Visualization Foundations

Prof. Dr. Renato Pajarola
Visualization and MultiMedia Lab
Department of Informatics
University of Zürich

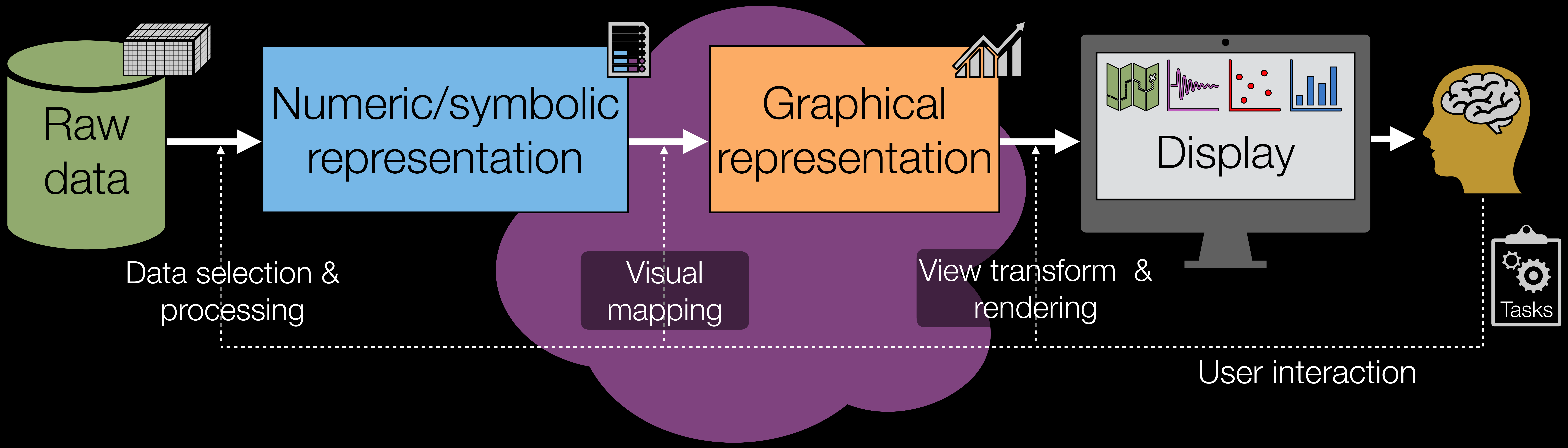
VISUALIZATIONAND
MULTIMEDIALAB

Copyrights

Overview

1. Review of the Visualization Process
2. Semiology of Graphical Symbols
3. The Eight Visual Variables

Where are we in the Visualization Pipeline?



- Data preprocessing and transformation
  - Selection of information and mapping to fundamental computer data types
  - Data cleaning, interpolation, sampling, filtering, aggregation, partitioning
- Mapping for visualization
  - Specific visual representation (geometry, color)
  - Embedding in Euclidean 2D/3D space
- Rendering transformations
  - Final image synthesis by 2D imaging and 3D graphics technology
  - Interactive data and view selection

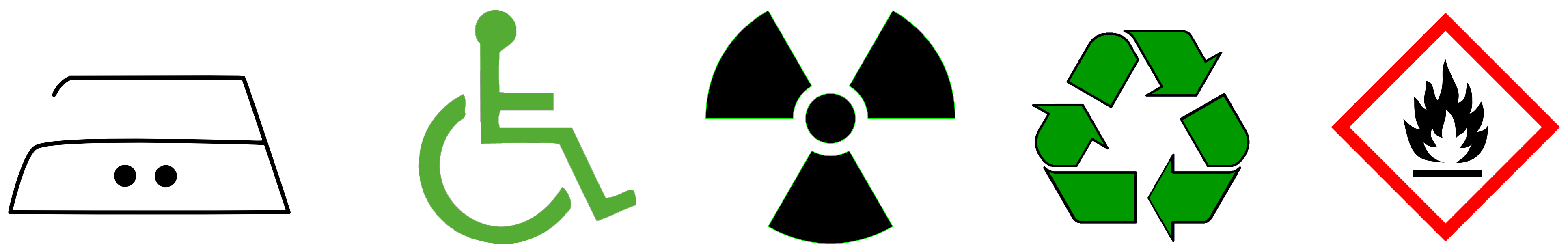
Semiology of Graphical Symbols

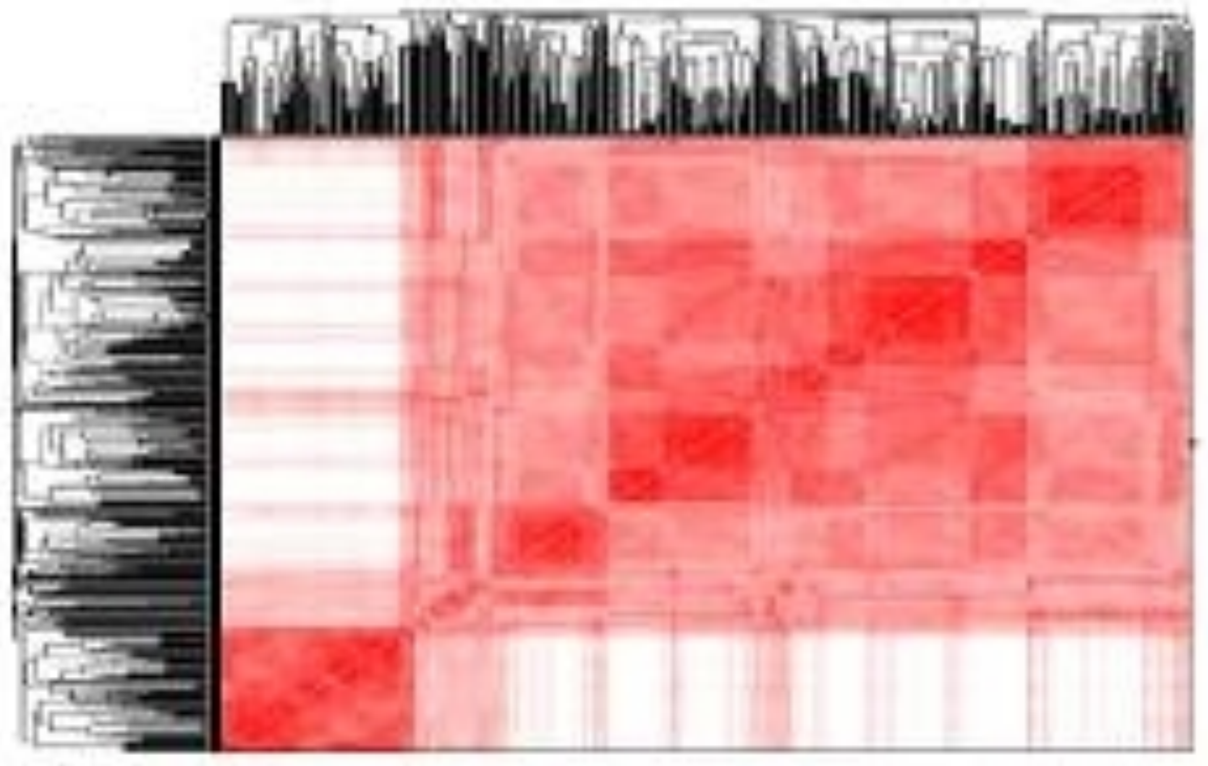
- Visualizations typically contain visual objects, graphical symbols
  - Semiology is the science of signs, graphical symbols, symbolism and communication
- Diagrams, networks, maps, plots all contain graphical symbols
- Good visualizations need well designed symbols and an understanding how they are perceived
  - Consider the common interpretation and semantic of symbols from everyday life

Symbol with obvious meaning

Symbols and Visualizations

- Some common symbols are universally recognizable
  - Perceived in one step through clearly associated meaning
  - May be preattentively recognizable
- Complex representations need multiple steps for understanding
  - Identify major graphical image elements
  - Identify relationships between them
- Human visual perception driven by physical interpretation
  - Visualization must have easily interpretable spatial (x,y,z) dimensions
- Any pattern, variation or order in the image must imply a pattern, variation or order in the data
  - Otherwise it is an artifact of the presentation



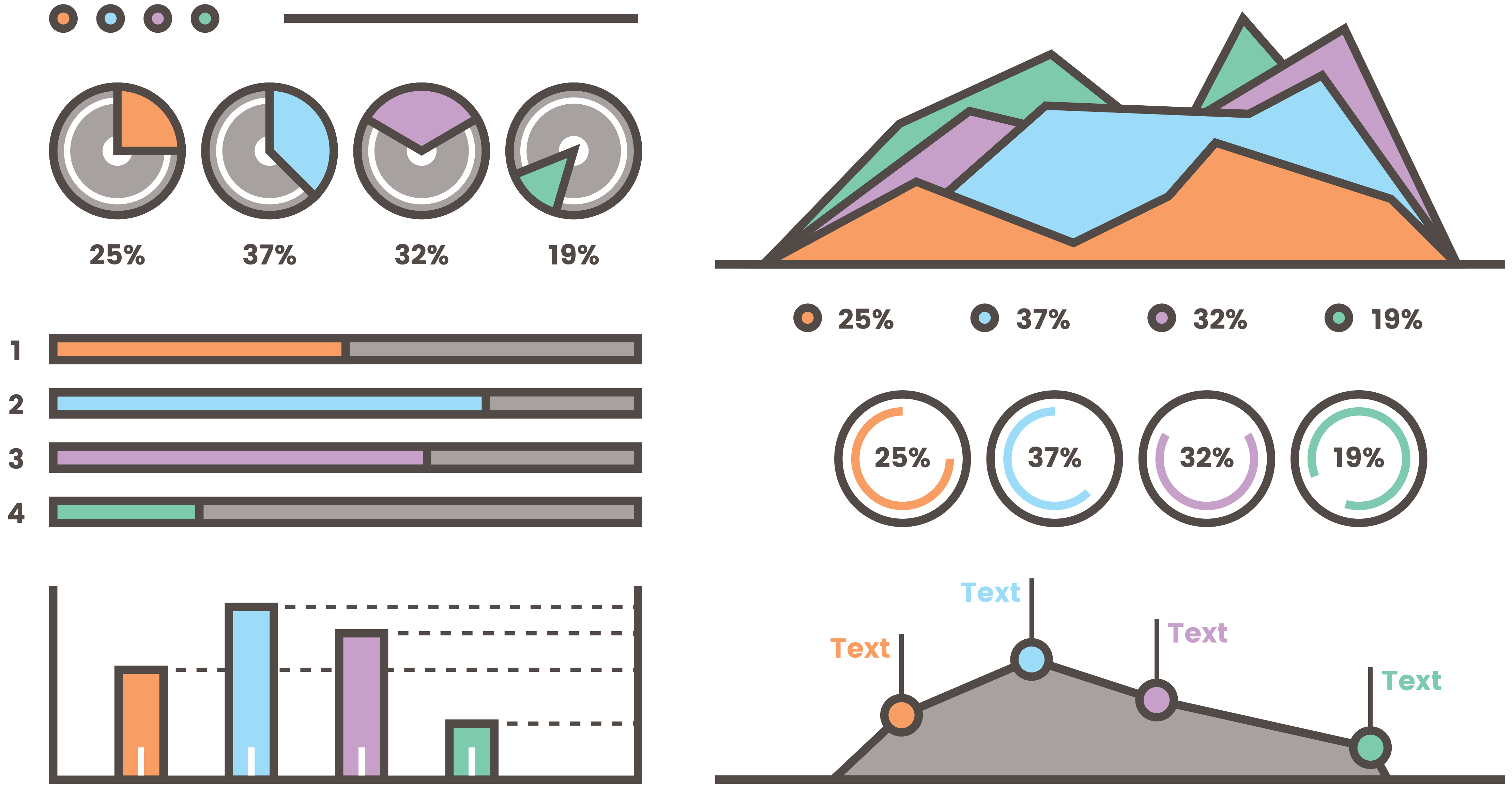


similarity in data ⇔ visual similarity of corresponding symbols

order between data items ⇔ visual order between corresponding symbols

Graphical Elements in Visualization

- Eventually a visualization is an image (on screen)
  - Consisting of objects, shapes, colors, patterns …
- Graphical elements are interpreted as flat (in 2D) or solid objects (in 3D)
  - x,y-position in (image) plane and size of symbol
- A visualization can contain large numbers of graphical elements, themselves potentially forming a texture
  - Desired or unwanted effect
- Groups of objects may be recognized
  - Followed by cognitively characterizing these groups



Marks

- A mark is made to represent some information other than itself
  - Represents a data point in a visualization
- Marks can be
  - **Points** are dimensionless locations on the plane, represented by signs that obviously need to have some size, shape or color for visualization
  - **Lines** represent information with a certain length, but no area and therefore no width, and are visualized by signs of some thickness
  - **Areas** have a length and a width and therefore a two-dimensional size
  - **Surfaces** are areas in a three-dimensional space, but with no thickness
  - **Volumes** have a length, a width and a depth, and are thus truly (filling) three-dimensional (space)

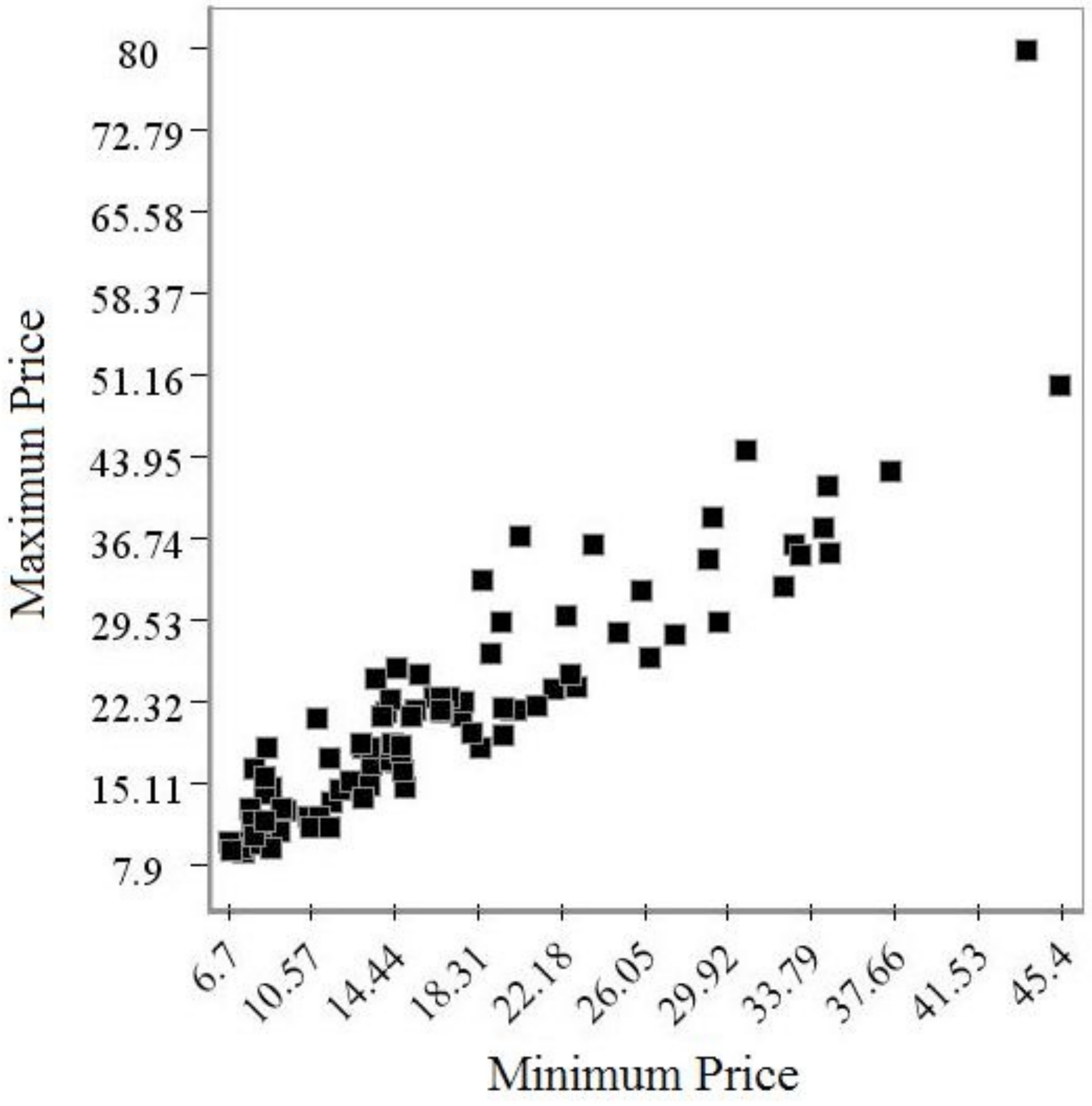
The Eight Visual Variables

1. Position
2. Shape (graphical primitive)
3. Size (length, area and volume)
4. Brightness
5. Color
6. Orientation
7. Texture
8. Motion

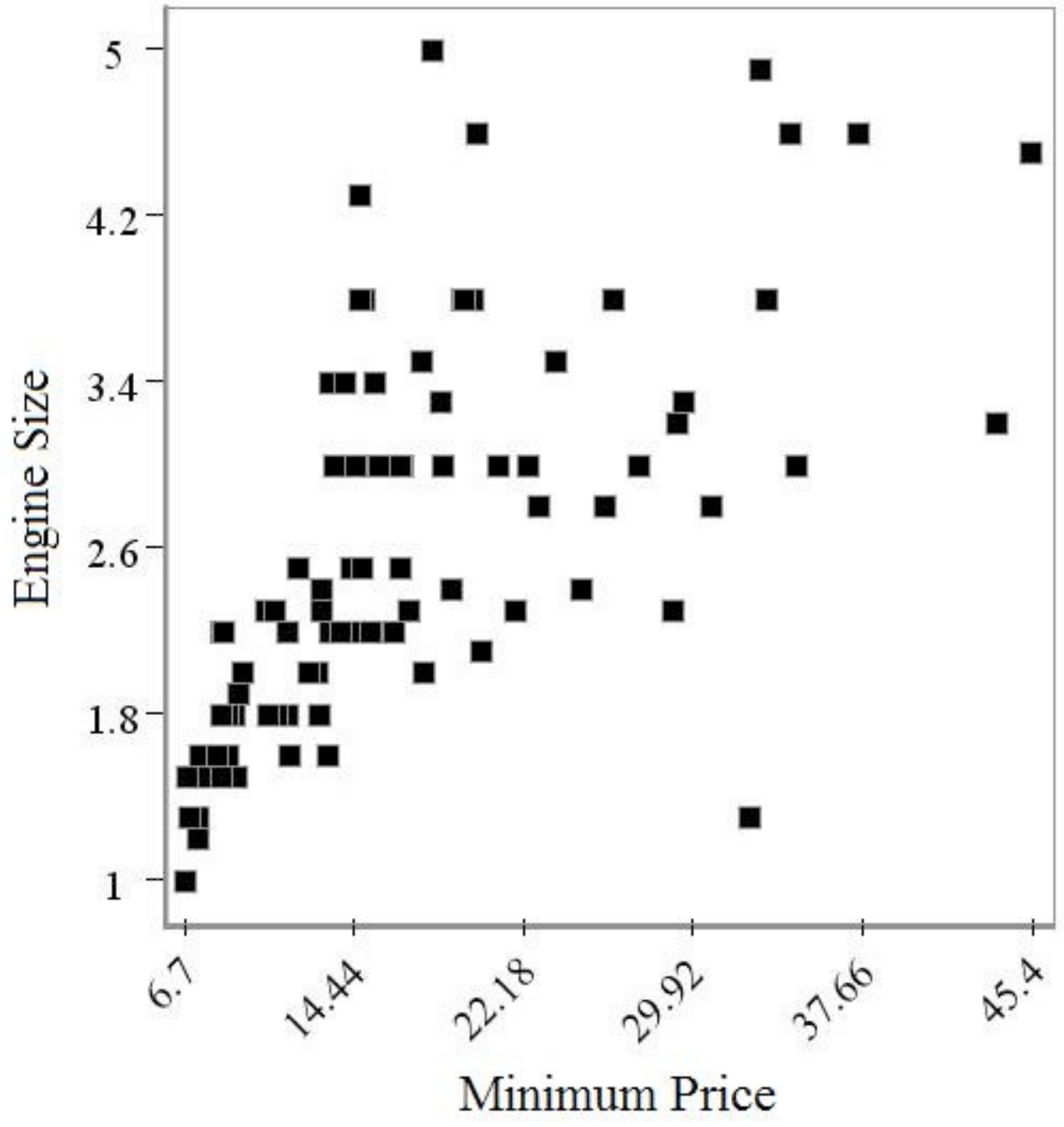
- Effects of Visual Variables

1. Position

- 1-, 2- or 3-dimensional
- Most important variable with greatest visual impact
- Space and pixels are limited
  - Cannot map multimillion data elements to unique positions on typical screens
- Linear and logarithmic scales can be applied
  - Consider distortion effects
- Mapping of two or more variables is generally a projection
  - From higher dimensional data to display space
    - eventually the image plane

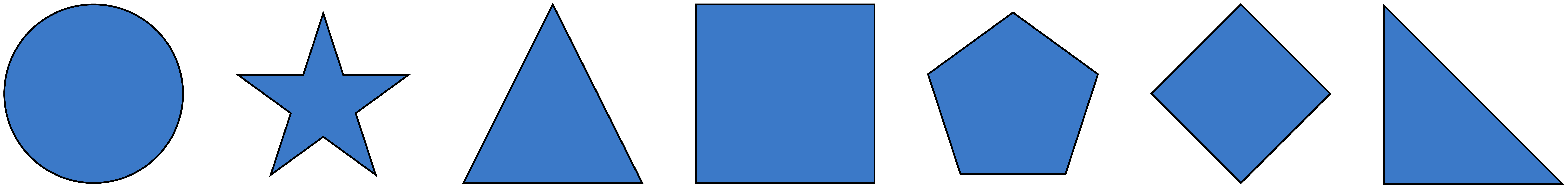


Display minimum price versus maximum price for cars with a 1993 model year. The spread of points appears to indicate a linear relationship between minimum and maximum price



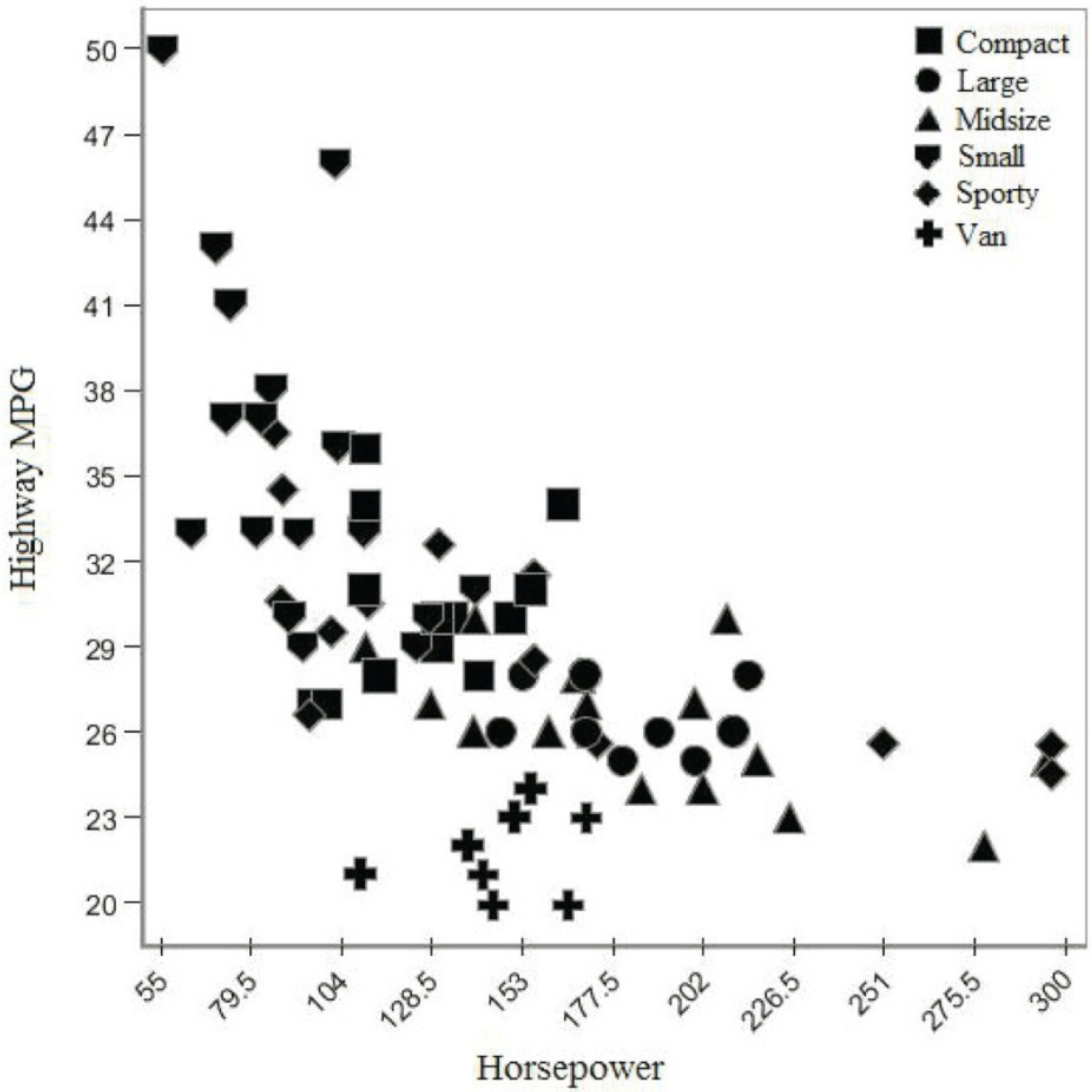
Compares minimum price with engine size for the 1993 cars data set. There does not appear to be a strong relationship between these two variables.

2. Shape

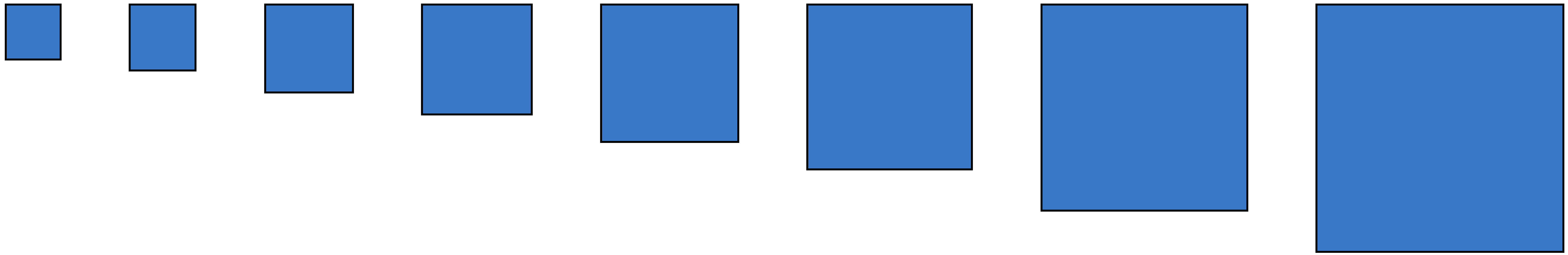


- The shape or *mark* can take on various forms
  - Points, lines, areas, volumes, complex 2D/3D shapes, symbols, letters, words
  - Care has to be taken with respect to differences in their relative appearance
    - covered area may be perceived as strength or some ordinal value
- Differentiability between marks important
  - Qualitative separation of data variable values
  - Hundreds or thousands of marks cannot be distinguished effortlessly

- Easy discernible marks
  - At least in low-density low-overlap areas
- Example: the mark is used to separate categories on top of a relational scatterplot
  - Comparing engine power and highway mileage for car types
  - Clusters are visible, as well as some outliers



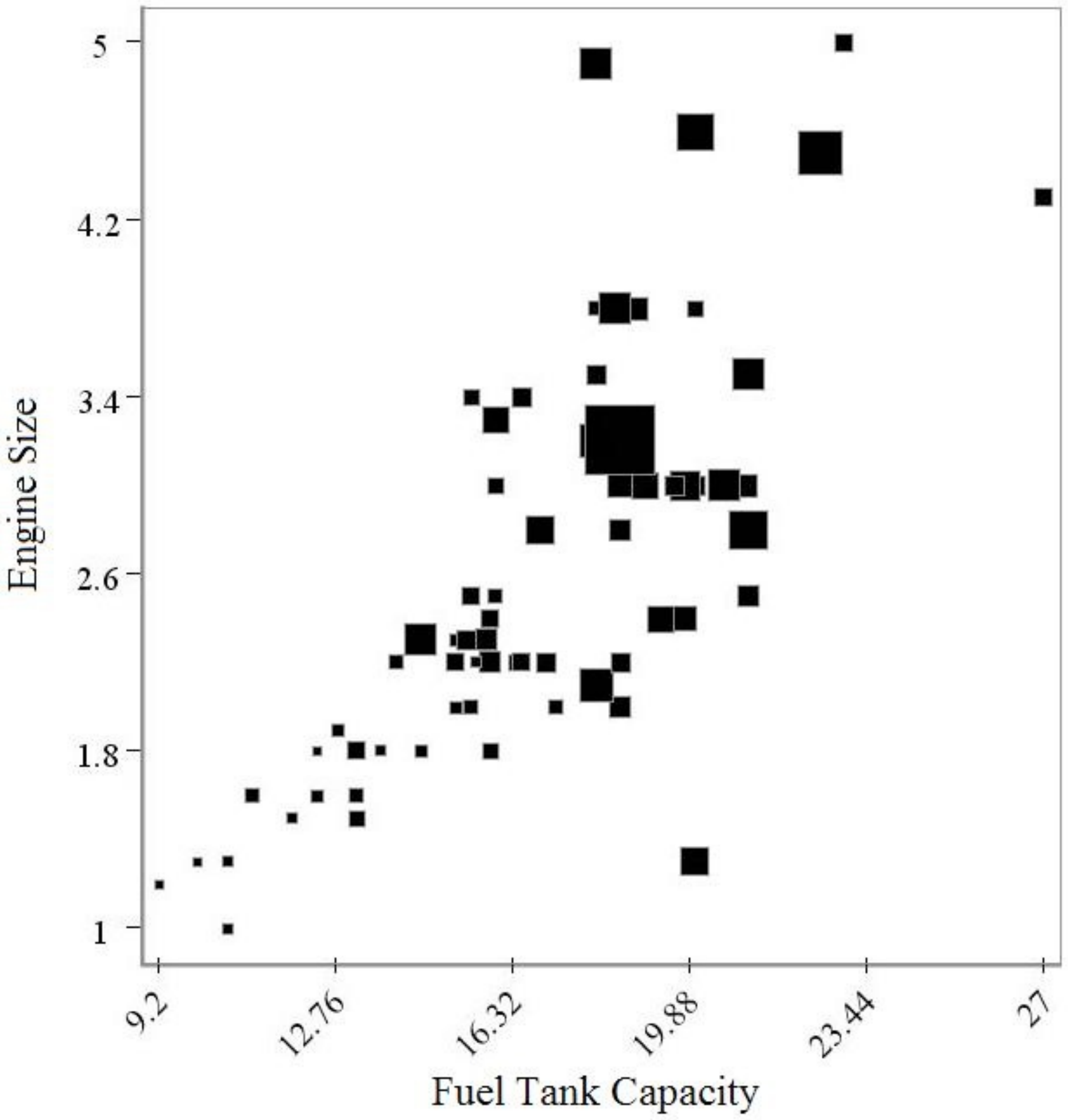
3. Size (Length, Area and Volume)



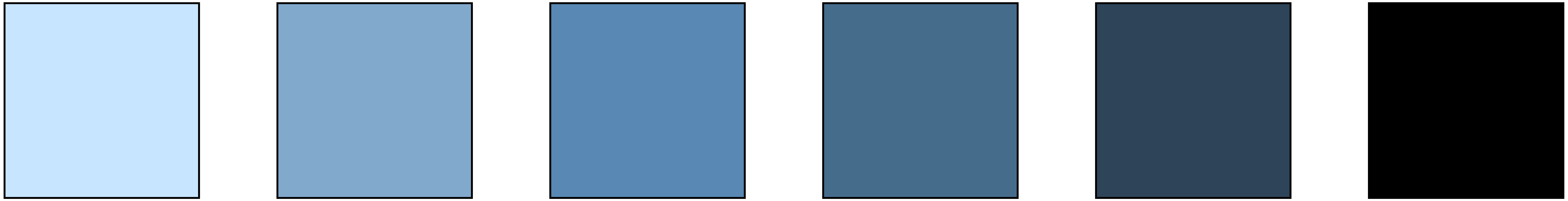
- Varying the size of the graphical element, mark is not mandatory
  - Position and mark are required though
- Additional attributes can be mapped to additional visual variables
- Size supports gradual relative ordering of data variables
- Quantitative separation aspect of size diminishes with sufficiently large area

Visualization of car models showing engine size versus fuel tank capacity.

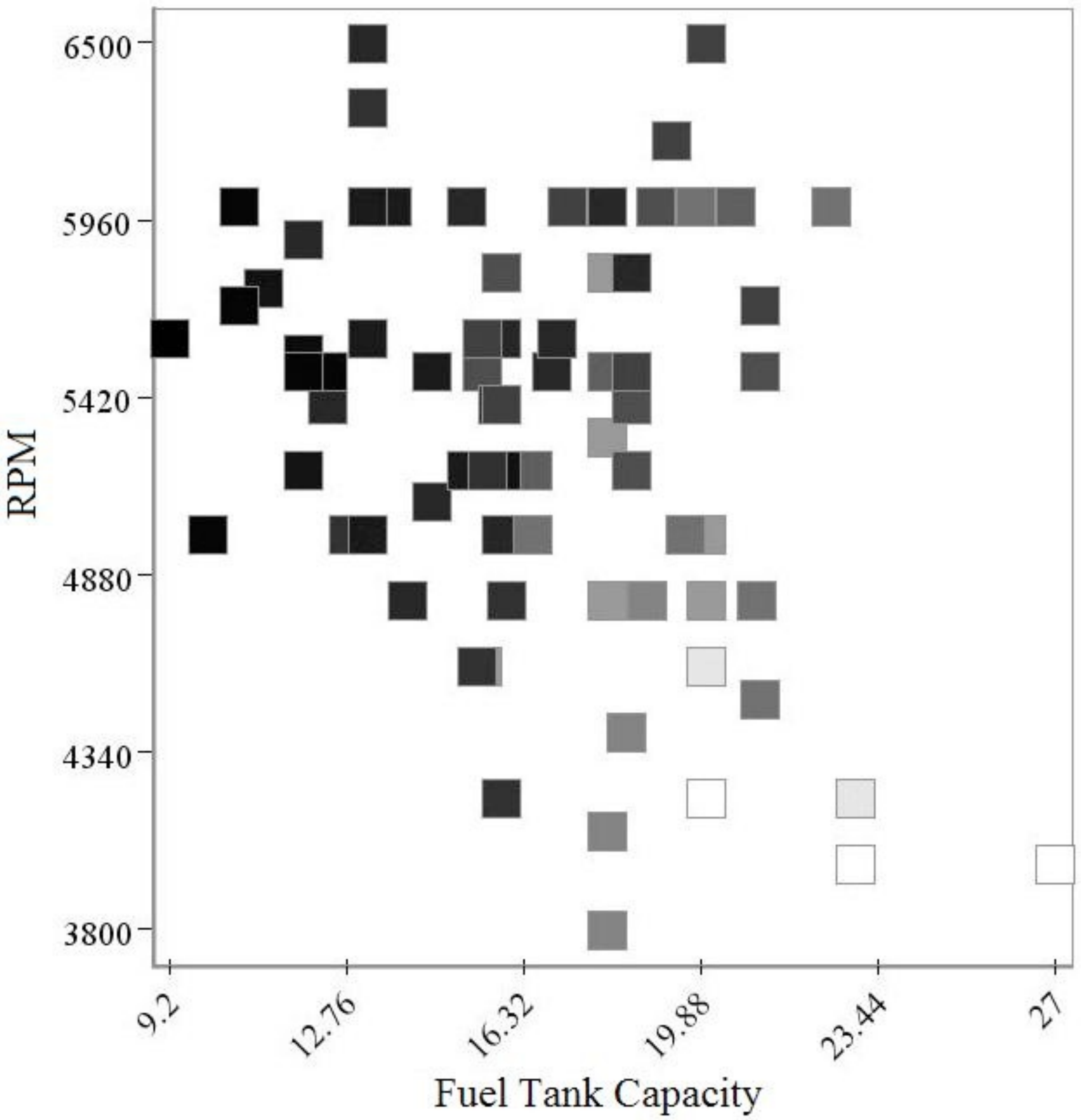
Size is mapped to maximum of price charged.
(larger is more expensive)



4. Brightness or Luminance



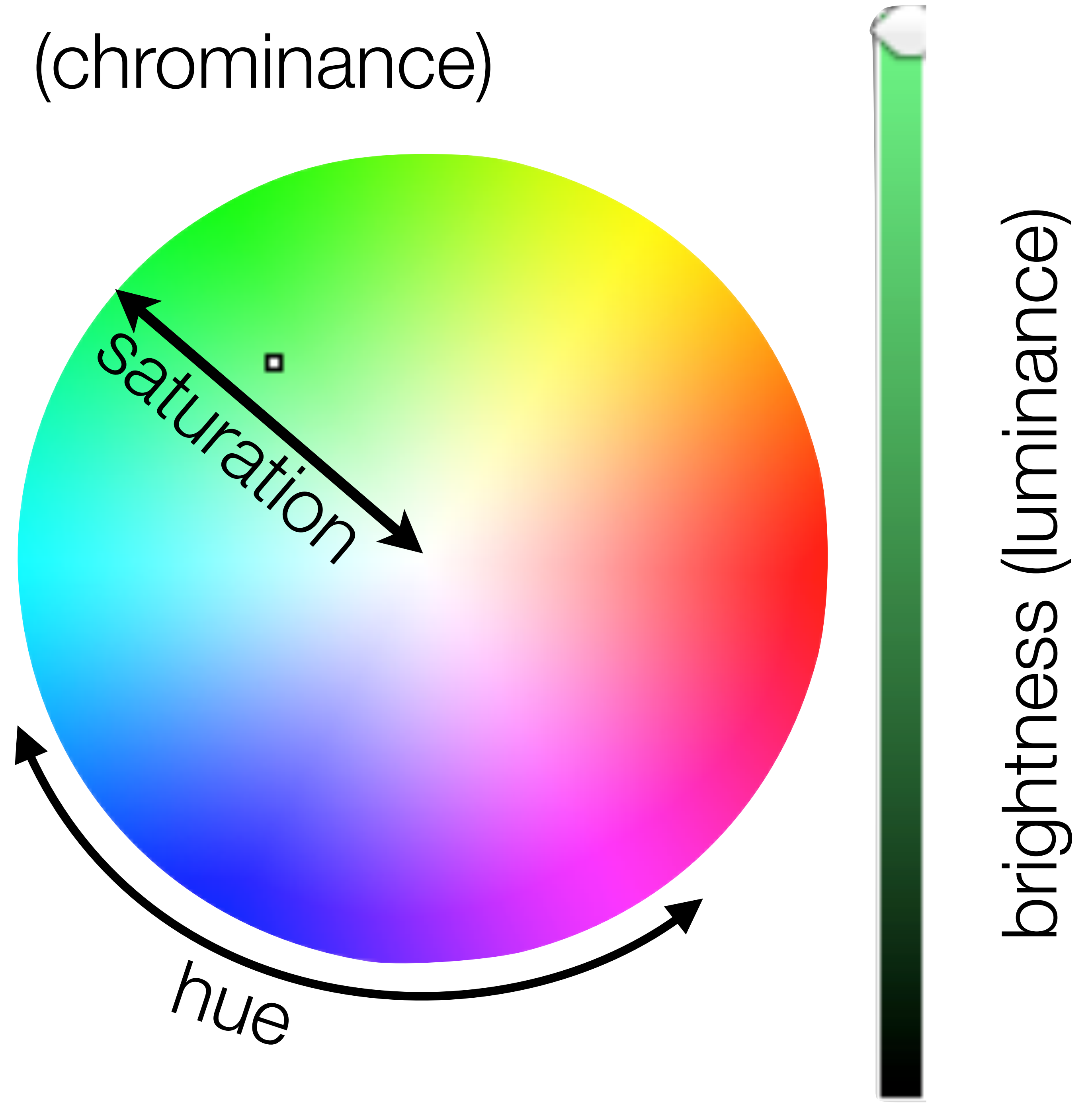
- Modify marks according to additional data variable
- Limited range of perceptually distinguishable brightness settings
  - Less critical for relative, in contrast to absolute differences of large interval and continuous data values
  - Reduced brightness scale can accurately identify data categories
- Perceptually linear brightness scales recommended
  - Use of gamma correction principle needed

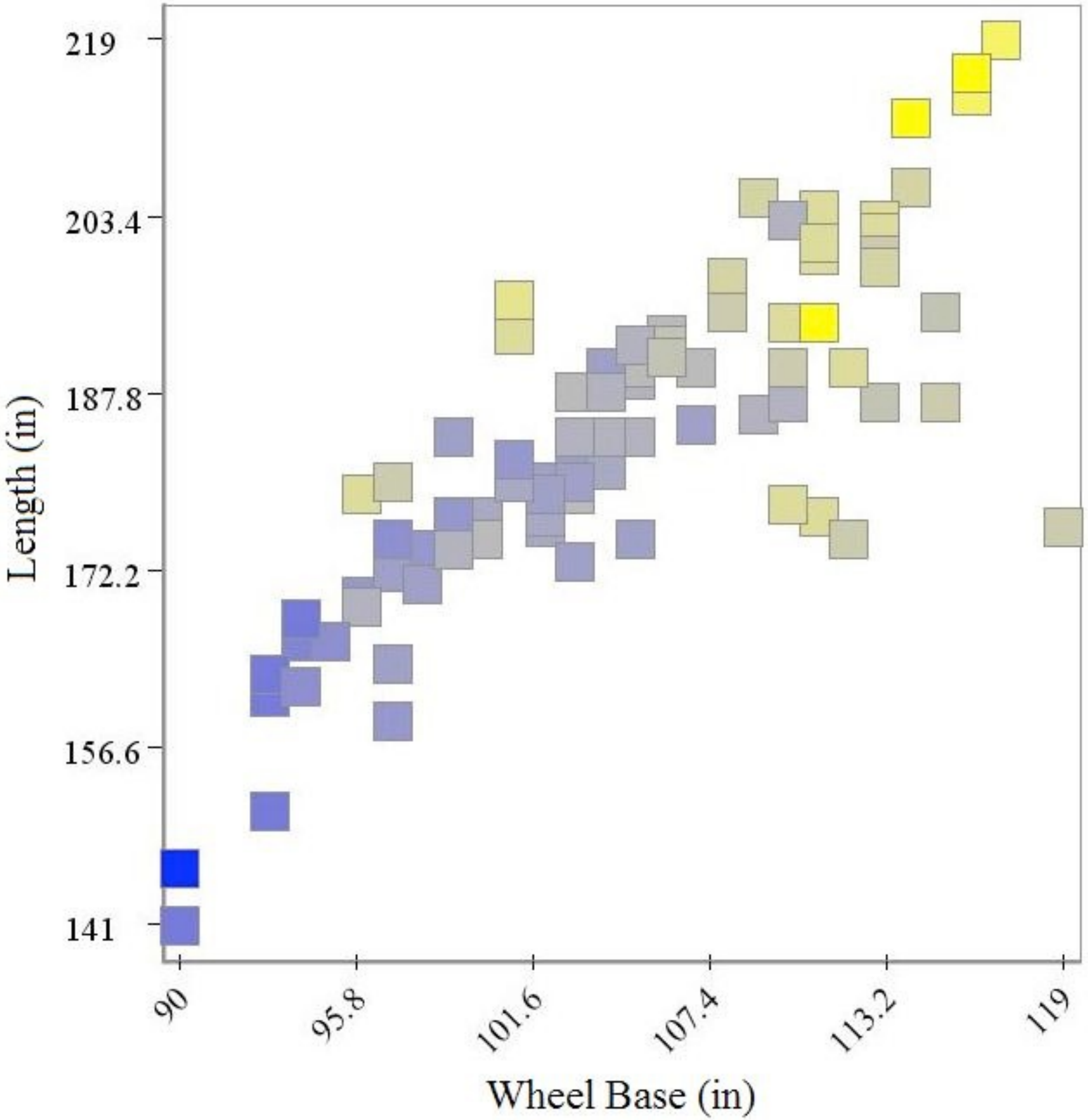


Visualization of the 1993 car models data set, illustrating the use of brightness to convey car width.
(the darker, the narrower the vehicle)

5. Color

- Color can be specified using different color models
  - Factoring out brightness (luminance), color can be specified using two (chrominance) parameters
- Hue defines the color's dominant wavelength in the visible spectrum
- Saturation indicates the intensity of the color
  - Relative to grey (or white for maximal brightness)
- Colormaps can be used to encode continuous/discrete (ordinal) as well as categorical (nominal) data variables



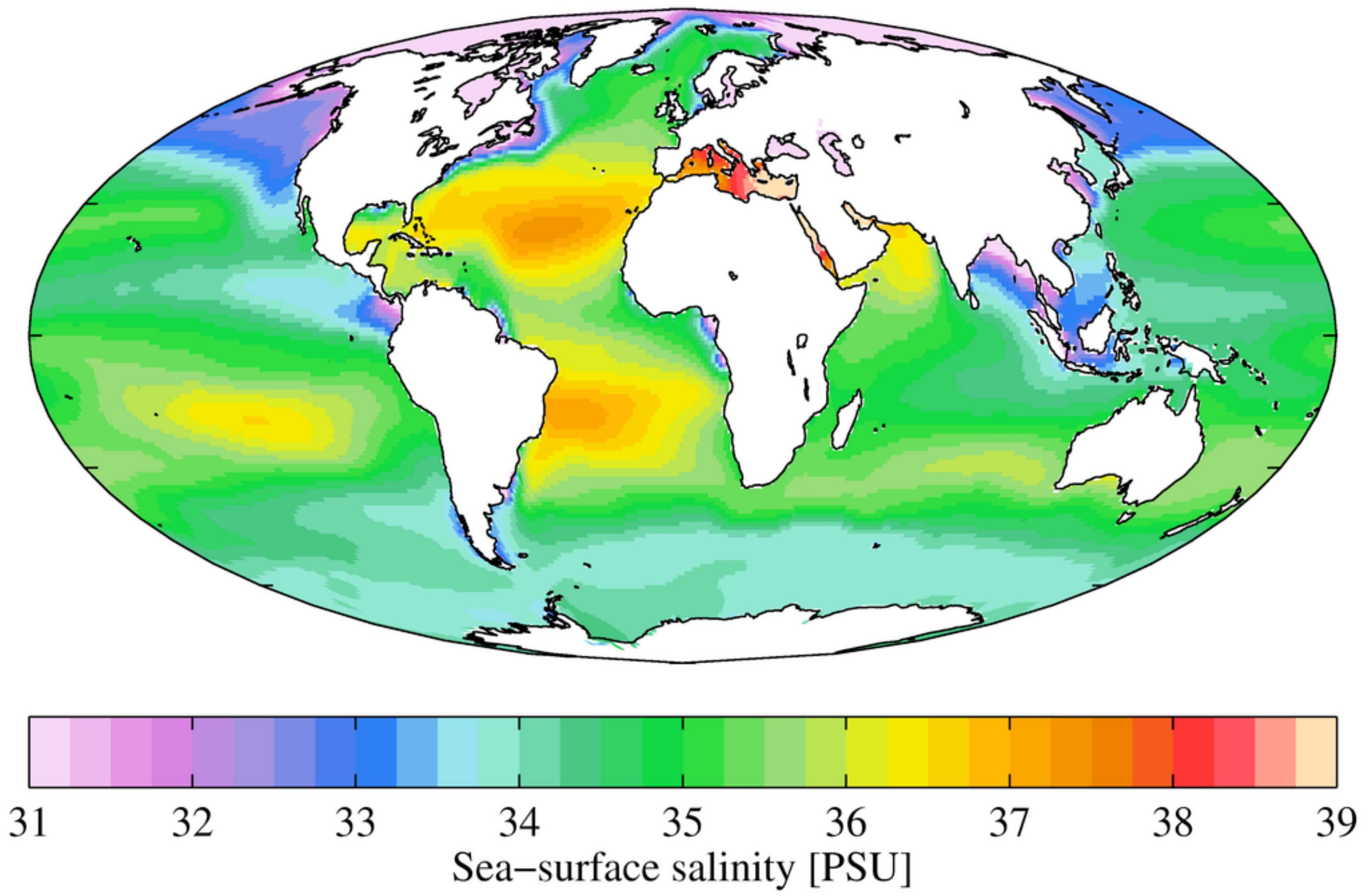


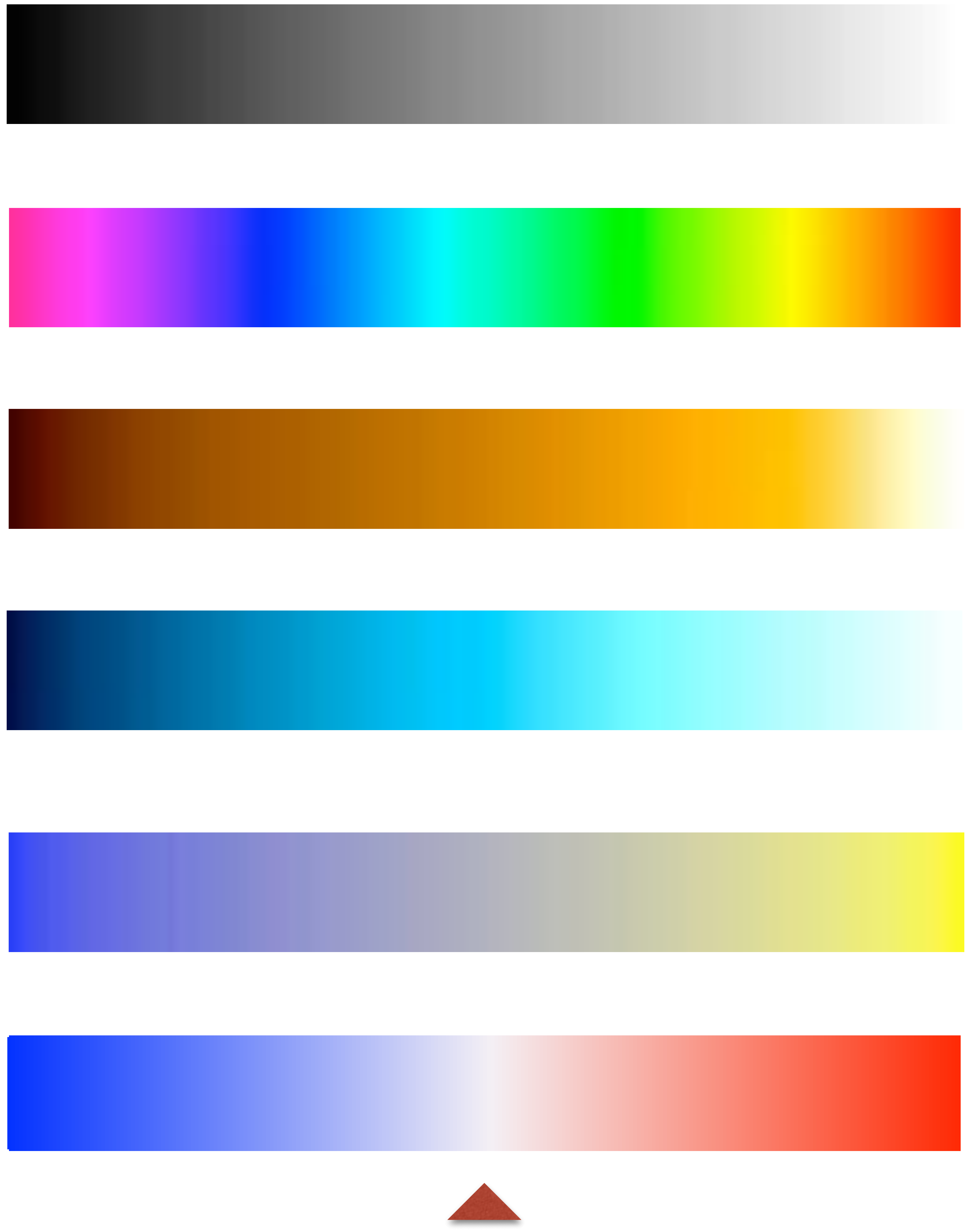
Visualization of the 1993 car models data set, showing the use of color to display an additional attribute.

Length is associated with the y-axis and is plotted against wheelbase on the x-axis.

Common Colormaps

- Use color gradients which are intuitive and distinguishable
- Use discretized color palettes that are perceptually well separated





standard linear gray scale

rainbow

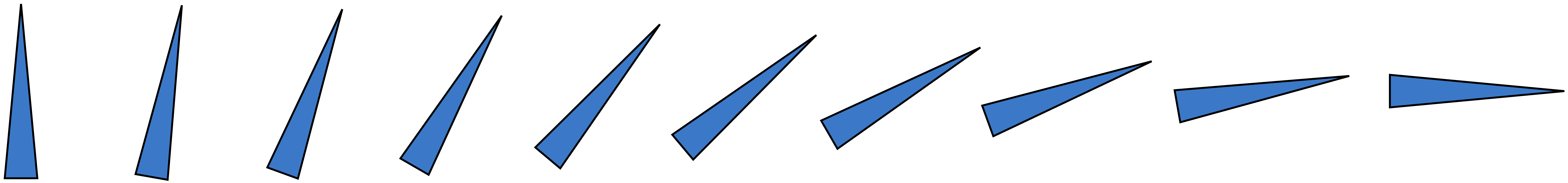
heated

blue to cyan

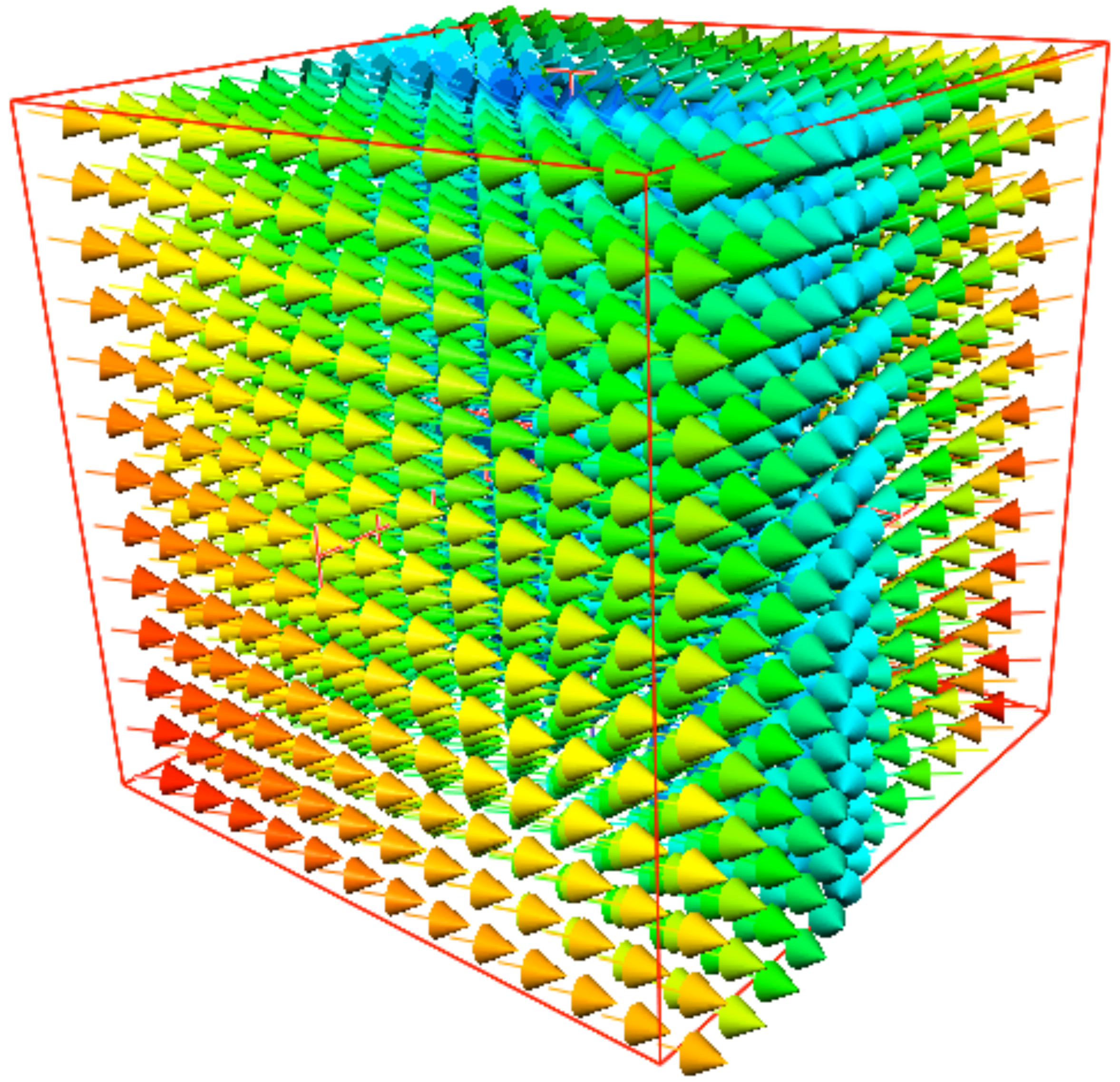
blue to yellow

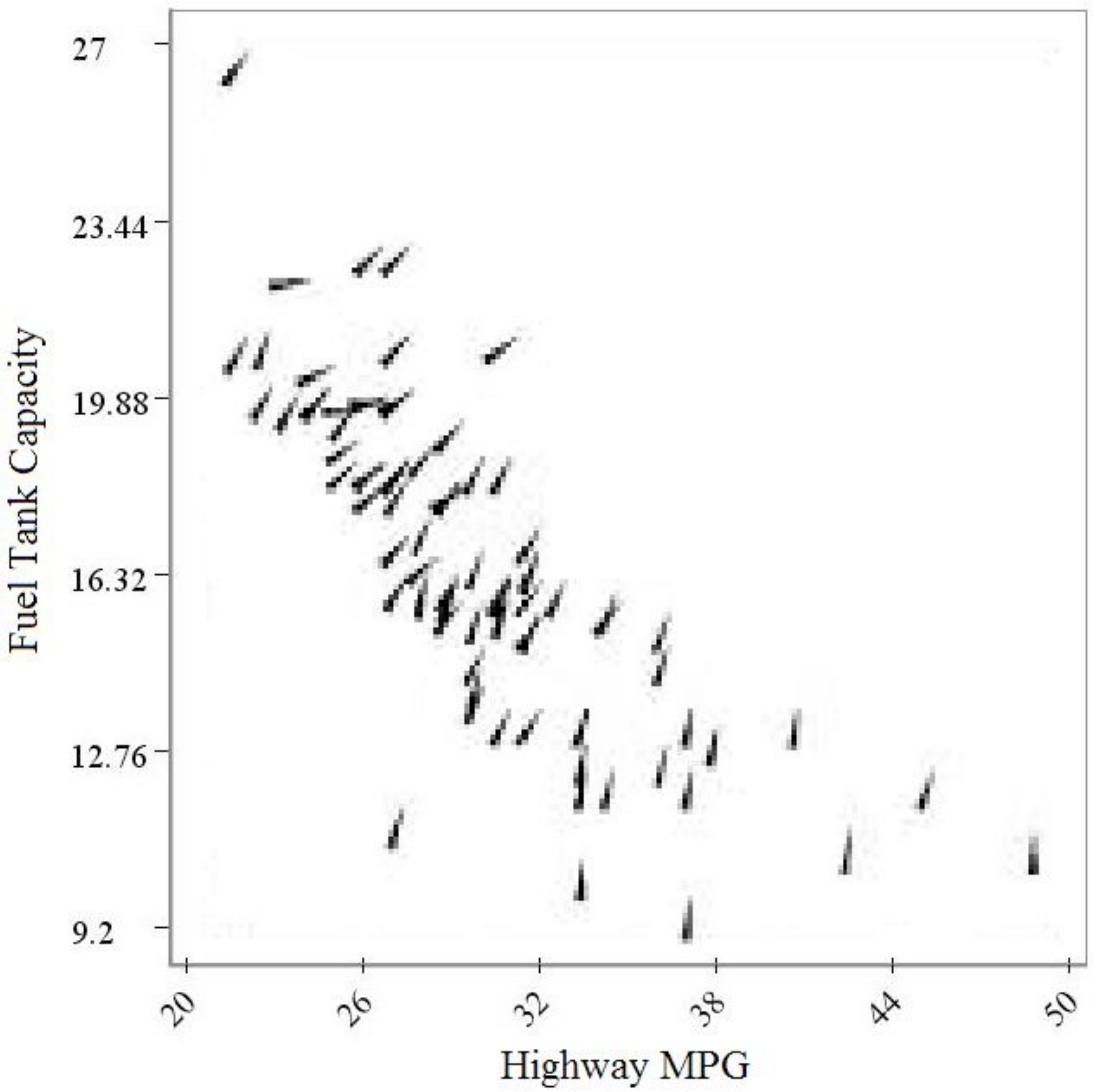
cold-hot/negative-positive

6. Orientation



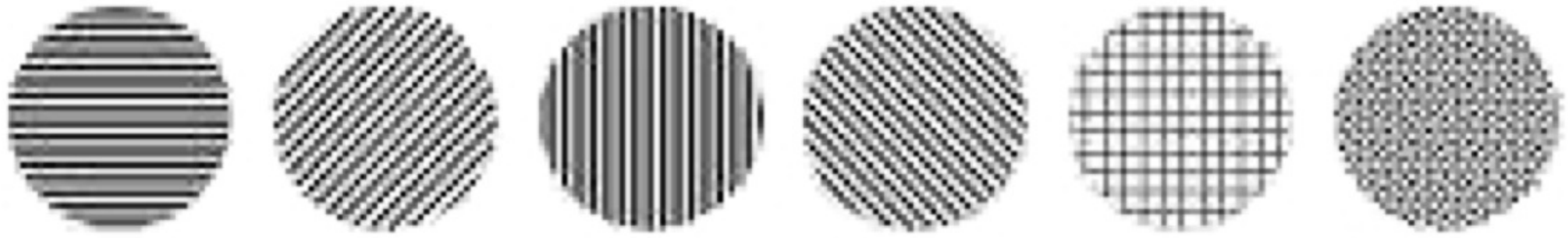
- Principle component of elongated marks with a clear symmetry axis
  - Rotational symmetry of mark shapes, e.g. a for a star, destroys orientability
- Rotational orientation supports relative ordering
  - Especially within a constrained but continuous range or interval
  - Limited absolute differentiability
    - cannot distinguish between 360 rotations





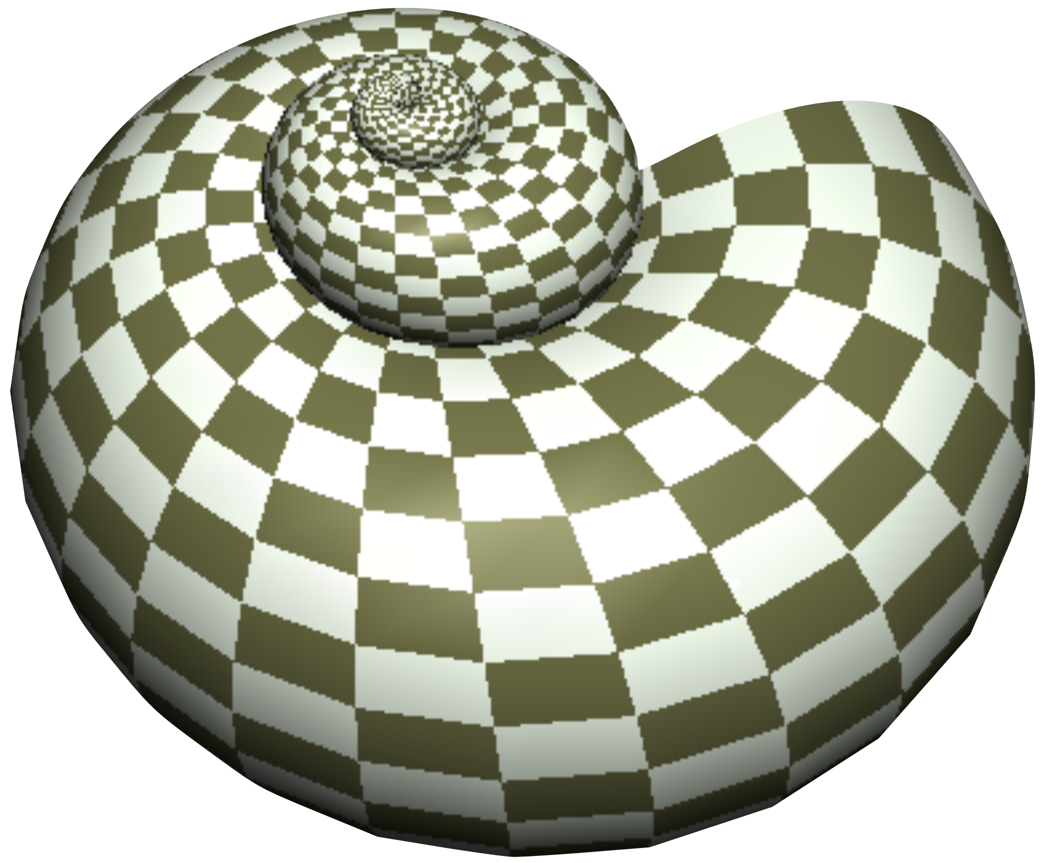
Visualization of the 1993 car models data set, using highway miles-per-gallon versus fuel tank capacity for position and the additional data variable, midrange price, used to adjust mark orientation.

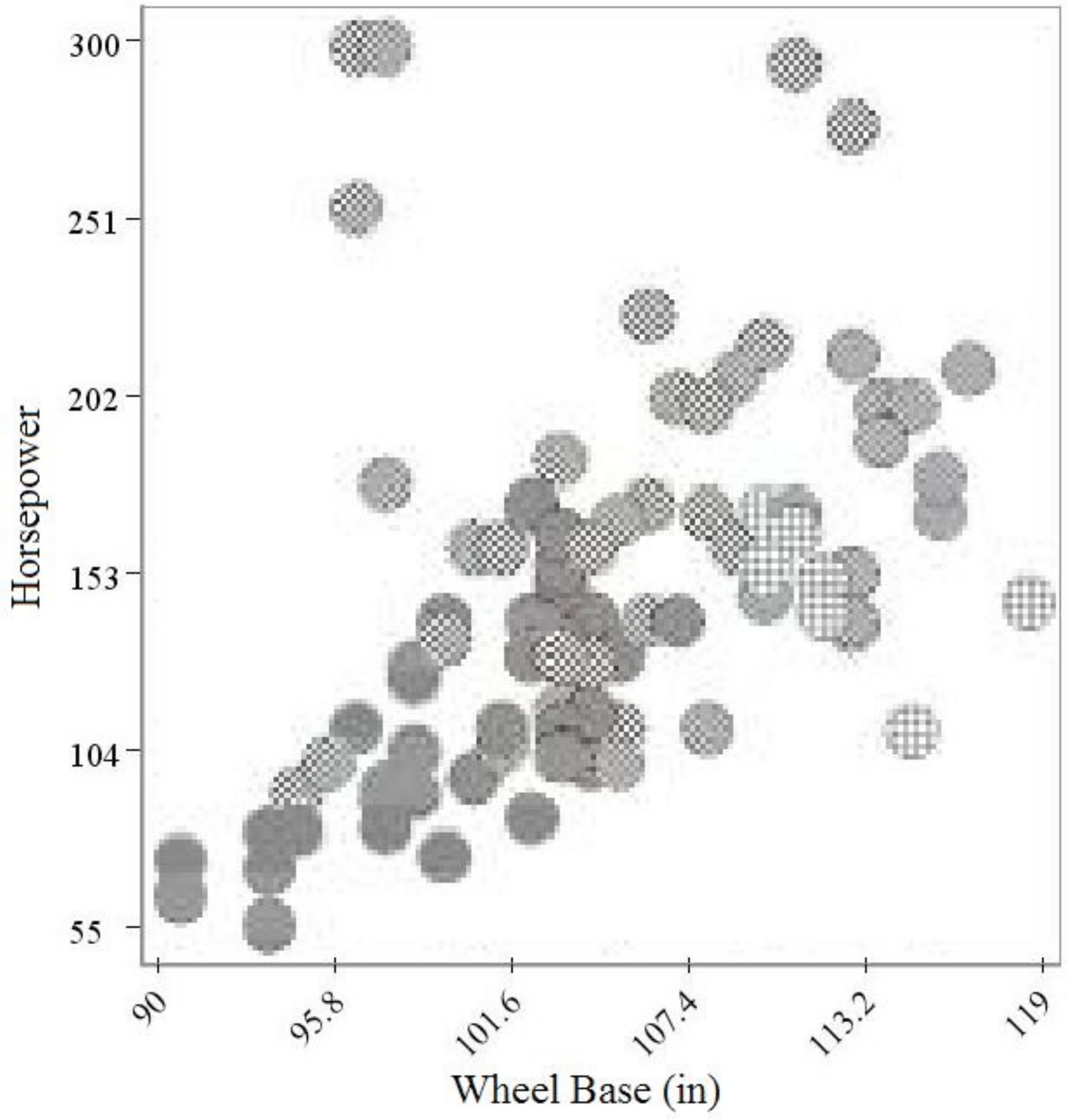
7. Texture



Six possible example textures that could be used to identify different data value categories.

- Similar to the mark as visual variable of a point sample, but also applicable over large spatial extents
  - Can differentiate qualitatively between categories
  - But may also provide relative quantitative information
    - e.g. density or spacing of patterns, color range
- Can also improve perception of 3D surfaces and shapes





Visualization using texture to provide additional information about the 1993 car models data set, showing the relationship between wheelbase versus horsepower for position as related to car types, depicted by different textures.

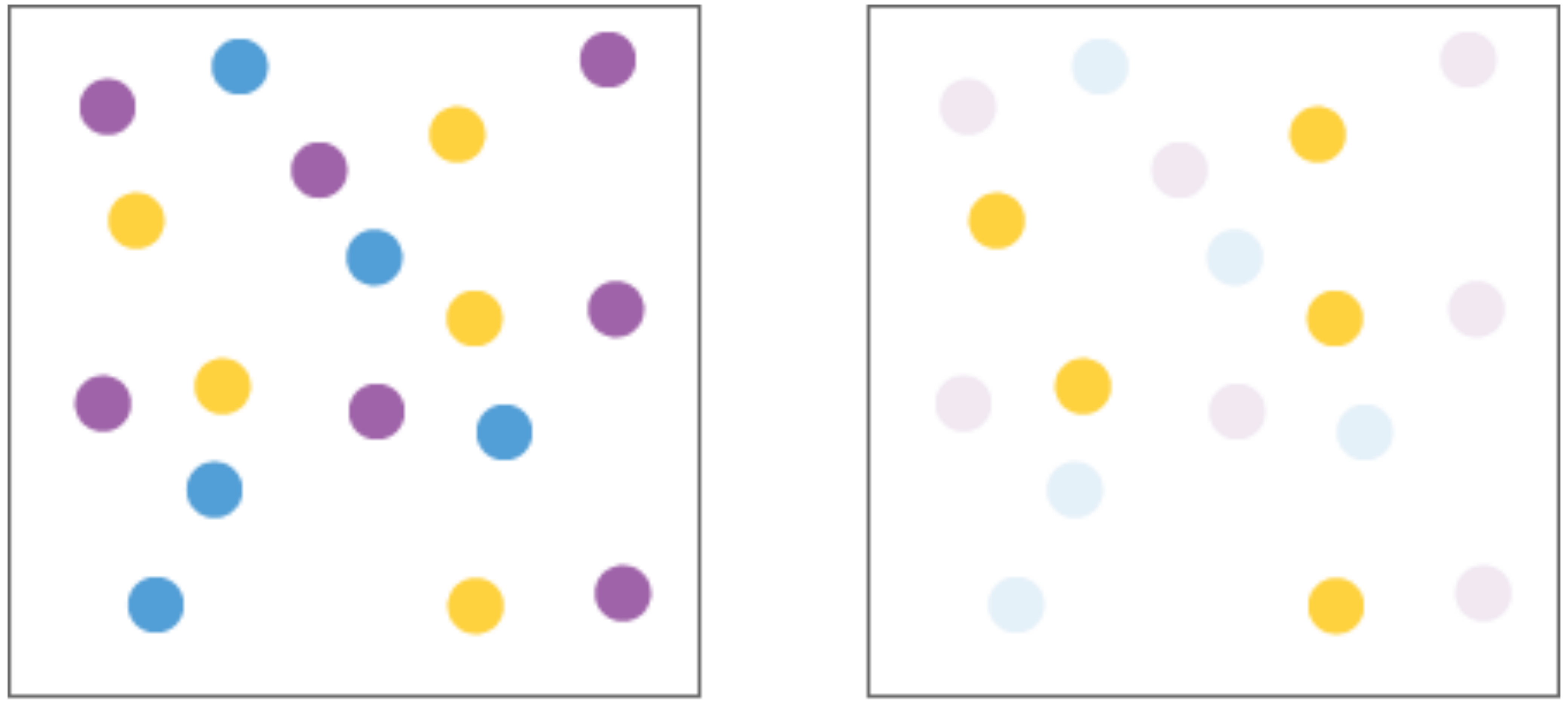
8. Motion

- Motion as change over time can be applied to all visual variables
  - Dynamically animated position
  - Deformation or transformation of mark shape
  - Change of color, size, orientation
- Rate of change can be indicative of a data variable
  - E.g. jittering of position or flashing (change of opacity)

Effects of Visual Variables

The different visual variables can be used for different purposes and perceptual tasks:

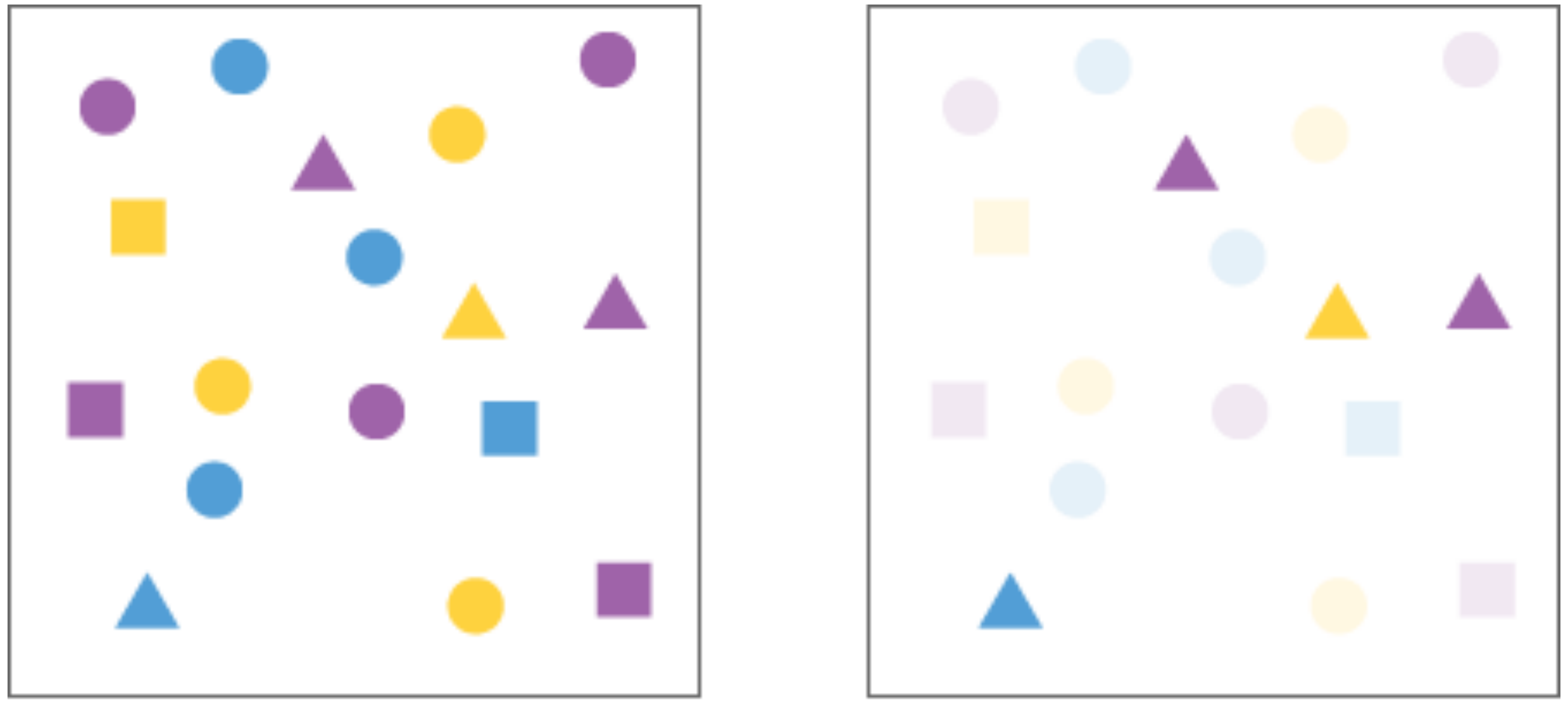
- Selective visual variables
  - Can divide different data values into distinguished groups based on a change in the variable
    - Color
    - Sample/texture
    - Size (length, area, volume)
    - Brightness
    - Direction/orientation



We can quickly and easily perceive a group of symbols based on color hue, e.g., the yellow symbols appear as a group. Therefore hue is *selective*.

Effects of Visual Variables

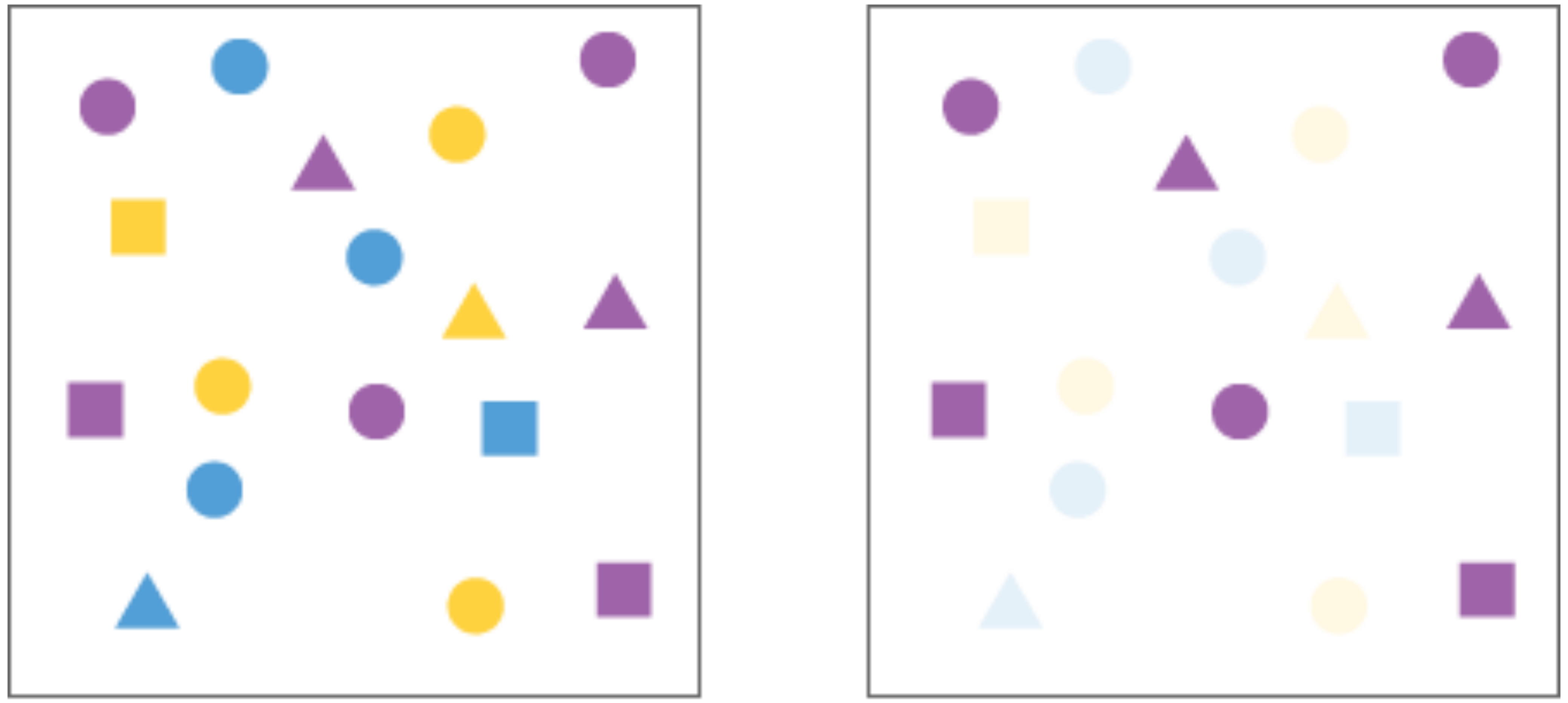
- Selectivity may be affected by the simultaneous use of other visual variables
  - See also interference of preattentive visual features when combining visual variables
    - visual features may mask each other



It is not easy to isolate symbols based on shape. Even with all the triangles pulled out at right, they still don't look like a group. Shape is not selective.

Effects of Visual Variables

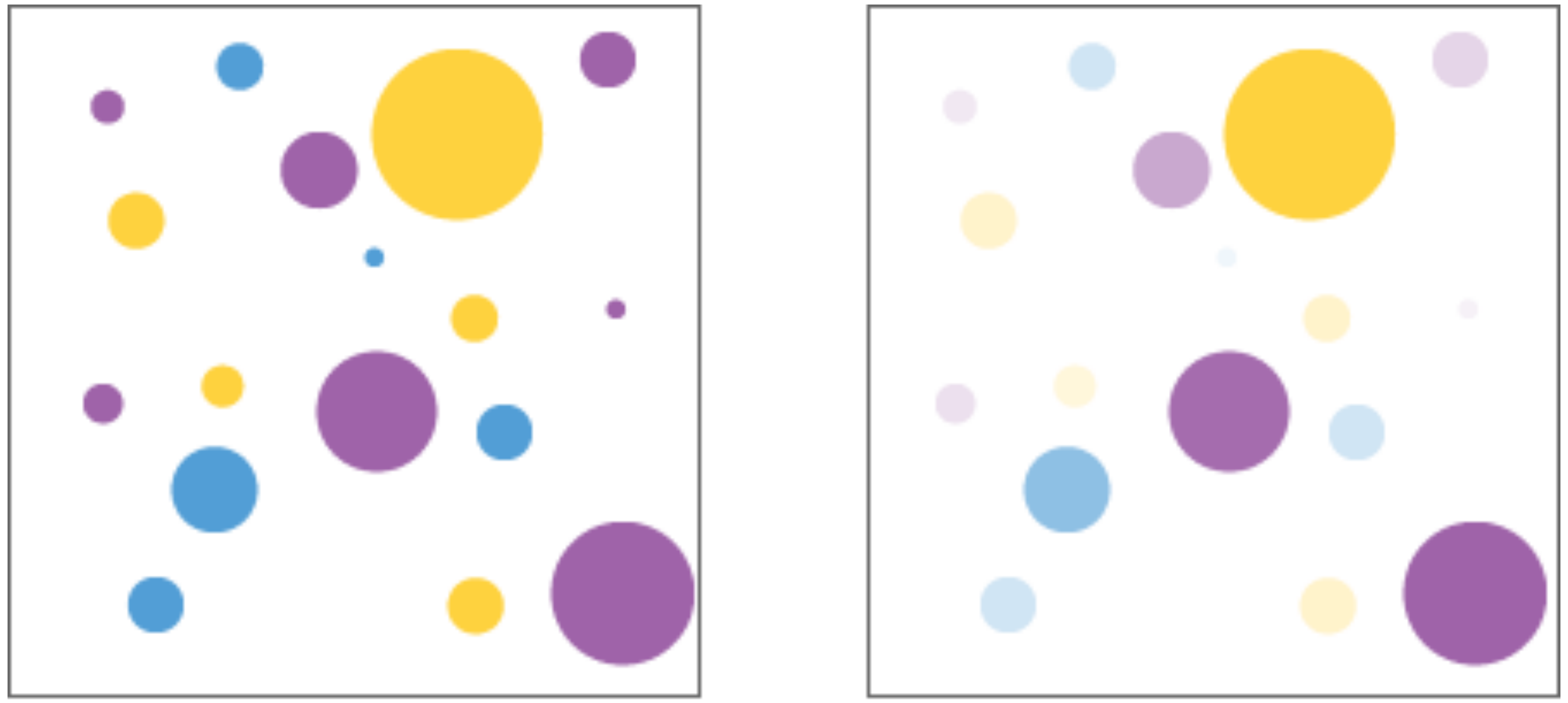
- Associative visual variables
  - Allows grouping across changes in that variable, can perceive symbols as a group despite differences in this variable
    - Shape
    - Color
    - Sample/texture
    - Direction/orientation



Symbols can be perceived as a group despite variation in shape, e.g., these similarly sized purple symbols appear as a group. Shape is therefore *associative*.

Effects of Visual Variables

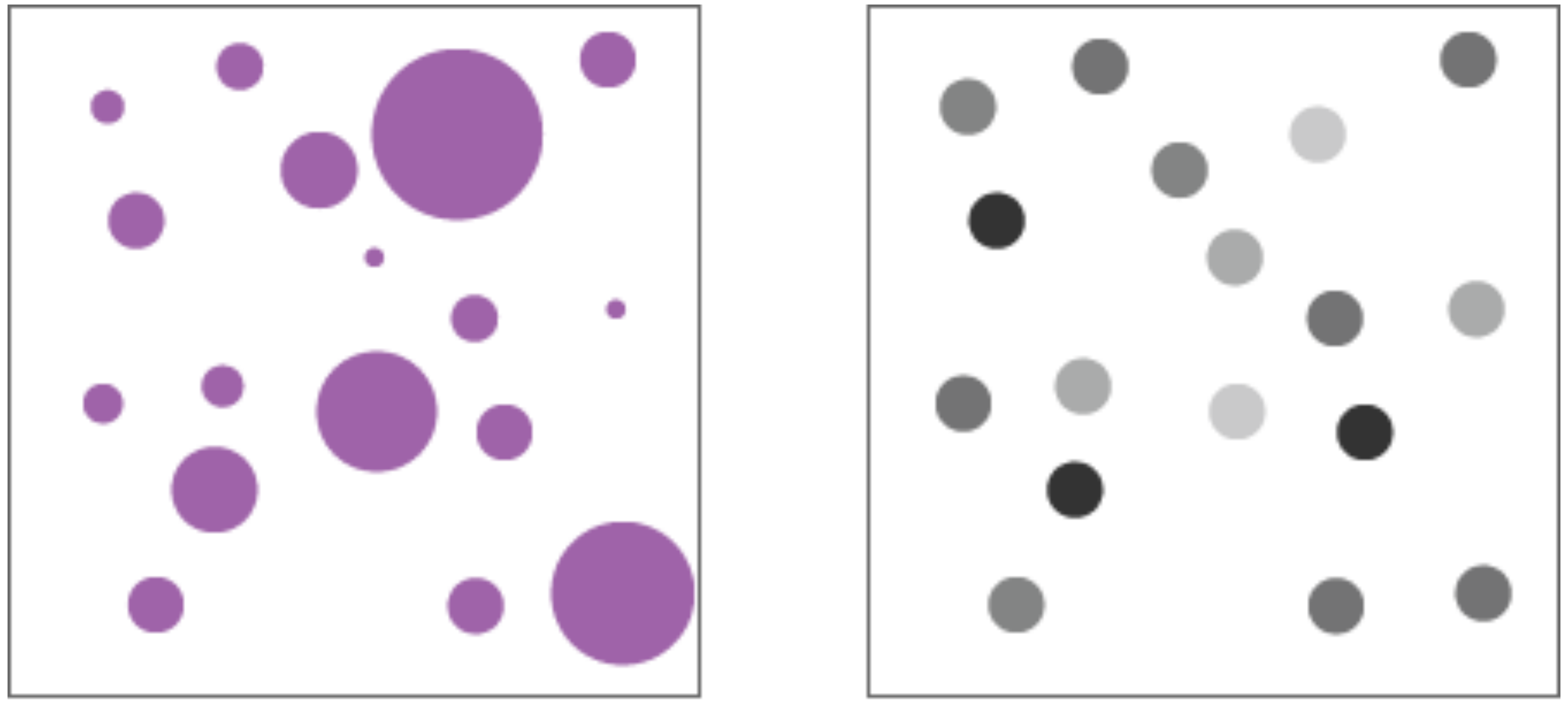
- Associativity can also be connected to the visibility of symbols
  - An associative variable does not cause the visibility of symbols to vary
  - A dissociative does affect the symbols visibility
    - Size (length, area, volume)
    - Brightness



Size variation affects the visibility of symbols and is impossible to ignore. We see larger symbols first, not a single group of circles. Size is *dissociative*.

Effects of Visual Variables

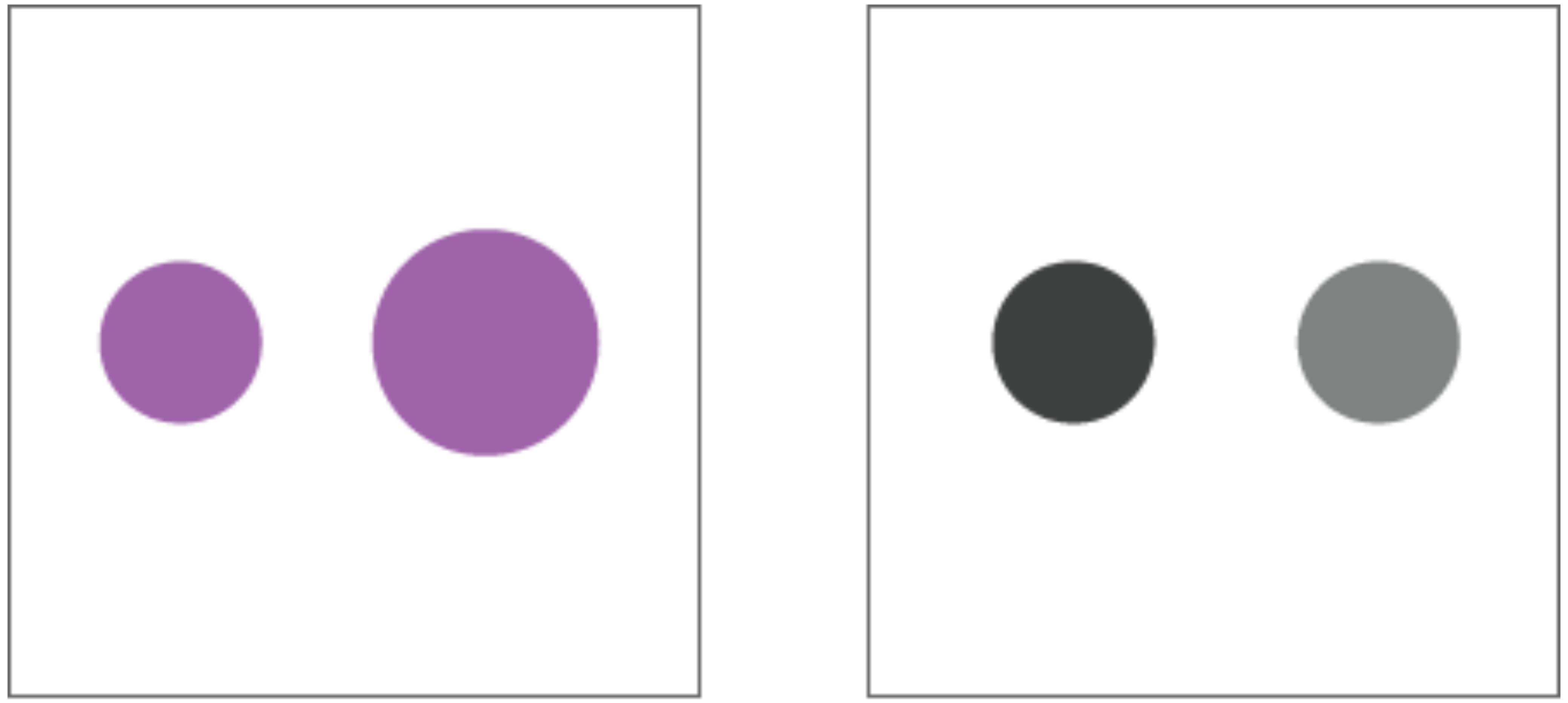
- Ordinal visual variables
  - Support easy ordering of data values (for visualizing ordinal and quantitative data)
    - Size (length, area, volume)
    - Brightness
    - Direction/orientation
    - Sample/texture



With both size and value, it is immediately obvious that there is some sequence to the symbols (small to large, light to dark). Size and value or *ordered*.

Effects of Visual Variables

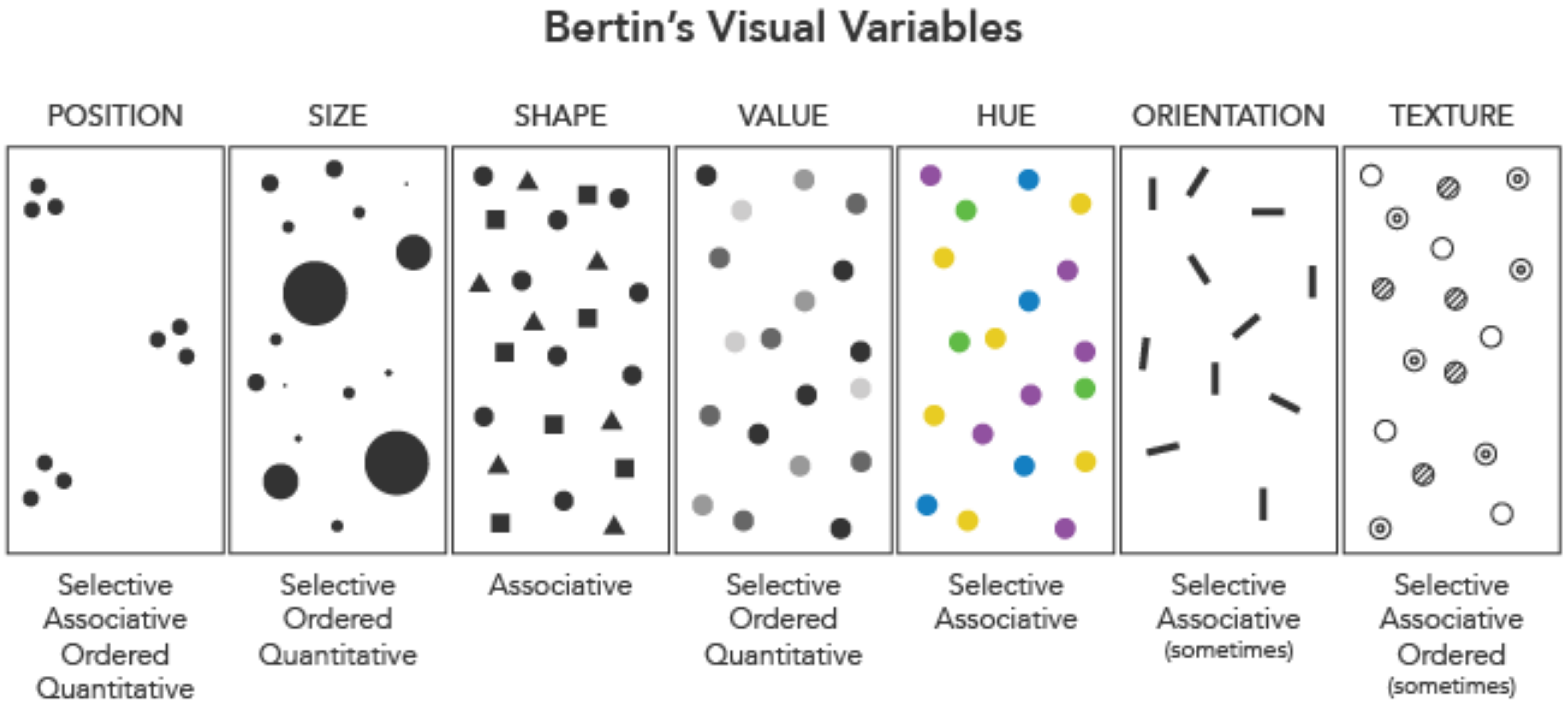
- Proportional visual variables
  - Allow direct association and comparison of relative size (for visualizing ordinal and quantitative data)
    - Size (length, area, volume)
    - Direction/orientation
    - Brightness



We can see that one purple circle is about twice the size of the other, but can't similarly measure a difference in lightness. Size is ***quantitative***.

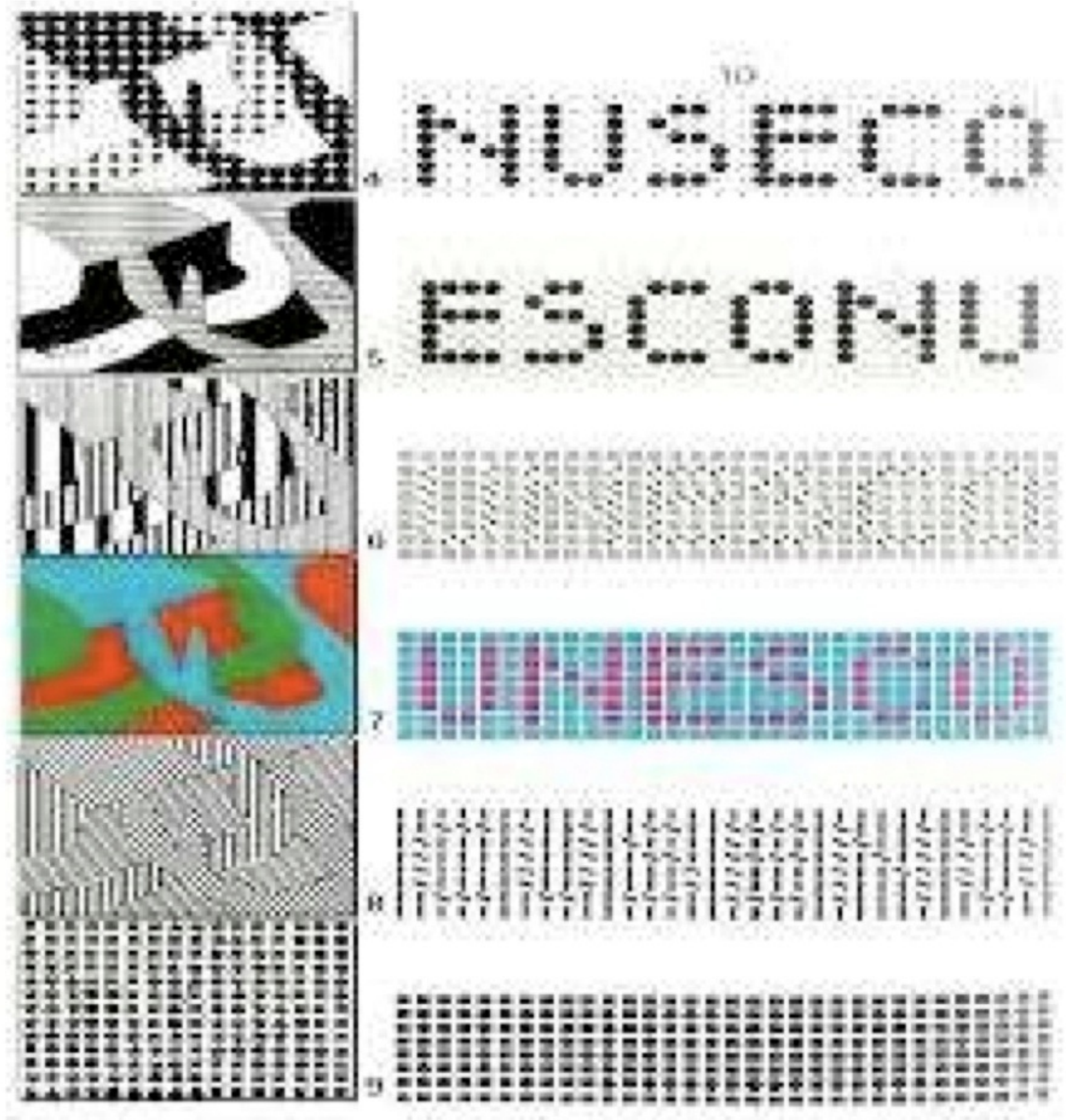
Bertin’s Visual Variables and Effects

- Extensions to Bertin’s original set of visual variables:
    - Motion
    - Saturation
    - Arrangement
    - Crispness
    - Resolution
    - Transparency



Effects of Visual Variables (Extension)

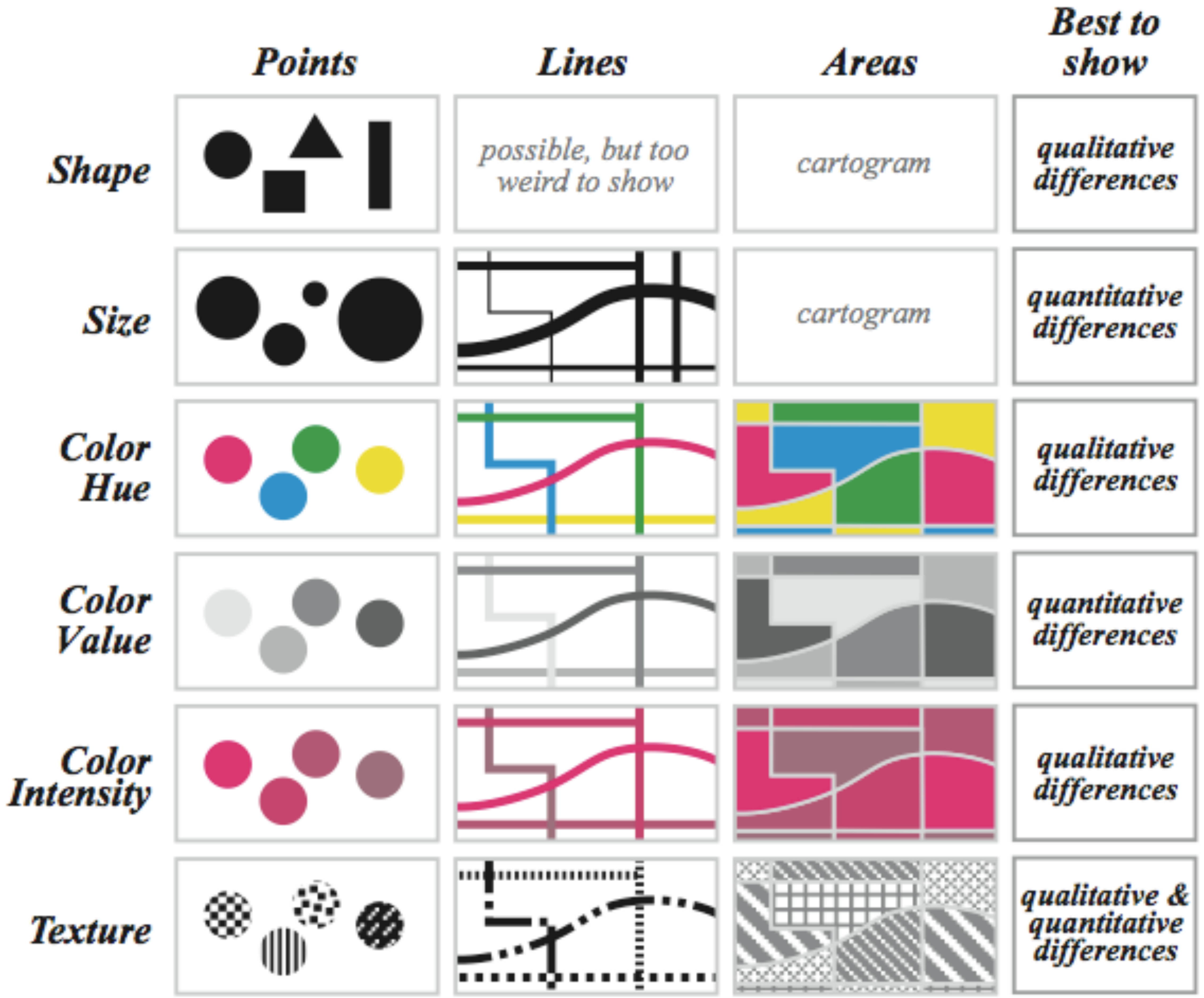
- Separating visual variables
  - All data elements are visible and separable
    - Sample/texture
    - Color
    - Direction/orientation
    - Shape



Separating texture

Visual Variables in Cartography

- Study of visual variables has been traditionally important in cartography
  - Based on Bertin's book *Semiology of Graphics*
  - Summarized in *Making Maps: A Visual Guide to Map Design for GIS* by John Krygier and Denis Wood
  - Krygier and Wood also show how to represent the attribute through points, lines, or areas



© Permission to use the above chart was granted by the designer.
http://understandinggraphics.com/visualizations/information-display-tips/

Recap

- **Semiology:** the study of signs and symbols, use of graphical elements in visualization
- **Visual variables:** the eight visual variables, their parameters, properties and effects

- Required textbook Chapter(s): 4

Related Readings

- **Bertin's Semiology of Graphics** [Diagrams, Networks, Maps. Madison,WI: University of Wisconsin Press, 1983. (trans. W. Berg)]
    - Essential reading for researchers in visualization. Currently out of print.
- **Wilkinson's The Grammar of Graphics** [Statistics and Computing, New York: Springer-Verlag, 2005]
    - Intriguing model for designing visualization techniques and systems. While mostly geared towards statistical graphics, it is applicable to a broader range of applications.
- **Ed Chi's Data State Reference Model for Information Visualization** ["A Taxonomy of Visualization Techniques Using the Data State Reference Model." In Proceedings of the IEEE Symposium on Information Vizualization 2000, p. 69. Washington, DC: IEEE Computer Society, 2000]
    - Seminal work, and several visualization packages have been developed based on it.