Week 4:
Manipulate, Facet, Reduce

Tamara Munzner
Department of Computer Science
University of British Columbia

JRNL 520H, Special Topics in Contemporary Journalism: Data Visualization
Week 4: 4 October 2016

http://www.cs.ubc.ca/~tmm/courses/journ16

Whereabouts

- Caitlin on travel this week and next week
  - don't expect email answers until she returns; email Tamara instead!
- Tamara on travel Thu Oct 6 - Mon Oct 10
  - in Portland Fri/Sat to give another keynote, will still be answering email
  - short office hours in Sing Tao next week: 12:30-1:30pm

News

- Assign 2 marks not out yet
  - stay tuned, just got back from Stanford late last night
- Today's format
  - interleave foundations & demos
    - Tamara will walk through Tableau demos
    - you follow along step by step on your own laptop
    - Tamara will take breaks to rove the room to help out folks who get stuck

Last Time

Demo 1: Stone Color Workbook

- Credit: Maureen Stone, Tableau Research
  - designer of Tableau color defaults, author of *A Field Guide to Digital Color*
  - workbook from Tableau Customer Conference 2014 talk
    Seriously Colorful: Advanced Color Principles & Practices

- Tableau Lessons
  - more visual encoding practice
  - color palettes, univariate & bivariate
  - discrete (categorical) vs continuous (quantitative)

- Big Ideas
  - Tableau has many built-in features to get color right, but care still needed

Demo 2: Intro to Maps

- Tableau Lessons
  - handling spatial data
  - multiple data sources
  - paths on maps
  - more on handling missing data: filtering

- Big Ideas
  - integrating visual encoding design choices with given spatial data

How?

Encode

Arrange

- Express
- Separate
- Order
- Align
- Use

Map

from categorical and ordered attributes

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

Direction, Rate, Frequency, ...

Manipulate

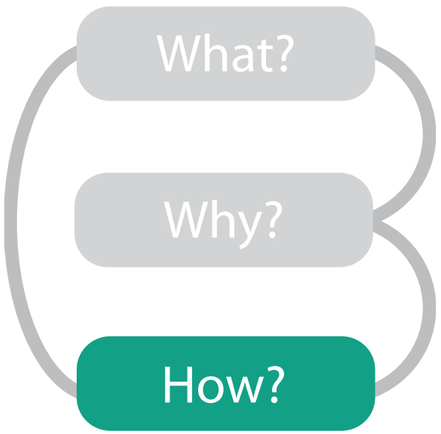
- Change
- Select
- Navigate

Facet

- Juxtapose
- Partition
- Superimpose

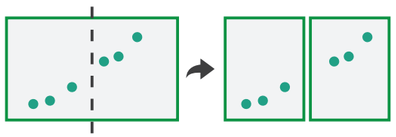
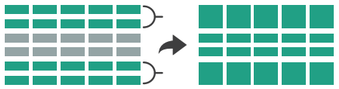
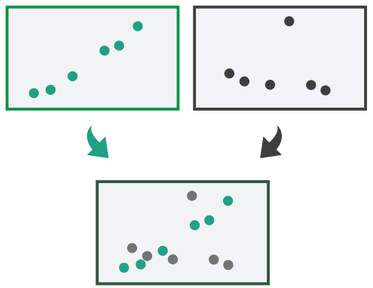
Reduce

- Filter
- Aggregate
- Embed



How to handle complexity: 1 previous strategy + 3 more

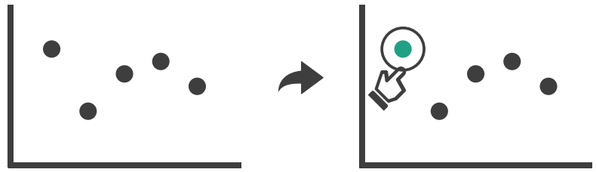
- derive new data to show within view
- change view over time
- facet across multiple views
- reduce items/attributes within single view

	Manipulate	Facet	Reduce
Derive 	Change 	Juxtapose 	Filter 
	Select 	Partition 	Aggregate 
	Navigate 	Superimpose 	Embed 

Manipulate

Change over Time

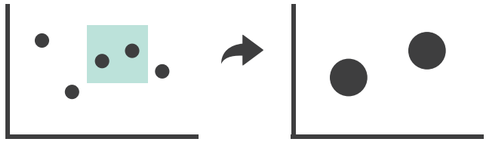
Select



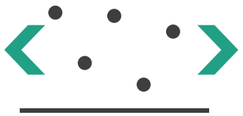
Navigate

Item Reduction

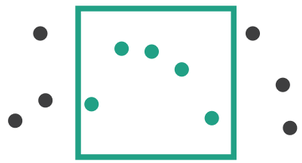
- *Zoom*
  Geometric or *Semantic*



- *Pan/Translate*

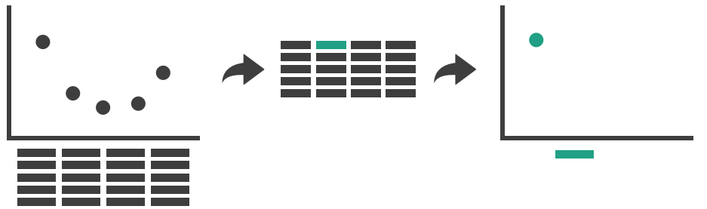


- *Constrained*

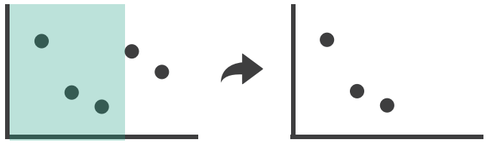


Attribute Reduction

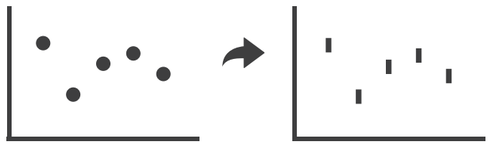
- *Slice*



- *Cut*



- *Project*

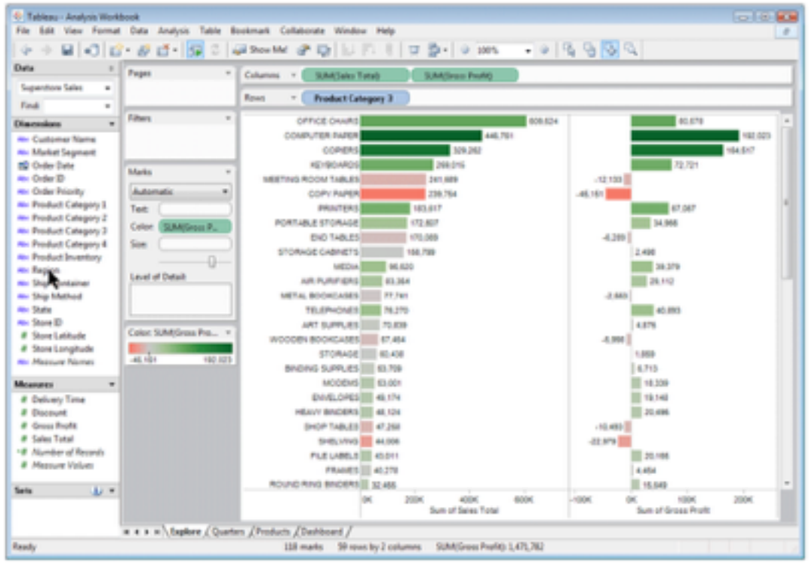


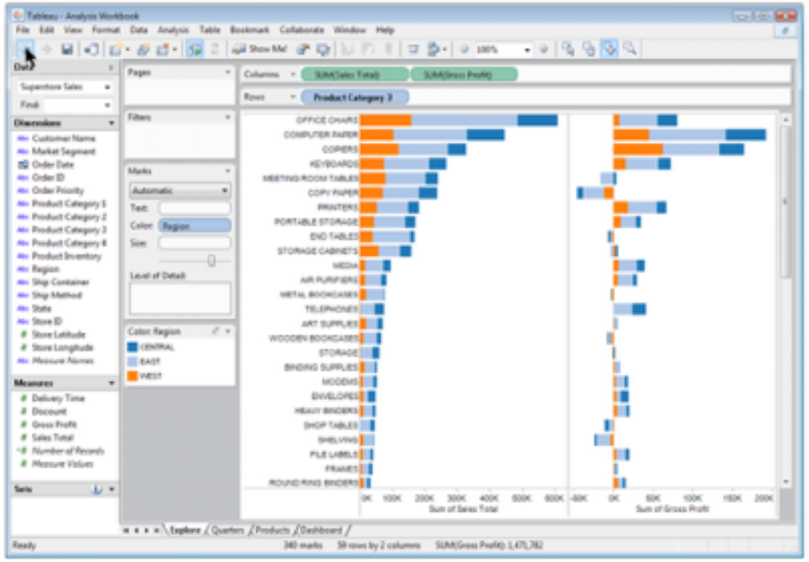
Change over time

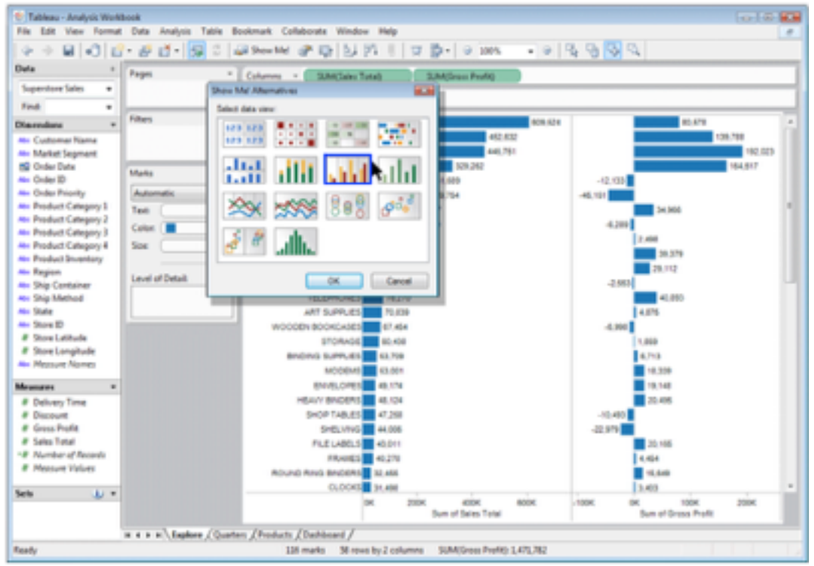
- change any of the other choices
  - encoding itself
  - parameters
  - arrange: rearrange, reorder
  - aggregation level, what is filtered...

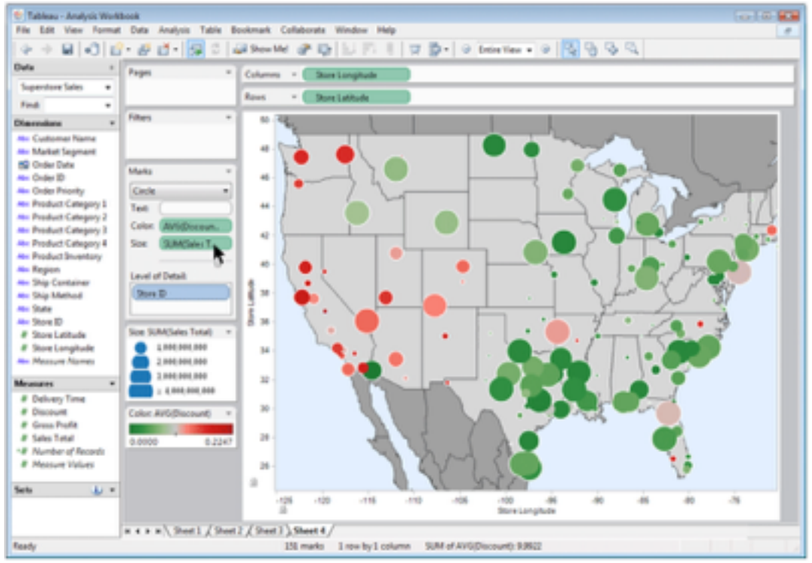
  - interaction entails change

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







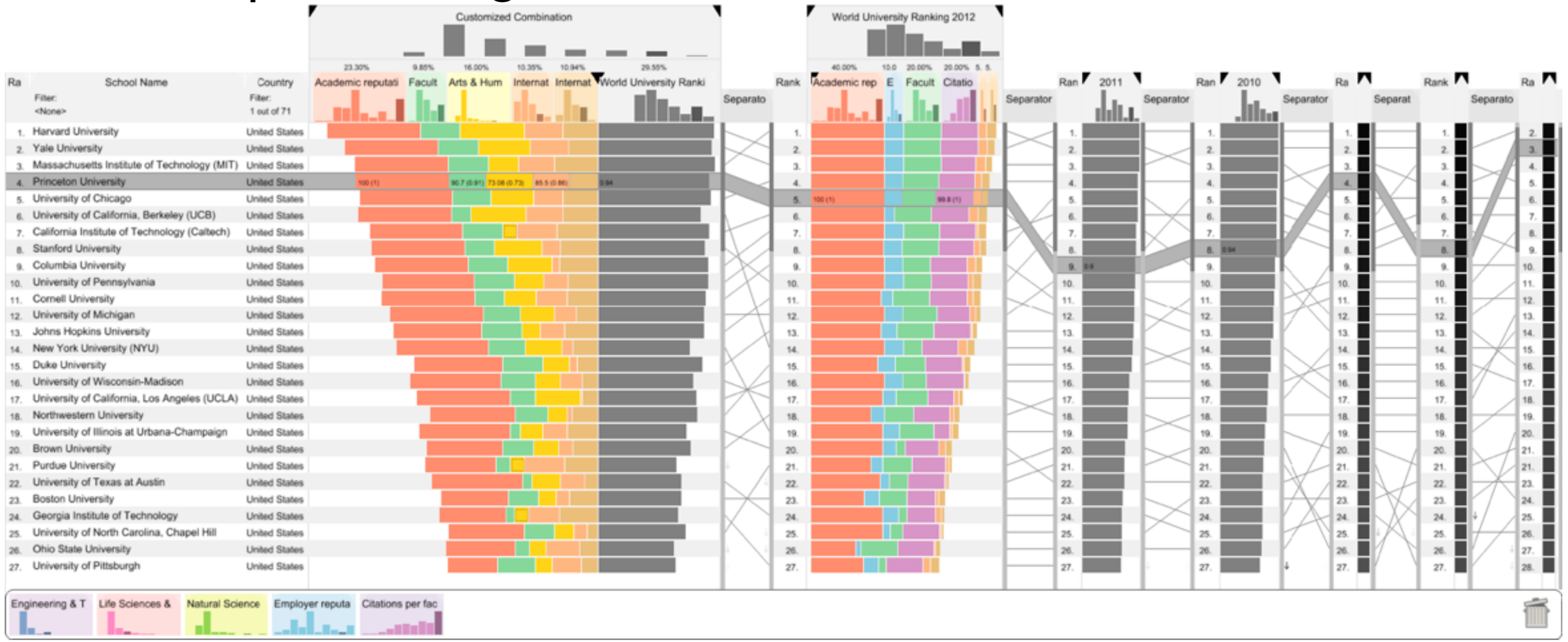


made using Tableau, http://tableausoftware.com

Idiom: **Reorder**

System: **LineUp**

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

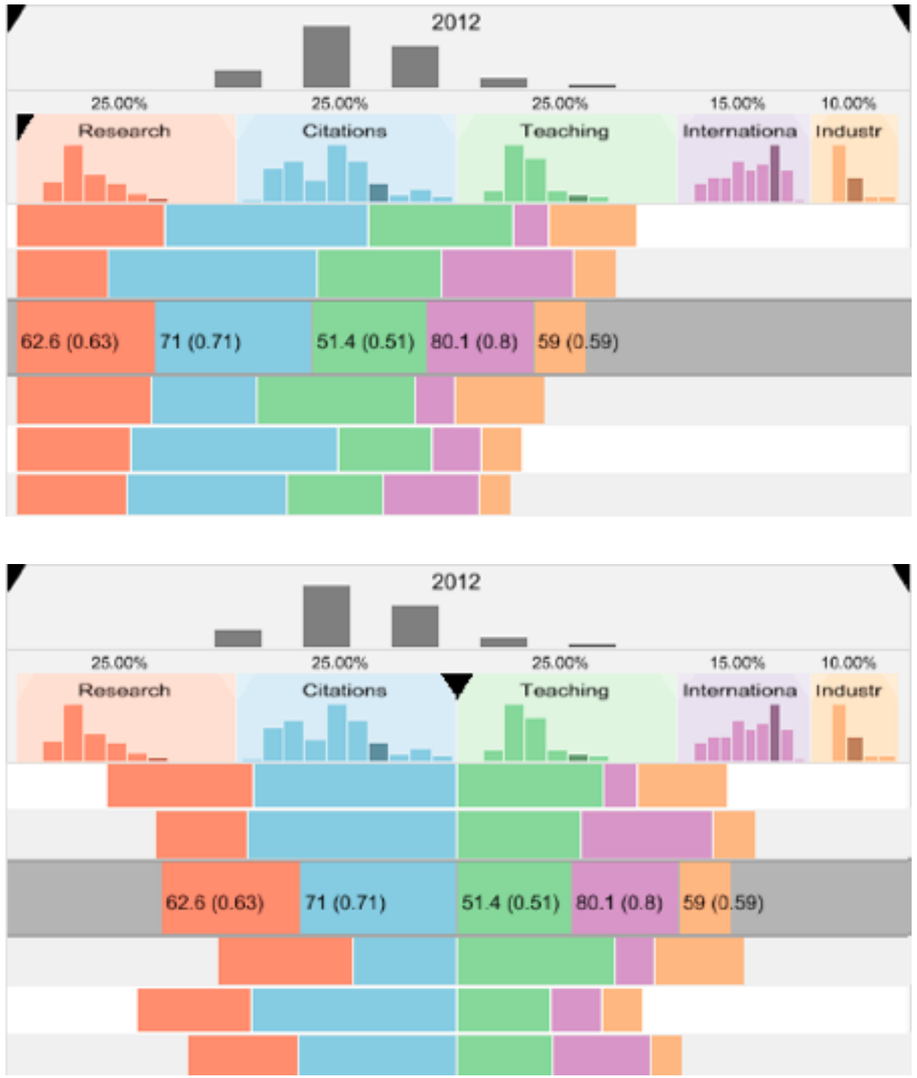


[LineUp:Visual Analysis of Multi-Attribute Rankings. Gratzl, Lex, Gehlenborg, Pfister, and Streit. IEEE Trans.Visualization and Computer Graphics (Proc. InfoVis 2013) 19:12 (2013), 2277–2286.]

Idiom: **Realign**

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

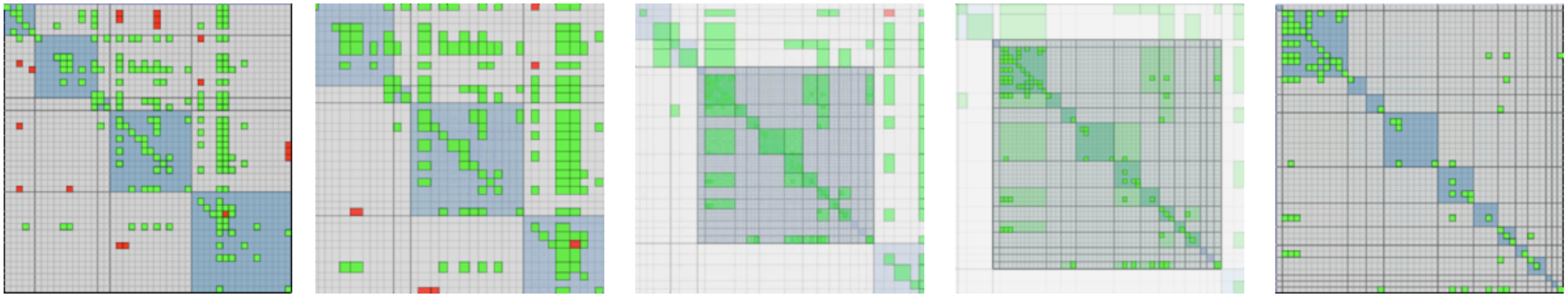
System: **LineUp**



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

Idiom: **Animated transitions**

- smooth 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.]

Select and highlight

- selection: basic operation for most interaction
- design choices
  - how many selection types?
    - click vs hover: heavyweight, lightweight
    - primary vs secondary: semantics (eg source/target)
- highlight: change visual encoding for selection targets
  - color
    - limitation: existing color coding hidden
  - other channels (eg motion)
  - add explicit connection marks between items

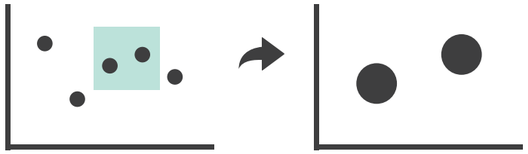
Navigate: Changing item visibility

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

Navigate

→ Item Reduction

→ *Zoom*
Geometric or *Semantic*



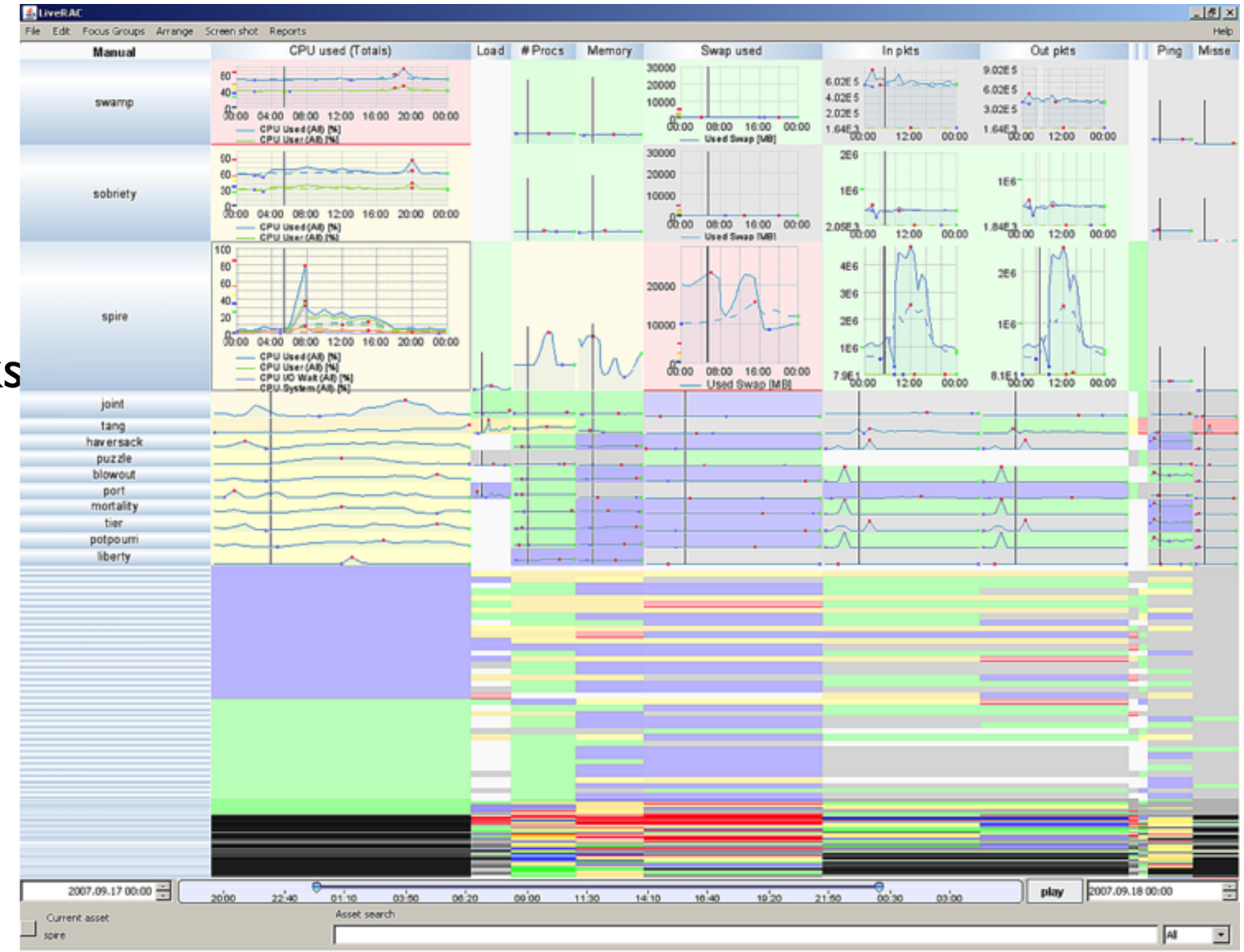
→ *Pan/Translate*

→ *Constrained*

Idiom: **Semantic zooming**

System: **LiveRAC**

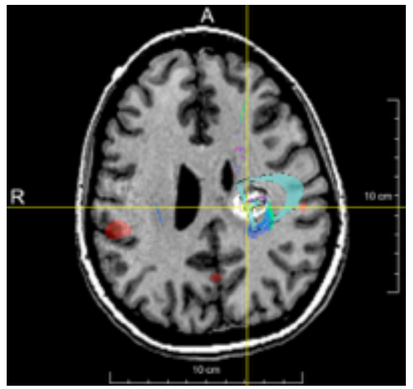
- visual encoding change
  - colored box
  - sparkline
  - simple line chart
  - full chart: axes and tickmarks

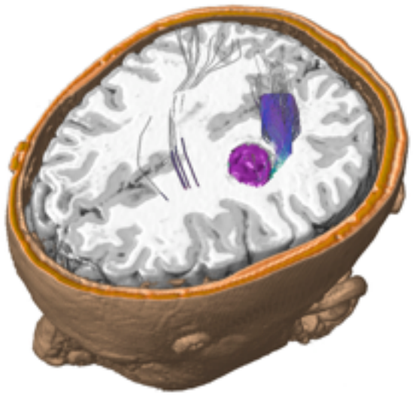


[LiveRAC - Interactive Visual Exploration of System Management Time-Series Data. McLachlan, Munzner, Koutsofios, and North. Proc. ACM Conf. Human Factors in Computing Systems (CHI), pp. 1483–1492, 2008.]

Navigate: Reducing attributes

- continuation of camera metaphor
  - slice
    - show only items matching specific value for given attribute: slicing plane
    - axis aligned, or arbitrary alignment
  - cut
    - show only items on far slide of plane from camera
  - project
    - change mathematics of image creation
      - orthographic
      - perspective
      - many others: Mercator, cabinet, ...

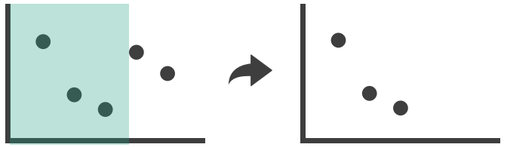




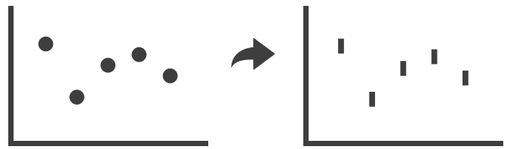
➔ Attribute Reduction

➔ *Slice*

➔ *Cut*



➔ *Project*



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

Previous Demos

- Tableau Lessons
  - changing visual encoding
  - changing ordering (sorting)
  - navigation
    - zoom/translate in maps

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

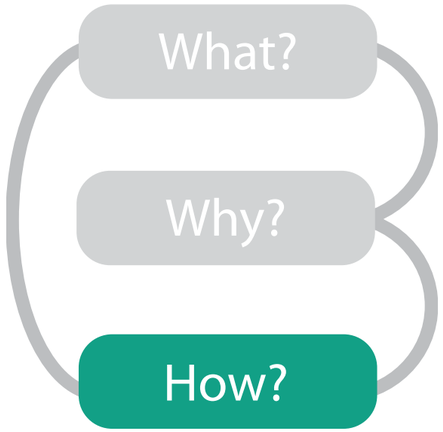
- Change
- Select
- Navigate

Facet

- Juxtapose
- Partition
- Superimpose

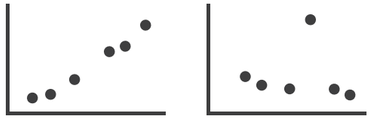
Reduce

- Filter
- Aggregate
- Embed

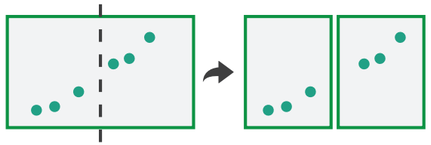


Facet

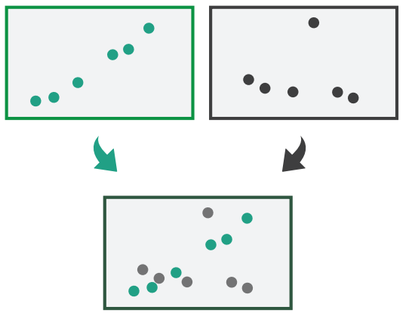
Juxtapose



Partition



Superimpose



Juxtapose and coordinate views

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

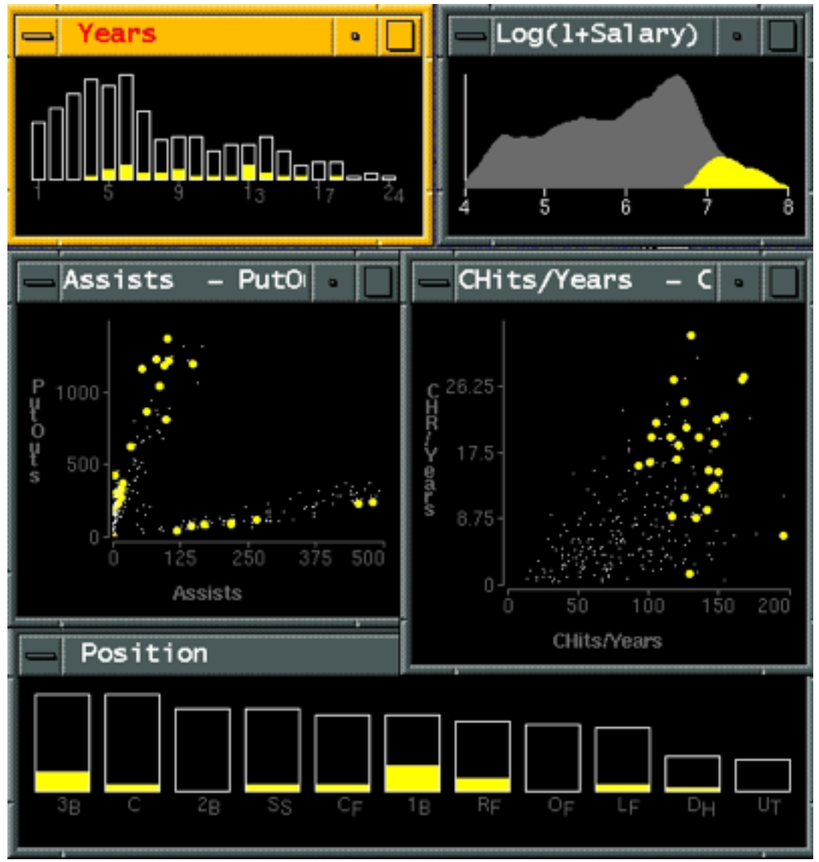
- Share Data: All/Subset/None

- Share Navigation

Idiom: **Linked highlighting**

System: **EDV**

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



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

Demo 1: Seattle Construction

- Credit: Ben Jones

- Tableau Lessons
  - linking views with actions: highlight on hover
  - global filtering

- Big Ideas
  - linking views possible but somewhat clunky in Tableau

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

System: **Google Maps**

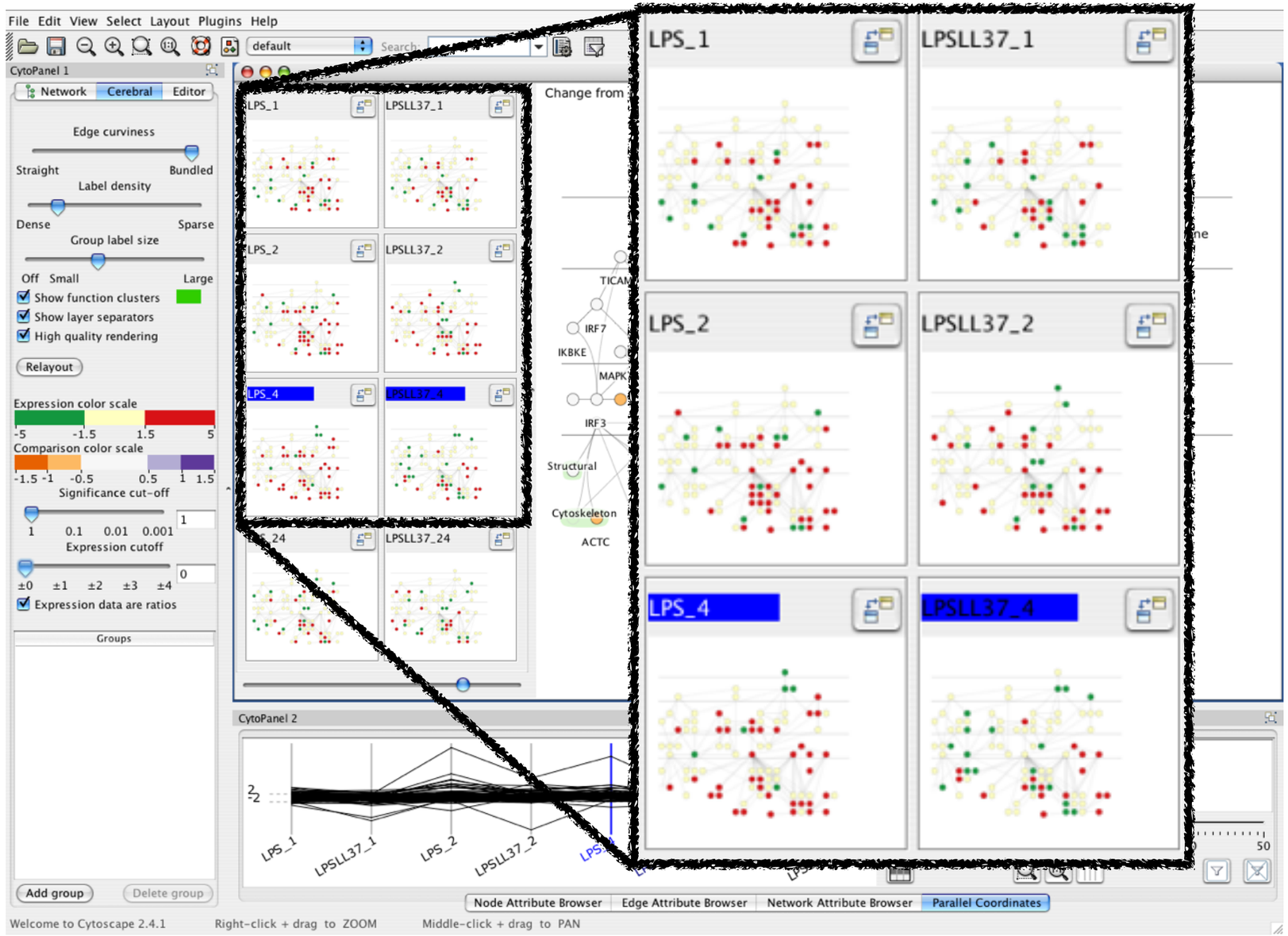
- encoding: same
- data: subset shared
- navigation: shared
  - bidirectional linking
- differences
  - viewpoint
  - (size)
- ***overview-detail***

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

Idiom: **Small multiples**

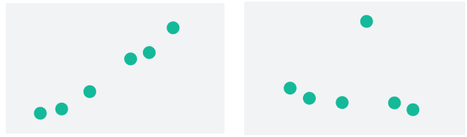
System: **Cerebral**

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



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

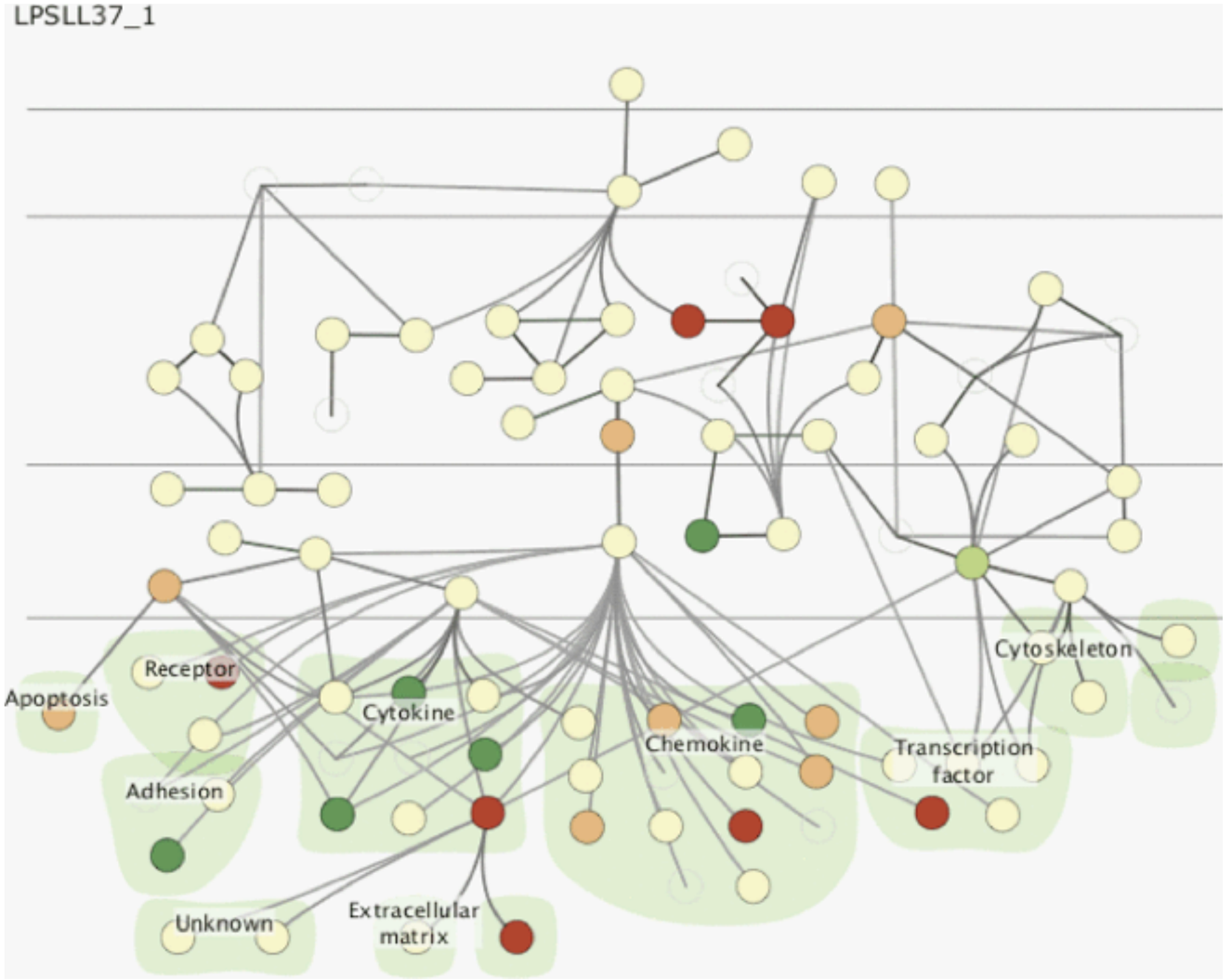
Coordinate views: Design choice interaction

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

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

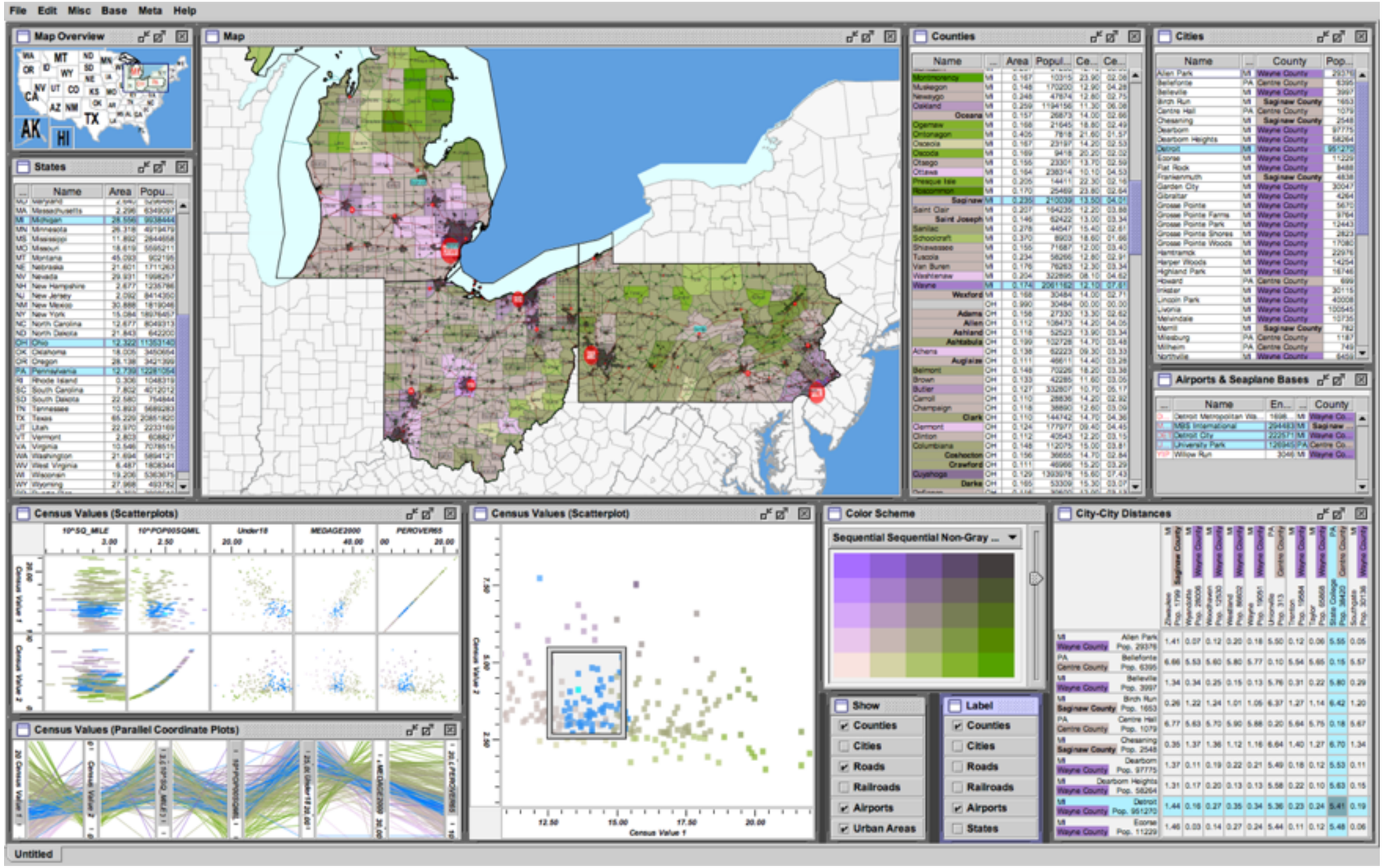
Why not animation?

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



System: **Improvise**

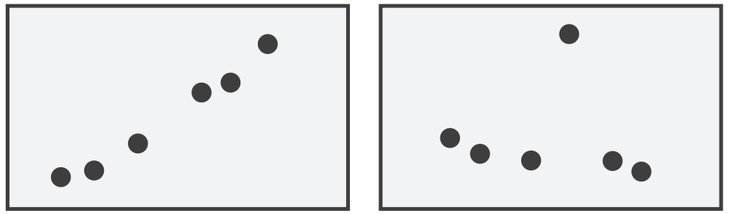
- investigate power of multiple views
  - pushing limits on view count, interaction complexity
  - how many is ok?
    - open research question
  - reorderable lists
    - easy lookup
    - useful when linked to other encodings

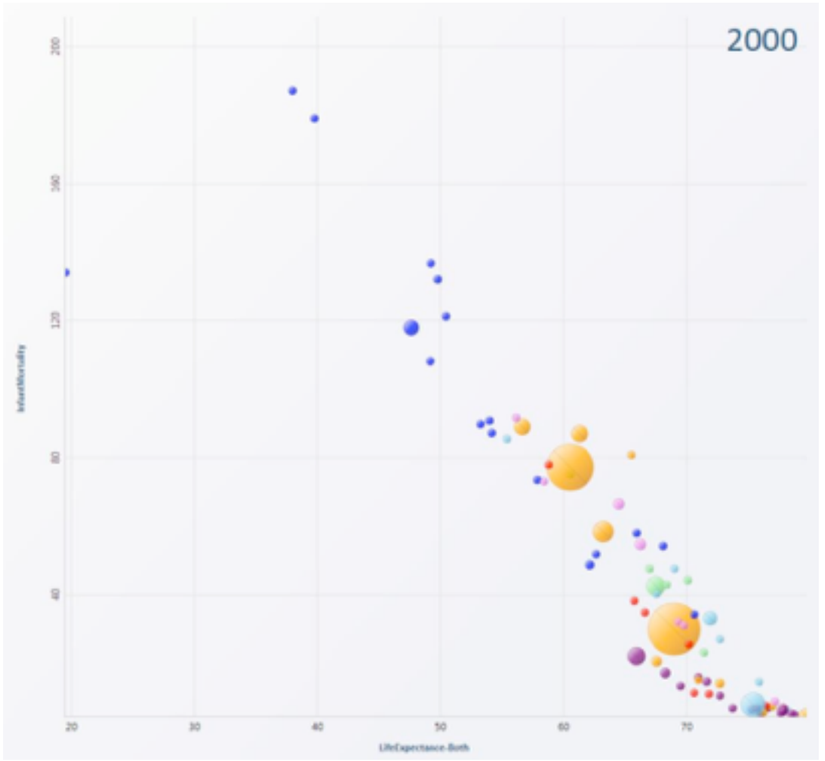


[Building Highly-Coordinated Visualizations In Improvise. Weaver. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 159–166, 2004.]

Partition into views

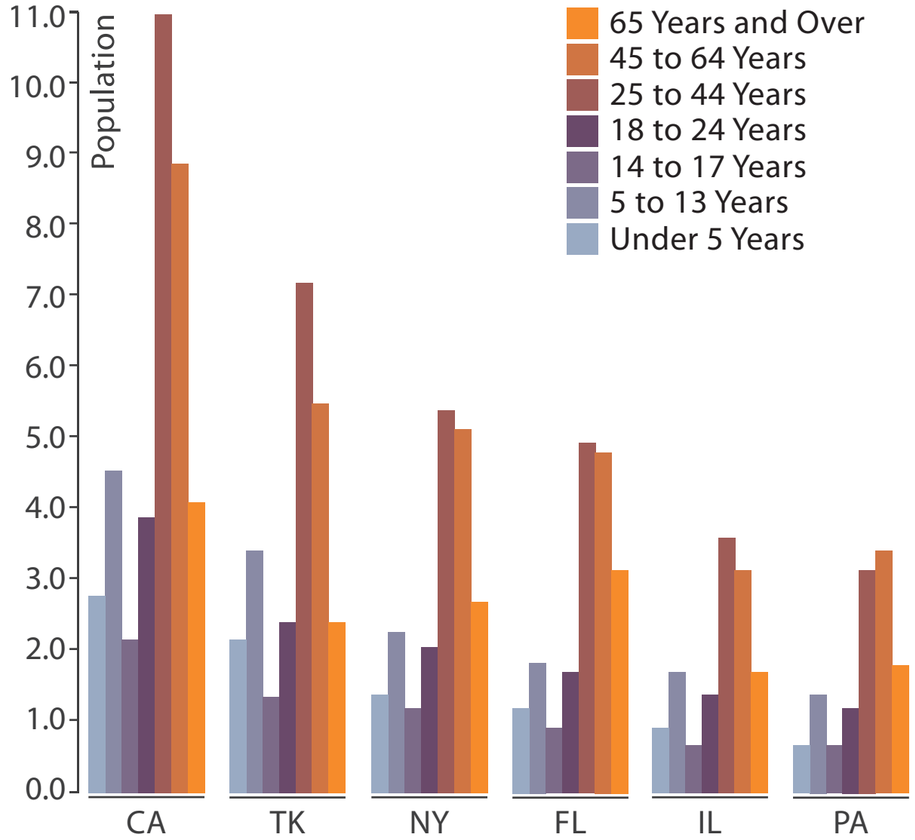
- how to divide data between views ➔ **Partition into Side-by-Side Views**
  - split into regions by attributes
  - encodes association between items using spatial proximity
  - order of splits has major implications for what patterns are visible
- no strict dividing line
  - ***view:*** *big/detailed*
    - contiguous region in which visually encoded data is shown on the display
  - ***glyph:*** *small/iconic*
    - object with internal structure that arises from multiple marks



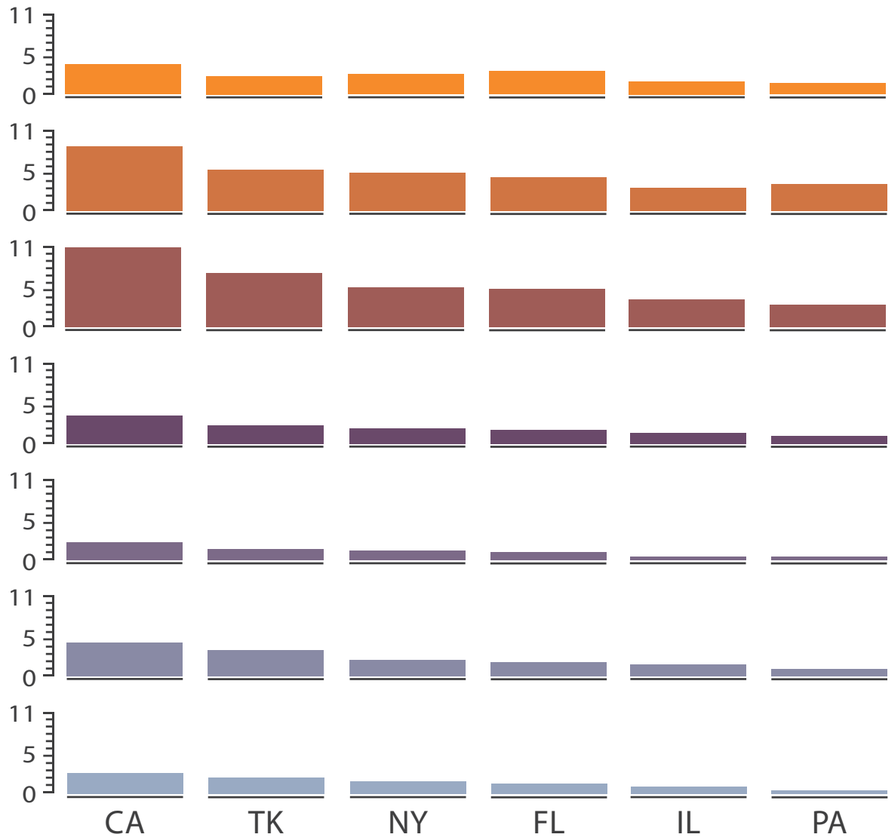


Partitioning: List alignment

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



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

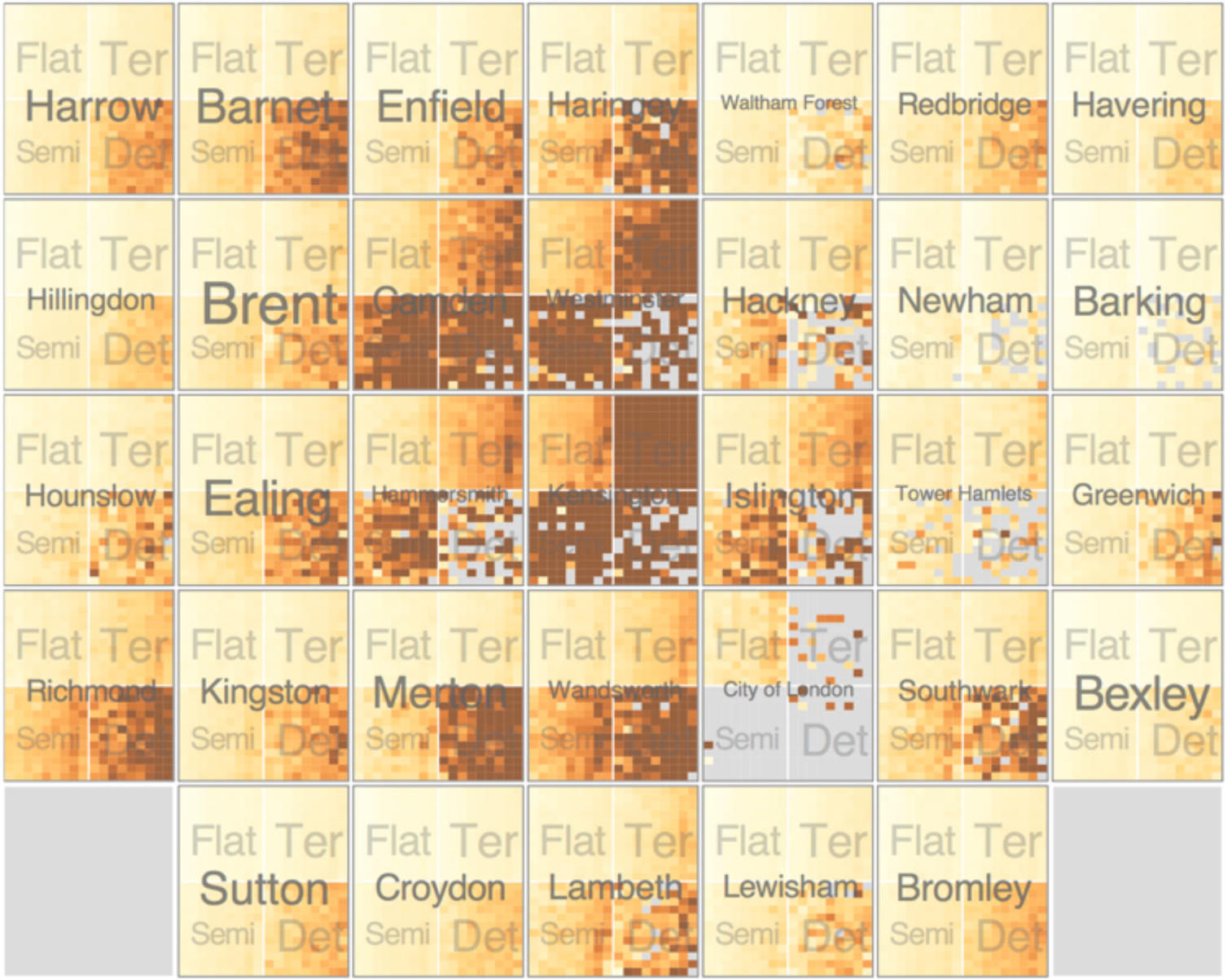


Partitioning: Recursive subdivision

System: **HIVE**

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

- neighborhood patterns
  - where it's expensive
  - where you pay much more for detached type



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

Partitioning: Recursive subdivision

System: **HIVE**

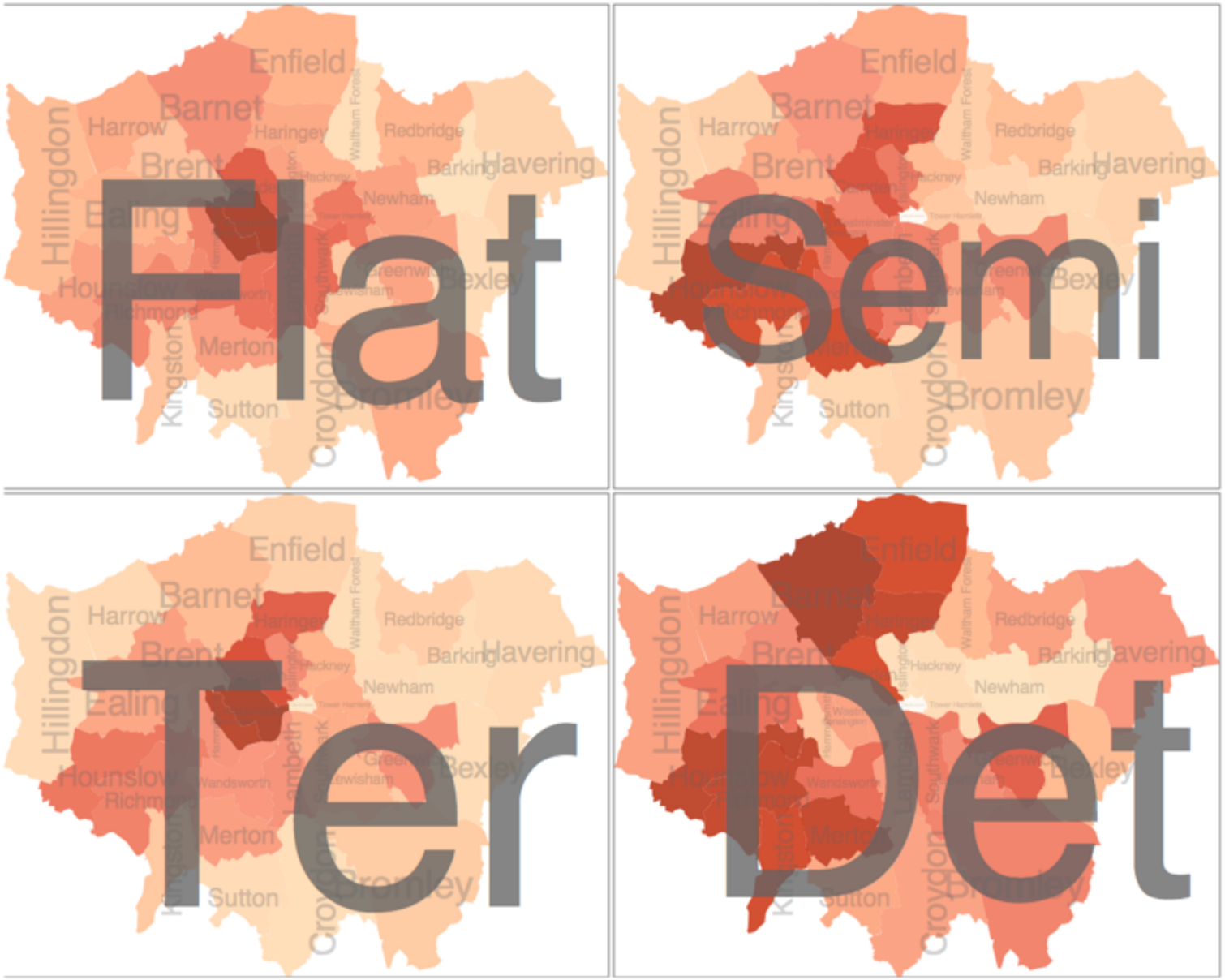
- switch order of splits
  - type then neighborhood
- switch color
  - by price variation
- type patterns
  - within specific type, which neighborhoods inconsistent

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

Partitioning: Recursive subdivision

System: **HIVE**

- different encoding for second-level regions
  - choropleth maps

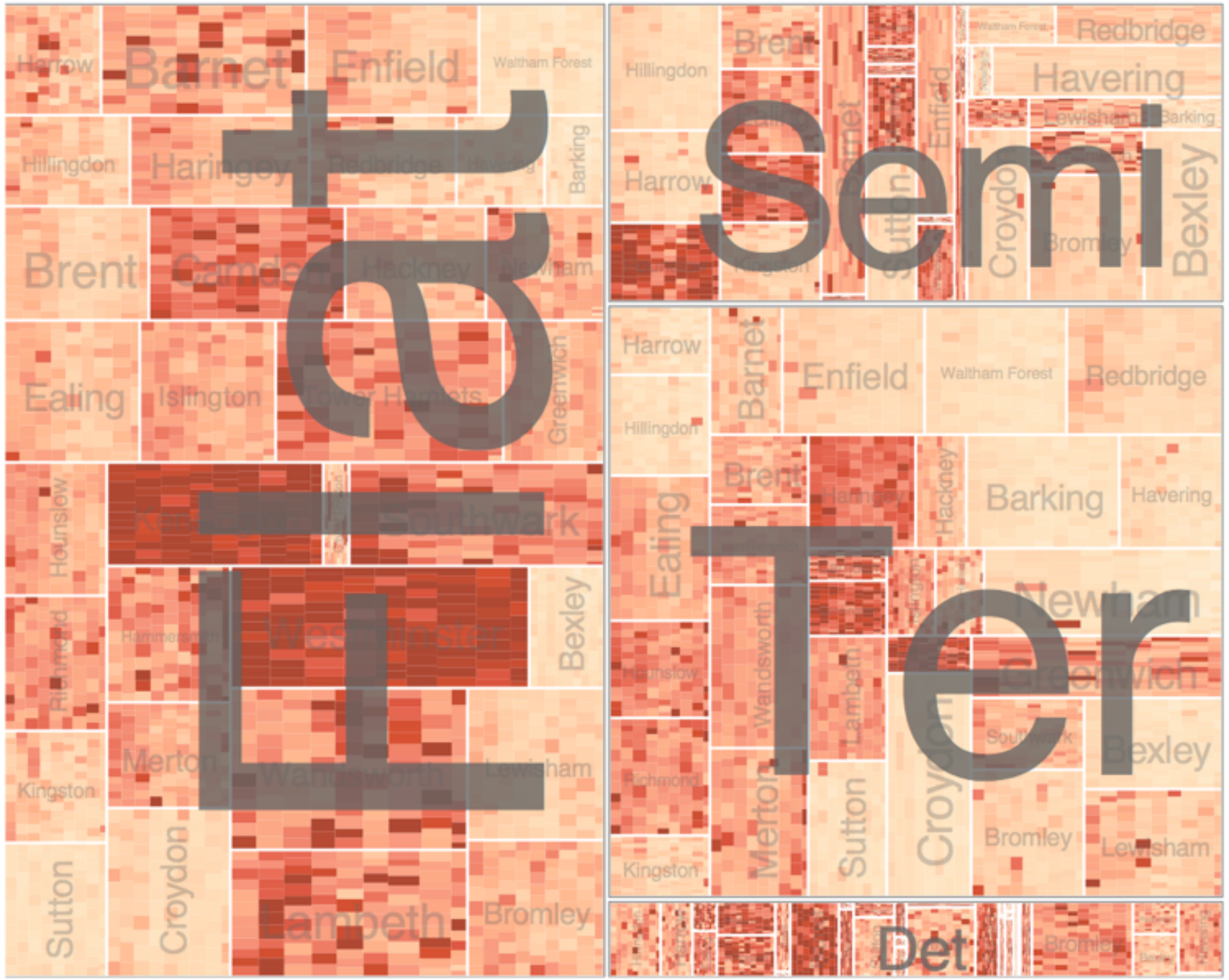


[Configuring Hierarchical Layouts to Address Research Questions. Slingsby, Dykes, and Wood. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2009) 15:6 (2009), 977–984.]

Partitioning: Recursive subdivision

System: **HIVE**

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



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

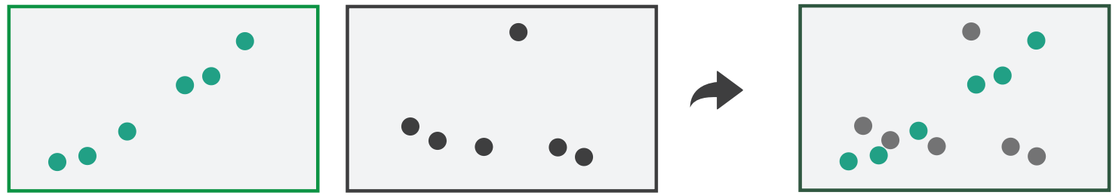
Previous Demos

- Tableau Lessons
  - partitioning: drag multiple pills into Row or Column
  - disaggregation: drag field into Detail/Color
    - aggregation is automatic and aggressive in Tableau

Superimpose layers

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

Superimpose Layers



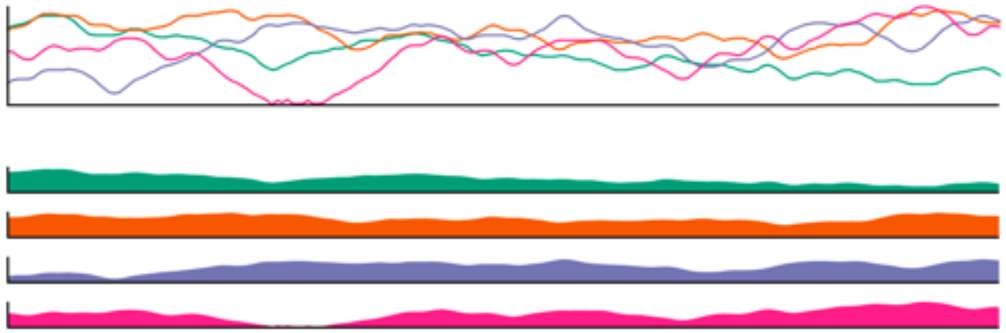
Static visual layering

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

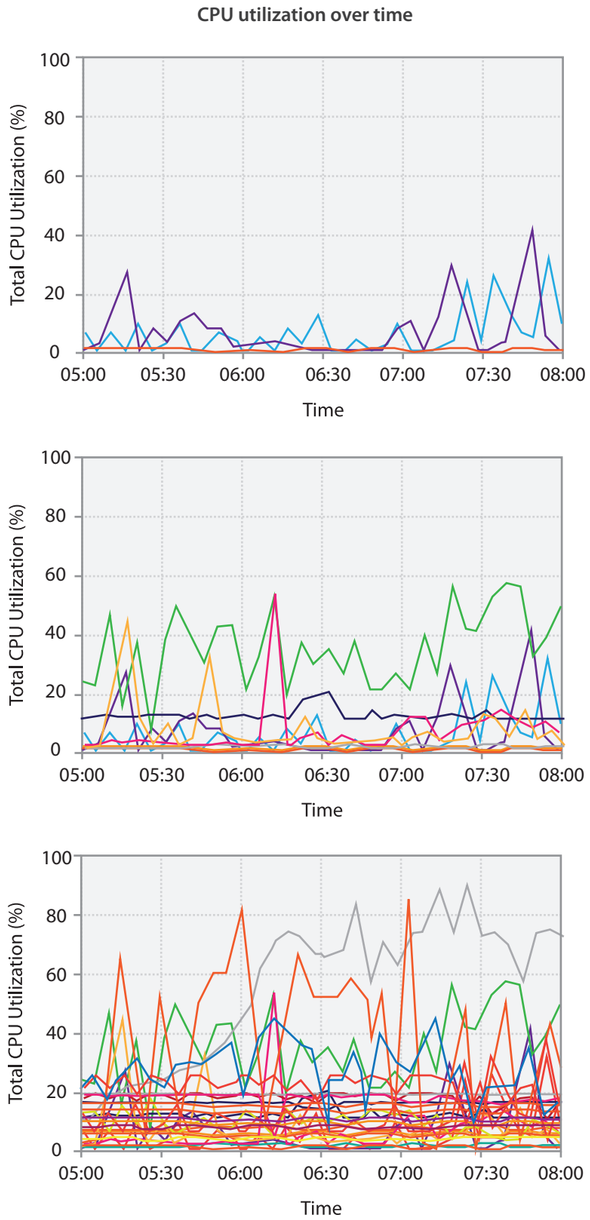
[Get it right in black and white. Stone. 2010. http://www.stonesc.com/wordpress/2010/03/get-it-right-in-black-and-white]

Superimposing limits

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

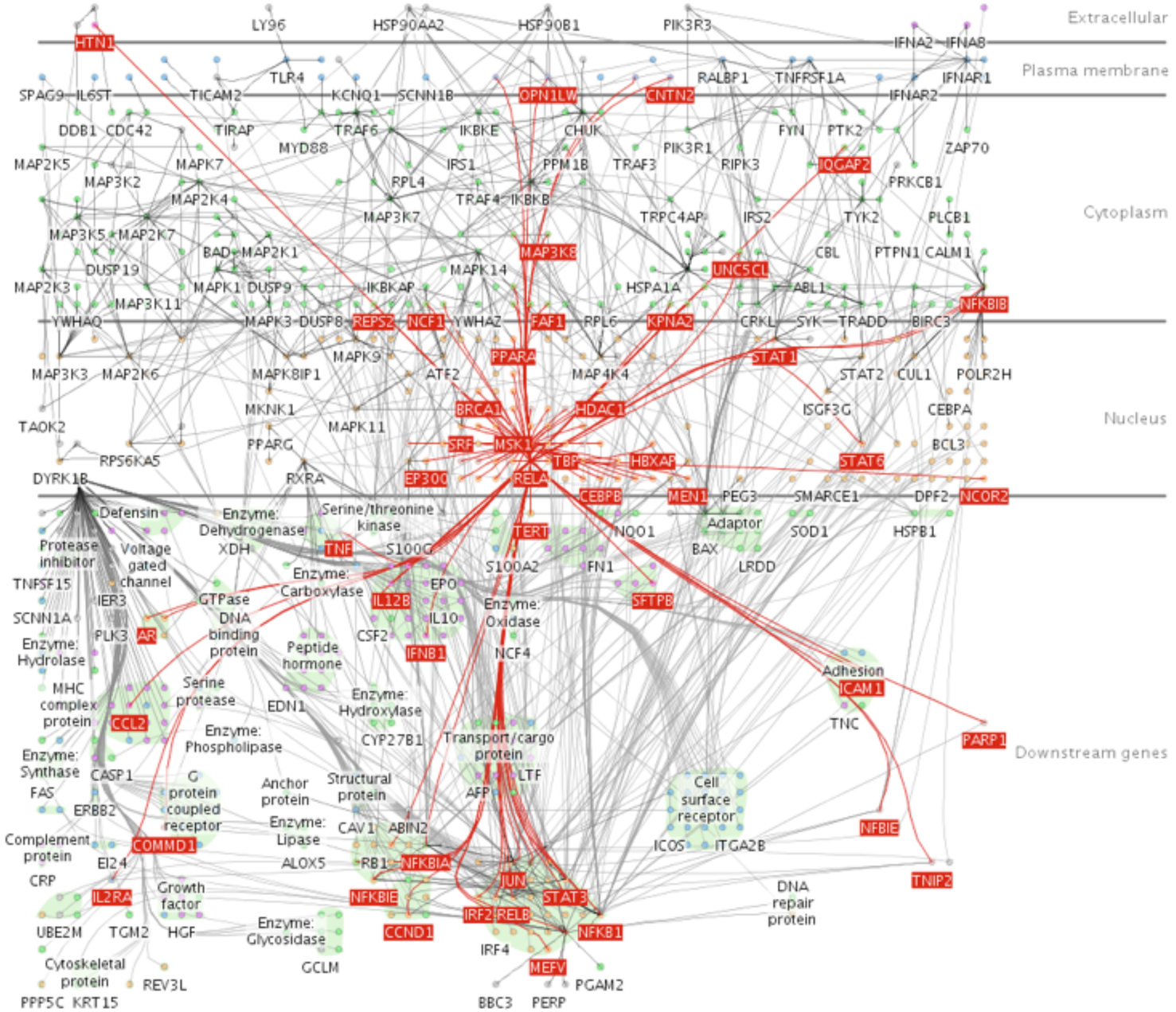


[Graphical Perception of Multiple Time Series. Javed, McDonnel, and Elmqvist. IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE InfoVis 2010) 16:6 (2010), 927–934.]



Dynamic visual layering

- 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.]

System: **Cerebral**

Reduce items and attributes

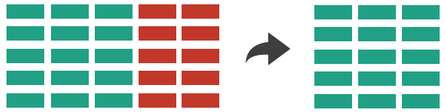
- reduce/increase: inverses
- filter
  - pro: straightforward and intuitive
    - to understand and compute
  - con: out of sight, out of mind
- aggregation
  - pro: inform about whole set
  - con: difficult to avoid losing signal
- not mutually exclusive
  - combine filter, aggregate
  - combine reduce,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

Reducing Items and Attributes

Filter

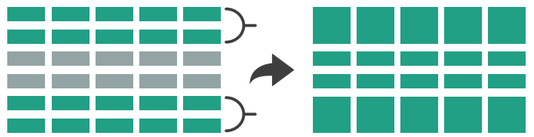
- Items

- Attributes

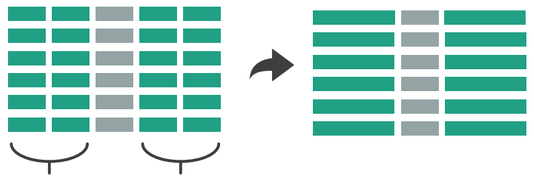


Aggregate

- Items



- Attributes



Reduce

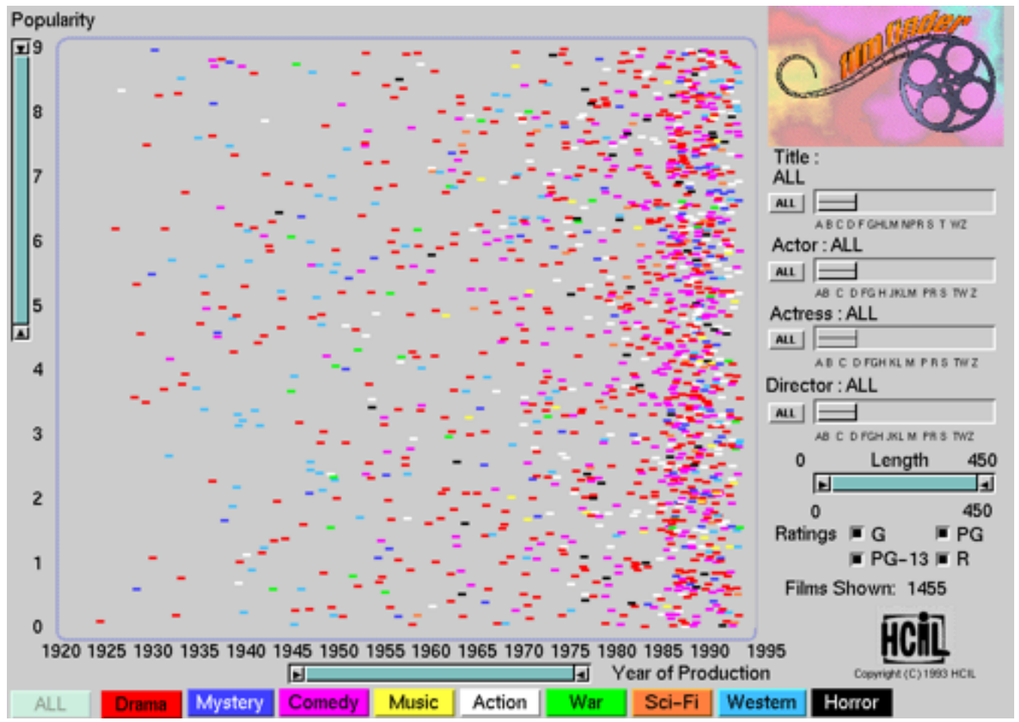
Filter

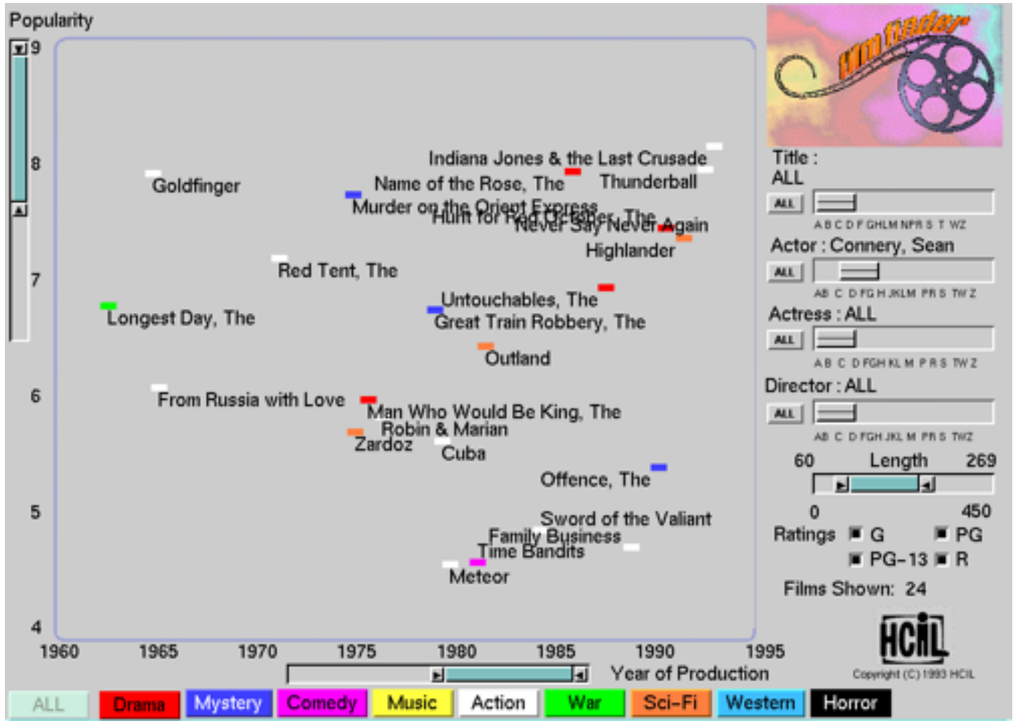
Aggregate

Embed

Idiom: **dynamic filtering** System: **FilmFinder**

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

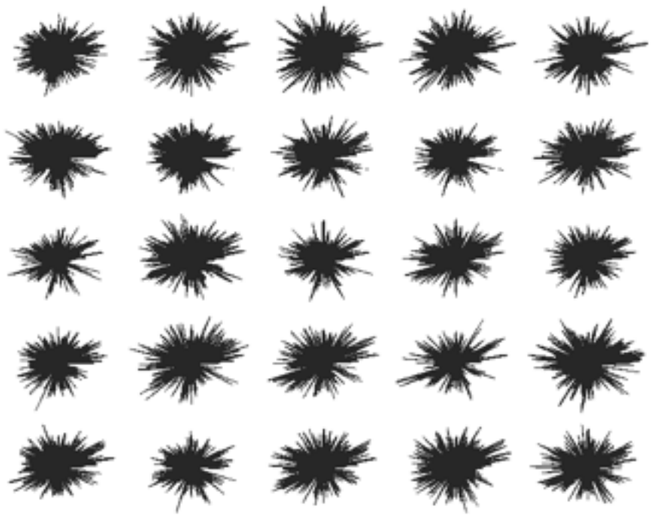


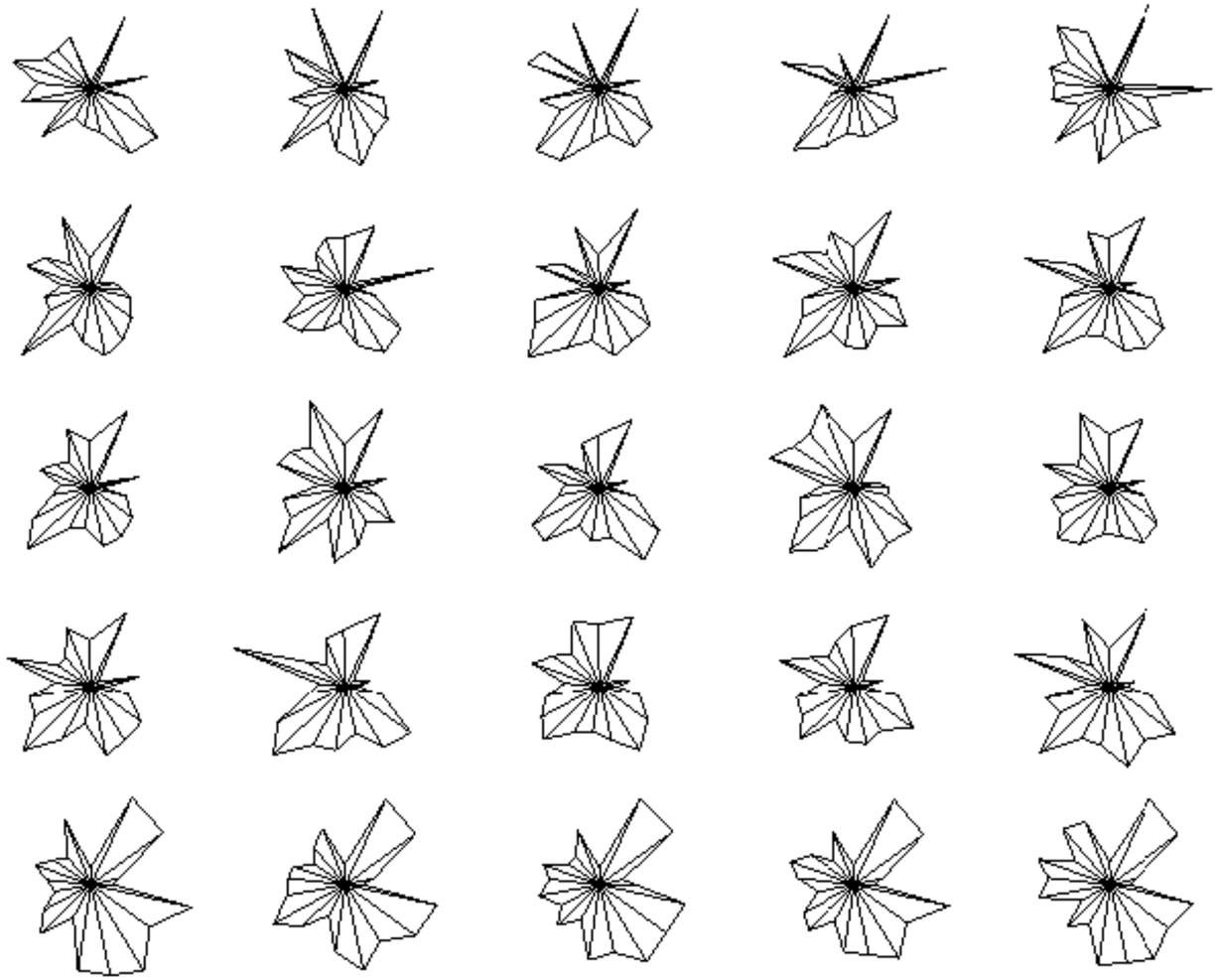


[Visual information seeking:Tight coupling of dynamic query filters with starfield displays. Ahlberg and Shneiderman. Proc.ACM Conf. on Human Factors in Computing Systems (CHI), pp. 313–317, 1994.]

Idiom: **DOSFA**

- attribute filtering
- encoding: star glyphs

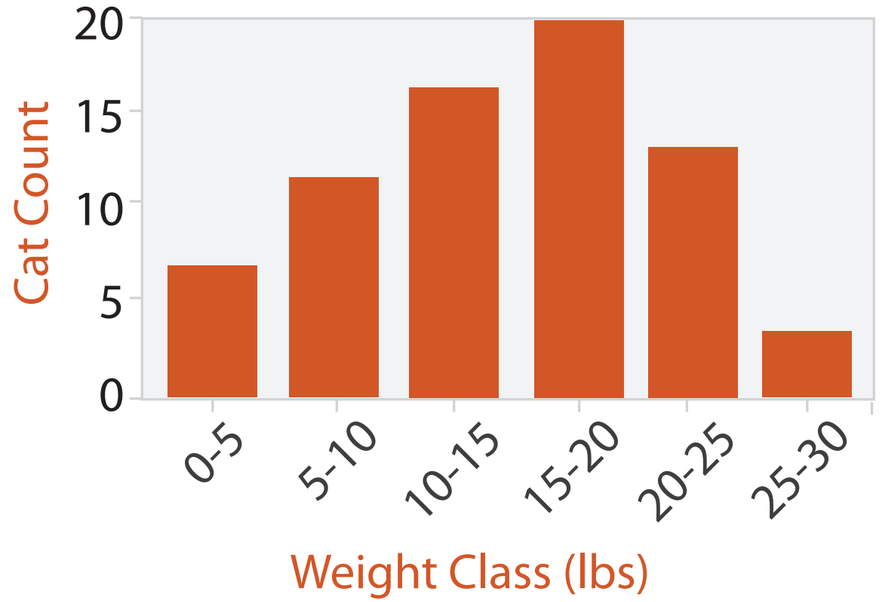




[Interactive Hierarchical Dimension Ordering, Spacing and Filtering for Exploration Of High Dimensional Datasets. Yang, Peng, Ward, and. Rundensteiner. Proc. IEEE Symp. Information Visualization (InfoVis), pp. 105–112, 2003.]

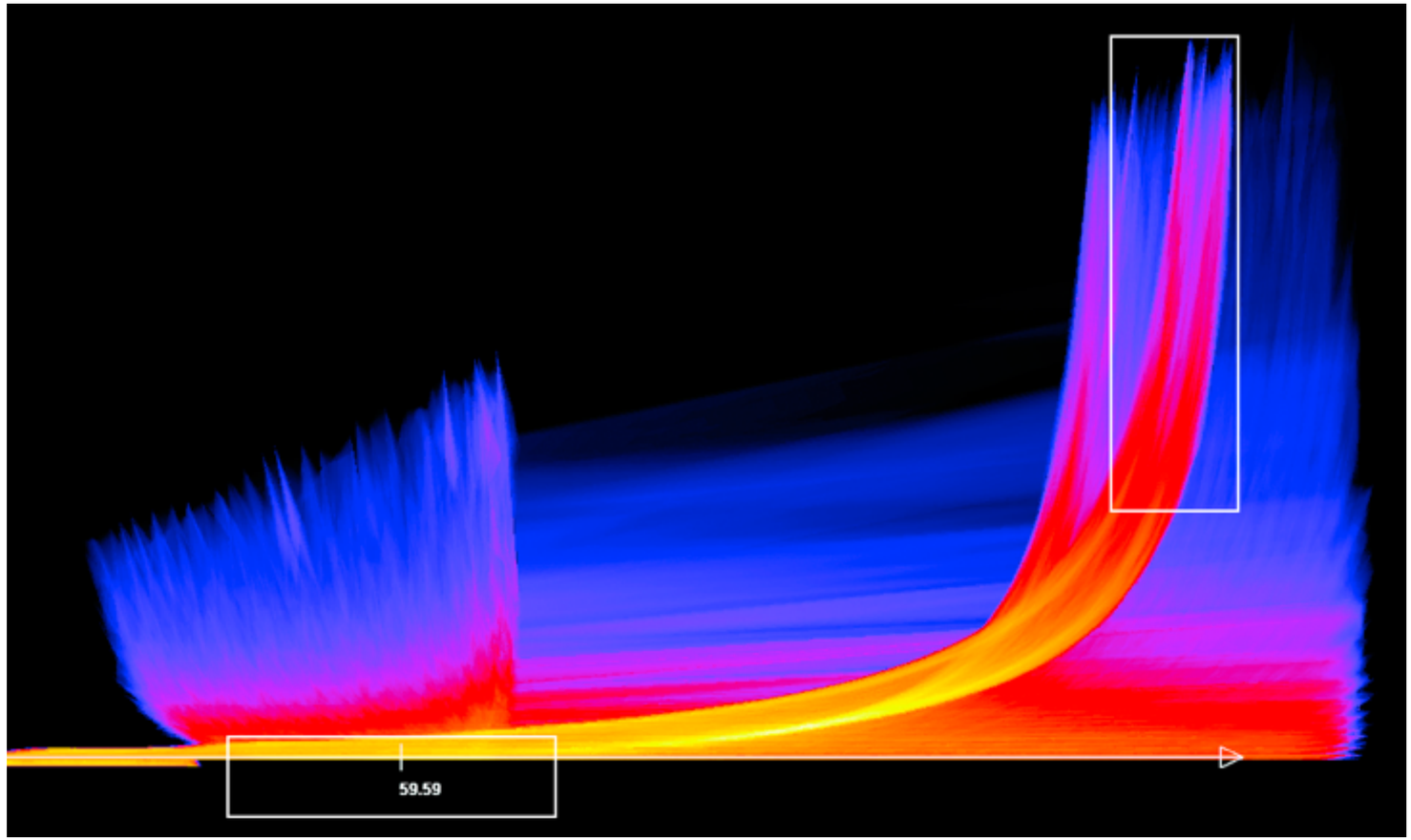
Idiom: **histogram**

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



Continuous scatterplot

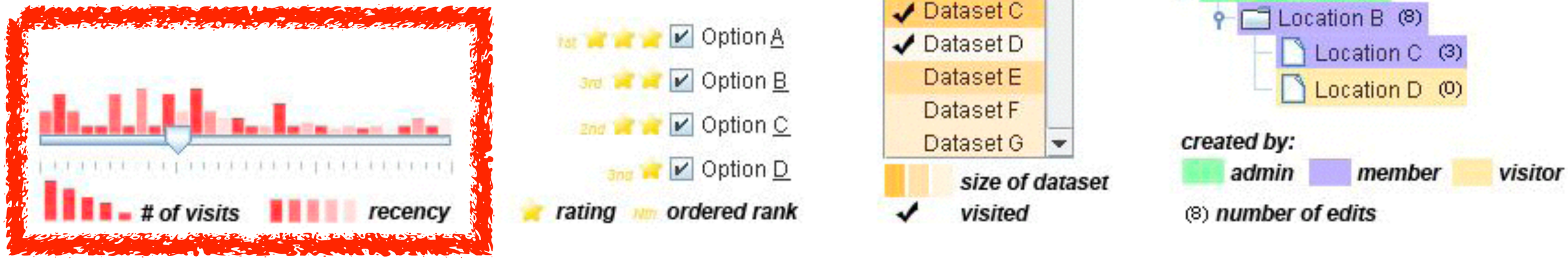
- static item aggregation
- data: table
- derived data: table
  - key attribs x,y for pixels
  - quant attrib: overplot density
- dense space-filling 2D matrix
- color: sequential categorical hue + ordered luminance colormap



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

Idiom: **scented widgets**

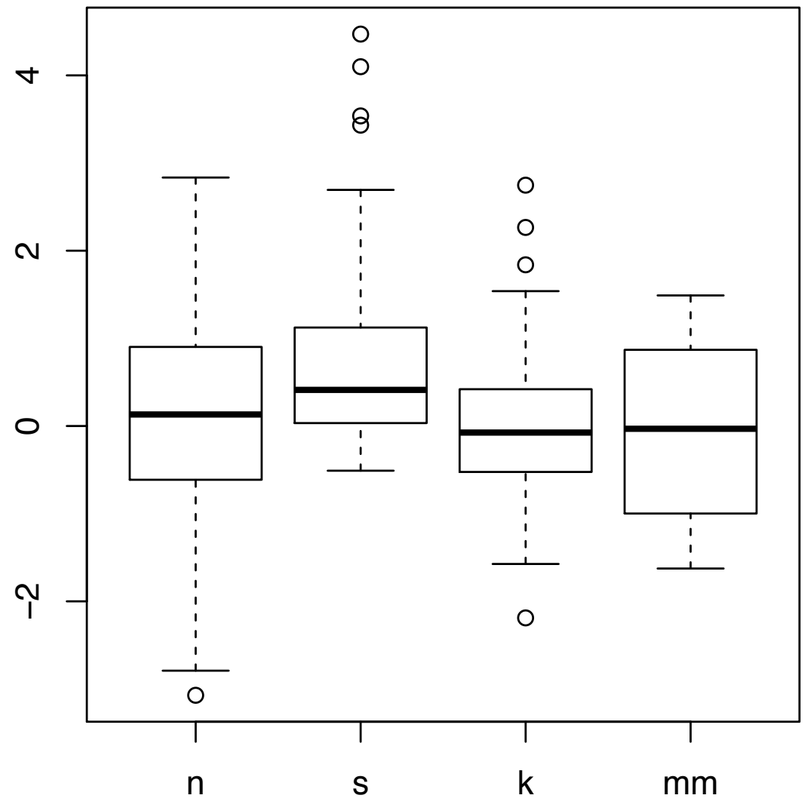
- augment widgets for filtering to show ***information scent***
  - cues to show whether value in drilling down further vs looking elsewhere
- concise, in part of screen normally considered control panel



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

Idiom: **boxplot**

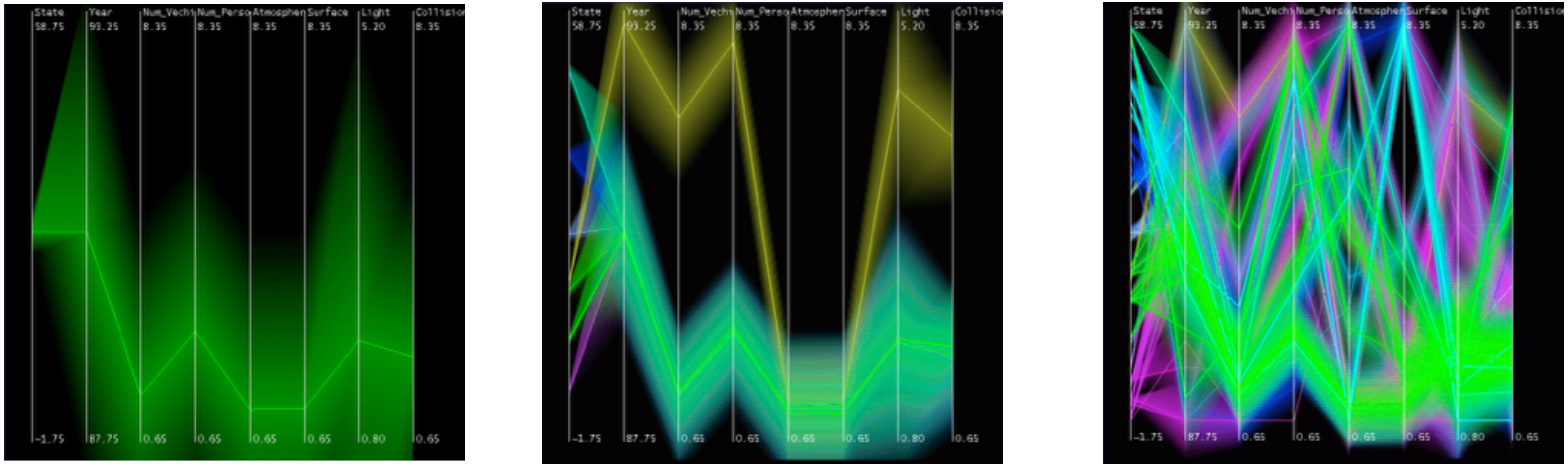
- static item aggregation
- task: find distribution
- data: table
- derived data
  - 5 quant attribs
    - median: central line
    - lower and upper quartile: boxes
    - lower upper fences: whiskers
      - values beyond which items are outliers
  - outliers beyond fence cutoffs explicitly shown



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

Idiom: **Hierarchical parallel coordinates**

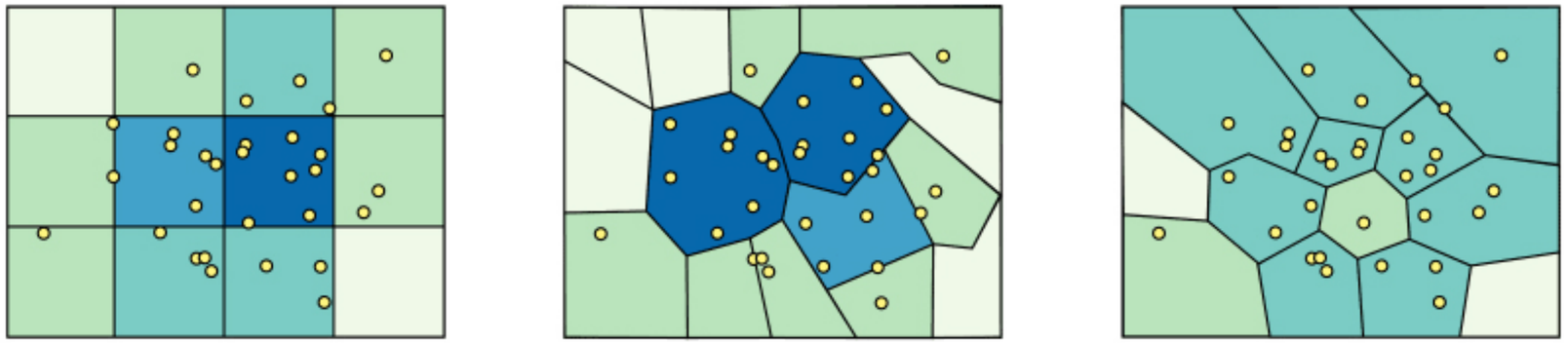
- dynamic item aggregation
- derived data: *hierarchical clustering*
- encoding:
  - –cluster band with variable transparency, line at mean, width by min/max values
  - –color by proximity in hierarchy



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

Spatial aggregation

- MAUP: Modifiable Areal Unit Problem
  - gerrymandering (manipulating voting district boundaries) is one example!



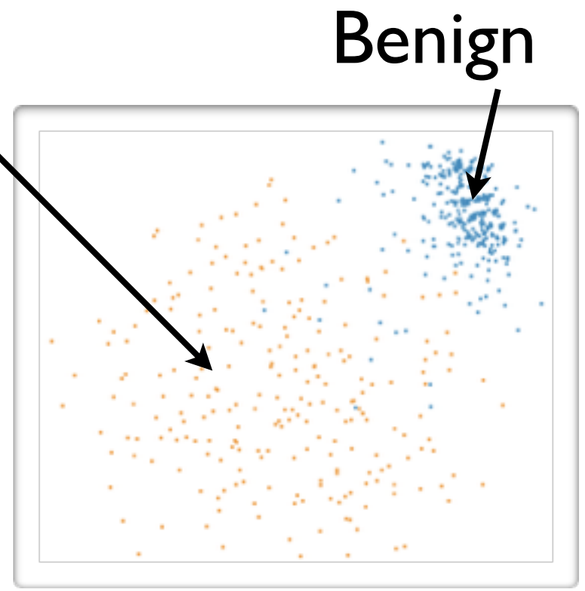
[http://www.e-education.psu/edu/geog486/l4_p7.html, Fig 4.cg.6]

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

Tumor
Measurement Data

→ DR →



data: 9D measured space

derived data: 2D target space

Idiom: **Dimensionality reduction for documents**

Task 1

In
HD data

Out
2D data

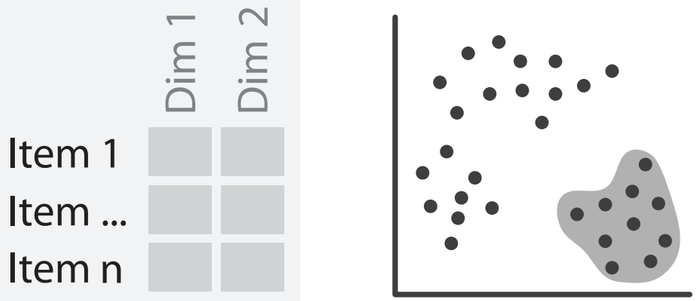
What?

- In High-dimensional data
- Out 2D data

Why?

- Produce
- Derive

Task 2



In
2D data

Out
Scatterplot
Clusters & points

What?

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

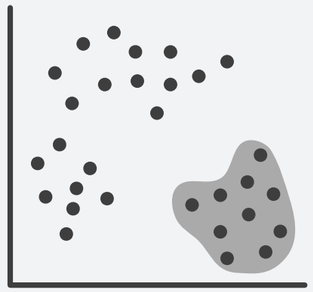
Why?

- Discover
- Explore
- Identify

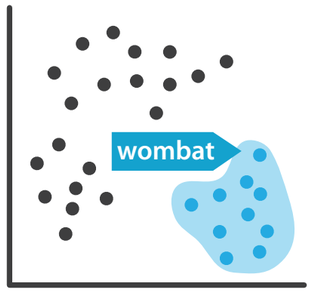
How?

- Encode
- Navigate
- Select

Task 3



In
Scatterplot
Clusters & points



Out
Labels for
clusters

What?

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

Why?

- Produce
- Annotate

Demo 2: Internet Use

- Credit: Ben Jones

- Tableau Lessons
  - more maps, dual axes
  - linked views (apply filter to selected worksheets)
  - actions: highlight/hover

- Big Ideas
  - Tableau interactivity defaults not necessarily what you want

Demo 3: House Price Index

- Credit: Robert Kosara, from TCC 2014 talk Recreating News Visualizations in Tableau

- Tableau Lessons
  - more calculated field practice
  - create parameter
  - reference lines
  - interactive sliders

- Big Ideas
  - calculated fields plus interactivity gives you a lot of power and flexibility

Assignment 4

- finish/review House Price Index workbook
- add interactivity to last week's story
  - update workbook
  - upload to Tableau Public
  - revise story to include embedded interactive
- final project proposal