

Data Mining

Youcef Messaoud (Git : ekehe)

Master Machine and Data Mining



Figure 1: The three best teams in the French championship this year

1 Introduction

My study is about the French championship. From the champion of the previous year and its characteristics, how the three best teams this year behave. The three best teams this year are Paris-Saint-Germain, Monaco and Nice. Last Year (2016), the champion was Paris-Saint-Germain. I would like to know who have a pace of champion. I will try to know thanks to three studies :

- Home results
- Away results
- number of points win per match

2 Dataset

First, I found an interesting dataset on internet. This dataset contains some informations : the matches between teams, the result, number of goals, number

of yellow/red cards per games, number of shoot per teams..... At the end of the csv file i added the last year results of the games of Paris-Saint-Germain for my study.

3 Home Results

I sorted my dataset and I obtain :

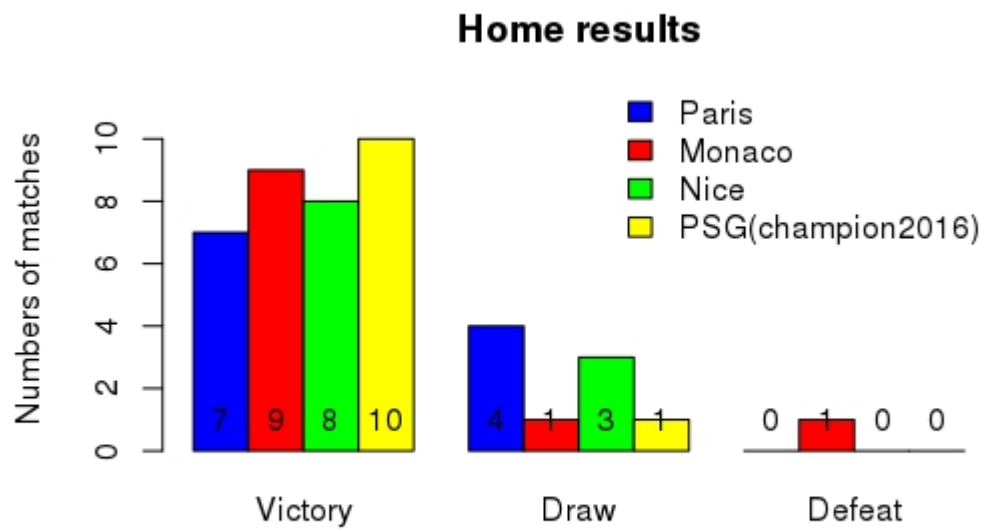


Figure 2: Home results on eleven games

Interpretations : as we can see :

- Paris have seven victories, four draws and zero defeat
- Monaco have nine victories, one draw and one defeat
- Nice have eight victories, one draws and zero defeat
- the 2016 Champion have ten victories, three draws and zero defeat

In Football, we know that when you win a match you win 3 points, if you draw you win 1 point and 0 point if you loose so :

- Paris won 25 points

- Monaco won 28 points
- Nice won 27 points
- Last year, at the same time the Champion won 31 points.

The results are close together. Only 3 points separate the first from the third. The number of point for Monaco is very close to the last Champion. Monaco win the first round.

4 Away Results

It's very difficult to win an away game because the team does not have his fans, the stadium is différent,travel fatigue..... Some experts says " the number of away victories can predict the Champions". So the results are :

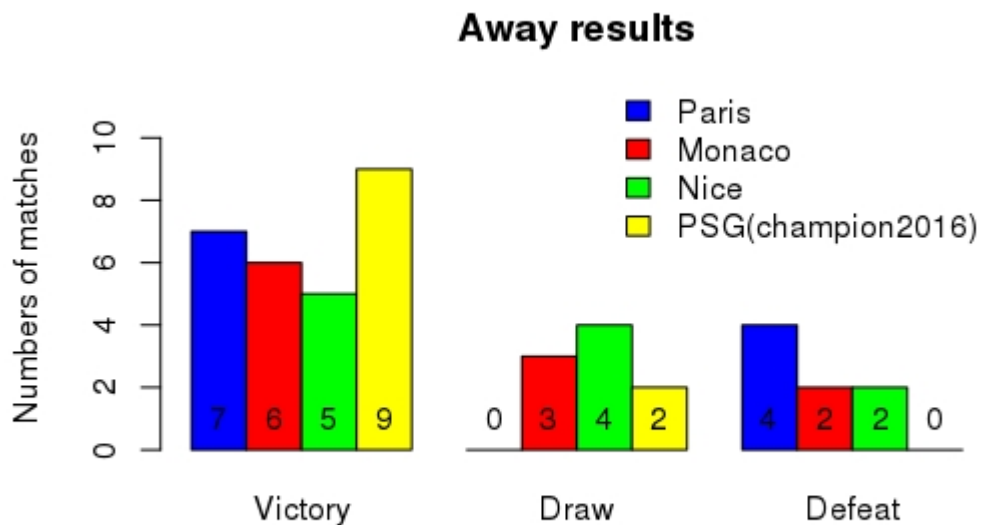


Figure 3: Away results on eleven games

Interpretations : we can see :

- Paris have seven victories, zero draw and four defeats
- Monaco have six victories, three draws and two defeats
- Nice have five victories, four draws and two defeats

- the 2016 Champion have eight victories, three draws and zero defeat

finally :

- Paris won 21 points
- Monaco won 21 points
- Nice won 19 points
- Last year, at the same time the Champion won 27 points.

The number of points for Monaco and Paris is far to the last Champion.
Nice lose this round and the battle.

5 Number of point win per match

In this part i will talk about the number point per match :

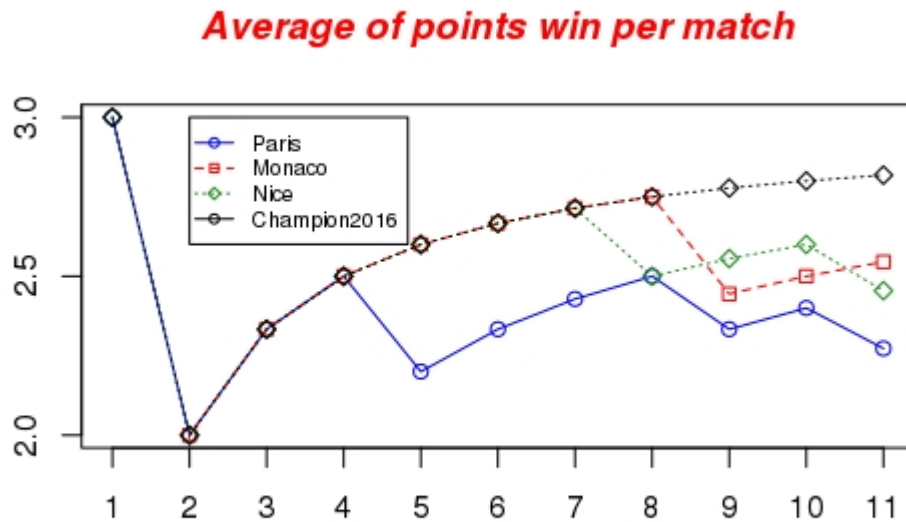


Figure 4: average of points win per match

The dataset is :

<i>Paris</i>	3 2.4	2 2.272727	2.333333	2.5	2.2	2.333333	2.428571	2.50	2.333333
<i>Monaco</i>	1 2.5	2 2.545455	2.333333	2.5	2.6	2.666667	2.714286	2.75	2.444444
<i>Nice</i>	3 2.6	2 2.454545	2.333333	2.5	2.6	2.666667	2.714286	2.50	2.555556
<i>Champion2016</i>	3 2.8	2 2.818182	2.333333	2.5	2.6	2.666667	2.714286	2.75	2.777778

Paris is not constant. from de second game to the eighth game, Monaco has the same result as the last Champion. Monaco and Nice are close together, but they are very far from the last Champion. Monaco win this round.

6 Conclusion



Figure 5: The Winner is Monaco

Thanks to my dataset, Monaco has the same behavior as the last Champion. So I think Monaco will become the Champion.

7 Avantages

I liked this project. Thanks to this project I improved in R.