Livellamento al Software R - Lezione 1



Docente: Andrea Gilardi

Contatto: andrea.gilardi@unimib.it

Politecnico di Milano

Aspetti organizzativi

- Il syllabus del corso è disponibile al seguente link, mentre la pagina web del corso è disponibile sul portale e-learning di Ateneo.
- Su e-learning potete trovare avvisi, orari ed aule aggiornati, slides e script utilizzati durante le lezioni, e informazioni relative alla modalità d'esame.
- Nel caso non riusciate a registrarvi alla pagina e-learning, potete trovare tutte le informazioni anche al seguente link: https://github.com/agila5/livellamenti-CLAMSES.

Aspetti organizzativi (cont)

- Orario:
 - Martedì 17/09, dalle 8.30 alle 11.30, Laboratorio 908;
 - Giovedì 19/09, dalle 13:30 alle 16:30;
 - Venerdì 20/09, dalle 8.30 alle 10.30;
 - Mercoledì 25/09, dalle 8.30 alle 11.30;
 - Lunedì 30/09, dalle 8.30 alle 11.30.
- Le lezioni (tranne quella del giorno 17/9) si terranno nel Laboratorio 715 in U7.
- E' necessario che portiate il vostro PC durante le lezioni in Laboratorio.

Esame Finale

- Per il conseguimento dei 2 CFU relativi a questo corso di livellamento è necessario il superamento di una prova finale in cui vi verrà chiesto di risolvere alcuni esercizi relativi agli argomenti visti durante il corso.
- La prova finale si svolgerà in laboratorio, indicativamente 3 settimane dopo la fine del corso.
- Non sono previsti ulteriori appelli oltre alla prova di Ottobre.

Riferimenti Bibliografici

- Testi di riferimento:
 - Gli script forniti a lezione.
 - De Micheaux, P.L., Drouilhet, R. and Liquet, B., 2013. The R software. Springer. New York. Disponibile online tramite la VPN di Ateneo.
 - Venables, W. N., Smith D. M. & the R Core Team (2021). An Introduction to R. II manuale è disponibile online.
- Testi di utile consultazione ed approfondimenti:
 - Wickham, H. and Grolemund, G., 2016. R for data science. O'Reilly Media. Disponibile online.
 - Chambers, J.M., 2008. Software for data analysis: programming with R (Vol. 2). New York: Springer. II libro (formato PDF o ebook)
 è disponibile gratuitamente online tramite la biblioteca di Ateneo.
 - Burns, P., 2011. The R inferno. Disponibile online.

Il software R

- R è un linguaggio di programmazione ed un software dedicato all'analisi statistica, il calcolo scientifico, la rappresentazione grafica e la manipolazione di dati.
- Il software è open source e può essere scaricato gratuitamente al seguente link: https://www.r-project.org/.
- Esistono alcune differenze di installazione in base all'OS. Nel dubbio, consultate il manuale.

Il software R (cont)

L'ultima versione rilasciata è la v4.4.1 (Race For Your Life);

Remark

A partire dalla v2.14.0, ad ogni versione di R viene associato un nickname, tipicamente basato sulle vignette dei Peanuts: https://www.youtube.com/watch?v=z1vTSdRolgI&ab_channel=RConsortium

- NB: R un linguaggio di programmazione scritto da statistici per statistici!
- Il software è estremamente duttile e versatile per analizzare dati e sviluppare nuovi modelli statistici.
- Inoltre, i metodi più comuni per l'analisi dati (e.g. ANOVA, regressione lineare, PCA, ...) sono implementati di default nel software (nel cosiddetto R - base).

Il software R (cont)

- Ciò nonostante, soprattutto per i primi tempi, il suo utilizzo potrebbe risultarvi leggermente ostico¹.
- La sintassi ed il codice presentano diverse stranezze, inconsistenze, e peculiarità di cui bisogna tener conto quando lavoriamo.
- La soluzione a tutti questi problemi è di non scoraggiarsi e provare a svolgere più esercizi ed esperimenti possibile per prendere familiarità con il software e il linguaggio.

¹A tal proposito, dopo aver completato questo livellamento (e magari anche qualche corso più avanzato di programmazione con R), vi suggerirei di leggere il resoconto riportato in questo link.

Rstudio

- Il codice R può essere scritto utilizzando la sua interfaccia nativa (Rgui) o anche un editor di testo (i.e. Notepad) e la linea di comando.
- Il linguaggio è integrato anche con gli editor di testo più comuni (e.g. Emacs/ESS o Visual Studio Code).
- In questo corso utilizzeremo un software chiamato Rstudio (Desktop) per interfacciarci con R. Questo è un software open-source scaricabile gratuitamente da qui.
- RStudio è infatti un IDE (Integrated Development Environment) che grazie alla sua interfaccia ne rende più agevole l'utilizzo.
- NB: Il software ed il linguaggio si chiamano R. Rstudio è l'IDE che fornisce solo l'interfaccia grafica.

Cominciamo ad aprire il nostro primo script!