

ESERCITAZIONE ACQUISIZIONE DATI

Corso di Data Analytics

15 Ottobre 2020

1. Caricare in un DataFrame Pandas il dataset contenuto del file "bank_dataset.csv". Stampare a video il numero di righe e di colonne.
2. Assicurarsi che il dataset NON contenga valori mancanti (NaN). Nel caso, rimuovere le righe associate ai valori mancanti. Stampare il numero di righe dopo l'operazione.
3. Assicurarsi che il dataset NON contenga righe duplicate, nel caso rimuoverle. Stampare il numero di righe dopo l'operazione.
4. Stampare età media, minima e massima degli utenti presenti nel dataset.
5. Contare gli utenti che hanno un grado di educazione (colonna education) pari a university.degree.
6. Stampare, per ogni categoria professionale (colonna job), il numero di utenti che hanno aperto un conto bancario (ossia per i quali l'ultima colonna è pari a y).
7. Trasformare il DataFrame effettuando l'one-hot encoding della variabile categorica education. Stampare il numero di colonne del DataFrame risultante.