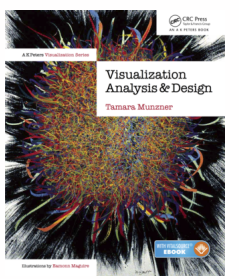


Visualization Analysis & Design



Tamara Munzner
Department of Computer Science
University of British Columbia

Data Visualization Masterclass: Principles, Tools, and Storytelling
June 8-20 15, VIVID, Sydney Australia

<http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney> @tamaramunzner

Outline

- **Session 1: Principles 9:30-11:00am**
 - Analysis: What, Why, How
 - Marks and Channels
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks and Trees
 - Map Color and Other Channels
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

<http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>

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.

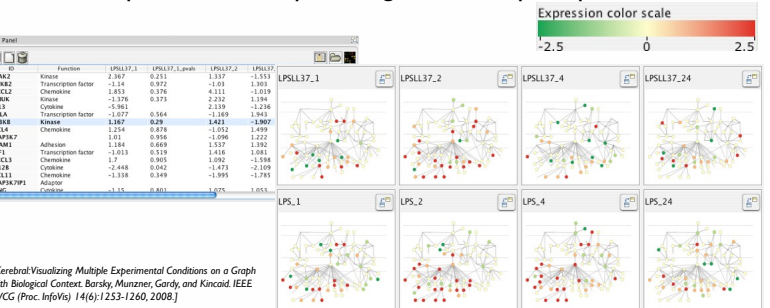
Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- 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



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

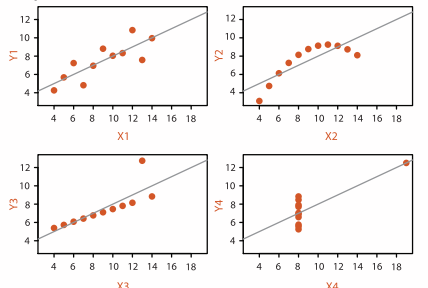
Why represent all the data?

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	8
y variance	4
x/y correlation	1



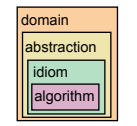
Why are there resource limitations?

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

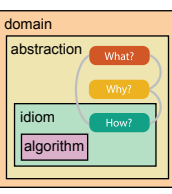
- computational limits
 - processing time
 - system memory
- human limits
 - human attention and memory
- display limits
 - pixels are precious resource, the most constrained resource
 - **information density**: ratio of space used to encode info vs unused whitespace
 - tradeoff between clutter and wasting 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. Behrer and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

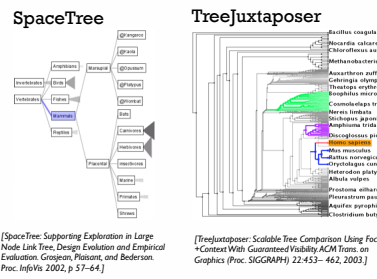
Validation methods from different fields for each level

anthropology/ ethnography	<ul style="list-style-type: none"> • Domain situation: Observe target users using existing tools • Data/task abstraction • Visual encoding/interaction idiom: Justify design with respect to alternatives • Algorithm: Measure system time/memory, Analyze computational complexity • Analyze results qualitatively, Measure human time with lab experiment (lab study) • Observe target users after deployment (field study), Measure adoption 	design
anthropology/ ethnography		computer science
		cognitive psychology

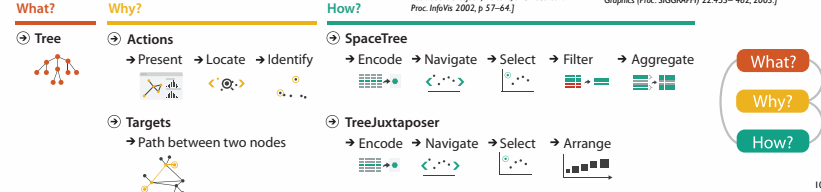
- mismatch: cannot show idiom good with system timings
- mismatch: cannot show abstraction good with lab study

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

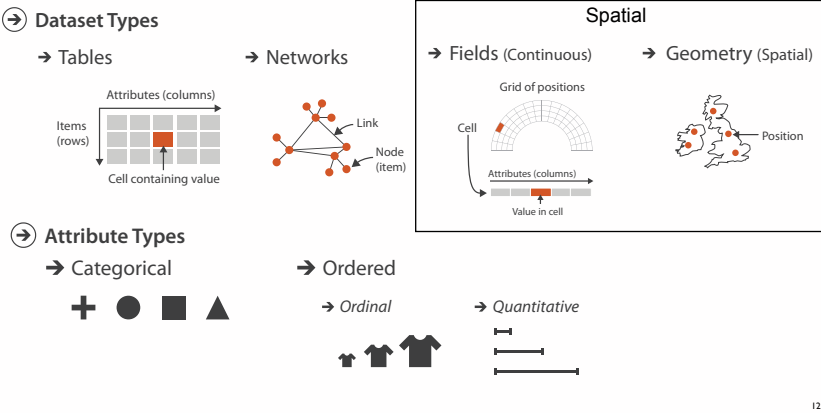


[SpaceTree: Supporting Exploration in Large Node Link Trees, Design Evolution and Empirical Evaluation. Garsner, Plaisant, and Bederson. Proc. InfoVis 2002, p. 57-64.]



What?	
Datasets	Attributes
<ul style="list-style-type: none"> • Data Types: Items, Attributes, Links, Positions, Grids • Data and Dataset Types: Tables, Networks & Trees, Fields, Geometry, Clusters, Sets, Lists • Dataset Types: Tables, Networks, Fields (Continuous), Geometry (Spatial) 	<ul style="list-style-type: none"> • Attribute Types: Categorical, Ordered, Quantitative • Ordering Direction: Sequential, Diverging, Cyclic • Dataset Availability: Static, Dynamic

Dataset and data types

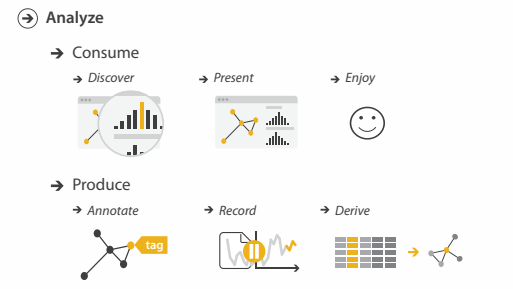


Why?	
Actions	Targets
<ul style="list-style-type: none"> • Analyze: Consume, Discover, Produce, Search, Query 	<ul style="list-style-type: none"> • All Data: Trends, Outliers, Features • Attributes: One, Many, Distribution, Extremes • Network Data: Topology, Paths • Spatial Data: Shape

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

Actions I: Analyze

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



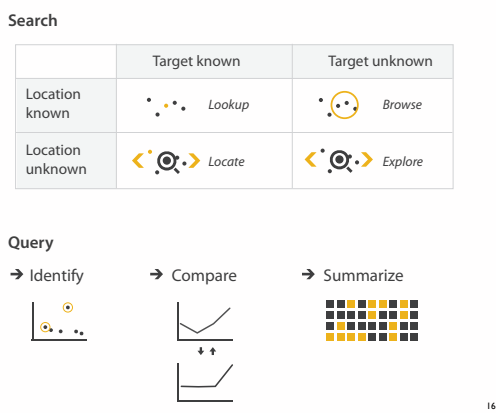
Actions II: Search

- what does user know?
 - target, location

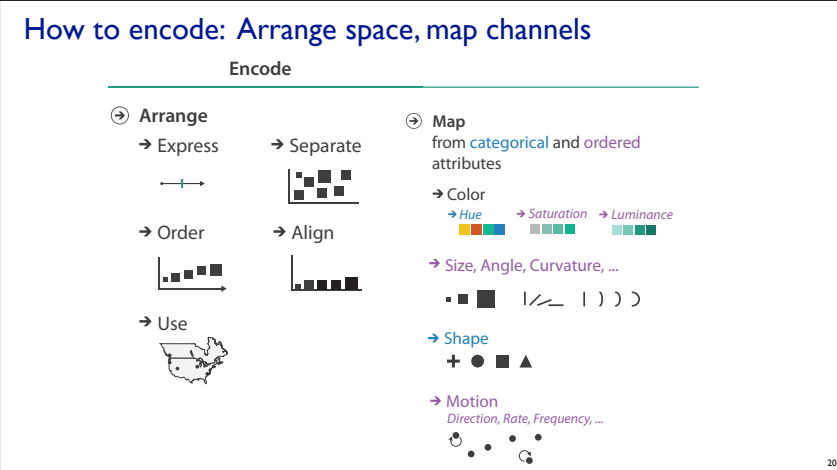
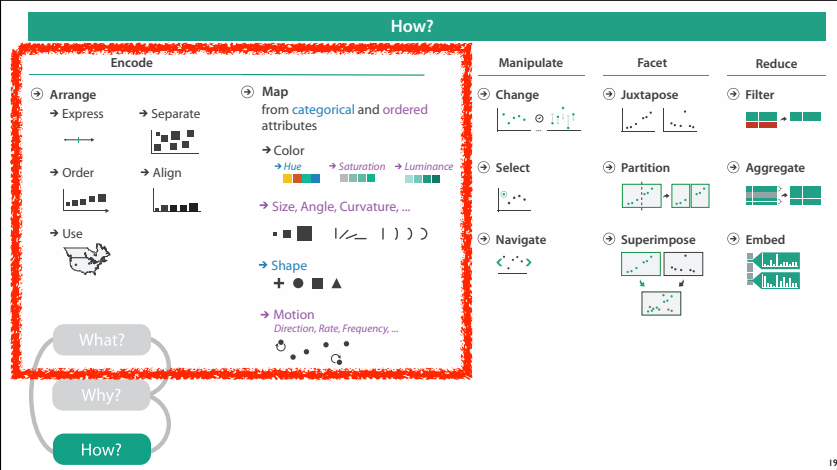
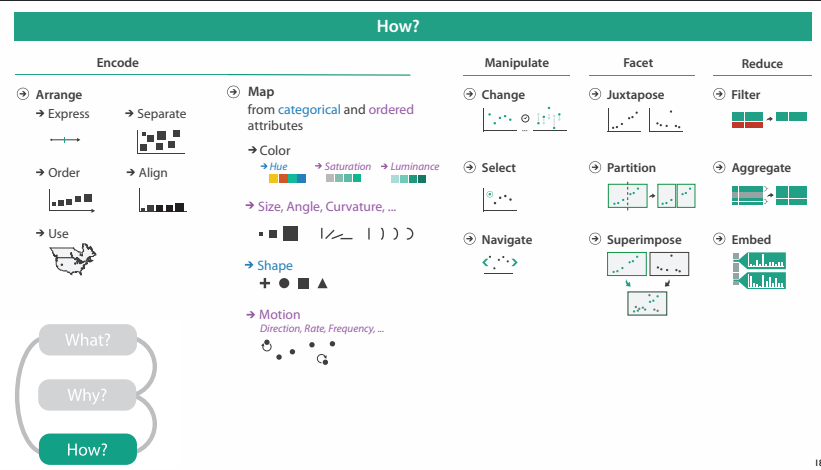
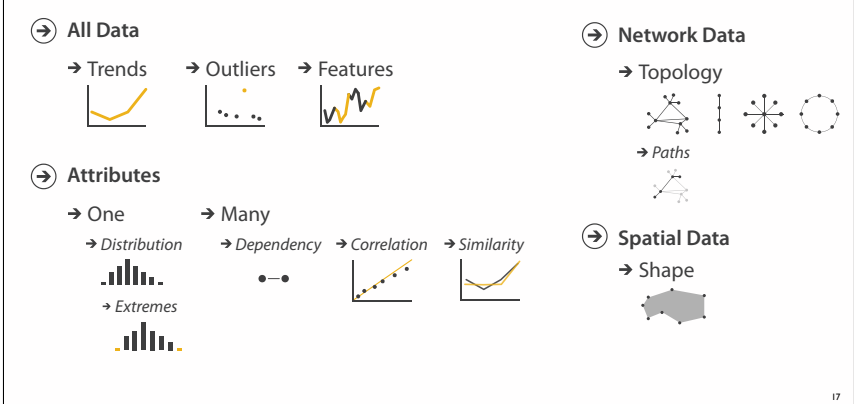
	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Actions III: Query

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

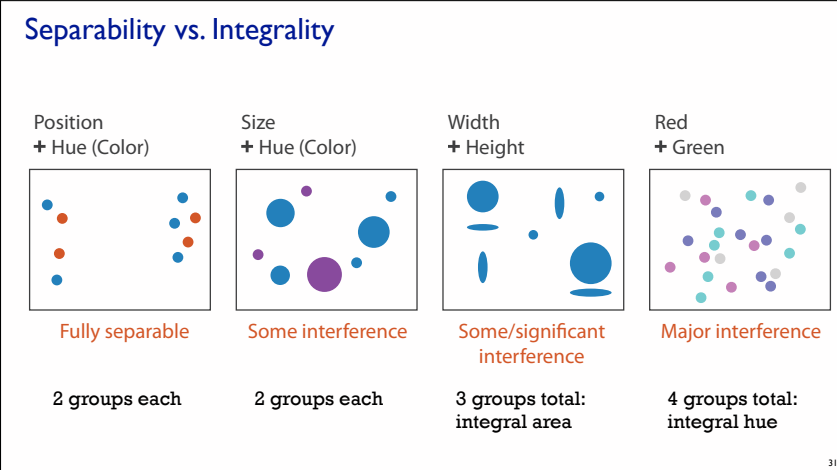
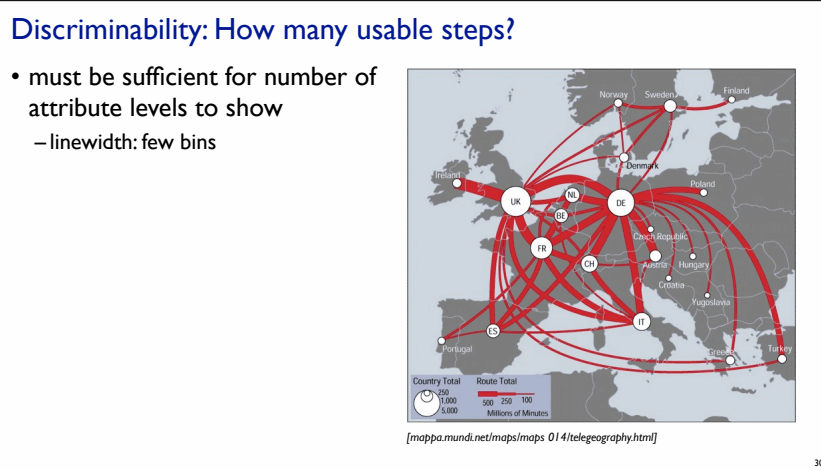
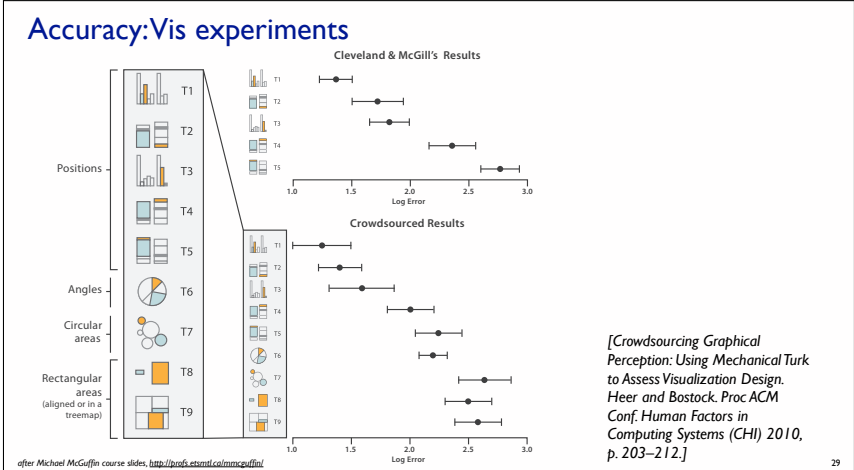
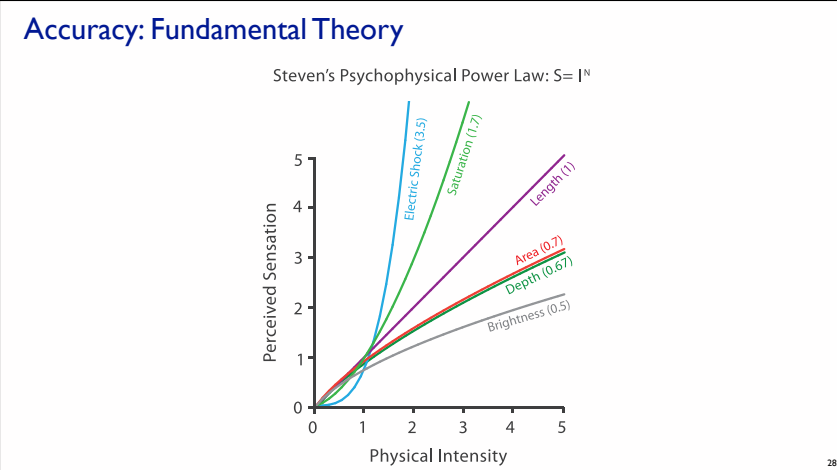
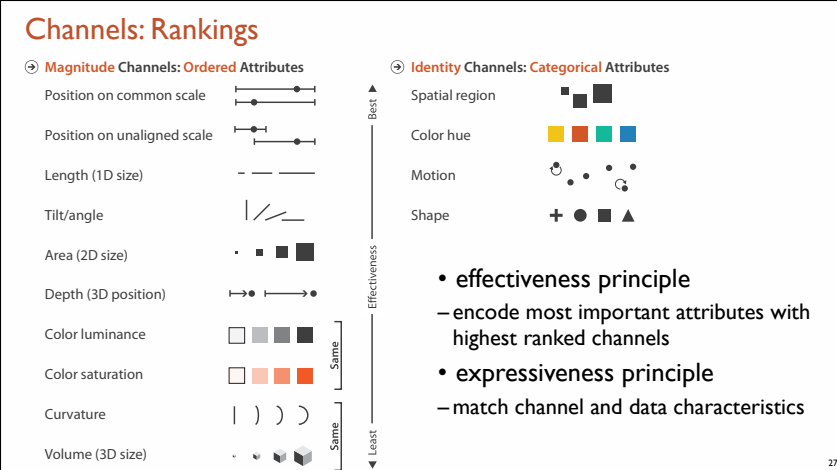
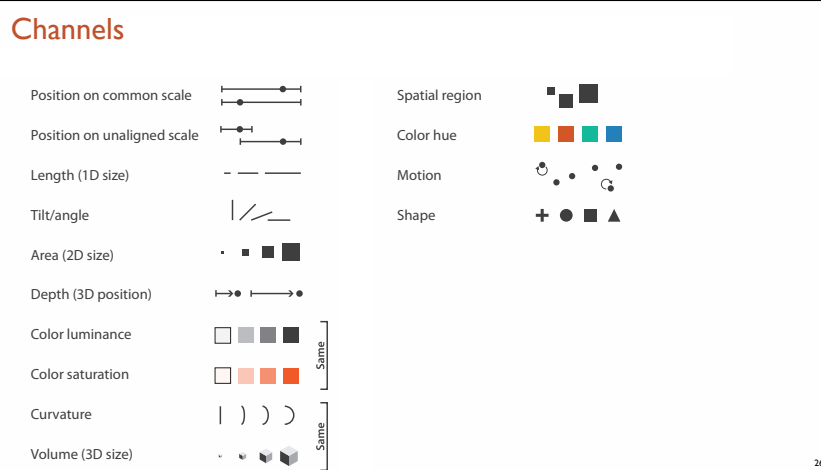
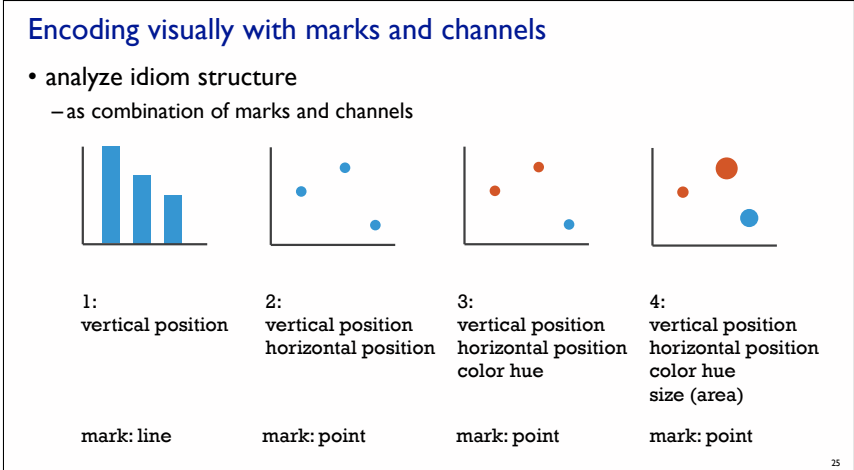
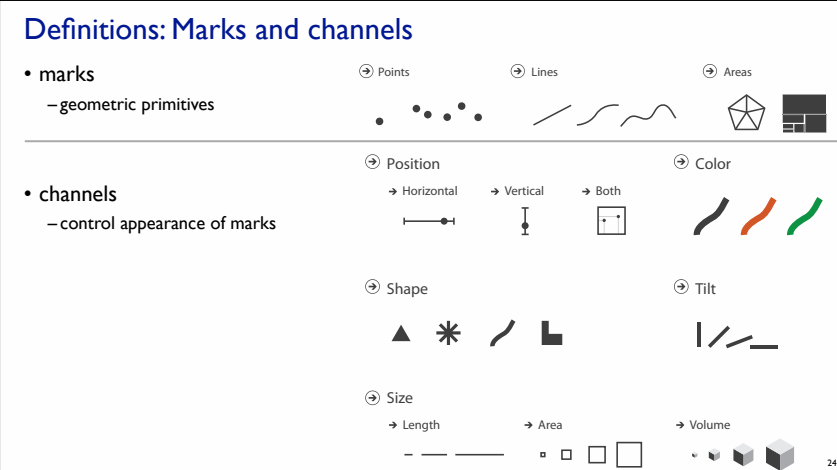
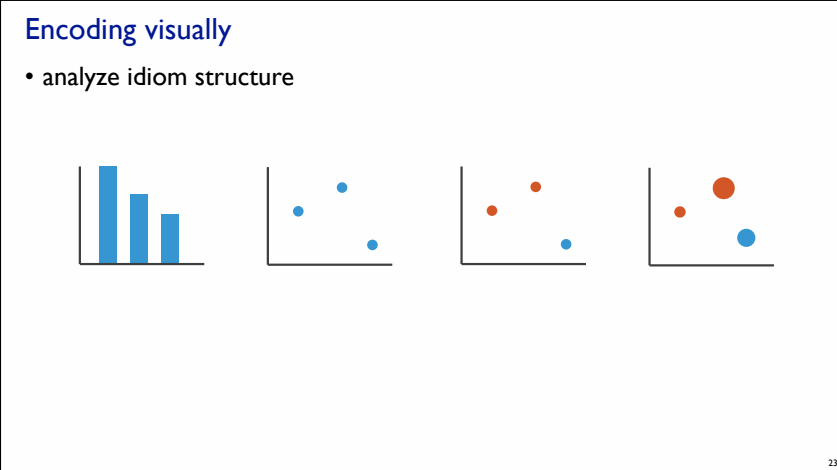


Targets



- ### 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 9:30-11:00am
 - Analysis: What, Why, How
 - Marks and Channels
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks and Trees
 - Map Color and Other Channels
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate
- <http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>



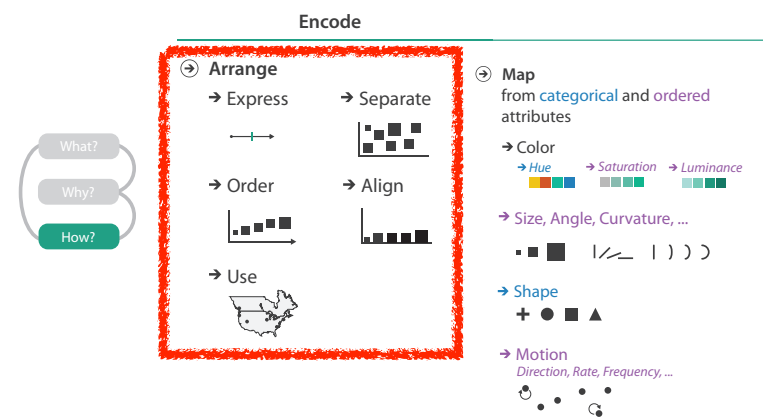
- ### Further reading
- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, Nov 2014.
 - Chap 5: Marks and Channels
 - On the Theory of Scales of Measurement. Stevens. Science 103:2684 (1946), 677–680.
 - Psychophysics: Introduction to its Perceptual, Neural, and Social 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

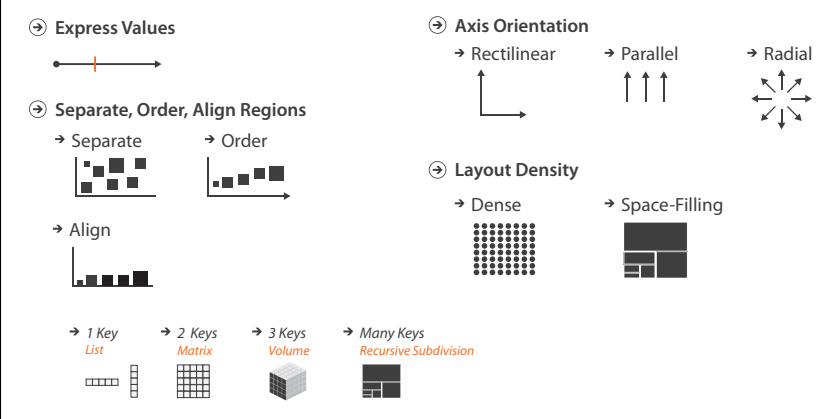
- **Session 1: Principles 9:30-11:00am**
 - Analysis: What, Why, How
 - Marks and Channels
 - **Arrange Tables**
 - Arrange Spatial Data
 - Arrange Networks and Trees
 - Map Color and Other Channels
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

<http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>

How to encode: Arrange position and region



Arrange tables

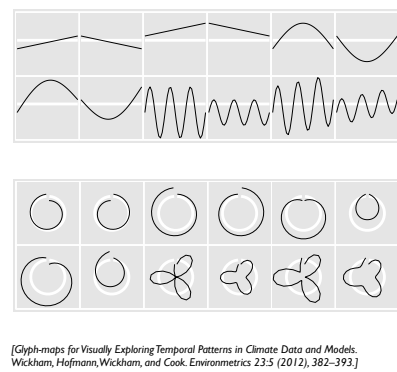
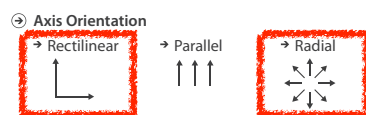


Idioms: dot chart, line chart

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

Idiom: glyphmaps

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



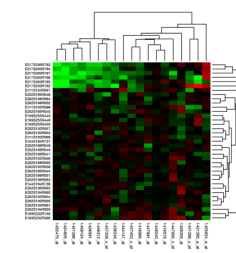
[Glyph-maps for Visually Exploring Temporal Patterns in Climate Data and Models. Wickham, Hofmann, Wickham, and Cook. *Environmetrics* 23.5 (2012), 382-393.]

Idiom: heatmap

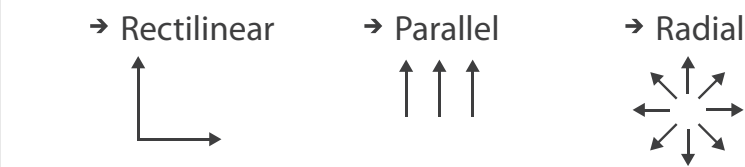
- two keys, one value
 - data
 - 2 categ attribs (gene, experimental condition)
 - 1 quant attrib (expression levels)
 - marks: area
 - separate and align in 2D matrix
 - indexed by 2 categorical attributes
 - channels
 - color by quant attrib
 - (ordered diverging colormap)
 - task
 - find clusters,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

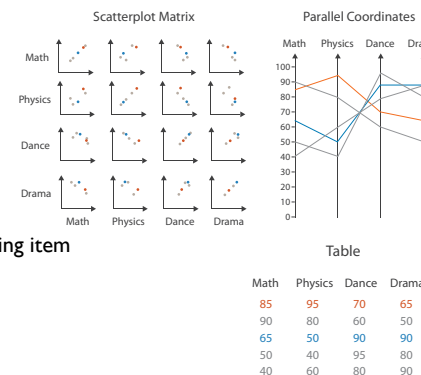


Axis Orientation



Idioms: scatterplot matrix, parallel coordinates

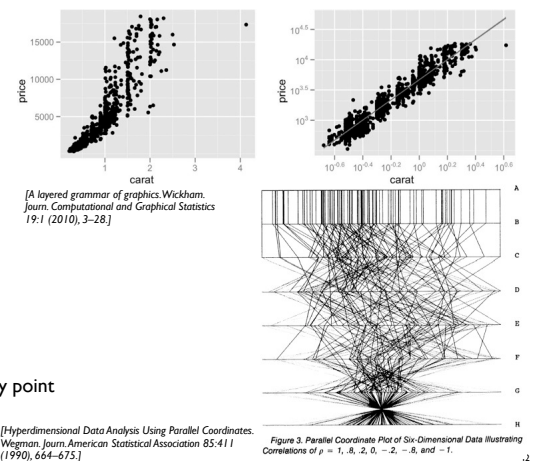
- scatterplot matrix (SPLOM)
 - rectilinear axes, point mark
 - all possible pairs of axes
 - scalability
 - one dozen attribs
 - dozens to hundreds of items
- parallel coordinates
 - parallel axes, jagged line representing item
 - rectilinear axes, item as point
 - axis ordering is major challenge
 - scalability
 - dozens of attribs
 - hundreds of items



offer [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



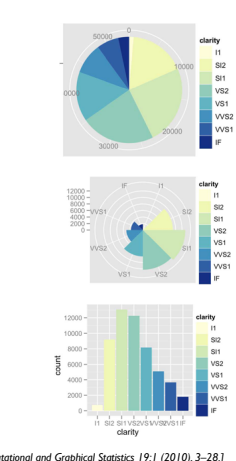
[A layered grammar of graphics. Wickham. *Journ. Computational and Graphical Statistics* 19:1 (2010), 3-28.]

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

Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of $\rho = 1, .8, .2, 0, -.2, -.4,$ and -1 .

Idioms: pie chart, polar area chart

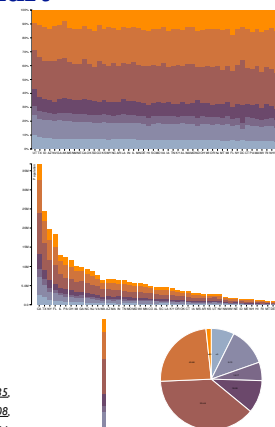
- pie chart
 - area marks with angle channel
 - accuracy: angle/area much less accurate than line length
- polar area chart
 - area marks with length channel
 - more direct analog to bar charts
- data
 - 1 categ key attrib, 1 quant value attrib
- task
 - part-to-whole judgements



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

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



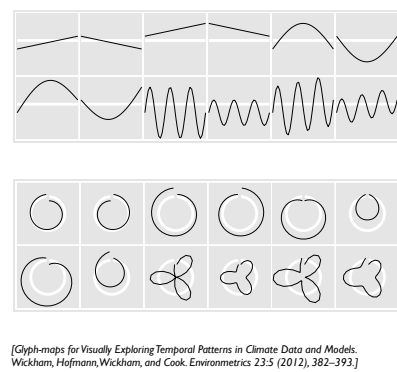
<http://tbllocks.org/imbstock/3887235>

<http://tbllocks.org/imbstock/3886708>

<http://tbllocks.org/imbstock/3886394>

Idiom: glyphmaps

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

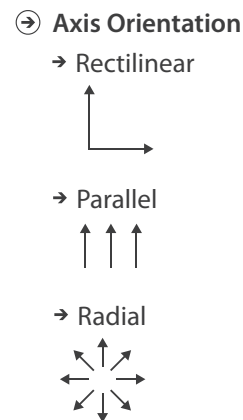


[Glyph-maps for Visually Exploring Temporal Patterns in Climate Data and Models. Wickham, Hofmann, Wickham, and Cook. *Environmetrics* 23.5 (2012), 382-393.]

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.]



Further reading

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

Outline

- **Session 1: Principles 9:30-11:00am**
 - Analysis: What, Why, How
 - Marks and Channels
 - Arrange Tables
 - **Arrange Spatial Data**
 - Arrange Networks and Trees
 - Map Color and Other Channels
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

<http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>

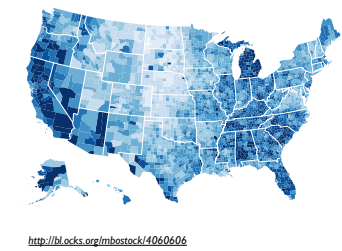
Arrange spatial data

- Use Given
 - Geometry
 - Geographic
 - Other Derived
 - Spatial Fields
 - Scalar Fields (one value per cell)
 - Isocontours
 - Direct Volume Rendering
 - Vector and Tensor Fields (many values per cell)
 - Flow Glyphs (local)
 - Geometric (sparse seeds)
 - Textures (dense seeds)
 - Features (globally derived)



Idiom: choropleth map

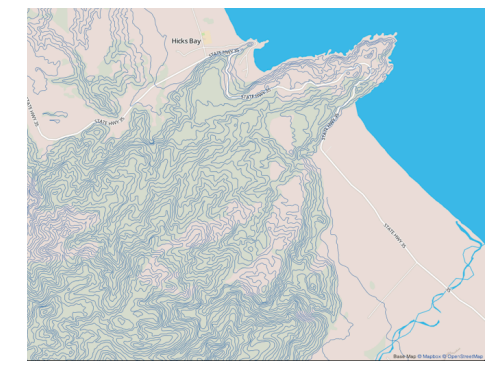
- use given spatial data
 - when central task is understanding spatial relationships
- data
 - geographic geometry
 - table with 1 quant attribute per region
- encoding
 - use given geometry for area mark boundaries
 - sequential segmented colormap



<http://bl.ocks.org/mbostock/14060606>

Idiom: topographic map

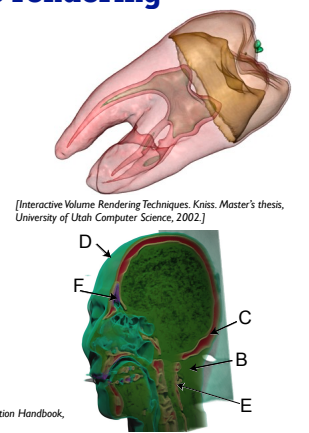
- data
 - geographic geometry
 - scalar spatial field
 - 1 quant attribute per grid cell
- derived data
 - isoline geometry
 - isocontours computed for specific levels of scalar values



Land Information New Zealand Data Service

Idioms: isosurfaces, direct volume rendering

- data
 - scalar spatial field
 - 1 quant attribute per grid cell
- task
 - shape understanding, spatial relationships
- isosurface
 - derived data: isocontours computed for specific levels of scalar values
- direct volume rendering
 - transfer function maps scalar values to color, opacity
 - no derived geometry

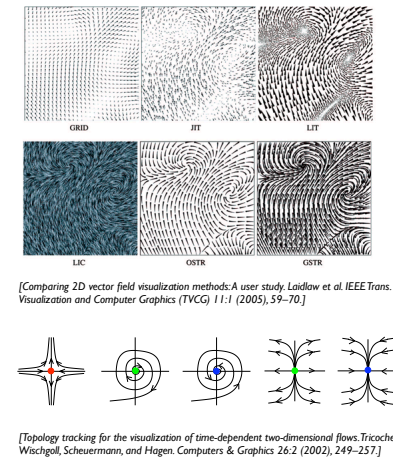


[Interactive Volume Rendering Techniques. Kniss. Master's thesis, University of Utah Computer Science, 2002.]

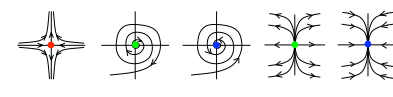
[Multidimensional Transfer Functions for Volume Rendering. Kniss, Kindmann, and Hansen. In The Visualization Handbook, edited by Charles Hansen and Christopher Johnson, pp. 189–210. Elsevier, 2005.]

Idioms: vector glyphs

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



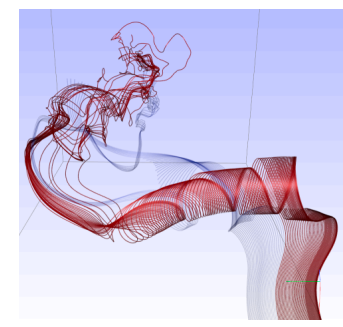
[Comparing 2D vector field visualization methods: A user study. Laidlaw et al. IEEE Trans. Visualization and Computer Graphics (TVCG) 11:1 (2005), 59–70.]



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

Idiom: similarity-clustered streamlines

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



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

Further reading

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

Outline

- Session 1: Principles 9:30-11:00am
 - Analysis: What, Why, How
 - Marks and Channels
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks and Trees
 - Map Color and Other Channels
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

<https://www.cs.ubc.ca/~trmm/talks.html#vad15sydney>

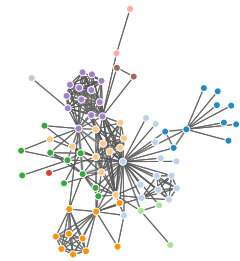
Arrange networks and trees

- Node-Link Diagrams
 - Connection Marks
 - NETWORKS
 - TREES
- Adjacency Matrix
 - Derived Table
 - NETWORKS
 - TREES
- Enclosure
 - Containment Marks
 - NETWORKS
 - TREES



Idiom: force-directed placement

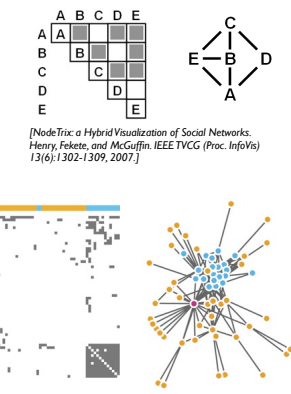
- visual encoding
 - link connection marks, node point marks
- considerations
 - spatial position: no meaning directly encoded
 - left free to minimize crossings
 - proximity semantics?
 - sometimes meaningful
 - sometimes arbitrary, artifact of layout algorithm
 - tension with length
 - long edges more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



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

Idiom: adjacency matrix view

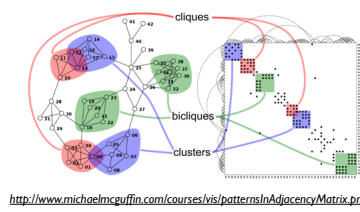
- data: network
 - transform into same data/encoding as heatmap
- derived data: table from network
 - 1 quant attrib
 - weighted edge between nodes
 - 2 categ attribs: node list x 2
- visual encoding
 - cell shows presence/absence of edge
- scalability
 - 1K nodes, 1M edges



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

Connection vs. adjacency comparison

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

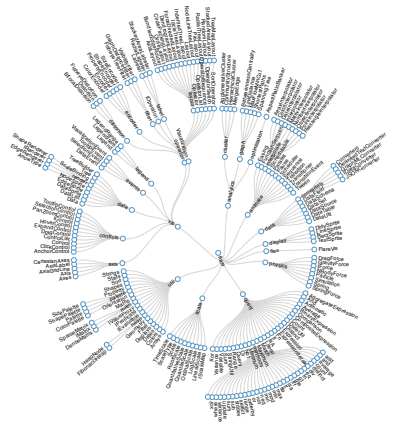


<http://www.michaelmcguffin.com/courses/vis/patterns/AdjacencyMatrix.png>

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

Idiom: radial node-link tree

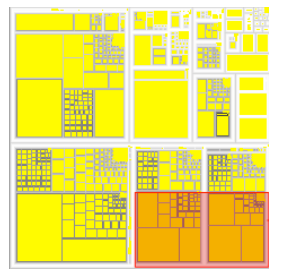
- data
 - tree
- encoding
 - link connection marks
 - point node marks
 - radial axis orientation
 - angular proximity: siblings
 - distance from center: depth in tree
- tasks
 - understanding topology, following paths
- scalability
 - 1K - 10K nodes



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

Idiom: treemap

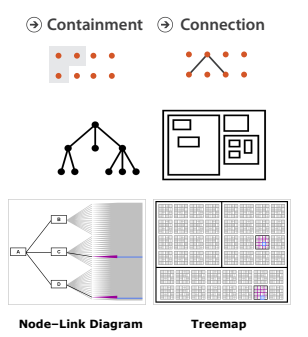
- data
 - tree
 - 1 quant attrib at leaf nodes
- encoding
 - area containment marks for hierarchical structure
 - rectilinear orientation
 - size encodes quant attrib
- tasks
 - query attribute at leaf nodes
- scalability
 - 1M leaf nodes



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

Connection vs. containment comparison

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



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

Further reading

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

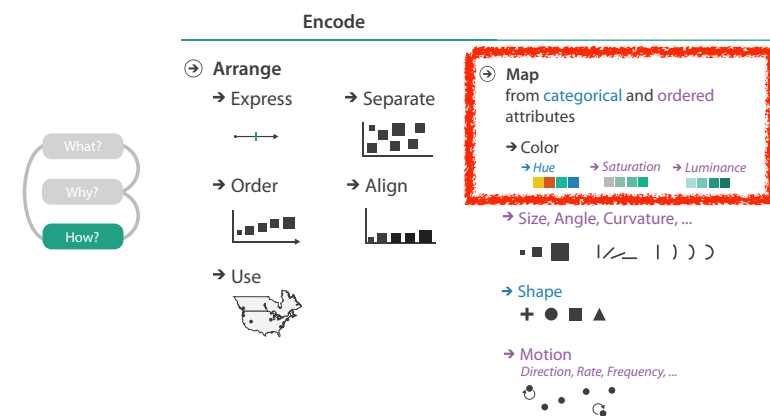
Outline

- **Session 1: Principles 9:30-11:00am**
 - Analysis: What, Why, How
 - Marks and Channels
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks and Trees
 - **Map Color and Other Channels**
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

<http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>

65

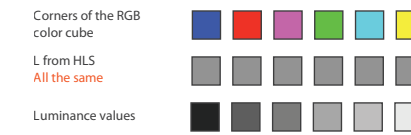
How to encode: Mapping color



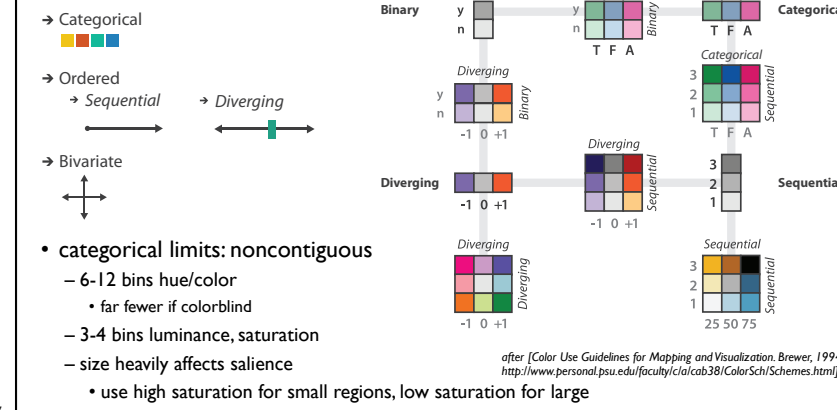
66

Color: Luminance, saturation, hue

- 3 channels
 - what/where for categorical
 - hue
 - how-much for ordered
 - luminance
 - saturation
- other common color spaces
 - RGB: poor choice for visual encoding
 - HSL: better, but beware
 - lightness ≠ luminance
- transparency
 - useful for creating visual layers
 - but cannot combine with luminance or saturation



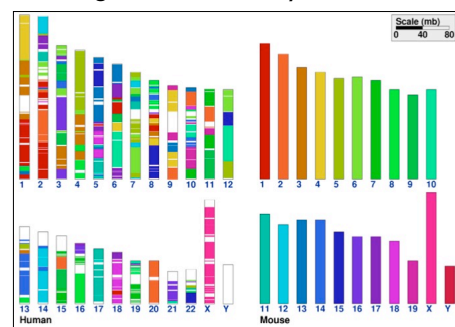
Colormaps



67

Categorical color: Discriminability constraints

- noncontiguous small regions of color: only 6-12 bins

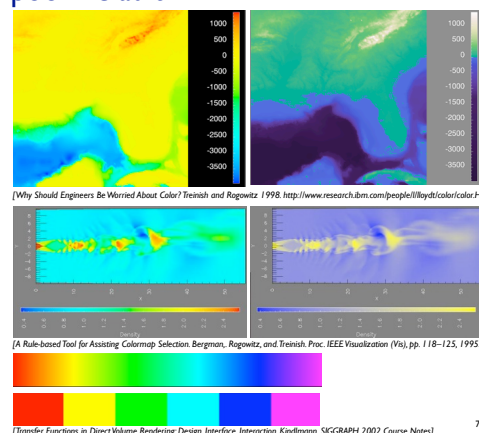


[Cinteny, flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. Sinha and Meller. BMC Bioinformatics, 8:82, 2007.]

69

Ordered color: Rainbow is poor default

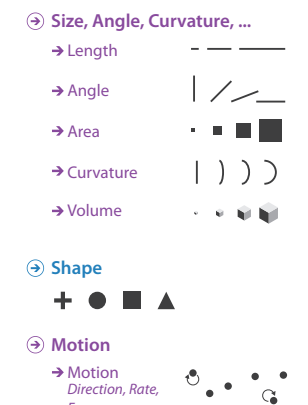
- problems
 - perceptually unordered
 - perceptually nonlinear
- benefits
 - fine-grained structure visible and nameable
- alternatives
 - fewer hues for large-scale structure
 - multiple hues with monotonically increasing luminance for fine-grained
 - segmented rainbows good for categorical, ok for binned



70

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



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. IEEEVis 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.

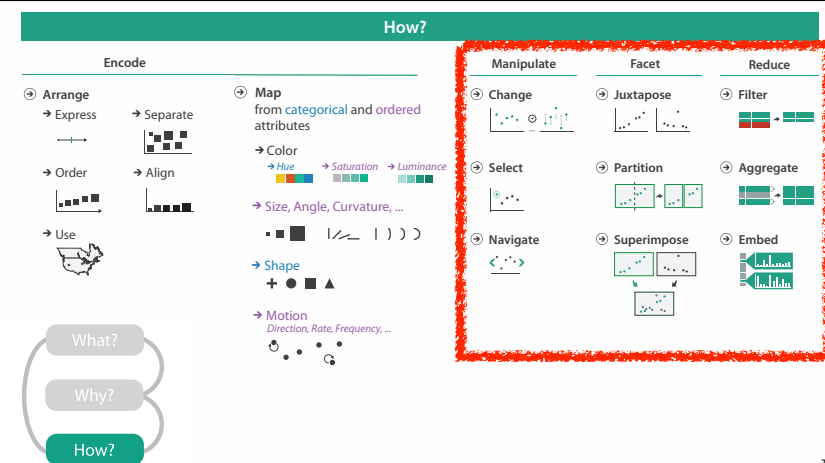
71

Outline

- **Session 1: Principles 9:30-11:00am**
 - Analysis: What, Why, How
 - Marks and Channels
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks and Trees
 - Map Color and Other Channels
 - **Manipulate: Change, Select, Navigate**
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

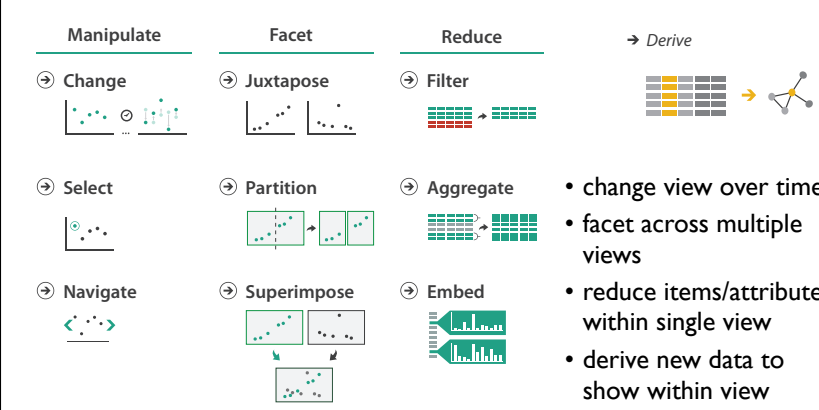
<http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>

73



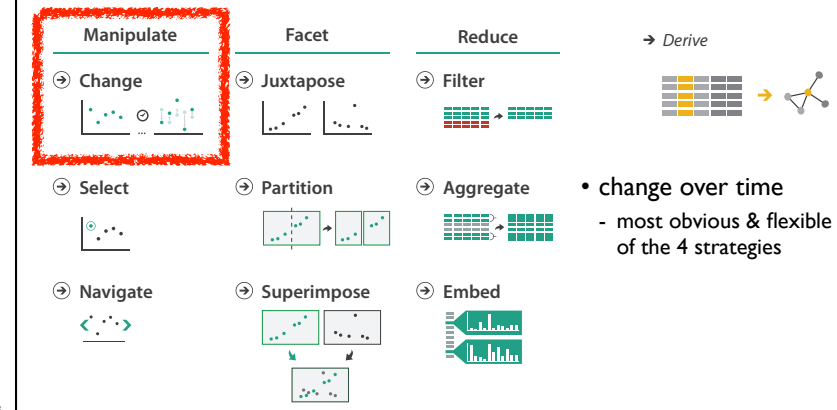
74

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

76

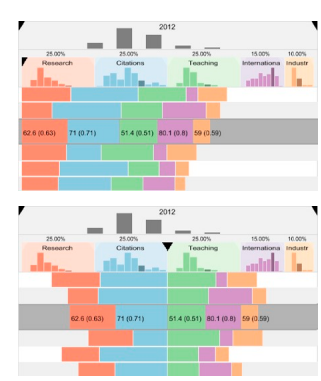
Change over time

- 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

- 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.]

78

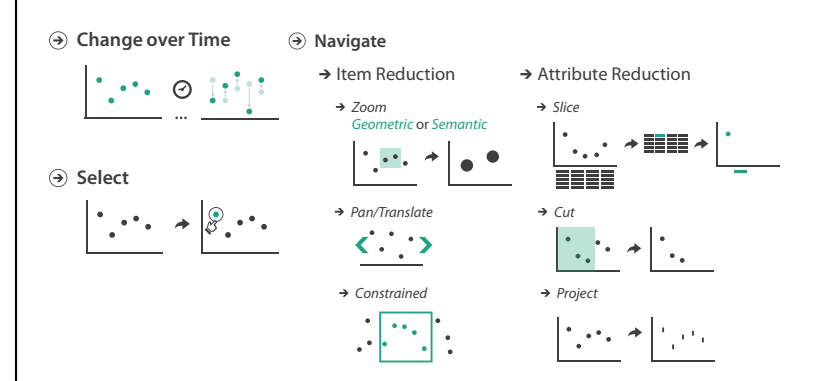
Idiom: Animated transitions

- 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
 - scope of what is shown narrows down
 - middle block stretches to fill space, additional structure appears within
 - other blocks squish down to increasingly aggregated representations



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

Manipulate

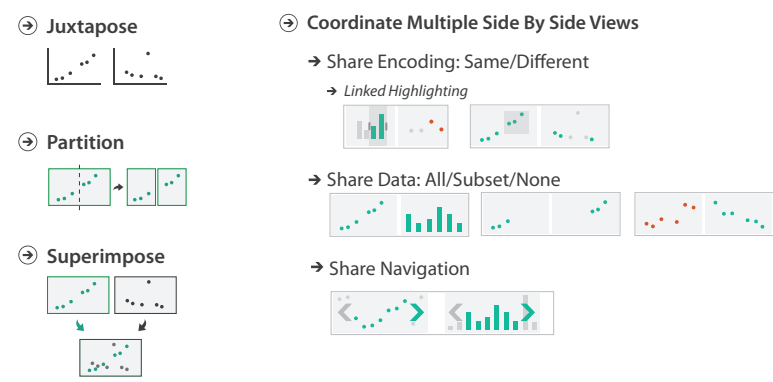


79

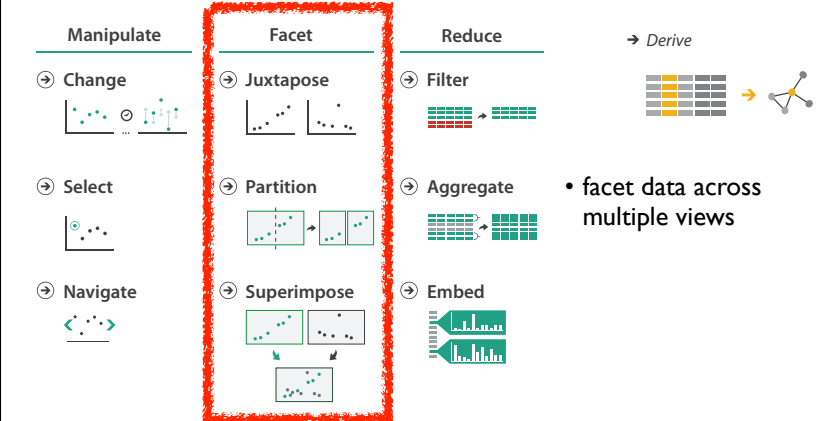
Further reading

- 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.

Facet



How to handle complexity: 3 more strategies + 1 previous



Idiom: Linked highlighting

• 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**

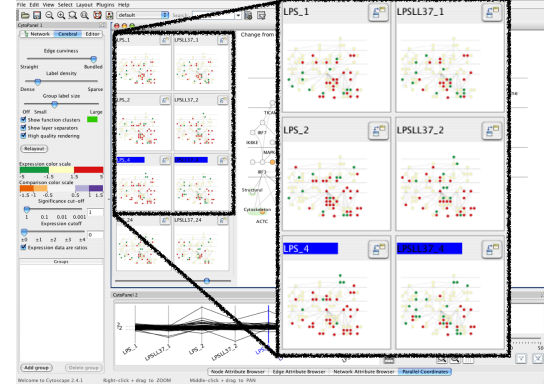


[A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. Cockburn, Karlson, and Bederson. ACM Computing Surveys 41:1 (2008), 1–31.]

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

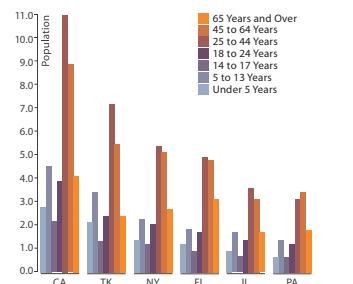
- why juxtapose views?
 - benefits: eyes vs memory
 - lower cognitive load to move eyes between 2 views than remembering previous state with single changing view
 - costs: display area, 2 views side by side each have only half the area of one view

Partition into views

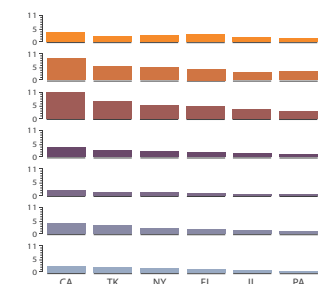
- how to divide data between views
 - encodes association between items using spatial proximity
 - major implications for what patterns are visible
 - split according to attributes
- design choices
 - how many splits
 - all the way down: one mark per region?
 - stop earlier, for more complex structure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

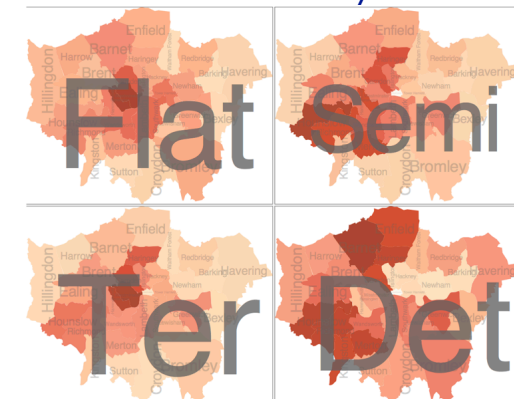


[Configuring Hierarchical Layouts to Address Research Questions. Slingsby, Dykes, and Wood. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2009) 15:6 (2009), 977–984.]

Partitioning: Recursive subdivision

System: HIVE

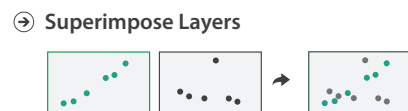
- different encoding for second-level regions
 - choropleth maps



[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 achievable, three with careful design



Static visual layering

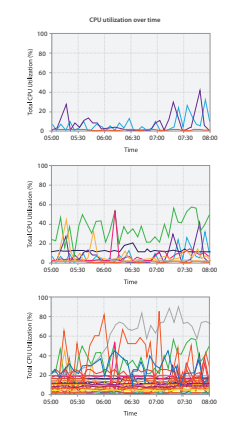
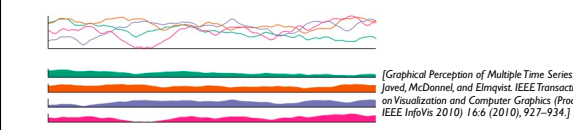
- foreground layer: roads
 - hue, size distinguishing main from minor
 - high luminance contrast from background
- background layer: regions
 - desaturated colors for water, parks, land areas
- user can selectively focus attention
- “get it right in black and white”
 - check luminance contrast with greyscale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

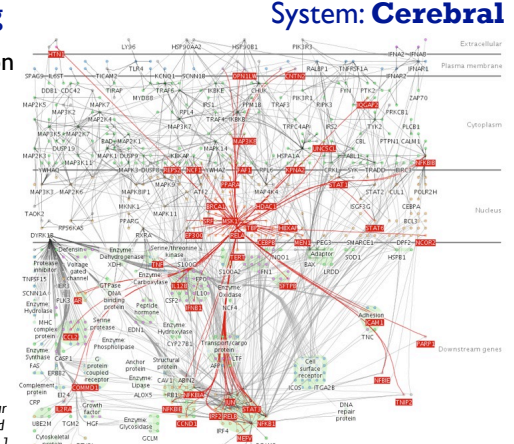
- few layers, but many lines
 - up to a few dozen
 - but not hundreds
- superimpose vs juxtapose: empirical study
 - superimposed for local visual, multiple for global
 - same screen space for all multiples, single superimposed
 - tasks
 - local: maximum, global: slope, discrimination



[Graphical Perception of Multiple Time Series. Javeli, McDonnell, and Elmquist. IEEE Transactions on Visualization and Computer Graphics (Proc. 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.]

Further reading

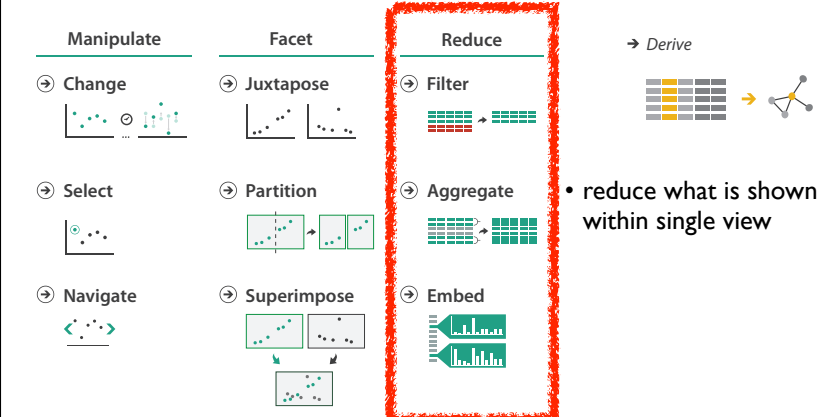
- Visualization Analysis and Design. Munzner. AK Peters Visualization Series, CRC Press, Nov 2014.
 - Chap 12: Facet Into Multiple Views
- A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. Cockburn, Karlson, and Bederson. ACM Computing Surveys 41:1 (2008), 1–31.
- A Guide to Visual 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 9:30-11:00am**
 - Analysis: What, Why, How
 - Marks and Channels
 - Arrange Tables
 - Arrange Spatial Data
 - Arrange Networks and Trees
 - Map Color and Other Channels
 - Manipulate: Change, Select, Navigate
 - Facet: Juxtapose, Partition, Superimpose
 - Reduce: Filter, Aggregate

<http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>

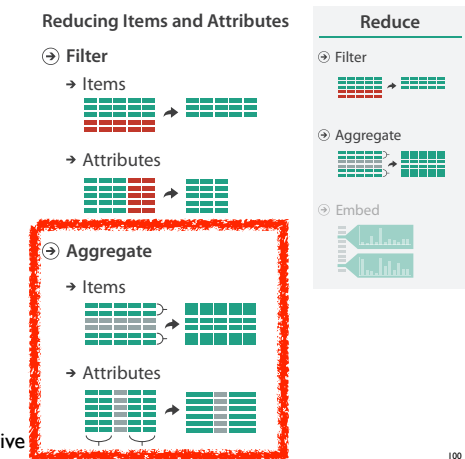
How to handle complexity: 3 more strategies + 1 previous



• reduce what is shown within single view

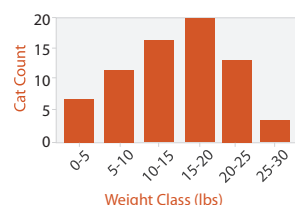
Reduce items and attributes

- 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



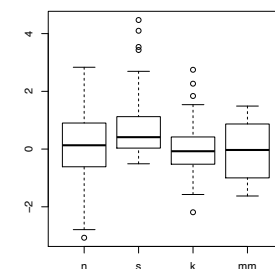
Idiom: histogram

- 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

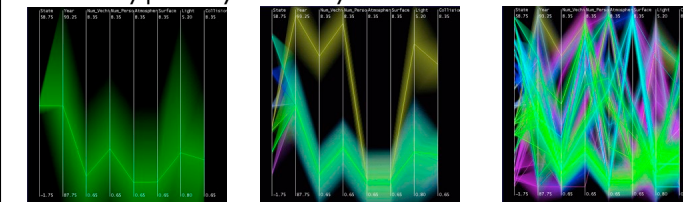
- static item aggregation
- task: find distribution
- data: table
- derived data
 - 5 quant attribs
 - median: central line
 - lower and upper quartile: boxes
 - lower upper fences: whiskers
 - values beyond which items are outliers
 - outliers beyond fence cutoffs explicitly shown



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

Idiom: Hierarchical parallel coordinates

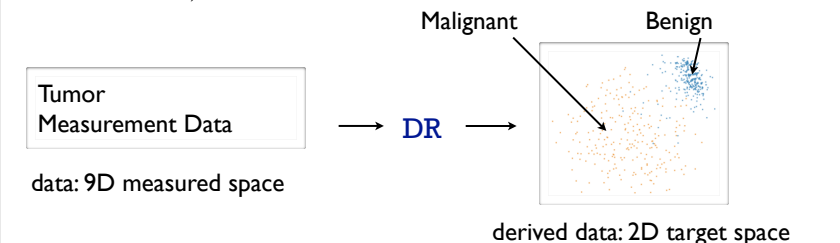
- dynamic item aggregation
- derived data: **hierarchical clustering**
- encoding:
 - cluster band with variable transparency, line at mean, width by min/max values
 - color by proximity in hierarchy



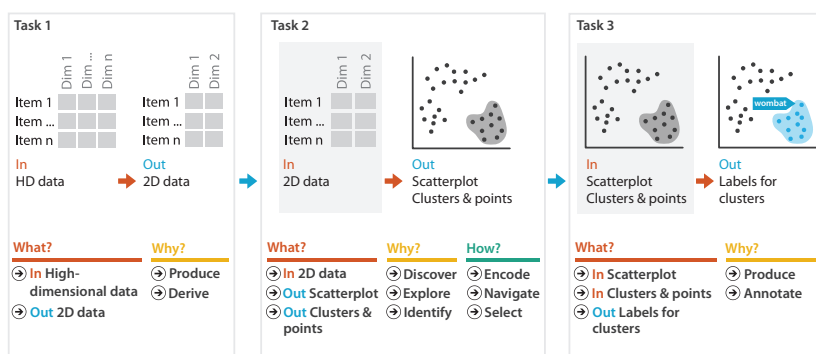
[Hierarchical Parallel Coordinates for Exploration of Large Datasets. Fua, Ward, and Rundensteiner. Proc. IEEE Visualization Conference (Vis '99), pp. 43–50, 1999.]

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

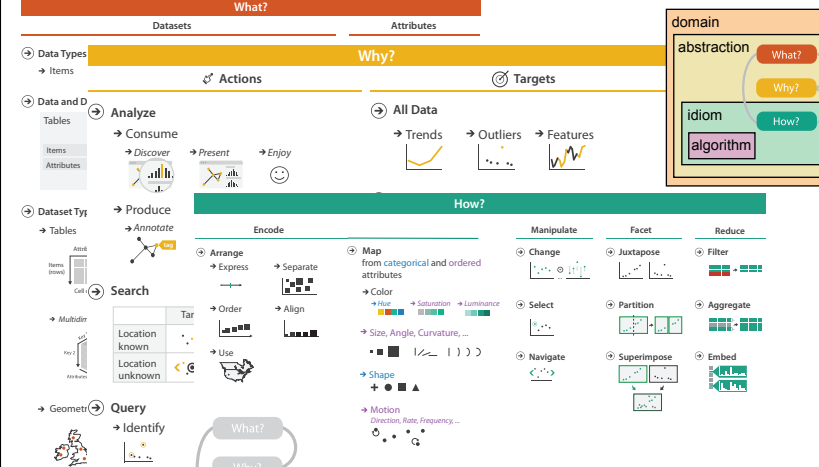


Idiom: Dimensionality reduction for documents



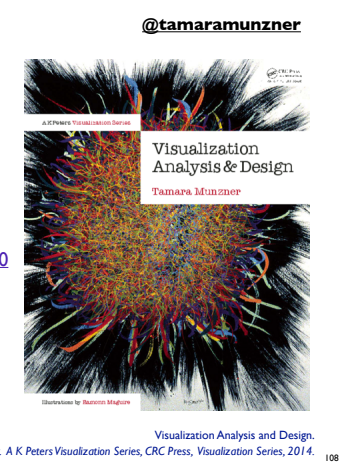
Further reading

- 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.



More Information

- this talk
 - <http://www.cs.ubc.ca/~tmm/talks.html#vad15sydney>
- 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.