

Principali informazioni sull'insegnamento	
Titolo insegnamento	Metodi Statistici Multivariati
Corso di studio	Strategie di Impresa e Management
Crediti formativi	8 CFU/ECTS
Denominazione inglese	Multivariate Data Analysis
Obbligo di frequenza	NO
Lingua di erogazione	Italiano (English on demand)

Docente responsabile	Nome Cognome	Indirizzo Mail
	Massimo Bilancia	Massimo.bilancia@uniba.it

Dettaglio credi formativi	Area	SSD	CFU/ETCS
		SECS-S/01 (Statistica)	8

Modalità di erogazione	
Periodo di erogazione	I° Semestre
Anno di corso	Secondo anno
Modalità di erogazione	Lezioni frontali + esercitazioni in laboratorio

Organizzazione della didattica	
Ore totali	200
Ore di corso	72 (1 CFU = 9h)
Ore di studio individuale	128

Calendario	
Inizio attività didattiche	13 settembre 2021
Fine attività didattiche	23 dicembre 2021

Syllabus	
Prerequisiti	Conoscenze di base di statistica e di matematica
Risultati di apprendimento previsti	<ul style="list-style-type: none"> • <i>Conoscenza e capacità di comprensione</i>: il corso si prefigge di fornire gli elementi di base delle tecniche di “data science” • <i>Conoscenza e capacità di comprensione applicate</i>: lo studente apprenderà a stimare ed utilizzare in pratica i modelli appresi durante la parte teorica del corso, mediante l'utilizzo del più diffuso software di analisi dei dati • <i>Autonomia di giudizio</i>: lo studente sarà in grado di decidere il modello più appropriato da utilizzare per estrarre nuova conoscenza negli svariati ambiti di applicazione previsti (applicazioni campo aziendale, business intelligence e marketing) • <i>Abilità comunicative</i>: lo studente, al termine del corso, avrà

	<p>acquisito la preparazione necessaria a generare reportistica, negli ambiti di applicazione previsti, sulla nuova conoscenza estratta</p> <ul style="list-style-type: none"> • <i>Capacità di apprendere</i>: il corso si prefigge di fornire gli elementi di base delle tecniche di “data science”, conoscenze sulle quali fondare la possibilità di seguire corsi di natura più avanzata mirati alla preparazione della figura di “data scientist”
Contenuti di insegnamento	<p>Modulo I. Elementi di probabilità ed inferenza statistica</p> <ol style="list-style-type: none"> 1. Introduzione di base 2. Probabilità elementare 3. Variabili aleatorie discrete e continue 4. Variabili aleatorie doppie e multiple 5. Elementi di inferenza statistica 6. Matrici dati <p>Modulo II. Modelli di Data Mining e Knowledge Discovery per l'azienda</p> <ol style="list-style-type: none"> 7. Apprendimento supervisionato e non supervisionato 8. Regole associative e Market Basket Analysis 9. Classificazione Naïve Bayes 10. Alberi di decisione 11. Clustering I: algoritmi gerarchici 12. Clustering II: k-means <p>Modulo III. Laboratorio</p> <p>Il laboratorio è parte integrante del corso. Sarà utilizzato il software di analisi dei dati R, liberamente disponibile sul network del CRAN all'indirizzo http://cran.r-project.org. Durante il corso di laboratorio sarà erogata un'introduzione approfondita ad R, e verranno mostrati casi pratici di studio relativi a tutti gli argomenti teorici riguardanti la seconda parte.</p>

Programma	
Testi di riferimento	M. Bilancia (2020) Dispense per il Corso di Metodi Statistici Multivariati – Versione 1.2 Settembre 2020. Dipartimento Jonico, liberamente distribuite sotto Licenza Creative Commons 4.0 CC BY-NC-ND
Note ai testi di riferimento	
Metodi didattici	Lezioni frontali in aula e lezioni pratiche al computer
Metodi di valutazione	Prova di laboratorio in R, più prova scritta finale sugli argomenti di carattere teorico del corso. Il docente si riserva la possibilità di cambiare le modalità di svolgimento degli esami in corso d'opera, in relazione all'andamento dell'emergenza sanitaria in atto
Criteri di valutazione	Idoneità per la prova di laboratorio, voto finale in 30/30 basato sulla prova scritta
Altro	