

Multivariate methods, spring 2012

Multivariate methods, spring 2012

Lecturer

[Kimmo Vehkalahti](#)

Type

Suitable for students of [REMS](#) as well as other students of [Social Sciences](#), including Statistics.

Prerequisites

You should have fairly good skills in the following topics **before** this course begins:

6 cu:

- basic skills of univariate data analysis using suitable software, such as **Survo, SPSS, R, Muste, or SAS**
- basic concepts of statistics and probability (e.g. Introduction to Statistics and Second Course in Statistics)

8 cu (only for Advanced studies of Statistics): as above, but also

- basic concepts of matrix algebra and mathematical analysis
- basic concepts of statistical inference and linear models

Description

The aim of the course is to learn the basics of **multivariate data analysis** and **multidimensional statistical modeling** in practice. **The focus will be on applications in Social Sciences.**

i For those who are looking for a more mathematical treatment of these topics, there is another course called [Unsupervised Machine Learning \(UML\)](#), which is organized jointly with the [Department of Computer Science](#) and the [Department of Mathematics and Statistics](#). On this course, you may consider the 8 cu option, i.e., doing a final report, see "Completion" below.

The focus of this course is on **practical working with real data** from [Social Sciences](#), learning to apply and interpret various multivariate methods, such as

- Factor Analysis
- Clustering methods
- Discriminant Analysis
- Multidimensional Scaling
- Correspondence Analysis

Schedule

Period IV (weeks 11-17), in **City Center Campus**.

⚠ PLEASE NOTE THE CHANGED SCHEDULE ON TUESDAYS!

- **Lectures:** Tue 8-10 and Wed 8-10 in [Unioninkatu 37, sh 4 \(2nd floor\)](#)
- **Computer class:** Wed 10-12 in [Aleksandria, K130](#)
 - **Exception:** on Mar 21 the computer class is in [Aleksandria, K131](#)

Completion

There is no exam. Instead, the course is completed (and graded 0-5) by active participation on lectures and by doing

1. **Exercises** (points depend on activity, **weekly deadlines**)
a shared workspace ([BSCW](#)) will be used
2. **Net poster** (compulsory, **deadline 29 April 2012**)
see [posters from previous courses](#)
3. **Final report** (for Advanced studies of Statistics only, **deadline 20 May 2012**)
(the topic of the report has to be agreed with the lecturer)

Registration

Note: the number of participants is limited to 24.

Data sets

The data sets will be found on BSCW. Own data sets may be used as well. Here you can see some general information of the data sets to be used in the exercises and posters:

- **European Social Survey**
 - <http://ess.nsd.uib.no/ess/round4/>
- **Economic Freedom**
 - <http://www.heritage.org/Index/>
- **Prices and Earnings around the Globe**
 - <http://www.ubs.com/research>
 - <http://www.macrofocust.com/public/products/infoscope/datasets/pricesandearnings/>

Books and websites (for 6 cu)

A selection of SUPPORTING material for the lectures and exercises:

- Hair Jr, Joseph F.; Anderson, Rolph E.; Tatham, Ronald L. & Black, William C. (1998). *Multivariate Data Analysis*. Fifth Edition, Prentice Hall.
- StatSoft, Inc. (2011). *Electronic Statistics Textbook*. StatSoft, Tulsa, Oklahoma.
- Stevens, James P. (2002). *Applied Multivariate Statistics for the Social Sciences*. Fourth Edition, Lawrence Erlbaum, Mahwah, New Jersey.

Suomeksi (in Finnish):

- [Menetelmäopetuksen tietovaranto](#), FSD (Yhteiskuntatieteellinen tietovarasto).
- Mustonen, Seppo (1995). *Tilastolliset monimuuttujamenetelmät*. Survo Systems, Helsinki.
- Nummenmaa, Tapio; Konttinen, Raimo; Kuusinen, Jorma & Leskinen, Esko (1996). *Tutkimusaineiston analyysi*. WSOY, Porvoo.
- Vehkalahti, Kimmo (2008). *Kyselytutkimuksen mittarit ja menetelmät*. Tammi, Helsinki.

Books and websites (for 8 cu)

Examples of ADDITIONAL material for writing the final report (Advanced Studies in Statistics):

- Chatfield, Christopher & Collins, Alexander J. (1980). *Introduction to Multivariate Statistics*. Chapman & Hall.
- Everitt, Brian (2005). *An R and S-PLUS Companion to Multivariate Analysis*. Springer.
- Everitt, Brian (2009). *Multivariable Modeling and Multivariate Analysis for the Behavioral Sciences*. Chapman & Hall/CRC.
- Flury, Bernard (1997). *A First Course in Multivariate Statistics*. Springer.
- Johnson, Richard A. & Wichern, Dean W. (2002). *Applied Multivariate Statistical Analysis*, Fifth Edition, Prentice Hall.
- Krzanowski, W. J. (2000). *Principles of Multivariate Analysis*. Revised Edition, Oxford University Press.
- Raykov, Tenko & Marcoulides, George A. (2008). *An Introduction to Applied Multivariate Analysis*. Routledge.
- Seber, George A. F. (2004). *Multivariate Observations*. Reprint of First Edition (1984). Wiley.
- StatSoft, Inc. (2011). *Electronic Statistics Textbook*. StatSoft, Tulsa, Oklahoma.
- Tabachnick, Barbara G. & Fidell, Linda S. (1996). *Using Multivariate Statistics*. Third Edition, HarperCollins.

Some more specialized books for Advanced Studies in Statistics:

- Cudeck, Robert & MacCallum, Robert C., eds. (2007). *Factor Analysis at 100: Historical Developments and Future*. Lawrence Erlbaum.
- Greenacre, Michael (2007). *Correspondence Analysis in Practice*, Second Edition, Chapman & Hall/CRC.
- Greenacre, Michael (2010). *Biplots in Practice*. BBVA Foundation, Madrid, Spain.
- Greenacre, Michael & Blasius, Jörg, eds. (2006). *Multiple Correspondence Analysis and Related Methods*. Chapman & Hall/CRC.
- Gower, J. C. & Hand, D. J. (1996). *Biplots*. Chapman & Hall.
- Heck, Ronald H. & Thomas, Scott L. (2009). *An Introduction to Multilevel Modeling Techniques*, Second Edition. Routledge.
- Mulaik, Stanley A. (2009). *Foundations of Factor Analysis*, Second Edition. Chapman & Hall/CRC.
- Seber, George A. F. (2008). *A Matrix Handbook for Statisticians*. Wiley.