# Syllabus

## *Course Description*

| | |
|---|---|
| **Course Title** | Statistical Methods |
| **Course Code** | 27502 |
| **Course Title Additional** | |
| **Scientific-Disciplinary Sector** | STAT-01/A |
| **Language** | English |
| **Degree Course** | Master in Data Analytics for Economics and Management |
| **Other Degree Courses (Loaned)** | M1 Statistical methods for business analysis is loaned from course 25559 – Master in Entrepreneurship and Innovation (LM-77 EI)<br>M2 Advanced statistics is loaned from course 73006 – Master in Computing for Data Science (LM-18) |
| **Lecturers** | Prof. Alessandro Casa,<br>Alessandro.Casa@unibz.it<br>https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46549 |
| **Teaching Assistant** | |
| **Semester** | Second semester |
| **Course Year/s** | 1 |
| **CP** | 12 |
| **Teaching Hours** | M1: 36 hours<br>M2: 40 hours |
| **Lab Hours** | M1: 18 hours<br>M2: 20 hours |
| **Individual Study Hours** | - |
| **Planned Office Hours** | M1: 18 hours<br>M2: 18 hours |
| **Contents Summary** | M1:<br>This module begins with a review of key principles of statistical inference and then introduces core concepts in statistical learning. Topics include linear regression and its extensions, advanced regression techniques such as decision trees, logistic regression, classification methods, model selection strategies, and |

| | |
|---|---|
| | unsupervised learning approaches like principal component analysis and clustering. Throughout the course, students will work hands-on in R, applying techniques to real-world datasets drawn from business scenarios. By the end, students will be able to choose suitable statistical models, apply them to a range of business problems, and effectively communicate their analytical insights<br><br>M2:<br>• Parameter estimation: maximum likelihood methods<br>• Parameter estimation: Bayesian inference<br>• Time series: components and forecasting<br>• Time series: causal relationship tests<br>• Missing data<br>• Elements of statistics for Big Data |
| Course Topics | M1:<br>- Review of statistical inference: random variables, confidence intervals, and hypothesis testing.<br>- Introduction to statistical learning concepts: basic vocabulary and notions, parametric and nonparametric approaches, predictive and inferential objectives, bias-variance trade off, supervised and unsupervised learning<br>- Linear regression and extensions: simple and multiple linear regression, model estimation and assessment, model assumptions, inferential tools, qualitative predictors, interaction effects, polynomial regression, basic notions on nonparametric regression<br>- Classification: introduction to classification, logistic regression, model estimation, evaluation of classifiers<br>- Other supervised learning techniques: trees, splines, additive models<br>- Model selection/assessment and evaluation of model complexity: resampling methods, cross-validation and information criteria<br>- Unsupervised learning: clustering tools such as k-means and hierarchical clustering, principal component analysis<br>- Applications with the R software<br><br>M2:<br>- Parameter estimation: maximum likelihood methods<br>- Parameter estimation: Bayesian inference<br>- Time series: components and forecasting<br>- Time series: causal relationship tests |

| | |
|---|---|
| | - Missing data<br>- Elements of statistics for Big Data |
| **Keywords** | Statistical learning, regression, classification, clustering, dimensionality reduction, model selection , statistical inference, Bayesian statistics, time series modelling, big data |
| **Recommended Prerequisites** | M1: No formal prerequisites are required. Nonetheless, knowledge of basic concepts in descriptive and inferential statistics is useful, and attending a pre-course in mathematics/statistics is recommended.<br><br>M2: the knowledge provided by a course in calculus and one in probability and statistics |
| **Propaedeutic Courses** | |
| **Teaching Format** | M1:<br>In-person lectures and computer labs. Whenever possible, lectures will be structured to prioritize in-class time for discussions, and practical applications.<br><br>M2:<br>Frontal lectures, discussions and exercises on computer. |
| **Mandatory Attendance** | Recommended, but not required. |
| **Specific Educational Objectives and Learning Outcomes** | Intended Learning Outcomes (ILO)<br>M1:<br><br>ILO 1 Knowledge and understanding:<br>ILO 1.1 The student acquires knowledge of the analytical techniques and tools required to understand and quantitatively analyse economic and business phenomena in order to support decision-making processes.<br>ILO 1.2 The student consolidates knowledge of statistical inference, linear models and their generalisations, linear algebra, and optimisation techniques.<br>ILO 1.3 The student acquires an in-depth knowledge of the main techniques of supervised and unsupervised statistical learning, which are instrumental in the development of analysis and visualisation of economic and business data.<br><br>ILO 2 Applying knowledge and understanding: |

| | |
|---|---|
| | ILO 2.1 Ability to apply and implement analysis techniques focusing on different types of datasets such as streaming data, tabular data, documents and images and analysis on joint datasets.<br>ILO 2.2 Ability to apply supervised and unsupervised learning, and knowledge modelling, extraction, integration, analysis and exploitation; these skills are declined in various application domains of interest to companies and public and private organisations.<br><br>ILO 3 Making judgements:<br>ILO 3.1 The student acquires the ability to apply acquired knowledge to interpret data in order to make directional and operational decisions in a business context.<br>ILO 3.2 The student acquires the ability to apply acquired knowledge to support processes related to production, management and risk promotion activities and investment choices through the organisation, analysis and interpretation of complex databases.<br><br>ILO4 Communication skills:<br>ILO 4.1 The student acquires the ability to communicate effectively in oral and written form the specialised content of the individual disciplines, using different registers, depending on the recipients and the communicative and didactic purposes, and to evaluate the formative effects of his/her communication.<br><br>ILO 5 Learning skills:<br>ILO 5.1 The student acquires knowledge of scientific research tools. He/she will also be able to make autonomous use of information technology to carry out bibliographic research and investigations both for his/her own training and for further education. Furthermore, through the curricular teaching and the activities related to the preparation of the final thesis, she will be able to acquire the ability<br>- to identify thematic connections and to establish relationships between methods of analysis and application contexts;<br>- to frame a new problem in a systematic manner and to implement appropriate analysis solutions; |

| | |
|---|---|
| | - to formulate general statistical-econometric models from the phenomena studied.<br><br>M2:<br><br>ILO 1 Knowledge and understanding:<br>ILO 1.1 The student acquires knowledge of the analytical techniques and tools required to understand and quantitatively analyse economic and business phenomena in order to support decision-making processes.<br>ILO 1.2 The student consolidates knowledge of statistical inference, linear models and their generalisations, linear algebra, and optimisation techniques.<br>ILO 1.3 The student acquires an in-depth knowledge of the main techniques of supervised and unsupervised statistical learning, which are instrumental in the development of analysis and visualisation of economic and business data.<br><br>ILO 2 Applying knowledge and understanding:<br>ILO 2.1 Ability to apply and implement analysis techniques focusing on different types of datasets such as streaming data, tabular data, documents and images and analysis on joint datasets.<br>ILO 2.2 Ability to apply supervised and unsupervised learning, and knowledge modelling, extraction, integration, analysis and exploitation; these skills are declined in various application domains of interest to companies and public and private organisations.<br><br>ILO 3 Making judgements:<br>ILO 3.1 The student acquires the ability to apply acquired knowledge to interpret data in order to make directional and operational decisions in a business context.<br>ILO 3.2 The student acquires the ability to apply acquired knowledge to support processes related to production, management and risk promotion activities and investment choices through the organisation, analysis and interpretation of complex databases.<br><br>ILO4 Communication skills: |

| | |
|---|---|
| | ILO 4.1 The student acquires the ability to communicate effectively in oral and written form the specialised content of the individual disciplines, using different registers, depending on the recipients and the communicative and didactic purposes, and to evaluate the formative effects of his/her communication.<br><br>ILO 5 Learning skills:<br>ILO 5.1 The student acquires knowledge of scientific research tools. He/she will also be able to make autonomous use of information technology to carry out bibliographic research and investigations both for his/her own training and for further education. Furthermore, through the curricular teaching and the activities related to the preparation of the final thesis, she will be able to acquire the ability<br>- to identify thematic connections and to establish relationships between methods of analysis and application contexts;<br>- to frame a new problem in a systematic manner and to implement appropriate analysis solutions;<br>- to formulate general statistical-econometric models from the phenomena studied. |
| **Specific Educational Objectives and Learning Outcomes (additional info.)** | |
| **Assessment** | The overall exam mark will be determined by the assessment of the two modules (M1+M2).<br><br>M1:<br>Assessment (for both attending and non-attending students):<br>- Written Exam: Exercises and review questions (65% of the final grade)(ILOs 1.1, 1.2, 1.3, 3.1, 3.2, 5.1)<br>- Data Analysis Project: Group project in which students select and analyze an interesting dataset using the tools learned in the course. Groups will present their work at the end of the course (35% of the final grade; optional)(ILOs 2.1, 2.2, 3.1, 3.2, 4.1, 5.1).<br><br>Notes:<br>- For students who do not complete the project, the written exam will count for 100% of the final grade. |

| | |
|---|---|
| | - Project grades remain valid for one academic year.<br><br>M2:<br>The assessment is based on class and lab participation (ILOs 1.1, 1.2, 1.3, 3.1, 3.2, 4.1), home-work exercises (2.1, 2.2, 3.1, 3.2, 4.1, 5.1) and a final written exam (ILOs 1.1, 1.2, 1.3, 3.1, 3.2, 5.1). The final written exam will include open questions and exercises to be worked out by the students as well as computational exercises to be solved with R. |
| **Evaluation Criteria** | M1:<br>- Written exam: understanding of statistical concepts, correct interpretation of results of statistical analyses, clarity and precision of explanations.<br>- Data Analysis Project: Quality and clarity of the presentation, adequacy and appropriateness of analyses with respect to dataset characteristics<br><br>M2:<br>For attending students the final grade will be determined by the evaluation of homeworks, class and lab participation (20%) and the evaluation of a final written exam (80%).<br>The homeworks and the final written exam are separately evaluated with a score expressed in 30/30.<br>For non-attending students the final grade will be determined by the evaluation of a final written exam (100%). The final written exam is evaluated with a score expressed in 30/30. |
| **Required Readings** | **M1:**<br><br>James, G., Witten, D., Hastie, T., Tibshirani, R. An Introduction to Statistical Learning with Applications in R. Springer, 2013. Freely available at http://www- bcf.usc.edu/~gareth/ISL/<br><br>Slides and lecture notes provided<br><br>**M2:**<br><br>Randall Pruim, 2018, *Foundations and Applications of Statistics An Introduction Using R.* American Mathematical Society, Providence. ISBN 9781470428488. From this book we discuss topics from |

| | |
|---|---|
| | chapters 4 and 5.<br><br>Robert Shumway and David Stoffer, 2019. *Time Series: A Data Analysis Approach Using R*. CRC Press, Boca Raton. ISBN 9780367221096. From this book we discuss chapters 1 to 4 and some optional topics from chapters 5 and 8. |
| Supplementary Readings | **M1:**<br><br>Bishop, C. M. (2006). Pattern recognition and machine learning. New York: Springer.<br><br>Agresti, A., Finlay, B. Statistica per le scienze sociali, Pearson, 2009.<br><br>Hyndman, R.J. and Athanasopoulos, G. Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, 2018.<br><br>Cicchitelli, Giuseppe. Statistica. Principi e metodi. Pearson, 2008.<br><br>Azzalini, Adelchi, and Bruno Scarpa. Data analysis and data mining: An introduction. OUP USA, 2012.<br><br>Grigoletto, Matteo, Laura Ventura, and Francesco Pauli. Modello lineare: teoria e applicazioni con R. G Giappichelli Editore, 2017.<br><br>Johnson, Richard A., and Dean W. Wichern. "Applied multivariate statistical analysis." New Jersey 405 (1992).<br><br>**M2:**<br><br>Additional material and readings provided in class by the lecturer. |
| Further Information | |

| | |
|---|---|
| **Sustainable Development Goals (SDGs)** | Good health and well-being, Climate action, Reduced inequalities, Decent work and economic growth |

# *Course Module*

| | |
|---|---|
| **Course Constituent Title** | M1 - Statistical methods for business analysis |
| **Course Code** | 27502A |
| **Scientific-Disciplinary Sector** | STAT-01/A |
| **Language** | English |
| **Lecturers** | Prof. Alessandro Casa,<br>Alessandro.Casa@unibz.it<br>https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46549 |
| **Teaching Assistant** | |
| **Semester** | Second semester |
| **CP** | 6 |
| **Responsible Lecturer** | |
| **Teaching Hours** | 36 |
| **Lab Hours** | 18 |
| **Individual Study Hours** | - |
| **Planned Office Hours** | 18 |
| **Contents Summary** | This module begins with a review of key principles of statistical inference and then introduces core concepts in statistical learning. Topics include linear regression and its extensions, advanced regression techniques such as decision trees, logistic regression, classification methods, model selection strategies, and unsupervised learning approaches like principal component analysis and clustering. Throughout the course, students will work hands-on in R, applying techniques to real-world datasets drawn from business scenarios. By the end, students will be able to choose suitable statistical models, apply them to a range of business problems, and effectively communicate their analytical insights |
| **Course Topics** | - Review of statistical inference: random variables, confidence intervals, and hypothesis testing.<br>- Introduction to statistical learning concepts: basic vocabulary and notions, parametric and nonparametric approaches, predictive and inferential objectives, bias-variance trade off, supervised and |

| | |
|---|---|
| | unsupervised learning<br>- Linear regression and extensions: simple and multiple linear regression, model estimation and assessment, model assumptions, inferential tools, qualitative predictors, interaction effects, polynomial regression, basic notions on nonparametric regression<br>- Classification: introduction to classification, logistic regression, model estimation, evaluation of classifiers<br>- Other supervised learning techniques: trees, splines, additive models<br>- Model selection/assessment and evaluation of model complexity: resampling methods, cross-validation and information criteria<br>- Unsupervised learning: clustering tools such as k-means and hierarchical clustering, principal component analysis<br>- Applications with the R software |
| **Teaching Format** | In-person lectures and computer labs. Whenever possible, lectures will be structured to prioritize in-class time for discussions, and practical applications. |
| **Required Readings** | James, G., Witten, D., Hastie, T., Tibshirani, R. An Introduction to Statistical Learning with Applications in R. Springer, 2013. Freely available at http://www- bcf.usc.edu/~gareth/ISL/<br><br>Slides and lecture notes provided |
| **Supplementary Readings** | Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.<br><br>Agresti, A., Finlay, B. Statistica per le scienze sociali, Pearson, 2009.<br><br>Hyndman, R.J. and Athanasopoulos, G. Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, 2018.<br><br>Cicchitelli, Giuseppe. Statistica. Principi e metodi. Pearson, 2008.<br><br>Azzalini, Adelchi, and Bruno Scarpa. Data analysis and data |

| | |
|---|---|
| | mining: An introduction. OUP USA, 2012.<br><br>Grigoletto, Matteo, Laura Ventura, and Francesco Pauli. Modello lineare: teoria e applicazioni con R. G Giappichelli Editore, 2017.<br><br>Johnson, Richard A., and Dean W. Wichern. "Applied multivariate statistical analysis." New Jersey 405 (1992). |

# *Course Module*

| | |
|---|---|
| **Course Constituent Title** | M2 - Advanced statistics |
| **Course Code** | 27502B |
| **Scientific-Disciplinary Sector** | STAT-01/A |
| **Language** | English |
| **Lecturers** | |
| **Teaching Assistant** | |
| **Semester** | Second semester |
| **CP** | 6 |
| **Responsible Lecturer** | |
| **Teaching Hours** | 40 |
| **Lab Hours** | 20 |
| **Individual Study Hours** | - |
| **Planned Office Hours** | 18 |
| **Contents Summary** | • Parameter estimation: maximum likelihood methods<br>• Parameter estimation: Bayesian inference<br>• Time series: components and forecasting<br>• Time series: causal relationship tests<br>• Missing data<br>• Elements of statistics for Big Data |
| **Course Topics** | |
| **Teaching Format** | Frontal lectures, discussions and exercises on computer. |
| **Required Readings** | Randall Pruim, 2018, *Foundations and Applications of Statistics An Introduction Using R.* American Mathematical Society, Providence. |

| | |
|---|---|
| | ISBN 9781470428488. From this book we discuss topics from chapters 4 and 5.<br><br>Robert Shumway and David Stoffer, 2019. *Time Series: A Data Analysis Approach Using R*. CRC Press, Boca Raton. ISBN 9780367221096. From this book we discuss chapters 1 to 4 and some optional topics from chapters 5 and 8. |
| Supplementary Readings | Additional material and readings provided in class by the lecturer. |