Basketball Bros
By: Josh Knull, Jitesh Motati, and Richard Wu
CGT 270 Fall 2019

Introduction

The topic we decide to explore relates to how various statistics used to measure performance in the NBA correlates with team success. The intended audience includes basketball enthusiasts and sports analysts as they can utilize the findings to draw conclusions in a more accurate manner. Prior work has been done on a topic similar to ours by Yuanhao Yang who studied the correlation between individual player's performance and their team's success level. The variables used in Yang's study revolves around normal statistics such as points, rebounds, and assists. Together as a group, we believe that these statistics do not accurately represent the value of individual players and their impact on the game. Therefore, we think our work is necessary as it provides more insight on team success by analyzing in-depth performance-related variables such as efficiency ratings and field goal percentages. The result of this project is significant to its intended audience because it provides insight on effective team compositions, assists analysts to provide more accurate analysis of teams, and helps organizations to trade players and formulate their teams in an efficient way.

Background

The variables that are used in this study includes efficiency ratings, shooting percentages, and r-squared values. The efficiency ratings are divided into three categories: offensive, defensive, and net efficiency. The formula for each is listed below.

Offensive Efficiency Rating = 100*Points Scored/Possessions(self)
Defensive Efficiency Rating = 100*Points Allowed/Possessions(opponent)
Net Efficiency Rating =
Offensive Efficiency Rating - Defensive Efficiency Rating

*Possession is defined as Field goal attempts - Offensive rebounds + Turnovers + .475*Free throw attempts

Shooting percentages are divided into three-point percentages, field goal percentages, and free-throw percentages. These variables measure the accuracy of each type of shot which can be calculated by
Number of shots made/Number of shots attempted.

R-squared values are statistical measures that describe the strength of a correlation between two variables. R-squared values ranges from 0 to 1 with 0 indicating no correlation and 1 indicating very strong correlation.

Figure 1

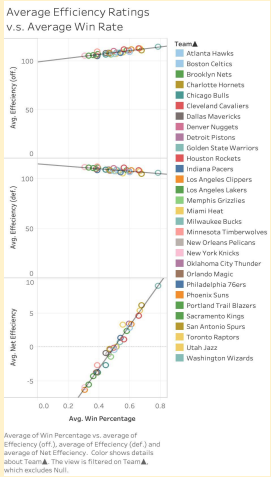
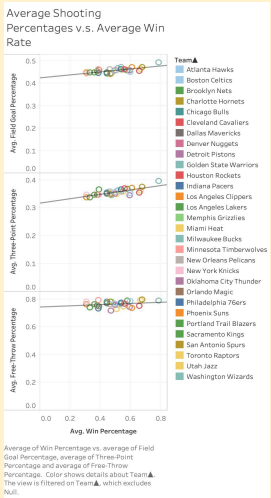


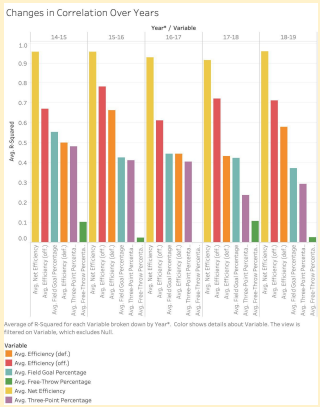
Figure 2



1. How do different stats correlate with team success? Which variable has the strongest correlation? Which has the weakest?

Figures 1 and 2 collectively addresses the question by analyzing two different sets of variables: efficient ratings and shooting percentages. Both visualizations are created using the average of all the values for each variable across five consecutive NBA regular seasons between the years of 2014 and 2019. The x-axis of each graph measures the average win ratio while the y-axis displays the average values for a specific variable. The data for each team is plotted in a different color and a line of best fit is created for each scatter plot using the functionalities provided by Tableau. These lines can be used to analyze the strength and direction of the correlation for a specific variable. A negative slope indicates that the variable negatively correlates with a team's success while a positive slope indicates the opposite. The steepness of the slope represents the strength of the correlation. A slope close to 1 indicates a strong correlation and a slope close to 0 indicates a weak correlation. As the two figures suggests, average net efficiency has the strongest correlation with team success due to its nearly-perfect diagonal slope. In contrast, the slope for average free-throw percentage reveals that it correlates the weakest with the successfulness of a team.

Figure 3



2. How do the correlations change over time?

Figure 3 addresses the question by comparing the r-squared values for each variable during different regular seasons. These r-squared values are calculated by dividing each scatter plot in Figures 1 and 2 into five separate ones with each representing a particular NBA season and analyzing the line of best fit descriptions generated by Tableau. The x-axis of Figure 3 displays the variables for each season while the y-axis provides a scale for the average r-squared values. The bars are color coded based on the variable they represent. The height of the bar suggests the strength of the correlation between a specific variable and the average win percentage. As the visualization suggests, average net efficiency has the highest r-squared value across all five season and average free-throw percentage has the lowest. In other words, average net efficiency has a consistent strong correlation with a team's success and average free-throw percentage has a consistent weak correlation. It is also noticed that the r-squared values for average defensive efficiency ratings fluctuates the most out of the six variables investigated in the study.

Conclusion

In our study, we explored the correlations between NBA teams' average win ratio over the five most recent regular seasons and six different variables that measure performance level. Based on our findings from the visualizations we created, we concluded that net efficient ratings correlates the strongest with a team's success in the league while free-throw percentage correlates the weakest. This finding is not surprising to us as we expected free-throw percentages to have the least impact on a team chances to win a game. We hope that the results of our study will highlight the significance of advanced statistical measures used in sports analysis such as efficiency ratings in comparison to regular stats (points per game, rebounds, etc.), which often fails to accurately evaluate a team's performance level. For future extensions to this study, we hope to incorporate more variables and increase the number of regular season used to generate a comprehensive version of the analysis we did in the current project. Furthermore, this will allow us to analyze any existing trends regarding the changes in the correlation strengths of each variable and apply it to the overall changes occuring in the league.

References

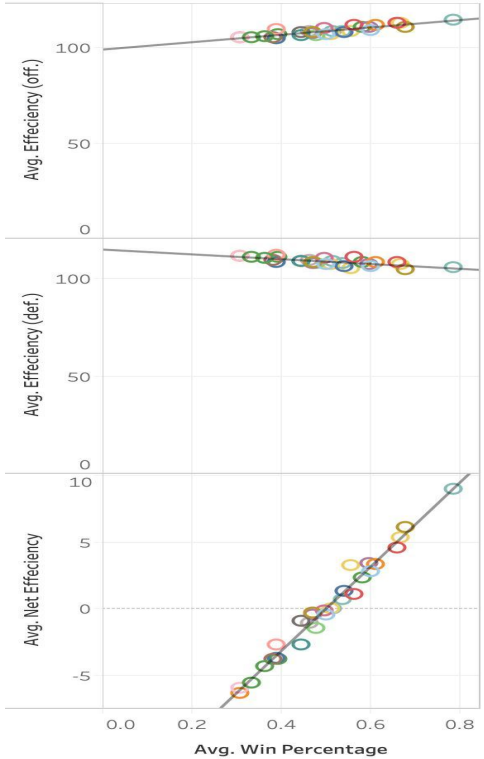
Data source: NBA & ABA League Index. (n.d.). Retrieved November 8, 2019, from https://www.basketball-reference.com/leagues/.

Related work: Yang, Y. S. (2015). *Predicting Regular Season Results of NBA Teams Based on Regression Analysis of Common Basketball Statistics* (Doctoral dissertation, PhD thesis, UC Berkeley).

Additional research:
Hayes, A. (2019, May 8). R-Squared. Retrieved November 28, 2019, from https://www.investopedia.com/terms/r/r-squared.asp.

Jones, K. (2018, November 4). The five factors behind the NCAA's NET ranking system. Retrieved November 8, 2019, from https://www.si.com/college/2018/11/04/college-basketball-rankings-net-system-explain.

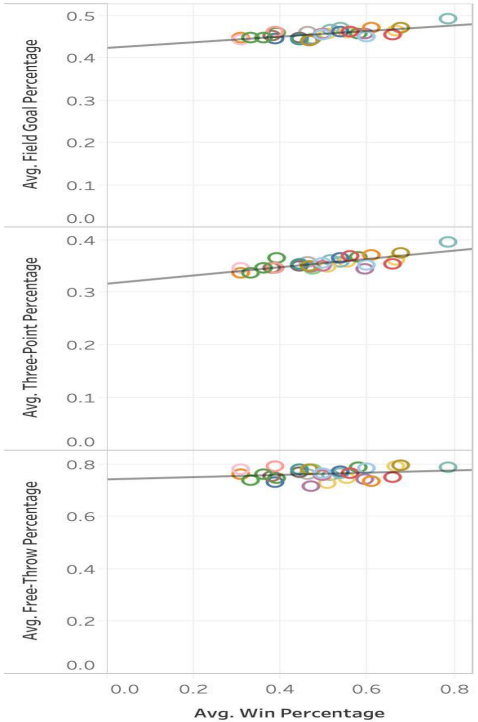
Average Efficiency Ratings v.s. Average Win Rate



Team▲
- Atlanta Hawks
- Boston Celtics
- Brooklyn Nets
- Charlotte Hornets
- Chicago Bulls
- Cleveland Cavaliers
- Dallas Mavericks
- Denver Nuggets
- Detroit Pistons
- Golden State Warriors
- Houston Rockets
- Indiana Pacers
- Los Angeles Clippers
- Los Angeles Lakers
- Memphis Grizzlies
- Miami Heat
- Milwaukee Bucks
- Minnesota Timberwolves
- New Orleans Pelicans
- New York Knicks
- Oklahoma City Thunder
- Orlando Magic
- Philadelphia 76ers
- Phoenix Suns
- Portland Trail Blazers
- Sacramento Kings
- San Antonio Spurs
- Toronto Raptors
- Utah Jazz
- Washington Wizards

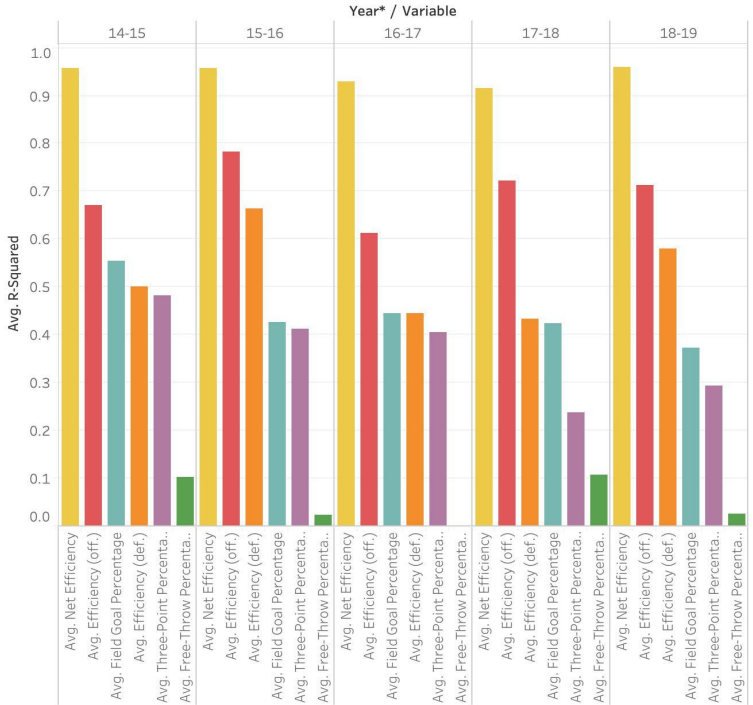
Average of Win Percentage vs. average of Effeciency (off.), average of Effeciency (def.) and average of Net Effeciency. Color shows details about Team▲. The view is filtered on Team▲, which excludes Null.

Average Shooting Percentages v.s. Average Win Rate



Average of Win Percentage vs. average of Field Goal Percentage, average of Three-Point Percentage and average of Free-Throw Percentage. Color shows details about Team▲. The view is filtered on Team▲, which excludes Null.

Changes in Correlation Over Years



Average of R-Squared for each Variable broken down by Year*. Color shows details about Variable. The view is filtered on Variable, which excludes Null.

Variable

- Avg. Efficiency (def.)
- Avg. Efficiency (off.)
- Avg. Field Goal Percentage
- Avg. Free-Throw Percentage
- Avg. Net Efficiency
- Avg. Three-Point Percentage