

Programma del modulo Statistica: analisi dei dati in R

Docente: Prof. Annamaria Porreca

Durata: 16 ore (8 lezioni da 2h)

Destinatari: Dottorandi in *Scienze e Tecnologie innovative per la salute e l'invecchiamento attivo* – XXXIX ciclo

Lezione 1 – Introduzione a R e concetti base (2h) - 11 ottobre 2025 ore 14:00-16:00

- R e RStudio: panoramica, differenze con altri software
- Script vs Console, gestione del workspace
- Oggetti in R: creazione, assegnazione e uso
- Pacchetti: installazione e caricamento
- Mini-esercizi: primi comandi e operazioni

Lezione 2 – Strutture di dati in R (2h) - 13 ottobre 2025 ore 14:00-16:00

- Tipi di variabili: numeriche, logiche, testuali, fattori
- Vettori e matrici: creazione, indicizzazione e operazioni
- Dataframe: selezione, aggiunta e rimozione di variabili
- Liste e loro utilizzo
- Esercizi guidati su dataset semplici

Lezione 3 – Manipolazione dei dati (2h) - 22 ottobre 2025 ore 14:00-16:00

- Importare dati da CSV, Excel, SPSS, database
- Esportare dati e risultati
- Gestione valori mancanti
- Creazione, modifica e ricodifica variabili
- Introduzione a `dplyr` per filtrare, selezionare e ordinare dati

Lezione 4 – Statistica descrittiva e visualizzazione (2h) - 24 ottobre 2025 ore 14:00-16:00

- Distribuzioni di frequenza
- Indici di posizione e variabilità
- Statistiche riassuntive in R (`summary`, `psych::describe`)
- Grafici di base: istogrammi, boxplot, scatterplot
- Interpretazione dei risultati

Lezione 5 – Test di ipotesi parametrici (2h) - 27 ottobre 2025 ore 14:00-16:00

- Logica del test di ipotesi e p-value
- Test t per una e due medie
- ANOVA a una via
- Visualizzazione dei risultati: grafici delle medie con IC
- Applicazioni a dataset reali

Lezione 6 – Test di ipotesi non parametrici (2h) - 30 ottobre 2025 ore 14:00-16:00

- Motivazione all'uso dei test non parametrici

- Test di Wilcoxon, Mann-Whitney, Kruskal-Wallis
- Confronto tra parametrici e non parametrici
- Visualizzazione con boxplot e violin plot

Lezione 7 – Correlazioni e associazioni (2h) - 31 ottobre 2025 ore 14:00-16:00

- Correlazioni: Pearson, Spearman, Kendall
- Matrici di correlazione e heatmap
- Associazioni per dati categoriali (chi-quadrato)
- Visualizzazione: scatterplot con retta di regressione

Lezione 8 – Regressione lineare e applicazioni finali (2h) - 4 novembre 2025 ore 14:00-16:00

- Regressione lineare semplice con `lm()`
- Interpretazione dei coefficienti e del R^2
- Diagnostica del modello (residui, grafici)
- Cenno alla regressione multipla
- Esercitazione completa: dal caricamento dati alla regressione finale