Information Visualisation

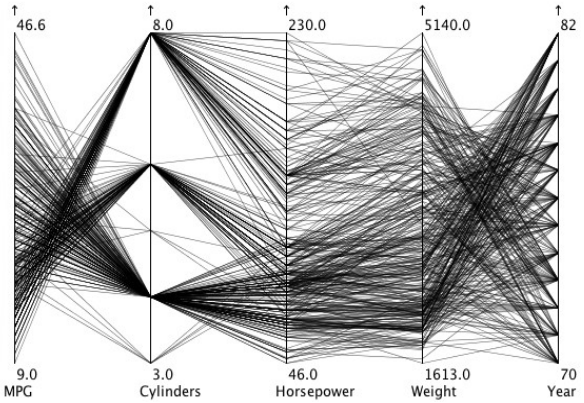
Computer Animation and Visualisation
Lecture 18

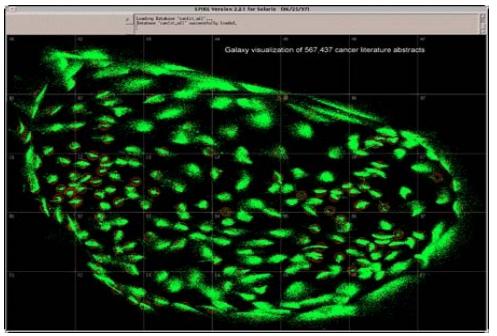
Taku Komura
tkomura@ed.ac.uk

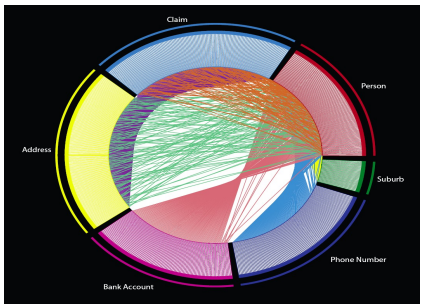
Institute for Perception, Action & Behaviour
School of Informatics

Overview

- **Information Visualisation**
  - **Univariate, bivariate, trivariate, multi-variate data**
  - **Relations visualized by lines, tree visualization**
  - **Document visualization**







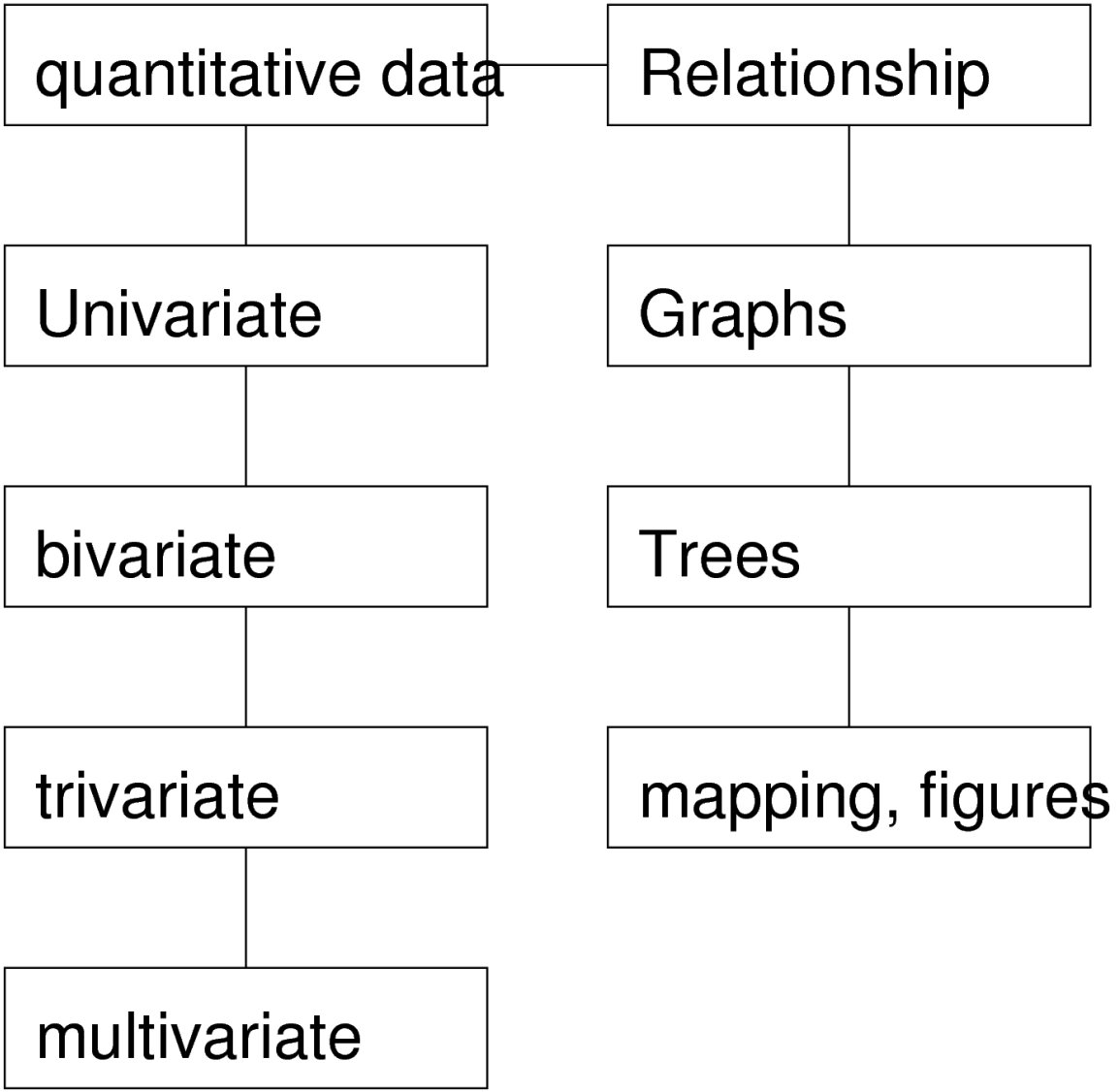
What is Information Visualisation ?

- Visualising ***discrete*** **data with no spatial information**

- Visualisation of important information contained in abstract data types
  - Needs to be intuitive
  - Such that people can easily and quickly understand

- *Tools for*
  - *Extraction of information from the data*
  - *Discovery of new knowledge*

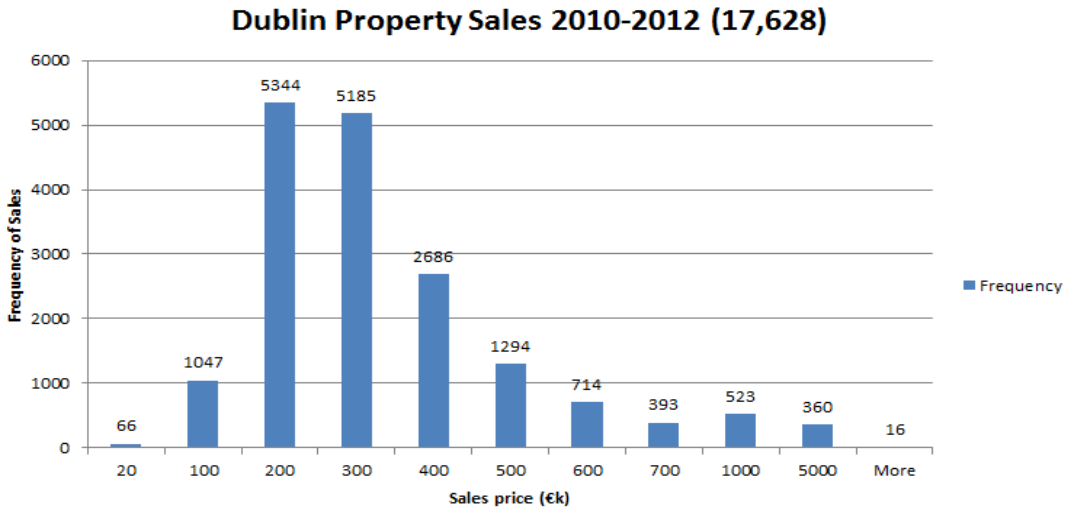
Data types



Document Visualization

Univariate Data, Bivariate Data

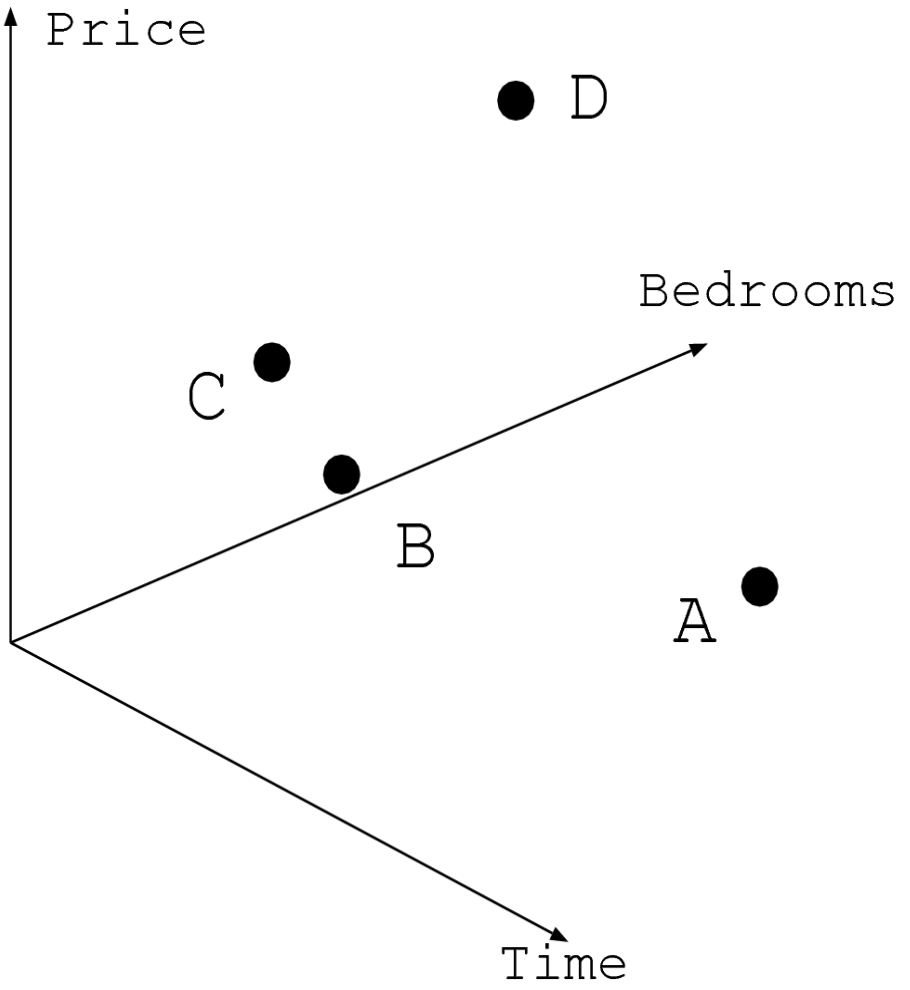
Can use scatter plots, histograms



Trivariate Data

Scatterplots

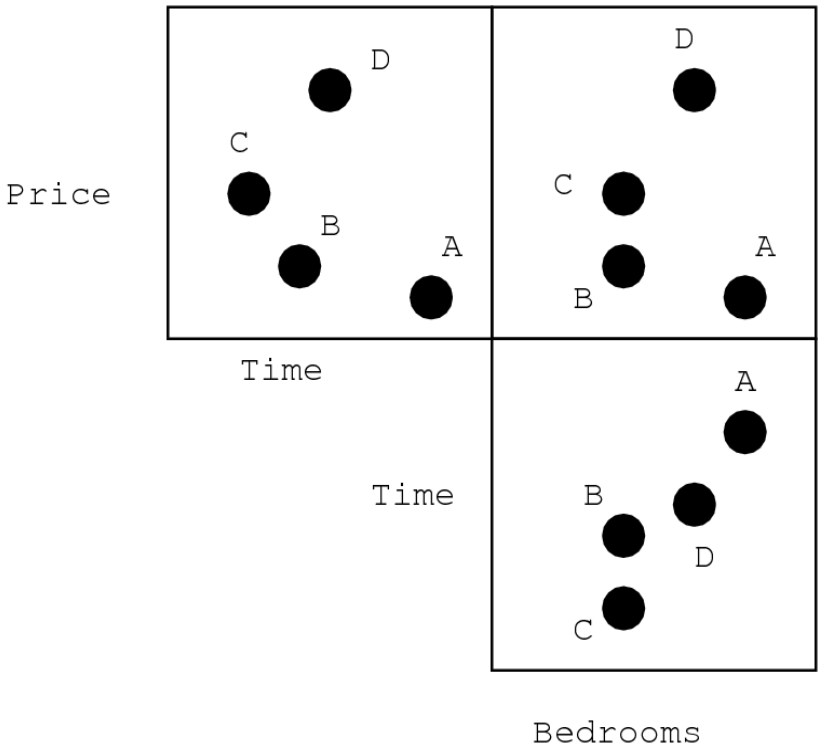
Scatterplot matrix

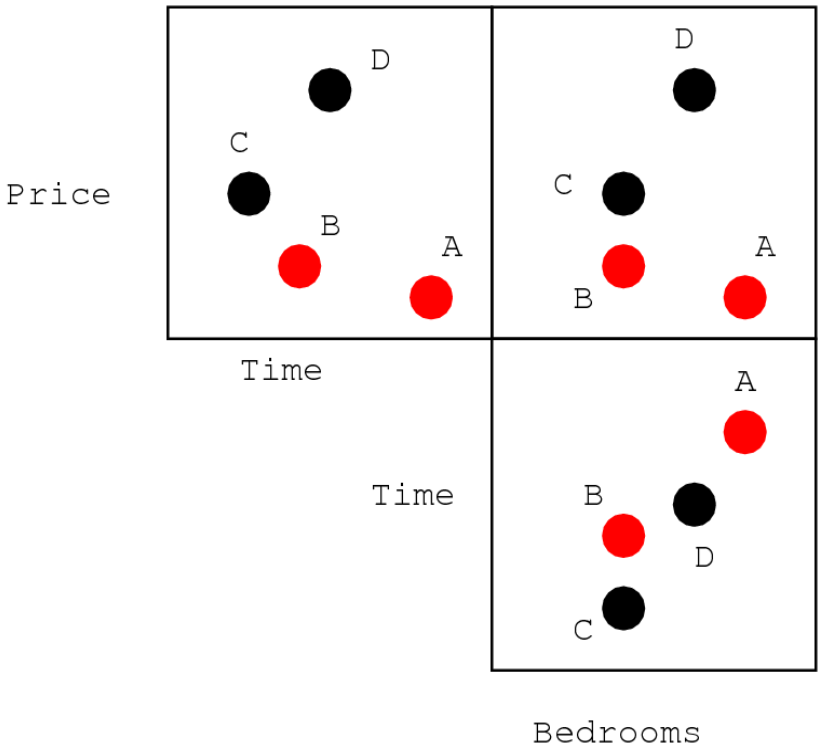


Not clear if D is more expensive
Than B and C

Trivariate Data

Scatterplot Matrix : Visualizing the relations of every two variables



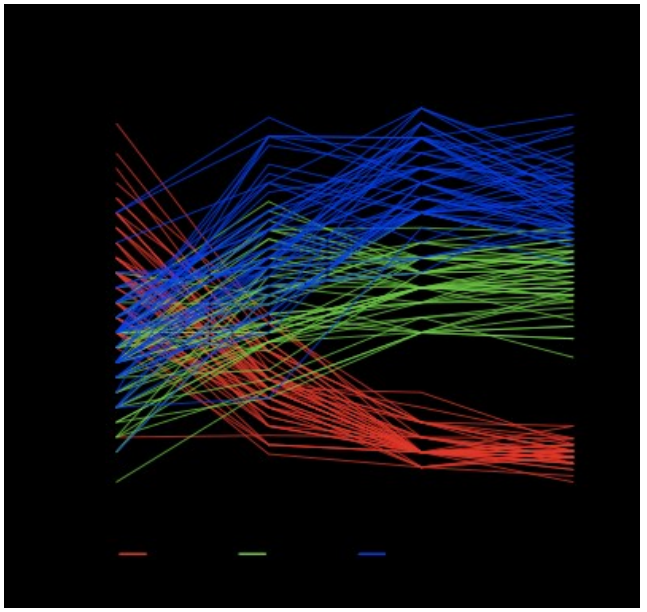


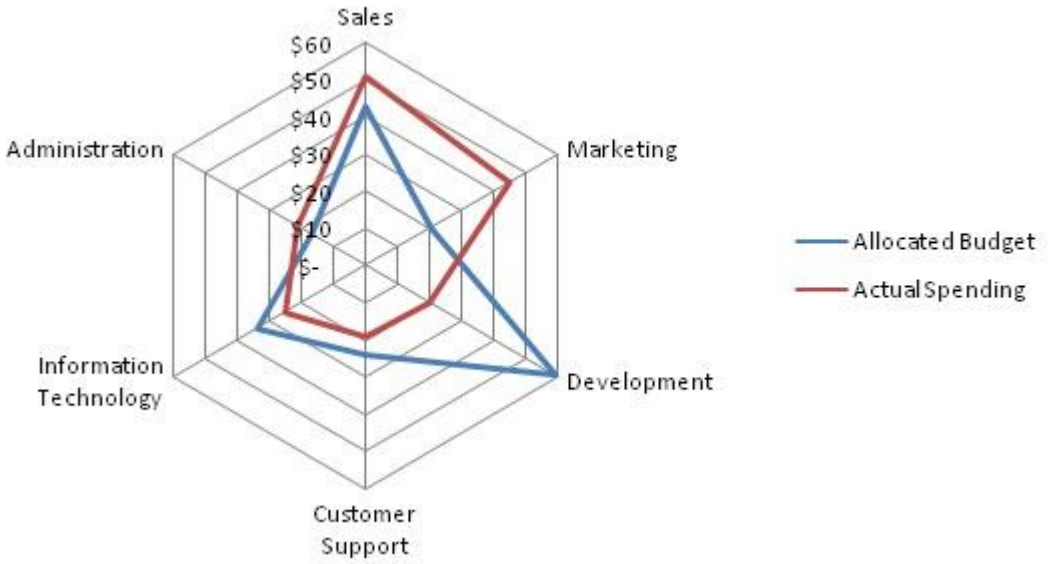
Multivariate Data

Parallel Coordinates

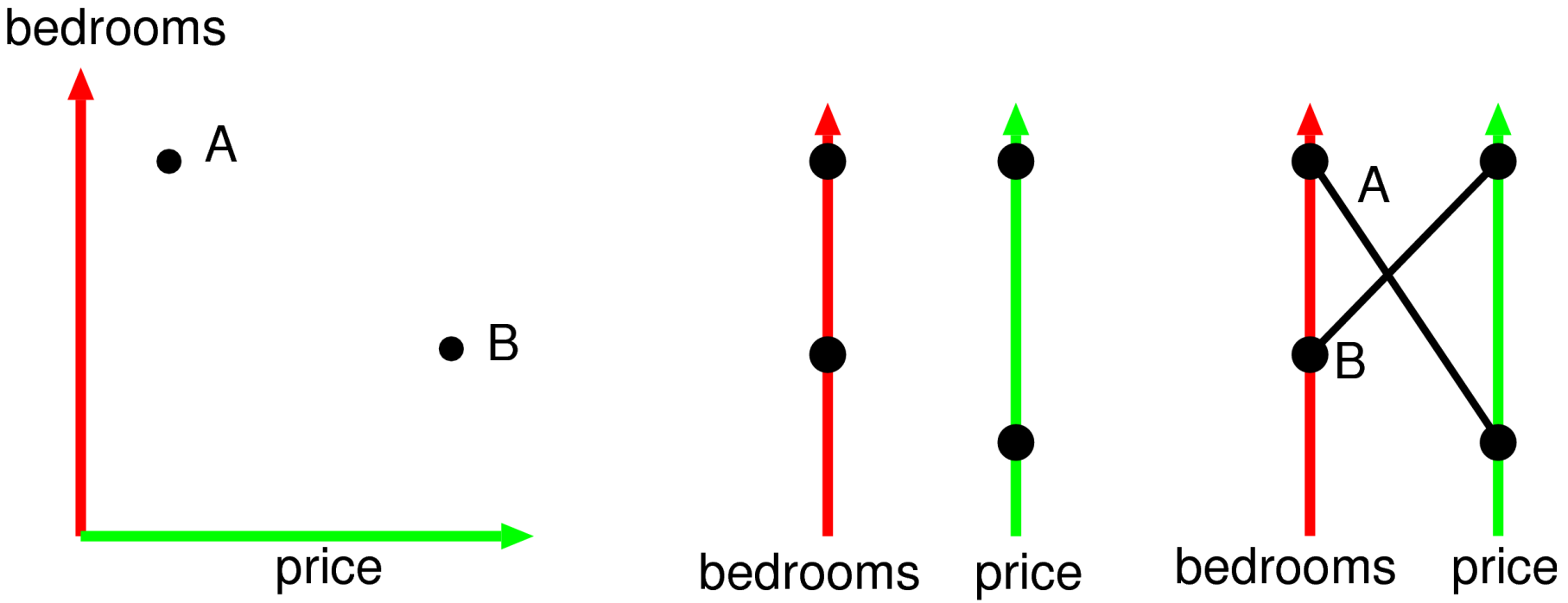
Star plots

Scattered plot matrix

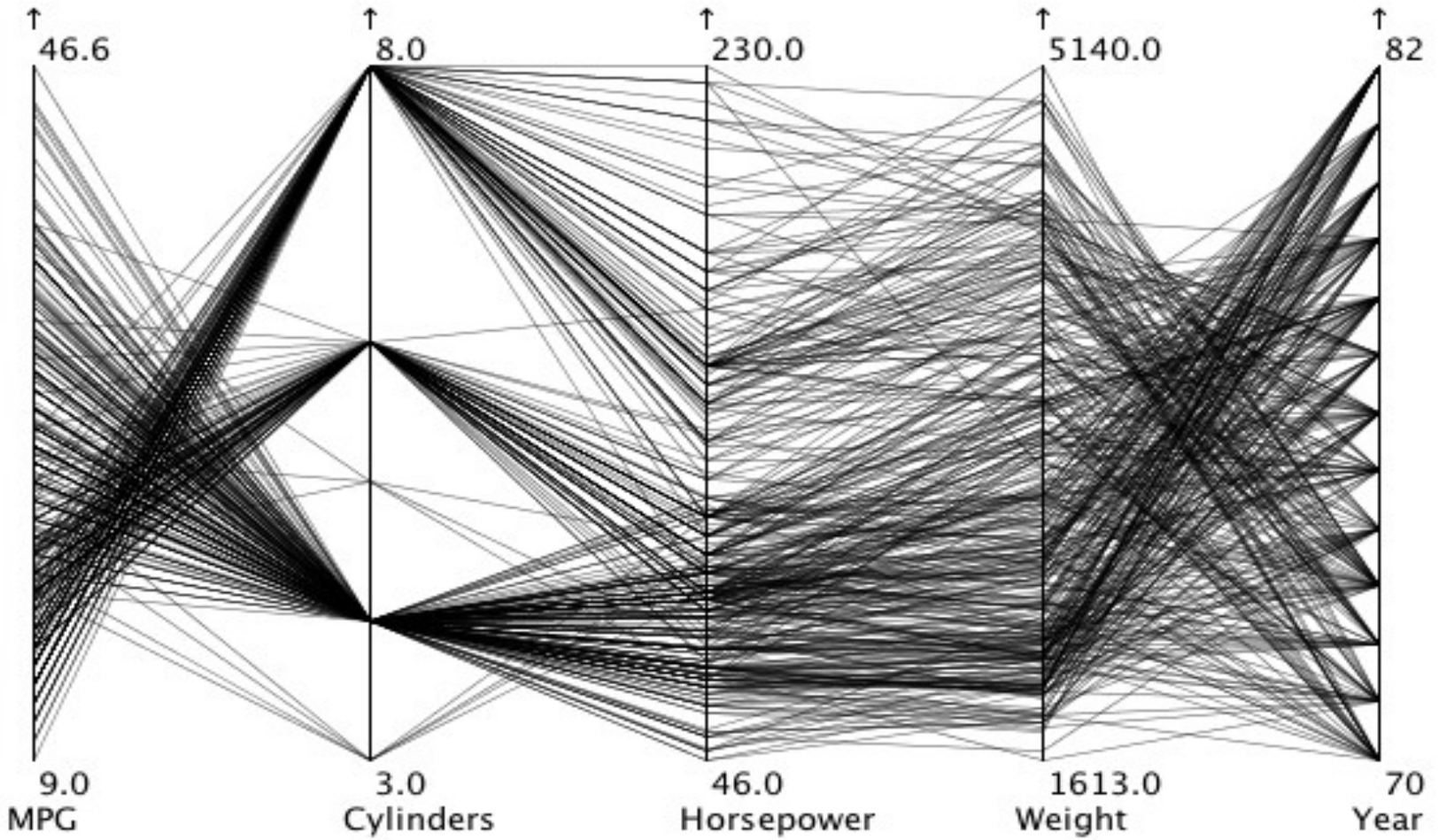




Parallel Coordinates



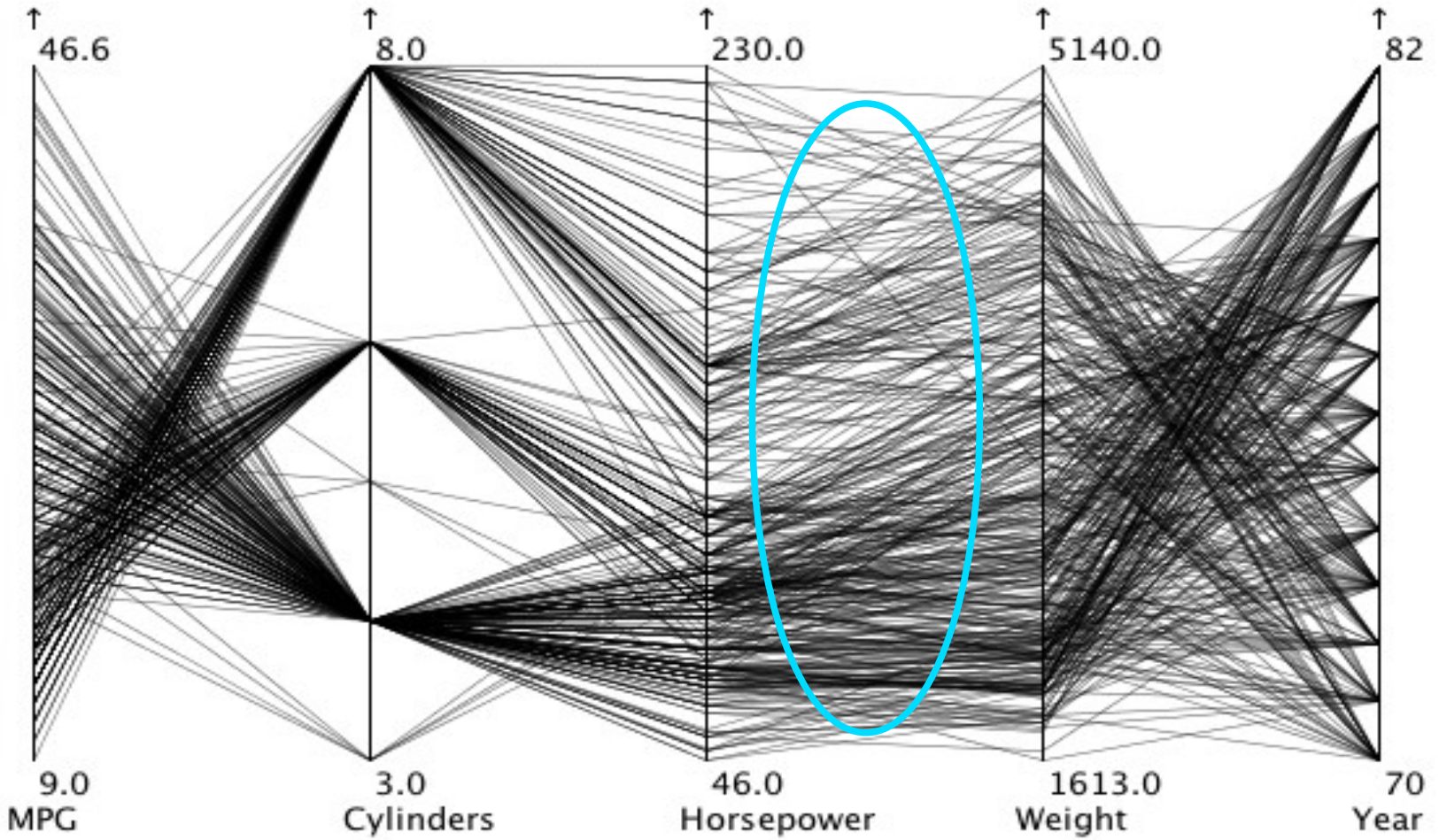
Parallel Coordinates



Car data :

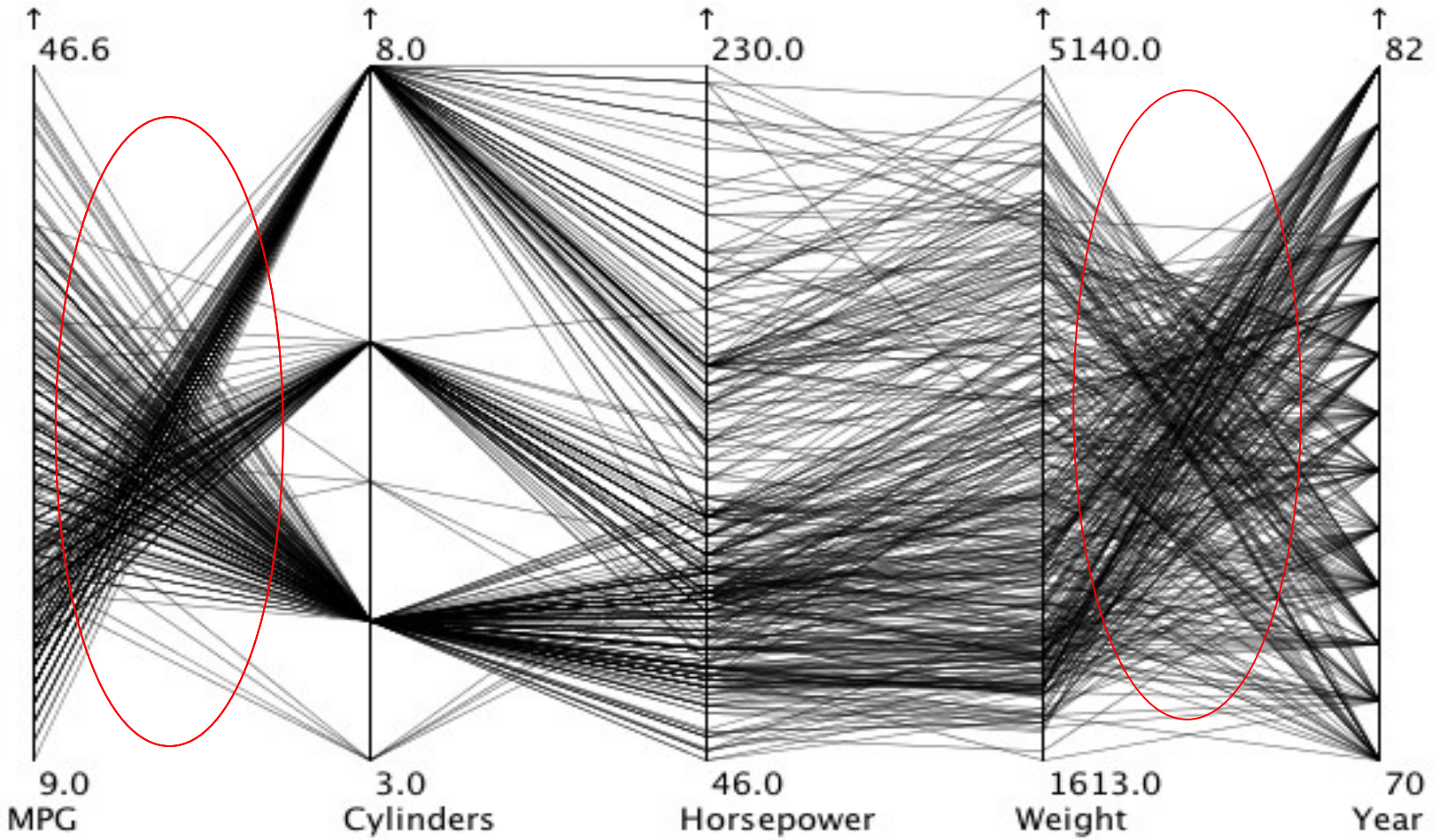
http://eagereyes.org/techniques/parallel-coordinates

Parallel Coordinates



Direct correlation

Parallel Coordinates

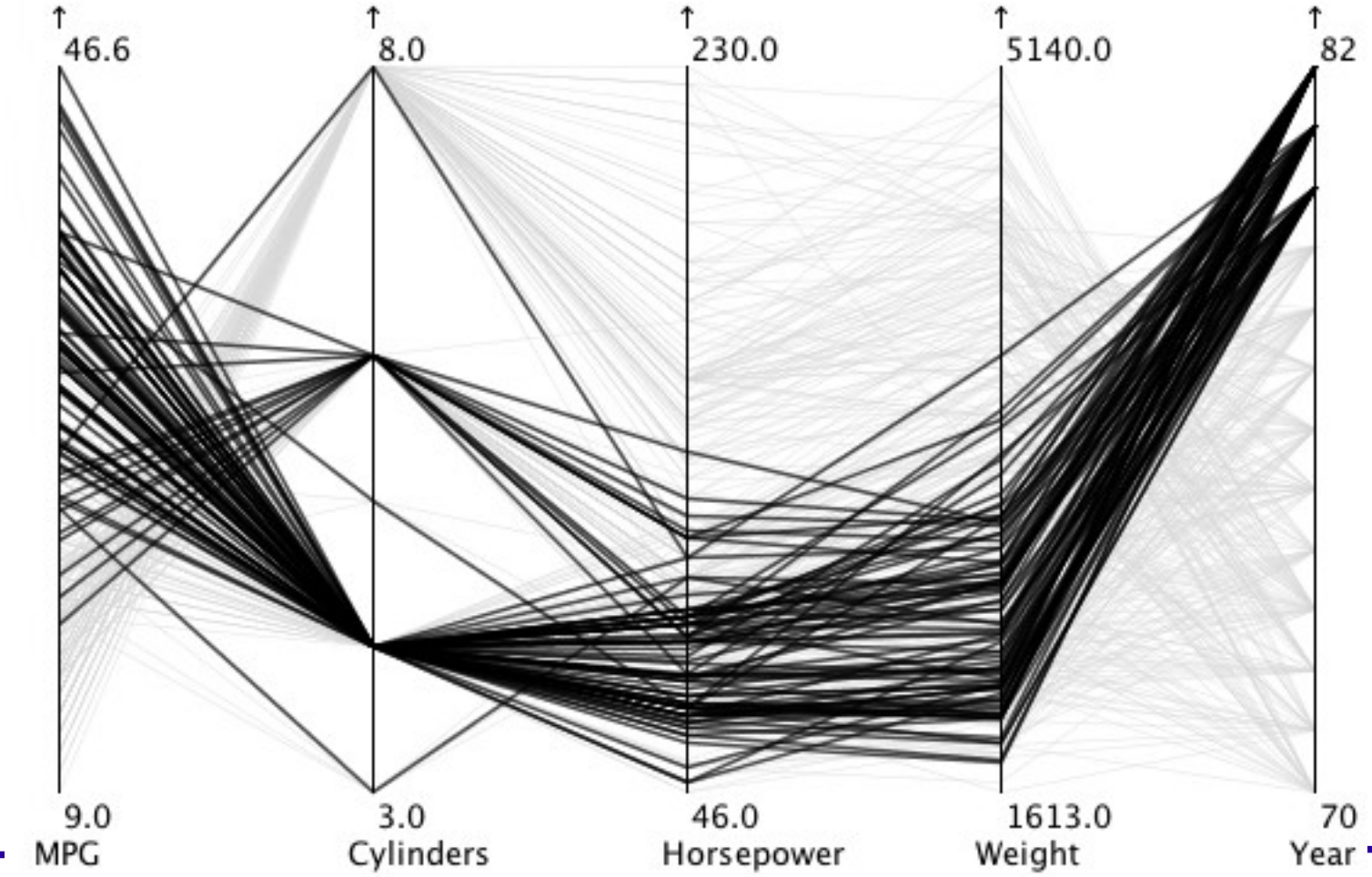


Inverse Relations

Brushing

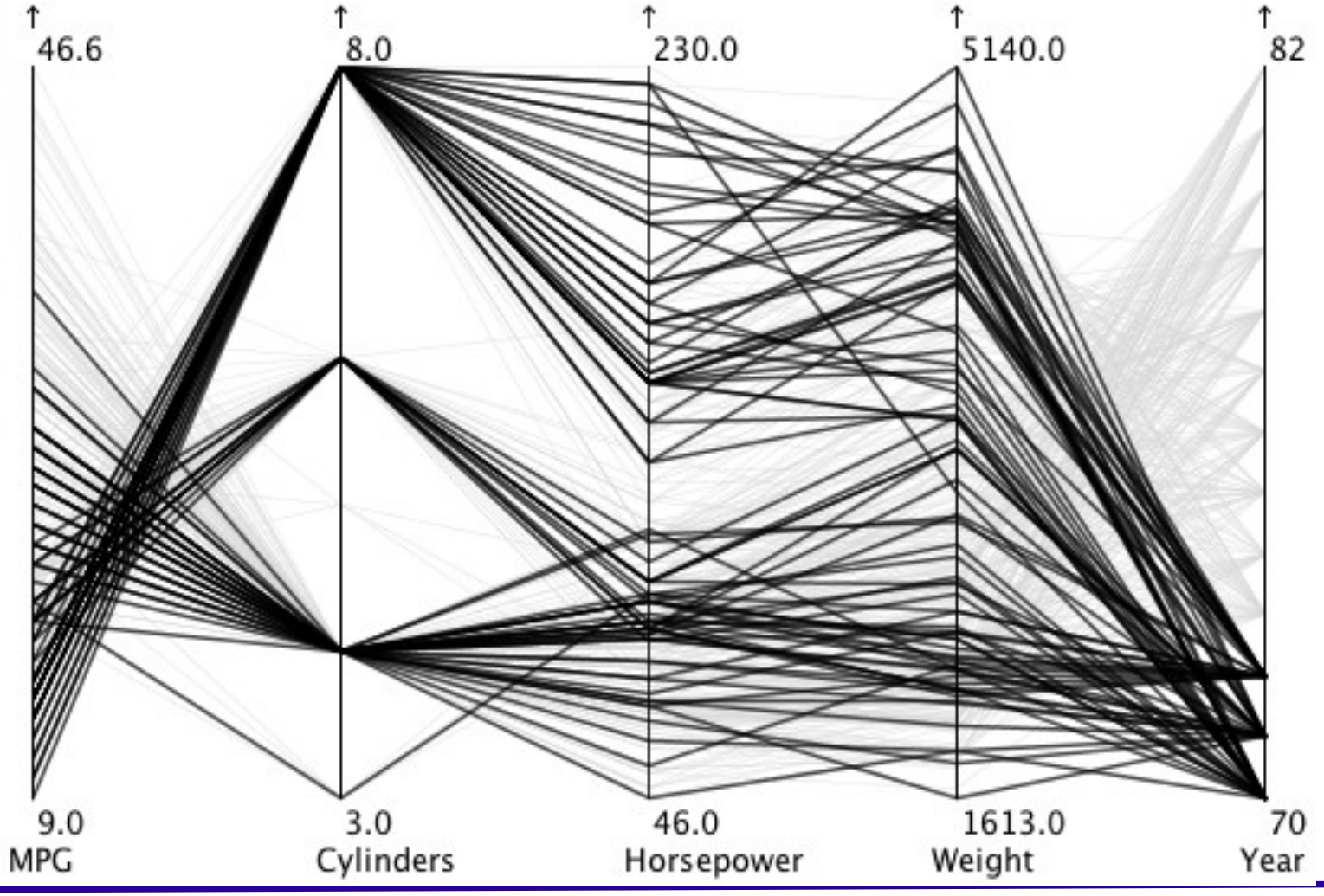
Select some data using one of the coordinates

Brushing years 1980 to 1982



Brushing

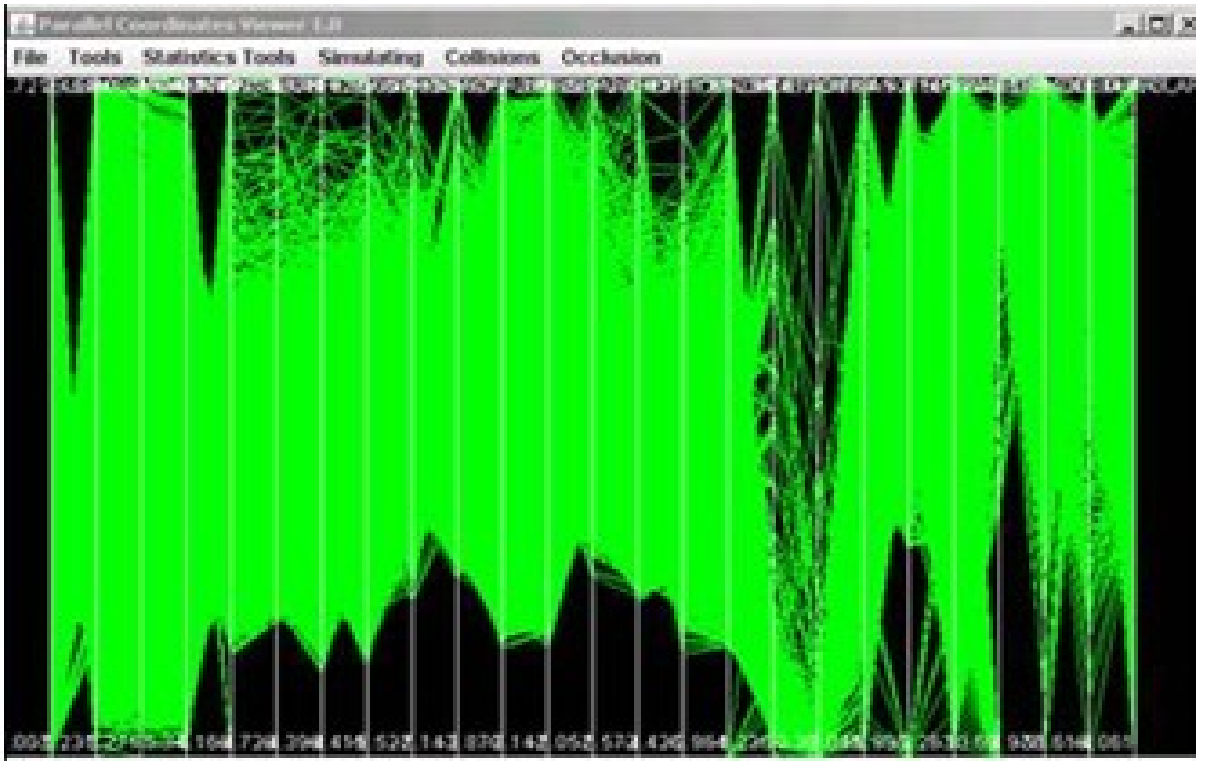
Brushing the years 1970 to 1972



Limitations

Visual clutter

- Many lines cluttered together making it impossible to see anything
- Too many dimensions make things difficult to see

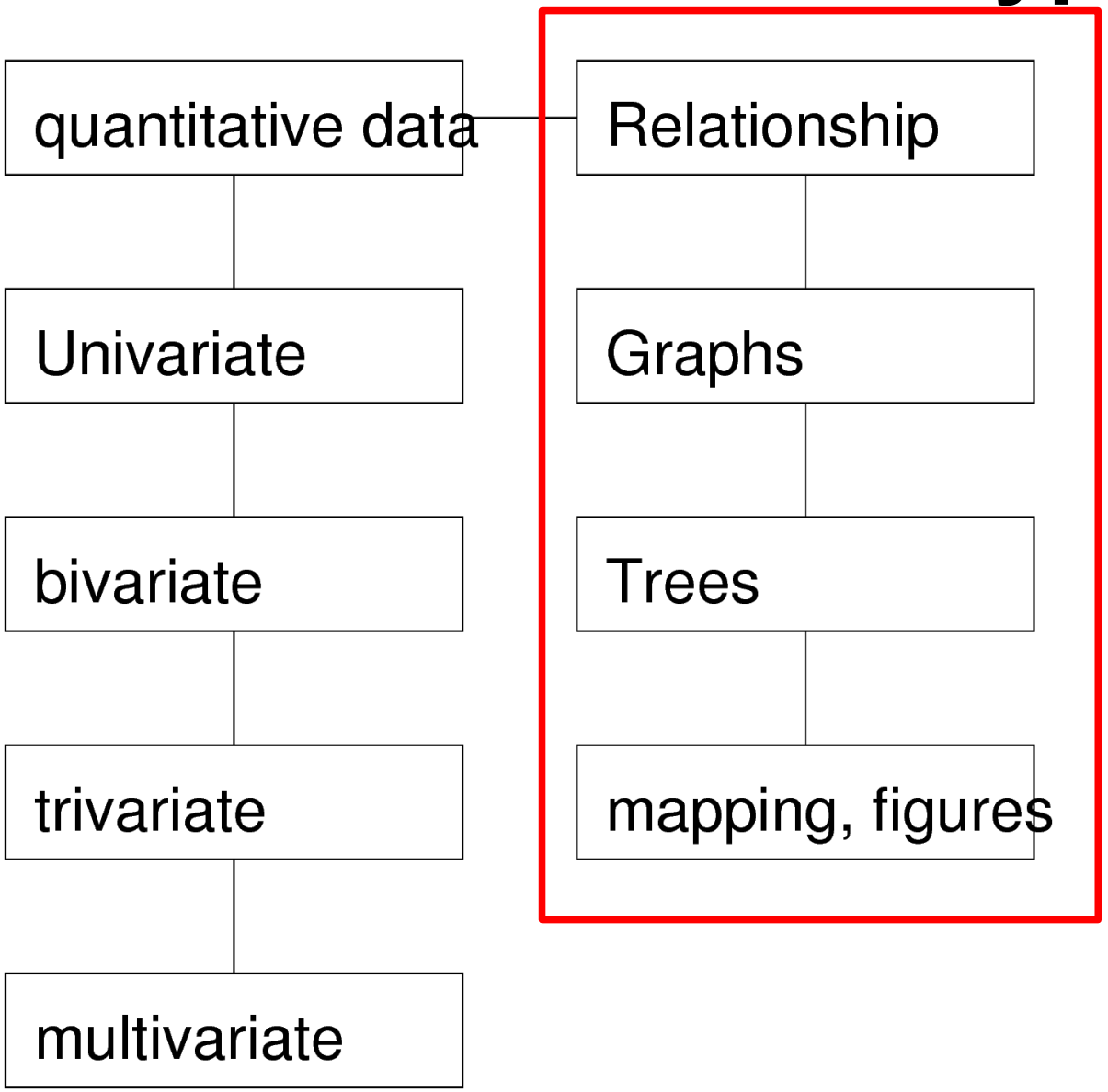


http://mbostock.github.io/protovis/ex/cars.html

Solutions

- Clustering the axes
- Re-ordering the axes
- But in general, cannot handle a large number of data

Data types

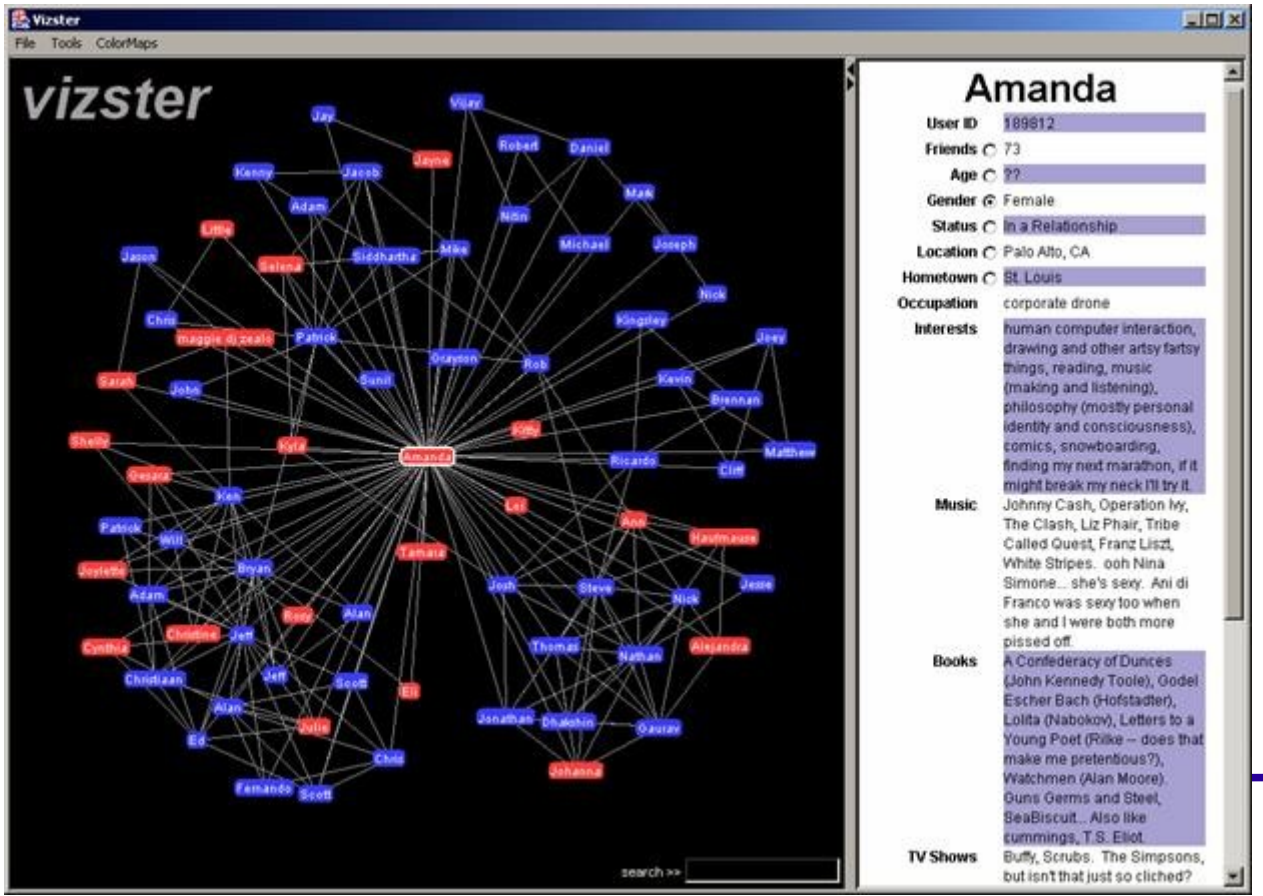


Document Visualization

Visualizing Relations

Relation: A logical or natural association between two or more things; relevance of one to another; connection

Usually use lines to represent the relations



Tree visualization

Trees have hierarchical structures

- No close loops

So many methods : see

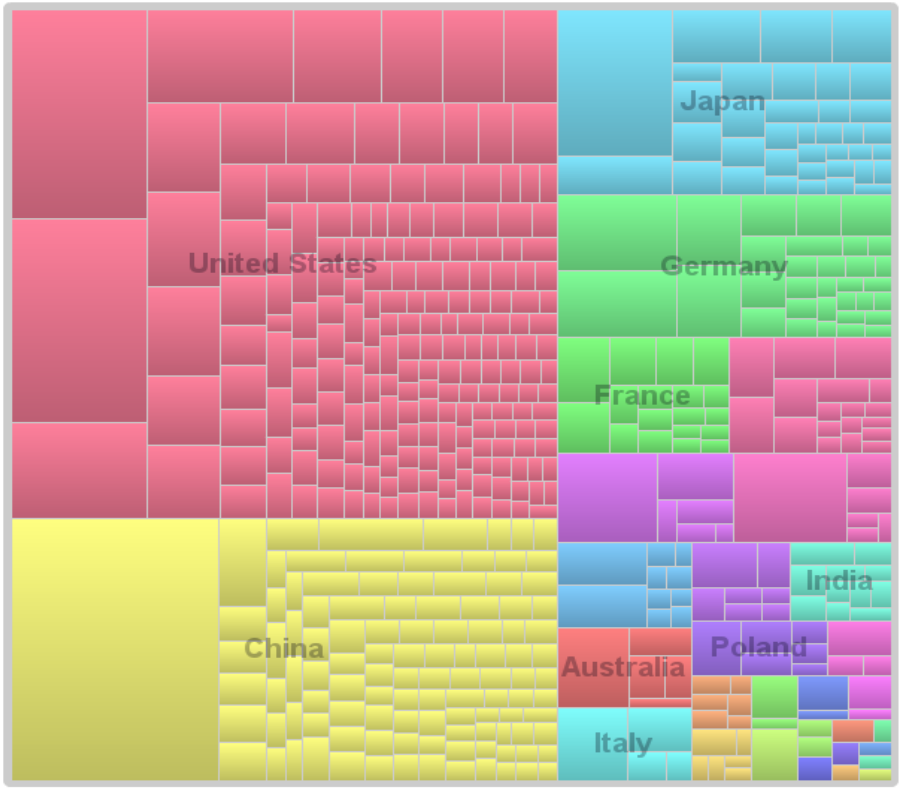
http://vcg.informatik.uni-rostock.de/~hs162/treeposter/poster.html

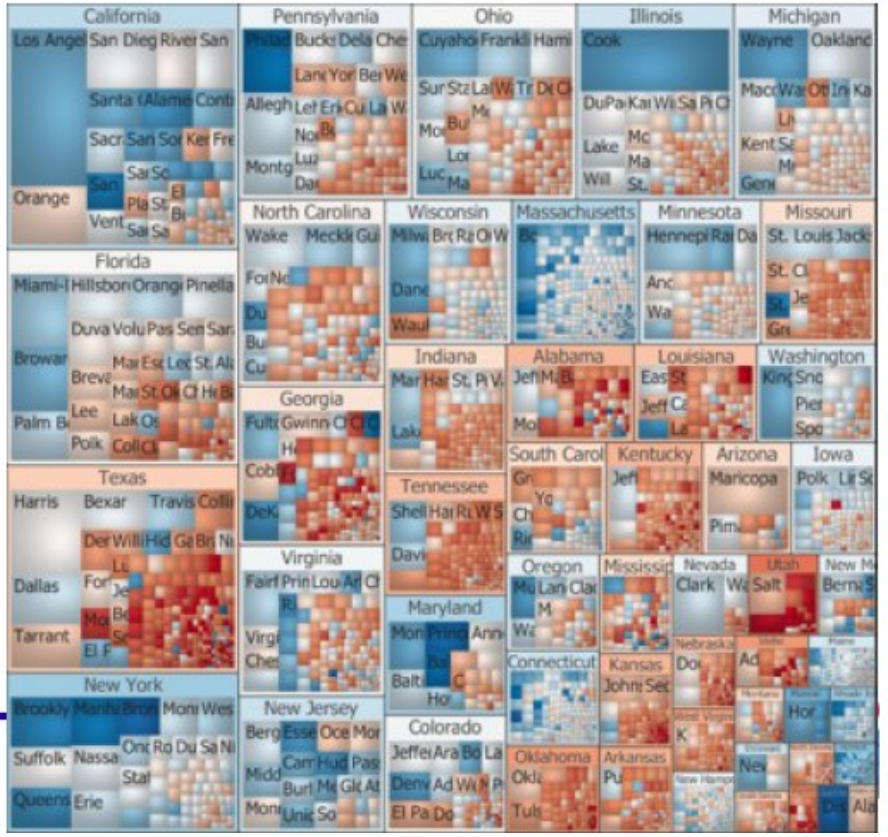
Treemaps

Display hierarchical (tree-structured) data as a set of nested rectangles

The area of the rectangles representing a scalar attribute

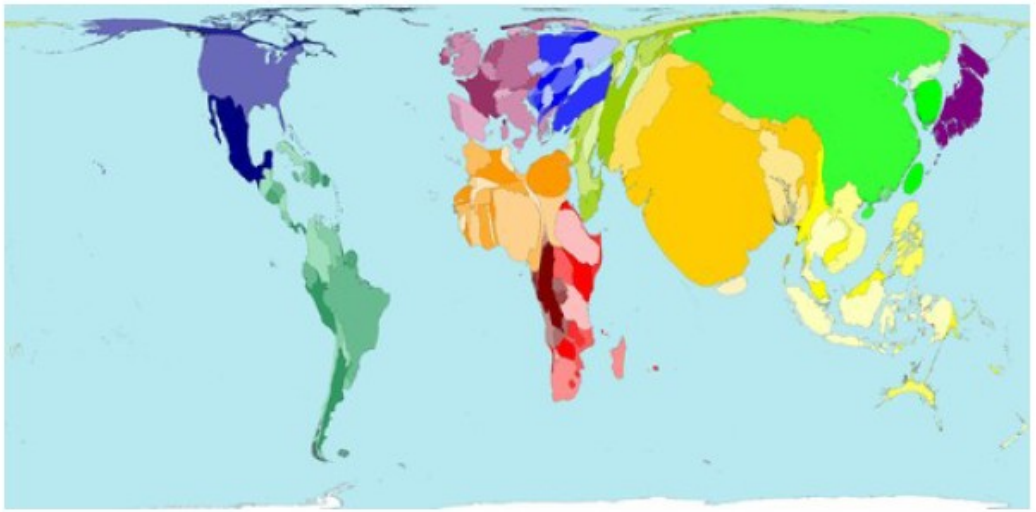
The leaf nodes are often colored to visualize another attribute data

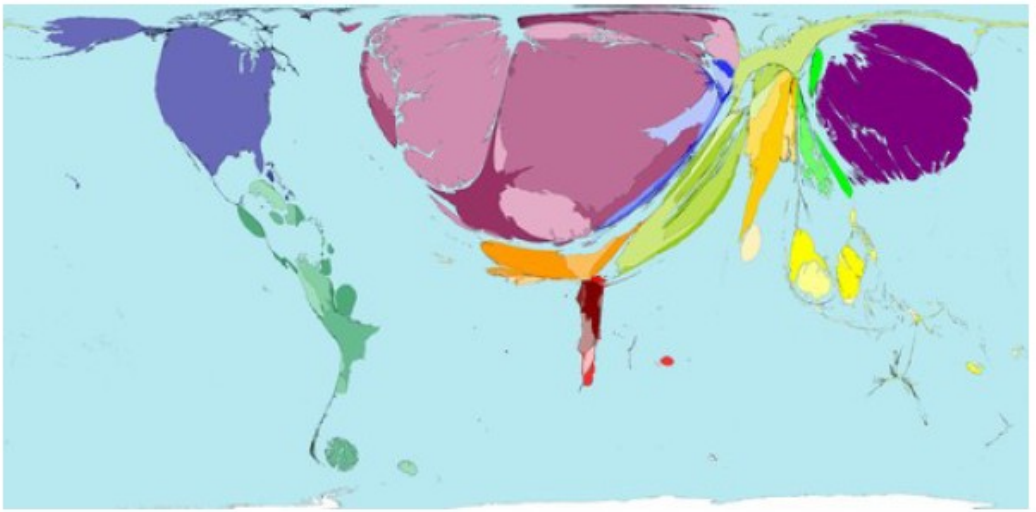




Worldmapper

http://sasi.group.shef.ac.uk/worldmapper/





Distorted maps according to numbers: **Cartograms**

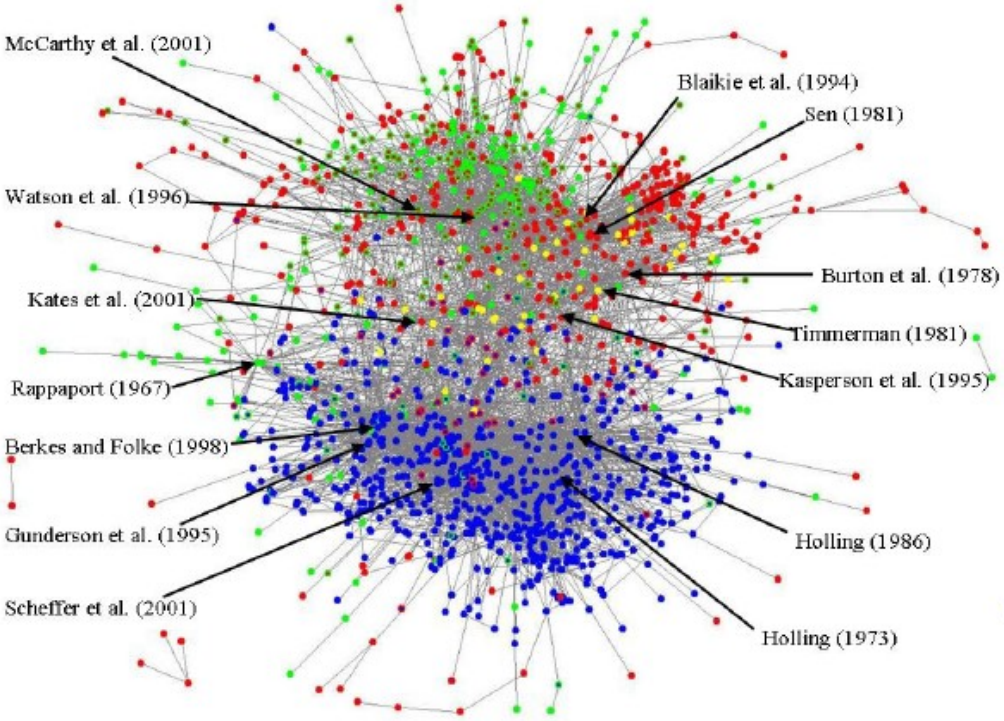
Graph Visualization

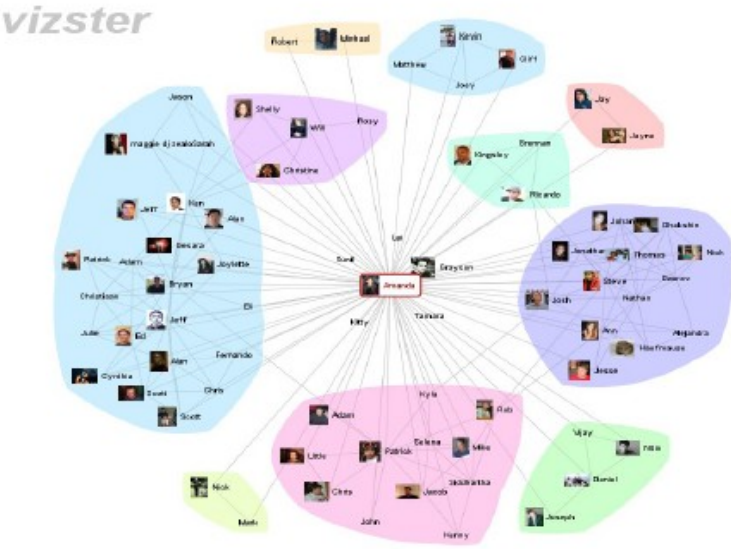
Visualizing correlation of different nodes

E.g.

- social networks
- citation networks

https://vimeo.com/19278513





Facebook relations

Facebook relations

“I defined weights for each pair of cities as a function of the Euclidean distance between them and the number of friends between them. Then I plotted lines between the pairs by weight, so that pairs of cities with the most friendships between them were drawn on top of the others. I used a color ramp from black to blue to white, with each line's color depending on its weight. I also transformed some of the lines to wrap around the image, rather than spanning more than halfway around the world."

Formal Aesthetics Metrics

Minimize node-node / node-edge occlusion

Minimize edge crossings

Minimize edge bends

Maximize symmetry

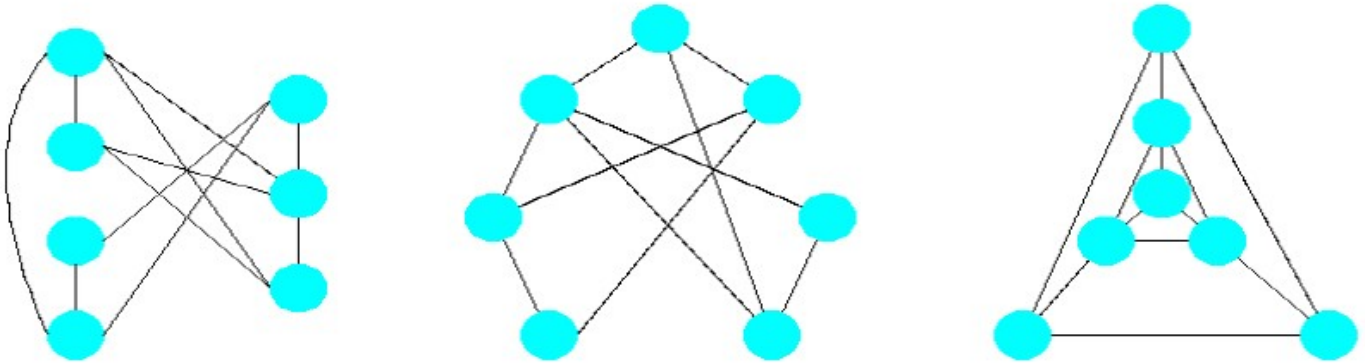
Maximize the minimum angle between neighbor edges

Maximize edge orthogonality

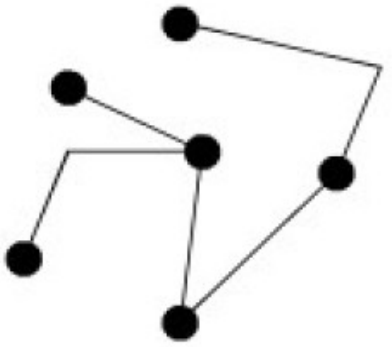
Maximize node orthogolnality

Formal Aesthetics Metrics

Minimize edge crossings

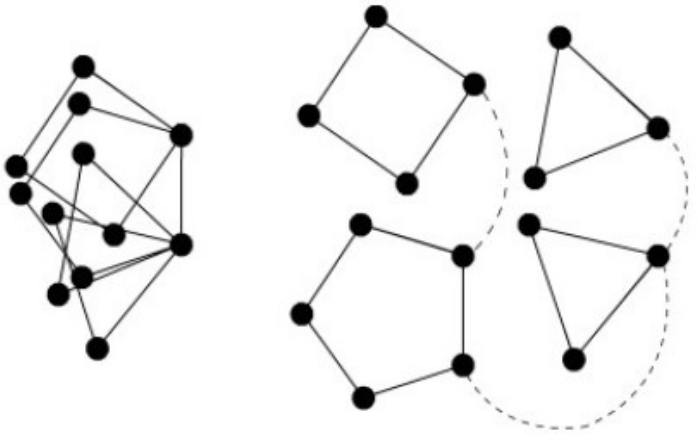


Minimize edge bends

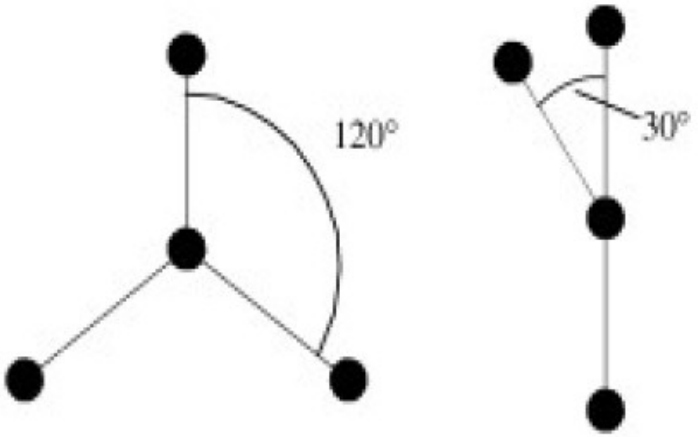


Formal Aesthetics Metrics

Maximizing symmetry

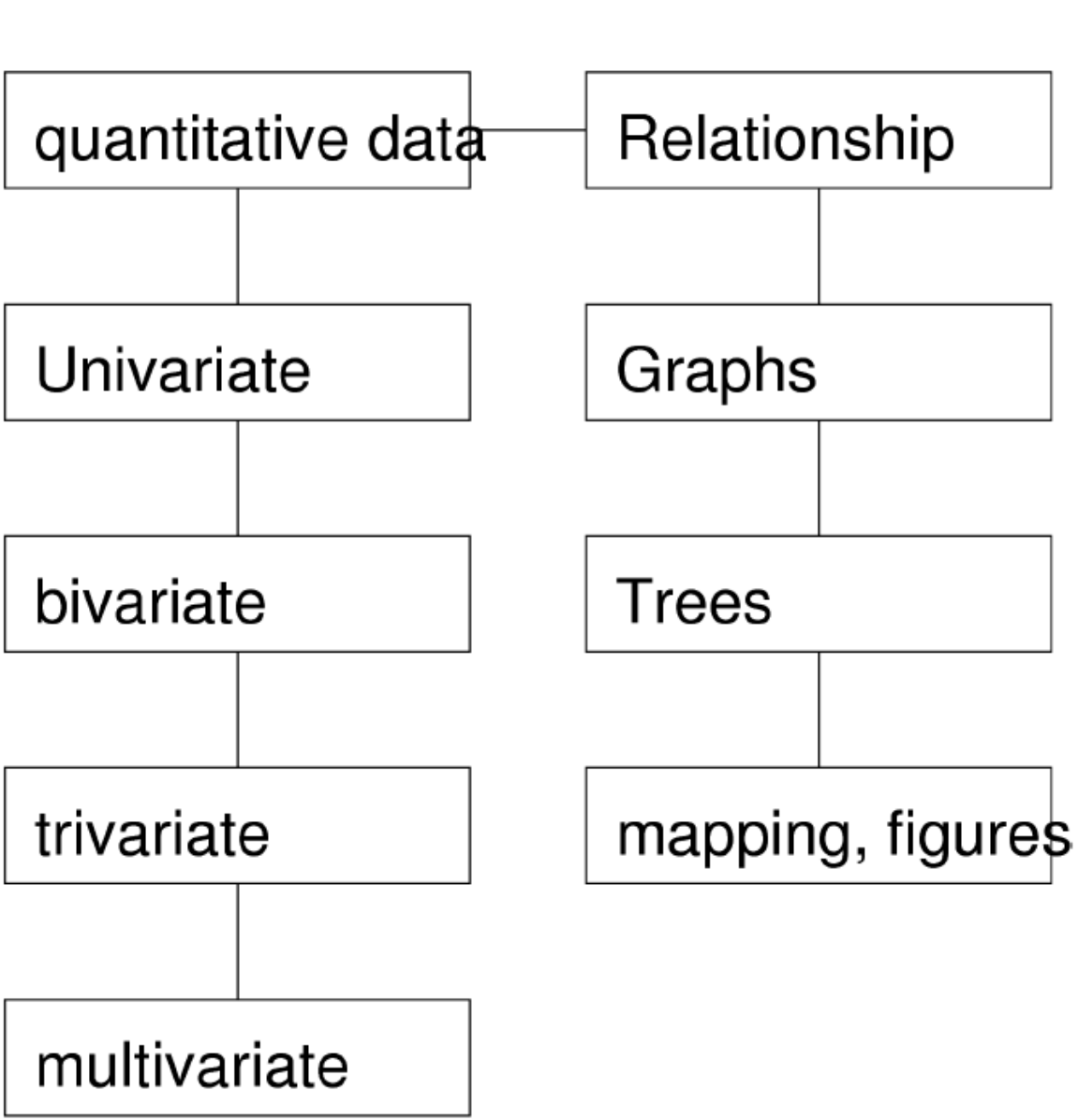


Maximizing the minimum angle between edges leaving a node



CAV : Lecture 18

Data types



Document Visualization

Document Visualisation

- Motivation:

Action	Units of Information transfer
Typing at 10 bytes per second	1
Mouse Operations.	2
Reading	3-40
Hearing	60
Visualisation and Pattern Recognition	12,500

Source :
Silicon Graphics Inc.

  - visualisation is considerably faster than hearing / reading!

Visualisation of Documents

- **Motivation : large bandwidth of human visual system**
  - 100s millions of documents available on-line
  - information only in textual form
- **'Visualising the non-visual'**
  - searching for scientific papers
  - analysing witness statements
  - awareness of events in news bulletins

Document Visualisation - Stages

- **Representation of results**
  - form high-dimensional vector (one for each word, ~10000+)
  - cluster documents based on vector similarity (e.g. Nearest-Neighbour)
- **Visualisation of clustered results**
  - projection to lower dimensional space
  - 3D “galaxy” / 2D “theme-scape” / 1D “theme-river”

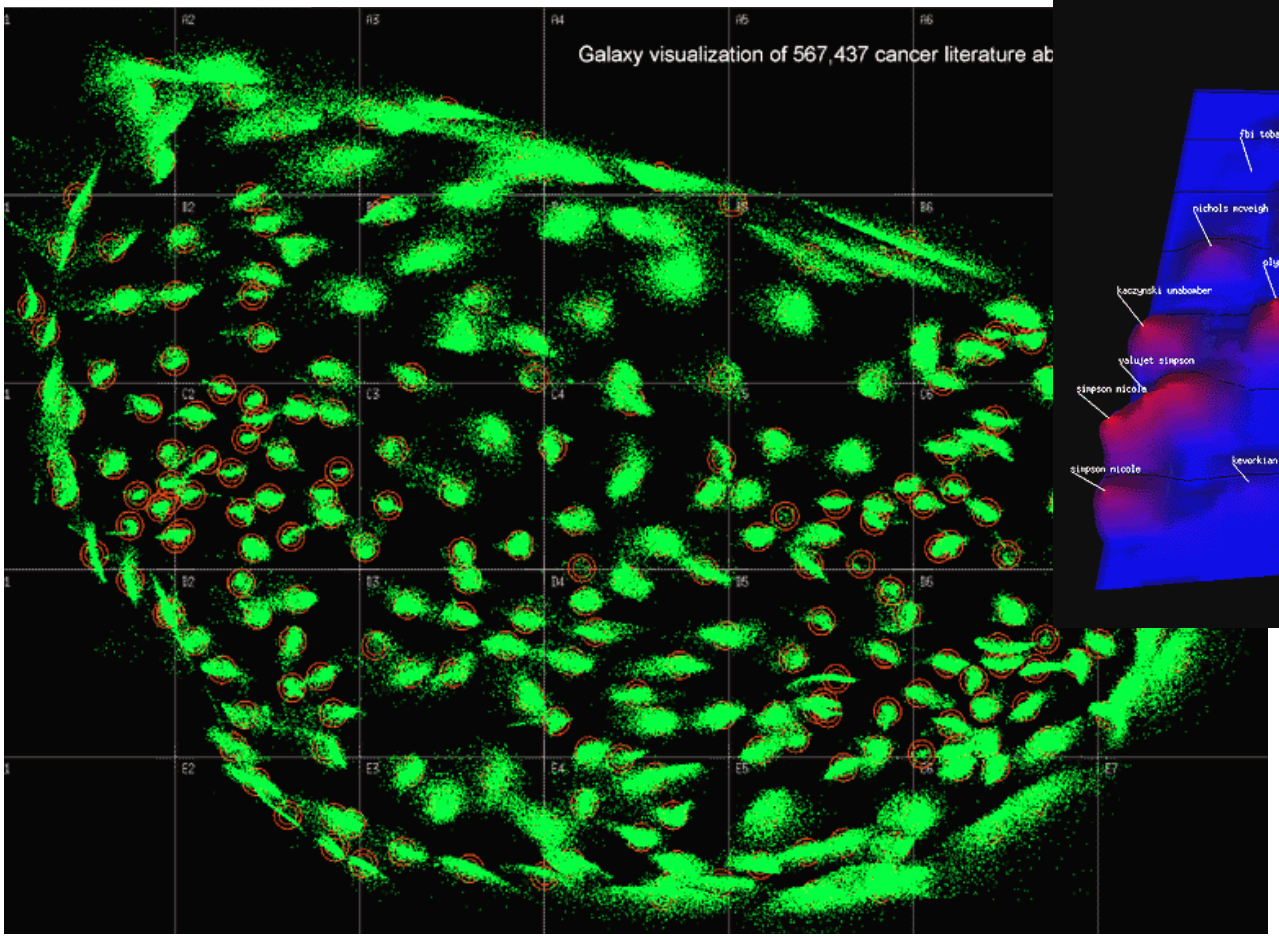
Query

“keywords” from user specification

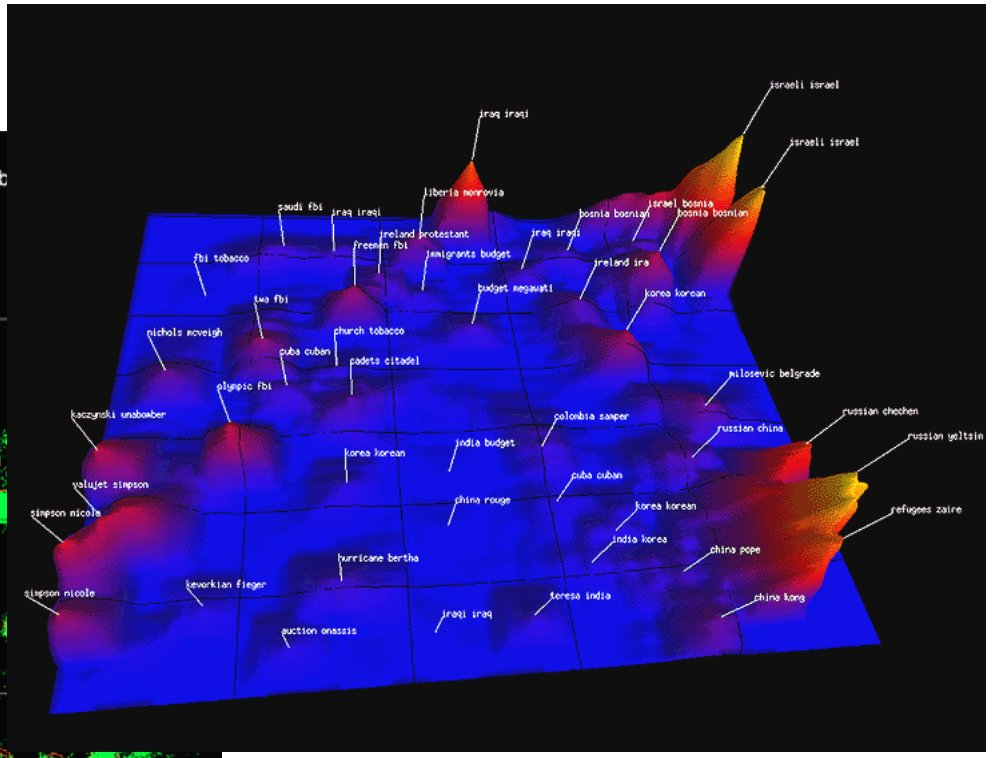
comparison to sample “reference” document

2D and 3D projections of documents

3D Visualisation of 567,000 cancer literature abstracts.



Pacific Northwest National Laboratory.

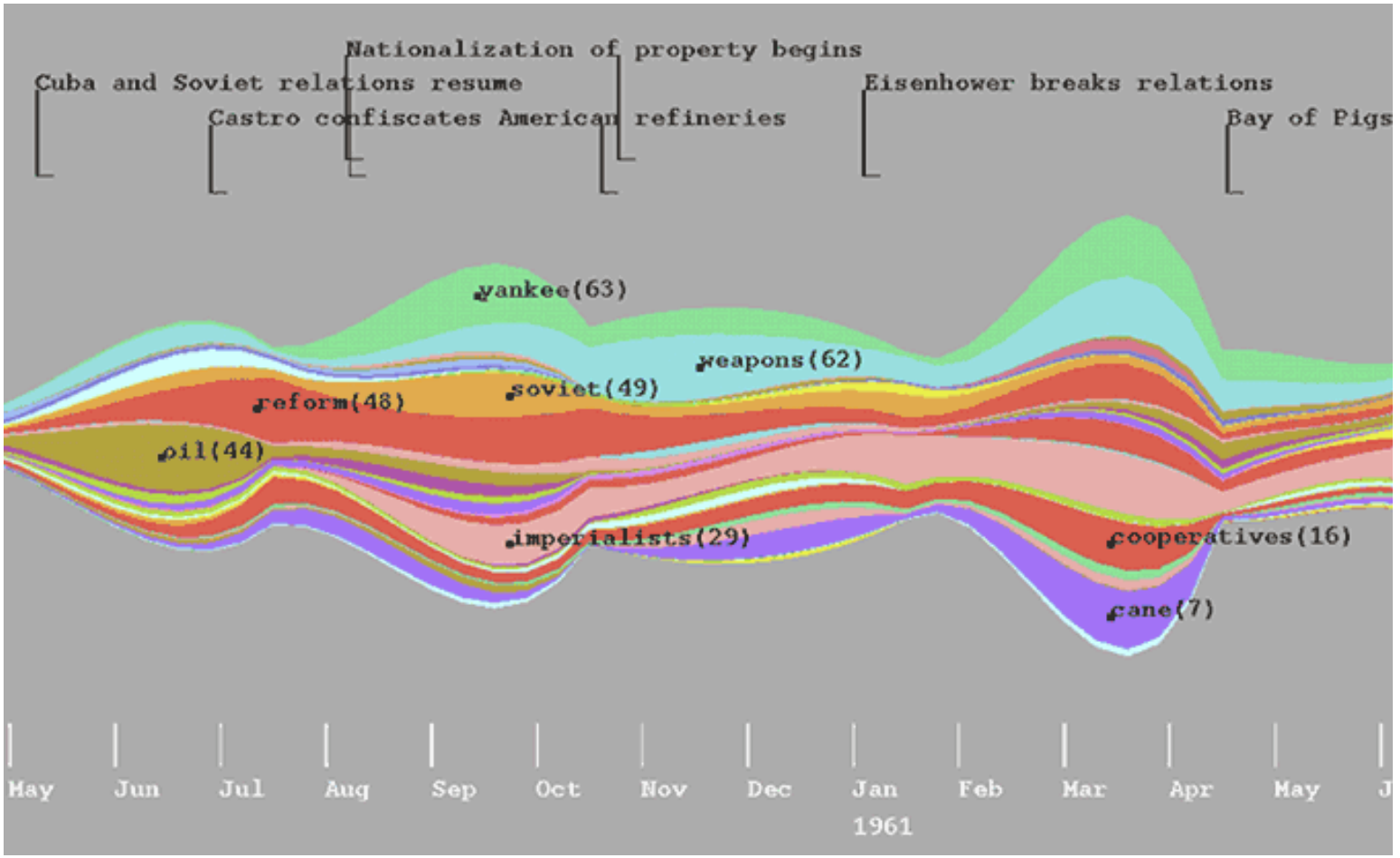


Articles in a collection of news items (2D).

1D visualisation of news articles

A 'Theme River' shows the relative importance of themes over the course of a year from press articles.

Pacific Northwest National Laboratory.

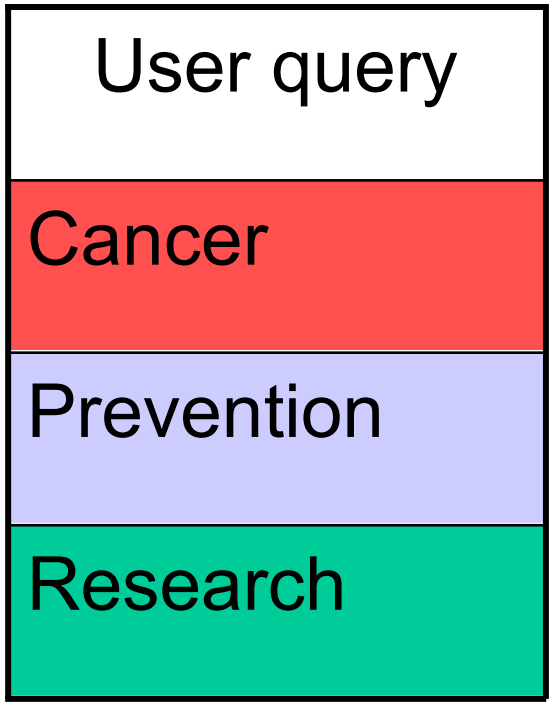


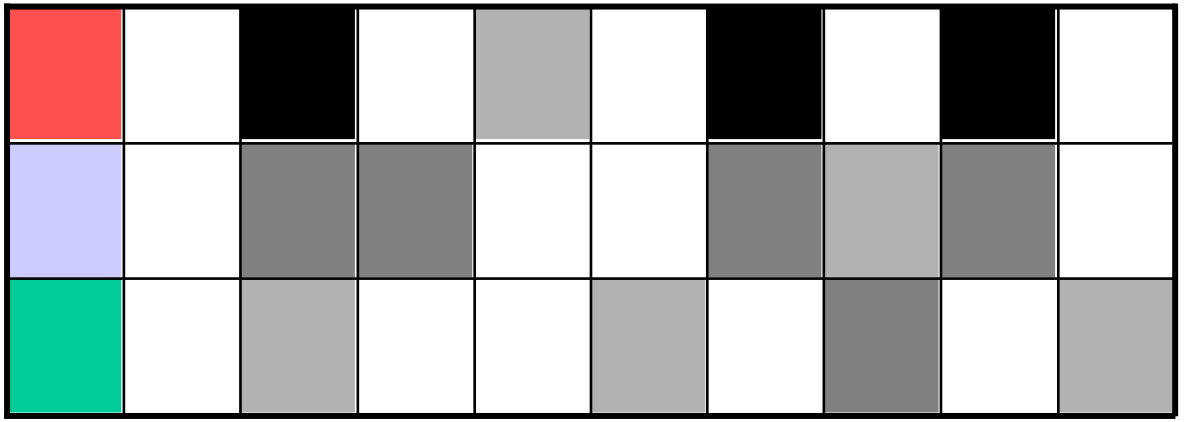
http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html?_r=0

Document Querying

- **Keyword search is problematic**
  - **ambiguity**
  - ~7-18% of people describe same concept with same word (Barnard '91)
- Interested in
  - **distribution of keywords in the document**
  - **related articles** to the keyword entered
- **Tile bar scheme** (Hearst 1995)
  - display a list of documents with a **tile bar**
  - tile bar **shows the occurrence of keywords** in document

Title Bar Method





Columns **represent** paragraphs or pages in a document.
Shade indicates relevance shown by word occurrence.
Shows length and likely relevance.
System allows **interactivity** by clicking on box.

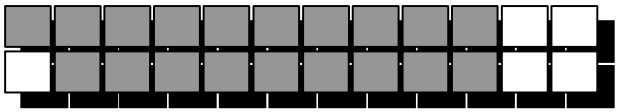
- Visualisation - Use of document topology / colour-mapping / interaction

Example : Title Bar Query / Result

Query terms: DBMS (Database Systems)

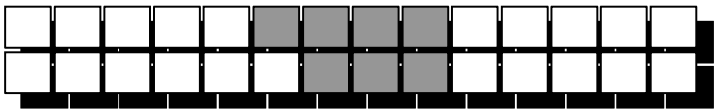
Reliability

What roles do they play in retrieved documents?

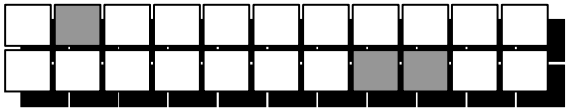


Mainly about both DBMS & reliability

Mainly about DBMS, discusses reliability



Mainly about, say, banking, with a subtopic discussion on DBMS/Reliability

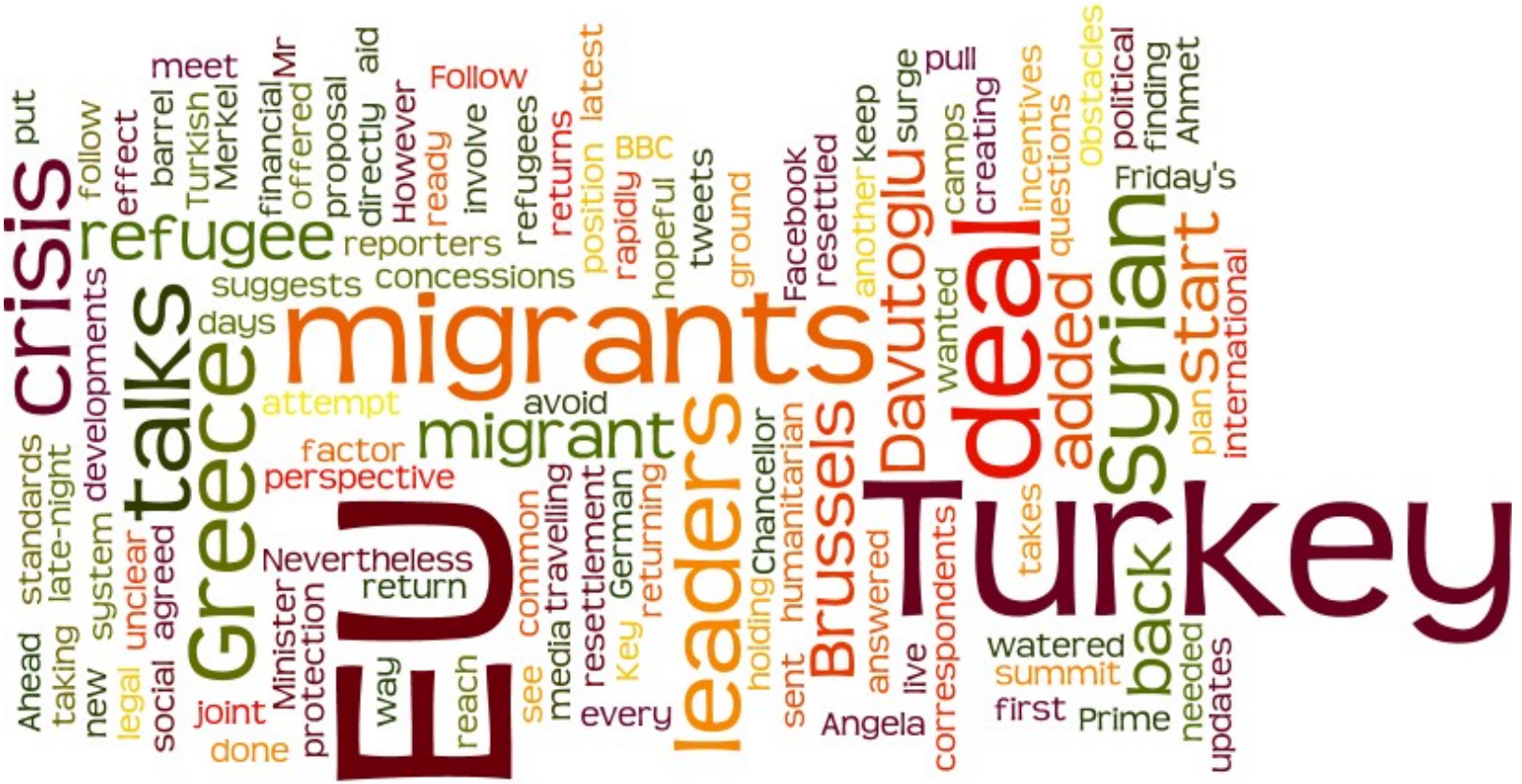


Mainly about something different

Wordle

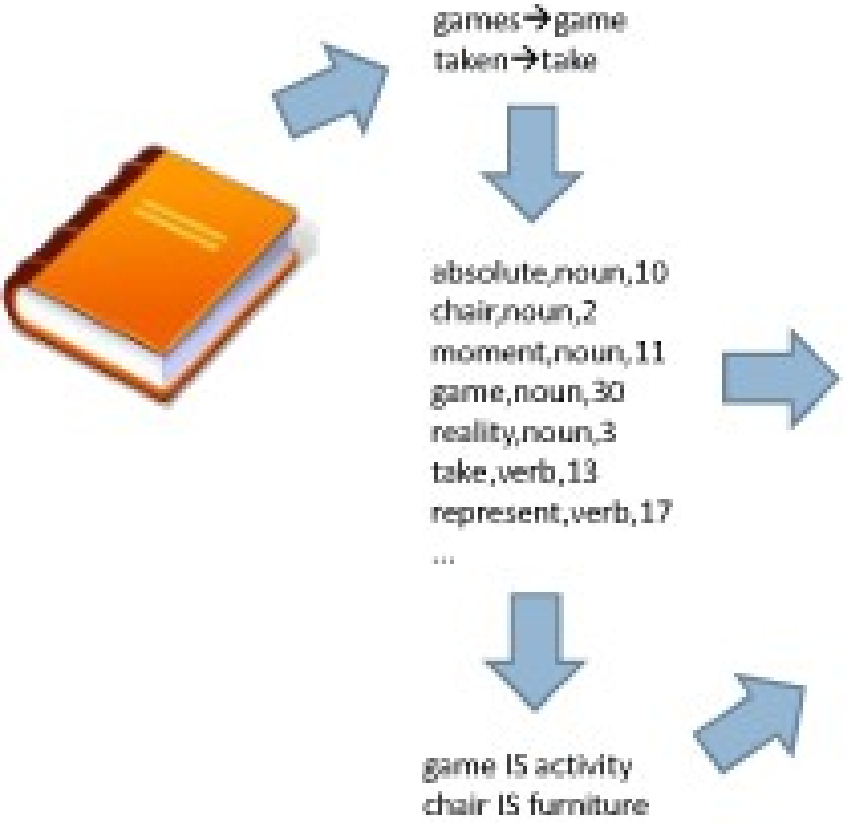
http://www.wordle.net/create

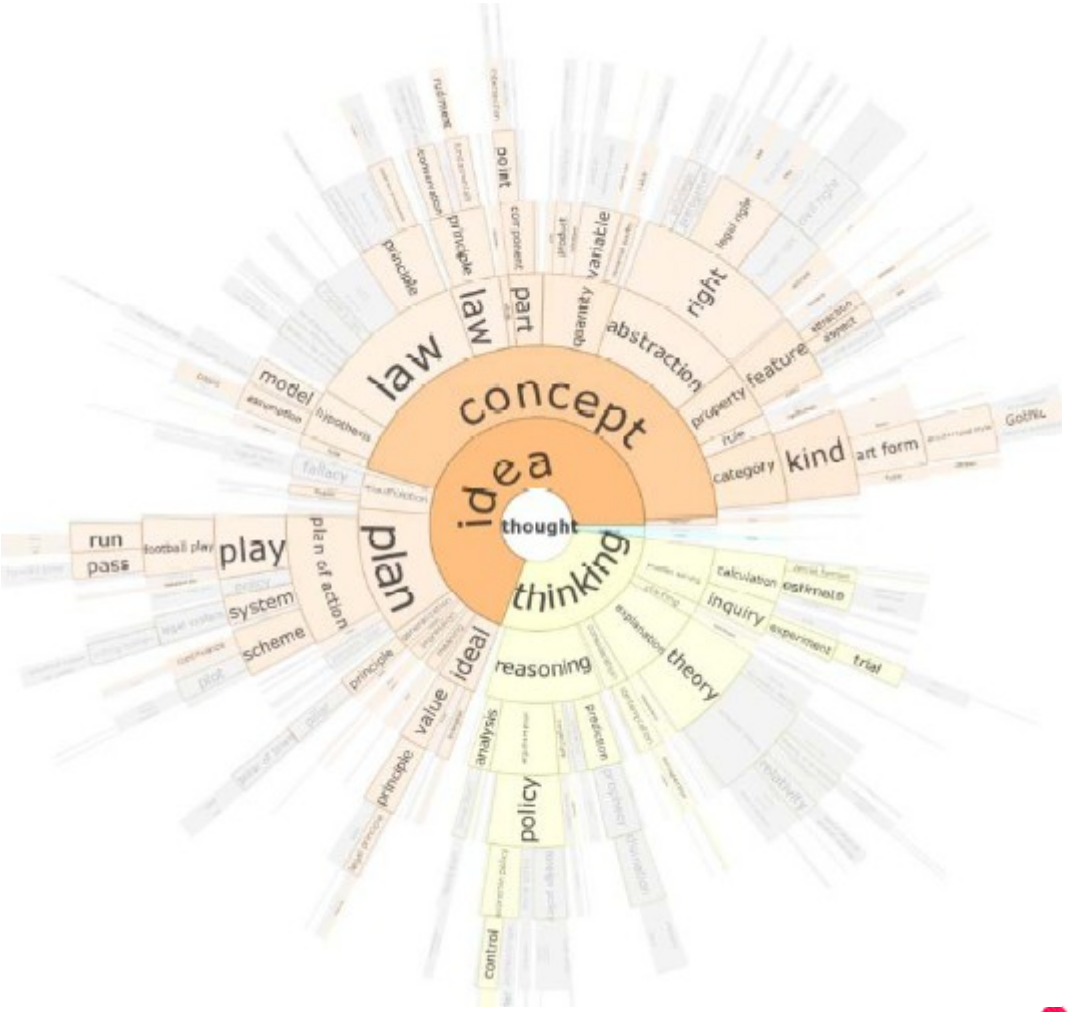
Produces a word cloud from a document



DocuBurst

- A radial, space-filling layout of hyponymy (IS-A relation)





Summary

- Information Visualisation
  - Univariate, bivariate, trivariate, multi-variate data
  - Relations visualized by lines, tree visualization
  - Document visualization

Reading

- *Marti A. Hearst* **TileBars: Visualization of Term Distribution Information in Full Text Information Access**
- **Collins, Christopher;** Carpendale, Sheelagh; and Penn, Gerald. DocuBurst: Visualizing Document Content using Language Structure. Computer Graphics Forum (Proceedings of Eurographics/IEEE-VGTC Symposium on Visualization (EuroVis '09)), 28(3): pp. 1039-1046, June, 2009
- http://faculty.uoit.ca/collins/research/docuburst/index.html
- http://searchuserinterfaces.com/book/sui_ch11_text_analysis_visualization.html