Data Visualization (CSCI 627/490)

Tabular Data

Dr. David Koop

Visual Encoding

- How do we encode data visually?
  - **Marks** are the basic graphical elements in a visualization
  - **Channels** are ways to control the appearance of the marks
- Marks classified by dimensionality:

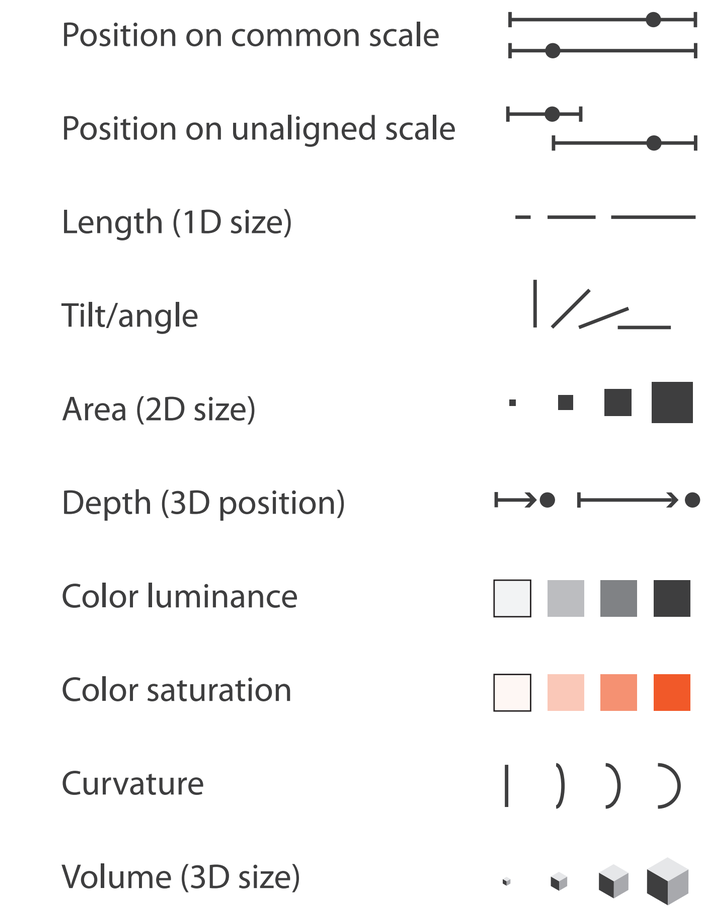
Points **Lines** **Areas**

- Also can have surfaces, volumes
- Think of marks as a mathematical definition, or if familiar with tools like Adobe Illustrator or Inkscape, the path & point definitions

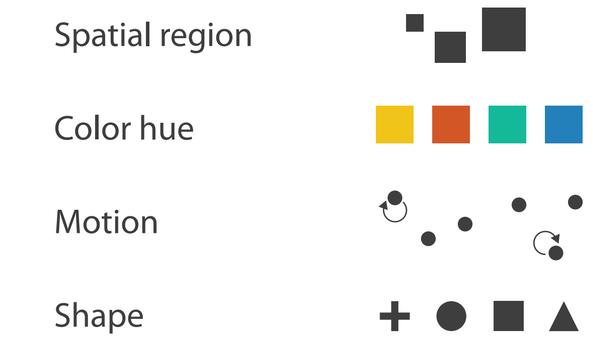
Channel Types

- Identity => what or where, Magnitude => how much

Magnitude Channels: Ordered Attributes

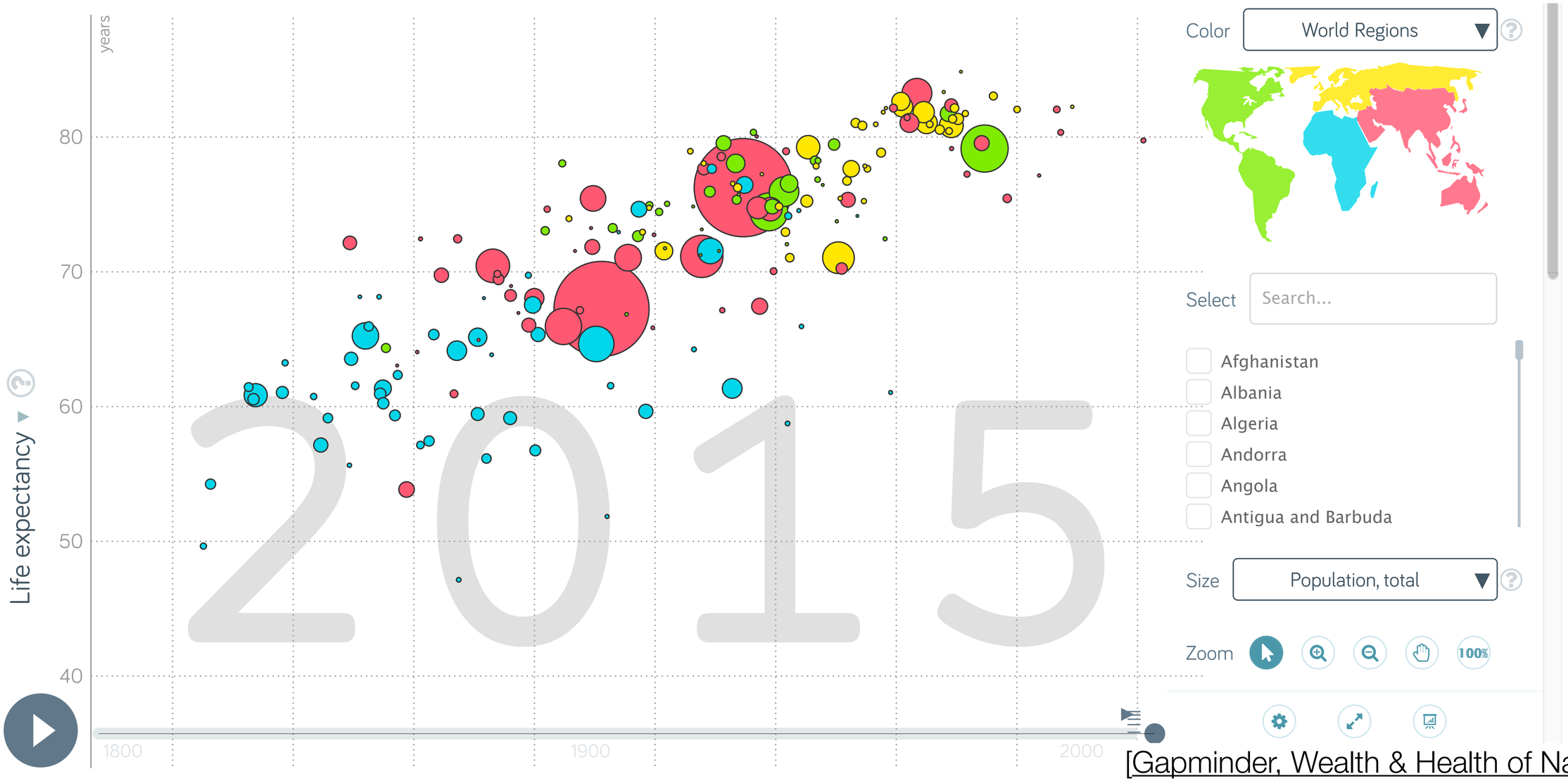


Identity Channels: Categorical Attributes



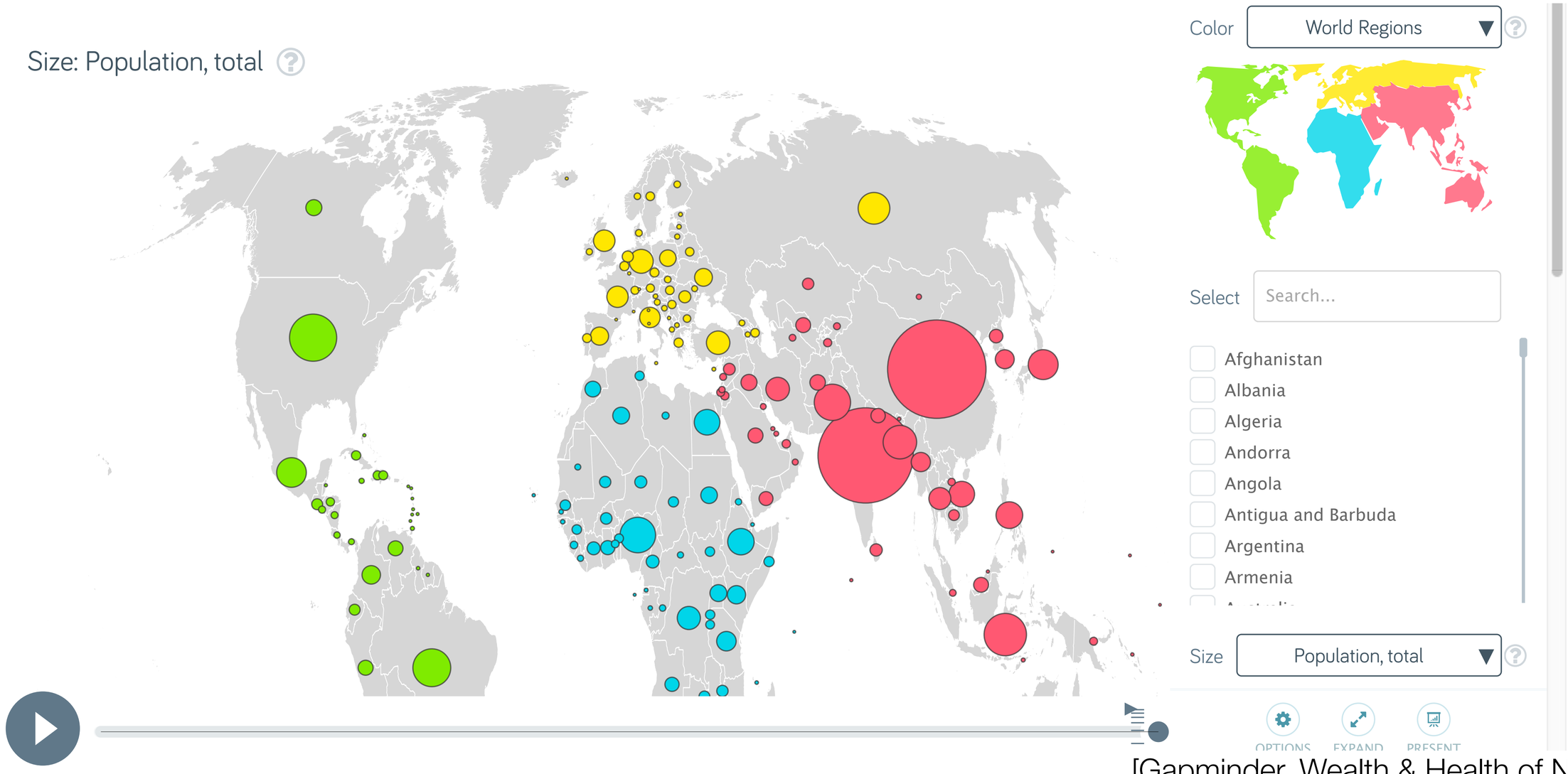
[Munzner (ill. Maguire), 2014]

Visual Encoding



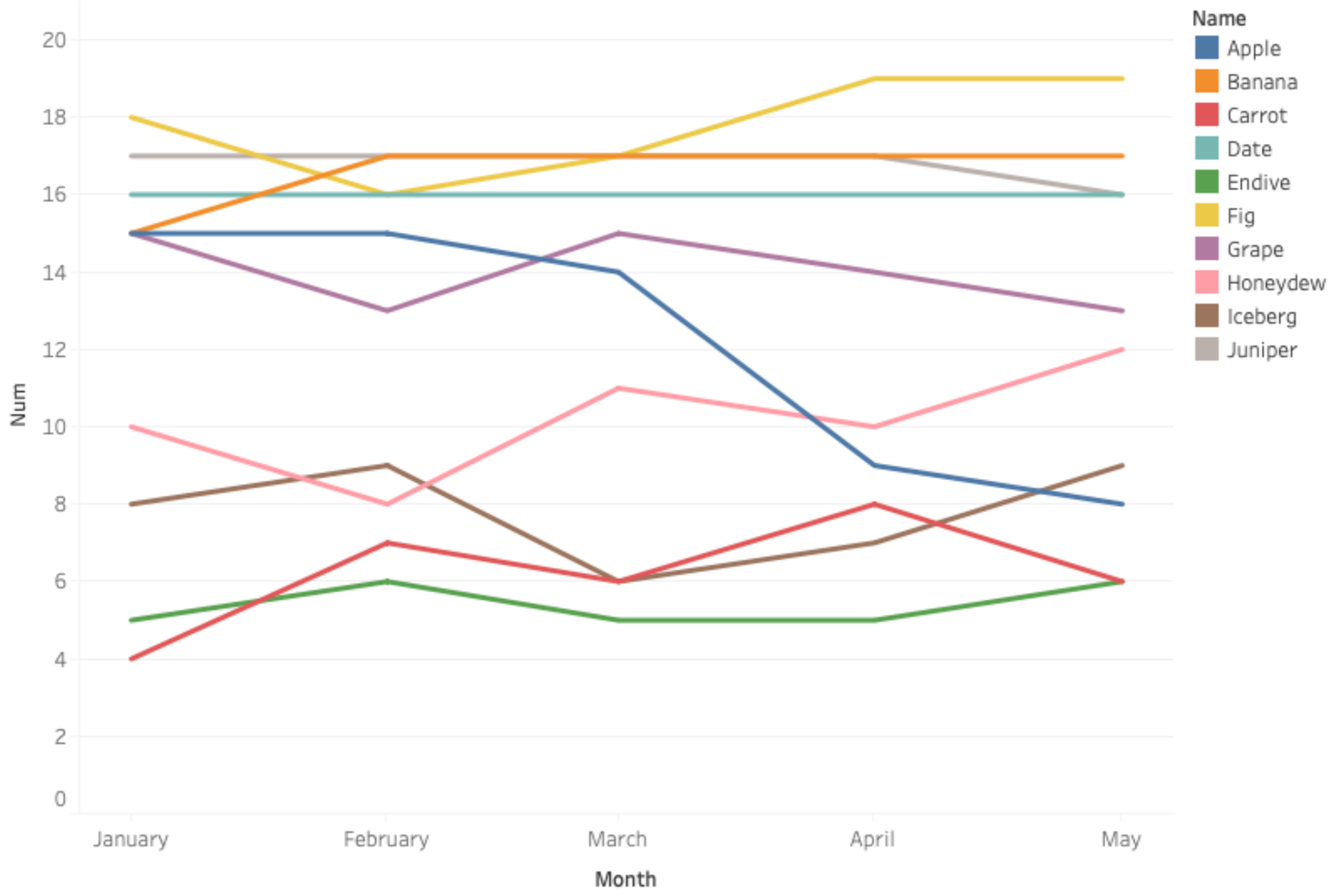
[Gapminder, Wealth & Health of Nations]

Another Encoding

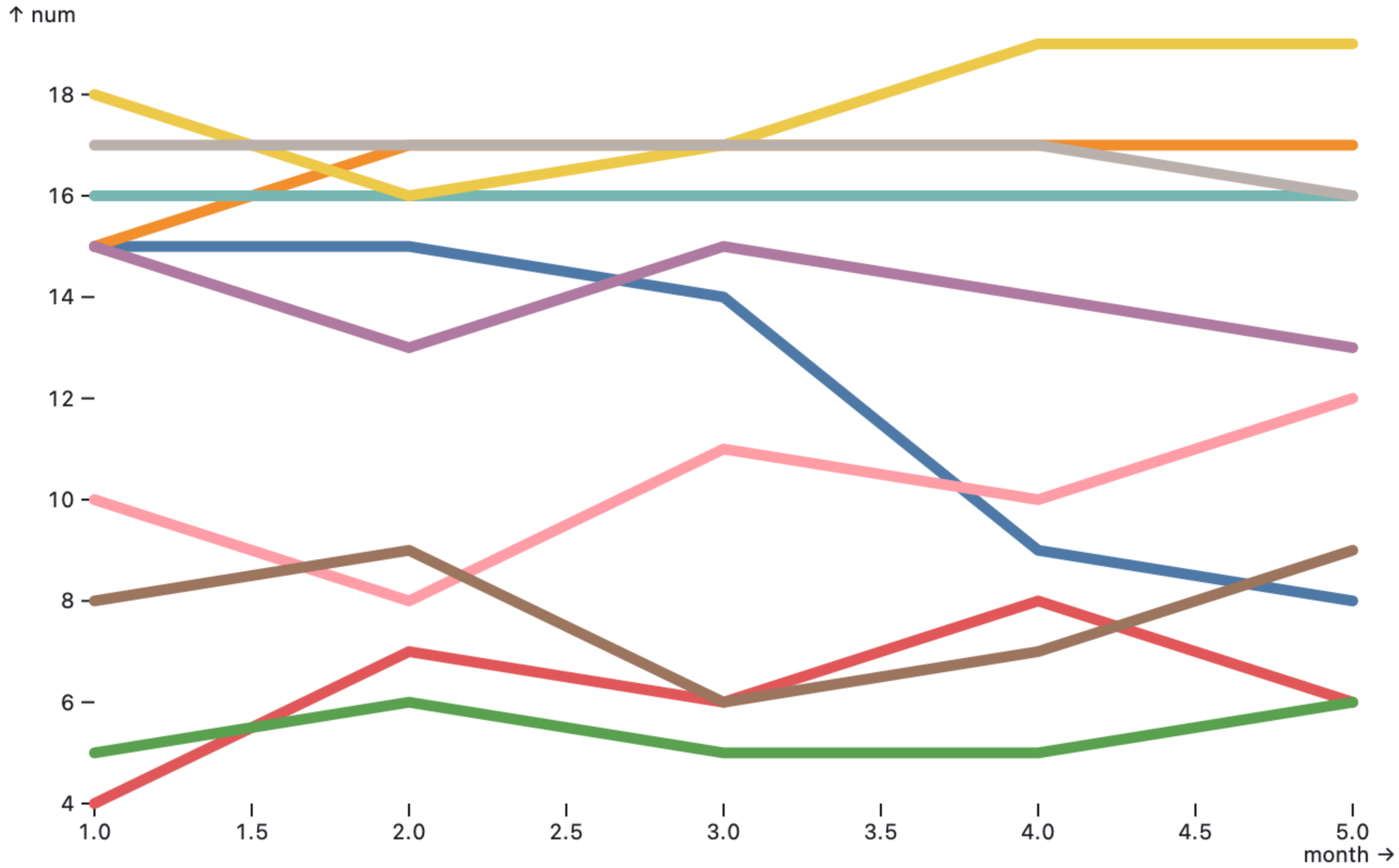


[Gapminder, Wealth & Health of Nations]

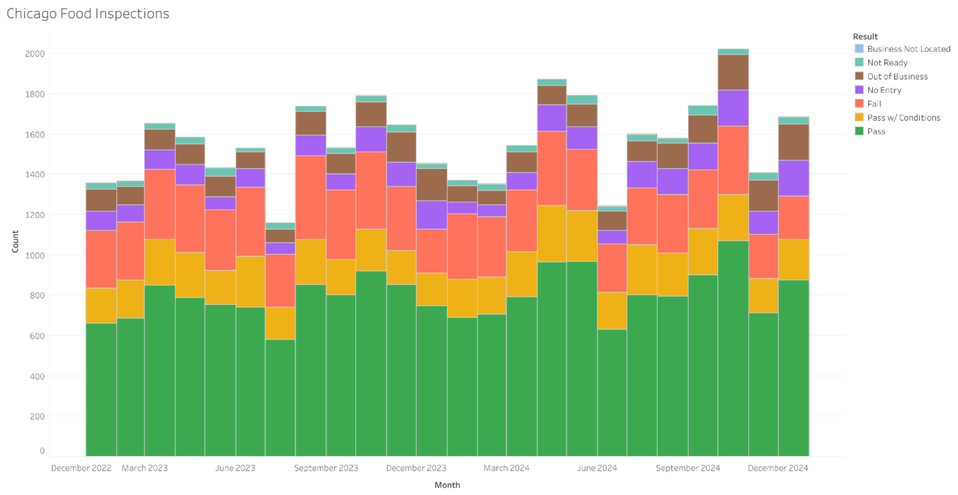
Tableau Example



Observable Plot Example



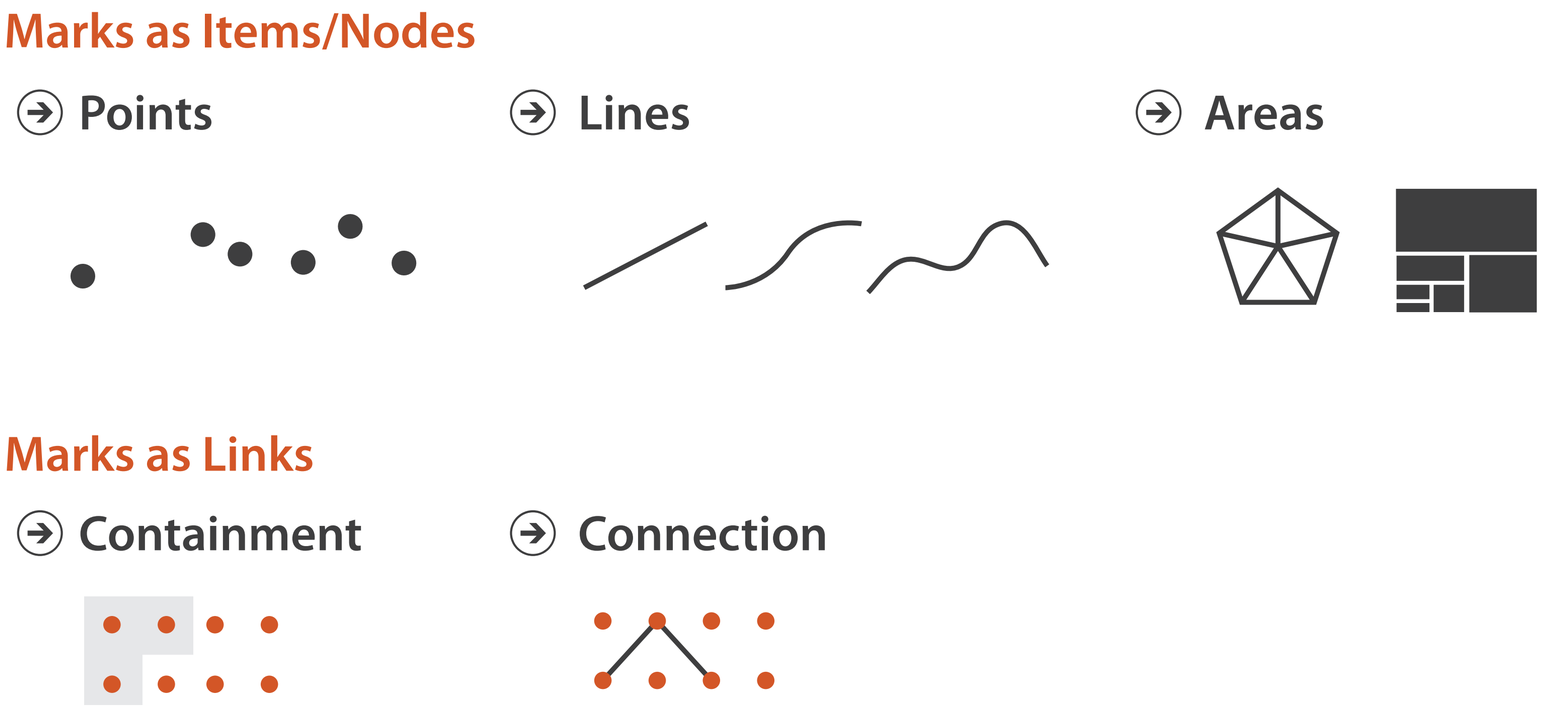
Assignment 3



- Food Inspections Data
- Create the same stacked bar chart using
  - Tableau Public
  - Observable Plot
  - D3
- D3 Stacked Bar Chart:
  - Required for CSCI 627 students
  - CSCI 490 students need not stack

Mark Types

- Can have marks for items and **links**
  - Connection => pairwise relationship
  - Containment => hierarchical relationship

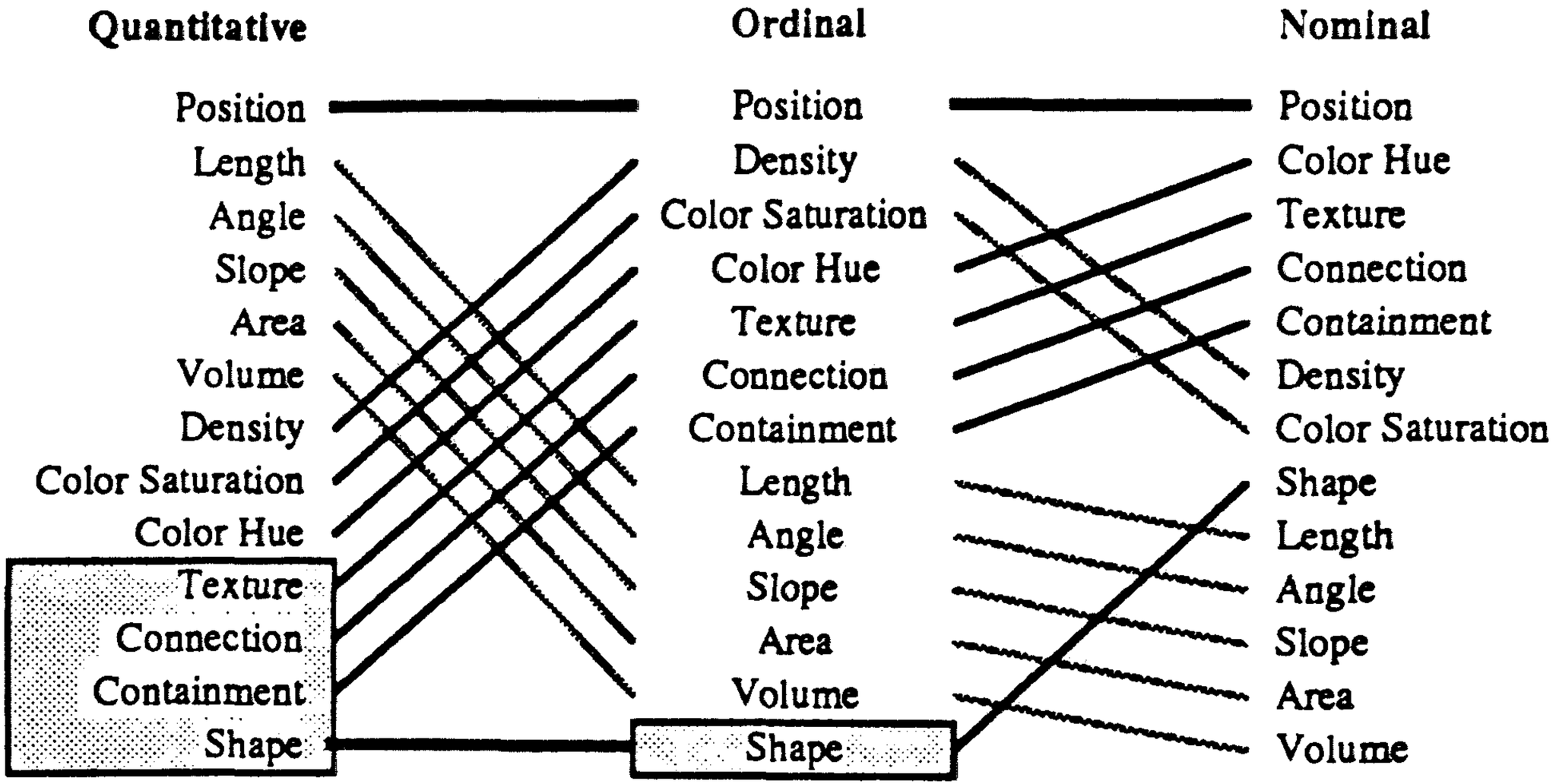


[Munzner (ill. Maguire), 2014]

Expressiveness and Effectiveness

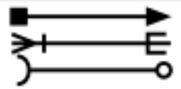
- Expressiveness Principle: all data from the dataset and nothing more should be shown
  - Do encode ordered data in an ordered fashion
  - Don’t encode categorical data in a way that implies an ordering
- Effectiveness Principle: the most important attributes should be the most **salient**
  - Saliency: how noticeable something is
  - How do the channels we have discussed measure up?

Mackinlay's Ranking of Perceptual Tasks



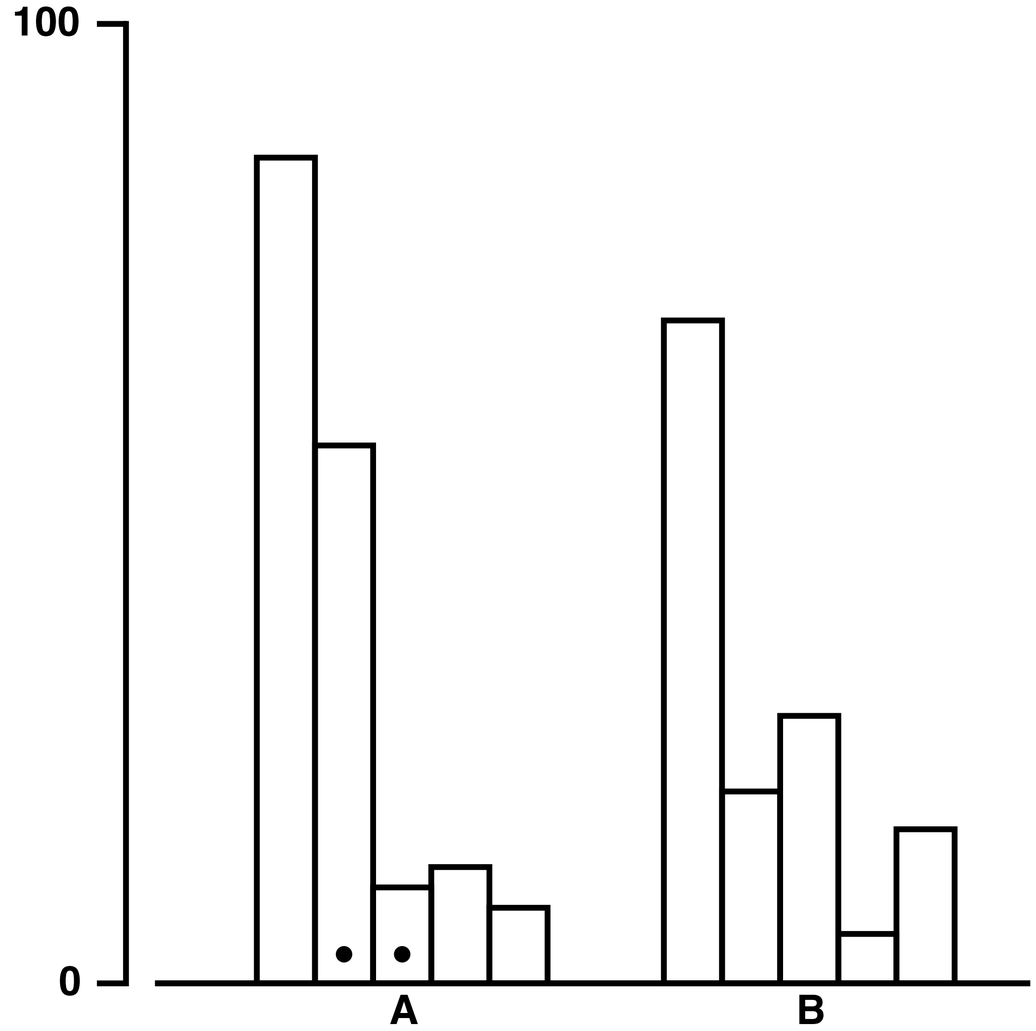
[Mackinlay,1986]

Iliinsky's Best Uses, +Ordering, +NumValues

Example	Encoding	Ordered	Useful values	Quantitative	Ordinal	Categorical	Relational
	position, placement	**yes**	**infinite**	Good	Good	Good	Good
1, 2, 3; A, B, C	text labels	**optional** (alphabetical or numbered)	**infinite**	Good	Good	Good	Good
	length	**yes**	**many**	Good	Good		
	size, area	**yes**	**many**	Good	Good		
	angle	**yes**	medium/few	Good	Good		
	pattern density	**yes**	few	Good	Good		
	weight, boldness	**yes**	few		Good		
	saturation, brightness	**yes**	few		Good		
	color	no	few (< 20)			Good	
	shape, icon	no	medium			Good	
	pattern texture	no	medium			Good	
	enclosure, connection	no	**infinite**			Good	Good
	line pattern	no	few				Good
	line endings	no	few				Good
	line weight	**yes**	few		Good		

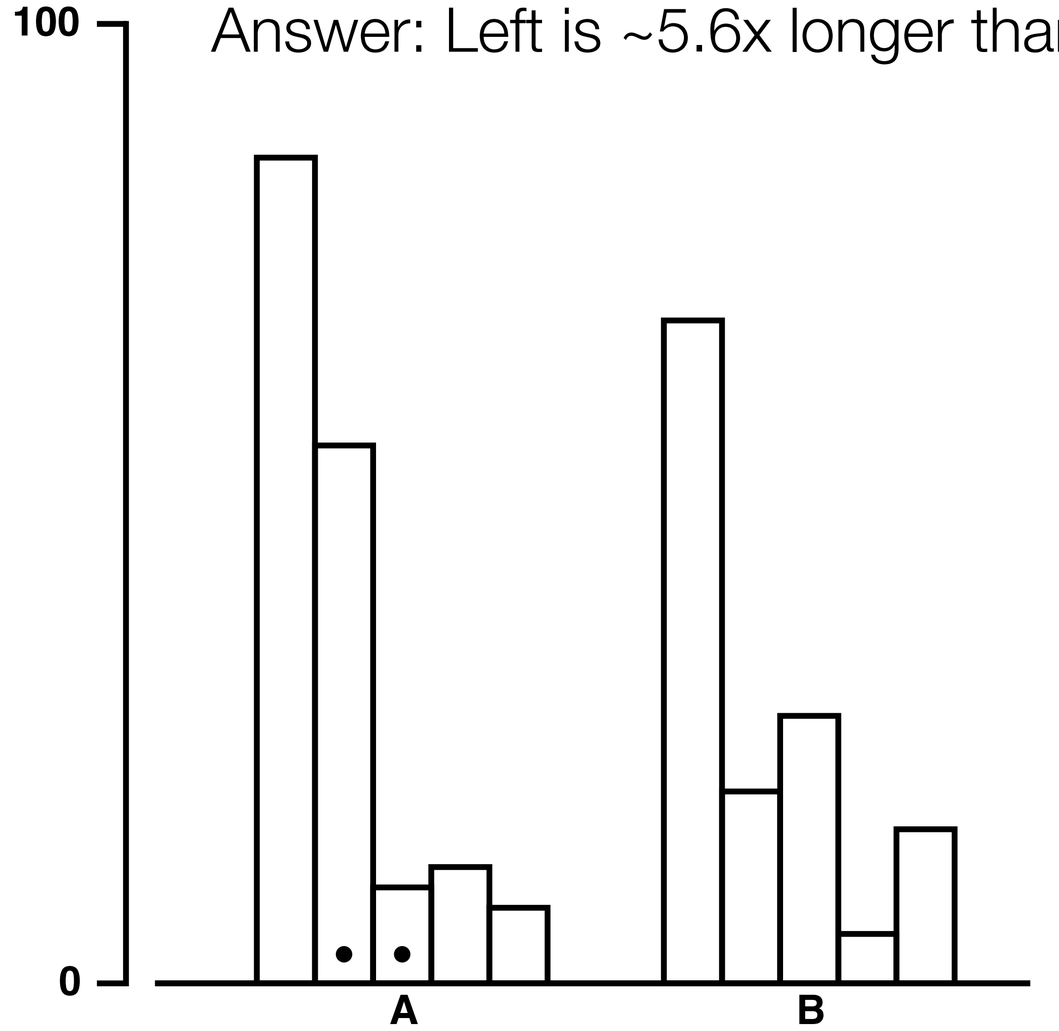
How do we get these rankings?

Test % difference in **length** between elements



[Heer & Bostock, 2010]

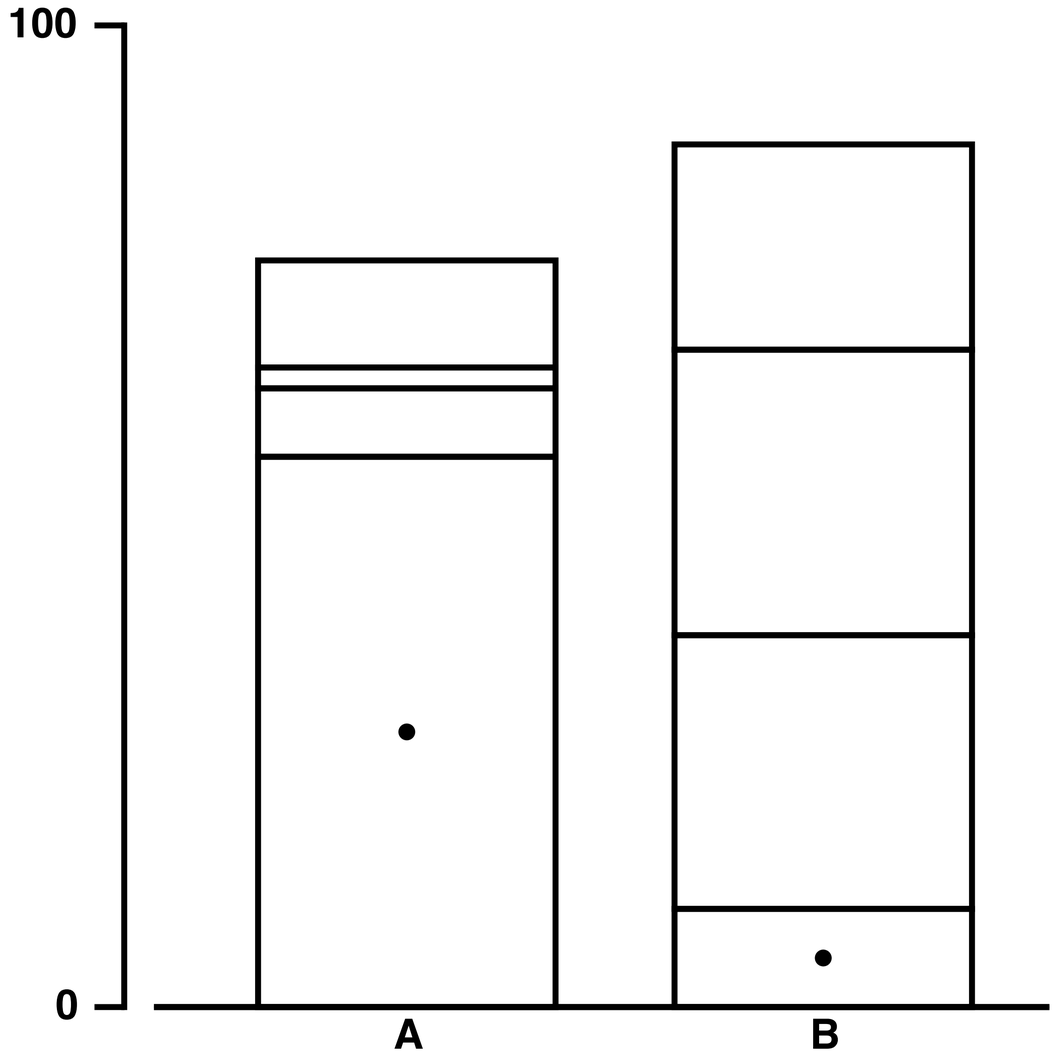
Test % difference in **length** between elements



Answer: Left is ~5.6x longer than Right

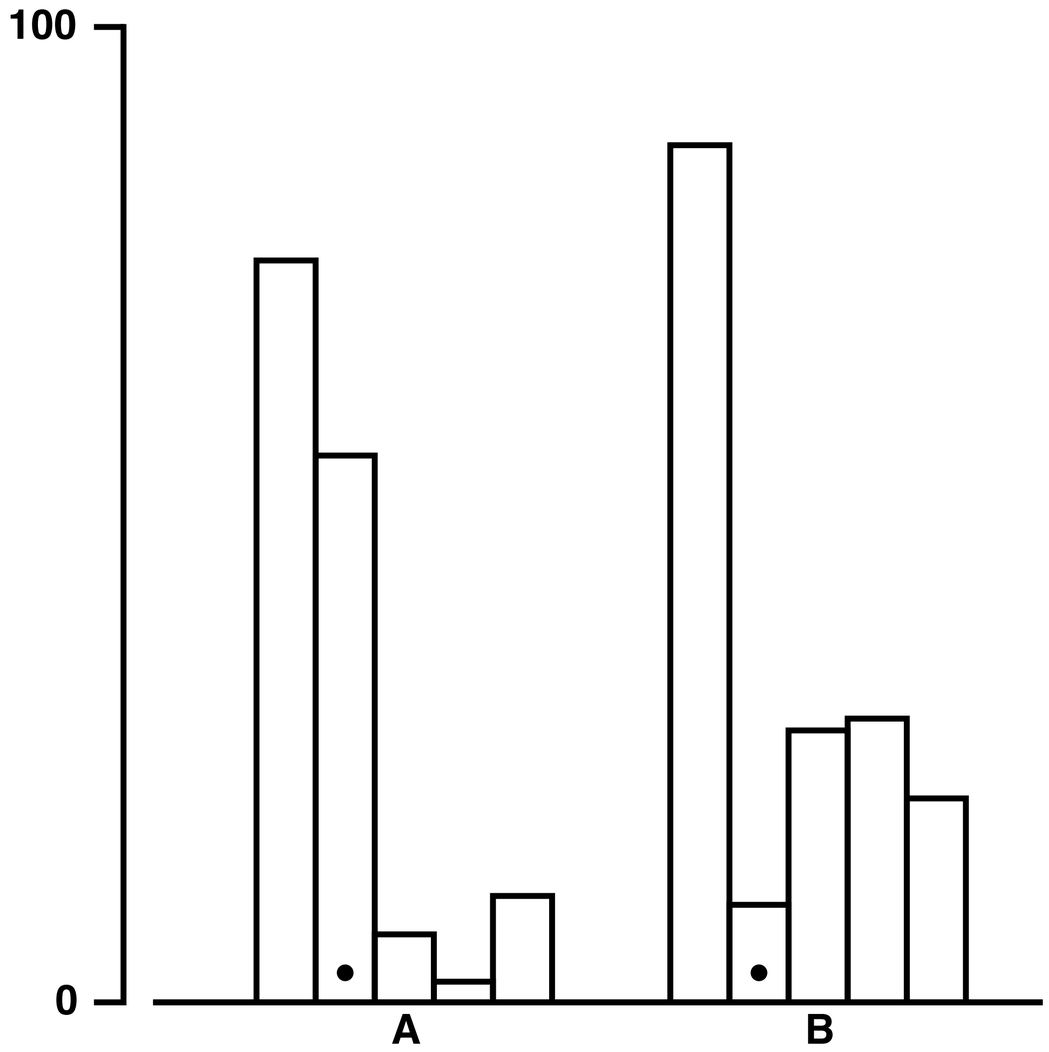
[Heer & Bostock, 2010]

Test % difference in **length** between elements



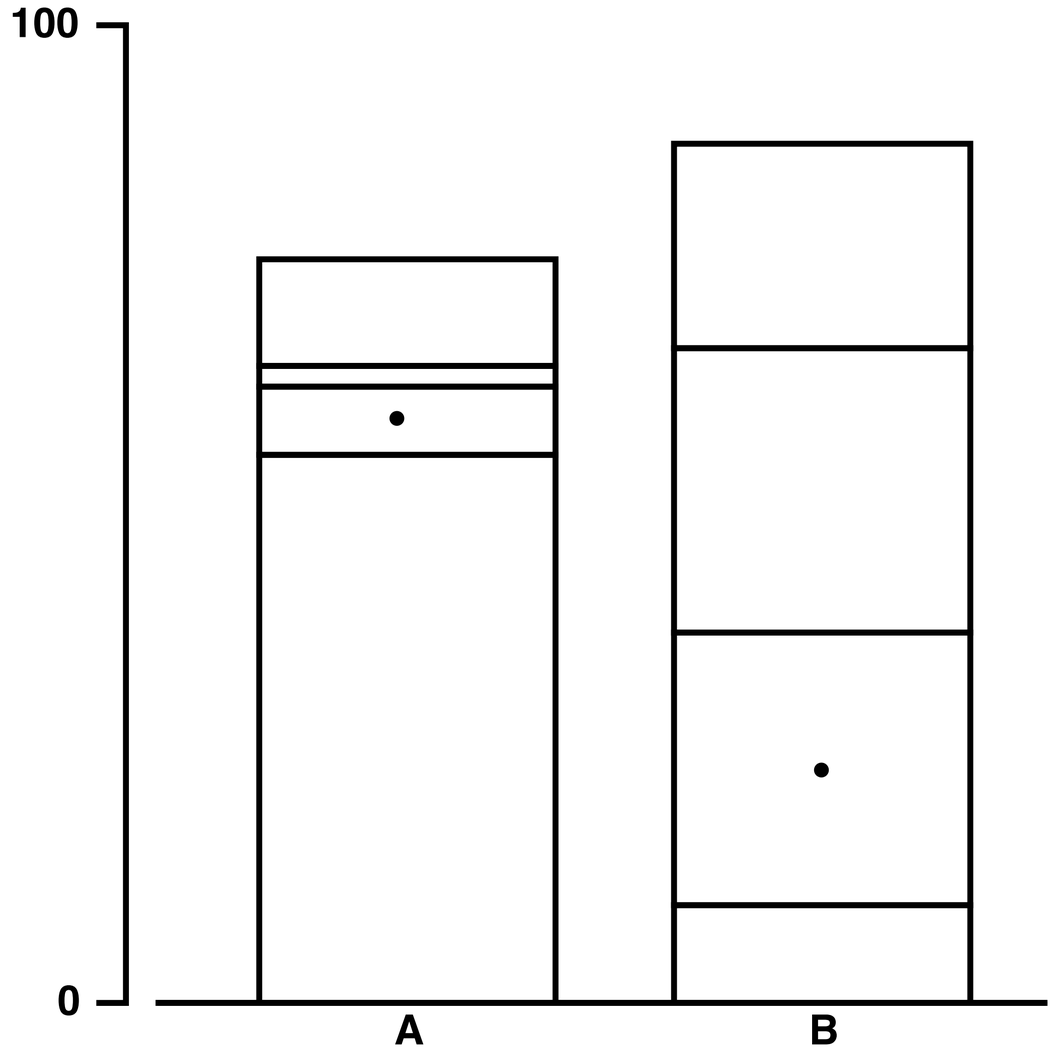
[Heer & Bostock, 2010]

Test % difference in **length** between elements



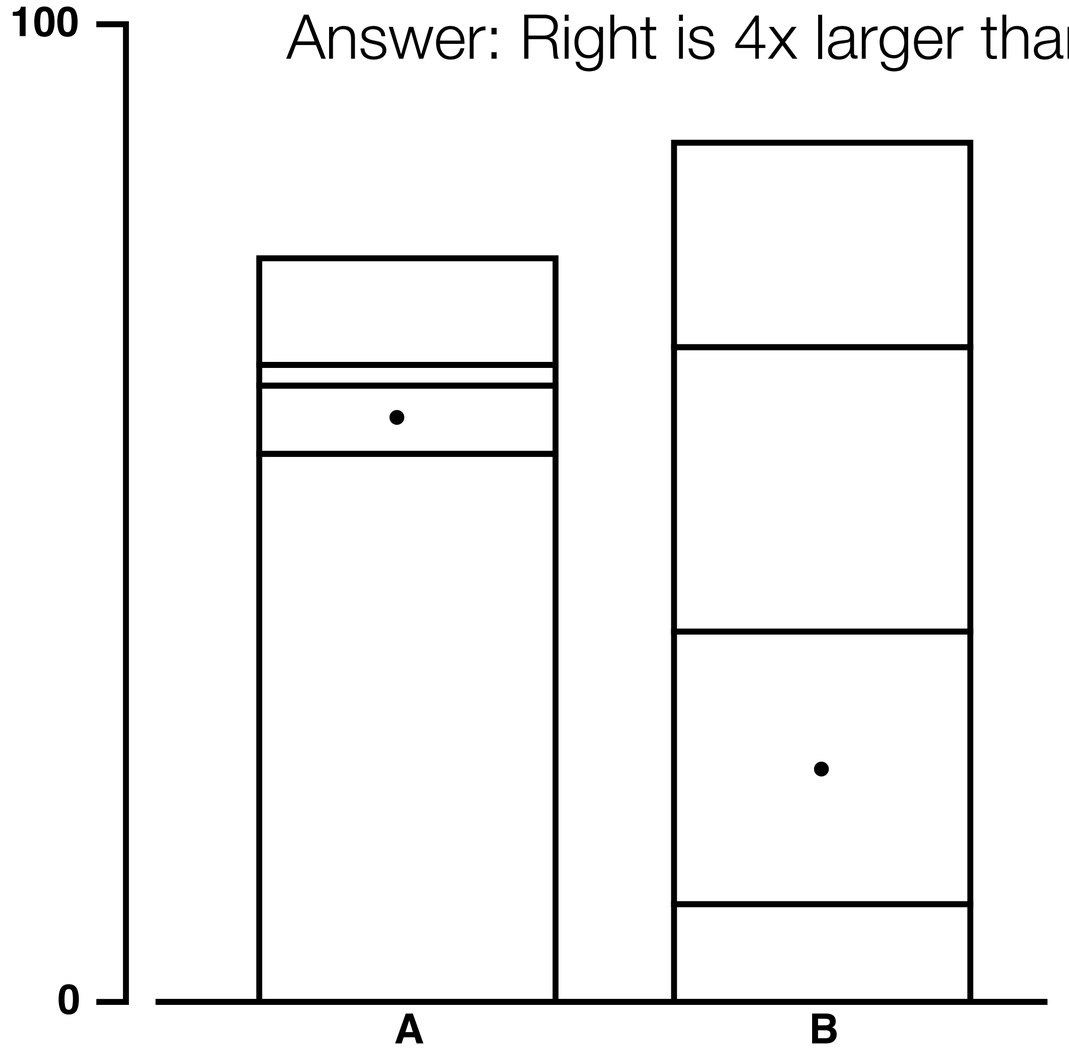
[Heer & Bostock, 2010]

Test % difference in **length** between elements



[Modified from Heer & Bostock, 2010]

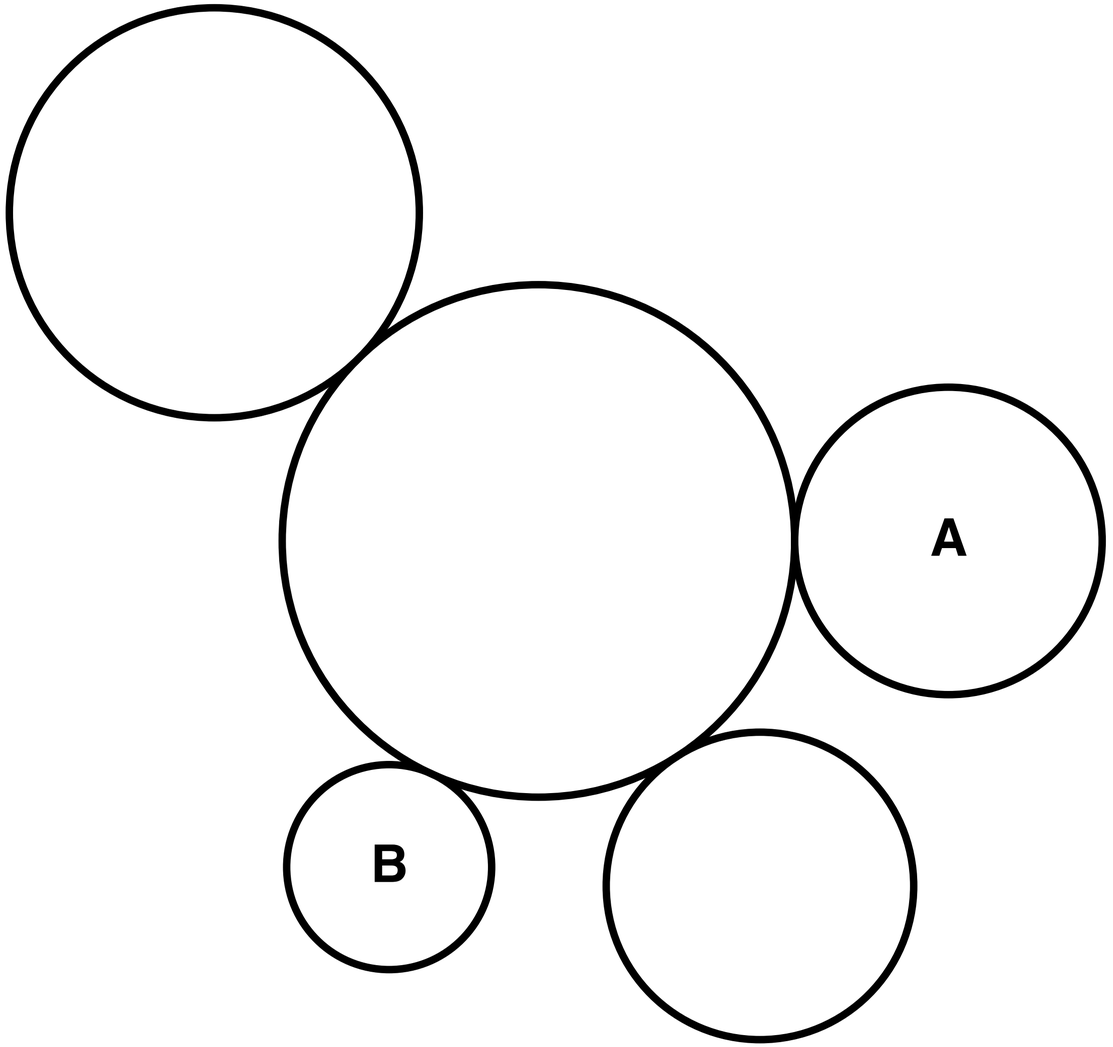
Test % difference in **length** between elements



Answer: Right is 4x larger than Left

[Modified from Heer & Bostock, 2010]

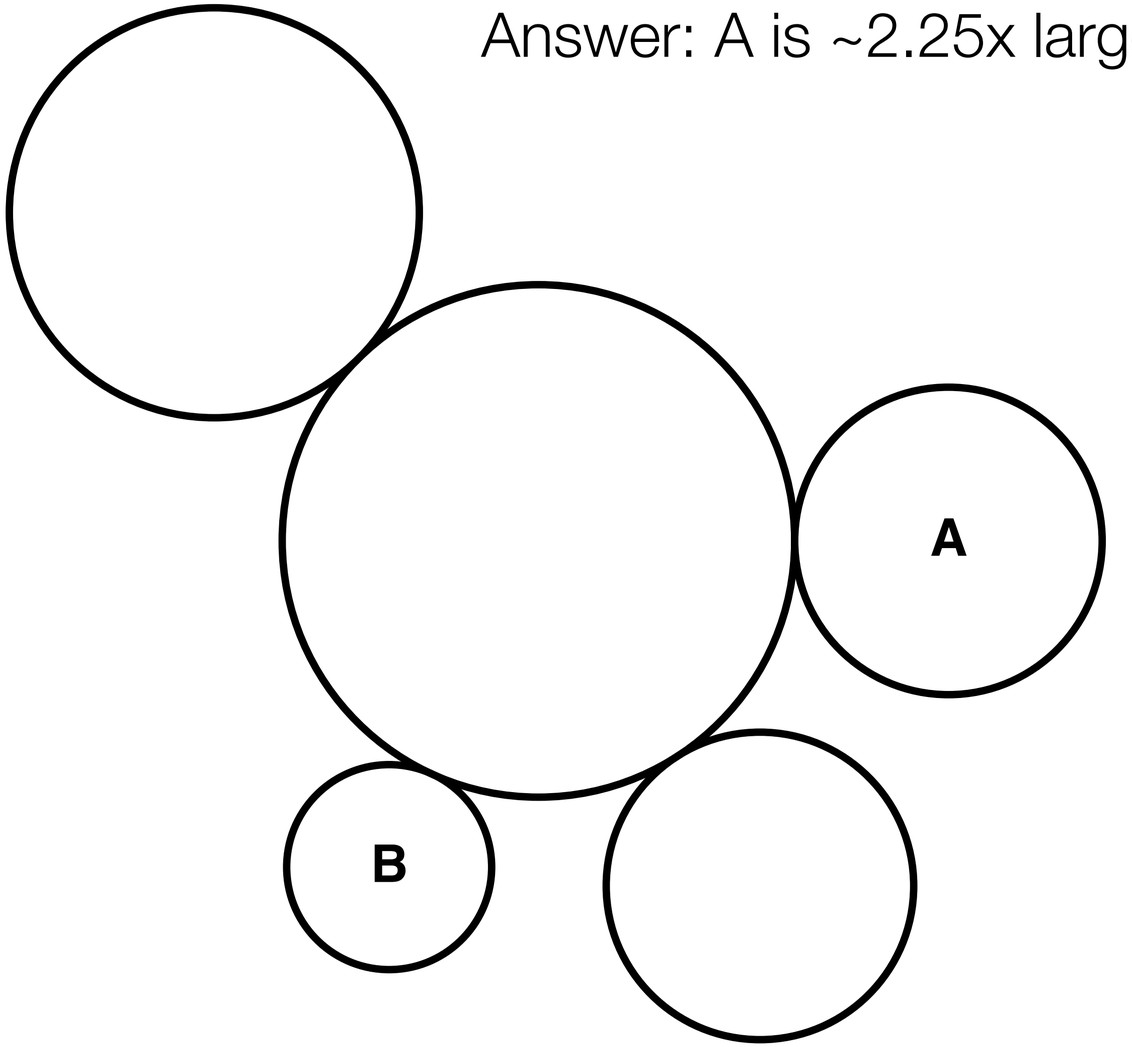
Test % difference in **area** between elements



[Heer & Bostock, 2010]

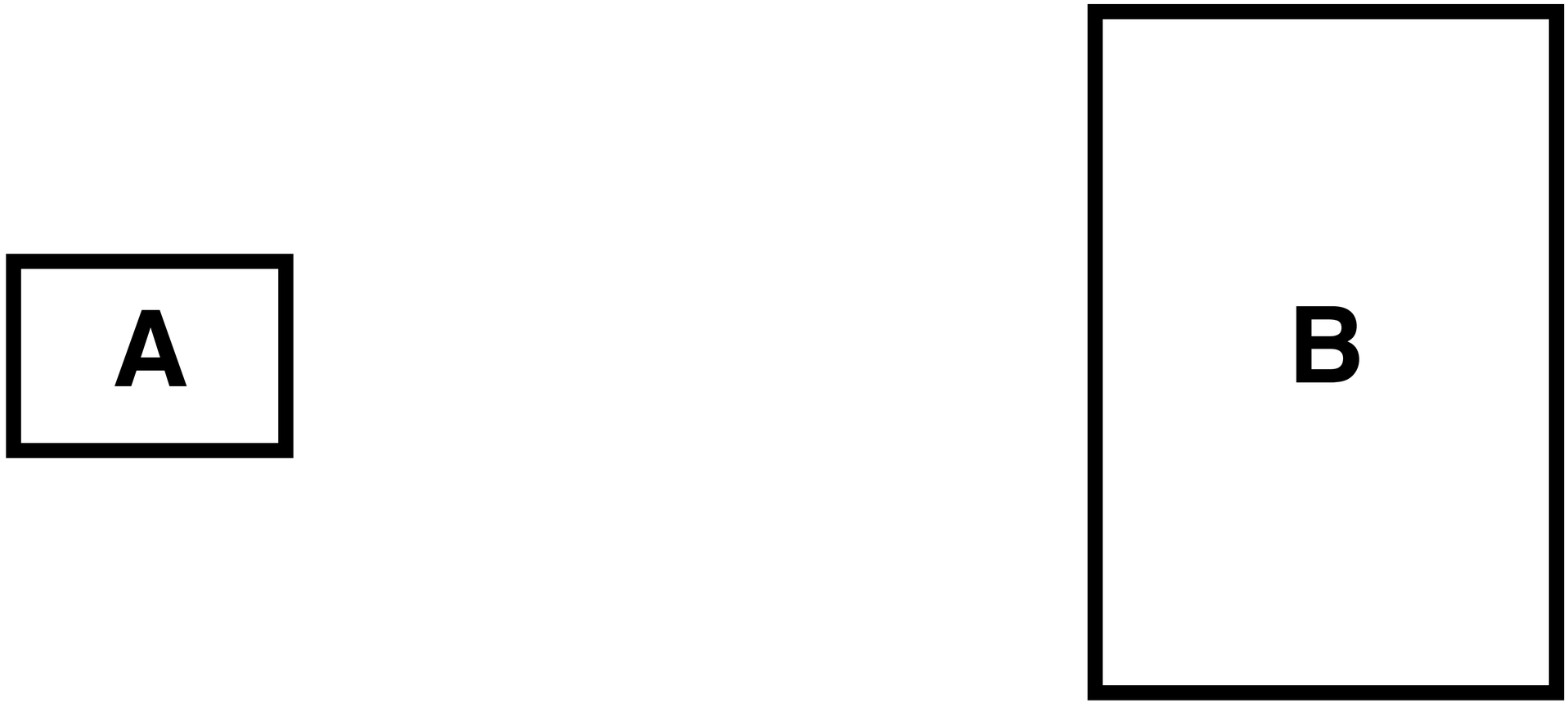
Test % difference in **area** between elements

Answer: A is ~2.25x larger (in area) than B



[Heer & Bostock, 2010]

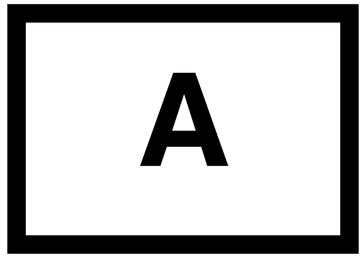
Test % difference in **area** between elements

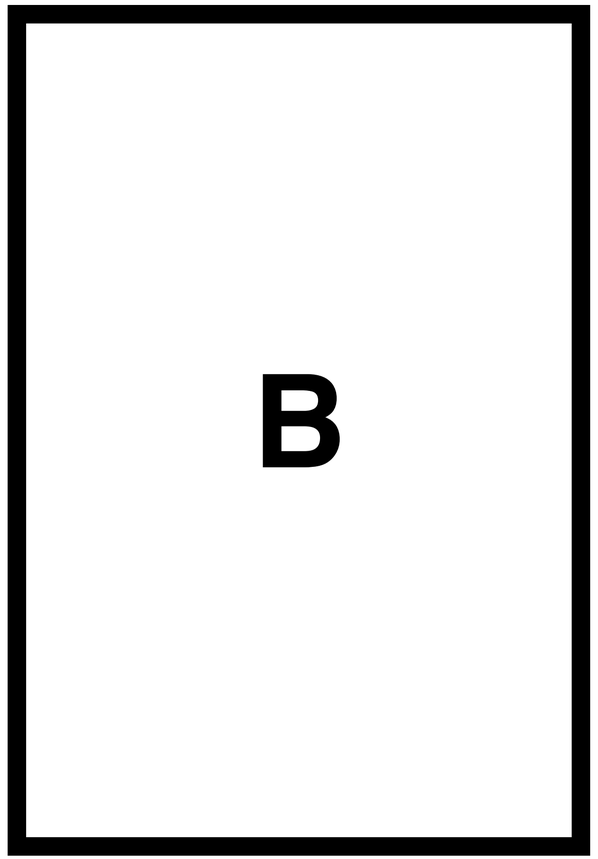


[Heer & Bostock, 2010]

Test % difference in **area** between elements

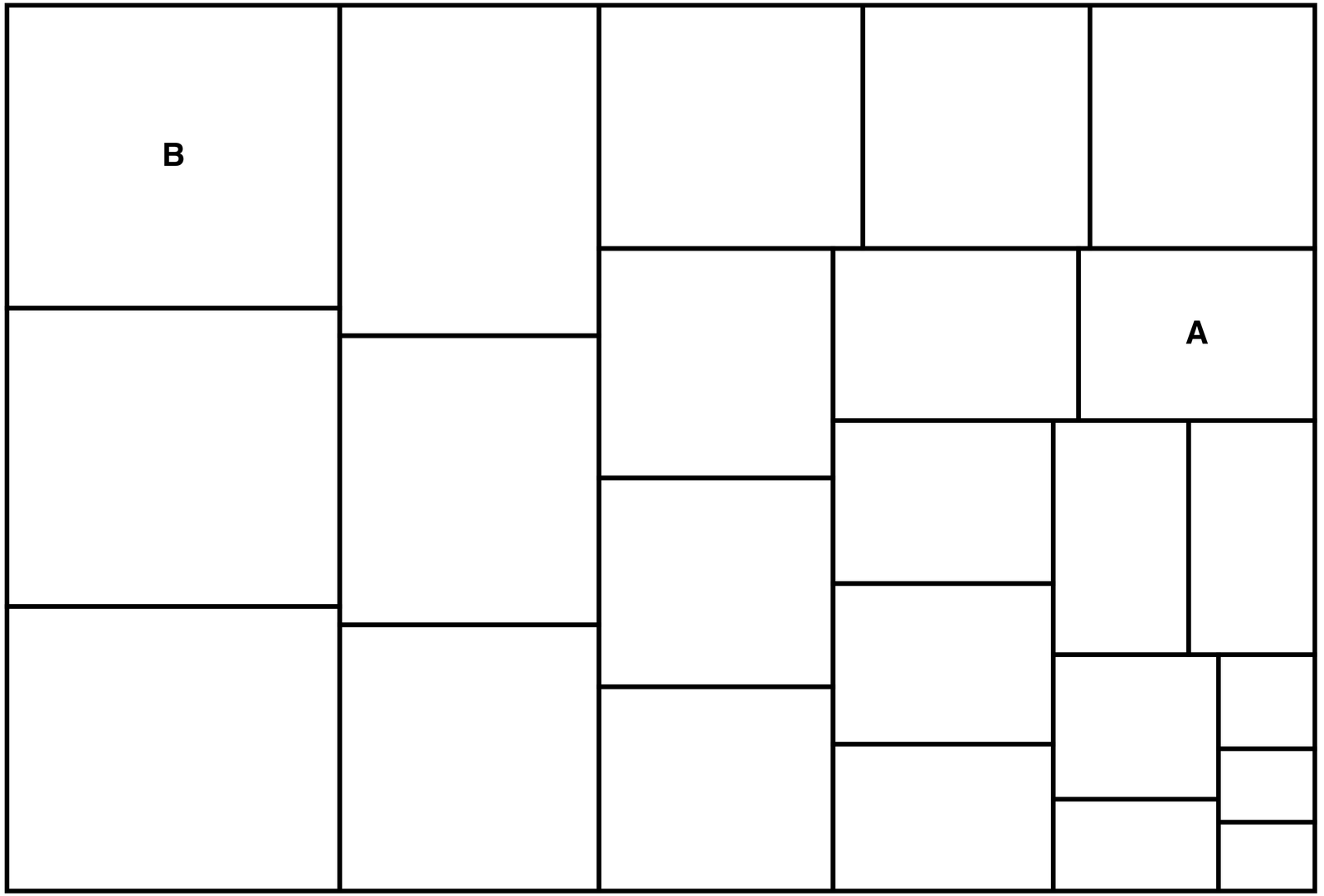
Answer: B is ~6.1x larger (in area) than A





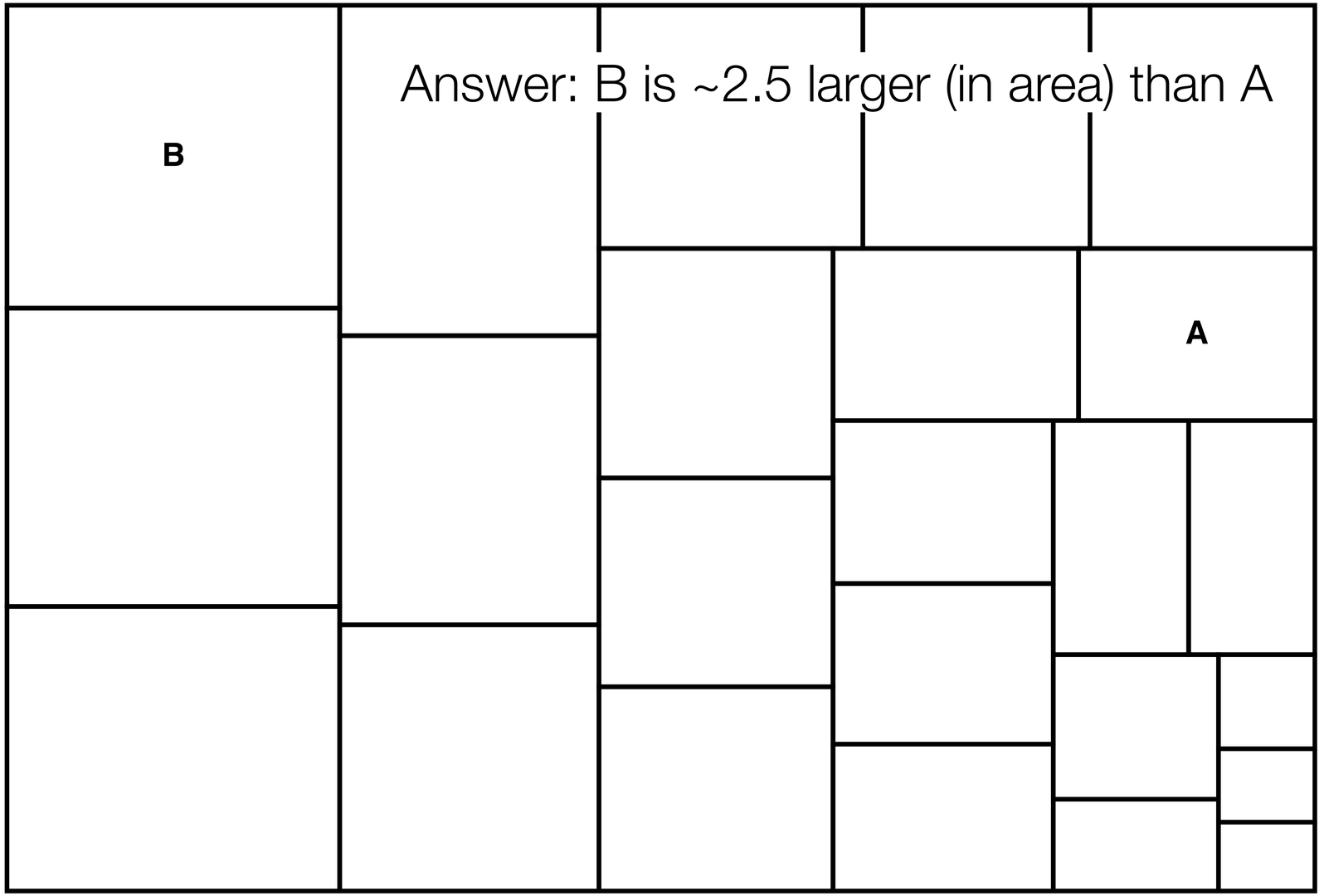
[Heer & Bostock, 2010]

Test % difference in **area** between elements



[Heer & Bostock, 2010]

Test % difference in **area** between elements



[Heer & Bostock, 2010]

Cleveland & McGill Experiments

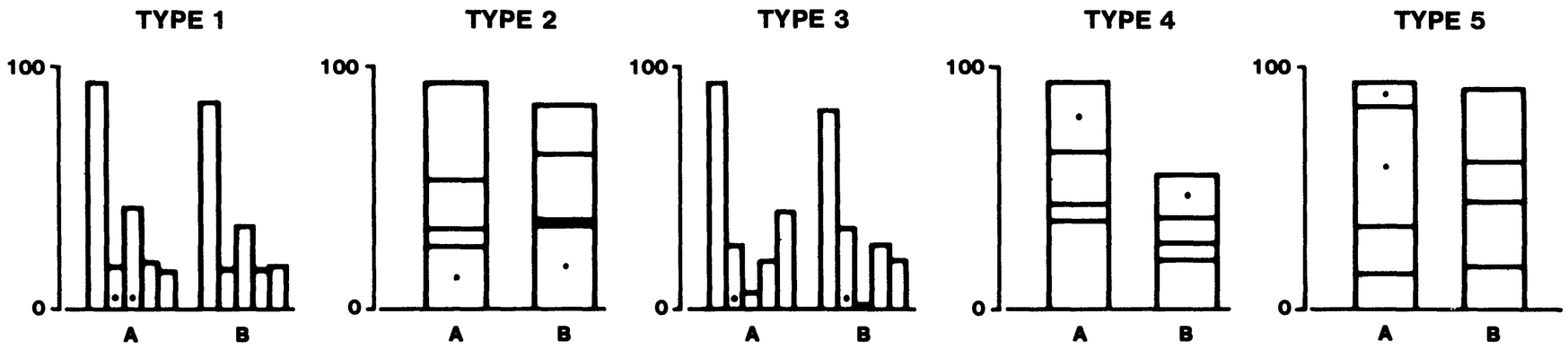


Figure 4. Graphs from position–length experiment.

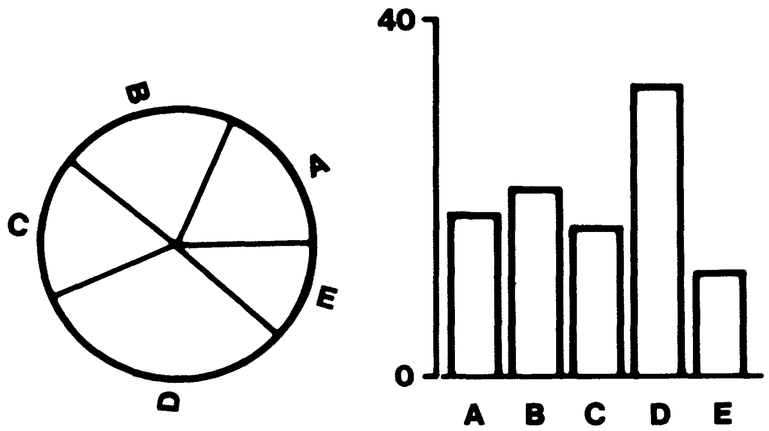


Figure 3. Graphs from position–angle experiment.

[Cleveland & McGill, 1984]

Heer & Bostock Experiments

- Rerun Cleveland & McGill's experiment using Mechanical Turk
- … with more tests

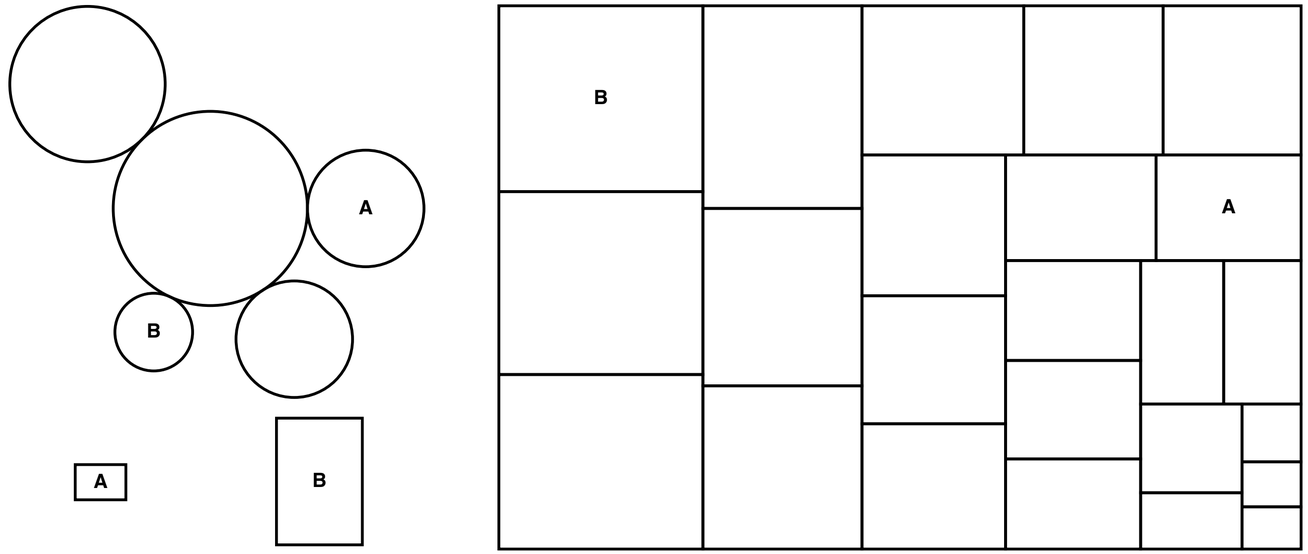
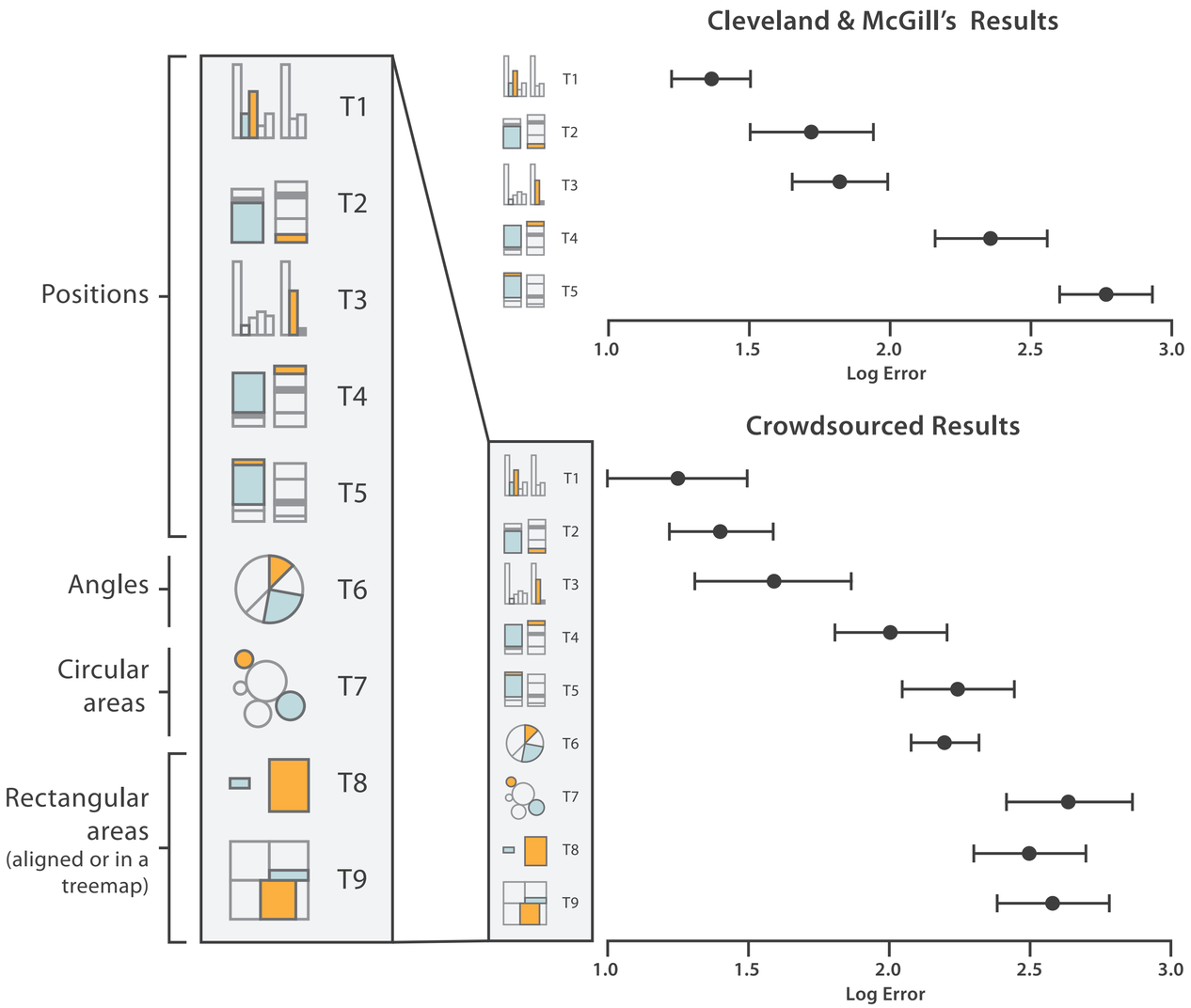


Figure 2: Area judgment stimuli. Top left: Bubble chart (T7), Bottom left: Center-aligned rectangles (T8), Right: Treemap (T9).

[Heer & Bostock, 2010]

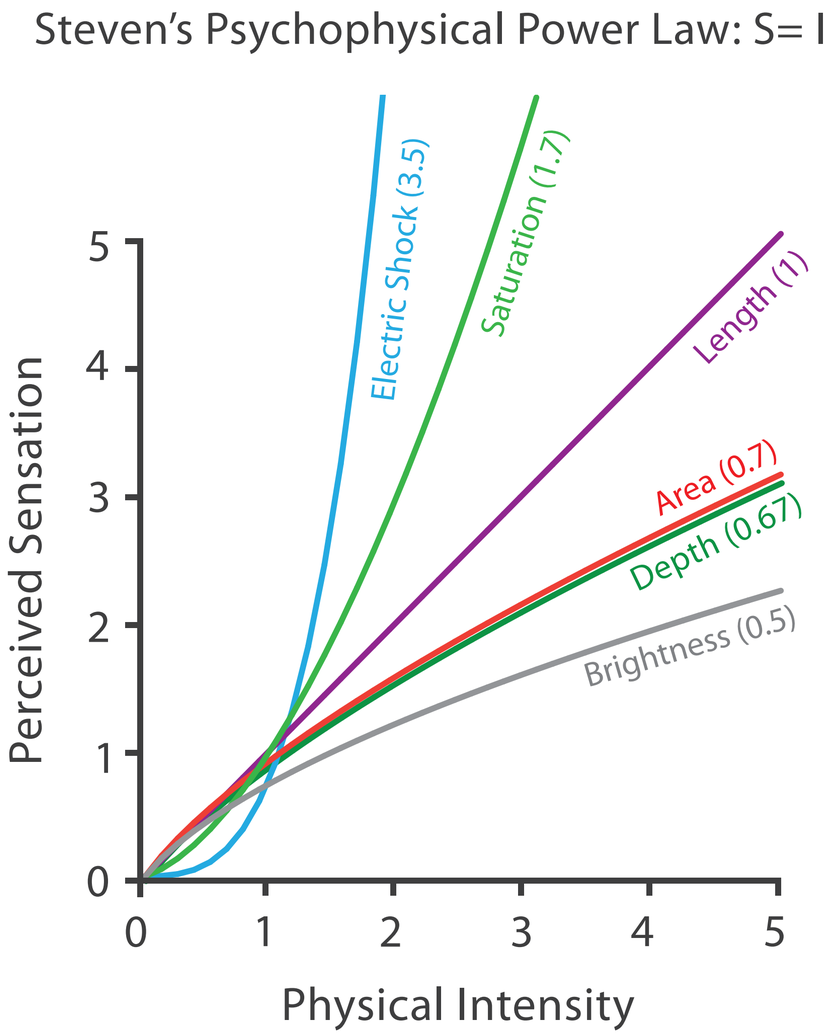
Results Summary



[Munzner (ill. Maguire) based on Heer & Bostock, 2014]

Psychophysics

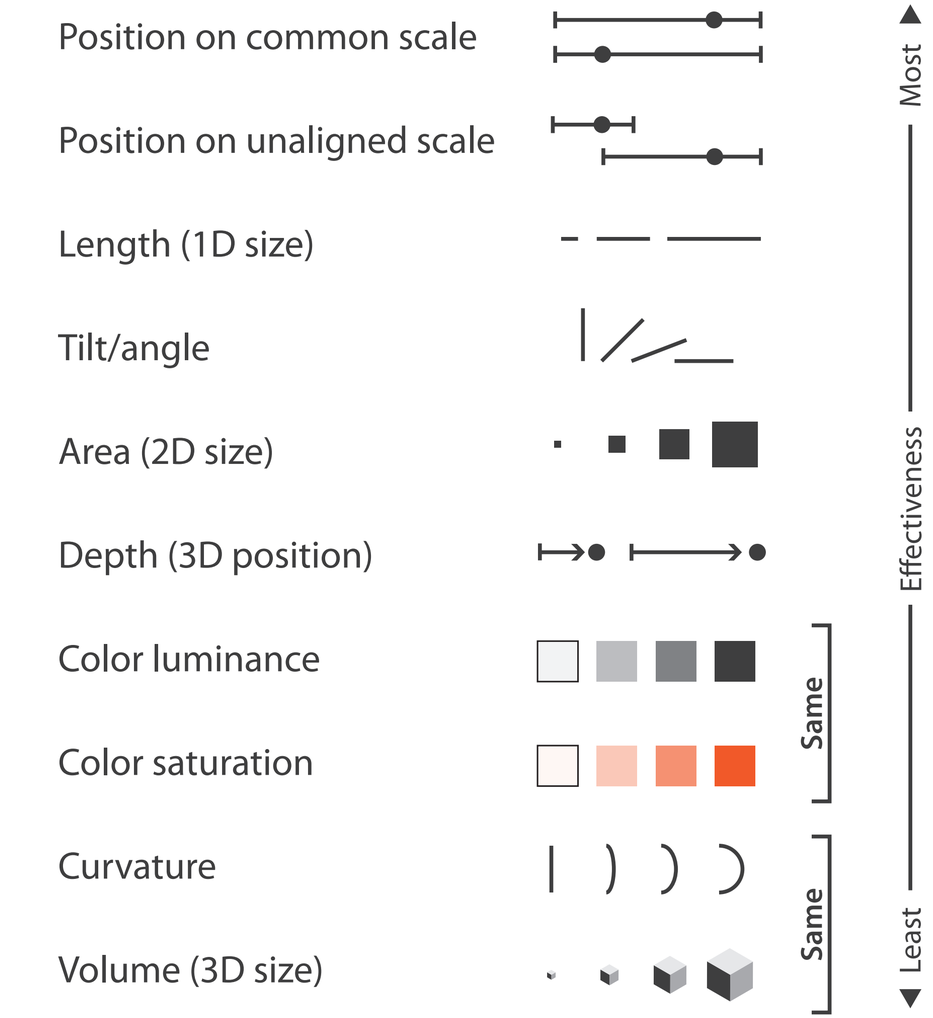
- How do we perceive changes in stimuli
- The Psychophysical Power Law [Stevens, 1975]: All sensory channels follow a power function based on stimulus intensity ($S = I^n$)
- Length is fairly accurate
- Magnified vs. compressed sensations



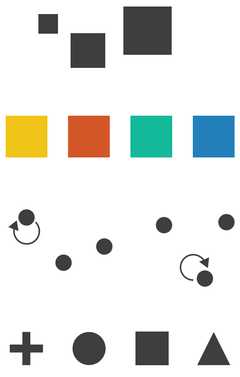
[Munzner (ill. Maguire), 2014]

Ranking Channels by Effectiveness

Magnitude Channels: Ordered Attributes



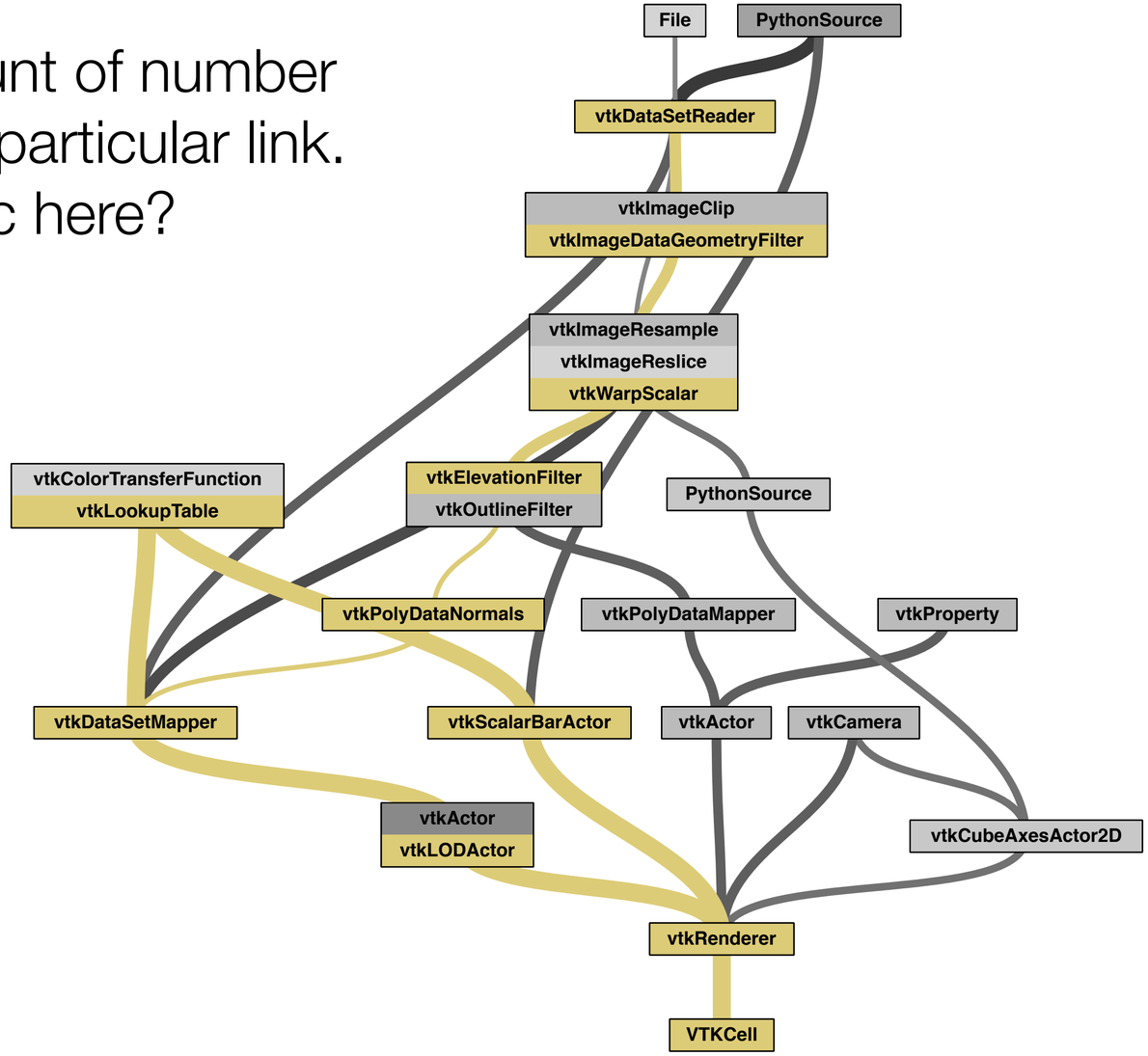
Identity Channels: Categorical Attributes



[Munzner (ill. Maguire), 2014]

Discriminability

- Width encodes count of number of networks with a particular link.
- What is problematic here?



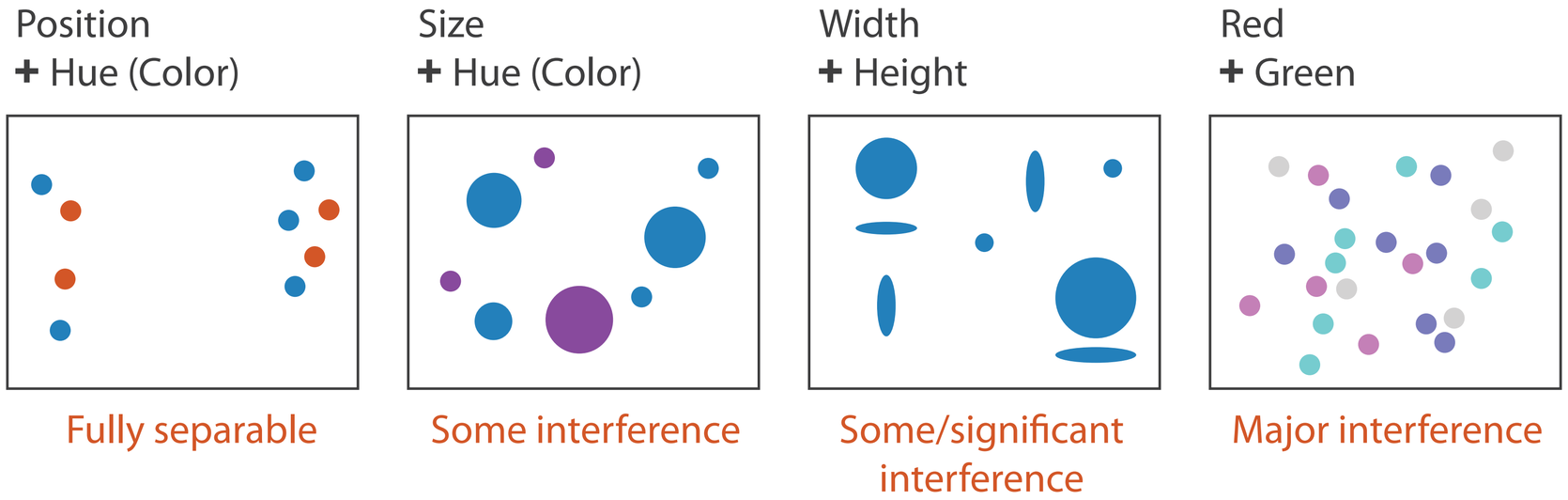
[Koop et al., 2013]

Discriminability

- Can someone tell the difference?
- How many values (bins) can be used so that a person can tell the difference?
- Example: Line width
  - Matching a particular width with a legend
  - Comparing two widths

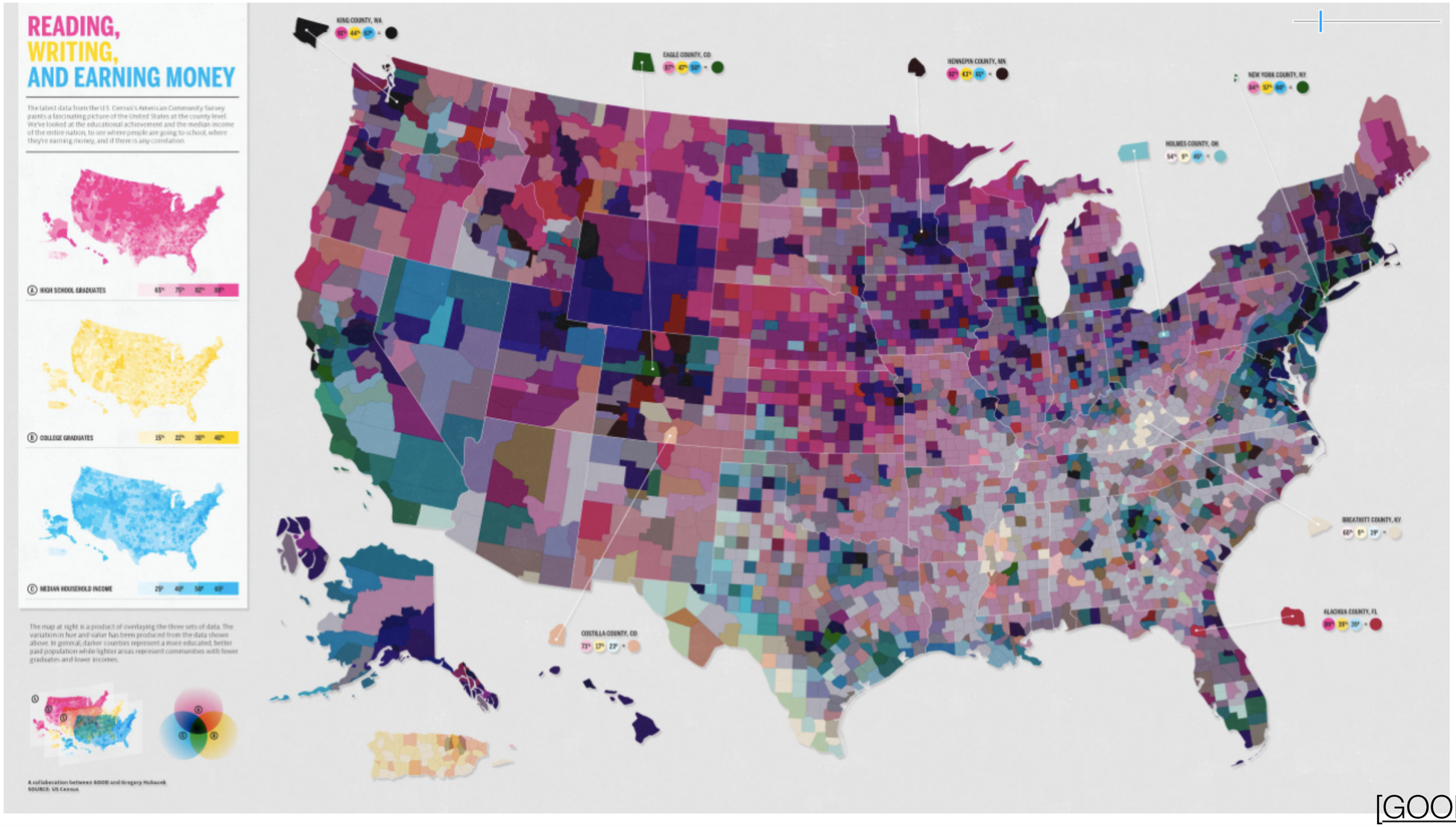
Separability

- Cannot treat all channels as independent!
- **Separable** means each individual channel can be distinguished
- **Integral** means the channels are perceived together



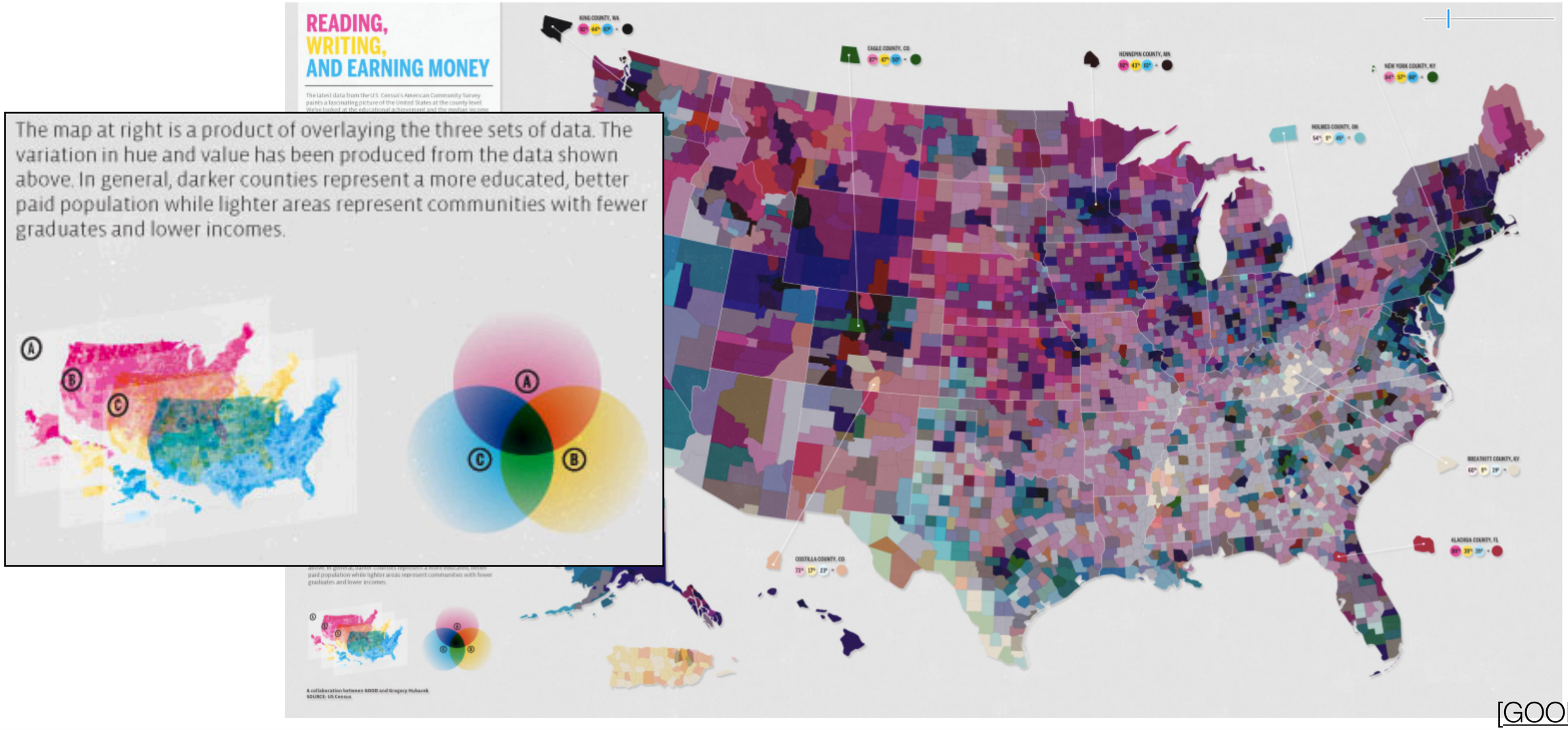
[Munzner (ill. Maguire) based on Ware, 2014]

Separable or Integral?



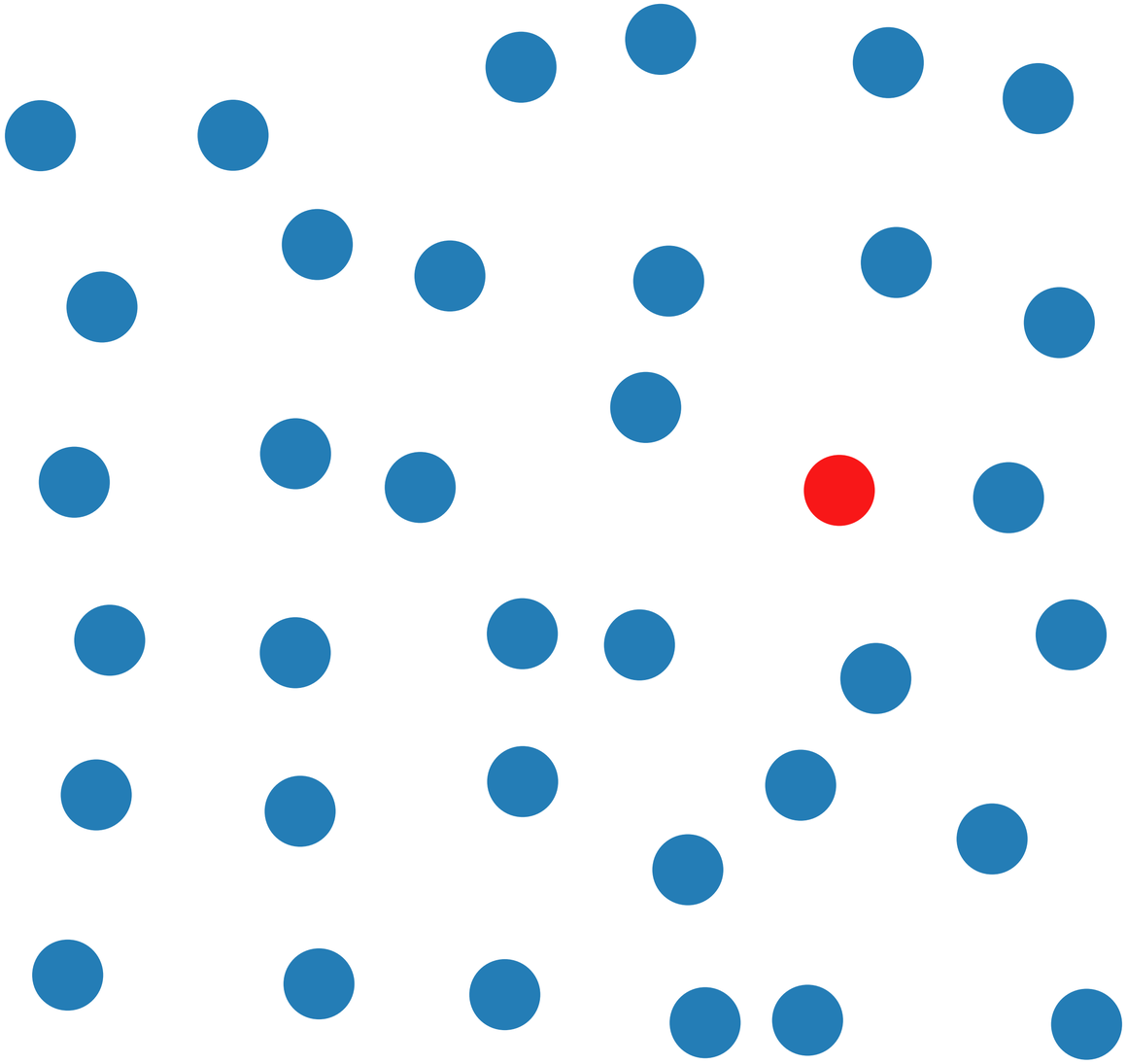
[GOOD]

Separable or Integral?



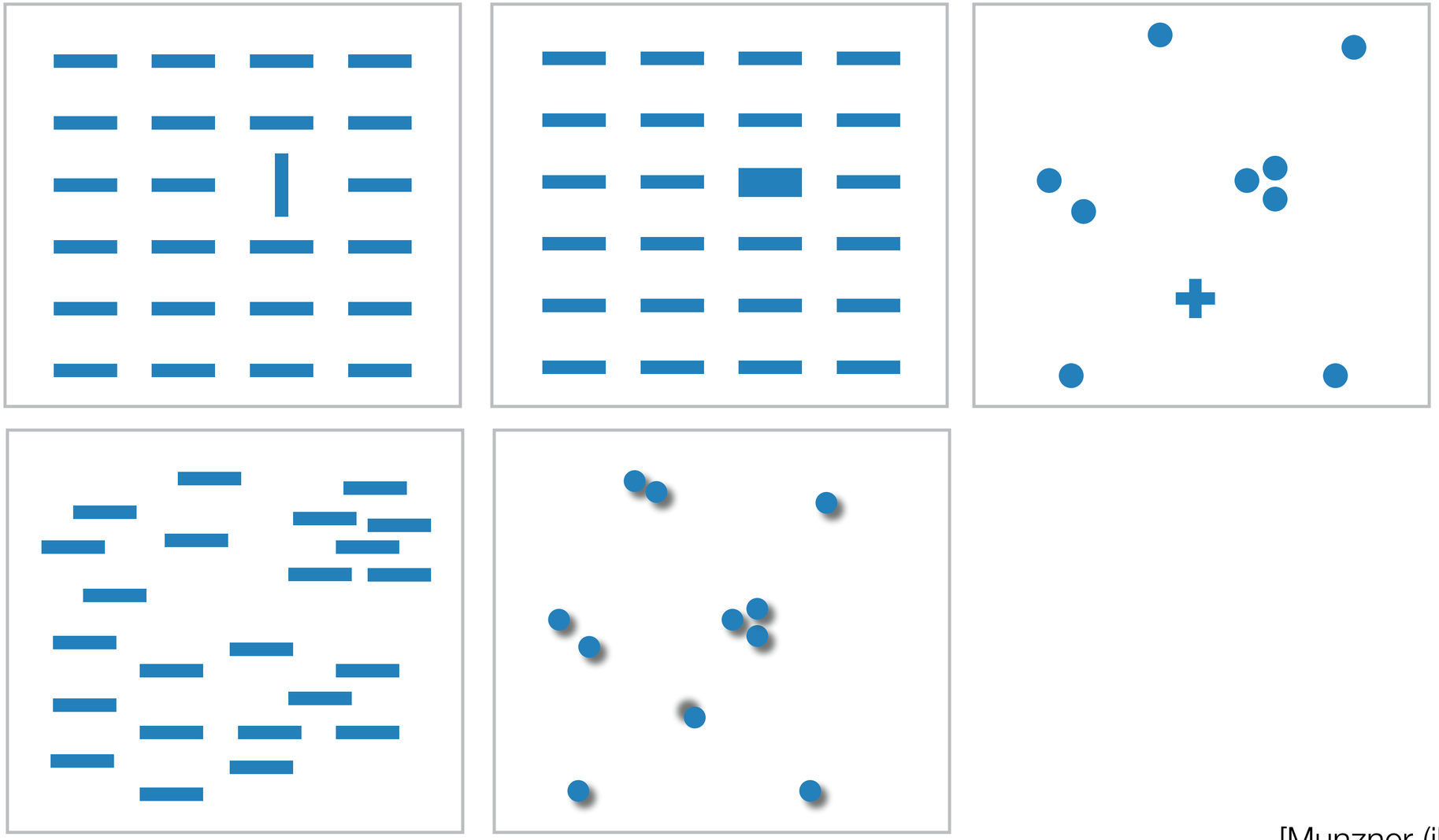
[GOOD]

Visual Popout



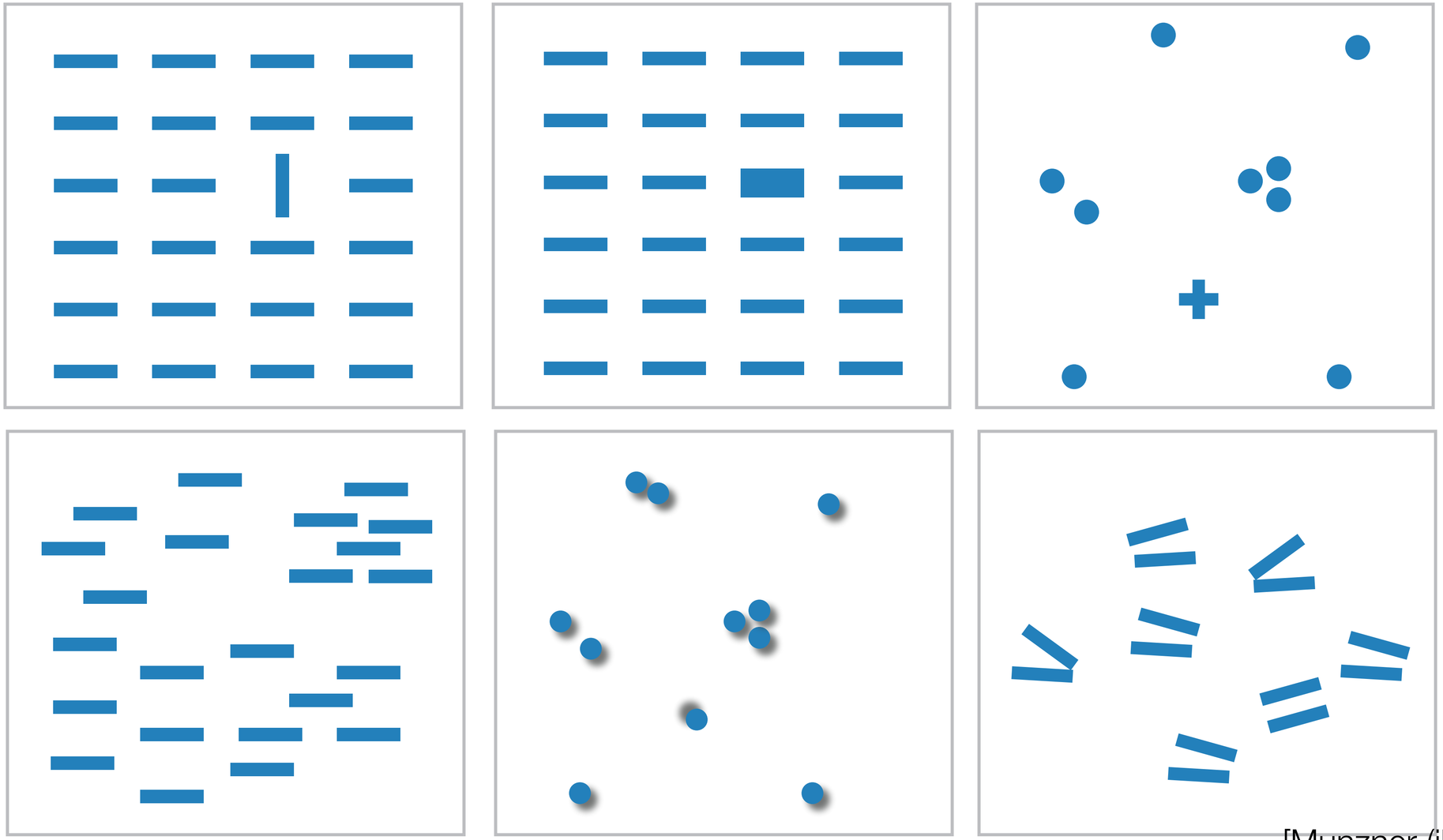
[C. G. Healey]

Visual Popout: Parallel Lines Require Search…



[Munzner (ill. Maguire), 2014]

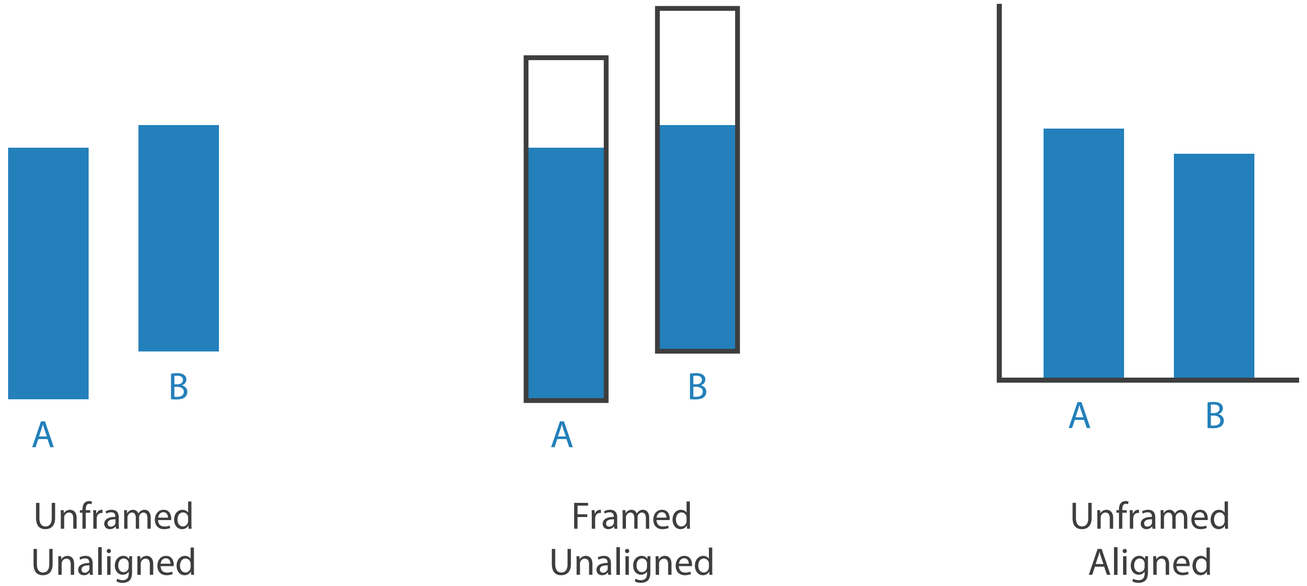
Visual Popout: Parallel Lines Require Search…



[Munzner (ill. Maguire), 2014]

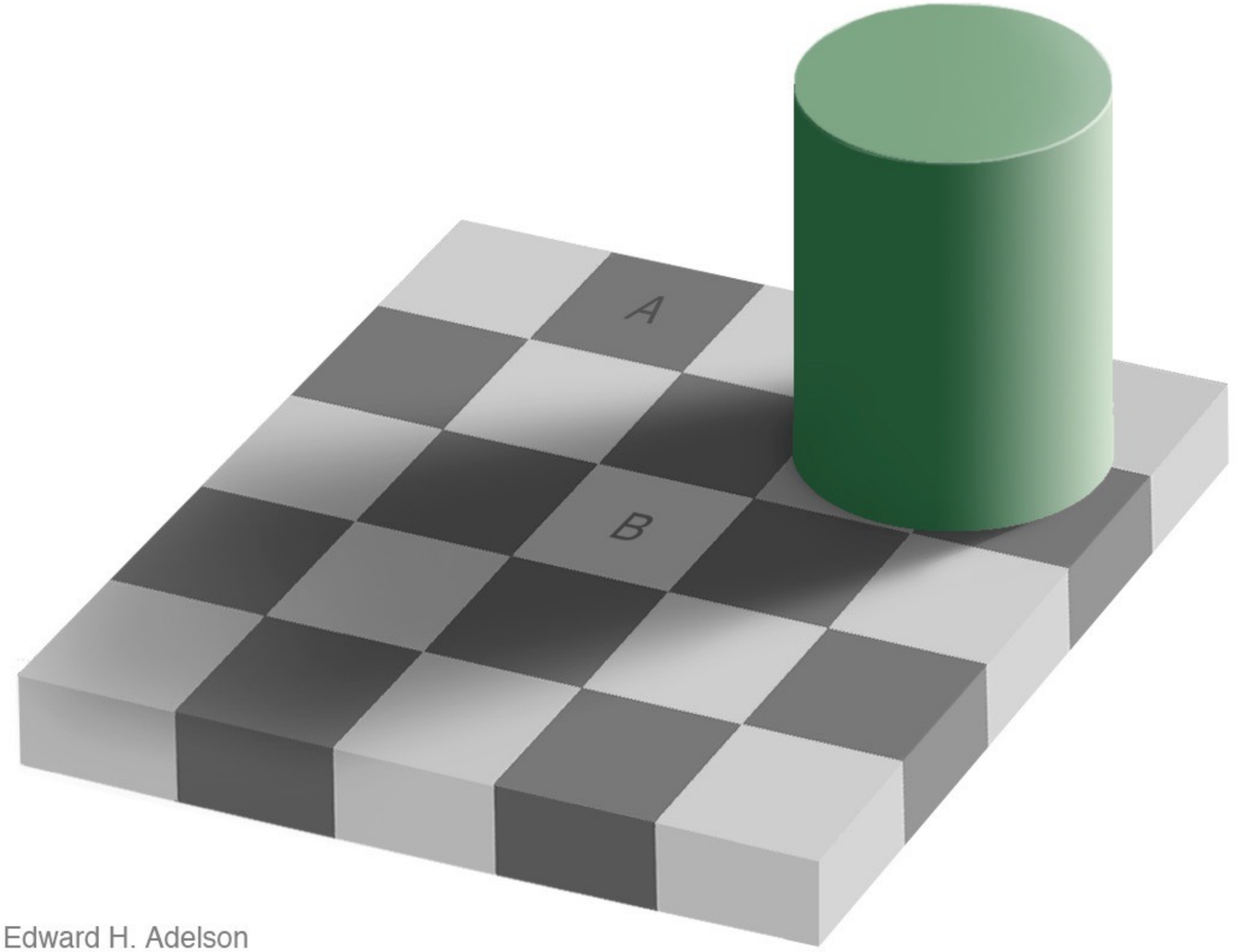
Relative vs. Absolute Judgments

- Weber's Law:
  - We judge based on relative (%-based) not absolute differences
  - The amount of perceived difference is relative to the object's magnitude!



[Munzner (ill. Maguire), 2014]

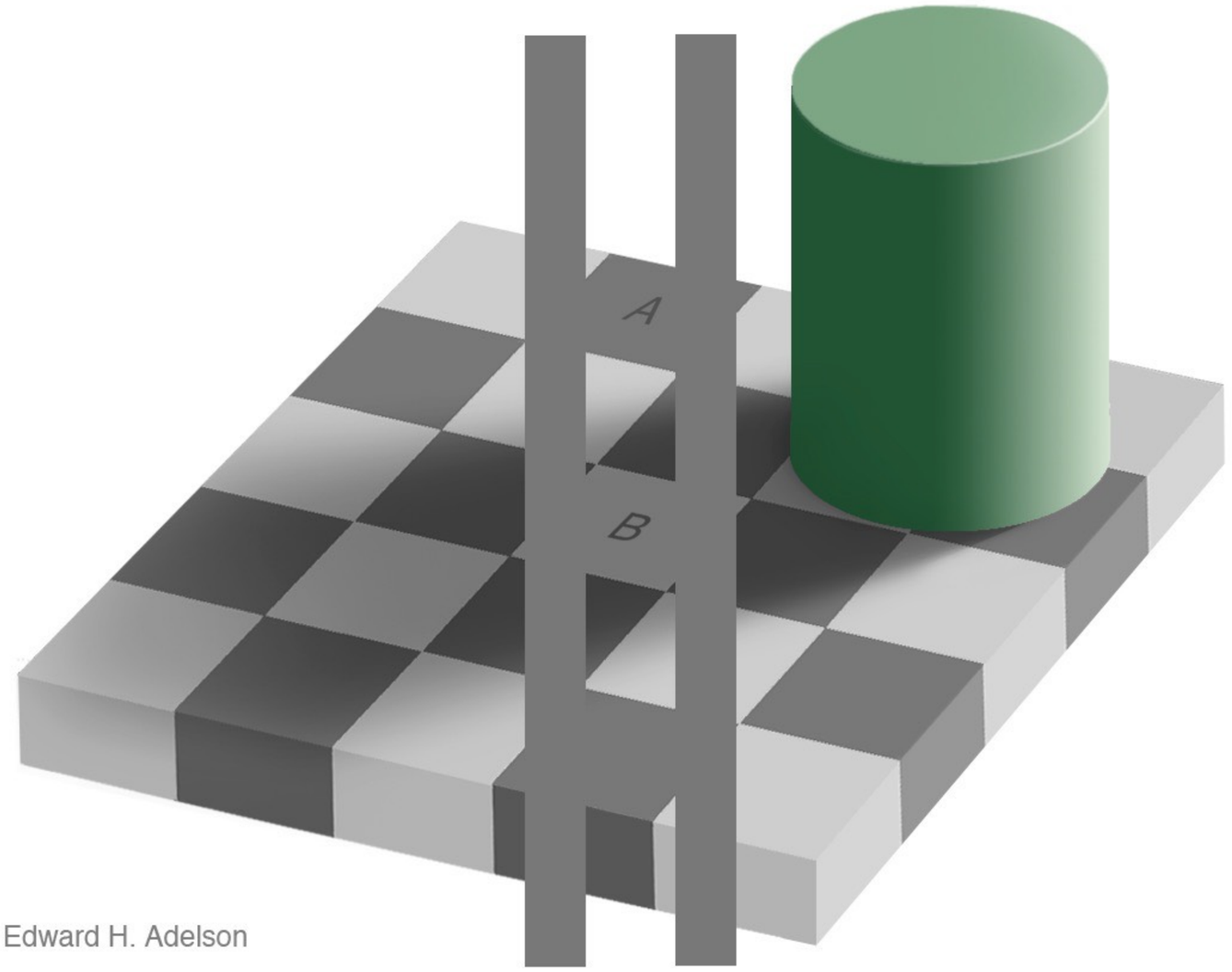
Luminance Perception



Edward H. Adelson

[E. H. Adelson, 1995]

Luminance Perception



[E. H. Adelson, 1995]

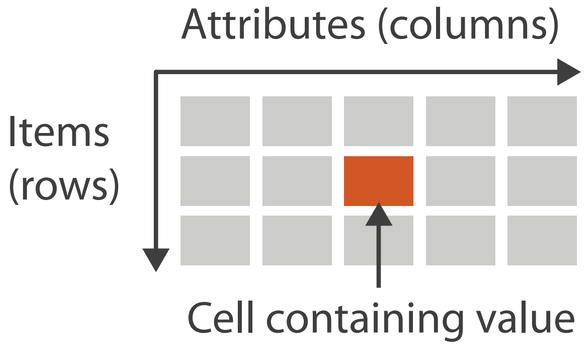
Visualizing Tabular Data

Tables

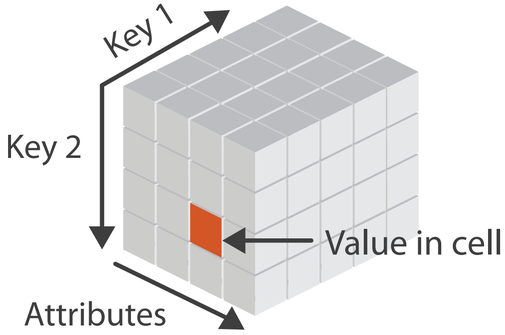
	REMOTE	STATION	FF	SEN/DIS	7-D AFAS UNL	D AFAS/RMF I	JOINT RR TKT	7-D UNL	30-D UNL
1	R011	42ND STREET & 8TH AVENUE	00228985	00008471	00000441	00001455	00000134	00033341	00071255
2	R170	14TH STREET-UNION SQUARE	00224603	00011051	00000827	00003026	00000660	00089367	00199841
3	R046	42ND STREET & GRAND CENTRAL	00207758	00007908	00000323	00001183	00003001	00040759	00096613
4	R012	34TH STREET & 8TH AVENUE	00188311	00006490	00000498	00001279	00003622	00035527	00067483
5	R293	34TH STREET - PENN STATION	00168768	00006155	00000523	00001065	00005031	00030645	00054376
6	R033	42ND STREET/TIMES SQUARE	00159382	00005945	00000378	00001205	00000690	00058931	00078644
7	R022	34TH STREET & 6TH AVENUE	00156008	00006276	00000487	00001543	00000712	00058910	00110466
8	R084	59TH STREET/COLUMBUS CIRCLE	00155262	00009484	00000589	00002071	00000542	00053397	00113966
9	R020	47-50 STREETS/ROCKEFELLER	00143500	00006402	00000384	00001159	00000723	00037978	00090745
10	R179	86TH STREET-LEXINGTON AVE	00142169	00010367	00000470	00001839	00000271	00050328	00125250
11	R023	34TH STREET & 6TH AVENUE	00134052	00005005	00000348	00001112	00000649	00031531	00075040
12	R029	PARK PLACE	00121614	00004311	00000287	00000931	00000792	00025404	00065362
13	R047	42ND STREET & GRAND CENTRAL	00100742	00004273	00000185	00000704	00001241	00022808	00068216

Visualization of Tables

- Items and attributes
- For now, attributes are not known to be positions
- Keys and values
  - **key** is an independent attribute that is unique and identifies item
  - **value** tells some aspect of an item
- Keys: categorical/ordinal
- Values: categorical/ordinal/quantitative
- Levels: unique *values* of categorical or ordered attributes



➔ *Multidimensional Table*



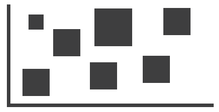
[Munzner (ill. Maguire), 2014]

Arrange Tables

Express Values

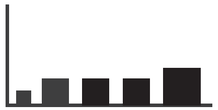
Separate, Order, Align Regions

- Separate



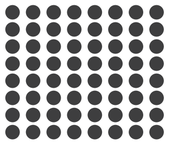
- Order

- Align



Layout Density

- Dense



- Space-Filling

- *1 Key*
  List

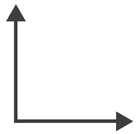
- *2 Keys*
  Matrix

- *3 Keys*
  Volume

- *Many Keys*
  Recursive Subdivision

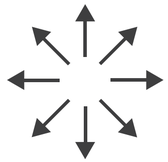
Axis Orientation

- Rectilinear



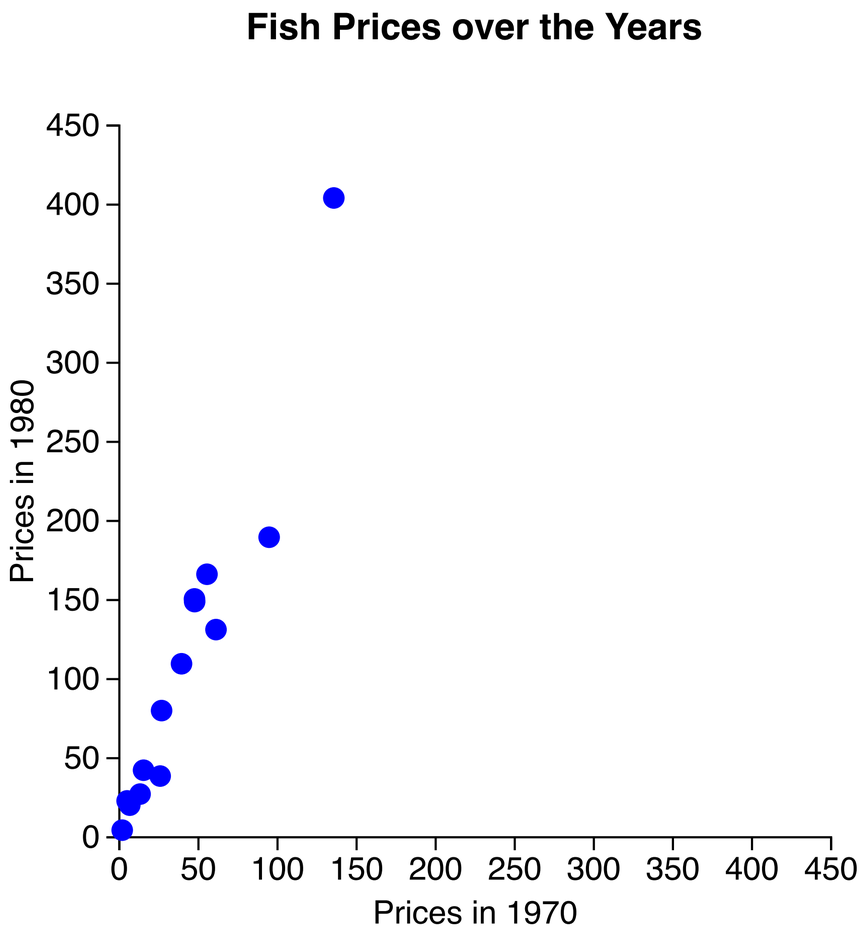
- Parallel

- Radial



[Munzner (ill. Maguire), 2014]

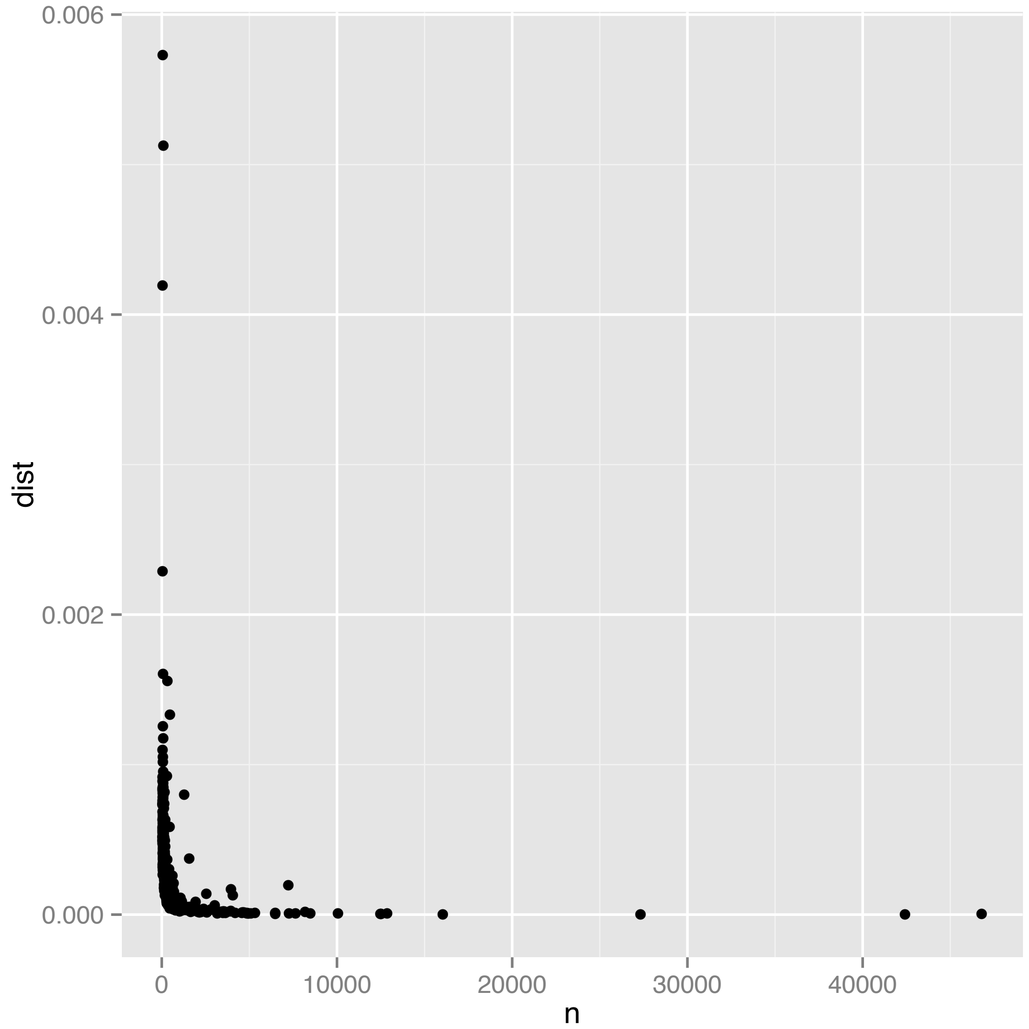
Express Values: Scatterplots



- Data: two quantitative values
- Task: find trends, clusters, outliers
- How: marks at spatial position in horizontal and vertical directions

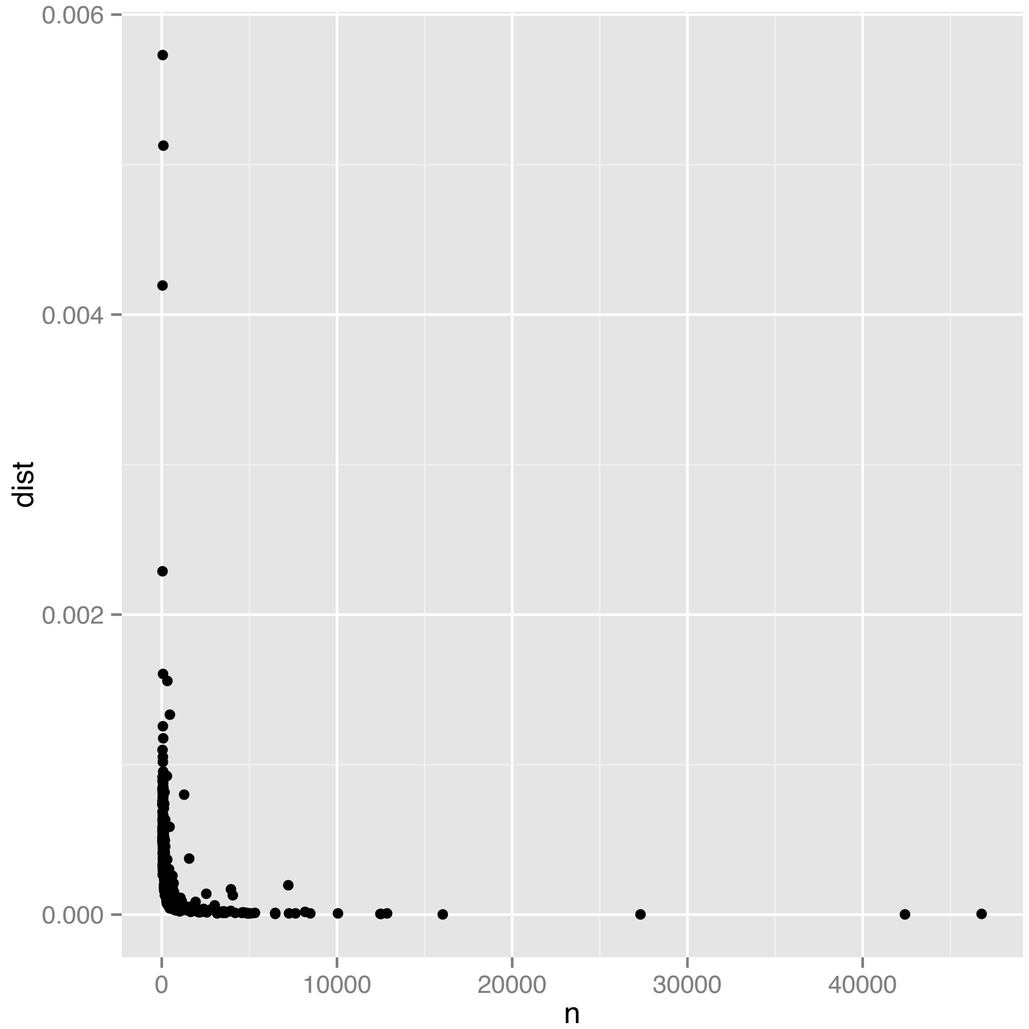
- Correlation: dependence between two attributes
  - Positive and negative correlation
  - Indicated by lines
- Coordinate system (axes) and labels are important!

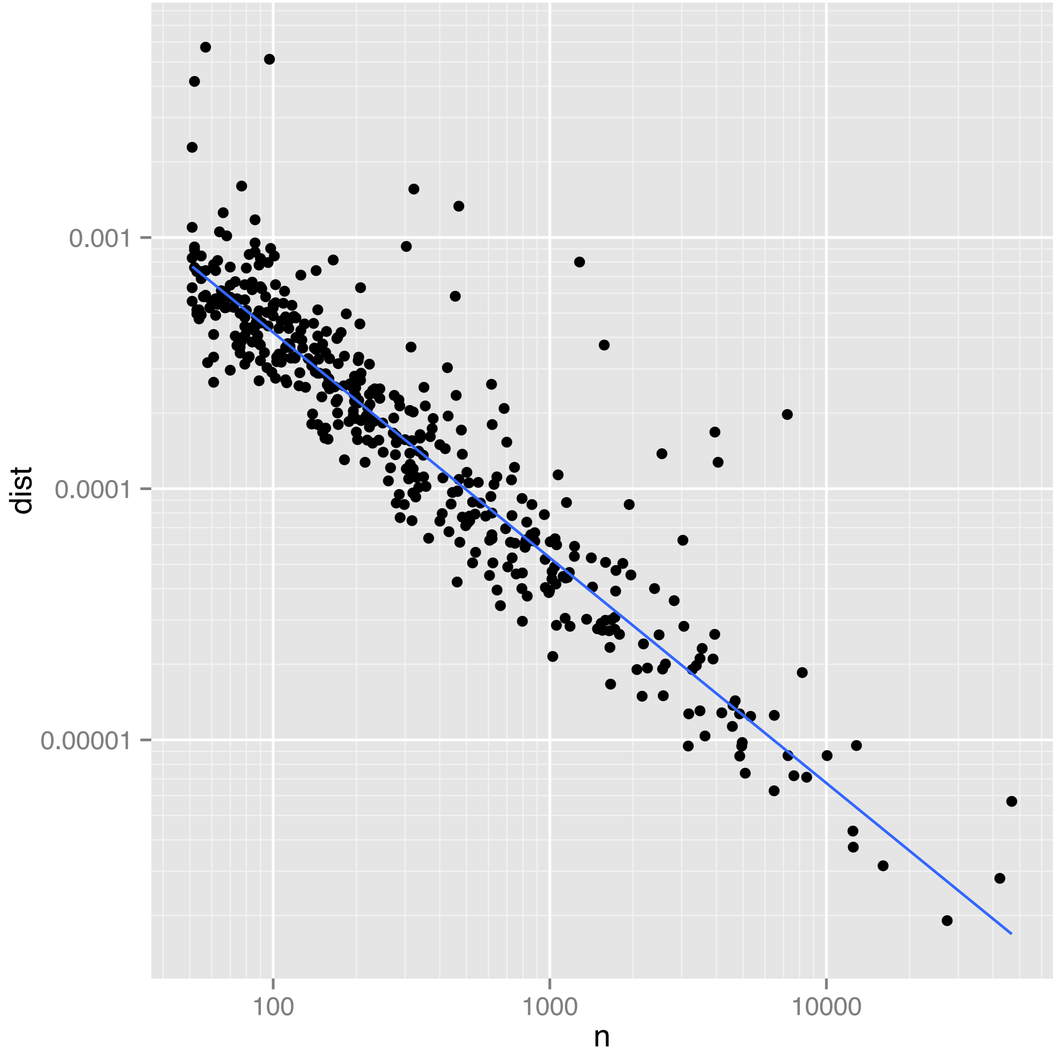
Coordinate Systems



[Wickham, 2014]

Coordinate Systems





[Wickham, 2014]