

# Práctica 6.1

## Reto de Transformación

---

# Implementa el merge/join para juntar los datos

```
[114] teams_df = pd.read_csv('/content/Teams.csv')
teams_df.head(3)
```

	team	ranking	games	wins	draws	losses	goalsFor	goalsAgainst	yellowCards	redCards
0	Brazil	1	5	3	1	1	9	4	7	2
1	Spain	2	6	5	0	1	7	2	3	0
2	Portugal	3	4	1	2	1	7	1	8	1

```
[115] players_df = pd.read_csv('/content/Players.csv')
players_df.head(3)
```

	surname	team	position	minutes	shots	passes	tackles	saves
0	Abdoun	Algeria	midfielder	16	0	6	0	0
1	Belhadj	Algeria	defender	270	1	146	8	0
2	Boudebouz	Algeria	midfielder	74	3	28	1	0

```
df = pd.merge(teams_df, players_df)
df.head(3)
```

	team	ranking	games	wins	draws	losses	goalsFor	goalsAgainst	yellowCards	redCards	surname	position	minutes	shots	passes	tackles	saves
0	Brazil	1	5	3	1	1	9	4	7	2	Baptista	midfielder	82	0	42	1	0
1	Brazil	1	5	3	1	1	9	4	7	2	Daniel Alves	defender	310	11	215	6	0
2	Brazil	1	5	3	1	1	9	4	7	2	Elano	midfielder	140	5	57	6	0



# Implementa al menos 2 técnicas de filtrado o selección

```
# obtenemos el total de shots para cada posición de cada equipo
shots_by_position = df.groupby(['team', 'position'])['shots'].sum().to_frame().reset_index()

# obtenemos un df con el total de shots para cada equipo
total_shots = shots_by_position.groupby('team')['shots'].sum().to_frame().reset_index()
total_shots = total_shots.rename(columns = {'shots': 'total_shots'}) # Renombramos la columna shots por total_shots

# unimos ambos dataframes
shots_by_position = pd.merge(shots_by_position, total_shots, how = 'inner', on = 'team')

# se obtiene el porcentaje de shots para cada posición de cada equipo
shots_by_position['ratio_shots'] = shots_by_position['shots'] / shots_by_position['total_shots']
shots_by_position = shots_by_position[['team', 'position', 'ratio_shots']]
shots_by_position
```

	team	position	ratio_shots
0	Algeria	defender	0.161290
1	Algeria	forward	0.225806
2	Algeria	goalkeeper	0.000000
3	Algeria	midfielder	0.612903
4	Argentina	defender	0.115942
...	...	...	...
123	USA	midfielder	0.490909
124	Uruguay	defender	0.071429
125	Uruguay	forward	0.657143
126	Uruguay	goalkeeper	0.000000

```
# Calculamos el porcentaje de juegos ganados para cada equipo
porcentaje_juegos_ganados = (df.groupby(['team'])['wins'].sum() / df.groupby(['team'])['games'].sum()).to_frame().reset_index()

# Filtramos para obtener aquellos equipos con un porcentaje de juegos ganados mayor al 50%
mejores Equipos_df = porcentaje_juegos_ganados[porcentaje_juegos_ganados[0] > 0.5 ]
mejores Equipos_shots_per_position = pd.merge(shots_by_position, mejores Equipos_df, how='inner')

# renombramos columnas
mejores Equipos_shots_per_position = mejores Equipos_shots_per_position.rename(columns = {0: 'ratio_wins'})
mejores Equipos_shots_per_position
```

	team	position	ratio_shots	ratio_wins
0	Argentina	defender	0.115942	0.800000
1	Argentina	forward	0.652174	0.800000
2	Argentina	goalkeeper	0.000000	0.800000
3	Argentina	midfielder	0.231884	0.800000
4	Brazil	defender	0.408451	0.600000
5	Brazil	forward	0.323944	0.600000
6	Brazil	goalkeeper	0.000000	0.600000
7	Brazil	midfielder	0.267606	0.600000
8	Germany	defender	0.045455	0.666667
9	Germany	forward	0.621212	0.666667
10	Germany	goalkeeper	0.000000	0.666667
11	Germany	midfielder	0.333333	0.666667



# Encuentra 3 insights a compartir

---

- Los cinco equipos con un porcentaje de juegos ganados mayor a 50% son: Argentina, Brazil, Germany, Netherlands, Spain los cuales se encuentran en el top 10 del Ranking.
- Estos equipos presentan un porcentaje arriba del 60% de shots realizados por las posiciones forward y midfielder. Más aún, el porcentaje de shots hechos por la posición forward es en cuatro de los cinco equipos aproximadamente el mismo que el porcentaje de shots hecho por la posición midfielder.
- Los equipos con mejor porcentaje de juegos ganados, presentan un porcentaje arriba del 70% de passes realizados por las posiciones defender y midfielder. Más aún, el porcentaje de passes hechos por ambas posiciones es aproximadamente el mismo.
- La posición midfielder parece ser una posición clave para tener un alto porcentaje de juegos ganados, por lo cual es recomendable que los equipos se enfoquen en entrenar a estos jugadores para tener las habilidades de hacer passes y shots.