

LABORATORIO DI INTRODUZIONE ALLA DATA SCIENCE

Cosa serve:

R-project → per scaricare l'ultima versione potete usare questo [link](#). Cliccate su "download R" e scegliete il mirror più vicino a voi (per l'Italia c'è l'università di Padova e quella di Milano).

RStudio → andate su questo [link](#) e scaricate la versione free di RStudio Desktop.

NOTA: per gli utenti Windows è utile scaricare ed installare anche l'ultima versione di Rtools che trovate a questo [link](#).

Programma

MODULO 1 (Prof. Riccardo Giubilei e Prof. Stefano Moro)

Prima parte – Introduzione all'analisi dei dati in R

1. Introduzione a R e RStudio
2. Primi passi in R: funzioni, oggetti, operatori logici, tipi di variabili, strutture dati
3. Statistica descrittiva: richiami teorici e implementazione in R
4. Variabili aleatorie e modelli probabilistici: richiami teorici e implementazione in R
5. Stimatori, intervalli di confidenza, test di ipotesi: richiami teorici e implementazione in R

Seconda Parte – Manipolazione e visualizzazione dati con Tidyverse

1. Data Visualization con tidyverse: ggplot2 e ggpubr
2. Data wrangling con tidyverse: dplyr e tidyr
3. Programmazione in R: cicli, funzioni e accenni di programmazione funzionale

MODULO 2 (Prof. Pierfrancesco Alaimo di Loro e Prof. Marco Mingione)

Terza Parte – Gli strumenti dello "Statistical Learning"

1. Modelli statistici e complessità
2. Tecniche di regressione lineare e non lineare
3. Approcci penalizzati alla stima: Ridge e Lasso
4. Valutazione, validazione e scelta del modello
5. *Extra: modelli basati su alberi e reti neurali*

Quarta Parte – Dalla teoria alla pratica (Advanced R programming)

1. Implementazione efficiente degli strumenti dello "Statistical Learning"
2. Tecniche per la reportistica: Rmarkdown
3. Costruzione di dashboard interattive: il pacchetto Shiny

References

- *Elements of Statistical Learning*, <https://hastie.su.domains/ElemStatLearn/>
- *R for Data Science*, Wickham e Grolemund, <https://r4ds.had.co.nz/index.html>
- *An Introduction to Statistical Learning*, Hastie e Tibshirani, <https://www.statlearning.com/>

- *Mastering Shiny*, Hadley Wickham, <https://mastering-shiny.org/>