In [ ]:
```python
import pandas as pd
import requests

request = requests.get("https://pt.wikipedia.org/wiki/La_Liga_de_2023%E2%80%9324")

tables = pd.read_html(request.text)

# for table in tables:
#     display(table)

league_table = tables[8]
game_schedule = tables[11]
display(league_table)
display(game_schedule)
```

	Pos	Equipe	Pts	J	V	E	D	GP	GC	SG	Classificação ou descenso
0	1	Girona	49	20	15	4	1	46	24	+22	Fase de grupos da Liga dos Campeões da UEFA de...
1	2	Real Madrid	48	19	15	3	1	40	11	+29	Fase de grupos da Liga dos Campeões da UEFA de...
2	3	Athletic Bilbao	41	20	12	5	3	38	20	+18	Fase de grupos da Liga dos Campeões da UEFA de...
3	4	Barcelona	41	19	12	5	2	36	22	+14	Fase de grupos da Liga dos Campeões da UEFA de...
4	5	Atlético de Madrid	38	19	12	2	5	39	23	+16	Fase de grupos da Liga Europa da UEFA de 2024–25
5	6	Real Sociedad	32	20	8	8	4	31	21	+10	Play-off da Liga Conferência Europa de 2024–25
6	7	Betis	31	20	7	10	3	21	19	+2	NaN
7	8	Las Palmas	28	20	8	4	8	19	17	+2	NaN
8	9	Getafe	26	19	6	8	5	24	25	−1	NaN
9	10	Valencia	26	19	7	5	7	22	23	−1	NaN
10	11	Rayo Vallecano	23	19	5	8	6	18	24	−6	NaN
11	12	Osasuna	22	19	6	4	9	22	29	−7	NaN
12	13	Alavés	20	20	5	5	10	18	27	−9	NaN
13	14	Mallorca	19	20	3	10	7	18	24	−6	NaN
14	15	Villarreal	19	20	5	4	11	27	41	−14	NaN
15	16	Celta de Vigo	17	20	3	8	9	21	30	−9	NaN
16	17	Sevilla	16	20	3	7	10	25	30	−5	NaN
17	18	Cádiz	15	19	2	9	8	14	26	−12	Zona de rebaixamento à Segunda Divisão Espanho...
18	19	Granada	11	20	2	5	13	22	41	−19	Zona de rebaixamento à Segunda Divisão Espanho...
19	20	Almería	6	20	0	6	14	19	43	−24	Zona de rebaixamento à Segunda Divisão Espanho...

	Casa \ Fora	ALA	ALM	ATH	ATM	BAR	BET	CAD	CEL	GET	...	GRA	LPA	MLL	OSA	RAY	RMA	RSO	SEV	VAL	VIL
0	Alavés	—	1–0	0–2	NaN	NaN	1–1	NaN	NaN	NaN	...	3–1	0–1	NaN	0–2	NaN	0–1	NaN	4–3	1–0	NaN
1	Almería	NaN	—	NaN	NaN	NaN	0–0	NaN	2–3	NaN	...	2–2	1–2	0–0	NaN	0–2	1–3	1–3	NaN	2–2	NaN
2	Athletic Bilbao	NaN	3–0	—	2–0	NaN	4–2	3–0	4–3	2–2	...	NaN	1–0	NaN	NaN	4–0	0–2	2–1	NaN	2–2	NaN
3	Atlético de Madrid	2–1	2–1	NaN	—	NaN	NaN	3–2	NaN	3–3	...	3–1	NaN	1–0	NaN	NaN	3–1	2–1	1–0	NaN	3–1
4	Barcelona	2–1	3–2	1–0	1–0	—	5–0	2–0	3–2	NaN	...	NaN	NaN	NaN	NaN	NaN	1–2	NaN	1–0	NaN	NaN
5	Betis	NaN	NaN	NaN	0–0	NaN	—	1–1	NaN	NaN	...	1–0	1–0	2–0	2–1	NaN	1–1	NaN	a	3–0	1–0
6	Cádiz	1–0	1–1	NaN	NaN	NaN	NaN	—	NaN	NaN	...	NaN	NaN	NaN	1–1	0–0	0–3	0–0	2–2	NaN	3–1
7	Celta de Vigo	1–1	NaN	NaN	0–3	NaN	2–1	1–1	—	2–2	...	1–0	NaN	0–1	0–2	NaN	0–1	NaN	1–1	NaN	NaN
8	Getafe	1–0	2–1	NaN	NaN	0–0	1–1	1–0	NaN	—	...	NaN	NaN	NaN	3–2	0–2	NaN	NaN	NaN	1–0	0–0
9	Girona	3–0	5–2	1–1	4–3	NaN	NaN	NaN	1–0	3–0	...	NaN	1–0	5–3	NaN	NaN	0–3	NaN	NaN	2–1	NaN
10	Granada	NaN	NaN	NaN	NaN	2–2	1–1	2–0	NaN	1–1	...	—	NaN	3–2	NaN	0–2	NaN	NaN	0–3	NaN	2–3
11	Las Palmas	NaN	NaN	1–1	2–1	1–2	NaN	1–1	2–1	2–0	...	1–0	—	1–1	NaN	0–1	NaN	0–0	NaN	NaN	3–0
12	Mallorca	0–0	NaN	0–0	NaN	2–2	NaN	NaN	1–1	0–0	...	NaN	NaN	—	3–2	NaN	NaN	NaN	1–0	1–1	0–1
13	Osasuna	NaN	1–0	0–2	0–2	1–2	NaN	NaN	NaN	NaN	...	2–0	1–1	NaN	—	1–0	NaN	1–1	0–0	NaN	NaN
14	Rayo Vallecano	2–0	NaN	NaN	0–7	1–1	NaN	NaN	0–0	NaN	...	NaN	NaN	2–2	NaN	—	NaN	2–2	NaN	0–1	1–1
15	Real Madrid	NaN	NaN	NaN	a	a	NaN	NaN	NaN	2–1	...	2–0	2–0	1–0	4–0	0–0	—	2–1	NaN	5–1	4–1
16	Real Sociedad	1–1	NaN	3–0	NaN	0–1	0–0	NaN	1–1	4–3	...	5–3	NaN	1–0	NaN	NaN	NaN	—	2–1	NaN	NaN
17	Sevilla	2–3	5–1	0–2	NaN	NaN	1–1	NaN	NaN	0–3	...	NaN	1–0	NaN	NaN	2–2	1–1	NaN	—	1–2	1–1
18	Valencia	NaN	NaN	NaN	3–0	1–1	NaN	2–0	0–0	NaN	...	1–0	1–0	NaN	1–2	NaN	NaN	0–1	NaN	—	3–1
19	Villarreal	1–1	2–1	2–3	NaN	3–4	1–2	NaN	3–2	NaN	...	NaN	1–2	NaN	3–1	NaN	NaN	0–3	NaN	a	—

20 rows × 21 columns

In [ ]:
```python
team_names = list(game_schedule["Casa \ Fora"])
nicknames = list(game_schedule.columns[1:])

print(team_names)
print(nicknames)

dict_teams = dict(zip(nicknames, team_names))
print(dict_teams)
```

```
['Alavés', 'Almería', 'Athletic Bilbao', 'Atlético de Madrid', 'Barcelona', 'Betis', 'Cádiz', 'Celta de Vigo', 'Getafe', 'Girona', 'Granada', 'Las Palmas', 'Mallorca', 'Osasuna', 'Rayo Vallecano', 'Real Madrid', 'Real Sociedad', 'Sevil
la', 'Valencia', 'Villarreal']
['ALA', 'ALM', 'ATH', 'ATM', 'BAR', 'BET', 'CAD', 'CEL', 'GET', 'GIR', 'GRA', 'LPA', 'MLL', 'OSA', 'RAY', 'RMA', 'RSO', 'SEV', 'VAL', 'VIL']
{'ALA': 'Alavés', 'ALM': 'Almería', 'ATH': 'Athletic Bilbao', 'ATM': 'Atlético de Madrid', 'BAR': 'Barcelona', 'BET': 'Betis', 'CAD': 'Cádiz', 'CEL': 'Celta de Vigo', 'GET': 'Getafe', 'GIR': 'Girona', 'GRA': 'Granada', 'LPA': 'Las Palm
as', 'MLL': 'Mallorca', 'OSA': 'Osasuna', 'RAY': 'Rayo Vallecano', 'RMA': 'Real Madrid', 'RSO': 'Real Sociedad', 'SEV': 'Sevilla', 'VAL': 'Valencia', 'VIL': 'Villarreal'}
```

In [ ]:
```python
game_schedule_adjusted = game_schedule.set_index('Casa \ Fora')
game_schedule_adjusted = game_schedule_adjusted.unstack().reset_index()
game_schedule_adjusted = game_schedule_adjusted.rename(columns={"level_0" : "away", "Casa \ Fora": "home", 0: "result"})

def adjust_team_nickname(line):
    nickname = line["away"]
    name = dict_teams[nickname]
    return name

game_schedule_adjusted["away"] = game_schedule_adjusted.apply(adjust_team_nickname, axis=1)
game_schedule_adjusted = game_schedule_adjusted[game_schedule_adjusted["away"] != game_schedule_adjusted["home"]]
display(game_schedule_adjusted)
```

	away	home	result
1	Alavés	Almería	NaN
2	Alavés	Athletic Bilbao	NaN
3	Alavés	Atlético de Madrid	2–1
4	Alavés	Barcelona	2–1
5	Alavés	Betis	NaN
...	...	...	...
394	Villarreal	Rayo Vallecano	1–1
395	Villarreal	Real Madrid	4–1
396	Villarreal	Real Sociedad	NaN
397	Villarreal	Sevilla	1–1
398	Villarreal	Valencia	3–1

380 rows × 3 columns

In [ ]:
```python
game_schedule_adjusted["result"] = game_schedule_adjusted["result"].fillna("to play")
completed_games_table = game_schedule_adjusted[game_schedule_adjusted["result"].str.contains("–")]
display(completed_games_table)
upcoming_games_table = game_schedule_adjusted[~game_schedule_adjusted["result"].str.contains("–")]
upcoming_games_table = upcoming_games_table.drop(columns=["result"])
display(upcoming_games_table)
```

	away	home	result
3	Alavés	Atlético de Madrid	2–1
4	Alavés	Barcelona	2–1
6	Alavés	Cádiz	1–0
7	Alavés	Celta de Vigo	1–1
8	Alavés	Getafe	1–0
...	...	...	...
392	Villarreal	Mallorca	0–1
394	Villarreal	Rayo Vallecano	1–1
395	Villarreal	Real Madrid	4–1
397	Villarreal	Sevilla	1–1
398	Villarreal	Valencia	3–1

195 rows × 3 columns

	away	home
1	Alavés	Almería
2	Alavés	Athletic Bilbao
5	Alavés	Betis
10	Alavés	Granada
11	Alavés	Las Palmas
...	...	...
384	Villarreal	Barcelona
387	Villarreal	Celta de Vigo
389	Villarreal	Girona
393	Villarreal	Osasuna
396	Villarreal	Real Sociedad

185 rows × 2 columns

In [ ]:
```python
completed_games_table[["home_goals", "away_goals"]] = completed_games_table["result"].str.split("–", expand=True)
completed_games_table = completed_games_table.drop(columns=["result"])
completed_games_table["home_goals"] = completed_games_table["home_goals"].astype(int)
completed_games_table["away_goals"] = completed_games_table["away_goals"].astype(int)
display(completed_games_table)
```

	away	home	home_goals	away_goals
3	Alavés	Atlético de Madrid	2	1
4	Alavés	Barcelona	2	1
6	Alavés	Cádiz	1	0
7	Alavés	Celta de Vigo	1	1
8	Alavés	Getafe	1	0
...	...	...	...	...
392	Villarreal	Mallorca	0	1
394	Villarreal	Rayo Vallecano	1	1
395	Villarreal	Real Madrid	4	1
397	Villarreal	Sevilla	1	1
398	Villarreal	Valencia	3	1

195 rows × 4 columns

In [ ]:
```python
home_goals_average = completed_games_table.groupby("home").mean(numeric_only=True)
home_goals_average = home_goals_average.rename(columns={"home_goals": "home_goals_scored", "away_goals": "home_goals_conceded"})
display(home_goals_average)
away_goals_average = completed_games_table.groupby("away").mean(numeric_only=True)
away_goals_average = away_goals_average.rename(columns={"home_goals": "away_goals_conceded", "away_goals": "away_goals_scored"})
display(away_goals_average)
```

home	home_goals_scored	home_goals_conceded
Alavés	1.111111	1.222222
Almería	0.900000	1.700000
Athletic Bilbao	2.454545	1.090909
Atlético de Madrid	2.300000	1.100000
Barcelona	2.100000	1.100000
Betis	1.300000	0.400000
Celta de Vigo	0.800000	1.300000
Cádiz	0.888889	1.000000
Getafe	1.000000	0.666667
Girona	2.500000	1.300000
Granada	1.444444	2.000000
Las Palmas	1.272727	0.727273
Mallorca	0.888889	0.777778
Osasuna	0.900000	1.200000
Rayo Vallecano	1.000000	1.777778
Real Madrid	2.444444	0.444444
Real Sociedad	1.800000	1.100000
Sevilla	1.363636	1.636364
Valencia	1.333333	0.555556
Villarreal	1.700000	2.100000

away	away_goals_conceded	away_goals_scored
Alavés	1.454545	0.727273
Almería	2.500000	0.900000
Athletic Bilbao	0.888889	1.222222
Atlético de Madrid	1.333333	1.777778
Barcelona	1.222222	1.666667
Betis	1.600000	0.900000
Celta de Vigo	1.700000	1.300000
Cádiz	1.777778	0.555556
Getafe	1.900000	1.500000
Girona	1.100000	2.100000
Granada	2.100000	0.700000
Las Palmas	1.000000	0.600000
Mallorca	1.600000	0.900000
Osasuna	1.888889	1.444444
Rayo Vallecano	0.777778	1.000000
Real Madrid	0.700000	1.800000
Real Sociedad	1.000000	1.300000
Sevilla	1.333333	1.111111
Valencia	1.800000	1.000000
Villarreal	1.909091	0.909091

In [ ]:
```python
statistics_table = home_goals_average.merge(away_goals_average, left_index=True, right_index=True)
statistics_table = statistics_table.reset_index()
statistics_table = statistics_table.rename(columns={"home": "team"})
display(statistics_table)
```

	team	home_goals_scored	home_goals_conceded	away_goals_conceded	away_goals_scored
0	Alavés	1.111111	1.222222	1.454545	0.727273
1	Almería	0.900000	1.700000	2.500000	0.900000
2	Athletic Bilbao	2.454545	1.090909	0.888889	1.222222
3	Atlético de Madrid	2.300000	1.100000	1.333333	1.777778
4	Barcelona	2.100000	1.100000	1.222222	1.666667
5	Betis	1.300000	0.400000	1.600000	0.900000
6	Celta de Vigo	0.800000	1.300000	1.700000	1.300000
7	Cádiz	0.888889	1.000000	1.777778	0.555556
8	Getafe	1.000000	0.666667	1.900000	1.500000
9	Girona	2.500000	1.300000	1.100000	2.100000
10	Granada	1.444444	2.000000	2.100000	0.700000
11	Las Palmas	1.272727	0.727273	1.000000	0.600000
12	Mallorca	0.888889	0.777778	1.600000	0.900000
13	Osasuna	0.900000	1.200000	1.888889	1.444444
14	Rayo Vallecano	1.000000	1.777778	0.777778	1.000000
15	Real Madrid	2.444444	0.444444	0.700000	1.800000
16	Real Sociedad	1.800000	1.100000	1.000000	1.300000
17	Sevilla	1.363636	1.636364	1.333333	1.111111
18	Valencia	1.333333	0.555556	1.800000	1.000000
19	Villarreal	1.700000	2.100000	1.909091	0.909091

In [ ]:
```python
from scipy.stats import poisson

def calculate_expected_score(line):
    home_team = line["home"]
    away_team = line["away"]

    home_lambda = (statistics_table.loc[statistics_table["team"] == home_team,"home_goals_scored"].iloc[0]
                  * statistics_table.loc[statistics_table["team"] == away_team, "away_goals_conceded"].iloc[0])

    away_lambda = (statistics_table.loc[statistics_table["team"] == away_team,"away_goals_scored"].iloc[0]
                  * statistics_table.loc[statistics_table["team"] == home_team, "home_goals_conceded"].iloc[0])

    pv_home = 0
    p_draw = 0
    pv_away = 0

    for home_goals in range(8):
        for away_goals in range(8):
            result_probability = poisson.pmf(home_goals, home_lambda) * poisson.pmf(away_goals, away_lambda)
            if home_goals == away_goals:
                p_draw += result_probability
            elif home_goals > away_goals:
                pv_home += result_probability
            elif home_goals < away_goals:
                pv_away += result_probability

    home_ev = pv_home * 3 + p_draw
    away_ev = pv_away * 3 + p_draw
    line["home_pts"] = home_ev
    line["away_pts"] = away_ev
    return line
```

In [ ]:
```python
upcoming_games_table = upcoming_games_table.apply(calculate_expected_score, axis=1)
display(upcoming_games_table)
```

	away	home	home_pts	away_pts
1	Alavés	Almería	1.417880	1.314788
2	Alavés	Athletic Bilbao	2.608721	0.218673
5	Alavés	Betis	2.482721	0.330989
10	Alavés	Granada	1.779706	1.005303
11	Alavés	Las Palmas	2.287667	0.502426
...	...	...	...	...
384	Villarreal	Barcelona	2.552657	0.216525
387	Villarreal	Celta de Vigo	1.610093	1.135991
389	Villarreal	Girona	2.443706	0.167442
393	Villarreal	Osasuna	1.798867	0.961277
396	Villarreal	Real Sociedad	2.509510	0.313253

185 rows × 4 columns

In [ ]:
```python
def adjust_team_name(line):
    for name in team_names:
        if name in line["Equipe"]:
            return name

league_table["team"] = league_table.apply(adjust_team_name, axis=1)
updated_league_table = league_table[["team", "Pts"]]
updated_league_table["Pts"] = updated_league_table["Pts"].astype(int)



display(updated_league_table)
```

	team	Pts
0	Girona	49
1	Real Madrid	48
2	Athletic Bilbao	41
3	Barcelona	41
4	Atlético de Madrid	38
5	Real Sociedad	32
6	Betis	31
7	Las Palmas	28
8	Getafe	26
9	Valencia	26
10	Rayo Vallecano	23
11	Osasuna	22
12	Alavés	20
13	Mallorca	19
14	Villarreal	19
15	Celta de Vigo	17
16	Sevilla	16
17	Cádiz	15
18	Granada	11
19	Almería	6

In [ ]:
```python
home_points_table = upcoming_games_table.groupby("home").sum(numeric_only=True)[["home_pts"]]
away_points_table = upcoming_games_table.groupby("away").sum(numeric_only=True)[["away_pts"]]
display(home_points_table)
display(away_points_table)
```

home	home_pts
Alavés	13.956284
Almería	9.216617
Athletic Bilbao	18.435976
Atlético de Madrid	19.253009
Barcelona	20.634058
Betis	20.222279
Celta de Vigo	9.766241
Cádiz	15.038689
Getafe	16.566093
Girona	19.919270
Granada	12.316965
Las Palmas	16.054298
Mallorca	15.842697
Osasuna	13.588612
Rayo Vallecano	12.285013
Real Madrid	25.655055
Real Sociedad	18.662283
Sevilla	12.028726
Valencia	20.895457
Villarreal	13.107468

away	away_pts
Alavés	4.884169
Almería	3.541928
Athletic Bilbao	14.506957
Atlético de Madrid	16.127048
Barcelona	13.670847
Betis	5.830258
Celta de Vigo	8.257772
Cádiz	4.953789
Getafe	10.178987
Girona	13.527192
Granada	3.630825
Las Palmas	8.242579
Mallorca	6.574153
Osasuna	8.906614
Rayo Vallecano	12.556053
Real Madrid	18.559057
Real Sociedad	12.935574
Sevilla	8.921009
Valencia	6.726365
Villarreal	4.983156

In [ ]:
```python
def update_prediction_score(line):
    team = line["team"]
    score = int(line["Pts"]) + float(home_points_table.loc[team, "home_pts"]) + float(away_points_table.loc[team, "away_pts"])
    return score


updated_league_table["Pts"] = updated_league_table.apply(update_prediction_score, axis=1)
updated_league_table = updated_league_table.sort_values(by="Pts", ascending=False)
updated_league_table = updated_league_table.reset_index(drop=True)
updated_league_table.index = updated_league_table.index + 1
display(updated_league_table)
```

	team	Pts
1	Real Madrid	92.214111
2	Girona	82.446461
3	Barcelona	75.304905
4	Athletic Bilbao	73.942933
5	Atlético de Madrid	73.380057
6	Real Sociedad	63.597856
7	Betis	57.052537
8	Valencia	53.621822
9	Getafe	52.745080
10	Las Palmas	52.296877
11	Rayo Vallecano	47.841066
12	Osasuna	44.495226
13	Mallorca	41.416850
14	Alavés	38.840453
15	Villarreal	37.090624
16	Sevilla	36.949735
17	Celta de Vigo	35.024013
18	Cádiz	34.992478
19	Granada	26.947791
20	Almería	18.758545