

Desafío - Introducción a Big Data

- Para realizar este desafío debes haber estudiado el material disponibilizado de la unidad.
- Crea una carpeta de trabajo y guarda todos los archivos correspondientes (notebook y csv).
- Una vez terminado el desafío, comprime la carpeta y sube el `.zip` a la sección correspondiente.

Ejercicio 1: Ingesta de datos semiestructurados

- Utilizando las librerías `requests` y `json`, realice una consulta a la API `BallDontLie` <https://www.balldontlie.io/>.
- Mediante esta API, genere una consulta de 100 juegos (busquen el endpoint `Get All Games` en la página, dentro de `Games`). Puede usar la opción `?per_page=100` en la dirección de la API para lograr esto.
- Guarde la consulta en una variable, y asegúrese que ésta sea semiestructurada. Consulte los datos y metadatos asociados.
- Dentro de la lista de datos, comente cuáles son las llaves de cada registro.

Ejercicio 2: Organización de los datos

- En base a los datos consultados en la API, genere un objeto `pd.DataFrame` con la siguiente estructura:

```
games_df.head(2)
```

| | season | period | home | home_conference | home_division | visitor | visitor_conference | visitor_division | home_score |
|---|--------|--------|----------------|-----------------|---------------|-------------------|--------------------|------------------|------------|
| 0 | 2018 | 4 | Boston Celtics | East | Atlantic | Charlotte Hornets | East | Southeast | 126 |
| 1 | 2018 | 4 | Boston Celtics | East | Atlantic | LA Clippers | West | Pacific | 112 |

| visitor | visitor_conference | visitor_division | home_score | visitor_score | home_differential | visitor_differential |
|-------------------|--------------------|------------------|------------|---------------|-------------------|----------------------|
| Charlotte Hornets | East | Southeast | 126 | 94 | 32 | -32 |
| LA Clippers | West | Pacific | 112 | 123 | -11 | 11 |

- Para obtener la estructura, **sólo** puede utilizar operaciones `map` y comprensiones de lista.

Ejercicio 3: El efecto de jugar de local

- Genere una columna en el `pd.DataFrame` que identifique si el equipo que jugó de local ganó (1) o no (0).
- Repita el procedimiento para identificar si el equipo que jugó de visitante ganó (1) o no (0).
- Reporte cuáles son los primeros y últimos 5 equipos en cuanto a desempeño por jugar local.
- Reporte cuáles son los primeros y últimos 5 equipos en cuanto a desempeño por jugar de visita.

Ejercicio 4: Obteniendo el porcentaje de ganar local y de visita

- Genere un nuevo objeto que guarde el porcentaje de juegos ganados como local **por equipo**.
- Repita lo mismo para los juegos donde el equipo fue visitante.
- ¿Qué equipos tienen iguales chances de ganar como visitante o local?