

Pràctica 1 - Tipologia i cicle de vida de les dades

Raúl Morcillo López

06/04/2022

Contents

Jugadors de futbol més valuosos segons Transfermarkt	2
Context	3
Descripció del dataset	3
Representació gràfica	3
Contingut	4
Agraïments	5
Inspiració	5
Llicència	5
Codi i dataset	6

Jugadors de futbol més valuosos segons Transfermarkt



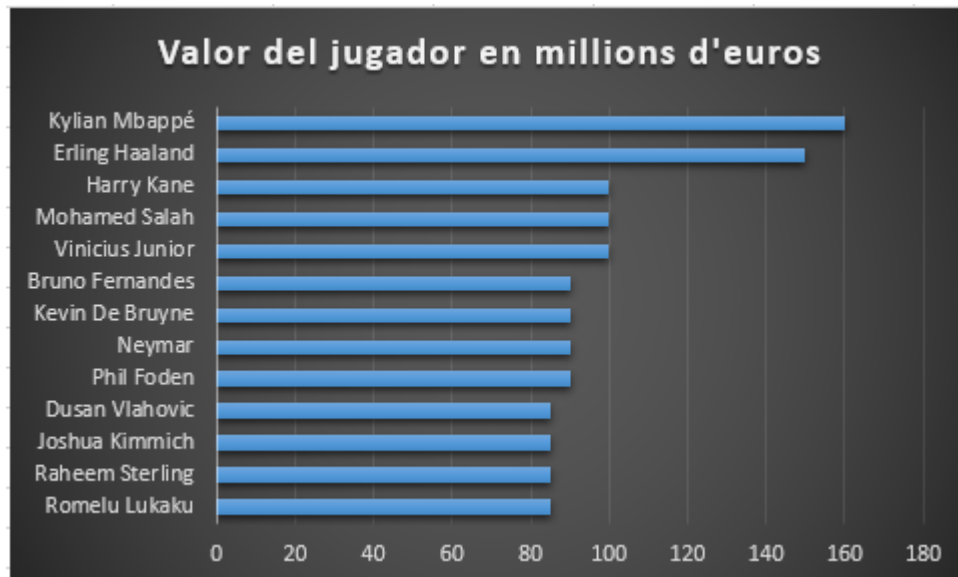
Context

Volem saber com es decideix el valor d'un jugador de futbol. Per poder tenir una idea de com passa això anem a buscar informació a la web Transfermarkt ([Link](#)). A aquesta web podem trobar informació sobre els jugadors de futbol, el valor actual i estadístiques actuals relacionades amb les seves intervencions en els partits del seu club.

Descripció del dataset

Dataset amb la informació dels 500 jugadors de futbol més valuosos segons Transfermarkt a data 30/03/2022. En aquest dataset es troba informació rellevant per entendre el valor de cada jugador segons les estadístiques dels partits jugats a la temporada actual (2021/2022) en l'àmbit de clubs.

Representació gràfica



Contingut

El dataset consta de 500 observacions i 16 atributs:

- **Nombre:** Nom del jugador.
- **Foto:** Link amb la fotografia del jugador.
- **Valor:** Valor del jugador (millions d'euros).
- **Posicion:** Posició que el jugador ocupa al camp.
- **Edad:** Edat del jugador.
- **Club:** Club de futbol al que pertany el jugador.
- **Escudo:** Link amb la imatge de l'escut del club.
- **Nacionalidad:** Nacionalitat del jugador (pot tenir més d'una).
- **Bandera:** Link amb la imatge de la bandera (pot tenir més d'una).
- **Partidos_jugador:** Partits jugats amb el seu club a la temporada 2021-2022 fins al 30/03/2022.
- **Goles:** Gols marcats amb el seu club a la temporada 2021-2022 fins al 30/03/2022 (excepte porters).
- **Asistencias:** Assistències realitzades amb el seu club a la temporada 2021-2022 fins al 30/03/2022 (excepte porters).
- **Min_gol:** Minuts per gol amb el seu club a la temporada 2021-2022 fins al 30/03/2022 (excepte porters).
- **Min_jugados:** Minuts jugats amb el seu club a la temporada 2021-2022 fins al 30/03/2022.
- **Goles_encajados:** Gols encaixats amb el seu club a la temporada 2021-2022 fins al 30/03/2022 (només porters).
- **Partidos_cero:** Total partits sense encaixar gols amb el seu club a la temporada 2021-2022 fins al 30/03/2022 (només porters).

Les dades fan referència a la temporada 2021-2022 fins al dia 30/03/2022 que va ser el dia que van ser recopilades. El primer pas va ser recopilar la llista dels 500 jugadors més valuosos que es troba en aquest link (**Link**). Un cop obtenia la informació d'aquests jugadors vam entrar al link de cada jugador (obtingut en l'operació anterior) per tal d'obtenir les estadístiques que duen durant la temporada actual en l'àmbit de clubs. Finalment, les dades han estat emmagatzemades en un fitxer csv. Tot això s'ha dut a terme mitjançant 'web scraping' amb Python.

Agraïments

Les dades han sigut descarregades de la web Transfermarkt ==> **Link**.

Inspiració

Conèixer quins jugadors de futbol són més valuosos i veure si hi ha alguna relació entre el seu valor i les estadístiques que duen amb el seu club. Si es descarrega informació durant diversos períodes de temps es pot veure com pugen i baixen de valor els jugadors segons la seva productivitat al camp.

Llicència

Una possible llicència per a aquest conjunt de dades podria ser CC BY-SA 4.0 License. L'elecció es basa en la idoneïtat de les clàusules que s'hi presenten en relació amb el treball realitzat on:

- Es prohibeix el nom del creador del conjunt de dades generat i se n'indiquen els canvis realitzats. D'aquesta manera, es reconeix el treball de tercers i en quina mesura es van fent aportacions pel que fa al treball original.
- Es permet l'ús comercial, cosa que incrementa les possibilitats que empreses puguin interessar-se en les dades generades, permetent així, la realització de nous projectes que reportin un reconeixement a l'autor original.
- Les noves contribucions han de ser publicades sota la mateixa llicència, cosa que permet que se'l reconegui a l'autor original en tot moment i sota els mateixos termes que van ser plantejades per a ell.

Codi i dataset

El codi python per poder descarregar les dades de Transfermarkt i el dataset de la descàrrega a 30/03/2022 es troben en aquest repositori:

Github ==> **Link**

També he pujat el dataset a la plataforma Zenodo accessible amb el següent **DOI**:

DOI ==> **Link**