



Il linguaggio Python per l'analisi dei dati

Un po' di storia

- Ideato da Guido van Rossum, primo rilascio ufficiale nel 1991.
- Linguaggio general-purpose
- Ad oggi ha una vastissima gamma di librerie disponibile
- Nel 2006 è rilasciata la prima versione di NumPy
- Nel 2008 di Pandas
- Nel 2010 di scikit-learn
- Nel 2014 di Seaborn
- Nel 2023 di Ollama

Come sarà organizzate le lezioni?

- Una parte relativa alle basi del linguaggio Python
- Una parte relativa alla Data Analysis
- Una parte relativa alle librerie di Python per la Data Analysis

Attività di analisi dei dati (con Python e non)

- Acquisizione dei dati da svariate fonti
- Pulizia e trasformazione dei dati
- Analisi dichiarative
- Presentazione grafica delle informazioni
- Analisi predittive e Intelligenza Artificiale

Acquisizione dei dati da svariate fonti

- CSV
- XLSX
- JSON
- File di testo strutturati
- Database relazionali
- Database non relazionali
- API

Pulizia e trasformazione dei dati

- Controllo dei tipi
- Pulizia di caratteri speciali
- Uniformità dei dizionari
- Rimozione dei duplicati
- Creazione di nuove colonne
- Verifica degli outlier
- Normalizzazione dei dati
- Filtrare righe non di interesse
- Combinare fonti dati differenti

Analisi dichiarative

- Filtrare i dati
- Raggruppare i dati e aggregarli
- Combinare fonti dati
- Ordinare i dati
- Creare nuove colonne

Ad esempio: quali sono le tre fatture di importo netto più grande tra quelle emesse nel 2024 a clienti della regione Lombardia?

Presentazione grafica delle informazioni

- Rappresentazione di una variabile qualitativa
- Rappresentazione di una variabile quantitativa
- Rappresentazione di più variabili qualitative
- Rappresentazione di più variabili quantitative
- Rappresentazioni miste

Analisi predittive e Intelligenza Artificiale

- Classificazione
- Regressione
- Clustering
- Analisi delle anomalie
- Regole di associazione

Ad esempio: creare un algoritmo di IA che permetta di classificare i prestiti in pagati o non pagati in base allo studio di uno storico esistente