CS49000-VIZ - Fall 2020

Introduction to Data Visualization

Tables

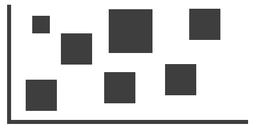
Lecture 8

September 21, 2020

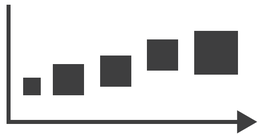
Express Values

Separate, Order, Align Regions

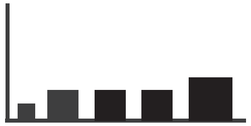
- Separate



- Order

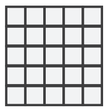


- Align



- *1 Key*
  List

- *2 Keys*
  Matrix

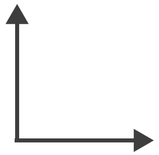


- *3 Keys*
  Volume

- *Many Keys*
  Recursive Subdivision

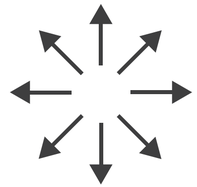
Axis Orientation

- Rectilinear



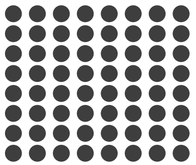
- Parallel

- Radial



Layout Density

- Dense



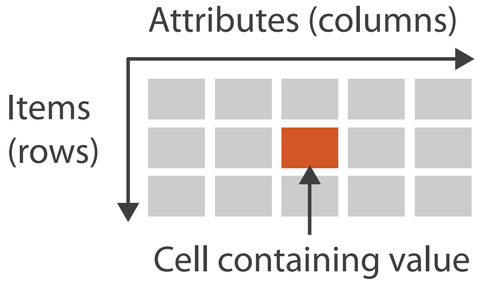
- Space-Filling

Keys and values

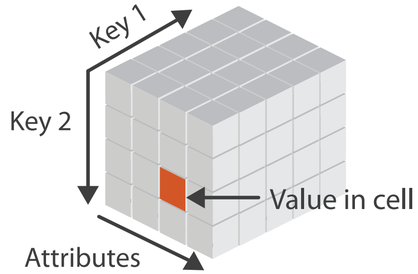
key

- independent attribute
- used as unique index to look up items
- simple tables: 1 key
- multidimensional tables: multiple keys

➔ Tables



➔ *Multidimensional Table*



Keys and values

value

- dependent attribute, value of cell.

Express Values

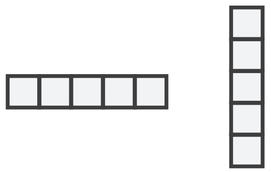
Keys and values

classify arrangements by key count

- 0, 1, 2, many...

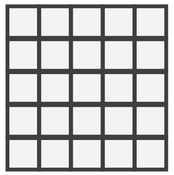
➔ *1 Key*

List



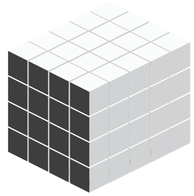
➔ *2 Keys*

Matrix



➔ *3 Keys*

Volume

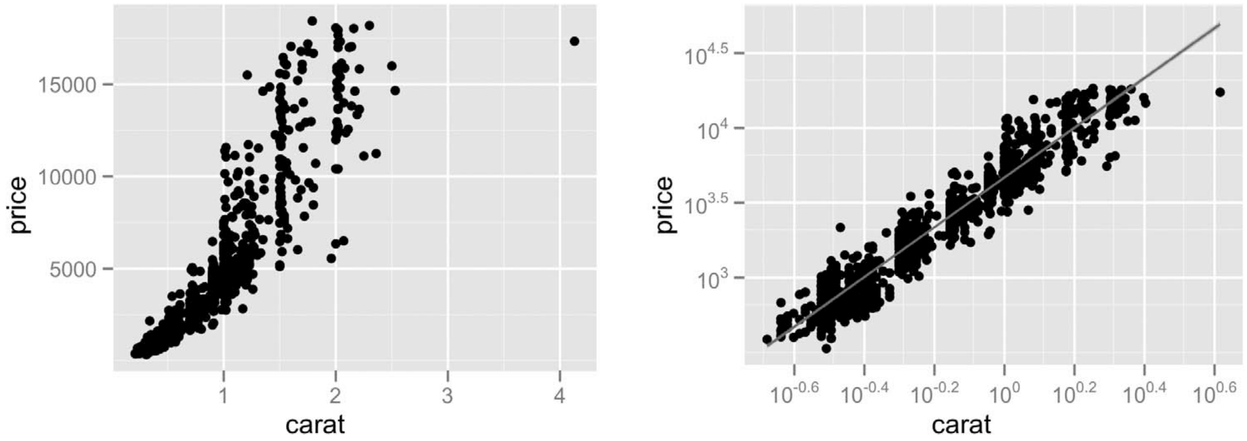


➔ *Many Keys*

Recursive Subdivision

Idiom: scatterplot

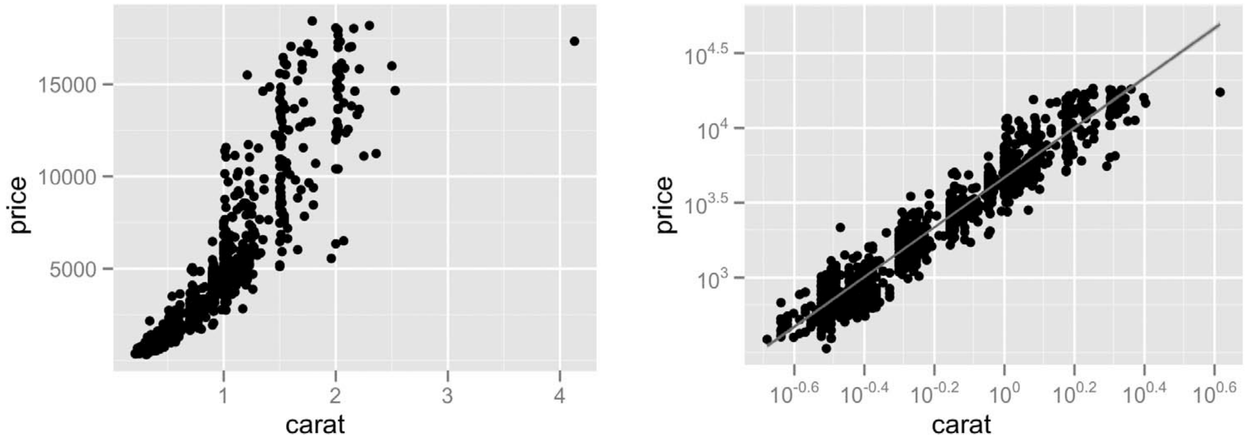
- ***express*** values
  - quantitative attributes



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

Idiom: scatterplot

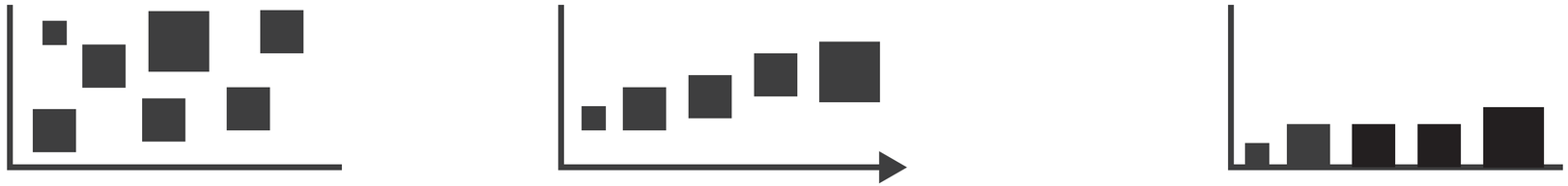
- ***express*** values
  - quantitative attributes
- no keys, only values
  - **data:** 2 quant attributes
  - **mark:** points
  - **channels:** horizontal + vertical position
  - **tasks:** find trends, outliers, distribution, correlation, clusters
  - **scalability:** hundreds of items



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

Some keys: Categorical regions

➔ Separate ➔ Order ➔ Align



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

➔ *1 Key* *List*

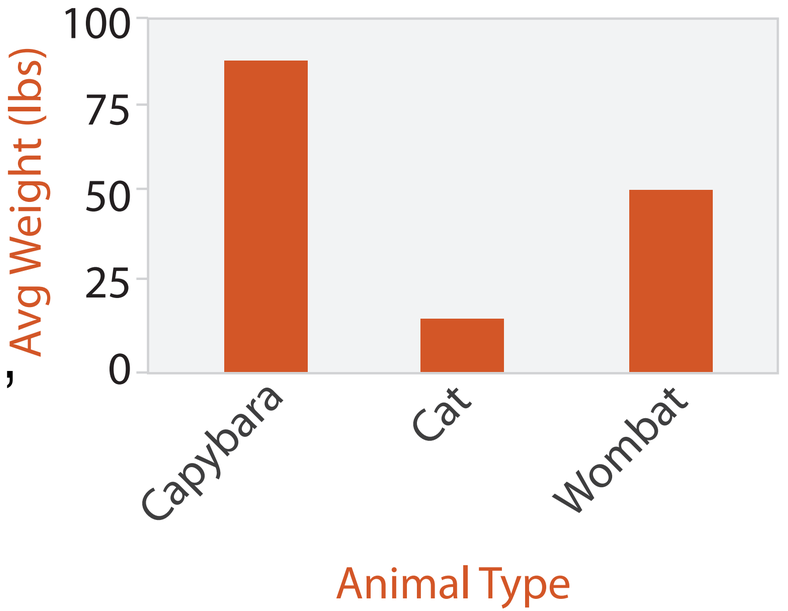
➔ *2 Keys* *Matrix*

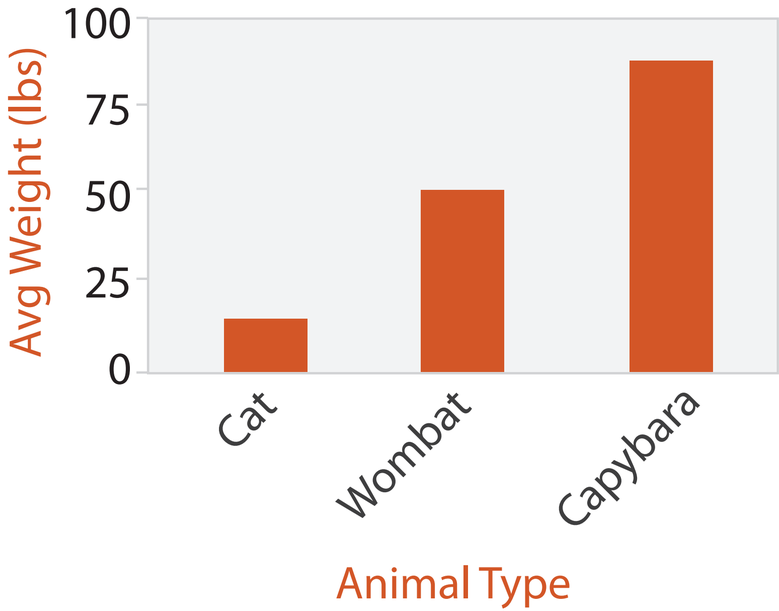
➔ *3 Keys* *Volume*

➔ *Many Keys* *Recursive Subdivision*

Bar Chart

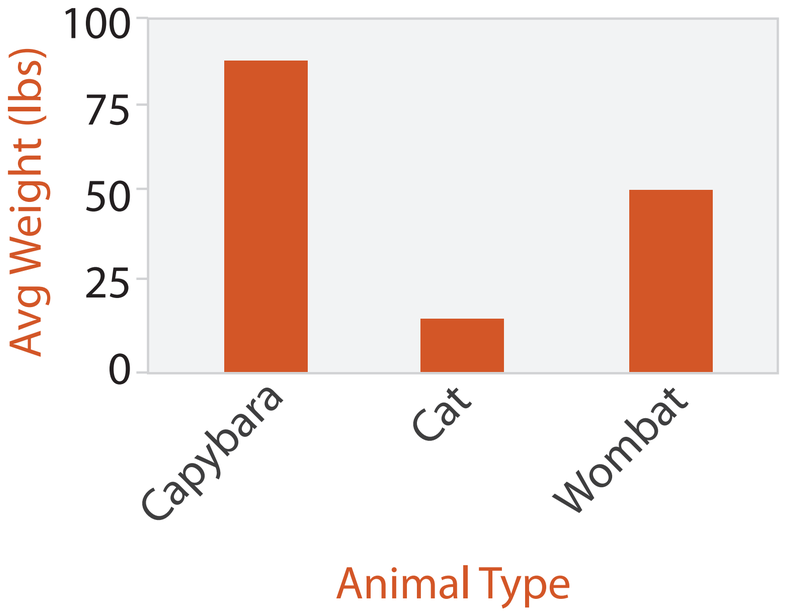
- 1key,1value
  - data:
    - 1 categ. attribute,
    - 1 quant. attribute

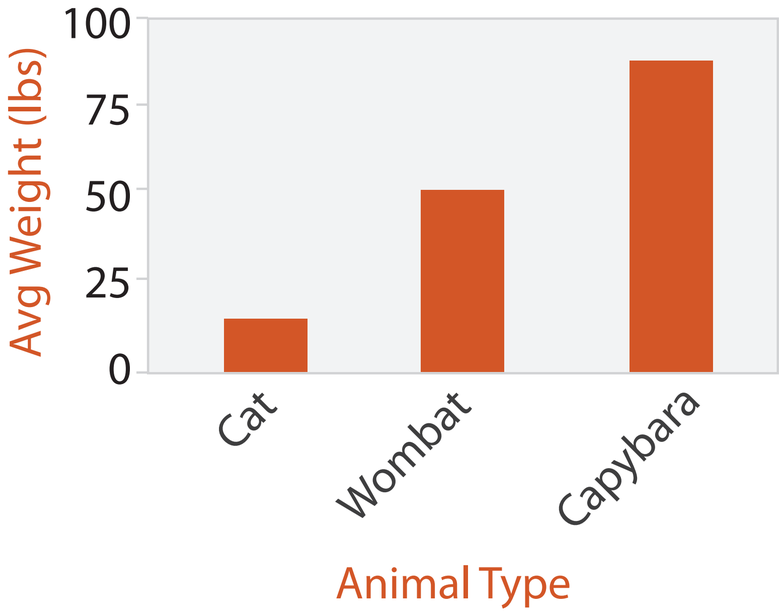




Bar Chart

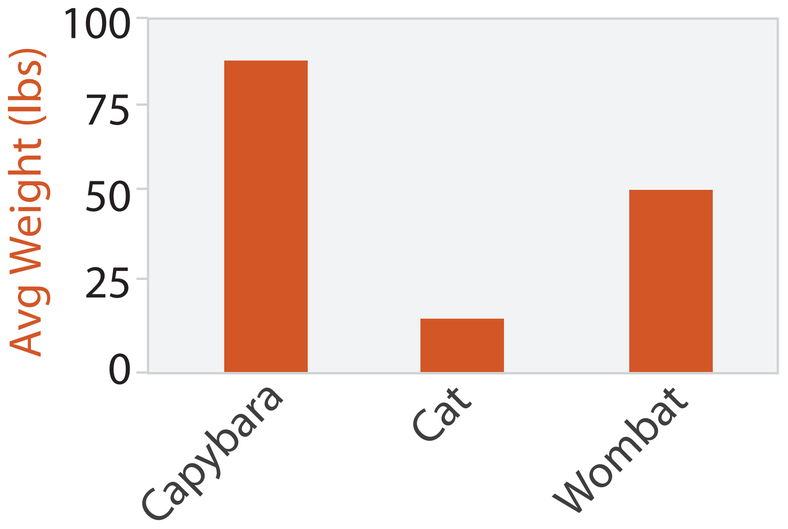
- 1key,1value
  - mark: lines

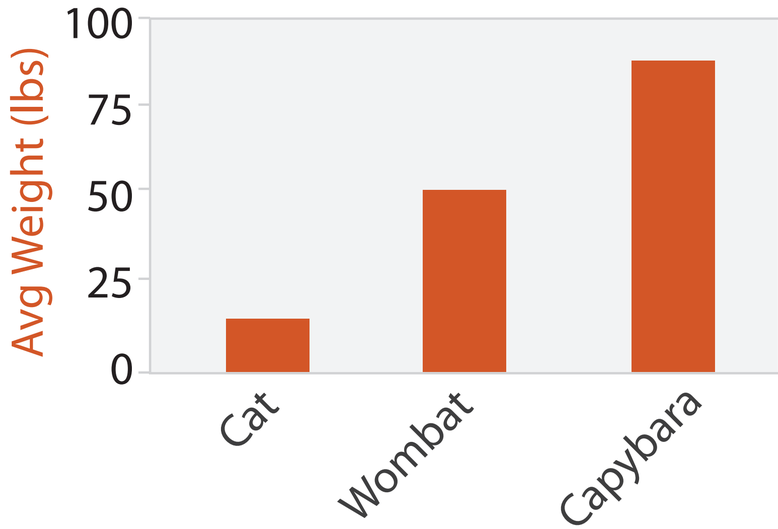




bar chart

- 1key,1value
  - channels

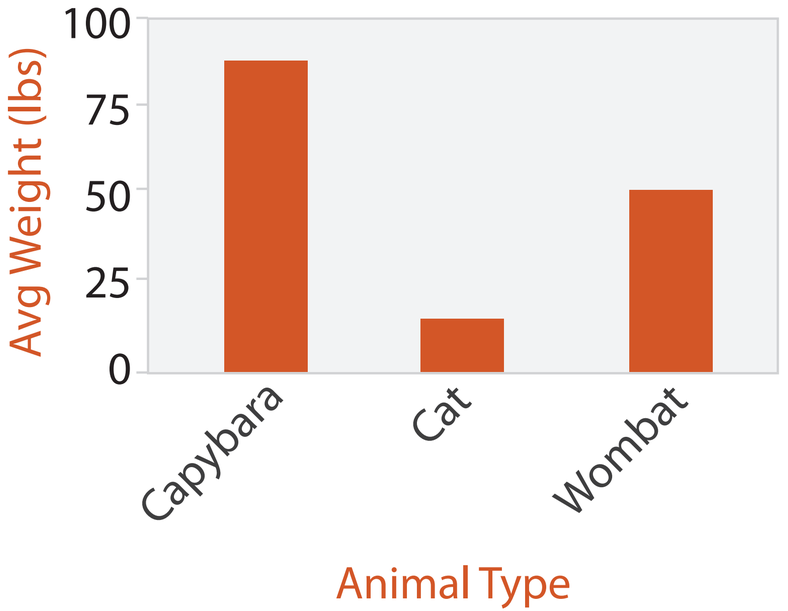


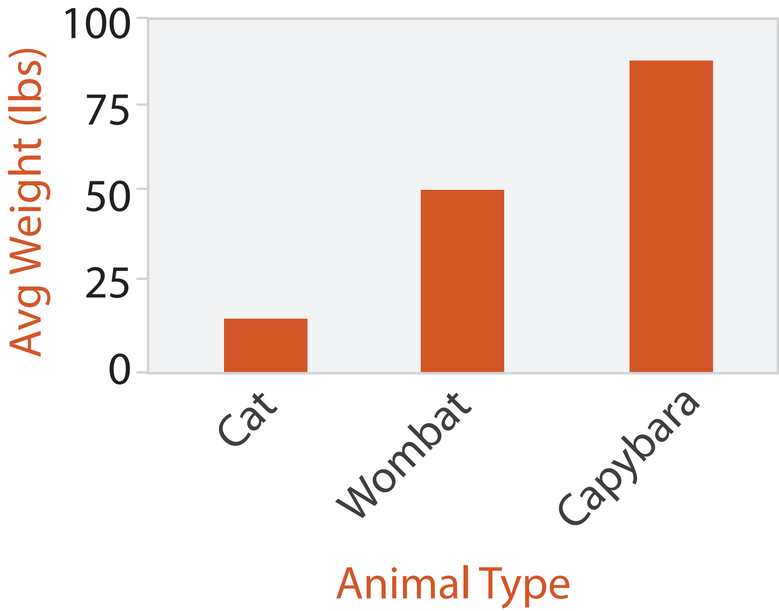


- length to express quant value
- spatial regions: one per mark
  - separated horizontally, aligned vertically
  - ordered by quant attrib
    - by label (alphabetical), by length attrib (data-driven)

Bar Chart

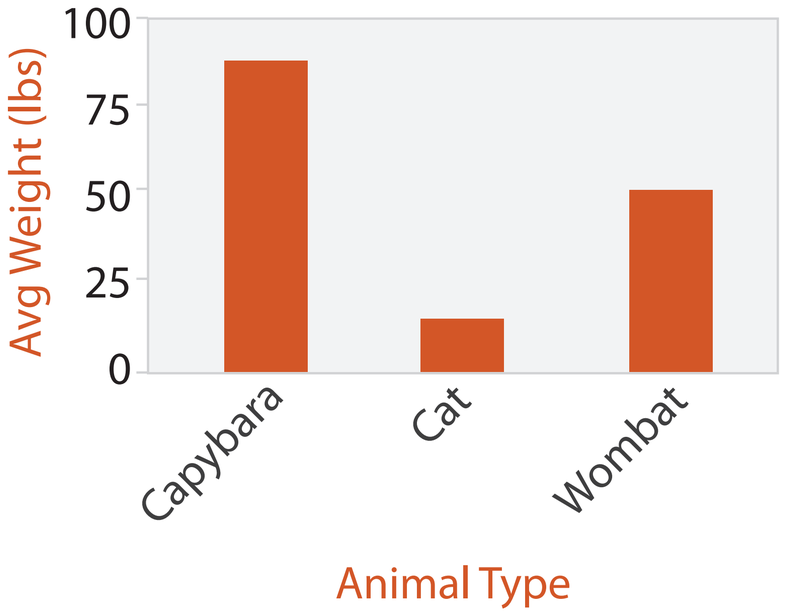
- 1key,1value
  - task
    - compare, lookup values

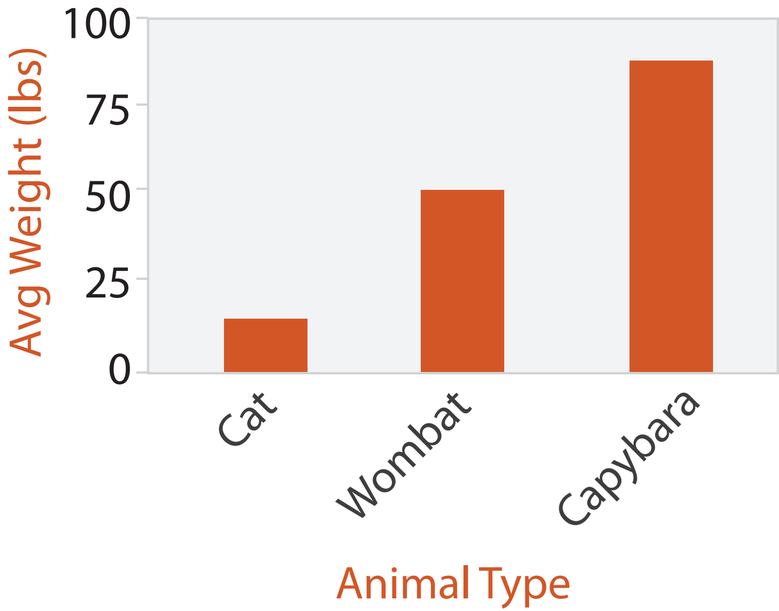




Bar Chart

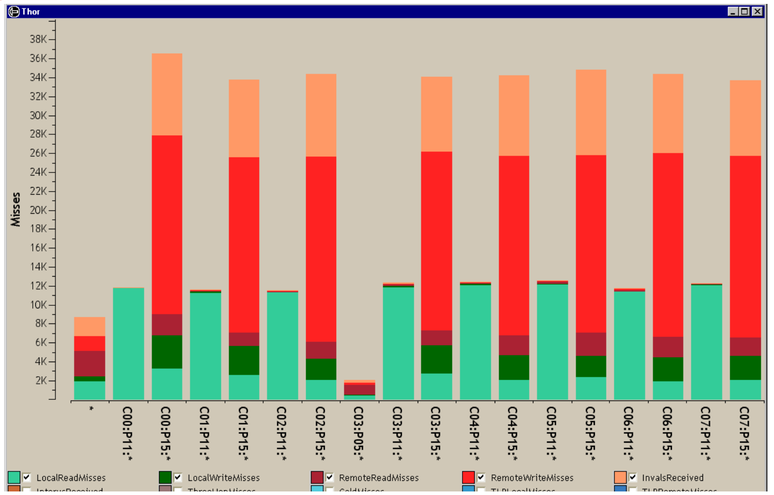
- 1key,1value
  - scalability
    - dozens to hundreds of levels for key attrib





Stacked Bar Chart

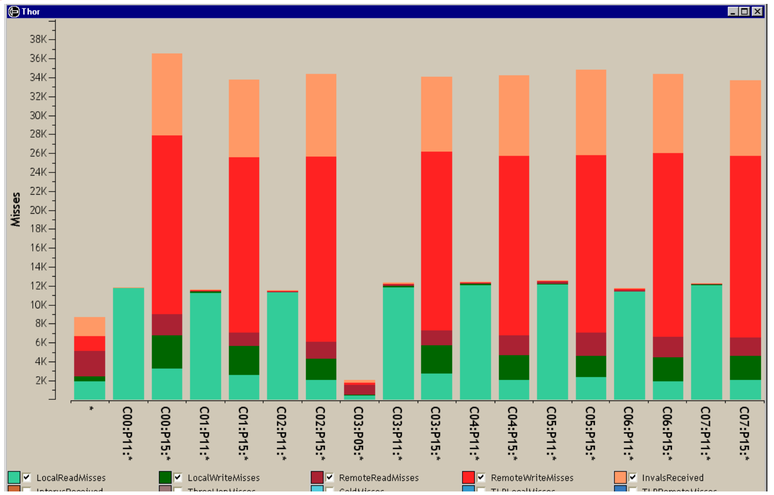
- one more key
  - data
    - 2 categ attrib, 1 quant attrib



[Using Visualization to Understand the Behavior of Computer Systems. Bosch. Ph.D. thesis, Stanford Computer Science, 2001.]

Stacked Bar Chart

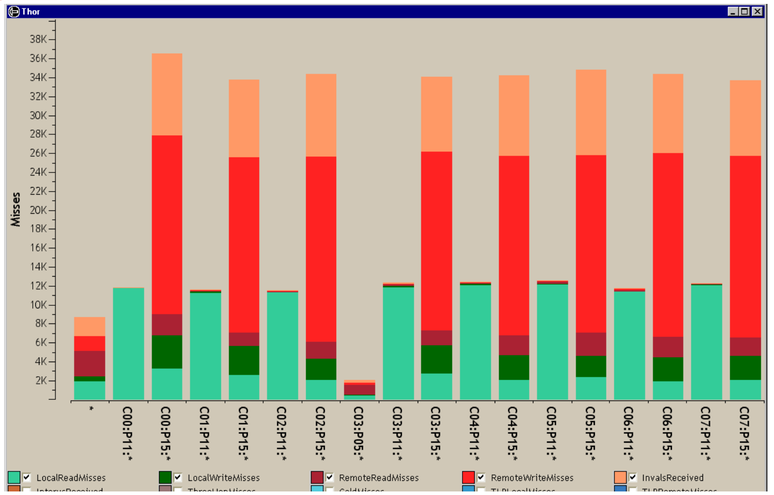
- one more key
  - mark: vertical stack of line marks
    - **glyph**: composite object, internal structure from multiple marks



[Using Visualization to Understand the Behavior of Computer Systems. Bosch. Ph.D. thesis, Stanford Computer Science, 2001.]

Stacked Bar Chart

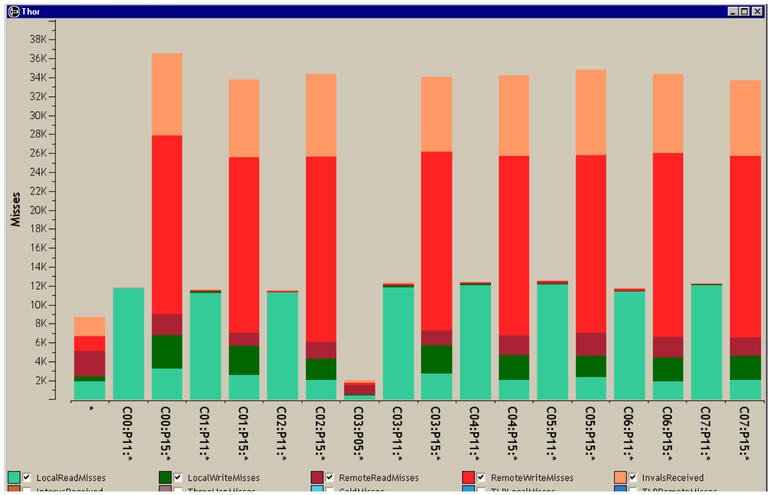
- one more key
  - channels
    - length and color hue
    - spatial regions: one per glyph
      - aligned: full glyph, lowest bar component
      - unaligned: other bar component



[Using Visualization to Understand the Behavior of Computer Systems. Bosch. Ph.D. thesis, Stanford Computer Science, 2001.]

Stacked Bar Chart

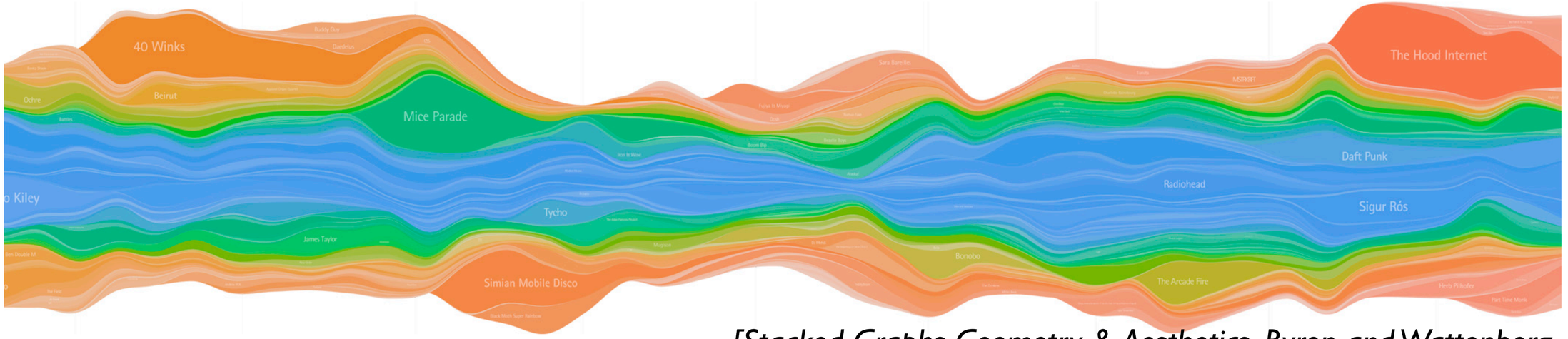
- one more key
  - task
    - part-to-whole relationship
  - scalability
    - several to one dozen levels for stacked attrib



[Using Visualization to Understand the Behavior of Computer Systems. Bosch. Ph.D. thesis, Stanford Computer Science, 2001.]

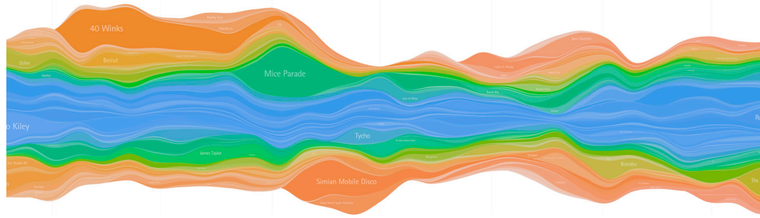
Streamgraph

- generalized stacked graph
  - emphasizing horizontal continuity
    - vs vertical items



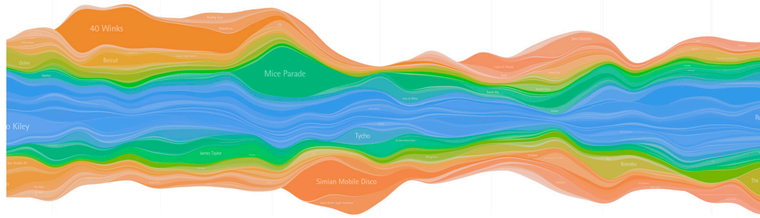
[Stacked Graphs Geometry & Aesthetics. Byron and Wattenberg. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2008) 14(6): 1245–1252, (2008).]

Streamgraph



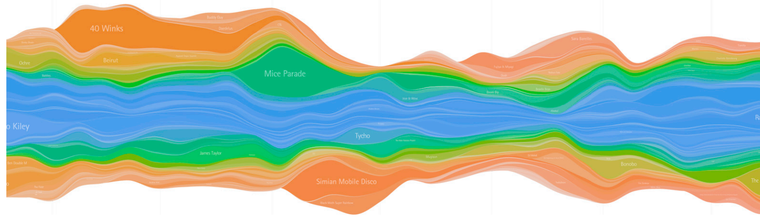
- generalized stacked graph
  - emphasizing horizontal continuity
  - data
    - 1 categ key attrib (artist)
    - 1 ordered key attrib (time)
    - 1 quant value attrib (counts)

Streamgraph



- generalized stacked graph
  - derived data
    - geometry: layers, where height encodes counts
    - 1 quant attrib (layer ordering)

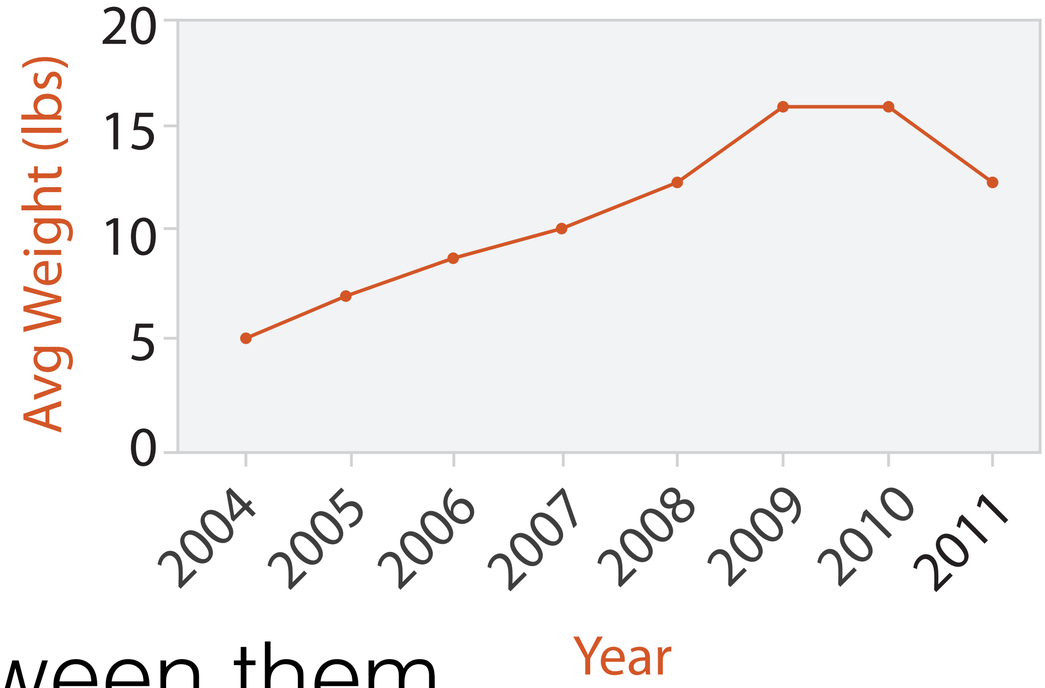
Streamgraph



- generalized stacked graph
- scalability
- hundreds of time keys
- dozens to hundreds of artist keys
- more than stacked bars, since most layers don’t extend across whole chart

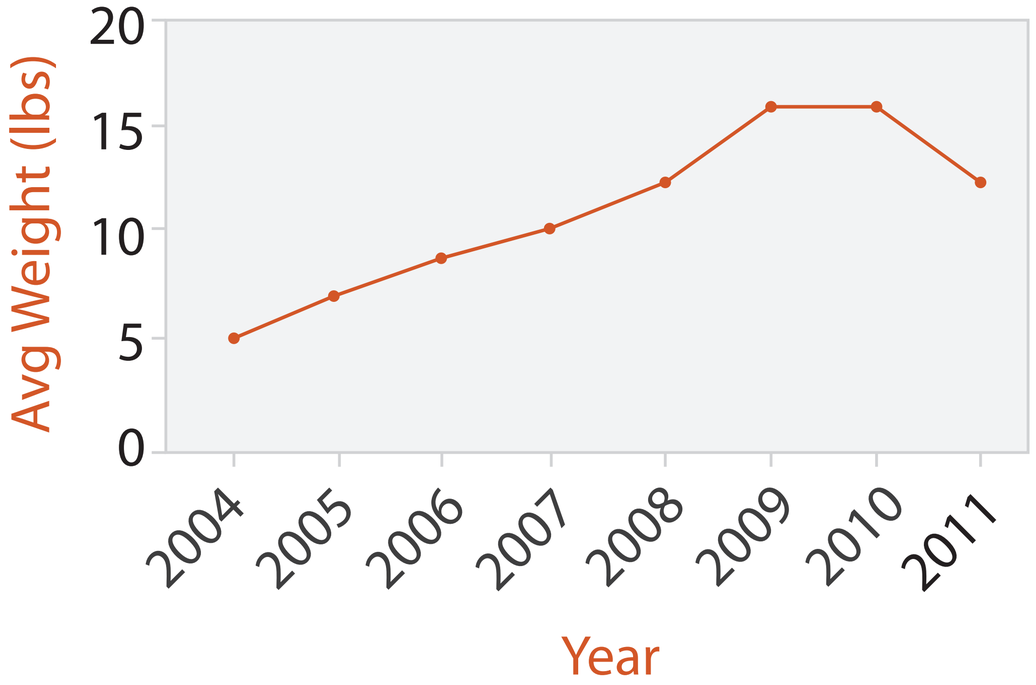
Line Chart

- one key, one value
  - data: 2 quant attribs
  - mark: points
    - line connection marks between them



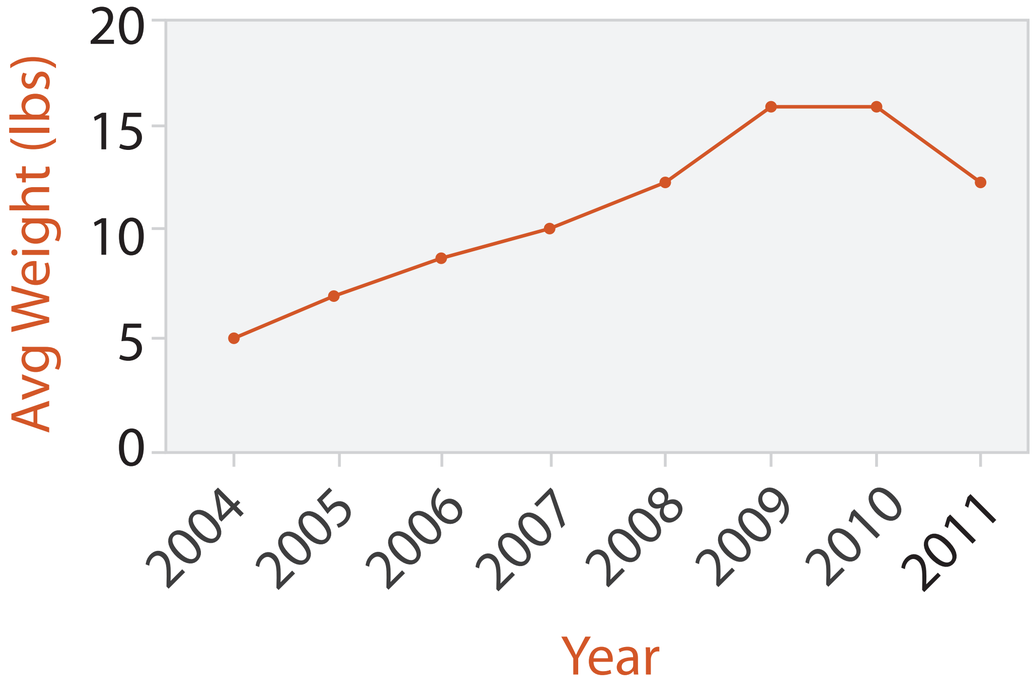
Line Chart

- one key, one value
  - channels
    - aligned lengths to express quant value
    - separated and ordered by key attrib into horizontal regions



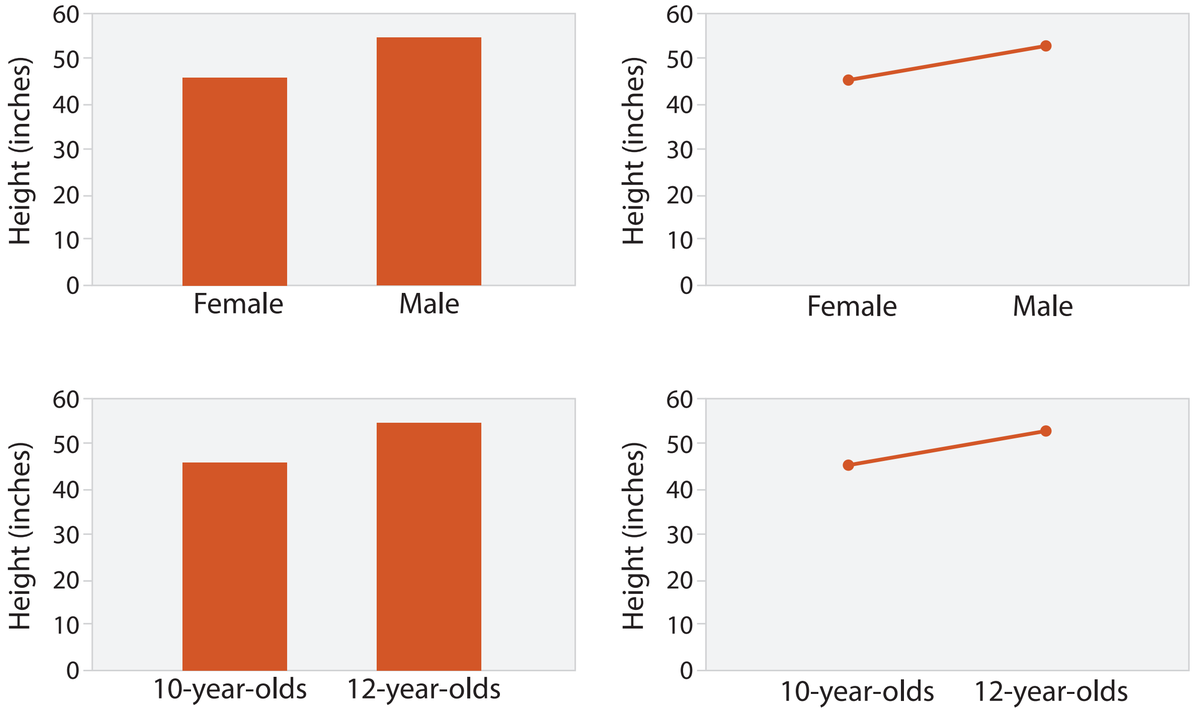
Line Chart

- one key, one value
  - task: find trend
    - connection marks emphasize ordering of items along key axis by explicitly showing relationship between one item and the next



Bar vs Line Charts

- depends on type of key attrib
  - bar charts if categorical
  - line charts if ordered

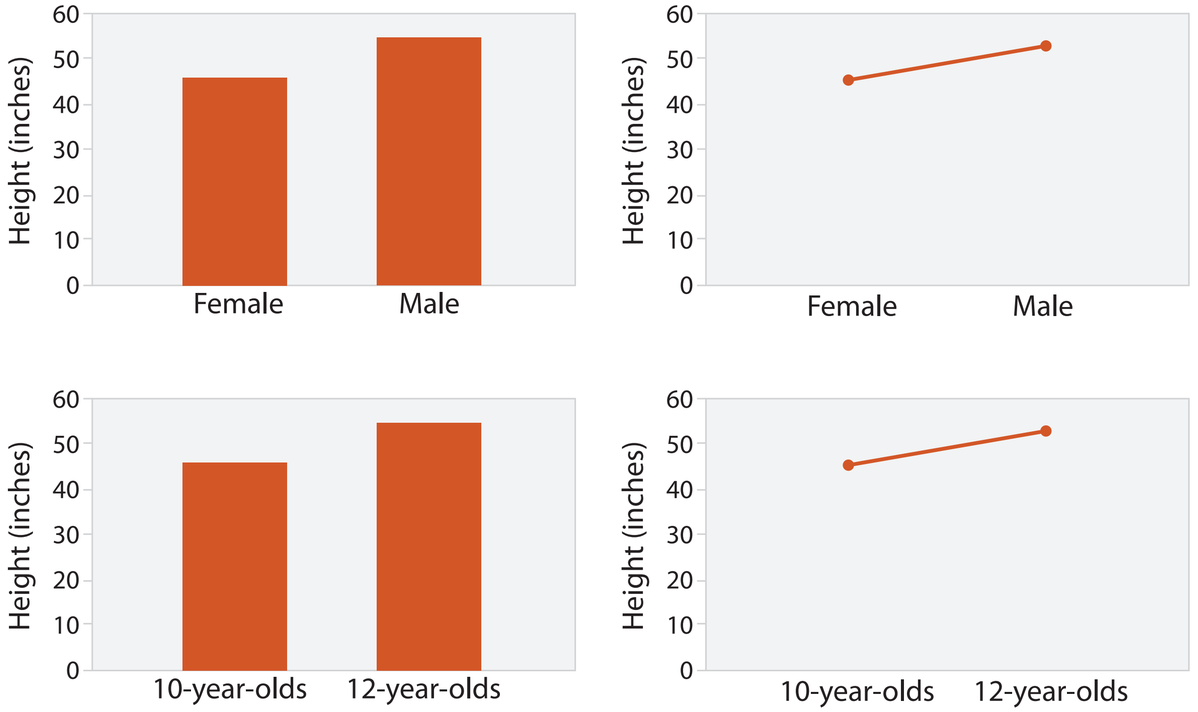


after [Bars and Lines: A Study of Graphic Communication. Zacks and Tversky. Memory and Cognition 27:6 (1999), 1073–1079.]

Bar vs Line Charts

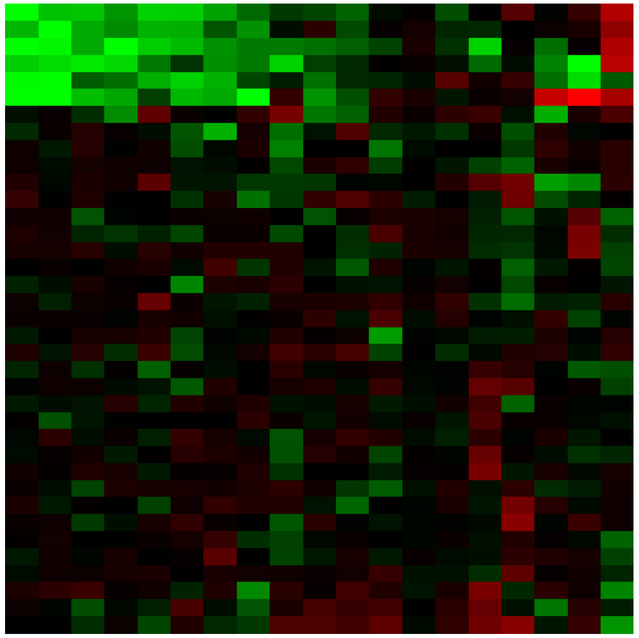
- do not use line charts for categorical key attribs
  - violates expressiveness principle

"*The more male a person is, the taller he/she is*"

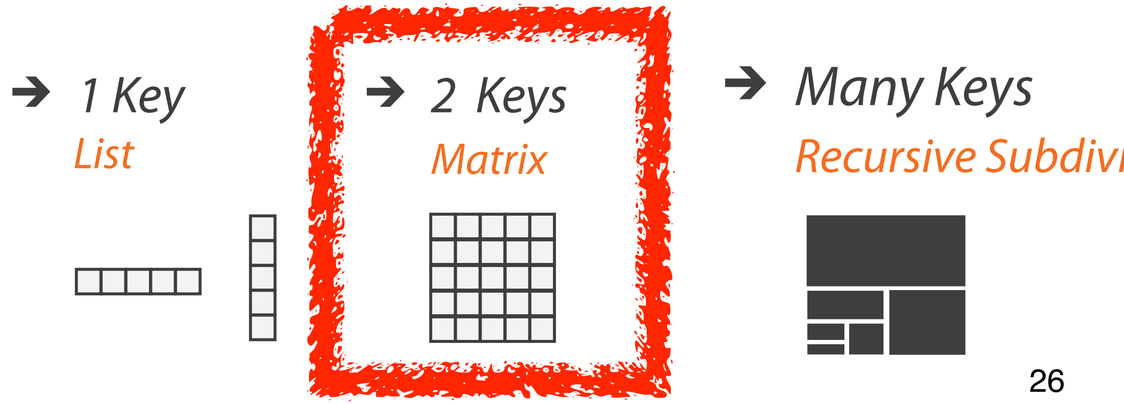


after [Bars and Lines: A Study of Graphic Communication. Zacks and Tversky. Memory and Cognition 27:6 (1999), 1073–1079.]

Heatmap

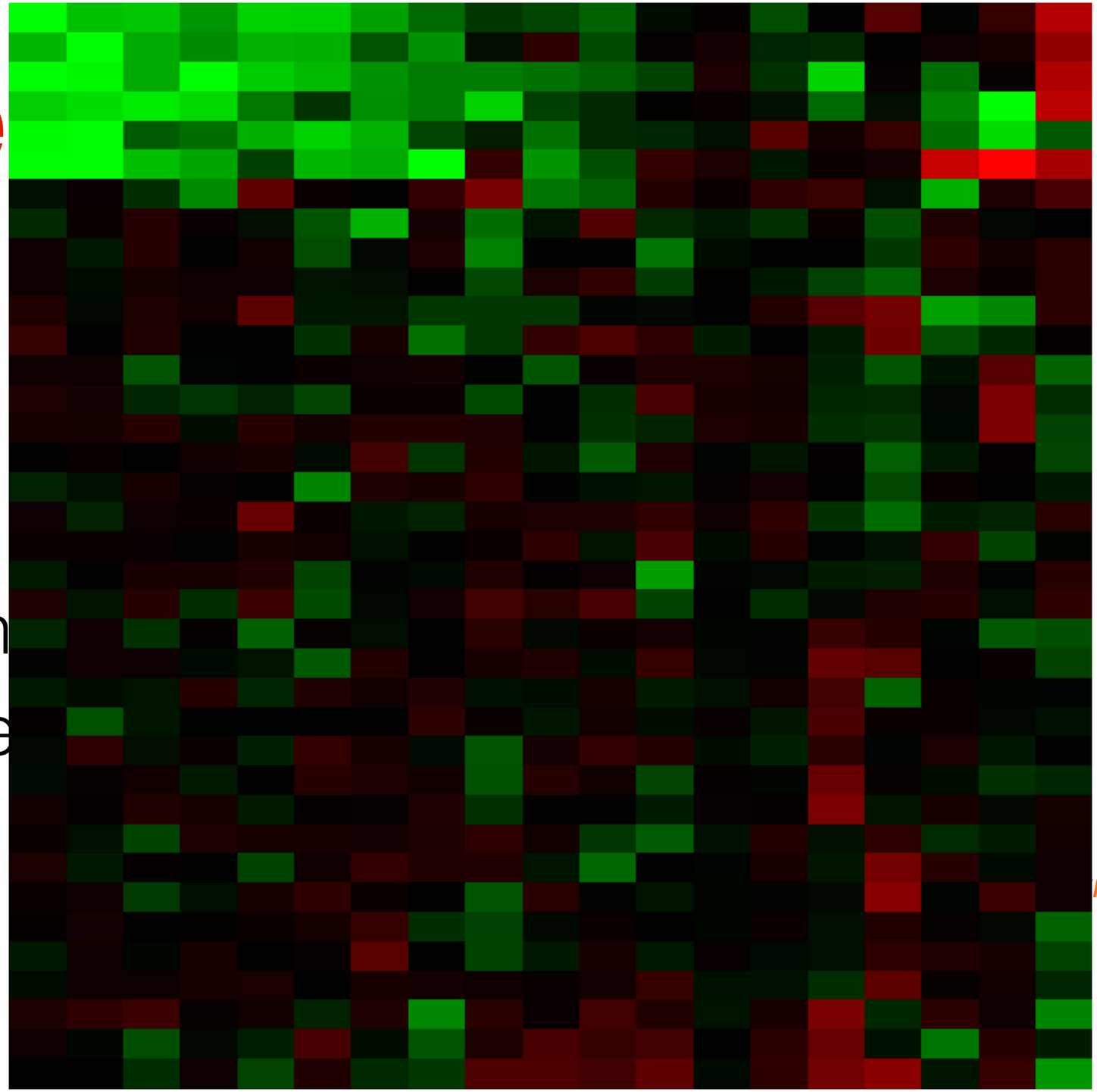


- two keys, one value
  - data: 2 categ., 1 quant.
  - marks: area
    - separate and align in 2D matrix
      - indexed by 2 categorical attributes

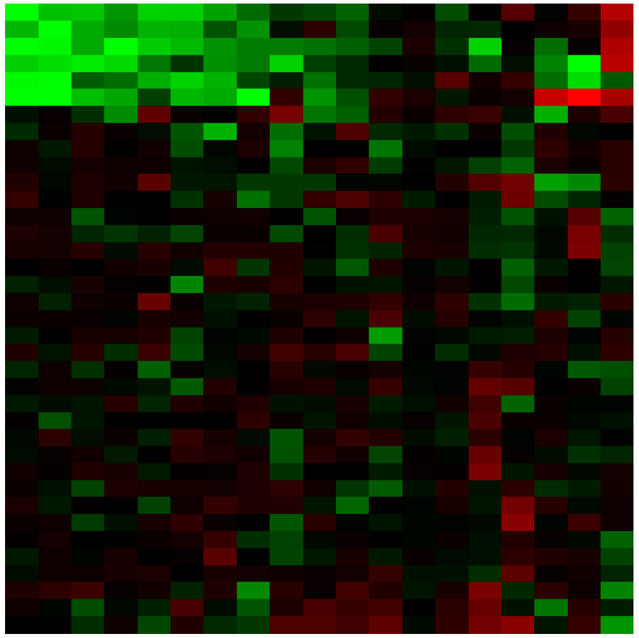


He

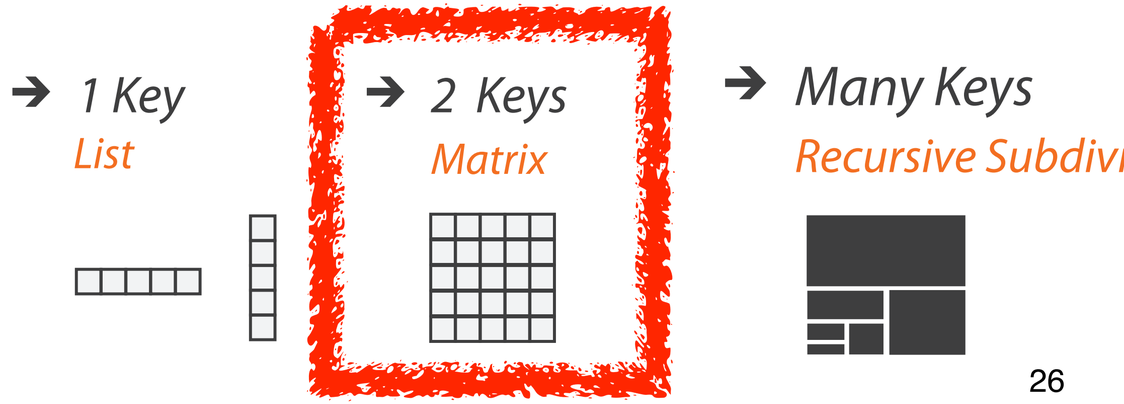
- two keys, one
  - data: 2 categ., 1
  - marks: area
    - separate and align
      - indexed by 2 cate



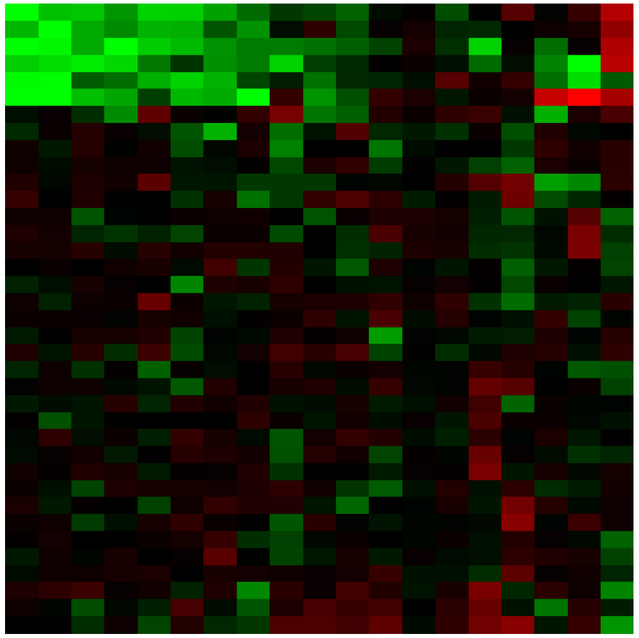
Heatmap



- two keys, one value
  - data: 2 categ., 1 quant.
  - marks: area
    - separate and align in 2D matrix
      - indexed by 2 categorical attributes

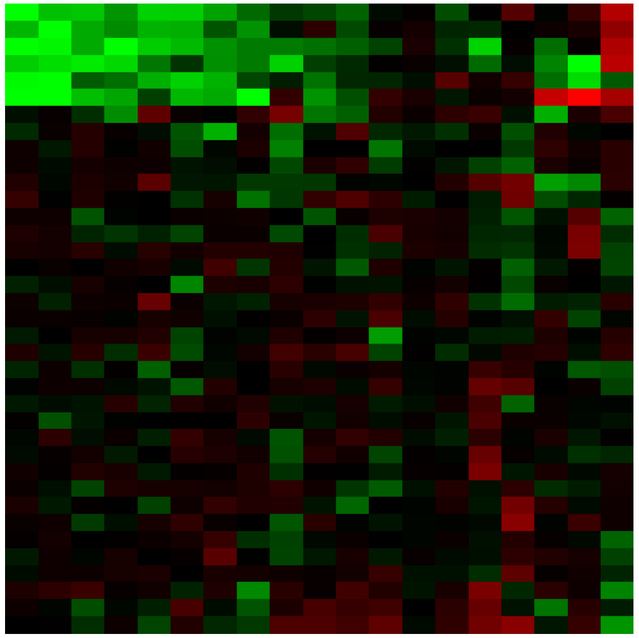


Heatmap



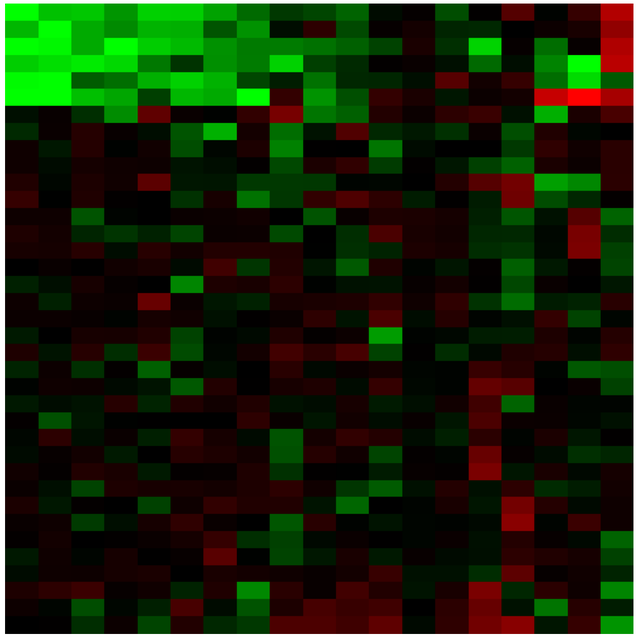
- two keys, one value
  - data: 2 categ.,1 quant.
  - marks: area
  - channels: color by quant attrib
    - (ordered diverging colormap)

Heatmap



- two keys, one value
  - data: 2 categ., 1 quant.
  - marks: area
  - channels: color by quant attrib
  - task: find clusters, outliers

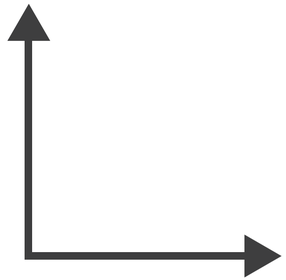
Heatmap



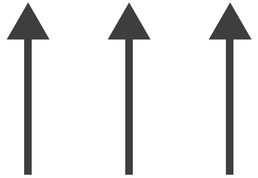
- two keys, one value
  - data: 2 categ., 1 quant.
  - marks: area
  - channels: color by quant attrib
  - task: find clusters, outliers
  - scalability:1M items, 100s of categ levels, ~10 quant attrib levels

Axis Orientation

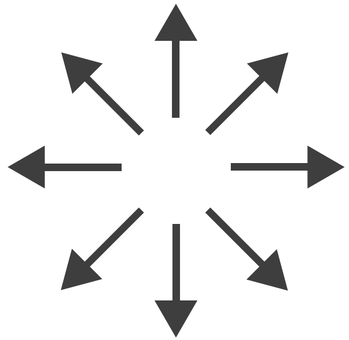
- Rectilinear



- Parallel



- Radial

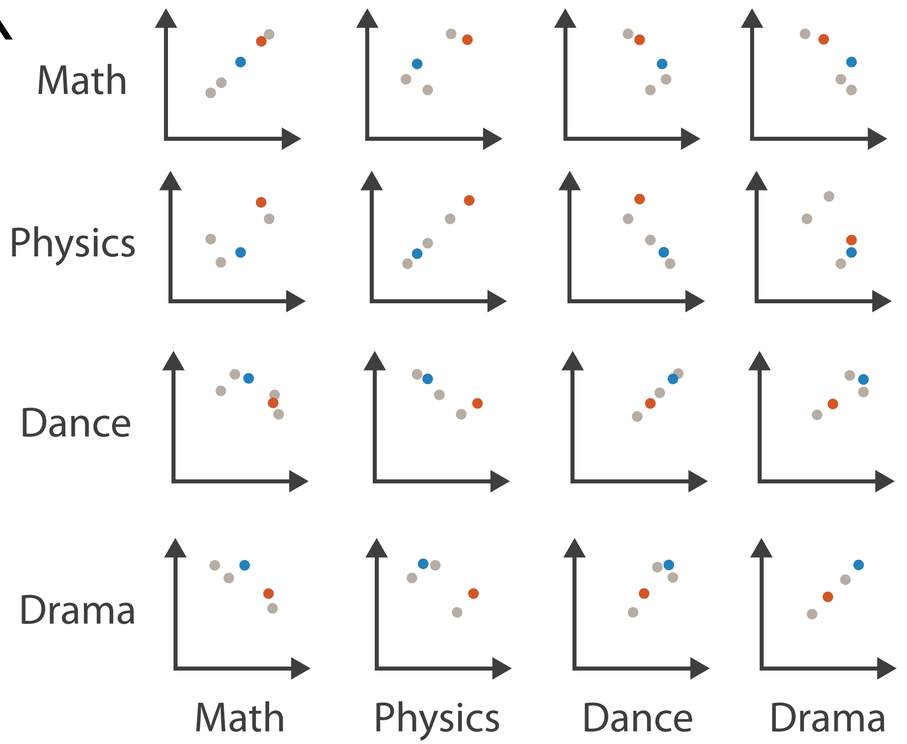


Scatterplot Matrix

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

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



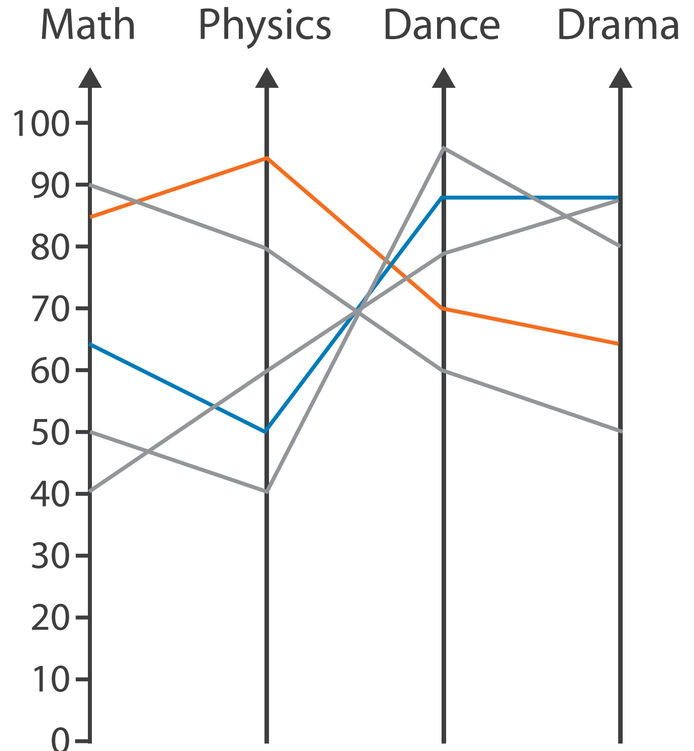
after [Visualization Course Figures. McGuffin, 2014. http://www.michaelmcguffin.com/courses/vis/]

Parallel Coordinates

Table

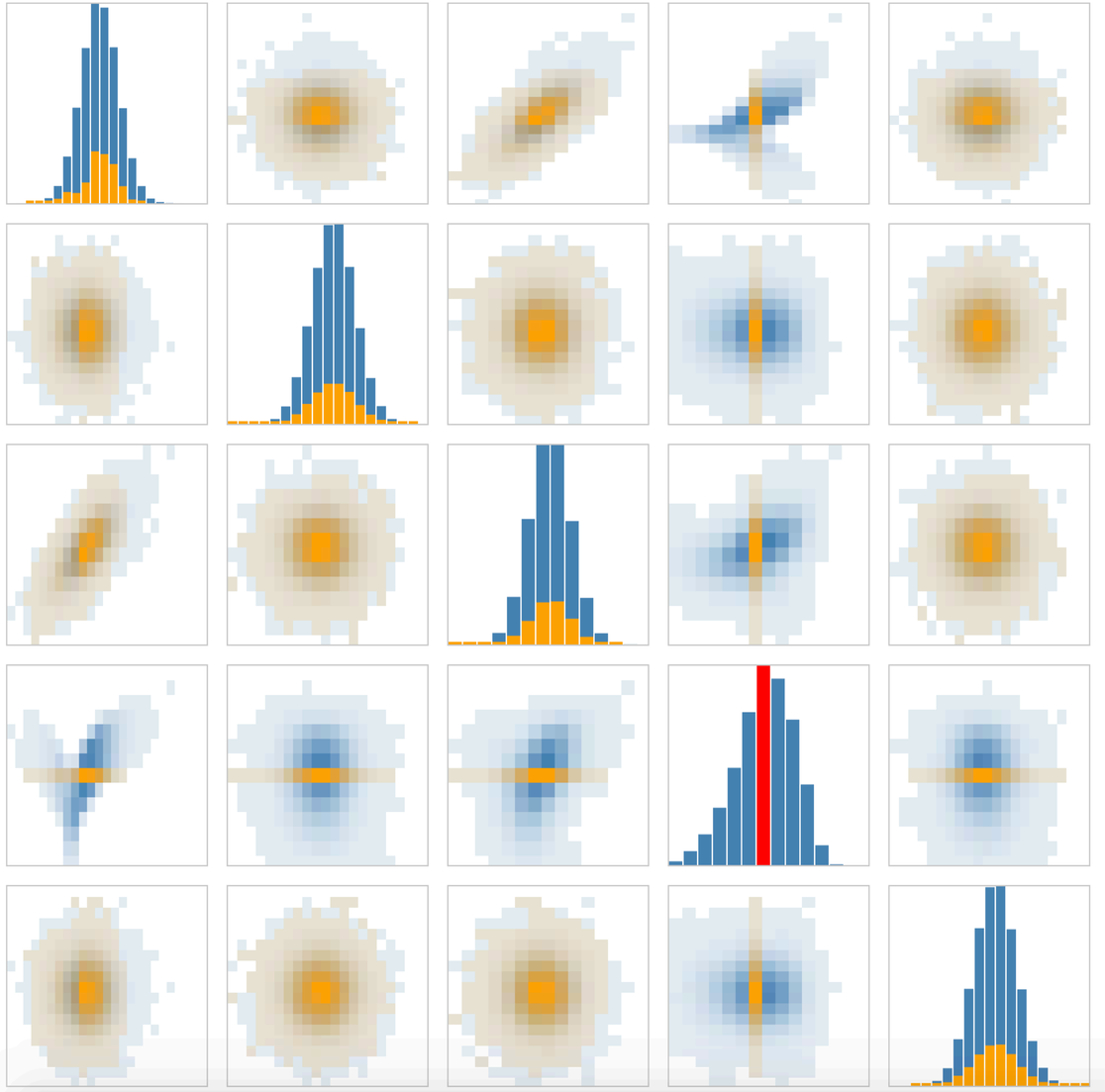
Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

- 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



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

Binned Scatter Plot



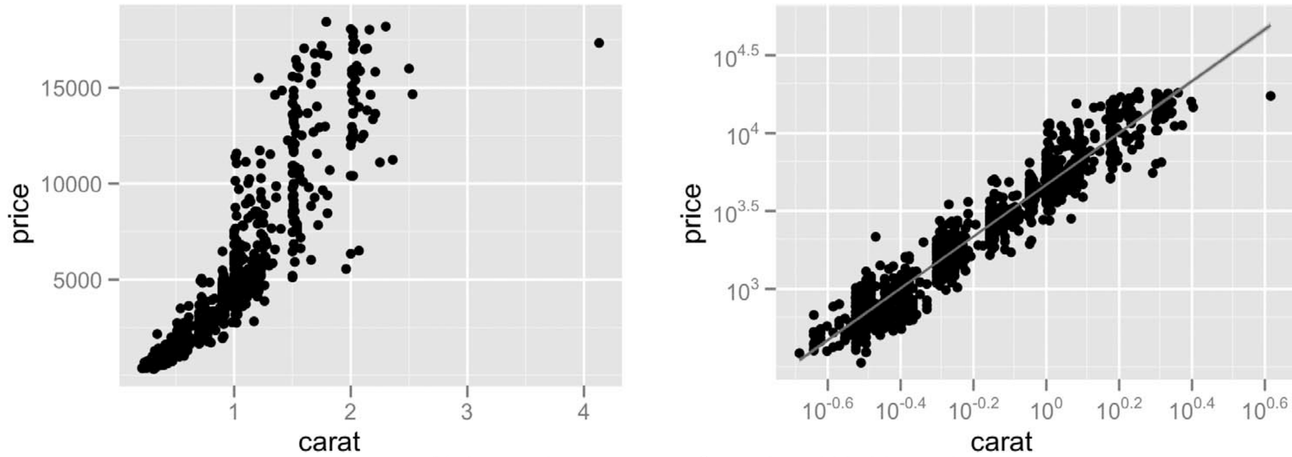
http://vis.stanford.edu/projects/datavore/splom/

But Also...

https://www.youtube.com/watch?v=iFqCBl4T8ks&feature=youtu.be

Task: Correlation

- scatterplot matrix
  - positive correlation
    - diagonal low-to-high
  - negative correlation
    - diagonal high-to-low
  - uncorrelated



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

Task: Correlation

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

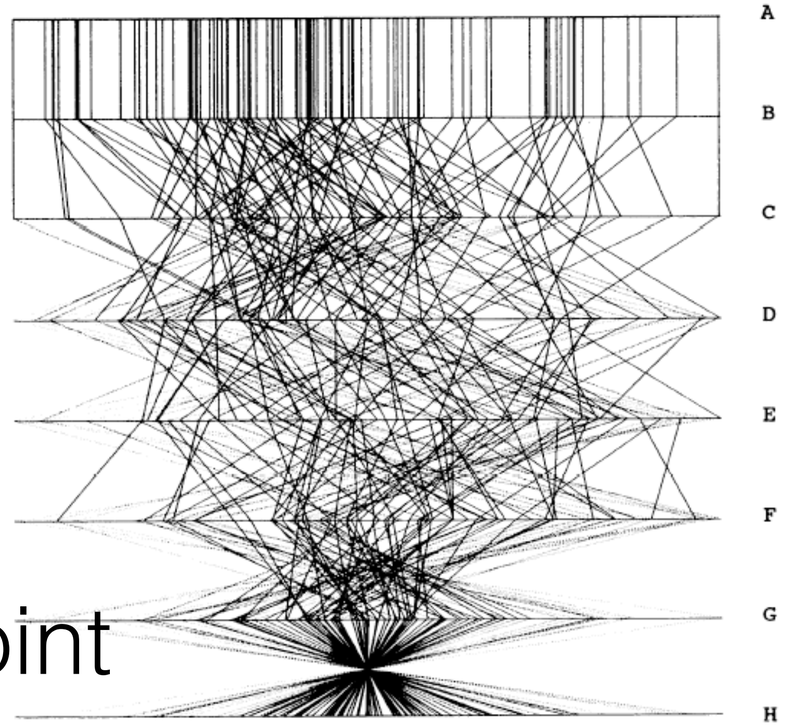


Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of $\rho = 1, .8, .2, 0, -.2, -.8,$ and -1.

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

Radial Bar Chart / Star Plot

- radial bar chart
  - radial axes meet at central ring, line mark

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

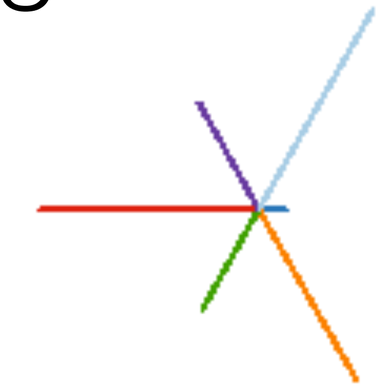
Radial Bar Chart / Star Plot

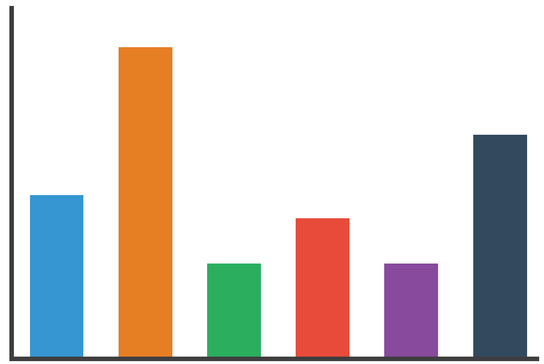
- radial bar chart
- star plot
  - radial axes, meet at central point, line mark

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

Radial Bar Chart / Star Plot

- radial bar chart
- star plot
- bar chart
  - rectilinear axes, aligned vertically

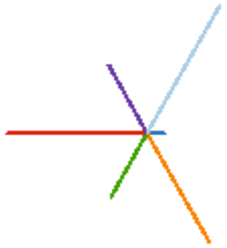




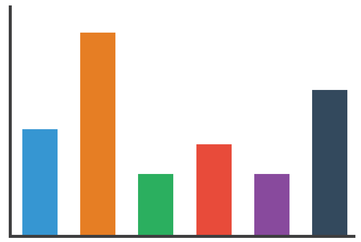
[Vismon: Facilitating Risk Assessment and Decision Making In Fisheries Management. Booshehrian, Möller, Peterman, and Munzner. Technical Report TR 2011-04, Simon Fraser University, School of Computing Science, 2011.]

Radial Bar Chart / Star Plot

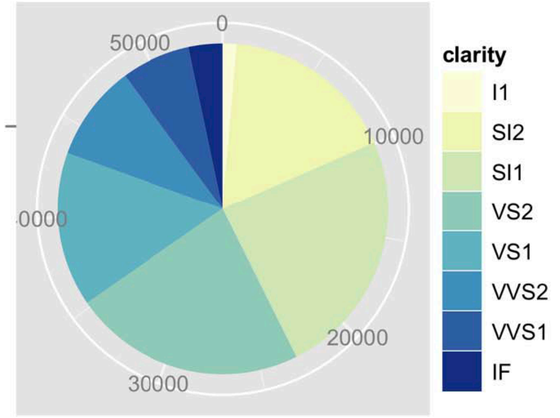
- radial bar chart
- star plot
- bar chart



- accuracy
  - length unaligned with radial
    - less accurate than aligned with rectilinear



Pie Charts

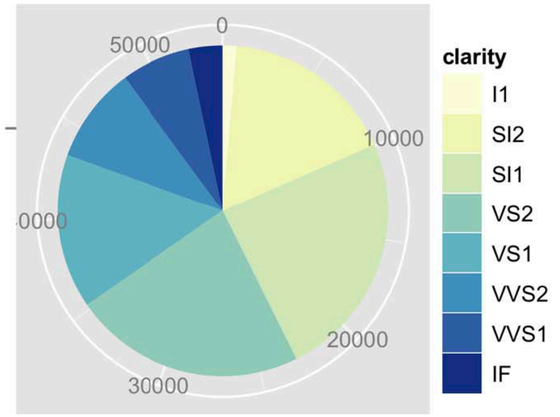


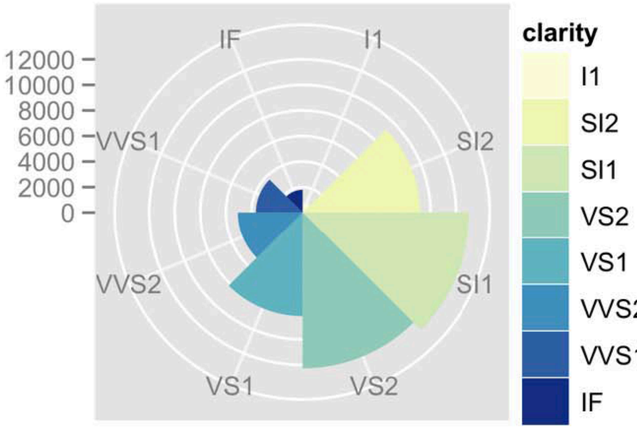
- Pie chart
  - area marks with angle channel
  - accuracy: angle/area much less accurate than line length

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

Pie Charts

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

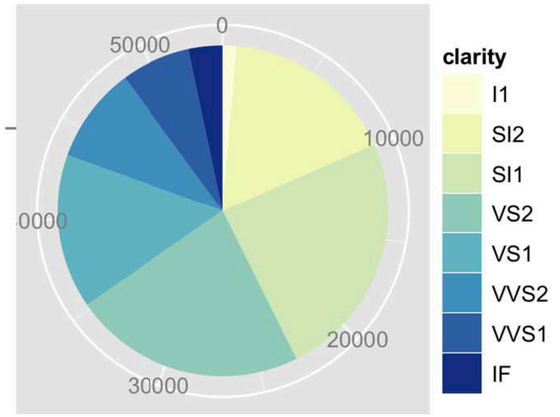


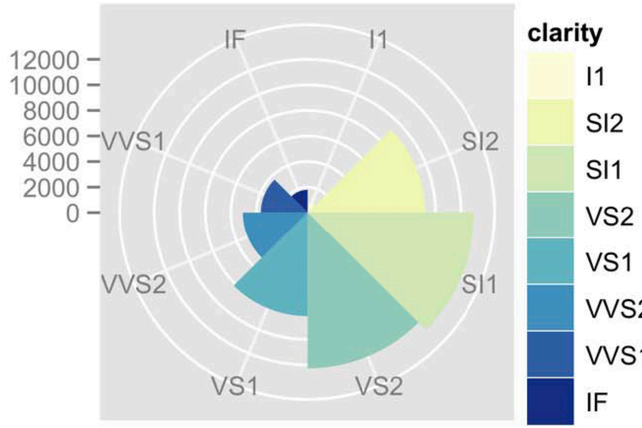


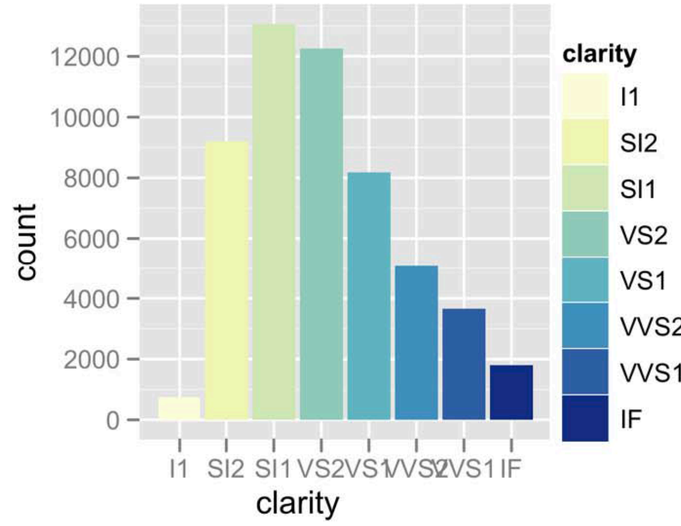
[A layered grammar of graphics. Wickham. Journ. Computational and Graphical Statistics 19:1 (2010), 3–28.]

Pie Chart

- Pie chart
- Polar area chart
- Data
  - 1 categorical key attribute
  - 1 quantitative value attribute

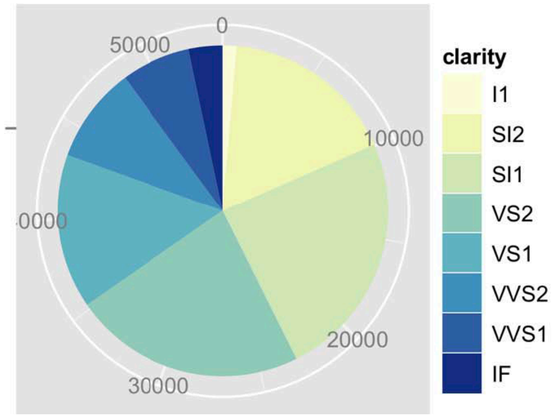


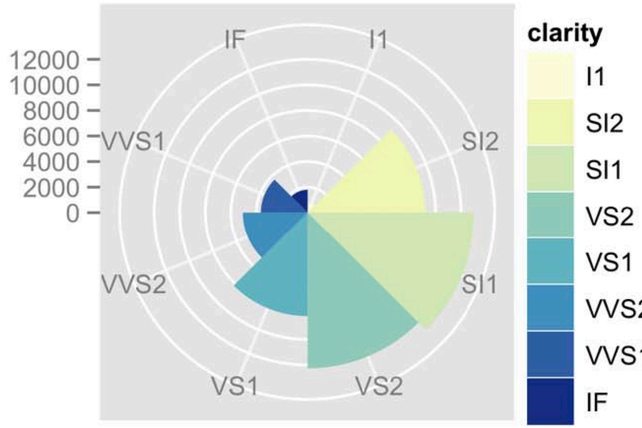


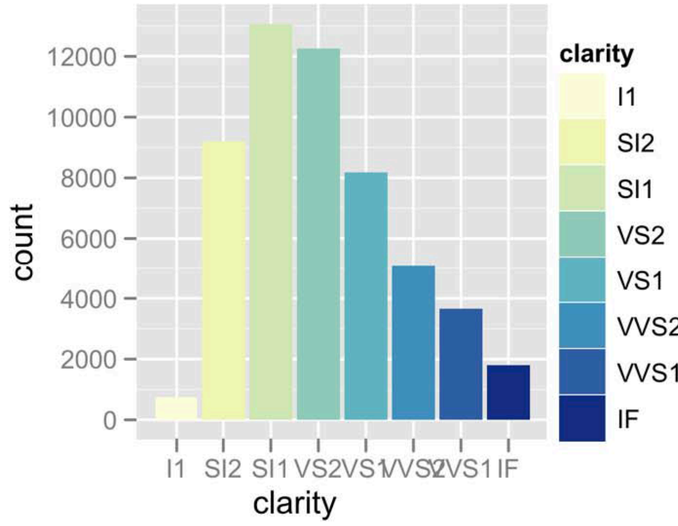


Pie Chart

- Pie chart
- Polar area chart
- Data
- Task
  - part-to-whole judgements

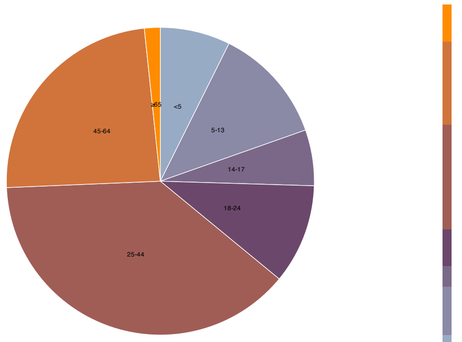


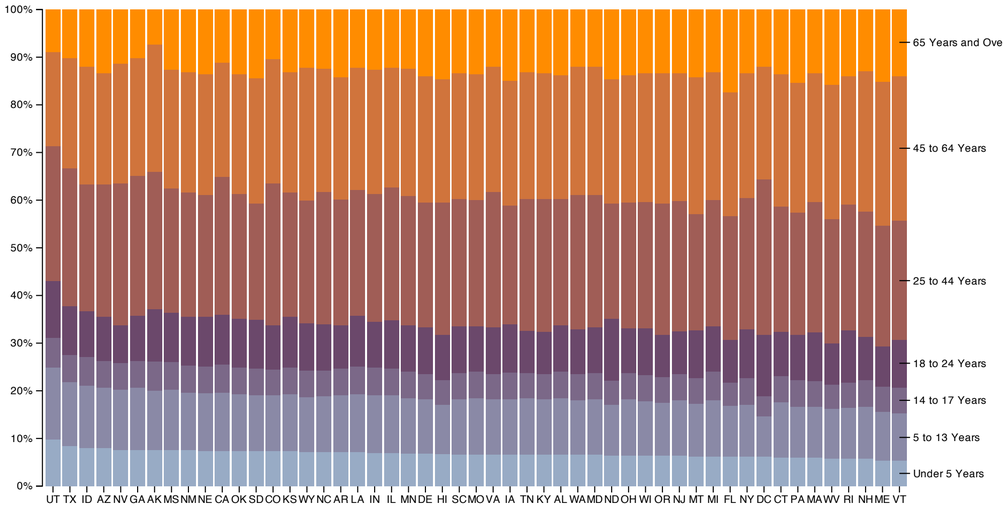


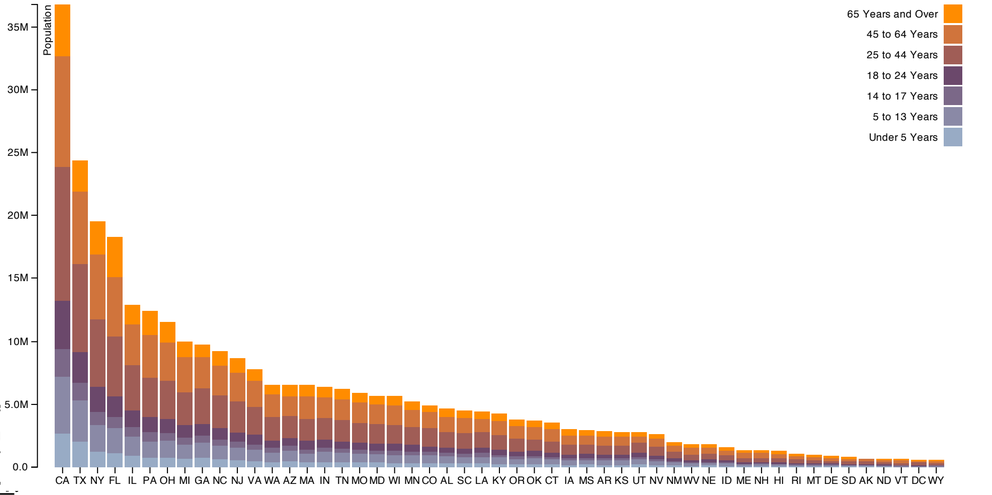


Normalized Stacked Bar Chart

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







http://bl.ocks.org/mbostock/388723
http://bl.ocks.org/mbostock/388620
http://bl.ocks.org/mbostock/388639

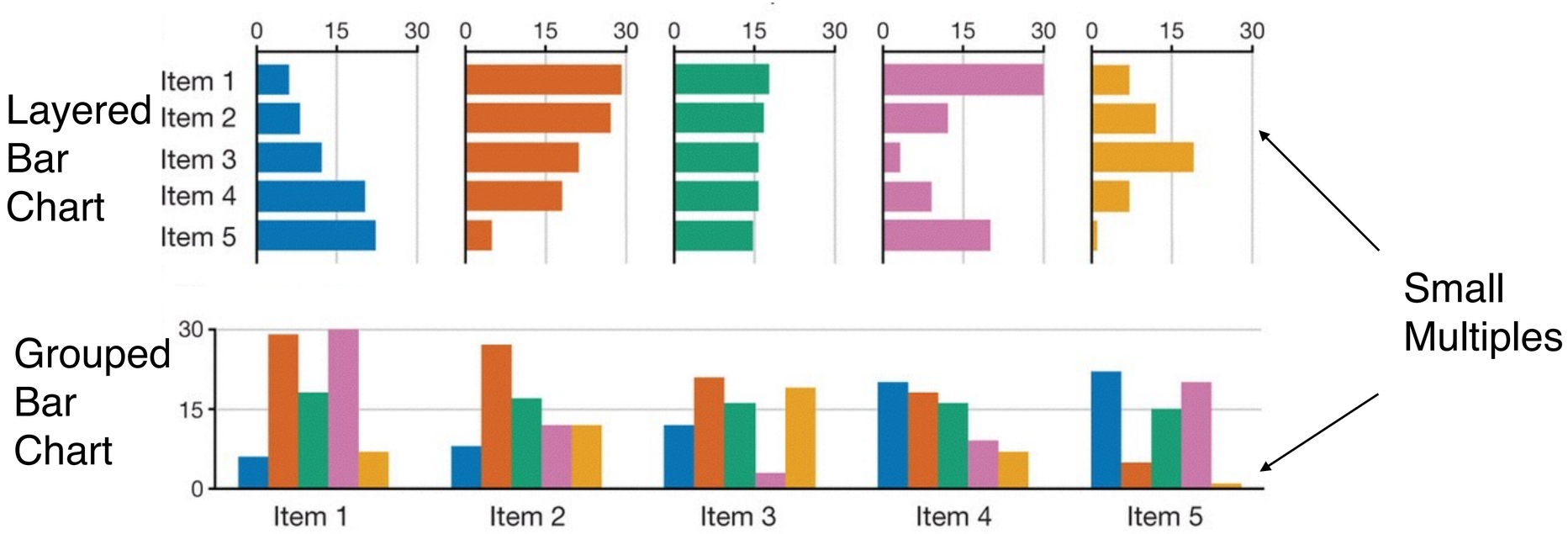
Charts Comparison

Pie Chart

Stacked bar chart

Layered Bar Chart

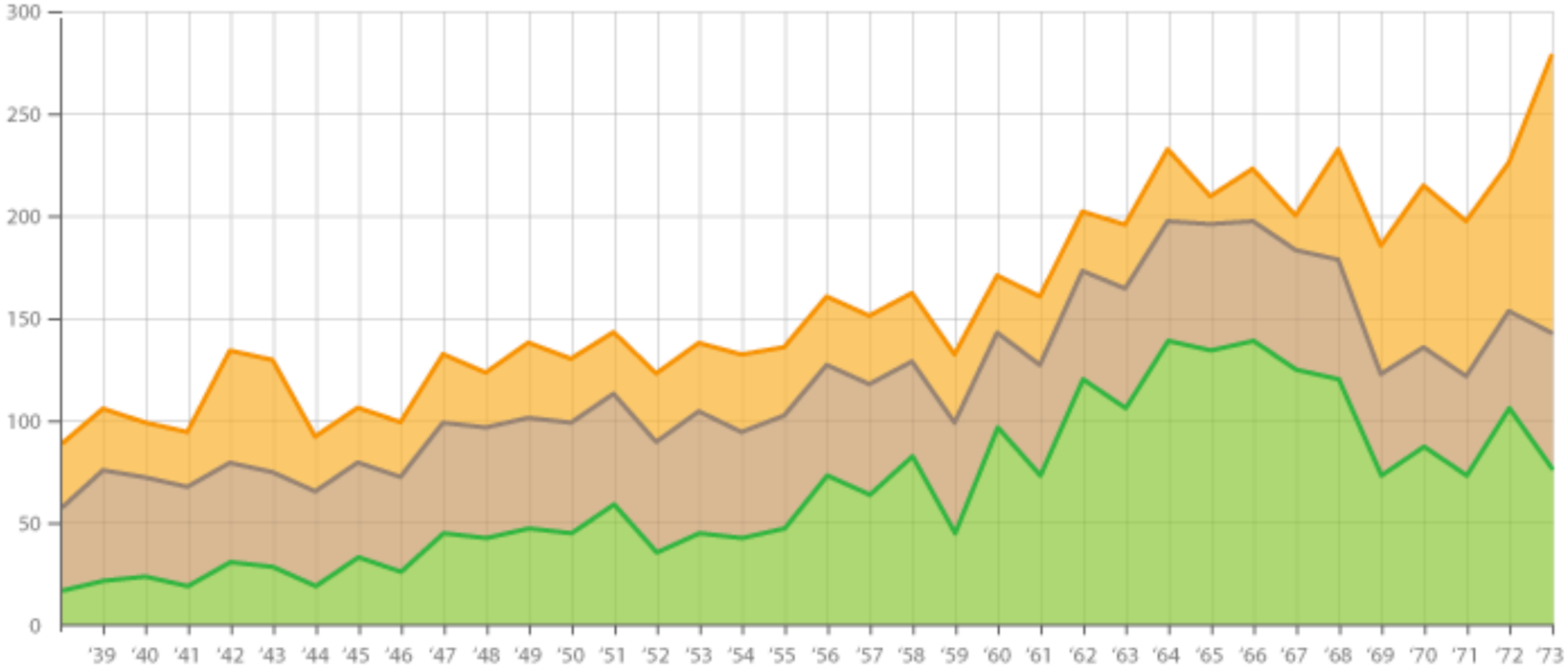
Grouped Bar Chart



Small Multiples

Streit & Gehlenborg, PoV, Nature Methods, 2014

Stacked Area Chart



https://datavizcatalogue.com/methods/stacked_area_graph.html

Graph Horizon

Cubism.js

Time Series Visualization