COURSE DESCRIPTION – ACADEMIC YEAR 2023/2024

Course title	**Advanced Statistics**
Course code	73066
Scientific sector	MAT/06
Degree	Master in Computational Data Science (LM-18)
Semester	2
Year	1
Credits	6
Modular	No

Total lecturing hours	40
Total lab hours	20
Attendance	Attendance of classes and labs is not compulsory but highly recommended. Weekly home-works on the topics discussed in class will be assigned. Non-attending students have to contact the lecturer at the start of the course to agree on the modalities of the independent study.
Prerequisites	The knowledge provided by a course in calculus and one in probability and statistics.
Course page	https://ole.unibz.it/

Specific educational objectives	The course belongs to the type "affini o integrative – formazione affine" in the curriculum "Data Analytics". A second course in statistics on selected topics of statistical inference, time series and computational statistics. Topics discussed include Montecarlo methods and the bootstrap; maximum likelihood and Bayesian estimation, likelihood ratio testing. ARMA and regression modelling for time series data and forecasting. Techniques for dealing with missing data. Nonparametric density estimation and goodness of fit testing (optional). The course alternates front classes and lab activity where the methodology discussed is applied on real and simulated data. The software R will be used. This course, by combining theory and computer simulations and applications, aims at providing deep understanding and operational knowledge of some core techniques of statistical analysis which can be exploited either for applied data analysis or theoretical research.

Lecturer	Emanuele Taufer
Contact	Piazza Domenicani 3, Emanuele.Taufer@unibz.it
Scientific sector of lecturer	SECS-S/01
Teaching language	English
Office hours	Arrange beforehand by email.
Lecturing Assistant (if any)	Laura Di Lucchio
Contact LA	Piazza Domenicani 3, laura.dilucchio@unibz.it
Office hours LA	Arrange beforehand by email.

List of topics	• 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
Teaching format	Frontal lectures, discussions and exercises on computer.

Learning outcomes	Knowledge and understanding: • D1.1 - Knowledge of the key concepts and technologies of data science disciplines • D1.8 - Knowledge of the mathematical-statistical principles required for data analysis Applying knowledge and understanding: • D2.1 - Practical application and evaluation of tools and techniques in the field of data science • D2.2 - Ability to address and solve a problem using scientific methods • D2.7 - Practical application of mathematical-statistical tools and methods from the field of data science Making judgments • D3.2 - Ability to autonomously select the documentation (in the form of books, web, magazines, etc.) needed to keep up to date in a given sector Communication skills • D4.1 - Ability to use English at an advanced level with particular reference to disciplinary terminology Learning skills • D5.3 - Ability to deal with problems in a systematic and creative way and to appropriate problem solving techniques.

Assessment	The assessment is based on class and lab participation, home-work exercises and a final written exam. 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.
Assessment language	English
Assessment Typology	Monocratic
Evaluation criteria and criteria for awarding marks	For attending students the final grade will be determined by the evaluation of home-works, class and lab participation (20%) and the evaluation of a final written exam (80%). The home-works and the final written exam are separately evaluated with a score expressed in 30/30. Both parts need to reach the minimum threshold of 18/30 in order to pass the exam. 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	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. Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it
Supplementary readings	Additional material and readings provided in class by the lecturer.
Software used	The software R (https://cran.r-project.org/) and RStudio (https://posit.co/) , freely available, will be used during the course, for the home-works and for the final written exam.