



Práctica 1 - Tipología y ciclo de vida de los datos

David Lucas Torres y Francisco Javier Cantero Zorita



Índice

1. Contexto
2. Descripción del dataset y representación gráfica
3. Contenido
4. Agradecimientos e Inspiración
5. Licencia y Dataset
6. Código

1. Contexto

Motivación: La importación de los datos a día de hoy y el auge de la estadística aplicada al deporte.

Objetivo: Obtener una herramienta capaz de obtener la estadísticas totales de lo jugadores en un periodo de tiempo.

Web de referencia: <https://www.basketball-reference.com/>

Herramientas utilizadas: Python | **Librerías:** BeautifulSoup, pythonrequest y pandas

Título del data set: NBAPlayerStatistics_20-21.csv

2. Descripción del dataset

Estadísticas totales de los jugadores durante la temporada 2020 - 2021

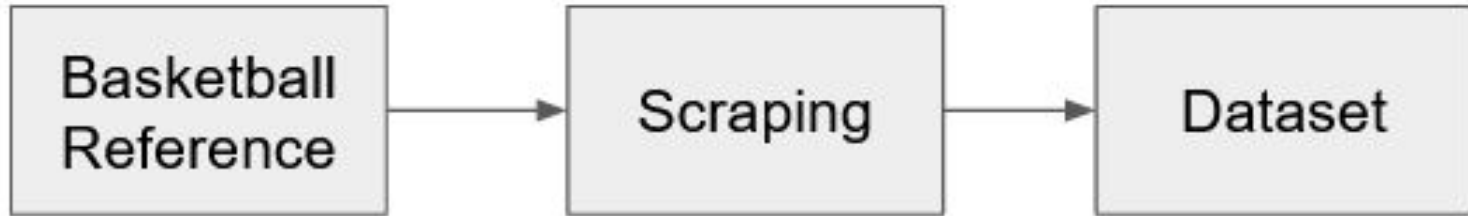
Inclusión de la variable Player Efficiency Rating (PER)

Objetivos del dataset:

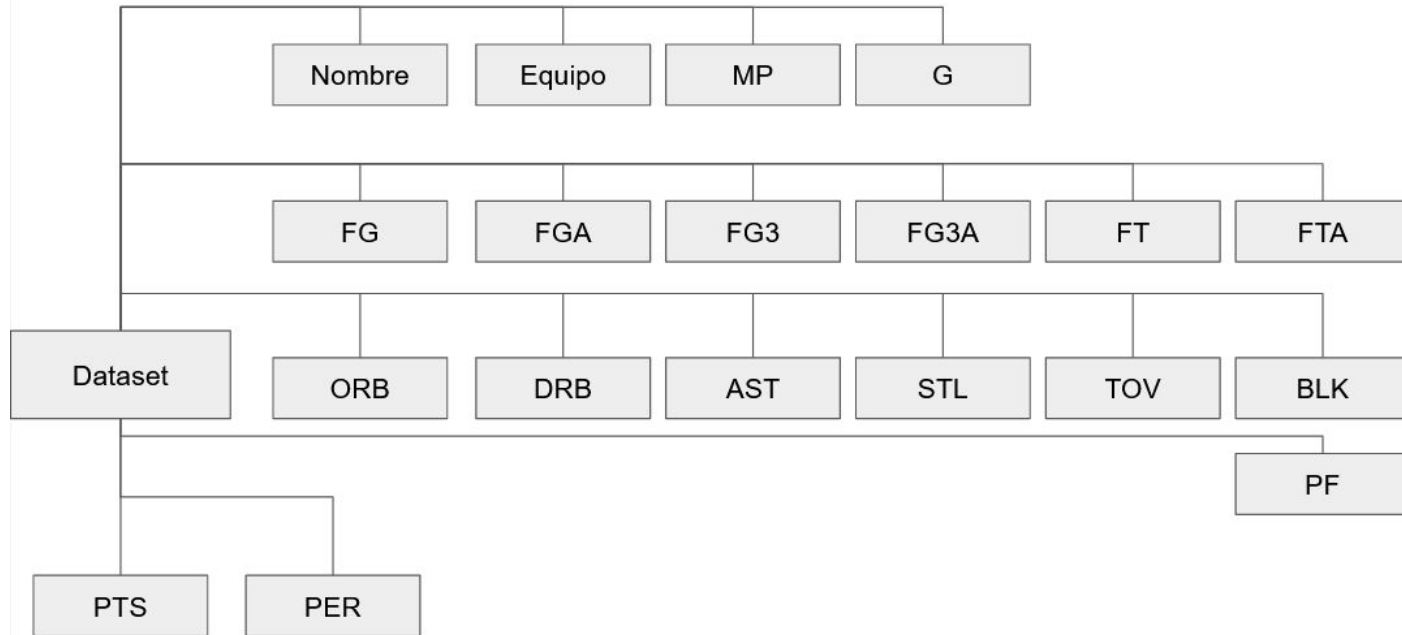
- ¿Cuales fueron las estadísticas totales de los jugadores durante la temporada 2020 - 2021
- ¿Qué jugadores han tenido mayor eficiencia durante la temporada 2020 - 2021?

3. Contenido I

Flujo de vida de los datos en el proyecto:



3. Contenido II



4. Agradecimientos e inspiración

Página de la que se han obtenido los datos: <https://www.basketball-reference.com/>

Proyectos similares:

- Basketball reference scraper: https://github.com/vishaalagartha/basketball_reference_scraper
- Basketball Reference Web Scraper: https://jaebradley.github.io/basketball_reference_web_scraper/

Objetivos del proyecto:

- ¿Cuales son las estadísticas totales de los jugadores en un periodo de tiempo?
- En un periodo de tiempo, ¿qué jugadores han tenido una mayor eficiencia?

5. Licencia y dataset

Released Under CC BY-SA 4.0 License



`10.5281/zenodo.6425685`

6. Código

Repositorio del proyecto: <https://github.com/dlucas98/PRA1-Tipologia>

Scripts del scraping:

- brScraper.py
- main.py

Ejecución de la herramienta:

- `python main.py --startDate [fecha de inicio] --endDate [fecha de final] --output [fichero de salida]`
- `python main.py -s [fecha de inicio] -e [fecha de final] -o[fichero de salida]`