COURSE DESCRIPTION – ACADEMIC YEAR 2024/2025

Course title	**Mathematics and Statistics for Data Science**
Course code	73065
Scientific sector	MAT/06
Degree	Master in Computing for Data Science (LM-18)
Semester	1
Year	1
Credits	6
Modular	No

Total lecturing hours	40
Total lab hours	20
Attendance	Generally, attendance is not compulsory, but recommended. Non-attending students must contact the lecturer at the start of the course to agree on the modalities of the independent study.
Prerequisites	Although the basic concepts and calculation techniques of integral and differential calculus and linear algebra are reviewed at the beginning of the course, it is necessary for the student to be familiar with the fundamental methodologies, methods and definitions of calculus and linear algebra.
Course page	https://ole.unibz.it/

Specific educational objectives	The course belongs to the type "affini o integrative – formazione affine". The main objective of the course is to provide students with a solid theoretical foundation in probability and statistics and the ability to solve problems in these two disciplines. The course consists of two parts. 1) In the **first part** of the course will cover the basic topics of probability, and will be preceded by a review of the concepts and mathematical tools of linear algebra and mathematical analysis necessary to understand the concepts and solve the problems of probability and statistics. 2) In the second **part** the course will deal with the basic topics of statistics. The course aims to give a solid theoretical foundation of the concepts of probability and statistics and to provide the tools and methods for solving practical problems that probability theory and statistics can solve in their applications to data science. At the end of the course, the student will have - **revised** the foundations of mathematical calculus necessary to approach probability and statistics problems

	- **acquired** the foundations of mathematical calculus, probability and statistics that will allow him/her to solve the most common problems of statistical **data processing and interpretation** that are common to many scientific fields such as computer science and software engineering, artificial intelligence, and data processing in numerous applications of these fields (e.g., biology, medicine and social sciences, etc.) - **acquired** the ability to solve problems that frequently arise in data science, such as, but not limited to, calculating the probability of events, modelling events data using probability distributions, testing hypotheses and predictive modelling of data.

Lecturer	Paola Lecca
Contact	Paola.Lecca@unibz.it, +39 0471 016162
Scientific sector of lecturer	INF/01
Teaching language	English
Office hours	Monday 10:00-11:00, appointment is requested by e-mail at least one day before.
Lecturing Assistant (if any)	Same as lecturer.
Contact LA	--
Office hours LA	--
List of topics	• Fundamentals of differential and integral calculus • Fundamentals of linear algebra • Probability theory • Data distribution models and analysis • Hypothesis tests • Regression analysis
Teaching format	Frontal lectures and labs with theoretical exercices.

Learning outcomes	**Knowledge and understanding** • Have a solid knowledge of the mathematical foundations of probability and statistics that are in support of the applications in computational data science. **Applying knowledge and understanding** • Be able to use the tools of mathematics to solve problems in data analysis. **Making judgments** • Be able to solve problems from a theoretical point of view, which is a indispensable prerequisite for later acquiring autonomy, discernment and judgement, as well ability to innovate in the use and implementation of problem-specific software solutions. **Communication skills** • Acquire the mathematical language to formalise the problem to be solved. **Ability to learn**

	• Ability to study and understand theoretical notions in order to recognize their applications.

Assessment	**Written exam** Assessment is based on 1) the evaluation of a review report on a theoretical topic assigned during the course 2) the evaluation of the final written exam consisting of exercises. Students not attending the course are asked to contact the lecturer during the last week of the course for the assignment of the theoretical topic review project.
Assessment language	English
Assessment Typology	Monocratic
Evaluation criteria and criteria for awarding marks	The examination is deemed passed if the student scores at least 18/30 on both the review report and the written examination. If a mark of 18/30 or higher is obtained in both parts (review reports and written exam) of the exams, the final mark will be calculated as the average of the two marks obtained.

Required readings	The course includes topics from different disciplinary areas of mathematics that are unlikely to be contained in a single textbook. It is therefore advisable that the student follows the notes and the didactical material that the lecturer will make available at each lecture and laboratory. The notes provided during the course can be deepened by referring to textbooks, as, for example: • Howie, John M., Real Analysis, Springer, 2001 • Maurits Kaptein , Edwin van den Heuvel, Statistics for Data Scientists. An Introduction to Probability, Statistics, and Data Analysis, Springer 2022 • Frederik Michel Dekking, Cornelis Kraaikamp, Hendrik Paul Lopuhaä, Ludolf Erwin Meester, A Modern Introduction to Probability and Statistics, Understanding Why and How, Springer 2005 • James, E. Gentle, Matrix Algebra: Theory, Computations and Applications in Statistics (Springer Texts in Statistics) 2nd ed. 2017. Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it

Supplementary readings	Suggested by the lecturer during the course if needed.
Software used	The course does not include programming labs. However, example scripts in R (www.r-project.org) may be shown. Communicate needed software and technical requirements in advance to cs-tech@inf.unibz.it