

53. Scrittura su file

Scrivete un programma che richiede in ingresso cognome, nome e matricola di 5 studenti.

Il programma scrive i dati su un file di testo, separando con una virgola il cognome, il nome e la matricola e andando a capo dopo aver inserito i dati di una persona. I dati salvati su file dovranno apparire in questo modo:

```
nome1,cognome1,matricola1,  
nome2,cognome2,matricola2,  
nome3,cognome3,matricola3
```

54. Lettura e scrittura su file

Scrivete un programma che richiede in input una sequenza di numeri interi diversi da zero (terminati da uno 0 che comunica che l'immissione è finita). Il programma deve scrivere questi numeri su un file di testo, separandoli da una virgola. Scrivete un secondo programma che legge il file scritto dal programma precedente, carica i valori numerici in una lista e la stampa a video.

55. File CSV

Troverete nella cartella di rete condivisa (e dopo il termine della lezione anche nel sito web del corso) un file dal nome "dataset_persone.zip", il quale all'interno contiene il file dataset_cognomi_nomi_eta_comunenascita.txt. Quest'ultimo è un file in formato CSV (comma separated value) che contiene al suo interno informazioni su un insieme di persone. Ogni riga del file riporta i dati di una singola persona, le informazioni presenti in una riga sono nell'ordine: nome, cognome, genere, età, comune di nascita.

Qui sotto potete vedere un estratto del file CSV

```
FABIO;GILARDI;M;22;ALATRI
```

```
STEFANIA;LICCIARDELLO;F;12;MONTENERO SABINO
```

```
ANNALENA;MAINI;F;43;CASALUCE
```

Utilizzando i dati contenuti nel file, scrivete dei programmi in python in grado di fornire dei risultati utili per rispondere alle domande riportate qui di seguito.

Importante: leggete le note finali prima di iniziare gli esercizi.

A. Quantità totali e distinte per genere

Quante sono le persone descritte nel file? Quante sono le persone di sesso maschile, quante le persone di sesso femminile?

B. Fasce d'età

Tutte le persone descritte nel file hanno un'età compresa tra 1 e 59 anni. Date 6 fasce d'età (1-9, 10-19, 20-29, 30-39, 40-49, 50-59), individuate quale è la numerosità (quante persone fanno parte) di ogni fascia d'età.

C. Numerosità cognomi

Per ogni cognome presente, stampate a video la numerosità associata al cognome. Ogni cognome deve apparire a video una sola volta.

D. Cognome, nome, luogo di nascita più diffusi

Individuate il cognome più diffuso e la sua numerosità (la numerosità delle persone che hanno tale cognome); individuate il nome più diffuso e la sua numerosità, individuate il comune di nascita più diffuso e la numerosità. Nota: si tratta di 3 esercizi distinti.

Note

- Copiate il file .zip in locale ed estraete il contenuto nella directory in cui lavorerete

- Il file contenente il dataset è stato generato casualmente a partire dall'elenco dei nomi e cognomi più diffusi in Italia nel 2006 e dall'elenco dei comuni italiani (entrambi disponibili sul sito web dell'istat)
- Per svolgere gli esercizi, occorre aver letto/studiato le slide della lezione sui file, in modo particolare è fondamentale utilizzare i metodi “strip” e “split”.

All'inizio, non vi conviene lavorare con il dataset completo dei dati. Createvi un file ridotto (per esempio con solo le prime 50 righe del file originale) e lavorate su quello. Quando avrete acquisito maggiore esperienza,

56. Esercizi a piacere

Svolgete delle analisi a vostro piacere sul dataset precedente

57. File CSV 2 (film)

Troverete nella cartella di rete condivisa (e dopo il termine della lezione anche nel sito web del corso) un file dal nome “catalogo_film.zip”, il quale all'interno contiene il file catalogo_film.xls. Quest'ultimo è un file excel, dal quale dovrete estrarre il corrispondente file in formato CSV (comma separated value). Ogni riga del file riporta i dati di un singolo film, le informazioni presenti in una riga sono nell'ordine: codice, genere, titolo, regista, attori, prezzo.

Utilizzando i dati contenuti nel file, svolgete delle analisi a vostro piacere, per es. quanti sono i film venduti ad un certo prezzo, divisi per genere; quali film di un certo regista, ecc.

Note

- Copiate il file .zip in locale ed estraete il contenuto nella directory in cui lavorerete
- Il file contenente il dataset è stato generato casualmente a partire dall'elenco dei film di una videoteca
- Attenzione che alcuni campi (soprattutto nelle colonne: regista, attori) possono essere vuoti. Inoltre le stringhe che riportano il titolo e il genere possono essere interrotte, o con parole abbreviate, tenetene conto nelle analisi che svolgerete