Research on Scoring Efficiency of NBA Players in Different Ways-Taking Stephen Curry as an Example

Hanbo Yang[*]

Beijing 101 high school, 11 Yiheyuan Road, Haidian District, Beijing, China, 100091.
Corresponding author: 403935922@qq.com

ABSTRACT

The National Basketball Association, known as NBA, has been loved by the public since 1949. With the popularity of basketball games, the NBA appears in the vision of more and more people. Besides that, those who play basketball also pay more attention to shooting and scoring. Scoring can be said to be the most important part in the basketball game. This study uses the method of mathematical statistics to pour all kinds of data of players' scores, such as the shooting hit rate and the number of shots in each area into the model for comparison, therefore, as to find out the most reasonable way and position of shots for NBA players, research should be focus on how to maximize the scoring efficiency. Because there are too many NBA players, in order to better perform this study, it should focus on a specific player. Therefore, this study mainly collects various data of Stephen Curry, calculates and compares the relevant values, and finally finds the most effective scoring method for Stephen Curry. Through this research, the result is that when Stephen Curry is shooting in the three-point area at the left 45-degree angle under tight defense, it is the most efficient scoring method.

Keywords: NBA Basketball, Stephen Curry, Mathematical statistics, shot efficiency, Shot percentage

1. INTRODUCTION

Stephen Curry's shot percentage changed over time. He started in January of 2022 with a career-low in field goal percentage, then returned to average in March. Therefore, the specific theme of this paper is the analysis of Stephen Curry's shooting efficiency and the comparison of his scoring efficiency in different positions and under different circumstances. One brilliant way to find the result is by dividing the research into several small specific research problems and solving them one by one. The first is where Stephen Curry has the highest scoring efficiency on the court, that is, the place with the highest expectations for shooting. The second question is under what kind of situation and defense Stephen Curry can get the highest shooting rate in the game. Combining these two questions can get the final conclusion. More importantly, it can find the most efficient scoring method for Stephen Curry himself and help the team win to a greater extent.

2. NBA PLAYERS' SCORING EFFICIENCY IN DIFFERENT WAYS

This paper mainly focuses on Stephen Curry's shooting efficiency. In order to compare and contrast his data, this study will use the Mathematical Statistics Method.

Since this research is based on Stephen Curry's shot percentage, the first thing to do is to collect his game data. According to the official website of the NBA, considering that the data would no longer be valid, this paper chooses the last four years of regular game shooting data for Stephen Curry for my calculations. Therefore, this study collects Stephen Curry's shooting data over the past four years [4][1]. Besides that, in order to find the best way for him to score most efficiently, the basketball court can be divided into 14 different zones, as shown in Figure 1:
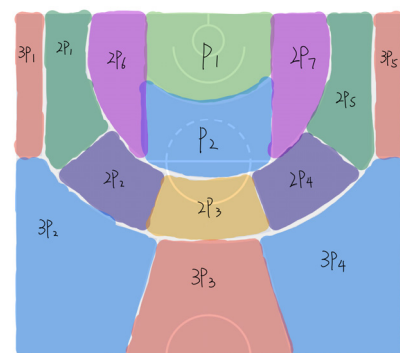

Figure 1. The shooting zones

From the graph above, giving each zone a specific name, for the zones outside the 3-point line, naming them the 3Pa, which indicates each shot made in those zones is worth 3 points; for the zones in the paint, naming them the Pa, which means the zones in the paint; naming the parts inside the 3-point line but outside the paint the 2Pa, which indicates each shot made in those zones is worth 2 points.

This study collected the data of the total number of shots he attempted in each area, and the data of the number of shots he made in each area in the past four years. This detailed data for each region is not easy to collect. However, by obtaining this data through the combination of Tencent sports and NBA official statistics data [4][1], the result is shown in the figure:
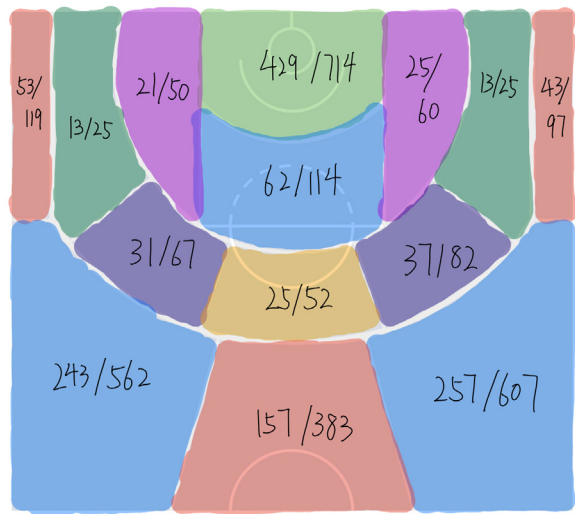

Figure 2. The attempts and made in different zones.

In Figure 2, there is a huge gap in the number of shots attempted in each zone. For example, in 2P1, he only had 25 attempts. However, in 3P4, it had 607 attempts. Therefore, the data have different sample sizes for each zone, in order to solve this problem, using a 95% confidence interval can help to predict expected score value for the different zones. In this way, the result is more convincing, and therefore it can support my research more sufficiently.

This is the part where mathematics formulation takes place:

$$P = \text{Hit percentage shots from different positions} \ldots (1)$$

$$Ps = \text{The percentage of shots shot in different situations} \ldots (2)$$

$$Pc = \text{95\% confidence interval of each } P \ldots (3)$$

$$n = \text{sample size} \ldots (4)$$

$$z^* = z_{\frac{A}{2}} \ldots (5)$$

$$A = 1 - x\% (\text{the confidece interval}) \ldots (6)$$

$$Pc = p\char`^ \pm z^* \times \sqrt{\frac{p\char`^(1 - p\char`^)}{n}} \ldots (7)$$

$$E(x) = pc \times value \ldots (8)$$

$$aveE(x) = \frac{\sum(x1 + x2 + \cdots + xn)}{n} \ldots (9)$$

In this case, each shot attempt has a huge gap. Thus, a 95% confidence interval can help a lot. A 95% confidence interval means that if the test takes many samples of the same size from this population, about 95% of them will result in an interval that captures the actual parameter value. Since the test is working on a 95% confidence interval, therefore: $A = 1 - 95\% = 5\% = 0.05$. By using the A value, the value of z^* should be: $z^* = z_{\frac{A}{2}} = z_{0.025} = 1.96$　　Thus:　　$Pc = p\char`^ \pm 1.96 \times \sqrt{\frac{p\char`^(1-p\char`^)}{n}}$

This research created a python program to calculate the 95% confidence interval of these data:

a=int(input("The sample size:"))

c=float(input("The shot made:"))

b=c/a

q=b+1.96*((b*(1-b)/a)**0.5)

w=b-1.96*((b*(1-b)/a)**0.5),

print("We are 95% confident that the true population proportion lies between",w,"and",q)

The table below shows the python programs' result:

Table 1. 95% confidence interval in different zones.

Zones	95% confidence interval
3P1	0.356 to 0.535 (average:0.445)
3P2	0.391 to 0.473 (average:0.526)
3P3	0.360 to 0.459 (average:0.410)
3P4	0.384 to 0.463 (average:0.423)
3P5	0.344 to 0.542 (average:0.443)
2P1	0.211 to 0.531 (average:0.520)
2P2	0.343 to 0.582 (average:0.463)
2P3	0.345 to 0.617 (average:0.480)
2P4	0.344 to 0.559 (average:0.451)
2P5	0.324 to 0.716 (average:0.520)
2P6	0.283 to 0.557 (average:0.420)
2P7	0.292 to 0.541 (average:0.417)
P1	0.564 to 0.637 (average:0.601)
P2	0.452 to 0.635 (average:0.544)

After getting the 95% confidence interval of the shot percentage of each zone, the next step is bringing the upper and lower limits of the interval into Python respectively, and calculate their corresponding expected values (E(x) = Pc × Value):

a=float(input("The lower shooting percentage:"))

m=float(input("The higher shooting percentage:"))

b=int(input("The score at this point:"))

c=a*b

f=m*b

print(c,"and",f)

The list below shows the result:

Table 2 The expected value in different zones.

Zones	Expected value
3P1	1.068 and 1.605 (average: 1.335)
3P2	1.173 and 1.419 (average: 1.578)
3P3	1.080 and 1.377 (average: 1.230)
3P4	1.152 and 1.389 (average: 1.269)
3P5	1.032 and 1.626 (average: 1.329)
2P1	0.422 and 1.062 (average: 1.040)
2P2	0.686 and 1.164 (average: 0.926)
2P3	0.690 and 1.234 (average: 0.960)
2P4	0.688 and 1.118 (average: 0.902)
2P5	0.648 and 1.432 (average: 1.040)
2P6	0.566 and 1.114 (average: 0.840)
2P7	0.584 and 1.082 (average: 0.834)
P1	1.128 and 1.274 (average: 1.202)
P2	0.904 and 1.270 (average: 1.088)

Make the corresponding bar chart after obtaining the expected values. Then, by comparing the data of these statistical charts, the conclusion about the shooting point with the highest expectation can be found:

import matplotlib.pyplot as plt

import numpy as np

m = ('3P1','3P2','3P3','3P4','3P5','2P1','2P2','2P3','2P4','2P5', '2P6','2P7','P1','P2')

Lower = [1.068,1.173,1.08,1.152,1.032,0.422,0.686,0.69,0.688,0.648,0.566,0.584,1.128,0.904]

Higher = [1.605,1.419,1.377,1.389,1.626,1.062,1.164,1.234,1.118,1.432,1.114,1.082,1.274,1.27]

bar_width = 0.3

index_Lower = np.arange(len(m))

index_Higher = index_Lower + bar_width

plt.bar(index_Lower, height=Lower, width=bar_width, color='b', label='Lower')

plt.bar(index_Higher, height=Higher, width=bar_width, color='y', label='Higher')

plt.legend()

plt.xticks(index_Lower + bar_width/2,m)

plt.ylabel('Expected value')

plt.xlabel('Type of shot')

plt.title('The expected value of both lower and higher percentage shot scoring')
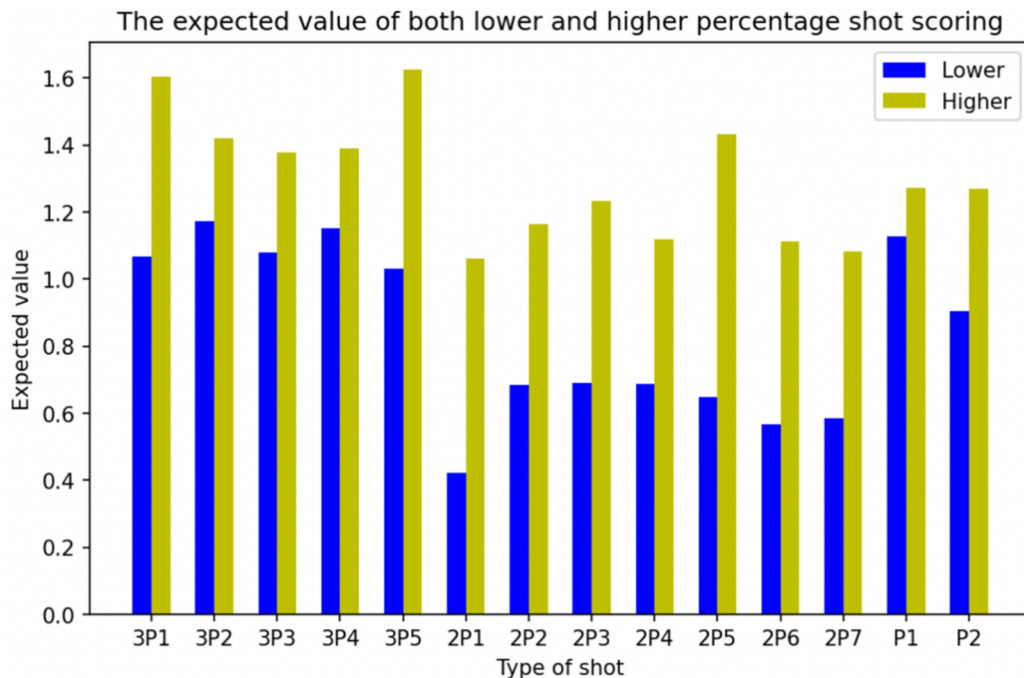
plt.show()

Figure 3. The expected value of both lower and higher percentage shot scoring

Here is the program for creating a bar graph of the average expected value of lower and higher percentage shot scoring for different zones:

```
import matplotlib.pyplot as plt

a = ('3P1','3P2','3P3','3P4','3P5','2P1','2P2','2P3','2P4','2P5',
     '2P6','2P7','P1','P2')

b = [1.335,1.578,1.23,1.269,1.329,1.04,0.926,0.96,0.902,1.04,0.84,0.834,1.202,1.088]

plt.bar(a,b,width=0.8)

plt.title('The expected value of average shot scoring')

plt.xlabel('Type of shoot')

plt.ylabel('Expected value')

plt.show()
```
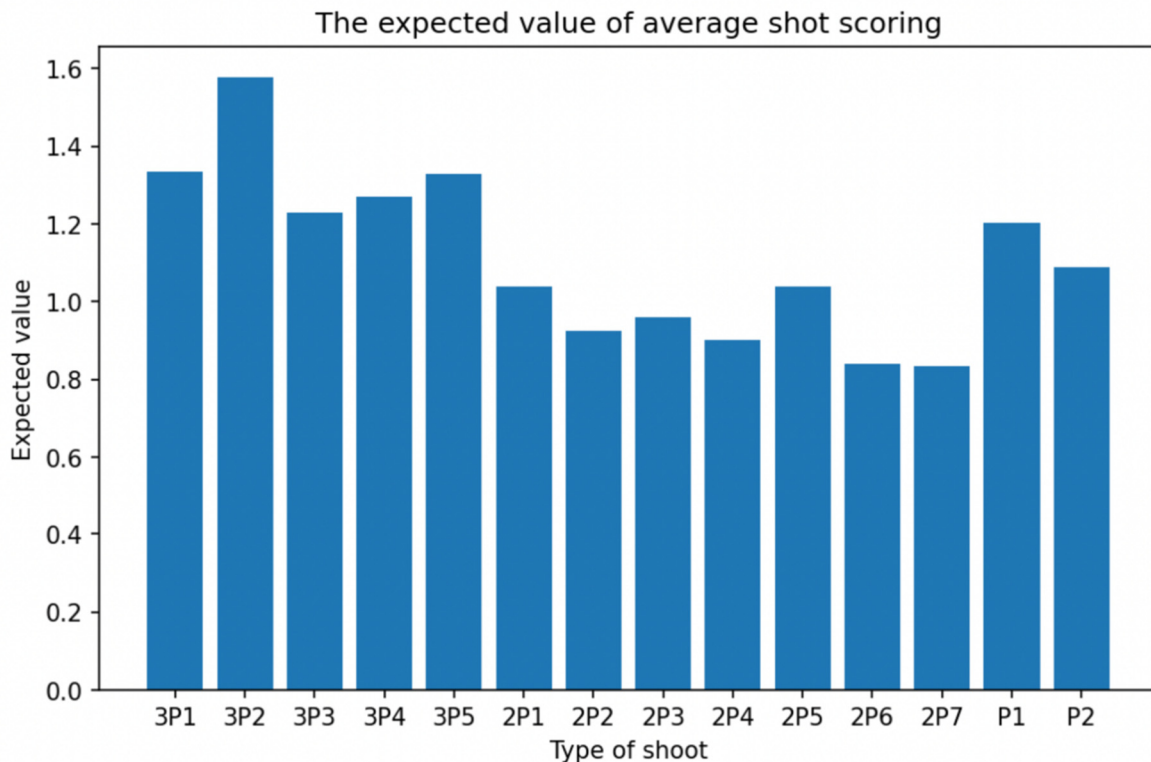

Figure 4. The expected value of average scoring.

After creating the graph of the average expected value of lower and higher percentage shot scores, the analysis of the mathematics part for how the different positions of shooting influence the shot percentage is finished. The next problem is to find out how the way he shot the ball influenced the shot percentage. In order to do that, the first thing is to collect Stephen Curry's different shot percentages under different defenses, and then made a relevant bar statistical chart.

Then compare this data to find the best shooting method of Stephen Curry under these conditions.

To better compare the shot percentage of different defenses on Stephen Curry, the first thing is to separate all kinds of defense into 4 kinds of the main type: Tight defense (defender distance 0-60cm). Close defense (60cm-1m 2). Big Space (1m 2-1m 8). No defense (above 1m 8).

Collecting data from website [2], by creating a python program can help to show all the data:

```
import matplotlib.pyplot as plt

a = ('Tight defense','close defence','Big sapce','no defence')

b = [0.545,0.483,0.468,0.489]

plt.bar(a,b,width=0.8)

plt.title('The Shooting percentage of different defence')

plt.xlabel('Type of defence')

plt.ylabel('shooting percentage')

plt.show()
```
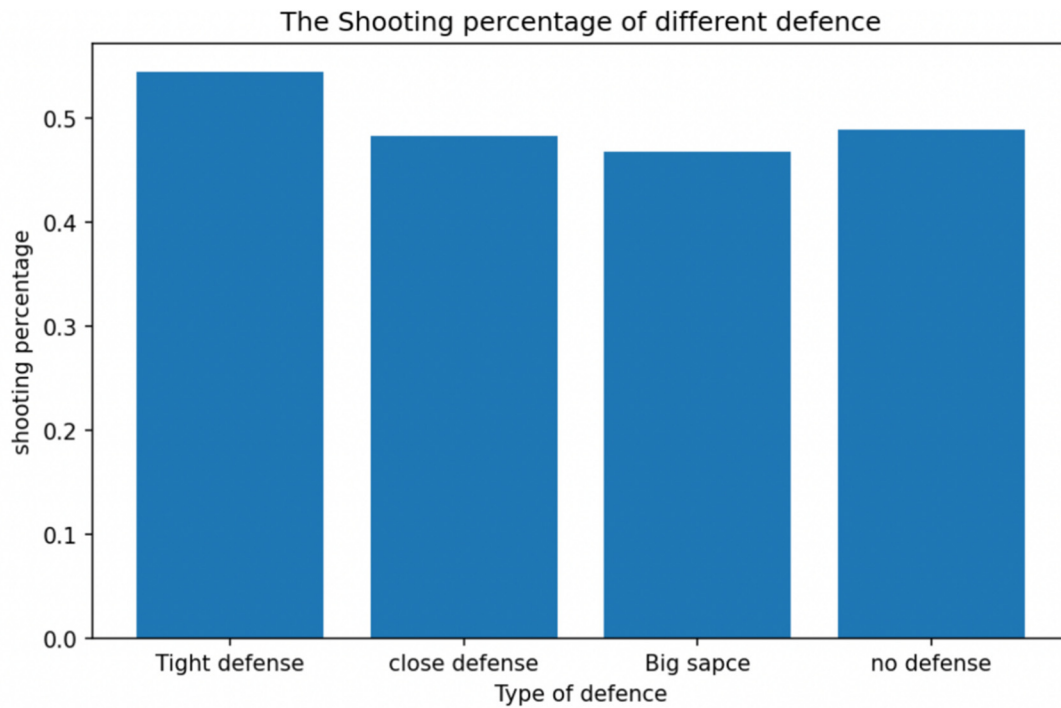
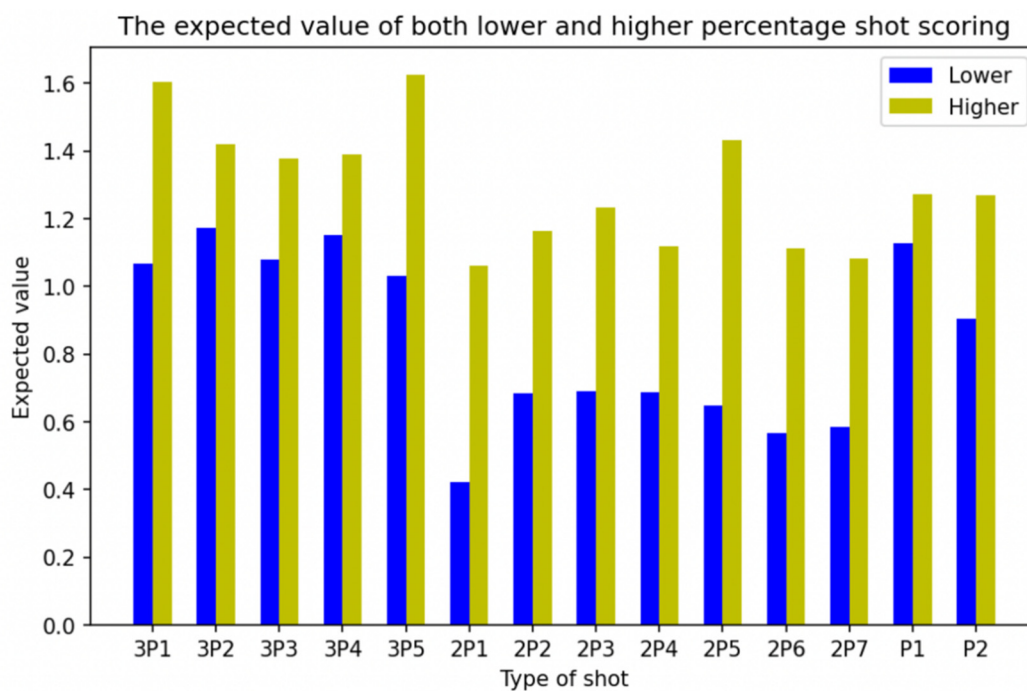**Figure 5.** The Shooting percentage under different defenses.

Figure 6. The expected value of both lower and higher percentage shot scoring

From the figures above, it is easy to conclude that the high expected value for higher bound percentage shot scoring took place when Stephen Curry made his shots at 3P1 and 3P5 and 2P5, the high expected value for lower bound percentage shot scoring took place when Stephen Curry made his shots at 3P2 and 3P4 and P1.
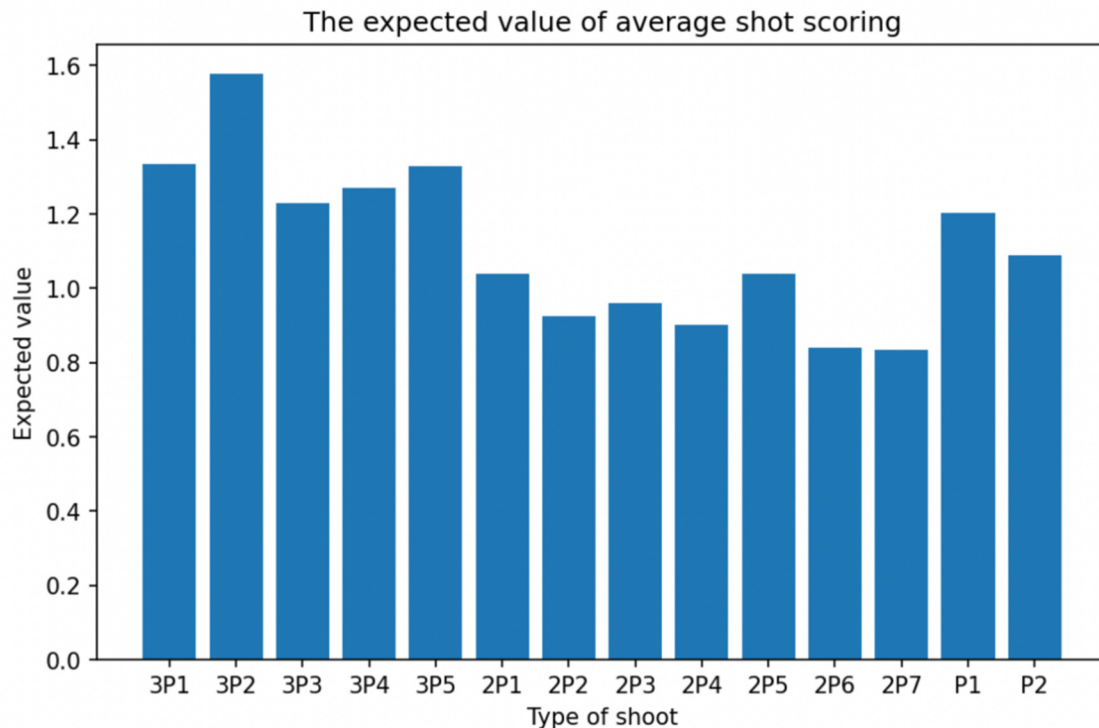
Figure 7. The expected value of average scoring.

By looking at the graph for the average expected value of lower and higher percentage shot scoring, it is easy to infer that, at 3P2, Stephen Curry has the highest average expected value of lower and higher percentage shot scoring.
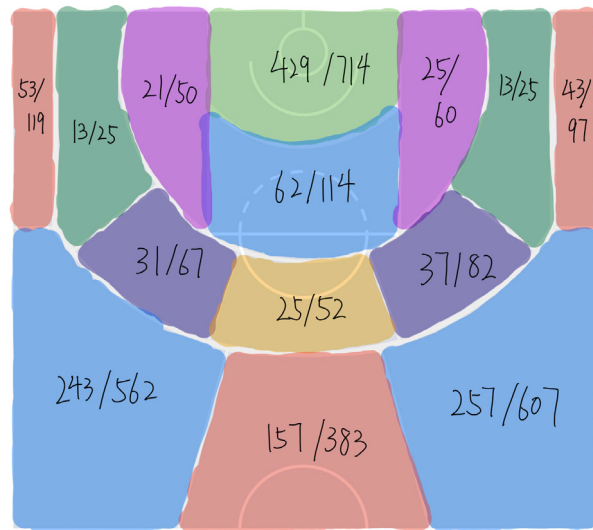
Figure 8. The attempts and made in different zones.

From this graph, it is easy to conclude that in the past four years, Stephen Curry took most of his shots in P1, 3P4, and 3P2. This probably indicates that he is more confident when he shot at those zones than other parts of the court.
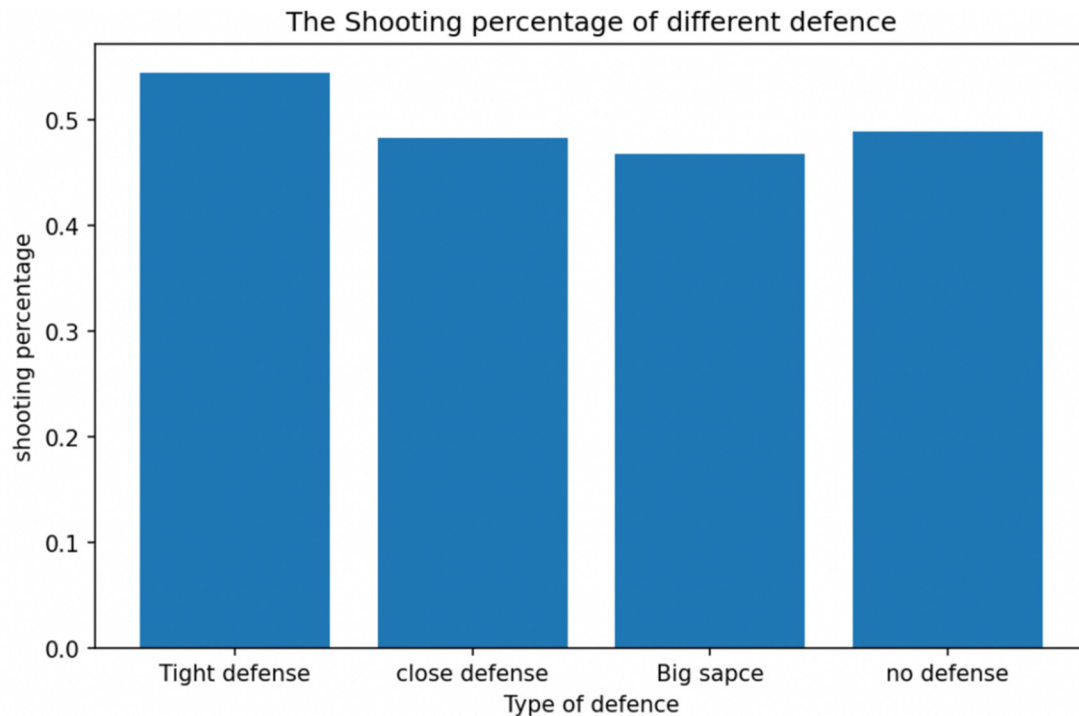
Figure 9. The Shooting percentage under different defenses.

From this graph, obviously, Stephen Curry's shot percentage is relatively the highest when he was under tight defense, even higher than no defense, and it's about 54.5%.
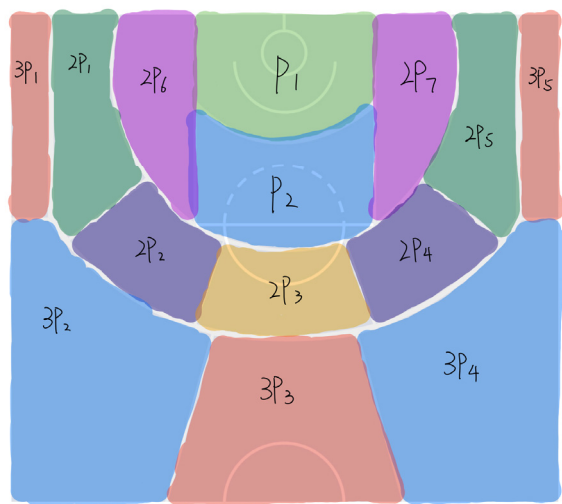

Figure 10. The shooting zones

3. CONCLUSION

Overall, according to the findings, zones3P2 has the greatest lower bound percentage shot scoring expected value and the fourth-highest higher bound percentage shot scoring expected value. Furthermore, the interval expectation is fairly constant. With the third most shot attempts and made shots, it also has the greatest average expected value of lower and higher percentage shot scoring. As a result, it can be determined that 3P2 is the most effective scoring zone for him [3]. Nevertheless,

while facing a tight defense, Stephen Curry has the greatest shooting percentage [2][5]. As a result, the conclusion is that Stephen Curry has had the greatest scoring efficiency in the last four years when shooting at 3P2 against the tight defense.

Stephen Curry's scoring efficiency and style were solely investigated in this study [3][5][6]. Nevertheless, because every player's status fluctuates from season to season, most likely due to age or injury, the collected data cannot eliminate outliers. In the future, it is hoped to develop a model that can be used to determine the appropriate scoring system and position based on any player's statistics. It would be preferable if outliers could be removed from the data.

REFERENCES

[1] Li. BoFang. (2020). "Stephen Curry's data. stat-nba." Retrieved from http://www.stat-nba.com/player/526.html

[2] Luo (2022). "Three points of high-order data in the library are released!" Retrieved from https://baijiahao.baidu.com/s?id=17208883419717 49060&wfr=spider&for=pc

[3] Qi.Tang. "Study on the characteristics of Kuri's accurate three-point Technology", Sport Science and Technology, page (1). (2018)

[4] Tencent sport (2022). Worriors-Stephen Curry. Retrieved from https://sports.qq.com/kbsweb/players.htm?pid=461 2

[5] Ju, Weisong. "On Curry's three-point ability and its importance in the game", Shandong Basketball Management Center, page (9). (2019)

[6] Yilin.Wan. "Research on Curry 's shooting ability of NBA Finals in 2016-2017", Science & Technology of Stationery & Sporting Goods, Page (10). (2019)