

Amsterdam Business Research Institute

Multivariate Data Analysis

For Business and Management Research

Course Manual (Version 1.0)

Academic Year 2025 – 2026



Faculty

Prof. dr. Hester van Herk h.van.herk@vu.nl Course coordinator & lecturer

Target groups

The course Multivariate Data Analysis for Business and Management Research (MDA-MBR) 2024-2025 is open to PhD candidates, research master students, postdocs, and junior faculty. This is an advanced course that assumes prior knowledge on quantitative research methods and R (e.g., the ABRI course R for Business and management Research). The course workload represents 5 ECTS.

Course Content

This course will emphasize understanding, implementation, and interpretation of multivariate statistical methods. The course will involve both lectures and work tutorials. First, we discuss how to clean data, recap some fundamental multivariate techniques such as analysis of (co)variance and regression analysis and go in-depth into factor analysis. Second, the main part of the course will be focused on learning some more advanced MDA techniques such as confirmatory factor analysis, structural equation modeling (SEM) and multi-level models (MLMs, also known as linear mixed models, hierarchical linear models, or mixed-effect models). We end with a short introduction into recent developments in the field, including network analysis.

In the course high-quality existing datasets will be introduced such as the European Social Survey (ESS), the European Values Survey (EVS), the World Values Survey (WVS) and the Dutch LISS panel data.

This course prepares the student for analyzing datasets using the freely available programming language R. R is a platform for which many scholars write packages. The basis enables you to manipulate data, clean data, and test hypotheses. The packages enable you to use advanced methods such as structural equation modeling and multilevel modeling. You will learn how to read various datasets (including files from SPSS or SAS) into R, and how to conduct manipulations such as transforming data.

Learning Objectives

Upon successful completion of the course, students will:

- Be able to evaluate and generate high-quality quantitative data, including making meaningful decisions in handling missing data and outliers.
- Critically assess the strengths and weaknesses of various multivariate analysis techniques, such as multiple regression, PCA, FA, structural equation modeling, and multilevel modeling.
- Perform multivariate analysis using various R-packages and apply them to their research projects.
- Effectively interpret and report the output of multivariate analyses.
- Demonstrate insights into reporting and publishing studies using these techniques in top-tier academic journals.

Course Design

The course consists of lectures, hands-on tutorials, student-led presentations, and discussions alternated with generic reflections on the materials when needed. It is very important that you apply what you have learned in the

lectures and the tutorials. Practice is vital to understand the techniques learned and to be able to apply it yourself. We will meet twice a week. Each week, there will be homework including conducting multivariate analyses and interpreting the results which will be presented and discussed during the tutorials. Please prepare the tutorials by reading the theory and the assignment, and by downloading the data in advance.

During the course you will work on two individual assignments. During week 1-2 you will work on the first assignment that will cover the content of the first two weeks. This assignment is handed in at the beginning of the third week (Monday 23:59) and will be discussed in the second class in the third week. The second assignment covers the contents of all weeks, with a focus on weeks 3-6. This assignment is handed in at the end of week 8 (Friday 23:59).

Assignment

Your overall course grade is based on two assignments and one presentation. During the course, you will work on the assignments. The assignments focus on performing the analyzes, interpreting the results, and being able to communicate about the results, i.e., writing a methods section in a scientific paper. The presentation requires you to interpret the methods and results session of selected papers and lead meaningful discussion. The theories behind the different models you estimate are not part of the assignments.

You submit the first assignment and the next lecture you present your results in class followed by a discussion of the findings. The first assignment will be graded by the lecturer. The second assignment is handed in after lectures and will be graded by the lecturer.

Grading

20%	Presentation and Interpretation of article results sections and R-output in lectures
30%	Final grade assignment 1
50%	Final grade assignment 2

The overall course grade needs to be 5.5 or higher.

The grading of the course consists of presentations and participation in class and two assignments. The final version of the second assignment is due Friday March 27, 2026.

Tuition Fee

Tuition fee (5 ECTS Course)	€1250
20% discount on early bird registration	€1000

Workload and credits

The estimated time participants spend on study activities is:

Attending Lectures and Interactive Assignment Sessions	24 hours
Studying Literature	60 hours
Peer review	16 hours
Completing the Assignments	40 hours
<i>Total</i>	<i>140 hours (5 ECTS)</i>

Session overview Year (in progress)

Week	Dates	Topic	Literature	R packages	Assignments due
1	3 Feb: 13:30 – 17:30 HG-08A37	Data cleaning, missing data analysis, assumptions. Regression analysis PCA and Factor analysis	T&F: Chapter 4 (data Cleaning) 5 (Regression) C&F: Chapters 1-3, 4,7, 8 T&F: Chapter 13 (PCA and FA) OR E&H: Chapters 3 and 5	tidyverse, psych, FactoMineR, factoextra	
2	10 & 12 Feb: 13:30–15:15 HG-02A16	Confirmatory Factor Analysis	Rosseel (2012 & 2020) E&H: Chapter 7 (7.1 – 7.3) C&F: Chapter 10 Beaujean (2014) Chapters 3	psych, lavaan, semPlot, GPArotation	Homework: PCA-CFA (Tuesday) Student presentation on article 1 (Thursday).
3	17 & 19 Feb: 13:30 – 15:15 HG-02A16	Structural Equation Modeling (SEM)	Rosseel (2012 & 2020) Iaccobucci (2009, 2010)	lavaan, semTools, semPlot	Hand in assignment 1 (Monday) Student presentation on article 2 (Thursday)
4	24 & 26 Feb: 13:30 – 15:15 HG-13A24	Structural Equation Modeling (SEM) Measurement invariance	Rosseel (2012 & 2020) Iaccobucci (2009, 2010) Beaujean (2014) Chapter 2, 3, 4	lavaan, semTools	Homework: CFA (Tuesday) Student presentation on article 3 (Thursday)
5	3 & 5 March: 13:30 – 15:15 HG-12A32	Multi-level models Longitudinal SEM	C&F: Chapter 9 (9.3.1 + 9.3.5-9.3.7) T&F: Chapter 15 (additional reading) Beaujean (2014) Chapter 5	lavaan, semTools	Homework: SEM modelling (Tuesday) Student presentation on article 4 (Thursday)
6	10 & 17 March: 13:30 – 15:15 HG-12A32	Related techniques and developments: PLS-SEM, Network Analysis	Selected articles by Hair and Borsboom (see further reading for some more introductory articles)	SEMinR, psychometrics	Homework: SEM modelling (Tuesday) Reflection and discussion on assignment 2 and the overall course (Thursday)
7		Office hour (on appointment)			Office hour on assignments
8	March 27	Deadline Assignments			Hand-in final assignment + revised assignment 1

In-class Presentations

Instructions for Presentations:

Two to three students will be assigned in one group to present one of the following articles. In the presentation Each group will present for 15 minutes and then lead a 15-minute discussion based on the presented materials. The presentation will focus on the methodological and reporting results of the data analysis sections of the paper.

Articles for student presentations (in order of presentation):

1. Ye, Q., Zhu, Y., Jin, Y., & Wang, D. (2023). Will there always be a return on investment? The effects of investment in employee development on employee entrepreneurship. *Journal of Vocational Behavior*, 141, 103843. <https://doi.org/10.1016/j.jvb.2023.103843>
2. Laguna, M., Mielniczuk, E., Razmus, W., Moriano, J. A., & J. Gorgievski, M. (2017). Cross-culture and gender invariance of the Warr (1990) job-related well-being measure. *Journal of Occupational and Organizational Psychology*, 90(1), 117-125. <https://doi.org/10.1111/joop.12166>
3. Farkas, B. C., Chambon, V., & Jacquet, P. O. (2022). Do perceived control and time orientation mediate the effect of early life adversity on reproductive behaviour and health status? Insights from the European Value Study and the European Social Survey. *Humanities and Social Sciences Communications*, 9(1), 52. <https://doi.org/10.1057/s41599-022-01066-y>
4. Leijen, I., van Herk, H., & Bardi, A. (2022). Individual and generational value change in an adult population, a 12-year longitudinal panel study. *Scientific Reports*, 12(1), 17844. <https://doi.org/10.1038/s41598-022-22862-1>

Literature

Books

Tabachnick, Barbara G., and Linda S. Fidell (2014). *Using Multivariate Statistics*, Pearson New International Edition. (for sale at bol.com; at AMAZON as printed book and as an e-book on Kindle); a good reference on MVA

Everitt, B., & Hothorn, T. (2011). *An introduction to applied multivariate analysis with r* (Ser. Use r!). Springer. <https://doi.org/10.1007/978-1-4419-9650-3> (e-book at VU library)

Beaujean, A. A. (2014). *Latent variable modeling using R: a step-by-step guide*. Routledge, Taylor & Francis Group. <https://doi-org.vu-nl.idm.oclc.org/10.4324/9781315869780> (e-book at Vu library)

Chapman, C., & Feit, E. M. D. (2019). *R for marketing research and analytics* (Second, Ser. Use R!). Springer. <https://link-springer-com.vu-nl.idm.oclc.org/book/10.1007%2F978-3-030-14316-9> (e-book at Vu library)

Articles (can all be downloaded from UBVU or Google Scholar)

Iacobucci, D. (2009). Everything you always wanted to know about SEM (structural equations modeling) but were afraid to ask. *Journal of Consumer Psychology*, 19(4), 673-680.

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20(1), 90-98.

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., ... & Chow, M. (2017). *Package 'lavaan'*. Retrieved July, 31, 2020.

Rosseel, Y. (2012). Llavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of statistical software*, 48(2), 1-36.

Additional reading (textbooks)

Hoyle, R. H. (2014). *Handbook of Structural Equation Modeling*. Guilford Publications.

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel modeling using R*. Crc Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). sage.

Further reading

- Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, 1(1), 58.
<https://doi.org/10.1038/s43586-021-00055-w>
- Fischer, R., & Karl, J. A. (2025). Measurement Invariance of Human Values Across Time and Countries—An Analysis Using 11 Rounds of the European Social Survey. *European Journal of Social Psychology*, n/a(n/a).
<https://doi-org.vu-nl.idm.oclc.org/10.1002/ejsp.70009>
- Graafland, J., & Schilpzand, A. (2025). Gender Equality and Life Satisfaction: A Mediation Model with Individual Autonomy, Income Per Capita and Trust. *Journal of Happiness Studies*, 26(3), 23. <https://doi.org/10.1007/s10902-024-00850-8>
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, 19(2), 139-152.
<https://doi.org/10.2753/MTP1069-6679190202>
- Henseler, J., Schuberth, F., Lee, N., & Kemény, I. (2025). Why researchers should be cautious about using PLS-SEM. *Industrial Marketing Management*.
<https://doi.org/https://doi.org/10.1016/j.indmarman.2024.01.017>
- Legate, A. E., Ringle, C. M., & Hair Jr, J. F. (2024). PLS-SEM: A method demonstration in the R statistical environment. *Human Resource Development Quarterly*, 35(4), 501-529.
<https://doi.org/https://doi.org/10.1002/hrdq.21517>