

COMS21202: Symbols, Patterns and Signals

Information Visualisation

Dima Damen

`Dima.Damen@bristol.ac.uk`

Bristol University, Department of Computer Science
Bristol BS8 1UB, UK

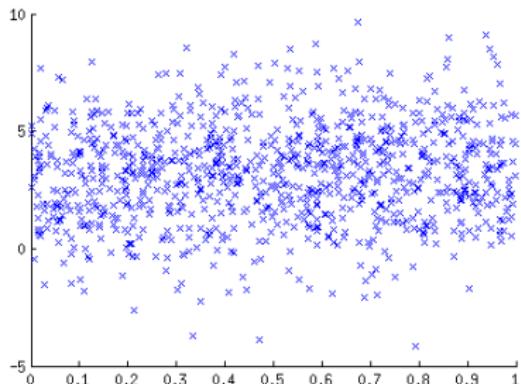
February 9, 2019

Information Visualisation

- ▶ Information visualisation, as a research discipline, has emerged over the last 20 years
- ▶ Driven by the volumes of data and the accessibility of *big data*
- ▶ Characterised by large quantities of data - not necessarily numbers

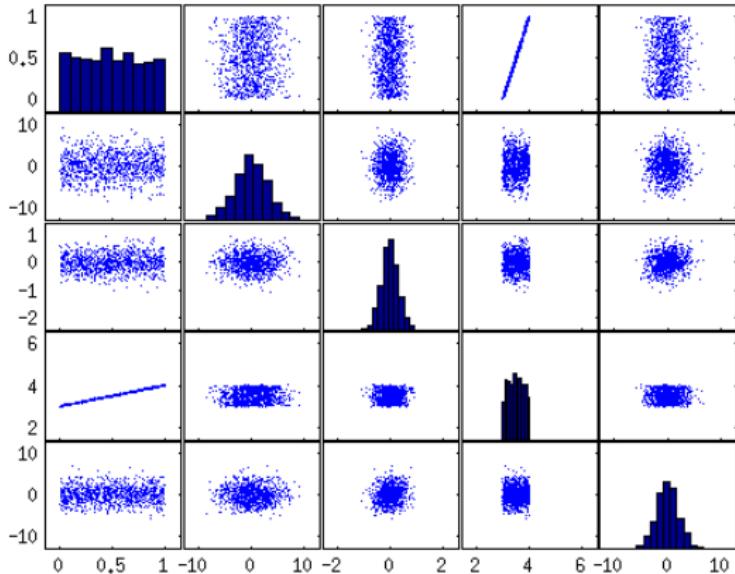
Information Visualisation - Simple Graphs

1. Scatter Plot



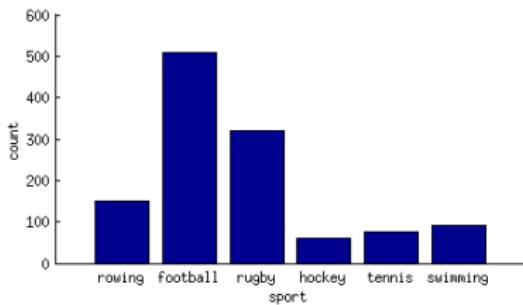
Information Visualisation - Simple Graphs

1. Scatter Plot
2. Scatter Plot Matrix



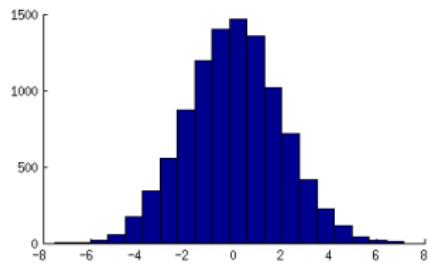
Information Visualisation - Simple Graphs

1. Scatter Plot
2. Scatter Plot Matrix
3. Histogram
 - ▶ Discrete Variable
(bar chart)



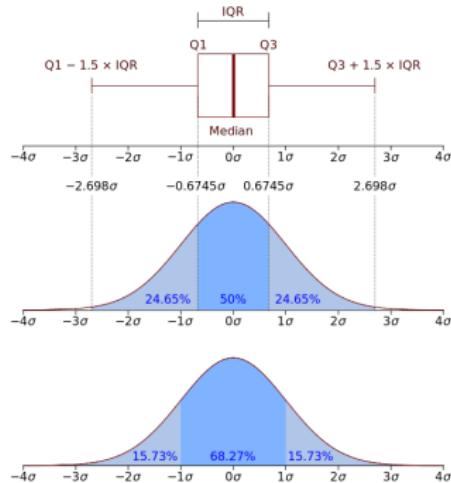
Information Visualisation - Simple Graphs

1. Scatter Plot
2. Scatter Plot Matrix
3. Histogram
 - ▶ Discrete Variable
 - ▶ Continuous Variable
 - Δv : width of bin
 - c : bin number; $0 \leq c \leq N$
 - $\min_x + c\Delta v \leq v < \min_x + (c + 1)\Delta v$



Information Visualisation - Simple Graphs

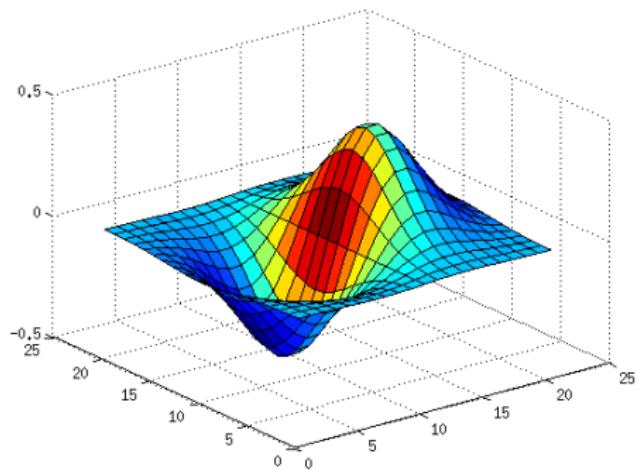
1. Scatter Plot
2. Scatter Plot Matrix
3. Histogram
4. Box plot



source:Wikipedia(2015)

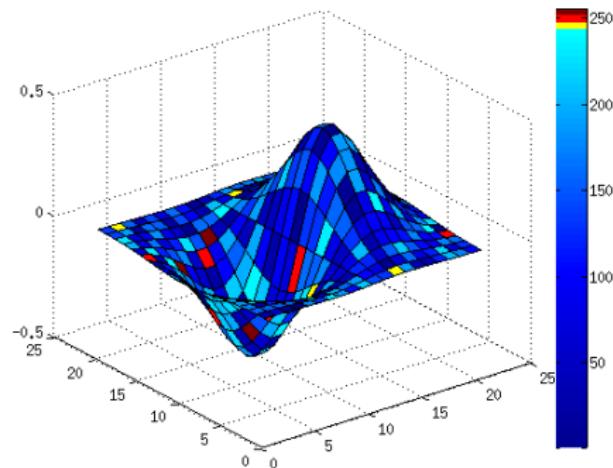
Information Visualisation - Simple Graphs

1. Scatter Plot
2. Scatter Plot Matrix
3. Histogram
4. Box plot
5. Surface
▶ 3D Data



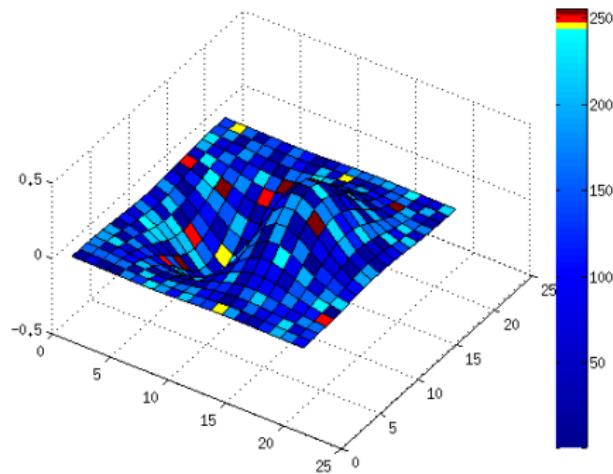
Information Visualisation - Simple Graphs

1. Scatter Plot
2. Scatter Plot Matrix
3. Histogram
4. Box plot
5. Surface
 - ▶ 3D Data
 - ▶ 4D Data



Information Visualisation - Simple Graphs

1. Scatter Plot
2. Scatter Plot Matrix
3. Histogram
4. Box plot
5. Surface
 - ▶ 3D Data
 - ▶ 4D Data



Information Visualisation

- ▶ Is that all?
- ▶ **Problem:** How to understand massive datasets?
- ▶ **Solution:** Convert information into a graphical representation to take advantage of human perception
- ▶ Information visualisation – what can you vary?
 - ▶ Colour / Colour Maps
 - ▶ Size / contour width
 - ▶ Shape / line stroke
 - ▶ Location (2 dimensions)
 - ▶ Transparency
- ▶ Information visualisation: “The use of computer-supported, interactive, visual representations of abstract data to simplify cognition.” (Card, Mackinlay, Shneiderman, 1999)
- ▶ But it existed **before** computers!

Historical Note - Ex1

Napoleon's disastrous Russian campaign (1812)

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 état-major.

Les nombreux hommes perdus sont représentés par les longueurs des lignes colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des pertes. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été pris dans les messages de M. M. Cibot, de Ségur, de Tocqueville, de Chambry ou le journal intime de Jacob, pharmacien de l'Armée depuis le 23 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai superposé que le corps de l'Armée décomme du Maréchal Davout, qui arriva de ses succès sur Moscou au régime vers Ossouïa ou Witebsk, auquel temps mesché avec l'Armée.

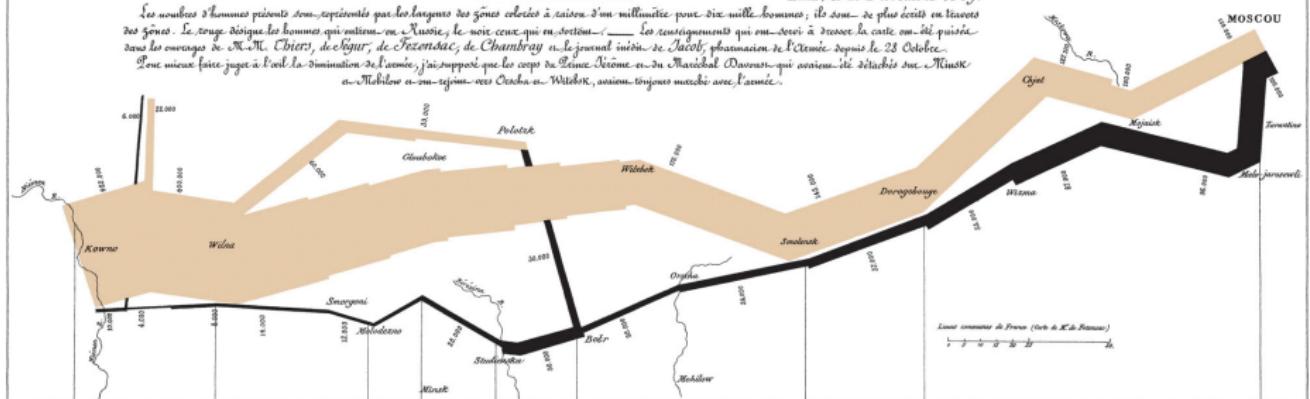
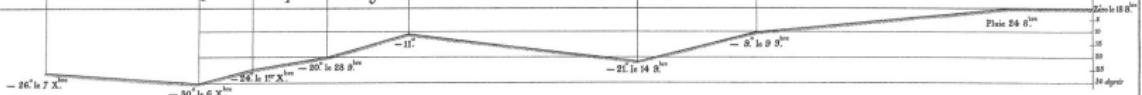


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Campagnes passent au gelé
le Nidement, juillet.

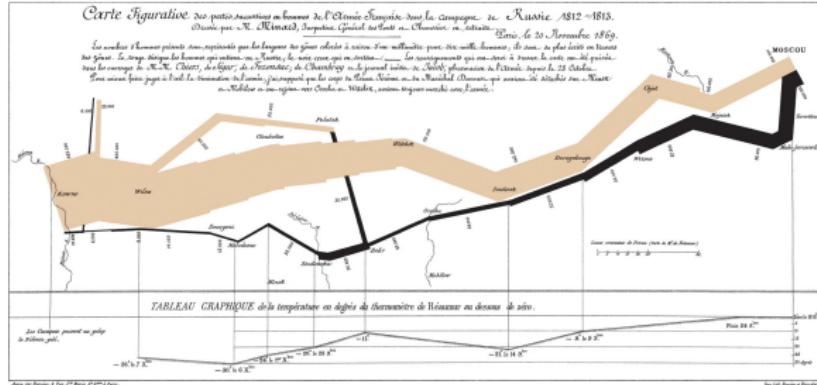


Abus. par Regnier, 8, Rue de la Paix, 25^e arrondissement de Paris.

Imp. L. Regnier à Bourges.

Further info <http://www.datavis.ca/gallery/re-minard.php>

Historical Note - Ex1



- ▶ Charles Minard
 - ▶ Six dimensions of data
 - ▶ Number of Napoleon's troops
 - ▶ Direction
 - ▶ Distance
 - ▶ Temperature
 - ▶ Location: Longitude and Latitude
 - ▶ Dates
 - ▶ *Acclaimed best statistical graphic ever drawn*

Dima Damen

Dima.Damen@bristol.ac.uk

COMS21202: Information Visualisation

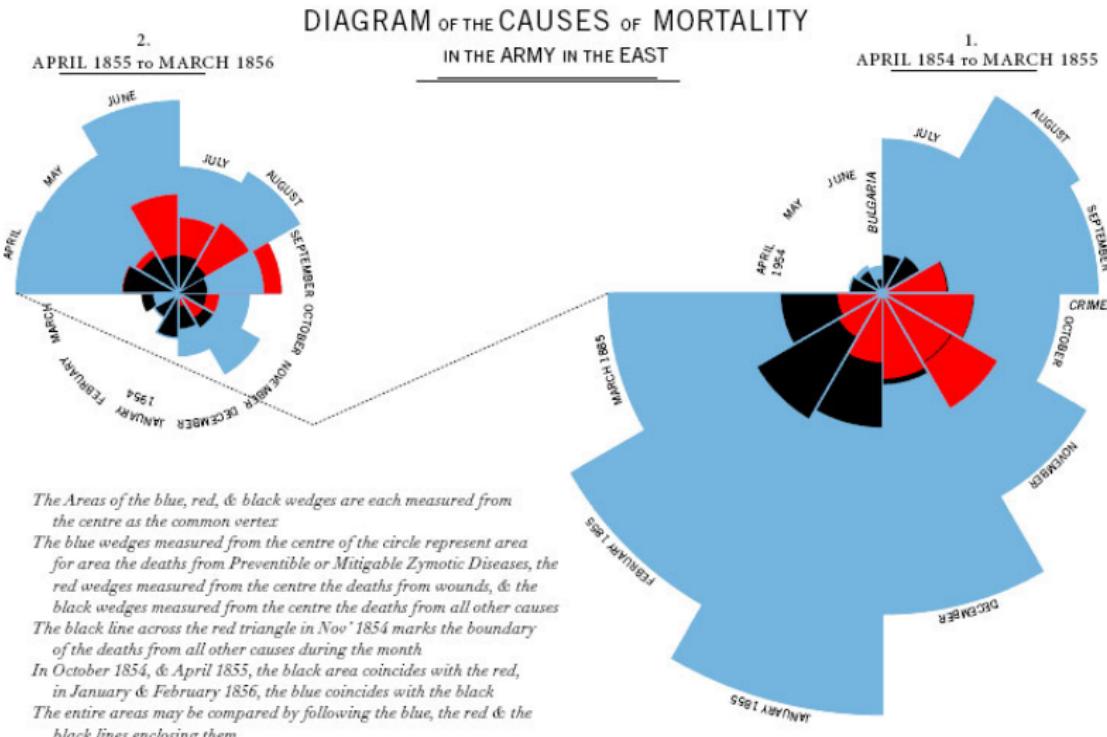
Information Visualisation

- ▶ Data is almost always multi-variate / multi-dimensional
- ▶ Visualisation should tell the truth about the data (or nearly)

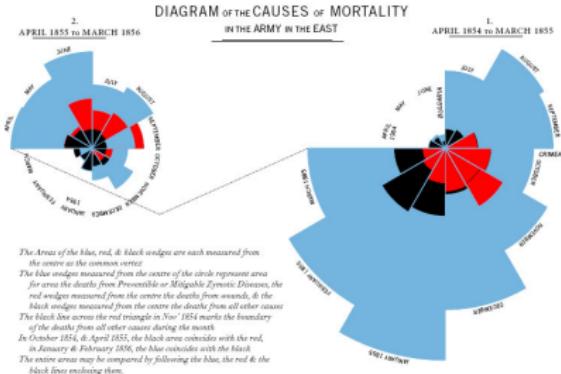
$$\text{Lie Factor (LF)} = \frac{\text{size of effect in visualisation}}{\text{size of effect in data}}$$

- ▶ $LF > 1 \rightarrow$ over stating
- ▶ $LF < 1 \rightarrow$ under stating
- ▶ $2 \leq LF \leq 5$ is common

Historical Note - Ex2



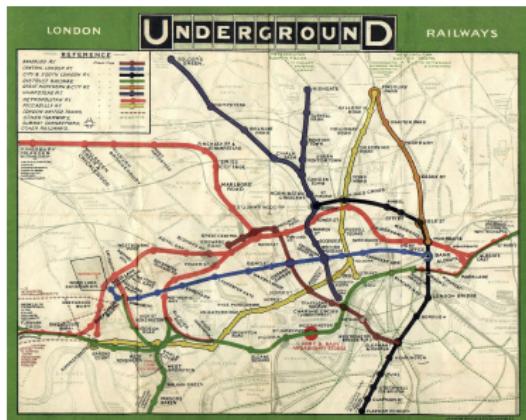
Historical Note - Ex2



- ▶ Florence Nightingale
 - ▶ Four dimensional data (date, disease1, disease2, disease3)
 - ▶ Histogram-style (area instead of height)
 - ▶ colours for the various dimensions (ordered max to min)

Historical Note - Ex3

- ▶ Harry Beck (1933)

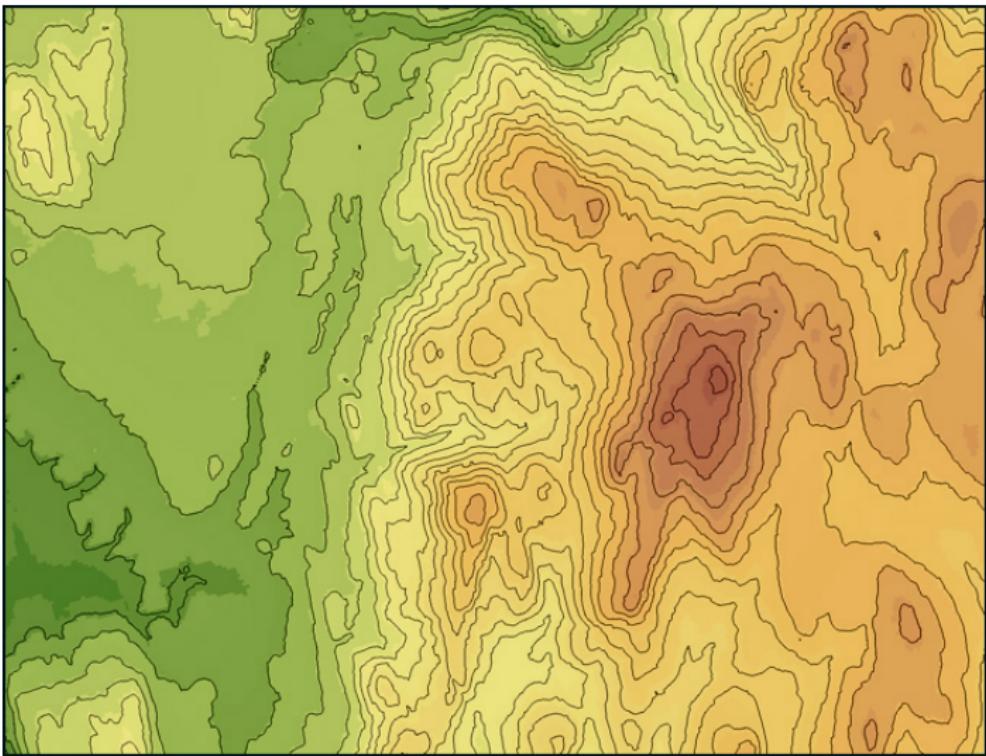


Dima Damen

Dima.Damen@bristol.ac.uk

COMS21202: Information Visualisation

Contour Maps

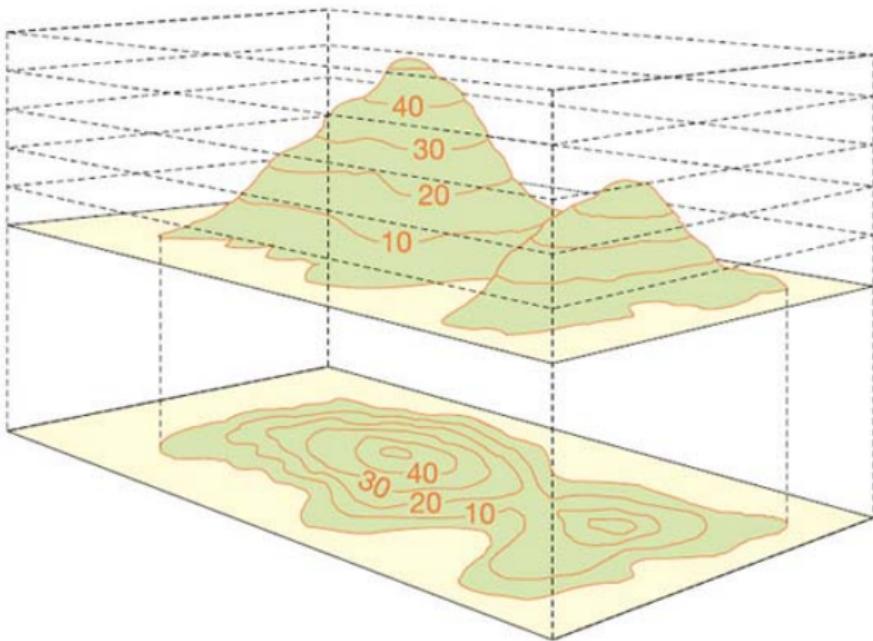


Dima Damen

Dima.Damen@bristol.ac.uk

COMS21202: Information Visualisation

Contour Maps



source: ordnancesurvey.co.uk

Flow Charts

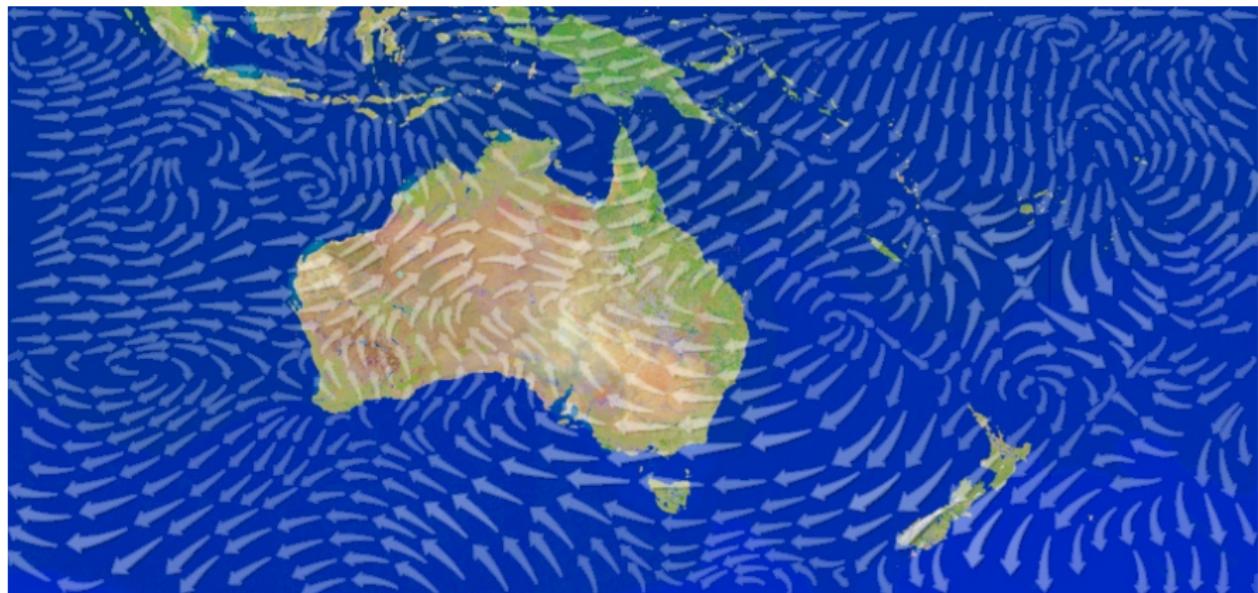
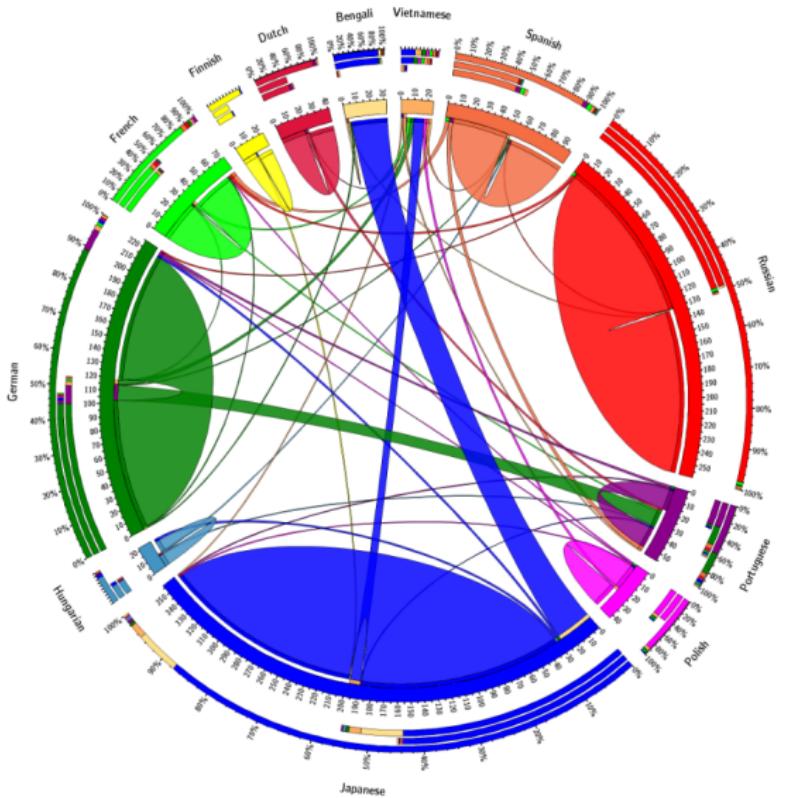


Image courtesy of Greg Turk

Circular Graphs



Source: AlNoamany et al (2013). Who and What Links to the Internet Archive

Dima Damen

Dima.Damen@bristol.ac.uk

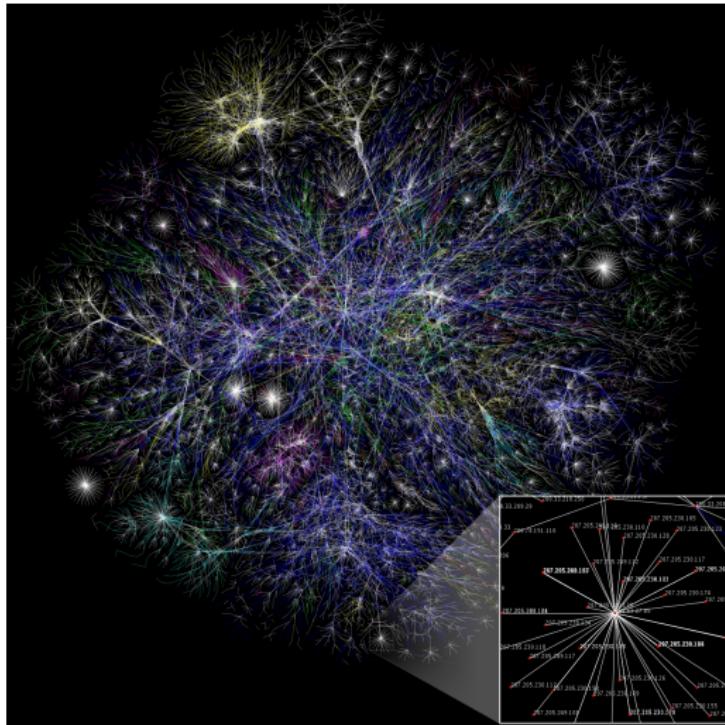
COMS21202: Information Visualisation

Wordles



Source: Wikipedia(2015)

Graphs and Hierarchies



Source: Wikipedia(2015)

Dima Damen

Dima.Damen@bristol.ac.uk

COMS21202: Information Visualisation

Further Reading

- ▶ **Visualization Handbook**
Hansen and Johnson (2004)