

# ANALISI DEI DATI PER L'IMPRESA

Introduzione al corso

M. Trevisani, R. Pappadà

Dipartimento di scienze economiche, aziendali, matematiche e statistiche B. de Finetti, UniTS

24 Febbraio 2025

## Docenti

- ▶ MATILDE TREVISANI (matildet@deams.units.it)
- ▶ ROBERTA PAPPADA' (rpappada@units.it)

## Orario delle lezioni

Durata del corso 45 ore (6 CFU)

- ▶ Calendario
  - ▶ lunedì 14:00-16:00
  - ▶ mercoledì 14:00-16:00
- ▶ Aula 4\_INF\_01

Possibili variazioni al calendario verranno comunicate attraverso Teams/Moodle

Periodo delle lezioni: 24 febbraio - 30 maggio 2025

Non si terrà lezione:

- ▶ il 21 e 23 aprile per la pausa pasquale

Si tratta di un calendario di massima, sono possibili variazioni che verranno comunicate in anticipo via Teams/Moodle

- ▶ Per l'ora e il giorno si veda la pagina della docente dal personale del dipartimento  
<https://deams.units.it/it/dipartimento/persone/personale-docente> e quindi dalla *scheda persona*
- ▶ più direttamente, inserendo il nome del docente in cerca docente di ateneo <https://units.coursecatalogue.cineca.it/cerca-docenti>
- ▶ Uffici
  - ▶ Roberta Pappadà Stanza 2.18, Il piano (Ed. D, P.le Europa)
  - ▶ Matilde Trevisani Stanza 2.16, Il piano (Ed. D, P.le Europa)

## Scopo del corso

- ▶ Si veda il [syllabus](#) del corso

Il corso mira a fornire agli studenti gli strumenti di base per la visualizzazione e l'analisi dei dati mediante l'uso del software statistico "R".

Al termine del corso, lo/a studente/ssa sarà in grado di:

- ▶ utilizzare R per il calcolo, l'organizzazione e la rappresentazione dei dati;
- ▶ utilizzare semplici funzioni per descrivere diversi tipi di variabili e sintetizzare i dati in maniera opportuna;
- ▶ utilizzare alcune tecniche statistiche e adattare modelli statistici utilizzando funzioni in R;
- ▶ ricorrere a strumenti di reporting per illustrare i risultati ottenuti.

Il corso riguarderà

- ▶ le principali strutture di dati in R, la lettura dei dati e la loro organizzazione in data frames e altre strutture
- ▶ la trasformazione dei dati e gli strumenti per l'analisi esplorativa e la visualizzazione dei dati
- ▶ le principali strutture di programmazione per lavorare con istruzioni, funzioni e strutture di controllo

Inoltre verranno trattati

- ▶ approfondimenti su alcuni aspetti di programmazione in R per l'analisi dei dati
- ▶ semplici modelli statistici (ad esempio, la regressione lineare multivariata) in R
- ▶ alcune tecniche statistiche per dati multivariati (ad esempio, tecniche di cluster analysis) in R

Verranno illustrati vari esempi di dati reali e proposti esercizi in aula

Materiale per apprendere R:

- ▶ Lucidi e appunti, file sorgenti di R e dati per le sessioni pratiche saranno forniti durante il corso.
- ▶ Ulteriore materiale presentato a lezione o utile per lo studio verrà reso disponibile via Moodle
- ▶ Ulteriori risorse sono disponibili nella pagina web del CRAN (The Comprehensive R Archive Network: <https://cran.r-project.org/>). Vedi ad es *Manuals* e *Contributed*

Testo di riferimento per le tecniche statistiche:

- ▶ James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to Statistical Learning, with application in R. New York: Springer. [8th printing 2017](#)

Dalla [pagina degli autori](#) è scaricabile una [seconda edizione \(corrected 2023\)](#)

### ▶ YARRR! THE PIRATE'S GUIDE TO R

Questa guida accessibile (e giocosa!) ti aiuterà ad analizzare immediatamente i dati. I capitoli 2-4, 9, 13, 15 ti prepareranno per la maggior parte di ciò che incontrerai nei corsi di statistica. Questo libro è ricco di informazioni utili facili e veloci da digerire.

### ▶ MODERN DIVE

Questo libro online fornisce un'introduzione equilibrata a R con una forte enfasi sulla discussione e sulla visualizzazione dei dati. Dopo aver seguito le prime due parti, sarai pronto per la maggior parte dei corsi R di alto livello.

### ▶ R FOR DATA SCIENCE

Questo è lo standard di riferimento per lo sviluppo di competenze di programmazione R, gestione dei dati e visualizzazione. Questo libro ha molti capitoli brevi. Anche solo leggere i capitoli 2-6 ti dà una familiarità di base con R.

per l'analisi dei dati:

- ▶ Elements of Data Analytic Style
- ▶ The Art of Data Science
- ▶ Exploratory Data Analysis
- ▶ Regression Modeling for Data Science in R

## Modalità di insegnamento

- ▶ Il corso prevede lezioni frontali che avranno tuttavia carattere principalmente pratico (si raccomanda di frequentare le lezioni)
- ▶ Ogni studente potrà utilizzare il proprio laptop (o una postazione dell'aula)
- ▶ E' importante ricordarsi di
  - ▶ installare R (<https://cran.mirror.garr.it/CRAN/>)
  - ▶ installare RStudio (<https://posit.co/download/rstudio-desktop/>)
- ▶ Durante le lezioni verranno talvolta proposti degli esercizi utili all'autovalutazione (attraverso, ad esempio, quiz di Wooclap o Mentimeter)

L'esame finale prevede una prova scritta e lo svolgimento di un lavoro di gruppo:

- ▶ il test scritto comprende domande a risposta chiusa (vero/falso e a scelta multipla) da svolgere con l'impiego di R.
- ▶ nel lavoro di **gruppo** (3/4 persone) verrà chiesto di sviluppare un progetto inerente alle tecniche di analisi dei dati discusse durante il corso e consegnare un report **entro la scadenza data che precede la sessione d'esame estiva**.
- ▶ Per valutare il lavoro di gruppo potrà essere approfondito il contributo individuale o richiesta l'esposizione di quanto presentato nel report

## Metodo di valutazione

- ▶ La valutazione si basa sul punteggio individuale ottenuto nella prova scritta (con un peso **0.7**) e sulla valutazione del lavoro di gruppo (con peso **0.3**).
- ▶ Lo studente che non partecipa al lavoro di gruppo può raggiungere un **punteggio massimo** di 24/30.
- ▶ Il voto finale è in trentesimi. Il voto minimo per superare l'esame è 18/30.
- ▶ Sono previsti tre appelli nella sessione di giugno-luglio, uno nella sessione di settembre.

Domande?