COURSE OF STUDY IN STRATEGIC MANAGEMENT AND DIGITAL MARKETING
ACADEMIC YEAR 2023 - 2024
DATA AND DECISIONS

General information	
Year of the course	I
Academic calendar (starting and ending date)	First semester (11/09/2023 – 15/12/2023)
Credits (CFU/ETCS):	6 CFU
SSD	Statistics, SECS-S/01
Language	Italian
Mode of attendance	Strongly recommended

Professor/ Lecturer	
Name and Surname	Crescenza Calculli
E-mail	crescenza.calculli@uniba.it
Telephone	(+39) 0805049266
Department and address	Room n. 18, 5th floor – Department of Economics and Finance @ ex-Facoltà di Economia, Largo Abbazia S. Scolastica - 70124 Bari, Italy
Virtual room	MS Teams "Ricevimento C. Calculli", MS Teams code: i45yfsh
Office Hours	Tuesday 11:30 - 13:30 and Friday 9.30 – 11.30 or by appointment

Work schedule			
Hours			
Total	Lectures	Hands-on (laboratory, workshops, working groups, seminars, field trips)	Out-of-class study hours/ Self-study hours
150	24	18	108
CFU/ETCS			
6			

Learning Objectives	The learning objectives of the course are synthesized into: • to understand the fundamental concepts of statistics and their role in data-driven decision-making processes. • to apply appropriate statistical techniques to analyze complex data and extract meaningful information from it. • to gain familiarity with statistical tools and methods used for data collection, organization, visualization, and cleaning. • to develop the ability to interpret and communicate statistical results clearly and convincingly. • to use statistical software and data analysis tools to conduct advanced quantitative analyses. • to acquire knowledge of best practices in data analysis. • to develop a solid theoretical foundation in statistics to critically address new challenges and complex problems in data-driven processes. • to integrate statistics with other disciplines such as artificial intelligence, machine learning, and data science to achieve an interdisciplinary and comprehensive perspective in data analysis and utilization.
Course prerequisites	Knowledge of the theoretical and methodological foundations of descriptive and inferential statistics acquired during the basic courses of undergraduate degrees in business, economics, statistics, and management engineering.

Teaching strategies	The Data and Decision course aims to provide students with the skills and knowledge necessary to tackle data-driven processes, using statistics effectively and rigorously as a key tool for understanding, analyzing, and transforming data into meaningful information to guide the business decision-making process. The lectures are accompanied by exercises in which theoretical topics are applied to real-world examples and case studies using the statistical software R. • The course consists of 24 hours of theoretical lectures and 18 hours of practical exercises using the R software. • Educational materials, exercises, and homework will be provided on the e-learning platform. Students are required to enroll in the course on the platform and will be asked to use a computer during the exercises. • Homework assignments (exercises to be completed after each lecture using the software) are provided to allow students to become familiar with the exam procedures. The assignments must be individually completed within a predetermined timeframe (generally before the next lecture). The results of the assignments contribute to assessing the overall commitment of the students enrolled in the course.
Expected learning outcomes in terms of	
Knowledge and understanding on:	The student will develop essential methodological knowledge to understand and critically evaluate different contexts within the field of data analysis to identify strategies for planning, organizing, and managing business resources. From a disciplinary perspective, the expected learning outcomes include acquiring competencies in various topics covered in the course curriculum. The objective is to provide students with the necessary tools to accurately interpret data and make informed decisions that promote business success.
Applying knowledge and understanding on:	The student is required to develop a strong mastery and autonomy in applying advanced methods and tools for data analysis, with particular attention to contexts related to strategic business management. In this regard, the use of the statistical software R represents an essential component of the educational journey. Students will be guided in learning the functionalities of R and will be able to use the software to perform statistical analyses, manipulate data, visualize results, and make data-driven decisions. The objective is to provide students with the necessary skills to address the complex challenges of data analysis in the business context and utilize R as a fundamental tool for analysis and information management.
• Soft skills	• *Making informed judgments and choices* The student is called upon to independently develop their own judgment skills in defining and evaluating issues related to data analysis in the context of business management. The development of critical thinking within the discipline will encourage an autonomous and informed approach to assessing and resolving issues related to data-driven decision-making processes. • *Communicating knowledge and understanding* The student will need to effectively communicate their opinions regarding the issues addressed in various analysis contexts and will be capable of presenting ideas and solutions for their treatment in a convincing manner. They will be encouraged to develop both written and verbal communication skills to persuasively present data analysis results and their interpretations to both specialists and non-specialists in the field. Through the practice of effective

	communication, the student will be able to adequately convey their ideas, promote discussion and collaboration, as well as facilitate understanding and implementation of proposed solutions within the realm of business management.
	• *Capacities to continue learning* The student will need to acquire the ability to thoroughly understand the multiple aspects of the study field to independently address different issues. This will be achieved through the utilization of tools provided during the course, deepening and updating their knowledge, as well as the ability to effectively locate and consult bibliographic material, utilizing both traditional methods and available online resources. These competencies will enable the student to navigate the field of study and approach challenges independently, making the best use of available resources, and maintaining a continuous approach to learning and staying updated.
Syllabus	

Content knowledge	**First part**
	• **Review of Descriptive Statistics and Probability**
	Types of data, frequency distributions, summary statistics. Basic concepts of probability, probability distributions (binomial, Poisson, normal, and t-Student distributions) and random variables.
	
	• **Inferential Problems: estimation, confidence intervals and hypothesis testing.**
	Goodness-of-fit tests, parametric tests (for one population, two populations, proportions, paired data tests). Group comparisons: one-way and two-way analysis of variance.
	
	Second part
	• **Statistical Learning Inference problems and prediction problems**
	Parametric and non-parametric methods. Classification and regression methods.
	
	• **Simple and multiple linear regression models**
	Motivation and definition of the linear regression model. Estimation and hypothesis testing of regression parameters. Model goodness-of-fit assessment. Residual analysis.
	
	• **Generalized linear models (logistic regression)**
	Motivation and definition of the logistic regression model. Estimation and hypothesis testing of model parameters. Application of logistic regression in classification problems.
	
	• **Other classification techniques (overview) and clustering methods**
	LDA and non-parametric methods, KNN; The ROC curve; Hierarchical clustering and k-means.
	
	Each topic of the program will be covered through the study of practical cases and with the aid of the statistical software R. The following are integral parts of the program:
	
	• **Introduction to the R statistical environment and the integrated development environment RStudio**
	    • Interaction with the software (loading data, creating working variables, objects, functions, packages)
	    • Use of real datasets, construction of graphical and tabular representations (basic and advanced graphical environments).

Texts and readings	For the first part CICCHITELLI, G., D'URSO, P., MINOZZO, M. (4° ed.). Statistica: principi e metodi. ISBN 8891915386 For the second part JAMES, G., WITTEN, D., HASTIE, T., TIBSHIRANI, R. *Introduzione all'apprendimento statistico con applicazioni in R*. Piccin Nuova Libraria S.p.A., Padova. ISBN 978-88-299-3094-4. Capitoli [1-5, 12] Edizione in inglese disponibile @ https://www.statlearning.com/ For the practical part WICKHAM, H., GROLEMUND, G. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data* (1st. ed.). O'Reilly Media, Inc. In-depth study PICCOLO, D. *Statistica per le decisioni (3° ed.).* Il Mulino, Bologna. ISBN 978-88-15-27220-1 SGARRO, A., Franzoi, L., Vicig, P. STATISTICA DI BASE. Idee e tecniche. Zanichelli, Bologna. ISBN 9788808620187
Notes, additional materials	The reference texts will be complemented with the lecturer's notes and teaching materials (slides and exercises) provided during the lectures and made available on the e-learning platform.
Repository	In e-learning mode, students will have access to slides and additional teaching materials, exercise guidelines, and data necessary for completing exercises and homework using the R software. The course address and password for the "Data and Decisions" course will be provided at the beginning of the course.

Assessment	
Assessment methods	For <u>attending students</u>, the evaluation of the educational activities is based on a final exam that includes both solving exercises using the statistical software R and an oral interview covering the entire course program. The practical exam using the software requires each student to submit a detailed report including commands and comments on the results. <u>Non-attending students</u> must take an exam covering the entire course program, which consists of: **1.** a multiple-choice test with 20 questions on all topics of the program, **2.** a practical exam conducted using the R software, and **3.** an oral interview. The multiple-choice test and the practical exam will be held a few days before the oral interview, following the instructions provided by the teacher.
Assessment criteria	• *Knowledge and understanding* The evaluation of the in-depth knowledge of the methodologies and tools presented during lectures and labs will be assessed based on the acquisition of all program contents during the examination. For <u>attending students</u>, the practical exam (2 hours) and the oral exam contribute equally to determining the final grade. Additionally, the final grade accounts the progress, continuity, and seriousness of the commitment demonstrated throughout the course, including the submission of homeworks assigned after each lesson. For <u>non-attending students</u>, the three specified tasks will contribute to determining the final grade, with weights of 20% (1.), 40% (2.), and 40% (3.). • *Applying knowledge and understanding* The ability of attending students to critically analyze the different proposed

case studies through the assigned exercises during the exam will be assessed in a 2-hour task. The examination method allows for verifying the skill of applying acquired knowledge. Similarly, non-attending students are required to perform data analysis on a set of data using the software R within a 2-hour timeframe, referring to the methodologies covered in the entire course program.

- *Autonomy of judgment*

The evaluation will focus on structured reasoning aimed at solving complex case studies through the independent selection of learned methods and tools. Comments, justifications made for solving the exercises, and the interpretation of results will be assessed.

- *Communication skills*

After completing and submitting their work using the software, each candidate will participate in a public oral discussion. For attending students, this discussion may involve questions related to the practical exam and one or two questions about the exam program. For non-attending students, the discussion will cover questions based on both the multiple-choice test answers and the practical exam. To evaluate the candidates' communication skills, the comments provided during the practical exam will also be taken into consideration.

- *Capacities to continue learning*

The learning outcomes will be assessed during the exam through both the practical test/multiple-choice test (only for non-attending students) and the oral discussion, aiming to evaluate the abilities developed by each candidate in terms of independent learning of methodologies and their application in different data analysis contexts. The monitoring of individual work required for the preparation of the final exam contributes to this assessment.

Final exam and grading criteria	For <u>attending students</u>: the correction and evaluation of assignments will take place in the days following the practical exam. The content will also be evaluated in terms of authenticity, and any instances of identical phrases or expressions in the interpretation of results or the same errors in commands will be automatically invalidated. Invalidating the practical exam requires the obligation to retake the exam. Each exam (practical and oral discussion) will be evaluated on a scale of 30. The proposed grade will be obtained by averaging the grades obtained in the two exams. Similarly, for <u>non-attending students</u>, the correction and evaluation of assignments will occur in the days following the practical exam with the software and the multiple-choice test. The content will also be evaluated for authenticity, and any instances of identical phrases or expressions in the interpretation of results or the same errors in R commands will be automatically invalidated. Invalidating the practical exam requires the obligation to retake the exam. Each exam will be evaluated on a scale of 30. The proposed grade will be obtained by averaging the grades obtained in the three tasks (test, practical, and oral), assigning weights of 20%, 40%, and 40%, respectively.
Further information	-----