# Programma del Seminario “Elementi di statistica e introduzione a R”

Il seminario proposto come attività formativa per il Collegio Borromeo presenta il **1° modulo** di un corso intensivo dal titolo “[Introduzione alla biostatistica con R](https://r4biostats.com/)” –– specificamente adattato ad un pubblico di studenti universitari. La proposta formativa si distingue per il taglio pratico, focalizzato sull’analisi di dati reali provenienti da studi in ambito biologico, medico e clinico.

Il programma prevede una revisione dei concetti fondamentali di statistica e metodologia della ricerca epidemiologica, unita a un’introduzione pratica all’uso di R, un software *open-source* sempre più diffuso nelle scienze della vita.

## Programma proposto

**LEZIONE: elementi di statistica descrittiva e introduzione a R**

Durata approssimativa: 2 ore

* Introduzione al linguaggio **R** utilizzando l’ambiente di sviluppo integrato (IDE) **RStudio**
* Principi e strumenti per l’analisi riproducibile dei dati
* Tipologie di variabili e livelli di misurazione
* Metodi di pulizia e manipolazione dei dati
* Cenni di statistica descrittiva
  + Misure di tendenza centrale, misure di variabilità (o dispersione) e distribuzioni di frequenza
* Criteri e strumenti per la visualizzazione dei dati quantitativi
  + La libreria **ggplot2** per la creazione di grafici in R

**ESERCITAZIONE con R**

Durata approssimativa: 2 ore

I concetti illustrati nella lezione vengono “rivisitati” in una corrispondente esercitazione pratica con il software R, basata su esempi di dati clinici reali. I materiali forniti (comprensivi di codice R) permetteranno ai partecipanti di consolidare quanto appreso e riutilizzare i contenuti in modo autonomo.

**APPROFONDIMENTI**

Durata approssimativa: 1 ora

* Introduzione al disegno degli studi epidemiologici
* Discussione sui principali approcci analitici:
  + Analisi causale e modelli predittivi
  + Esempi introduttivi di algoritmi di Machine Learning
* Conclusioni e domande

## Prerequisiti

Non vi sono prerequisiti specifici, ma il seminario si rivolge idealmente a studenti che abbiano familiarità con analisi quantitative nell’ambito delle scienze della vita.

Per approfittare al massimo del tempo disponibile, l’installazione di R e RStudio dovrà essere effettuata prima dell’inizio del seminario (a tal fine, verranno fornite istruzioni dettagliate).