Multivariable Data Analysis

Code: 101148
ECTS Credits: 6

2025/2026

Degree	Type	Year
Sociology	OB	3

Contact

Name: Jose Pedro Lopez Roldan

Email: pedro.lopez.roldan@uab.cat

Teaching groups languages

You can view this information at the end of this document.

Prerequisites

In order to be able to take this course, it is advisable to have successfully followed up the Quantitative Methods of Social Research and Analysis Methods.

Objectives and Contextualisation

This is an introductory course to the techniques of multivariate statistical data analysis that is proposed as a continuation of the quantitative perspective of social research initiated in the degree. The procedures, methods and techniques already discussed so far will be expanded to consider what we can generally call the transition from bivariate analysis procedures to multivariate analysis procedures.

In the context of the itinerary of technical and methodological subjects, which seek to offer a complete overview of the different procedures of the sociological scientific activity, and given the extension and variety of the analysis procedures in the field of social sciences, it entails directing the teaching towards the selection of a few topics or instruments considered as some of the most fundamental and of greatest interest in the practice of sociological research.

Specifically, the subject aims to:

1) From the point of view of the students, the construction of their learning will be carried out from:

- Knowledge and understanding of the main concepts associated with the multivariate analysis of statistical data, exemplified by sociological concepts.
- The ability to apply technical instruments for the advanced analysis of statistical data considered in the course.
- Know how to use statistical software for statistical analysis bivariate and multivariable.
- Know how to interpret the statistical results of a data analysis from the technical and substantive point of view according to some knowledge and study objectives of the social reality.

2) From the general conditions of a subject of this type in relation to the use of students it is about:

- Facilitate the understanding, management and interpretation of a basic algebraic and statistical conceptual system to assimilate the use of techniques that involve the quantification and formalization of social

phenomena.
- Framing in a balanced, comprehensive and integratingway the contents ofthis subject within the set of the usual methods in sociology.

Competences

- Applying the main quantitative and qualitative methods and techniques of social research to a specific topic.
- Describing social phenomena in a theoretically relevant way, bearing in mind the complexity of the involved factors, its causes and its effects.
- Designing a social research project by defining a comprehensive theoretical framework with clearly defined concepts, formulating consistent and significant hypothesis, choosing suitable investigation techniques for the adopted concepts, and analysing the empirical results obtained with those techniques.
- Developing critical thinking and reasoning and communicating them effectively both in your own and other languages.
- Developing self-learning strategies.
- Enumerating the methodology and investigation techniques that support the main hypothesis about social relationships, the positions and practices of individuals in a social structure and the social changes.
- Searching for documentary sources starting from concepts.
- Students must be capable of assessing the quality of their own work.
- Students must be capable of managing their own time, planning their own study, managing the relationship with their tutor or adviser, as well as setting and meeting deadlines for a work project.
- Working in teams and networking in different situations.

Learning Outcomes

1. Defining concepts of analysis.
2. Developing critical thinking and reasoning and communicating them effectively both in your own and other languages.
3. Developing self-learning strategies.
4. Explaining the methodological basis of these quantitative and qualitative methods and techniques.
5. Formulating a hypothesis with these concepts.
6. Identifying the main quantitative and qualitative methods and techniques.
7. Indicating their dimensions, their possible quantitative indicators and the significant qualitative evidence in order to empirically observe them.
8. Measuring a social phenomenon with these instruments on the basis of a theoretical framework of analysis.
9. Mentioning the main concepts of sociology.
10. Obtaining conclusions from the information obtained with this tool.
11. Preparing an analytical tool that is significant to this hypothesis.
12. Relating them with the different approaches of sociology.
13. Searching for documentary sources starting from concepts.
14. Students must be capable of assessing the quality of their own work.
15. Students must be capable of managing their own time, planning their own study, managing the relationship with their tutor or adviser, as well as setting and meeting deadlines for a work project.
16. Using the advanced multivariate statistical tools.
17. Using the appropriate software to the advanced multivariate statistical tools.
18. Using the appropriate software to the basic multivariate statistical tools.
19. Using the appropriate software to the univariate statistical tools.
20. Using the basic multivariate statistical tools.
21. Using the univariate statistical tools.
22. Working in teams and networking in different situations.

Content

General introduction
- Objectives of the subject, contents, course dynamics and evaluation
- Multivariate analysis: characteristics and classification of techniques
- Software for the analysis of statistical data

PART I. Analysis of interdependence with qualitative variables

Unit 1. Analysis of contingency tables
- Classic analysis of multidimensional contingency tables

Unit 2. Log-linear analysis
- General linear logarithmic analysis

PART II. The dependence analysis

Unit 3. Analysis of variance
- One-way analysis of variance
- Analysis of multivariate variance

Unit 4. Regression analysis
- Simple regression analysis
- Multiple regression analysis

PART III. The analysis of interdependence for the construction of typologies

Unit 5. Factor analysis
- Mathematical foundations of multivariate data analysis
- Factor analysis of principal components
- Factor analysis of correspondences

Unit 6. Cluster analysis
- Cluster analysis and the construction of typologies
- Automatic cluster analysis

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Classroom practices	16	0.64	1, 11, 5, 8, 10, 18, 20, 16
Individual preparation of practical exercises	30	1.2	3, 4, 8, 10, 18, 20, 16
Master classes	30	1.2	1, 11, 9, 4, 5, 6, 7, 8, 18, 20, 16
Type: Supervised			
Programmed individual tutorials	2	0.08	14, 13, 1, 11, 5, 15, 8, 10, 22, 18, 20, 16
Type: Autonomous			

Individual work	30	1.2	13, 1, 11, 5, 15, 8, 10, 22, 18, 20, 16
Readings	30	1.2	4, 6, 8, 18, 20, 16

The course is presented with a continuous dynamic of teaching and learning, which implies tracking the rhythms of the course and the various contents that have been designed in accordance with the different scheduled activities. The contents of each unit have a thread linked to the research process and the continuity of the learning of concepts and instruments that are incorporated progressively, as well as the resolution of problems and questions, which are based in the assimilation and practice of each previous topic of each unit.

Since the objective of the training is that students learn to research sociology using advanced statistical techniques, the teaching methodology and the training activities of the subject result from the combination of expositive sessions with problem solving exercises and practices in the classroom that allow to apply the acquired concepts and explained techniques, as well as tutorials of follow-up and autonomous work.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Analysis works	90,0%	0	0	14, 13, 1, 3, 2, 11, 9, 4, 5, 15, 6, 7, 8, 10, 12, 22, 17, 18, 19, 20, 16, 21
Practical exercices	10,0%	12	0.48	14, 13, 1, 3, 2, 11, 9, 4, 5, 15, 6, 7, 8, 10, 12, 22, 17, 18, 19, 20, 16, 21

The course is continuously evaluated. Regular attendance at class sessions is important to ensure correct learning and assessment.

A final weighted average mark of 5 out of 10 or more is required for the evaluation activities.

The assessment combines three aspects

1) The analysis papers (88%): 2 works, carried out in groups of 2 people, of sociological analysis of quantitative data with a database chosen by the students and with the help of the software, in relation to the following topics:

1. Analysis of multidimensional and log-linear contingency tables
a) Work tutorial (2%)
b) Selection of data and formulation of hypotheses (7%)
c) Complete analysis work (35%)

2. Typological analysis combining factorial and classification analysis
a) Work tutorial (2%)
b) Selection of data and formulation of hypotheses (7%)
c) Complete analysis work (35%)
- A minimum grade of 5 out of 10 of each work is required.
- They will be elaborated in content and form as an academic research article according to a defined format

and with a maximum extension of 3,500 words of writing, separate annexes.
- It will be necessary to do a mandatory tutorial, face-to-face or by Teams, with the two members of the group, before the first delivery of each work, in order to guide its correct realization and comment on the chosen data, the hypotheses formulated and the first analysis based on the previous preparation work of the students of the group.
- The works with a grade lower than 5 can be recovered at any time and before the date set in the evaluation calendar. The work will have to be re-presented with the corrections made and a page will be attached at the end of the work document where the corrections made in relation to the work initially presented will be briefly explained. The recovery will be scored at a maximum of 7.
- Failure to present the work, the presentation after the deadline without justification prior to delivery, the incomplete presentation of the analysis or not doing the tutoring of the work, will mean a non-presentation of the work and, depending on the case, the non-evaluation or the suspension of the subject.

2) The practices (12%): there will be 6 practical sessions that will consist of the realization in the computer room of exercises of application of the techniques of analysis of data treated:
1. Statistical software
2. Analysis of multidimensional and log-linear contingency tables
3. Variance analysis
4. Regression Analysis
5. Factorial analysis of main components and multiple correspondences
6. Classification analysis
The evaluation of the activity will be the result of the grade obtained in a questionnaire of questions about each practical exercise. The practices will be done in pairs (the two people of the work groups). If the practice is not done in the classroom on the scheduled day for a justified reason, it can be recovered before a week passes and the grade will be 5. Outside of this period, it will be considered not completed and the grade will be 0.

3) Additionally, if a follow-up of the subject is higher than 75% in relation to all the activities proposed throughout the course, including the specific activities of each day of class, you can choose to add up to 0.5 points in the final grade of the subject.

Criteria by which "Not evaluable" will be assigned: the non-presentation of the two works.

In this subject, the use of Artificial Intelligence (AI) technologies is allowed as an integral part of the development of the work, provided that the final result reflects a significant contribution of the student in the analysis and personal reflection. The student must clearly identify which parts have been generated with this technology, specify the tools used and include a critical reflection on how these have influenced the process and the final result of the activity. The lack of transparency in the use of AI will be considered a lack of academic honesty and may lead to a penalty in the grade of the activity, or greater sanctions in serious cases.

Bibliography

Basic bibliography
López-Roldán, P.; Fachelli, S. (2015). *Metodología de la investigación social cuantitativa*. Bellaterra (Barcelona): Dipòsit Digital de Documents, Universitat Autònoma de Barcelona. 1a. edición.
http://ddd.uab.cat/record/129382 | http://pagines.uab.cat/plopez/content/misc
López-Roldán, P. (2015). *Recursos para la investigación social*. Dipòsit Digital de Documents. Bellaterra (Barcelona): Universitat Autònoma de Barcelona.
http://ddd.uab.cat/record/89349 | http://pagines.uab.cat/plopez

Further reading
The manual *Metodología de la investigación social cuantitativa* (MISC) contains in each chapter a list of specific bibliographic references that complement the basic bibliography..

Selected bibliographical references:

Aldas, J.; Uriel, E. (2017). *Análisis multivariante aplicado con R* (2.ª ed.). Madrid: Paraninfo

Ato García, M.; López García, J. J. (1996). *Análisis estadístico para datos categóricos*. Madrid: Síntesis.

Bailey, K. D. (1994). *Typologies and Taxonomies. An Introduction to Classification Techniques*. Thousand Oaks (California): Sage.

Brown, B. L.; Hendrix, S. B.; Hedges, D. W.; Smith, T. B. (2011). *Multivariate analysis for the biobehavioral and social sciences. Agraphical approach*. Hoboken: John Wiley & Sons.

Cea d'Ancona, M. A. (2002/2014). *Anàlisis multivariable. Teoría y práctica en la investigación social*. Madrid: Síntesis.

Christensen, R. R. (1997). *Log-linear models and logistic regression.* New York: Springer-Verlag.

Correa Piñero, A. D. (2002). *Análisis logarítmico lineal*. Madrid: La Muralla.

Everitt, B.; Hothorn, T. (2011). *An introduction to applied multivariate analysis with R*. New York: Springer.

Greenacre, M. J. (2008). *La práctica del análisis de correspondencias*. Madrid: Fundación BBVA.

http://www.fbbva.es/TLFU/tlfu/esp/publicaciones/libros/fichalibro/index.jsp?codigo=300

García Ferrando, M. (1987). *Socioestadística. Introducción a la estadística en sociología*. 2a edició amp. Madrid: Alianza. Alianza Universidad Textos, 96.

Guillén, M. F. (1992). *Análisis de regresión múltiple*. Madrid: Centro de Investigaciones Sociológicas.

Hahs-Vaughn, D. L. (2017). *Applied multivariate statistical concepts*. Nueva York: Routledge.

Hair, J. F., Black, W. C.; Babin, B. J.; Anderson, R. E. (2013). *Multivariate data analysis*. Pearson new international edition (7.ª ed.). Harlow: Pearson.

Hernández Encinas, L. (2001). *Técnicas de taxonomía numérica*. Madrid: La Muralla.

Harlow, L. L. (2014). *The essence of multivariate thinking. Basic themes and methods* (2.ª ed.). Nueva York: Routledge.

Joaristi Olariaga, L.; Lizasoain Hernandez, L. (1999). *Análisis de correspondencias*. Madrid: La Muralla.

Lévy Mangin, J. P.; Varela Mallou, J. (2003/2008) *Análisis multivariables para las ciencias sociales*. Madrid. Pearson-Prentice Hall.

López-Roldán, P.; Fachelli, S. (2018). *Metodología de construcción de tipologías para el análisis de la realidad social*. Bellaterra (Cerdanyola del Vallès): Dipòsit Digital de Documents, Universitat Autònoma de Barcelona. 2a. edición.

MacFarland, T. W. (2012). *Two-Way Analysis of Variance: Statistical Tests and Graphics Using R*. New York: Springer.

Marradi, A. (1990). Classification, typology, taxonomy. *Quality & Quantity*, 24, 129-157.

Mateos-Aparicio, G.; Hernandez Estrada, A. (2021). An*alisis multivariante de datos: Cómo buscar patrones de comportamiento en Big Data*. Madrid: Pirámide.

Meneses, J. (2019). *Introducción al análisis multivariante*. Barcelona: UOC

Miller, J. E. (2013). *The Chicago guide to writing about multivariate analysis* (2.ª ed.). Chicago: The University of Chicago Press.

Pituch, K. A.; Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences* (6.ª ed.). Nueva York: Routledge.

Powers, D. A.; Xie, Y. (2008). *Statistical Methods for Categorical Data Analysis*. Bingley, U.K.: Emerald. 2a. edició.

Sánchez Carrión, J.J. (1999). *Manual de análisis estadístico de los datos*. Madrid: Alianza. Manuales, 055.

Sánchez Carrión, J. J. (Ed.) (1984). *Introducción a las técnicas de multivariable aplicadas a las ciencias sociales*. Madrid: Centro de Investigaciones Sociológicas.

Sánchez Carrión, J. J. (1989). *Análisis de tablas de contingencia. El uso de los porcentajes en ciencias sociales*. Madrid: Centro de Investigaciones Sociológicas-Siglo XXI.

Tabachnick, B. G.; Fidell, L. S. (2019). *Using multivariate statistics* (7.ª ed.). Nueva York: Pearson.

Tejedor, F. J. (1999). *Análisis de varianza: introducción conceptual y diseños básicos.* Madrid: La Muralla.

VV.AA. (1996). La construcció de tipologies. Exemples. Monogràfic de *Papers. Revista de Sociologia*, 48.

http://ddd.uab.cat/search?cc=papers&f=issue&p=02102862n48&rg=100&sf=fpage&so=a&ln=en

Software

The course will use the IBM SPSS Statistics software for statistical data analysis.

In addition, Moodle, MS-Office (Word, Excel) and Adobe Acrobat will be used.

Groups and Languages

Please note that this information is provisional until 30 November 2025. You can check it through this link. To consult the language you will need to enter the CODE of the subject.

Name	Group	Language	Semester	Turn
(SEM) Seminars	1	Catalan	first semester	morning-mixed
(SEM) Seminars	51	Catalan	first semester	afternoon
(TE) Theory	1	Catalan	first semester	morning-mixed
(TE) Theory	51	Catalan	first semester	afternoon