

Team: Sigma Seekers

Members: Navika Maglani, Rishika Vala, Chin-Ya Yueh, Aaron Lint

DSC 423

March 6, 2023

DATA ANALYSIS AND REGRESSION

Final Project Report

Summary

Basketball is a world wide sport that offers fans an opportunity to see their favorite players score points and win games against their rivals. Beyond the excitement of live matches, player statistics including numbers and patterns therein can offer immense insight into the very essence of player scoring. This report seeks to unravel the impacts of several basketball statistics in their predictability on player scoring.

We aim to enhance player performance, team strategies, and possible game outcomes through regression analysis. To understand the statistical significance of player statistics and metrics, we used a comprehensive regression analysis strategy to identify key performance indicators. Our analysis is based on well known basketball statistics- such as points per game, assists, rebounds, and blocks- and the examination of such variables to inform basketball strategies. Overall, we aim to bridge the gap between the collection of raw data into actionable knowledge which may contribute to better team performance on the court. We will examine each position on the court- the point guard, shooting guard, small forward, power forward, and center- and their separate contributions to scoring.

Findings

Point Guard

1. Age and the number of games played have a relatively smaller but still significant impact on scoring. This implies that experience and physical condition play a role in a Point Guard's scoring performance.
2. 3-point attempts and successful 2-point shots have a positive impact on scoring, a higher effective field goal percentage is strongly associated with increased scoring with a high coefficient of 1.692 suggesting the importance of efficient shooting for the point guard.
3. Point guards should excel on the defensive end by contributing to defensive rebounds because defensive rebounding, whether offensive or defensive, has a positive impact on scoring with a coefficient of 0.1168, suggesting the importance of defensive skills for Point Guards. It indicates that point guards who actively contribute to rebounding, especially on the defensive end, tend to score higher.
4. Assists play a significant role in the model. While it might seem counterintuitive that more assists are associated with higher scores, it may indicate that the Point Guard is more productive when assisting team members. This aligns with the diverse play styles

exhibited by different point guards. Team synergy and chemistry would be a factor to influence a Point Guard's scoring behavior. However, the interaction terms and second-order terms indicate that there is a trade-off for assists. It reflects the decision-making required in balancing playmaking responsibilities with other aspects.

5. Variables that do not contribute: Min.. played, field goals, Field goals attempted, FG, X3 point made, X2 point attempted, Free throw made, Total rebounds, Blocks, Turnovers.

Shooting Guard

1. The analysis for shooting guard identified several significant predictors on points scored including age, games played, games scored in, free throw percentage, minutes played, three point percentage, effective field goal percentage, offensive rebounds, assists, steals, turnovers, and personal fouls. Both linear and quadratic terms for minutes played were significant indicating a non-linear relationship with points scored.
2. Age is significant when predicting points scored through an inverse relationship suggesting an one unit increase in player's age is associated with a decrease in .193 points in scoring.
3. Both free throw percentage (FT%) and effective field goal percentage have a positive significant impact on scoring suggesting that these variables are critical to scoring points. Effective field goal percentage has a very high coefficient of 1.29, suggesting its importance on scoring.
4. The interaction between games scored and free throw percentage indicates that the impact of games scored on points is more pronounced as free throw percentage increases.
5. Minutes played has a significant positive relationship on points scored; however, its squared term also indicates that, while playing time is crucial for scoring, there is a level at which a player experiences diminishing returns on points scored with minutes played increasing.
6. Turnovers have a significant positive effect on scoring. This may indicate that a player who handles the ball often has more opportunities to score while having more opportunities for turnovers.
7. Defensive rebounds and blocks were insignificant when predicting points scored for the shooting guard which may indicate that defensive statistics are not relevant to an offensive player's scoring.

Small Forward

1. Free throws made and the square of free throws made have a positive relationship with points scored, with coefficients of .68 and .08 respectively. This may indicate that each free throw made contributes to scoring but with a diminishing impact.
2. The negative coefficient for games played (-0.0024) and the positive coefficient for its squared term (.00002) indicates that, while more games might not directly translate to more points scored, over time points scored will increase with experience.
3. Points scored increases by 2.42 points for every unit of increase in effective field goal percentage- a powerful predictor in points scored. However, the negative coefficient for

field goal percentage (-1.093) indicates that a higher scoring percentage might be associated with less scoring.

4. Turnovers have a positive relationship (.3134) with points scored indicating a small forward who handles the ball more, despite turnovers, may score more points. Blocks (-.17) have a negative relationship with points scored indicating that blocks, which are typically made under the basket, may not be a reliable position for the small forward.
5. Assists (.1223) and defensive rebounds (.3137) are significant and positive relationships with points scored. This may indicate that the small forward is more productive when contributing to play on the court other than just shooting.
6. Three point percentage (.136) also shows a strong and positive relationship when predicting points scored, indicating the importance of three points scoring.
7. 2 point percentage, minutes played, and offensive rebounds were insignificant when predicting points scored for the small forward.

Power Forward

1. The coefficients of free throws attempted is highly statistically significant in predicting the response variable. The corresponding p-value is "< 2.20-16," which is incredibly low. Strong evidence that the variable "Free Throw Attempted" is a significant predictor in the model is suggested by the low p-value. It suggests that in the context of basketball data, a team's or individual's total number of free throw attempts has a significant influence on the response variable under study in your regression model.
2. The coefficient of three pointers has a p-value that is very significant statistically. This suggests that a key factor in predicting the response variable is the quantity of 3-pointers made. The response variable taken into account in your regression analysis is significantly impacted by a team's or individual's total number of successful 3-pointers made. More effective 3-point shooting teams or individuals typically have a discernible impact on the result under study.
3. With an estimated coefficient of 5.237e-03, the variable "Offensive rebounds" is shown to be statistically significant in predicting the outcome of the analysis. This suggests that there is a strong positive correlation between an increase in offensive rebounds and the response variable.
4. The coefficient of defensive rebound is extremely significant statistically in predicting the result of the analysis. It suggests that the outcome or performance being modeled is often significantly improved by teams or individuals who have a higher number of defensive rebounds. Analysts and coaches could think about emphasizing defensive rebounding tactics as a way to boost the performance of the squad as a whole.
5. There is a strong statistical correlation between the variable "Field Goals" and the predicted outcome of the analysis. This suggests that a significant and very influential component determining the result or performance being modeled is the quantity of field goals made. Players or teams with a greater percentage of made field goals have a tendency to significantly improve the final result.

Center

1. The coefficients of assists and turnovers suggest that these variables have a moderate but statistically significant impact on players' points. Assists contribute positively, indicating that facilitating scoring opportunities for teammates enhances overall point production. Conversely, turnovers have a negative impact on points, highlighting the importance of ball security and minimizing errors.
2. The coefficient of three-pointers made (3P) indicates that successful three-point shooting contributes positively to players' points. However, the coefficient for three-point attempts (3PA) is not statistically significant, suggesting that the frequency of attempting three-point shots may not significantly impact overall scoring.
3. The coefficients of field goals attempted (FGA) and field goals made (FG) suggest that increasing the number of attempted field goals and the field goal percentage (FG%) positively affects the points scored by players. This highlights the importance of shooting accuracy and efficiency in contributing to overall scoring.

Data Selection and Preparation

The dataset was retrieved from Kaggle named NBA_Player_Stats.csv. There are 14,575 rows and 31 variables of data where each row indicates a player's stats for the year. The data ranges from 1998 through 2022. Before splitting the data among team members, we removed variables Rk, Player, Tm, and Year as these variables are irrelevant to scoring according to position. An explanation for each variable is listed in the table below. The position variable was then sorted and split among team members for the analysis. It is important to note that certain variables are a description of the average for a year's time frames as shown below.

Features	Description
Rk	Player Identifier
Player	Player
Pos	Player's position
Age	Player's age
Tm	Player's team for the year
G	Games played in the year
GS	# of games scored in
MP	Average minutes played per game
FG	Average field goals made per game
FGA	Average field goals attempted per game
FG%	Field goal percentage
3P	Average 3 point shots made per game
3PA	Average 3 points attempts per game
3P%	3 point field goal percentage made
2P	Average 2 point shots made per game
2PA	Average 2 point attempts per game
2P%	2 point field goal percentage made
eFG%	Adjusted field goal percentage accounting for 3 pointers
FT	Average free throws made per game
FTA	Average free throw attempts per game
FT%	Free throws made percentage
ORB	Average offensive rebounds per game
DRB	Average defensive rebounds per game
TRB	Average total rebounds per game
AST	Average assists per game
STL	Average steals per game
BLK	Average blocks per game
TOV	Average turnovers per game
PF	Average personal fouls per game
PTS	Average points per game
Year	Year of player statistics

Methods

The method for this analysis follows a distinct process. The process for this analysis:

- Create Dummy Variables
- Transform the Variables
- Check for Multicollinearity
- Perform Feature Selection
- Second Order and Interaction Terms
- Perform Residual Analysis
- Create Train-Test Cross Validation
- Evaluate Final Model

Visualize the Dataset & Transform the Variables

First, we analyzed by visualizing our dataset and creating distribution plots and histograms to check if the data follows assumptions such as normality, identify outliers, and gain a very first sense of data. We identified from the plots below (see appendix) that most of the variables have a right-skewed distribution. Therefore, we transformed the variables by using log, sqrt, square, or cube functions to normalize the distribution.

Variables of note that were consistently right-skewed across our data sets were (See Appendix for further details):

- Age
- Points
- Games Scored
- 3 Point Attempted
- 2 Point Made
- Free Throw Attempted
- Free Throw Made
- Offensive Rebounds
- Defensive Rebounds
- Assists
- Steals
- Turnovers

Check for Multicollinearity

Next, check if the data followed assumptions such as linearity and no multicollinearity by fitting a regression model to identify any potential issues that might impact the reliability of the model. Independent variables that are highly correlated can lead to overfitting of the analysis, where our model may not be accurate on unseen data. Given the data set has many variables, it is difficult to view scatterplots for all variables clearly and generate all correlations, so we checked multicollinearity by using VIF in car package, where a VIF value of greater than or equal to 10 should be removed due to its influence. After removing high VIF one by one, the results for all

variables were lower than 10, which suggests that there is no severe multicollinearity issue for now. Variables with high VIF values removed by position (See Appendix for further details):

- Point Guard: Minutes played, Field Goals (made), Field Goals Attempted, FG. (%), 3 Point Made, 2 Point Attempted, Free Throw Made, Total Rebounds.
- Shooting Guard: Field Goals (made), Field Goals Attempted, Free Throw Made, Total Rebounds, 3 Point Attempted, 3 Point Made, 2 Point Attempted, 2 Point Made.
- Small Forward: Field Goals (made), Field Goals Attempted, Free Throw Attempted, Total Rebounds, and Defensive Rebounds.
- Power Forward: 2 point made, rebounds, field goals, 2 points attempted, assist, 3 points attempted, free throw made, field goals, min played, age and steals.
- Center: Field Goals, Total Rebounds, 2 point made, free throw made, effective field goal, 3 point attempted, minute played, and steals

From a theoretical standpoint, field goals made, field goals attempted, 2 point made, 2 point attempted, 3 point made, 3 point attempted, free throws made, and free throws attempted could be removed because each has a variable for the efficiency expressed as a percentage. However, as variables are removed one at a time due to high VIF, other VIF dropped below 10 due to the absence of correlation between independent variables.

Perform Feature Selection

After transforming relevant variables and checking for multicollinearity, we proceeded to select features for the model. We conducted this by using both backward elimination and forward selection methods for each position. The primary goal in this section is to identify influential variables and create a first model that balances between explanatory power and simplicity.

- **Point Guard:** The model shows an impressive explanatory power of 98.72% as indicated by the adjusted R-squared. All variables present P-values below 0.05 (assuming that the significance level is 0.05), which means they are statistically significant at the conventional 0.05 level. It emphasizes the model's effectiveness in predicting scores. The selection of Assists and Defensive Rebounds are aligned with the distinctive role of a Point Guard. The model ensures simplicity without compromising predictive power.

Selected features: Age, Games, Games.Scored, X3.point.attempted, X3P, X2.Point.made, X2P, Effective.field.goal., Free.Throw.Attempted, Free.Throw., Offensive.rebounds, Defensive.rebounds, Assists, Personal.Fouls.

- **Shooting Guard:** The model shows an excellent explanatory ability of 91.53%. The P-values for most of the variables present are below 0.05, except for Offensive rebounds and Personal fouls. It emphasizes the model's effectiveness in predicting scores. The selecting method ensures the model captures the unique characteristics of a Shooting Guard - Effective Field Goal and Steals.

Selected features: Age, Games, Games.Scored, Min..Played, X3P., Effective.field.goal., FT., Offensive.rebounds, Defensive.rebounds, Assists, Steals, Blocks, Turnovers, Personal.Fouls.

- **Small Forward:** The chosen features are mostly significant (P-values < 0.05) with an exception of Offensive Rebounds, including Free Throw Made and Turnovers which are aligned with the unique characteristics of a Small Forward. The model achieves an excellent explanatory power of 92.02%, suggesting its ability to predict scores effectively.

Selected features: Games, Game.Scored, X3P., Effective.field.goal., Free.Throw.made, Offensive.rebounds, Defensive.rebounds, Turnovers, FG., Assists, Personal.Fouls, Blocks, Steals.

- **Power Forward:** Most of the variables show that P-values are below 0.05, except for Blocks. The explanatory power is impressive at 99.25%, emphasizing the model's effectiveness in prediction scores. Variables like Field Goals and Offensive Rebounds are expected and tailored to the Power Forward role, suggesting scoring and rebounding capabilities.

Selected features: Games, Games.Scored, Field.Goals, X3.Point.made, X3P., X2P., Effective.field.goal., Free.Throw.Attempted, Free.Throw., Offensive.rebounds, Defensive.rebounds, Blocks, Turnovers, Personal.Fouls.

- **Center:** Most of the variables show that P-values are below 0.05, except for Games and Personal fouls. The overall model passed the F-statistic with a P-value of < 2.2e-16. With an adjusted R-squared of 0.998, the model explains an impressive 99.8% of the variance in the dependent variable, which indicates an extremely high level of explanation provided by the model. The selections are tailored to Center role, such as Field Goals Attempted and Free Throw Attempted.

Selected features: Games, Field.Goals.Attempted, FG., X3.Point.made, X3P., X2P., Free.Throw.Attempted, FT., Assists, Games_FG, Games_X3P, FGA_FTA, Assist_FGA, Personal.Fouls.

Based on the result, only Shooting Guard selected Mins. Played, suggesting playing time duration may have a more significant impact on scoring for Shooting Guard. Only Small Forward not selected Free Throw, suggesting the ability to score through free throws might not be as critical for Small Forward. Only Power Forward not selected assist, indicating Power Forward may have less playmaking responsibilities. Only Center not selected Effective.field.goal, Offensive.rebounds, Defensive.rebounds, suggesting scoring efficiency and rebounding are not as directly linked to their scoring performance.

Variables that are selected by all positions are Games, X3P., and Personal Fouls. Games is likely to be selected universally because the number of games played is a fundamental metric for assessing player performance and contribution. Including X3P. suggests that successful 3-point shooting is considered crucial across all positions. It

indicates that long-range scoring is an important ability. All positions selected Personal Fouls could imply that referees' judgments and players' behavior in avoiding fouls are critical aspects for all positions.

This feature selection ensures our models capture the importance of each position while avoiding overfitting. The selected features represent key factors contributing to performance, providing valuable insights for predicting different basketball positions.

Second Order and Interaction Terms

After performing backward and feature selection on our models, we found second order and interaction terms for all positions except for the center. Since basketball is a sport where the player's performance can be influenced by several variables, we looked for non-linear relationships between independent and dependent variables- where a non-linear relationship may suggest that there may be a diminishing or increasing effect on points scored as the value increases. We looked for interaction terms where independent variables may jointly influence points scored. We found that:

- **Point Guard:** Assists and Games are squared terms. While assists contribute to points scored up to a certain level, the negative term shows there may be a diminishing return on points as assists increase. Games squared shows the potential that players may increase their points scored over time. Both the interaction between games/defensive rebounds and assist/X3P (negative) shows, when combined, the point guard's ability to score decreases.
- **Shooting Guard:** The shooting guard showed a significant minutes played squared term (negative) suggesting that as minutes played initially predicts an increase in points scored, there are diminishing returns as minutes played increased. The significant interaction between games scored and free throw percentage suggests both variables enhance their predictability on points scored.
- **Small Forward:** The small forward has significant variables for both games played and free throws made squared (both positive). Games squared reflects that playing more games predicts scoring points at an increasing and changing rate. However, beyond a certain point, the effect starts to diminish. Free throws made has the same effect.
- **Power Forward:** The power forward has three interaction terms between X3P/Def. Reb., Turnovers/Personal fouls, and Games/Field goals. These interaction terms have negative coefficient, suggesting that when combined, the increase in points scored is smaller than the increase that each variable has on points scored.

Perform Residual Analysis

A crucial stage in assessing a regression model's validity and performance is residual analysis. The variation between the values observed and the values the model predicts are known as residuals. The distribution of residuals may be evaluated with the use of a graphical depiction. Patterns or trends can be seen by plotting the residuals against each predictor variable or

against the expected values. Detailed residual analysis guarantees that the regression model fits the data well and that the assumptions of the model are fulfilled. In order to improve the model's predicted accuracy and dependability, it is imperative that any problems found during the study are fixed.

- **Point Guard :** The residuals seem to have a curved pattern. This shows that the data may contain heteroscedasticity, or non-constant variance of residuals, or non-linear associations, which would go against the principles of linear regression. There are two clear outliers with substantial positive and negative residuals, 1877 and 725, respectively. These could be significant findings that have a big influence on the regression's outcomes. As the fitted values grow, the bulk of the residuals cluster around zero, with increasing dispersion (variability). This further suggests heteroscedasticity because of the uneven variation among fitted values.
- **Shooting Guard:** Across the fitted value range, the residuals seem to be scattered around the zero horizontal line. This implies that the linearity and homoscedasticity (constant variance of residuals) conditions of linear regression are plausibly satisfied. The absence of a clear pattern or trend in the residuals suggests that the data do not have any problems with heteroscedasticity or non-linear connections. Further evidence for homoscedasticity comes from the residuals' apparent consistent distribution over the fitted value range. There don't seem to be any notable outliers that may significantly affect the regression findings, despite a few bigger residual values. To fulfill the residuals' normality assumption, the histogram of the residuals in the bottom panel should roughly have a normal distribution.
- **Small Forward:** For the majority of fitted values, the residuals often show a random scatter pattern around the zero line, suggesting that the linearity assumption is fairly satisfied. At higher fitted values, there appears to be a small fanning or megaphone pattern, with the residuals' variability rising in spread. This may indicate residual heteroscedasticity or non-constant variance. Although the majority of residuals are clustered close to zero, certain bigger residual values—both positive and negative—may be outliers that have an impact on the model's fit.
- **Power Forward:** The red line, which represents a nonlinear pattern, indicates that there may not be a fully linear relationship between the predictors and the result variable. A prominent curve in the figure suggests that the model might not be appropriate. Potential breaches of the linearity and homoscedasticity assumptions of the linear regression model are indicated by the residuals vs. fitted plot. Additionally, outliers might have an impact on the model. It is advised to think about using a non-linear model, adding interaction terms, or changing the variables.
- **Center:** The residuals' random dispersion around the horizontal zero line across the fitted value range suggests that the linear regression model's linearity condition is probably satisfied. Homoscedasticity, or the constant variance of residuals, appears to

be satisfied since the dispersion of the residuals appears to be generally consistent throughout the fitted values.

Create Train-Test Cross-Validation

Cross-validation is a crucial tool in regression analysis for assessing model performance and ensuring that the model generalizes well to new data. The process involves partitioning the dataset into multiple subsets, training the model on some of these subsets, and then evaluating its performance on the remaining subset(s).

Cross-validation working:

Splitting the Data: The dataset is divided into a training set and a testing set. Common splits include a 70-30 or 80-20 ratio, where the larger portion is used for training and the smaller for testing.

Training the Model: The regression model is trained on the training set. This involves fitting the model to the training data and adjusting its parameters to minimize prediction errors.

Testing the Model: The trained model is then used to make predictions on the testing set. The predicted values are compared to the actual values in the testing set to evaluate the model's performance.

Technique:

K-Fold Cross-Validation: The dataset is divided into k equal-sized folds. The model is trained k times, each time using $k-1$ folds for training and the remaining fold for testing. The performance metrics are averaged over the k iterations.

Assessing Performance: Common metrics used to assess the performance of a regression model include mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and R-squared (coefficient of determination).

Position	PG	SG	SF	PF	C
RMSE	0.062	0.183	0.167	0.122	0.026
Adjusted R squared	0.989	0.921	0.922	0.959	0.998
correlation coefficient	0.994	0.9556	0.922		0.998

RMSE (Root Mean Squared Error): the values are relatively low across all positions, suggesting that the model's predictions are generally close to the actual values.

Adjusted R-squared: The values for all positions range from 0.921 to 0.998, which indicates that the model explains a significant amount of the variability.

Performed on center

Model	RMSE	Rsquared	MAE
Linear Regression	0.02630964	0.9980099	0.01736849
Ridge Regression	0.02630931	0.9980102	0.01735987
Lasso Regression	0.02680864	0.9979380	0.01761774

3 rows

All three regression techniques perform well on the given dataset, with very low errors and high R-squared values. Ridge Regression and Lasso Regression provide regularization benefits, potentially improving generalization to unseen data. Ridge Regression offers results very similar to Linear Regression, while Lasso Regression shows slight differences due to its feature selection properties.

Evaluate Final Model

When evaluating the final models for the positions being analyzed, we examine the model's key aspects to determine model accuracy, fit, and overall significance. Indicators we look for are:

1. Residual standard error which measures the standard deviation of the residuals, the differences between observed and predicted values. A lower number indicates less variance in the residuals and a better-fit model.
 2. F-statistic and its p-value is examined to test the overall significance of the model. A p-value less than .05 suggests that we can reject the null hypothesis that all coefficients are equal to 0.
 3. T-statistic p-value for the predictors tells us whether our predictor is statistically significant when predicting point scored. Using an alpha of .05, a p-value less than .05 suggests a meaning relationship with the variable and points scored.
 4. Adjusted r-squared is evaluated to determine how much of the variance in points scored is explained by the model's independent variables. A higher r-squared value indicates a better fit of the data.
- **Point Guard:** RSE .06262, F-stat p-value < 2.2e-16, r-squared .9895. This model is a very strong fit with a low average distance between observed and predicted values and at least one predictor that is significant when predicting points.
 - **Shooting Guard:** RSE .1786, F-stat p-value < 2.2e-16, r-squared .9209. This model is a very strong fit with a high r-squared value. The average distance between observed and predicted values is also low at .1786, and all predictors are in the .05 significance range.

- **Small Forward:** RSE .1722, F-stat p-value < 2.2e-16, r-squared .921. The small forward model is also a very strong fit with an r-squared of .921. An average distance between observed and predicted values is .1722, relatively small for this model. All predictors are below the .05 significance level and the f-stat p-value shows that all coefficients are not equal to 0.
- **Power Forward:** RSE .04832, F-stat: 2.181e+04, and the p-value: < 2.2e-16. The power forward model is a very strong fit with r square .9925 with low average between observed and predicted values.
- **Center:** RSE .02952, the high F-stat value of 5.715e+04 and corresponding p-value (< 2.2e-16) suggest that the overall regression model is significant and the adjusted R-squared value of .9974 analysis demonstrates a highly significant and robust model for predicting basketball players' points based on various performance metrics.

Conclusion

Our regression analysis reveals statistically significant insights into the various positions- points guard, shooting guard, small forward, power forward, and center- that player's perform and what impacts their ability to score more points. Key findings show that shooting efficiency, especially effective field goal percentage, has a significant impact on scoring points across all positions. Contrary to what one might think for an offensive position, the point guard could score more with higher defensive rebounding stats, and assisting significantly contributes to points scored. The shooting guard, another primarily offensive position, sees a boost in points scored with more efficient shooting and experience, where free throw percentage and games played can be an important factor. Small forwards scoring is heavily impacted by free throw accuracy and three point scoring. Power forwards, more of a defensive position, may look to draw more fouls, where more free throw attempts leads to more scoring, and defensive rebounds are critical. Centers should minimize turnovers and put more focus on offensive plays like three point shots.

Overall, each position can benefit from a unique set of attributes when scoring more points, and coaches and teams alike could align their strategies to reflect the significant impact of variables mentioned in this report.

Appendix

Point Guard

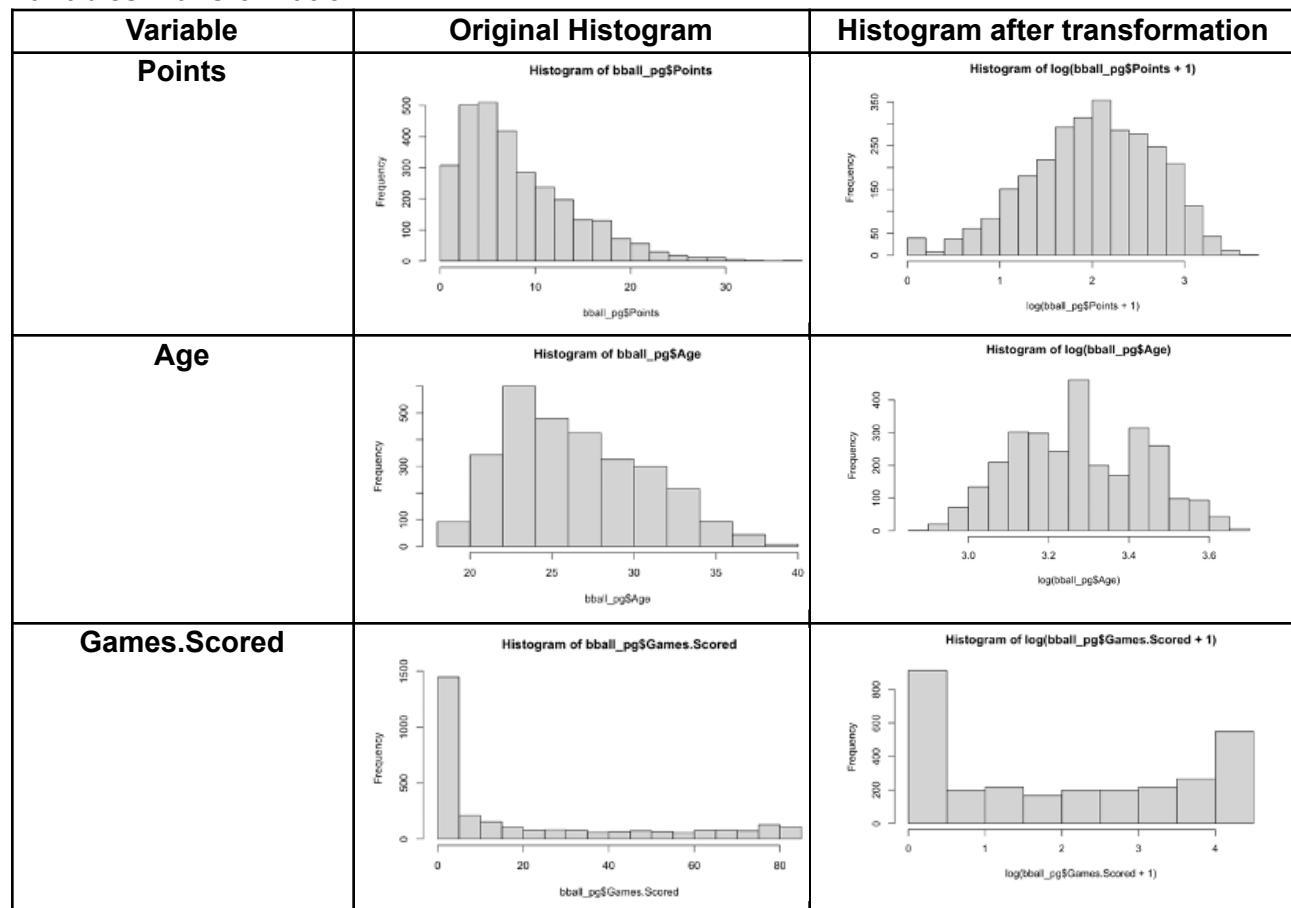
Initial VIF

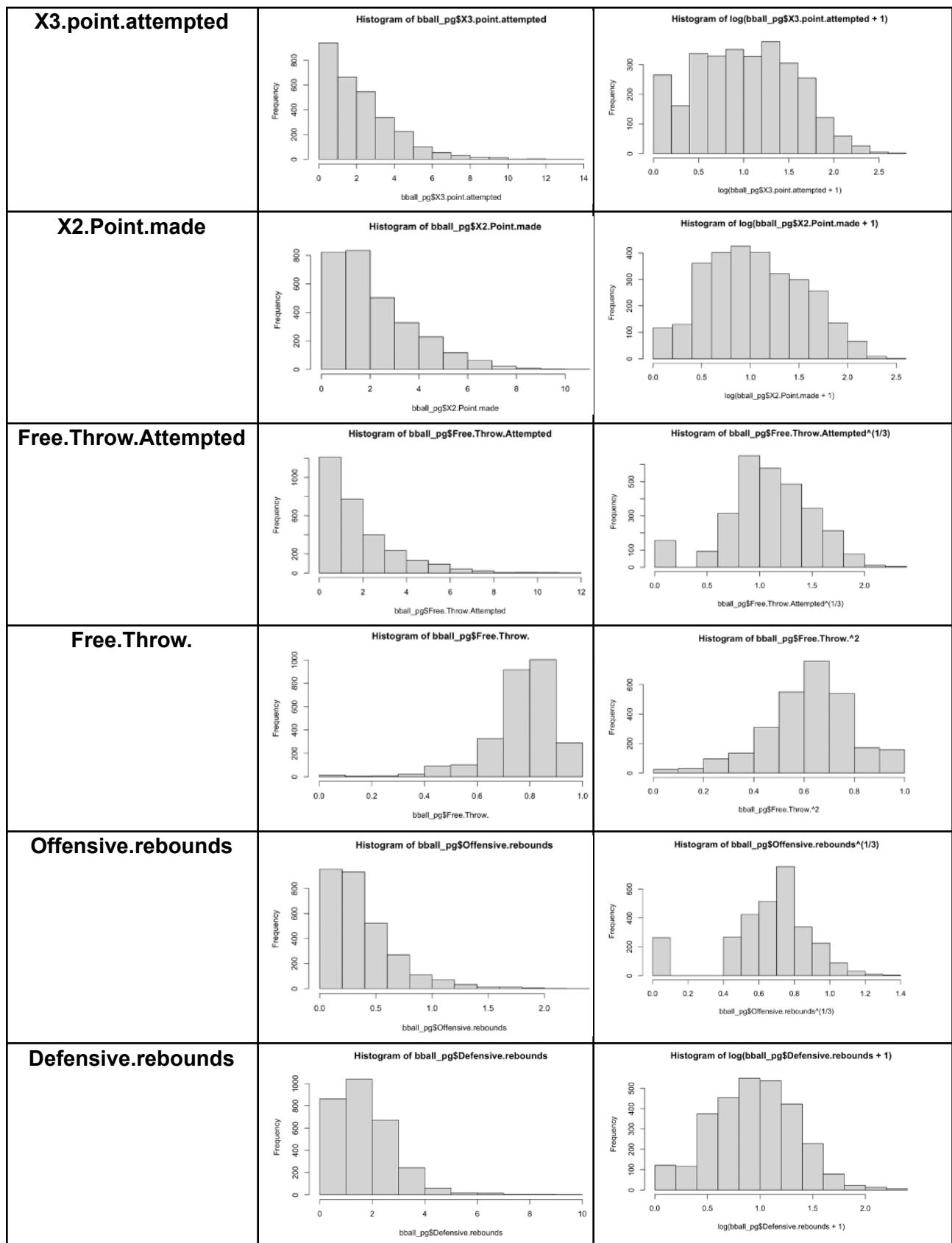
	Age	Games	Games.Scored	Min..Played
	1.165090	1.956263	4.685335	19.188900
Field.Goals	Field.Goals.Attempted		FG.	X3.Point.made
518.079133	372.418021		159.362447	59.401050
X3.point.attempted	X3P.		X2.Point.made	X2.Point.Attempted
58.855089	3.901643		297.742104	298.147851
X2P. Effective.field.goal..				
12.351847	132.944706		66.787632	64.335075
Free.Throw.	Offensive.rebounds		Free.Throw.made	Free.Throw.Attempted
2.961556	7.589057		103.378073	Total.Rebounds
Assists	Steals		Blocks	Turnovers
7.867159	3.405131		1.643507	6.769025
Personal.Fouls				
	2.744168			

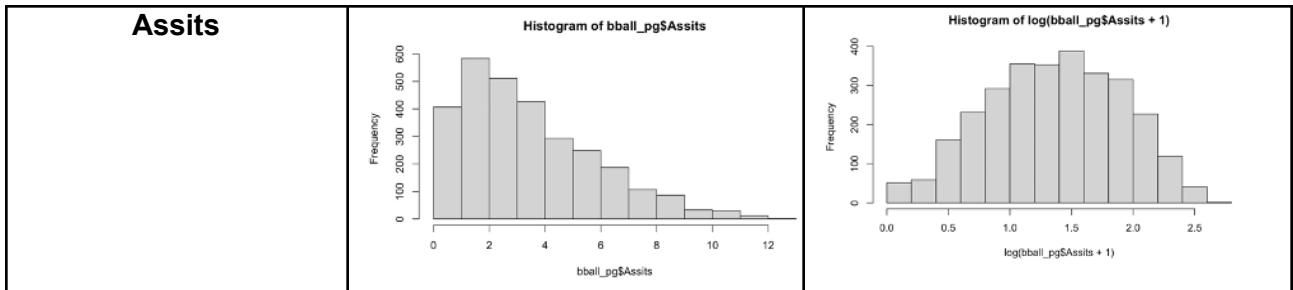
Final VIF result for the model

	Age	Games	Games.Scored	X3.point.attempted
	1.129351	1.867289	3.769149	1.964766
	X3P.	X2.Point.made	X2P. Effective.field.goal..	
	2.339669	6.758880	3.908322	5.404048
Free.Throw.Attempted	Free.Throw.	Offensive.rebounds	Defensive.rebounds	
3.826875	1.115568	2.048391	5.023711	
Assists	Steals	Blocks	Turnovers	
7.023532	3.240511	1.632738	6.728954	
Personal.Fouls				
	2.470862			

Variables Transformation







Result for backward elimination & forward selection

Call:

```
lm(formula = Points ~ Age + Games + Games.Scored + X3.point.attempted +
X3P. + X2.Point.made + X2P. + Effective.field.goal.. + Free.Throw.Attempted +
Free.Throw. + Offensive.rebounds + Defensive.rebounds + Assists +
Personal.Fouls, data = bball_pg)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.50254	-0.03431	0.00261	0.03741	0.32668

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.252e-01	3.041e-02	-4.118	3.93e-05 ***
Age	4.164e-02	8.861e-03	4.700	2.73e-06 ***
Games	4.151e-04	7.251e-05	5.725	1.15e-08 ***
Games.Scored	-5.111e-03	1.537e-03	-3.325	0.000896 ***
X3.point.attempted	3.164e-01	3.457e-03	91.509	< 2e-16 ***
X3P.	-1.105e-01	1.683e-02	-6.565	6.19e-11 ***
X2.Point.made	7.465e-01	7.141e-03	104.537	< 2e-16 ***
X2P.	-9.824e-01	3.255e-02	-30.184	< 2e-16 ***
Effective.field.goal..	1.768e+00	4.138e-02	42.737	< 2e-16 ***
Free.Throw.Attempted	3.247e-01	7.714e-03	42.088	< 2e-16 ***
Free.Throw.	1.108e-01	7.869e-03	14.076	< 2e-16 ***
Offensive.rebounds	3.552e-02	8.233e-03	4.315	1.66e-05 ***
Defensive.rebounds	2.104e-02	7.642e-03	2.754	0.005932 **
Assists	2.838e-02	5.694e-03	4.984	6.62e-07 ***
Personal.Fouls	2.327e-02	2.858e-03	8.142	5.87e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Residual standard error: 0.06887 on 2708 degrees of freedom

(201 observations deleted due to missingness)

Multiple R-squared: 0.9872, Adjusted R-squared: 0.9872

F-statistic: 1.497e+04 on 14 and 2708 DF, p-value: < 2.2e-16

Final Model

Call:

```
lm(formula = Points ~ Age + Games + Games.Scored + X3.point.attempted +
  X3P. + X2.Point.made + X2P. + Effective.field.goal.. + Free.Throw.Attempted +
  Free.Throw. + Offensive.rebounds + Defensive.rebounds + Assists +
  Personal.Fouls + Games_DefReb + Assists_X3P + Assists2 +
  Games2, data = bball_pg)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.48732	-0.03260	-0.00388	0.03063	0.34177

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)		
(Intercept)	-2.894e-01	2.910e-02	-9.945	< 2e-16 ***		
Age	3.611e-02	8.076e-03	4.471	8.10e-06 ***		
Games	1.347e-03	2.426e-04	5.554	3.07e-08 ***		
Games.Scored	2.640e-03	1.440e-03	1.834	0.06681 .		
X3.point.attempted	3.181e-01	3.350e-03	94.949	< 2e-16 ***		
X3P.	-1.048e-02	2.888e-02	-0.363	0.71674		
X2.Point.made	7.390e-01	6.508e-03	113.566	< 2e-16 ***		
X2P.	-9.293e-01	3.028e-02	-30.692	< 2e-16 ***		
Effective.field.goal..	1.692e+00	3.876e-02	43.656	< 2e-16 ***		
Free.Throw.Attempted	3.560e-01	7.164e-03	49.689	< 2e-16 ***		
Free.Throw.	1.114e-01	7.162e-03	15.550	< 2e-16 ***		
Offensive.rebounds	1.657e-02	7.589e-03	2.183	0.02912 *		
Defensive.rebounds	1.168e-01	9.920e-03	11.772	< 2e-16 ***		
Assists	2.221e-01	1.439e-02	15.436	< 2e-16 ***		
Personal.Fouls	1.111e-02	2.663e-03	4.172	3.12e-05 ***		
Games_DefReb	-1.929e-03	1.702e-04	-11.333	< 2e-16 ***		
Assists_X3P	-9.135e-02	2.154e-02	-4.241	2.30e-05 ***		
Assists2	-6.395e-02	4.612e-03	-13.867	< 2e-16 ***		
Games2	7.723e-06	2.691e-06	2.869	0.00415 **		

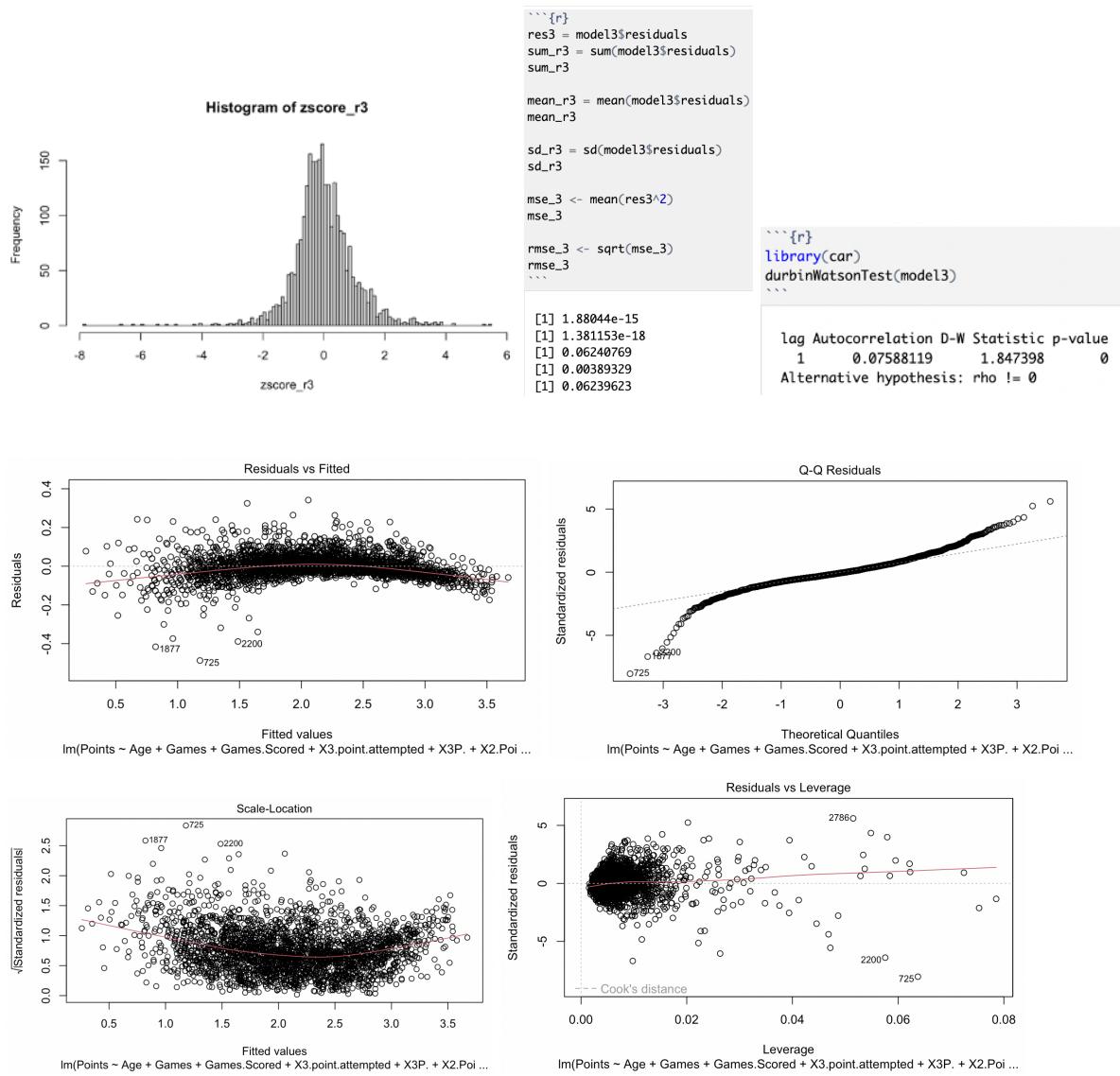
Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘ ’	1

Residual standard error: 0.06262 on 2704 degrees of freedom

Multiple R-squared: 0.9895, Adjusted R-squared: 0.9894

F-statistic: 1.412e+04 on 18 and 2704 DF, p-value: < 2.2e-16

Result for residual analysis



Result for cross-validation

```

```{r}
partition2 <- sample(2, nrow(bball_pg), replace = TRUE, prob = c(0.80, 0.20))
head(partition2)
train2 <- bball_pg[partition == 1 ,]
test2 <- bball_pg[partition == 2 ,]
```

```

Shooting Guard

Final Model

```

Call:
lm(formula = Points ~ Age + Games + Games.Scored + Games.Scored *
    FT. + Min. Played + I(Min..Played^2) + X3P. + Effective.field.goal.. +
    FT. + Offensive.rebounds + Assists + Steals + Turnovers +
    Personal.Fouls, data = SGProject_clean)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.78101 -0.11080  0.00363  0.12001  0.75002 

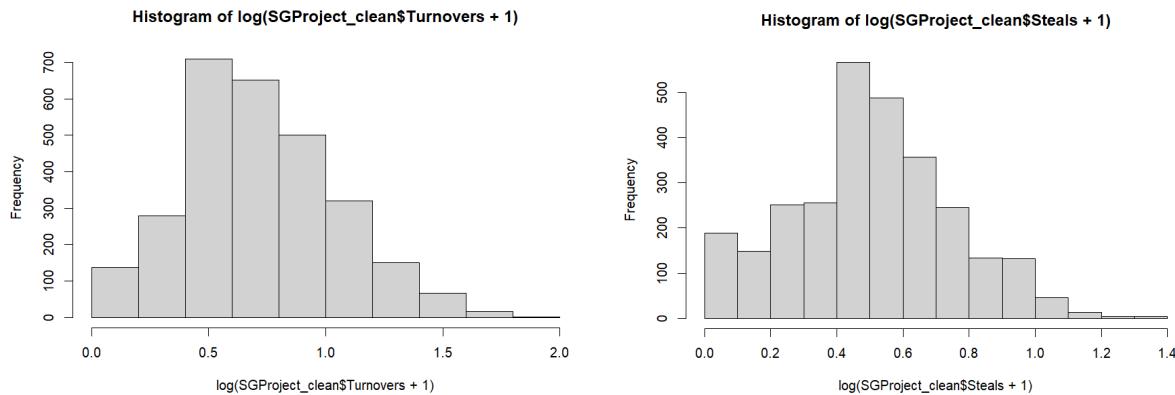
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.4618377  0.0806953  5.723 1.16e-08 ***
Age          -0.1934678  0.0239663 -8.072 1.01e-15 ***
Games        0.0003744  0.0001860  2.013 0.044239 *  
Games.Scored -0.1233295  0.0161657 -7.629 3.21e-14 ***
FT.          0.2271905  0.0315433  7.203 7.55e-13 ***
Min.Played   0.0810745  0.0023362 34.703 < 2e-16 ***
I(Min..Played^2) -0.0006344  0.0000436 -14.552 < 2e-16 ***
X3P.         0.1337392  0.0414062  3.230 0.001252 ** 
Effective.field.goal.. 1.2979439  0.0619456 20.953 < 2e-16 ***
Offensive.rebounds 0.0672507  0.0201246  3.342 0.000843 *** 
Assists       -0.0767228  0.0168247 -4.560 5.33e-06 *** 
Steals        -0.0773150  0.0241140 -3.206 0.001360 ** 
Turnovers     0.5441962  0.0225986 24.081 < 2e-16 ***
Personal.Fouls -0.1660342  0.0221586 -7.493 8.96e-14 *** 
Games.Scored:FT. 0.1008069  0.0202410  4.980 6.73e-07 *** 

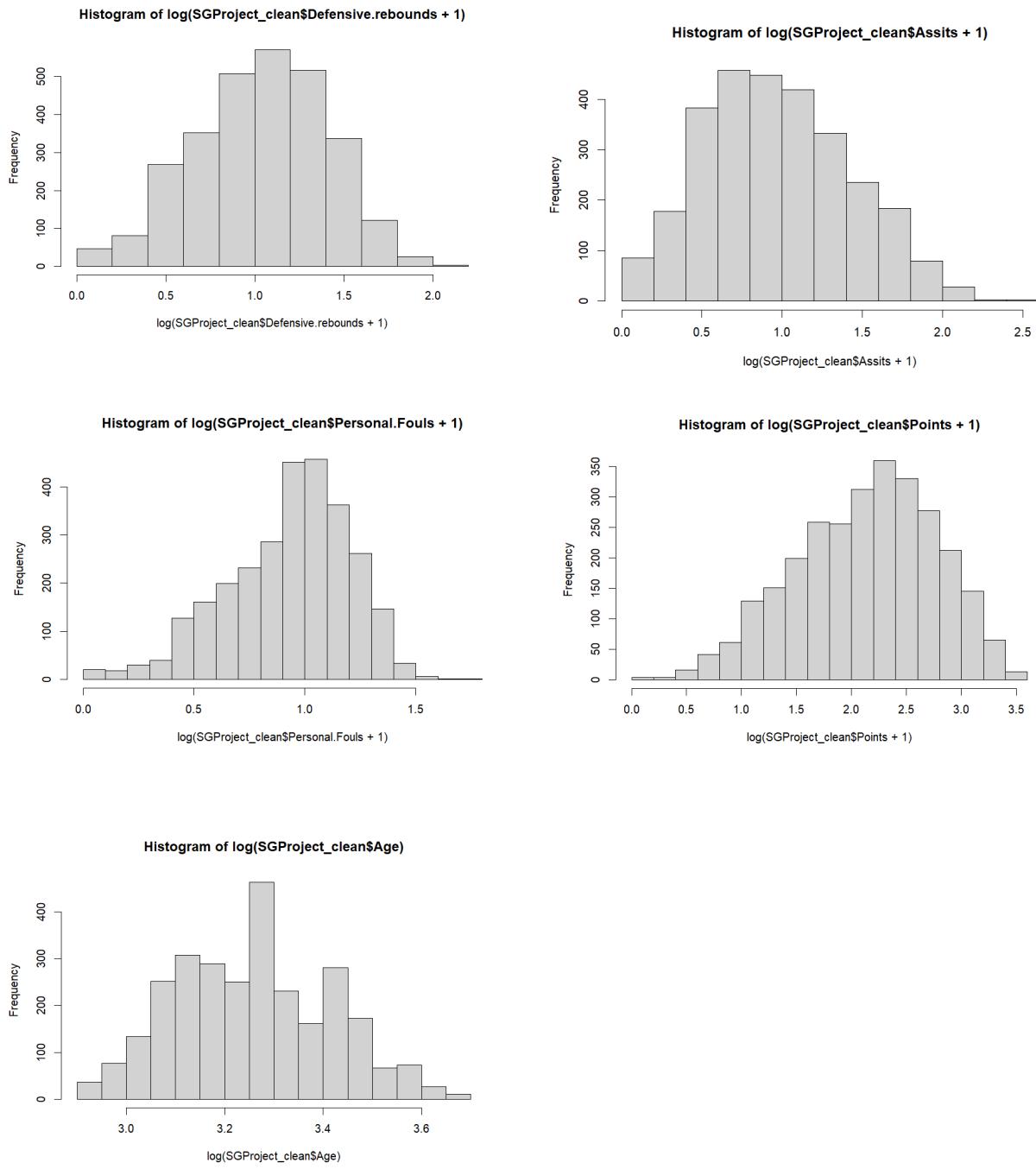
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1786 on 2816 degrees of freedom
Multiple R-squared:  0.9213, Adjusted R-squared:  0.9209 
F-statistic: 2354 on 14 and 2816 DF, p-value: < 2.2e-16

```

Histograms





Initial VIF

vif(model)

| Age | Games | Games.Scored | Min..Played | Field.Goals |
|-----------------------|--------------------|--------------------|------------------------|-----------------|
| 1.180881 | 1.786479 | 3.197179 | 14.727634 | 2111.379528 |
| Field.Goals.Attempted | FG. | X3.Point.made | X3.point.attempted | X3P. |
| 9298.538830 | 19.750962 | 275.193369 | 1516.956659 | 2.979446 |
| X2.Point.made | X2.Point.Attempted | X2P. | Effective.field.goal.. | Free.Throw.made |
| 1437.598165 | 5642.836651 | 3.671103 | 21.649085 | 124.515076 |
| Free.Throw.Attempted | FT. | Offensive.rebounds | Defensive.rebounds | Total.Rebounds |
| 129.881860 | 1.847291 | 65.687624 | 523.854049 | 800.621321 |
| Assits | Steals | Blocks | Turnovers | Personal.Fouls |
| 4.834327 | 3.012012 | 1.612050 | 7.075091 | 2.921075 |

Final VIF

vif(model1)

| Age | Games | Games.Scored | Min..Played | FG. |
|--------------------|----------|------------------------|-------------|--------------------|
| 1.144484 | 1.770236 | 2.928522 | 9.179361 | 7.579918 |
| X3P. | X2P. | Effective.field.goal.. | FT. | Offensive.rebounds |
| 2.706359 | 2.769456 | 9.926750 | 1.184323 | 2.157737 |
| Defensive.rebounds | Assits | Steals | Blocks | Turnovers |
| 3.844213 | 4.703069 | 2.939463 | 1.598655 | 5.099181 |
| Personal.Fouls | | | | |
| 2.738520 | | | | |

Backward and Forward Model

```

Call:
lm(formula = Points ~ Age + Games + Games.Scored + Min.Played +
X3P. + Effective.field.goal... + FT. + Offensive.rebounds. +
Defensive.rebounds. + Assists + Steals + Blocks + Turnovers +
Personal.Fouls, data = SGProject_clean)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.90818 -0.11469  0.00473  0.12601  0.73011 

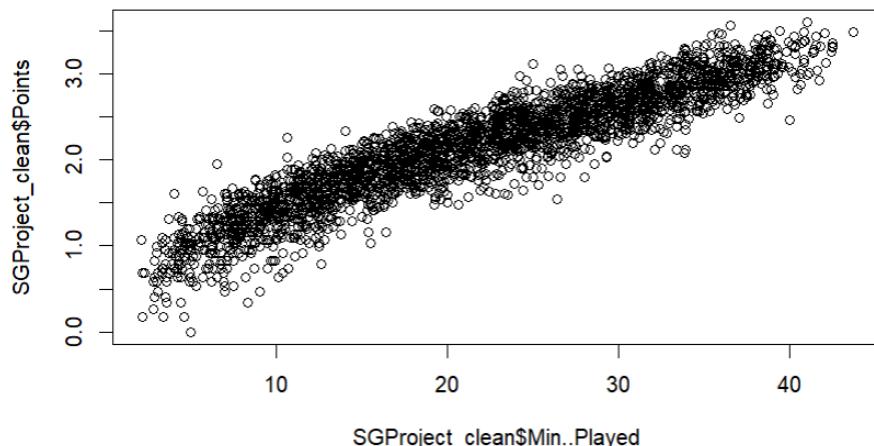
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.3285894  0.0819202   4.011 6.20e-05 ***
Age         -0.1453323  0.0245843  -5.912 3.80e-09 ***
Games        0.0007716  0.0001902   4.058 5.09e-05 ***
Games.Scored -0.0490121  0.0041391  -11.841 < 2e-16 ***
Min.Played   0.0496682  0.0011631   42.704 < 2e-16 ***
X3P.          0.1542296  0.0478696   3.598 0.000327 ***
Effective.field.goal... 1.4057702  0.0641154   21.926 < 2e-16 ***
FT.           0.3382916  0.0277734   12.180 < 2e-16 ***
Offensive.rebounds 0.0400750  0.0212982   1.882 0.059991 .
Defensive.rebounds 0.0660996  0.0196418   3.365 0.000775 ***
Assists       -0.0563818  0.0173975  -3.241 0.001206 ** 
Steals         -0.0844044  0.0252609  -3.341 0.000845 ***
Blocks         -0.0653101  0.0301771  -2.164 0.030531 *  
Turnovers      0.4964893  0.0231186   21.476 < 2e-16 ***
Personal.Fouls -0.0373498  0.0209808  -1.780 0.075152 . 

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

Residual standard error: 0.1848 on 2816 degrees of freedom
Multiple R-squared:  0.9157,  Adjusted R-squared:  0.9153
F-statistic: 2185 on 14 and 2816 DF,  p-value: < 2.2e-16

```

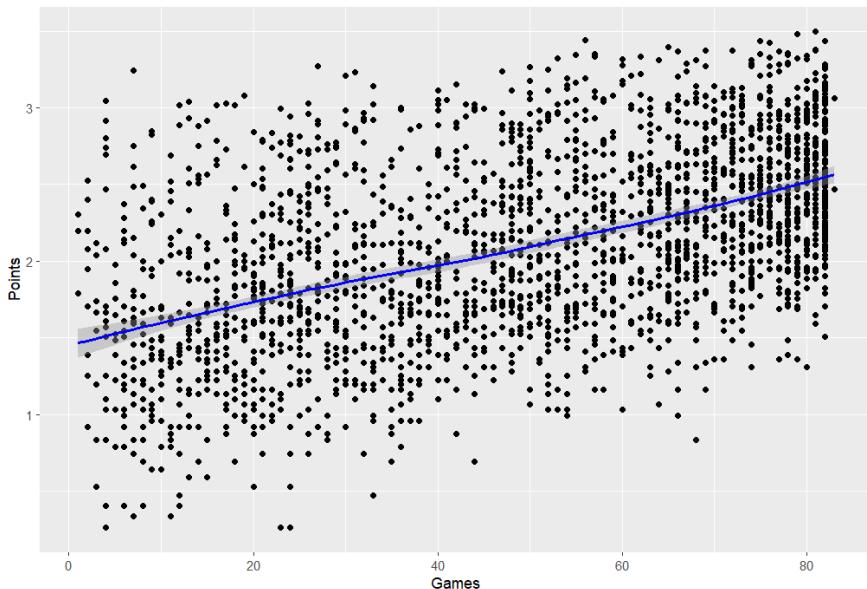
Minutes played on Points



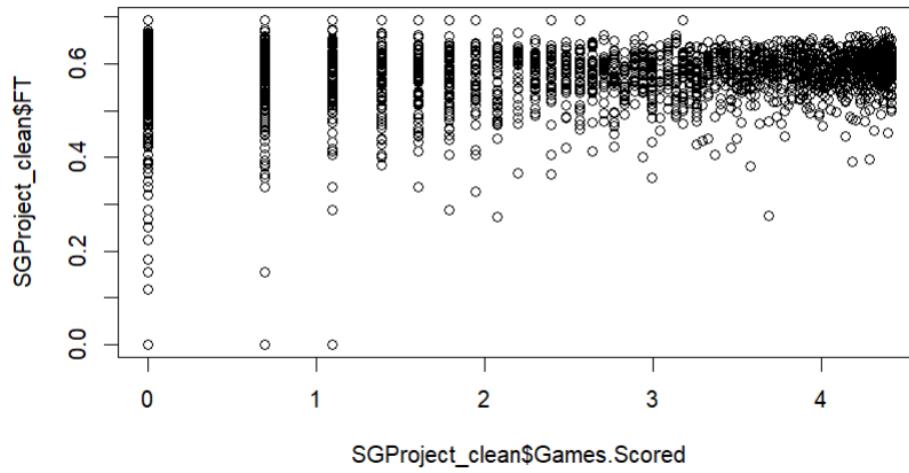
Squaring Minutes Played

```
Call:  
lm(formula = Points ~ Age + Games + Games.Scored + Min.Played +  
  I(Min.Played^2) + X3P. + Effective.field.goal.. + FT. +  
  Offensive.rebounds + Defensive.rebounds + Assists + Steals +  
  Blocks + Turnovers + Personal.Fouls, data = SGProject_clean)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-0.80126 -0.10945  0.00343  0.12095  0.76060  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 3.886e-01 7.961e-02 4.881 1.11e-06 ***  
Age          -1.876e-01 2.407e-02 -7.797 8.87e-15 ***  
Games        3.821e-04 1.868e-04  2.045 0.04094 *  
Games.Scored -4.502e-02 4.027e-03 -11.180 < 2e-16 ***  
Min.Played   7.946e-02 2.512e-03 31.632 < 2e-16 ***  
I(Min.Played^2) -5.848e-04 4.405e-05 -13.276 < 2e-16 ***  
X3P.         1.356e-01 4.162e-02  3.259 0.00113 **  
Effective.field.goal.. 1.312e+00 6.261e-02 20.958 < 2e-16 ***  
FT.          3.075e-01 2.705e-02 11.370 < 2e-16 ***  
Offensive.rebounds 5.623e-02 2.070e-02  2.716 0.00664 **  
Defensive.rebounds 1.049e-03 1.968e-02  0.053 0.95750  
Assists      -7.454e-02 1.694e-02 -4.401 1.12e-05 ***  
Steals       -8.028e-02 2.451e-02 -3.275 0.00107 **  
Blocks       -4.363e-02 2.933e-02 -1.488 0.13689  
Turnovers    5.438e-01 2.271e-02 23.942 < 2e-16 ***  
Personal.Fouls -1.602e-01 2.236e-02 -7.164 9.97e-13 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.1793 on 2815 degrees of freedom  
Multiple R-squared:  0.9207, Adjusted R-squared:  0.9202  
F-statistic: 2178 on 15 and 2815 DF, p-value: < 2.2e-16
```

Check for Games to be Squared



Interaction of FT and Games Scored



Interacting FT and Games Scored

```

Call:
lm(formula = Points ~ Age + Games + Games.Scored + Games.Scored *
    FT. + Min..Played + I(Min..Played^2) + X3P. + Effective.field.goal.. +
    FT. + Offensive.rebounds + Defensive.rebounds + Assists +
    Steals + Blocks + Turnovers + Personal.Fouls, data = SGProject_clean)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.77807 -0.11072  0.00344  0.12003  0.74194 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 4.630e-01  8.076e-02  5.733 1.09e-08 ***
Age          -1.946e-01  2.401e-02 -8.103 7.90e-16 ***
Games        3.695e-04  1.861e-04  1.986 0.047165 *  
Games.Scored -1.216e-01  1.625e-02 -7.484 9.57e-14 ***
FT.          2.276e-01  3.155e-02  7.213 7.00e-13 ***
Min..Played  8.096e-02  2.521e-03 32.115 < 2e-16 ***
I(Min..Played^2) -6.322e-04  4.494e-05 -14.066 < 2e-16 ***
X3P.         1.322e-01  4.146e-02  3.188 0.001450 ** 
Effective.field.goal.. 1.302e+00  6.240e-02 20.859 < 2e-16 ***
Offensive.rebounds  7.004e-02  2.081e-02  3.365 0.000775 *** 
Defensive.rebounds 1.873e-03  1.960e-02  0.096 0.923854  
Assists       -7.689e-02  1.687e-02 -4.557 5.42e-06 ***
Steals        -7.294e-02  2.446e-02 -2.982 0.002889 ** 
Blocks        -3.132e-02  2.932e-02 -1.068 0.285402  
Turnovers     5.448e-01  2.262e-02 24.082 < 2e-16 ***
Personal.Fouls -1.636e-01  2.228e-02 -7.340 2.77e-13 *** 
Games.Scored:FT.  9.886e-02  2.033e-02  4.864 1.22e-06 *** 

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1786 on 2814 degrees of freedom
Multiple R-squared:  0.9213, Adjusted R-squared:  0.9209 
F-statistic: 2059 on 16 and 2814 DF, p-value: < 2.2e-16

```

Deleting Blocks and Defensive Rebounds due to Insignificance

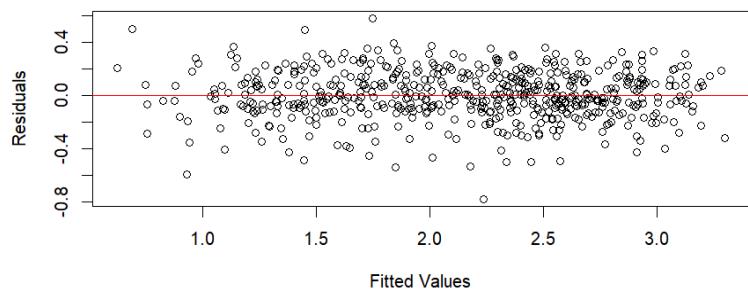
```
Call:  
lm(formula = Points ~ Age + Games + Games.Scored + Games.Scored *  
  FT. + Min..Played + I(Min..Played^2) + X3P. + Effective.field.goal.. +  
  FT. + Offensive.rebounds + Assists + Steals + Turnovers +  
  Personal.Fouls, data = SGProject_clean)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-0.78101 -0.11080  0.00363  0.12001  0.75002  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) 0.4618377  0.0806953  5.723 1.16e-08 ***  
Age          -0.1934678  0.0239663 -8.072 1.01e-15 ***  
Games         0.0003744  0.0001860  2.013 0.044239 *  
Games.Scored -0.1233295  0.0161657 -7.629 3.21e-14 ***  
FT.           0.2271905  0.0315433  7.203 7.55e-13 ***  
Min..Played   0.0810745  0.0023362 34.703 < 2e-16 ***  
I(Min..Played^2) -0.0006344  0.0000436 -14.552 < 2e-16 ***  
X3P.          0.1337392  0.0414062  3.230 0.001252 **  
Effective.field.goal.. 1.2979439  0.0619456 20.953 < 2e-16 ***  
Offensive.rebounds 0.0672507  0.0201246  3.342 0.000843 ***  
Assists        -0.0767228  0.0168247 -4.560 5.33e-06 ***  
Steals          -0.0773150  0.0241140 -3.206 0.001360 **  
Turnovers       0.5441962  0.0225986 24.081 < 2e-16 ***  
Personal.Fouls  -0.1660342  0.0221586 -7.493 8.96e-14 ***  
Games.Scored:FT. 0.1008069  0.0202410  4.980 6.73e-07 ***  
---  
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
Residual standard error: 0.1786 on 2816 degrees of freedom  
Multiple R-squared:  0.9213, Adjusted R-squared:  0.9209  
F-statistic: 2354 on 14 and 2816 DF, p-value: < 2.2e-16
```

Cross Validation Results

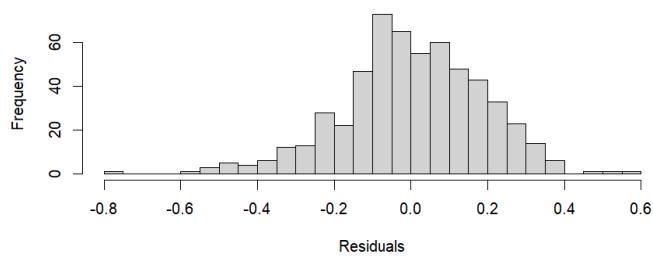
```
# Load the necessary library  
> library(caret)  
> # Setting a seed for reproducibility  
> set.seed(123)  
> # Splitting the data into training (80%) and testing (20%) sets  
> trainIndex <- createDataPartition(SGProject_clean$Points, p = .8,  
+                                     list = FALSE,  
+                                     times = 1)  
> trainSet <- SGProject_clean[trainIndex, ]  
> testSet <- SGProject_clean[-trainIndex, ]  
> # Fitting the model on the training set  
> trainTestModel <- lm(Points ~ Age + Games + Games.Scored + Games.Scored * F  
T. +  
+               Min..Played + I(Min..Played^2) + X3P. + Effective.field.goal..  
+               FT. + Offensive.rebounds + Assists + Steals + Turnovers +  
+               Personal.Fouls,  
+               data = trainSet)  
> # Predicting on the test set  
> predictions <- predict(model15, newdata = testSet)  
> # Calculating RMSE for the test set predictions  
> rmse_test <- sqrt(mean((predictions - testSet$Points)^2))  
> # Print RMSE  
> print(paste("RMSE on Test Set:", rmse_test))  
[1] "RMSE on Test Set: 0.183316171693796"  
  
# Assuming 'actuals' contains the actual Points from the test set  
# and 'predictions' contains the predicted Points from your model  
> actuals <- testSet$Points  
> predictions <- predict(model15, newdata = testSet)  
> # Calculate correlation  
> correlation_coefficient <- cor(actuals, predictions)  
> print(paste("Correlation coefficient:", correlation_coefficient))  
[1] "Correlation coefficient: 0.955694812691538"
```

Residual Analysis

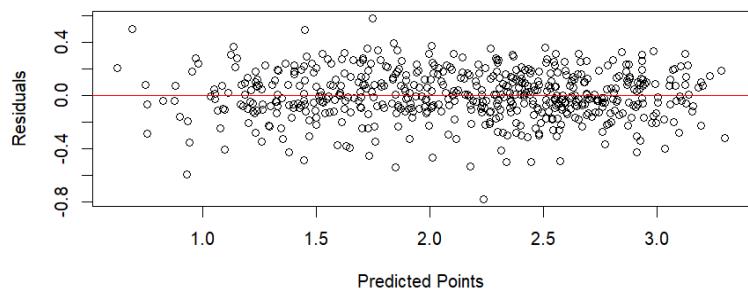
Residuals vs. Fitted Values



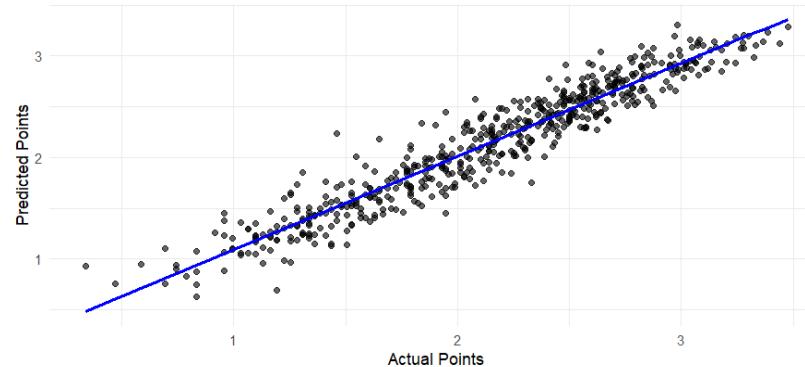
Histogram of Residuals



Residuals vs. Predicted



Actual vs. Predicted Points



Small Forward

Final Model

```

Call:
lm(formula = Points ~ Free.Throw.made + Games + I(Games^2) +
   I(Free.Throw.made^2) + Effective.field.goal.. + Turnovers +
   FG. + Games.Scored + Blocks + Personal.Fouls + Steals + Assists +
   Defensive.rebounds + X3P., data = sfclean)

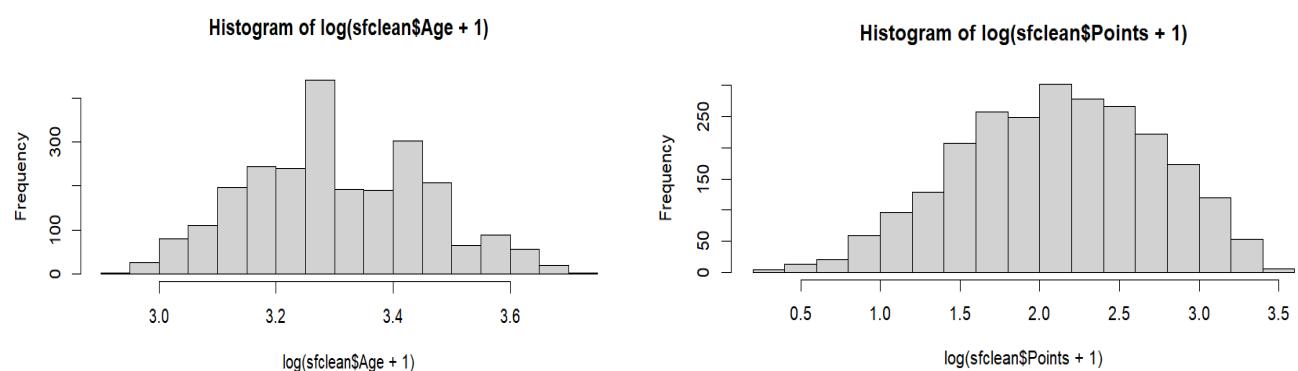
Residuals:
    Min      1Q  Median      3Q     Max 
-0.87087 -0.11036  0.00265  0.10930  0.75260 

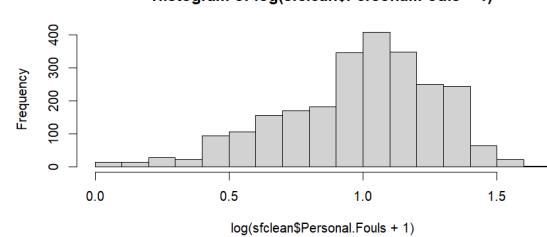
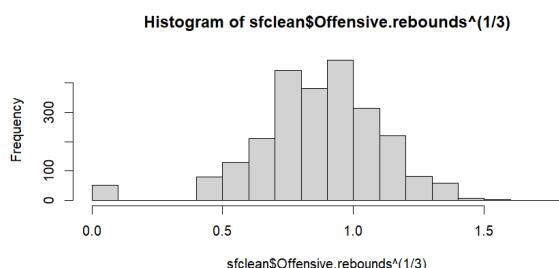
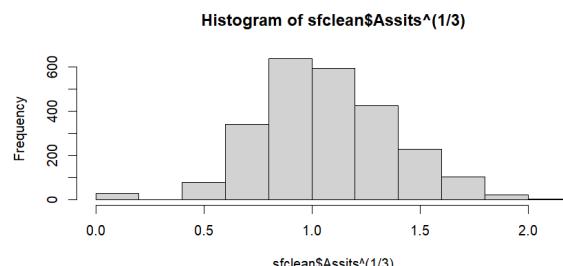
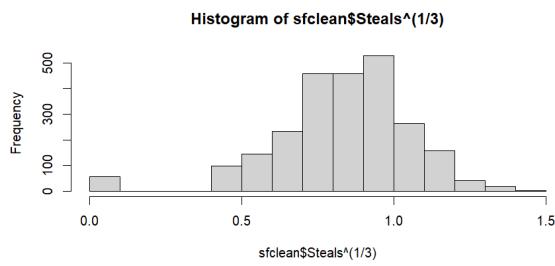
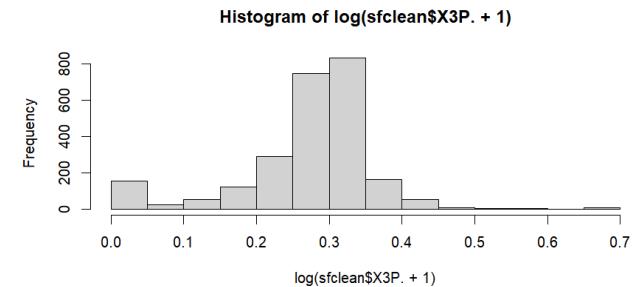
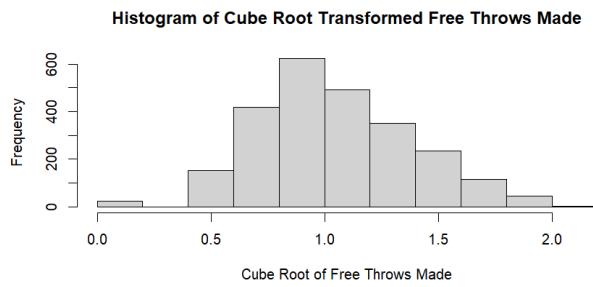
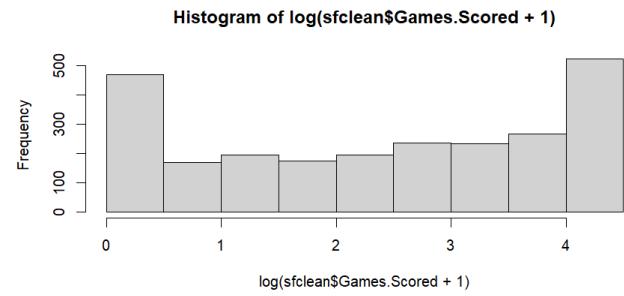
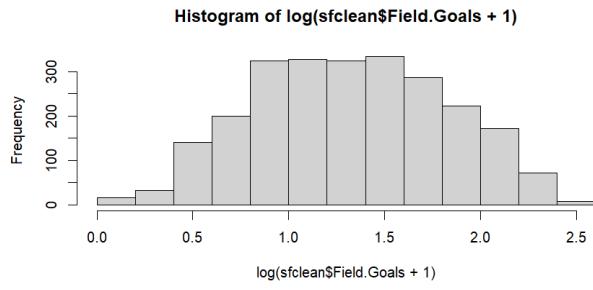
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -4.386e-01 3.552e-02 -12.346 < 2e-16 ***
Free.Throw.made 6.791e-01 4.936e-02 13.759 < 2e-16 ***
Games         -2.426e-03 7.001e-04 -3.465 0.000540 ***  
I(Games^2)    2.222e-05 7.160e-06 3.104 0.001932 **  
I(Free.Throw.made^2) 8.181e-02 2.137e-02 3.829 0.000132 *** 
Effective.field.goal.. 2.424e+00 1.243e-01 19.504 < 2e-16 ***
Turnovers      3.134e-01 2.938e-02 10.665 < 2e-16 ***
FG.           -1.093e+00 1.319e-01 -8.285 < 2e-16 ***
Games.Scored   3.666e-02 4.027e-03 9.105 < 2e-16 ***
Blocks          -1.695e-01 4.341e-02 -3.905 9.70e-05 ***  
Personal.Fouls  5.000e-02 7.409e-03 6.749 1.85e-11 ***  
Steals          7.877e-02 2.530e-02 3.113 0.001873 **  
Assists         1.223e-01 2.224e-02 5.499 4.21e-08 ***  
Defensive.rebounds 3.137e-01 1.755e-02 17.878 < 2e-16 ***
X3P.          1.361e-01 4.990e-02 2.727 0.006432 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

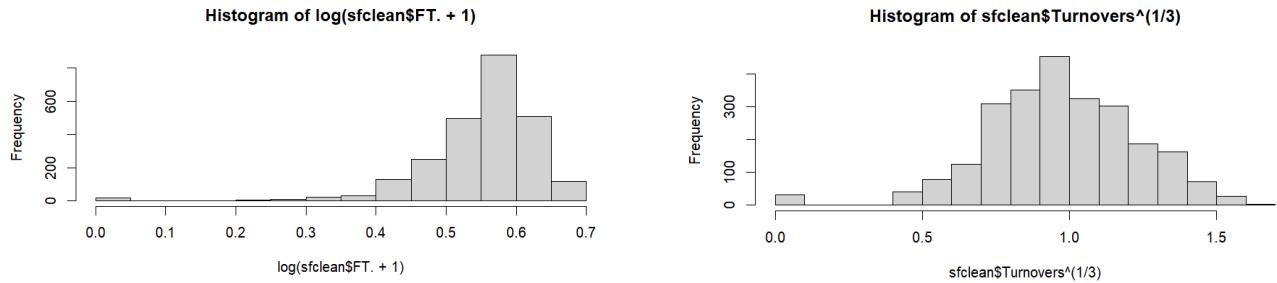
Residual standard error: 0.1722 on 2443 degrees of freedom
Multiple R-squared:  0.9214,    Adjusted R-squared:  0.921 
F-statistic: 2047 on 14 and 2443 DF,  p-value: < 2.2e-16

```

Histograms

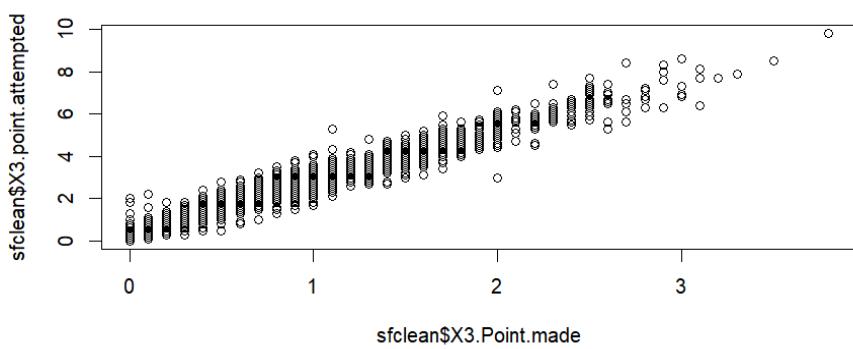






Initial VIF

| Age | Games | Games Scored | Min. Played |
|--------------------|-------------|-----------------------|--------------------|
| 1.174695 | 1.880057 | 3.608075 | 14.538302 |
| X3P | Field Goals | Field Goals Attempted | FG. |
| 2.032767 | 97.096969 | 97.163958 | 9.783745 |
| Free Throw Attempt | X2P. | Effective Field Goal | Free Throw Made |
| 86.298270 | 2.696811 | 8.483729 | 84.350685 |
| Total Rebounds | FT. | Offensive Rebounds | Defensive Rebounds |
| 1430.743128 | 1.749748 | 135.305825 | 888.262985 |
| Turnovers | Assists | Steals | Blocks |
| 7.859468 | 4.796956 | 2.720961 | 1.677246 |
| Personal Fouls | | | |
| 2.898103 | | | |





Final VIF

| | Age | Games | Games.Scored |
|------------------------|----------|-----------------|--------------------|
| Min...Played | 1.137537 | 1.968587 | 4.190867 |
| 11.883200 | | FG. | X2P. |
| Effective.field.goal.. | 8.398891 | 2.063485 | 2.643865 |
| 9.029676 | | Free.Throw.made | Offensive.rebounds |
| Defensive.rebounds | 4.576864 | FT. | |
| 5.267475 | | 1.382191 | 2.468373 |
| Assists | | Steals | Blocks |
| Turnovers | 4.400189 | 2.709604 | 1.788613 |
| 4.541305 | | | |
| Personal.Fouls | 2.674298 | | |

Backward and Forward Models

```

Call:
lm(formula = Points ~ Free.Throw.made + Defensive.rebounds +
    Effective.field.goal.. + Turnovers + Games.Scored + FG. +
    Assists + Personal.Fouls + Blocks + X3P. + Steals + Games +
    Offensive.rebounds, data = sfclean)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.84433 -0.11128  0.00142  0.10833  0.88606 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.5567070  0.0287037 -19.395 < 2e-16 ***
Free.Throw.made 0.8626003  0.0187767  45.940 < 2e-16 ***
Defensive.rebounds 0.3065685  0.0183064  16.747 < 2e-16 ***
Effective.field.goal.. 2.4927623  0.1338496  18.624 < 2e-16 ***
Turnovers 0.3113492  0.0293566  10.606 < 2e-16 ***
Games.Scored 0.0368730  0.0040364   9.135 < 2e-16 ***
FG. -1.2049063  0.1449867  -8.310 < 2e-16 ***
Assists 0.1380034  0.0223752   6.168 8.08e-10 ***
Personal.Fouls 0.0459075  0.0074817   6.136 9.85e-10 ***
Blocks -0.1527681  0.0435667  -3.507 0.000462 ***
X3P. 0.1323356  0.0501972   2.636 0.008434 **  
Steals 0.0554555  0.0253200   2.190 0.028605 *  
Games -0.0004302  0.0001988  -2.209 0.027270 *  
Offensive.rebounds 0.0323304  0.0228533   1.415 0.157286
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Residual standard error: 0.173 on 2444 degrees of freedom
 Multiple R-squared: 0.9207, Adjusted R-squared: 0.9202
 F-statistic: 2181 on 13 and 2444 DF, p-value: < 2.2e-16

```

summary(sfbackward_mode)

Call:
lm(formula = Points ~ Games + Games.Scored + FG. + X3P. + Effective.field.goa
    +
    Free.Throw.made + Offensive.rebounds + Defensive.rebounds +
    Assists + Steals + Blocks + Turnovers + Personal.Fouls, data = sfclean)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.84433 -0.11128  0.00142  0.10833  0.88606 

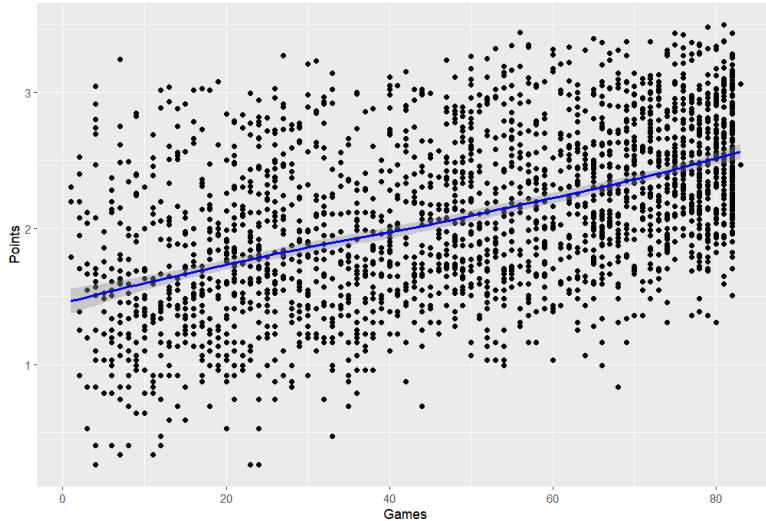
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.5567070  0.0287037 -19.395 < 2e-16 ***
Games        -0.0004392  0.0001988  -2.209 0.027270 *  
Games.Scored  0.0368730  0.0040364   9.135 < 2e-16 *** 
FG.          -1.2049063  0.1449867  -8.310 < 2e-16 *** 
X3P.         0.1323356  0.0501972   2.636 0.008434 ** 
Effective.field.goal... 2.4972623  0.1338496  18.624 < 2e-16 *** 
Free.Throw.made 0.8626003  0.0187767  45.940 < 2e-16 *** 
Offensive.rebounds 0.0323304  0.0228533   1.415 0.157286    
Defensive.rebounds 0.3065685  0.0183064  16.747 < 2e-16 *** 
Assists       0.1380034  0.0223752   6.168 8.08e-10 *** 
Steals         0.0554555  0.0253200   2.190 0.028605 *  
Blocks        -0.1527681  0.0435667  -3.507 0.000462 *** 
Turnovers     0.3113492  0.0293566  10.606 < 2e-16 *** 
Personal.Fouls 0.0459075  0.0074817   6.136 9.85e-10 *** 

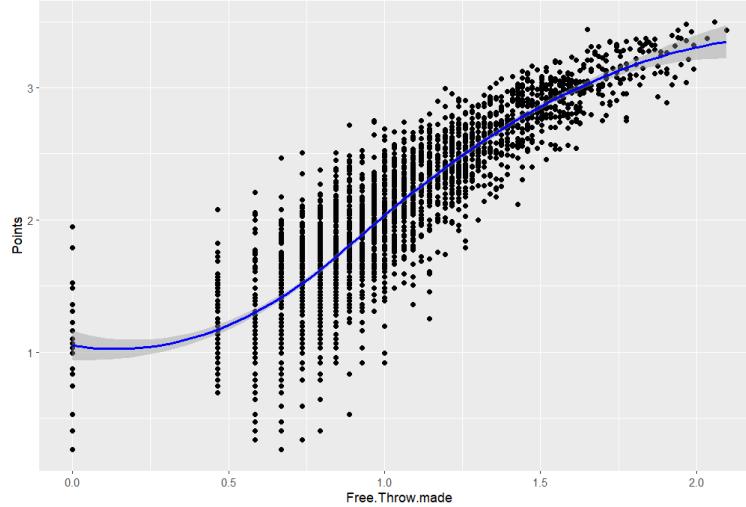
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.173 on 2444 degrees of freedom
Multiple R-squared:  0.9207, Adjusted R-squared:  0.9202 
F-statistic: 2181 on 13 and 2444 DF, p-value: < 2.2e-16

```

Squaring Games and Free Throw Made





Model after Squared Terms and Deleting Insignificant Offensive Rebounds

```

Call:
lm(formula = Points ~ Free.Throw.made + Games + I(Games^2) +
    I(Free.Throw.made^2) + Effective.field.goal.. + +Turnovers +
    FG. + Games.Scored + Blocks + Personal.Fouls + Steals + Assists +
    Defensive.rebounds + X3P., data = srcLean)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.87087 -0.11036  0.00265  0.10930  0.75260 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 4.386e-01 3.552e-02 -12.346 < 2e-16 ***  
Free.Throw.made 6.791e-01 4.936e-02 13.759 < 2e-16 ***  
Games        -2.426e-03 7.001e-04 -3.465 0.000540 ***  
I(Games^2)   2.222e-05 7.160e-06 3.104 0.001932 **  
I(Free.Throw.made^2) 8.181e-02 2.137e-02 3.829 0.000132 ***  
Effective.field.goal.. 2.424e+00 1.243e-01 19.504 < 2e-16 ***  
Turnovers    3.134e-01 2.938e-02 10.665 < 2e-16 ***  
FG.          -1.093e+00 1.319e-01 -8.285 < 2e-16 ***  
Games.Scored 3.666e-02 4.027e-03 9.105 < 2e-16 ***  
Blocks       -1.695e-01 4.341e-02 -3.905 9.70e-05 ***  
Personal.Fouls 5.000e-02 7.409e-03 6.749 1.85e-11 ***  
Steals        7.877e-02 2.530e-02 3.113 0.001873 **  
Assists      1.223e-01 2.224e-02 5.499 4.21e-08 ***  
Defensive.rebounds 3.137e-01 1.755e-02 17.878 < 2e-16 ***  
X3P.         1.361e-01 4.990e-02 2.727 0.006432 **  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1722 on 2443 degrees of freedom
Multiple R-squared:  0.9214, Adjusted R-squared:  0.921 
F-statistic:  2047 on 14 and 2443 DF, p-value: < 2.2e-16

```

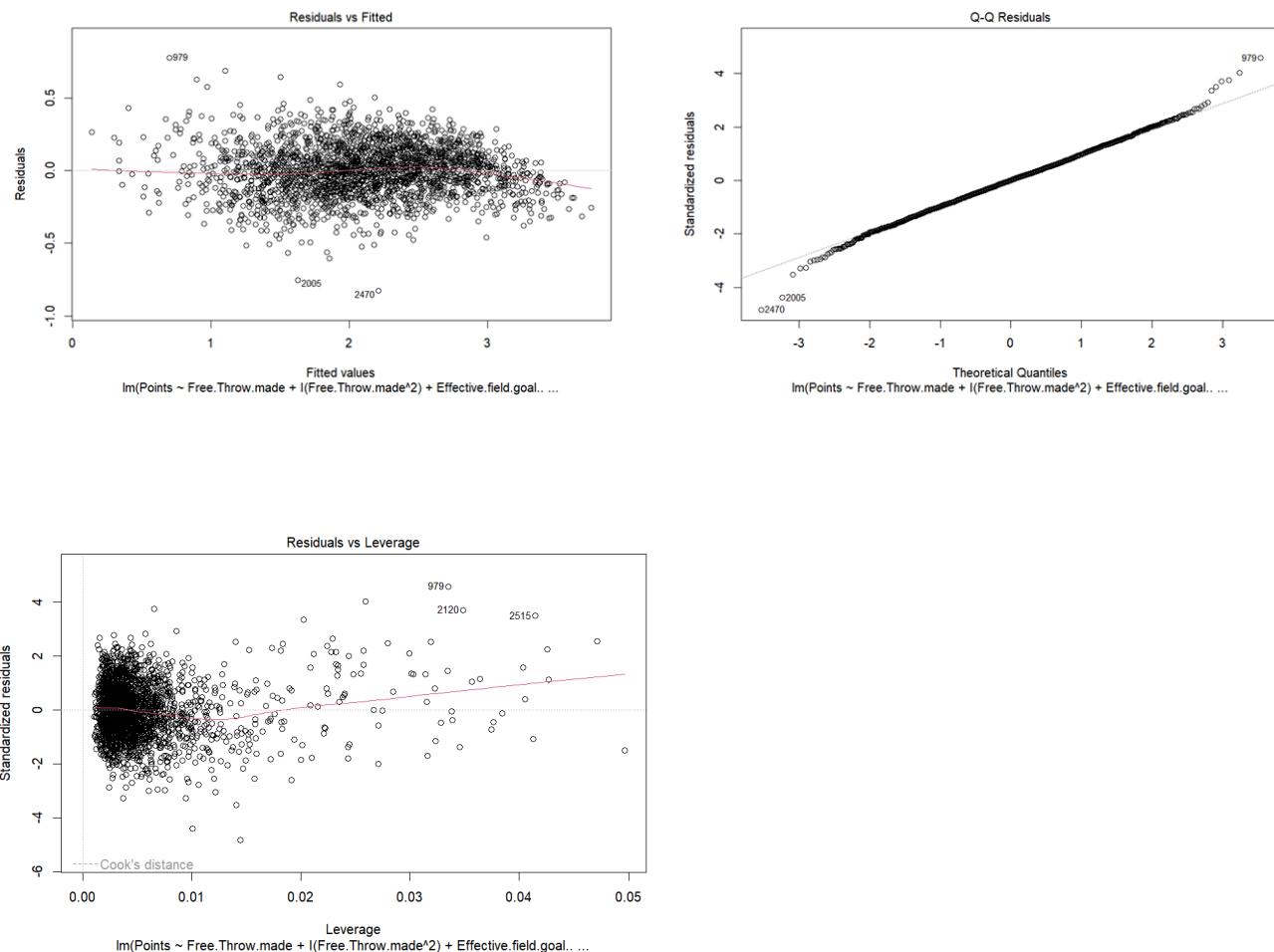
Cross Validation

```

Predict on the test data
> sfpredictions <- predict(sftrainmodel, newdata = sfTestData)
> # Calculate the mean squared error (MSE) on the test data
> mse <- mean((sfTestData$Points - sfpredictions)^2)
> # Print the MSE
> print(mse)
[1] 0.02806707
> # Alternatively, calculate other performance metrics like RMSE
> sfrmse <- sqrt(mse)
> print(sfrmse)
[1] 0.1675323
test_r_squared <- 1 - (sum((sfpredictions - sfTestData$Points)^2) / sum((mean(sftraindata$Points) - sfTestData$Points)^2))
> print(test_r_squared)
[1] 0.9227619

```

Residual Analysis



Power Forward

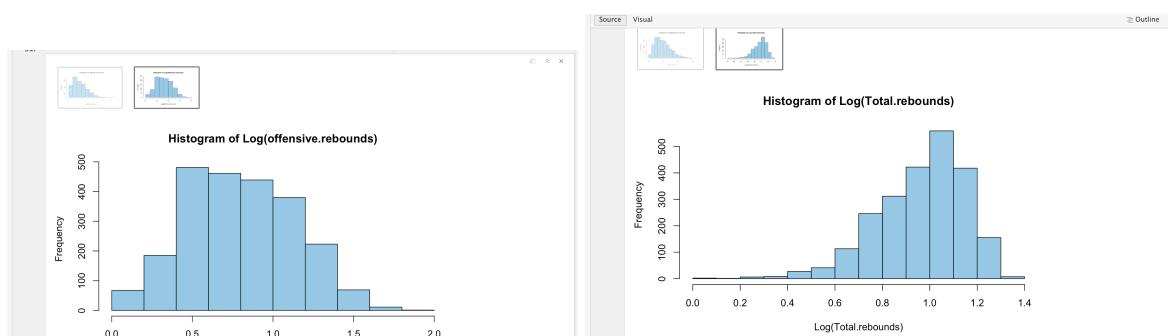
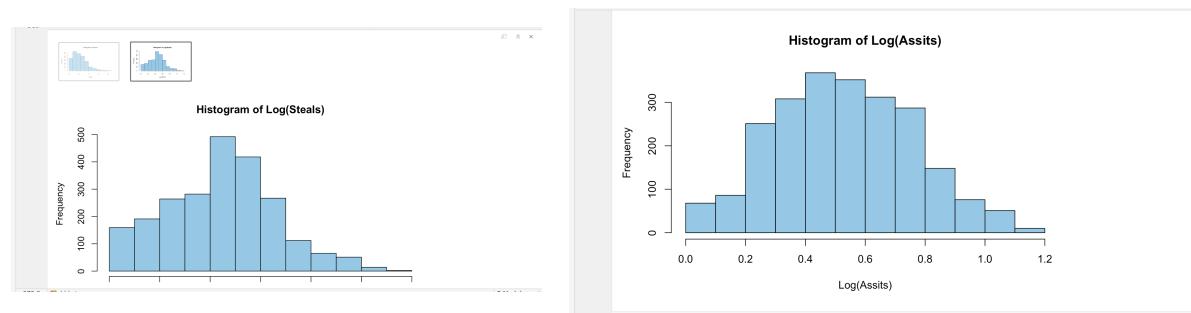
Initial VIF

| | Age | Games | Games.Scored | Min..Played | Field.Goals |
|-----------------------|------------|--------------------|-----------------------------|--------------------|-----------------|
| Field.Goals.Attempted | 1.184793 | 1.984725 | 4.019014 | 18.577488 | 978.342507 |
| X2.Point.made | 595.731482 | 124.527798 | 67.338241 | X3.point.attempted | X3P. |
| Free.Throw.Attempted | 612.241353 | X2.Point.Attempted | X2P. Effective.field.goal.. | 61.673899 | 120.085168 |
| Assists | 4.303206 | 417.597842 | 8.555688 | 86.432704 | Free.Throw.made |
| X3.X3P. | 66.030179 | FT. | Offensive.rebounds | Defensive.rebounds | 106.730633 |
| | 132.675969 | Steals | Blocks | Turnovers | Total.Rebounds |
| | | 60.464432 | 1.771925 | 5.688923 | Personal.Fouls |
| | | | | | 3.160702 |

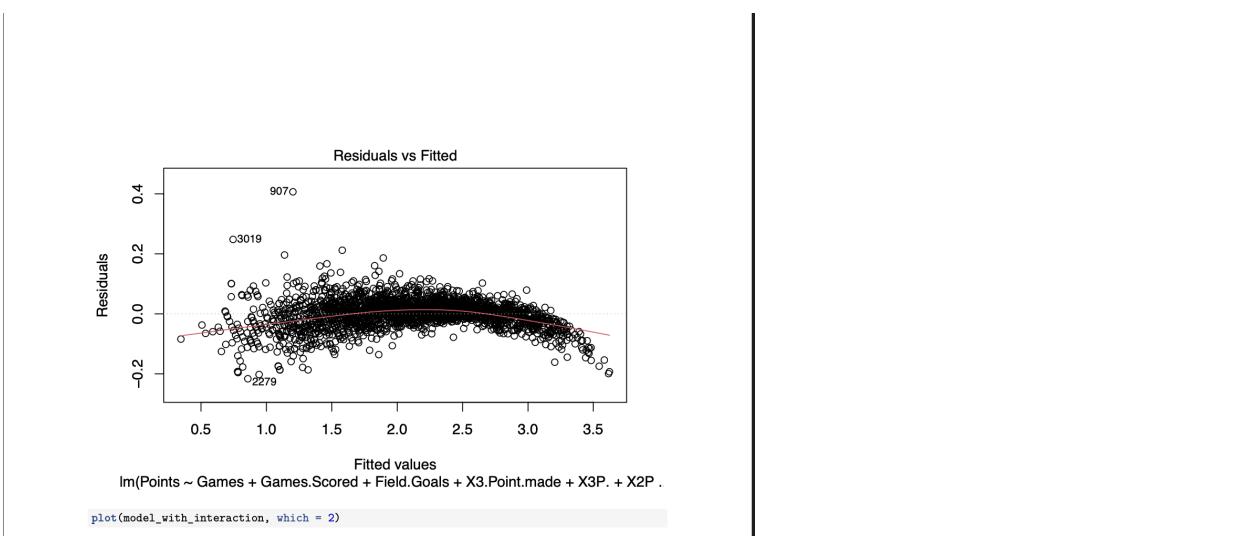
Final VIF

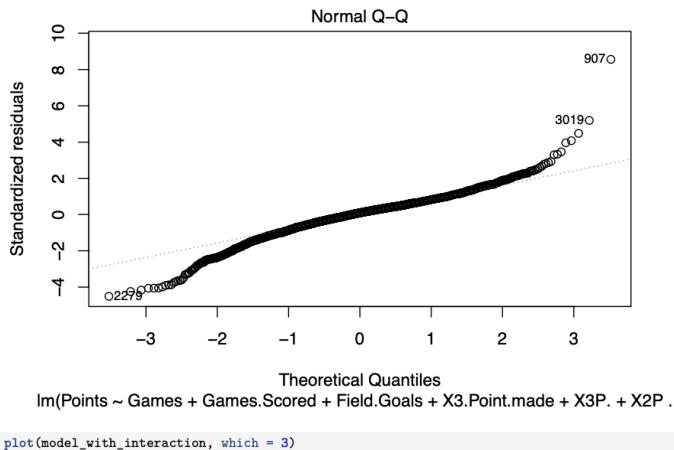
| | Age | Games | Games.Scored | Field.Goals | X3.Point.made |
|--------------------|----------|-----------------------------|--------------|----------------------|---------------|
| | 1.126441 | 1.771829 | 2.809583 | 8.697570 | 2.430657 |
| | X3P. | X2P. Effective.field.goal.. | | Free.Throw.Attempted | FT. |
| | 1.579816 | 2.554378 | 3.106086 | 3.819317 | 1.218447 |
| Offensive.rebounds | | Defensive.rebounds | Steals | Blocks | Turnovers |
| 3.347134 | | 6.330660 | 2.147108 | 1.745043 | 5.474498 |
| Personal.Fouls | | | | | |
| 2.781088 | | | | | |

HISTOGRAM



Residual analysis





Forward and backward elimination

```

Residuals:
    Min      1Q   Median      3Q     Max
-0.26013 -0.03061  0.00345  0.03285  0.35178

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.827e-01 1.131e-02 25.006 < 2e-16 ***
Games        6.306e-04 6.191e-05 10.187 < 2e-16 ***
Games.Scored -1.095e-03 6.491e-05 -16.874 < 2e-16 ***
Field.Goals  1.028e+00 6.481e-03 158.696 < 2e-16 ***
X3.Point.made 1.382e-01 4.651e-03 29.725 < 2e-16 ***
X3P.        -1.629e-02 7.594e-03 -2.146 0.03201 *
X2P.        -1.694e-01 2.265e-02 -7.482 1.04e-13 ***
Effective.field.goal.. 2.645e-01 2.843e-02 9.303 < 2e-16 ***
Free.Throw.Attempted 3.702e-02 1.242e-03 29.816 < 2e-16 ***
FT.          2.125e-01 9.397e-03 22.612 < 2e-16 ***
Offensive.rebounds 5.237e-03 2.433e-03 2.152 0.03149 *
Defensive.rebounds 5.039e-02 6.103e-03 8.256 2.50e-16 ***
Blocks       -8.105e-03 5.462e-03 -1.484 0.13803
Turnovers    -2.277e-02 8.085e-03 -2.817 0.00489 **
Personal.Fouls 2.442e-02 2.292e-03 10.655 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05291 on 2302 degrees of freedom
Multiple R-squared:  0.9925,    Adjusted R-squared:  0.9925
F-statistic: 2.181e+04 on 14 and 2302 DF,  p-value: < 2.2e-16

```

679:9 # `` `

R Markdown ▾

Cross validation

```
+ Fold01: intercept=TRUE
- Fold01: intercept=TRUE
+ Fold02: intercept=TRUE
- Fold02: intercept=TRUE
+ Fold03: intercept=TRUE
- Fold03: intercept=TRUE
+ Fold04: intercept=TRUE
- Fold04: intercept=TRUE
+ Fold05: intercept=TRUE
- Fold05: intercept=TRUE
+ Fold06: intercept=TRUE
- Fold06: intercept=TRUE
+ Fold07: intercept=TRUE
- Fold07: intercept=TRUE
+ Fold08: intercept=TRUE
- Fold08: intercept=TRUE
+ Fold09: intercept=TRUE
- Fold09: intercept=TRUE
+ Fold10: intercept=TRUE
- Fold10: intercept=TRUE
Aggregating results
Fitting final model on full training set
>
> # View the cross-validation results
> print(model_cv)
Linear Regression

2317 samples
 14 predictor

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 2085, 2085, 2086, 2085, 2085, 2084, ...
Resampling results:

  RMSE      Rsquared      MAE
0.1226165  0.9597456  0.09136524

Tuning parameter 'intercept' was held constant at a value of TRUE
>
```

Final model

```

Call:
lm(formula = Points ~ Games + Games.Scored + Field.Goals + X3.Point.made +
   X3P. + X2P. + Effective.field.goal.. + Free.Throw.Attempted +
   FT. + Offensive.rebounds + Defensive.rebounds + Blocks +
   Turnovers + Personal.Fouls + X3P_Def_Reb + Turnovers_Personal_Fouls +
   Games_Field_Goals, data = basket1)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.21639 -0.02457  0.00435  0.02730  0.40682 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.775e-01 1.147e-02 15.475 < 2e-16 ***
Games        2.153e-03 1.552e-04 13.871 < 2e-16 ***
Games.Scored -4.353e-04 6.948e-05 -6.266 4.41e-10 ***
Field.Goals   1.082e+00 7.892e-03 137.097 < 2e-16 ***
X3.Point.made 1.490e-01 4.439e-03 33.558 < 2e-16 ***
X3P.          1.371e-01 1.913e-02 7.169 1.02e-12 ***
X2P.          -9.792e-02 2.158e-02 -4.538 5.99e-06 ***
Effective.field.goal.. 1.362e-01 2.772e-02 4.913 9.61e-07 ***
Free.Throw.Attempted 4.522e-02 1.204e-03 37.543 < 2e-16 ***
FT.           2.058e-01 8.588e-03 23.960 < 2e-16 ***
Offensive.rebounds 4.617e-03 2.227e-03 2.073 0.0382 *  
Defensive.rebounds 4.934e-02 6.546e-03 7.537 6.85e-14 ***
Blocks         -6.022e-03 4.997e-03 -1.205 0.2282    
Turnovers      9.249e-02 1.187e-02 7.792 9.92e-15 ***
Personal.Fouls 5.874e-02 3.944e-03 14.896 < 2e-16 ***
X3P_Def_Reb   -1.204e-01 1.423e-02 -8.457 < 2e-16 ***
Turnovers_Personal_Fouls -5.637e-02 4.540e-03 -12.416 < 2e-16 ***
Games_Field_Goals -1.373e-03 1.164e-04 -11.791 < 2e-16 ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

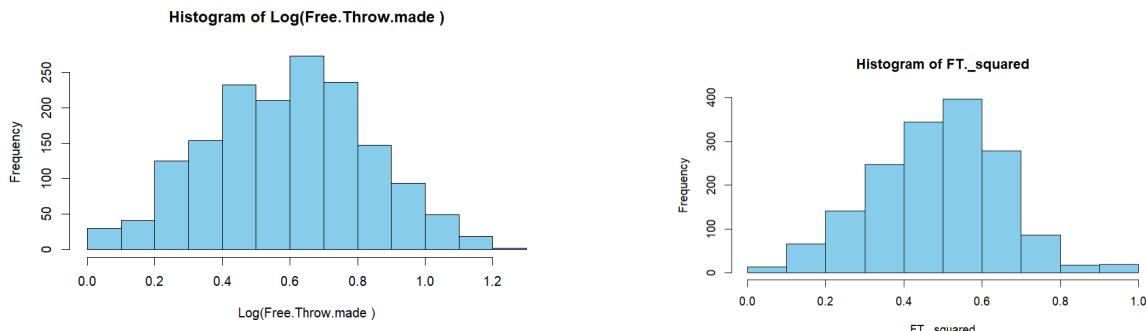
Residual standard error: 0.04832 on 2299 degrees of freedom
Multiple R-squared:  0.9938, Adjusted R-squared:  0.9937 
F-statistic: 2.156e+04 on 17 and 2299 DF,  p-value: < 2.2e-16

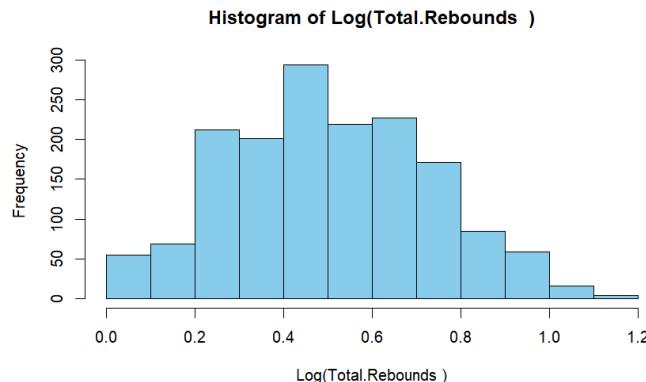
```

> |

Center

Histograms





Initial VIF

| | | | Age | Games | Games.Scored |
|-----------------------|-------------|-----------------------------|------------|----------------|-----------------|
| Min.Played | Field.Goals | | 3.690368 | 14.148999 | 1565.671893 |
| Field.Goals.Attempted | FG. | X3.Point.made | 42.407585 | 42.675862 | X3P. |
| 754.635355 | 161.072325 | X2.Point.Attempted | 19.518848 | 135.321369 | 1.597436 |
| X2.Point.made | 609.525687 | X2P. Effective.field.goal.. | 45.442028 | 220.072214 | Free.Throw.made |
| 1353.795706 | FT. | offensive.rebounds | 220.072214 | 86.631690 | 86.631690 |
| Free.Throw.Attempted | 80.859124 | Defensive.rebounds | 4.415183 | 409.714943 | Total.Rebounds |
| 80.859124 | Assists | Blocks | 2.056276 | Personal.Fouls | 409.714943 |
| Assists | 21.287245 | Turnovers | 2.179904 | Personal.Fouls | 2.936089 |
| 17.785627 | | | 6.099547 | | |

Final VIF

| Age | Games | Games.Scored | Field.Goals.Attempted | FG. |
|--------------------|--------------------|--------------|-----------------------|-----------|
| 1.121773 | 1.654343 | 3.167660 | 6.552424 | 8.110136 |
| X3.Point.made | X3P. | X2P. | Free.Throw.Attempted | FT. |
| 2.494353 | 1.401656 | 7.229124 | 4.415183 | 1.350304 |
| Offensive.rebounds | Defensive.rebounds | Assists | Blocks | Turnovers |
| 4.496886 | 7.523412 | 2.448352 | 2.056276 | 5.857592 |
| Personal.Fouls | | | | |
| 2.641261 | | | | |

Interaction

```
#interaction created
bas$Games_FG <- bas$Games * bas$FG.
bas$Games_X3P <- bas$Games * bas$X3P.
bas$Assists_Turnovers <- bas$Assists * bas$Turnovers
bas$FGA_FTA <- bas$Field.Goals.Attempted * bas$Free.Throw.Attempted
bas$FGA_Personal_Fouls <- bas$Field.Goals.Attempted * bas$Personal.Fouls
bas$Assists_FGA <- bas$Assists * bas$Field.Goals.Attempted
bas$Games_Personal_Fouls <- bas$Games * bas$Personal.Fouls
```

```

Call:
lm(formula = Points ~ Games + Field.Goals.Attempted + FG. + X3.Point.made +
    X3P. + X2P. + Free.Throw.Attempted + FT. + Assists + Turnovers +
    Personal.Fouls + Games_FG + Games_X3P + Assists_Turnovers +
    FGA_FTA + FGA_Personal_Fouls + Assists_FGA + Games_Personal_Fouls,
    data = bas)

Residuals:
    Min      1Q   Median      3Q     Max 
-0.216468 -0.010932  0.000359  0.011241  0.188978 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -6.104e-01  9.120e-03 -66.927 < 2e-16 ***
Games        -3.924e-06  2.183e-06 -1.797 0.072498 .  
Field.Goals.Attempted 8.950e-01  5.646e-03 158.506 < 2e-16 ***
FG.          1.158e+00  2.661e-02  43.506 < 2e-16 ***
X3.Point.made 1.216e-01  3.701e-03  32.860 < 2e-16 *** 
X3P.         4.674e-02  7.566e-03  6.177 8.26e-10 *** 
X2P.         1.415e-01  3.107e-02  4.554 5.66e-06 *** 
Free.Throw.Attempted 3.336e-01  8.094e-03 41.224 < 2e-16 *** 
FT.          1.692e-01  4.363e-03  38.779 < 2e-16 *** 
Assists      2.081e-02  4.403e-03  4.727 2.48e-06 *** 
Turnovers    2.198e-03  6.720e-03  0.327 0.743687    
Personal.Fouls -1.713e-03  4.206e-03 -0.407 0.683870    
Games_FG     9.072e-06  3.890e-06  2.332 0.019809 *  
Games_X3P    -8.718e-06  1.770e-06 -4.924 9.34e-07 *** 
Assists_Turnovers -6.279e-04  3.803e-03 -0.165 0.868869    
FGA_FTA      -4.371e-02  3.742e-03 -11.682 < 2e-16 *** 
FGA_Personal_Fouls -8.480e-04  2.296e-03 -0.369 0.711861    
Assists_FGA   -9.092e-03  2.693e-03 -3.376 0.000753 *** 
Games_Personal_Fouls 4.933e-07  4.717e-07  1.046 0.295819  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02603 on 1593 degrees of freedom
Multiple R-squared:  0.998,    Adjusted R-squared:  0.998 
F-statistic: 4.496e+04 on 18 and 1593 DF,  p-value: < 2.2e-16
```

 R Markdown 

Forward selection and Backward Elimination

```

Call:
lm(formula = Points ~ Games + Field.Goals.Attempted + FG. + X3.Point.made +
    X3P. + X2P. + Free.Throw.Attempted + FT. + Assists + Games_FG +
    Games_X3P + FGA_FTA + Assists_FGA + Personal.Fouls, data = bas)

Residuals:
    Min      1Q   Median      3Q     Max 
-0.215310 -0.010887  0.000394  0.011333  0.190090 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -6.092e-01  7.867e-03 -77.443 < 2e-16 ***
Games        -2.888e-06  1.960e-06 -1.473 0.141    
Field.Goals.Attempted 8.934e-01  3.604e-03 247.882 < 2e-16 ***
FG.          1.155e+00  2.606e-02  44.340 < 2e-16 *** 
X3.Point.made 1.213e-01  3.667e-03  33.075 < 2e-16 *** 
X3P.         4.742e-02  7.522e-03  6.304 3.74e-10 *** 
X2P.         1.424e-01  3.082e-02  4.620 4.14e-06 *** 
Free.Throw.Attempted 3.336e-01  7.318e-03 45.581 < 2e-16 *** 
FT.          1.693e-01  4.335e-03 39.061 < 2e-16 *** 
Assists      2.107e-02  4.281e-03  4.922 9.47e-07 *** 
Games_FG     9.523e-06  3.831e-06  2.486 0.013 *  
Games_X3P    -8.837e-06  1.765e-06 -5.007 6.13e-07 *** 
FGA_FTA      -4.353e-02  3.176e-03 -13.707 < 2e-16 *** 
Assists_FGA   -9.467e-03  1.783e-03 -5.309 1.26e-07 *** 
Personal.Fouls -1.775e-03  1.241e-03 -1.431 0.153    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.026 on 1597 degrees of freedom
Multiple R-squared:  0.998,    Adjusted R-squared:  0.998 
F-statistic: 5.791e+04 on 14 and 1597 DF,  p-value: < 2.2e-16
```

Cross Validation

```
Warning: package 'lattice' was built under R version 4.1.3+ Fold01: intercept=TRUE
- Fold01: intercept=TRUE
+ Fold02: intercept=TRUE
- Fold02: intercept=TRUE
+ Fold03: intercept=TRUE
- Fold03: intercept=TRUE
+ Fold04: intercept=TRUE
- Fold04: intercept=TRUE
+ Fold05: intercept=TRUE
- Fold05: intercept=TRUE
+ Fold06: intercept=TRUE
- Fold06: intercept=TRUE
+ Fold07: intercept=TRUE
- Fold07: intercept=TRUE
+ Fold08: intercept=TRUE
- Fold08: intercept=TRUE
+ Fold09: intercept=TRUE
- Fold09: intercept=TRUE
+ Fold10: intercept=TRUE
- Fold10: intercept=TRUE
Aggregating results
Fitting final model on full training set
Linear Regression

1612 samples
 14 predictor

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1451, 1451, 1450, 1451, 1450, 1452, ...
Resampling results:

  RMSE     Rsquared     MAE
  0.02630964  0.9980099  0.01736849

Tuning parameter 'intercept' was held constant at a value of TRUE
```

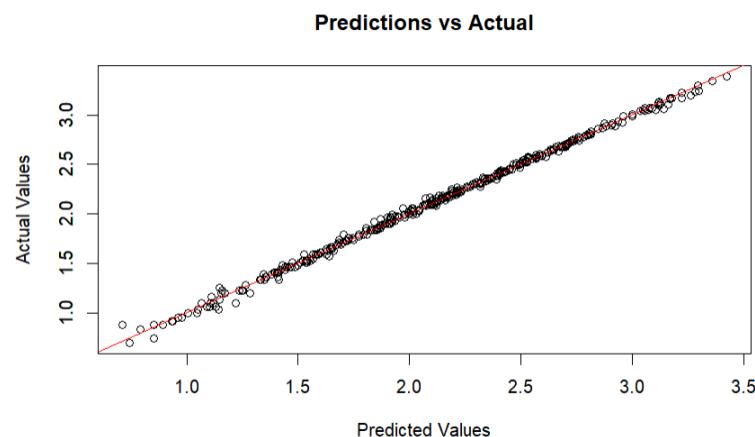
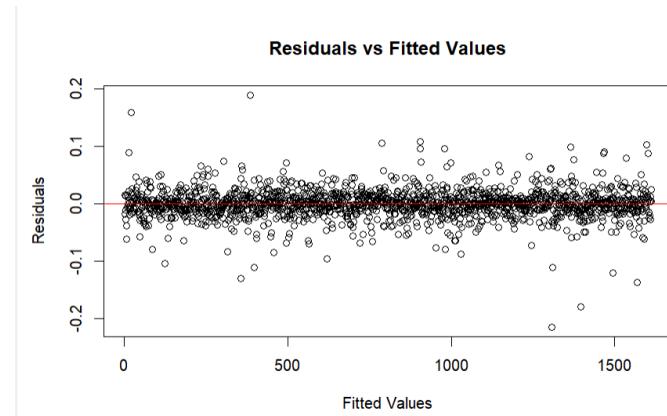
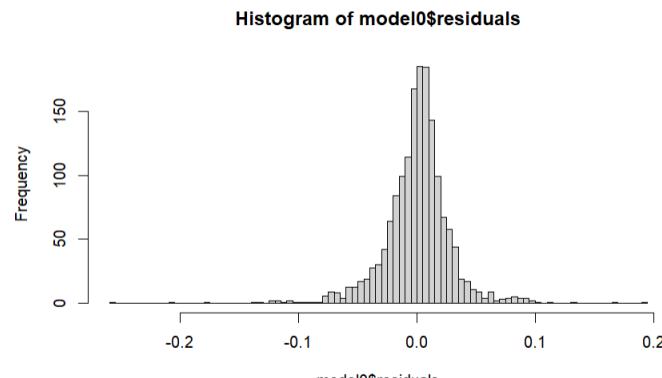
Testing

```
;`{r}
} prediction <- predict(model0, test)
} head(prediction)
)```
8 24 27 29 31 41
2.6116744 1.1486859 1.6040005 2.5099296 0.7099302 1.3451646

;`{r}
} actual = test$Points
! head(actual)
)```
[1] 2.5726122 1.2527630 1.6094379 2.5095993 0.8754687 1.3862944

;`{r}
} cor(prediction, actual)
)```
[1] 0.9987288
```

Residual



Final model

```
Call:
lm(formula = Points ~ Games + Field.Goals.Attempted + FG. + X3.Point.made +
X3P. + X2P. + Free.Throw.Attempted + FT. + Assists + Turnovers +
Personal.Fouls, data = bas)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.259750 -0.012782  0.001634  0.013746  0.191017 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -5.449e-01  6.611e-03 -82.430 < 2e-16 ***
Games        1.448e-06  3.920e-07   3.693 0.000229 ***
Field.Goals.Attempted 8.506e-01  3.212e-03  264.795 < 2e-16 ***
FG.          1.248e+00  2.637e-02   47.310 < 2e-16 ***
X3.Point.made 1.209e-01  4.075e-03   29.668 < 2e-16 ***
X3P.         1.561e-02  5.219e-03   2.992 0.002813 **  
X2P.         1.047e-01  3.452e-02   3.034 0.002452 **  
Free.Throw.Attempted 2.348e-01  3.127e-03   75.093 < 2e-16 ***
FT.          1.705e-01  4.927e-03   34.612 < 2e-16 ***
Assists      -3.926e-03  1.102e-03  -3.562 0.000379 *** 
Turnovers    -1.571e-02  5.764e-03  -2.726 0.006483 **  
Personal.Fouls 4.242e-03  1.460e-03   2.907 0.003705 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02952 on 1600 degrees of freedom
Multiple R-squared:  0.9975,    Adjusted R-squared:  0.9974 
F-statistic: 5.715e+04 on 11 and 1600 DF,  p-value: < 2.2e-16
```