

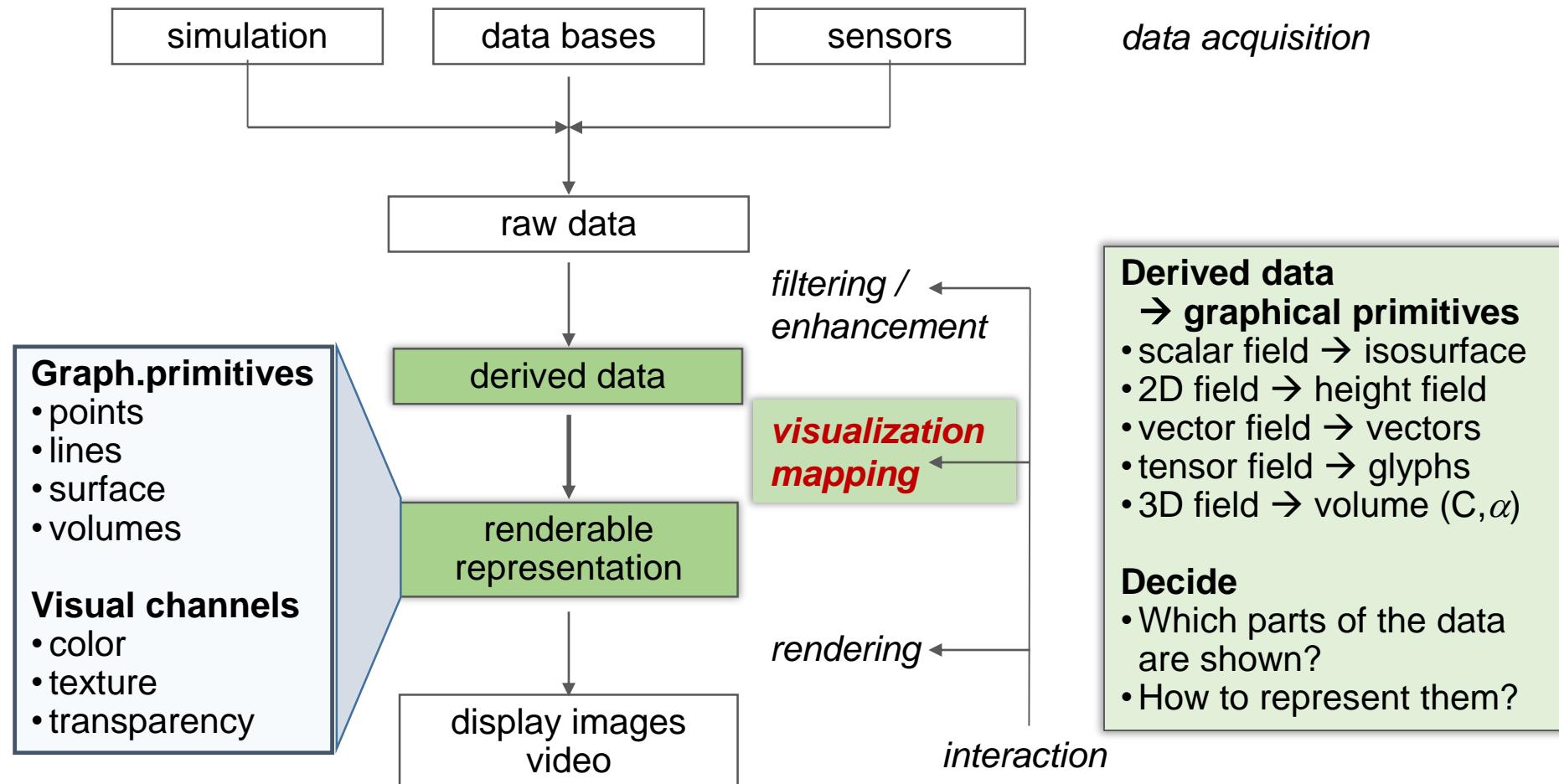
# Visualization

## – Elementary Methods

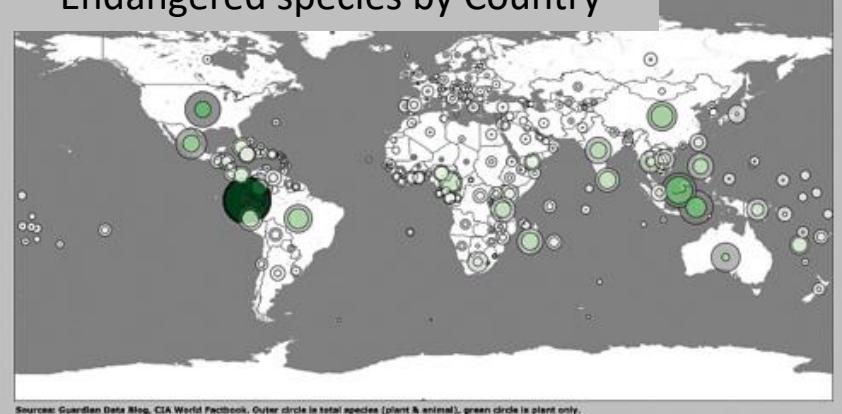
---

J.-Prof. Dr. habil. Kai Lawonn

# Mapping techniques



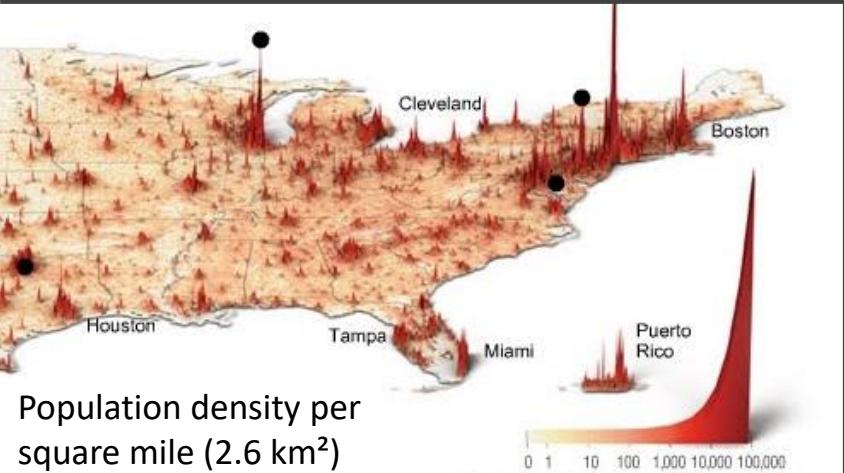
Endangered species by Country



Sources: Guardian Data Blog, CIA World Factbook. Outer circle is total species (plant & animal), green circle is plant only.

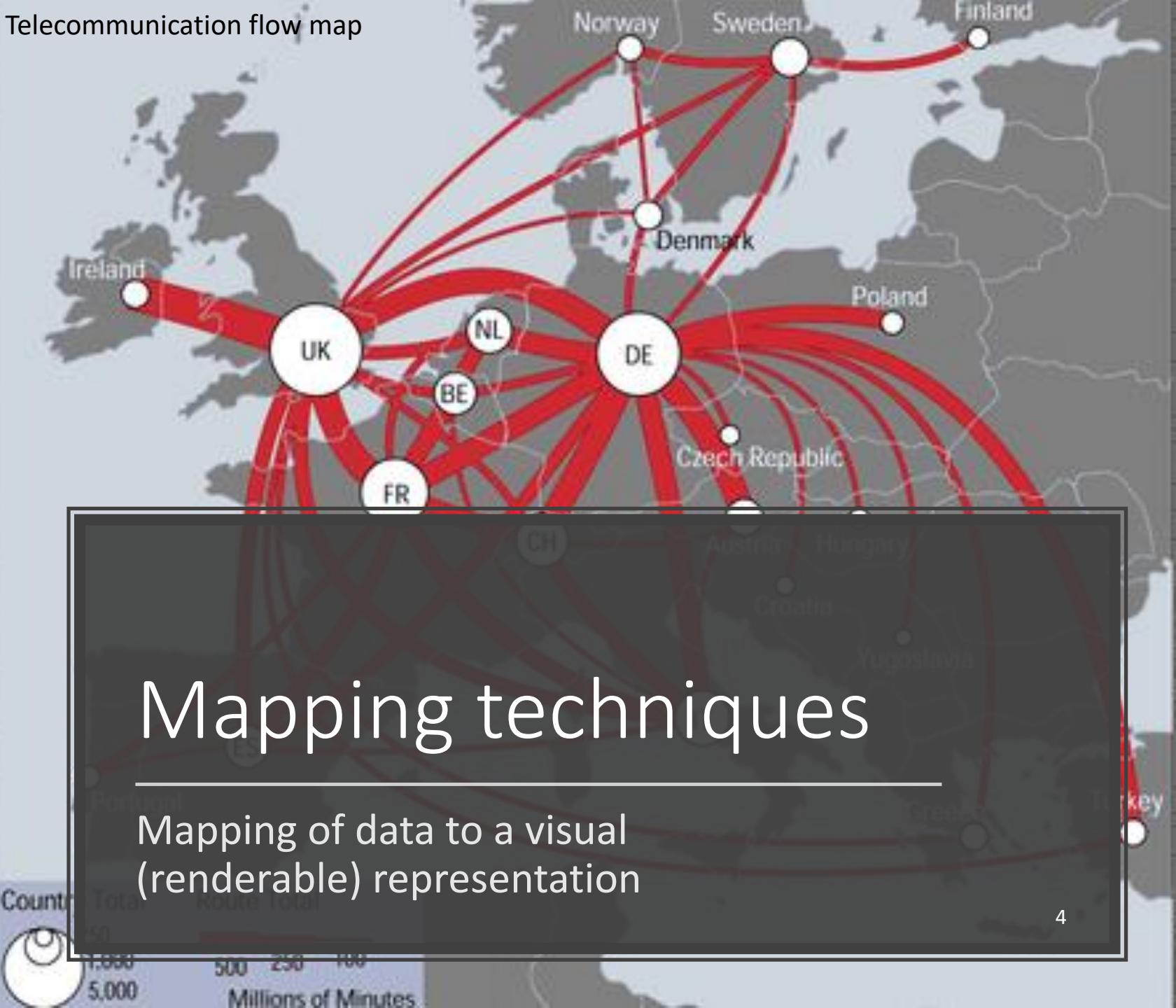


Military expenses worldwide



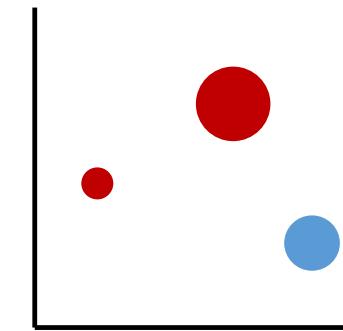
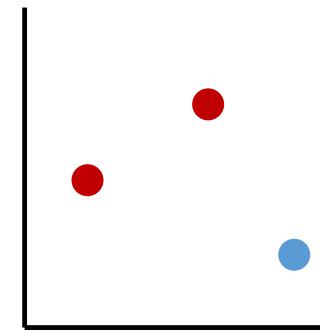
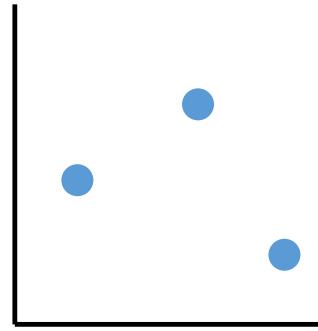
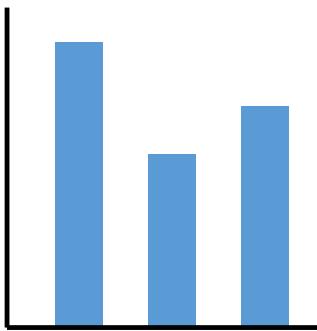
Population density per square mile ( $2.6 \text{ km}^2$ )

Telecommunication flow map



# Mapping techniques

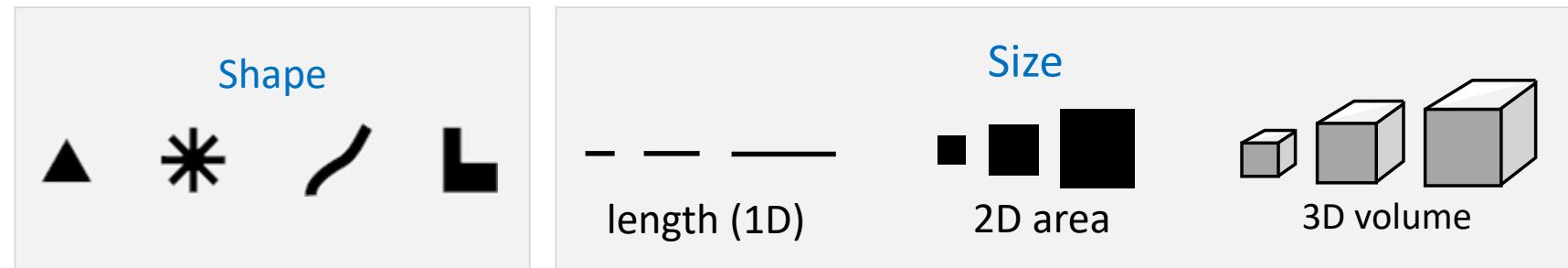
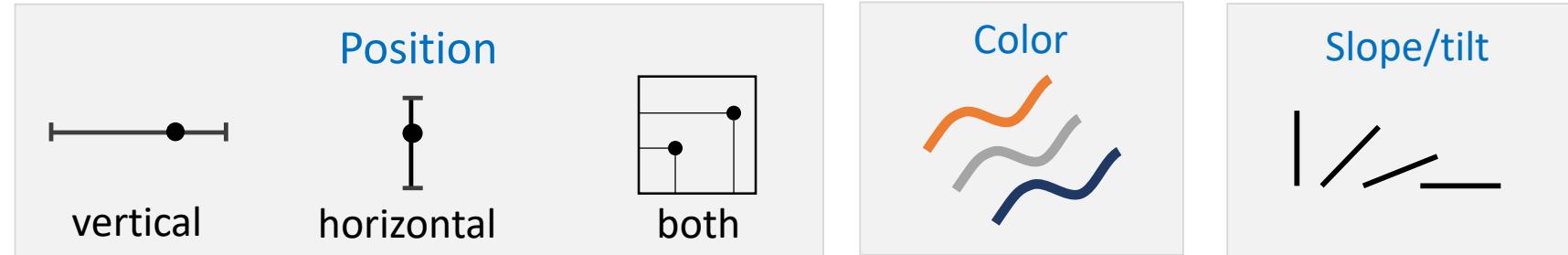
Analyzing visual mappings



# Mapping techniques

Mapping of data to a visual (**renderable**) representation consists of:

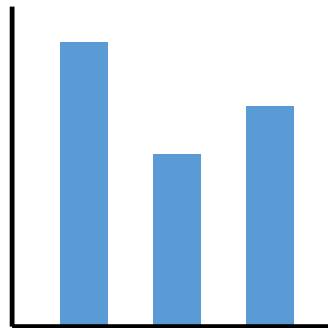
- **Graphical primitives** that can be displayed
- **Visual channels** that control appearance of graph. primitives and are used to encode data values



# Mapping techniques

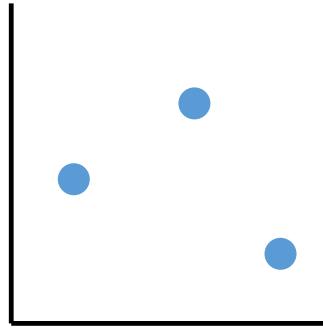
## Analyzing visual mappings

- combination of graphical primitives and visual channels

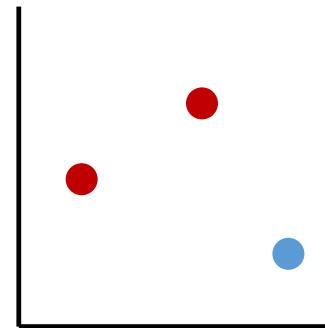


**Visual channel:**

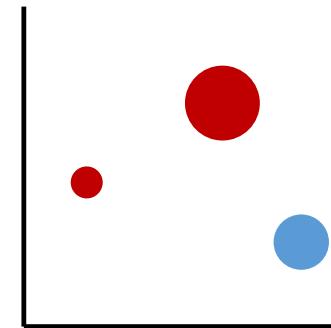
1:  
Length/position



2:  
Vertical position  
horizontal position



3:  
Vertical position  
horizontal position  
color hue



4:  
Vertical position  
horizontal position  
color hue  
size (area)

**Graphical primitive:**

Line/bar

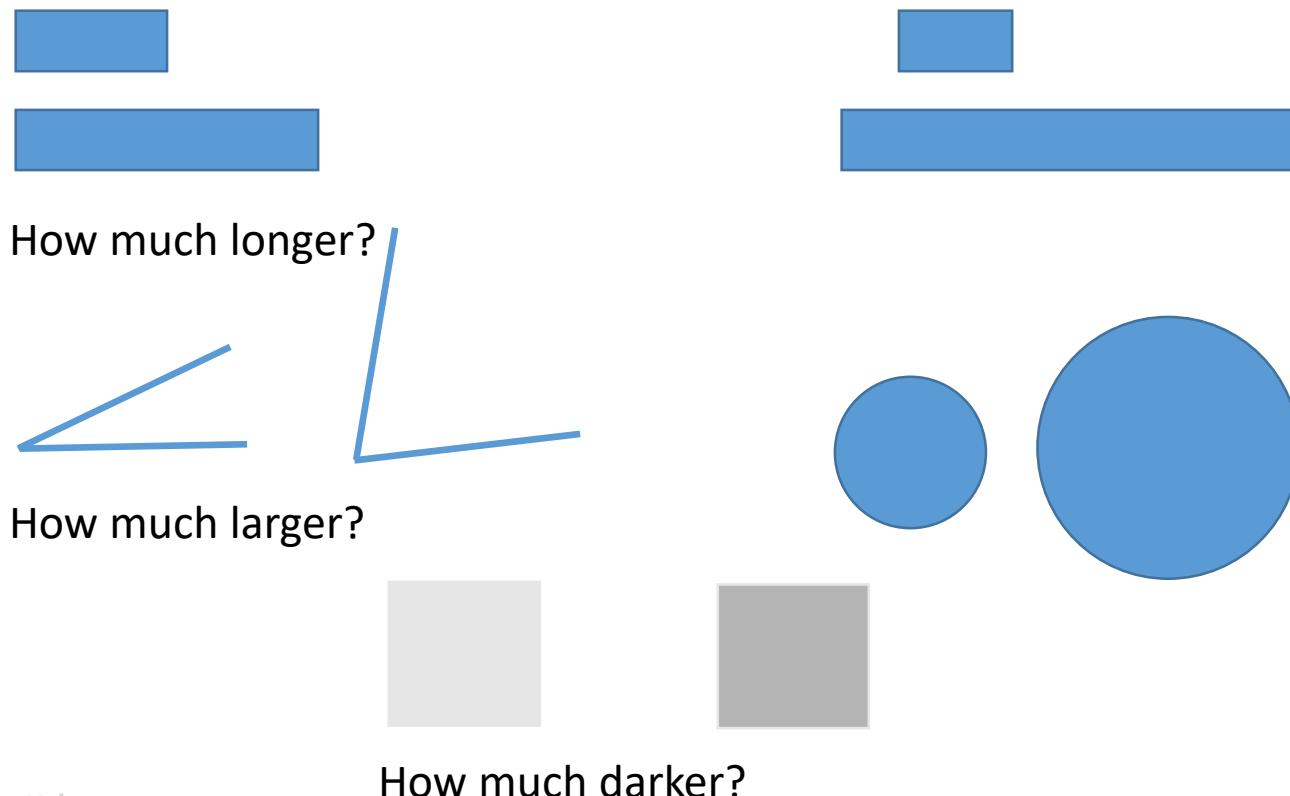
Point

Point

Point

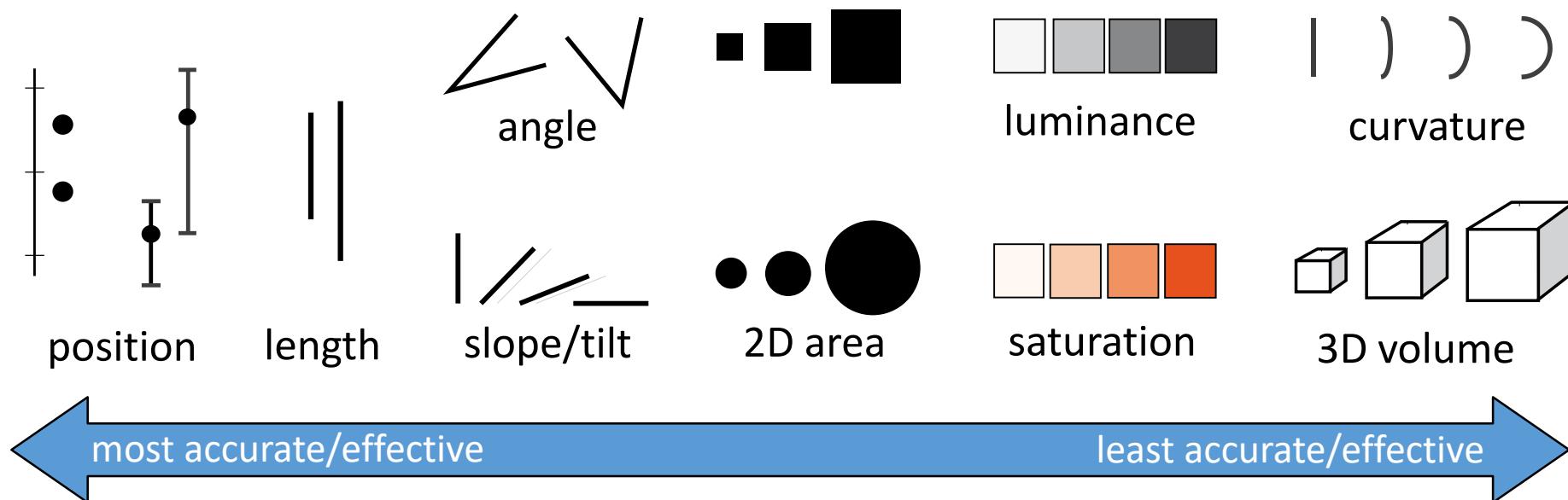
# Mapping techniques

- Accuracy/effectiveness:  
Some visual channels can be compared more accurately



# Mapping techniques

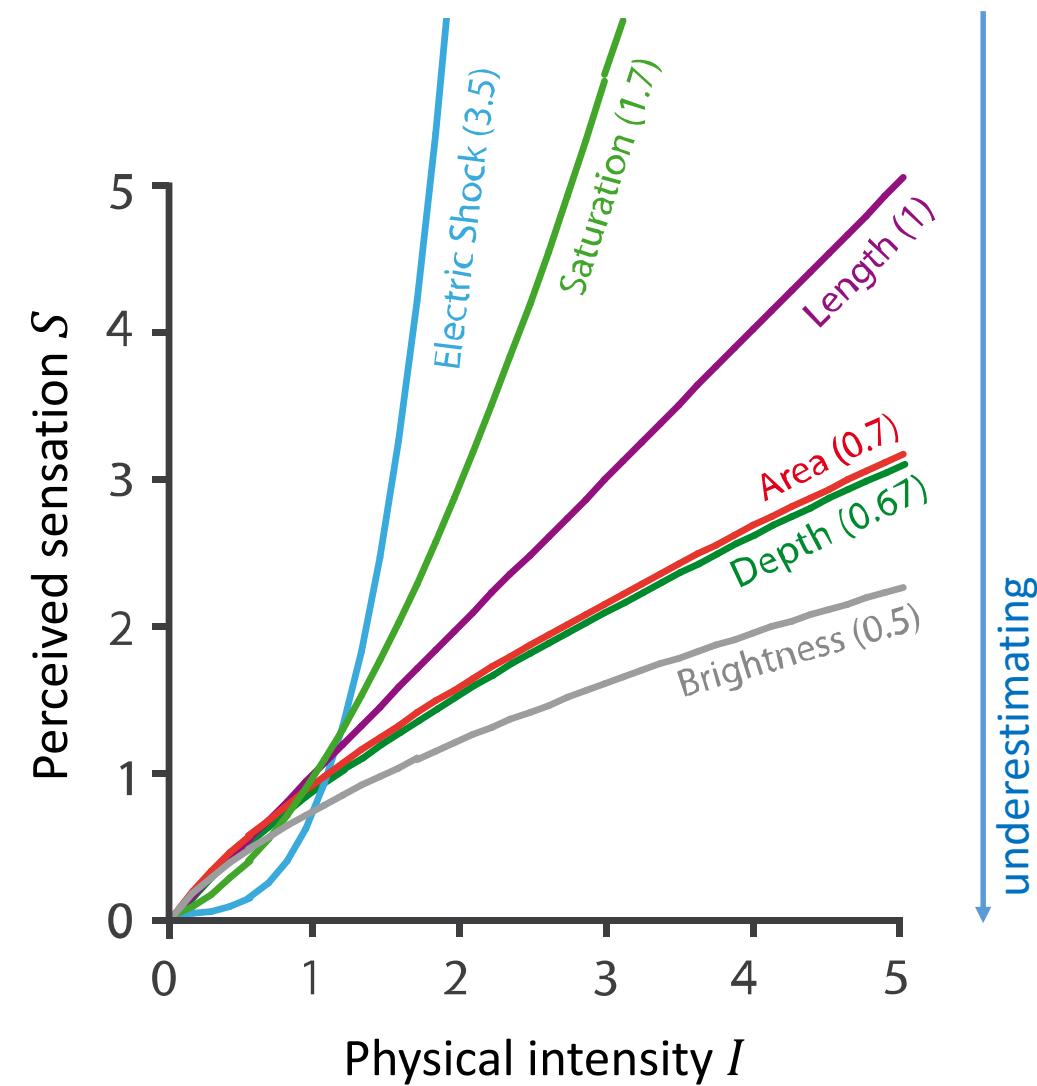
- Accuracy/effectiveness
  - Ranking derived from empirical studies



[Cleveland & McGill 84,  
Heer and Bostock 10,  
Munzner 14]

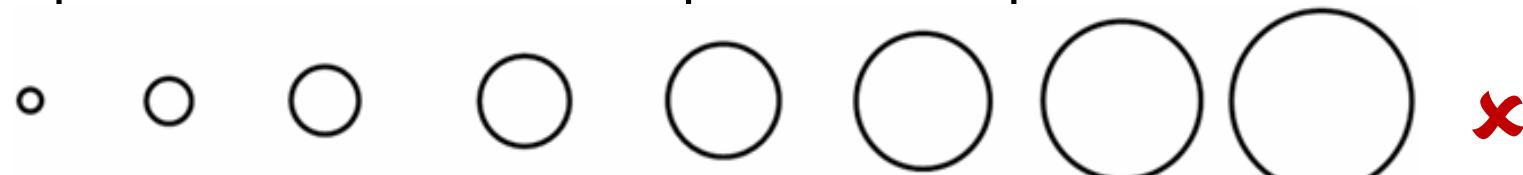
# Mapping techniques

- Accuracy/effectiveness
  - Steven's psychophysical power law:  $S = I^\gamma$



# Mapping techniques

- Area judgement
  - Equal steps in data should look equal in the representation



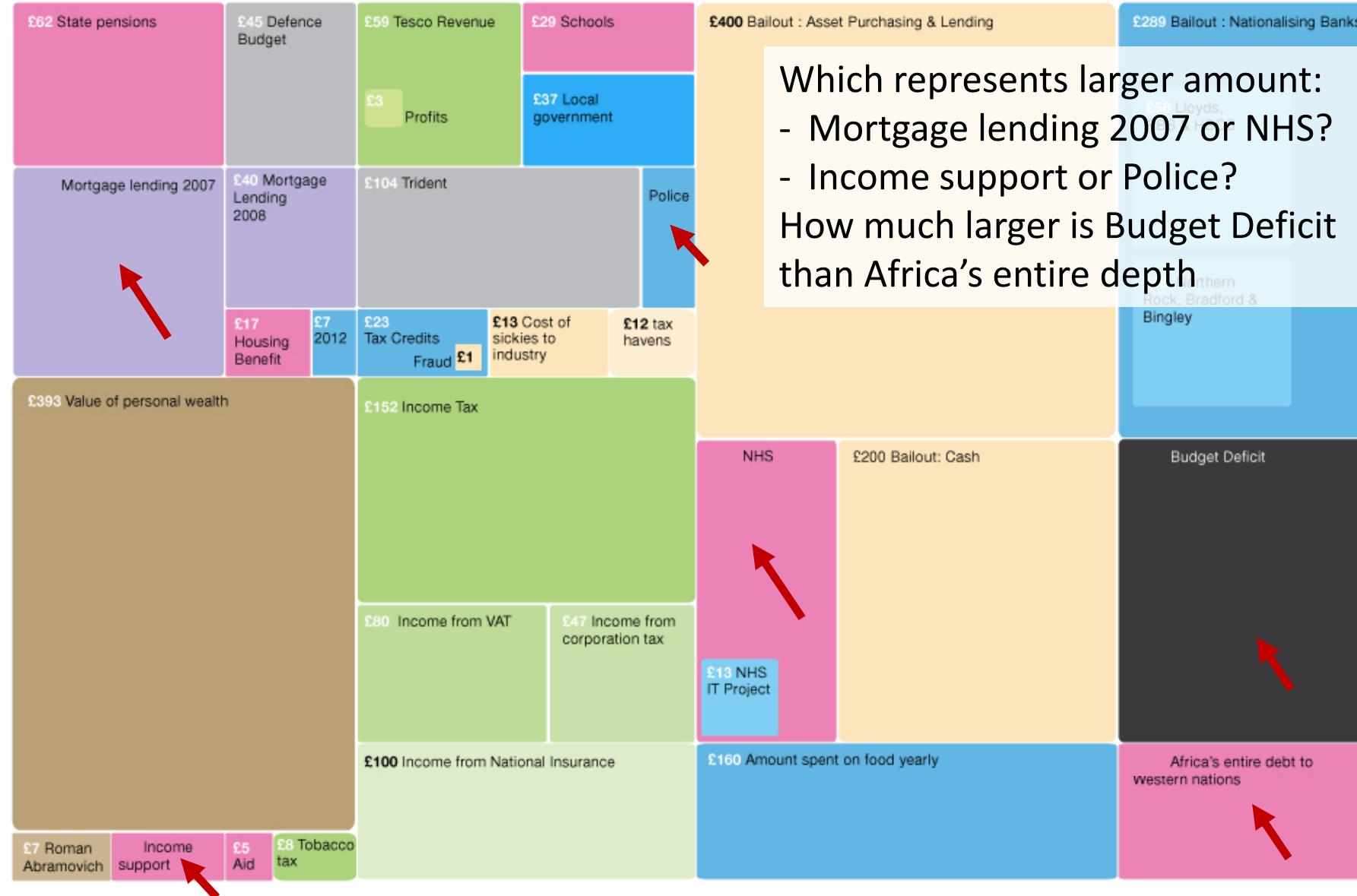
Attribute → radius



Attribute → area:  $I = \alpha r^2$



Steven's area judgment scale:  $I = \alpha(r^2)^{0.7} = \alpha r^{1.4}$



The Billion Pound-O-Gram

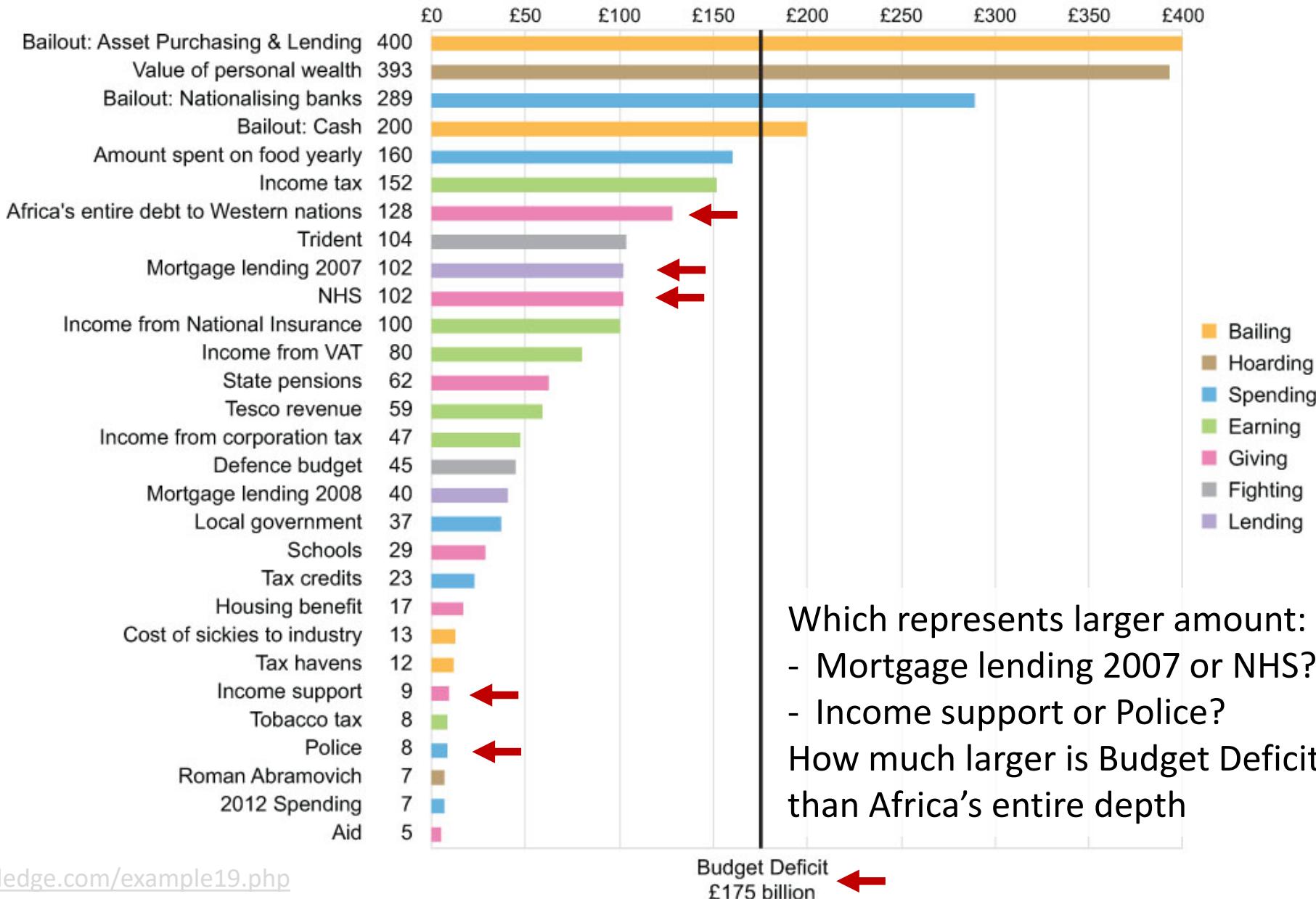
David McCandless / [InformationIsBeautiful.net](http://InformationIsBeautiful.net)

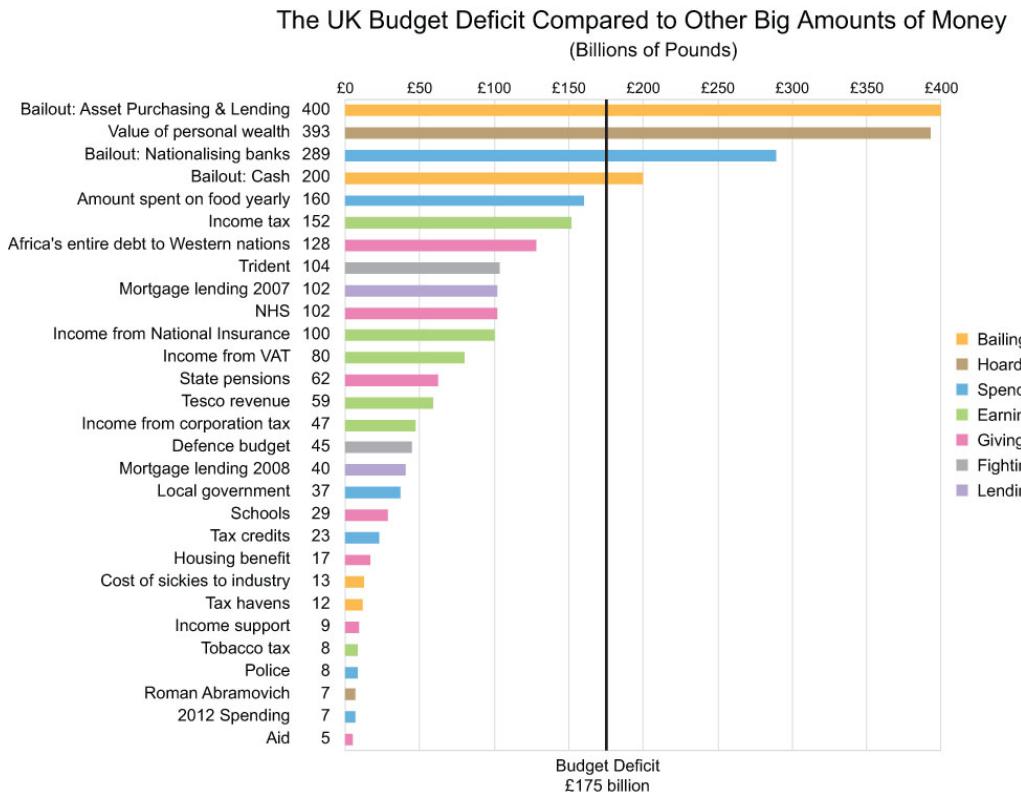
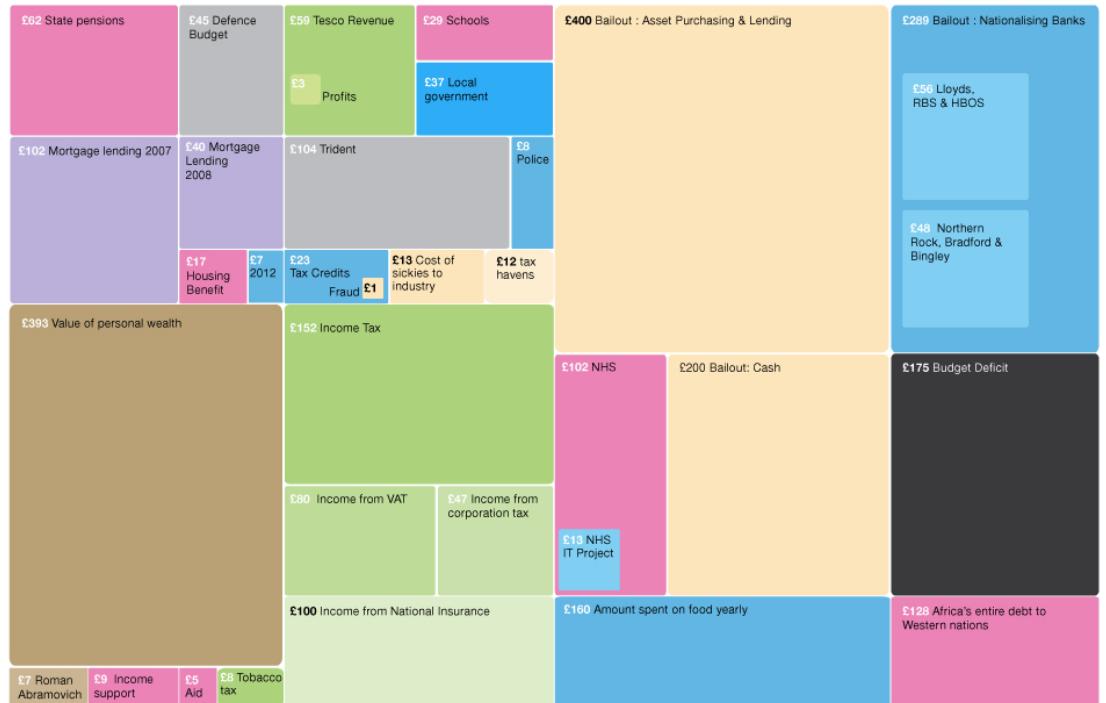
■ Giving ■ Spending ■ Fighting ■ Hoarding ■ Lending ■ Bailing ■ Earning

Source: UK Treasury, Guardian

# Redesign by Stephen Few

The UK Budget Deficit Compared to Other Big Amounts of Money  
(Billions of Pounds)





# Compare

- What do you prefer?

# Mapping techniques

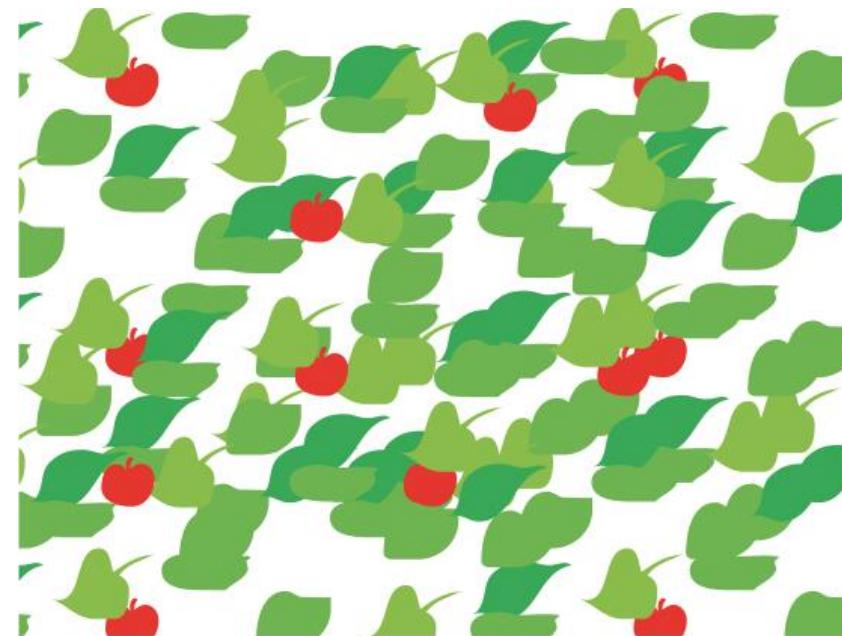
- Effectiveness principle
  - Encode most important attributes with most effective/accurate channels
- Many visualization mappings don't work
  - Accuracy
  - Pop-out (emphasize important information)
  - Discriminability (how many usable steps?)
  - Separability (judge each channel independently)
  - Relative vs. absolute judgement

# Mapping techniques

- Some visual channels “pop out”



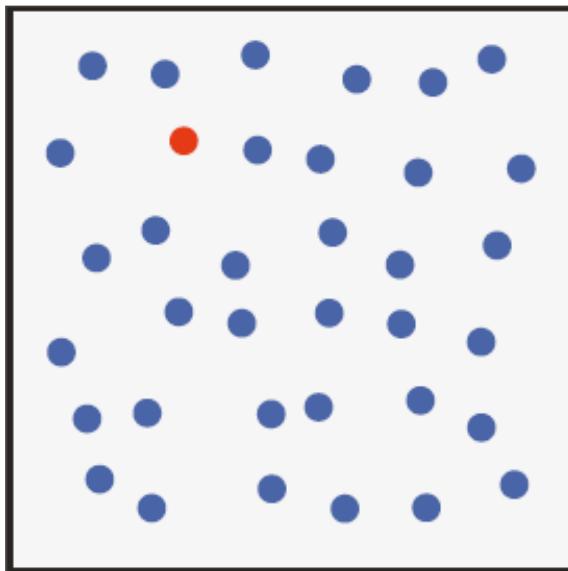
Where are the cherries?



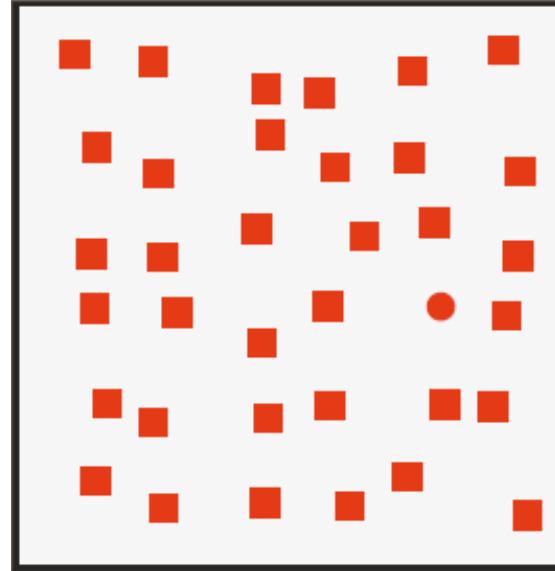
Hunters & gatherers

# Mapping techniques

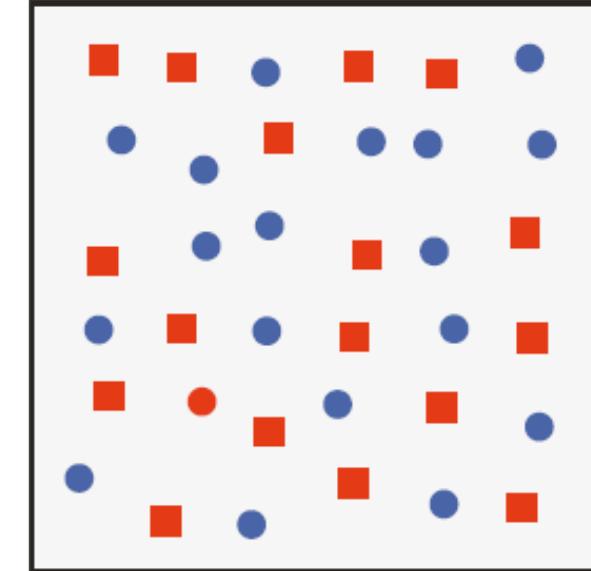
- Pop-out
  - Parallel processing on many individual channels
  - Speed independent of distractor count



Color



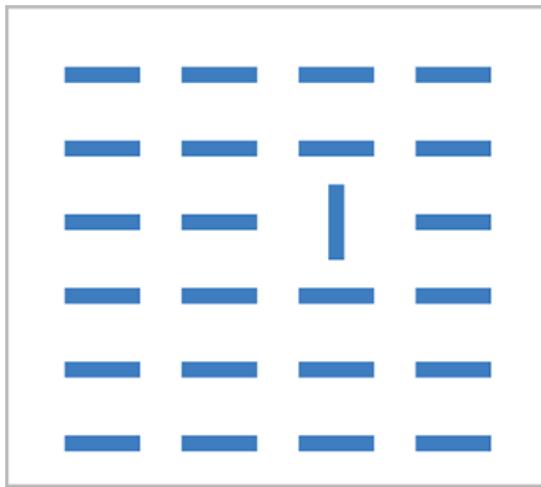
Shape



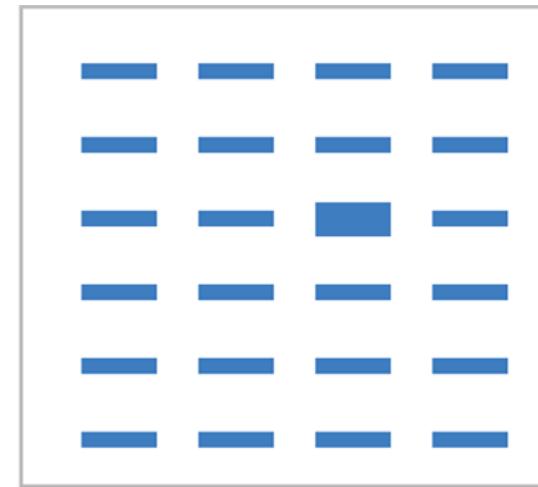
Combination of channels  
usually requires serial search

# Mapping techniques

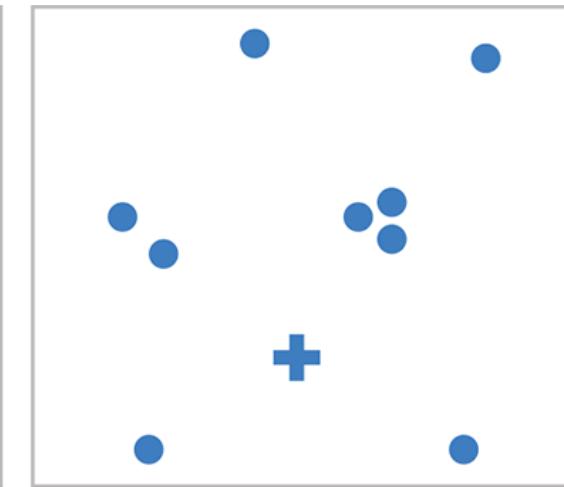
- Pop-out



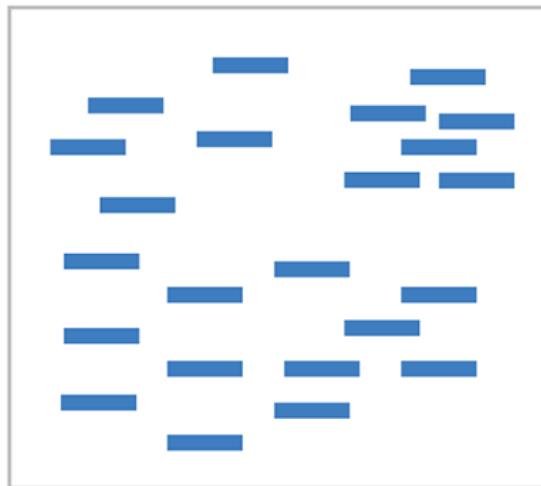
Tilt



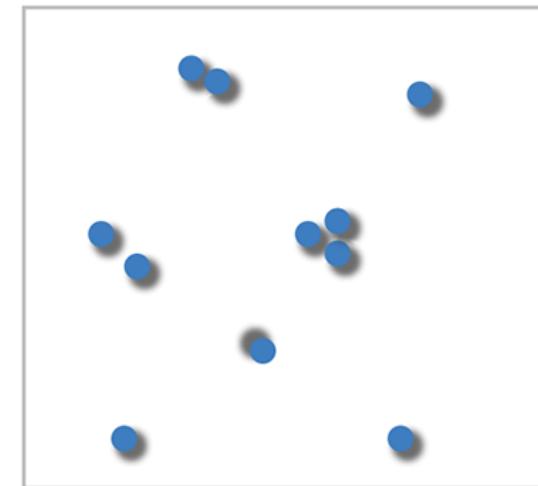
Size



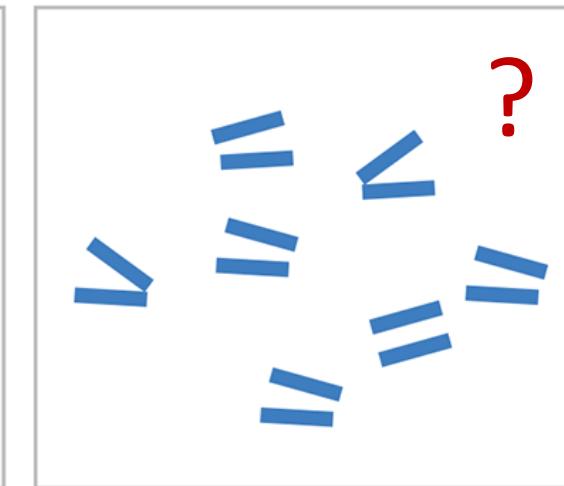
Shape



Proximity



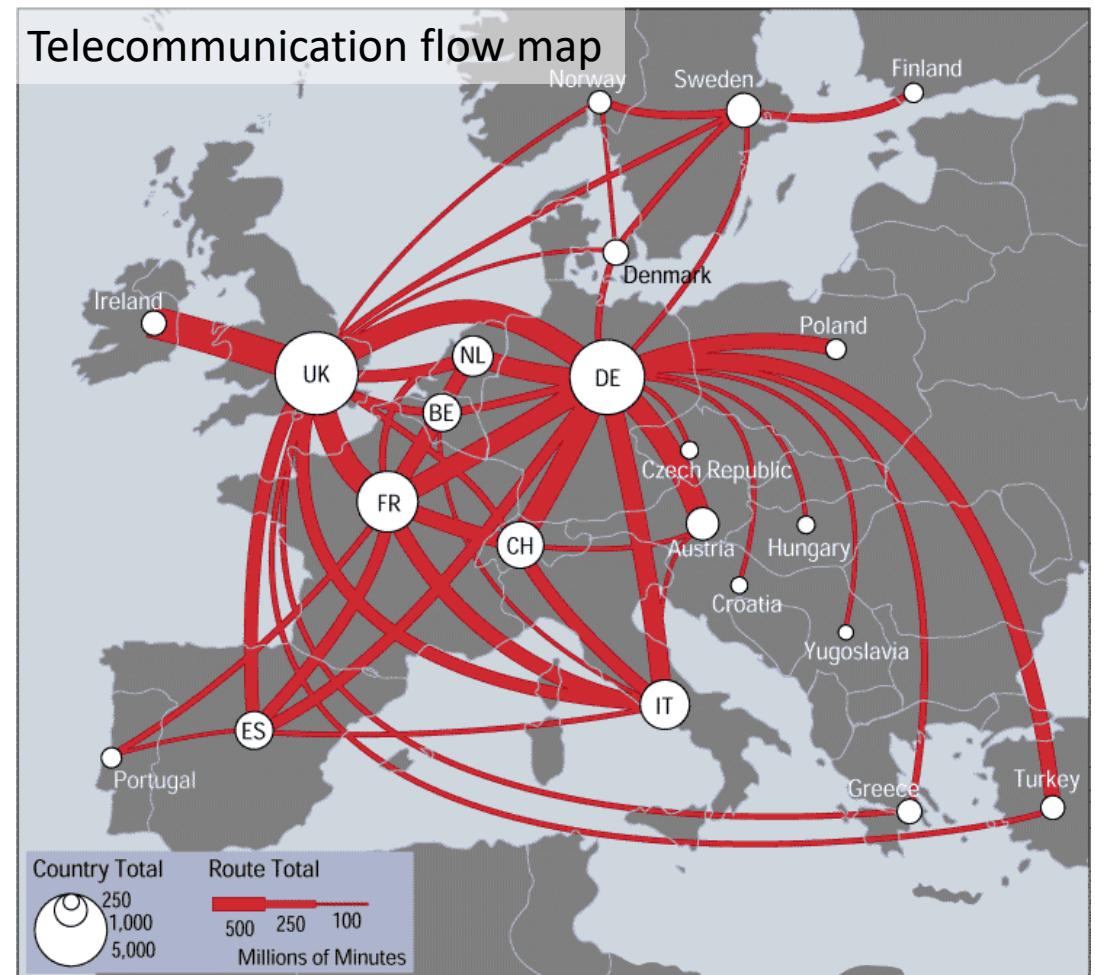
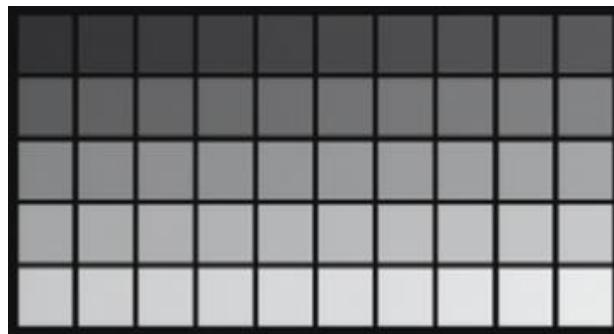
Shadow direction



Parallel lines (serial search)

# Mapping techniques

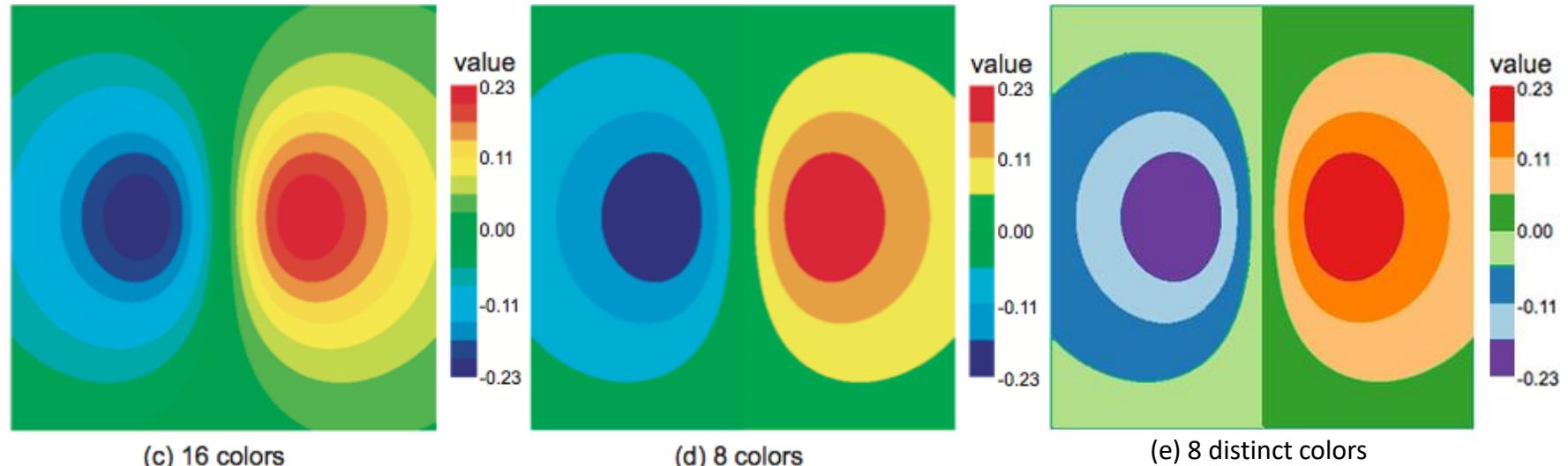
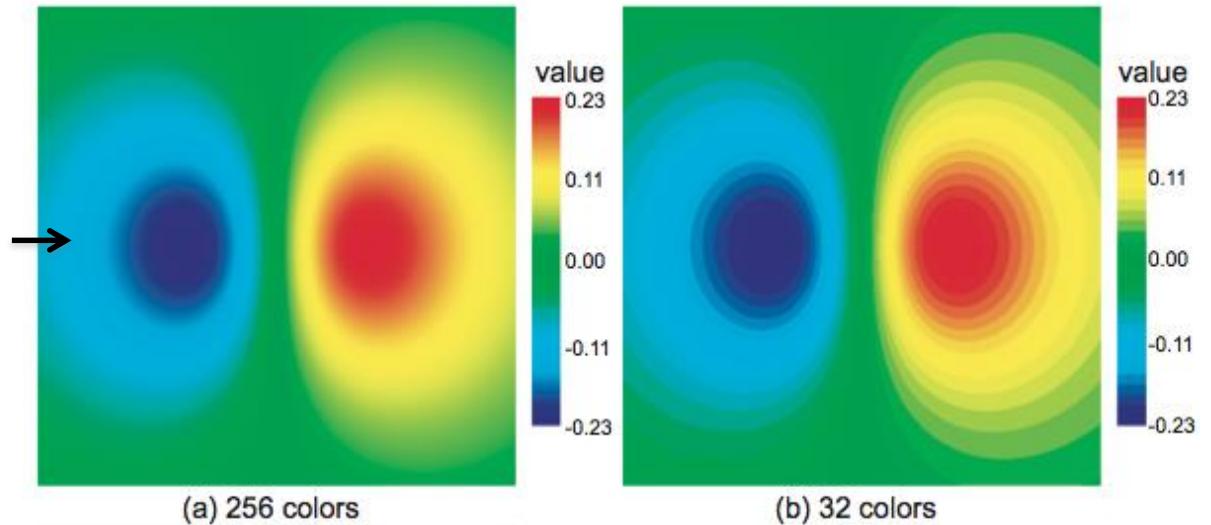
- Discriminability: How many usable steps?
  - Must be sufficient for number of discriminable bins



# Mapping techniques

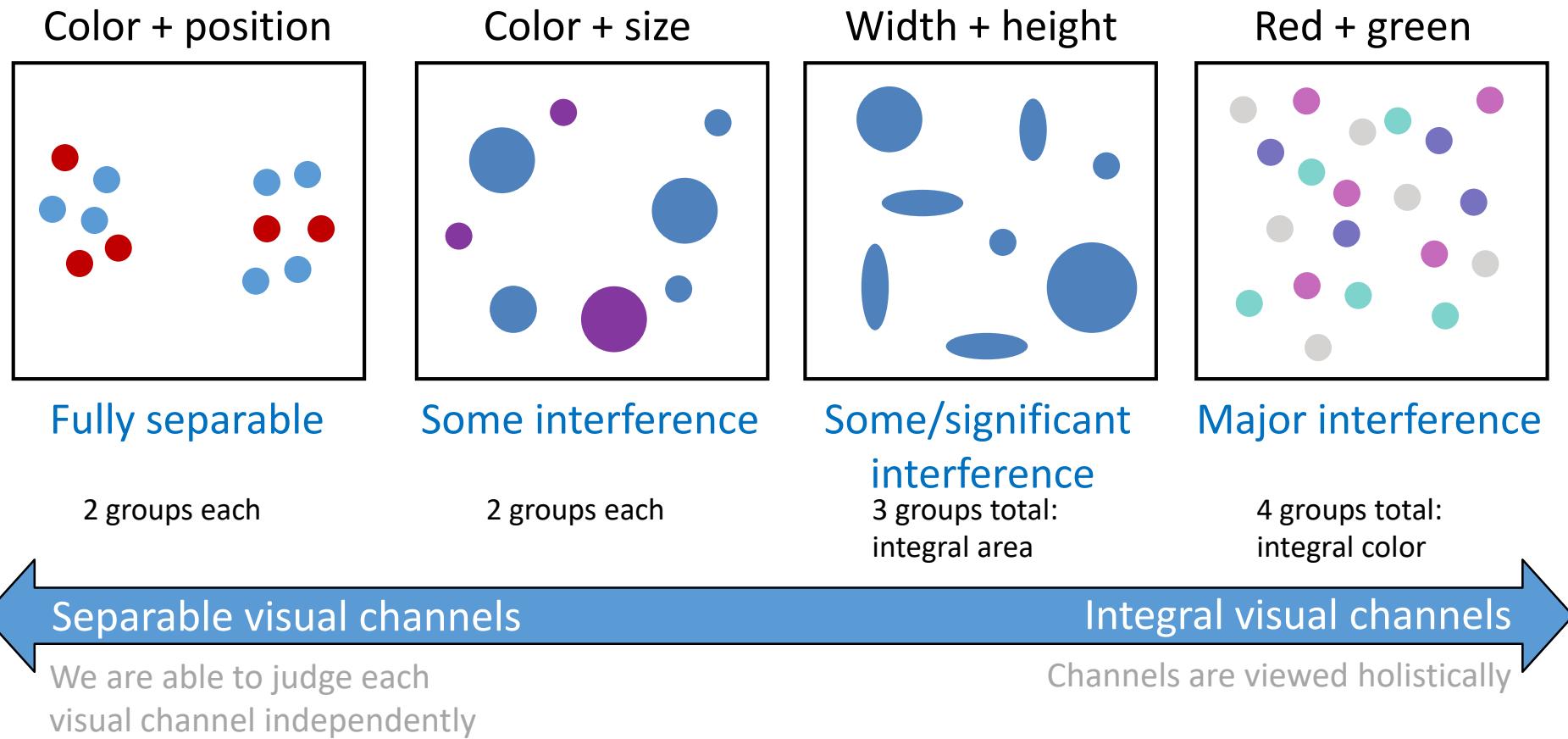
- Discriminability:  
How many  
usable steps?

Smoother results  
with more colors



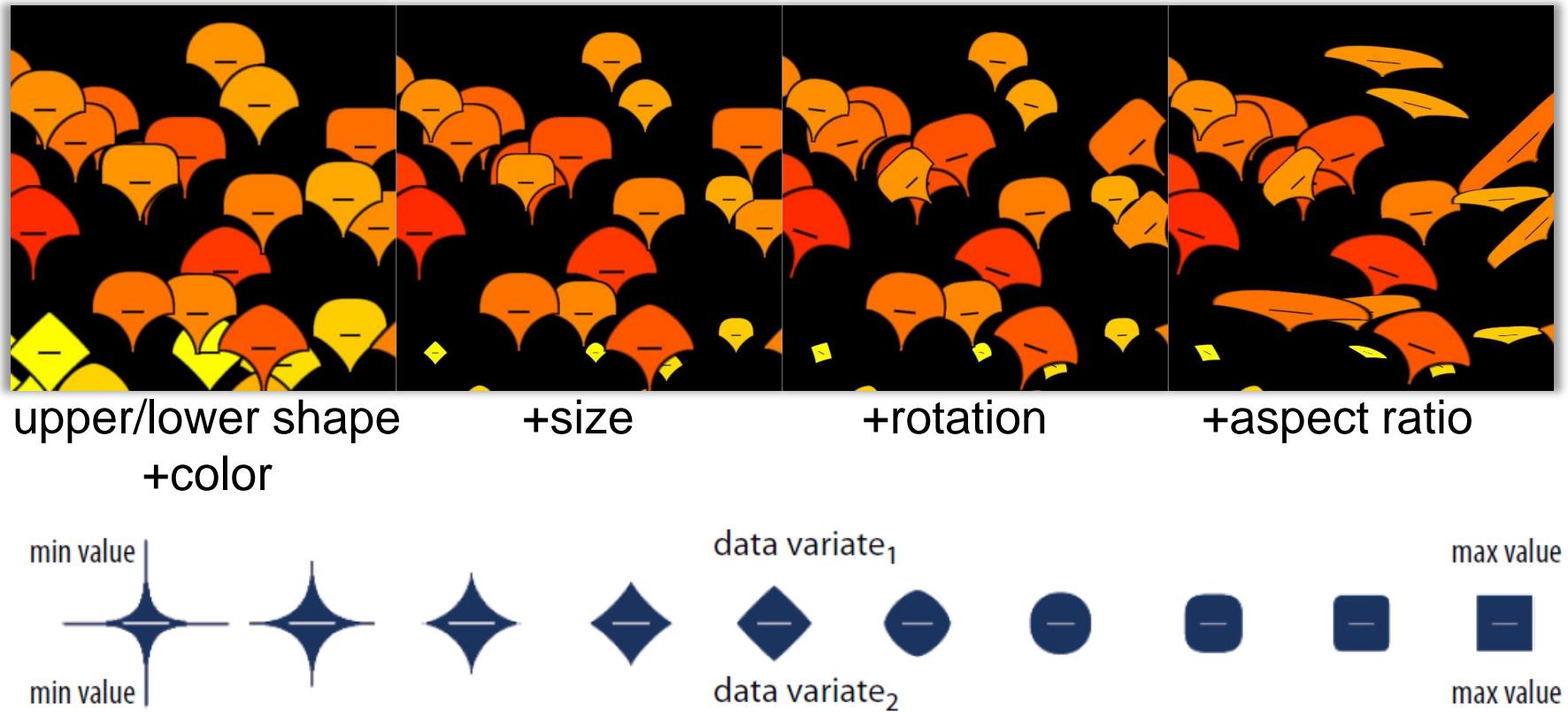
# Mapping techniques

- Separable vs. integral visual channels



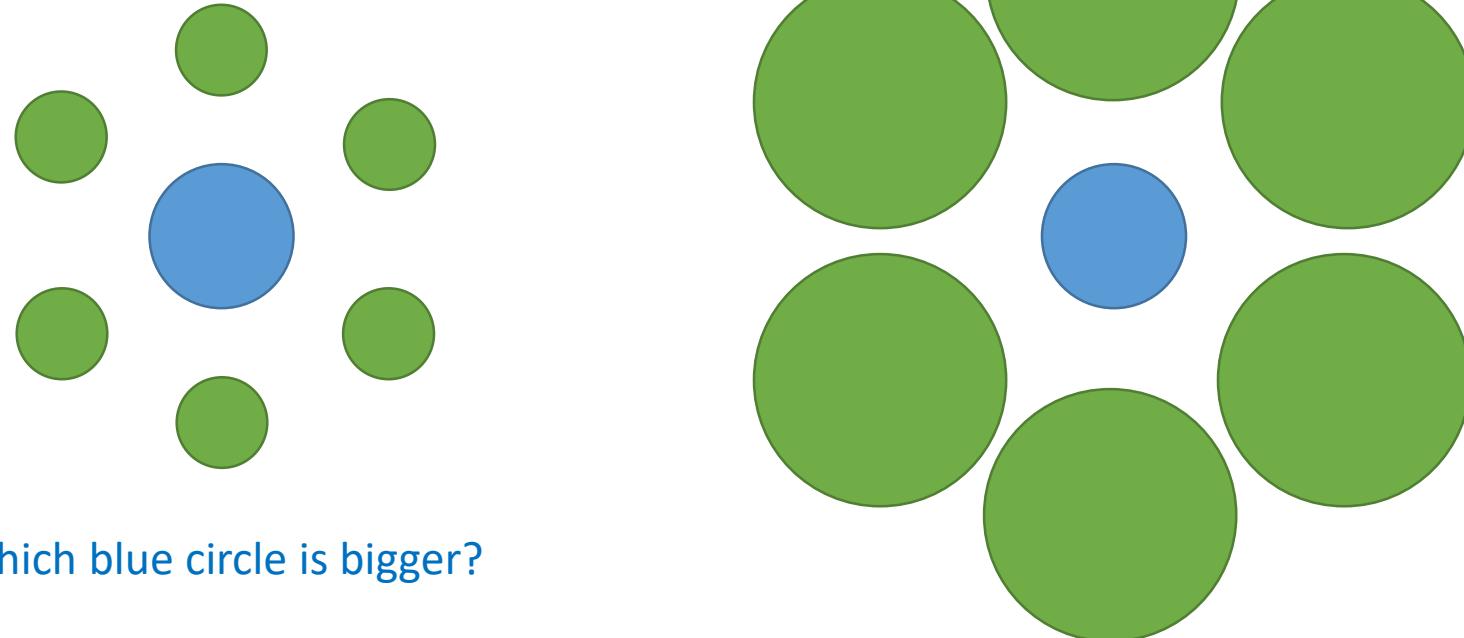
# Mapping techniques

- Separable vs. integral visual channels
  - Perceive each visual channel independently



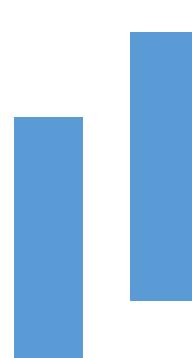
# Mapping techniques

- Relative vs. absolute judgments
  - Perception highly context-dependent
  - Perceptual system mostly operates with **relative judgments**, not absolute ones



# Mapping techniques

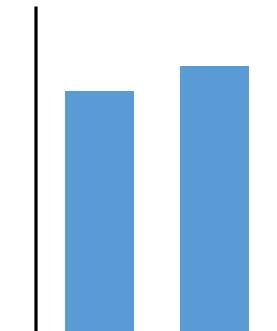
- Relative vs. absolute judgments
  - **Weber's Law:** just-noticeable difference is a fixed percentage of the magnitude of stimuli (e.g., bar length)
    - filled rectangles differ in length by 1:9 → difficult judgment
    - white rectangles differ in length by 1:2 → easy judgment



length



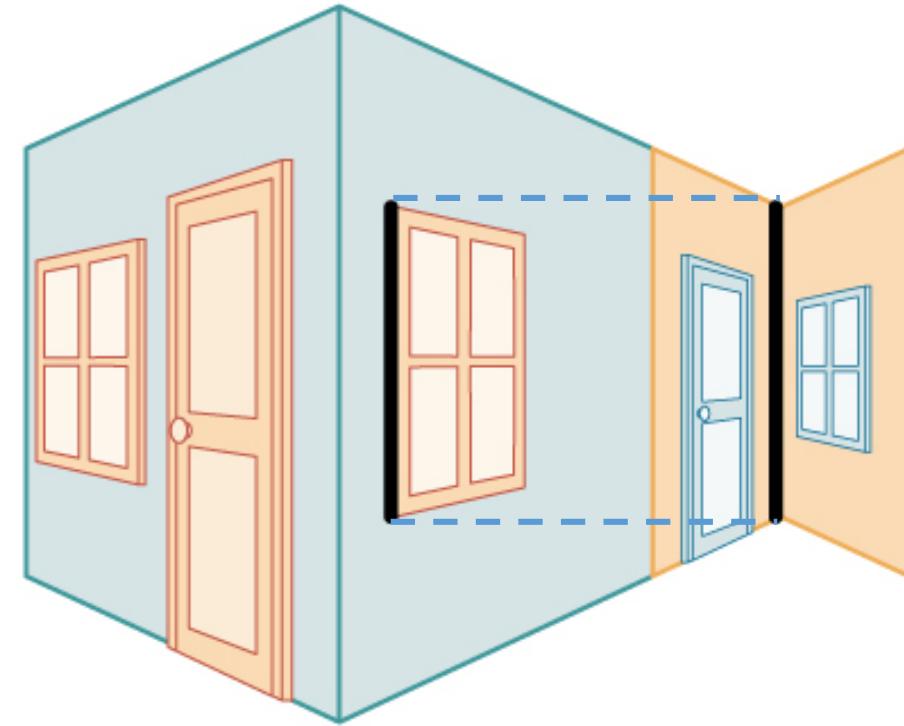
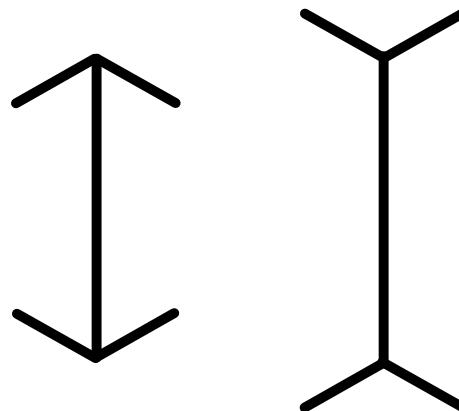
position along  
unaligned  
common scale



position along  
aligned scale

# Mapping techniques

- Relative vs. absolute judgments (Müller-Lyer illusion)

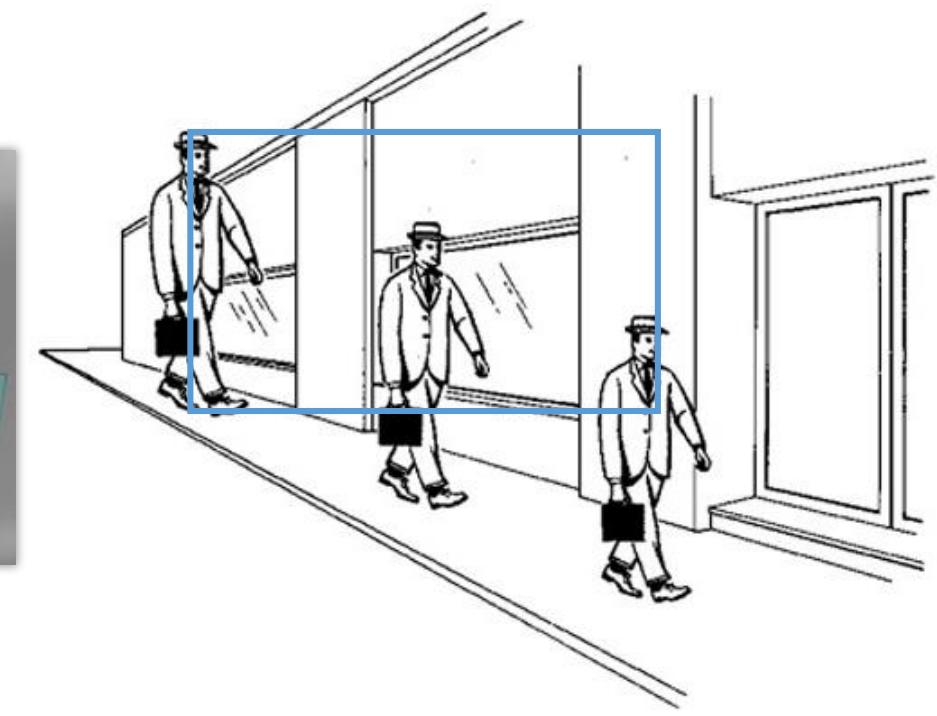


# Mapping techniques

- Relative vs. absolute judgments

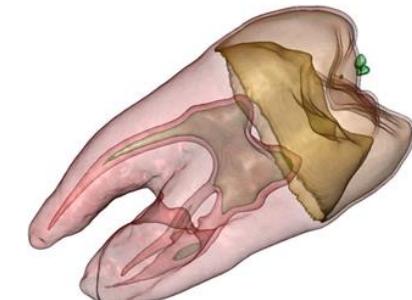
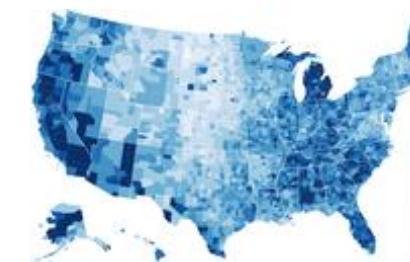
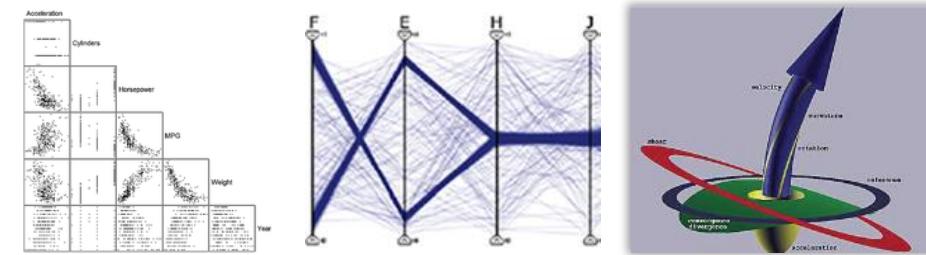
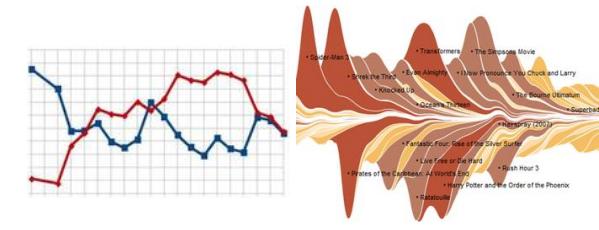
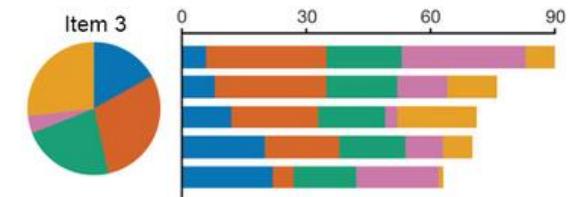


Ames Illusion



# Mapping data to primitives

- Categorical + quantitative data
  - Bar/pie chart, stacked bars
- Time-dependent data
  - Line graph, ThemeRiver, Horizon graph
- Single and multiple variables
  - Histogram, scatterplot, parallel coordinates, glyphs
  - Color mapping

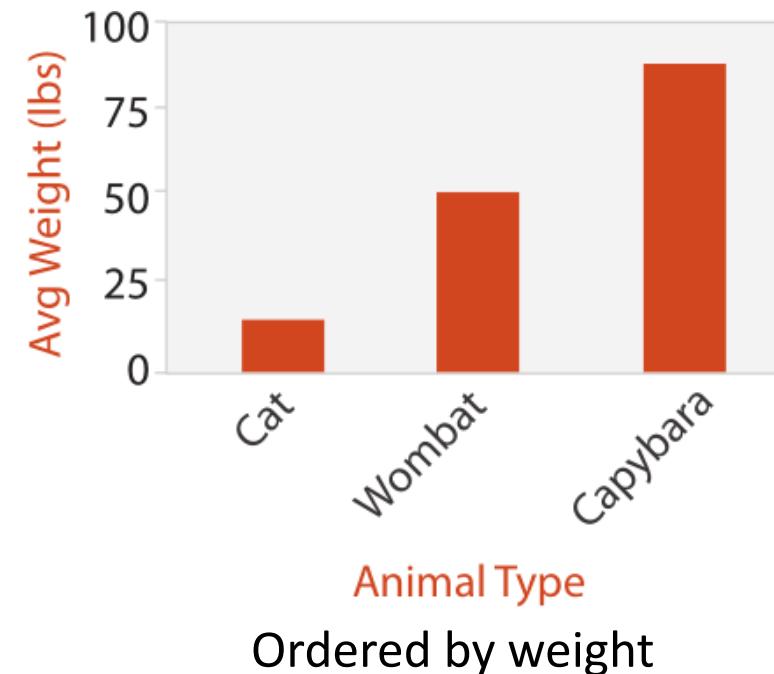
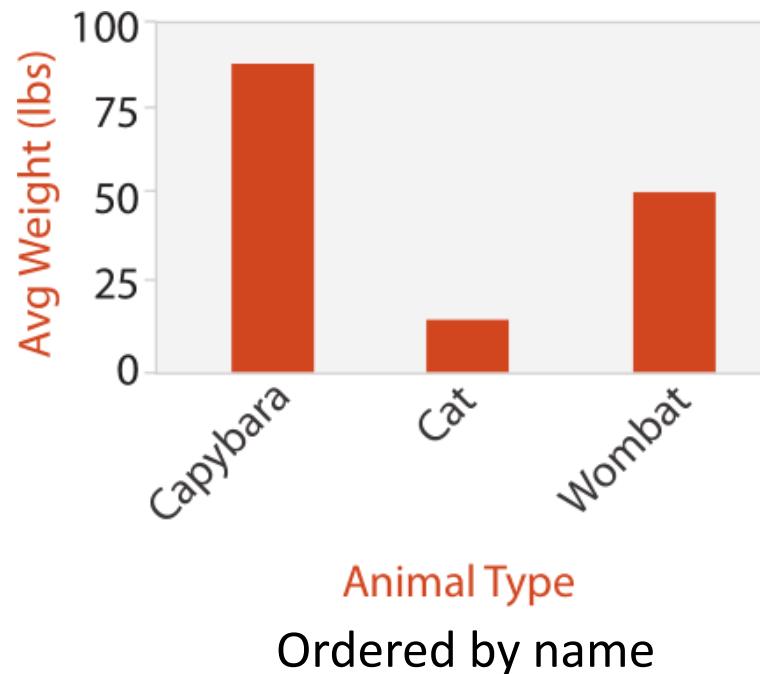


- Categorical + quantitative data
  - Bar/pie chart, stacked bars
- Time-dependent data
  - Line graph, ThemeRiver, Horizon graph
- Single and multiple variables
  - Histogram, scatterplot,  
parallel coordinates, glyphs
  - Color mapping
  - Iso-lines/surfaces

# Quantitative Data

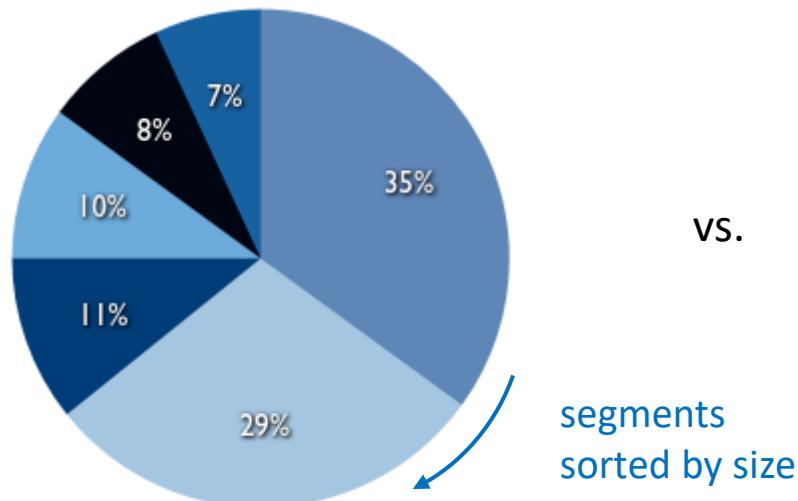
# Diagram techniques

- Bar chart
  - Attrib. 1: categorical → horizontal position
  - Attrib. 2: quantitative (dependent) → length/vertical position

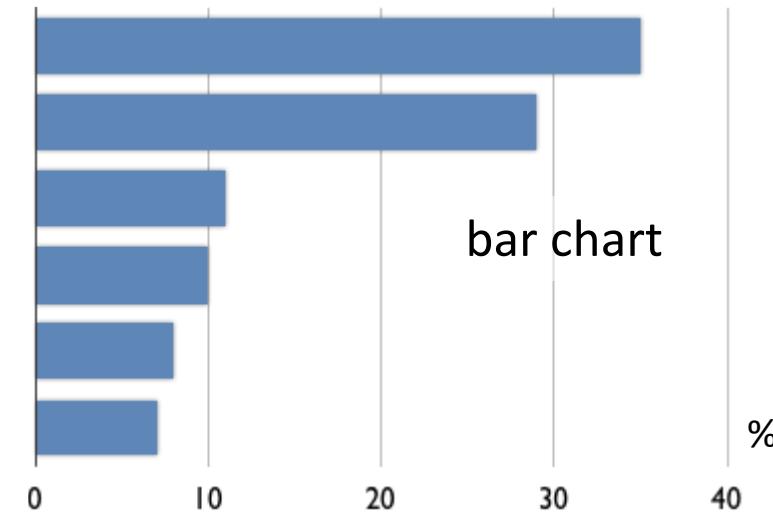


# Diagram techniques

- Pie chart
  - Pie chart splits population (100%) into parts
  - Attrib. 1: categorical → color
  - Attrib. 2: quantitative (dependent) → angle
  - However, angle/area less accurate than bar length

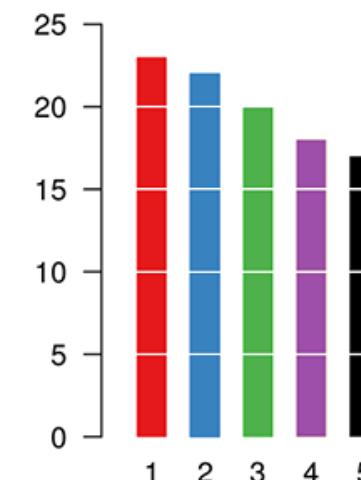
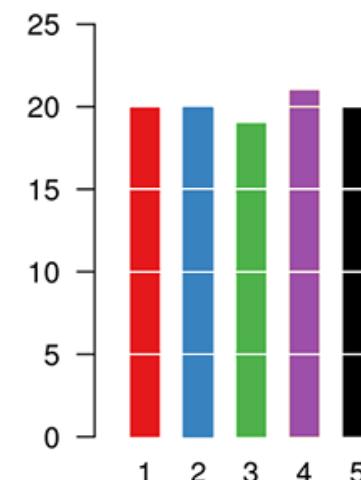
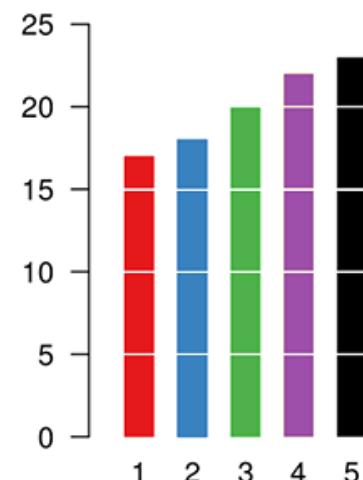
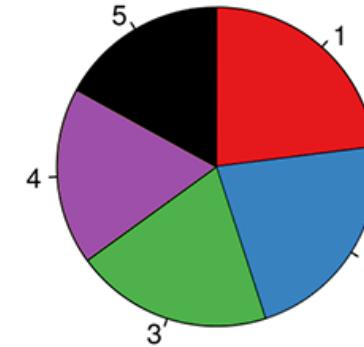
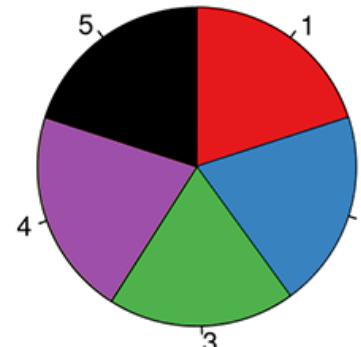
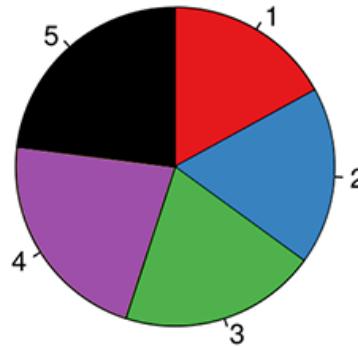


vs.



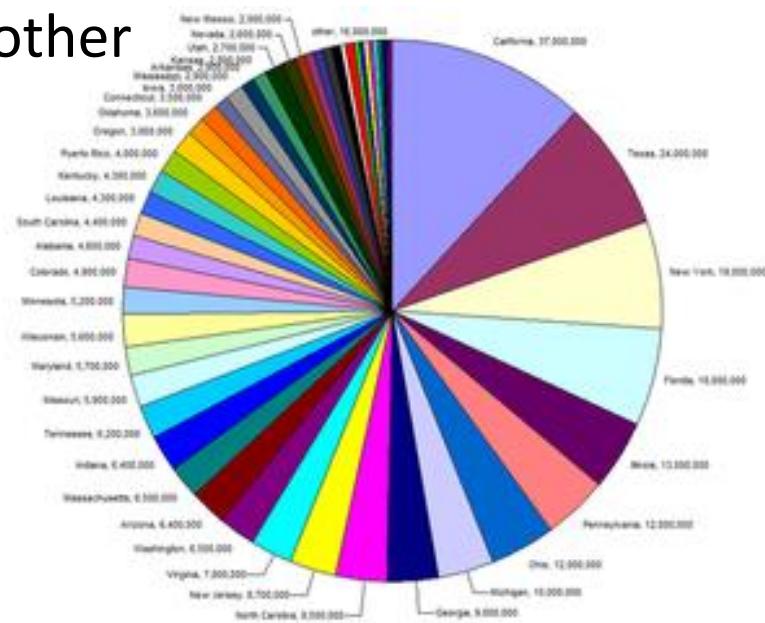
# Diagram techniques

- Pie chart vs. Bar chart
  - Angle/area judgment less accurate than bar length



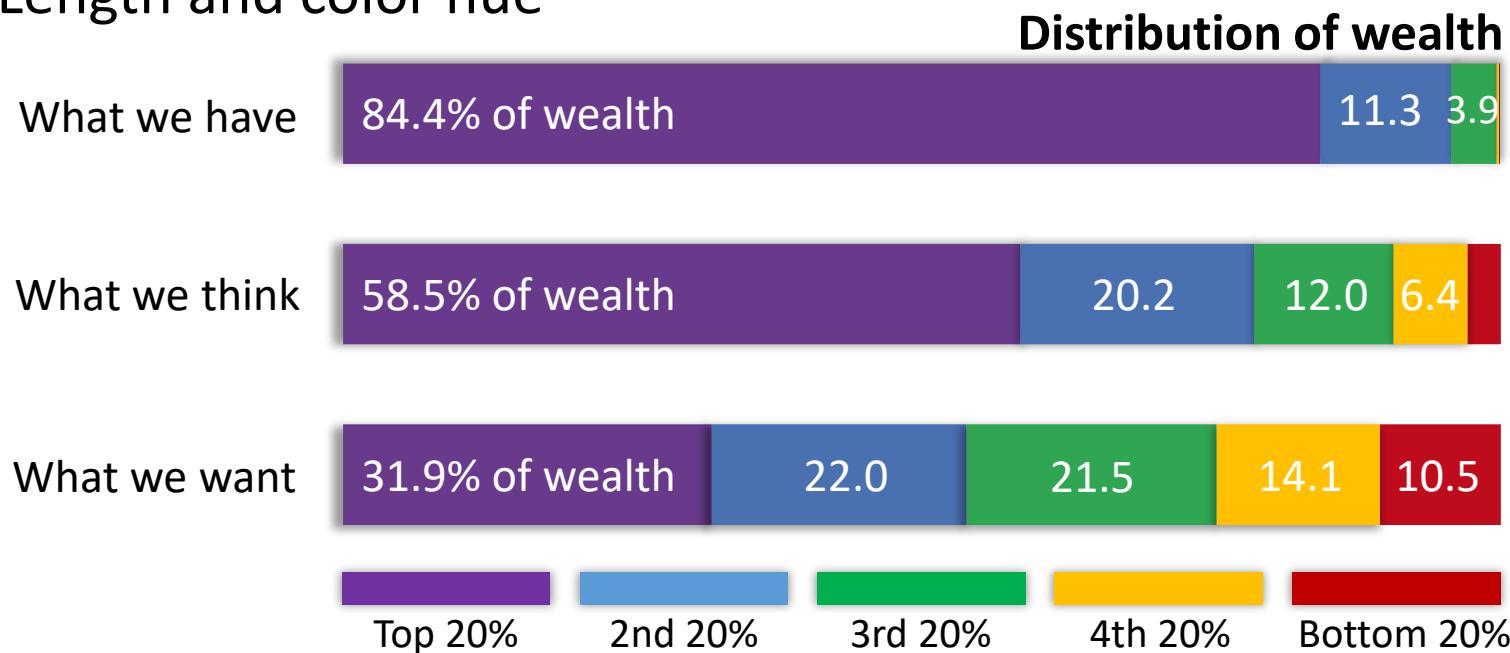
# Diagram techniques

- When to use a pie chart?
  - Often bar chart is a better choice!
  - Do the parts make up a meaningful whole?
  - Are the parts mutually exclusive?
  - Do you want to compare the parts to each other or the parts to the whole?
  - How many parts do you have?



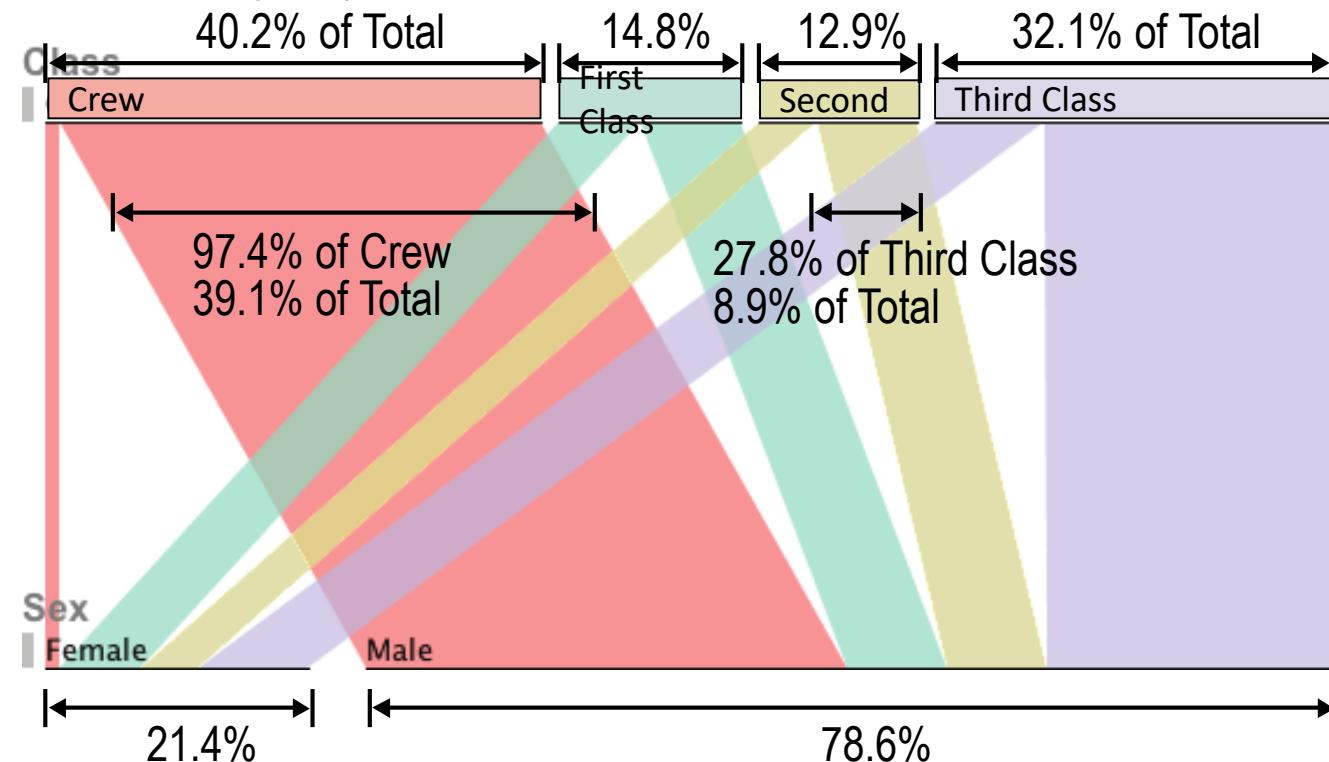
# Diagram techniques

- Stacked bar chart
  - Quantitative data wrt two categorical vars (horizontal & vertical)
  - Investigate part-to-whole relationship (100%)
  - Length and color hue



# Diagram techniques

- Parallel sets
  - Quantitative data wrt. multiple categorical attributes
  - Shows connections and proportions



# European asylum decisions in 2014

Origin

alpha » size »

Syria

Serbia

Eritrea

Afghanistan

Pakistan

Albania

Kosovo

First-instance Decisions

Accepted

alpha »

size »

Rejected

Destination

Germany

alpha »

size »

Sweden

France

Nether...

Switze...

Unite...

Italy

Gr...

Au...

Based on the infographics "Seeking safety." The Economist, 2015

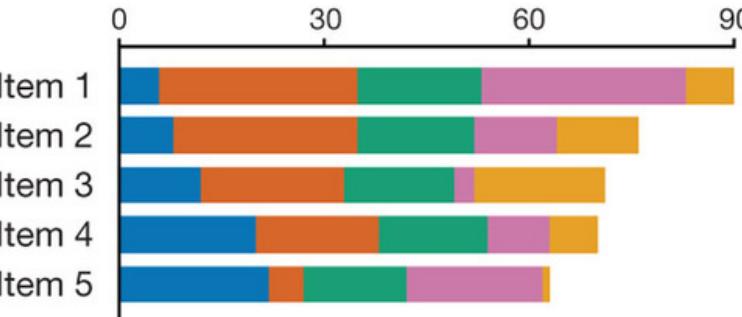
Data: Eurostat  
\*Austria: data from 2013

[multivis.net/lecture/parallel-sets.htm](http://multivis.net/lecture/parallel-sets.htm)

# Diagram techniques

# Variations of pie & bar charts

- Category 1
- Category 2
- Category 3
- Category 4
- Category 5

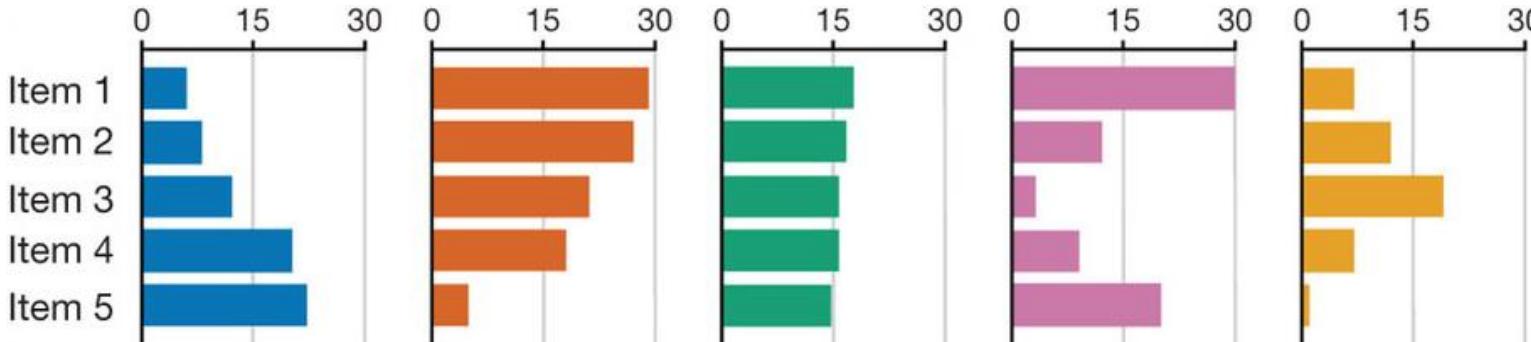


**Pie chart:** compare values in different categories

**Stacked bar charts:** compare overall values across items, but also show contribution per category



**Grouped bar charts:** compare values across categories within each item

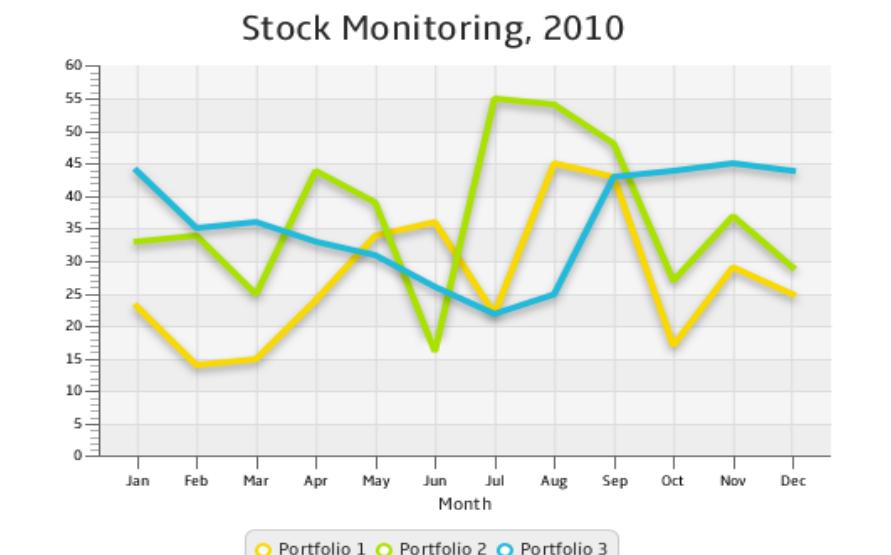
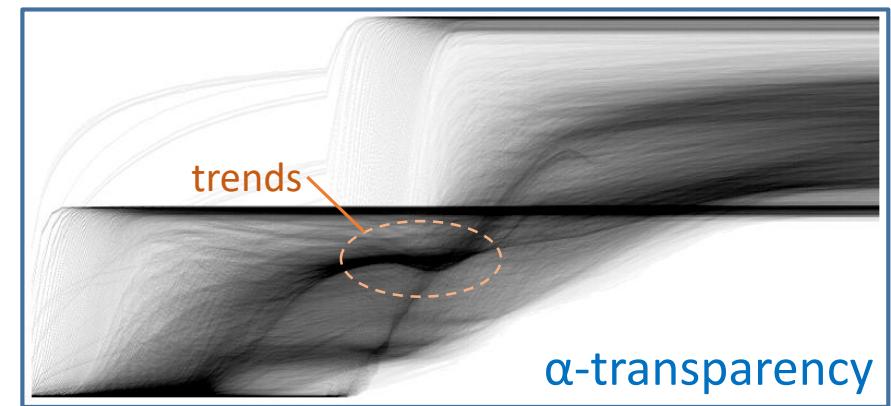
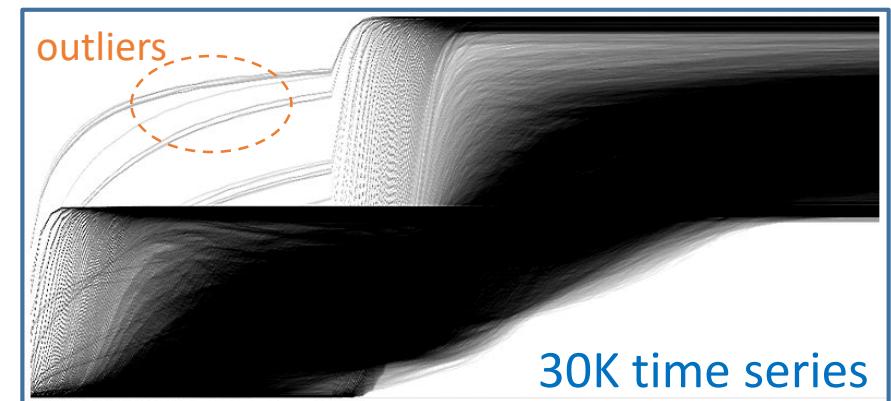


**Layered bar charts:** compare values within categories

# Time-dependent Data

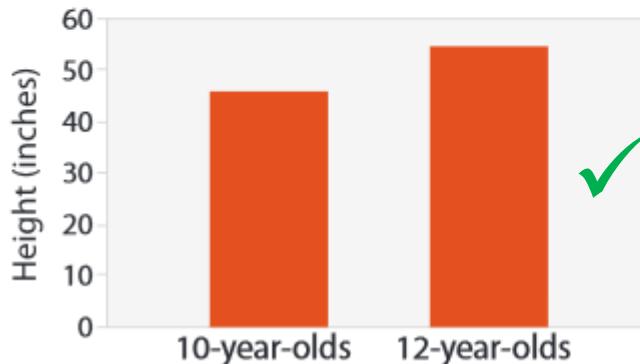
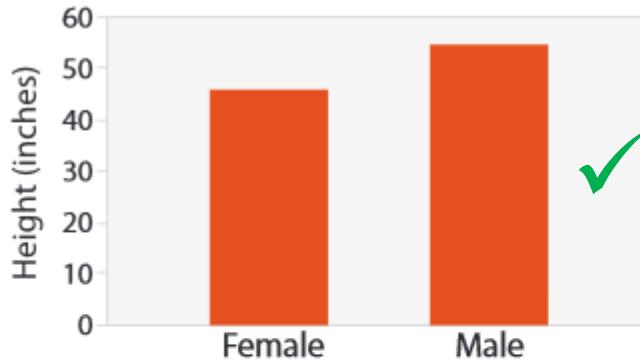
# Diagram techniques

- Line graphs
  - Quantitative data on common scale(s) wrt. time
  - Connection between points – trends, structures, groups



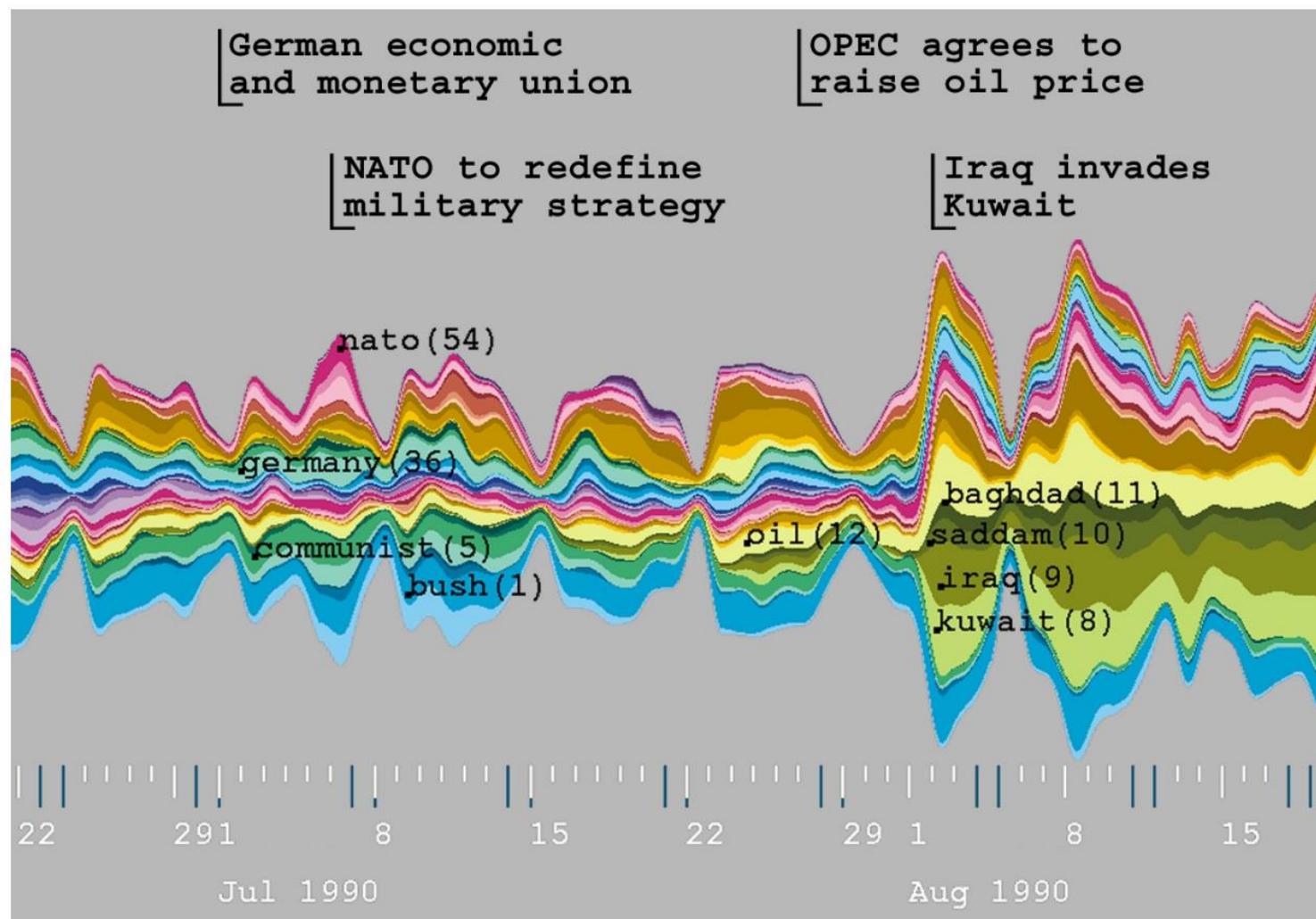
# Diagram techniques

- Choosing bar vs. line charts
  - Bars support comparison
  - Lines imply trends

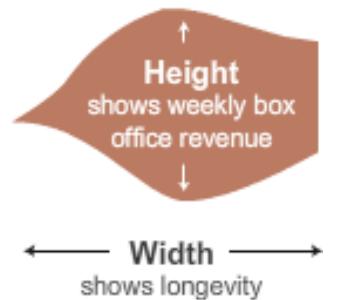
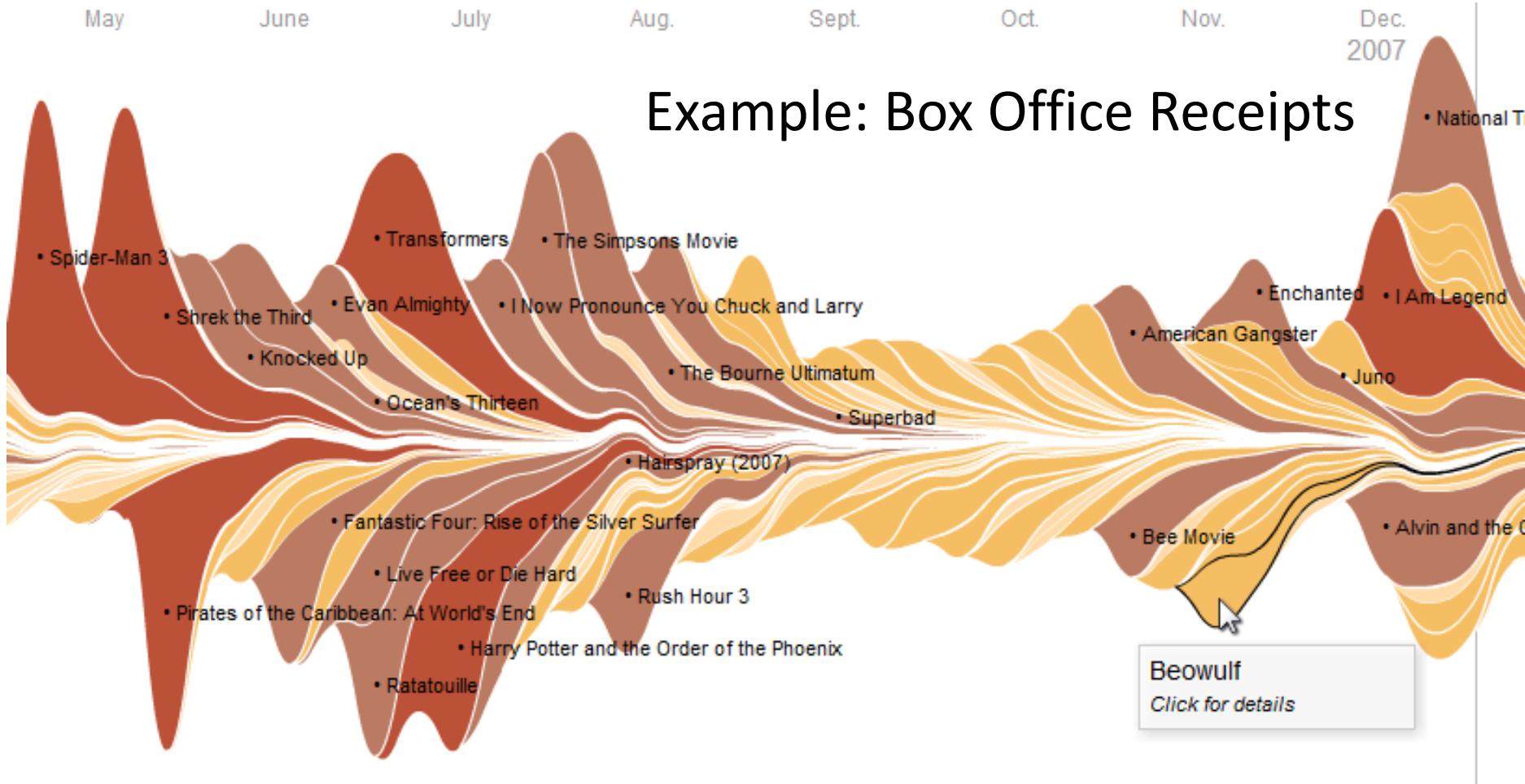


# Diagram techniques

- ThemeRiver
  - Thematic changes in documents
  - Occurrence per topic mapped to width of bands
  - Less distorted around center  
→ rearrange bands



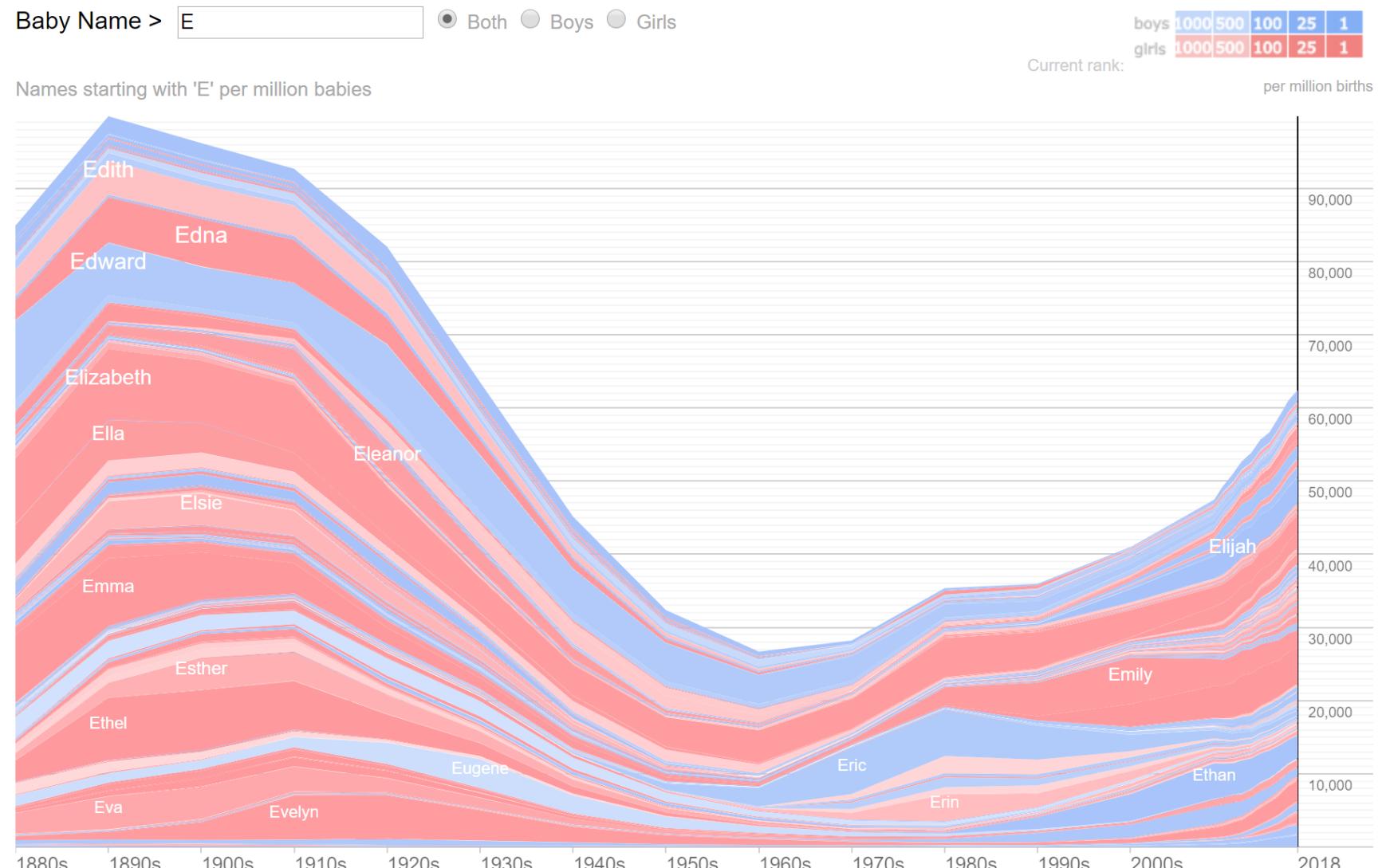
# Diagram techniques



Each shape shows how one film did at the box office

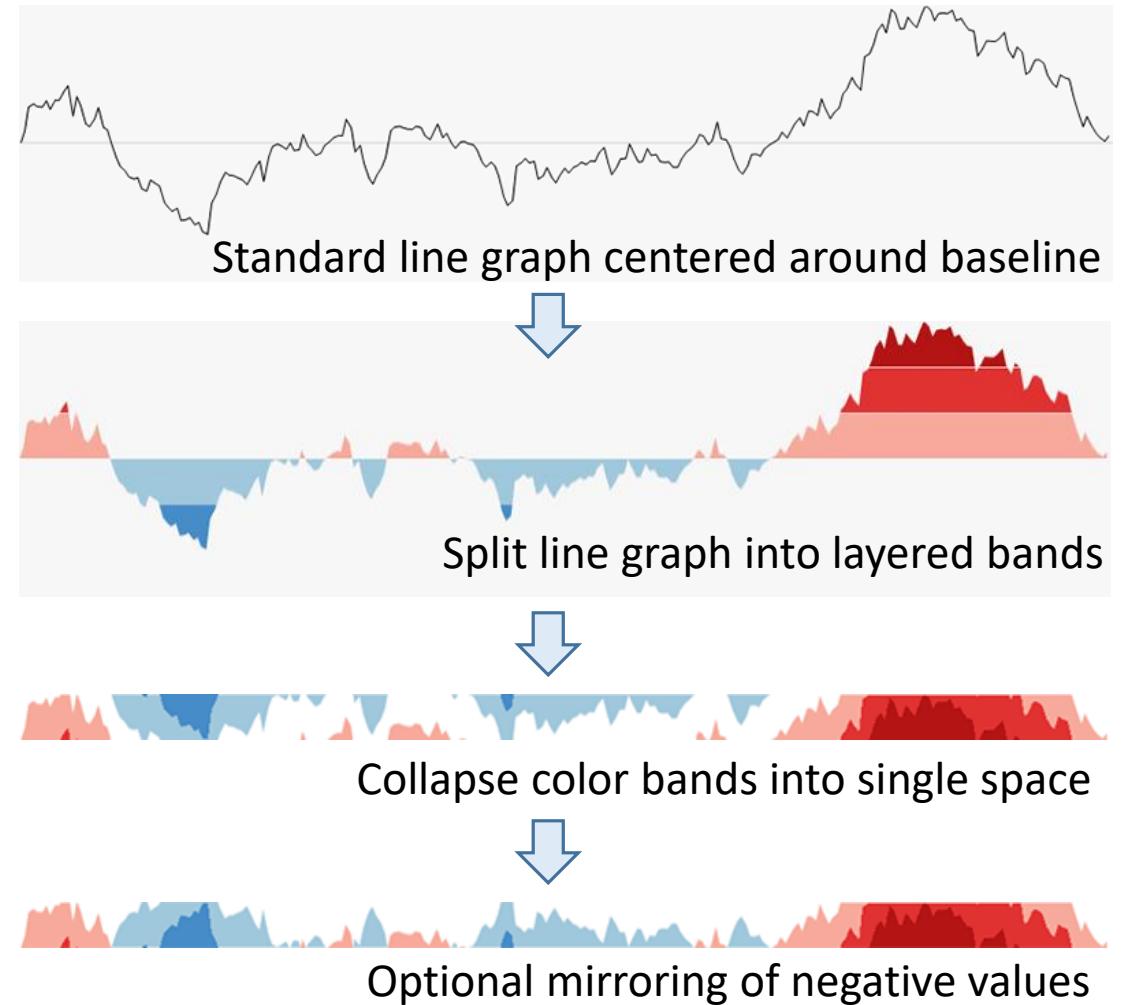
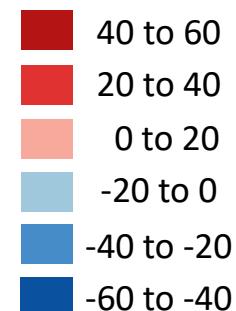
The **area** of the shape (and its **color**) correspond to the film's total domestic gross, through Feb. 21.

# Diagram techniques



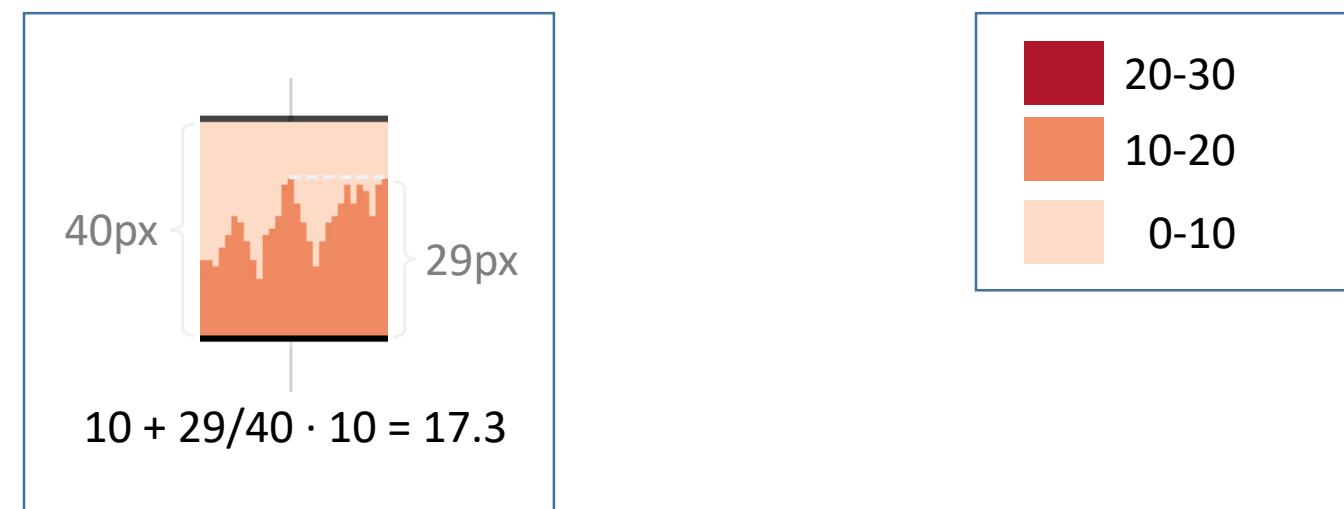
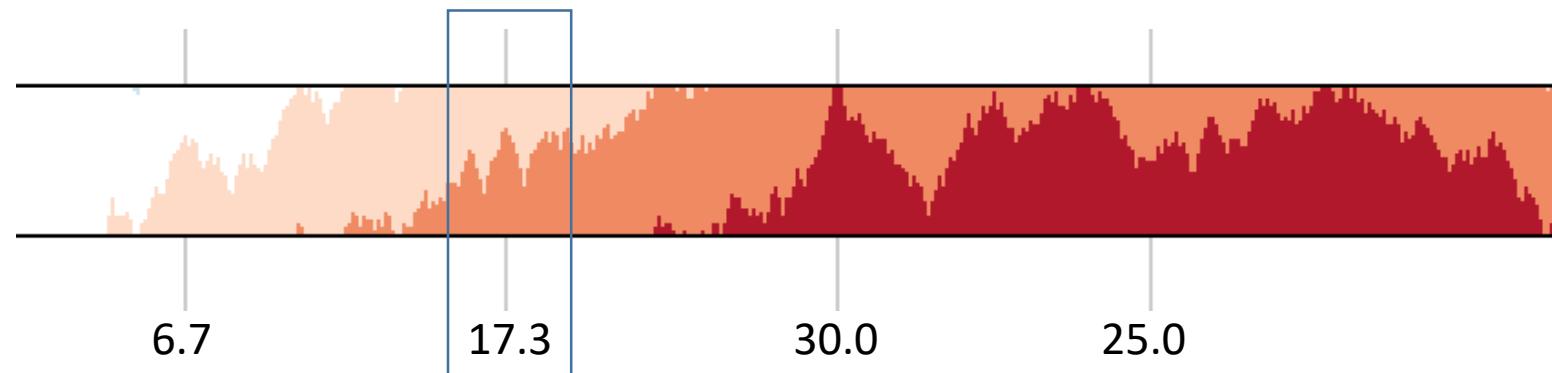
# Diagram techniques

- Horizon graphs
  - Reduces vertical space without losing precision



# Diagram techniques

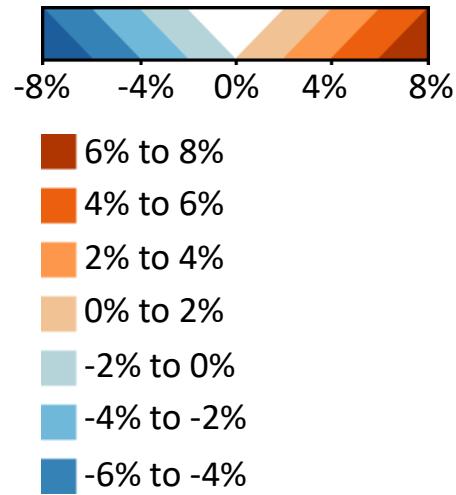
- Example



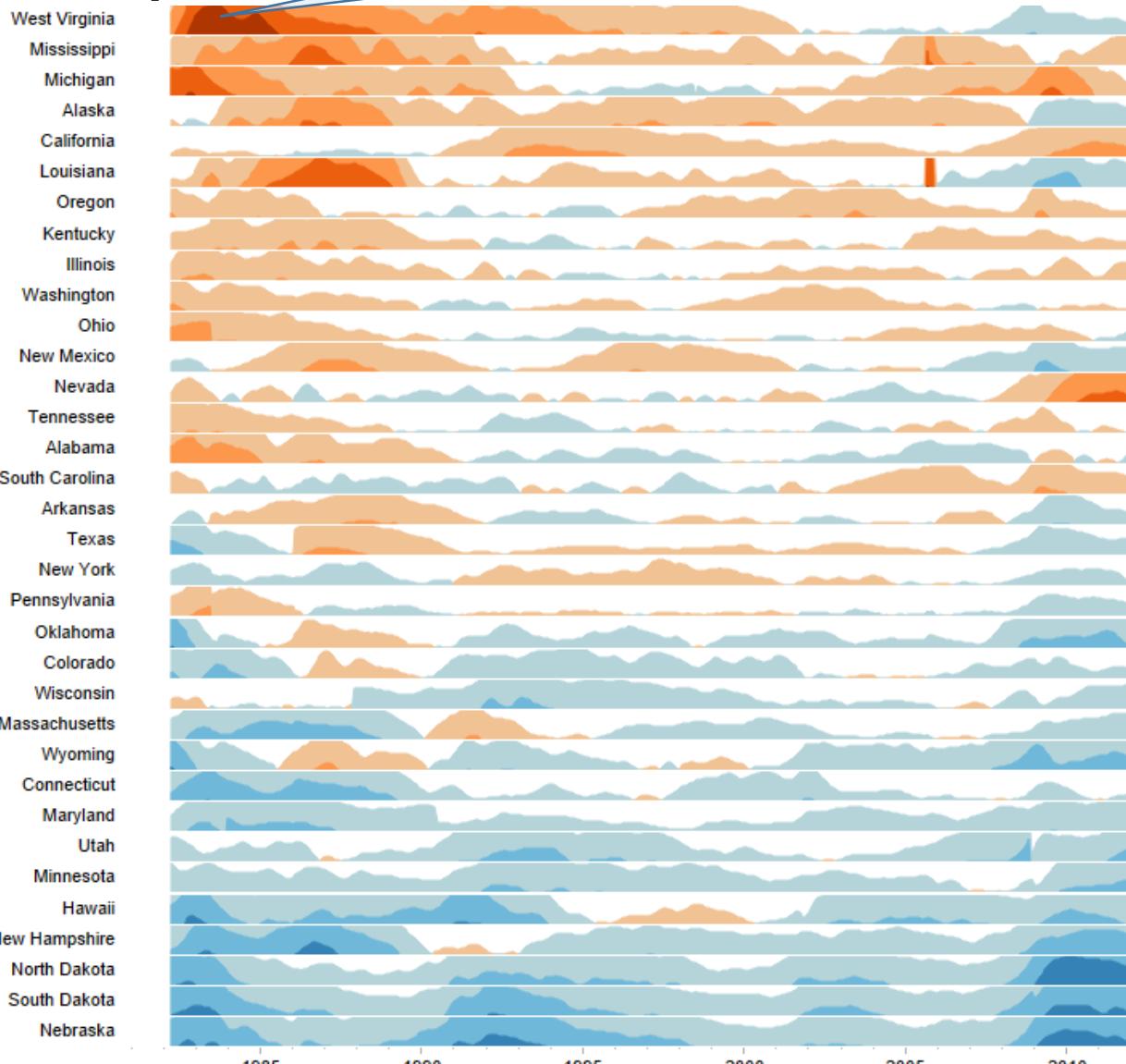
# Diagram techniques

Unemployment  
rate (1982-2012)

Difference from average



[www.tableau.com/learn/gallery/unemployment-horizon-chart](http://www.tableau.com/learn/gallery/unemployment-horizon-chart)



# Diagram techniques

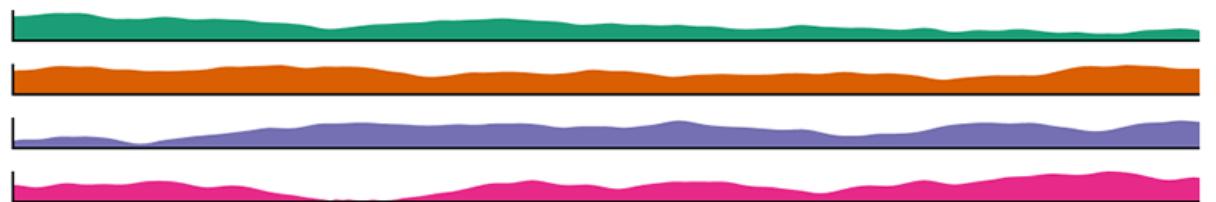
- Looking at multiple time series



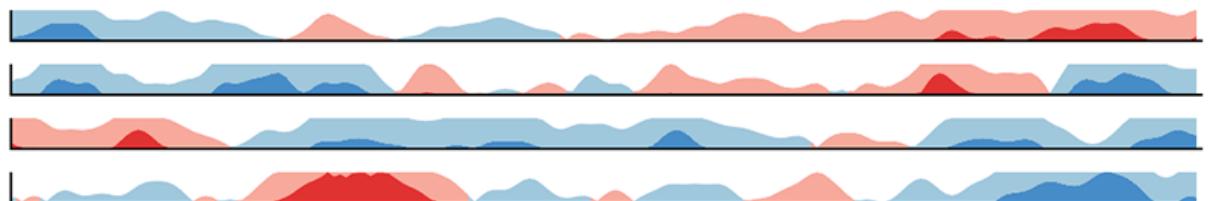
Simple line graphs



Stacked graphs



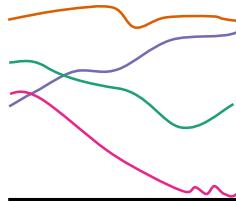
Small multiples



Horizon graphs

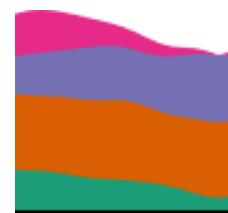
# Diagram techniques

## Simple line graphs



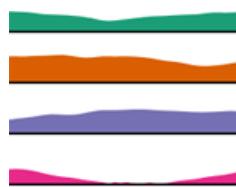
- + direct comparison
- + observe fine details/slopes
- difficult to distinguish different time series
- visual clutter

## Stacked graphs



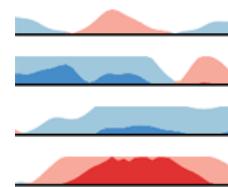
- + easy to distinguish time series
- less vertical resolution
- difficult to compare time series due to distortion

## Small multiples



- + easy to distinguish time series
- less vertical resolution per time series
- difficult to compare time series that are far apart

## Horizon graphs

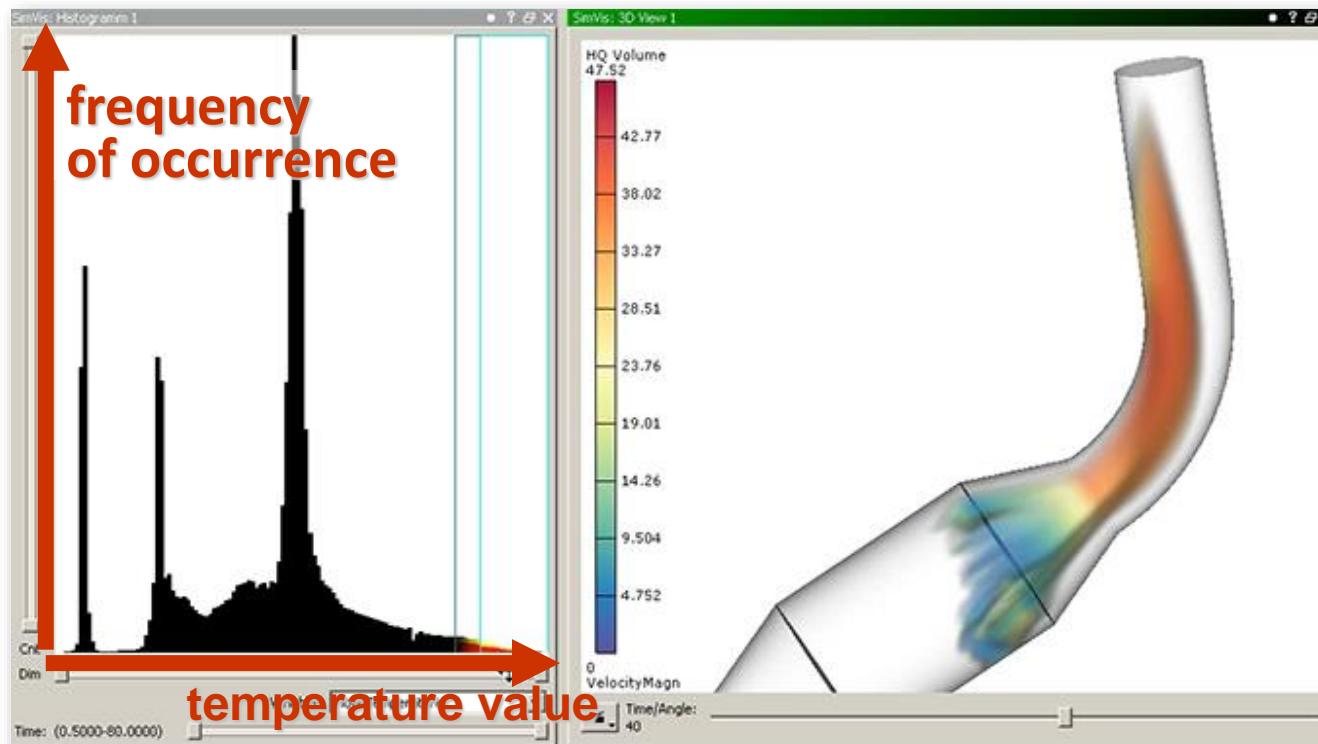


- + preserve vertical resolution
- + quickly spot min./max. values
- steep learning curve
- difficult comparison between time series

# Single and Multiple Variables

# Diagram techniques

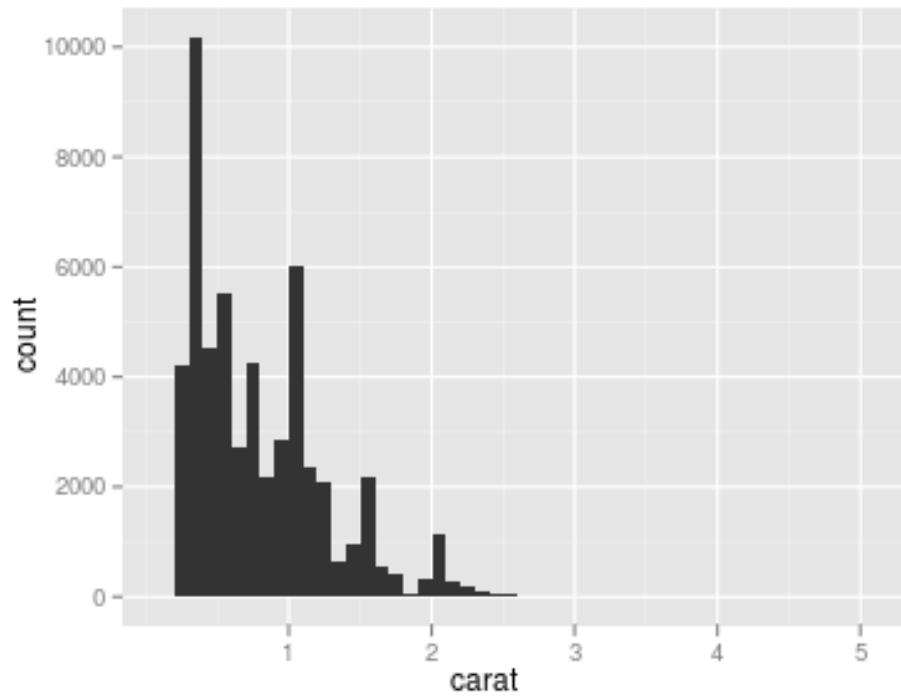
- Histogram (graphical display of statistics)
  - Special kind of bar chart
  - Show frequency of occurrence of values for one variable



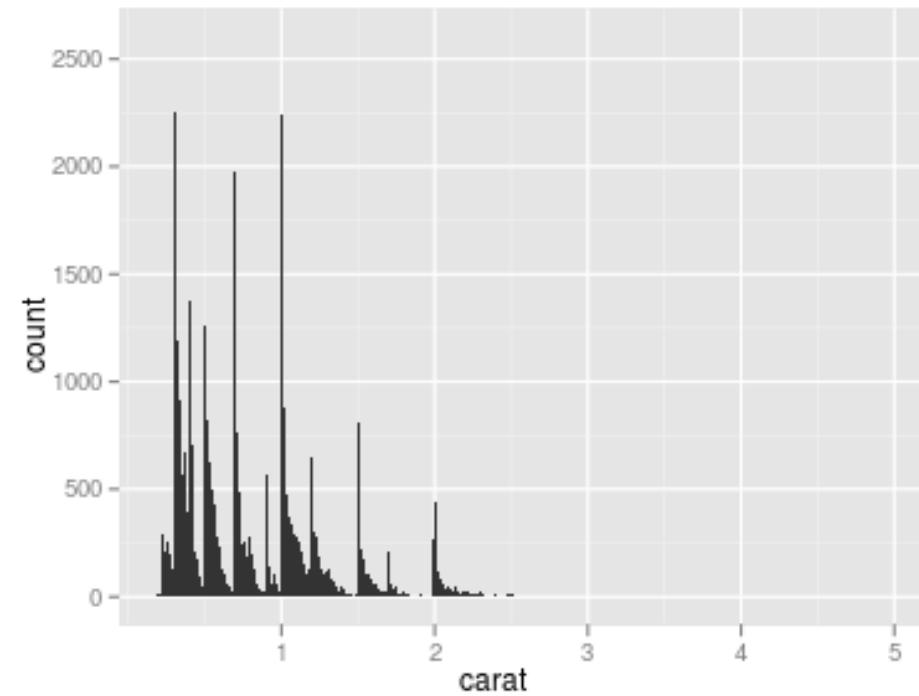
Diesel particulate filter

# Diagram techniques

- Histogram
  - Binning: group values into equally spaced intervals (bins)
  - Bin width affects representation



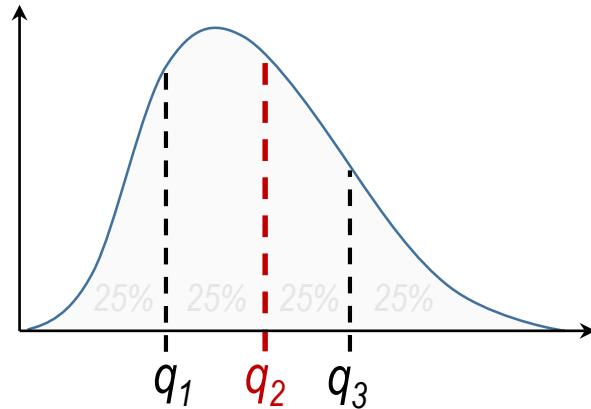
bin width = 0.1



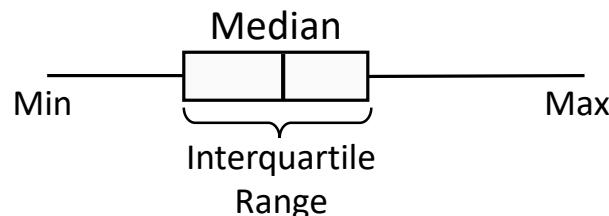
bin width = 0.01

# Diagram techniques

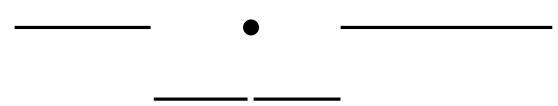
- Box plot variations
  - Shows summary statistics of a distribution (1 variable)



Probability density function (PDF)  
 $q_1$  ... lower quartile  
 $q_2$  ... median  
 $q_3$  ... upper quartile



Tukey's box plot



Tufte's quartile plot variations

# Diagram techniques

$q_1$  ... lower quartile

$q_2$  ... median

$q_3$  ... upper quartile

# Diagram techniques

$q_1$  ... lower quartile

$q_2$  ... median

$q_3$  ... upper quartile

2,3,4,1,5,2,6,4,7,4,7,9,2,3,4,3,5,6,5,2,3,5

# Diagram techniques

$q_1$  ... lower quartile

$q_2$  ... median

$q_3$  ... upper quartile

2,3,4,1,5,2,6,4,7,4,7,9,2,3,4,3,5,6,5,2,3,5

Sort

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.
1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5	6	6	7	7	9

# Diagram techniques

$q_1$  ... lower quartile

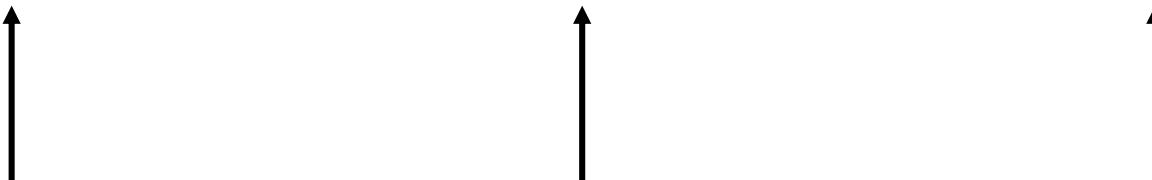
$q_2$  ... median

$q_3$  ... upper quartile

2,3,4,1,5,2,6,4,7,4,7,9,2,3,4,3,5,6,5,2,3,5

Sort

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.
1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5	6	6	7	7	9



# Diagram techniques

$$q_1 = 3$$

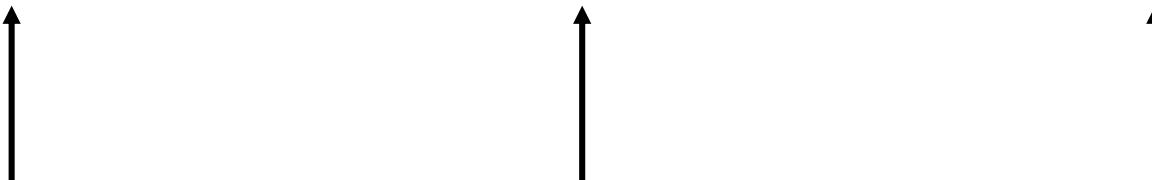
$$q_2 = 4$$

$$q_3 = 5$$

2,3,4,1,5,2,6,4,7,4,7,9,2,3,4,3,5,6,5,2,3,5

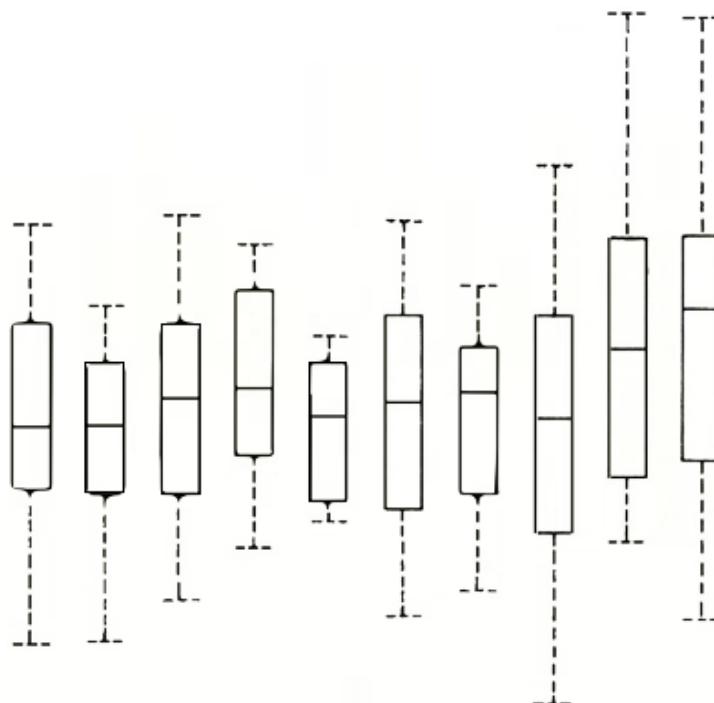
Sort

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.
1	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5	6	6	7	7	9

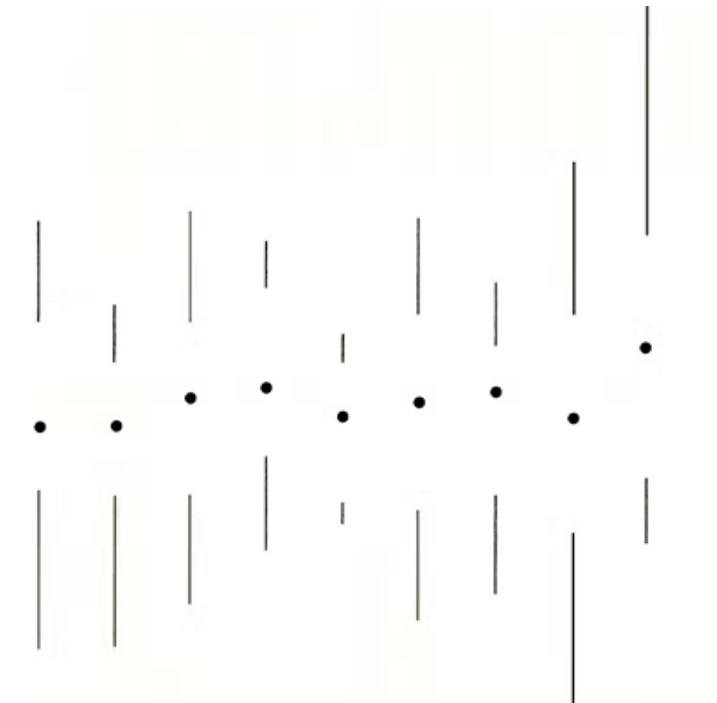


# Diagram techniques

- Box plot variations



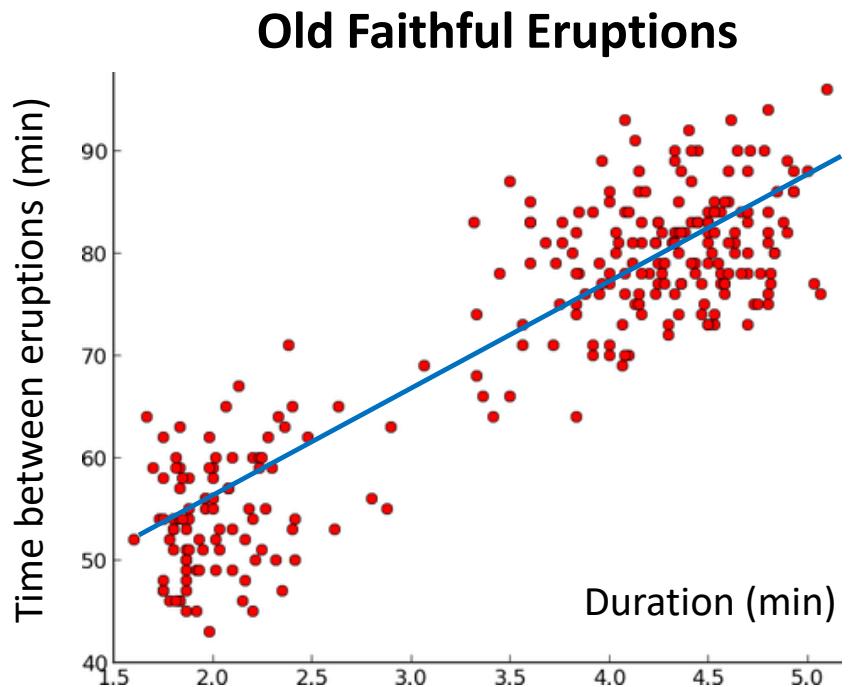
Tukey's box plot



Tufte's quartile plot

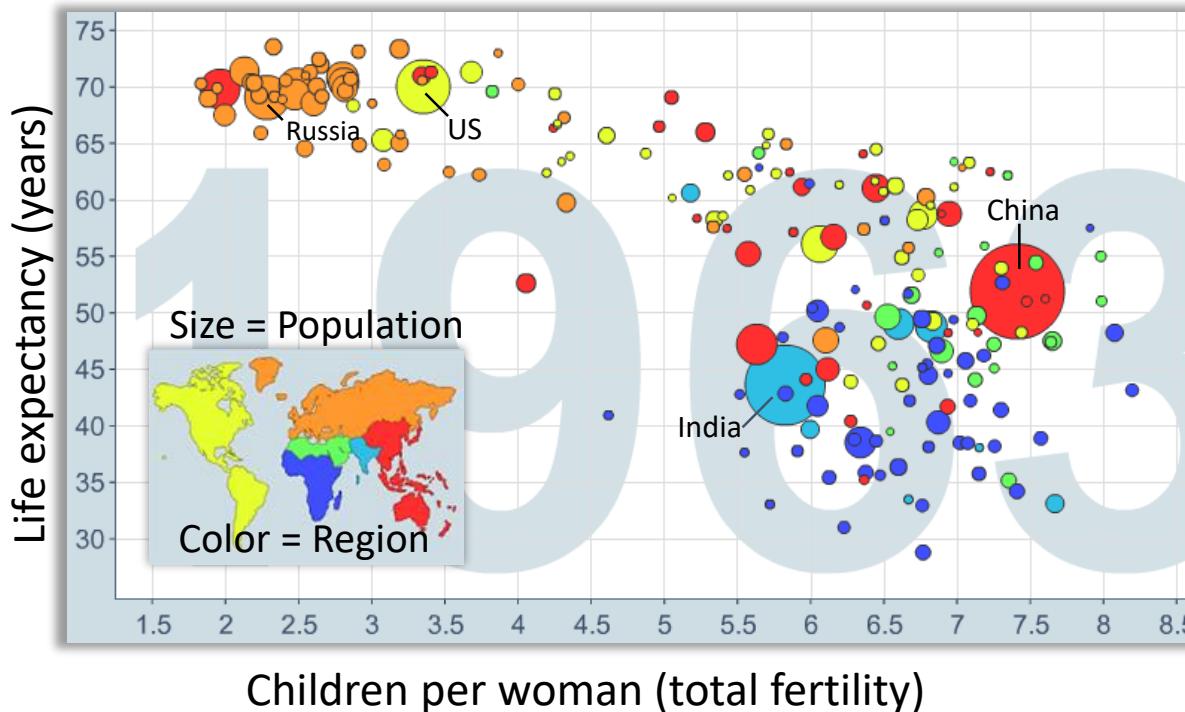
# Diagram techniques

- Scatterplots
  - Show correlations between 2 dependent variables
  - Typically quantitative (measurable) data attrib.
  - Find trends, outliers, distributions, correlations, clusters, ...



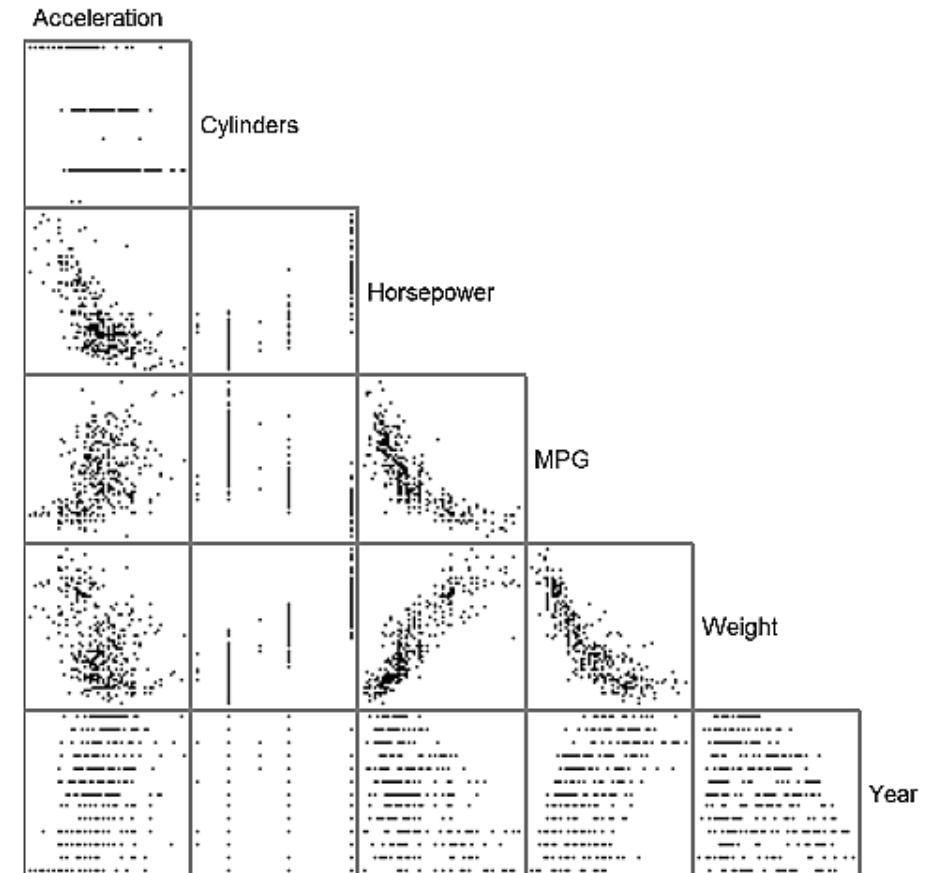
# Diagram techniques

- Scatterplot variations
  - Encode additional attributes by size, color, shape, ...



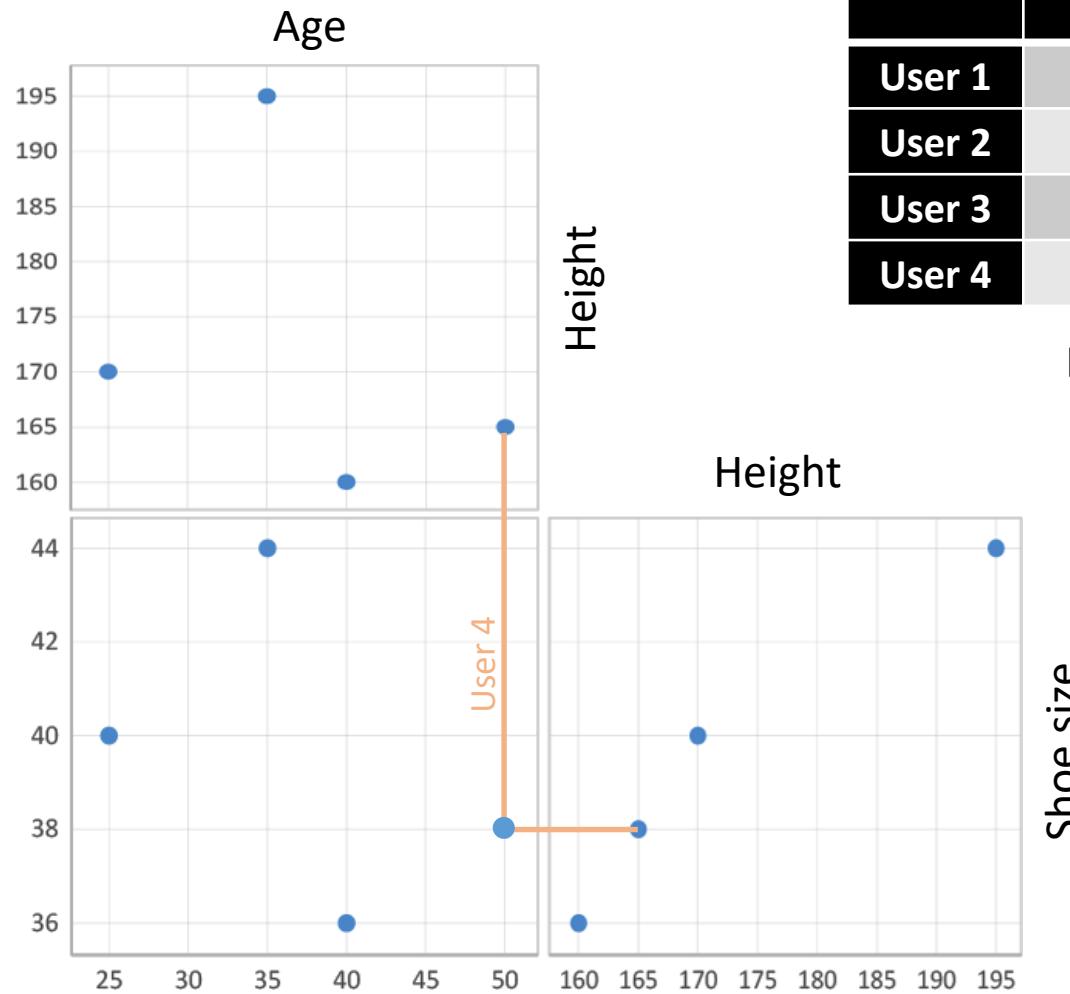
# Diagram techniques

- Scatterplot matrix
  - Show (all possible) combinations of attributes in a scatterplot matrix
  - Each row/column is one attribute
  - Overview of correlation and patterns between data attributes



# Diagram techniques

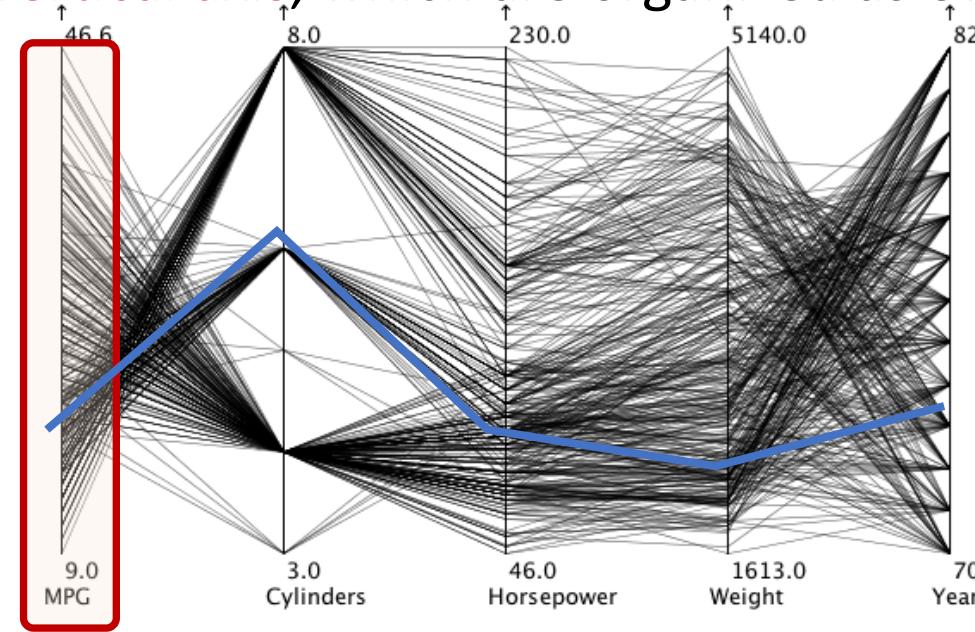
- Scatterplot matrix



	Age	Height	Shoe size
User 1	35	195	44
User 2	25	170	40
User 3	40	160	36
User 4	50	165	38

# Diagram techniques

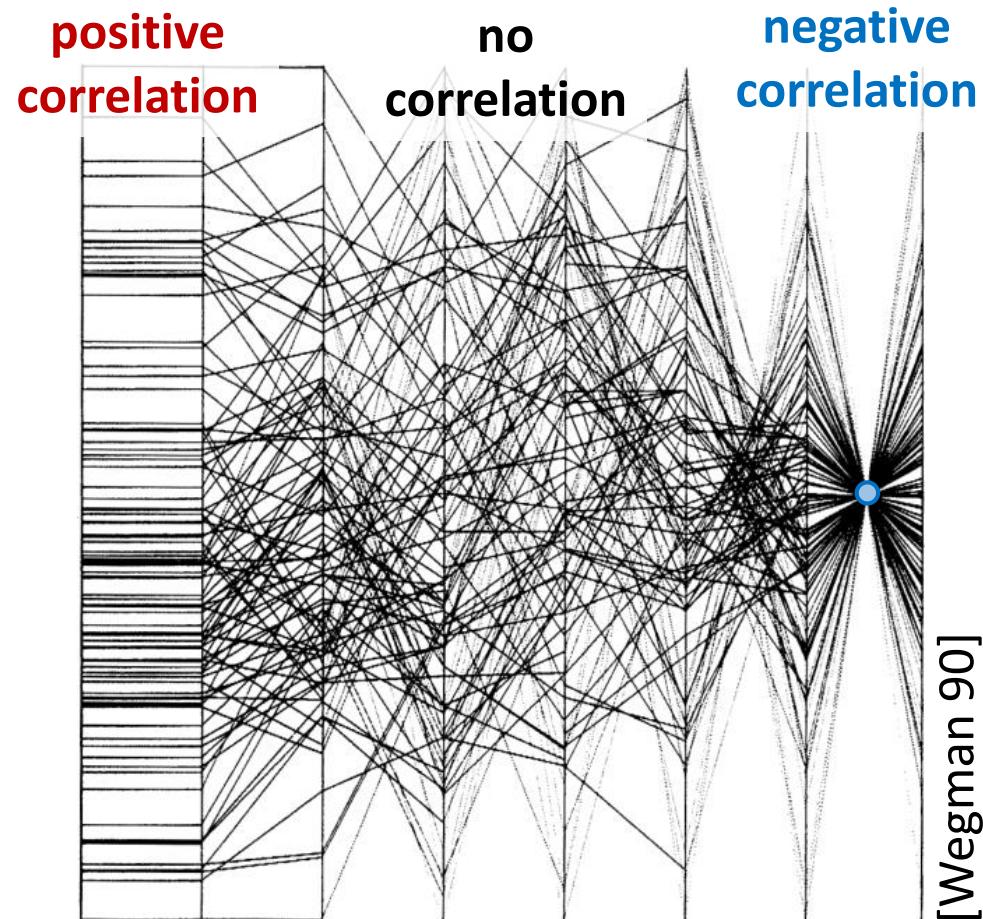
- Parallel coordinates
  - Represent multiple data variables
  - Each variable is represented by a **vertical axis**, which are organized as evenly spaced parallel lines
  - Data on each axis is normalized to min/max
  - One data sample is represented by a **connected set of points**, one on each axis



**Attribute / Dimension**

# Diagram techniques

- Parallel coordinates
  - Recognize patterns between adjacent axes
  - Steep learning curve for novices
  - Brushing (mark interesting data subset)



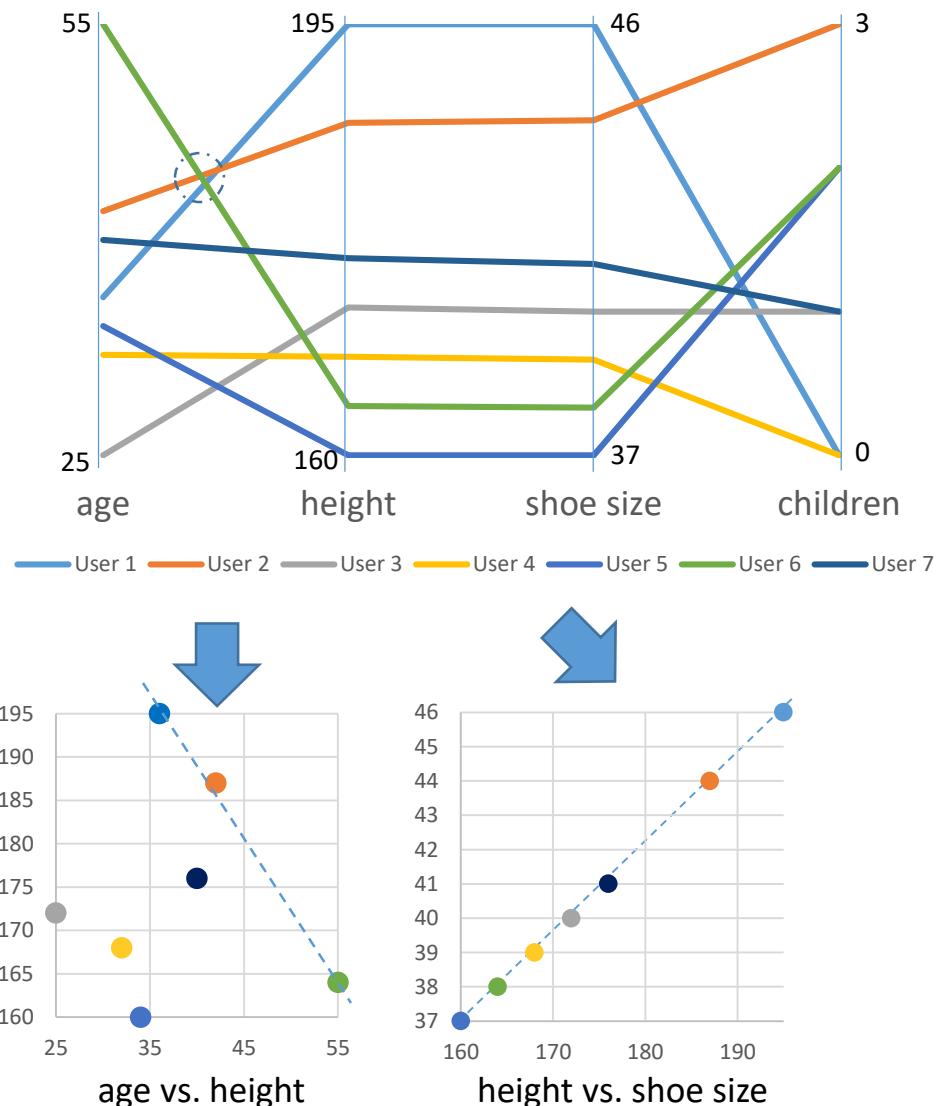
Parallel coordinates illustrating correlation  
of  $\rho = 1, 0.8, 0.2, 0, -0.2, -0.8, -1$

# Diagram techniques

- Parallel coordinates

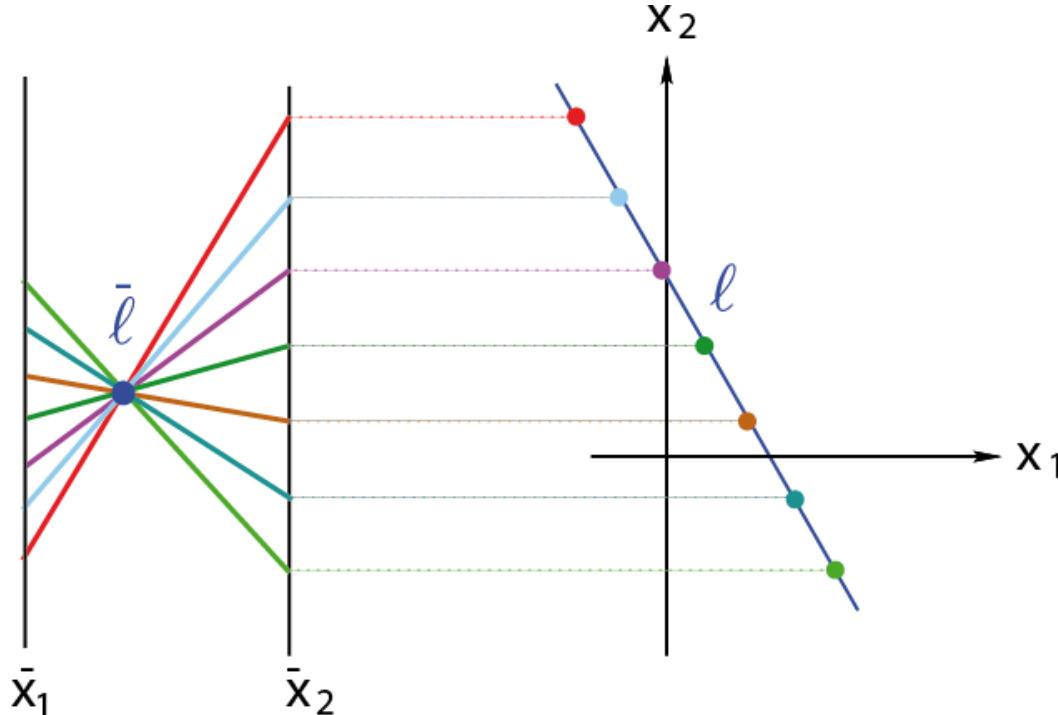
	Age	Height	Shoe size	No. of children
User 1	36	195	46	0
User 2	42	187	44	3
User 3	25	172	40	1
User 4	32	168	39	0
User 5	34	160	37	2
User 6	55	164	38	2
User 7	40	176	41	1

Normalize to min/max



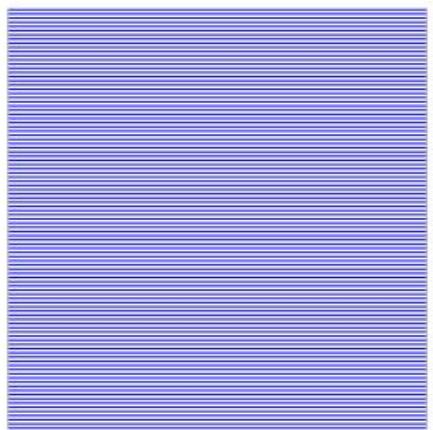
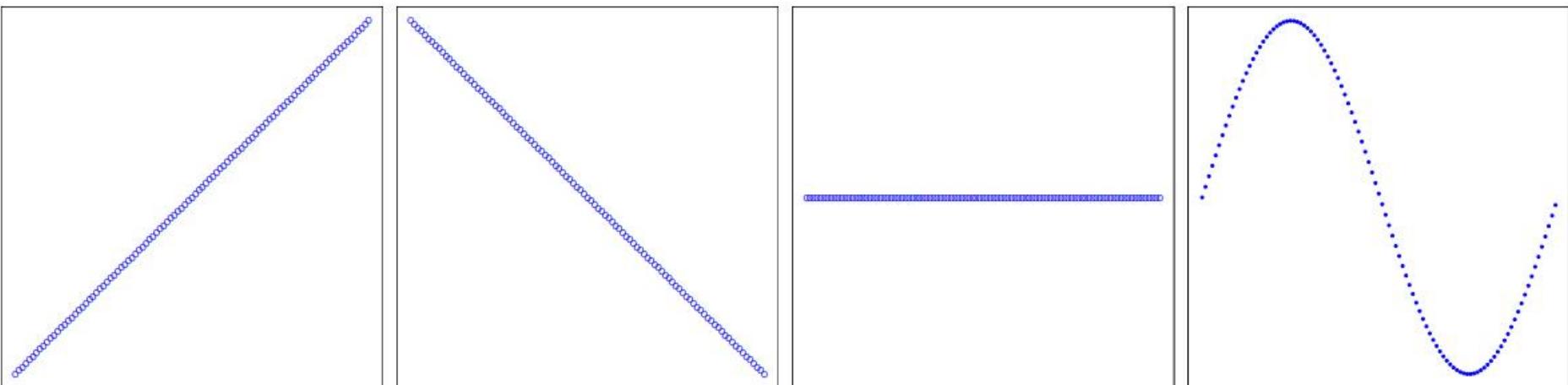
# Diagram techniques

- Parallel coordinates
  - Line point duality
    - Line in parallel coordinates maps to point in scatterplot
    - Line in scatterplot maps to point in parallel coordinates

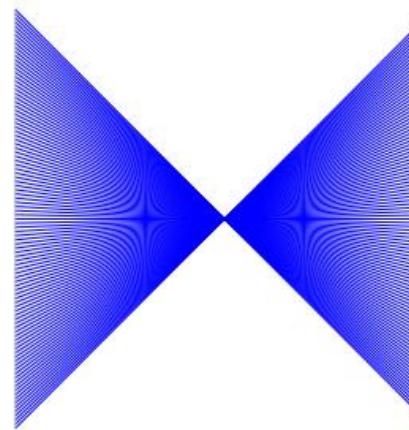


# Diagram techniques

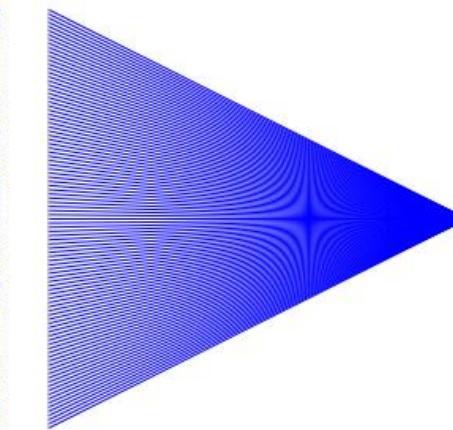
- Duality



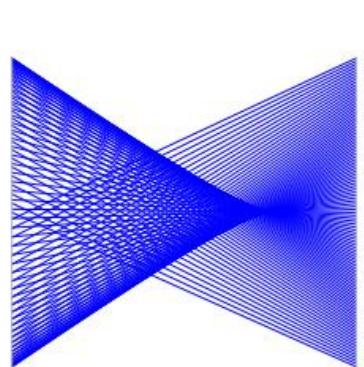
(a)  $y = x$



(b)  $y = -x$



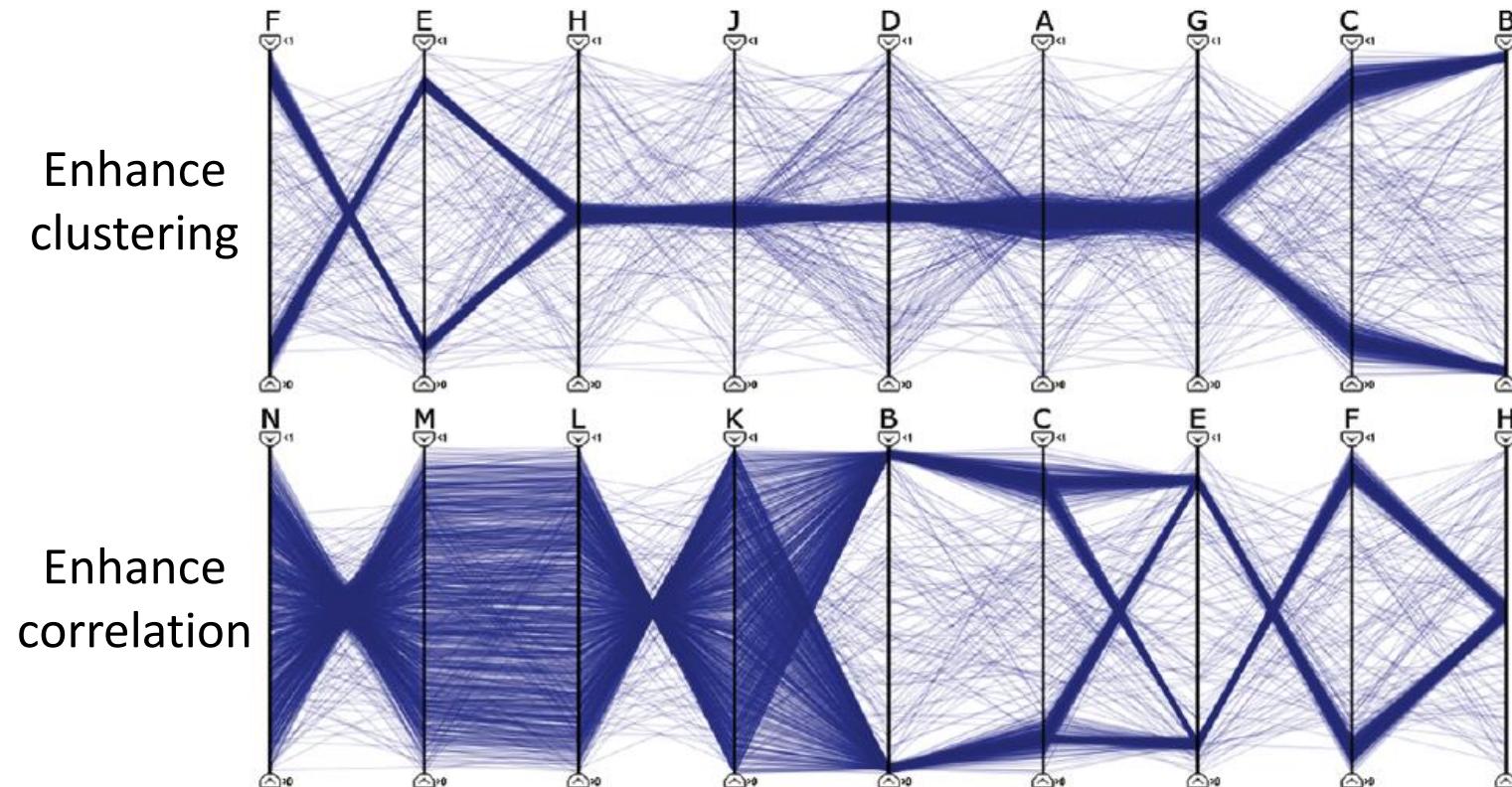
(c)  $y = 0$



(d)  $y = \sin(x)$

# Diagram techniques

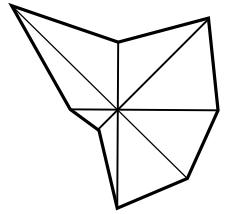
- Parallel coordinates
  - Axis ordering is a major challenge
  - Order by quality metrics



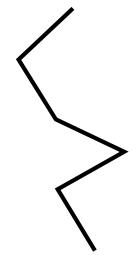
# Single and Multiple Variables – Glyphs

# Glyphs and icons

## 2D glyphs



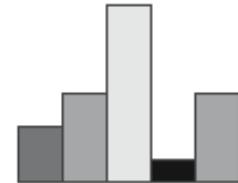
Star glyphs



Stick figures

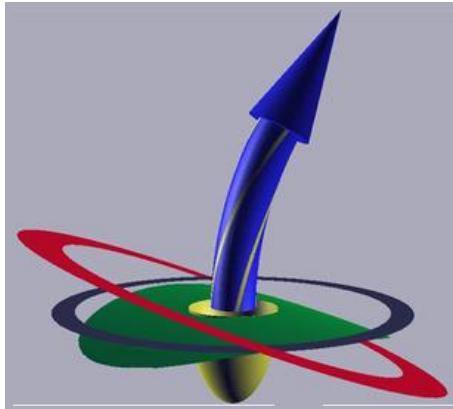


Chernoff  
faces

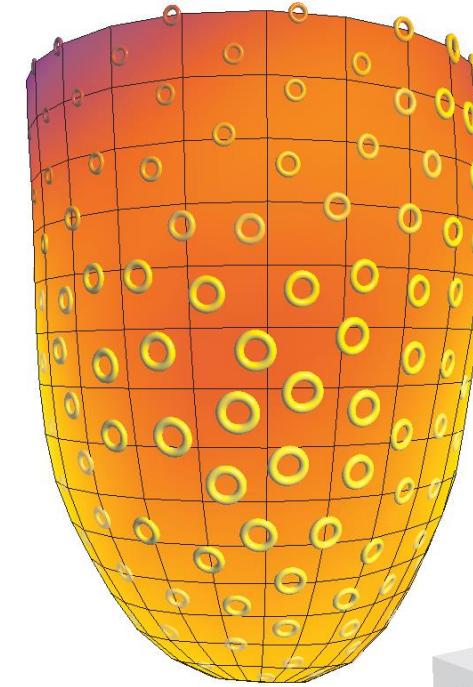


Profile glyphs

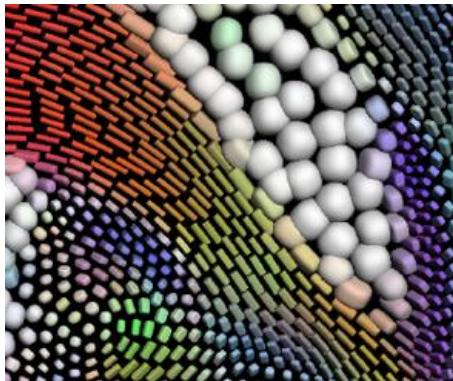
## 3D glyphs



## Surface glyphs



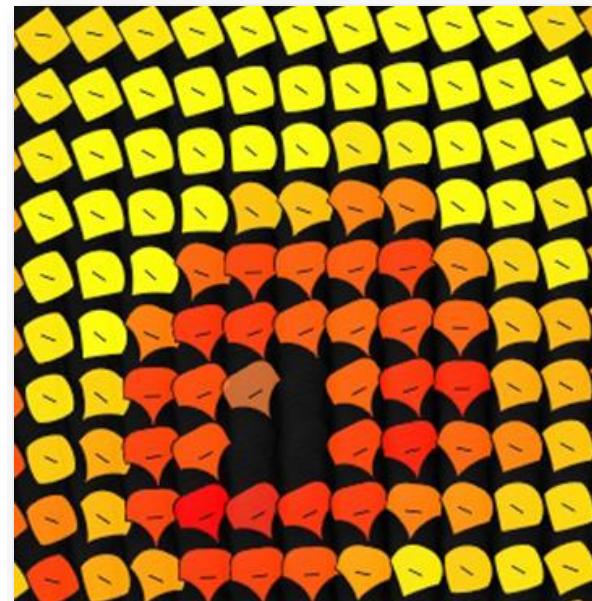
[Meyer-Spradow et al. 08]



[de Leeuw&van Wijk 93],  
[Kindlmann&Westin 06]

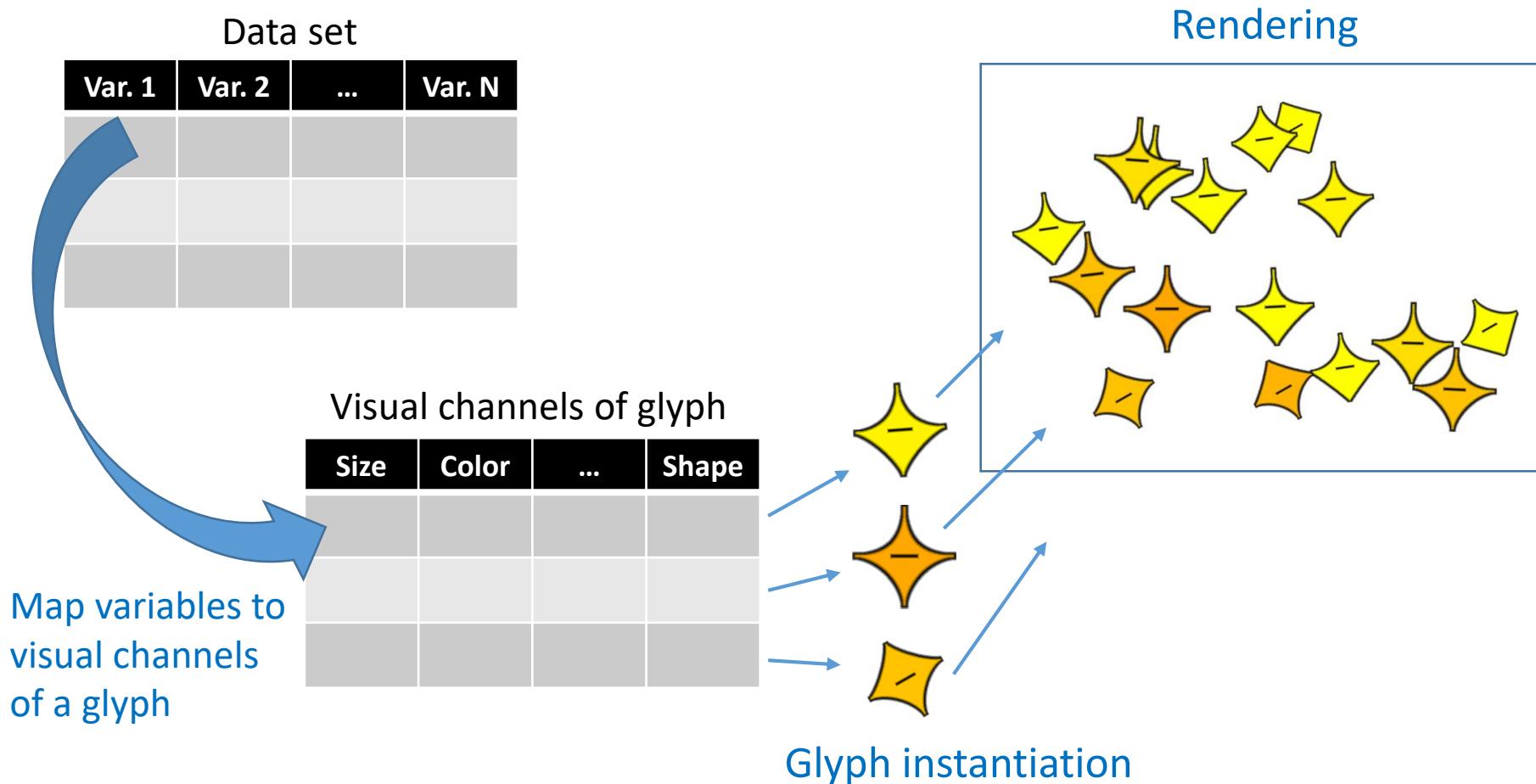
# Glyphs and icons

- Glyphs: Small independent visual objects that depict multiple attributes of a data record
  - Discretely placed in a display space
  - Data attributes are represented by/mapped to different visual channels of glyph (e.g., shape, color, size, orientation)
  - Visual channels should be easy to distinguish and combine
  - Mainly used for multivariate data



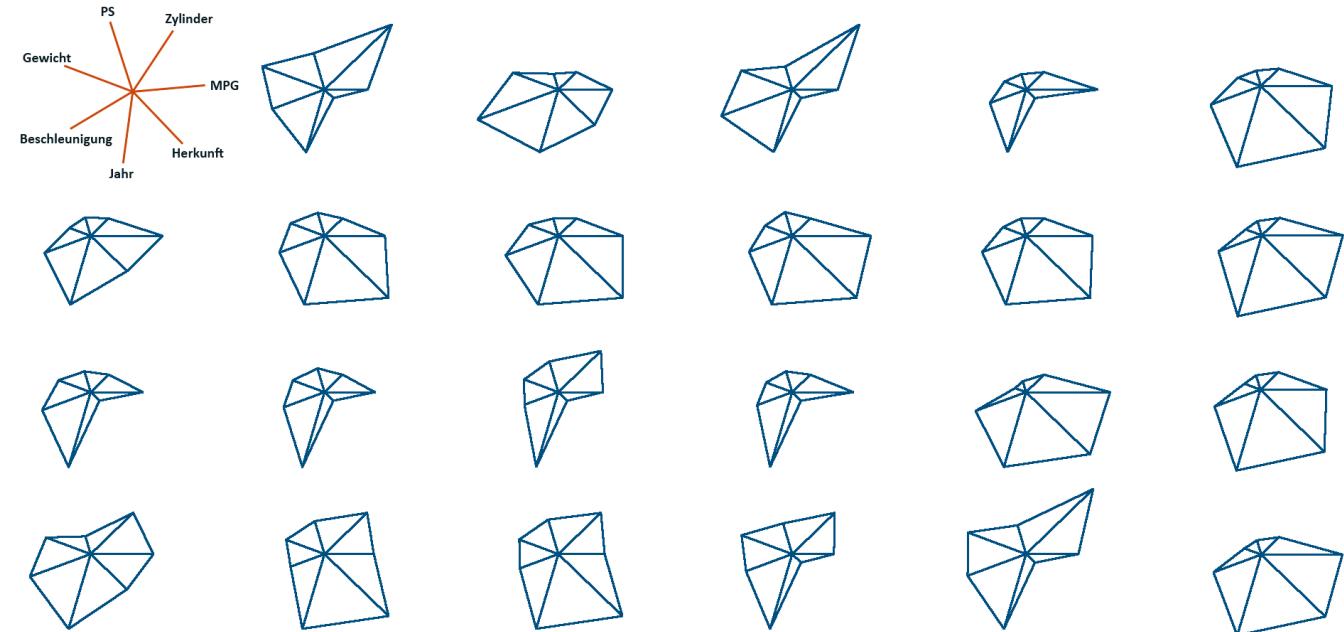
# Glyphs and icons

- Glyph creation



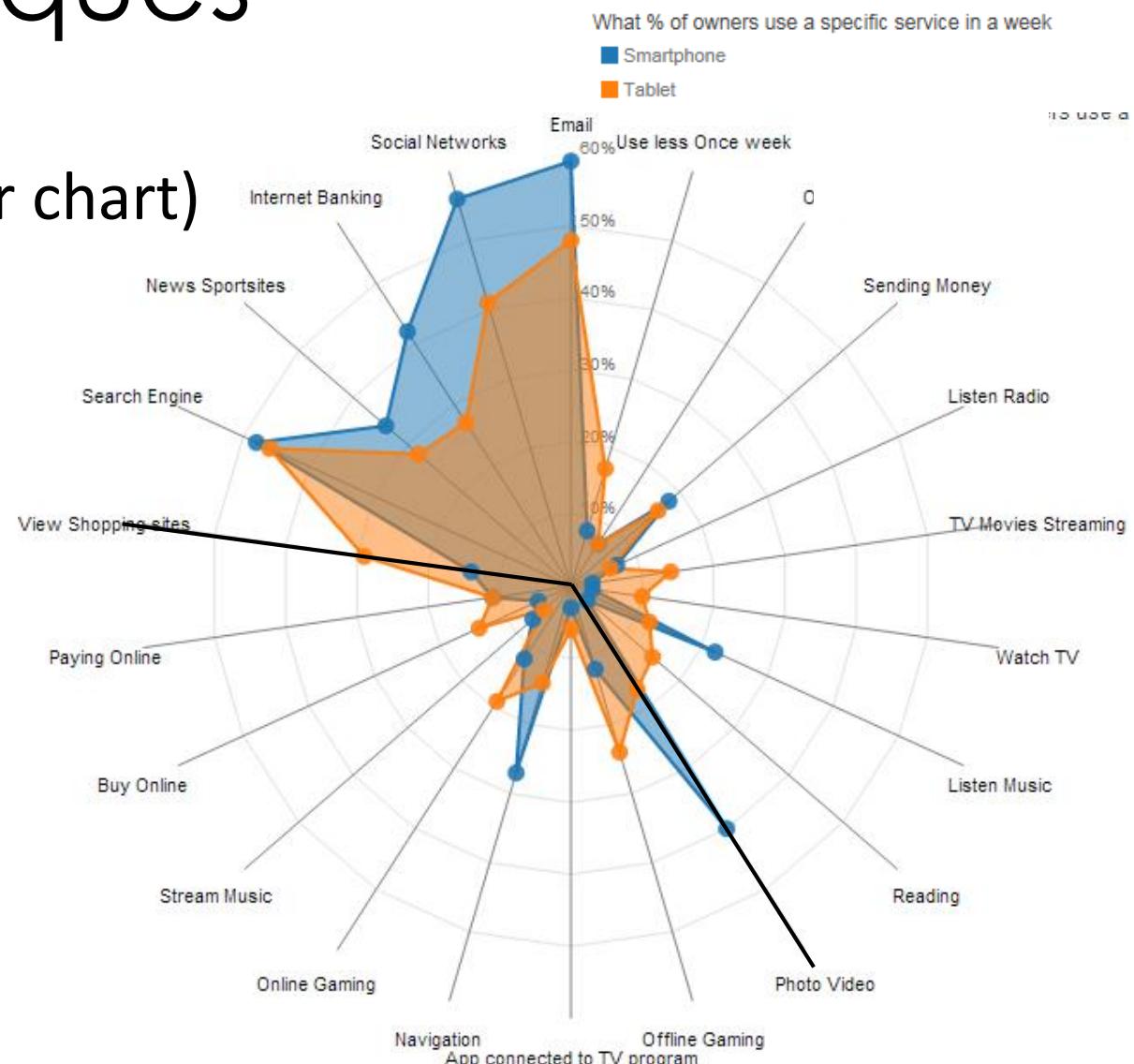
# Glyphs and icons

- Example: Star glyphs
  - Star is composed of equally spaced spikes, originating from center
  - Length of spikes represents value of respective attribute
  - End of rays connected by line



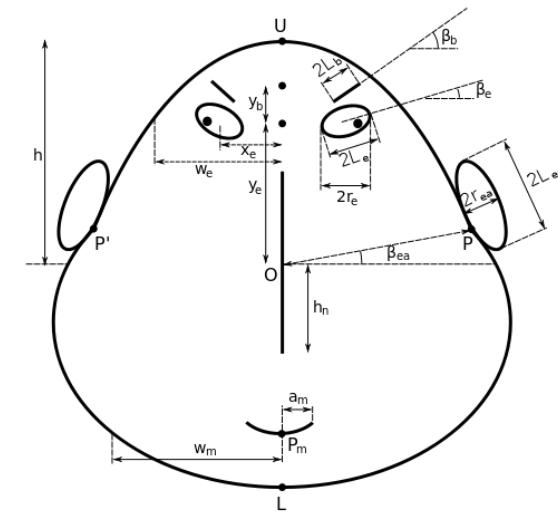
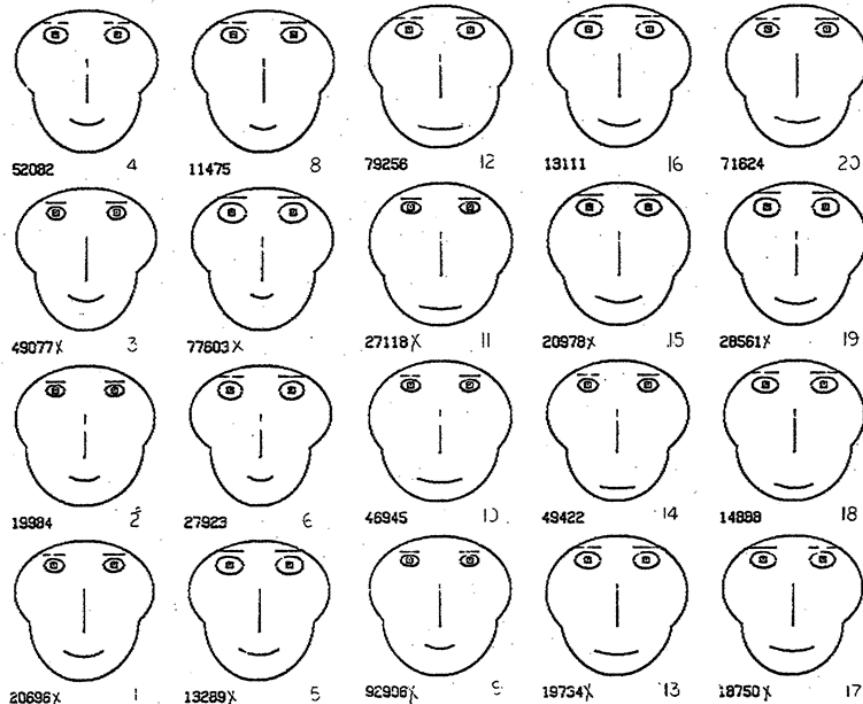
# Diagram techniques

- Radar chart (star plot, spider chart)
  - Radial axes arrangement
  - Items are polylines



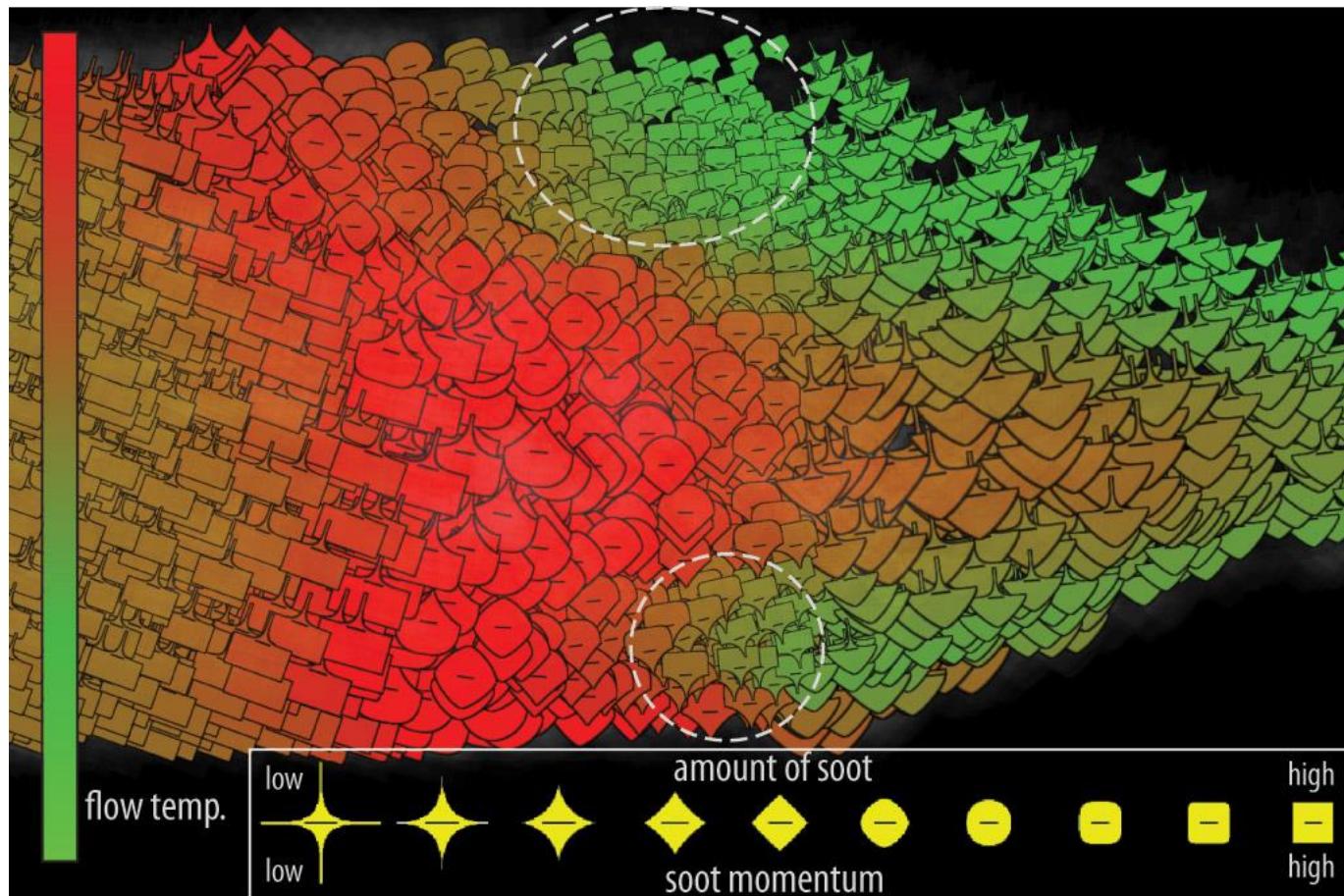
# Glyphs and icons

- Example: Chernoff faces
  - Data attributes represented by features of a face  
(eye position, nose length, mouth form, etc.)



- Faces are perceived holistically  
- Efficiency?

# Diesel Particulate Filter



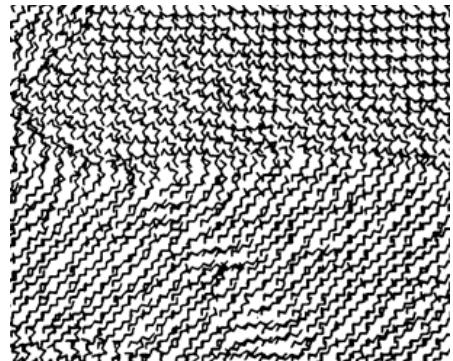
Size & color: flow temp.

Glyph rotation (-45°, 45°): O<sub>2</sub> fraction

# Glyphs and icons

- Glyph complexity vs. placement density

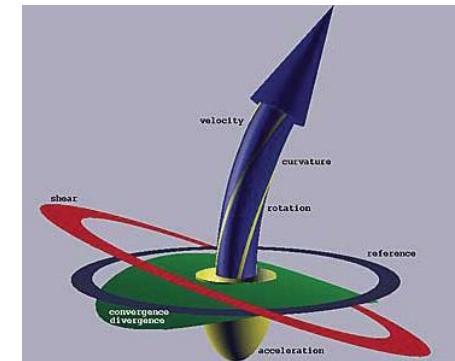
Dense placement  
& simple glyphs



Stick figures

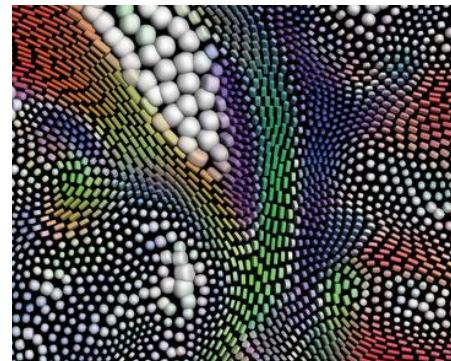
[Pickett&Grinstein 88]

Sparse placement  
& complex glyphs



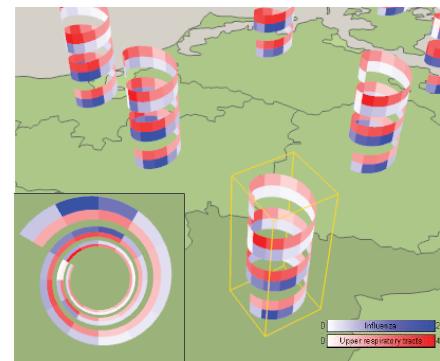
Local flow probe

[de Leeuw&van Wijk 93]



Glyph packing

[Kindlmann&Westin 06]

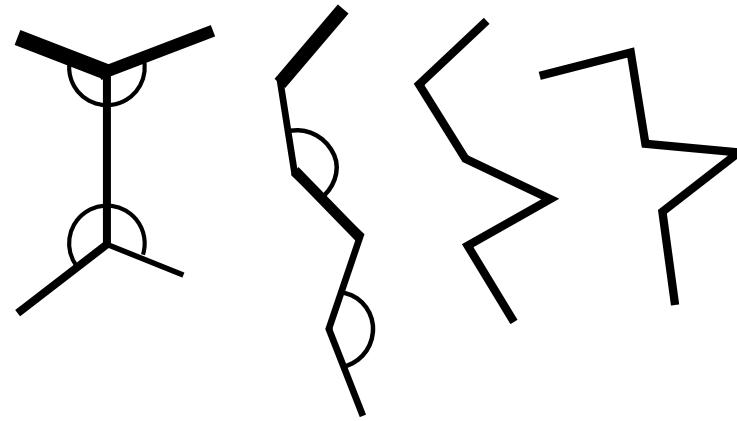


Helix glyphs

[Tominski et al. 05]

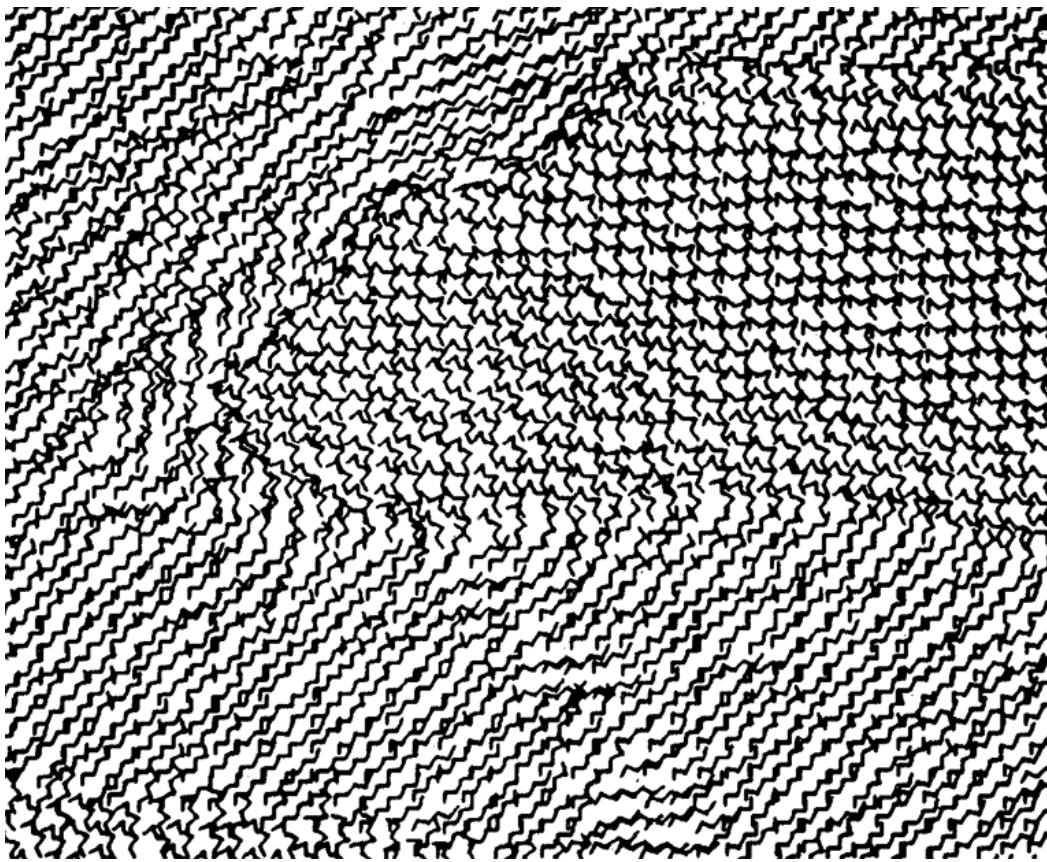
# Glyphs and icons

- Stick figures
  - 2D figure with limbs
  - Data encoded by
    - length
    - line thickness
    - angle between lines



# Glyphs and icons

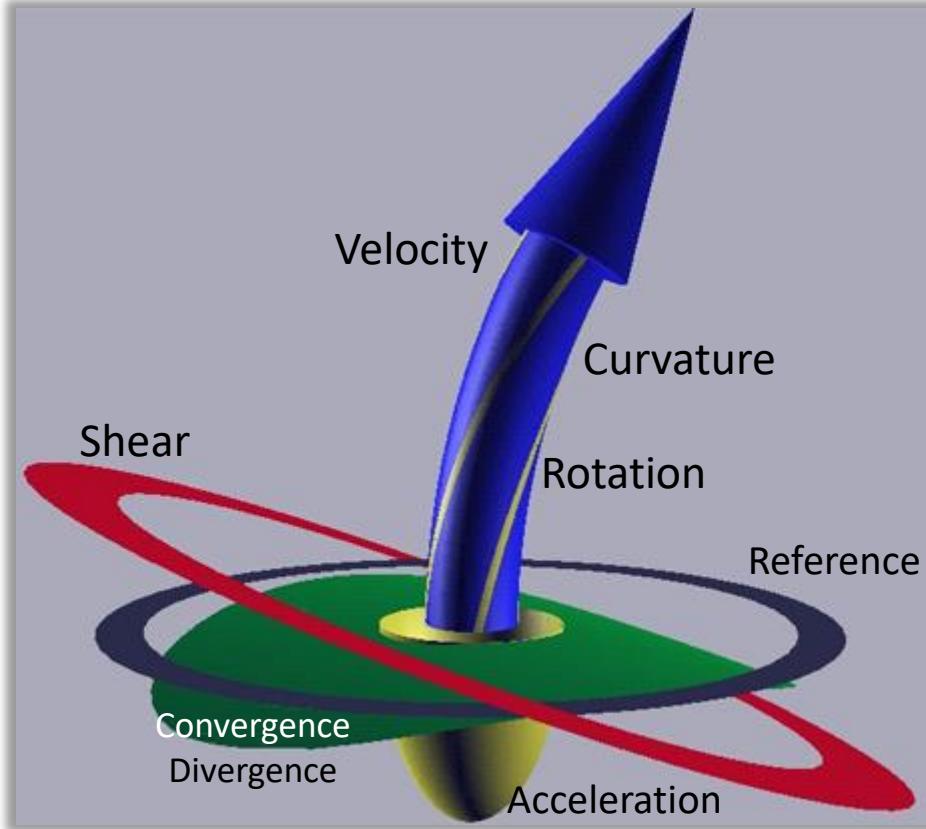
- Stick figures
  - 2D figure with limbs
  - Data encoded by
    - length
    - line thickness
    - angle between lines
- Texture patterns show certain data characteristics



5 channels from multispectral satellite image,  
Lake Ontario, US

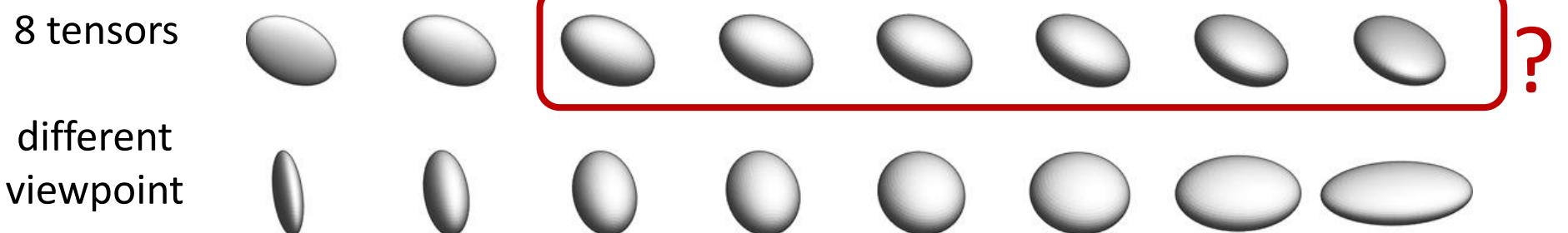
# Glyphs and icons

- Local flow probe
  - Depicts multiple flow characteristics
  - Large & complex glyphs need to be placed sparsely to avoid occlusion

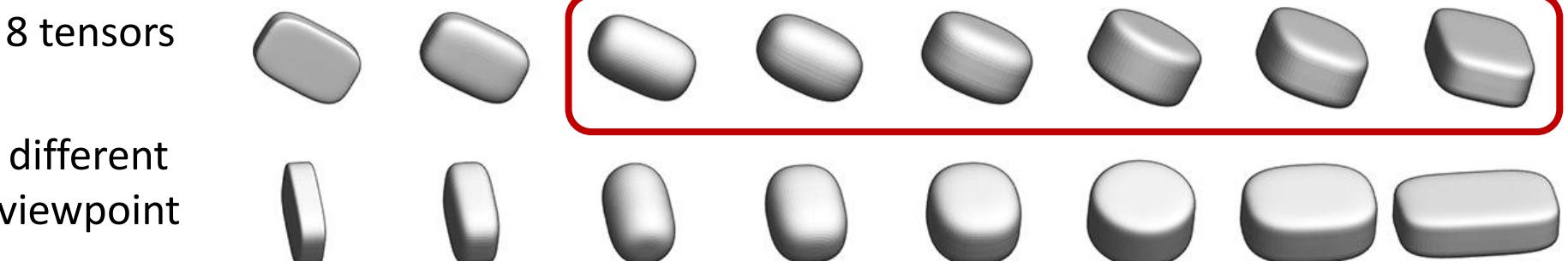


# Glyphs and icons

- View-point independence
  - Ellipsoid glyphs

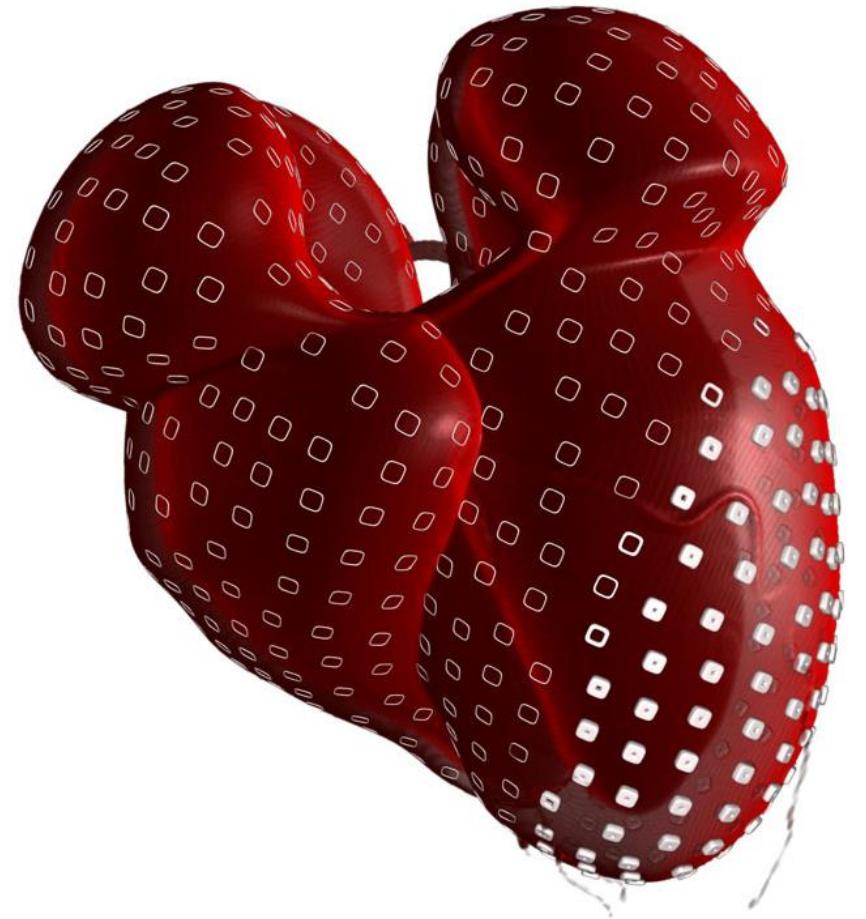


- Superquadric glyphs



# Glyphs and icons

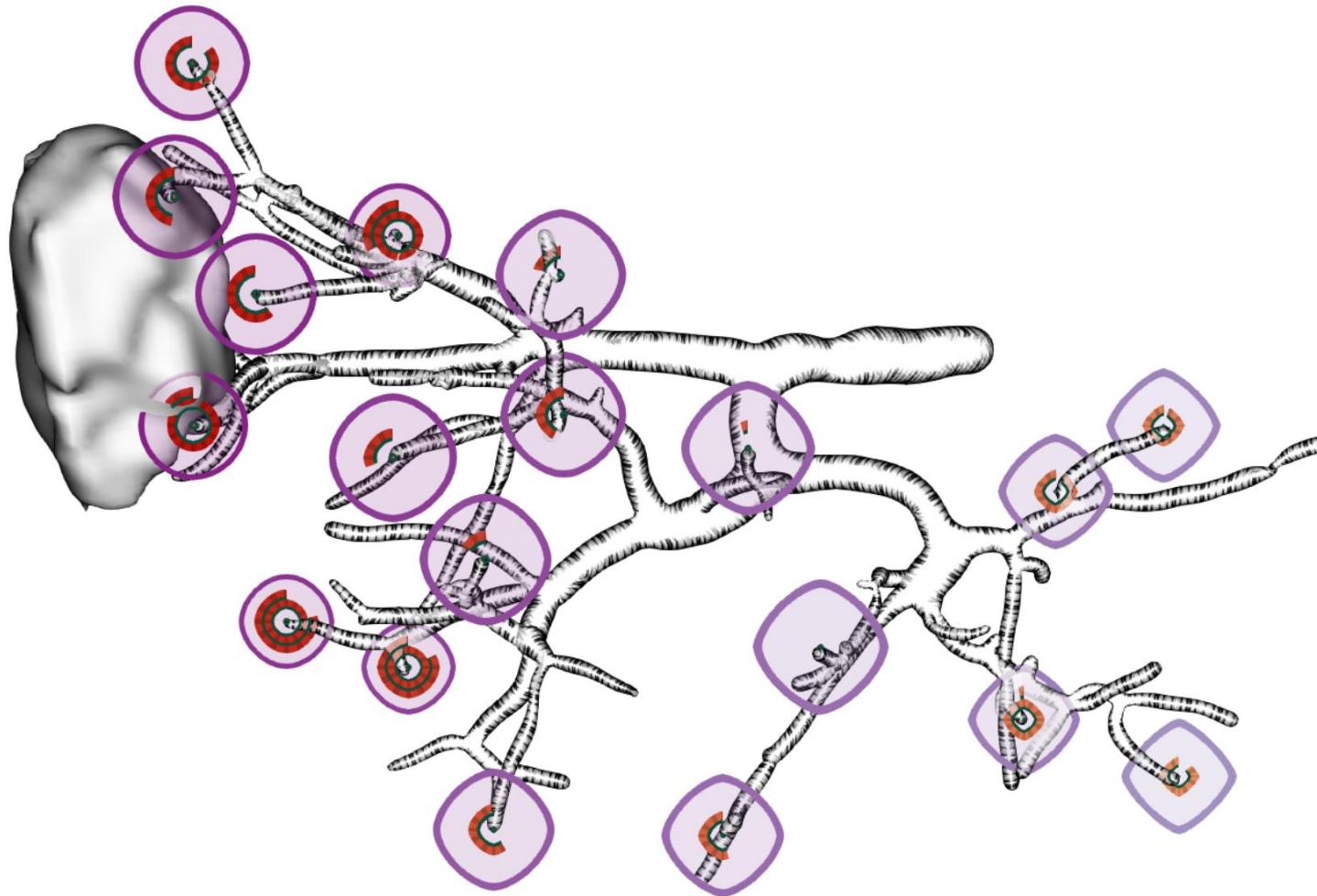
- Importance-based Mapping
  - Emphasize important attributes
  - Guide the user's attention (e.g., color, size)



PET activity → thinness



# Glyphs and icons



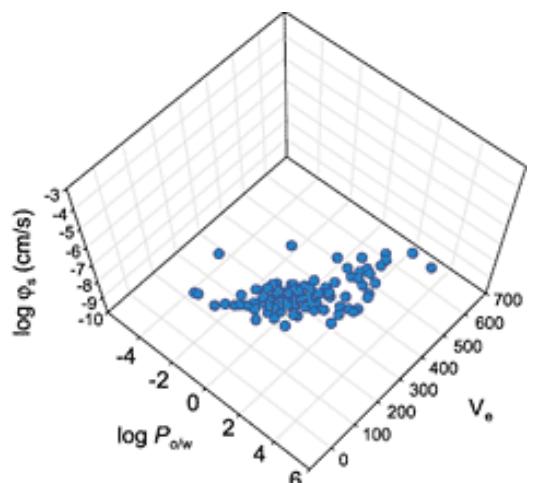
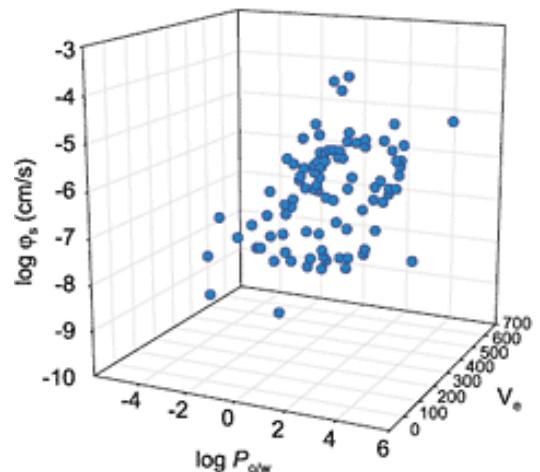
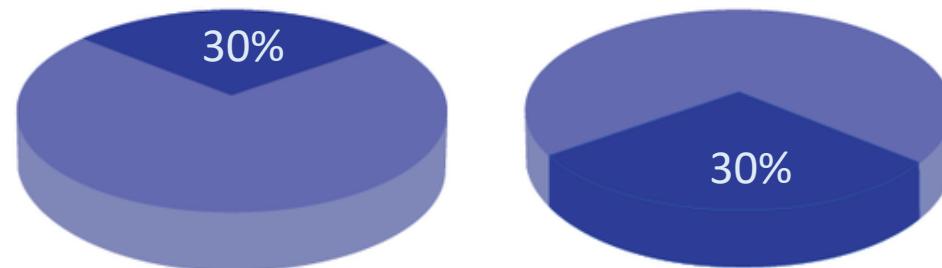
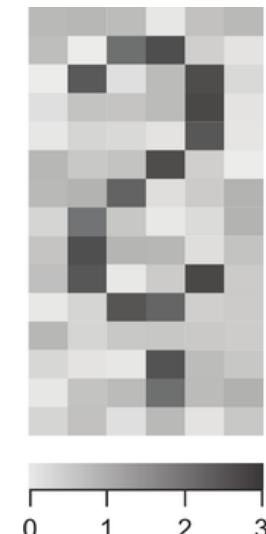
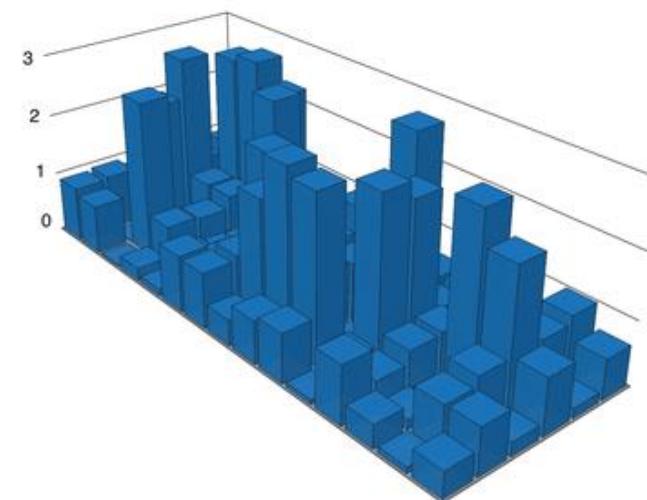
# Glyphs and icons

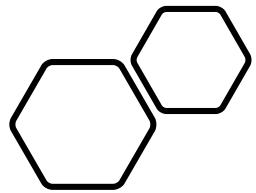
- Summary
  - Just combining visual channels is not enough
  - Design considerations (e.g., separability, semantics, density, complexity, etc.)
  - Local & compact representation of many data attributes
  - Glyph design restricted by perceptual limits

# Diagram techniques

- 3D Pitfalls: Occlusion and Perspective

Which one is the tallest bar?  
What is the pattern in the data?





# Questions???