

Lecture 8

Tables

DTS204TC Data Visualisation



Outline

- Keys
- Axis orientation
- Dense Layout



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 keys used
 - 0, 1, 2, ...

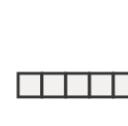
→ 0 Keys

④ Express Values



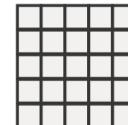
→ 1 Key

List

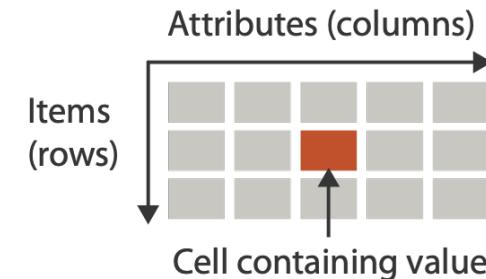


→ 2 Keys

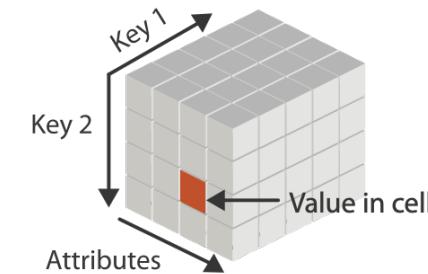
Matrix



→ Tables



→ Multidimensional Table



0 keys: Idiom: scatterplot

⇒ Express Values



- **express** values (magnitudes)

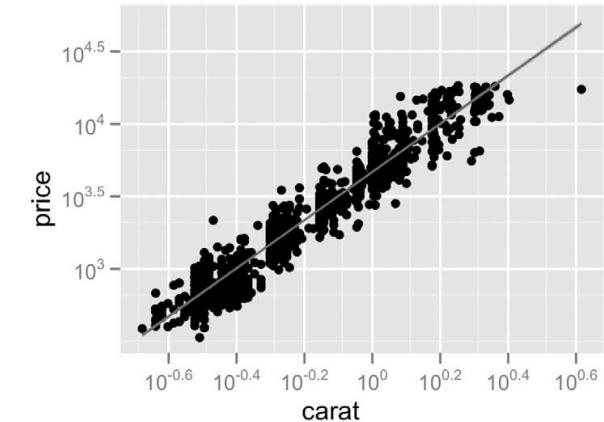
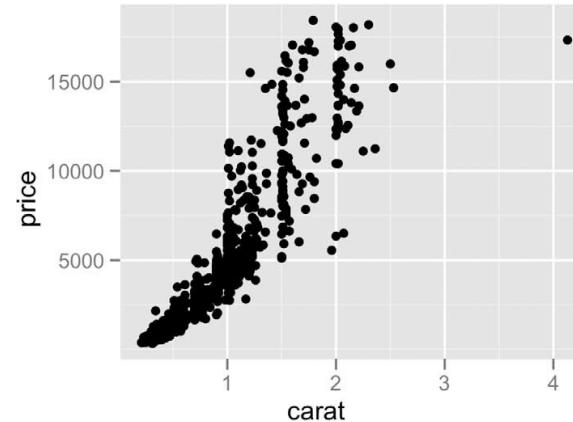
- quantitative attributes

- no keys, only values

- data
 - 2 quant attrs

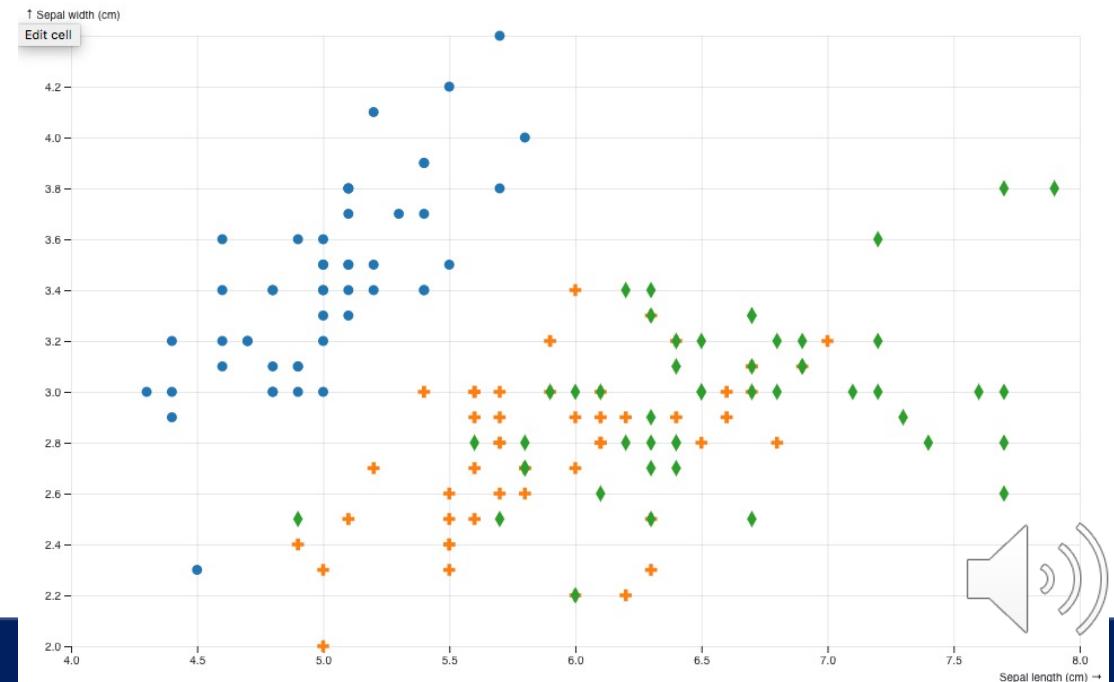
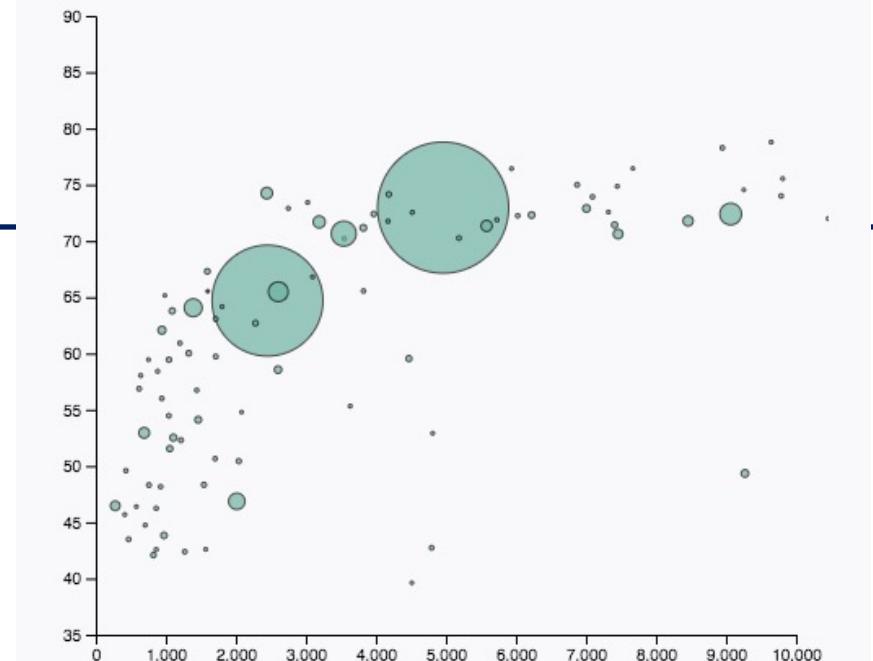
- mark: points
 - channels
 - horiz + vert position

- tasks
 - find trends, outliers, distribution, correlation, clusters
 - scalability
 - hundreds of items



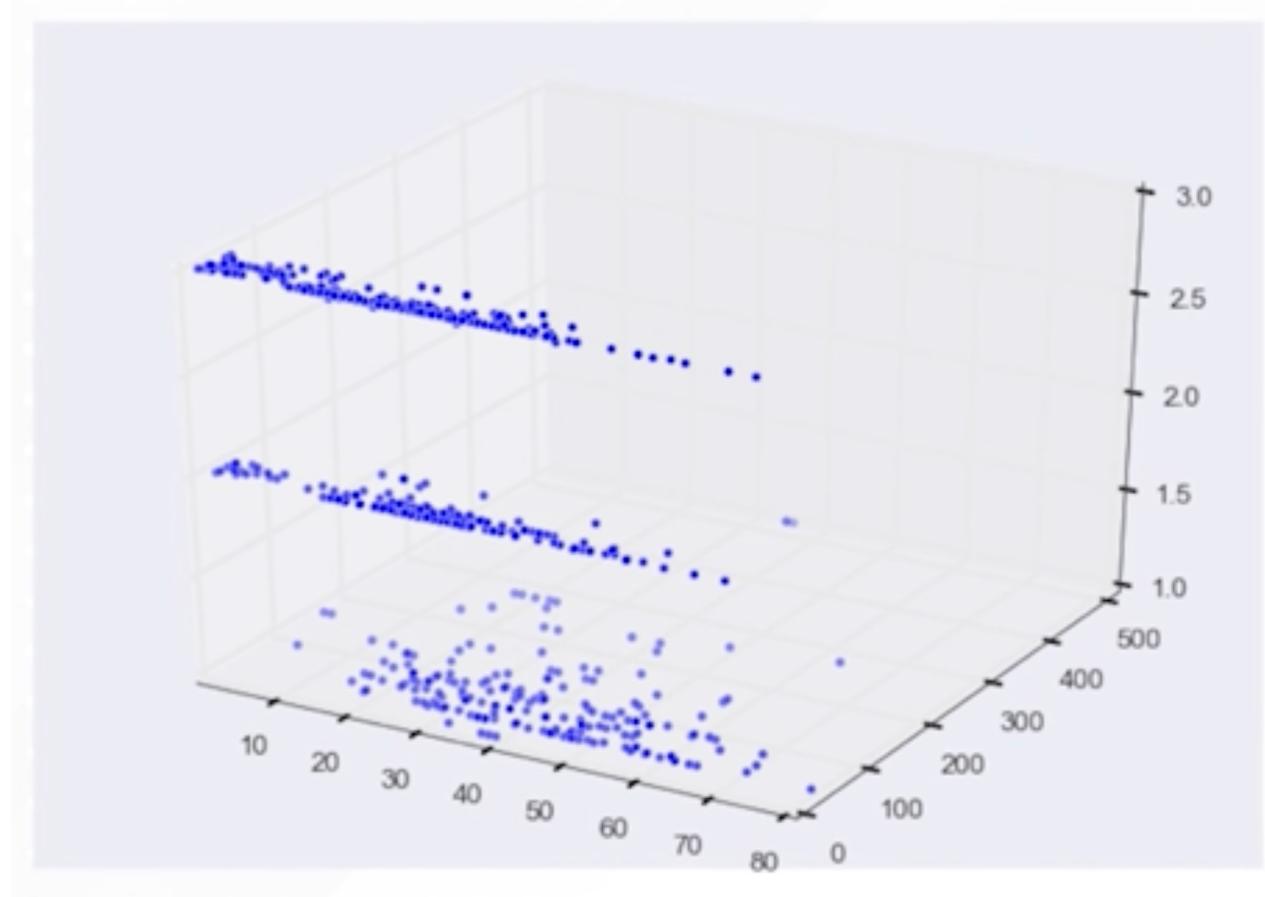
Idiom: scatterplot

- Encoding more channels
 - additional channels viable since using point marks
 - colour
 - size (1 quant attribute, used to control 2D area)
 - note radius would mislead, take square root since area grows quadratically
 - shape



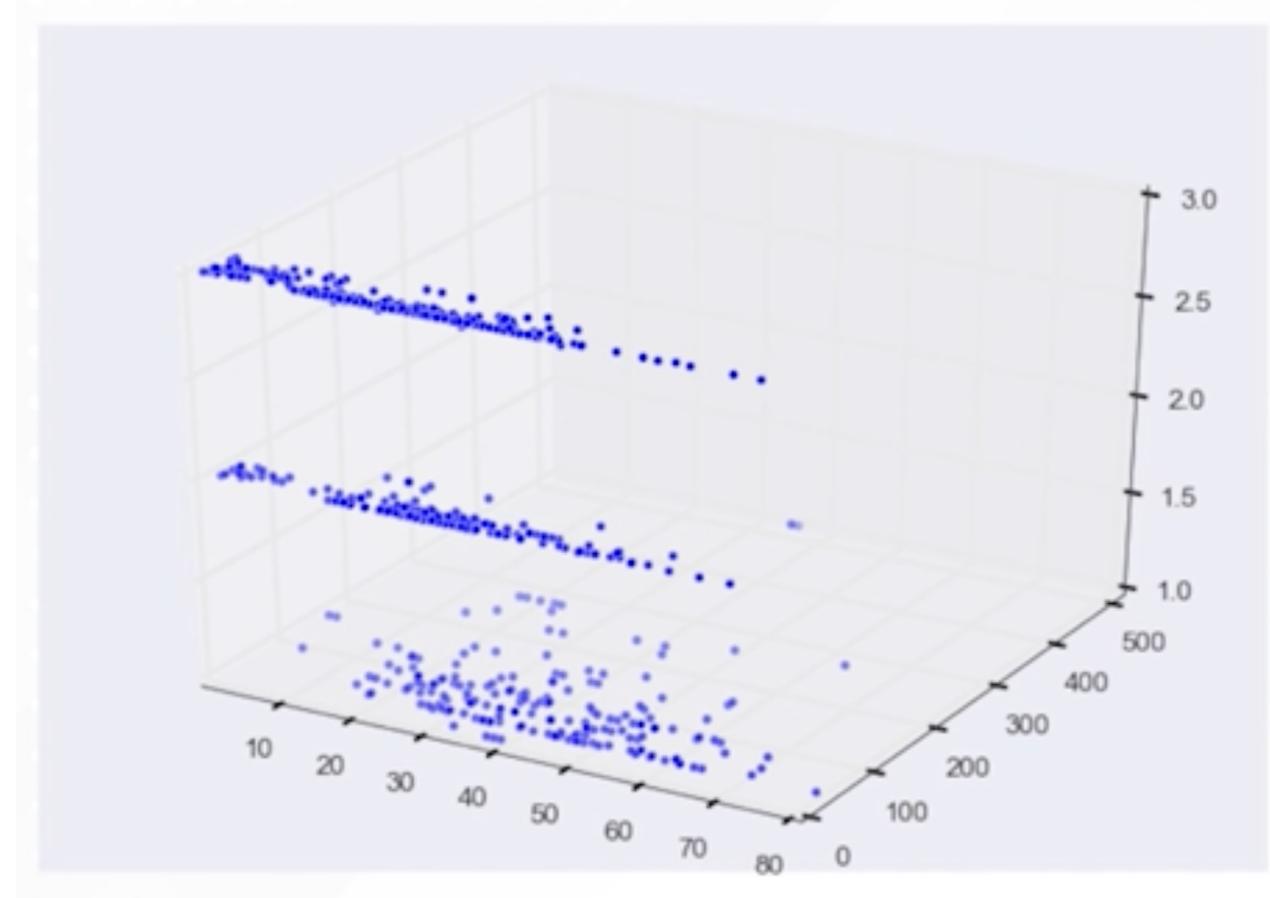
Idiom: scatterplot

- Encoding more channels
 - 3D?



Idiom: scatterplot

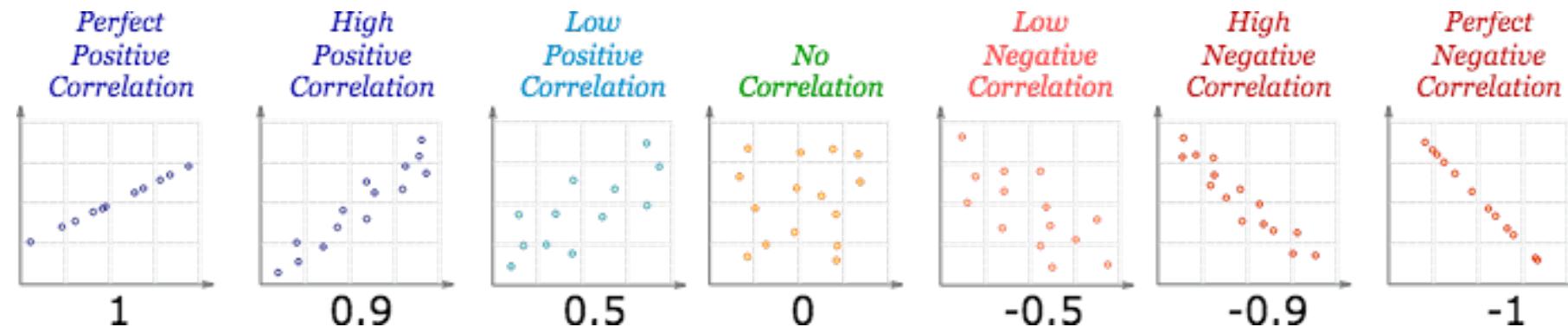
- Encoding more channels
 - 3D? NO!



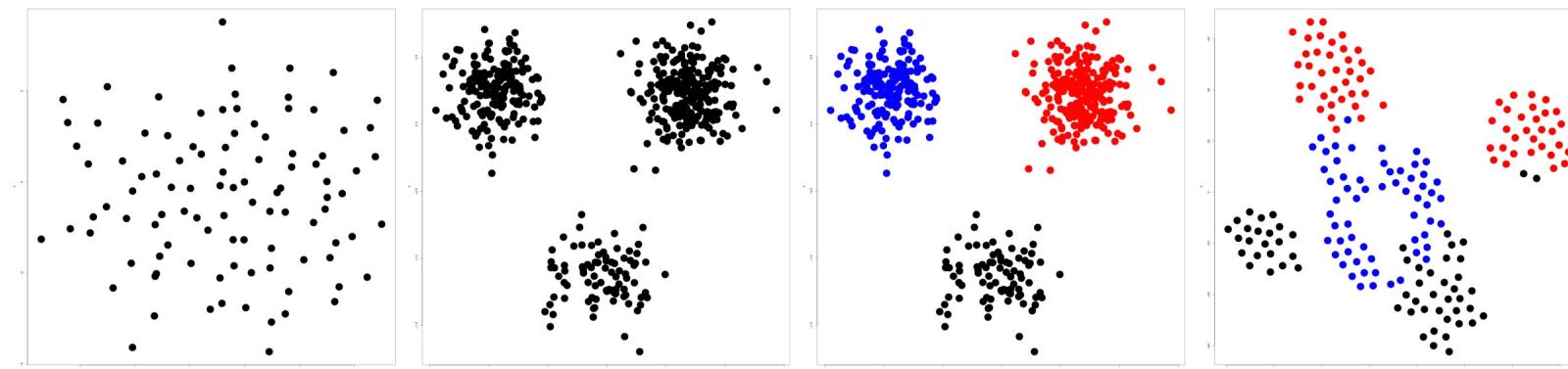
Idiom: scatterplot

- Scatterplot tasks

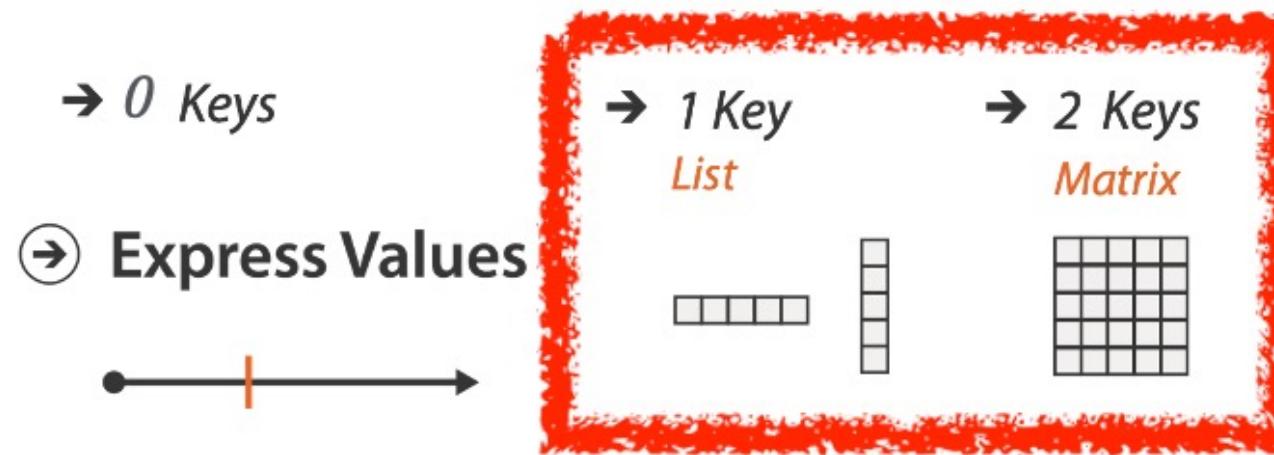
- correlation



- clusters/groups, and clusters vs classes



Some Keys



Some keys: Categorical regions

- Regions: Separate, order, align

→ Separate



→ Order



→ Align

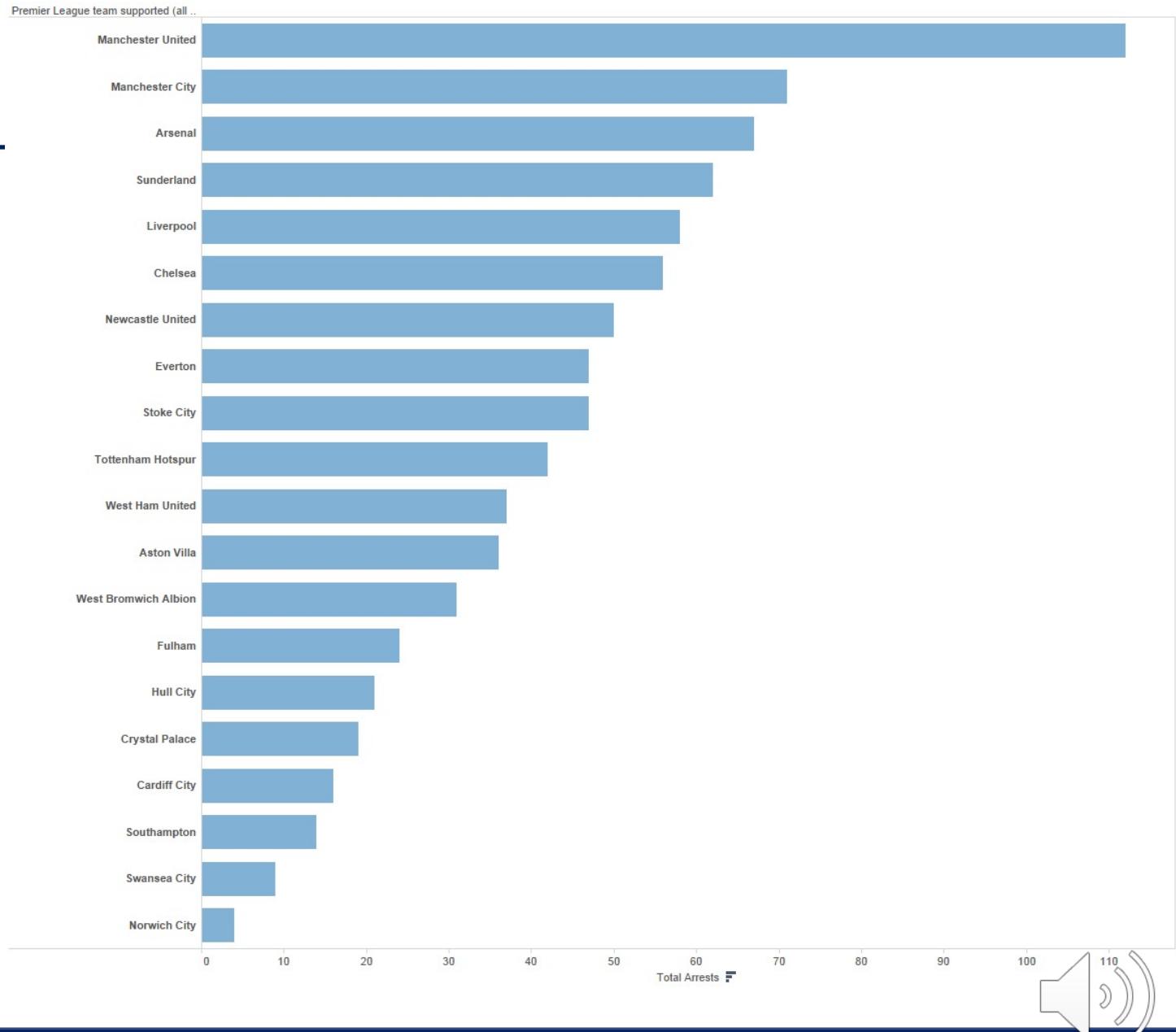


- regions: contiguous bounded areas distinct from each other
 - separate into spatial regions: one mark per region (for now)
- use categorical or ordered attribute to separate into regions
 - no conflict with expressiveness principle for categorical attributes
- use ordered attribute to order and align regions



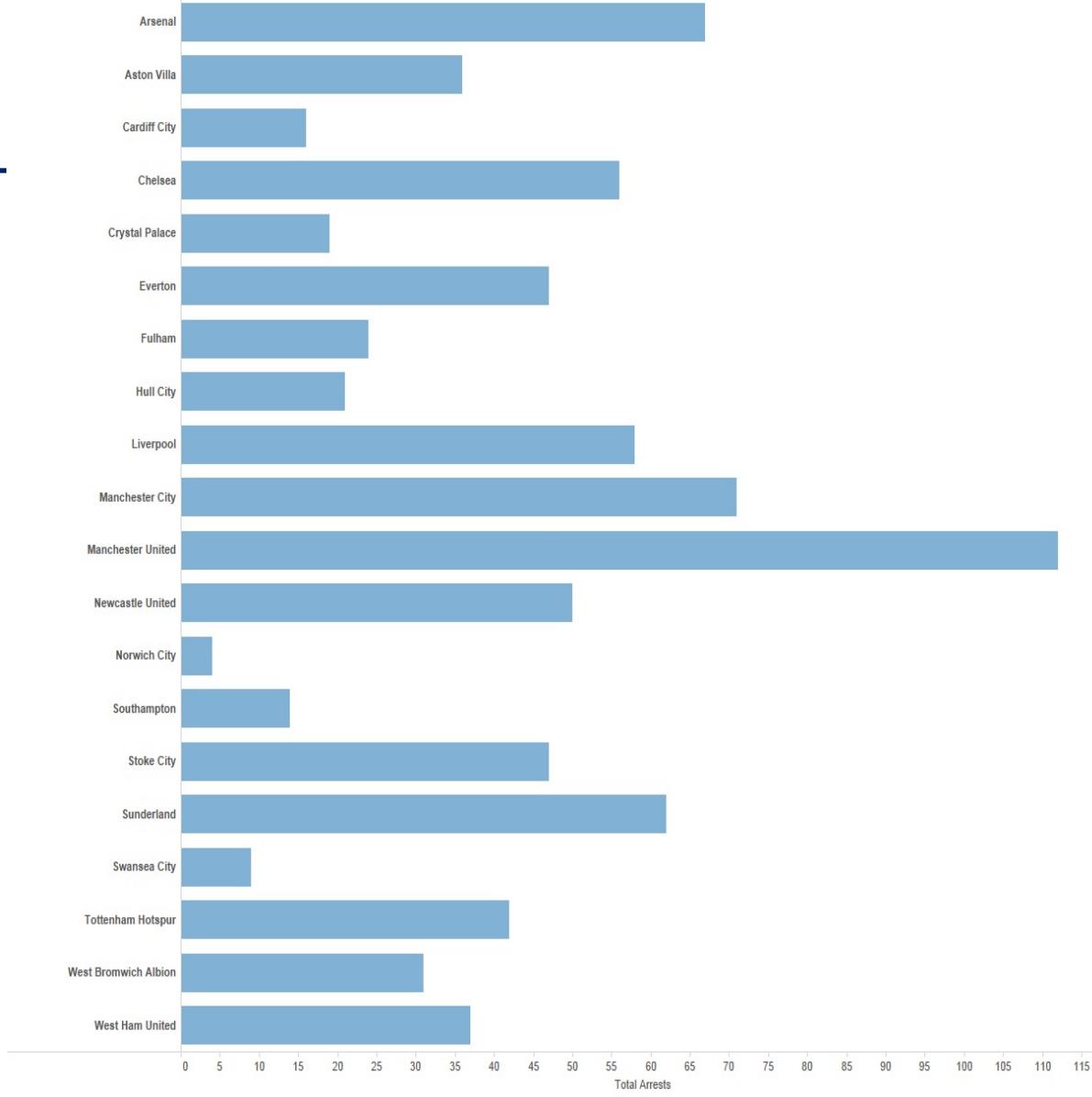
Some keys

- Separated and aligned and ordered
 - best case



Some keys

- Separated and aligned but not ordered
 - limitation: hard to know rank.
what's 4th? what's 7th?



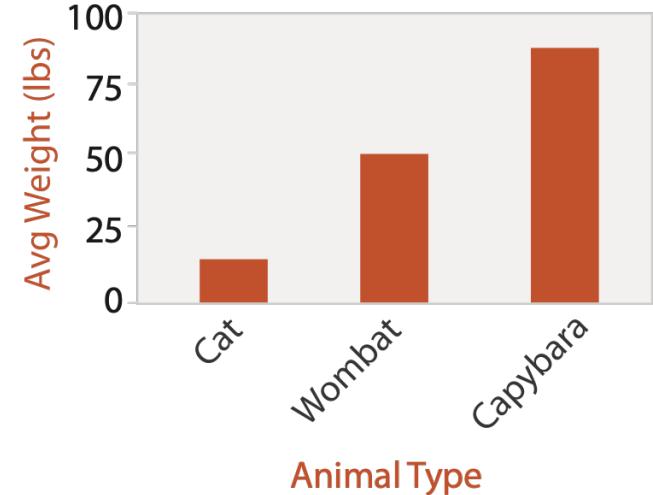
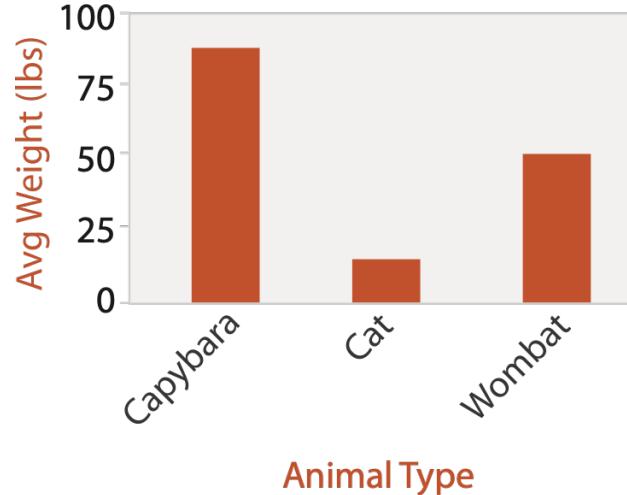
Some keys

- Separated but not aligned or ordered
 - limitation: hard to make comparisons with size (vs aligned position)



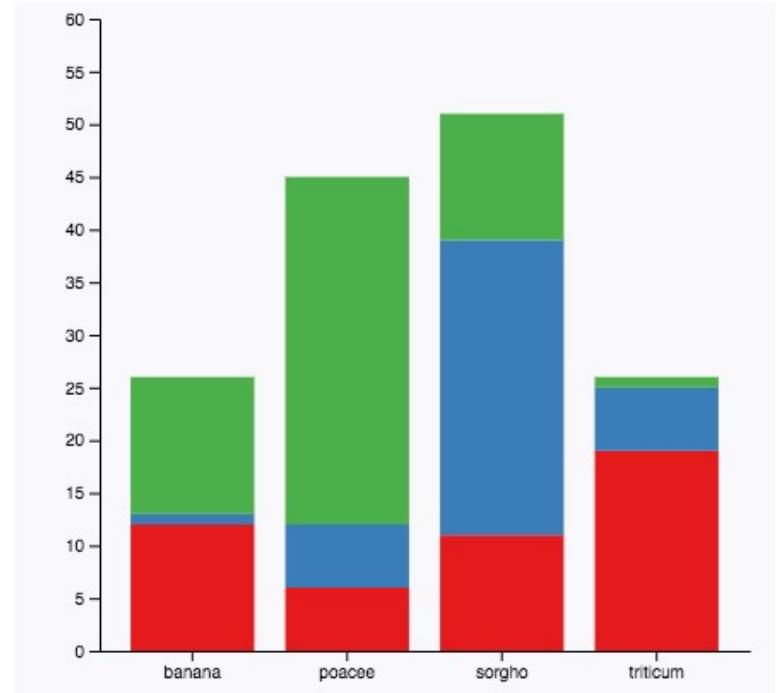
Some keys: Idiom: bar chart

- one key, one value
 - data
 - 1 categorical attribute, 1 quantitative attribute
 - 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 [bars], hundreds for values



Some keys: 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: asymmetric
 - for *stacked* key attrib, 10-12 levels [segments]
 - for *main* key attrib, dozens to hundreds of levels [bars]

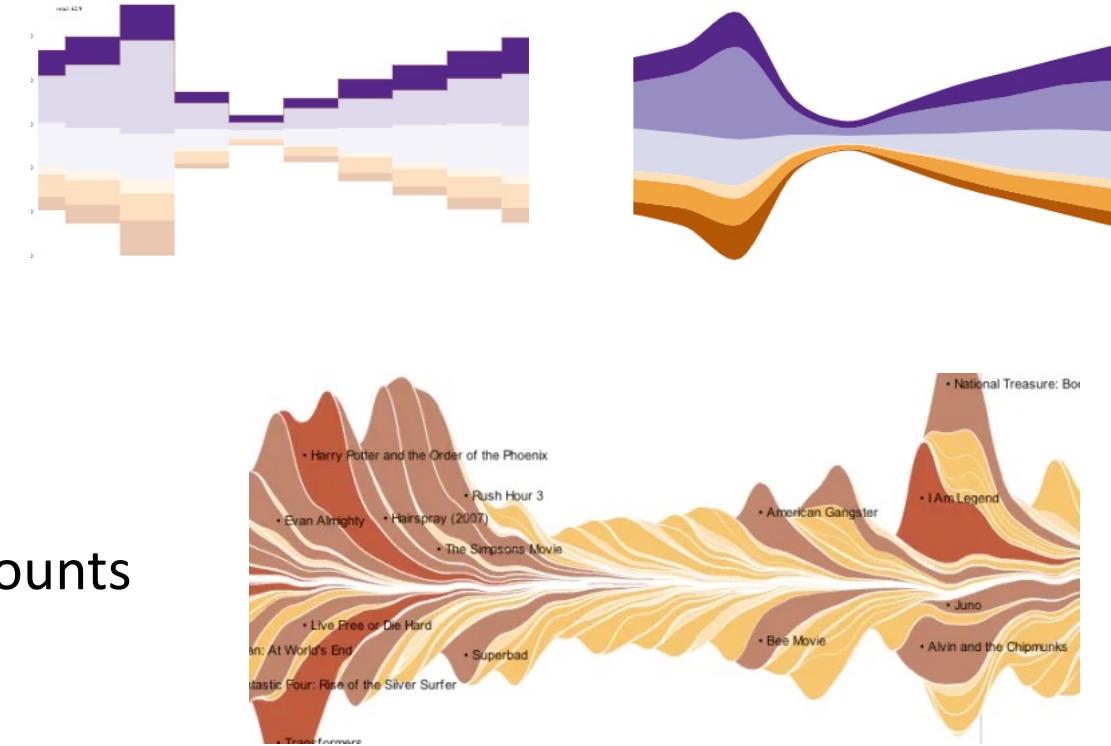


https://www.d3-graph-gallery.com/graph/barplot_stacked_basicWide.html



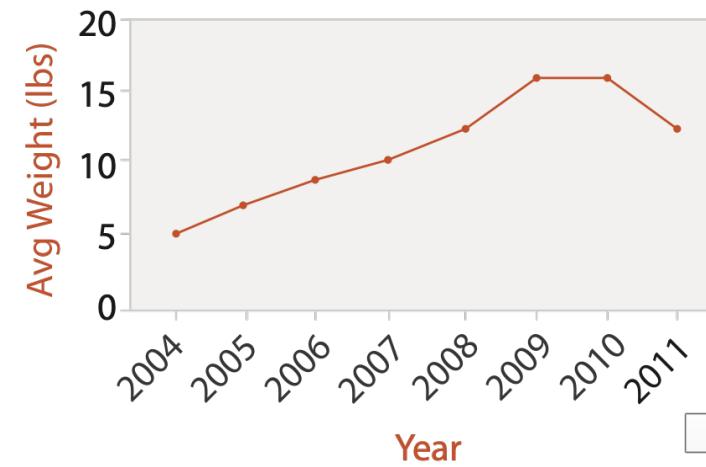
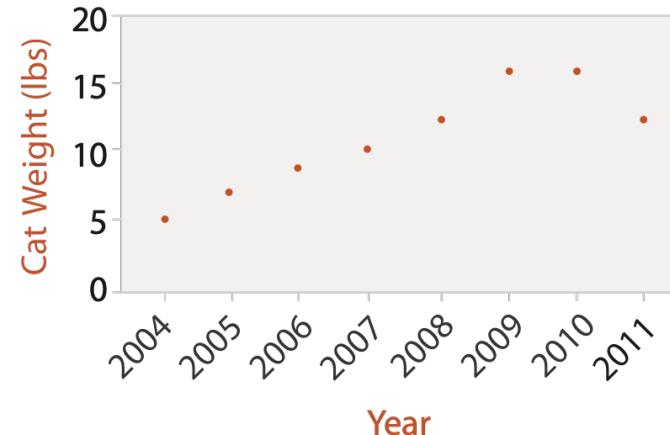
Some keys: Idiom: streamgraph

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



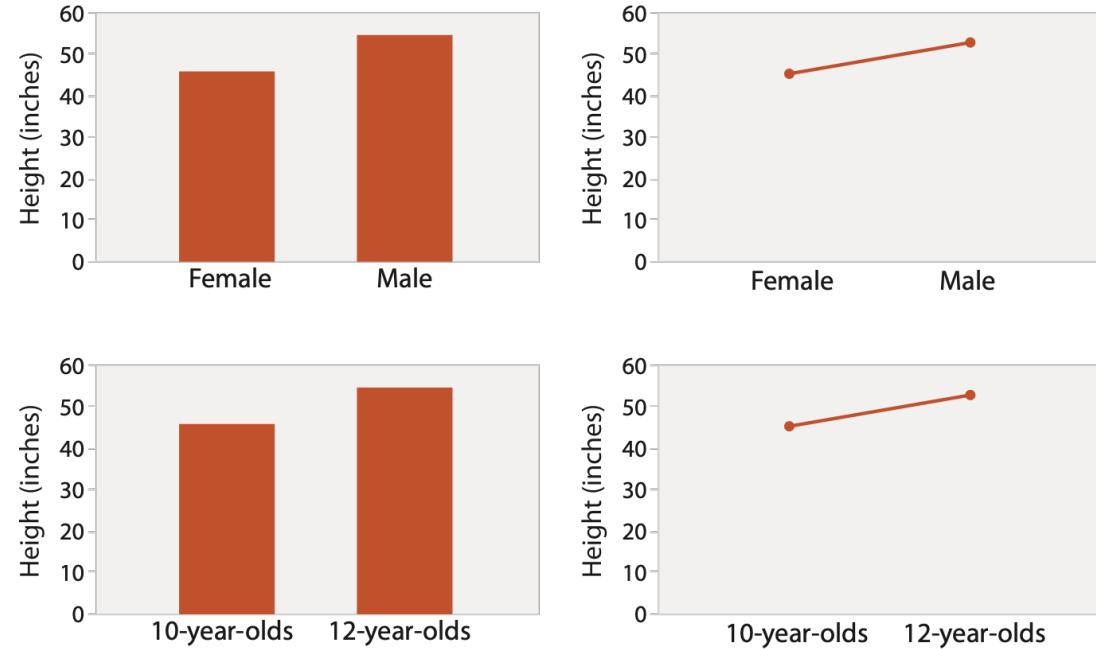
Some keys: Idiom: dot / line chart

- one key, one value
 - data
 - 2 quant attribs
 - mark: points
AND 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



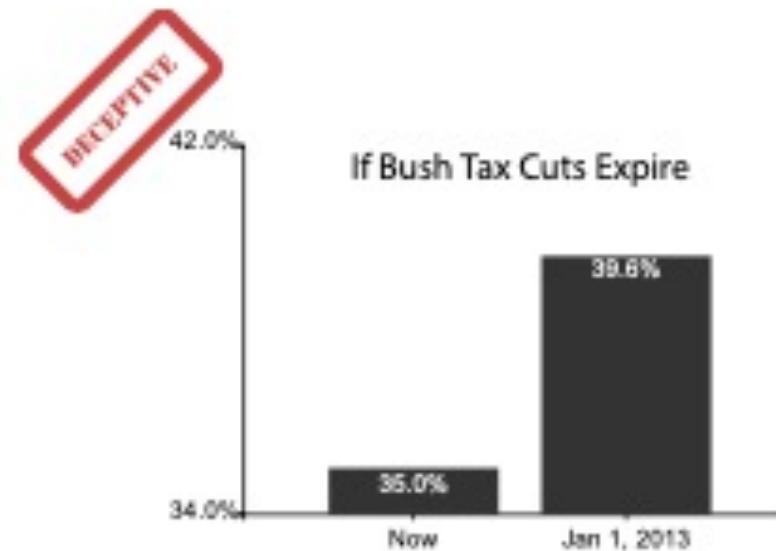
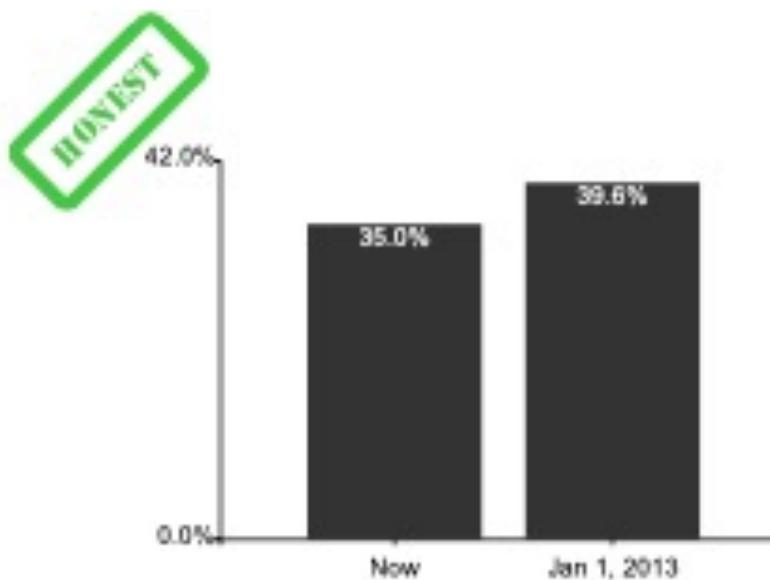
Some keys: Idiom: dot / line chart

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



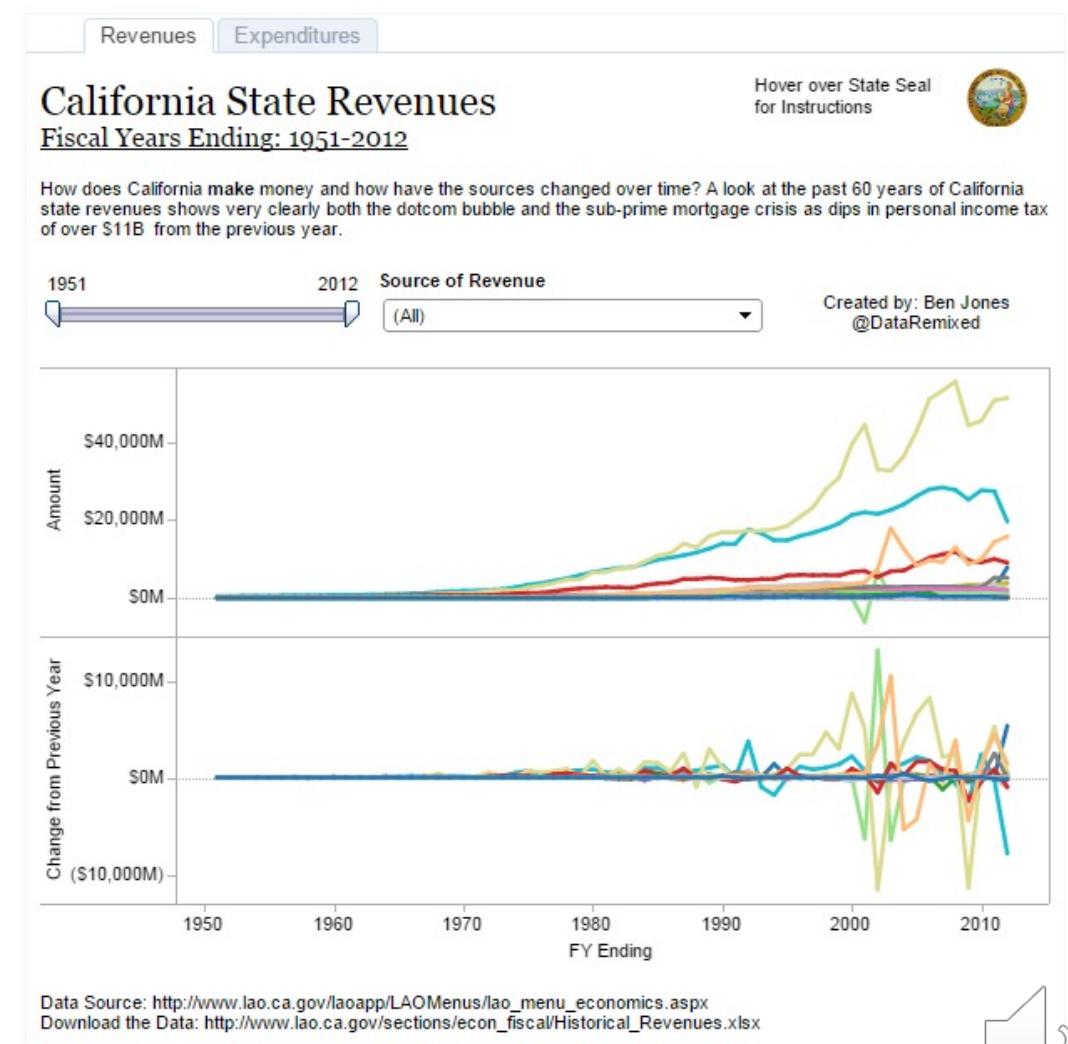
Some keys: Chart Axes

- best practice to label
 - few exceptions: individual small multiple views could share axis label
- avoid cropping y axis
 - some exceptions (arbitrary 0, small change matters)



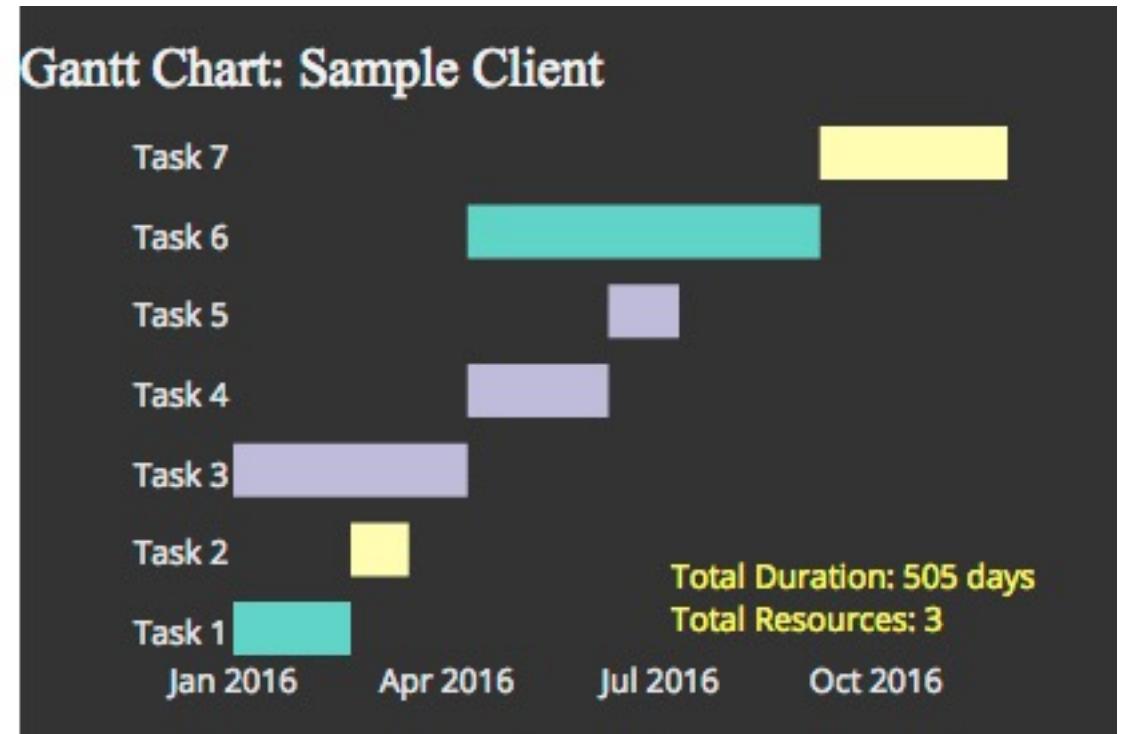
Some keys: Idiom: Indexed line charts

- data: 2 quant attrs
 - 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



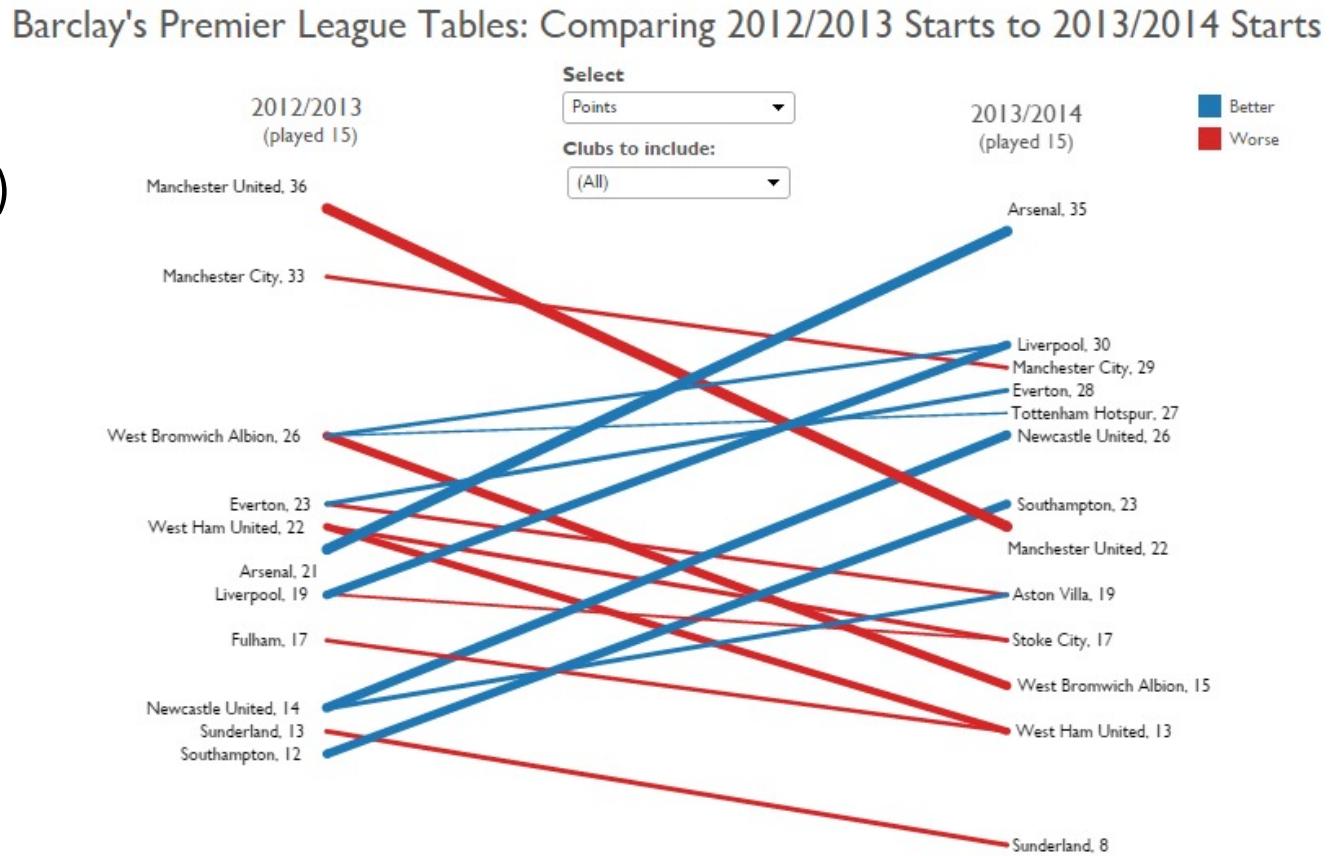
Some keys: Idiom: Gantt charts

- one key, two (related) values
 - data
 - 1 categ attrib, 2 quant attrs
 - mark: line
 - length: duration
 - channels
 - horiz position: start time
(+end from duration)
 - task
 - emphasize temporal overlaps & start/end dependencies between items
 - scalability
 - dozens of key levels [bars]
 - hundreds of value levels [durations]



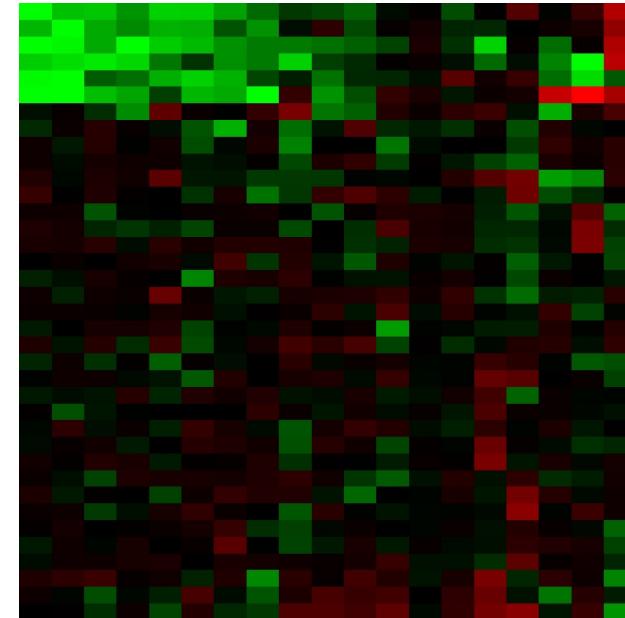
Some keys: Idiom: Slope graphs

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



Some keys: Idiom: heatmap

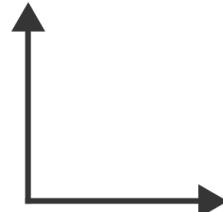
- two keys, one value
 - data
 - 2 categ attrs (gene, experimental condition)
 - 1 quant attrib (expression levels)
 - marks: point
 - separate and align in 2D matrix
 - indexed by 2 categorical attributes
 - channels
 - color by quant attrib
 - (ordered diverging colormap)
 - task
 - find clusters,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



Axis Orientation

→ Axis Orientation

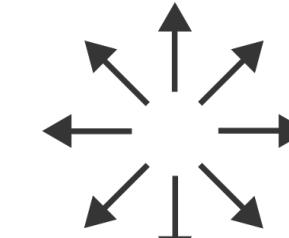
→ Rectilinear



→ Parallel

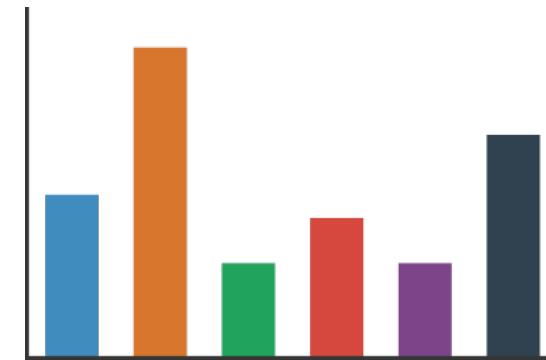
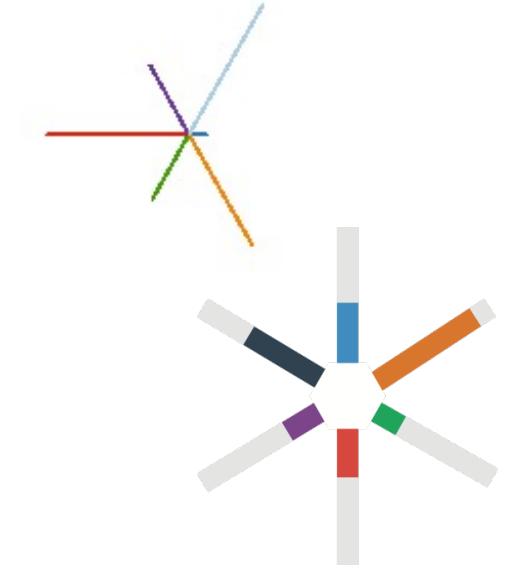


→ Radial



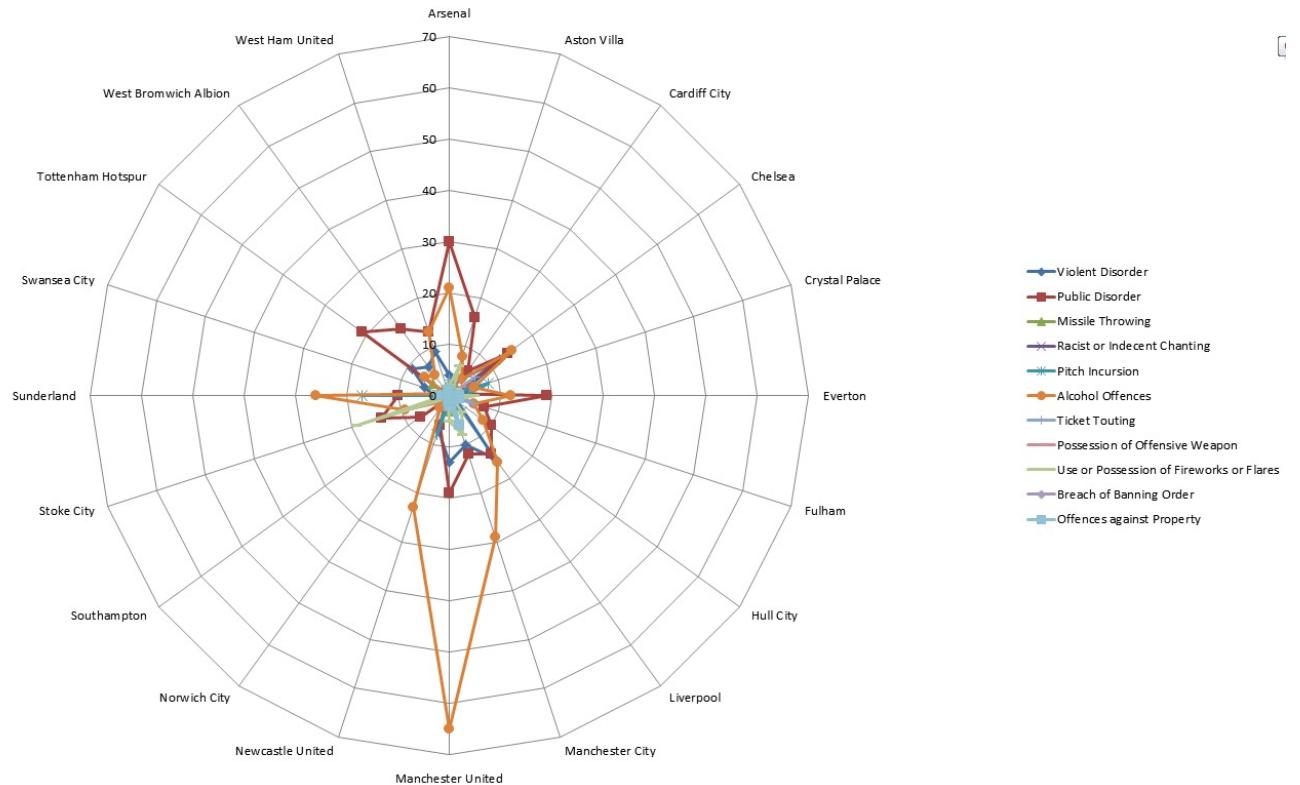
Axis Orientation: Idioms: **radial bar chart, star plot**

- star plot
 - line mark, radial axes meet at central point
- radial bar chart
 - line mark, radial axes meet at central ring
 - channels: length, angle/orientation
- bar chart
 - rectilinear axes, aligned vertically
- accuracy
 - length not aligned with radial layouts
 - less accurately perceived than rectilinear aligned

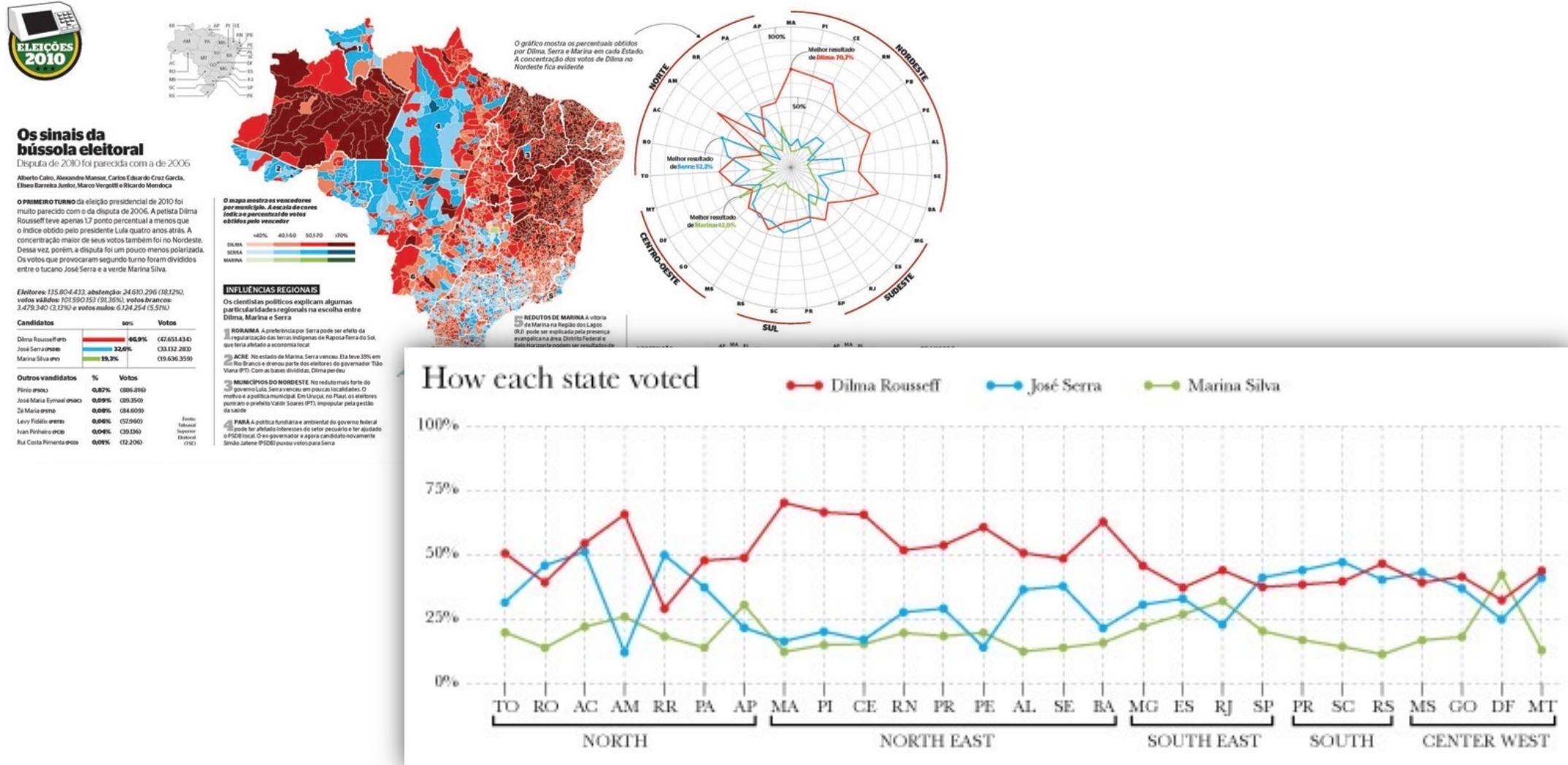


Axis Orientation: Idiom: radar plot

- radial line chart
 - point marks, radial layout
 - connecting line marks
- avoid unless data is cyclic

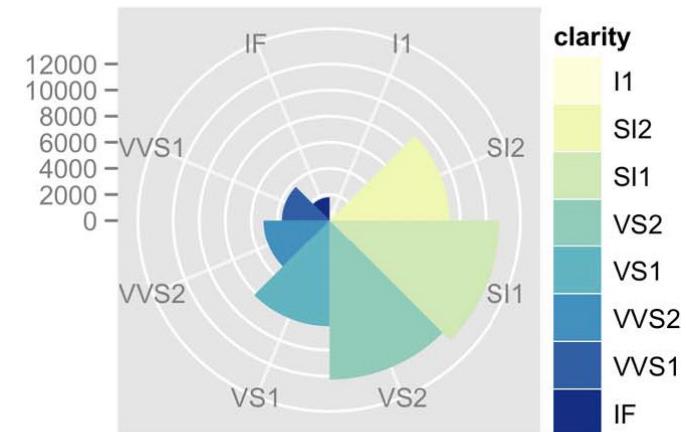
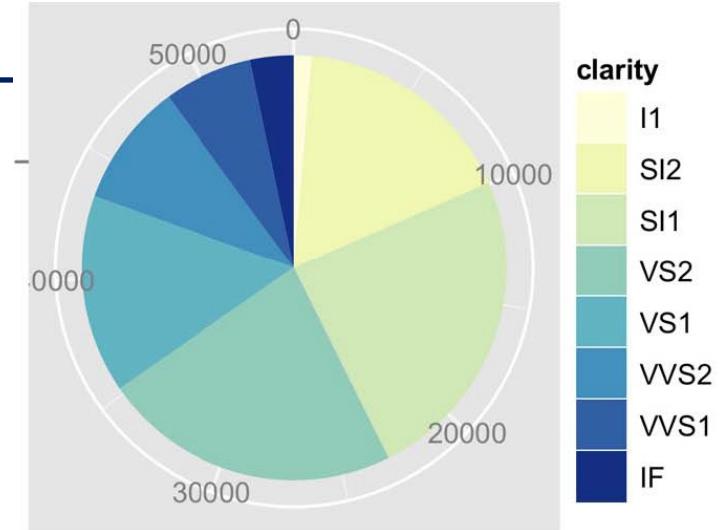


Axis Orientation: Idiom: radar plot



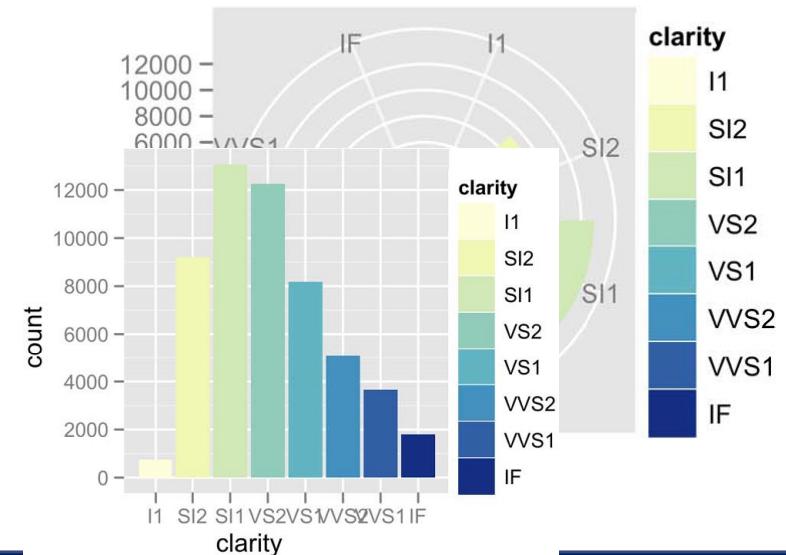
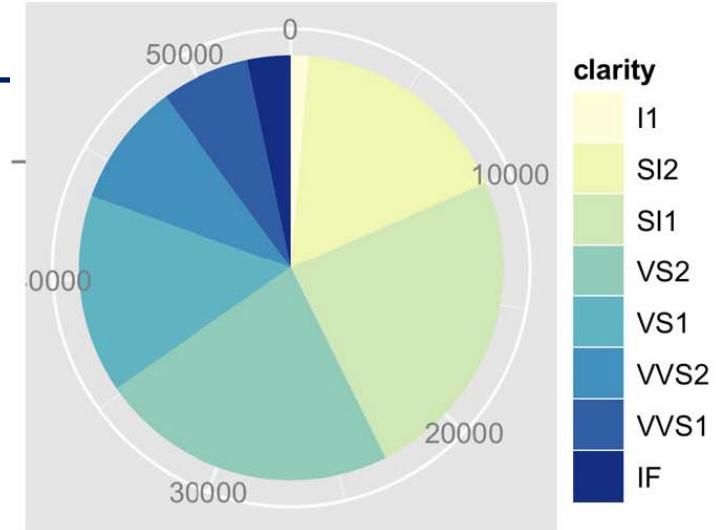
Axis Orientation: Idioms: pie chart, coxcomb chart

- pie chart
 - **interlocking area marks with angle channel: 2D area varies**
 - separated & ordered radially, uniform height
 - accuracy: area less accurate than rectilinear aligned line length
 - **task: part-to-whole judgements**
- coxcomb chart
 - line marks with length channel: **1D length varies**
 - separated & ordered radially, uniform width
 - direct analog to radial bar charts
- data
 - 1 categ key attrib, 1 quant value attrib



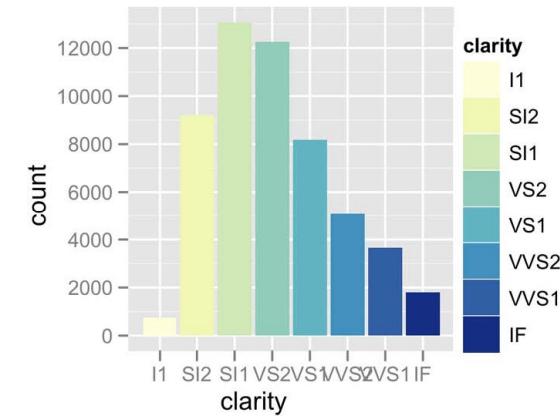
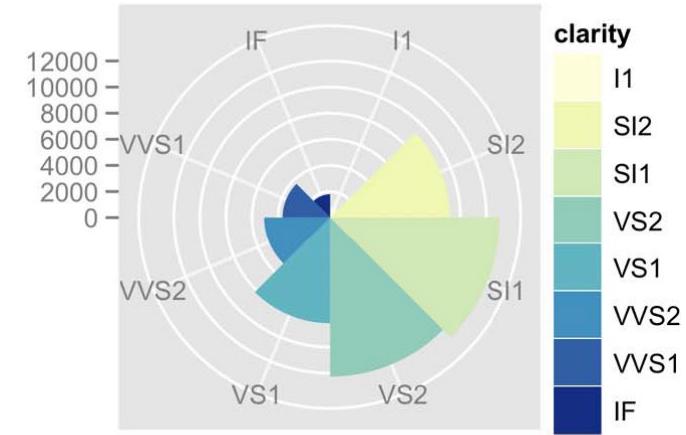
Axis Orientation: Idioms: pie chart, coxcomb chart

- pie chart
 - **interlocking area marks with angle channel: 2D area varies**
 - separated & ordered radially, uniform height
 - accuracy: area less accurate than rectilinear aligned line length
 - **task: part-to-whole judgements**
- coxcomb chart
 - line marks with length channel: **1D length varies**
 - separated & ordered radially, uniform width
 - direct analog to radial bar charts
- data
 - 1 categ key attrib, 1 quant value attrib



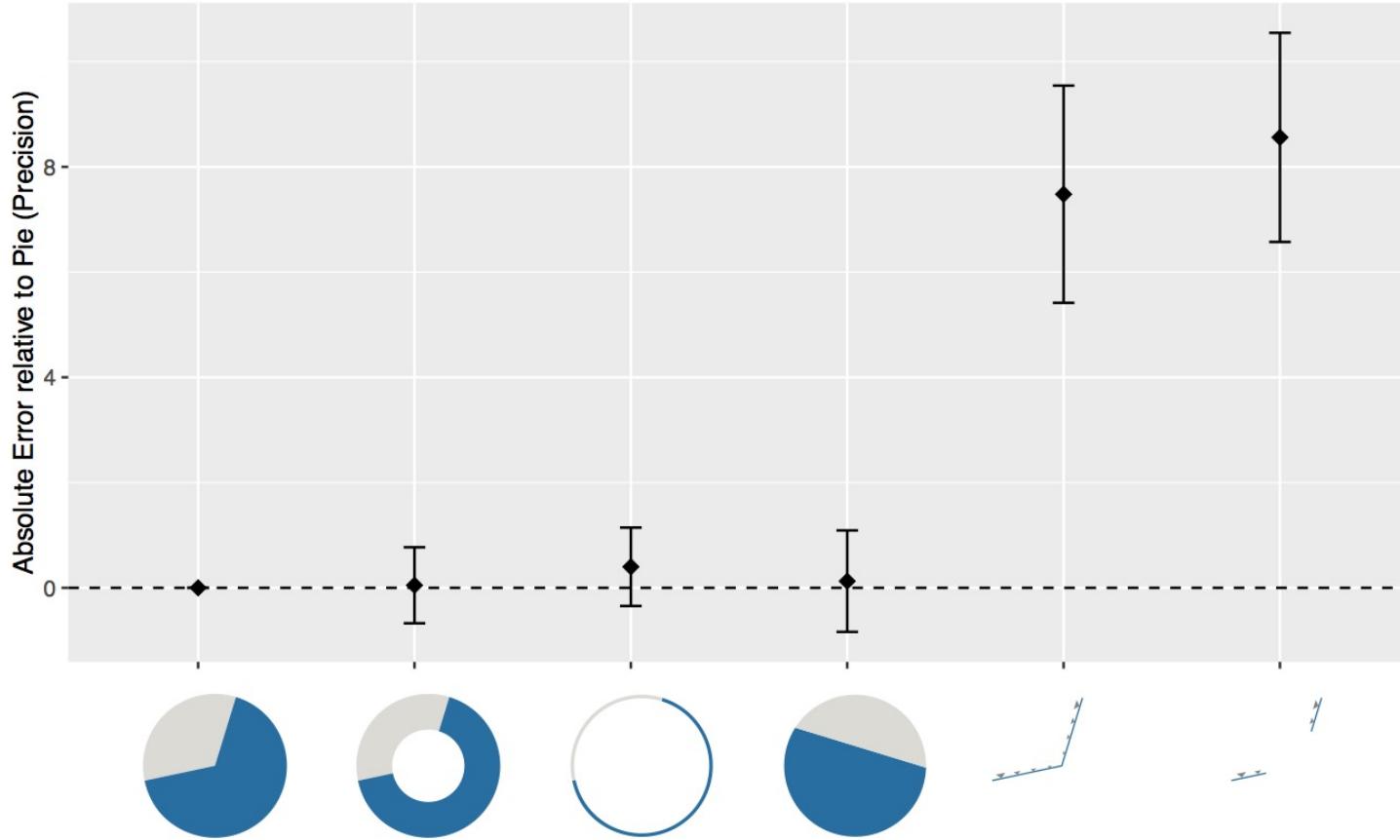
Axis Orientation: Idioms: coxcomb chart

- encode: **1D length**
- decode/perceive: **2D area**
- nonuniform line/sector width as length increases
 - so area variation is nonlinear wrt line mark length!
- bar chart safer: uniform width, so area is linear with line mark length
 - both radial & rectilinear cases



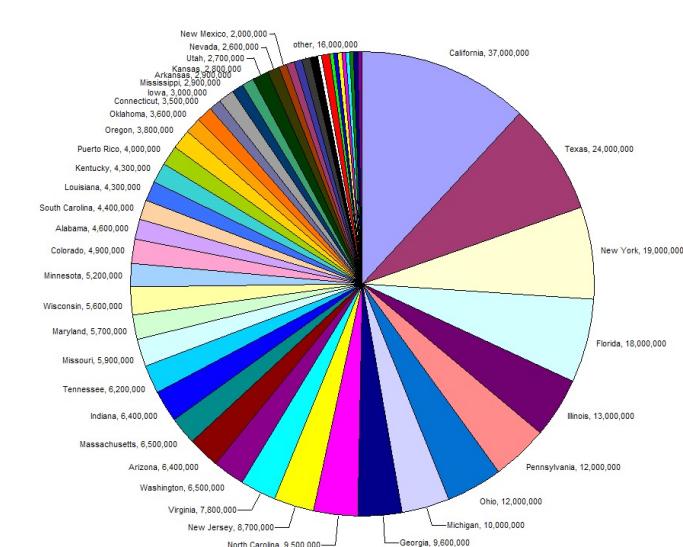
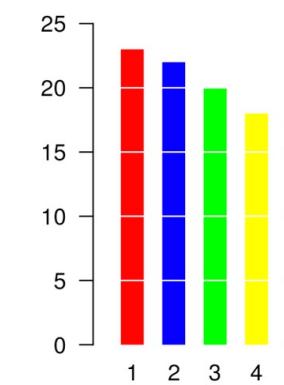
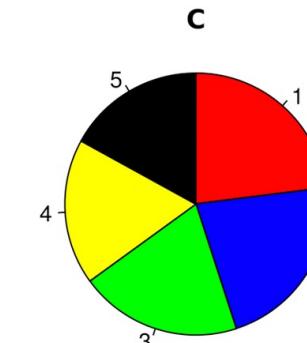
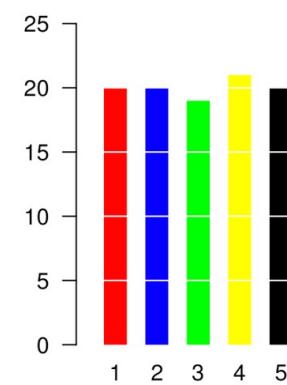
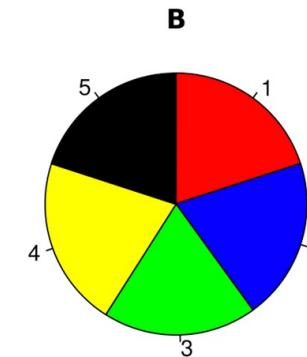
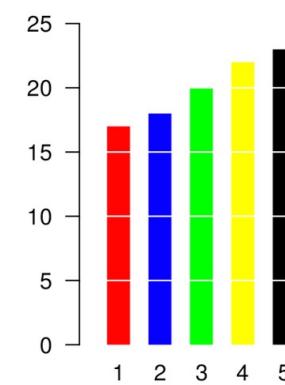
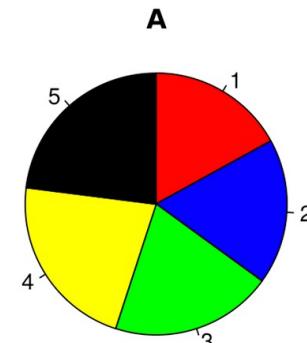
Axis Orientation: Idioms: Pie chart

- some empirical evidence that people respond to arc length
 - decode/perceive: not angles
 - maybe also areas?...
- donut charts no worse than pie charts



Axis Orientation: Idioms: Pie chart

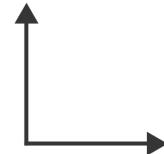
- not so bad for two (or few) levels, for part-to-whole task
- dubious for several levels if details matter
- terrible for many levels



Axis Orientation

④ Axis Orientation

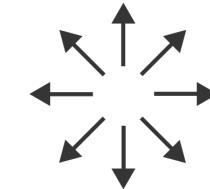
→ Rectilinear



→ Parallel

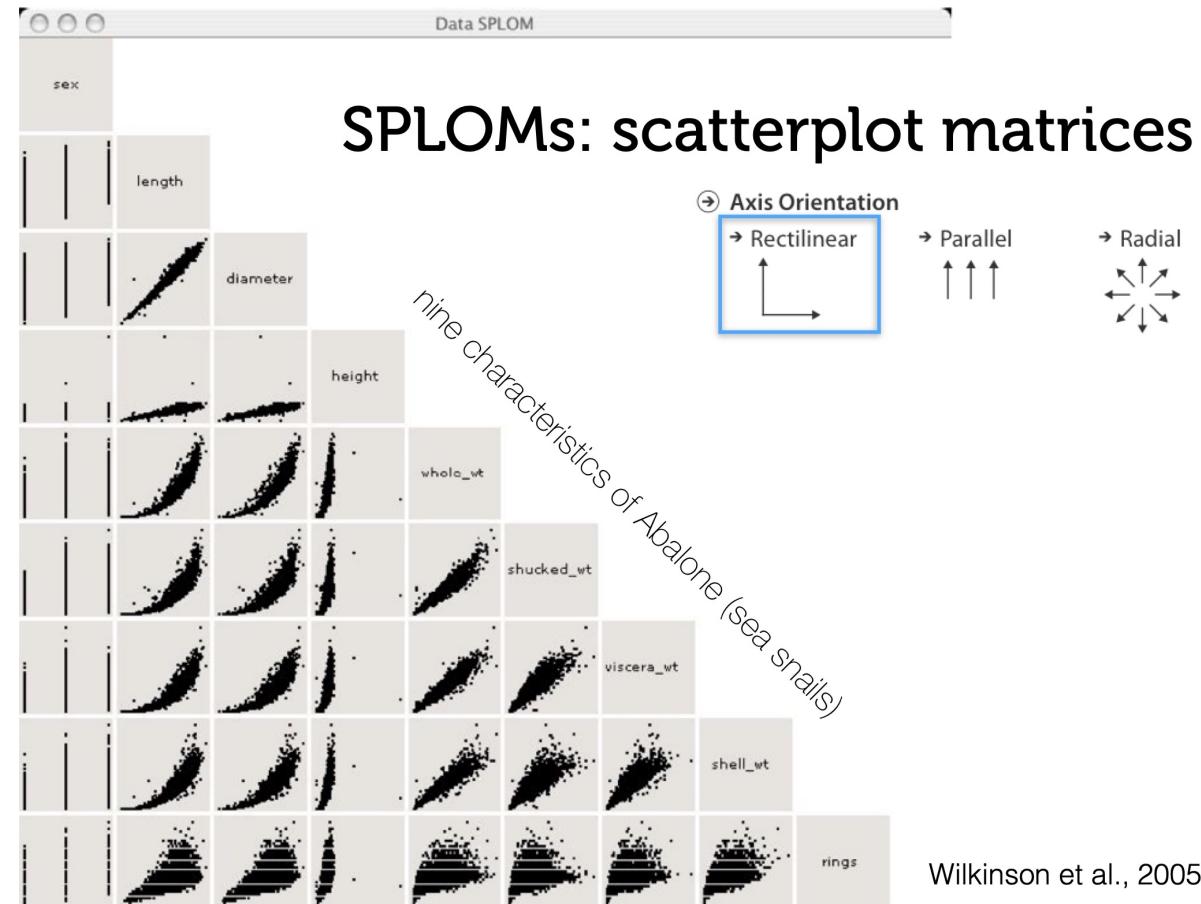


→ Radial



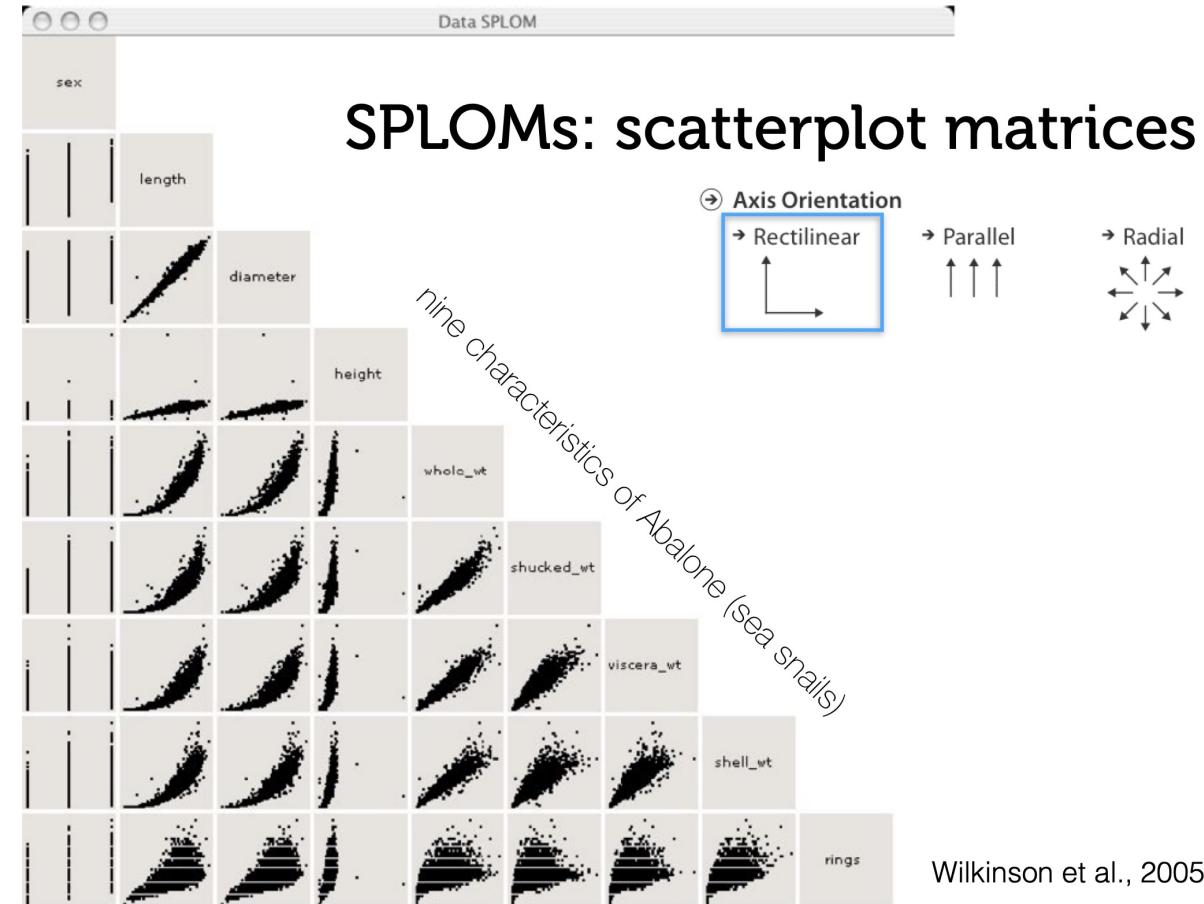
Axis Orientation: Idiom: SPLOM

- scatterplot matrix (SPLOM)
 - rectilinear axes,
point mark
 - all possible pairs of axes
 - scalability
 - one dozen attrs
 - dozens to hundreds of items



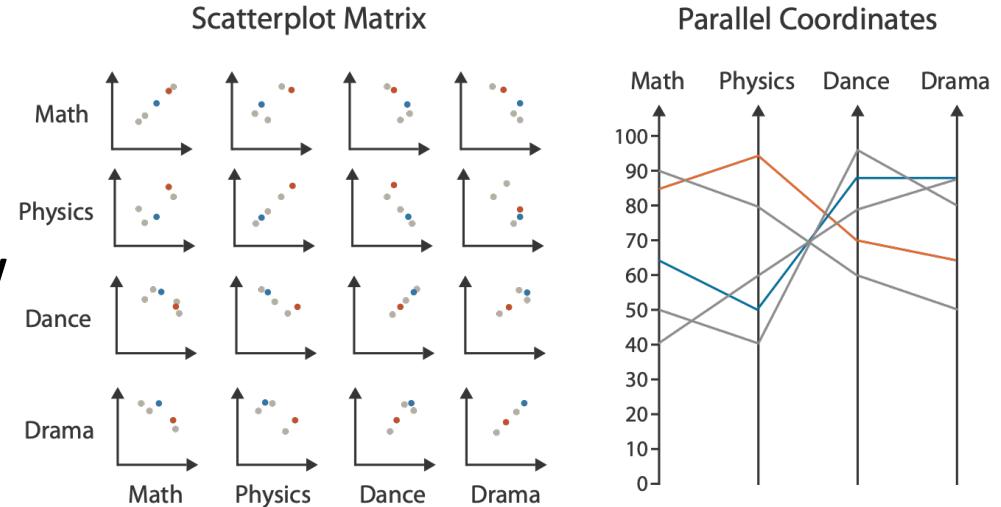
Axis Orientation: Idiom: SPLOM

- tasks
 - positive correlation
 - diagonal low-to-high
 - negative correlation
 - diagonal high-to-low
 - uncorrelated: spread out



Axis Orientation: Idiom: parallel coordinates

- scatterplot limitation
 - visual representation with orthogonal axes
 - can show only two attributes with spatial position channel
- alternative: line up axes in parallel to show many attributes with position
 - item encoded with a line with n segments
 - n is the number of attributes shown
- parallel coordinates
 - parallel axes, jagged line for item
 - rectilinear axes, item as point
 - axis ordering is major challenge
 - scalability
 - dozens of attrs
 - hundreds of items



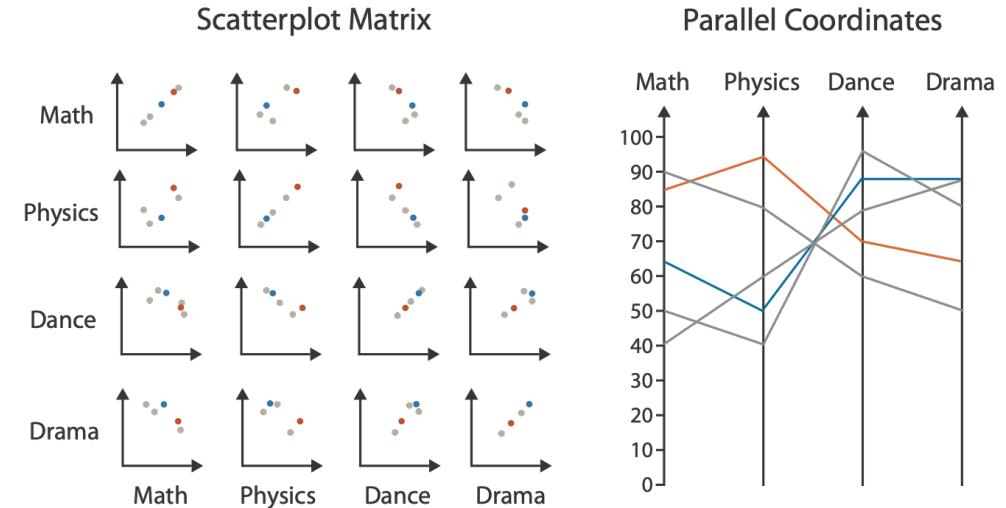
Table

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



Axis Orientation: Idiom: parallel coordinates

- tasks
 - positive correlation
 - parallel line segments
 - negative correlation
 - all segments cross at halfway point
 - uncorrelated
 - scattered crossings



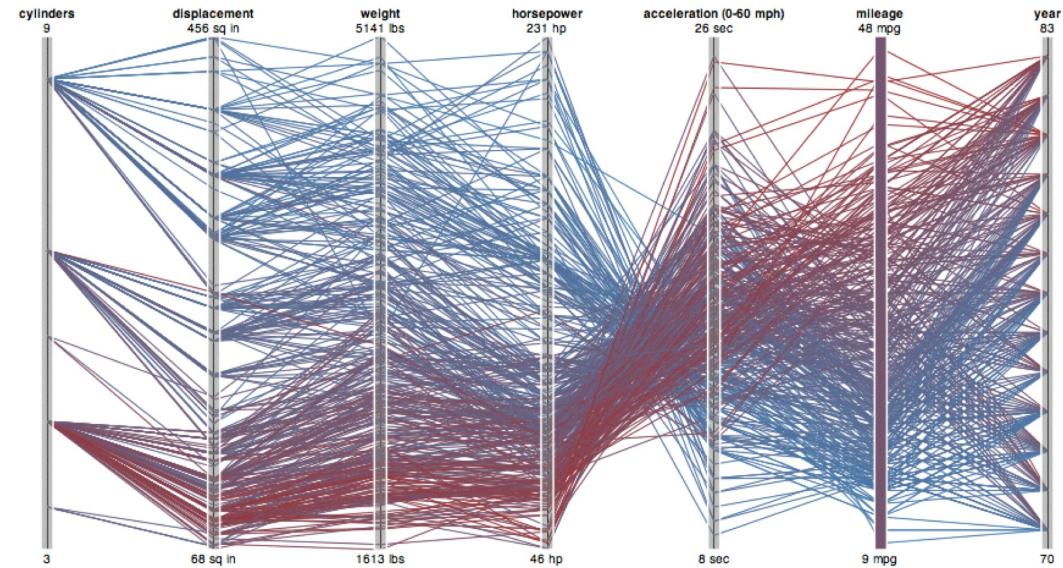
	Math	Physics	Dance	Drama
85	95	70	65	
90	80	60	50	
65	50	90	90	
50	40	95	80	
40	60	80	90	



Axis Orientation: Idiom: parallel coordinates

- **limitations**

- visible patterns only between neighbouring axis pairs
- how to pick axis order?
 - usual solution: reorderable axes, interactive exploration
 - same weakness as many other techniques
 - downside of interaction: human-powered search
 - some algorithms proposed, none fully solve



Axis Orientation: limitations

- rectilinear: scalability wrt #axes

- 2 axes best, 3 problematic, 4+ impossible

- parallel: unfamiliarity, training time

- radial: perceptual limits

- polar coordinate asymmetry

- angles lower precision than length
 - nonuniform sector width/size depending on radial distance

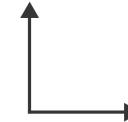
- frequently problematic

- but sometimes can be deliberately exploited!

- for 2 attrs of very unequal importance

⇒ Axis Orientation

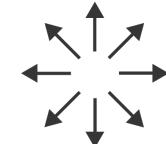
→ Rectilinear



→ Parallel



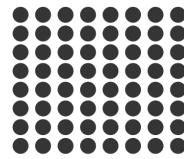
→ Radial



Layout Density

➔ Layout Density

→ Dense

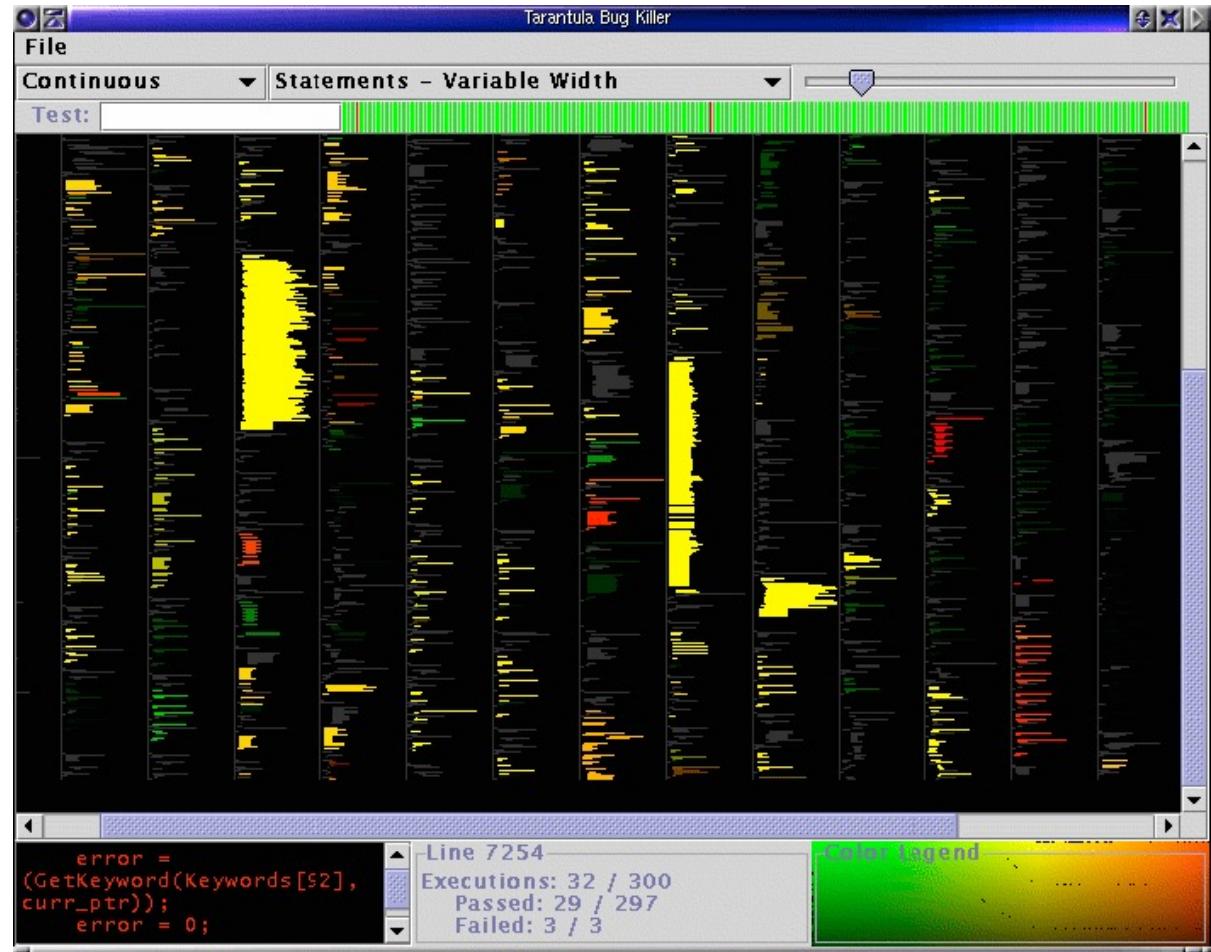


→ Space-Filling



Layout Density: Idiom: Dense software overviews

- data: text
 - text + 1 quant attrib per line
- derived data:
 - one pixel high line
 - length according to original
- colour line by attrib
- scalability
 - 10K+ lines



Summary

- Keys
- Axis orientation
- Dense Layout

