Visualisation

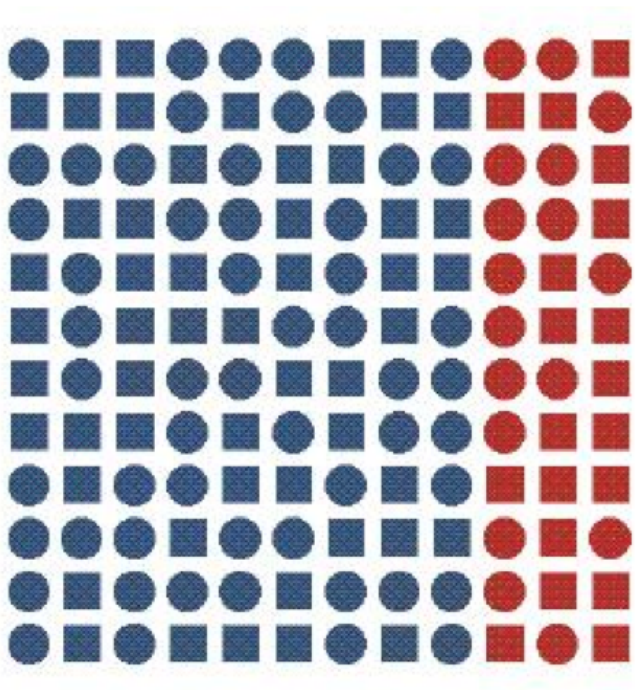
info-20002: foundations of informatics

Why Visualisation

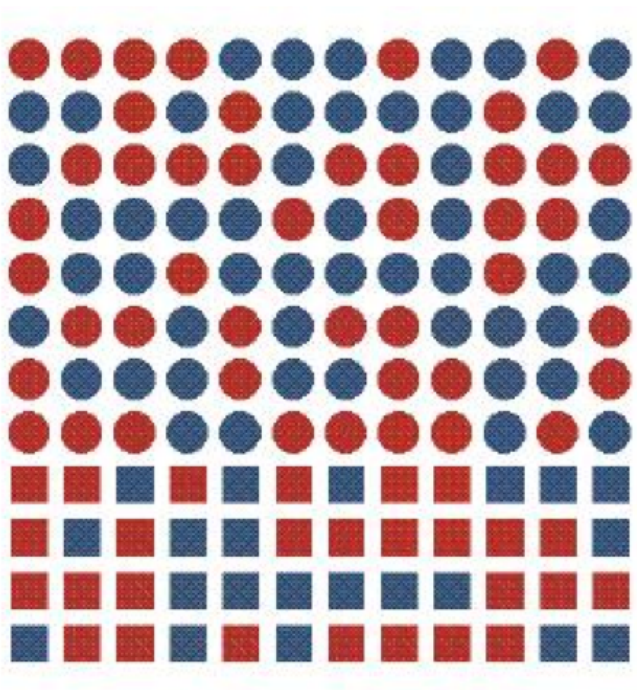
Pre-attentive Attention

Preattentive processing of visual information is performed automatically on the entire visual field detecting basic features of objects in the display. Such basic features include colors, closure, line ends, contrast, tilt, curvature and size. These simple features are extracted from the visual display in the preattentive system and later joined in the focused attention system into coherent objects. Preattentive processing is done quickly, effortlessly and in parallel without any attention being focused on the display. [Treisman, 1985; Treisman, 1986]

Treisman, A. (1985). **Preattentive Processing in Vision**. Computer Vision, Graphics, and Image Processing (31):2, pp. 157–177



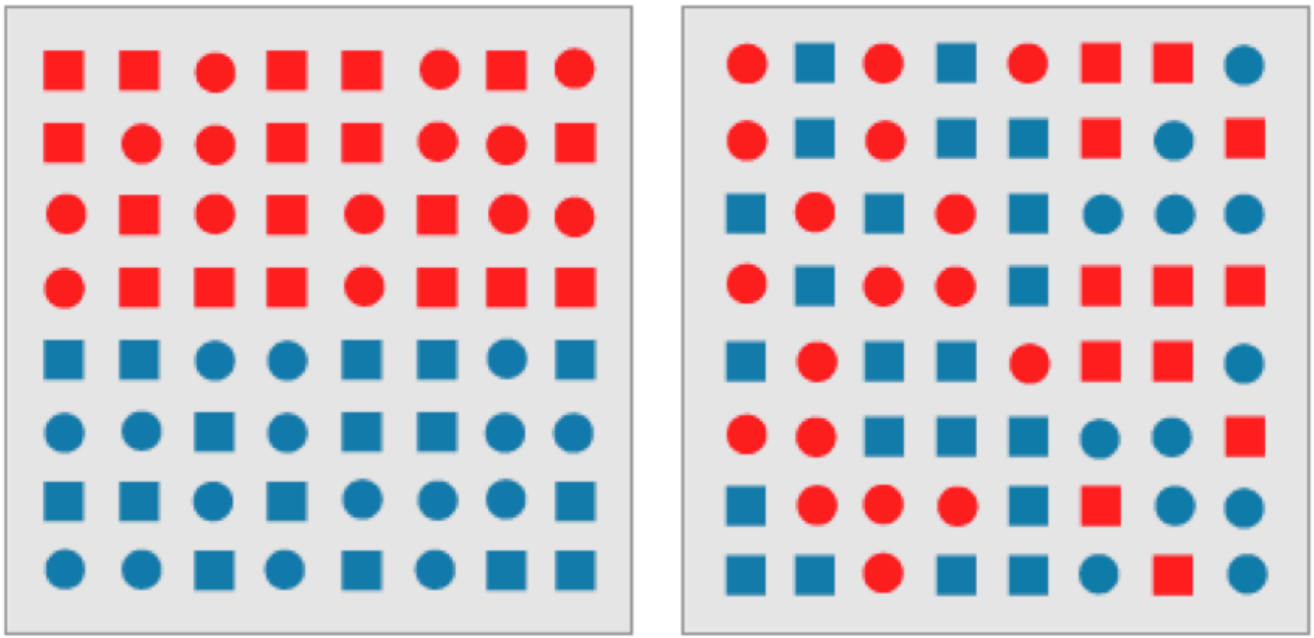
(a)



(b)

Treisman, A. (1985). [Preattentive Processing in Vision](). Computer Vision, Graphics, and Image Processing 31(2): 157–177

Why Visualisation?



(a) (b)

An example of a boundary detection from Treisman's experiments: (a) a boundary defined by a unique feature hue (red circles and red squares on the top, blue circles and blue squares on the bottom) is preattentively classified as horizontal; (b) a boundary defined by a conjunction of features (red circles and blue squares on the left, blue circles and red squares on the right) cannot be preattentively classified as vertical.

Treisman, A. (1985). **Preattentive Processing in Vision**. Computer Vision, Graphics, and Image Processing (31):2, pp. 157–177

Which city is the hottest?

city/month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Melbourne	41.2	35.5	37.4	29.3	23.9	16.8	18.2	25.7	22.3	33.5	36.9	41.1
Brisbane	31.3	40.2	37.9	29	30	26.7	26.7	28.8	31.2	34.1	31.1	31.2
Darwin	34	34	33.2	34.5	34.8	33.9	32	34.3	36.1	35.4	37	35.5
Perth	41.9	41.5	42.4	36	26.9	24.5	23.8	24.3	27.6	30.7	39.8	44.2
Adelaide	42.1	38.1	39.7	33.5	26.3	16.5	21.4	30.4	30.2	34.9	37.1	42.2
Canberra	35.8	29.6	35.1	26.5	22.4	15.3	15.7	21.9	22.1	30.8	33.4	35
Hobart	35.5	34.1	30.7	26	20.9	15.1	17.5	21.7	20.9	24.2	30.1	33.4
Sydney	30.6	29	35.1	27.1	28.6	20.7	23.4	27.7	28.6	34.8	26.4	30.2

Which city is the hottest?

city/month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Melbourne	41.2	35.5	37.4	29.3	23.9	16.8	18.2	25.7	22.3	33.5	36.9	41.1
Brisbane	31.3	40.2	37.9	29	30	26.7	26.7	28.8	31.2	34.1	31.1	31.2
Darwin	34	34	33.2	34.5	34.8	33.9	32	34.3	36.1	35.4	37	35.5
Perth	41.9	41.5	42.4	36	26.9	24.5	23.8	24.3	27.6	30.7	39.8	44.2
Adelaide	42.1	38.1	39.7	33.5	26.3	16.5	21.4	30.4	30.2	34.9	37.1	42.2
Canberra	35.8	29.6	35.1	26.5	22.4	15.3	15.7	21.9	22.1	30.8	33.4	35
Hobart	35.5	34.1	30.7	26	20.9	15.1	17.5	21.7	20.9	24.2	30.1	33.4
Sydney	30.6	29	35.1	27.1	28.6	20.7	23.4	27.7	28.6	34.8	26.4	30.2

Why Visualisation?

- Four datasets that have nearly identical simple statistical properties
- Descriptive Statistics of Anscombe's Quartet Data Table

Property	**Value**
Mean of x of each data set	9 (exact)
Variance of x in each data set	11 (exact)
Mean of y in each data set	7.50 (to 2 decimal places)
Variance of y in each data set	4.122 or 4.127 (to 3 decimal places)
Correlation between x and y in each data set	0.816 (to 3 decimal places)
Linear regression line for each data set	y = 3.00 + 0.500x (to 2 and 3 decimal places, respectively)

Anscombe, F. J. (1973). **"Graphs in Statistical Analysis"**. American Statistician 27(1): 17–21.

Why Visualisation?

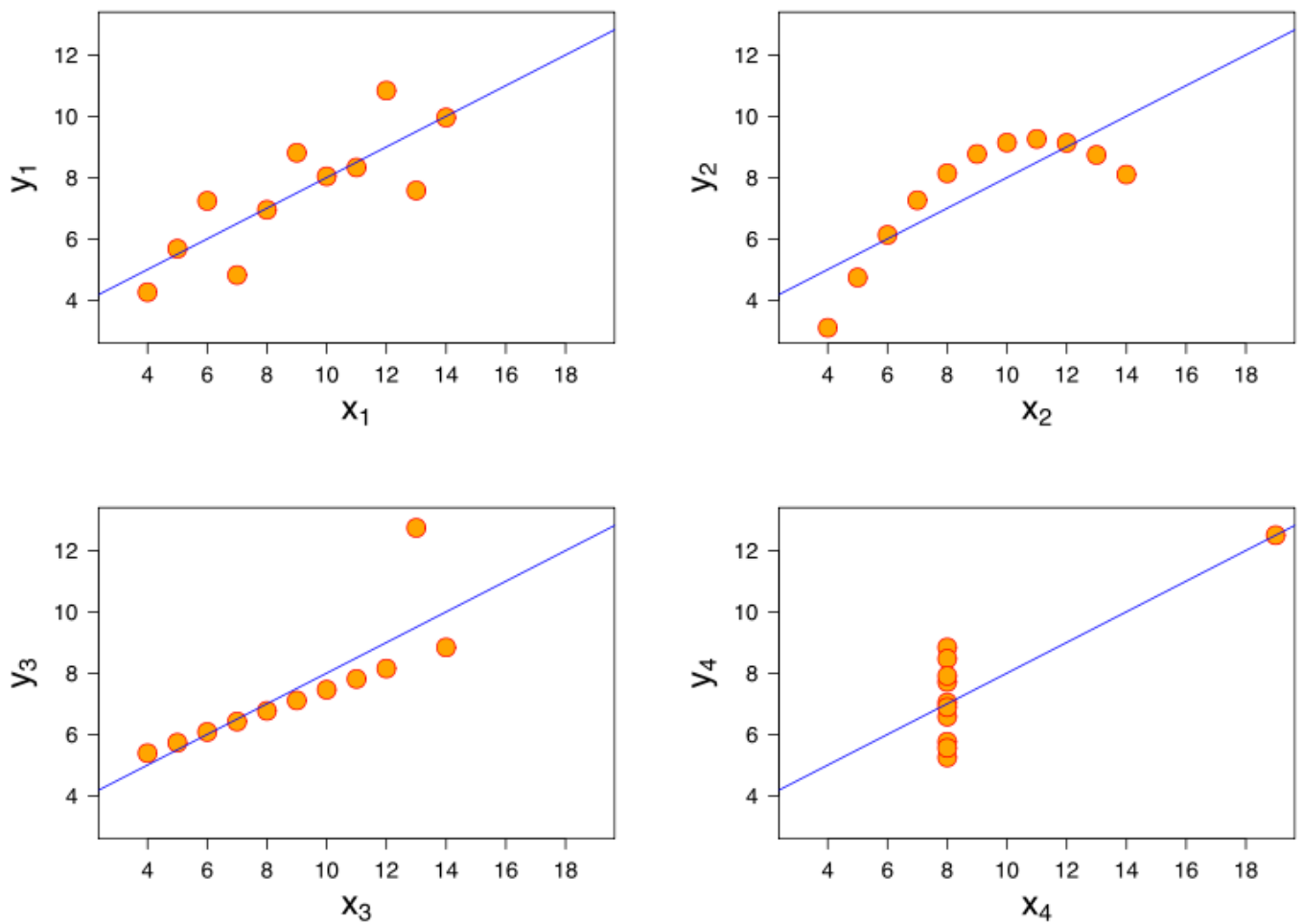
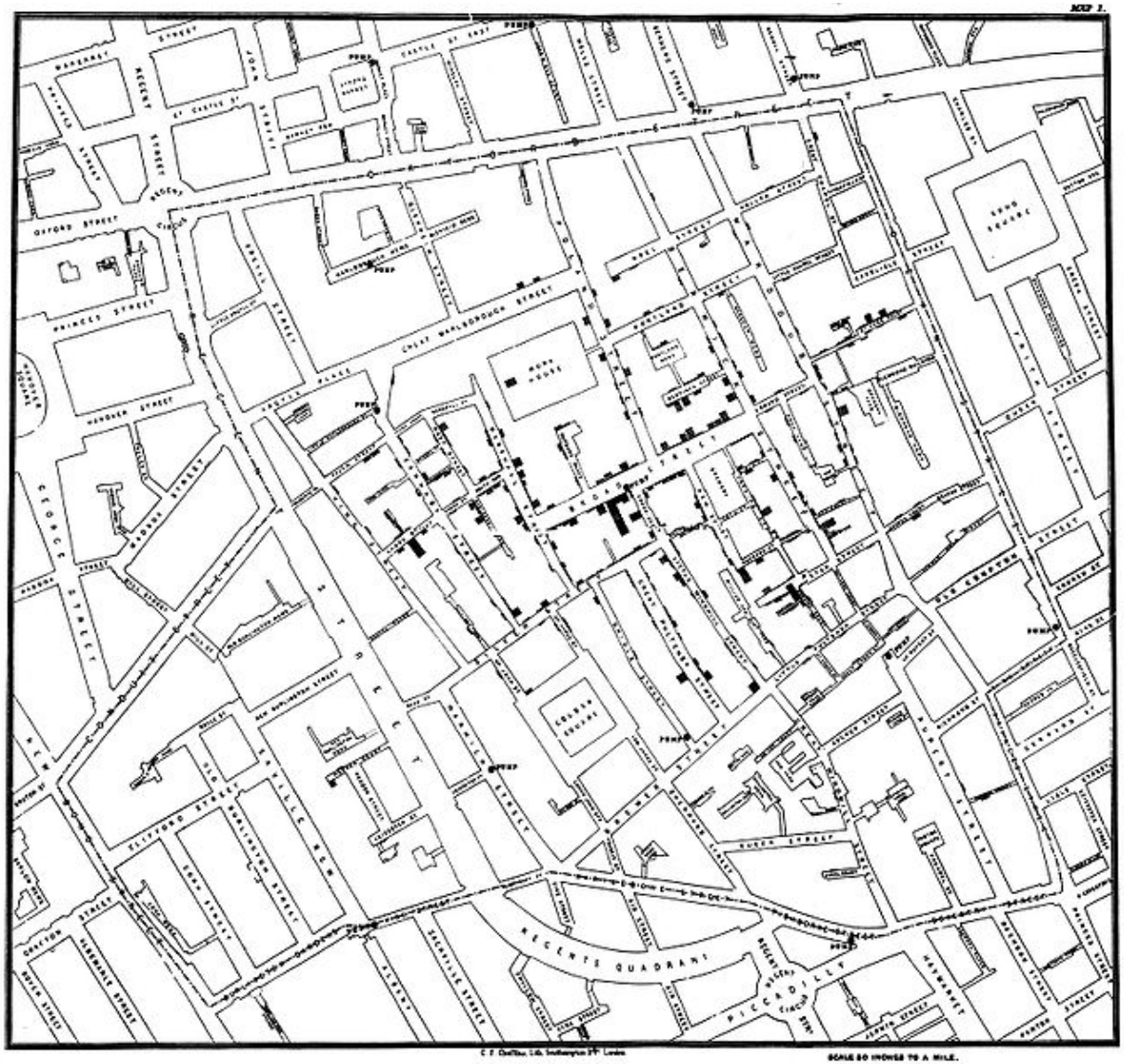


image source: Wikimedia Commons.

Visualization is

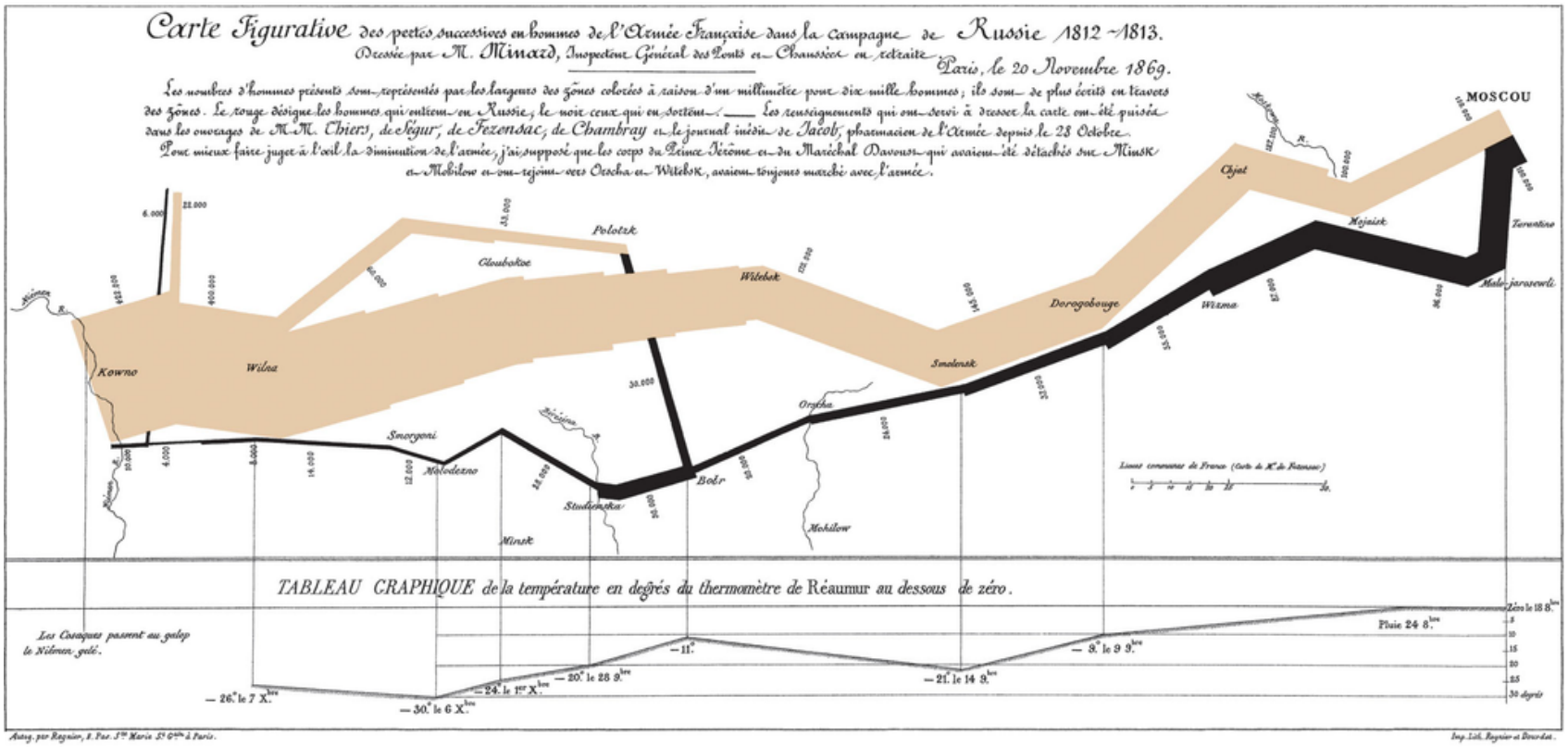
- Way of seeing meaning/relationships in data – prosthetic to our mental models
- Visualisations amplify cognition
- Visualisations provide a space for exploring data
- **Exploratory and explanatory**
  - “discover the unexpected; describe and explain the expected” (National Visualization and Analytics Center, Pacific Northwest Laboratory)

John Snow's Dot Distribution Map of Broad Street Cholera



Visualization in story telling

Minard's map of Napoleon's 1812 Russian Campagin



Visualization is

a mapping from *data variables* to *visual variables*

Examples:

- Temperature → brightness of colour
- City → hue of colour
- City → location (x coordinate)

Data variables

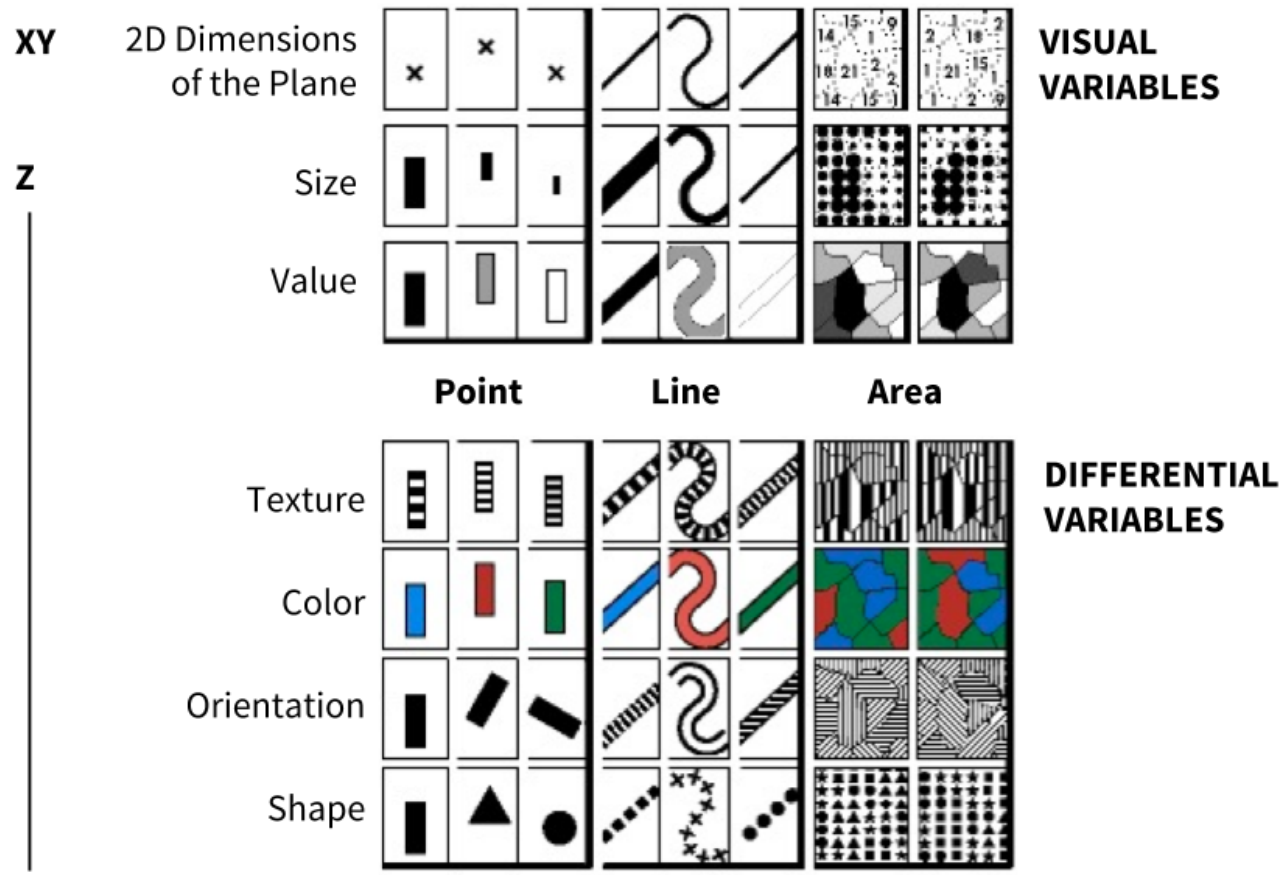
- Continuous and discrete variables
  - **Continuous**
    Numerical or quantitative (Mackinlay, 2086)
    Numbers or anything that can have a range
  - **Discrete**
    Anything else
- Discrete/categorical
  - **Ordinal**
    Ordered/ranked set
    Example: (N, P, H3, H2B, H2A, H1)
  - **Nominal**
    Unordered set
    Examples: (sydney,melbourne,perth,...) and (male,female)

Data variables - Steven's Types (Scales)

- **Nominal scale**
- **Ordinal scale**
- **Interval scale**
  - Has no meaningful zero point (no origin)
  - Express difference (interval) between two measurements
  - Examples: dates (AD), latitude
  - Celcius temperature scale
  - To say that "2°C temperature is twice as hot as 1°C" is NOT meaningful
- **Ratio scale**
  - Has a meaningful, unique and non-arbitrary zero point.
  - Examples: mass, length, duration, angle.
  - Kelvin temperature scale contains absolute zero point.
  - It is meaningful to say that "a 2-metre object has "twice the length" of a 1-metre object"

Stevens, S. S. (1946). **"On the Theory of Scales of Measurement"**.
Science 103 (2684): 677–680.

Visual Variables



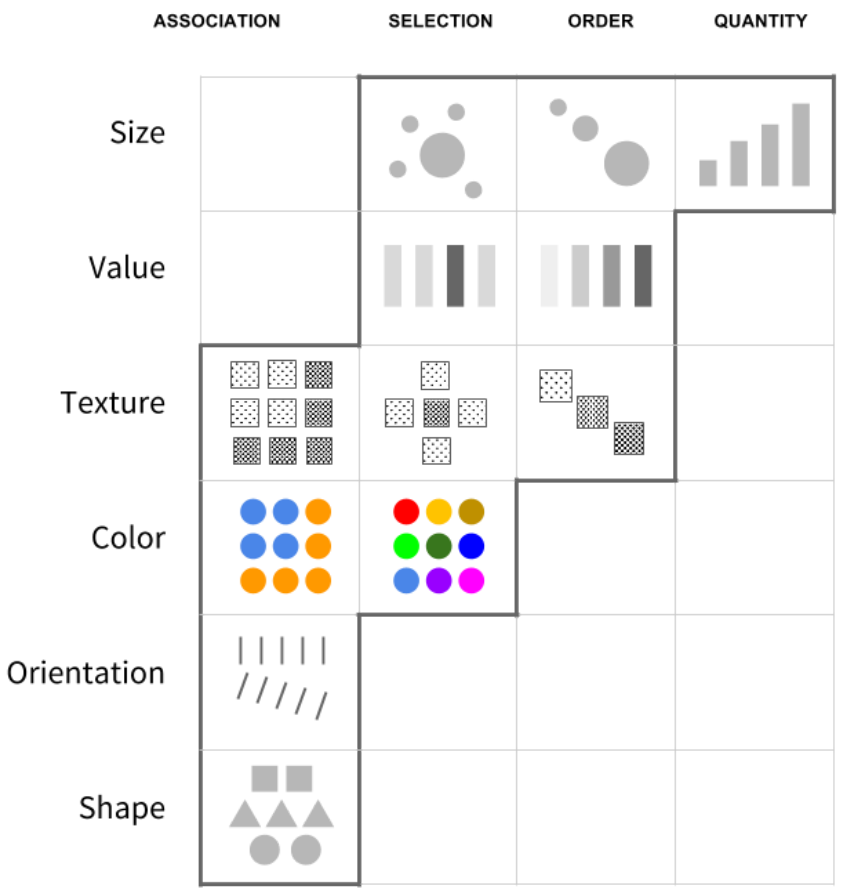
Bertin, J. (1967). **"Semiology of Graphics, Semiology of Graphics: Diagrams, Networks, Maps"**. 1983 edition. University of Wisconsin Press.

Functions of Visualisation

- **Associative function**
  grouping of all correspondences differentiated by a variable
  perceived as *similar*
- **Selective function**
  isolation of correspondences differentiated by a variable
  perceived as *different*, forming families
- **Ordering function**
  a spontaneous visual perception of rank
  perceived as *ordered*
- **Quantitative function**
  a visual difference between two categories expressed as a numerical ratio
  perceived as *proportional*

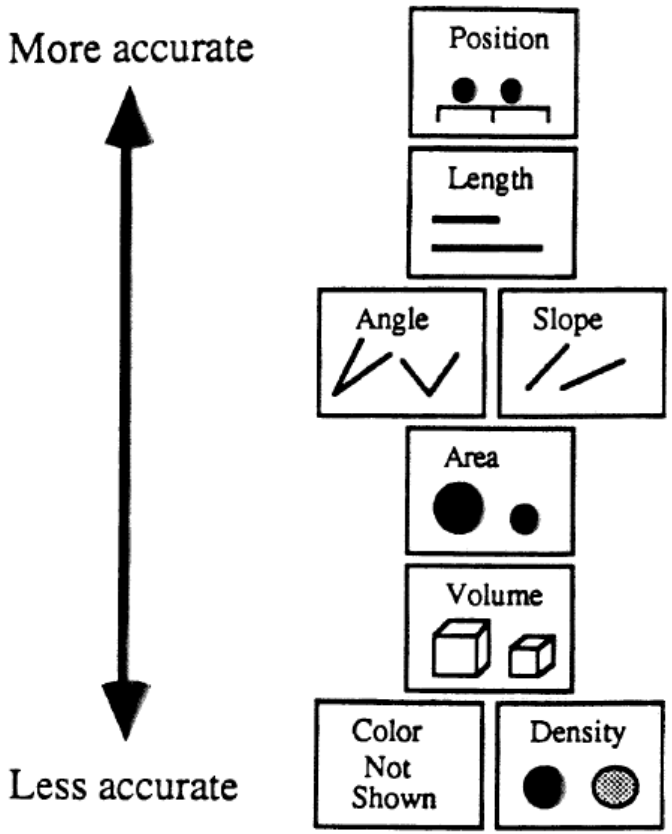
Bertin, J. (1967). **"Semiology of Graphics, Semiology of Graphics: Diagrams, Networks, Maps"**. 1983 edition. University of Wisconsin Press.

Variables vs Functions



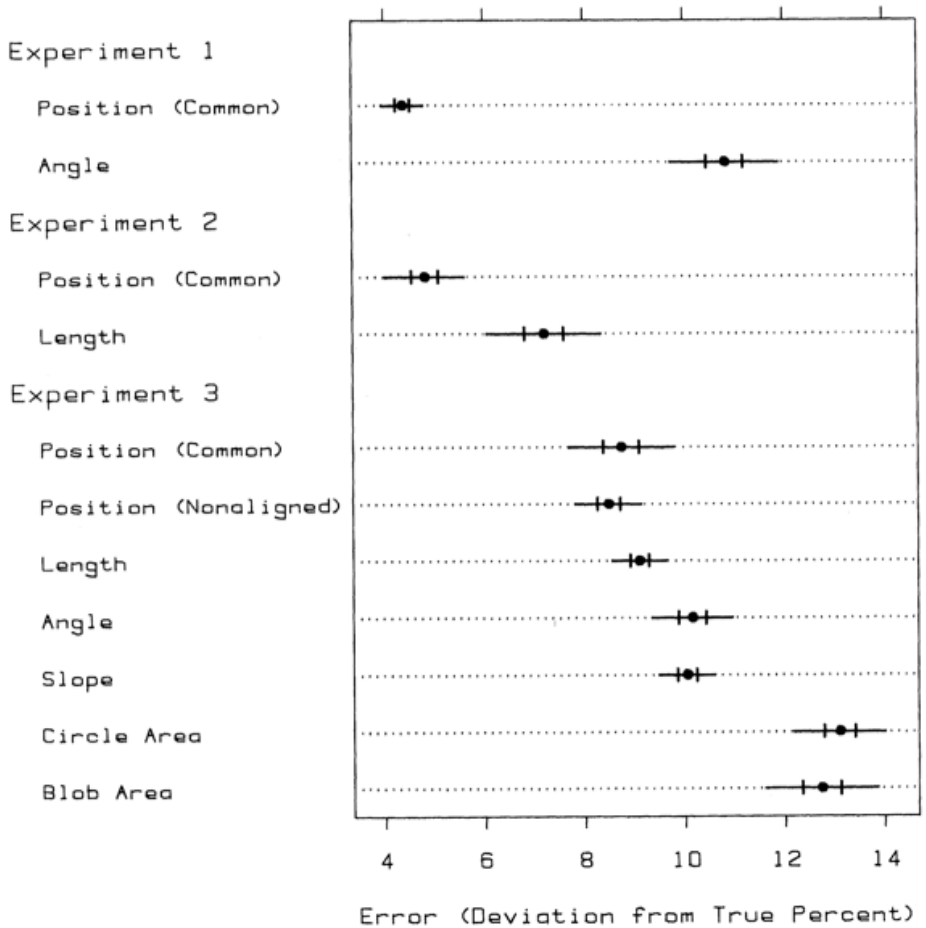
Bertin, J. (1967). **"Semiology of Graphics, Semiology of Graphics: Diagrams, Networks, Maps"**. 1983 edition. University of Wisconsin Press.

Accuracy of Visual Variables



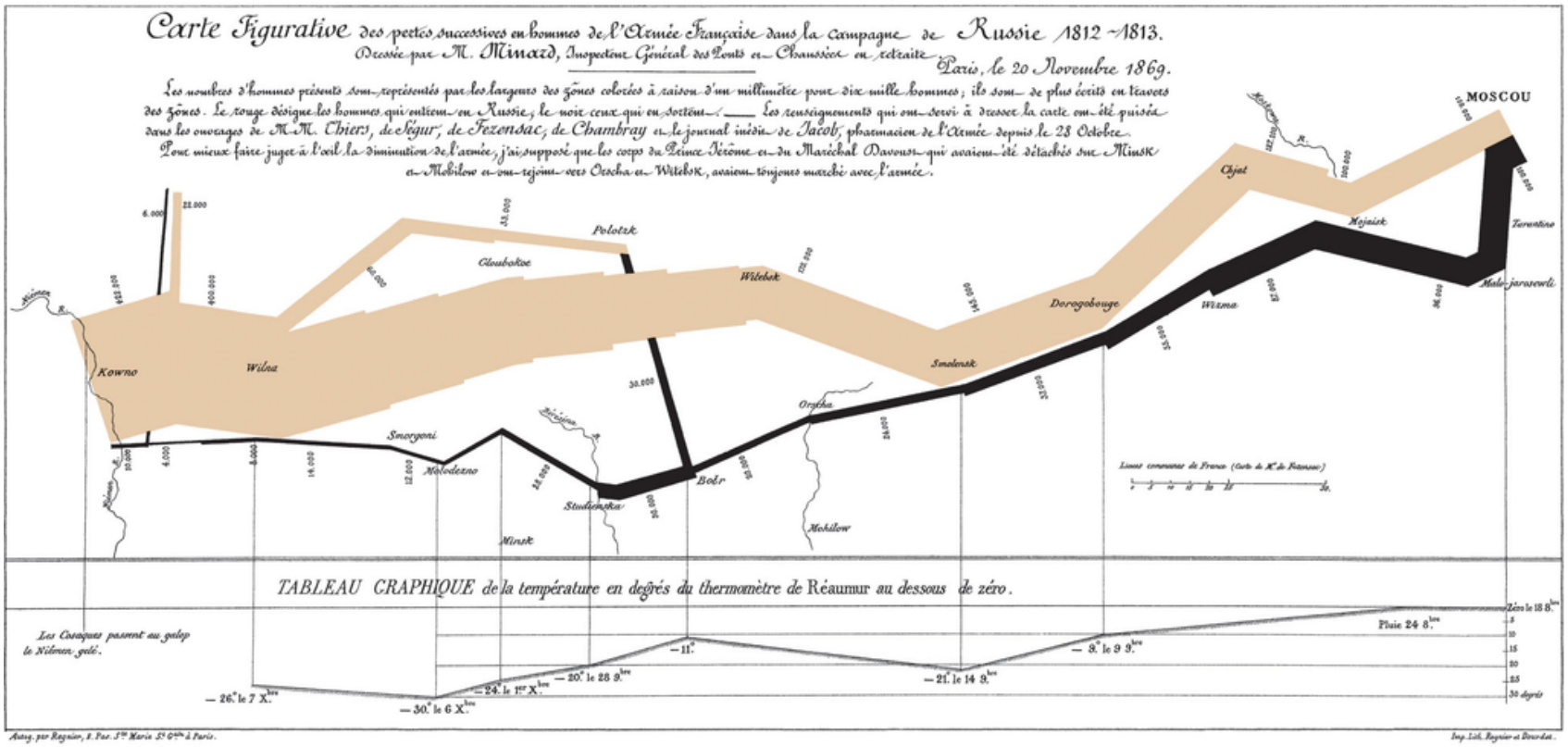
Mackinlay, J. (1986). **"Automating the design of graphical presentations of relational information"**. ACM Trans. Graph 5(2): 110-141.

Accuracy of Visual Variables

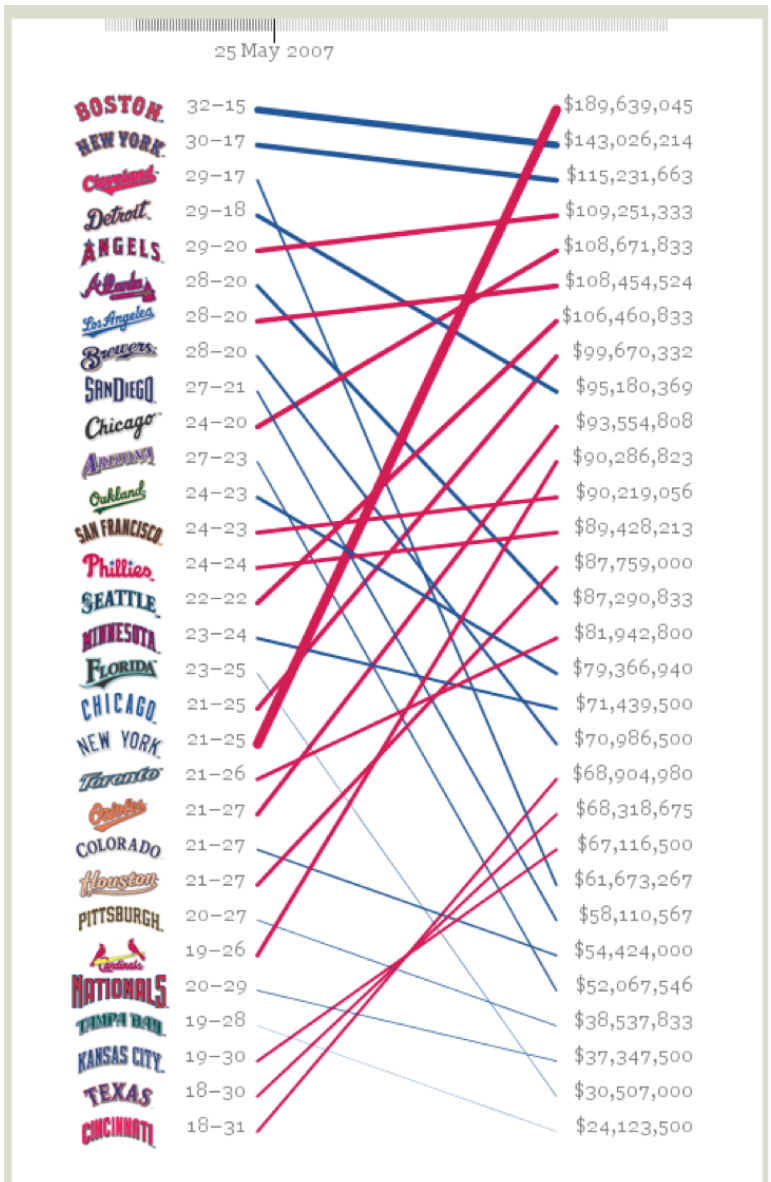


W. S. Cleveland and R. McGill (1984). **Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods"**. Journal of the American Statistical Association, 79:531-554.

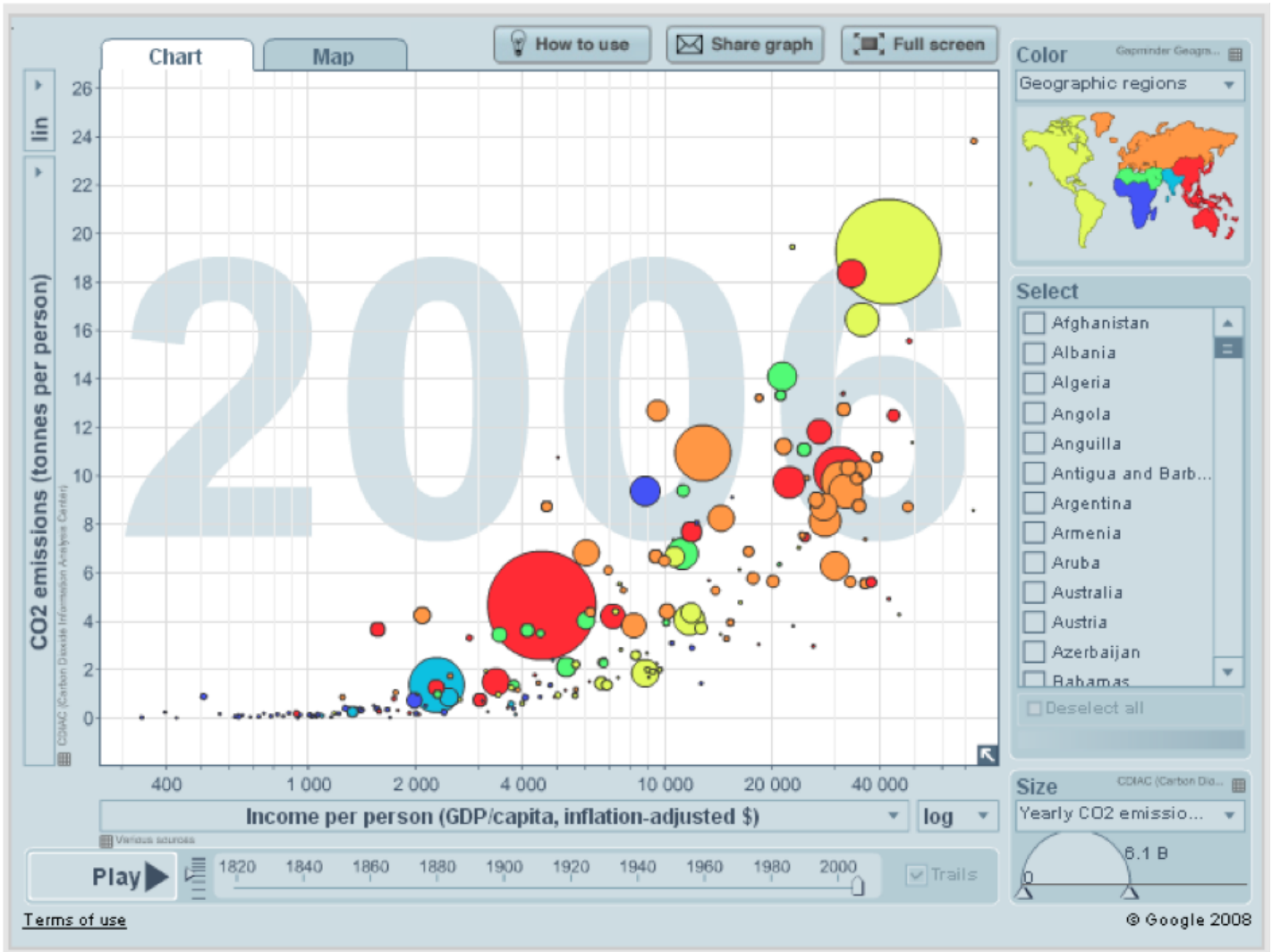
What are the data/visual variables?



What are the data/visual variables?



What are the data/visual variables?



http://gapminder.org