

Analisi dei dati

Domenico De Stefano

a.a. 2018/2019

Indice

1 Dettagli sul corso

Orario

	Lun	Mar	Mer	Gio	Ven
11-13		Lezione Aula 425,Ed.A	Lezione Aula 425,Ed.A		
14-17				Pratica Aula C,Ed. F	

La parte pratica inizierà indicativamente la settimana del 15 ottobre, verrà data comunicazione via **Moodle** e messaggistica.

Programma del corso

Due parti: **indagini statistiche e dati in campo sociale (social surveys)** e **analisi dei dati (data analysis/data mining)** e **modelli statistici**

1 Indagini statistiche

- **Indagini statistiche e dati statistici in campo sociale** I dati. Dati e scienze sociali. Le indagine statistiche. Esempi di indagine statistiche a livello nazionale (in particolare indagini ISTAT). Popolazione (Universo) e campione. Il campionamento (concetti base). Studi osservazionali e sperimentali. Dati da social network. Tecniche di rilevazione dati. Nozioni di campionamento statistico e alcune tra le principali indagini campionarie in Italia ed Europa. Matrice dati e variabili. Indicatori statistici.

Programma del corso (2)

2 Analisi dei dati

- ▶ **Tecniche di analisi dei dati e alcuni modelli statistici.**

Introduzione all'analisi dei dati da indagine. Tecniche di analisi dei dati multidimensionali. La matrice di correlazione. Elementi di analisi fattoriale (analisi in componenti principali e analisi delle corrispondenze). Analisi dei gruppi (segmentazione). Introduzione ai modelli statistici con particolare attenzione al modello di regressione lineare multiplo.

- ▶ **Parte pratica** Il software *R* (<https://cran.r-project.org/>)

Libri di testo e riferimento

■ Indagini statistiche

- ▶ Parti scelte da ISTAT, **Manuale di tecniche di indagine**, liberamente scaricabile dal sito dell'ISTAT (url fornito dal docente via *moodle2* (moodle2.units.it))

■ Analisi dei dati e modelli statistici

- ▶ Capitoli da: Marcello Gallucci, Luigi Leone e Manuela Berlingeri. **Modelli statistici per le scienze sociali** 2/Ed. Pearson Italia

■ Per entrambe le parti: Slides, appunti dalle lezioni e materiale didattico integrativo fornito dal docente → *moodle2* (moodle2.units.it)

■ Parte pratica

- ▶ Stefano M. Iacus e Guido Masarotto. **Laboratorio di statistica con R**, 2/Ed. McGraw-Hill Education

Esame

La valutazione finale consisterà in 2 parti: una prova scritta ed un lavoro di gruppo (prova pratica)

- La **prova scritta** consiste nella risposta ad alcuni quesiti, nell'interpretazione di output di modelli e nella soluzione ad alcuni esercizi che dovrete risolvere, illustrando il procedimento adottato
 - ▶ 80% della valutazione complessiva (100% per coloro che non svolgeranno il lavoro di gruppo)
- Il **lavoro di gruppo** consiste nell'analizzare i dati mediante il software R su un particolare tema di analisi che verrà fornito dal docente.
 - ▶ restante 20% della valutazione complessiva
 - ▶ La valutazione di questa parte avverrà a seguito di una presentazione svolta dai vari gruppi in una giornata appositamente dedicata.