Computation and Visualization for Analytics

Spring 2021

Week 4.2

Visual Encoding

- Marks and Channels (*representation*)
- Channel Expressiveness and Effectiveness (*presentation*)
- Relative vs. Absolute Judgements (*presentation*)

Why Visualize?

- Explore
- Explain
  - ⇢ Memory limits
  - ⇢ Visual limits
  - ⇢ Choice limits

Explanatory Visualization

- Representation
  Involves making decisions about how you are going to portray your data visually so that the subject understanding it offers can be made accessible to your audience. It is all about charts and the act of selecting the right chart to show the features of your data that you think are most relevant. (*Andy Kirk*)

- Presentation
  Involves effective presentation to audience taking into consideration limits of visual perception

Representation (Marks)

Points

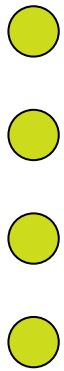
Lines

Areas

Munzner, T. (2014). *Visualization analysis and design*. CRC press.

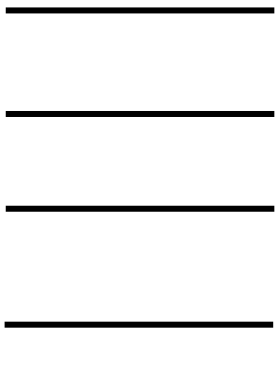
Representation (Marks-points)

Age
10
20
30
40



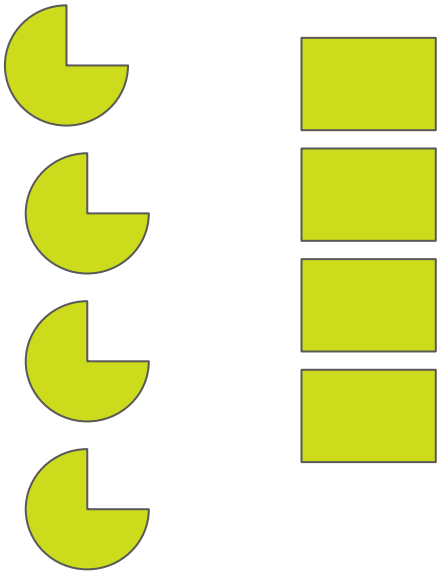
Representation (Marks-lines)

State
MA
NH
NY
GA



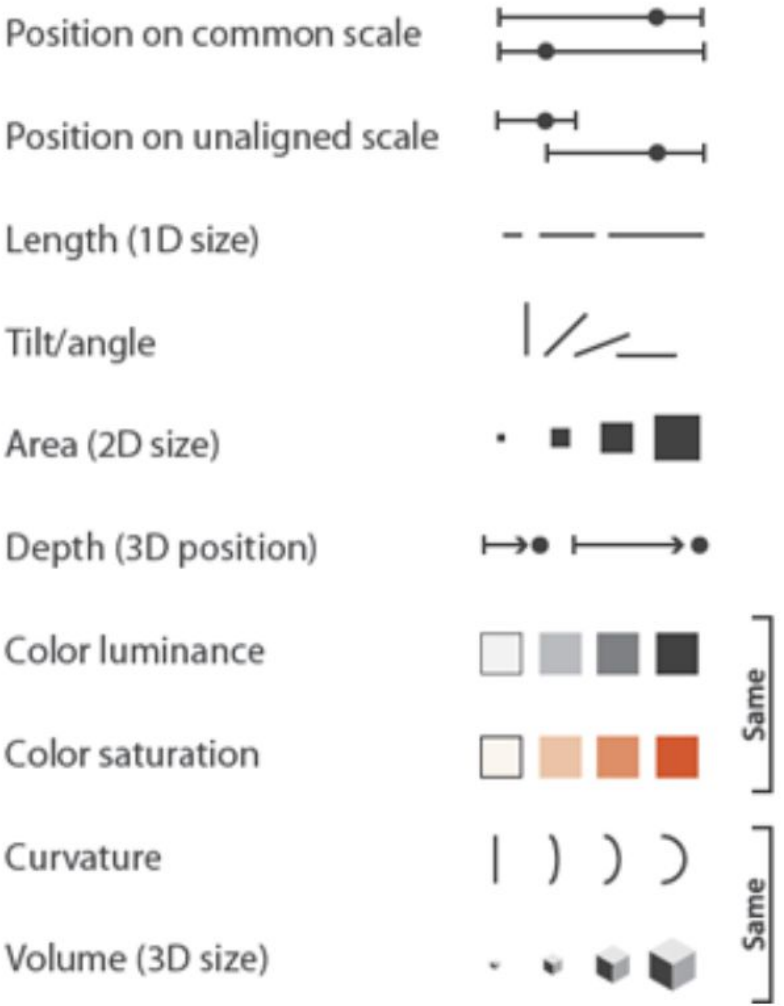
Representation (Marks-area, shapes)

State
MA
NH
NY
GA



Representation (Channels)

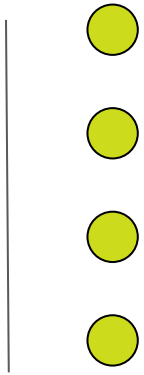
Position
Shape
Tilt
Color
Size



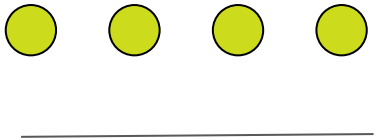
Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Representation (Mark-point, Channel-position)

Age
10
20
30
40



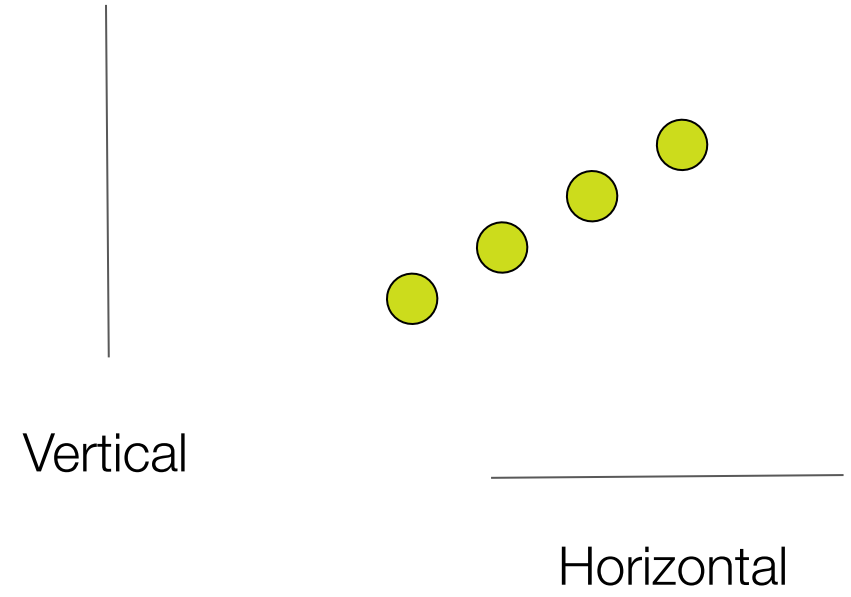
Vertical



Horizontal

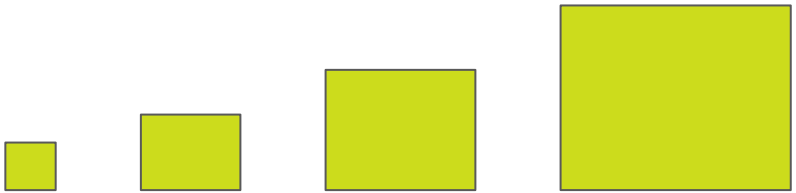
Representation (Mark-point, Channel-position)

Age	Height
10	20
20	40
30	60
40	80



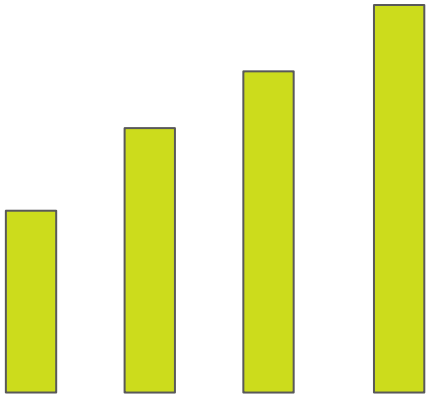
Representation (Mark-area, Channel-size)

Age
10
20
30
40



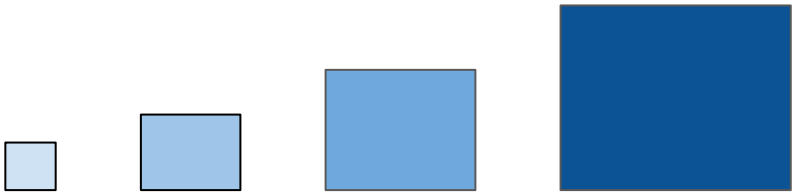
Representation (Mark-line, Channel-size)

Age
10
20
30
40



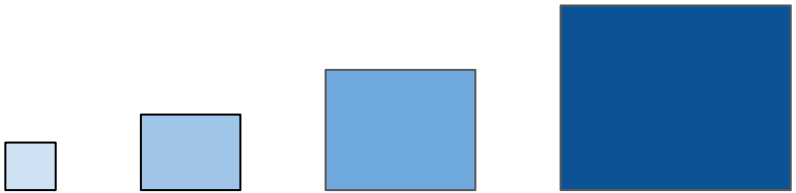
Representation (Mark-area, Channel-color saturation)

Age
10
20
30
40



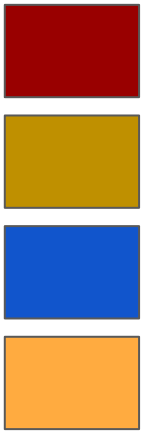
Representation (Mark-area, Channel-color saturation)

Income
10000
20000
30000
40000



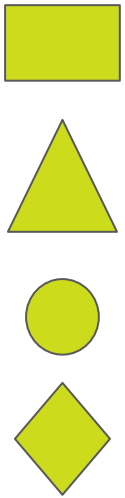
Representation (Mark-area, Channel-color hue)

State
MA
NH
NY
GA



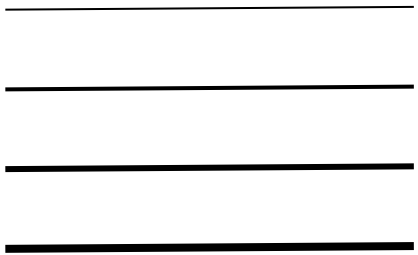
Representation (Mark-area, Channel-shape)

State
MA
NH
NY
GA



Representation (Mark-line, Channel-thickness)

Source	Destination	#Flights
BOS	LGA	1000
BOS	DTW	2000
BOS	ATL	3000
BOS	ORD	4000



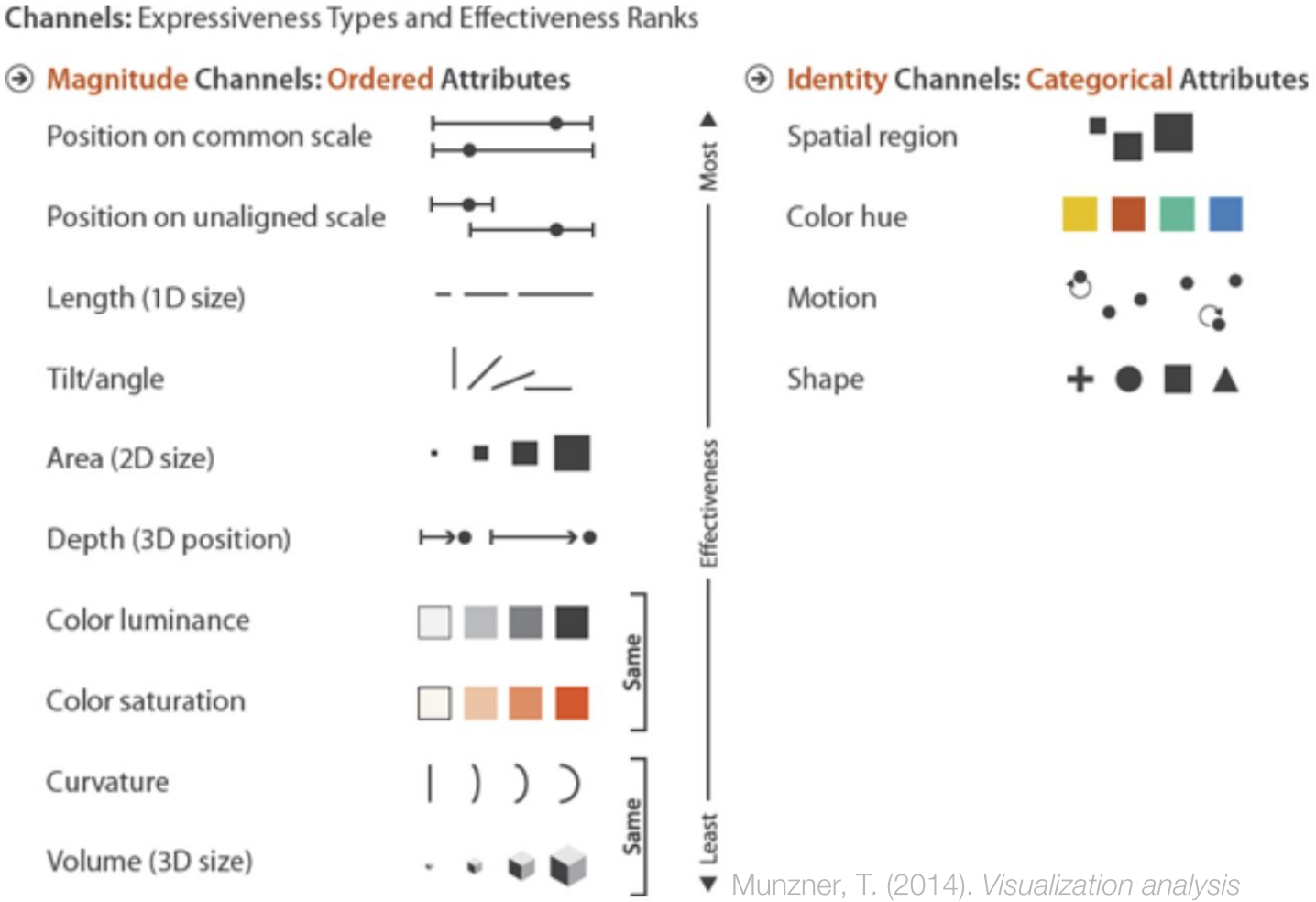
Marks and Channels

Type of variable	Examples	Appropriate scale	Description
quantitative/numerical continuous	1.3, 5.7, 83, $1.5x10^{-2}$	continuous	Arbitrary numerical values. These can be integers, rational numbers, or real numbers.
quantitative/numerical discrete	1, 2, 3, 4	discrete	Numbers in discrete units. These are most commonly but not necessarily integers. For example, the numbers 0.5, 1.0, 1.5 could also be treated as discrete if intermediate values cannot exist in the given dataset.
qualitative/categorical unordered	dog, cat, fish	discrete	Categories without order. These are discrete and unique categories that have no inherent order. These variables are also called *factors*.
qualitative/categorical ordered	good, fair, poor	discrete	Categories with order. These are discrete and unique categories with an order. For example, “fair” always lies between “good” and “poor”. These variables are also called *ordered factors*.
date or time	Jan. 5 2018, 8:03am	continuous or discrete	Specific days and/or times. Also generic dates, such as July 4 or Dec. 25 (without year).
text	The quick brown fox jumps over the lazy dog.	none, or discrete	Free-form text. Can be treated as categorical if needed.

https://serialmentor.com/dataviz/aesthetic-mapping.html

Continuous Variables

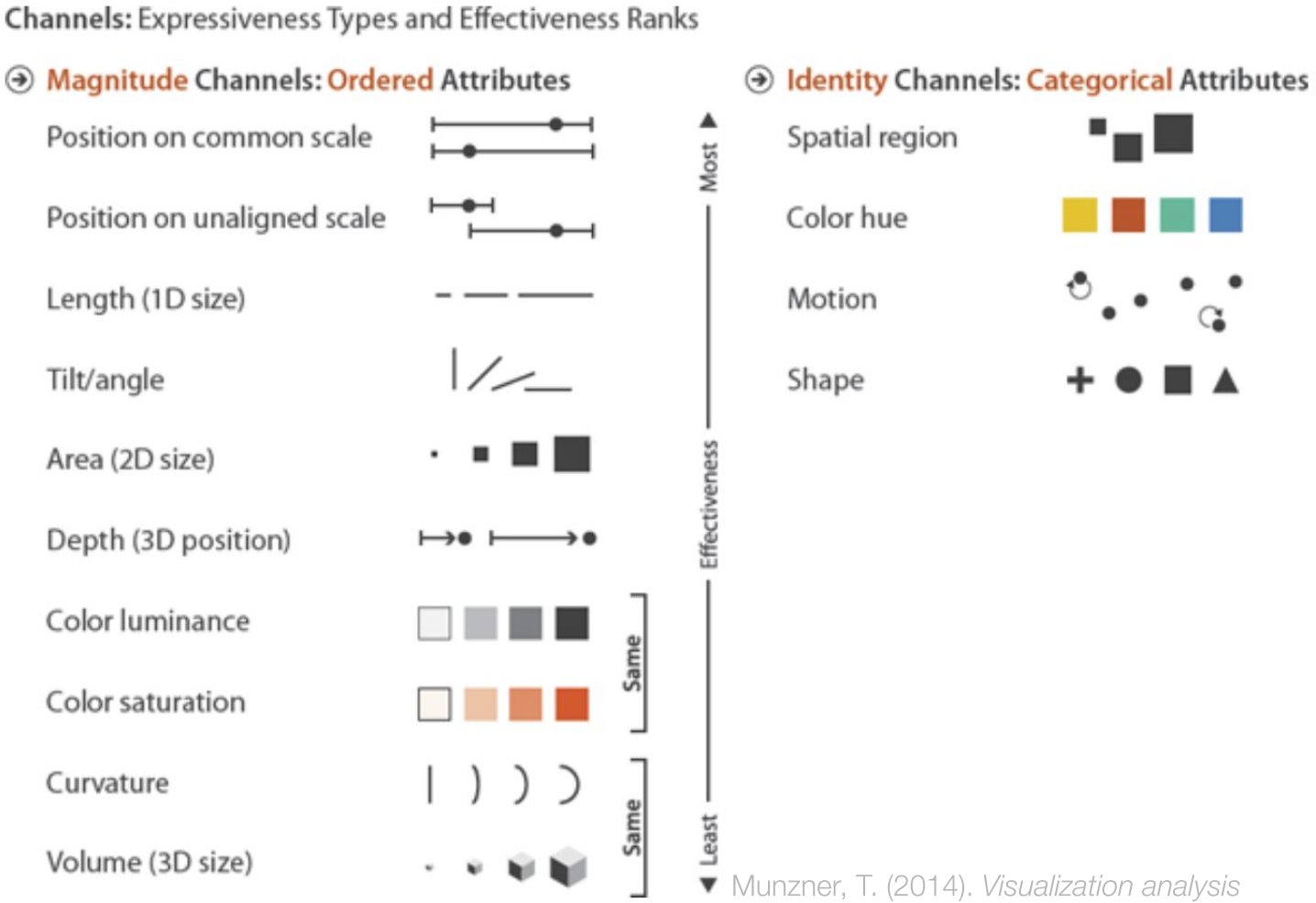
#Flights
1000
2000
3000
4000



Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Categorical Variables

State
MA
NH
NY
GA

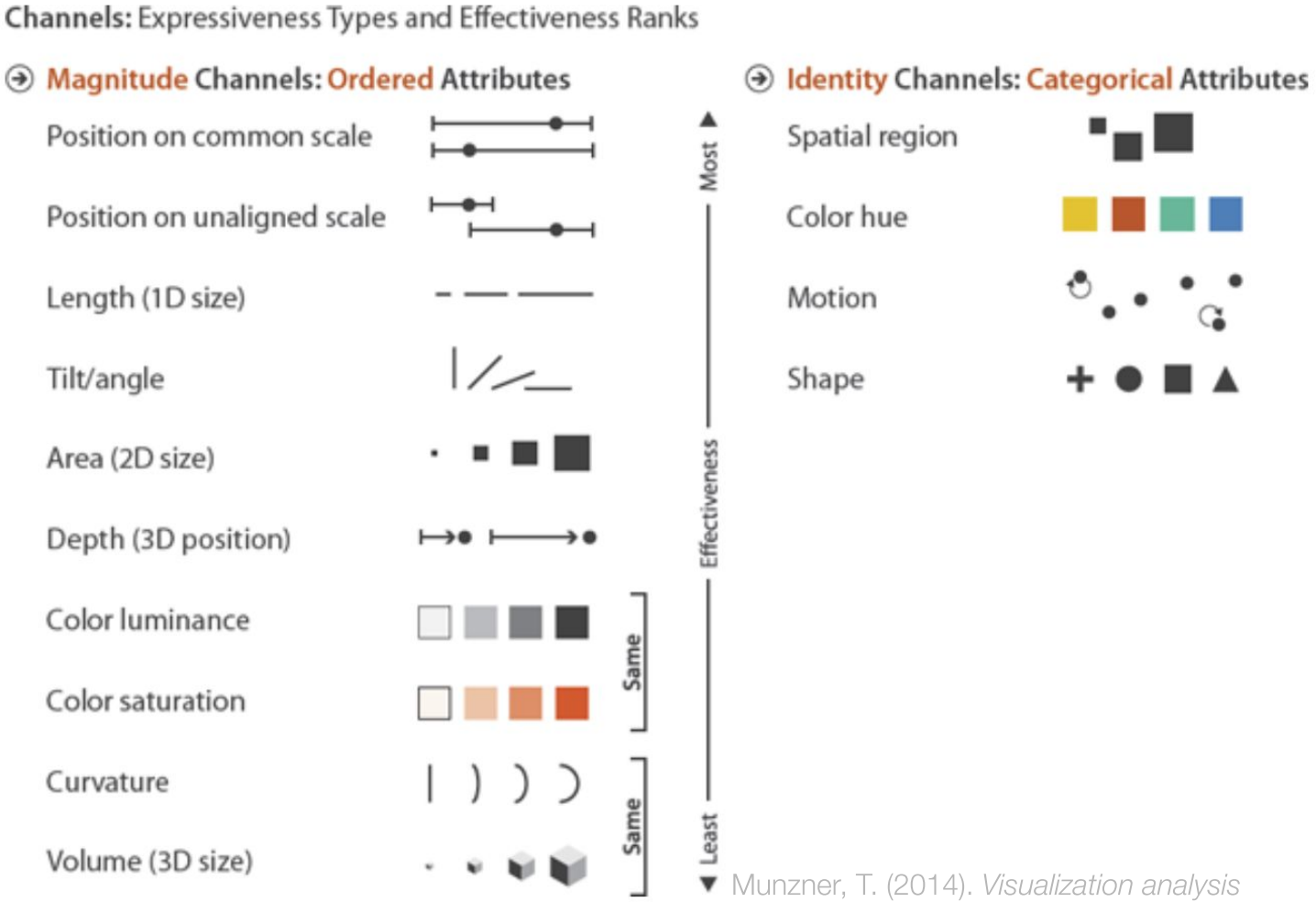


Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Categorical Variables

Income
Low
Med
High

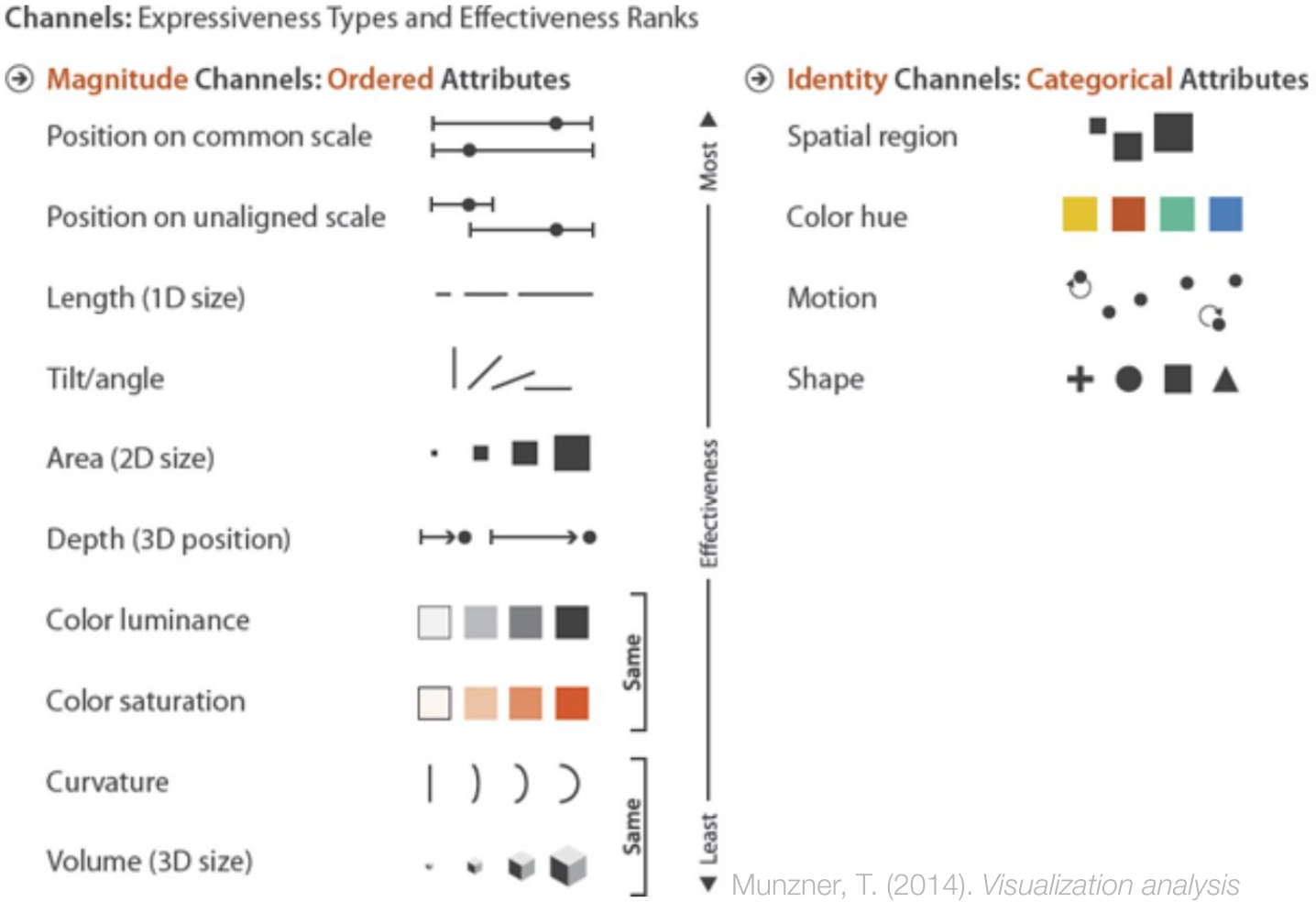
Performance
Poor
Satisfactory
Good



Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Date Variables

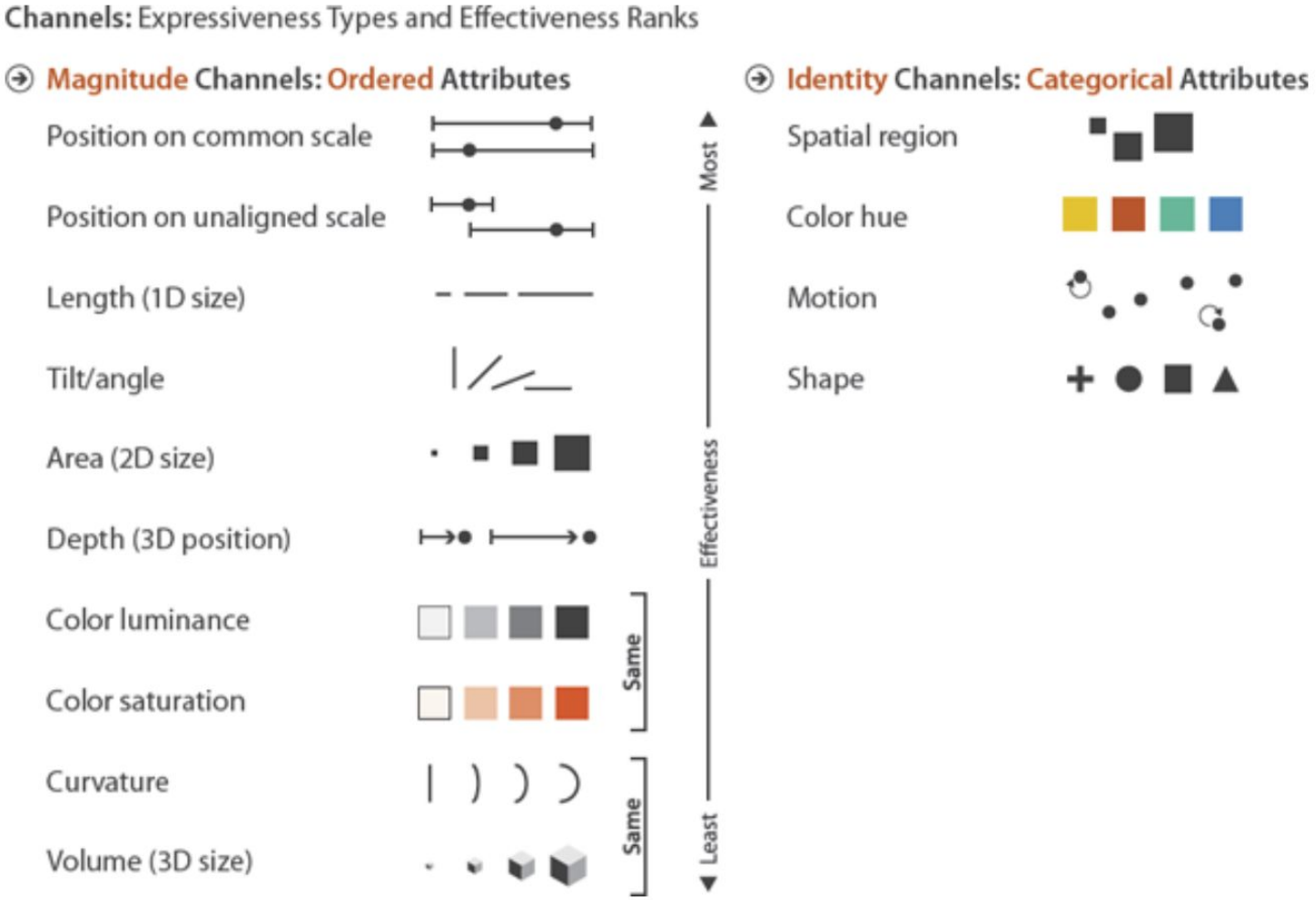
Date
1/1/2021
1/2/2021
1/3/2021
1/4/2021



Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Date Variables

Day
Sunday
Monday
Tuesday
Wednesday



Multivariate Encoding

Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Expressiveness and Effectiveness

- The **expressiveness** principle dictates that the visual encoding should express all of, and only, the information in the dataset attributes.*

- The **effectiveness** principle dictates that the importance of the attribute should match the salience of the channel; that is, its noticeability.*

- In other words, the most important attributes should be encoded with the most effective channels in order to be most noticeable, and then decreasingly important attributes can be matched with less effective channels.*

* Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Accuracy)

- **Accuracy:** how close is human perceptual judgement to some objective measurement of the stimulus?*

- We perceive different visual channels with different levels of accuracy; they are not all equally distinguishable.*

* Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Accuracy)

- **Psychophysics**, the subfield of psychology devoted to the systematic measurement of general human perception

- “Our responses to the sensory experience of magnitude are characterizable by power laws, where the exponent depends on the exact sensory modality: most stimuli are magnified or compressed, with few remaining unchanged” - (*Tamara Munzner*)

* Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Accuracy)

- **Steven's Power law** $\psi(I) = kI^a$

where ***I*** is the intensity or strength of the stimulus in physical units (energy, weight, pressure, mixture proportions, etc.), **ψ(*I*)** is the magnitude of the sensation evoked by the stimulus, ***a*** is an exponent that depends on the type of stimulation or sensory modality, and ***k*** is a proportionality constant that depends on the units used

* Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Accuracy)

- **Steven's Power law**

$$\psi(I) = kI^a$$

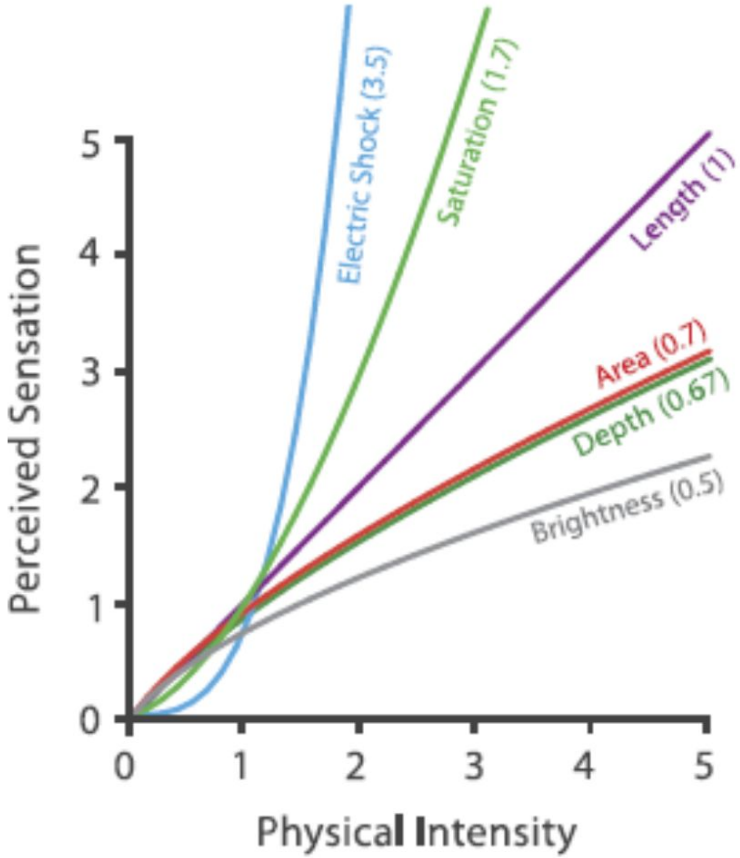
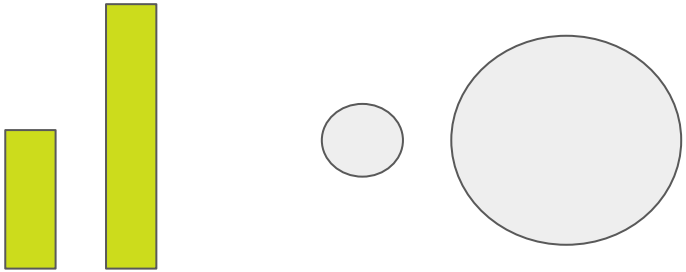


Image from Munzner, T. (2014). *Visualization analysis and design*. CRC press.
https://en.wikipedia.org/wiki/Stevens%27s_power_law

Channel Effectiveness (Accuracy)

- **Steven's Power law**

$$\psi(I) = kI^a$$



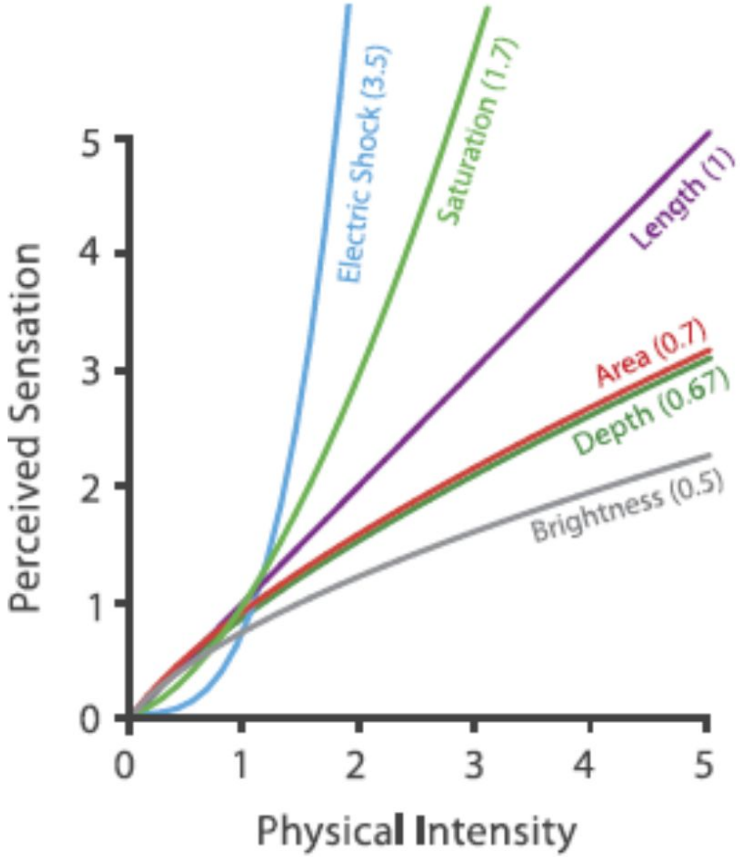
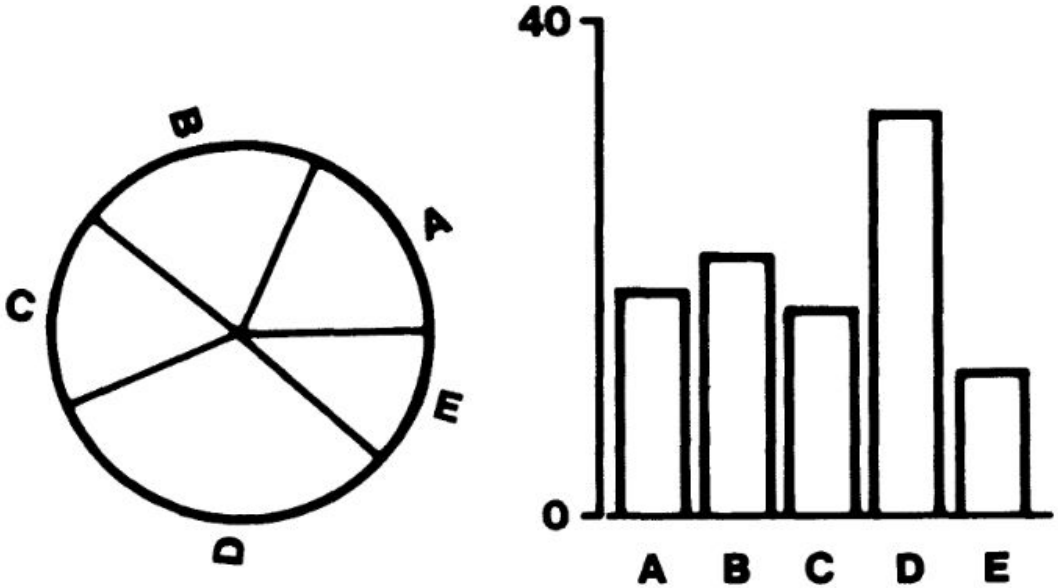


Image from Munzner, T. (2014). *Visualization analysis and design*. CRC press.
https://en.wikipedia.org/wiki/Stevens%27s_power_law

Channel Effectiveness (Accuracy)

Cleveland McGill: Ranking Accuracy, 1984



https://info.sice.indiana.edu/~katy/S637-S11/cleveland84.pdf

Channel Effectiveness (Accuracy)

Cleveland McGill: Ranking Accuracy, 1984

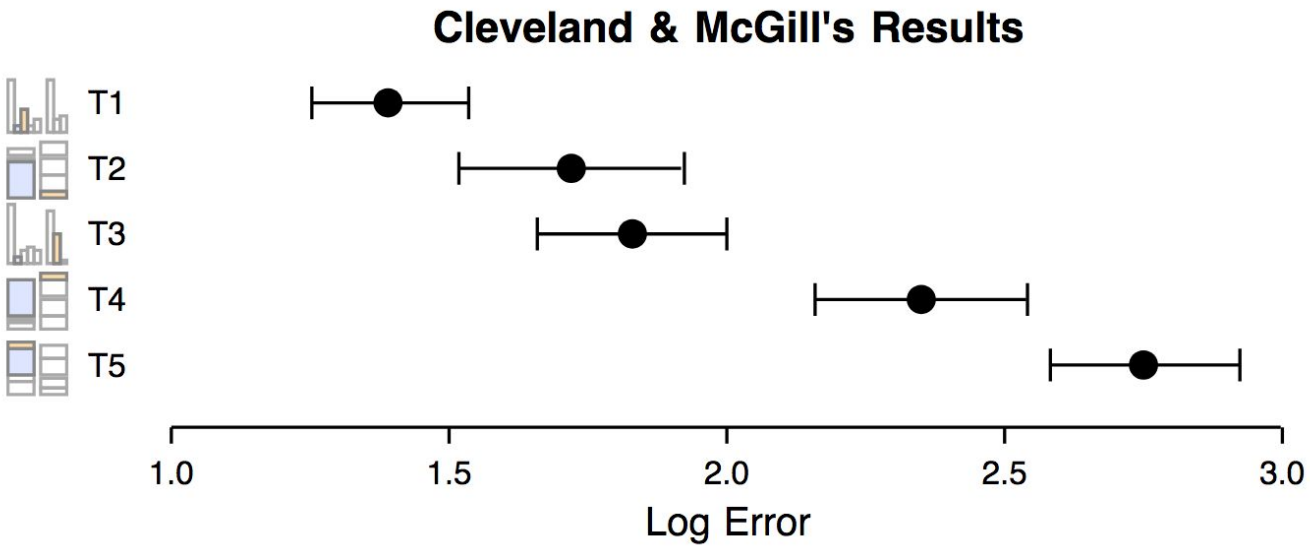


Image Source: https://www.cs.ubc.ca/~tmm/courses/533-11/slides/eval.jessica.pdf

Channel Effectiveness (Accuracy)

Heer and Bostock 2010

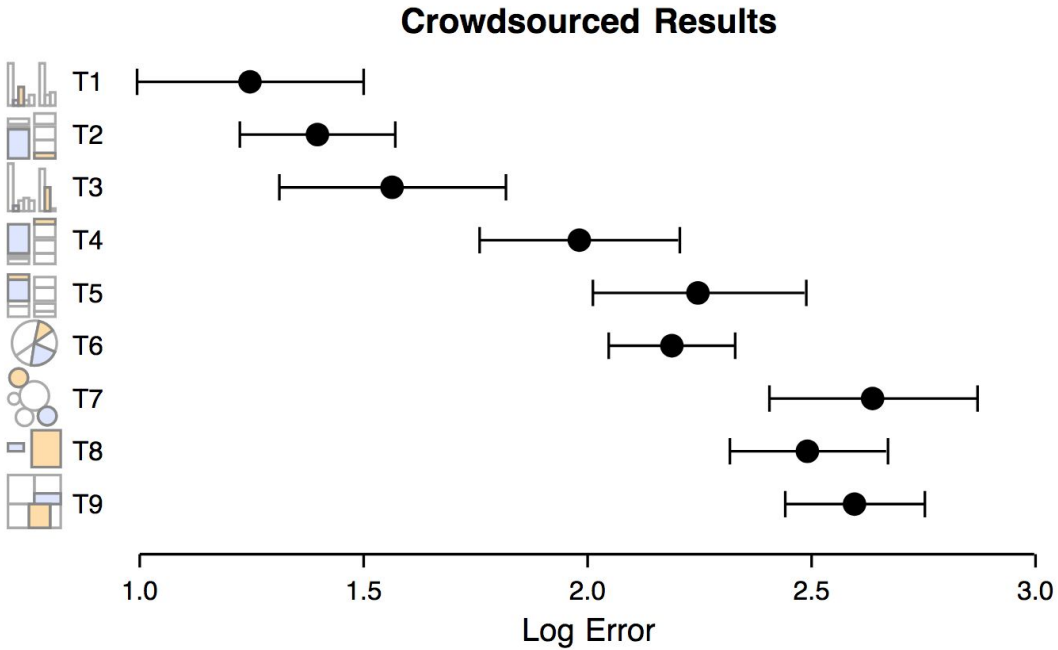


Image Source: https://www.cs.ubc.ca/~tmm/courses/533-11/slides/eval.jessica.pdf
Conference Article: https://dl.acm.org/doi/10.1145/1753326.1753357

Channel Effectiveness (Discriminability)

- If you encode data using a particular visual channel, are the differences between items perceptible to the human as intended?*

- The characterization of visual channel thus should quantify the number of bins that are available for use within a visual channel, where each bin is a distinguishable step or level from the other * - (*Tamara Munnzer*)

* Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Discriminability)

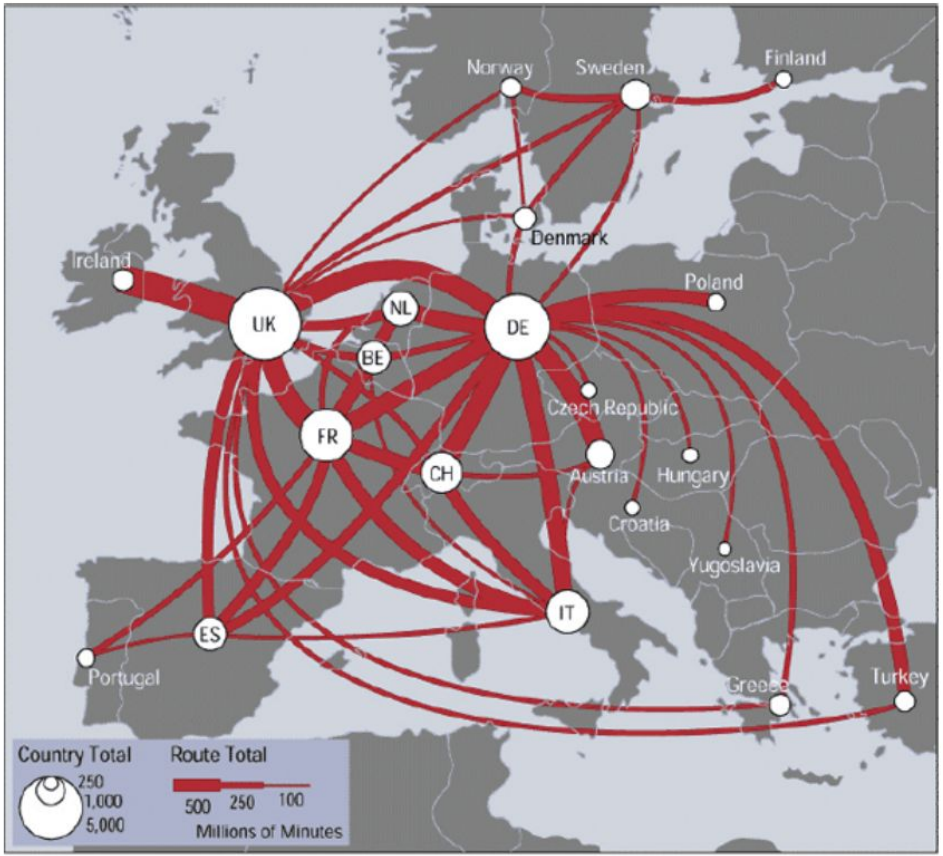


Image Source: Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Separability)

- One cannot treat all visual channels as completely independent from each other, because some have dependencies and interactions with others. -(*Tamara Munzner*)

- One must consider a continuum of potential interactions between channels for each pair, ranging from the orthogonal and independent separable channels to the inextricably combined integral channels.-(*Tamara Munzner*)

* Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Separability)

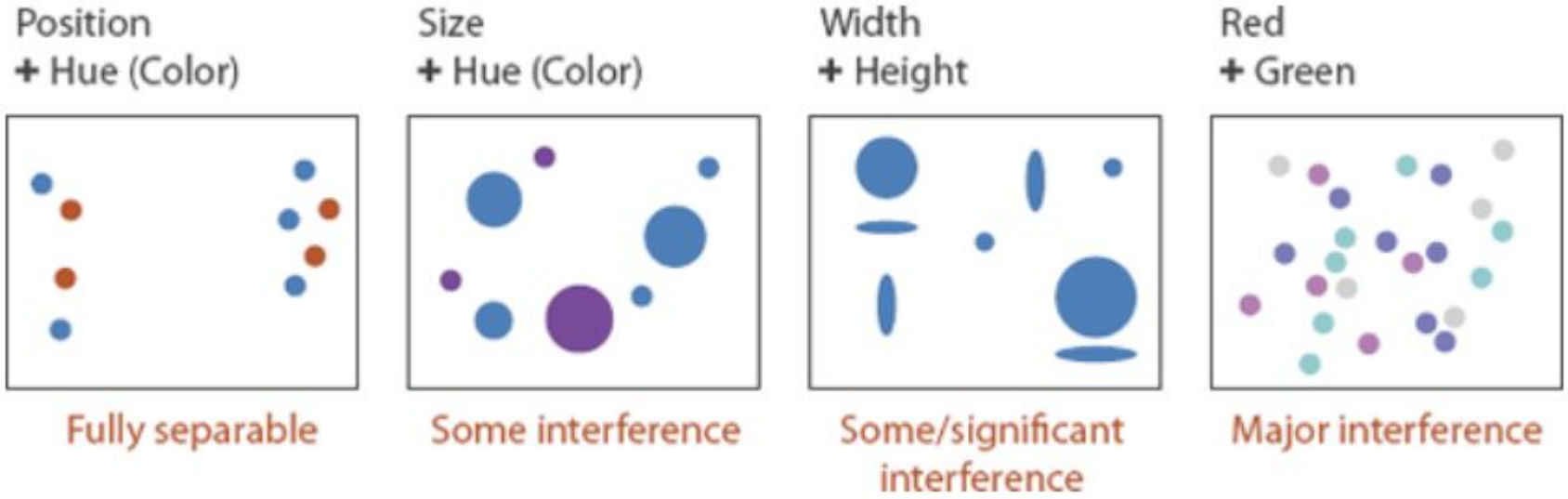


Image Source: Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Popout)

- Visual popout is often called preattentive processing or tunable detection - (*Tamara Munzer*)

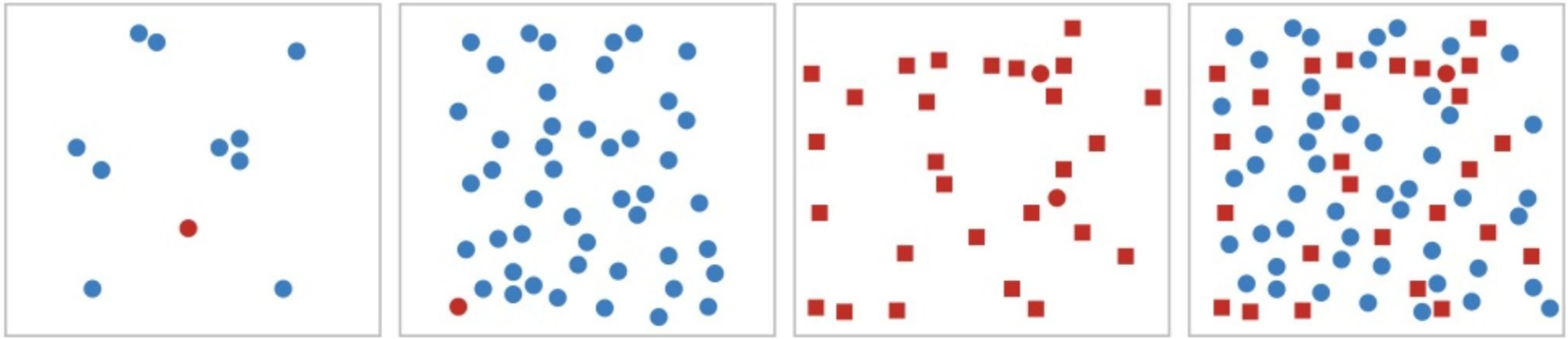


Image Source: Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Popout)

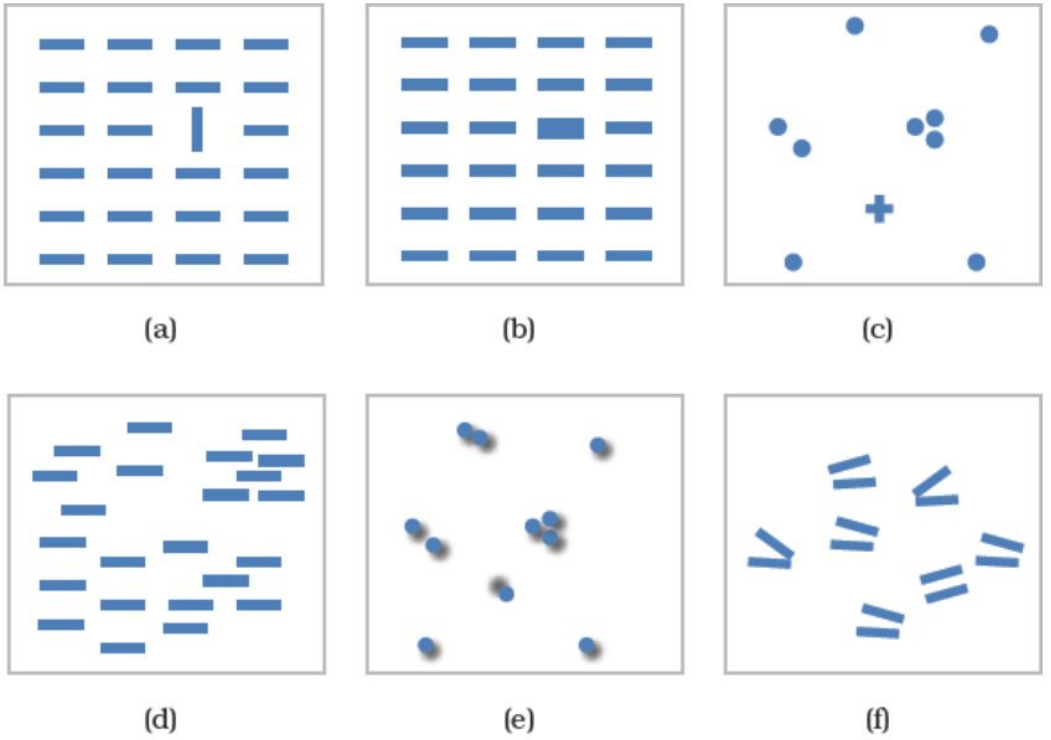
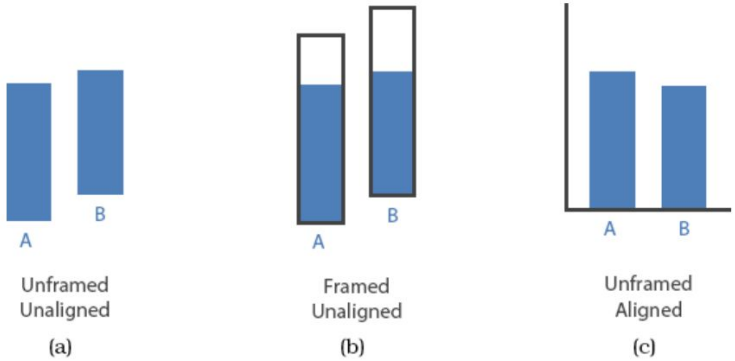


Image Source: Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channel Effectiveness (Relative vs. Absolute)

- The human perceptual system is fundamentally based on relative judgements, not absolute ones this principle is known as Weber's Law - *Tamara Munzner*
- For instance, the amount of length difference we can detect is a percentage of the object's length -*Tamara Munzner*



Weber's Law states that we judge based on relative, not absolute differences. (a) The lengths of unframed, unaligned rectangles of slightly different sizes are hard to compare. (b) Adding a frame allows us to compare the very different sizes of the unfilled rectangles between the bar and frame tops. (c) Aligning the bars also makes the judgement easy. Redrawn and extended after [Cleveland and McGill 84a, Figure 12].

Image Source: Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Channels: Expressiveness Types and Effectiveness Ranks

Magnitude Channels: Ordered Attributes

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

Most ▲ Effectiveness ▼ Least

Same: Color luminance, Color saturation

Same: Curvature, Volume (3D size)

Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape

Next Week (Week 5.1)

- Assignment A2 will be posted today (02/11)
- Quiz 1 (Introduction to data vis, visual encoding))
- Week 5.1 (Visualizing amounts)