

# Article 1

*Justin Yee*

*10/25/2018*

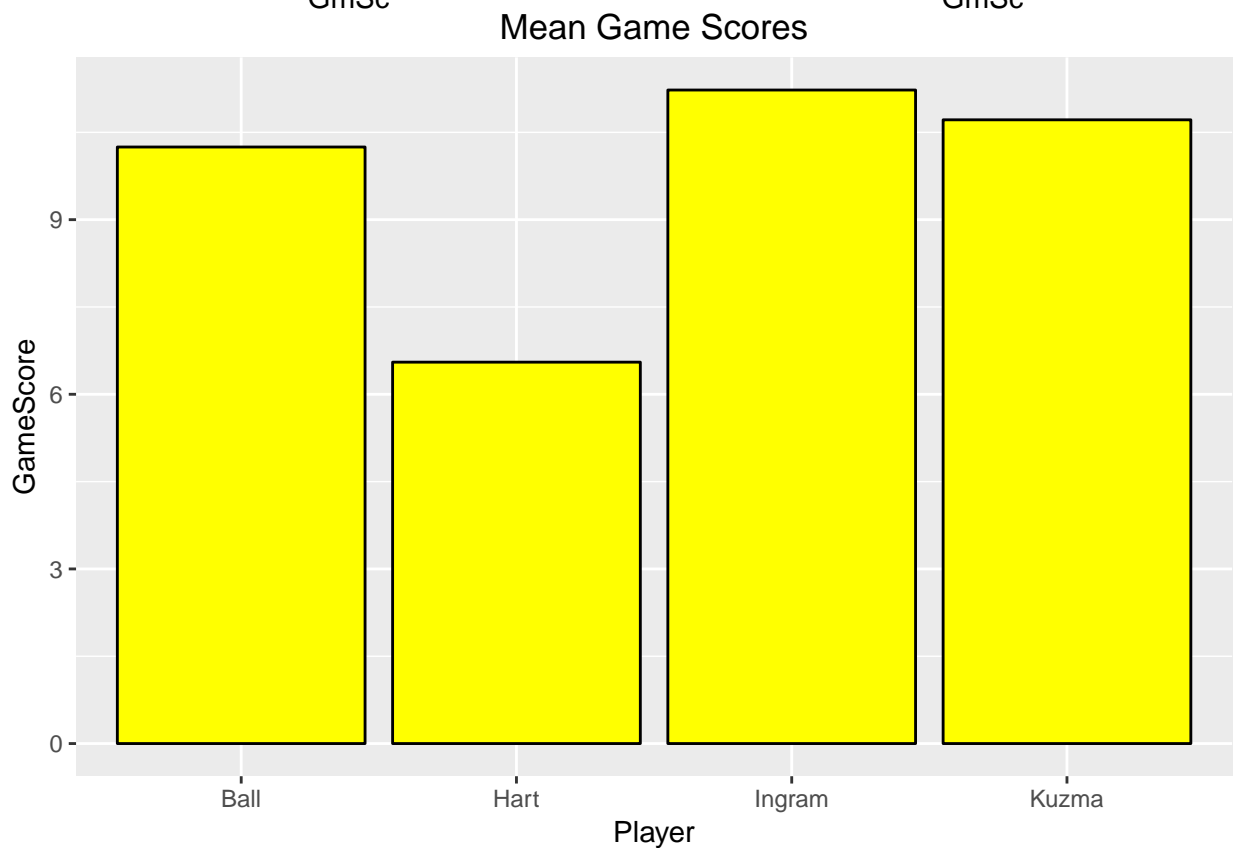
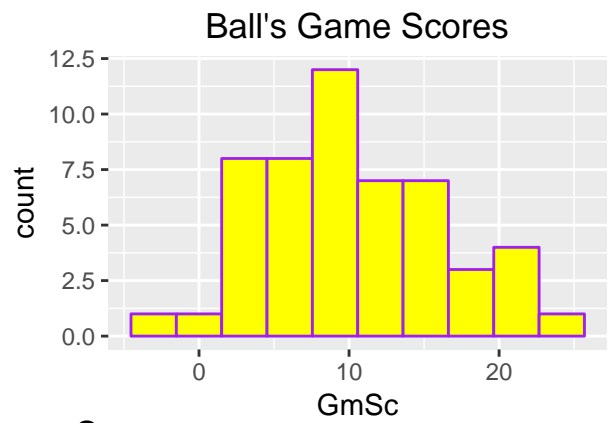
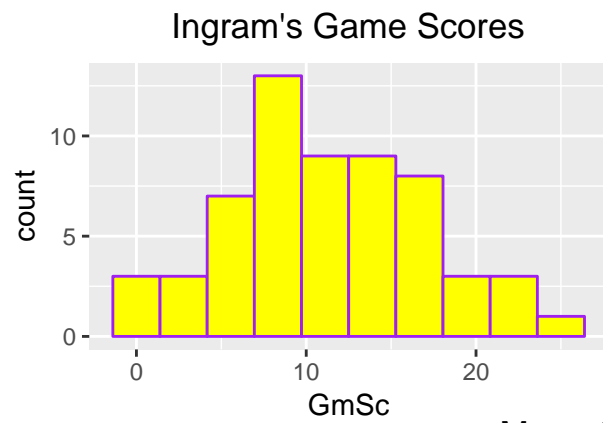
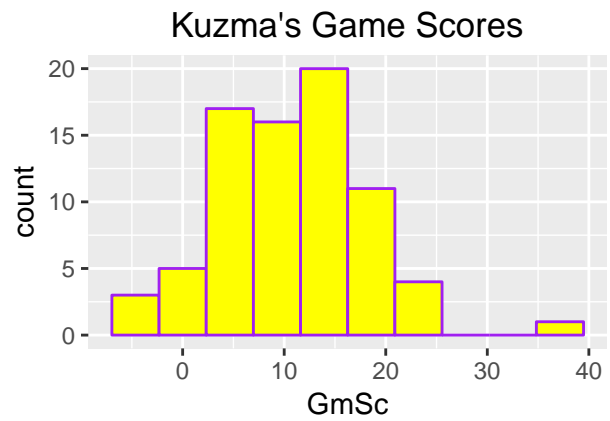
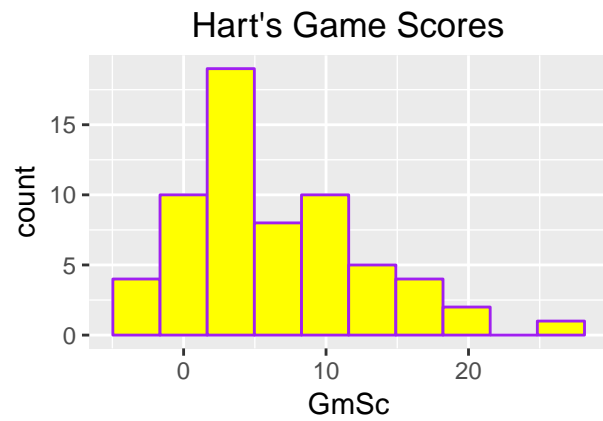
## **Keys to Success for the Young Lakers Core**

(INCLUDE THE LINK TO GAMESCORE STAT CALC)

With Lebron coming to Los Angeles, the expectations surrounding the Lakers haven't been this high since the Kobe Bryant, Dwight Howard, and Steve Nash team-up. After missing the playoffs for blank consecutive years, the Lakers look to make a playoff run and possibly contend for a championship title with the addition of the NBA's most prominent star. In addition to Lebron James, the Lakers added notable veterans such as Rajon Rondo, Javale McGee, and Lance Stephenson. However, the Lakers remain at their core a very young team, with blossoming players with high potentials. These young players who look to be the future of the franchise are Brandon Ingram, Kyle Kuzma, Josh Hart, and Lonzo Ball. In this article, I will focus on these young players, analyzing their individual game log data I retrieved from Basketball-Reference.com and using John Hollinger's Game Score statistic as a metric in determining the success each player had the last 2017-18 season in each game they played in. Game Score is a more complex metric that was designed to roughly estimate the success of a player's single game. If you are interested in knowing how Game Score is exactly calculated, you can find the informaion in this link. My goal in analyzing this data is to highlight the strengths and weaknesses of each player from the perspective of exactly *how* each succeeded relative to their average performances of the season.

### **Exploring the Data**

To start my stastical analysis, I began by exploring the data. First, we can take a quick look at the distributions of Game Score for each player, as well as comparing the mean Game Score for each player in the following charts.



From these charts, we can notice that the distribution of the Game Score's were roughly similiar to normal

distributions. We want to capture what the greatest predictors are of each individual's Game Score relative to their own average performance on the year. In other words, we want to find out exactly *what* that player did exceptionally well or exceptionally poorly to merit a good or bad game score.

## Methodology and Rationale

From a statistical standpoint, my primary methodology predicting the Game Score statistic using variables from the game logs as my predictive variables through a multiple linear regression model. The rationale behind choosing such a model is for interperatable purposes, as it is clear from the output of this model the effects of each variable in predicting the statistic Game Score.

## Methodolgy of Normalization

(FINISH THIS PART)

An essential part of building these models is the question of normalizing the data. Should I normalize the data according to standard deviations of the stats of the individual player, according to league averages, or other method? YES, we want to normalize by individual since this will tell us what the individual should work on more for their own personal game, not compared to everyone else, or likewise, tell the individual what aspects of their game are negatively affecting their Gamescore and what they must improve upon.

## Building the Models

To build the models, I tested the subsets up to a total of five predictors through an exhaustive search, choosing the model with the highest adjusted  $R^2$ , a statistical measure which describes the explanatory power of the model in predicting Game Score. Since Game Score most heavily weights points in its calculation, I forced the Points variable out of consideration for building the models to tease out the other predictive powers of the more evenly weighted variables. The following charts below summarize my results for these models.

Table 1: Model Coefficients for Each Player

	Hart	Kuzma	Ingram	Ball
Intercept	6.5523810	10.715584	11.227119	10.248077
FG	7.0236542	6.851481	7.039416	6.503627
FGA	-3.4908755	-2.411566	-2.872533	-2.696985
FT	1.5759834	2.284202	1.497359	0.000000
AST	1.6627760	1.618518	1.747599	1.969441
STL	0.7679265	0.000000	0.000000	1.245747
TOV	0.0000000	-1.717221	-1.428057	-1.316189

## Interpretation and Breakdown of the Models

Looking across the various models, we can notice some differences. These results are somewhat suprising, as they give us a different perspective than looking at a player's overall season statistics, such as Josh Hart's 47% FG% on the season telling us that he is an efficient scorer. This compared to Lonzo Ball's shockingly-low 36% FG% from last season. Lonzo also shot at a much higher clip than Hart... However, we must keep in mind that these statistics are normalized, so they are adjusted for in respect to that player's average performance. THIS IS THE KEY, as to why the results are the way they are.

For Hart, his field goal atempts most drastically affects his Gamescore negatively, meaning he should focus on better shot attempts.

For Kuzma, his turnovers most adversely affect his Gamescore negatively, meaning he should focus on being more careful with the ball. In addtion, his field goal attempts do not dramatically decrease his Gamescore

compared to his fellow teammates.

For Ingram, his field goals most positively improve his performance, so he should focus on staying aggressive and being a primary scoring option.

For Ball, his assists and steals most positively affect his Gamescore relative to the other three players. His made field goals contribute the least to his Gamescore relative to the other three players. For him to succeed, he must maintain aggressive on defensive possession as well as improve on his scoring.

## **Conclusion and Further Remarks**

There are two key aspects of our analysis that differentiate it from the typical, more basic analysis of a player's performance. From the perspective of looking at singular games rather than a whole season, we can find different results that may give us deeper insight into the success of these players solely on a game-to-game basis. Additionally, choosing to normalize based on that player's own performances, we can see statistical differences in their own games rather than comparing it to a metric of other players, such as league averages for box-score statistics. However, despite the uniqueness of our analysis, there are still some important concerns we must keep in mind.

With so few data points (only about 60-82 games), there are some difficulties that arise. Due to this small sample size, we are unable to train the model with a higher number of predictive variables, which is the reason I limited the models to a maximum of five predictive variables to try and avoid overfitting. However, as a result,

Another problem we came across was trying to evaluate individual success based upon the single metric of Gamescore. This decision to use Game Score as our metric effectively constrained the success of our analysis to the accuracy of Game Score as a barometer for success. To further answer these questions, we would have to dive into the robustness and validity of Game Score. While Basketball-Reference.com provides us with a great open-source database, there still lies constraints in our ability to capture game-by-game data of other metrics besides Game Score.

Additionally, because Game Score is a statistic that is calculated using box score statistics, and our predictive variables themselves are box score statistics, there arises an issue of inherent dependencies between our predictive variables and our response variable of Game Score.

Despite these potential pitfalls and constraints on our analysis, the models constructed for this analysis provide some extra insight into the factors that determine success for the young Lakers core. In the world of data analysis, and especially sports, there is almost never a flawless interpretation of the the analysis. However, there is still much to gain and learn from by taking a closer look at the numbers and potentially uncovering insights made possible through data analytics.