Visual perception

Data Management and Visualization

http://softeng.polito.it

Version 2.1.1

GRAPHICAL INTEGRITY

Principles of integrity

PUC

- Proportionality
  - Representation as physical quantities should be proportional to the represented numbers
- Utility
  - Graphical element should convey useful information
- Clarity
  - Labeling should counter graphical distortion and ambiguity

Proportionality

PUC

- The magnitude of visual attributes should represent faithfully the magnitude of measures
- They should allow
  - Discrimination: are they different?
  - Comparison: which is larger?
  - Magnitude Assessment: how much larger?

Lie Factor

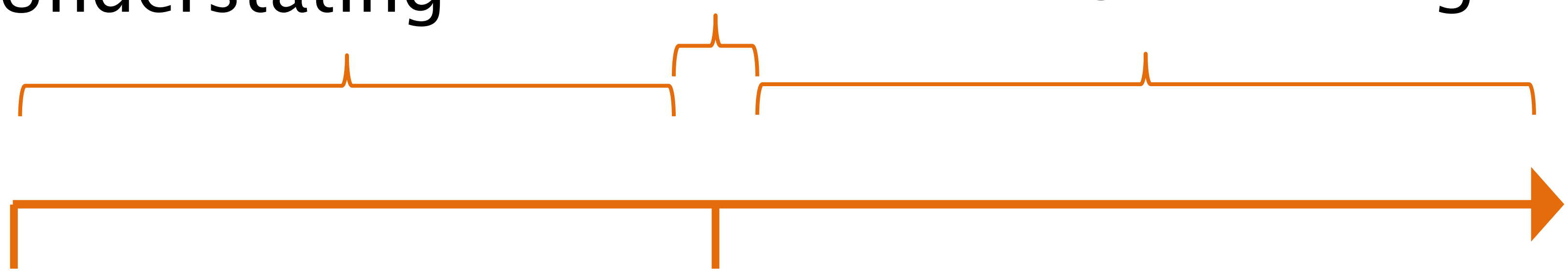
PUC

$$LF = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

- Overstating
  - $LF > 1 \Leftrightarrow Log(LF) > 0$
- Understating
  - $LF < 1 \Leftrightarrow Log(LF) < 0$
- Fair
  - $LF = 1 \Leftrightarrow Log(LF) = 0$

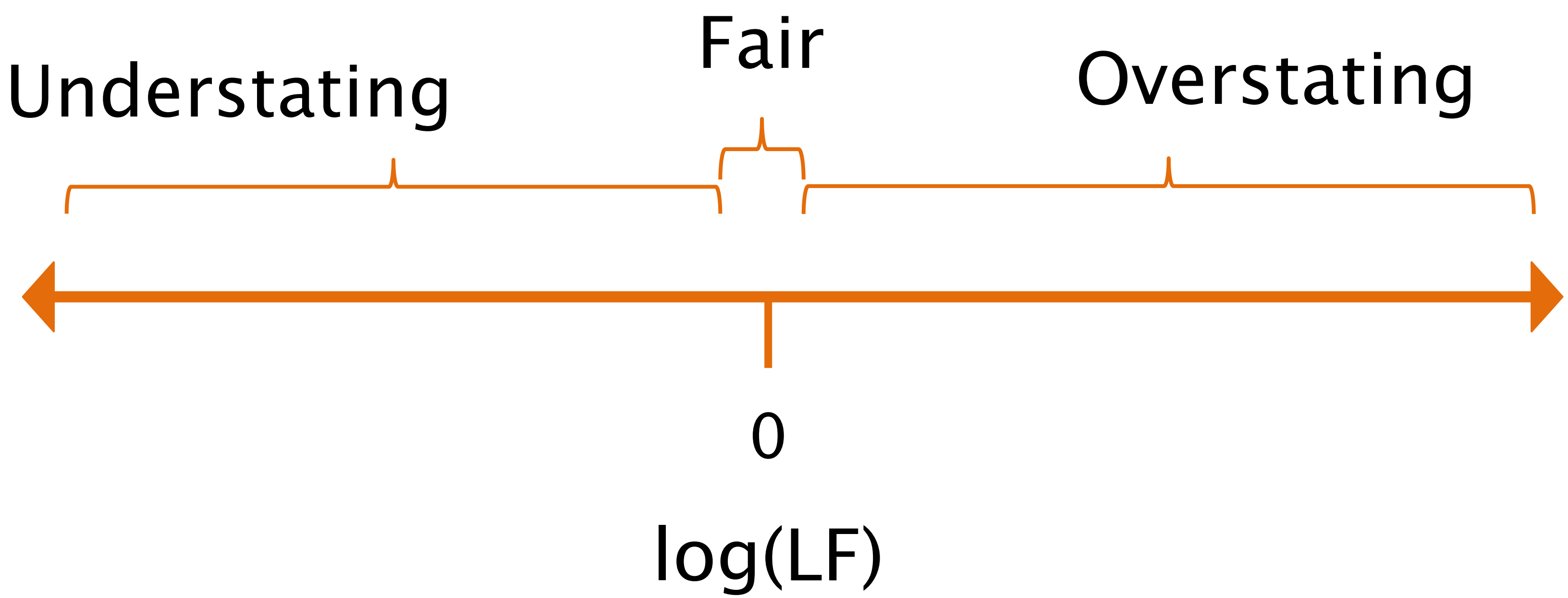
Lie Factor

$$LF = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$



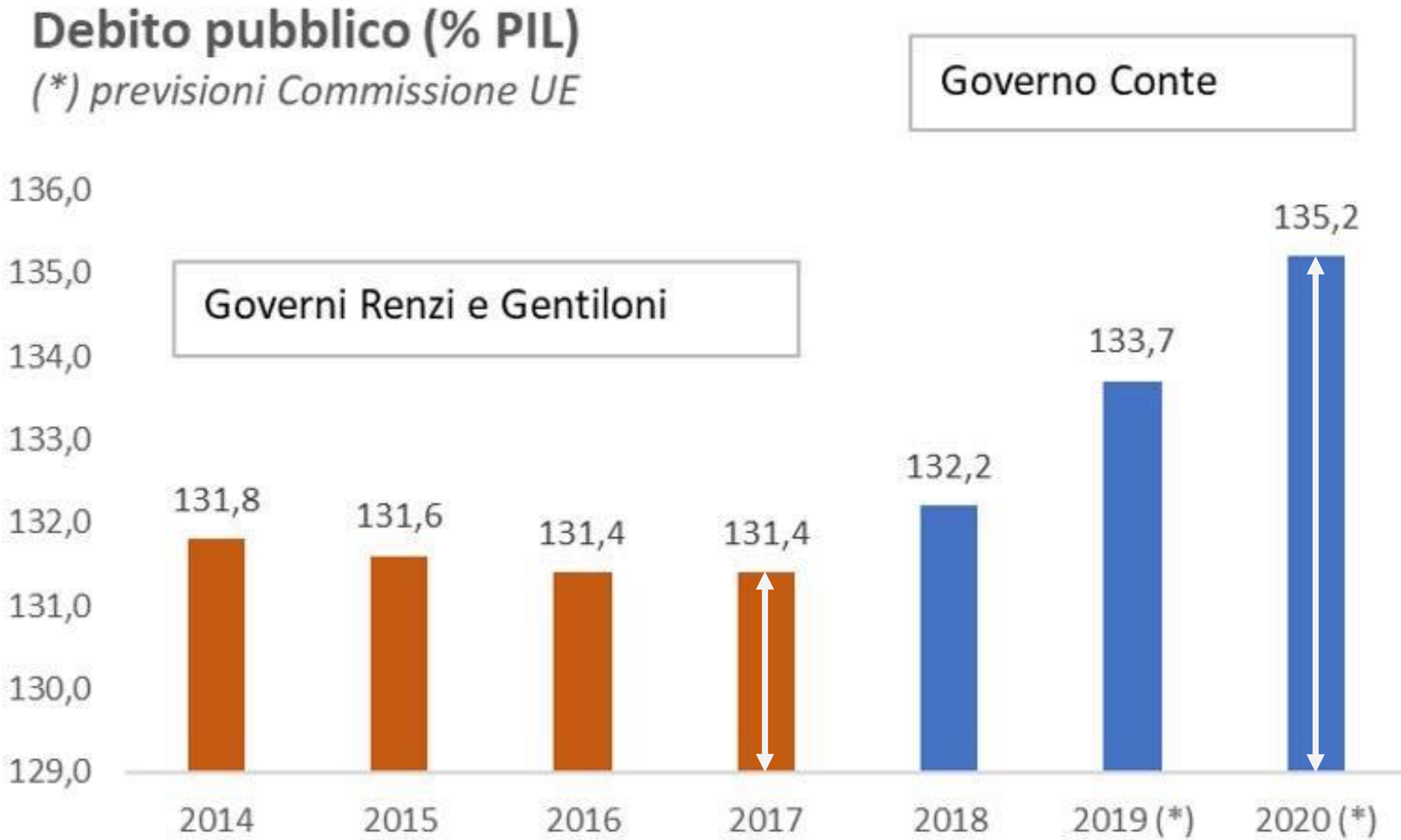
Lie Factor

$$LF = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$



Example

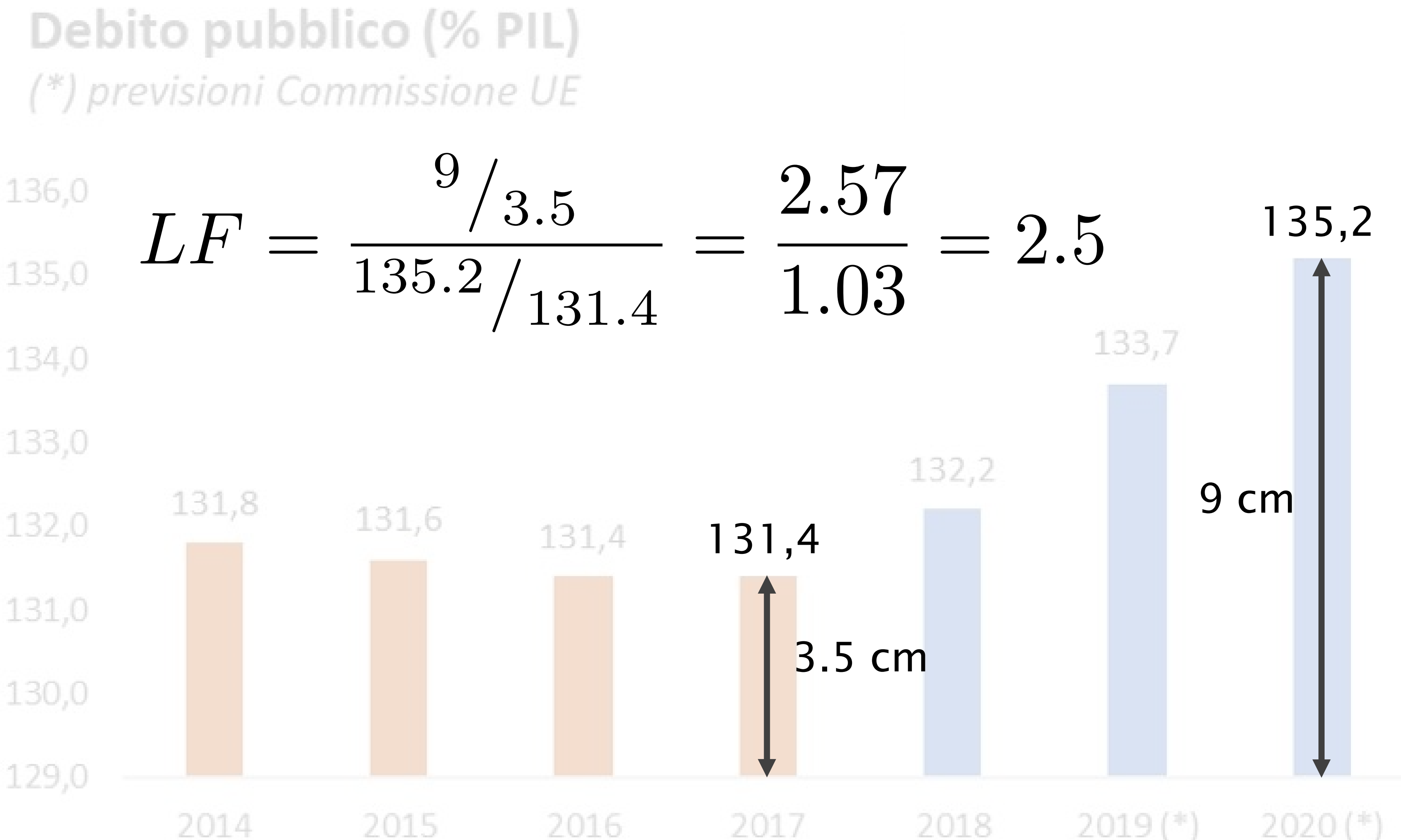
PUC


Debito pubblico (% PIL)
(*) previsioni Commissione UE
Governo Conte
Governi Renzi e Gentiloni
136,0
135,0
134,0
133,0
132,0
131,0
130,0
129,0
131,8
131,6
131,4
131,4
132,2
133,7
135,2
2014
2015
2016
2017
2018
2019 (*)
2020 (*)

Example – Lie Factor

PUC

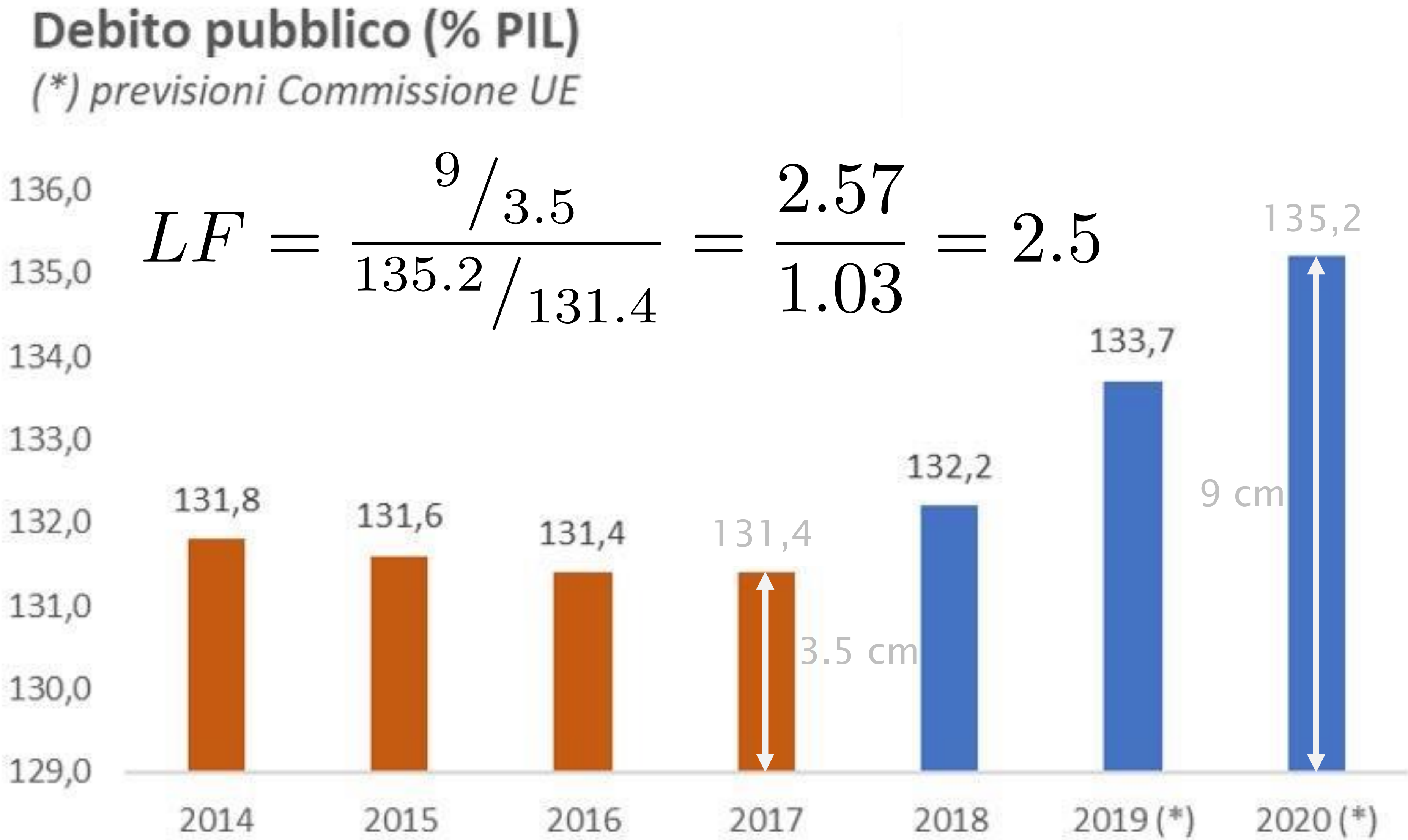
$$LF = \frac{{}^{9}/_{3.5}}{{}^{135.2}/_{131.4}} = \frac{2.57}{1.03} = 2.5$$



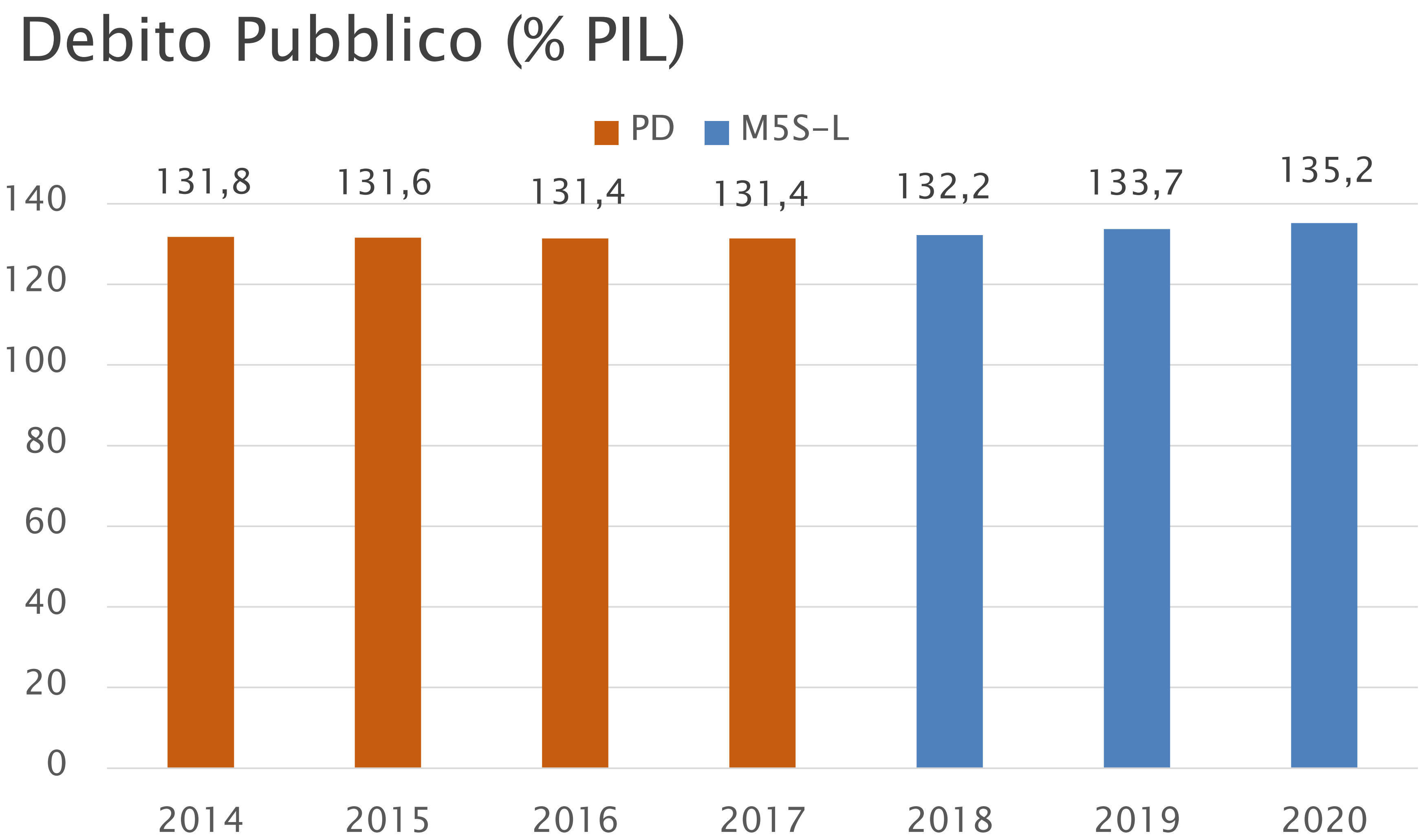
Example – Lie Factor

PUC

$$LF = \frac{^{9}/_{3.5}}{^{135.2}/_{131.4}} = \frac{2.57}{1.03} = 2.5$$

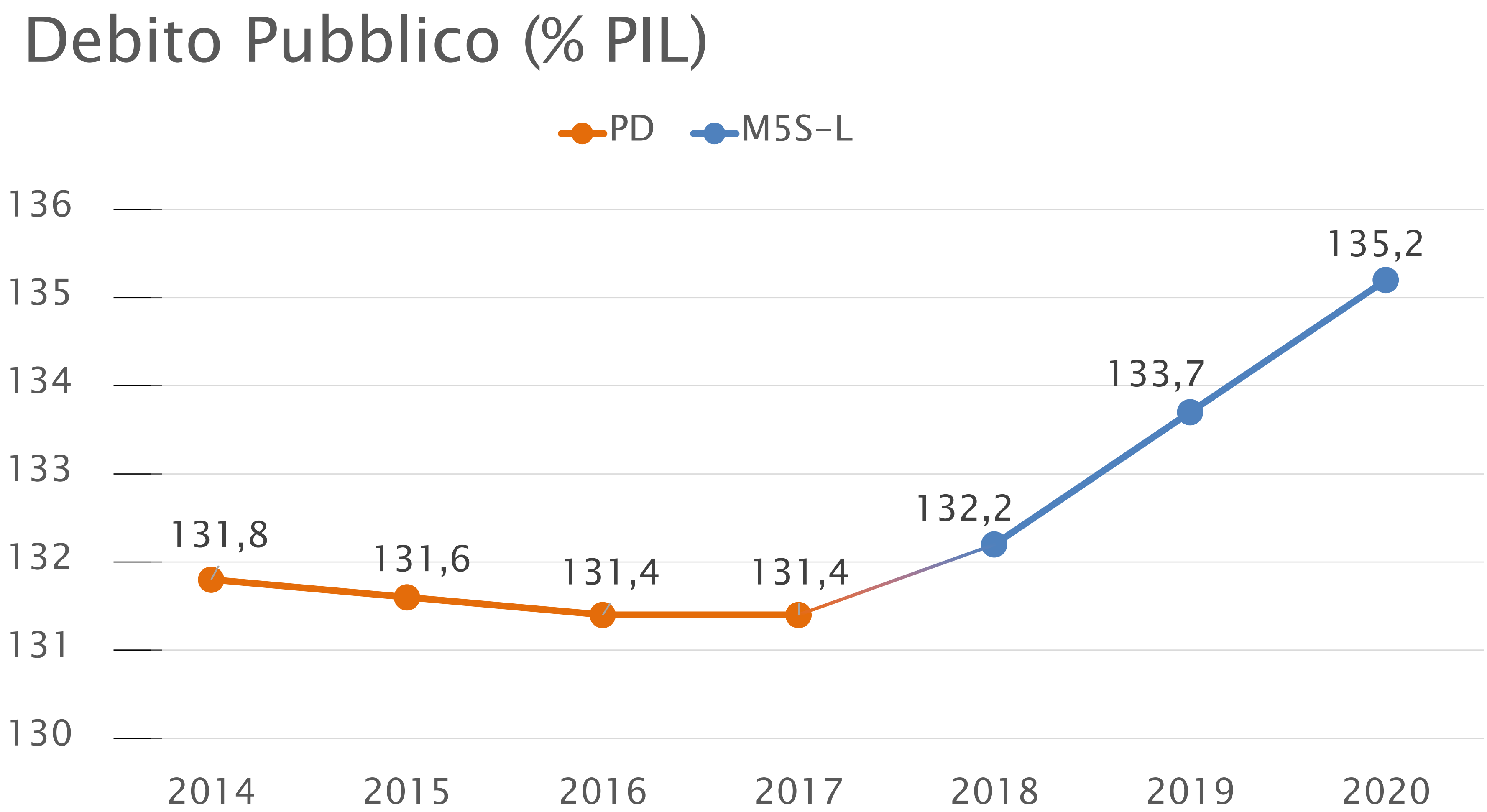


Example – Redesign



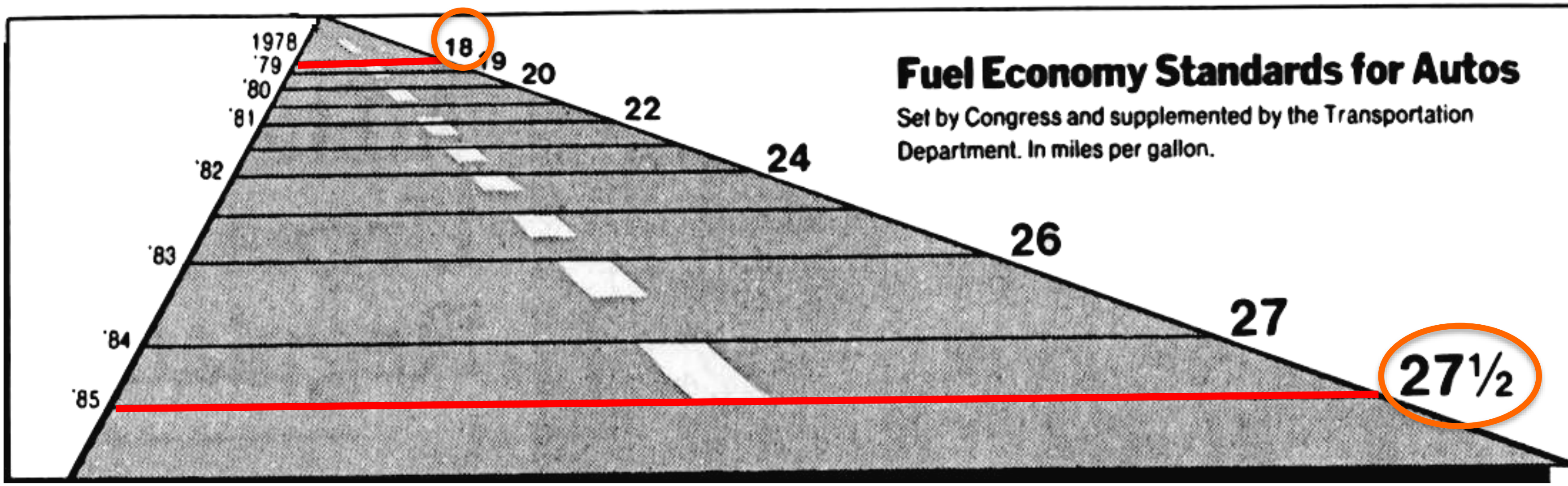
Example – Redesign

PUC



Lie Factor

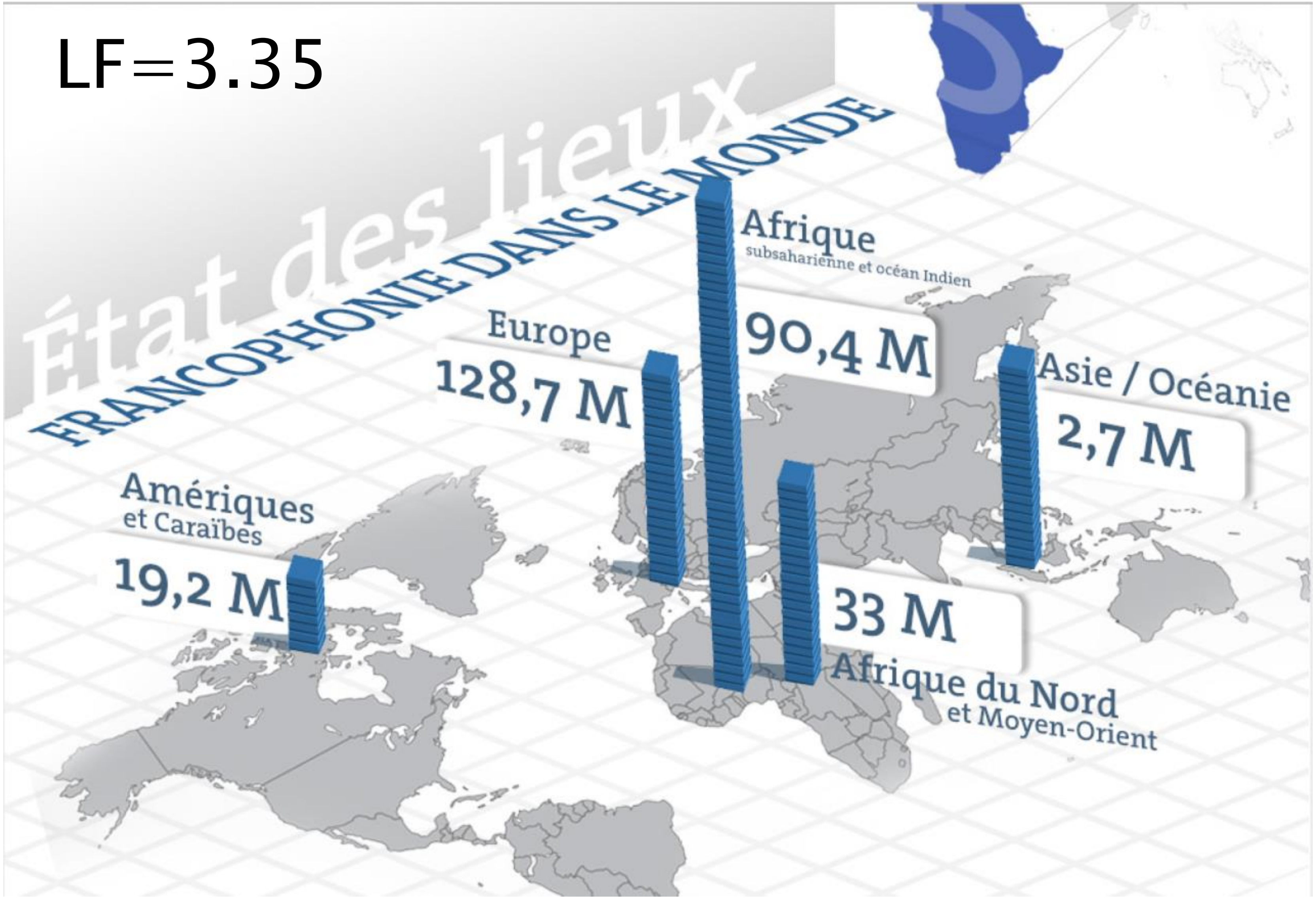
PUC



$$\frac{18.7}{2.2} = 8.5 \text{ on graphic} \qquad \frac{27.5}{18} = 1.52 \text{ in data}$$

LF = 8.5 / 1.52 = 5.59

Lie Factor


LF=3.35
État des lieux
FRANCOPHONIE DANS LE MONDE
Afrique
subsaharienne et océan Indien
90,4 M
Europe
128,7 M
Asie / Océanie
2,7 M
Amériques
et Caraïbes
19,2 M
33 M
Afrique du Nord
et Moyen-Orient

Principle of design

- Keep the physical Lie Factor = 0
- Limit the perceptual Lie Factor as much as possible
  - Per Steven's law, avoid area comparisons

Utility

PUC

- Every element should convey useful information
- Unnecessary visual objects or attributes distract from the message
  - Different attributes trigger a search for a rationale (e.g. random colors)

Data-ink

PUC

$$\text{Data-ink ratio} = \frac{\text{data ink}}{\text{total ink used to print the graphic}}$$

- Proportion of a graphic's ink devoted to the non-redundant display of data information
  - Or:

$$1 - \frac{\text{ink that can be erased without loss of information}}{\text{total ink used to print the graphic}}$$

Principles of design

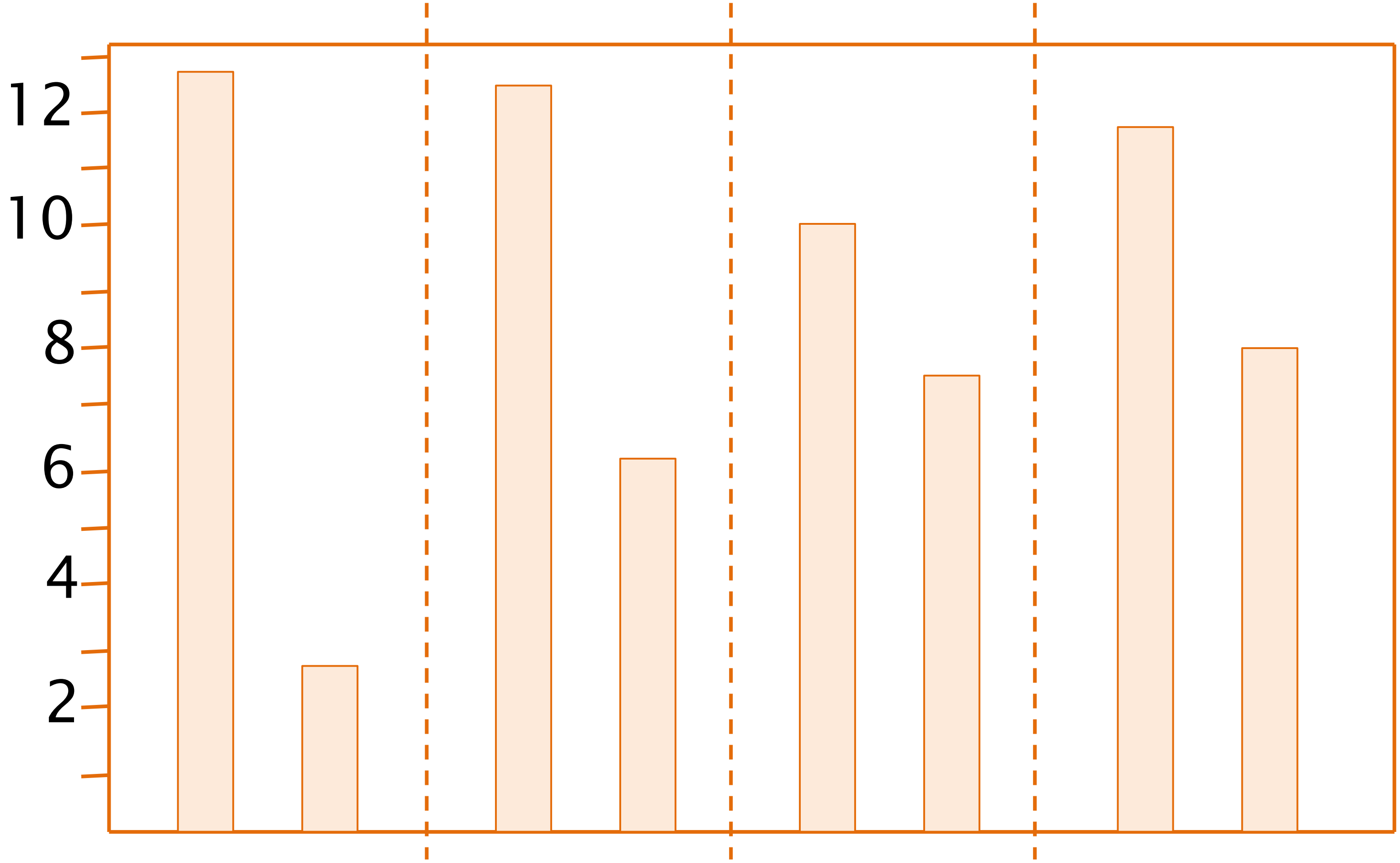
- Maximize data-ink ratio
  - Erase non-data-ink
  - Erase redundant data-ink

- "Within reason"

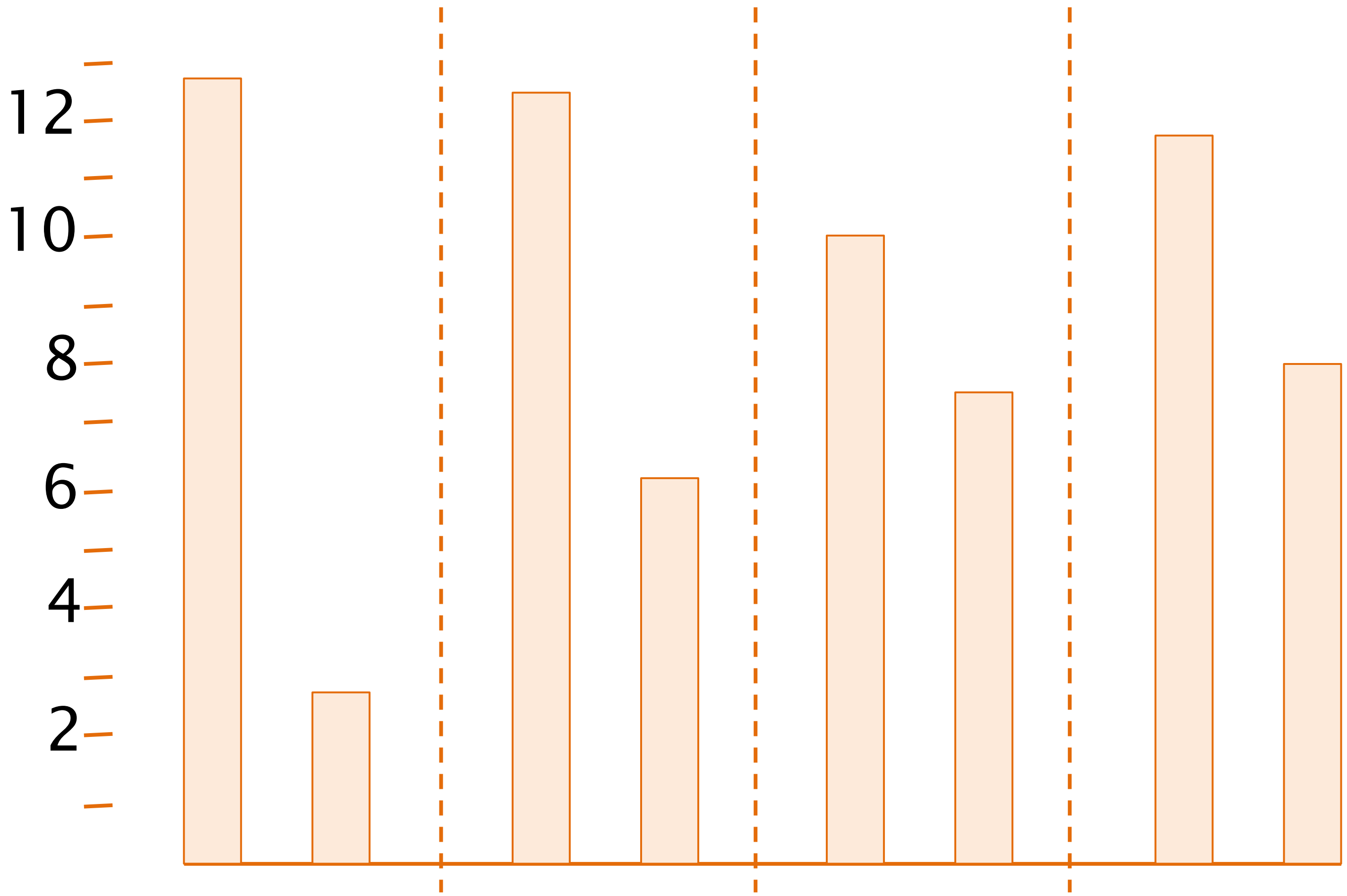
Above all else show the data

E.Tufte

Data-ink

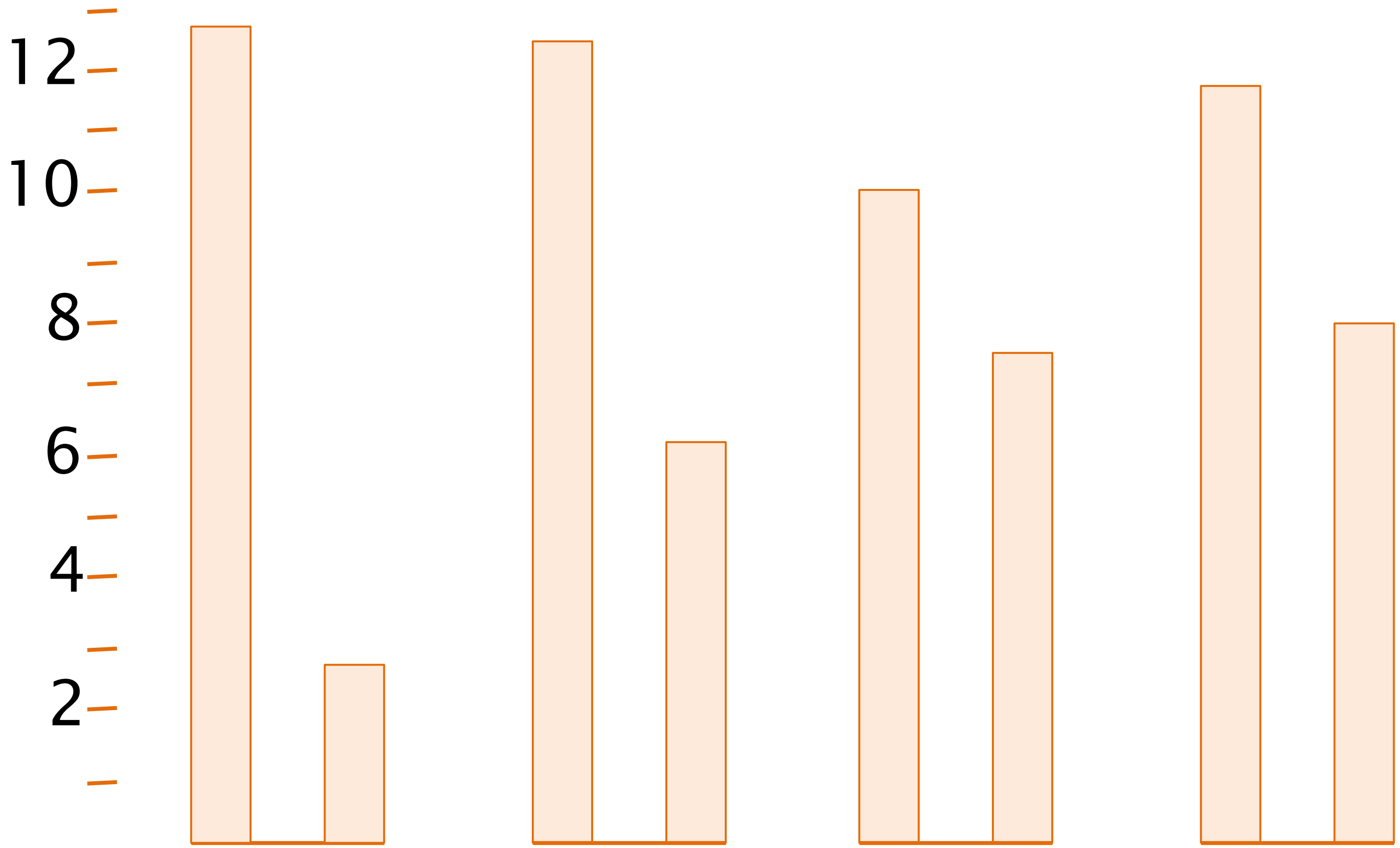


Data-ink



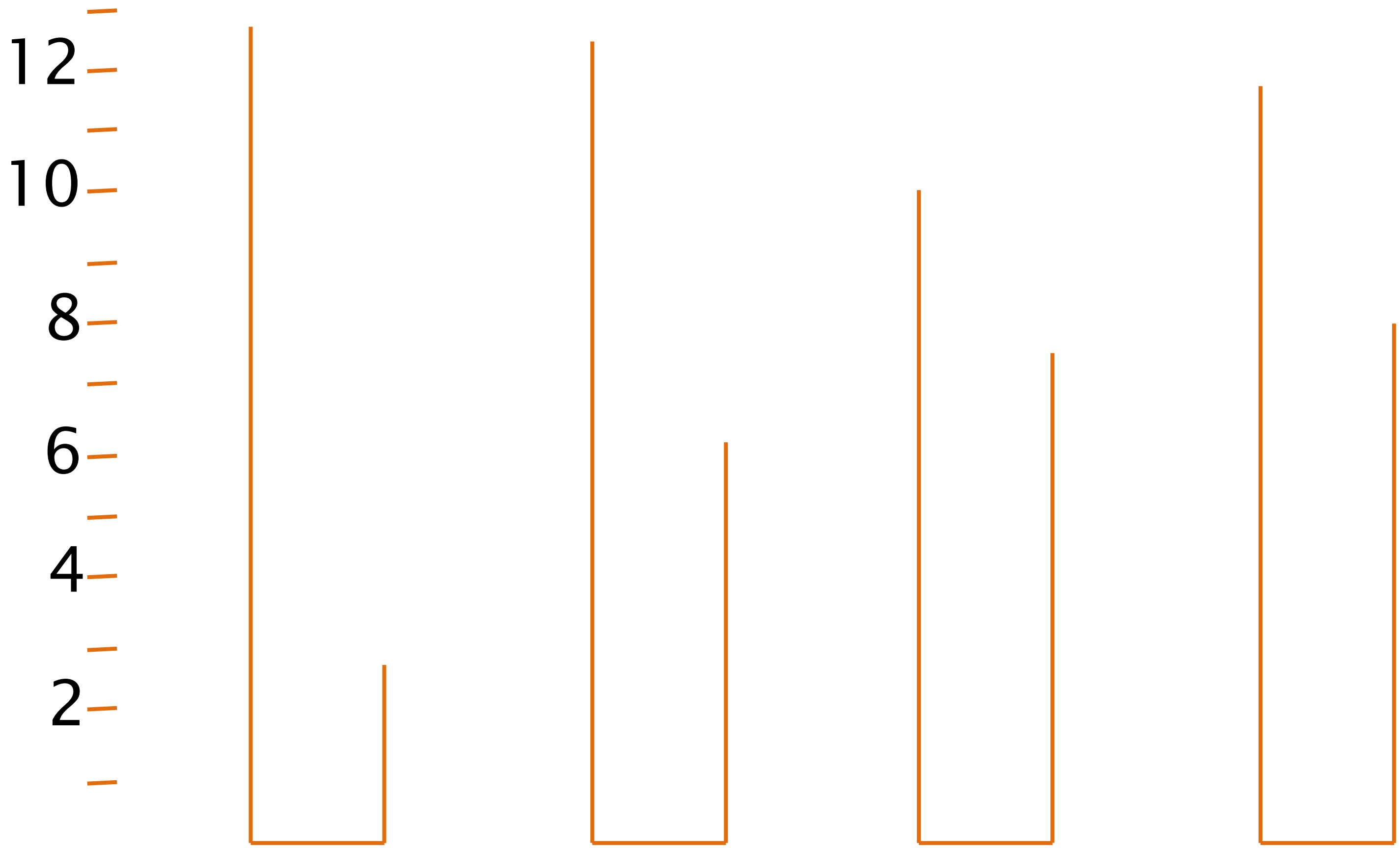
Data-ink

PUC



Data-ink

PUC



Tufte's proposed redesign

Use of contrast

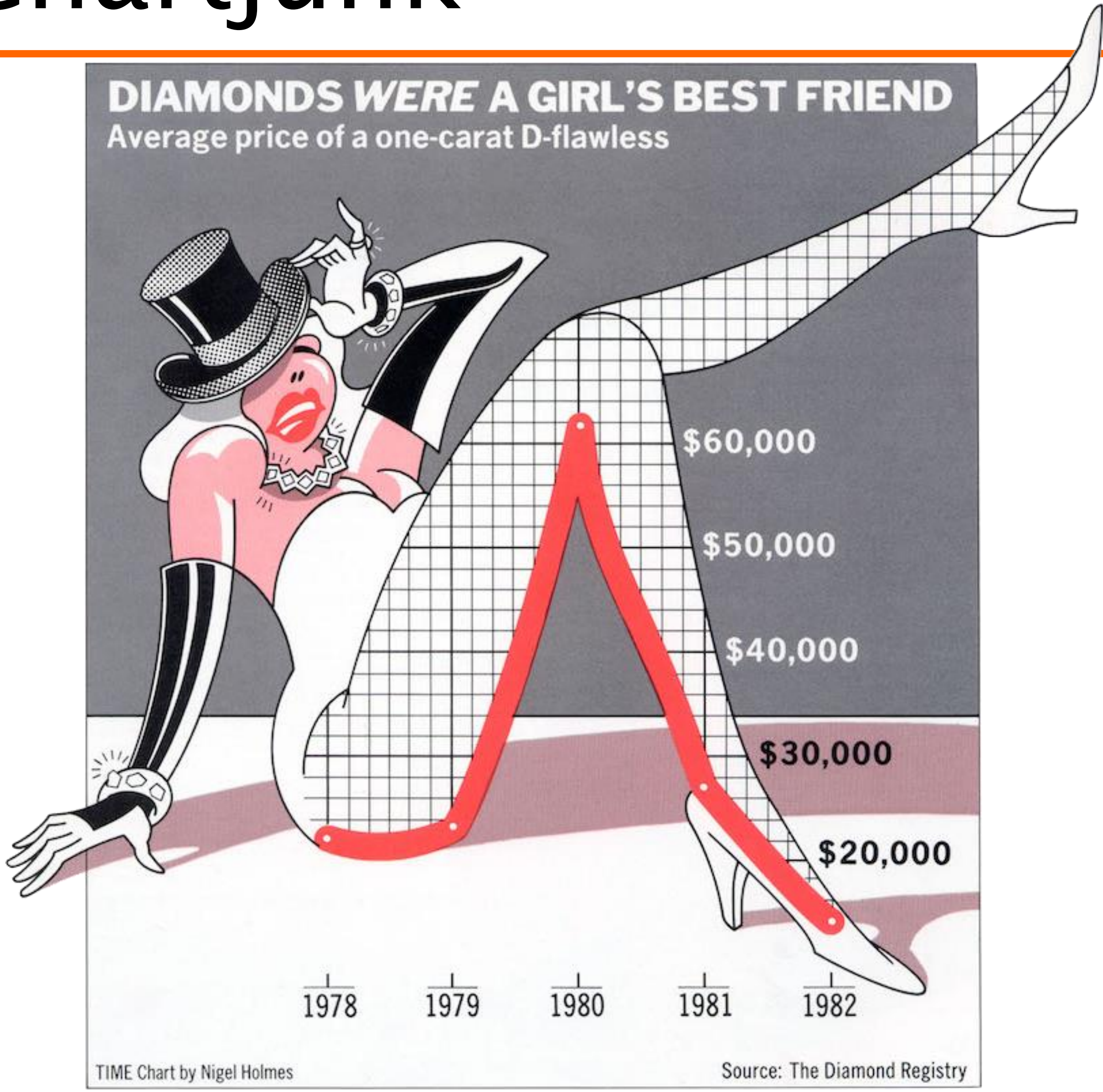
PUC

- Include differences corresponding to actual differences
- Effective when one item is different in a context of other items that are the same
  - Bright saturated color among mid colors

Chartjunk

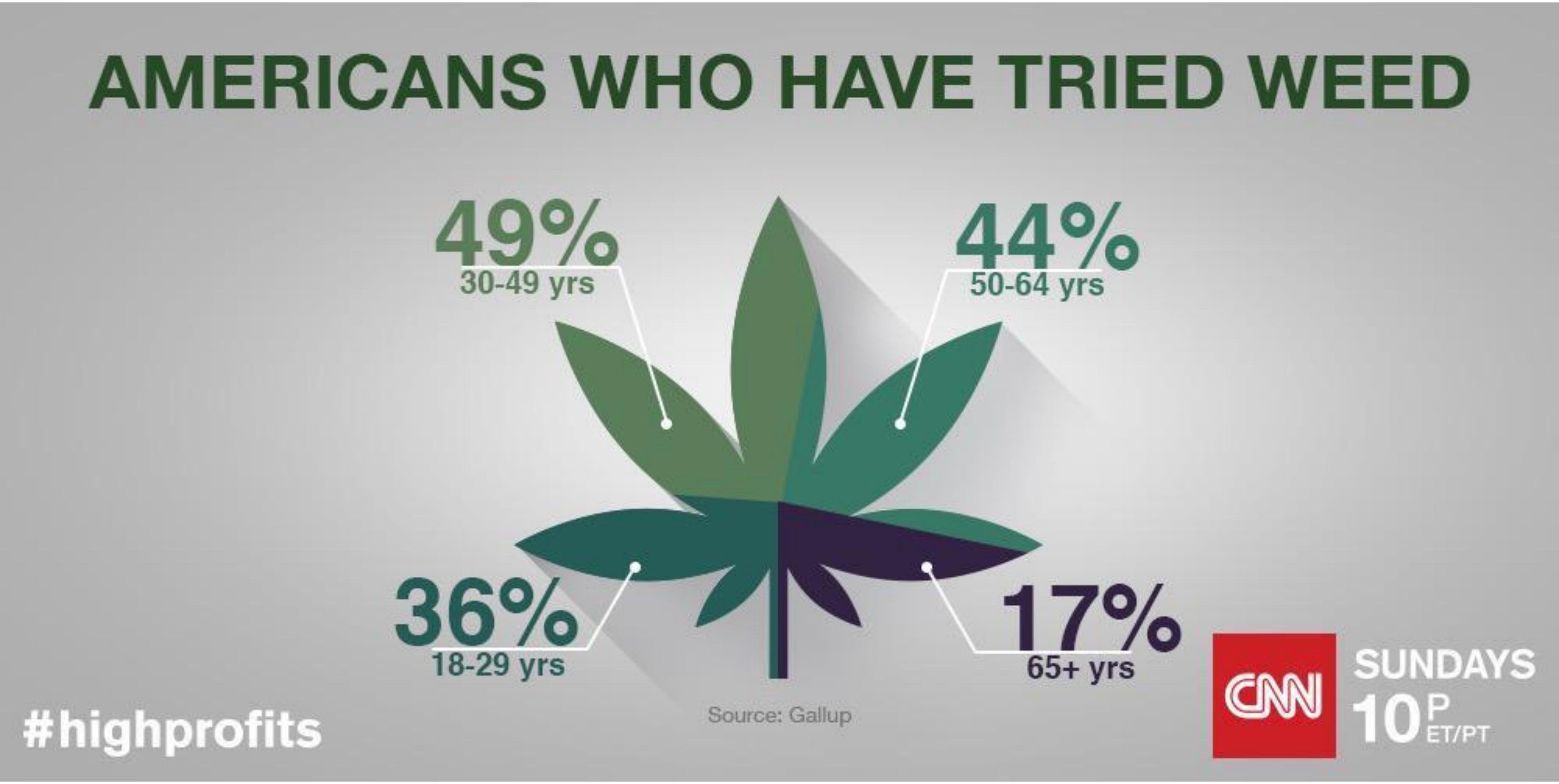
- The presence of unnecessary elements that distract or hide the message conveyed by the diagram

Chartjunk



Nigel Holmes:
http://nigelholmes.com

Chartjunk


AMERICANS WHO HAVE TRIED WEED
49%
30-49 yrs
44%
50-64 yrs
36%
18-29 yrs
17%
65+ yrs
#highprofits
Source: Gallup
CNN
SUNDAYS
10P ET/PT

Clarity

- Visual encoding and layout should make perception tasks easy and effortless
- Textual and support elements should provide effective support to understanding the information
- Any variation in the graph should represent useful information otherwise it is noise obfuscating the message

Clarity

PUC

- Textual elements should provide effective support to understanding
  - Hierarchical
    - Size and position reflects importance
  - Readable
    - Large enough
  - Horizontal
  - Close to data (avoid legends)
- Always label the axes

Colors

- Get it right in black and white
- Use medium hues or pastels
  - Bright colors distract and tire out
- Use color only when needed to serve a particular communication goal

Detection and Separation

Efficiency and efficacy of perception tasks is affected by:

- Detection

  The capability to visually identify the objects that represent the data to be compared

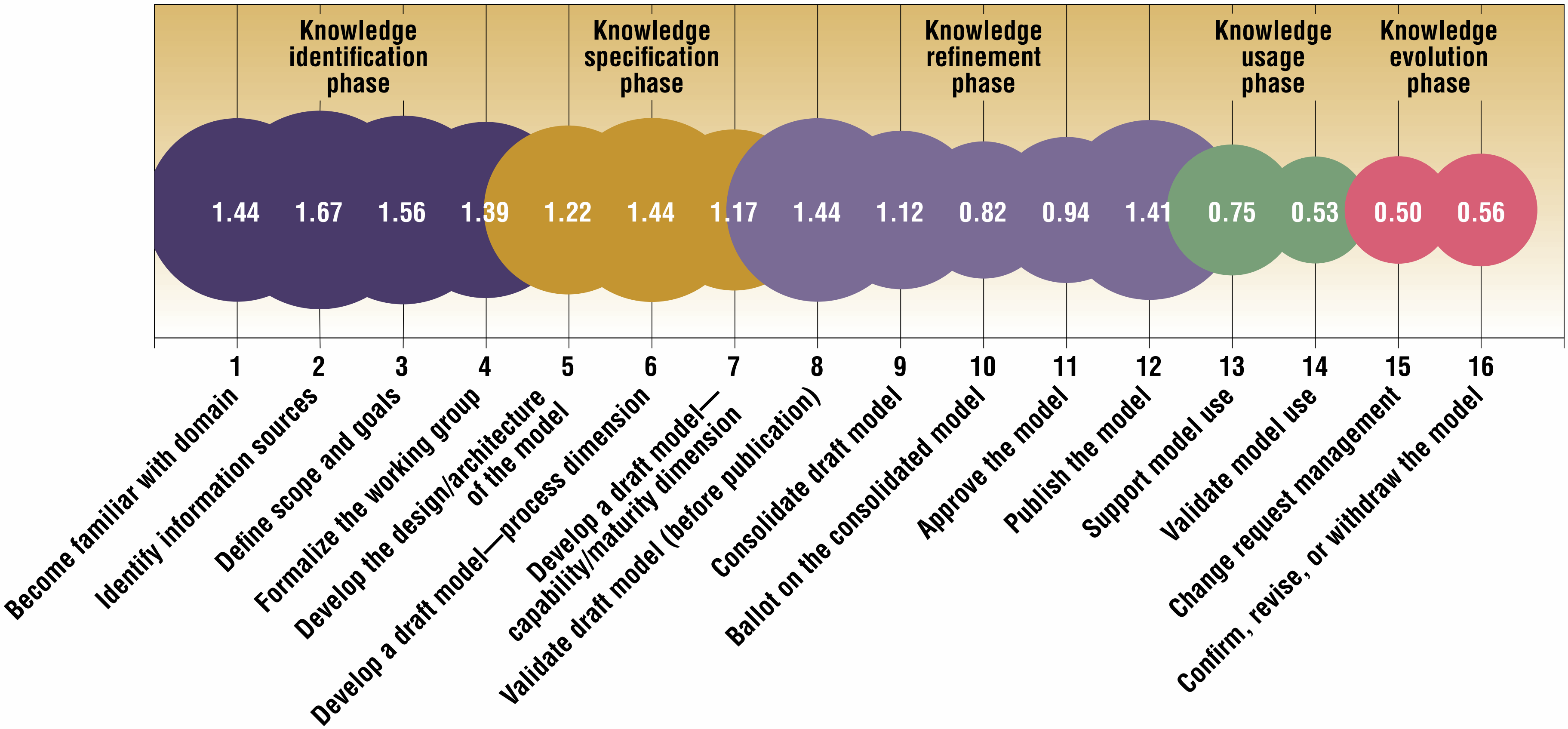
- Separation

  The distance between the objects to be compared

  - affects negatively the accuracy

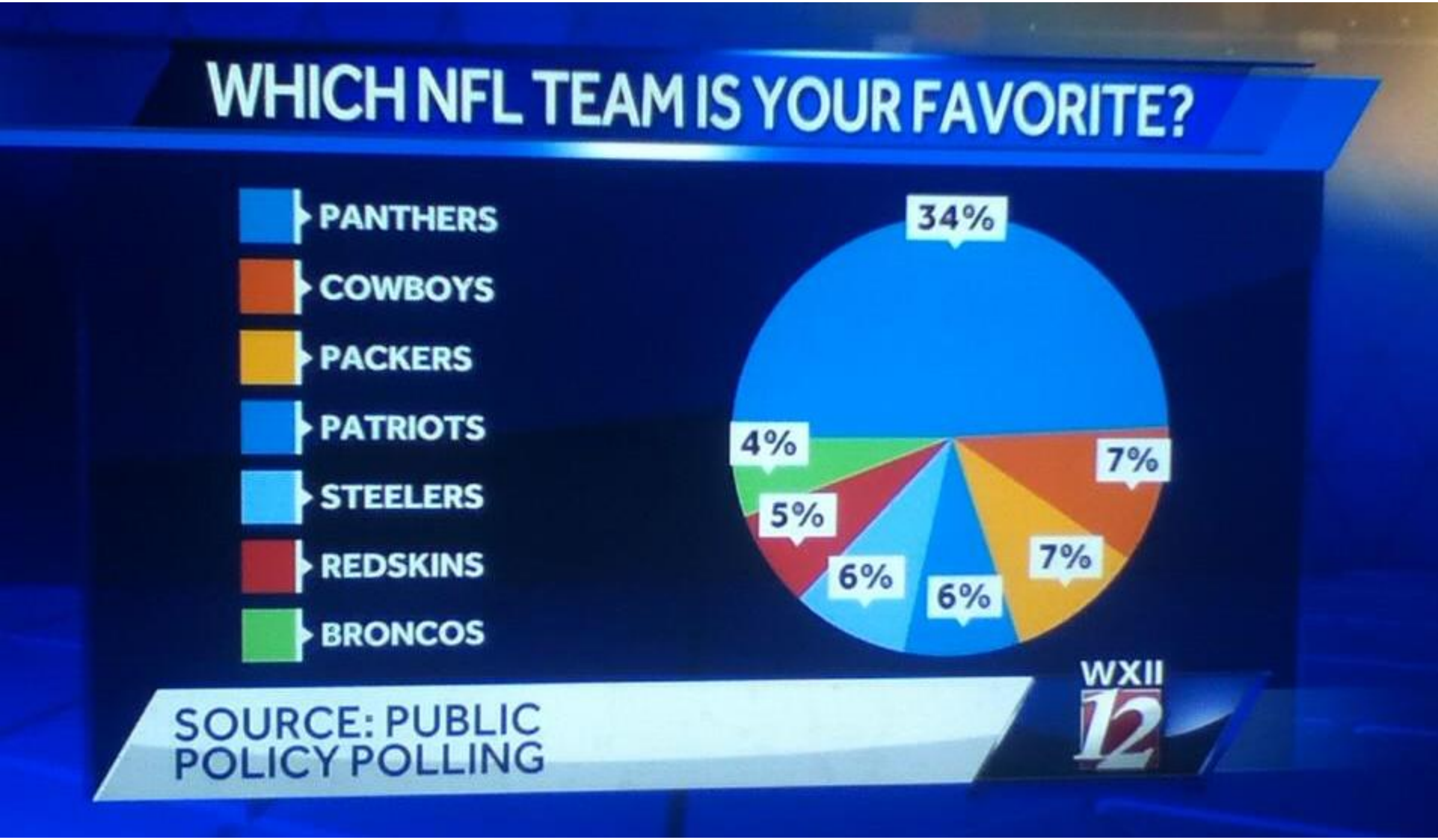
Clarity

PUC



Knowledge identification phase
Knowledge specification phase
Knowledge refinement phase
Knowledge usage phase
Knowledge evolution phase
1.44
1.67
1.56
1.39
1.22
1.44
1.17
1.44
1.12
0.82
0.94
1.41
0.75
0.53
0.50
0.56
1 Become familiar with domain
2 Identify information sources
3 Define scope and goals
4 Formalize the working group
5 Develop the design/architecture of the model
6 Develop a draft model—process dimension
7 Develop a draft model—capability/maturity dimension
8 Validate draft model (before publication)
9 Consolidate draft model
10 Ballot on the consolidated model
11 Approve the model
12 Publish the model
13 Support model use
14 Validate model use
15 Change request management
16 Confirm, revise, or withdraw the model

Case study


WHICH NFL TEAM IS YOUR FAVORITE?
PANTHERS
COWBOYS
PACKERS
PATRIOTS
STEELERS
REDSKINS
BRONCOS
34%
4%
7%
5%
7%
6%
6%
SOURCE: PUBLIC POLICY POLLING
WXII 12

Assessment

- Question:
  - Is there one (or more) question addressed by the visualization?
- Data:
  - Is the data quality appropriate?
- Visual Integrity:
  - Are the visual features appropriate?

Visual Integrity

- Proportionality:
  - Are the values encoded in a uniformly proportional way?
- Utility:
  - All the elements in the graph convey useful information?
- Clarity:
  - Are the data in the graph identifiable and understandable (properly described)?

Question

- What are the most popular/favorite NFL teams in our audience?
- …

Data

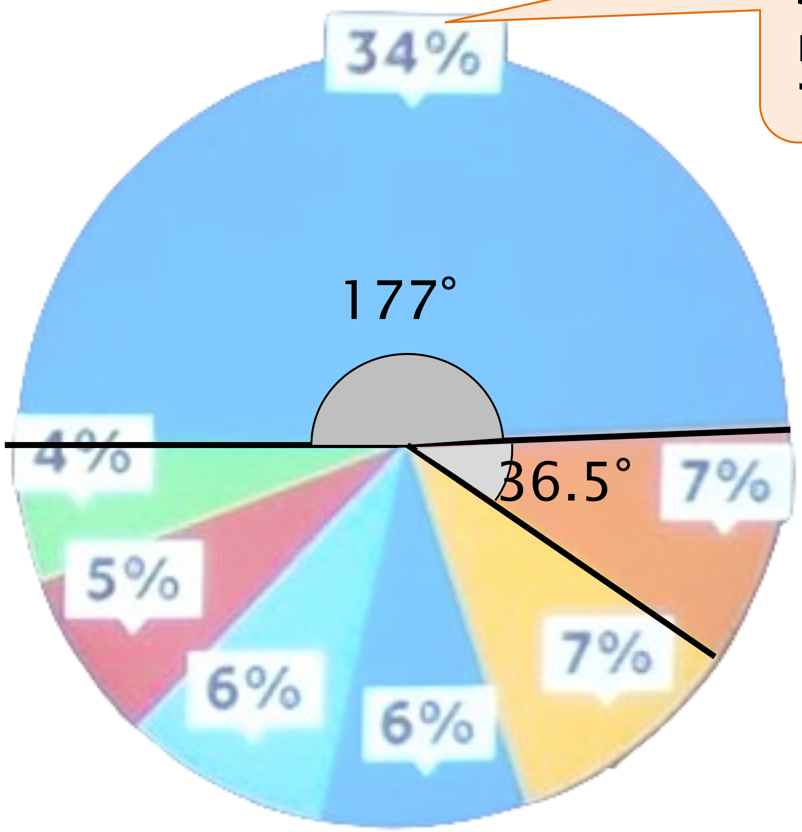
WXII-TV is an NBC-affiliated television station serving North Carolina: home of Panthers

Team	Preferences
Panthers	34%
Cowboys	7%
Packers	7%
Patriots	6%
Steelers	6%
Redskins	5%
Broncos	4%
Total:	69%

Full data

Team	Preferences
Panthers	34%
Cowboys	7%
Packers	7%
Patriots	6%
Steelers	6%
Redskins	5%
Broncos	4%
Other	31%
Total:	100%

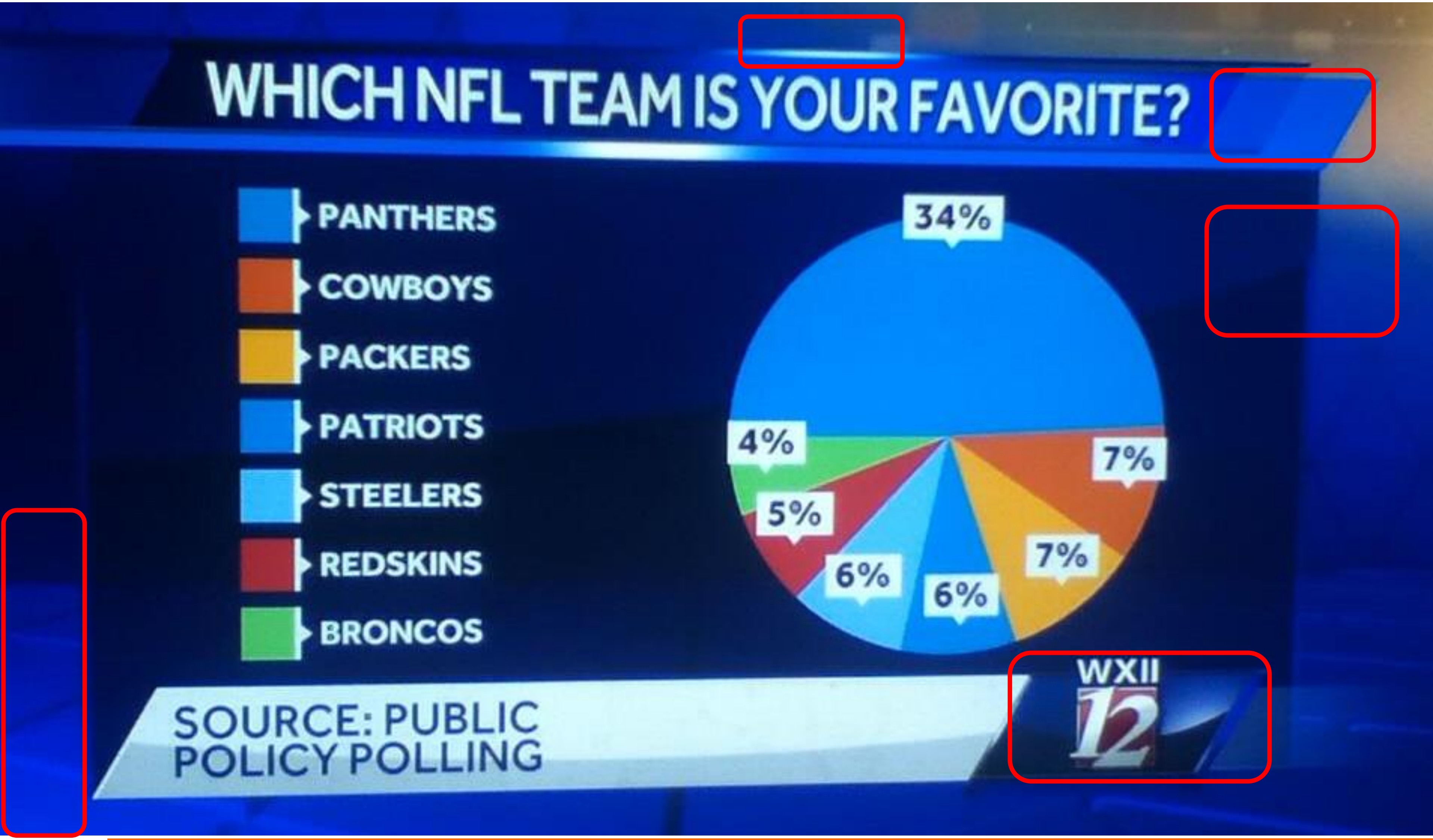
Integrity – Proportionality



34% corresponds to 50% of the pie!

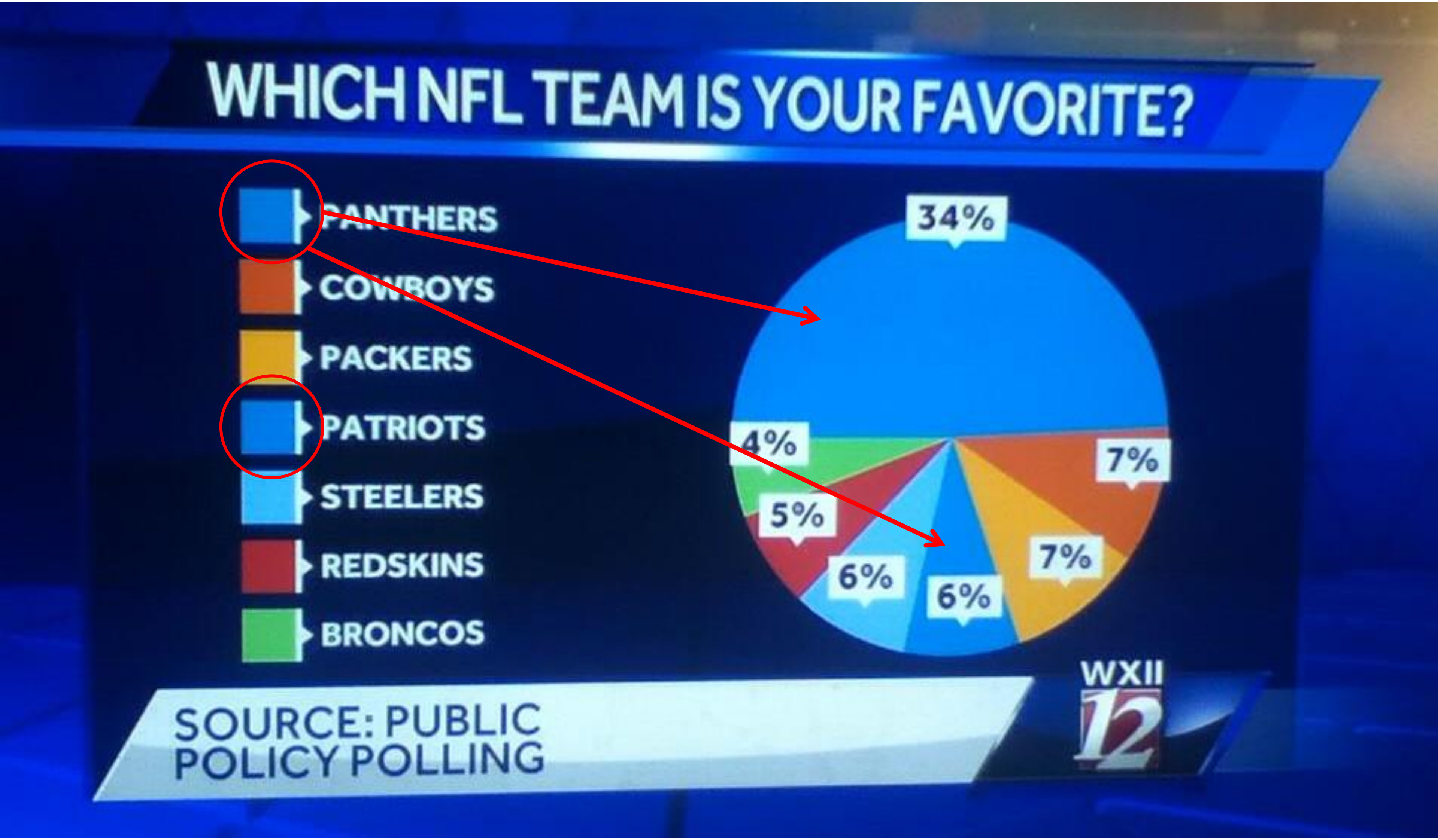
177/36.5 = 4.8
34/7 = 4.8

Utility



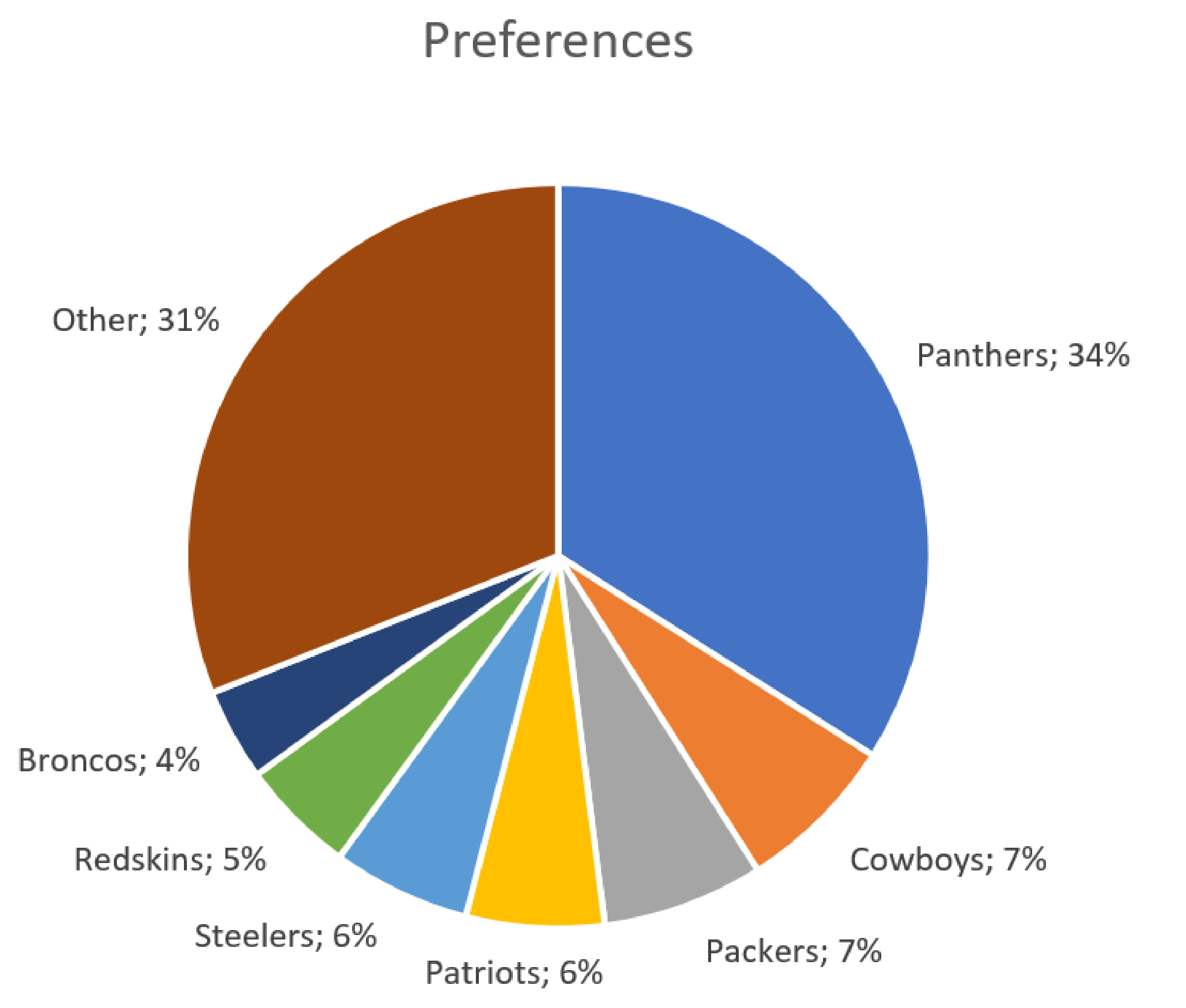
WHICH NFL TEAM IS YOUR FAVORITE?
PANTHERS
COWBOYS
PACKERS
PATRIOTS
STEELERS
REDSKINS
BRONCOS
34%
4%
7%
5%
7%
6%
6%
SOURCE: PUBLIC POLICY POLLING
WXII 12

Clarity

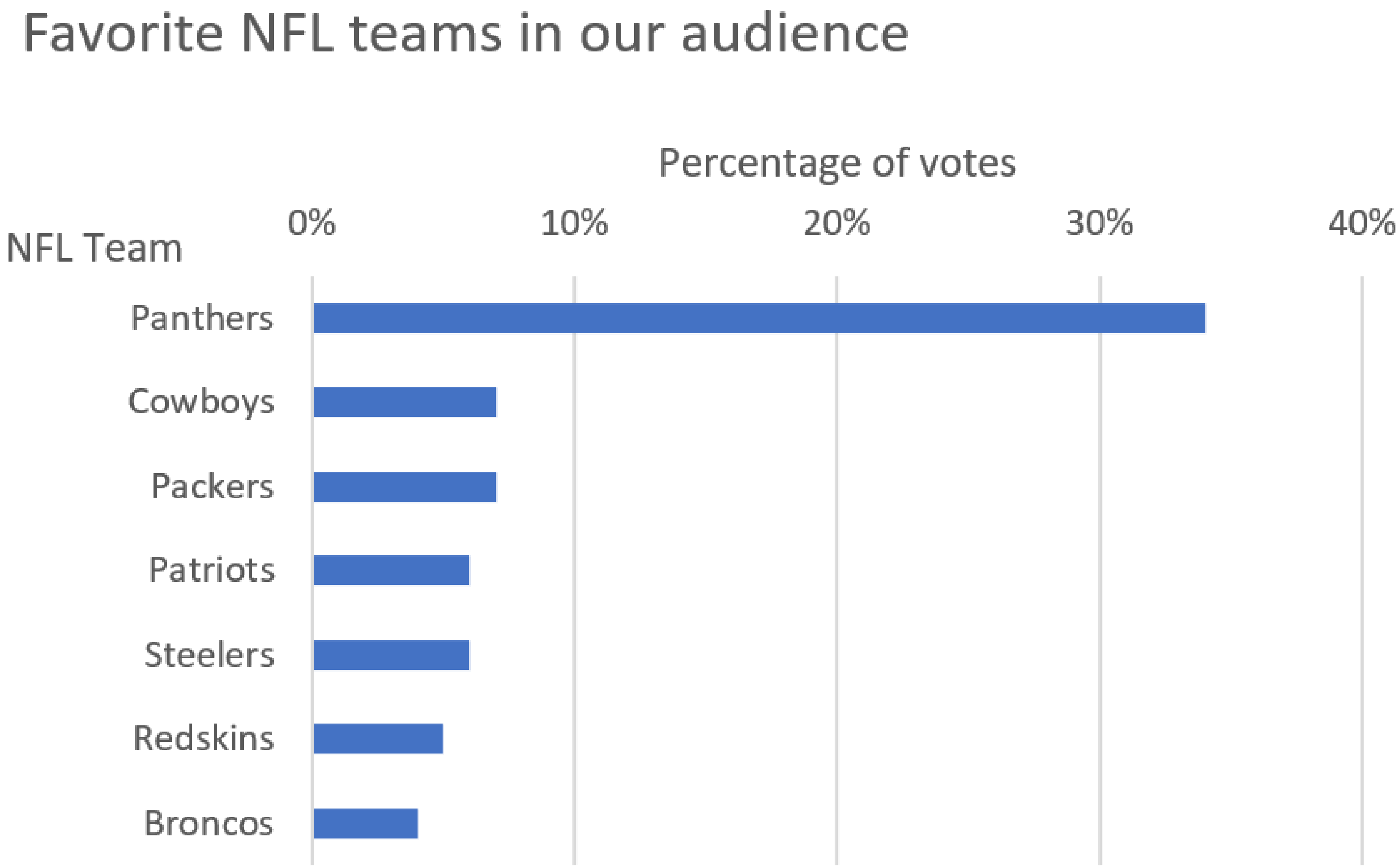


WHICH NFL TEAM IS YOUR FAVORITE?
PANTHERS
COWBOYS
PACKERS
PATRIOTS
STEELERS
REDSKINS
BRONCOS
34%
4%
7%
5%
7%
6%
6%
SOURCE: PUBLIC POLICY POLLING
WXII 12

Redesign #1

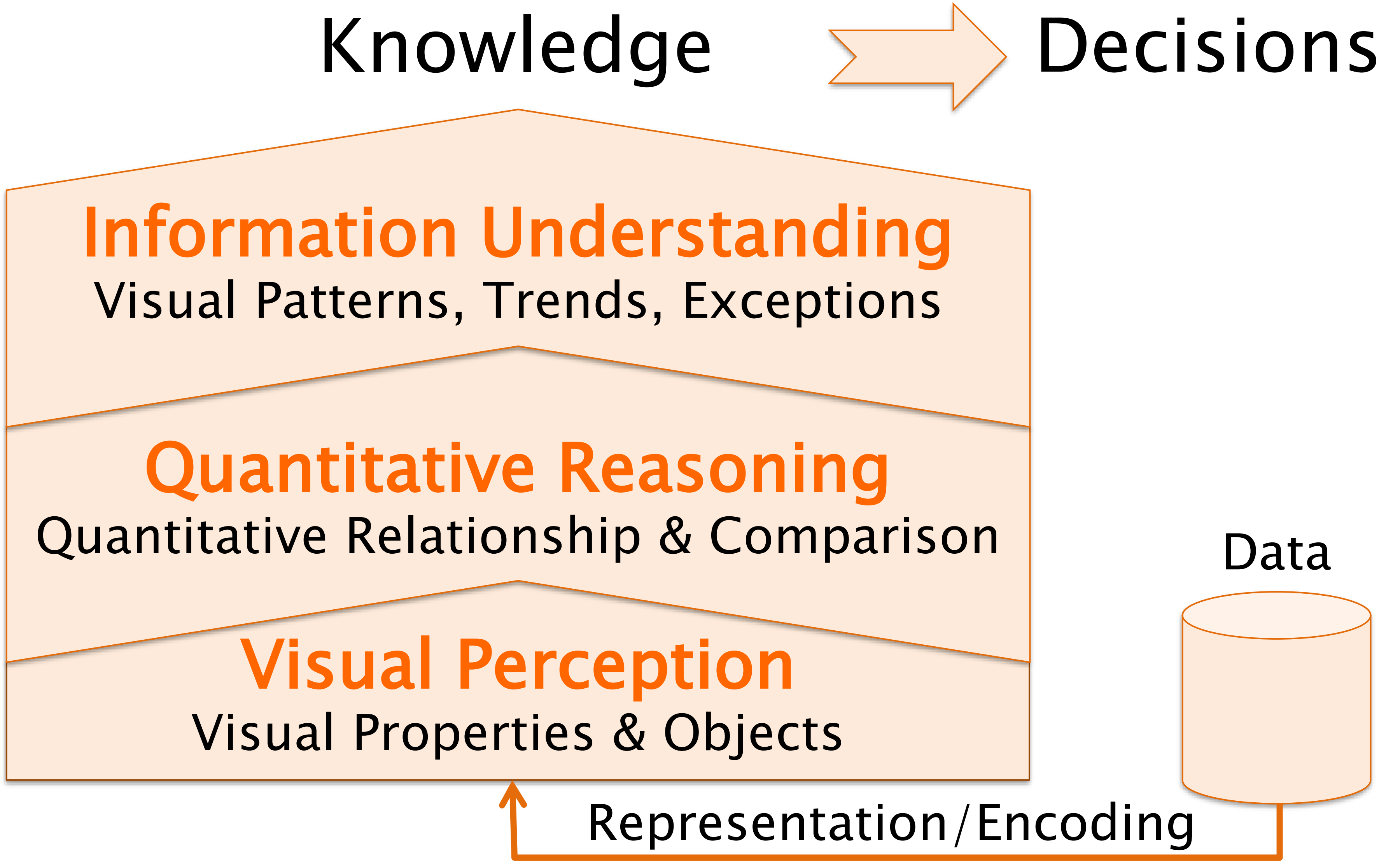

Preferences
Other; 31%
Panthers; 34%
Broncos; 4%
Redskins; 5%
Cowboys; 7%
Steelers; 6%
Patriots; 6%
Packers; 7%

Redesign #2


Favorite NFL teams in our audience
Percentage of votes
NFL Team
0%
10%
20%
30%
40%
Panthers
Cowboys
Packers
Patriots
Steelers
Redskins
Broncos

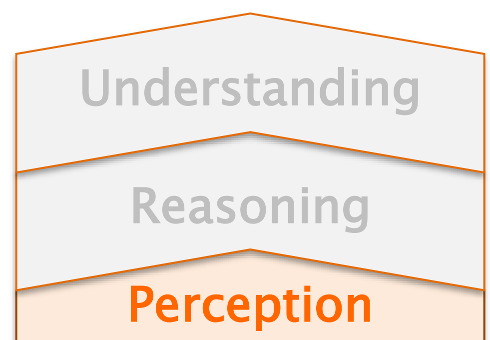
VISUALIZATION PIPELINE

Visualization Pipeline



Visual Perception

- Any variable (measure) must be visually encoded, i.e. we need to identify:
  - Visual object to represent entity
  - Visual attribute to represent the measure



Example

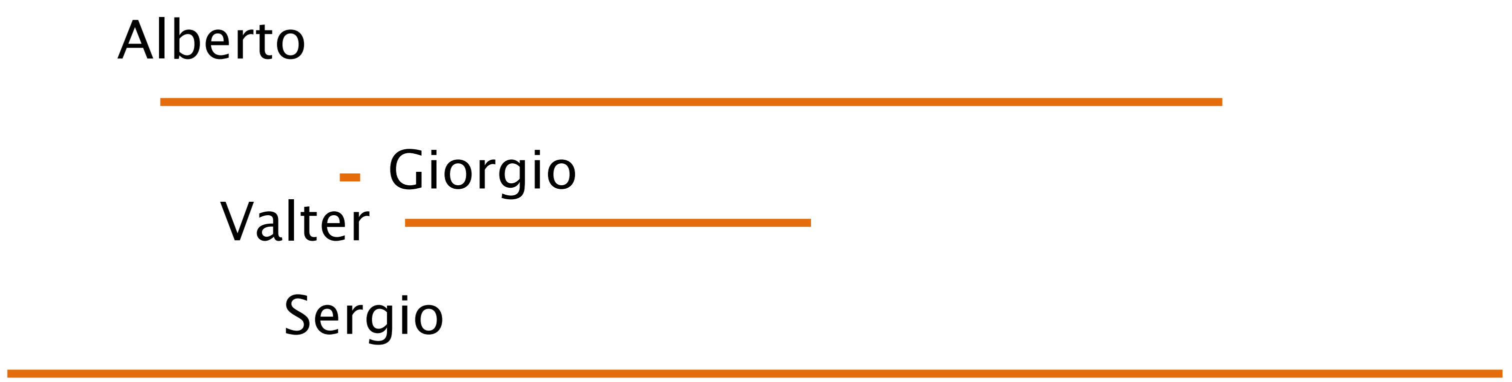
Votes received by four candidates in recent elections

Candidate	Votes	Proportion
Sergio	197800	50.09%
Alberto	140545	35.59%
Giorgio	53748	13.61%
Valter	2759	0.70%

http://www.comune.torino.it/elezioni/2019/regionali/presidente/citta/

Encoding

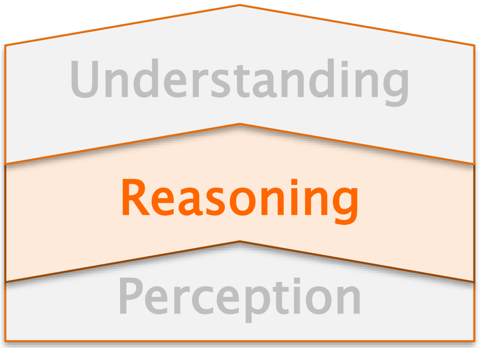
- Visual object: line
- Visual attribute: length



Visual Reasoning

Layout and visual attributes allow:

- Discrimination
  - Distinguish visual objects or group of –
- Comparison
  - Place visual objects in order
- Magnitude assessment
  - Evaluate the (relative) magnitude of visual objects



Reasoning

Alberto

- Giorgio

Valter

Sergio

Reasoning

- Discrimination

Reasoning

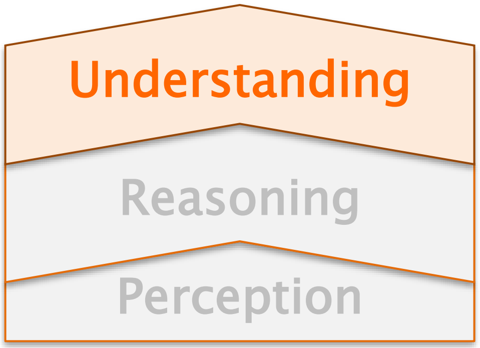
- Comparison

Reasoning

- Assessment

Understanding

- Variation within quantitative measures
  - Distribution
  - Deviation
  - Correlation
- Variation within category
  - Ranking
  - Part-to-whole
  - Time
  - Space
- Multivariate



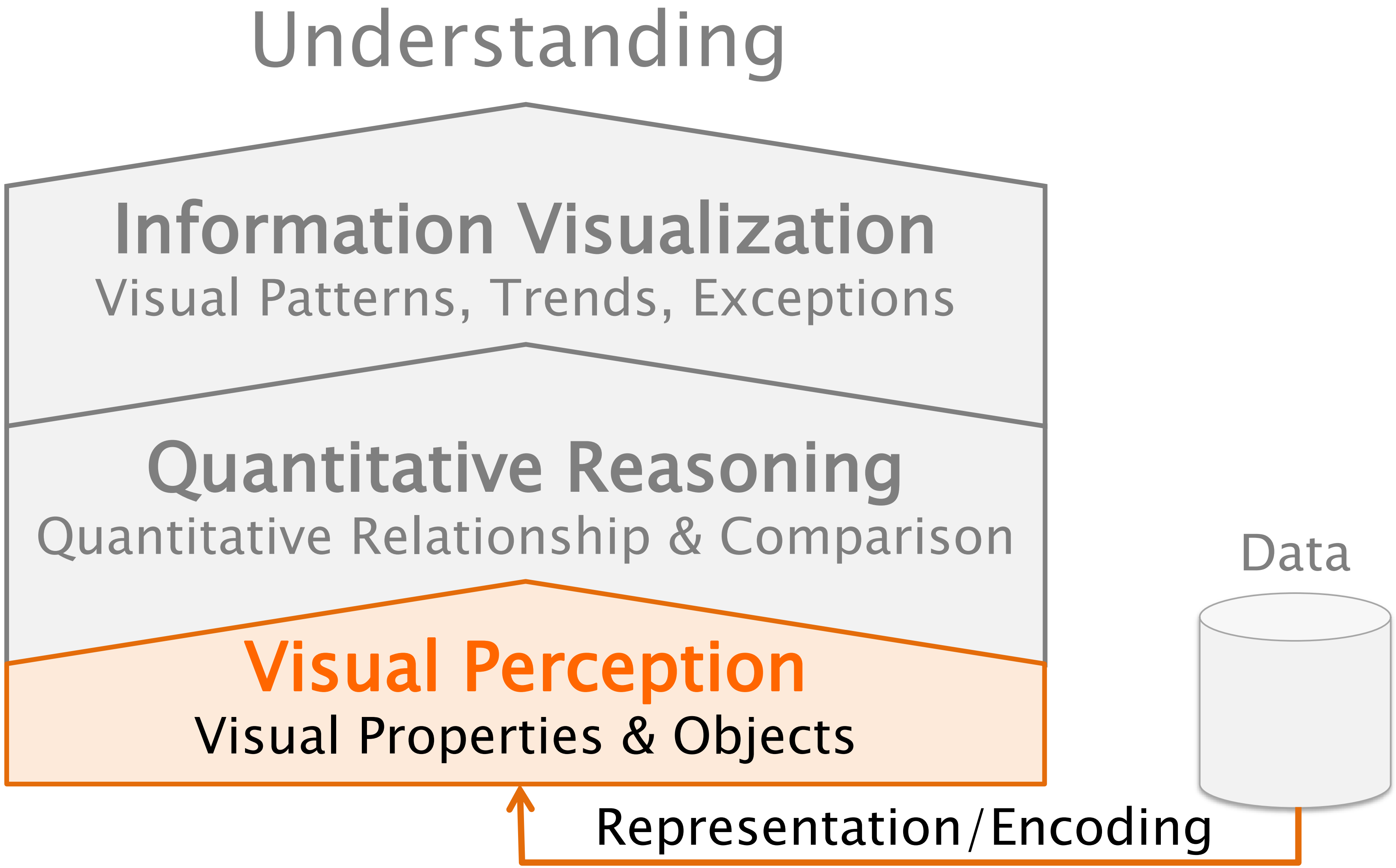
Understanding

Understanding

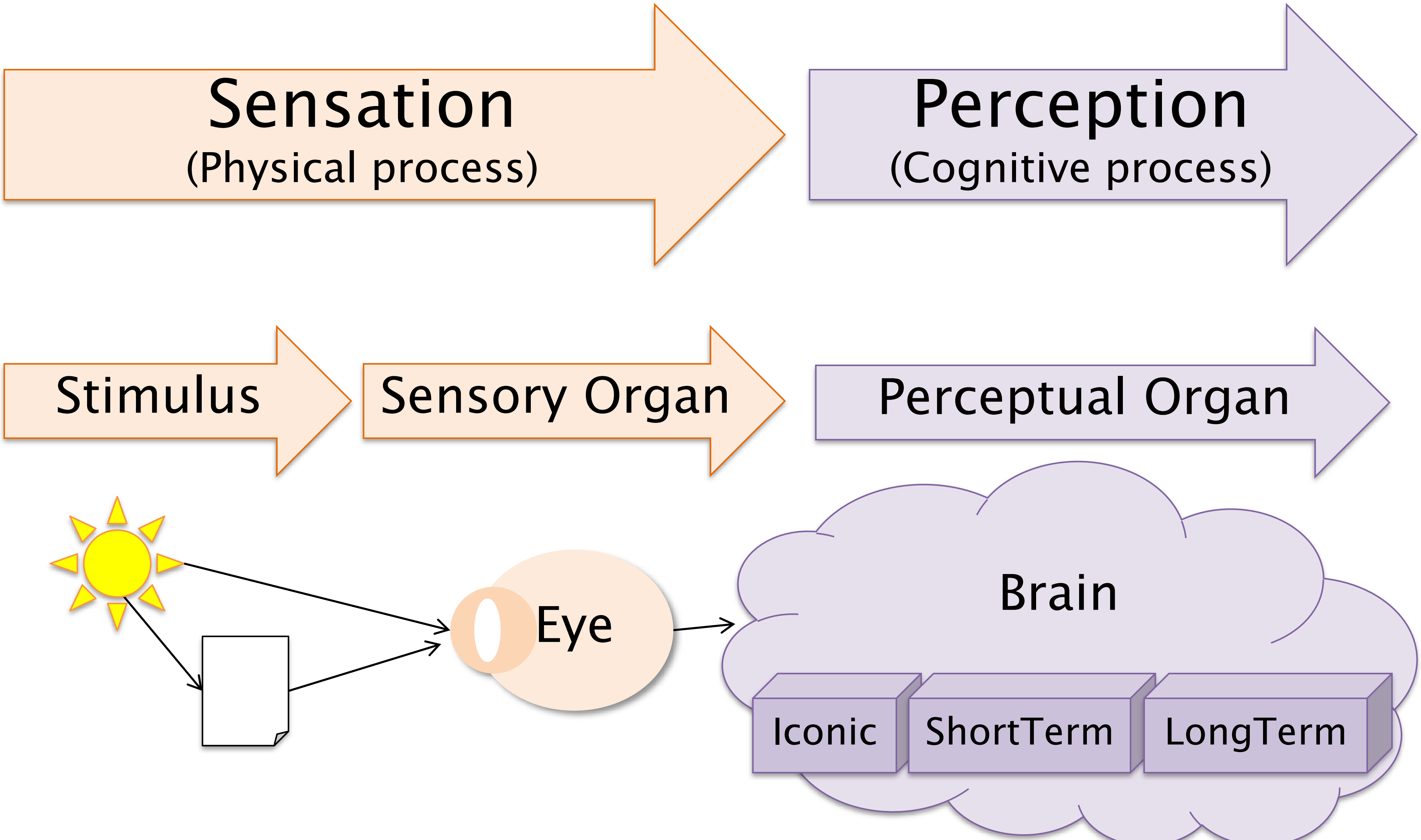
- Ranking

VISUAL PERCEPTION

Data Visualization



Visual perception



Memory Hierarchy

- Iconic memory (visual sensory register)
  - Pre-attentive processing
  - Detects a limited number of attributes
- Short-term memory (working memory)
  - Store visual chunks
  - Limited number
- Long-term memory
  - Store high-level knowledge

Simplified Model

- The three levels of memory represent a simplified model
  - does not correspond to “real” physical brain structure
- Useful to explain a few phenomena
  - The 7 ± 2 rule
  - Change blindness

Change blindness

http://www2.psych.ubc.ca/~rensink/flicker/download/index.html

Change blindness

http://www2.psych.ubc.ca/~rensink/flicker/download/index.html

Pre-Attentive Attributes

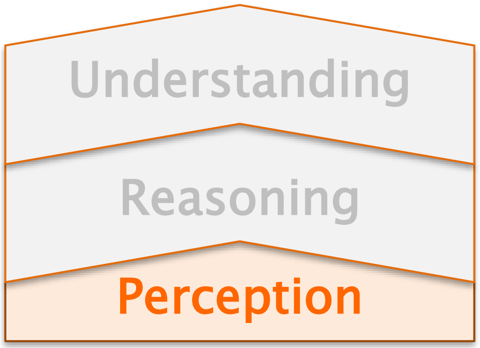
5 7 8 4 9 8 3 1 1 0 6 8 8 2 1 1 5 2 6 6 5
9 5 1 8 4 6 8 4 9 3 0 4 5 3 4 9 2 5 8 5 8
5 0 5 4 6 2 6 5 7 3 7 8 6 5 3 7 2 6 3 1 5
5 8 6 6 8 3 7 6 5 0 9 6 3 4 6 1 9 5 6 6 4
1 6 7 3 9 9 2 8 3 4 0 3 5 1 6 3 5 3 9 3 4
8 6 9 7 5 4 2 4 7 4 9 5 8 5 3 0 7 6 0 6 7
0 3 1 5 3 2 3 5 6 7 2 8 9 8 5 3 7 8 8 2 4
5 5 3 4 8 1 5 6 2 3 5 5 1 2 1 0 8 7 2 6 3
7 4 3 8 4 8 2 6 7 9 5 6 2 3 6 7 8 0 8 3 6
4 9 5 6 7 2 2 2 8 3 1 1 0 1 8 6 2 6 2 1 4

Pre-Attentive Attributes

5 7 8 4 9 8 3 1 1 0 6 8 8 2 1 1 **5** 2 6 6 **5**
9 **5** 1 8 4 6 8 4 9 3 0 4 **5** 3 4 9 2 **5** 8 **5** 8
5 0 **5** 4 6 2 6 **5** 7 3 7 8 6 **5** 3 7 2 6 3 1 **5**
5 8 6 6 8 3 7 6 **5** 0 9 6 3 4 6 1 9 **5** 6 6 4
1 6 7 3 9 9 2 8 3 4 0 3 **5** 1 6 3 **5** 3 9 3 4
8 6 9 7 **5** 4 2 4 7 4 9 **5** 8 **5** 3 0 7 6 0 6 7
0 3 1 **5** 3 2 3 **5** 6 7 2 8 9 8 **5** 3 7 8 8 2 4
5 **5** 3 4 8 1 **5** 6 2 3 **5** **5** 1 2 1 0 8 7 2 6 3
7 4 3 8 4 8 2 6 7 9 **5** 6 2 3 6 7 8 0 8 3 6
4 9 **5** 6 7 2 2 2 8 3 1 1 0 1 8 6 2 6 2 1 4

Encoding

- Encoding is the key to enable visual perception
  - Visual object to represent entity
  - Visual attribute to represent the measure
- Two main types
  - Quantitative (different properties)
  - Categorical (ordinal or not)



Pre-Attentive attributes

Category	Attribute
Form	Orientation Length/distance Line width Size Shape Curvature Added marks Enclosure
Color	Hue Intensity
Spatial position	2-D position
Motion	Flicker Direction Speed

Perception task

Visual attributes allow:

- Discrimination
  - Distinguish visual objects
- Comparison
  - Place visual objects in order
- Magnitude assessment
  - Evaluate the (relative) magnitude of visual objects

Just noticeable difference

- Given a physical dimension (length, brightness, etc.) x
- *d* is the just noticeable difference if:
  - difference between *x* and *x+d* is perceivable
  - but not smaller differences
- *d* depends on many factors:
  - Subject
  - Environment
  - Physical dimension

Weber's law

- Just noticeable difference *d* is:

$$d_p(x) = k_p \cdot x$$

- Where
  - x: dimension
  - $d_p(x)$: just noticeable difference
  - k_p: constant
    - Subjective
    - Environmental

Consequences of Weber's law

- It is easier to compare lengths that differ by a large percentage
- The same difference is easier to notice between smaller measures
  - More likely to be larger than just noticeable difference

$$x < y \implies d_p(x) < d_p(y)$$

- Length of non-aligned objects is harded to compare
  - Double comparison

Non-aligned objects lengths

- Additional references my help comparison
  - They provide alternative possible comparisons
- If lenghts range between 0 and a maximum (L) e.g percentages
- Comparing l_1 and l_2 (close to L) that differ by a small amount d
  - Difference $L-l_1$ vs. $L-l_2$ easier to notice than l_1 vs. l_2

Stevens's law

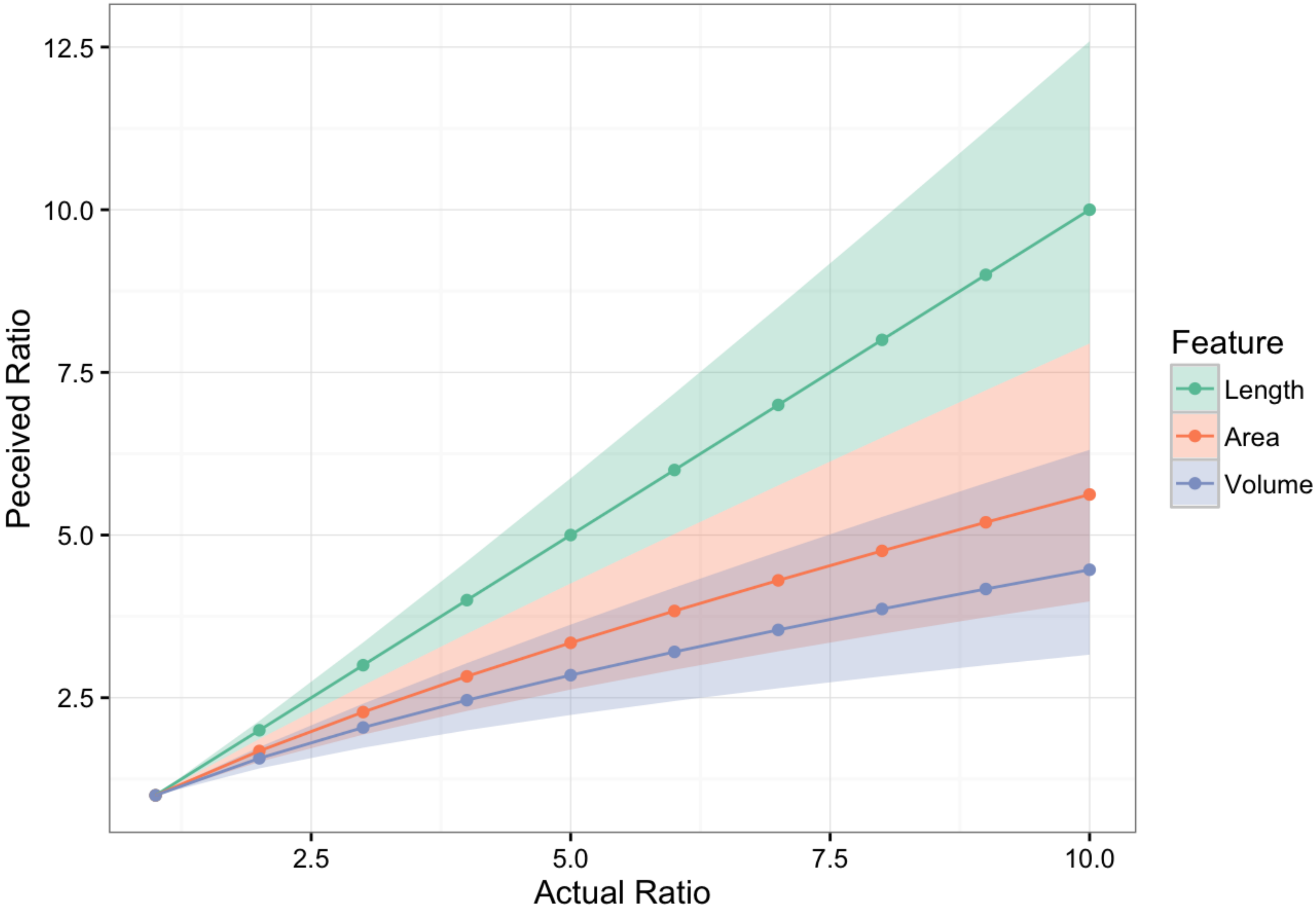
- Perceive scale (magnitude ratio)

$$p(x) = c \cdot x^{\beta}$$

- Where β depends on spatial dimension
  - 1D: Length → β in [0.9, 1.1]
  - 2D: Area → β in [0.6, 0.9]
  - 3D: Volume → β in [0.5, 0.8]

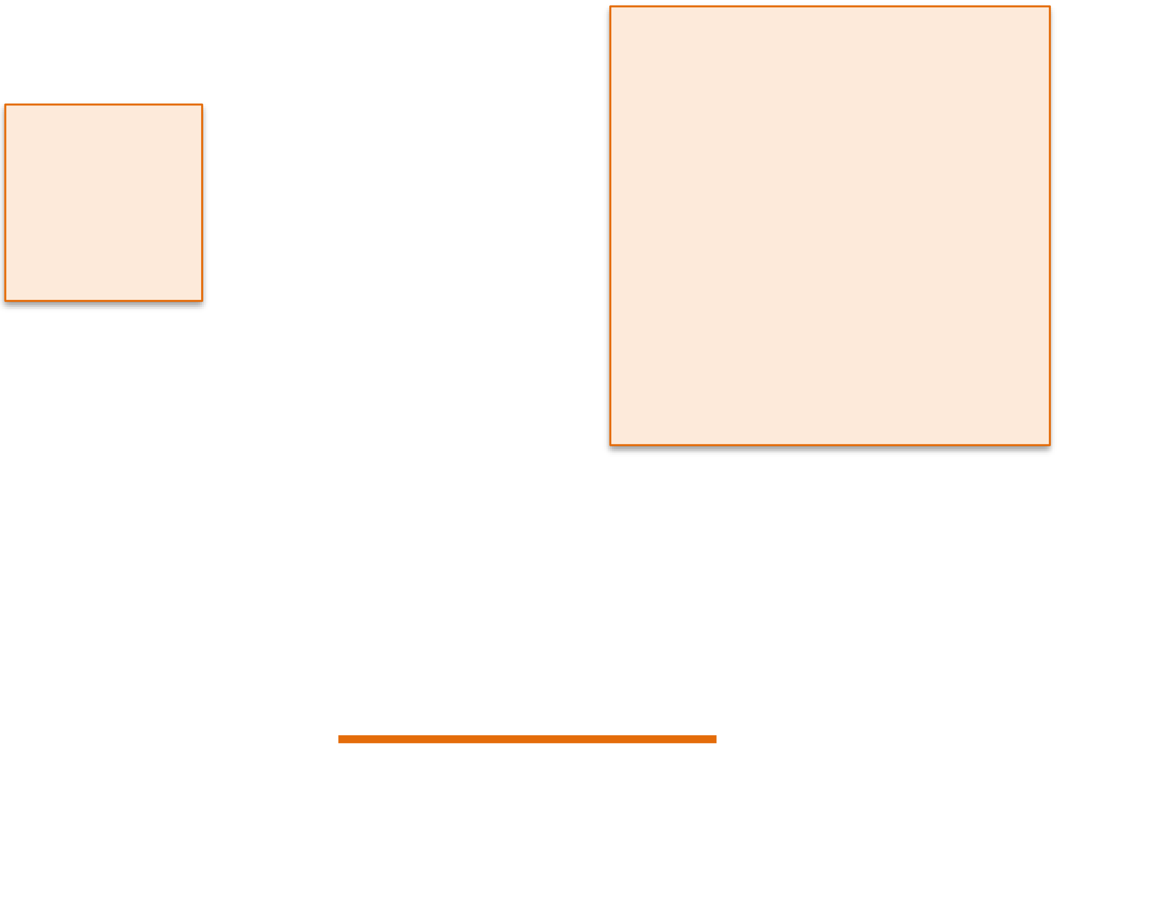
Stevens S. S. (1975). Psychophysics, John Wiley & Sons.

Steven's law



12.5
10.0
7.5
5.0
2.5
Peceived Ratio
2.5
5.0
7.5
10.0
Actual Ratio
Feature
Length
Area
Volume

Steven's law

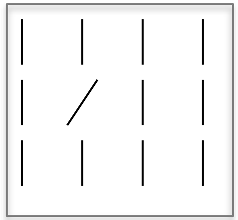


Consequences

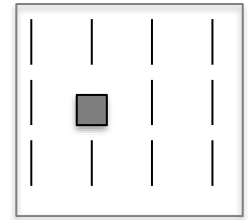
- Prefer comparing lengths
- Avoid comparison between areas
  - Except for ordinal measures
- Never-ever make volume comparisons

Attributes of form

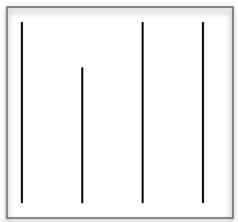
Orientation



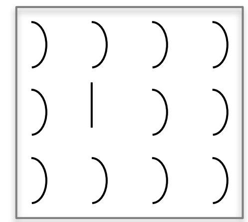
Shape



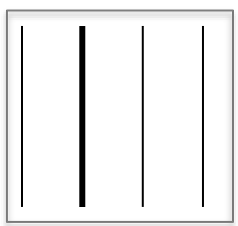
Line Length



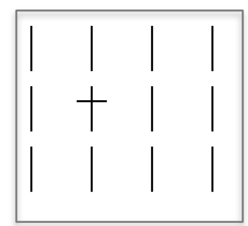
Curvature



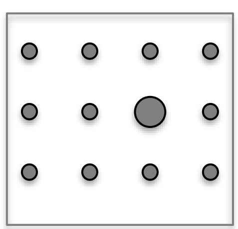
Line Width



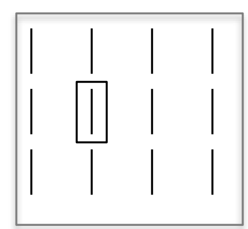
Added mark



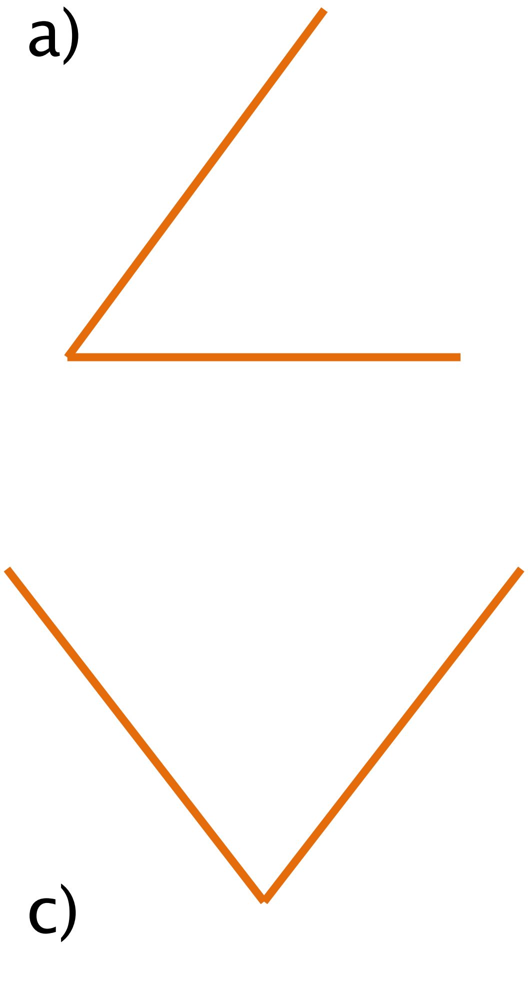
Size



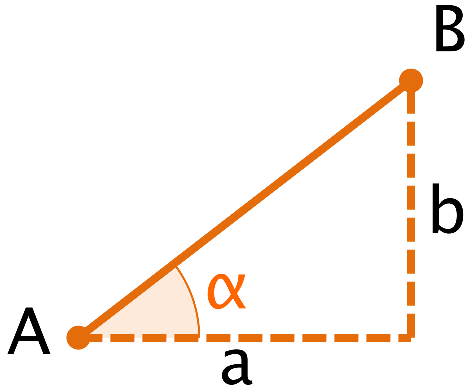
Enclosure



Orientation (angle or slope)



Angle vs. Slope



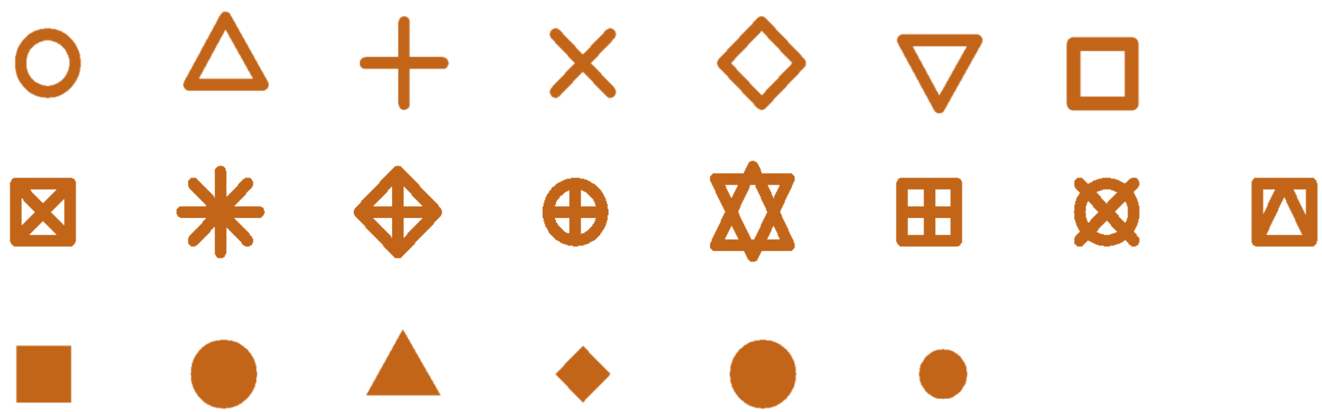
- Slope of A–B is b/a
  - tan(α)
- Slope judgment typically falls back to an angle judgment
  - Given an error ϵ in the angle judgment
  - It is reflected in a slope error

$$tan(\alpha + \epsilon) - tan(\alpha) = \epsilon \cdot tan'(\alpha) = \frac{\epsilon}{cos^2(\alpha)}$$

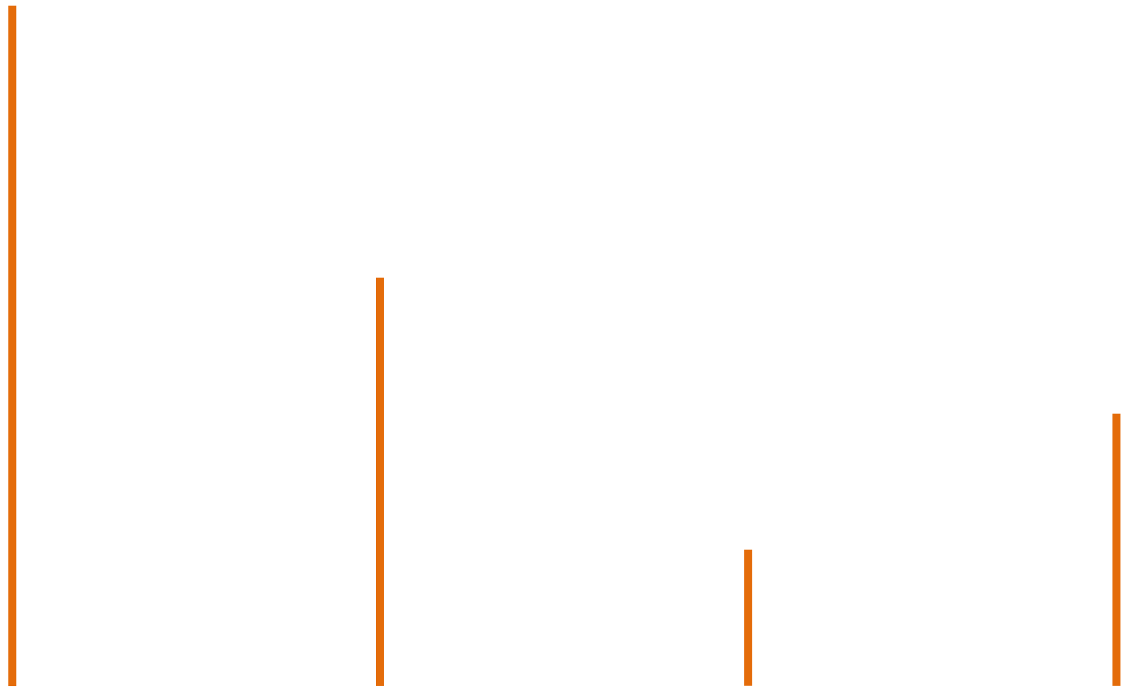
– Getting infinite as α approaches to π/2

Shape

- There is no common quantitative semantics for the shapes
  - – Unless they are characters…
  - ◆ Fill textures are shapes too



Length



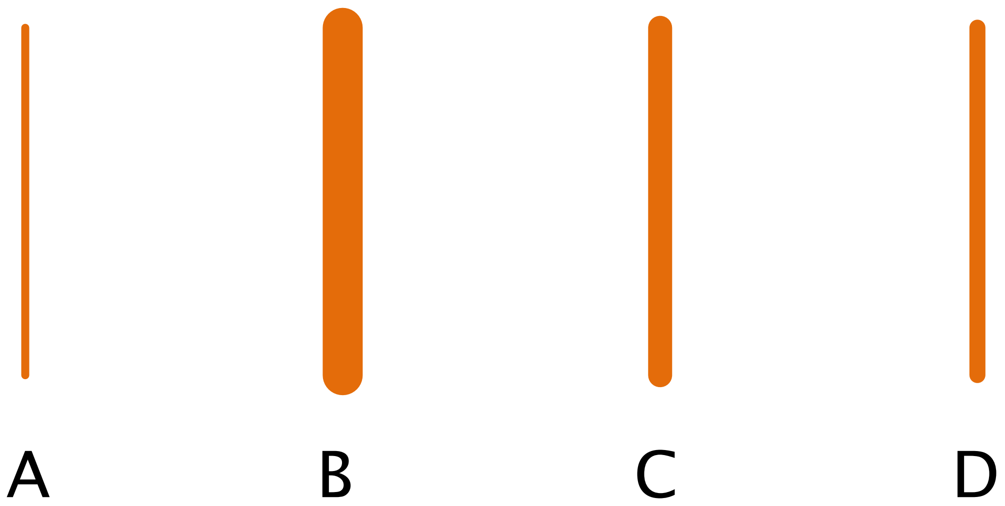
Effect of context

Curvature

- There is no common magnitude assessment for the curvature

Width

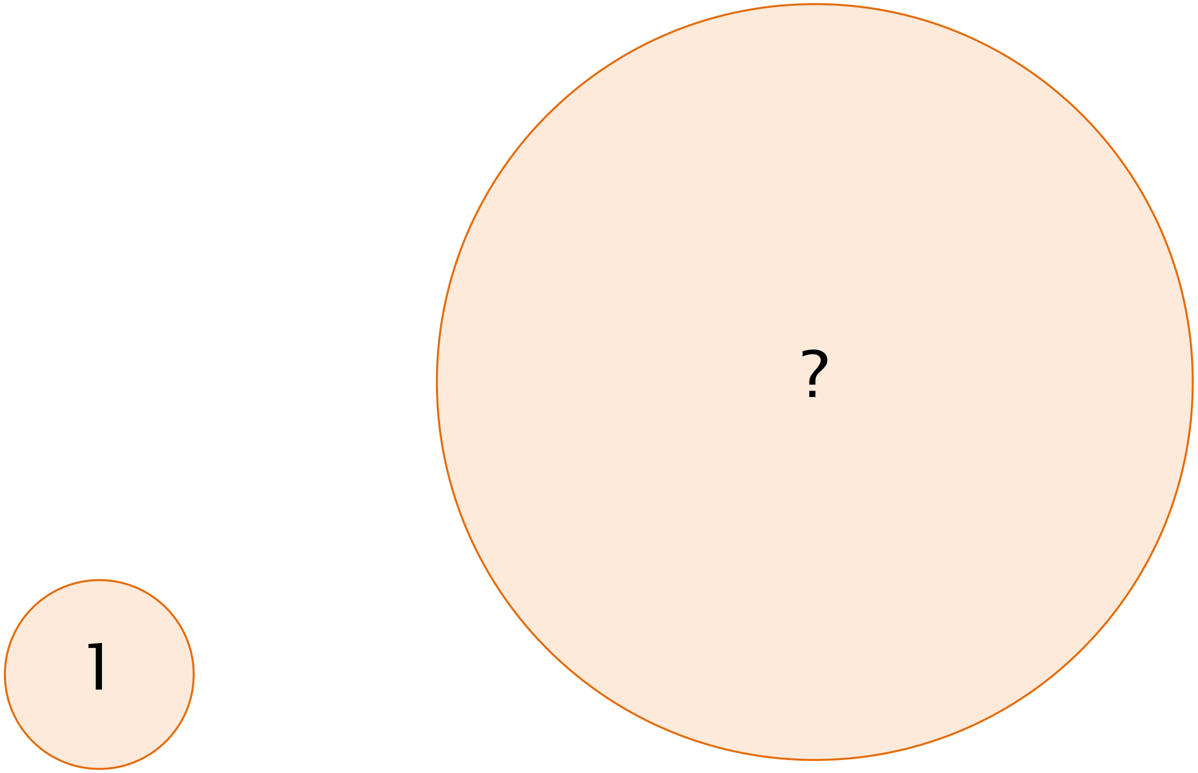
- Order can be identified
  - Difficult to appreciate actual magnitude



Mark

- No common quantitative semantics of marks
- Number of marks could encode a natural number
  - Harder to read than a cipher

Size / Area

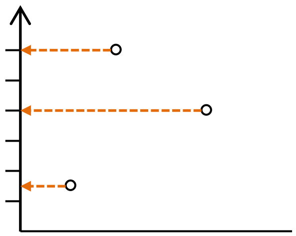


Enclosure

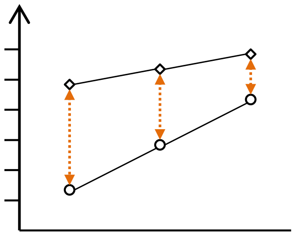
- No common quantitative semantics for enclosure
  - Except counting items enclosed

Spatial Position

- Position along axis
  - Common scale
  - Distinct identical scales
    - Possibly un-aligned

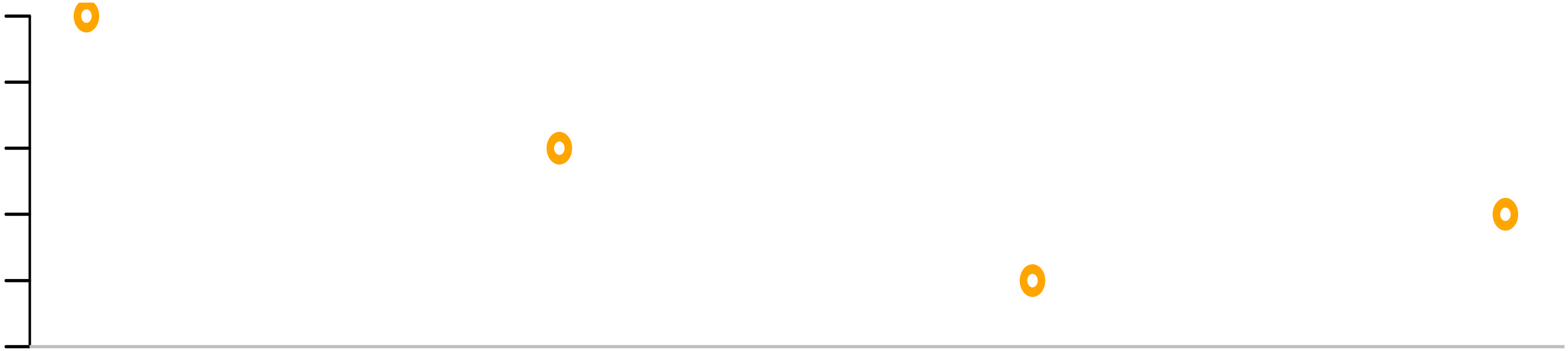


- Distance

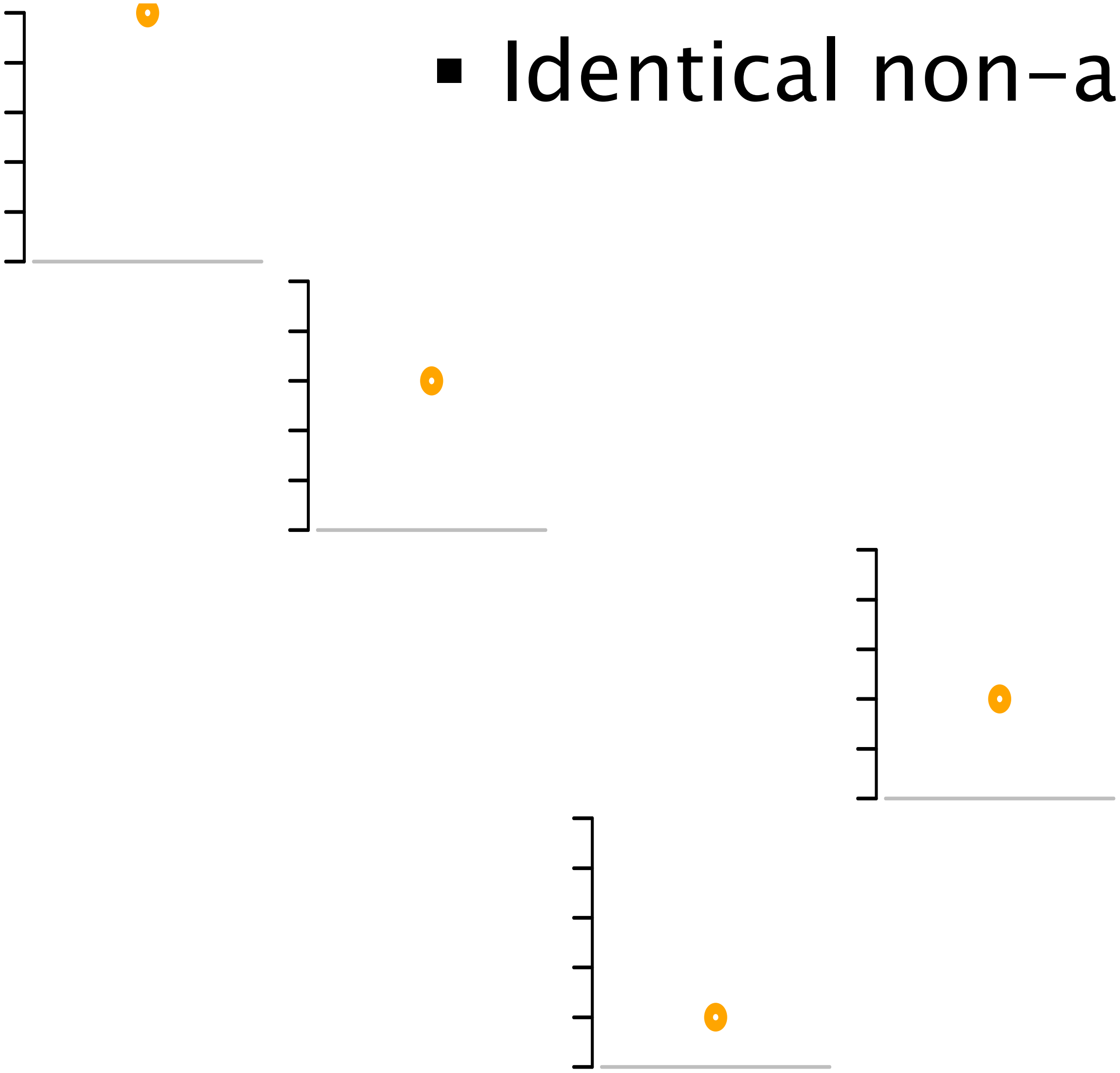


Position

- A common scale



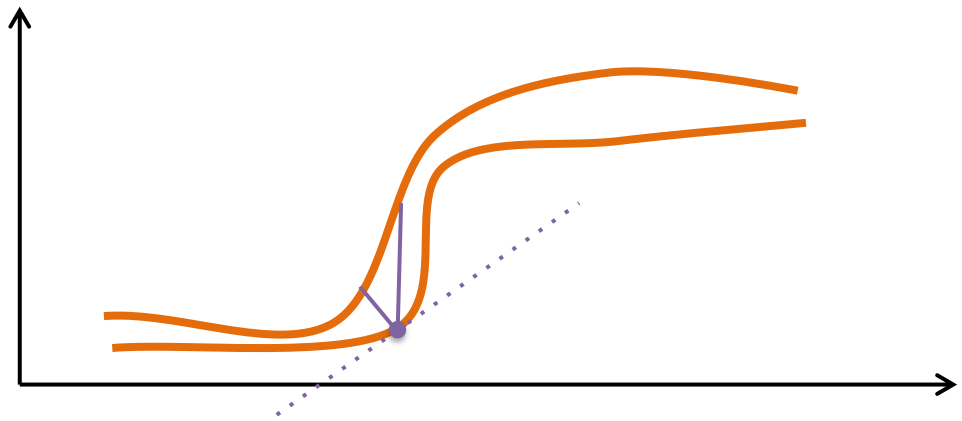
Position



- Identical non-aligned scales

Distance

- Points
  - Use length of imaginary connecting lines
- Lines
  - Distance orthogonal to tangent
    - Not what is meant in xy plots



Detection and Separation

Comparison is affected by:

- Detection
  - The capability to visually identify the objects that represent the data to be compared
- Separation
  - The distance between the objects to be compared
    - affects negatively the accuracy

Attributes of color

- Hue

- Saturation

- Intensity

  - Luminance
  - Value

Hue

- There is no common ordering semantics for hues
  - High spatial frequencies are perceived through intensity changes
  - Often perceived as separated into bands of almost constant hue, with sharp transitions between hues
- Nominal values can be represented by suitably spaced values

Intensity

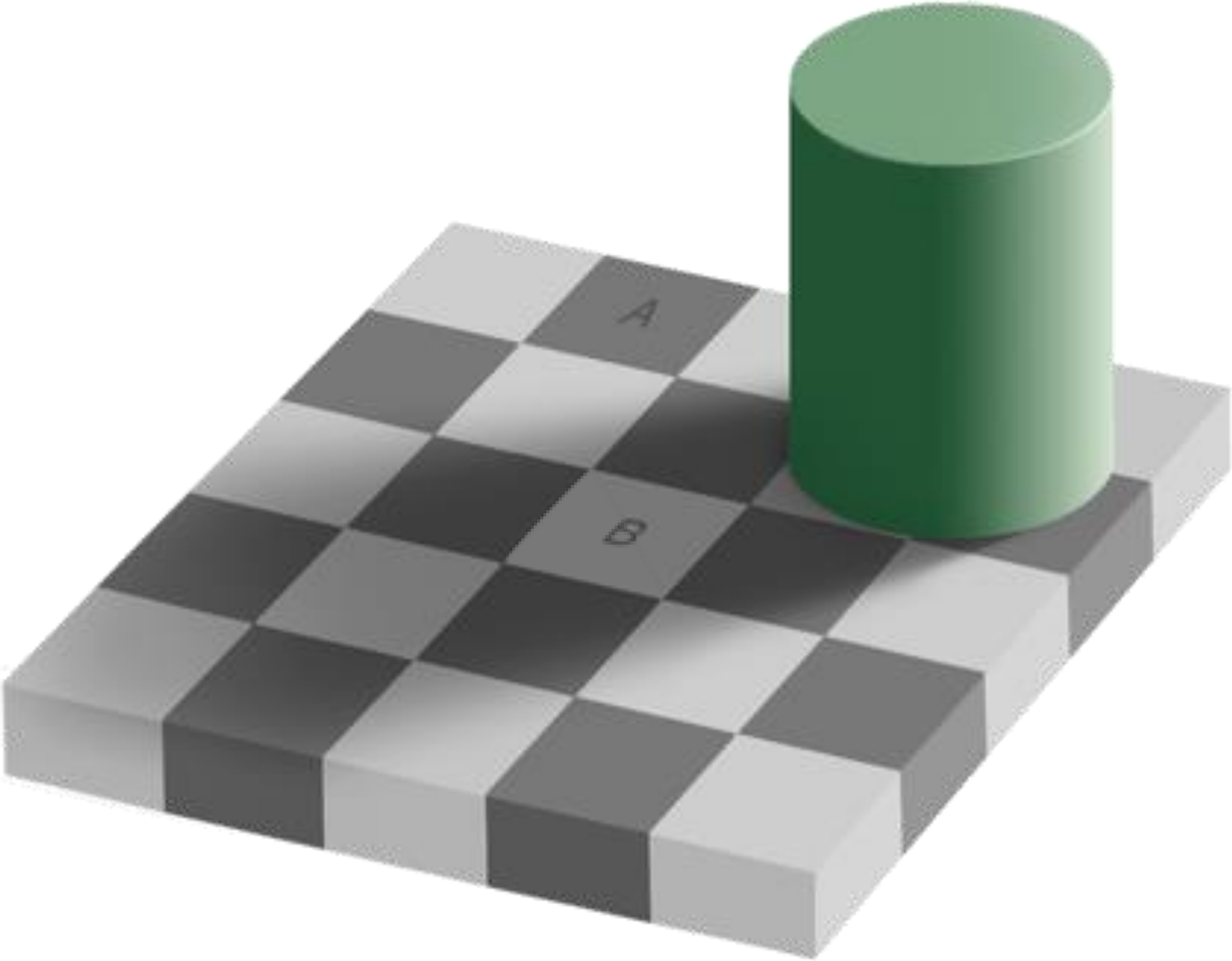
- a.k.a. Luminance, Value

- Provides a perceptually unambiguous ordering
  - Context can affect accuracy

Saturation

- Perceptually difficult to associate an ordered semantics
  - Can be combined with hue to increase discrimination

Effect of Context

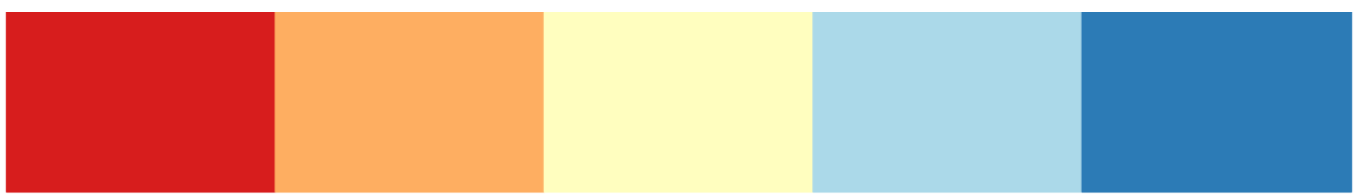


Effect of Context

- Use uniform background
  - To make distinct visual objects for the same feature look the same
- Use a background color that is contrasting enough with the visual objects' color
  - To make visual objects easily seen
- Avoid non-uniform background

Color usage

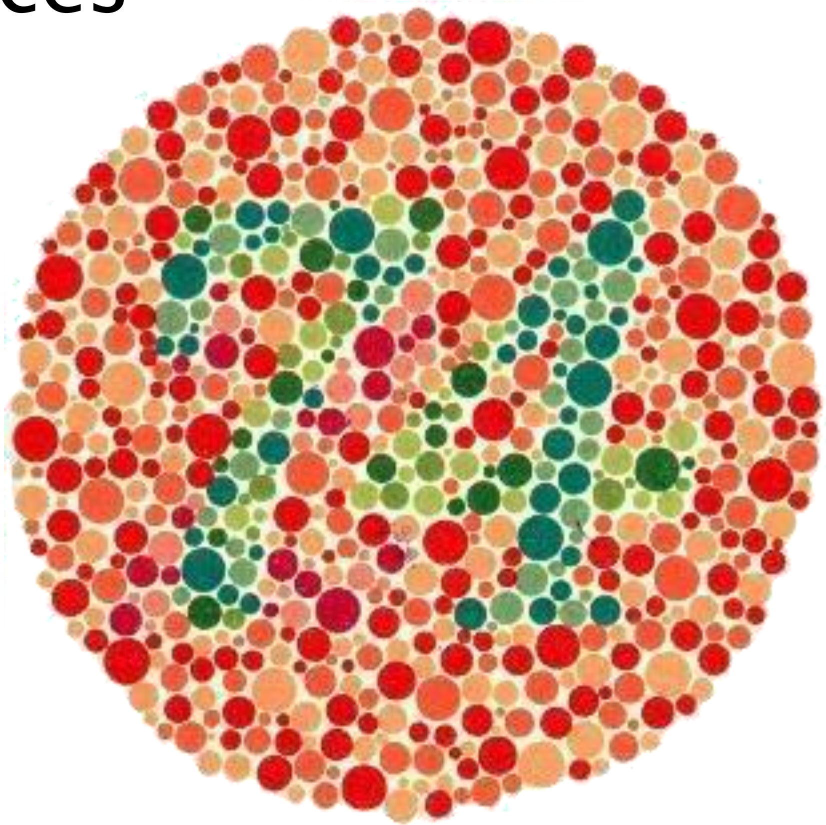
- Ordinal measure should be mapped to increasing saturation and intensity
  - Avoid rainbow palette
- Use sequential or diverging palette
  - E.g.



    – http://colorbrewer2.org/

Color Blindness

- Inability so see colors or perceive color differences



http://www.color-blindness.com

Visual Encoding: Quantitative

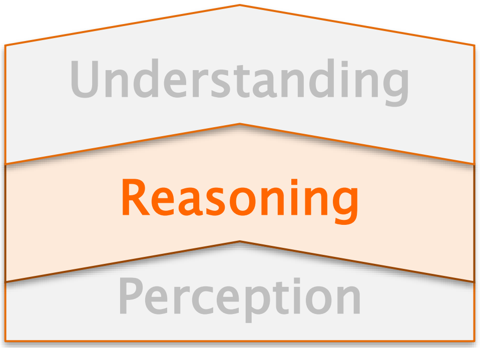
Object	Attribute
Point	Position (w.r.t. axis/axes)
Line	Length Position (w.r.t. axis/axes) Slope
Bar	Length
Shape	Size (area) Count

Visual Encoding: Categorical

Attribute	
Position	
Size	
Color	Intensity Saturation Hue
Shape	
Fill pattern	
Line style	

Ordinal (Position, Size, Color: Intensity, Saturation)

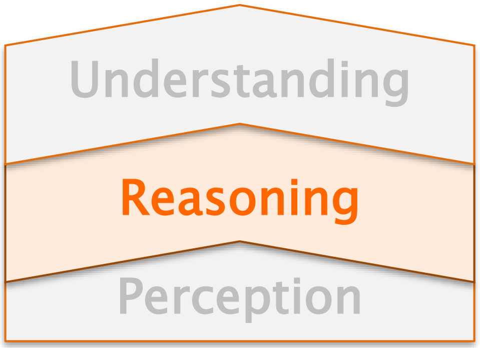
VISUAL REASONING



Graph layout

Layout and visual attributes allow:

- Discrimination
  - Distinguish visual objects or group of –
- Comparison
  - Place visual objects in order
- Magnitude assessment
  - Evaluate the (relative) magnitude of visual objects

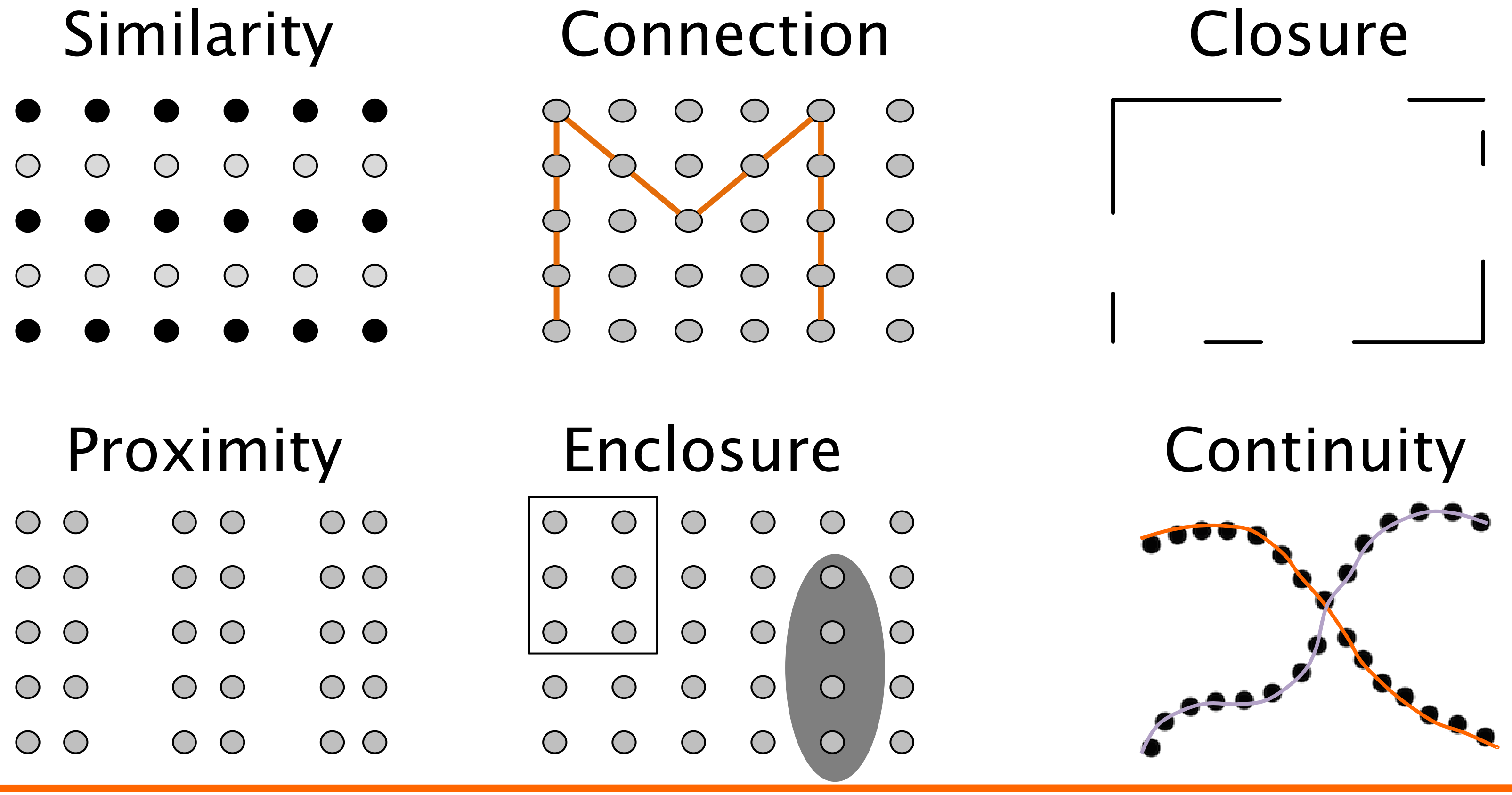


Gestalt principles

- Visual features that lead us to group visual objects together
  - Proximity
  - Similarity
  - Enclosure
  - Closure
  - Continuity
  - Connection

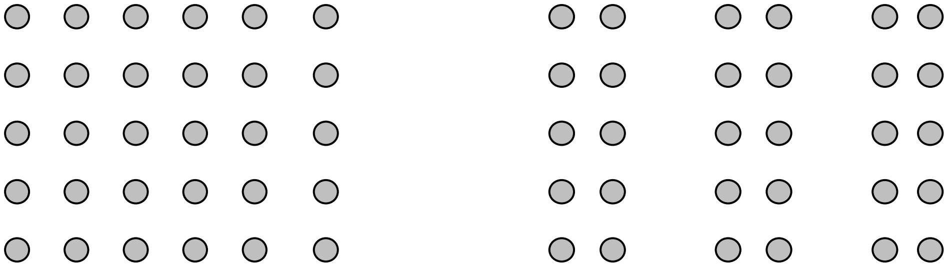
Gestalt principles

- Visual features that lead the viewer to group visual objects together



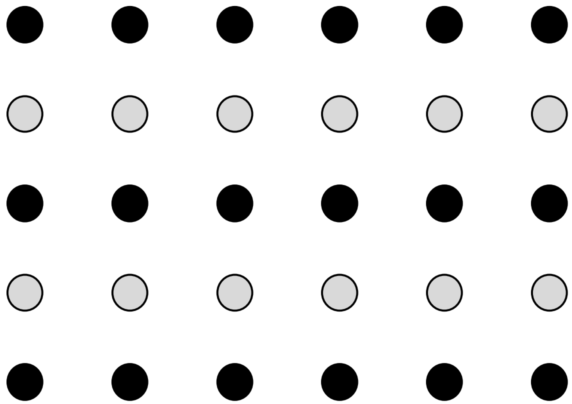
Gestalt principles

- Visual attributes/patterns that lead observer to group objects together
  - **Proximity**
  - Similarity
  - Enclosure
  - Closure
  - Continuity
  - Connection



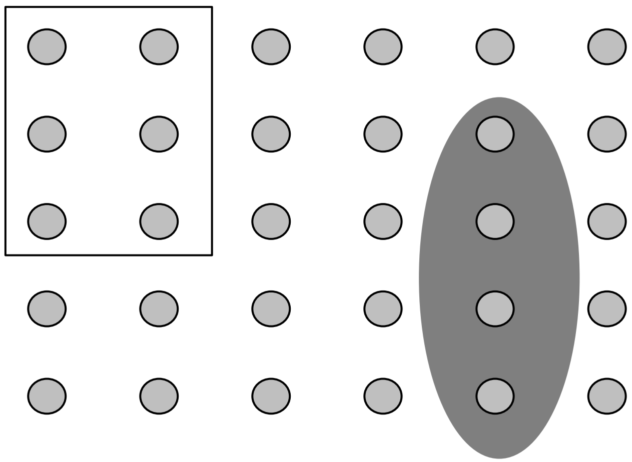
Gestalt principles

- Visual attributes/patterns that lead observer to group objects together
  - Proximity
  - **Similarity**
  - Enclosure
  - Closure
  - Continuity
  - Connection



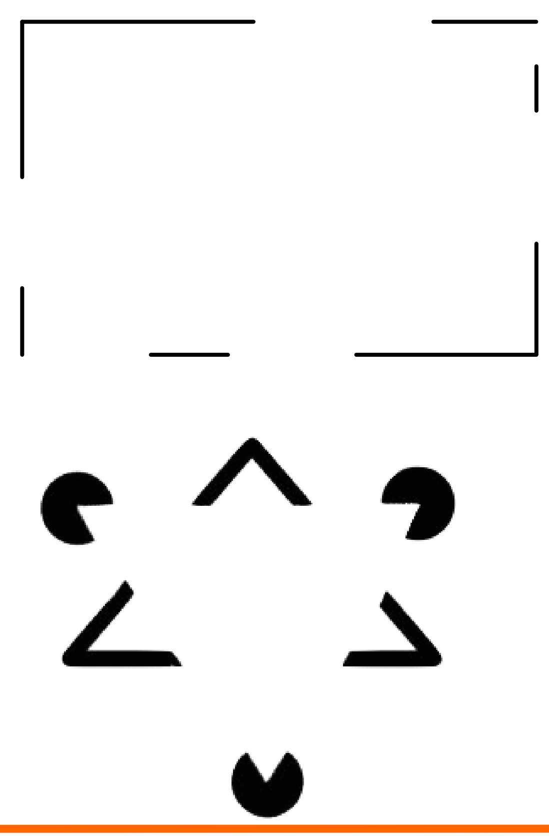
Gestalt principles

- Visual attributes/patterns that lead observer to group objects together
  - Proximity
  - Similarity
  - **Enclosure**
  - Closure
  - Continuity
  - Connection



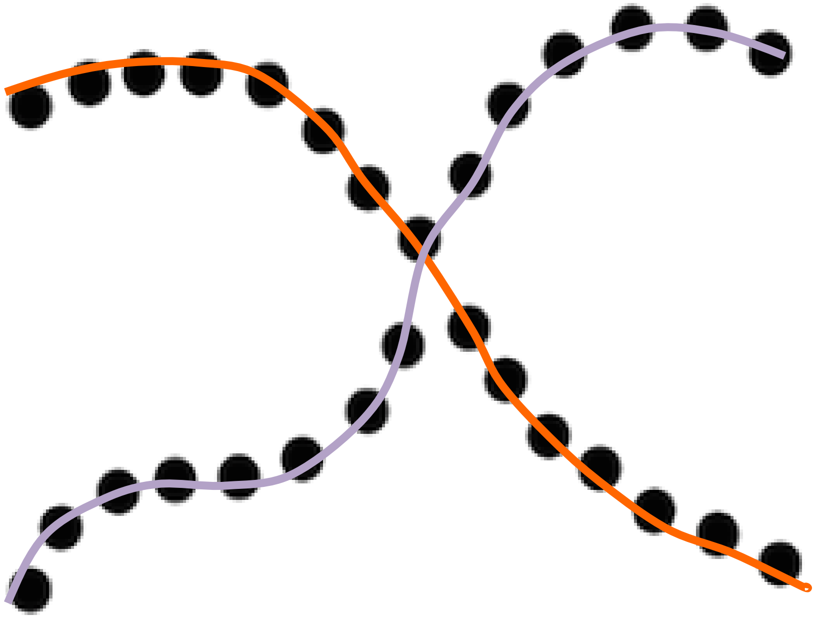
Gestalt principles

- Visual attributes/patterns that lead observer to group objects together
  - Proximity
  - Similarity
  - Enclosure
  - **Closure**
  - Continuity
  - Connection



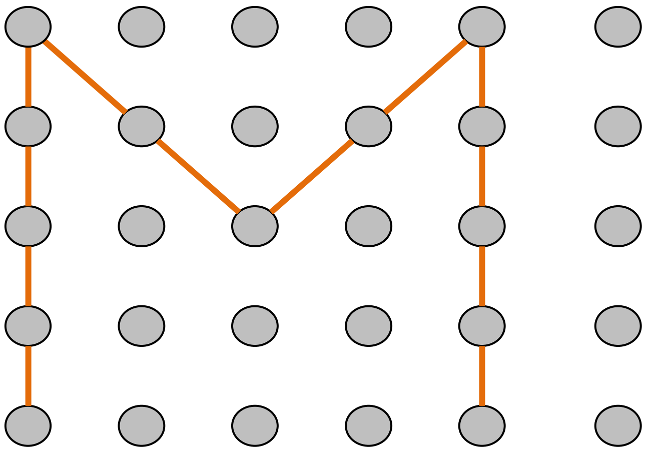
Gestalt principles

- Visual attributes/patterns that lead observer to group objects together
  - Proximity
  - Similarity
  - Enclosure
  - Closure
  - **Continuity**
  - Connection

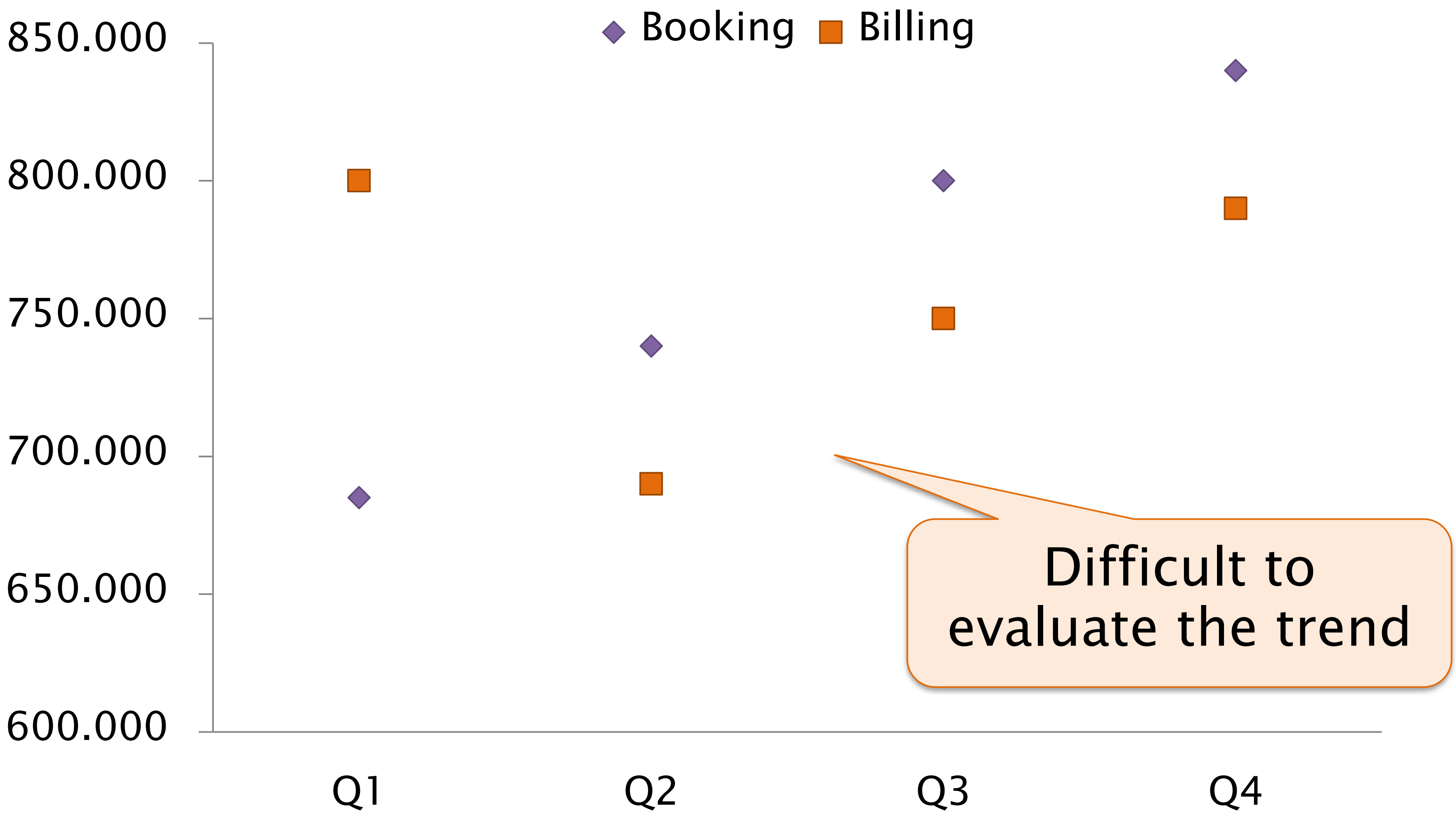


Gestalt principles

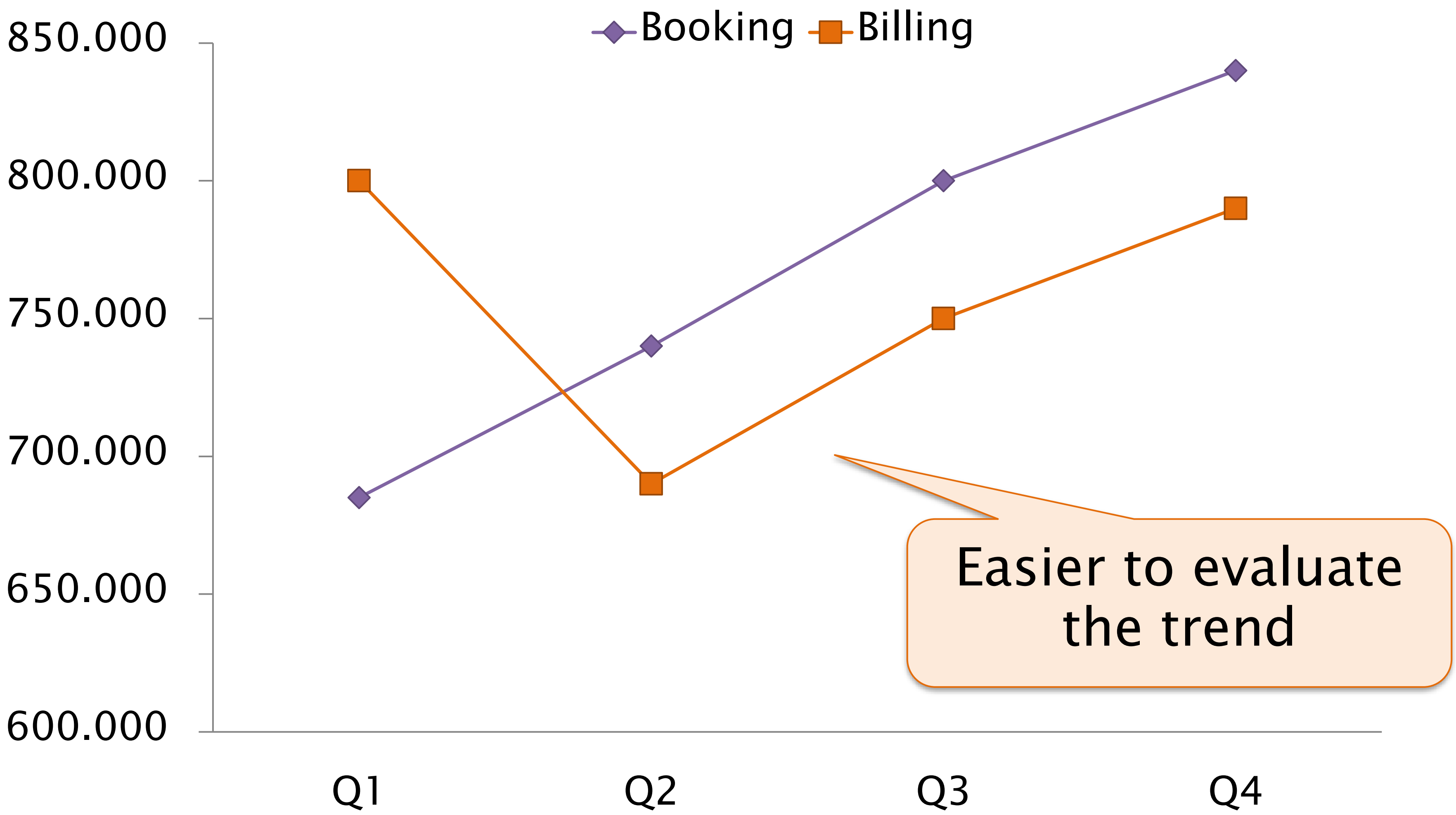
- Visual attributes/patterns that lead observer to group objects together
  - Proximity
  - Similarity
  - Enclosure
  - Closure
  - Continuity
  - **Connection**



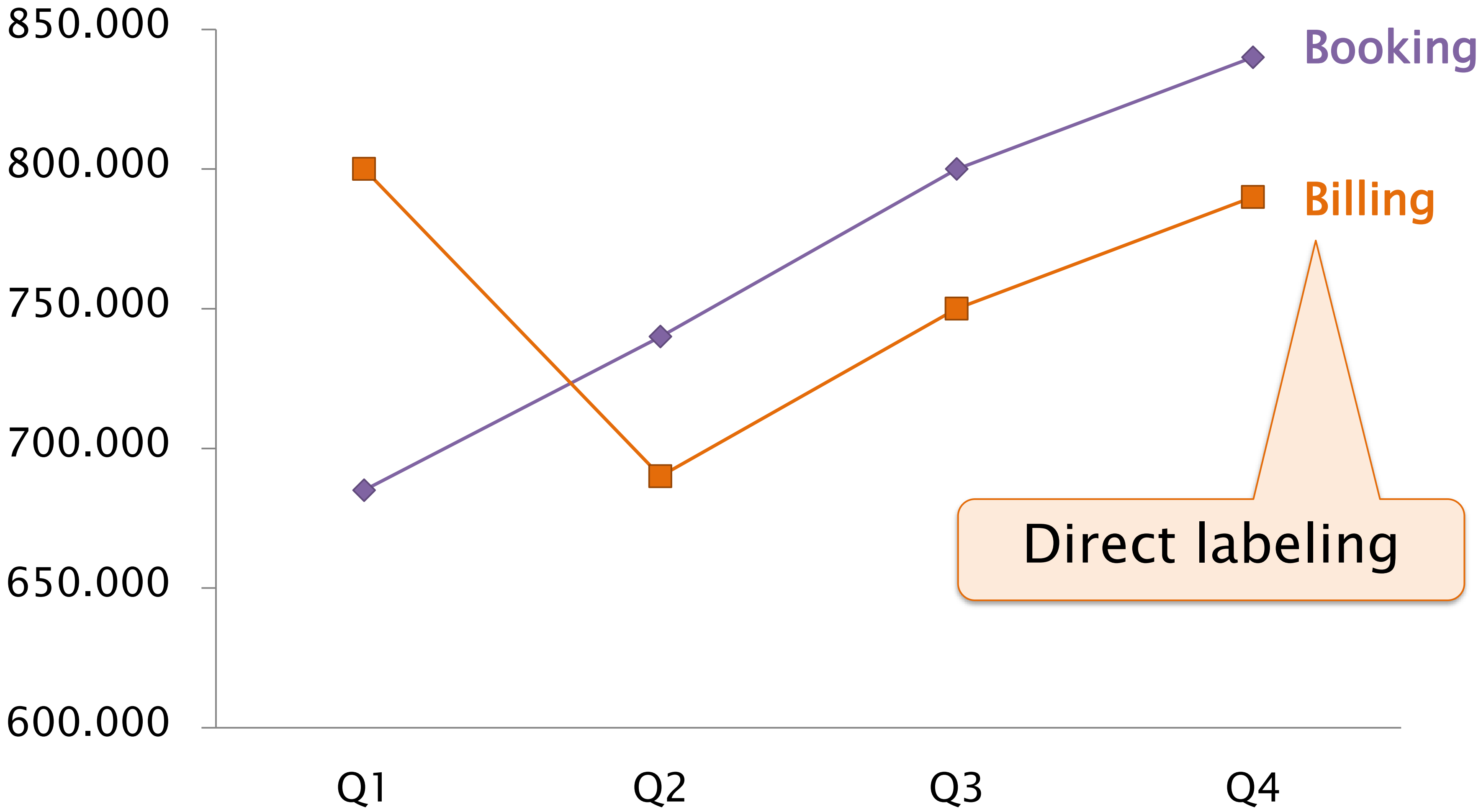
Similarity in Shape & Color



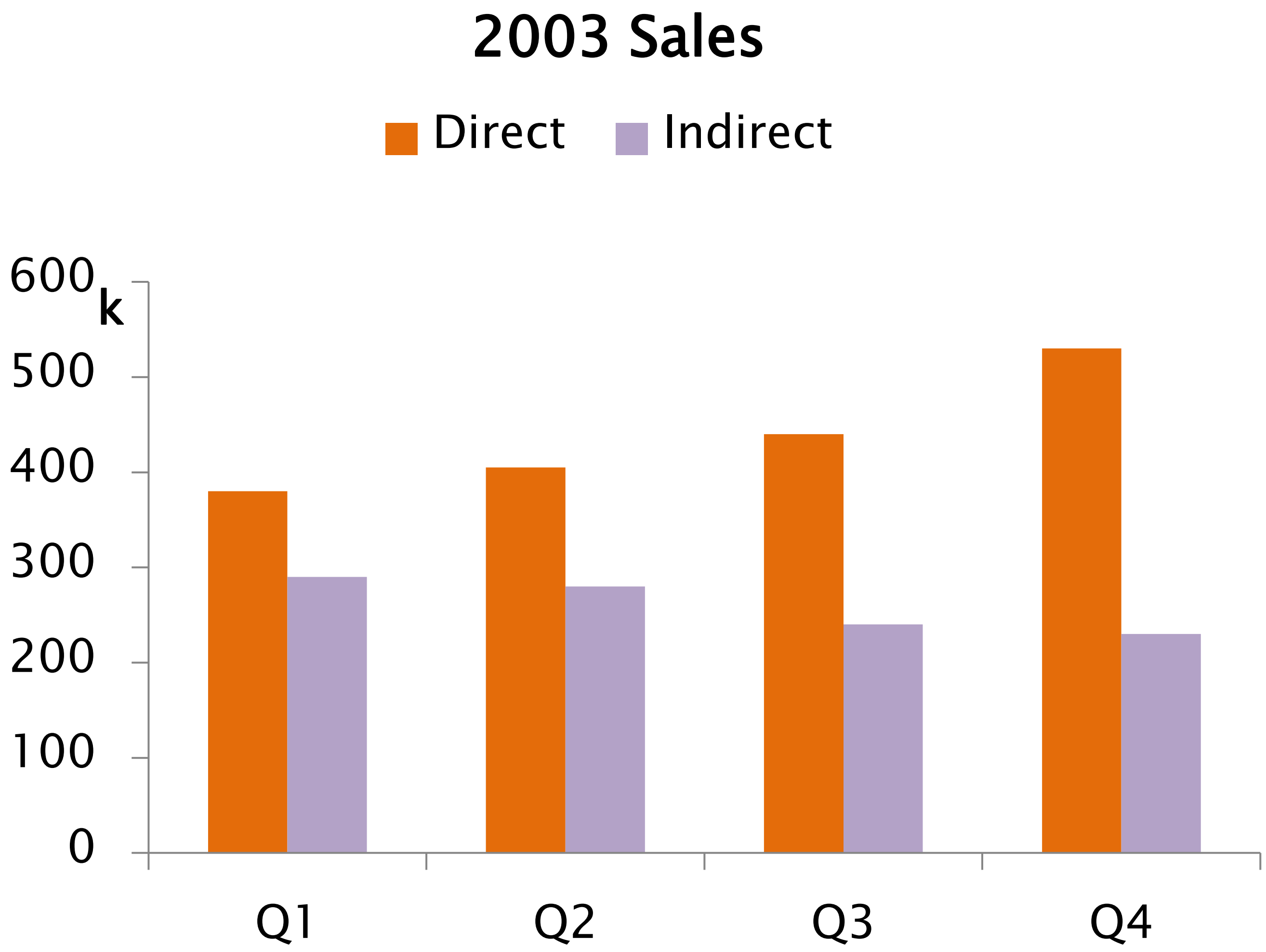
Similarity+Connection



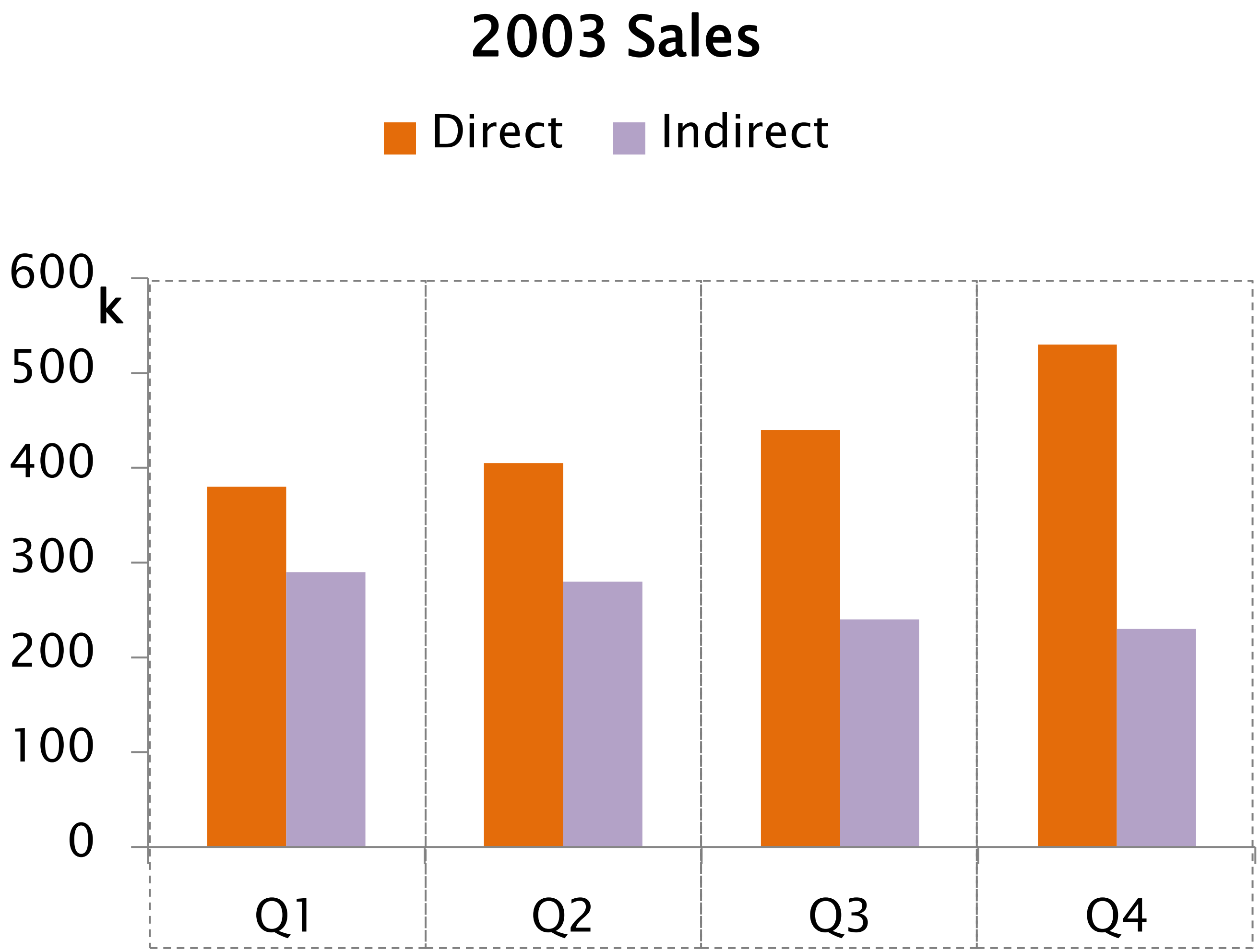
Similarity+Connection+Proximity



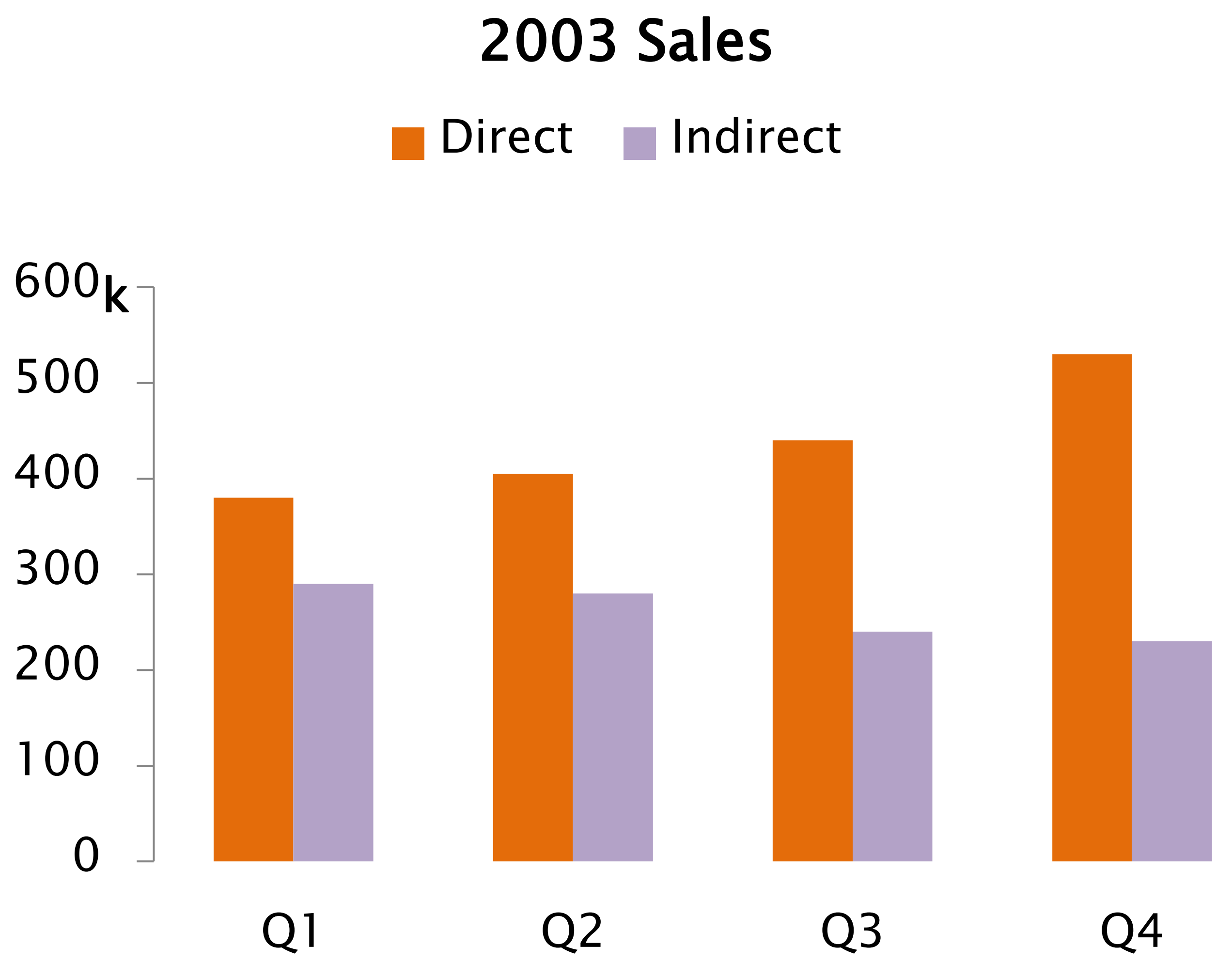
Similarity × Proximity



Similarity × Proximity & Enclosure



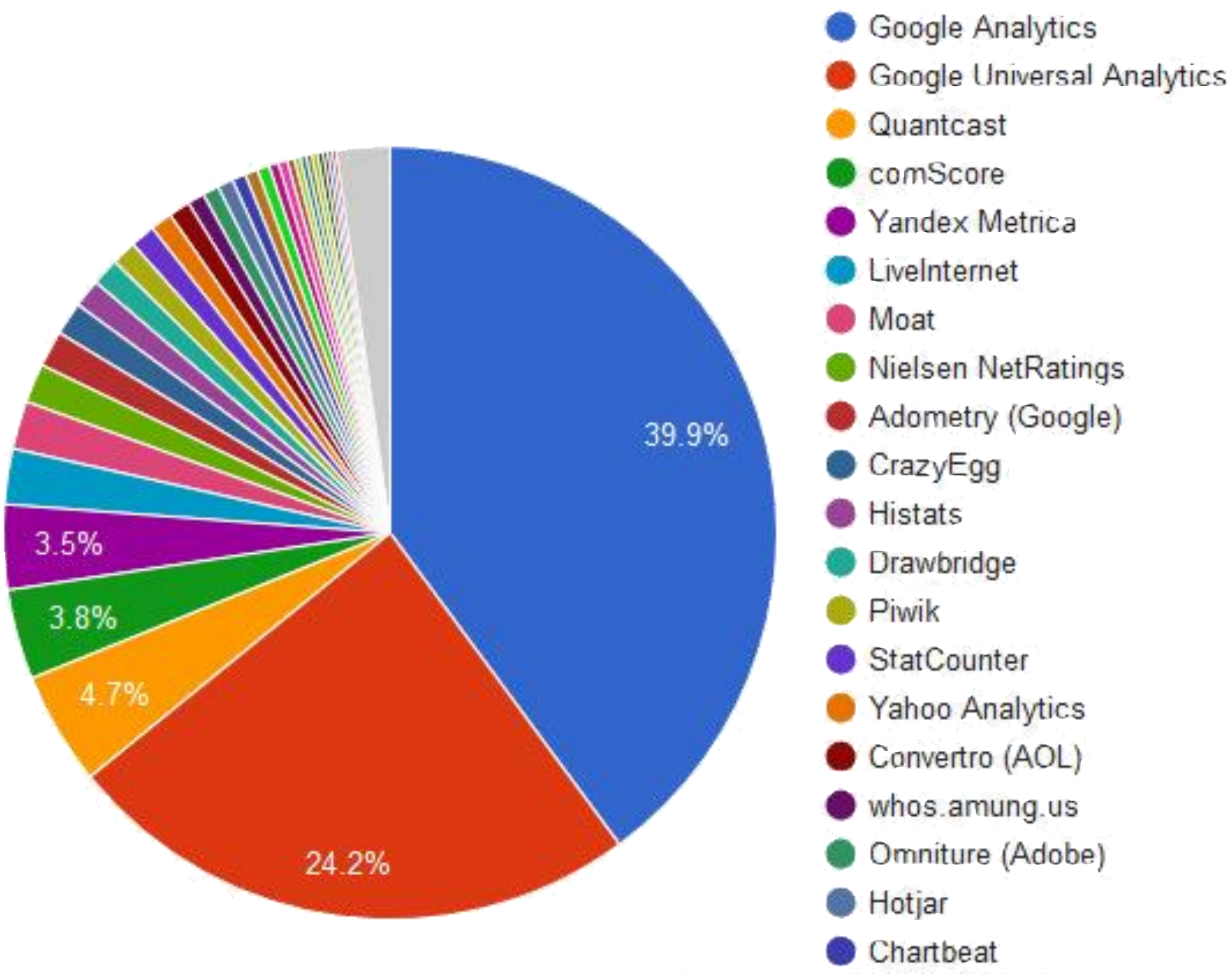
Continuity replaces axis



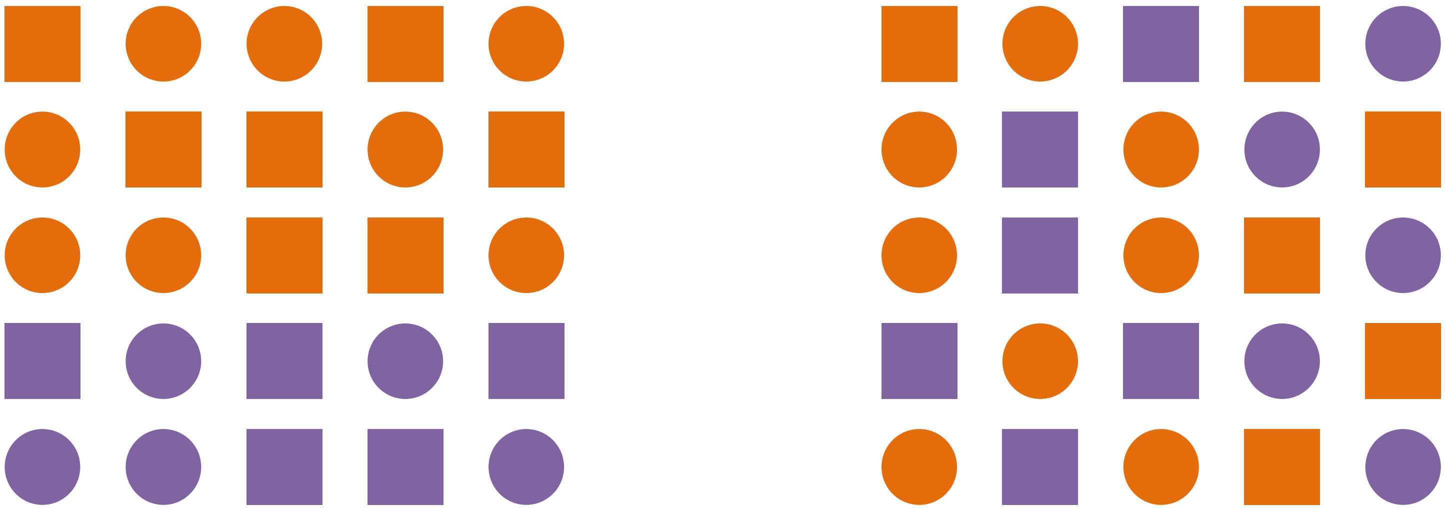
Distinct perceptions

- The immediacy of any pre-attentive cue declines as the variety of alternative patterns increases
  - Even if all the distracting patterns are individually distinct from the target
  - For each single attribute no more than four distinct levels are discernible

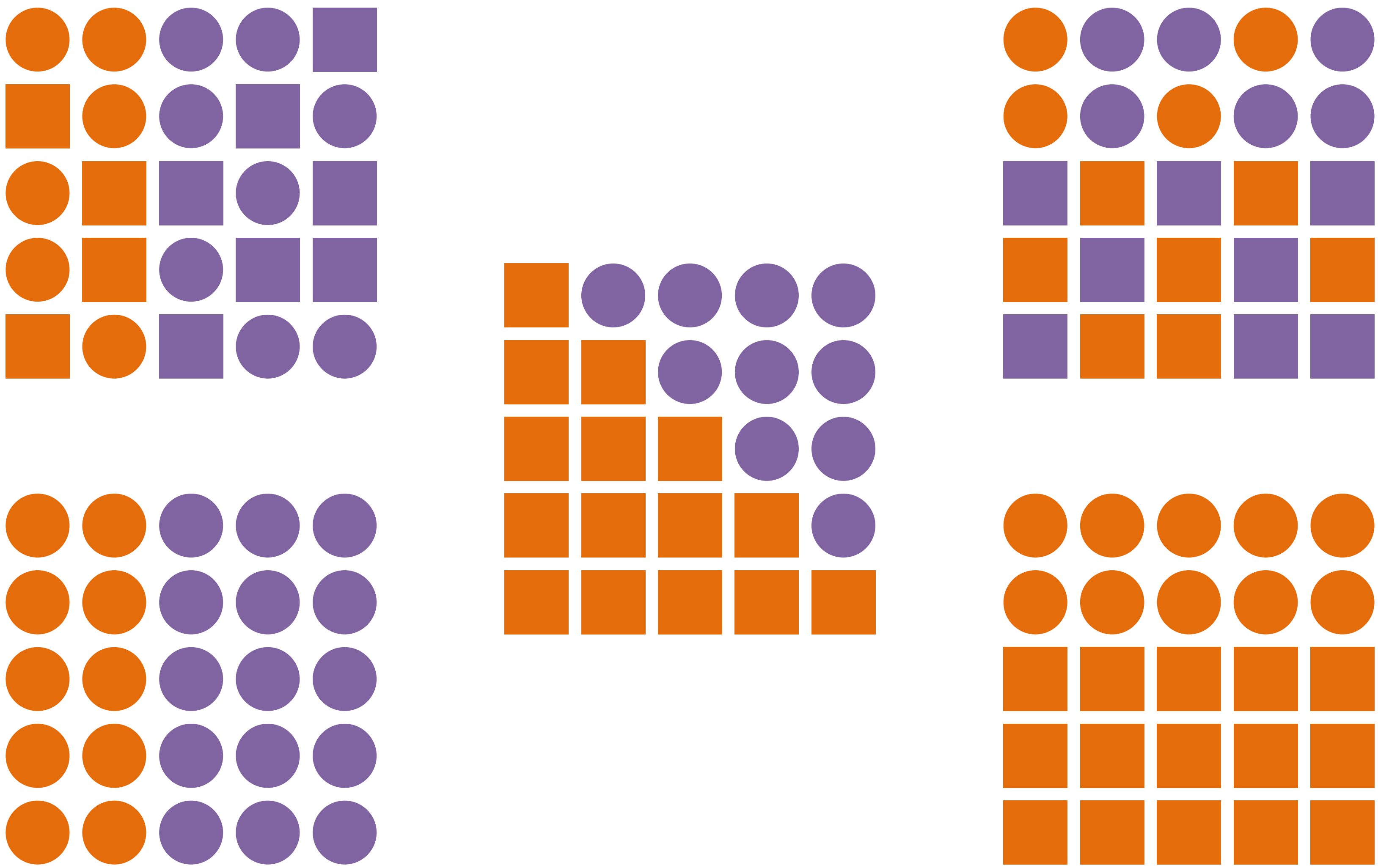
Rainbow Pies



Attribute Interference



Attribute Interference



Cultural conventions

- Reading proceed from left to right and from top to bottom
  - At least in western culture
- What is at the top (on the left) precedes what is at the bottom (on the right) in terms of
  - Importance
  - Ordering
  - Time

Emphasis

Attribute	Tables	Graphs
Line width	Boldface text	Thicker lines
Size	Bigger tables Larger fonts	Bigger graphs Wider bars Bigger symbols
Color intensity	Darker or brighter colors	
2-D position	Positioned at the top Positioned at the left Positioned in the center	

Colors

- Red
- Green
- Yellow
- Blue
- Black
- White
- Pink
- Cyan
- Gray
- Orange
- Brown
- Purple

References

- C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers, Inc., San Francisco, California, 2000
- C. Healey, and J. Enns. *Attention and Visual Memory in Visualization and Computer Graphics*. IEEE Transactions on Visualization and Computer Graphics, 18(7), 2012
- I. Inbar, N. Tractinsky and J.Meyer. Minimalism in information visualization: attitudes towards maximizing the data-ink ratio.
  - http://portal.acm.org/citation.cfm?id=1362587

References

- S.Few, “Practical Rules for Using Color in Charts”
  - http://www.perceptualedge.com/articles/visual_business_intelligence/rules_for_using_color.pdf
- D. Borland and R. M. Taylor Ii, "Rainbow Color Map (Still) Considered Harmful," in *IEEE Computer Graphics and Applications*, vol. 27, no. 2, pp. 14-17, March-April 2007.
  - http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4118486
- http://www.color-blindness.com
- http://www.csc.ncsu.edu/faculty/healey/PP/index.html