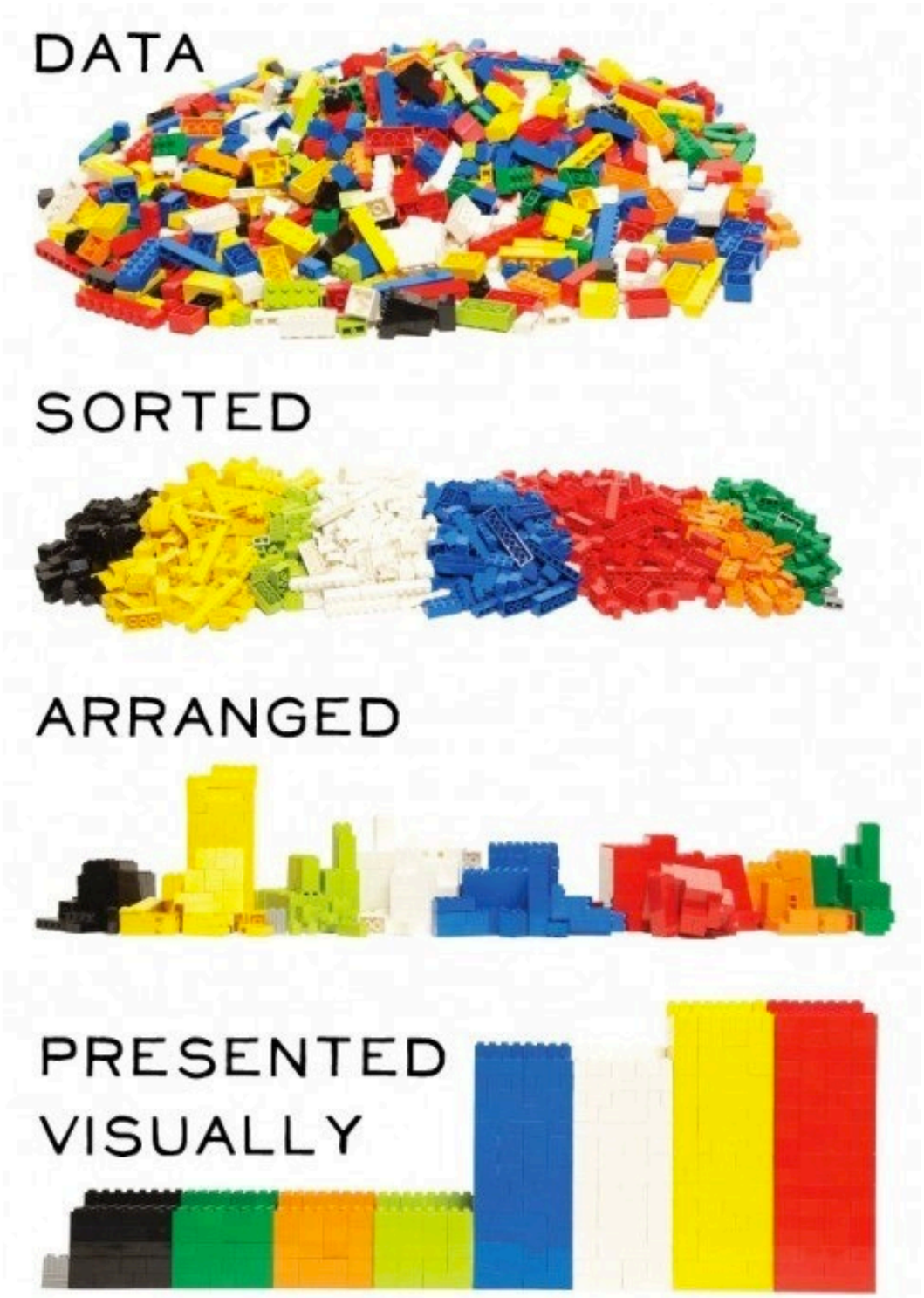
CSC2537 / STA2555 - INFORMATION VISUALIZATION

# DATA MODELS

Fanny CHEVALIER

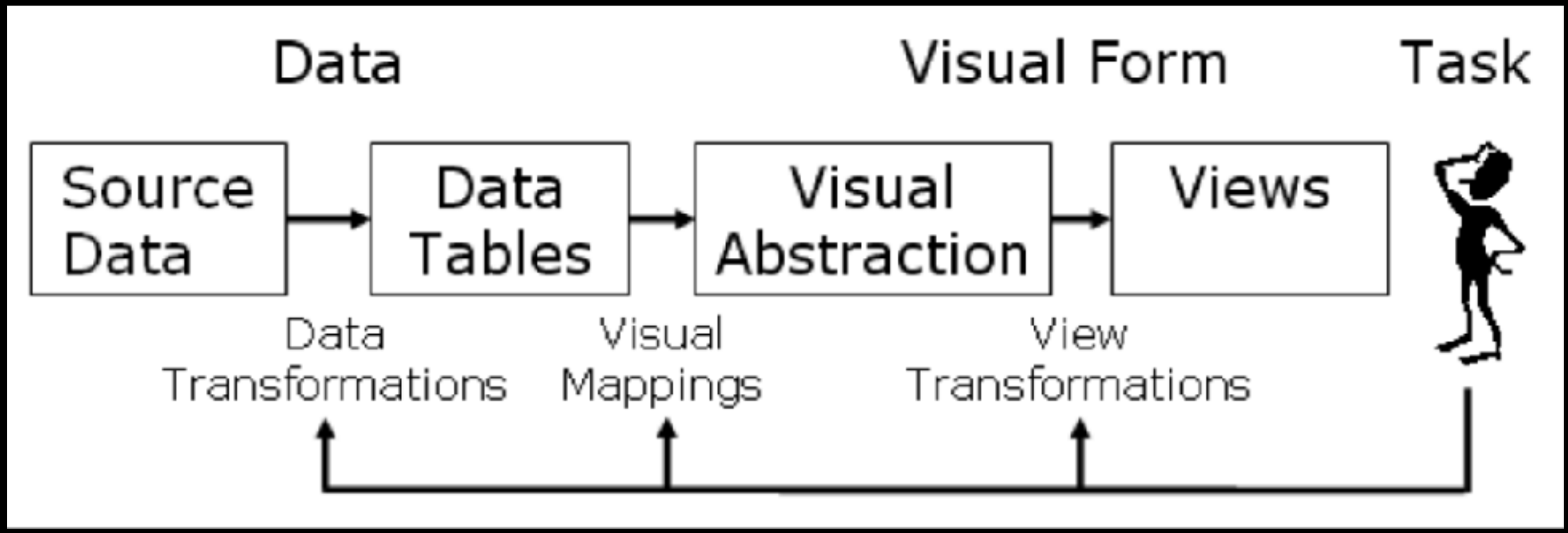
UNIVERSITY OF TORONTO




Source: http://www.hotbutterstudio.com/

# THE INFOVIS REFERENCE MODEL

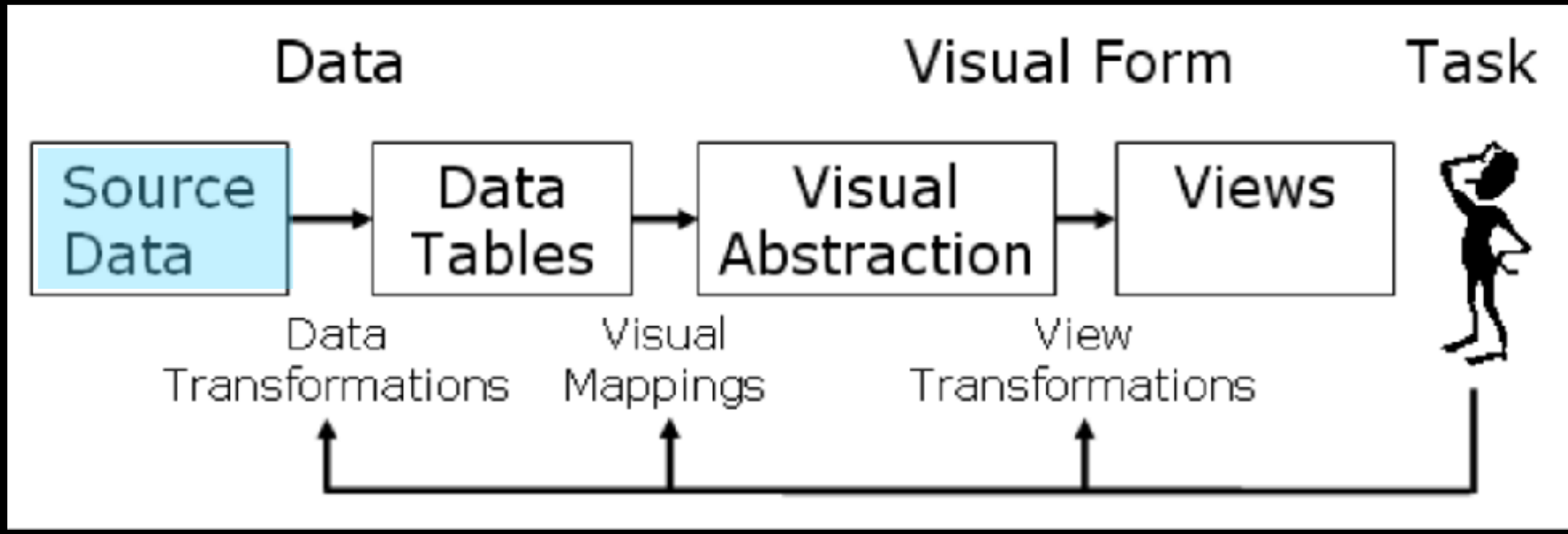
## aka infovis pipeline, data state model [Chi99]



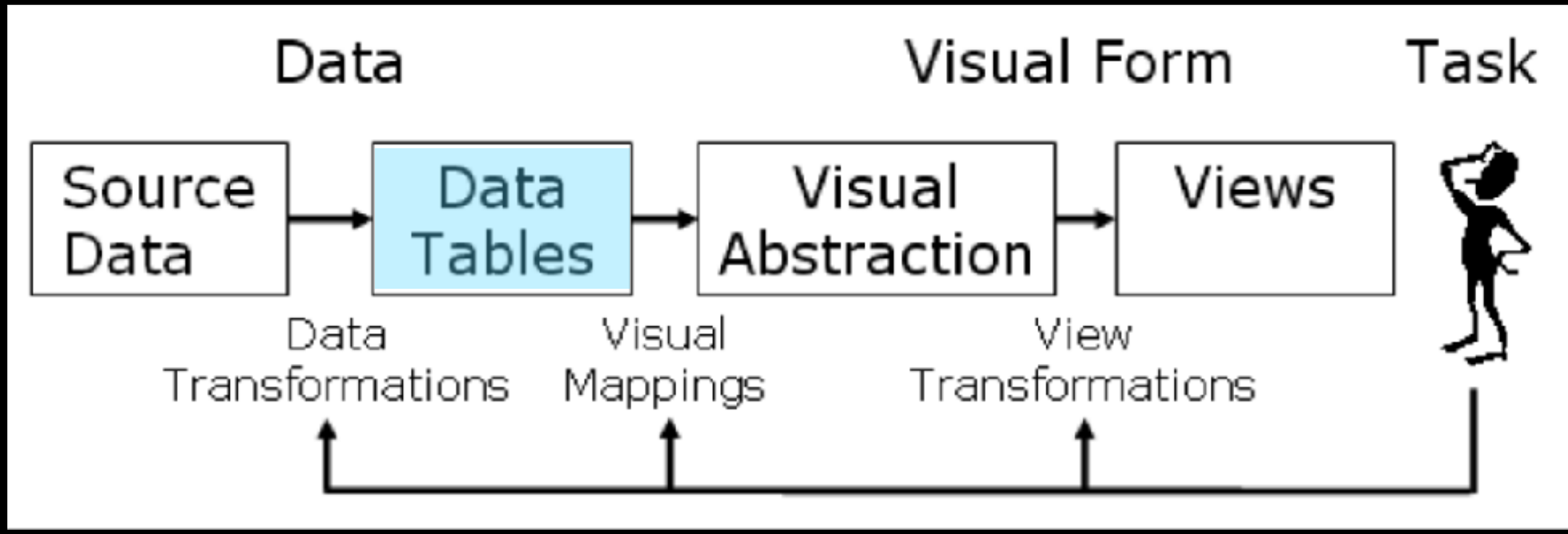

Ed Chi. A Framework for Information visualisation spreadsheets.
PhD Thesis, University of Minnesota, 1999.

Image from: Card, Mackinlay, and Shneiderman. Readings in Information Visualization: Using Vision To Think, Chapter 1. Morgan Kaufmann, 1999

# THE INFOVIS REFERENCE MODEL

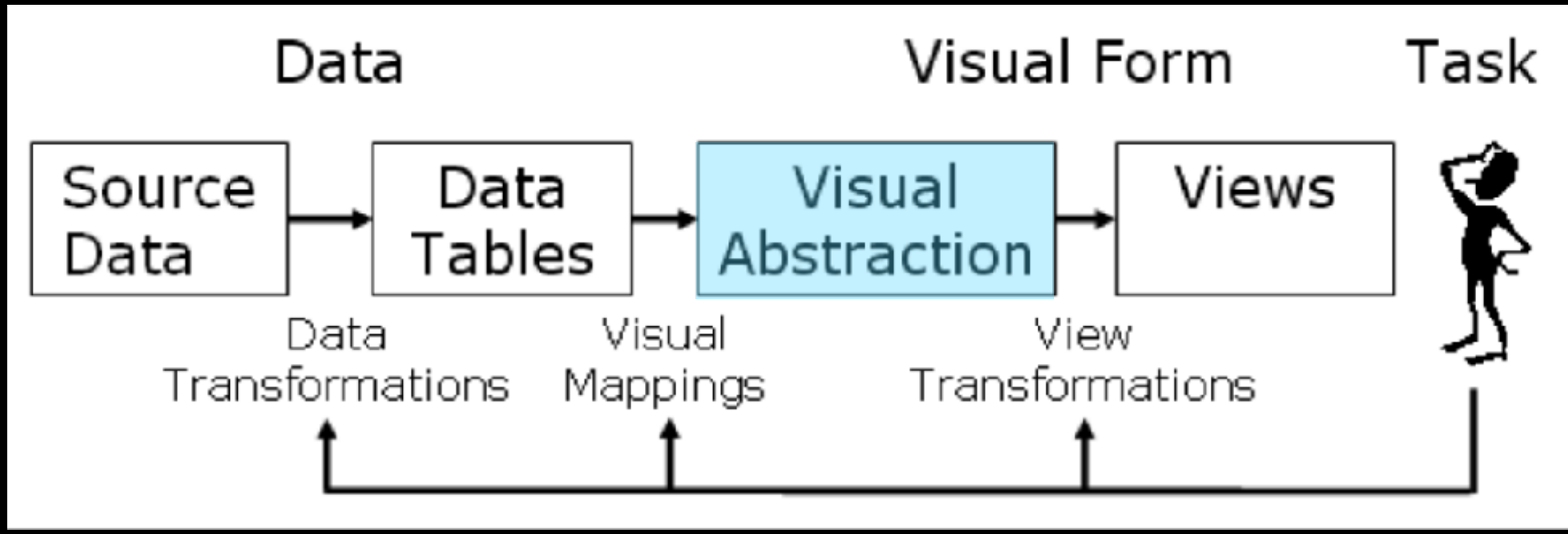



# THE INFOVIS REFERENCE MODEL

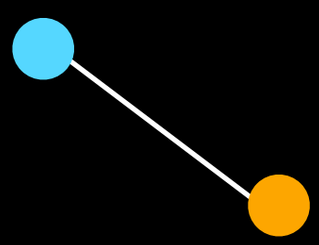



Fanny, Thomas
Fanny, Géry
Fanny, Nicolas
Fanny, Laurent
Fanny, Bruno
Fanny, Laëtitia
Thomas, Géry
Thomas, Nicolas
Laëtitia, Laurent
Laëtitia, Mathieu
Laëtitia, Julie
Bruno, Mathieu
…

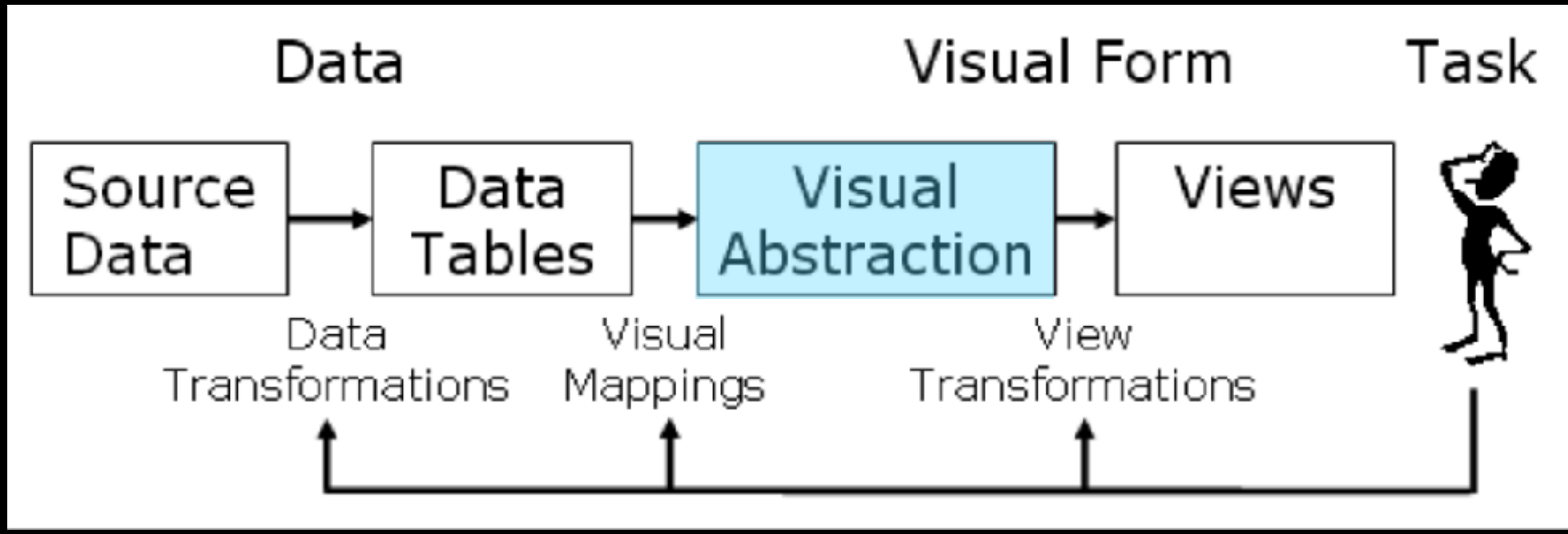
# THE INFOVIS REFERENCE MODEL



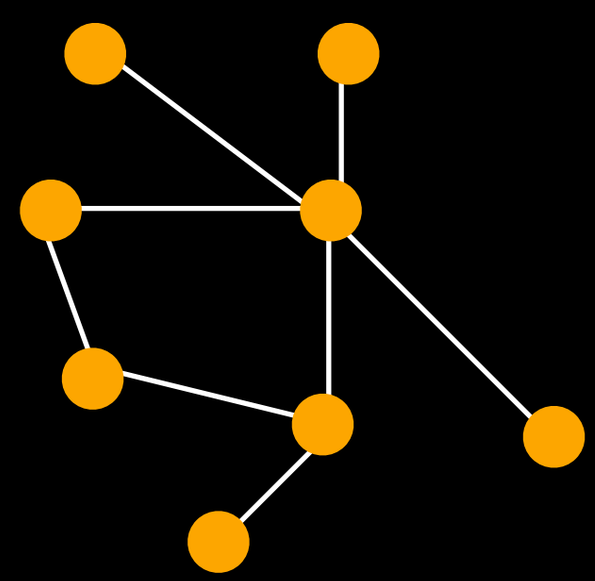

**Fanny**, **Thomas**
Fanny, Géry
Fanny, Nicolas
Fanny, Laurent
Fanny, Bruno
Fanny, Laëtitia
Thomas, Géry
Thomas, Nicolas
Laëtitia, Laurent
Laëtitia, Mathieu
Laëtitia, Julie
Bruno, Mathieu
…



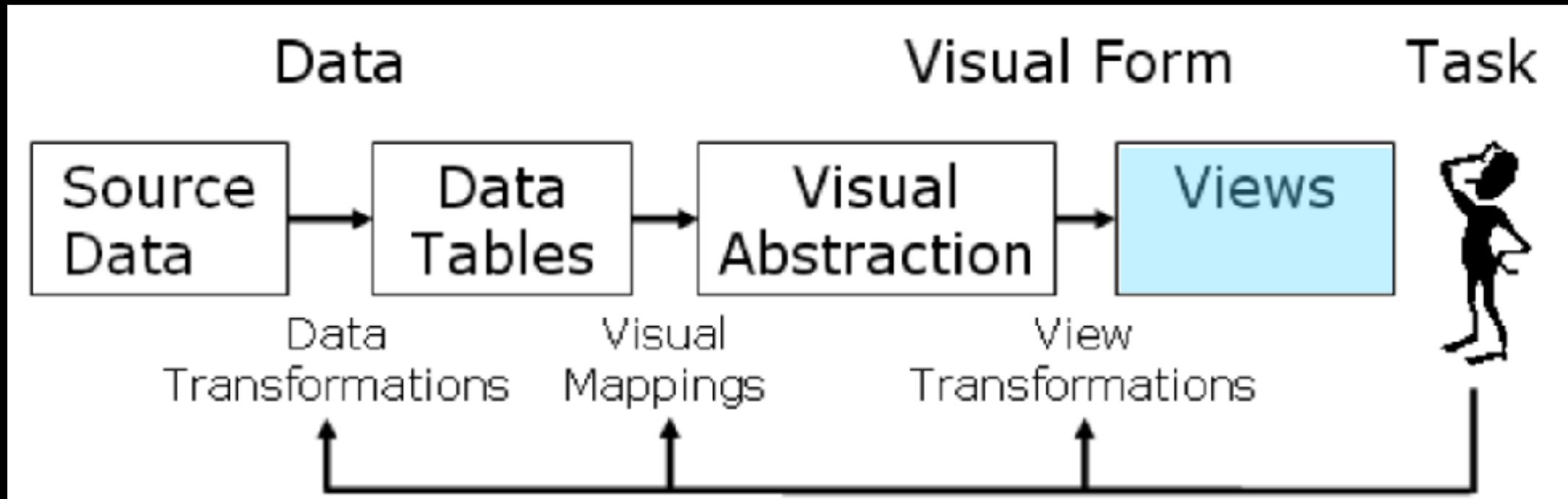
# THE INFOVIS REFERENCE MODEL



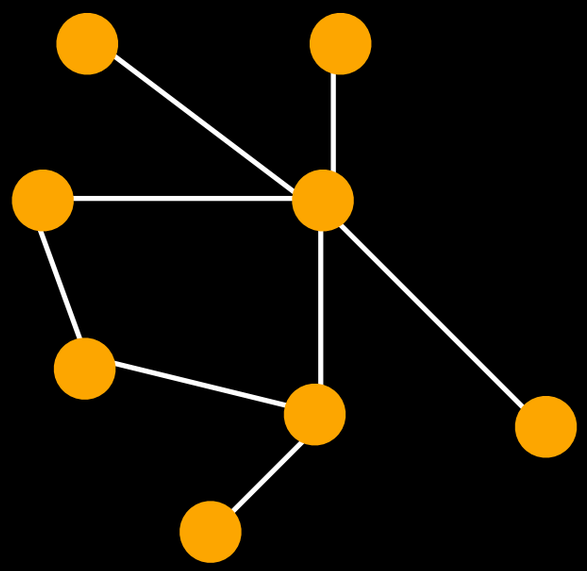

**Fanny, Thomas**
Fanny, Géry
Fanny, Nicolas
Fanny, Laurent
Fanny, Bruno
Fanny, Laëtitia
Thomas, Géry
Thomas, Nicolas
Laëtitia, Laurent
Laëtitia, Mathieu
Laëtitia, Julie
Bruno, Mathieu
…

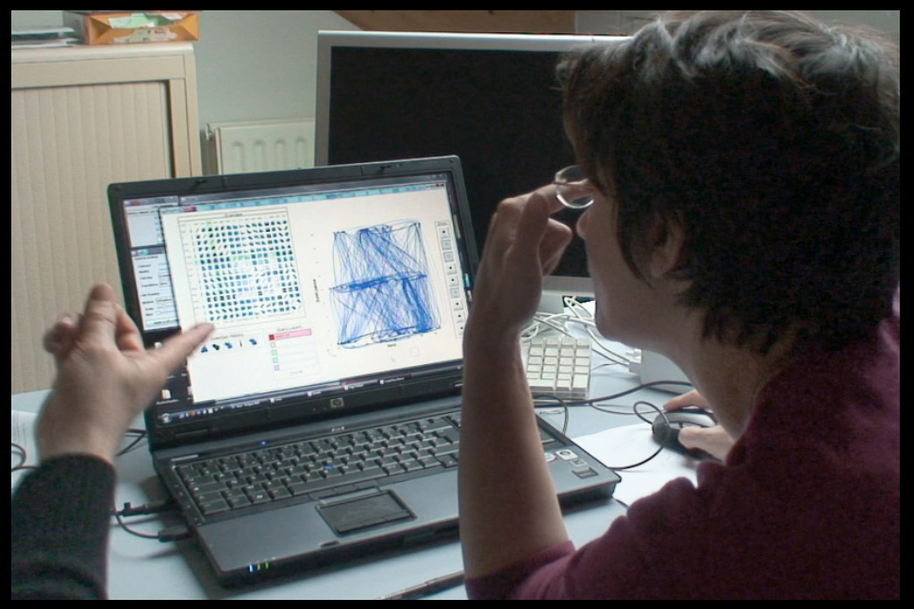


# THE INFOVIS REFERENCE MODEL

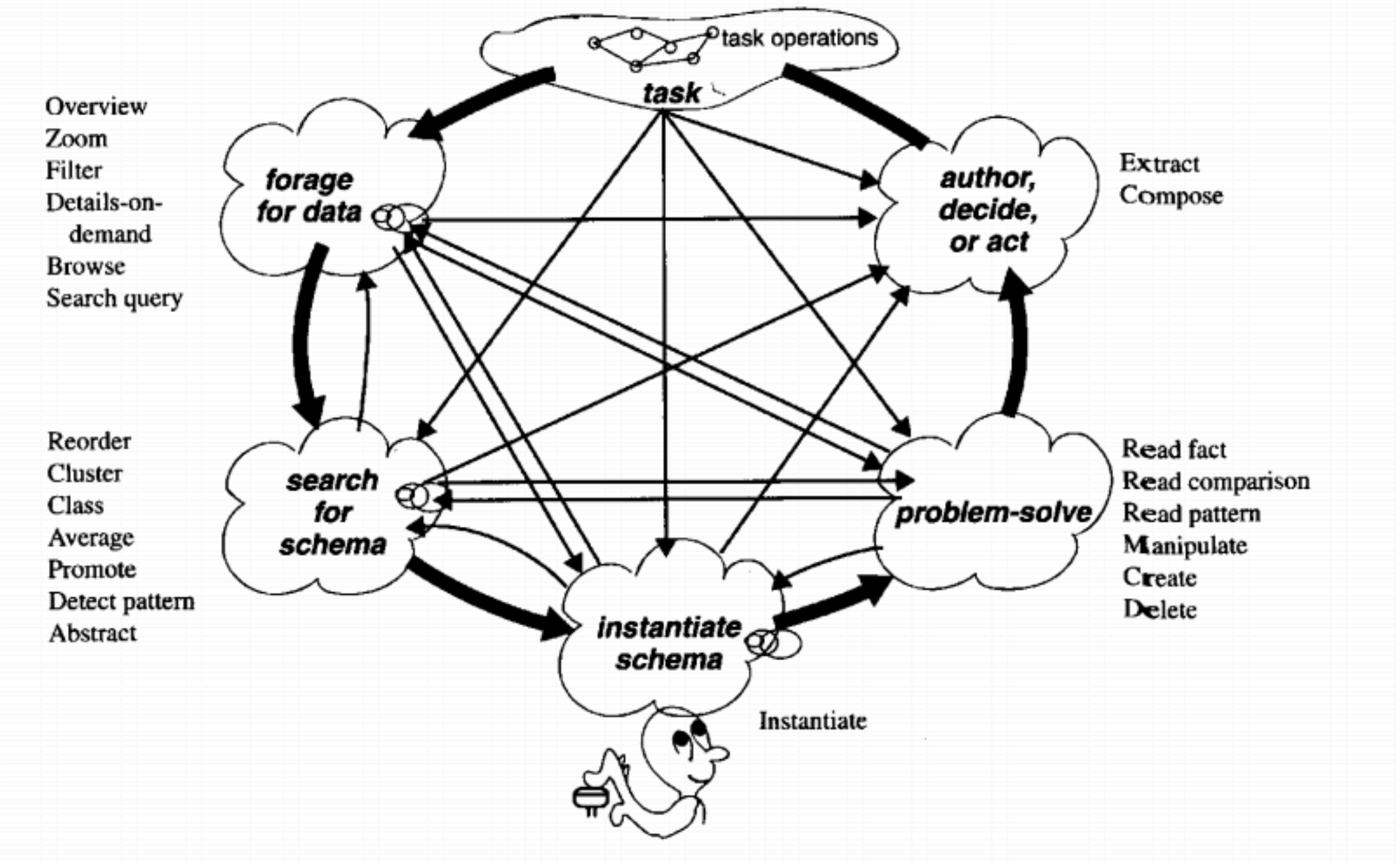



**Fanny, Thomas**
Fanny, Géry
Fanny, Nicolas
Fanny, Laurent
Fanny, Bruno
Fanny, Laëtitia
Thomas, Géry
Thomas, Nicolas
Laëtitia, Laurent
Laëtitia, Mathieu
Laëtitia, Julie
Bruno, Mathieu
…





# KNOWLEDGE CRISTALIZATION PROCESS



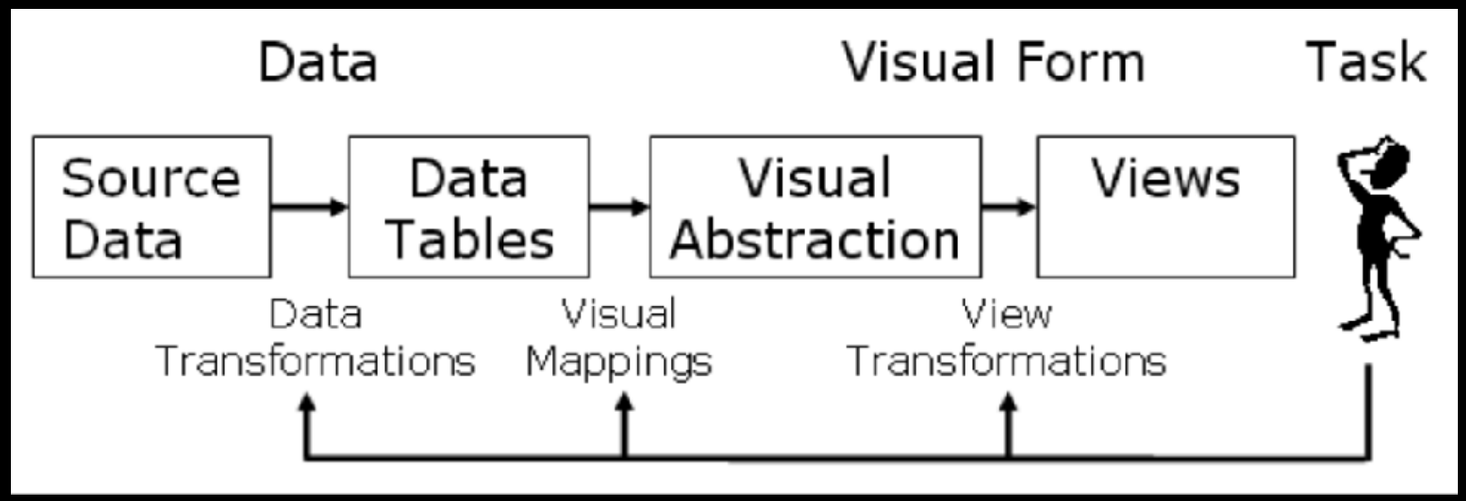

**WORKING WITH VISUALIZATIONS IS NOT A LINEAR PROCESS**

# THE VISUAL INFORMATION-SEEKING MANTRA

Overview first, Zoom and filter, then Details-on-demand.

Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations.
In Proc. Visual Languages, 336–343, 1996.

# CHALLENGES




- Collect the right data
- Choose the right data structure
- Not discard important data
- Choose the right representation
- Develop appropriate interaction mechanisms

# DATA TYPES

Taxonomies of **data types** stem from Steven's scale of measurement

- **Nominal** (identity)
- **Ordinal** (comparison)
- **Quantitative** (differences, ratios)

S.S. Stevens, On the theory of scales of measurements, 1946

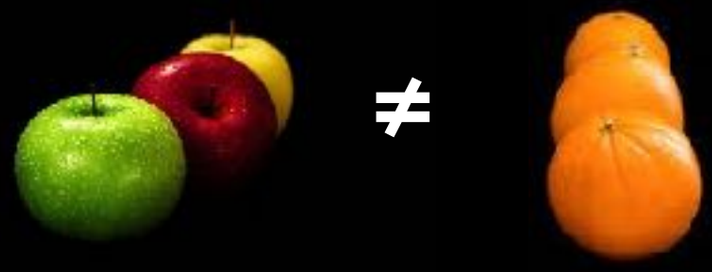
See also:
S. Card and J. Mackinlay. The Structure of the Information Visualization Design Space. In proc. InfoVis'97, 92–99, 1997.

# DATA TYPES

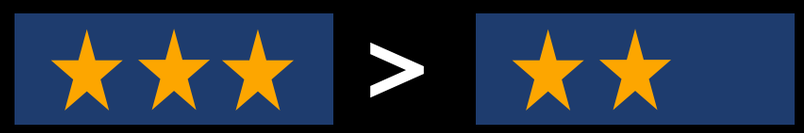
- **Nominal** (labels)
  - Fruits: apples, oranges
- **Ordinal**
  - Energy class: A, B, C, D, E
  - Meat quality: grade A, AA, AAA
  - Can be counted and compared, but not measured
- **Quantitative** : Interval
  - No absolute zero (or arbitrary)
  - E.g., dates, longitude, latitude
- **Quantitative** : Ratio
  - Meaningful origin
  - Physical measures (temperature, mass, length)
  - Accounts

# DATA TYPES

- **Nominal** (labels)
  - Operations: =, ≠




- **Ordinal**
  - Operations: =, ≠, <, >




- **Quantitative** : Interval
  - Operations: =, ≠, <, >, -, +
  - Distance measure possible

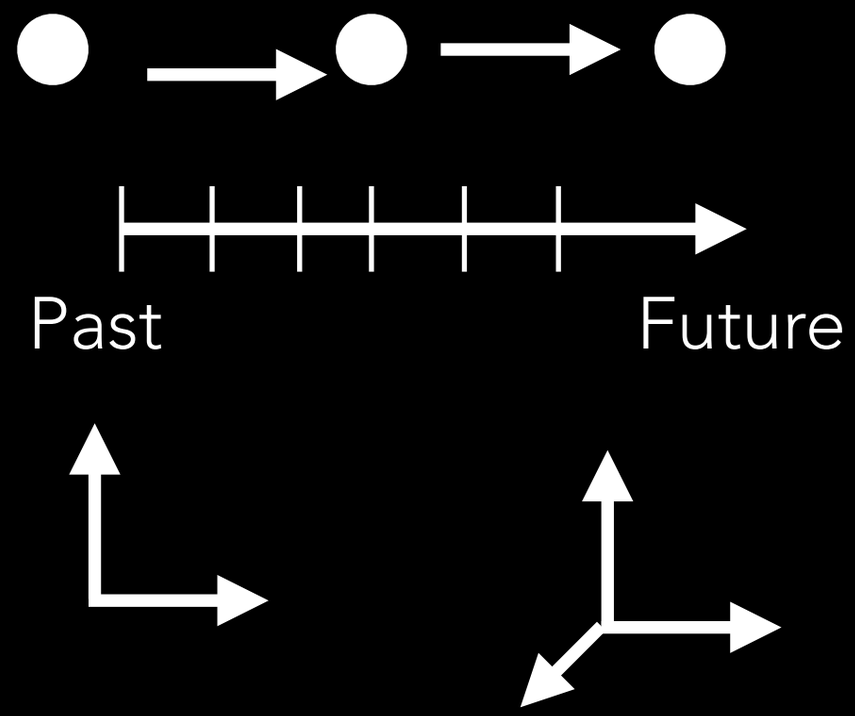
[1989 - 1999] **+** [2002 - 2012]

- **Quantitative** : Ratio
  - Operations : =, ≠, <, >, -, +, x, /
  - Ratio or proportion measure possible

10kg **/** 5kg

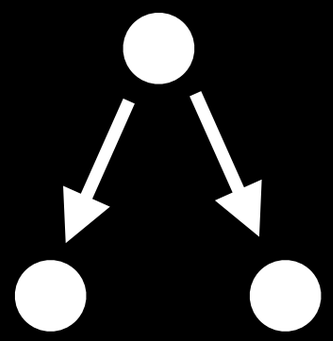
# DATA TYPES

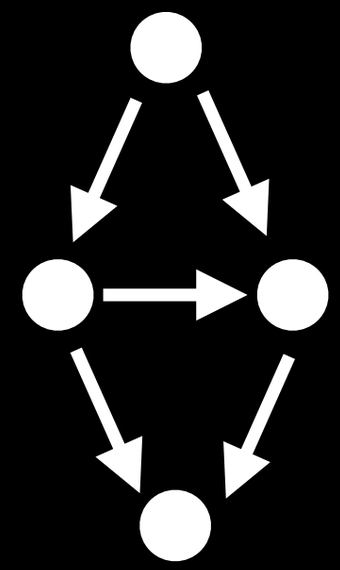
- **1D** (linear)
- **Temporal**
- **2D** (map)
- **3D**
- **nD** (relational)
- **Tree** (hierarchical)
- **Network** (graphs)









# Why is it important?

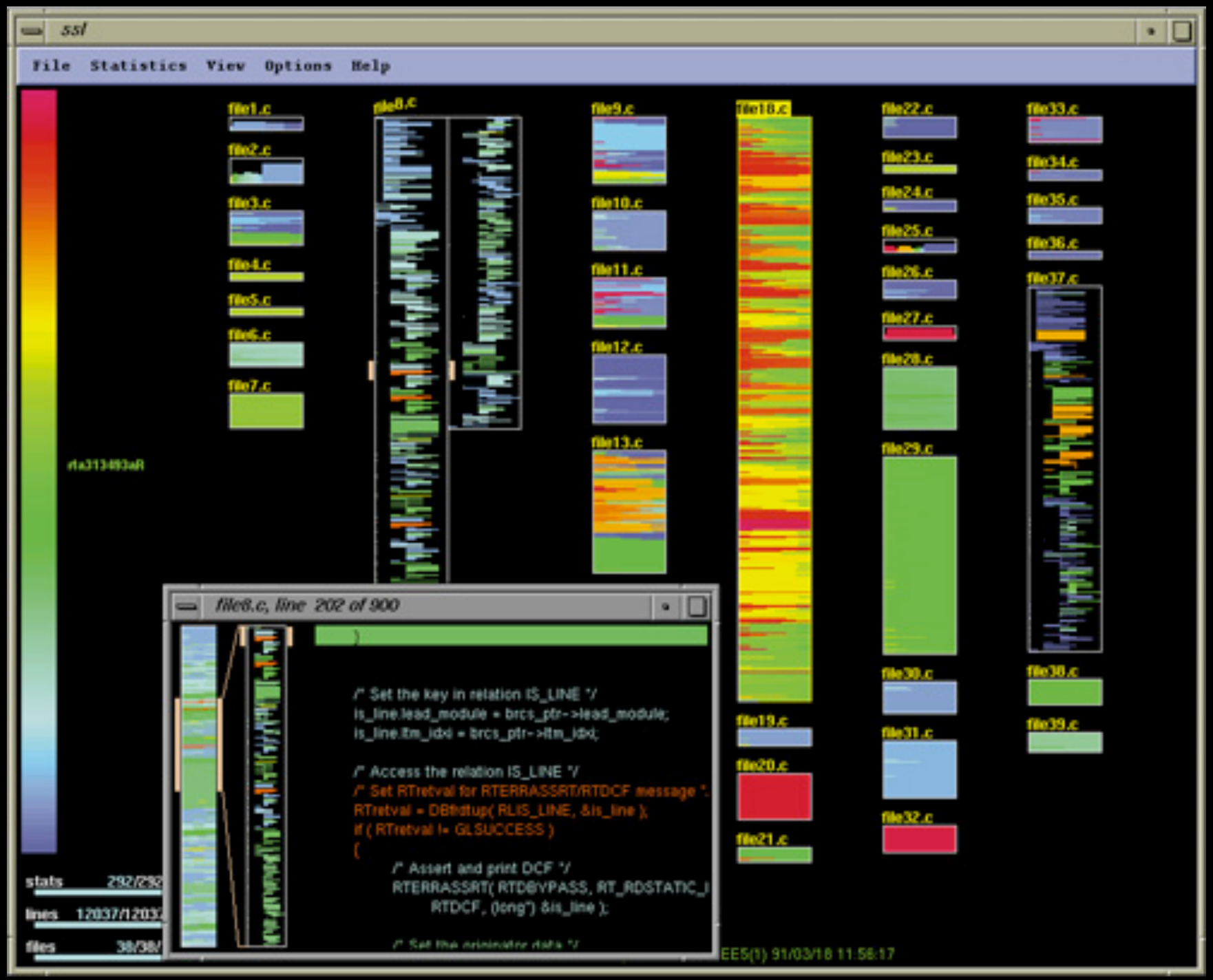
- The most appropriate visual representation for different data types (ordinal, nominal, quantitative) are different

- Different data types are often tied to specific tasks
    - temporal data: compare events
    - hierarchical data: understand parent-child relationships

**But :**

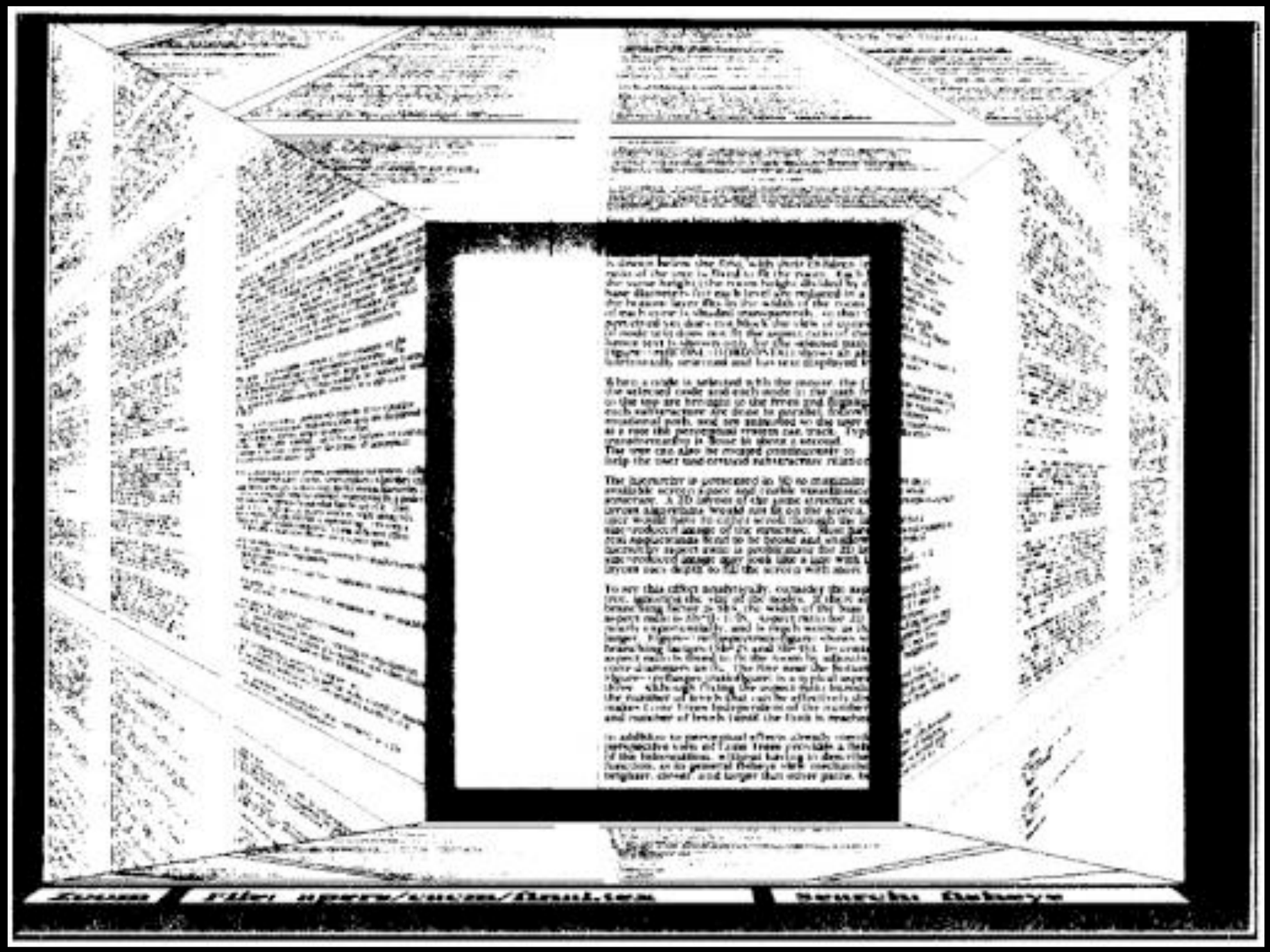
**Each data type (1D, 2D, …) can be represented in multiple ways**

# LINEAR DATA




Seesoft [Eick, IEEE TVCG 1992]

# LINEAR DATA



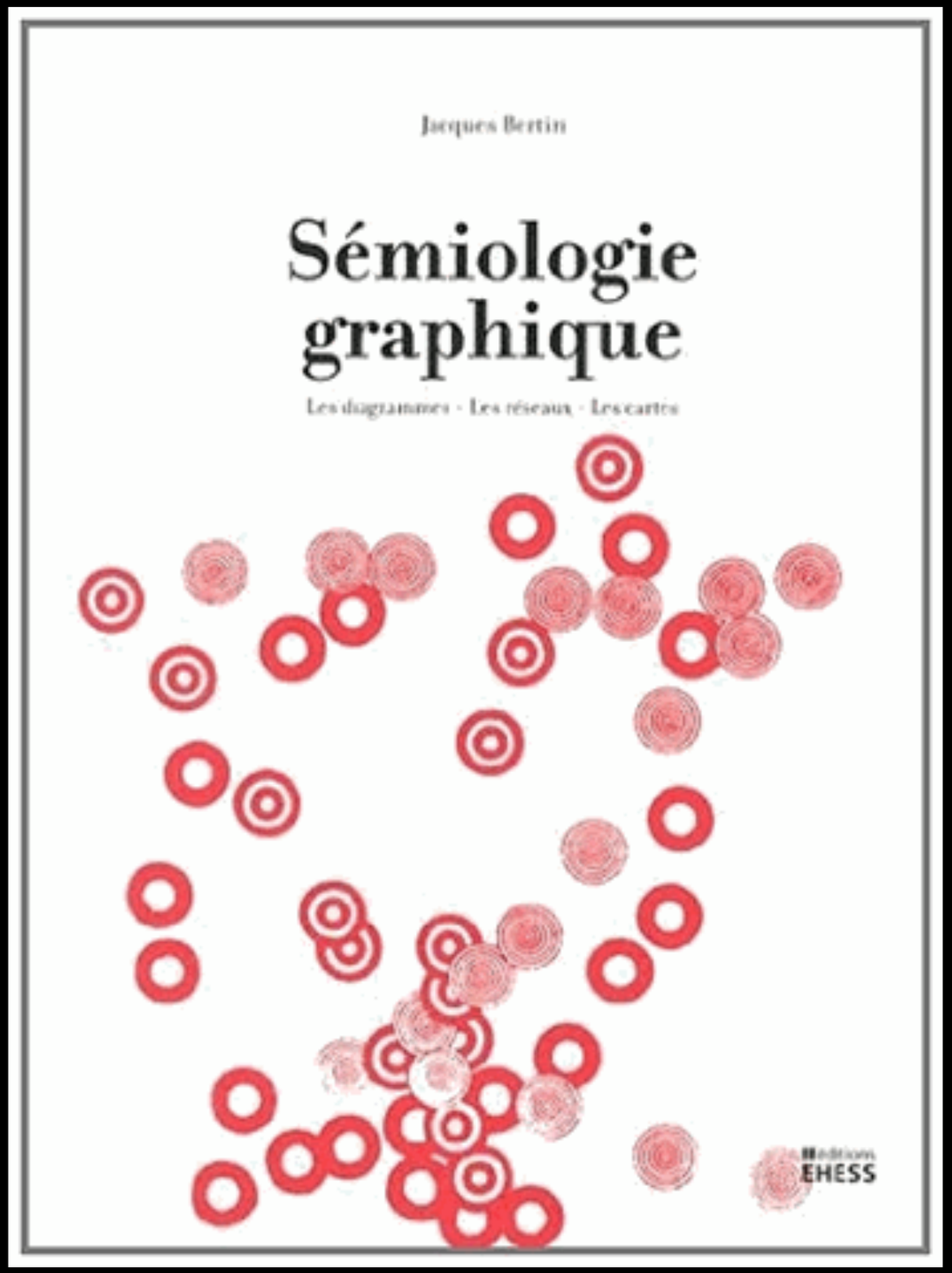
The Document Lens [Robertson & Mackinlay, UIST'93]

# LINEAR DATA


See also: http://www.bewitched.com/song.html

Arc Diagrams [Wattenberg, Infovis 2002]

# GRAPHIC SEMIOLOGY

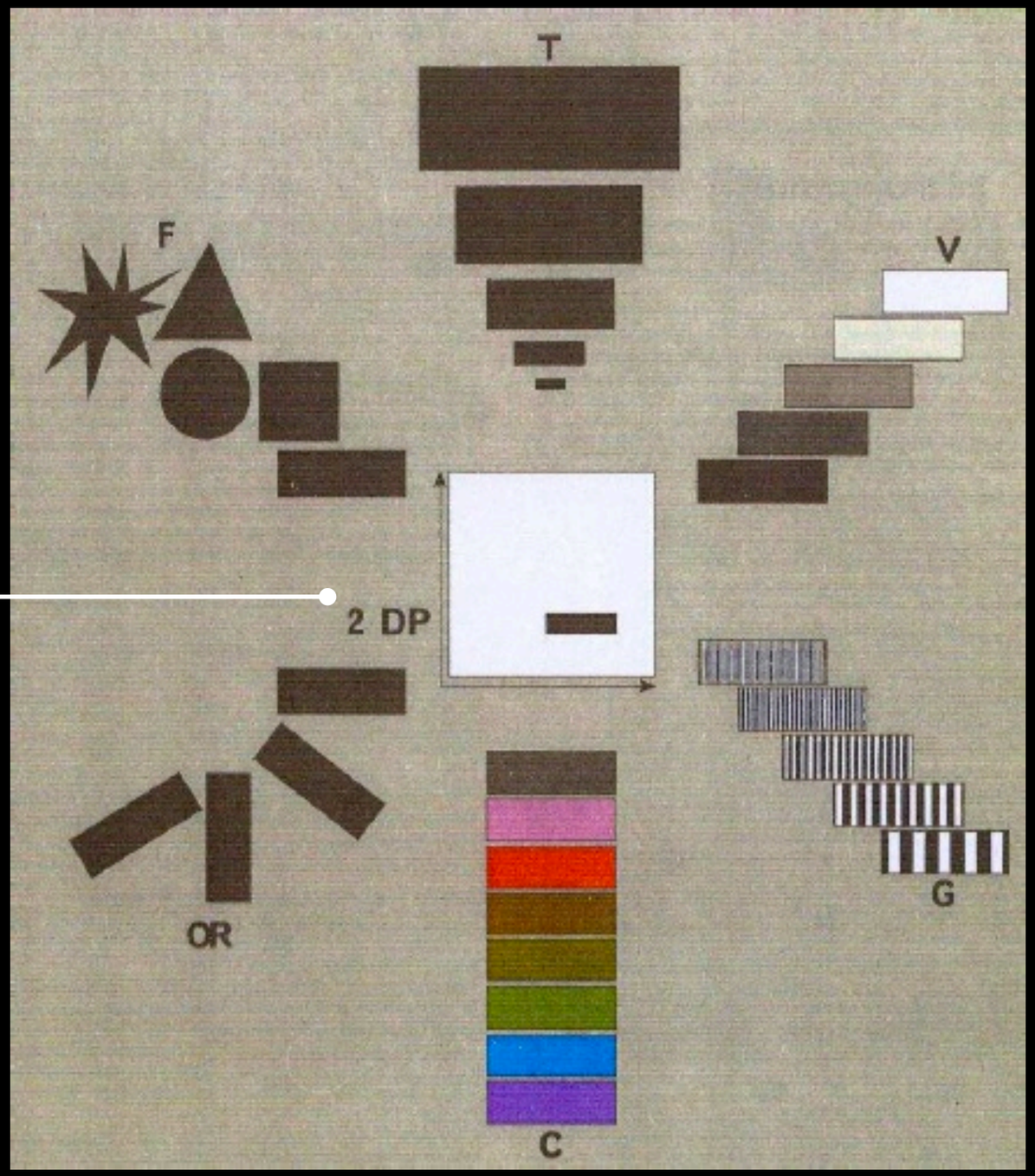



Jacques Bertin (1918-2010)

# VISUAL VARIABLES (aka Retinal variables)

Size

Shape




Value

2D position

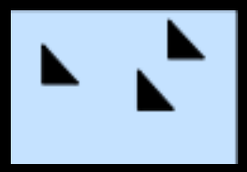
Orientation

Texture

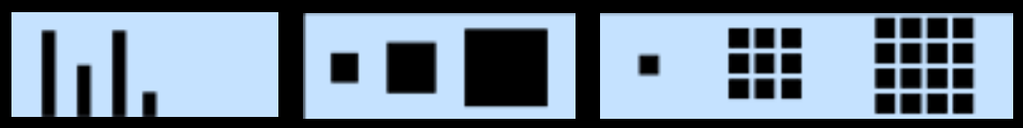
Color (Hue)

# VISUAL VARIABLES: ATTRIBUTES

- **position**
  changes in the x, y (z) location

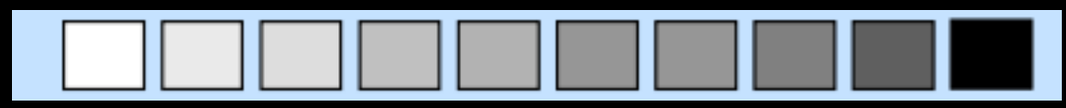


- **size**
  changes in length, area or repetition

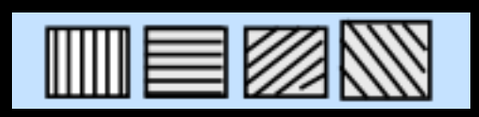


- **shape**
  infinite number of shapes

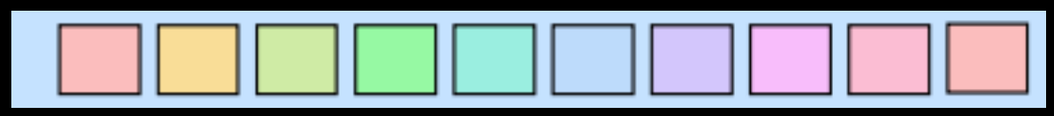
- **value**
  changes from light to dark



- **orientation**
  changes in alignment



- **colour**
  changes in hue at a given value



- **texture**
  variation in pattern

- **(motion)**

# VISUAL VARIABLES : CHARACTERISTICS

- **selective**
  is a change in this variable enough to allow us to select it from a group?

- **associative**
  is a change in this variable enough to allow us to perceive them as a group?

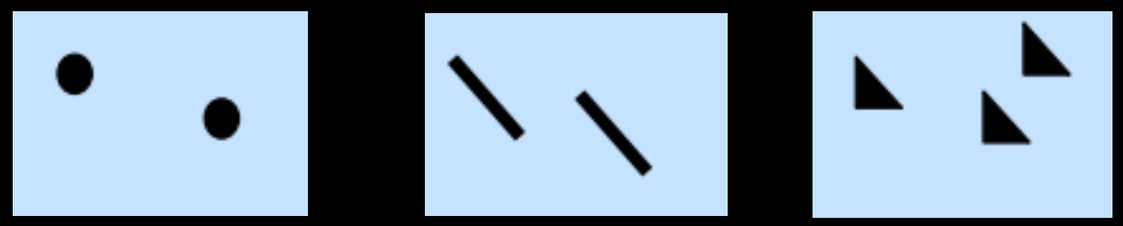
- **quantitative**
  is there a numerical reading obtainable from changes in this variable?

- **order**
  are changes in this variable perceived as ordered?

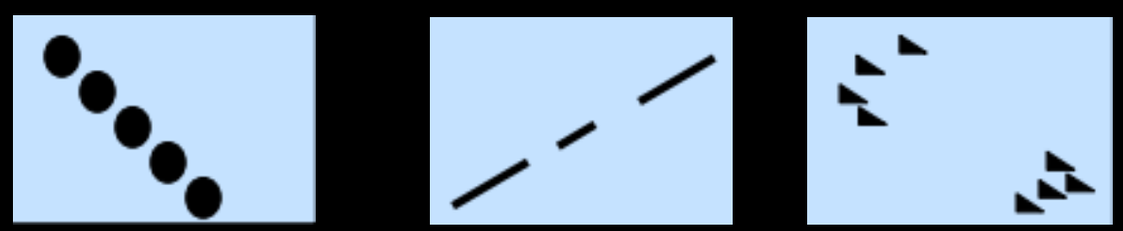
- **length**
  across how many changes in this variable are distinctions perceptible?

# POSITION

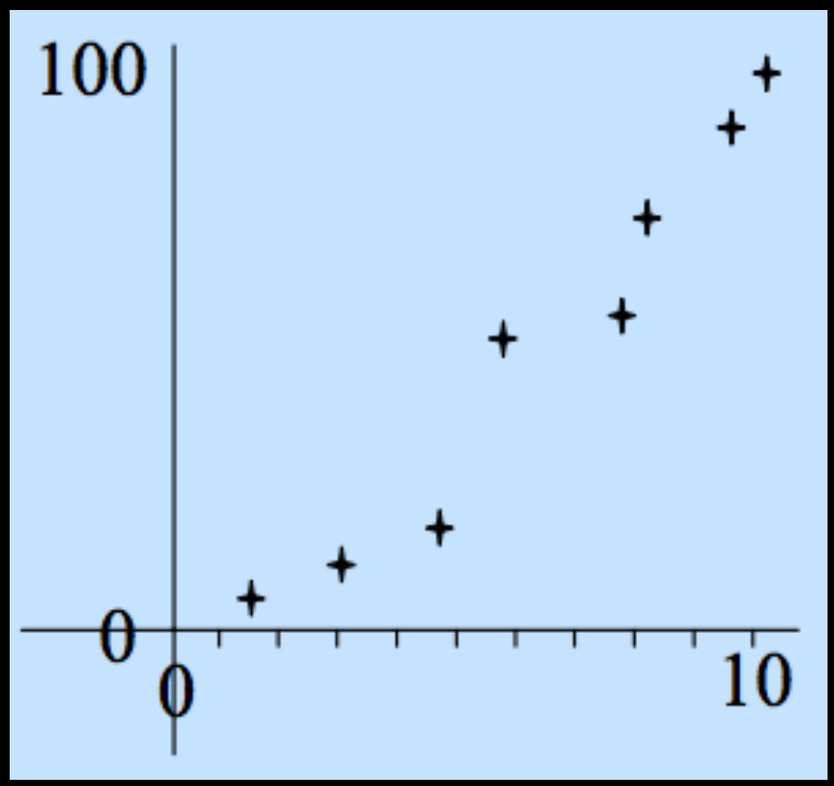
- ✓ **selective**



- ✓ **associative**

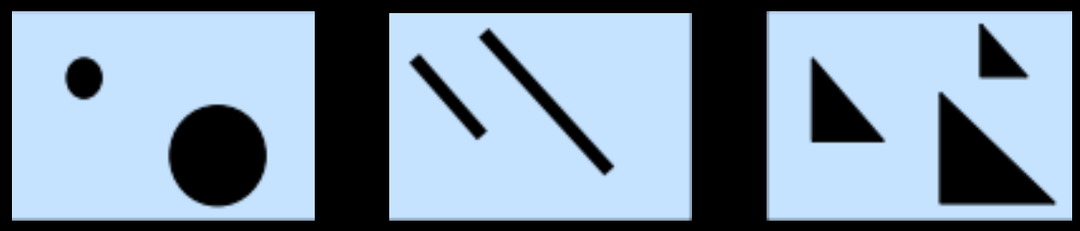


- ✓ **quantitative**
- ✓ **order**
- ✓ **length**

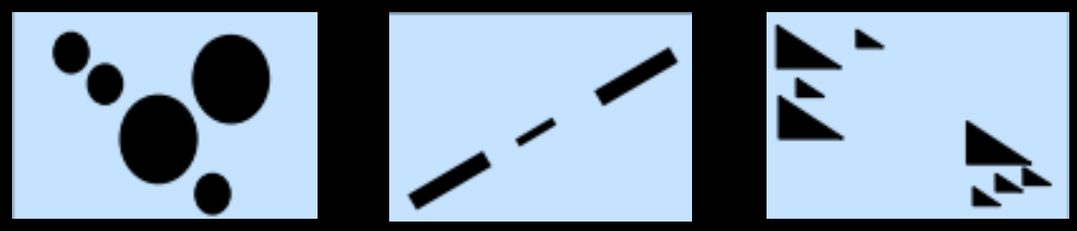


# SIZE

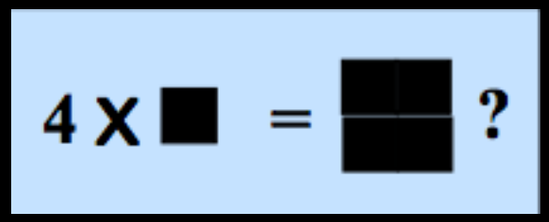
✓ **selective**



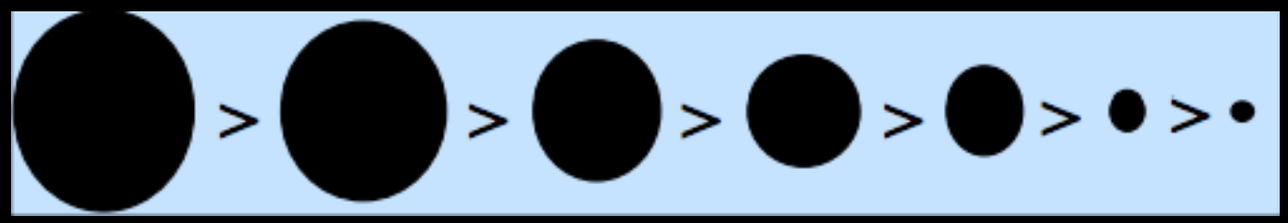
✓ **associative**



≈ **quantitative**




✓ **order**

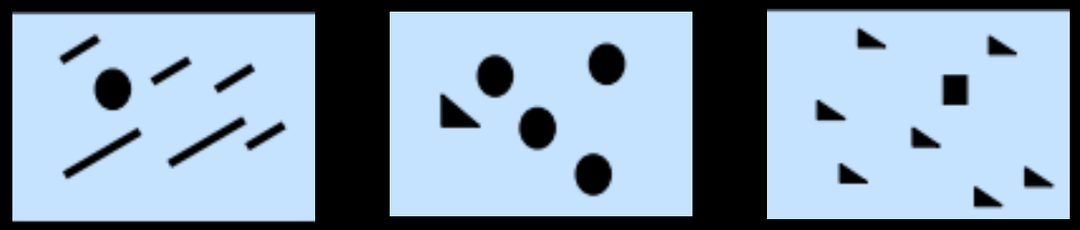


**length**

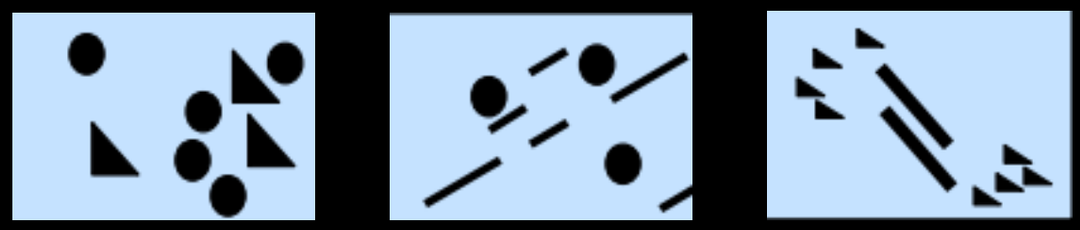
✓
- theoretically infinite but practically limited
- association and selection ~5 and distinction ~ 20

# SHAPE

≈ **selective**

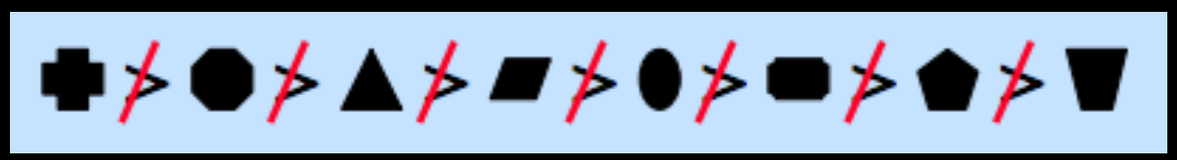


≈ **associative**



✗ **quantitative**

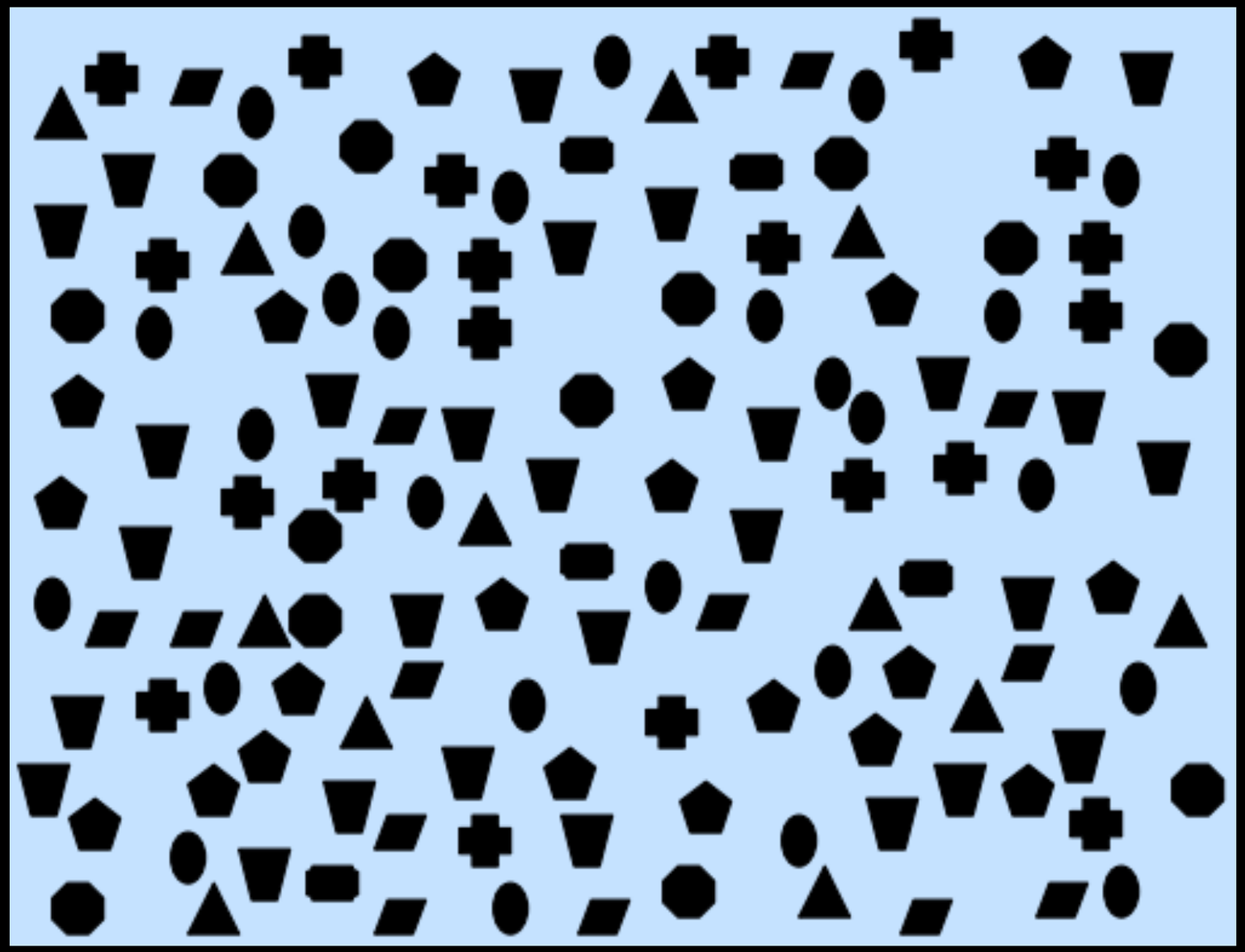
✗ **order**



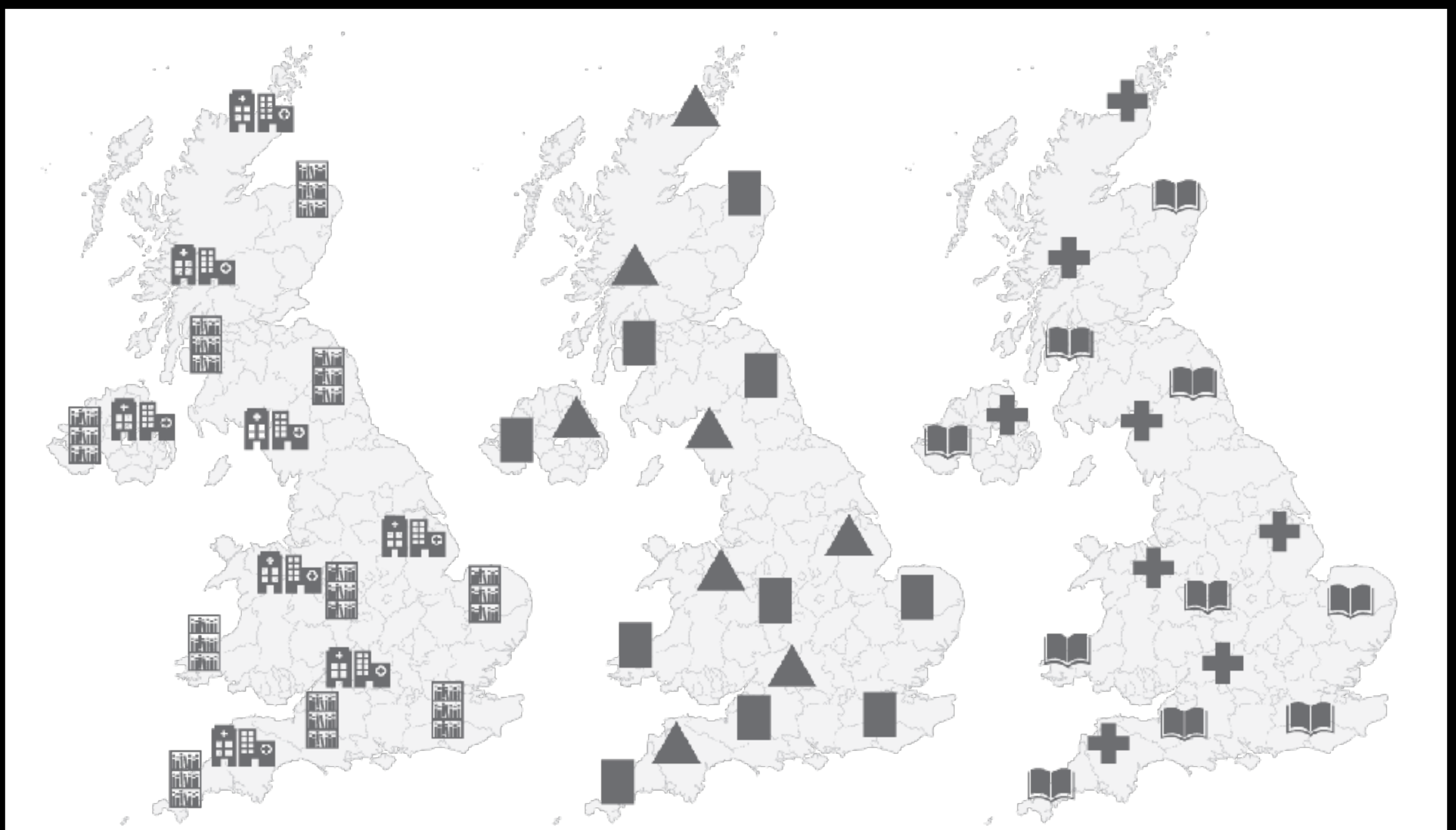
✓ **length**

- infinite variations

# SHAPE

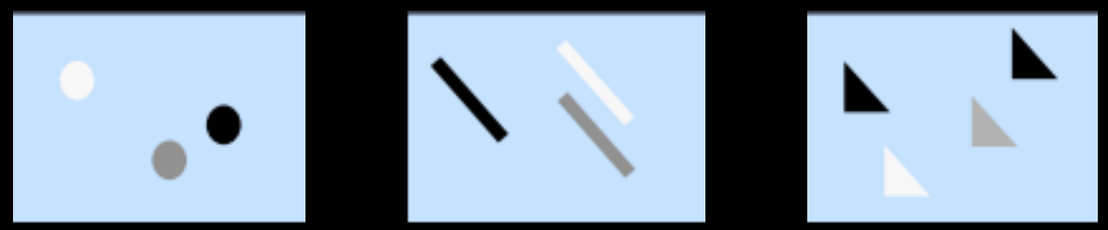


# SHAPE

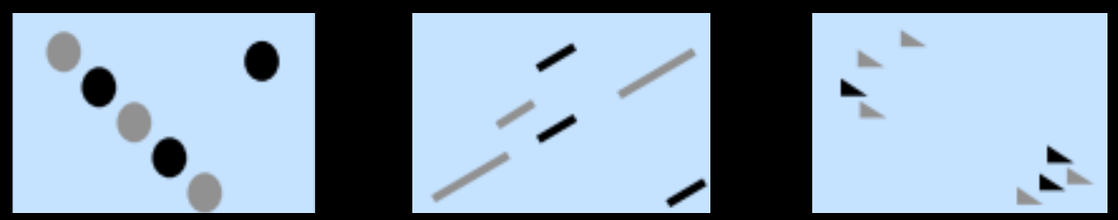


# VALUE

✓ **selective**

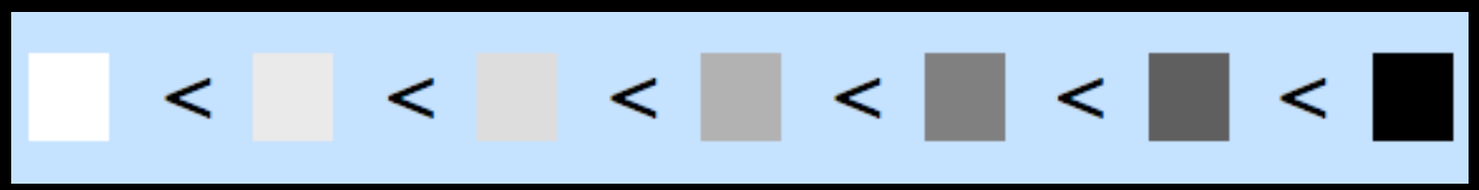


✓ **associative**



✗ **quantitative**

✓ **order**

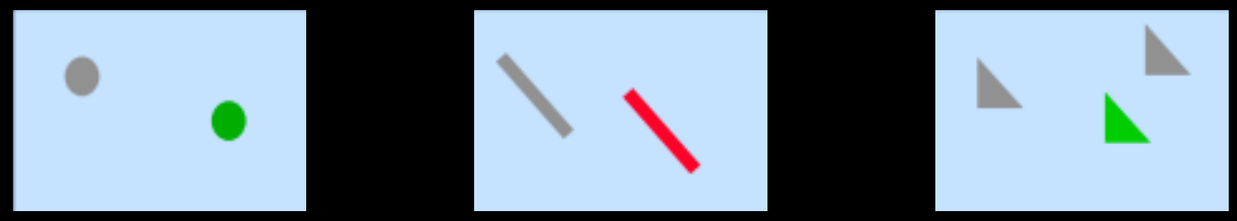



✓ **length**

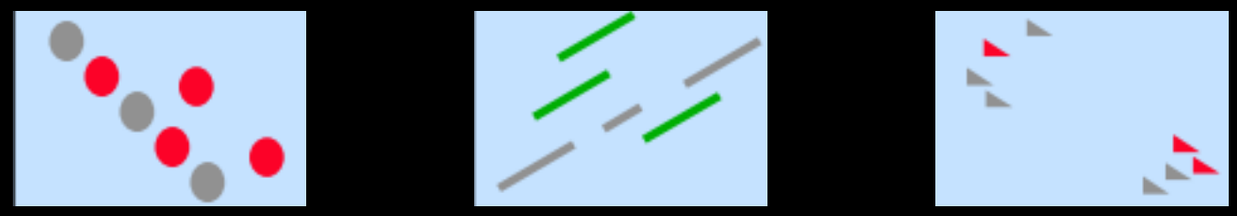
- theoretically infinite but practically limited
- association and selection < ~7 and distinction ~10

# COLOUR

✓ **selective**



✓ **associative**



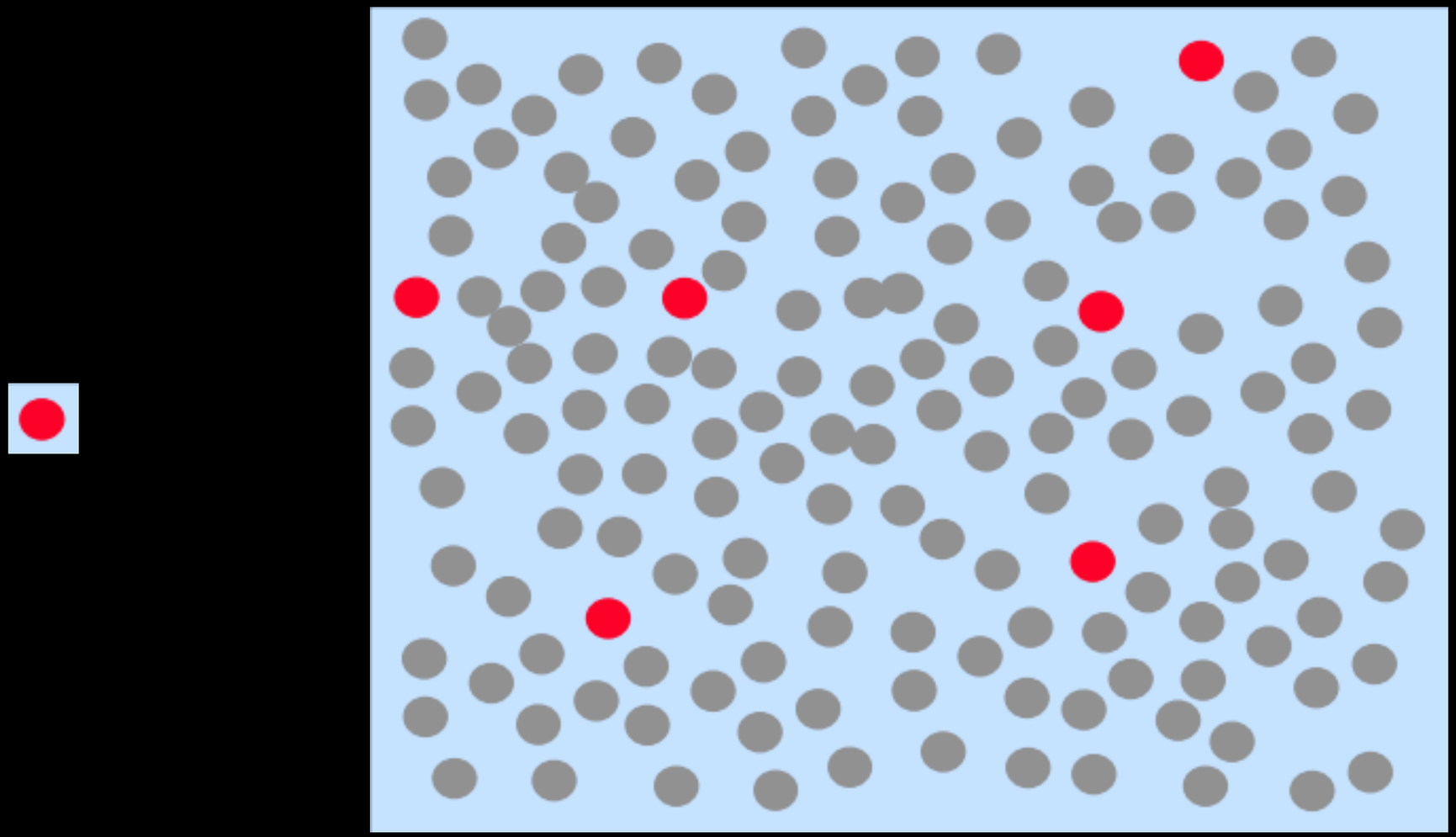
✗ **quantitative**

✗ **order**

✓ **length**

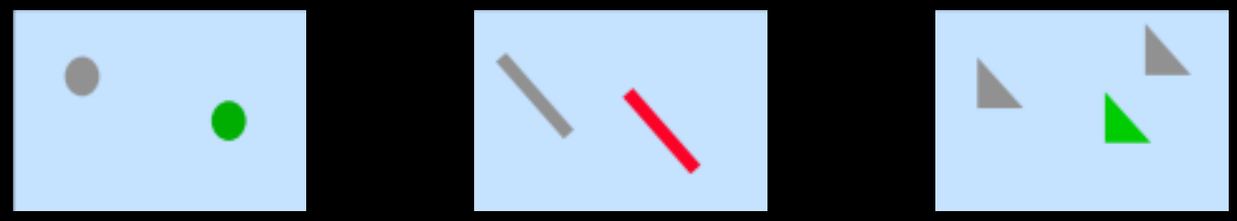
- theoretically infinite but practically limited
- association and selection < ~7 and distinction ~10

# COLOUR

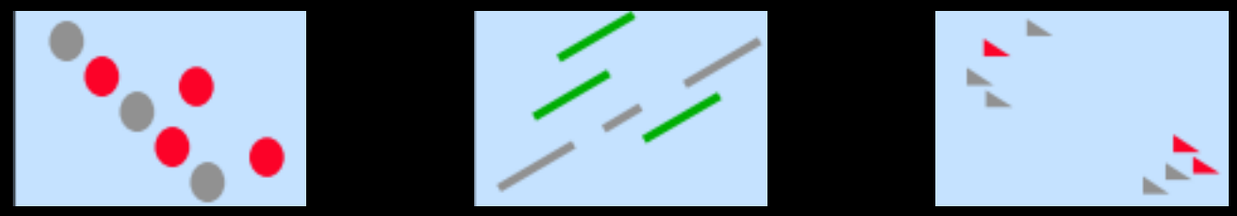


# COLOUR

✓ **selective**



✓ **associative**



✗ **quantitative**

✗ **order**

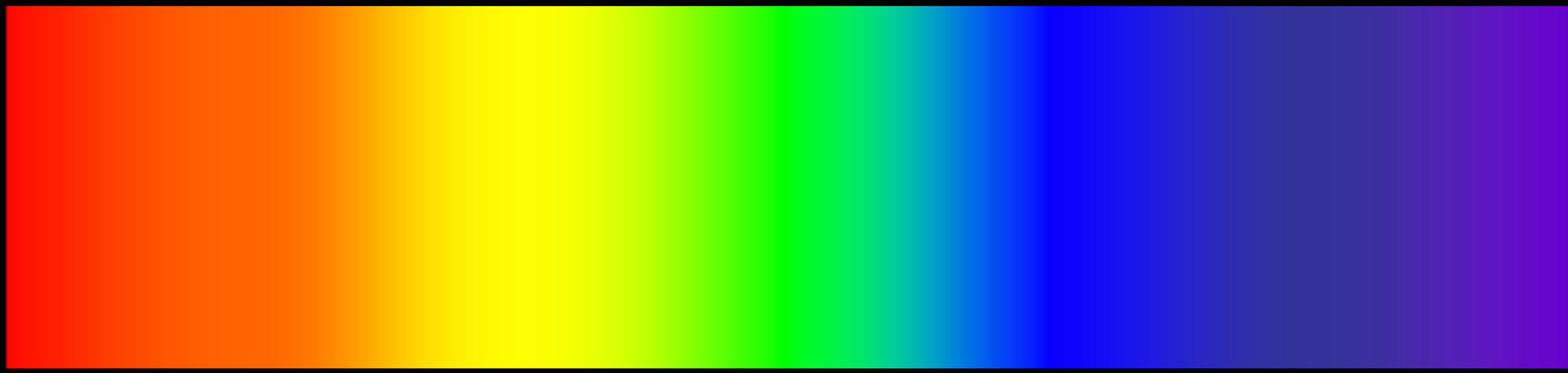
✓ **length**

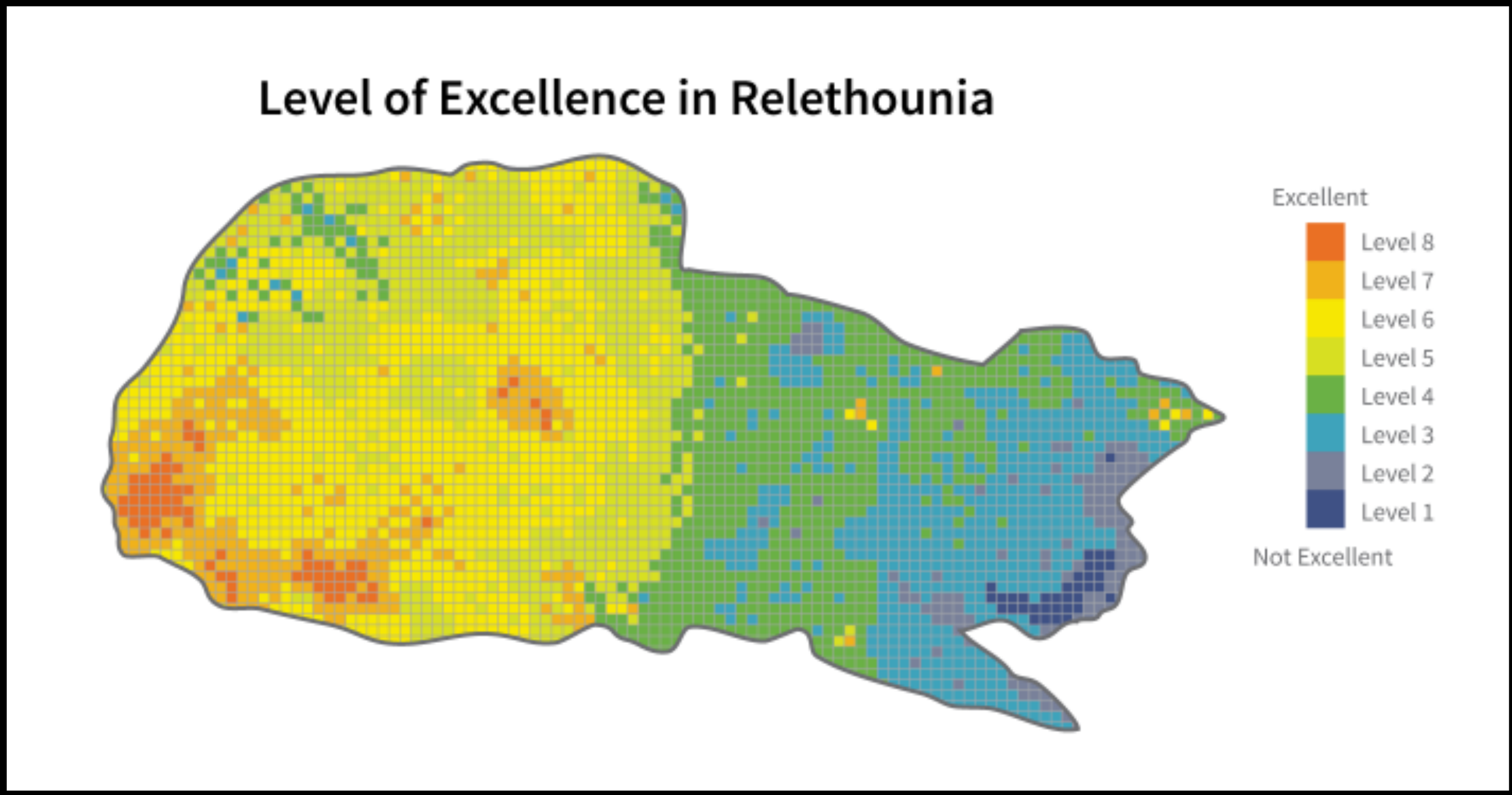
- theoretically infinite but practically limited
- association and selection < ~7 and distinction ~10

# ENCODING

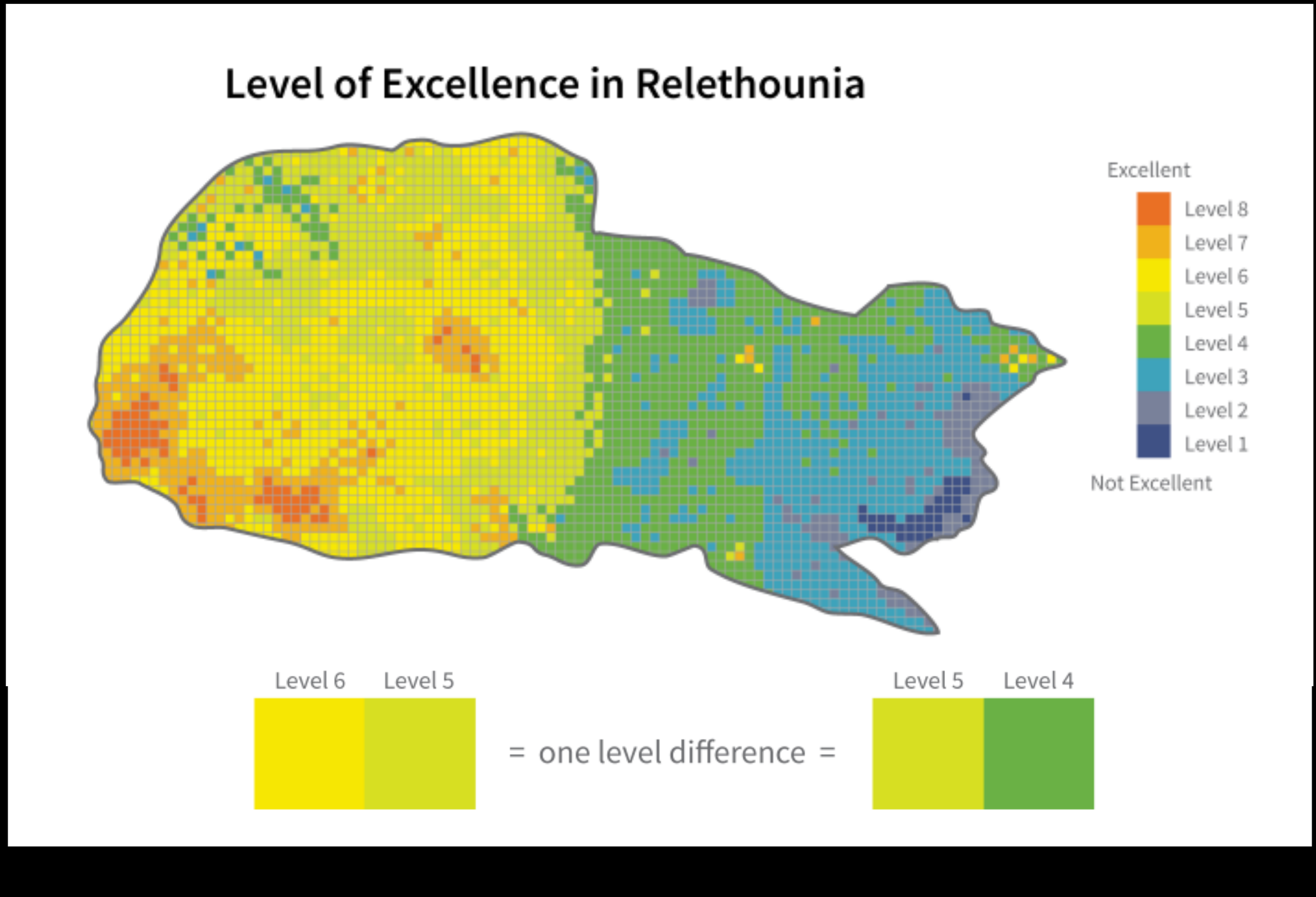
Common advice says use a rainbow scale

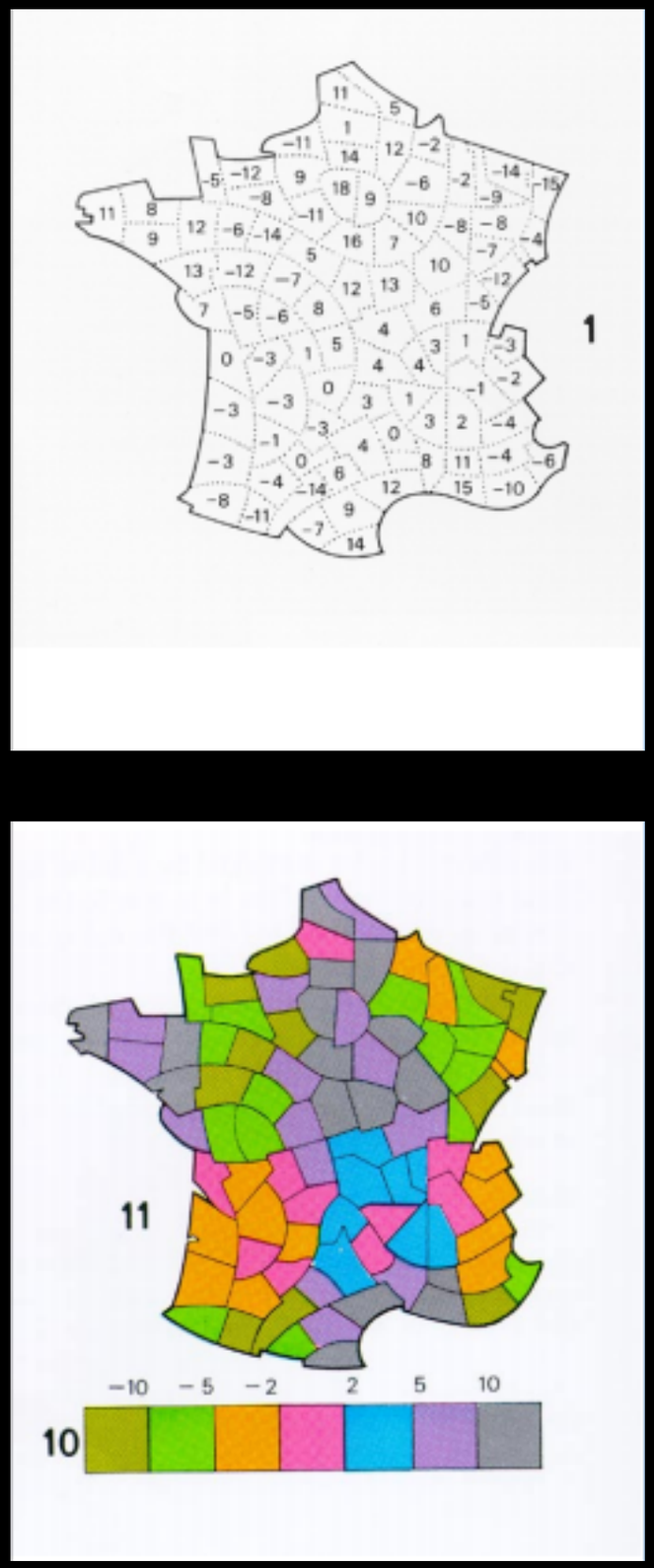
- Marcus, Murch, Healey
- strong problems with rainbows

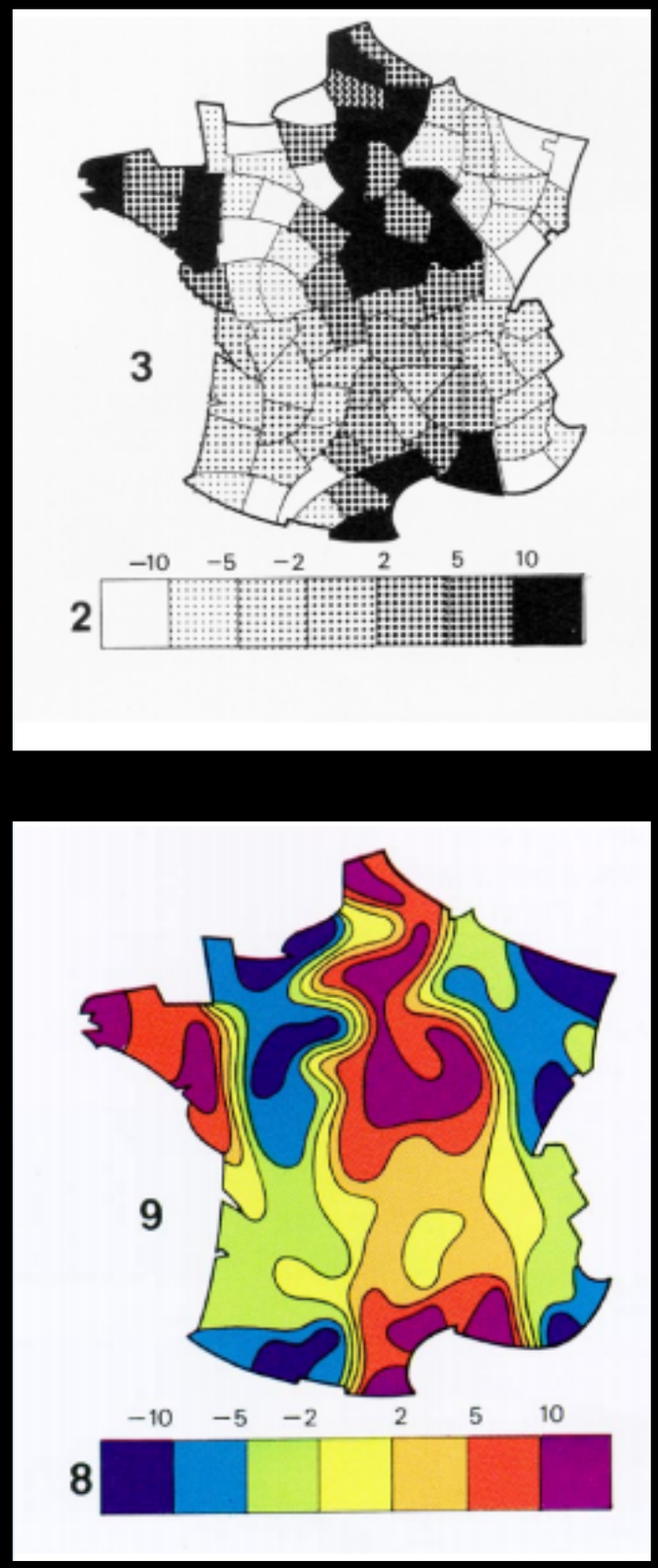


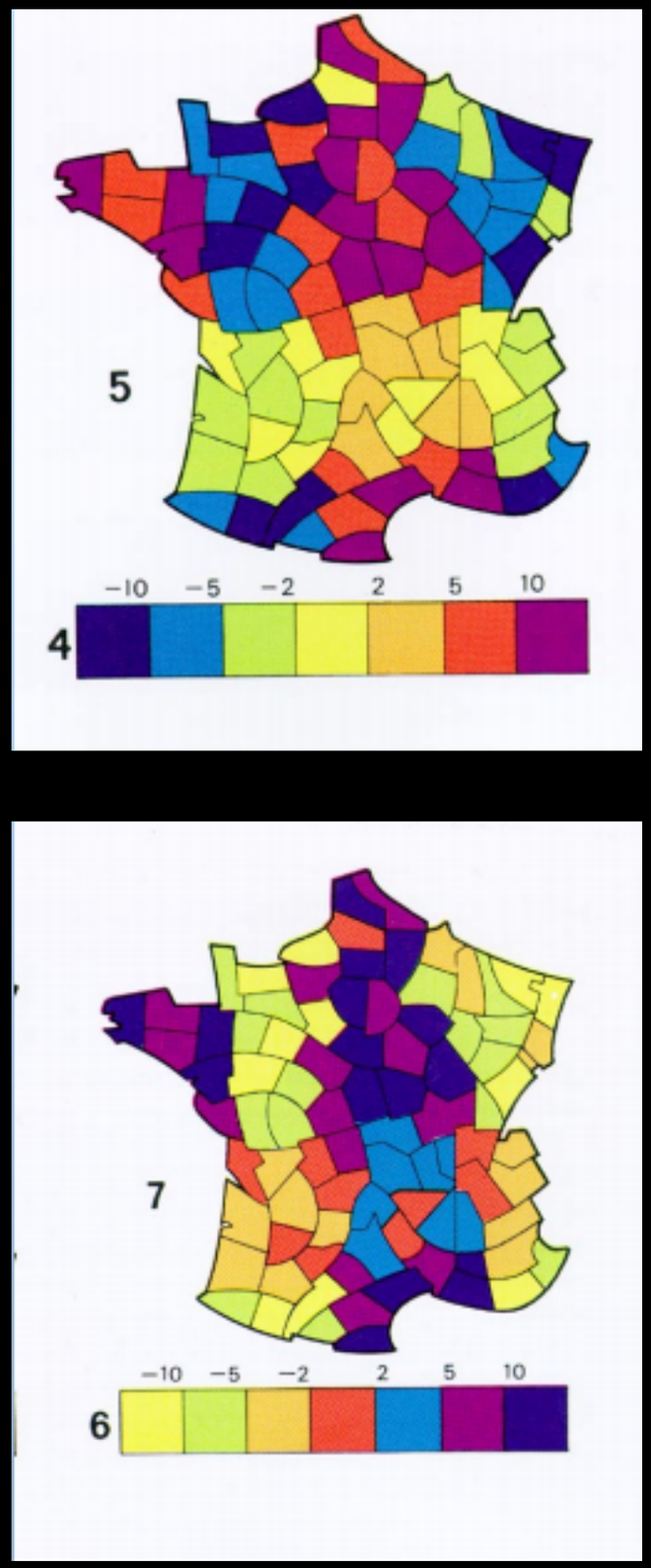



Which stands out to you? Do you see a division?


Level of Excellence in Relethounia
Excellent
Level 8
Level 7
Level 6
Level 5
Level 4
Level 3
Level 2
Level 1
Not Excellent
Level 6
Level 5
= one level difference =
Level 5
Level 4


1

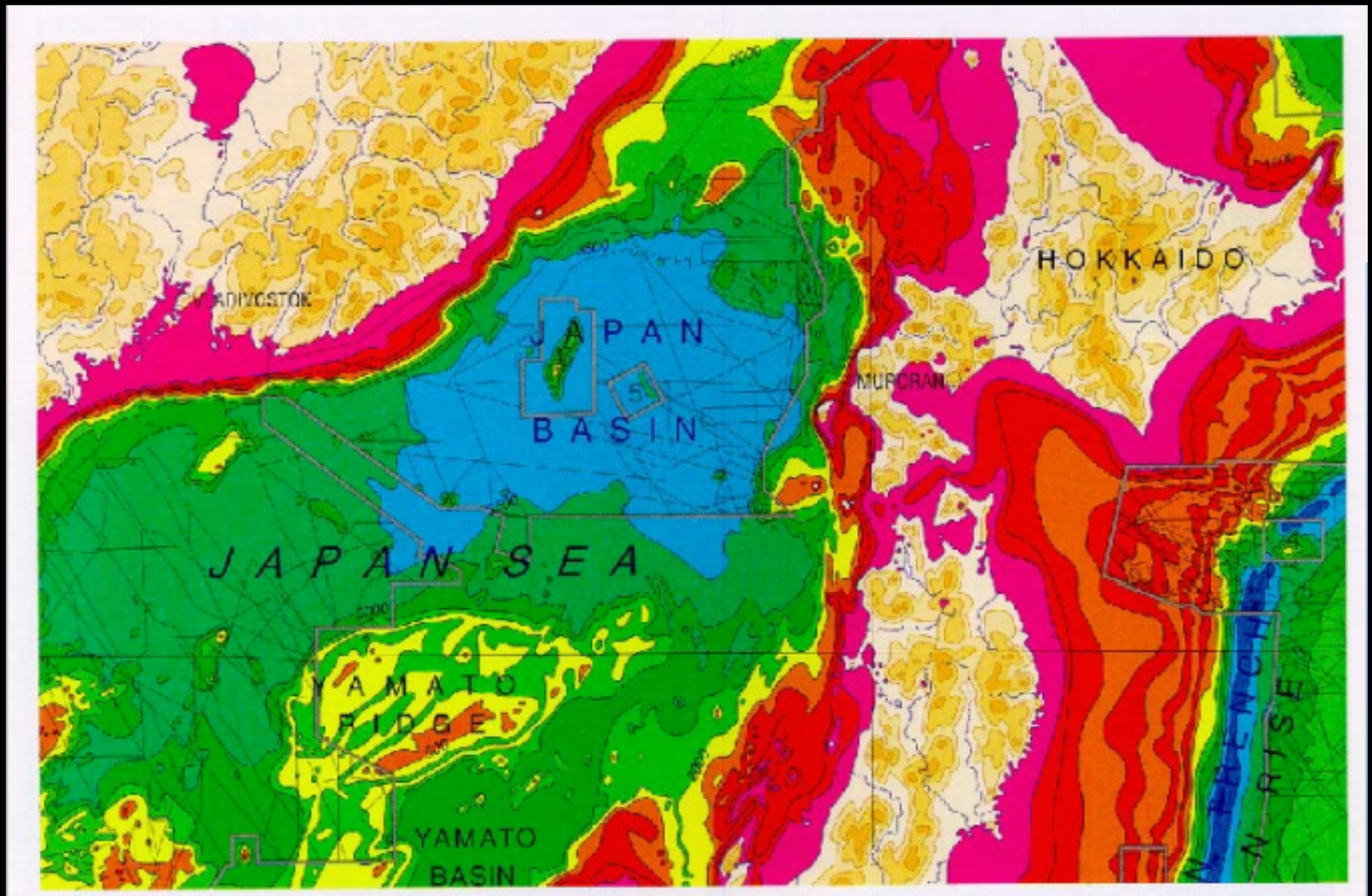

3
2
−10 −5 −2 2 5 10

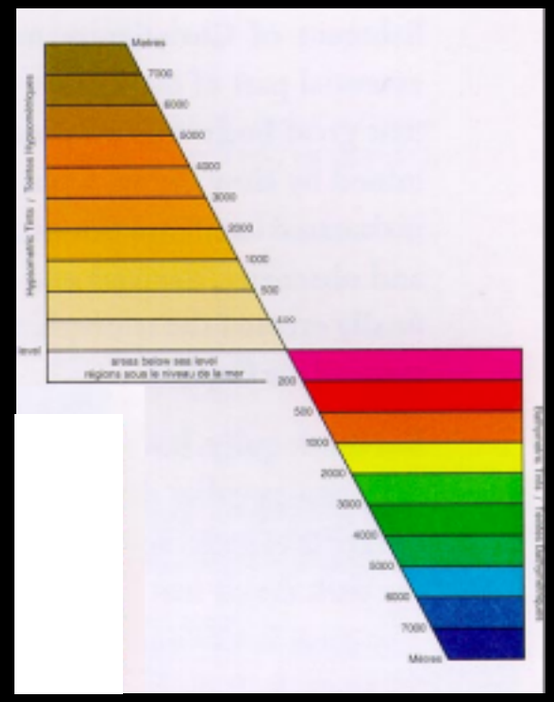

5
4
−10 −5 −2 2 5 10

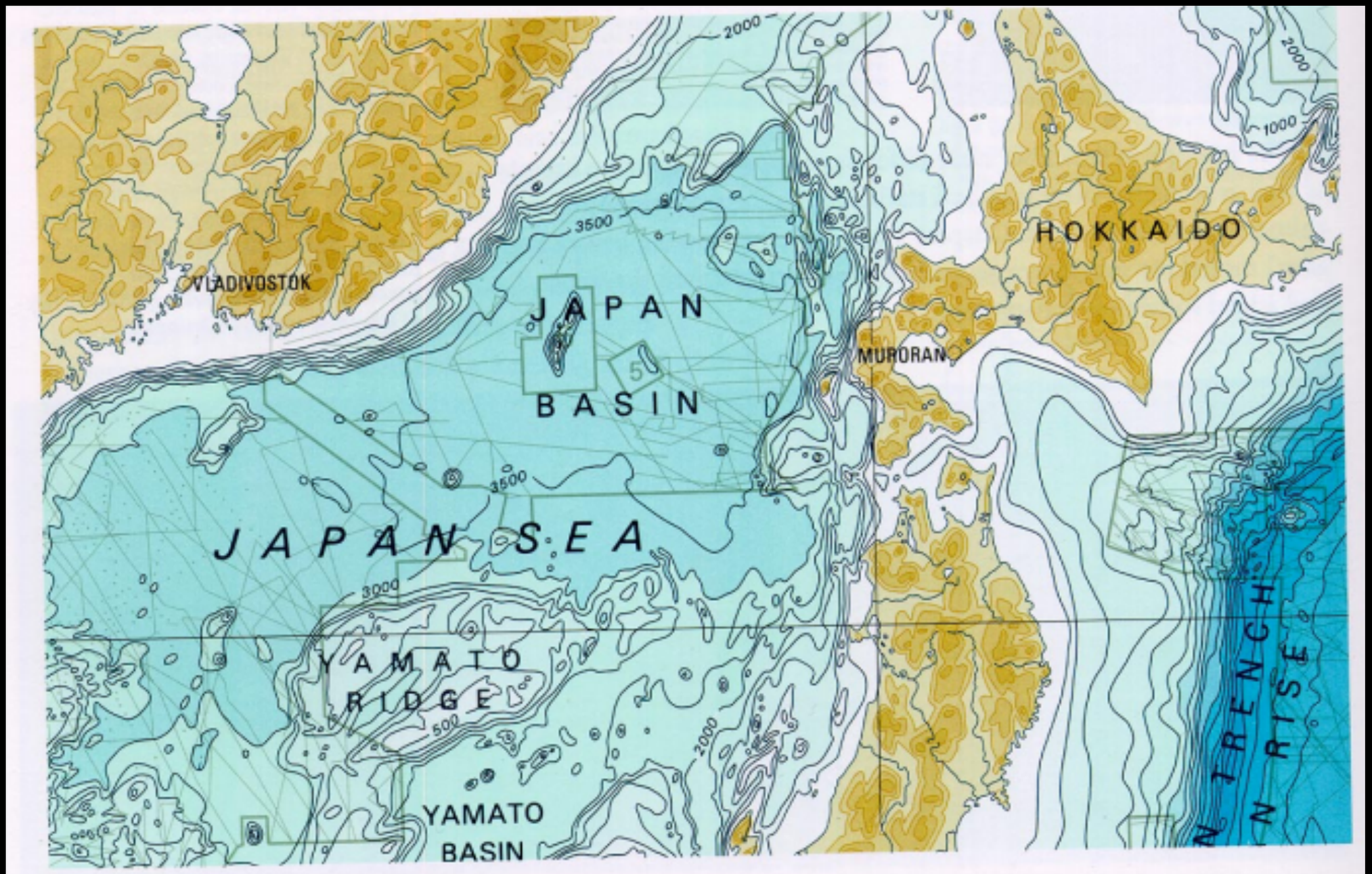

11
10
−10 −5 −2 2 5 10

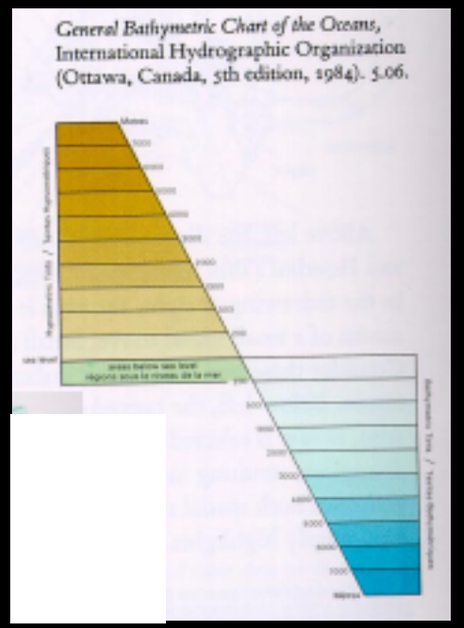

9
8
−10 −5 −2 2 5 10


7
6
−10 −5 −2 2 5 10


HOKKAIDO
VLADIVOSTOK
JAPAN
BASIN
MURORAN
JAPAN SEA
YAMATO
RIDGE
YAMATO
BASIN
KURILE TRENCH


Metres
7000
6000
5000
4000
3000
2000
1000
500
200
sea level
areas below sea level
régions sous le niveau de la mer
Metres


2000
3500
1000
VLADIVOSTOK
JAPAN
BASIN
HOKKAIDO
MURORAN
JAPAN SEA
3500
3000
YAMATO
RIDGE
500
2000
YAMATO
BASIN
TRENCH
RISE


General Bathymetric Chart of the Oceans,
International Hydrographic Organization
(Ottawa, Canada, 5th edition, 1984). 5.06.

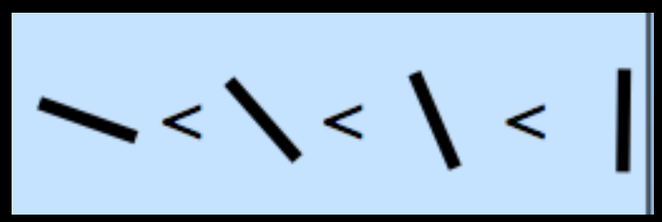
# ORIENTATION

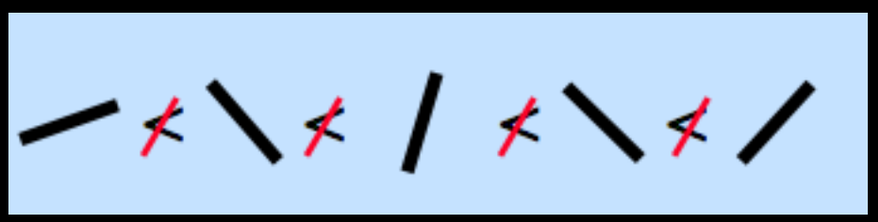
**selective**

**associative**

**quantitative**

**order**



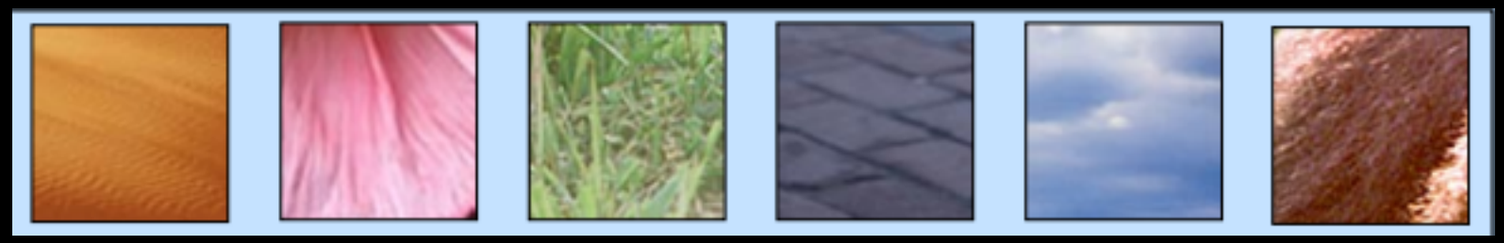



**length**

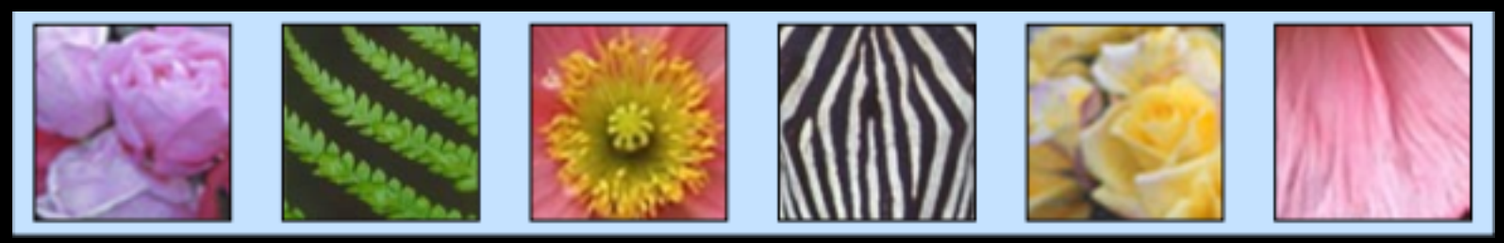
- ~5 in 2D; ? in 3D

# TEXTURE

✓ **selective**



✓ **associative**



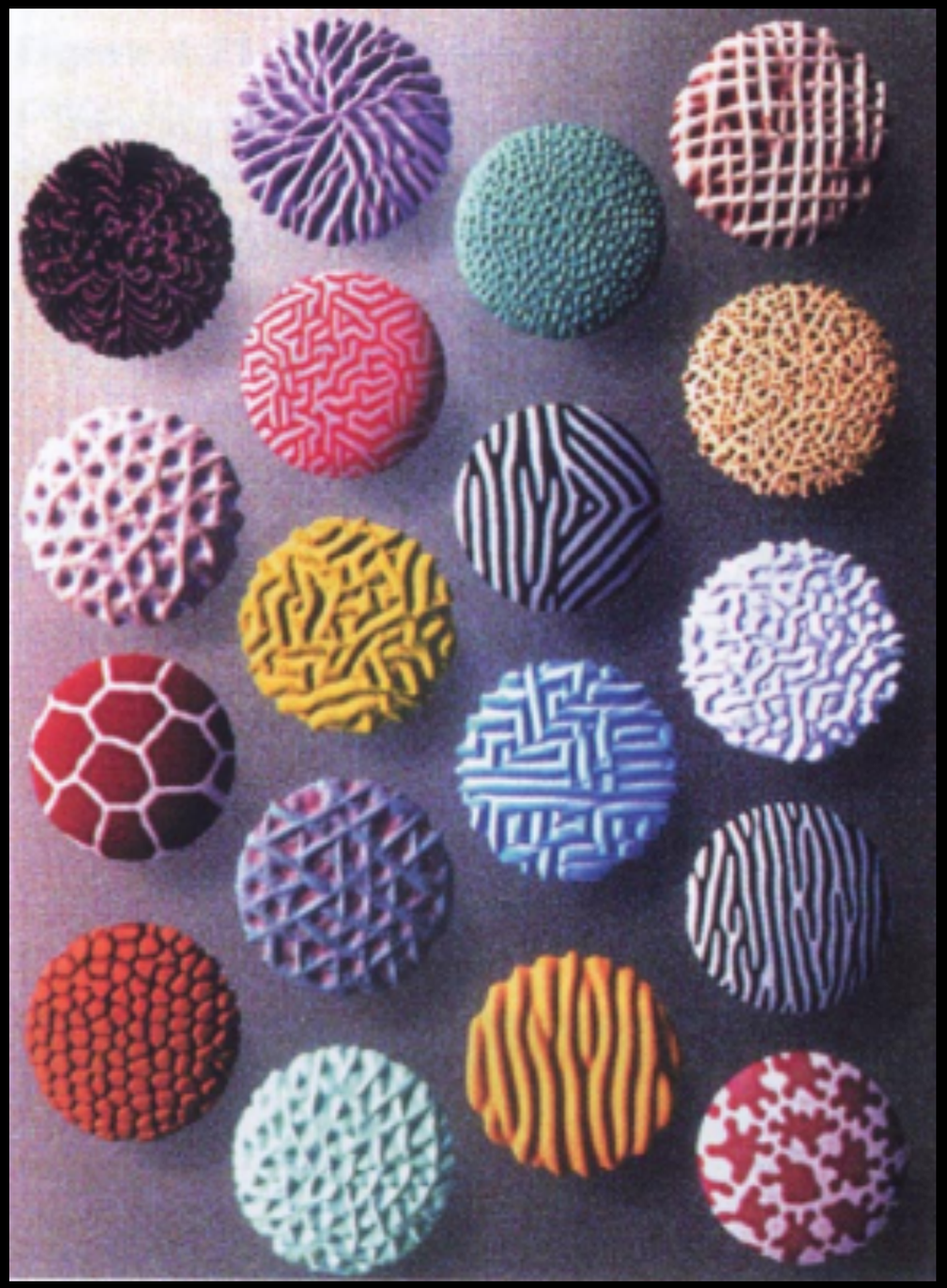
✗ **quantitative**

✗ **order**

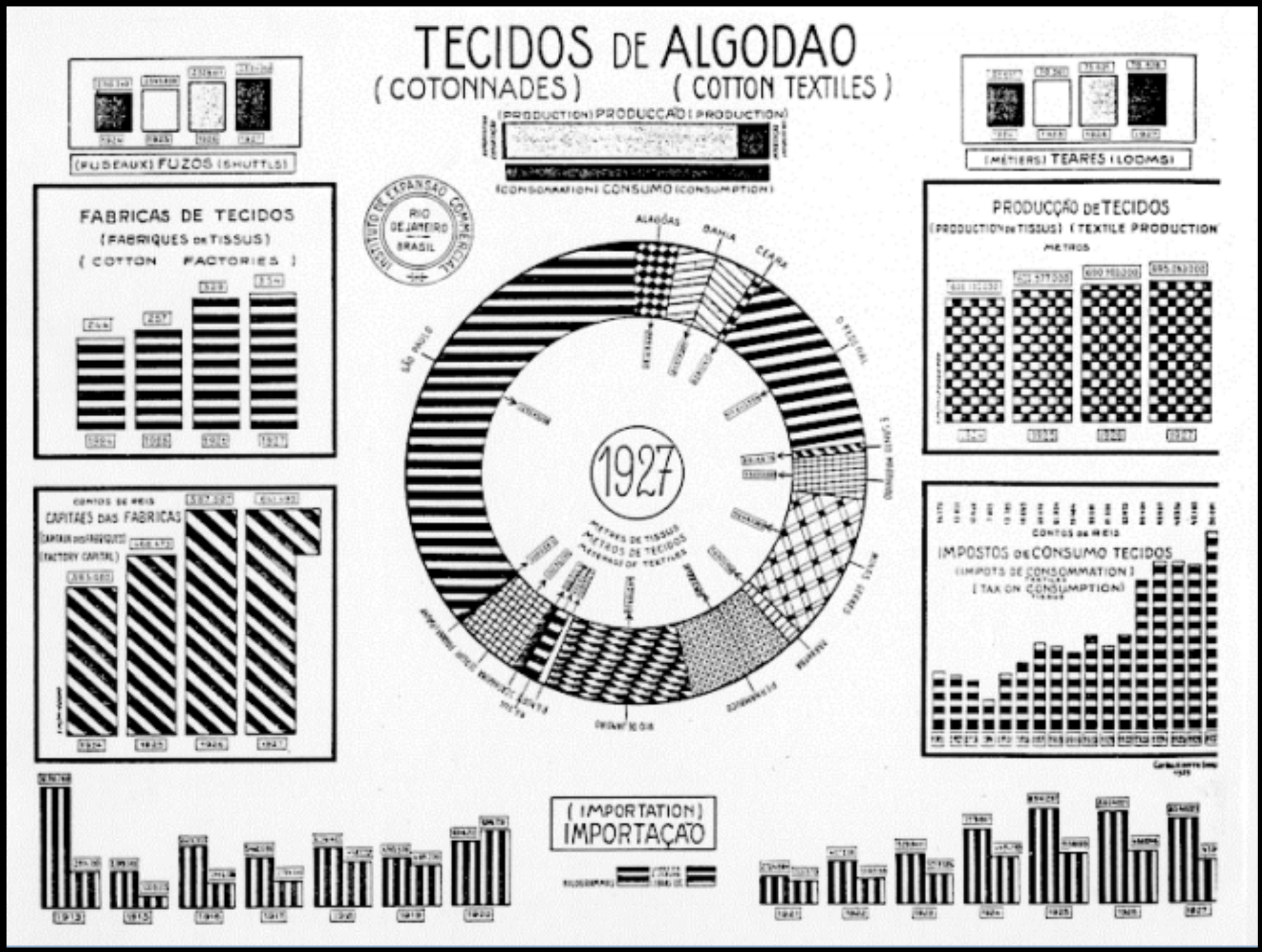
✓ **length**

- ~5 in 2D; ? in 3D

# TEXTURE

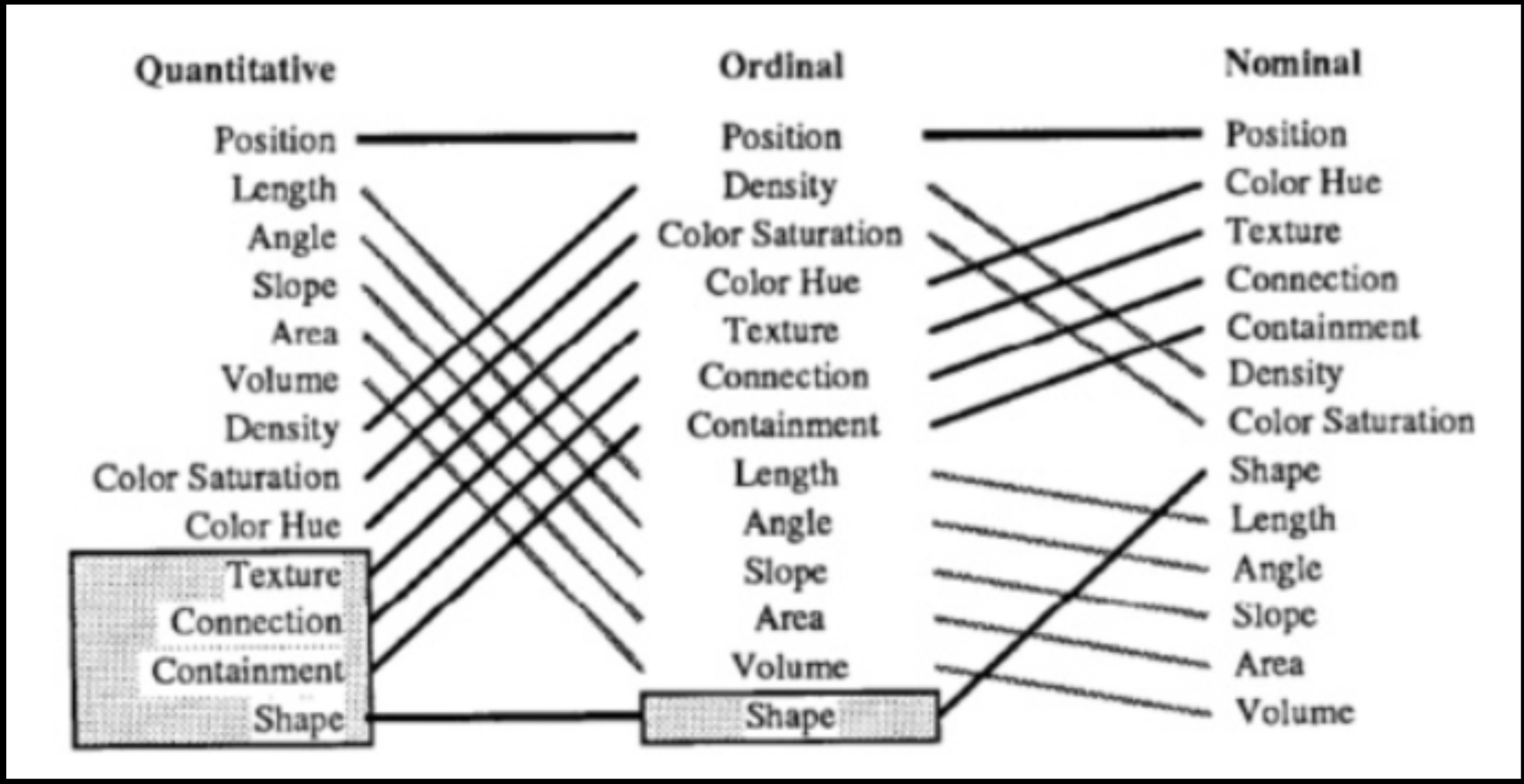


# TEXTURE




Cotton production in Brazil, 1927

# GUIDELINES FOR MAPPING




W. S. Cleveland and R. McGill. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. Journal of the American Statistical Association. 79(387). 1984

J. Mackinlay. Automating the Design of Graphical Presentations of Relational Information. ACM Trans. Graph. 5(2): 110–141, 1986.

# INFORMATION VISUALIZATION

Graphics should reveal the data

- show the data
- not get in the way of the message
- avoid distortion
- present many numbers in a small space
- make large data sets coherent
- encourage comparison between data
- supply both a broad overview and fine detail
- serve a clear purpose

*E. Tufte*
*Visual Display of Quantitative Information*