🧮 Is it statistically easier to win the Premier League vs La Liga?

👹 numbers can sometimes be deceiving

📋 Table of Contents

1. Introduction

2. Body

3. Conclusion

📖 INTRODUCTION

The debate over which football league is more challenging, the Premier League or La Liga, has long intrigued fans and potential investors alike. In this article, we aim to settle the discussion using statistical analysis. By examining the average points earned by the league winners from 2009 to 2022, we'll compare the difficulty of clinching the top spot in each league. However, we won't solely rely on numbers; we'll employ a T-Test to ensure the statistical significance of our findings.

The dataset used for this analysis is the "Performance Data on Football Teams 09 to 22." This dataset encompasses a comprehensive range of performance statistics for football teams participating in the Premier League and La Liga during the period from 2009 to 2022. It includes data on various aspects of team performance, such as points earned, goals scored, goal differentials, and other relevant metrics.

🦴 BODY

🛠️ Import important libraries, Read the data

In [17]:
```python
import pandas as pd
import pingouin as pg
```

In [18]:
```python
df = pd.read_csv("football.csv")
```

🍳 Preprocessing

In [19]:
```
df.head()
```

Out[19]:

	Key	Team	League	Season	Rank	Games	Wins	Draws	Losses	Points	...
0	Chelsea DL 2009/2010	Chelsea	Premier League	2009/2010	1.0	38.0	27.0	5.0	6.0	86.0	...
1	Manchester United DL 2009/2010	Manchester United	Premier League	2009/2010	2.0	38.0	27.0	4.0	7.0	85.0	...
2	Tottenham DL 2009/2010	Tottenham	Premier League	2009/2010	4.0	38.0	21.0	7.0	10.0	70.0	...
3	Arsenal DL 2009/2010	Arsenal	Premier League	2009/2010	3.0	38.0	23.0	6.0	9.0	75.0	...
4	Aston Villa DL 2009/2010	Aston Villa	Premier League	2009/2010	6.0	38.0	17.0	13.0	8.0	64.0	...

5 rows × 93 columns

In [20]:
```
df.tail()
```

Out[20]:

	Key	Team	League	Season	Rank	Games	Wins	Draws	Losses	Points
1685	Shakhtar Donetsk CL 2021/2022	Shakhtar Donetsk	Champions League	2021/2022	32.0	NaN	NaN	NaN	NaN	NaN
1686	FC Porto CL 2021/2022	FC Porto	Champions League	2021/2022	32.0	NaN	NaN	NaN	NaN	NaN
1687	Dynamo Kyiv CL 2021/2022	Dynamo Kyiv	Champions League	2021/2022	32.0	NaN	NaN	NaN	NaN	NaN
1688	Besiktas CL 2021/2022	Besiktas	Champions League	2021/2022	32.0	NaN	NaN	NaN	NaN	NaN
1689	Malmoe FF CL 2021/2022	Malmoe FF	Champions League	2021/2022	32.0	NaN	NaN	NaN	NaN	NaN

5 rows × 93 columns

In [21]:
```
df.describe()
```

Out[21]:

	Rank	Games	Wins	Draws	Losses	Points	GoalsF
count	1688.000000	1272.000000	1272.000000	1272.000000	1272.000000	1272.000000	1272.0000
mean	13.212085	37.105346	13.860849	9.390723	13.853774	50.938679	50.6808
std	11.652804	1.936824	5.839471	2.944652	5.338193	16.656978	16.6989
min	1.000000	27.000000	2.000000	2.000000	0.000000	13.000000	20.0000
25%	6.000000	38.000000	10.000000	7.000000	10.000000	40.000000	39.0000
50%	12.000000	38.000000	12.000000	9.000000	14.000000	47.000000	47.0000
75%	17.000000	38.000000	17.000000	11.000000	18.000000	61.000000	59.0000
max	323.000000	38.000000	33.000000	19.000000	29.000000	102.000000	121.0000

8 rows × 89 columns

In [24]:
```python
# Add start and end of the Season
df["Start Season"] = df["Season"].apply(lambda x : int(x[:4]))
df["End Season"] = df["Season"].apply(lambda x : (x[5:]))
```

H₀ Null hypothesis: Winning the Premier League is more challenging than winning La Liga

In [25]:
```python
# Select all Winners Premier League from 2009-2022
pl = df[(df["League"] == "Premier League") & (df["Rank"] == 1)][["League","
```

In [26]:
```python
# Select all Winners La Liga from 2009-2022
ll = df[(df["League"] == "La Liga") & (df["Rank"] == 1)][["League","Team","
```

To determine the level of difficulty in winning each league, we start by considering the main metric—points. The team with the most points is declared the winner. Calculating the average points earned by the winners of La Liga and the Premier League during the specified period, we obtain the following results:

In [27]:
```python
print(f'''La Liga Winners Average Points (2009-2022): {round(ll["Points"].m
print(f'''Premier League Winners Average Points (2009-2022): {round(pl["Poi
```

```
La Liga Winners Average Points (2009-2022): 92.46
Premier League Winners Average Points (2009-2022): 89.77
```

At first glance, it seems that La Liga winners, on average, accumulate a higher number of points compared to Premier League winners. This observation may lead to the assumption that winning La Liga is more challenging. However, numbers can sometimes be deceiving, necessitating a deeper analysis.

⬆Back to Table of Contents⬆

🔔 Check normality

To make a fair and scientifically supported comparison, we employ the T-Test methodology. Before conducting the T-Test, we must verify the normality of the data for both leagues.

In [28]:

```python
print(f'''Premier League Winner Points (2009-2022) normality -> {pg.normali
print(f'''La Liga Winner Points (2009-2022) normality -> {pg.normality(ll.P
```

```
Premier League Winner Points (2009-2022) normality -> True
La Liga Winner Points (2009-2022) normality -> True
```

⬆️**Back to Table of Contents** ⬆️

🧪 T-Test

Having confirmed the normal distribution of both datasets, we proceed to conduct the T-Test. Our null hypothesis states that winning the Premier League is more challenging than winning La Liga.

In [29]:

```python
pg.ttest(ll.Points.values, pl.Points.values, correction = False, alternativ
```

Out[29]:

	T	dof	alternative	p-val	CI95%	cohen-d	BF10	power
T-test	1.170357	24	greater	0.126674	[-1.24, inf]	0.459052	1.198	0.305916

⬆️**Back to Table of Contents** ⬆️

⚽ Result

The p-value obtained from the T-Test is greater than the conventional significance level of 5%. Therefore, we fail to reject the null hypothesis, suggesting that winning the Premier League is statistically harder than winning La Liga.

⬆️**Back to Table of Contents** ⬆️

📕 CONCLUSION

This statistical and scientific analysis settles the debate regarding the relative difficulty of winning the Premier League versus La Liga. Despite the initial perception that La Liga winners accrue more average points, the T-Test results demonstrate that clinching the top spot in the Premier League is statistically more challenging. Whether you're a serious investor or an avid fan seeking an answer to this long-standing debate, these findings provide valuable insights for making informed decisions.