

# Esame di Linguistica Computazionale e Laboratorio - Modulo A

Simulazione 12 dicembre 2023

Docente: Daniele Francesco Santamaria

## Istruzioni

Il codice Python deve trovarsi all'interno di un unico file chiamato **Cognome\_Nome.txt**. Il file va inviato all'indirizzo **daniele.santamaria@unict.it**. Non è possibile usare appunti, esercizi svolti, altro materiale o comunicare con altra candidate.

## Testo della prova

Si consideri il file CSV disponibile all'indirizzo

<https://pastebin.com/raw/RtwHi7yN>

che descrive alcune rilevazioni su opere letterarie.

Sfruttando il linguaggio Python, si risolvano i seguenti punti.

1. Scrivere una funzione che legga i dati nel csv e restituisca una lista di liste, una per ogni riga del CSV. Ogni elemento della sottolista contiene una tripla costituita dal nome della colonna, dal numero di **colonna** e dal valore della cella della riga data. Escludere la riga contenente i nomi delle colonne. (MAX 5 punti).
2. Scrivere una funzione che presa la struttura al punto 1 restituisca la lista delle opere (colonna "book\_or\_manuscript") pubblicate dopo il 1960 (colonna "year\_of\_publish"). Notare che è possibile trovare nella colonna "year\_of\_publish" valori nella forma "1460-1470" o nella forma "1960s". (MAX 5 punti).
3. Scrivere una funzione che presa la struttura al punto 1 restituisca la lista delle opere (colonna "book\_or\_manuscript") con autore sconosciuto (colonna "author"). (MAX 5 punti).
4. Scrivere una funzione che presa la struttura al punto 1 restituisca i nomi degli autori (colonna "author") che hanno venduto copie per più di un milione di dollari (colonna "sale\_price\_in\_millions\_of"). Notare che i valori della colonna "sale\_price\_in\_millions\_of" contengono il simbolo "\$" e il simbolo ".". (MAX 10 punti).
5. Scrivere una funzione che presa la struttura al punto 1 e il nome di un file, salvi sul file con il nome dato tutte le righe della struttura, riga per riga. (MAX 5 punti).