L'esercizio è mirato a prendere confidenza con il concetto di dato e informazione. Scegli un argomento di tuo interesse. Riporta su un documento testuale quali informazioni ti piacerebbe derivare cercando su internet i potenziali dataset strutturati e non strutturati (almeno 3 del primo tipo, 2 del secondo) utilizzabili. Per ogni dataset strutturato individua gli attributi e l’identificativo; per ogni dataset non strutturato descrivi le analisi che svolgeresti su di esso e l’obiettivo di queste analisi.

Argomento scelto: Stagione calcistica Serie A 2023/2024

Dataset strutturati:

* Classifica del campionato
* Classifica dei marcatori
* Classifica degli assist man

Dataset non strutturati:

* Post su Instagram della pagina della Lega Serie A
* Post su Facebook della Roma

<https://www.legaseriea.it/it/serie-a/classifica>

Identificativo: Classifica del campionato

Attributi: Posizione, Nome della squadra, Punteggio, Partite Giocate, Partite Vinte, Partite Perse, Partite Pareggiate, Gol Fatti, Gol Subiti, Differenza Reti, Trend

<https://www.repubblica.it/sport/dirette/calcio/serie-a-2023/marcatori>

Identificativo: Classifica dei marcatori

Attributi: NR, Giocatore, Squadra, Gol Fatti, Rigori

<https://www.kickest.it/it/serie-a/statistiche/giocatori/assist/2023-2024>

Identificativo: Classifica degli assist man

Attributi: #, Giocatore, Posizione, Squadra, Assist Fatti

Per entrambi i dataset non strutturati andrei a svolgere un’analisi sia sul post in sé (cercando di ricavare dati testuali se presenti e classificarli per tipologia) che sui commenti espressi (in quel caso andrei a catalogare quando sono stati fatti, la tipologia del commento) al fine di poter avere un’immagine di come i vari post sono stati percepiti dagli utenti