NBA Team Performance

Noah Cyr, Dominic Ranelli, Matt Maguire, Pedro Bandeira, J.R. Newman

Motivation

Basketball is a complex sport that contains a statistic for almost every aspect of the game. There are many ways to pair and compare these stats to investigate possible themes and relationships between them. We've selected certain game stats to test assumptions about the NBA and the results of their competitions, as well as the importance of the different components that make up the sport of basketball.

The Dataset

- First thing to be done: find a robust dataset.
- From the 2004 season to 2022 season, the data set had data about:
    - Games
    - Details of Games
    - Players
    - Ranking
    - Teams

Conjectures

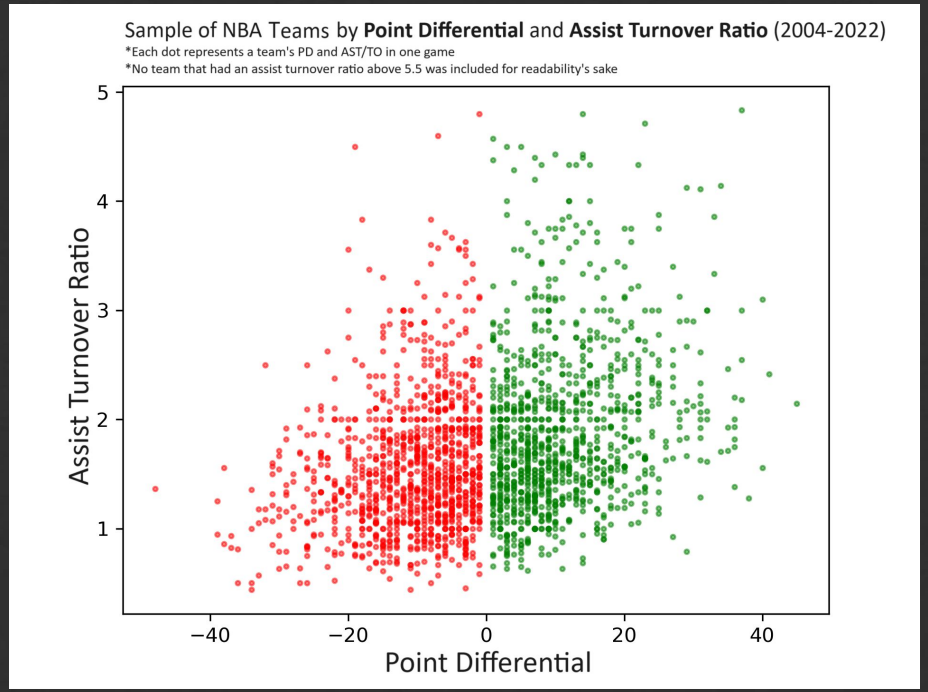
1. A higher assist to turnover ratio deems a team more likely to win.
2. The home team of a competition is more likely to win.
3. Model to predict the amount of points scored in a game.

Conjecture 1

A higher assist to turnover ratio deems a team more likely to win

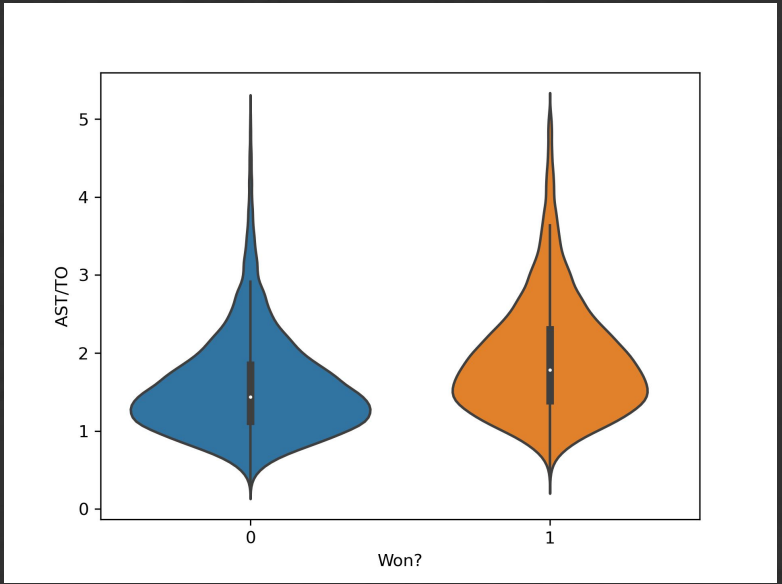
Assist/Turnover

- We looked at teams assist to turnover ratio and compared it to their point differential to their opponent.
- The scatter plot shows the correlation between a sample of the two stats. (ratio of 5 or less)



Assist/Turnover

- Here's a violin plot of all teams who had an assist turnover ratio less than 5.5 between 2004 and 2022, grouped by if they won or loss (0 for loss, 1 for win)
- Mean AST/TO of losing teams = 1.55
- Mean AST/TO of winning teams = 1.9
- This shows that the team with a high AST/TO does not always end up winning.



Assist/Turnover Conclusion

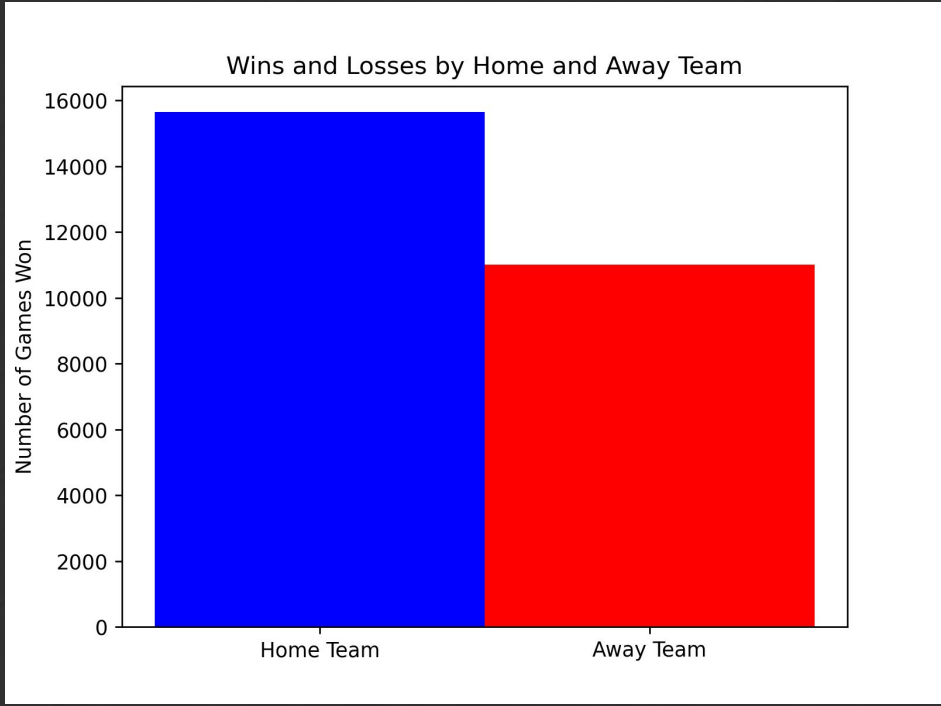
- If a team has an assist to turnover ratio of 3 or higher then it is very likely that the team will win.
- This ratio does not fully predict the outcome of a game and a team still has a good chance to win even with a low assist to turnover ratio.

Conjecture 2

The home team of a competition is more likely to win

Home Court Advantage

- Home team win percentage: 58.7%
- Away team win percentage: 41.3%

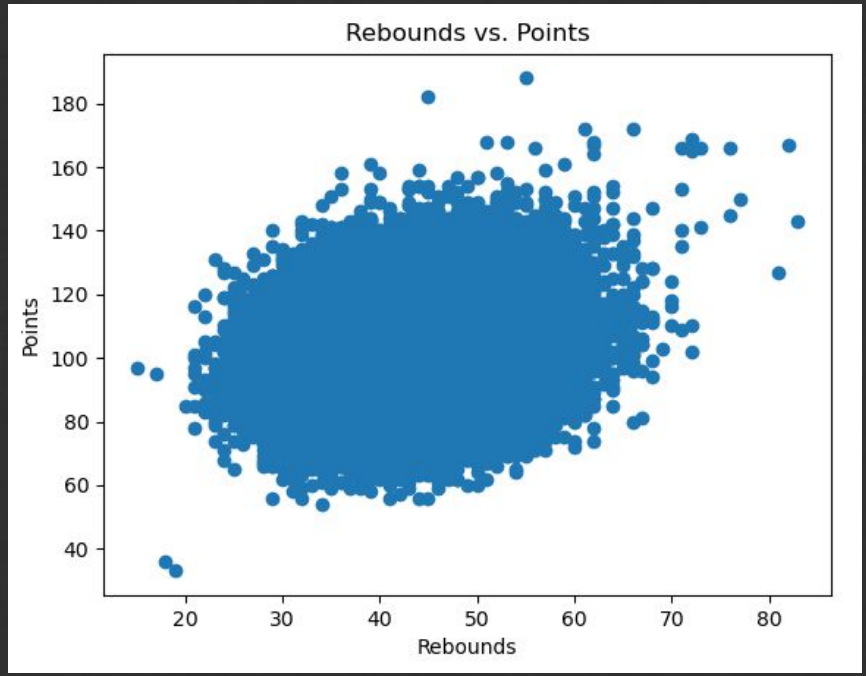


Conjecture 3

Can you make a model using statistics like a team's field goal percentage, free throw percentage, rebounds, assists, etc., to predict the amount of points scored in a game?

Predicting Points Scored - Rebounds

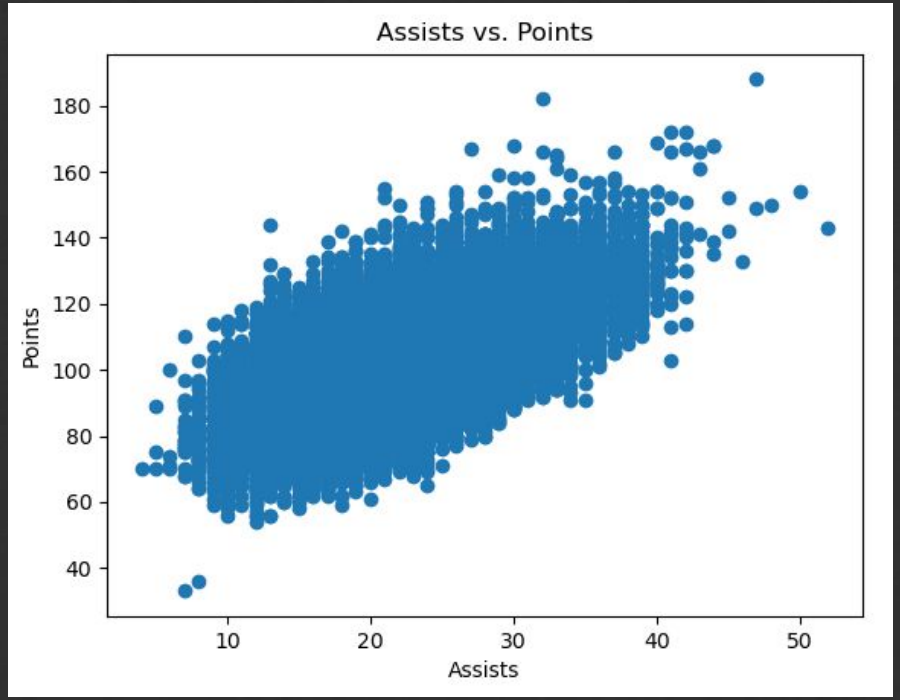
Both offensive and defensive rebounds are thought to be crucial to winning games. By gaining more possessions it is assumed that you will score more points. However, by observing the scatter plot, we can see that there is no clear relationship between the amount of rebounds a team records and the amount of points they score.



Correlation (r) = 0.19

Predicting Points Scored - Assists

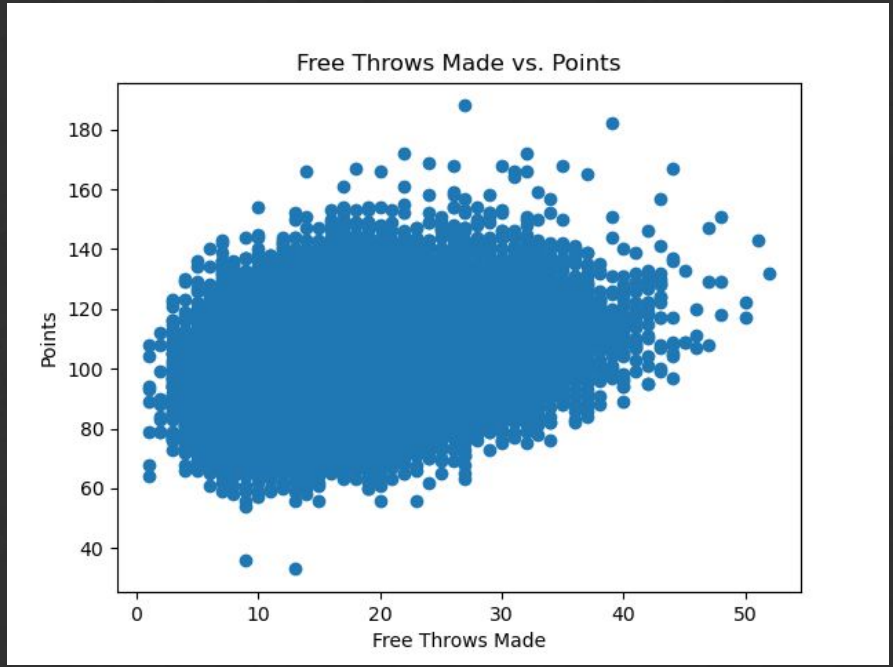
Assists are the result of teammates sharing the ball to get others baskets. There seems to be truth to the fact that assists can be a predictor of points scored in a game. Here we see that there is a sufficiently strong, positive correlation between the amount of assists in a game and the points scored in a game.



Correlation (r) = 0.61

Predicting Points Scored - Free Throws

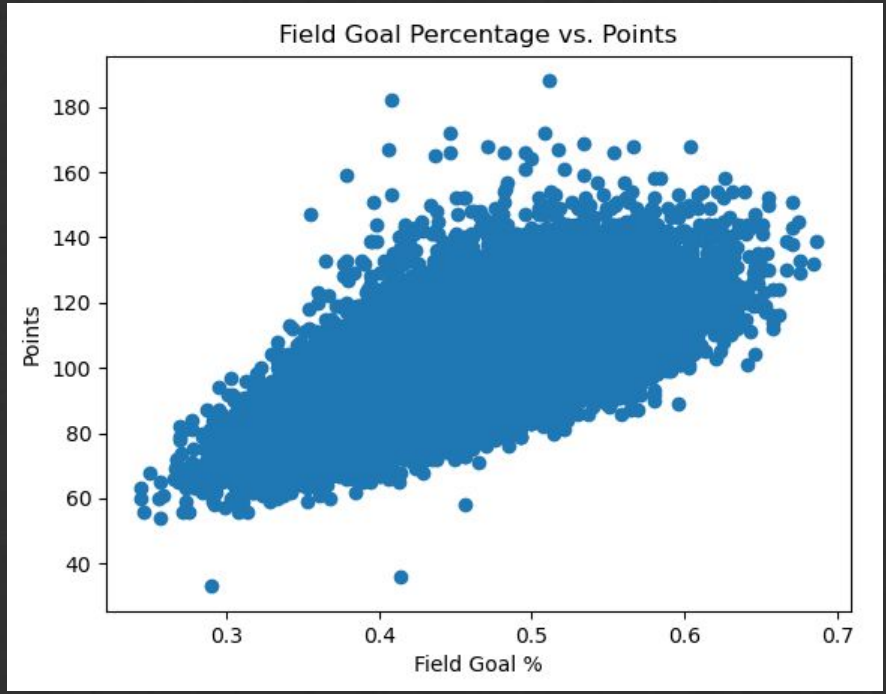
Free throws give teams the ability to gain unguarded points without losing time on the clock. Even still, there doesn't look to be any clear relationship between free throw percentage and points scored. We might have seen a more positive and stronger correlation if teams were taking a higher magnitude of free throws within these games.



Correlation (r) = 0.27

Predicting Points Scored - Field Goal Percentage

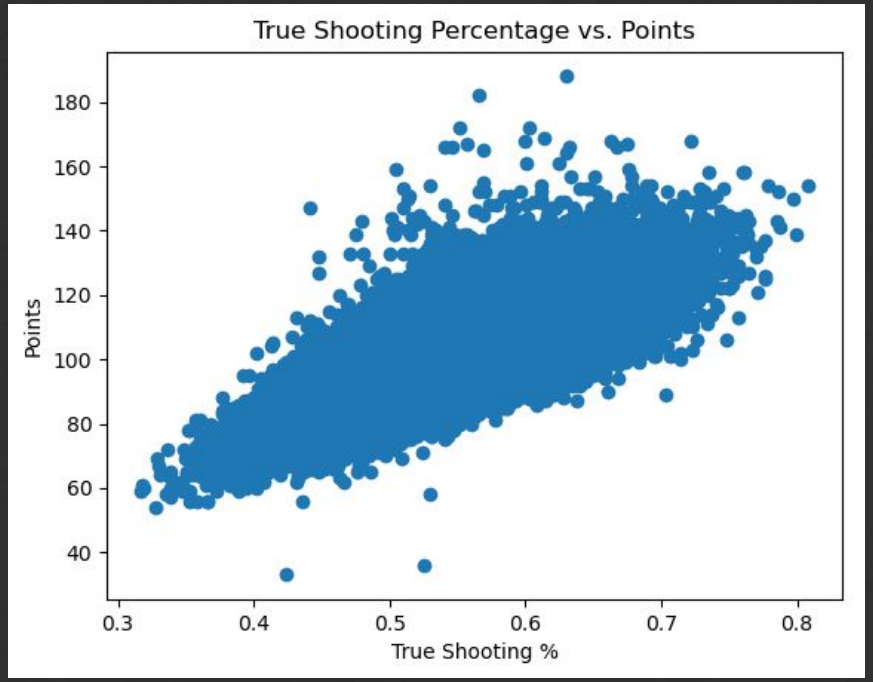
Field goal percentage is the amount of shots made divided by the amount of shots taken. The scatter plot shows us a moderately strengthed relation, but positive correlation between field goal percentage and points scored. An explanation for the relationship not being stronger could be that teams are not taking enough shots in a game to take advantage of a higher field goal percentage.



Correlation (r) = 0.66

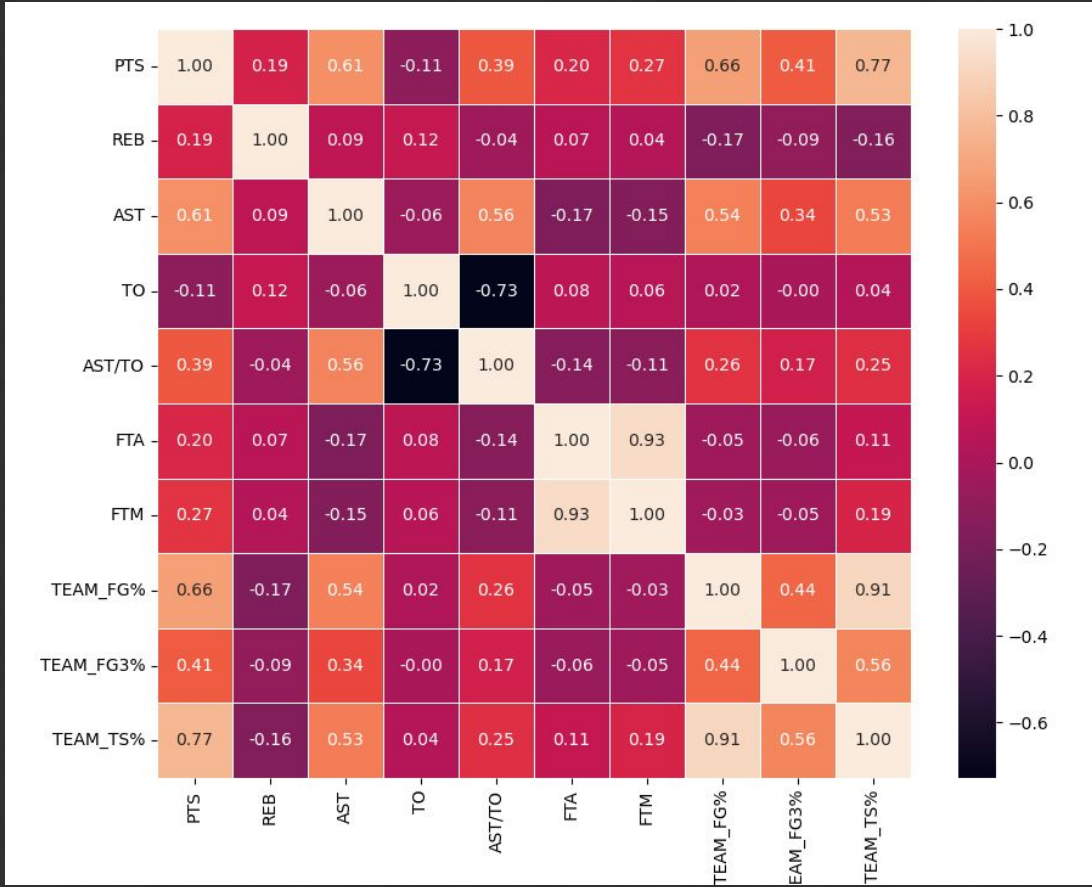
Predicting Points Scored - True Shooting Percentage

True shooting percentage is an advanced shooting statistic the factors in a team's performance for all types of shots-field goals, three pointers, and free throws. As you can see in the scatterplot, the data points are a little more narrow and linear then the field goal percentage scatterplot. This corresponds to the correlation coefficients of the variables, with FG% at 0.66 and TS% at 0.77.



Correlation (r) = 0.77

Correlation Matrix



Our Regression Model

- We used Multiple Linear Regression to fit a model
- Training set size: 81% of our data
- Testing set size: 19% of our data

Our formula:

$$PTS = -16.9 + (135.9*TS\%) + (0.72*AST) + (0.51*REB) + (0.38*FTM)$$

- Coefficient of determination (R^2) for our model: 0.662
  - 66.2% of the variation of a team's points in a game is accounted for by the independent variables TS%, AST, REB, and FTM
  - Good R^2 value considering NBA's fluidity and unpredictability

Our Regression Model in Action

DATE OF GAME	TEAM	ACTUAL PTS	MODEL'S PTS
December 7, 2007	Toronto Raptors	84 pts	83.94 pts
	Boston Celtics	112 pts	123.39 pts
February 2, 2018	Miami Heat	97 pts	96.20 pts
	Philadelphia 76ers	103 pts	103.91 pts
April 2, 2021	Oklahoma City Thunder	103 pts	93.87 pts
	Phoenix Suns	140 pts	132.41 pts

Thanks for listening