Ch 7/10: Tables, Color
Paper: D3

Tamara Munzner
Department of Computer Science
University of British Columbia

CPSC 547, Information Visualization
Week 5: 10 October 2017

http://www.cs.ubc.ca/~tmm/courses/547-17F

This Time

- paper: ArteryViz (carryforward from last time)
- chapters: Tables, Color
  - some new material, not just backup slides
- paper: D3
  - system context
- 3 shorter in-class exercises
  - Two Numbers
  - Bars/Radial
  - Color Palettes

Next Time

- to read
  - VAD Ch. 8: Arrange Spatial Data
  - VAD Ch. 9: Arrange Networks
  - paper: ABySS-Explorer: visualizing genome sequence assemblies.. Cydney B. Nielsen, Shaun D. Jackman, Inanc Birol, Steven J.M. Jones. TVCG 15(6):881-8, 2009 (Proc. InfoVis 2009).
    - [paper type: design study]
  - paper: Interactive Visualization of Genealogical Graphs. Michael J. McGuffin, Ravin Balakrishnan. Proc. InfoVis 2005, pp 17-24.
    - [paper type: technique]
- to prepare
  - project pitches (3 min each)

Ch 7: Arrange Tables

VAD Ch 7: Arrange Tables

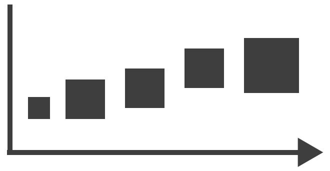
Encode

Arrange

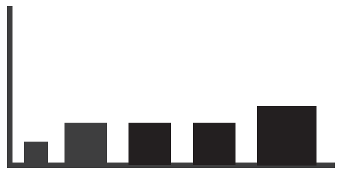
- Express

- Separate

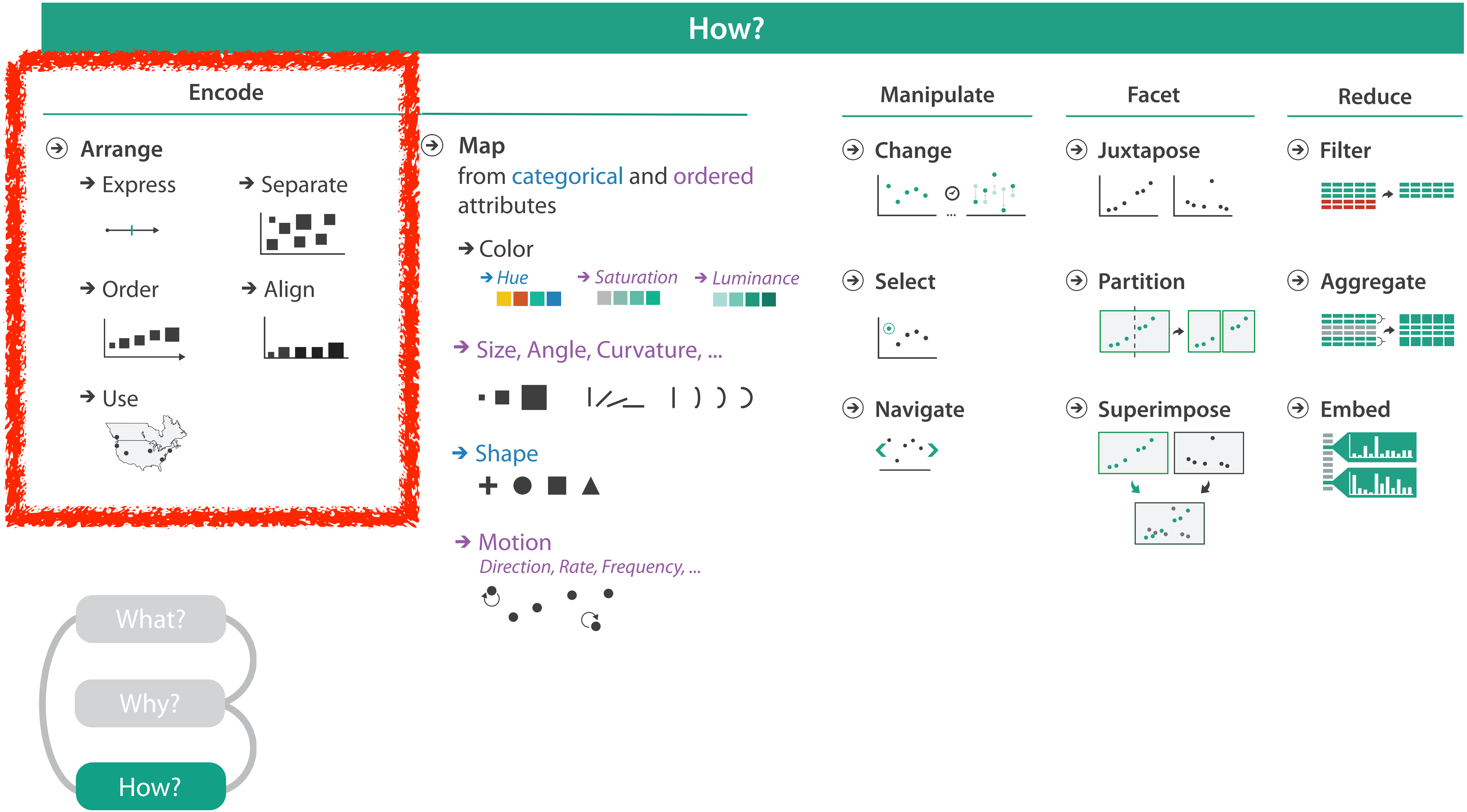
- Order



- Align



- Use


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?

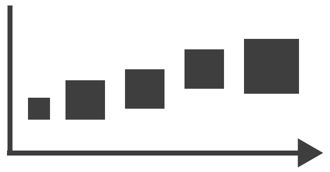
Encode tables: Arrange space

Encode

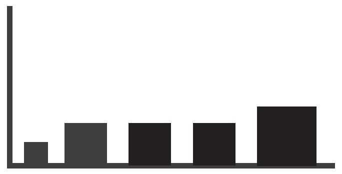
- **Arrange**
  - Express

  - Separate

  - Order



  - Align

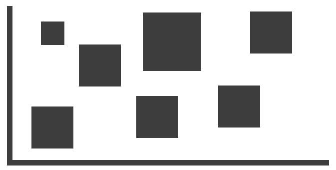


Arrange tables

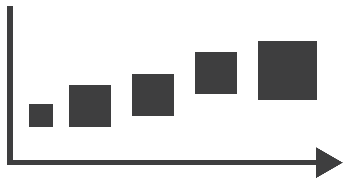
Express Values

Separate, Order, Align Regions

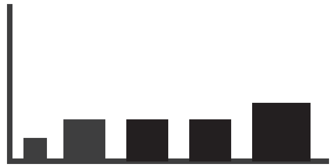
- Separate



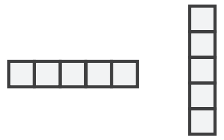
- Order



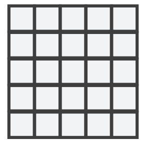
- Align



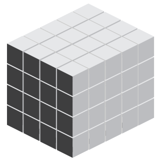
- *1 Key*
  List



- *2 Keys*
  Matrix



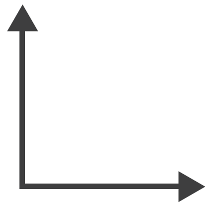
- *3 Keys*
  Volume



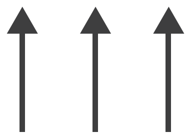
- *Many Keys*
  Recursive Subdivision

Axis Orientation

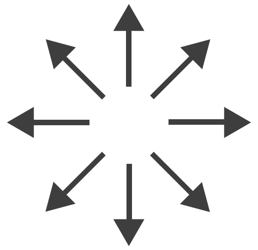
- Rectilinear



- Parallel

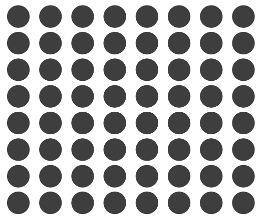


- Radial



Layout Density

- Dense

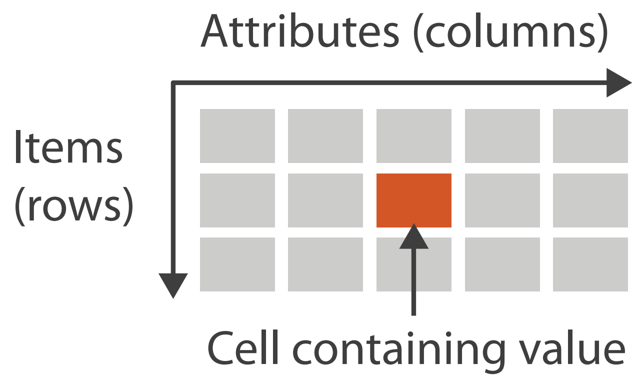


- Space-Filling

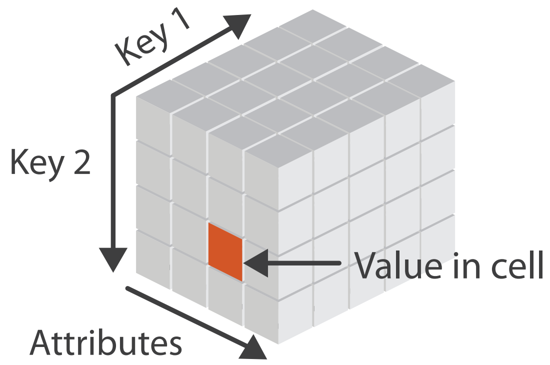
Keys and values

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

→ Tables

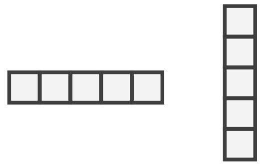


→ *Multidimensional Table*

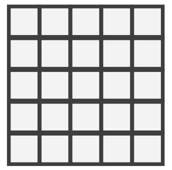


Express Values

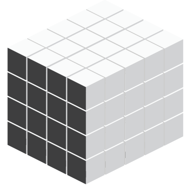
→ *1 Key*
List



→ *2 Keys*
Matrix

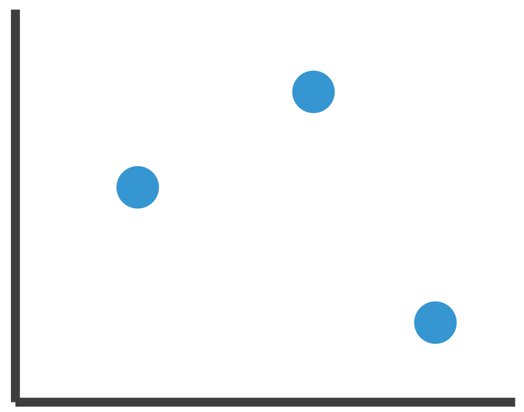


→ *3 Keys*
Volume

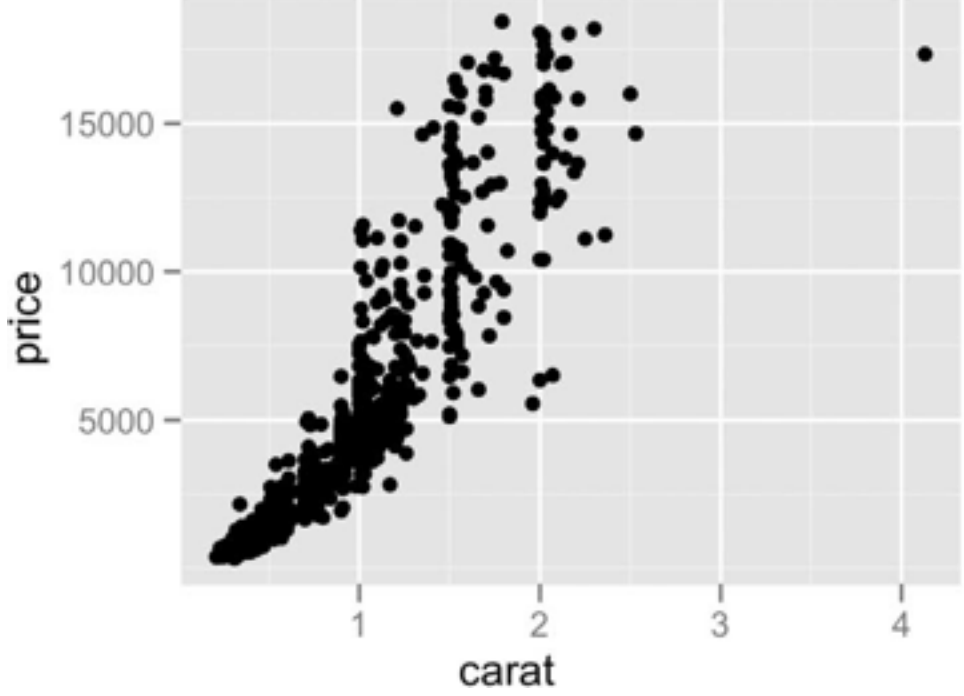


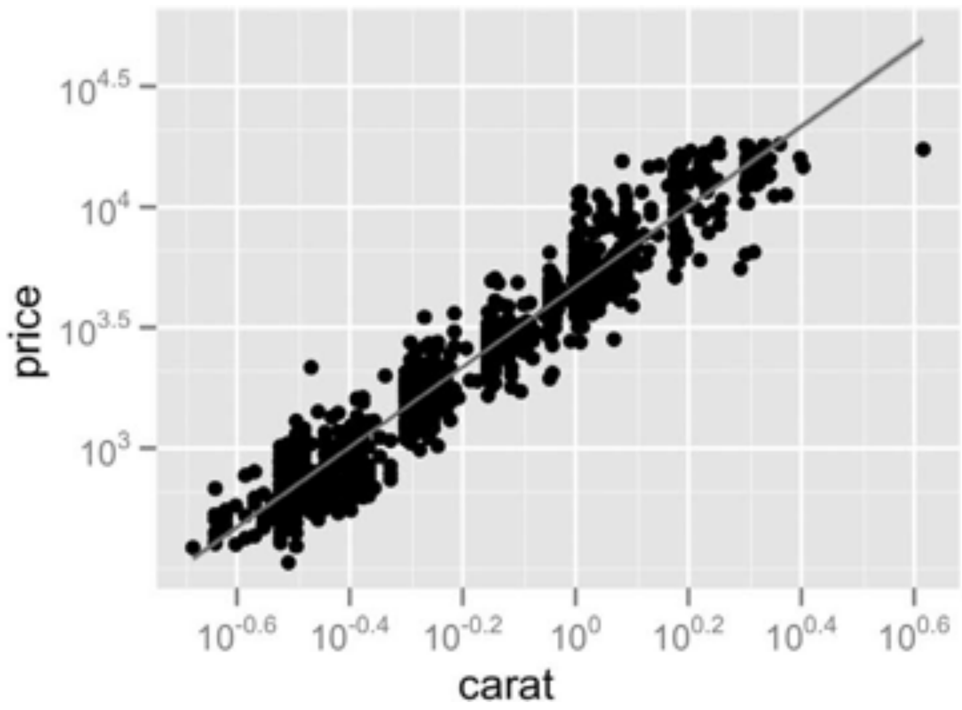
→ *Many Keys*
Recursive Subdivision

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



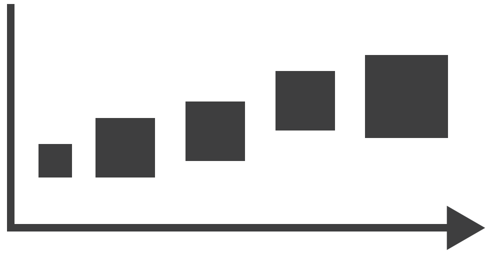


[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

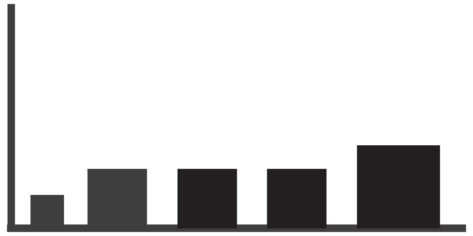
Some keys: Categorical regions

➔ Separate

➔ Order

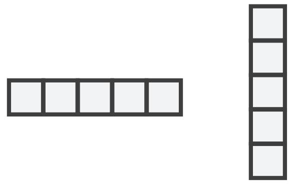


➔ Align

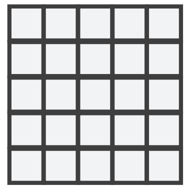


- **regions**: contiguous bounded areas distinct from each other
  - using space to ***separate*** (proximity)
  - following expressiveness principle for categorical attributes
- use ordered attribute to ***order*** and ***align*** regions

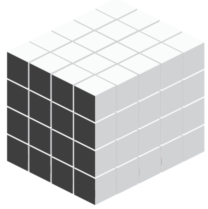
➔ *1 Key*
List



➔ *2 Keys*
Matrix



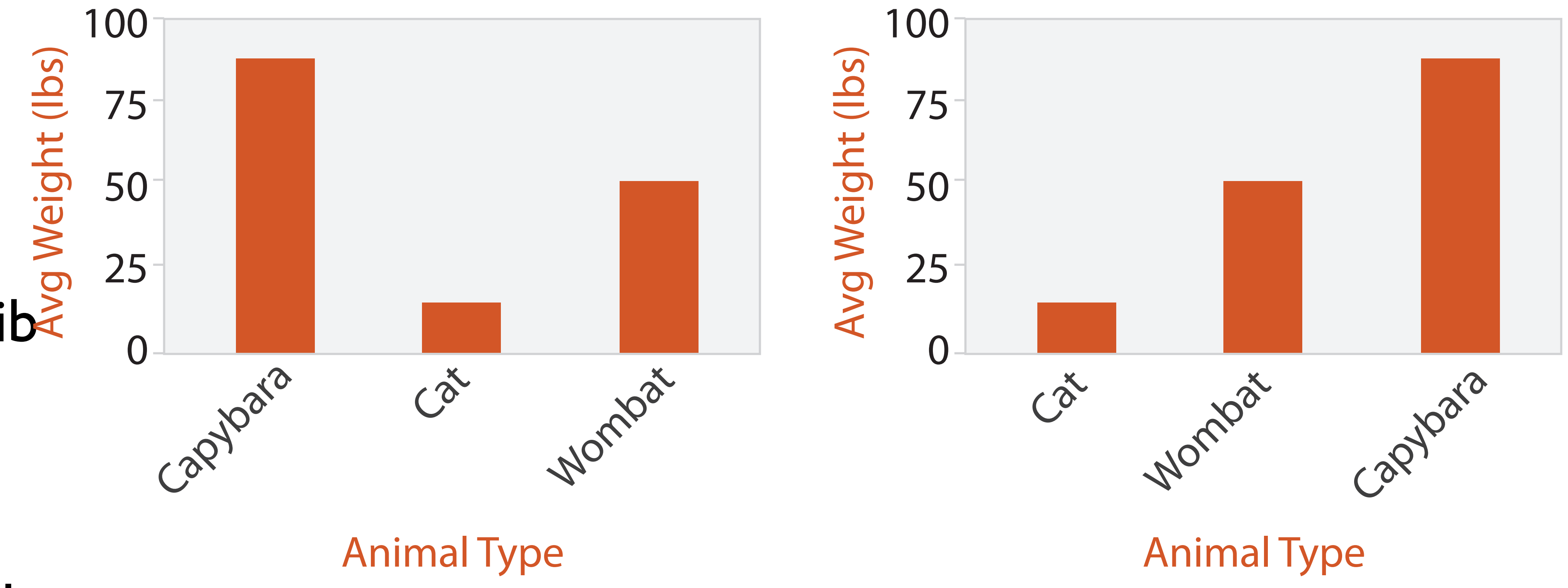
➔ *3 Keys*
Volume



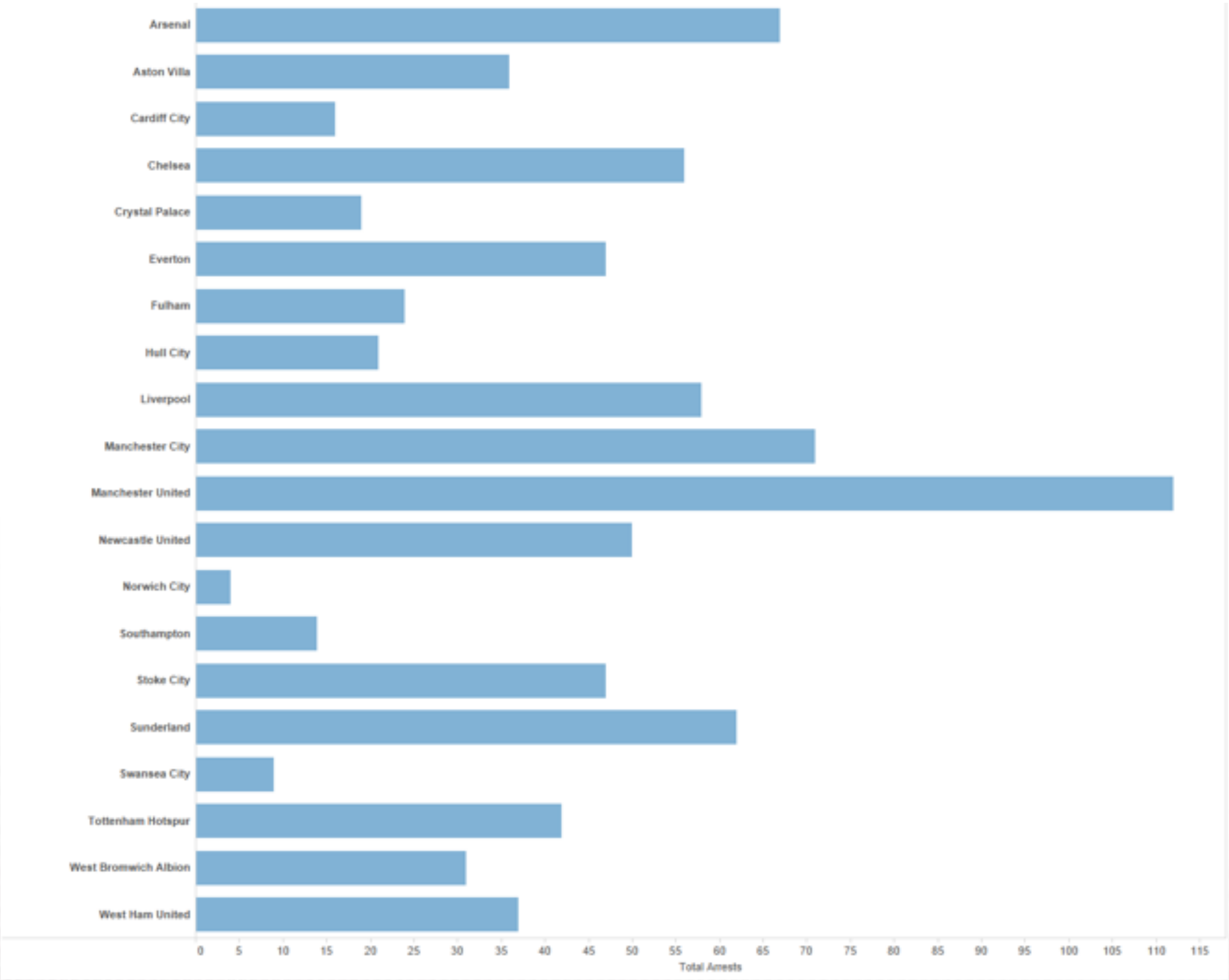
➔ *Many Keys*
Recursive Subdivision

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



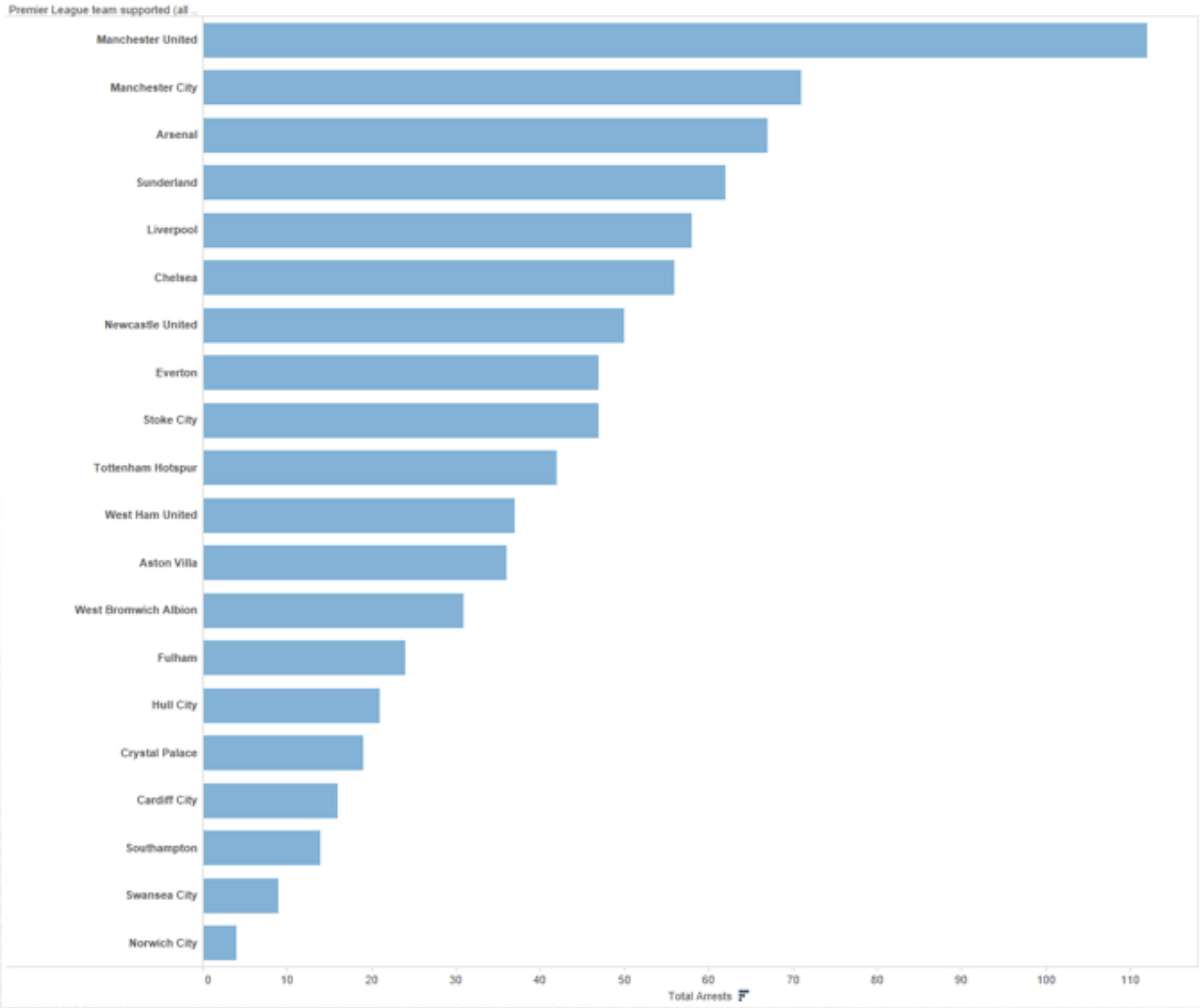
Separated and Aligned but not Ordered



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

[Slide courtesy of Ben Jones]

Separated, Aligned and Ordered



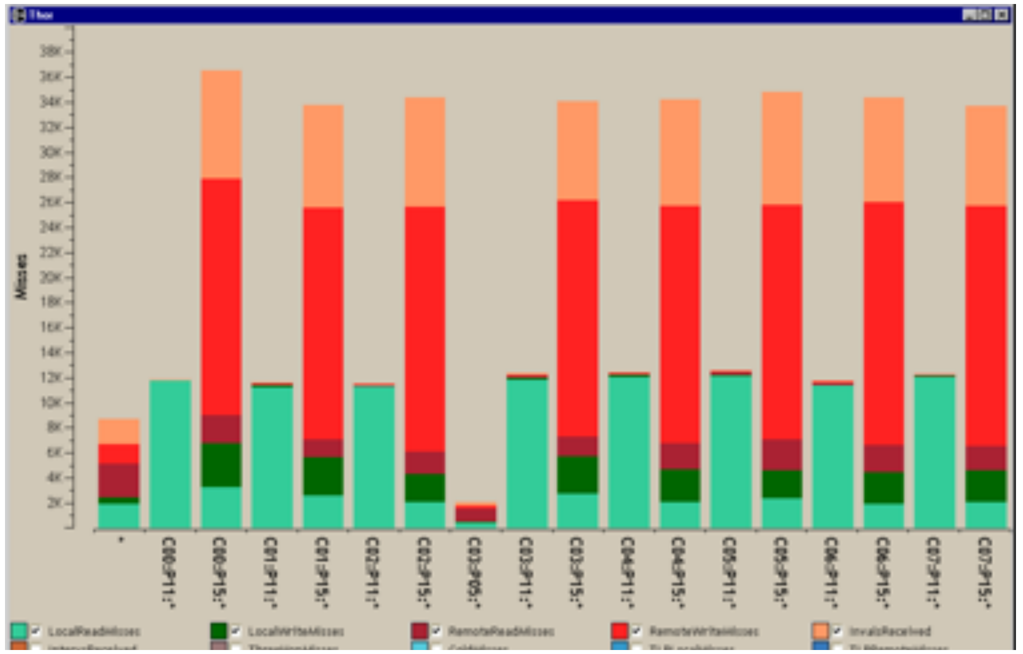
[Slide courtesy of Ben Jones]

Separated but not Ordered or Aligned

LIMITATION: Hard to make comparisons

[Slide courtesy of Ben Jones]

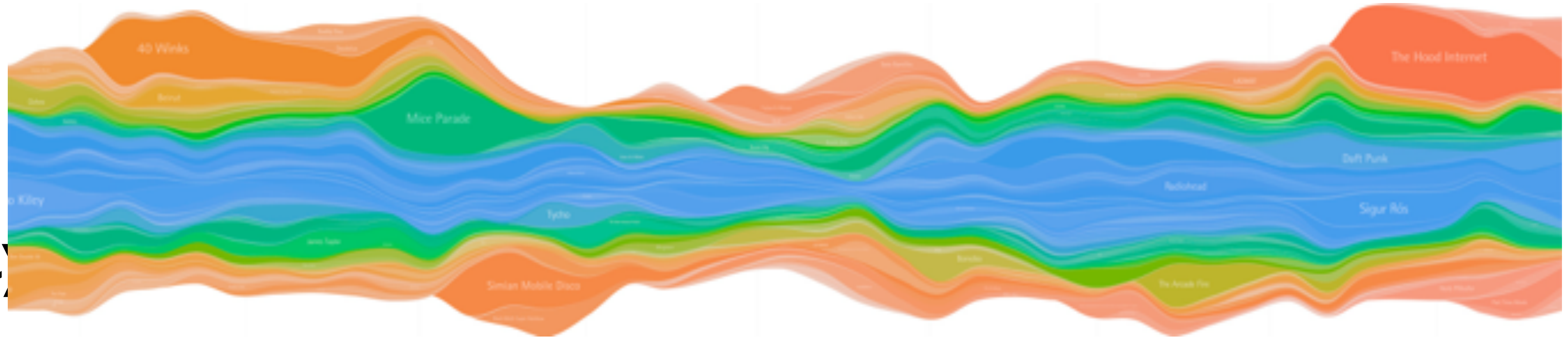
Idiom: **stacked bar chart**



- 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.]

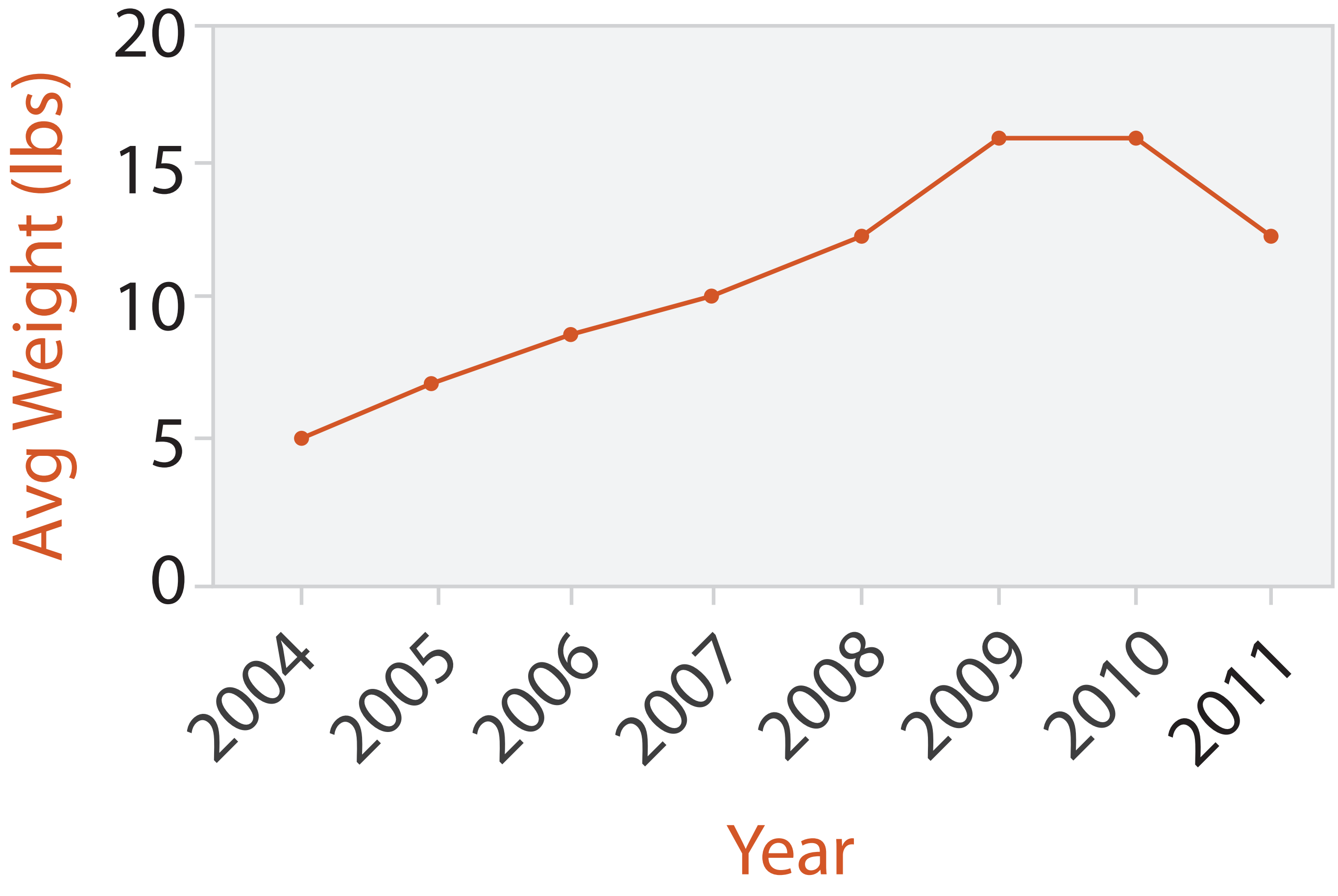
Idiom: **streamgraph**



[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

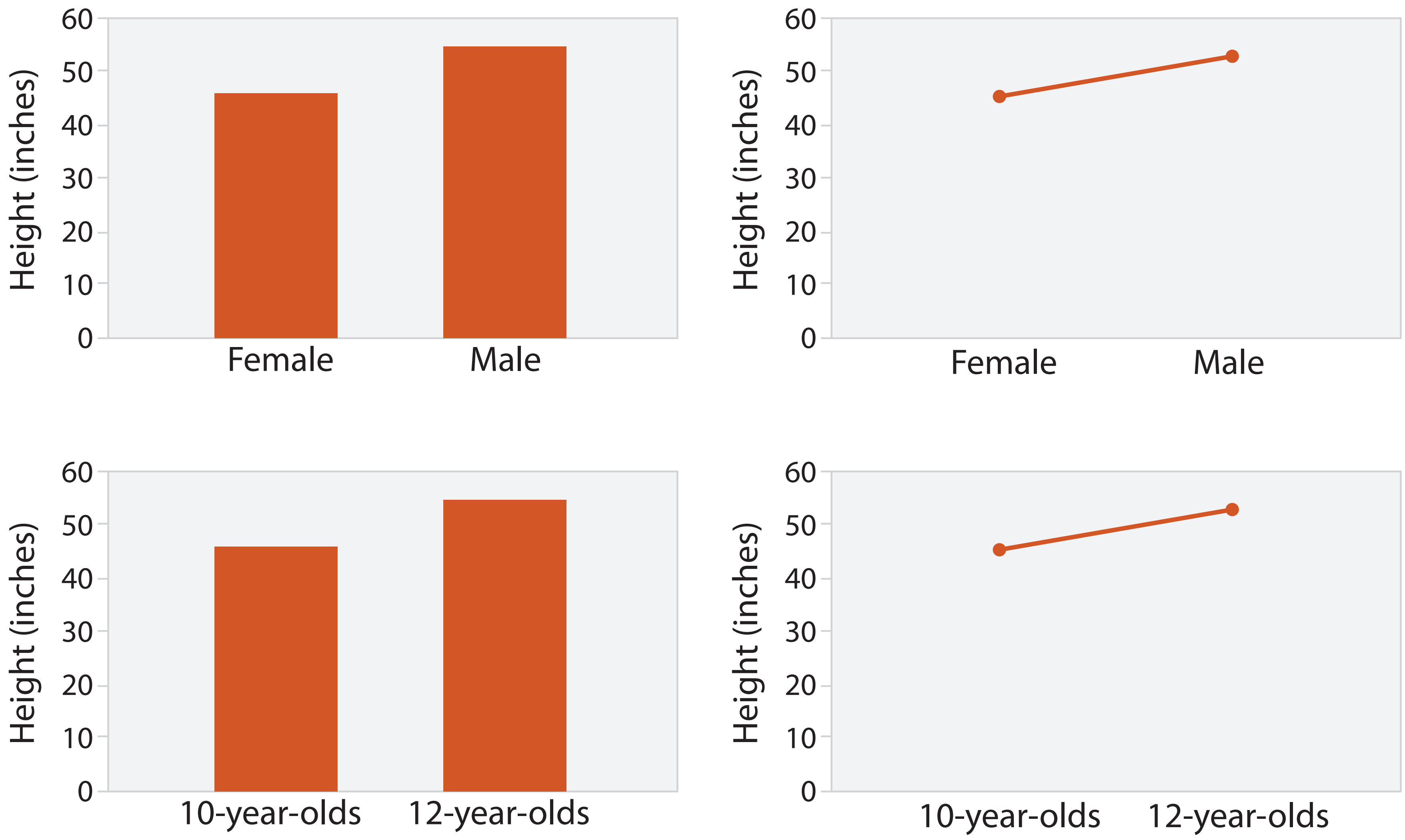
Idiom: **line chart / dot plot**



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

Choosing bar vs line charts

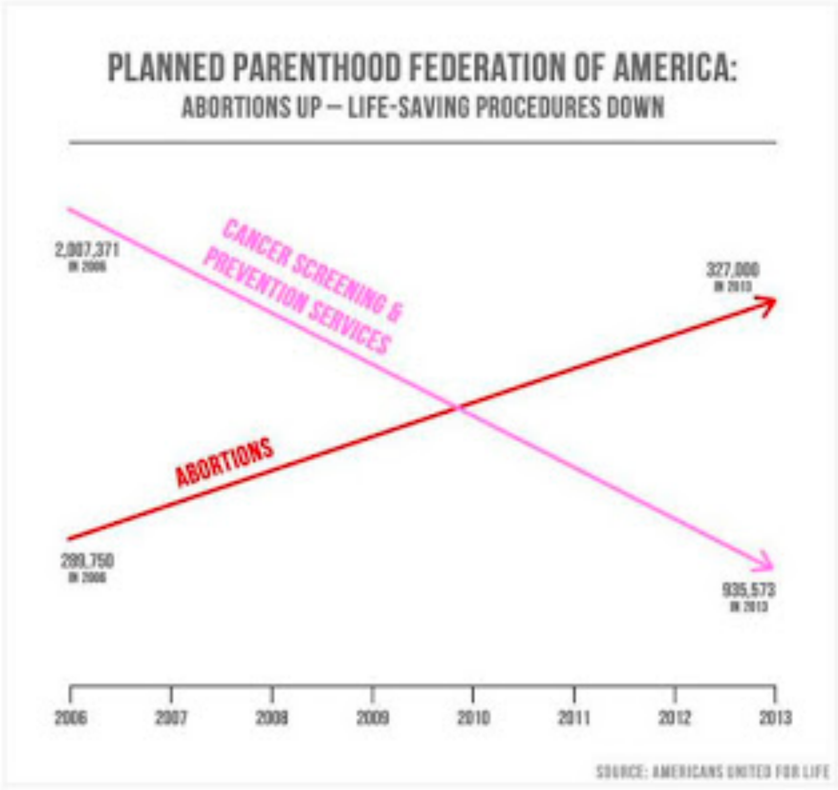
- depends on type of key attrib
  - –bar charts if categorical
  - –line charts if ordered
- do not use line charts for categorical key attribs
  - –violates expressiveness principle
    - implication of trend so strong that it overrides semantics!
      - – “The more male a person is, the taller he/she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

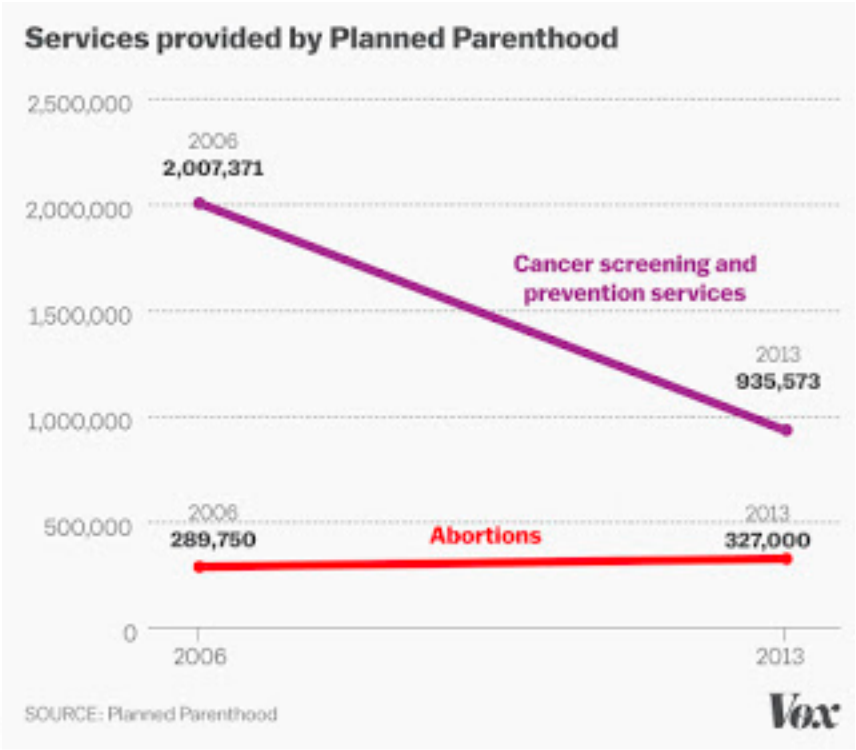


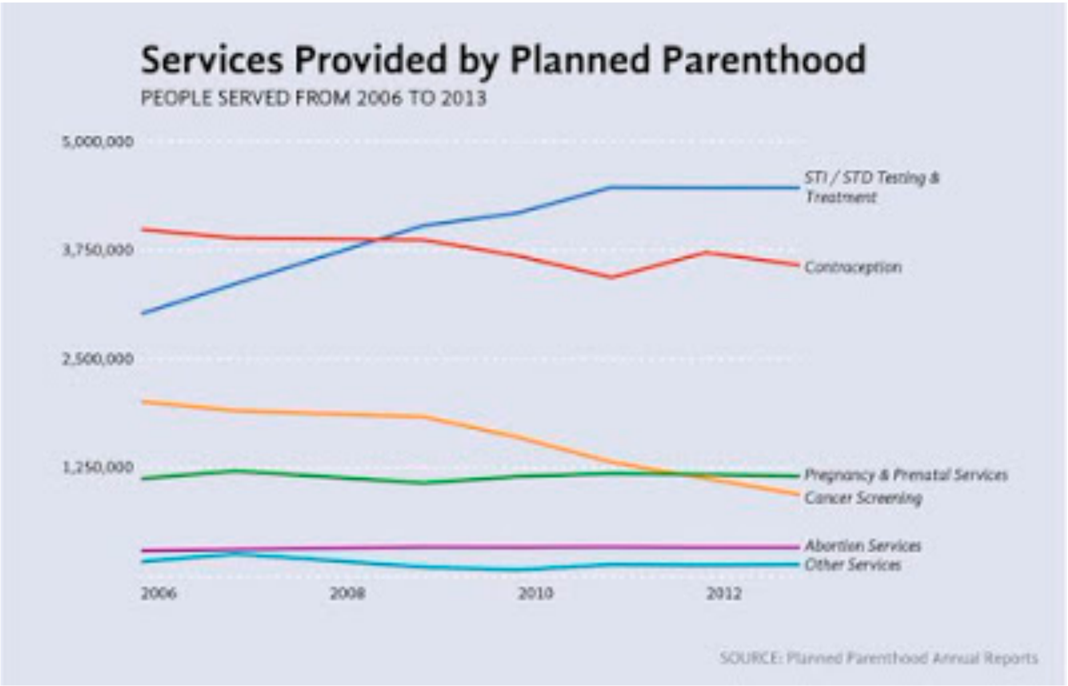
after [Bars and Lines: A Study of Graphic Communication. Zacks and Tversky. Memory and Cognition 27:6 (1999), 1073–1079.]

Chart axes

- labelled axis is critical
- avoid cropping y-axis
  - include 0 at bottom left
  - or slope misleads
- dual axes controversial
  - acceptable if commensurate
  - beware, very easy to mislead!



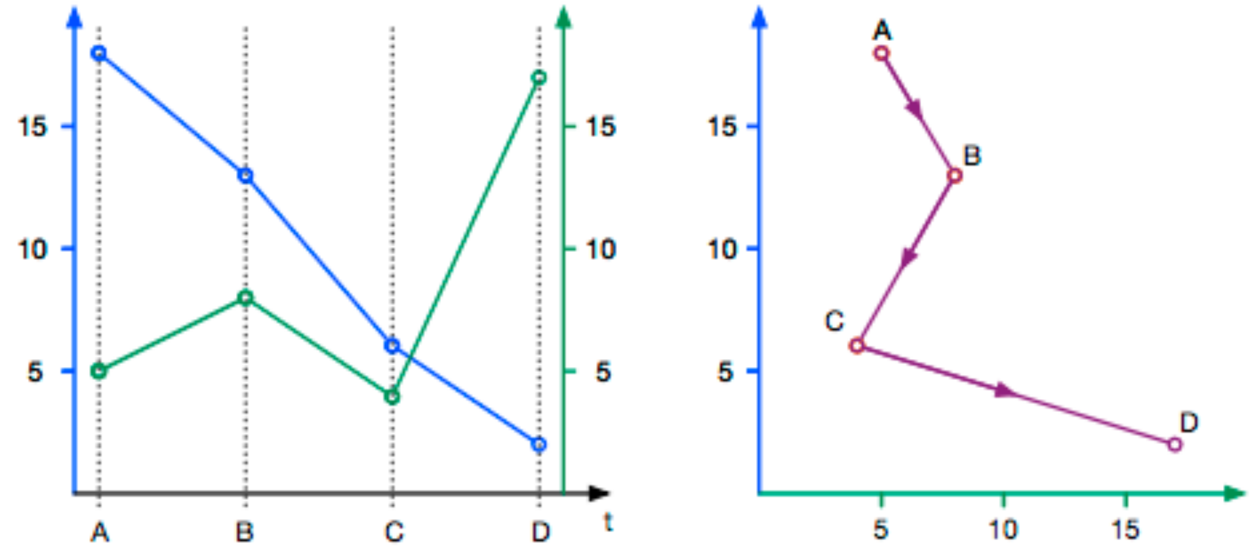


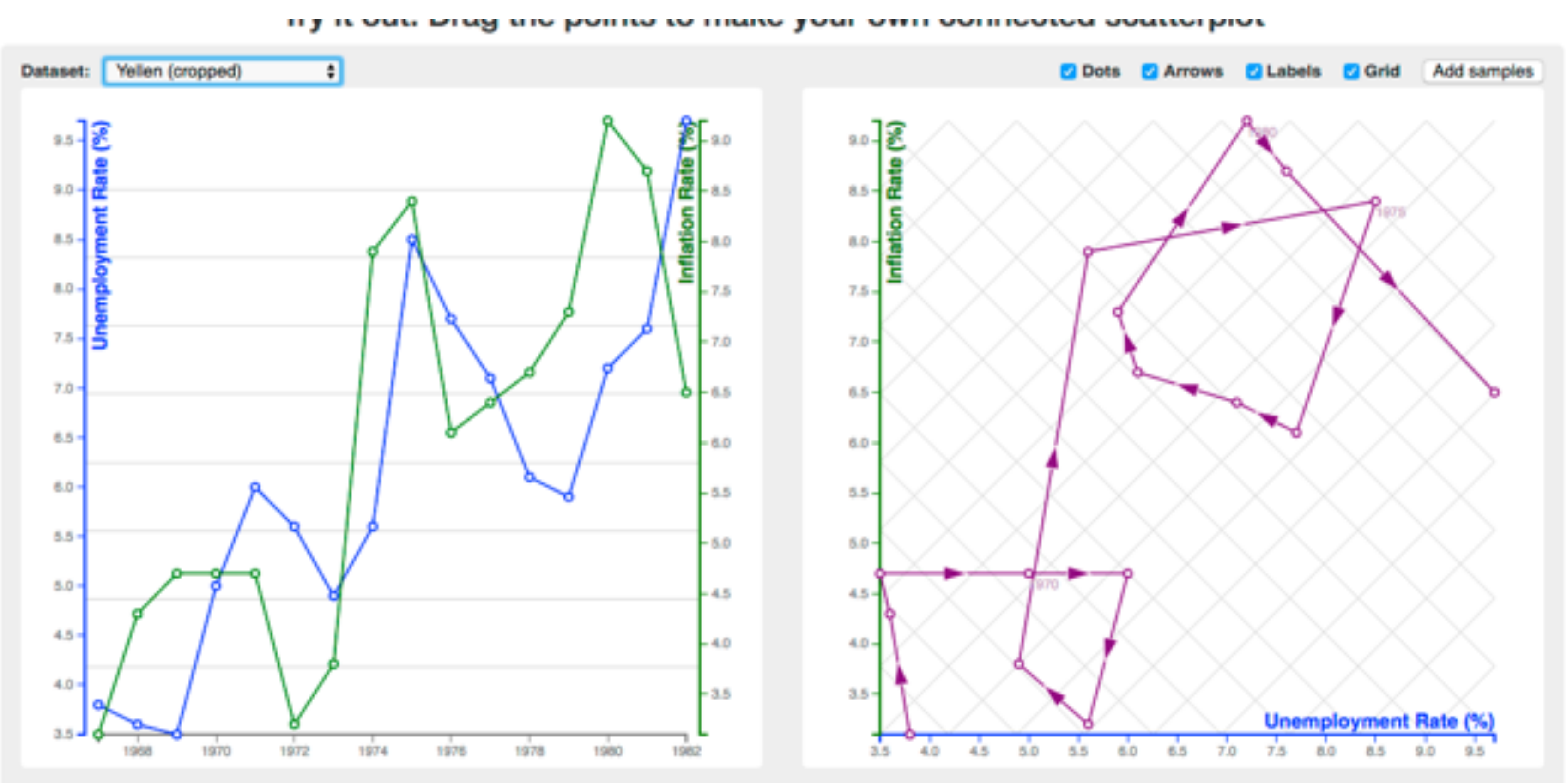


http://www.thefunctionalart.com/2015/10/if-you-see-bullshit-say-bullshit.html

Idiom: **connected scatterplots**

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



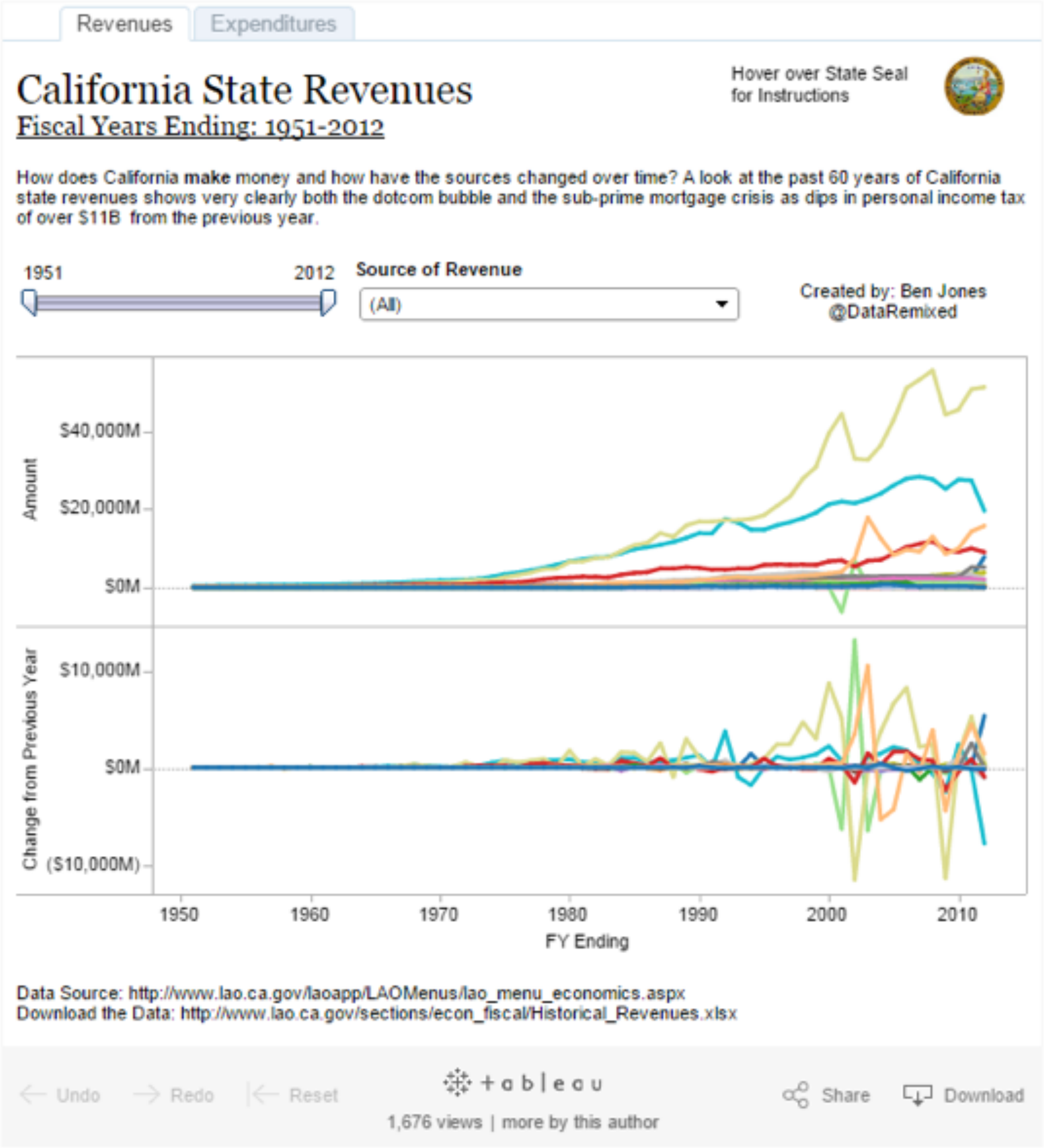


[The Connected Scatterplot for Presenting Paired Time Series. Haroz, Kosara and Franconeri. IEEE TVCG 22(9):2174-86, 2016.]

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

Idiom: **Indexed line charts**

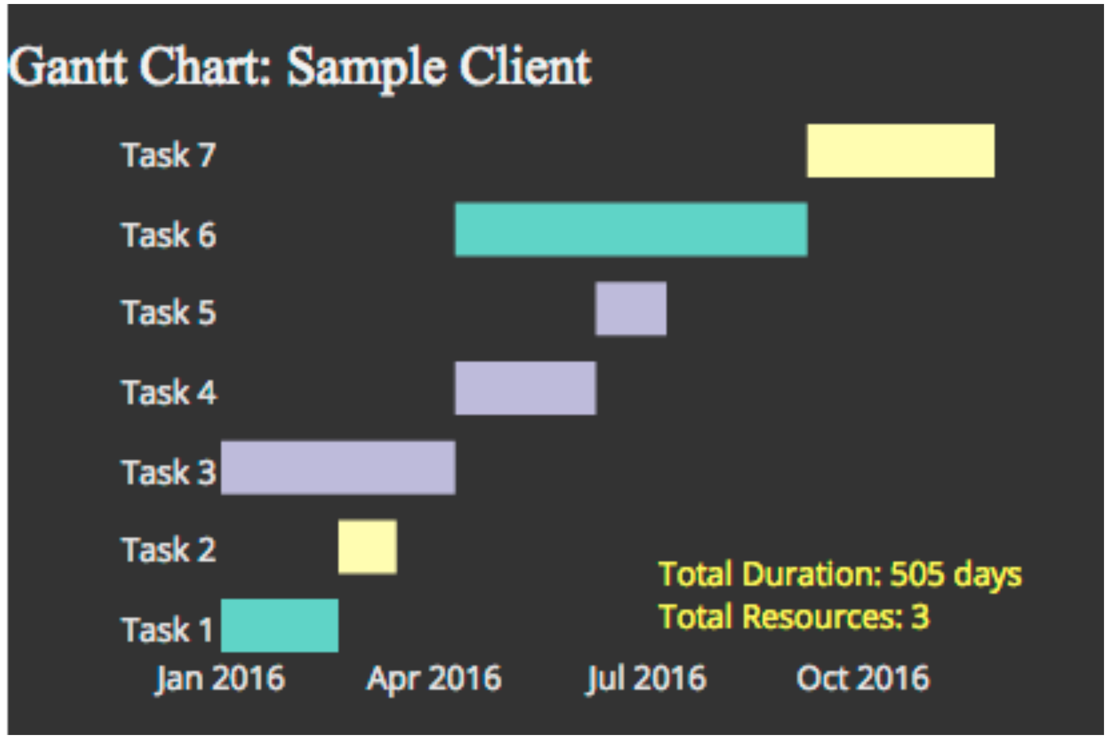
- data: 2 quant attires
  - 1 key + 1 value
- derived data: new quant value attrib
  - index
  - plot instead of original value
- task: show change over time
  - principle: normalized, not absolute
- scalability
  - same as standard line chart



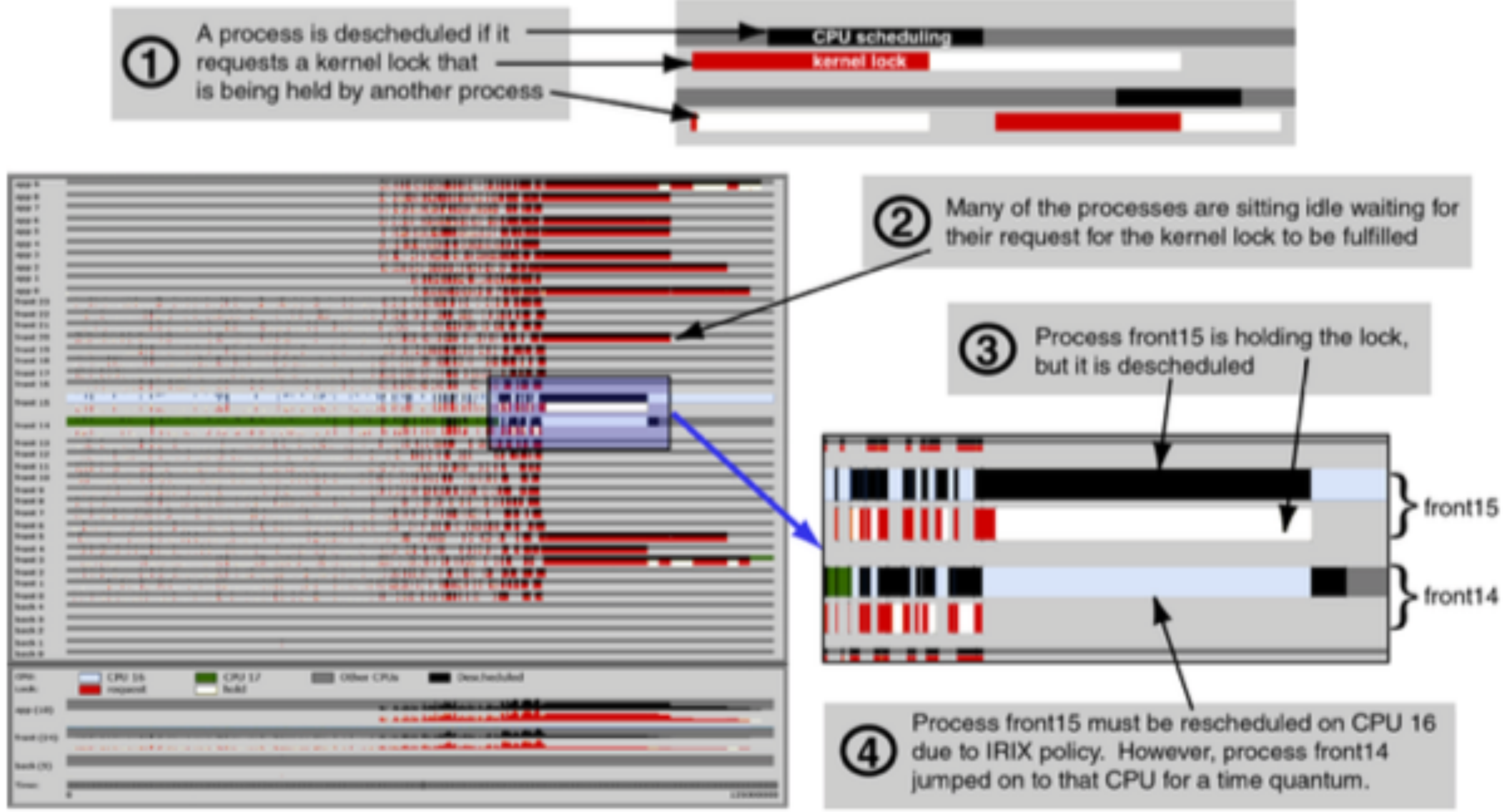
https://public.tableau.com/profile/ben.jones#!/vizhome/CAStateRevenues/Revenues

Idiom: **Gantt charts**

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



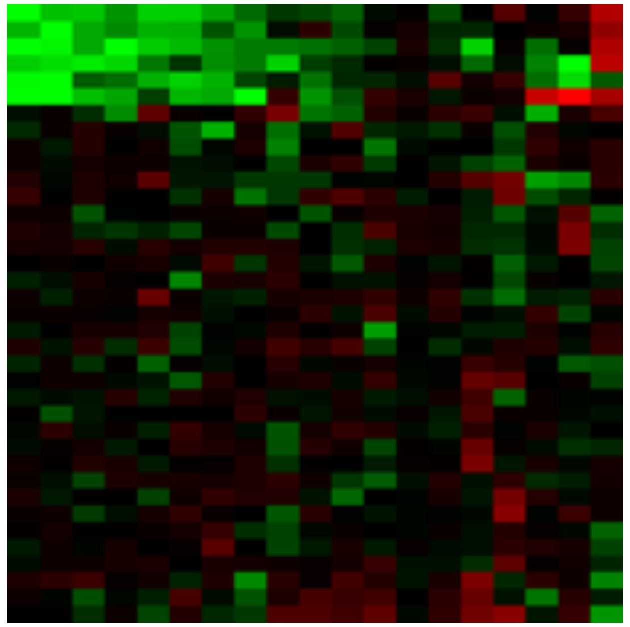
https://www.r-bloggers.com/gantt-charts-in-r-using-plotly/

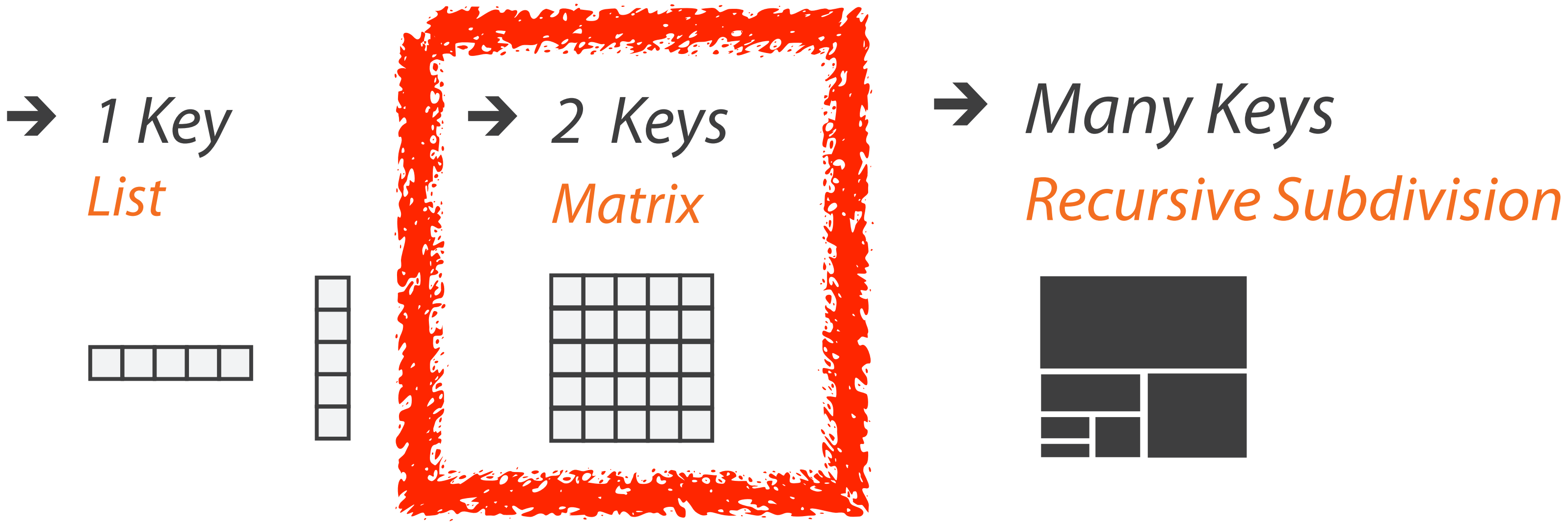


[Performance Analysis and Visualization of Parallel Systems Using SimOS and Rivet: A Case Study. Bosch, Stolte, Stoll, Rosenblum, and Hanrahan. Proc. HPCA 2000.]

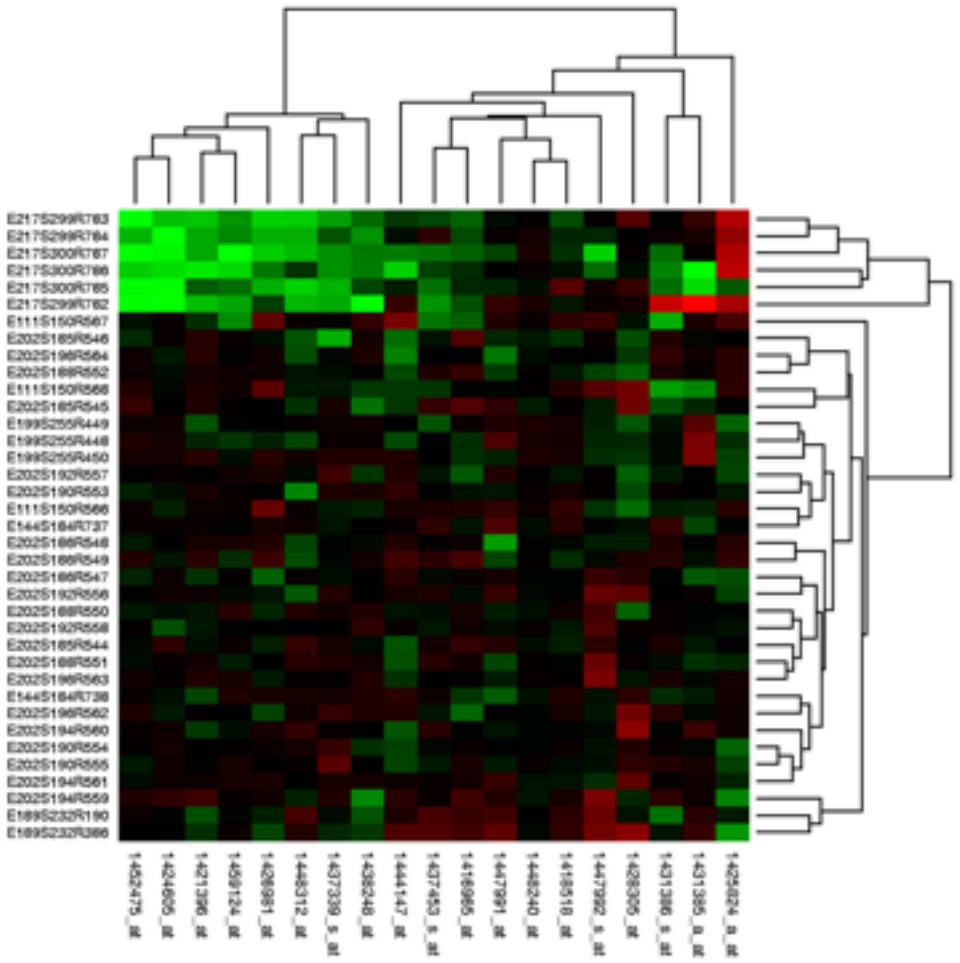
Idiom: **heatmap**

- 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





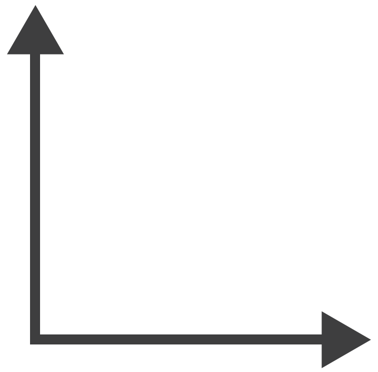
Idiom: **cluster heatmap**



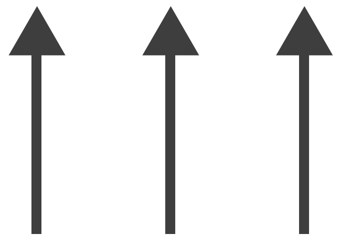
- in addition
  - derived data
    - 2 cluster hierarchies
  - dendrogram
    - parent-child relationships in tree with connection line marks
    - leaves aligned so interior branch heights easy to compare
  - heatmap
    - marks (re-)ordered by cluster hierarchy traversal

Axis Orientation

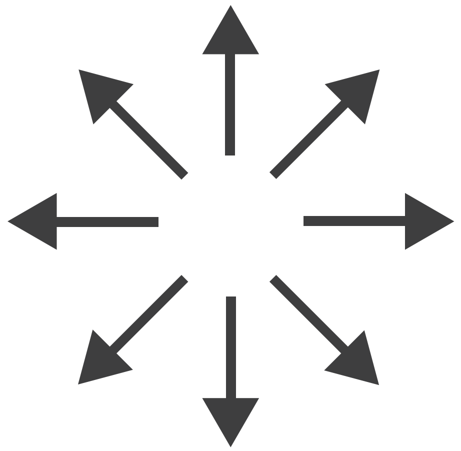
- Rectilinear



- Parallel

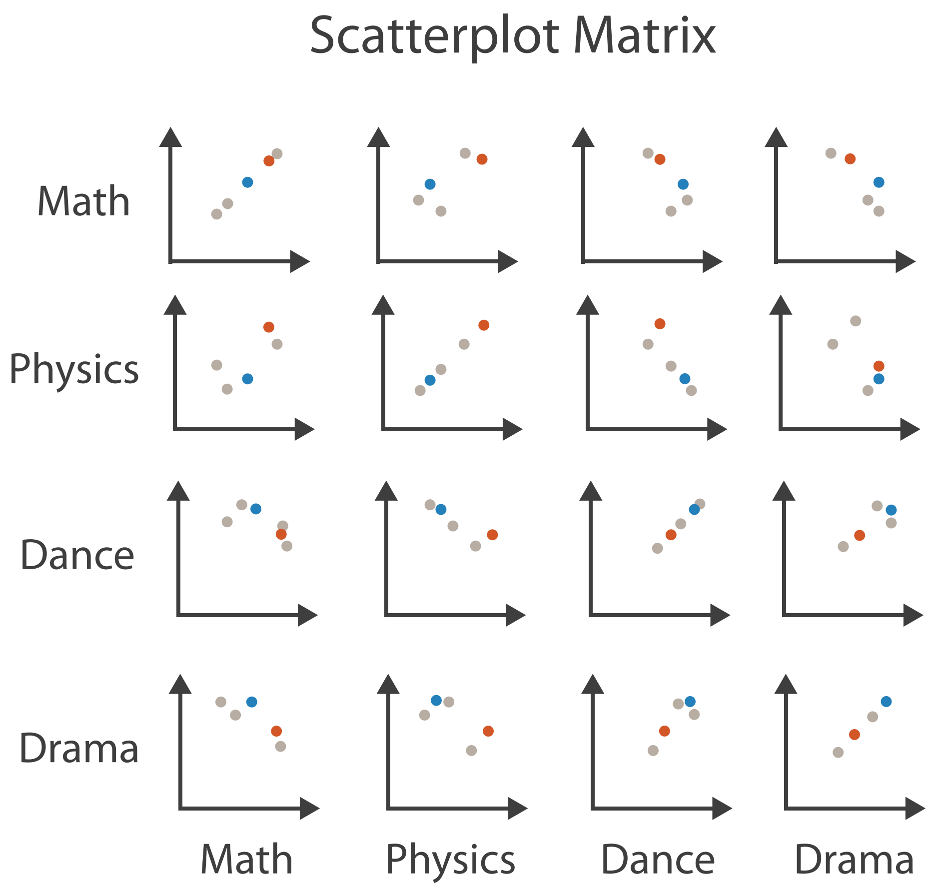


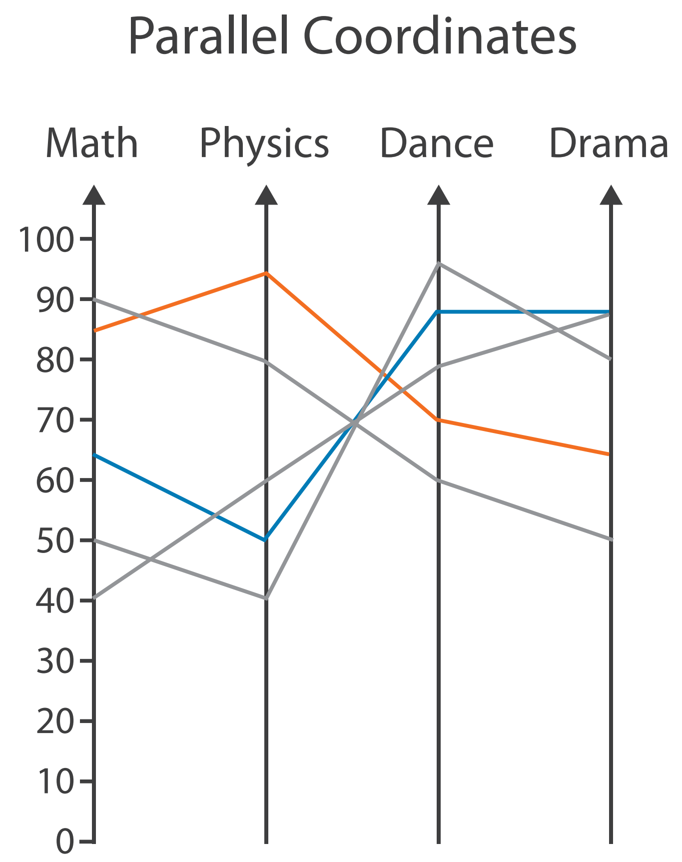
- Radial



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





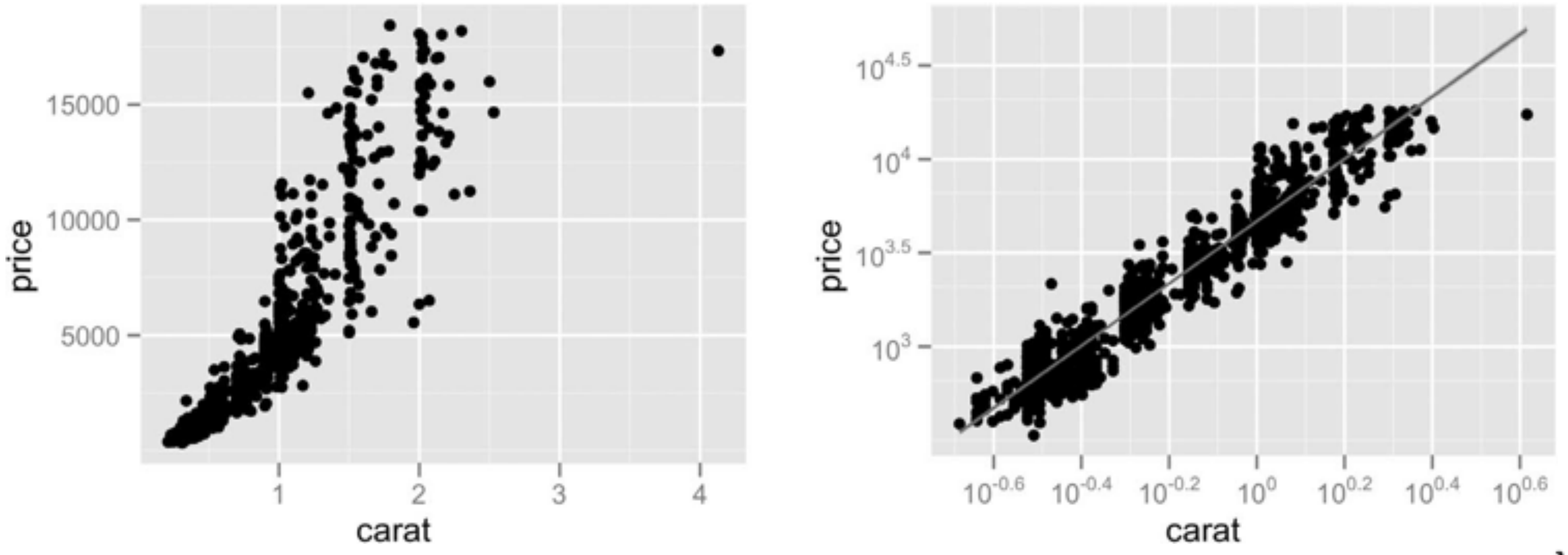
Table

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
- 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.]

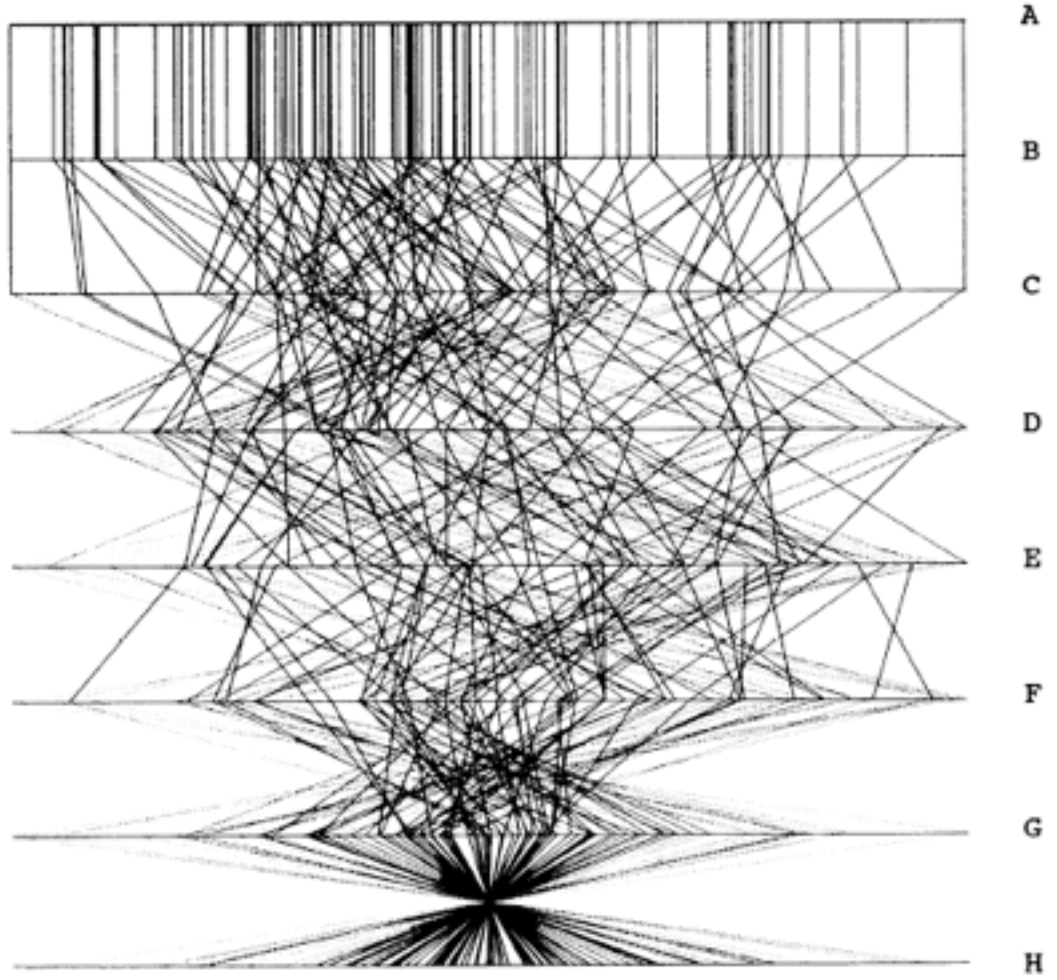


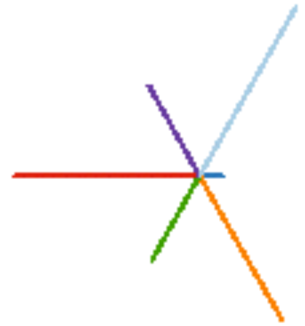
Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of ρ = 1, .8, .2, 0, −.2, −.8, and −1.

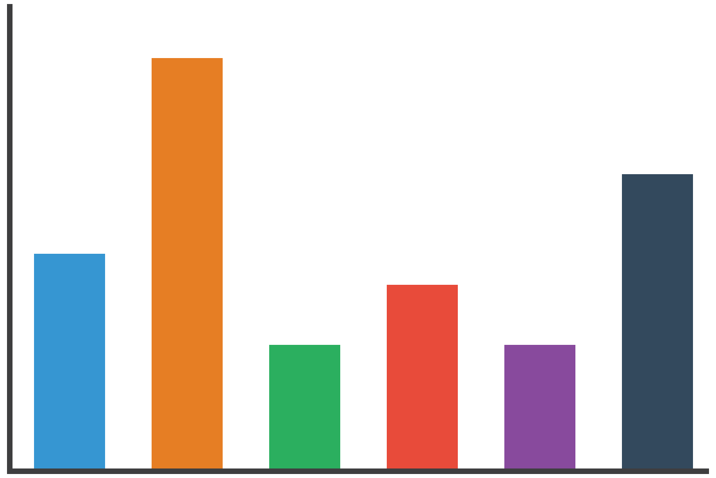
[Hyperdimensional Data Analysis Using Parallel Coordinates. Wegman. Journ. American Statistical Association 85:411 (1990), 664–675.]

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



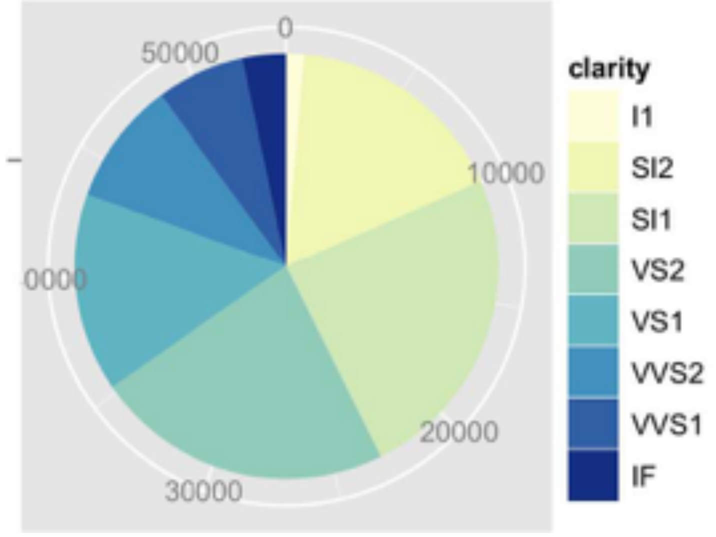


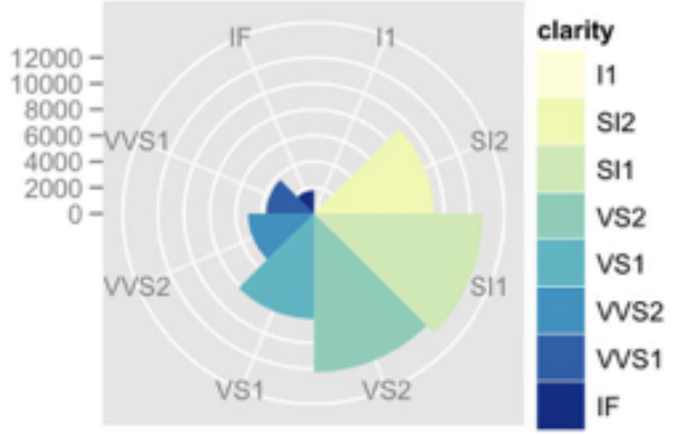
[Vismon: Facilitating Risk Assessment and Decision Making In Fisheries Management. Booshehrian, Möller, Peterman, and Munzner. Technical Report TR 2011-04, Simon Fraser University, School of Computing Science, 2011.]

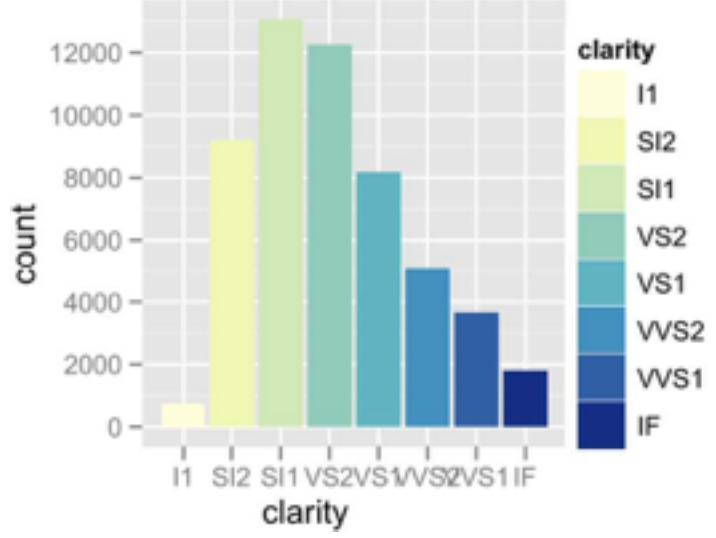
Idioms: **pie chart, polar area chart**

- pie chart
  - area marks with angle channel
  - accuracy: angle/area less accurate than line length
    - arclength also less accurate than line length
- polar area chart
  - area marks with length channel
  - more direct analog to bar charts

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



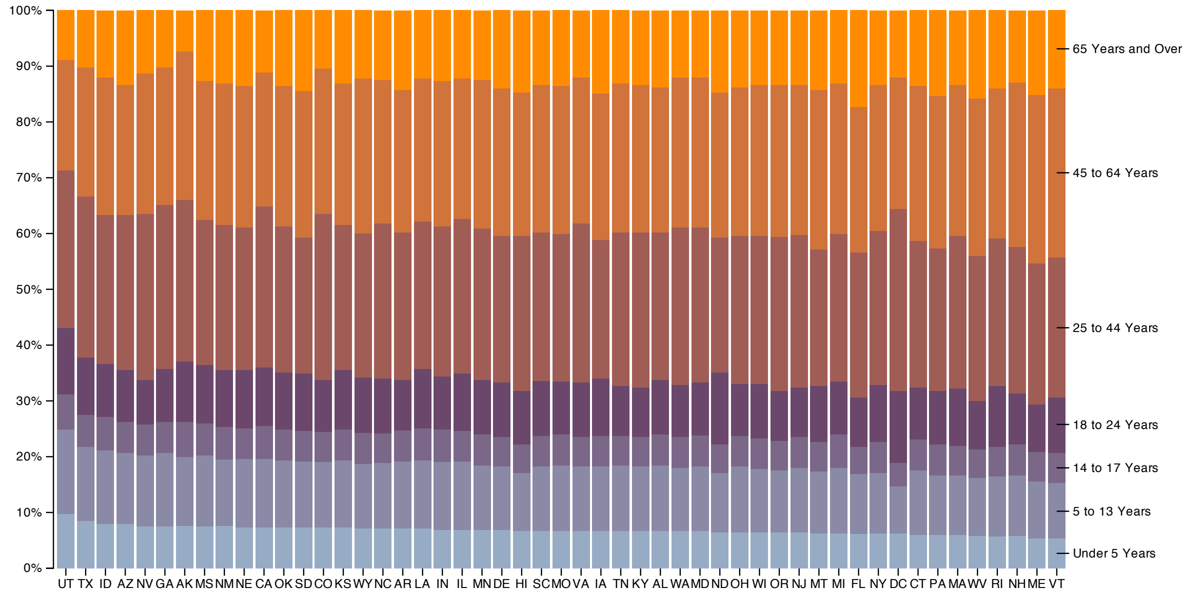


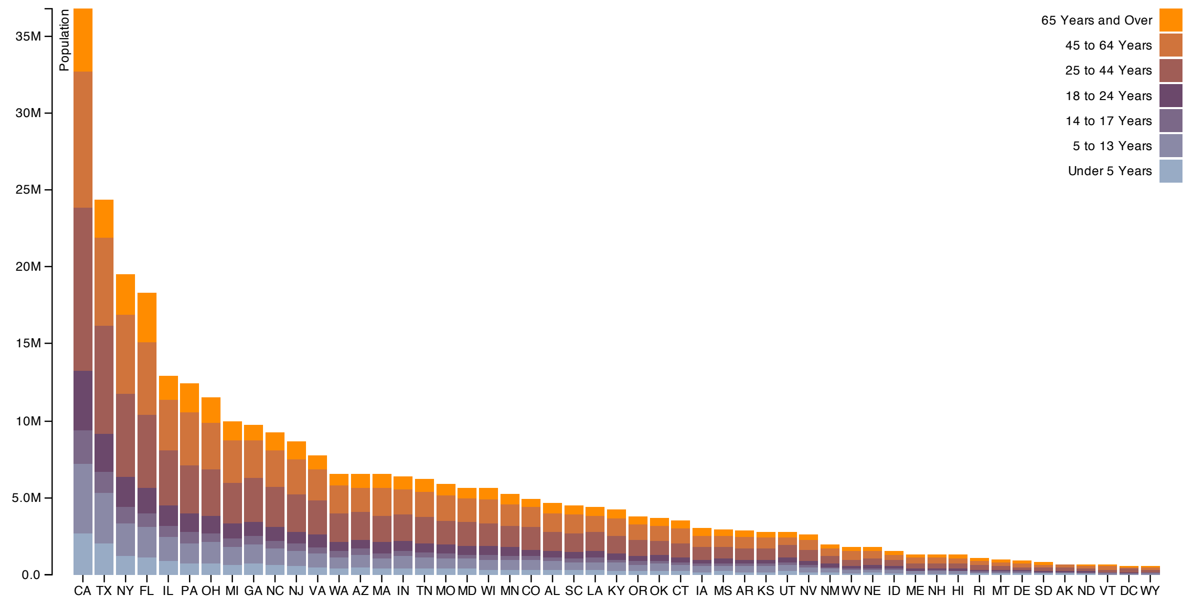


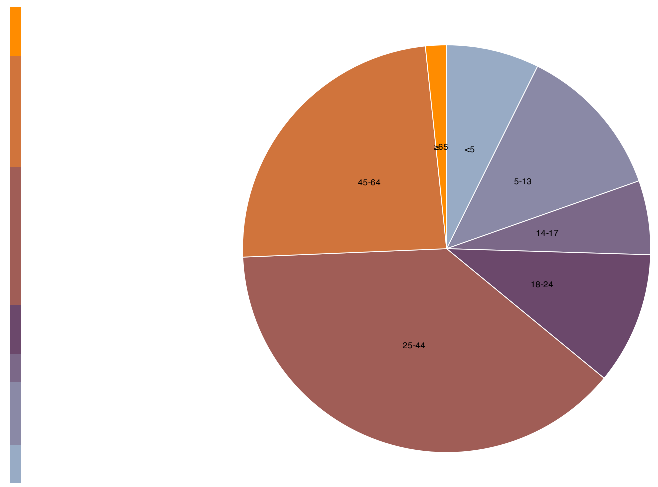
[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

Idioms: **normalized stacked bar chart**

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







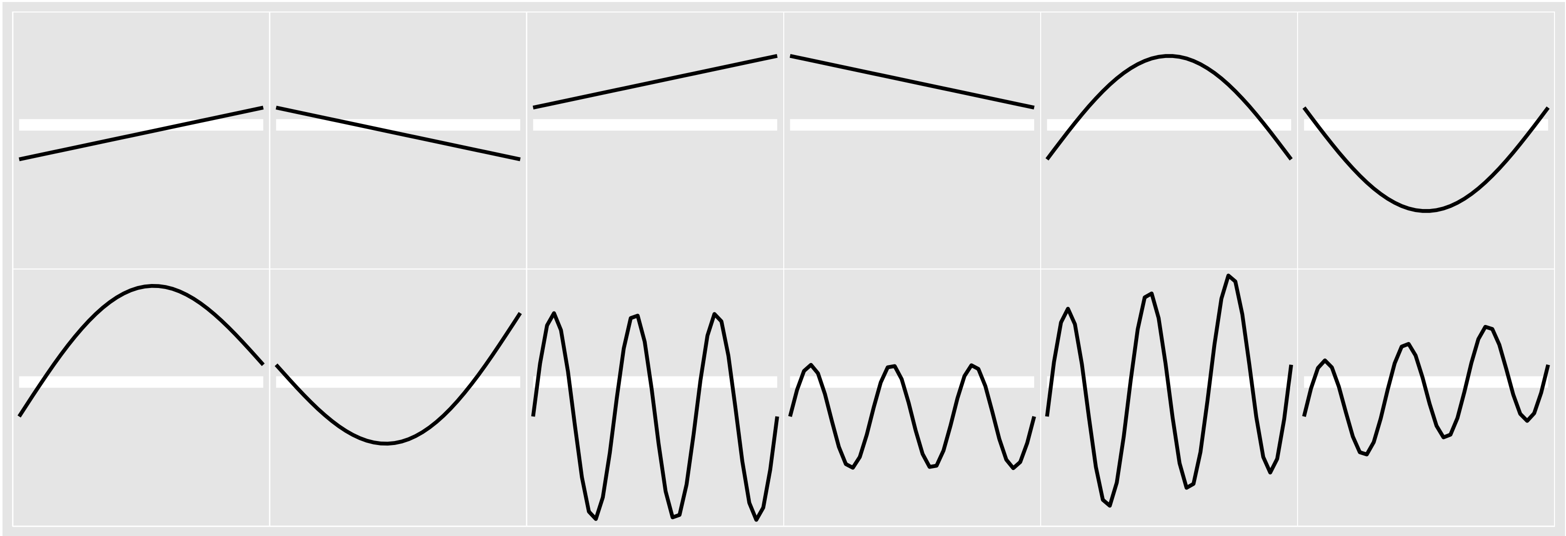
http://bl.ocks.org/mbostock/3887235,

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

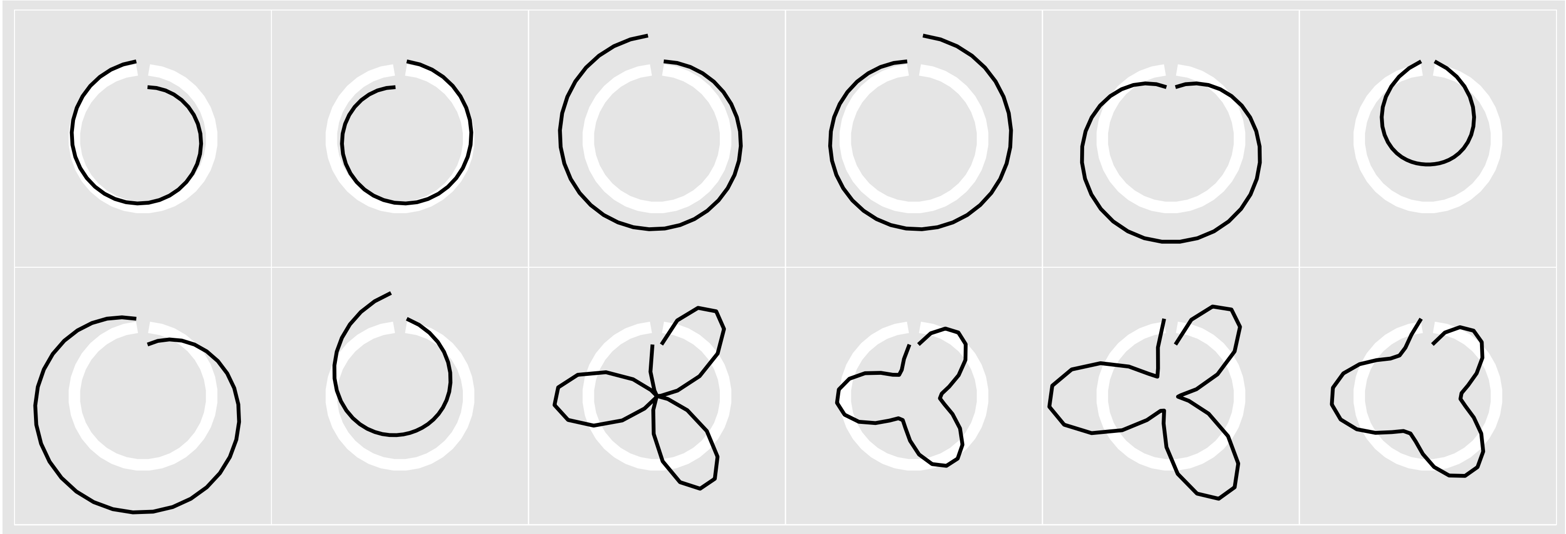
http://bl.ocks.org/mbostock/3886394.

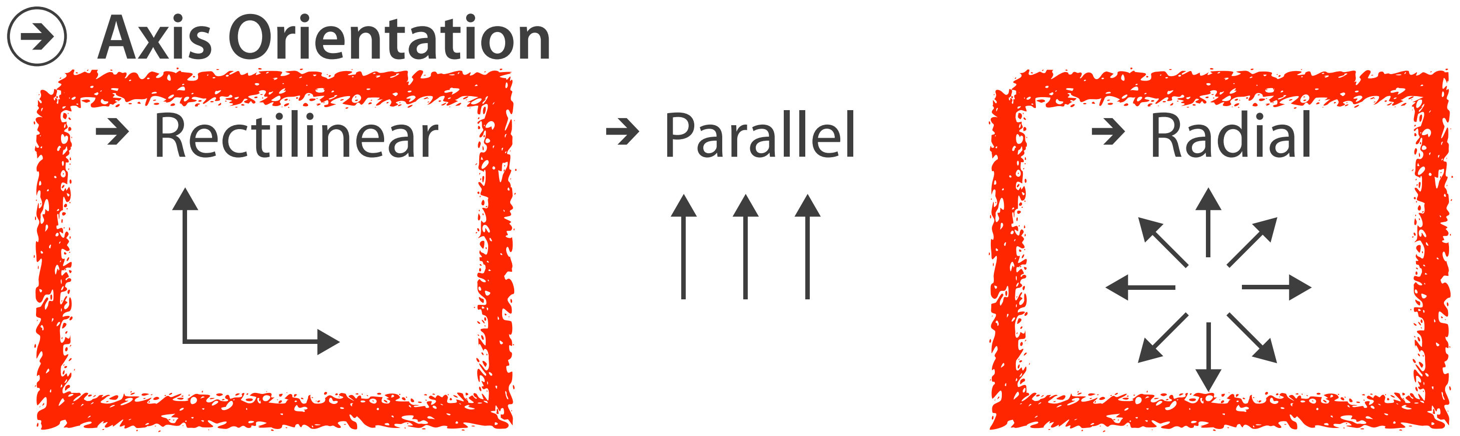
Idiom: **glyphmaps**

- rectilinear good for linear vs nonlinear trends



- radial good for cyclic patterns



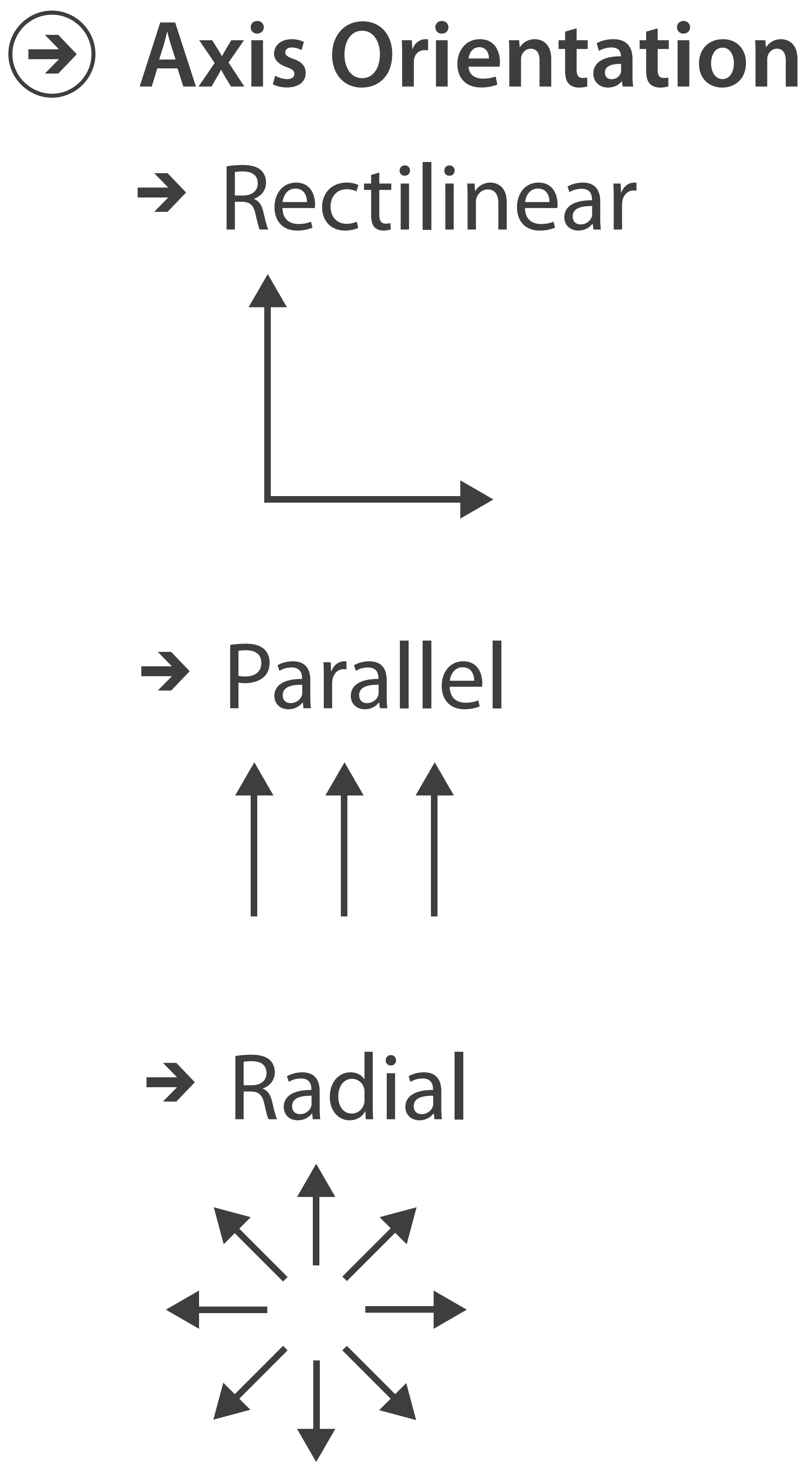


[Glyph-maps for Visually Exploring Temporal Patterns in Climate Data and Models. Wickham, Hofmann, Wickham, and Cook. Environmetrics 23:5 (2012), 382–393.]

Orientation limitations

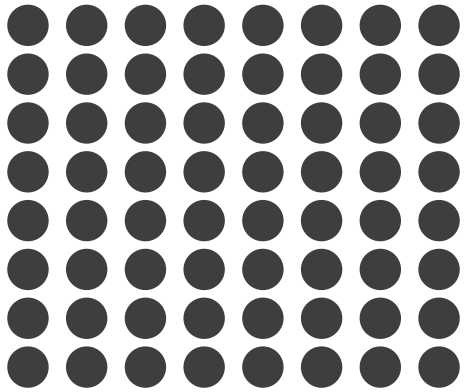
- rectilinear: scalability wrt #axes
  - 2 axes best
  - 3 problematic
    - more in afternoon
  - 4+ impossible
- parallel: unfamiliarity, training time
- radial: perceptual limits
  - angles lower precision than lengths
  - asymmetry between angle and length
    - can be exploited!

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

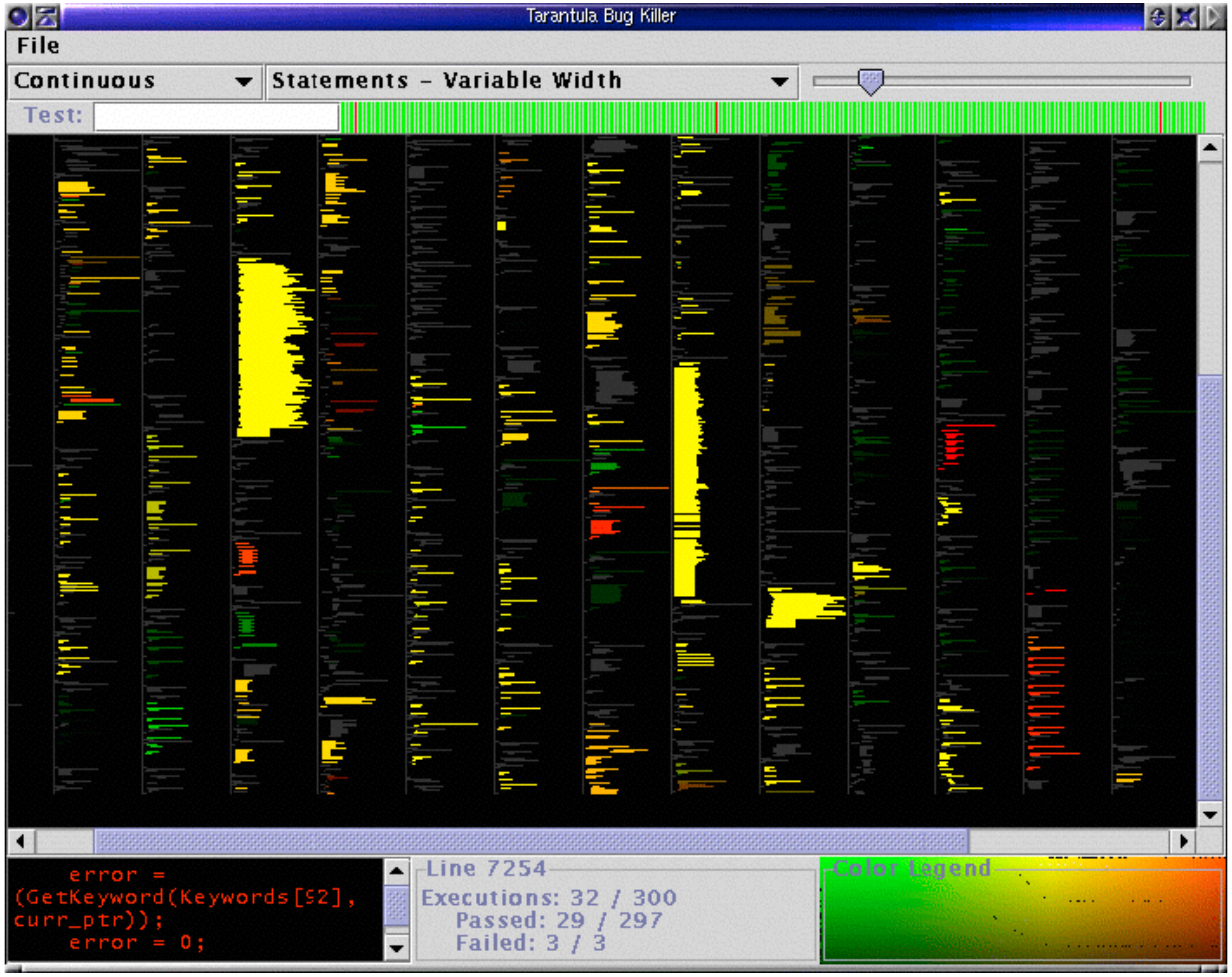


Layout Density

- Dense



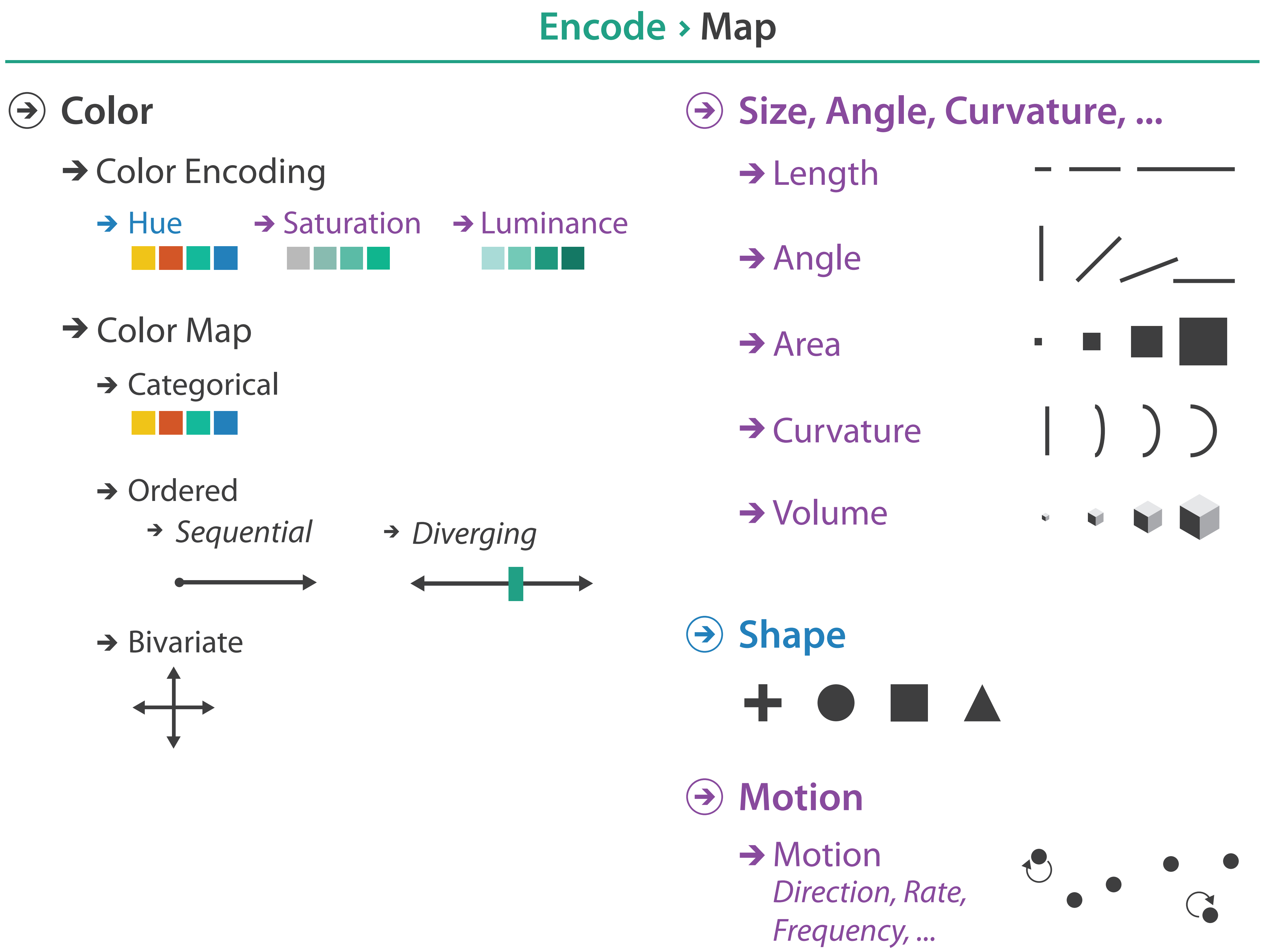
dense software overviews



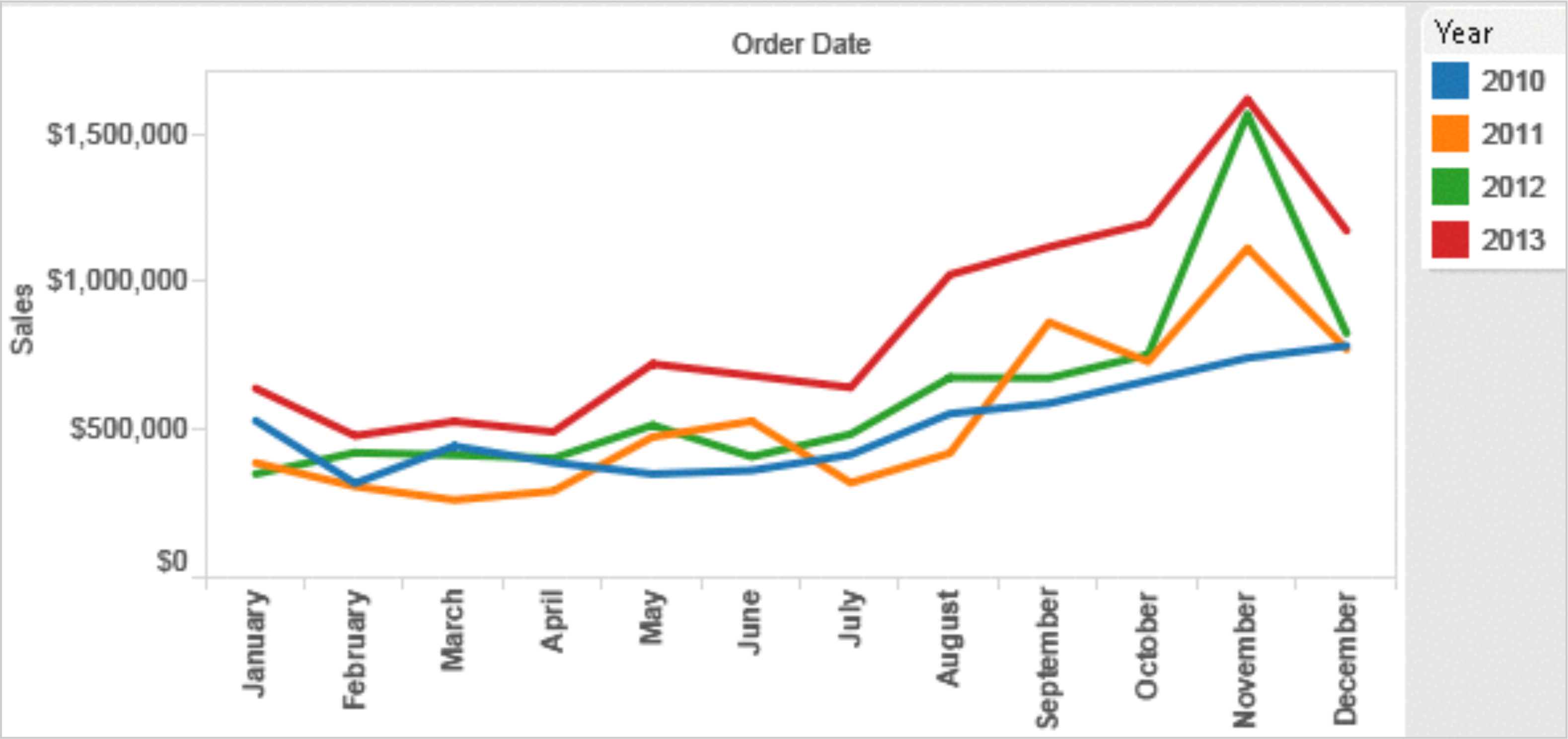
[Visualization of test information to assist fault localization. Jones, Harrold, Stasko. Proc. ICSE 2002, p 467-477.]

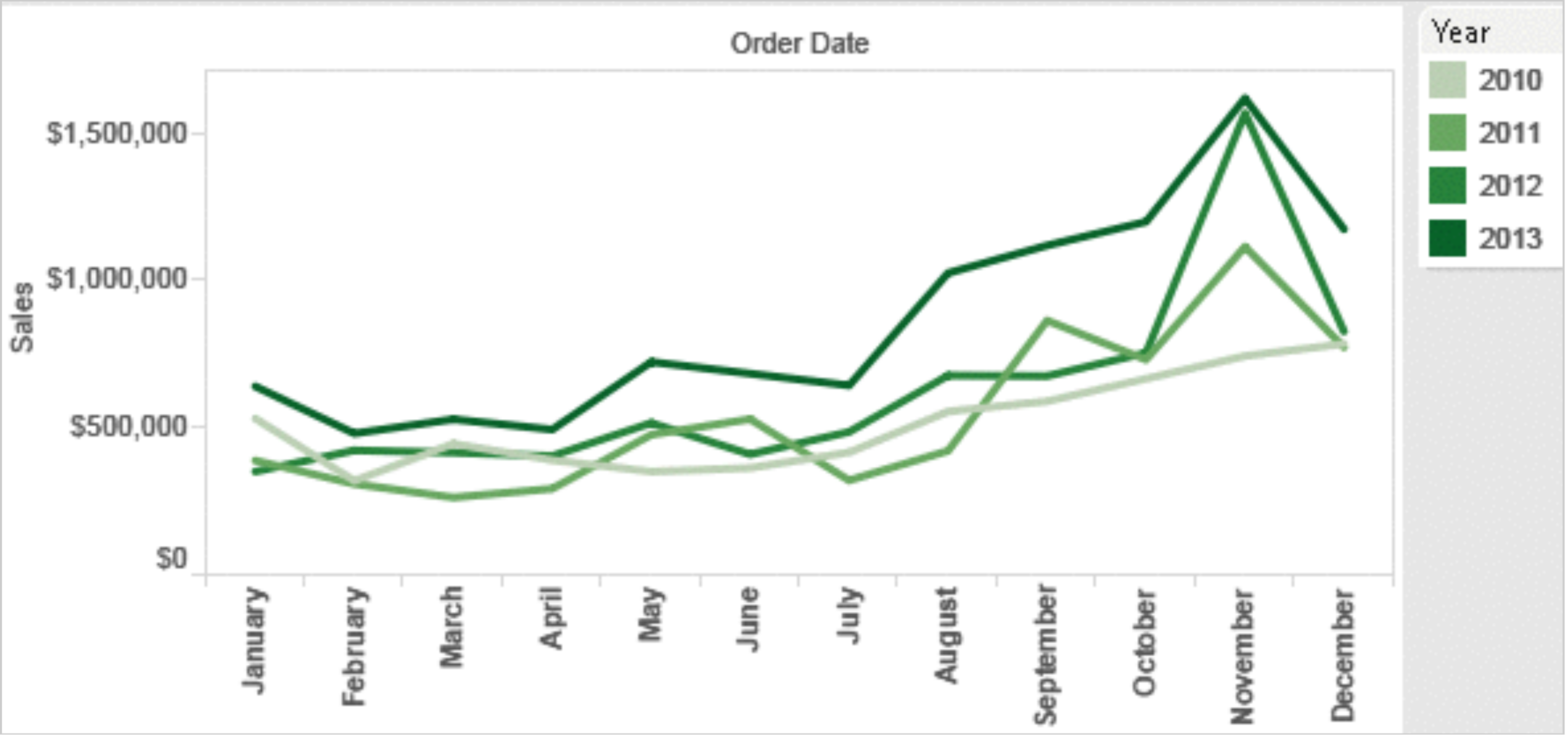
Ch 10: Map Color and Other Channels

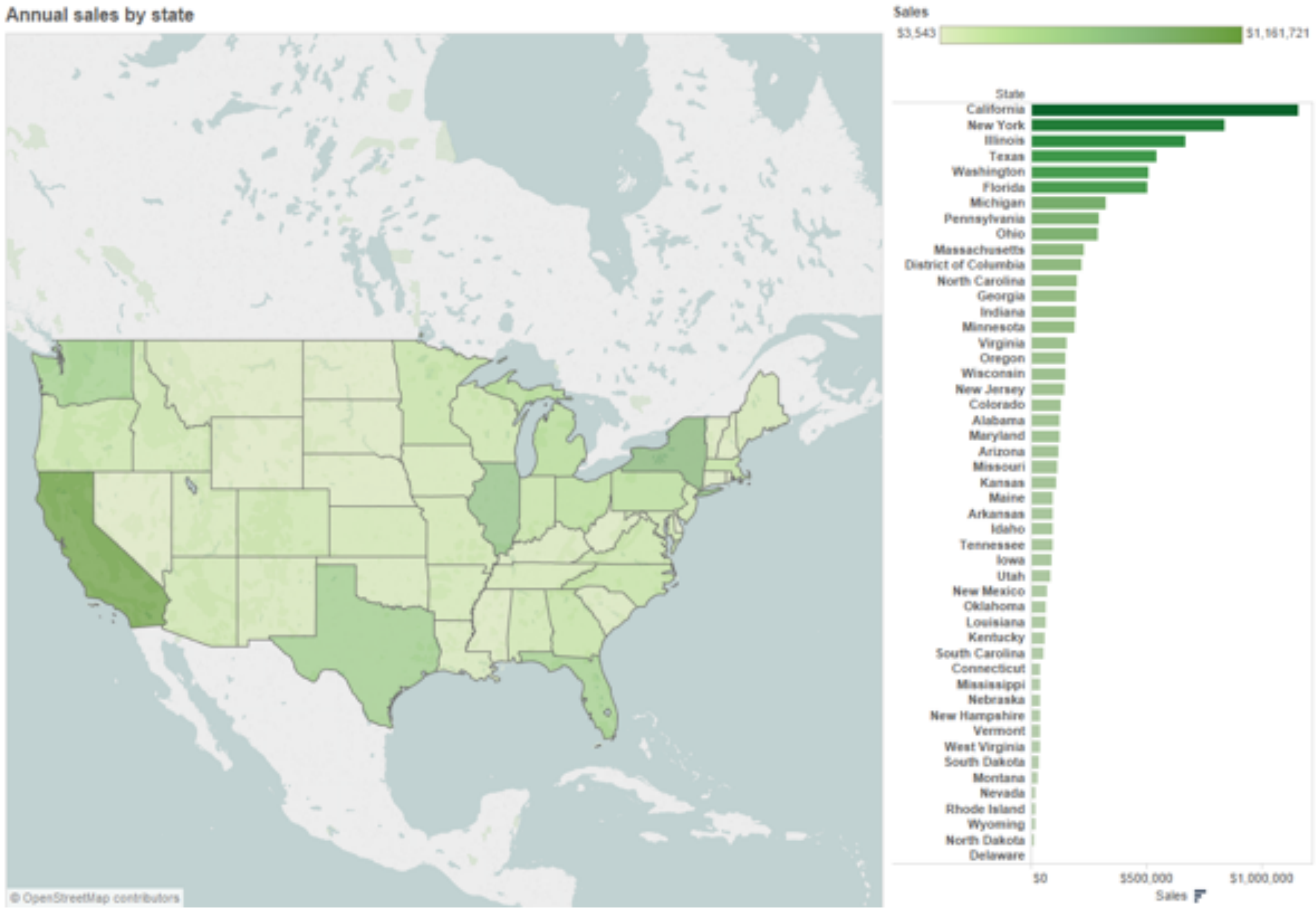
VAD Chap 10: Map Color and Other Channels



Categorical vs ordered color





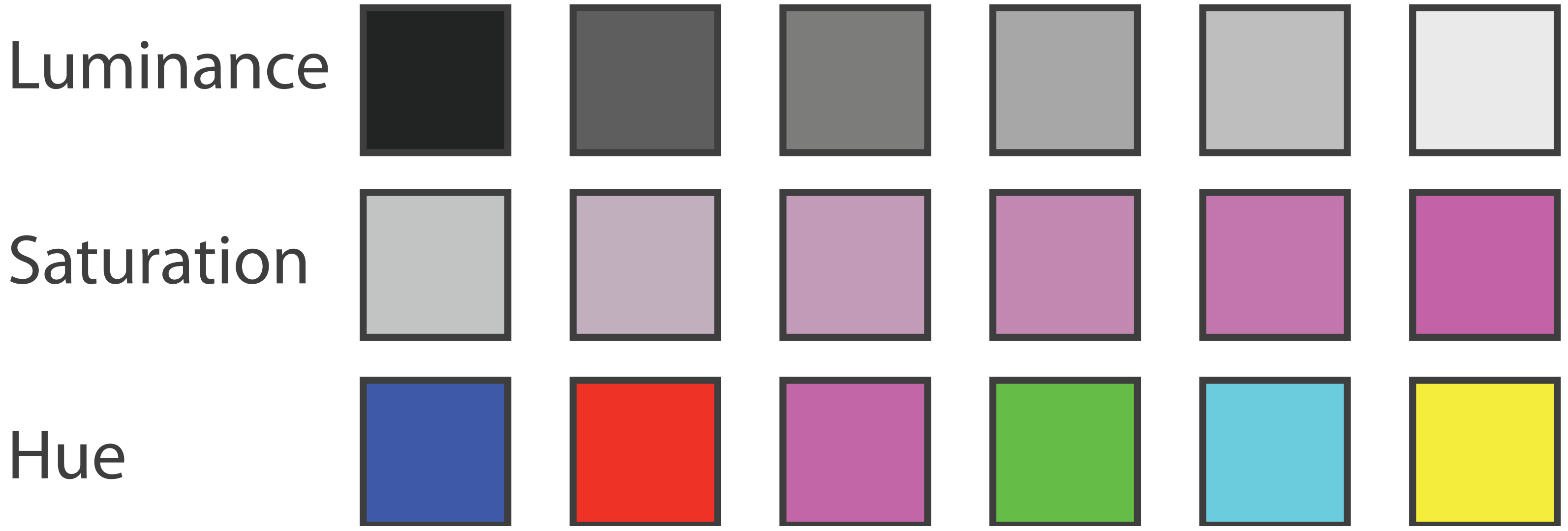


[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

Decomposing color

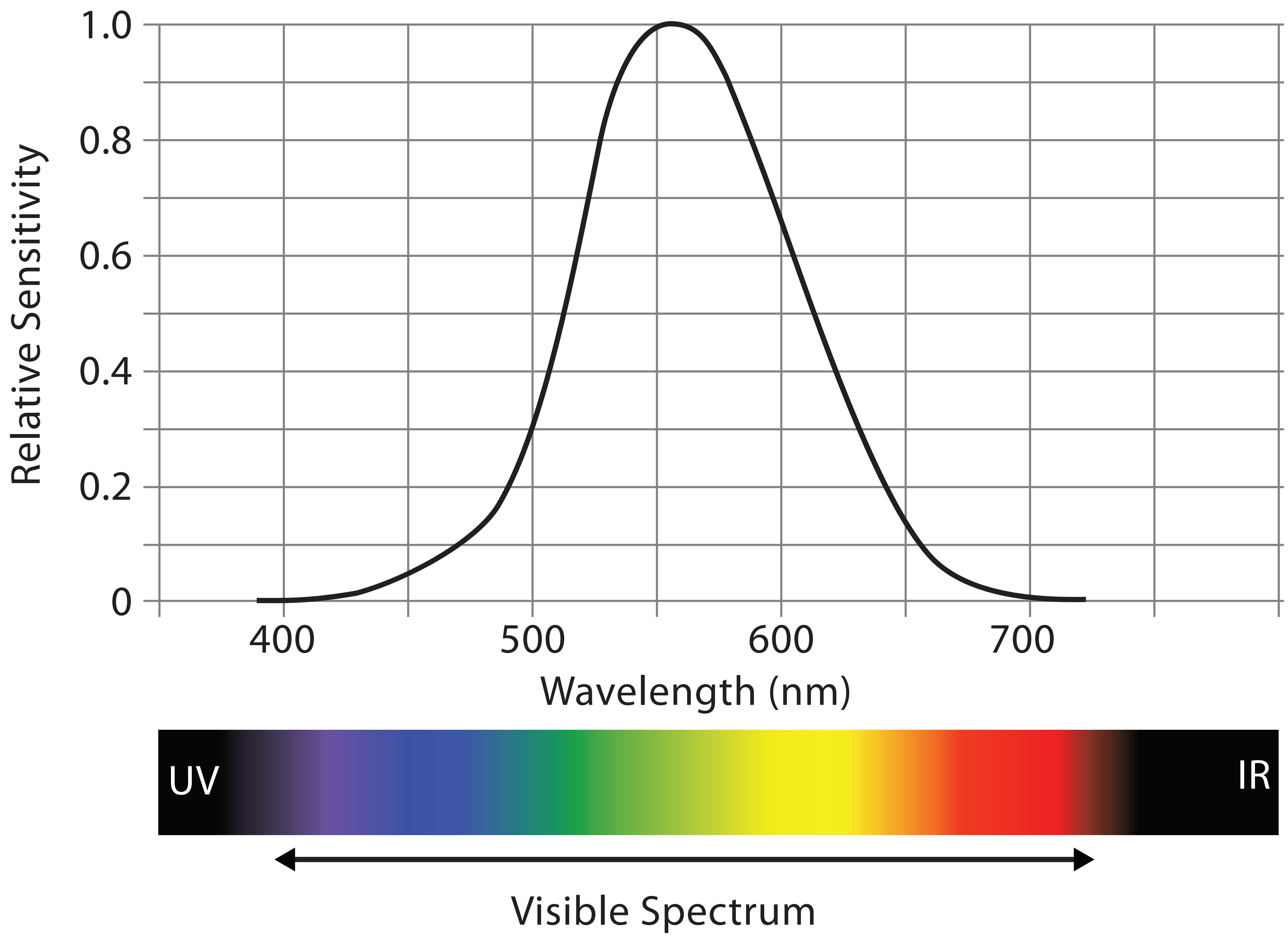
- first rule of color: do not talk about color!
  - color is confusing if treated as monolithic

- 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?

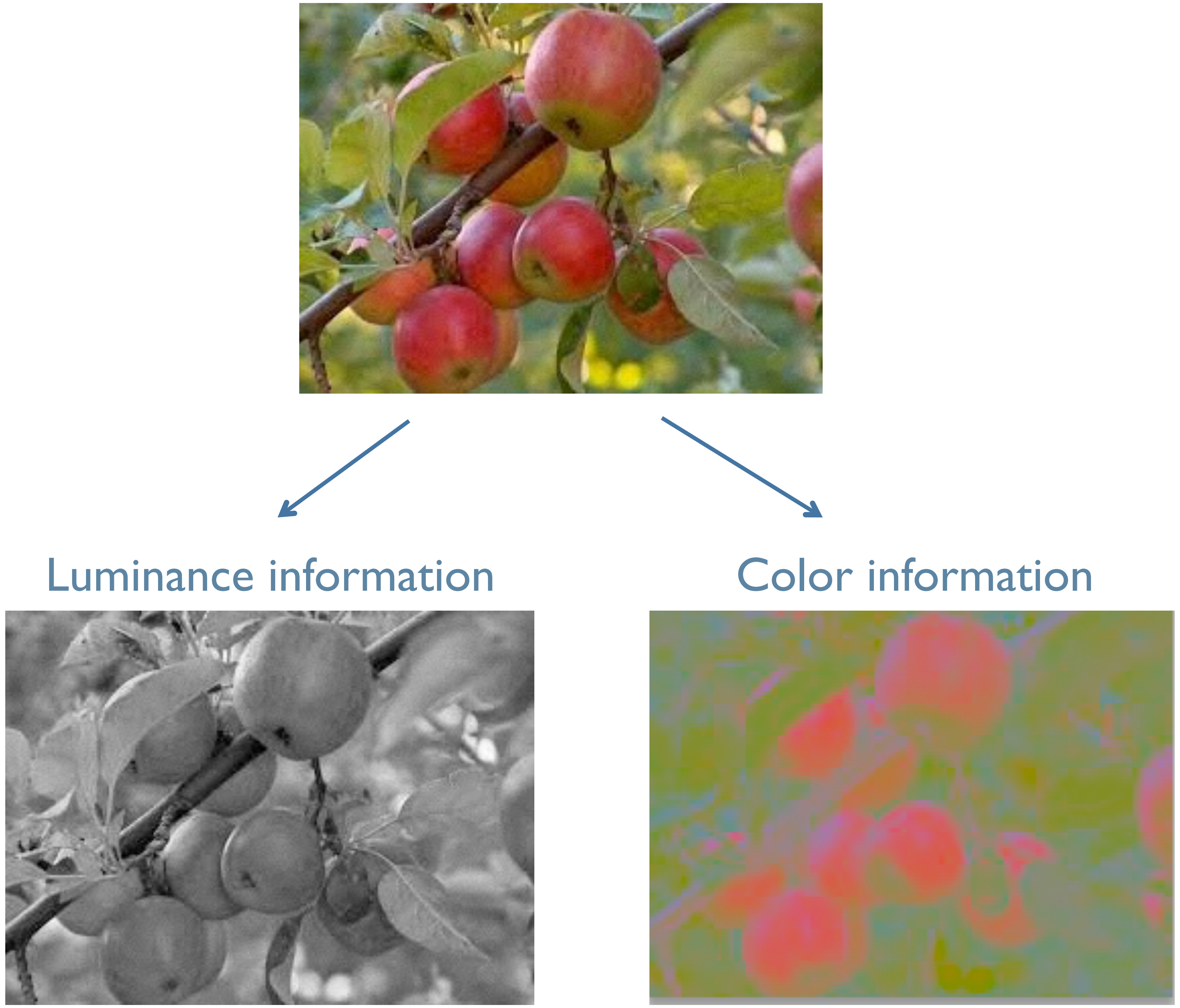
Spectral sensitivity



Luminance

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

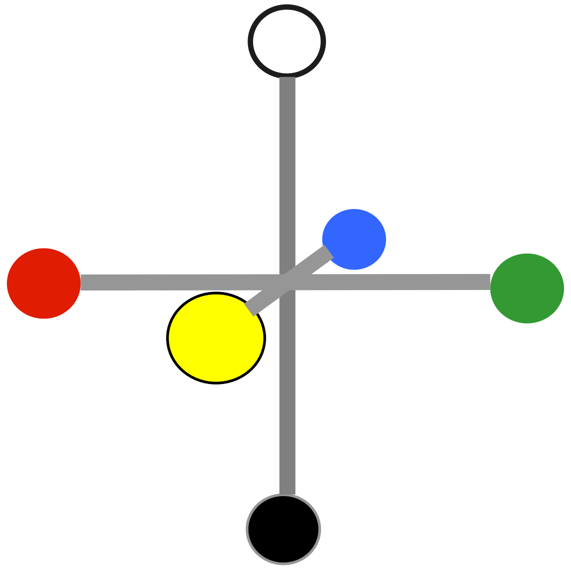
- intrinsic perceptual ordering

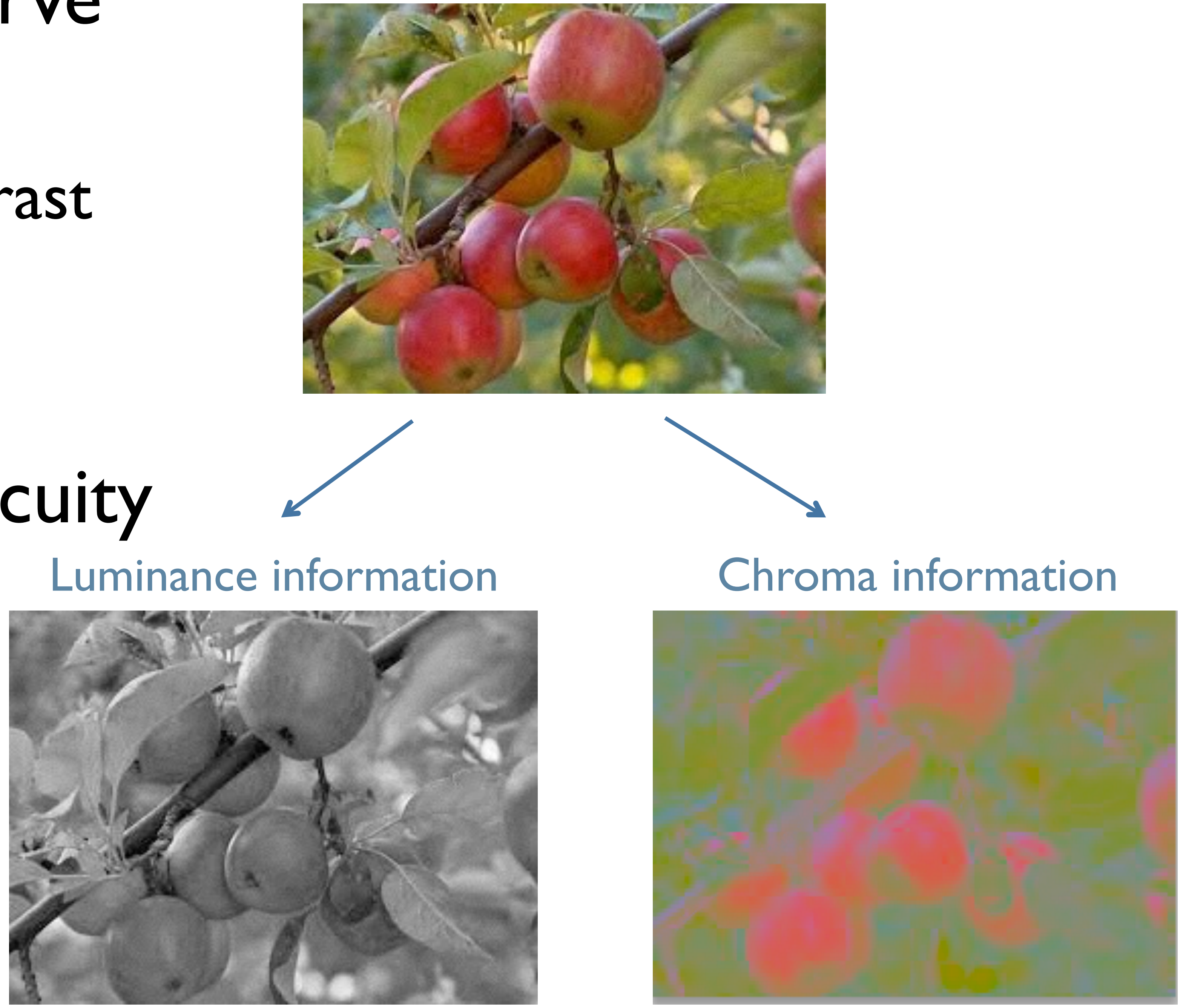


[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

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

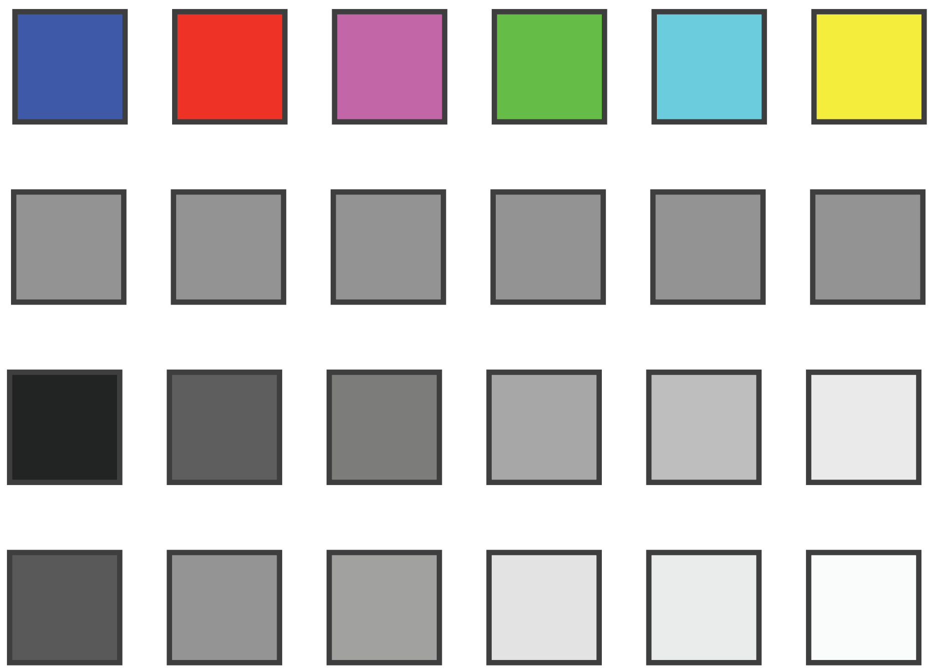
- CIE L*a*b*: good for computation
  - L* intuitive: perceptually linear luminance
  - a*b* axes: perceptually linear but nonintuitive
- RGB: good for display hardware
  - poor for encoding
- HSL/HSV: somewhat better for encoding
  - hue/saturation wheel intuitive
  - beware: only pseudo-perceptual!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

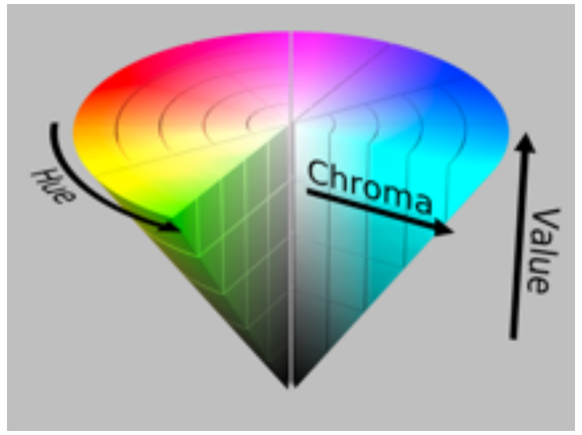
Corners of the RGB color cube

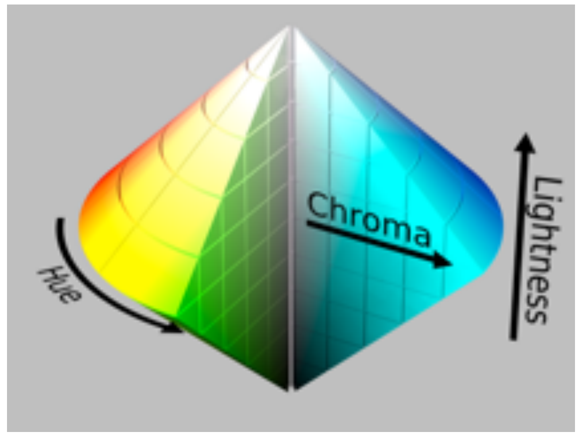
L from HLS
All the same

Luminance values

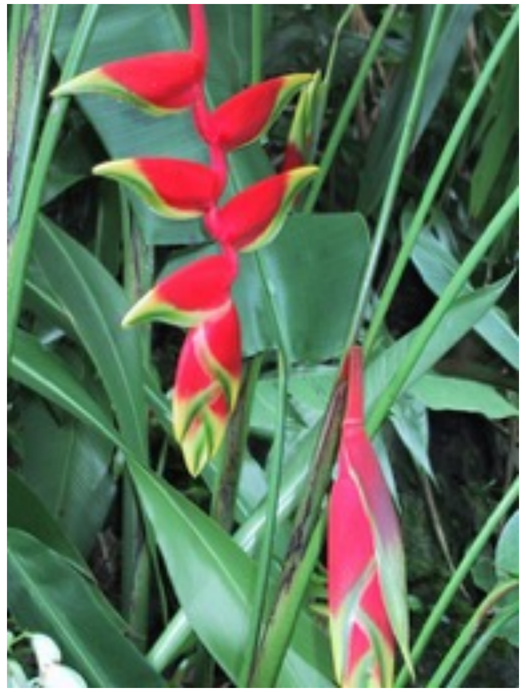
L* values



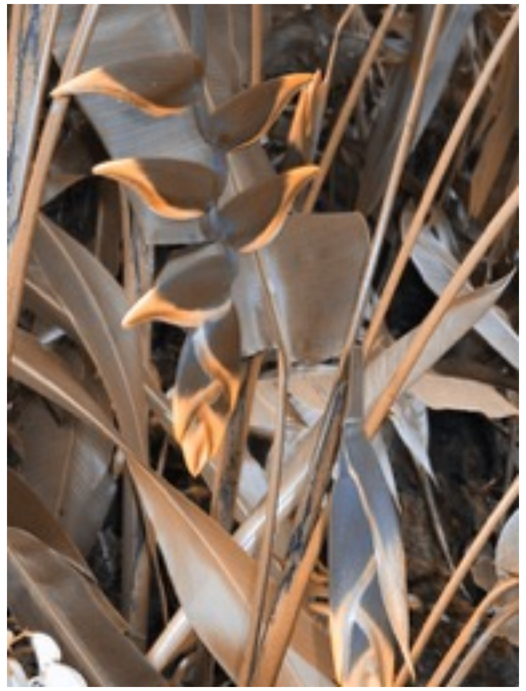




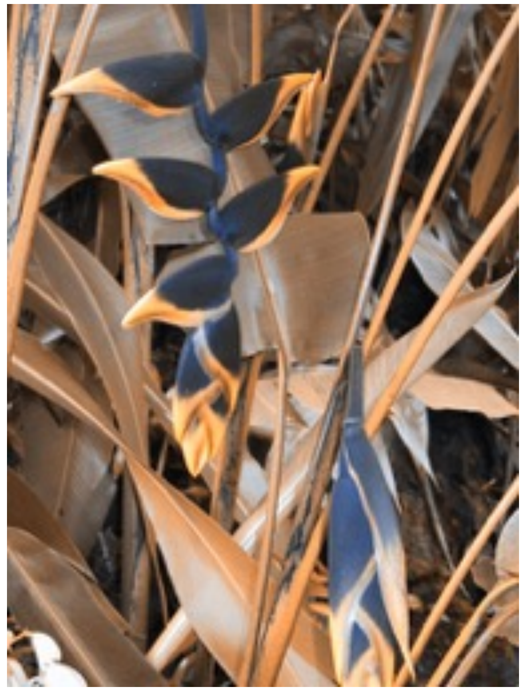
Designing for color deficiency: Check with simulator



Normal vision

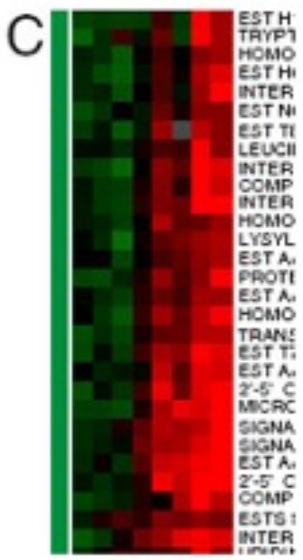


Deuteranope



Protanope

Tritanope

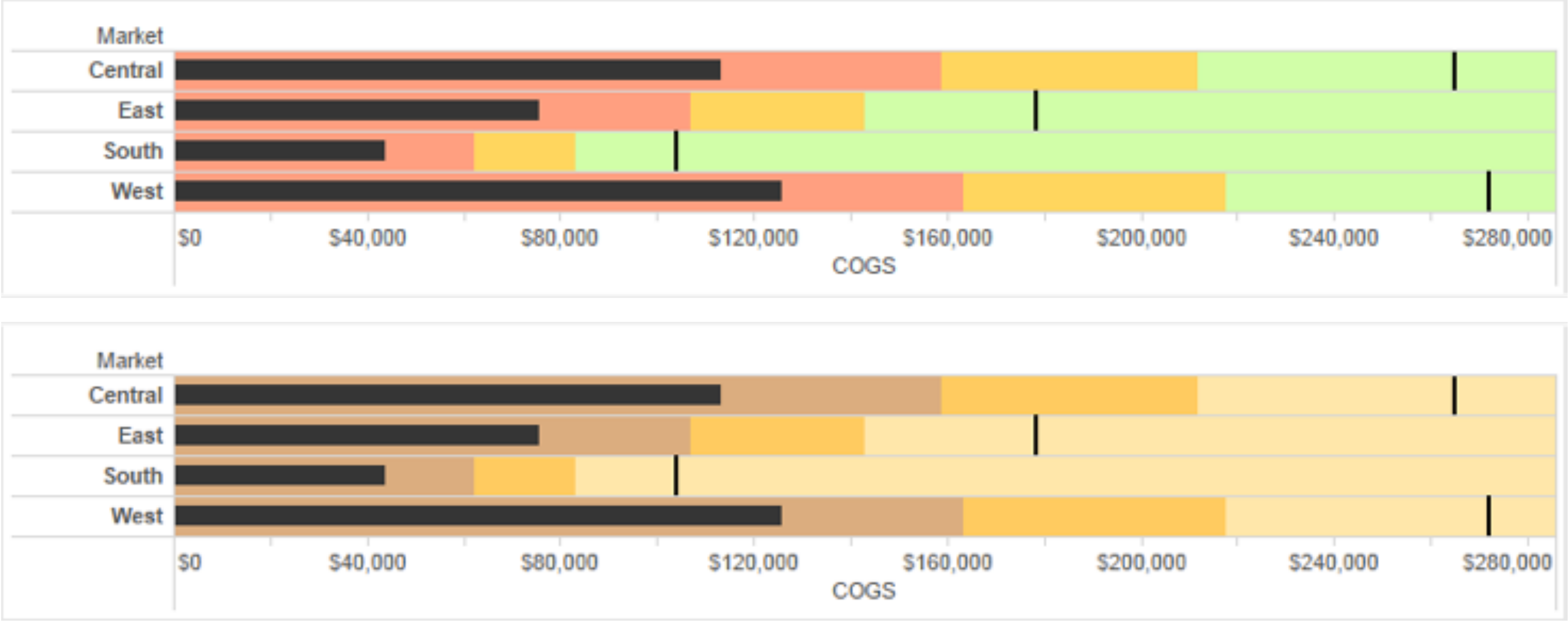


http://rehue.net

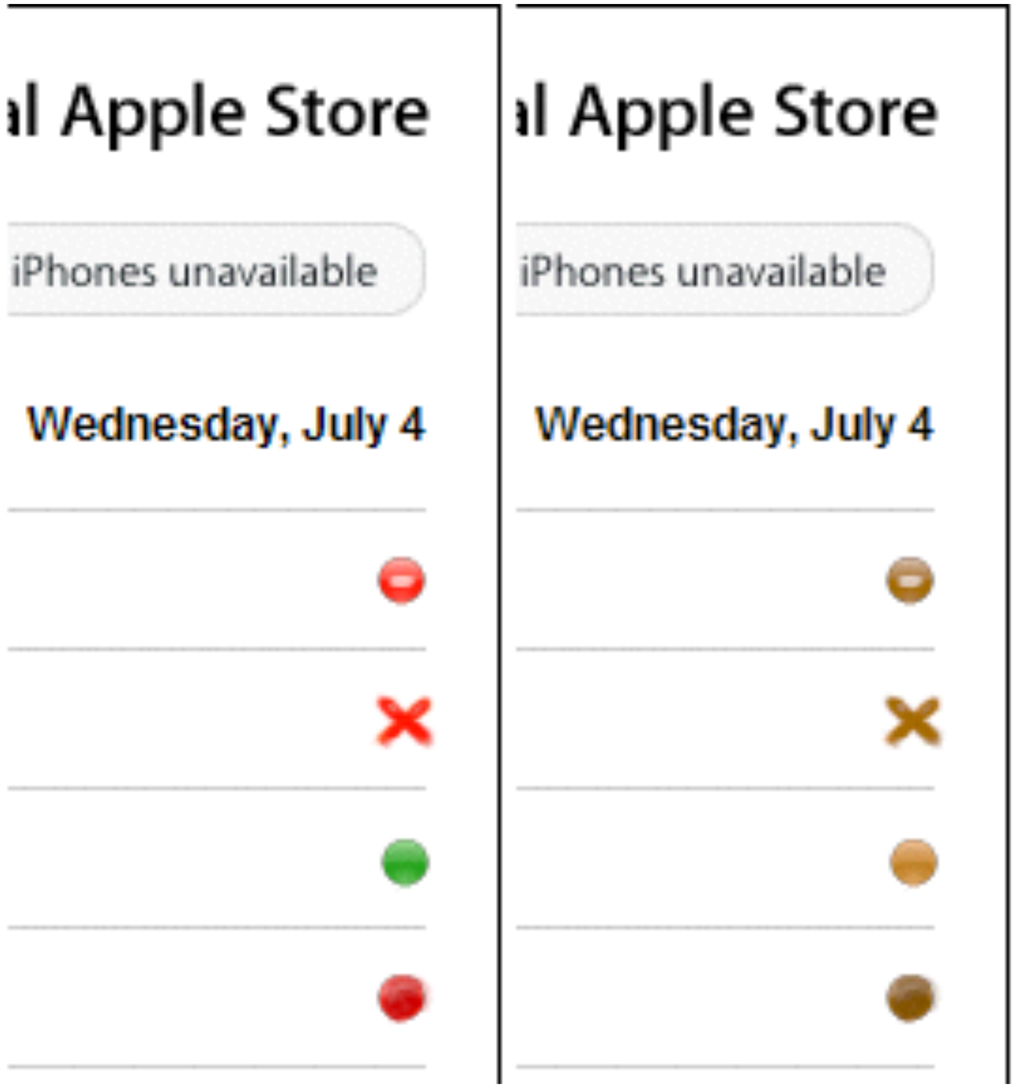
[Seriously Colorful: Advanced Color Principles & Practices. Stone. Tableau Customer Conference 2014.]

Designing for color deficiency: Avoid encoding by hue alone

- redundantly encode
  - vary luminance
  - change shape



Deuteranope simulation

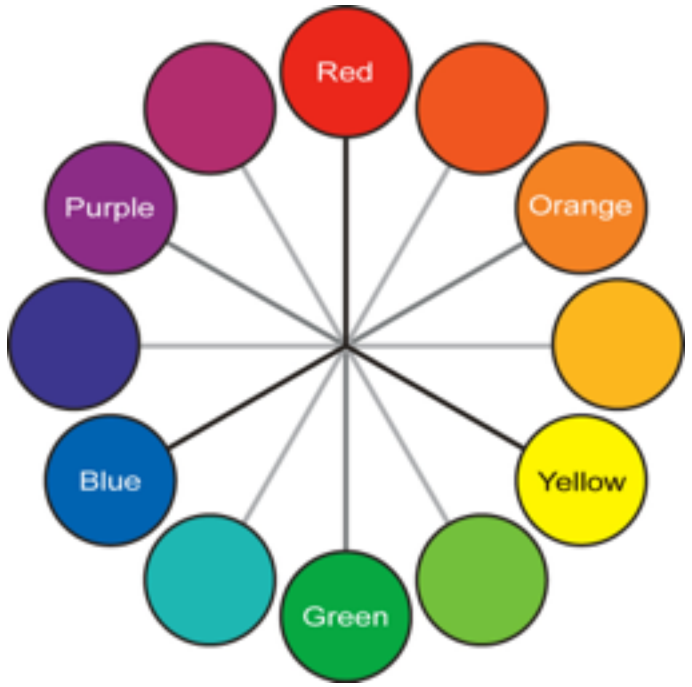


Change the shape

Vary luminance

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

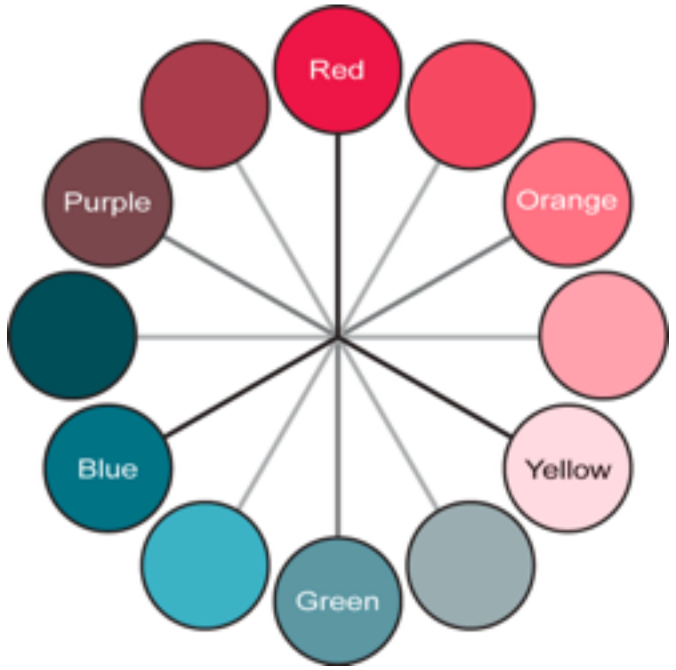
Color deficiency: Reduces color to 2 dimensions



Normal

Protanope

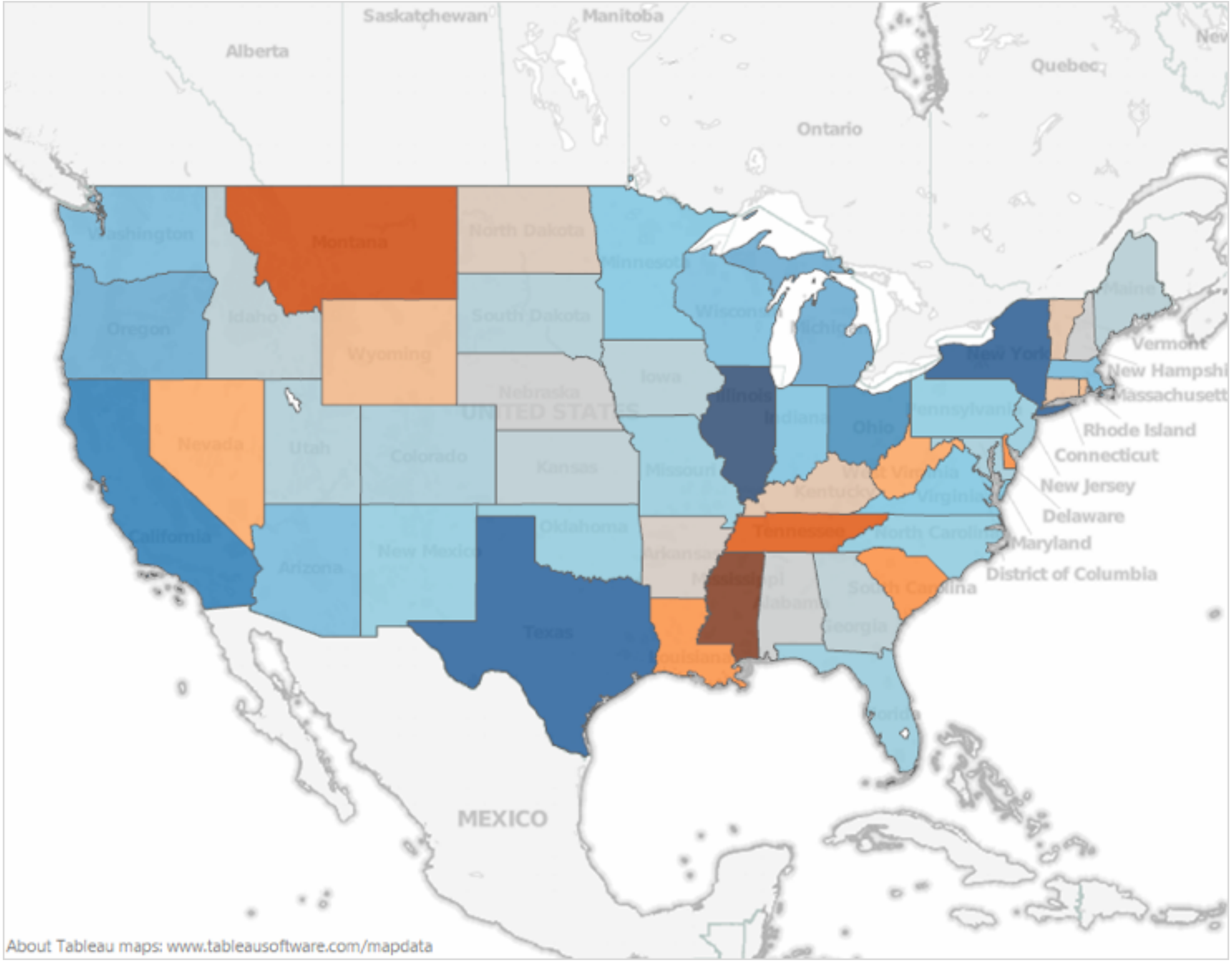
Deuteranope



Tritanope

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

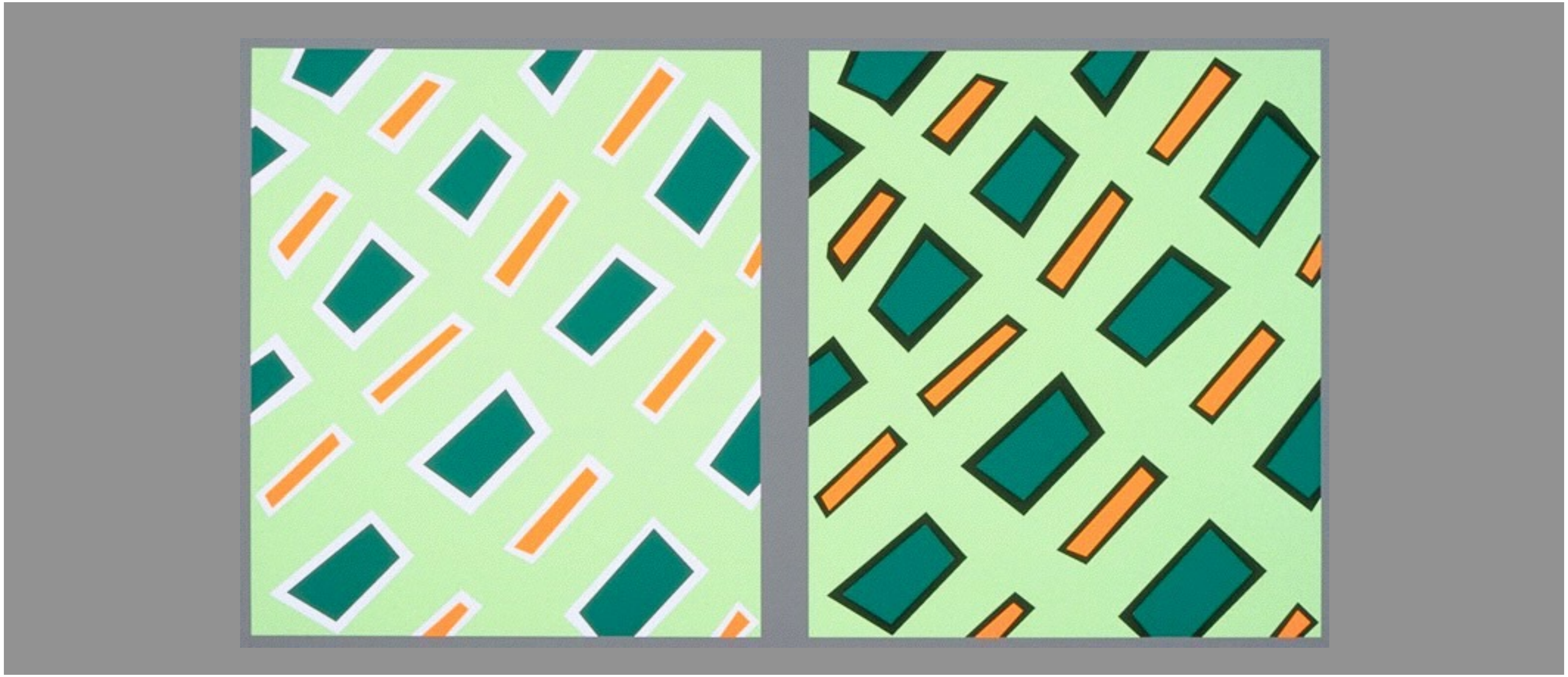
Designing for color deficiency: Blue-Orange is safe



[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.]

Color/Lightness constancy: Illumination conditions

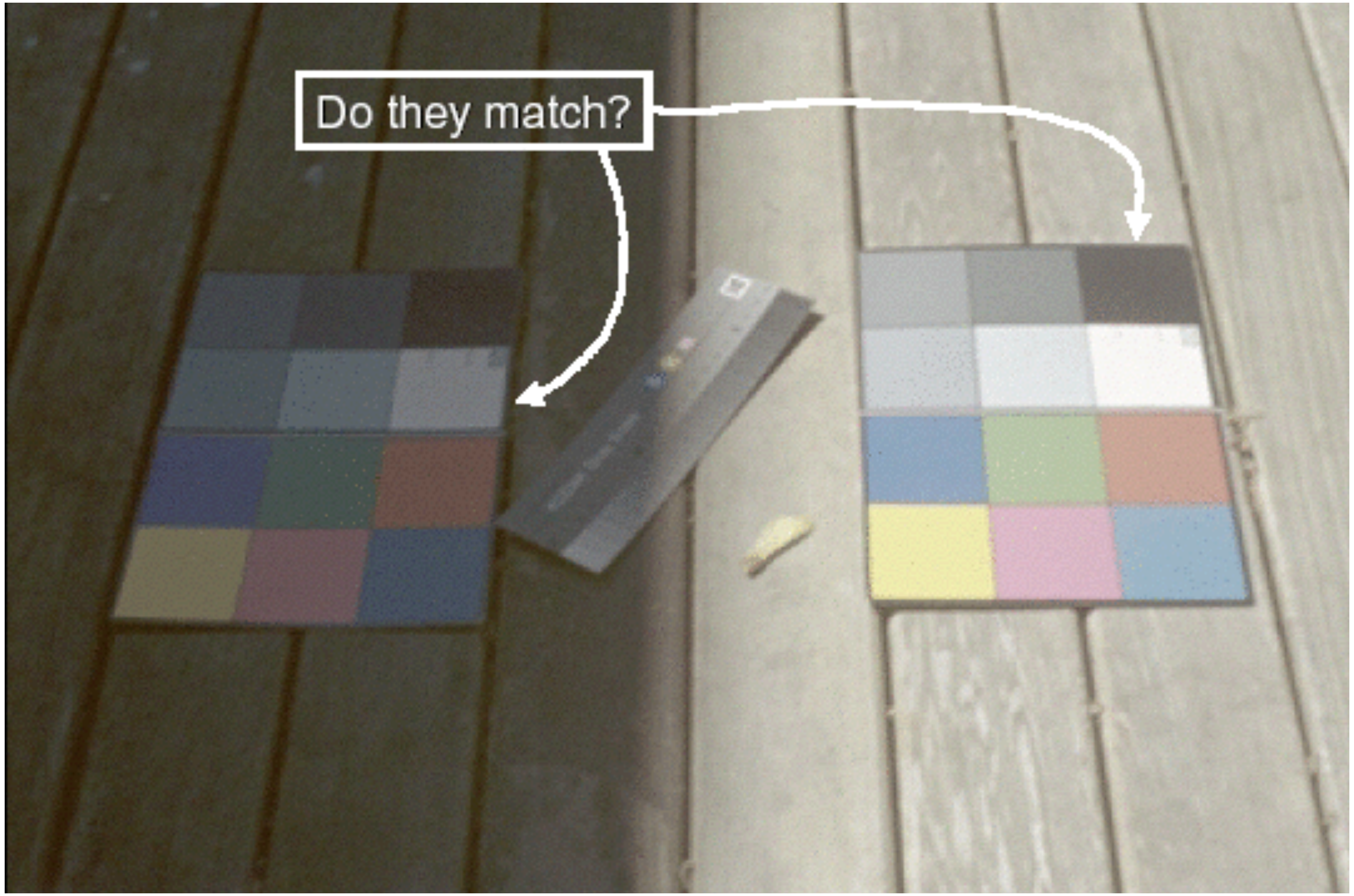


Image courtesy of John McCann

Color/Lightness constancy: Illumination conditions

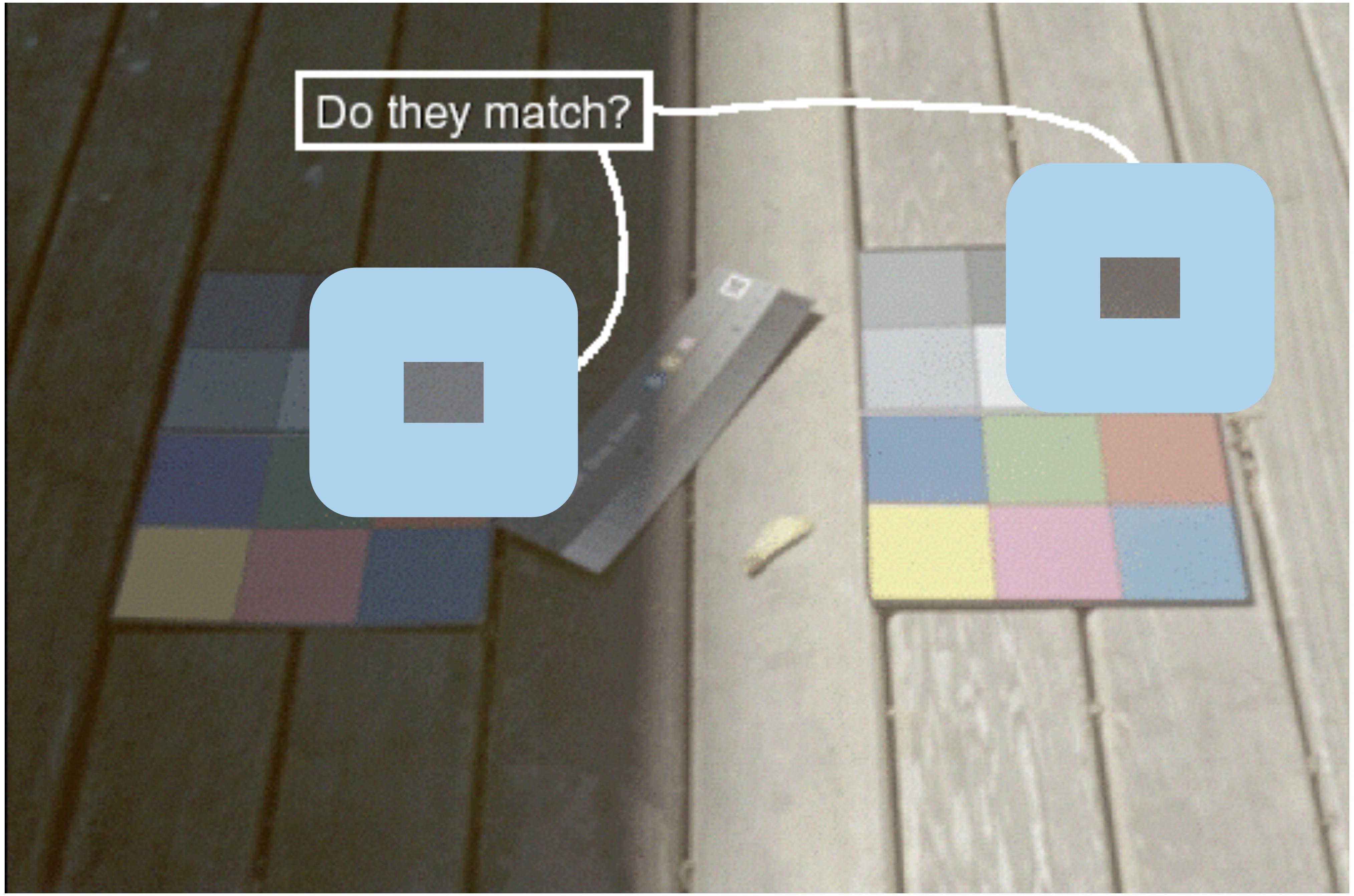
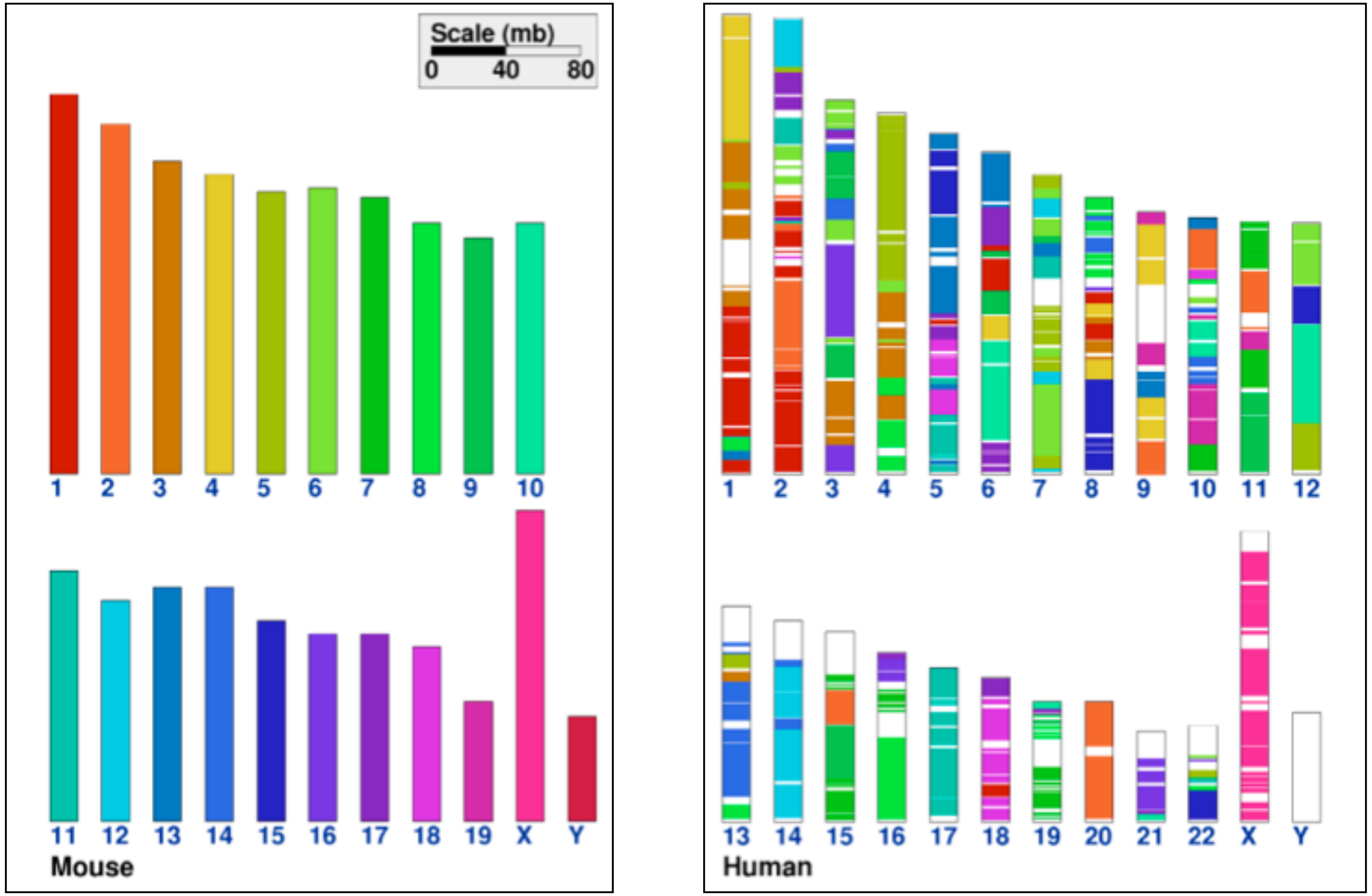


Image courtesy of John McCann

Categorical color: limited number of discriminable bins

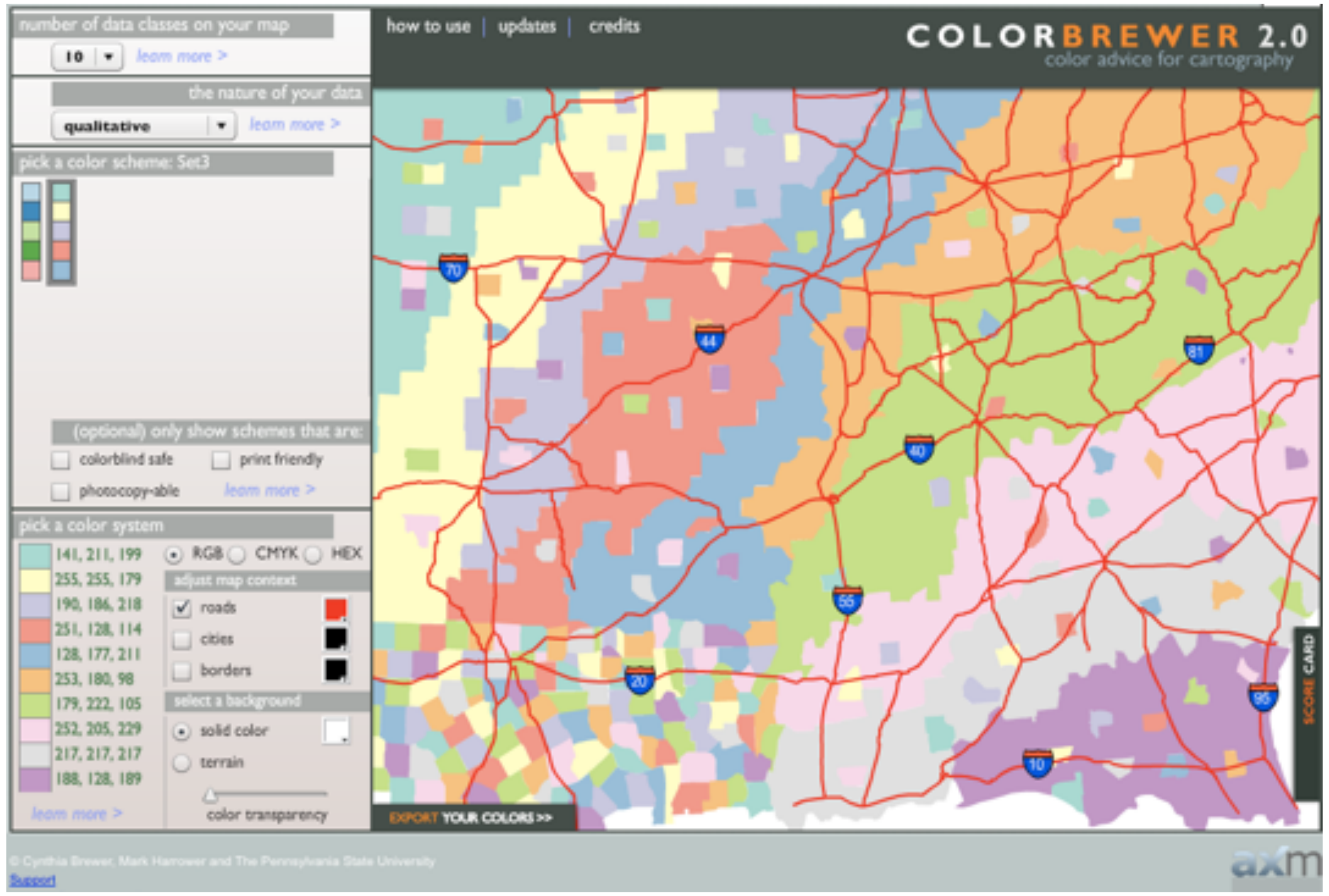
- human perception built on relative comparisons
  - great if color contiguous
  - surprisingly bad for absolute comparisons
- noncontiguous small regions of color
  - fewer bins than you want
  - rule of thumb: 6-12 bins, including background and highlights

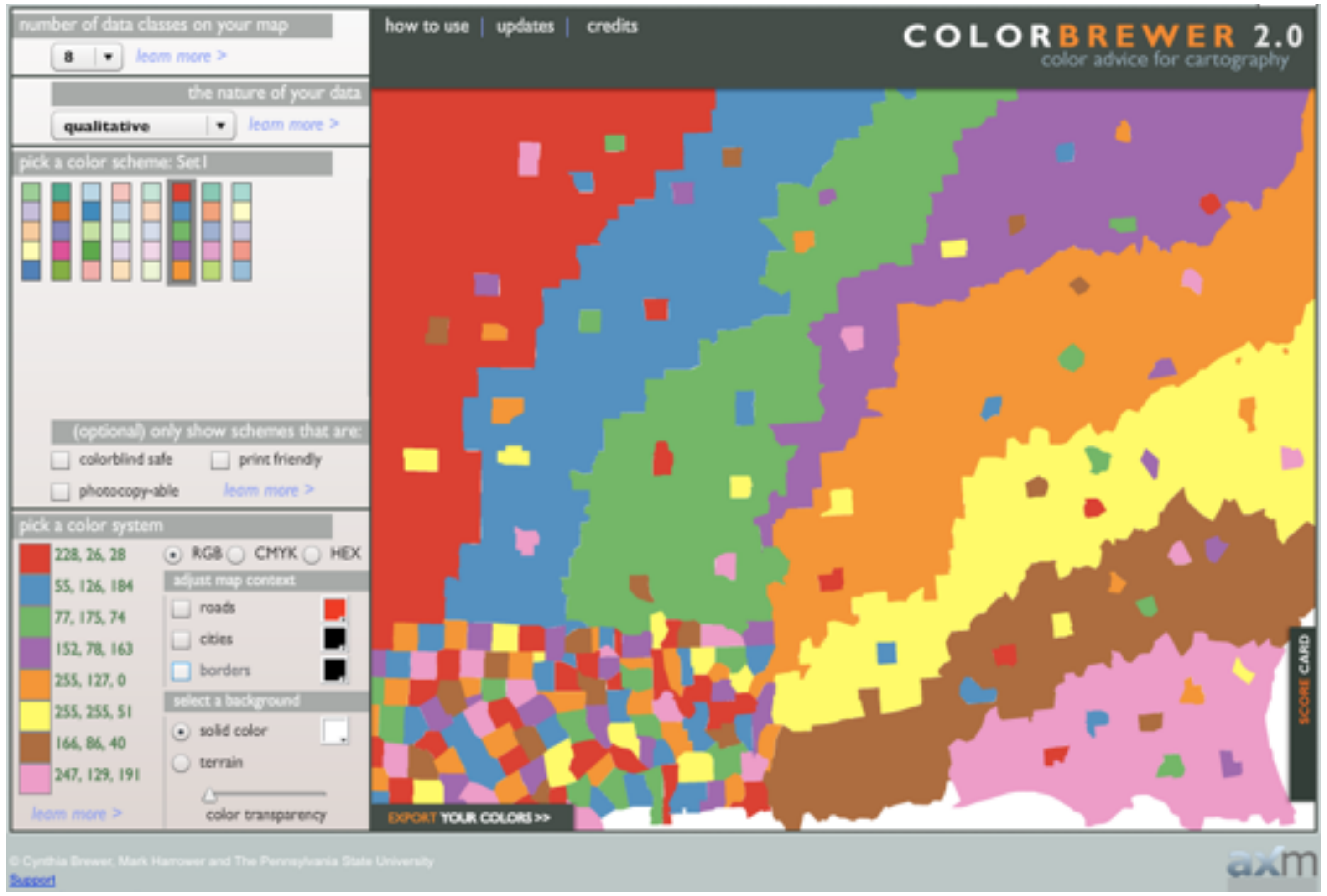


[Cinteny: flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. Sinha and Meller. BMC Bioinformatics, 8:82, 2007.]

ColorBrewer

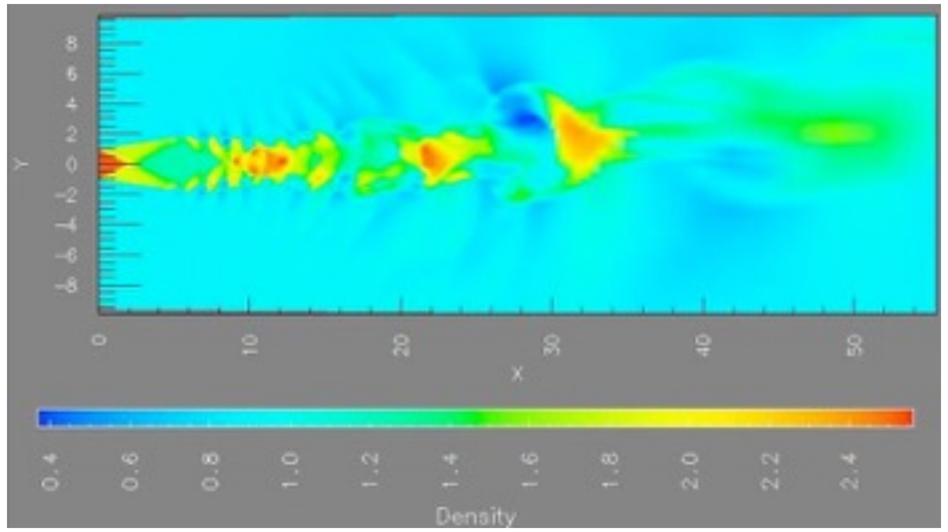
- http://www.colorbrewer2.org
- saturation and area example: size affects salience!



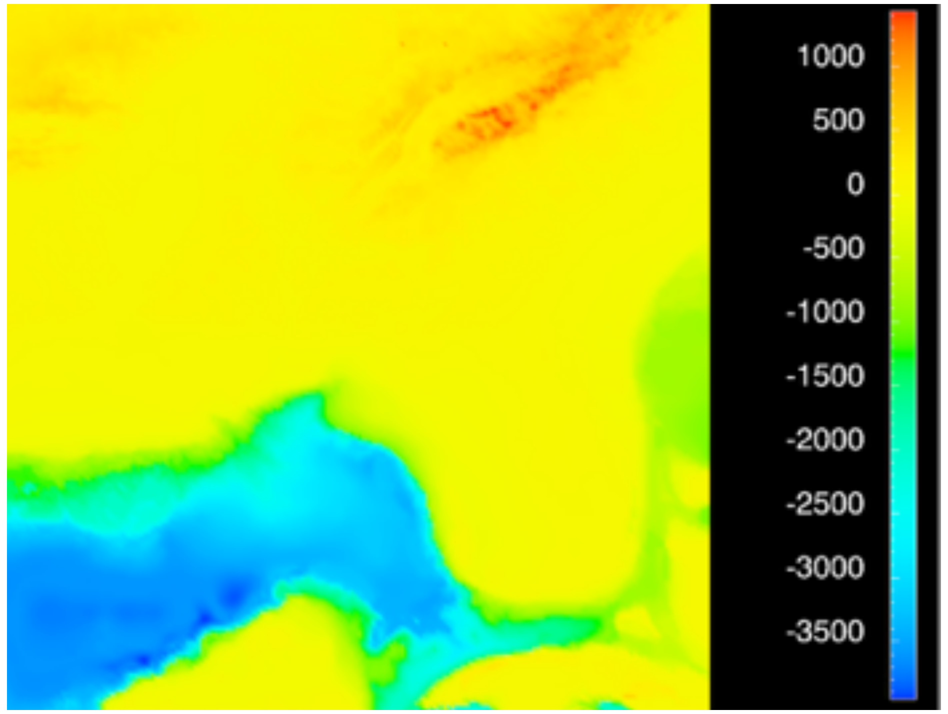


Ordered color: Rainbow is poor default

- 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.]

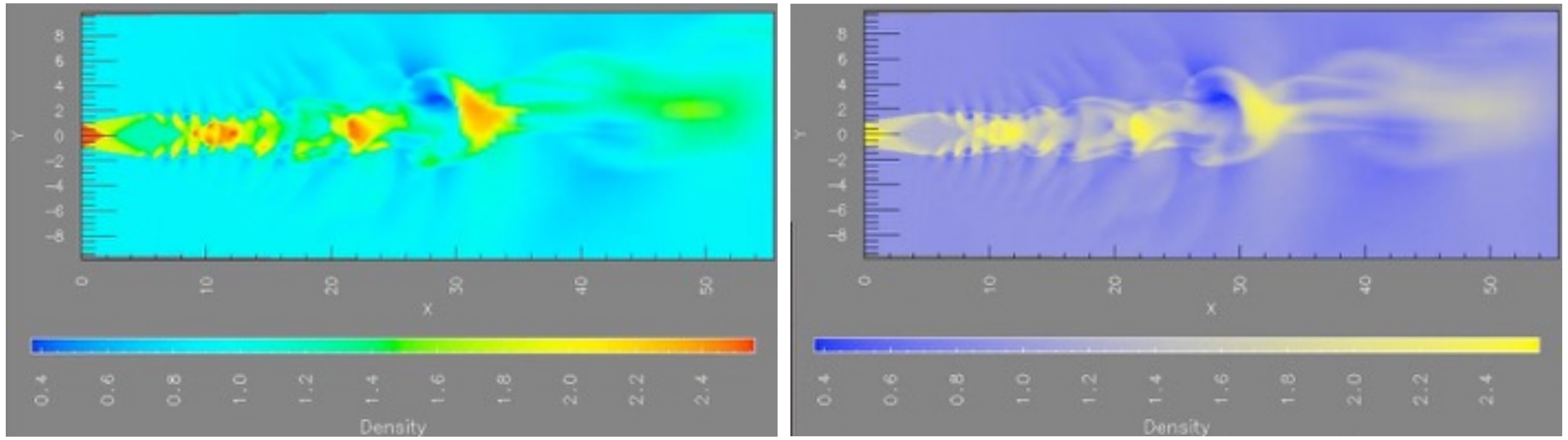


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

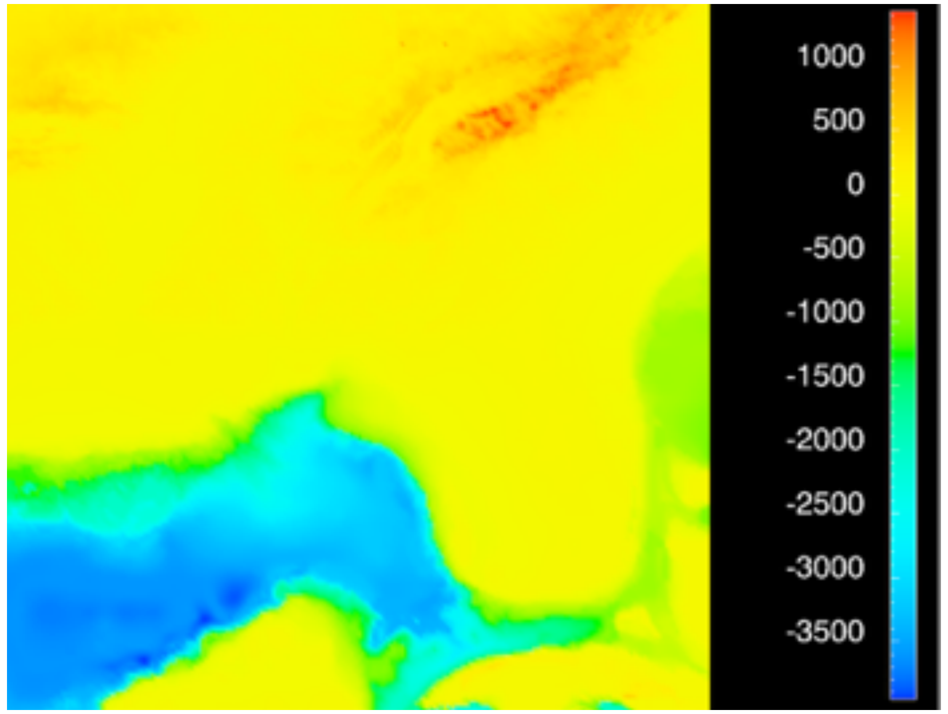
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

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.]

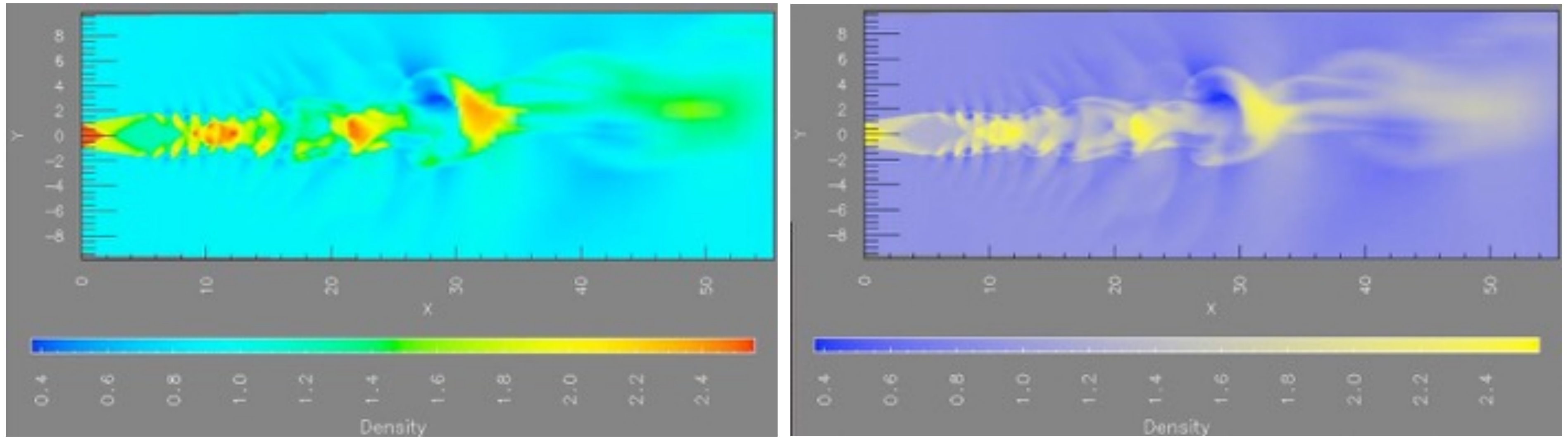


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

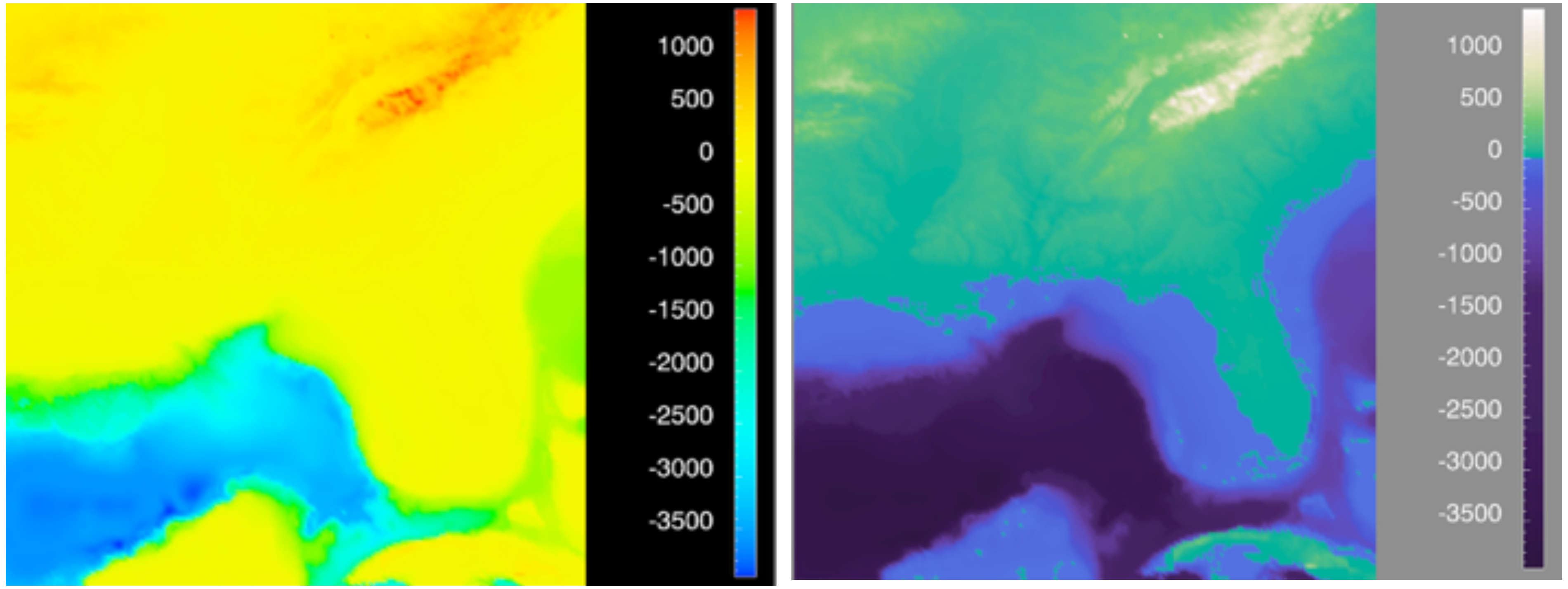
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

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.]

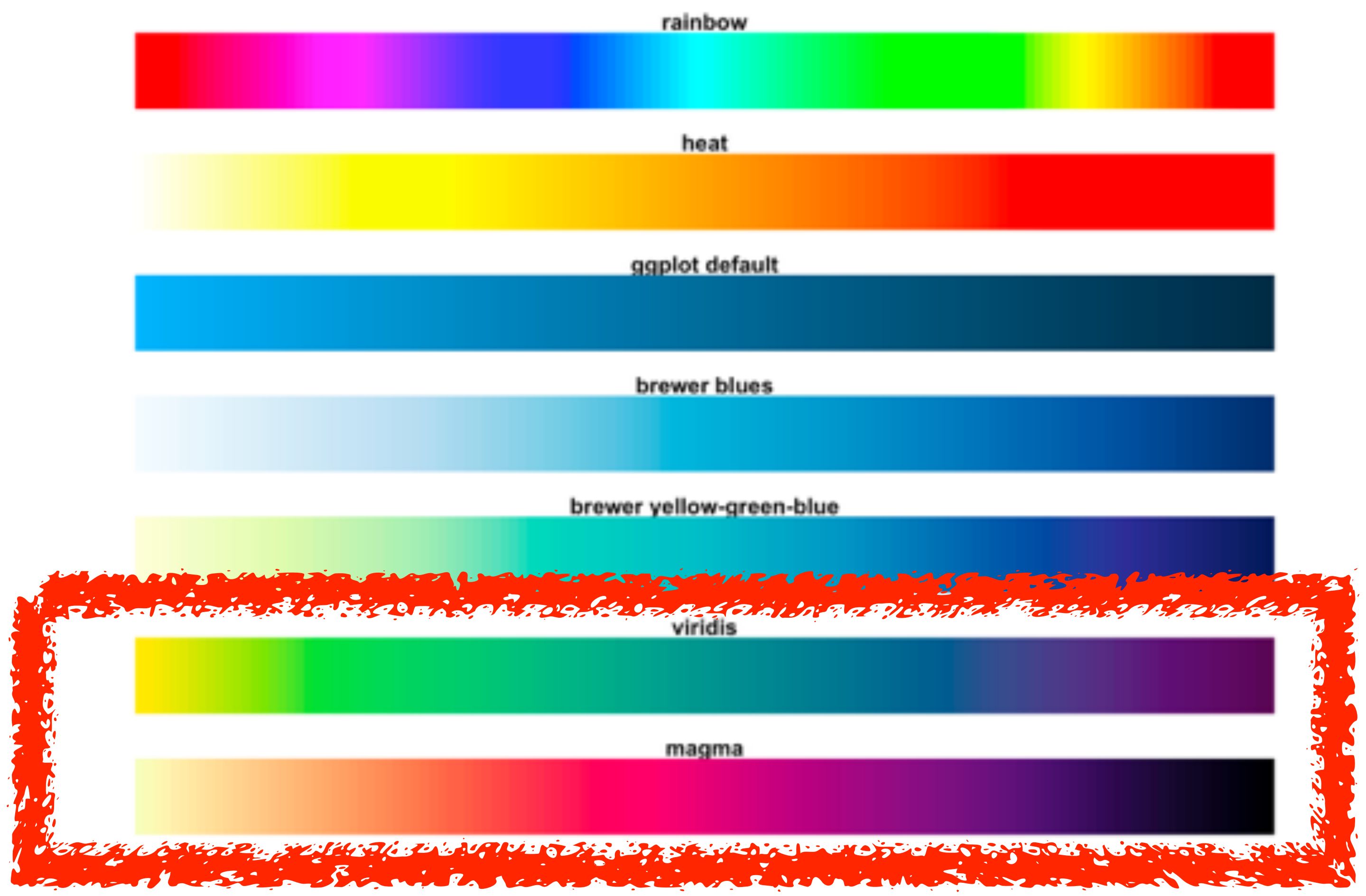


[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998. http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

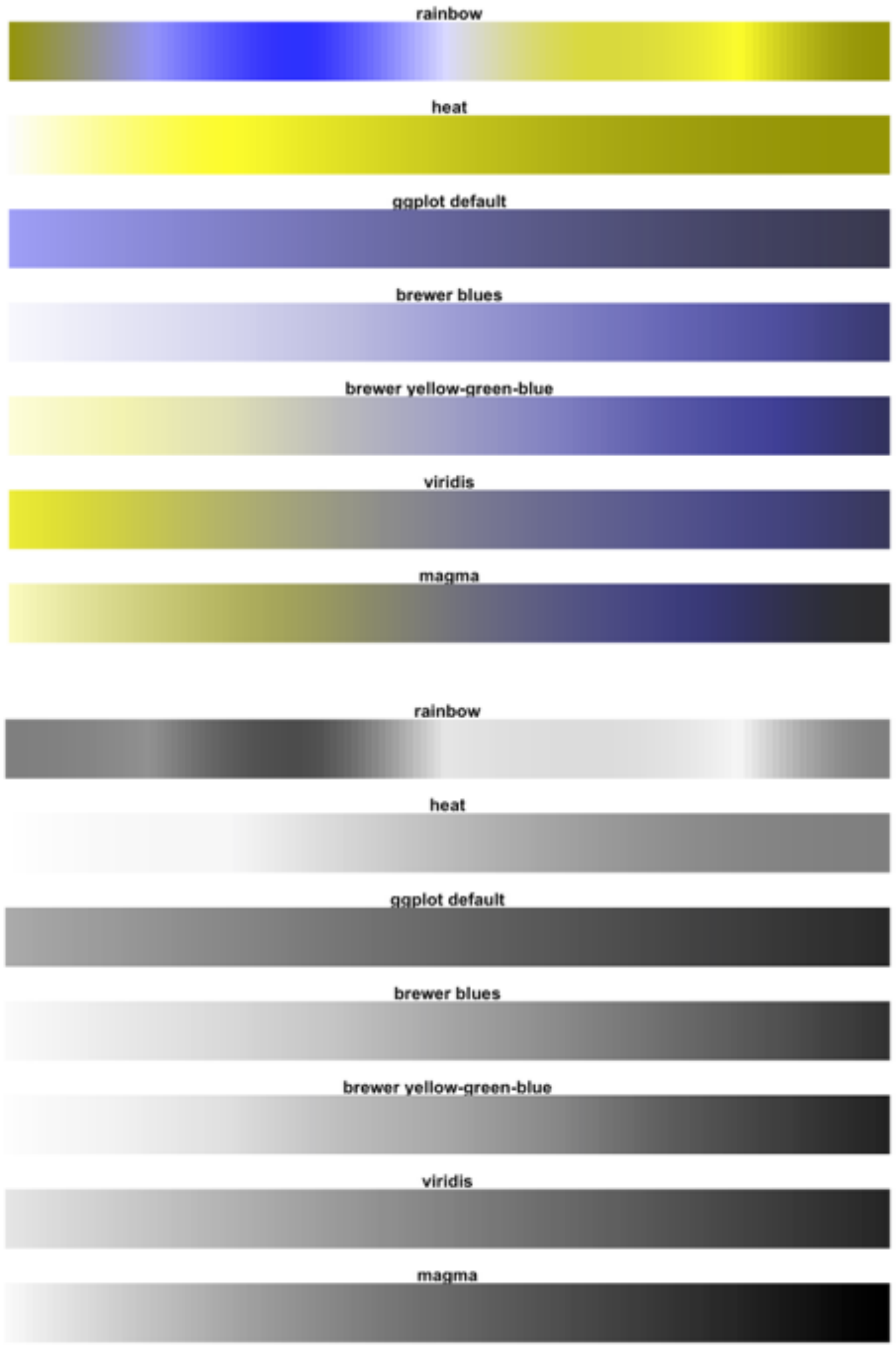
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]

Viridis

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

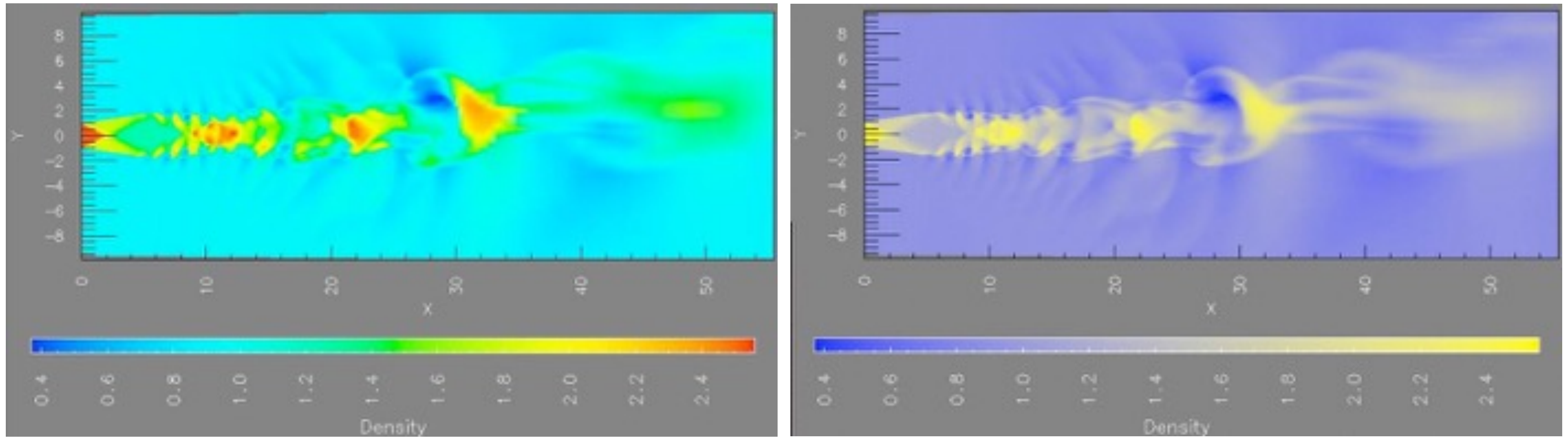


https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html

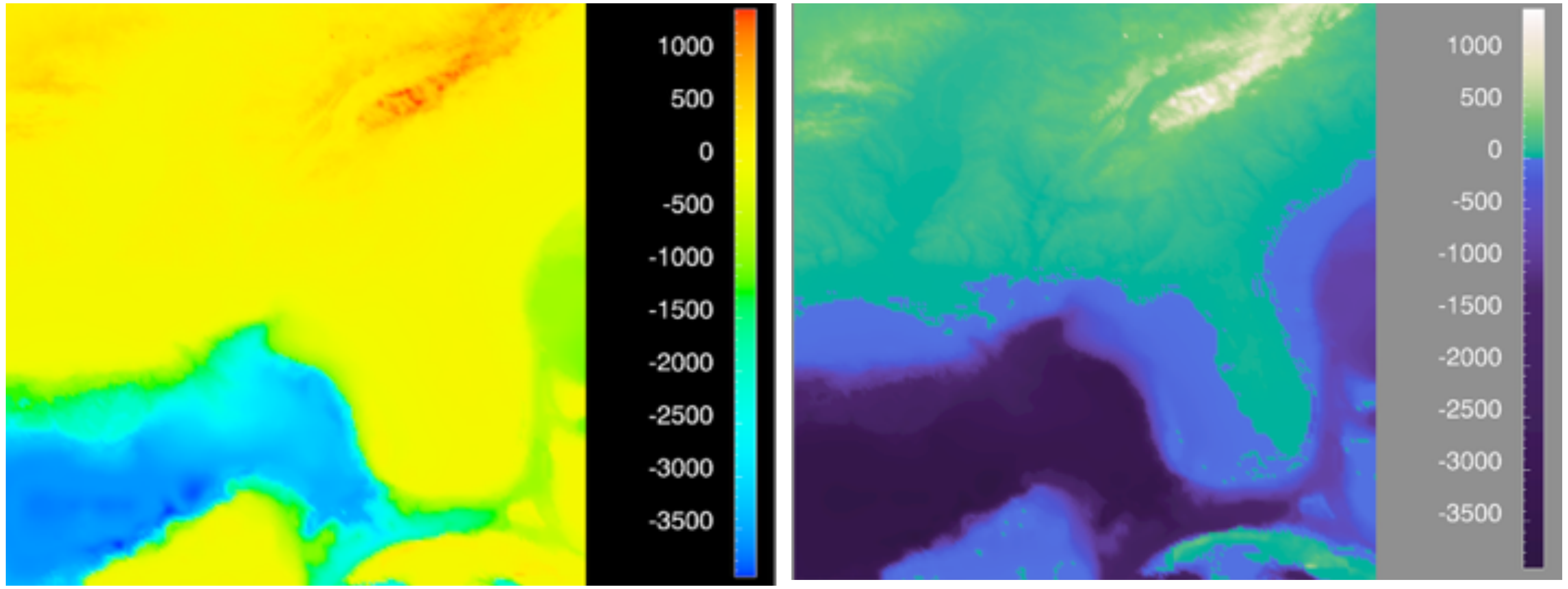


Ordered color: Rainbow is poor default

- 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

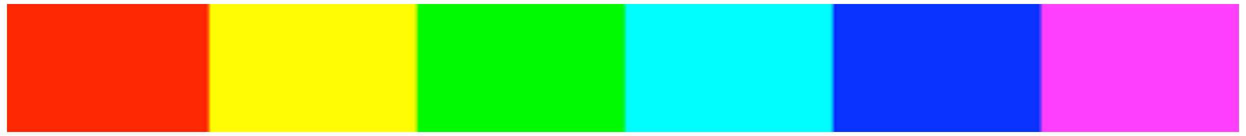


[A Rule-based Tool for Assisting Colormap Selection. Bergman,. Rogowitz, and. Treinish. Proc. IEEE Visualization (Vis), pp. 118–125, 1995.]



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

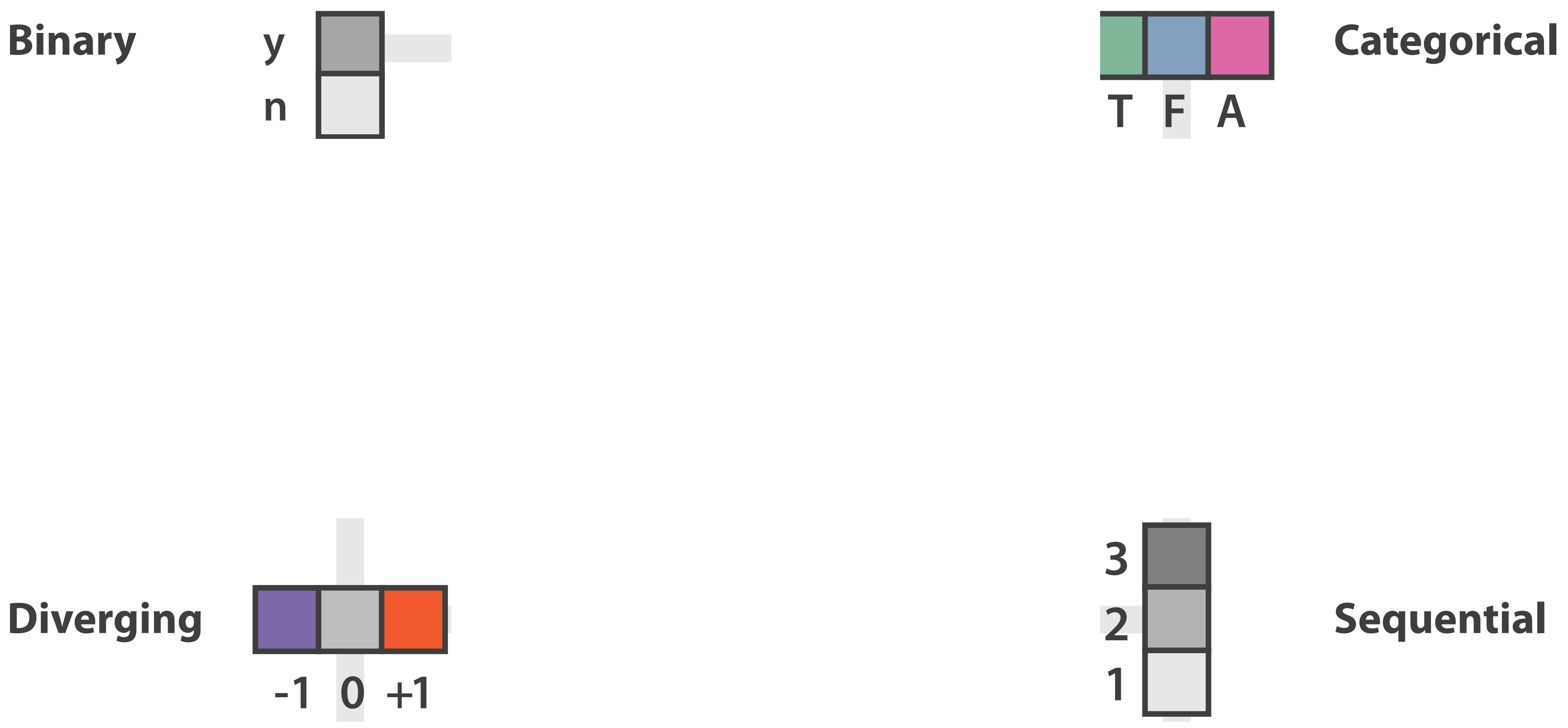
[Transfer Functions in Direct Volume Rendering: Design, Interface, Interaction. Kindlmann. SIGGRAPH 2002 Course Notes]



Colormaps

- ➔ Categorical

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



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

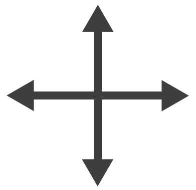
Colormaps

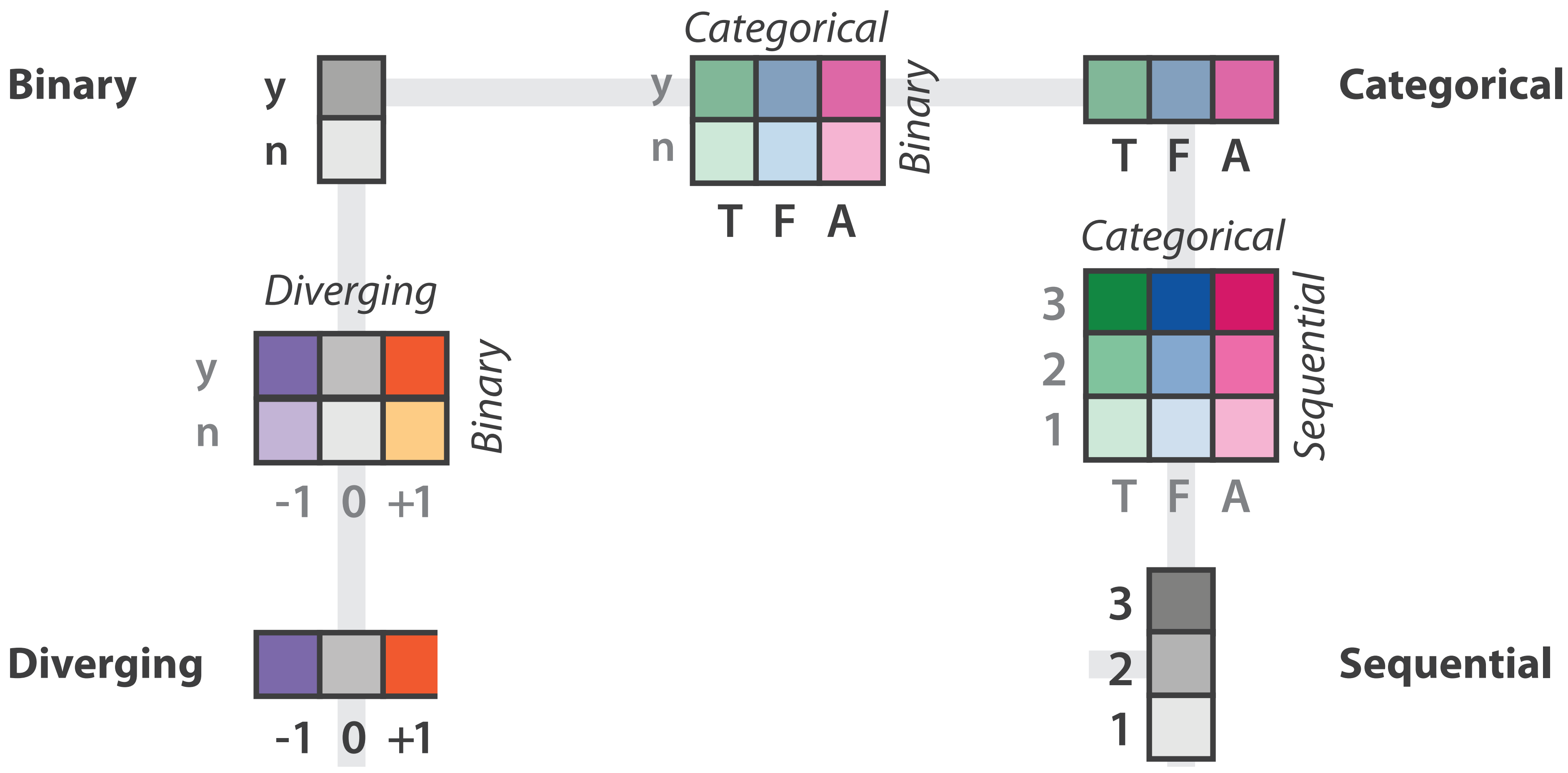
- ➔ Categorical

- ➔ Ordered
  - ➔ *Sequential*

  - ➔ *Diverging*

- ➔ Bivariate



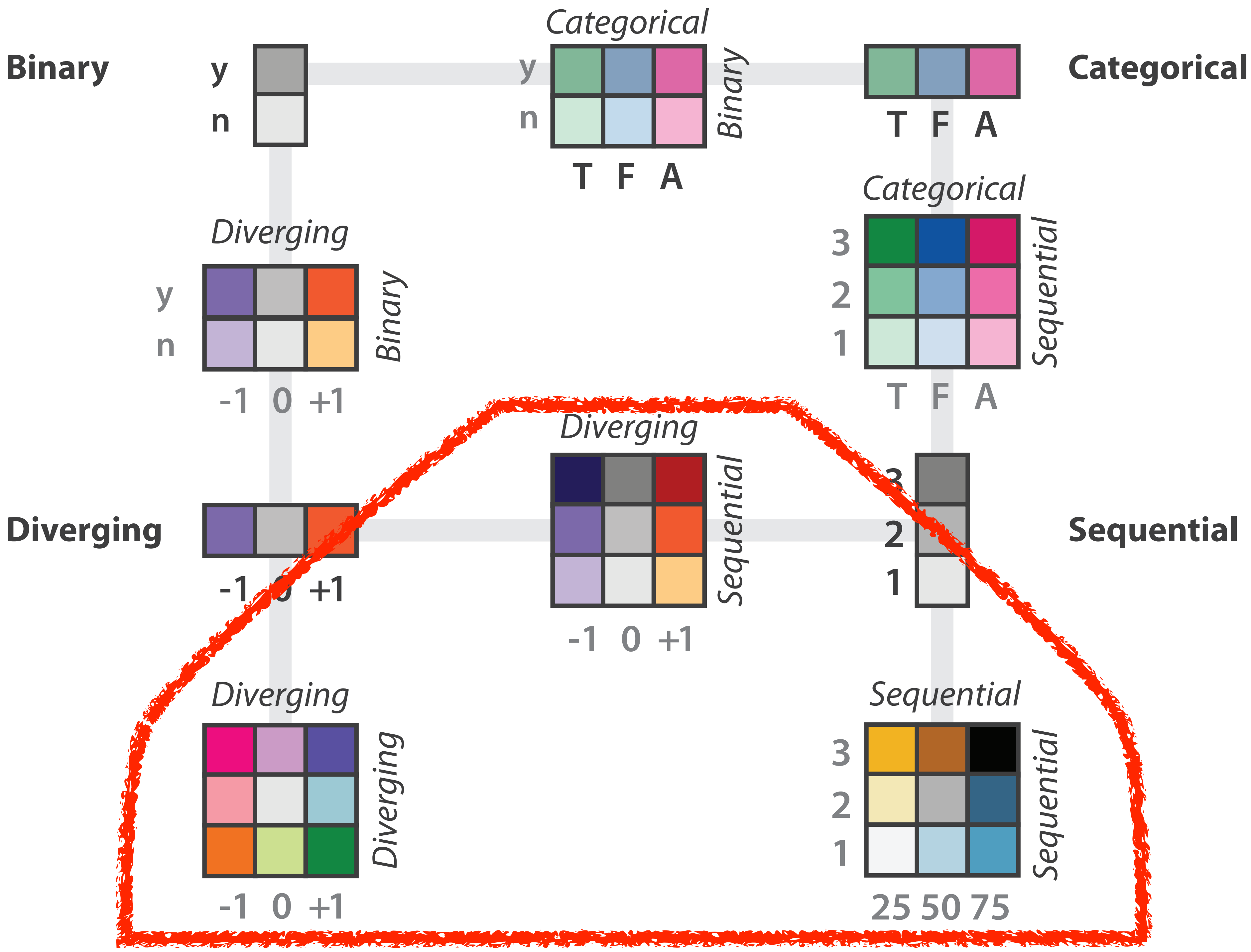


after [Color Use Guidelines for Mapping and Visualization. Brewer, 1994. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html]

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

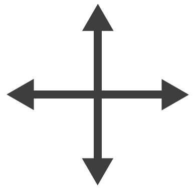
→ Categorical

→ Ordered

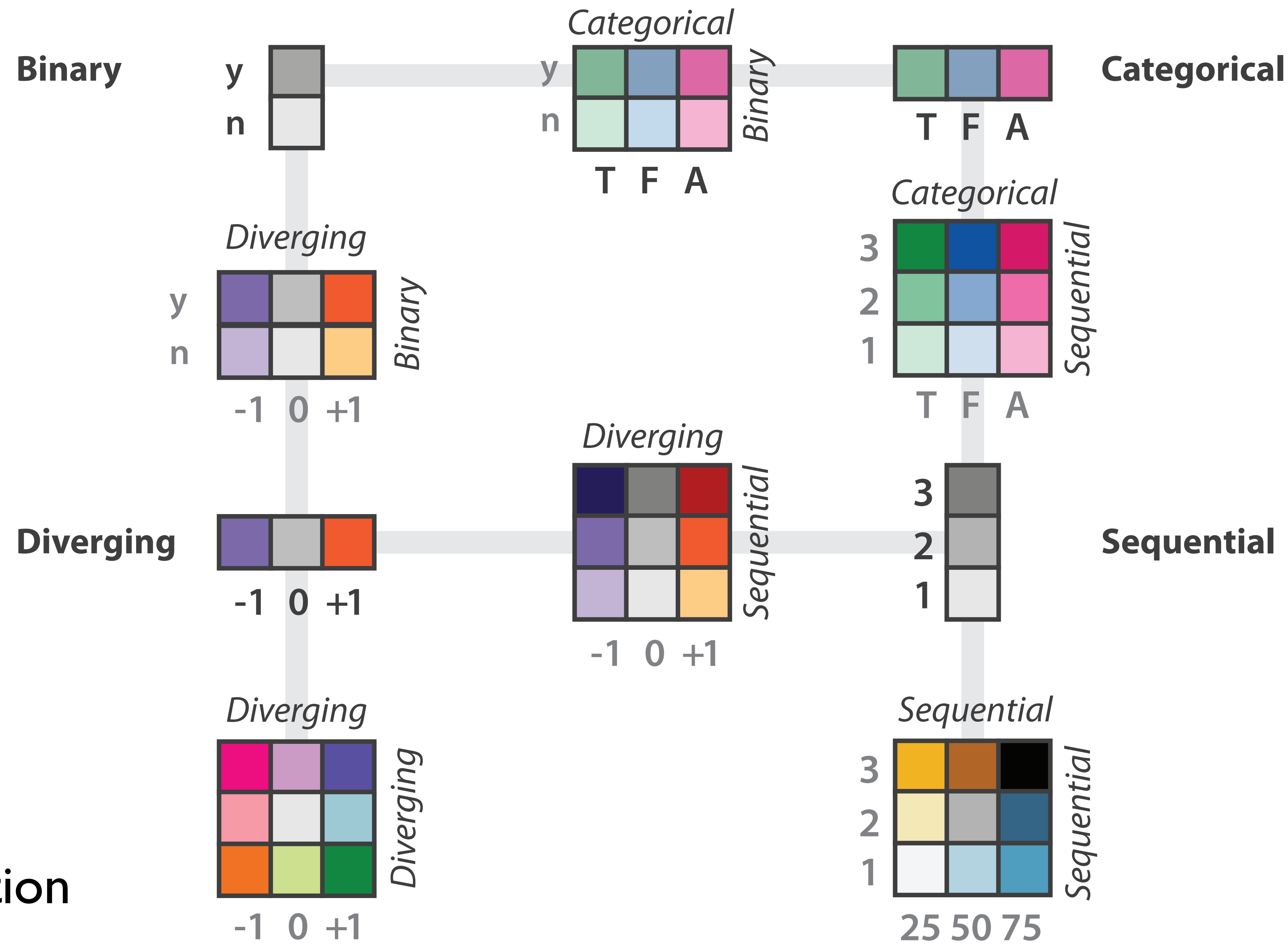
→ *Sequential*

→ *Diverging*

→ Bivariate



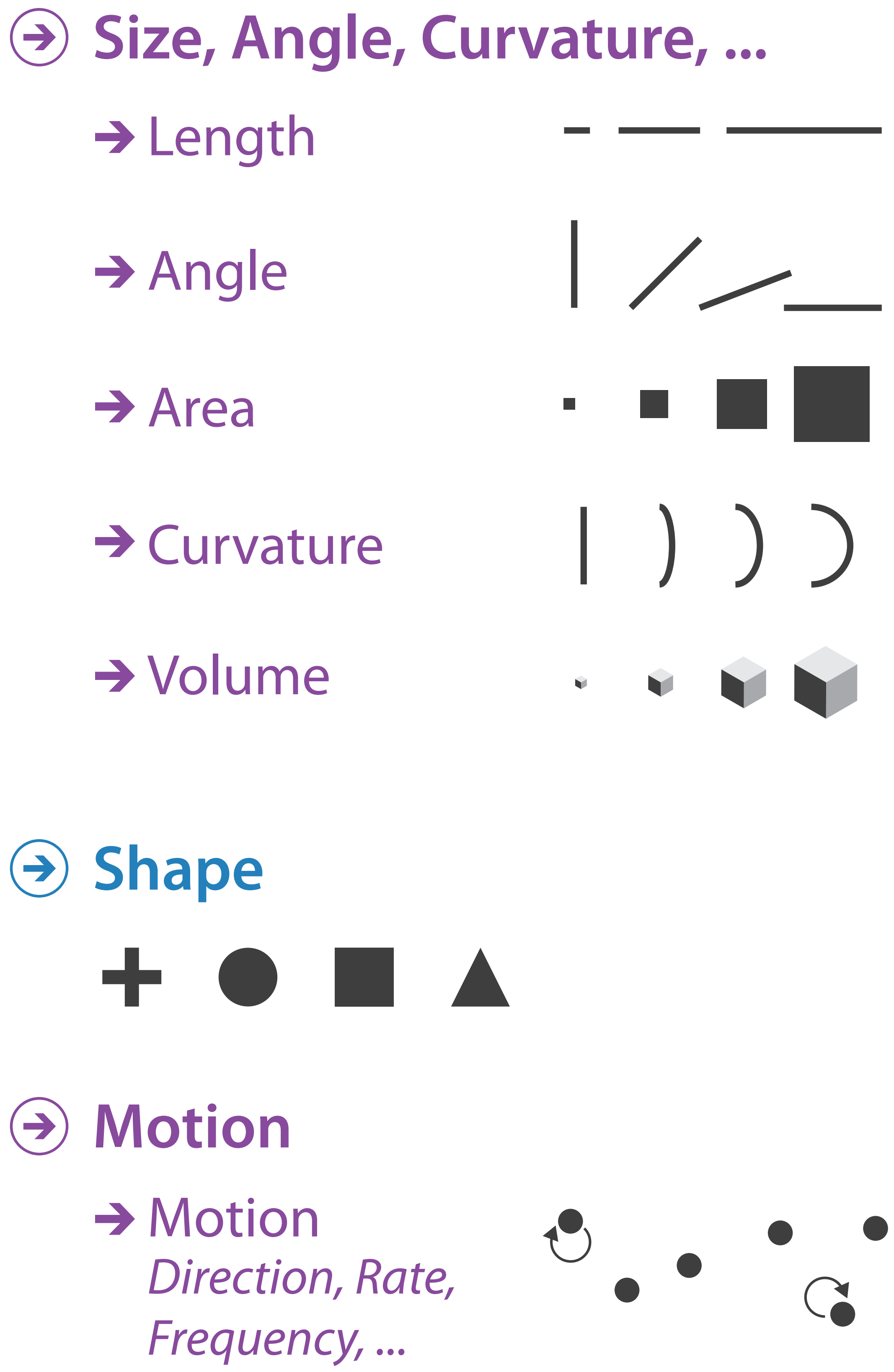
- color channel interactions
  - size heavily affects salience
    - small regions need high saturation
    - large need low saturation
  - saturation & luminance: 3-4 bins max
    - also not separable from transparency



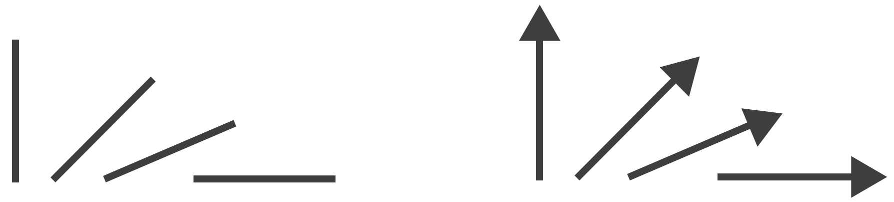
after [Color Use Guidelines for Mapping and Visualization. Brewer, 1994. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html]

Map other channels

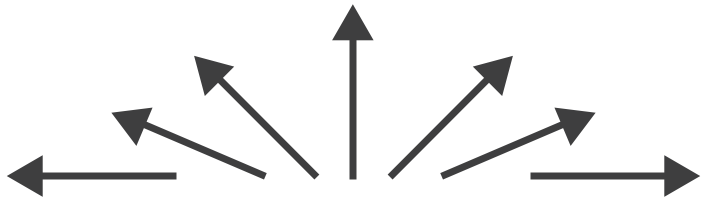
- size
  - –length accurate, 2D area ok, 3D volume poor
- angle
  - –nonlinear accuracy
    - horizontal, vertical, exact diagonal
- shape
  - –complex combination of lower-level primitives
  - –many bins
- motion
  - –highly separable against static
    - binary: great for highlighting
  - –use with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



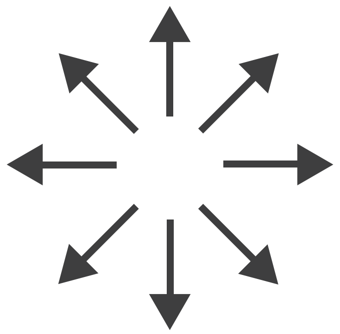
Angle



Sequential ordered
line mark or arrow glyph



Diverging ordered
arrow glyph



Cyclic ordered
arrow glyph

Paper: D3 System

Paper: D3

- paper types
  - design studies
  - technique/algorithm
  - evaluation
  - model/taxonomy
  - **system**

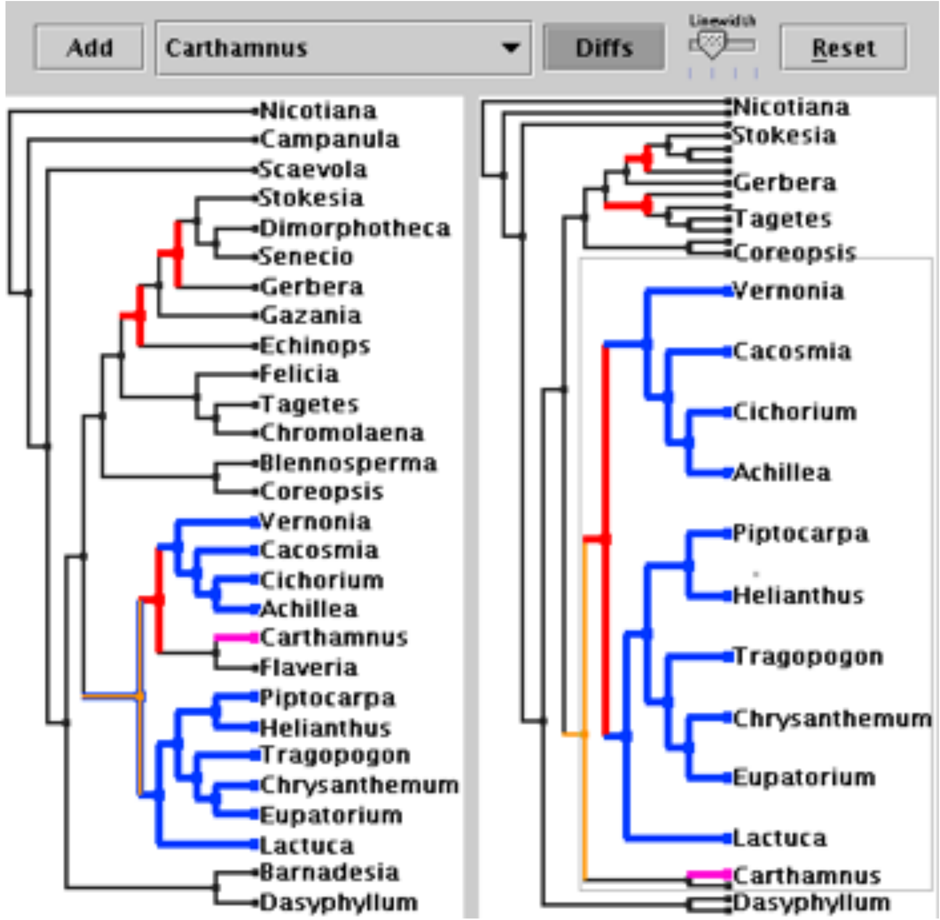
[D3: Data-Driven Documents. Bostock, Ogievetsky, Heer. IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis), 2011.]

Toolkits

- imperative: how
  - low-level rendering: Processing, OpenGL
  - parametrized visual objects: prefuse
    - also flare: prefuse for Flash
- declarative: what
  - Protoviz, D3, ggplot2
  - separation of specification from execution
- considerations
  - expressiveness
    - can I build it?
  - efficiency
    - how long will it take?
  - accessibility
    - do I know how?

WebGL/OpenGL

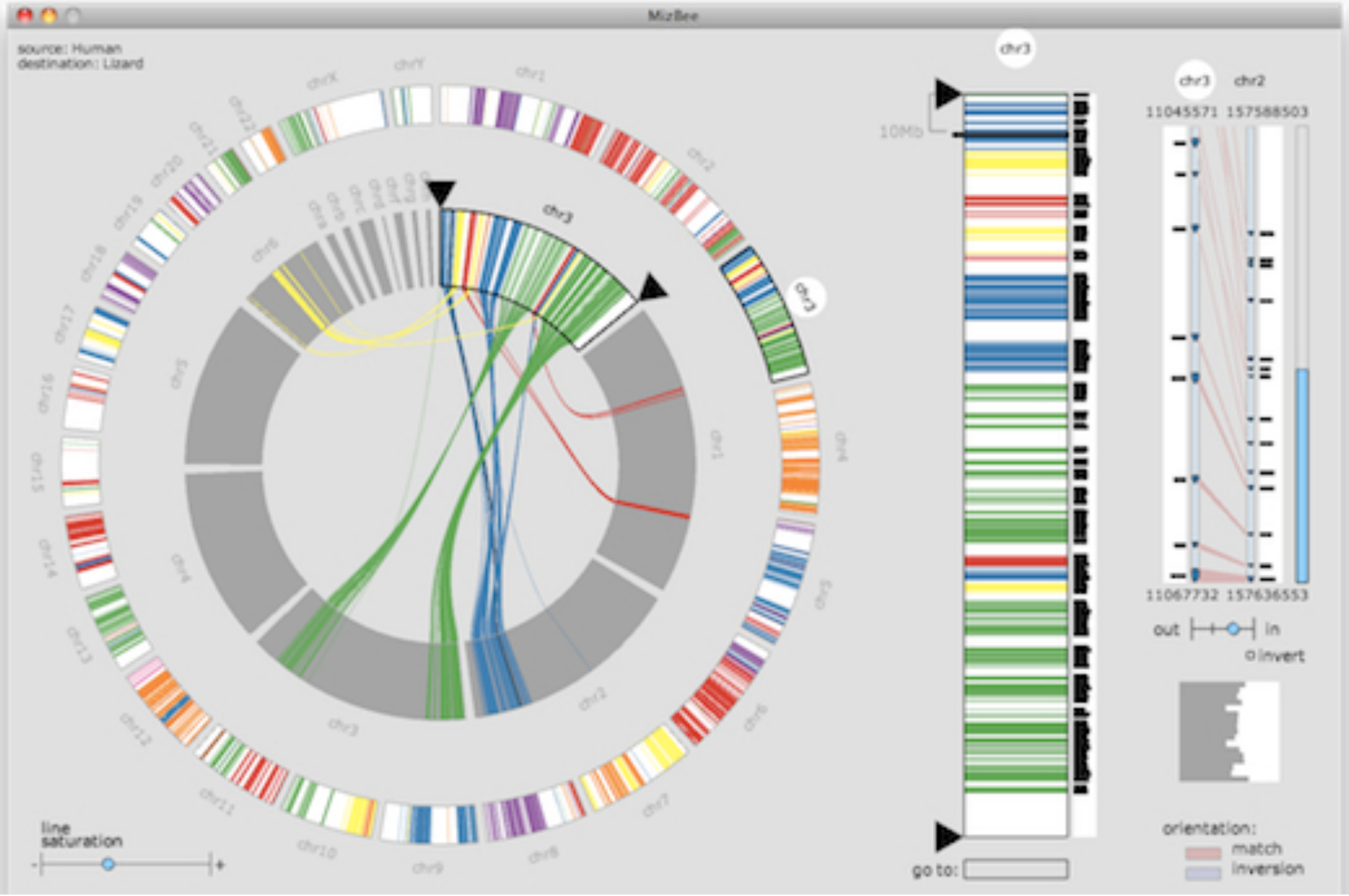
- graphics library
  - pros
    - power and flexibility, complete control for graphics
    - hardware acceleration
    - many language bindings: js, C, C++, Java (w/ JOGL)
  - cons
    - big learning curve if you don't know already
    - no vis support, must roll your own everything
  - example app: TreeJuxtaposer (OpenGL)



[Fig 5. Munzner et al. TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. Proc SIGGRAPH 2003, pp 453-462.]

Processing / p5.js

- layer on top of Java/OpenGL, Javascript/WebGL
- visualization esp. for artists/designers
- pros
  - great sandbox for rapid prototyping
  - huge user community, great documentation
- cons
  - poor widget library support
- example app: MizBee



[Fig 1. Meyer et al. MizBee: A Multiscale Synteny Browser. Proc. InfoVis 2009.]

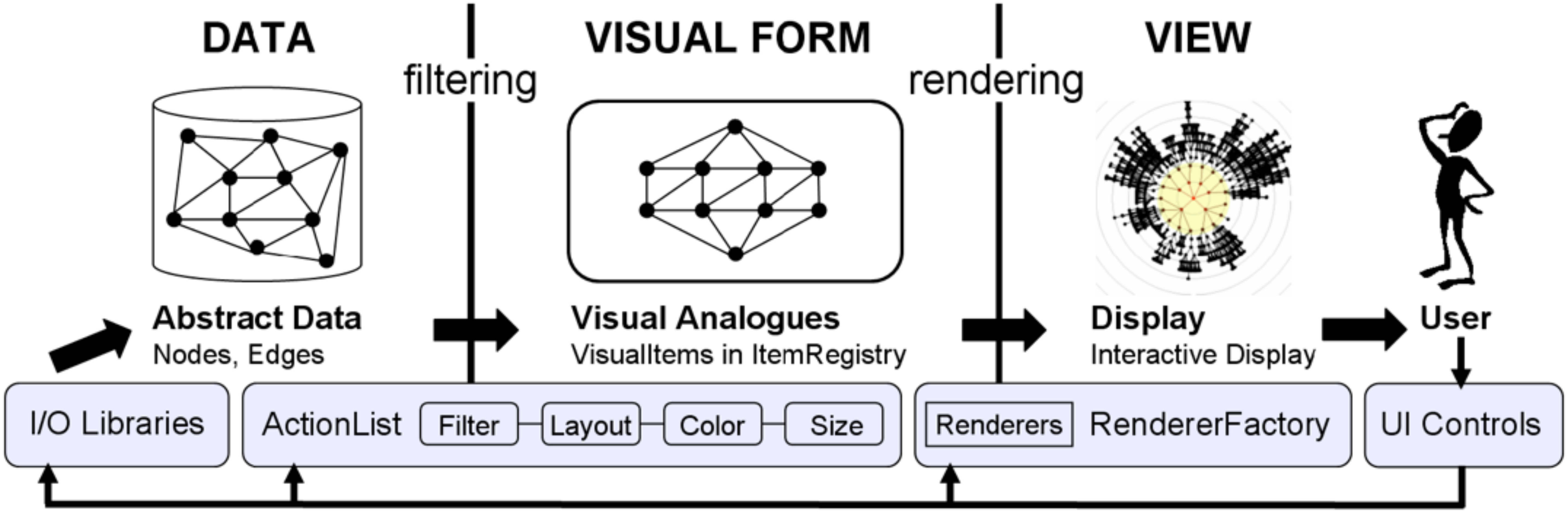
prefuse

- infovis toolkit, in Java
- fine-grained building blocks for tailored visualizations
- pros
  - heavily used (previously)
  - very powerful abstractions
  - quickly implement most techniques covered so far
- cons
  - no longer active
  - nontrivial learning curve
- example app: DOITrees Revisited

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

prefuse

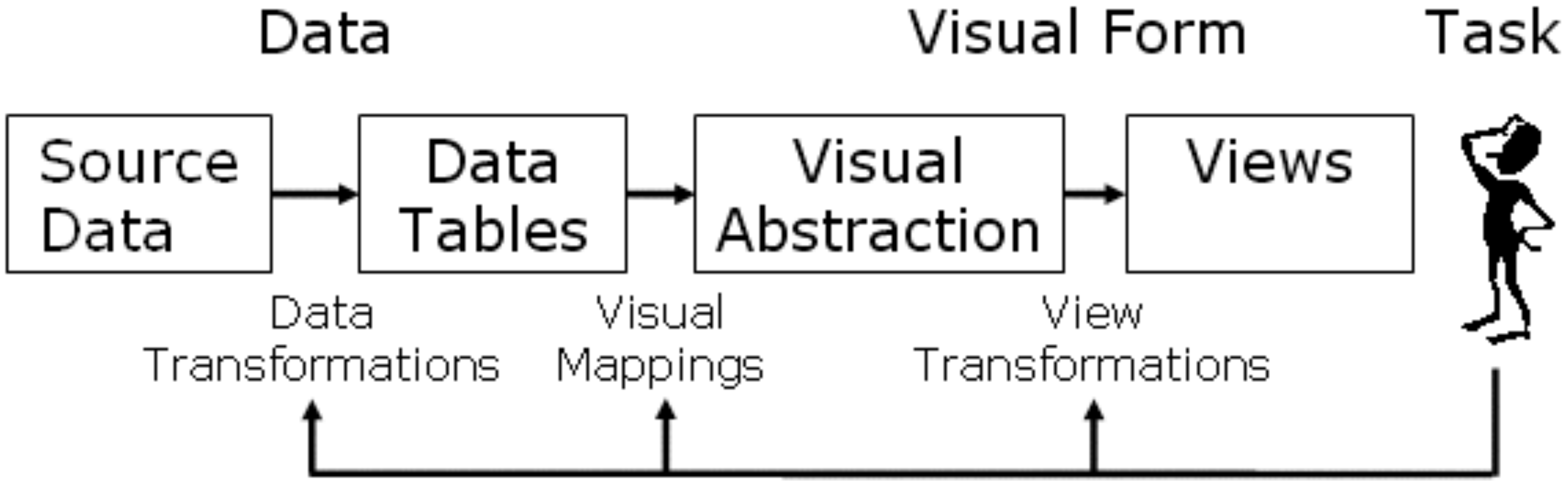
- separation: abstract data, visual form, view
  - data: tables, networks
  - visual form: layout, color, size, ...
  - view: multiple renderers



[Fig 2. Heer, Card, and Landay. Prefuse: A Toolkit for Interactive Information Visualization. Proc. CHI 2005, 421-430]

InfoVis Reference Model

- conceptual model underneath design of prefuse and many other toolkits
- heavily influenced much of infovis (including nested model)
  - aka infovis pipeline, data state model



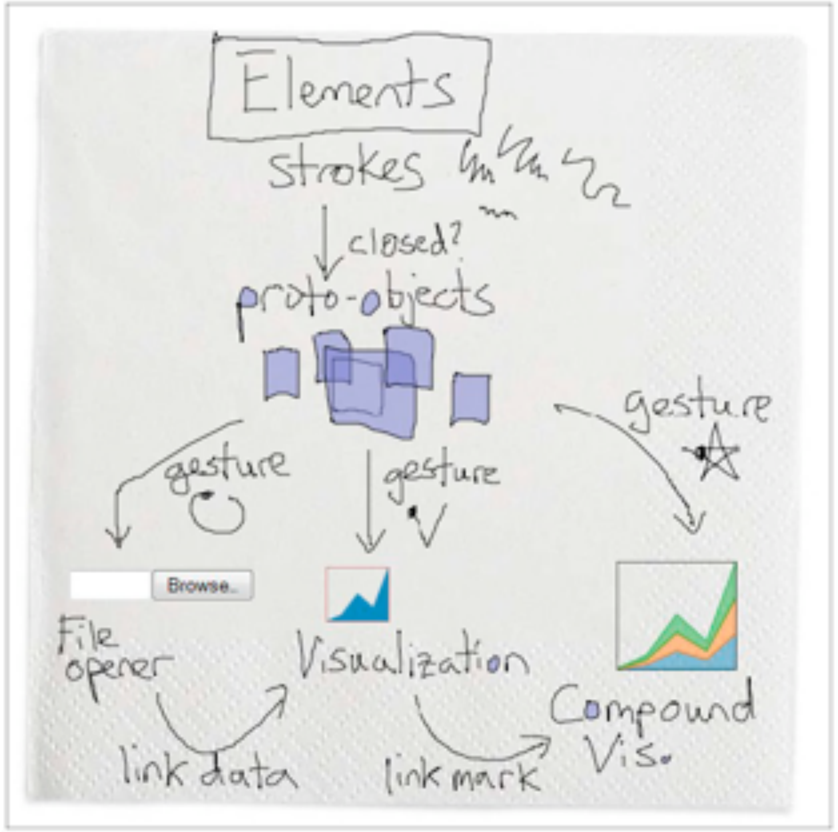
[Redrawn Fig 1.23. Card, Mackinlay, and Shneiderman. Readings in Information Visualization: Using Vision To Think, Chapter 1. Morgan Kaufmann, 1999.]

Declarative toolkits

- imperative tools/libraries
  - say exactly **how** to do it
  - familiar programming model
    - OpenGL, prefuse, ...
- declarative: other possibility
  - just say **what** to do
  - Protovis, D3

Protovis

- declarative infovis toolkit, in Javascript
  - –also later Java version
- marks with inherited properties
- pros
  - –runs in browser
  - –matches mark/channel mental model
  - –also much more: interaction, geospatial, trees,...
- cons
  - –not all kinds of operations supported
- example app: NapkinVis (2009 course project)



[Fig 1, 3. Chao. NapkinVis. http://www.cs.ubc.ca/~tmm/courses/533-09/projects.html#will]

Protovis Validation

- wide set of old/new app examples
  - expressiveness, effectiveness, scalability
  - accessibility
- analysis with cognitive dimensions of notation
  - closeness of mapping, hidden dependencies
  - role-expressiveness visibility, consistency
  - viscosity, diffuseness, abstraction
  - hard mental operations

[Cognitive dimensions of notations. Green (1989). In A. Sutcliffe and L. Macaulay (Eds.) People and Computers V. Cambridge, UK: Cambridge University Press, pp 443-460.]

D3

- declarative infovis toolkit, in Javascript
- Protovis meets Document Object Model
- pros
  - seamless interoperability with Web
  - explicit transforms of scene with dependency info
  - massive user community, many thirdparty apps/libraries on top of it, lots of docs
- cons
  - even more different from traditional programming model
- example apps: many

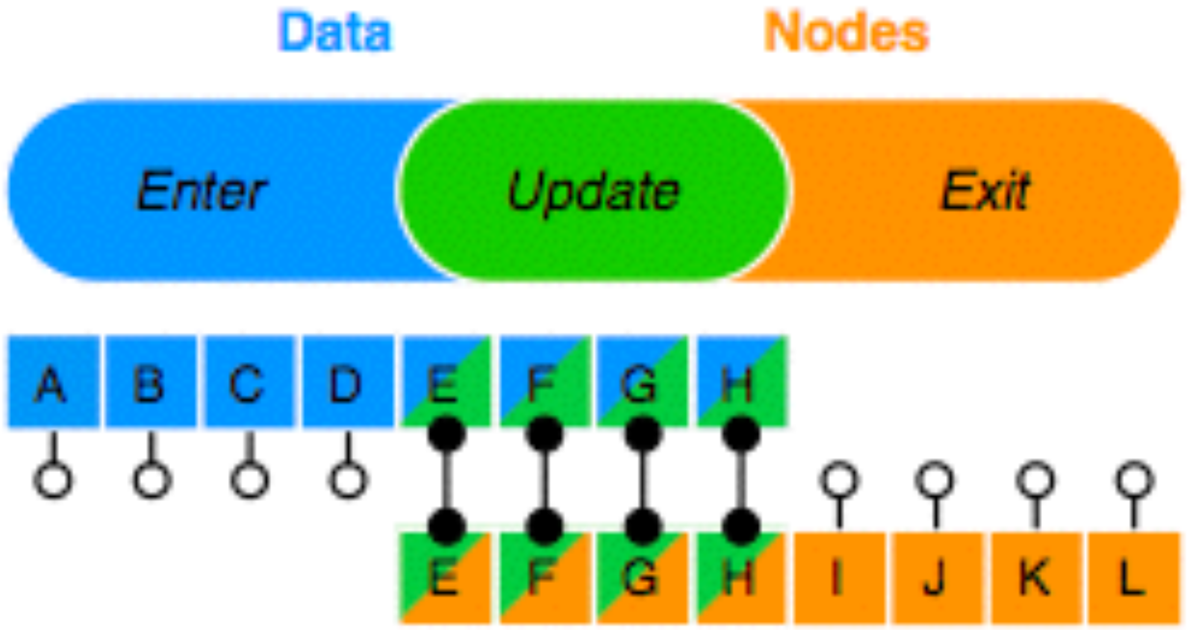
D3

- objectives
  - compatibility
  - debugging
  - performance
- related work typology
  - document transformers
  - graphics libraries
  - infovis systems
    - general note: all related work sections are a mini-taxonomy!

[D3: Data-Driven Documents. Bostock, Ogievetsky, Heer. IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis), 2011.]

D3 capabilities

- query-driven selection
  - selection: filtered set of elements queries from the current doc
    - also partitioning/grouping!
  - operators act on selections to modify content
    - instantaneous or via animated transitions with attribute/style interpolators
    - event handlers for interaction
- data binding to scenegraph elements
  - data joins bind input data to elements
  - enter, update, exit subselections
  - sticky: available for subsequent re-selection
  - sort, filter



[D3: Data-Driven Documents. Bostock, Ogievetsky, Heer. IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis), 2011.]

D3 Features

- document transformation as atomic operation
  - scene changes vs representation of scenes themselves
- immediate property evaluation semantics
  - avoid confusing consequences of delayed evaluation

- validation
  - performance benchmarks
    - page loads, frame rate
  - accessibility
    - everybody has voted with their feet by now!

Next Time

- to read
  - VAD Ch. 8: Arrange Spatial Data
  - VAD Ch. 9: Arrange Networks
  - paper: ABySS-Explorer: visualizing genome sequence assemblies.. Cydney B. Nielsen, Shaun D. Jackman, Inanc Birol, Steven J.M. Jones. TVCG 15(6):881-8, 2009 (Proc. InfoVis 2009).
    - [paper type: design study]
  - paper: Interactive Visualization of Genealogical Graphs. Michael J. McGuffin, Ravin Balakrishnan. Proc. InfoVis 2005, pp 17-24.
    - [paper type: technique]
- to prepare
  - project pitches (3 min each)