Visual Encoding with SCOPeS© Visual Variables/Properties

IT 7113 Data Visualization

http://idi.kennesaw.edu/it7113/

Jack Zheng

Fall 2023

https://www.edocr.com/v/631d1wpb/jgzheng/SCOPeS-Visual-Properties
https://www.edocr.com/user/jgzheng/collection/datavisualizationlecturenotes

Visual Properties: SCOPeS

- Visual variable or property is the "decoration" applied to visual elements to represent data values
- A visual property is used to encode different values of a particular dimension of data
- There are six basic visual properties - easier to be remembered as "SCOPeS" (Dr. Jack's term)

Size

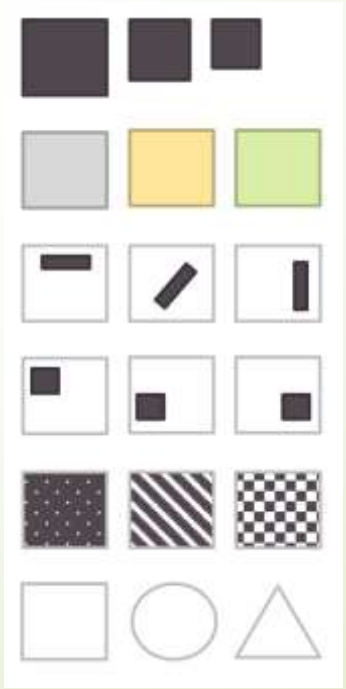
Color

Orientation

Position

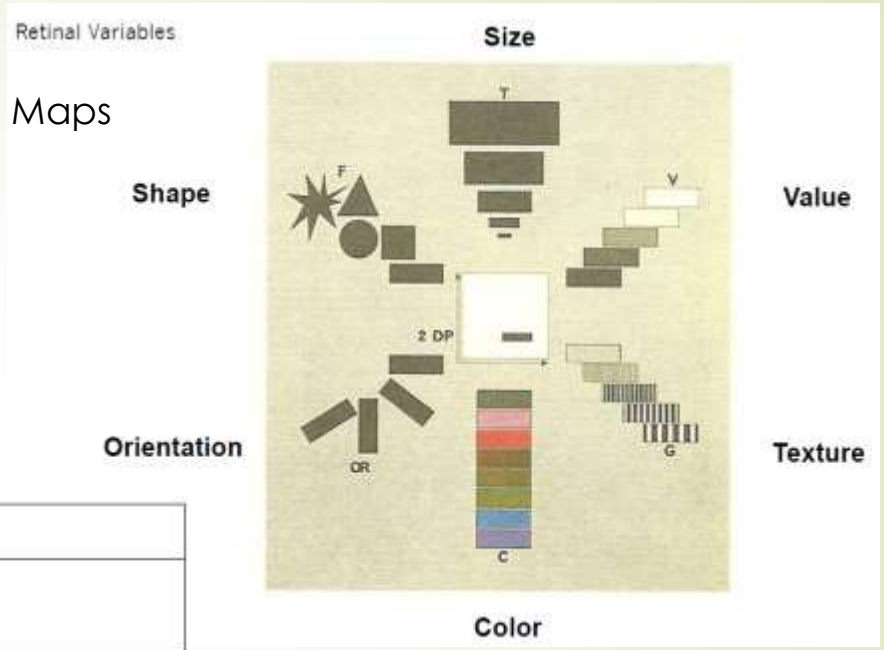
T**e**xture

Shape



Bertin's Original Version

Semiology of Graphics: Diagrams, Networks, Maps
1st Edition, by Jacques Bertin,
https://www.amazon.com/dp/1589482611



https://infovis-wiki.net/wiki/Visual_Variables

Bertin's Original Visual Variables

Position changes in the x, y location	
Size change in length, area or repetition	
Shape infinite number of shapes	
Value changes from light to dark	
Colour changes in hue at a given value	
Orientation changes in alignment	
Texture variation in 'grain'	

Data Characteristics

- Data encoding or mapping is determined both by the kind of visual property and the type of data. The encoding process is the fit of the two.
- The value of a data item is mapped to the value of a particular visual property based on the types of data

Continuous quantitative data	Numerical values. Example: sales amount, age, height, etc. usually they can be aggregated (sum, average, etc.)
Ordinal data	Discrete data, but with an order; often qualitative (for example month in a calendar year) but can be quantitative (ranked or ranged, like age groups/intervals).
Nominal data	The data is a collection of non-numerical and non-ordered data (discrete): categorical. Example: departments in the college

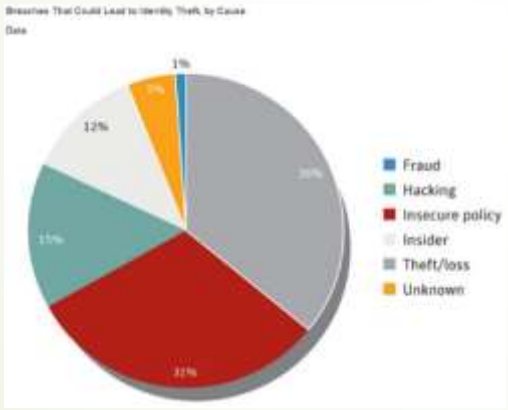
- Some properties can be more effectively represent values of certain data types than others

Expanded readings on data characteristics https://dwbi.org/pages/18

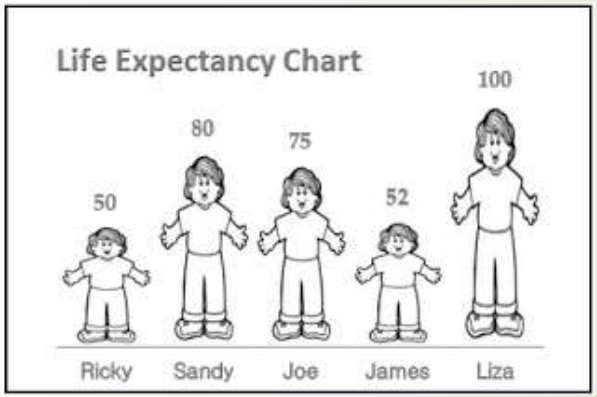
Visual Property: **Size**

- Size is a physical measures of the visual element like length, width, height, area, angle, quantity of items, etc. It is commonly used for continuous data values.
- Examples

Area size represents percentages



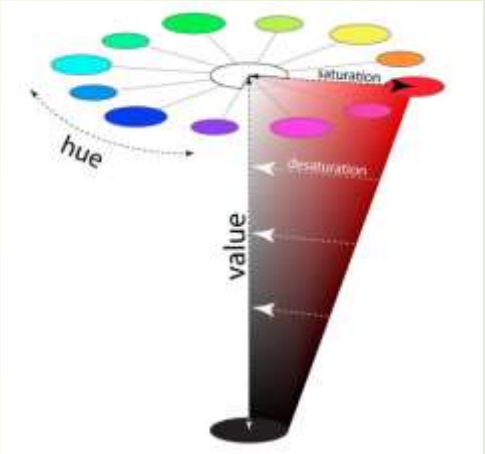
Person symbol height represents number of years



- Scaling issue
  - For various reasons, it is common that the size property does not directly and truly represent the underlying value. In these cases, it must be very careful to design the size property, because unreasonable distortions will impact human perception.

Visual Property: Color

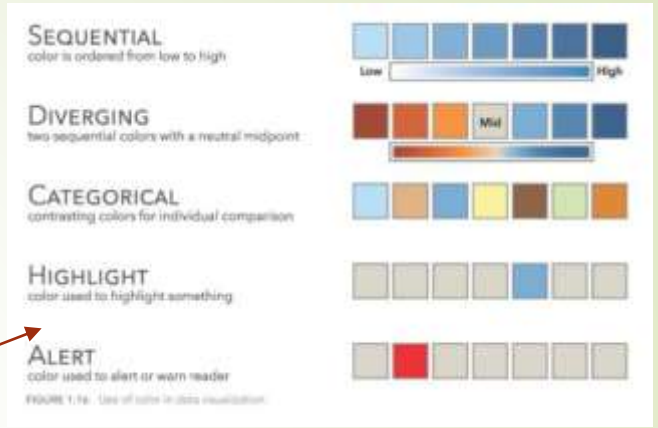
- Color is the most common visual property that can be used for both dimensions (categorical) and measures (continuous or discrete)
- Color has three major properties (HSV)
  - hue (color spectrum)
  - saturation
  - brightness (value)

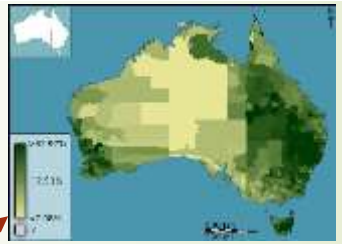


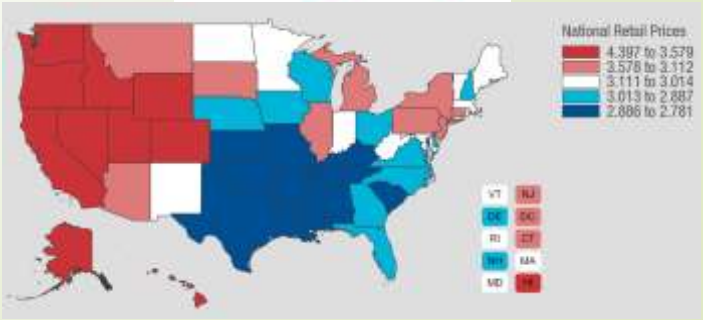
- Extended reading:
  - http://learn.leighcotnoir.com/artspeak/elements-color/hue-value-saturation/
  - Color and color schemes: https://blogs.ifgi.de/digital-cartography/symbols/color-and-color-schemes/
  - https://en.wikipedia.org/wiki/HSL_and_HSV

Color coding based on data characteristics

- Two basic ways of using colors: categorical and continuous
- Categorical data (dimensions or measures) – using hues
  - Represent and differentiate individual data item - http://en.wikipedia.org/wiki/Pie_chart (colors in pie chart commonly represent countries, which is nominal or categorical) – however, color is ineffective if there are more than 7 items or categories.
  - Group a set of similar items or data points (dimensions or measures)
  - Highlight (dimensions or measures): alert, distinguish, etc.
- Continuous data (measures) – using values or saturations
  - Sequential system: using one color's value range
  - Diverging system: two color gradient (or step) system, often used in heat maps where two colors represent opposite directions. (Example: https://www.theinformationlab.co.uk/2014/12/03/geospatial-heat-maps-tableau-via-alteryx/)
  - Gradient or stepped:
    - gradient is a smooth transition, often used for continuous measure
    - Stepped usually categorize measures in ranges and uses more discrete color values, https://covid.cdc.gov/covid-data-tracker/#global-counts-rates; https://gasprices.aaa.com/
- More references
  - https://spectrum.adobe.com/page/color-for-data-visualization/







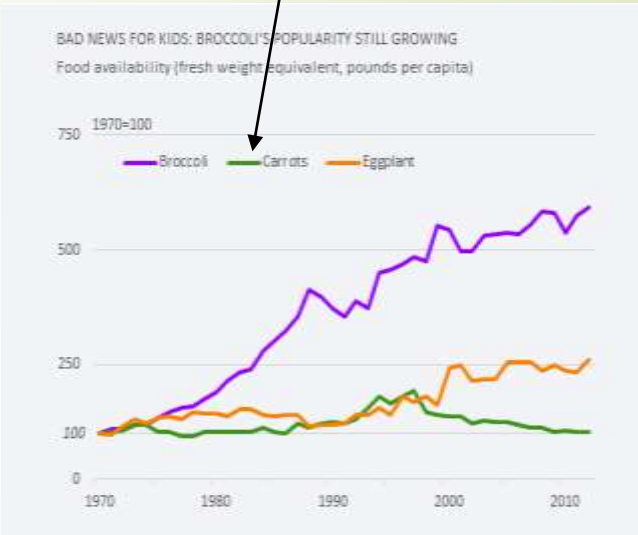
Color Systems, Cases and Best Practices

- Traffic light system in performance management
  - Uses green, yellow, and red to represent different levels of performances (good, warning, bad); three colors corresponds to target values; it is a hue-based system used ranged measures.
  - https://citoolkit.com/articles/traffic-light-assessment/
- Temperature system in maps (heatmap) – a bipolar hot/cold system: red for hotter temperature and blue for colder
  - Heat maps (https://en.wikipedia.org/wiki/Heat_map) use color to represent target values, either continuous or in groups/ranges (ordinal)
  - Originally used for data related to temperature and weather https://weatherboy.com/north-america-chills/
  - Also for data that has a temperature metaphor, such as click map (ttps://www.ometrics.com/heat-maps/), population density (https://andyarthur.org/thematic-map-county-population-density.html), lighting, salary, etc.
- More best practices and examples of using colors in data visualization:
  - https://cambridge-intelligence.com/choosing-colors-for-your-data-visualization/
  - https://www.dataquest.io/blog/what-to-consider-when-choosing-colors-for-data-visualization/
  - https://colorbrewer2.org (a good color choice tool)

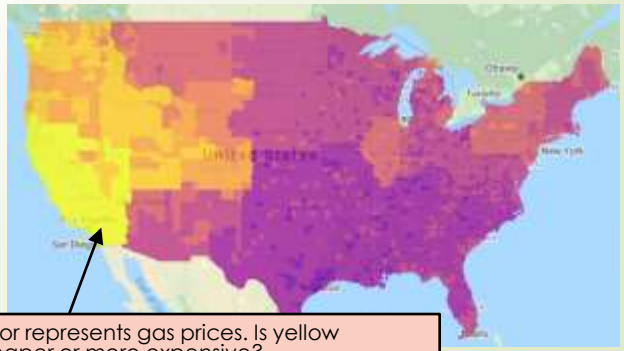
Other considerations when selecting color

- Color symbolism
  - Match the color of visual element to the real-world counterpart (things or phenomenon) it represents.
  - CI color: use the color consistent with the company image to represent the company or organization.
- Reasonably consider aesthetics that contributes to human's emotional perception
  - temperature - warm vs. cool
  - harmony - sharp vs. soft
  - Use beautiful colors: https://blog.datawrapper.de/beautifulcolors/
- Color blindness
  - Prepare for a second alternative, usually using textual instead of color.

Mismatch of natural color of the vegetable with the line color that represents the vegetable.



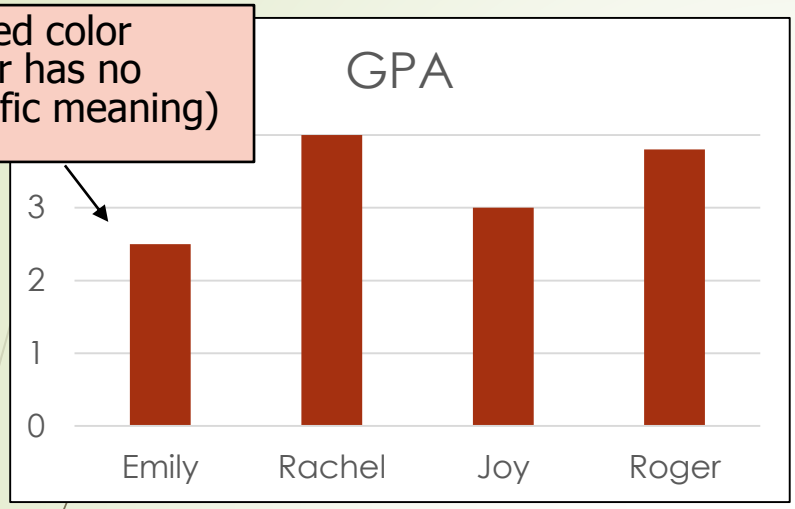
http://www.dataatworkbook.com/data-work-03-beyond-visual-perception/



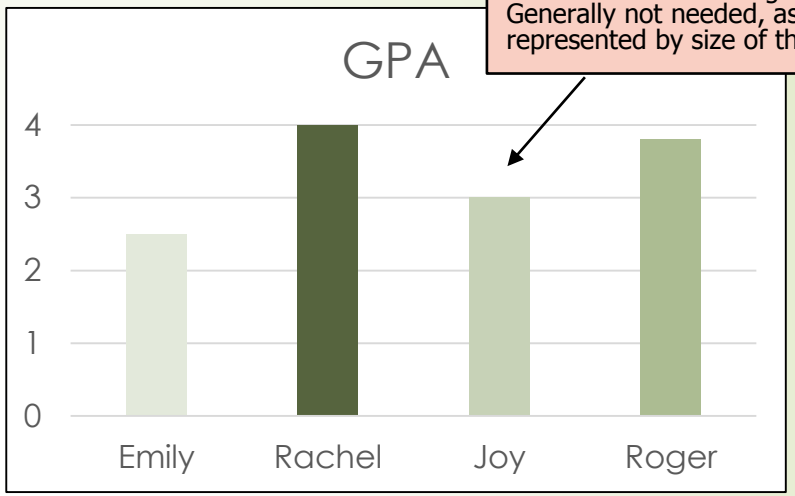
Color represents gas prices. Is yellow cheaper or more expensive? https://www.gasbuddy.com/GasPriceMap

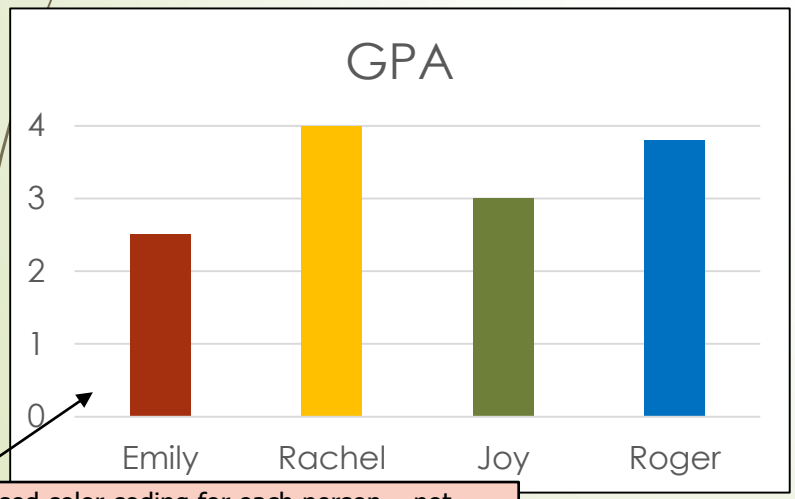
Which color system is better in the following charts? Any other better designs?

Unified color (color has no specific meaning)

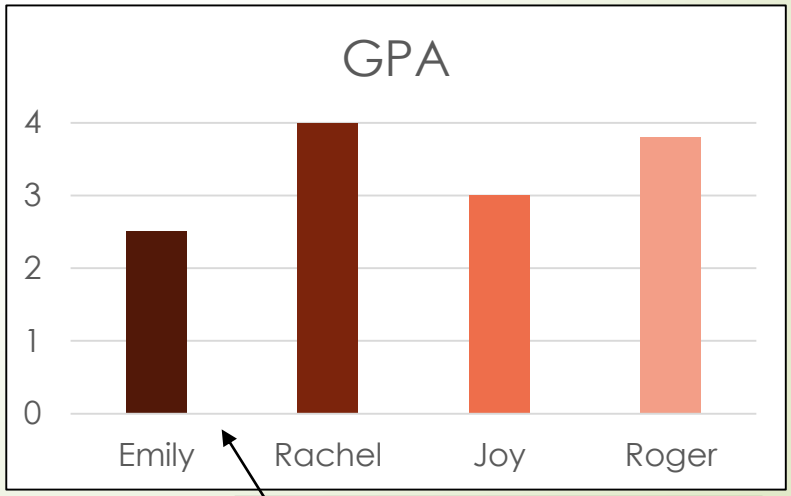


Can we use value-based color coding to represent GPA values? I.e. darker color for higher GPA? Generally not needed, as GPA is represented by size of the bars.





Hue based color coding for each person – not necessary, as each person is represented by position. Color may be used to represent a second dimension, like school or class.



Value based color coding for each person – not recommended same reason as the one on the left.

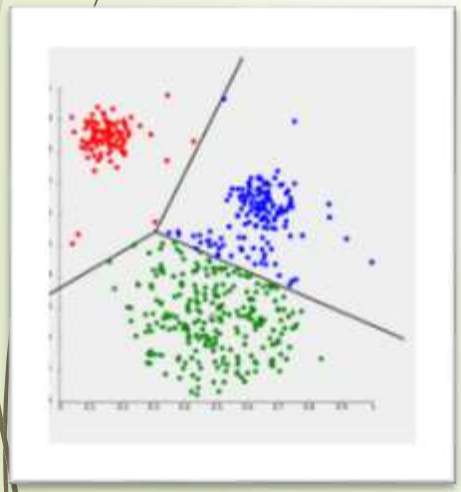
Visual Property: **Orientation**

- Orientation is about direction. It can be seen as variations of a particular shape or pattern pointing to different directions. It can also be seen as a variation of position (pointing to different positions)
- A common example is the use of arrows or hands pointing to different directions.
- Examples
  - Gauge chart

FIFA World Cup			Dow Jones			KLCI		
Tournament	Time	Host	May	Jun	Jul	May	Jun	Jul
1990	8 Jun - 8 Jul	Italy	→	→	↘	↗	→	↗
1994	17 Jun - 17 Jul	USA	↗	↗	↘	→	→	↗
1998	10 Jun - 12 Jul	France	↗	↘	↘	↘	↘	↘
2002	31 May - 30 Jun	Korea/Japan	↘	↘	↘	↘	↘	→
2006	9 Jun - 9 Jul	Germany	→	→	↗	↘	→	↗
2010	11 Jun - 11 Jul	South Africa	↘	↗	↘	↘	↗	↗
2014	12 Jun - 13 Jul	Brazil	?	?	?	?	?	?

http://voyager8.blogspot.com/2014/01/the-historical-relationship-between.html

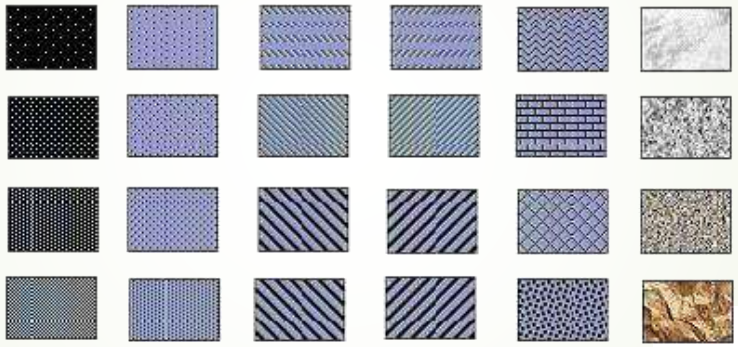
Visual Property: **Position**

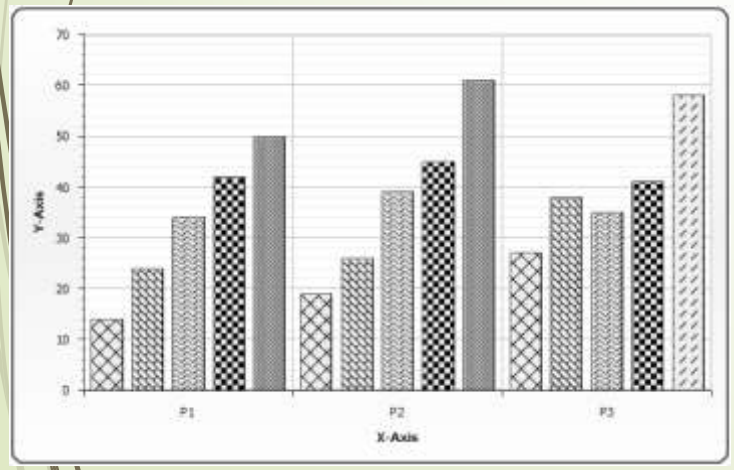


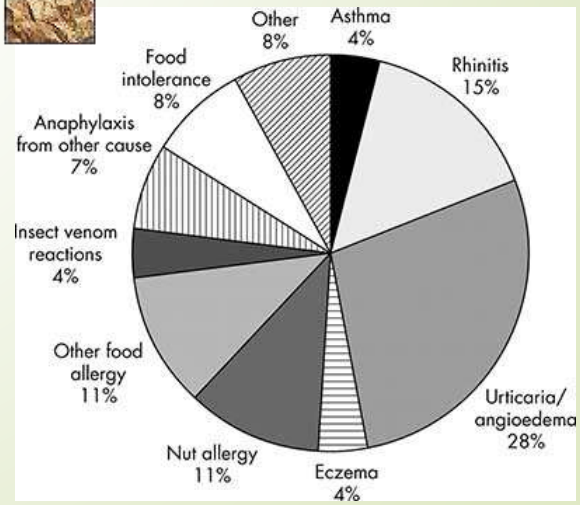
- Position refers to the location where a data item is placed. Data values can be visualized as absolute positions in the visualization, or as the relative distance between elements.
- Position is commonly used in
  - Dot plot – a variation of column chart
  - Strip plot (one-dimension scatter plot)
  - The placement of data items against a pre-established scheme (such as a Cartesian coordinate system)
  - Spatial distances (especially used with maps, geo maps, or any spatial locations like building, stadium, campus, etc.)
- Examples
  - http://www.gartner.com/technology/research/methodologies/research_mq.jsp (top left chart)
  - http://en.wikipedia.org/wiki/Cluster_analysis (bottom left chart)

Visual Property: **Texture**

- Texture is used much like colors but seldom used for continuous data.
- It is important when color sensitivity is an issue. Implementations include fill patterns, border patterns, shadow, etc.
- Examples





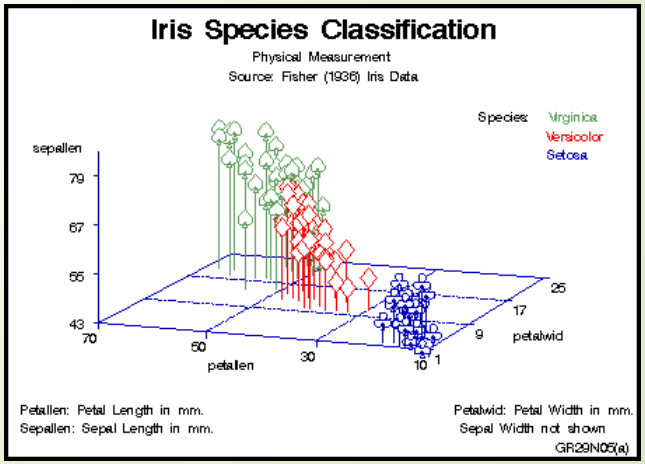


Visual Property: Shape

- Type of shapes
  - Shapes can be formed using simple shapes: square, triangle, etc.
  - More complex shapes also can be formed by combination of simple shapes: icon, marker, etc.
- Shapes are usually used to represent different type of things, or nominal/discrete data (e.g., category). Shapes are rarely used for continuous data.
- Shapes are usually used for dimensions but sometimes used for measures as well.
- Examples
  - http://www.masters.com/en_US/scores/ - in this example, shapes are used for measures (ordinal) actually

Hole	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
Par	4	5	4	3	4	3	4	5	4	4	4	3	5	4	5	3	4	4	72
R1	4	4	4	3	4	3	4	4	5	4	5	3	4	3	4	3	4	4	69
R2	3	4	3	4	4	3	3	4	3	5	4	3	4	3	4	4	4	4	66
R3	4	5	5	3	3	3	4	4	3	3	4	4	3	4	3	4	4	4	67
R4	5	5	3	3	4	4	3	5	4	4	5	2	5	3	5	3	4	4	71

  - http://v8doc.sas.com/sashtml/gref/z15-ex.htm - shapes are used for species



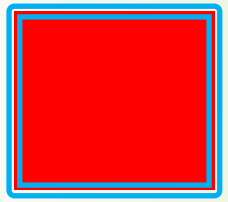
Motion - on Top of SCOPeS

- Motion is the movement or transformation of the basic visual properties, which can represent richer meanings and grab greater attention.
  - Dynamic change of positions, shape, color, orientation, texture, size
  - Note: simply animating an object in the visualization can just be decorative, but not data driven
- Examples:
  - Simple position change: move one position to the other
  - Growing or shrinking size, and/or accompanies by shape shifting
  - Moving direction (of position change): used like orientation
  - Flickering/blinking/flashing/spinning pattern: can be used like color or texture for categorical data
- Speed (of movement and changing) can be added as additional properties in motion driven visualizations
- Typical usages and examples
  - Show trends along a period of time - https://www.linkedin.com/feed/update/activity:6464536253118373888/
  - Operational process simulation: e.g., disk optimization
  - http://metrocosm.com/global-immigration-map/
  - http://www.storybench.org/role-motion-visualizations/
  - https://www.gapminder.org/tools/#$chart-type=bubbles&url=v1
  - https://www.digitalattackmap.com

We do not cover motion in depth in IT 7113, but it can be a very good research topic for class project.

Composition of Multiple Properties

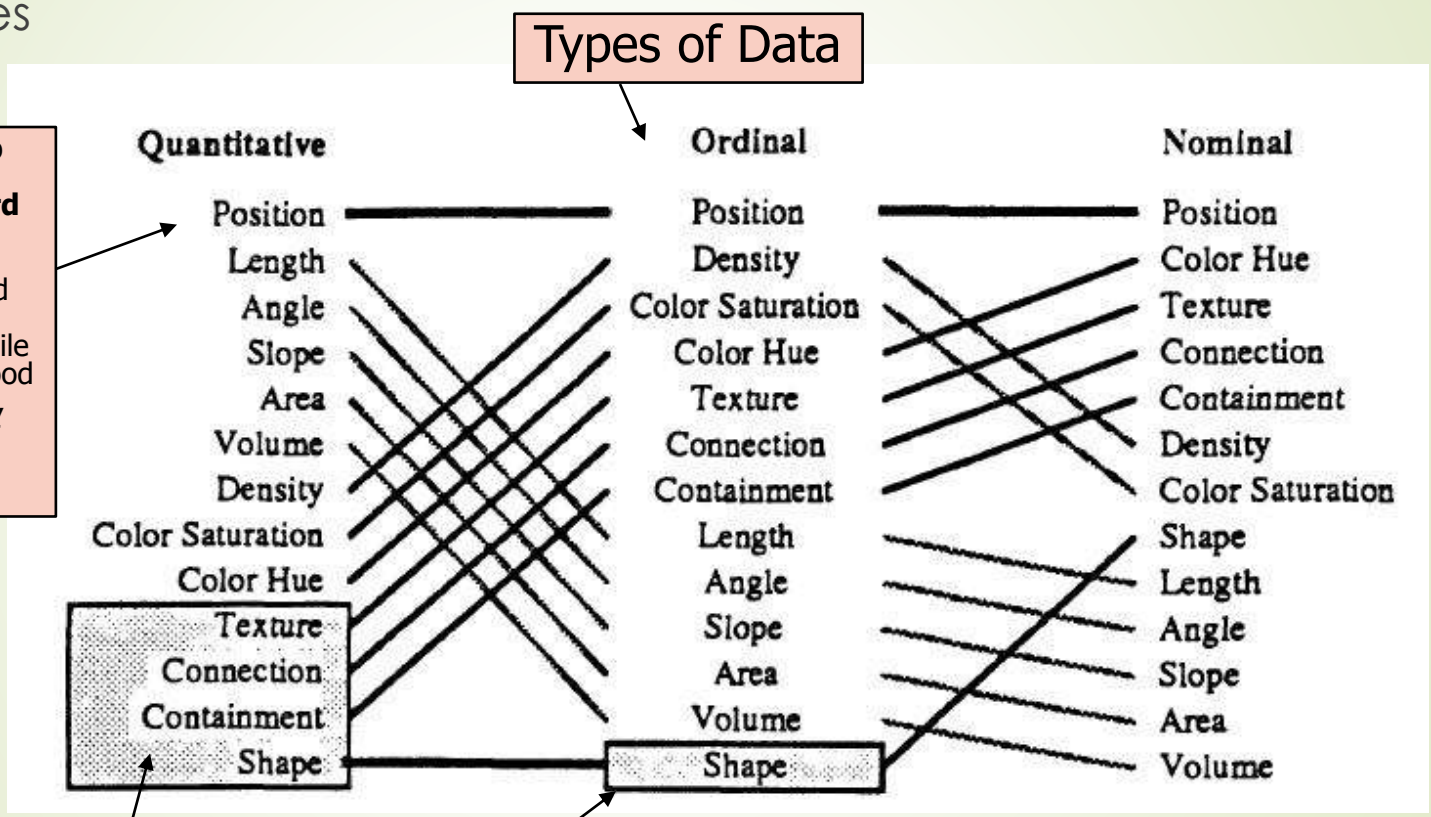
- Combinations of these properties can be used to represent multi-dimensional data in the same visualization.
- Example: here is one product visualized as an object with small (size) square (shape), red color (color), and double-line border (texture).
  - Size: representing sales amount
  - Shape: representing product type
  - Color: representing profit/loss
  - Line textual: representing sales channel



- More complex visual elements (such as icons and symbols) can be built based on the basic elements and properties discussed above.
- See more examples of the visual property use at http://innovis.cpsc.ucalgary.ca/innovis/uploads/Courses/InformationVisualizationDetails/09Bertin.pdf

Visual Properties Choice

- Choose appropriate visual properties to code data based on data types



Best choice on top and then less appropriate toward the bottom. For example, position property can be used to code all types of data really good; while area is somewhat good for quantitative data, but it is not good for ordinal and nominal data.

The shaded box indicates most inappropriate.

For example, do not use shape for ordinal type of data, like month, or class level, etc.

Refer to https://infovis-wiki.net/wiki/Visual_Variables

Also see https://www.qlik.com/blog/visual-encoding

Visual Properties used for Dimensions and Measures

- Choose one visual property to represent one dimension. E.g., use colors for products.
- Use different visual properties, instead of different values of the same visual property, for different dimensions (or dimension attributes). E.g. if colors are used for products, then do not use color for countries again. Choose position or shape for countries.

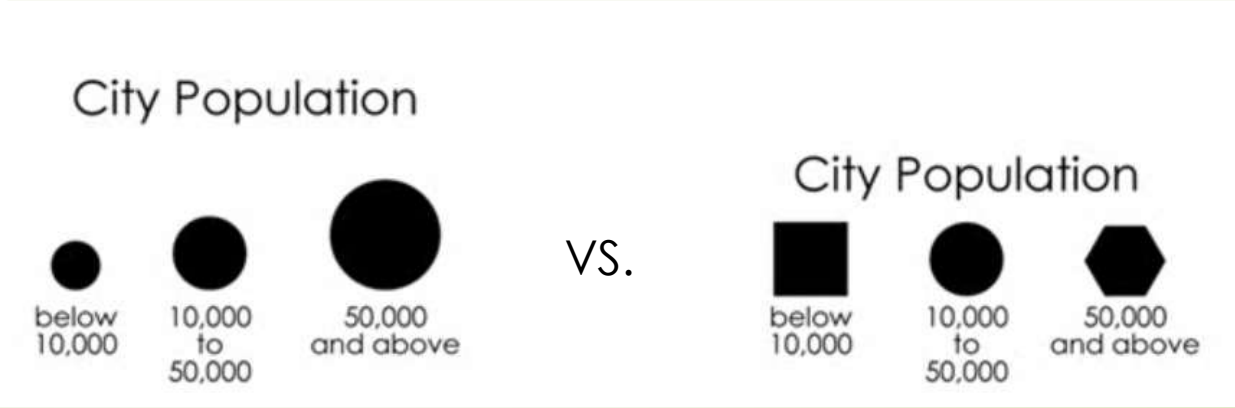
	Measure (continuous)	Measure (ordinal)	Dimension (categorical: ordinal or nominal)
Shape	X	*	***
Color (Hue)	*	*	***
Color (Value)	**	**	*
Orientation	**	**	**
Position	***	**	**
Texture	X	*	**
Size	***	**	*

* Occasionally, only in specific cases
** Sometimes, depending on more factors
*** Very often and appropriate

Exercise: which visual coding method is better?

- Representing product categories

- Representing population



Readings and Resources

- Visual variable/property basics
  - https://infovis-wiki.net/wiki/Visual_Variables
  - http://innovis.cpsc.ucalgary.ca/innovis/uploads/Courses/InformationVisualizationDetails/09Bertin.pdf
- Choosing visual encoding
  - https://www.qlik.com/blog/visual-encoding
  - Choose Appropriate Visual Encodings https://www.oreilly.com/library/view/designing-data-visualizations/9781449314774/ch04.html
  - For maps https://www.e-education.psu.edu/geog486/node/594
- More articles and books
  - http://wiki.gis.com/wiki/index.php/Visual_variable
  - Considering Visual Variables as a Basis for Information Visualisation, https://prism.ucalgary.ca/handle/1880/45758
  - Roth, R. E. 2017. Visual Variables. The International Encyclopedia of Geography. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118786352.wbieg0761
  - Semiology of Graphics: Diagrams, Networks, Maps 1st Edition, by Jacques Bertin, https://www.amazon.com/dp/1589482611