

OSDA-4 - LINGUAGGIO R

Categoria: **Data & Big Data**

INFORMAZIONI SUL CORSO



Durata:
3 Giorni



Categoria:
Data & Big Data



Qualifica Istruttore:
Docente Senior (min.
5 anni)



Dedicato a:
Sviluppatore



Produttore:
PCSNET

OBIETTIVI

R è un ambiente Open Source per l'analisi statistica dei dati, la manipolazione dei dati, il calcolo e la loro rappresentazione grafica.

E' un sistema integrato di utilities che può essere installato su piattaforme MS Windows, Linux e MacOS.

Il corso intende fornire una competenza teorica e applicativa di base del software R.

Attraverso applicazioni ed esercitazioni pratiche permetterà di acquisire le nozioni fondamentali per l'utilizzo dell'ambiente R e fornirà ai discenti una solida base di istruzioni operative e procedure per l'analisi statistica descrittiva e inferenziale.

PREREQUISITI

Per la partecipazione è richiesta una conoscenza di base dei concetti di statistica descrittiva e inferenziale.

CONTENUTI

Introduzione a R

- La filosofia del software R, punti di forza e obiettivi
- Caratteristiche principali
- Pannelli e barre degli strumenti
- La console di R e i workspace
- Acquisizione di un file dati esterno al sistema
- Uso dell'help e ricerca di comandi, documentazione utile
- Cenni sui packages sviluppati in ambiente R

Utilizzo del software R

- Download, installazione e aggiornamento di R
- Download, installazione e aggiornamento di R Studio
- Muoversi in R Studio: tour dell'interfaccia e dei 4 pannelli
- Installazione e caricamento dei package
- Personalizzare le sessioni R Studio con .Rprofile
- Salvataggio ed esportazione dei dati e dei risultati
- Esercizi pratici in aula

Importazione ed esportazione di dati in R

- Lettura dei dati da file o DB
- Scrittura dei dati su file o DB
- Importazione dei dati (vari formati)
- Esportazione dei dati (vari formati)
- Funzioni di esame dei dati
- Esercizi pratici in aula

Fondamenti del software R

- Classi, oggetti e metodi
- I tipi di dati
- Gli scalari
- I vettori
- Le matrici
- Le liste
- I fattori
- Le array
- I dataframe
- Creazione degli oggetti
- Indicizzazione (metodi)
- Mascheramento degli errori
- Operatori di assegnazione
- Esercizi pratici in aula

Introduzione all'uso delle tecniche statistiche per l'analisi dei dati in R

- Struttura dei dati
- Natura delle variabili
- Costruzione di un vettore e una matrice dati in ambiente R
- Sintesi numerica delle informazioni: indici di posizione e di variabilità
- Sintesi grafica delle informazioni e principali comandi di R
- Cenni sulla trasformazione di variabili
- La standardizzazione

Elaborazione dei dati in R

- Vettorizzazione
- Sequenze regolari
- Date
- Trasformazioni numeriche
- Stringhe e loro manipolazione
- Booleani e confronti
- Gli insiemi vuoti
- I dati mancanti
- Sort & order
- Manipolazione di matrici
- Esercizi

Le funzioni in R

- Cosa sono
- La struttura

- Come crearle
- Esercizi

INFO

Materiale didattico: Materiale didattico e relativo prezzo da concordare

Costo materiale didattico: NON incluso nel prezzo del corso

Natura del corso: Operativo (previsti lab su PC)