678 Final Project

Paul Moon

Abstract

In the movie Moneyball, 'Brad Pitt' uses pure statistics to select which MLB players should play for the Athletics, disregarding what other teams and analysts say about their skill based on other parameters, such as "looks or personality." He used on-base percentage (OBP) and slugging percentage with the main goal of "getting on base" because "getting on base" meant "runs" and the more "runs" you score, the higher the chances of winning was.

In my final project, my goal is to do something similar but with the NBA. The NBA mainly consists of 3 statistics: points, rebounds, and assists (PRAs). First, I will find the average PRAs a playoff NBA team records throughout an entire 82-game season and compare it to an NBA team that does not make the playoffs. That will be the basis of our project. Find players who will get you the amount of PRAs you need to reach the playoffs.

In addition to PRAs, I will also be looking into Field Goals Made (FGM), Field Goals Attempted (FGA), and 3-Point Statistics. The reason for this comes from Wayne 'The Great One' Gretzky's quote "You miss 100% of the shots you don't take." I believe that the more shots a player attempts, the more likely they will score more points, resulting in more wins.

Finally, I will be looking into Games Started (GS) and Minutes Played (MIN) due to the fact that you want to draft/trade for players who know how to stay healthy and are less prone to injury.

The goal is to look at statistics of the past 3 years: the 2021-22, 2022-23, and 2023-24 NBA Seasons. This is due to the fact that the 2019-20 NBA Season was cut short due to Covid-19 and the 2020-21 NBA Season only had 72 instead of the regular 82 games due to the same reason. Most importantly, the new playoff format was updated in the 2020-21 NBA Season with the introduction of the play-in tournament.

Introduction

With the introduction of the play-in tournament, after the conclusion of the 82 game season, 6 teams per conference (total of 12 teams) will automatically make the playoffs. In addition

to that, 4 more teams per conference (total of 8 teams) will have the opportunity to still make it into the playoffs following this format:

1*

Data of both play-in tournament requirements and playoff requirements will be analyzed in this section; however, in this project, I will calculate the playoff scenario where play-in tournament teams have a 50% chance in making the playoffs. This makes sense since every year there are 8 play-in tournament teams and only 4 teams make it into the playoffs.

2021-22

Eastern Conference

	W	L	PCT	GB	HOME	AWAY	DIV	CONF	PPG	OPP PPG	DIFF	STRK	L10
z – Miami Heat	53	29	.646	-	29-12	24-17	13-3	35-17	110.0	105.6	+4.4	L1	6-4
y – Boston Celtics	51	31	.622	2	28-13	23-18	9-7	33-19	111.8	104.5	+7.3	W1	7-3
y – Milwaukee Bucks	51	31	.622	2	27-14	24-17	12-4	33-19	115.5	112.1	+3.4	L1	6-4
x – Philadelphia 76ers	51	31	.622	2	24-17	27-14	6-10	32-20	109.9	107.3	+2.6	W2	6-4
x – Toronto Raptors	48	34	.585	5	24-17	24-17	10-6	30-22	109.4	107.1	+2.3	L1	8-2
x – Chicago Bulls	46	36	.561	7	27-14	19-22	10-6	29-23	111.6	112.0	-0.4	W1	4-6
xp – Brooklyn Nets	44	38	.537	9	20-21	24-17	10-6	31-21	112.9	112.1	+0.8	W4	6-4
xp – Atlanta Hawks	43	39	.524	10	27-14	16-25	9-7	26-26	113.9	112.4	+1.5	W1	7-3
pb – Cleveland Cavaliers	44	38	.537	9	25-16	19-22	10-6	27-25	107.8	105.7	+2.1	W1	3-7
pb – Charlotte Hornets	43	39	.524	10	22-19	21-20	8-8	27-25	115.3	114.9	+0.4	W3	6-4
e – New York Knicks	37	45	.451	16	17-24	20-21	5-11	22-30	106.5	106.6	-0.1	W2	7-3
e – Washington Wizards	35	47	.427	18	21-20	14-27	7-9	24-28	108.6	112.0	-3.4	L3	5-5
e – Indiana Pacers	25	57	.305	28	16-25	9-32	2-14	11-41	111.5	114.9	-3.4	L10	0-10
e – Detroit Pistons	23	59	.280	30	13-28	10-31	6-10	18-34	104.8	112.5	-7.7	L3	4-6
e – Orlando Magic	22	60	.268	31	12-29	10-31	3-13	12-40	104.2	112.2	-8.0	W1	3-7

Western Conference

	W	L	PCT	GB	HOME	AWAY	DIV	CONF	PPG	OPP PPG	DIFF	STRK	L10
* – Phoenix Suns	64	18	.780	-	32-9	32-9	10-6	39-13	114.8	107.4	+7.4	L1	6-4
y – Memphis Grizzlies	56	26	.683	8	30-11	26-15	11-5	36-16	115.6	109.9	+5.7	L1	7-3
x – Golden State Warriors	53	29	.646	11	31-10	22-19	12-4	33-19	111.0	105.5	+5.5	W5	6-4
x – Dallas Mavericks	52	30	.634	12	29-12	23-18	14-2	36-16	108.0	104.7	+3.3	W4	8-2
y – Utah Jazz	49	33	.598	15	29-12	20-21	15-1	33-19	113.6	107.6	+6.0	W1	4-6
x – Denver Nuggets	48	34	.585	16	23-18	25-16	6-10	29-23	112.7	110.4	+2.3	L1	6-4
xp – Minnesota Timberwolves	46	36	.561	18	26-15	20-21	12-4	32-20	116.0	113.3	+2.7	L1	4-6
xp – New Orleans Pelicans	36	46	.439	28	19-22	17-24	6-10	25-27	109.3	110.3	-1.0	L2	6-4
pb – LA Clippers	42	40	.512	22	25-16	17-24	9-7	26-26	108.4	108.4	0.0	W5	6-4
pb – San Antonio Spurs	34	48	.415	30	16-25	18-23	6-10	24-28	113.2	113.0	+0.2	L3	6-4
e – Los Angeles Lakers	33	49	.402	31	21-20	12-29	3-13	18-34	112.1	115.1	-3.0	W2	2-8
e – Sacramento Kings	30	52	.366	34	16-25	14-27	6-10	20-32	110.3	115.8	-5.5	W1	5-5
e – Portland Trail Blazers	27	55	.329	37	17-24	10-31	1-15	11-41	106.2	115.1	-8.9	L11	0-10
e – Oklahoma City Thunder	24	58	.293	40	12-29	12-29	6-10	17-35	103.7	111.8	-8.1	L3	4-6
e – Houston Rockets	20	62	.244	44	11-30	9-32	3-13	11-41	109.7	118.2	-8.5	L7	3-7

2*

In the 2021-2022 NBA Season, if you were in the Eastern Conference, a team needed 38 wins to qualify for the play-in tournament. To instantly qualify for the playoffs, a team needed 45 wins. However, in the Western Conference, a team needed only 34 wins to qualify for the play-in tournament; however, a team needed 47 wins to instantly qualify for the playoffs.

After the conclusion of the play-in tournament, playoff qualifying teams passed the average threshold of 44 wins in the Eastern Conference and 39 wins in the Western Conference (The average threshold was found by averaging the lowest playoff team to the highest non-playoff team).

2022-23

Eastern Conference

	W	L	PCT	GB	HOME	AWAY	DIV	CONF	PPG	OPP PPG	DIFF	STRK	L10
* -- Milwaukee Bucks	58	24	.707	-	32-9	26-15	11-5	35-17	116.9	113.3	+3.6	L2	6-4
y -- Boston Celtics	57	25	.695	1	32-9	25-16	11-5	34-18	117.9	111.4	+6.5	W3	8-2
x -- Philadelphia 76ers	54	28	.659	4	29-12	25-16	10-6	34-18	115.2	110.9	+4.3	W2	5-5
x -- Cleveland Cavaliers	51	31	.622	7	31-10	20-21	13-3	34-18	112.3	106.9	+5.4	L1	7-3
x -- New York Knicks	47	35	.573	11	23-18	24-17	8-8	32-20	116.0	113.1	+2.9	L2	5-5
x -- Brooklyn Nets	45	37	.549	13	23-18	22-19	7-9	30-22	113.4	112.5	+0.9	L1	6-4
xp -- Atlanta Hawks	41	41	.500	17	24-17	17-24	8-8	26-26	118.4	118.1	+0.3	L2	5-5
y -- Miami Heat	44	38	.537	14	27-14	17-24	10-6	24-28	109.5	109.8	-0.3	W1	6-4
pb -- Chicago Bulls	40	42	.488	18	22-19	18-23	7-9	27-25	113.1	111.8	+1.3	W2	6-4
pb -- Toronto Raptors	41	41	.500	17	27-14	14-27	4-12	26-26	112.9	111.4	+1.5	W1	6-4
e -- Indiana Pacers	35	47	.427	23	20-21	15-26	7-9	24-28	116.3	119.5	-3.2	W1	3-7
e -- Washington Wizards	35	47	.427	23	19-22	16-25	8-8	21-31	113.2	114.4	-1.2	L1	3-7
e -- Orlando Magic	34	48	.415	24	20-21	14-27	7-9	20-32	111.4	114.0	-2.6	L4	5-5
e -- Charlotte Hornets	27	55	.329	31	13-28	14-27	7-9	15-37	111.0	117.2	-6.2	W1	5-5
e -- Detroit Pistons	17	65	.207	41	9-32	8-33	2-14	8-44	110.3	118.5	-8.2	L1	1-9

Western Conference

	W	L	PCT	GB	HOME	AWAY	DIV	CONF	PPG	OPP PPG	DIFF	STRK	L10
z – Denver Nuggets	53	29	.646	-	34-7	19-22	10-6	34-18	115.8	112.5	+3.3	W1	5-5
y – Memphis Grizzlies	51	31	.622	2	35-6	16-25	13-3	30-22	116.9	113.0	+3.9	L1	6-4
y – Sacramento Kings	48	34	.585	5	23-18	25-16	9-7	32-20	120.7	118.1	+2.6	L3	5-5
x – Phoenix Suns	45	37	.549	8	28-13	17-24	9-7	30-22	113.6	111.6	+2.0	L2	7-3
x – LA Clippers	44	38	.537	9	23-18	21-20	9-7	27-25	113.6	113.1	+0.5	W3	6-4
x – Golden State Warriors	44	38	.537	9	33-8	11-30	7-9	30-22	118.9	117.1	+1.8	W3	8-2
xp – Los Angeles Lakers	43	39	.524	10	23-18	20-21	6-10	27-25	117.2	116.6	+0.6	W2	8-2
xp – Minnesota Timberwolves	42	40	.512	11	22-19	20-21	8-8	29-23	115.8	115.8	0.0	W3	7-3
pb – Oklahoma City Thunder	40	42	.488	13	24-17	16-25	9-7	25-27	117.5	116.4	+1.1	W2	4-6
pb – New Orleans Pelicans	42	40	.512	11	27-14	15-26	11-5	29-23	114.4	112.5	+1.9	L1	7-3
e – Dallas Mavericks	38	44	.463	15	23-18	15-26	9-7	28-24	114.2	114.1	+0.1	L2	2-8
e – Utah Jazz	37	45	.451	16	23-18	14-27	6-10	24-28	117.1	118.0	-0.9	L1	2-8
e – Portland Trail Blazers	33	49	.402	20	17-24	16-25	7-9	23-29	113.4	117.4	-4.0	L4	1-9
e – Houston Rockets	22	60	.268	31	14-27	8-33	4-12	12-40	110.7	118.6	-7.9	W3	4-6
e – San Antonio Spurs	22	60	.268	31	14-27	8-33	3-13	10-42	113.0	123.1	-10.1	W1	3-7

3*

In the 2022-2023 NBA Season, if you were in the Eastern Conference, a team needed 36 wins to qualify for the play-in tournament. To instantly qualify for the playoffs, a team needed 45 wins. In the Western Conference, a team needed 39 wins to qualify for the play-in tournament and 44 wins to instantly qualify for the playoffs.

After the conclusion of the play-in tournament, playoff qualifying teams passed the average threshold of 41 wins in the Eastern Conference and 42 wins in the Western Conference.

2023-24

Eastern Conference

	W	L	PCT	GB	HOME	AWAY	DIV	CONF	PPG	OPP PPG	DIFF	STRK	L10
* -- Boston Celtics	64	18	.780	-	37-4	27-14	15-2	41-11	120.6	109.2	+11.4	W2	7-3
x -- New York Knicks	50	32	.610	14	27-14	23-18	12-5	35-17	112.8	108.2	+4.6	W5	6-4
y -- Milwaukee Bucks	49	33	.598	15	31-11	18-22	10-7	34-18	119.0	116.4	+2.6	L2	3-7
x -- Cleveland Cavaliers	48	34	.585	16	26-15	22-19	11-5	31-21	112.6	110.2	+2.4	L1	4-6
y -- Orlando Magic	47	35	.573	17	29-12	18-23	9-7	32-20	110.5	108.5	+2.0	W1	5-5
x -- Indiana Pacers	47	35	.573	17	26-15	21-20	11-6	32-20	123.3	120.2	+3.1	W1	7-3
xp -- Philadelphia 76ers	47	35	.573	17	25-16	22-19	8-8	31-21	114.6	111.5	+3.1	W8	8-2
xp -- Miami Heat	46	36	.561	18	22-19	24-17	13-3	32-20	110.1	108.4	+1.7	W2	7-3
pb -- Chicago Bulls	39	43	.476	25	20-21	19-22	7-9	22-29	112.3	113.7	-1.4	L1	5-5
pb -- Atlanta Hawks	36	46	.439	28	21-20	15-26	8-8	22-30	118.3	120.5	-2.2	L6	3-7
e -- Brooklyn Nets	32	50	.390	32	20-21	12-29	5-11	24-28	110.4	113.3	-2.9	L2	5-5
e -- Toronto Raptors	25	57	.305	39	14-27	11-30	1-15	18-34	112.4	118.8	-6.4	L4	2-8
e -- Charlotte Hornets	21	61	.256	43	11-30	10-31	6-10	14-38	106.6	116.8	-10.2	W1	3-7
e -- Washington Wizards	15	67	.183	49	7-34	8-33	4-12	11-41	113.7	123.0	-9.3	L6	1-9
e -- Detroit Pistons	14	68	.171	50	7-33	7-35	2-14	10-41	109.9	119.0	-9.1	L1	2-8

Western Conference

	W	L	PCT	GB	HOME	AWAY	DIV	CONF	PPG	OPP PPG	DIFF	STRK	L10
z -- Oklahoma City Thunder	57	25	.695	-	33-8	24-17	12-4	36-16	120.1	112.7	+7.4	W5	7-3
x -- Denver Nuggets	57	25	.695	-	33-8	24-17	10-6	33-19	114.9	109.6	+5.3	W1	6-4
x -- Minnesota Timberwolves	56	26	.683	1	30-11	26-15	12-4	37-15	113.0	106.5	+6.5	L1	6-4
y -- LA Clippers	51	31	.622	6	25-16	26-15	9-7	30-22	115.6	112.3	+3.3	L3	6-4
y -- Dallas Mavericks	50	32	.610	7	25-16	25-16	11-5	31-21	117.9	115.6	+2.3	L2	7-3
x -- Phoenix Suns	49	33	.598	8	25-16	24-17	9-9	29-23	116.2	113.2	+3.0	W3	7-3
xp -- Los Angeles Lakers	47	35	.573	10	28-14	19-21	7-10	27-25	118.0	117.4	+0.6	W2	7-3
xp -- New Orleans Pelicans	49	33	.598	8	21-19	28-14	9-7	30-22	115.1	110.7	+4.4	L1	5-5
pb -- Sacramento Kings	46	36	.561	11	24-17	22-19	10-7	30-22	116.6	114.8	+1.8	W1	4-6
pb -- Golden State Warriors	46	36	.561	11	21-20	25-16	7-9	26-26	117.8	115.2	+2.6	W1	8-2
e -- Houston Rockets	41	41	.500	16	27-14	14-27	9-7	28-24	114.3	113.2	+1.1	W2	4-6
e -- Utah Jazz	31	51	.378	26	21-20	10-31	5-11	16-36	115.7	120.5	-4.8	L1	2-8
e -- Memphis Grizzlies	27	55	.329	30	9-32	18-23	8-8	14-37	105.8	112.8	-7.0	L5	3-7
e -- San Antonio Spurs	22	60	.268	35	12-29	10-31	3-13	14-37	112.1	118.6	-6.5	W2	6-4
e -- Portland Trail Blazers	21	61	.256	36	11-30	10-31	1-15	8-44	106.4	115.4	-9.0	L5	2-8

4*

In the 2022-2023 NBA Season, if you were in the Eastern Conference, a team needed 33 wins to qualify for the play-in tournament. To instantly qualify for the playoffs, a team needed 47 wins. In the Western Conference, a team needed 42 wins to qualify for the play-in tournament and 48 wins to instantly qualify for the playoffs.

After the conclusion of the play-in tournament, playoff qualifying teams passed the average threshold of 40 wins in the Eastern Conference and 47 wins in the Western Conference.

To summarize, Eastern Conference teams needed 38, 36, 33 wins to make the play-in tournament and the Western Conference teams needed 34, 39, and 42 wins to make the play-in tournament in 2022, 2023, and 2024 respectively.

According to NBA play-in tournament Wikipedia (5*), the average number of wins required to secure the 10th seed typically falls between 38 and 42 wins. For a more detailed breakdown, The Eastern Conference is historically less competitive, requiring around 38-40 wins whereas the more competitive Western Conference requires around 40-42 wins.

Combining this information with the data that we have, let's set the weight of each of these 50/50 (csv data and online data), the "magic" number that NBA teams are trying to reach to secure a spot in the play-in tournament league-wide will be set to 39.

In order to secure a NBA playoff spot without having to go into the play-in tournament, in previous years, Eastern Conference teams needed 45, 45, 47 wins and the Western Conference teams needed 47, 44, 48 wins in 2022, 2023, 2024 respectively.

According to the NBA's official website (6*), the Eastern Conference requires around 42-45 wins and the Western Conference requires around 44-47 wins to achieve this feat, with the league average being around 45 wins.

Again, combining our csv data and online data, the "magic" number that NBA teams are trying to reach to secure a top 6 spot to go straight into the playoffs will be set to 46.

Combining the 50% chance that teams have in making the playoffs after qualifying for the play-in tournament, this project: the EDA and Modeling will revolve around achieving one single number:

43 wins

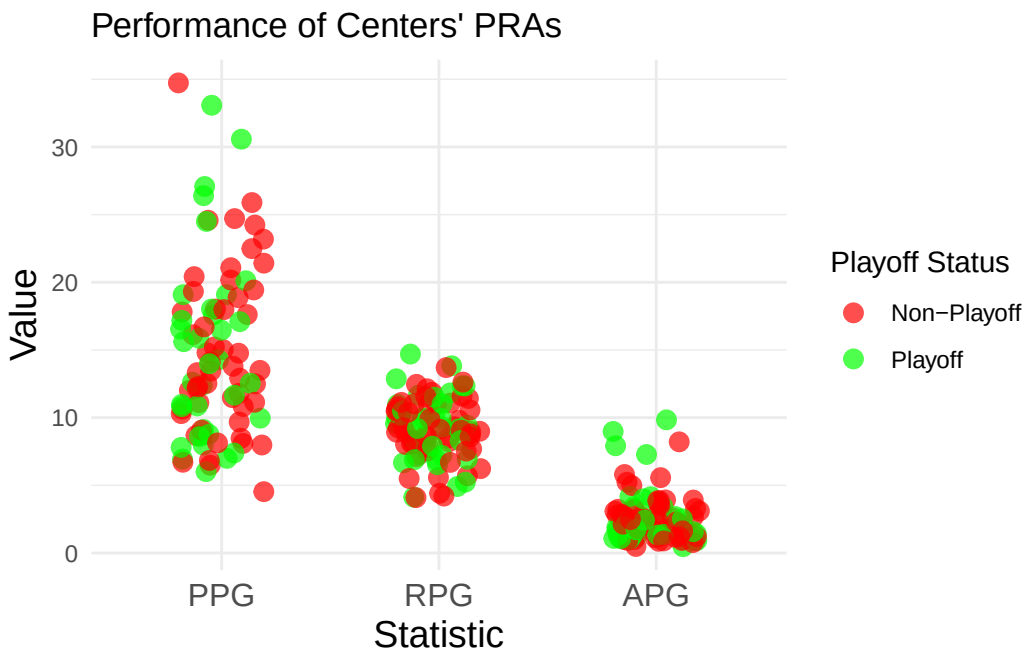
EDA

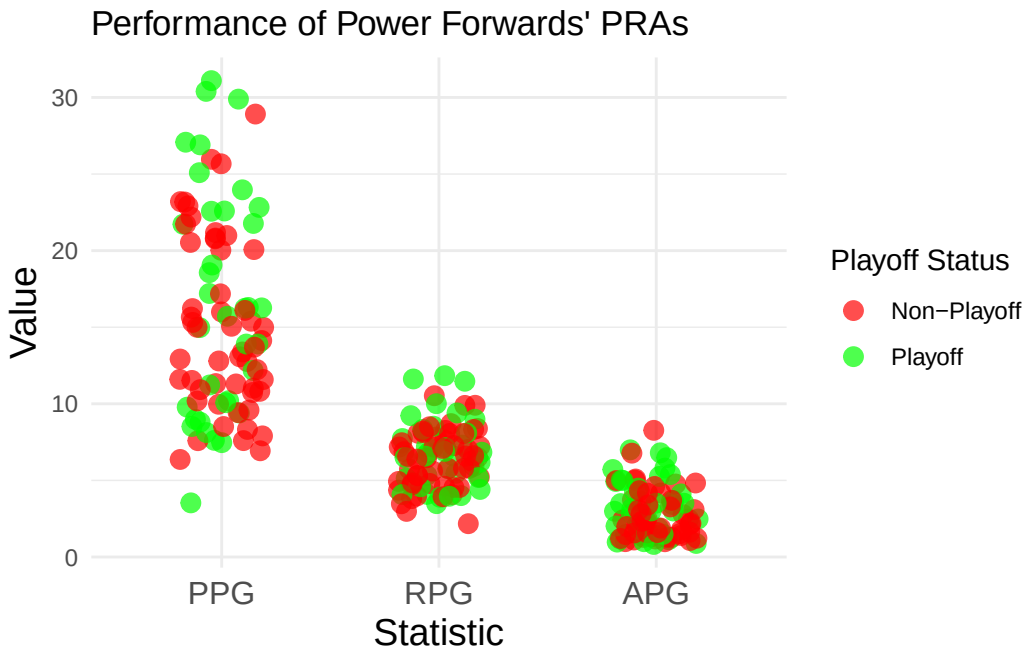
In our EDA, we will visualize the data to see if anything catches our eyes before we dive into the Modeling section. We will be combining and looking at all 3 years of data.

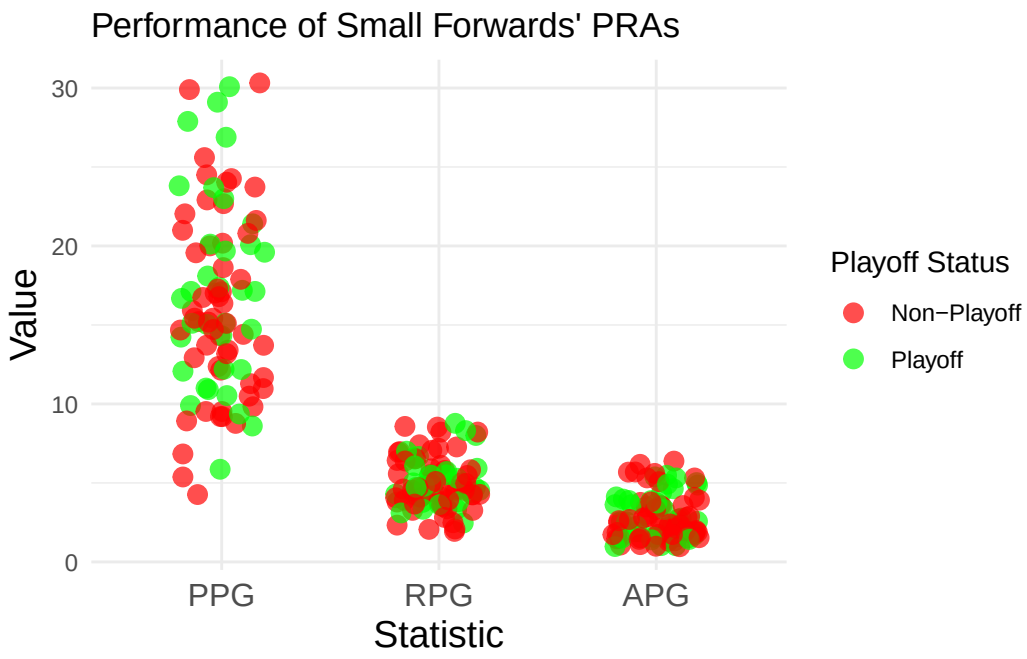
PRAs

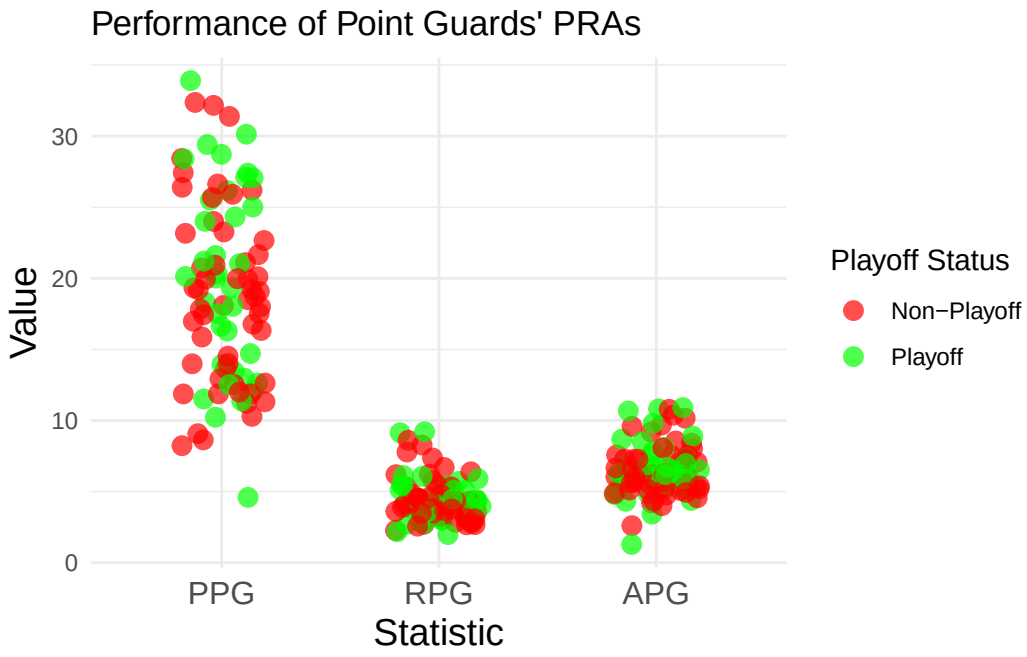
Playoff vs Non-Playoffs

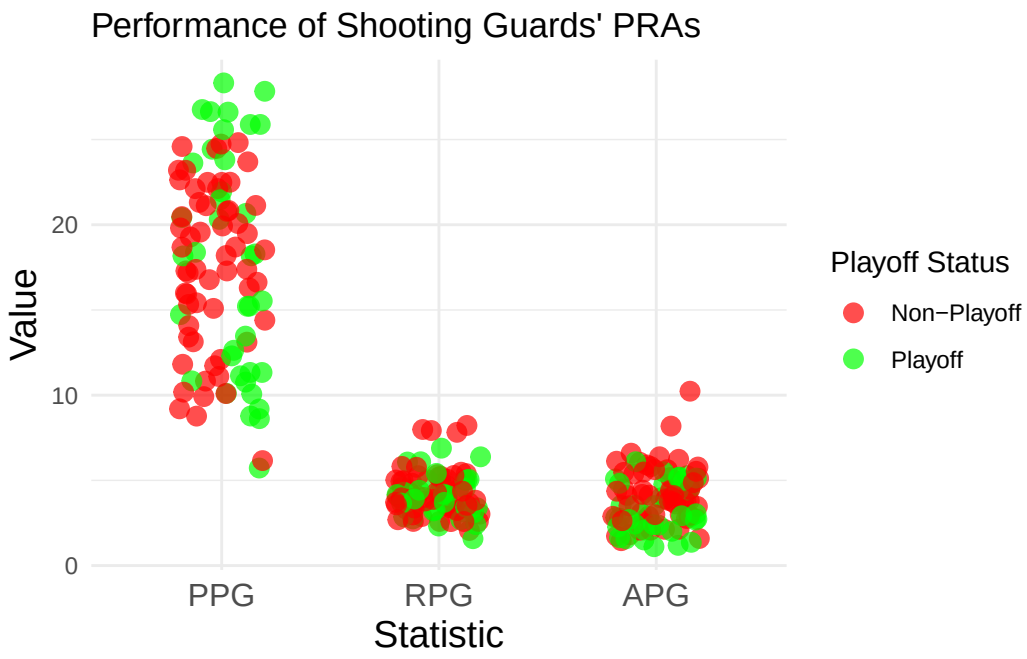
We will first look to see if PRA averages between playoff clinching players vs players who did not clinch the playoffs differ.


Performance of Centers' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Playoff Status
Non−Playoff
Playoff


Performance of Power Forwards' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Playoff Status
Non−Playoff
Playoff

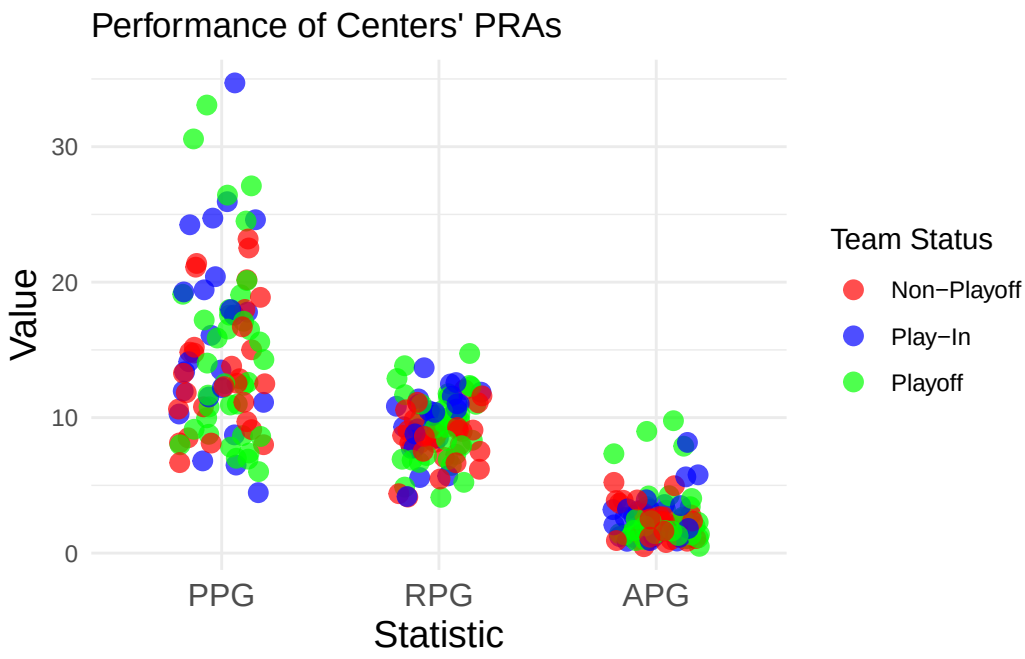

Performance of Small Forwards' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Playoff Status
Non−Playoff
Playoff

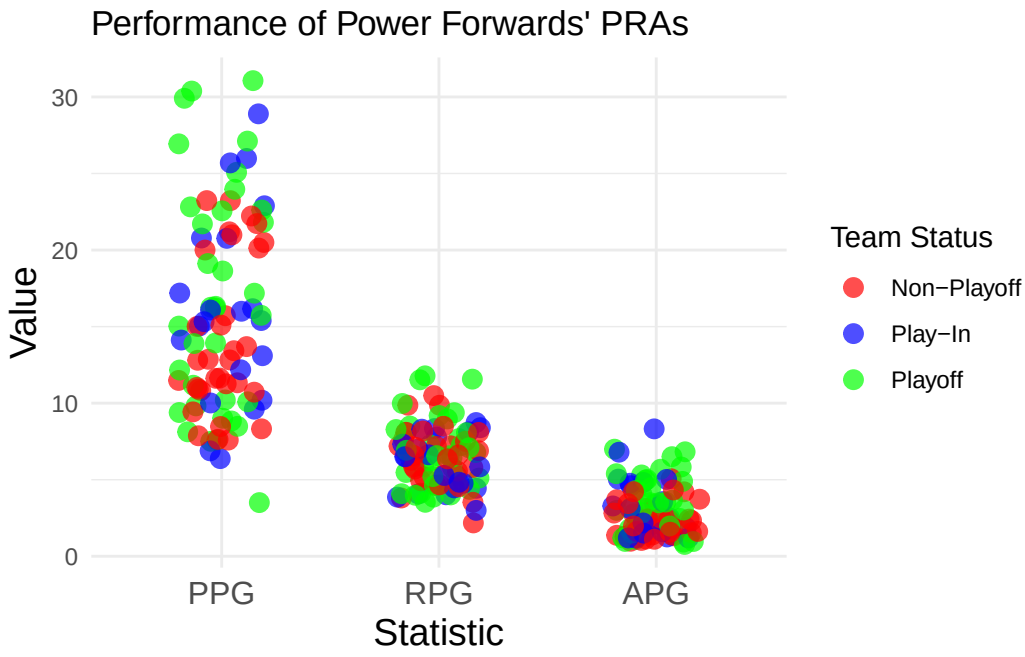

Performance of Point Guards' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Playoff Status
Non−Playoff
Playoff

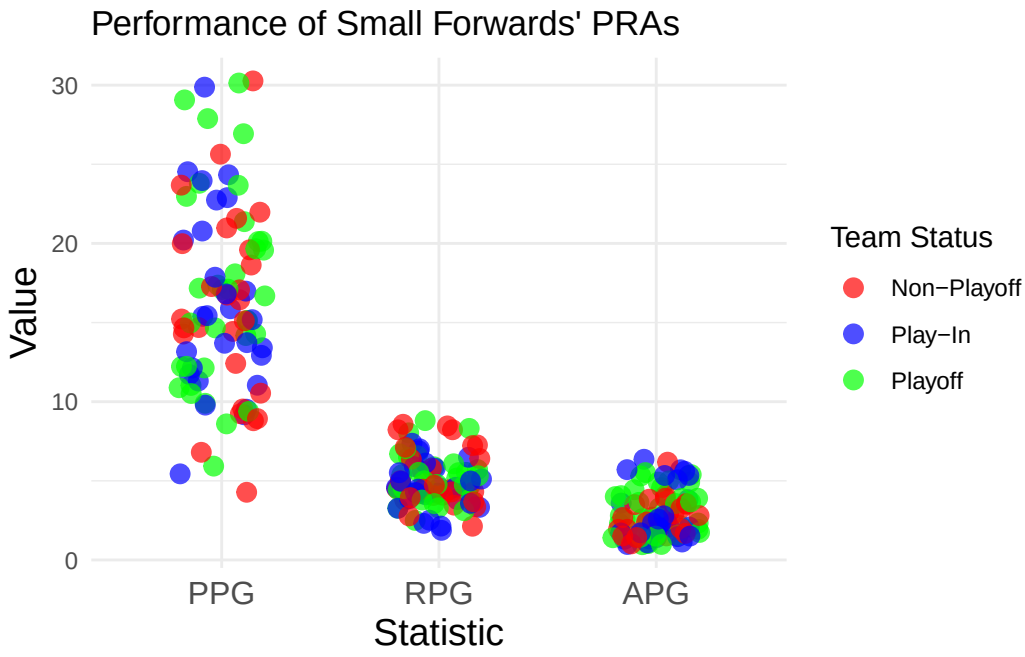


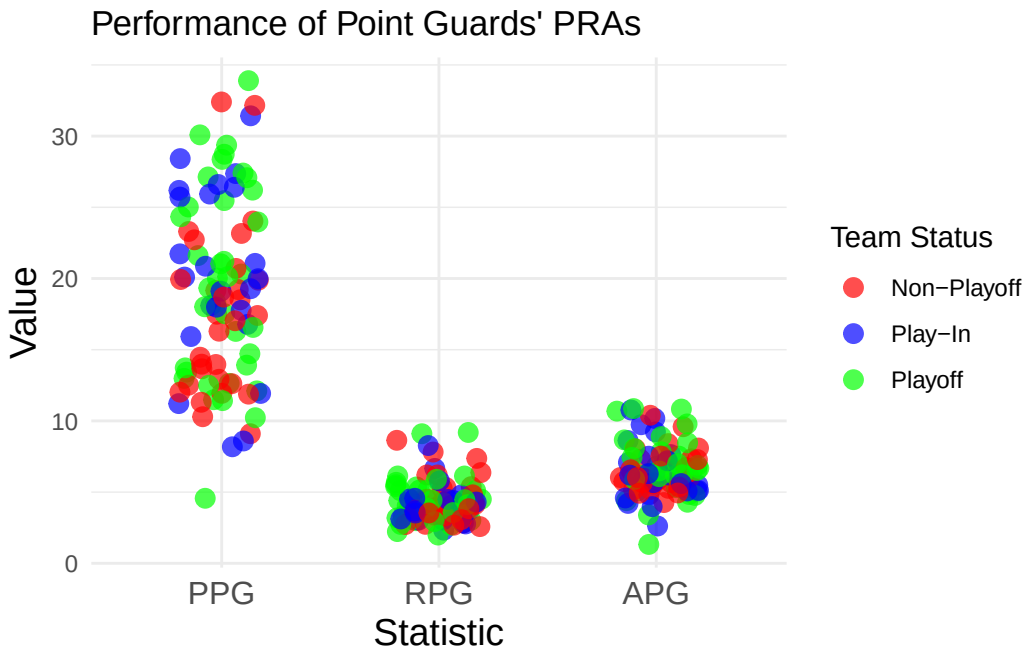
Looking at the data, the green and red dots seem to be scattered "randomly" with only some of the apex dots (players) being green. If playoff teams were "better" than non-playoff teams, then we would expect to see more green dots at the top and the red dots at the bottom. Maybe players on teams that are close to making the playoffs are the red dots at the top; therefore, let's do the same plots again but with play-in tournament players shown in blue.

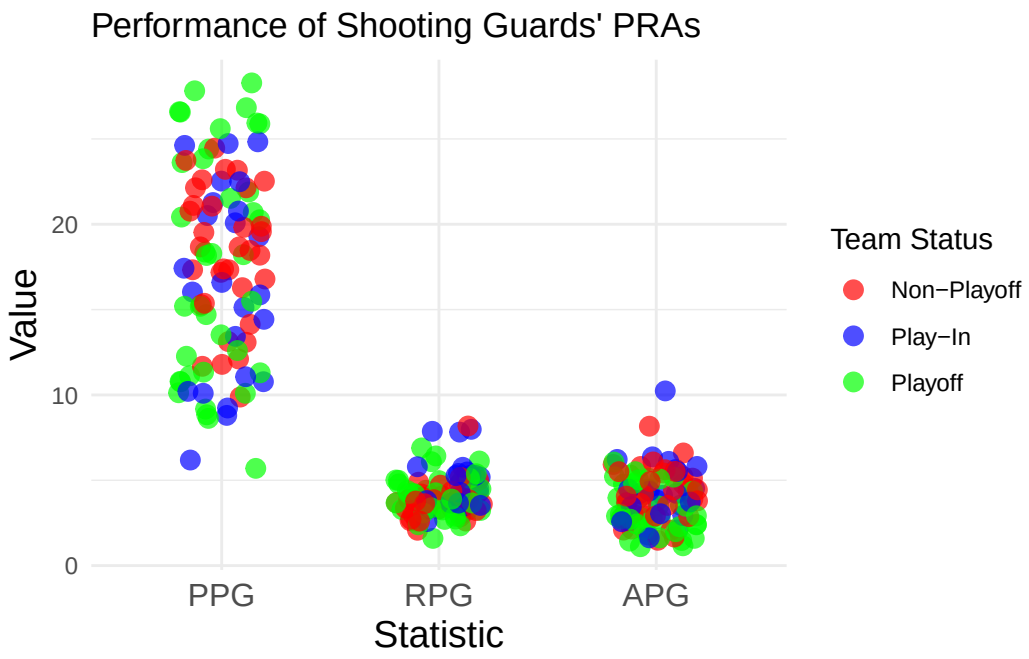
Playoffs vs Play-Ins vs Non-Playoffs


Performance of Centers' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Power Forwards' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Small Forwards' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Point Guards' PRAs
Value
30
20
10
0
PPG
RPG
APG
Statistic
Team Status
Non−Playoff
Play−In
Playoff



Based on these dot plots, we can see that there are fewer red dots near the top. It is mostly filled with green and blue dots, which is what we are aiming for.

Conclusion

It seems as if RPGs, more specifically APGs, are not as important as PPGs due to the low amount players record in general. However, it is important to notice that certain positions will be recording higher counts of either PPGs, RPGs, or APGs than other positions simply due to the play-style of the game. Therefore, this is something that we will be looking at in the Modeling section.

Shooting

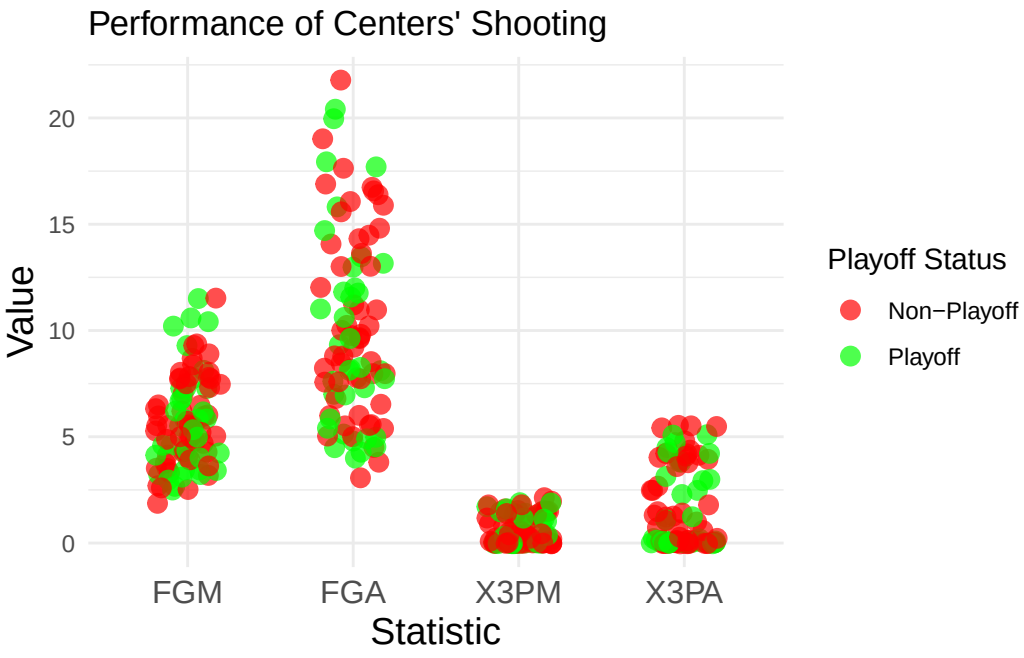
Playoff vs Non-Playoffs

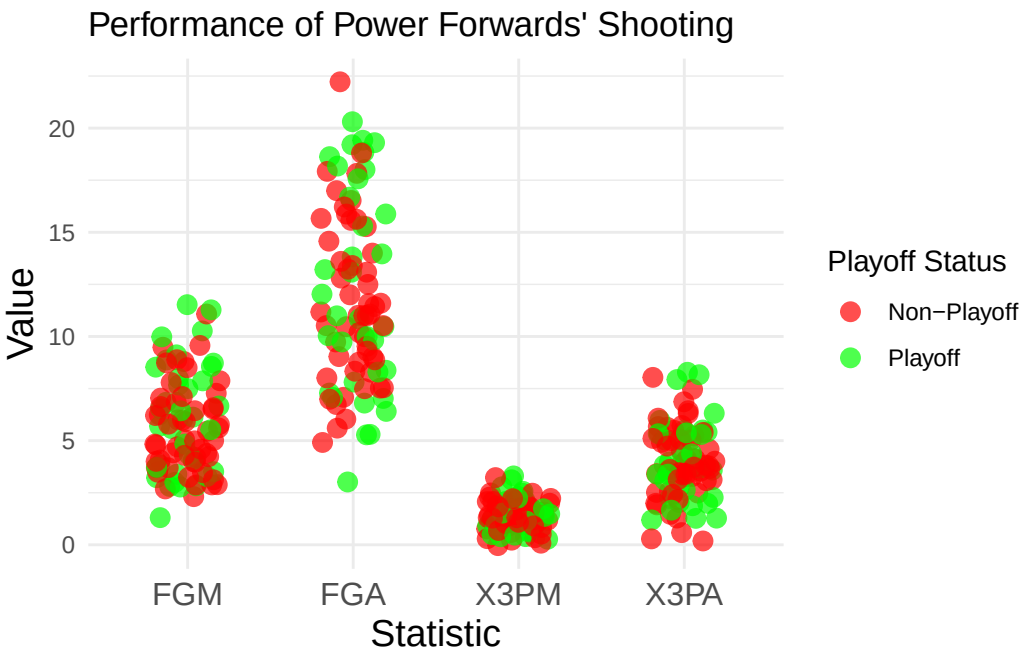
Now, let's see if we see differences in shooting within playoffs vs non-playoffs teams.

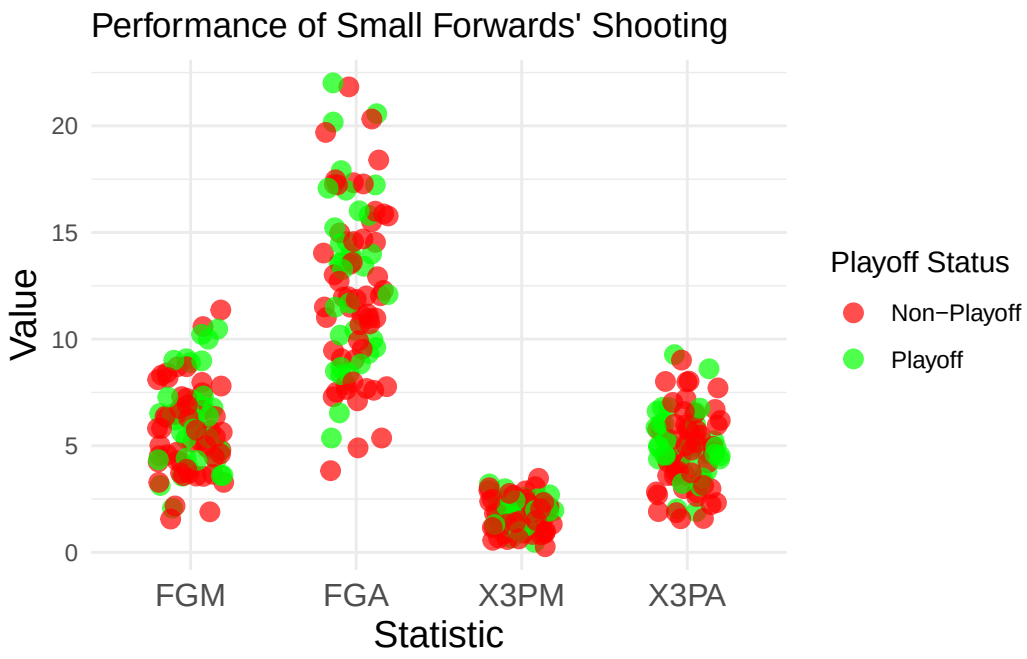
I have two predictions:

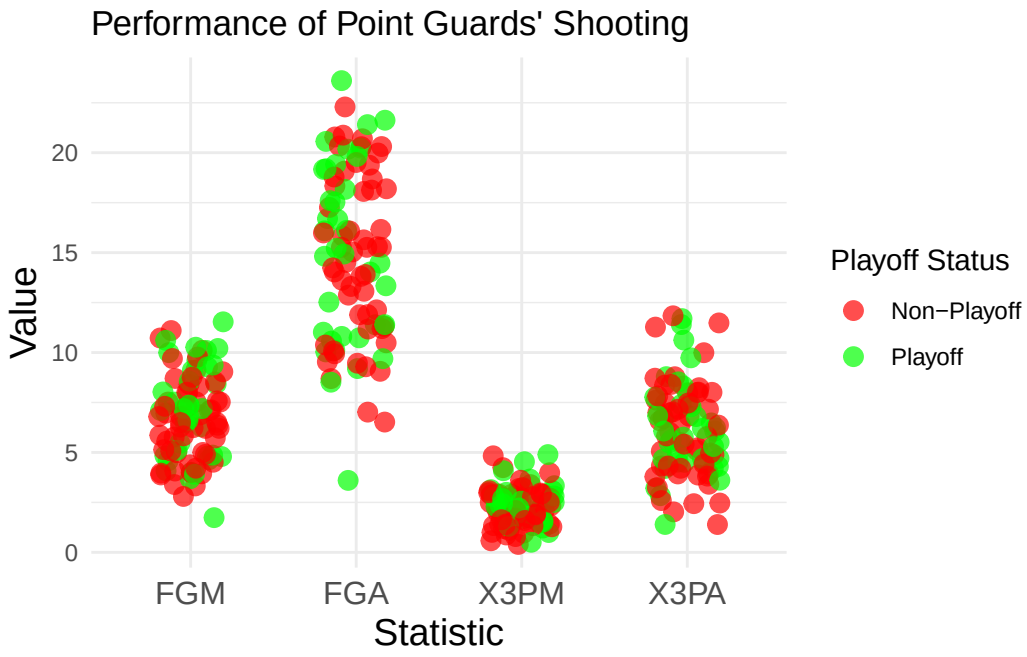
- "Better" teams will have players with higher shooting percentages. In this case, it will be more green dots near the top of FGM and X3PM but lower in FGA and X3PA

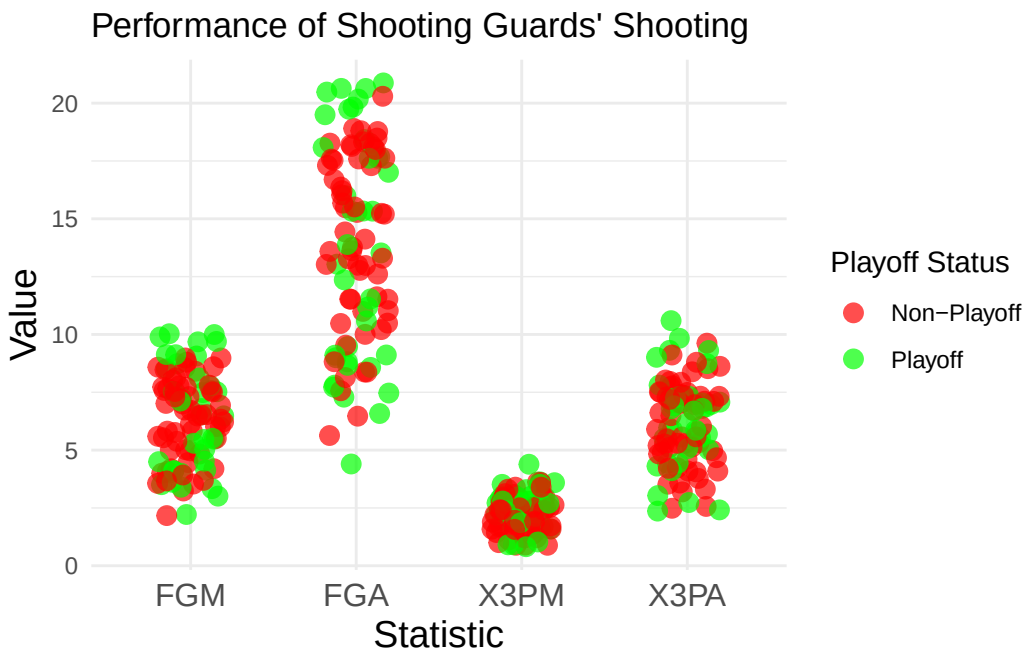
- "Better" teams will have players with higher FGM, FGA, X3PM, and X3PA due to the fact that these players are creating more shooting opportunities and thus with normal shooting percentages, are scoring more.





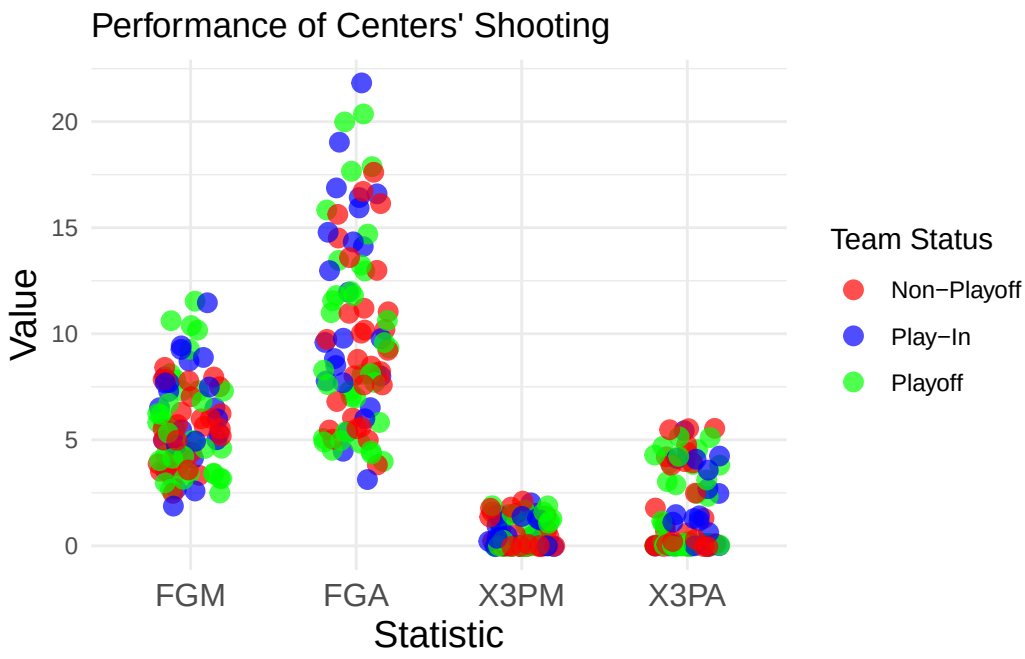

Performance of Small Forwards' Shooting
Value
20
15
10
5
0
FGM
FGA
X3PM
X3PA
Statistic
Playoff Status
Non−Playoff
Playoff

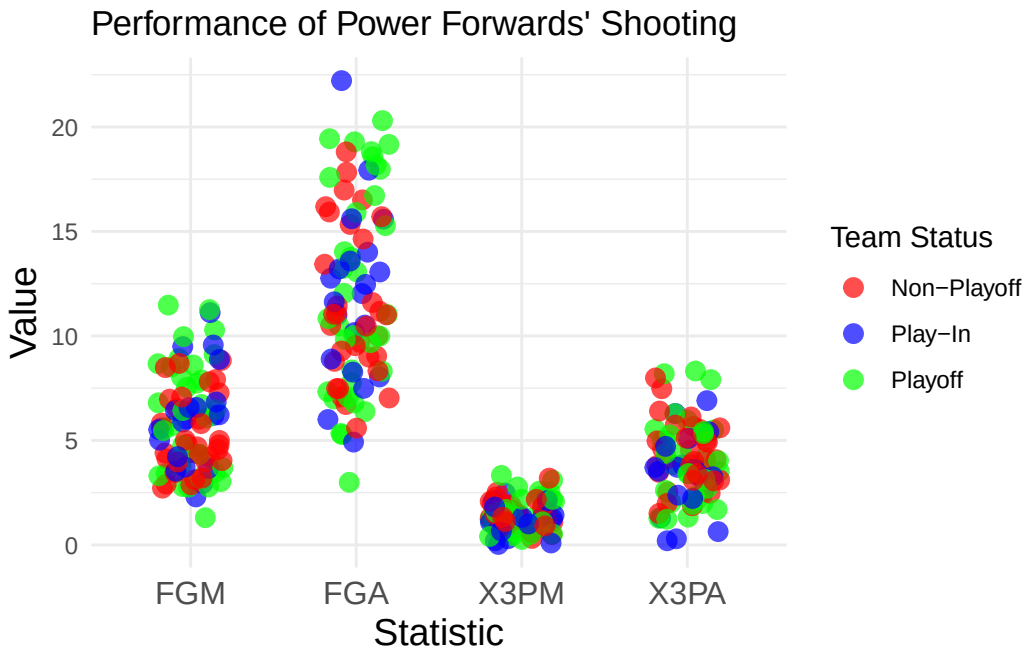

Performance of Point Guards' Shooting
Value
20
15
10
5
0
FGM
FGA
X3PM
X3PA
Statistic
Playoff Status
Non−Playoff
Playoff

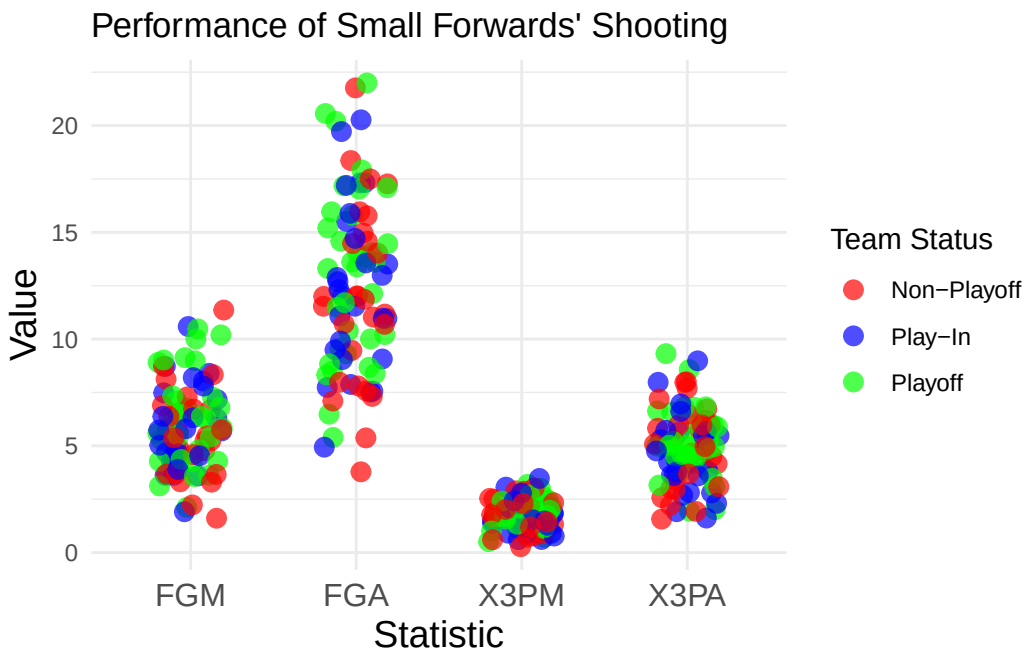


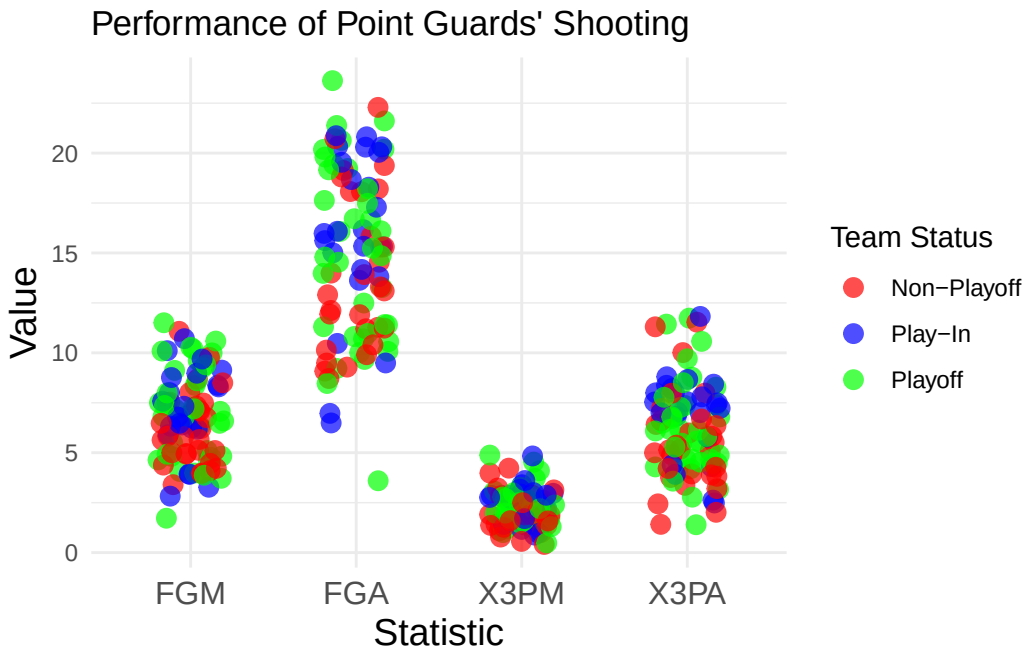
Again, for a clearer visual, let's try adding in the play-in teams to see if the higher red dots in the dot plots are from teams who are close to making the playoffs.

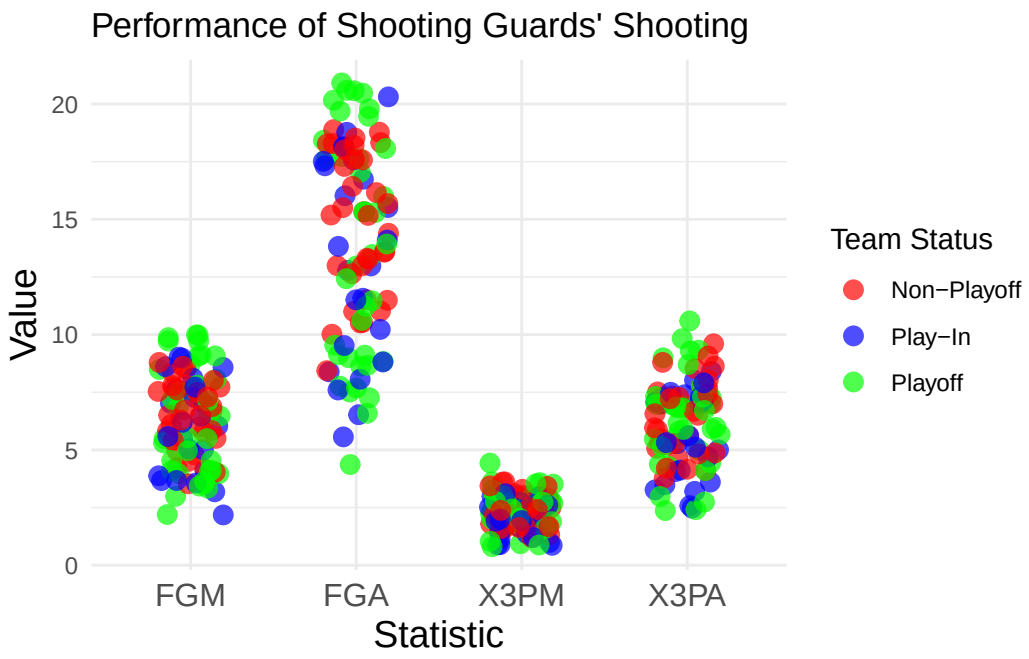
Playoffs vs Play-Ins vs Non-Playoffs


Performance of Centers' Shooting
Value
20
15
10
5
0
FGM
FGA
X3PM
X3PA
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Power Forwards' Shooting
Value
20
15
10
5
0
FGM
FGA
X3PM
X3PA
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Small Forwards' Shooting
Value
20
15
10
5
0
FGM
FGA
X3PM
X3PA
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Point Guards' Shooting
Value
20
15
10
5
0
FGM
FGA
X3PM
X3PA
Statistic
Team Status
Non−Playoff
Play−In
Playoff



Conclusion

It seems like my second prediction was more accurate. Based on the dot plot, we can see that in both sections (FGM and FGA), there are more greens and blues at the top. However, when it comes to three pointers, there was not really a big difference between playoffs and non-playoffs teams. This might be due to the fact that some teams have players who are just better long range shooters or are known for their long range shot, like Stephen Curry, and some teams aren't.

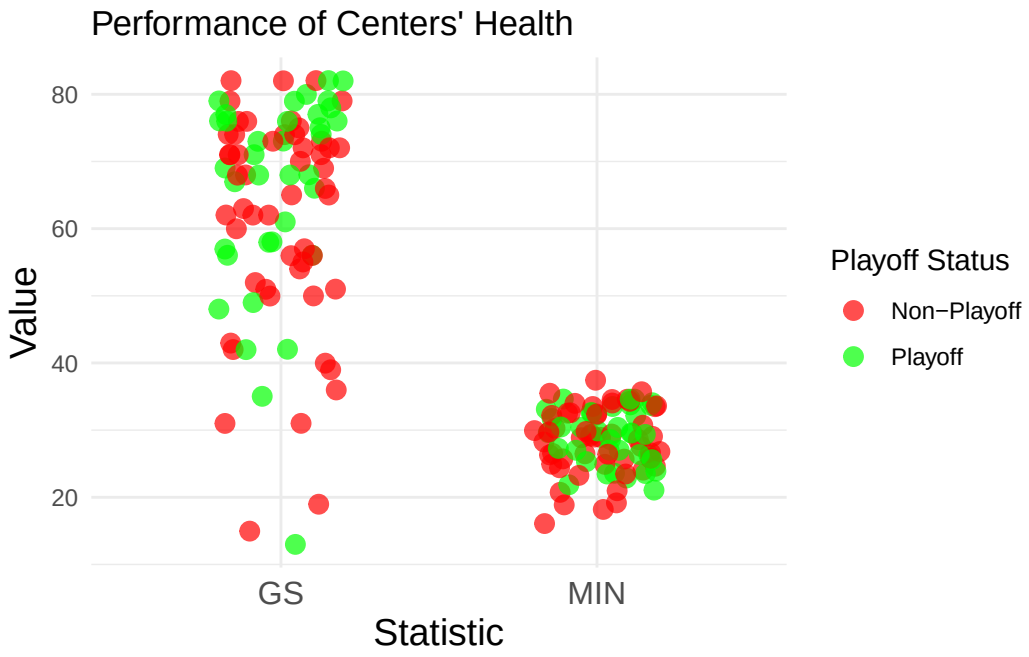
Health

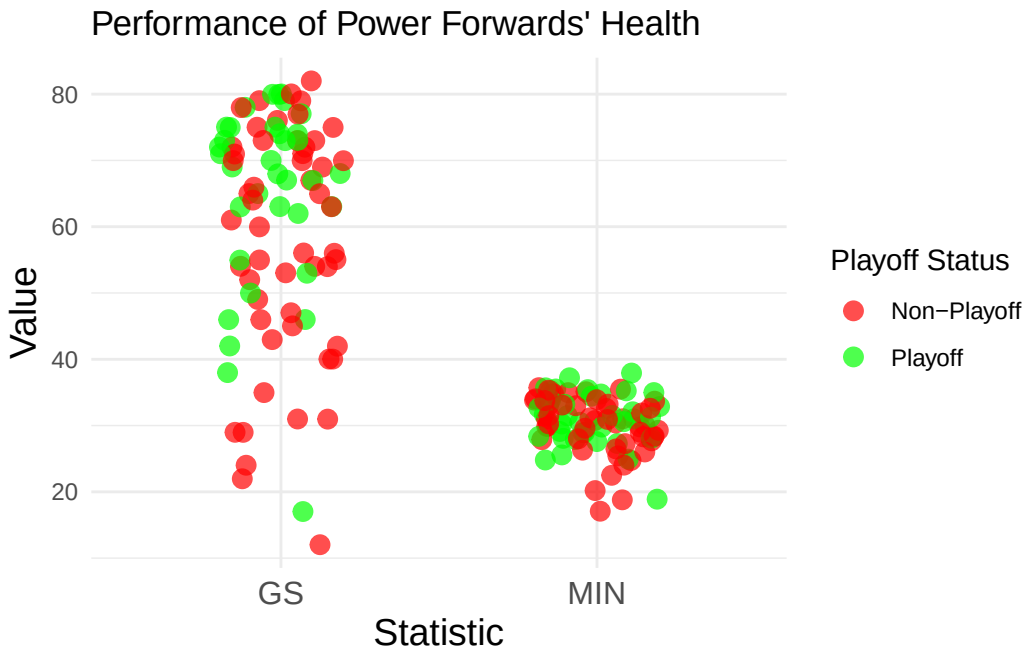
Now, when drafting or trading for a player, it is very important to assess their health. Are they injury prone? Have they missed numerous games due to illnesses or other personal issues? Usually, when a player is good enough to be a starter, they will stay a starter whenever they are in good condition. When players are considered "active" for the game, but start on the bench, it means that there is something wrong with the player. Therefore, we will look at players' games started (GS) and minutes played (MIN) in order to see how much they will be able to contribute to the team during the season.

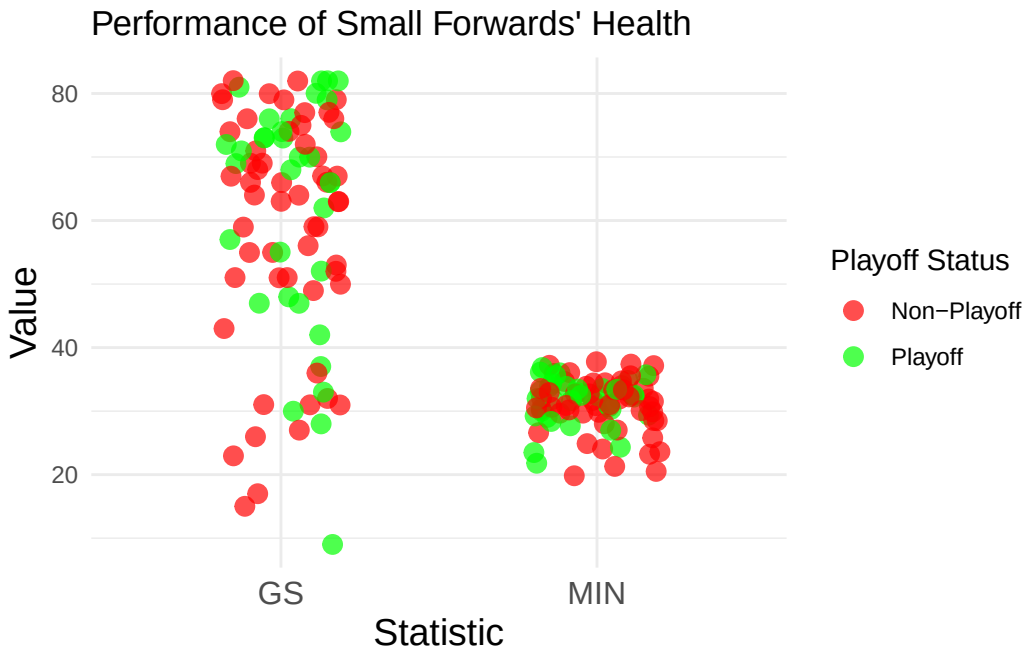
Playoff vs Non-Playoffs

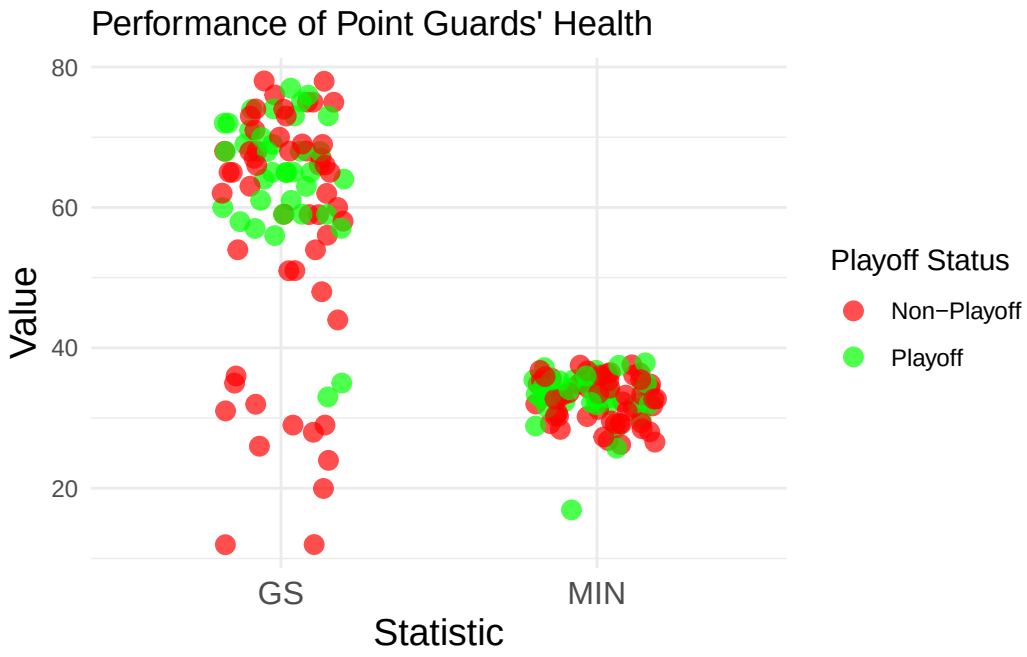
First, we will look at the just playoff teams vs non-playoff teams with play-in qualifying teams being considered non-playoff teams like before.

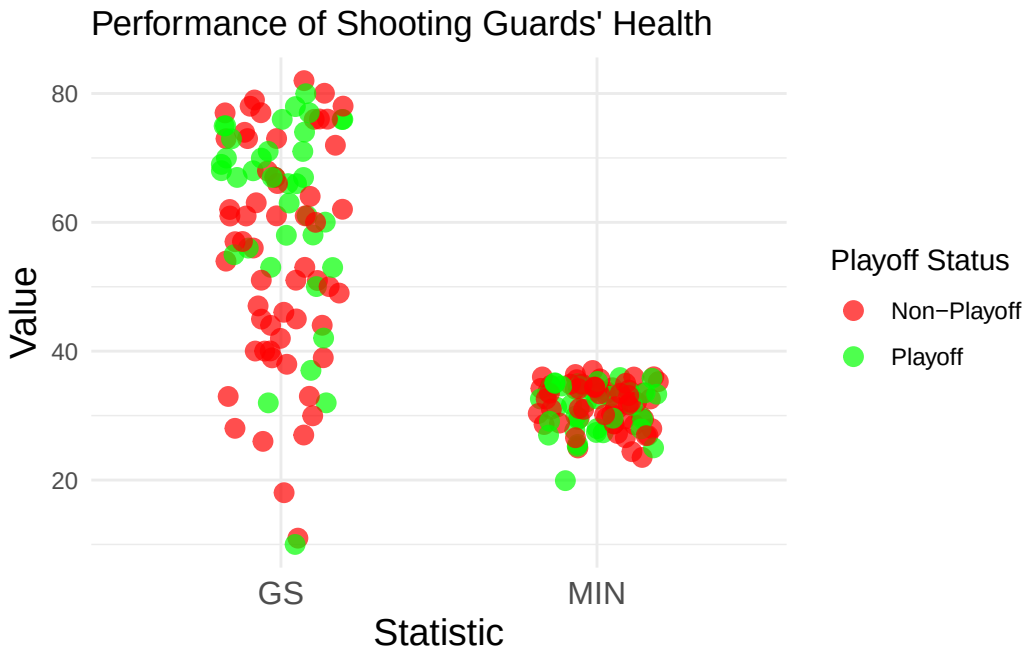
My prediction is that "better" teams will have players who are not injury prone (GS) and have better stamina (MIN).




Performance of Power Forwards' Health
Value
80
60
40
20
GS
MIN
Statistic
Playoff Status
Non−Playoff
Playoff

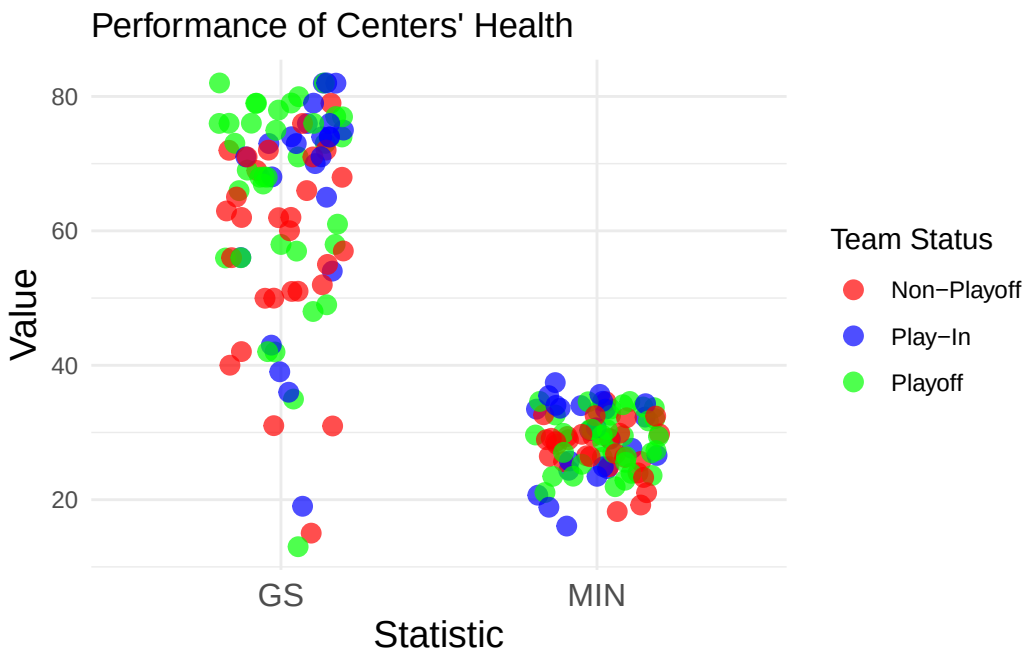

Performance of Small Forwards' Health
Value
80
60
40
20
GS
MIN
Statistic
Playoff Status
Non−Playoff
Playoff

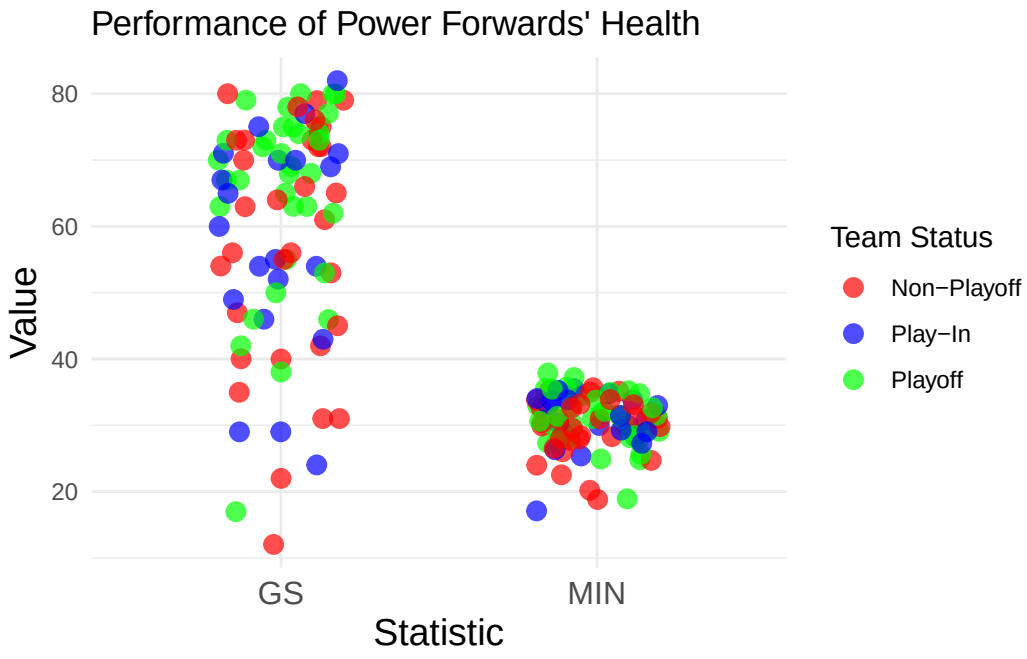


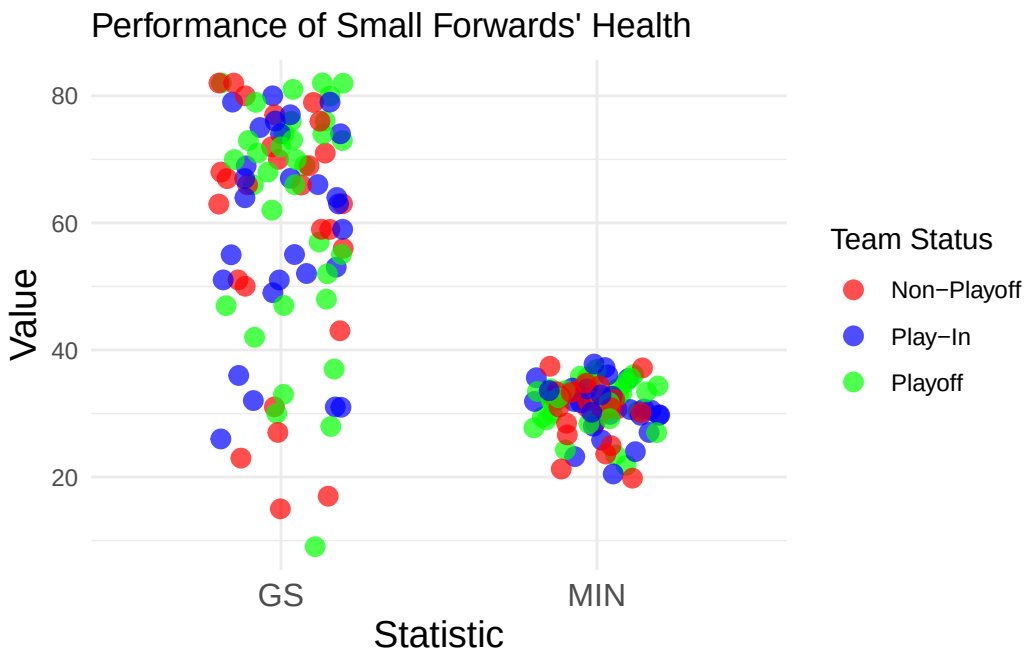


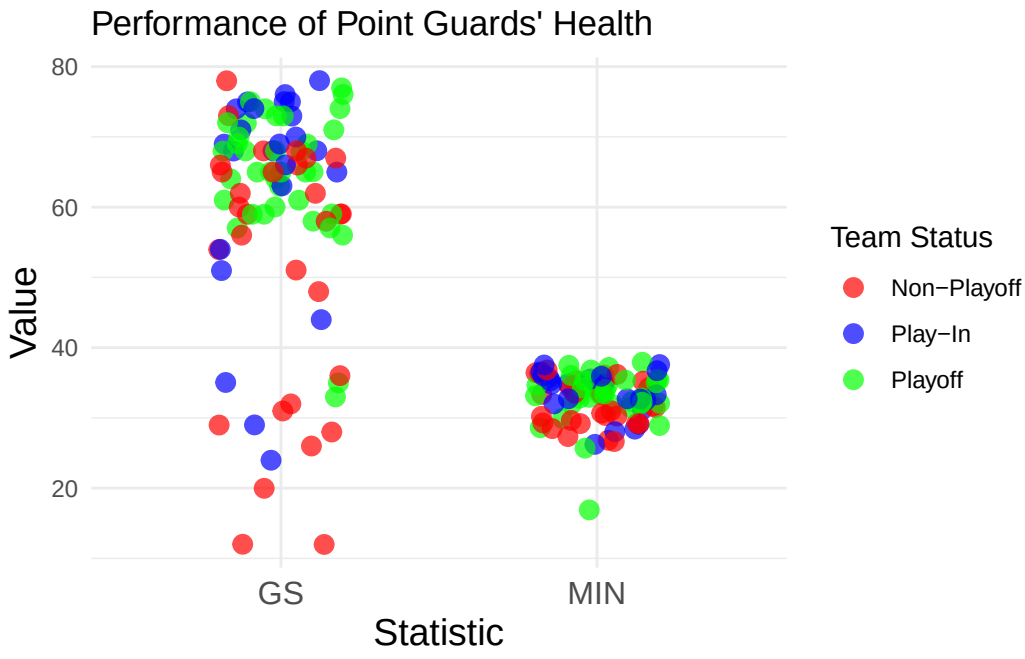
Playoffs vs Play-Ins vs Non-Playoffs

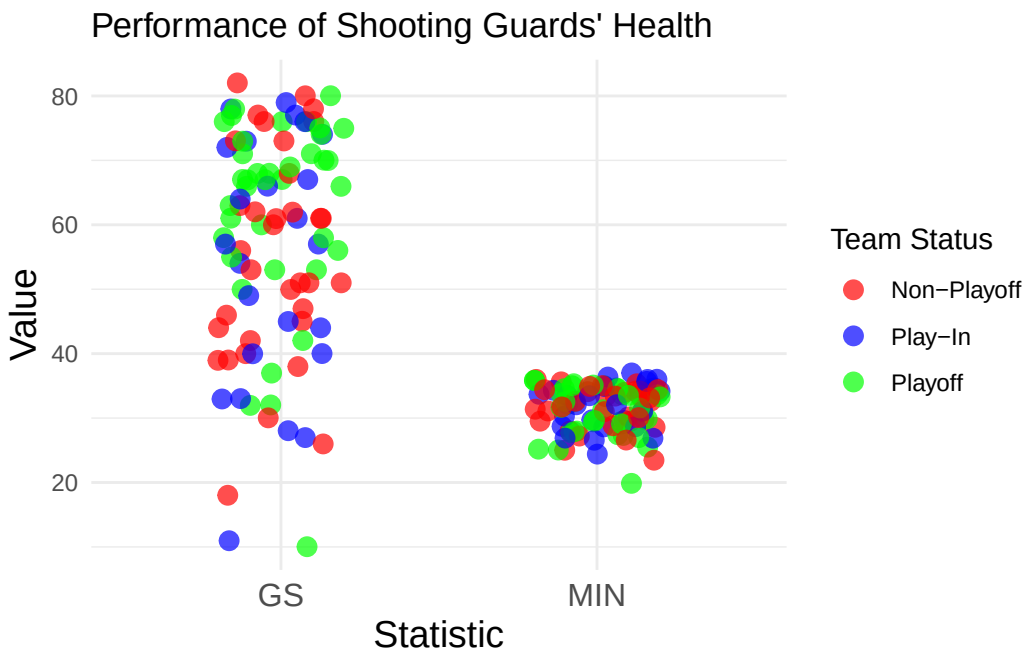
Now, let us look at when play-in teams are differentiated.


Performance of Centers' Health
Value
80
60
40
20
GS
MIN
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Power Forwards' Health
Value
80
60
40
20
GS
MIN
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Small Forwards' Health
Value
80
60
40
20
GS
MIN
Statistic
Team Status
Non−Playoff
Play−In
Playoff


Performance of Point Guards' Health
Value
80
60
40
20
GS
MIN
Statistic
Team Status
Non−Playoff
Play−In
Playoff



Conclusion

It's clear in both dot plots that in the amount of games started, playoff teams had players who started most, if not all 82, games of the season. Other than some outliers, my initial prediction was correct. However, for the number of minutes a player plays during the game, it seems as if there was no correlation between playoff teams and non-playoff teams. The reason for this could be the fact that players on both "good" and "bad" teams are still playing all 48 minutes of the game.

Modeling

We now have an idea of what teams have to do in order to secure a playoff spot. However, just eyeballing our data is not good enough for scouts to use and pick out specific players to draft or trade, especially with all the scattered range in our dot plots. Therefore, we will do three different models: the Null Model, Linear Model, and Multilinear Model. Within each of these models, we will do: Complete Pooling, No Pooling, and Partial Pooling.

As for the predictors from the EDA, we have decided to only test: GS, PPG, RPG, and APG. We already established that MIN, 3PM, and 3PA are not really important and for FGM and FGA, we decided not to use this since more shots and more field goals will directly translate to PPG, RPG, and APG.

The Null Model

Games Started

Complete Pooling

Playoff Teams

```
Call:
lm(formula = GS ~ 1, data = playoff_data)

Residuals:
    Min      1Q  Median      3Q     Max
-54.911  -5.911   4.089  10.089  18.089

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   63.911      1.093   58.48   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.66 on 179 degrees of freedom
```

Play-In Teams

```
Call:
lm(formula = GS ~ 1, data = playin_data)

Residuals:
    Min      1Q  Median      3Q     Max
-49.099  -9.099   5.901  13.901  21.901

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    60.10       1.55   38.77   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17.05 on 120 degrees of freedom
```

Non-Playoff Teams

```
Call:
lm(formula = GS ~ 1, data = non_playoff_data)

Residuals:
    Min      1Q  Median      3Q     Max
-43.967 -10.967   4.033  13.283  26.033

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   55.967      1.438   38.93   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 17.72 on 151 degrees of freedom
```

No Pooling

Playoff Teams

```
# A tibble: 5 x 5
  Position mean_gs std_err t_value  p_value
  <chr>      <dbl>   <dbl>   <dbl>    <dbl>
1 C           65.7    2.56    25.7 2.78e-24
2 PF          65.2    2.38    27.3 9.97e-25
3 PG          64.5    1.56    41.5 5.50e-32
4 SF          61.9    3.18    19.5 1.13e-19
5 SG          62.3    2.46    25.3 5.57e-25
```

Play-In Teams

```
# A tibble: 5 x 5
  Position mean_gs std_err t_value  p_value
  <chr>      <dbl>   <dbl>   <dbl>    <dbl>
1 C           66.1    3.37    19.6 7.24e-16
2 PF          57.7    3.61    16.0 7.35e-13
3 PG          63.1    3.15    20.0 4.80e-16
4 SF          59.1    3.08    19.2 2.85e-17
5 SG          54.4    3.93    13.8 1.20e-12
```

Non-Playoff Teams

```
# A tibble: 5 x 5
  Position mean_gs std_err t_value  p_value
  <chr>      <dbl>   <dbl>   <dbl>    <dbl>
1 C           57.6    2.78    20.8 1.54e-18
2 PF          57.4    3.24    17.7 8.85e-18
3 PG          51.2    3.45    14.8 4.47e-15
4 SF          58.6    3.87    15.1 2.11e-14
5 SG          55.3    2.81    19.7 8.15e-20
```

Partial Pooling

Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: GS ~ 1 + (1 | Position)
   Data: playoff_data

REML criterion at convergence: 1474.5

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.7452 -0.4032  0.2789  0.6881  1.2338

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept)   0       0.00
 Residual             215      14.66
Number of obs: 180, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   63.911      1.093   58.48
optimizer (nloptwrap) convergence code: 0 (OK)
boundary (singular) fit: see help('isSingular')
```

Play-In Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: GS ~ 1 + (1 | Position)
```

```
   Data: playin_data

REML criterion at convergence: 1025.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.7684 -0.5158  0.3791  0.7491  1.3577

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept)   9.023   3.004
 Residual             283.539  16.839
Number of obs: 121, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   60.091      2.039   29.47
```

Non-Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: GS ~ 1 + (1 | Position)
   Data: non_playoff_data

REML criterion at convergence: 1301.8

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.4806 -0.6188  0.2275  0.7494  1.4687

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept)   0.0     0.00
 Residual             314.2    17.72
Number of obs: 152, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   55.967      1.438   38.93
optimizer (nloptwrap) convergence code: 0 (OK)
boundary (singular) fit: see help('isSingular')
```

Points Per Game

Complete Pooling

Playoff Teams

```
Call:
lm(formula = PPG ~ 1, data = playoff_data)

Residuals:
     Min       1Q   Median       3Q      Max
-13.7111  -5.9111  -0.7111   5.3889  16.6889

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  17.2111     0.5177   33.24   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.946 on 179 degrees of freedom
```

Play-In Teams

```
Call:
lm(formula = PPG ~ 1, data = playin_data)

Residuals:
     Min       1Q   Median       3Q      Max
-12.5157  -5.0157  -0.8157   4.0843  17.6843

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  17.0157     0.5759   29.55   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.335 on 120 degrees of freedom
```

Non-Playoff Teams

```
Call:
lm(formula = PPG ~ 1, data = non_playoff_data)

Residuals:
     Min       1Q   Median       3Q      Max
-11.6618  -4.1618  -0.7618   4.0382  16.4382

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  15.9618     0.4338    36.8   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.348 on 151 degrees of freedom
```

No Pooling

Playoff Teams

```
# A tibble: 5 x 5
  Position mean_ppg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            14.8    1.15    12.9 7.15e-15
2 PF           16.7    1.28    13.1 7.84e-15
3 PG           19.8    1.13    17.5 3.75e-19
4 SF           17.1    1.05    16.3 2.56e-17
5 SG           17.6    1.08    16.3 1.70e-18
```

Play-In Teams

```
# A tibble: 5 x 5
  Position mean_ppg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            16.1    1.47    10.9 1.36e-10
2 PF           16.1    1.35    11.9 1.53e-10
3 PG           20.3    1.28    15.9 7.01e-14
4 SF           16.1    1.09    14.8 1.83e-14
5 SG           16.5    1.15    14.4 5.73e-13
```

Non-Playoff Teams

```
# A tibble: 5 x 5
  Position mean_ppg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            13.8   0.879    15.8 1.87e-15
2 PF           14.2   0.896    15.9 2.00e-16
3 PG           17.4   1.06     16.4 3.06e-16
4 SF           15.9   1.17     13.6 2.38e-13
5 SG           18.2   0.671    27.1 3.86e-24
```

Partial Pooling

Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: PPG ~ 1 + (1 | Position)
   Data: playoff_data

REML criterion at convergence: 1204.8

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.06971 -0.89628 -0.07599  0.78056  2.53686

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept)  1.925   1.388
 Residual             46.686   6.833
Number of obs: 180, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)  17.1970     0.8029   21.42
```

Play-In Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: PPG ~ 1 + (1 | Position)
   Data: playin_data

REML criterion at convergence: 787.1

Scaled residuals:
```

```
     Min       1Q   Median       3Q      Max
-1.93830 -0.73335 -0.09513  0.67603  2.91363

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept)  1.731   1.316
 Residual             38.742   6.224
Number of obs: 121, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)  17.0226     0.8172   20.83
```

Non-Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: PPG ~ 1 + (1 | Position)
   Data: non_playoff_data

REML criterion at convergence: 932.7

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.2627 -0.7224 -0.1032  0.7092  2.9922

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept)  2.868   1.694
 Residual             26.231   5.122
Number of obs: 152, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)  15.9213     0.8643   18.42
```

Rebounds Per Game

Complete Pooling

Playoff Teams

```
Call:
lm(formula = RPG ~ 1, data = playoff_data)

Residuals:
    Min      1Q  Median      3Q     Max
-4.3633 -1.8633 -0.7633  1.2367  8.7367

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.9633     0.1997   29.85   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.68 on 179 degrees of freedom
```

Play-In Teams

```
Call:
lm(formula = RPG ~ 1, data = playin_data)

Residuals:
    Min      1Q  Median      3Q     Max
-3.9413 -1.9413 -0.8413  1.8587  7.8587

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.8413     0.2437   23.97   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.681 on 120 degrees of freedom
```

Non-Playoff Teams

```
Call:
lm(formula = RPG ~ 1, data = non_playoff_data)

Residuals:
```

```
    Min      1Q  Median      3Q     Max
-3.5776 -1.8776 -0.6776  1.8224  6.4224

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.6776     0.1897   29.93   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.339 on 151 degrees of freedom
```

No Pooling

Playoff Teams

```
# A tibble: 5 x 5
  Position mean_rpg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            9.38   0.414    22.6 1.74e-22
2 PF           6.81   0.386    17.6 1.12e-18
3 PG           4.48   0.257    17.4 4.18e-19
4 SF           5.17   0.253    20.4 2.77e-20
5 SG           4.10   0.192    21.4 1.91e-22
```

Play-In Teams

```
# A tibble: 5 x 5
  Position mean_rpg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            9.82   0.476    20.6 2.42e-16
2 PF           6.1    0.390    15.6 1.11e-12
3 PG           4.11   0.272    15.1 1.97e-13
4 SF           4.57   0.279    16.4 1.43e-15
5 SG           4.85   0.304    16.0 6.24e-14
```

Non-Playoff Teams

```
# A tibble: 5 x 5
  Position mean_rpg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            8.46   0.356    23.8 4.23e-20
```

```
2 PF           6.55   0.335    19.5 5.70e-19
3 PG           4.48   0.290    15.5 1.54e-15
4 SF           5.27   0.356    14.8 3.49e-14
5 SG           3.86   0.191    20.2 3.62e-20
```

Partial Pooling

Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: RPG ~ 1 + (1 | Position)
   Data: playoff_data

REML criterion at convergence: 751

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.79908 -0.58552 -0.03962  0.47916  2.89797

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 4.576    2.139
 Residual             3.462    1.861
Number of obs: 180, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   5.9877     0.9667   6.194
```

Play-In Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: RPG ~ 1 + (1 | Position)
   Data: playin_data

REML criterion at convergence: 489.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.2370 -0.6716  0.0017  0.5441  2.4274

Random effects:
```

```
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 5.264    2.294
 Residual             2.925    1.710
Number of obs: 121, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)    5.889      1.038   5.674
```

Non-Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: RPG ~ 1 + (1 | Position)
   Data: non_playoff_data

REML criterion at convergence: 605.5

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.5660 -0.6152 -0.1167  0.5313  2.5444

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 3.254    1.804
 Residual             2.841    1.686
Number of obs: 152, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   5.7237     0.8183   6.994
```

Assists Per Game

Complete Pooling

Playoff Teams

```
Call:
lm(formula = APG ~ 1, data = playoff_data)
```

```
Residuals:
    Min      1Q  Median      3Q     Max
-3.3772 -1.9022 -0.3772  1.4228  7.0228

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.8772     0.1724   22.49   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.313 on 179 degrees of freedom
```

Play-In Teams

```
Call:
lm(formula = APG ~ 1, data = playin_data)

Residuals:
    Min      1Q  Median      3Q     Max
-3.0008 -2.0008 -0.4008  1.5992  6.8992

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.9008     0.2106   18.52   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.317 on 120 degrees of freedom
```

Non-Playoff Teams

```
Call:
lm(formula = APG ~ 1, data = non_playoff_data)

Residuals:
    Min      1Q  Median      3Q     Max
-3.1092 -1.6342 -0.5092  1.5908  6.7908

Coefficients:
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.6092     0.1663    21.7   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.051 on 151 degrees of freedom
```

No Pooling

Playoff Teams

```
# A tibble: 5 x 5
  Position mean_apg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            2.73   0.382    7.15 2.45e- 8
2 PF           3.45   0.306   11.3  4.83e-13
3 PG           6.68   0.329   20.3  2.53e-21
4 SF           3.04   0.246   12.3  6.58e-14
5 SG           3.38   0.228   14.8  3.72e-17
```

Play-In Teams

```
# A tibble: 5 x 5
  Position mean_apg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            2.78   0.364    7.63 9.49e- 8
2 PF           3.23   0.437    7.39 3.86e- 7
3 PG           6.45   0.424   15.2  1.76e-13
4 SF           2.88   0.328    8.76 2.25e- 9
5 SG           4.27   0.386   11.1  1.10e-10
```

Non-Playoff Teams

```
# A tibble: 5 x 5
  Position mean_apg std_err t_value  p_value
  <chr>       <dbl>   <dbl>   <dbl>    <dbl>
1 C            2.29   0.240    9.54 2.69e-10
2 PF           2.31   0.200   11.6  8.73e-13
3 PG           6.45   0.259   24.9  3.90e-21
4 SF           2.69   0.213   12.6  1.39e-12
5 SG           4.18   0.262   15.9  4.59e-17
```

Partial Pooling

Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: APG ~ 1 + (1 | Position)
   Data: playoff_data

REML criterion at convergence: 741

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.9020 -0.7527 -0.2077  0.5791  3.8599

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 2.488    1.577
 Residual             3.314    1.820
Number of obs: 180, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   3.8563     0.7184   5.368
```

Play-In Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: APG ~ 1 + (1 | Position)
   Data: playin_data

REML criterion at convergence: 509.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.9459 -0.7564 -0.2347  0.6416  3.1449

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 2.199    1.483
 Residual             3.586    1.894
Number of obs: 121, groups:  Position, 5
```

```
Fixed effects:
            Estimate Std. Error t value
(Intercept)   3.9175     0.6853   5.717
```

Non-Playoff Teams

```
Linear mixed model fit by REML ['lmerMod']
Formula: APG ~ 1 + (1 | Position)
   Data: non_playoff_data

REML criterion at convergence: 532

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.03495 -0.78500 -0.06839  0.62289  3.06616

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 3.117    1.766
 Residual             1.725    1.313
Number of obs: 152, groups:  Position, 5

Fixed effects:
            Estimate Std. Error t value
(Intercept)   3.5845     0.7968   4.499
```

The Linear Model

Complete Pooling

Playoff Teams

```
Call:
glm(formula = Playoff_Prob ~ GS + PPG + RPG + APG, family = binomial,
    data = nba_data)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.597662   0.487366   1.226   0.2201
GS           0.014001   0.006780   2.065   0.0389 *
```

```
PPG          0.002854   0.023055   0.124   0.9015
RPG         -0.021522   0.046566  -0.462   0.6440
APG         -0.016067   0.063483  -0.253   0.8002
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 367.78  on 452  degrees of freedom
Residual deviance: 363.43  on 448  degrees of freedom
AIC: 576.92

Number of Fisher Scoring iterations: 4
```

No Pooling

Playoff Teams

```
# A tibble: 5 x 6
  Position Intercept  GS_coef PPG_coef RPG_coef APG_coef
  <chr>        <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
1 C            1.07   0.0234  -0.00243  -0.154    0.0891
2 PF           1.17   0.00898 -0.101    -0.0106   0.471
3 PG          -0.215  0.0334   0.0234   -0.0181  -0.127
4 SF           1.12  -0.00392  0.0928   -0.233    0.0213
5 SG           0.224  0.0104  -0.0578    0.866   -0.480
```

Partial Pooling

Playoff Teams

```
Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: Playoff_Prob ~ GS + PPG + RPG + APG + (1 | Position)
   Data: nba_data

     AIC      BIC   logLik deviance df.resid
   561.0    585.7   -274.5    549.0      447

Scaled residuals:
    Min      1Q  Median      3Q     Max
```

```
-1.8529 -0.3877  0.6020  0.6662  0.8940

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 0        0
Number of obs: 453, groups:  Position, 5

Fixed effects:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.314758   0.448338   0.702   0.4826
GS           0.012277   0.006217   1.975   0.0483 *
PPG          0.009169   0.020667   0.444   0.6573
RPG         -0.031408   0.041613  -0.755   0.4504
APG         -0.044555   0.056590  -0.787   0.4311
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
    (Intr) GS     PPG    RPG
GS  -0.620
PPG -0.263 -0.131
RPG -0.342 -0.149 -0.218
APG -0.106 -0.036 -0.580  0.180
optimizer (Nelder_Mead) convergence code: 0 (OK)
boundary (singular) fit: see help('isSingular')
```

The Multilinear Model

Complete Pooling

```
Call:
lm(formula = Playoff_Prob ~ GS + PPG + RPG + APG + PPG:RPG, data = nba_data)

Residuals:
    Min      1Q  Median      3Q     Max
-0.8697 -0.2552  0.1777  0.2319  0.4192

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.9330929  0.1236742   7.545 2.56e-13 ***
GS           0.0027561  0.0010317   2.671  0.00783 **
```

```
PPG         -0.0170693  0.0070862  -2.409  0.01641 *
RPG         -0.0513466  0.0182734  -2.810  0.00517 **
APG         -0.0023775  0.0092597  -0.257  0.79748
PPG:RPG      0.0027778  0.0009876   2.813  0.00513 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3541 on 447 degrees of freedom
Multiple R-squared:  0.03037,   Adjusted R-squared:  0.01953
F-statistic:   2.8 on 5 and 447 DF,  p-value: 0.0167
```

No Pooling

```
# A tibble: 5 x 8
  Position Intercept   GS_coef PPG_coef RPG_coef APG_coef PPG_RPG_coef R_squared
  <chr>        <dbl>     <dbl>    <dbl>    <dbl>    <dbl>        <dbl>     <dbl>
1 C             1.29   0.00521  -0.0460   -0.0933 -0.00215      0.00491    0.0689
2 PF            1.47   0.00236  -0.0599   -0.118   0.0526       0.00670    0.122
3 PG            0.690  0.00644  -0.00794  -0.0594 -0.0174       0.00255    0.0924
4 SF            1.04  -0.000190 -0.00468  -0.0989  0.00385      0.00354    0.0776
5 SG            1.48   0.00115  -0.0553   -0.0959 -0.0704       0.0117     0.176
```

Partial Pooling

```
Linear mixed model fit by REML ['lmerMod']
Formula: Playoff_Prob ~ GS + PPG + RPG + APG + PPG:RPG + (1 | Position)
   Data: nba_data

REML criterion at convergence: 395

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.4562 -0.7206  0.5020  0.6549  1.1838

Random effects:
 Groups   Name        Variance Std.Dev.
 Position (Intercept) 0.0000   0.0000
 Residual             0.1254   0.3541
Number of obs: 453, groups:  Position, 5

Fixed effects:
```

```
              Estimate Std. Error t value
(Intercept)  0.9330929  0.1236742   7.545
GS           0.0027561  0.0010317   2.671
PPG         -0.0170693  0.0070862  -2.409
RPG         -0.0513466  0.0182734  -2.810
APG         -0.0023775  0.0092597  -0.257
PPG:RPG      0.0027778  0.0009876   2.813

Correlation of Fixed Effects:
        (Intr) GS     PPG    RPG    APG
GS      -0.316
PPG     -0.774 -0.143
RPG     -0.814 -0.136  0.781
APG     -0.058 -0.022 -0.282  0.057
PPG:RPG  0.796  0.093 -0.881 -0.929  0.009
optimizer (nloptwrap) convergence code: 0 (OK)
boundary (singular) fit: see help('isSingular')
```

Conclusion

Now, since we created and took a look at all these different models, we can analyze our results to reach our "magic" number of 43 wins.

From the null model, we were able to determine the overall average GS, PPG, RPG, and APG of players and the average of players based on their position in the NBA. With the p-value being statistically significant, we were able to use weighted average calculations to determine the average threshold players needs to create a successful playoff team.

Since only 50% of play-in teams make it to the playoffs, we said half of their performance contributes to the playoff picture; therefore, using this formula:

Weighted Statistic for Playoffs = (Playoff Statistic) + 0.5 * (Play-In Statistic - Playoff Statistic)

We will be using the No Pooling Null Model to determine the average for each position and will be testing both Linear and Multilinear Model in order to calculate the chances the average playoff qualifying player had to make the playoffs.

Center

Null Model

- Games Played: 65.88

- Points Per Game: 15.47
- Rebounds Per Game: 9.60
- Assists Per Game: 2.75

Linear Model

With this model, the team with this center will have a 79.1% chance to make the playoffs.

Multilinear Model

With this model, the team with this center will have a 68.0% chance to make the playoffs.

Power Forward

Null Model

- Games Played: 59.44
- Points Per Game: 16.41
- Rebounds Per Game: 6.48
- Assists Per Game: 3.34

Linear Model

With this model, the team with this center will have a 82.5% chance to make the playoffs.

Multilinear Model

With this model, the team with this power forward will have a 68.0% chance to make the playoffs.

Small Forward

Null Model

- Games Played: 60.49
- Points Per Game: 17.56
- Rebounds Per Game: 5.161
- Assists Per Game: 2.96

Linear Model

With this model, the team with this center will have a 79.8% chance to make the playoffs.

Multilinear Model

With this model, the team with this power forward will have a 68.3% chance to make the playoffs.

Point Guard

Null Model

- Games Played: 64.81
- Points Per Game: 20.0
- Rebounds Per Game: 4.29
- Assists Per Game: 6.56

Linear Model

With this model, the team with this center will have a 81.8% chance to make the playoffs.

Multilinear Model

With this model, the team with this center will have a 69.0% chance to make the playoffs.

Shooting Guard

Null Model

- Games Played: 59.33
- Points Per Game: 17.07
- Rebounds Per Game: 4.47
- Assists Per Game: 4.84

Linear Model

With this model, the team with this center will have a 80.3% chance to make the playoffs.

Multilinear Model

With this model, the team with this center will have a 67.4% chance to make the playoffs.

Error Analysis

We can see that there is a huge difference in playoff percentage between the linear and multilinear model, even with the same statistics. Let's dive into why this might be the case.

When viewing the p-values of the linear model, we can see that all predictors except for GS were statistically insignificant whereas in the multilinear model, all predictors except APG were statistically significant; therefore, it will be better to accept the multilinear model over the linear model. Although, there were a lot of large p-values, we still decided to run the tests to see whether or not the percentages were something out of the ordinary or if they made sense.

References

1* - https://www.nba.com/news/nba-play-in-tournament

2* - https://www.espn.com/nba/standings/_/season/2022

3* - https://www.espn.com/nba/standings/_/season/2023

4* - https://www.espn.com/nba/standings/_/season/2024

5* - https://en.wikipedia.org/wiki/NBA_play-in_tournament?utm_source=chatgpt.com

6* - https://www.nba.com/news/playoff-primer-seeding-matchup-postseason-history-all-16-teams

The websites I used to create my csv files are:

https://www.nbastuffer.com/nba-stats/player/

https://www.statmuse.com/

https://www.basketball-reference.com/teams/