

# Spatial Layout: Arrange Tables



# How?

## Encode

- ④ Arrange
  - Express
  - Separate
- Order
  - Align
- Use

- ④ Map from **categorical** and ordered attributes
  - Color
    - Hue
    - Saturation
    - Luminance
  - Size, Angle, Curvature, ...
    - ■ ■ ■   | / \   | ) ) )
  - Shape
    - + ● ■ ▲
  - Motion
    - Direction, Rate, Frequency, ...

## Manipulate

- ④ Change

## Facet

- ④ Juxtapose

## Reduce

- ④ Filter

## Aggregate



## Select



## Partition



## Navigate



## Superimpose



## Embed



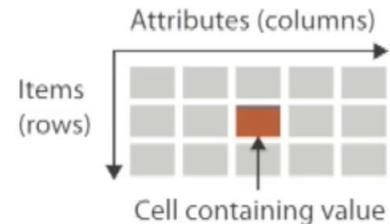
What?

Why?

How?



## → Tables



# Key and Value

## ○ In visualization

- Value:
  - dependent variables
  - Value you want to show in the visualization
- Key: independent variables
  - To index a value (Usually, categorical or ordinal)

1	day	apricot	blueberry	cherry
2	Mon	120	180	100
3	Tue	60	185	105
4	Wed	100	215	110
5	Thu	80	230	105
6	Fri	120	240	105

2 keys (2D): combination of (day, fruitName)  
can point to a value

## → Express Values

→ 1 Key

List



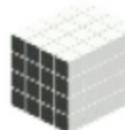
→ 2 Keys

Matrix



→ 3 Keys

Volume



→ Many Keys

Recursive Subdivision





# Key and Value

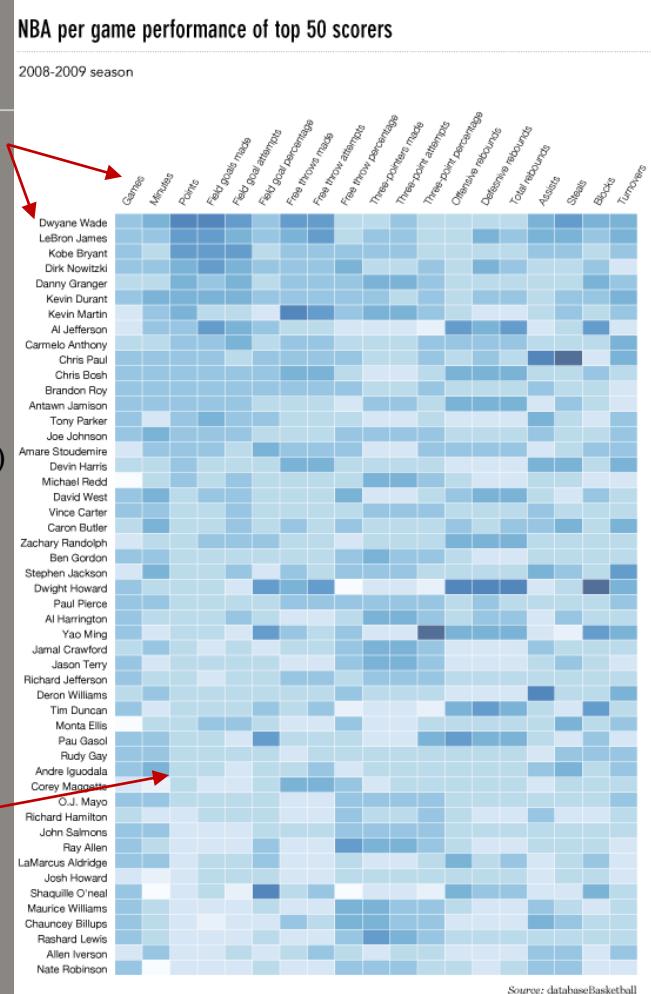


## In visualization

- Value:
  - dependent variables
  - Value you want to show in the visualization
- Key: independent variables
  - To index a value (Usually, categorical or ordinal)

Player	Games	Minutes	Points	Rebounds	.....
James	80	44	28	9	...
Wade	75	32	25	6	...
...	...	...	...	...	...

2 keys





S05-01



# Arrange Table Data

## ④ Express Values



## ④ Separate, Order, Align Regions

→ Separate



→ Order



→ Align



## ④ Axis Orientation

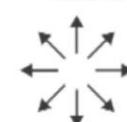
→ Rectilinear



→ Parallel



→ Radial



## ④ Layout Density

→ Dense



→ Space-Filling



→ 1 Key

*List*



→ 2 Keys

*Matrix*



→ 3 Keys

*Volume*



→ Many Keys

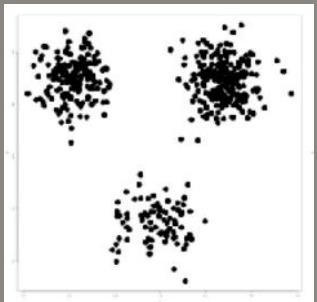
*Recursive Subdivision*



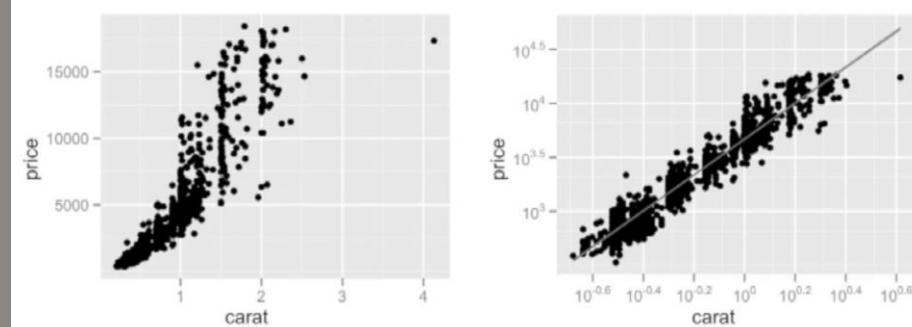


# Idiom: scatterplot

- Express values
  - Quantitative attribute
- No keys, only values
  - Data
    - 2 quantitative attributes
  - Mark:
    - points
  - Channels
    - Horizontal + vertical position
  - Tasks
    - Find trends, outliers, distribution, correlation, clusters
  - Scalability
    - Hundreds of items



**Scalability:** how many data you can show before this visualization is ineffective



you do not care which date of the data points,  
But correlation, outlier.....

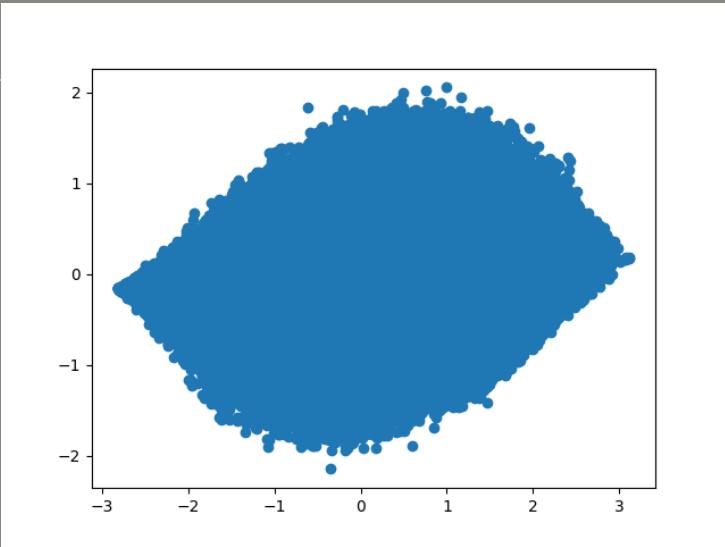
date	carat	price
Jan-20	1	100
Jan-20	2	250
Jan-20	3	400
Jan-20	4	800
20-Feb	1	125
20-Feb	2	250
20-Feb	3	700
20-Feb	4	1200
...	...	...



## Idiom: scatterplot

- 🟡 Express values
  - Quantitative attribute
- 🟡 No keys, only values
  - Data
    - 🟡 2 quantitative attributes
  - Mark:
    - 🟡 points
  - Channels
    - 🟡 Horizontal + vertical position
  - Tasks
    - 🟡 Find trends, outliers, distribution, correlation, clusters
  - Scalability
    - 🟡 Hundreds of items

**Scalability:** how many data you can show before this visualization is ineffective



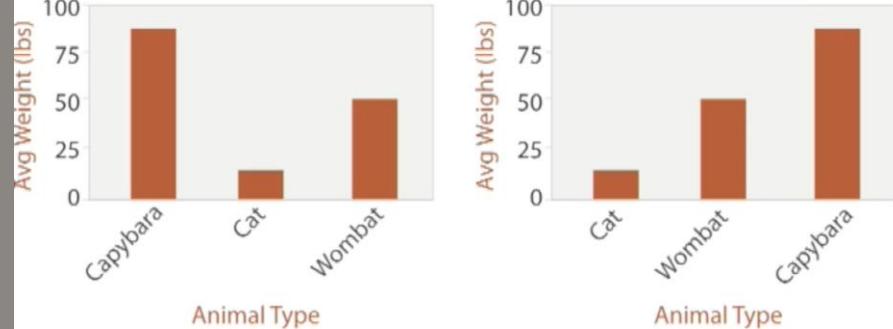
Too many points. It can not show more details



## Idiom: bar chart



- One key, one value
  - Data
    - 1 categorical attribute (key), 1 quantitative attribute
  - Mark: lines
  - Channels
    - Length to express quantitative value
    - Spatial region: one per mark
      - Separated horizontally, aligned vertically
      - Ordered by quantitative attribute
        - By label (alphabetical) , by length attribute(data-driven)
  - Task
    - Compare, lookup values
  - Scalability
    - Dozens to hundreds of levels for key attribute





### How do we order the x-axis??

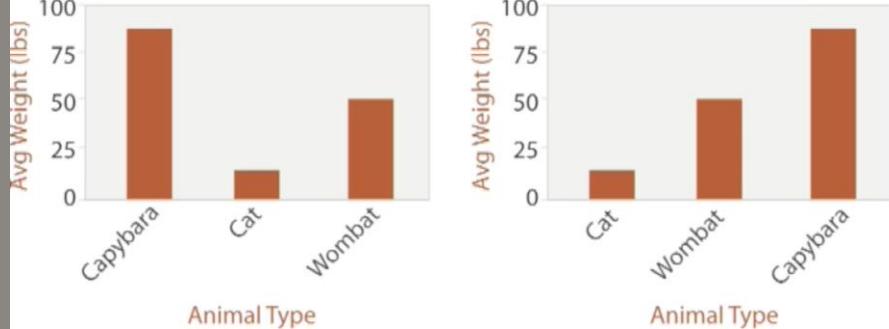
Visual channel x-axis: magnitude channel  
Data attribute on x-axis: categorical

## Idiom: bar chart



### One key, one value

- Data
  - 1 categorical attribute (key), 1 quantitative attribute
- Mark: lines
- Channels
  - Length to express quantitative value
  - Spatial region: one per mark
    - Separated horizontally, aligned vertically
    - Ordered by quantitative attribute
      - By label (alphabetical), by length attribute(data-driven)
- Task
  - Compare, lookup values
- Scalability
  - Dozens to hundreds of levels for key attribute



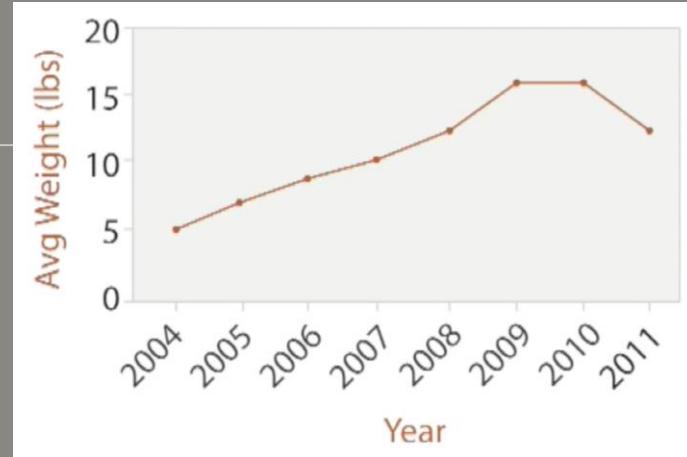


S05-02



## Idiom: line chart / dot plot

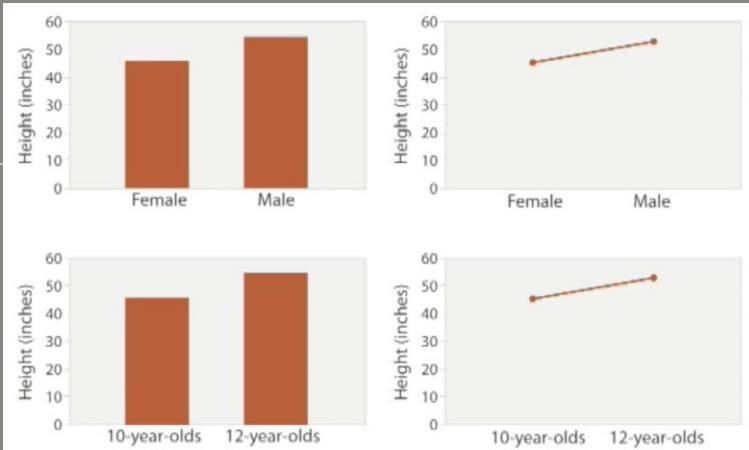
- One key, one value
  - Data
    - 2 quantitative attributes
  - Mark:
    - points
    - Line connection marks between them
  - Channels
    - Aligned lengths to express quantitative value
    - Separated and ordered by key attribute into horizontal regions
  - Task
    - Find trend: connection marks (perceptually) **emphasize** ordering of items along key axis by explicitly showing relationship one item and the next
  - Scalability
    - Hundreds of key levels, hundreds of value levels





## Bar Chart vs Line Chart

- ➊ Depends on type of key attribute
  - Bar charts if categorical
  - Line chart if ordered
- ➋ 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??

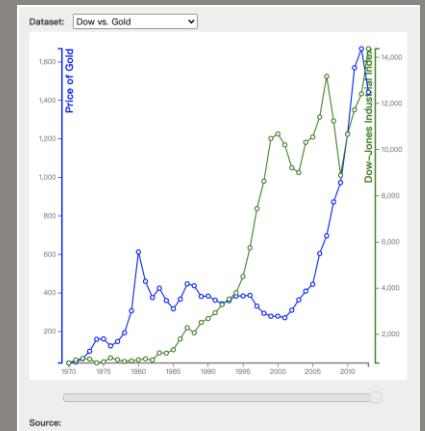
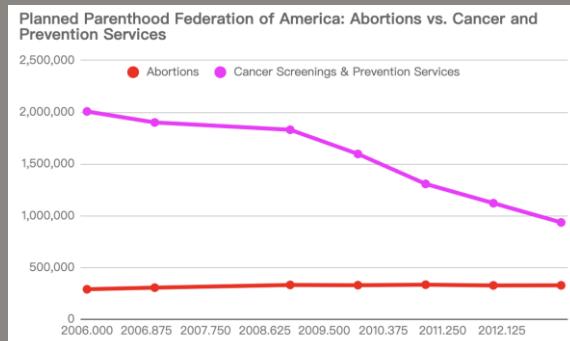
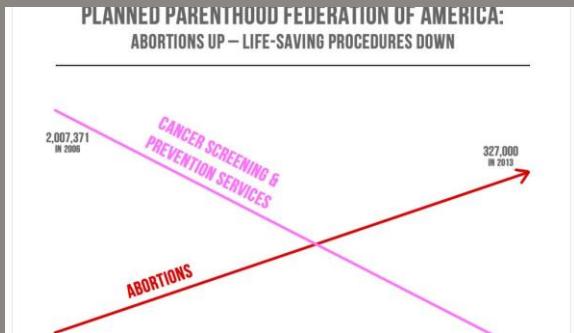


The line char implicitly interpolate the values between two points



# Dual Axis Line Chart

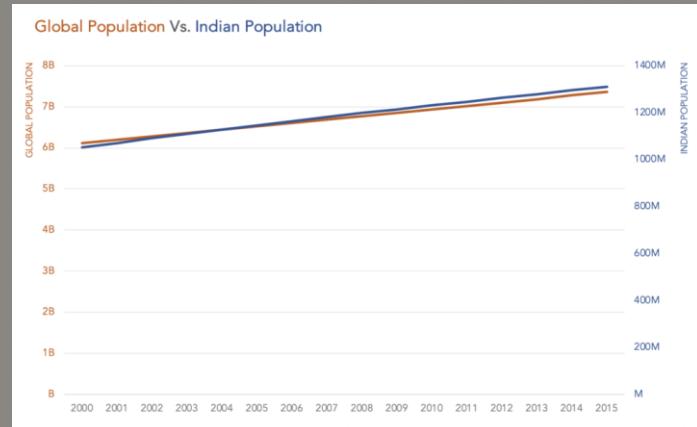
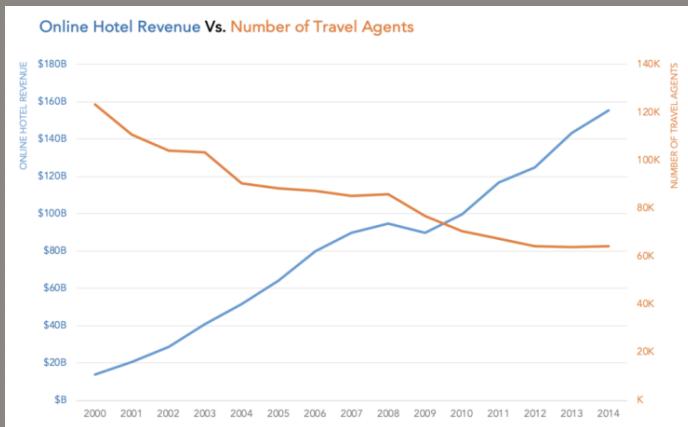
⌚ Axes are critical!!!





# Dual Axis Line Chart

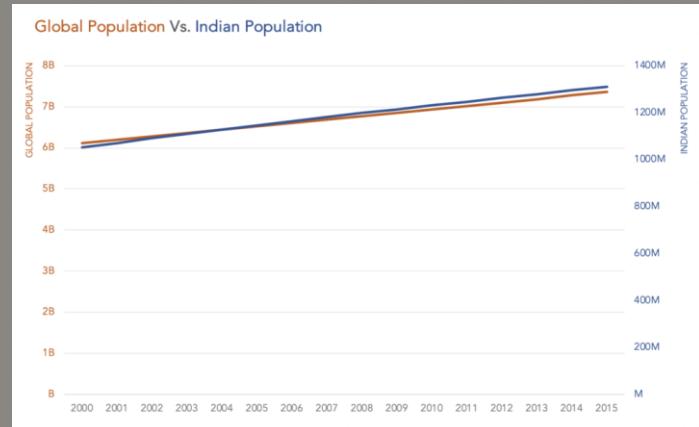
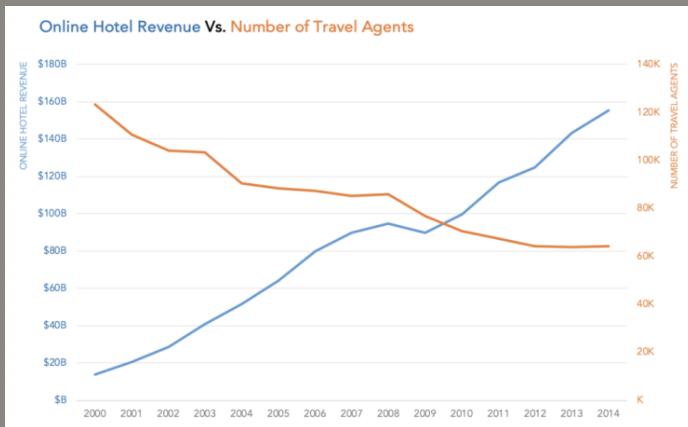
- Axes are critical!!!
- For the axes
  - Same unit?
  - Same scale?





# Dual Axis Line Chart

- ➊ What users may focus on when they look at 2 lines in a line chart?
  - Crossing point
  - Correlation/trend/rate





S05-03



## Dual Axis Line Chart

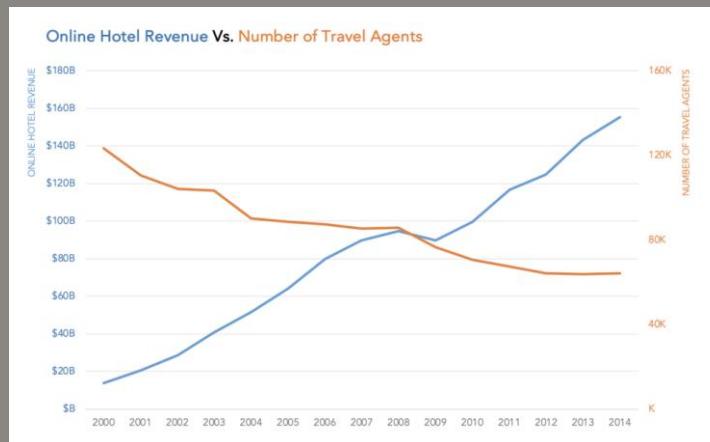
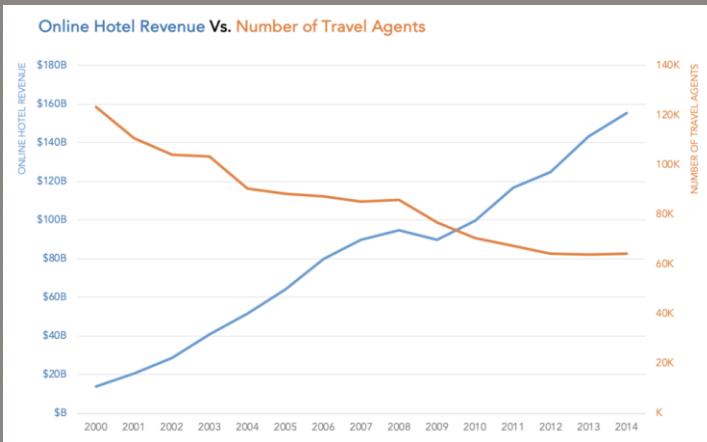
- ⌚ What users may focus on when they look at 2 lines in a line chart?
  - Crossing point
  - Correlation/trend/rate





# Dual Axis Line Chart

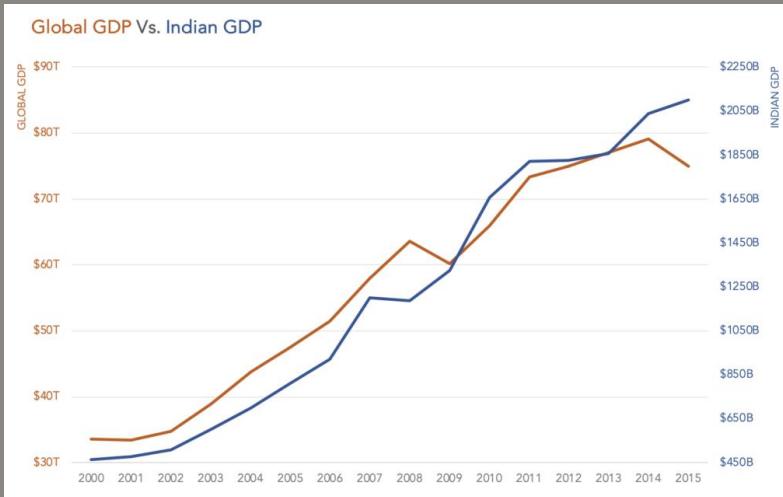
- What users may focus on when they look at 2 lines in a line chart?
  - **Crossing point:** we can simply change the scale of axis to get different crossing point (because unit/scale are different in this case)
  - Correlation/trend/rate





## Dual Axis Line Chart

- ⌚ What users may focus on when they look at 2 lines in a line chart?
  - Crossing point
  - Correlation/trend/rate

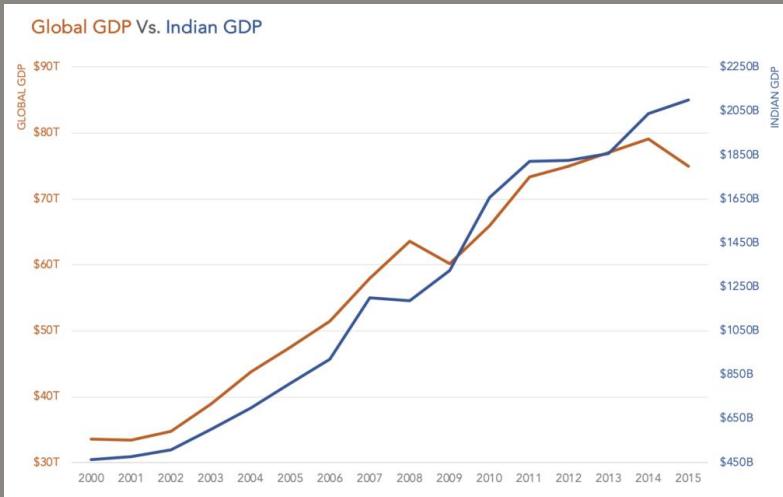


Global GDP and Indian GDP have almost same growing rate?



## Dual Axis Line Chart

- ⌚ What users may focus on when they look at 2 lines in a line chart?
  - Crossing point
  - Correlation/trend/rate



Global GDP and Indian GDP have almost same growing rate?

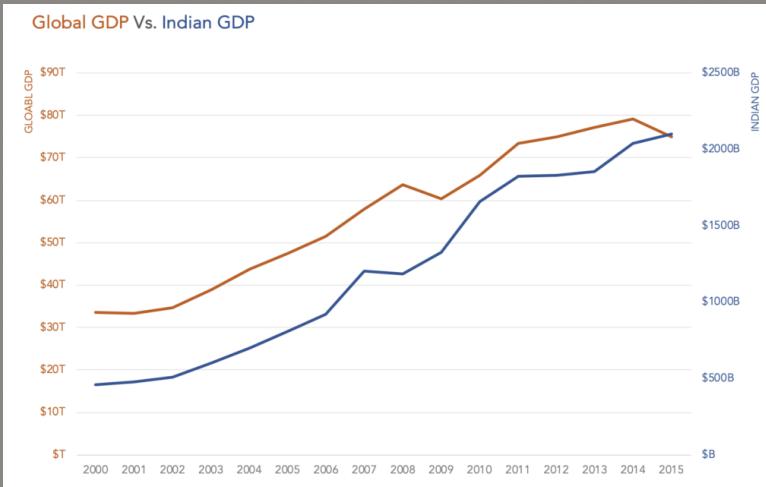
No, Indian has a much higher growing rate

Note: the baselines are not zero



## Dual Axis Line Chart

- ⌚ What users may focus on when they look at 2 lines in a line chart?
  - Crossing point
  - Correlation/trend/rate



Zero baseline?

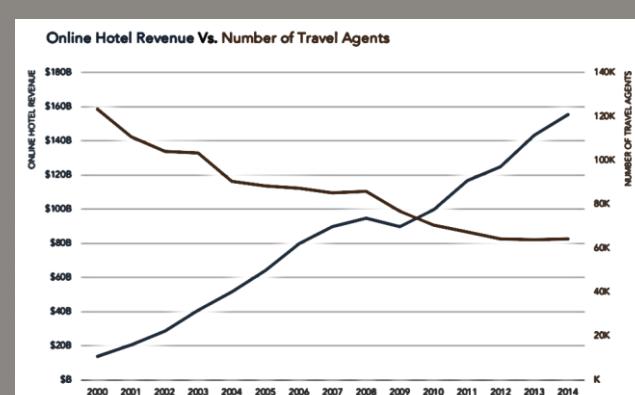
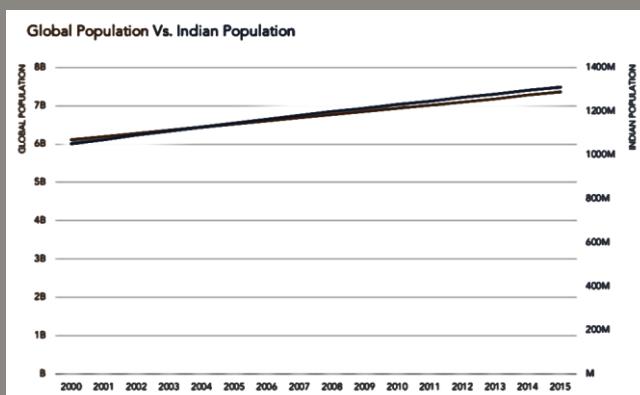
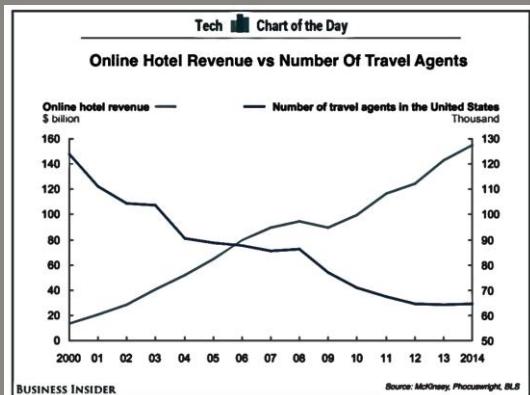
Does the crossing point mean anything?  
Indian GDP catches up global GDP in 2015?

Log scale on y-axis?



# Dual Axis Line Chart

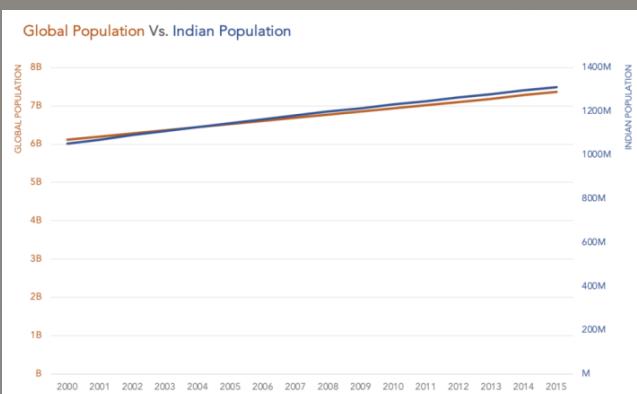
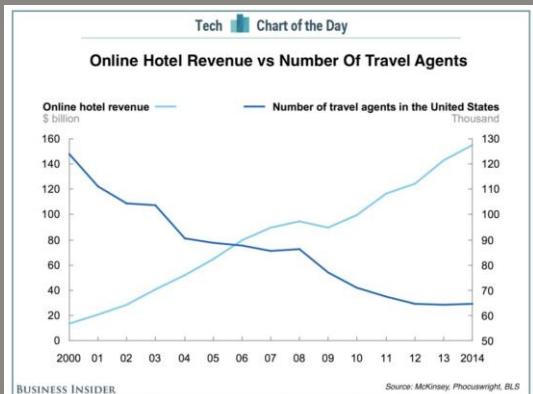
- 🟡 Not intuitive to know a line belongs to which axis?
- 🟡 We could use different color to pair them





# Dual Axis Line Chart

- 🟡 Not intuitive to know a line belongs to which axis?
- 🟡 We could use different color to pair them





## Dual Axis Line Chart

- ◉ When you look at a dual axis line chart
  - Too many ways to mislead you (un/intentionally)
  - Check the axes first: unit, scale, zero baseline or not
  
- ◉ When you draw a dual axis line chart
  - Too many ways to mislead users
  - If a crossing point exist, does the crossing points convey the proper information
  - If the line chart shows a special pattern (e.g. correlation), is that what you want?
  - Visually pair the line to the corresponding axis



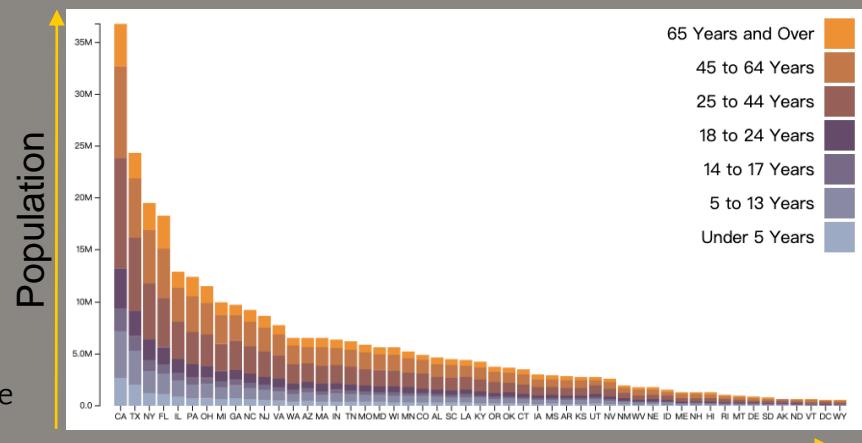
# Idiom: Stacked bar chart



- One more key
  - Data
    - 2 categorical/ordinal attr, 1 quantitative attr
  - 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
    - Compare bars at bottom (close to the baseline)
  - Scalability
    - Several to one dozen levels for stacked attribute

	< 5	5 - 13	14 - 17	...
AZ	...	...	...	...
CA	...	...	...	...
TX	...	...	...	...
...	.....	.....	...	...

Different color bars: age group



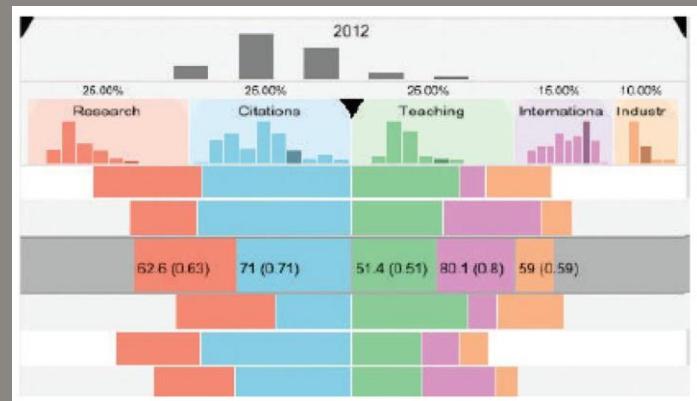
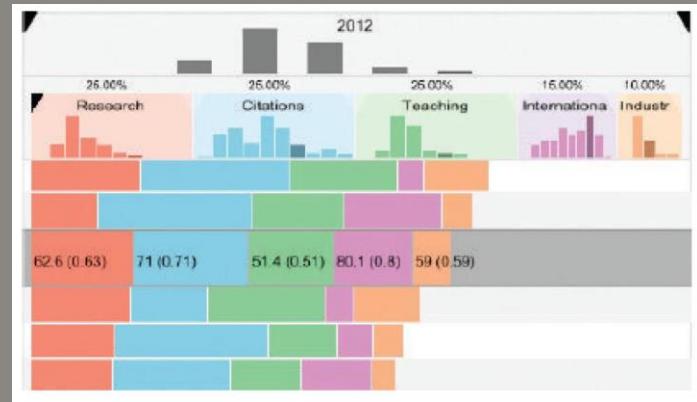


# Idiom: Stacked bar chart



- One more key
  - Data
    - 2 categorical/ordinal attr, 1 quantitative attr
  - 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
    - Compare bars at bottom (close to the baseline)
  - Scalability
    - Several to one dozen levels for stacked attribute

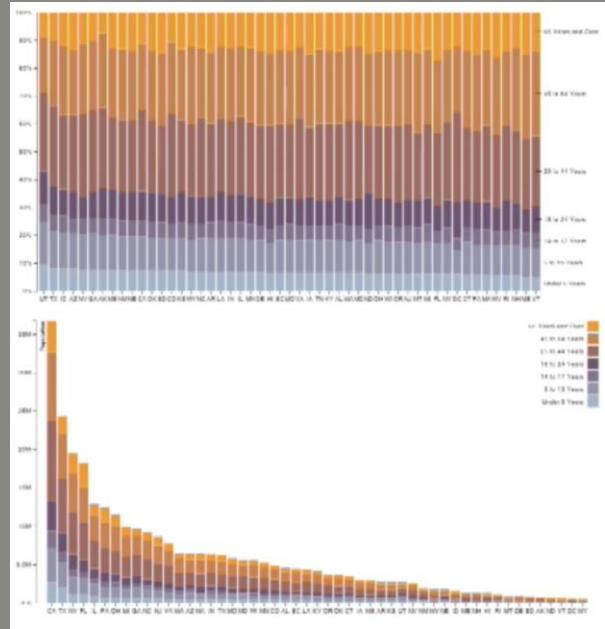
Choice of the baseline?





# Idiom: Normalized Stacked Barchart

- Task
  - Part-to-whole judgements
  - compare ratio among same categories
- Normalized stacked bar chart
  - Stacked bar chart, normalized to full vertical height



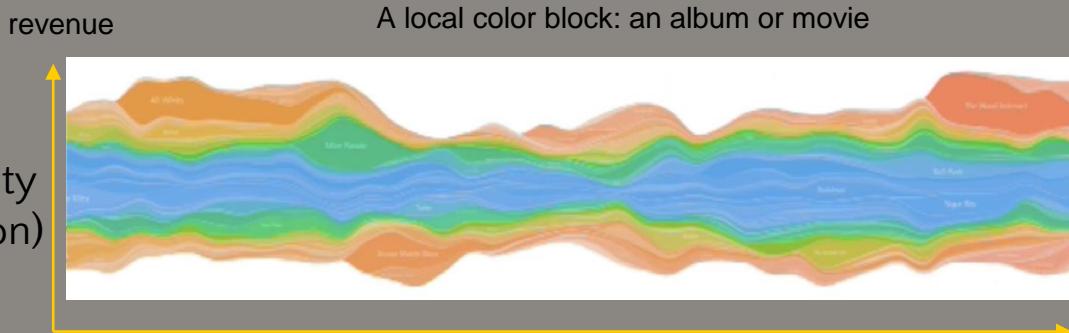


# Idiom: Streamgraph



## Generalized stacked graph

- Emphasizing horizontal continuity  
(create block size for comparison)
  - vs vertical item



- Data
  - 1 categorical key attribute
  - 1 ordered key attribute
  - 1 quantitative value attribute

- Task
  - Part-to-whole relationship over x-axis

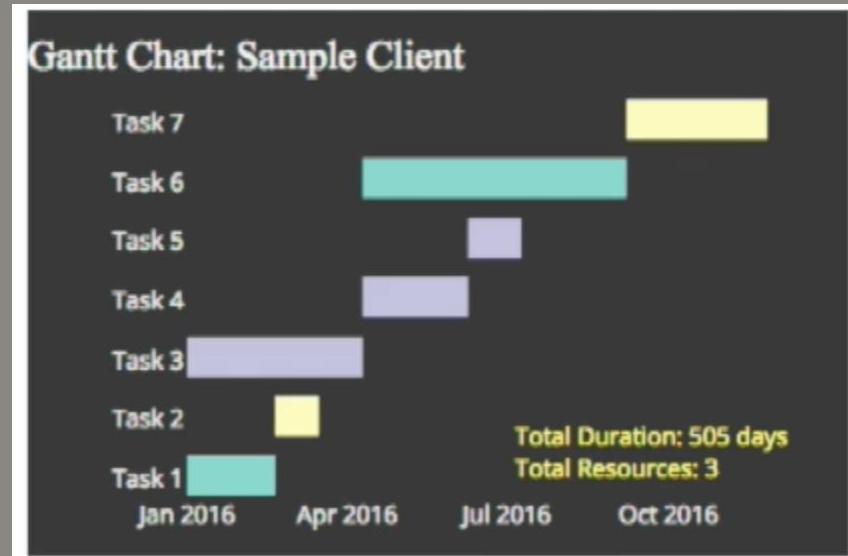
- Scalability
  - Hundreds of keys (x-axis direction)
  - Dozens to hundreds of keys (y-axis direction)
    - More than stacked bars, since most layers do not extend across whole chart

D3: <https://www.d3-graph-gallery.com/streamgraph>



## Idiom: Gantt charts

- ➊ One key, two (related) values
  - Data
    - 1 categorical attr, 2 quantitative attr
  - Mark: line
    - Length: duration
  - Channels
    - Horizontal position: start/end time
    - Horizontal length: duration
  - Task
    - Emphasize **temporal overlaps**, start/end dependencies between items
  - Scalability
    - Dozens of key levels
    - Hundreds of value levels



D3: <http://bl.ocks.org/dk8996/5449641>



# Idiom: Slopegraph



## Two values

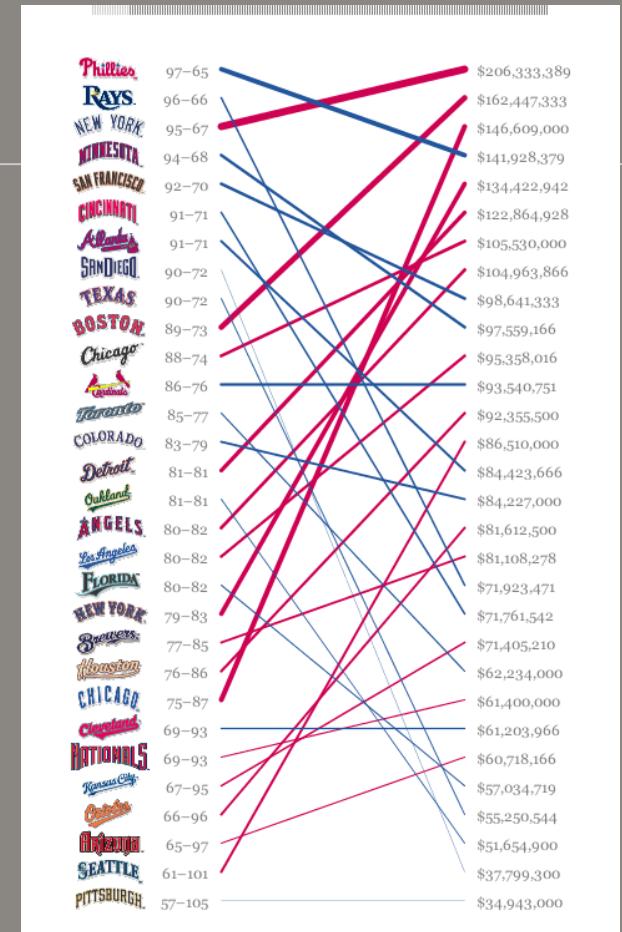
- Data
  - 2 quantitative values attributes
  - 1 derived attribute: change magnitude

- Mark: point, line
  - Line connecting mark between points

- Channels
  - 2 vertical position: express attribute value
  - linewidth/size, color

- Task
  - Use the “slope” to emphasize changes in rank/value

- Scalability
  - Hundreds of value levels



Use the slope of lines to emphasize the ranks difference of two attributes of the same item



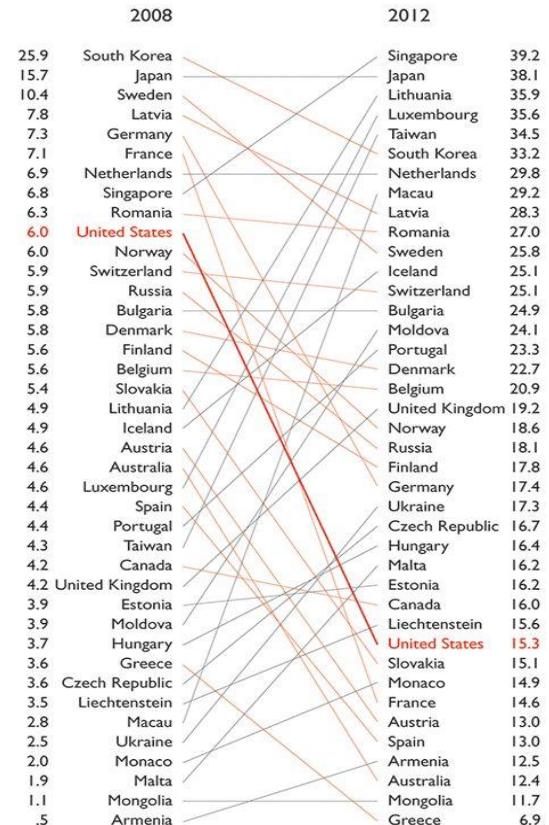
# Idiom: Slopegraph



## Two values

- Data
  - 2 quantitative values attributes
  - 1 derived attribute: change magnitude
- Mark: point, line
  - Line connecting mark between points
- Channels
  - 2 vertical position: express attribute value
  - linewidth/size, color
- Task
  - Use the “slope” to emphasize changes in rank/value
- Scalability
  - Hundreds of value levels

Internet download speeds by country



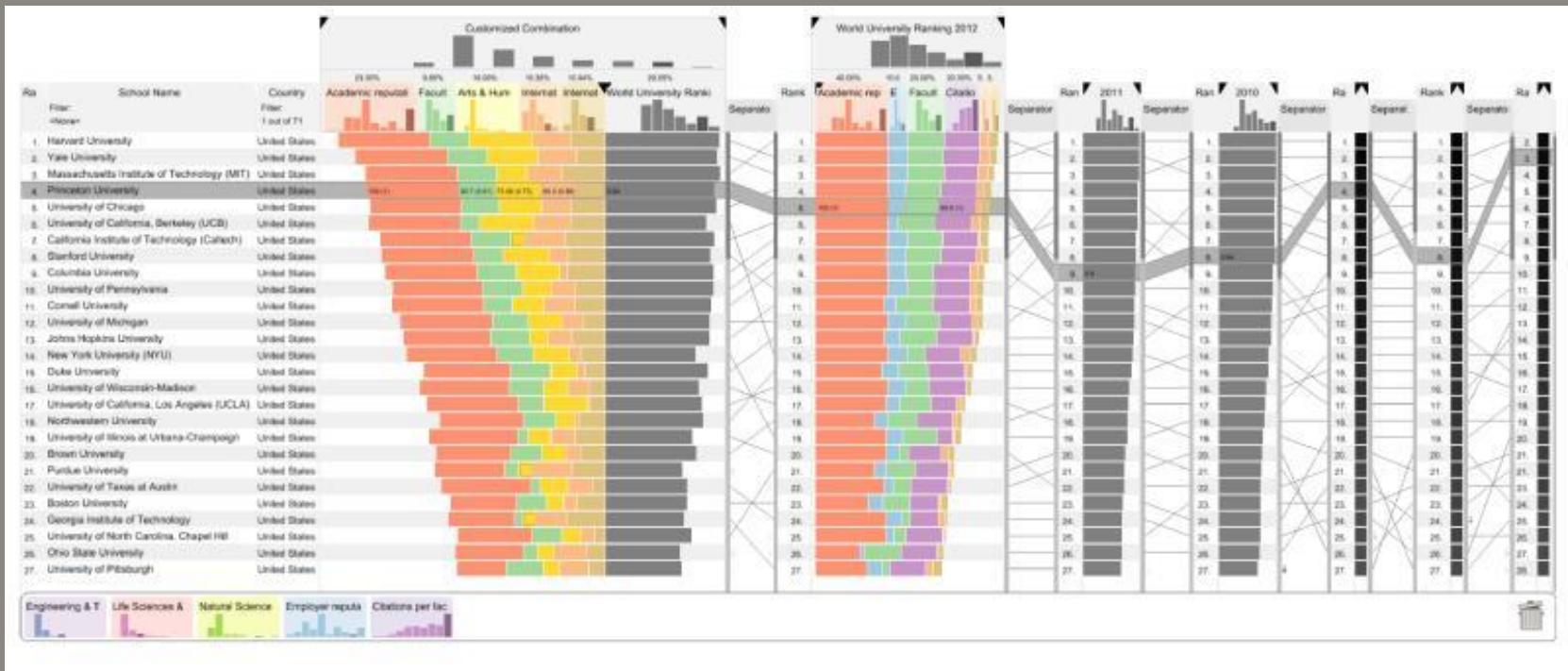
Source: <http://www.netindex.com/download/allcountries/>

Use the slope of lines to emphasize the ranks difference of two attributes of the same item



# Idiom: Slopegraph

## ○ Advanced application of slopegraph



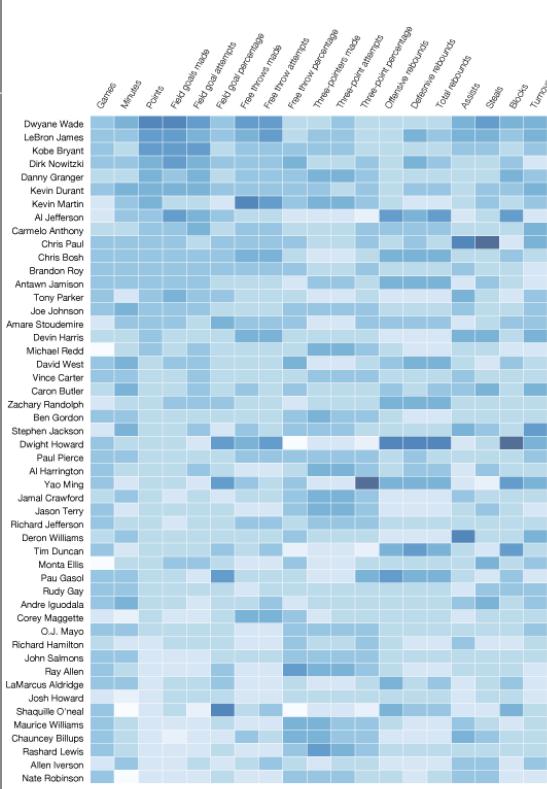


## Idiom: heatmap

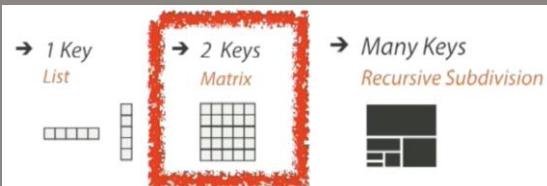
- Two keys, one value
  - Data
    - 2 categorical attributes
    - 1 quantitative attribute
  - marks: area
    - Separate and align in 2D matrix
    - Indexed by 2 categorical attributes
  - Channels
    - Color by quantitative value
  - Task
    - Find patterns, clusters, outlier
  - Scalability
    - 1M items, 100s of categorical levels, -10 quantitative attribute levels

NBA per game performance of top 50 scorers

2008-2009 season



Source: databaseBasketball





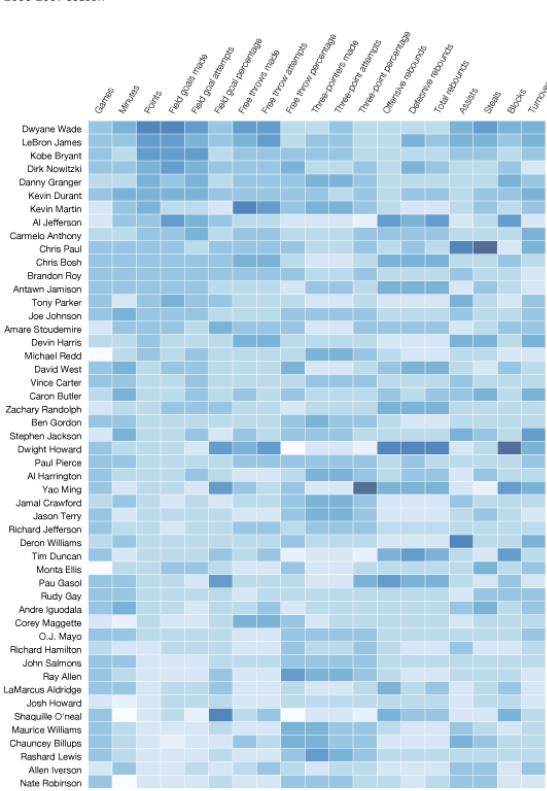
## Idiom: heatmap

Again! How to determine the order along x- and y axis

- Two keys, one value
  - Data
    - 2 categorical attributes
    - 1 quantitative attribute
  - marks: area
    - Separate and align in 2D matrix
    - Indexed by 2 categorical attributes
  - Channels
    - Color by quantitative value
  - Task
    - Find patterns, clusters, outlier
  - Scalability
    - 1M items, 100s of categorical levels, -10 quantitative attribute levels

NBA per game performance of top 50 scorers

2008-2009 season



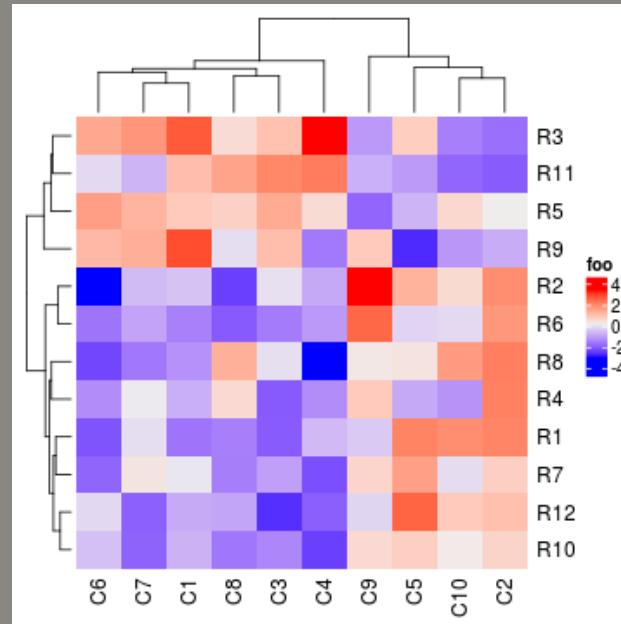
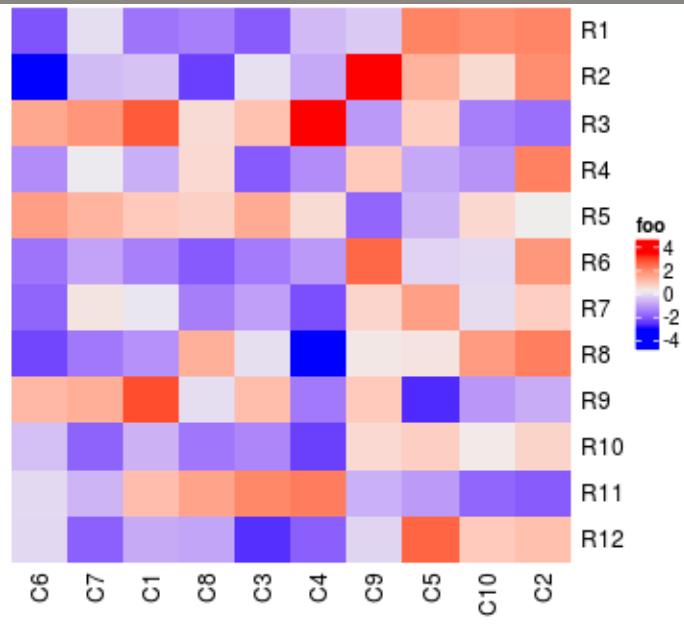
Source: databaseBasketball





## Idiom: heatmap

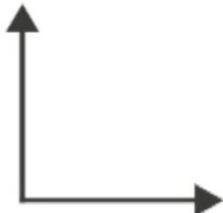
- Clustering and reordering the axis by similarity could be the key to show the patterns in the dataset



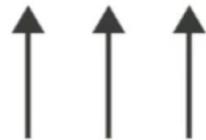
Reorder the row and column by hierarchical clustering

## → Axis Orientation

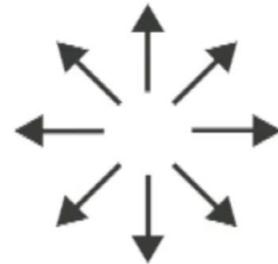
→ Rectilinear



→ Parallel



→ Radial





## Multi-variant Data Visualization

- Usually, we want show a visualization for users to discover the **correlation** among variables
- Multi-variant data

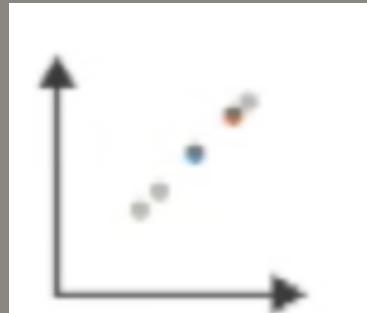
Table			
Math	Physics	Dance	Drama
V1	V2	V3	V4
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

4 variables: each data item is described by 4 values



# Multi-variant Data Visualization

- Usually, we want show a visualization for users to discover the correlation among variables
- Multi-variant data
- Correlation between variables
  - Math vs. Math



Perfect positive correlation

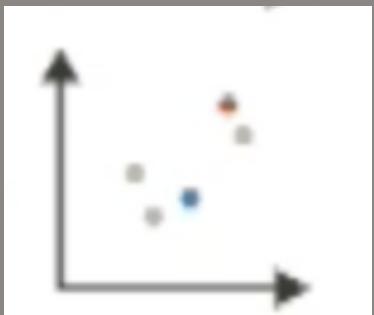
Table				
Math	Physics	Dance	Drama	
V1	V2	V3	V4	
85	95	70	65	
90	80	60	50	
65	50	90	90	
50	40	95	80	
40	60	80	90	

4 variables: each data item is described by 4 values



# Multi-variant Data Visualization

- Usually, we want show a visualization for users to discover the correlation among variables
- Multi-variant data
- Correlation between variables
  - Math vs. Physics
  - Dance vs. Drama



Positive correlation

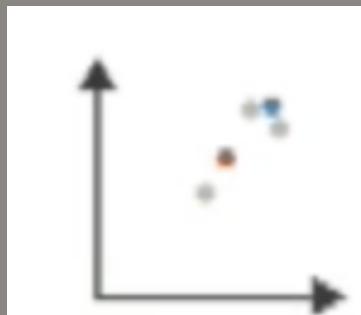


Table			
Math	Physics	Dance	Drama
V1	V2	V3	V4
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

4 variables: each data item is described by 4 values



# Multi-variant Data Visualization

- ➊ Usually, we want show a visualization for users to discover the correlation among variables
- ➋ Multi-variant data
- ➌ Correlation between variables
  - Physics vs. Dance



Negative correlation

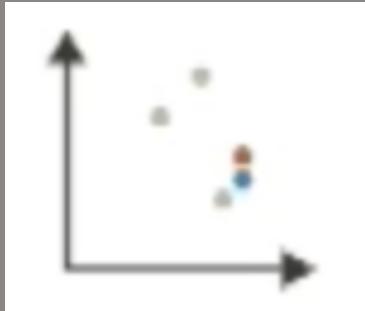
Table				
	Math	Physics	Dance	Drama
V1	85	95	70	65
V2	90	80	60	50
V3	65	50	90	90
V4	50	40	95	80
V5	40	60	80	90

4 variables: each data item is described by 4 values



# Multi-variant Data Visualization

- ➊ Usually, we want show a visualization for users to discover the correlation among variables
- ➋ Multi-variant data
- ➌ Correlation between variables
  - Physics vs. Dance



Uncorrelation?

Table				
Math	Physics	Dance	Drama	
V1	V2	V3	V4	
85	95	70	65	
90	80	60	50	
65	50	90	90	
50	40	95	80	
40	60	80	90	

4 variables: each data item is described by 4 values

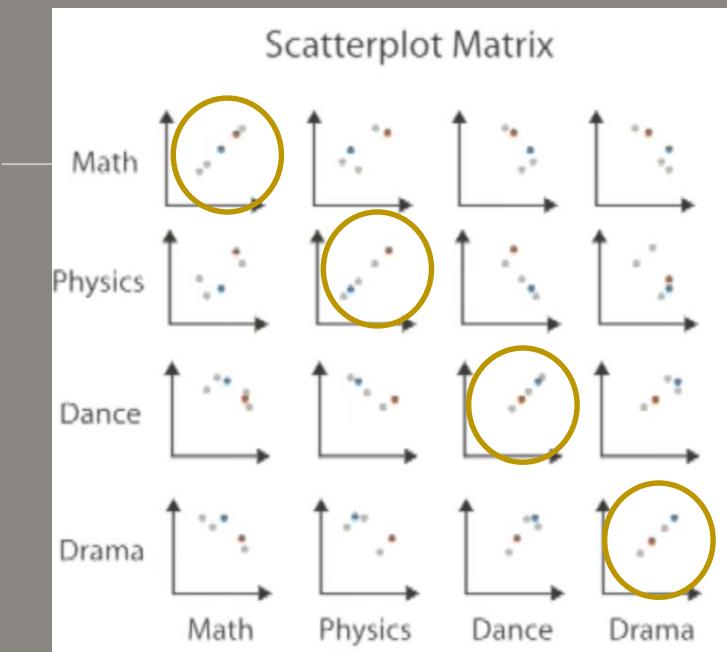


## Idioms: scatterplot matrix



### Scatterplot matrix (SPLOM)

- Rectilinear axes, point mark
- All possible pairs of axes
- Scalability
  - **One dozen attributes**
  - Dozens to hundreds of items



Diagonal plots must be perfect correlation.  
We can consider use the diagonal space  
to show other information

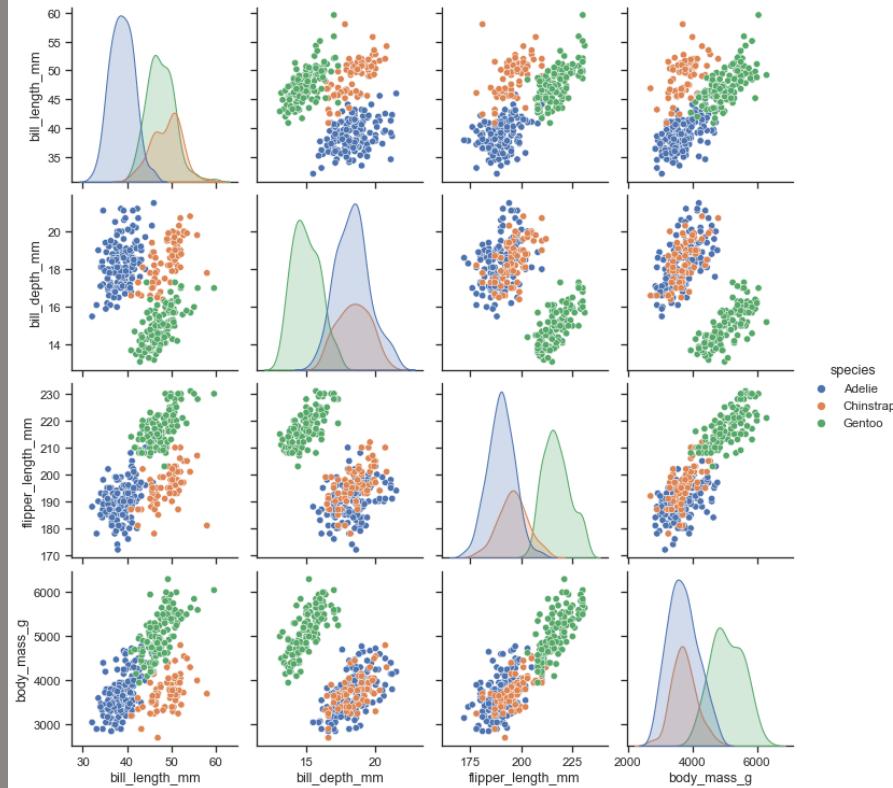


## Idioms: scatterplot matrix



### Scatterplot matrix (SPLOM)

- Rectilinear axes, point mark
- All possible pairs of axes
- Scalability
  - One dozen attributes
  - Dozens to hundreds of items



Diagonal plots must be perfect correlation.  
We can consider use the diagonal space  
to show other information

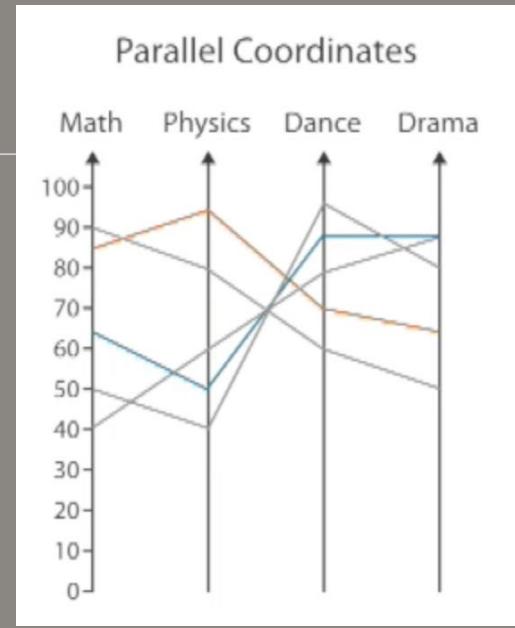


## Idioms: parallel coordinates



### Parallel coordinates

- Parallel axes, jagged line representing item
- Rectilinear axes, item as point
  - Axis ordering is major challenge
- Scalability
  - **Dozens of attributes**
  - Hundreds of items



Table

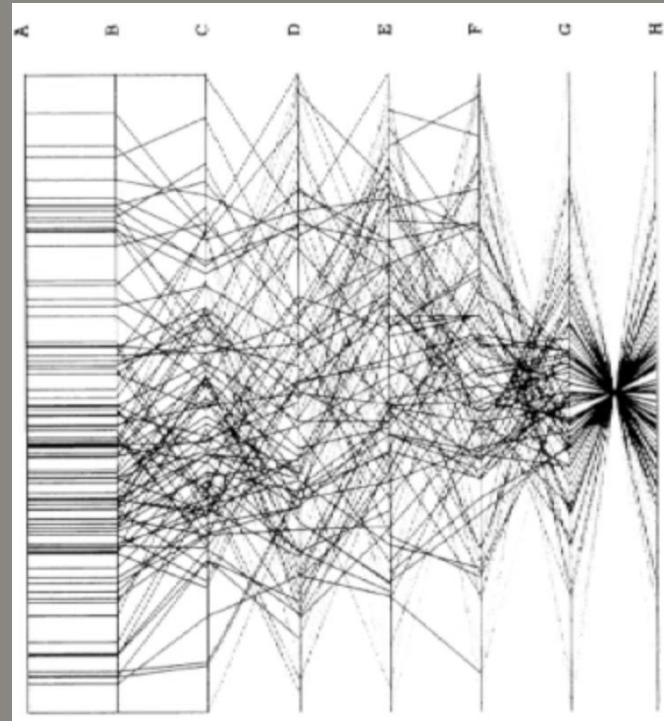
	Math	Physics	Dance	Drama
Item 1 (Orange)	85	95	70	65
Item 2 (Blue)	60	80	60	50
Item 3 (Green)	40	50	95	80
Item 4 (Red)	90	70	70	65



# Idioms: parallel coordinates

## Parallel coordinates

- Positive correlation
  - Parallel line segments
  - E.g. A vs. B
- Negative correlation
  - All segments cross at halfway point
  - E.g. G vs. H
- Uncorrelated
  - Scattered crossings
  - E.g. D vs E

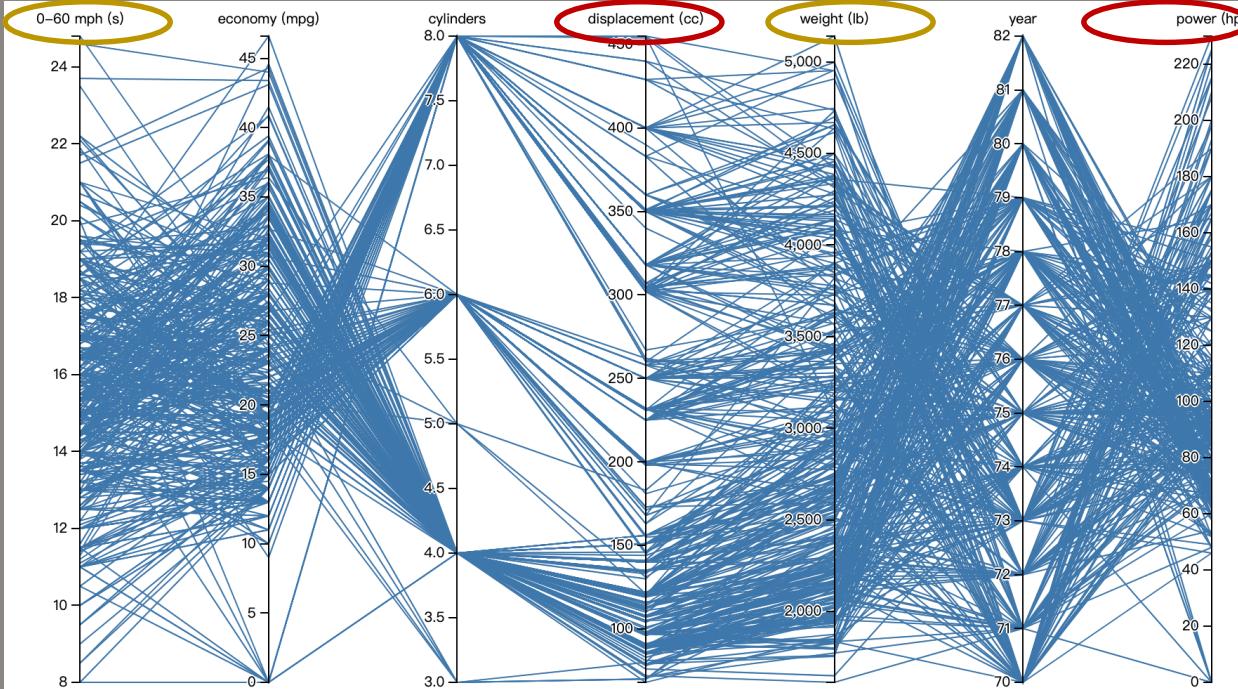




# Idioms: parallel coordinates



- Parallel coordinates
  - Order of axes matter
  - Correlation between
    - Mph vs. weight
    - Power vs. disp.

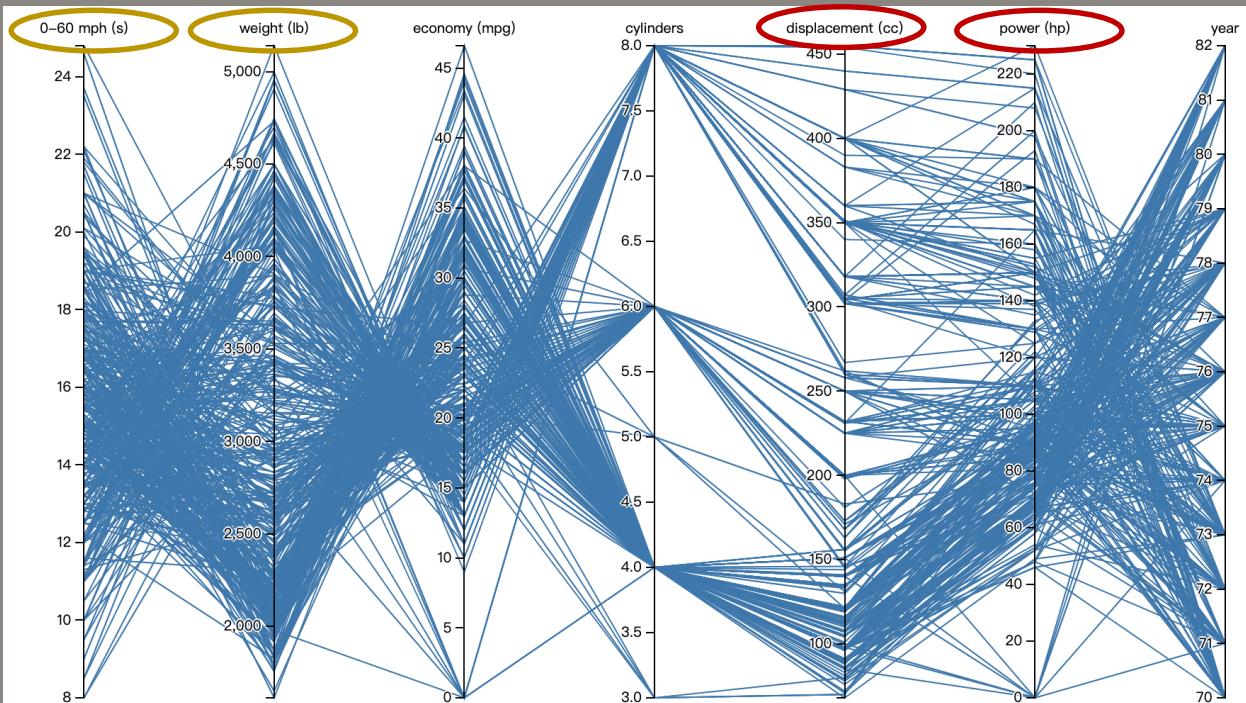




# Idioms: parallel coordinates



- Parallel coordinates
  - Order of axes matter
  - Correlation between
    - Mph vs. weight
    - Power vs. disp.
  - Reorder them
  - Usually, the interaction is required in PCP



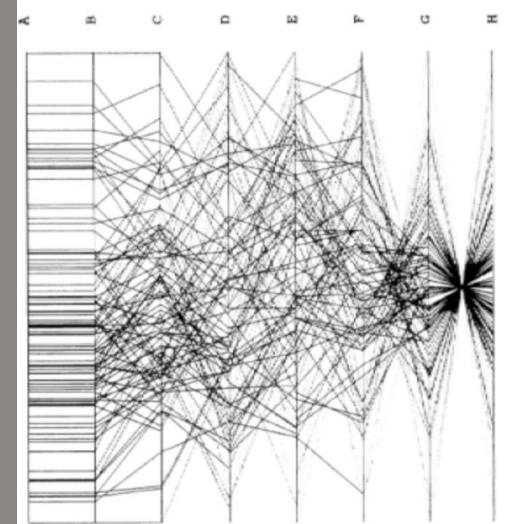
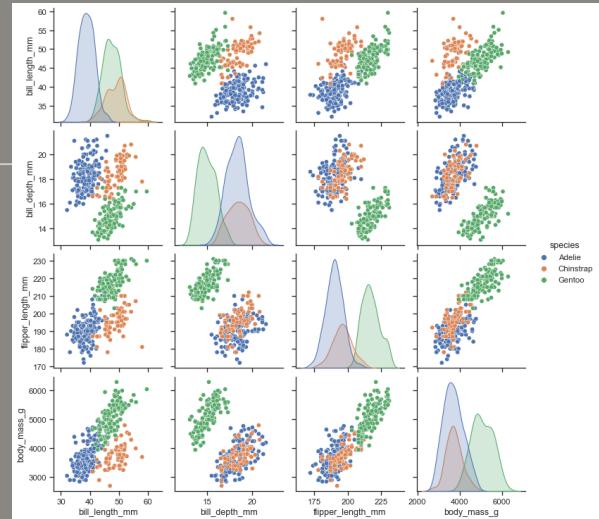


S05-04



# Pros and Cons

- Scatterplot matrix
  - Pros: easy to understand
  - Cons: scalability
    - The number of the sub-scatter plots grows exponentially
  
- Parallel coordinates
  - Pros: scalability
    - It can show much more variables
  - Cons:
    - Most people know how to interpret parallel coordinate plot.
    - People need to be trained to interpret it.
    - Not good to use it on news paper (interaction)



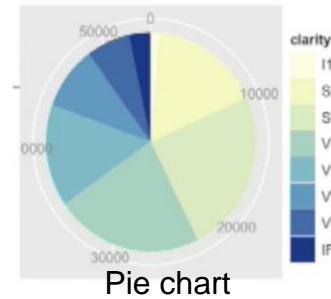


## Idiom: Radial Layout

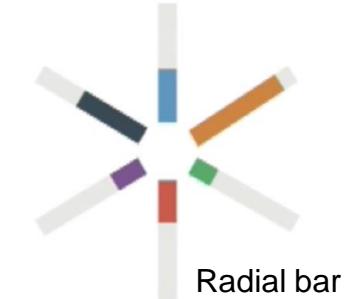


### Pie chart (classic/controversial)

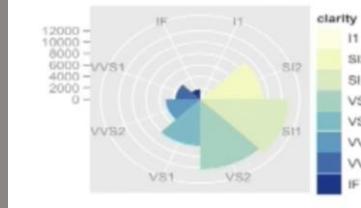
- Area mark with angle channel
- Data: 1 categorical key attribute, 1 quantitative value attribute
- Task: Part-to-whole judgements



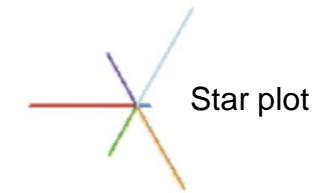
Pie chart



Radial bar



Radius pie chart



Star plot

D3: [https://www.d3-graph-gallery.com/circular\\_barplot](https://www.d3-graph-gallery.com/circular_barplot)

### → Axis Orientation

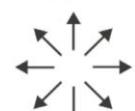
→ Rectilinear

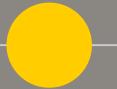


→ Parallel



→ Radial





## Idiom: Radial Layout

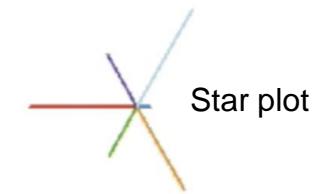
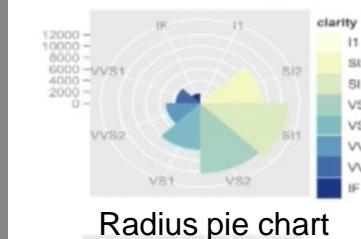
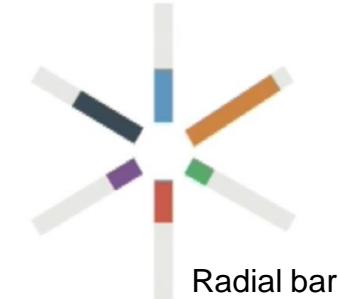
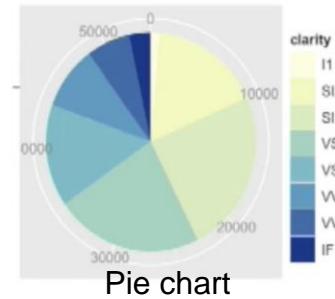


### Pie chart (classic/controversial)

- Area mark with angle channel
- Data: 1 categorical key attribute, 1 quantitative value attribute
- Task: Part-to-whole judgements



### Advantage: visually appealing and **ok for Part-to-whole judgements**



D3: [https://www.d3-graph-gallery.com/circular\\_barplot](https://www.d3-graph-gallery.com/circular_barplot)

### → Axis Orientation

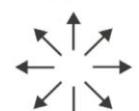
→ Rectilinear



→ Parallel



→ Radial





S05-05



# Idiom: Radial Layout



## Pie chart (classic/controversial)

- Area mark with angle channel
- Data: 1 categorical key attribute, 1 quantitative value attribute
- Task: Part-to-whole judgements

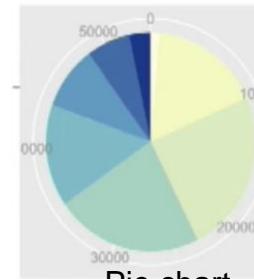


Advantage: visually appealing and ok for Part-to-whole judgements



Many alternative visualization is better than radial layout in terms of data understanding purpose

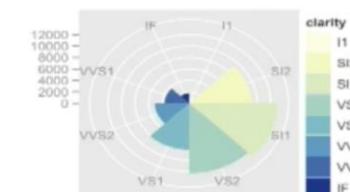
- Accuracy: angle/area less accurate than line length
  - Arc length also less accurate than line length



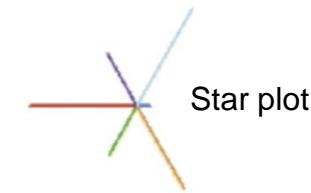
Pie chart



Radial bar



Radius pie chart



Star plot

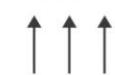
D3: [https://www.d3-graph-gallery.com/circular\\_barplot](https://www.d3-graph-gallery.com/circular_barplot)

## → Axis Orientation

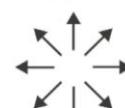
→ Rectilinear



→ Parallel



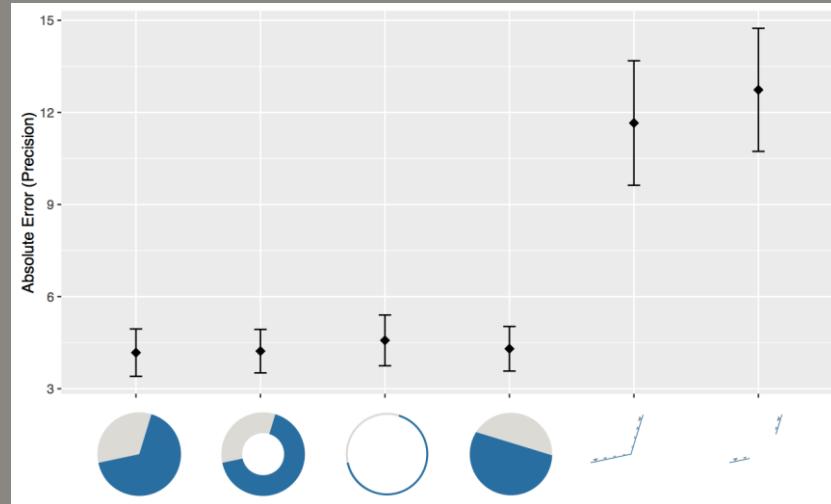
→ Radial





# Pie Chart Perception

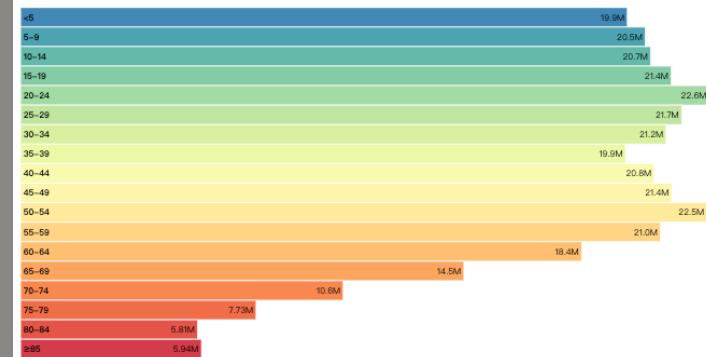
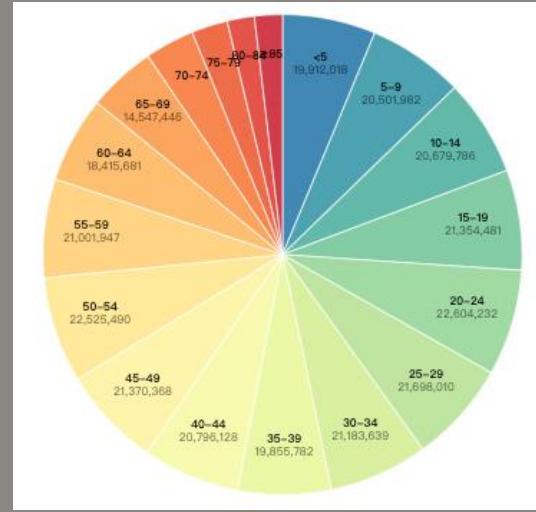
- ➊ Some empirical evidence that people respond to arc length
  - Not angle
  - Maybe also area
- ➋ Donut chart no worse than pie charts



the U.S. population from 2015,  
divided into 5-year age buckets

## Pie Chart vs Bar Chart

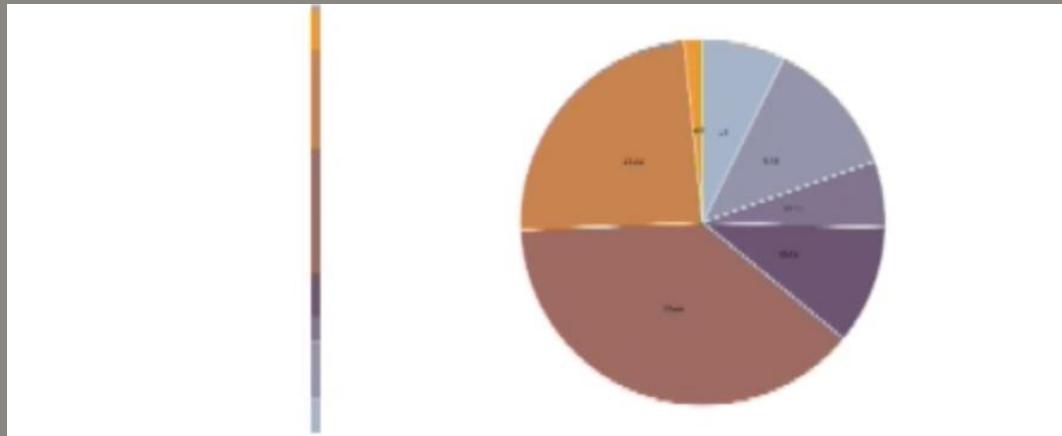
- ➊ Impression of this pie chart: “each age group is roughly the same size, decreasing for the older groups.”
- ➋ Check the same data shown by bar chart
  - humans are better at comparing lengths than angles





## Pie Chart vs One Stacked Bar Chart

- Single stacked bar equivalent to full pie
  - High information density: requires narrow rectangle
- Pie chart
  - Information density: requires large circle





S05-06

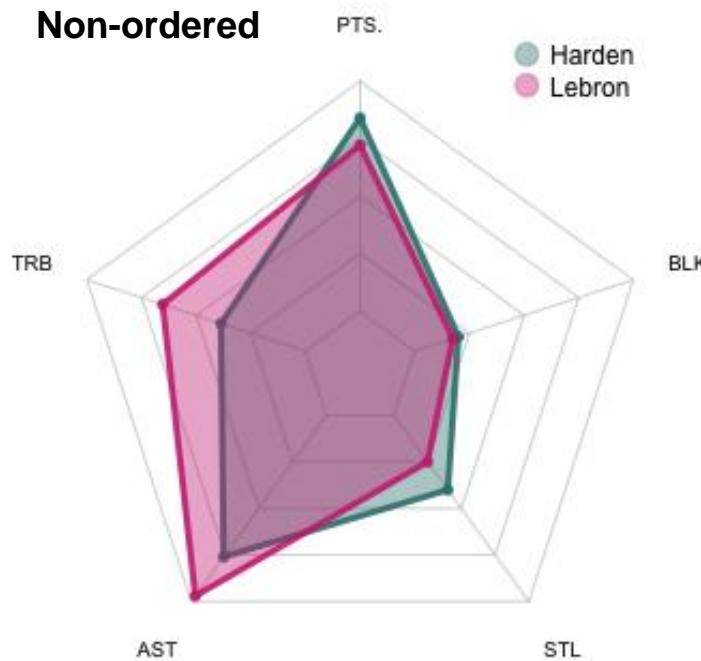


# Radar Plot



Not so proper when categories are not ordered/cyclic

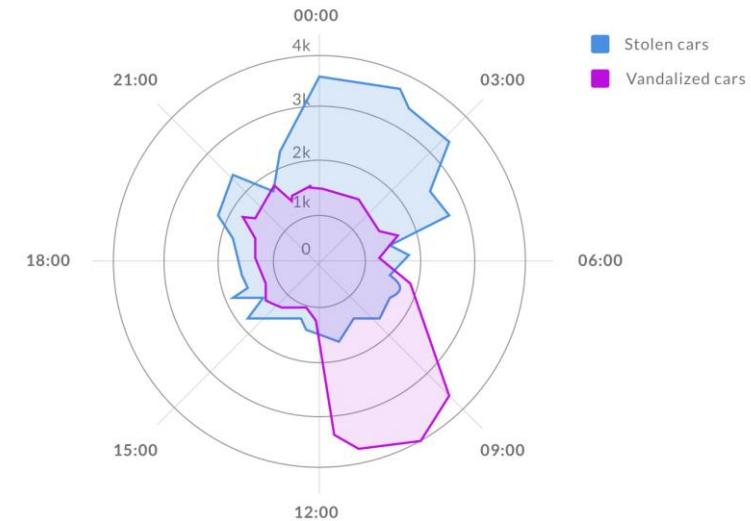
**Non-ordered**



Auto crime by time of day

< 08 Jan 2019 >

**Ordered and cyclic**



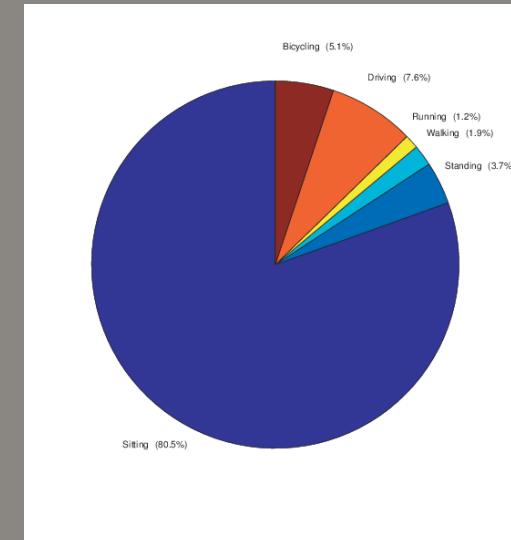
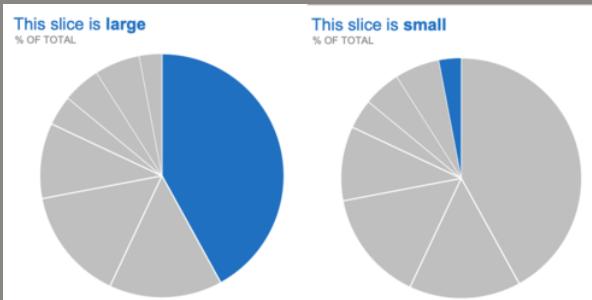


# When to Use Radial Layout



Is pie chart or radial layout always bad choice?

- Part-to-whole/one segment ratio/precision is less important
  - If you want your audience to have a **general sense of the part-to-whole relationship in your data** and comparing the precise sizes of the slices is less important.
  - To convey that one segment of the total is relatively **small or large**.



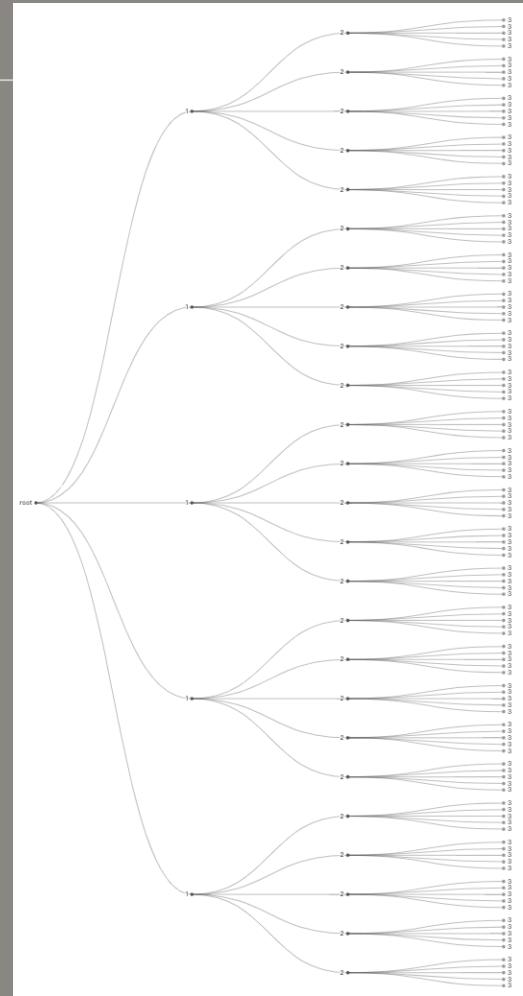
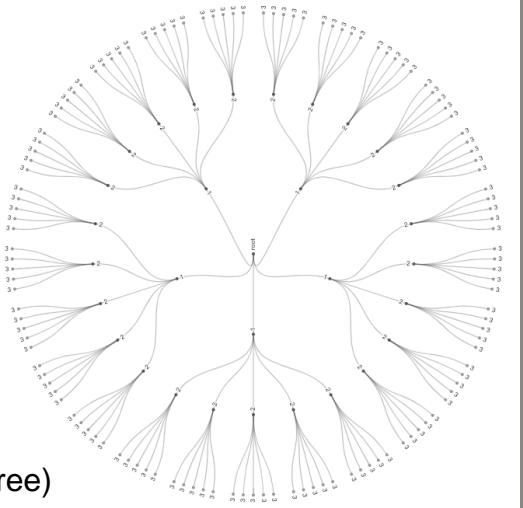
You can use it if no need to compare an arc with another arc



# When to Use Radial Layout



- Is pie chart or radial layout always bad choice?
  - Part-to-whole/one segment ratio/precision is less important
  - Compact
    - Tree: give leaf nodes the same size for interaction
    - We have display limit



D3: <https://bl.ocks.org/d3noob/8375092> (tree)

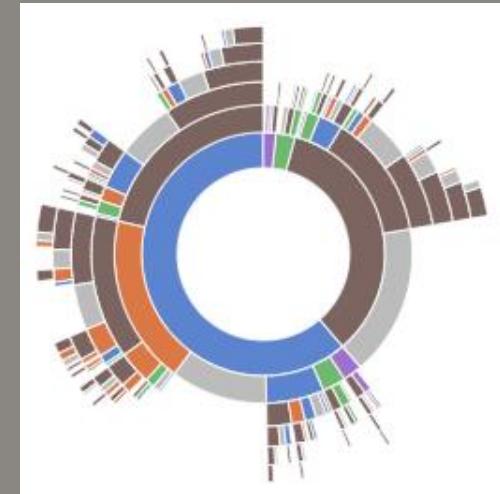
D3: <https://observablehq.com/@d3/radial-tidy-tree> (radial tree)



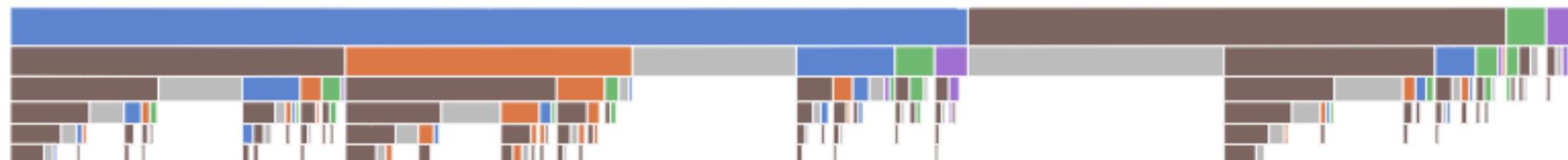
# When to Use Radial Layout

- 🟡 Is pie chart or radial layout always bad choice?
  - Part-to-whole/one segment ratio/precision is less important
  - Compact
    - Tree: give leaf nodes the same size for interaction
    - We have display limit

sunburst



D3: <https://bl.ocks.org/kerryrodden/7090426>





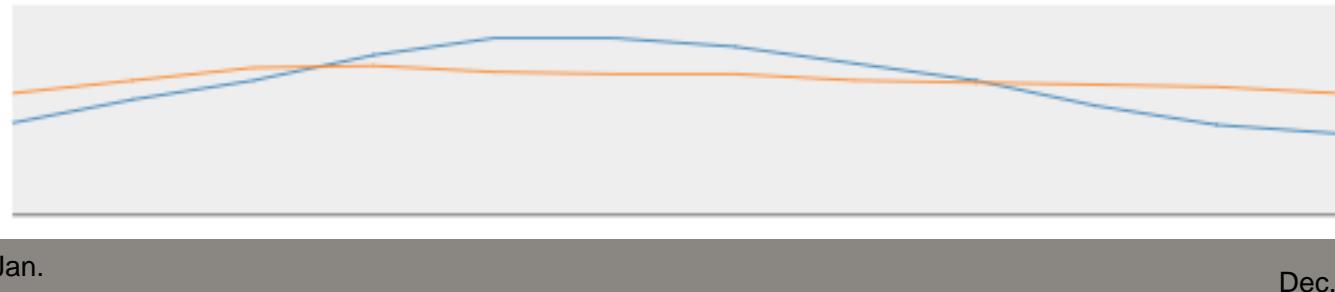
# When to Use Radial Layout



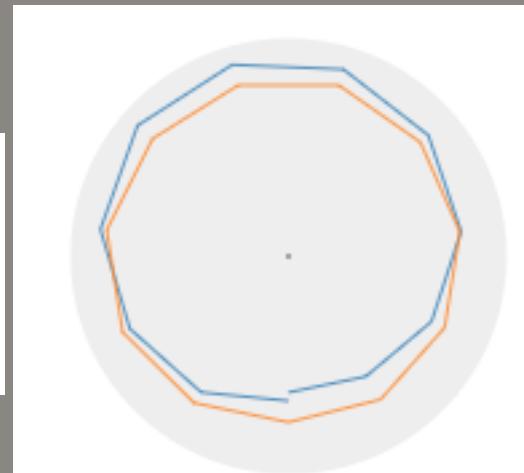
- Is pie chart or radial layout always bad choice?
  - Part-to-whole/one segment ratio/precision is less important
  - Compact
  - Cyclic data
    - e.g. continuity between the winter months

temperature

Line chart is still easy to interpret



New York (blue) vs LA (orange)  
average temperature each month





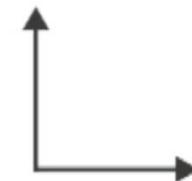
## Limitations of Orientation

- ◉ Rectilinear: scalability (axes)
  - 2 axes is best
  - 3 problematic
  - 4+ impossible
- ◉ Parallel
  - unfamiliarity, training time
- ◉ Radial: perceptual limits
  - Angles lower precision than lengths
  - Asymmetry between angle and length



### Axis Orientation

#### → Rectilinear



#### → Parallel



#### → Radial

