3B · VISUAL VARIABLES

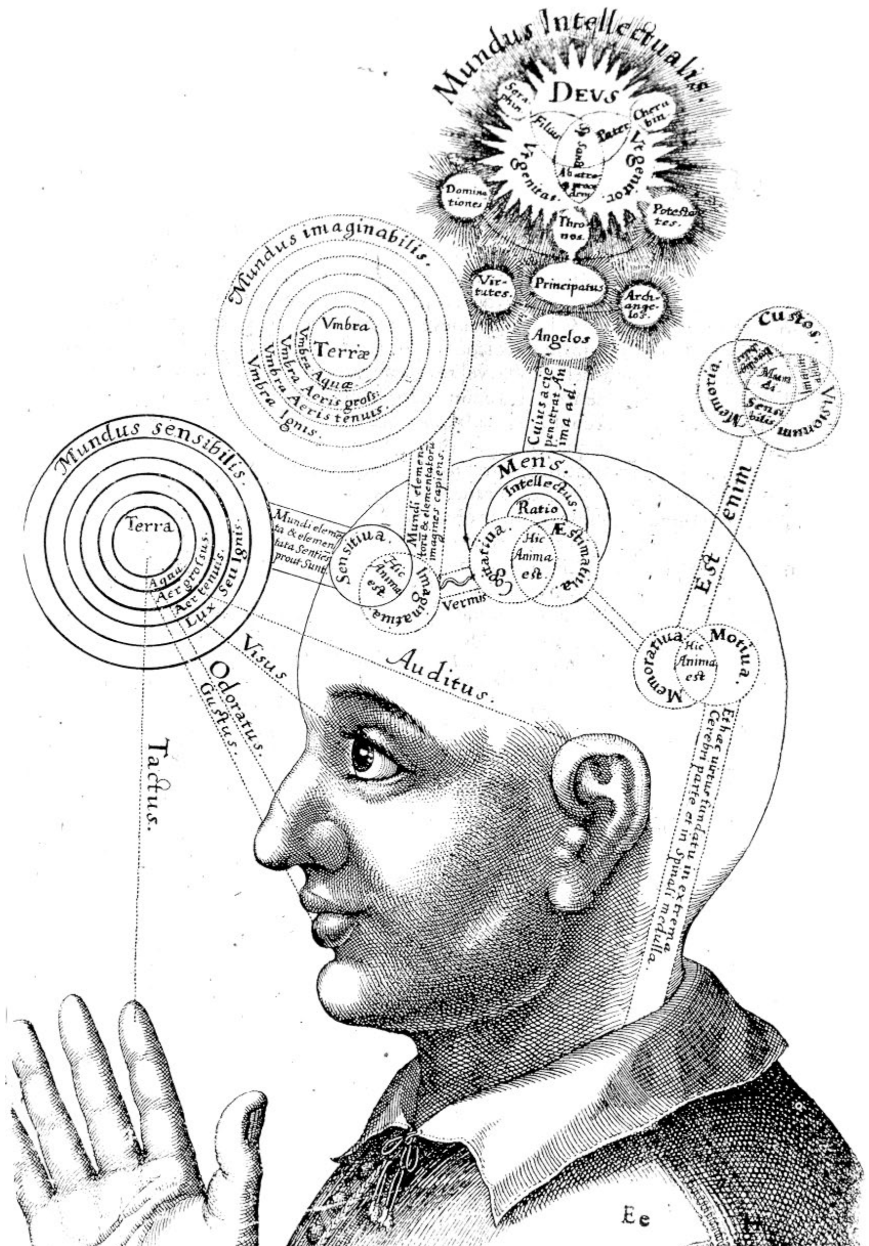
TEACHER

Salvo Rinzivillo

University of Pisa · Department of Computer Science

Course: Visual Analytics (602AA)

Academic Year 2024–2025



Data Types

- **Nominal (N)**
  - Equality relation
  - Apples, bananas, pears...
- **Ordinal (O)**
  - Ordering relation
  - Small, medium, large, darker, light...
- **Quantitative (Q)**
  - Arithmetic relations
  - 10m, 32 degrees, 2 bars...

Data Types - Operators

- **Nominal**
  - ≠, =
- **Ordinal**
  - ≠, =, >, <
- **Quantitative Interval** (no reference point)
  - ≠, =, >, <, +, -
  - Dates, Location
  - Distances: A is 3 degrees hotter than B
- **Quantitative Ratio** (reference point)
  - ≠, =, >, <, +, -, ×, ÷
  - Length, mass
  - Proportions: A is twice as large as B

From Data to Conceptual Model

- **Data Model:** low-level representation of data and operations
- **Conceptual Model:** mental and semantic construction

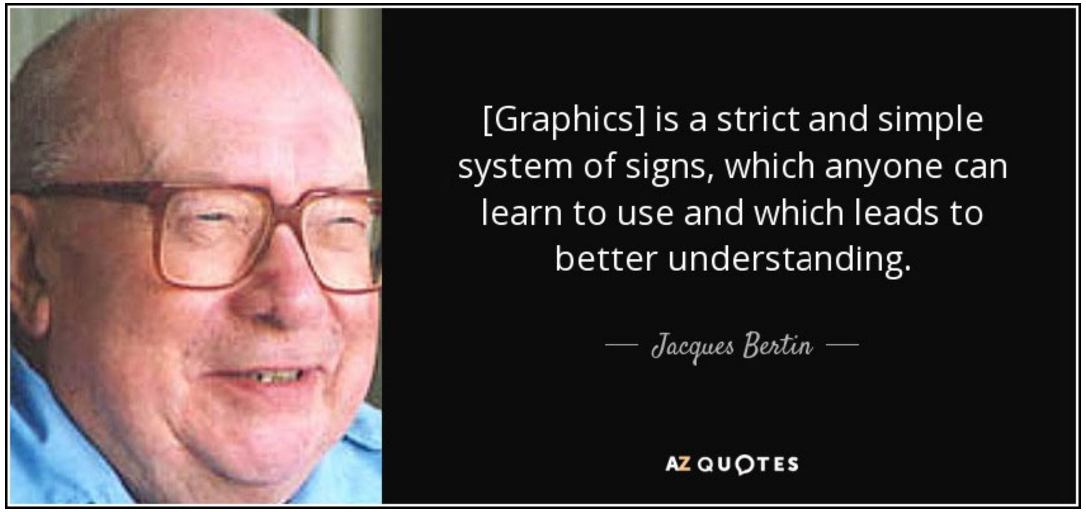
Data	Concept
1D number	Temperature
2D numbers	Geographic Coordinate
3D numbers	Spatio-temporal position

From Data to Conceptual Model

- From **data model**...
  - 70.8, 27.2, -10.2...
- ...using **conceptual model**...
  - Temperature
- ...to **data type**
  - Continuous variation
  - Warm, hot, cold
  - Burned vs not burned

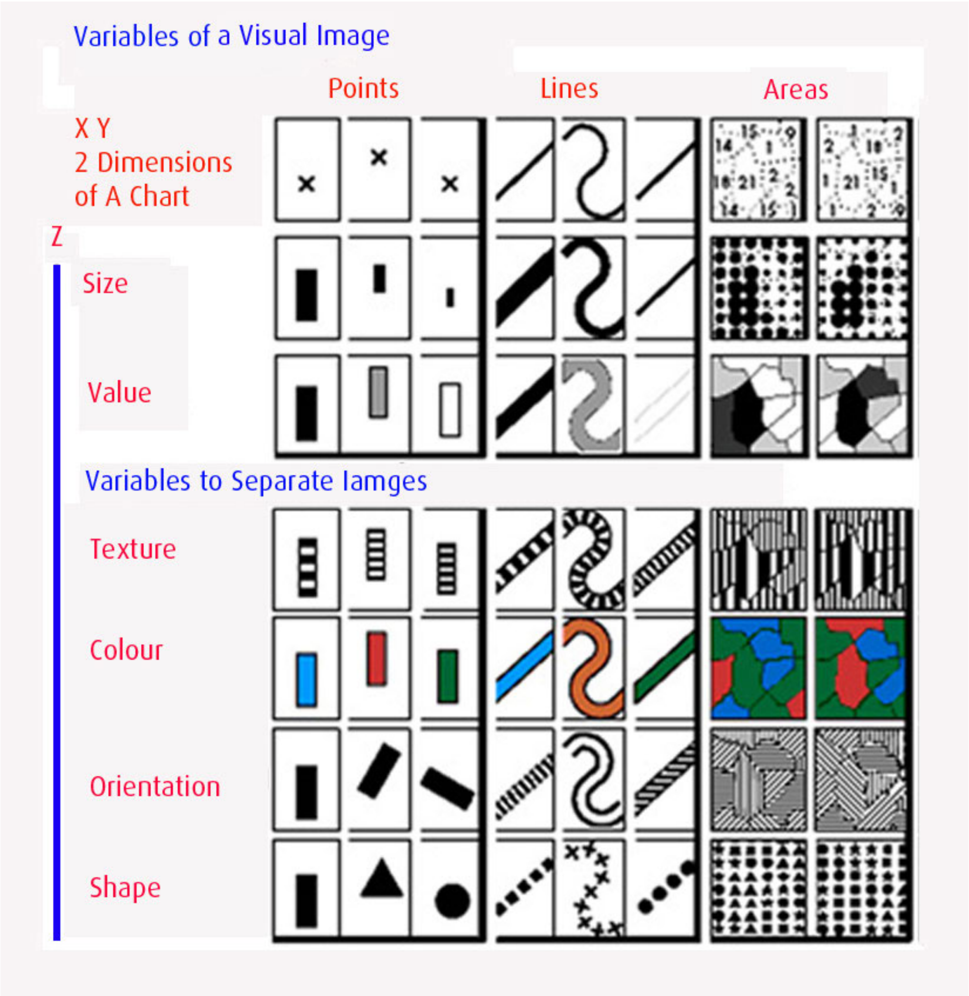
Visual Variables

- Jacques Bertin (1918-2010), cartographer
- Theoretical principles of visual encodings
- Semiology of Graphics (1967)



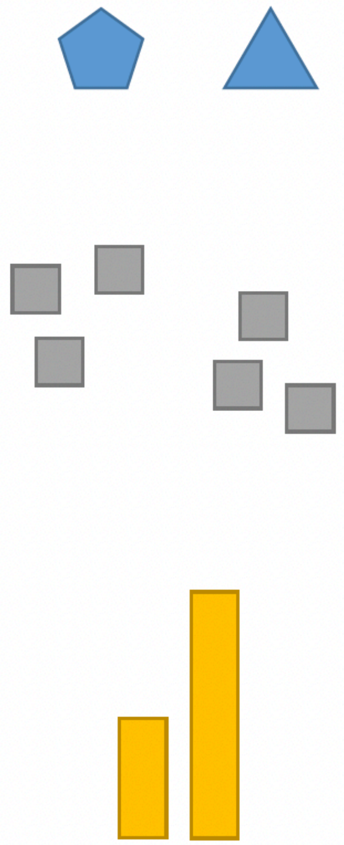
Bertin's Visual Variables

This is a visual summary of Bertin's classification of visual variables

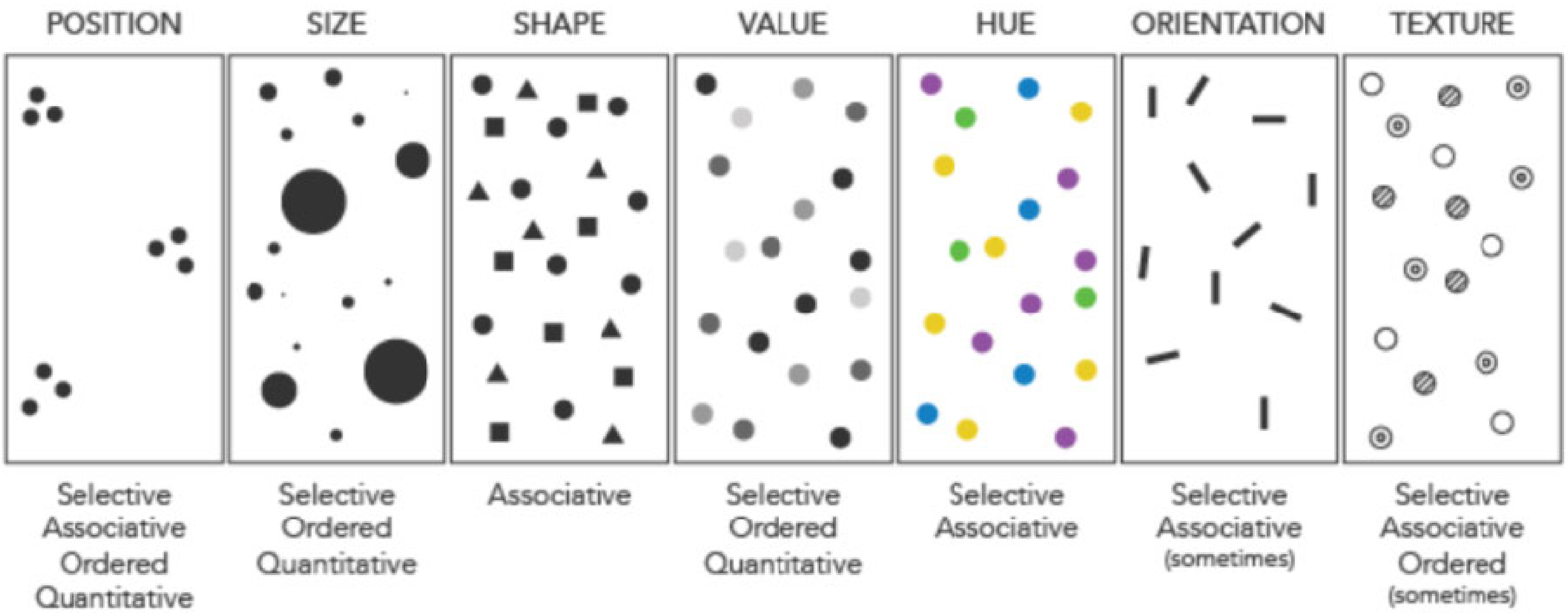


Characteristics of Visual Variables

- **Selective**
  - May I distinguish a symbol from the others?
- **Associative**
  - May I identify groups?
- **Quantitative**
  - May I quantify the difference of two values?
- **Order**
  - May I identify an ordering?

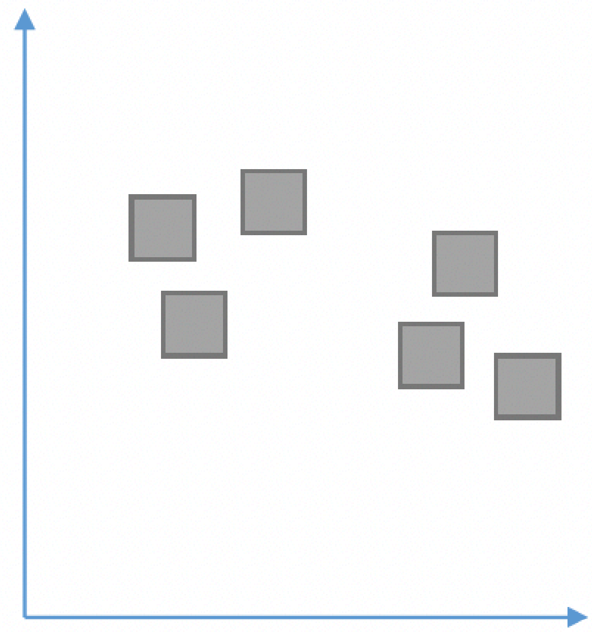


Characteristics of Visual Variables



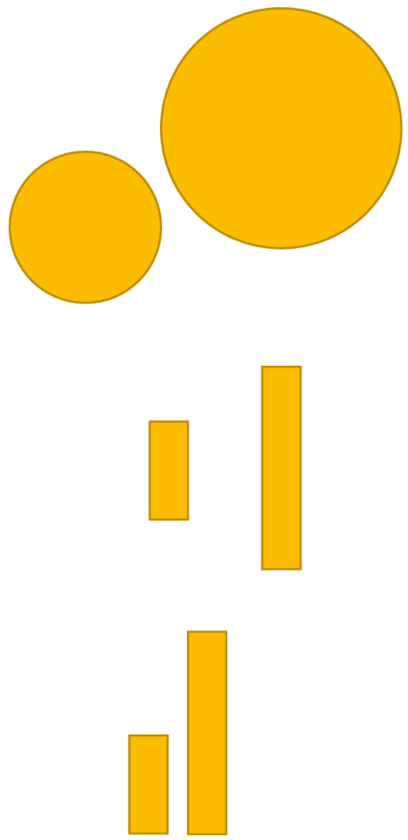
VV: Position

- Strongest visual variable
- Compatible for all data types
- **Cons**:
  - Not always applicable (e.g. nD data)
  - Cluttering



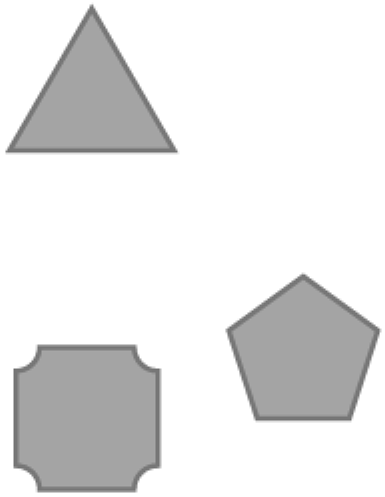
VV: Size and Length

- Easy to compare dimensions
- Grouping
- Estimate differences
  - Quantitative encoding
  - Changes in lengths
  - Worse for change in area



VV: Shapes

- Strong for nominal encoding
- No ordering
- No grouping



VV: Value (Intensity)

- Quantitative representation (when size and length are used)
- Limited number of shades
- Support grouping

VV: Color (Tint)

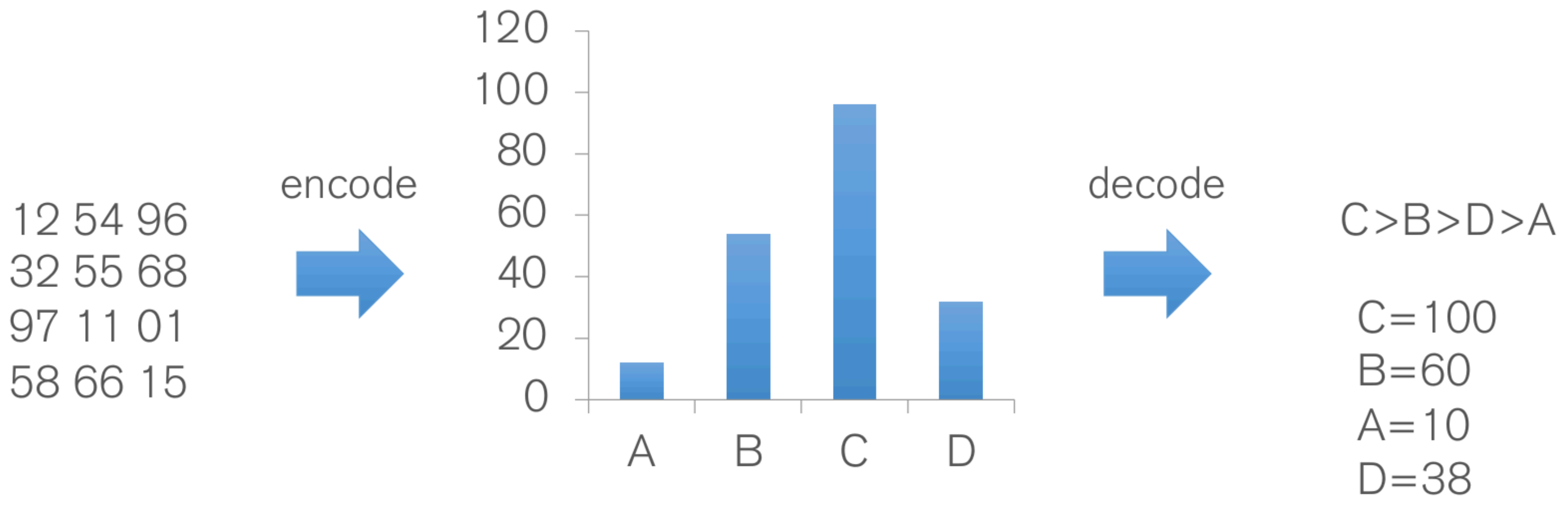
- Good for qualitative data
- Limited number of classes (!!!)
- Not good for quantitative data
- Be careful!!

Bertin Visual Variables

	Nominal	Ordinal	Quantitative
Position	√	√	√
Size	√	√	~
Value (intensity)	√	√	~
Texture	√	~	✗
Color	√	✗	✗
Orientation	√	✗	✗
Shape	√	✗	✗

Visual Encoding/Decoding

- A graph encodes a set of information as a set of graphical attributes
- The observer has to decode the graphical attributes to extract the original information



TAXONOMY OF VISUAL VARIABLES

Cleveland McGill [1984]

William S. Cleveland; Robert McGill, "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." 1984

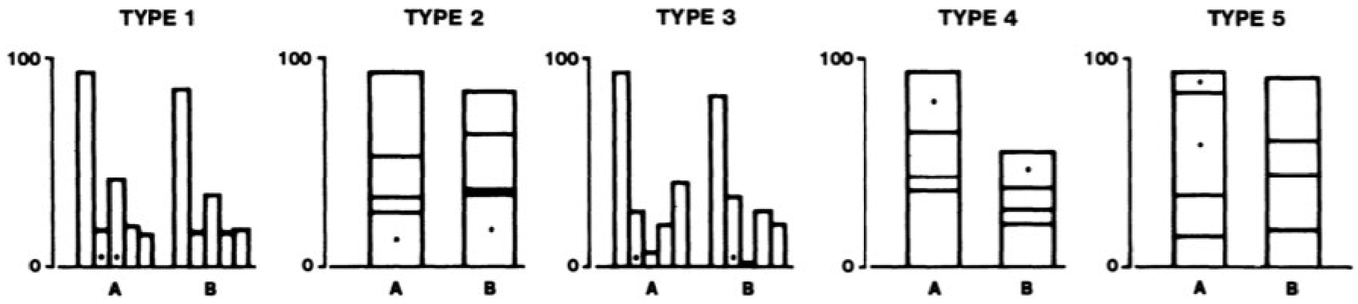


Figure 4. Graphs from position–length experiment.

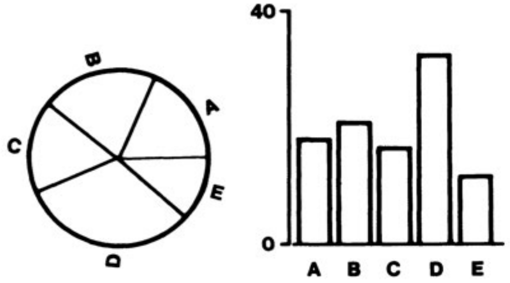


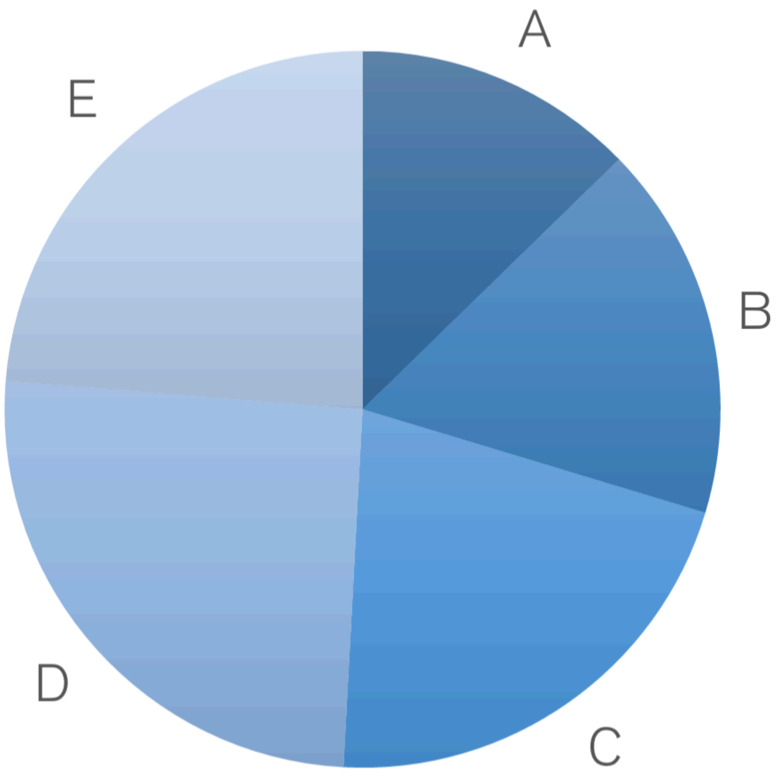
Figure 3. Graphs from position–angle experiment.

Cleveland & McGill: Graphical Encodings

- Angle
- Area
- Color Hue
- Color Saturation
- Density
- Length
- Position on a common scale
- Position on non-aligned scale
- Slope
- Volume

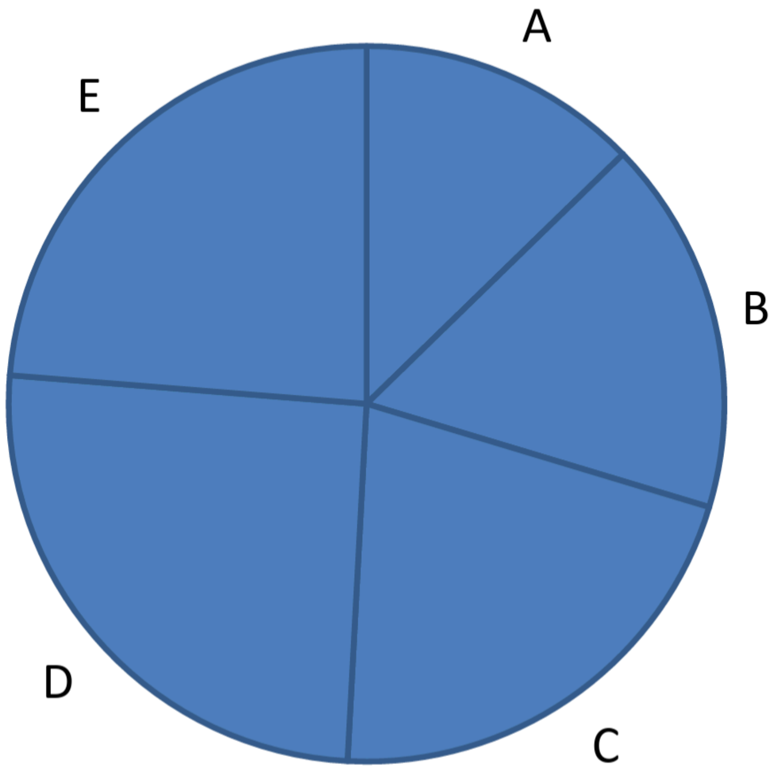
Angle Decoding

- It is difficult to compare angles
  - Underestimation of acute angles
  - Overestimation of obtuse angles
  - Easier if bisectors are aligned
- Area estimation helps



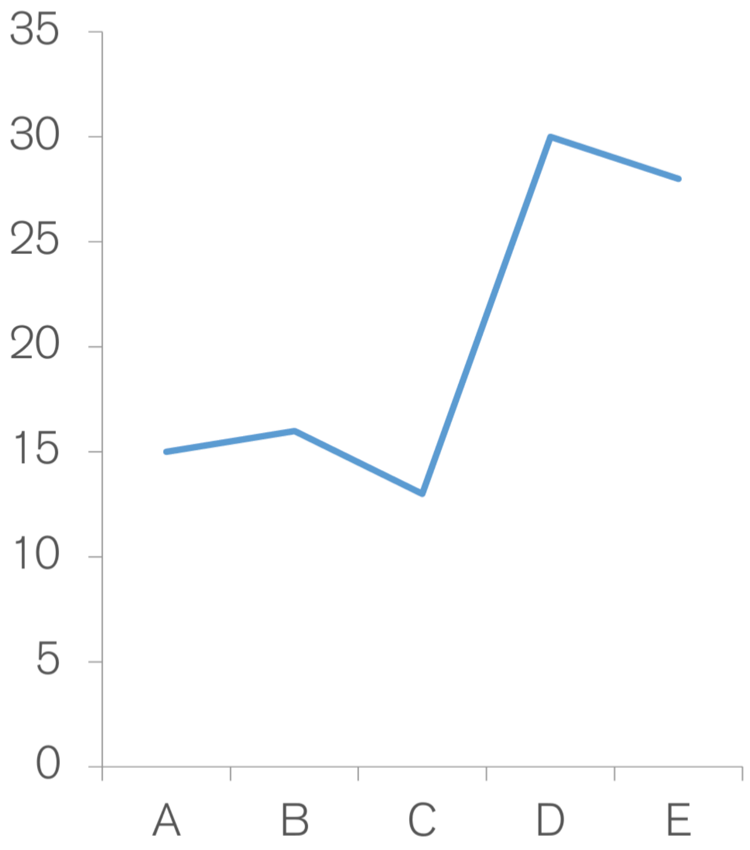
Angle Decoding

- It is difficult to compare angles
  - Underestimation of acute angles
  - Overestimation of obtuse angles
  - Easier if bisectors are aligned



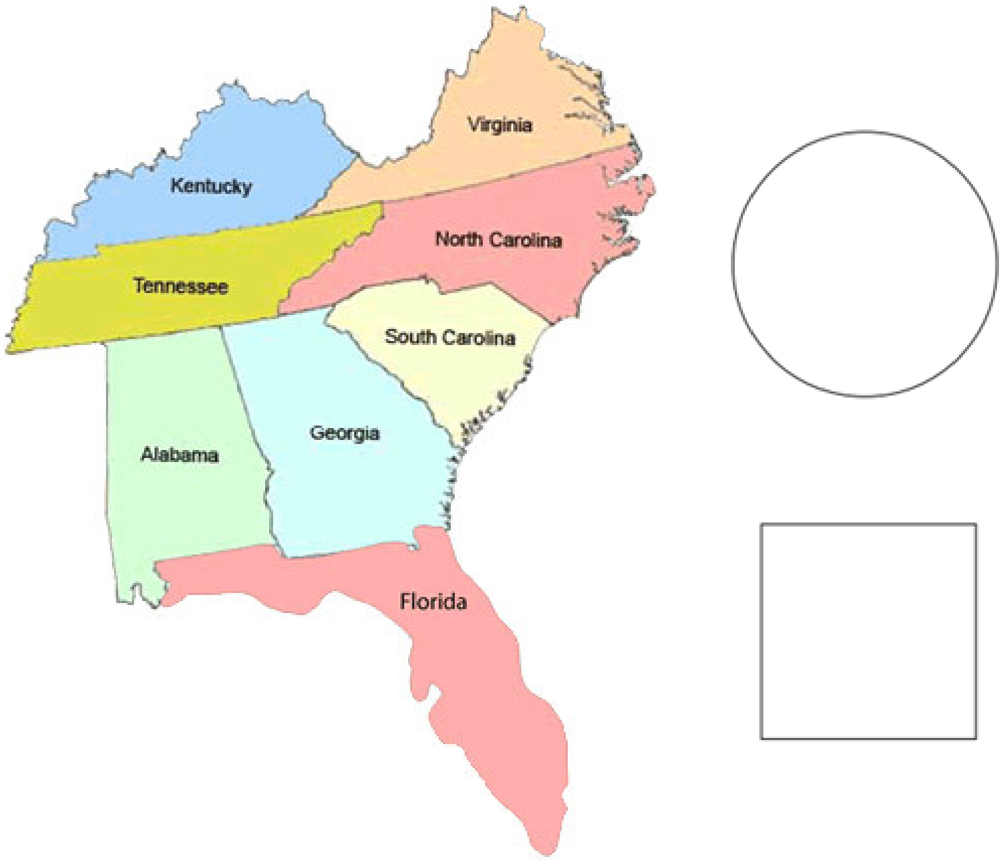
Slopes Decoding

- Same difficulties as angles
- Easier task since one branch is aligned with x-axis



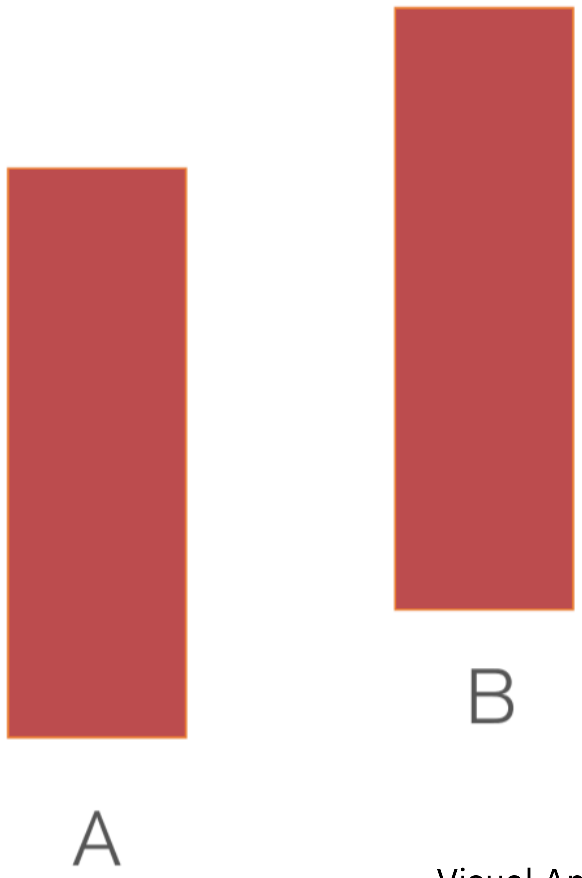
Area Decoding

- Area is not well decoded
- Different regular shapes
- Irregular shapes
- Context influences (thin area within compact thick area)



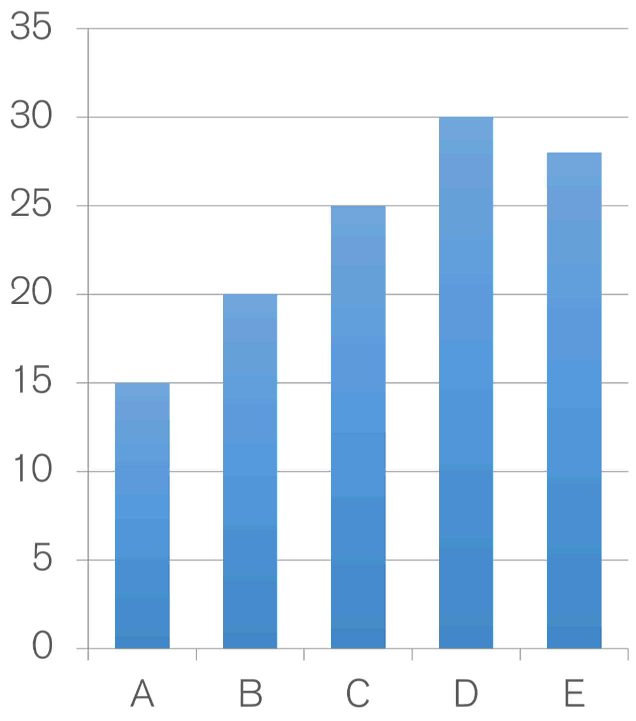
Length Decoding

- Straight forward to encode numerical values
- Difficulties with relative lengths



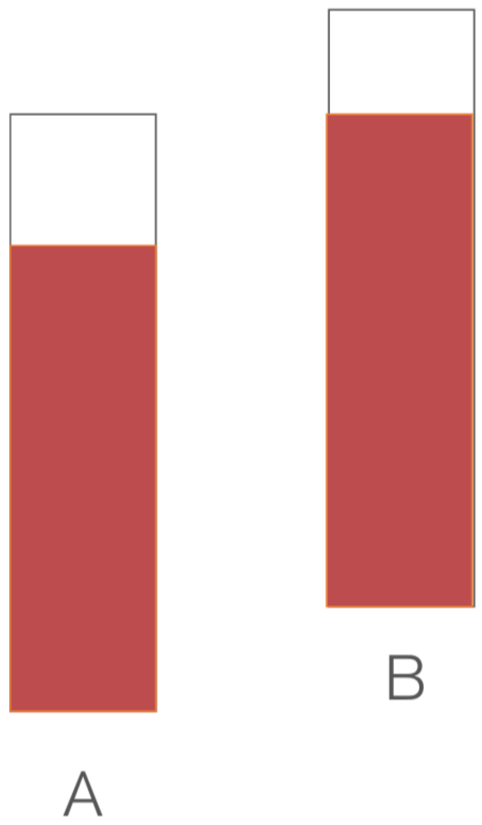
Position on a Common Scale

- Widely used in statistical charts



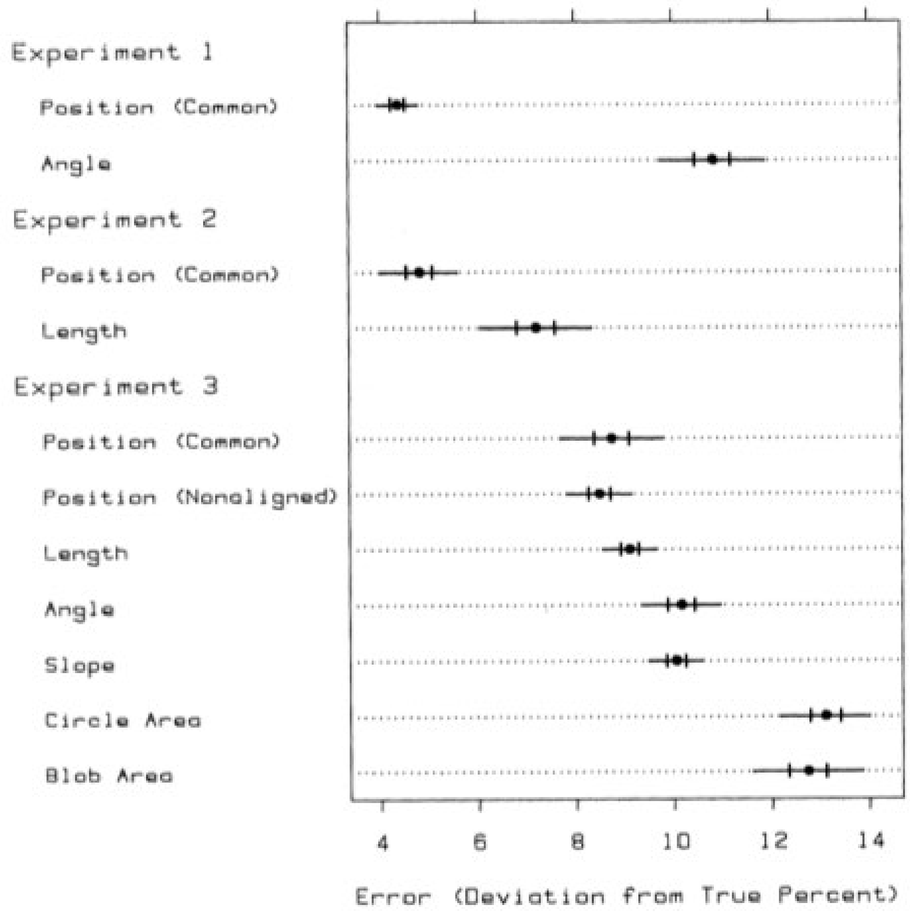
Position on Non-Aligned Scale

- Not as good as common scale
- Still acceptable

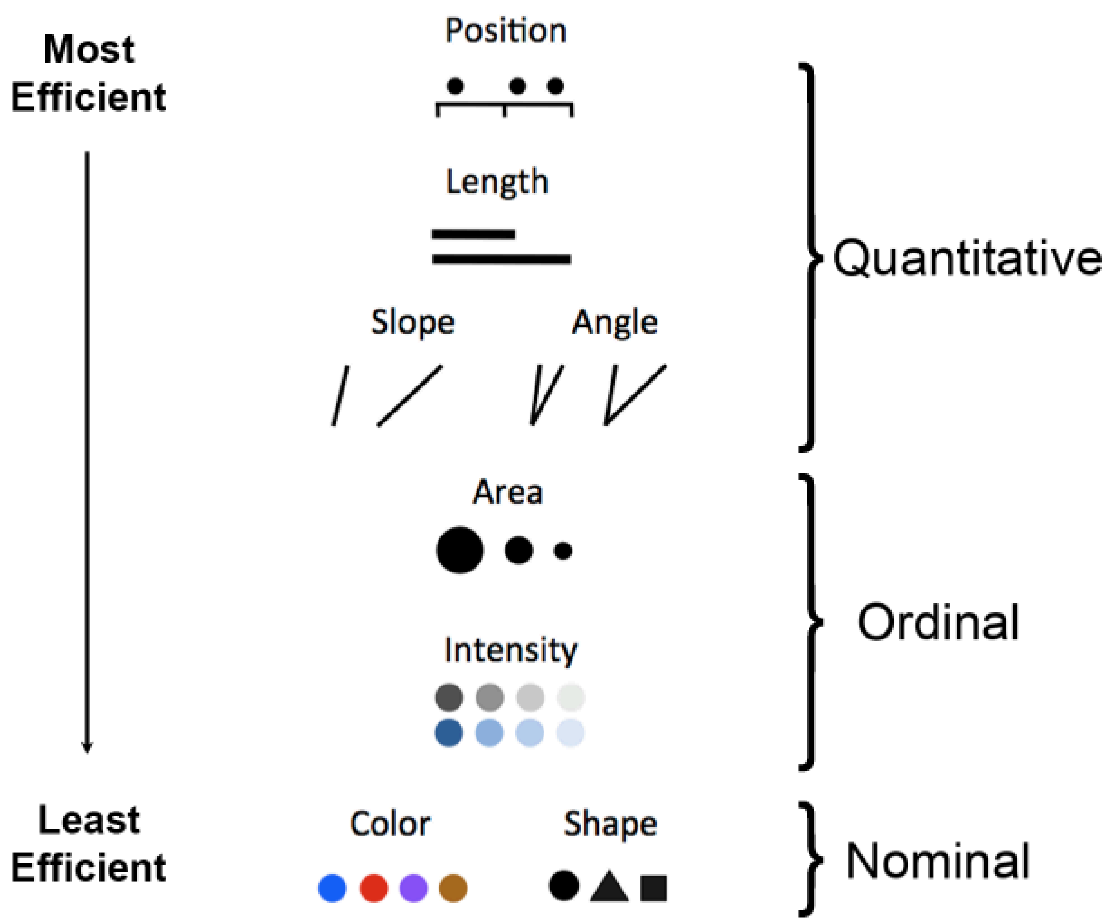


Designing Effective Visualizations

- If possible, use graphical encoding that is easily decoded
- Graphical Attributes ordered (Cleveland & McGill):
   i. Position along a common scale
   ii. Position on non-aligned scales
   iii. Length
   iv. Angle and Slope
   v. Area
   vi. Volume, density, color saturation
   vii. Color Hue



Most Efficient to Least Efficient



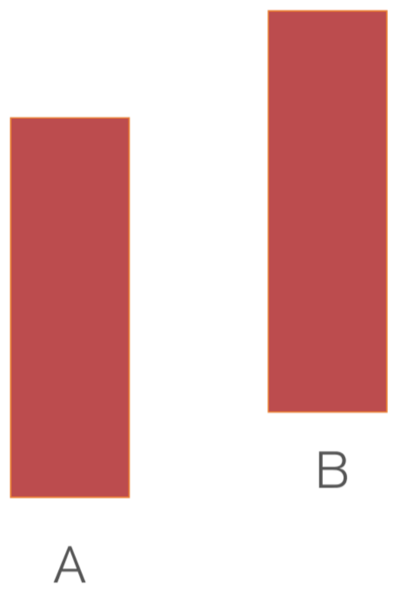
PERCEPTION LAWS

Weber's Law

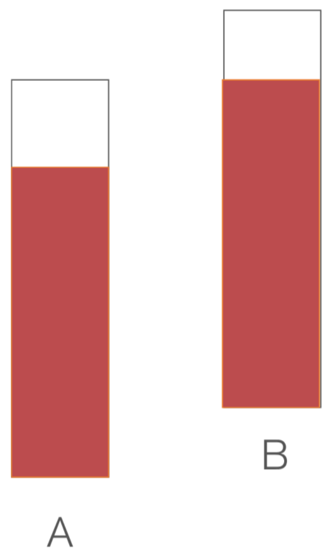
- Just-noticeable difference between two stimuli is proportional to their magnitudes
- Case study on length:
  - Given two lines with lengths x and $x+w$
  - If w is small, it is difficult to notice difference between the two lines
  - If w is larger, it is easier to catch the difference
- How large should w be?
  - The probability of detecting the change is proportional to the relative value w/x

Weber's Law

- Given values (90, 92)
- Detect with probability of 2/90



- Given values (90, 92)
- Detect with probability of 2/10

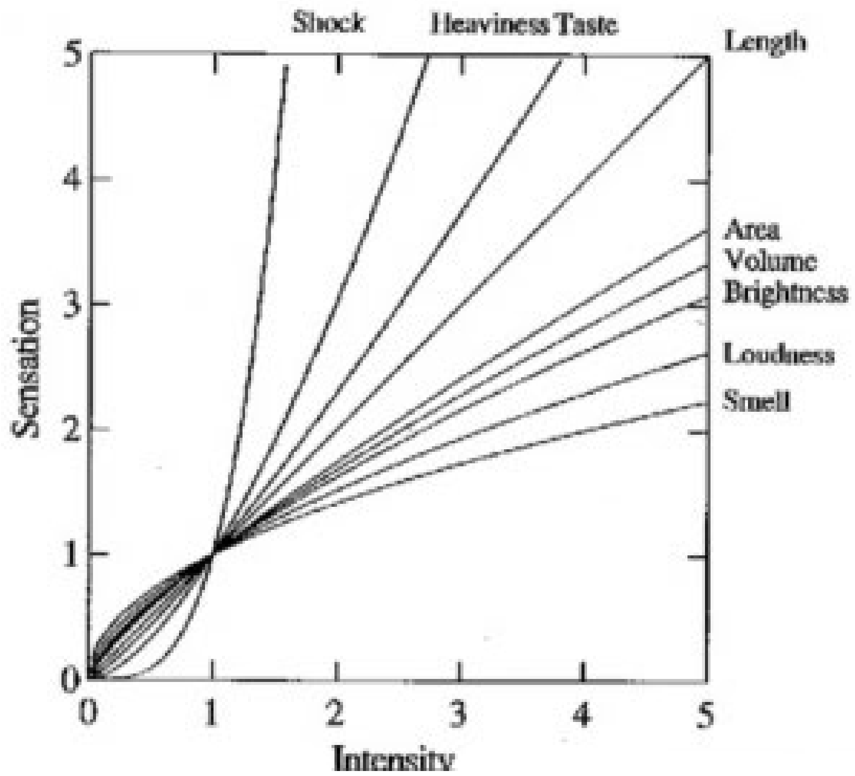


Stevens' Law

- Model the relation between a stimulus and its perceived intensity
- Given a stimulus x encoded with a visual attribute
- An observer decodes a perceived value $p(x)$
- Stevens' law states that:
  - $p(x) = kx^{\beta}$
  - where k is constant and
  - β is a constant that depends on the nature of stimulus

Stevens' Law

- Better effectiveness when $p(x) = kx^{\beta}$ is linear
- Linearity depends only on β
- Different visual encodings yields typical ranges for β:
  - Lengths: 0.9 – 1.1
  - Area: 0.6 – 0.9
  - Volume: 0.5 – 0.8



Stevens' Law graph showing over/underestimation of values

Weber and Stevens' Laws

- Given two values x_1 and x_2
- Let the perceived values be $p(x_1)$ and $p(x_2)$

$$\frac{p(x_1)}{p(x_2)} = \left(\frac{x_1}{x_2}\right)^{\beta}$$

Weber and Stevens' Laws: Areas

- For areas β=0.7
- Let x_1=2 and x_2=1
- The perceived difference will be:

$$\frac{p(2)}{p(1)} = \left(\frac{2}{1}\right)^{0.7} = 1.6245$$

- For areas β=0.7
- Let x_1=0.5 and x_2=1
- The perceived difference will be:

$$\frac{p(\frac{1}{2})}{p(1)} = \left(\frac{1/2}{1}\right)^{0.7} = 0.6155$$

Weber and Stevens' Laws: Areas vs Lengths

- For areas β=0.7
- Let $x_2=x_1+w$
- The perceived difference will be:

$$\left(\frac{x+w}{x}\right)^{0.7} \approx 1 + \frac{0.7w}{x}$$

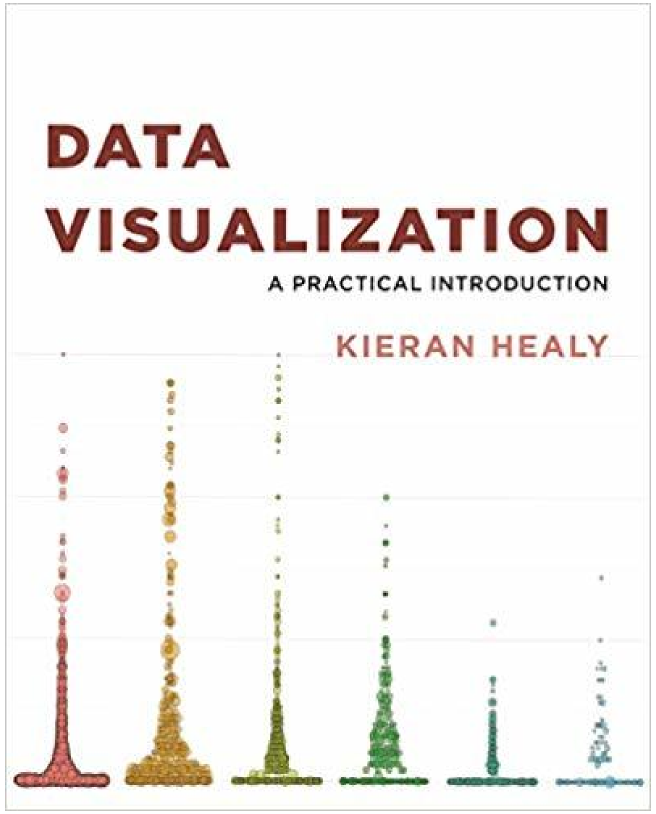
- For lengths β=1
- Let $x_2=x_1+w$
- The perceived difference will be:

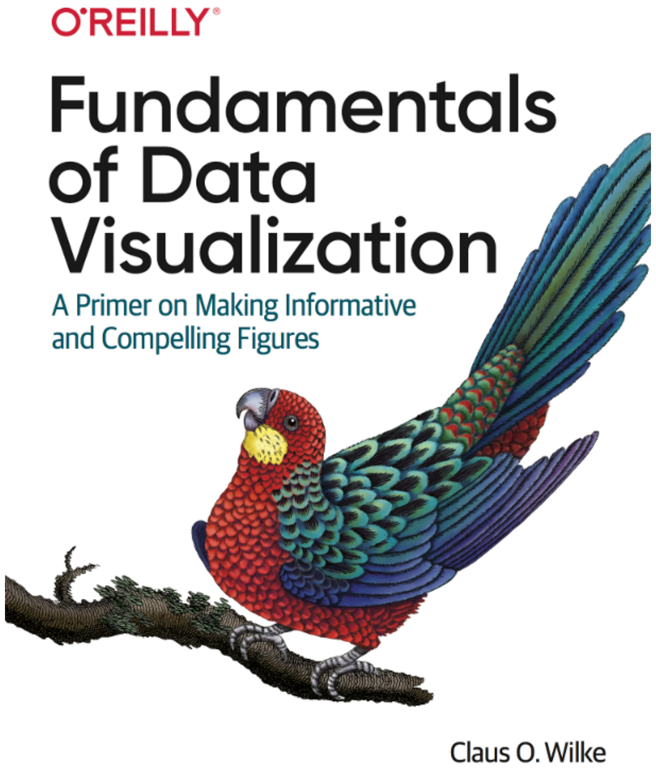
$$\left(\frac{x+w}{x}\right)^{1} = 1 + \frac{w}{x}$$

Takeaway Messages

- Data type for entities and relationships
- Visual variables for representation
- Mapping of types to VVs
- Some VVs are more appropriate for specific data types

Suggested Readings

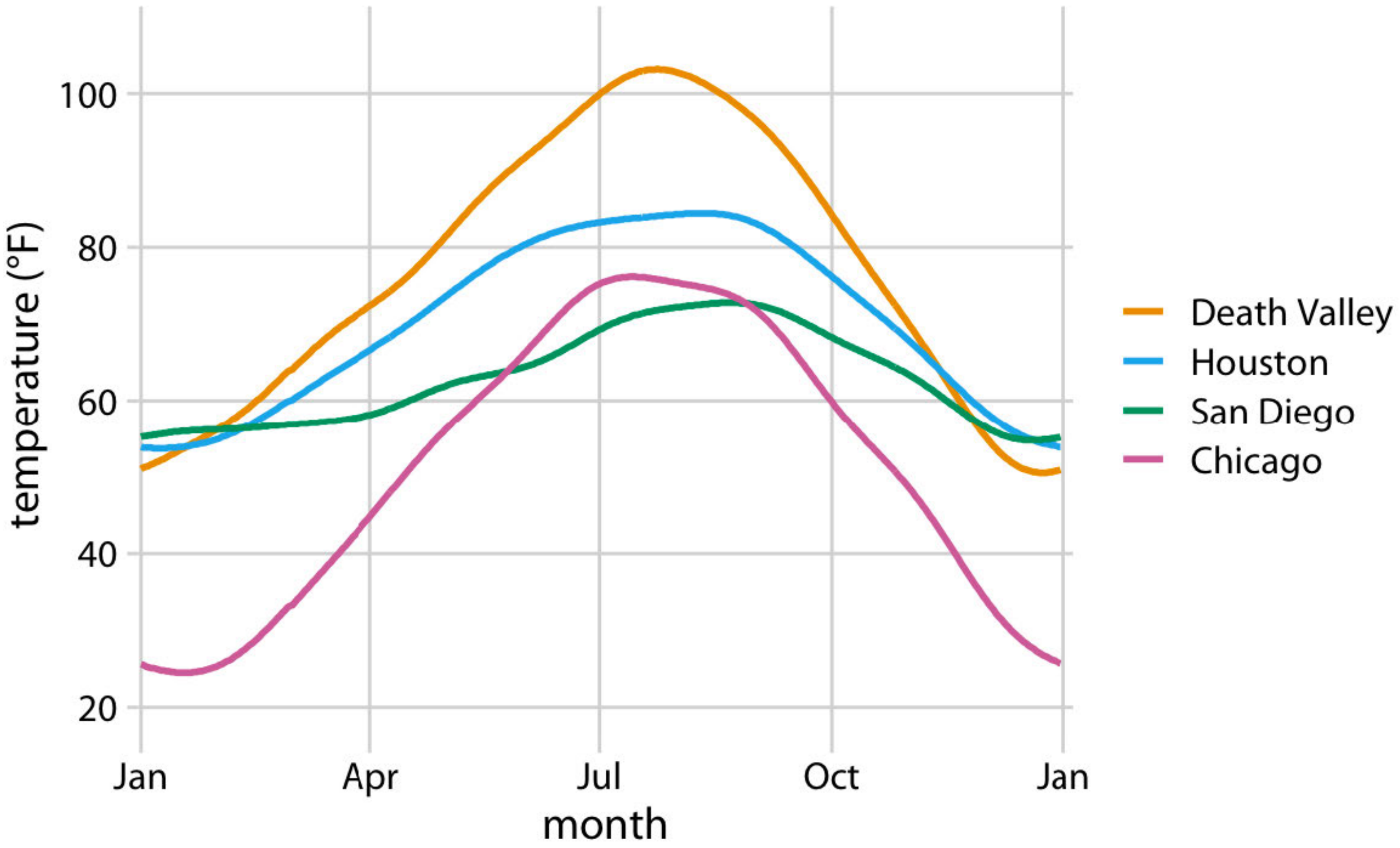




Example

Month	Day	Location	Station ID	Temperature
Jan	1	Chicago	USW00014819	25.6
Jan	1	San Diego	USW00093107	55.2
Jan	1	Houston	USW00012918	53.9
Jan	1	Death Valley	USC00042319	51.0
Jan	2	Chicago	USW00014819	25.5
Jan	2	San Diego	USW00093107	55.3
Jan	2	Houston	USW00012918	53.8
Jan	2	Death Valley	USC00042319	51.2
Jan	3	Chicago	USW00014819	25.3
Jan	3	San Diego	USW00093107	55.3
Jan	3	Death Valley	USC00042319	51.3
Jan	3	Houston	USW00012918	53.8

Visual Solution (1)



Visual Solution (2)