

# Coding e Big Data

2025-2026

Vincenzo Nardelli



[vincenzo.nardelli@unicatt.it](mailto:vincenzo.nardelli@unicatt.it)

# Analisi di big data e programmazione

## Efficienza

- ▶ Le interfacce punta e clicca non sono efficienti in termini di tempo
- ▶ Automatizzare significa velocizzare le operazioni
- ▶ Riproducibilità
- ▶ Crescente necessità di fornire dati, materiali ed analisi insieme ad i risultati
- ▶ Assicura la possibilità di controllare i risultati e le procedure
- ▶ Rende possibile effettuare analisi in produzione

# R & Python

- ▶ Open source
- ▶ Gratuito e liberamente utilizzabile
- ▶ Strumenti avanzati
- ▶ Pacchetti e librerie per ogni tipo di analisi
- ▶ Documentazione e comunità
- ▶ Nessun supporto cliente a pagamento ma comunità!

# R vs Python

## R

- ▶ Data analytics, statistica
- ▶ Usato da statistici e dalla ricerca
- ▶ 12000 package on CRAN
- ▶ Semplice comunicazione (visualization, reporting and dashboard)

## Python

- ▶ Deployment and production
- ▶ Usato da programmatori e sviluppatori
- ▶ Integrazione con diversi sistemi operativi
- ▶ Algoritmi complessi e struttura ad oggetti

# R e RStudio

## Linguaggio



Motore

## IDE

(integrated development environment)



Cruscotto

# Python

## Linguaggio



Motore

## IDE

(integrated development environment)



Visual Studio Code



SPYDER



Cruscotto

# Curva di apprendimento

## EXCEL

La maggior parte delle persone probabilmente ha già appreso almeno alcune nozioni di base in Microsoft Excel. Questo è un vantaggio sostanziale dell'utilizzo di Excel: la curva di apprendimento iniziale è piuttosto minima e la maggior parte delle analisi può essere eseguita puntando e facendo clic sul pannello superiore. Una volta che un utente ha importato i propri dati nel programma, non è eccessivamente difficile creare grafici e diagrammi di base.

# Curva di apprendimento

R

R è un linguaggio di programmazione, tuttavia, il che significa che la curva di apprendimento iniziale è più ripida. Ci vorranno almeno alcune settimane per familiarizzare con l'interfaccia e padroneggiare le varie funzioni. Fortunatamente, l'uso di R può diventare rapidamente una seconda natura con la pratica.



# Perché programmare invece di usare Excel?

- ▶ R aiuta a leggere qualsiasi tipo di dato disponibile.
- ▶ L'automazione è molto più semplice rispetto a Excel.
- ▶ Supporta set di dati più grandi
- ▶ Offre anche un calcolo più veloce
- ▶ Le capacità di manipolazione dei dati sono potenti rispetto a Excel.
- ▶ Organizzazione del prodotto più semplice
- ▶ R può anche aiutare a rilevare qualsiasi tipo di errore (riproducibilità)
- ▶ È gratuito e open source. Quindi non è necessario pagare come Excel.
- ▶ R fornisce funzionalità statistiche avanzate.

# Installazione e configurazione di R e RStudio

- ▶ [https://www.vincnardelli.com/cbd/guide/installazione\\_r\\_rstudio](https://www.vincnardelli.com/cbd/guide/installazione_r_rstudio)
- ▶ [https://www.vincnardelli.com/cbd/guide/usare\\_r\\_rstudio](https://www.vincnardelli.com/cbd/guide/usare_r_rstudio)