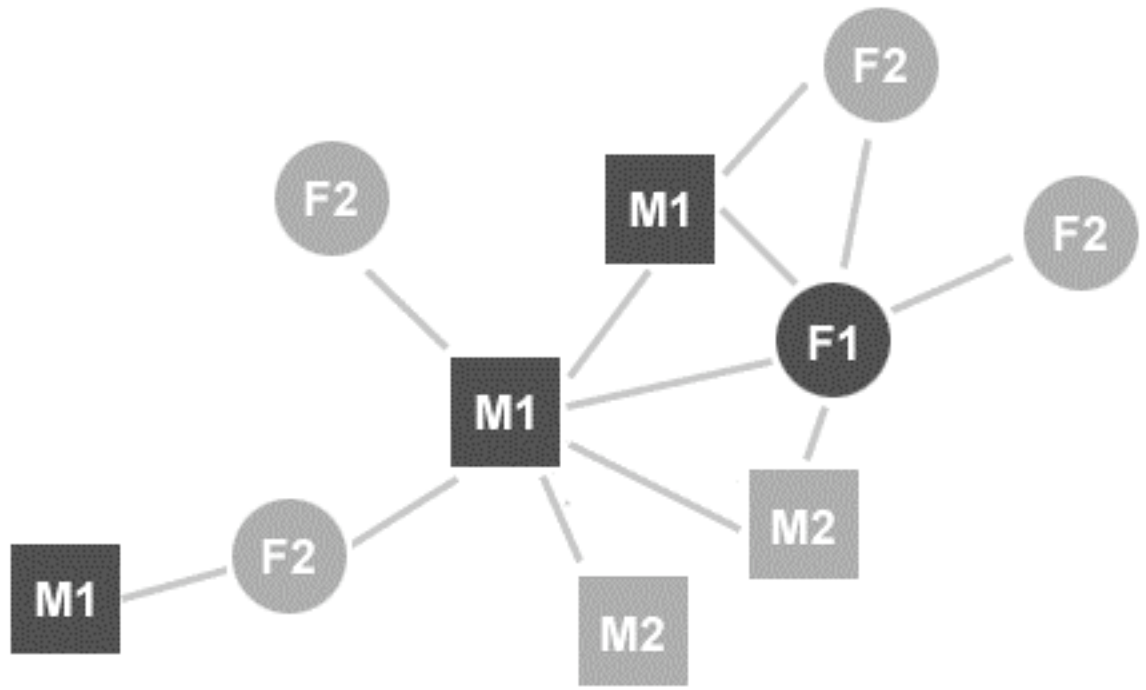
[InterRing: An Interactive Tool for Visually Navigating and Manipulating Hierarchical Structures. Yang, Ward, Rundensteiner. Proc. InfoVis 2002, p 77-84.]

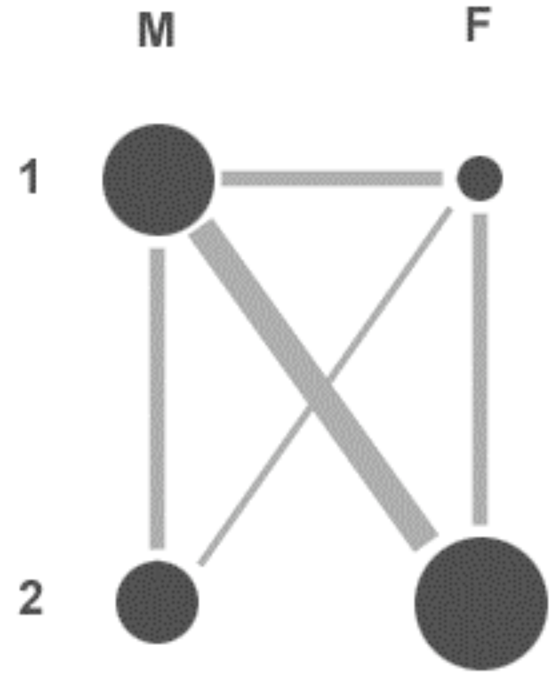
InterRing Analysis

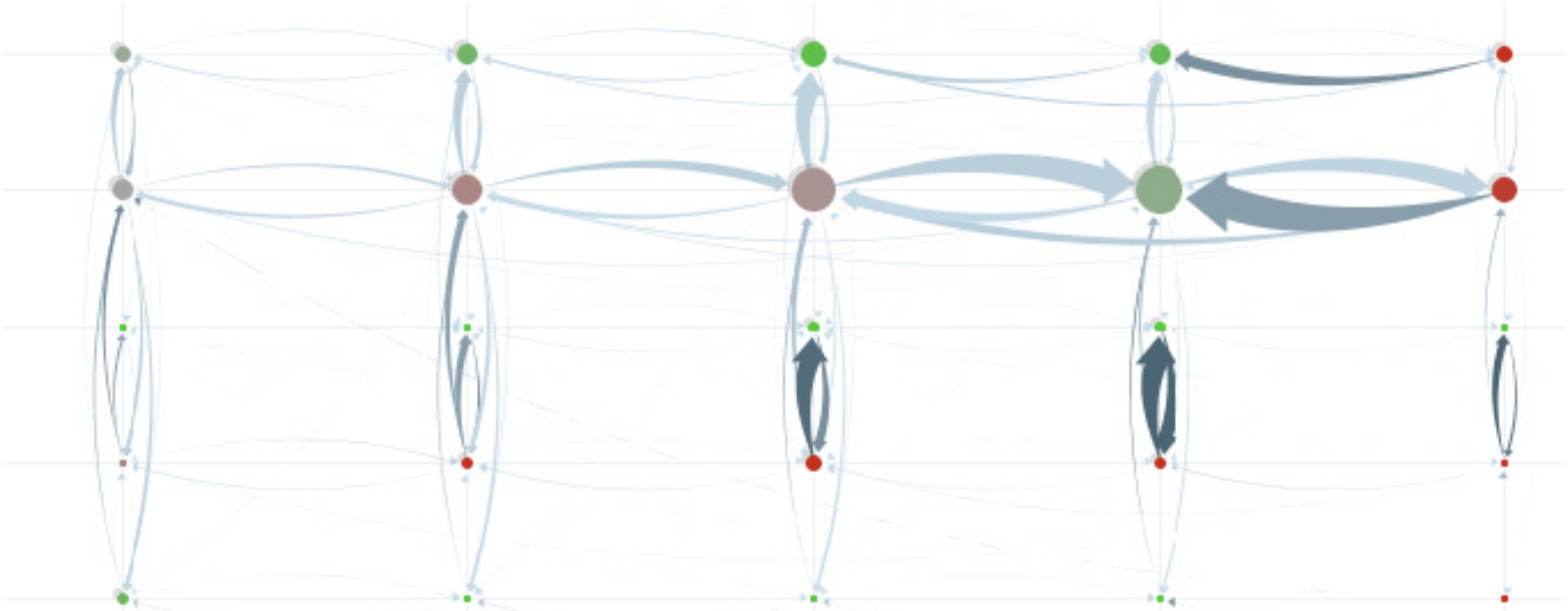
System	**InterRing**
What: Data	Tree.
Why: Tasks	Selection, rollup/drilldown, hierarchy editing.
How: Encode	Radial, space-filling layout. Color by tree structure.
How: Facet	Linked coloring and highlighting.
How: Reduce	Embed: distort; multiple foci.
Scale	Nodes: hundreds if labeled, thousands if dense. Levels in tree: dozens.

PivotGraph

- derived rollup network

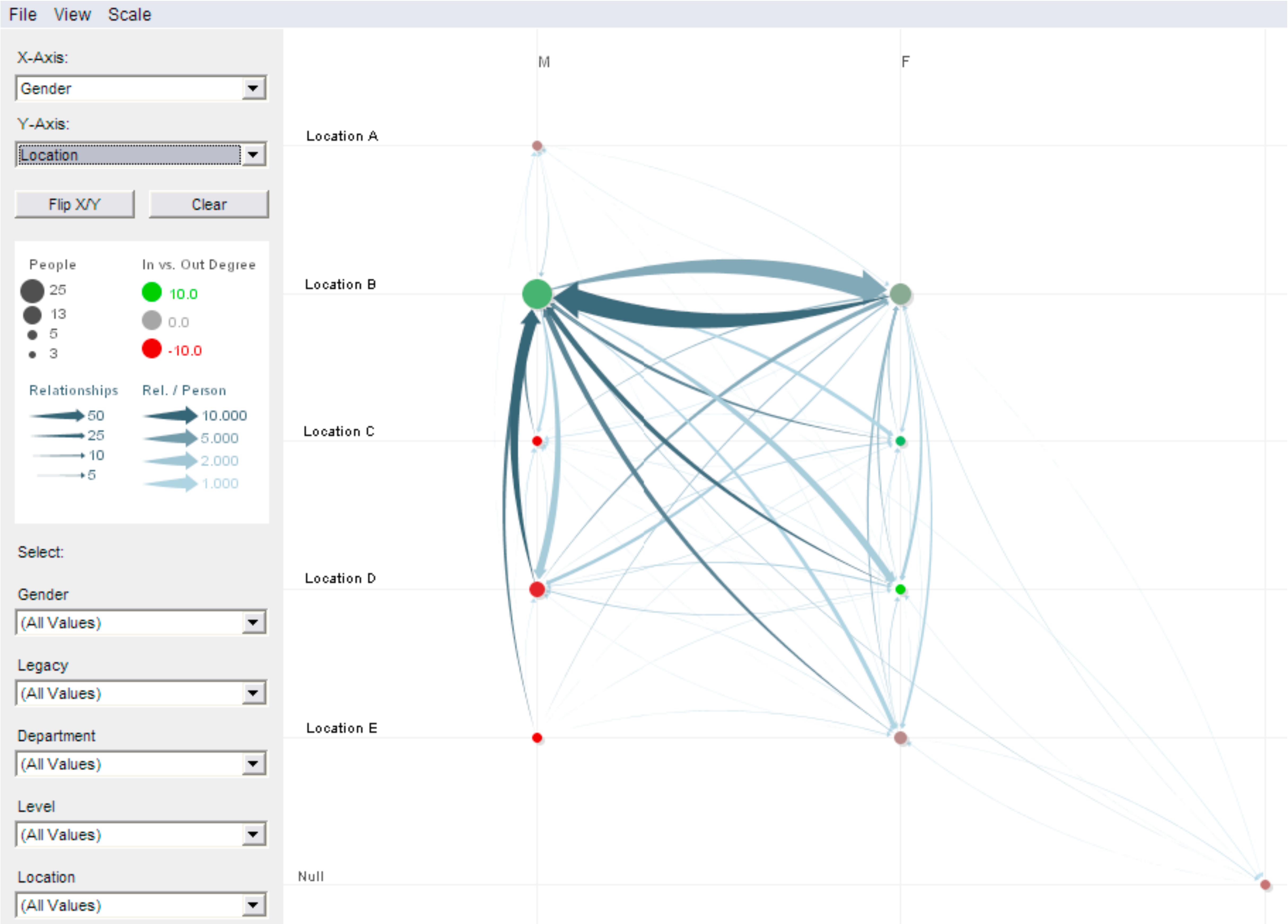






[Visual Exploration of Multivariate Graphs, Martin Wattenberg, CHI 2006.]

PivotGraph



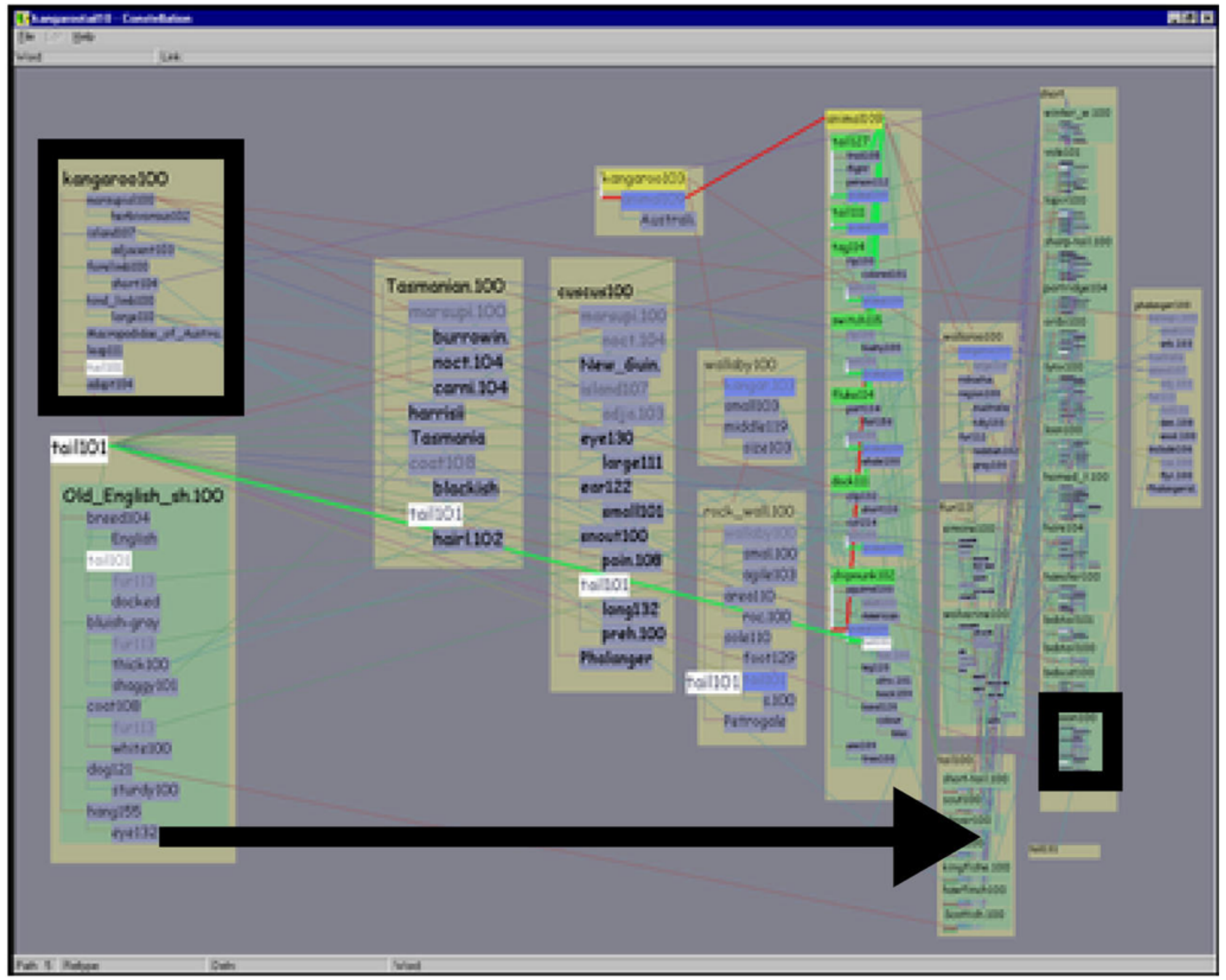
[Visual Exploration of Multivariate Graphs, Martin Wattenberg, CHI 2006.]

PivotGraph Analysis

Idiom	**PivotGraph**
What: Data	Network.
What: Derived	Derived network of aggregate nodes and links by roll-up into two chosen attributes.
Why: Tasks	Cross-attribute comparison of node groups.
How: Encode	Nodes linked with connection marks, size.
How: Manipulate	Change: animated transitions.
How: Reduce	Aggregation, filtering.
Scale	Nodes/links in original network: unlimited. Roll-up attributes: 2. Levels per roll-up attribute: several, up to one dozen.

Analysis example: Constellation

- data
  - multi-level network
    - node: word
    - link: words used in same dictionary definition
    - subgraph for each definition
      - not just hierarchical clustering
  - paths through network
    - query for high-weight paths between 2 nodes
      - quant attrib: plausibility

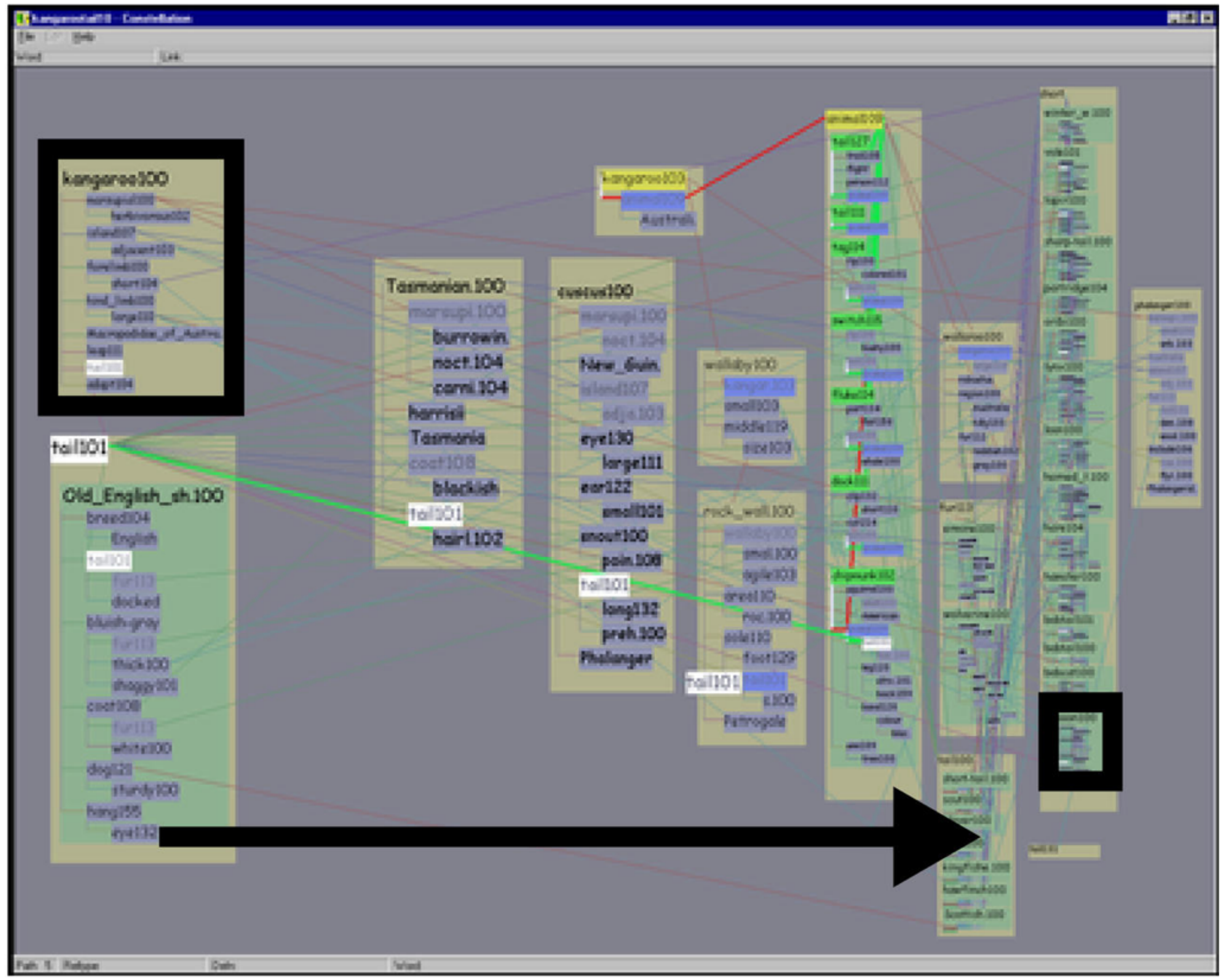


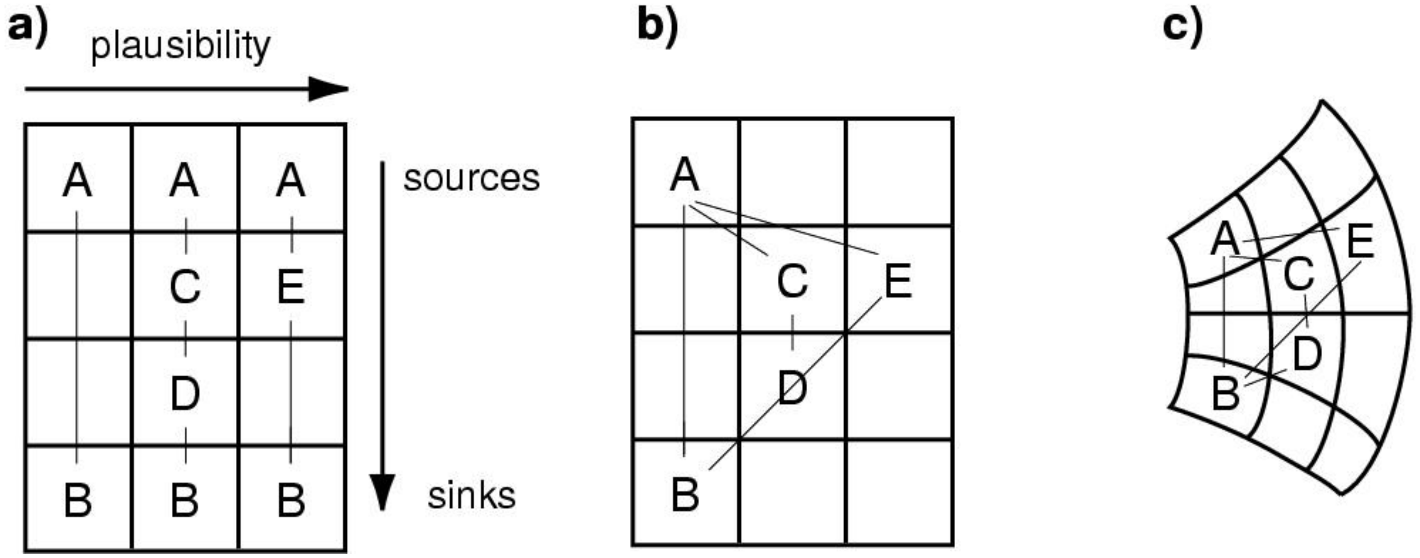
[Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]

[Constellation: A Visualization Tool For Linguistic Queries from MindNet. Munzner, Guimbretière and Robertson. Proc. IEEE Symp. InfoVis 1999, p. 132-135.]

Using space: Constellation

- visual encoding
  - link connection marks between words
  - link containment marks to indicate subgraphs
  - encode plausibility with horiz spatial position
  - encode source/sink for query with vert spatial position
- spatial layout
  - curvilinear grid: more room for longer low-plausibility paths

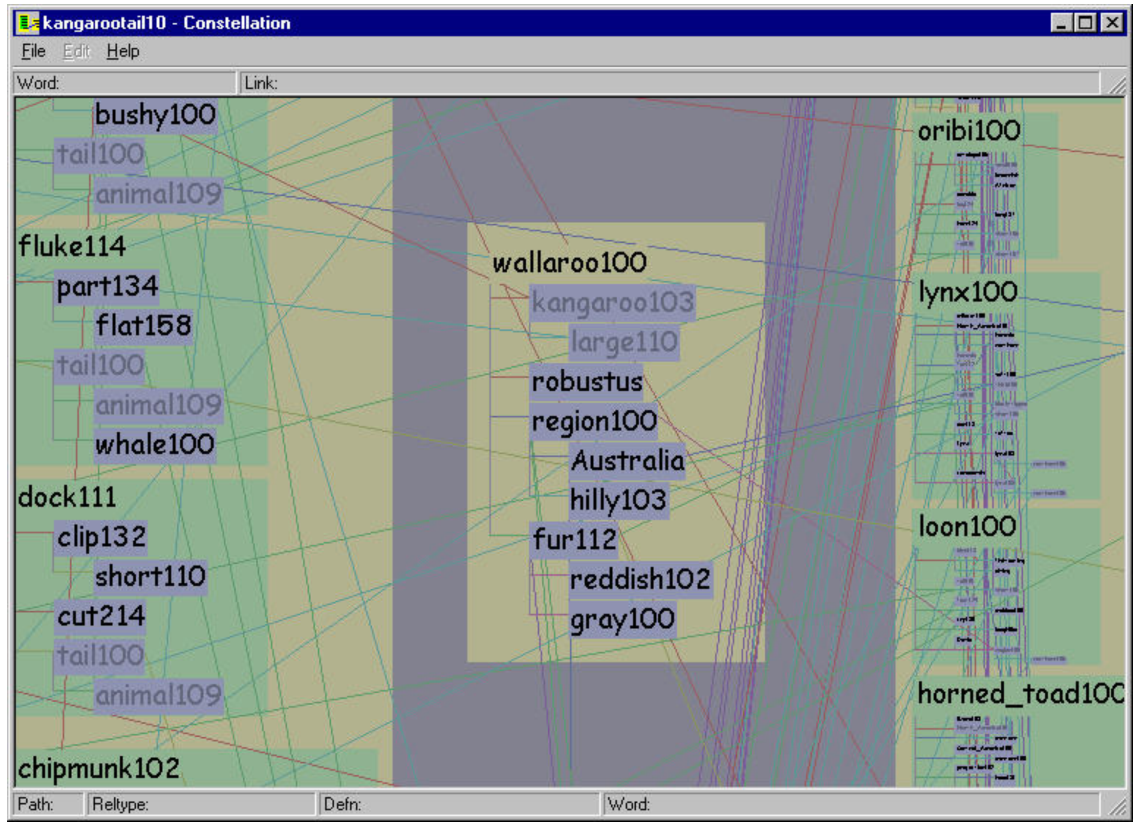


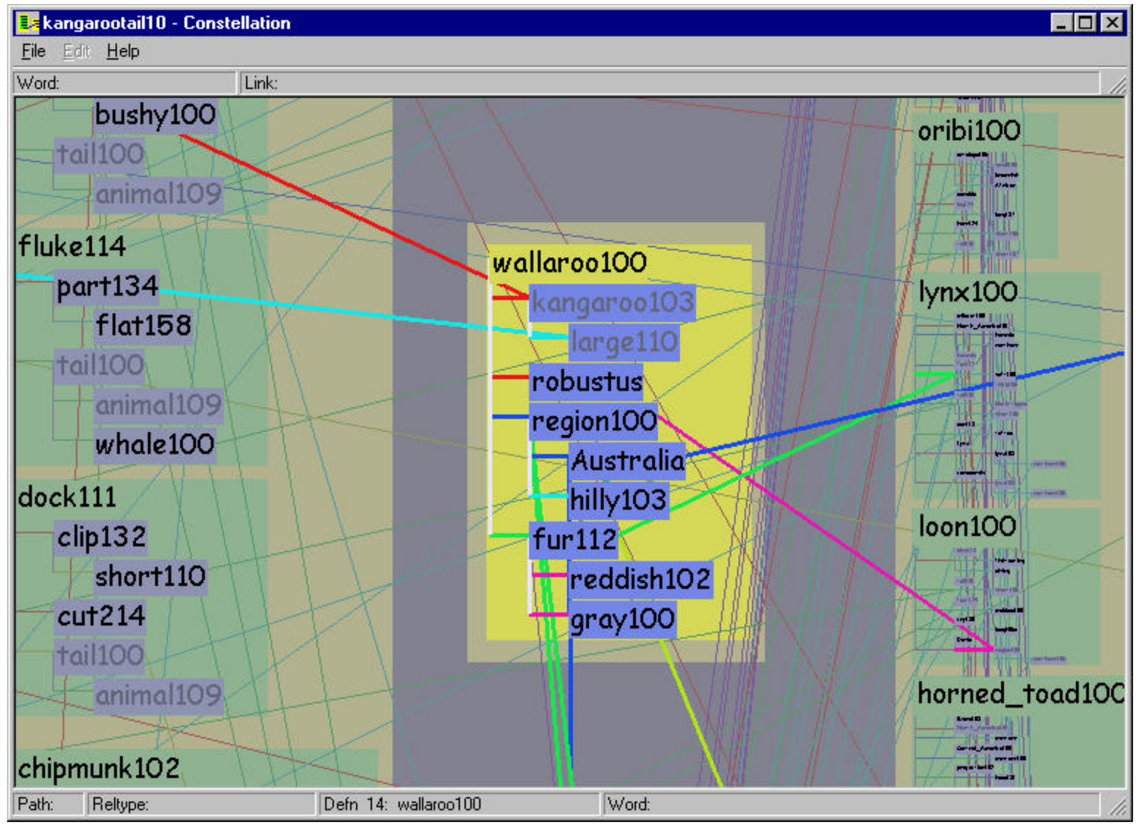


[Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]

Using space: Constellation

- edge crossings
  - cannot easily minimize instances, since position constrained by spatial encoding
  - instead: minimize perceptual impact
- views: superimposed layers
  - dynamic foreground/background layers on mouseover, using color
  - four kinds of constellations
    - definition, path, link type, word
      - not just 1-hop neighbors





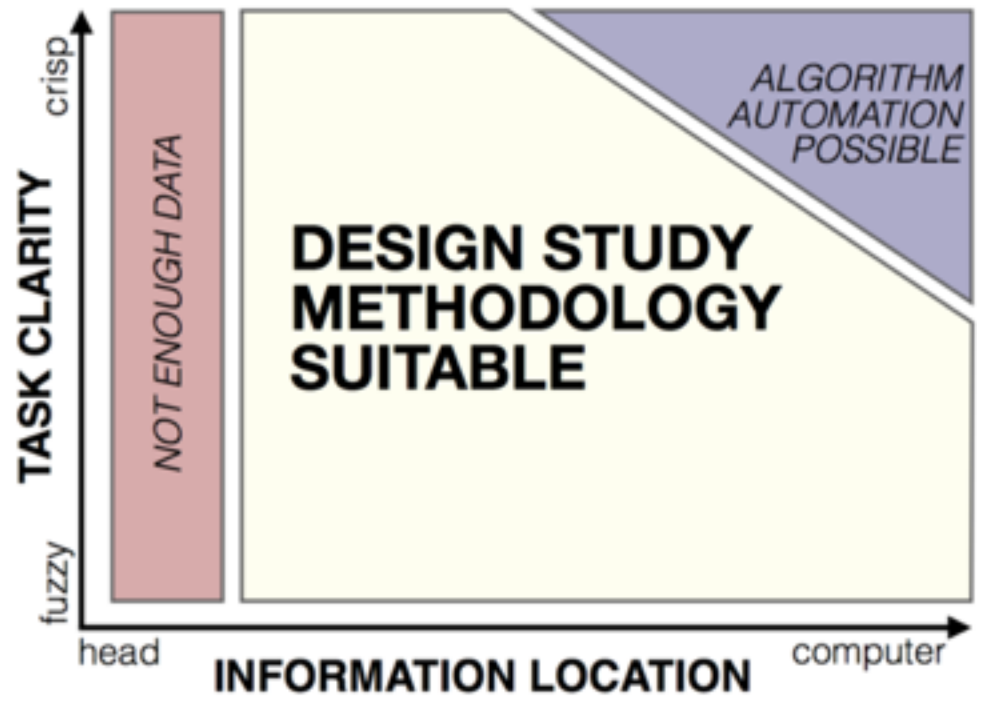
https://youtu.be/7sJC3QVpSkQ

[Interactive Visualization of Large Graphs and Networks. Munzner. Ph.D. Dissertation, Stanford University, June 2000.]

Constellation Analysis

System	**Constellation**
What: Data	Three-level network of paths, subgraphs (definitions), and nodes (word senses).
Why: Tasks	Discover/verify: browse and locate types of paths, identify and compare.
How: Encode	Containment and connection link marks, horizontal spatial position for plausibility attribute, vertical spatial position for order within path, color links by type.
How: Manipulate	Navigate: semantic zooming. Change: Animated transitions.
How: Reduce	Superimpose dynamic layers.
Scale	Paths: 10–50. Subgraphs: 1–30 per path. Nodes: several thousand.

Design Study Methodology



Design Study Methodology

Reflections from the Trenches and from the Stacks

Michael Sedlmair

Miriah Meyer

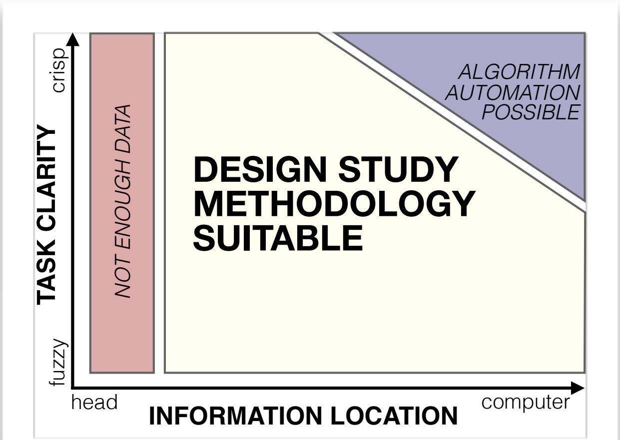
http://www.cs.ubc.ca/labs/imager/tr/2012/dsm/

Tamara Munzner
@tamaramunzner

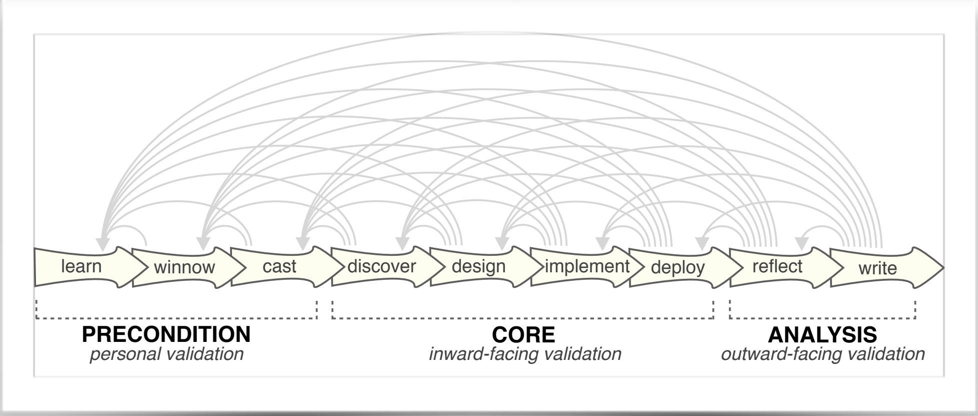
Design Study Methodology: Reflections from the Trenches and from the Stacks.
Sedlmair, Meyer, Munzner. IEEE Trans. Visualization and Computer Graphics 18(12): 2431-2440, 2012 (Proc. InfoVis 2012).

Methodology for problem-driven work

- definitions



- 9-stage framework

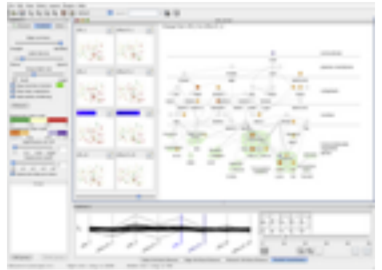


- 32 pitfalls & how to avoid them

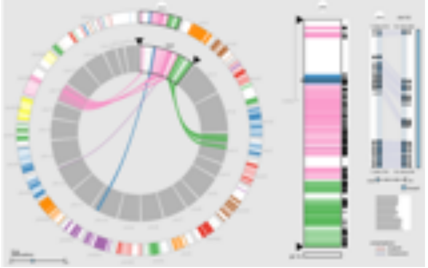
PF-1	premature advance: jumping forward over stages	general
PF-2	premature start: insufficient knowledge of vis literature	learn
PF-3	premature commitment: collaboration with wrong people	winnow
PF-4	no real data available (yet)	winnow
PF-5	insufficient time available from potential collaborators	winnow
PF-6	no need for visualization: problem can be automated	winnow
PF-7	researcher expertise does not match domain problem	winnow
PF-8	no need for research: engineering vs. research project	winnow
PF-9	no need for change: existing tools are good enough	winnow

- comparison to related methodologies

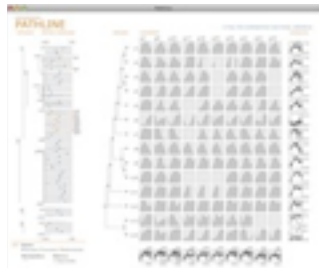
Lessons learned from the trenches: 21 between us



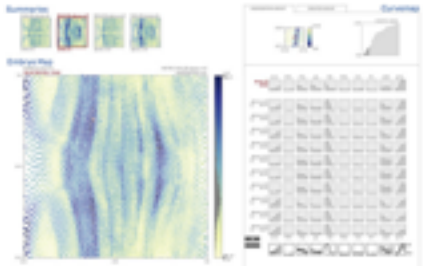
Cerebral
genomics



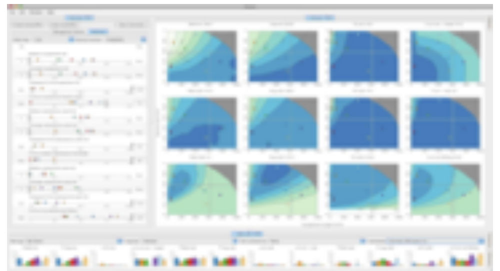
MizBee
genomics



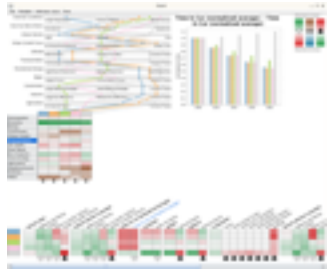
Pathline
genomics



MulteeSum
genomics



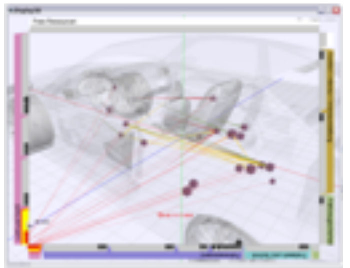
Vismon
fisheries management



QuestVis
sustainability

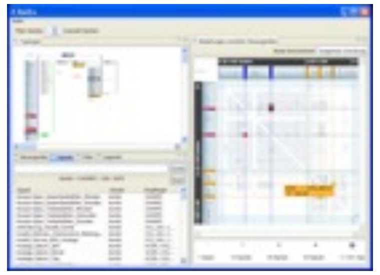
WiKeVis
in-car networks

MostVis
in-car networks

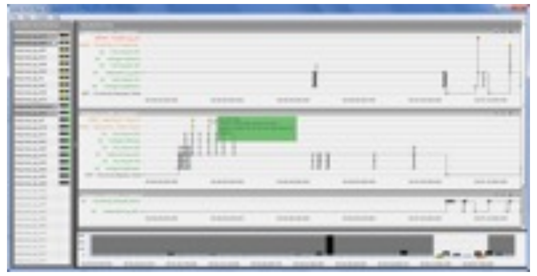


Car-X-Ray
in-car networks

ProgSpy2010
in-car networks



RelEx
in-car networks

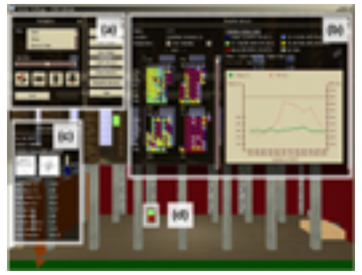


Cardiogram
in-car networks

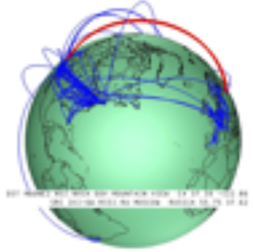
AutobahnVis
in-car networks

VisTra
in-car networks

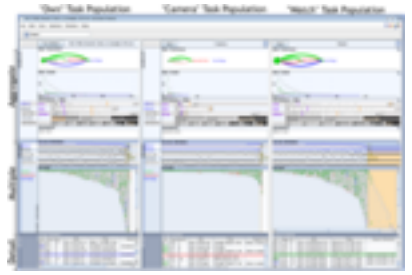
Constellation
linguistics



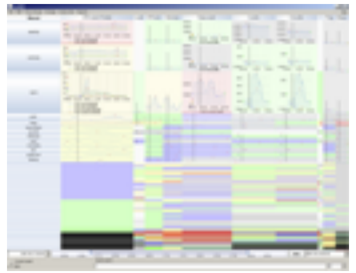
LibVis
cultural heritage



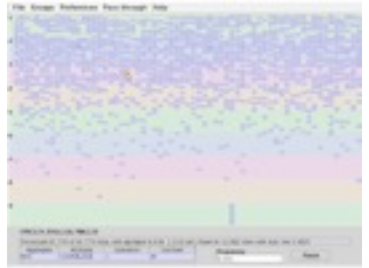
Caidants
multicast



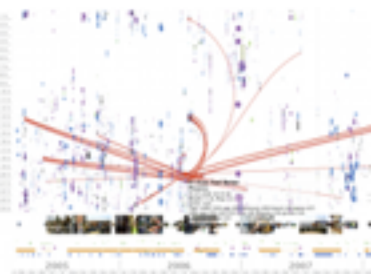
SessionViewer
web log analysis



LiveRAC
server hosting

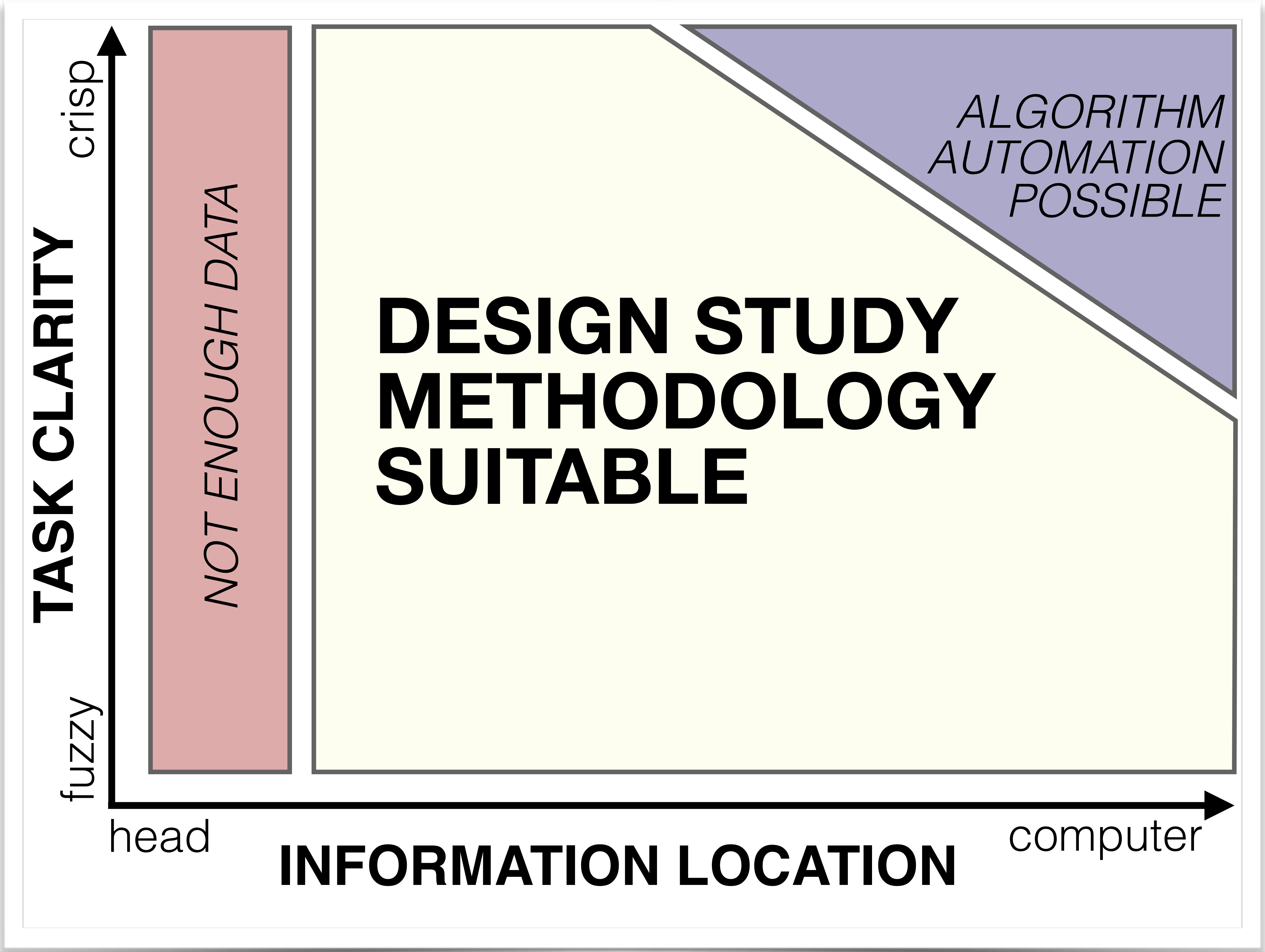


PowerSetViewer
data mining

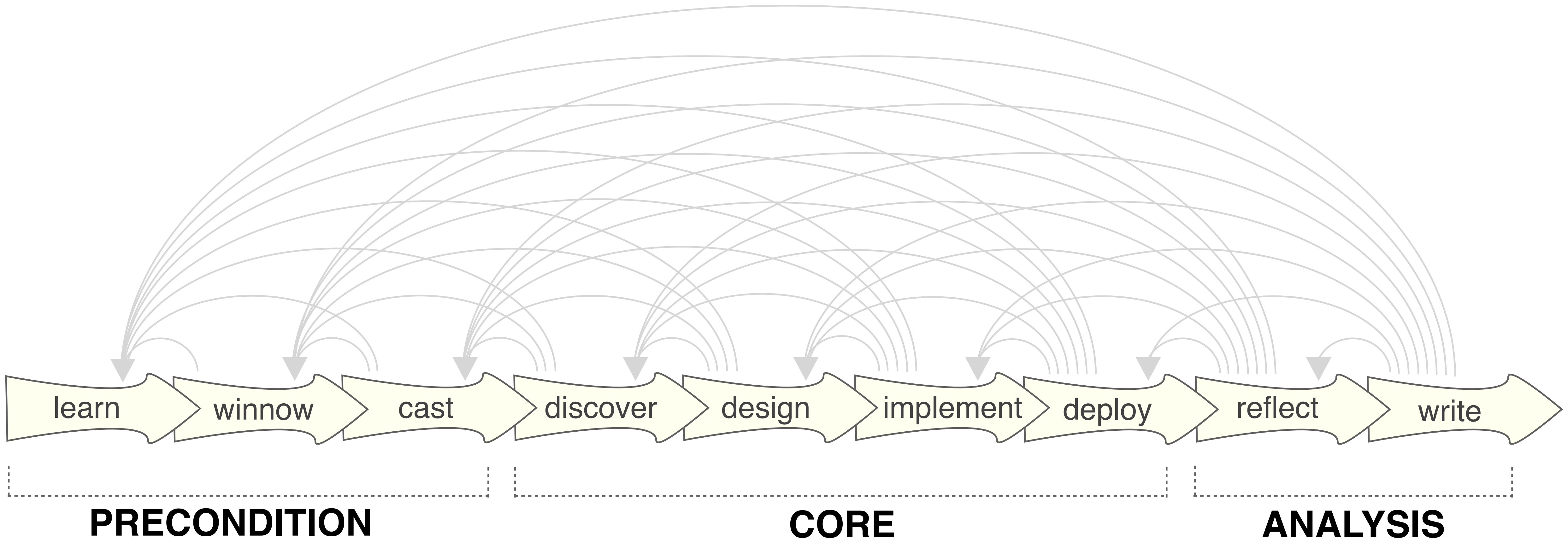


LastHistory
music listening

Design study methodology: definitions

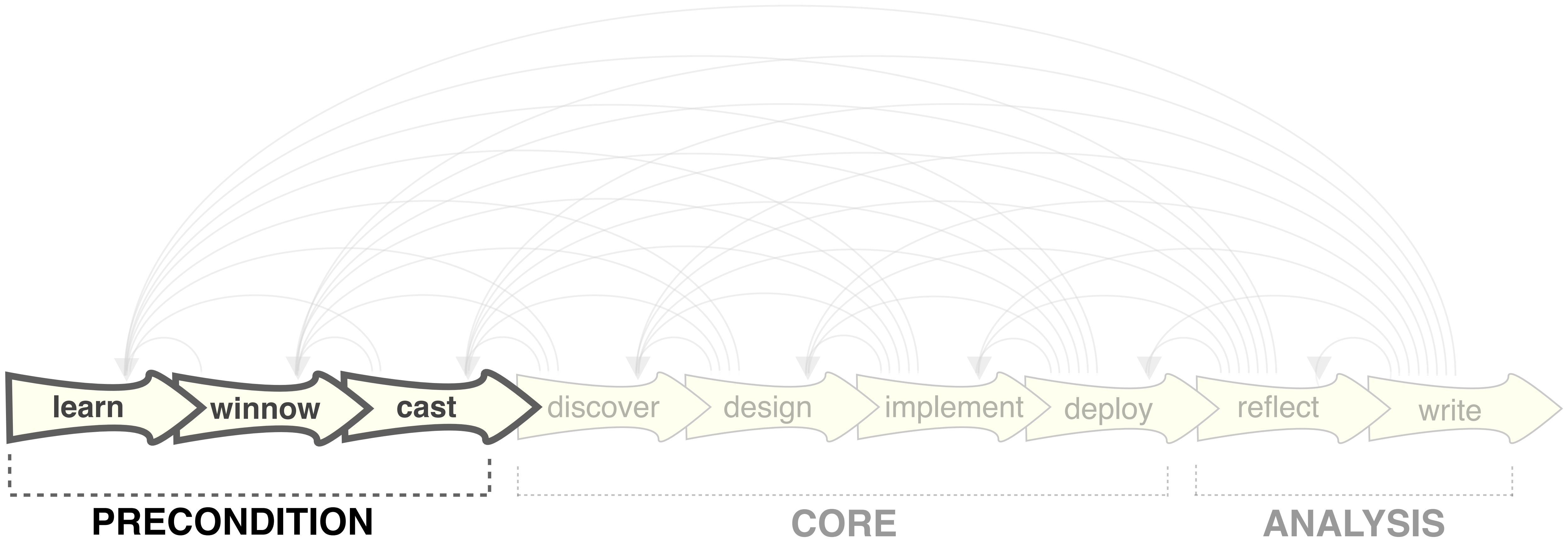


9 stage framework



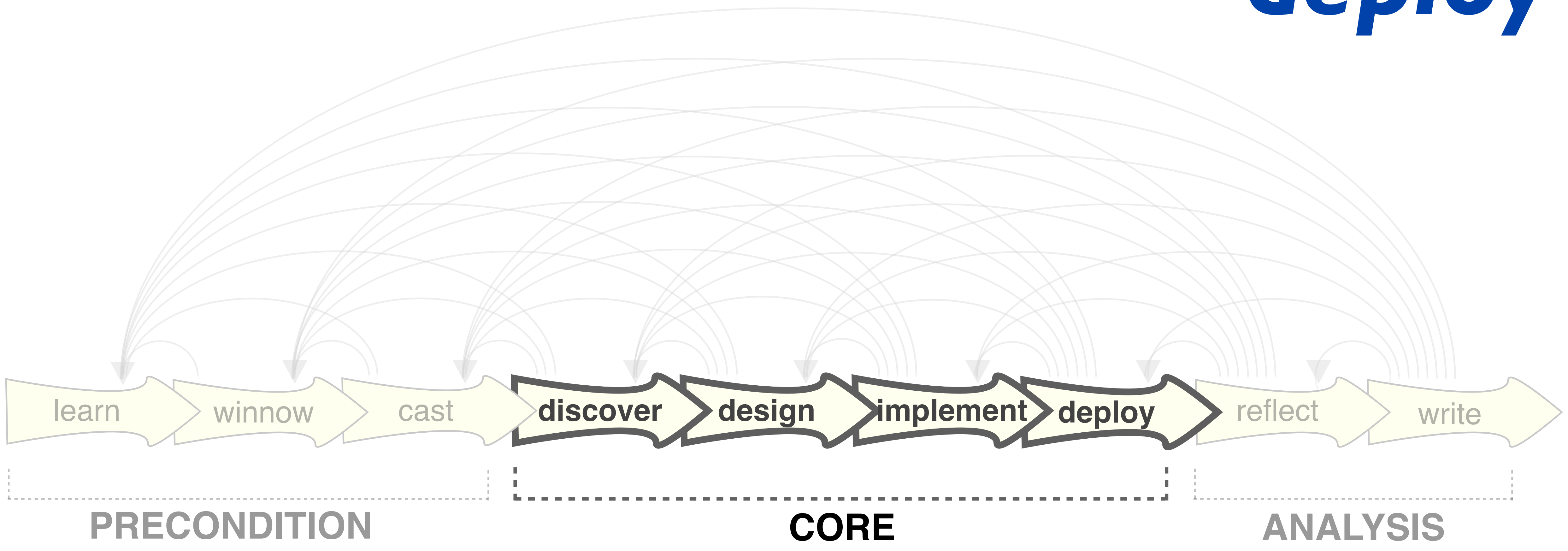
9-stage framework

learn
winnow
cast



9-stage framework

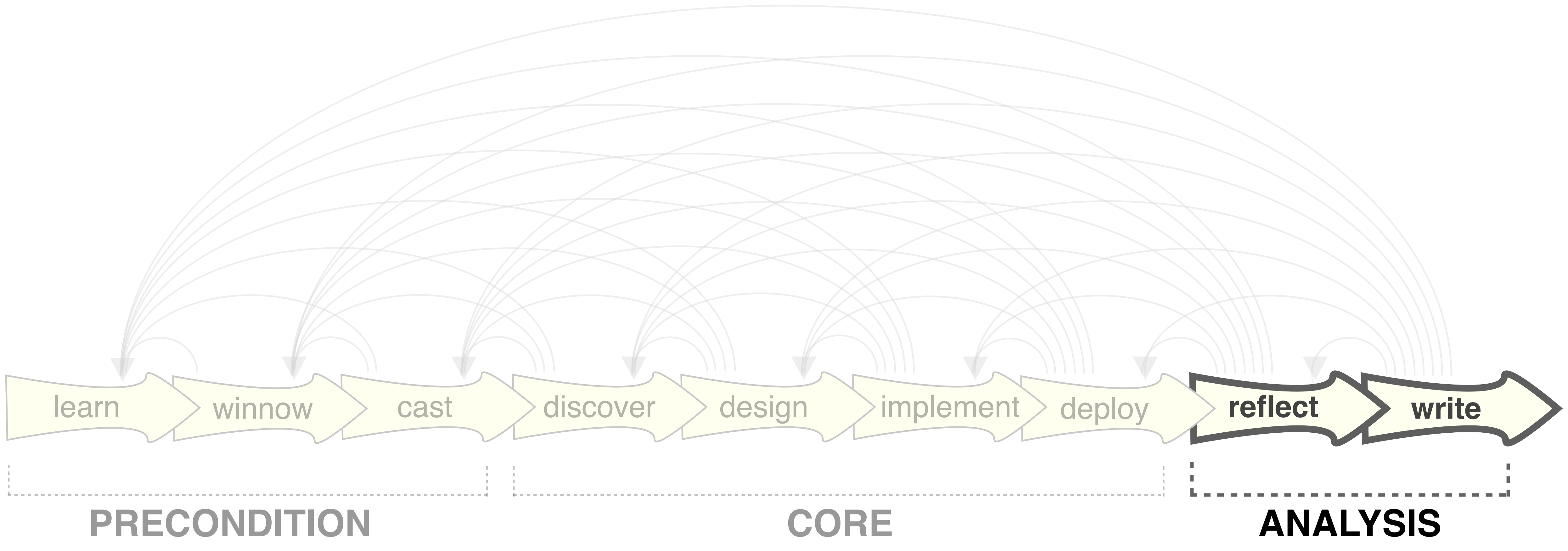
discover
design
implement
deploy



9-stage framework

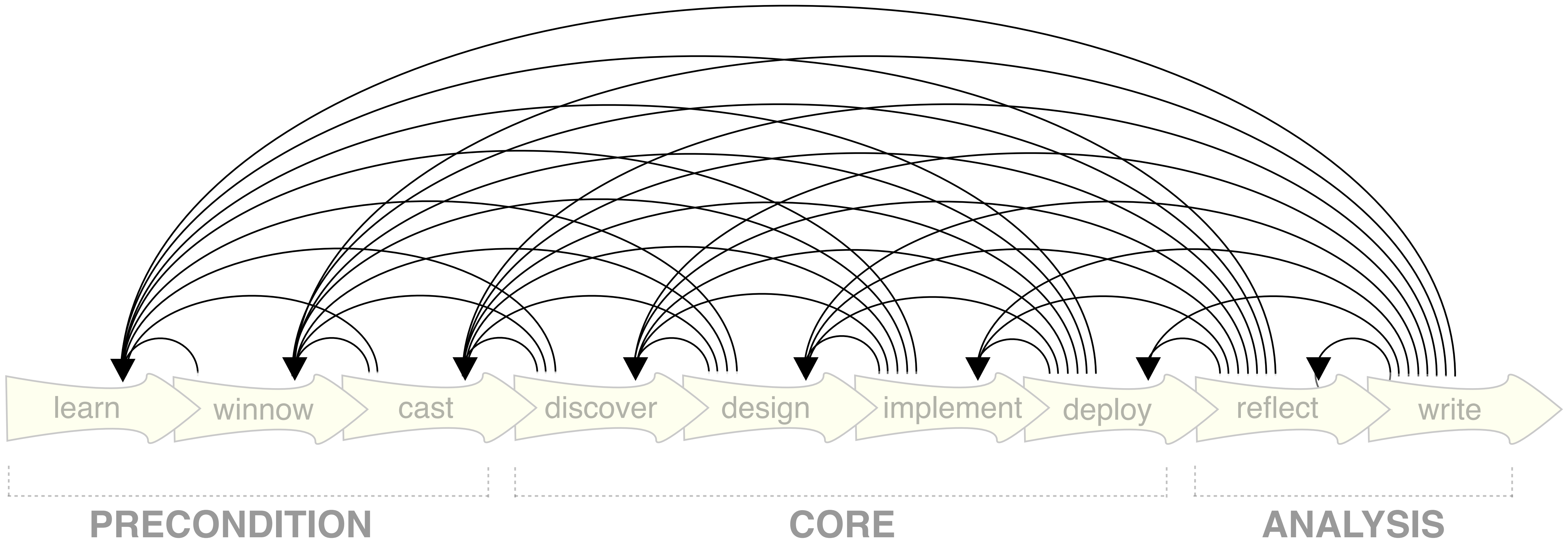
reflect
write

- guidelines: confirm, refine, reject, propose



9-stage framework

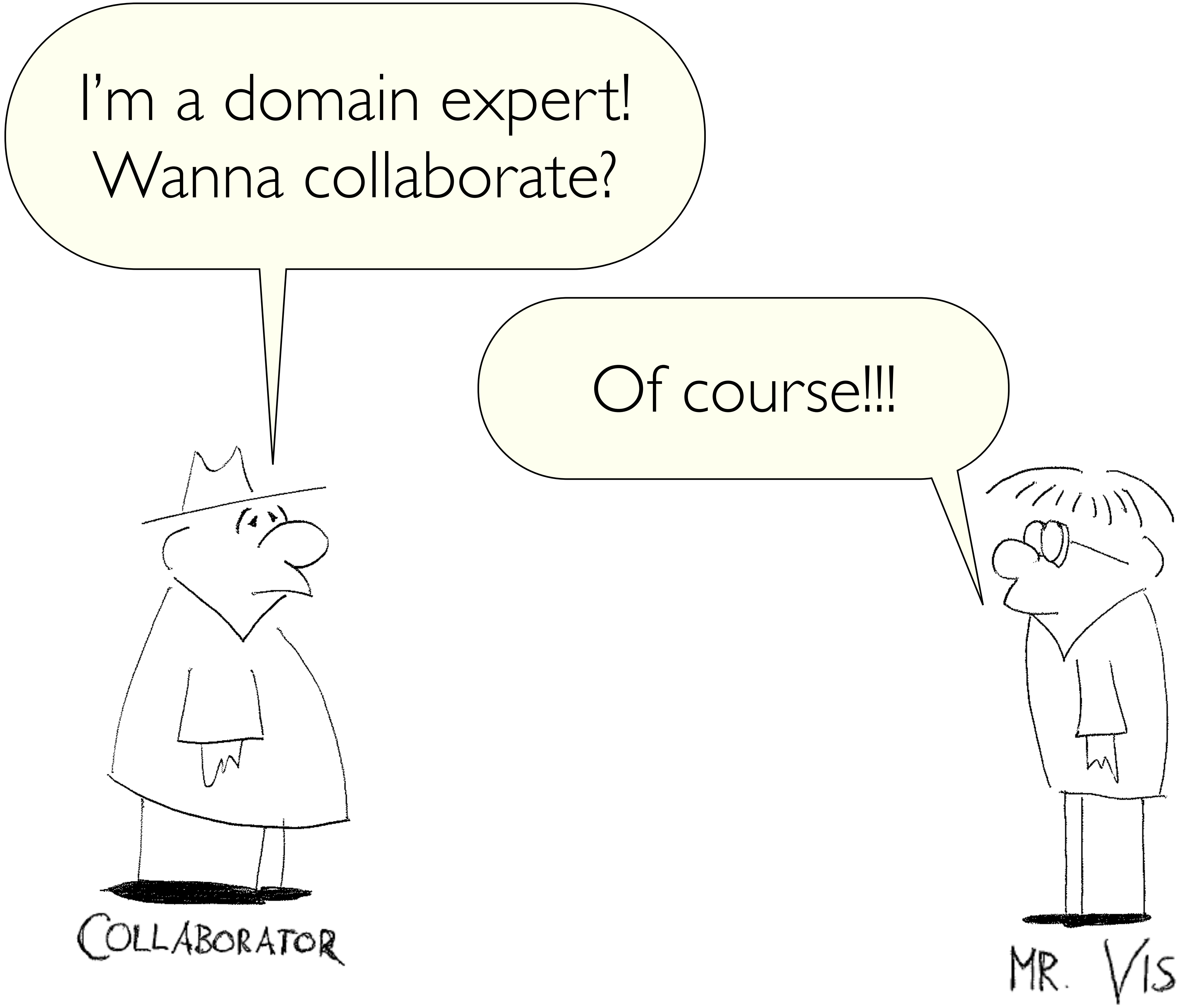
iterative



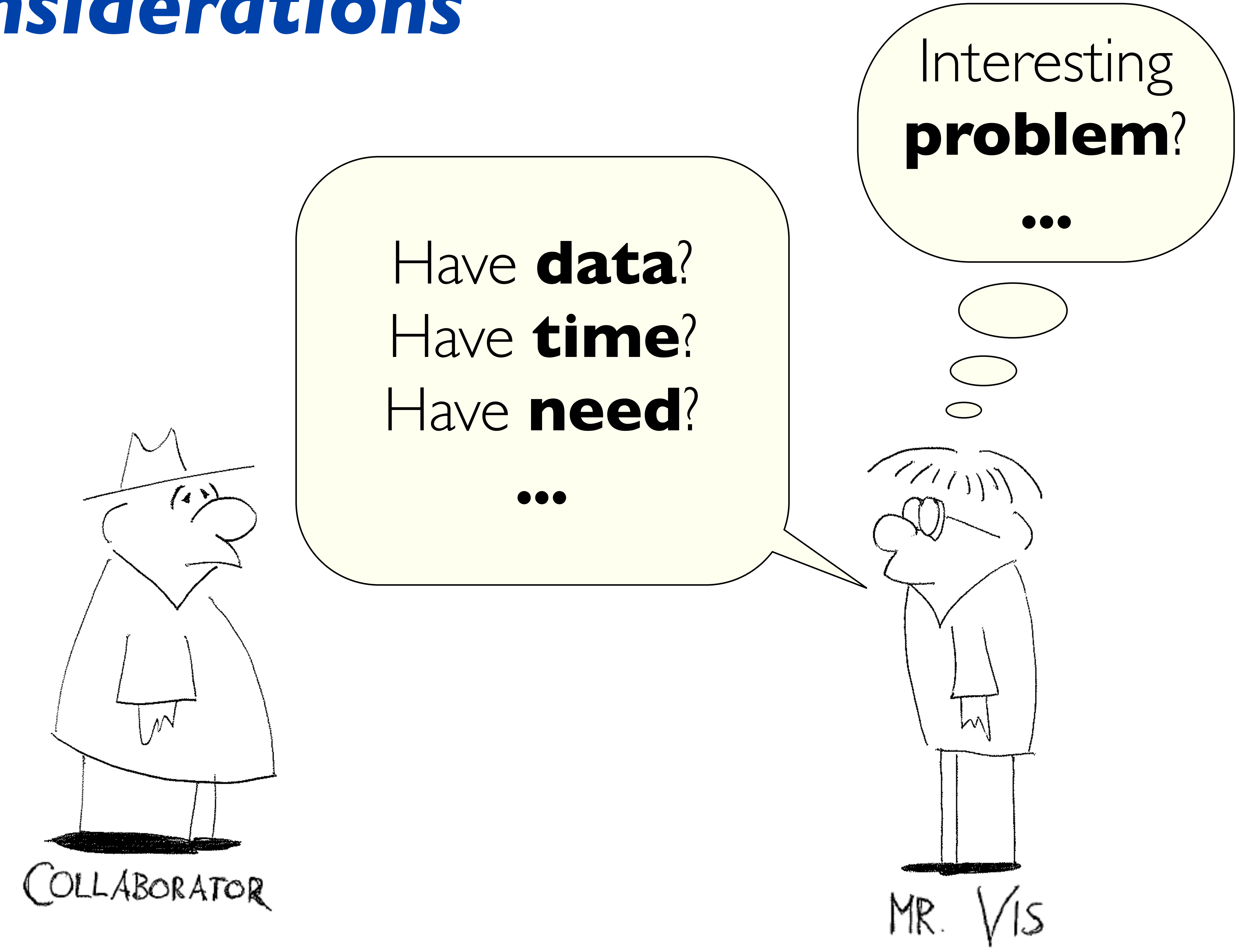
Design study methodology: 32 pitfalls

- and how to avoid them

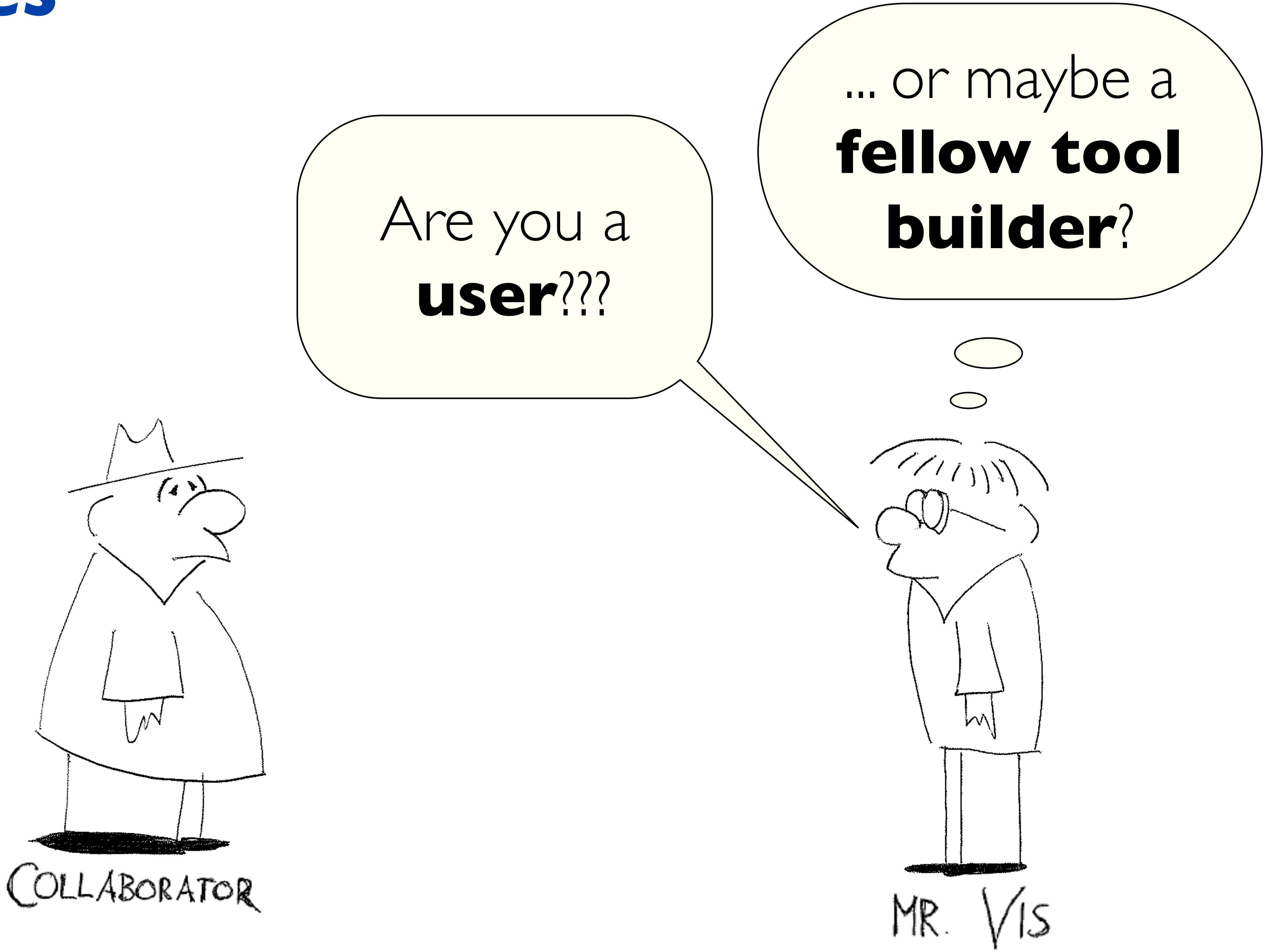
PF-1	premature advance: jumping forward over stages	general
PF-2	premature start: insufficient knowledge of vis literature	learn
PF-3	premature commitment: collaboration with wrong people	winnow
PF-4	no real data available (yet)	winnow
PF-5	insufficient time available from potential collaborators	winnow
PF-6	no need for visualization: problem can be automated	winnow
PF-7	researcher expertise does not match domain problem	winnow
PF-8	no need for research: engineering vs. research project	winnow
PF-9	no need for change: existing tools are good enough	winnow


I'm a domain expert!
Wanna collaborate?
Of course!!!
COLLABORATOR
MR. VIS

considerations

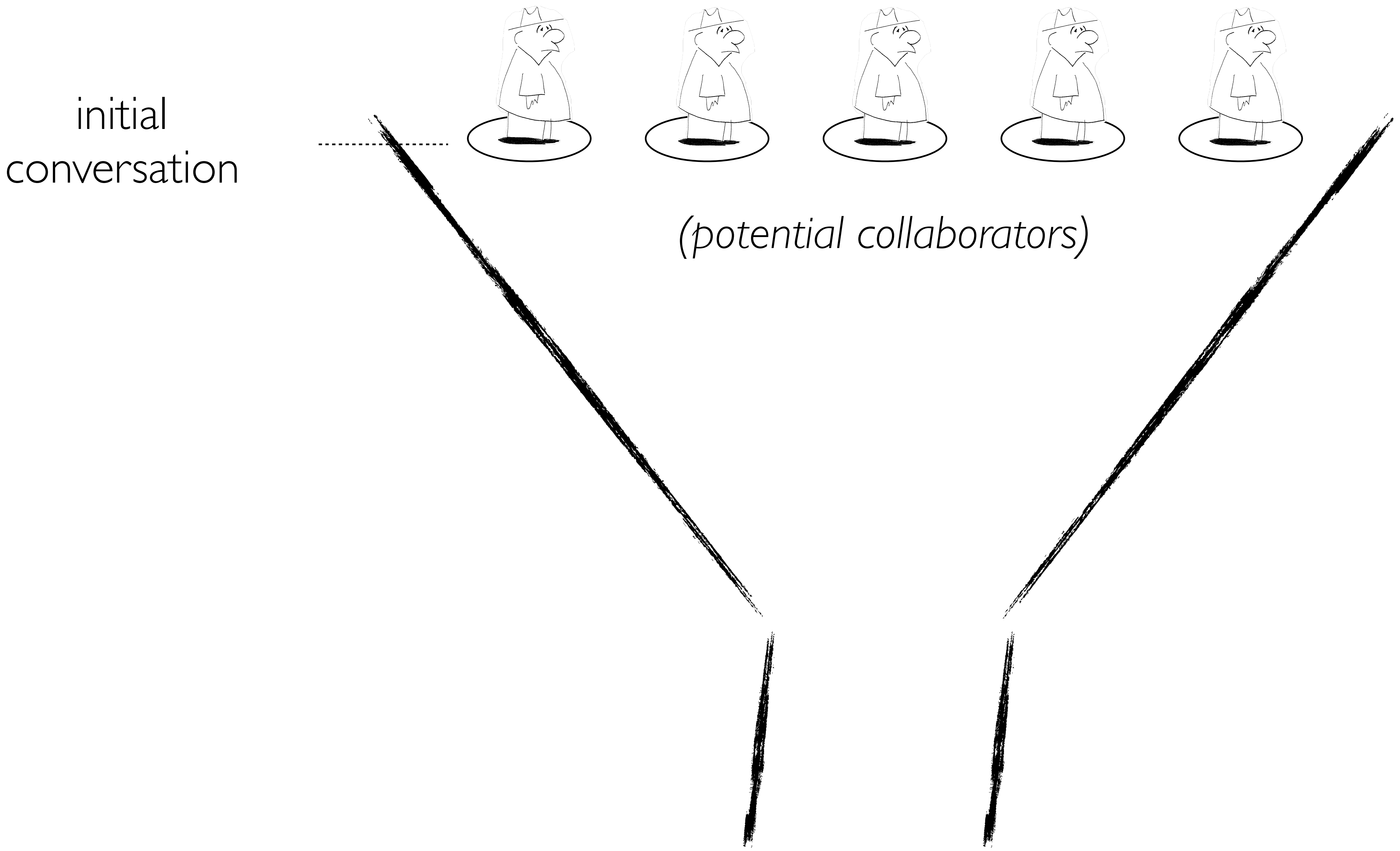


roles

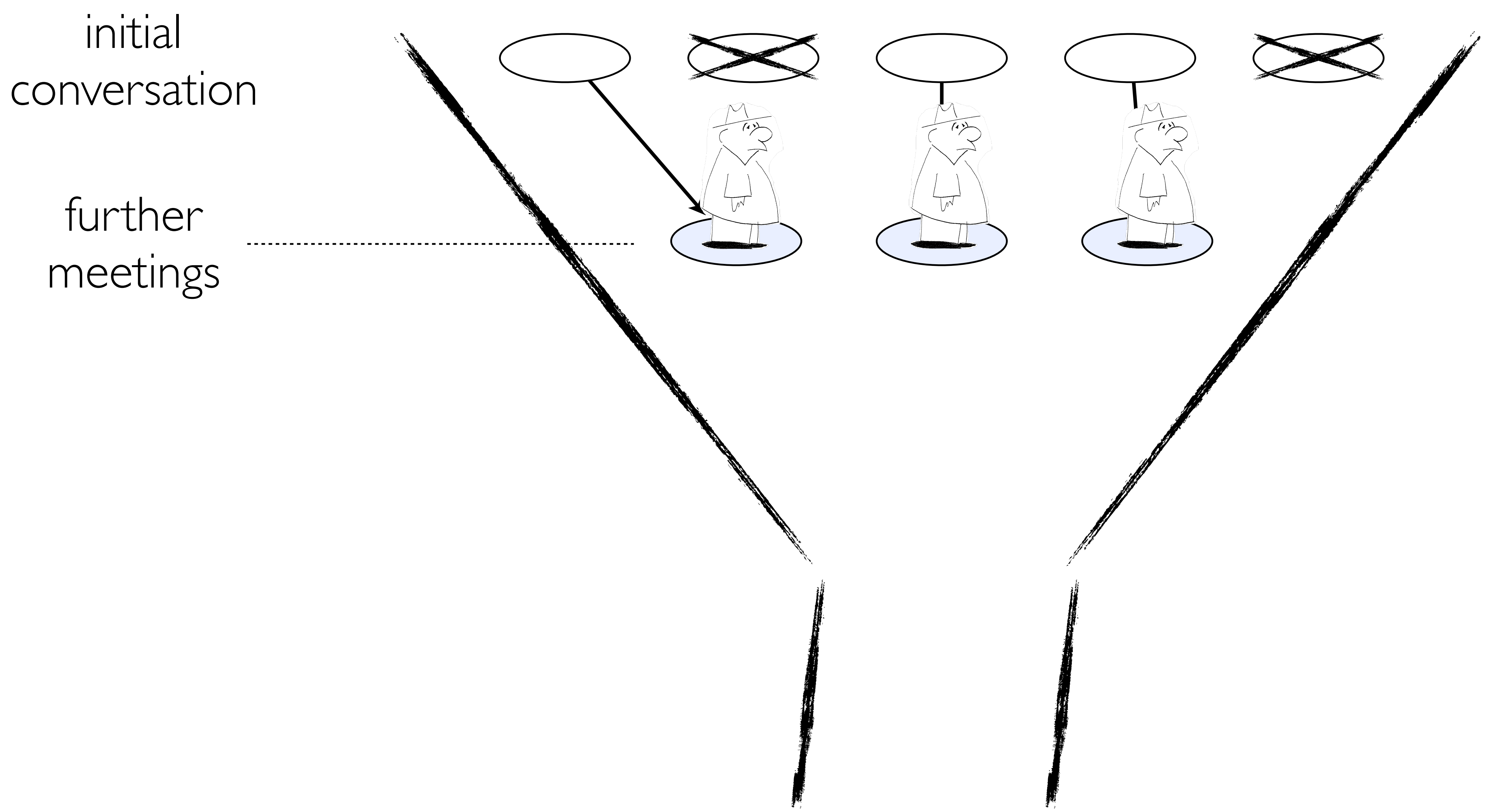

Are you a **user**???
... or maybe a **fellow tool builder**?
COLLABORATOR
MR. VIS

METAPHOR
Winnowing

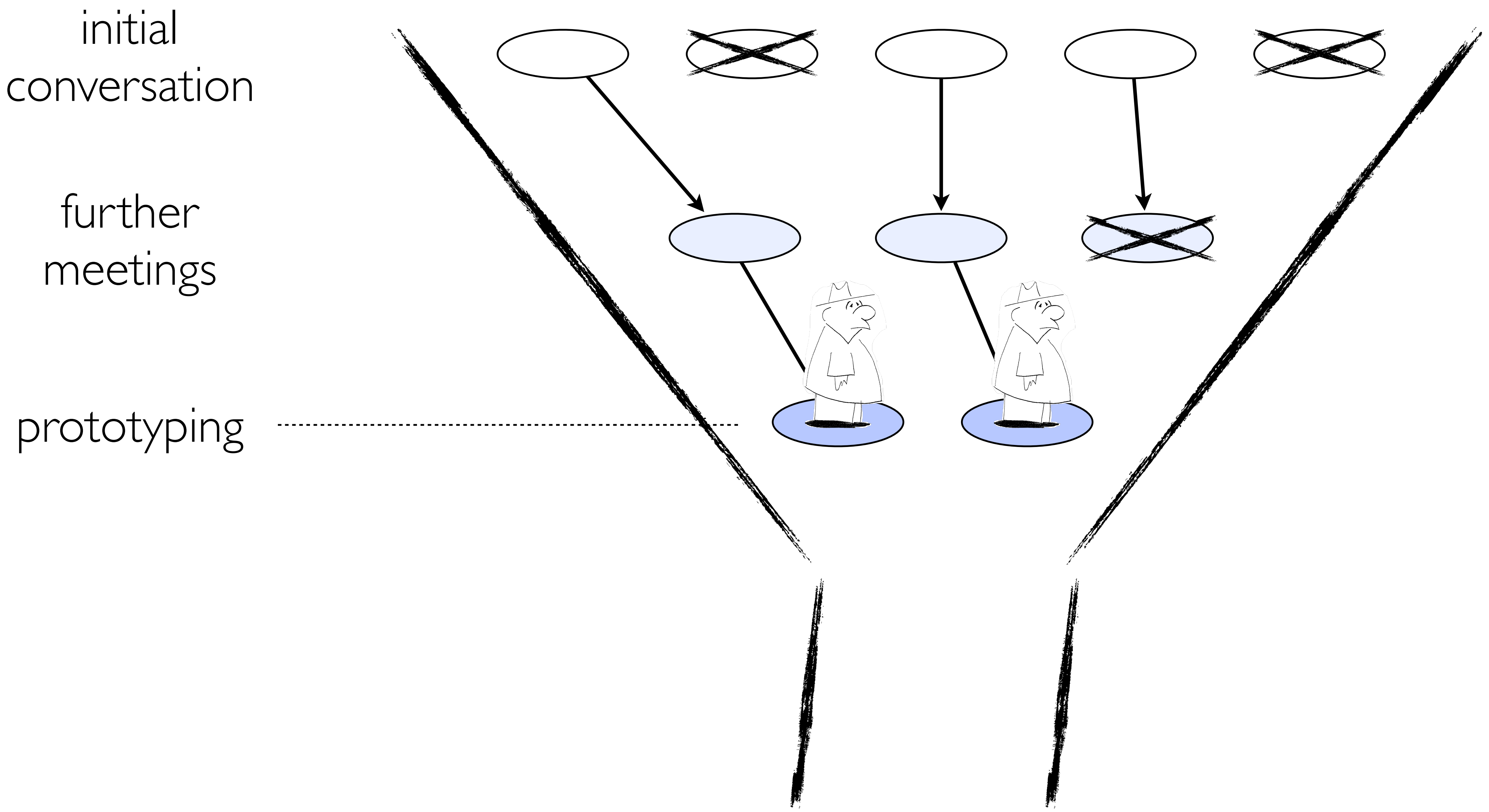
Collaborator winnowing



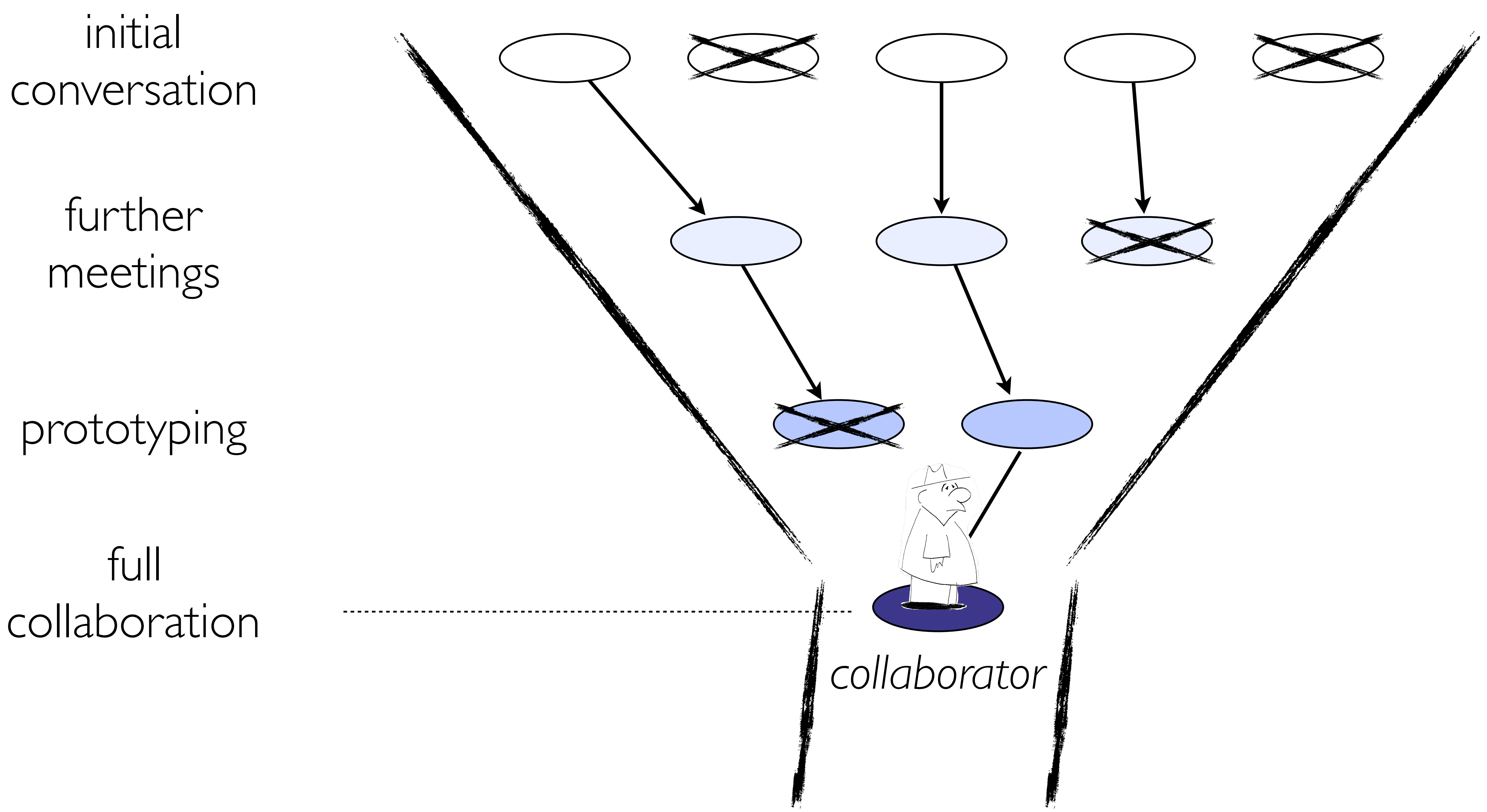
Collaborator winnowing



Collaborator winnowing



Collaborator winnowing



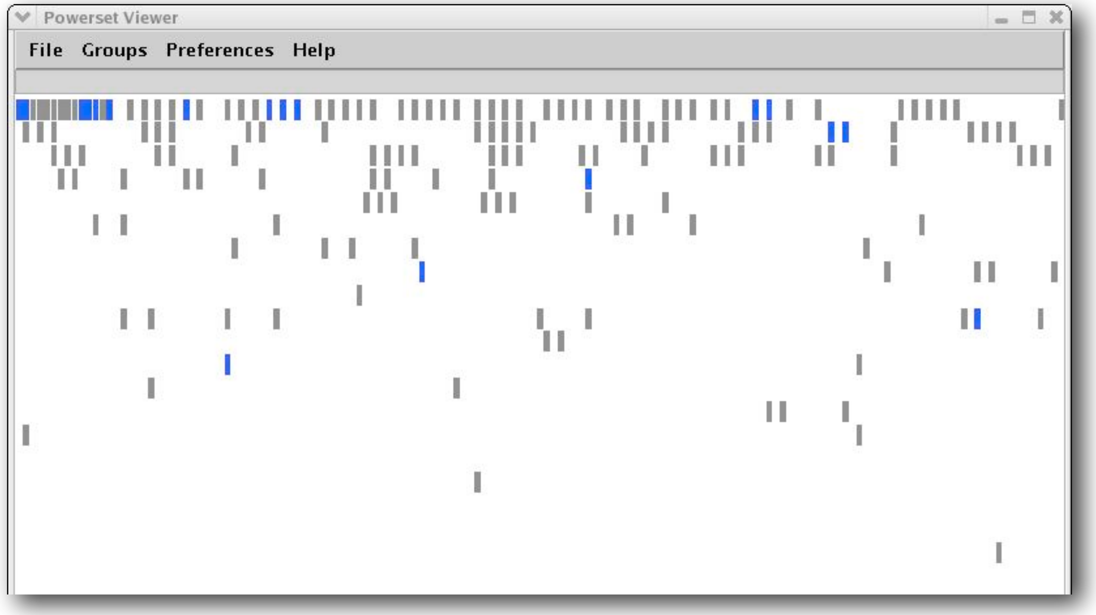
Collaborator winnowing

Talk with many, stay with few!

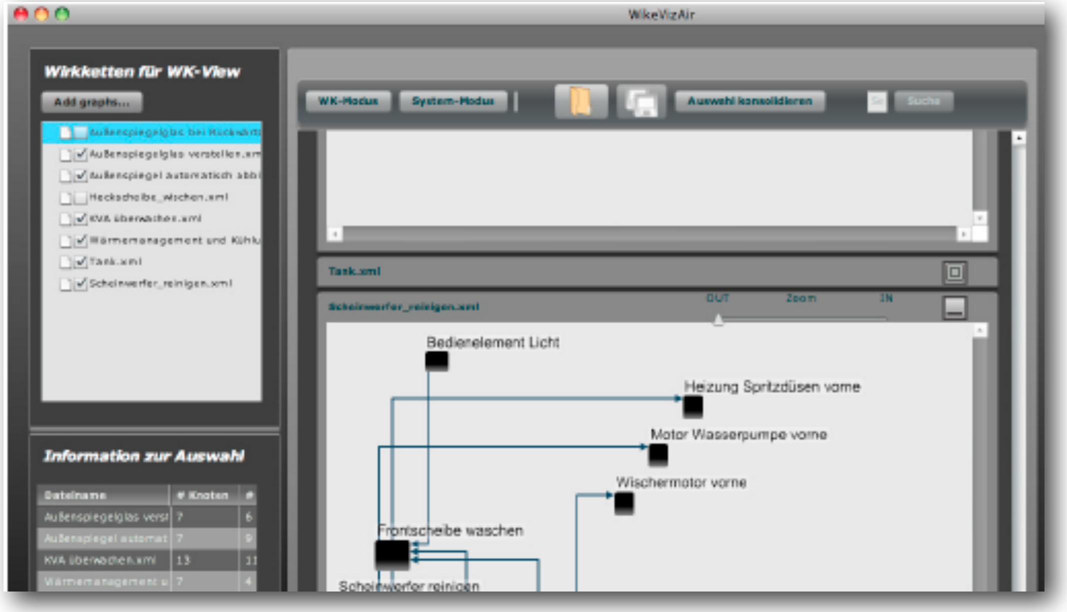
Example From The Trenches

Premature Collaboration!

PowerSet Viewer
2 years / 4 researchers



WikeVis
0.5 years / 2 researchers



EXAMPLE FROM THE TRENCHES

Premature Collaboration!

PowerSet Viewer
2 years / 4 researchers

WikeVis
0.5 years / 2 researchers

- Fellow tool builders

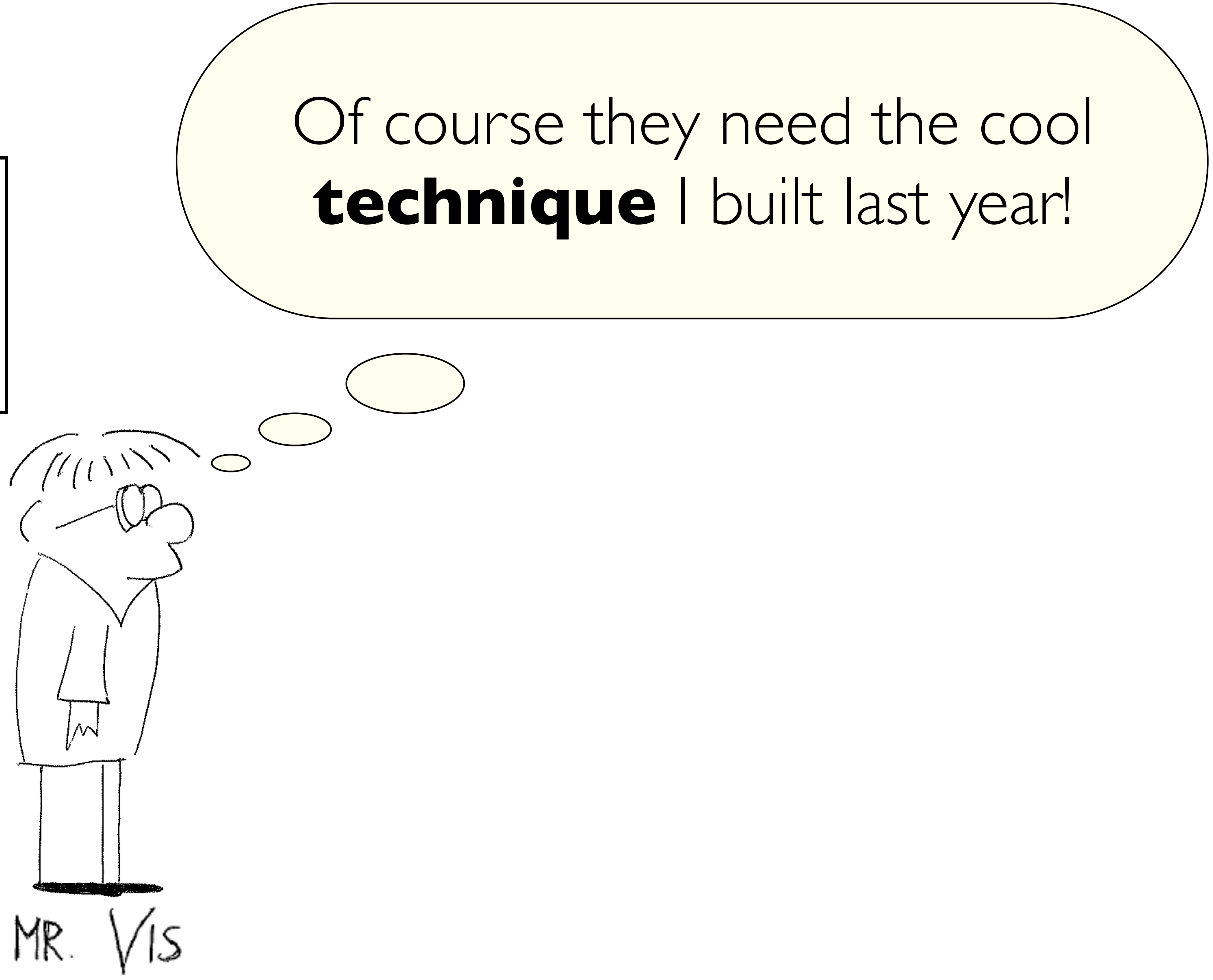
- Data promised

Design study methodology: 32 pitfalls

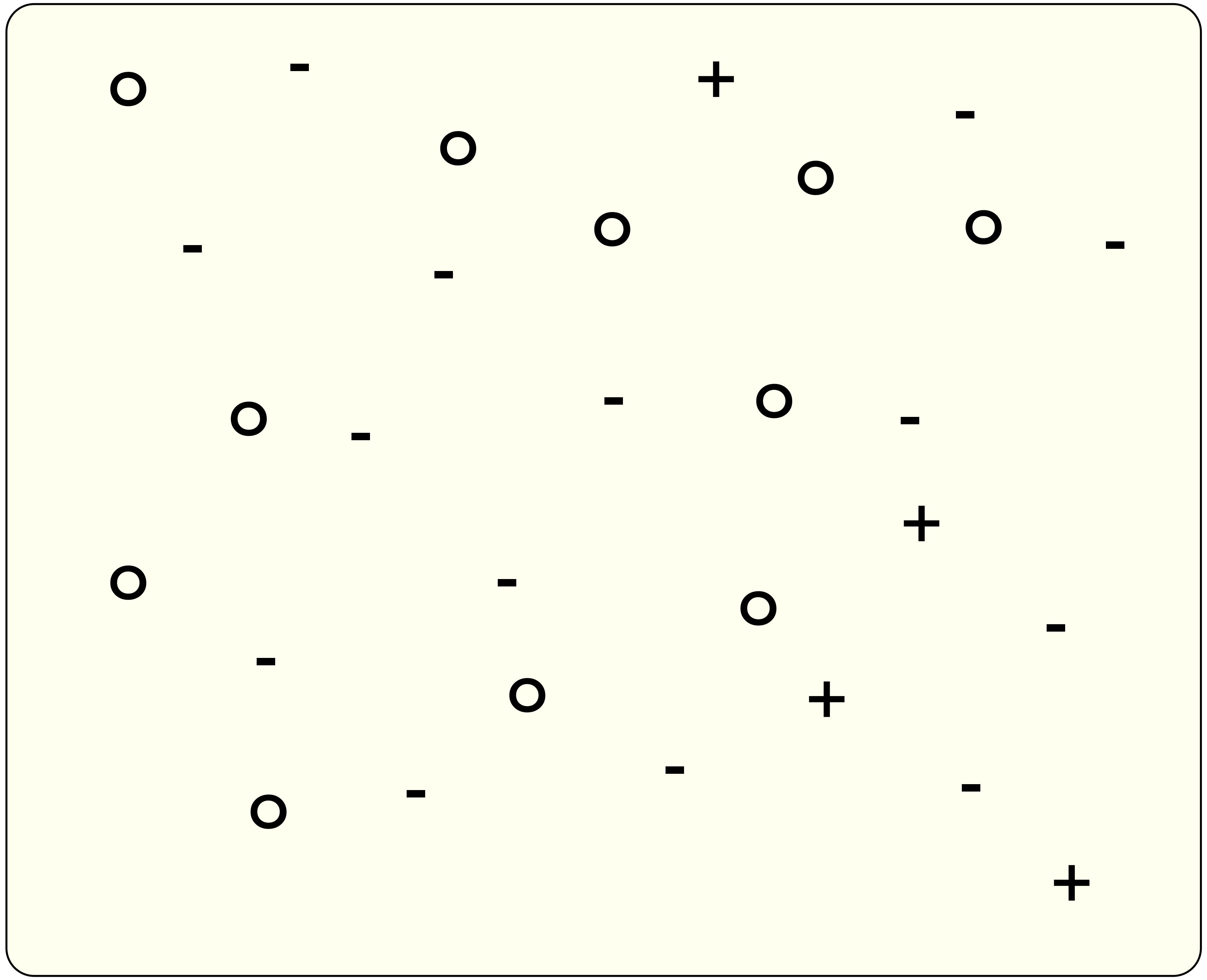
PF-10	no real/important/recurring task	winnow
PF-11	no rapport with collaborators	winnow
PF-12	not identifying front line analyst and gatekeeper before start	cast
PF-13	assuming every project will have the same role distribution	cast
PF-14	mistaking fellow tool builders for real end users	cast
PF-15	ignoring practices that currently work well	discover
PF-16	expecting *just talking* or *fly on wall* to work	discover
PF-17	experts focusing on visualization design vs. domain problem	discover
PF-18	learning their problems/language: too little / too much	discover
PF-19	abstraction: too little	design
PF-20	premature design commitment: consideration space too small	design

Pitfall

Premature Design Commitment



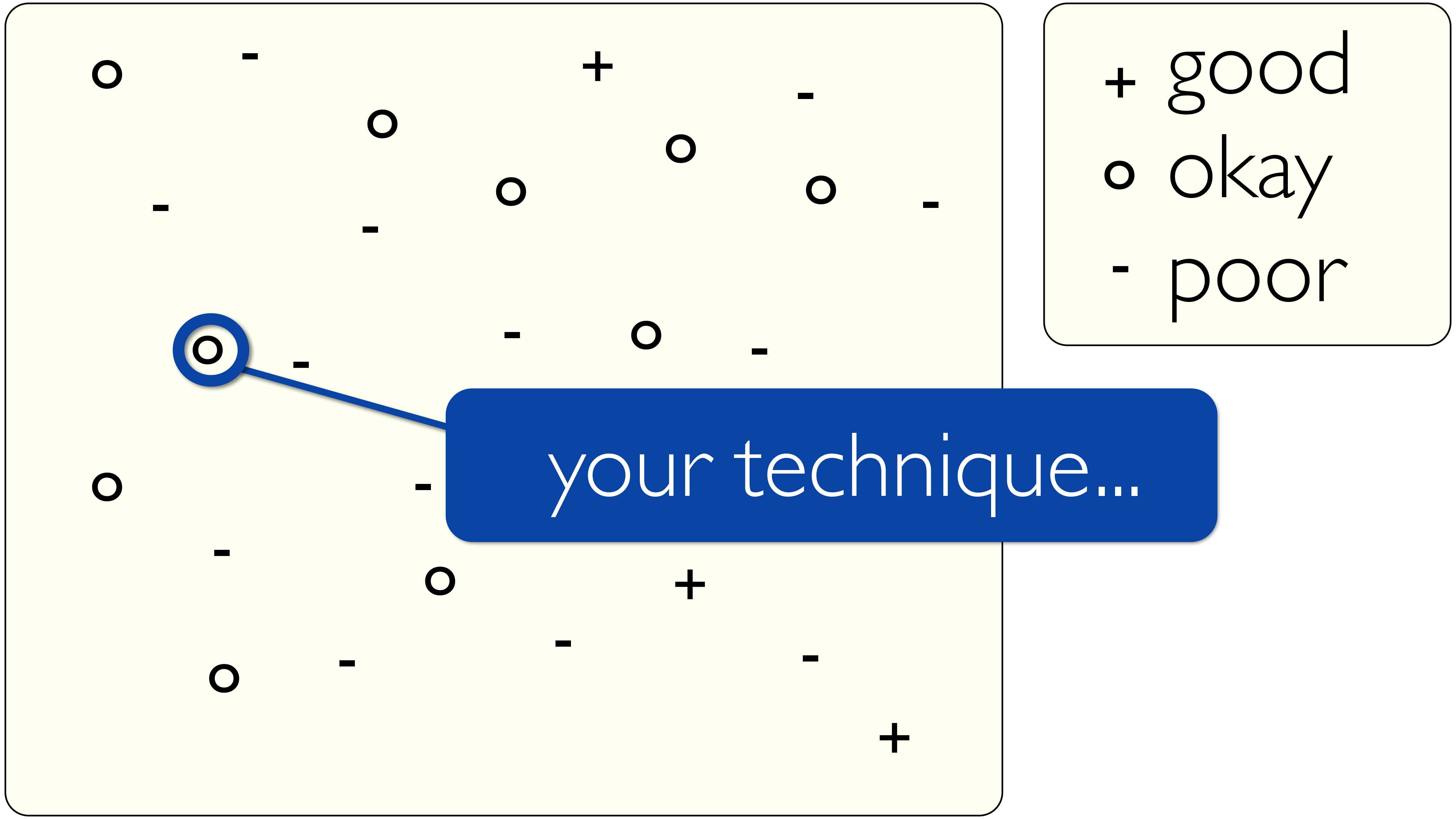
METAPHOR
Design Space



+ good
o okay
- poor

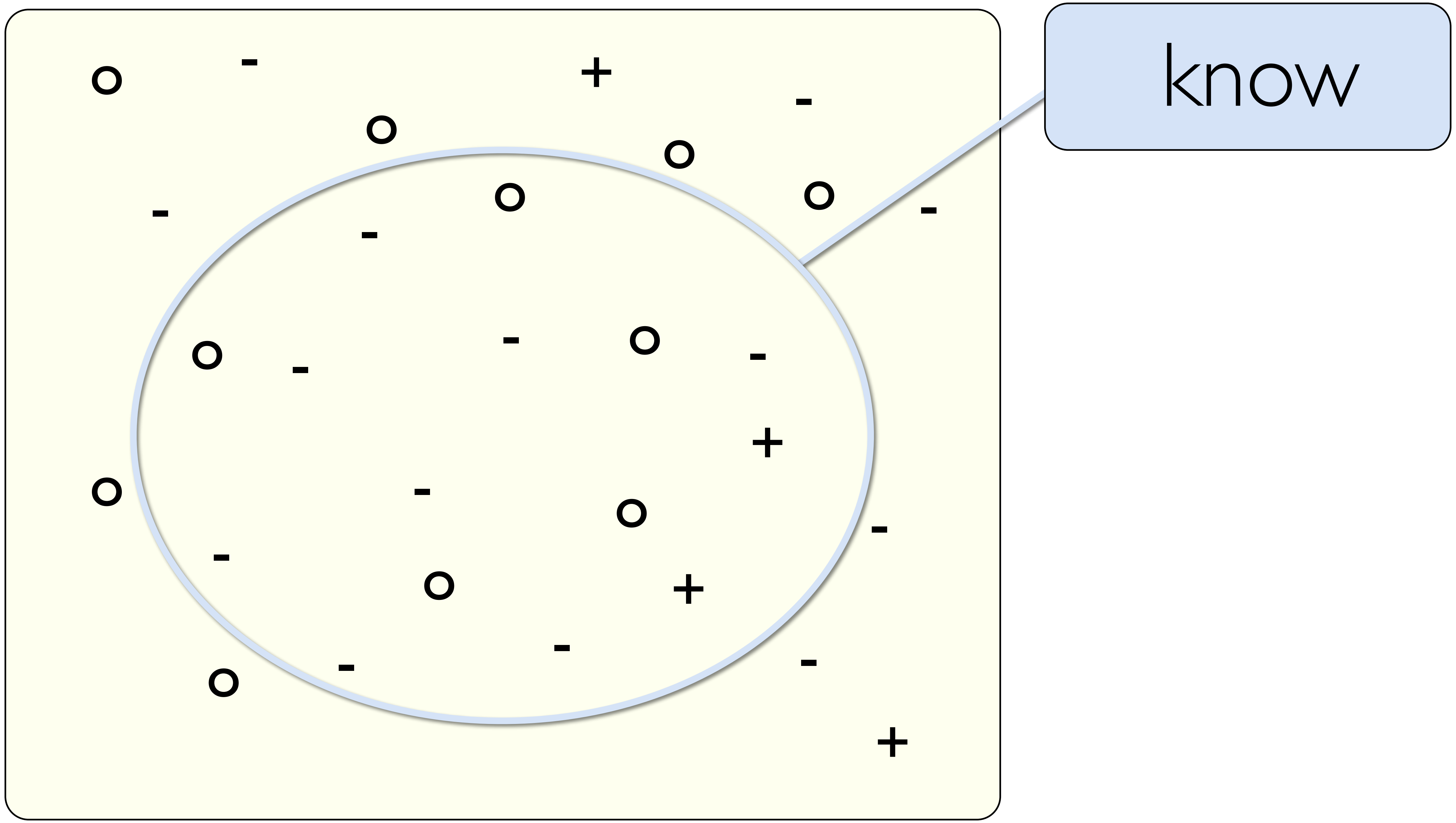
METAPHOR

Design Space

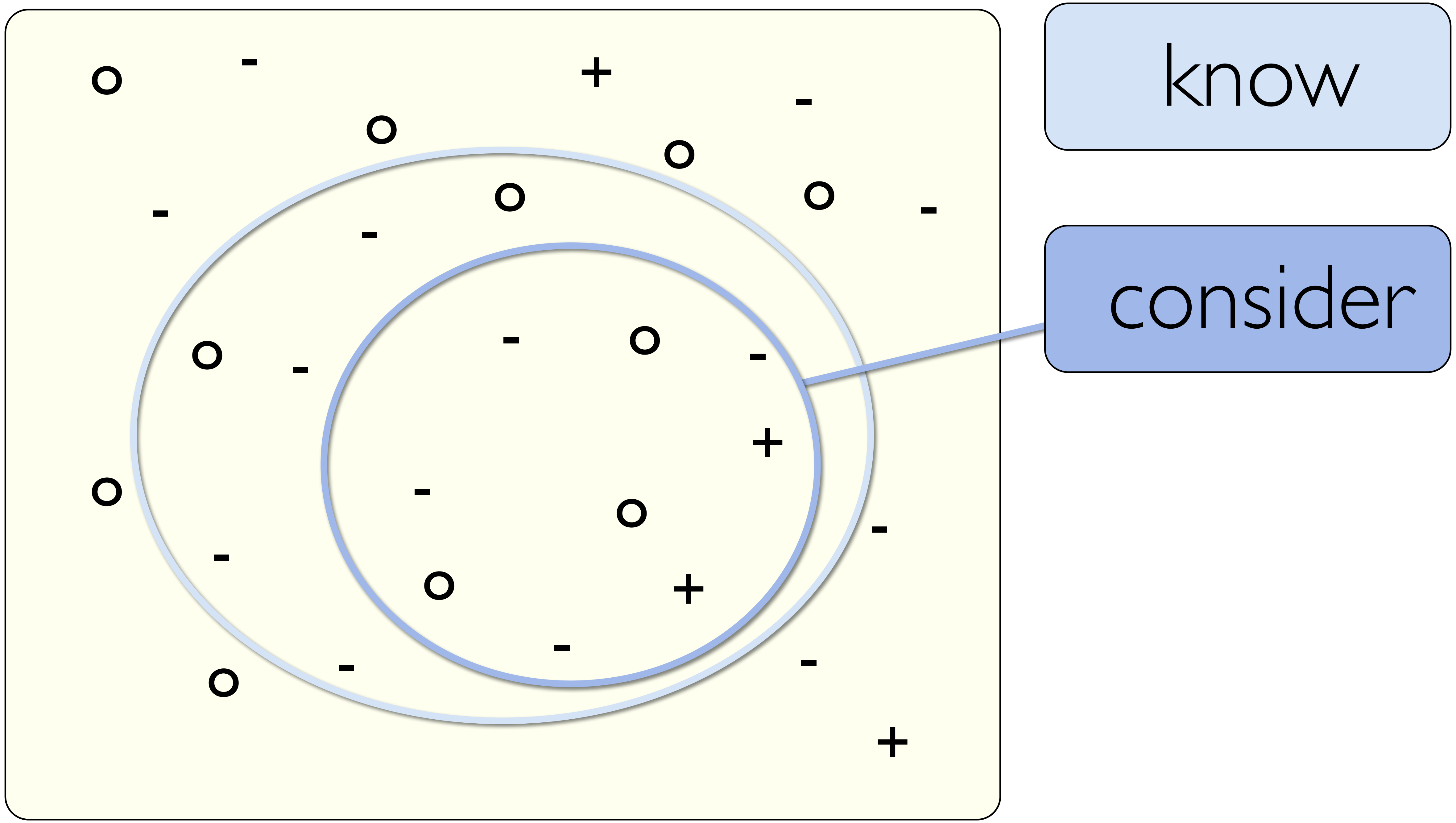


METAPHOR

Design Space

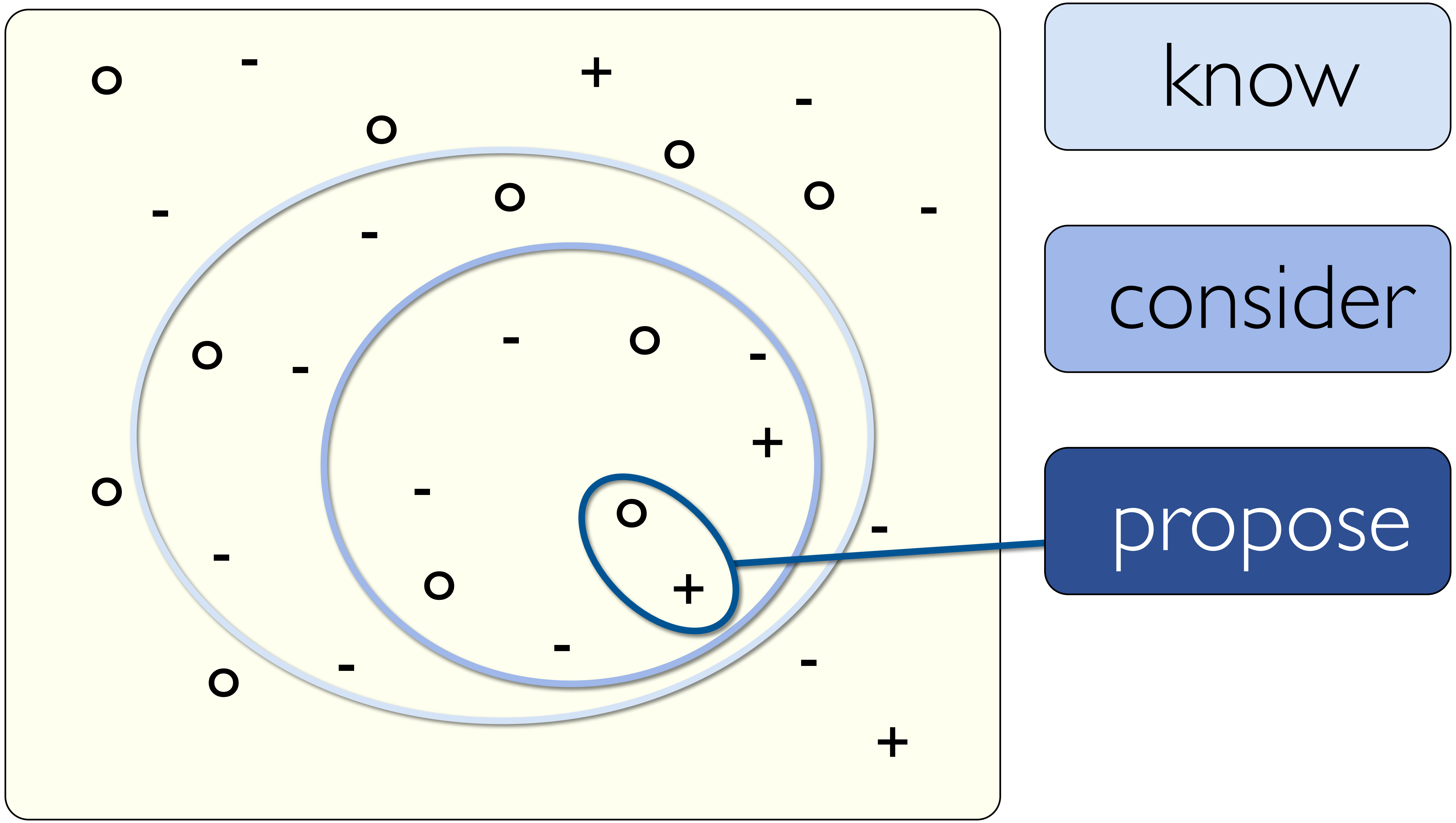


Metaphor
Design Space

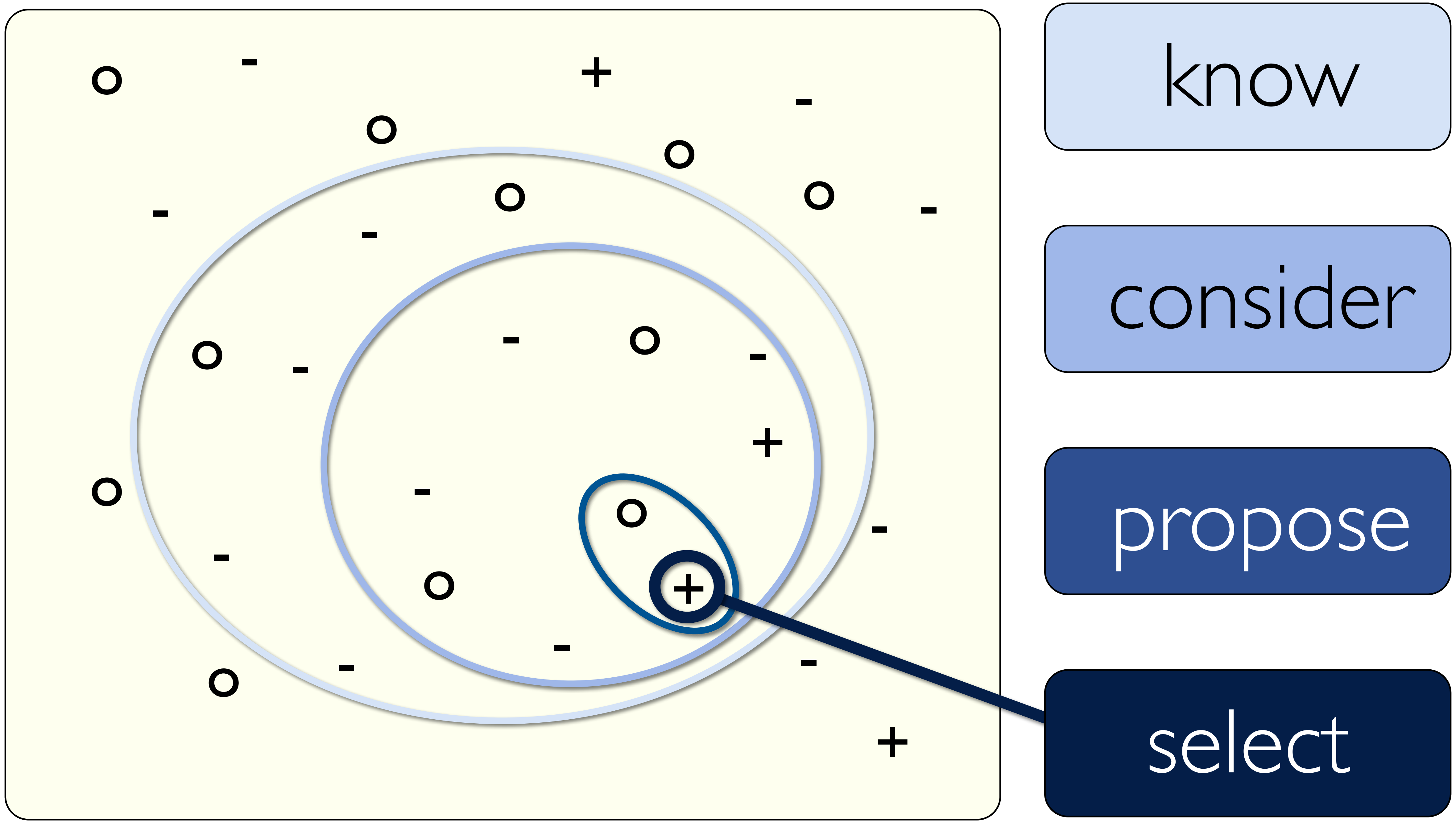


Metaphor

Design Space



Metaphor
Design Space



METAPHOR
Design Space

Think broad!

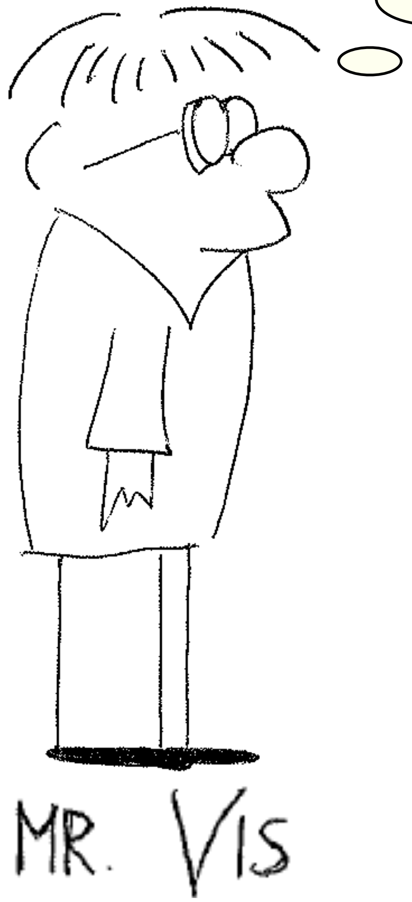
Design study methodology: 32 pitfalls

PF-21	mistaking technique-driven for problem-driven work	design
PF-22	nonrapid prototyping	implement
PF-23	usability: too little / too much	implement
PF-24	premature end: insufficient deploy time built into schedule	deploy
PF-25	usage study not case study: non-real task/data/user	deploy
PF-26	*liking* necessary but not sufficient for validation	deploy
PF-27	failing to improve guidelines: confirm, refine, reject, propose	reflect
PF-28	insufficient writing time built into schedule	write
PF-29	no technique contribution $\neq$ good design study	write
PF-30	too much domain background in paper	write
PF-31	story told chronologically vs. focus on final results	write
PF-32	premature end: win race vs. practice music for debut	write

PITFALL

PREMATURE PUBLISHING

I can write a design study **paper** in a week!



"writing is research"

[Wolcott: Writing up qualitative research, 2009]

Metaphor

Horse Race vs. Music Debut

Must be first!

Am I ready?

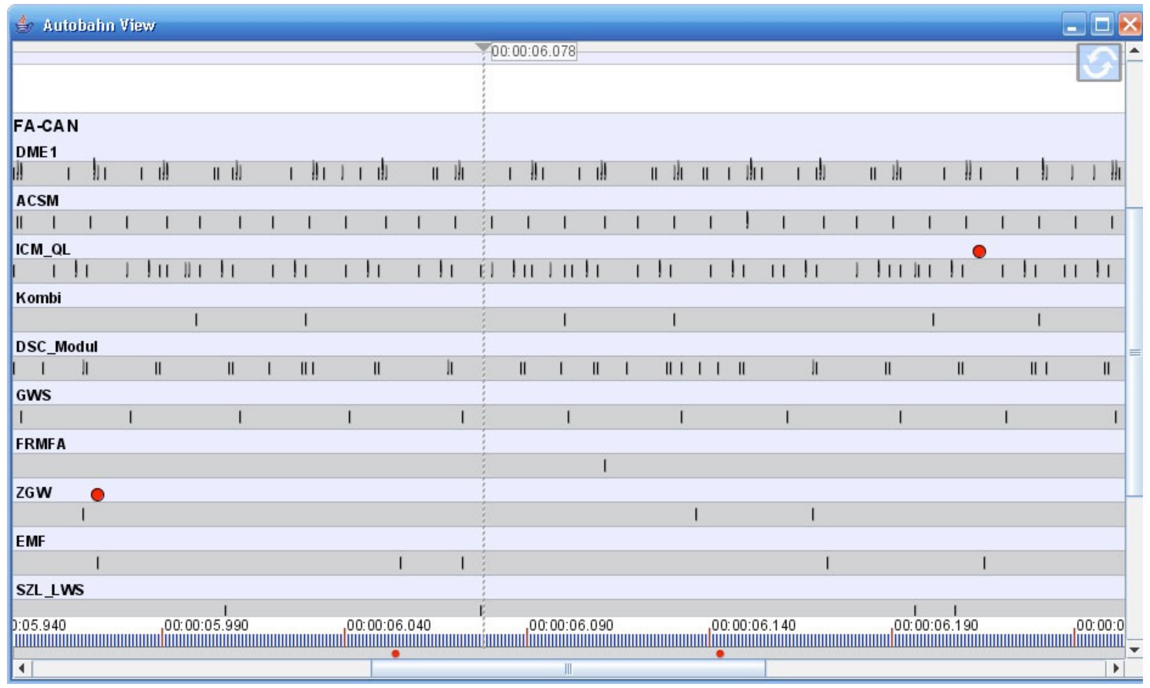
technique-driven

problem-driven

Example From The Trenches

Don’t step on your own toes!

First design round published



AutobahnVis 1.0

[Sedlmair et al., Smart Graphics, 2009]

Subsequent work not stand-alone paper

AutobahnVis 2.0

[Sedlmair et al., Information Visualization 10(3), 2011]

Reflections from the stacks: Wholesale adoption inappropriate

- ethnography
  - rapid, goal-directed fieldwork
- grounded theory
  - not empty slate: vis background is key
- action research
  - aligned
    - intervention as goal
    - transferability not reproducibility
    - personal involvement is key
  - opposition
    - translation of participant concepts into visualization language
    - researcher lead not facilitate design
    - orthogonal to vis concerns: participants as writers, adversarial to status quo, postmodernity

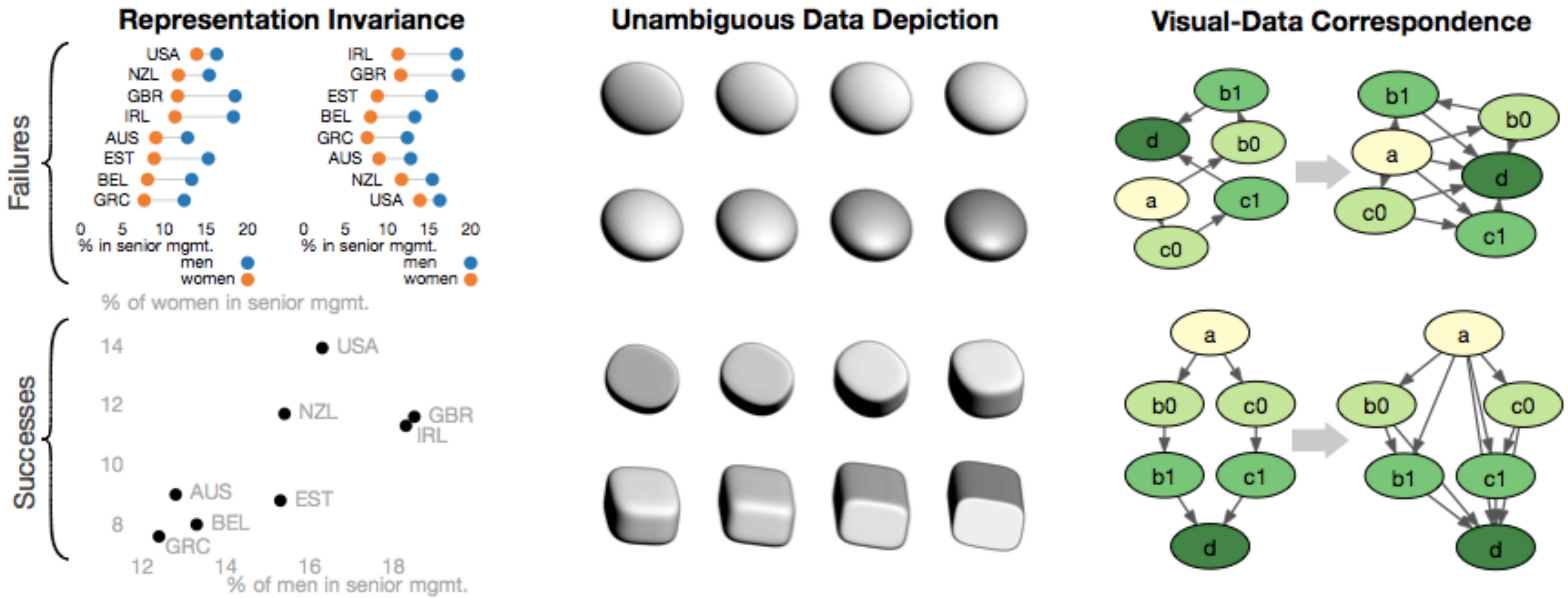
Next Steps

What-Why-How Analysis

- this approach is not the only way to analyze visualizations!
  - one specific framework intended to help you think
  - other frameworks support different ways of thinking
    - following: one interesting example

Algebraic Process for Visualization Design

- which mathematical structures in data are preserved and reflected in vis
  - negation, permutation, symmetry, invariance



(a) Employment rates
(b) Tensor Glyphs
(c) Directed Graphs

[Fig 1. An Algebraic Process for Visualization Design. Carlos Scheidegger and Gordon Kindlmann. IEEE TVCG (Proc. InfoVis 2014), 20(12):2181-2190.]

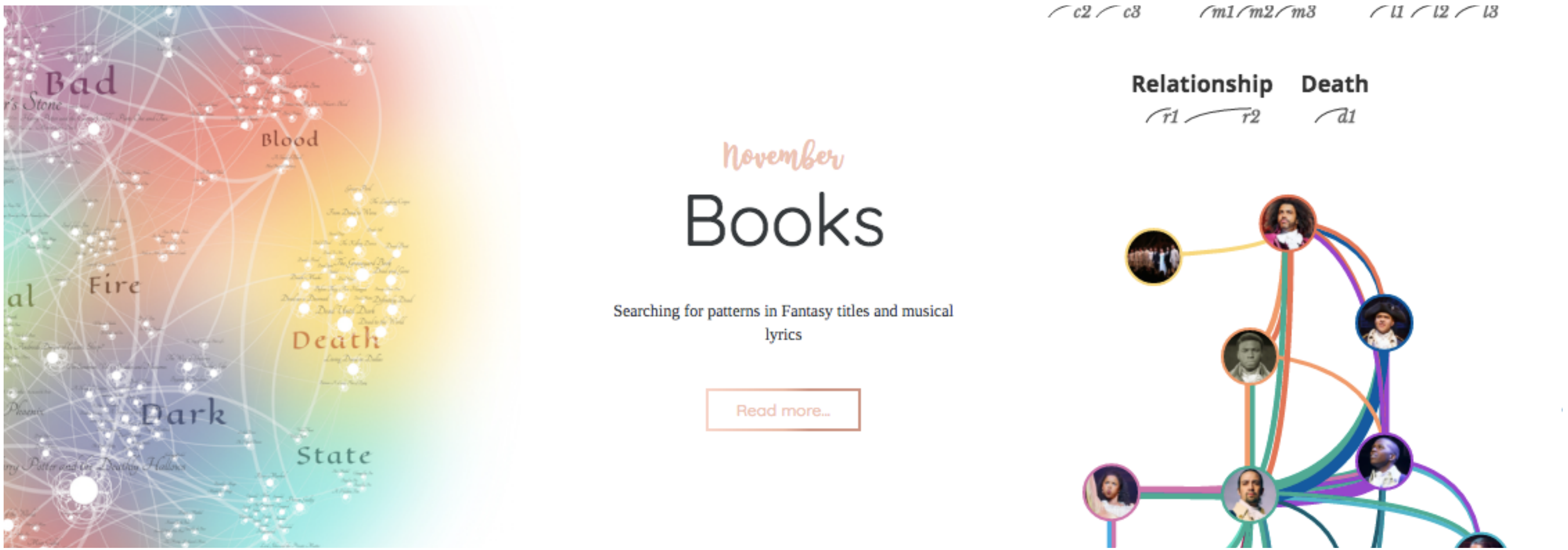
Algebraic process: Vocabulary

- **invariance** violation: single dataset, many visualizations
  - **hallucinator**
- **unambiguity** violation: many datasets, same vis
  - data change invisible to viewer
    - **confuser**
- **correspondence** violation:
  - can't see change of data in vis
    - **jumbler**
  - salient change in vis not due to significant change in data
    - **misleader**
  - match mathematical structure in data with visual perception
- we can X the data; can we Y the image?
  - are important data changes well-matched with obvious visual changes?

Visual Design Process In Depth: **Dear Data**

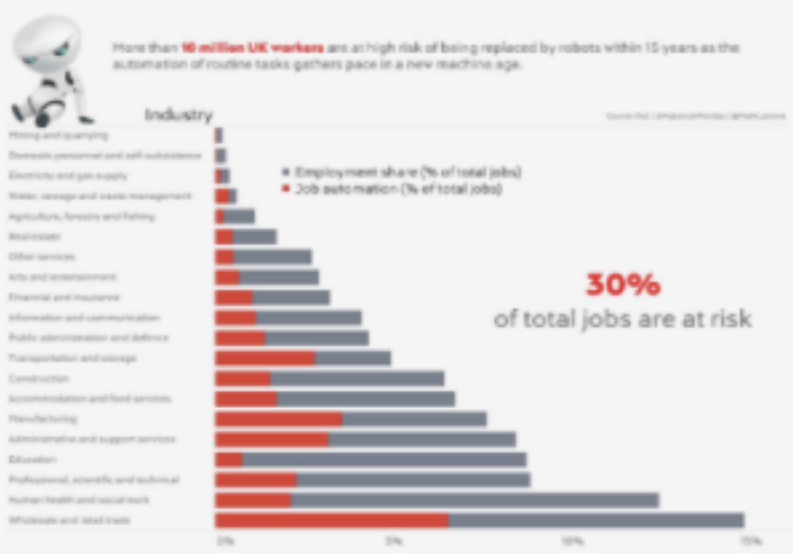
http://www.dear-data.com/by-week/

Visual Design Process In Depth: **Data Sketches**



http://www.datasketch.es/

Redesign En Masse: **Makeover Mondays**



Week 14 – Millions of UK workers at risk of being replaced by robots

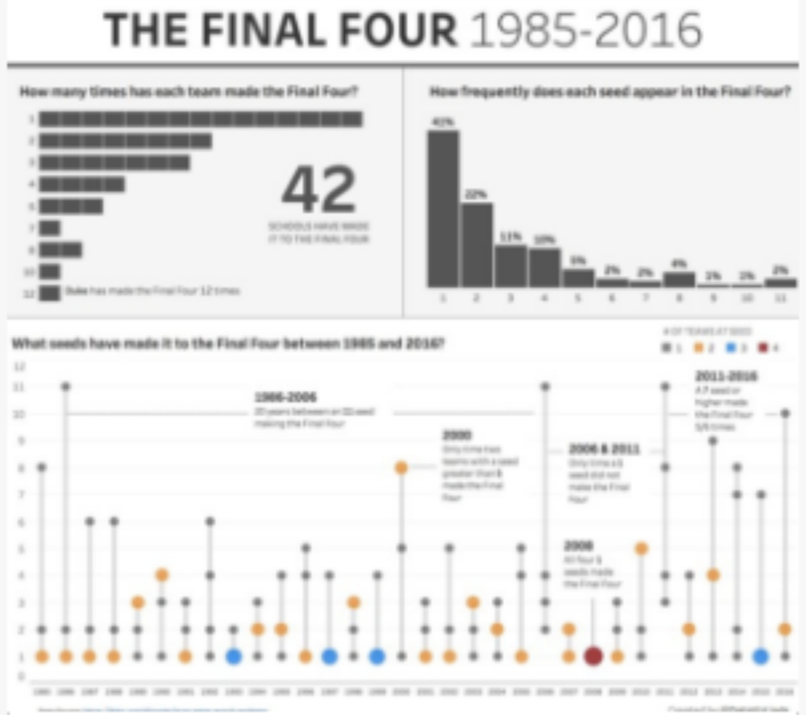
Apr 7, 2017

During week 14 we looked at job automation and the potential impact of robots and AI on the UK employment market.

Week 13 – The Secret of Success

Mar 31, 2017

Week 13 took a look at a Russian survey about the secret of success. Dot plot, bump charts, bar charts, radar charts. This week had it all! Plus seven lessons to take on board.



Week 12 – March Madness

Mar 24, 2017

We looked at March Madness data for week 12, highlighting the phenomenon that is US college basketball. Quite a few vizzes showed the passion that

http://www.makeovermonday.co.uk/blog/

In-Class Exercise

Scenario

- data: room occupancy rates
  - 1 room
  - occupancy measured every 5 min, duration 1 day
- task: characterize space usage pattern

- design
  - propose idioms (visual encoding, interaction)
  - justify idiom choice

Consider

- **what’s the cardinality of the data?**
- **is a single static chart good enough?**
- **should you derive any useful additional data?**

Cardinality

- Marshall: 68 cities * 40 years * 4 crime types = 10,880
- Wine: 130K * 4 = 650,000
  - spatial (hierarchical), quantitative, categorical, free-form text

Scenario

- data: room occupancy rates
  - 20 rooms
  - measured every 5 min, duration 1 day
- task: compare space usage patterns between rooms
- design
  - propose idioms (visual encoding, interaction)
  - justify idiom choice

Consider

- what's the cardinality of the data?
- is a single static chart good enough?
- should you derive any useful additional data?

- **what are trade-offs between**
  - **filtering to see one chart at a time**
  - **showing all side by side with small multiples**
  - **superimposing all on top of each other**

Scenario

- data: room occupancy rates in building
  - 1 building: 200 rooms across 4 floors
  - measured every 5 min, duration 1 day
  - time series + floor plans
- task: characterize space usage patterns
  - trends, outliers
- design
  - propose & justify idioms

Consider

- what's the cardinality of the data?
- is a single static chart good enough?
- should you derive any useful additional data?
- what are trade-offs between
  - filtering to see one chart at a time
  - showing side by side with small multiples
  - superimposing on top of each other

- **multi-scale structure to exploit? aggregate, zoom, slice/ dice, filter?**

Scenario

- data: room occupancy rates in building
  - 1 building: 200 rooms across 4 floors
  - measured every 5 min, duration 1 **year**
  - time series + floor plans **+ room sizes**
- task: characterize space usage patterns
  - trends, outliers
- design
  - propose & justify idioms

Consider

- what's the cardinality of the data?
- is a single static chart good enough?
- should you derive any useful additional data?
- what are trade-offs between
  - filtering to see one chart at a time
  - showing side by side with small multiples
  - superimposing on top of each other
- multi-scale structure to exploit? aggregate, zoom, slice/dice, filter?
- **can you normalize the data? should you - always vs on demand?**
- **how to handle multi-scale space and multi-scale time?**

Scenario

- data: currency exchange rates
  - 30 countries (each against CAD)
  - measured every 5 min, duration 5 years
  - time series + country names + continent names (+ map shapefiles) + country populations
- task: find groups of similarly-performing currencies
- design
  - propose & justify idioms

Consider

- what's the cardinality of the data?
- is a single static chart good enough?
- should you derive any useful additional data?
- what are trade-offs between
  - filtering to see one chart at a time
  - showing side by side with small multiples
  - superimposing on top of each other
- multi-scale structure to exploit? aggregate, zoom, slice/dice, filter?
- can you normalize the data? should you - always vs on demand?
- how to handle multi-scale space and multi-scale time?
- **is spatial information germane or extraneous?**

Scenario

- data: CPU usage across many machines
  - 100 machines, belonging to 20 companies
  - measured every 5 min, duration 1 month
  - time series + company name + company location (country)
- task: capacity planning for machine room
- design
  - propose & justify idioms

Scenario

- data: many metrics across many machines
  - 100 machines, belonging to 20 companies
  - 4 metrics measured every 5 min, duration 1 month
    - CPU, memory, disk I/O, network traffic
  - time series + company name + company sector (finance/tech/entertainment/other)
- task: forensic analysis to determine possible causes of crashes
- design
  - propose & justify idioms

Big Picture & Other Synthesis

Ch 2

What?

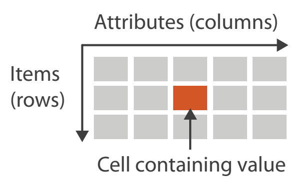
Datasets

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

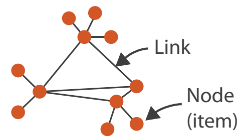
Data and Dataset Types

Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

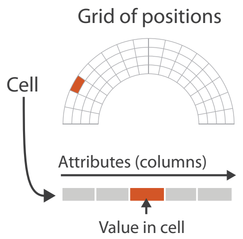
Dataset Types
- Tables



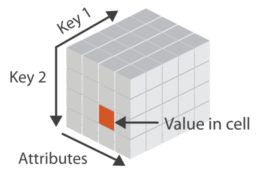
- Networks



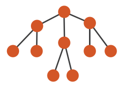
- Fields (Continuous)



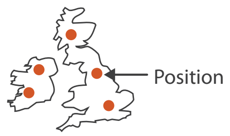
- *Multidimensional Table*



- *Trees*



- Geometry (Spatial)



Dataset Availability
- Static

- Dynamic

Attributes

Attribute Types
- Categorical

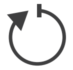
- Ordered
  - *Ordinal*

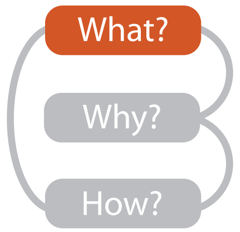
  - *Quantitative*

Ordering Direction
- Sequential

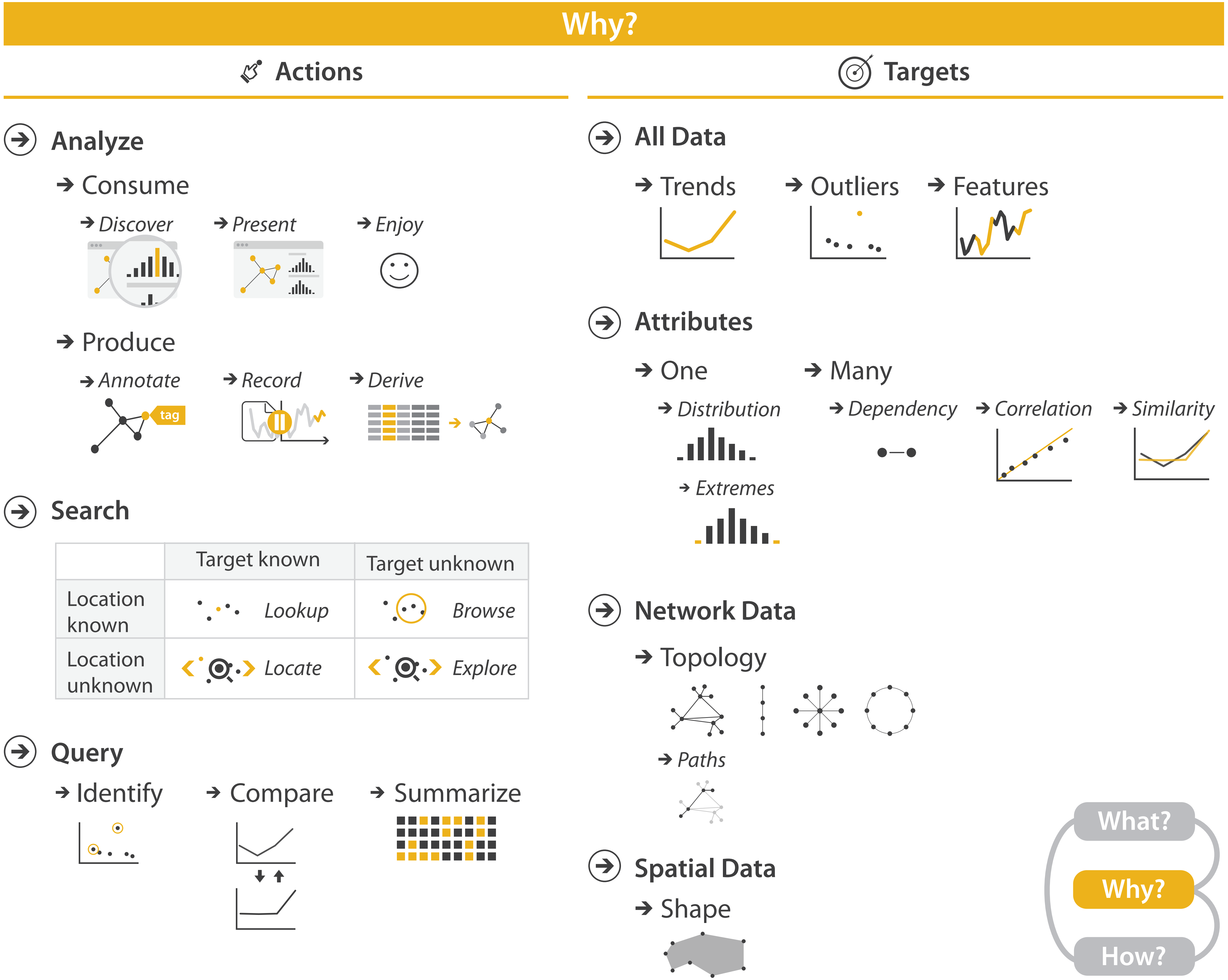
- Diverging

- Cyclic





Ch 3


Why?
Actions
Targets
Analyze
Consume
Discover
Present
Enjoy
Produce
Annotate
tag
Record
Derive
Search
Target known
Target unknown
Location known
Lookup
Browse
Location unknown
Locate
Explore
Query
Identify
Compare
Summarize
All Data
Trends
Outliers
Features
Attributes
One
Many
Distribution
Dependency
Correlation
Similarity
Extremes
Network Data
Topology
Paths
Spatial Data
Shape
What?
Why?
How?

Domain situation
Observe target users using existing tools

Data/task abstraction

Visual encoding/interaction idiom
Justify design with respect to alternatives

Algorithm
Measure system time/memory
Analyze computational complexity

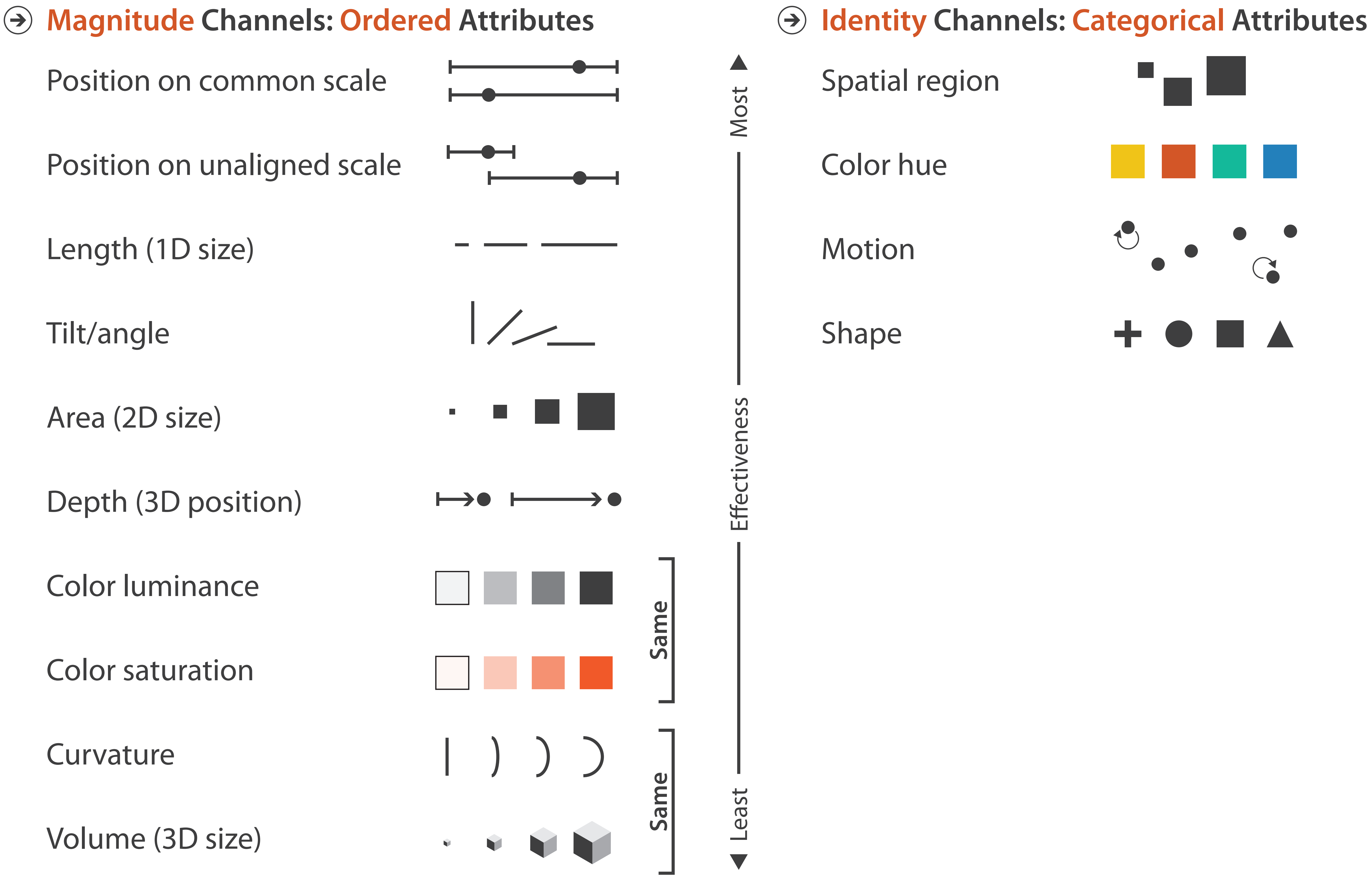
Analyze results qualitatively

Measure human time with lab experiment (*lab study*)

Observe target users after deployment (*field study*)

Measure adoption

Channels: Expressiveness Types and Effectiveness Ranks



Ch 6

- No unjustified 3D
  - Power of the plane
  - Disparity of depth
  - Occlusion hides information
  - Perspective distortion dangers
  - Tilted text isn't legible
- No unjustified 2D
- Eyes beat memory
- Resolution over immersion
- Overview first, zoom and filter, details on demand
- Responsiveness is required
- Function first, form next

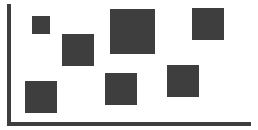
Ch 7

Arrange Tables

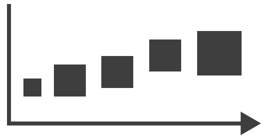
Express Values

Separate, Order, Align Regions

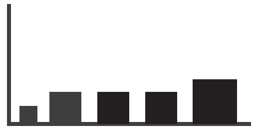
→ Separate



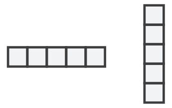
→ Order



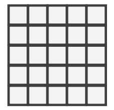
→ Align



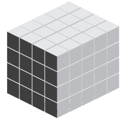
→ *1 Key*
List



→ *2 Keys*
Matrix



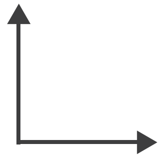
→ *3 Keys*
Volume



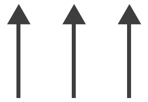
→ *Many Keys*
Recursive Subdivision

Axis Orientation

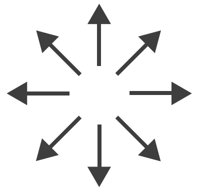
→ Rectilinear



→ Parallel

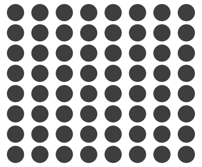


→ Radial



Layout Density

→ Dense



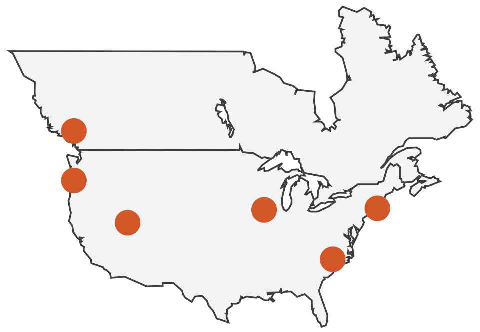
→ Space-Filling

Ch 8

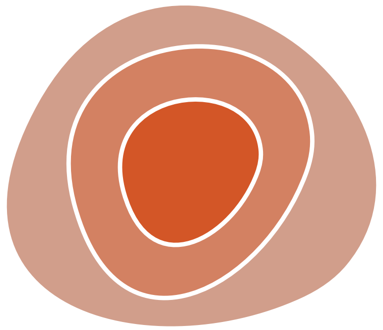
Arrange Spatial Data

➔ Use Given

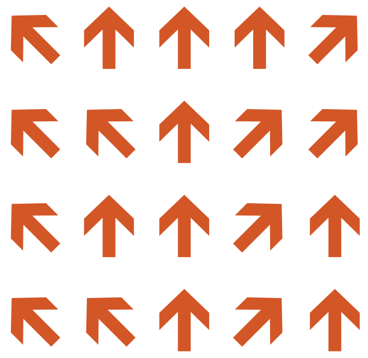
- ➔ Geometry
  - ➔ *Geographic*
  - ➔ *Other Derived*



- ➔ Spatial Fields
  - ➔ *Scalar Fields (one value per cell)*
    - ➔ *Isocontours*
    - ➔ *Direct Volume Rendering*



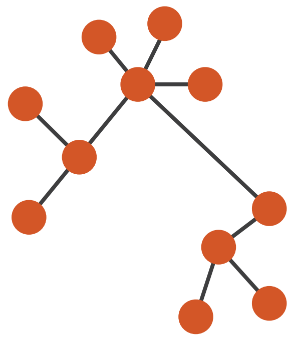
  - ➔ *Vector and Tensor Fields (many values per cell)*
    - ➔ *Flow Glyphs (local)*
    - ➔ *Geometric (sparse seeds)*
    - ➔ *Textures (dense seeds)*
    - ➔ *Features (globally derived)*



Arrange Networks and Trees

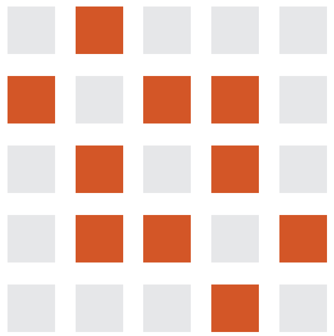
Node–Link Diagrams
Connection Marks

✓ NETWORKS ✓ TREES



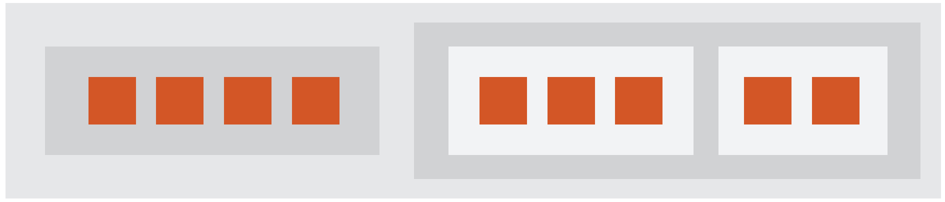
Adjacency Matrix
Derived Table

✓ NETWORKS ✓ TREES



Enclosure
Containment Marks

✗ NETWORKS ✓ TREES



Encode › Map

Color

Color Encoding

- Hue
- Saturation
- Luminance

Color Map

- Categorical
- Ordered
  - *Sequential*
  - *Diverging*
- Bivariate

Size, Angle, Curvature, ...

- Length
- Angle
- Area
- Curvature
- Volume

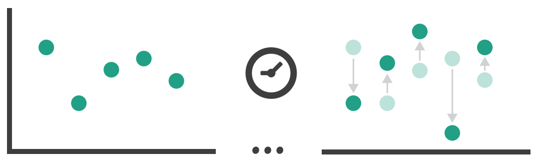
Shape

Motion

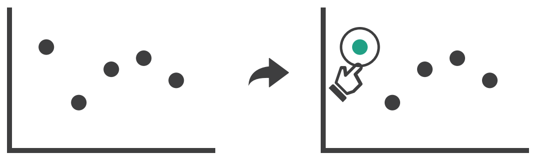
- Motion
  Direction, Rate, Frequency, ...

Manipulate

Change over Time



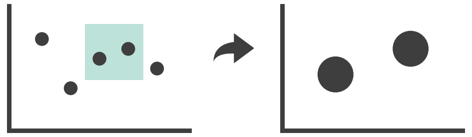
Select



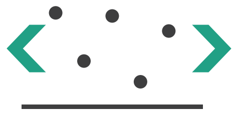
Navigate

Item Reduction

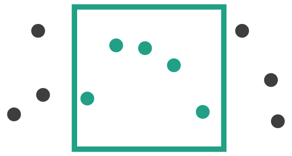
- *Zoom*
  Geometric or *Semantic*



- *Pan/Translate*



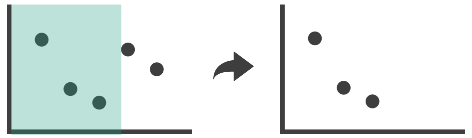
- *Constrained*



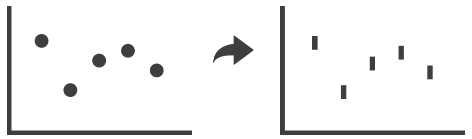
Attribute Reduction

- *Slice*

- *Cut*



- *Project*



Ch 12

Facet

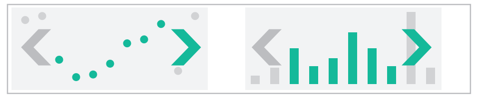
Juxtapose and Coordinate Multiple Side-by-Side Views

→ Share Encoding: Same/Different

→ *Linked Highlighting*

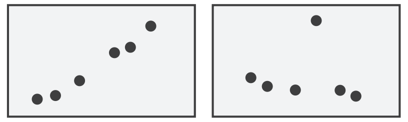
→ Share Data: All/Subset/None

→ Share Navigation

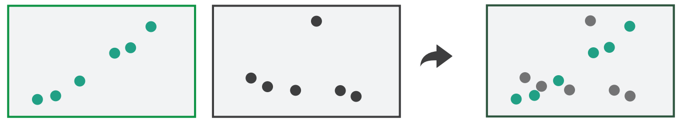


		Data		
		All	Subset	None
Encoding	Same	Redundant	Overview/ Detail	Small Multiples
	Different	Multiform	Multiform, Overview/ Detail	No Linkage

Partition into Side-by-Side Views



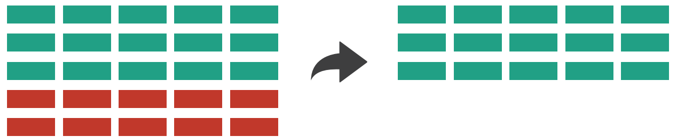
Superimpose Layers



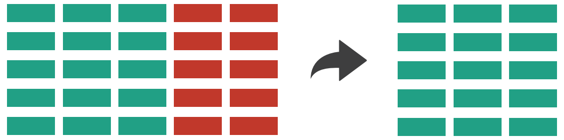
Reducing Items and Attributes

Filter

→ Items

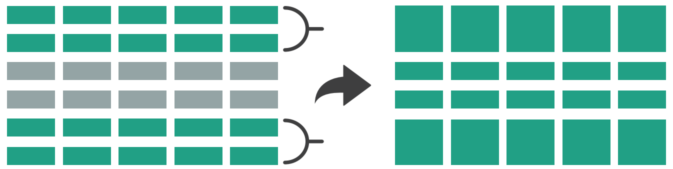


→ Attributes

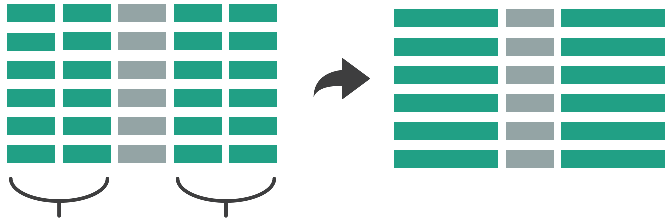


Aggregate

→ Items

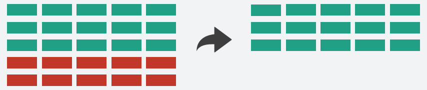


→ Attributes

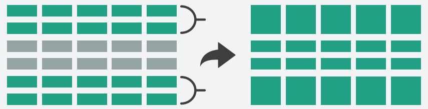


Reduce

Filter



Aggregate



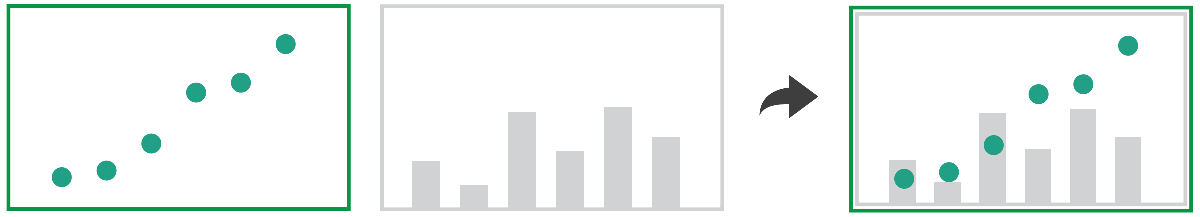
Embed

Ch 14

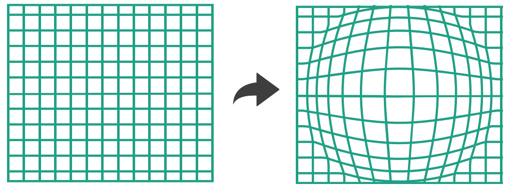
Embed

Elide Data

Superimpose Layer

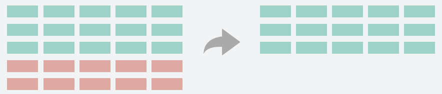


Distort Geometry

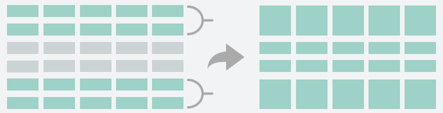


Reduce

Filter



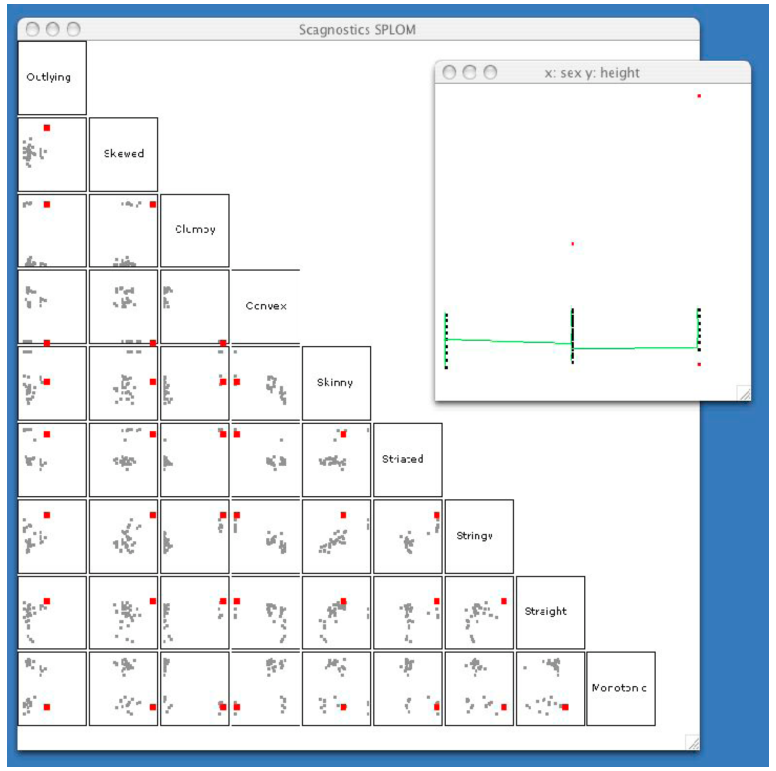
Aggregate



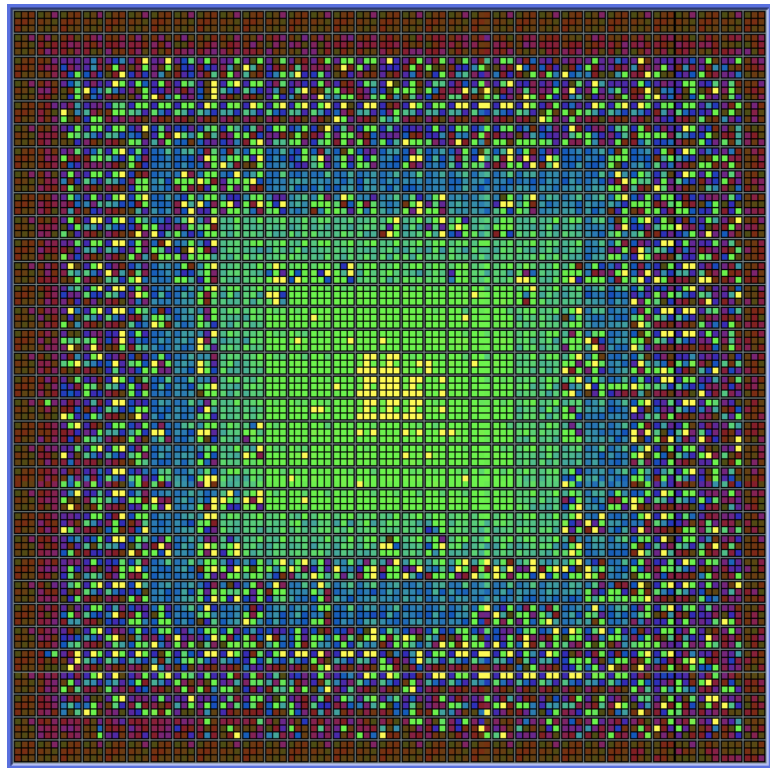
Embed

Ch 15

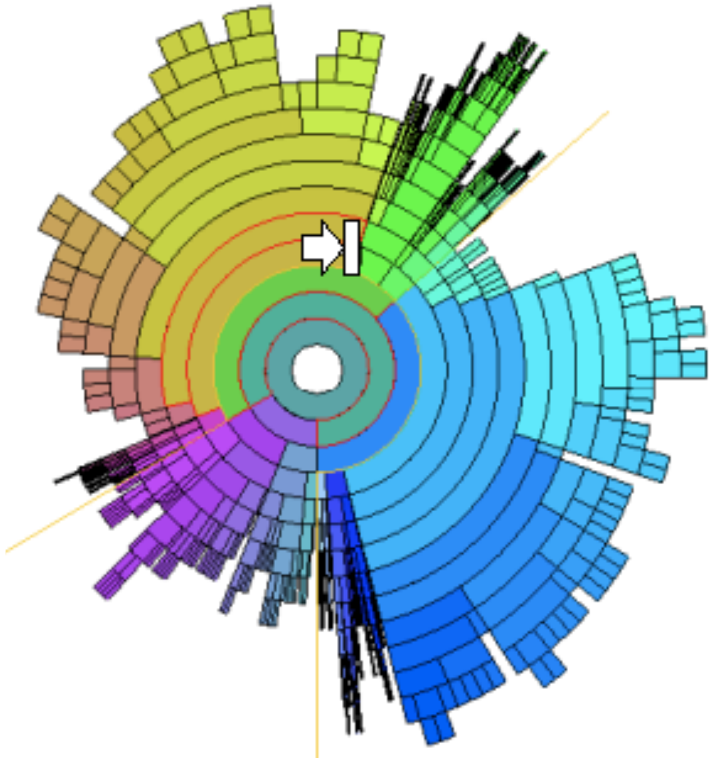
Scagnostics



VisDB

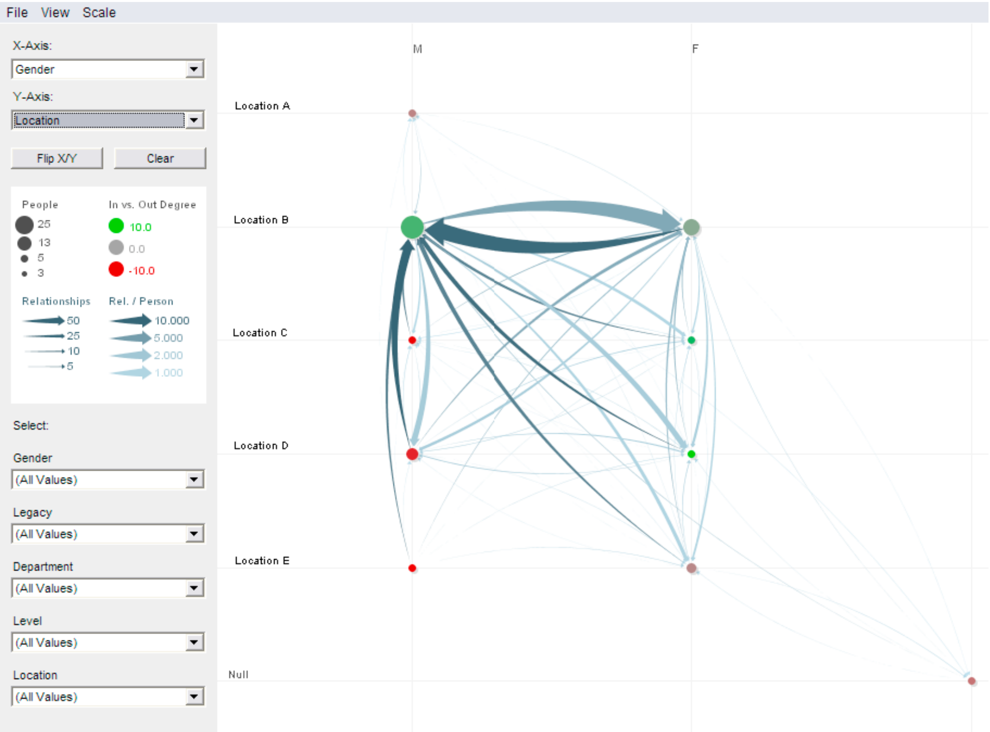


InterRing

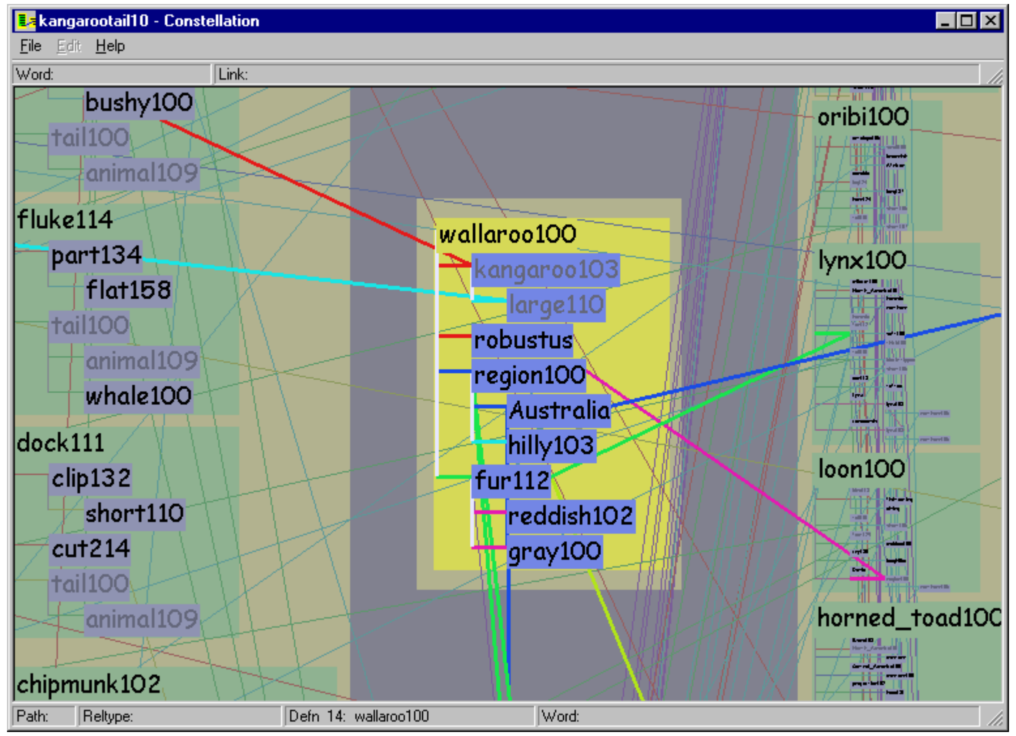


HCE

PivotGraph

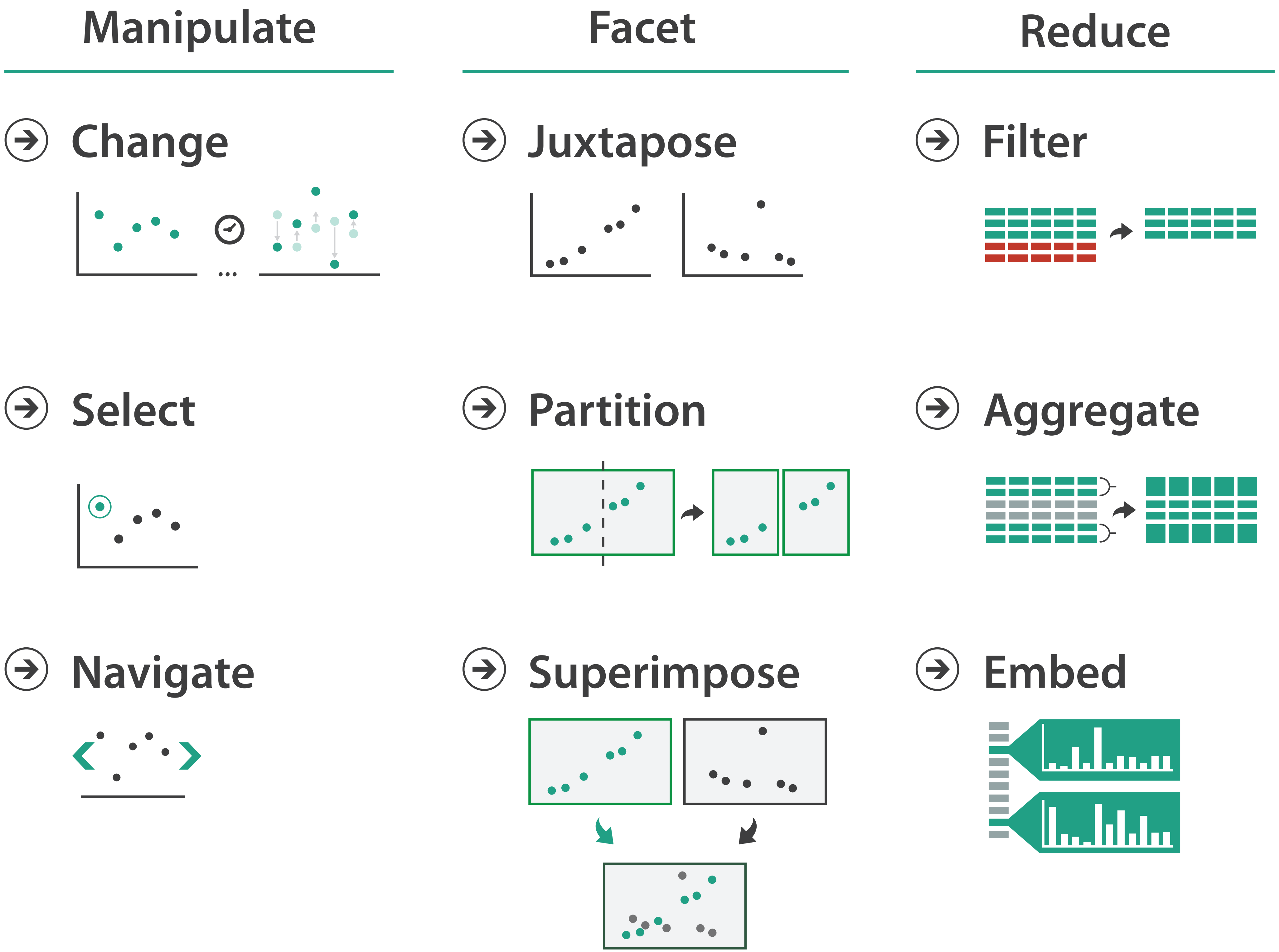


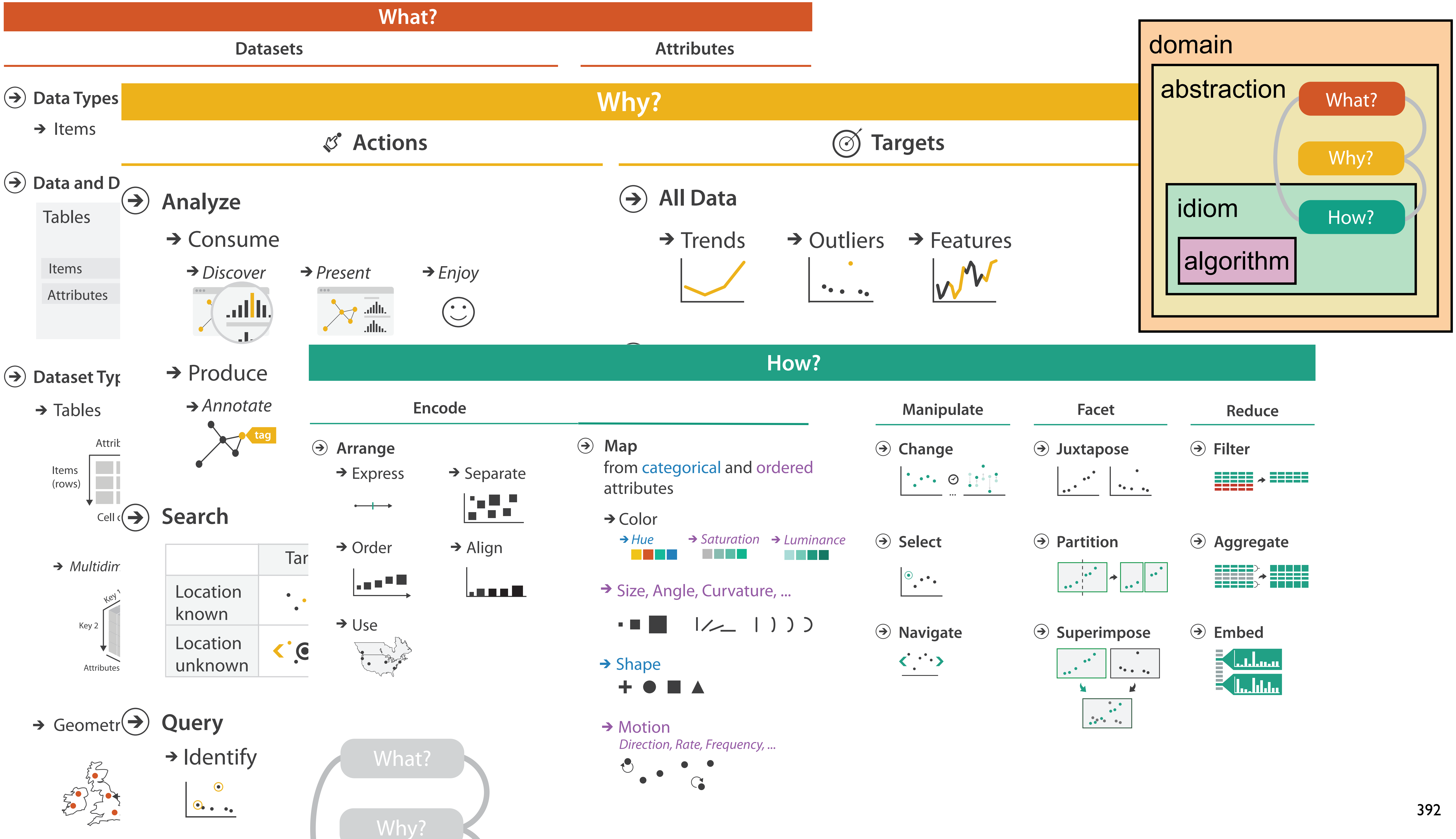
Constellation



How to handle complexity: 4 families of strategies

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




What?
Datasets
Attributes
Data Types
Items
Data and D
Tables
Items
Attributes
Dataset Typ
Tables
Attrib
Items (rows)
Cell c
Multidim
Key 2
Attributes
Geometr
Why?
Actions
Targets
Analyze
Consume
Discover
Present
Enjoy
Produce
Annotate
tag
Search
Tar
Location known
Location unknown
Query
Identify
All Data
Trends
Outliers
Features
domain
abstraction
What?
Why?
idiom
How?
algorithm
How?
Encode
Manipulate
Facet
Reduce
Arrange
Express
Separate
Order
Align
Use
Map
from categorical and ordered attributes
Color
Hue
Saturation
Luminance
Size, Angle, Curvature, ...
Shape
Motion
Direction, Rate, Frequency, ...
Change
Select
Navigate
Juxtapose
Partition
Superimpose
Filter
Aggregate
Embed
What?
Why?

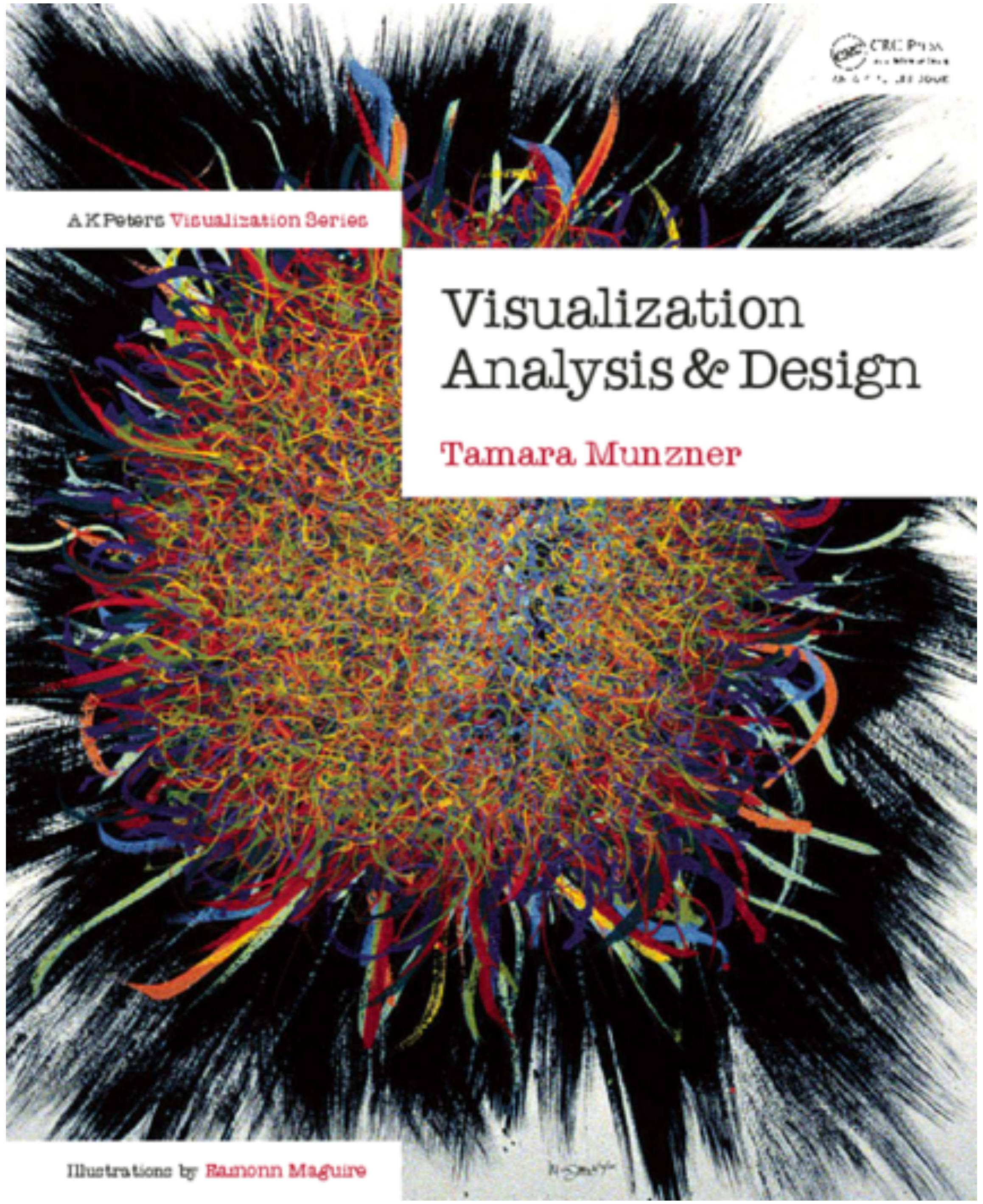
More Information

@tamaramunzner

- these slides
  http://www.cs.ubc.ca/~tmm/talks.html#vadallslides

- book page (including other 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.