Predicting whether the Los Angeles Lakers will have enough wins to guarantee them a spot in the 2022 – 2023 NBA Playoffs.

Purpose of the Project

The purpose of this project is to predict whether the Los Angeles Lakers will have enough wins to guarantee them a spot in the Playoffs for the 2022 – 2023 season by analyzing the current team's NRtg (net rating). This will allow the team to determine if changes are needed so that they will be guaranteed a spot in the Playoffs. Although this project looks specifically at the Los Angeles Lakers and in the 2022 – 2023 season, this model can be implemented to any of the other 29 teams in the NBA (National Basketball Association) and for future seasons.

About the Project

The NBA can be defined as having 2 distinguished periods throughout a season. The first, is the regular season which consists of 30 teams that play 82 games, and the second, a postseason tournament called the Playoffs to determine the league champion. Prior to the 2019 – 2020 season, teams that finished within the top 8 seeds in their respective conference, the East and the West, played in the Playoffs. However, with the introduction of the play-in tournament[1], the top 6 teams from each conference are now guaranteed a spot in the Playoffs while teams that finish the season between 7th and 10th place compete for the final 2 spots. Therefore, this analysis will look at the statistics of an NBA team to determine whether a team will achieve enough wins to finish the season within the top 6 places from each conference, guaranteeing the team a spot in the Playoffs.

This analysis will look specifically at the Los Angeles Lakers.

Throughout this report, all basketball statistics are related to a team as a whole, not an individual player.

[1] The play-ins, a preliminary tournament that determines the final two playoff seeds in the Eastern Conference and Western Conference, was introduced in the 2020 – 2021 season.

Key Questions

1. How many wins does the Los Angeles Lakers need in the regular season to be guaranteed a spot in the Playoffs?
2. What does it take to win enough games in the regular season to be guaranteed a spot in the Playoffs?
3. Will the current roster of the Los Angeles Lakers achieve enough wins to guarantee them a spot in the Playoffs?
4. What can the Los Angeles Lakers do to improve their regular season record to be guaranteed a spot in the Playoffs?

Key Insights

1. The Los Angeles Lakers should have a goal of winning approximately 54 games by the end of the season to finish within the top 6 places in the Western Conference to be guaranteed a spot in the Playoffs.

2. The Los Angeles Lakers need to have a NRtg of around 5.9 to win approximately 54 games by the end of the season.

3. No, the Los Angeles Lakers will not have enough wins to be guaranteed a spot in the Playoffs for the 2022– 2023 season. The Los Angeles Lakers, with a NRtg of -1.7, is predicted to win 36 games by the end of the 2022 – 2023 NBA season.

4. The Los Angeles Lakers should look to change the rotation of their current roster to reach a more optimal NRtg or bring in key players that will have a positive impact towards increasing the NRtg in order to win more games in the regular season to be guaranteed a spot in the Playoffs.

About the Data

The data contains season statistics for the 30 teams in the NBA (National Basketball Association) from the seasons that represent the Modern Era[2], 2013 – present. The data that represents the 2023 season was used after the statistical model was created to predict the number of wins the Los Angeles Lakers;s current roster will have by the end of the 2022 – 2023 season. From the data sources, I used the tables labeled “Conference Standings,” “Per Game Stats,” and

[2] 2013 to present, a period characterized with more efficient shot selections by rewarding three-point shots compared to the two-point shots (https://content.iospress.com/articles/journal-of-sports-analytics/jsa200525)

"Advanced Stats." I combined the data sources into 3 different tables in Excel that contained information regarding the teams' standings at the end of the season, advanced team statistics, and per game statistics. In R, the 3 tables were combined to create one large dataframe with a total of 58 columns.

Data Sources:
https://www.basketball-reference.com/leagues/NBA_2023.html
https://www.basketball-reference.com/leagues/NBA_2022.html
https://www.basketball-reference.com/leagues/NBA_2021.html
https://www.basketball-reference.com/leagues/NBA_2020.html
https://www.basketball-reference.com/leagues/NBA_2019.html
https://www.basketball-reference.com/leagues/NBA_2018.html
https://www.basketball-reference.com/leagues/NBA_2017.html
https://www.basketball-reference.com/leagues/NBA_2016.html
https://www.basketball-reference.com/leagues/NBA_2015.html
https://www.basketball-reference.com/leagues/NBA_2014.html

Tools Used:
I used R Studio to create the statistical models using these packages:

1. tidyverse (dplyr, ggplot2)
2. ggpubr

In Depth Analysis

1. How many wins does the Los Angeles Lakers need in the regular season to be guaranteed a spot in the Playoffs?

To determine how many wins the Los Angeles Lakers need in the regular season to be guaranteed a spot in the Playoffs, the average percentage of games a team in the Western Conference that finishes in the top 6 places wins was calculated because the Los Angeles Lakers is a part of the Western Conference.

All 30 teams were split into their respective conferences and 2 new data frames that contained all the data from teams that finished within the top 6 in previous seasons were created. The column with the total number of wins wasn't used to determine the average number of games that a team in the top 6 places win because there may be seasons where teams play a varying number of games in the regular season than the standard 82 games, such as the 2019 – 2020 and 2020 – 2021 seasons where the number of games in the regular season were shortened because of COVID-19.

Instead, the percentage of games that teams that finish within the top 6 places win was calculated by dividing the total wins by the total number of games played for each team from the 2013 – 2022 seasons.

$$\text{Win Percentage} = (\text{Total Wins} / \text{Total Games}) \times 100$$

Then, the average of the win percentages from all 54 teams (9 seasons x 6 teams) was calculated.

$$\text{Average Win Percentage} = (\text{Total Win Percentage}/54)$$

Teams that finish within the top 6 places in Western Conference win, on average, 66.4% of their games in the regular season.

Therefore, by multiplying the average win percentage for the Western Conference by the total number of games that will be played in the 2022 – 2023 regular season, 82 games, the total number of wins necessary to finish within the top 6 can be calculated.

$$\text{Total Wins} = \text{Average Western Conference Win Percentage} \times 82$$

$$\text{Total Wins} = 0.664 \times 82$$

$$\text{Total Wins} = 54.448$$

The Los Angeles Lakers should have a goal of winning approximately 54 games by the end of the season to finish within the top 6 places in the Western Conference to be guaranteed a spot in the Playoffs.

2. What does it take to win enough games in the regular season to be guaranteed a spot in the Playoffs?

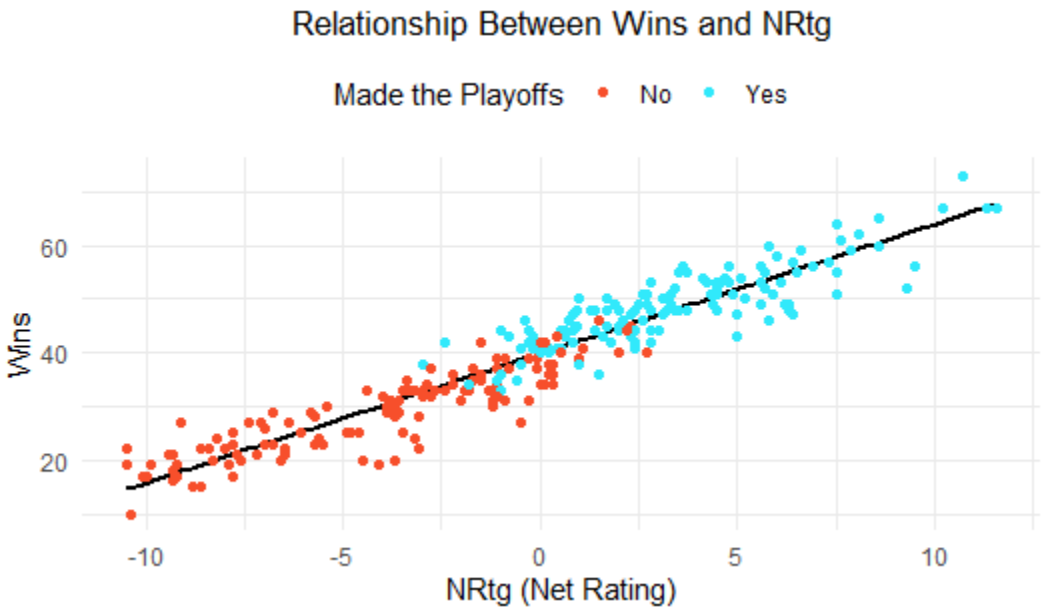
The Los Angeles Lakers need to win about 54 games in the regular season to be guaranteed a spot in the Playoffs. So, how does a basketball team win games?

The basketball team that scores more points than their opponent by the end of regulation wins. Therefore, the basketball statistic that represents the score and also is related to winning was found by performing a Pearson correlation test between columns that represent the total number of wins to 46 other variables that represent the offensive statistics and defensive statistics from all 30 teams in the association, across 9 seasons.

From these calculations, the NRtg (net rating) has the highest correlation with 0.95. The NRtg is the difference in the score per 100 possessions as a combination of 5 players. It will display the quality of play and the quality of the team.

In order to determine the necessary NRtg to win 54 games in the regular season, a linear regression model was created (Figure 1).

Figure 1



Call:
lm(formula = W ~ NRtg, data = df)

Residuals:

Min	1Q	Median	3Q	Max
-11.6338	-2.3582	0.1805	2.6591	9.0763

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	39.83787	0.23705	168.06	<2e-16 ***
NRtg	2.40815	0.04943	48.72	<2e-16 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.895 on 268 degrees of freedom
Multiple R-squared: 0.8986, Adjusted R-squared: 0.8982
F-statistic: 2374 on 1 and 268 DF, p-value: < 2.2e-16

Figure 1 shows a positive relationship between the NRtg and wins, and the linear regression model shows that the results are highly significant. This means that teams with a higher NRtg are more likely to win.

Our regression equation to calculate the number of wins is:

$$Wins = 39.83787 + 2.40815(NRtg)$$

To be guaranteed a spot in the Playoffs, a team needs approximately 55 wins during the regular season. So, the necessary NRtg can be calculated by using the regression model created:

$$54 = 39.83787 + 2.40815(NRtg)$$
$$NRtg = 5.88091$$

The Los Angeles Lakers need to have a NRtg of around 5.9 to win 54 games by the end of the season.

3. Will the current roster of the Los Angeles Lakers achieve enough wins to guarantee them a spot in the Playoffs?

To determine if the current roster of the Los Angeles Lakers will have enough wins to guarantee them a spot in the Playoffs, the NRtg of the current team needs to be calculated.

This is done by taking the difference of the ORtg (offensive rating) from the DRtg (defensive rating):

$$NRtg = ORtg - DRtg$$

From the previous analysis (Figure 1), it was determined that a team with a positive NRtg of 5.9 was needed to win enough games to guarantee them a spot in the Playoffs. So, how does a team achieve a positive NRtg?

A team will need an ORtg that is higher than the DRtg.

To determine which variables have a statistical relationship with the ORtg and the DRtg, two Pearson correlation tests were performed. These 2 correlations compare the ORtg and the DRtg from all 30 teams, across 9 seasons, to 46 different variables that represent the offensive and defensive statistics.

Variables highly correlated to the ORtg:

- TS% has a positive correlation of 0.903844287
    - TS% = a measure of shooting efficiency that takes into account 2-pointers, 3-pointers and free throws
- eFG% has a positive correlation of 0.879423785
    - eFG% = Statistics adjusted for the fact that a 3-point field goal is worth one more point than a 2-point field goal percentage

Variables highly correlated to the DRtg:

- OeFG% has a positive correlation of 0.91730576
    - OeFG% =Opponent's statistics adjusted for the fact that a 3-point field goal is worth one more point than a 2-point field goal percentage
- PA/G% has a positive correlation of 0.86309331
    - PA/G% = Opponent's points per game

ORtg (Offensive Rating) Calculation

First, the variables that correlate with the ORtg were fitted into a multivariate regression model to understand how these variables combined affect the ORtg of a team.

Figure 2

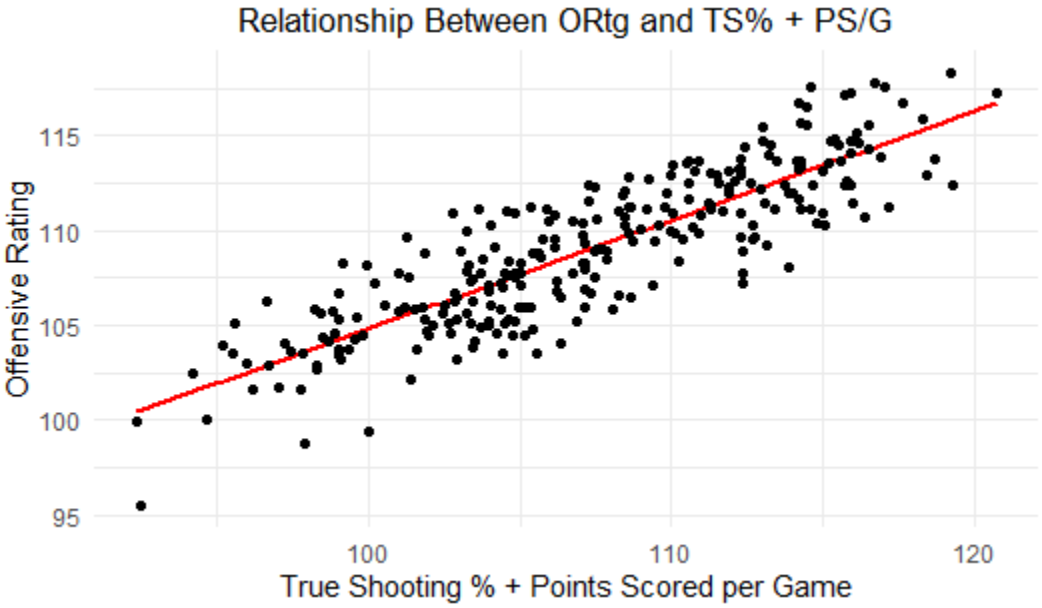
```
Call:
lm(formula = ORtg ~ TS. + eFG., data = df1)

Residuals:
    Min      1Q  Median      3Q     Max
-4.5668 -1.1866 -0.1145  1.1928  5.8777

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   12.390      3.121   3.970 9.25e-05 ***
TS.          195.288     24.274   8.045 2.83e-14 ***
eFG.         -22.341     22.608  -0.988    0.324
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.72 on 267 degrees of freedom
Multiple R-squared:  0.8176,  Adjusted R-squared:  0.8162
F-statistic: 598.4 on 2 and 267 DF,  p-value: < 2.2e-16
```

From the regression model, we see that the eFG% is not significant, meaning that there is no statistical relationship between the eFG% and ORtg.

Instead, another multivariate regression model was created by substituting the eFG% with the PS/G (points scored per game), the next highest correlation to the ORtg.

PS/G has a positive correlation of 0.858321231.

Figure 3



Call:
lm(formula = ORtg ~ TS. + PS.G, data = df1)

Residuals:

Min	1Q	Median	3Q	Max
-4.3440	-1.0509	-0.0255	1.0124	4.5629

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	19.41223	2.56342	7.573	5.98e-13 ***
TS.	117.25455	8.04738	14.571	< 2e-16 ***
PS.G	0.23086	0.02836	8.140	1.51e-14 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.542 on 267 degrees of freedom
Multiple R-squared: 0.8533, Adjusted R-squared: 0.8522
F-statistic: 776.7 on 2 and 267 DF, p-value: < 2.2e-16

From Figure 3, the results show that the “signif. code” (p-value) is between 0 and 0.001. This shows a very high significance between the ORtg, TS%, and PS/G. It also shows a positive relationship between the variables ORtg and the TS% + PS/G.

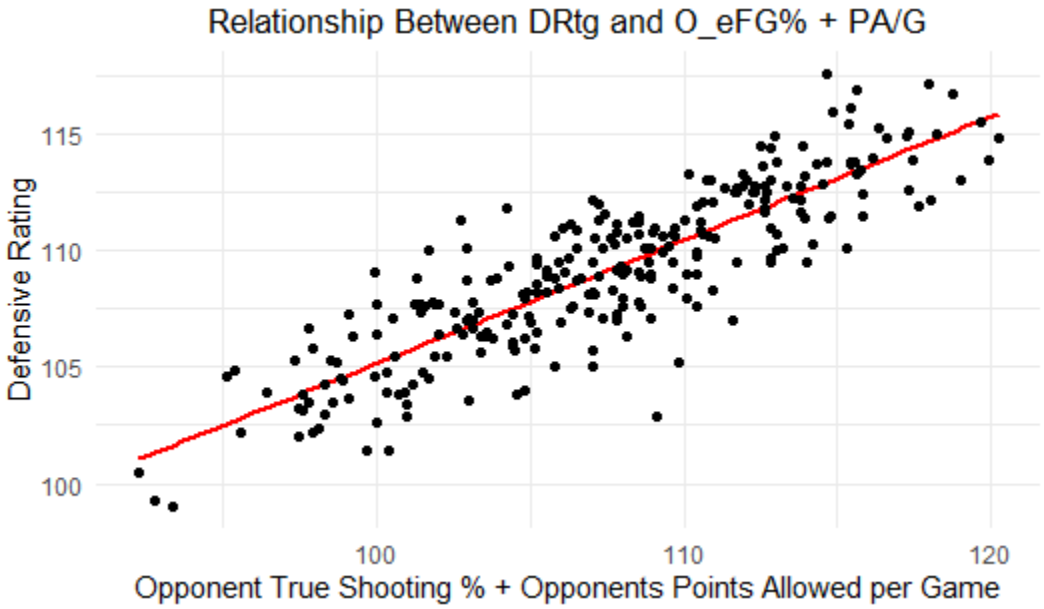
From Figure 3, the Los Angeles Lakers’s ORtg can be calculated:

ORtg = 19.41223 + 117.25455(TS%) + 0.23086(PS/G)

DRtg (Defensive Rating) Calculation

Another multivariate regression model was created to understand the relationship between the O_eFG% and PA/G to the DRtg of a team.

Figure 4



Call:
lm(formula = DRtg ~ O_eFG. + PA.G, data = df1)

Residuals:

Min	1Q	Median	3Q	Max
-3.5615	-0.7996	-0.0682	0.8688	3.1836

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	28.47503	1.95587	14.56	<2e-16 ***
O_eFG.	112.27513	6.46080	17.38	<2e-16 ***
PA.G	0.21032	0.02259	9.31	<2e-16 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.243 on 267 degrees of freedom
Multiple R-squared: 0.8803, Adjusted R-squared: 0.8794
F-statistic: 981.9 on 2 and 267 DF, p-value: < 2.2e-16

From Figure 4, the results show that the “signif. code” (p-value) is between 0 and 0.001. This shows a very high significance between the DRtg, O_eFG%, and PA/G. It also shows a positive relationship between the variables DRtg and the O_eFG% + PA/G.

From Figure 4, the Los Angeles Lakers’s DRtg can be calculated:

$$DRtg = 28.47503 + 112.27513(O\_eFG\%) + 0.21032(PA/G)$$

With Figure 3 and 4, the Los Angeles Lakers’s NRtg can be calculated:

Los Angeles Lakers’s Basketball Statistics as of March 2, 2023:

$TS\% = 0.557$

$PS/G = 116.9$

$O\_eFG\% = 0.536$

$PA/G = 117.9$

$$NRtg = ORtg - DRtg$$

$$ORtg = 19.41223 + 117.25455(TS\%) + 0.23086(PS/G)$$

$$DRtg = 28.47503 + 112.27513(O\_eFG\%) + 0.21032(PA/G)$$

$$NRtg = [19.41223 + 117.25455(TS\%) + 0.23086(PS/G)] - [28.47503 + 112.27513(O\_eFG\%) + 0.21032(PA/G)]$$

$$NRtg = [19.41223 + 117.25455(0.557) + 0.23086(116.9)] - [28.47503 + 112.27513(0.536) + 0.21032(117.9)]$$

$$NRtg = -1.740679$$

With the calculated NRtg, the number of wins a the Los Angeles Lakers will have by the end of the season can be predicted by using Figure 1:

$$Wins = 39.83787 + 2.40815(NRtg)$$

$$Wins = 39.83787 + 2.40815(-1.740679)$$

$$Wins = 35.64605$$

The Los Angeles Lakers, with a NRtg of -1.7, is predicted to win about 36 games by the end of the 2022 – 2023 NBA season. According to the predicted number of wins, the Los Angeles

Lakers will not have enough wins to be guaranteed a spot in the Playoffs for the 2022– 2023 season.

4. What can the Los Angeles Lakers do to improve their regular season record to be guaranteed a spot in the Playoffs?

The Los Angeles Lakers currently have an NRtg of -1.7, which is short from the estimated NRtg of 5.9 needed to win approximately 54 games in the regular season in order to be guaranteed a spot in the Playoffs. Since the NRtg equals the ORtg subtracted by the DRtg, the Los Angeles Lakers should look to find methods to increase their ORtg and lower their DRtg. Possible recommendations are to change the rotation of their current roster to reach a more optimal NRtg or bring in key players that will have a positive impact towards increasing the NRtg in order to win more games in the regular season to be guaranteed a spot in the Playoffs.