

# 07: Arrange Tables

**Enrico Puppo**

Department of Computer Science, Bioengineering, Robotics and Systems Engineering  
**University of Genova**

**90529 Data Visualization**

5-9 November 2020

**<https://2020.aulaweb.unige.it/course/view.php?id=4293>**

**Credits:**

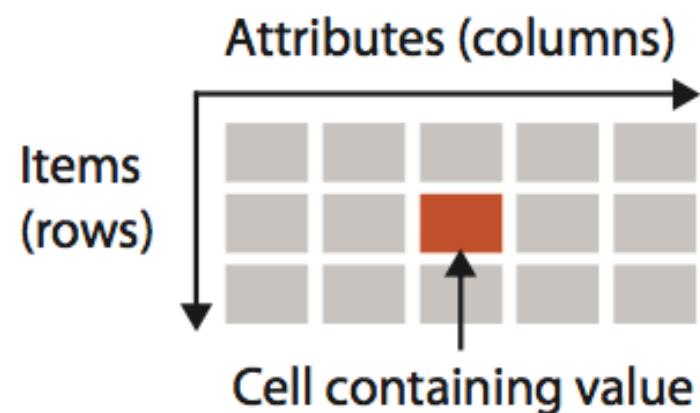
material in these slides is partially taken from

- T. Munzner, University of British Columbia
- A. Lex, University of Utah

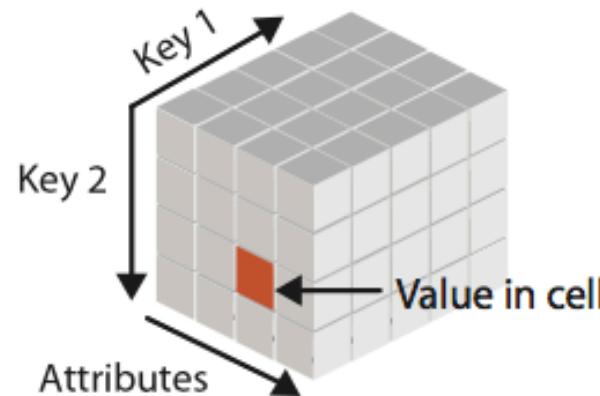
other credits in the slides

# Dataset types

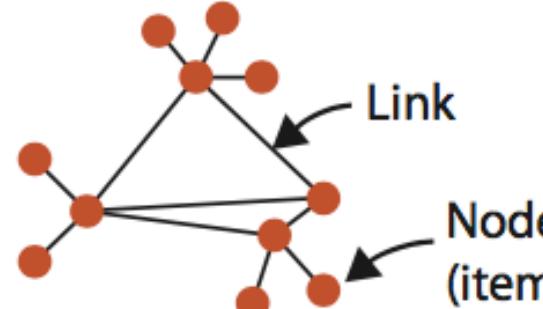
## → Tables



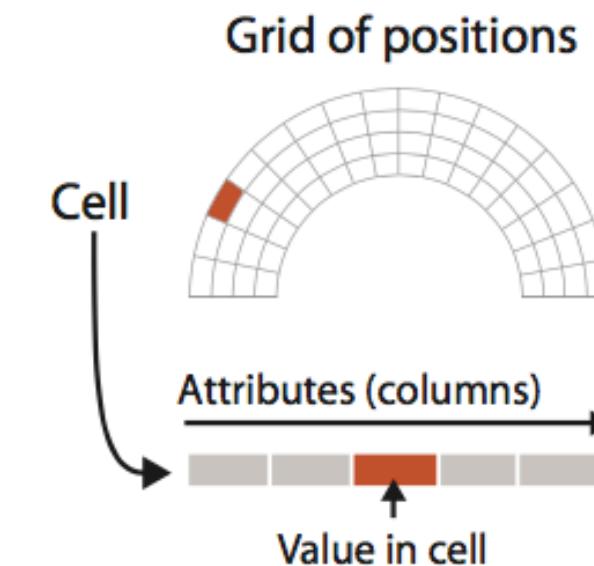
## → Multidimensional Table



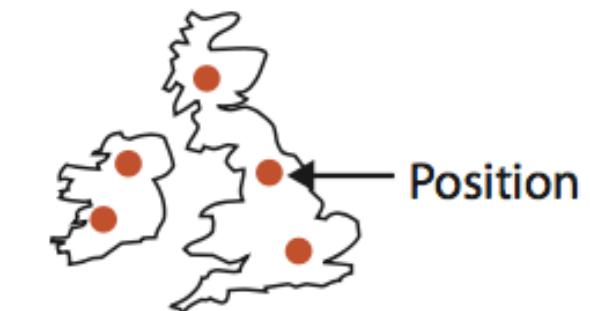
## → Networks



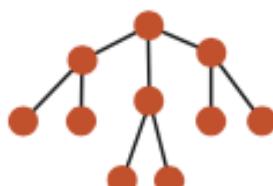
## → Fields (Continuous)



## → Geometry (Spatial)



## → Trees



# Arrange tables

- How to arrange *marks* & *channels* to represent *tabular data*
- Reminder:
  - items
  - attributes
    - keys (independent; lookup items; categorical & ordinal) - they identify items
    - values (dependent; categorical, ordinal & quantitative) - they provide information on items
- Choice of idiom depends (also) on # items / # keys / # values
- Categorical & ordinal attributes need a certain number of *levels*
  - *levels* = # values an attribute can take
  - how many keys? how many levels per key?
  - how many values?
  - (remark: quantitative attributes assume values on a continuous scale)

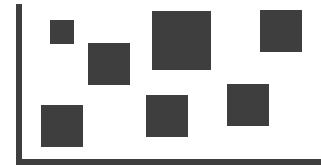
# Arrange tables

## → Express Values

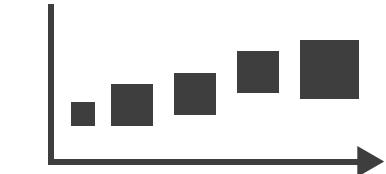


## → Separate, Order, Align Regions

→ Separate



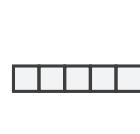
→ Order



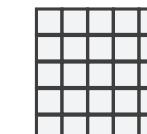
→ Align



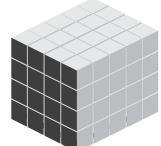
→ 1 Key  
*List*



→ 2 Keys  
*Matrix*

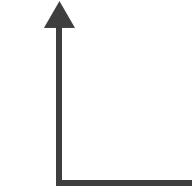


→ 3 Keys  
*Volume*

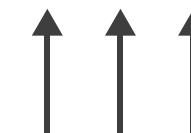


## → Axis Orientation

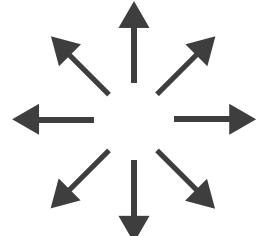
→ Rectilinear



→ Parallel

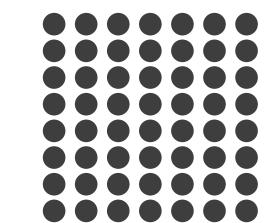


→ Radial

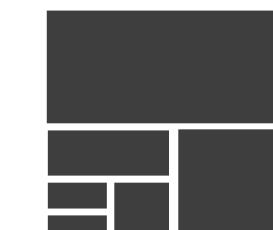


## → Layout Density

→ Dense



→ Space-Filling



→ Many Keys  
*Recursive Subdivision*



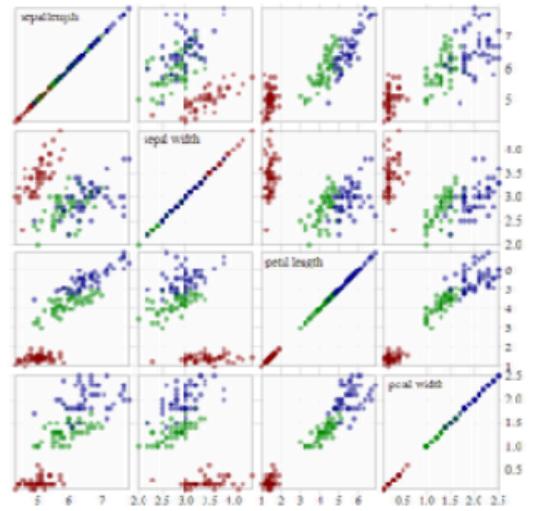
# Scale of tables

- Need different approaches for “normal” and “high-dimensional” tables
  - How many dimensions?
  - ~50 – tractable with “just” Vis
  - ~1000 – need analytical methods
- How many items?
  - ~ 1000 – “just” Vis is fine
  - >> 10,000 – need analytical methods
- Homogeneity
  - Same data type?
  - Same scales?

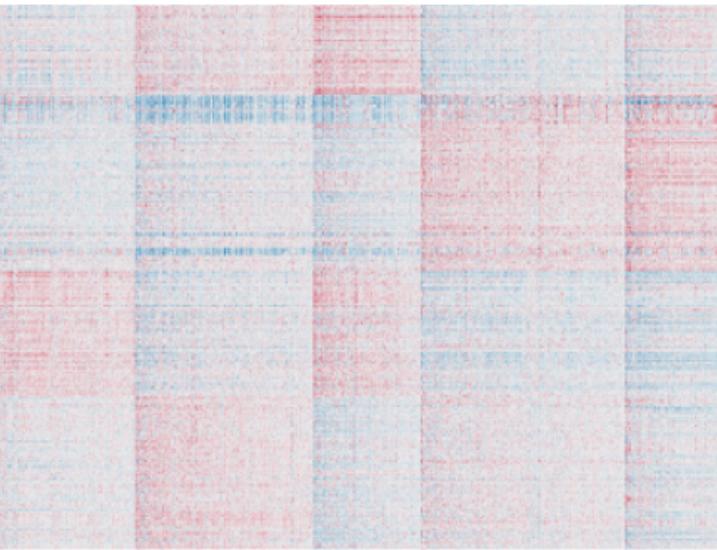
	<b>Age</b>	<b>Gender</b>	<b>Height</b>
<b>Bob</b>	<b>25</b>	<b>M</b>	<b>181</b>
<b>Alice</b>	<b>22</b>	<b>F</b>	<b>185</b>
<b>Chris</b>	<b>19</b>	<b>M</b>	<b>175</b>

	<b>BPM 1</b>	<b>BPM 2</b>	<b>BPM 3</b>
<b>Bob</b>	<b>65</b>	<b>120</b>	<b>145</b>
<b>Alice</b>	<b>80</b>	<b>135</b>	<b>185</b>
<b>Chris</b>	<b>45</b>	<b>115</b>	<b>135</b>

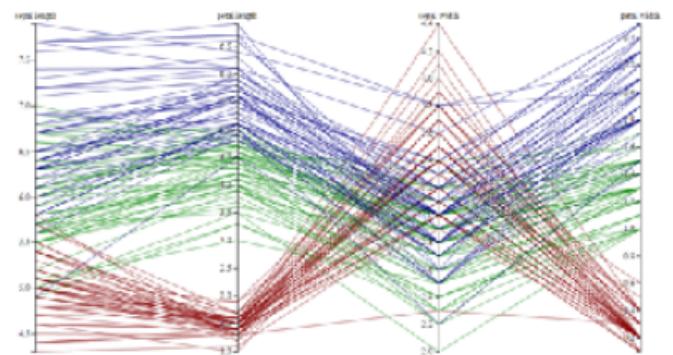
# Analytic component



**Scatterplot Matrices**  
[Bostock]



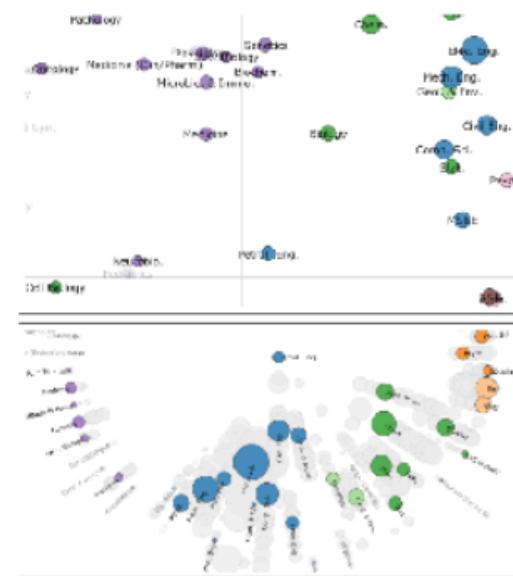
**Pixel-based visualizations /  
heat maps**



**Parallel Coordinates**  
[Bostock]



**Multidimensional Scaling**  
[Doerk 2011]



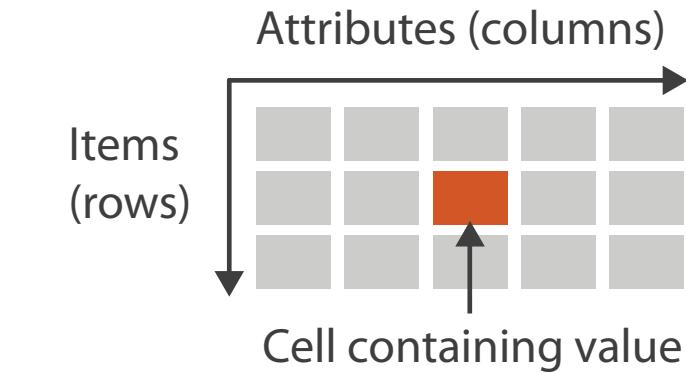
[Chuang 2012]

**no / little analytics**

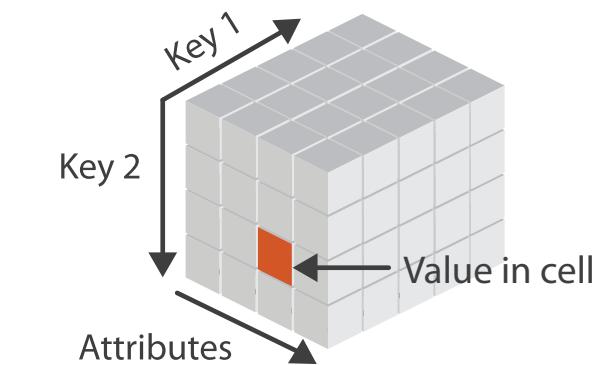
**strong analytics  
component**

# Keys and values

→ Tables



→ *Multidimensional Table*



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

→ Express Values

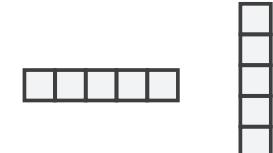
→ 1 Key

*List*



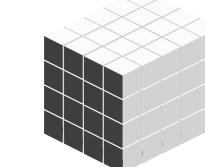
→ 2 Keys

*Matrix*



→ 3 Keys

*Volume*



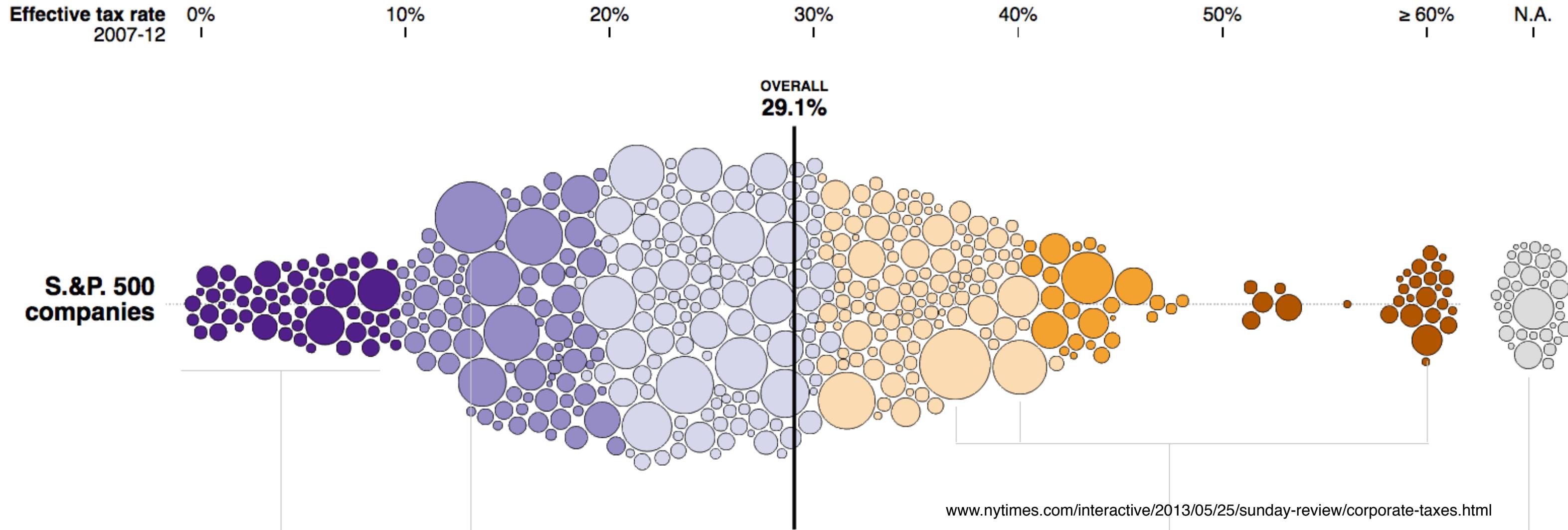
→ Many Keys

*Recursive Subdivision*



# Example

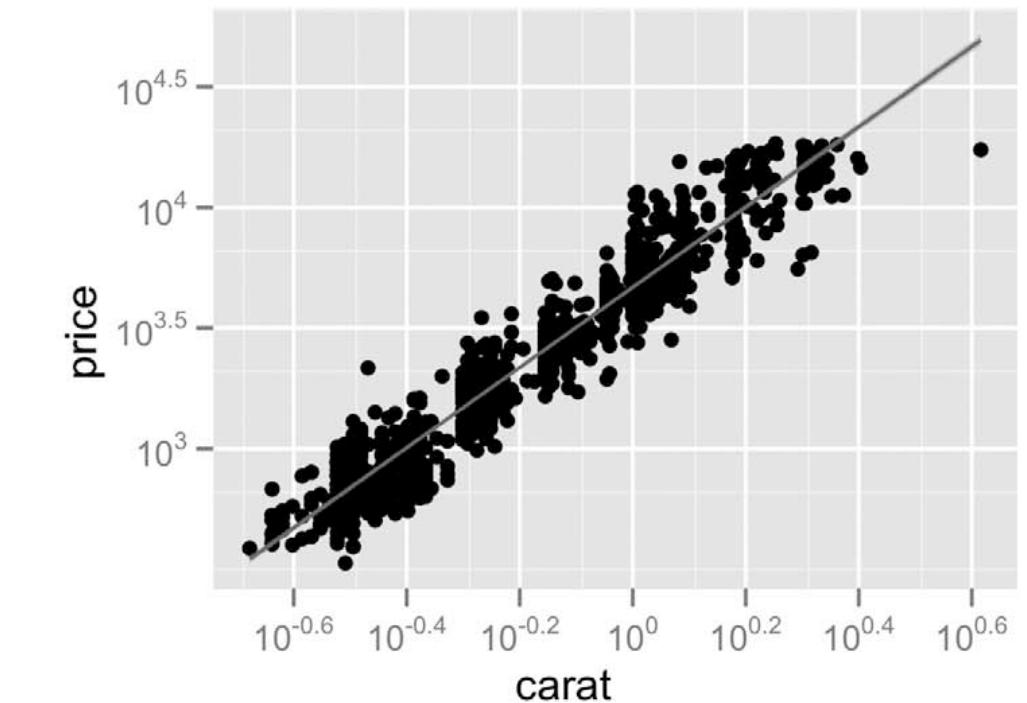
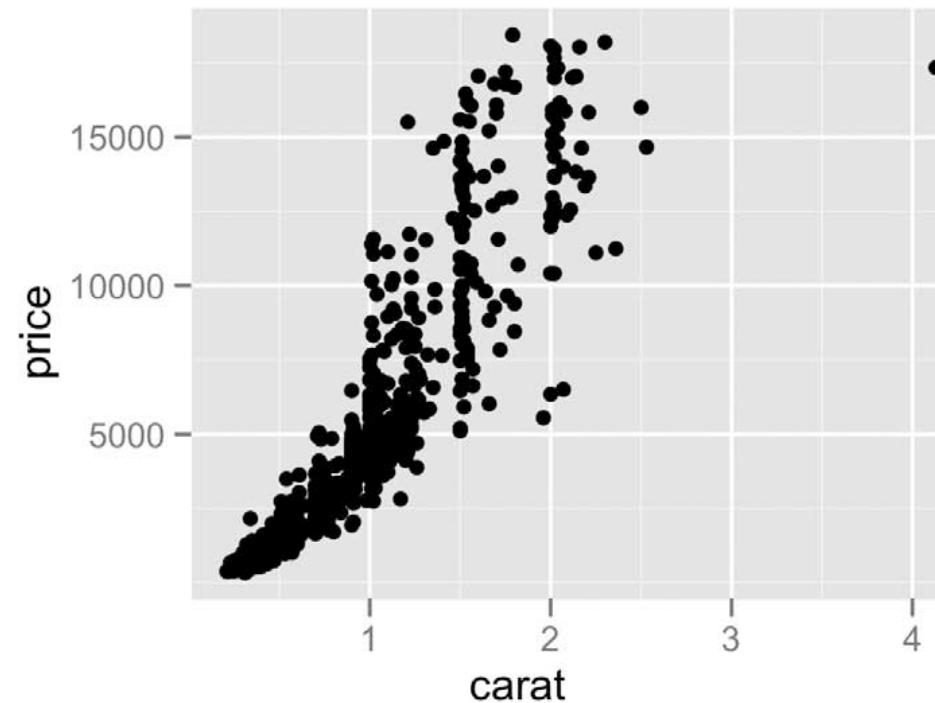
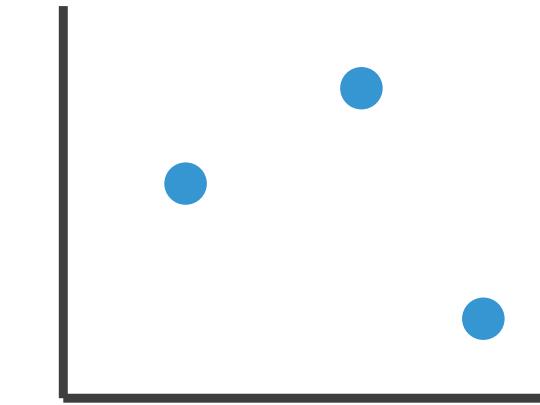
- No keys, one (fundamental) quantitative attribute per item:
  - arrange items by position along an axis
  - for additional attributes: color, size, gliph
    - *Gliph: composite object made of multiple marks forming a single shape (e.g.: arrow)*



# Idiom: scatterplot

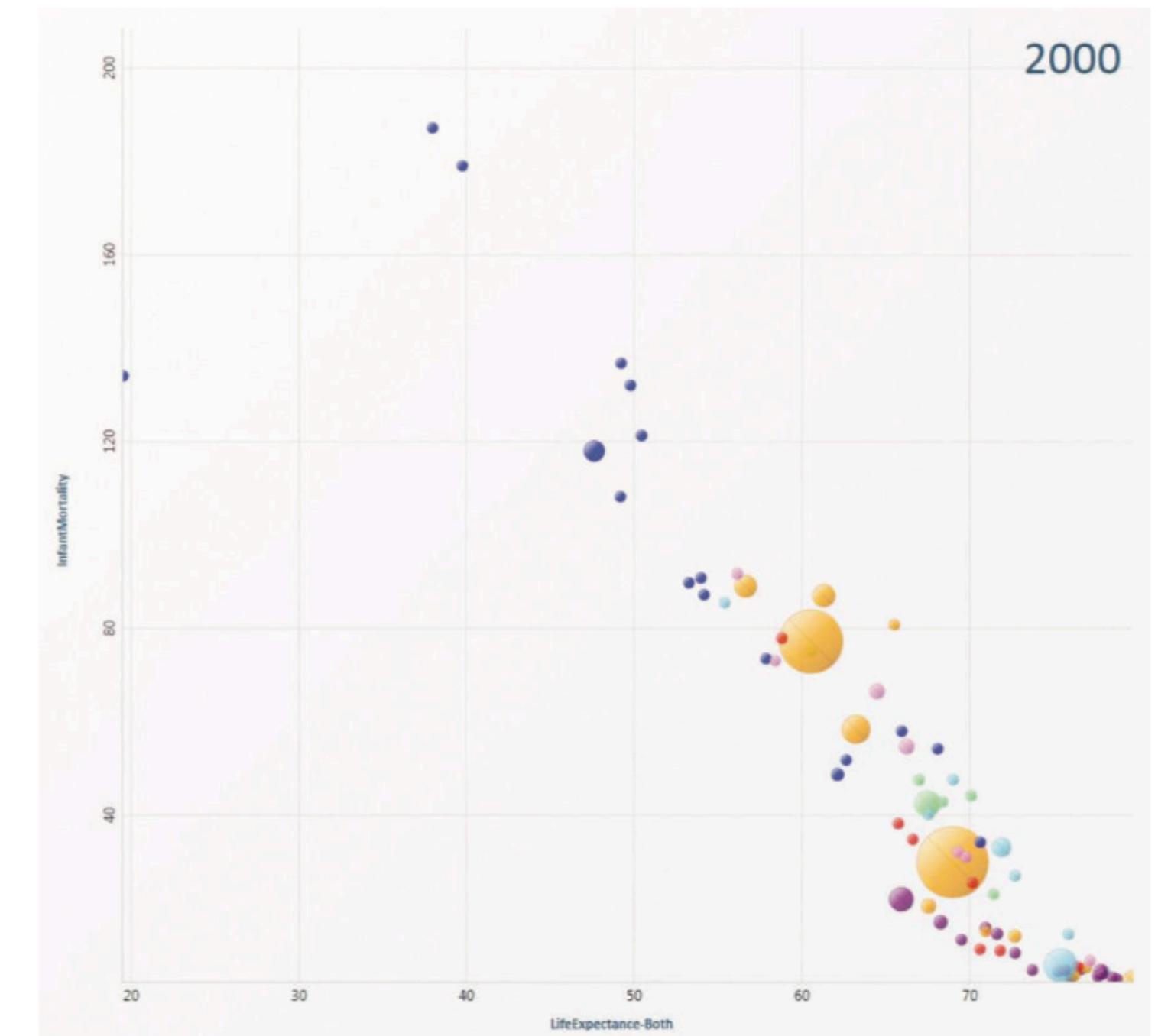
- **express** values
  - quantitative attributes
- no keys, two values
  - data
    - 2 quant attrs per item
  - mark: points (0D)
  - channels
    - horiz + vert position
  - tasks
    - find trends, outliers, distribution, correlation, clusters
  - scalability
    - hundreds of items

→ Express Values



# Idiom: scatterplot

- Additional attributes:
  - size: quantitative/ordinal
  - intensity/saturation: quantitative (ordinal); lower relevance
  - hue: categorical (ordinal)
  - shape: categorical (ordinal)
  - glyph: composite



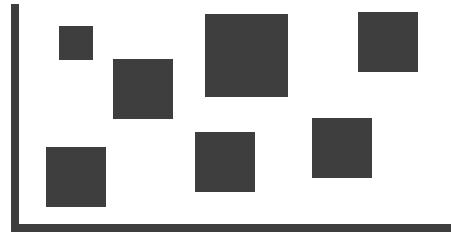
**Figure 7.2.** Scatterplot. Each point mark represents a country, with horizontal and vertical spatial position encoding the primary quantitative attributes of life expectancy and infant mortality. The color channel is used for the categorical country attribute and the size channel for quantitative population attribute. From [Robertson et al. 08, Figure 1c].

# Idiom: scatterplot

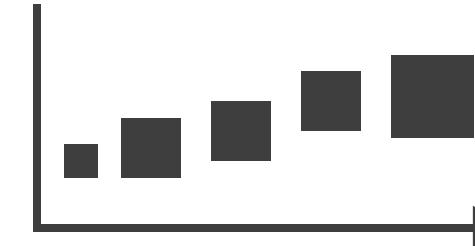
Idiom	Scatterplots
What: Data	Table: two quantitative value attributes.
How: Encode	Express values with horizontal and vertical spatial position and point marks.
Why: Task	Find trends, outliers, distribution, correlation; locate clusters.
Scale	Items: hundreds.

# Some keys: Categorical regions

→ Separate



→ Order



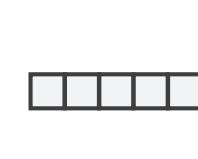
→ Align



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

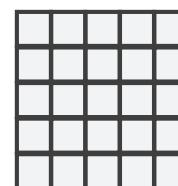
→ 1 Key

*List*



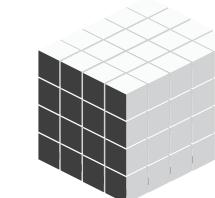
→ 2 Keys

*Matrix*



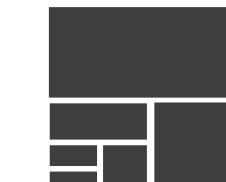
→ 3 Keys

*Volume*



→ Many Keys

*Recursive Subdivision*



# Some keys: Categorical regions

- item + categorical attribute → region
- where to put regions?
  - separate data into regions (based on categorical attribute)
  - align regions (based on ordered attribute)
  - order regions (based on ordered attribute)
- spatial proximity can be used to encode similarity
  - unless spatial position is used to represent attributes

# One key

- List alignment
  - one region per item (e.g.: row, column)
  - items span one dimension
  - attributes listed on the other dimension

- Example:

- textual, spreadsheet-like table
- not a vis idiom

A	B	C	D	E	F	G	H	I	J	K	L
Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Postal Code	City	State	Country
2	40098 CA-2014-AB10015140-41954	11/11/14	13/11/14	First Class	AB-100151402	Aaron Bergman	Consumer	73120	Oklahoma City	Oklahoma	United States
3	26341 IN-2014-JR162107-41675	05/02/14	07/02/14	Second Class	JR-162107	Justin Ritter	Corporate		Wollongong	New South Wales	Australia
4	25330 IN-2014-CR127307-41929	17/10/14	18/10/14	First Class	CR-127307	Craig Reiter	Consumer		Brisbane	Queensland	Australia
5	13524 ES-2014-KM1637548-41667	28/01/14	30/01/14	First Class	KM-1637548	Katherine Murray	Home Office		Berlin	Berlin	Germany
6	47221 SG-2014-RH9495111-41948	05/11/14	06/11/14	Same Day	RH-9495111	Rick Hansen	Consumer		Dakar	Dakar	Senegal
7	22732 IN-2014-JM156557-41818	28/06/14	01/07/14	Second Class	JM-156557	Jim Mitchum	Corporate		Sydney	New South Wales	Australia
8	30570 IN-2012-TS2134092-41219	06/11/12	08/11/12	First Class	TS-2134092	Toby Swindell	Consumer		Porirua	Wellington	New Zealand
9	31192 IN-2013-MB1808592-41378	14/04/13	18/04/13	Standard Class	MB-1808592	Mick Brown	Consumer		Hamilton	Waikato	New Zealand
10	40099 CA-2014-AB10015140-41954	11/11/14	13/11/14	First Class	AB-100151402	Aaron Bergman	Consumer	73120	Oklahoma City	Oklahoma	United States
11	36258 CA-2012-AB10015140-40974	06/03/12	07/03/12	First Class	AB-100151404	Aaron Bergman	Consumer	98103	Seattle	Washington	United States
12	36259 CA-2012-AB10015140-40974	06/03/12	07/03/12	First Class	AB-100151404	Aaron Bergman	Consumer	98103	Seattle	Washington	United States
13	28879 ID-2013-AJ107801-41383	19/04/13	22/04/13	First Class	AJ-107801	Anthony Jacobs	Corporate		Kabul	Kabul	Afghanistan
14	45794 SA-2012-MM7260110-41269	26/12/12	28/12/12	Second Class	MM-7260110	Magdelene Morse	Consumer		Jizan	Jizan	Saudi Arabia
15	4132 MX-2013-VF2171518-41591	13/11/13	13/11/13	Same Day	VF-2171518	Vicky Freymann	Home Office		Toledo	Parana	Brazil
16	27704 IN-2014-PF191207-41796	06/06/14	08/06/14	Second Class	PF-1912027	Peter Fuller	Consumer		Mudanjiang	Heilongjiang	China
17	13779 ES-2015-BP1118545-42216	31/07/15	03/08/15	Second Class	BP-1118545	Ben Peterman	Corporate		Paris	Ile-de-France	France
18	39519 CA-2012-AB10015140-40958	19/02/12	25/02/12	Standard Class	AB-100151402	Aaron Bergman	Consumer	76017	Arlington	Texas	United States
19	12069 ES-2015-PJ1883564-42255	08/09/15	14/09/15	Standard Class	PJ-1883564	Patrick Jones	Corporate		Prato	Tuscany	Italy
20	22096 IN-2015-JS156857-42035	31/01/15	01/02/15	First Class	JS-156857	Jim Sink	Corporate		Townsville	Queensland	Australia
21	49463 TZ-2015-RH9555129-42343	05/12/15	07/12/15	Second Class	RH-9555129	Rita Hightower	Consumer		Uvinza	Kigoma	Tanzania
22	46630 PL-2013-AB600103-41494	08/08/13	10/08/13	First Class	AB-600103	Ann Blume	Corporate		Bytom	Silesia	Poland
23	36260 CA-2012-AB10015140-40974	06/03/12	07/03/12	First Class	AB-100151404	Aaron Bergman	Consumer	98103	Seattle	Washington	United States
24	21586 IN-2012-JK1532527-41030	01/05/12	02/05/12	First Class	JK-1532527	Jason Klamczynski	Corporate		Suzhou	Anhui	China
25	13528 ES-2014-LB16795139-41697	27/02/14	01/03/14	Second Class	LB-16795139	Laurel Beltran	Home Office		Edinburgh	Scotland	United Kingdom
26	1570 US-2015-NP1832582-42216	31/07/15	01/08/15	First Class	NP-1832582	Naresj Patel	Consumer		Juárez	Chihuahua	Mexico
27	3484 MX-2015-VD2167039-42252	05/09/15	08/09/15	First Class	VD-2167039	Valerie Dominguez	Consumer		Soyapango	San Salvador	El Salvador
28	30191 IN-2012-PB19210127-41259	16/12/12	19/12/12	First Class	PB-19210127	Phillip Breyer	Corporate		Taipei	Taipei City	Taiwan
29	11645 ES-2012-EB1411048-40981	13/03/12	16/03/12	Second Class	EB-1411048	Eugene Barchas	Consumer		Leipzig	Saxony	Germany
30	38460 CA-2012-AH10030140-41020	21/04/12	23/04/12	Second Class	AH-100301406	Aaron Hawkins	Corporate	12180	Troy	New York	United States
31	22999 IN-2013-BP1123058-41329	24/02/13	24/02/13	Same Day	BP-1123058	Benjamin Patterson	Consumer		Surat	Gujarat	India
32	220 US-2012-RR1952536-41270	27/12/12	29/12/12	Second Class	RR-1952536	Rick Reed	Corporate		Santo Domingo	Santo Domingo	Dominican Republic
33	10648 ES-2013-BS1136545-41472	17/07/13	19/07/13	First Class	BS-1136545	Bill Shonely	Corporate		Saint-Brieuc	Brittany	France
34	40977 CA-2013-AH10030140-41635	27/12/13	31/12/13	Standard Class	AH-100301404	Aaron Hawkins	Corporate	94122	San Francisco	California	United States
35	21286 IN-2012-DP131057-41215	02/11/12	04/11/12	Second Class	DP-131057	Dave Poirier	Corporate		Gold Coast	Queensland	Australia
36	40976 CA-2013-AH10030140-41635	27/12/13	31/12/13	Standard Class	AH-100301404	Aaron Hawkins	Corporate	94122	San Francisco	California	United States
37	47905 CG-2012-AH1003033-41165	13/09/12	14/09/12	First Class	AH-1003033	Aaron Hawkins	Corporate		Kamina	Katanga	Democratic Republic of the Congo
38	36651 CA-2012-AH10030140-41041	12/05/12	18/05/12	Standard Class	AH-100301404	Aaron Hawkins	Corporate	90004	Los Angeles	California	United States
39	40100 CA-2012-AH10030140-41273	30/12/12	31/12/12	First Class	AH-100301406	Aaron Hawkins	Corporate	10035	New York City	New York	United States

# One key - all quantitative attributes

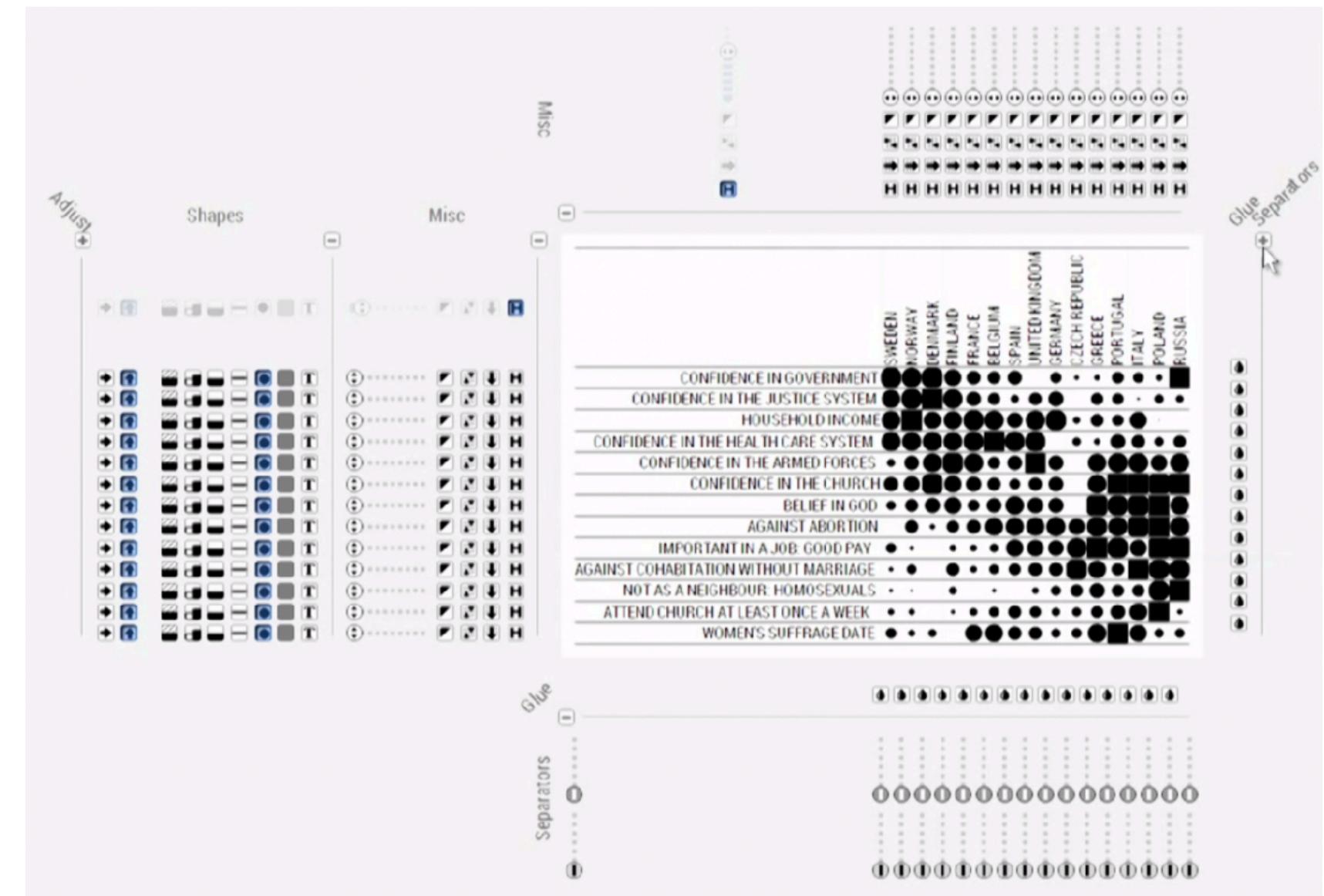
- Table cells can contain vis marks

- Examples:

- *Bertifier*
- *Polaris*
- *Tableau*

- Task:

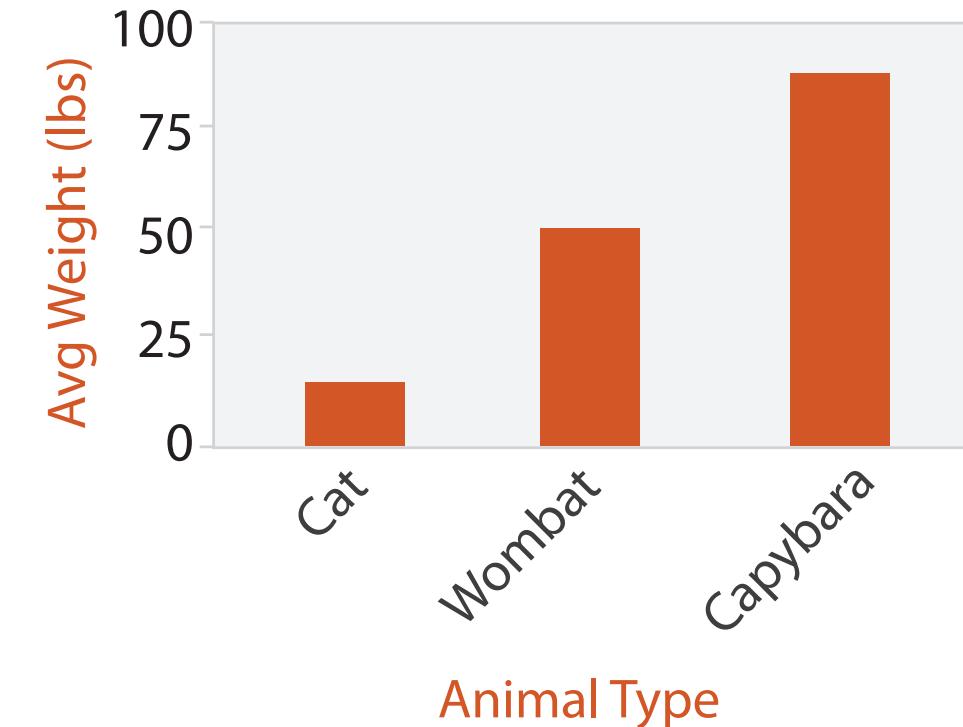
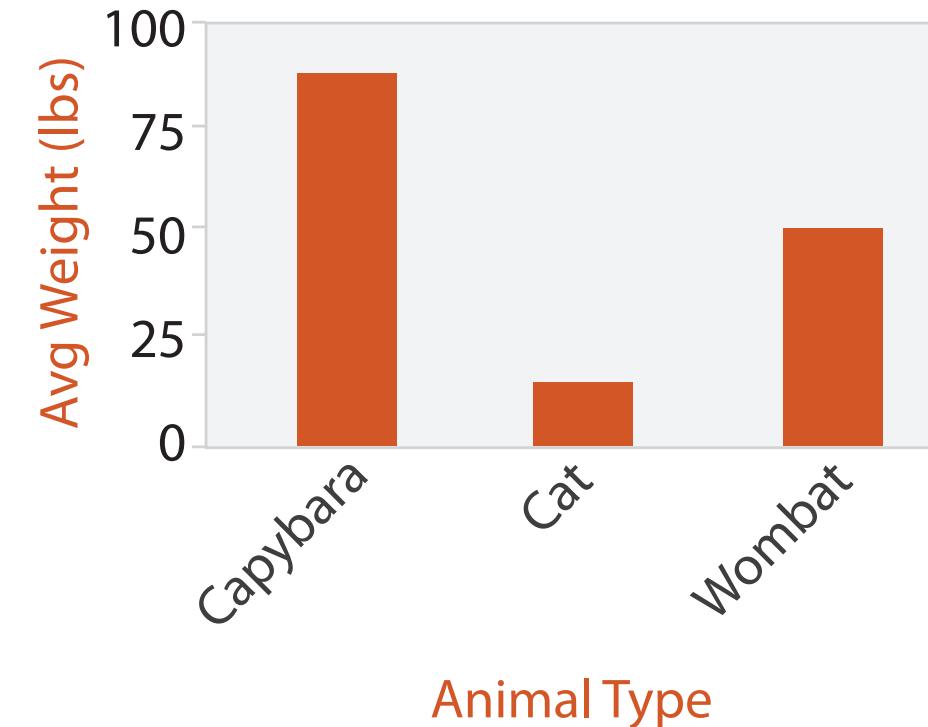
- discover correlation & similarity
- major challenge: sort rows & cols



[Charles Perin, Pierre Dragicevic, Jean-Daniel Fekete (2014). Revisiting Bertin's Matrices: New Interactions for Crafting Tabular Visualizations. TVCG, VIS' 2014.]

# Idiom: bar chart

- one key, one value
  - data
    - 1 categor attrib, 1 quant attrib
  - mark: lines
  - channels
    - length to express quant value
    - spatial regions: one per mark
      - separated horizontally, aligned vertically (or vice-versa)
      - ordered by quantitative attribute
        - » by label (alphabetical), by length attribute (data-driven)
  - task
    - compare, lookup values
  - scalability
    - dozens to hundreds of items / levels of key - constrained by screen space



# Idiom: bar chart

Idiom

What: Data

How: Encode

Why: Task

Scale

## Bar Charts

Table: one quantitative value attribute, one categorical key attribute.

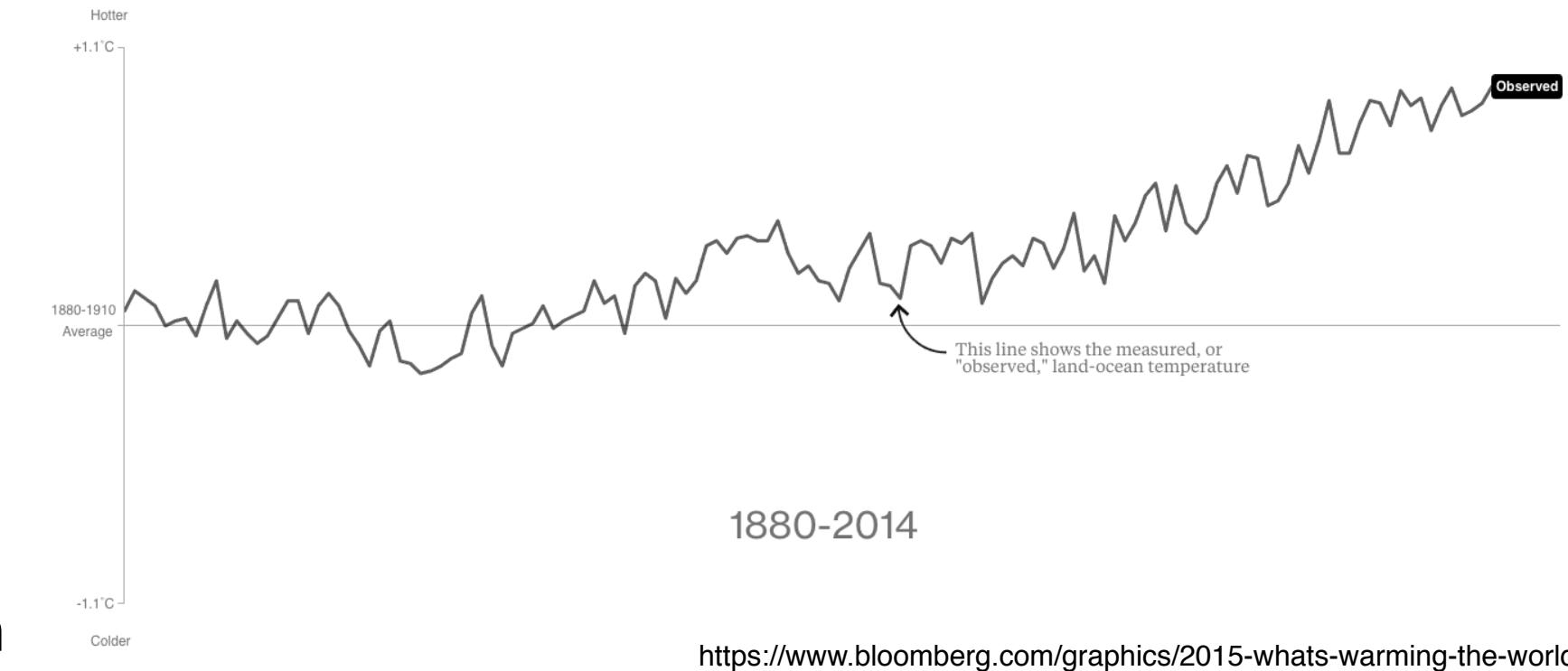
Line marks, express value attribute with aligned vertical position, separate key attribute with horizontal position.

Lookup and compare values.

Key attribute: dozens to hundreds of levels.

# Idiom: line chart

- one key, one value
  - emphasizing horizontal continuity
  - data
    - 2 quantitative attributes
  - mark: points
    - line connection marks between them
  - channels
    - aligned lengths, separated and ordered by key attribute
  - task
    - find trend
      - connection marks emphasize ordering of items along key axis by explicitly showing relationship between one item and the next
  - scalability
    - continuous scales in both dimensions: ~1000 levels (constrained by resolution)



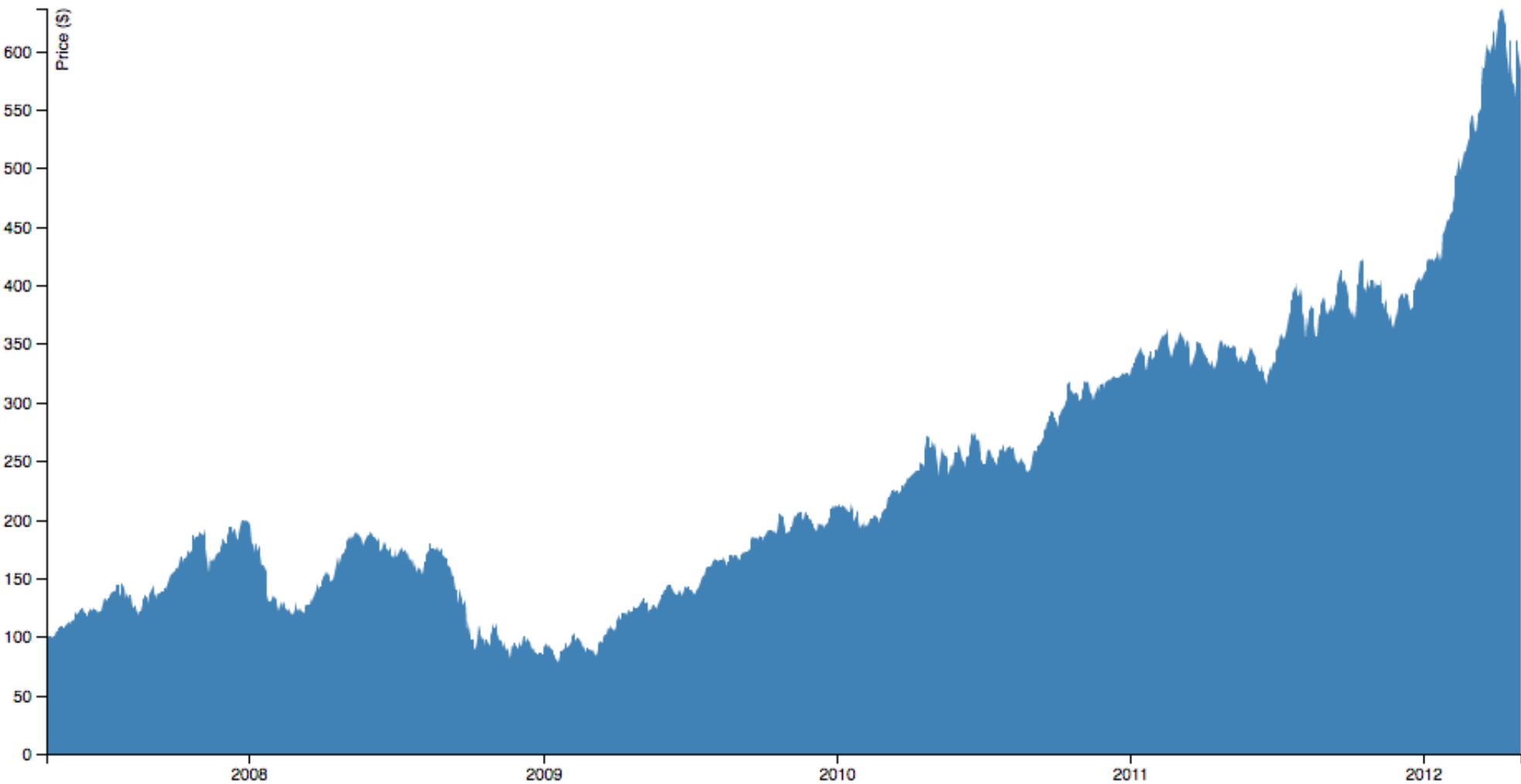
<https://www.bloomberg.com/graphics/2015-whats-warming-the-world/>

# Idioms: line chart & dot chart

Idiom	Dot Charts
What: Data	Table: one quantitative value attribute, one ordered key attribute.
How: Encode	Express value attribute with aligned vertical position and point marks. Separate/order into horizontal regions by key attribute.
Idiom	Line Charts
What: Data	Table: one quantitative value attribute, one ordered key attribute.
How: Encode	Dot chart with connection marks between dots.
Why	Show trend.
Scale	Key attribute: hundreds of levels.

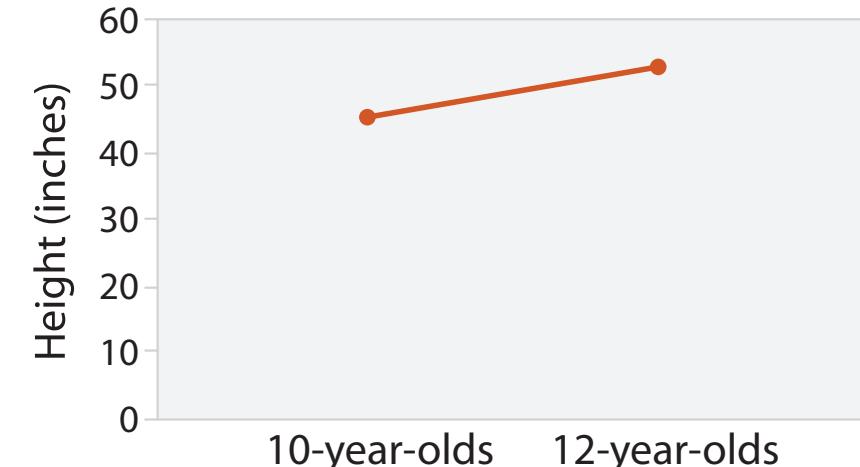
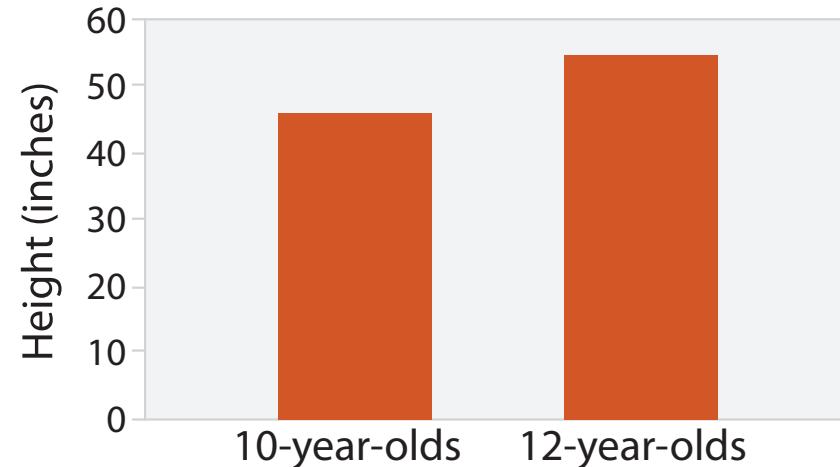
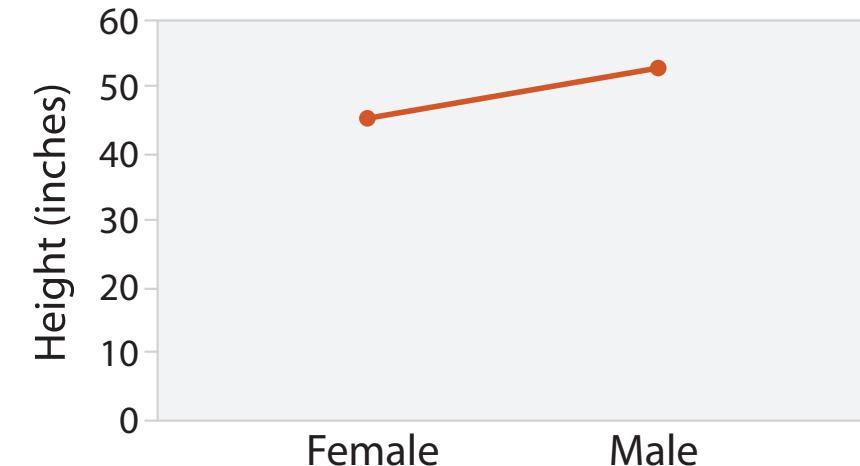
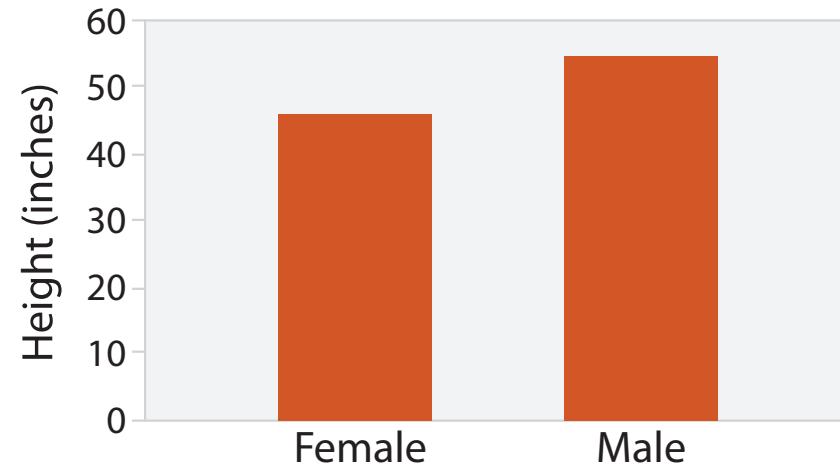
# Idiom: area chart

- Variant of line chart
  - same kind of data
  - highlights area below chart
    - integral properties
  - task
    - summarize trend & distribution



# Choosing bar vs line charts

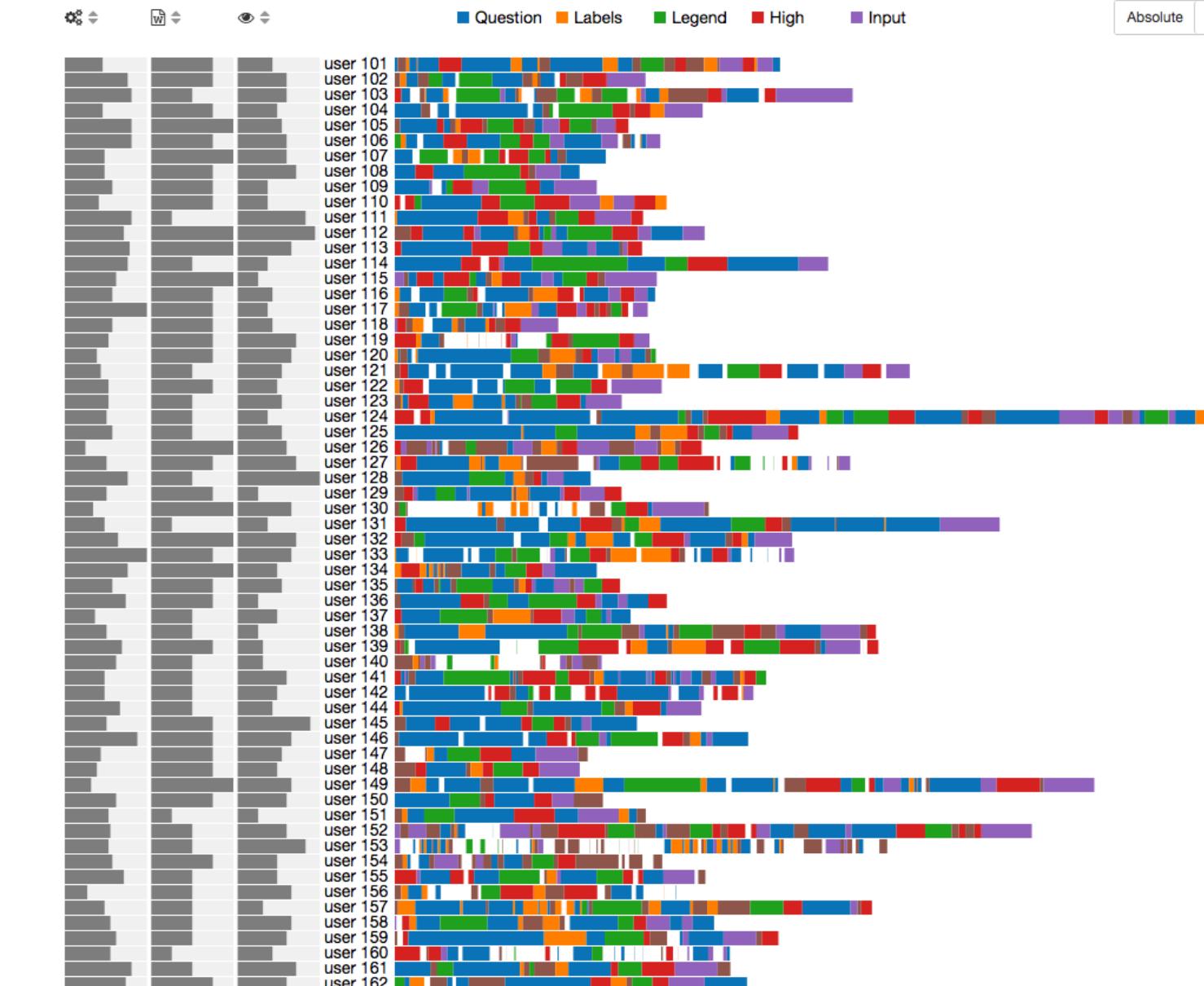
- depends on type of key attribute
  - bar charts if categorical
  - line charts if ordered
- do not use line charts for categorical key attributes
  - violates expressiveness principle
    - implication of trend so strong that it overrides semantics!
    - “The more male a person is, the taller he/she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



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

# Idiom: stacked bar chart

- one more key
  - data
    - 2 categorical attrs, 1 quantitative attrib
  - mark: vertical stack of line marks
    - *glyph*: internal structure from multiple marks
  - channels
    - length and color hue
    - spatial regions: one per glyph
      - aligned: full glyph, lowest bar component
      - unaligned: other bar components
  - task
    - part-to-whole relationship
  - scalability
    - several to one dozen levels for stacked attribute



<https://www.cs.ubc.ca/labs/imager/tr/2015/SEQIT/demo/>

# Idiom: stacked bar chart - alternative view

- one key to identify items

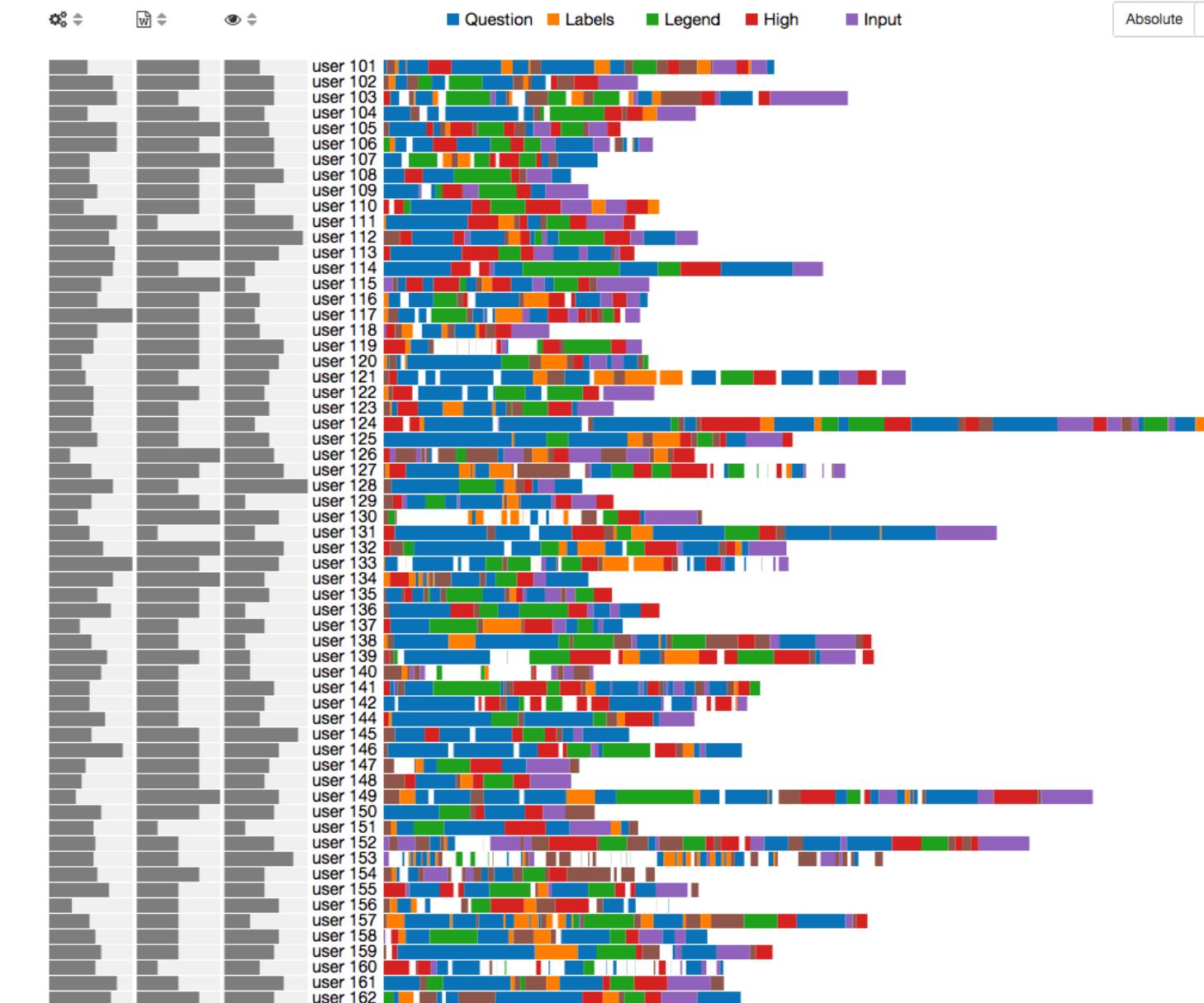
- data

- several *homogeneous* quantitative attributes
    - items are listed along one dimension (key)
    - attributes are stacked along the other dimension

- scalability

- ~100 items; several to one dozen attributes

- the rest is the same



# Idiom: stacked bar chart

Idiom

## Stacked Bar Charts

What: Data

Multidimensional table: one quantitative value attribute, two categorical key attributes.

How: Encode

Bar glyph with length-coded subcomponents of value attribute for each category of secondary key attribute. Separate bars by category of primary key attribute.

Why: Task

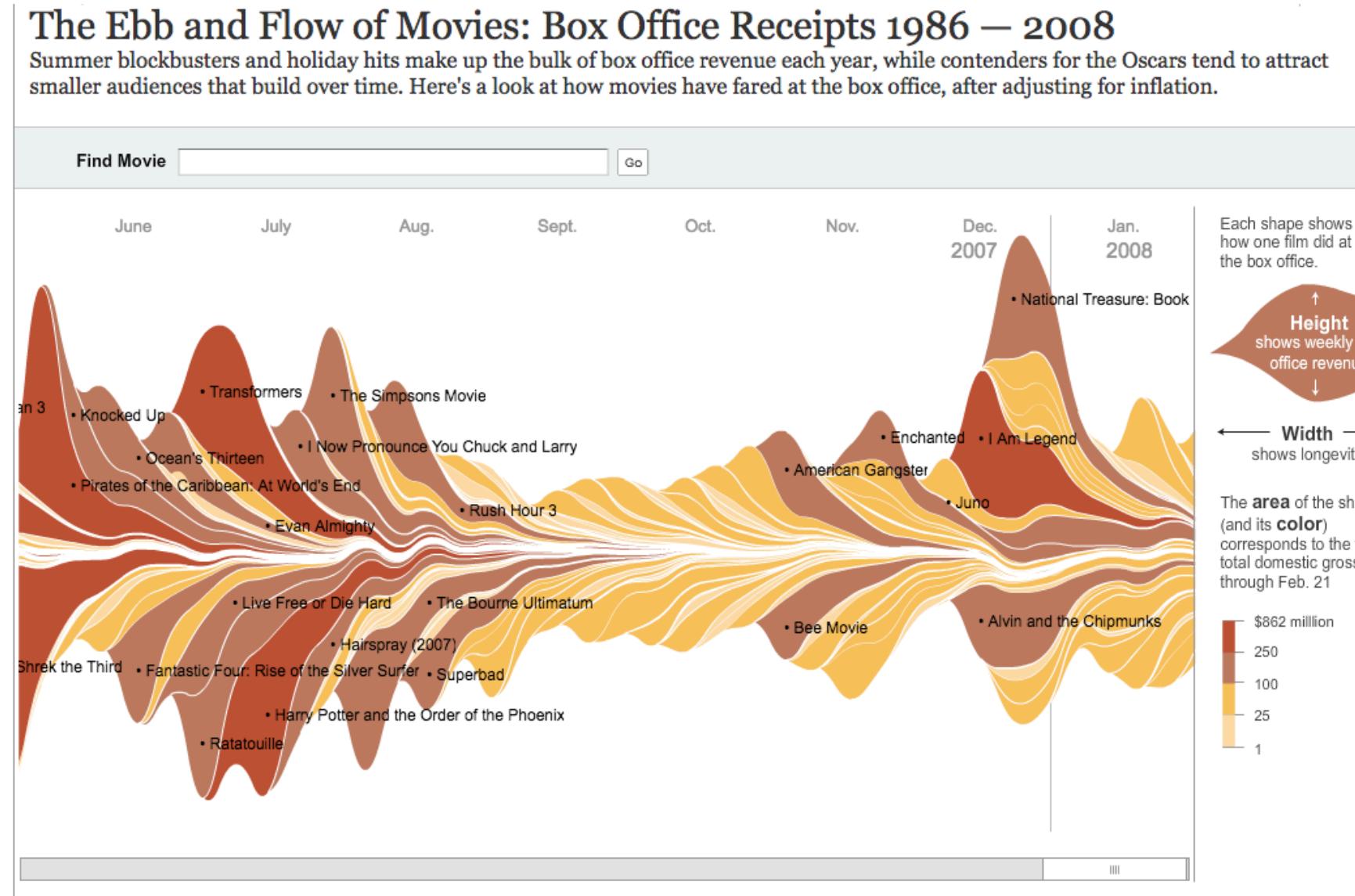
Part-to-whole relationship, lookup values, find trends.

Scale

Key attribute (main axis): dozens to hundreds of levels. Key attribute (stacked glyph axis): several to one dozen

# Idiom: streamgraph

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



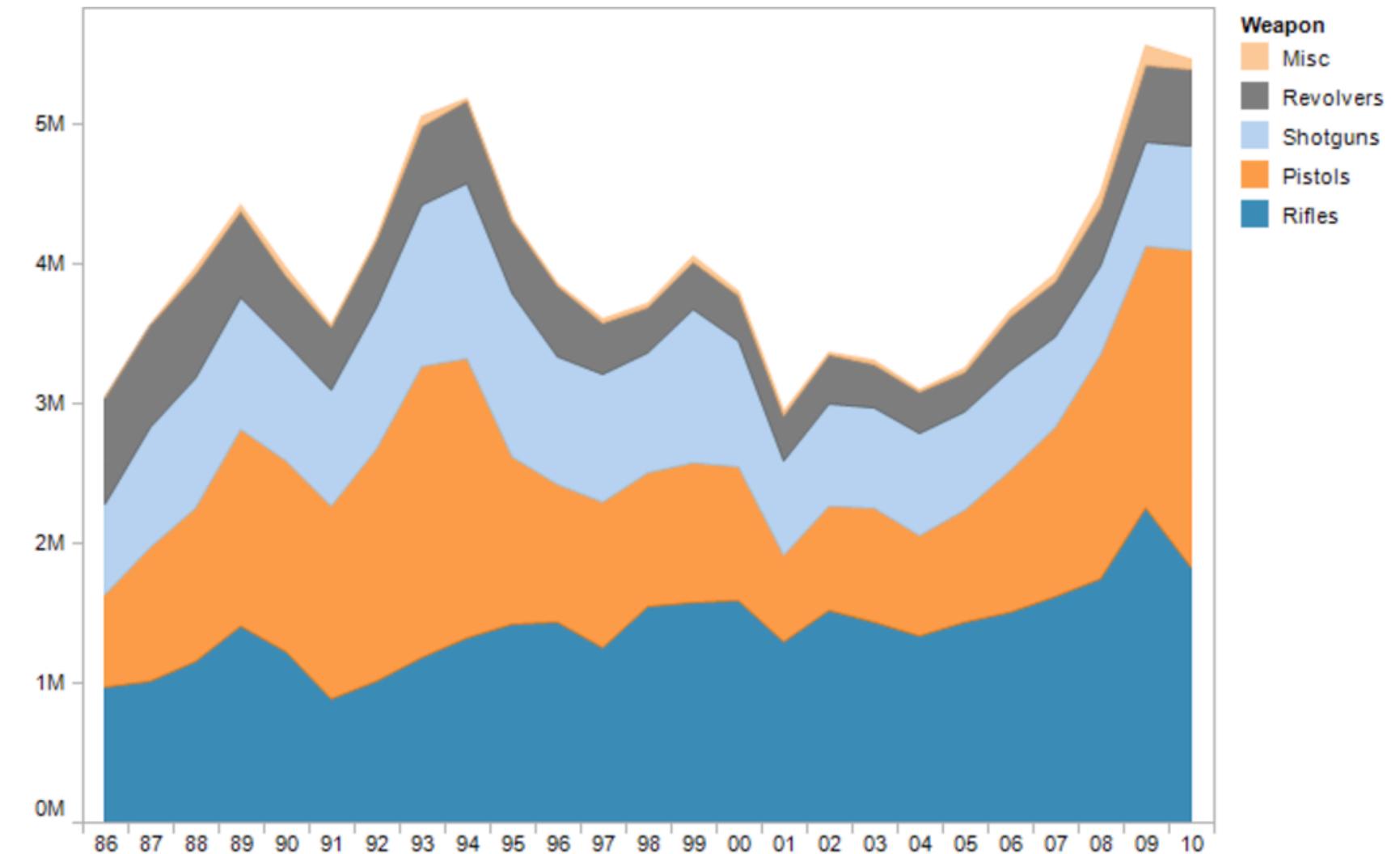
[www.nytimes.com/interactive/2008/02/23/movies/20080223\\_REVENU GRAPHIC.html](http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENU GRAPHIC.html)

# Idiom: streamgraph

Idiom	Streamgraphs
What: Data	Multidimensional table: one quantitative value attribute (counts), one ordered key attribute (time), one categorical key attribute (artist).
What: Derived	One quantitative attribute (for layer ordering).
How: Encode	Use derived geometry showing artist layers across time, layer height encodes counts.
Scale	Key attributes (time, main axis): hundreds of time points. Key attributes (artists, short axis): dozens to hundreds

# Idiom: stacked area chart

- Continuous generalization of stacked bar chart
  - similar to streamgraph
  - alignment to reference frame
  - task
    - find trend (of total)
    - compare different attrs (not aligned)
  - scalability
    - continuous scales in both dimensions
    - limited number of categories



# Two keys

- Grid alignment
  - one region of a 2D grid per item (grid cell)
  - keys index the two dimensions of the grid
  - each cell contains information on data identified by the pair of keys
    - attribute
    - glyph representing several attributes
    - derived diagram

# Idiom: heatmap

- two keys, one value

- data

- 2 categorical keys (gene, experimental condition)
    - 1 quantitative attribute (expression levels)

- marks: area

- separate and align in 2D matrix
      - indexed by 2 categorical attributes (keys)

- channels

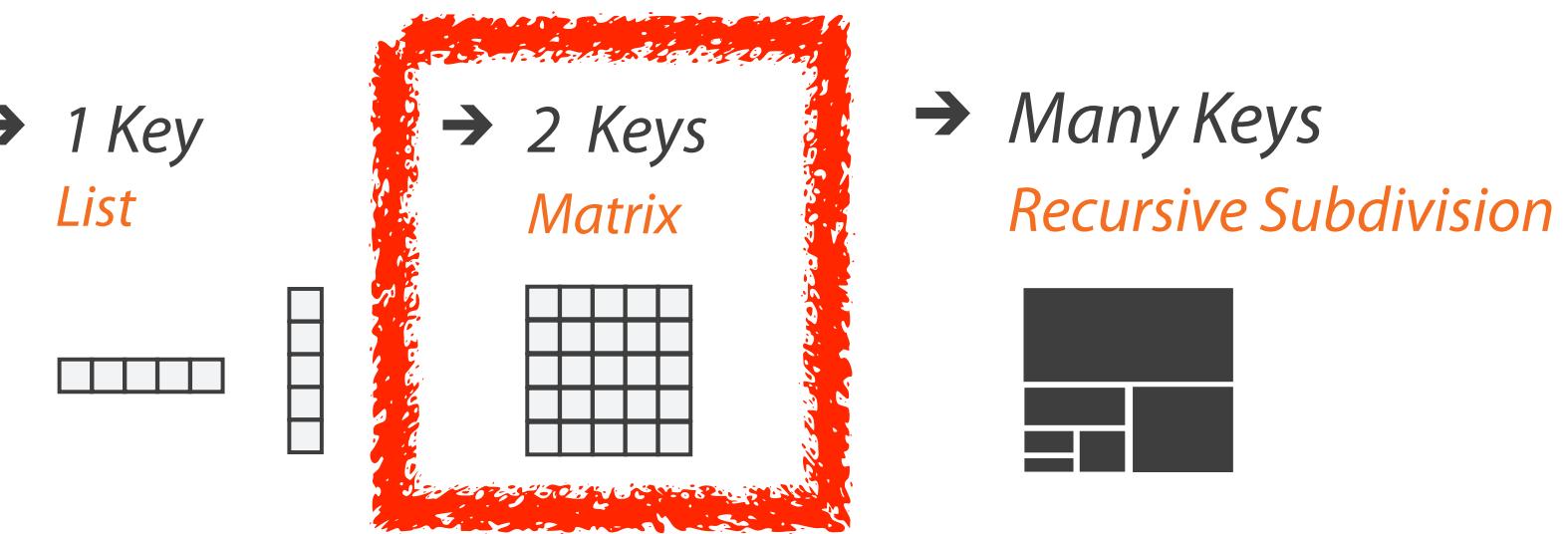
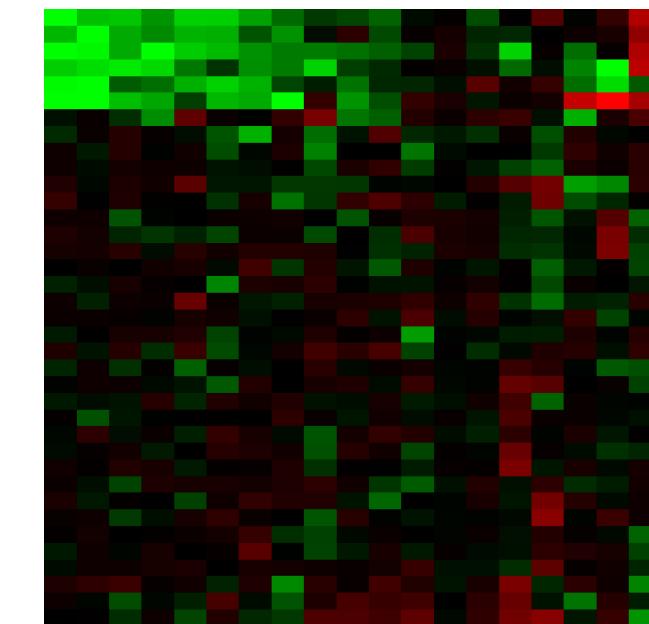
- color by quantitative attribute
      - (ordered diverging colormap)

- task

- find clusters, outliers

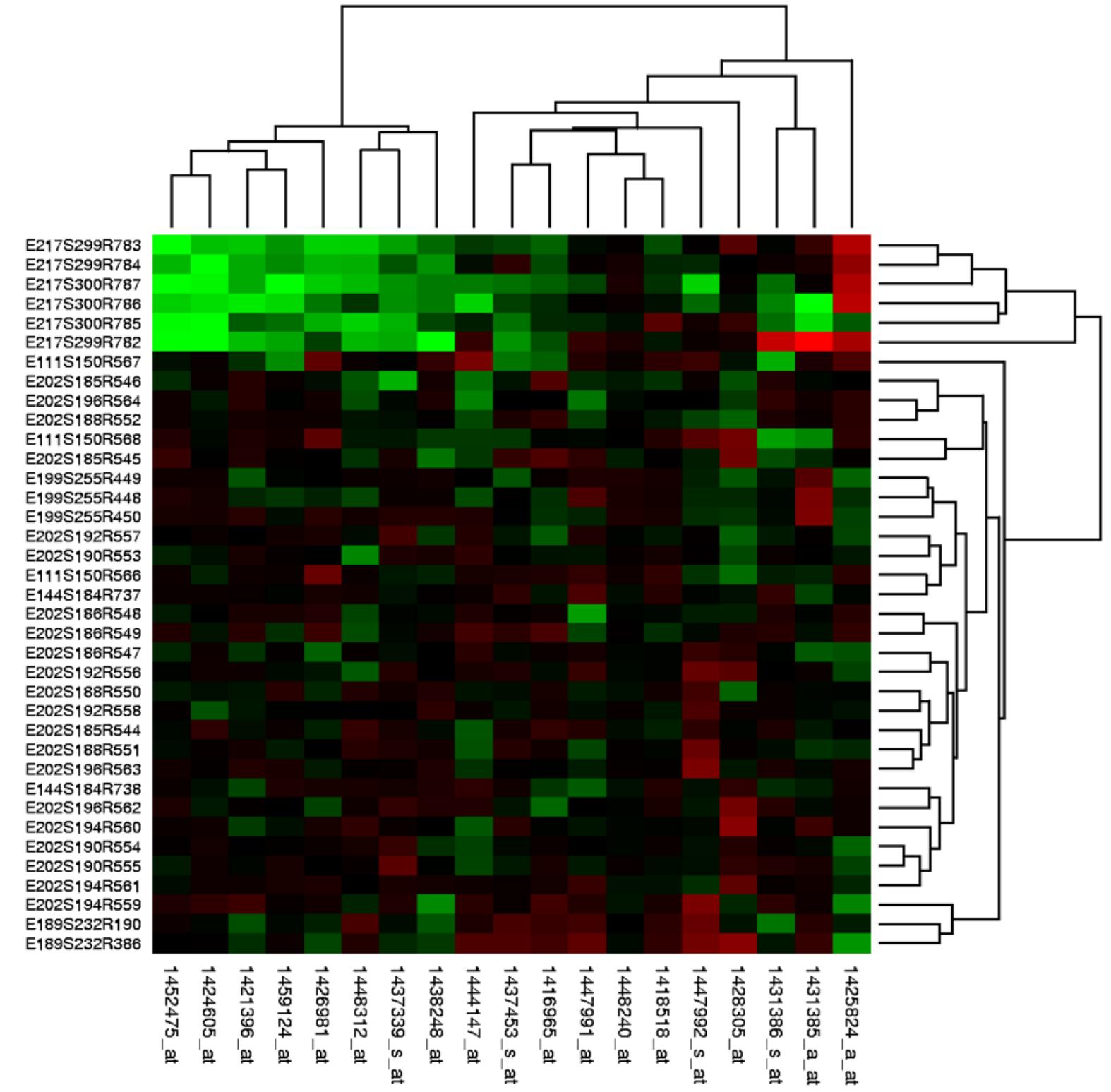
- scalability

- limit ~1M items, 100s of categorical levels, ~10 quantitative attribute 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 are easy to compare
  - heatmap
    - marks (re-)ordered by cluster hierarchy traversal

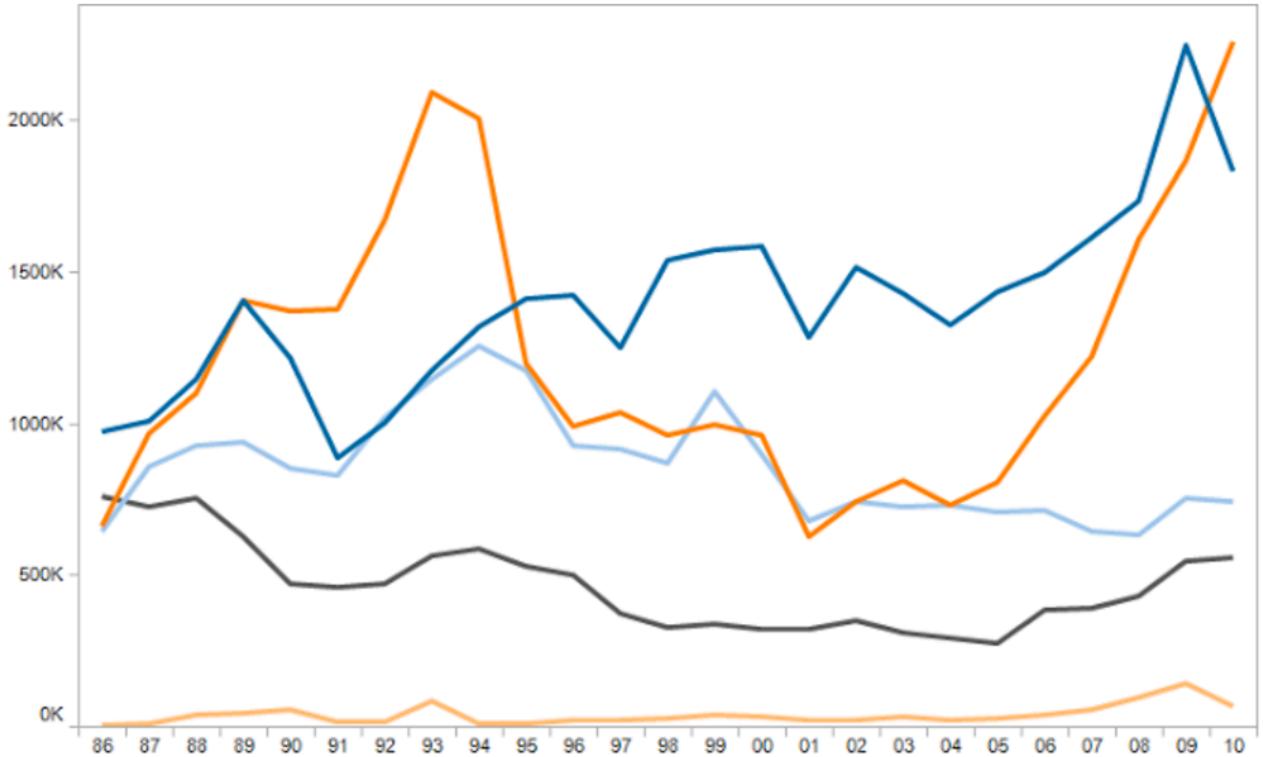


# Idioms: heatmap & cluster heatmap

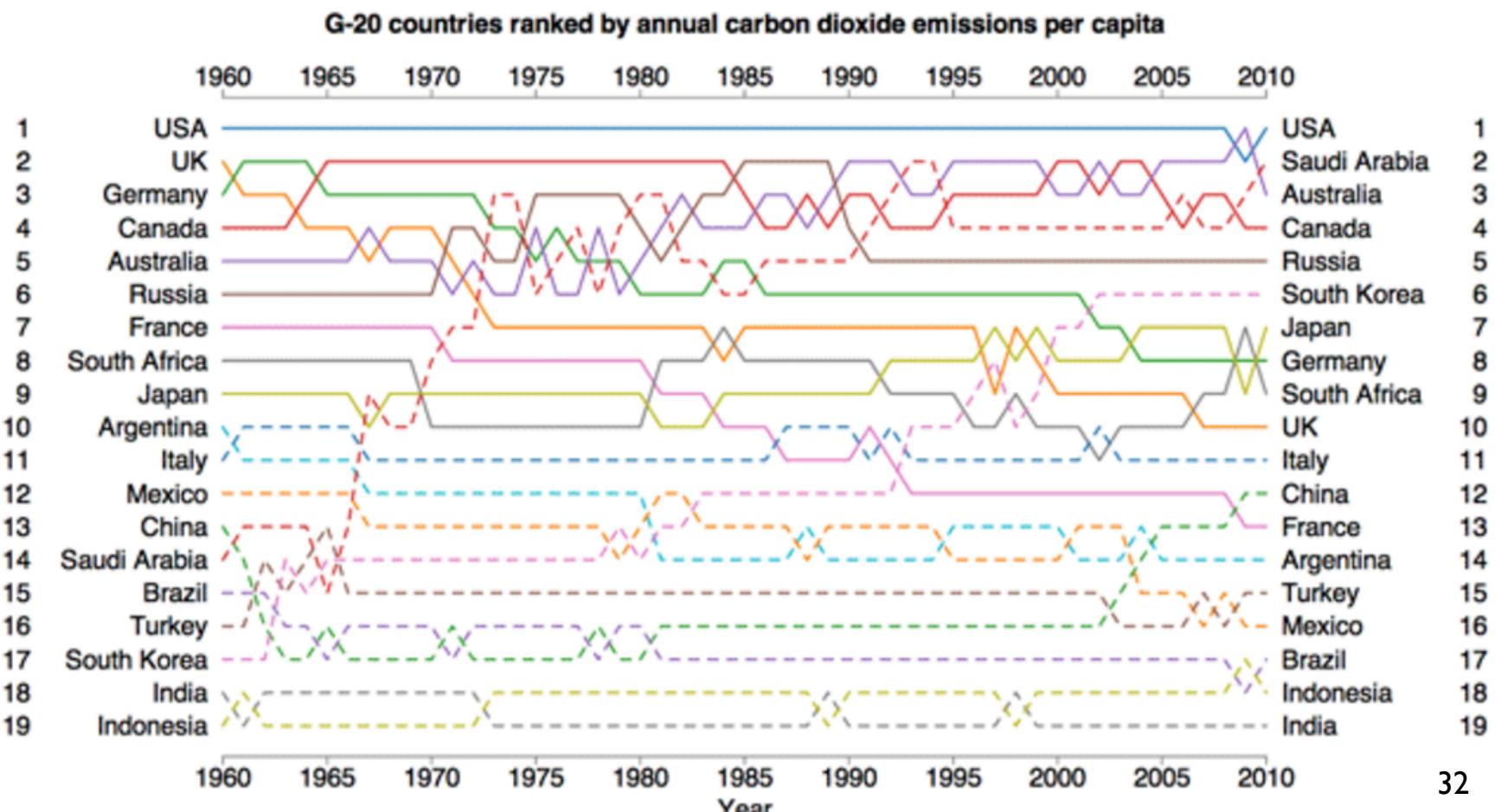
Idiom	Heatmaps
What: Data	Table: two categorical key attributes (genes, conditions), one quantitative value attribute (activity level for gene in condition).
How: Encode	2D matrix alignment of area marks, diverging color-map.
Why: Task	Find clusters, outliers; summarize.
Scale	Items: one million. Categorical attribute levels: hundreds. Quantitative attribute levels: 3–11.
Idiom	Cluster Heatmaps
What: Derived	Two cluster hierarchies for table rows and columns.
How: Encode	Heatmap: 2D matrix alignment, ordered by both cluster hierarchies. Dendrogram: connection line marks for parent–child relationships in tree.

# Two keys

- one categorical + one quantitative
- one value quantitative or ordinal
- examples: time series data
  - value measured per item & time: *multiple line chart*
  - rank measured per item & time: *bump chart*

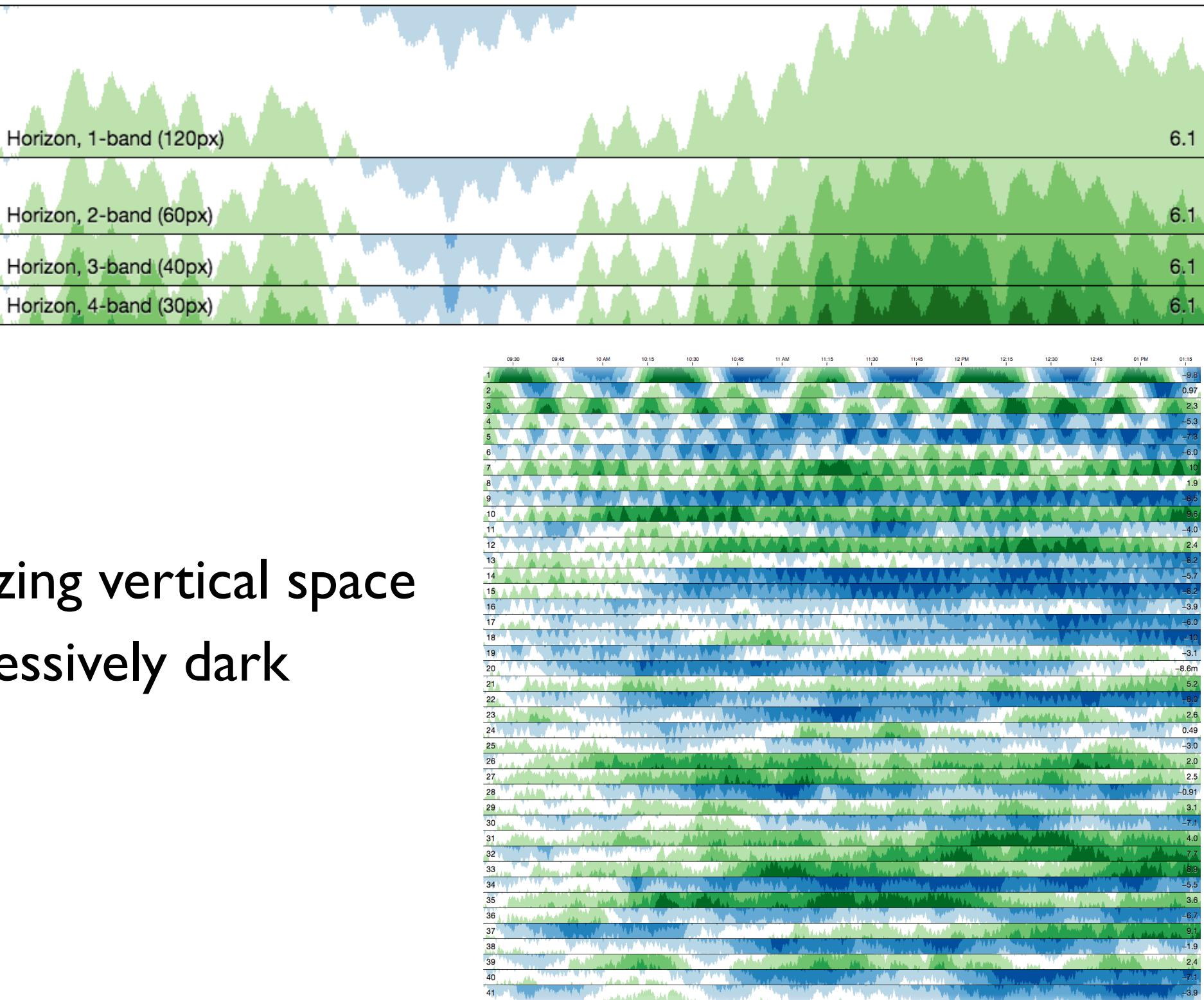


- mark:
  - one polyline per item
- scalability:
  - dozens of items
  - multiple line charts: continuous scales
  - bump chart: dozens rank levels



# Multiple time series - Idiom: **horizon chart**

- Challenges:
  - many items/series/signals
  - different scale for each item
  - cannot show in common frame!
- Horizon chart:
  - variant of area chart
  - preserves resolution when squeezing vertical space
  - larger values over plotted in successively dark colors
  - negative values descend from top

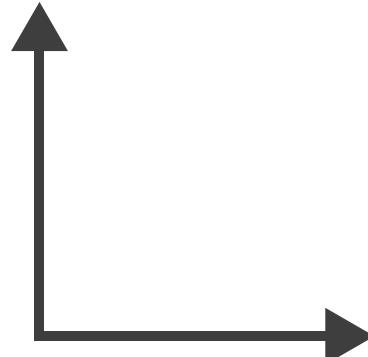


# Multi-dimensional data

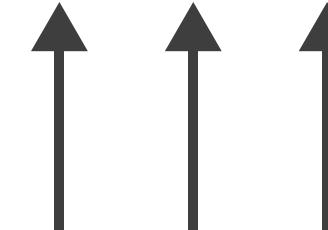
- Several quantitative/ordered attributes per item
  - Each item can be seen as a vector in a multi-dimensional space
  - Vis space has just two dimensions
  - How to arrange data?

## → Axis Orientation

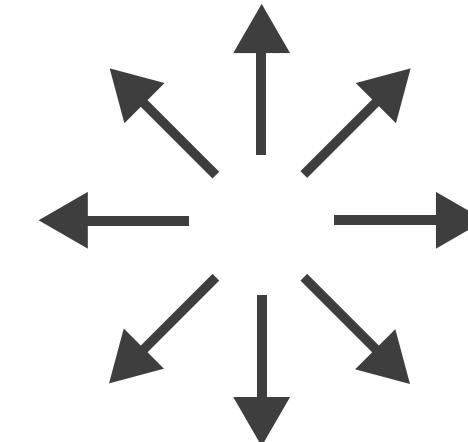
→ Rectilinear



→ Parallel

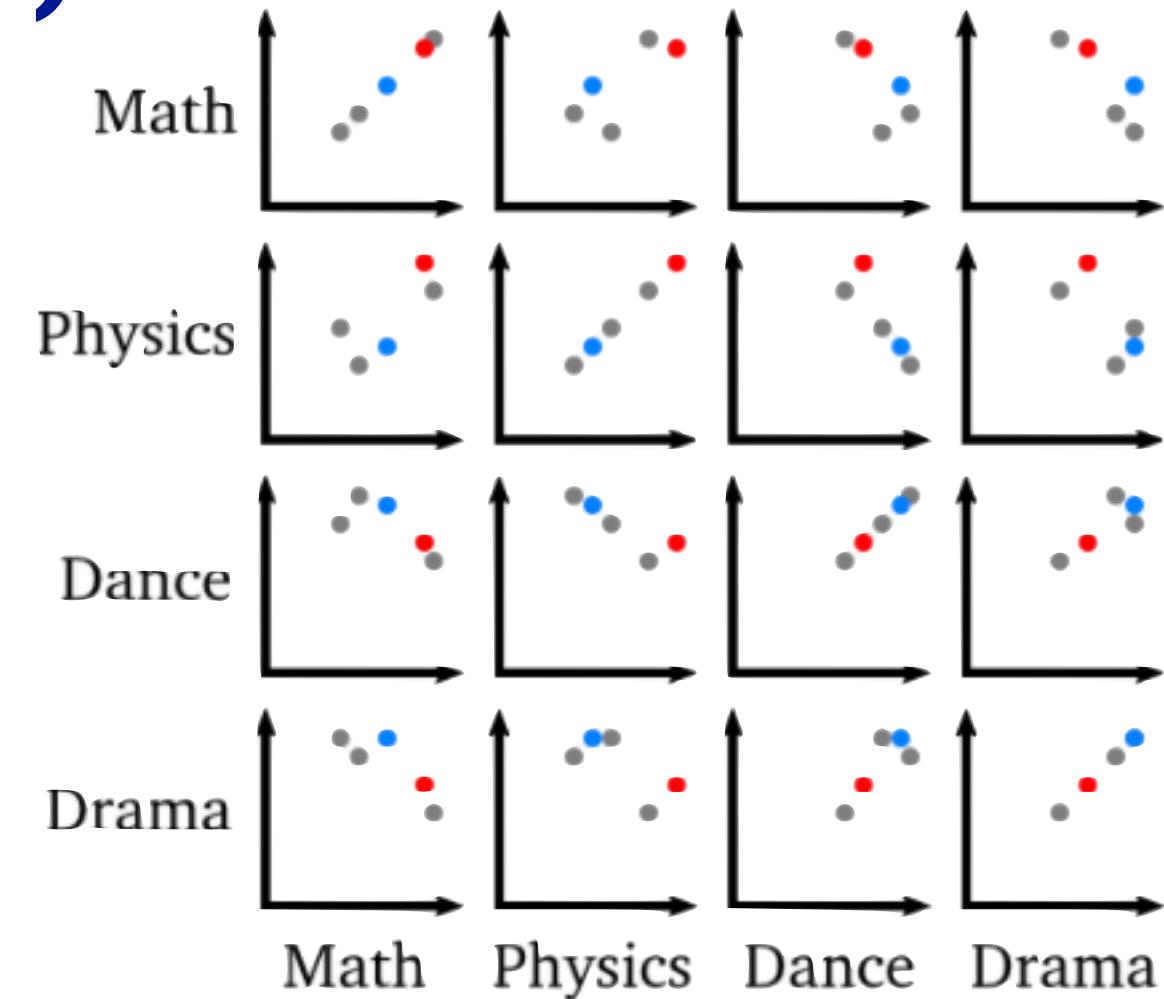


→ Radial



# Idiom: scatterplot matrix (SPLOM)

- several dependent values per item
  - rectilinear axes, point mark
  - all possible pairs of axes
  - each entry in a matrix is a scatterplot
    - projection on a 2D hyperplane
  - tasks
    - find trends & outliers
    - find correlation & similarity
  - scalability
    - one dozen attributes
    - dozens to hundreds of items



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

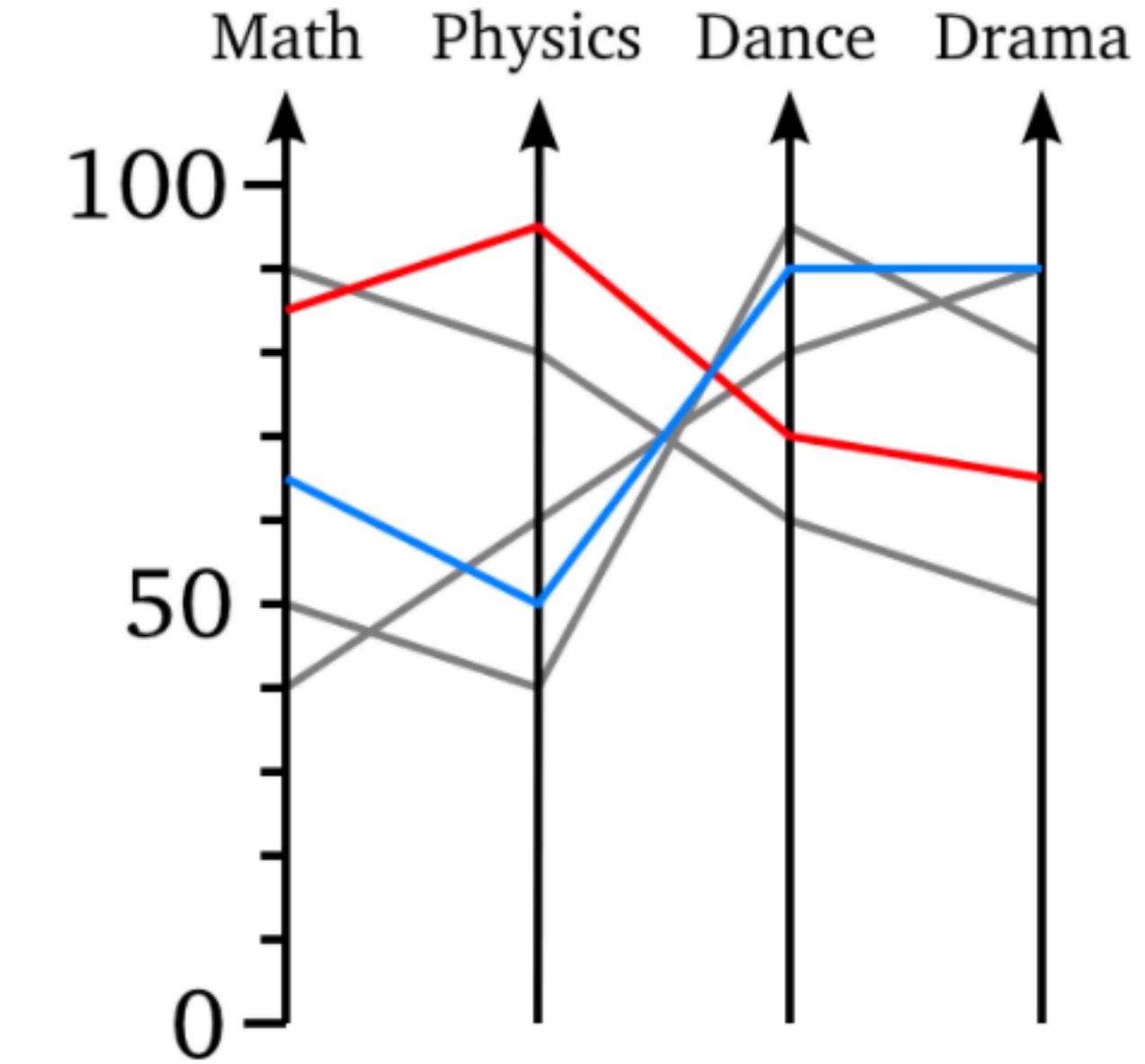
after [Visualization Course Figures. McGuffin, 2014. <http://www.michaelmcguffin.com/courses/vis/>]

# Idiom: scatterplot matrix (SPLOM)

Idiom	Scatterplot Matrix (SPLOM)
What: Data	Table.
What: Derived	Ordered key attribute: list of original attributes.
How: Encode	Scatterplots in 2D matrix alignment.
Why: Task	Find correlation, trends, outliers.
Scale	Attributes: one dozen. Items: dozens to hundreds.

# Idiom: parallel coordinates

- several dependent values per item
  - parallel axes, jagged line representing item
  - rectilinear axes, value of attribute per item as point
    - axis ordering is major challenge
- tasks
  - overview of attributes
  - find range of attributes
  - browse & explore
  - select based on attributes
- scalability
  - dozens of attrs
  - hundreds of items

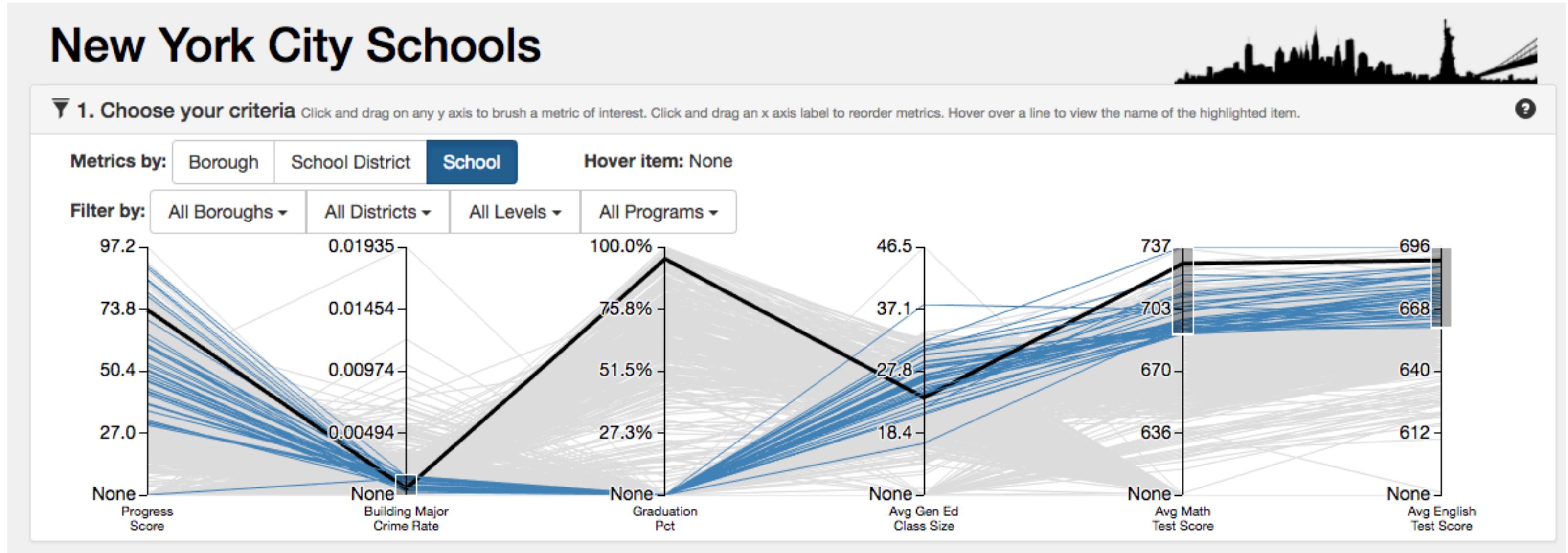


	Math	Physics	Dance	Drama
1	85	95	70	65
2	90	80	60	50
3	65	50	90	90
4	50	40	95	80
5	40	60	80	90

after [Visualization Course Figures. McGuffin, 2014. <http://www.michaelmcguffin.com/courses/vis/>]

# Idiom: parallel coordinates

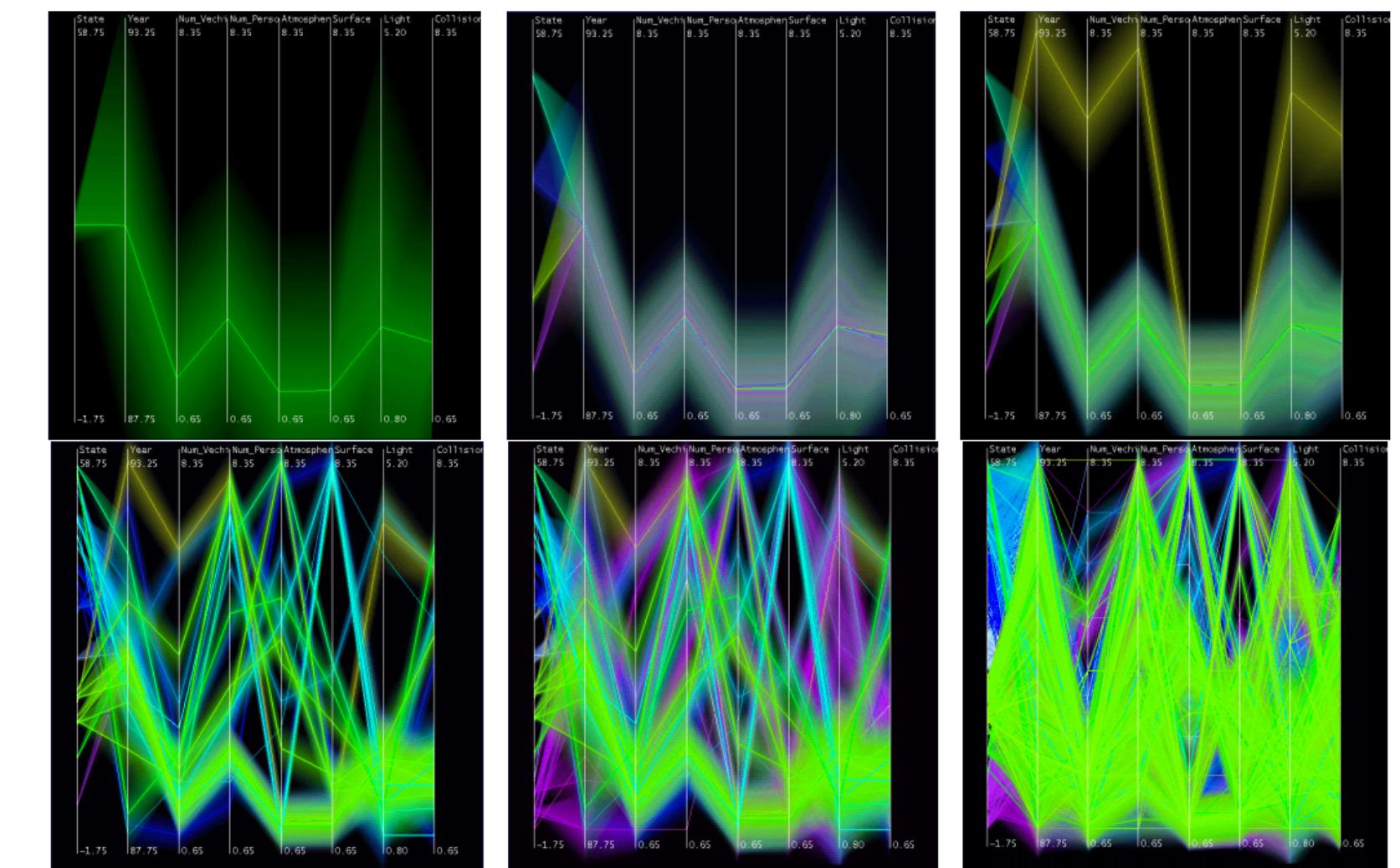
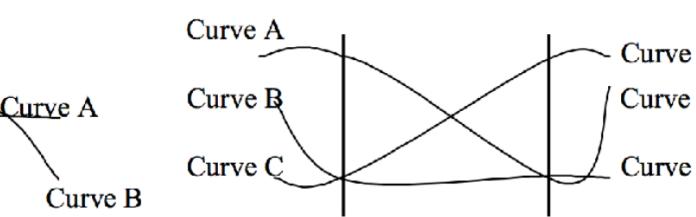
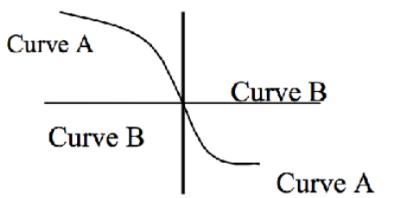
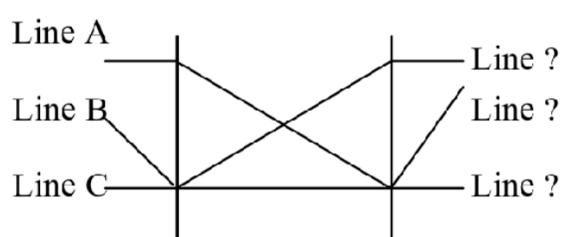
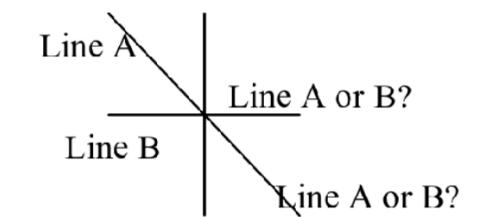
- Example



[itisaasta.com/nycs/nycs.html](http://itisaasta.com/nycs/nycs.html)

# Idiom: parallel coordinates

- Limitations
  - Correlation only between adjacent axes
    - Solution: brushing & reorder
  - Ambiguity
    - Solution: curves, brushing, selection
  - Scalability to many items
    - Solution: brushing, selection, filtering
    - Solution: hierarchy by clustering



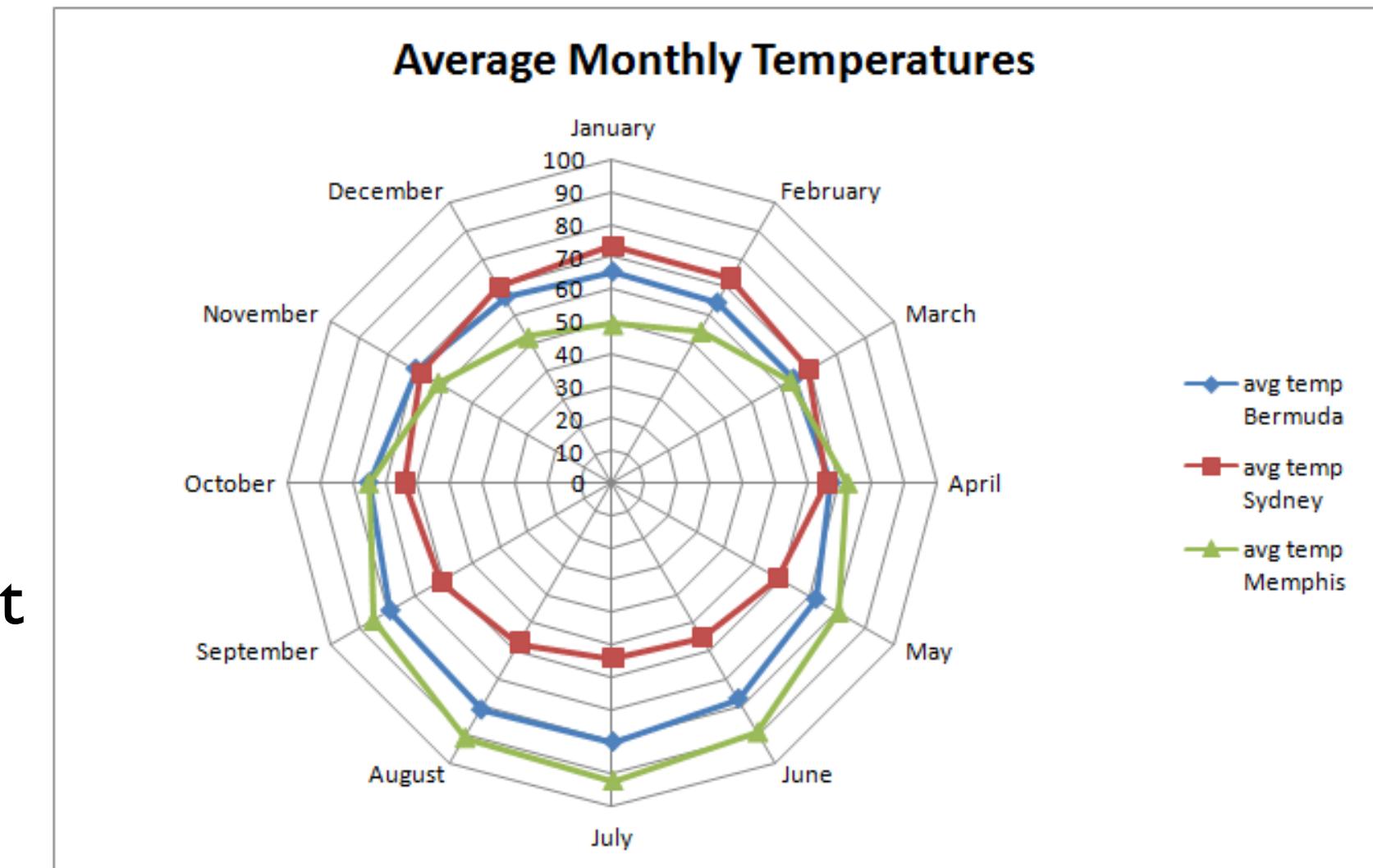
# Idiom: parallel coordinates

Idiom	Parallel Coordinates
What: Data	Table: many value attributes.
How: Encode	Parallel layout: horizontal spatial position used to separate axes, vertical spatial position used to express value along each aligned axis with connection line marks as segments between them.
Why: Tasks	Find trends, outliers, extremes, correlation.
Scale	Attributes: dozens along secondary axis. Items: hundreds.

# Idiom: **radar chart** (a.k.a **spider chart**)

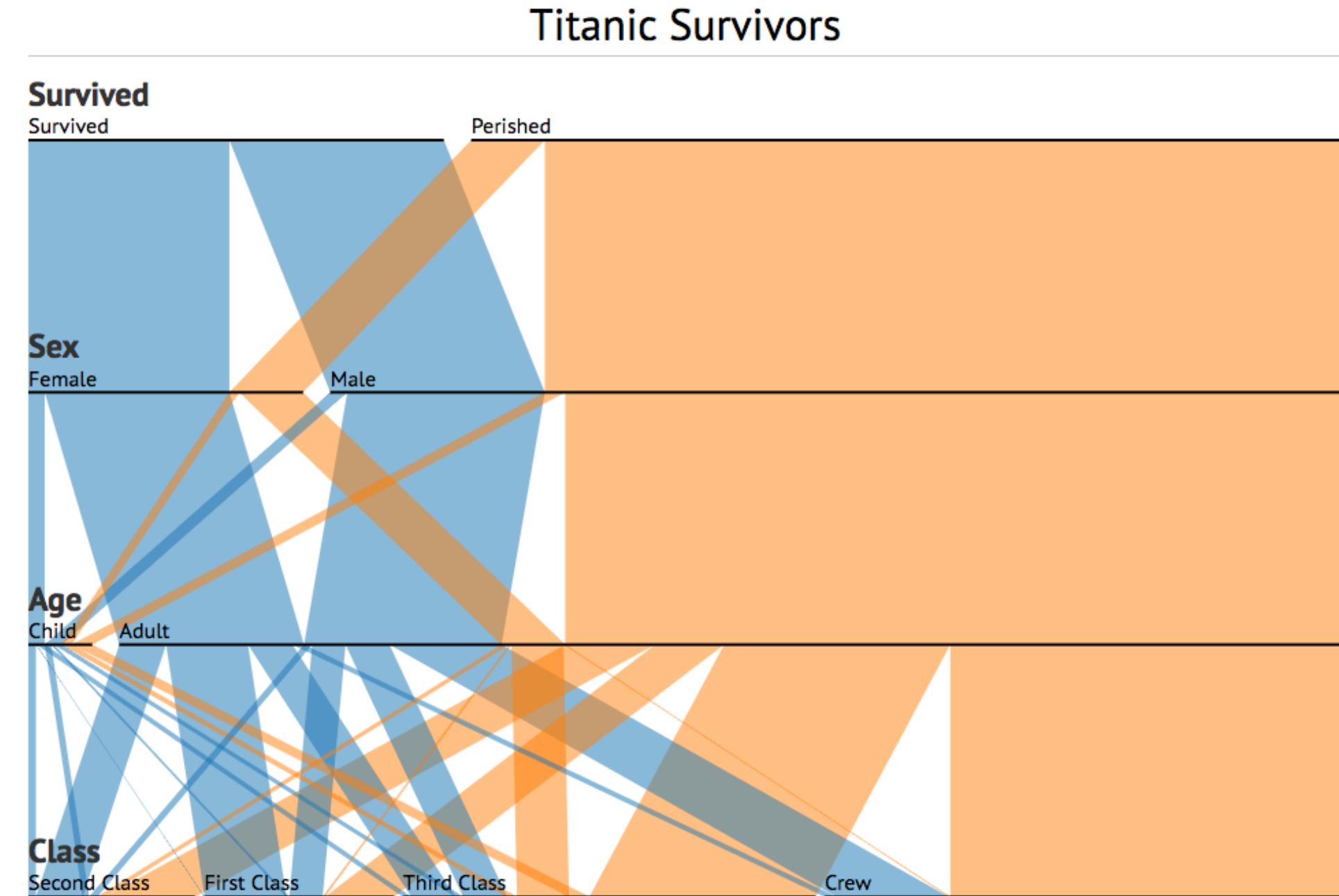
- Same as parallel coordinates, but radial axis orientation
  - when categories are radially arranged
  - when categories give uniform scores/costs

- Pros:
  - make cyclic structure explicit
  - internal area provides overall score/cost
    - integral property
- Cons:
  - harder to read



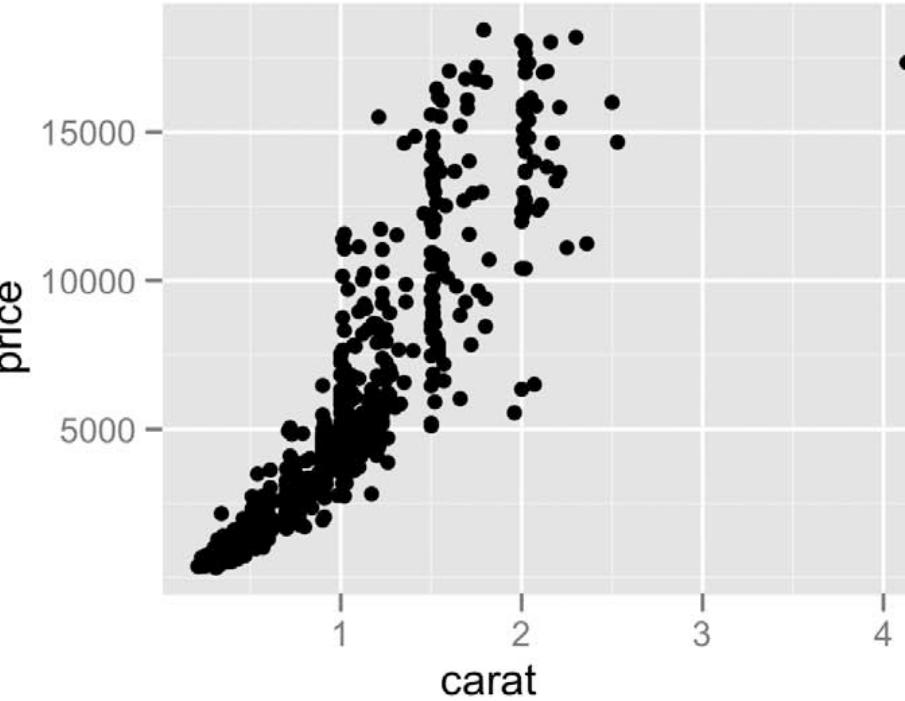
# Idiom: parallel sets

- builds on parallel coordinates to better handle categorical data
  - discrete
  - small number of values
  - no implied ordering
- task:
  - find relationships
- interaction driven technique

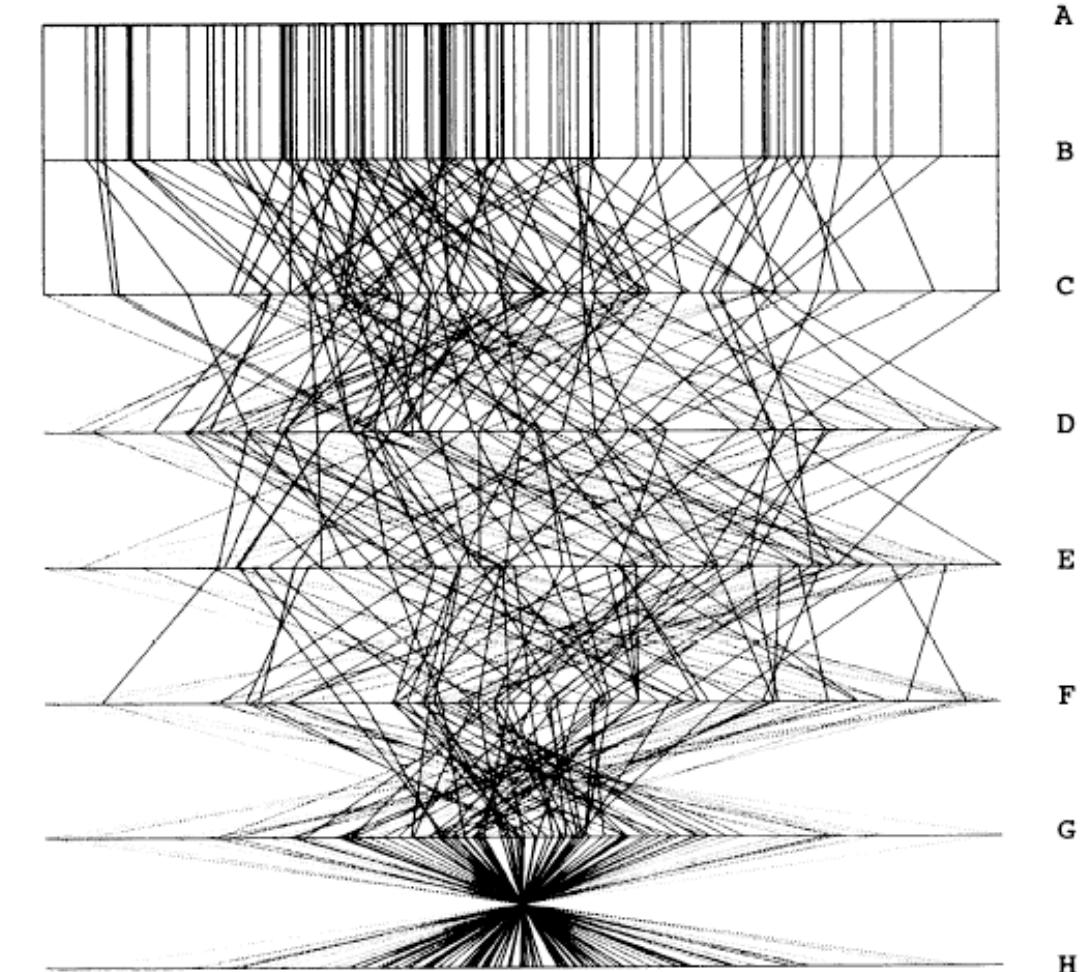
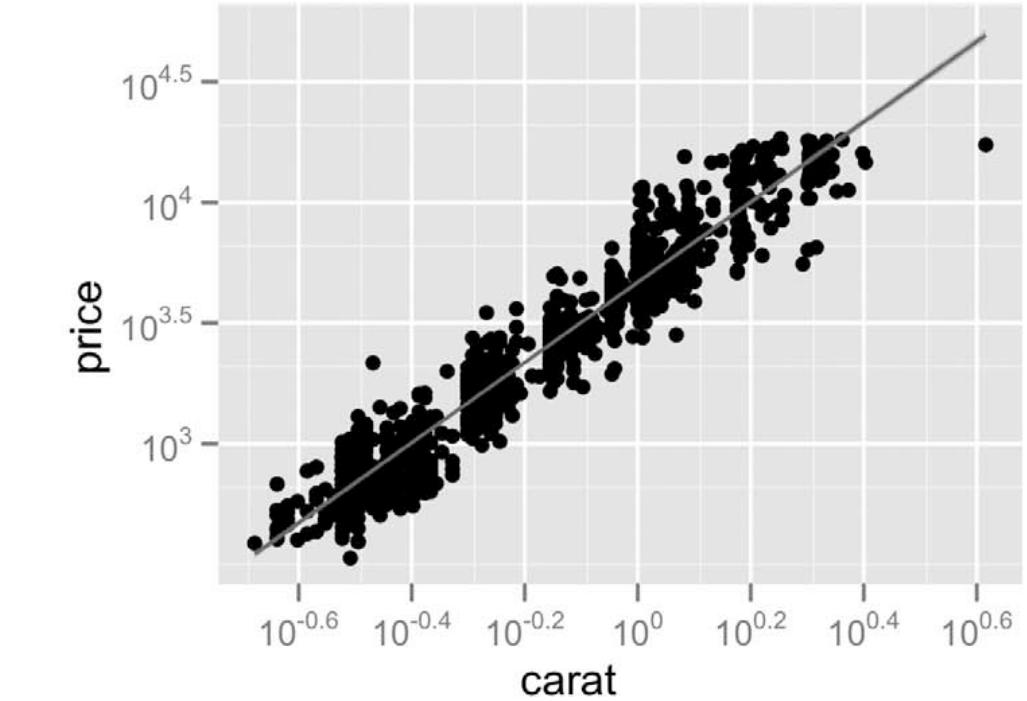


# Task: Correlation

- 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.]

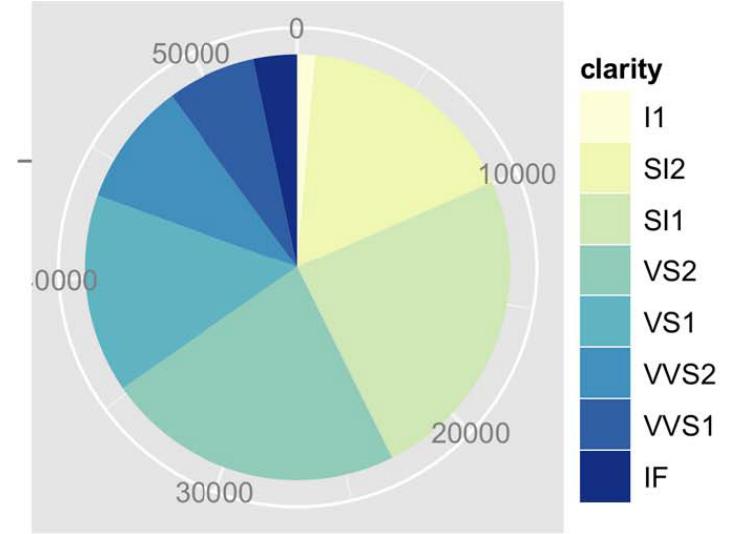


[Hyperdimensional Data Analysis Using Parallel Coordinates. Wegman. Journ. American Statistical Association 85:411 (1990), 664–675.]

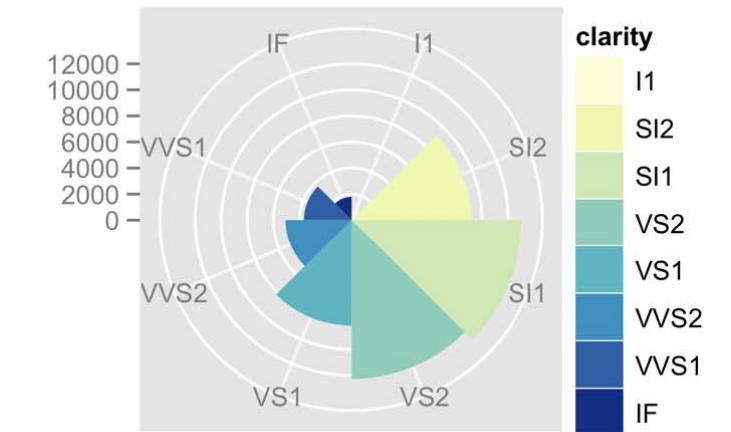
Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of  $\rho = 1, .8, .2, 0, -.2, -.8$ , and  $-1$ .

# Idioms: pie chart, polar area chart

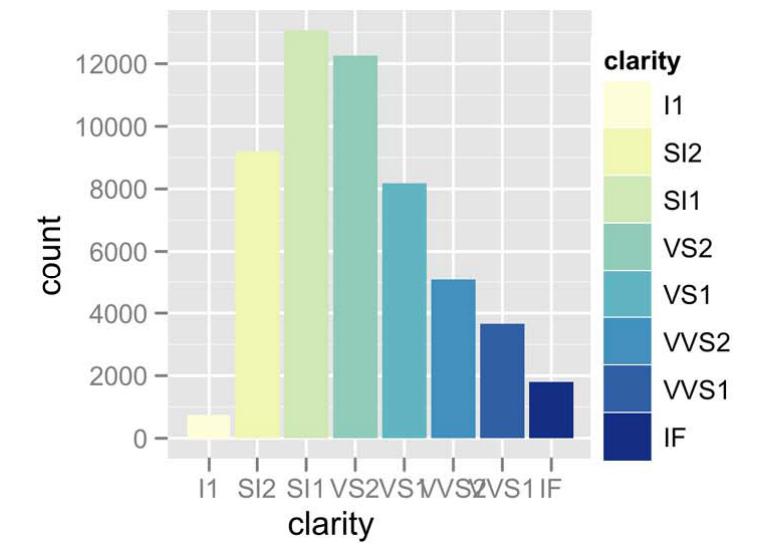
- pie chart
  - area marks with angle channel
  - accuracy: angle/area much less accurate than line length



- polar area chart
  - area marks with length channel
  - more direct analog to bar charts

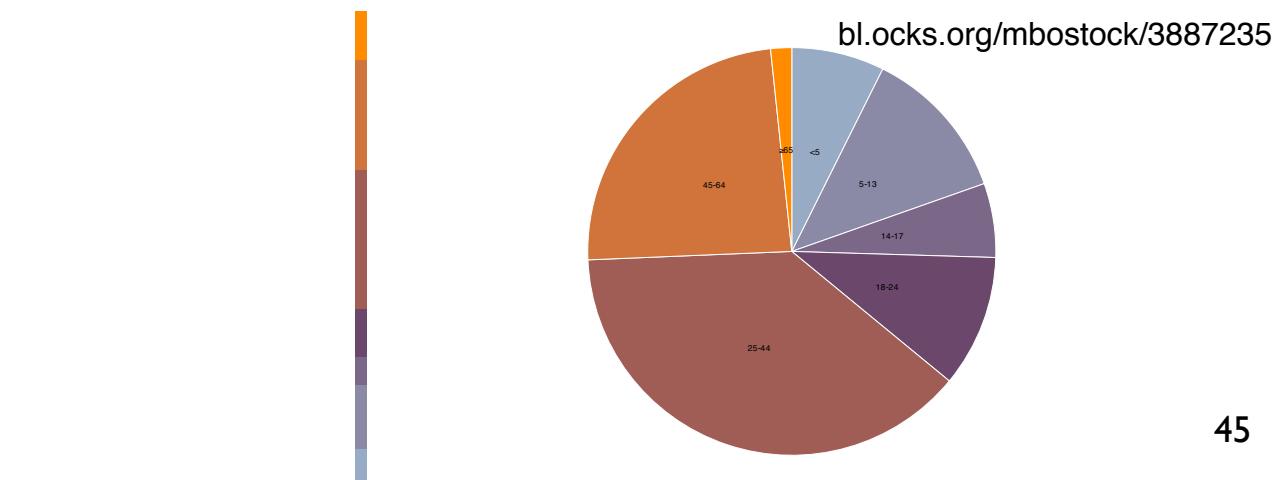
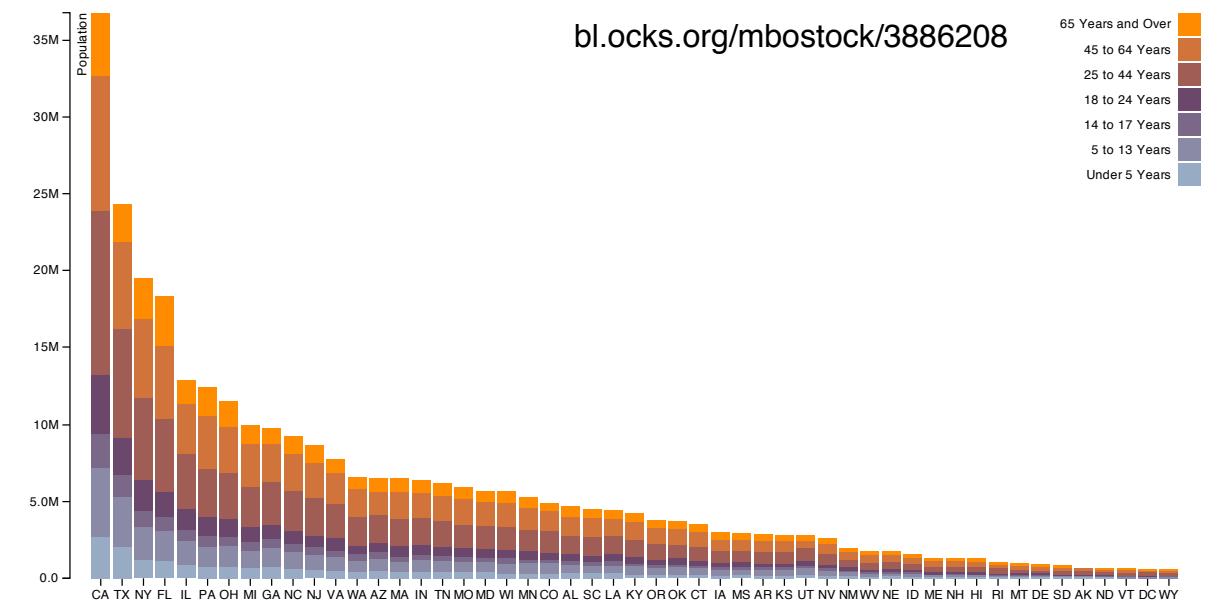
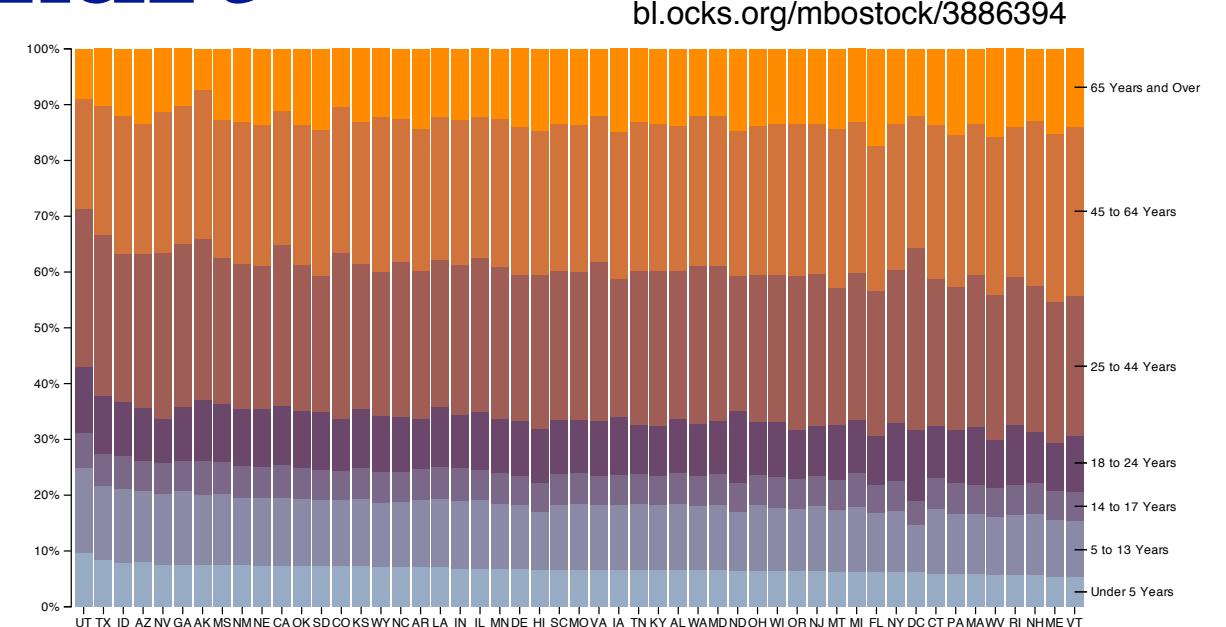


- data
  - 1 categor key attrib, 1 quant value attrib
- task
  - part-to-whole judgements

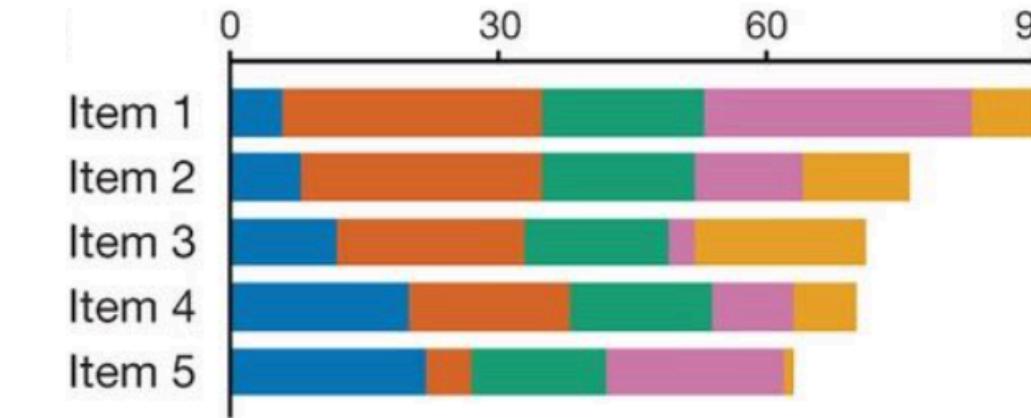
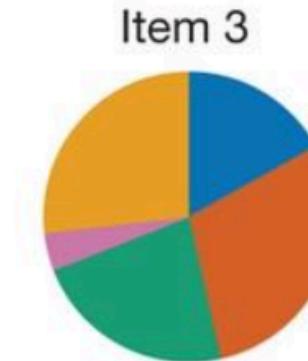


# Idioms: normalized stacked bar chart

- 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



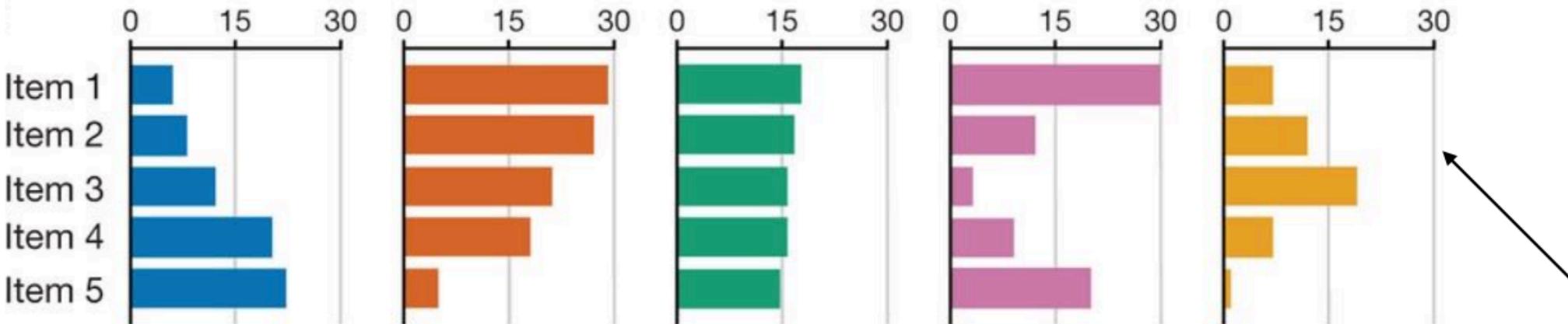
# Pie charts vs bar chart variants



Pie Chart

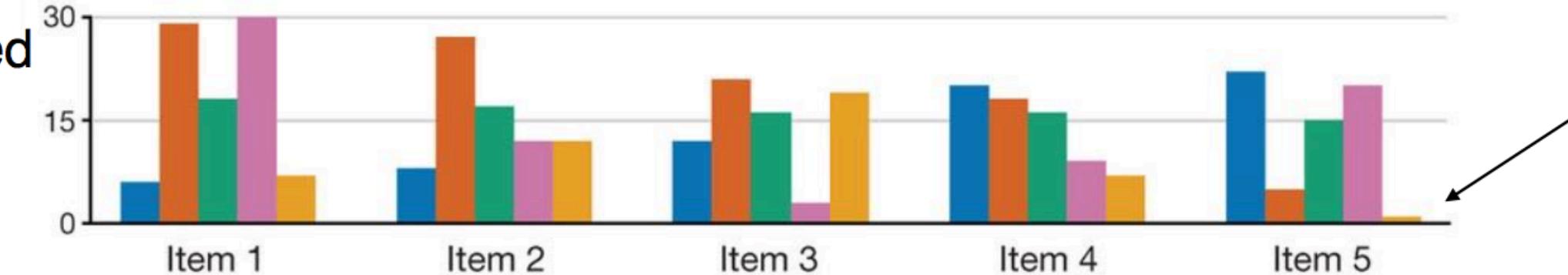
Stacked bar chart

Layered Bar Chart



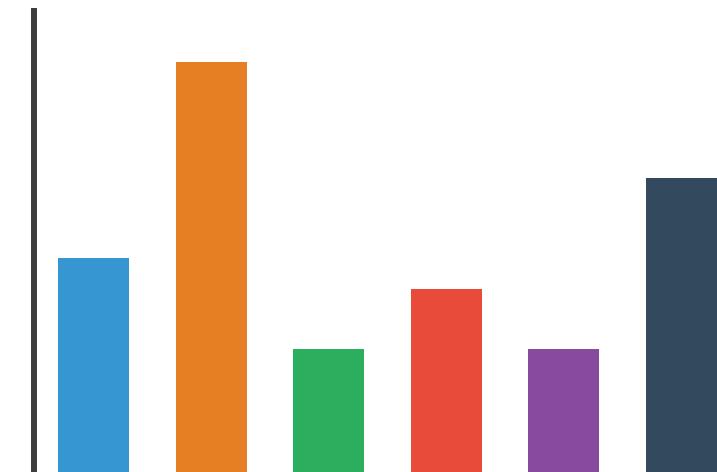
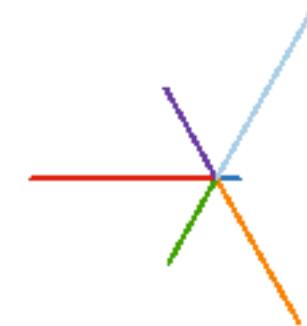
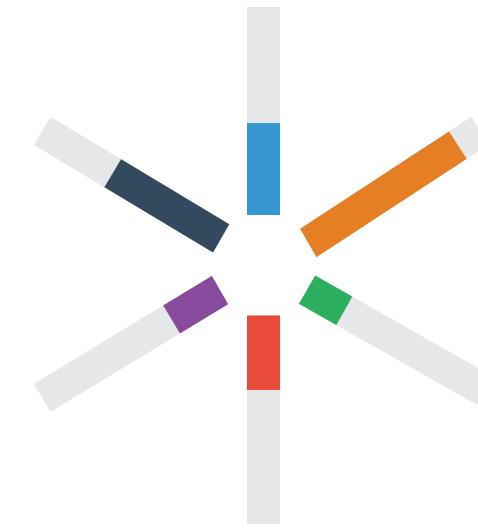
Small Multiples

Grouped Bar Chart



# Idioms: **radial bar chart**, **star plot**

- radial bar chart
  - radial axes meet at central ring, line mark
- star plot
  - radial axes, meet at central point, line mark
- bar chart
  - rectilinear axes, aligned vertically
- accuracy
  - length unaligned with radial
    - less accurate than aligned with rectilinear

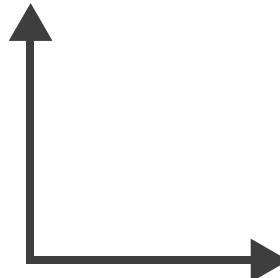


# Orientation limitations

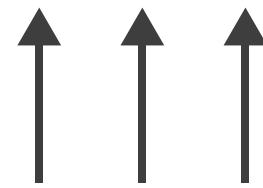
- rectilinear: scalability wrt # axes
  - 2 axes best
  - 3 problematic
  - 4+ impossible
- parallel: unfamiliarity, training time
- radial: perceptual limits
  - angles lower precision than lengths
  - asymmetry between angle and length
    - can be exploited!

→ Axis Orientation

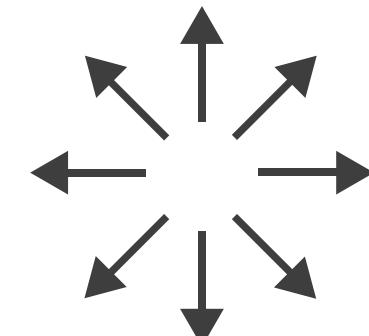
→ Rectilinear



→ Parallel



→ Radial



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

## Next Time

- to read
  - VAD Ch. 8: Arrange Spatial Data (only 8.1-8.3 included)