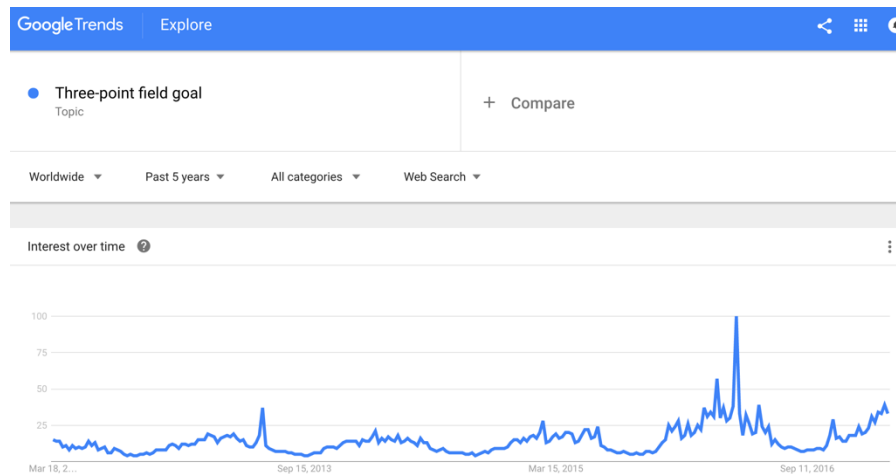


NBA Data Exploratory – Is 3-Point Shooting Just a Hype?

Introduction

Over the past 3 years or so, because of one team, Golden State Warriors, and one man, Stephen Curry, 3-point shooting has become a buzz word. Headlines everywhere covering how great 3-point shooting is, media hyping the all-star Stephen Curry, and sport analysts talking about how teams should build around threes, and the league is shooting more threes than ever in the NBA history.



As a data scientist, we would like to discover the truth about rumors and hype. Thus, this leads to some interesting questions to answer:

- Do winning teams take more 3-point shots?
 - Do winning teams take less short ranged shots?
- Have all the teams shifted to long range shooting strategy?
 - Are there teams getting left behind because they didn't adapt?
- Are 3 point shooting as effective as the hype says?
 - How effective comparing 3-point vs 2-point shots?

In this data exploratory project, I will answer these questions using facts and insights from data, and I will validate my points using statistics and mathematics.

Data

Important notes about the data:

- Data was obtained from two sources:
 1. The official NBA website, using the advanced stats function (<http://stats.nba.com/teams/advanced/>)
 2. NBA Miner website, for the shot distance usage data (<http://www.nbaminer.com/shot-distances/>)
- Two datasets were joined by matching the “**Season**” and “**Team Name**”
- From season 2004~2005 to season 2015~2016
- Seasonal average of team data

- Only regular season data was used. Did not use any playoff data
- The data table looks like the following:

Season	Team	GP	W	Less than 8ft. usage %	8-16 feet usage %	16-24 feet usage %	24+ feet usage %	Avg. Shot Dis.(ft.)	Offensive Rating
2015-2016	Atlanta Hawks	82	48	41.75	10.71	13.84	33.45	12.76	103
2015-2016	Boston Celtics	82	48	42.73	12.51	15.55	29	12.43	103.9
2015-2016	Brooklyn Nets	82	21	43.68	16.32	18.1	21.68	11.66	100.9
2015-2016	Charlotte Hornets	82	48	36.4	13.29	15.52	34.64	13.73	105.1
2015-2016	Chicago Bulls	82	42	41.5	15.43	18.51	24.11	12.32	102.1
2015-2016	Cleveland Cavaliers	82	57	39.39	12.83	12.52	35.09	13.33	108.1
2015-2016	Dallas Mavericks	82	42	35.79	14.47	15.87	33.6	13.71	104.8

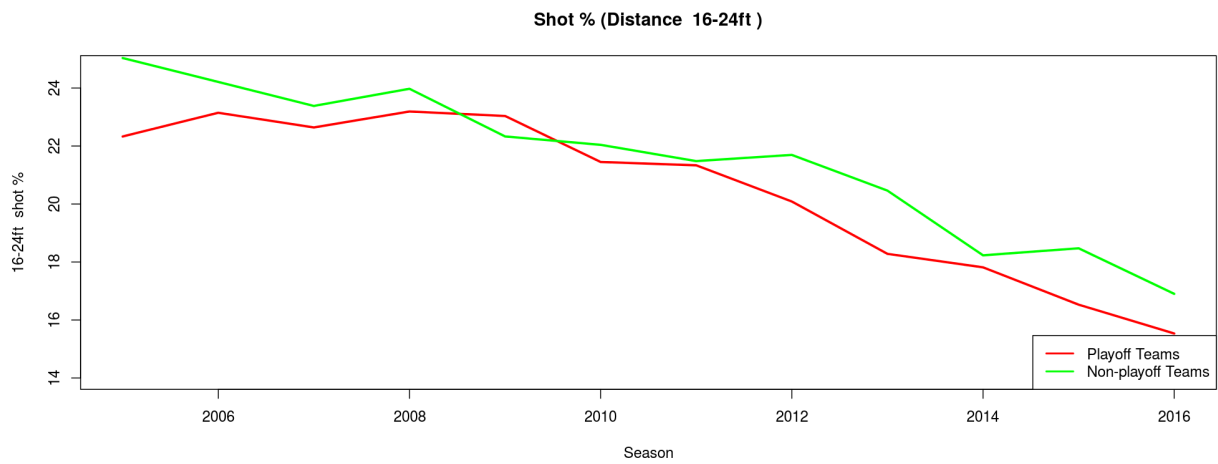
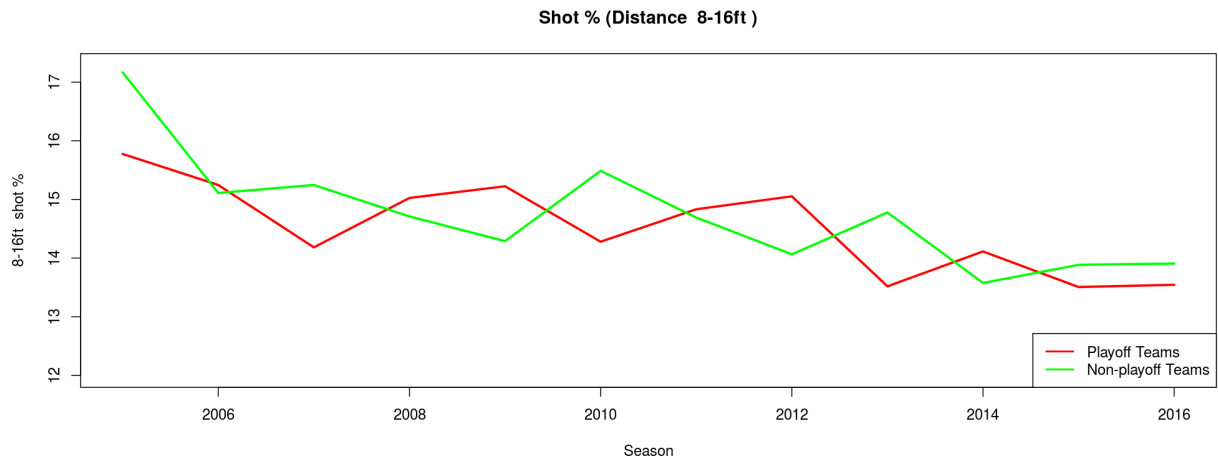
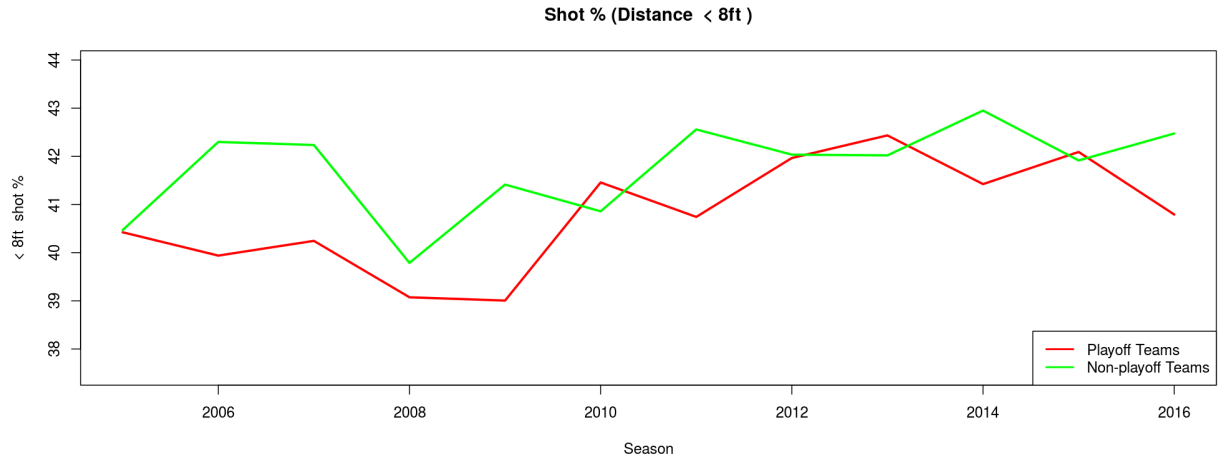
Data Exploration

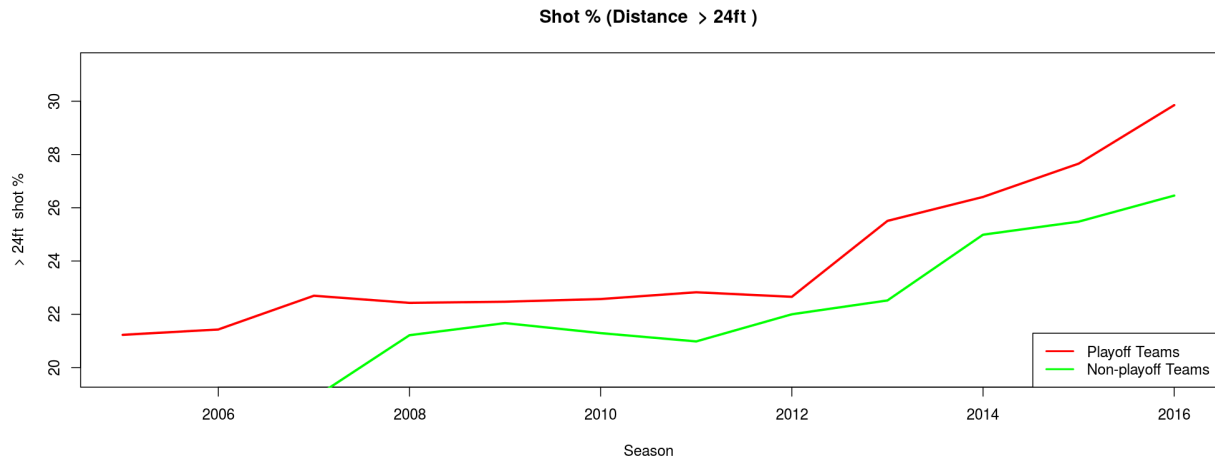
The teams are divided into:

- 1) Teams got into playoff
- 2) Team did not get into playoff

The shot usage % for different distances are potted for each season. Note the shot usage % for all distances should sum up to 100%:

$$(< 8ft.shot\ usage\ \%) + (8 - 16ft.shot\ usage\ \%) + (16 - 24ft.shot\ usage\ \%) + (> 24\ ft.shot\ usage\ \%) = 100\%$$

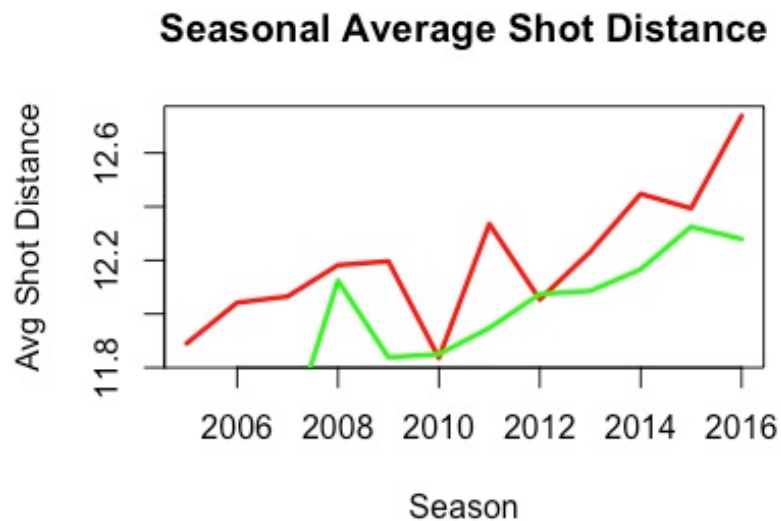




Important features for these 4 graphs are:

- Shot usage % for distance “< 8ft.” and “16-24ft.”, playoff teams tend to shoot less than non-playoff teams, while for distance “> 24ft.”, playoff teams shoot more.
- The whole league is taking more 3-point shots since 2012, as can be observed from the upward trend in the “> 24ft.” plot. All teams are sacrificing shots from “16-24 ft.” for the 3-point shots.
- There is barely any difference for shot distance “8-16 ft.”.

An additional plot below confirms again, teams that do well take longer range shots:



Going back to our hypothesis:

- Do winning teams take more 3-point shots?
 - Do winning teams take less short ranged shots?
- Have all the teams shifted to long range shooting strategy?
 - Are there teams getting left behind because they didn't adapt?

We can answer loud and clear, teams got into playoff definitely take more 3-point shots and less short ranged shots. However, surprisingly, the whole league is adapting the 3-point shooting strategy, not just Golden State Warriors. Now we are down to one last hypothesis to validate:

- Are 3 point shooting as effective as the hype says?
 - How effective comparing 3-point vs 2-point shots?

Before we can validate, we need to look at one advanced metric, offensive rating:

$$Offensive\ Rating = \frac{100 \times Points}{0.5} \times \left\{ \begin{aligned} &\left[FGA + (0.4 \times FTA) - \left(1.07 \times \left(\frac{ORB}{ORB + Opp. ORB} \right) \times (FGA - FG) \right) + TOV \right] + \\ &\left[Opp. FGA + (0.4 \times Opp. FTA) - \left(1.07 \times \left(\frac{Opp. ORB}{ORB + Opp. ORB} \right) \times (Opp. FGA - Opp. FG) \right) + TOV \right] \end{aligned} \right\}$$

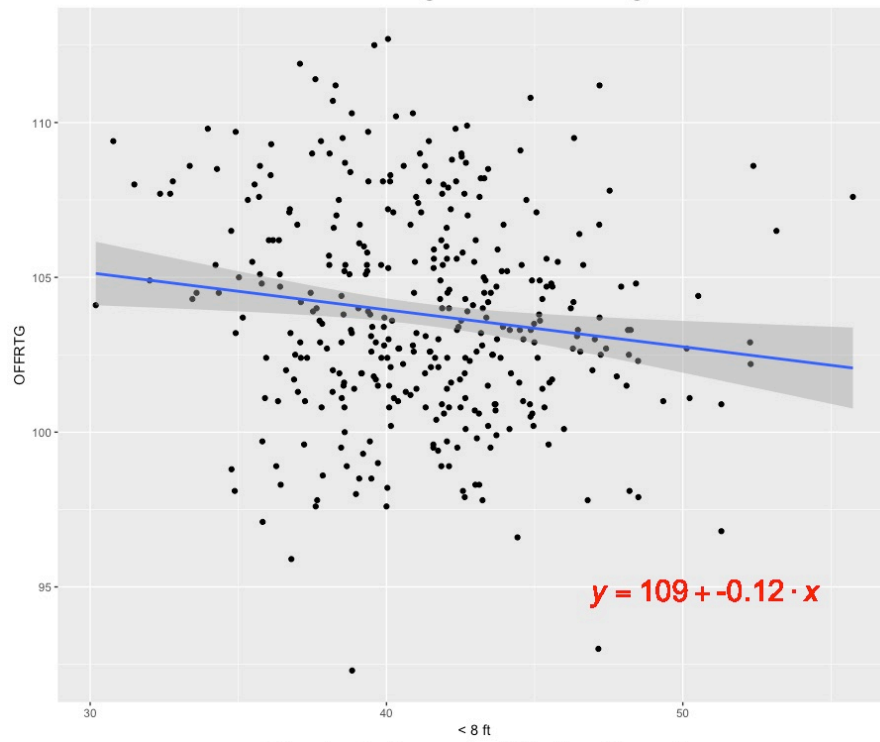
Offensive rating is developed by a statistician, Dean Oliver. It is a rather complex equation, so we will look at a simpler version that is more understandable:

$$Offensive\ Rating = 100 \times \frac{Points\ Produced}{Total\ Possession}$$

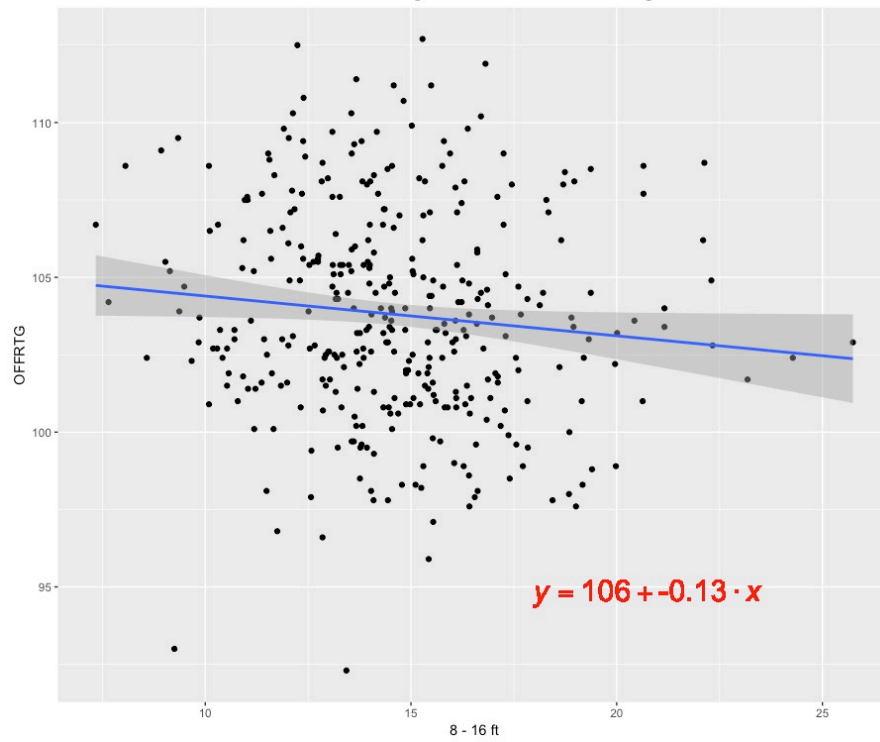
$$Points\ Produced = f(FGA, FTA, ORB, TOV)$$

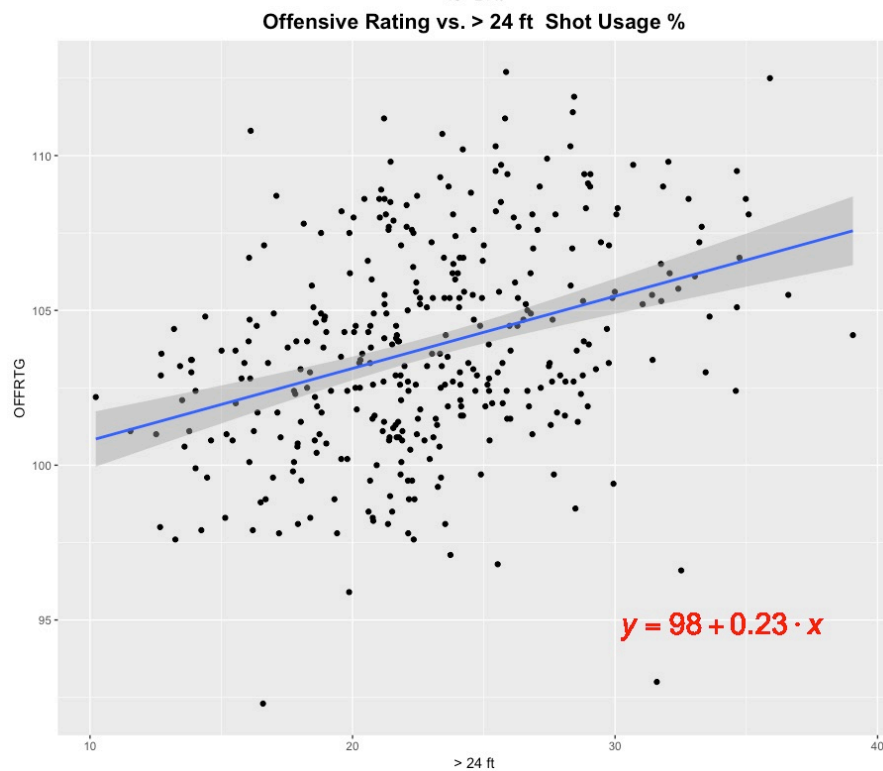
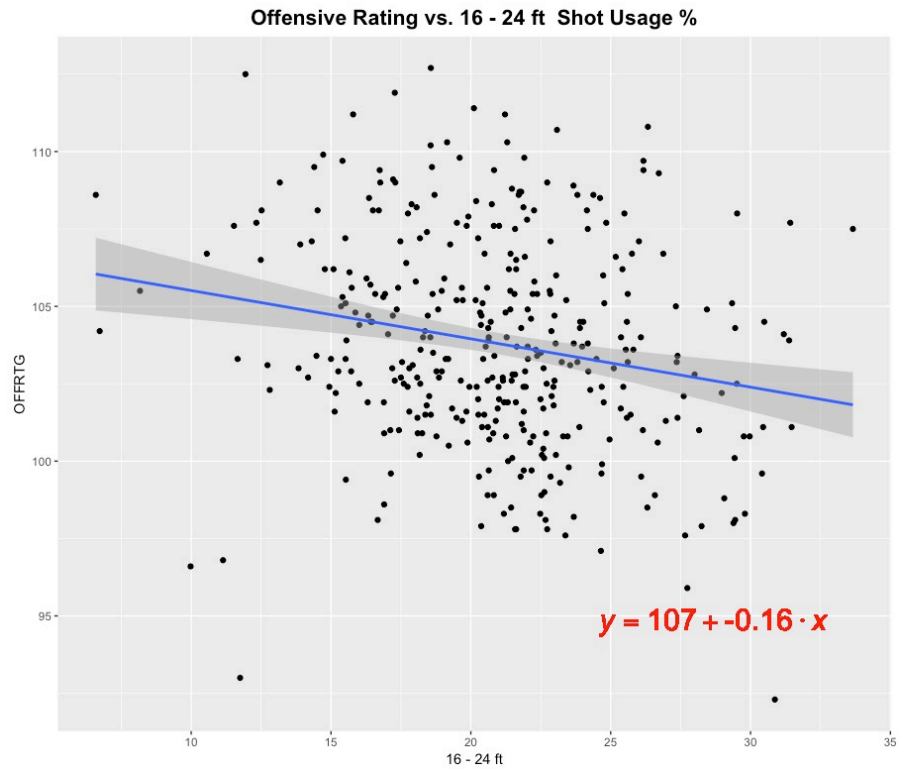
With offensive rating, we can measure how effective is 3-point shooting compared to other shot types, thus, offensive rating vs. different shot distance usage are plotted:

Offensive Rating vs. < 8 ft Shot Usage %



Offensive Rating vs. 8 - 16 ft Shot Usage %



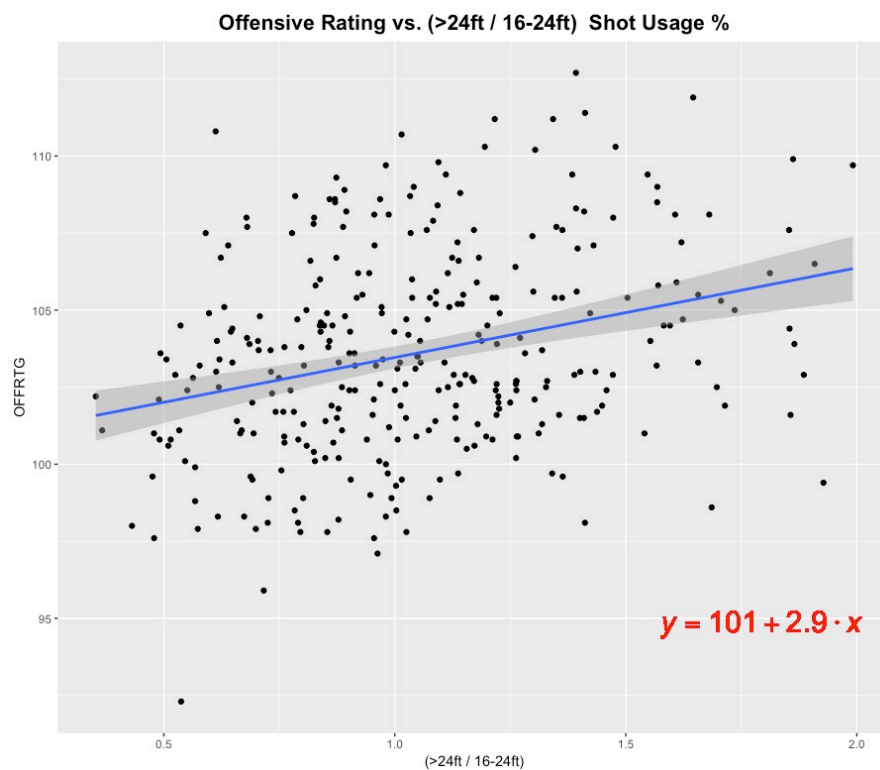


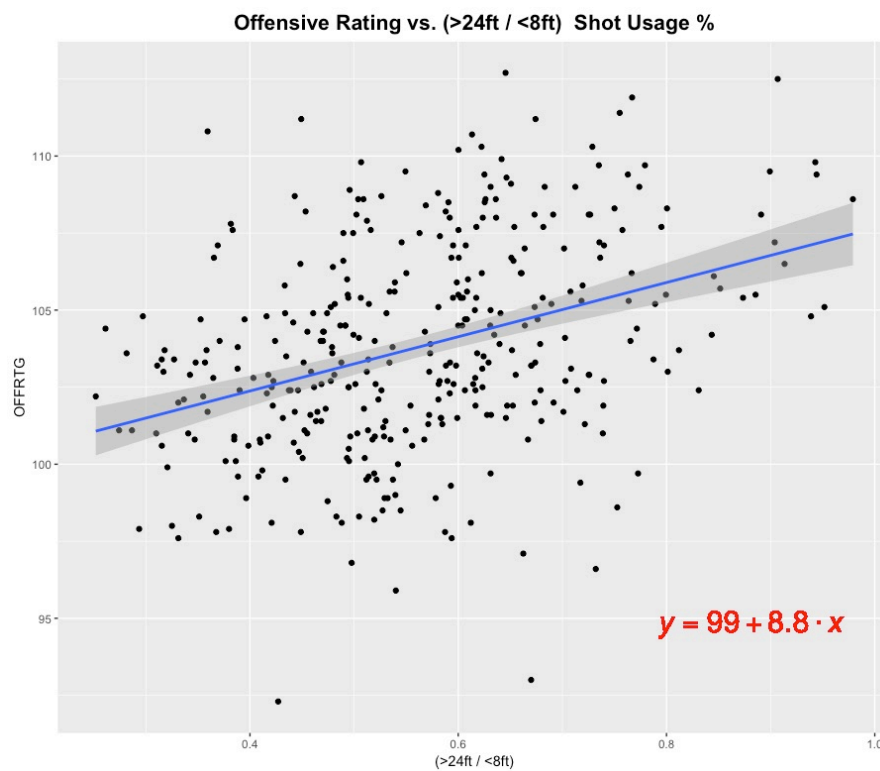
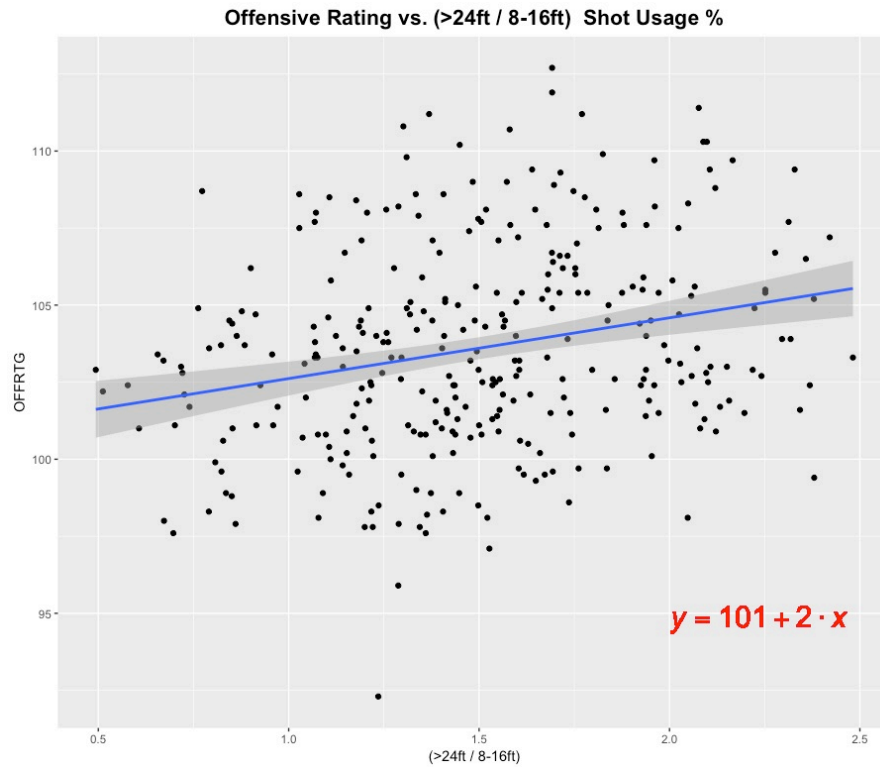
As we can observe from these 4 plots, all shot types except for “> 24ft.” has negative correlation with offensive rating, i.e. the slope is negative. Only “> 24ft.” has a positive slope, thus, we may conclude that the more 3-point shots are taken, the higher offensive rating will be.

However, we may realize the slope for “Offensive Rating vs. > 24ft.” is quite small, 0.23, meaning that the “> 24ft.” variable alone may not be a good indicator of how well the teams will do offensively. At this point, we need some new indicator to tell us how well the teams we will do offensively in terms of shot distance. A useful metric would be the ratio of shot types. Coaches often tune the team by mixing shooters and drivers (guards that can penetrate defense and get the ball close to the basket), therefore, the ratio between 3-point shot vs. other shooting distance usage may be a good indicator for offensive rating. We came up with 3 new variables here:

1. (3-point shot) / (16-24ft. shot)
2. (3-point shot) / (8-16ft. shot)
3. (3-point shot) / (< 8ft. shot)

Offensive rating vs. these 3 new variables are plotted below:





We can see the slope for “Offensive Rating vs. (3-point shot / <8ft. shot)” has the largest slope of 8.8, meaning for 1 point increase in the (3-point shot / <8ft. shot) ratio, we gain 8 points

in offensive rating. If we further run a linear model using the shot distance variables to predict offensive rating, we get the following summary:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	291.3102	188.2017	1.548	0.12268
cleaned_rating_data.ratio_24_0ft	37.4477	13.8355	2.707	0.00718 **
cleaned_rating_data.ratio_24_8ft	2.2929	2.0082	1.142	0.25445
cleaned_rating_data.ratio_24_16ft	0.4886	2.5386	0.192	0.84751
cleaned_rating_data.Less.than.8ft..usage..	-1.7226	1.9037	-0.905	0.36624
cleaned_rating_data.X8.16.feet.usage..	-1.8803	1.8844	-0.998	0.31915
cleaned_rating_data.X16.24.feet.usage..	-2.1789	1.8861	-1.155	0.24889
cleaned_rating_data.X24..feet.usage..	-2.9917	1.8823	-1.589	0.11300

Looking at the last column (p-values) of the row with double asterisk mark, the value shows 0.00718, it is telling us:

The linear model is 99.282% confident that the (3-point shot / <8ft. shot) ratio is a good indicator of how well the team will do offensively.

Recall our seasonal shot distance plot, teams are sacrificing “16-24ft.” shot for “> 24ft.” shot:



Our insight tells us, trade-off between 3-points shot and < 8ft. shot might be a better choice for offensive, because for 1 point gain in this ratio, we gain 8.8 points in offensive rating, while other trade-off only gain 2 ~ 3 points in offensive rating. The teams may want to rethink their strategy of which shot types to sacrifice for 3-point shot. Now to answer our hypothesized question:

- Are 3-point shooting as effective as the hype says?
 - How effective comparing 3-point vs 2-point shots?

Our answer is 3-point shot is the most effective shot, the league has over underestimate its power before. Also, it is worth to revisit the shot combination strategy to see which shot types should the team sacrifice for 3-point shots.

Conclusion

The 3-point shooting hype is real, as all the teams are shooting more threes than ever before. We do see playoff teams are shooting more longer range shots, and we realize the power of 3-point shots are definitely underestimated before by the league. We also found that shot combination is a very important variable to tell us how well the teams will do offensively, and the teams may want to revisit their shot combination strategy. In this project, we have barely scratched the surface of data exploratory for NBA statistics. More variables may be looked at such as rebounds, turn overs, defensive ratings, and game logs data etc., to make prediction of how well the team will do during the season.