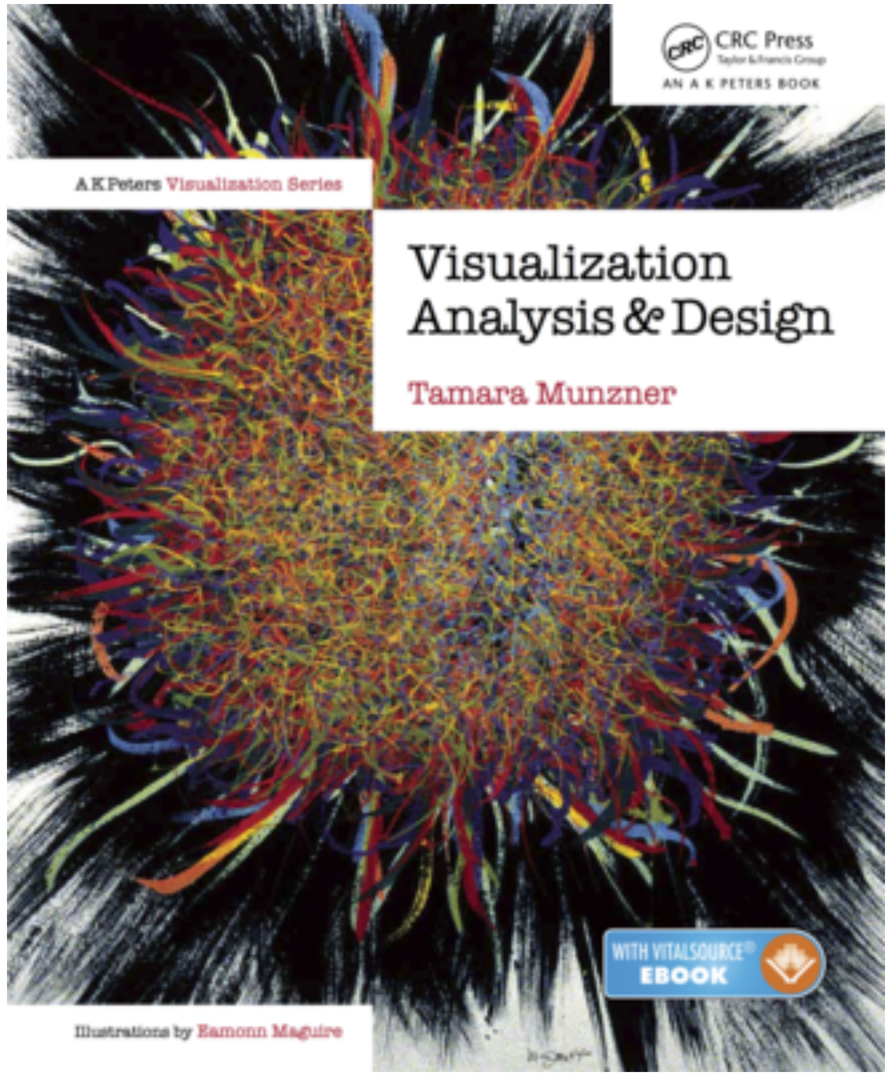
Visualization Analysis & Design

Tamara Munzner
Department of Computer Science
University of British Columbia



Data Visualization Masterclass: Principles, Tools, and Storytelling
June 13 2017, VIZBI/VIVID, Sydney Australia

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

@tamaramunzner

Outline

- **Session 1: Principles** *9:15-10:30am*
  - Analysis: What, Why, How
  - Marks and Channels, Perception
  - Color
- **Session 2: Techniques for Scaling** *10:50-11:40am*
  - Manipulate: Change, Select, Navigate
  - Facet: Juxtapose, Partition, Superimpose
  - Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

Defining visualization (vis)

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

Why?...

Why have a human in the loop?

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

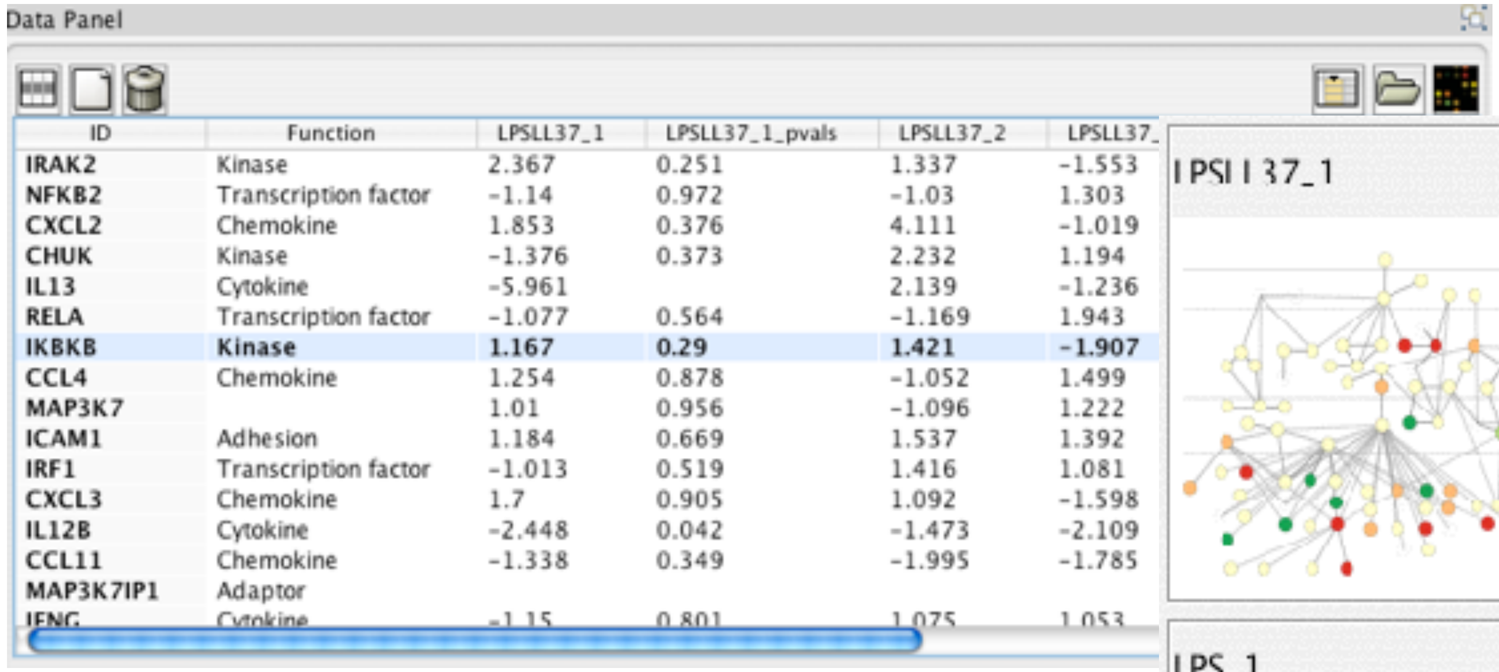
Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- don't need vis when fully automatic solution exists and is trusted
- many analysis problems ill-specified
  - don't know exactly what questions to ask in advance
- possibilities
  - long-term use for end users (e.g. exploratory analysis of scientific data)
  - presentation of known results
  - stepping stone to better understanding of requirements before developing models
  - help developers of automatic solution refine/debug, determine parameters
  - help end users of automatic solutions verify, build trust

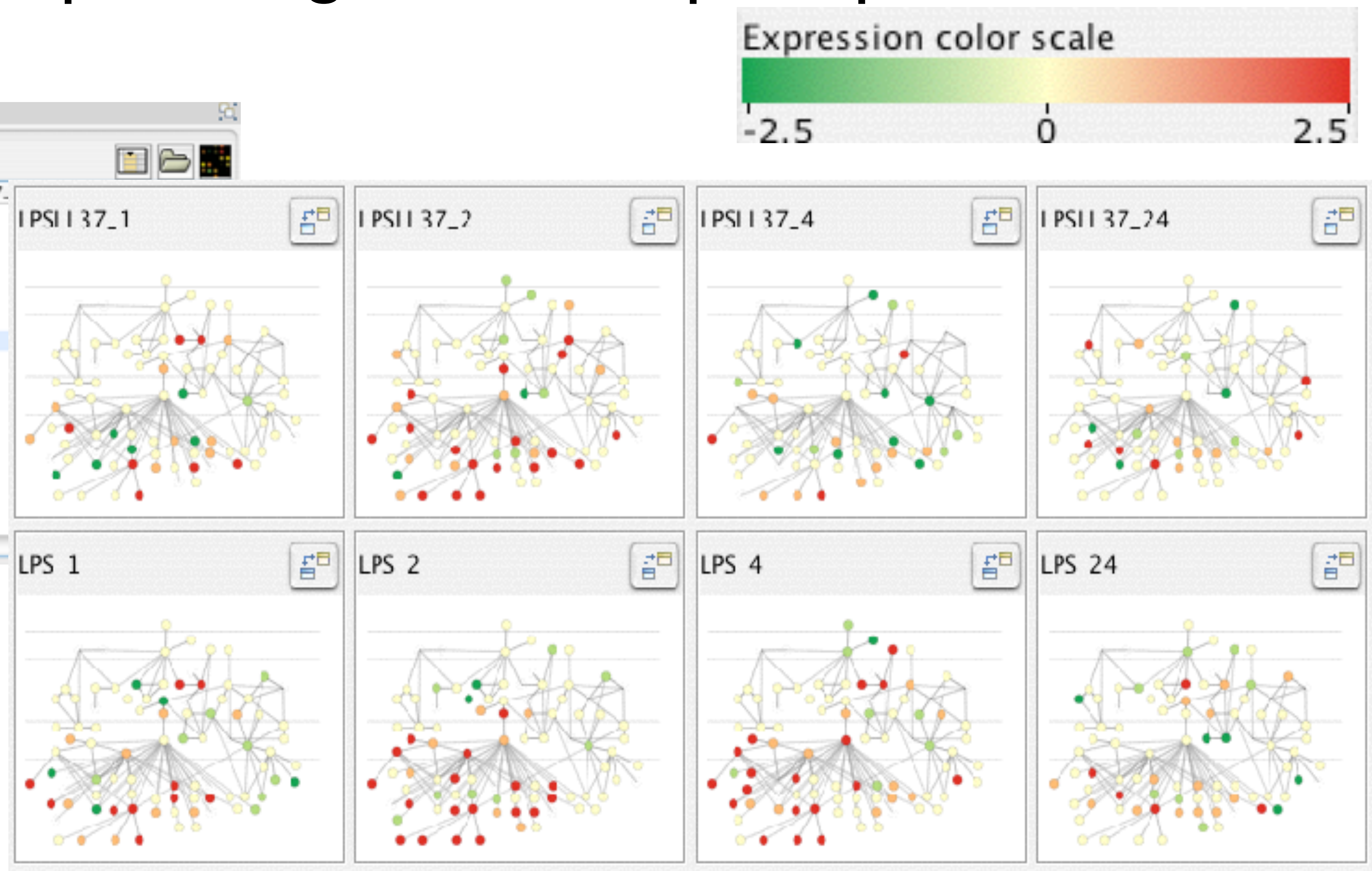
Why use an external representation?

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

- external representation: replace cognition with perception


Data Panel

ID	Function	LPSLL37_1	LPSLL37_1_pvals	LPSLL37_2	LPSLL37_
IRAK2	Kinase	2.367	0.251	1.337	-1.553
NFKB2	Transcription factor	-1.14	0.972	-1.03	1.303
CXCL2	Chemokine	1.853	0.376	4.111	-1.019
CHUK	Kinase	-1.376	0.373	2.232	1.194
IL13	Cytokine	-5.961		2.139	-1.236
RELA	Transcription factor	-1.077	0.564	-1.169	1.943
IKBKB	Kinase	1.167	0.29	1.421	-1.907
CCL4	Chemokine	1.254	0.878	-1.052	1.499
MAP3K7		1.01	0.956	-1.096	1.222
ICAM1	Adhesion	1.184	0.669	1.537	1.392
IRF1	Transcription factor	-1.013	0.519	1.416	1.081
CXCL3	Chemokine	1.7	0.905	1.092	-1.598
IL12B	Cytokine	-2.448	0.042	-1.473	-2.109
CCL11	Chemokine	-1.338	0.349	-1.995	-1.785
MAP3K7IP1	Adaptor				
IFNG	Cytokine	-1.15	0.801	1.075	1.053



[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 represent all the data?

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

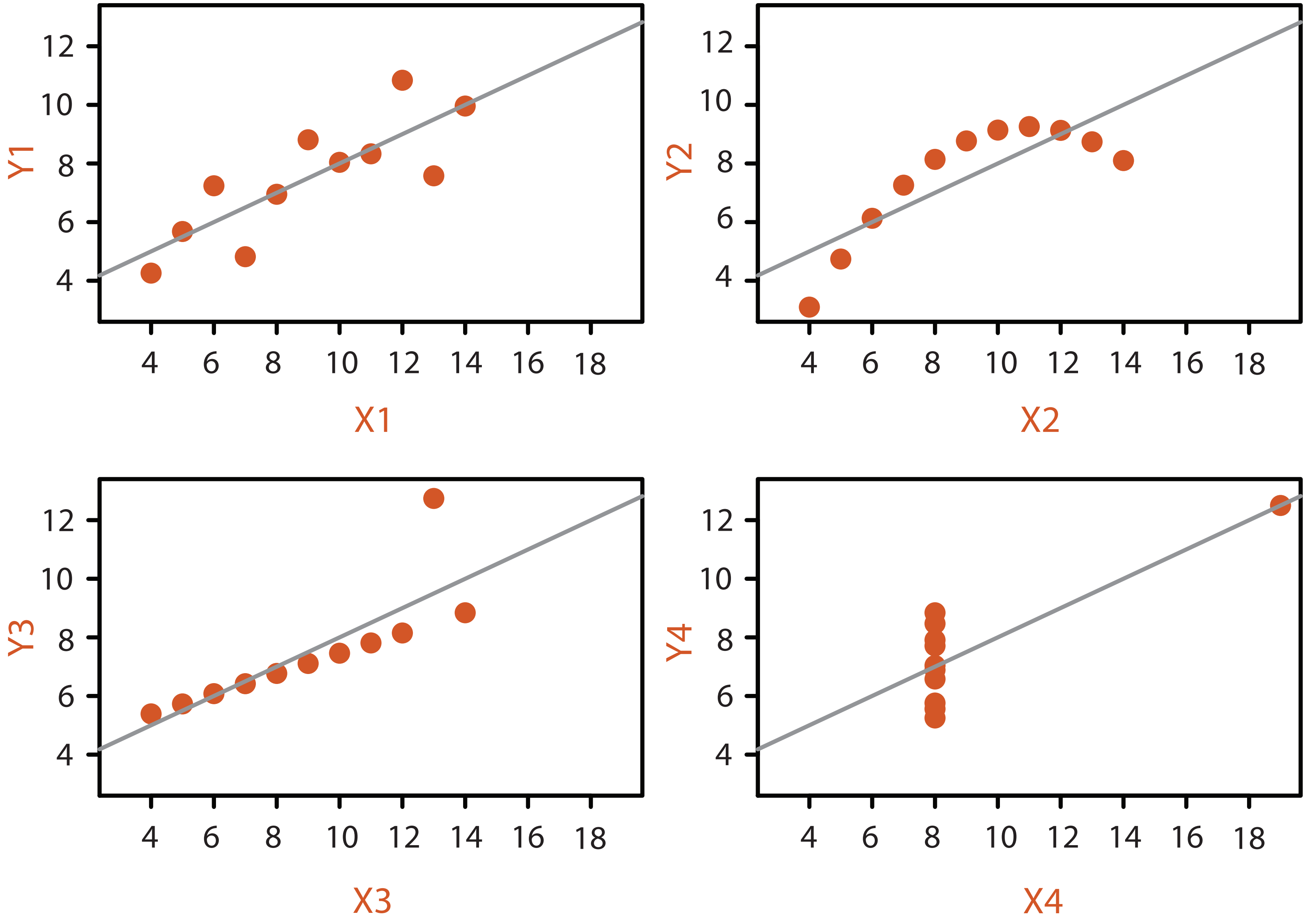
- summaries lose information, details matter
  - confirm expected and find unexpected patterns
  - assess validity of statistical model

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



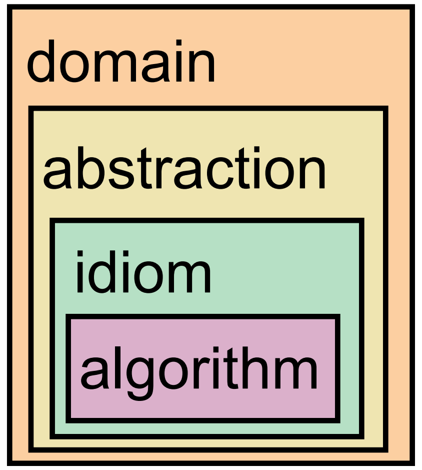
Why are there resource limitations?

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

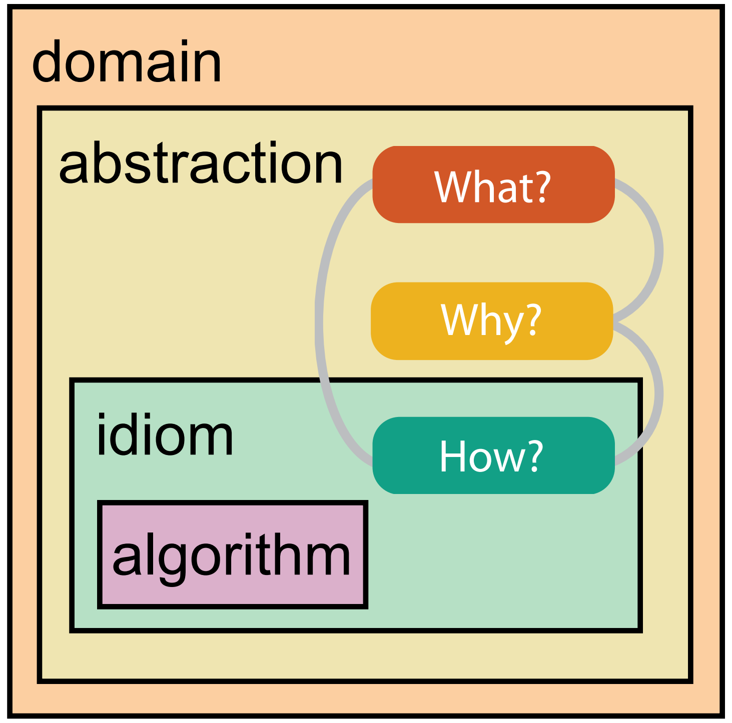
- computational limits
  - processing time
  - system memory
- human limits
  - human attention and memory
- display limits
  - pixels are precious resource, the most constrained resource
  - **information density**: ratio of space used to encode info vs unused whitespace
    - tradeoff between clutter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

Analysis framework: Four levels, three questions

- *domain* situation
  – who are the target users?
- *abstraction*
  – translate from specifics of domain to vocabulary of vis
- **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 Nested Model of Visualization Design and Validation.
Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]



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

Validation methods from different fields for each level

anthropology/
ethnography

Domain situation
Observe target users using existing tools

Data/task abstraction

Visual encoding/interaction idiom
Justify design with respect to alternatives

design

Algorithm
Measure system time/memory
Analyze computational complexity

computer
science

Analyze results qualitatively

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

cognitive
psychology

Observe target users after deployment (*field study*)

anthropology/
ethnography

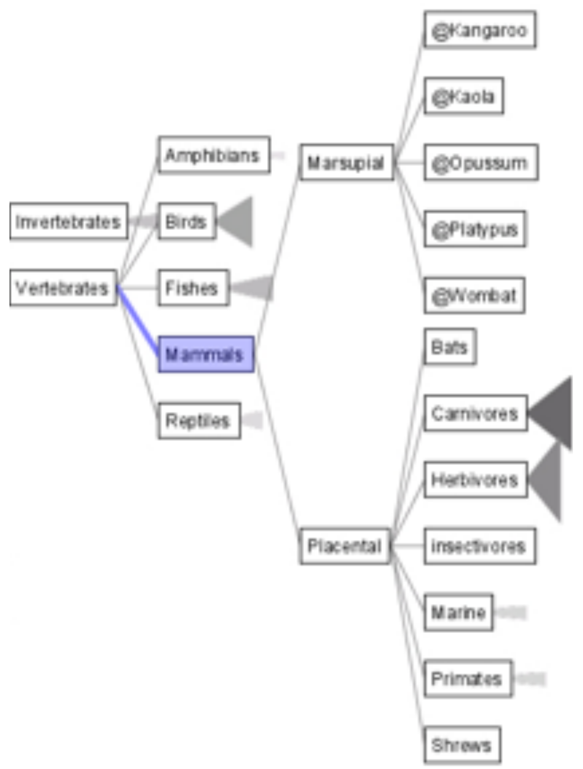
Measure adoption

- mismatch: cannot show idiom good with system timings
- mismatch: cannot show abstraction good with lab study

Why analyze?

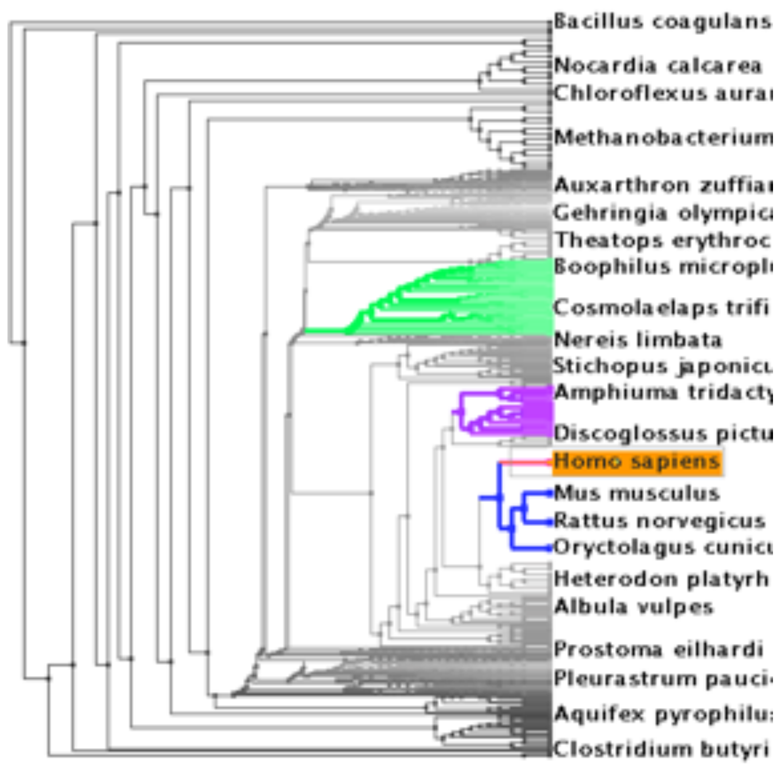
- imposes a structure on huge design space
  - scaffold to help you think systematically about choices
  - analyzing existing as stepping stone to designing new

SpaceTree



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

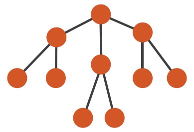
TreeJuxtaposer



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

What?

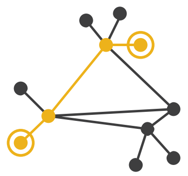
- Tree



Why?

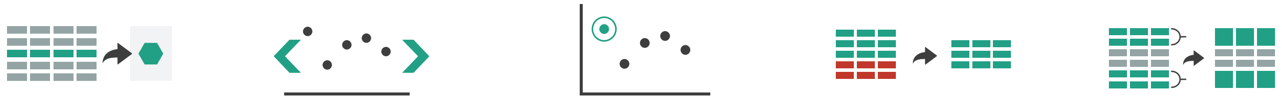
- Actions
  - → Present → Locate → Identify

- Targets
  - → Path between two nodes

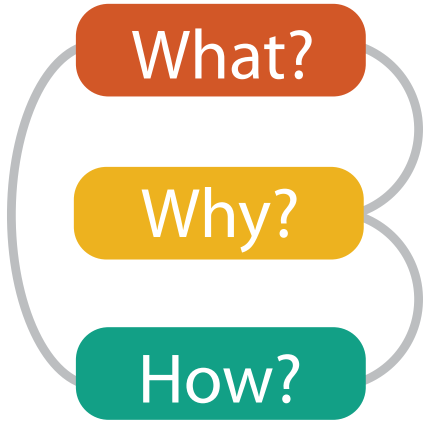


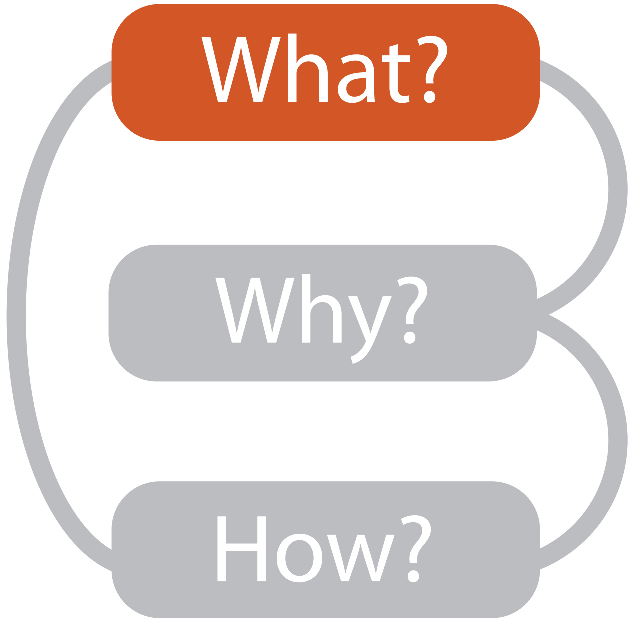
How?

- SpaceTree
  - → Encode → Navigate → Select → Filter → Aggregate



- TreeJuxtaposer
  - → Encode → Navigate → Select → Arrange





What?

Datasets

Data Types

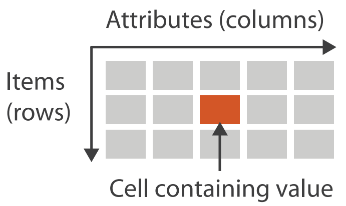
- Items
- Attributes
- Links
- Positions
- Grids

Data and Dataset Types

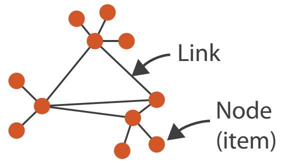
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

Dataset Types

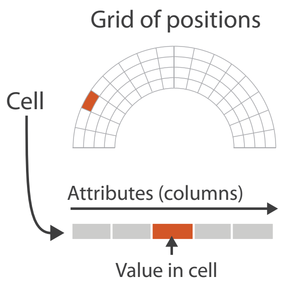
- Tables



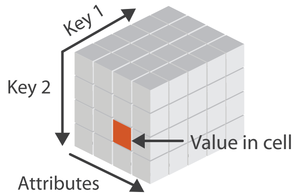
- Networks



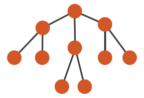
- Fields (Continuous)



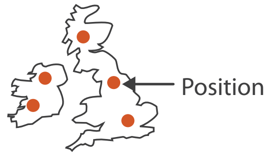
- *Multidimensional Table*



- *Trees*



- Geometry (Spatial)



Dataset Availability

- Static

- Dynamic

Attributes

Attribute Types

- Categorical

- Ordered
  - *Ordinal*

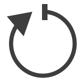
  - *Quantitative*

Ordering Direction

- Sequential

- Diverging

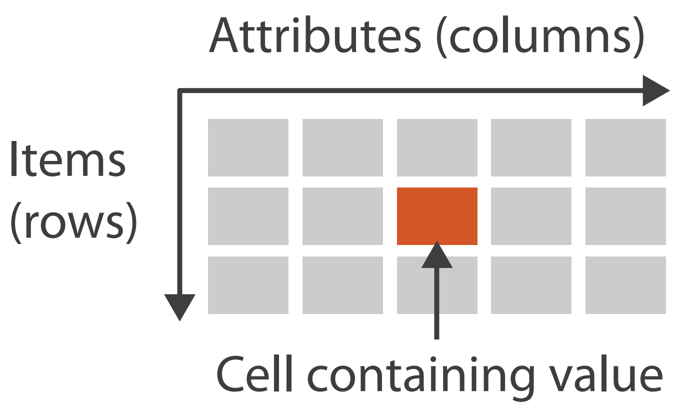
- Cyclic



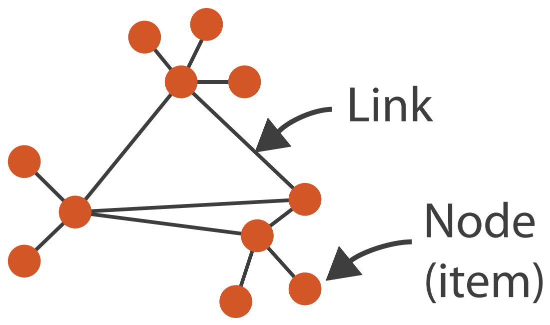
Dataset and data types

➔ Dataset Types

➔ Tables

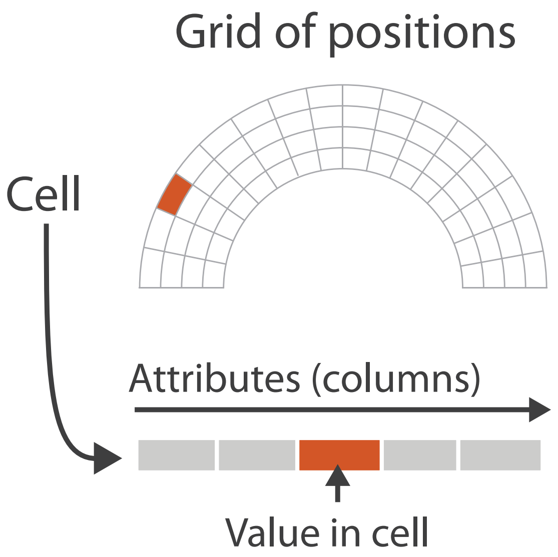


➔ Networks

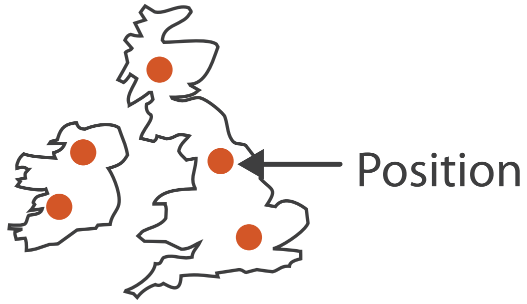


Spatial

➔ Fields (Continuous)



➔ Geometry (Spatial)

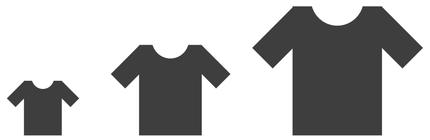


➔ Attribute Types

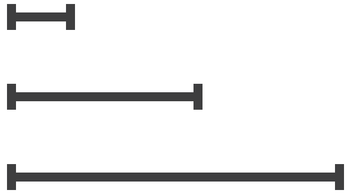
➔ Categorical

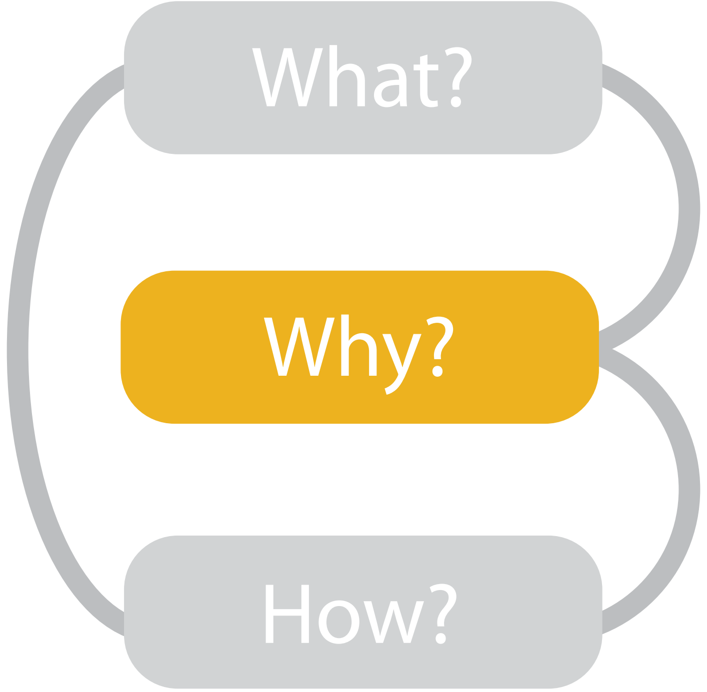
➔ Ordered

➔ *Ordinal*

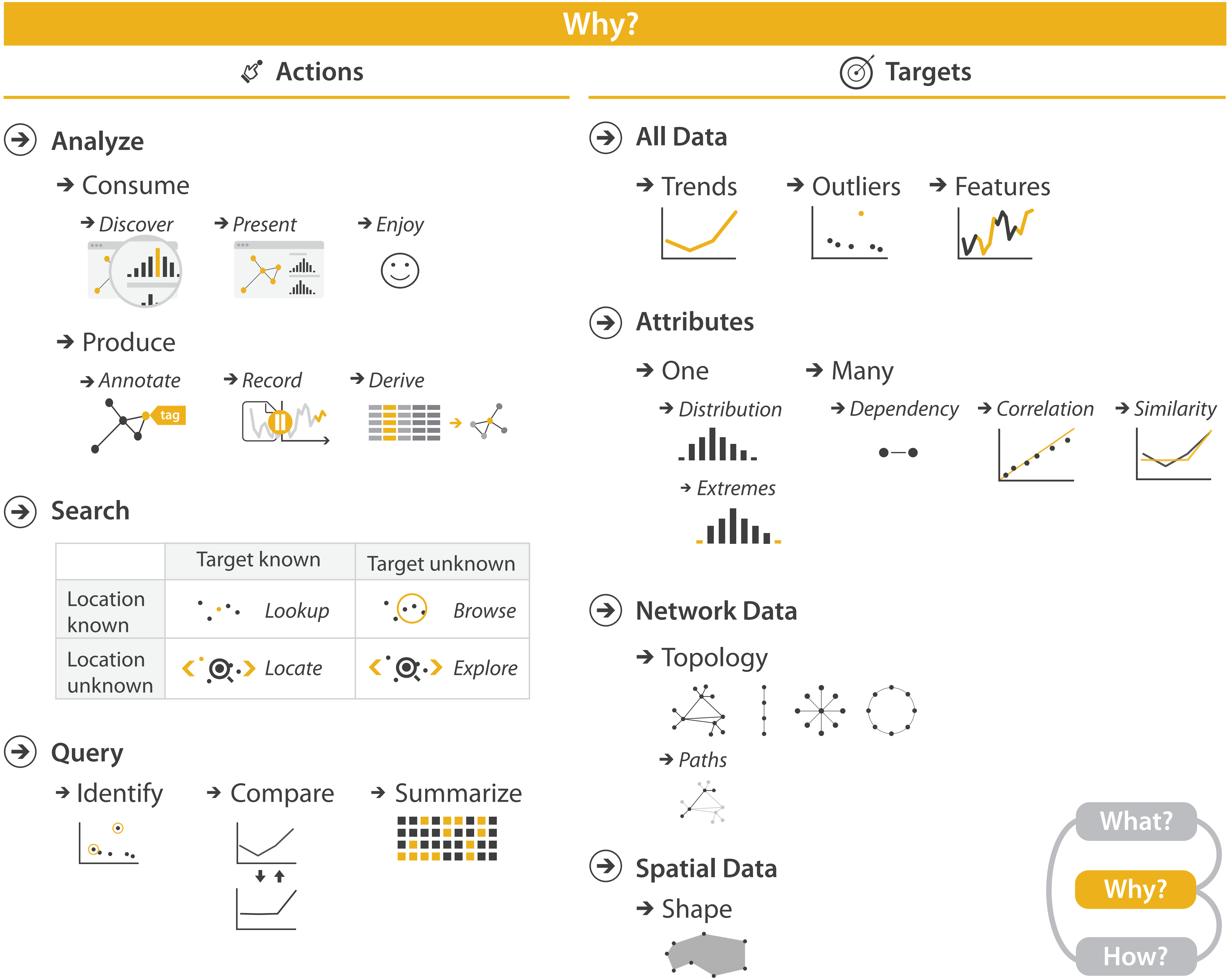


➔ *Quantitative*



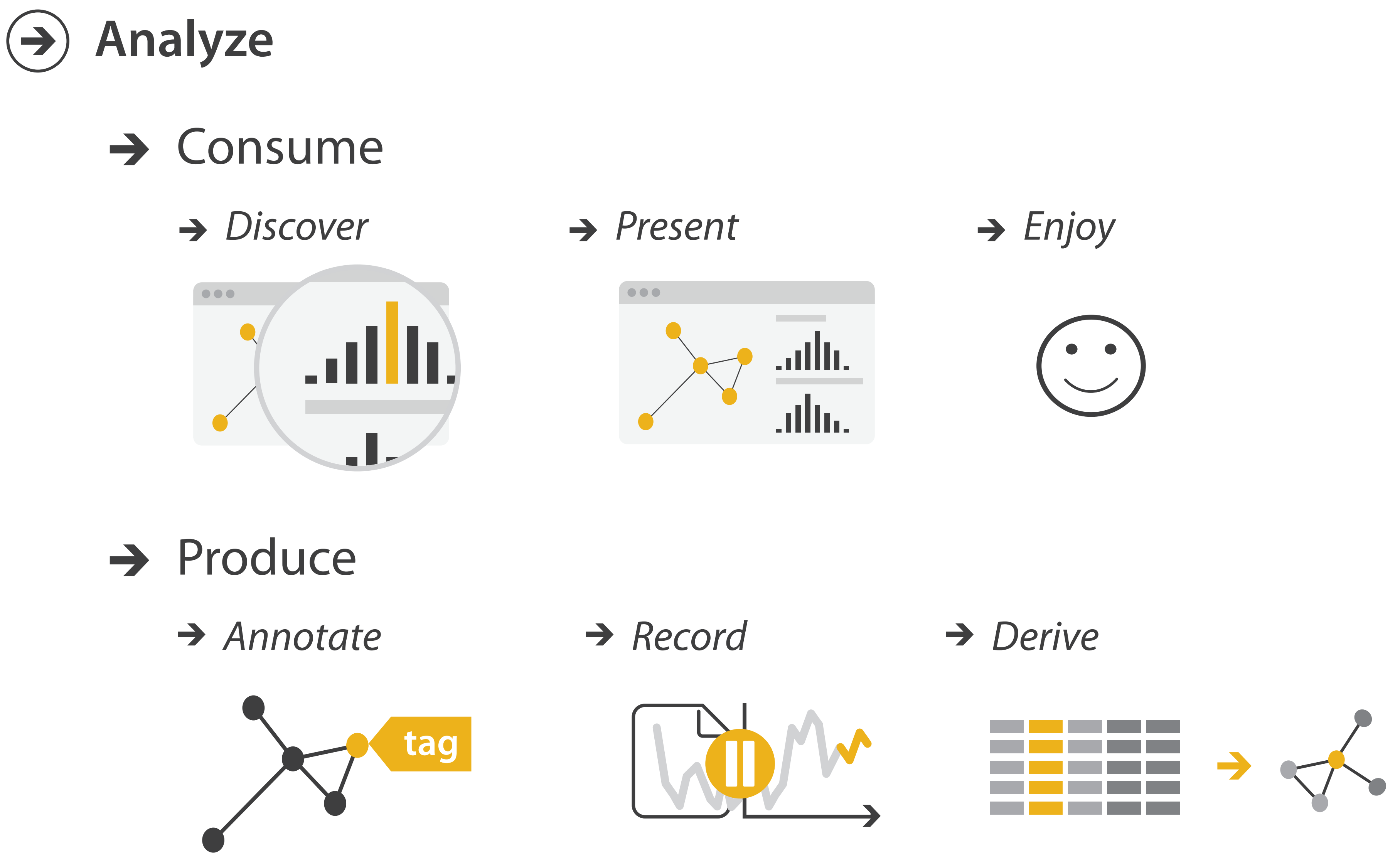


- {action, target} pairs
  - *discover distribution*
  - *compare trends*
  - *locate outliers*
  - *browse topology*



Actions 1: Analyze

- consume
  - discover vs present
    - classic split
    - aka explore vs explain
  - enjoy
    - newcomer
    - aka casual, social
- produce
  - annotate, record
  - derive
    - crucial design choice



Actions II: Search

- what does user know? ➔ **Search**
  - –target, location

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

Actions III: Query

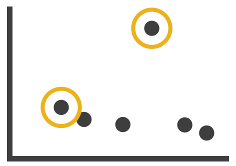
- what does user know?
  - –target, location
- how much of the data matters?
  - –one, some, all

➔ **Search**

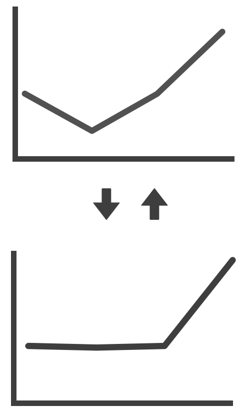
	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

➔ **Query**

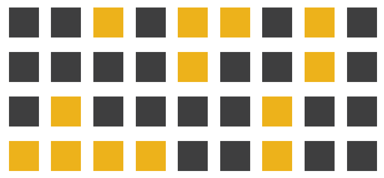
→ Identify



→ Compare



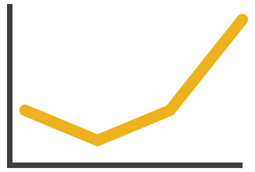
→ Summarize



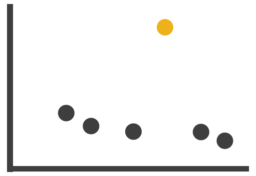
Targets

➔ All Data

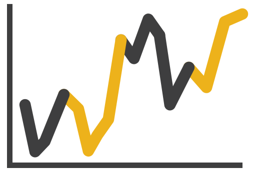
→ Trends



→ Outliers



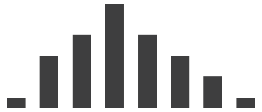
→ Features



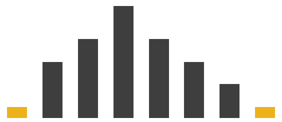
➔ Attributes

→ One

→ *Distribution*



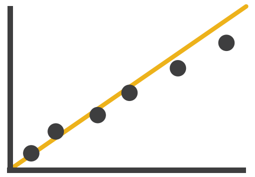
→ *Extremes*



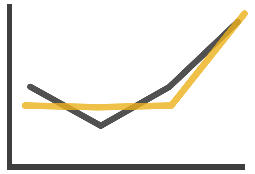
→ Many

→ *Dependency*

→ *Correlation*

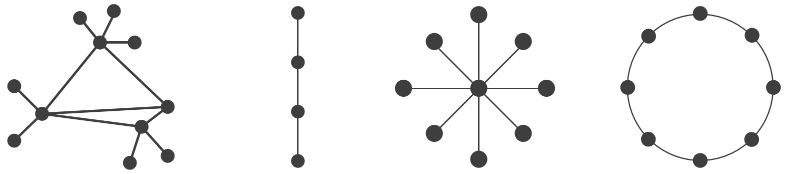


→ *Similarity*

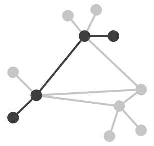


➔ Network Data

→ Topology

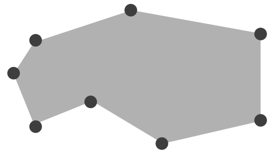


→ *Paths*

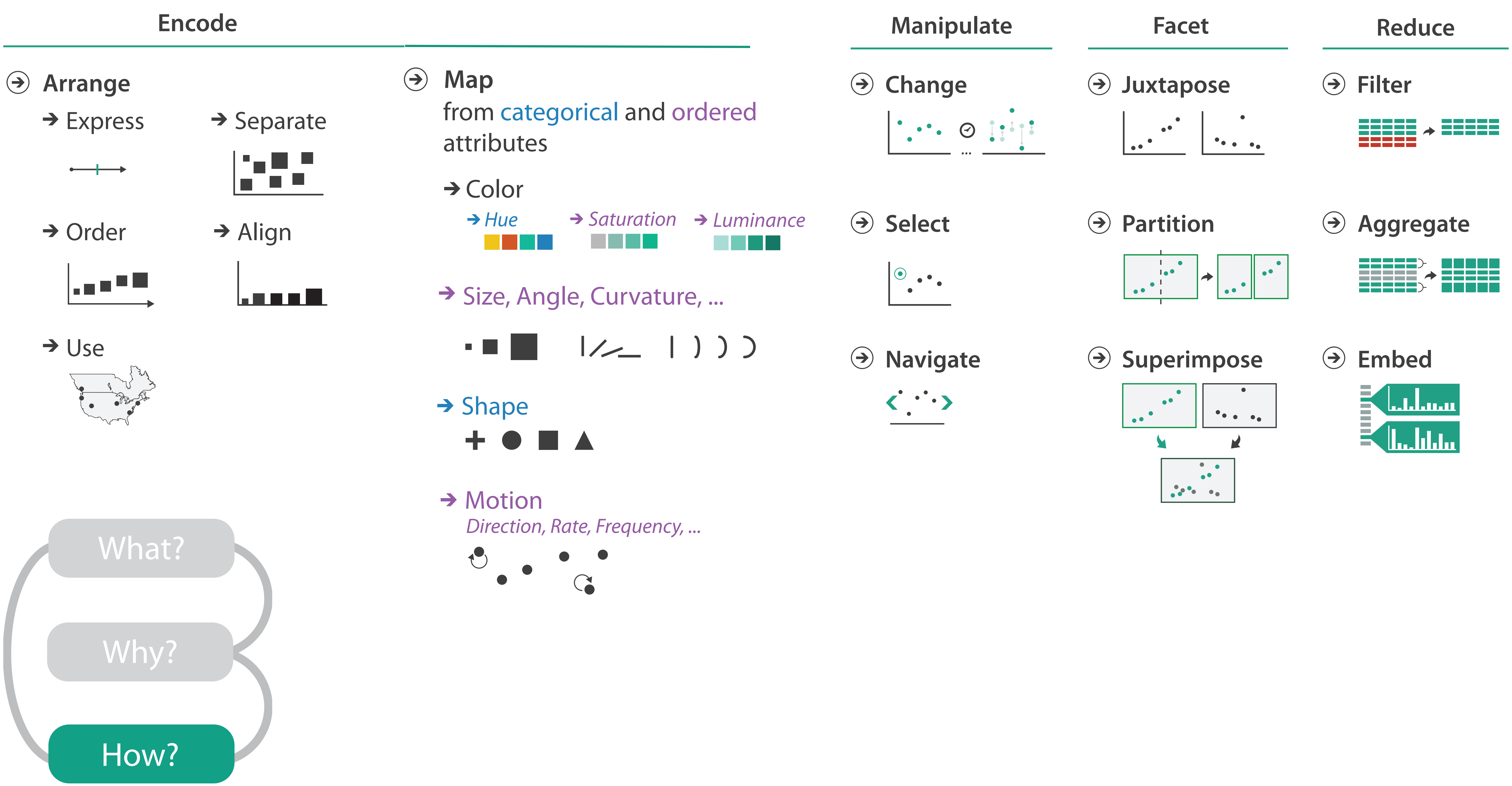


➔ Spatial Data

→ Shape



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?

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, Nov 2014.
    - *Chap 1: What's Vis, and Why Do It?*
  - *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.

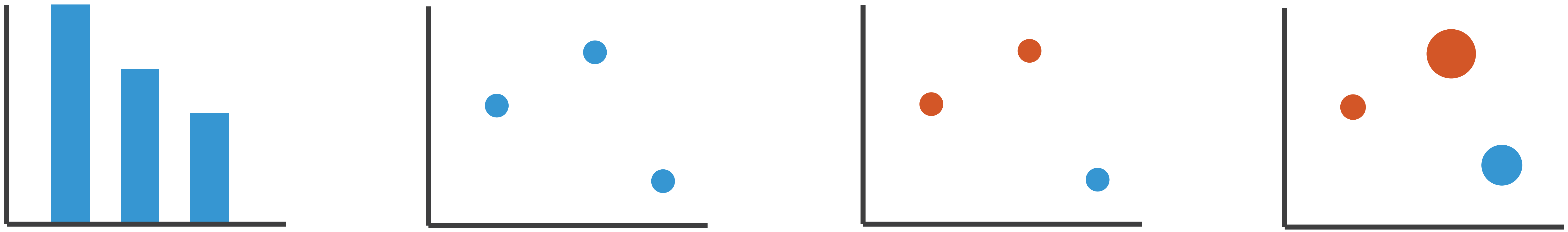
Outline

- **Session 1: Principles** *9:15-10:30am*
  - Analysis: What, Why, How
  - Marks and Channels, Perception
  - Color
- **Session 2: Techniques for Scaling** *10:50-11:40am*
  - Manipulate: Change, Select, Navigate
  - Facet: Juxtapose, Partition, Superimpose
  - Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

Encoding visually

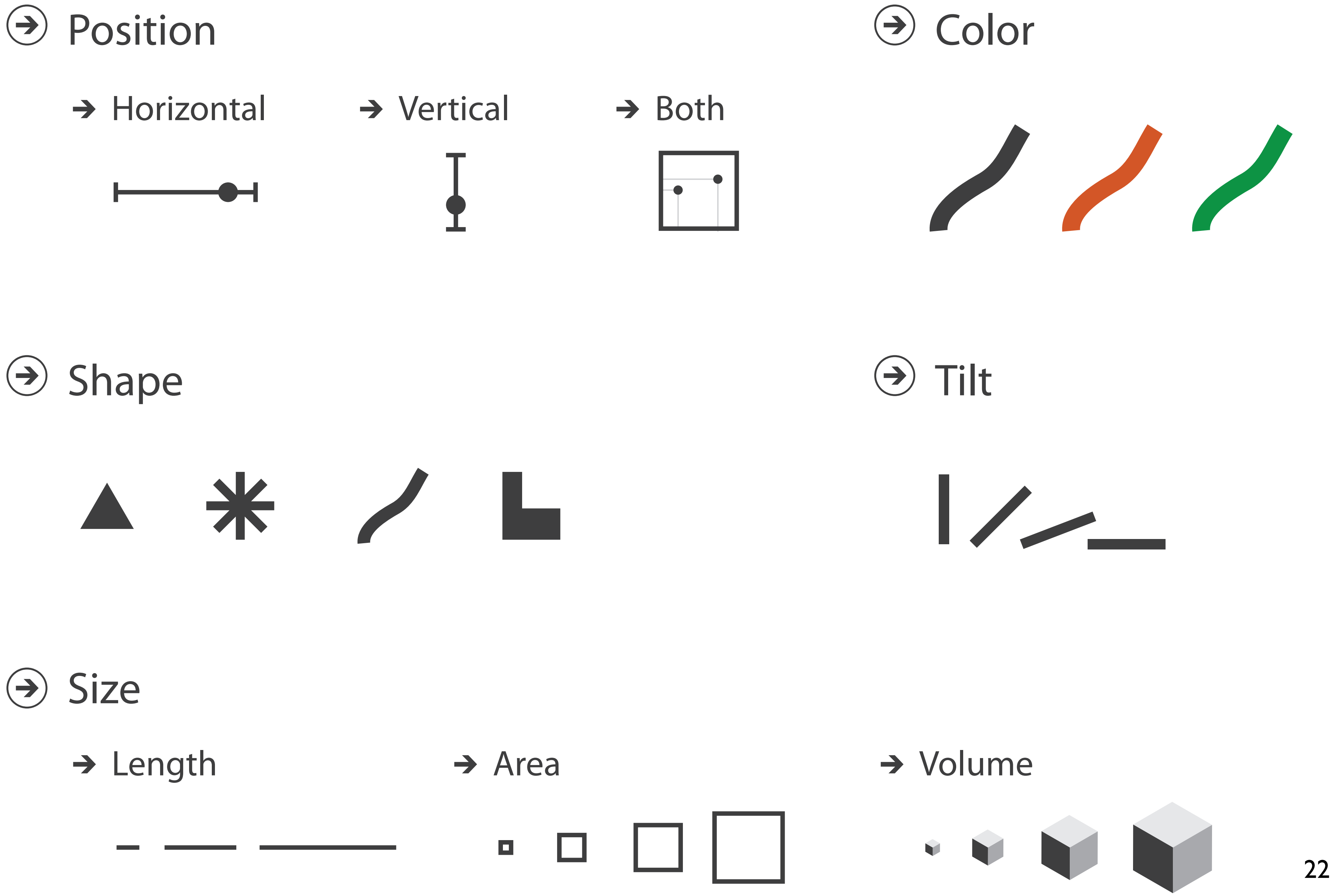
- analyze idiom structure



Definitions: Marks and channels

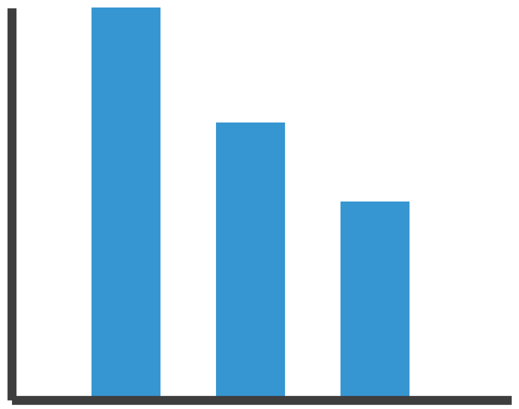
- marks
  - geometric primitives

- channels
  - control appearance of marks



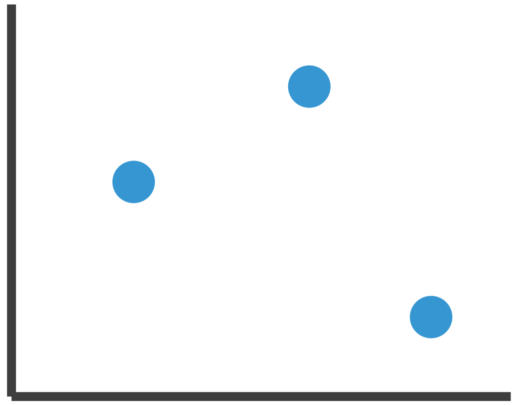
Encoding visually with marks and channels

- analyze idiom structure
  - as combination of marks and channels



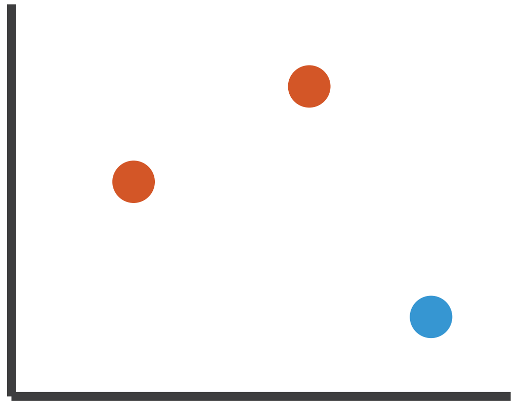
1:
vertical position

mark: line



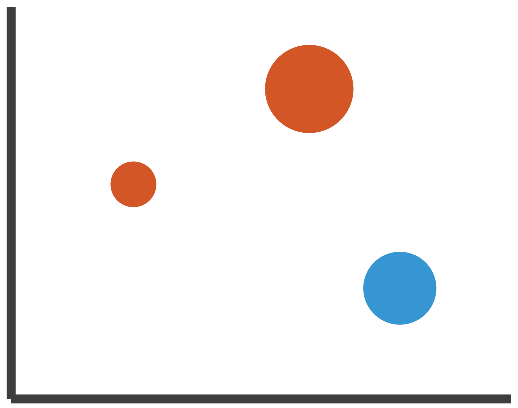
2:
vertical position
horizontal position

mark: point



3:
vertical position
horizontal position
color hue

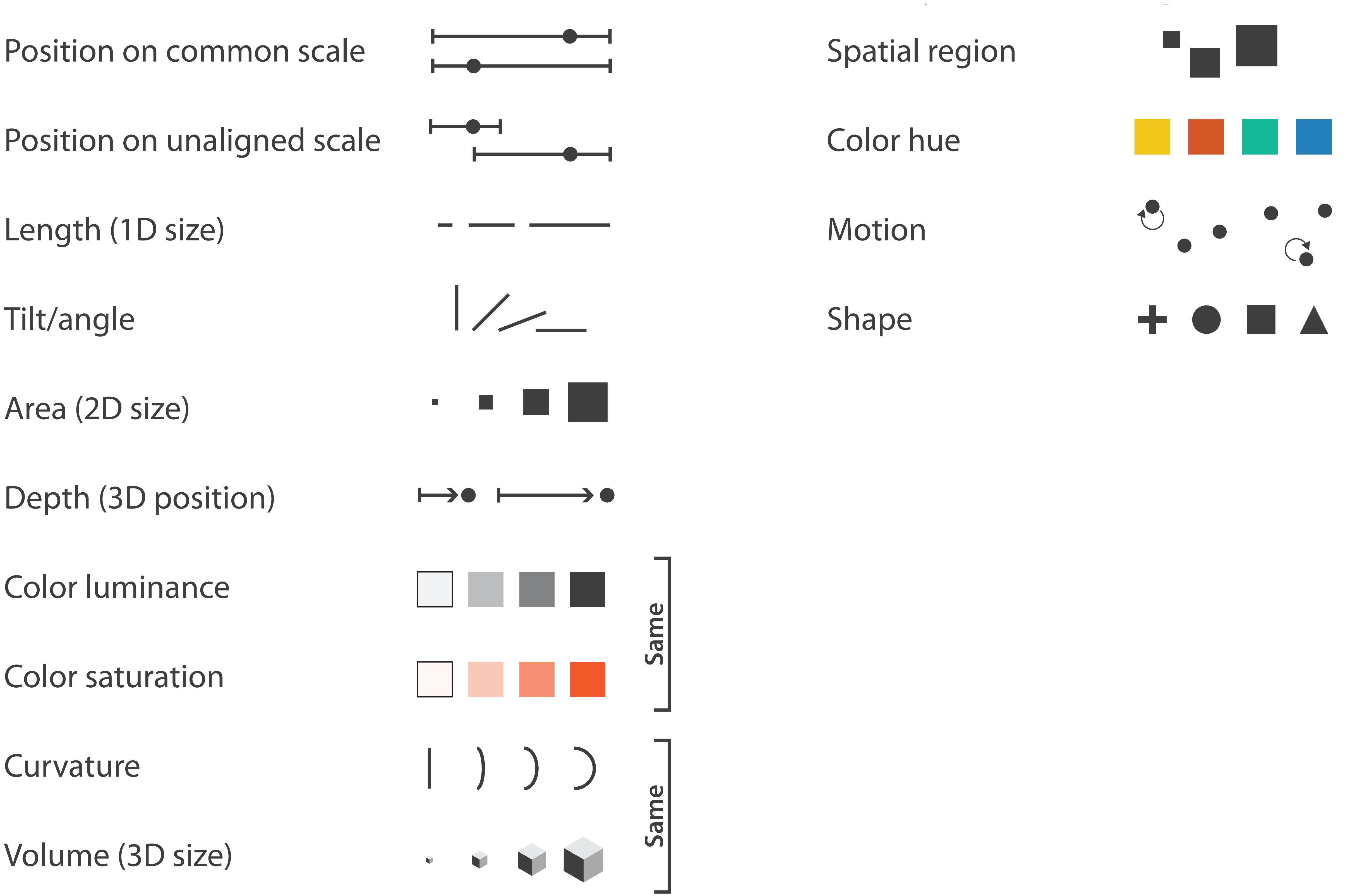
mark: point



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

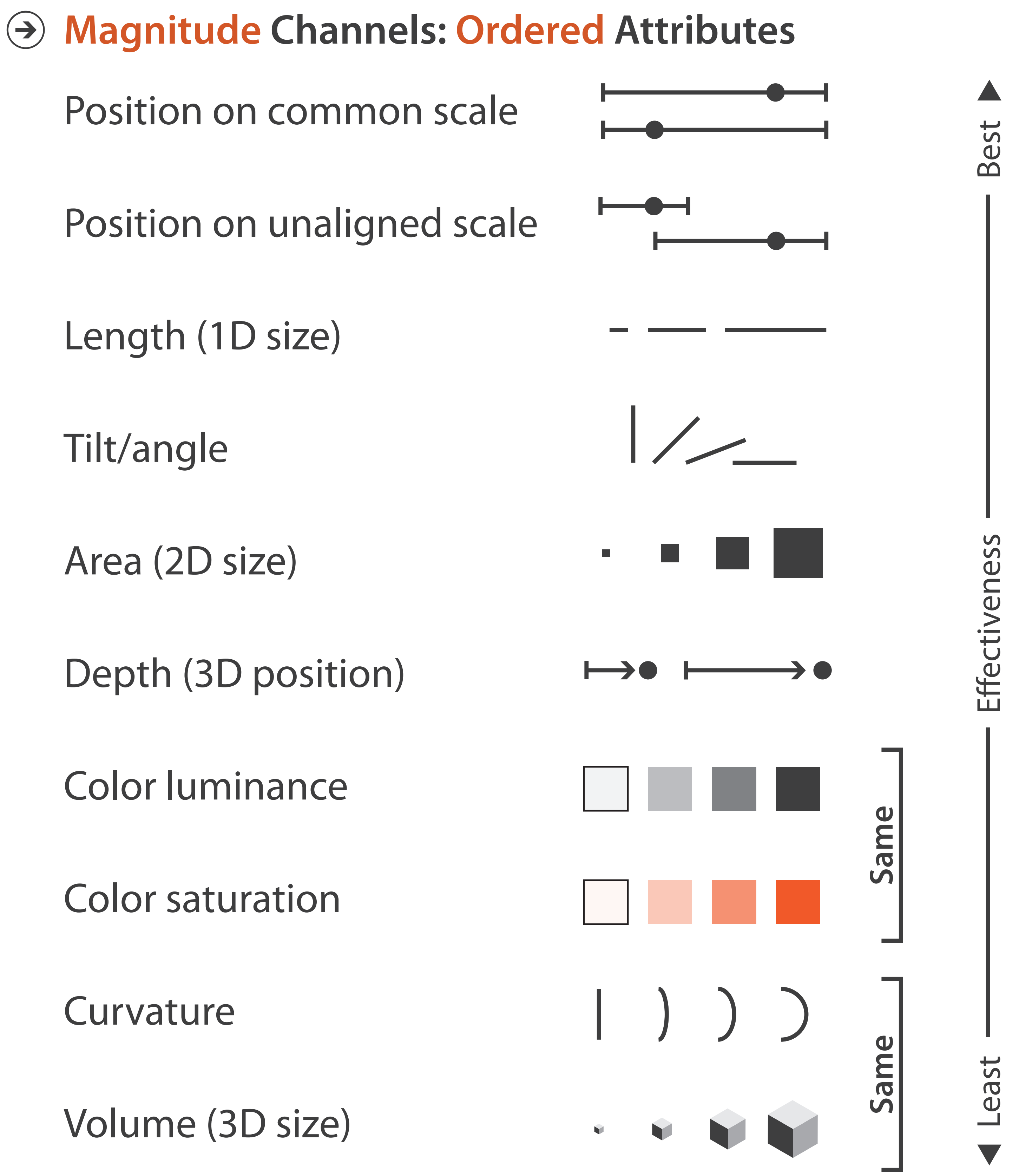
mark: point

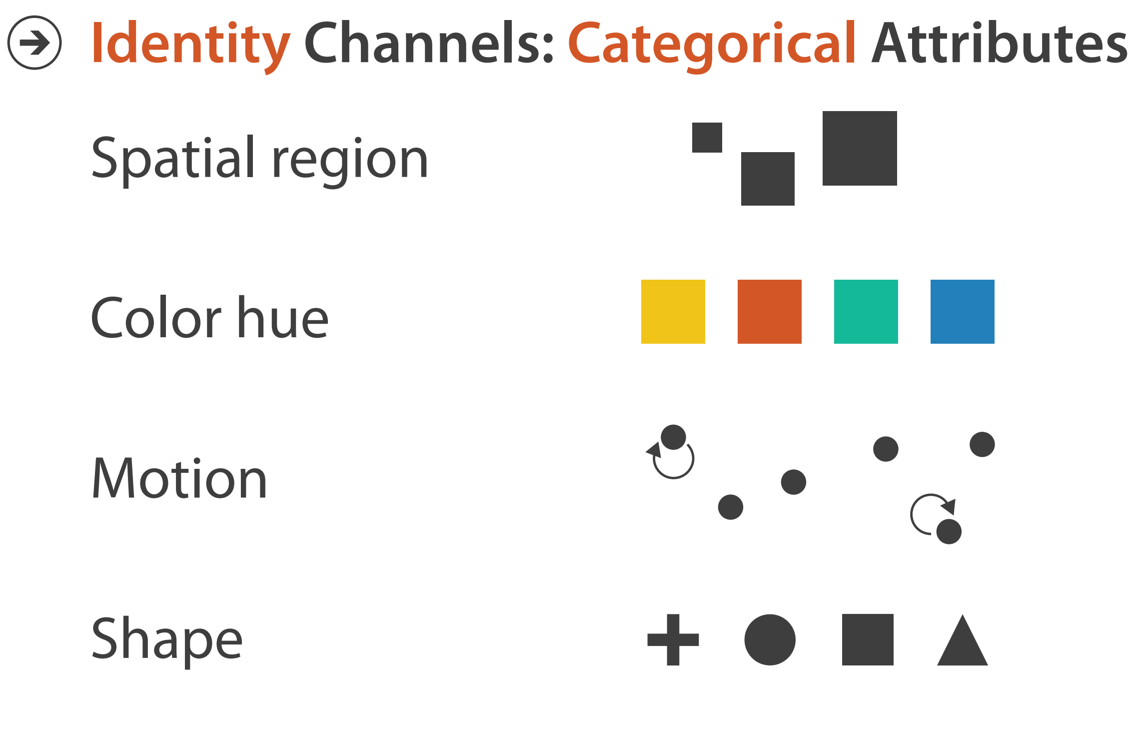
Channels



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: Rankings

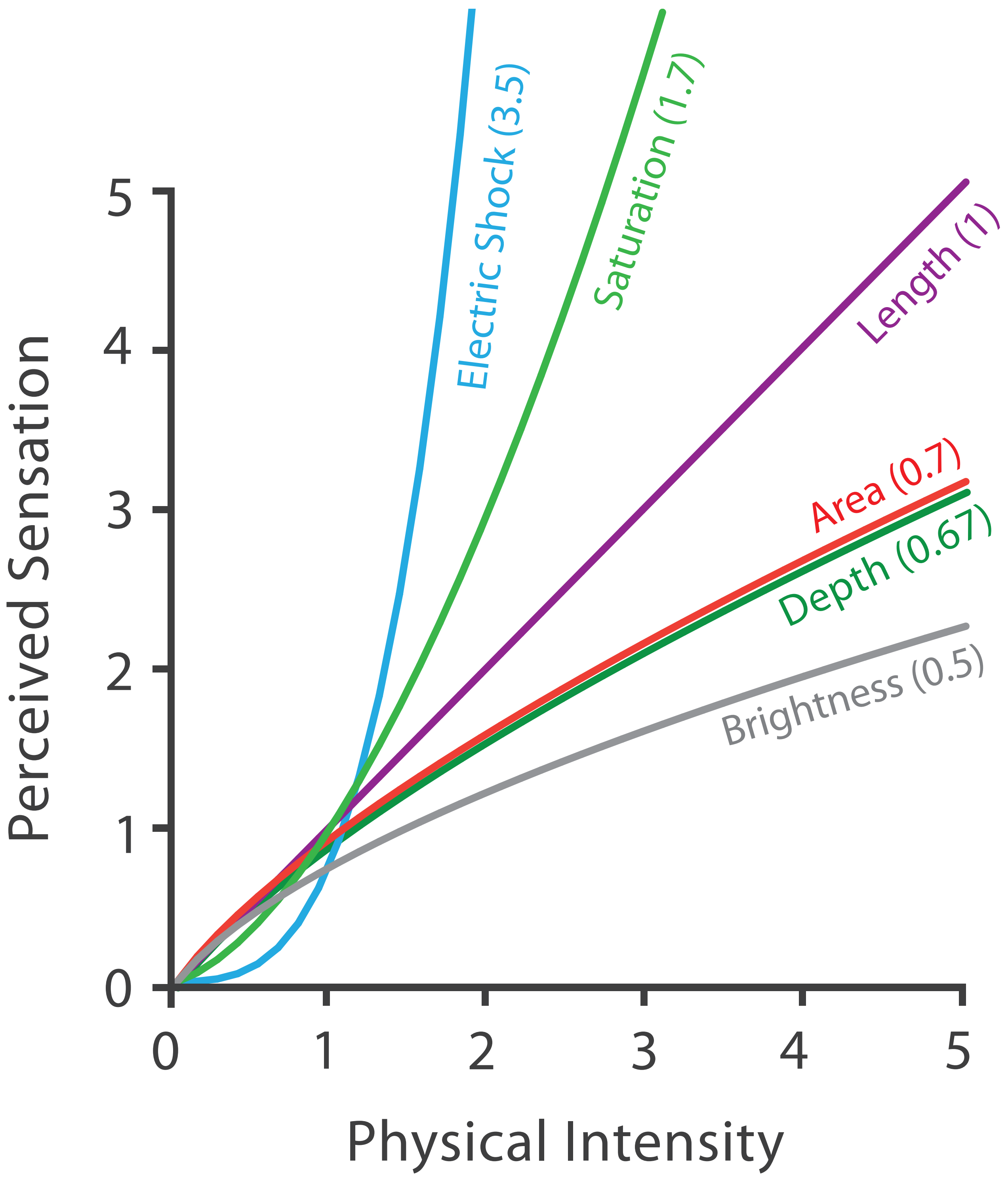




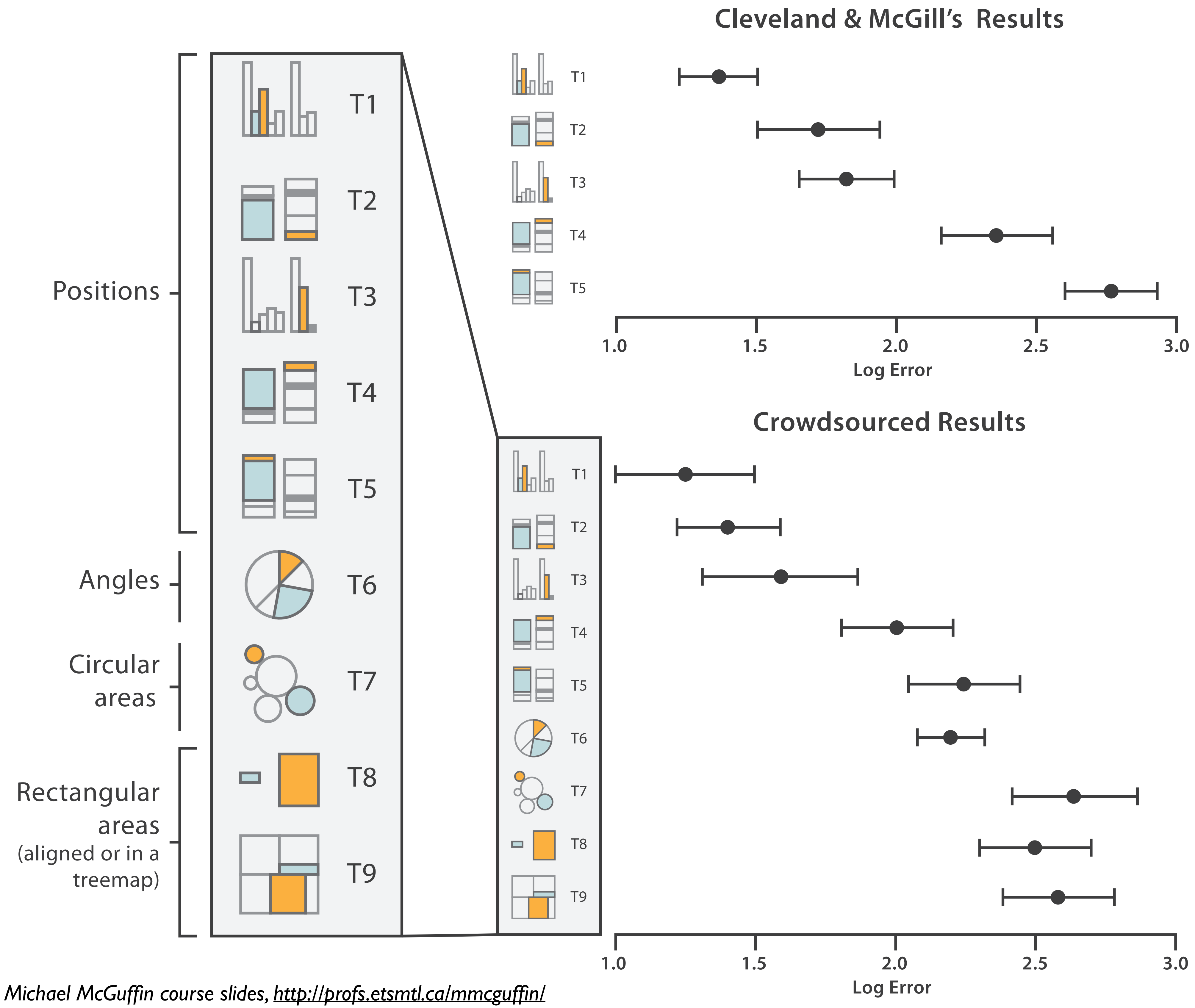
- effectiveness principle
  - encode most important attributes with highest ranked channels
- expressiveness principle
  - match channel and data characteristics

Accuracy: Fundamental Theory

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



Accuracy: Vis experiments

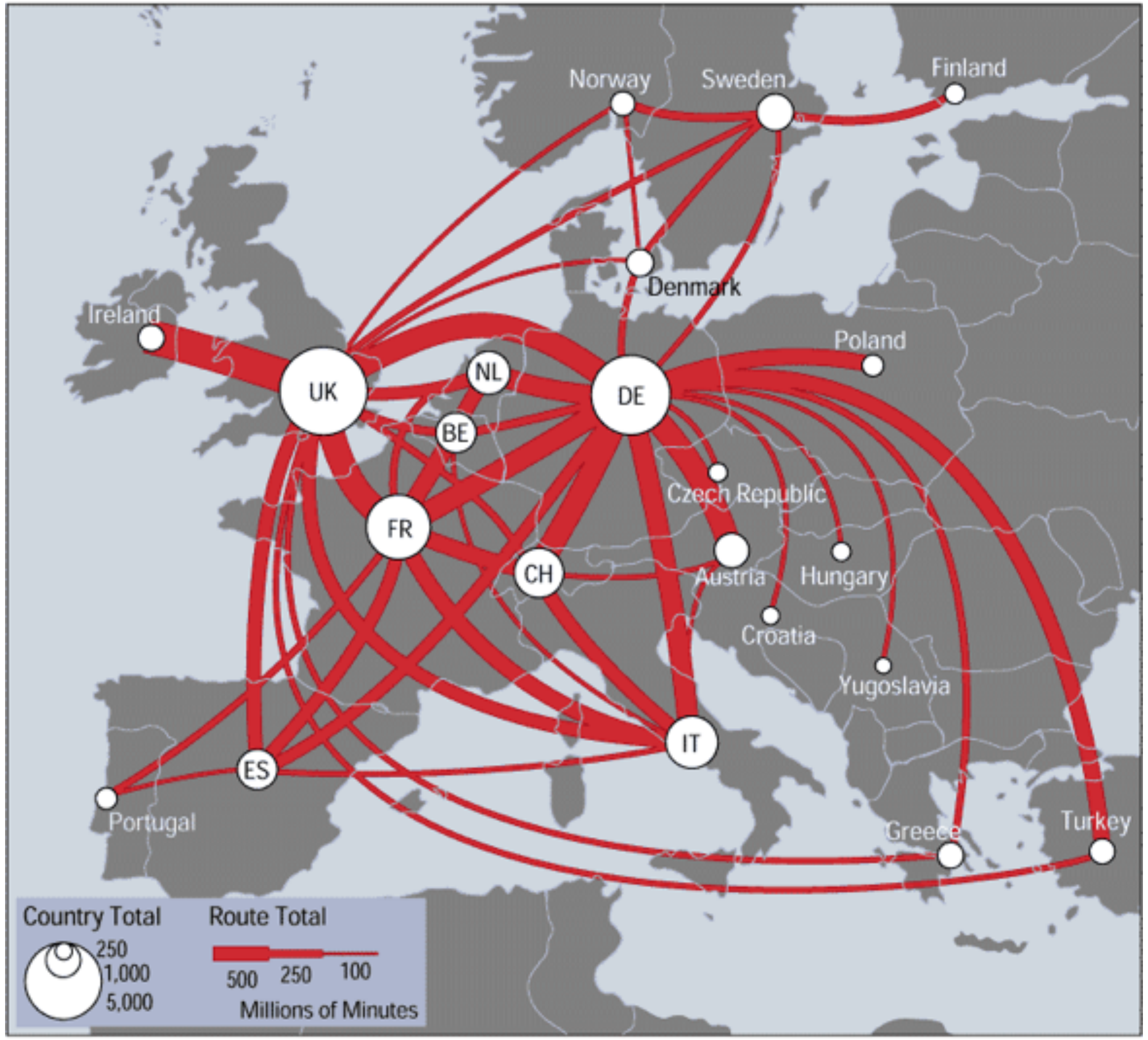


[Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. Heer and Bostock. Proc ACM Conf. Human Factors in Computing Systems (CHI) 2010, p. 203–212.]

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

Discriminability: How many usable steps?

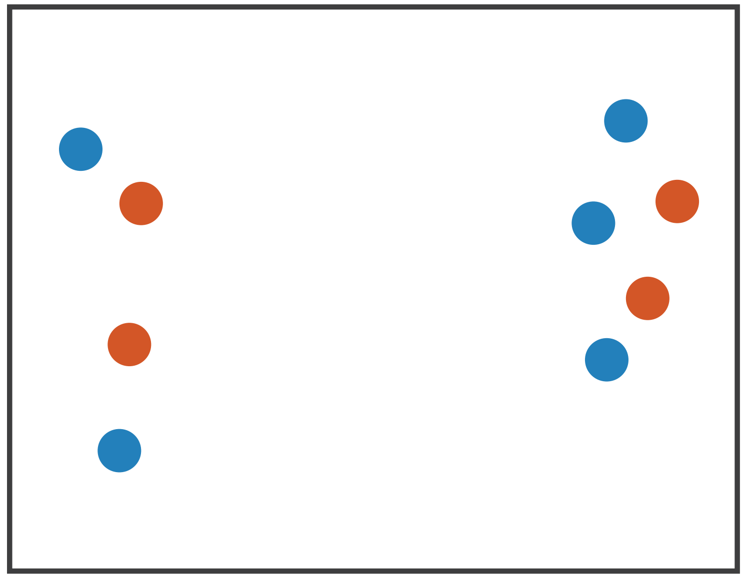
- must be sufficient for number of attribute levels to show
  - linewidth: few bins



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

Separability vs. Integrality

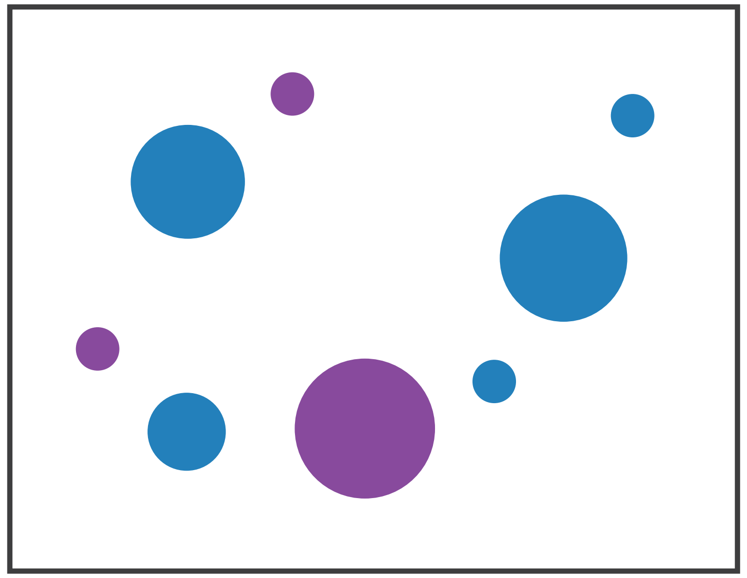
Position
+ Hue (Color)



Fully separable

2 groups each

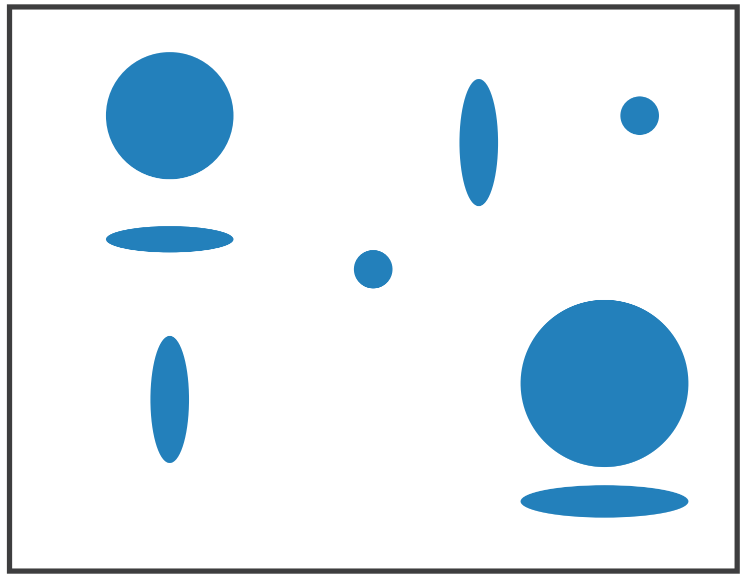
Size
+ Hue (Color)



Some interference

2 groups each

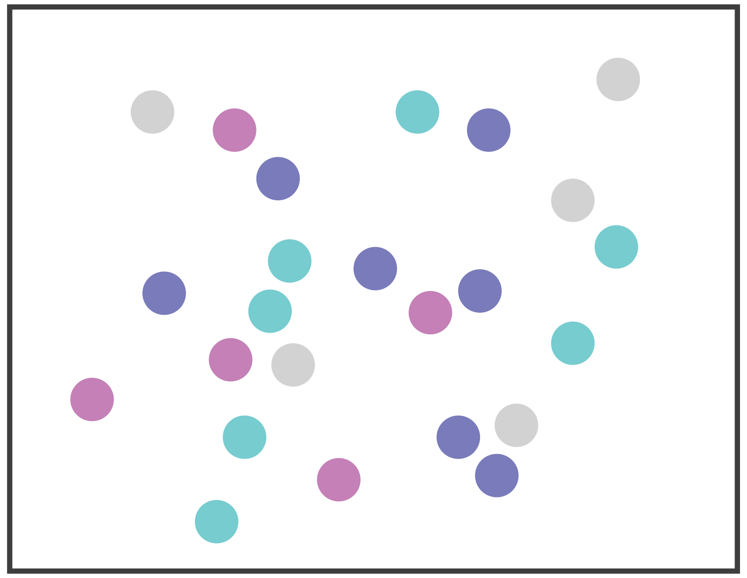
Width
+ Height



Some/significant interference

3 groups total: integral area

Red
+ Green



Major interference

4 groups total: integral hue

Further reading

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

Outline

- Session 1: Principles *9:15-10:30am*
  - Analysis: What, Why, How
  - Marks and Channels, Perception
  - Color
- Session 2: Techniques for Scaling *10:50-11:40am*
  - Manipulate: Change, Select, Navigate
  - Facet: Juxtapose, Partition, Superimpose
  - Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

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

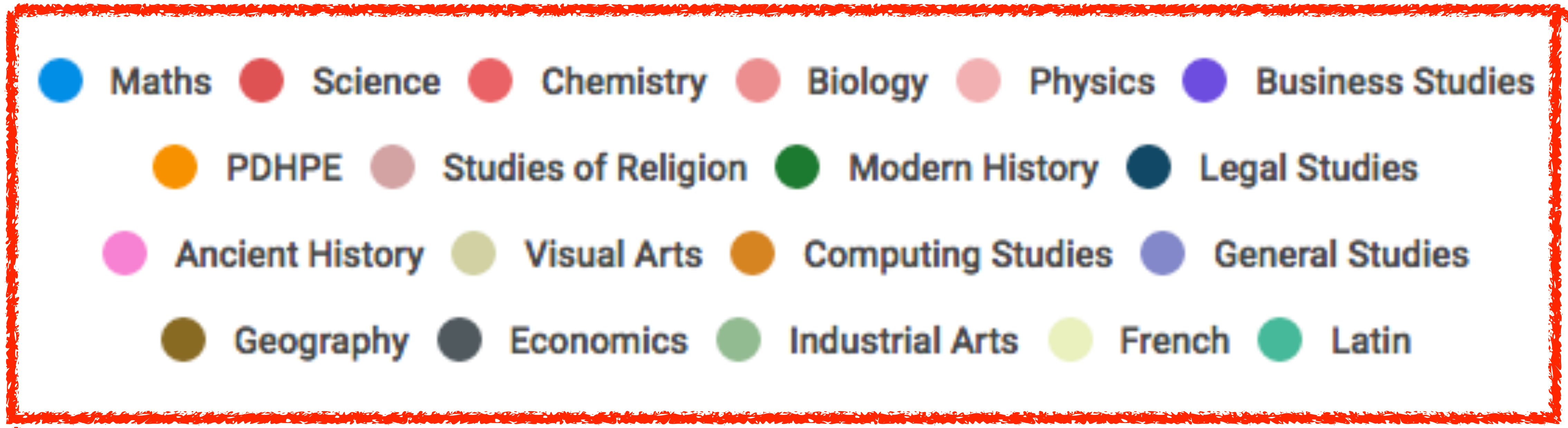
What?

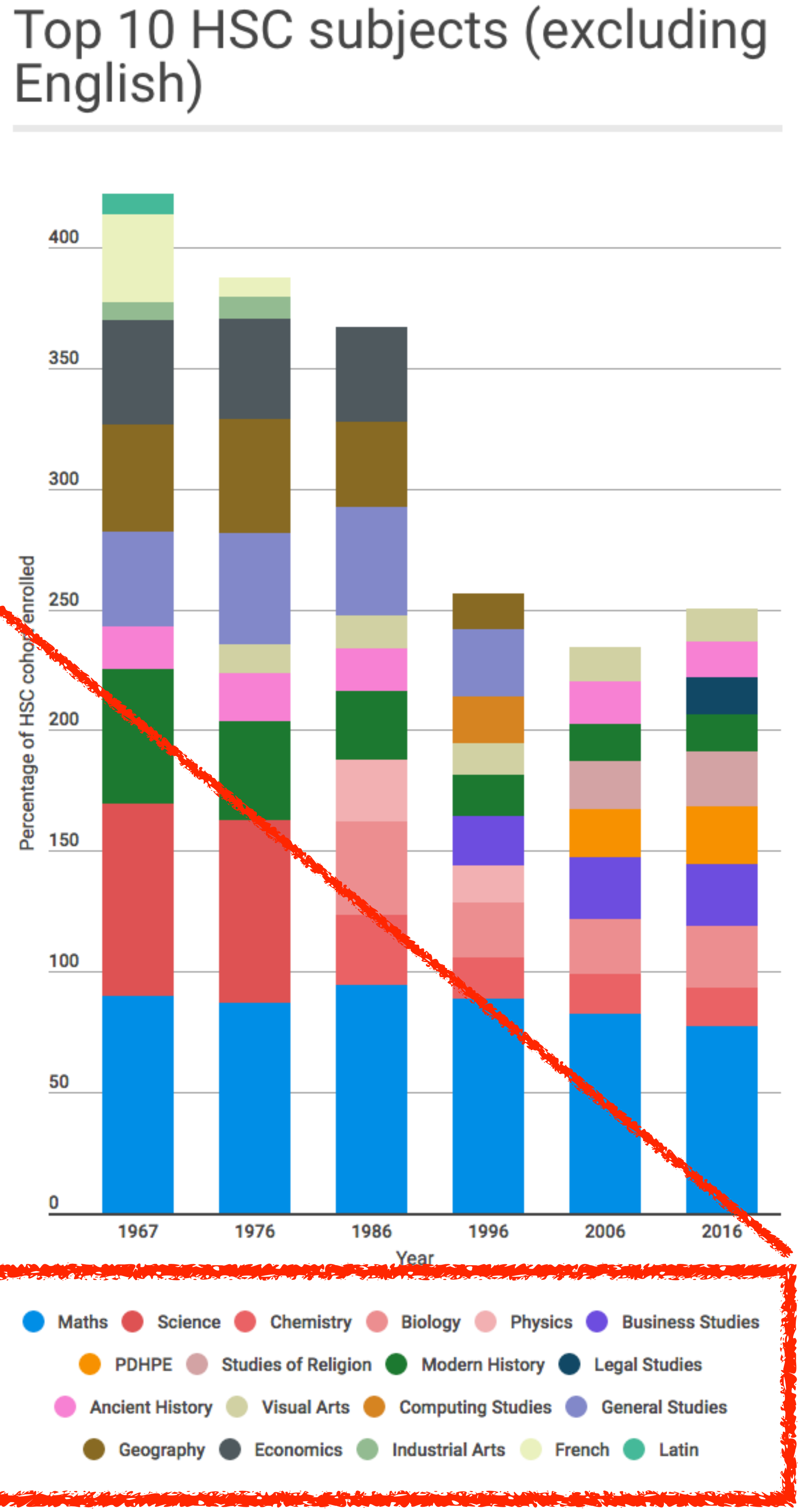
Why?

How?

Challenges of Color

- what is wrong with this picture?



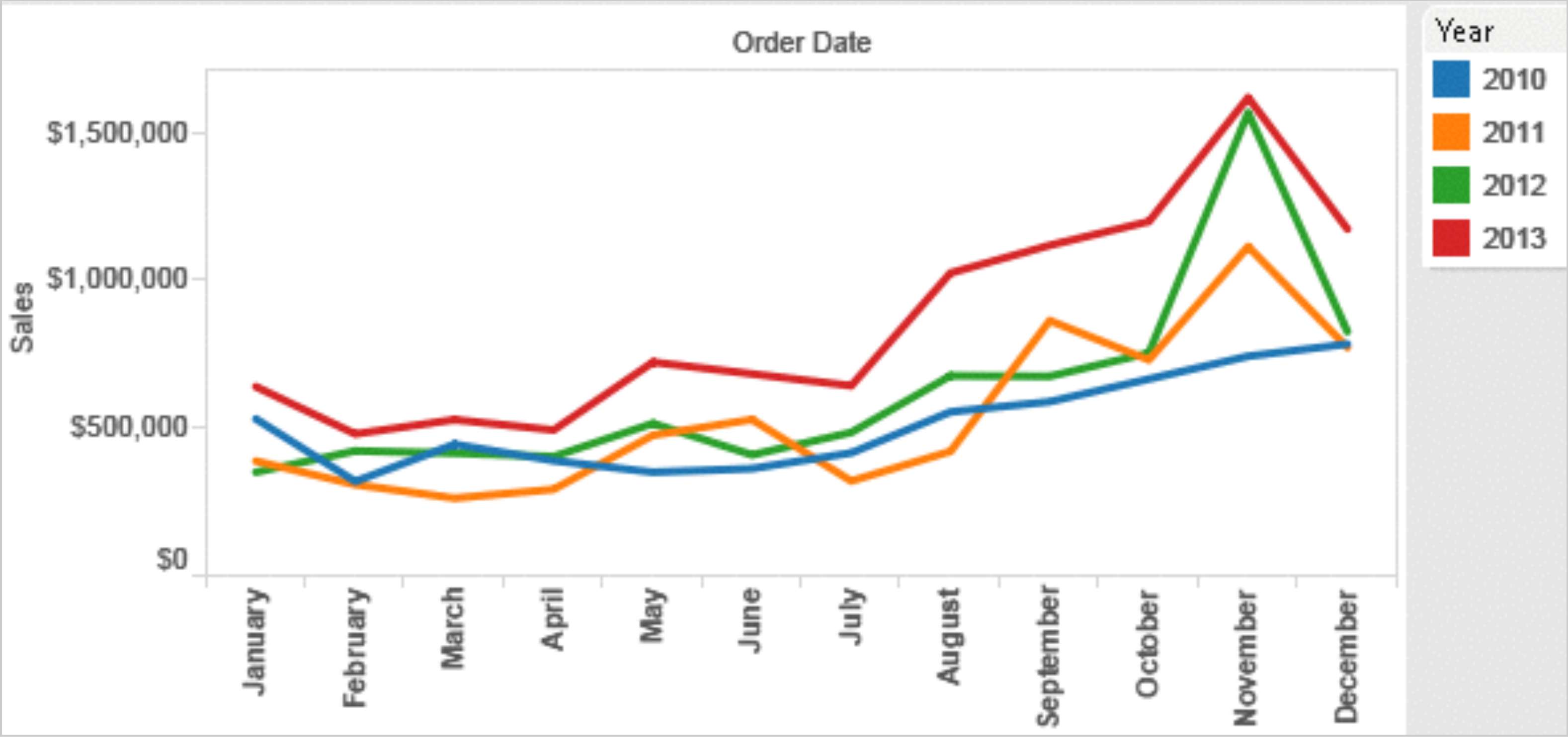


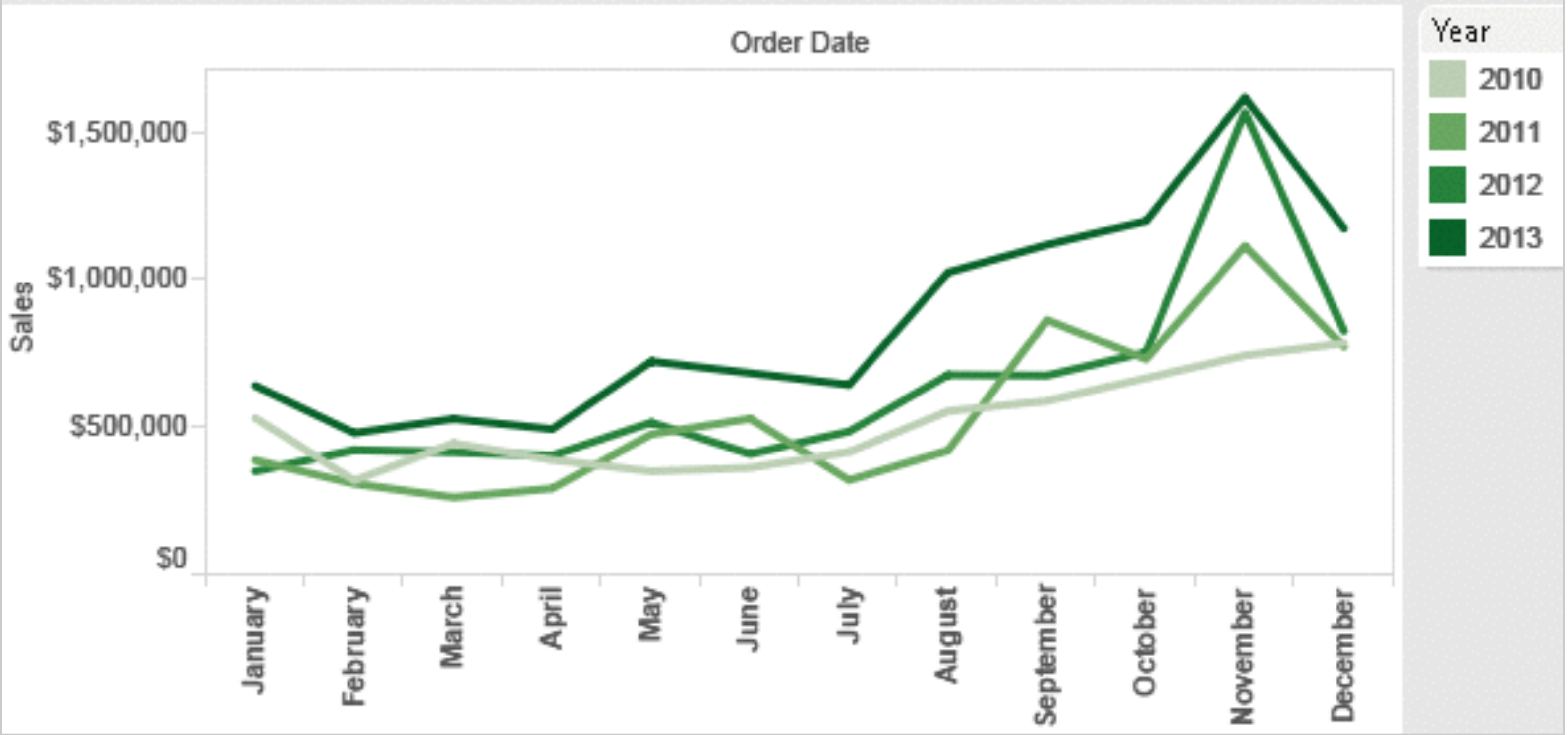
@WTFViz

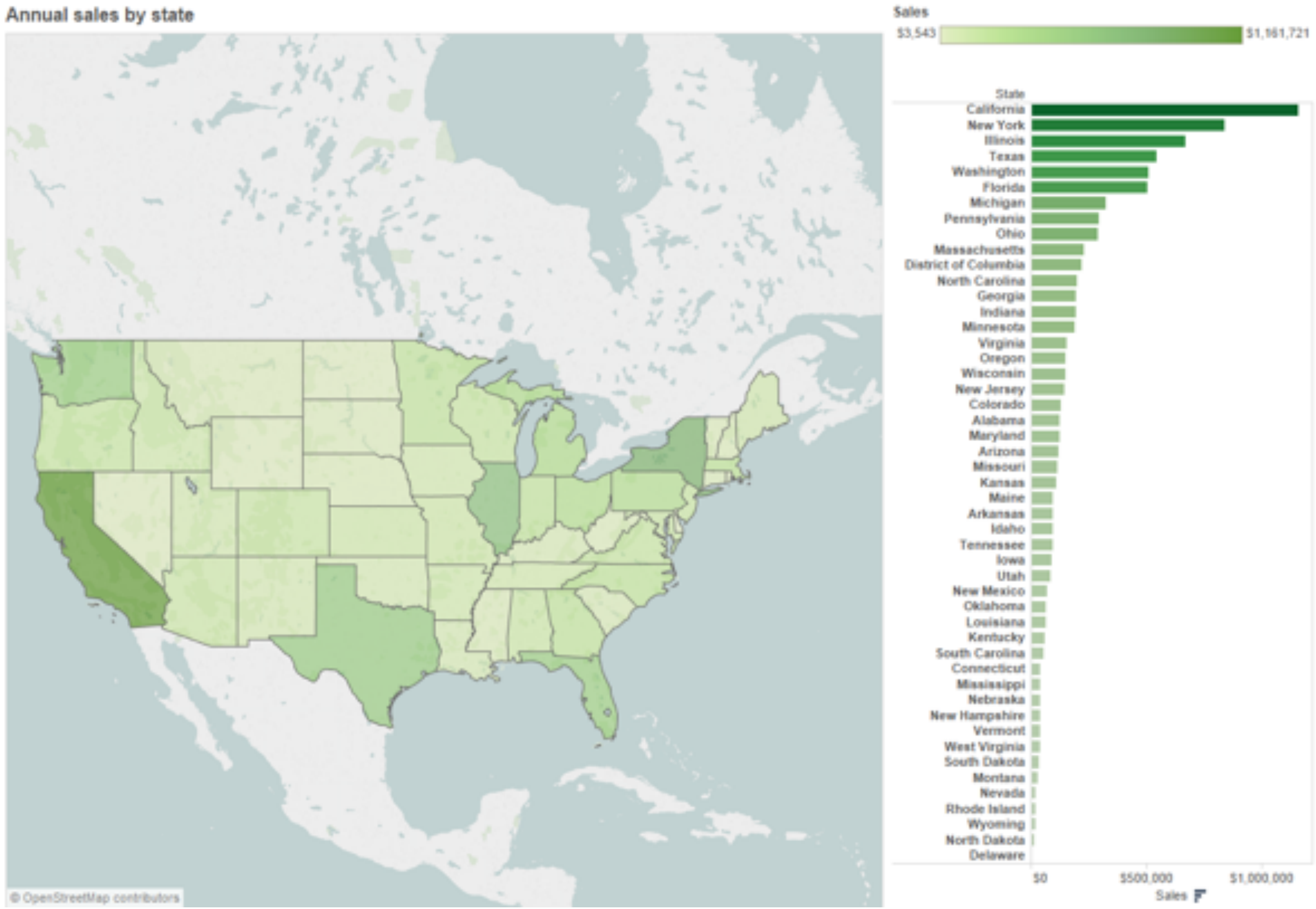
"visualizations that make no sense"

http://viz.wtf/post/150780948819/maths-enrolments-drop-to-lowest-rate-in-50-years

Categorical vs ordered color





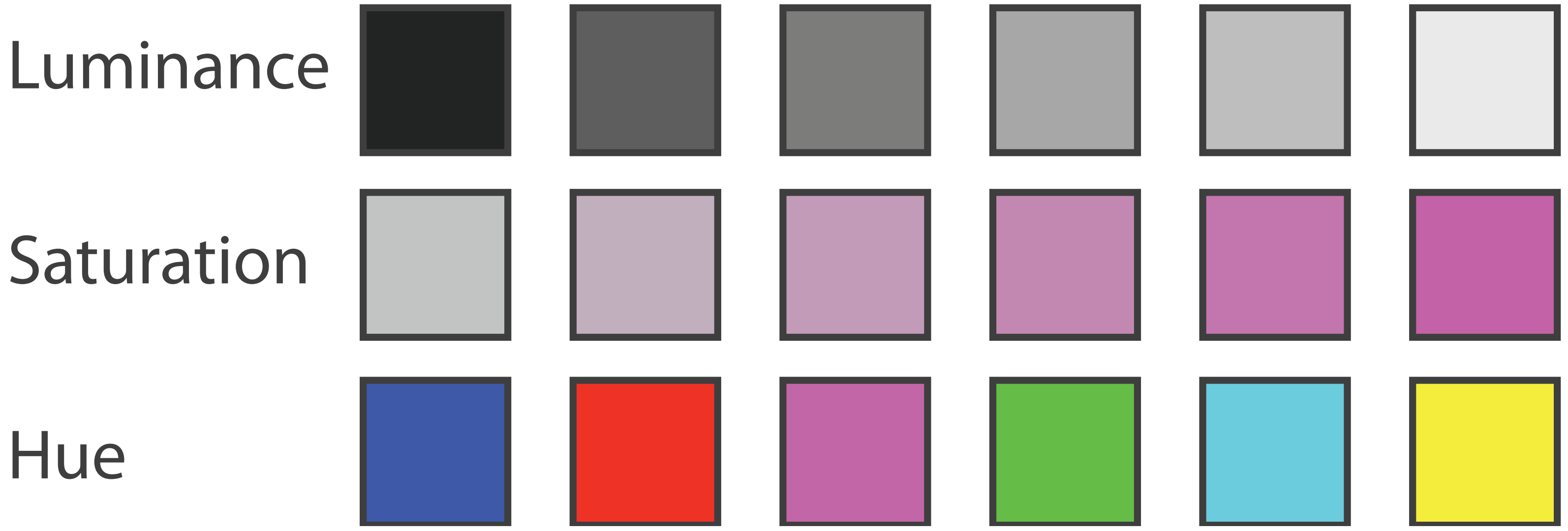


[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

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
    - saturation
  - categorical can show identity
    - hue

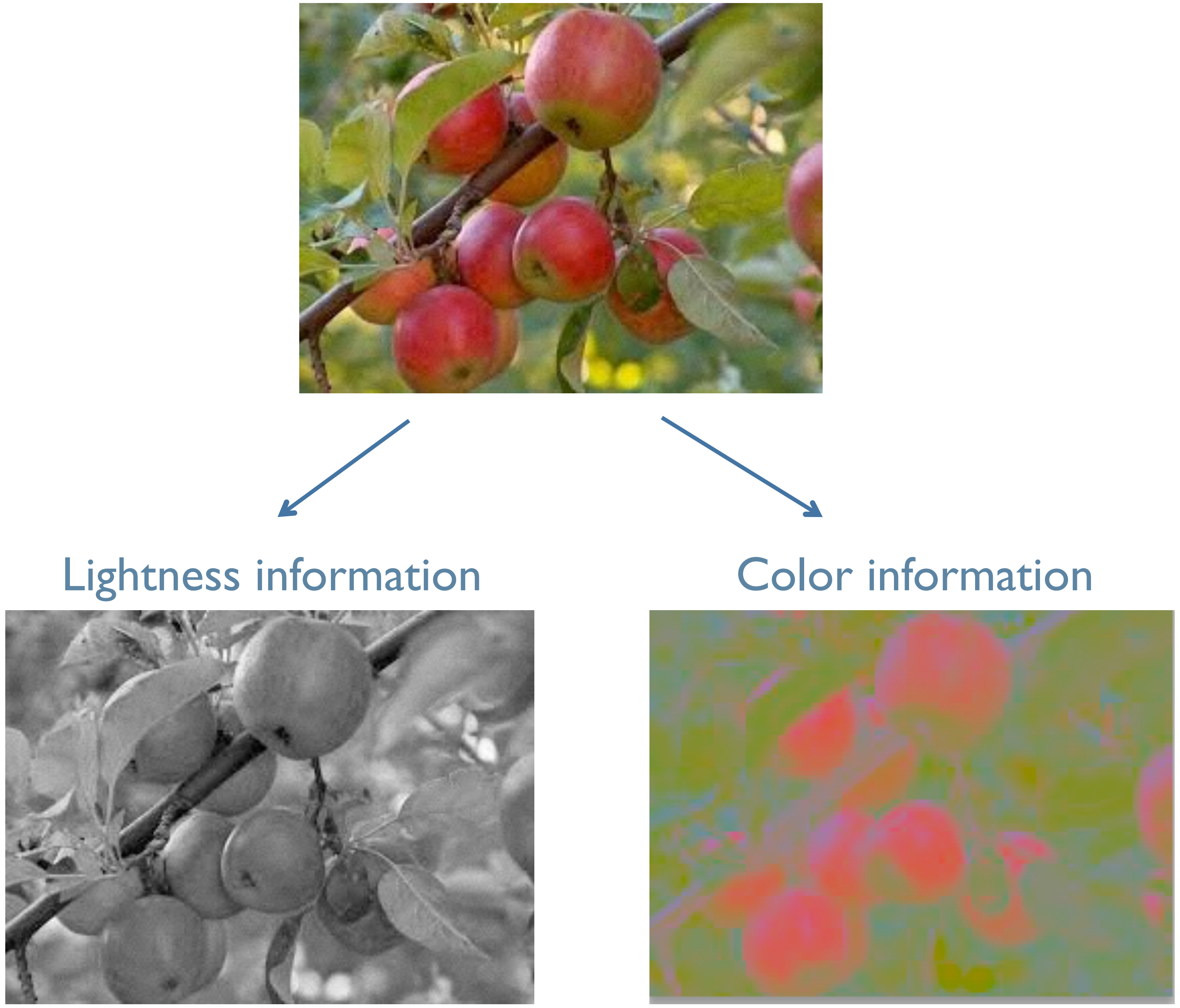


- channels have different properties
  - what they convey directly to perceptual system
  - how much they can convey: how many discriminable bins can we use?

Luminance

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

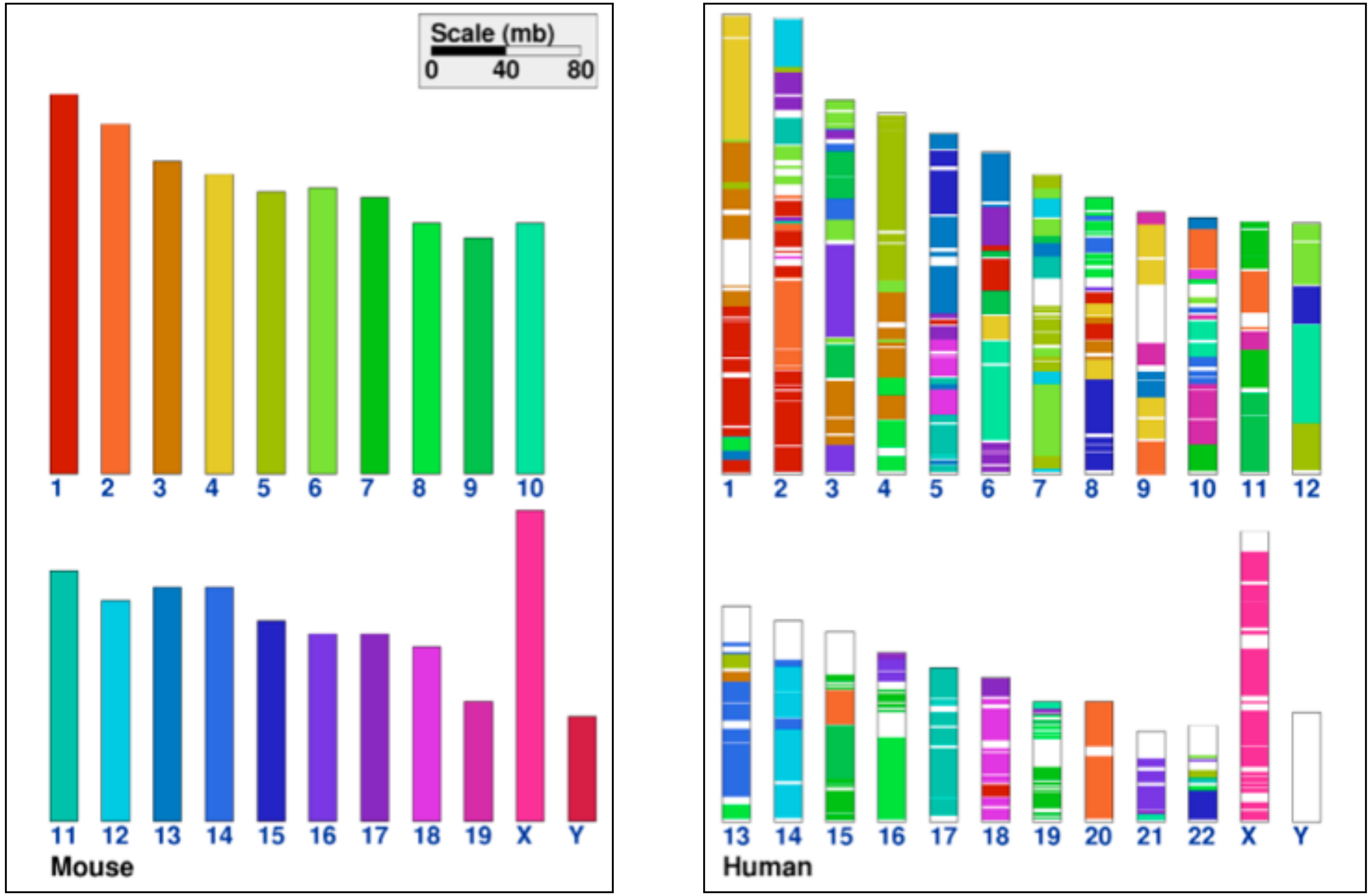
- intrinsic perceptual ordering



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

Categorical color: limited number of discriminable bins

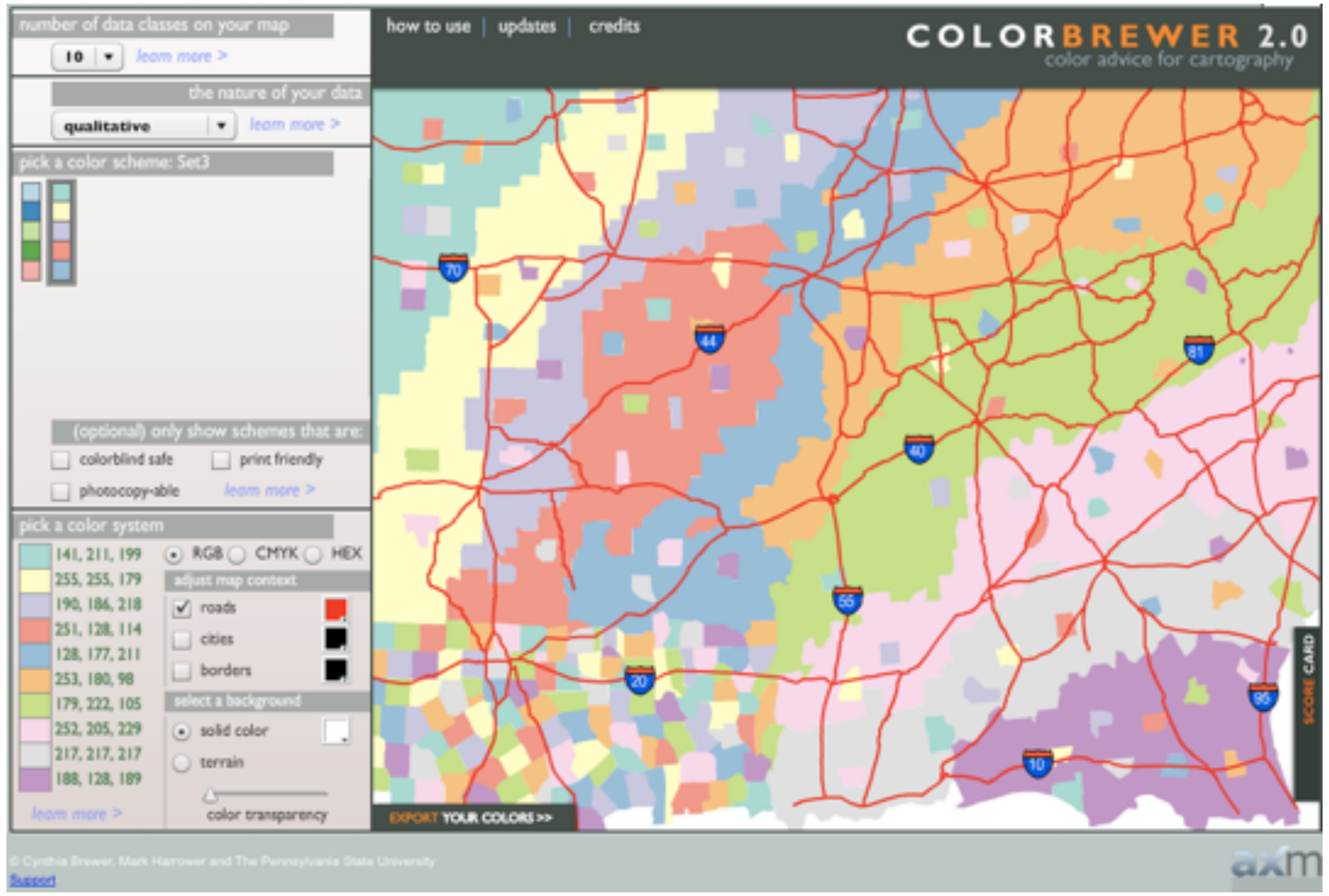
- human perception built on relative comparisons
  - great if color contiguous
  - surprisingly bad for absolute comparisons
- noncontiguous small regions of color
  - fewer bins than you want
  - rule of thumb: 6-12 bins, including background and highlights

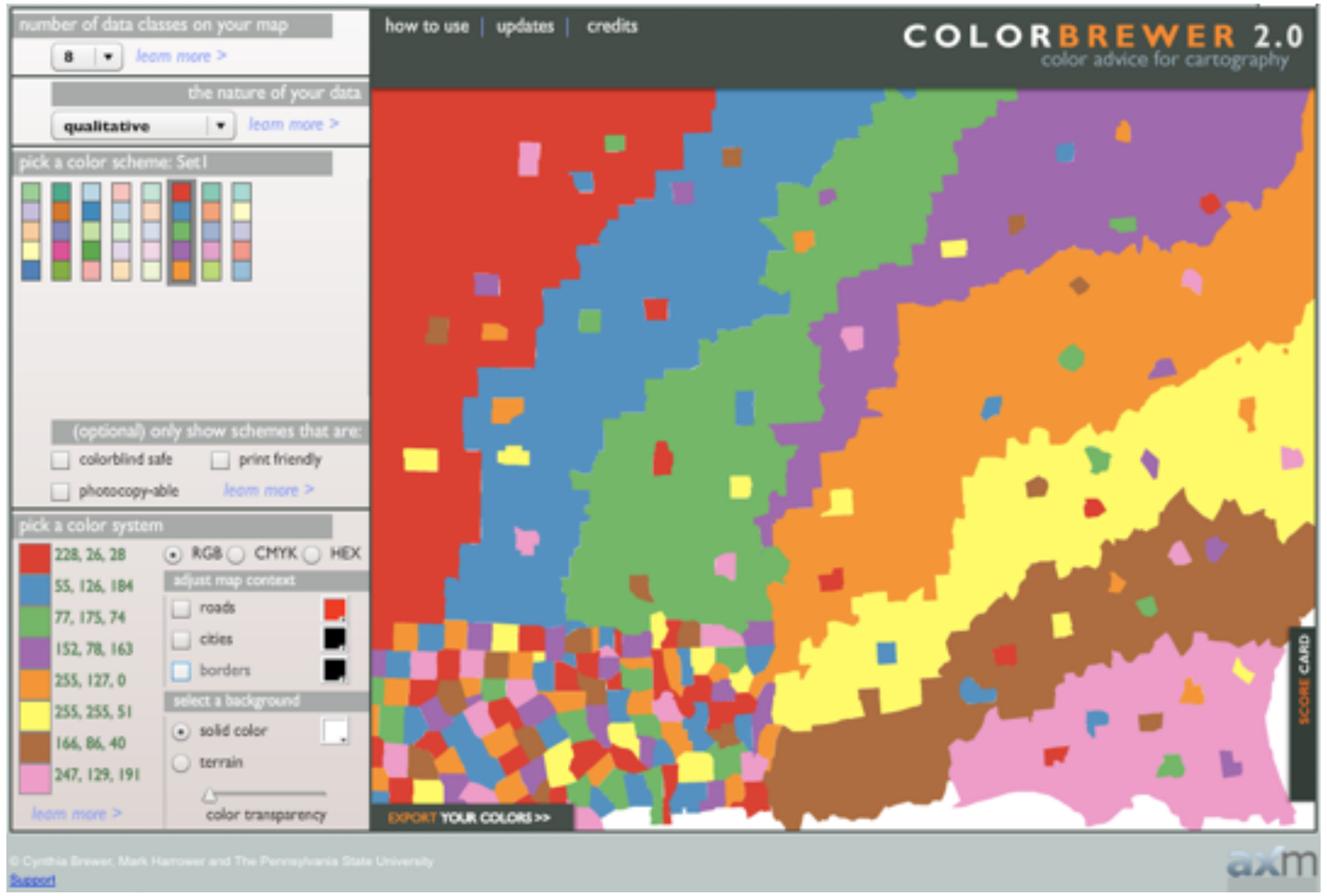


[Cinteny: flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. Sinha and Meller. BMC Bioinformatics, 8:82, 2007.]

ColorBrewer

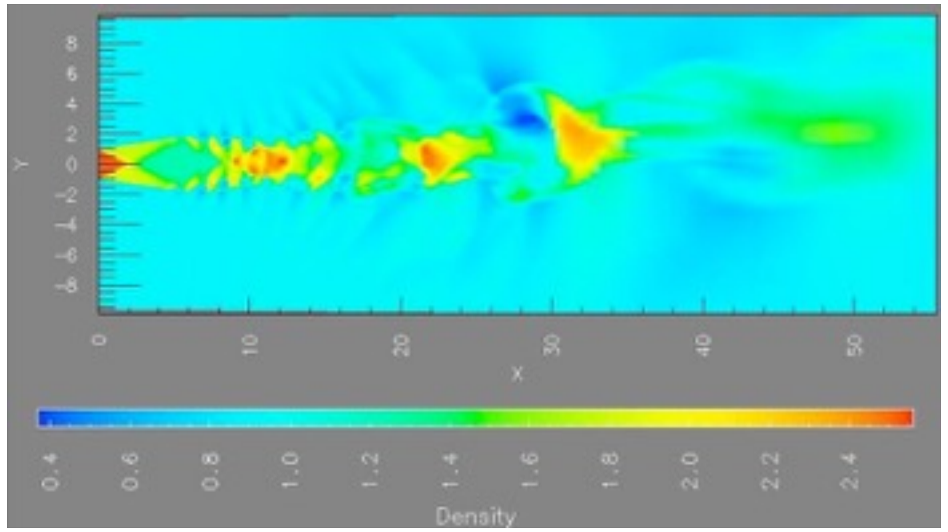
- http://www.colorbrewer2.org
- saturation and area example: size affects salience!



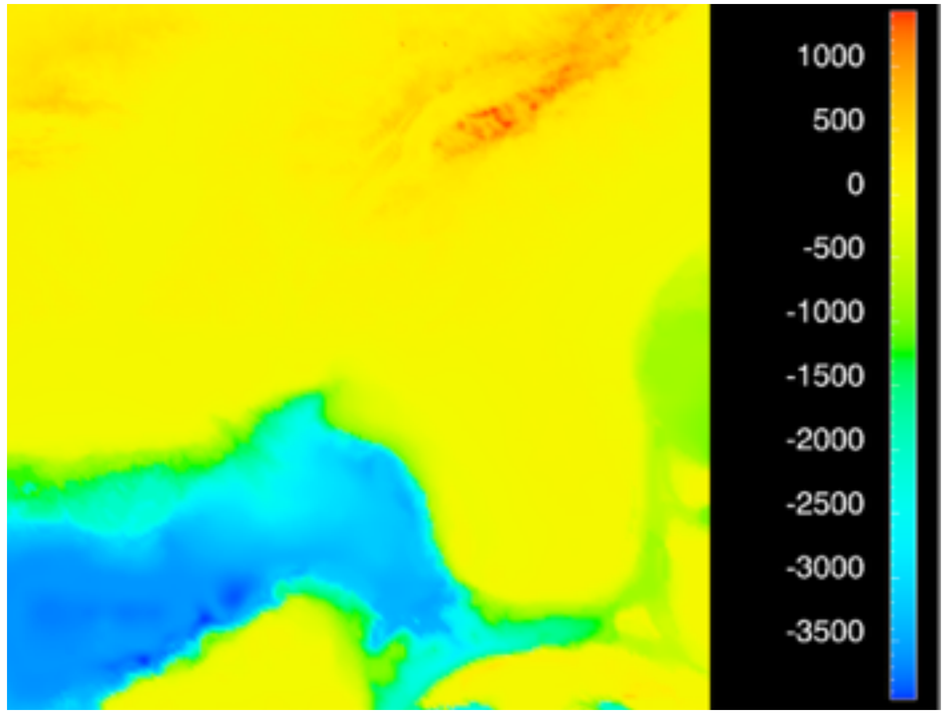


Ordered color: Rainbow is poor default

- 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.]

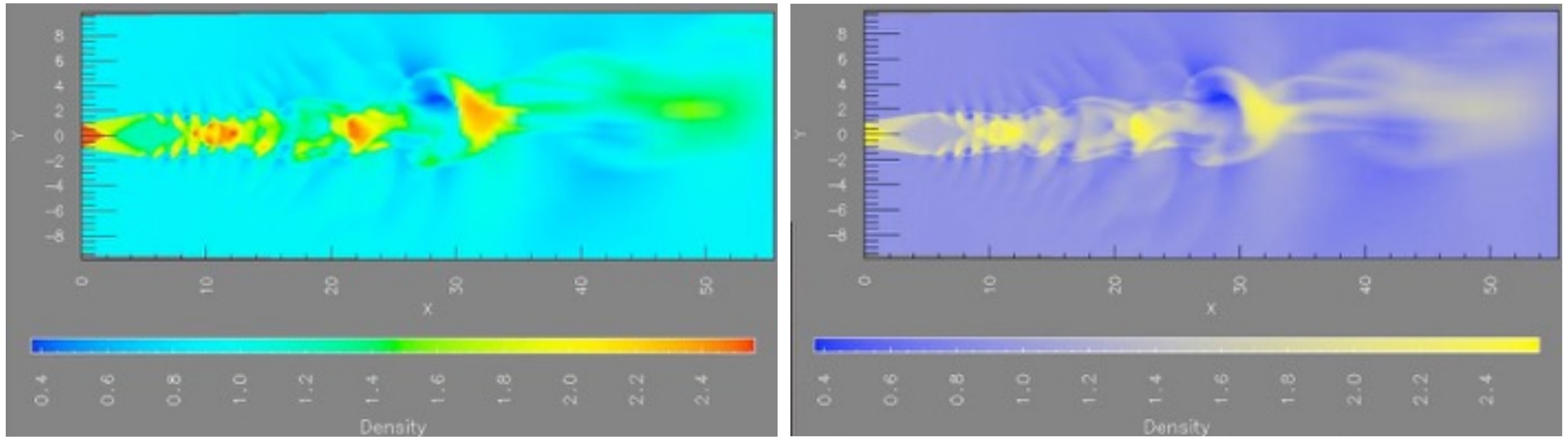


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

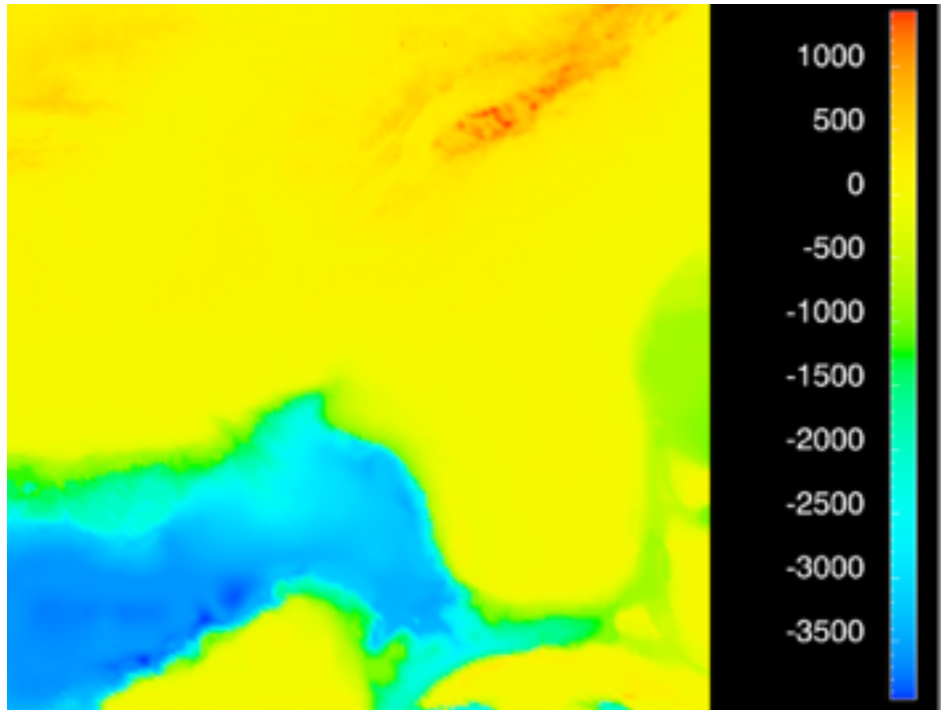
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

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.]

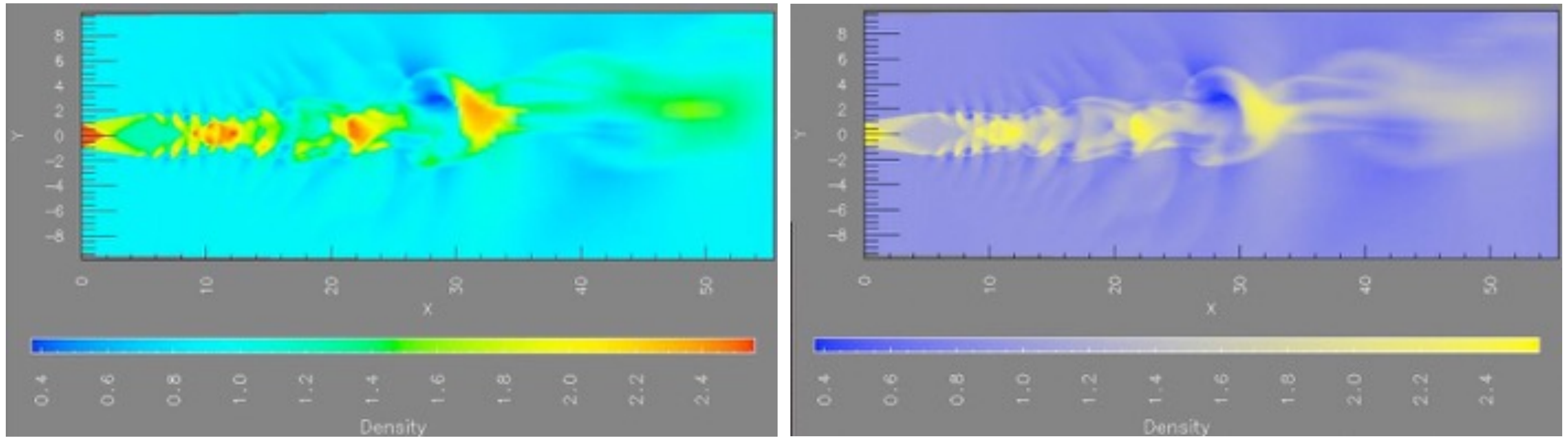


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

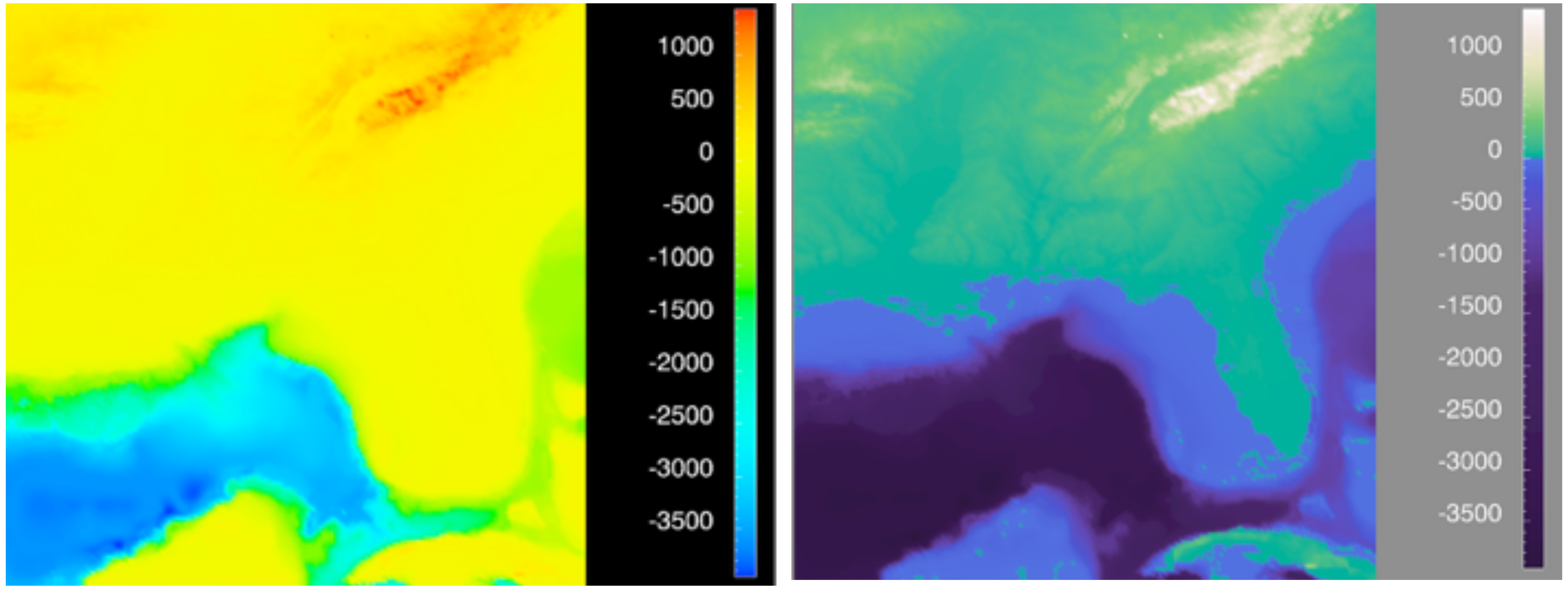
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

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.]

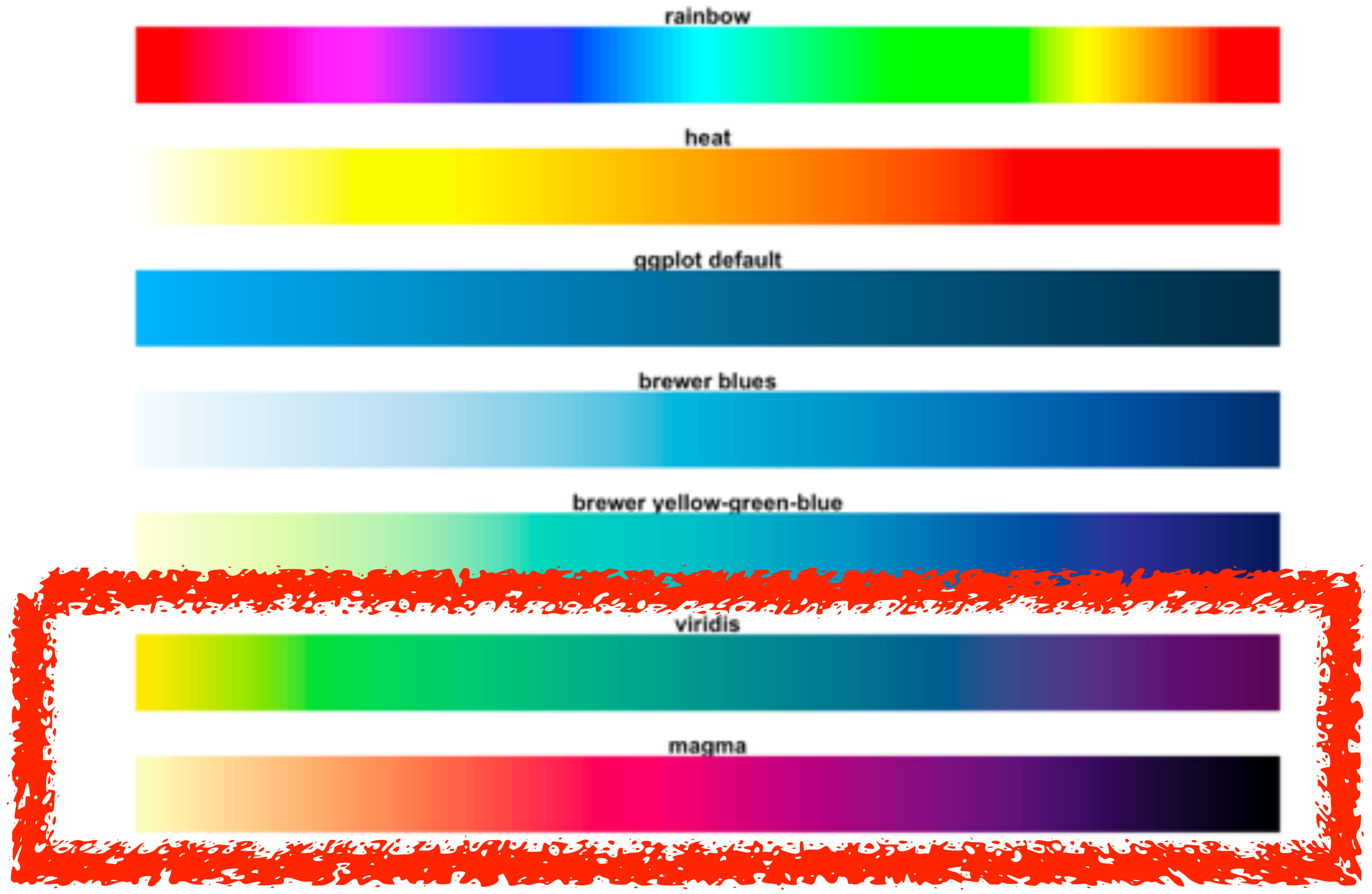


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

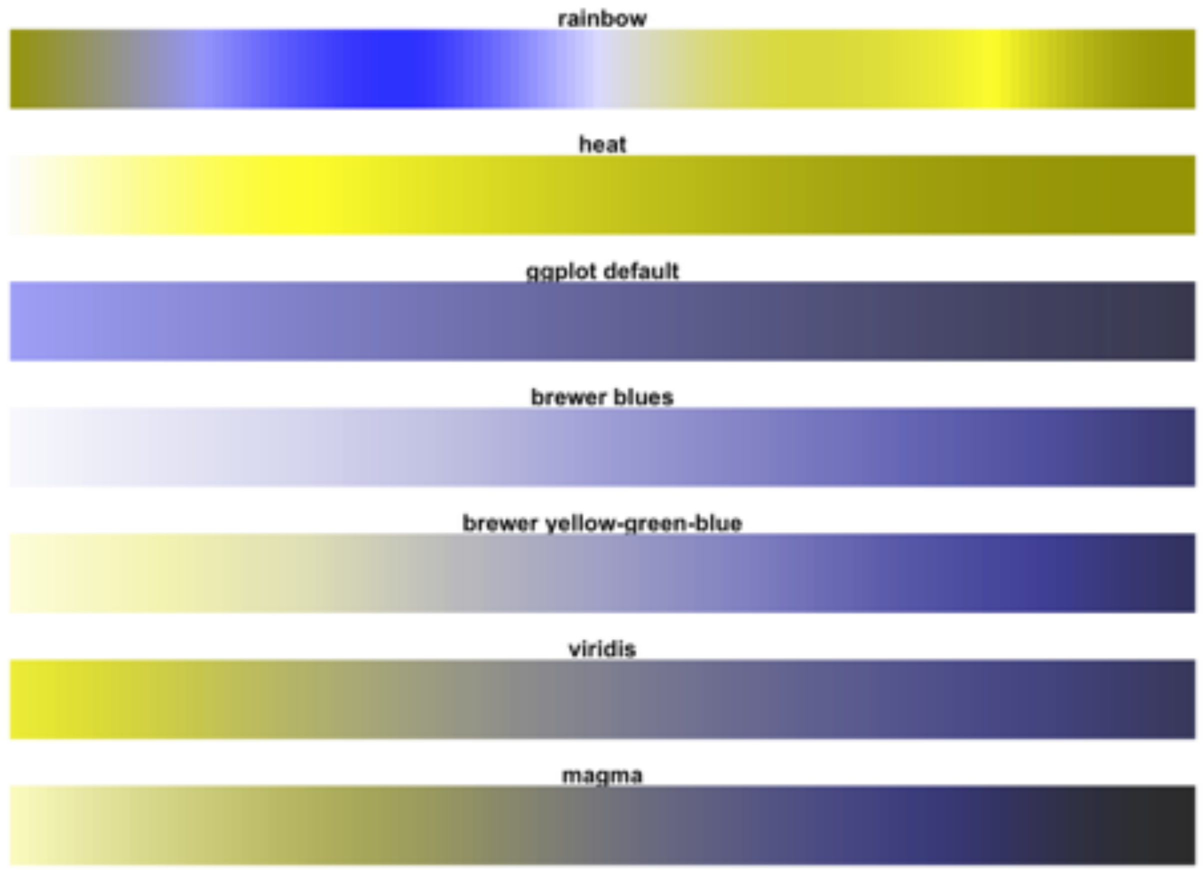
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

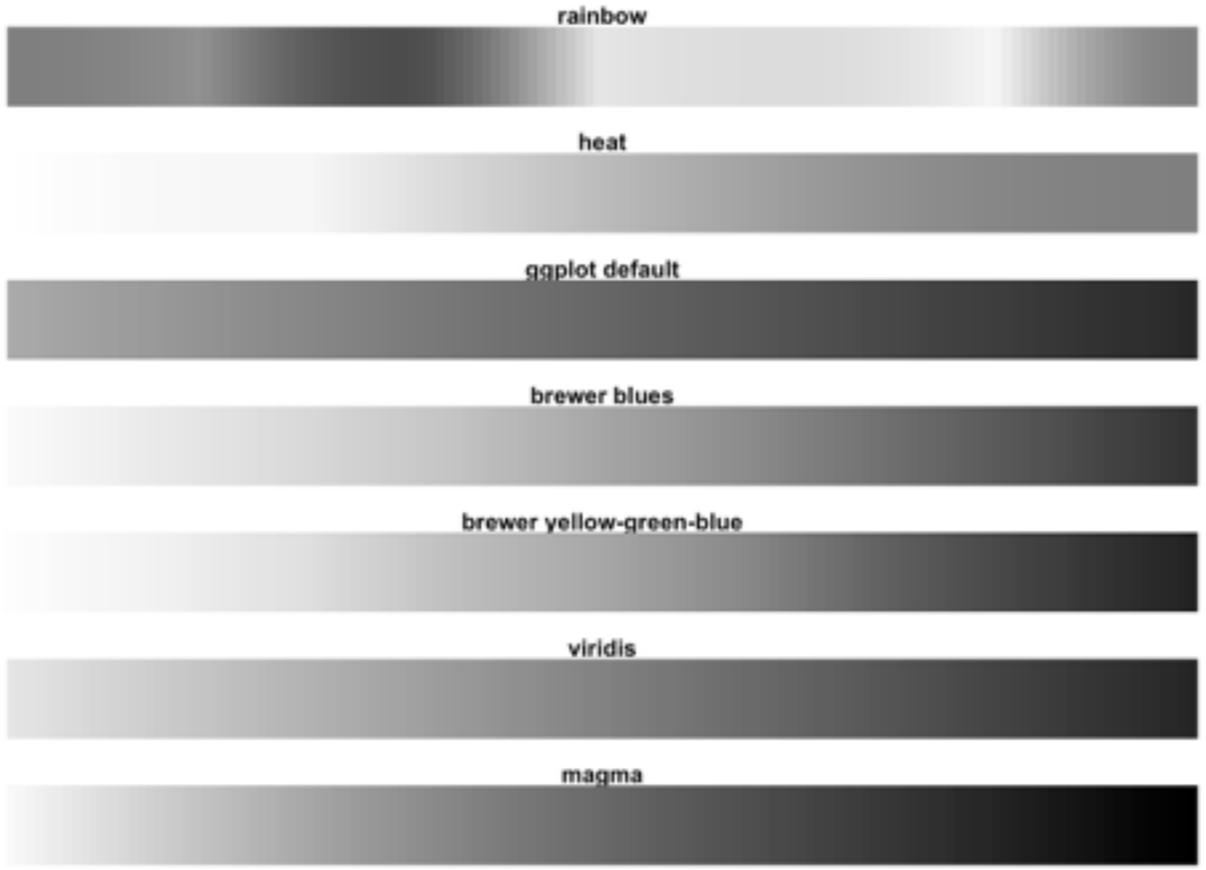
Viridis

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



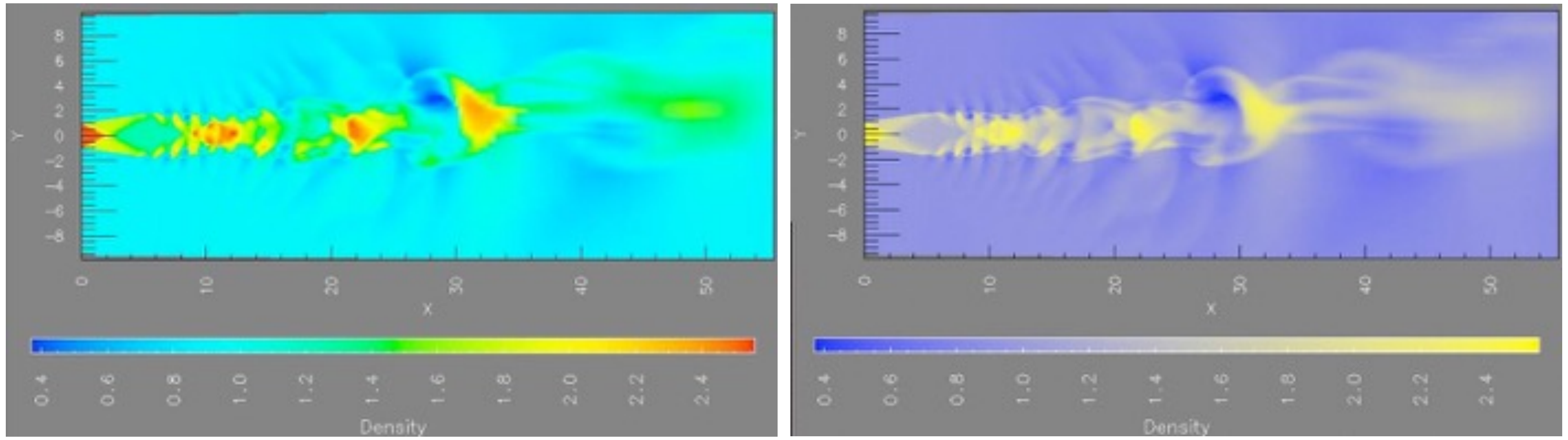
https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html



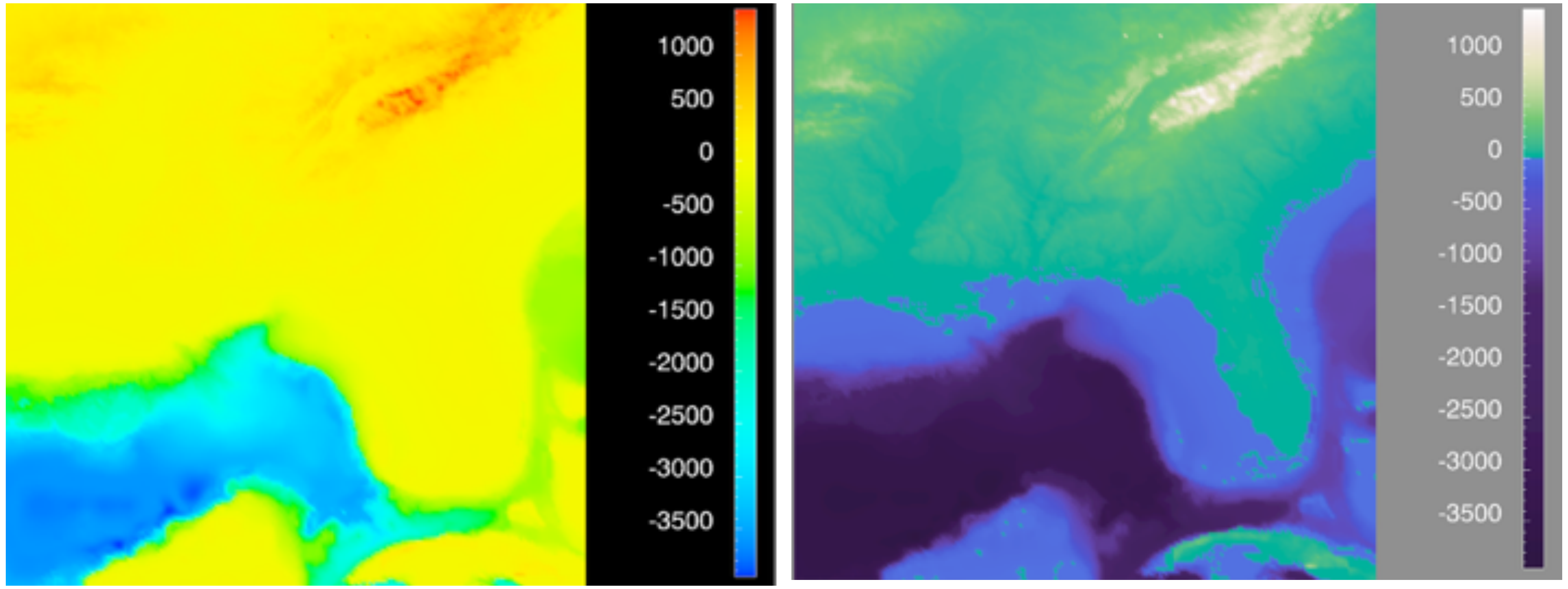


Ordered color: Rainbow is poor default

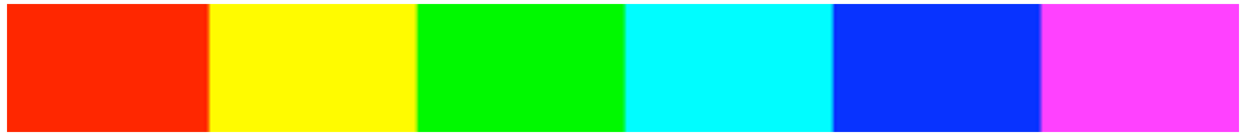
- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable
- alternatives
  - large-scale structure: fewer hues
  - fine structure: multiple hues with monotonically increasing luminance [eg viridis R/python]
  - segmented rainbows for binned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

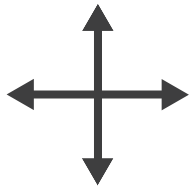
Colormaps

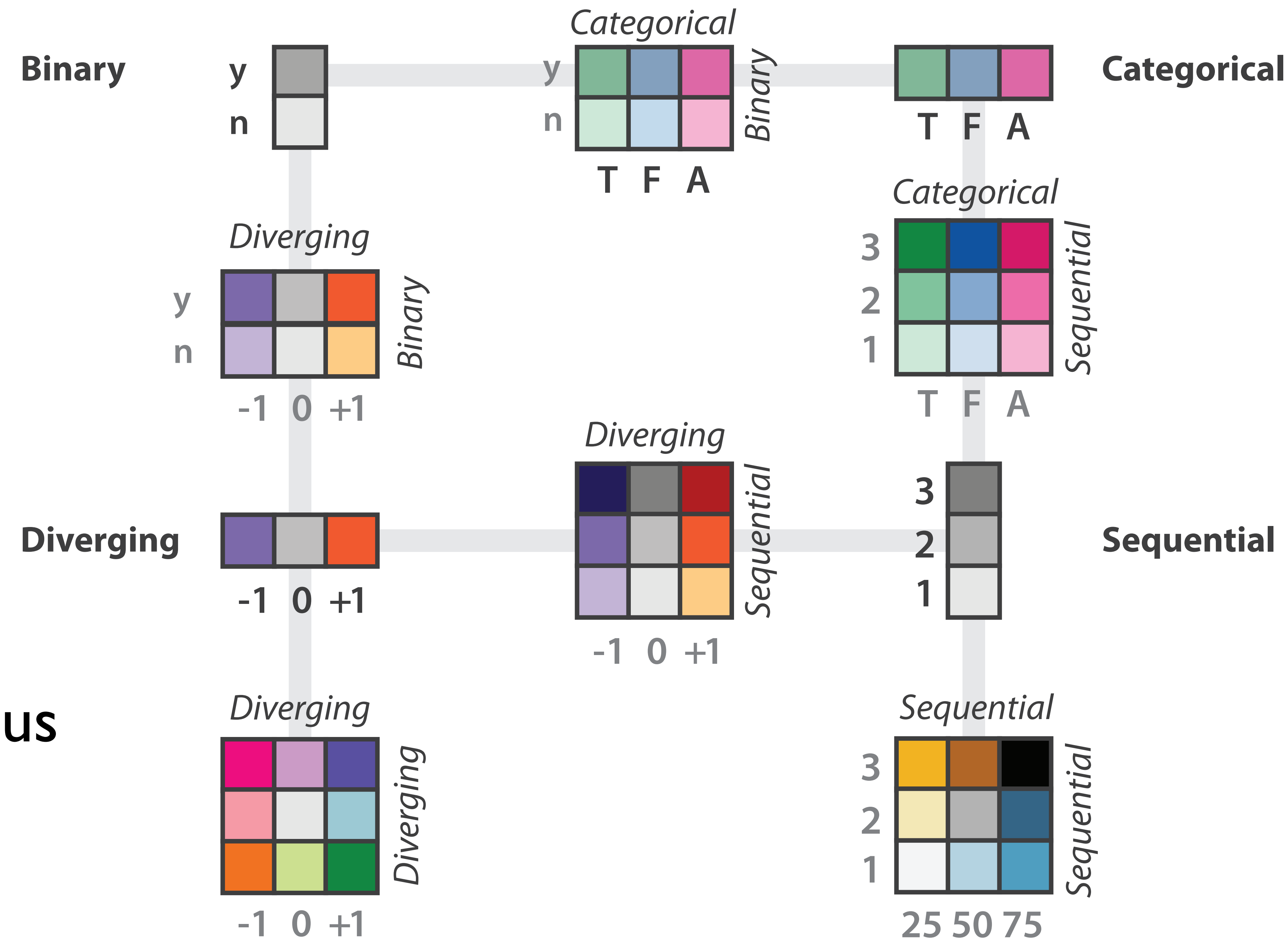
- Categorical

- Ordered
  - *Sequential*

  - *Diverging*

- Bivariate





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

- categorical limits: noncontiguous
  - 6-12 bins hue/color
    - far fewer if colorblind
  - 3-4 bins luminance, saturation
  - size heavily affects salience
    - use high saturation for small regions, low saturation for large

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, Nov 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.

Outline

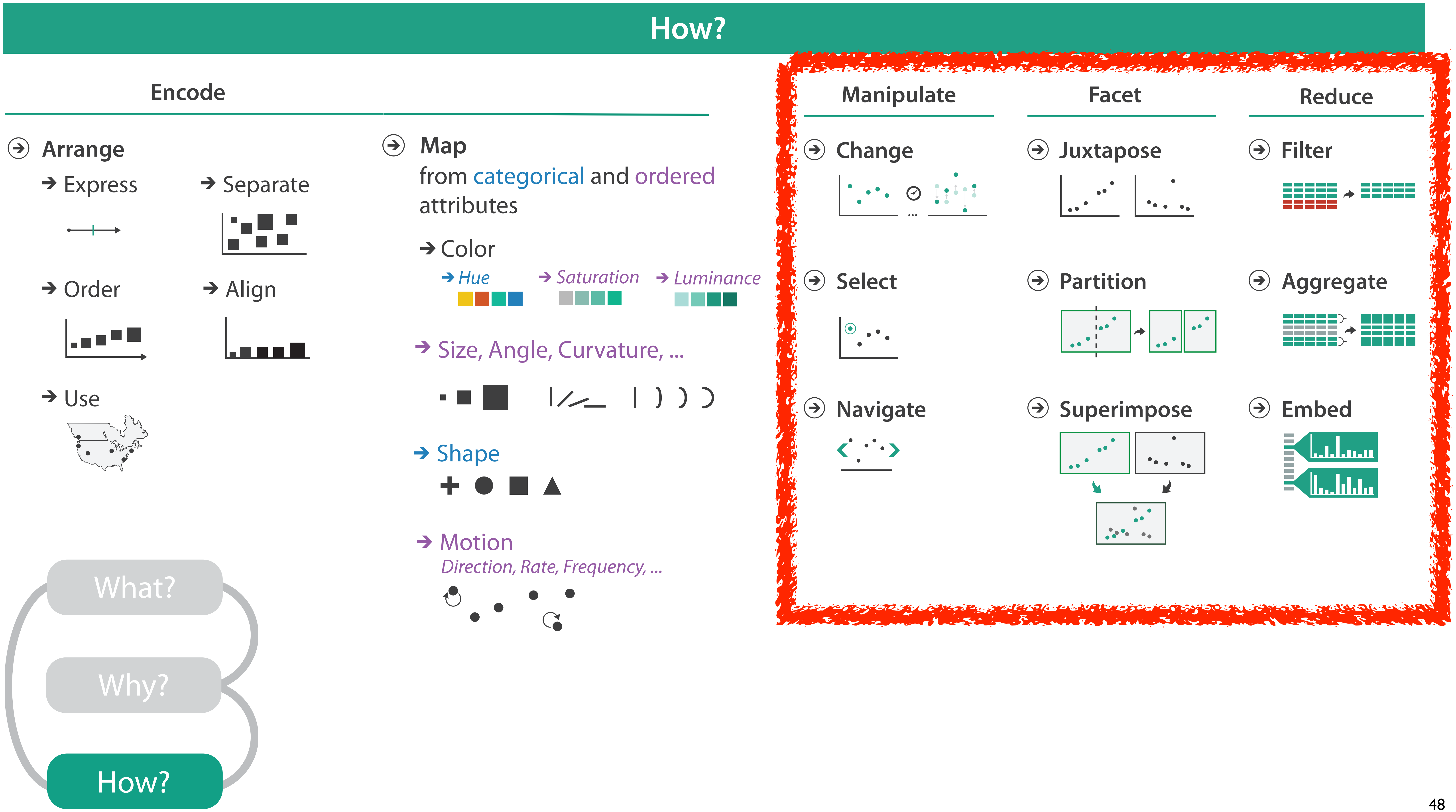
- Session 1: Principles *9:15-10:30am*
  - Analysis: What, Why, How
  - Marks and Channels, Perception
  - Color
- Coffee Break *10:30-10:50am*
- Session 2: Techniques for Scaling *10:50-11:40am*
  - Manipulate: Change, Select, Navigate
  - Facet: Juxtapose, Partition, Superimpose
  - Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

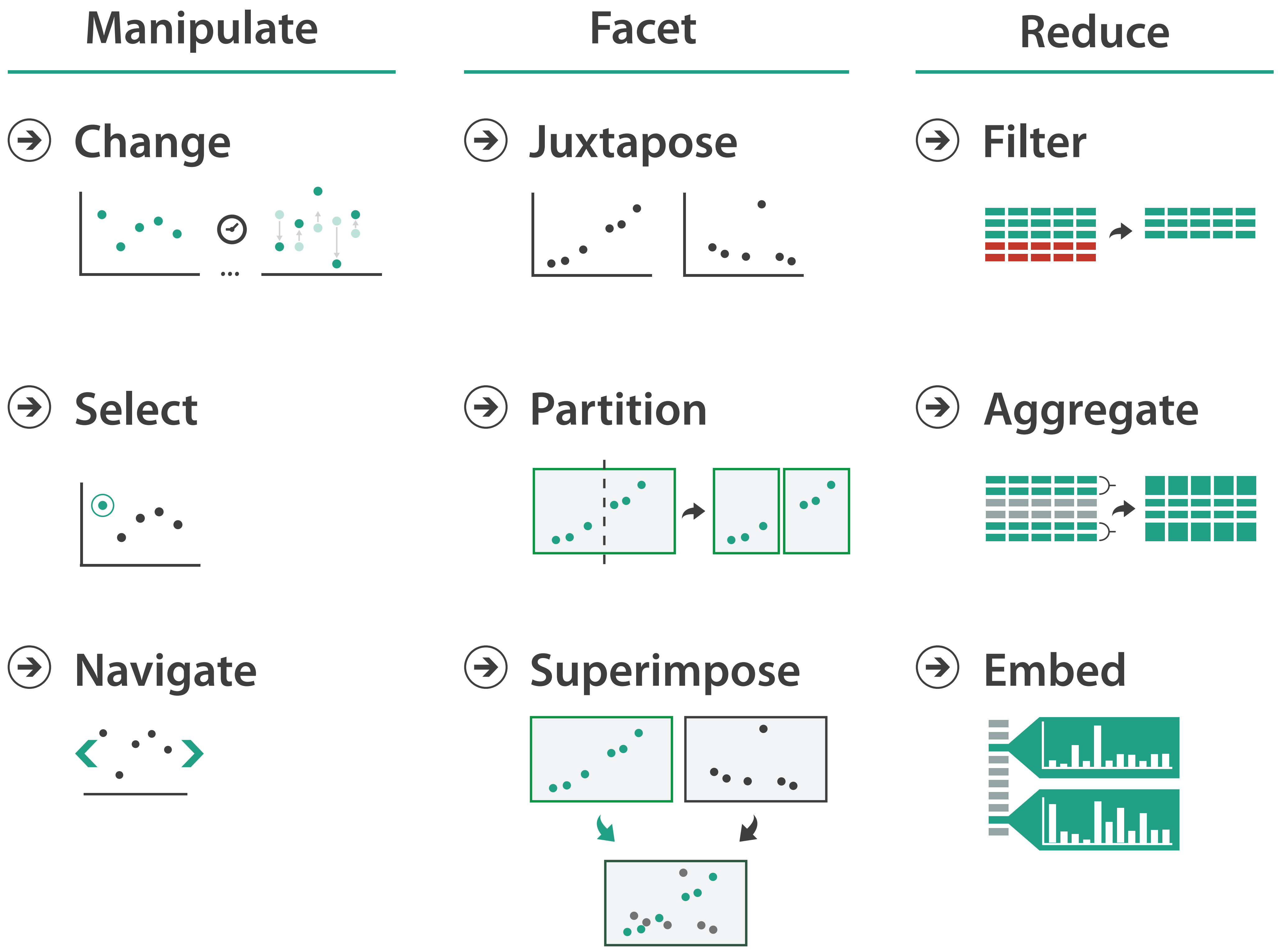
Outline

- **Session 1: Principles** *9:15-10:30am*
  - Analysis: What, Why, How
  - Marks and Channels, Perception
  - Color
- **Session 2: Techniques for Scaling** *10:50-11:40am*
  - Manipulate: Change, Select, Navigate
  - Facet: Juxtapose, Partition, Superimpose
  - Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

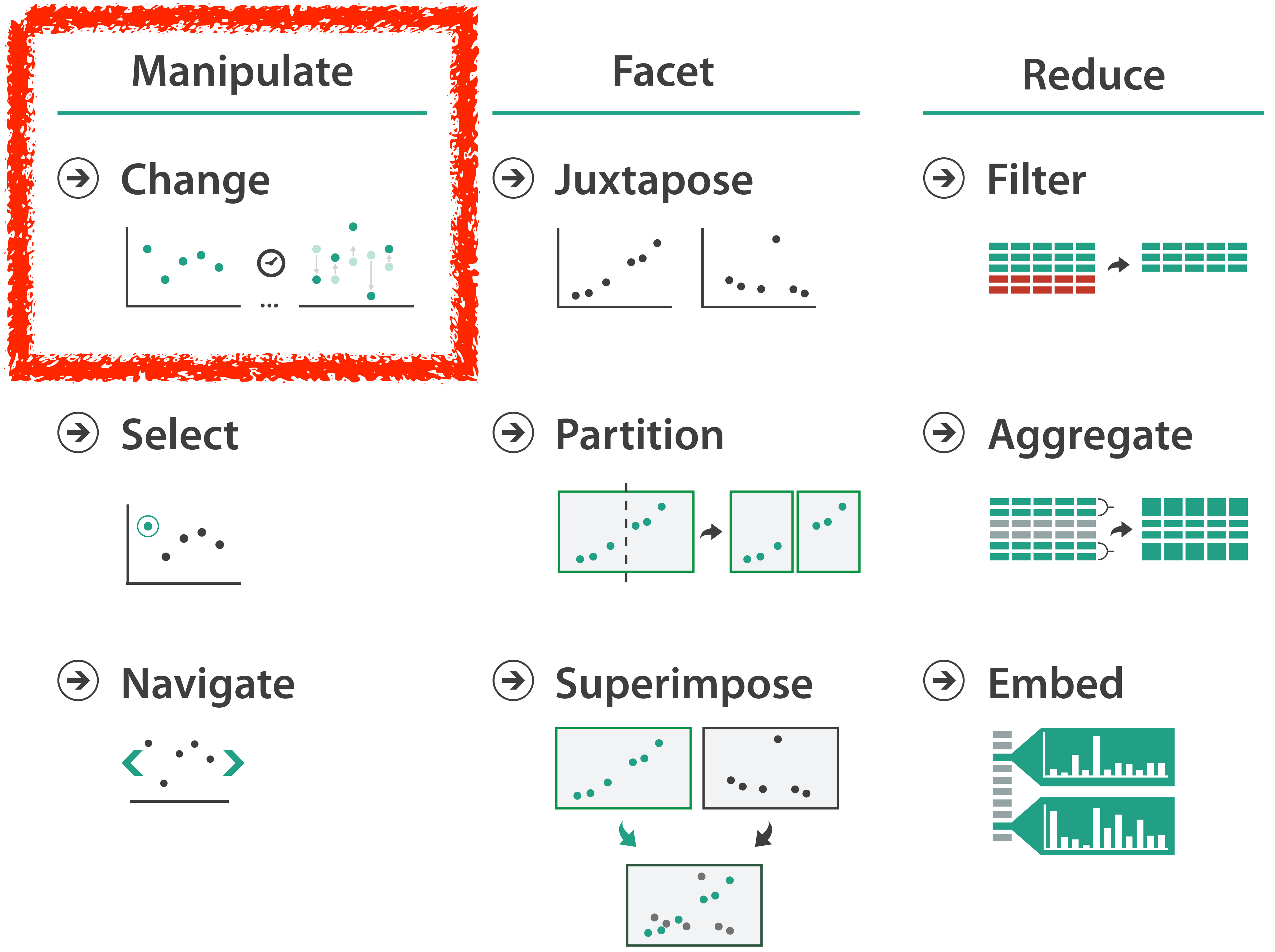

How?
Encode
Arrange
Express
Separate
Order
Align
Use
Map
from categorical and ordered attributes
Color
Hue
Saturation
Luminance
Size, Angle, Curvature, ...
Shape
Motion
Direction, Rate, Frequency, ...
Manipulate
Change
Select
Navigate
Facet
Juxtapose
Partition
Superimpose
Reduce
Filter
Aggregate
Embed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

How to handle complexity: 3 more strategies + 1 previous



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

How to handle complexity: 3 more strategies + 1 previous



- change over time
  - most obvious & flexible of the 4 strategies

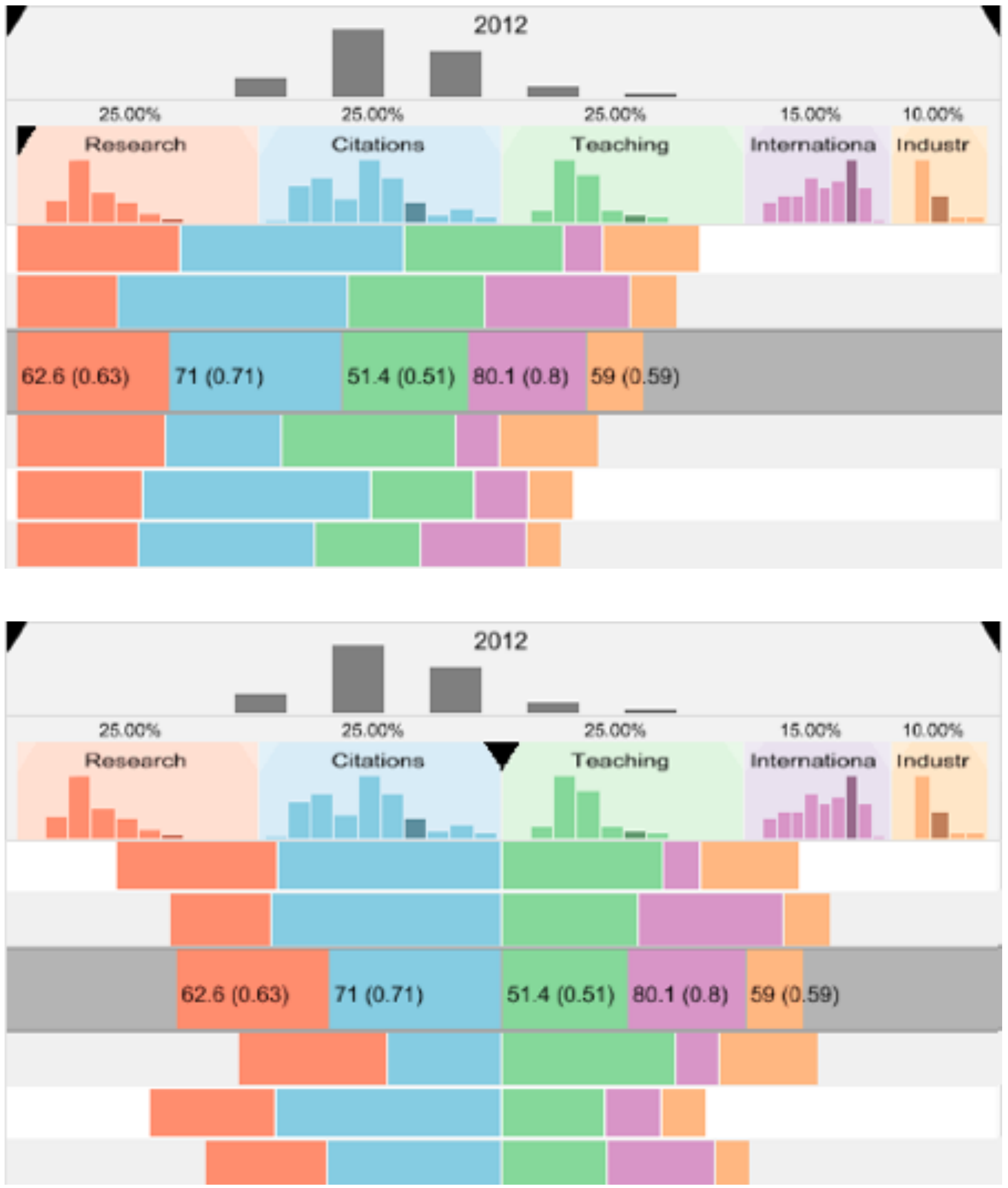
Change over time

- change any of the other choices
  - encoding itself
  - parameters
  - arrange: rearrange, reorder
  - aggregation level, what is filtered...
- why change?
  - one of four major strategies
    - change over time
    - facet data by partitioning into multiple views
    - reduce amount of data shown within view
      - embedding focus + context together
  - most obvious, powerful, flexible
  - interaction entails change

Idiom: **Realign**

System: **LineUp**

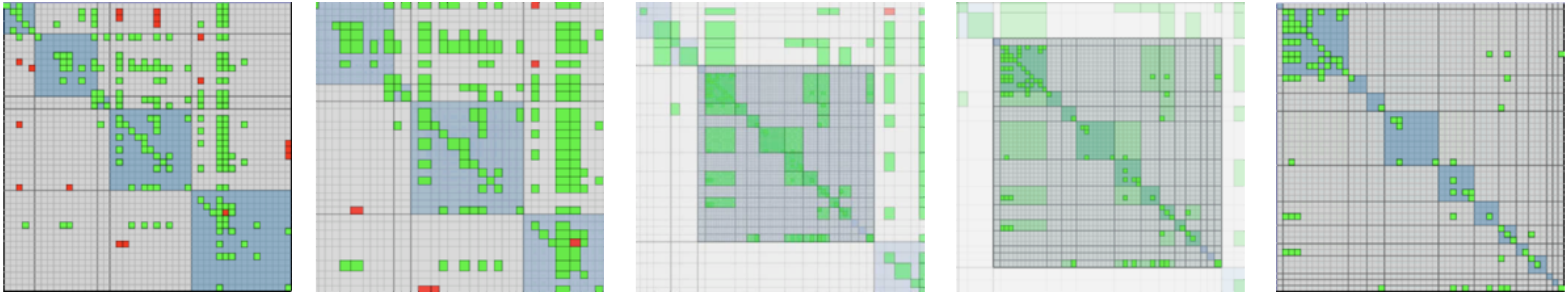
- stacked bars
  - easy to compare
    - first segment
    - total bar
- align to different segment
  - supports flexible comparison



[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.]

Idiom: **Animated transitions**

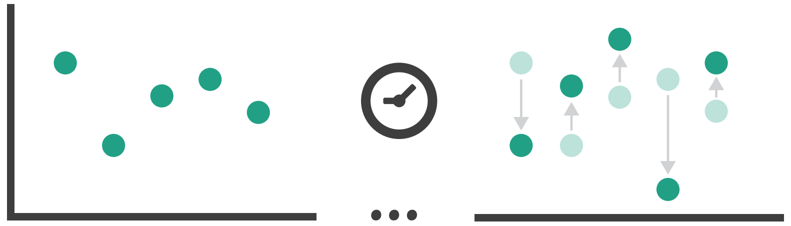
- smooth transition from one state to another
  - alternative to jump cuts
  - support for item tracking when amount of change is limited
- example: multilevel matrix views
- example: animated transitions in statistical data graphics
  - https://vimeo.com/19278444



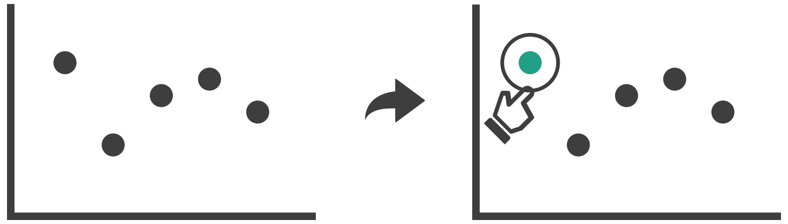
[Using Multilevel Call Matrices in Large Software Projects. van Ham. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 227–232, 2003.]

Manipulate

➔ Change over Time



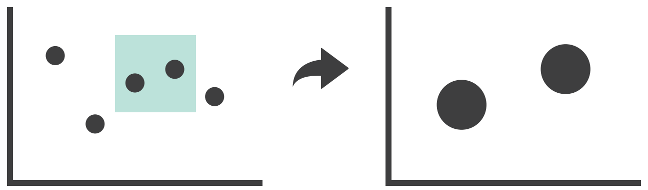
➔ Select



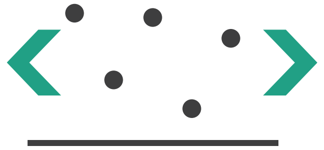
➔ Navigate

➔ Item Reduction

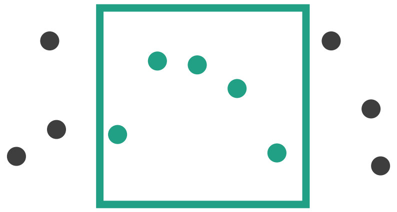
➔ *Zoom*
Geometric or *Semantic*



➔ *Pan/Translate*



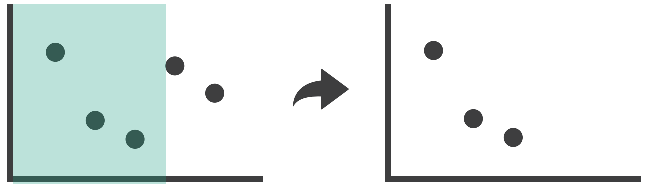
➔ *Constrained*



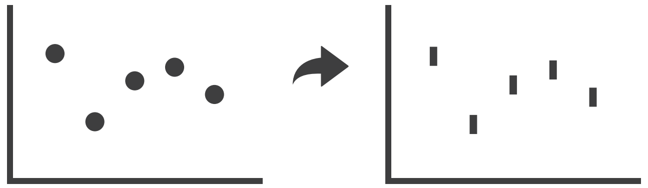
➔ Attribute Reduction

➔ *Slice*

➔ *Cut*



➔ *Project*



Further reading

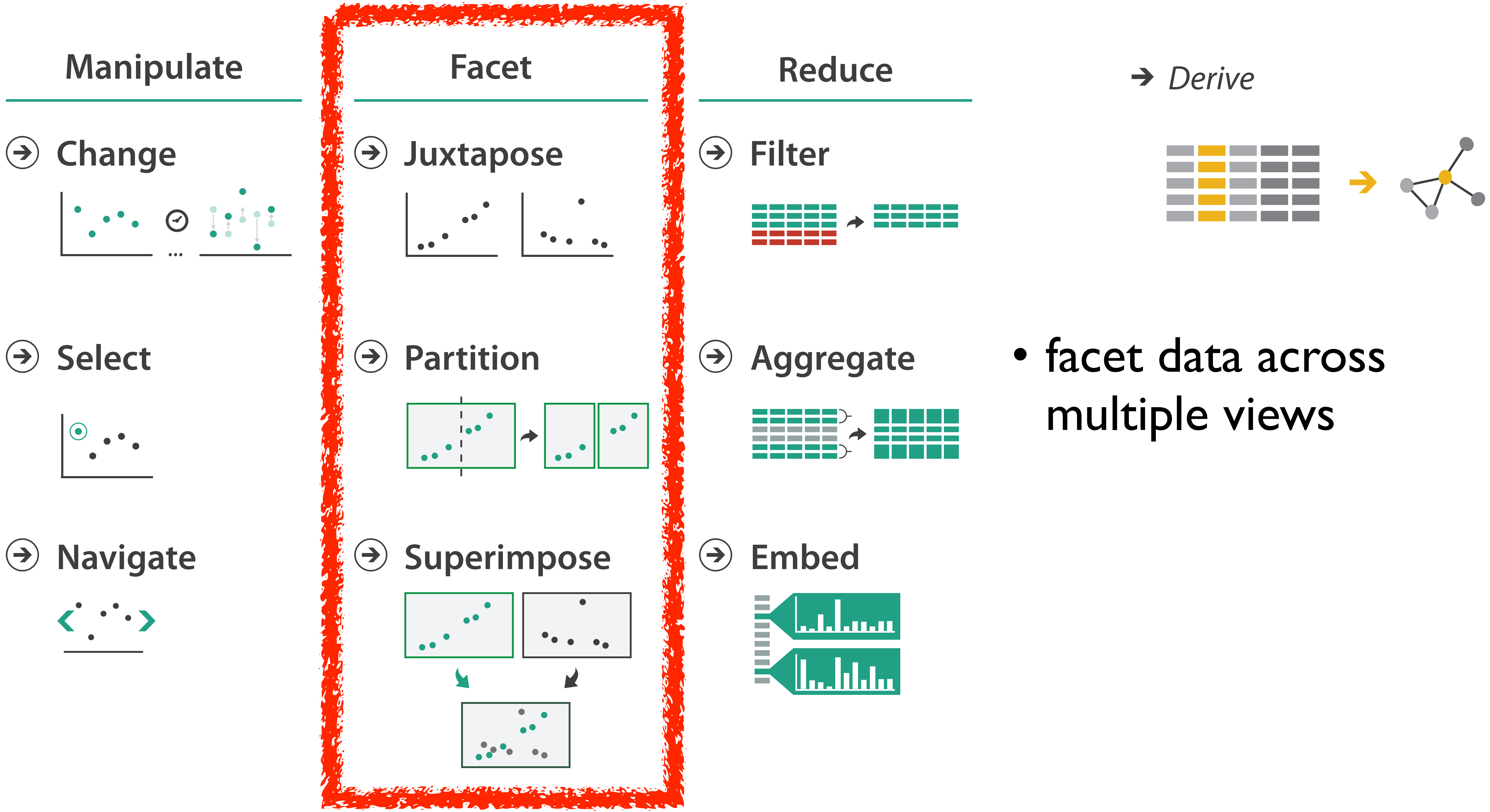
- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, Nov 2014.
  - *Chap 11: Manipulate View*
- *Animated Transitions in Statistical Data Graphics*. Heer and Robertson. IEEE Trans. on Visualization and Computer Graphics (Proc. InfoVis07) 13:6 (2007), 1240–1247.
- *Selection: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

Outline

- **Session 1: Principles** *9:15-10:30am*
  - Analysis: What, Why, How
  - Marks and Channels, Perception
  - Color
- **Session 2: Techniques for Scaling** *10:50-11:40am*
  - Manipulate: Change, Select, Navigate
  - Facet: Juxtapose, Partition, Superimpose
  - Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

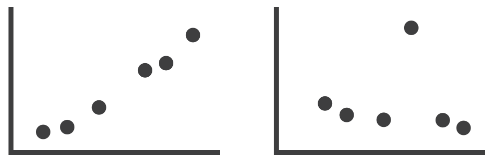
How to handle complexity: 3 more strategies + 1 previous



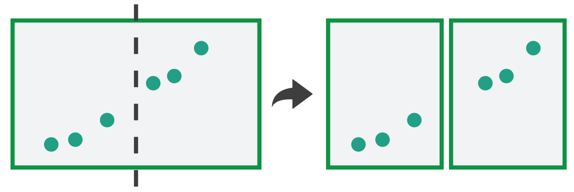
- facet data across multiple views

Facet

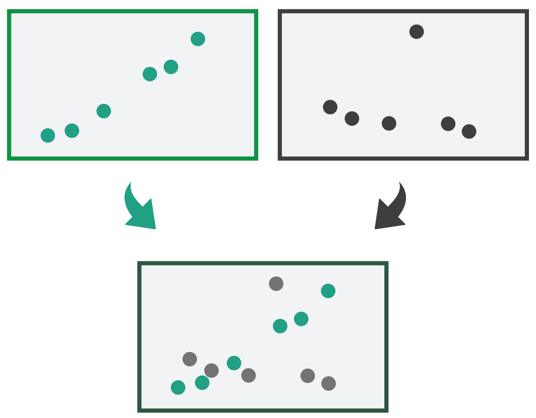
- **Juxtapose**



- **Partition**



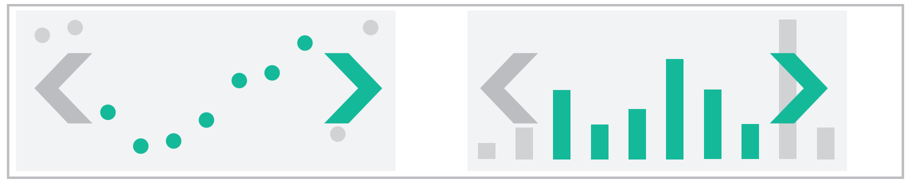
- **Superimpose**



- **Coordinate Multiple Side By Side Views**
  - Share Encoding: Same/Different
    - *Linked Highlighting*

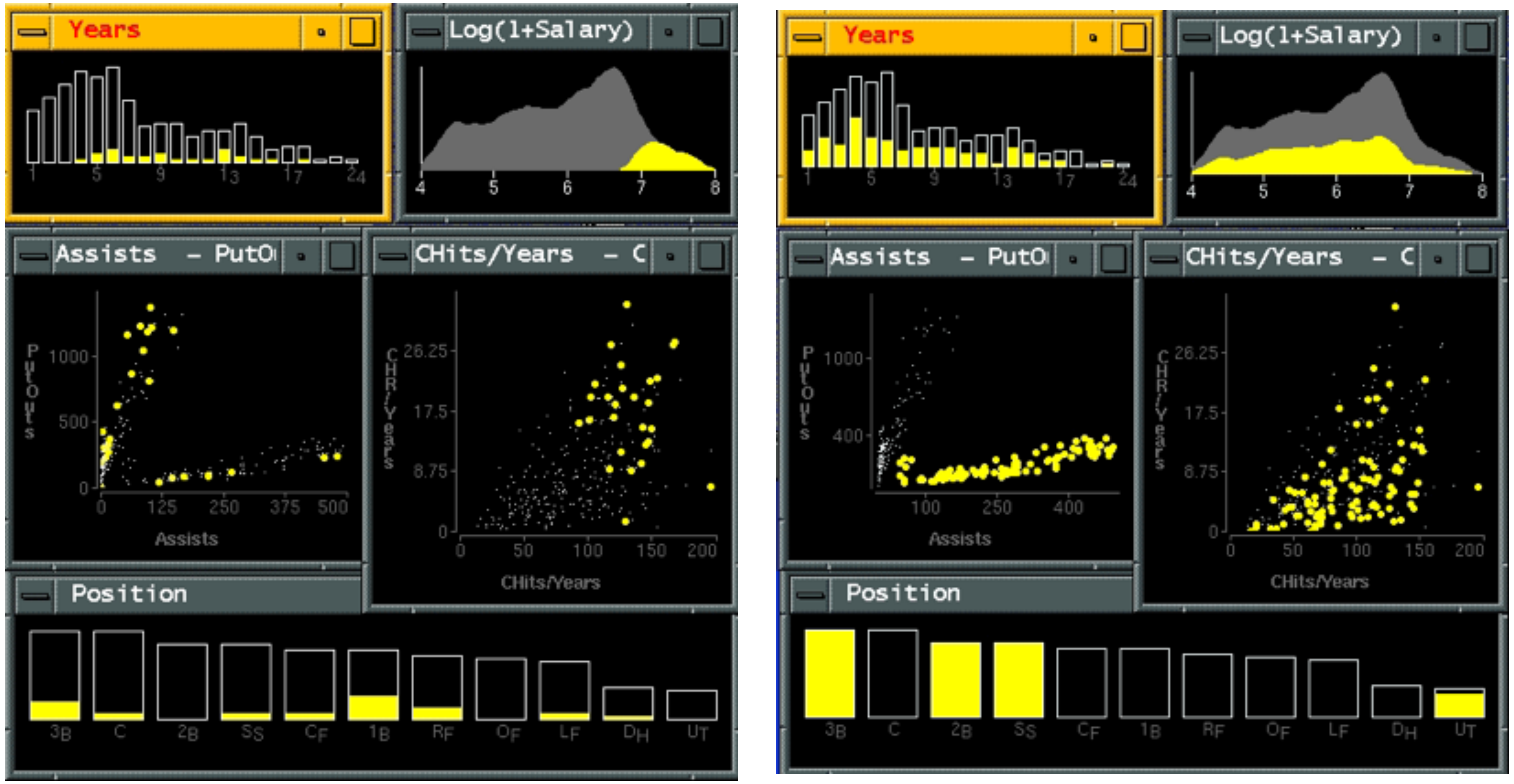
  - Share Data: All/Subset/None

  - Share Navigation



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

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



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

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

System: **Google Maps**

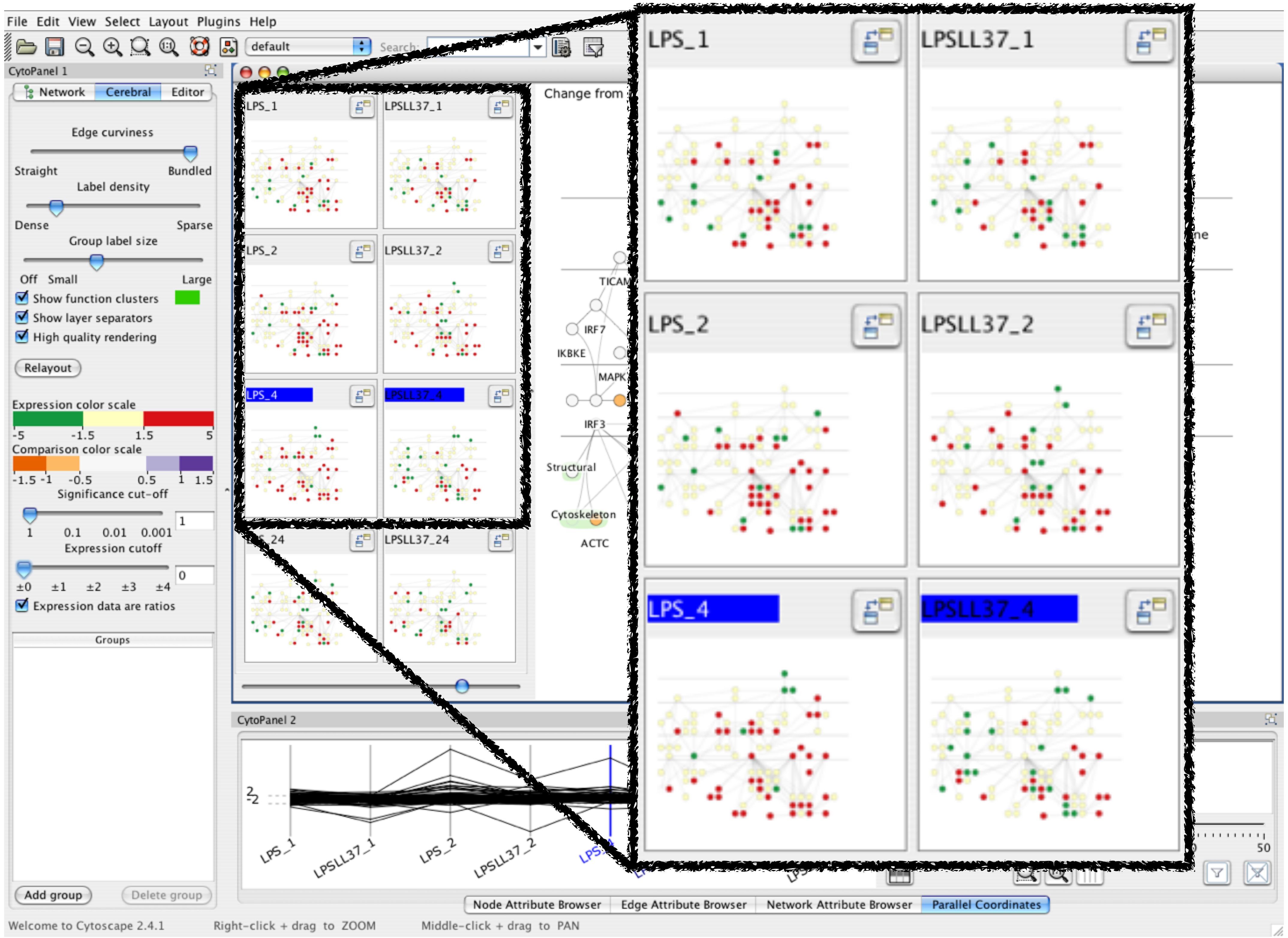
- encoding: same
- data: subset shared
- navigation: shared
  - bidirectional linking
- differences
  - viewpoint
  - (size)
- ***overview-detail***

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

Idiom: **Small multiples**

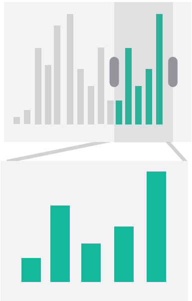
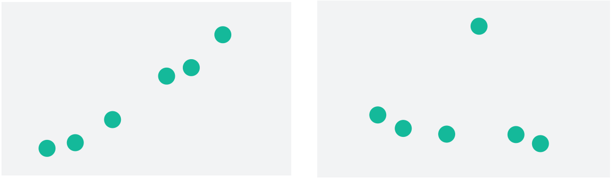
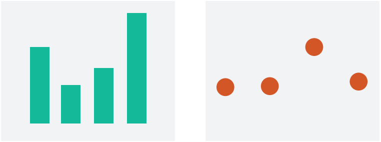
System: **Cerebral**

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



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

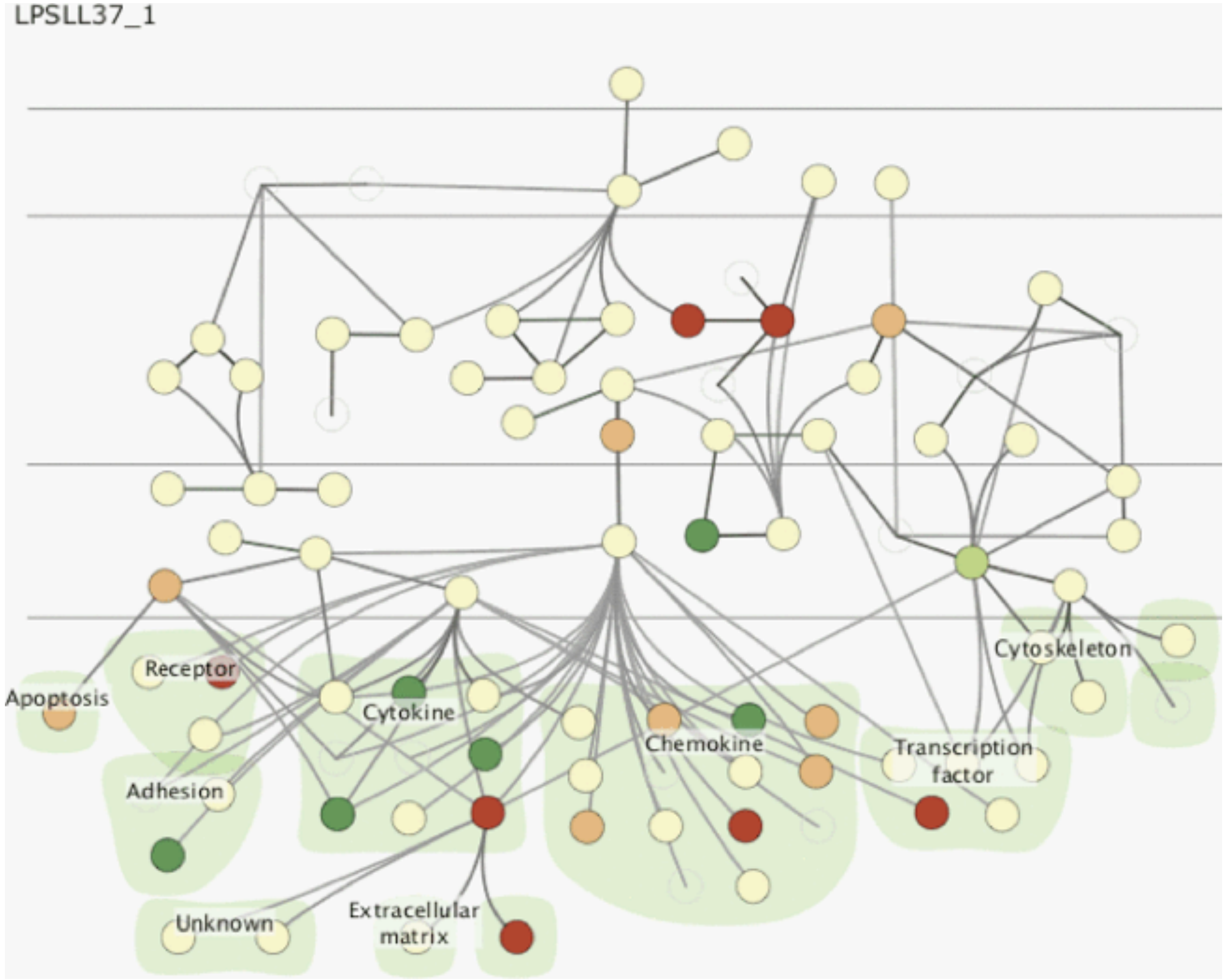
Coordinate views: Design choice interaction

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

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

Idiom: Animation (change over time)

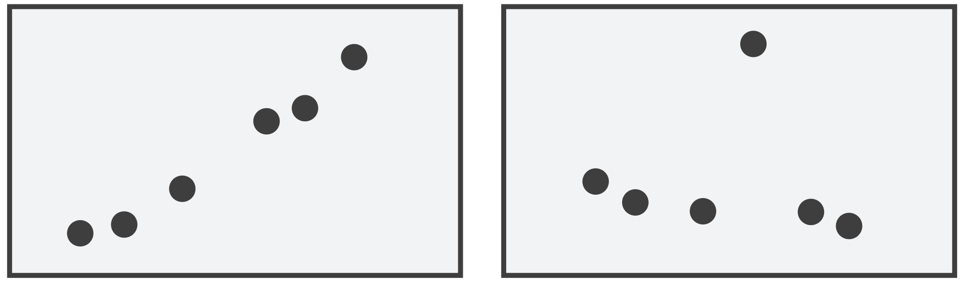
- weaknesses
  - widespread changes
  - disparate frames
- strengths
  - choreographed storytelling
  - localized differences between contiguous frames
  - animated transitions between states



Partition into views

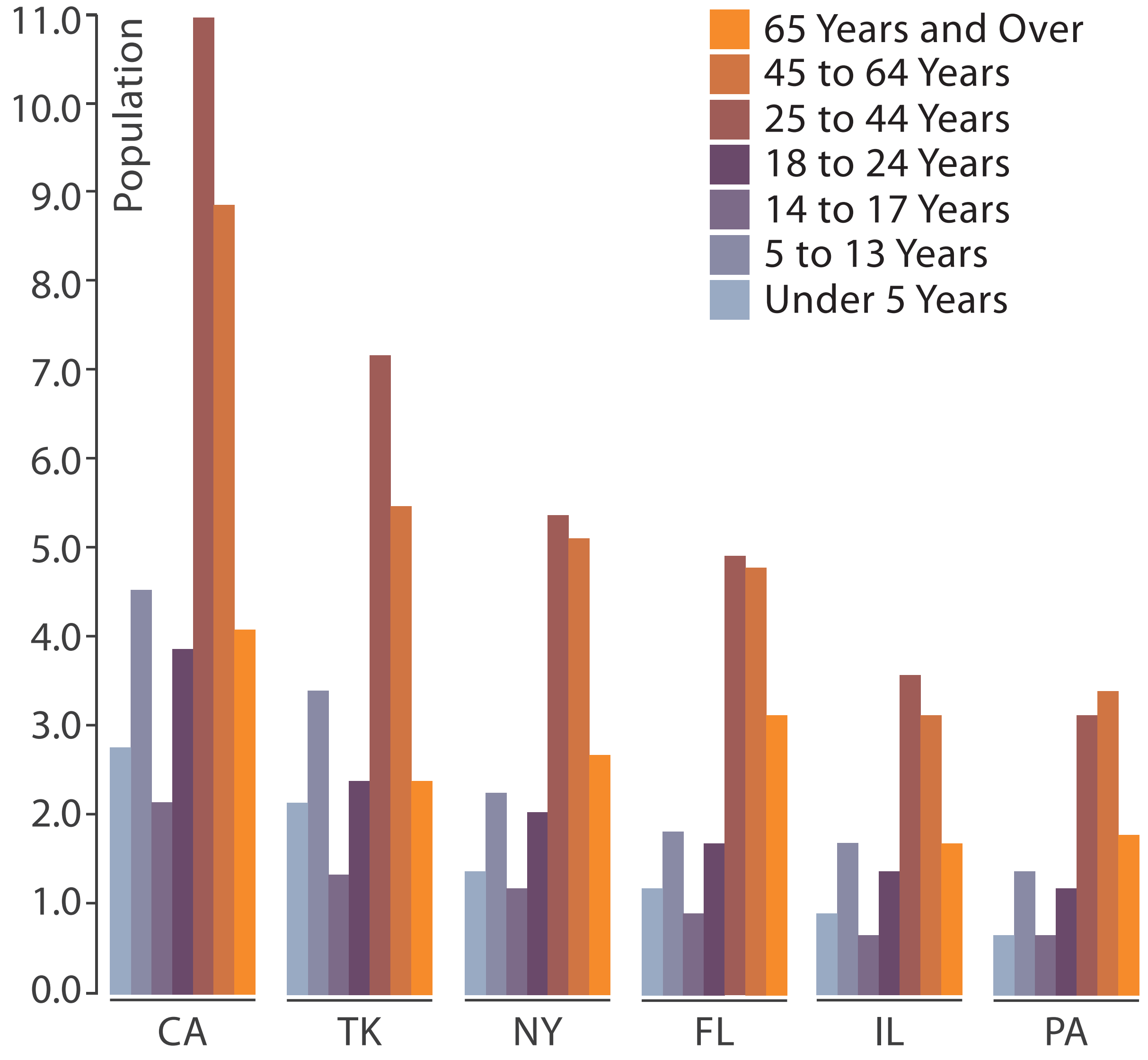
- how to divide data between views
  - encodes association between items using spatial proximity
  - major implications for what patterns are visible
  - split according to attributes
- design choices
  - how many splits
    - all the way down: one mark per region?
    - stop earlier, for more complex structure within region?
  - order in which attribs used to split

Partition into Side-by-Side Views

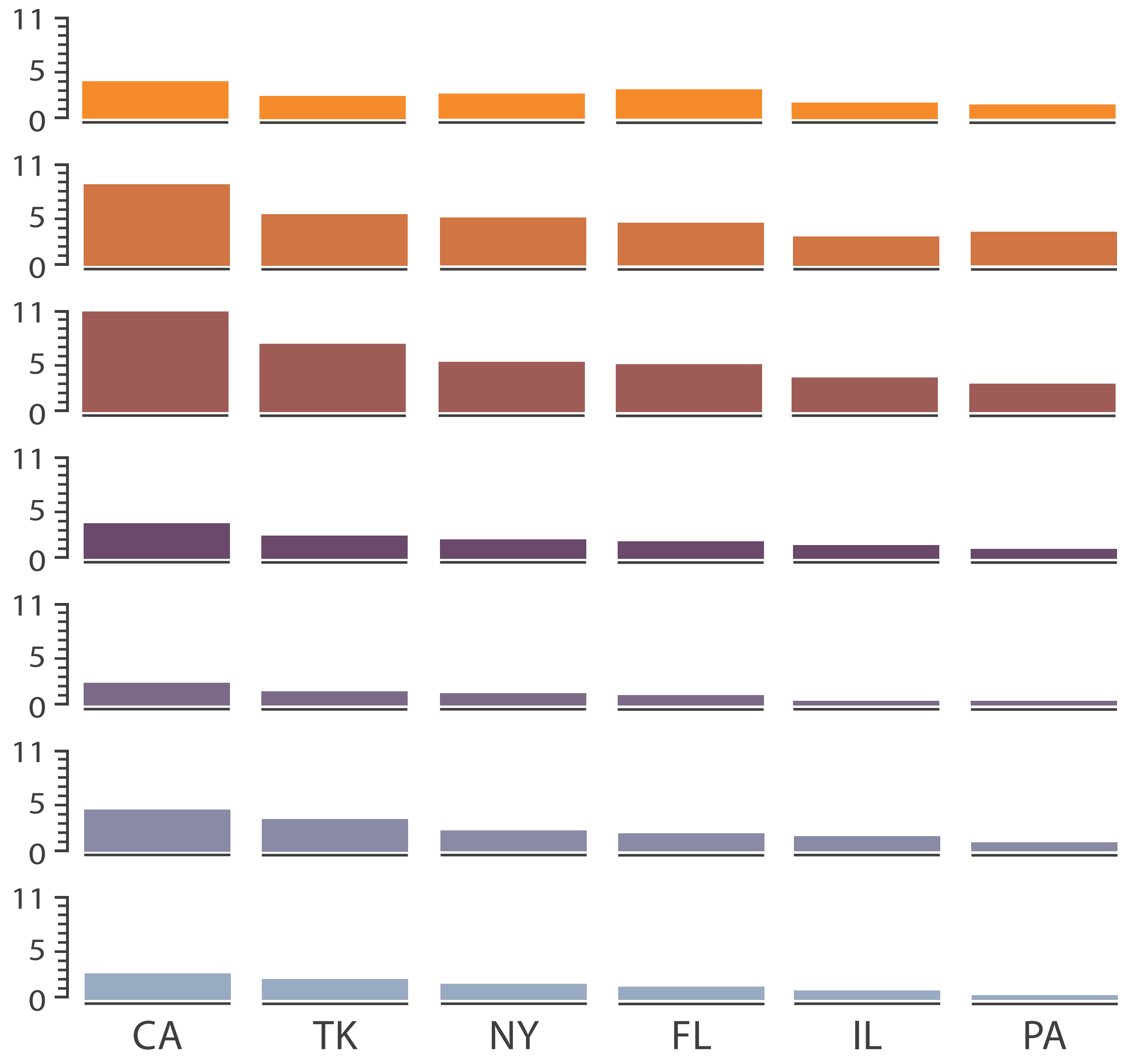


Partitioning: List alignment

- 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**

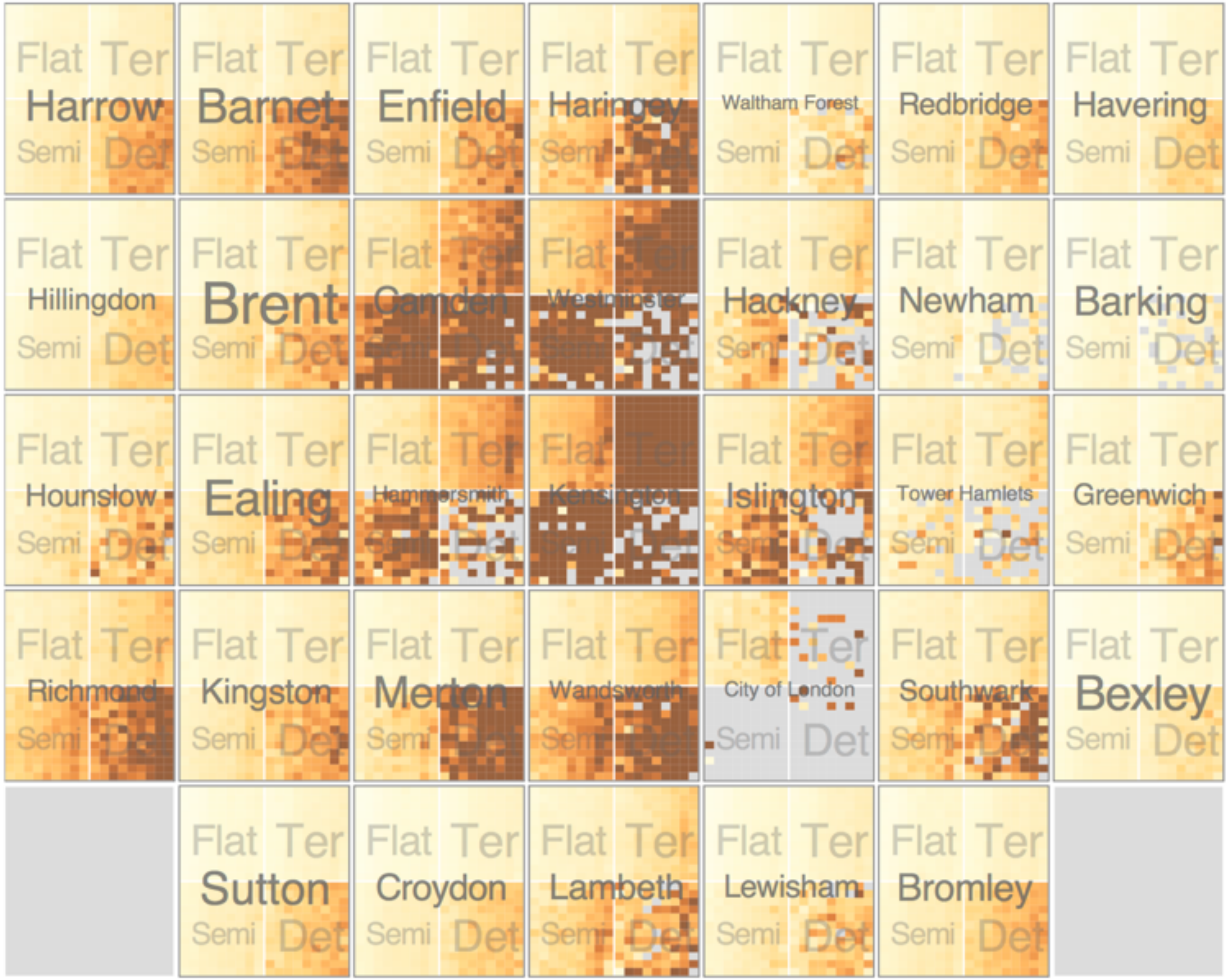
- split by type
- then by neighborhood
- then time
  - –years as rows
  - –months as columns

[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
  - neighborhood then type
- very different patterns

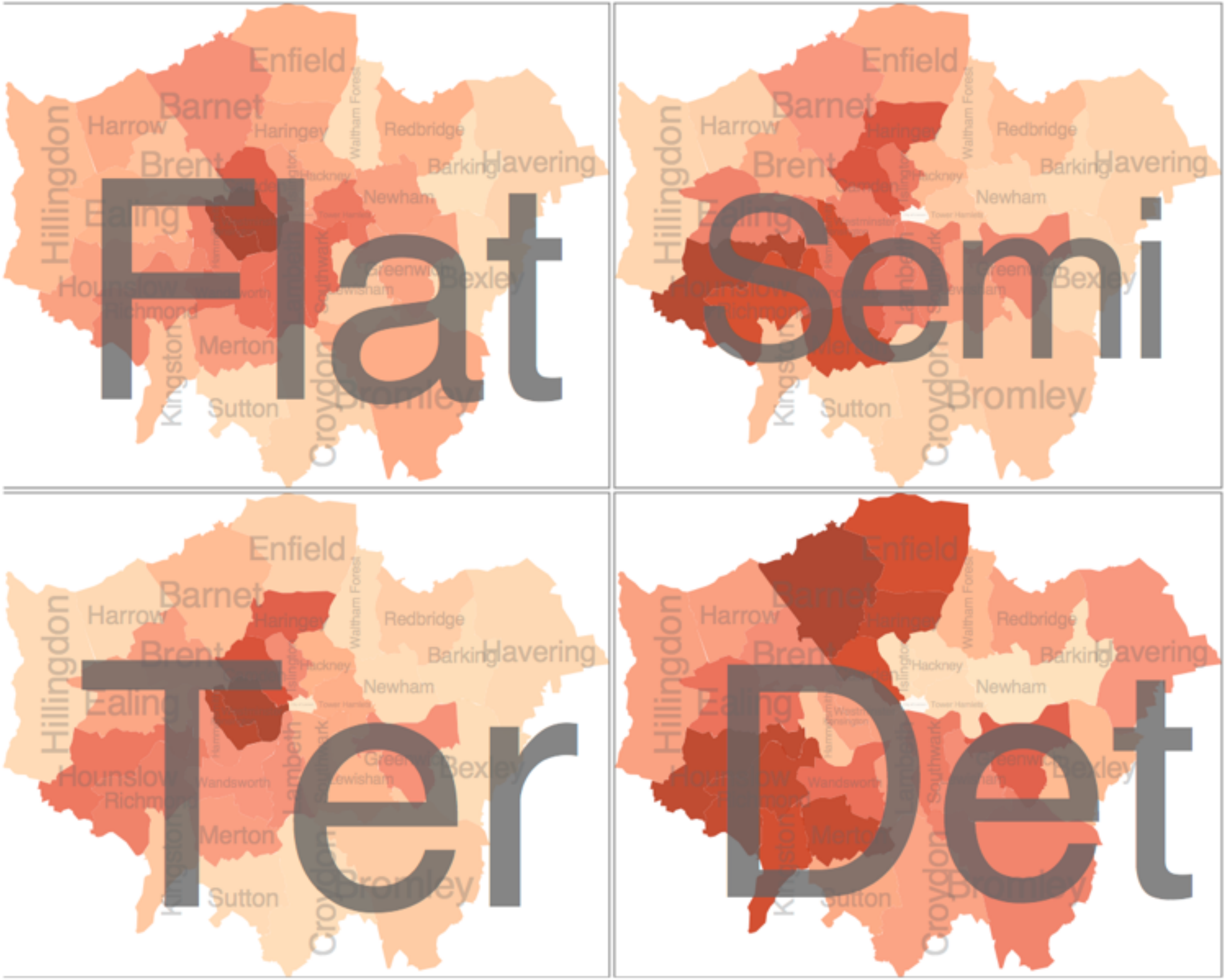


[*Configuring Hierarchical Layouts to Address Research Questions.* Slingsby, Dykes, and Wood. *IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2009) 15:6 (2009), 977–984.*]

Partitioning: Recursive subdivision

System: **HIVE**

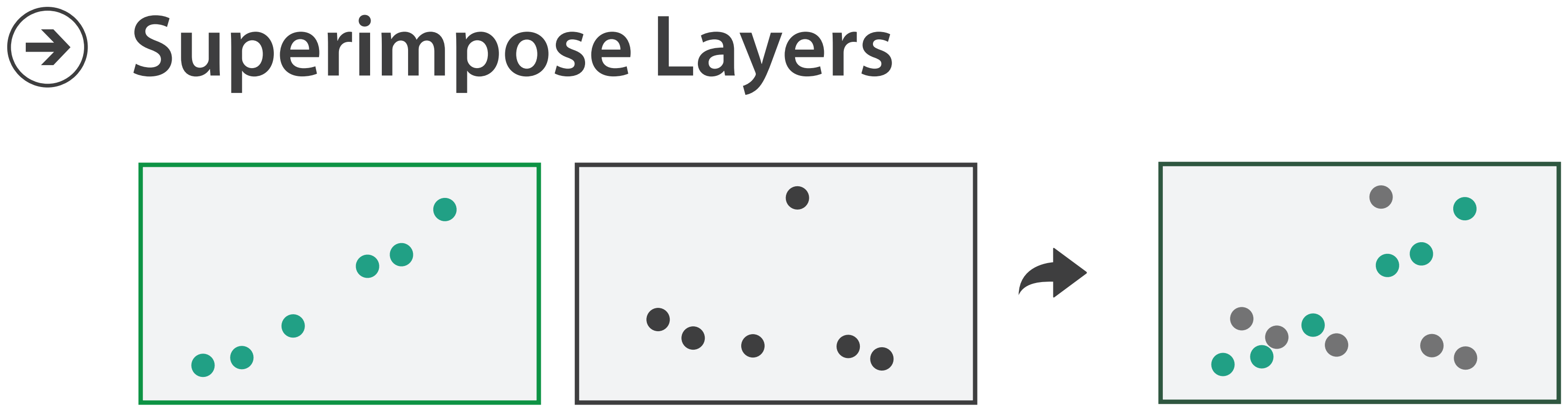
- different encoding for second-level regions
  - choropleth maps



[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.]

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 are layers distinguished?
  - small static set or dynamic from many possible?
  - how partitioned?
    - heavyweight with attribs vs lightweight with selection
- distinguishable layers
  - encode with different, nonoverlapping channels
    - two layers achieveable, three with careful design



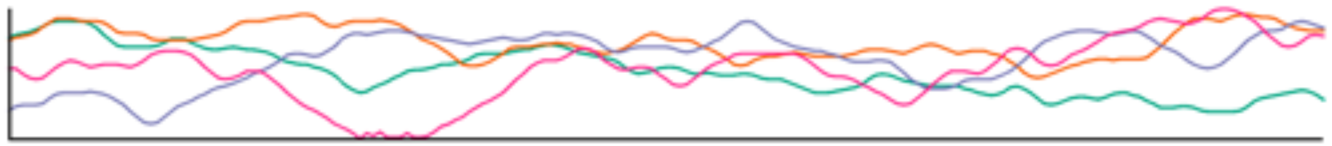
Static visual layering

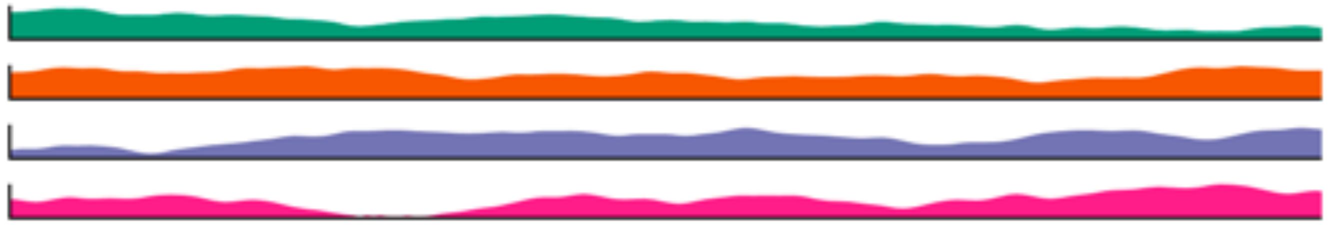
- foreground layer: roads
  - hue, size distinguishing main from minor
  - high luminance contrast from background
- background layer: regions
  - desaturated colors for water, parks, land areas
- user can selectively focus attention
- "get it right in black and white"
  - check luminance contrast with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

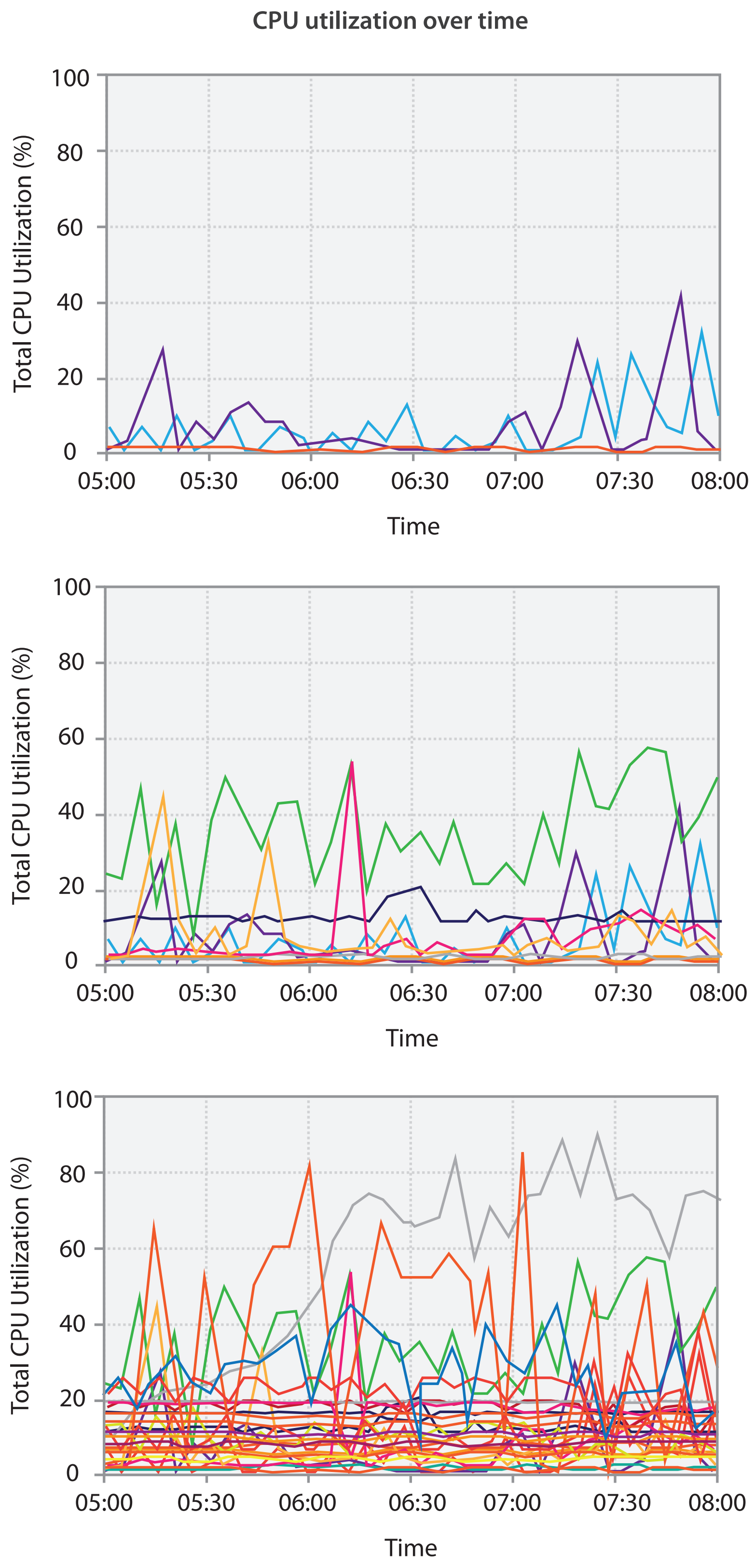
Superimposing limits

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





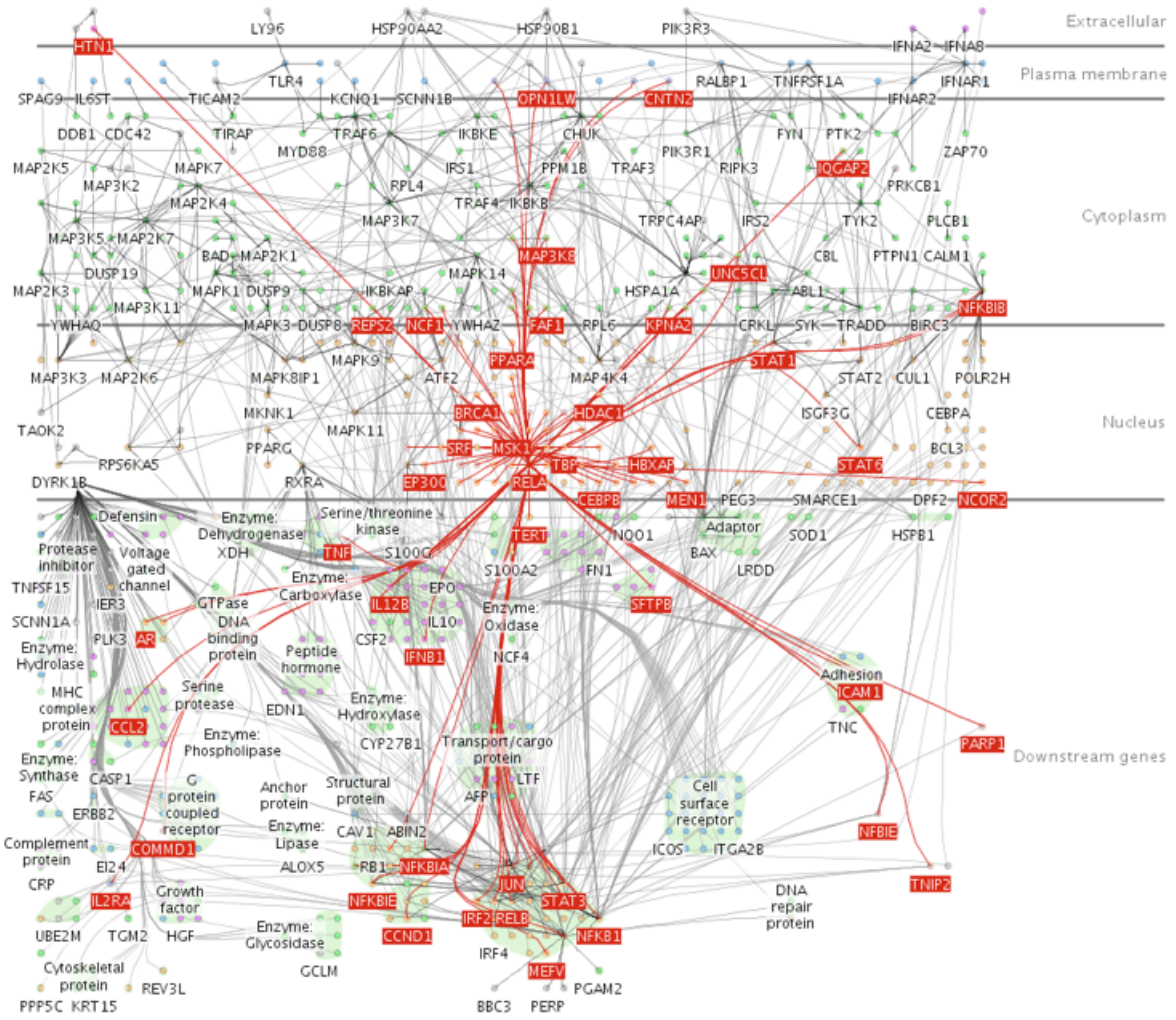
[Graphical Perception of Multiple Time Series. Javed, McDonnel, and Elmqvist. IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE InfoVis 2010) 16:6 (2010), 927–934.]



Dynamic visual layering

System: **Cerebral**

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



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

Further reading

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

Outline

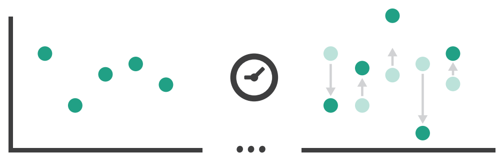
- **Session 1: Principles** *9:15-10:30am*
  - Analysis: What, Why, How
  - Marks and Channels, Perception
  - Color
- **Session 2: Techniques for Scaling** *10:50-11:40am*
  - Manipulate: Change, Select, Navigate
  - Facet: Juxtapose, Partition, Superimpose
  - Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

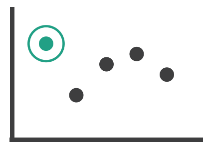
How to handle complexity: 3 more strategies + 1 previous

Manipulate

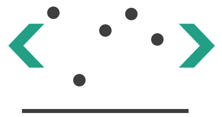
→ Change



→ Select

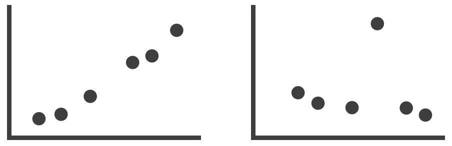


→ Navigate

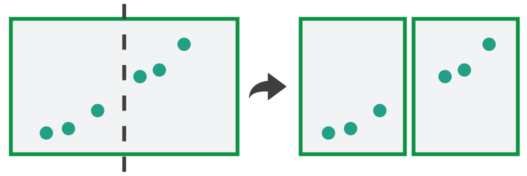


Facet

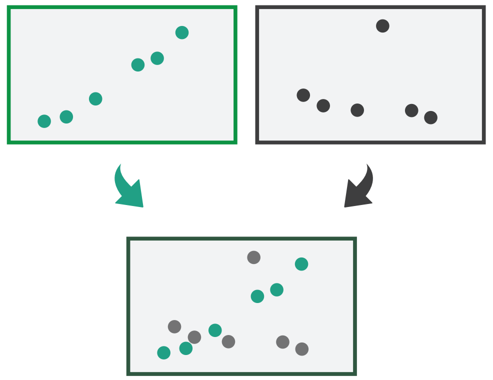
→ Juxtapose



→ Partition

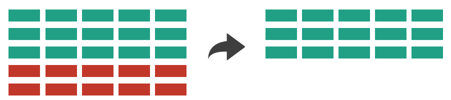


→ Superimpose

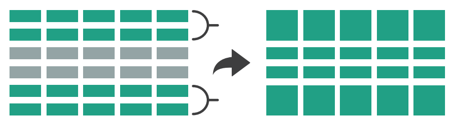


Reduce

→ Filter

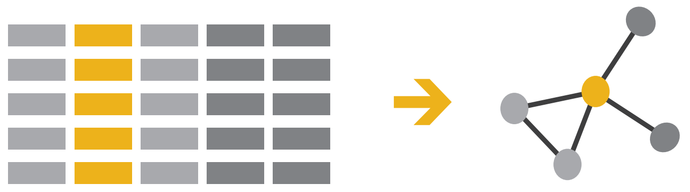


→ Aggregate



→ Embed

→ *Derive*

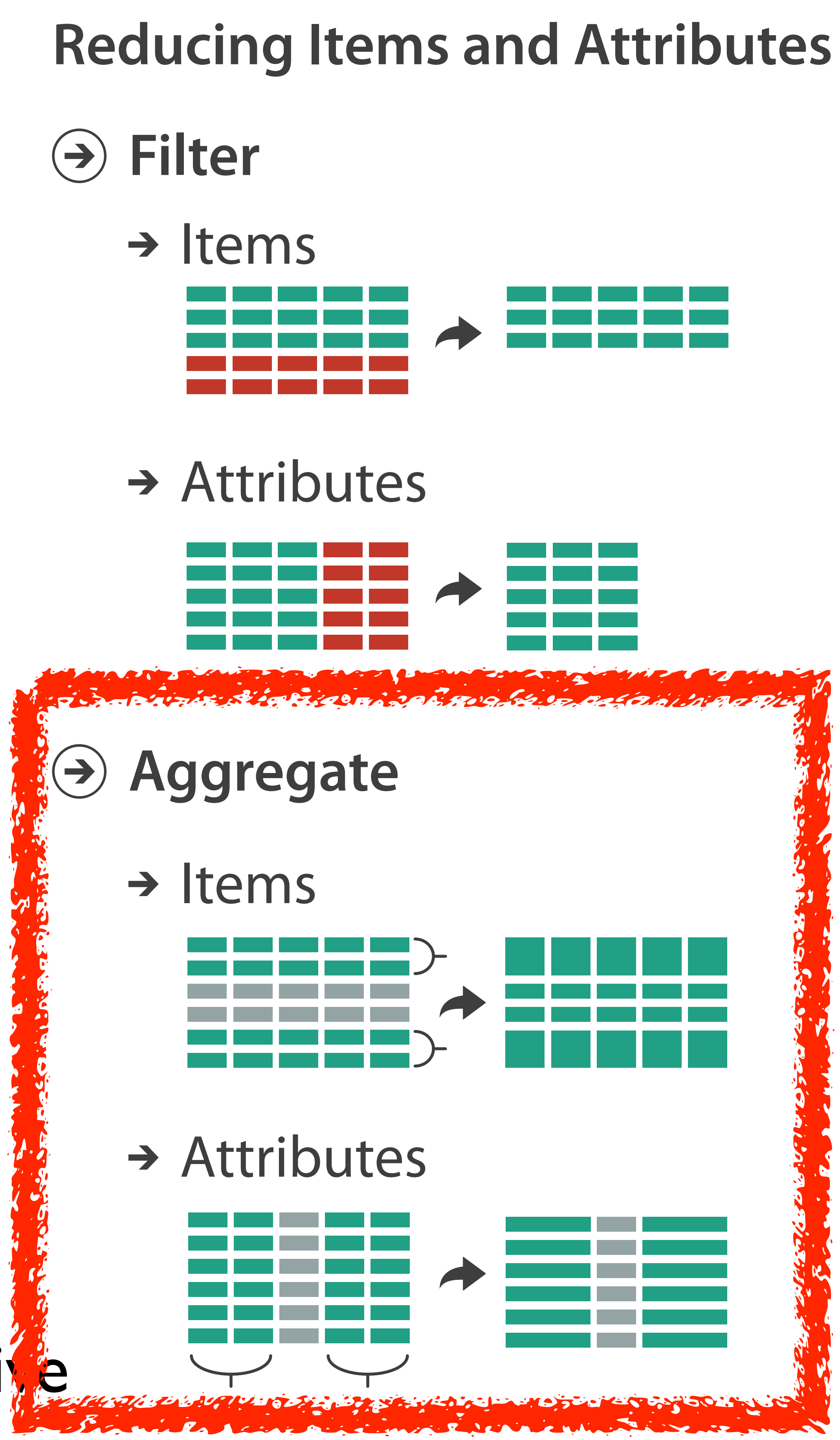


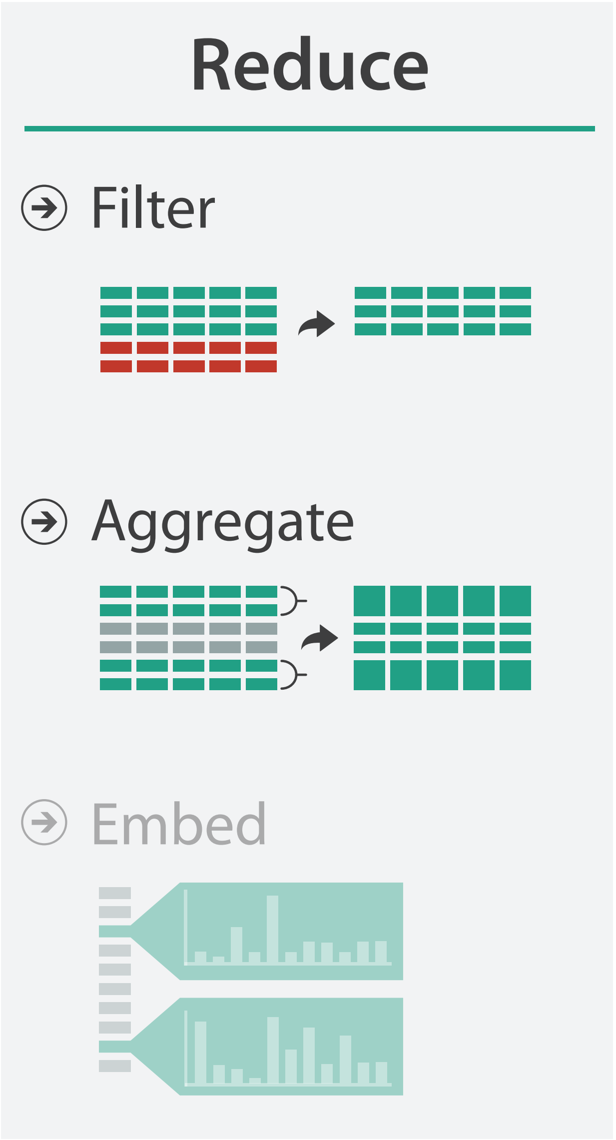
- reduce what is shown within single view

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

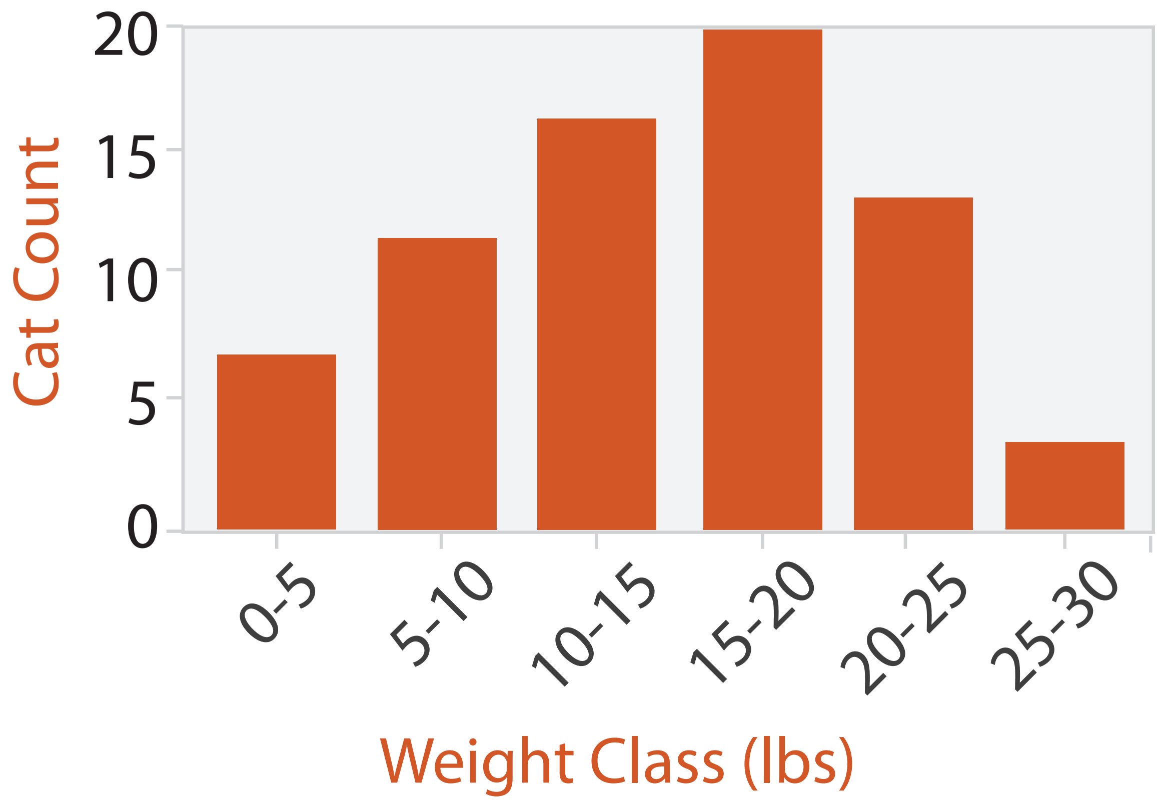
- not mutually exclusive
  - –combine filter, aggregate
  - –combine reduce, facet, change, derive





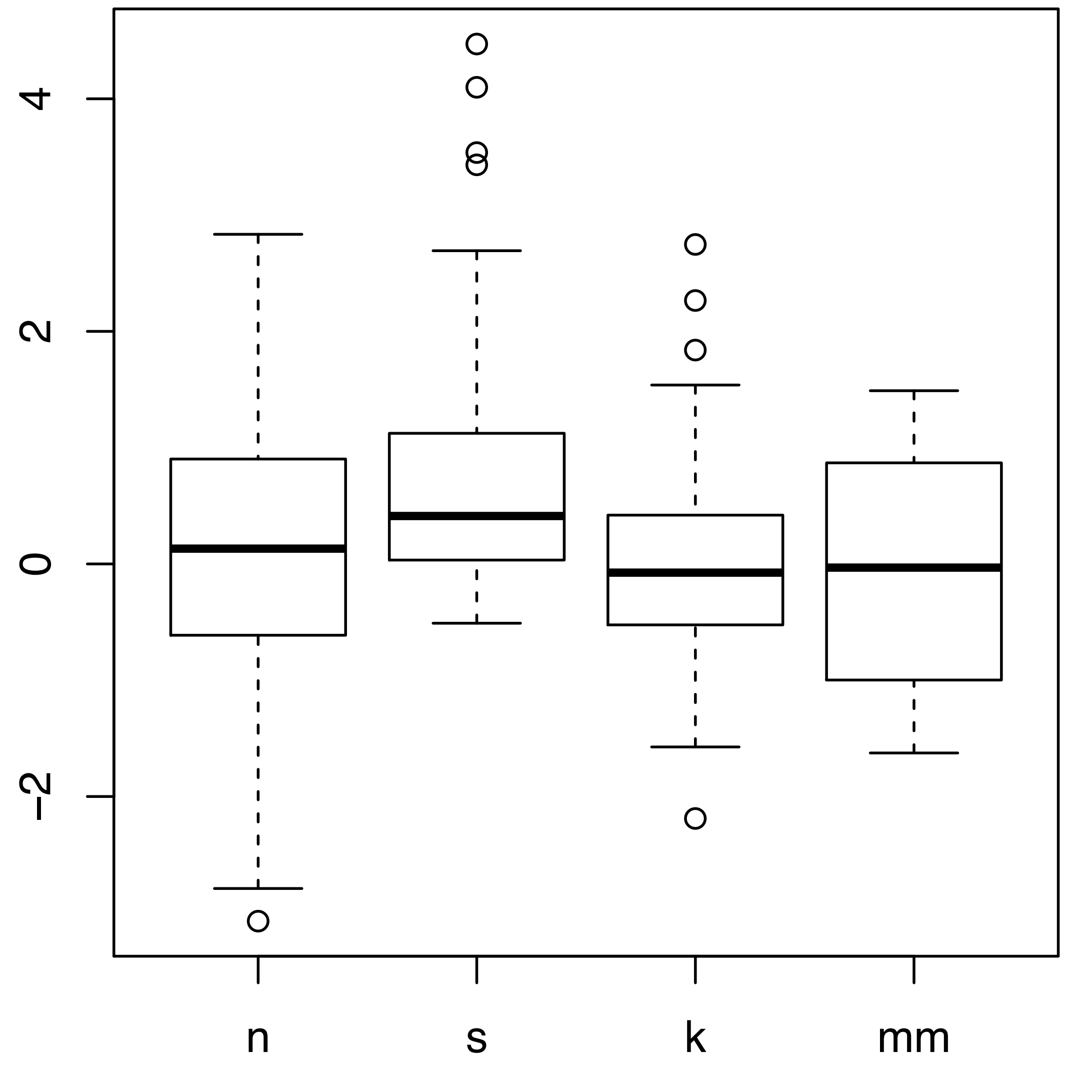
Idiom: **histogram**

- 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: **boxplot**

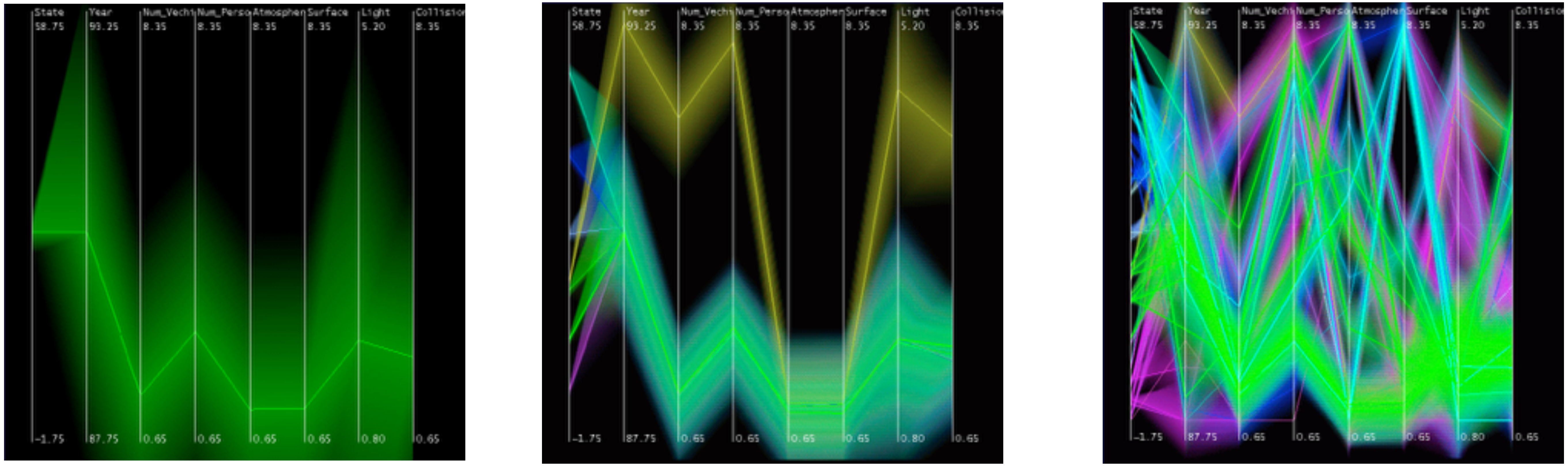
- static item aggregation
- task: find distribution
- data: table
- derived data
  - 5 quant attribs
    - median: central line
    - lower and upper quartile: boxes
    - lower upper fences: whiskers
      - values beyond which items are outliers
  - outliers beyond fence cutoffs explicitly shown



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

Idiom: **Hierarchical parallel coordinates**

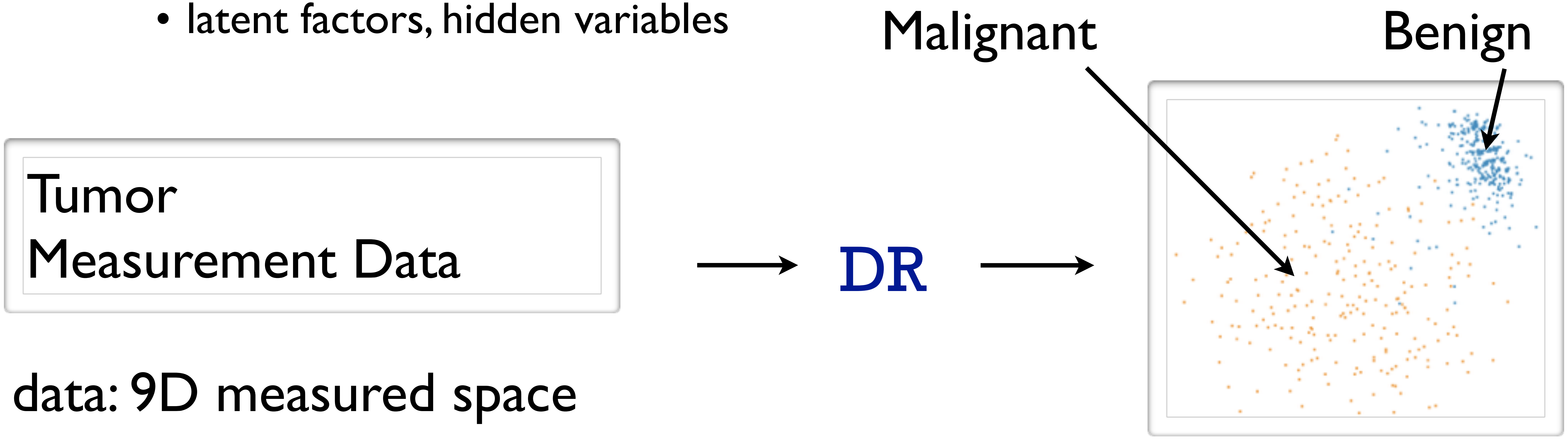
- dynamic item aggregation
- derived data: *hierarchical clustering*
- encoding:
  - cluster band with variable transparency, line at mean, width by min/max values
  - color by proximity in hierarchy



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

Dimensionality reduction

- attribute aggregation
  - derive low-dimensional target space from high-dimensional measured space
  - 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

Idiom: Dimensionality reduction for documents

Task 1

	Dim 1	Dim ...	Dim n
Item 1			
Item ...			
Item n			

	Dim 1	Dim 2
Item 1		
Item ...		
Item n		

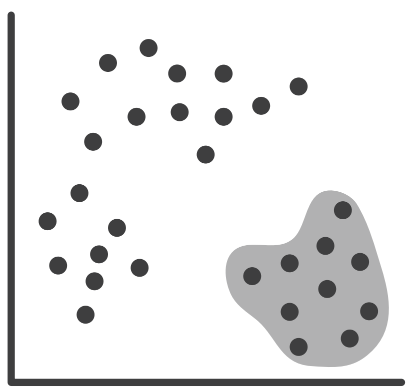
In
HD data → Out
2D data

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

Why?
- Produce
- Derive

Task 2

	Dim 1	Dim 2
Item 1		
Item ...		
Item n		



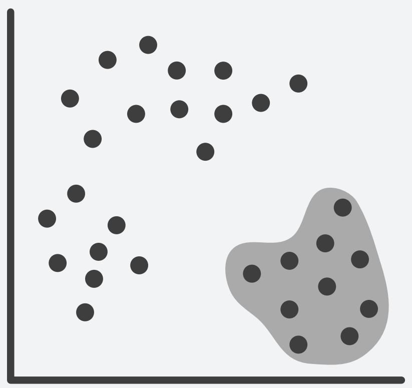
In
2D data → Out
Scatterplot
Clusters & points

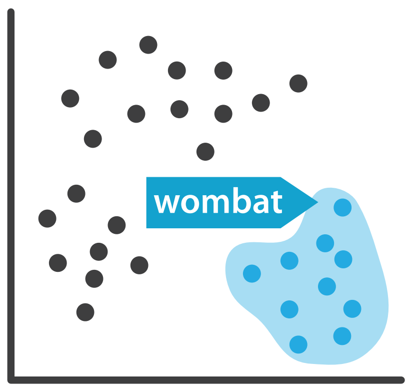
What?
- In 2D data
- Out Scatterplot
- Out Clusters & points

Why?
- Discover
- Explore
- Identify

How?
- Encode
- Navigate
- Select

Task 3





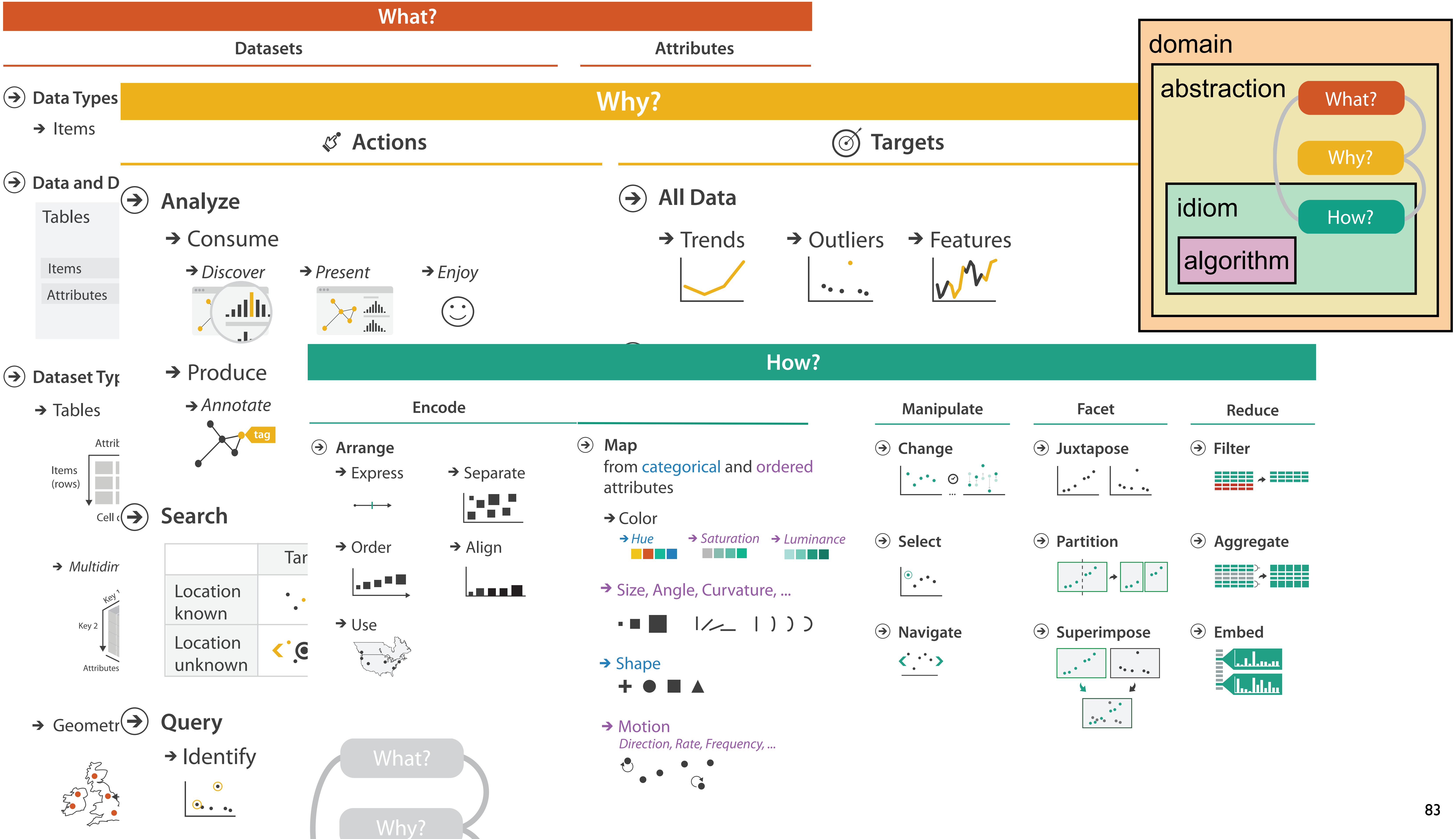
In
Scatterplot
Clusters & points → Out
Labels for clusters

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

Why?
- Produce
- Annotate

Further reading

- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, Nov 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.


What?
Datasets
Attributes
Data Types
Items
Data and D
Tables
Items
Attributes
Dataset Typ
Tables
Items (rows)
Cell
Multidir
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
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?

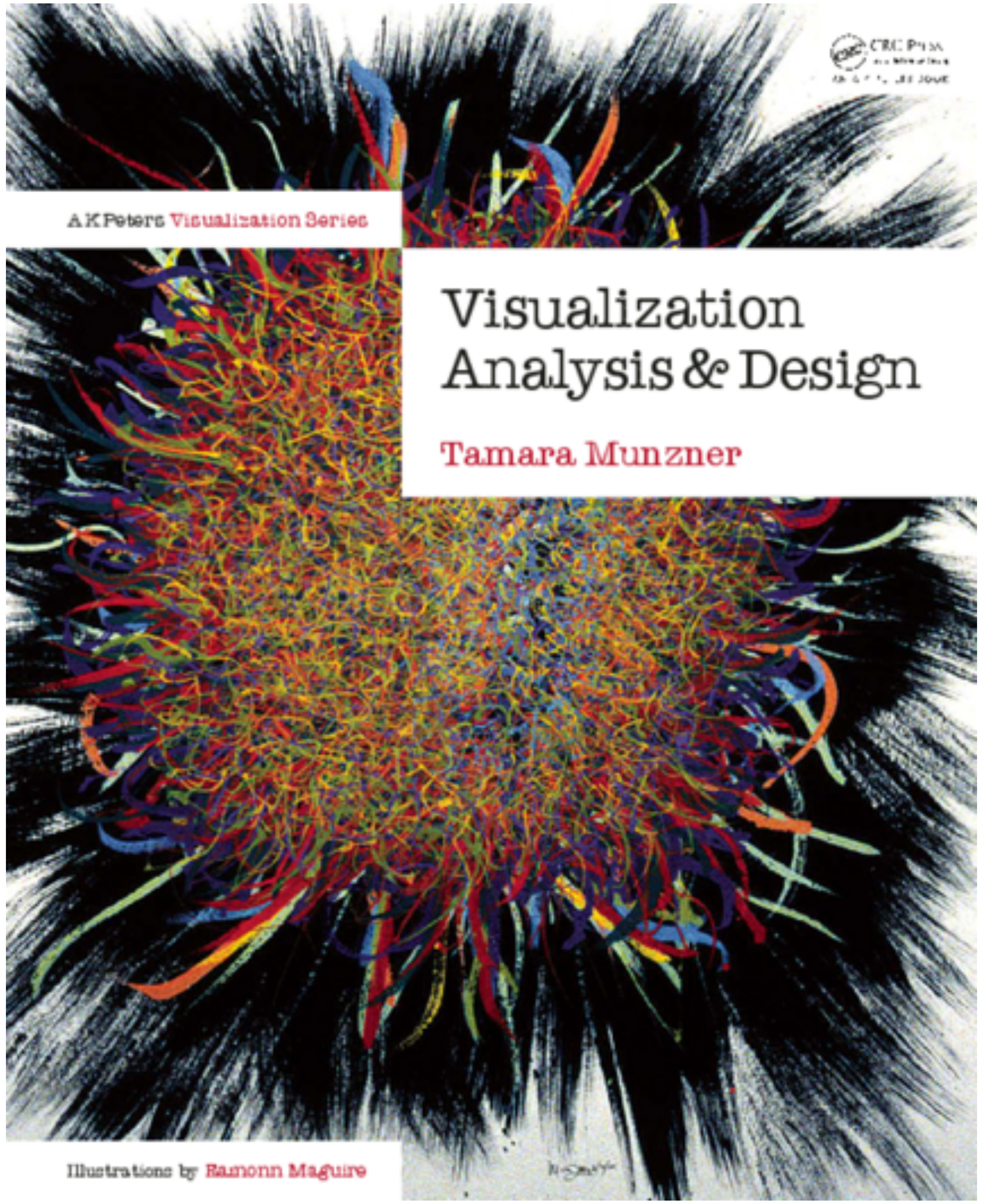
More Information

@tamaramunzner

- this talk
  http://www.cs.ubc.ca/~tmm/talks.html#vad17sydney

- book page (including tutorial lecture slides)
  http://www.cs.ubc.ca/~tmm/vadbook
  - 20% promo code for book+ebook combo: HVN17
  - http://www.crcpress.com/product/isbn/9781466508910
  - illustrations: Eamonn Maguire

- papers, videos, software, talks, full 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.