



THE UNIVERSITY OF
MELBOURNE



**GEOM90007
SPATIAL VISUALISATION**

LECTURE 4:

DATA GRAPHICS 1 -
VISUAL VARIABLES

REVIEW

Excellence in statistical graphics consists of complex ideas communicated with:

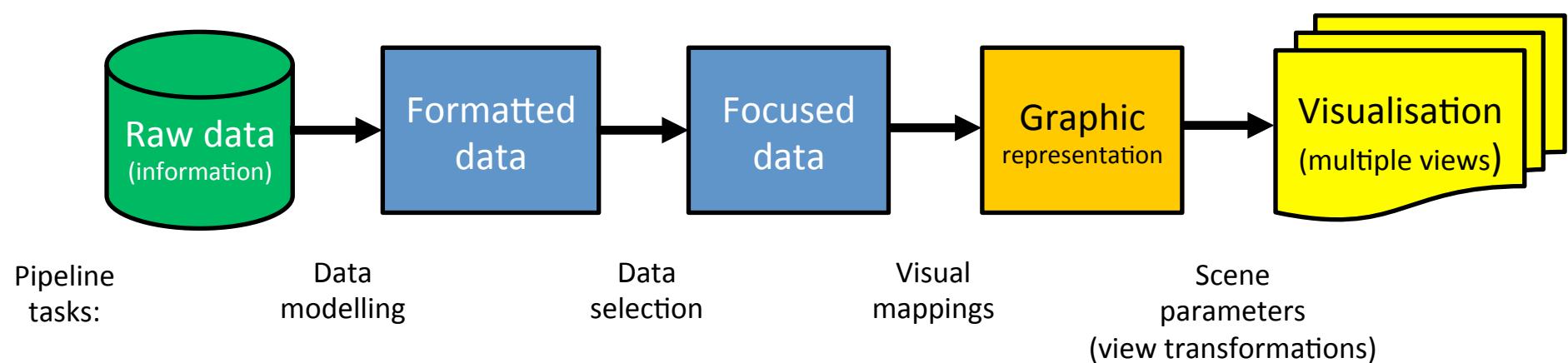
- Clarity
- Precision
- Efficiency

Graphics *reveal* data.

(Tufte, 2001)

REVIEW

Visualisation pipeline (Lecture 1)



REVIEW

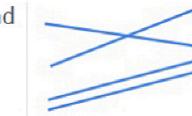
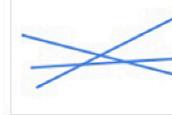
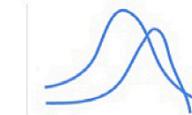
Sensing (Lecture 2)

- Stimuli
- Human visual system
 - Colour deficiencies (diversity)
 - Acuity
 - Contrast

Perception (Lecture 3)

- Light (luminance, brightness, contrast)
- Steven's law and perceived sensation
- Colour vision (deficiencies, models)
- Pre-attentive processing (subconscious) “pop out” effects
 - Categories (form, colour, movement, spatial position)
- Post-attentive processing (conscious) serial searching
- Glyphs (discrete), texture (continuous)

Patterns – Taking Pre-attentive Processing a Step Farther

Pattern	Example	Pattern	Example
High, low and in between		Non-intersecting and intersecting	
Going up, going down and remaining flat		Symmetrical and skewed	
Steep and gradual		Wide and narrow	
Steady and fluctuating		Clusters and gaps	
Random and repeating		Tightly and loosely distributed	
Straight and curved		Normal and abnormal	

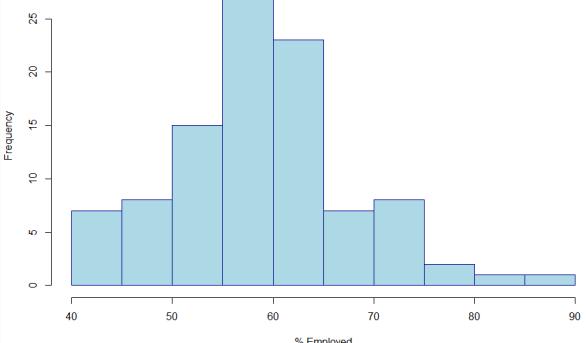
REVIEW

Statistical foundation (Lecture 4)

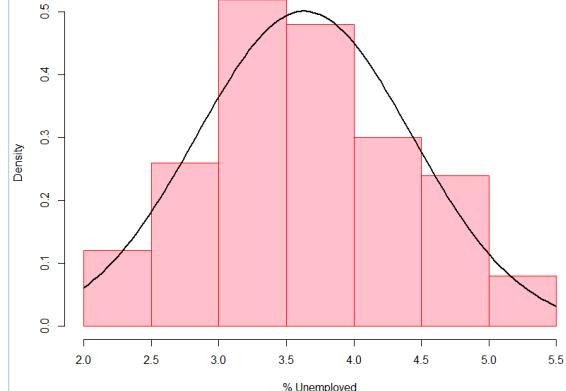
- Types of data
 - Numerical
 - Categorical
 - Ordinal
 - Nominal

Understanding your data is essential for designing a visualisation

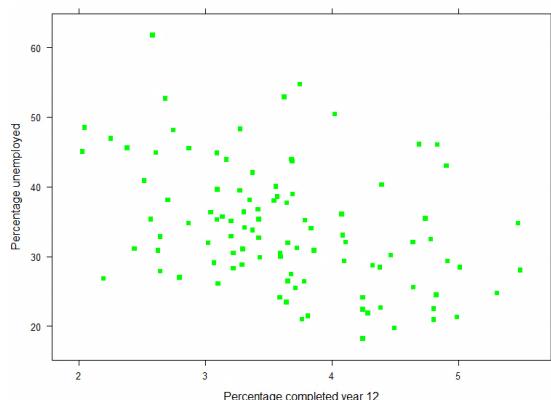
Histogram of employment in the top 100 cities of Australia (ABS Census 2011)



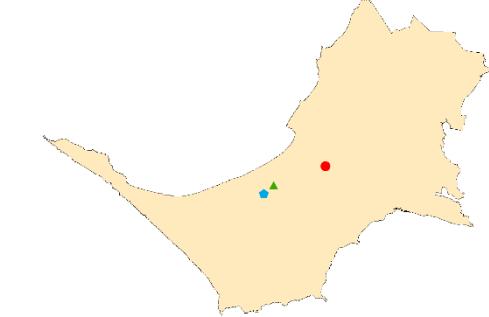
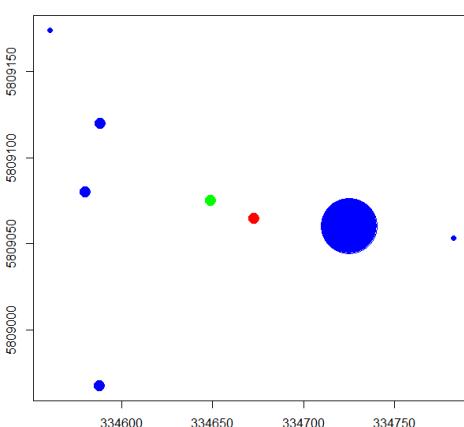
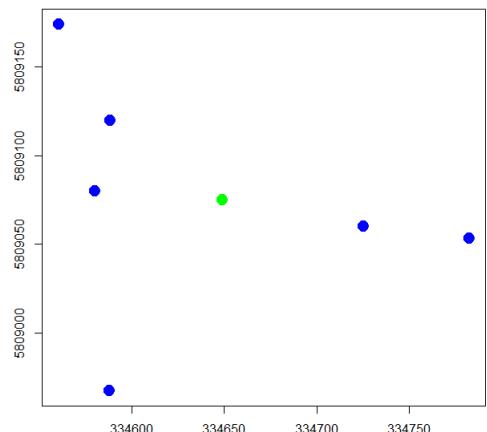
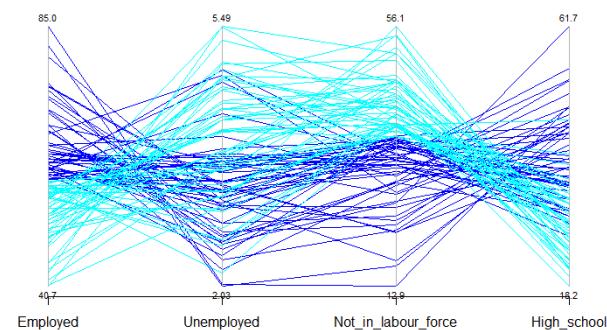
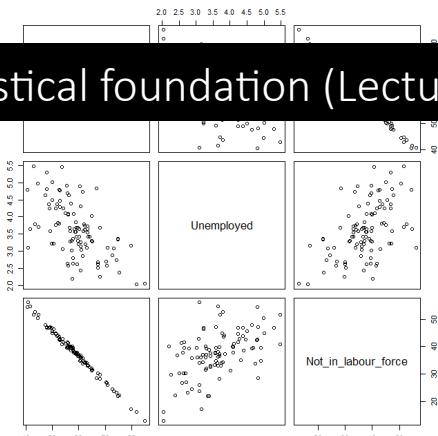
Histogram of unemployment in the top 100 cities of Australia (ABS Census 2011)



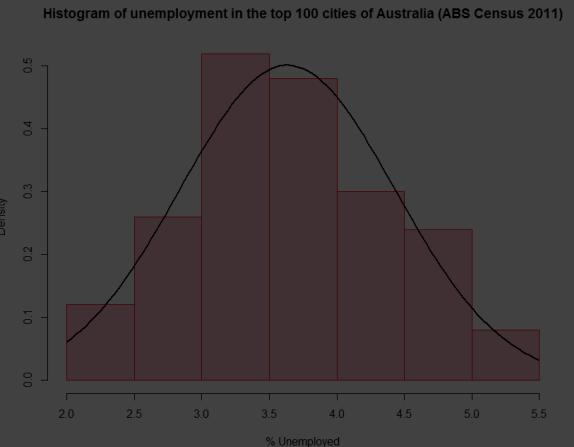
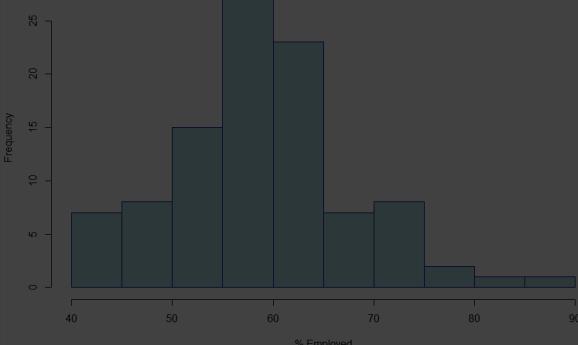
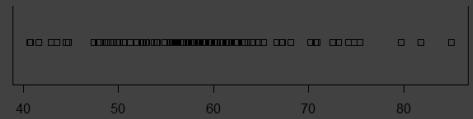
Percentage unemployed against percentage completed year 12
top 100 cities of Australia (ABS Census 2011)



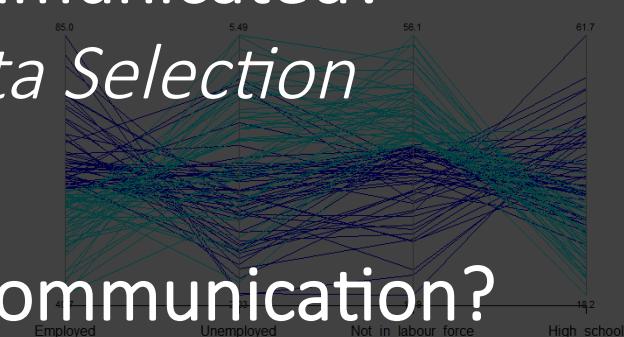
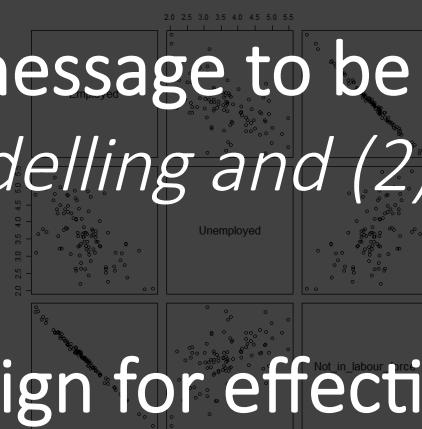
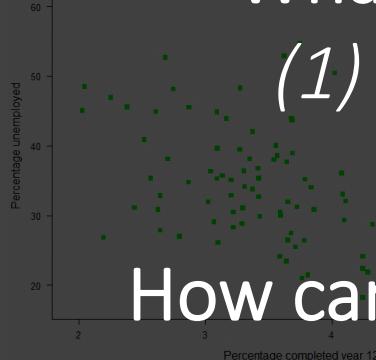
Statistical foundation (Lecture 3)



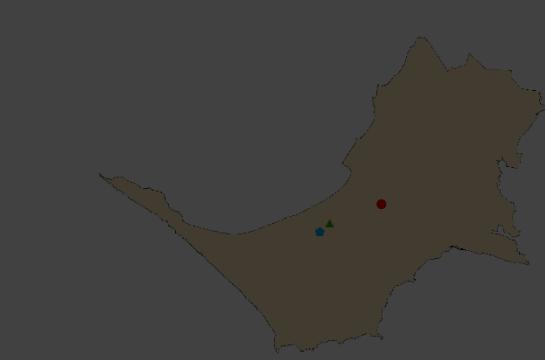
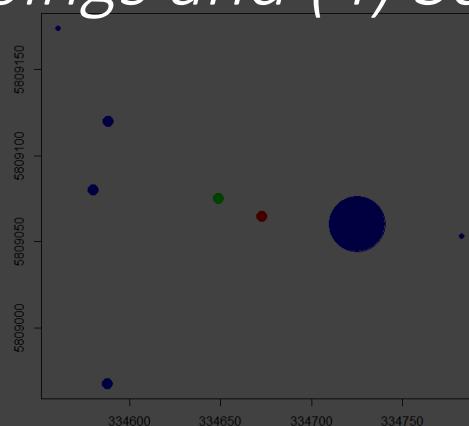
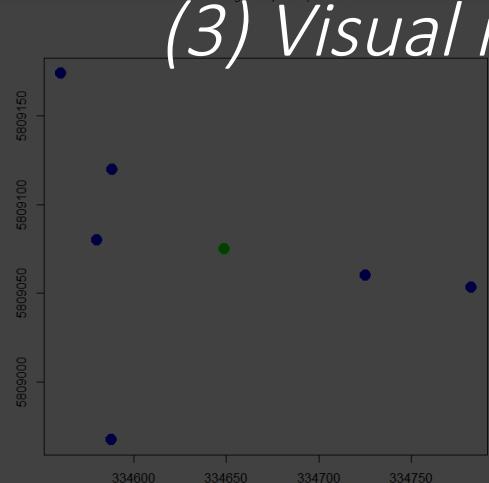
Histogram of employment in the top 100 cities of Australia (ABS Census 2011)

Percentage unemployed against percentage completed year 12
top 100 cities of Australia (ABS Census 2011)

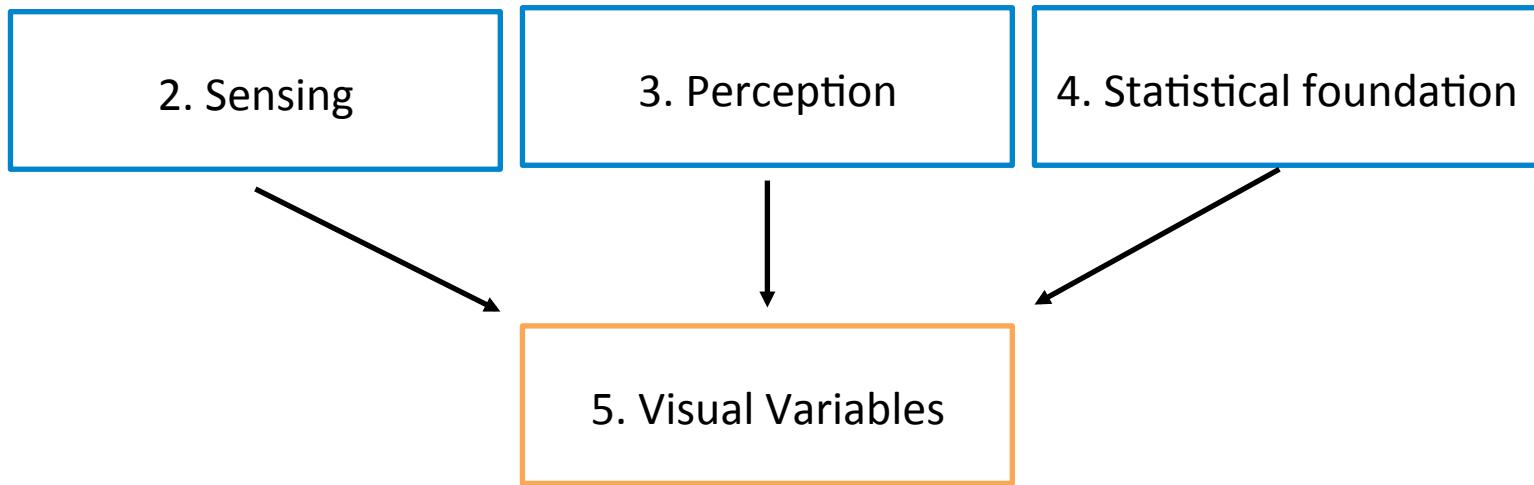
What is the message to be communicated?
(1) Data modelling and (2) Data Selection



How can you design for effective communication?
(3) Visual mappings and (4) Scene parameters



FOCUSING



- A. Semiology and visual language
- B. Visual variables
- C. Value message

A. SEMIOLOGY AND GRAPHIC SYMBOLS

In data graphics, a visual object is called a *graphical symbol*, the science of which is called *semiology*

The meaning of common symbols is learned



No entry symbol
Image: www.clipartbest.com

A. VISUAL VARIABLES

Using graphics to communicate information requires an understanding of primitives and their properties

Graphic primitives (basic units) used in data graphics are referred to as **marks** that can vary using **visual variables**

“Our usual way of communicating is with words. Written words consist of single symbols (letters)... If there are basic visual symbols arranged in a particular way, can they be used to convey information in a similar manner?”

www.infovis-wiki.net

A. EVOLUTION OF VISUALISATION

Jacques Bertin

- Semiology of graphics, 1967
 - 2D
 - Three graphic primitives (points, lines, areas)
 - Six retinal variables



Jock Mackinlay, 1986

- Proposed an automatic graphic presentation tool
 - 1D/2D/3D
 - Two additional primitives (surfaces, volumes)
 - Temporal (animation)

...

A. VISUAL LANGUAGE AND COMMUNICATION

Semiology within semiotics is the science of signs and symbols

Each mark describes a piece of data (single observation)

Visual variables to describe extents (domain and range) of data

Mackinlay's two criteria:

Expressiveness communicating the desired information

Effectiveness of a representing salient features using the display
considering the human visual system



A. DIFFERENT DISPLAY SPACES

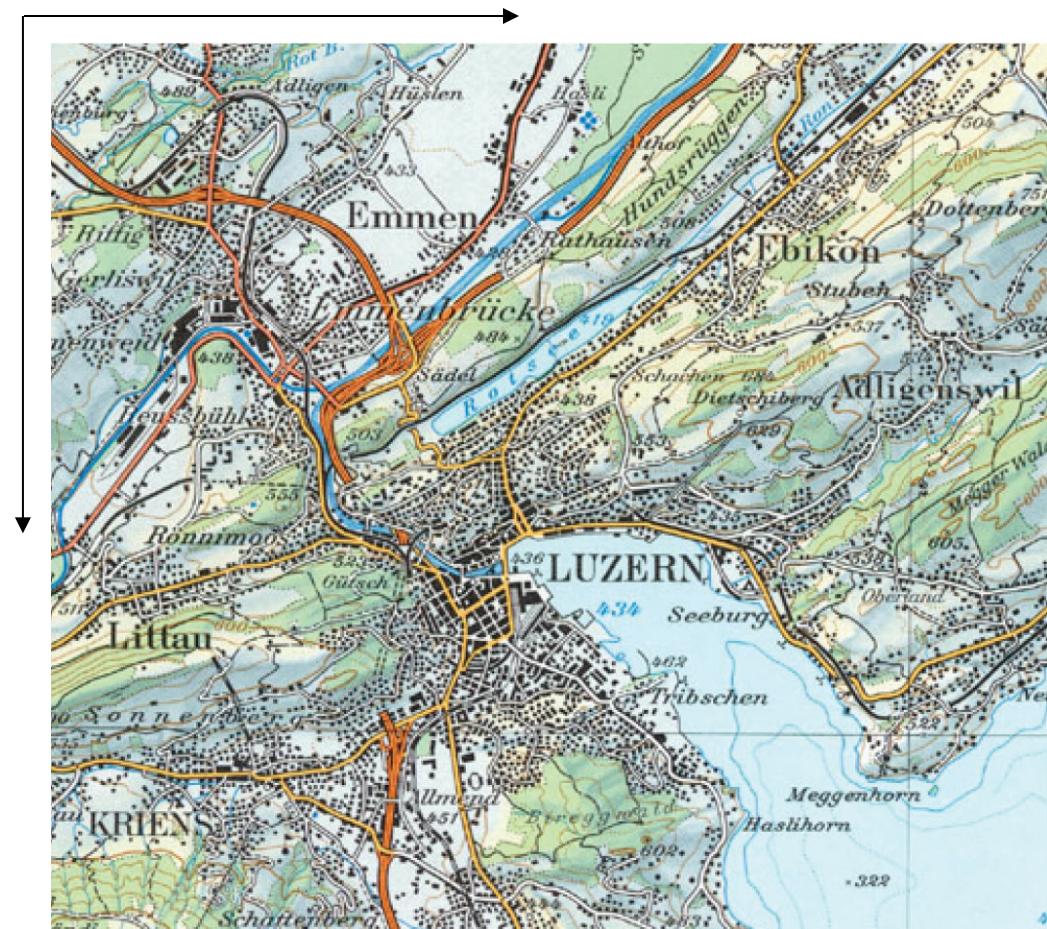
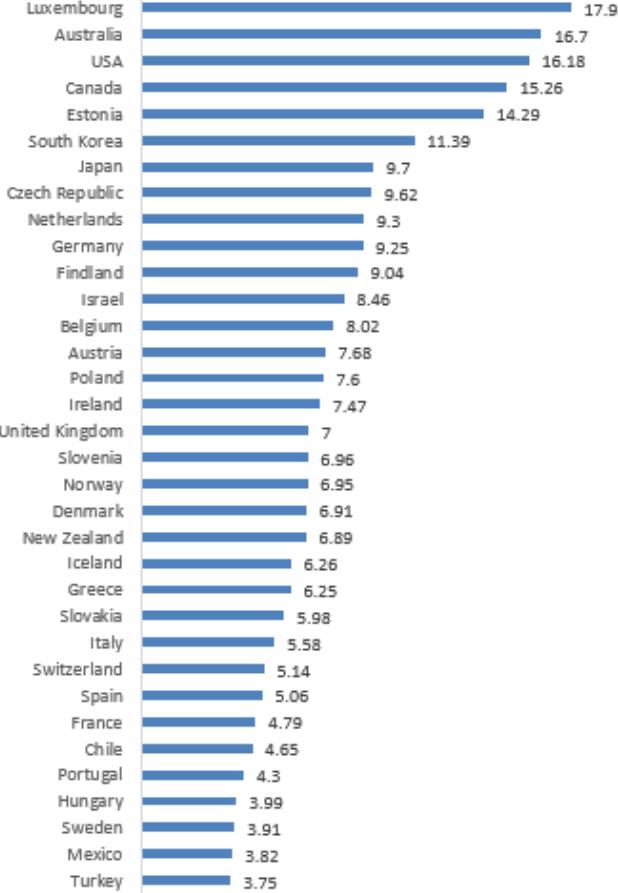


Image: Swisstopo (2002)

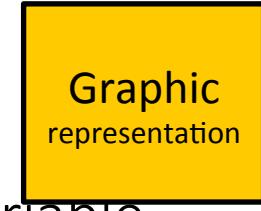
A. MAPPING VARIABLES



Data variable
(following data modelling
and selection tasks)



Map



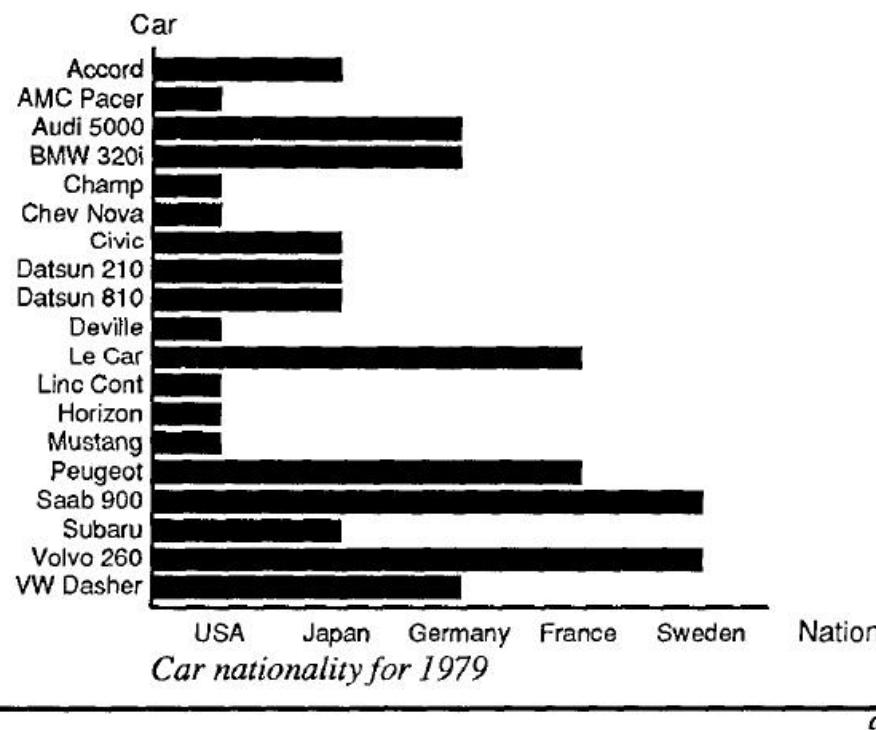
Visual variable

1. Position
2. Shape
3. Size
4. Brightness
5. Colour
6. Orientation
7. Pattern
8. Motion

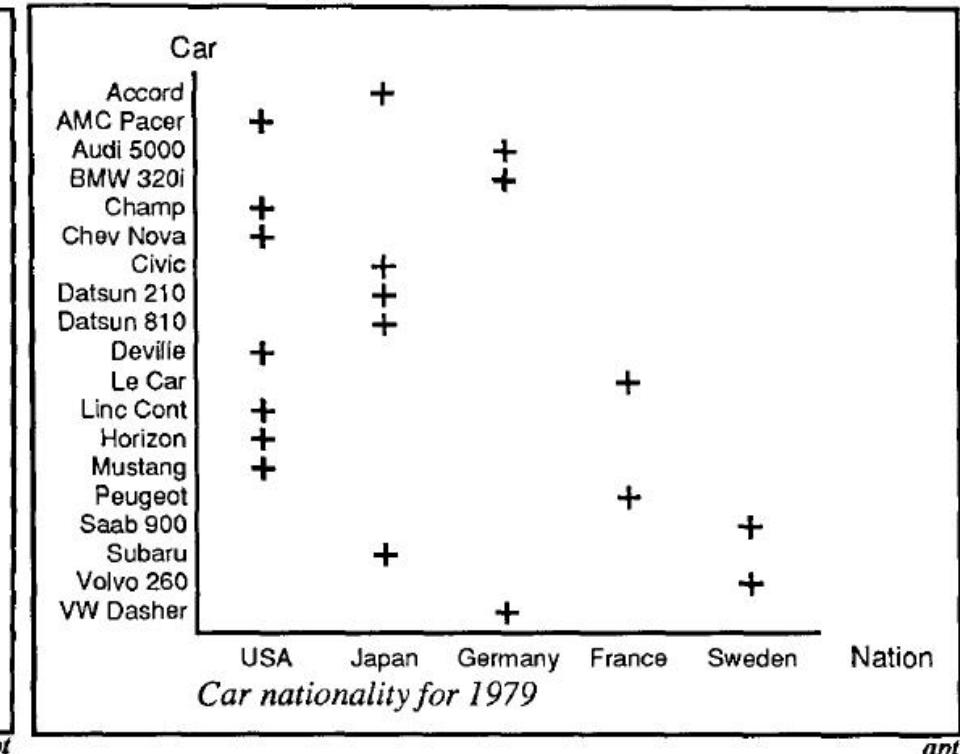
Challenge:

The wrong mapping can mislead
or confuse the viewer!

A. LOGIC IS NECESSARY



Poor use of visual variables (bar chart)



Better use of visual variables (scatter plot)

A.1 POSITION

The location of geometry within the display, 1D/2D/3D

Selection of variables is key for communication:

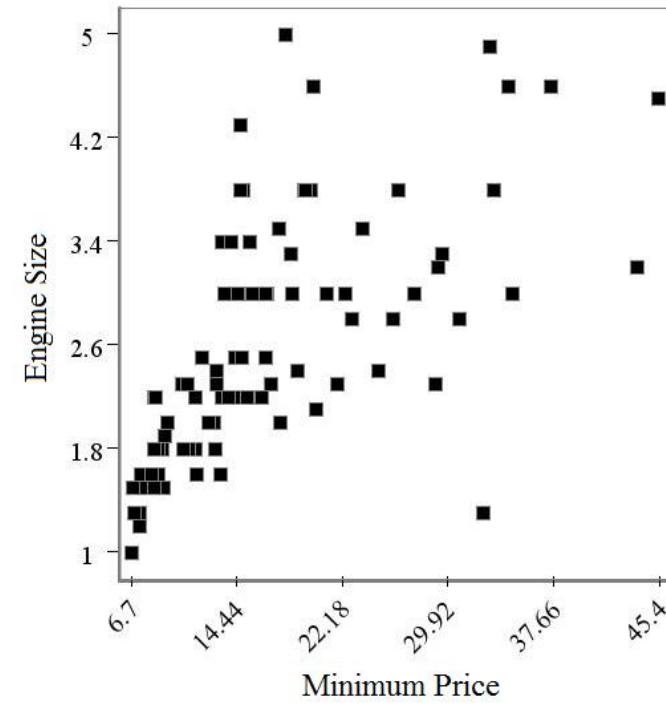
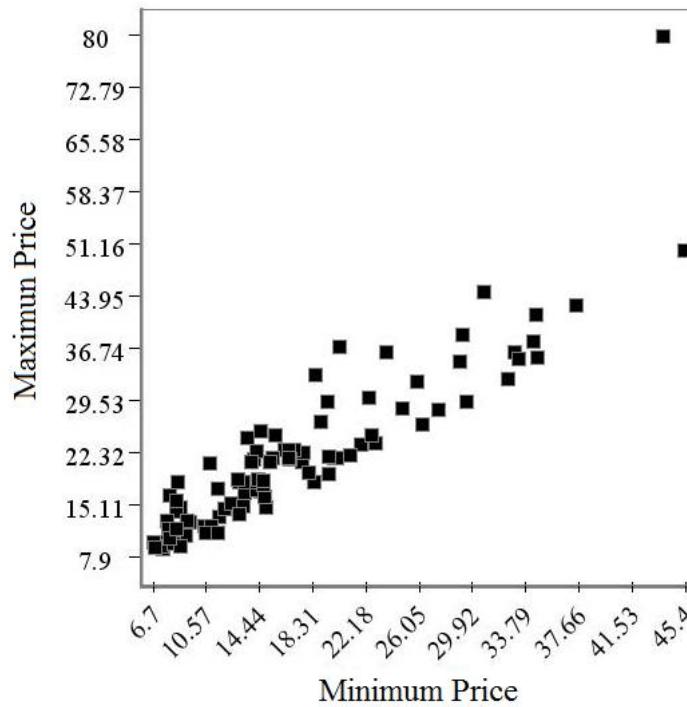
- Distribution of data
- Choice of scales

Best schema maps each variable to unique positions without overlap!



Image: Ward et al (2010)

A.1 POSITION



Example visualizations: (left) using position to convey information. Displayed here is the minimum price versus the maximum price for cars with a 1993 model year. The spread of points appears to indicate a linear relationship between minimum and maximum price; (right) another visualization using a different set of variables. This figure compares minimum price with engine size for the 1993 cars data set. Unlike (left), there does not appear to be a strong relationship between these two variables.

A.2 SHAPE

Marks (glyphs) and their representation

- Graphic primitives (points, lines, areas, surfaces, volumes)

Geometric vs pictographic symbols (beware semantic issues)



Image: Ward et al (2010)

A.3 SIZE

Use of length, area or volume to encode quantitative attributes

- Data's extents (range and domain)

Easily maps to continuous numerical data, size supports gradual increments. Caution is advised when using categorical data.

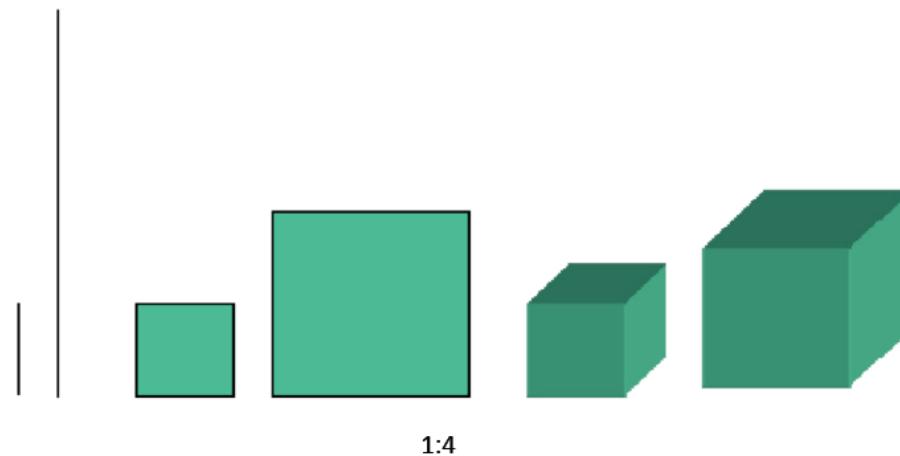
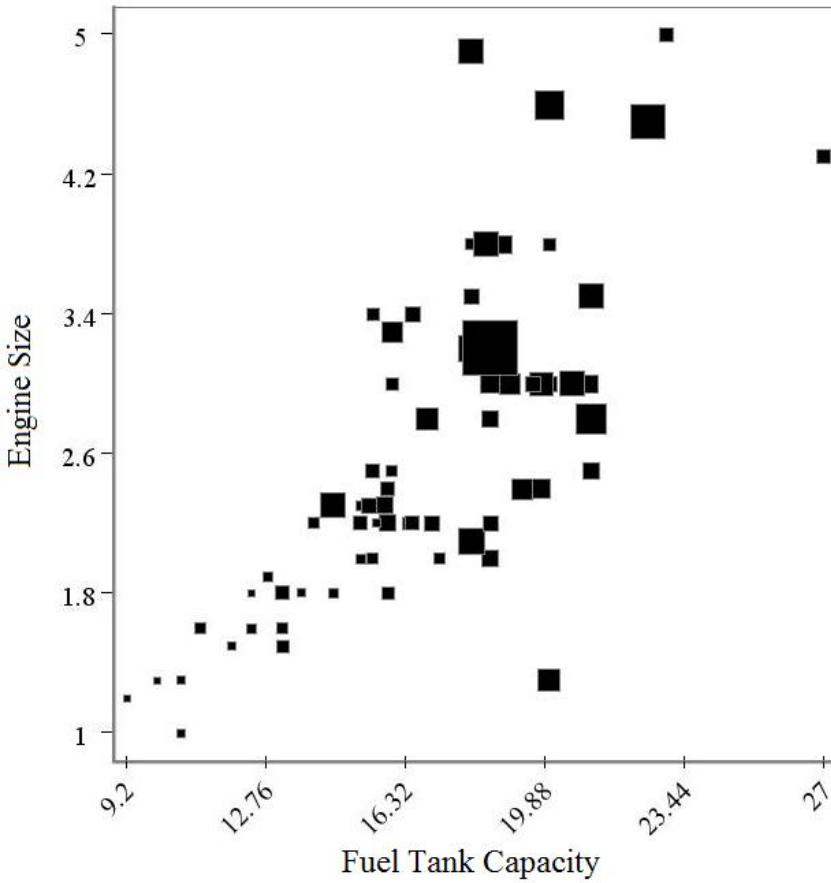


Image: Ward et al (2010)

A.3 SIZE



This is a visualization of the 1993 car models data set, showing engine size versus fuel tank capacity. Size is mapped to maximum price charged.



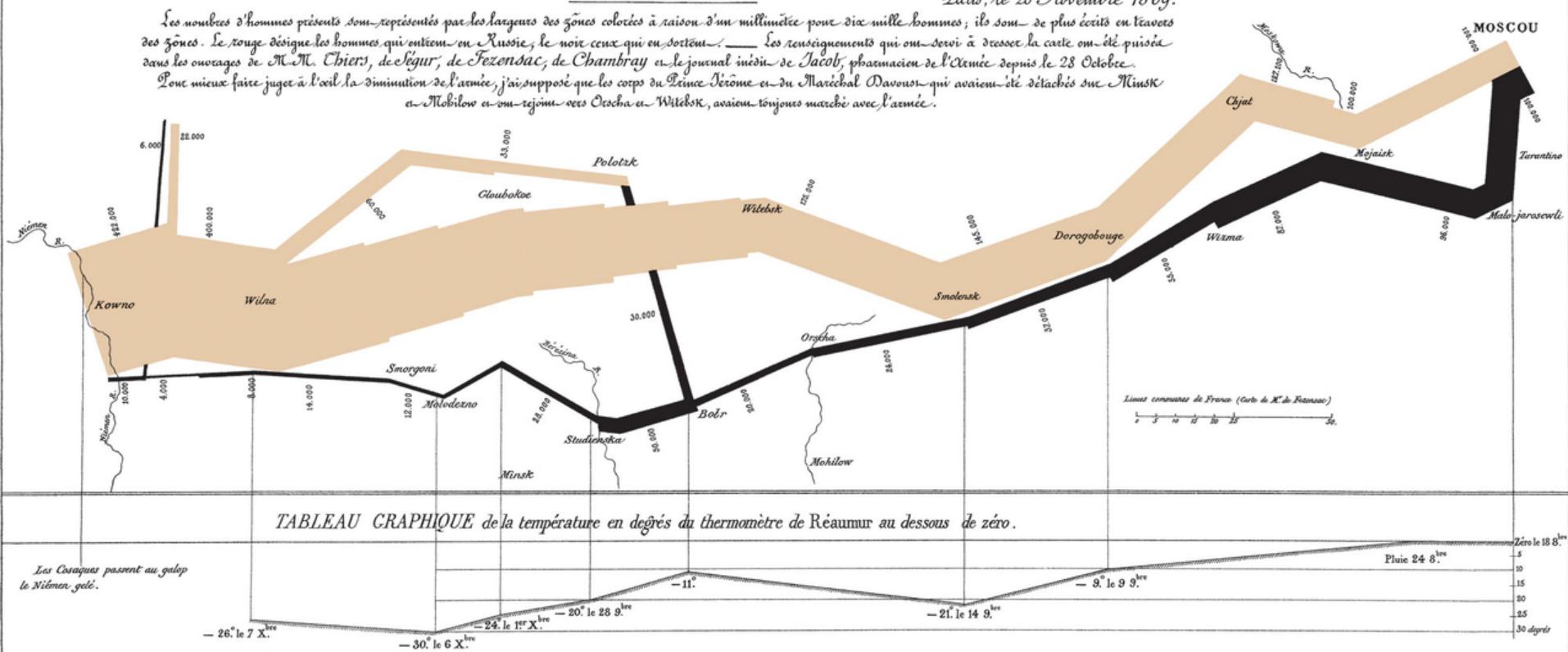
A.3 SIZE

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dessiner la carte ont été puisés dans les ouvrages de M. Chiers, de l'Égur, de Fezensac, de Chambray et le journal médical de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés de Minsk au Mohilow et se rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.



A.4 BRIGHTNESS

Perceived amount of light or value: 0 = absence of light (black)

Grey scale ramp represent relative differences in numerical data

Beware the ‘simultaneous brightness contrast’ issue!

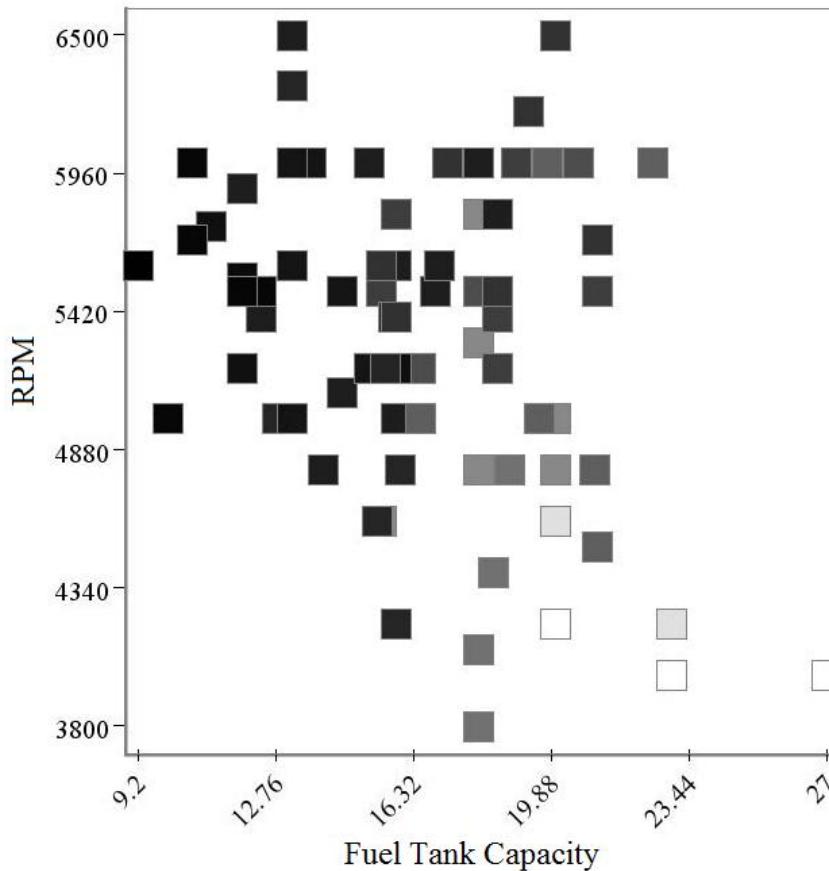
6 classes



Ramp



A.4 BRIGHTNESS



Another visualization of the 1993 car models data set, this time illustrating the use of brightness to convey car width (the darker the points, the wider the vehicle).

A.5 COLOUR

Typically HSV or HSL are used as they are more intuitive than RGB*

- Colour hue
 - Ideal for categorical data
 - Poor for quantitative data
- Colour saturation
 - 100 = colourful (saturated)
 - 0 = colourless (grey)



Image: Microsoft's 2D colour picker

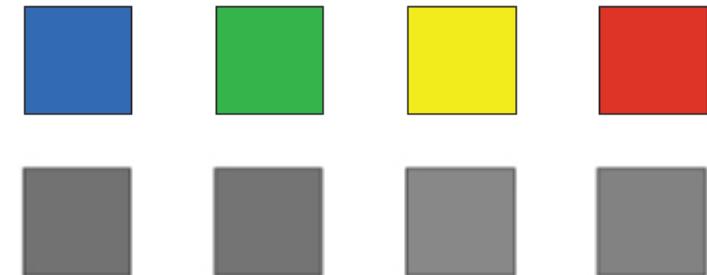


Image: Ward et al (2010)

A.5 COLOUR

Hue: An angular measurement (analogous to position around a colour wheel), independent of intensity or lightness

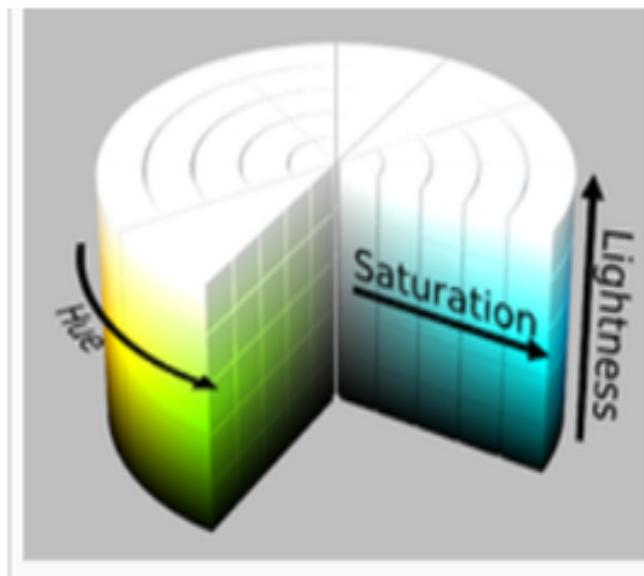
Value: Brightness. A value of 0 represents the absence of light, or black. A maximum value means that the colour is at its brightest

Saturation: Colour intensity. A value of 0 means that the colour is “colourless” (grey); a maximum value means colour at maximum “colourfulness” for hue angle and brightness.

Lightness: Maximum value means colour is white (regardless of the current values of the hue and saturation components).

A.5 COLOUR – HSL/HSV

HSL



HSV

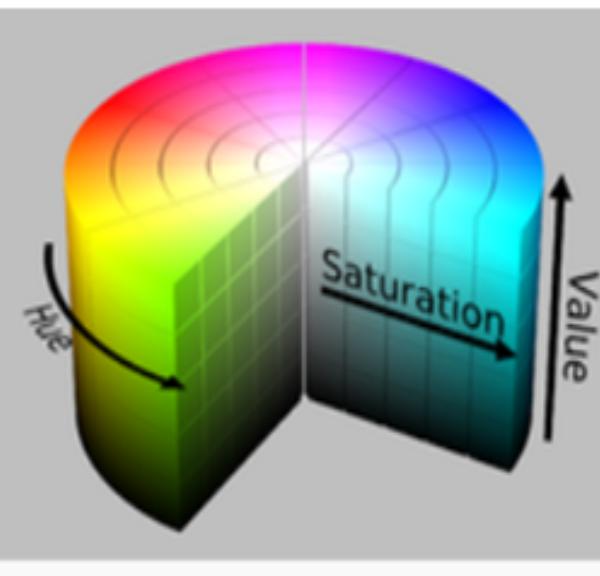
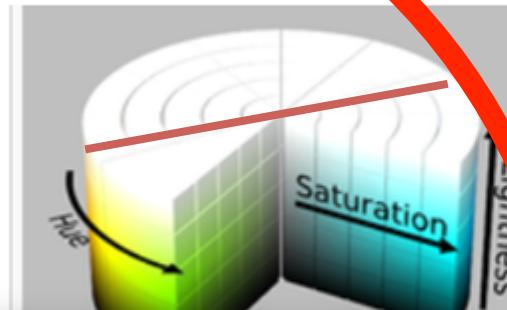


Image: Wikimedia commons

A.5 COLOUR HSL/HSV

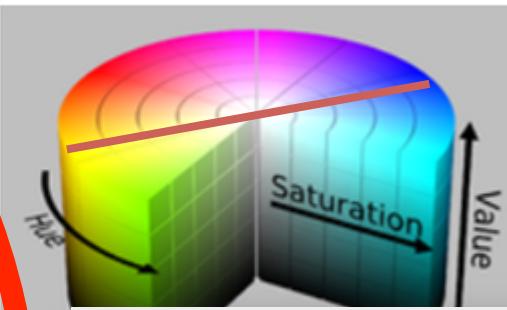
HSL



$H = 240^\circ$
(Blue)



HSV



$H = 240^\circ$
(Blue)

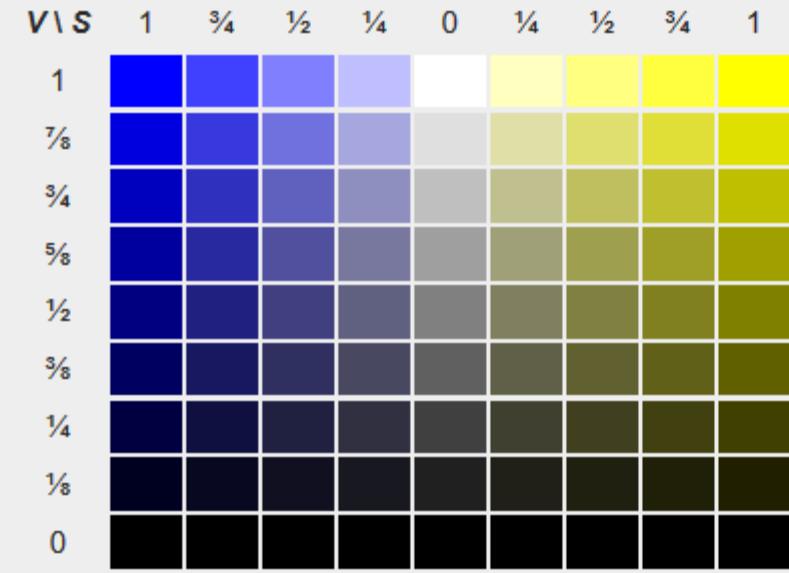


Image: Wikimedia commons

A.5 USING THE HSL SCHEME

Variables should first and foremost be ordered by lightness, “people are not accustomed to accurately comparing saturation levels, especially between hues”
 (Brewer, 1999)

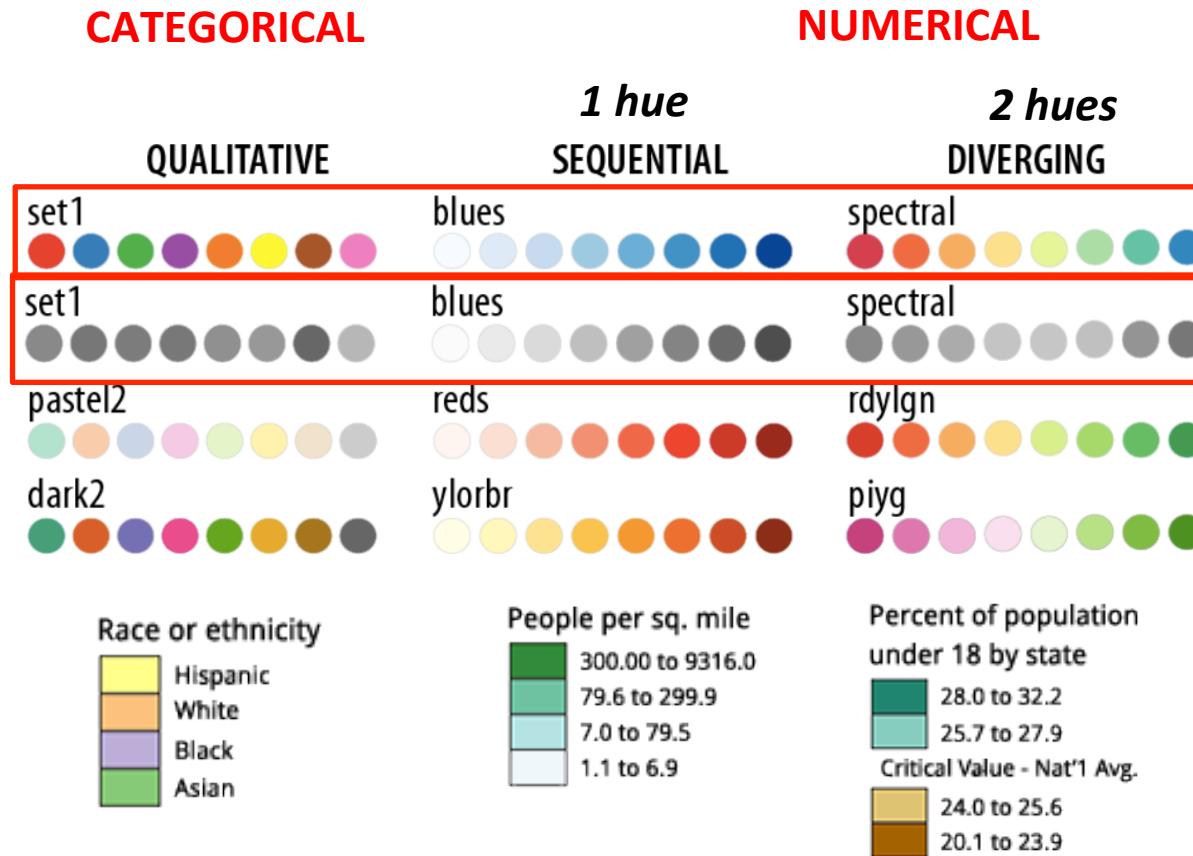


Image: Krzywinski
<http://mkweb.bcgsc.ca/brewer/>

Images: Brewer
www.colorbrewer2.org/

More information:
http://colorbrewer2.org/learnmore/schemes_full.html

A.5 COMMON COLOUR MAPS



standard linear gray scale



rainbow



heated

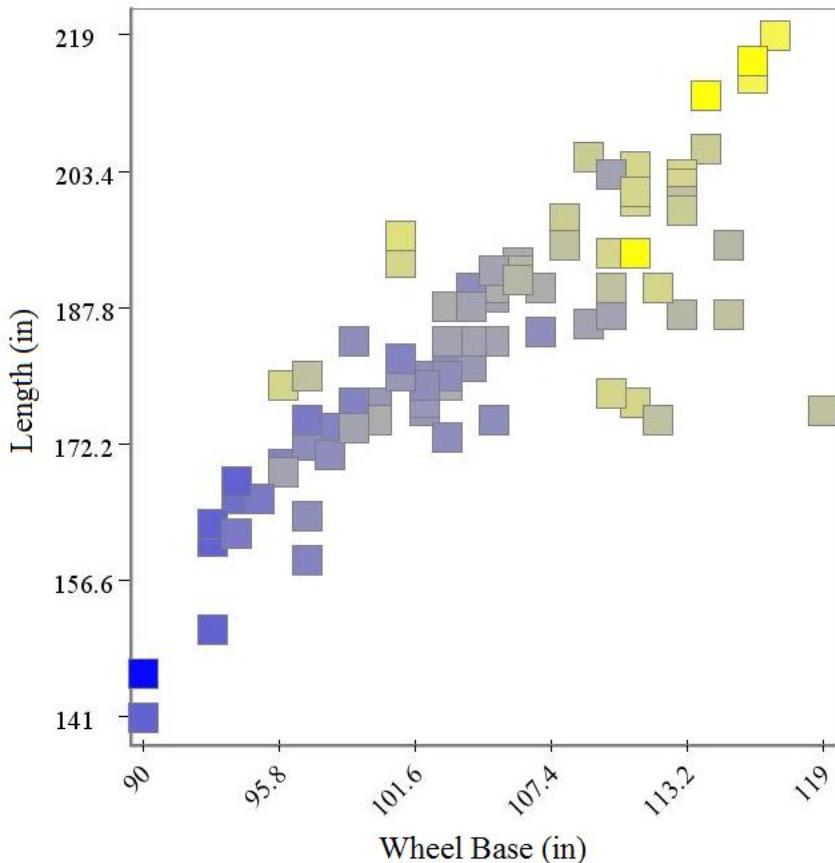


blue to cyan



blue to yellow

A.5 COLOUR



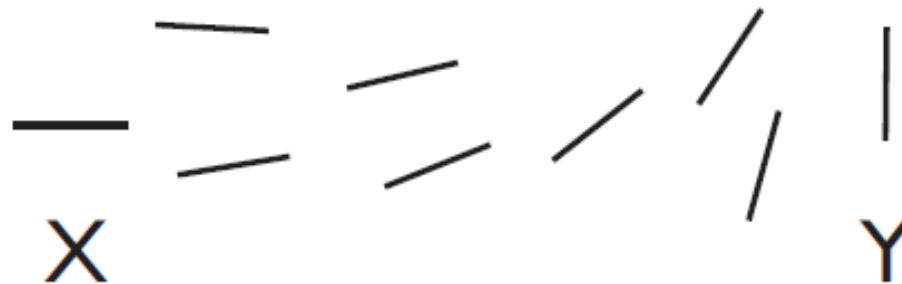
A visualization of the 1993 car models, showing the use of color to display the car's length. Here length is also associated with the y-axis and is plotted against wheelbase. In this figure, blue indicates a shorter length, while yellow indicates a longer length.

A.6 ORIENTATION

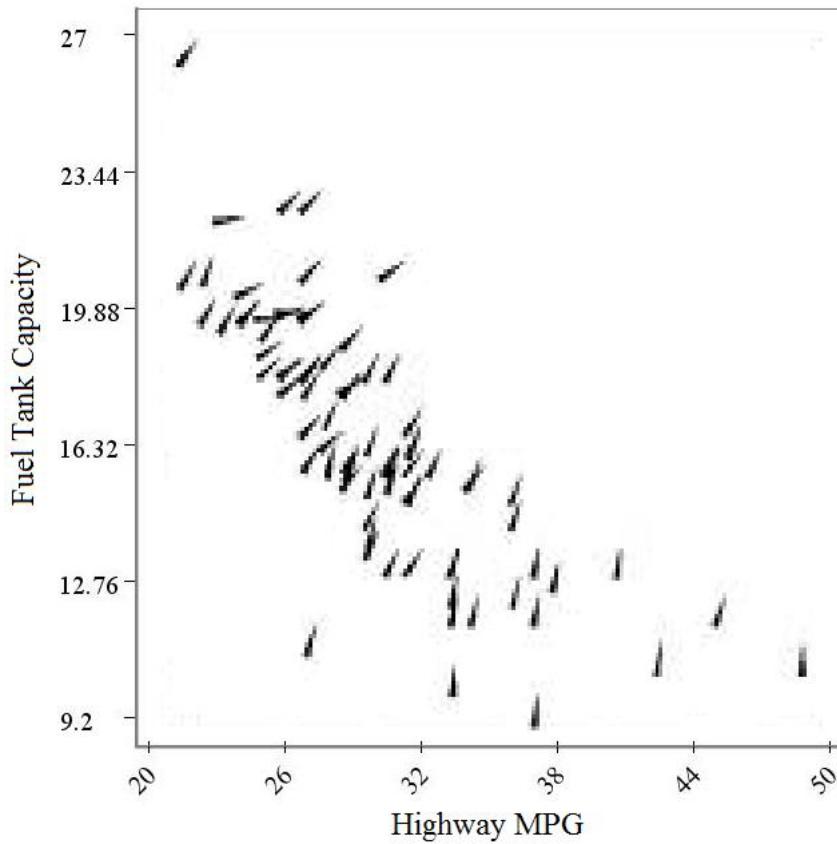
Orientation (or direction) is directly tied to pre-attentive vision

Suitable point marks have a single major axis (e.g., eq. triangle)

Variable can be potentially mapped to 360°



A.6 ORIENTATION



Sample visualization of the 1993 car models data set depicting using highway milesper-gallon versus fuel tank capacity (position) with the additional data variable, midrange price, used to adjust mark orientation.

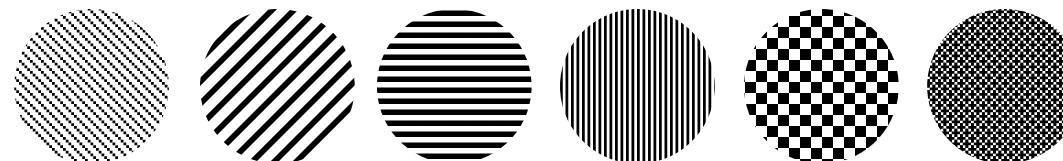
A.7 PATTERN / TEXTURE

Texture, focus and arrangement of marks

Focus and arrangement (e.g., clusters)

Texture can be considered variation in grain, a combination of the other variables (refer pre-attentive)

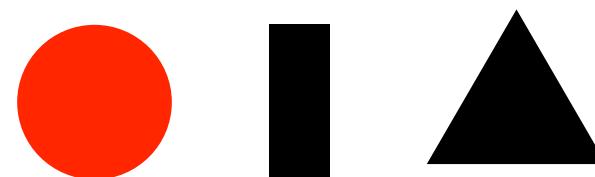
1. Size
2. Contrast
3. Orientation
4. Colour hue



A.8 MOTION

Can be associated with any of the previous variables to communicate some change over time

e.g., colour (saturation), orientation, size



IMPORTANT NOTES

BEWARE OF SEMANTICS



SOME OTHER COLOUR ISSUES (THERE ARE MANY)

1. Issues using a spectral (rainbow) colour ramp for quantitative data
 1. Colour deficiency issues
 2. Changes very hard to see – how to order colour?



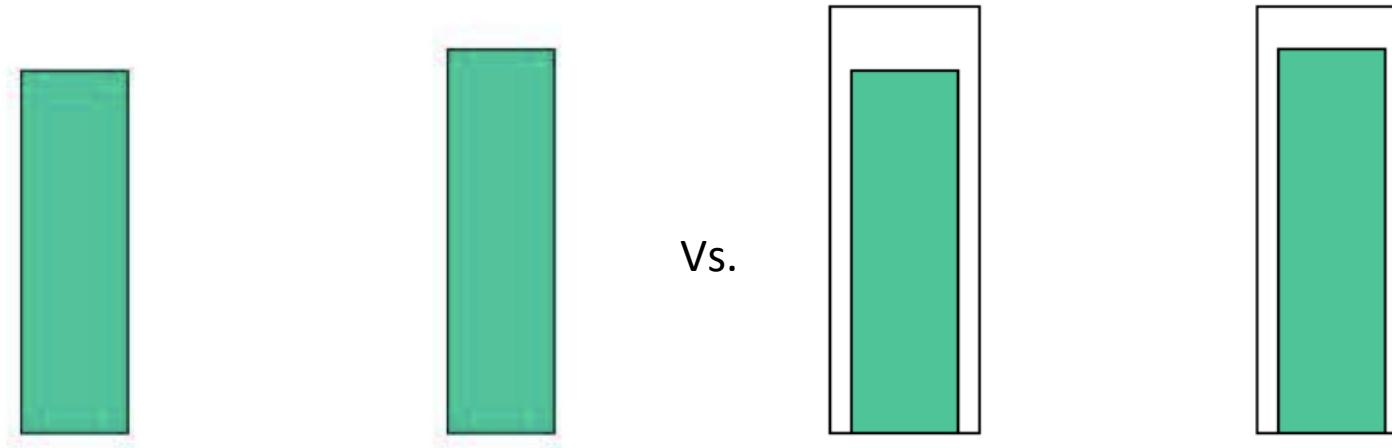
Reading: http://www.personal.psu.edu/cab38/Pub_scans/Brewer_1997_Spectral_CaGIS.pdf

2. Social/cultural/contextual perspectives (e.g., red)

Judgments in perception: Weber's Law

Detecting change is proportional to the relative, not the absolute change

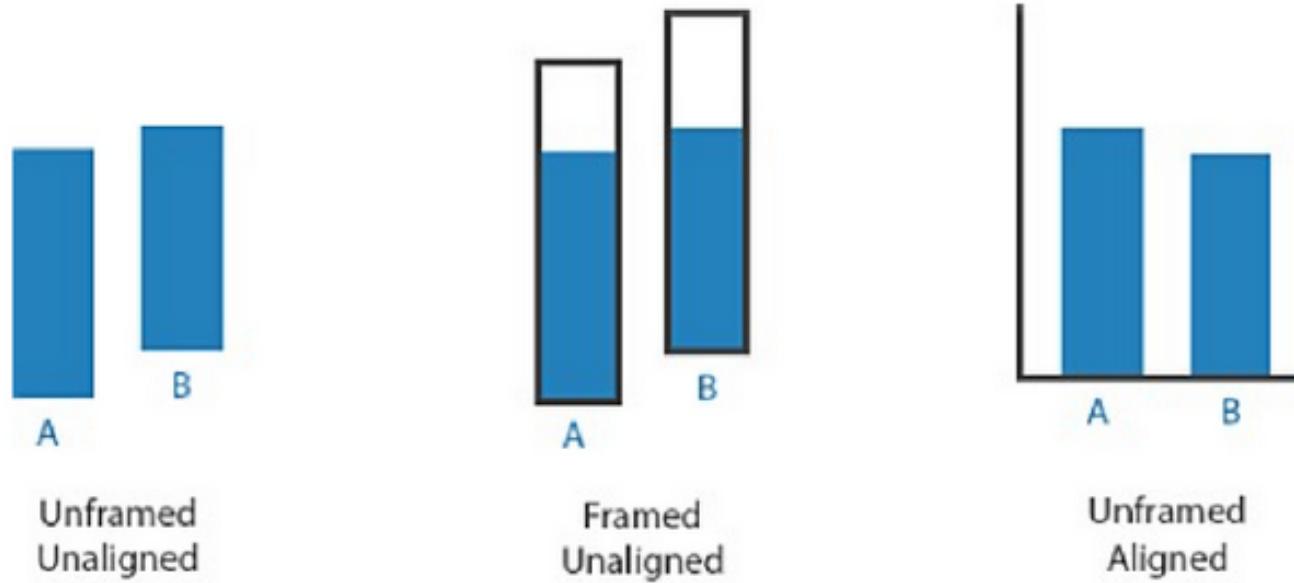
$$25 \text{ vs. } 26\text{cm} = 2.5 \text{ vs. } 2.6\text{cm}$$



Judgments in perception: Weber's Law

Detecting change is proportional to the relative, not the absolute change

$$25 \text{ vs. } 26\text{cm} = 2.5 \text{ vs. } 2.6\text{cm}$$



Errors in perception: Steven's Law

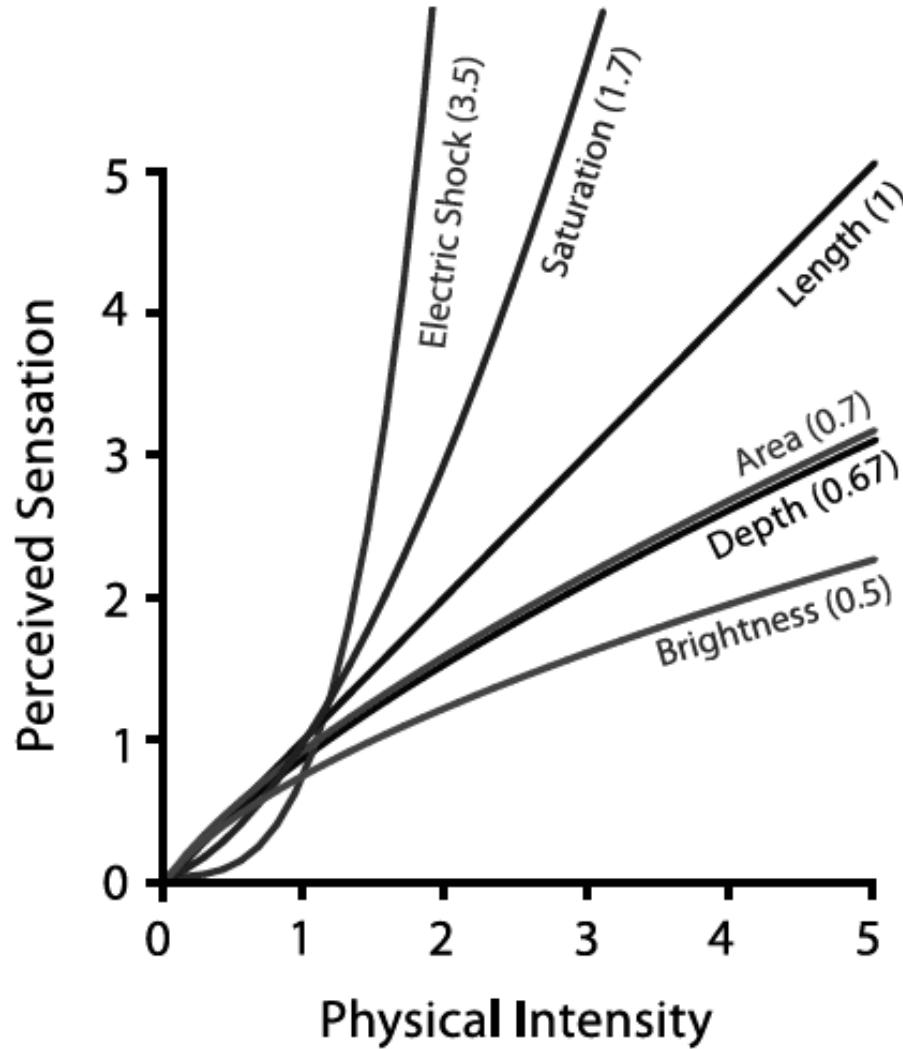
Perceived sensation S is proportional to the stimulus intensity I raised to a power of n

$$\text{Perceived scale} = (\text{Actual scale})^n$$

Continuum	Exponent (approx.)
Colour saturation	1.7
Length	1.0
Area	0.7
Volume	0.5
Brightness	0.5

Errors in perception: Steven's Law

Image: Munzner (2015)



Errors in perception, Cleveland and McGill (1985)

More accurate



position



length



angle



slope



area



Don't use areas (or volume) to show one dimensional data

volume



{# of information carrying dimensions \leq # of data dimensions}

density



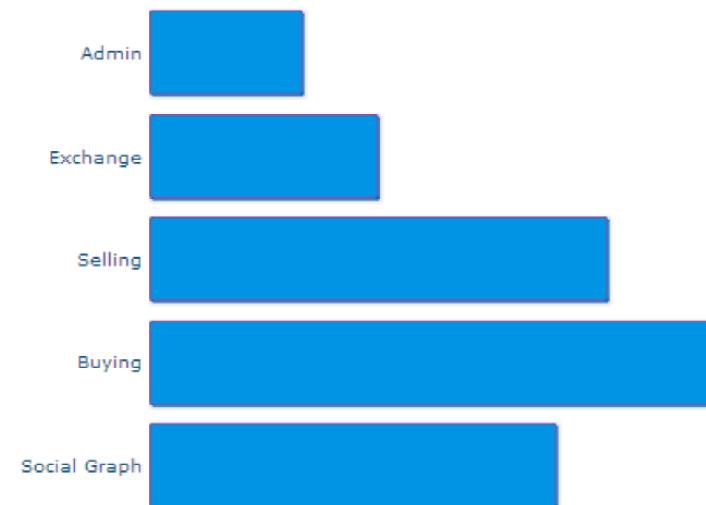
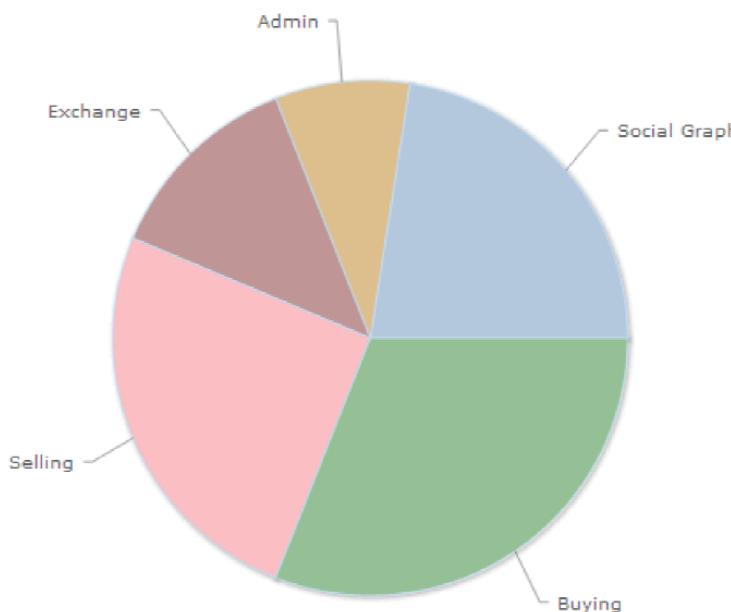
colour



Less accurate

Errors in perception, Cleveland and McGill (1985)

- *Example: angles vs lengths*



VALUE MESSAGE

Can your message be effectively communicated using your chosen visual variables?

	Numerical	Ordinal	Nominal
Position	Good	Good	Good
Size	Good	Good	Good
Pattern	Moderate	Moderate	Good
Color hue	Moderate	Moderate	Good
Color saturation	Moderate	Good	Poor
Color value	Moderate	Good	Poor
Orientation	Poor	Good	Poor
Shape	Poor	Poor	Good

Another graphic: Suitability of visual variables

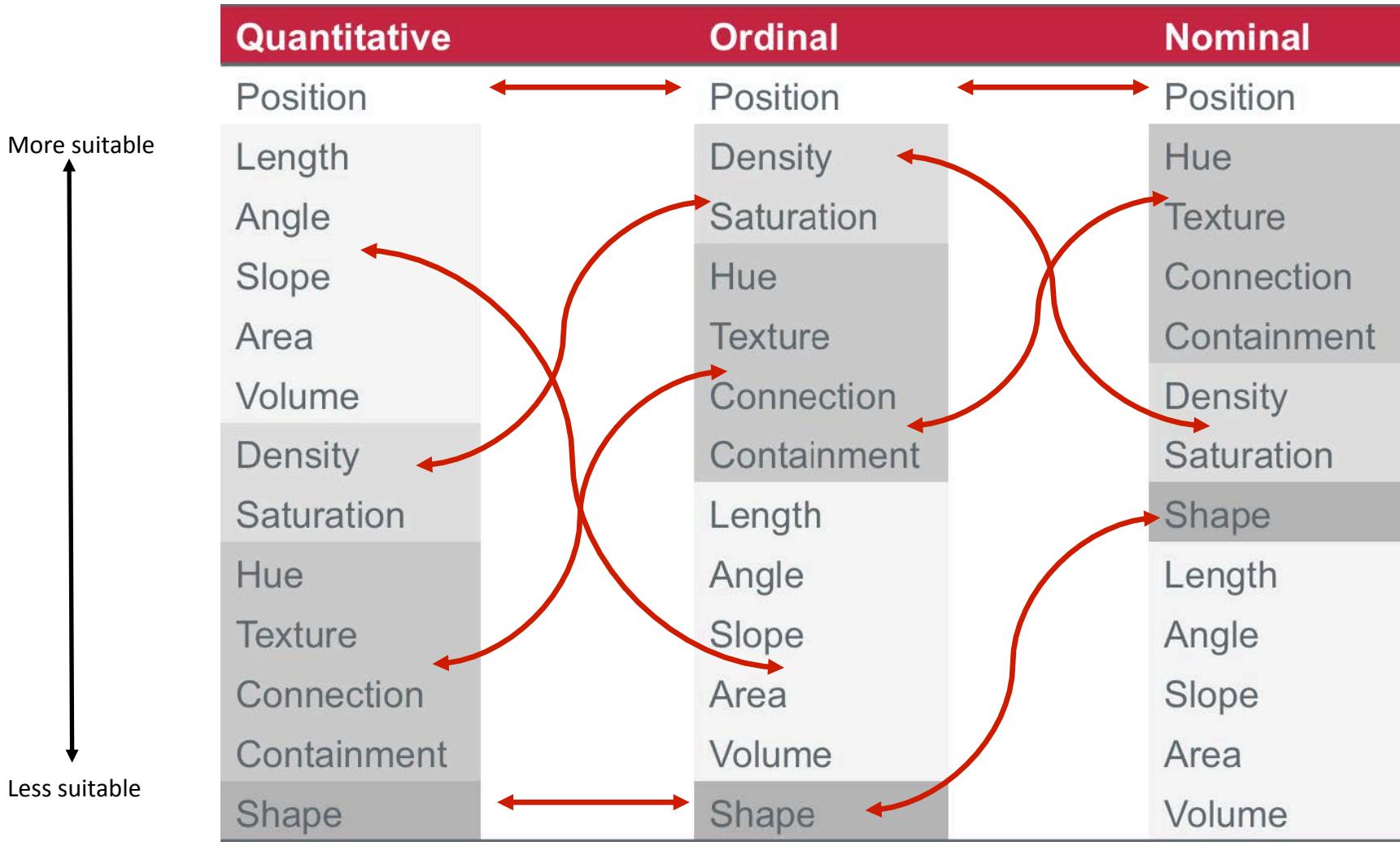


Image: Mackinlay (1986)

GESTALT LAWS

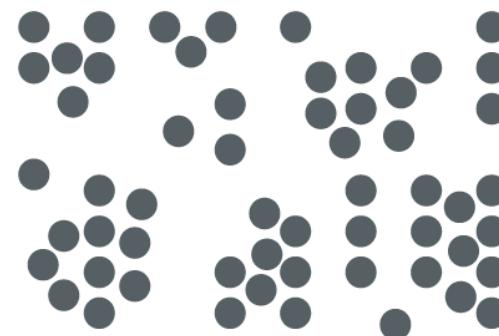
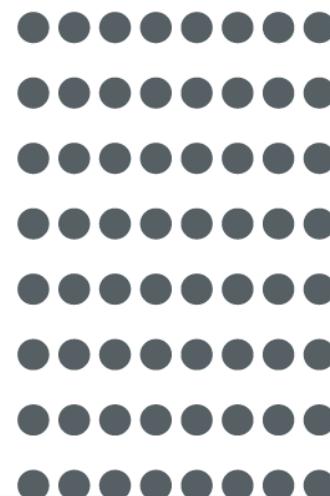
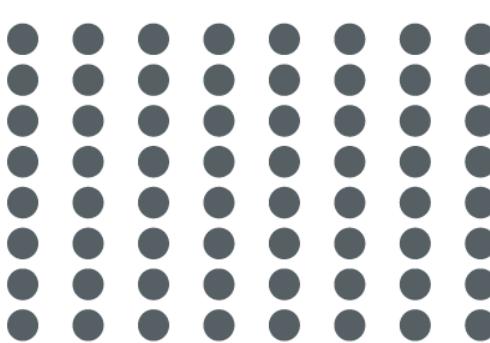
Gestalt principles describe how our mind organizes individual elements into groups. We can use these principles to highlight patterns that are important, and downplay other patterns.

- Visual system tries to structure what we see into patterns
- Gestalt is the interplay between the parts and the whole

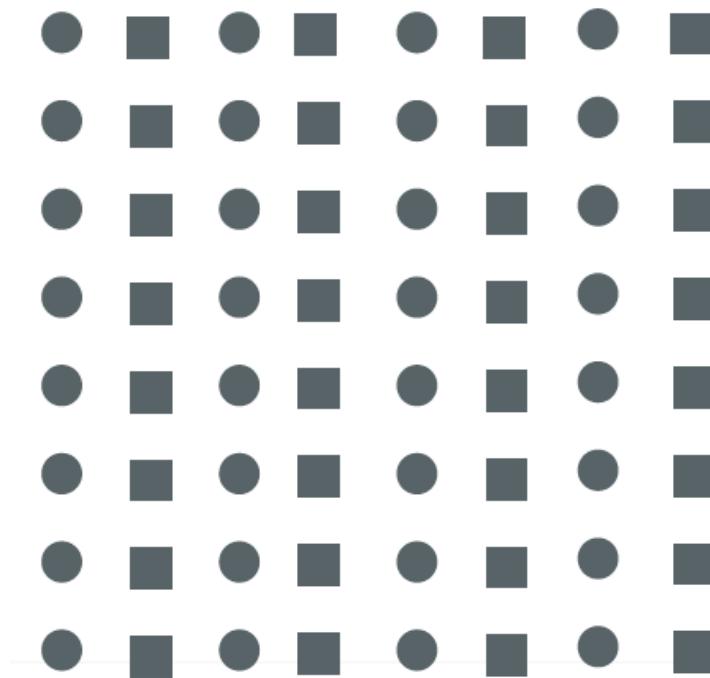
The Laws:

- Proximity
- Similarity
- Closure
- Connection
- Continuity
- Familiarity
- Symmetry
- Figure & Ground

Proximity

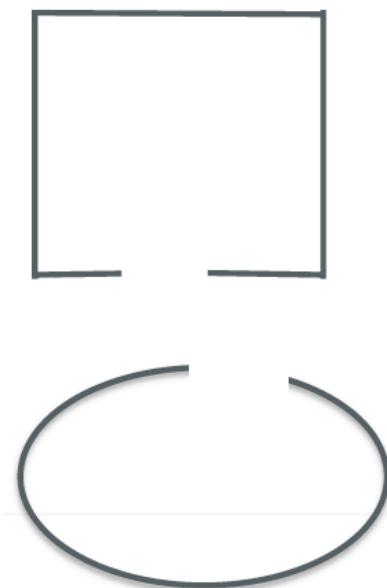


Similarity



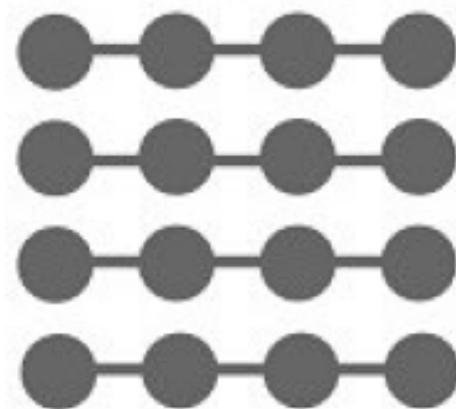


Closure



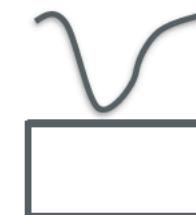
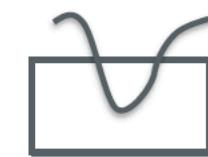
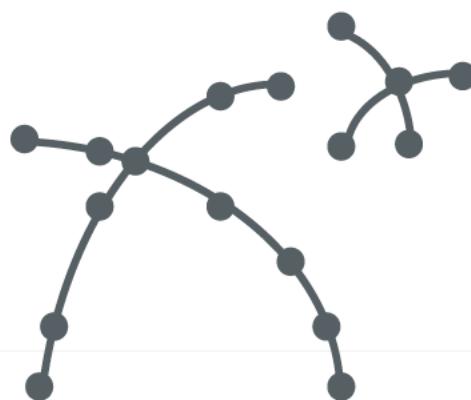
Illusory

Connectivity





Continuity

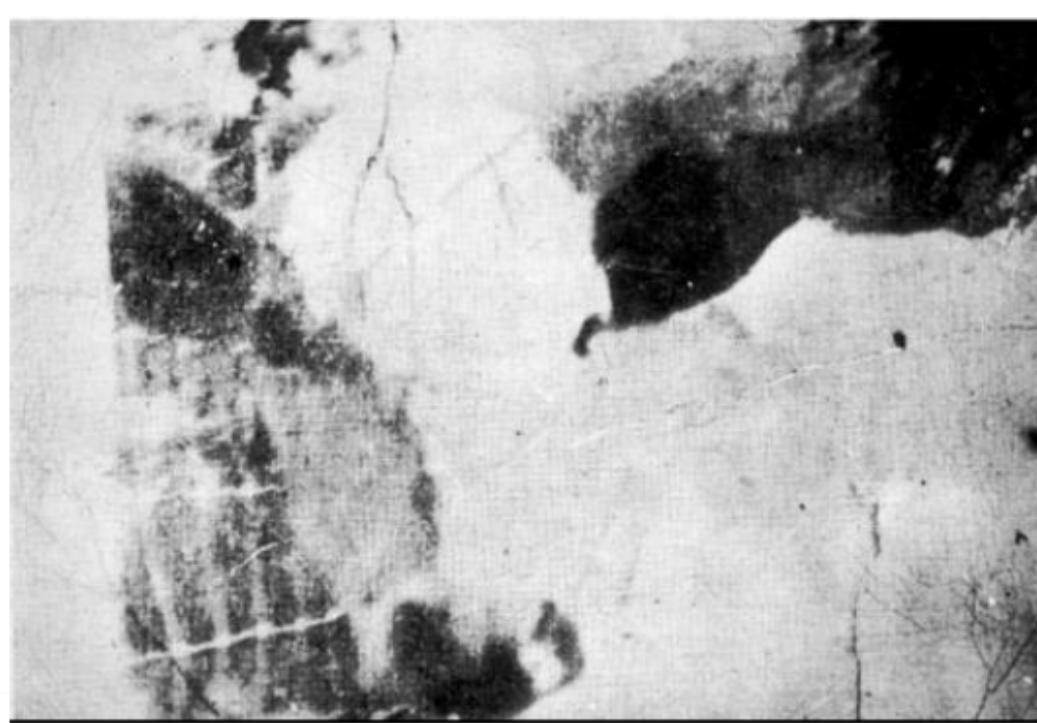


✓



✗

Familiarity



Symmetry

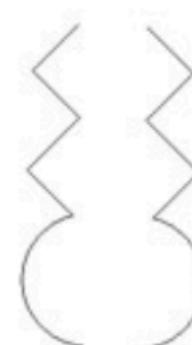
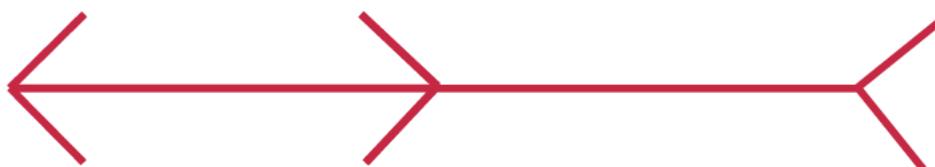


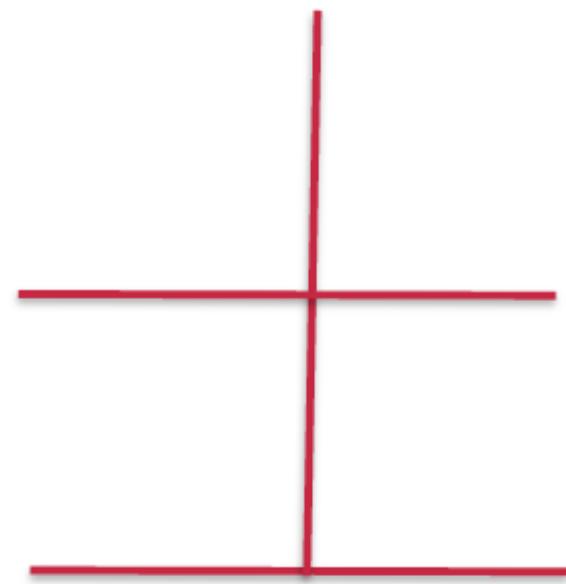
Figure & Ground



AMBIGUOUS INFORMATION

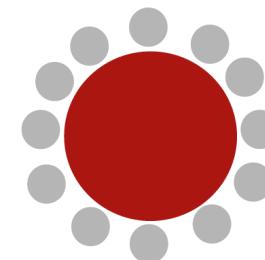
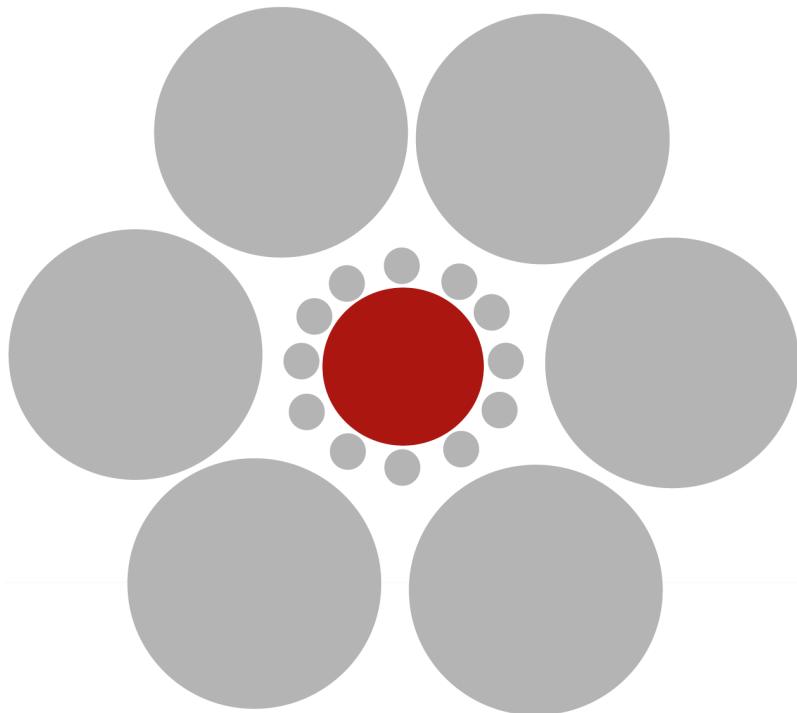


Length



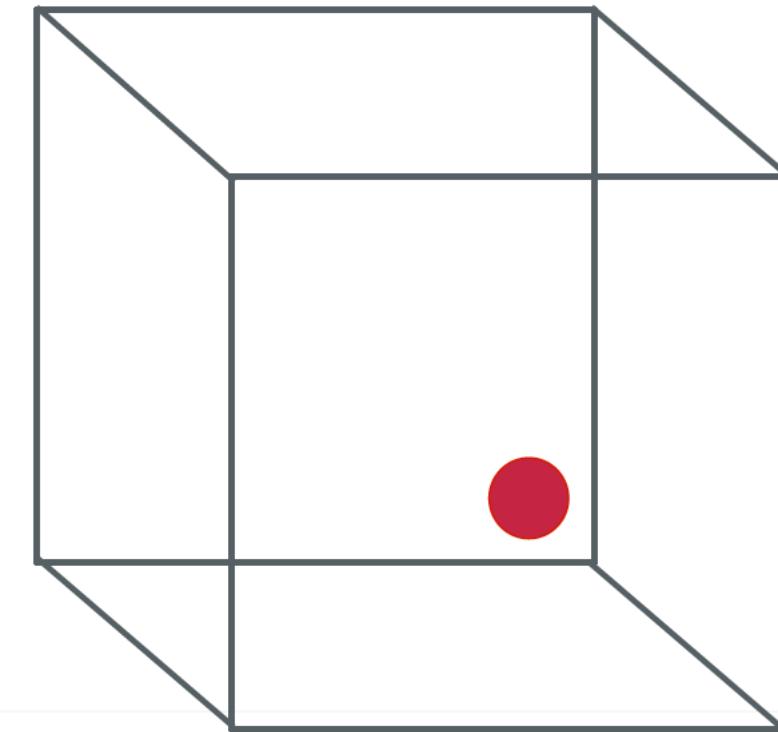
H-V illusion

AMBIGUOUS INFORMATION (Context)



Area

AMBIGUOUS INFORMATION



Position in 2D space

NEXT

- Data graphics 2

READING

- Munzner, T. (2015) Chapter 5: Marks and channels. In *Visualization Analysis and Design*. CRC Press, Boca Raton, FL, USA.

Access:

<http://www.crcnetbase.com.ezp.lib.unimelb.edu.au/doi/pdfplus/10.1201/b17511-6>