

## **MA/MSc Digital Journalism & MA/MSc Creating Social Media - Goldsmiths, University of London**

### **Digital Research Methods: Statistical Data Mining (SDM)**

#### **Dr Daniel Stamate, Department of Computing**

This segment of the course introduces students to elements of statistical computing including summary statistics (mean, mode, median, quartiles, variance, standard deviation), visual data representation for data repartition and patterns (stem and leaf plots, box plots, histograms, scatter plots, pie charts), measures of association in data (covariance, correlation, determination), linear predictive and estimation models (simple and multiple linear regression), and elements of hypothesis testing.

Students are also exposed to data mining concepts and techniques including data pre-processing (treatment of missing values, data transformation and feature reduction) classification (decision trees), and clustering (k-means, hierarchical methods).

The statistical and data mining concepts and techniques are illustrated with specific software when introduced in class. Analytic applications that support investigative activity (e.g. summarising, analysing and understanding data, finding interesting hidden patterns in data that can inspire a story, etc) and that employ introduced quantitative methods, are illustrated and tackled in practical sessions using specific software.

#### **Bibliography**

##### Essential material:

1. Data Mining: A Tutorial Based Primer, by R. Roiger et al., Addison Wesley, 2002
2. Discovering Statistics using SPSS, 3rd edition, by A. Field, Sage, 2009
3. SDM Module website, by D. Stamate, 2012 (online teaching resources)

##### Further reading:

4. SPSS QuickStarts, by N. Salkind et al., Prentice Hall, 2011
5. Data Mining: Concepts and Techniques, by J. Han et al., Morgan Kaufmann, 2005 (2nd edition) or 2011 (3rd edition)
6. Data Miner UK: Finding Stories in Data <http://datamineruk.wordpress.com/>
7. Introductory Statistics, by S. Ross, Academic Press, 2005 (2nd edition) or 2011 (3rd edition)

*SDM Module design, © Daniel Stamate, 2011*