Introduction to Data Exploration

Visual Variables

Ira A. Fulton Schools of Engineering
Arizona State University

Objectives

Objective

Define the properties of Bertin’s visual variables

Data Types

Nominal:

Data whose categories have no implied ordering.

Ordinal:

Data that has a specified order, but no specified distance metric.

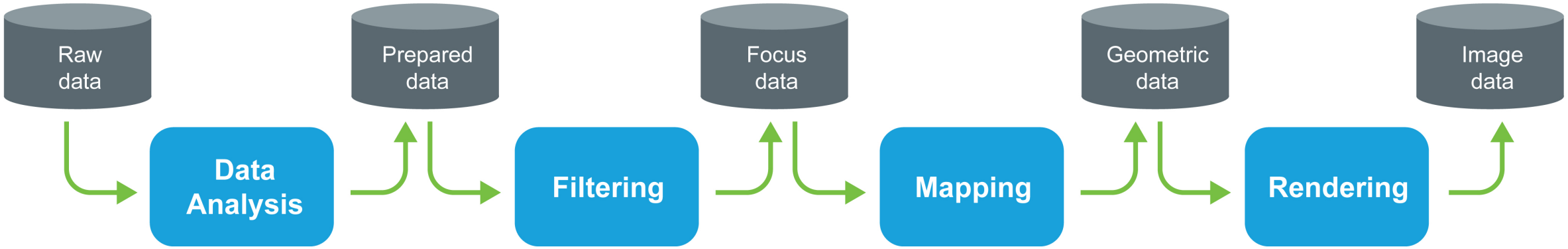
Interval:

Data that has measurable distances.

Ratio:

Same as interval, but include a zero point.

Visualization Pipeline



We want to take these different data types and map them to an appropriate visual representation

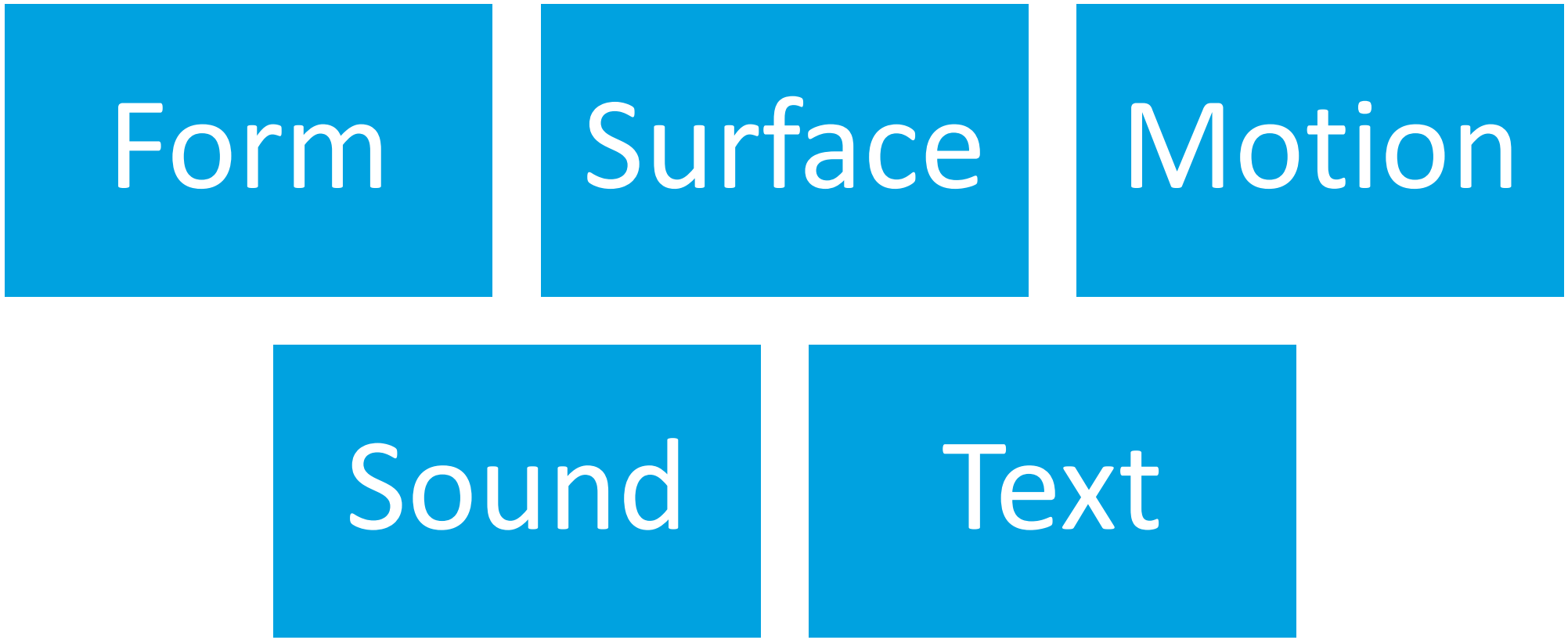
Data Analysis
Data are prepared for visualization (smooth, interpolate, transform)

Filtering
A subset of the data is selected for visualization

Mapping
Data are mapped to geometric primitives and their attributes

Rendering
Geometric data are transformed to image data

Mapping Data: Aesthetic Attributes



L Wilkinson (2005) *The Grammar of Graphics*

Aesthetic Attributes

- Must be capable of representing both continuous and categorical variables
  - Continuous variable: an attribute must vary primarily on **one** psychophysical dimension
  - Multidimensional attributes: must scale them on a single dimension
- Does not imply a linear perceptual scale

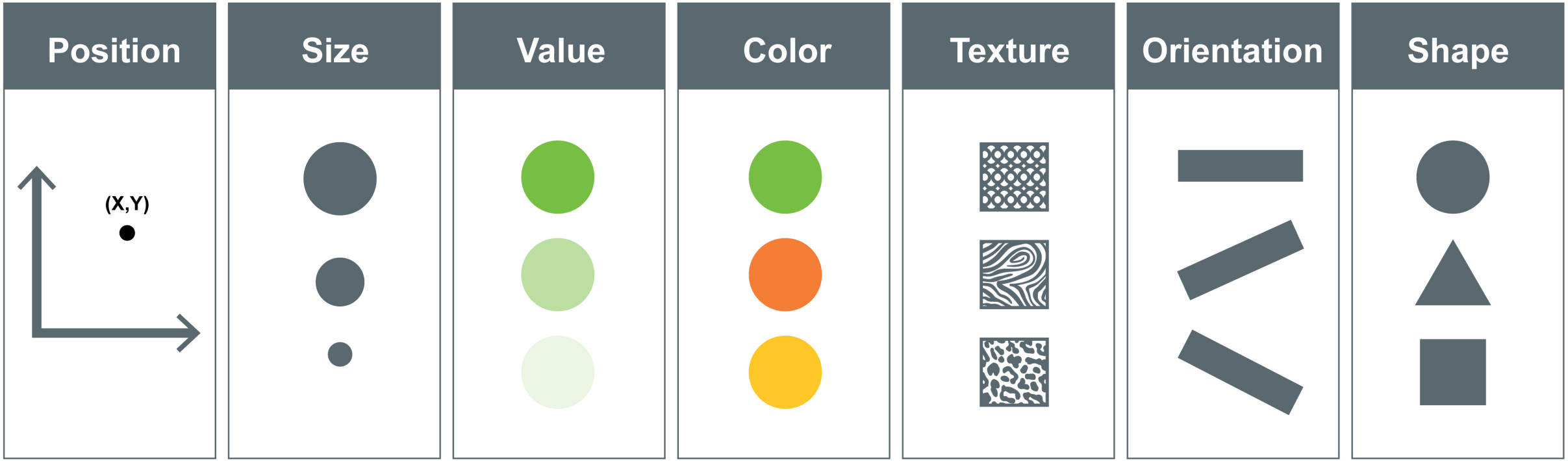
Graphic Design

Much of the skill in graphic design is knowing what combination of attributes should be avoided.

1-SM Kosslyn (1994), *The Elements of Graph Design*

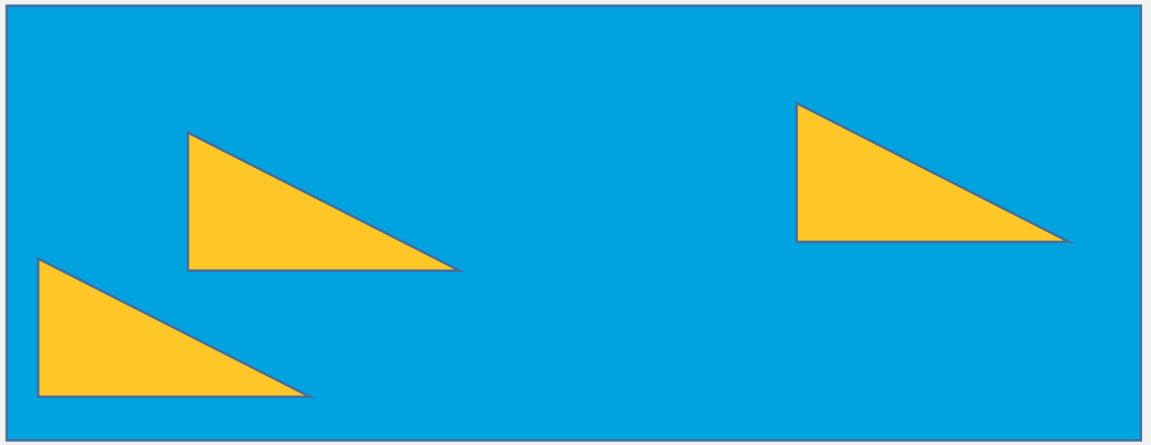
Bertin's Visual Variables

Visualization is concerned primarily with a mapping to visual form



Position

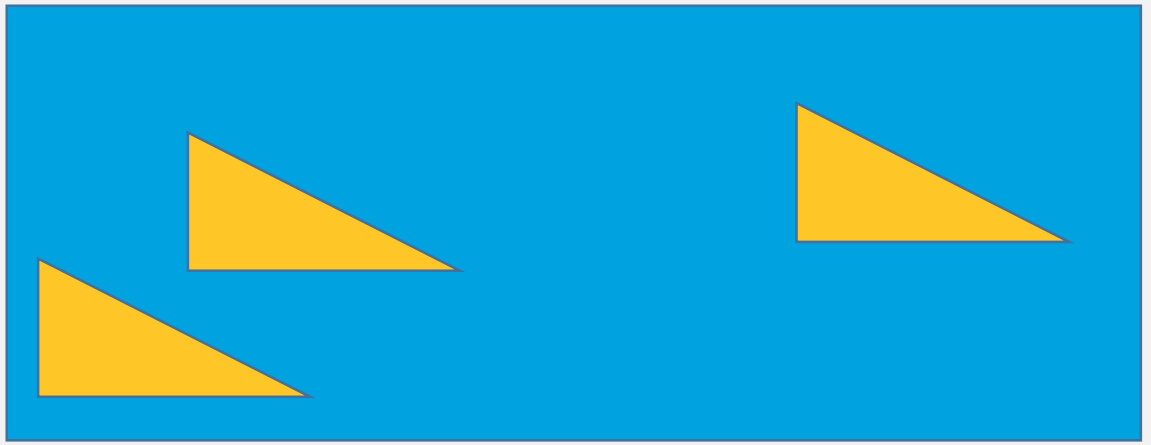
- A location in a multi-dimensional space
- **Continuous variables** map to densely distributed locations
- **Categorical variables** map to a lattice
- Ordering may or may not have meaning in terms of what is being measured



J Bertin (1967), *The Semiology of Graphics*

Position

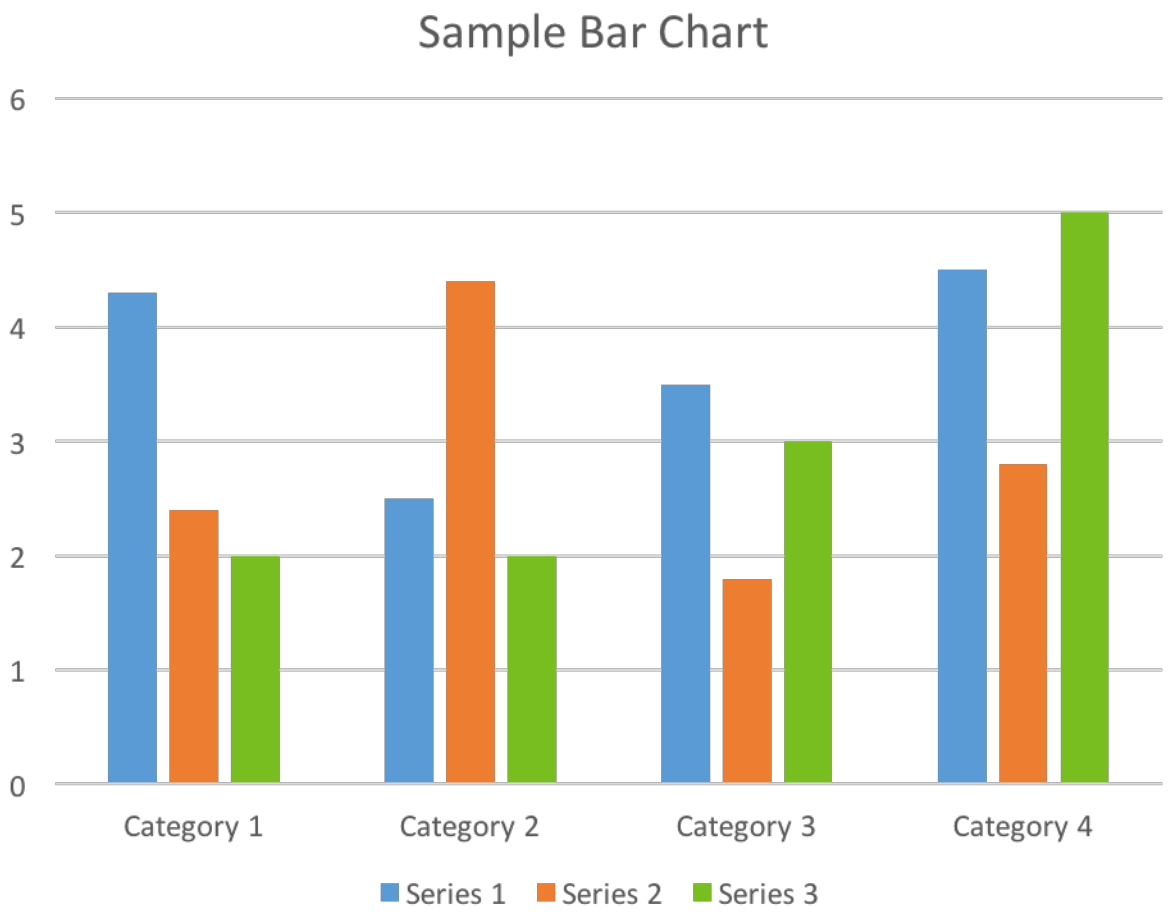
- Best way to represent a quantitative dimension visually
- Points or line lengths placed adjacent to a common axis enable judgments with the least bias or error



J Bertin (1967), *The Semiology of Graphics*

Size

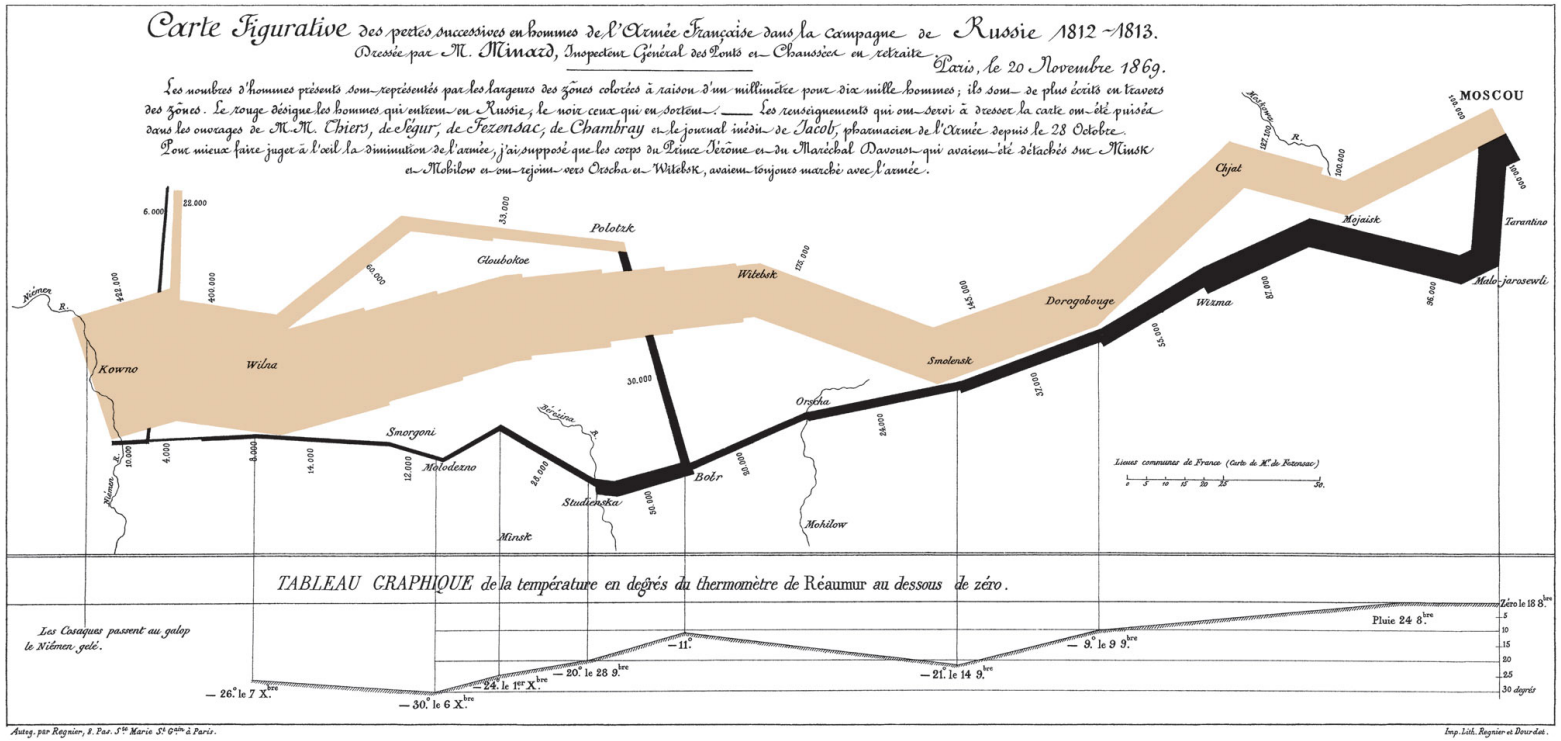
- The variation in terms of length or area
- In three dimensions, includes volume
- Area and volume representations among the **worst attributes** to use for graphing data



J Bertin (1967), *The Semiology of Graphics*

Size

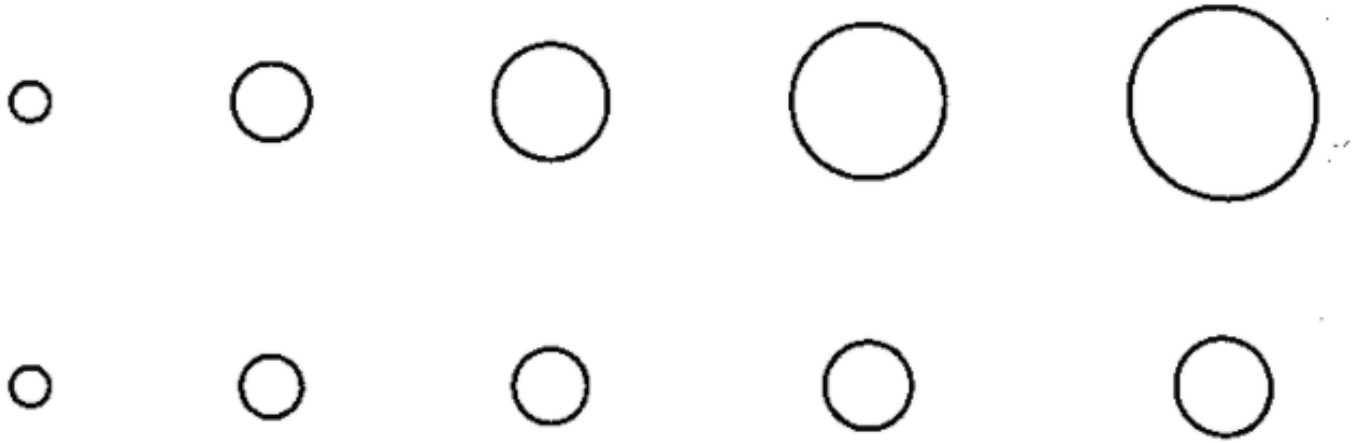
- Size for lines is usually equivalent to thickness
  - less likely to induce perceptual distortion
- Size can be used to great effect with path



1 – Charles Joseph Minard: Mapping Napoleon's March, 1861 by John Corbett, Center for Spatially Integrated Social Science

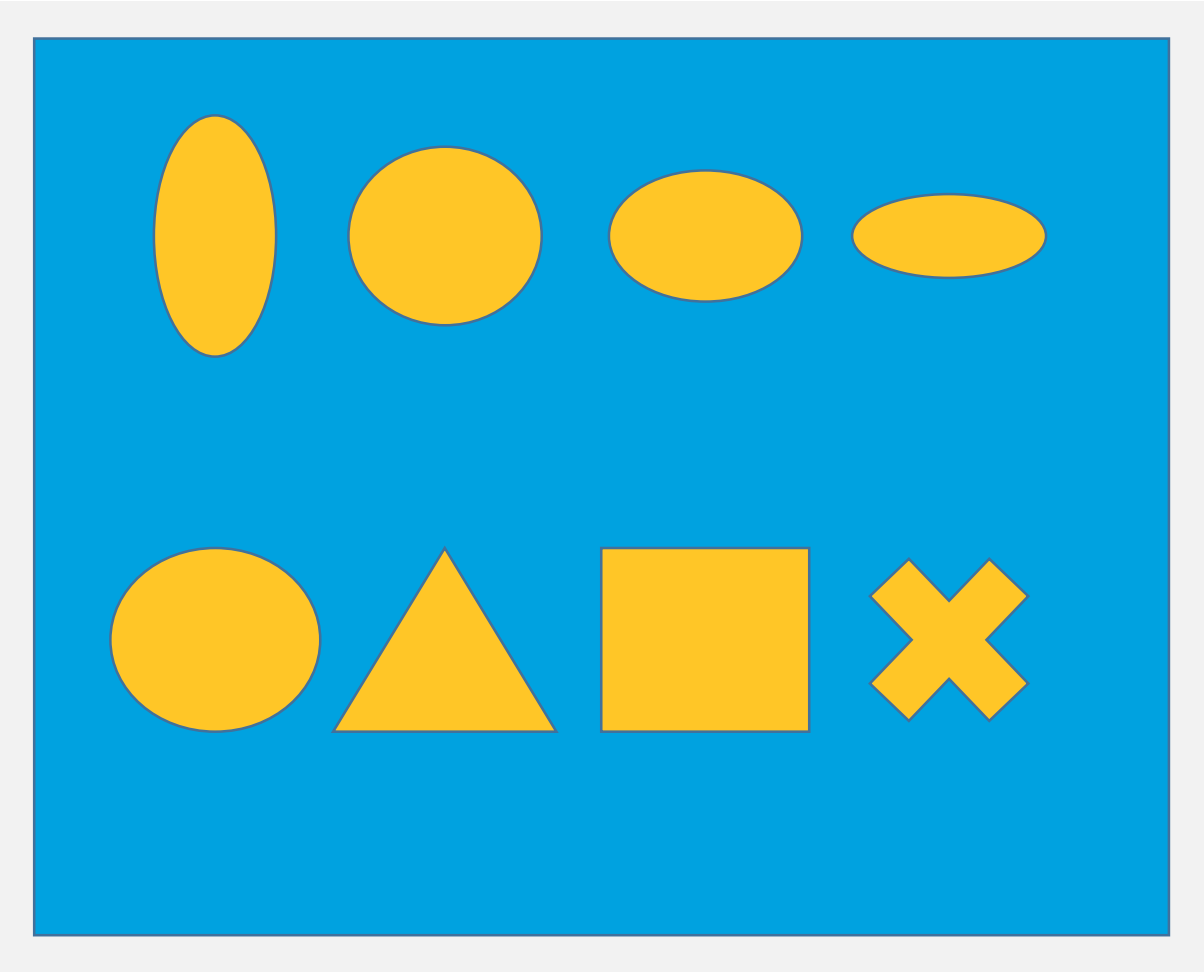
- For objects with rotational symmetry, map size to the diameter rather than area

- Representing data through area or volume should probably be confined to positively skewed data that can benefit from the perceptual equivalent of the square root transformation



Shape

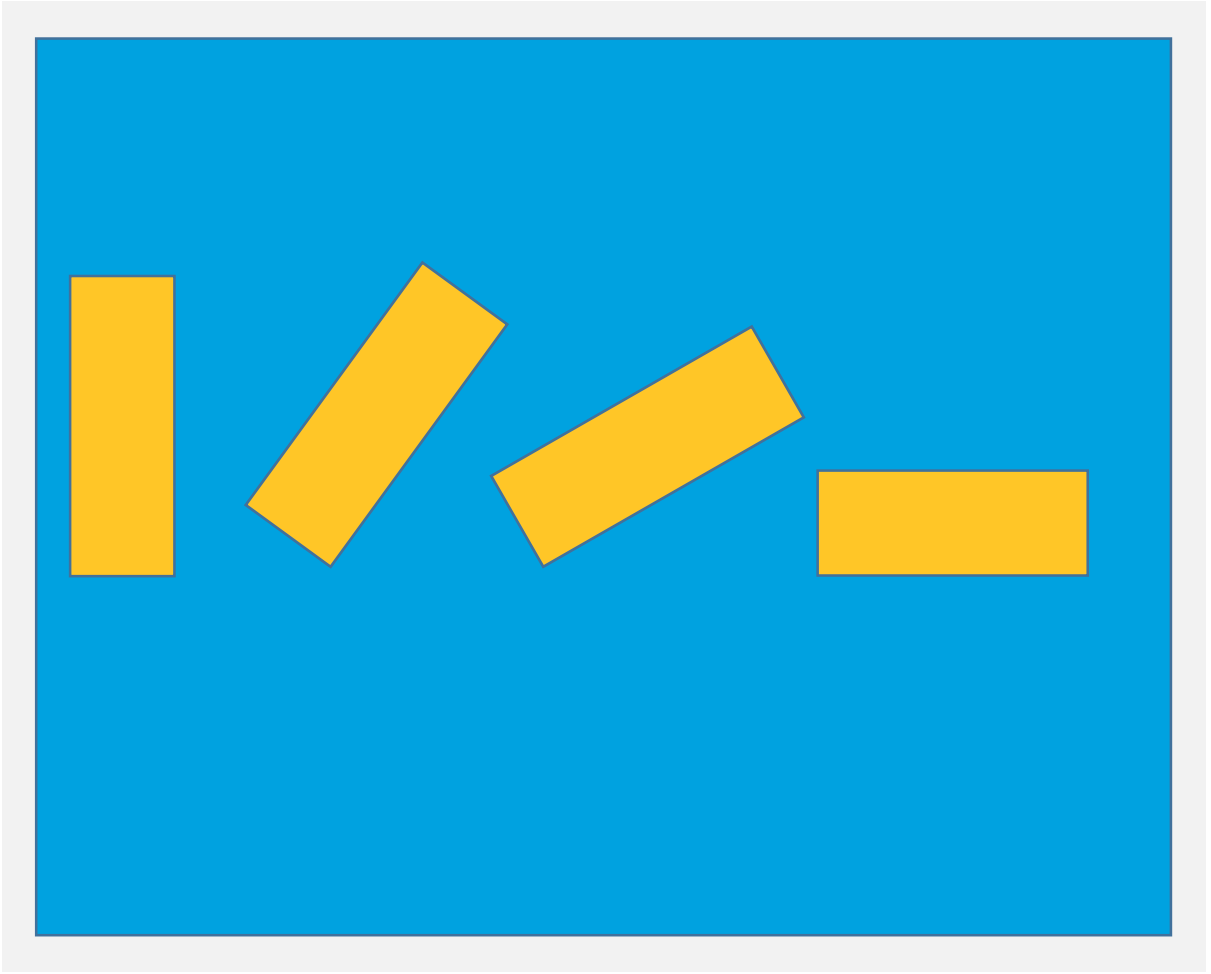
- The shape or boundary of an object
- Shape must vary without affecting size, rotation and other attributes
- **Example:**
  - Map symbols



J Bertin (1967), *The Semiology of Graphics*

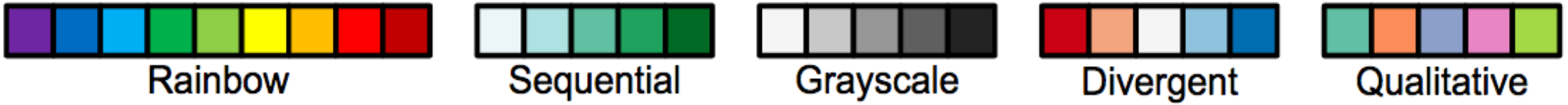
Rotation

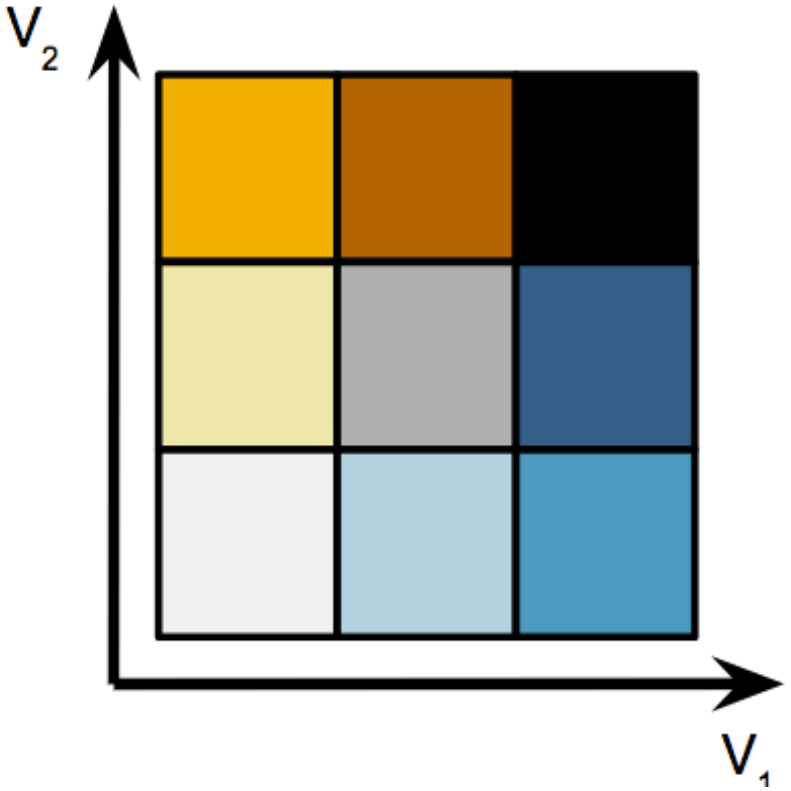
- Rotational angle of the graphic primitive
- Lines, areas and surfaces can only rotate if they are positionally unconstrained

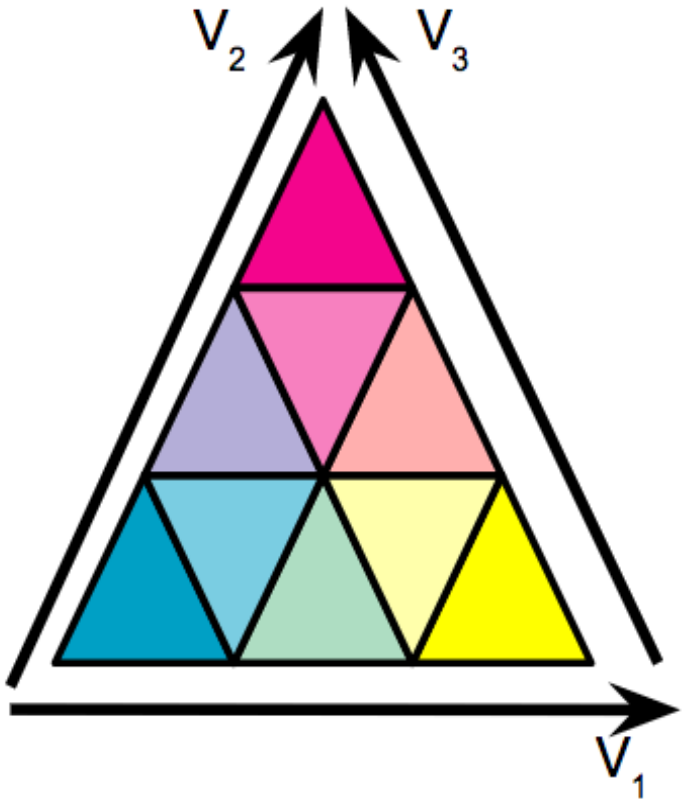


J Bertin (1967), *The Semiology of Graphics*

Color


Rainbow
Sequential
Grayscale
Divergent
Qualitative


V2
V1


V2
V3
V1

Texture

- Includes pattern, granularity and orientation
  - Granularity
    - repetition of a pattern per unit of area
  - Orientation
    - Angle of the pattern
- Texture alone can be a basis for perception

J Bertin (1967), *The Semiology of Graphics*

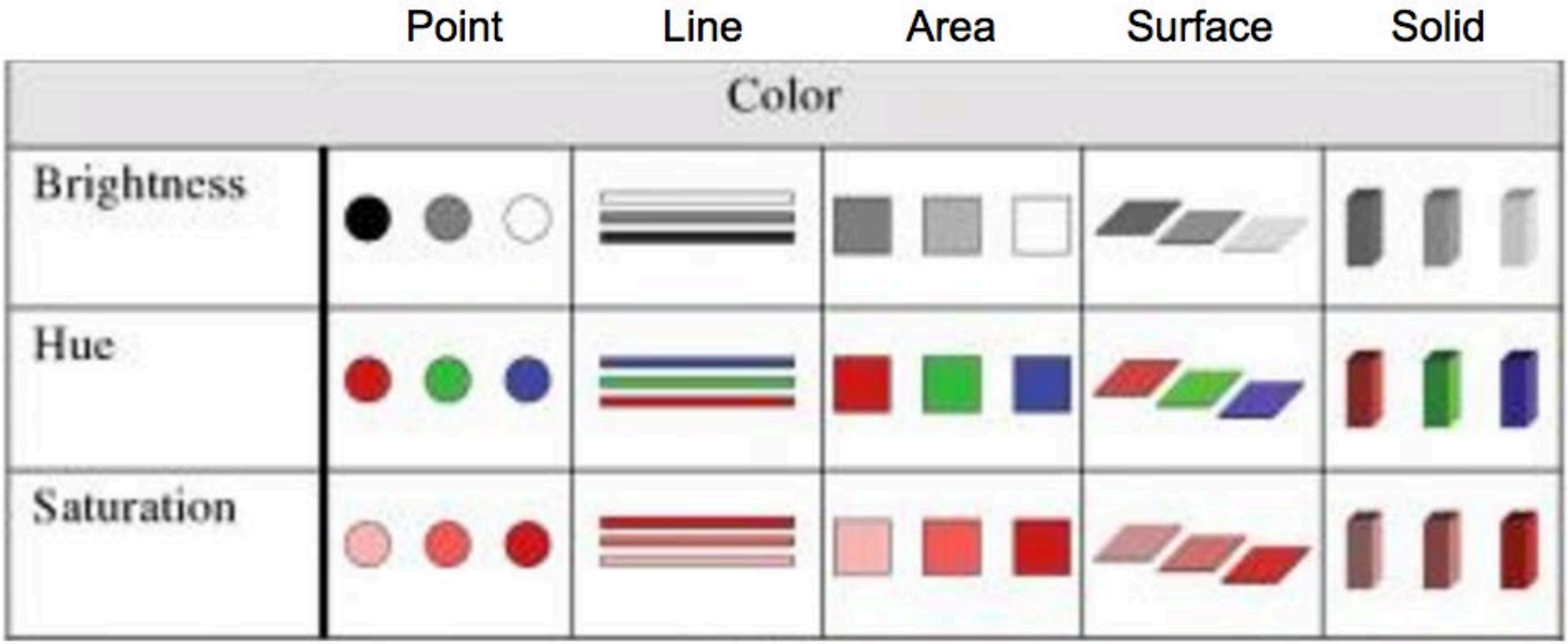
Texture

Textures can be described in a variety of ways

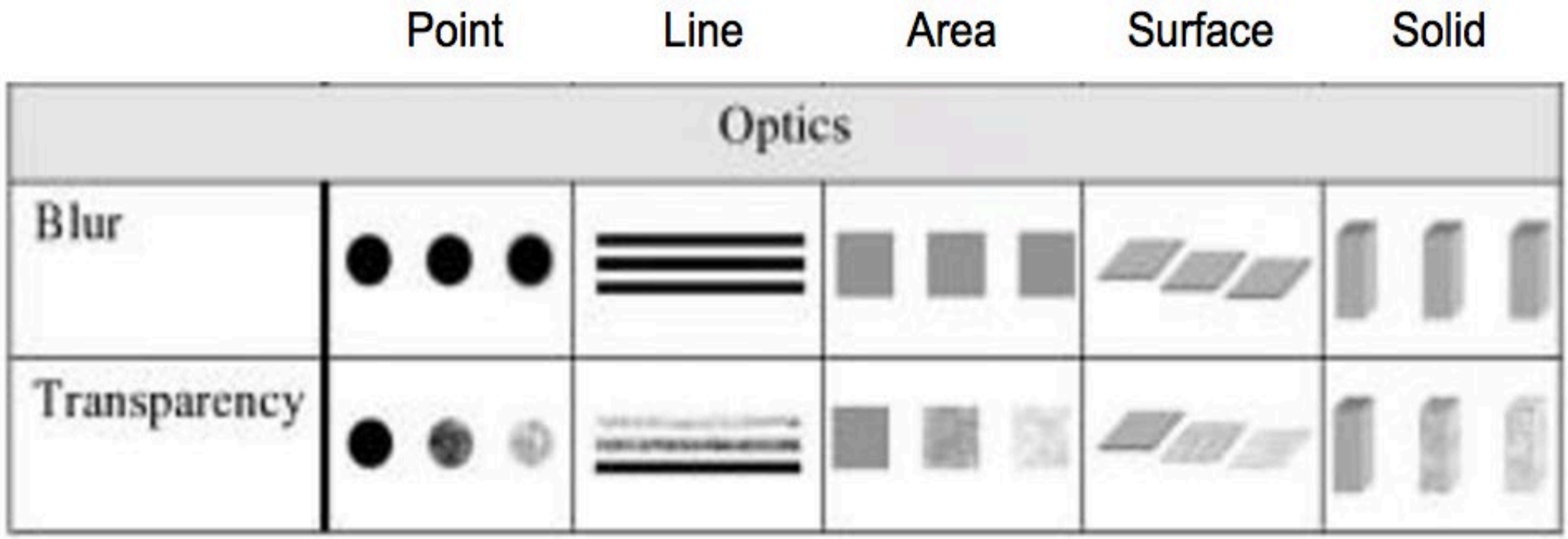
- **Fourier transform** – decomposes a grid of brightness values into sums of trigonometric components
- **Auto-correlogram** – characterize the spatial moments of a texture

J Bertin (1967), *The Semiology of Graphics*

	Point	Line	Area	Surface	Solid
Form					
Size					
Shape					
Rotation					


Point
Line
Area
Surface
Solid
Color
Brightness
Hue
Saturation

	Point	Line	Area	Surface	Solid
Texture					
Granularity					
Pattern					
Orientation					


Point
Line
Area
Surface
Solid
Optics
Blur
Transparency