

Insegnamento di: Statistica e Data Science per il Marketing					
Classe di laurea: LM-40 - Matematica	Corso di Laurea in: Matematica	Anno accademico: 2019/2020			
Denominazione inglese insegnamento: Statistics and Data Science for Marketing	Tipo di insegnamento: d) attività a scelta	Anno: 2	Semestre: 2		
Tipo attività formativa: A scelta	Ambito disciplinare:	Settore scientifico-disciplinare: MAT/06	CFU totali: 7 di cui CFU lezioni: 6.5 CFU ese/lab/tutor: 0.5		
Modalità di erogazione, ore di didattica assistita ed ore dedicate allo studio individuale					
ore di lezione: 52	ore di esercitazione/laboratorio/tutorato: 8				
totale ore didattica assistita: 60					
totale ore di studio individuale: 90					
Lingua di erogazione: Italiano	Obbligo di frequenza: no				
Docente: Marcello De Giosa	Tel: +39 080 5442707 e-mail: marcello.degiosa@uniba.it	Ricevimento studenti: Dip. Matematica piano 4 , stanza 12	Giorni e ore ricevimento: Lunedì 11-13, o previo appuntamento		
Conoscenze preliminari: Calcolo delle Probabilità e Statistica, come generalmente offerto in un corso della classe L-35					
Obiettivi formativi: Acquisire padronanza dei concetti e delle tecniche e del software R per le applicazioni del Data Science e della Statistica ai problemi moderni del Marketing.					
Risultati di apprendimento previsti	<p>Conoscenza e capacità di comprensione: Gli studenti dovrebbero acquisire i concetti e le tecniche moderne fondamentali del Data Science e della Statistica per il Marketing. Gli studenti dovrebbero anche acquisire abilità avanzate per l'uso del software R per la soluzione di tipici problemi pratici di Marketing.</p> <p>Conoscenza e capacità di comprensione applicate: Le conoscenze teoriche e le abilità con il software R sono utili nell'affrontare i problemi pratici tipici del marketing.</p> <p>Autonomia di giudizio: Capacità di affrontare problematiche di Marketing con le appropriate tecniche statistiche. L'abilità nel risolvere problemi dovrebbe essere supportata dalla capacità di implementare tecniche di Data Science e Analisi di Dati e protocolli consistenti con l'approccio teorico.</p> <p>Abilità comunicative: Acquisizione del linguaggio e formalismo della statistica necessari per comprendere e tradurre i problemi pratici in problemi di analisi di dati, per risolverli con i protocolli e i software corretti, per spiegare e presentare la conoscenza ed i risultati acquisiti ad una ampia platea.</p> <p>Capacità di apprendere: Gli studenti dovrebbero acquisire i metodi di apprendimento e le capacità necessarie per interpretare e risolvere anche problemi pratici non direttamente affrontati nel corso.</p>				
Programma del corso					
1. Panoramica sul linguaggio R e sui pacchetti usati durante il corso.					
2. Alcuni Modelli Lineari Avanzati. Regressione Multipla. Regressione Logistica. Modelli Lineari Gerarchici.					
3. Ridurre la dimensione e la complessità dei dati. Analisi delle componenti principali (PCA). Mappe percettive. Analisi fattoriale esplorativa (EFA). Rotazioni. Approfondimenti matematici su PCA e EFA.					

4. Analisi Fattoriale Confermativa e Modelli ad Equazioni Strutturali..

5. Segmentazione. Clustering e Classificazioni.

Clustering Gerarchico. Clustering basato sulla media. Clustering basato sul modello. Classificazioni.

Approfondimenti matematici su Clustering e Classificazioni.

6. Regole associative ed analisi del cesto della spesa.

Metriche. Individuare e visualizzare regole. Esplorare segmenti.

Approfondimenti matematici su Associazioni ed Analisi del Cesto della Spesa.

7. Modelli di Scelta

Analisi congiunta basata sulla scelta. Simulare dati di scelta. Adattare un modello di scelta. Aggiungere eterogeneità dei consumatori ai modelli di scelta. Modelli di scelta gerarchici bayesiani. Pianificazione di una indagine su modelli di scelta.

8. Analisi dei Social Network.

Network Diretti contro Indiretti. Visualizzare ed Analizzare Networks. Metriche e Tassonomia dei Social Data. Usare Metriche di Network in Previsioni e Classificazioni. Collezionare Dati sui Social Network con R.

9. Text Mining.

La Rappresentazione Tabulare del Testo: Matrice Termine-Dокументo e “Borsa-di-parole”. “Borsa-di-parole” contro Estrazione di significato a Livello di Documento. Estrazione di Significati a Livello di Documento. Preprocessare il Testo. Implementare Metodi di Text Mining.

Metodi di insegnamento:

Lezioni frontali ed esercitazioni con il software R.

Supporti alla didattica:

Note e slides distribuiti dal docente.

Controllo dell'apprendimento e modalità d'esame:

Prova orale e di laboratorio con R.

Testi di riferimento principali:

1) G. Shmueli, P.C. Bruce, I. Yahav, N.R. Patel, K.C. Lichtendahl (2018) – Data Mining for Business Analytics.

Concepts, Techniques, and Applications in R. - Wiley.

2) C. Chapman, E.McDonnel Feit (2019) - R for Marketing and Analytics. - Springer.

Altre letture

3) W.W. Piegorsch (2015) - Statistical Data Analytics. Foundations for Data Mining, Informatics, and Knowledge Discovery. - Wiley

4) T. W. Miller (2015) - Marketing Data Science - Pearson.

5) M. Sarstedt, E. Mooi (2019) - A Concise Guide to Market Research - Springer.