

### 1 . Título del dataset

La liga de las temporadas 2012/13 a 2017/18.

### 2. Subtítulo del dataset.

Resultados de los encuentros de fútbol de los equipos de la primera división española desde la temporada 2012/13 a 2017/18

### 3. Imagen



#### 4 . Contexto. ¿Cuál es la materia del conjunto de datos?

El conjunto de datos se corresponde con los resultados de las 38 jornadas disputadas por los equipos de la 1ª división de la liga española desde la temporada 2012/13 hasta la temporada 2017/2018.

#### 5. Contenido. ¿Qué campos incluye? ¿Cuál es el periodo de tiempo de los datos y cómo se ha recogido?

A.-Los campos donde se ha almacenado la información en un fichero .csv son:

- País:** País donde se celebra la competición. En este caso: España.
- Competición:** Nombre dado a la categoría en la que se juegan las ligas de ese país.  
En este caso: La Liga
- Categoría:** Nivel de calidad donde juegan los equipos. En este caso: 1ª división.
- Temporada:** Indica en qué años se han jugado los encuentros de fútbol.
- Jornada:** número de la jornada en la que se disputó el encuentro.
- Equipo 1:** Nombre del equipo que juega en casa este encuentro.
- Resultado:** Resultado del choque deportivo correspondiendo la primera cantidad al Equipo 1 y la segunda al Equipo 2.
- Equipo 2:** Nombre del equipo que juega fuera de casa este encuentro.

B.-Período de tiempo: Desde el año 2012 hasta el 2018.

C.-Los datos han sido recogidos utilizando técnicas de web scraping del sitio web <https://www.sport.es/es/>

#### 6. Agradecimientos. ¿Quién es propietario del conjunto de datos? Incluid citas de investigación o análisis anteriores.

Los datos son de dominio público y no tienen propietario pues se corresponden con resultados deportivos. No obstante, agradecer al diario Sport el hecho de haberlos podido recopilar de su página web.

#### 7. Inspiración. ¿Por qué es interesante este conjunto de datos? ¿Qué preguntas le gustaría responder la comunidad?

Es interesante para aquellas personas aficionadas al fútbol y a las estadísticas el poder obtener los resultados deportivos de los últimos 6 temporadas aunado en un solo fichero.

También puede ser interesante para obtener modelos predictivos en el mundo de las apuestas deportivas.

8. Licencia. Seleccionad una de estas licencias y decid porqué la habéis seleccionado:

La licencia elegida es Released Under CC0: Public Domain License debido a que los datos son hechos deportivos de dominio público que no tienen dueño. Por tanto, cualquier persona que lo desee puede hacer uso de ellos libremente.

9. Código:

El código realizado en Python se encuentra en el repositorio de GitHub:

[https://github.com/ernperalta/Prac\\_1\\_WebScraping/tree/master/src](https://github.com/ernperalta/Prac_1_WebScraping/tree/master/src).

10. Dataset: Dataset en formato CSV

El dataset con los datos obtenidos están en el fichero "estadisticasFutbol.csv" que se encuentra en el repositorio de GitHub:

[https://github.com/ernperalta/Prac\\_1\\_WebScraping](https://github.com/ernperalta/Prac_1_WebScraping)

Componentes del grupo:

Esta práctica ha sido realizada de manera individual por Ernesto Peralta Macías.