14D007

3 ECTS

Data Visualization

Instructors Michael Greenacre & Ioannis Arapakis

Overview and Objectives

The course deals with the visualization of large data sets as a means of communicating relevant data patterns in the form of graphical displays, in order to interpret and understand the data. The course aims

1. to increase students' visual literacy, to sensitize students to the graphical elements that are used to create data visualizations, and to root the choice of graphical display in the context of the data and the audience for which it is intended,
2. to show methods of visualizing different data types: temporal data, spatial data, spatio-temporal data, categorical and textual data
3. to introduce a wide spectrum of multivariate statistical methods that extract essential patterns from high-dimensional data
4. to demonstrate web-based, interactive data visualization techniques for large-scale graphical applications,
5. to consider the various software options available for creative graphics.

Throughout the course various software options will be considered, particularly the R programming language.

Course Outline

Week 1

• Introduction to visual literacy • Simple charts, plots and information graphics • Graphical elements and colour theory • Aspects of design • Tufte's principles of scientific graphics

Week 2

• Visualization of temporal data • Visualization of spatial data • Visualization of spatio-temporal data

Week 3

• Multivariate distance and proximity • Clusters, trees and networks • Multidimensional scaling
• Dimension reduction and data maps • Data transformations • Principal component methods

Week 4

• Correspondence analysis • Categorical and textual data • Fundamentals of data-driven information visualization.

14D007

3 ECTS

Data Visualization

Week 5

• Dynamic and interactive visualization techniques • Big data visualization

Required Activities

Attendance at classes, and submission of homeworks.

Evaluation

• Homework during the course (50%) • A practical project at the end of the course (50%).

Pre-course reading

Read the short extract "Aesthetics and Technique in Data Graphical Design" from Edward Tufte's highly recommended book *The Visual Display of Quantitative Information* (Cheshire Press) at this link:

www.econ.upf.edu/~michael/visualdata/tufte-aesthetics_and_technique.pdf

Materials

A website dedicated to the course will provide links to relevant material on the web (visualization examples, readings, videos) as well as the class material and assignments.

For an introduction to the visualization of multivariate data, you can consult the following books, available for free download:

Greenacre, M. (2010) *Biplots in Practice*. BBVA Foundation, Madrid. Download from

www.multivariatestatistics.org

James, G., Witten, D, Hastie, T. and Tibshirani, R. (2013). *An Introduction to Statistical Learning, with Applications in R*. Download from

www.stanford.edu/~hastie/local.ftp/Springer/ISLR_print1.pdf

Supporting reading

Zettl, H. (2011). *Sight Sound Motion. Applied Media Aesthetics. Sixth Edition*. Wadsworth.

Blasius, J. and Greenacre, M. (2014). *Visualization and Verbalization of Data.* Chapman & Hall / CRC.

Greenacre, M. (2007). *Correspondence Analysis in Practice, 2nd edition.* Chapman & Hall / CRC.

Cook, D. and Swayne, D.F. (2007). *Interactive and Dynamic Graphics for Data Analysis.* Springer UseR! Series

Ward, M.O., Grinstein, G. and Keim, D. (2015). *Interactive Data Visualization, 2nd edition.* Chapman & Hall/CRC.

14D007

Data Visualization

3 ECTS

Murray, S. (2013). *Interactive Data Visualization for the Web*. O'Reilly.

Zhu, N.-Q. (2013). *Data Visualization with D3.js Cookbook.* Packt Publishing.