Syllabus

Descrizione corso

Titolo insegnamento	Statistical Methods
Codice insegnamento	27502
Titolo aggiuntivo	
Settore Scientifico-Disciplinare	STAT-01/A
Lingua	Inglese
Corso di Studio	Corso di laurea magistrale in Data Analytics for Economics and Management
Altri Corsi di Studio (mutuati)	M1 Statistical methods for business analysis is loaned from course 25559 – Master in Entrepreneurship and Innovation (LM-77 EI) M2 Advanced statistics is loaned from course 73006 – Master in Computing for Data Science (LM-18)
Docenti	prof. Alessandro Casa, Alessandro.Casa@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46549
Assistente	
Semestre	Secondo semestre
Anno/i di corso	1
CFU	12
Ore didattica frontale	M1: 36 hours M2: 40 hours
Ore di laboratorio	M1: 18 hours M2: 20 hours
Ore di studio individuale	-
Ore di ricevimento previste	M1: 18 hours M2: 18 hours
Sintesi contenuti	M1: 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 M2: • Parameter estimation: maximum likelihood methods • Parameter estimation: Bayesian inference • Time series: components and forecasting • Time series: causal relationship tests • Missing data • Elements of statistics for Big Data
Argomenti dell'insegnamento	M1: - Review of statistical inference: random variables, confidence intervals, and hypothesis testing. - 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 - 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 - Classification: introduction to classification, logistic regression, model estimation, evaluation of classifiers - Other supervised learning techniques: trees, splines, additive models - Model selection/assessment and evaluation of model complexity: resampling methods, cross-validation and information criteria - Unsupervised learning: clustering tools such as k-means and hierarchical clustering, principal component analysis - Applications with the R software M2: - Parameter estimation: maximum likelihood methods - Parameter estimation: Bayesian inference

	- Time series: components and forecasting - Time series: causal relationship tests - Missing data - Elements of statistics for Big Data
Parole chiave	Statistical learning, regression, classification, clustering, dimensionality reduction, model selection , statistical inference, Bayesian statistics, time series modelling, big data
Prerequisiti	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. M2: the knowledge provided by a course in calculus and one in probability and statistics
Insegnamenti propedeutici	
Modalità di insegnamento	M1: In-person lectures and computer labs. Whenever possible, lectures will be structured to prioritize in-class time for discussions, and practical applications. M2: Frontal lectures, discussions and exercises on computer.
Obbligo di frequenza	Recommended, but not required.
Obiettivi formativi specifici e risultati di apprendimento attesi	Intended Learning Outcomes (ILO) M1: ILO 1 Knowledge and understanding: 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. ILO 1.2 The student consolidates knowledge of statistical inference, linear models and their generalisations, linear algebra, and optimisation techniques. 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.

	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. 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. ILO 3 Making judgements: 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. 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. 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. ILO 5 Learning skills: 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 - to identify thematic connections and to establish relationships between methods of analysis and application contexts;

	- 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. M2: ILO 1 Knowledge and understanding: 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. ILO 1.2 The student consolidates knowledge of statistical inference, linear models and their generalisations, linear algebra, and optimisation techniques. 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. 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. 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. ILO 3 Making judgements: 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. 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.

	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. ILO 5 Learning skills: 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 - to identify thematic connections and to establish relationships between methods of analysis and application contexts; - 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.
Obiettivi formativi specifici e risultati di apprendimento attesi (ulteriori info.)	
Modalità di esame	The overall exam mark will be determined by the assessment of the two modules (M1+M2). M1: Assessment (for both attending and non-attending students): - Written Exam: Exercises and review questions (65% of the final grade)(ILOs 1.1, 1.2, 1.3, 3.1, 3.2, 5.1) - 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). Notes:

	- 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. M2: 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.
Criteri di valutazione	M1: - Written exam: understanding of statistical concepts, correct interpretation of results of statistical analyses, clarity and precision of explanations. - Data Analysis Project: Quality and clarity of the presentation, adequacy and appropriateness of analyses with respect to dataset characteristics M2: 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%). The homeworks and the final written exam are separately evaluated with a score expressed in 30/30. 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.
Bibliografia obbligatoria	**M1:** 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/ Slides and lecture notes provided **M2:**

	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. 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.
Bibliografia facoltativa	**M1:** Bishop, C. M. (2006). Pattern recognition and machine learning. New York: Springer. Agresti, A., Finlay, B. Statistica per le scienze sociali, Pearson, 2009. Hyndman, R.J. and Athanasopoulos, G. Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, 2018. Cicchitelli, Giuseppe. Statistica. Principi e metodi. Pearson, 2008. Azzalini, Adelchi, and Bruno Scarpa. Data analysis and data mining: An introduction. OUP USA, 2012. Grigoletto, Matteo, Laura Ventura, and Francesco Pauli. Modello lineare: teoria e applicazioni con R. G Giappichelli Editore, 2017. Johnson, Richard A., and Dean W. Wichern. "Applied multivariate statistical analysis." New Jersey 405 (1992). **M2:**

	Additional material and readings provided in class by the lecturer.
Altre informazioni	
Obiettivi di Sviluppo Sostenibile (SDGs)	Buona salute, Lotta contro il cambiamento climatico, Ridurre le disuguaglianze, Buona occupazione e crescita economica

Modulo del corso

Titolo della parte costituente del corso	M1 - Statistical methods for business analysis
Codice insegnamento	27502A
Settore Scientifico-Disciplinare	STAT-01/A
Lingua	Inglese
Docenti	prof. Alessandro Casa, Alessandro.Casa@unibz.it https://www.unibz.it/en/faculties/economics-management/academic-staff/person/46549
Assistente	
Semestre	Secondo semestre
CFU	6
Docente responsabile	
Ore didattica frontale	36
Ore di laboratorio	18
Ore di studio individuale	-
Ore di ricevimento previste	18
Sintesi contenuti	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

Argomenti dell'insegnamento	- Review of statistical inference: random variables, confidence intervals, and hypothesis testing. - 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 - 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 - Classification: introduction to classification, logistic regression, model estimation, evaluation of classifiers - Other supervised learning techniques: trees, splines, additive models - Model selection/assessment and evaluation of model complexity: resampling methods, cross-validation and information criteria - Unsupervised learning: clustering tools such as k-means and hierarchical clustering, principal component analysis - Applications with the R software
Modalità di insegnamento	In-person lectures and computer labs. Whenever possible, lectures will be structured to prioritize in-class time for discussions, and practical applications.
Bibliografia obbligatoria	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/ Slides and lecture notes provided
Bibliografia facoltativa	Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer. Agresti, A., Finlay, B. Statistica per le scienze sociali, Pearson, 2009. Hyndman, R.J. and Athanasopoulos, G. Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, 2018.

	Cicchitelli, Giuseppe. Statistica. Principi e metodi. Pearson, 2008. Azzalini, Adelchi, and Bruno Scarpa. Data analysis and data mining: An introduction. OUP USA, 2012. Grigoletto, Matteo, Laura Ventura, and Francesco Pauli. Modello lineare: teoria e applicazioni con R. G Giappichelli Editore, 2017. Johnson, Richard A., and Dean W. Wichern. "Applied multivariate statistical analysis." New Jersey 405 (1992).

Modulo del corso

Titolo della parte costituente del corso	M2 - Advanced statistics
Codice insegnamento	27502B
Settore Scientifico-Disciplinare	STAT-01/A
Lingua	Inglese
Docenti	
Assistente	
Semestre	Secondo semestre
CFU	6
Docente responsabile	
Ore didattica frontale	40
Ore di laboratorio	20
Ore di studio individuale	-
Ore di ricevimento previste	18
Sintesi contenuti	• Parameter estimation: maximum likelihood methods • Parameter estimation: Bayesian inference • Time series: components and forecasting • Time series: causal relationship tests • Missing data

	• Elements of statistics for Big Data
Argomenti dell'insegnamento	
Modalità di insegnamento	Frontal lectures, discussions and exercises on computer.
Bibliografia obbligatoria	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. 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.
Bibliografia facoltativa	Additional material and readings provided in class by the lecturer.