

GRUPPO K

Il dataset scelto è stato reperito da <https://vincentarelbundock.github.io/Rdatasets/datasets.html>.

Qui, dopo aver osservato i molti altri dataset presenti, ne è stato scelto uno che riguarda 25 qualità della personalità che sono state analizzate su un campione di 2800 persone. Il titolo è “25 Personality items representing 5 factors”.

Gli attributi sono 25 caratteristiche della personalità di una persona, le quali a grandi linee riescono a coprire tutte le qualità caratteriali e personali di una persona. A questi sono inoltre aggiunti tre attributi demografici come età, sesso e livello di istruzione.

La documentazione fornisce informazioni riguardo il tipo di contenuto del dataset ed emerge infatti, dopo averlo collegato con python e fatto stampare, che i dati sono tutti attribuiti in formato numerico intero. Più precisamente da 1 a 6 per le 25 qualità caratteriali, 1 e 2 per il sesso e da 1 a 5 per il livello di istruzione.

Può essere interessante analizzare il dataset per rilevare delle eventuali correlazioni tra le emozioni di una persona in funzione con l'età, verificando ad esempio se la persona ha delle caratteristiche intrinseche oppure viene influenzata dal contesto in cui vive. Si può inoltre verificare se il livello di istruzione può alterare o meno la peculiarità di una persona e ovviamente verificare se il sesso ha un peso rilevante nella personalità di una persona oppure se la società ne ha uno maggiore.

Come accennato sopra, il dataset è stato già testato su python ed è stato verificato il contenuto. Per queste operazioni non sono stati riscontrati problemi, ma è stato sufficiente applicare le tecniche per collegare un normale dataset, come visto fino ad ora.