General Information	
Academic subject	Multivariate Data Analysis
Degree course	Strategia d'Impresa e Management
Curriculum	
ECTS credits	8 CFU/ECTS
Compulsory attendance	No
Language	Italian (English on demand)

Subject teacher	Name Surname	Mail address	SSD
	Massimo Bilancia	massimo.bilancia@uniba.it	SECS-S/01

ECTS credits details	Area		CFU/ETCS
Basic teaching activities			

Class schedule	
Period	First Semester
Year	Second Year
Type of class	Classroom lectures and lab exercises

Time management	
Hours	200
In-class study hours	72 (1 ECTS = 9h)
Out-of-class study hours	128

Academic calendar	
Class begins	September 13, 2021
Class ends	December 23, 2021

Syllabus	
Prerequisites/requirements	Basic knowledge of statistics and mathematics at undergraduate level
Expected learning outcomes	*Knowledge and understanding on:* Students learn the basics of data mining techniques *Applying knowledge and understanding on:* Students will be able to apply data science techniques using R software. *Making informed judgements and choices:* Students will be able to apply the correct data analysis technique depending on the context of data analysis (e.g.: business intelligence and marketing) *Communicating knowledge and understanding*: Students will be able to create appropriate reporting of data analysis results *Capacities to continue learning*: This course can become an essential prerequisite for advanced data science courses
Contents	**Part I – Basic probability and statistical inference** 1. Introduction 2. Elementary probability 3. Discrete and continuous random variables 4. Double and multivariate random variables 5. The elements of statistical inference

	6. Data matrices **Part II – Data Mining and knowledge discovery** 1. Supervised and unsupervised learning 2. Association rules and Market Basket Analysis 3. Naïve Bayes classification 4. Decision trees 5. Clustering I: hierarchical clustering 6. Clustering II: k-means clustering **Part III – Lab: Introduction to R and Data Mining libraries**
Course program	
Bibliography	M. Bilancia (2020) Dispense per il Corso di Metodi Statistici Multivariati – Versione 1.2 Febbraio 2020. Freely available under Creative Commons 4.0 CC BY-NC-ND Licence.
Notes	
Teaching methods	Classroom lectures and lab exercises
Assessment methods	PC skills test on a real data set, plus a written exam on the theoretical aspects of the course
Evaluation criteria	Eligibility for PC skill test, grade in 30/30 for the written exam
Further information	