

Programación para Data Science

Unidad 4: Librerías científicas en Python - pandas - Ejercicios

Ejercicio 1

Cargad los datos del fichero `historical_projections.csv` en un *dataframe*. Este conjunto de datos recoge resultados históricos (hasta 2015) sobre el *draft* de la NBA (https://en.wikipedia.org/wiki/NBA_draft) junto con valores probabilísticos extraídos de un modelo para predecir el éxito de los jugadores.

Mostrad el número de filas del *dataframe* y las etiquetas de los ejes.

```
In [1]: # Respuesta
```

Ejercicio 2

Agrupad los datos cargados en el ejercicio 1 por años. Para cada año, mostrad el número de jugadores que hay en el *draft* y la media de la probabilidad de que un jugador sea una superestrella (la probabilidad se encuentra en la columna *Superstar*).

```
In [2]: # Respuesta
```

Ejercicio 3

Mostrad los datos de los jugadores que aparecen en el *draft* de 2015 con un valor «Projected SPM» superior a 0.2 y juegan o bien en la posición escolta (en inglés, *shooting guard* o SG) o bien en la posición ala-pívot (en inglés, *power forward* o PF).

```
In [3]: # Respuesta
```

Ejercicio 4

Contad el número de jugadores (diferentes) que aparecen en el *dataframe* utilizando las funciones de la librería pandas. ¿Existen jugadores que aparecen múltiples veces en el *dataframe*?

```
In [4]: # Respuesta
```

Ejercicio 5

Añadid una nueva columna al *dataframe* con un valor booleano indicando si el jugador será o no una superestrella. Considerad que un jugador será una superestrella si la probabilidad de que lo sea es superior a 0.1.

```
In [5]: # Respuesta
```