The Impact of Scheduling on NBA Team Performance

Author: Zachary P. Malosh[*]
Reader: Michael Magazine, PhD [†]
Reader: Tom Zentmeyer, MBA[‡]

30 November 2017

Abstract

Every year, the NBA releases their league schedule for the coming year. The construction of the schedule contains many potential schedule-based factors (such as rest, travel, and home court) that can impact each game. Understanding the impact of these factors is possible by creating a regression model that quantifies the team performance in a particular game in terms of final score and fouls committed. Ultimately, rest, distance, attendance, and time in the season had direct impact on the final score of the game while the attendance at a game led to an advantage in fouls called against the home team. The quantification of the impact of these factors can be used to anticipate variations in performance to improve accuracy in a Monte Carlo simulation.

[*]University of Cincinnati - Masters Candidate, Fox Sports - Predictions Team Lead (zachary.malosh@foxsports.com)
[†]University of Cincinnati - Professor and Ohio Eminent Scholar
[‡]Fox Sports - Vice President of Technology

Contents

Introduction

In the multi-billion dollar industry of professional sports, there is an often overlooked, yet fundamentally necessary, requirement which drives the competition that excites fans: both teams showing up to the game. Before every season, the league offices decide who will play where and when. There are many decisions that go into the final schedule that fans look forward to receiving every off-season. While the schedule is a clear necessity, it introduces many game-altering factors for both teams.

There are a wide range of game-changing factors introduced by the schedule. Four important factors are if a team is home or away, how far a team travels, how much rest each team has, and the attendance of the game. Identifying how exactly these factors influence the game for both teams will allow for potential advantages to be found and the expectations for team performance in those games to be changed accordingly. The successful development of a model (or series of models) that defines the impact of these factors can then be used in other systems, such as a Monte Carlo simulation, in order to improve the quality of those systems.

Ultimately, models focusing on points scored and fouls committed will be created for both the home and away teams.

Data

Data Definition

The core observation unit in this analysis is the game. Each game includes the participating teams, the game result, derived distance values based on team and game locations, final game statistics, and statistics for games that fall into the window of 4 games prior to the observed game. This 4 game window, referred to as the *past game window*, was used as a baseline for a team's performance in the games leading up to an observed game. The window width of 4 games was chosen after replicating the data set multiple times with differing game windows and seeing similar results when 4 or more games were used and very different values when less games were used.

Data Source

All data utilized in this analysis comes from the Sports Data system at Fox Sports. This system processes and stores raw data from Stats, Inc. Direct access to this proprietary system is available for select employees of Fox Sports. Any additional information regarding the data used is available upon request.

Game Selection

Standard Criteria

The following criteria were used to determine which games would be included in this analysis:

- The game must be from any of the seasons between 2005 and 2016, inclusive.
- The game must be in the regular season.
- The game must be preceded by at least 4 regular seasons game for both teams in a given season.
- The game must have a status of final. Any game that has a status of forfeited will not be selected.
- A game with a status of final can have a forfeited or canceled game in its past game window. In that case, the forfeited/canceled game will be omitted from the past game window. This will result in only 3 games in the past game window.

Neutral Site Games

Given the lack of a true home team, neutral site games are considered a separate concern when looking at the impact of the schedule on NBA team performance. With this in mind, all games played at neutral sites were removed from the data set. This means that 22 of 13596 games (0.162%) were removed. As a result, all values for the distance that the home team is playing from their home arena is 0.

Extended Breaks Between Games

The NBA has multiple reasons which may cause a team to have extended breaks between games. The most common reason in the regular season is the All-Star break. Further exploration of this, as discussed in the data summary, show that the number of qualified games where a team has 4-or-more days of rest is quite small. As such, the rest days data was winsorized to a maximum value of 4 days.

The frequency table leading to the decision to winsorize at 99% can be found below.

	0	1	2	3	4	5	6	7	8	9	10
Freq	6517	15154	3951	985	158	217	69	37	51	6	3
Density	0.24	0.56	0.15	0.04	0.01	0.01	0	0	0	0	0
Cum Density	0.24	0.8	0.94	0.98	0.99	0.99	1	1	1	1	1

Variable Definitions

The following variables are used in this analysis and are the key focus points to be analyzed.

- *Season* - factor
  - The year in which the season containing this game began. If the first regular season game for a sport is in October of 2015 and the championship for that sport is determined in June of 2016, the value will be 2015.
- *AwayTeam/HomeTeam* - factor
  - A factor based on the preferred short display form for a particular season. This value may be different across years for a particular team based on the team changing cities or names. (For example, the Seattle Supersonics became the Oklahoma City Thunder prior to the 2008 season. The team is *'SEA'* in 2007 and *'OKC'* in 2008.)
- *AwayScore/HomeScore* - int
  - Number of points scored *in regulation* by the associated team.
- *GameDate* - date
  - The date in the Eastern Time Zone that the game was scheduled to start. If a game was rescheduled, this date represents the new scheduled date. The initial game will not appear in the data.
- *Attendance* - int
  - Game attendance in thousands
- *AwayRest/HomeRest* - int
  - Number of calendar days between the game being referenced and the previous game. If the referenced game is Thursday and the previous game was Monday, the value will be *2* as Tuesday and Wednesday were rest days.
- *AwayMilesTraveled/HomeMilesTraveled* - int
  - Travel distance in hundreds of miles
  - The number of miles between the venue of the previous game and the venue of the referenced game. This value is calculated using the longitude and latitude values for each of the venues and, as such, represents the direct distance between venues and not the driving distance.
- *AwayMilesAway/HomeMilesAway* - int
  - Distance away from home stadium in hundreds of miles
  - If the game is not at a neutral site, the value will be 0 for the home team. This value is calculated using the longitude and latitude values for each of the venues and, as such, represents the direct distance between venues and not the driving distance.
- *HomeMargin* - int
  - The difference between the *HomeScore* and the *AwayScore* ($HomeScore - AwayScore$)
- ***Statistics***
  - The following statistics are also included for both teams in each game as well as in the past game window (both for the team and against): Score, Rebounds, Blocks, Assists, Steals, Turnovers, ThreePointMakes, Fouls, FreeThrowAtts

Summary Statistics

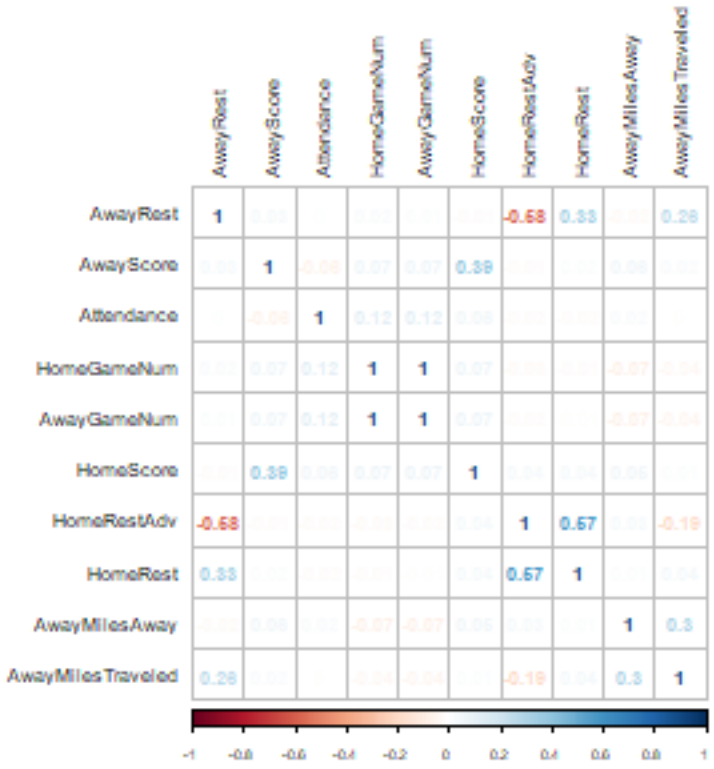
A table with all game and schedule summary statistics can be found in the appendix.

Data Exploration

Correlation between Variables

Schedule and Result Based

The correlation between schedule-based and result variables was investigated to explore the impact of scheduling. The resulting correlation matrix can be found below.



The above matrix shows that there is only 1 sets of strongly correlated ($r > 0.7$) variables in the schedule-based/result variable set. That is the pair of the *HomeGameNumber* and *AwayGameNumber* variables. These variables have a perfect correlation of r=1. As such, *AwayGameNumber* was dropped moving forward.

Past Game Window Statistics

The past game window statistics (namely score, rebounds, blocks, assists, steals, turnovers, fouls, and free throw attempts) were evaluated to determine if there was a correlation between statistics. The full correlation matrix can be found in the appendix. The following pairs were determined to be strongly correlated:

1) *TeamTurnovers* - *OppSteals*: The most frequent turnover for a team is when the other team records a steal.
2) *Fouls* - *OppFouls*: This correlation indicates that the number of fouls called against each team increases at an approximately equal rate. There are multiple factors which can contribute to the number of fouls in a game for a given team. Examples of these factors include the officials, being the home or away team, and the pace of the game. At first glance, it is good to see that the number of fouls between the two teams is highly correlated. Further analysis is necessary, and would later be performed in this analysis, to determine if there is a statistical difference between the number of fouls called against the home and away teams.

Home Court Advantage

To show there is a statistically significant difference between home score and away score, a paired two-sample t-test for the difference of means was performed. The results and other relevant values can be found below.

Table 2: Paired Two-Sample t-Test for Home Court Advantage

	t.value	df	p.value	CI	Mean.Est
t-Test	26.33	13573	< 2.2e-16	(2.74,3.18]	2.9636

The p-values for the mean difference between the home team and the away team for the NBA is < 2.2e-16. This shows that there is a statistically significant difference between home team scoring and away team scoring, thus statistically verifying the existance of home court advantage. Additionally, the advantage, or mean difference, is approximately 2.96 for the NBA with a 95% confidence interval of (2.74,3.18], showing that the advantage is approximately one scoring play.

Home Advantage in Personal Fouls

Many sports fans believe that there is a major impact on the officiating of a game based on which team is home and which team is away. There are many ways that the officiating crew impacts the game. The best direct measure for the impact of officiating in the game is the number of calls made. This includes both fouls and possession-costing violations. It should be noted that no-call decisions are just as important. A no-call can show up in the final box score as a turnover, as points, or as a rebound. Given that there is no reliable way to determine where no-call decisions were made for the full duration of each game during the analyzed period, the best way to determine the home court impact on officiating is to analyze the number of fouls called against each team. In order to test if there is a difference in the number of fouls called based on which team is home and which team is away, a paired two-sample t-test on the difference of the mean number of personal fouls was performed. The results can be found below.

Table 3: Paired Two-Sample t-Test for Home Advantage in Fouls

	t.value	df	p.value	CI	Mean.Est
t-Test	-17.38	13573	< 2.2e-16	(-0.87,-0.69]	-0.7778

This shows that there is, in fact, a difference in the number of fouls called based on the home/away status of a team. This advantage is less than one foul per game and, while statistically significant, can arguably be seen as negligible on the final outcome of the game.

Distance Summary

The distance of travel, both between games and from a team's home court, is an important aspect of the impact of scheduling on NBA team performance. With teams scattered across North America, it is important to know how far apart teams may be. The greatest and shortest distances between 2016 home arenas can be found below. (A full chart of stadium-to-stadium distances can be found in the appendix.)

Table 4: Summary Statistics for Arena-to-Arena Distance

	Min	Q1	Med	Mean	Q3	Max
Distance	0	635	1047	1164.95	1635.5	2706

Table 5: Bottom 8 Distances Between 2016 Home Arenas in Miles

Stadium1	Stadium2	Distance
Staples Center (LAC)	Staples Center (LAL)	0
Barclays Center (BKN)	Madison Square Garden (NY)	4
Golden 1 Center (SAC)	ORACLE Arena (GS)	68
BMO Harris Bradley Center (MIL)	United Center (CHI)	81
Barclays Center (BKN)	Wells Fargo Center (PHI)	82
Madison Square Garden (NY)	Wells Fargo Center (PHI)	85
Quicken Loans Arena (CLE)	The Palace of Auburn Hills (DET)	115
Verizon Center (WSH)	Wells Fargo Center (PHI)	120

Table 6: Top 8 Distances Between 2016 Home Arenas in Miles

Stadium1	Stadium2	Distance
AmericanAirlines Arena (MIA)	Moda Center at the Rose Quarter (POR)	2706
ORACLE Arena (GS)	TD Garden (BOS)	2691
Golden 1 Center (SAC)	TD Garden (BOS)	2631
Staples Center (LAC)	TD Garden (BOS)	2598
Staples Center (LAL)	TD Garden (BOS)	2598
AmericanAirlines Arena (MIA)	ORACLE Arena (GS)	2585
Barclays Center (BKN)	ORACLE Arena (GS)	2565
Madison Square Garden (NY)	ORACLE Arena (GS)	2563

The shortest distance between venues is listed as the Staples Center (LAC) to the Staples Center (LAL). While this is the same venue, it is viewed as two different venues for the purpose of home court advantage due to the differing atmospheres based on which franchise (LAL or LAC) is acting as host. As such, this anomoly was determined to be correct and was not adjusted. The BKN-NY-PHI team group stands out as having many short-distance pairs, potentially providing an advantage to those teams. On the other side, two teams (BOS and GS) each are part of at least four of the largest distances between arenas.

Method

Desired Modeling Outcomes

Given the ultimate desire to use the resulting models in a Monte Carlo simulation, each analyzed statistic (such as score or fouls) will be broken into two models, one for the home team and one for the away team. This was chosen in favor of one model for statistic total and one for team margin.

Variable Selection

The impact of scheduling on the outcome of game performance was analyzed using the relevant variables from the previous game stat window along with the schedule-based variables for the analyzed game. The statistics from the previous game stat window include the statistic value for the team and the allowed statistic value for the opponent. For example, the baseline statistics used when analyzing HomeScore will be the average score for the home team in their past 4 games and the average score allowed in the past 4 games for the away team. The schedule-based statistics used in model building are *AwayRest*, *HomeRest*, *AwayMilesTraveled*, *HomeMilesTraveled*, *AwayMilesAway*, *Attendance*, and *HomeGameNum*.

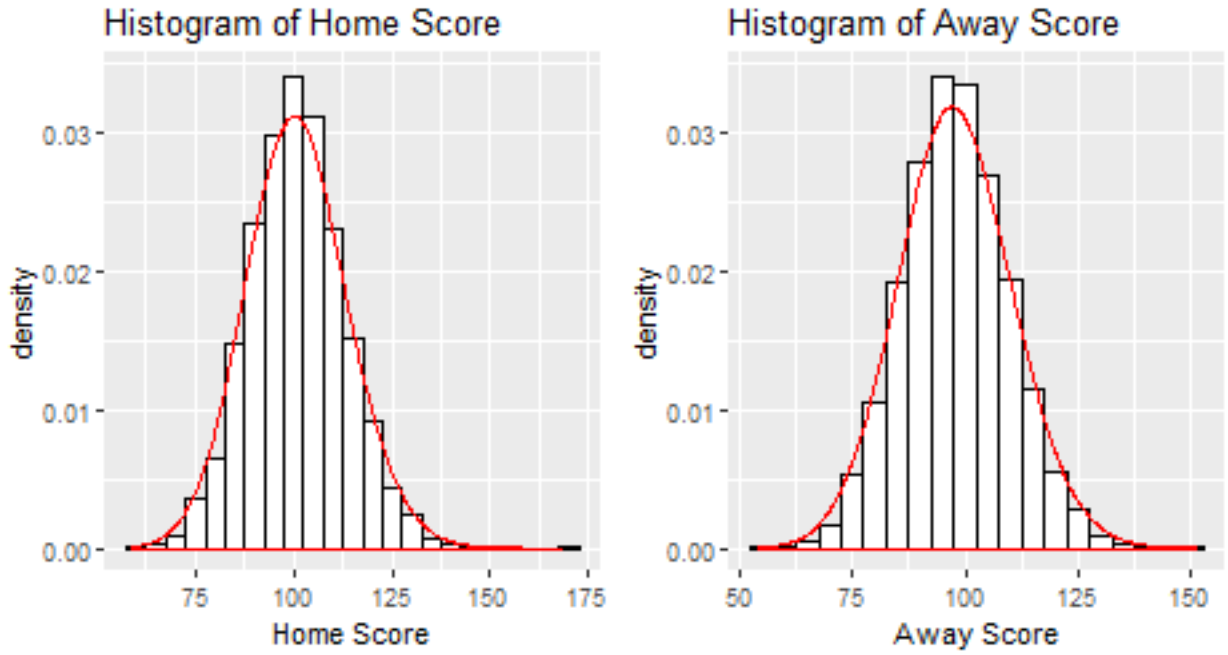
Model Creation

For each of the desired models, a simple linear regression was performed using a model that was created using the stepwise fitting method. The entry and exit p-value for this process was set at 0.01. This p-value translated to a one-tailed chi-square value of 6.635 for the k parameter of the *step* method. After each model was created, its fitness was analyzed using residual, normal Q-Q, scale-location, and residual vs leverage plots produced from the data.

Analysis

Points Scored

The goal of every NBA game is to outscore the opponent. With this in mind, the score of a game is the most important metric of the performance of an NBA team. An effective model of the impact of scoring on each of the schedule-based metrics is the highest modeling priority. Visualizations of the dependent variables (*HomeScore* and *AwayScore*) can be found below.



Model Creation

Models were created by stepwise fitting a linear regression model using the baseline and schedule-based variables described previously. The summaries of the resulting models can be found below.

Table 7: Summary of Home Score Model

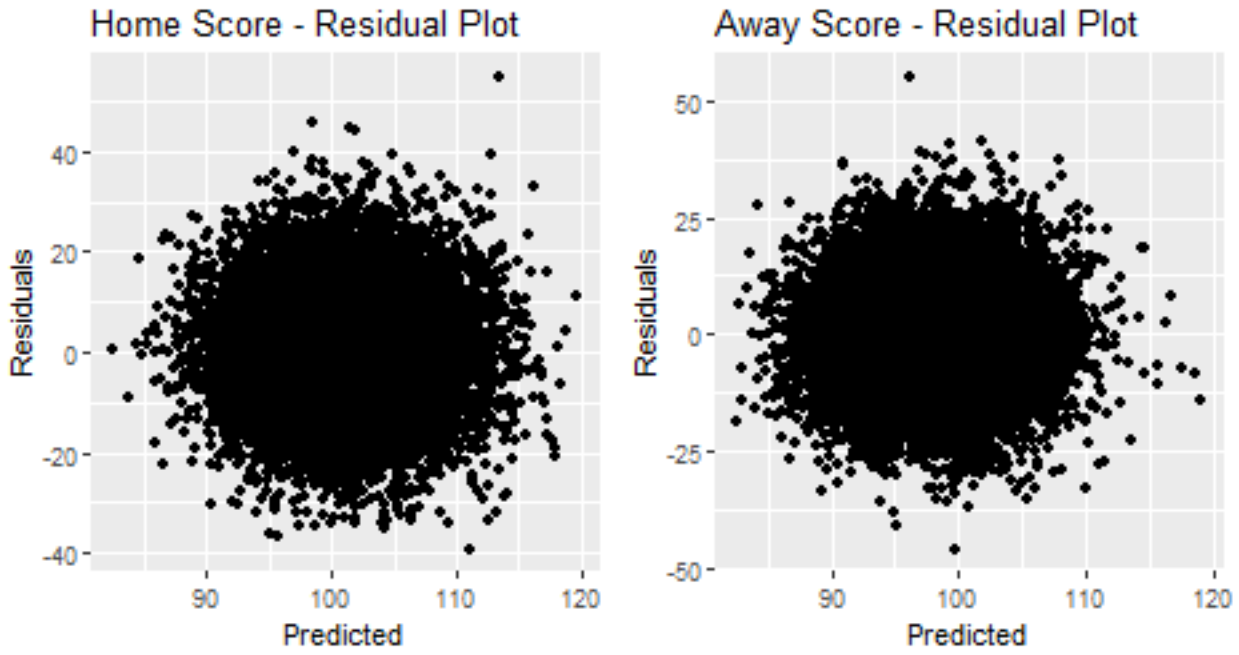
	Estimate	Std.Err	t.Value	p.Value
(Intercept)	15.2463343	1.6790428	9.080372	< 2.2e-16
PrevHomeScore	0.4482854	0.0127318	35.209987	< 2.2e-16
PrevOppAwayScore	0.4069630	0.0127243	31.983153	< 2.2e-16
HomeRest	0.5574520	0.1164067	4.788830	1.695e-06
AwayRest	-0.3998061	0.1153787	-3.465164	0.0005315
AwayMilesAway	0.0385659	0.0145692	2.647083	0.0081283

Table 8: Summary of Away Score Model

	Estimate	Std.Err	t.Value	p.Value
(Intercept)	22.5501739	1.7825879	12.650245	< 2.2e-16
PrevOppHomeScore	0.4051257	0.0126768	31.958016	< 2.2e-16
PrevAwayScore	0.3813972	0.0124921	30.531001	< 2.2e-16
Attendance	-0.2519021	0.0359033	-7.016135	2.389e-12
AwayMilesAway	0.0836904	0.0143464	5.833530	5.550e-09
AwayRest	0.4619506	0.1109430	4.163855	3.149e-05
HomeMilesTraveled	-0.0456558	0.0166999	-2.733901	0.006267
HomeGameNum	0.0115209	0.0042661	2.700547	0.006931

Model Diagnostics

Plotting the predicted home scores versus the residual values produced by the models resulted in the following residual plots.



The above plots, in conjunction with additional diagnostic plots available in the appendix, show that the regression is a solid fit and does not show characteristics which would tend to indicate issues in the model.

Model Interpretation

One of the key aspects of the two points models is that the baseline statistics are kept in both models. This means that the two baseline calibration variables in the previous game window are still extremely important to fitting the ultimate regression model. There are five other variable sets that occur across the two models that are statistically significant to the understanding of schedule-based influence on scoring.

Rest Days

The home score model sees 0.557 additional points for each day of rest (up to 4) and 0.4 fewer points for each day of away rest (up to 4). Meanwhile, the away score model sees 0.462 additional points per day of away rest. The presence of away rest days in both models means that the number of days of away rest results in an impact on the ultimate home game margin of 0.862 points. This means that the number of away days of rest, completely ignoring home days of rest, can represent a swing of up to 3.448 points in favor of the away team in a single game. This value, which is more than what can be achieved in a single possession, indicates that a game-altering scoring change can be achieved just by the number of rest days scheduled between games for the away team by the league office. It should be noted that the away team having 4 days rest and the home team having no rest is a situation that would likely be avoided in the interest of fair competition. If the home team has 2 days of rest compared to 4 days rest for the away team, the expected difference in score as produced by the scheduled rest days would be 2.332 points. This is still greater than the two points for a standard field goal, meaning that a decisive difference may come from just from the number of rest days scheduled.

Away Miles Away

An interesting aspect of the two scoring models is that the number of points goes up for both the home and away scores based on the number of miles (by the hundred) away the away team is from their home court for a given game. The number of points increases by 0.039 for the home team and 0.084 for the away team based on the away team miles away from home. Based on both teams having positive values for this measure, the pace of the game can be said to increase by 0.122 per game per hundred miles away. While there may be

many explanations for this simultaneous increase for both teams, one of the strongest possibilities for this increase in pace is the decrease in concentration that results from extended travel. A reduced concentration leads to more possessions for both teams and, as such, an increased total score for the game.

Attendance

The away team scores -0.252 fewer points per 1000 fans in attendance per game. This means that the middle 50% of games (based on attendance) see the away team scoring between -4.01 and -4.928 fewer points per game. This indicates that having a strong attendance as the home team has a statistically significant impact on the final score of the game, even if it does only have a direct (and negative) impact on the final score of the away team.

Home Miles Traveled

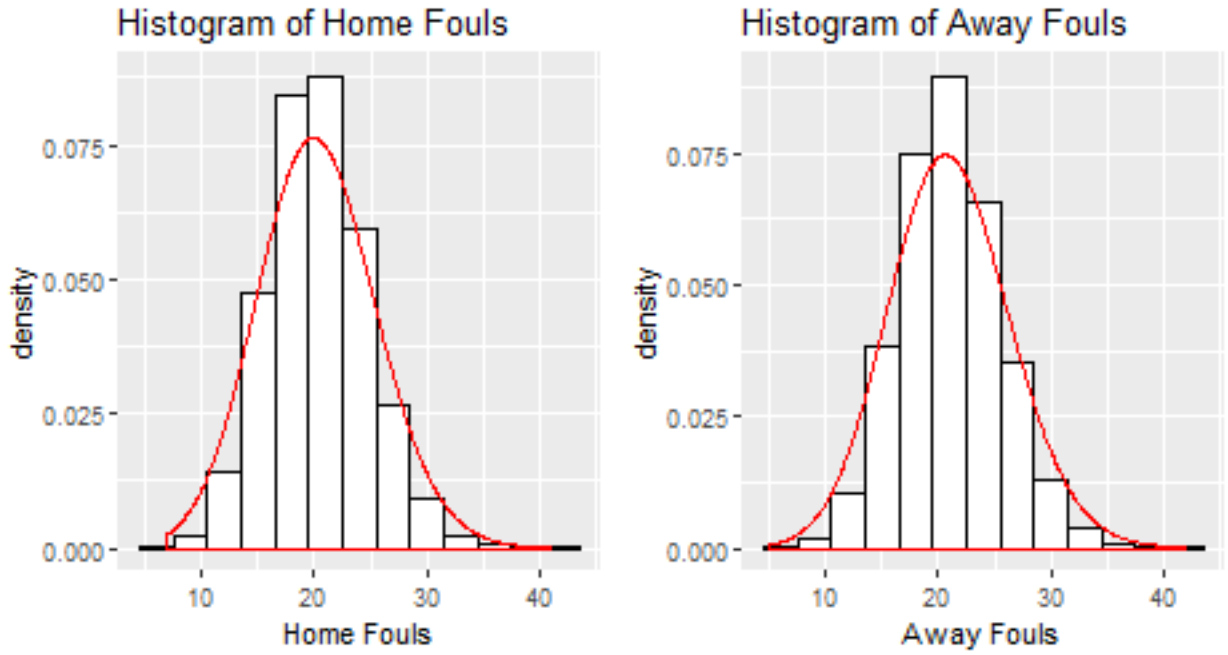
If a home team is returning from a road trip, the away team is expected to score -0.046 more points per hundred miles on the return trip for the home team. The impact of this travel, if the home team traverses the greatest distance between venues on their return trip, can result in the away team scoring as many as 1.235 more points in a given game. This also means that it is advantageous for an away team to play in the first game of the home stand of their hosts. It is interesting to note that the number of points scored by the home team is not impacted by their own travel.

Home Game Number

The home game number, which represents how far into the season both teams are, has a positive coefficient of 0.012 for the away team only. This means that an away team playing in the last scheduled game for the home team will score 0.945 points more than they would score, with all else equal, if it were the first game of the home team. As with the home miles traveled, there is no impact introduced by the *HomeGameNum* on the home team score.

Personal Fouls

The number of fouls was proven earlier as being advantageous toward the home team (by less than one foul per game). This advantage having roots in the schedule was further analyzed.



Model Creation

Models were created by stepwise fitting a linear regression model using the baseline and schedule-based variables described previously. The summaries of the resulting models can be found below.

Table 9: Summary of Home Fouls Model

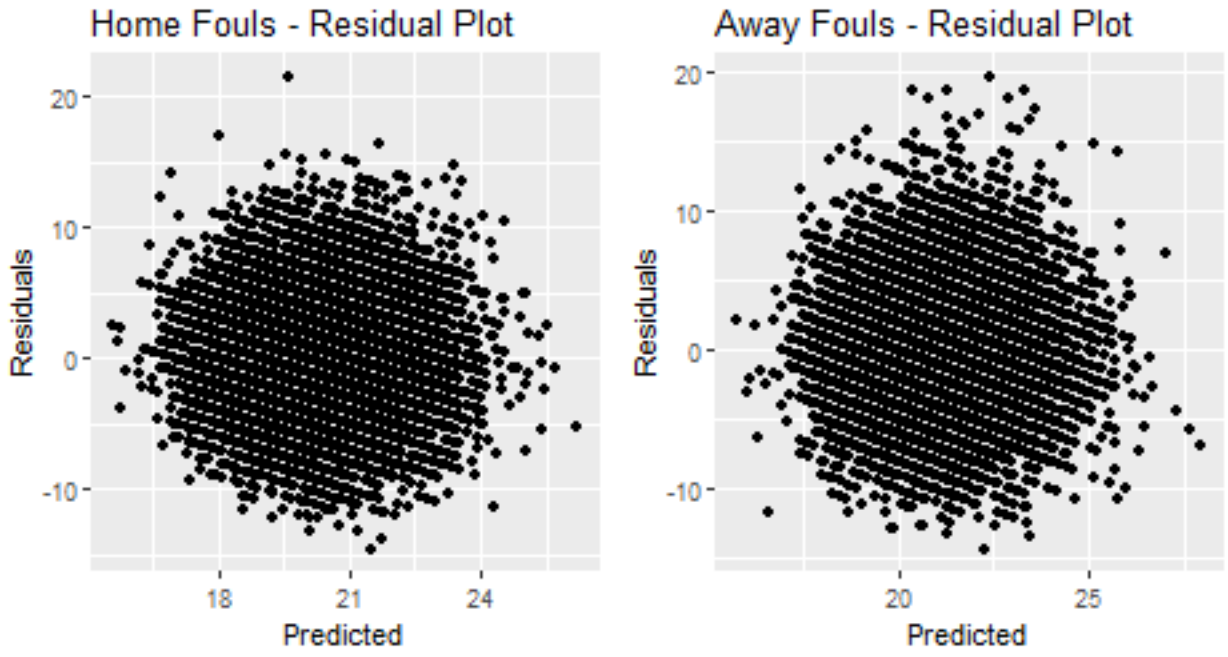
	Estimate	Std.Err	t.Value	p.Value
(Intercept)	7.9975801	0.4489155	17.815334	< 2.2e-16
PrevHomeFouls	0.3356573	0.0133218	25.196071	< 2.2e-16
PrevOppAwayFouls	0.3101760	0.0134584	23.046994	< 2.2e-16
Attendance	-0.0640932	0.0135272	-4.738083	2.179e-06

Table 10: Summary of Away Fouls Model

	Estimate	Std.Err	t.Value	p.Value
(Intercept)	5.8760115	0.3823149	15.36956	< 2.2e-16
PrevAwayFouls	0.3724664	0.0139313	26.73595	< 2.2e-16
PrevOppHomeFouls	0.3601459	0.0138982	25.91311	< 2.2e-16

Model Diagnostics

Plotting the actual home scores versus the residual values produced by the models resulted in the following residual plots.



The above plots show that the regression is a solid fit and does not show characteristics which would tend to indicate issues in the model. Additional diagnostic plots, including a scale-location plot, can be found in the appendix. Both scale-location plots have definitive lattice-like patterns for plotted standardized residual values (y-axis) less than approximately 0.8. While finding patterns in diagnostic plots can indicate underlying issues in the model, the key use of this plot is to test for homoscedasticity. Given that the plot does not

widen as the predicted value increases as well as the discrete nature of the dependent variable, this pattern is considered to be acceptable and not a cause for concern for these regression models.

Model Interpretation

Both regression models ultimately included both baseline stats in the final variable selection. In fact, both foul models have approximately equal estimates between the team baseline stat (*PrevHomeFouls*/*PrevAwayFouls*) and the opponent allowed baseline stat (*PrevOppAwayFouls*/*PrevOppHomeFouls*). The home foul model includes a negative coefficient estimate of approximately -0.064 for each 1000 fans in attendance at a particular game. This means that with approximately 16000 fans, the home team has one foul less called against them. Given that the Q1 to Q3 range for attendance (in thousands) is [15.92, 19.56] and the mean value is 17.51, most games have the home team aided by having one less foul called during the game than the away team. This also means that teams that have a smaller attendance will be less likely to have the reduction in fouls for the home team. The flip of that is also true: a home team that has extremely high attendance is more likely to have this reduction in fouls amplified. One potential explanation for this is that the officials are more leniant toward the home team in order to please the crowd. The officials, in general, call a similiar-but-not-equal number of fouls between the two teams. This model potentially identifies the reason for the statistical inequality previously tested and confirmed. Alternatively, this may be explained by the home team playing higher quality basketball due to the energy of the crowd. Regardless of the reason why, the size of the home crowd is advantageous for the home team in terms of fouls called during a home game.

Conclusion

The impact of scheduling can be seen in the performance of teams in the NBA. From the number of rest days to the distance that the away team is from home, the scheduling impact can be seen directly in the score line of any NBA game. Additionally, the only scheduling variable which impacted both points and fouls in a game was the number of fans that a given team was able to have for a given game. With these advantages measured, the impact of scheduling can be efficiently included in a Monte Carlo simulation to predict the expected outcome of NBA games in the currently scheduled season.

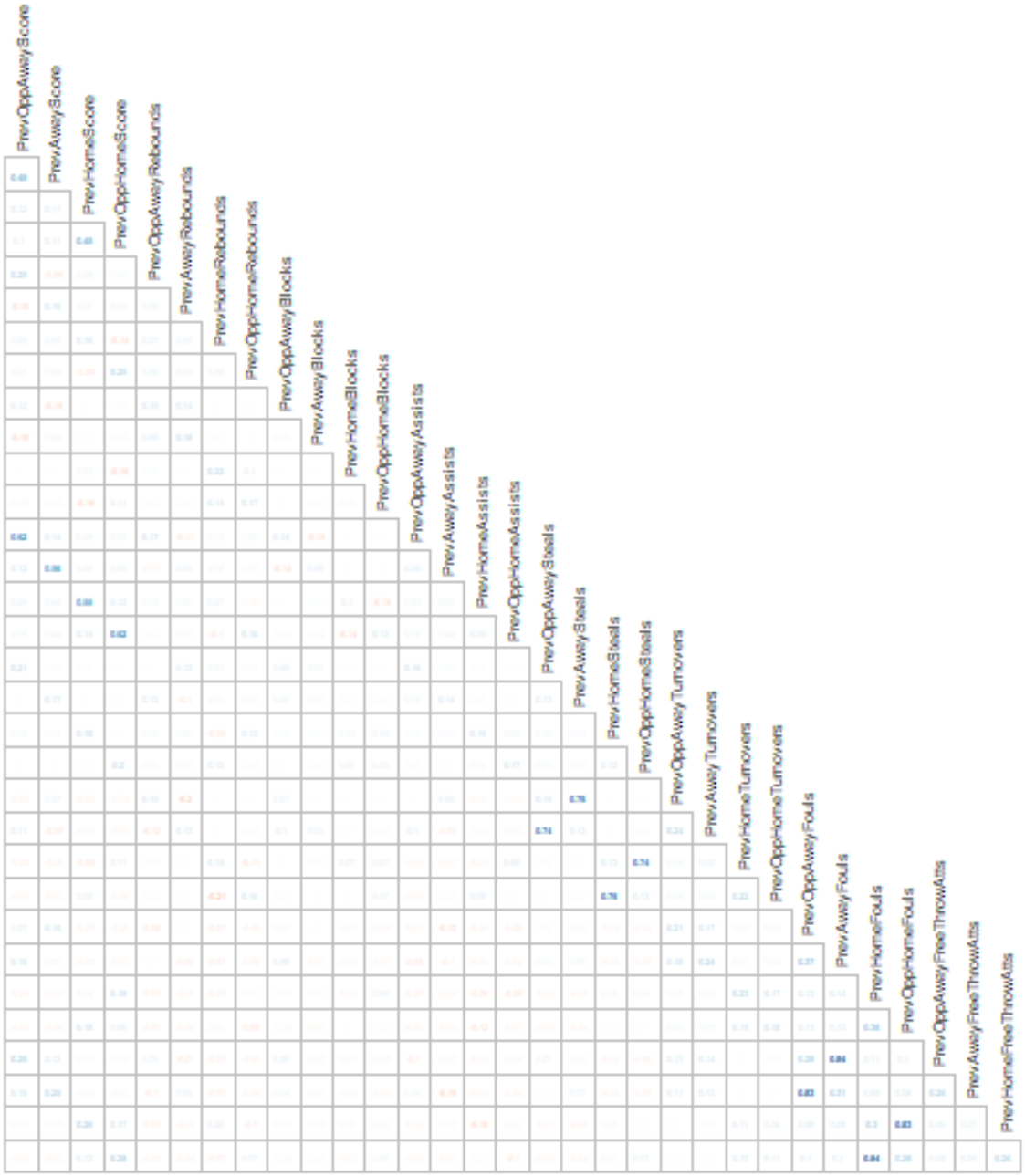
Appendix

Summary Statistics Table

Table 11: Summary Statistics for Game and Schedule Variables

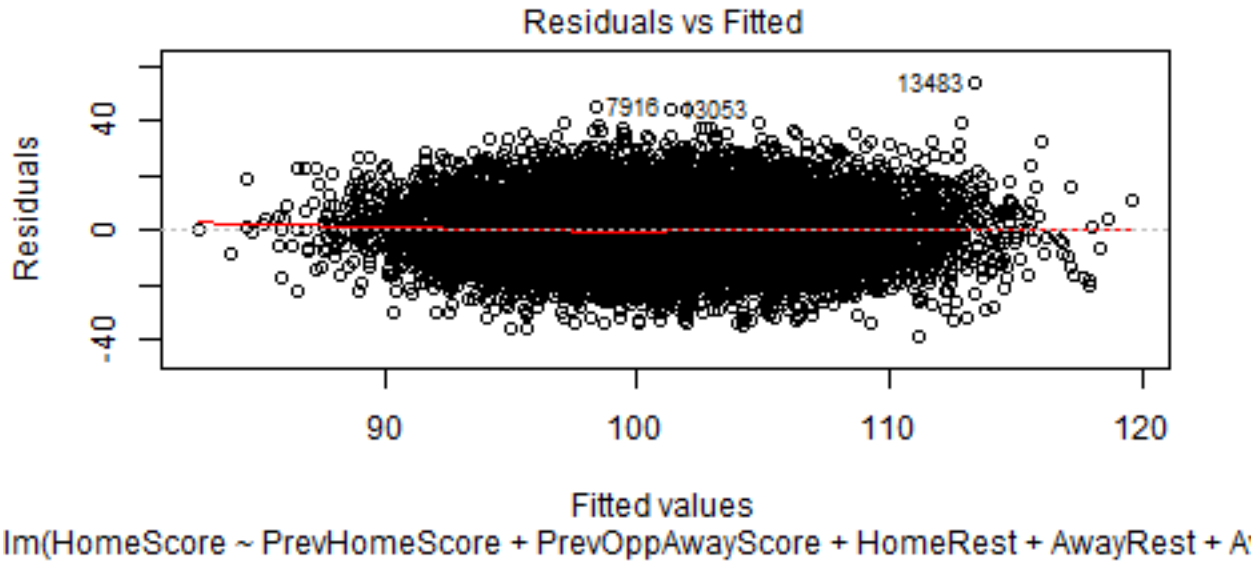
	Min	Q1	Mean	Median	Q3	Max	Std.Dev
AwayScore	54.000	90.000	97.83	98.00	105.00000	151.000	11.70
HomeScore	59.000	93.000	100.80	101.00	109.00000	168.000	11.98
HomeMargin	-51.000	-6.000	2.96	3.00	11.00000	55.000	13.11
AwayRest	0.000	0.000	0.90	1.00	1.00000	4.000	0.83
HomeRest	0.000	1.000	1.18	1.00	1.00000	4.000	0.83
HomeRestAdv	-9.000	0.000	0.28	0.00	1.00000	9.000	1.10
AwayMilesAway	0.000	5.330	10.57	9.39	14.94000	27.320	6.48
AwayMilesTraveled	0.000	3.610	7.02	5.98	9.36000	42.210	4.72
HomeMilesTraveled	0.000	0.000	4.07	0.00	6.95750	47.030	5.55
Attendance	1.557	15.918	17.51	18.20	19.56175	23.152	2.63
AwayAssists	4.000	17.000	20.94	21.00	24.00000	43.000	4.98
HomeAssists	6.000	19.000	22.41	22.00	26.00000	47.000	5.09
AwaySteals	0.000	5.000	7.43	7.00	9.00000	20.000	2.87
HomeSteals	0.000	5.000	7.52	7.00	9.00000	22.000	2.90
AwayTurnovers	3.000	11.000	13.86	14.00	16.00000	28.000	3.86
HomeTurnovers	2.000	11.000	13.51	13.00	16.00000	29.000	3.79
AwayRebounds	20.000	37.000	41.50	41.00	46.00000	81.000	6.43
HomeRebounds	17.000	38.000	42.79	43.00	47.00000	72.000	6.51
AwayBlocks	0.000	3.000	4.54	4.00	6.00000	17.000	2.44
HomeBlocks	0.000	3.000	5.09	5.00	7.00000	19.000	2.64
AwayFouls	5.000	18.000	21.06	21.00	24.00000	42.000	4.51
HomeFouls	7.000	17.000	20.29	20.00	23.00000	41.000	4.33
AwayFreeThrowAtts	1.000	18.000	23.41	23.00	28.00000	64.000	7.45
HomeFreeThrowAtts	4.000	19.000	24.53	24.00	29.00000	63.000	7.81
AwayFreeThrowMakes	1.000	13.000	17.72	17.00	22.00000	52.000	6.09
HomeFreeThrowMakes	1.000	14.000	18.61	18.00	23.00000	48.000	6.32
AwayThreePointMakes	0.000	5.000	7.07	7.00	9.00000	25.000	3.34
HomeThreePointMakes	0.000	5.000	7.20	7.00	9.00000	24.000	3.41

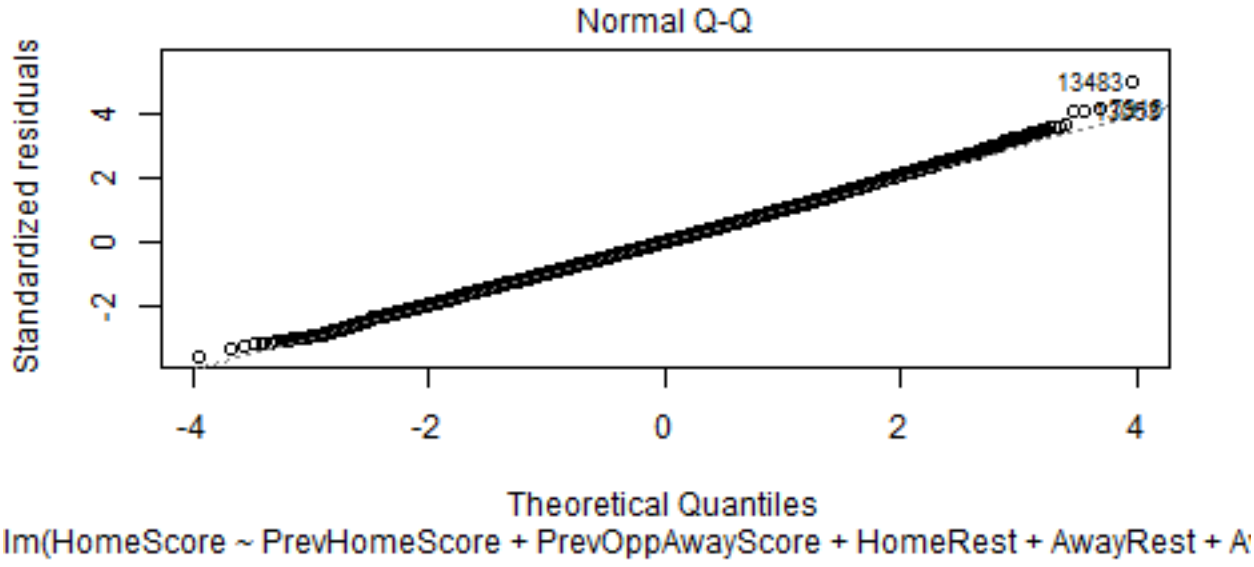
Game Statistics Correlation Matrix

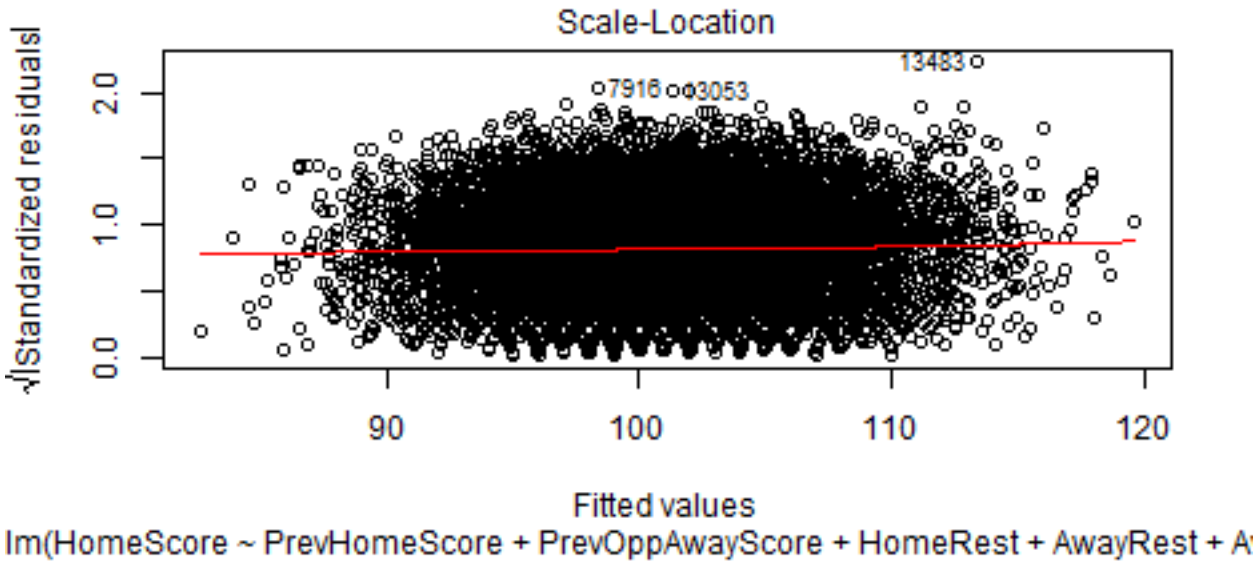


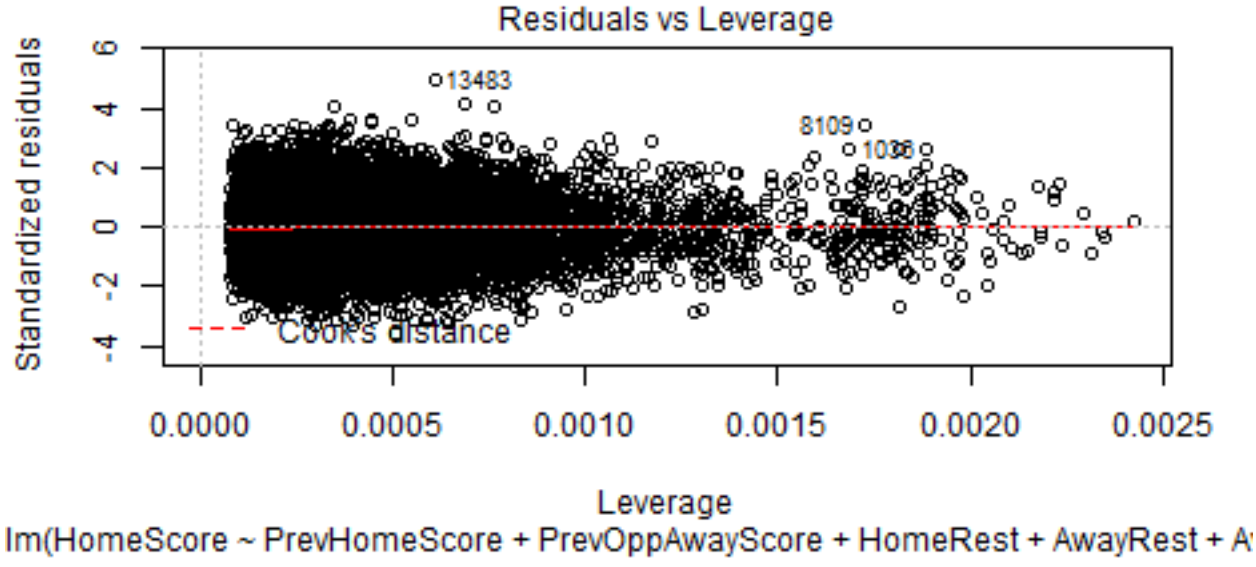
Arena Distance Chart

Home Score - Diagnostic Plots

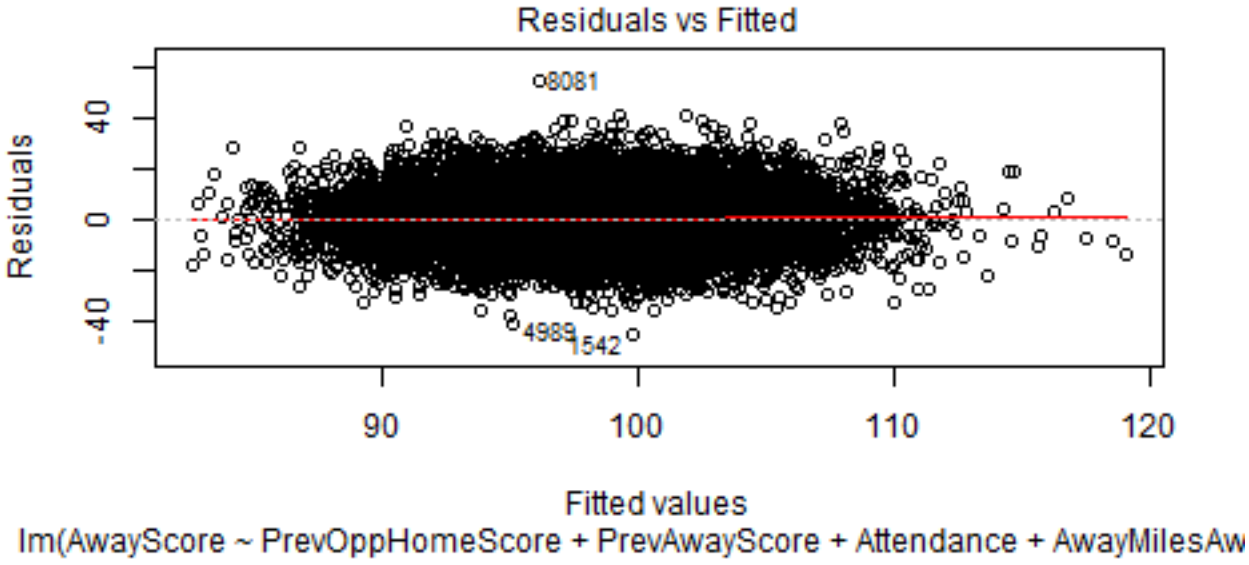


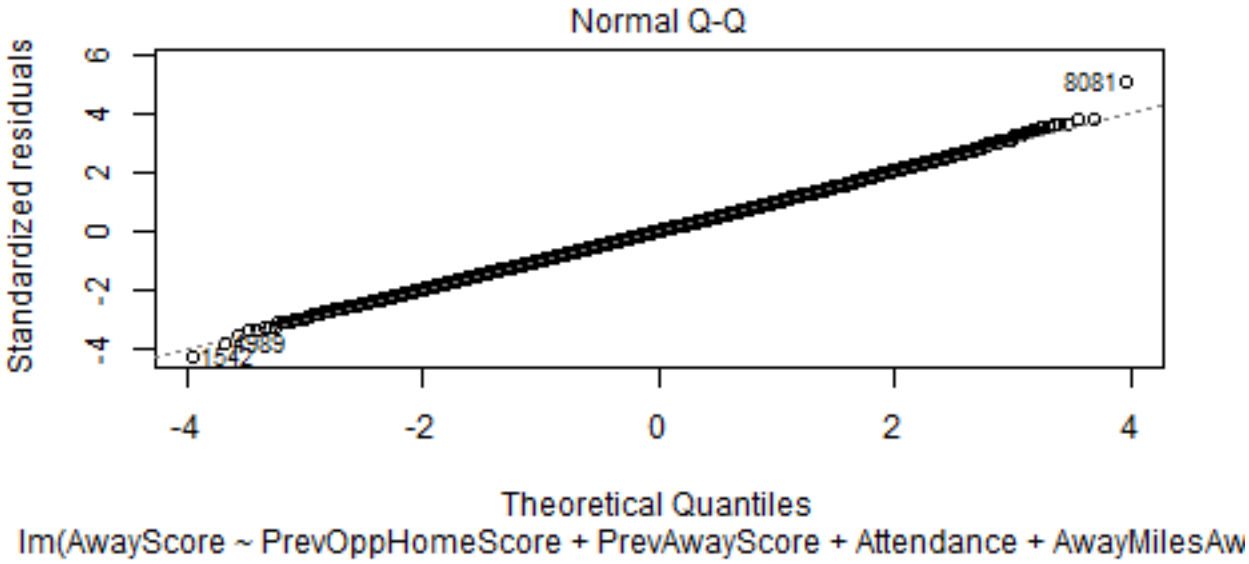


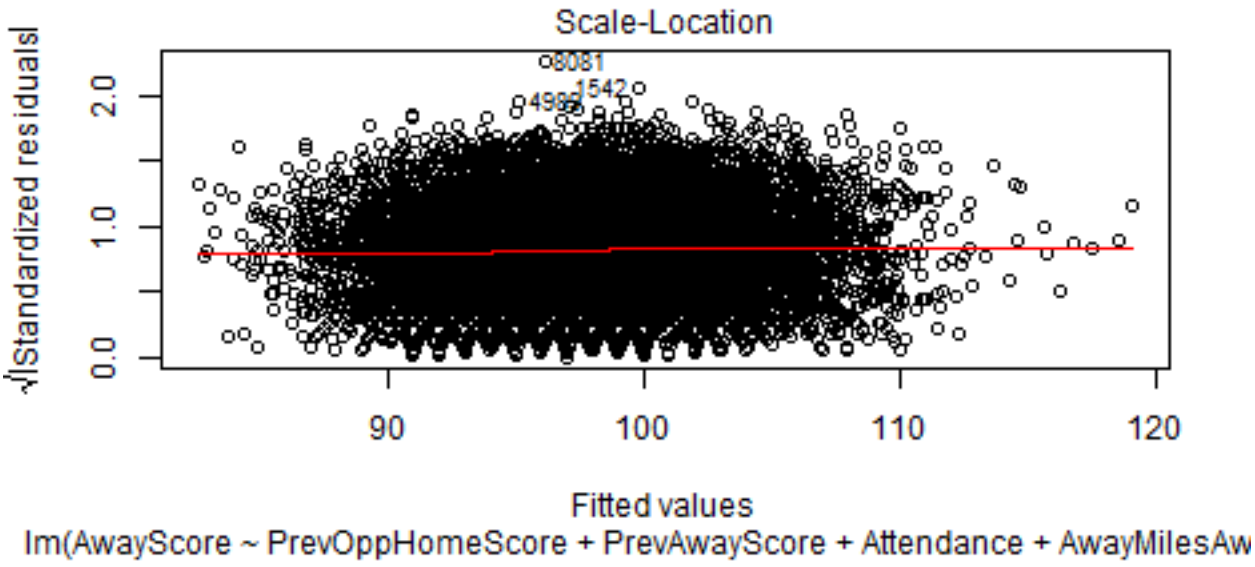

Scale-Location
√|Standardized residuals|
0.0
1.0
2.0
7916
3053
13483
90
100
110
120
Fitted values
lm(HomeScore ~ PrevHomeScore + PrevOppAwayScore + HomeRest + AwayRest + A

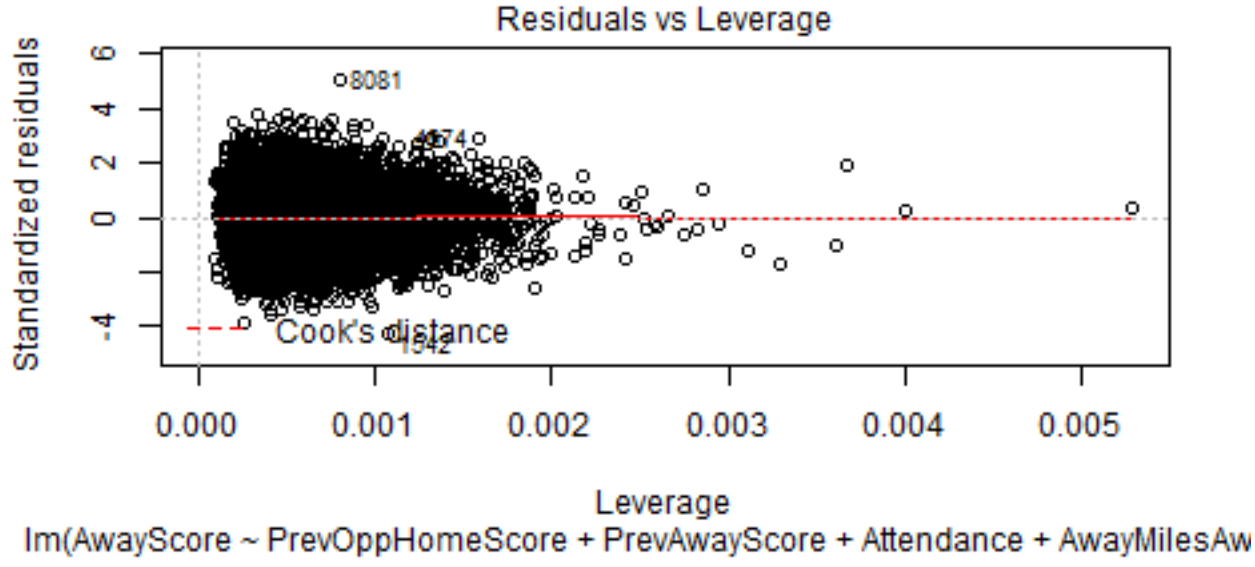

Residuals vs Leverage
Standardized residuals
-4
-2
0
2
4
6
13483
8109
1036
Cook's distance
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
Leverage
lm(HomeScore ~ PrevHomeScore + PrevOppAwayScore + HomeRest + AwayRest + A

Away Score - Diagnostic Plots

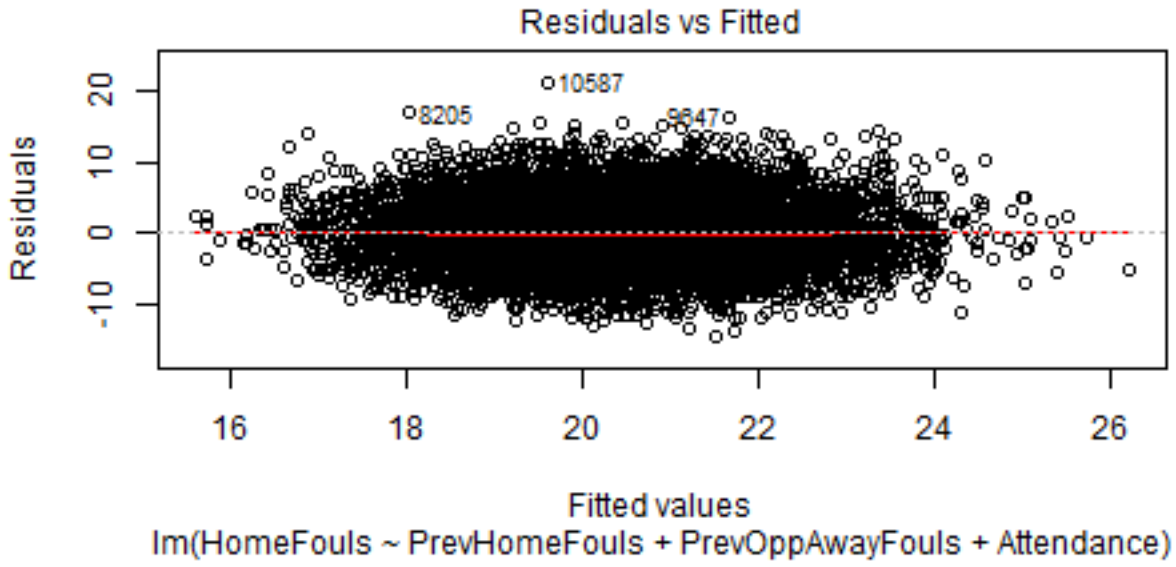


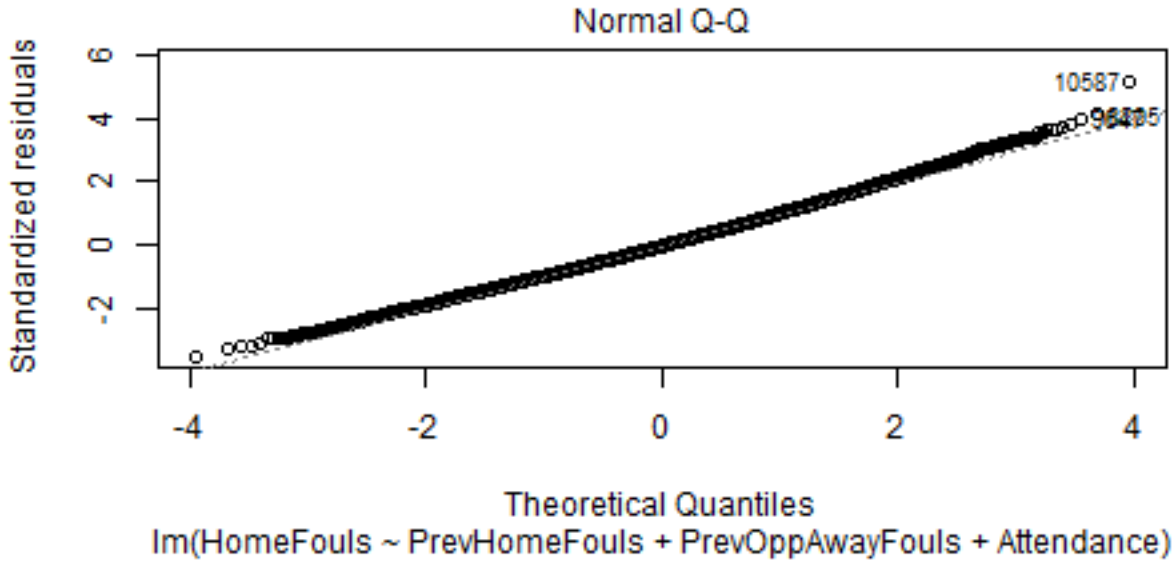


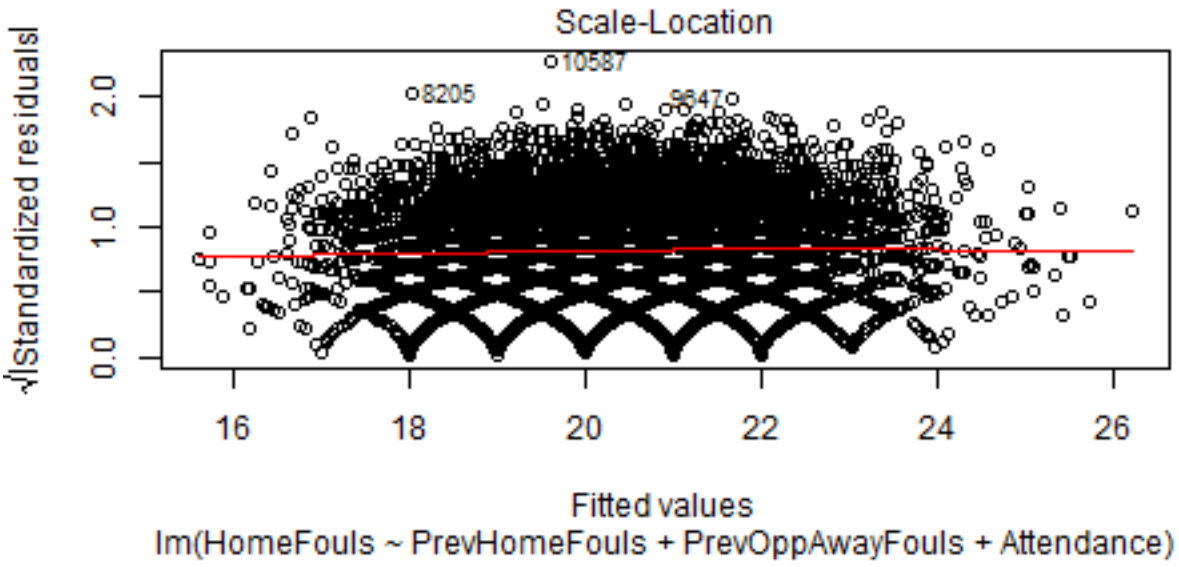

Scale-Location
√|Standardized residuals|
8081
1542
4982
0.0
1.0
2.0
90
100
110
120
Fitted values
lm(AwayScore ~ PrevOppHomeScore + PrevAwayScore + Attendance + AwayMilesAw

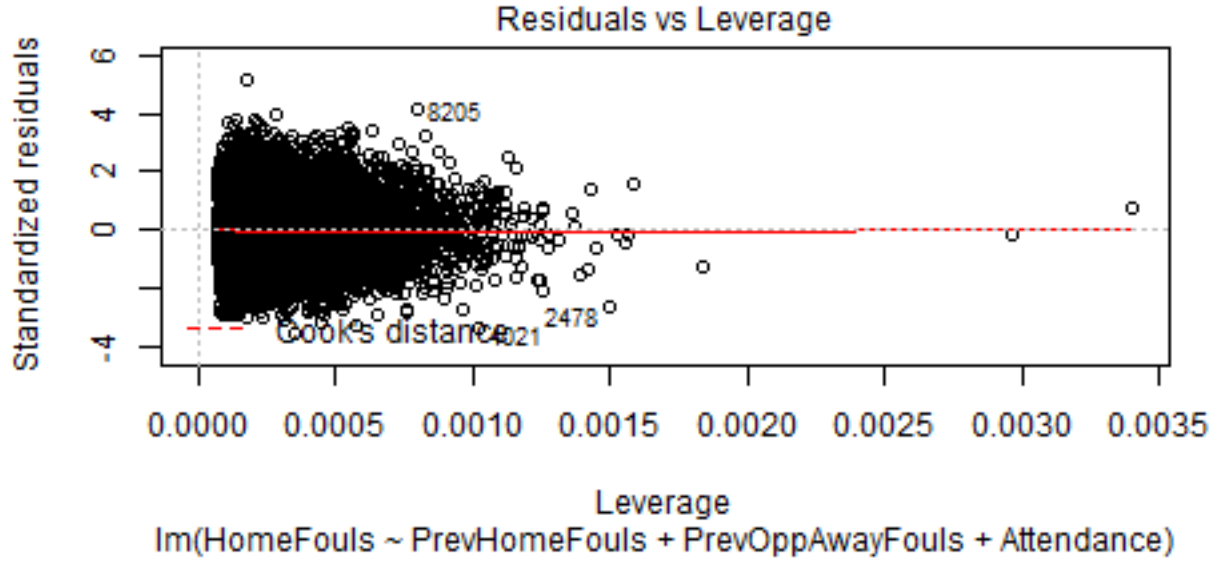

Residuals vs Leverage
Standardized residuals
8081
4674
1542
Cook's distance
-4
0
2
4
6
0.000
0.001
0.002
0.003
0.004
0.005
Leverage
lm(AwayScore ~ PrevOppHomeScore + PrevAwayScore + Attendance + AwayMilesAw

Home Fouls - Diagnostic Plots

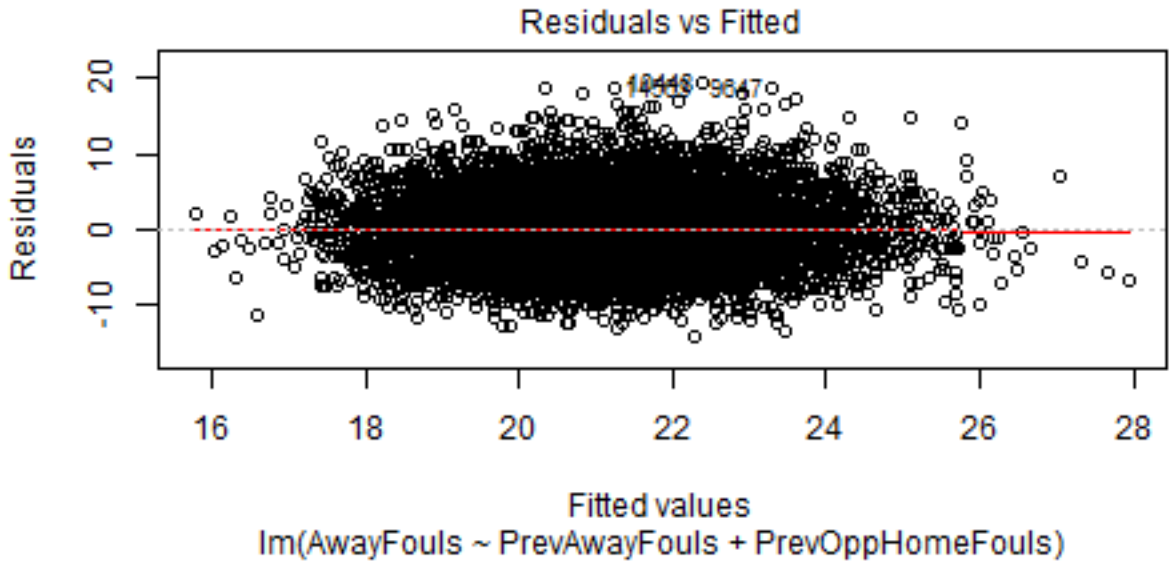


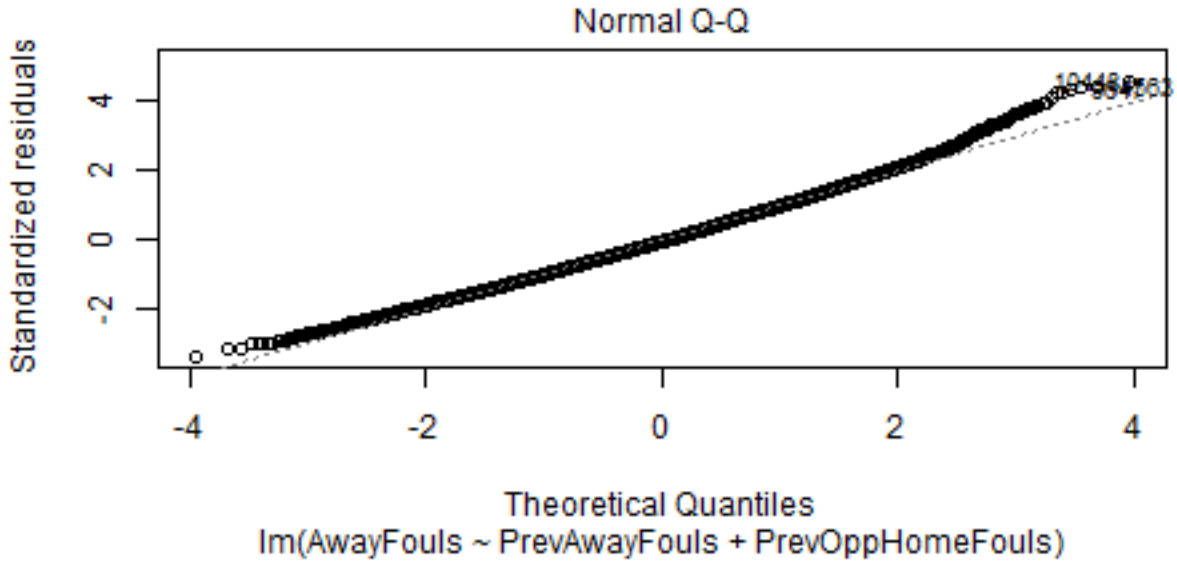


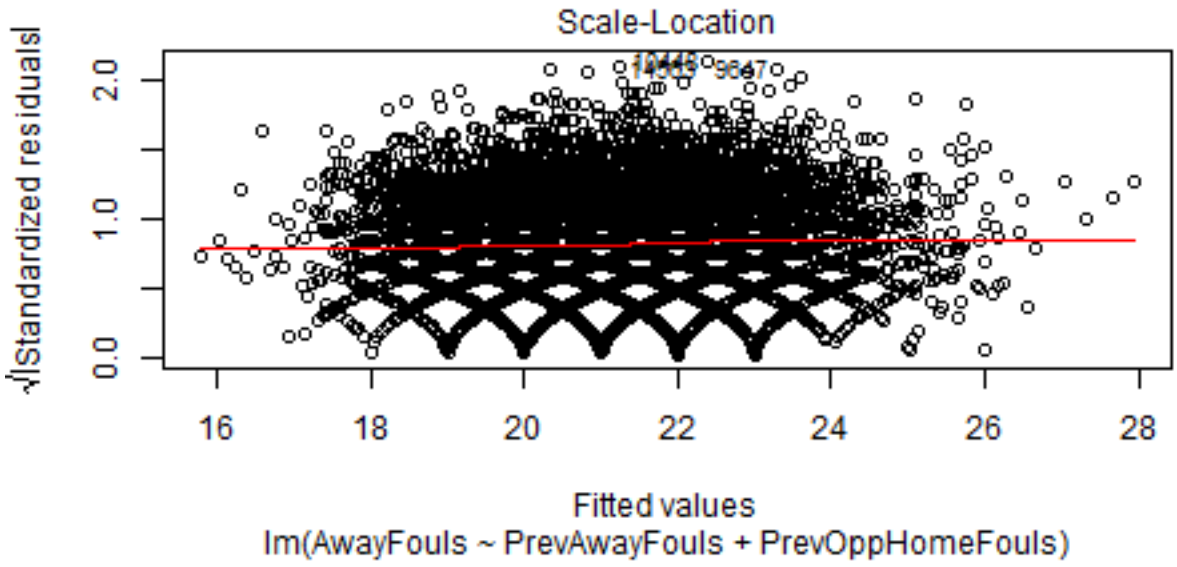

Scale-Location
10587
8205
9647
√|Standardized residuals|
0.0
1.0
2.0
16
18
20
22
24
26
Fitted values
lm(HomeFouls ~ PrevHomeFouls + PrevOppAwayFouls + Attendance)

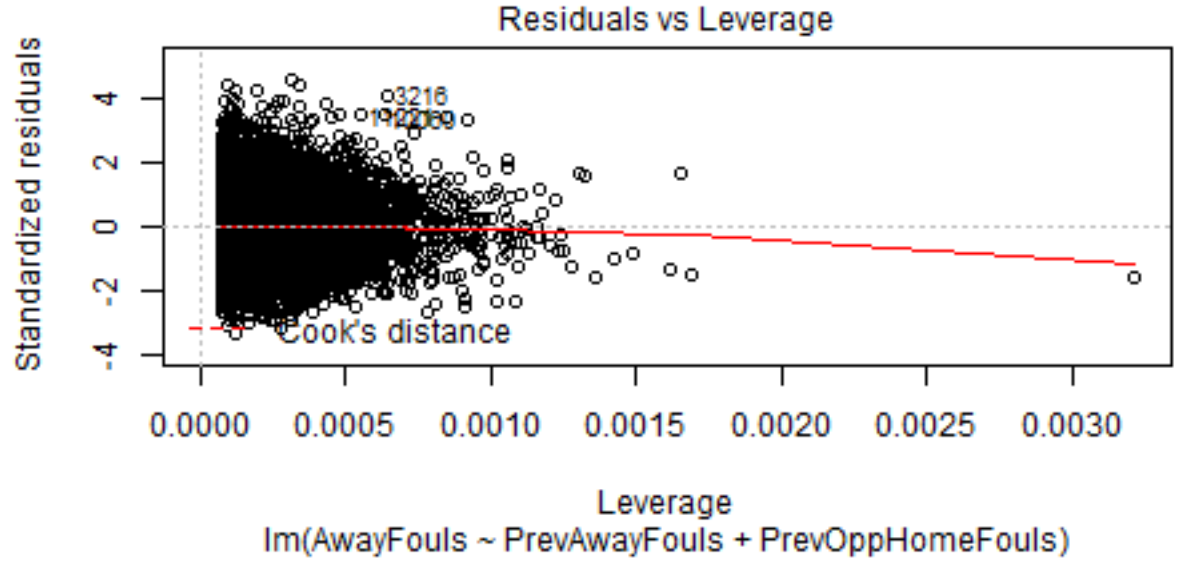

Residuals vs Leverage
8205
2478
4021
Cook's distance
Standardized residuals
-4
0
2
4
6
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
0.0035
Leverage
lm(HomeFouls ~ PrevHomeFouls + PrevOppAwayFouls + Attendance)

Away Fouls - Diagnostic Plots






Scale-Location
√|Standardized residuals|
0.0
1.0
2.0
16
18
20
22
24
26
28
Fitted values
lm(AwayFouls ~ PrevAwayFouls + PrevOppHomeFouls)


Residuals vs Leverage
Standardized residuals
-4
-2
0
2
4
3216
Cook's distance
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
Leverage
lm(AwayFouls ~ PrevAwayFouls + PrevOppHomeFouls)

NbaGames.sql - Query for NBA game data used

```
## USE FoxSports_SportsData_NBA
## GO
##
## DECLARE @regSeasonGameTypeID INT = 1;
## DECLARE @playoffsGameTypeID INT = 2;
## DECLARE @preseasonGameTypeID INT = 3;
## DECLARE @finalGameStatusID INT = 2;
## DECLARE @forfeitByAwayStatusID INT = 4;
## DECLARE @forfeitByHomeStatusID INT = 5;
## DECLARE @forfeitByBothStatusID INT = 7;
## DECLARE @srid INT = 4326; -- SRID Geography TYPE. DEFINED BY EXTERNAL STANDARD
## DECLARE @metersToMiles REAL = (0.62137119/1000);
## DECLARE @pastGameCount INT = 4;
## DECLARE @minSeason INT = 2005;
## DECLARE @maxSeason INT = 2016;
##
## WITH RegulationScores_CTE AS (
##  SELECT sch.GameID as 'GameID'
##      , CAST(REPLACE(LEFT(AwayScoreByQuarter, 2), ';', '') AS TINYINT) +
##      CAST(REPLACE(SUBSTRING(AwayScoreByQuarter,
##                  CHARINDEX(';', AwayScoreByQuarter,1) + 1, 2), ';','') AS TINYINT) +
##      CAST(REPLACE(SUBSTRING(AwayScoreByQuarter,
##                  CHARINDEX(';', AwayScoreByQuarter,
##                  CHARINDEX(';', AwayScoreByQuarter, 1) + 1) + 1,2), ';','') AS TINYINT) +
##      CAST(REPLACE(SUBSTRING(AwayScoreByQuarter,
##                  CHARINDEX(';', AwayScoreByQuarter,
##                  CHARINDEX(';', AwayScoreByQuarter, 1) +
##                  CHARINDEX(';', AwayScoreByQuarter, 1) + 1) + 1,2), ';','') AS TINYINT)
##          as 'AwayScore'
##      , CAST(REPLACE(LEFT(HomeScoreByQuarter, 2), ';', '') AS TINYINT) +
##      CAST(REPLACE(SUBSTRING(HomeScoreByQuarter,
##                  CHARINDEX(';', HomeScoreByQuarter,1) + 1, 2), ';','') AS TINYINT) +
##      CAST(REPLACE(SUBSTRING(HomeScoreByQuarter,
##                  CHARINDEX(';', HomeScoreByQuarter,
##                  CHARINDEX(';', HomeScoreByQuarter, 1) + 1) + 1,2), ';','') AS TINYINT) +
##      CAST(REPLACE(SUBSTRING(HomeScoreByQuarter,
##                  CHARINDEX(';', HomeScoreByQuarter,
##                  CHARINDEX(';', HomeScoreByQuarter, 1) +
##                  CHARINDEX(';', HomeScoreByQuarter, 1) + 1) + 1,2), ';','') AS TINYINT)
##          as 'HomeScore'
##  FROM GameSchedules as sch
## ),
## HomeStadiums_CTE AS (
##  SELECT x.Season, x.TeamID, x.StadiumID, s.Latitude, s.Longitude
##  FROM (
##      SELECT t.Season
##          , t.TeamID
##          , t.StadiumID
##          , ROW_NUMBER() OVER (PARTITION BY t.Season, t.TeamID ORDER BY GameCount desc)
##              as 'StadiumRank'
##      FROM (
##          SELECT ts.Season, ts.TeamID, sch.StadiumID, COUNT(*) as 'GameCount'
```

```
##          FROM TeamSeasons as ts
##          INNER JOIN GameSchedules as sch ON ts.Season = sch.Season
##              AND ts.TeamID = sch.HomeTeamID
##          WHERE sch.GameTypeID IN (@playoffsGameTypeID, @regSeasonGameTypeID)
##          GROUP BY ts.Season, ts.TeamID, sch.StadiumID
##      ) as t
##  ) as x
##  INNER JOIN Stadiums as s ON x.StadiumID = s.StadiumID
##  WHERE x.StadiumRank = 1
## ),
## SeasonGameOrder_CTE AS (
##  SELECT sch.GameID
##      , sch.GameDate
##      , ts.Season
##      , ts.TeamID
##      , ROW_NUMBER() OVER (PARTITION BY ts.Season, ts.TeamID ORDER BY GameDate asc)
##          as 'GameOrder'
##  FROM TeamSeasons as ts
##  INNER JOIN GameSchedules as sch ON ts.Season = sch.Season
##      AND ts.TeamID IN (sch.HomeTeamID, sch.AwayTeamID)
##  WHERE sch.GameTypeID IN (@playoffsGameTypeID, @regSeasonGameTypeID)
##      AND sch.StatusID IN (@finalGameStatusID
##          , @forfeitByAwayStatusID
##          , @forfeitByBothStatusID
##          , @forfeitByHomeStatusID)
## ),
## TeamPrevGame_CTE AS (
##  SELECT ts.TeamID
##      , gmOrd_cur.GameID as 'CurrentGameID'
##      , gmOrd_lst.GameID as 'PrevGameID'
##      , ABS(DATEDIFF(DAY, gmOrd_cur.GameDate, gmOrd_lst.GameDate)) as 'RestDays'
##      , gmOrd_cur.GameOrder as 'CurGameNumber'
##  FROM TeamSeasons as ts
##  INNER JOIN SeasonGameOrder_CTE as gmOrd_cur ON ts.Season = gmOrd_cur.Season
##      AND ts.TeamID = gmOrd_cur.TeamID
##  INNER JOIN SeasonGameOrder_CTE as gmOrd_lst ON gmOrd_cur.Season = gmOrd_lst.Season
##      AND gmOrd_cur.TeamID = gmOrd_lst.TeamID
##      AND gmOrd_cur.GameOrder = gmOrd_lst.GameOrder + 1
## ),
## TeamGameBoxscores_CTE AS (
##  SELECT tBox.GameID
##      , tBox.TeamID
##      , tBox.Season
##      , CASE WHEN (tBox.TeamID = sch.HomeTeamID AND sch.HomeScore > sch.AwayScore)
##                          OR (tBox.TeamID = sch.AwayTeamID AND sch.HomeScore < sch.AwayScore)
##              THEN 1
##              WHEN sc.HomeScore = sc.AwayScore
##              THEN 0.5
##              ELSE 0
##              END as 'GameResult'
##      , CASE WHEN (tBox.TeamID = sch.AwayTeamID AND sc.HomeScore > sc.AwayScore)
##                          OR (tBox.TeamID = sch.HomeTeamID AND sc.HomeScore < sc.AwayScore)
##              THEN 1
##              WHEN sc.HomeScore = sc.AwayScore
```

```
##              THEN 0.5
##              ELSE 0
##              END as 'OppGameResult'
##      , CASE WHEN tBox.TeamID = sch.HomeTeamID
##              THEN sc.HomeScore
##              ELSE sc.AwayScore END as 'Score'
##      , CASE WHEN tBox.TeamID = sch.AwayTeamID
##              THEN sc.HomeScore
##              ELSE sc.AwayScore END as 'OppScore'
##      , tBox.Assists as 'Assists'
##      , tOppBox.Assists as 'OppAssists'
##      , tBox.Rebounds as 'Rebounds'
##      , tOppBox.Rebounds as 'OppRebounds'
##      , tBox.Steals as 'Steals'
##      , tOppBox.Steals as 'OppSteals'
##      , tBox.Blocks as 'Blocks'
##      , tOppBox.Blocks as 'OppBlocks'
##      , tBox.FreeThrowsAttempted as 'FreeThrowAtts'
##      , tOppBox.FreeThrowsAttempted as 'OppFreeThrowAtts'
##      , tBox.FreeThrowsMade as 'FreeThrowMakes'
##      , tOppBox.FreeThrowsMade as 'OppFreeThrowMakes'
##      , tBox.ThreePointFieldGoalsMade as 'ThreePointMakes'
##      , tOppBox.ThreePointFieldGoalsMade as 'OppThreePointMakes'
##      , tBox.PersonalFouls as 'Fouls'
##      , tOppBox.PersonalFouls as 'OppFouls'
##      , tBox.Turnovers as 'Turnovers'
##      , tOppBox.Turnovers as 'OppTurnovers'
##  FROM TeamBoxscoreStats as tBox
##  INNER JOIN GameSchedules as sch ON tBox.GameID = sch.GameID
##  INNER JOIN TeamBoxscoreStats as tOppBox ON tBox.GameID = tOppBox.GameID
##      AND tBox.TeamID <> tOppBox.TeamID
##  INNER JOIN RegulationScores_CTE as sc ON sch.GameID = sc.GameID
##  WHERE sch.GameTypeID IN (@playoffsGameTypeID, @regSeasonGameTypeID)
##      AND sch.StatusID = @finalGameStatusID
##      AND sch.Season >= @minSeason
##      AND sch.Season <= @maxSeason
## ),
## GameWindow_CTE as (
##  SELECT cur.TeamID, cur.Season, cur.GameID as 'CurGameID', prev.GameID as 'PrevGameID'
##  FROM SeasonGameOrder_CTE as cur
##  INNER JOIN SeasonGameOrder_CTE as prev ON cur.Season = prev.Season
##      AND cur.TeamID = prev.TeamID
##      AND cur.GameOrder > prev.GameOrder
##      AND (cur.GameOrder - @pastGameCount) <= prev.GameOrder
##  WHERE cur.GameOrder > @pastGameCount
## ),
## PrevGameStats_CTE as (
##  SELECT sch.GameID as 'GameID'
##      , ts.TeamID as 'TeamID'
##      , COUNT(*) as 'GameCnt'
##      , AVG(CAST(pBox_past.GameResult AS REAL)) as 'GameResult'
##      , AVG(CAST(pBox_past.OppGameResult AS REAL)) as 'OppGameResult'
##      , AVG(CAST(pBox_past.Score AS REAL)) as 'Score'
##      , AVG(CAST(pBox_past.OppScore AS REAL)) as 'OppScore'
```

```
##      , AVG(CAST(pBox_past.Assists AS REAL)) as 'Assists'
##      , AVG(CAST(pBox_past.OppAssists AS REAL)) as 'OppAssists'
##      , AVG(CAST(pBox_past.Rebounds AS REAL)) as 'Rebounds'
##      , AVG(CAST(pBox_past.OppRebounds AS REAL)) as 'OppRebounds'
##      , AVG(CAST(pBox_past.Steals AS REAL)) as 'Steals'
##      , AVG(CAST(pBox_past.OppSteals AS REAL)) as 'OppSteals'
##      , AVG(CAST(pBox_past.Blocks AS REAL)) as 'Blocks'
##      , AVG(CAST(pBox_past.OppBlocks AS REAL)) as 'OppBlocks'
##      , AVG(CAST(pBox_past.FreeThrowAtts AS REAL)) as 'FreeThrowAtts'
##      , AVG(CAST(pBox_past.OppFreeThrowAtts AS REAL)) as 'OppFreeThrowAtts'
##      , AVG(CAST(pBox_past.FreeThrowMakes AS REAL)) as 'FreeThrowMakes'
##      , AVG(CAST(pBox_past.OppFreeThrowMakes AS REAL)) as 'OppFreeThrowMakes'
##      , AVG(CAST(pBox_past.Fouls AS REAL)) as 'Fouls'
##      , AVG(CAST(pBox_past.OppFouls AS REAL)) as 'OppFouls'
##      , AVG(CAST(pBox_past.Turnovers AS REAL)) as 'Turnovers'
##      , AVG(CAST(pBox_past.OppTurnovers AS REAL)) as 'OppTurnovers'
##      , AVG(CAST(pBox_past.ThreePointMakes AS REAL)) as 'ThreePointMakes'
##      , AVG(CAST(pBox_past.OppThreePointMakes AS REAL)) as 'OppThreePointMakes'
##  FROM TeamSeasons as ts
##  INNER JOIN GameSchedules as sch ON ts.Season = sch.Season
##      AND ts.TeamID IN (sch.HomeTeamID, sch.AwayTeamID)
##  INNER JOIN GameWindow_CTE as win ON sch.GameID = win.CurGameID
##      AND ts.Season = win.Season
##      AND ts.TeamID = win.TeamID
##  INNER JOIN TeamGameBoxscores_CTE as pBox_past ON win.PrevGameID = pBox_past.GameID
##      AND pBox_past.TeamID = ts.TeamID
##  INNER JOIN TeamGameBoxscores_CTE as pOppBox_past ON pBox_past.GameID = pOppBox_past.GameID
##      AND pBox_past.TeamID <> pOppBox_past.TeamID
##  WHERE sch.GameTypeID IN (@playoffsGameTypeID, @regSeasonGameTypeID)
##      AND sch.StatusID = @finalGameStatusID
##      AND sch.Season >= @minSeason
##      AND sch.Season <= @maxSeason
##  GROUP BY sch.GameID, ts.TeamID
## )
##
## SELECT sch.GameID as 'GameID'
##  , sch.Season
##  , CASE WHEN sch.GameTypeID = @regSeasonGameTypeID THEN 1
##      WHEN sch.GameTypeID = @playoffsGameTypeID THEN 2
##      WHEN sch.GameTypeID = @preseasonGameTypeID THEN 3
##      ELSE NULL END as 'GameTypeID'
##  , ats.Alias as 'AwayTeam'
##  , hts.Alias as 'HomeTeam'
##  , sc.AwayScore as 'AwayScore'
##  , sc.HomeScore as 'HomeScore'
##  , CAST(sch.GameDate AS DATE) as 'GameDate'
##  , CASE WHEN sch.StadiumID = hHmStadLookup.StadiumID
##          AND sch.HomeTeamID = hHmStadLookup.TeamID
##          AND sch.Season = hHmStadLookup.Season
##      THEN 0
##      ELSE 1
##      END as 'IsNeutralSite'
##  , sch.TotalQuarters - 4 as 'OvertimePeriods'
##  , sch.Attendance as 'Attendance'
```

```
##   , aPrevGmLookup.RestDays - 1 as 'AwayRest'
##   , hPrevGmLookup.RestDays - 1 as 'HomeRest'
##   , aPrevGmLookup.CurGameNumber as 'AwayGameNum'
##   , hPrevGmLookup.CurGameNumber as 'HomeGameNum'
##   , CAST(ABS(
##          GEOGRAPHY::Point(stad.Latitude, stad.Longitude, @srid)
##          .STDistance(GEOGRAPHY::Point(aPrevGmStad.Latitude, aPrevGmStad.Longitude, @srid)
##       )*@metersToMiles) AS INT) as 'AwayMilesTraveled'
##   , CAST(ABS(
##          GEOGRAPHY::Point(stad.Latitude, stad.Longitude, @srid)
##          .STDistance(GEOGRAPHY::Point(hPrevGmStad.Latitude, hPrevGmStad.Longitude, @srid)
##       )*@metersToMiles) AS INT) as 'HomeMilesTraveled'
##   , CAST(ABS(
##          GEOGRAPHY::Point(stad.Latitude, stad.Longitude, @srid)
##          .STDistance(GEOGRAPHY::Point(aHmStadLookup.Latitude, aHmStadLookup.Longitude, @srid)
##       )*@metersToMiles) AS INT) as 'AwayMilesAway'
##   , CAST(ABS(
##          GEOGRAPHY::Point(stad.Latitude, stad.Longitude, @srid)
##          .STDistance(GEOGRAPHY::Point(hHmStadLookup.Latitude, hHmStadLookup.Longitude, @srid)
##       )*@metersToMiles) AS INT) as 'HomeMilesAway'
##   , CASE WHEN ats.ConferenceID = hts.ConferenceID
##       THEN 1
##       ELSE 0
##       END as 'AreSameConference'
##   , CASE WHEN ats.ConferenceID = hts.ConferenceID AND ats.DivisionID = hts.DivisionID
##       THEN 1
##       ELSE 0
##       END as 'AreSameDivision'
##   , CASE WHEN sch.HomeScore > sch.AwayScore THEN 1 ELSE 0 END as 'HomeResult'
##   , CASE WHEN sch.AwayScore > sch.HomeScore THEN 1 ELSE 0 END as 'AwayResult'
##   , (hPrevGmLookup.RestDays - 1) - (aPrevGmLookup.RestDays - 1) as 'HomeRestAdv'
##   , hBox.Assists as 'HomeAssists'
##   , hBox.Rebounds as 'HomeRebounds'
##   , hBox.Steals as 'HomeSteals'
##   , hBox.Turnovers as 'HomeTurnovers'
##   , hBox.PersonalFouls as 'HomeFouls'
##   , hBox.FreeThrowsAttempted as 'HomeFreeThrowAtts'
##   , hBox.FreeThrowsMade as 'HomeFreeThrowMakes'
##   , hBox.Blocks as 'HomeBlocks'
##   , hBox.ThreePointFieldGoalsMade as 'HomeThreePointMakes'
##   , aBox.Assists as 'AwayAssists'
##   , aBox.Rebounds as 'AwayRebounds'
##   , aBox.Steals as 'AwaySteals'
##   , aBox.Turnovers as 'AwayTurnovers'
##   , aBox.PersonalFouls as 'AwayFouls'
##   , aBox.FreeThrowsAttempted as 'AwayFreeThrowAtts'
##   , aBox.FreeThrowsMade as 'AwayFreeThrowMakes'
##   , aBox.Blocks as 'AwayBlocks'
##   , aBox.ThreePointFieldGoalsMade as 'AwayThreePointMakes'
##   , hpgs.Score as 'PrevHomeScore'
##   , hpgs.Assists as 'PrevHomeAssists'
##   , hpgs.Rebounds as 'PrevHomeRebounds'
##   , hpgs.Steals as 'PrevHomeSteals'
##   , hpgs.Turnovers as 'PrevHomeTurnovers'
```

```
##  , hpgs.Fouls as 'PrevHomeFouls'
##  , hpgs.FreeThrowAtts as 'PrevHomeFreeThrowAtts'
##  , hpgs.FreeThrowMakes as 'PrevHomeFreeThrowMakes'
##  , hpgs.Blocks as 'PrevHomeBlocks'
##  , hpgs.ThreePointMakes as 'PrevHomeThreePointMakes'
##  , apgs.Score as 'PrevAwayScore'
##  , apgs.Assists as 'PrevAwayAssists'
##  , apgs.Rebounds as 'PrevAwayRebounds'
##  , apgs.Steals as 'PrevAwaySteals'
##  , apgs.Turnovers as 'PrevAwayTurnovers'
##  , apgs.Fouls as 'PrevAwayFouls'
##  , apgs.FreeThrowAtts as 'PrevAwayFreeThrowAtts'
##  , apgs.FreeThrowMakes as 'PrevAwayFreeThrowMakes'
##  , apgs.Blocks as 'PrevAwayBlocks'
##  , apgs.ThreePointMakes as 'PrevAwayThreePointMakes'
##  , hpgs.OppScore as 'PrevOppHomeScore'
##  , hpgs.OppAssists as 'PrevOppHomeAssists'
##  , hpgs.OppRebounds as 'PrevOppHomeRebounds'
##  , hpgs.OppSteals as 'PrevOppHomeSteals'
##  , hpgs.OppTurnovers as 'PrevOppHomeTurnovers'
##  , hpgs.OppFouls as 'PrevOppHomeFouls'
##  , hpgs.OppFreeThrowAtts as 'PrevOppHomeFreeThrowAtts'
##  , hpgs.OppFreeThrowMakes as 'PrevOppHomeFreeThrowMakes'
##  , hpgs.OppBlocks as 'PrevOppHomeBlocks'
##  , hpgs.OppThreePointMakes as 'PrevOppHomeThreePointMakes'
##  , apgs.OppScore as 'PrevOppAwayScore'
##  , apgs.OppAssists as 'PrevOppAwayAssists'
##  , apgs.OppRebounds as 'PrevOppAwayRebounds'
##  , apgs.OppSteals as 'PrevOppAwaySteals'
##  , apgs.OppTurnovers as 'PrevOppAwayTurnovers'
##  , apgs.OppFouls as 'PrevOppAwayFouls'
##  , apgs.OppFreeThrowAtts as 'PrevOppAwayFreeThrowAtts'
##  , apgs.OppFreeThrowMakes as 'PrevOppAwayFreeThrowMakes'
##  , apgs.OppBlocks as 'PrevOppAwayBlocks'
##  , apgs.OppThreePointMakes as 'PrevOppAwayThreePointMakes'
## FROM GameSchedules as sch
## INNER JOIN RegulationScores_CTE as sc ON sch.GameID = sc.GameID
## INNER JOIN Stadiums as stad ON sch.StadiumID = stad.StadiumID
## INNER JOIN StadiumSeasons as stadSea ON stad.StadiumID = stadSea.StadiumID
##  AND sch.Season = stadSea.Season
##
## INNER JOIN TeamSeasons as ats ON sch.Season = ats.Season
##  AND sch.AwayTeamID = ats.TeamID
## INNER JOIN TeamPrevGame_CTE as aPrevGmLookup ON sch.GameID = aPrevGmLookup.CurrentGameID
##  AND sch.AwayTeamID = aPrevGmLookup.TeamID
## INNER JOIN GameSchedules as aPrevGm ON aPrevGmLookup.PrevGameID = aPrevGm.GameID
## INNER JOIN Stadiums as aPrevGmStad ON aPrevGm.StadiumID = aPrevGmStad.StadiumID
## INNER JOIN HomeStadiums_CTE as aHmStadLookup ON sch.AwayTeamID = aHmStadLookup.TeamID
##  AND sch.Season = aHmStadLookup.Season
## INNER JOIN TeamBoxscoreStats as aBox ON ats.TeamID = aBox.TeamID AND sch.GameID = aBox.GameID
## INNER JOIN PrevGameStats_CTE as apgs ON sch.GameID = apgs.GameID
##  AND sch.AwayTeamID = apgs.TeamID
##
## INNER JOIN TeamSeasons as hts ON sch.Season = hts.Season
```

```
##  AND sch.HomeTeamID = hts.TeamID
## INNER JOIN TeamPrevGame_CTE as hPrevGmLookup ON sch.GameID = hPrevGmLookup.CurrentGameID
##  AND sch.HomeTeamID = hPrevGmLookup.TeamID
## INNER JOIN GameSchedules as hPrevGm ON hPrevGmLookup.PrevGameID = hPrevGm.GameID
## INNER JOIN Stadiums as hPrevGmStad ON hPrevGm.StadiumID = hPrevGmStad.StadiumID
## INNER JOIN HomeStadiums_CTE as hHmStadLookup ON sch.HomeTeamID = hHmStadLookup.TeamID
##  AND sch.Season = hHmStadLookup.Season
## INNER JOIN TeamBoxscoreStats as hBox ON hts.TeamID = hBox.TeamID AND sch.GameID = hBox.GameID
## INNER JOIN PrevGameStats_CTE as hpgs ON sch.GameID = hpgs.GameID
##  AND sch.HomeTeamID = hpgs.TeamID
##
## WHERE sch.GameTypeID IN (@playoffsGameTypeID, @regSeasonGameTypeID)
##  AND sch.StatusID = @finalGameStatusID
##  AND sch.Season >= 2005
##  AND sch.Season < 2017
##  AND hPrevGmLookup.CurGameNumber > 1
##  AND aPrevGmLookup.CurGameNumber > 1
```

NbaArenaDistance.sql - Query for NBA arena distance analysis

```
## USE FoxSports_SportsData_NBA
## GO
##
## DECLARE @regSeasonGameTypeID INT = 1;
## DECLARE @playoffsGameTypeID INT = 2;
## DECLARE @preseasonGameTypeID INT = 3;
## DECLARE @finalGameStatusID INT = 2;
## DECLARE @forfeitByAwayStatusID INT = 4;
## DECLARE @forfeitByHomeStatusID INT = 5;
## DECLARE @forfeitByBothStatusID INT = 7;
## DECLARE @srid INT = 4326; -- SRID Geography TYPE. DEFINED BY EXTERNAL STANDARD
## DECLARE @metersToMiles REAL = (0.62137119/1000);
## DECLARE @pastGameCount INT = 4;
## DECLARE @minSeason INT = 2005;
## DECLARE @maxSeason INT = 2016;
##
## WITH HomeStadiums_CTE AS (
##  SELECT x.Season, x.TeamID, x.StadiumID, s.Latitude, s.Longitude
##  FROM (
##      SELECT t.Season
##          , t.TeamID
##          , t.StadiumID
##          , ROW_NUMBER() OVER (PARTITION BY t.Season, t.TeamID ORDER BY GameCount desc)
##              as 'StadiumRank'
##      FROM (
##          SELECT ts.Season, ts.TeamID, sch.StadiumID, COUNT(*) as 'GameCount'
##          FROM TeamSeasons as ts
##          INNER JOIN GameSchedules as sch ON ts.Season = sch.Season
##              AND ts.TeamID = sch.HomeTeamID
##          WHERE sch.GameTypeID IN (@playoffsGameTypeID, @regSeasonGameTypeID)
##              AND sch.Season >= @minSeason
##              AND sch.Season <= @maxSeason
##          GROUP BY ts.Season, ts.TeamID, sch.StadiumID
##      ) as t
##  ) as x
##  INNER JOIN Stadiums as s ON x.StadiumID = s.StadiumID
##  WHERE x.StadiumRank = 1
## ),
## ValidStadiums_CTE AS (
##  SELECT ts.Alias as 'TeamAlias'
##      , ts.TeamID
##      , CASE WHEN LEFT(ss.[Name], LEN('Nassau')) = 'Nassau'
##          THEN 'Nassau Coliseum'
##          ELSE ss.[Name]
##          END as 'StadiumName'
##      , st.Season
##      , st.Longitude
##      , st.Latitude
##  FROM (
##      SELECT hst.TeamID
##          , hst.StadiumID
##          , hst.Longitude
```

```
##          , hst.Latitude
##          , hst.Season
##          , ROW_NUMBER() OVER (PARTITION BY hst.TeamID, hst.StadiumID ORDER BY Season desc)
##              as 'StadiumOrder'
##      FROM HomeStadiums_CTE as hst
##  ) as st
##  INNER JOIN TeamSeasons as ts ON st.Season = ts.Season AND st.TeamID = ts.TeamID
##  INNER JOIN StadiumSeasons as ss ON st.Season = ss.Season AND st.StadiumID = ss.StadiumID
##  WHERE st.StadiumOrder = 1 AND st.Season = @maxSeason
## )
##
## SELECT a.StadiumName + ' (' + a.TeamAlias + ')' as 'A'
##  , b.StadiumName + ' (' + b.TeamAlias + ')' as 'B'
##  , CAST(ABS(
##          GEOGRAPHY::Point(a.Latitude, a.Longitude, @srid)
##          .STDistance(GEOGRAPHY::Point(b.Latitude, b.Longitude, @srid)
##      )*@metersToMiles) AS INT) as 'StadDistMiles'
## FROM ValidStadiums_CTE as a
## CROSS JOIN ValidStadiums_CTE as b
```