

# MADAB DATA VISUALIZATION

## 1c. SPATIAL LAYOUT

Instructor: Rossano Schifanella

# Encode tables: Arrange space

## Encode

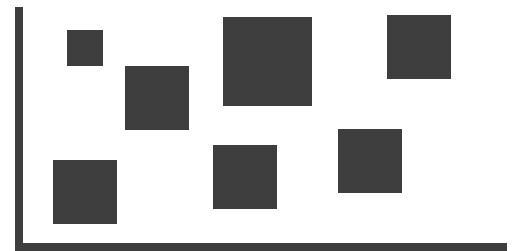
---

→ Arrange

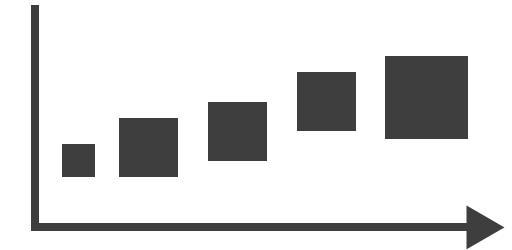
→ Express



→ Separate



→ Order



→ Align



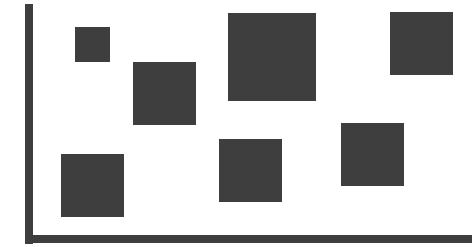
# Arrange tables

## → Express Values



## → Separate, Order, Align Regions

→ Separate



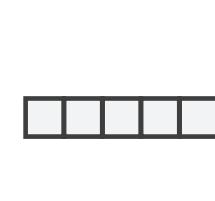
→ Order



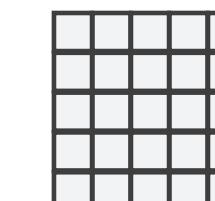
→ Align



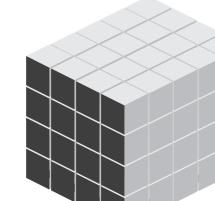
→ 1 Key  
*List*



→ 2 Keys  
*Matrix*

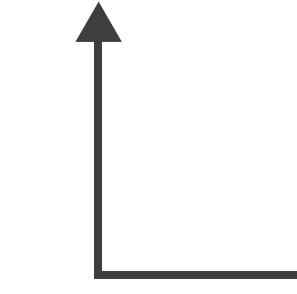


→ 3 Keys  
*Volume*

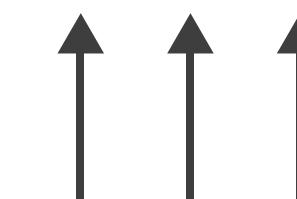


## → Axis Orientation

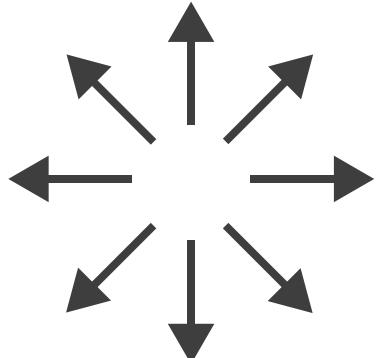
→ Rectilinear



→ Parallel

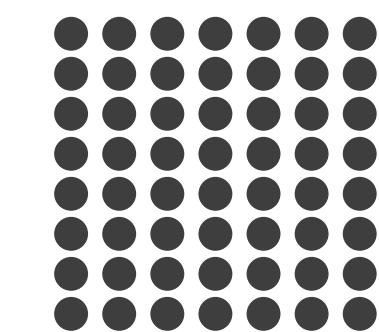


→ Radial



## → Layout Density

→ Dense



→ Space-Filling



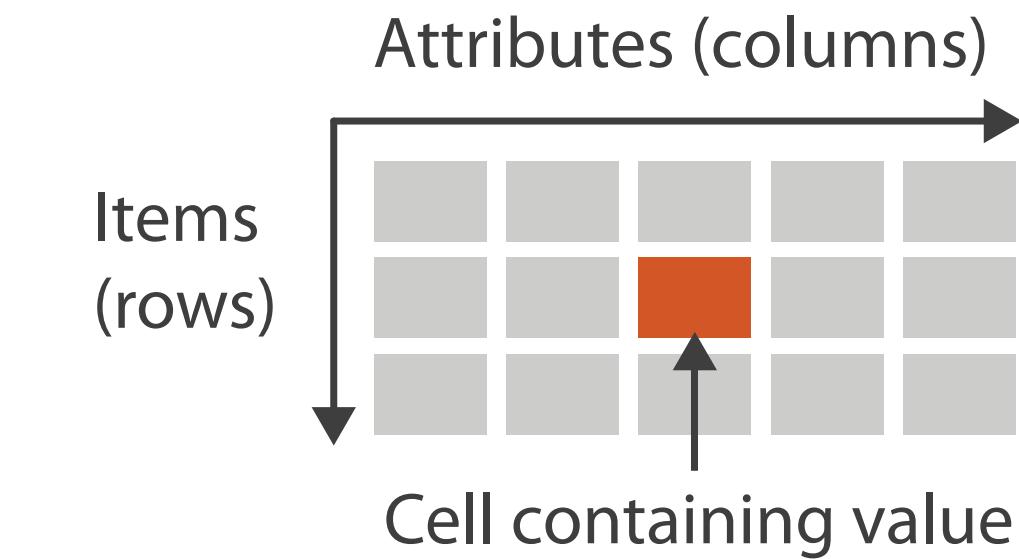
→ Many Keys  
*Recursive Subdivision*



# Keys and values

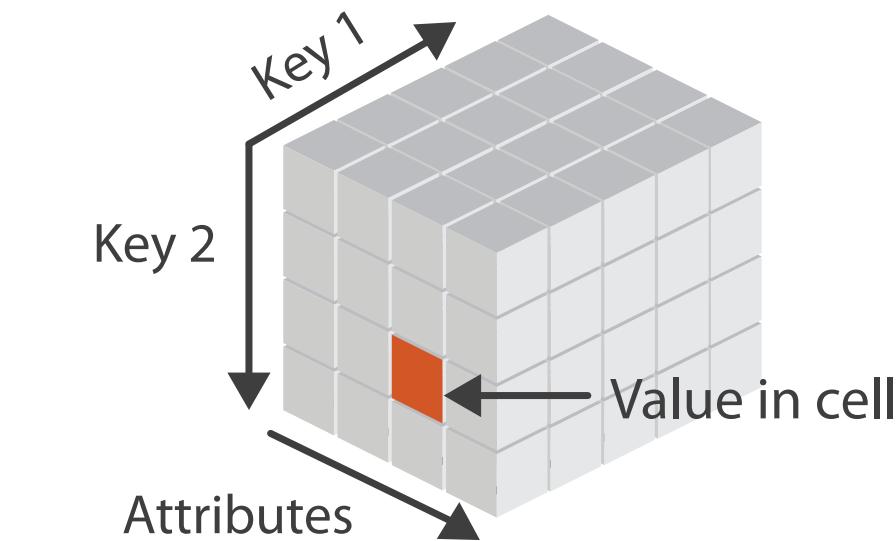
→ Tables

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

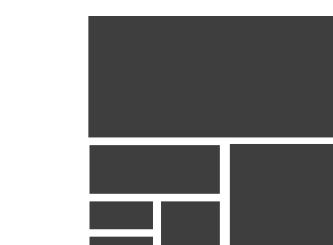
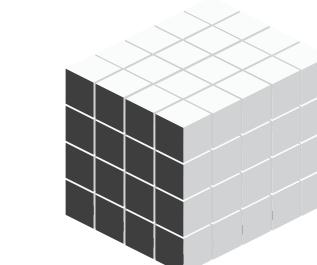
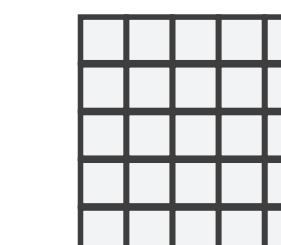
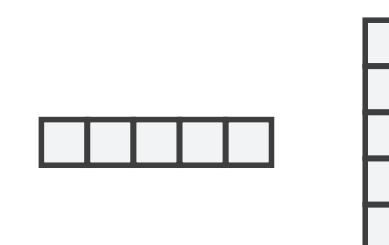


→ *Multidimensional Table*

- **value**
  - dependent attribute, value of cell
- **classify arrangements by key count**
  - 0, 1, 2, many...



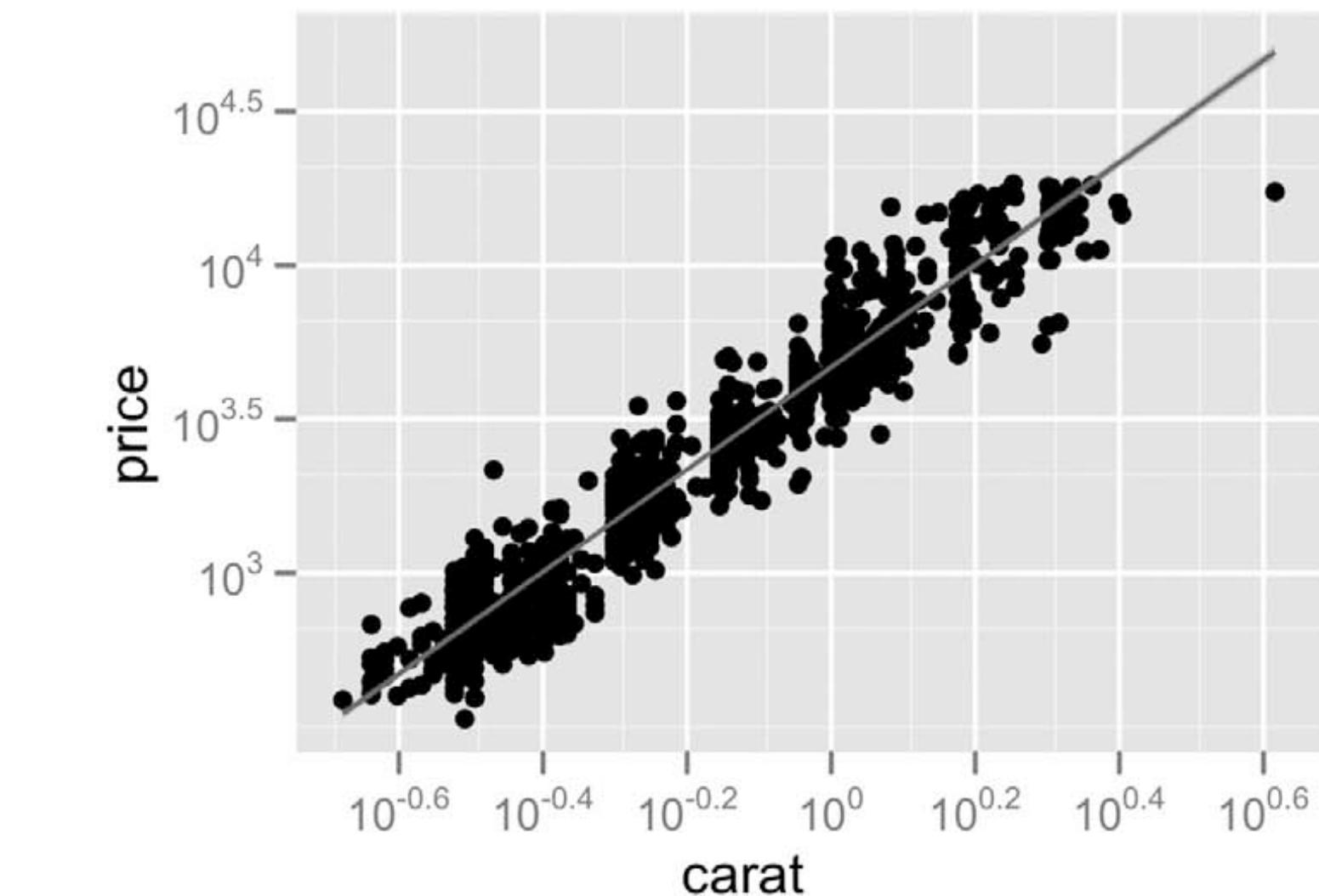
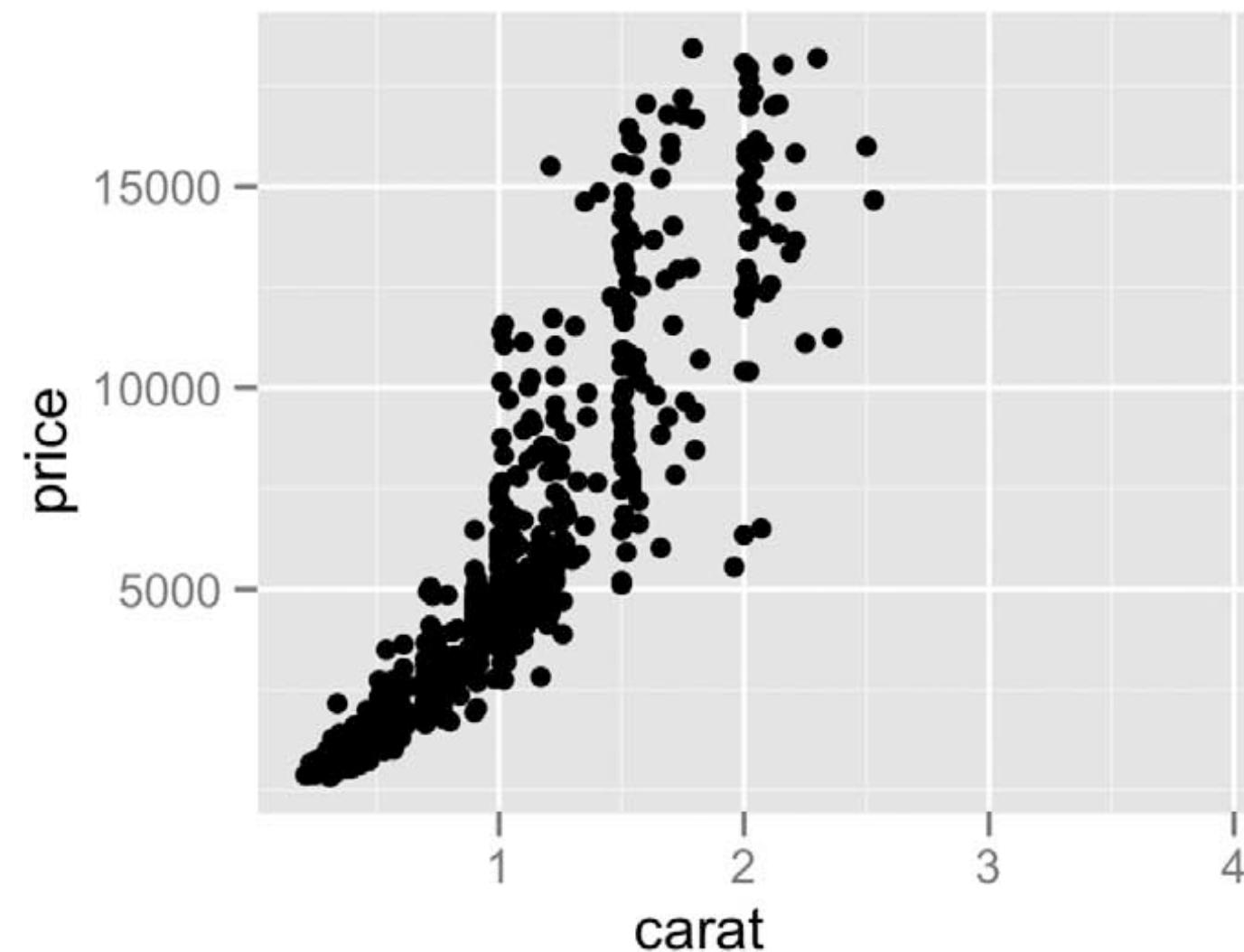
→ Express Values → 1 Key → 2 Keys → 3 Keys → Many Keys  
*List*                    *Matrix*                    *Volume*                    *Recursive Subdivision*



# Idiom: scatterplot

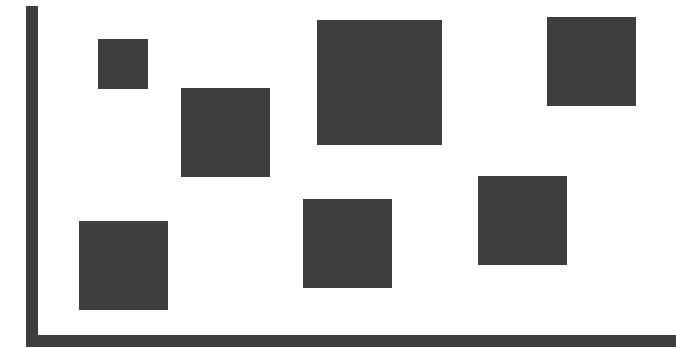
→ Express Values

- **express** values
  - quantitative attributes
- no keys, only values
  - data
    - 2 quant attrs
  - mark: points
  - channels
    - horiz + vert position
  - tasks
    - discover trends, locate outliers, distribution, correlation, clusters
  - scalability
    - hundreds of items



# Some keys: Categorical regions

→ Separate



→ Order

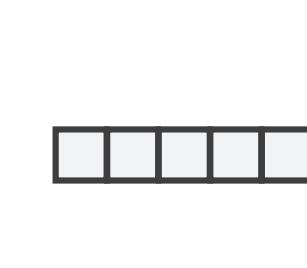


→ Align

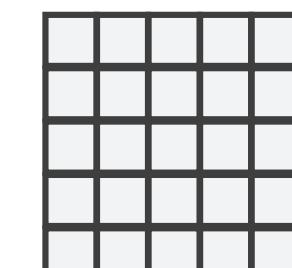


- **regions**: contiguous bounded areas distinct from each other
  - using space to **separate** (proximity)
  - following expressiveness principle for categorical attributes
- use ordered attribute to **order** and **align** regions

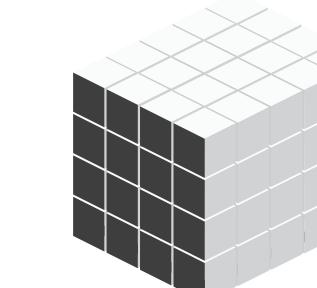
→ 1 Key  
*List*



→ 2 Keys  
*Matrix*



→ 3 Keys  
*Volume*

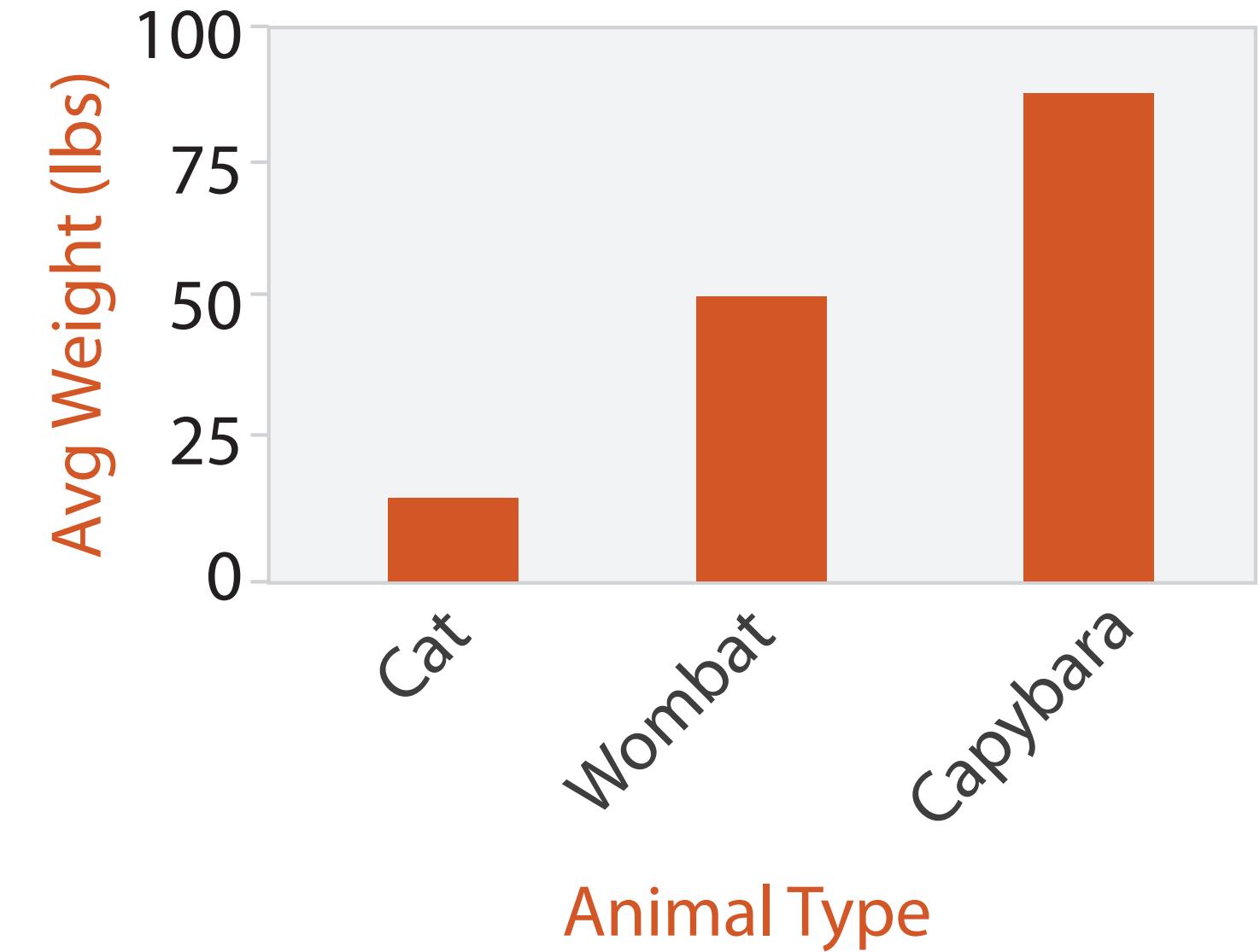
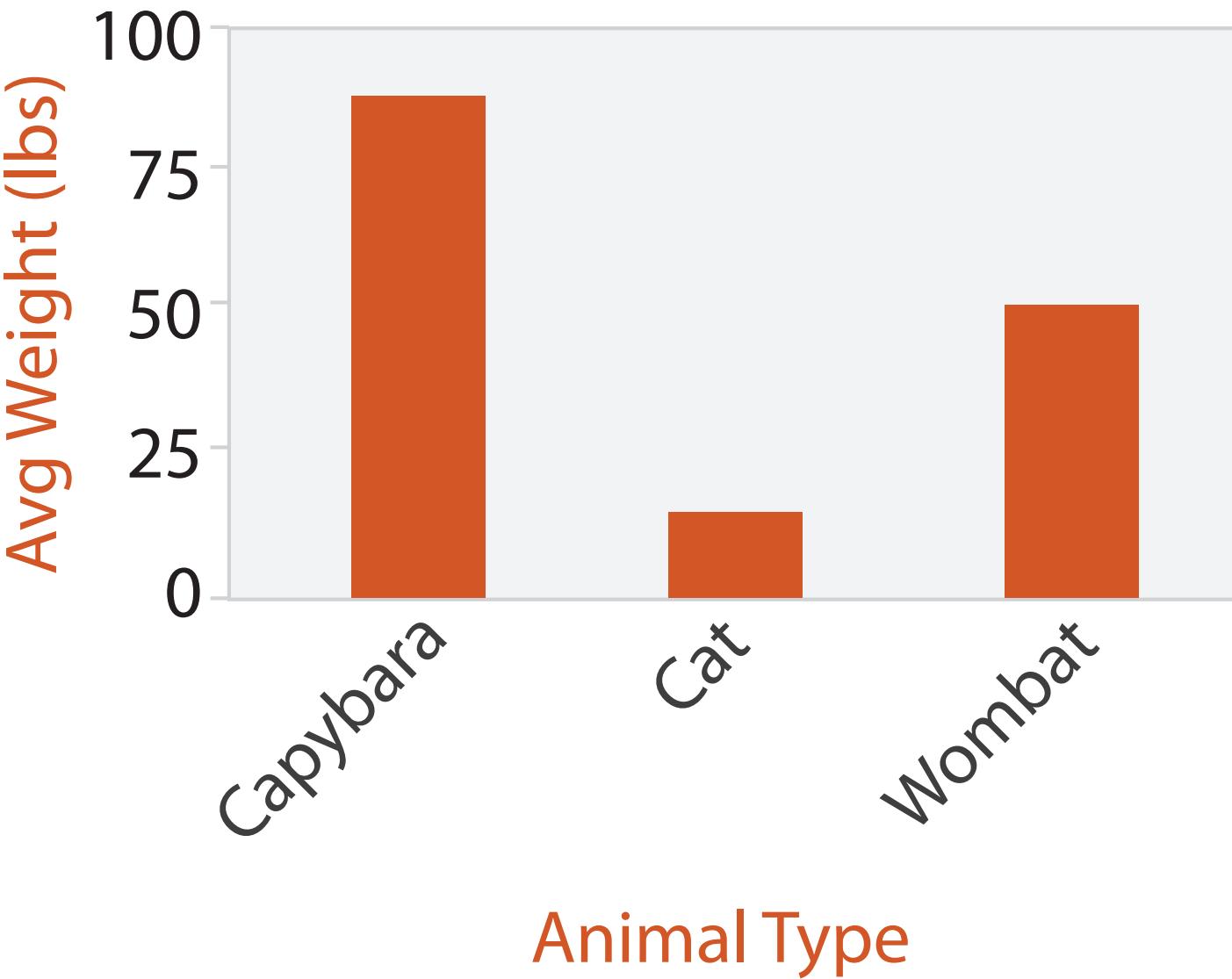


→ Many Keys  
*Recursive Subdivision*

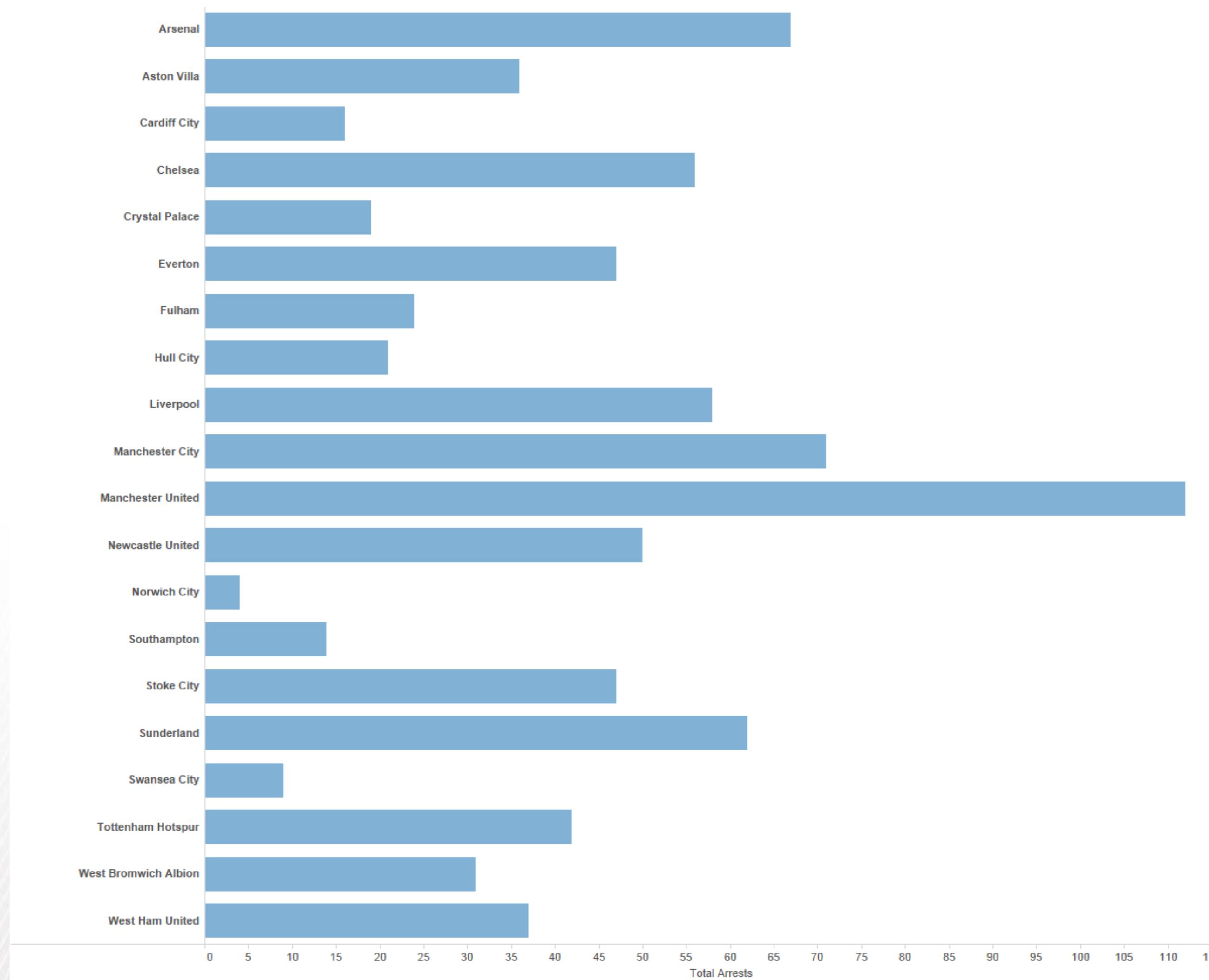


# Idiom: bar chart

- one key, one value
  - data
    - 1 categ attrib, 1 quant attrib
  - 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

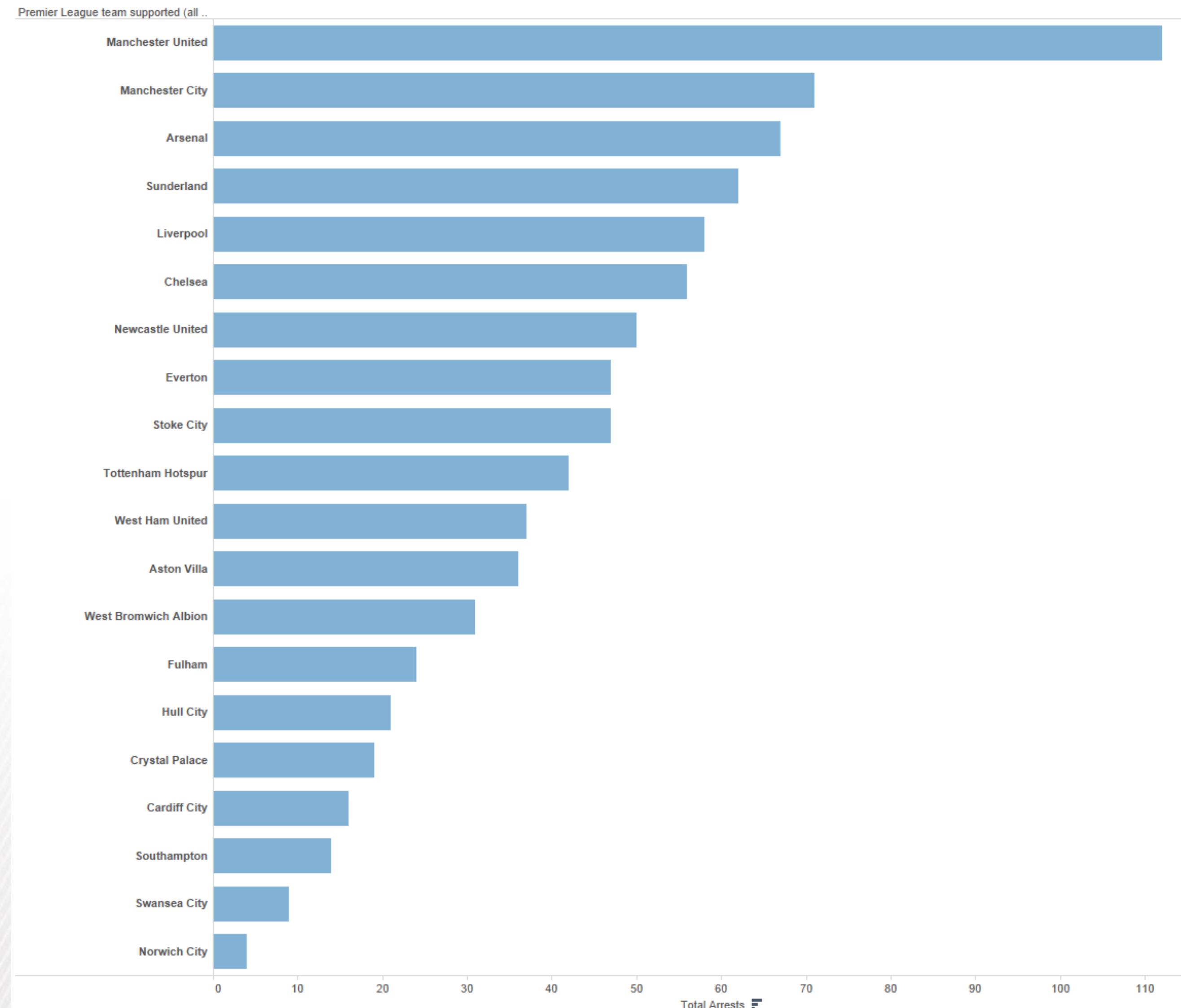


# Separated and Aligned but not Ordered



**LIMITATION:** Hard to know rank. What's the 4<sup>th</sup> most? The 7<sup>th</sup>?

# Separated, Aligned and Ordered



[Slide courtesy of Ben Jones]

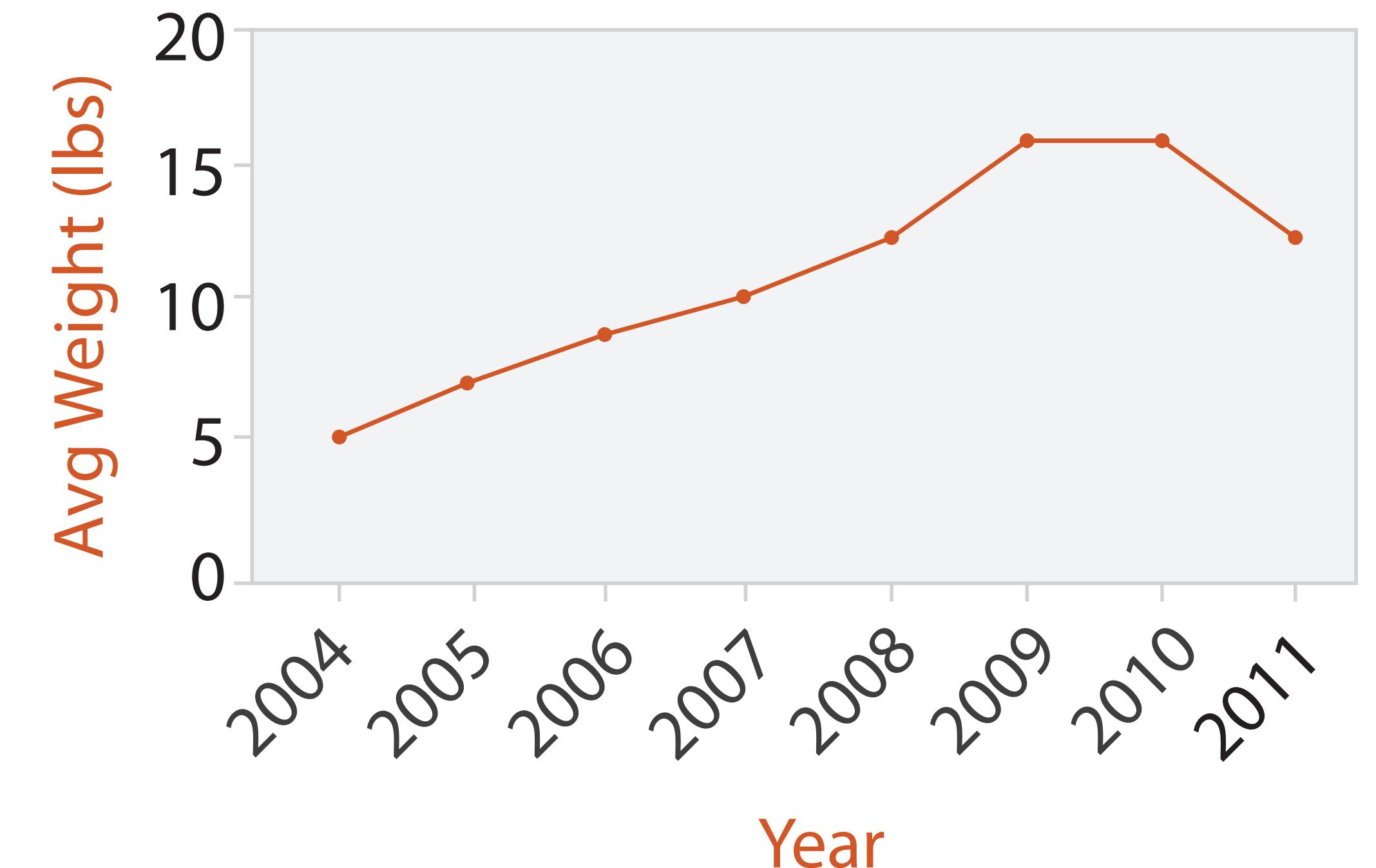
# Separated but not Ordered or Aligned



**LIMITATION:** Hard to make comparisons

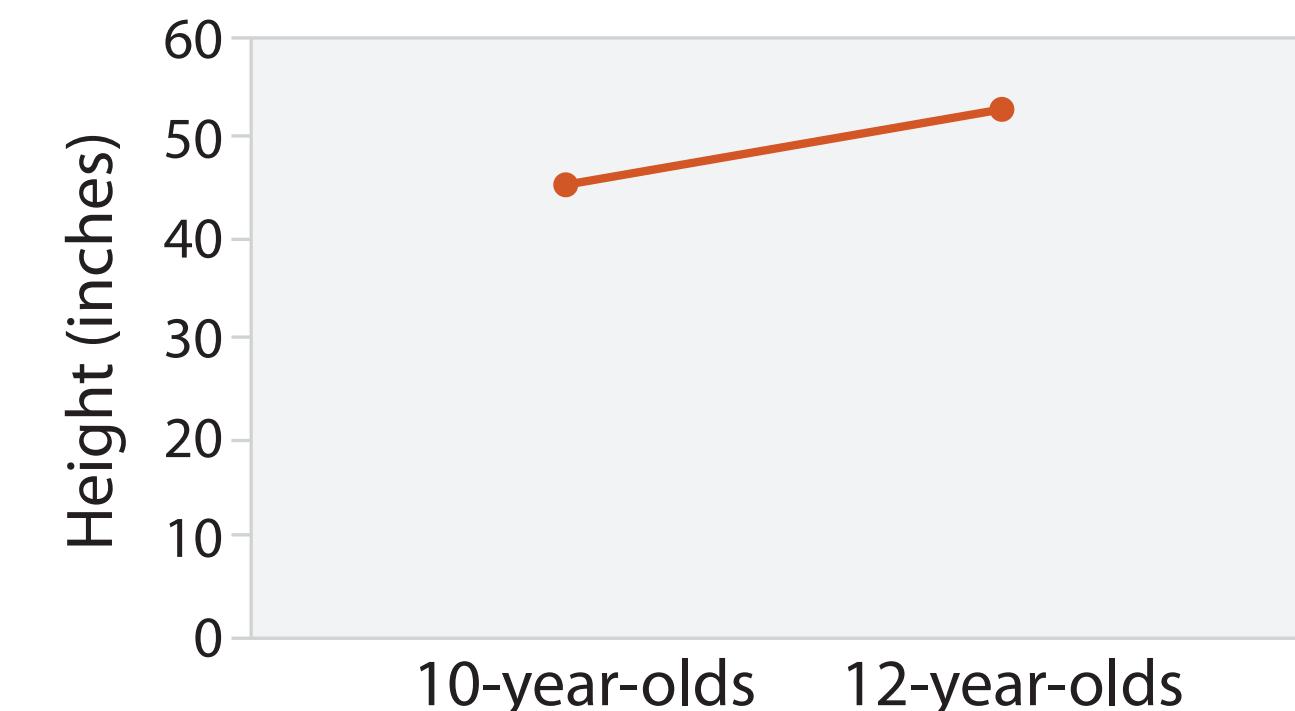
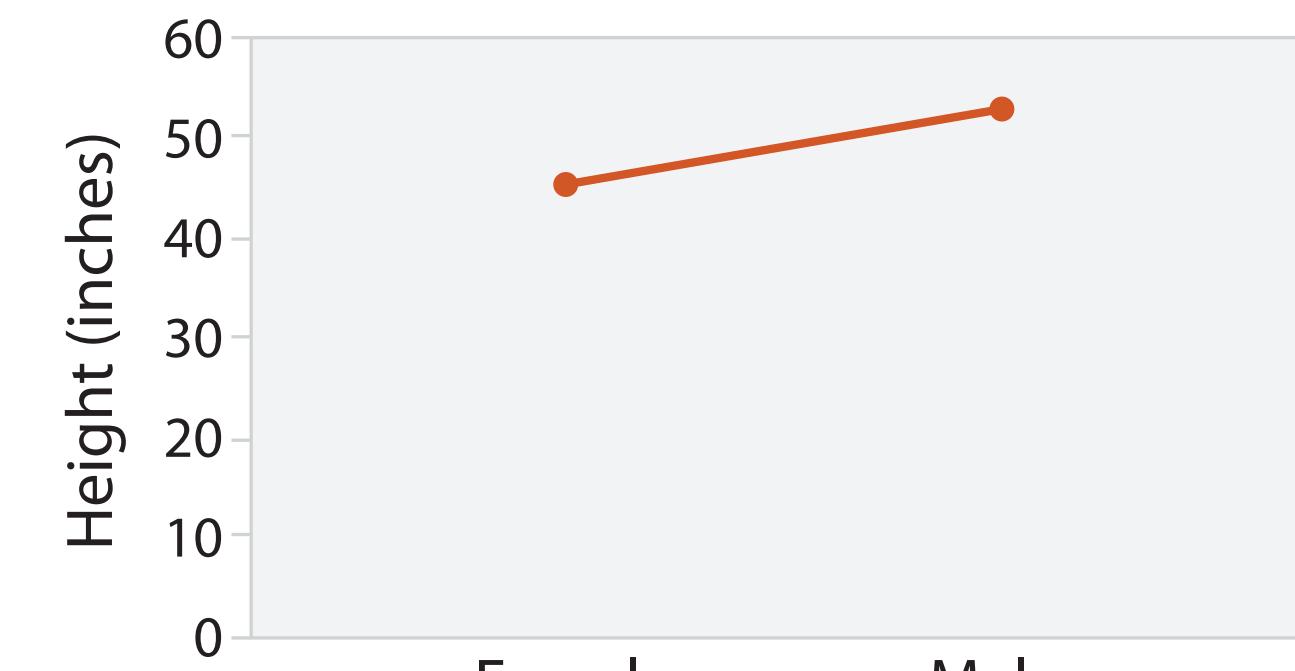
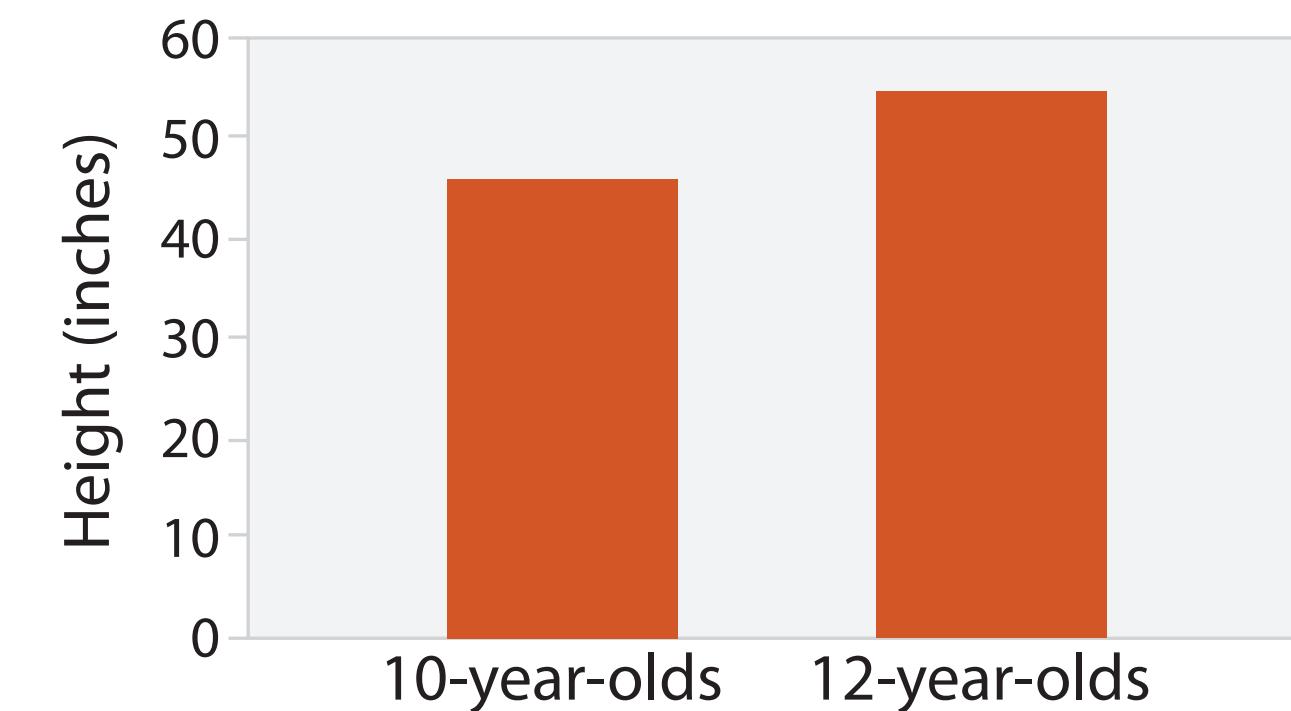
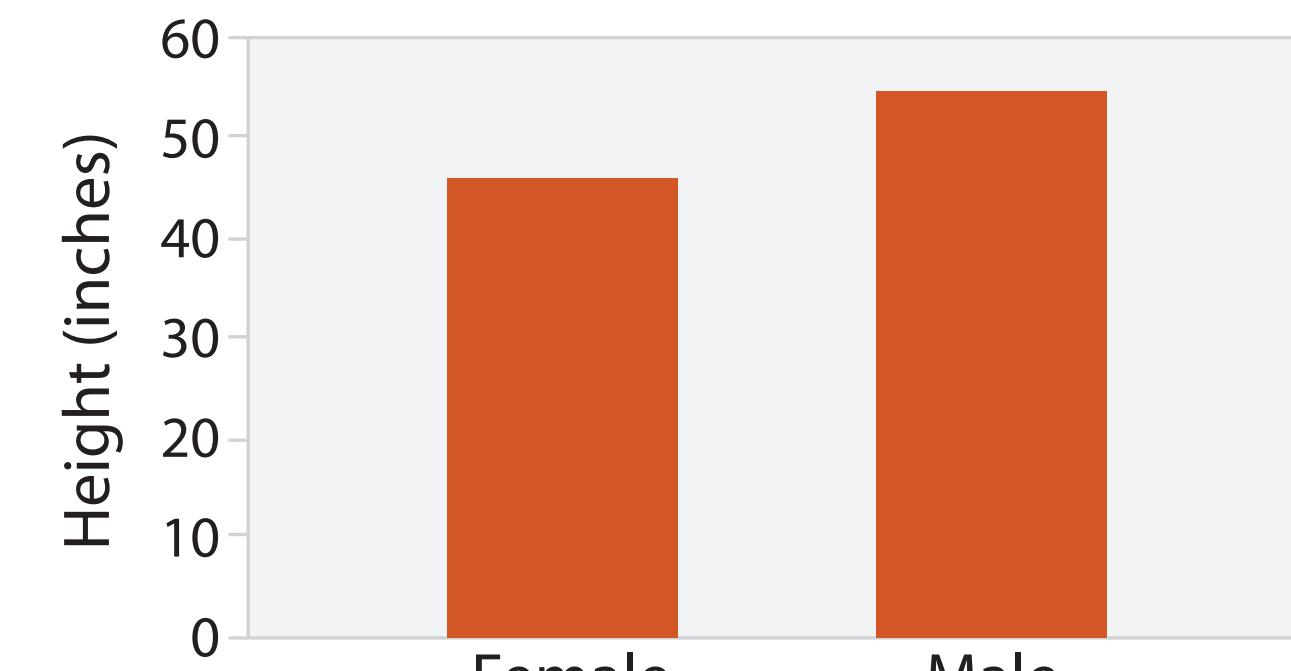
# Idiom: line chart / dot plot

- one key, one value
  - data
    - 2 quant attrs
  - mark: points
    - 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



# Choosing bar vs line charts

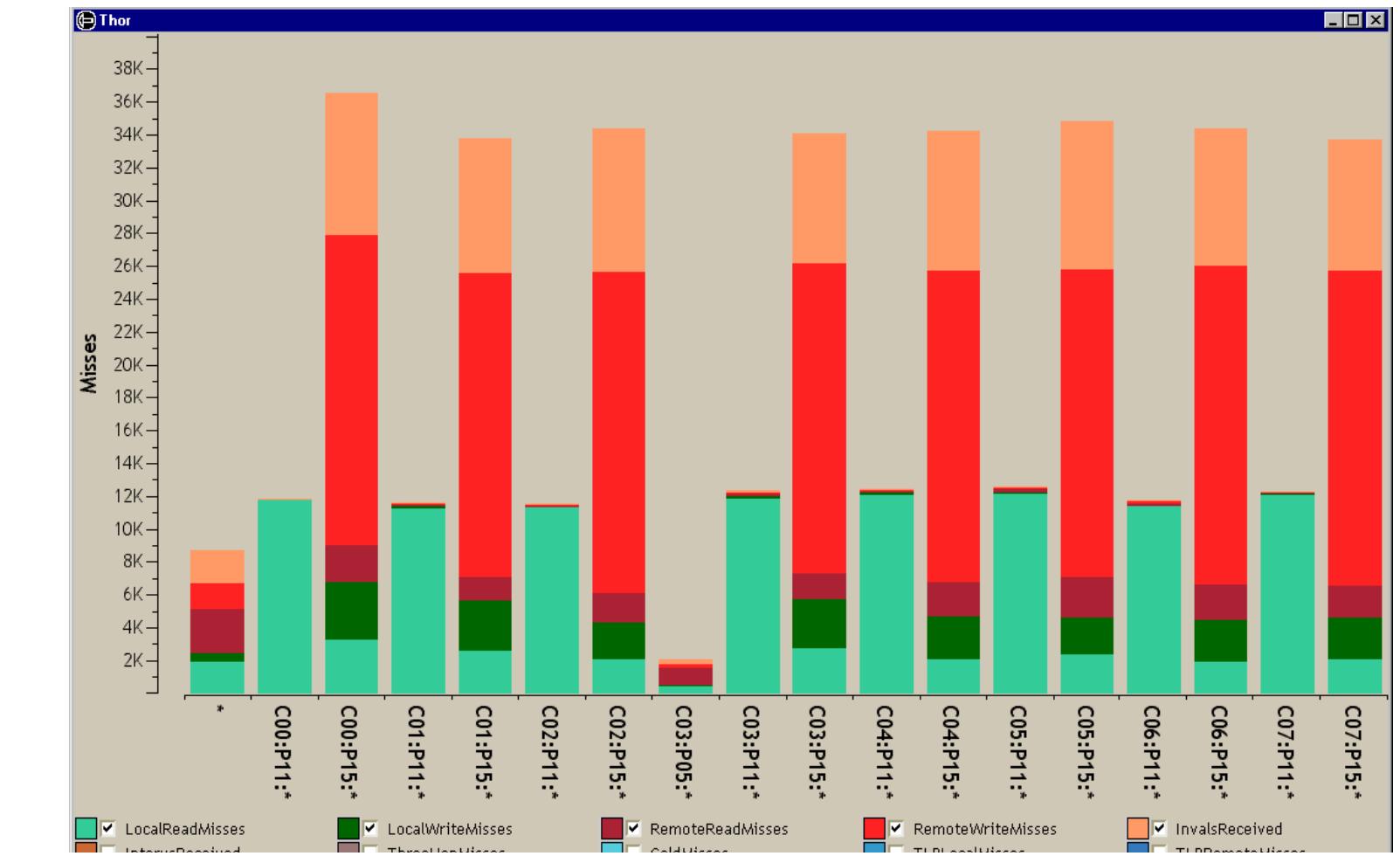
- depends on type of key attrib
  - bar charts if categorical
  - line charts if ordered
- do not use line charts for categorical key attrs
  - violates expressiveness principle
    - implication of trend so strong that it overrides semantics!
      - “The more male a person is, the taller he/she is”



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

# Idiom: stacked bar chart

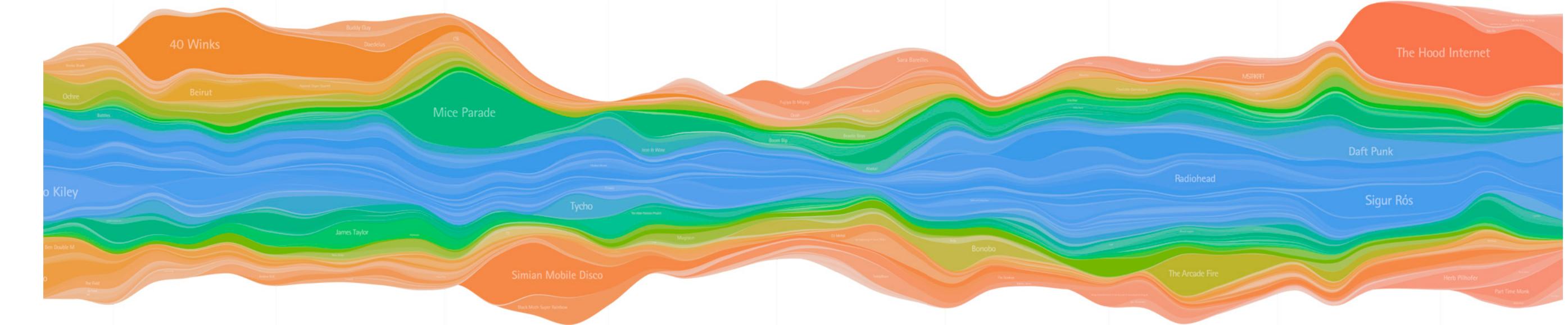
- one more key
  - data
    - 2 categ attrib, 1 quant attrib
  - mark: vertical stack of line marks
    - **glyph**: composite object, internal structure from multiple marks
  - channels
    - length and color hue
    - spatial regions: one per glyph
      - aligned: full glyph, lowest bar component
      - unaligned: other bar components
  - task
    - part-to-whole relationship
  - scalability
    - 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.]

# Idiom: streamgraph

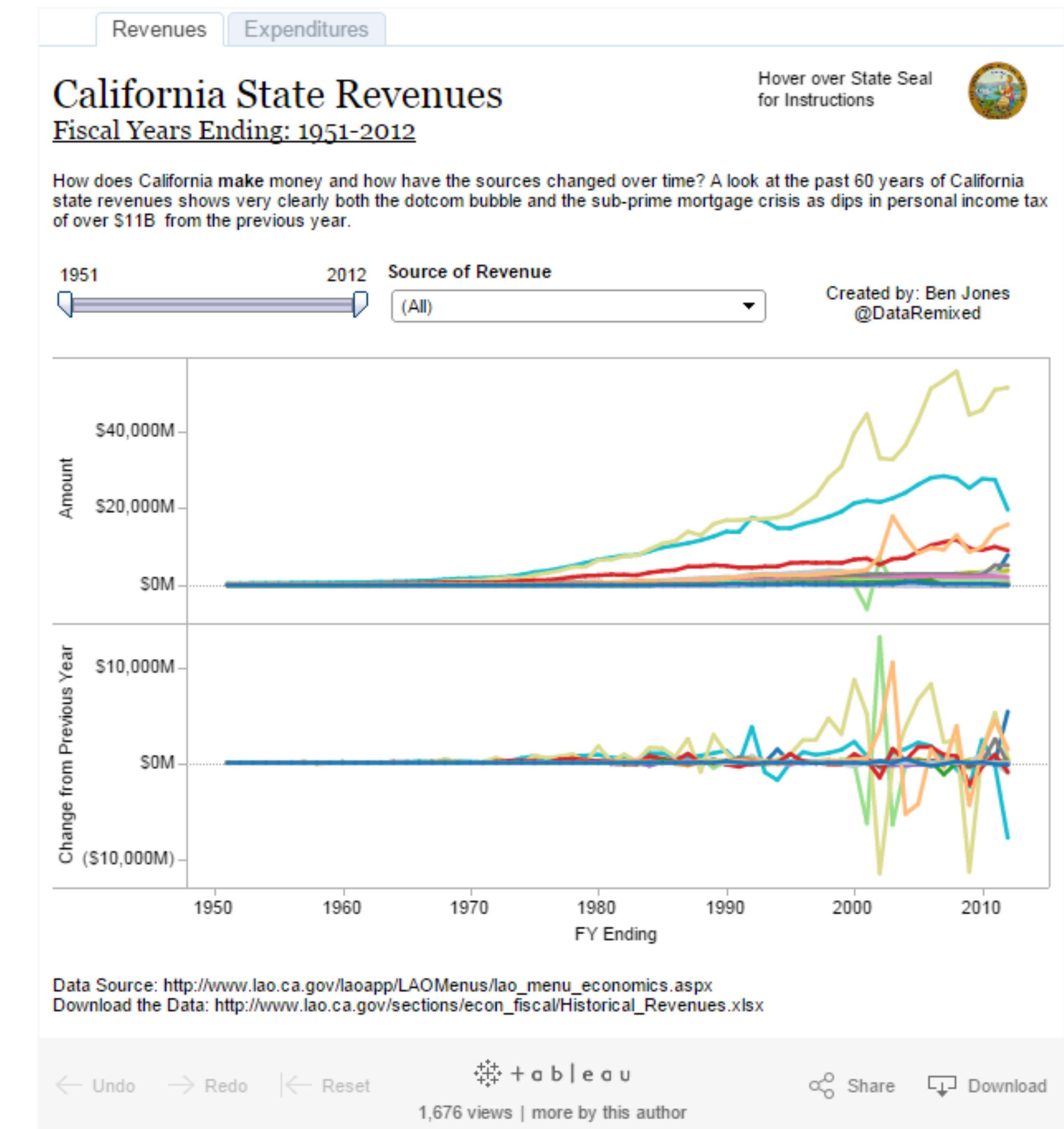
- generalized stacked graph
  - emphasizing horizontal continuity
    - vs vertical items
  - data
    - 1 categ key attrib (artist)
    - 1 ordered key attrib (time)
    - 1 quant value attrib (counts)
  - derived data
    - geometry: layers, where height encodes counts
    - 1 quant attrib (layer ordering)
  - scalability
    - hundreds of time keys
    - dozens to hundreds of artist keys
      - more than stacked bars, since most layers don't extend across whole chart



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

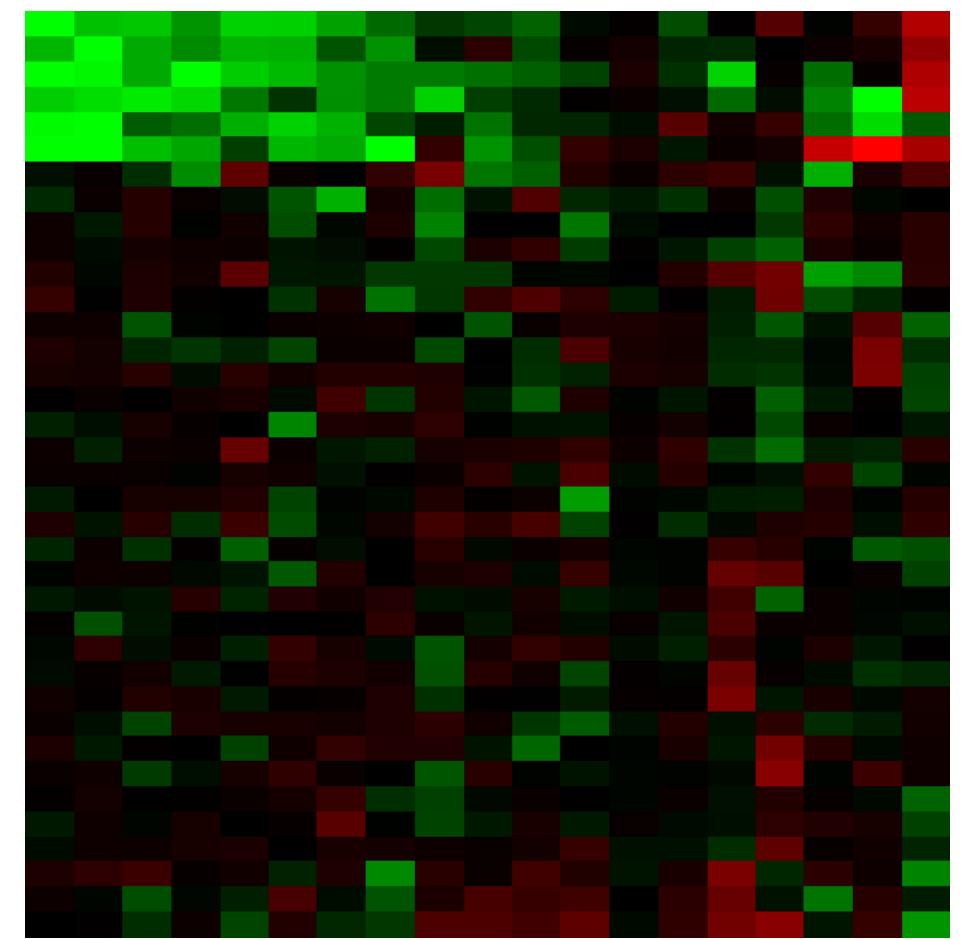
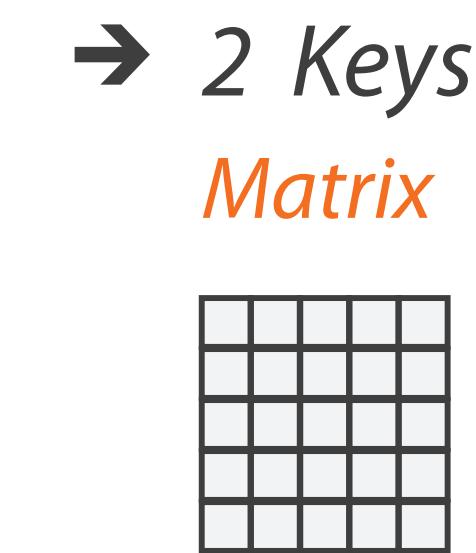
# Idiom: Indexed line charts

- data: 2 quant attires
  - 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



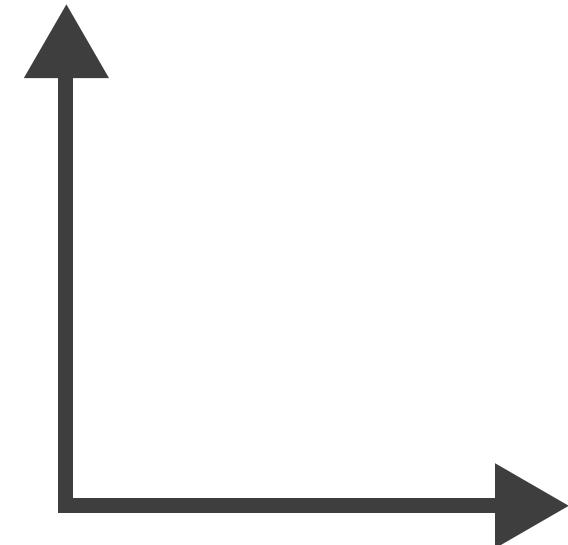
# Idiom: heatmap

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

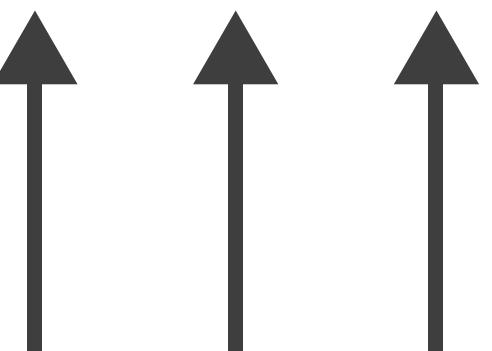


## → Axis Orientation

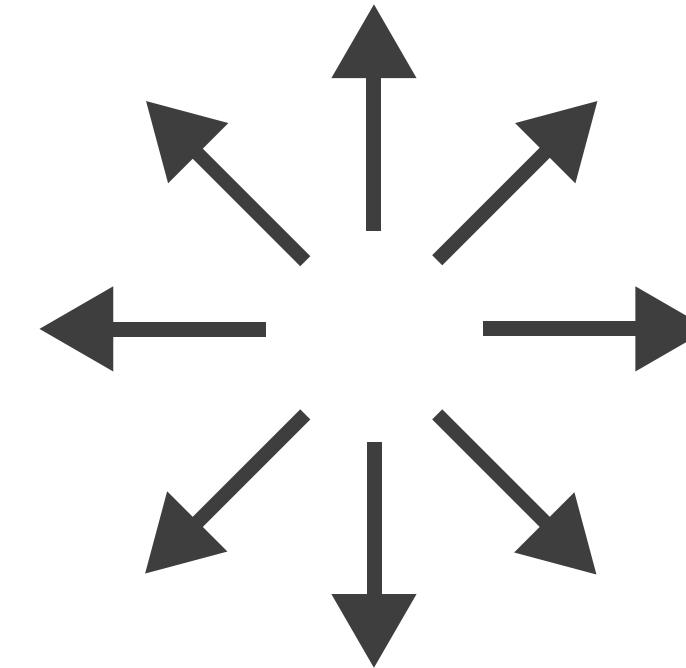
→ Rectilinear



→ Parallel

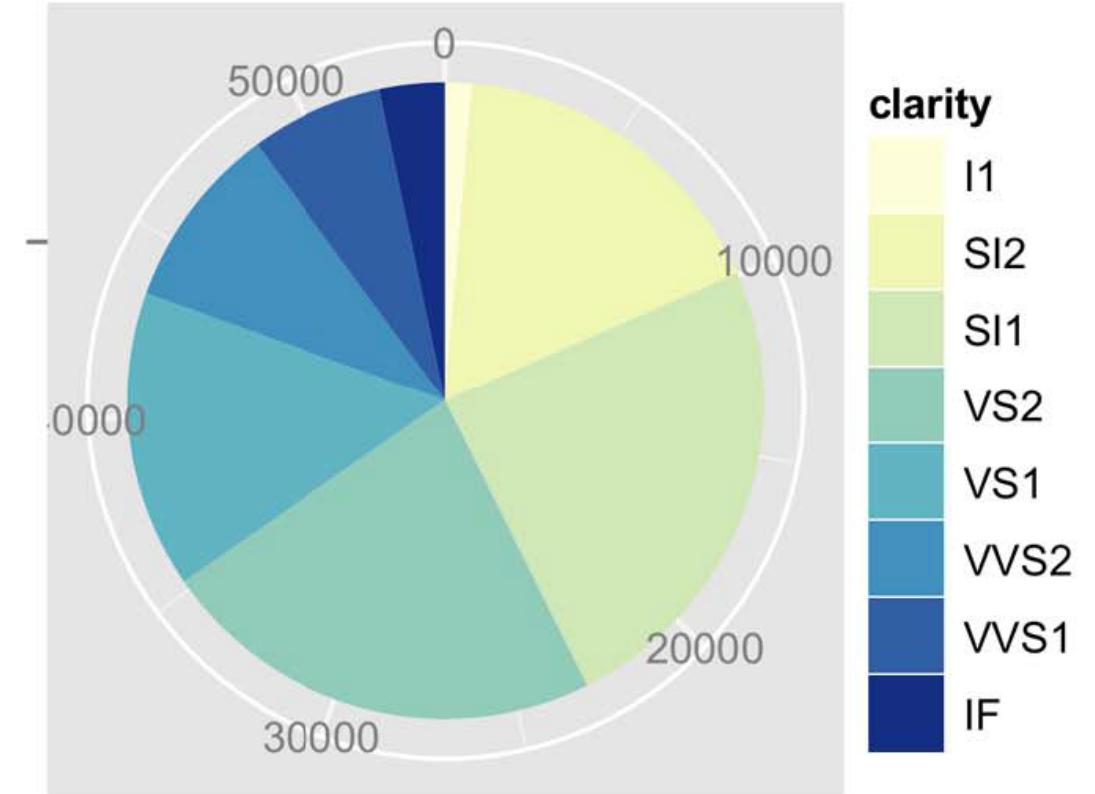


→ Radial

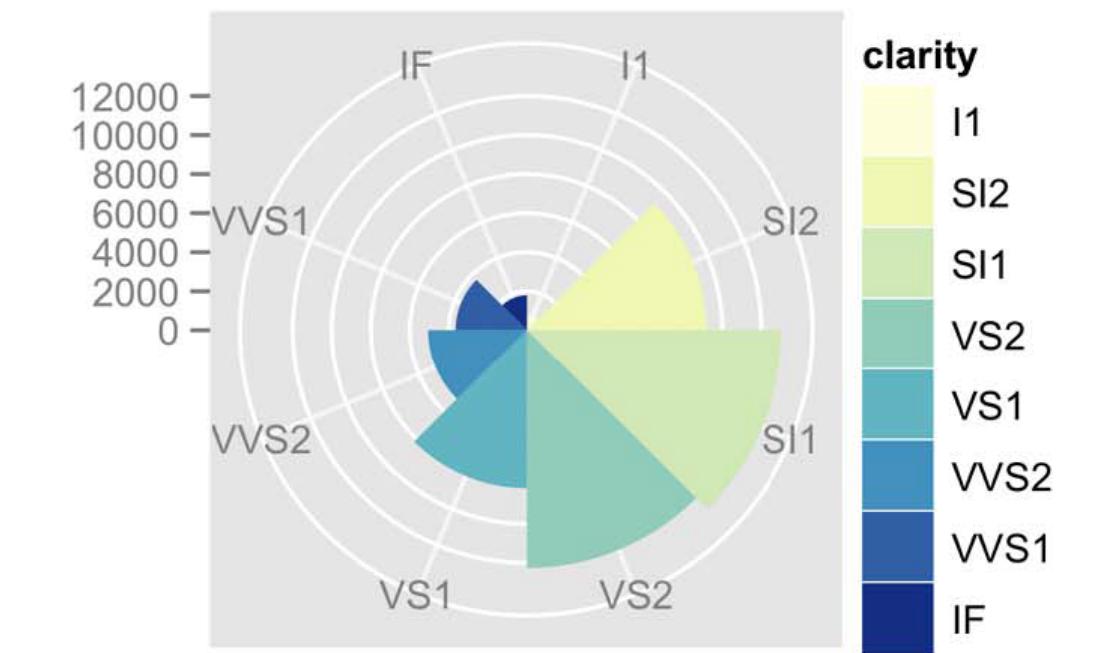


# Idioms: pie chart, polar area chart

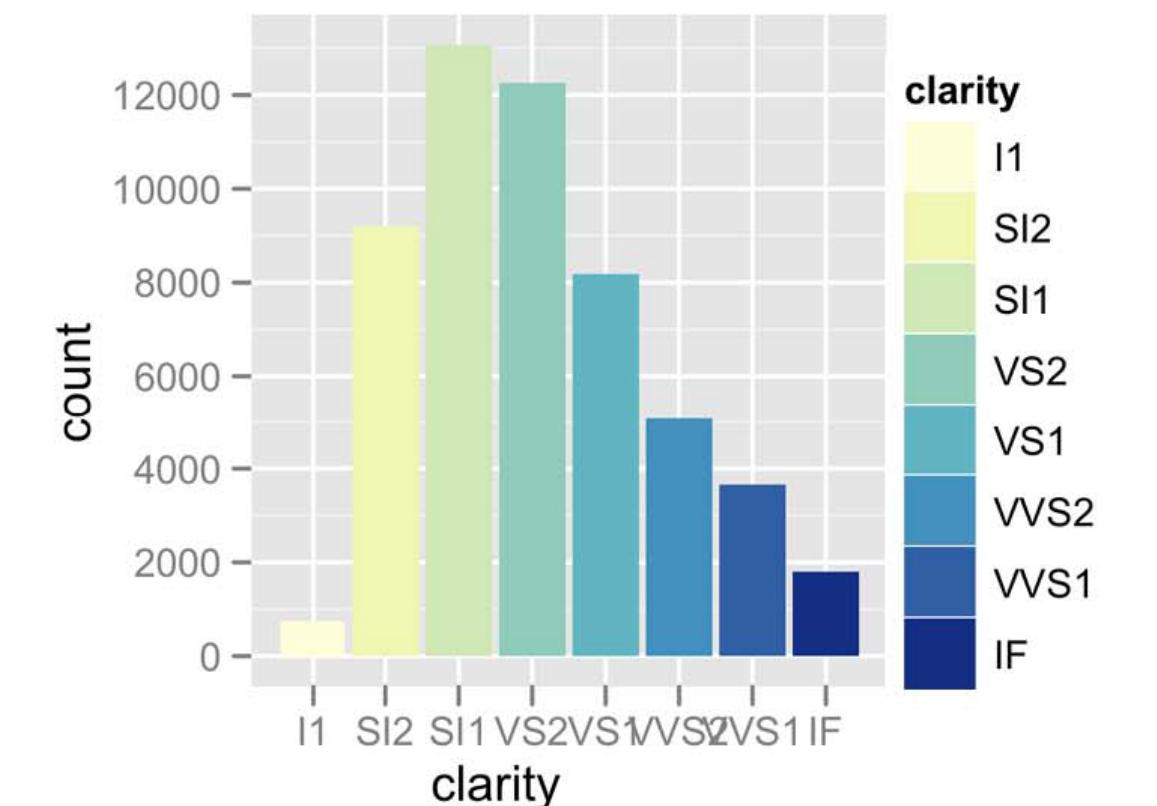
- pie chart
  - area marks with angle channel
  - accuracy: angle/area less accurate than line length
    - arclength also less accurate than line length



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

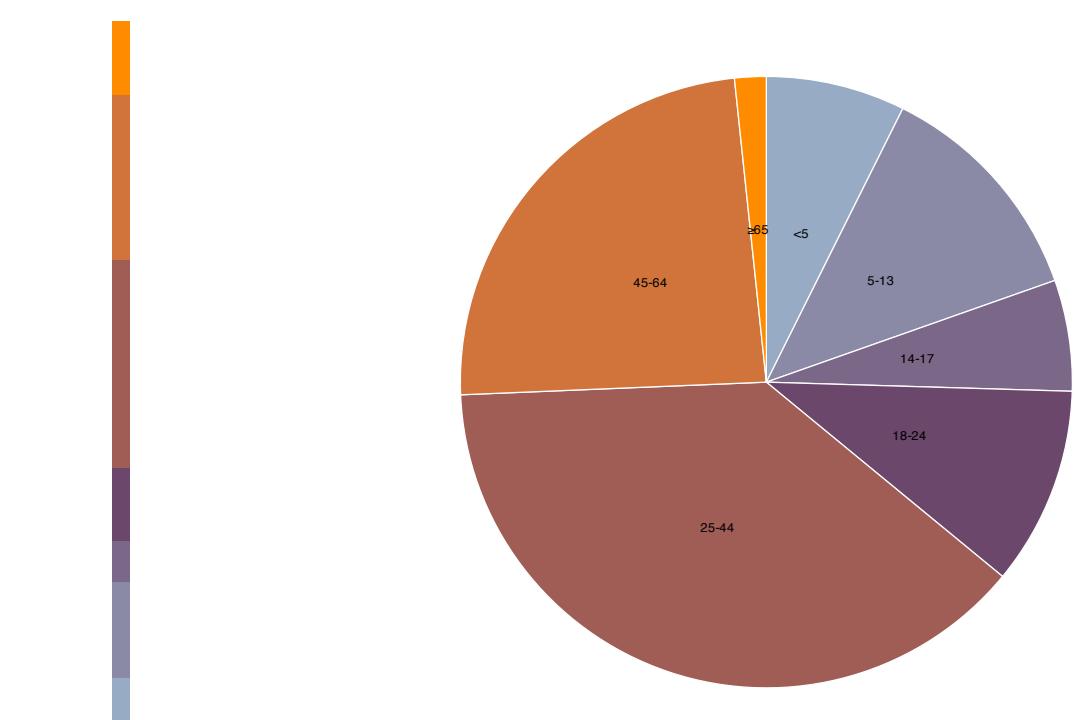
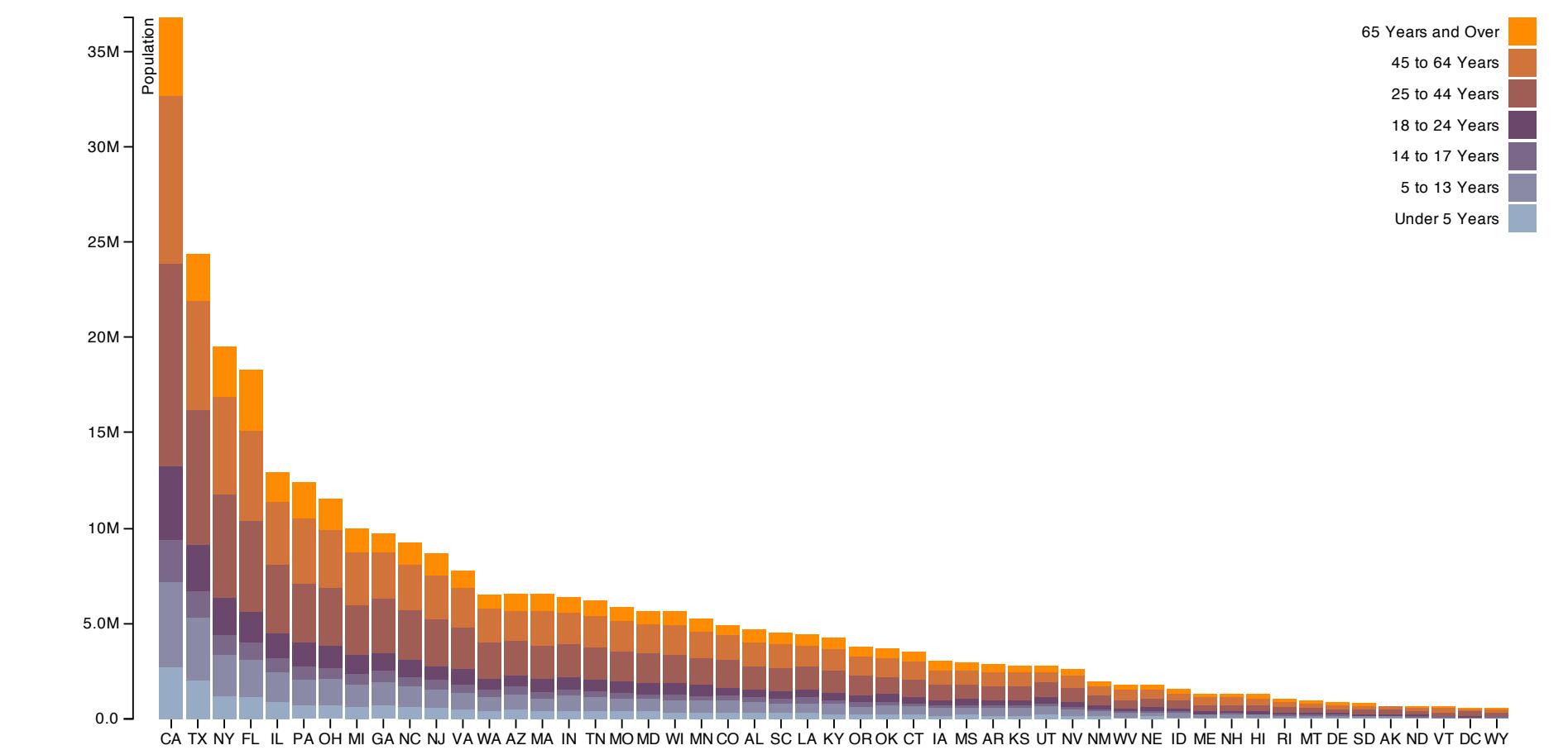
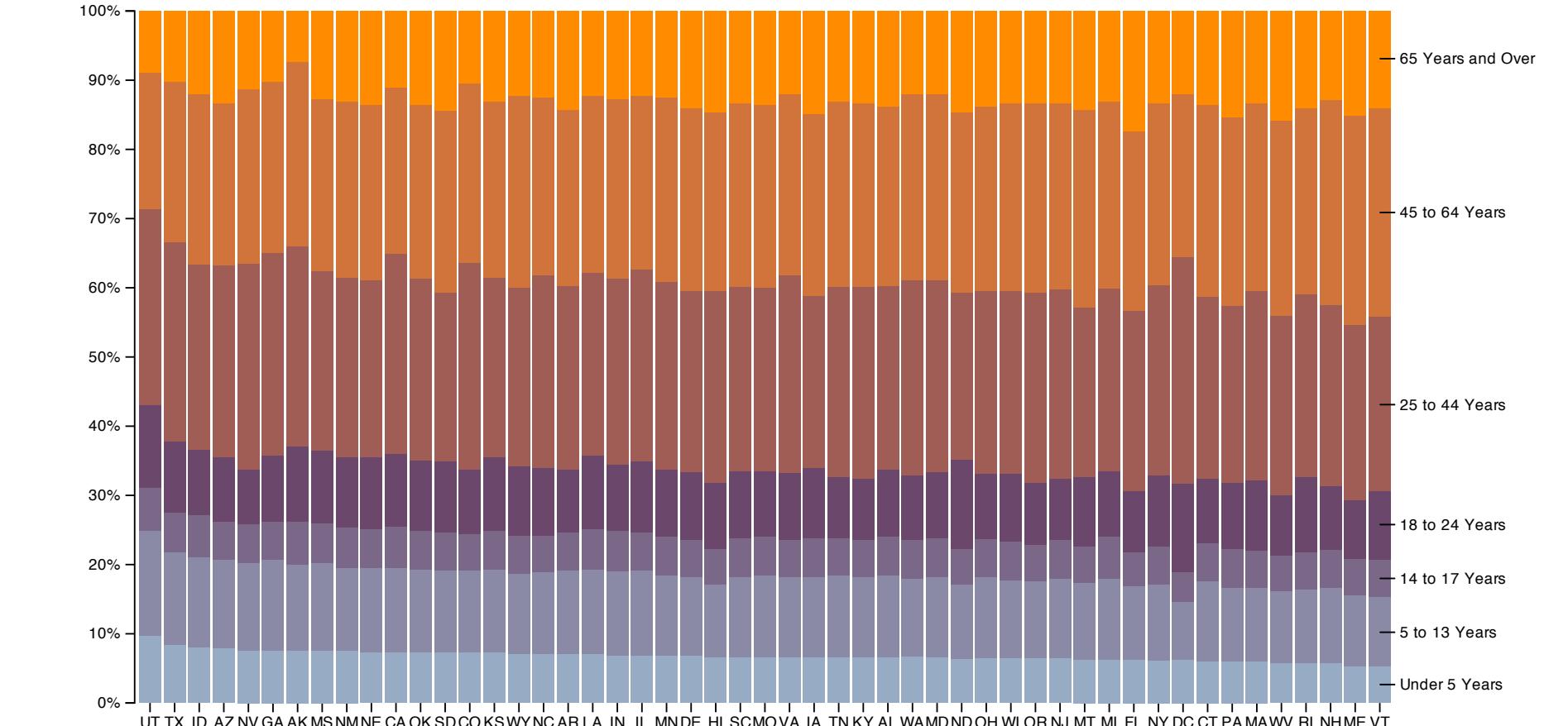


- data
  - 1 categ key attrib, 1 quant value attrib
- task
  - part-to-whole judgements



# Idioms: normalized stacked bar chart

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



<http://bl.ocks.org/mbostock/3887235>,  
<http://bl.ocks.org/mbostock/3886208>,  
<http://bl.ocks.org/mbostock/3886394>.

## Further reading

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

# Questions?

:::::::



@rschifan



schifane@di.unito.it



<http://www.di.unito.it/~schifane>