

```
In [13]: # Agafa un conjunt de dades de tema esportiu que t'agradi i selecciona un atribut del conjunt de dades. Calcula la moda,
# la mediana, la desviació estàndard i la mitjana aritmètica.
# Continuant amb les dades de tema esportiu, selecciona dos atributs i calcula'n la seva correlació.
```

```
In [14]: import pandas as pd
df = pd.read_csv("england-premier-league-matches-2018-to-2019-stats.csv")
df
```

Out[14]:

	timestamp	date_GMT	status	attendance	home_team_name	away_team_name	referee	Game Week	Pre-Match PPG (Home)	Pre-Match PPG (Away)	...	odds_ft_home_team_win	odds_ft_draw	odds_ft_away_team_win	odds_ft
0	1533927600	Aug 10 2018 - 7:00pm	complete	74439	Manchester United	Leicester City	Andre Marriner	1	0.00	0.00	...	1.37	4.98		9.81
1	1533987000	Aug 11 2018 - 11:30am	complete	51749	Newcastle United	Tottenham Hotspur	Martin Atkinson	1	0.00	0.00	...	4.51	3.77		1.84
2	1533996000	Aug 11 2018 - 2:00pm	complete	10353	AFC Bournemouth	Cardiff City	Kevin Friend	1	0.00	0.00	...	2.03	3.51		3.96
3	1533996000	Aug 11 2018 - 2:00pm	complete	24821	Fulham	Crystal Palace	Mike Dean	1	0.00	0.00	...	2.31	3.44		3.26
4	1533996000	Aug 11 2018 - 2:00pm	complete	24121	Huddersfield Town	Chelsea	Chris Kavanagh	1	0.00	0.00	...	7.47	4.27		1.51
...
375	1557669600	May 12 2019 - 2:00pm	complete	53331	Liverpool	Wolverhampton Wanderers	Martin Atkinson	38	2.78	1.28	...	1.32	5.80		9.00
376	1557669600	May 12 2019 - 2:00pm	complete	74457	Manchester United	Cardiff City	Jonathan Moss	38	2.00	0.61	...	1.28	6.20		10.00
377	1557669600	May 12 2019 - 2:00pm	complete	30367	Southampton	Huddersfield Town	Lee Probert	38	1.22	0.33	...	1.41	4.90		7.50
378	1557669600	May 12 2019 - 2:00pm	complete	60124	Tottenham Hotspur	Everton	Andre Marriner	38	2.06	1.06	...	2.10	3.50		3.55
379	1557669600	May 12 2019 - 2:00pm	complete	20067	Watford	West Ham United	Chris Kavanagh	38	1.50	1.00	...	2.20	3.70		3.10

380 rows x 64 columns

```
In [15]: df2= df['attendance']
df2
```

```
Out[15]: 0      74439
1      51749
2     10353
3     24821
4     24121
...
375    53331
376    74457
377    30367
378    60124
379     20067
Name: attendance, Length: 380, dtype: int64
```

```
In [16]: import statistics
mediana = statistics.median(df2)
print("La mitjana d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de " + str(mediana))
```

La mitjana d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de 31948.0

```
In [17]: moda = statistics.mode(df2)
print("La moda d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de " + str(moda))
```

La moda d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de 40491

```
In [18]: desviacio= statistics.pstdev(df2)
print("La desviació estàndard d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de " + str(desviacio))
```

La desviació estàndard d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de 16698.319833314097

```
In [49]: import math
mitjana= statistics.mean(df2)

x = (math.ceil(mitjana))

res="La mitjana d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de " + str(x)

print(res)
```

La mitjana d'assistents als estadis de la premier league de la temporada 2018 - 2019 és de 38187

```
In [20]: # Ex3 Nivell 2
import numpy as np

df3 = df['odds_ft_over45']

correlacio = np.corrcoef (df2, df3)
correlacio
```

```
Out[20]: array([[ 1.          , -0.35331693],
       [-0.35331693,  1.          ]])
```

```
In [23]: df4 = df[(df['home_team_name'] == 'Manchester United') | (df['home_team_name'] == 'Manchester City')]
df4

df5= df[['timestamp', 'attendance', 'Game Week', 'Pre-Match PPG (Home)', 'Pre-Match PPG (Away)',
'home_ppg', 'away_ppg', 'home_team_goal_count' ]]
y= df5.sample(20)

y
```

Out[23]:

	timestamp	attendance	Game Week	Pre-Match PPG (Home)	Pre-Match PPG (Away)	home_ppg	away_ppg	home_team_goal_count
245	1549119600	39380	25	1.75	1.36	1.79	1.21	1
31	1535810400	30526	4	3.00	0.00	1.21	0.26	2
108	1541347200	40407	11	2.20	1.20	2.21	1.53	3
213	1547305200	30725	22	1.18	0.50	1.05	0.37	0
26	1535286600	20014	3	3.00	3.00	1.42	1.53	2
20	1535196600	31222	3	1.00	3.00	1.79	2.32	1
334	1555354800	20480	34	1.73	1.27	1.42	1.32	0
60	1538220600	56938	7	0.33	2.00	1.63	1.58	3
104	1541257200	56862	11	0.80	0.80	1.63	0.89	4
200	1546345800	39052	21	1.80	1.40	1.79	1.32	0
314	1554230700	17301	33	1.64	0.13	1.42	0.26	4
266	1550943000	31778	27	1.17	1.08	1.42	1.53	1
359	1556465400	74526	36	2.06	1.65	1.89	1.58	1
338	1555768800	23957	35	0.47	1.18	0.47	1.21	1
75	1538834400	20139	8	2.25	1.00	1.42	0.84	0
212	1547305200	19316	22	1.00	0.18	1.21	0.26	2
40	1537011000	80188	5	3.00	3.00	2.00	2.32	1
190	1546095600	30597	20	1.67	1.00	1.21	1.05	1
149	1544040000	33012	15	1.80	0.57	2.00	0.84	3
161	1544886000	24738	17	0.71	1.38	1.05	1.32	1

```
In [22]: z =y.corr(method="pearson")
z
```

Out[22]:

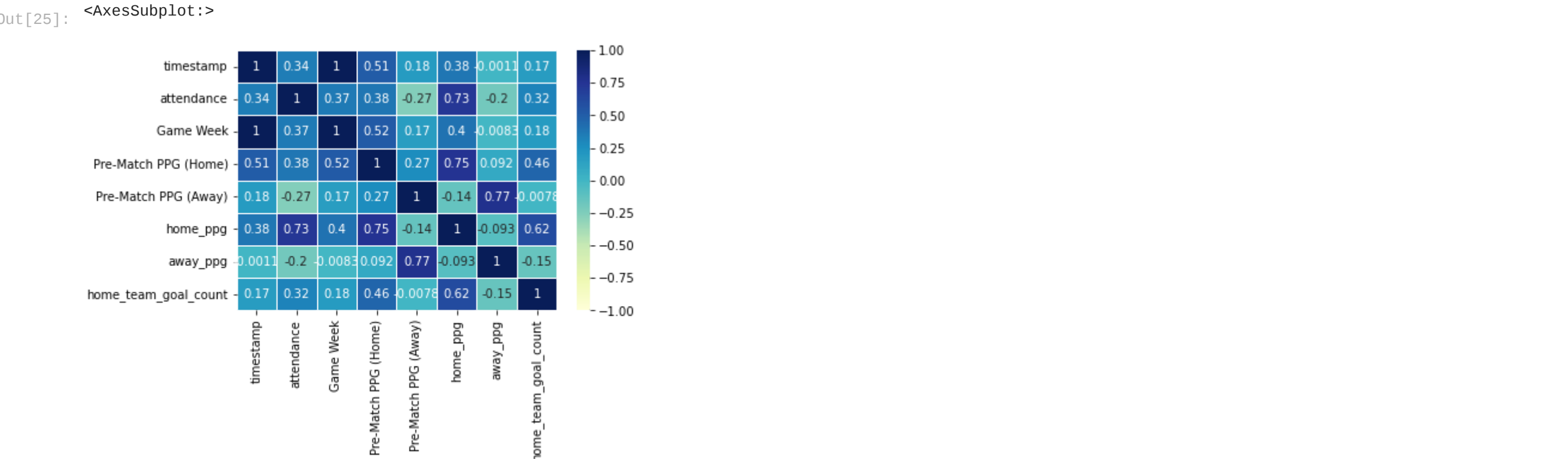
	timestamp	attendance	Game Week	Pre-Match PPG (Home)	Pre-Match PPG (Away)	home_ppg	away_ppg	home_team_goal_count
timestamp	1.000000	0.344056	0.997141	0.509205	0.180729	0.375974	-0.001080	0.174875
attendance	0.344056	1.000000	0.366698	0.381896	-0.265045	0.731073	-0.196131	0.317092
Game Week	0.997141	0.366698	1.000000	0.518226	0.169370	0.402555	-0.008302	0.183949
Pre-Match PPG (Home)	0.509205	0.381896	0.518226	1.000000	0.268546	0.750445	0.091661	0.461732
Pre-Match PPG (Away)	0.180729	-0.265045	0.169370	0.268546	1.000000	-0.142784	0.774515	-0.007811
home_ppg	0.375974	0.731073	0.402555	0.750445	-0.142784	1.000000	-0.093436	0.615492
away_ppg	-0.001080	-0.196131	-0.008302	0.091661	0.774515	-0.093436	1.000000	-0.150703
home_team_goal_count	0.174875	0.317092	0.183949	0.461732	-0.007811	0.615492	-0.150703	1.000000

```
In [24]: import matplotlib.pyplot as plt
plt.matshow(y.corr())
```



```
In [25]: import seaborn as sns

sns.heatmap(z, annot= True, vmin= -1, vmax= 1, linewidths=.5, cmap="YlGnBu")
```



```
In [47]: # Nivell 3 Ex 4
#Continuant amb les dades de tema esportiu, selecciona un atribut i calcula la mitjana geomètrica i la mitjana harmònica.
import statistics
mitjana_aritmetica = round(statistics.geometric_mean(y['attendance']),1)
print("La mitjana geomètrica dels assistens a camp de futbol de la premier league durant la temporada 2017 - 2018 és de " + str(mitjana_aritmetica))
```

La mitjana geomètrica dels assistens a camp de futbol de la premier league durant la temporada 2017 - 2018 és de 32608.5

```
In [51]: mitjana_harmonica = round(statistics.harmonic_mean(y['attendance']),1)
print("La mitjana harmònica dels assistens a camp de futbol de la premier league durant la temporada 2017 - 2018 és de " + str(mitjana_harmonica))
```

La mitjana harmònica dels assistens a camp de futbol de la premier league durant la temporada 2017 - 2018 és de 29906.6