

Cognome e Nome	Matr.
----------------	-------

(*) **Tempo: 40 minuti**(**) Gli studenti con un **numero di matricola DISPARI** usino il dataset “mutanti_dispari.txt”, mentre gli studenti con un **numero di matricola PARI** usino “mutanti_pari.txt”(***) Sul foglio riportate per ogni domanda, la risposta e i comandi utilizzati per rispondere.

(****) Per rispondere potete usare anche il retro dei fogli.

Nel dataset **mutanti** sono raccolte le informazioni riguardanti la popolazione del villaggio di Thamos che è stato recentemente invaso dall'esercito di Slymers formato da mutanti antropomorfi.

Il CISM (Centro Italiano Studi sui Mutanti) ha condotto un'indagine sulla popolazione di Thamos per analizzare le differenze tra i Mutanti e i Non Mutanti, in particolare ha raccolto dati sulla lunghezza degli arti (cm), sulla lunghezza dei capelli (cm) e sulla forza fisica (midichlorian) di 100 abitanti del villaggio. Queste misurazioni sono raccolte nel dataset **mutanti**, contenente “**braccio**”, “**capelli**”, “**forza**” e “**tipo**”; in particolare **tipo** = “NM” per gli individui Non Mutanti e **tipo** = “M” per i Mutanti)

Il CISM pensa che i Mutanti possano essere affetti da Sansonite e chiede il vostro aiuto per confermare o smentire questa ipotesi. Per trarre una conclusione adeguata rispondete alle seguenti domande riportando i principali comandi utilizzati e le opportune considerazioni e/o conclusioni statistiche.

D 0 Importare il dataset contenuto nel file “mutanti_pari.txt” o “mutanti_dispari.txt” a seconda del vostro numero di matricola.

Risposta:

D 1 Calcolare gli indici di posizione che conoscete della variabile “**forza**”

Risposta:

D 2 Rappresentare l'istogramma della “**forza**” con 12 classi (intervalli) di uguale lunghezza nell'intervallo (0,600) (attenzione: 0 e 600 non sono il minimo e il massimo della variabile). Indicare le 5 classi con la maggiore frequenza assoluta e riportare il valore di tale frequenza. Qual è la classe modale?

Risposta:

D 3 Riportare la tabella delle frequenze relative della variabile categorica “**tipo**”

Risposta:

D 4 Proporre una rappresentazione grafica della distribuzione della variabile “**forza**” distinguendo in base alla categoria “**tipo**”. Cosa si può concludere?

Risposta:

D 5 Calcolare la correlazione tra la “**forza**” e la lunghezza dei “**capelli**” per il dataset completo.

Risposta:

D 6 Calcolare la correlazione tra la “**forza**” e la lunghezza dei “**capelli**” distinguendo tra Mutanti e Non mutanti. Commentare i risultati.

Risposta:

D 7 Analizzare la normalità dei dati riguardanti la lunghezza del “**braccio**”

Risposta:

D 8 Costruire un modello di regressione per spiegare il legame lineare tra la “forza” e tutte le altre variabili (“braccio”, “capelli” e la variabile categorica “tipo”). Effettuare l’opportuna selezione dei predittori influenti. Commentare la significatività del modello finale e darne un’interpretazione alla luce dell’ipotesi del CISM.

Risposta:

D 9 Analizzare i residui del modello finale selezionato al punto 8.

Risposta:

D 10 Prevedere sia puntualmente che con un intervallo di livello 90% il nuovo valore della Forza di un individuo Mutante con le seguenti caratteristiche:

Lunghezza braccio = 45 cm e Lunghezza capelli = 35 cm

Risposta: