

Si consideri il seguente scenario di classificazione binaria in cui si vuole predire se una persona si iscriverà al corso di Inglese di livello B2

Età	Data di Nascita	Titolo di Studio	Diplomato?	Spesa Corsi	Parteciperà?
39	14/12/1784	Laurea Magistrale	Si	5000	Si
50	01/07/1973	Diploma	Si	2000	Si
15	02/09/2008	Licenza Media	No		No
23	19/04/2000	Diploma		1275	No
	30/03/1992	Diploma	Si	2000	No
42	15/11/1981	Laurea	Si		Si
17	28/10/2006	Licenza Media	No	750	Si

Esercizio 1

Indica se applicando il criterio di cancellazione delle righe si ottiene una distorsione nei dati. Motivare la risposta.

Esercizio 2

Applica tutti i possibili metodi deduttivi per ottenere alcuni dei valori mancanti. Riporta il dataset risultante.

Esercizio 3

Prova ad imputare i valori mancanti dell'attributo Spesa Corsi, sostituendo con i valori di media o moda. N.B.: devono essere riportate entrambe le tabelle risultati.