Home Work Assignment - 01

Critical Thinking Group 5

Contents


```
library(ggplot2)
library(MASS)
library(knitr)
library(xtable)
library(dplyr)
library(psych)
library(stringr)

moneyballvars <- read.csv("https://raw.githubusercontent.com/kishkp/data621-ctg5/master/HW1/moneyballva

moneyballvars <- moneyballvars[moneyballvars[,1]!="INDEX",]

moneyball<- read.csv("https://raw.githubusercontent.com/kishkp/data621-ctg5/master/HW1/moneyball-traini

moneyball2<- select(moneyball, -(INDEX))
```

Data Exploration

In this section we will explore how the data looks like. The goal of data exploration is to look at summaries / descriptives for each variable, shape of the distribution, identify variables of interest, decide on how to treat missing values and outliers.

First lets look at the variables.

```
moneyballvars <- read.csv("https://raw.githubusercontent.com/kishkp/data621-ctg5/master/HW1/moneyballva

kable(moneyballvars)
```

VARIABLE_NAME	DEFINITION	THEORETICAL_EFFECT
INDEX	Identification Variable (do not use)	None
TARGET_WINS	Number of wins	Target
TEAM_BATTING_H	Base Hits by batters (1B,2B,3B,HR)	Positive Impact on Wins
TEAM_BATTING_2B	Doubles by batters (2B)	Positive Impact on Wins
TEAM_BATTING_3B	Triples by batters (3B)	Positive Impact on Wins
TEAM_BATTING_HR	Homeruns by batters (4B)	Positive Impact on Wins
TEAM_BATTING_BB	Walks by batters	Positive Impact on Wins
TEAM_BATTING_HBP	Batters hit by pitch (get a free base)	Positive Impact on Wins
TEAM_BATTING_SO	Strikeouts by batters	Negative Impact on Wins
TEAM_BASERUN_SB	Stolen bases	Positive Impact on Wins
TEAM_BASERUN_CS	Caught stealing	Negative Impact on Wins
TEAM_FIELDING_E	Errors	Negative Impact on Wins
TEAM_FIELDING_DP	Double Plays	Positive Impact on Wins
TEAM_PITCHING_BB	Walks allowed	Negative Impact on Wins
TEAM_PITCHING_H	Hits allowed	Negative Impact on Wins
TEAM_PITCHING_HR	Homeruns allowed	Negative Impact on Wins
TEAM_PITCHING_SO	Strikeouts by pitchers	Positive Impact on Wins

We notice that all variables are numeric. The variable names seem to follow certain naming pattern to highlight certain arithmetic relationships. In other words, we can compute the number of '1B' hits by taking the difference between overall hits and '2B', '3B', 'HR'. Although such naming and construct is not recommended in normalized database design (as it violates third normal form), it is very frequent practice in the data analytics.

Our predictor input is made of 15 variables. And our dependent variable is one variable called TARGET_WINS.

Below are the variable that have been identified and their respective type and category:

Next we start with a summary of the variables and see what we can infer from the same. The goal is to look at measures of central tendancy and dispersion to see how the variables are currently placed in their structure.

Summary / Descriptives / Correlation

```
ds_stats <- psych::describe(moneyball2, skew = FALSE, na.rm = TRUE)[c(3:6)]
ds_stats <- cbind(VARIABLE_NAME = rownames(ds_stats), ds_stats)
#rownames(ds_stats) <- NULL

Variable<- rownames(ds_stats)

fun <- function(x) sum(!complete.cases(x))
Missing <- sapply(moneyball2[Variable], FUN = fun)

#ds_stats <- cbind(ds_stats, Missing)


# fun <- function(x) mean(x, na.rm=T)
# Mean <- sapply(moneyball2[Variable], FUN = fun)
```

```
fun <- function(x, y) cor(y, x, use = "na.or.complete")
Correlation <- sapply(moneyball2[Variable], FUN = fun, y=moneyball2$TARGET_WINS)

ds_stats <- data.frame(cbind(ds_stats, Missing, Correlation))
ds_stats <- left_join(ds_stats, moneyballvars, by="VARIABLE_NAME")
kable(ds_stats)
```

VARIABLE_NAME	mean	sd	median	trimmed	Missing	Correlation	DEFINITION
TARGET_WINS	80.79086	15.75215	82.0	81.31229	0	1.0000000	Number of wins
TEAM_BATTING_H	1469.26977	144.59120	1454.0	1459.04116	0	0.3887675	Base Hits by batter
TEAM_BATTING_2B	241.24692	46.80141	238.0	240.39627	0	0.2891036	Doubles by batters
TEAM_BATTING_3B	55.25000	27.93856	47.0	52.17563	0	0.1426084	Triples by batters (
TEAM_BATTING_HR	99.61204	60.54687	102.0	97.38529	0	0.1761532	Homeruns by batte
TEAM_BATTING_BB	501.55888	122.67086	512.0	512.18331	0	0.2325599	Walks by batters
TEAM_BATTING_SO	735.60534	248.52642	750.0	742.31322	102	-0.0317507	Strikeouts by batte
TEAM_BASERUN_SB	124.76177	87.79117	101.0	110.81188	131	0.1351389	Stolen bases
TEAM_BASERUN_CS	52.80386	22.95634	49.0	50.35963	772	0.0224041	Caught stealing
TEAM_BATTING_HBP	59.35602	12.96712	58.0	58.86275	2085	0.0735042	Batters hit by pitch
TEAM_PITCHING_H	1779.21046	1406.84293	1518.0	1555.89517	0	-0.1099371	Hits allowed
TEAM_PITCHING_HR	105.69859	61.29875	107.0	103.15697	0	0.1890137	Homeruns allowed
TEAM_PITCHING_BB	553.00791	166.35736	536.5	542.62459	0	0.1241745	Walks allowed
TEAM_PITCHING_SO	817.73045	553.08503	813.5	796.93391	102	-0.0784361	Strikeouts by pitche
TEAM_FIELDING_E	246.48067	227.77097	159.0	193.43798	0	-0.1764848	Errors
TEAM_FIELDING_DP	146.38794	26.22639	149.0	147.57789	286	-0.0348506	Double Plays

Based on the table for the variables listed above, there are some things that stand out:

1. Some of the variables like TEAM_PITCHING_H, TEAM_PITCHING_SO and TEAM_FIELDING_E seem to have outliers which is evident from the mean, median and trimmed mean values.

2. TEAM_BATTING_HBP and TEAM_BASERUN_CS seems to be missing a lot of values which casts doubt on its usefulness as a predictor. Maybe a flag for presense or absense of TEAM_BATTING_HBP and TEAM_BASERUN_CS might be a better predictor. Also given the fact that there is low correlation, we decided to exclude these 2 variables from any missing value or outlier treatment.

3. Most of the variables seem to indicate a positive / negative correlation in line with the theoretical effect. However, the following stand out as they show a correlation opposite to the theoretical impact: TEAM_BASERUN_CS, TEAM_PITCHING_HR, TEAM_PITCHING_BB, TEAM_PITCHING_SO and TEAM_FIELDING_DP. Lets evaluate these variables further once we fix any missing values or outliers.

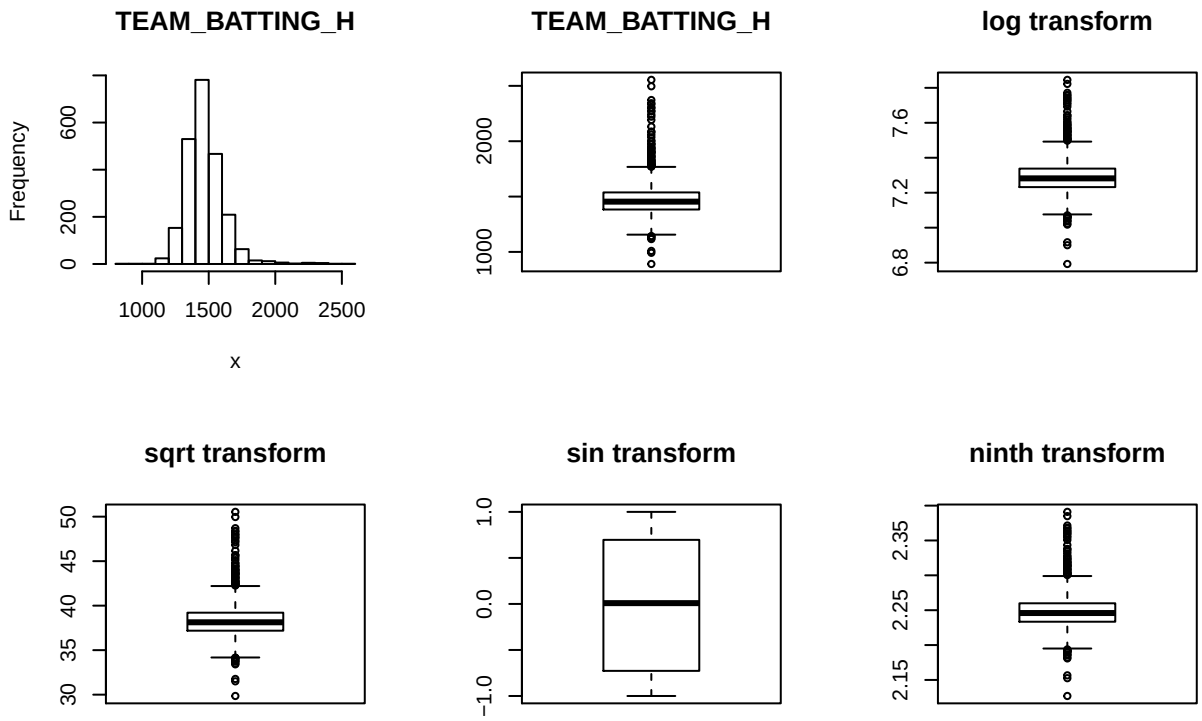
4. We will impute the missing values in TEAM_BATTING_SO, FIELDING_DP, BASERUN_SB and TEAM_PITCHING_SO since it has lesser missing values even though there is low correlation. So we will create new variables that will have the respective missing values handled.

Distribution and Correlation

In this section we look at boxplots to determine the outliers in variables and decide on whether to act on the outliers.

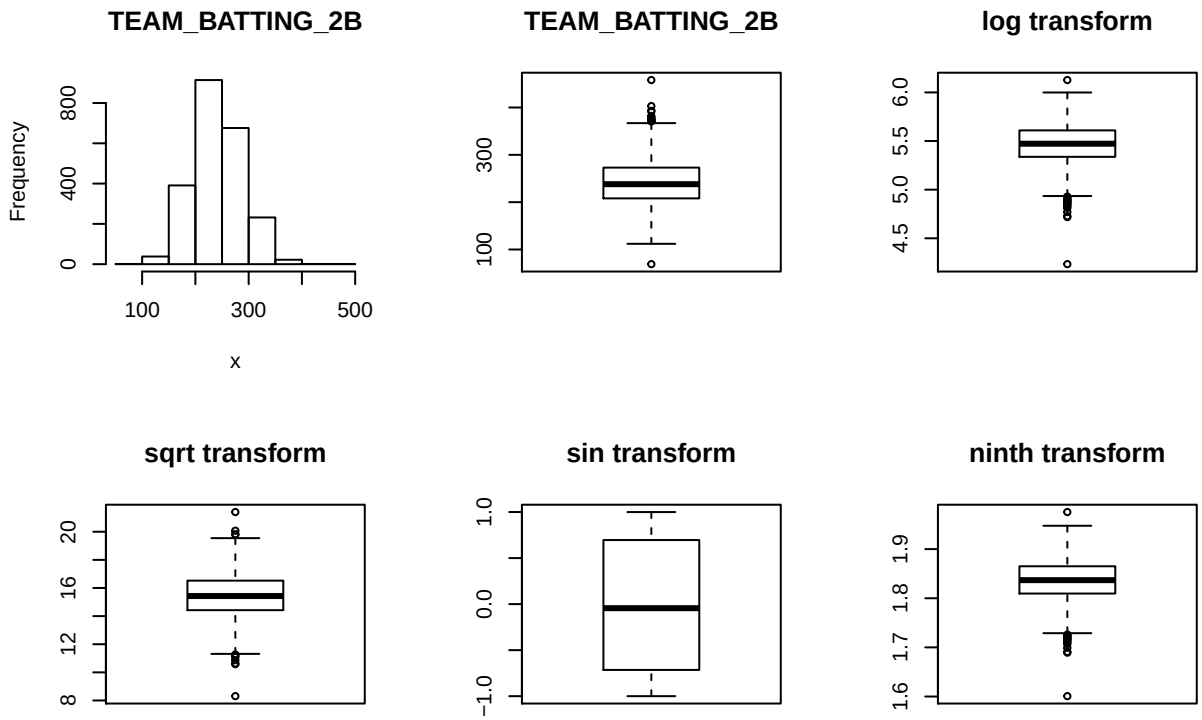
Lets do some univariate analysis. We will look at the Histogram and Boxplot for each variable to detect outliers if any and treat it accordingly.

```
show_charts <- function(x, ...) {

    par(mfrow=c(2,3))

    xlabel <- unlist(str_split(deparse(substitute(x)), pattern = "\\$"))[2]
#    ylabel <- unlist(str_split(deparse(substitute(y)), pattern = "\\$"))[2]

    hist(x,main=xlabel)
    boxplot(x,main=xlabel)

    y<-log(x)
    boxplot(y,main='log transform')
    y<-sqrt(x)
    boxplot(y,main='sqrt transform')
    y<-sin(x)
    boxplot(y,main='sin transform')
    y<-(x)^(1/9)
    boxplot(y,main='ninth transform')
}

#show_charts(moneyball2$TEAM_BATTING_H,moneyball2$TARGET_WINS)

show_charts(moneyball2$TEAM_BATTING_H)
```



For TEAM_BATTING_H, we can see that there are quite a few outliers, both at the upper and lower end. Accordingly, we decide to create a new variable that will have the outlier fixed.

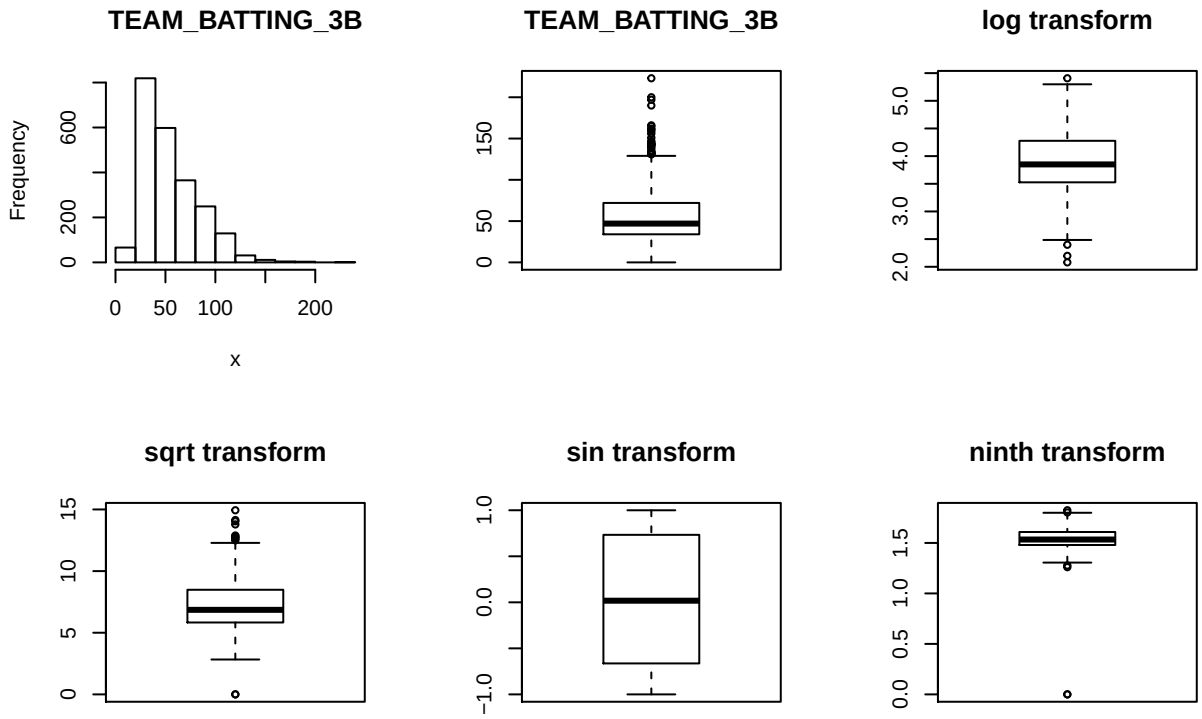
```
show_charts(moneyball2$TEAM_BATTING_2B)
```



For TEAM_BATTING_2B, we can see that there are quite a few outliers, both at the upper and a single outlier at the lower end. For this variable we decide to create a new variable that will have the outliers fixed.

```
show_charts(moneyball2$TEAM_BATTING_3B)
```

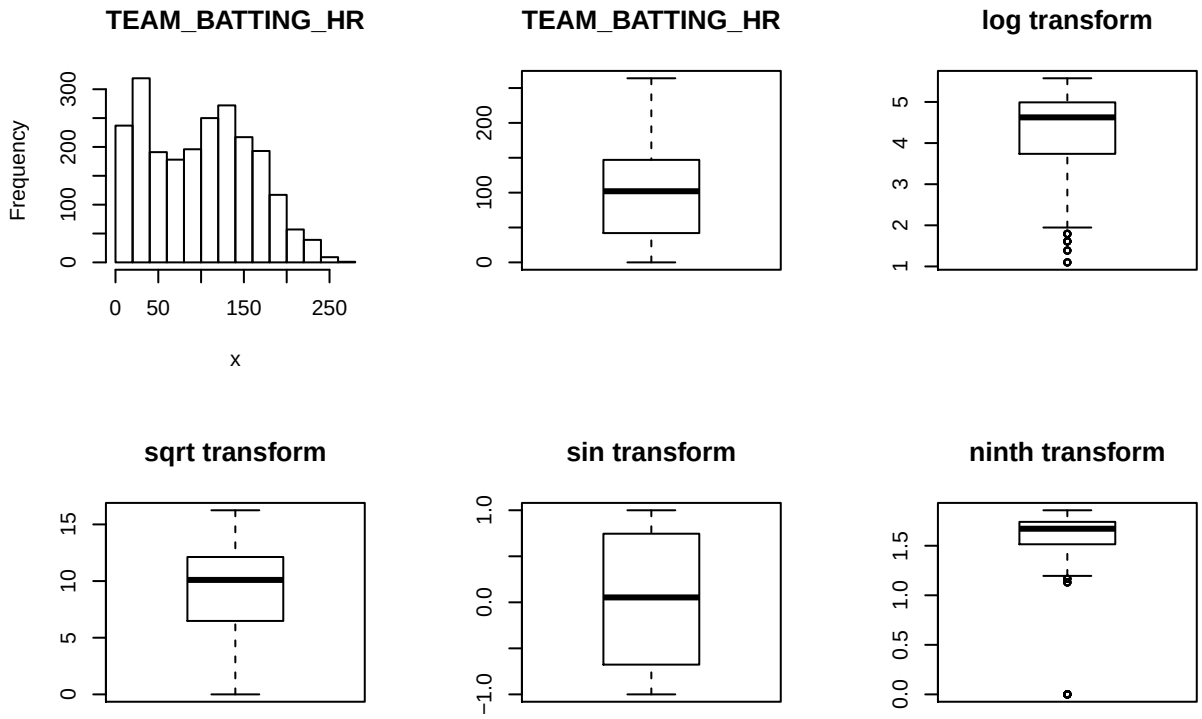
```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



For TEAM_BATTING_3B, we can see that there are quite a few outliers at the upper end. For this variable we decide to create a new variable that will have the outliers fixed.

```
show_charts(moneyball2$TEAM_BATTING_HR)
```

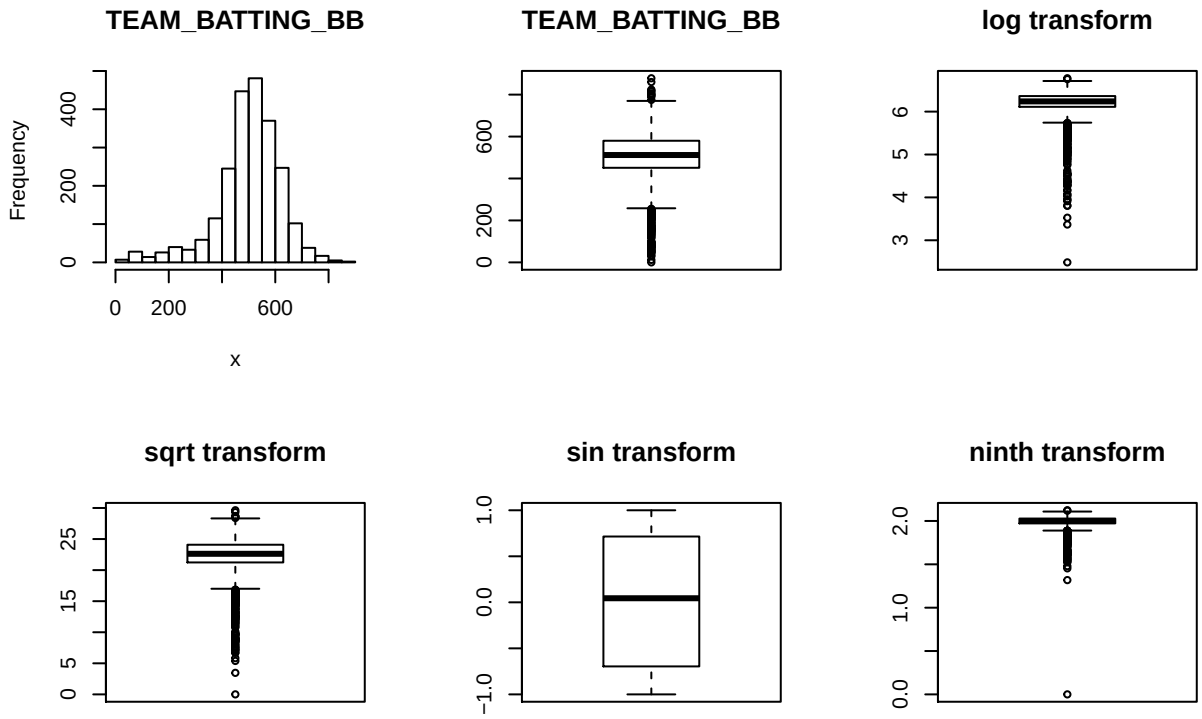
```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



For TEAM_BATTING_HR, we can see that there are no outliers.

```
show_charts(moneyball2$TEAM_BATTING_BB)
```

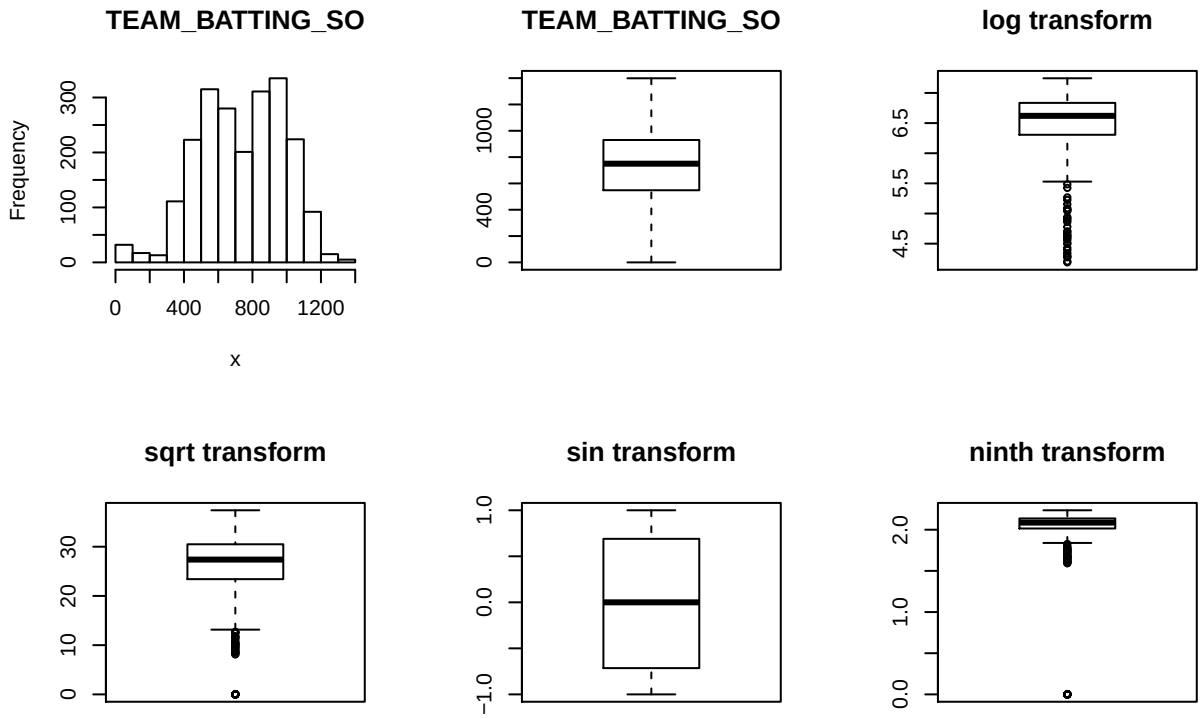
```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



For TEAM_BATTING_BB, we can see that there are quite a few outliers, both at the upper and lower end. For this variable we decide to create a new variable that will have the outlier fixed.

```
show_charts(moneyball2$TEAM_BATTING_SO)
```

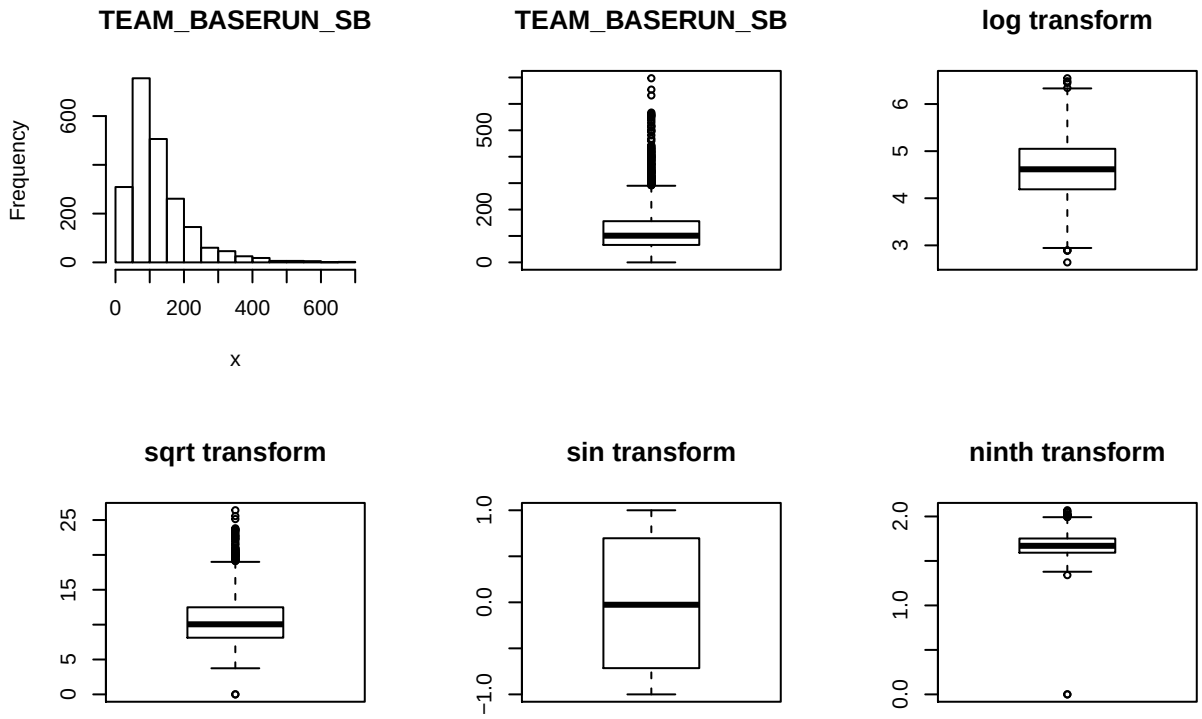
```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



For TEAM_BATTING_SO, we can see that there are no outliers. No further action needed for this variable.

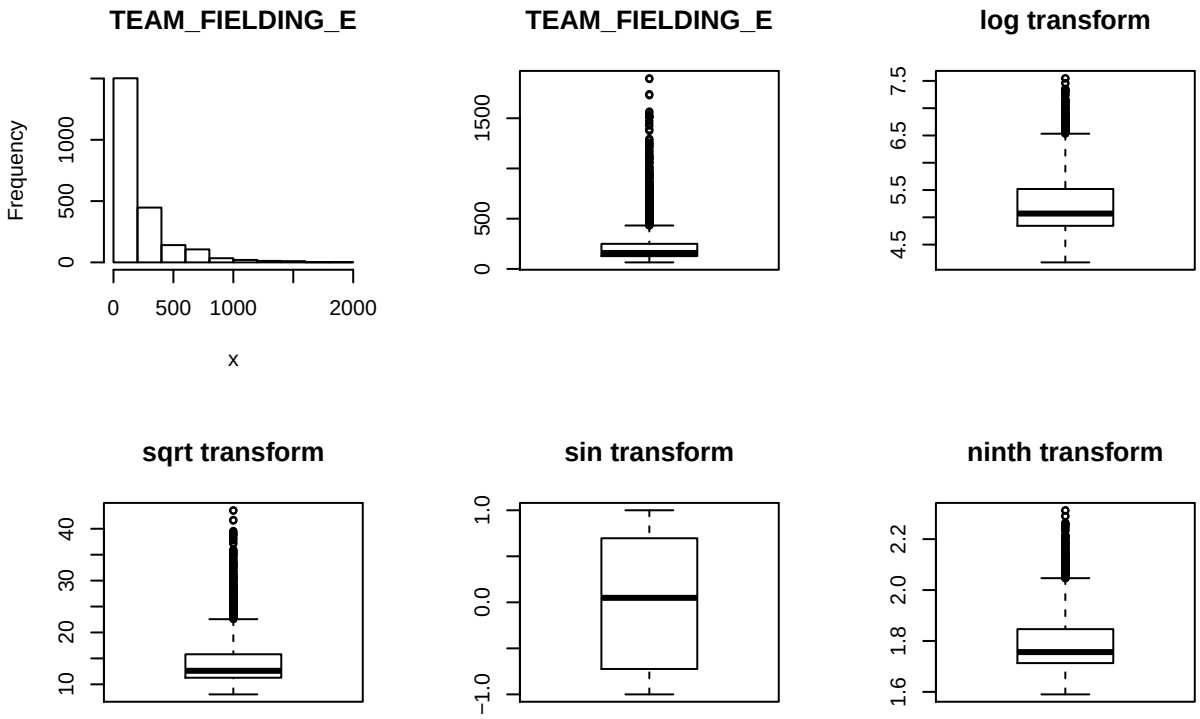
```
show_charts(moneyball2$TEAM_BASERUN_SB)
```

```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



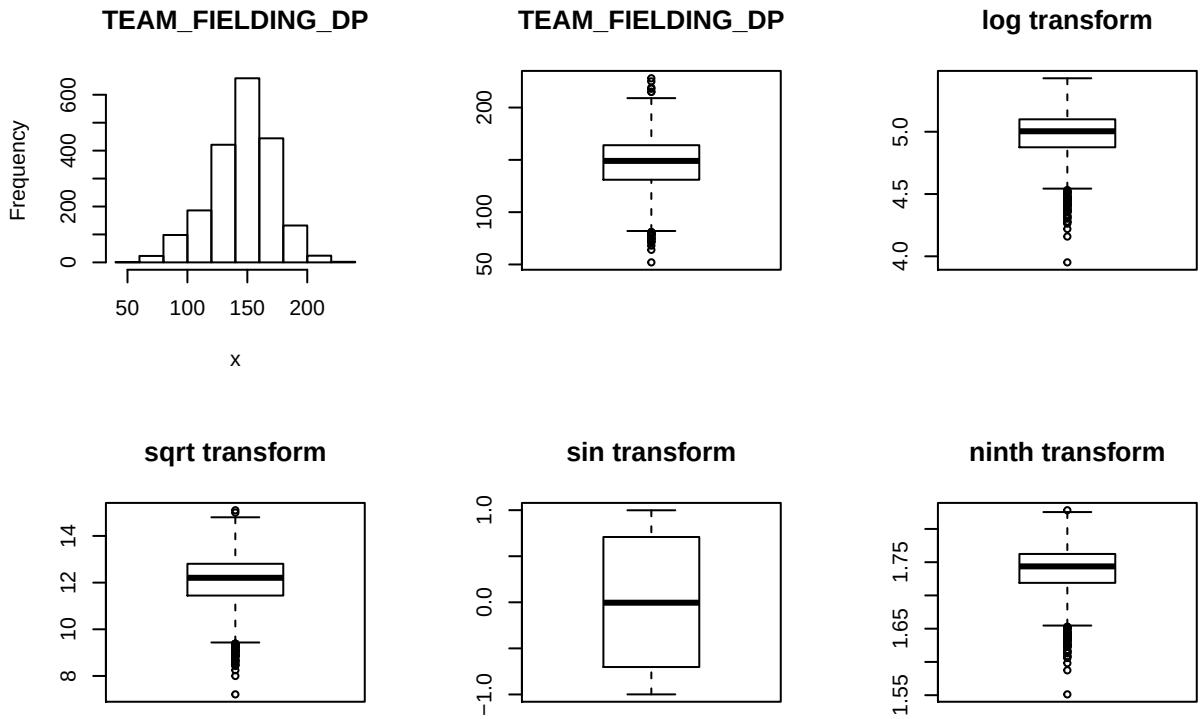
For TEAM_BASERUN_SB, we can see that there are quite a few outliers at the upper end. For this variable we decide to create a new variable that will have the outlier fixed.

```
show_charts(moneyball2$TEAM_FIELDING_E)
```



For TEAM_FIELDING_E, we can see that there are quite a few outliers at the upper end. For this variable we decide to create a new variable that will have the outlier fixed.

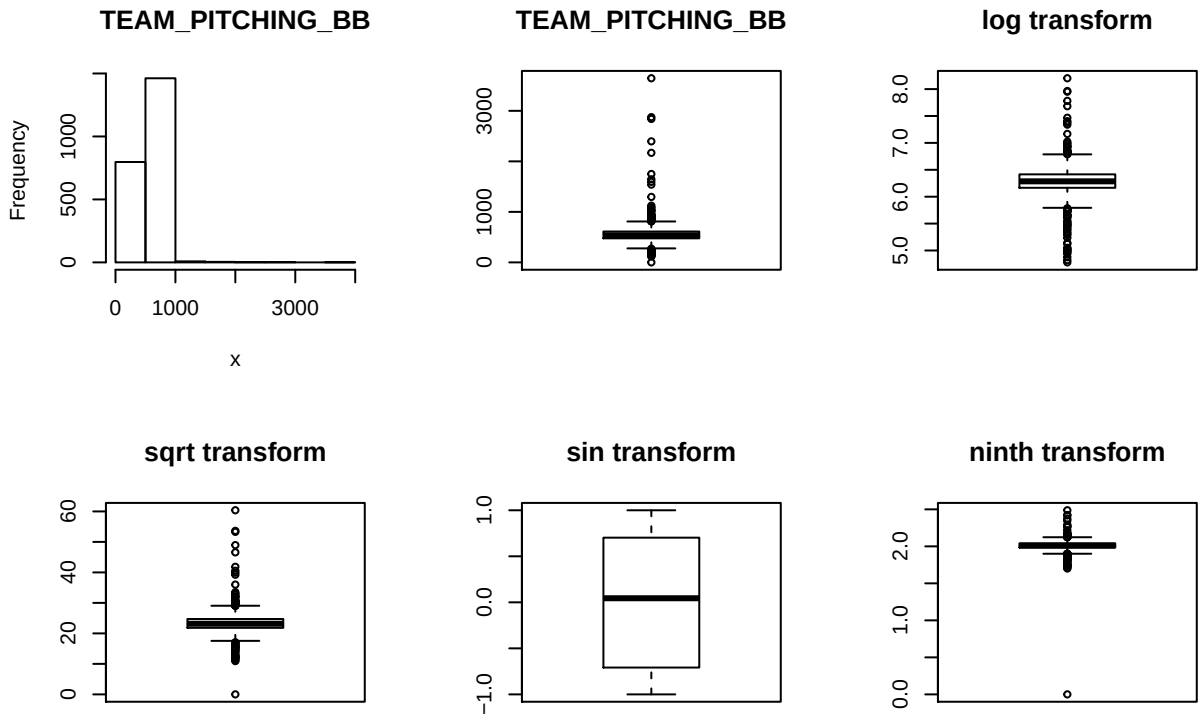
```
show_charts(moneyball2$TEAM_FIELDING_DP)
```



For TEAM_FIELDING_DP, we can see that there are quite a few outliers, both at the upper and lower end. For this variable we decide to create a new variable that will have the outlier fixed.

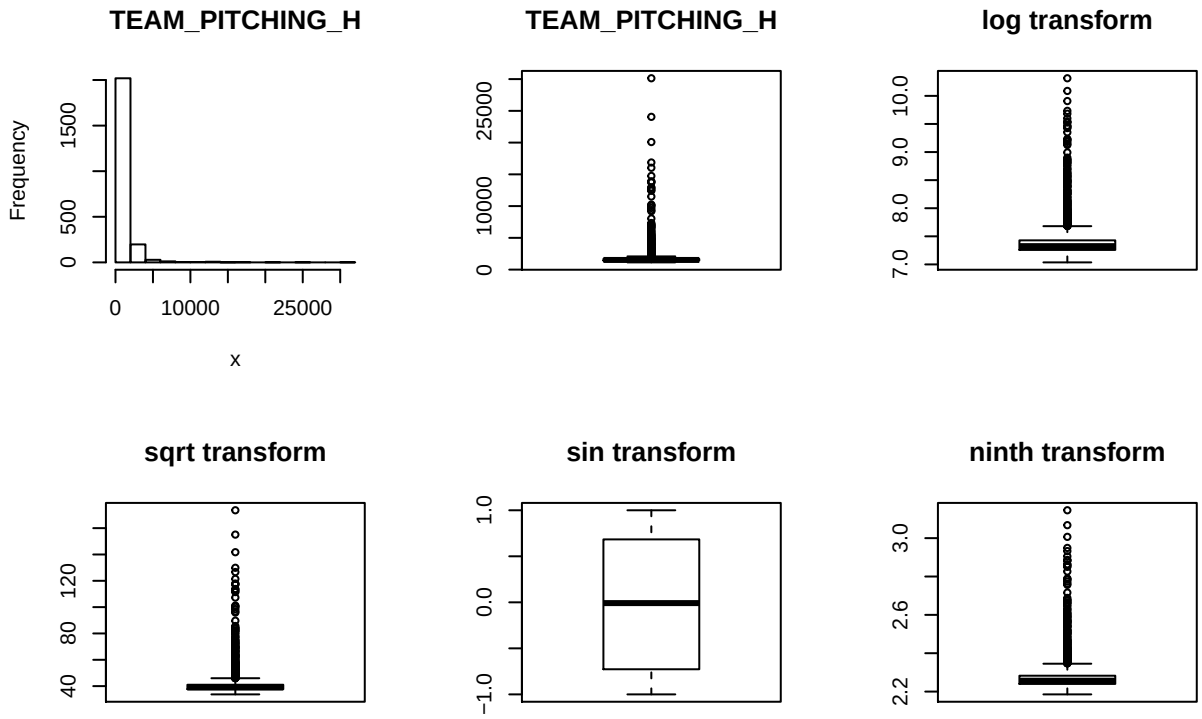
```
show_charts(moneyball2$TEAM_PITCHING_BB)
```

```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



For TEAM_PITCHING_BB, we can see that there are quite a few outliers, both at the upper and lower end. For this variable we decide to create a new variable that will have the outlier fixed.

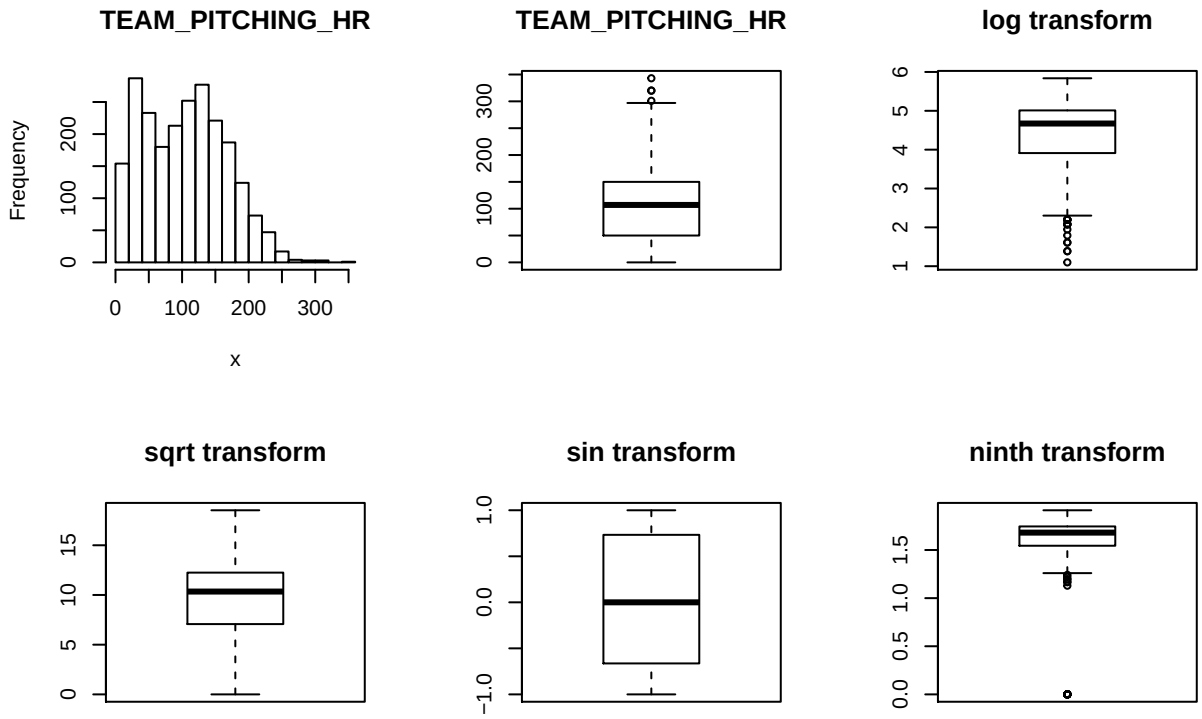
```
show_charts(moneyball2$TEAM_PITCHING_H)
```



For TEAM_PITCHING_H, we can see that there are quite a few outliers at the upper end. For this variable we decide to create a new variable that will have the outlier fixed.

```
show_charts(moneyball2$TEAM_PITCHING_HR)
```

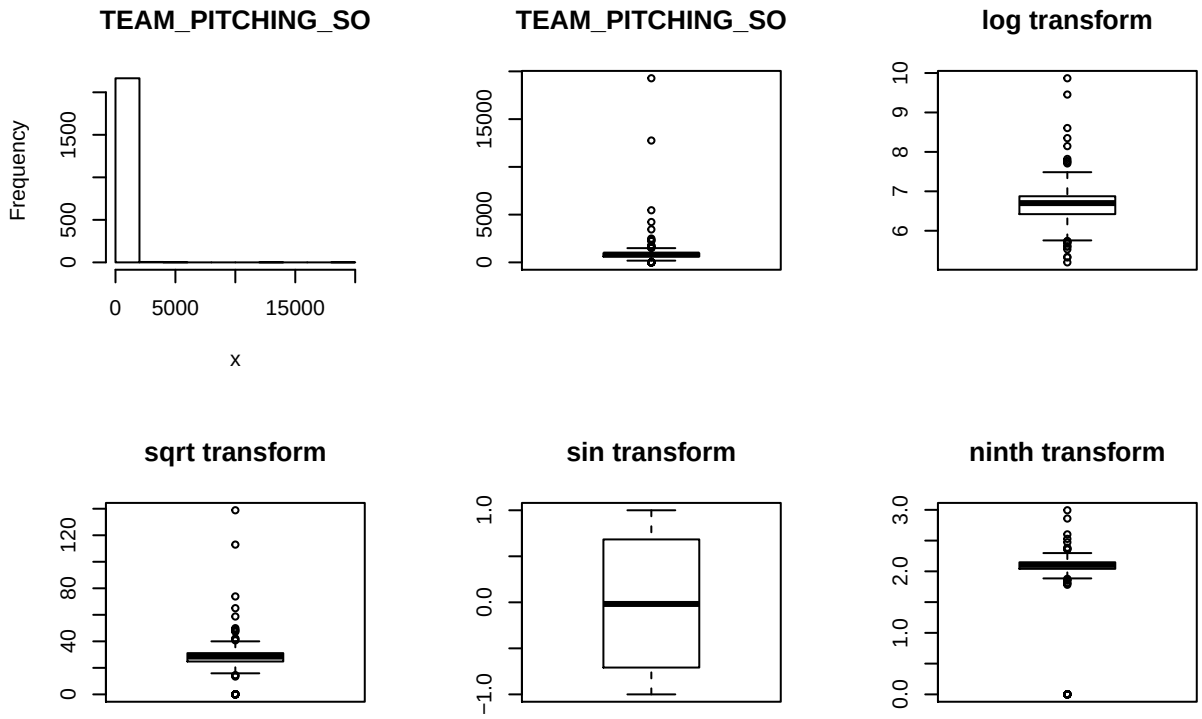
```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



For TEAM_PITCHING_HR, we can see that there only 3 outliers at the upper end. For this variable we decide to create a new variable that will have the outlier fixed.

```
show_charts(moneyball2$TEAM_PITCHING_SO)
```

```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



For TEAM_PITCHING_SO, we can see that there are quite a few outliers at the upper and a single outlier on the lower end. For this variable we decide to create a new variable that will have the outlier fixed.

** In most of the cases above, we see that a SIN transformation seems to work well to take care of the outliers. We will go ahead and create these new variables respectively.**

Data Preparation

Now that we have the preliminary analysis ready, we will go ahead and carry out the necessary transformations to the data.

This will primarily take care of Missing Values, Handle Outliers and create some additional variables.

Outliers

For outliers, we will create 2 sets of variables.

The first set uses the capping method. In this method, we will replace all outliers that lie outside the 1.5 times of IQR limits. We will cap it by replacing those observations less than the lower limit with the value of 5th %ile and those that lie above the upper limit with the value of 95th %ile.

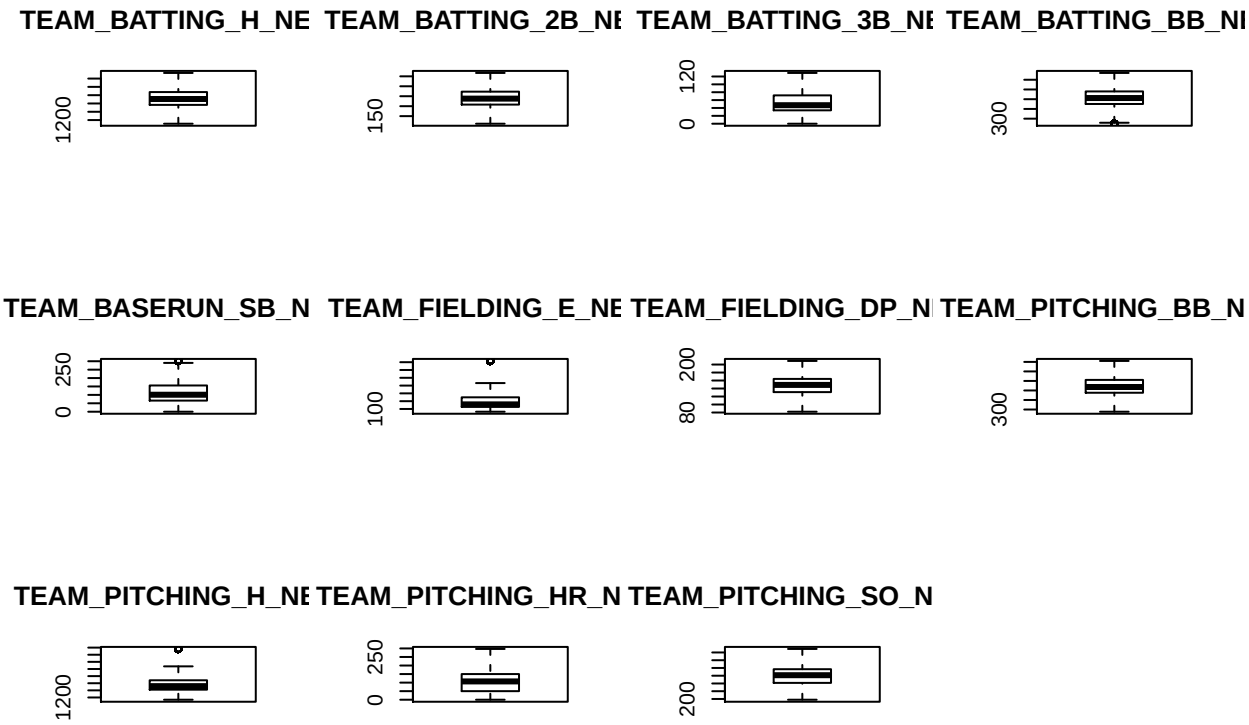
Accordingly we create the following new variables while retaining the original variables as is.

TEAM_BATTING_H_NEW TEAM_BATTING_2B_NEW TEAM_BATTING_3B_NEW TEAM_BATTING_BB_NEW
TEAM_BASERUN_SB_NEW TEAM_FIELDING_E_NEW TEAM_FIELDING_DP_NEW
TEAM_PITCHING_BB_NEW TEAM_PITCHING_H_NEW TEAM_PITCHING_HR_NEW
TEAM_PITCHING_SO_NEW

```
# function for removing outliers - http://r-statistics.co/Outlier-Treatment-With-R.html

treat_outliers <- function(x) {
qnt <- quantile(x, probs=c(.25, .75), na.rm = T)
caps <- quantile(x, probs=c(.05, .95), na.rm = T)
H <- 1.5 * IQR(x, na.rm = T)
x[x < (qnt[1] - H)] <- caps[1]
x[x > (qnt[2] + H)] <- caps[2]

return(x)
}

moneyball2$TEAM_BATTING_H_NEW <- treat_outliers(moneyball2$TEAM_BATTING_H)
moneyball2$TEAM_BATTING_2B_NEW <- treat_outliers(moneyball2$TEAM_BATTING_2B)
moneyball2$TEAM_BATTING_3B_NEW <- treat_outliers(moneyball2$TEAM_BATTING_3B)
moneyball2$TEAM_BATTING_BB_NEW <- treat_outliers(moneyball2$TEAM_BATTING_BB)
moneyball2$TEAM_BASERUN_SB_NEW <- treat_outliers(moneyball2$TEAM_BASERUN_SB)
moneyball2$TEAM_FIELDING_E_NEW <- treat_outliers(moneyball2$TEAM_FIELDING_E)
moneyball2$TEAM_FIELDING_DP_NEW <- treat_outliers(moneyball2$TEAM_FIELDING_DP)
moneyball2$TEAM_PITCHING_BB_NEW <- treat_outliers(moneyball2$TEAM_PITCHING_BB)
moneyball2$TEAM_PITCHING_H_NEW <- treat_outliers(moneyball2$TEAM_PITCHING_H)
moneyball2$TEAM_PITCHING_HR_NEW <- treat_outliers(moneyball2$TEAM_PITCHING_HR)
moneyball2$TEAM_PITCHING_SO_NEW <- treat_outliers(moneyball2$TEAM_PITCHING_SO)
```

Lets see how the new variables look in boxplots.

```
par(mfrow=c(3,4))

boxplot(moneyball2$TEAM_BATTING_H_NEW,main="TEAM_BATTING_H_NEW")
boxplot(moneyball2$TEAM_BATTING_2B_NEW,main="TEAM_BATTING_2B_NEW")
boxplot(moneyball2$TEAM_BATTING_3B_NEW,main="TEAM_BATTING_3B_NEW")
boxplot(moneyball2$TEAM_BATTING_BB_NEW,main="TEAM_BATTING_BB_NEW")
boxplot(moneyball2$TEAM_BASERUN_SB_NEW,main="TEAM_BASERUN_SB_NEW")
boxplot(moneyball2$TEAM_FIELDING_E_NEW,main="TEAM_FIELDING_E_NEW")
boxplot(moneyball2$TEAM_FIELDING_DP_NEW,main="TEAM_FIELDING_DP_NEW")
boxplot(moneyball2$TEAM_PITCHING_BB_NEW,main="TEAM_PITCHING_BB_NEW")
boxplot(moneyball2$TEAM_PITCHING_H_NEW,main="TEAM_PITCHING_H_NEW")
boxplot(moneyball2$TEAM_PITCHING_HR_NEW,main="TEAM_PITCHING_HR_NEW")
boxplot(moneyball2$TEAM_PITCHING_SO_NEW,main="TEAM_PITCHING_SO_NEW")
```



In the second set, we will use the sin transformation and create the following variables:

TEAM_BATTING_H_SIN TEAM_BATTING_2B_SIN TEAM_BATTING_3B_SIN TEAM_BATTING_BB_SIN TEAM_BASERUN_SB_SIN TEAM_FIELDING_E_SIN TEAM_FIELDING_DP_SIN TEAM_PITCHING_BB_SIN TEAM_PITCHING_H_SIN TEAM_PITCHING_HR_SIN TEAM_PITCHING_SO_SIN

```
moneyball2$TEAM_BATTING_H_SIN <- sin(moneyball2$TEAM_BATTING_H)
moneyball2$TEAM_BATTING_2B_SIN <- sin(moneyball2$TEAM_BATTING_2B)
moneyball2$TEAM_BATTING_3B_SIN <- sin(moneyball2$TEAM_BATTING_3B)
moneyball2$TEAM_BATTING_BB_SIN <- sin(moneyball2$TEAM_BATTING_BB)
moneyball2$TEAM_BASERUN_SB_SIN <- sin(moneyball2$TEAM_BASERUN_SB)
moneyball2$TEAM_FIELDING_E_SIN <- sin(moneyball2$TEAM_FIELDING_E)
moneyball2$TEAM_FIELDING_DP_SIN <- sin(moneyball2$TEAM_FIELDING_DP)
moneyball2$TEAM_PITCHING_BB_SIN <- sin(moneyball2$TEAM_PITCHING_BB)
moneyball2$TEAM_PITCHING_H_SIN <- sin(moneyball2$TEAM_PITCHING_H)
moneyball2$TEAM_PITCHING_HR_SIN <- sin(moneyball2$TEAM_PITCHING_HR)
moneyball2$TEAM_PITCHING_SO_SIN <- sin(moneyball2$TEAM_PITCHING_SO)
```

Missing Values

Next we impute missing values. Since we have handled outliers, we can go ahead and use the mean as impute values. As with outliers, we will go ahead and create new variables for the following:

TEAM_BATTING_SO_NEW

We will re-use the already created new variables for fixing the missing values for the below:

TEAM_PITCHING_SO_NEW TEAM_BASERUN_SB_NEW TEAM_FIELDING_DP_NEW

```
moneyball2$TEAM_BATTING_SO_NEW <- moneyball2$TEAM_BATTING_SO
moneyball2$TEAM_BATTING_SO_NEW[is.na(moneyball2$TEAM_BATTING_SO_NEW)] <- mean(moneyball2$TEAM_BATTING_SO

moneyball2$TEAM_PITCHING_SO_NEW[is.na(moneyball2$TEAM_PITCHING_SO_NEW)] <- mean(moneyball2$TEAM_PITCHING
moneyball2$TEAM_BASERUN_SB_NEW[is.na(moneyball2$TEAM_BASERUN_SB_NEW)] <- mean(moneyball2$TEAM_BASERUN_SB
moneyball2$TEAM_FIELDING_DP_NEW[is.na(moneyball2$TEAM_FIELDING_DP_NEW)] <- mean(moneyball2$TEAM_FIELDING
```

Additional Variables

Lets now create some additional variables that might help us in out analysis.

Missing Flags

First we create flag variables to indicate whether TEAM_BATTING_HBP and TEAM_BASERUN_CS and missing. If the value is missing, we code it with 1 and if the value is present we code it with 0.

```
moneyball2$TEAM_BATTING_HBP_Missing <- ifelse(complete.cases(moneyball2$TEAM_BATTING_HBP),1,0)
moneyball2$TEAM_BASERUN_CS_Missing <- ifelse(complete.cases(moneyball2$TEAM_BASERUN_CS),1,0)
```

Ratios

Next we create some additional variables, that we think may be useful with the prediction. Here we create the following ratios:

```
moneyball2$Hits_R <- moneyball2$TEAM_BATTING_H/moneyball2$TEAM_PITCHING_H
moneyball2$Walks_R <- moneyball2$TEAM_BATTING_BB/moneyball2$TEAM_PITCHING_BB
moneyball2$HomeRuns_R <- moneyball2$TEAM_BATTING_HR/moneyball2$TEAM_PITCHING_HR
moneyball2$Strikeout_R <- moneyball2$TEAM_BATTING_SO/moneyball2$TEAM_PITCHING_SO
```

Calculated Variables

Finally, we create some calculated variables as below:

1. TEAM_BATTING_EB (Extra Base Hits) = 2B + 3B + HR
2. TEAM_BATTING_1B (Singles by batters) = TEAM_BATTING_H - TEAM_BATTING_EB

```
moneyball2$TEAM_BATTING_EB <- moneyball2$TEAM_BATTING_2B + moneyball2$TEAM_BATTING_3B + moneyball2$TEAM_

moneyball2$TEAM_BATTING_1B <- moneyball2$TEAM_BATTING_H - moneyball2$TEAM_BATTING_EB
```

Correlation for new variables

Lets see how the new variables stack up against wins.

```
fun <- function(x, y) cor(y, x, use = "na.or.complete")
Correlation <- sapply(moneyball2[, 40:47], FUN = fun, y=moneyball2$TARGET_WINS)
Correlation
```

```
## TEAM_BATTING_HBP_Missing  TEAM_BASERUN_CS_Missing                  Hits_R
##              0.002610647              0.004864215             0.095800033
##                  Walks_R               HomeRuns_R             Strikeout_R
##              0.083660245              0.013440964             0.063193881
##          TEAM_BATTING_EB          TEAM_BATTING_1B
##              0.344958150              0.217430135
```

All new variables seem to have a positive correlation with wins. However, some of them do not seem to have a strong correlation. Lets see how they perform while modeling.

Build Models

We will now build 4 models that will use different variables in the dataset. The following are the 4 models along with the variables that will be used in the respective model:

```
modelvars <- read.csv("https://raw.githubusercontent.com/kishkp/data621-ctg5/master/HW1/ModelVars.csv")
kable(modelvars)
```

VARIABLE_NAME	Comments	Theoretical.Effect	Model1	Model2	Model3	Model4
TEAM_BATTING_H	Given	Positive	Y			Y
TEAM_BATTING_2B	Given	Positive	Y			Y
TEAM_BATTING_3B	Given	Positive	Y			Y
TEAM_BATTING_HR	Given	Positive	Y			Y
TEAM_BATTING_BB	Given	Positive	Y			Y
TEAM_BATTING_HBP	Given	Positive	Y			
TEAM_BATTING_SO	Given	Negative	Y			Y
TEAM_BASERUN_SB	Given	Positive	Y			Y
TEAM_BASERUN_CS	Given	Negative	Y			
TEAM_FIELDING_E	Given	Negative	Y			Y
TEAM_FIELDING_DP	Given	Positive	Y			Y
TEAM_PITCHING_BB	Given	Negative	Y			Y
TEAM_PITCHING_H	Given	Negative	Y			Y
TEAM_PITCHING_HR	Given	Negative	Y			Y
TEAM_PITCHING_SO	Given	Positive	Y			Y
TEAM_BATTING_H_NEW	Derived	Positive		Y		Y
TEAM_BATTING_2B_NEW	Derived	Positive		Y		Y
TEAM_BATTING_3B_NEW	Derived	Positive		Y		Y
TEAM_BATTING_BB_NEW	Derived	Positive		Y		Y
TEAM_BASERUN_SB_NEW	Derived	Positive		Y		Y
TEAM_FIELDING_E_NEW	Derived	Negative		Y		Y
TEAM_FIELDING_DP_NEW	Derived	Positive		Y		Y
TEAM_PITCHING_BB_NEW	Derived	Negative		Y		Y
TEAM_PITCHING_H_NEW	Derived	Negative		Y		Y
TEAM_PITCHING_HR_NEW	Derived	Negative		Y		Y

VARIABLE_NAME	Comments	Theoretical.Effect	Model1	Model2	Model3	Model4
TEAM_PITCHING_SO_NEW	Derived	Positive		Y		Y
TEAM_BATTING_H_SIN	Derived	Positive			Y	Y
TEAM_BATTING_2B_SIN	Derived	Positive			Y	Y
TEAM_BATTING_3B_SIN	Derived	Positive			Y	Y
TEAM_BATTING_BB_SIN	Derived	Positive			Y	Y
TEAM_BASERUN_SB_SIN	Derived	Positive			Y	Y
TEAM_FIELDING_E_SIN	Derived	Negative			Y	Y
TEAM_FIELDING_DP_SIN	Derived	Positive			Y	Y
TEAM_PITCHING_BB_SIN	Derived	Negative			Y	Y
TEAM_PITCHING_H_SIN	Derived	Negative			Y	Y
TEAM_PITCHING_HR_SIN	Derived	Negative			Y	Y
TEAM_PITCHING_SO_SIN	Derived	Positive			Y	Y
TEAM_BATTING_HBP_Missing	Derived				Y	Y
TEAM_BASERUN_CS_Missing	Derived				Y	Y
Hits_R	Derived				Y	Y
Walks_R	Derived				Y	Y
HomeRuns_R	Derived				Y	Y
Strikeout_R	Derived				Y	Y
TEAM_BATTING_EB	Derived				Y	Y
TEAM_BATTING_1B	Derived				Y	Y

We will now go ahead and work on the 4 models. For each model we will do a stepwise selection and stop at a point where we retain only those variables that have a significant p value.

Model1: This is the first model.

```
model1<-lm(TARGET_WINS~TEAM_BATTING_H+TEAM_BATTING_2B+TEAM_BATTING_3B+TEAM_BATTING_HR+TEAM_BATTING_BB+T

summary(model1)
```

```
##
## Call:
## lm(formula = TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B +
##     TEAM_BATTING_3B + TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_HBP +
##     TEAM_BATTING_SO + TEAM_BASERUN_SB + TEAM_BASERUN_CS + TEAM_FIELDING_E +
##     TEAM_FIELDING_DP + TEAM_PITCHING_BB + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_SO, data = na.omit(moneyball2))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -19.8708  -5.6564  -0.0599   5.2545  22.9274
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      60.28826   19.67842   3.064  0.00253 **
## TEAM_BATTING_H    1.91348    2.76139   0.693  0.48927
## TEAM_BATTING_2B   0.02639    0.03029   0.871  0.38484
## TEAM_BATTING_3B  -0.10118    0.07751  -1.305  0.19348
## TEAM_BATTING_HR  -4.84371   10.50851  -0.461  0.64542
## TEAM_BATTING_BB  -4.45969    3.63624  -1.226  0.22167
```

```
## TEAM_BATTING_HBP  0.08247    0.04960   1.663  0.09815 .
## TEAM_BATTING_SO   0.34196    2.59876   0.132  0.89546
## TEAM_BASERUN_SB   0.03304    0.02867   1.152  0.25071
## TEAM_BASERUN_CS  -0.01104    0.07143  -0.155  0.87730
## TEAM_FIELDING_E  -0.17204    0.04140  -4.155 5.08e-05 ***
## TEAM_FIELDING_DP -0.10819    0.03654  -2.961  0.00349 **
## TEAM_PITCHING_BB  4.51089    3.63372   1.241  0.21612
## TEAM_PITCHING_H  -1.89096    2.76095  -0.685  0.49432
## TEAM_PITCHING_HR  4.93043   10.50664   0.469  0.63946
## TEAM_PITCHING_SO -0.37364    2.59705  -0.144  0.88577
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.467 on 175 degrees of freedom
## Multiple R-squared:  0.5501, Adjusted R-squared:  0.5116
## F-statistic: 14.27 on 15 and 175 DF,  p-value: < 2.2e-16
```

```
# coefficients(model1) # model coefficients
# confint(model1, level=0.95) # CIs for model parameters
# fitted(model1) # predicted values
# residuals(model1) # residuals
# anova(model1) # anova table
# vcov(model1) # covariance matrix for model parameters
# influence(model1) # regression diagnostics
```

Lets now step thru this model and retain only those variables that have the most impact.

```
step1 <- step(model1,direction="backward",test="F")
```

```
## Start:  AIC=831.31
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_HBP + TEAM_BATTING_SO +
##     TEAM_BASERUN_SB + TEAM_BASERUN_CS + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_PITCHING_BB + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_SO    1      1.24 12547 829.33  0.0173  0.895462
## - TEAM_PITCHING_SO   1      1.48 12547 829.33  0.0207  0.885767
## - TEAM_BASERUN_CS    1      1.71 12548 829.34  0.0239  0.877303
## - TEAM_BATTING_HR    1     15.23 12561 829.54  0.2125  0.645420
## - TEAM_PITCHING_HR   1     15.79 12562 829.55  0.2202  0.639462
## - TEAM_PITCHING_H    1     33.63 12580 829.82  0.4691  0.494317
## - TEAM_BATTING_H     1     34.42 12580 829.83  0.4802  0.489267
## - TEAM_BATTING_2B    1     54.41 12600 830.14  0.7590  0.384844
## - TEAM_BASERUN_SB    1     95.22 12641 830.76  1.3281  0.250708
## - TEAM_BATTING_BB    1    107.84 12654 830.95  1.5042  0.221675
## - TEAM_PITCHING_BB   1    110.48 12656 830.99  1.5411  0.216120
## - TEAM_BATTING_3B    1    122.16 12668 831.16  1.7040  0.193477
## <none>                           12546 831.31
## - TEAM_BATTING_HBP   1    198.21 12744 832.31  2.7647  0.098152 .
## - TEAM_FIELDING_DP   1    628.49 13174 838.65  8.7667  0.003494 **
```

```
## - TEAM_FIELDING_E   1   1237.79 13784 847.28 17.2657 5.076e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=829.33
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_HBP + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_PITCHING_BB +
##     TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BASERUN_CS   1      1.59 12549 827.35  0.0223  0.881502
## - TEAM_BATTING_HR   1     15.82 12563 827.57  0.2219  0.638192
## - TEAM_PITCHING_HR  1     16.39 12564 827.58  0.2299  0.632189
## - TEAM_BATTING_2B   1     53.47 12601 828.14  0.7501  0.387638
## - TEAM_PITCHING_H   1     88.45 12636 828.67  1.2408  0.266844
## - TEAM_BATTING_H    1     90.30 12637 828.70  1.2666  0.261934
## - TEAM_BASERUN_SB   1     94.19 12641 828.76  1.3212  0.251935
## - TEAM_BATTING_BB   1    107.95 12655 828.97  1.5142  0.220136
## - TEAM_PITCHING_BB  1    110.60 12658 829.01  1.5514  0.214590
## - TEAM_BATTING_3B   1    122.20 12669 829.18  1.7140  0.192168
## <none>                          12547 829.33
## - TEAM_BATTING_HBP  1    197.11 12744 830.31  2.7649  0.098135 .
## - TEAM_FIELDING_DP  1    630.68 13178 836.70  8.8466  0.003349 **
## - TEAM_FIELDING_E   1   1240.80 13788 845.34 17.4048 4.738e-05 ***
## - TEAM_PITCHING_SO  1   1312.89 13860 846.34 18.4161 2.925e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=827.35
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_HBP + TEAM_BASERUN_SB +
##     TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_PITCHING_BB + TEAM_PITCHING_H +
##     TEAM_PITCHING_HR + TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_HR   1     16.06 12565 825.60  0.2265  0.634713
## - TEAM_PITCHING_HR  1     16.64 12565 825.61  0.2347  0.628689
## - TEAM_BATTING_2B   1     53.05 12602 826.16  0.7483  0.388185
## - TEAM_PITCHING_H   1     90.24 12639 826.72  1.2729  0.260759
## - TEAM_BATTING_H    1     92.13 12641 826.75  1.2995  0.255845
## - TEAM_BATTING_BB   1    110.31 12659 827.03  1.5559  0.213918
## - TEAM_PITCHING_BB  1    113.00 12662 827.07  1.5938  0.208444
## - TEAM_BASERUN_SB   1    123.42 12672 827.22  1.7408  0.188740
## - TEAM_BATTING_3B   1    129.33 12678 827.31  1.8242  0.178534
## <none>                          12549 827.35
## - TEAM_BATTING_HBP  1    197.23 12746 828.33  2.7819  0.097101 .
## - TEAM_FIELDING_DP  1    635.62 13184 834.79  8.9655  0.003145 **
## - TEAM_PITCHING_SO  1   1311.88 13861 844.35 18.5041 2.798e-05 ***
## - TEAM_FIELDING_E   1   1322.05 13871 844.49 18.6476 2.614e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=825.6
```

```
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_BB + TEAM_BATTING_HBP + TEAM_BASERUN_SB + TEAM_FIELDING_E +
##     TEAM_FIELDING_DP + TEAM_PITCHING_BB + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_2B    1     55.48 12620 824.44  0.7859 0.3765268
## - TEAM_PITCHING_H    1     89.26 12654 824.95  1.2645 0.2623142
## - TEAM_BATTING_H     1     91.97 12657 824.99  1.3029 0.2552256
## - TEAM_BATTING_BB    1    104.58 12669 825.18  1.4815 0.2251501
## - TEAM_PITCHING_BB   1    107.19 12672 825.22  1.5185 0.2194792
## <none>                           12565 825.60
## - TEAM_BATTING_3B    1    137.48 12702 825.68  1.9476 0.1645807
## - TEAM_BASERUN_SB    1    146.90 12712 825.82  2.0811 0.1508875
## - TEAM_BATTING_HBP   1    200.36 12765 826.62  2.8384 0.0937878 .
## - TEAM_FIELDING_DP   1    628.95 13194 832.93  8.9101 0.0032349 **
## - TEAM_PITCHING_HR   1    853.54 13418 836.15 12.0918 0.0006365 ***
## - TEAM_PITCHING_SO   1   1316.68 13882 842.63 18.6529 2.601e-05 ***
## - TEAM_FIELDING_E    1   1333.15 13898 842.86 18.8862 2.328e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=824.44
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_3B + TEAM_BATTING_BB +
##     TEAM_BATTING_HBP + TEAM_BASERUN_SB + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_PITCHING_BB + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_PITCHING_H    1     84.47 12705 823.71  1.1981 0.2751776
## - TEAM_BATTING_H     1     87.79 12708 823.76  1.2452 0.2659657
## - TEAM_BATTING_BB    1     98.92 12719 823.93  1.4031 0.2377794
## - TEAM_PITCHING_BB   1    101.48 12722 823.97  1.4394 0.2318250
## - TEAM_BASERUN_SB    1    109.27 12730 824.09  1.5499 0.2147828
## <none>                           12620 824.44
## - TEAM_BATTING_3B    1    147.01 12767 824.65  2.0851 0.1504896
## - TEAM_BATTING_HBP   1    204.39 12825 825.51  2.8990 0.0903707 .
## - TEAM_FIELDING_DP   1    649.12 13269 832.02  9.2068 0.0027703 **
## - TEAM_PITCHING_HR   1    812.92 13433 834.36 11.5301 0.0008435 ***
## - TEAM_PITCHING_SO   1   1262.90 13883 840.66 17.9124 3.690e-05 ***
## - TEAM_FIELDING_E    1   1379.34 14000 842.25 19.5640 1.685e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=823.71
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_3B + TEAM_BATTING_BB +
##     TEAM_BATTING_HBP + TEAM_BASERUN_SB + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_PITCHING_BB + TEAM_PITCHING_HR + TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_BB    1     32.85 12738 822.21  0.4655 0.4959485
## - TEAM_PITCHING_BB   1     43.42 12748 822.37  0.6151 0.4338894
## - TEAM_BASERUN_SB    1    105.16 12810 823.29  1.4899 0.2238228
## <none>                           12705 823.71
## - TEAM_BATTING_3B    1    153.13 12858 824.00  2.1696 0.1425129
```

```
## - TEAM_BATTING_HBP  1    183.82 12888 824.46  2.6043 0.1083238
## - TEAM_BATTING_H    1    504.11 13209 829.15  7.1423 0.0082203 **
## - TEAM_FIELDING_DP  1    602.80 13308 830.57  8.5404 0.0039191 **
## - TEAM_PITCHING_HR  1    850.25 13555 834.09 12.0463 0.0006496 ***
## - TEAM_PITCHING_SO  1   1259.72 13964 839.77 17.8476 3.797e-05 ***
## - TEAM_FIELDING_E   1   1419.39 14124 841.94 20.1098 1.299e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=822.21
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_3B + TEAM_BATTING_HBP +
##     TEAM_BASERUN_SB + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_PITCHING_BB +
##     TEAM_PITCHING_HR + TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BASERUN_SB   1    109.99 12848 821.85  1.5630 0.2128380
## <none>                          12738 822.21
## - TEAM_BATTING_3B   1    156.45 12894 822.54  2.2232 0.1376961
## - TEAM_BATTING_HBP  1    186.58 12924 822.98  2.6513 0.1052022
## - TEAM_BATTING_H    1    485.67 13223 827.35  6.9013 0.0093520 **
## - TEAM_FIELDING_DP  1    623.19 13361 829.33  8.8555 0.0033213 **
## - TEAM_PITCHING_HR  1    843.83 13581 832.46 11.9907 0.0006672 ***
## - TEAM_PITCHING_SO  1   1267.25 14005 838.32 18.0075 3.509e-05 ***
## - TEAM_FIELDING_E   1   1395.02 14133 840.06 19.8231 1.483e-05 ***
## - TEAM_PITCHING_BB  1   2364.81 15102 852.73 33.6038 2.951e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=821.85
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_3B + TEAM_BATTING_HBP +
##     TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_PITCHING_BB + TEAM_PITCHING_HR +
##     TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_3B   1    133.47 12981 821.82  1.8908  0.170802
## <none>                          12848 821.85
## - TEAM_BATTING_HBP  1    177.11 13025 822.46  2.5090  0.114934
## - TEAM_BATTING_H    1    566.11 13414 828.09  8.0196  0.005149 **
## - TEAM_FIELDING_DP  1    737.46 13585 830.51 10.4469  0.001459 **
## - TEAM_PITCHING_HR  1    756.49 13604 830.78 10.7166  0.001271 **
## - TEAM_PITCHING_SO  1   1257.91 14106 837.69 17.8197 3.829e-05 ***
## - TEAM_FIELDING_E   1   1330.40 14178 838.67 18.8466 2.347e-05 ***
## - TEAM_PITCHING_BB  1   2371.12 15219 852.20 33.5895 2.947e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=821.82
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HBP + TEAM_FIELDING_E +
##     TEAM_FIELDING_DP + TEAM_PITCHING_BB + TEAM_PITCHING_HR +
##     TEAM_PITCHING_SO
##
##                     Df Sum of Sq   RSS    AIC F value    Pr(>F)
## <none>                          12981 821.82
## - TEAM_BATTING_HBP  1    228.70 13210 823.16  3.2241 0.0742109 .
```

```
## - TEAM_BATTING_H    1    449.87 13431 826.33  6.3420 0.0126483 *
## - TEAM_FIELDING_DP  1    813.17 13794 831.43 11.4637 0.0008685 ***
## - TEAM_PITCHING_HR  1    990.20 13971 833.86 13.9593 0.0002495 ***
## - TEAM_PITCHING_SO  1   1316.56 14298 838.27 18.5602 2.683e-05 ***
## - TEAM_FIELDING_E   1   1334.60 14316 838.52 18.8145 2.377e-05 ***
## - TEAM_PITCHING_BB  1   2583.00 15564 854.49 36.4137 8.663e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coefficients(step1)
```

```
##      (Intercept)   TEAM_BATTING_H TEAM_BATTING_HBP  TEAM_FIELDING_E
##      60.95453718       0.02541359       0.08711968      -0.17218039
## TEAM_FIELDING_DP TEAM_PITCHING_BB TEAM_PITCHING_HR TEAM_PITCHING_SO
##      -0.11904326       0.05672227       0.08944977      -0.03136311
```

Model2: This is the second model.

```
model2<-lm(TARGET_WINS~TEAM_BATTING_H_NEW+TEAM_BATTING_2B_NEW+TEAM_BATTING_3B_NEW+TEAM_BATTING_BB_NEW+T

summary(model2)
```

```
##
## Call:
## lm(formula = TARGET_WINS ~ TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW +
##     TEAM_BATTING_3B_NEW + TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW +
##     TEAM_FIELDING_E_NEW + TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW +
##     TEAM_PITCHING_H_NEW + TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW,
##     data = na.omit(moneyball2))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.8141  -6.3893  -0.0595   5.0336  22.0504
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           58.83398   19.34512   3.041 0.002710 **
## TEAM_BATTING_H_NEW    -0.10194    0.20504  -0.497 0.619664
## TEAM_BATTING_2B_NEW    0.02566    0.03072   0.835 0.404644
## TEAM_BATTING_3B_NEW   -0.12553    0.07569  -1.658 0.098993 .
## TEAM_BATTING_BB_NEW    0.03674    0.08499   0.432 0.666031
## TEAM_BASERUN_SB_NEW    0.03137    0.02271   1.381 0.168873
## TEAM_FIELDING_E_NEW   -0.17714    0.04048  -4.376 2.05e-05 ***
## TEAM_FIELDING_DP_NEW -0.10377    0.03657  -2.838 0.005070 **
## TEAM_PITCHING_BB_NEW   0.01763    0.08317   0.212 0.832365
## TEAM_PITCHING_H_NEW    0.12603    0.20539   0.614 0.540252
## TEAM_PITCHING_HR_NEW   0.09054    0.02564   3.532 0.000525 ***
## TEAM_PITCHING_SO_NEW -0.02961    0.00731  -4.051 7.59e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.469 on 179 degrees of freedom
## Multiple R-squared:  0.5396, Adjusted R-squared:  0.5113
## F-statistic: 19.07 on 11 and 179 DF,  p-value: < 2.2e-16
```

```
# coefficients(model1) # model coefficients
# confint(model1, level=0.95) # CIs for model parameters
# fitted(model1) # predicted values
# residuals(model1) # residuals
# anova(model1) # anova table
# vcov(model1) # covariance matrix for model parameters
# influence(model1) # regression diagnostics
```

Lets now step thru this model and retain only those variables that have the most impact.

```
step2 <- step(model2,direction="backward",test="F")
```

```
## Start:  AIC=827.73
## TARGET_WINS ~ TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_PITCHING_BB_NEW  1      3.22 12843 825.78  0.0449 0.8323653
## - TEAM_BATTING_BB_NEW   1     13.41 12853 825.93  0.1869 0.6660312
## - TEAM_BATTING_H_NEW    1     17.73 12858 826.00  0.2472 0.6196643
## - TEAM_PITCHING_H_NEW   1     27.01 12867 826.13  0.3765 0.5402520
## - TEAM_BATTING_2B_NEW   1     50.05 12890 826.48  0.6978 0.4046437
## <none>                              12840 827.73
## - TEAM_BASERUN_SB_NEW   1    136.88 12977 827.76  1.9083 0.1688727
## - TEAM_BATTING_3B_NEW   1    197.27 13037 828.65  2.7502 0.0989930 .
## - TEAM_FIELDING_DP_NEW  1    577.56 13417 834.14  8.0519 0.0050705 **
## - TEAM_PITCHING_HR_NEW  1    894.67 13734 838.60 12.4727 0.0005252 ***
## - TEAM_PITCHING_SO_NEW  1   1177.21 14017 842.49 16.4116 7.586e-05 ***
## - TEAM_FIELDING_E_NEW   1   1373.69 14213 845.15 19.1508 2.049e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=825.78
## TARGET_WINS ~ TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_H_NEW + TEAM_PITCHING_HR_NEW +
##     TEAM_PITCHING_SO_NEW
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_H_NEW    1     20.13 12863 824.08  0.2822 0.5959495
## - TEAM_PITCHING_H_NEW   1     29.95 12873 824.23  0.4198 0.5178756
## - TEAM_BATTING_2B_NEW   1     50.77 12894 824.53  0.7115 0.4000676
## <none>                              12843 825.78
## - TEAM_BASERUN_SB_NEW   1    144.37 12987 825.92  2.0234 0.1566187
## - TEAM_BATTING_3B_NEW   1    199.12 13042 826.72  2.7907 0.0965485 .
## - TEAM_FIELDING_DP_NEW  1    577.04 13420 832.18  8.0874 0.0049733 **
## - TEAM_PITCHING_HR_NEW  1    898.17 13741 836.69 12.5882 0.0004951 ***
## - TEAM_PITCHING_SO_NEW  1   1178.49 14022 840.55 16.5171 7.194e-05 ***
## - TEAM_FIELDING_E_NEW   1   1382.32 14225 843.31 19.3738 1.838e-05 ***
```

```
## - TEAM_BATTING_BB_NEW   1   2247.06 15090 854.58 31.4936 7.452e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=824.08
## TARGET_WINS ~ TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW + TEAM_BATTING_BB_NEW +
##     TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW + TEAM_FIELDING_DP_NEW +
##     TEAM_PITCHING_H_NEW + TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_2B_NEW   1     48.07 12911 822.79  0.6765 0.4118873
## <none>                              12863 824.08
## - TEAM_BASERUN_SB_NEW   1    146.71 13010 824.25  2.0644 0.1525008
## - TEAM_BATTING_3B_NEW   1    205.02 13068 825.10  2.8849 0.0911321 .
## - TEAM_PITCHING_H_NEW   1    253.75 13117 825.81  3.5706 0.0604107 .
## - TEAM_FIELDING_DP_NEW  1    598.69 13462 830.77  8.4243 0.0041626 **
## - TEAM_PITCHING_HR_NEW  1    885.75 13749 834.80 12.4636 0.0005263 ***
## - TEAM_PITCHING_SO_NEW  1   1172.81 14036 838.75 16.5029 7.228e-05 ***
## - TEAM_FIELDING_E_NEW   1   1364.22 14227 841.33 19.1963 1.994e-05 ***
## - TEAM_BATTING_BB_NEW   1   2264.52 15128 853.05 31.8646 6.293e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=822.79
## TARGET_WINS ~ TEAM_BATTING_3B_NEW + TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW +
##     TEAM_FIELDING_E_NEW + TEAM_FIELDING_DP_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BASERUN_SB_NEW   1    113.32 13024 822.46  1.5974 0.2078878
## <none>                              12911 822.79
## - TEAM_BATTING_3B_NEW   1    218.04 13129 823.99  3.0736 0.0812590 .
## - TEAM_PITCHING_H_NEW   1    526.52 13438 828.43  7.4220 0.0070712 **
## - TEAM_FIELDING_DP_NEW  1    603.21 13514 829.51  8.5031 0.0039914 **
## - TEAM_PITCHING_HR_NEW  1    847.68 13759 832.94 11.9492 0.0006805 ***
## - TEAM_PITCHING_SO_NEW  1   1125.64 14037 836.76 15.8674 9.815e-05 ***
## - TEAM_FIELDING_E_NEW   1   1435.65 14347 840.93 20.2373 1.217e-05 ***
## - TEAM_BATTING_BB_NEW   1   2290.03 15201 851.98 32.2809 5.211e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=822.46
## TARGET_WINS ~ TEAM_BATTING_3B_NEW + TEAM_BATTING_BB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_H_NEW + TEAM_PITCHING_HR_NEW +
##     TEAM_PITCHING_SO_NEW
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## <none>                              13024 822.46
## - TEAM_BATTING_3B_NEW   1    189.69 13214 823.22  2.6653 0.1042808
## - TEAM_PITCHING_H_NEW   1    612.89 13637 829.24  8.6113 0.0037688 **
## - TEAM_FIELDING_DP_NEW  1    715.75 13740 830.68 10.0566 0.0017802 **
## - TEAM_PITCHING_HR_NEW  1    759.30 13784 831.28 10.6686 0.0013011 **
## - TEAM_PITCHING_SO_NEW  1   1118.70 14143 836.20 15.7182 0.0001053 ***
## - TEAM_FIELDING_E_NEW   1   1369.82 14394 839.56 19.2466 1.936e-05 ***
```

```
## - TEAM_BATTING_BB_NEW   1   2282.86 15307 851.31 32.0752 5.664e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coefficients(step2)
```

```
##          (Intercept)  TEAM_BATTING_3B_NEW  TEAM_BATTING_BB_NEW
##          59.66418057          -0.12207345           0.05500336
##  TEAM_FIELDING_E_NEW TEAM_FIELDING_DP_NEW  TEAM_PITCHING_H_NEW
##          -0.17423573          -0.11230653           0.03134957
## TEAM_PITCHING_HR_NEW TEAM_PITCHING_SO_NEW
##           0.08093837          -0.02841523
```

Model3: This is the third model.

```
model3<-lm(TARGET_WINS~TEAM_BATTING_H_SIN+TEAM_BATTING_2B_SIN+TEAM_BATTING_3B_SIN+TEAM_BATTING_BB_SIN+T

summary(model3)
```

```
##
## Call:
## lm(formula = TARGET_WINS ~ TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN +
##     TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     TEAM_BATTING_HBP_Missing + TEAM_BASERUN_CS_Missing + Hits_R +
##     Walks_R + HomeRuns_R + Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B,
##     data = na.omit(moneyball2))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -23.0633  -7.2221   0.1263   6.9949  24.0791
##
## Coefficients: (2 not defined because of singularities)
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               4.557e+02  4.035e+02   1.129  0.26031
## TEAM_BATTING_H_SIN       -6.147e-01  2.083e+00  -0.295  0.76823
## TEAM_BATTING_2B_SIN       8.235e-02  1.085e+00   0.076  0.93960
## TEAM_BATTING_3B_SIN       5.884e-01  1.144e+00   0.514  0.60772
## TEAM_BATTING_BB_SIN      -2.319e+00  2.291e+00  -1.012  0.31295
## TEAM_BASERUN_SB_SIN      -2.182e+00  1.104e+00  -1.978  0.04957 *
## TEAM_FIELDING_E_SIN       5.054e-01  1.099e+00   0.460  0.64625
## TEAM_FIELDING_DP_SIN      2.355e+00  1.115e+00   2.113  0.03602 *
## TEAM_PITCHING_BB_SIN      4.716e-01  2.246e+00   0.210  0.83392
## TEAM_PITCHING_H_SIN       7.726e-01  2.068e+00   0.374  0.70920
## TEAM_PITCHING_HR_SIN     -1.696e+00  1.101e+00  -1.541  0.12526
## TEAM_PITCHING_SO_SIN      7.777e-01  1.113e+00   0.699  0.48564
## TEAM_BATTING_HBP_Missing         NA         NA      NA       NA
## TEAM_BASERUN_CS_Missing          NA         NA      NA       NA
## Hits_R                    4.799e+02  9.617e+03   0.050  0.96026
## Walks_R                  -1.007e+04  5.217e+03  -1.930  0.05524 .
## HomeRuns_R                3.948e+03  2.006e+03   1.968  0.05068 .
```

```
## Strikeout_R               5.172e+03  8.565e+03   0.604  0.54673
## TEAM_BATTING_EB           1.020e-01  1.756e-02   5.812  2.9e-08 ***
## TEAM_BATTING_1B           4.287e-02  1.298e-02   3.303  0.00116 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.14 on 173 degrees of freedom
## Multiple R-squared:  0.3622, Adjusted R-squared:  0.2995
## F-statistic: 5.778 on 17 and 173 DF,  p-value: 2.536e-10
```

```
# coefficients(model1) # model coefficients
# confint(model1, level=0.95) # CIs for model parameters
# fitted(model1) # predicted values
# residuals(model1) # residuals
# anova(model1) # anova table
# vcov(model1) # covariance matrix for model parameters
# influence(model1) # regression diagnostics
```

Lets now step thru this model and retain only those variables that have the most impact.

```
step3 <- step(model3,direction="backward",test="F")
```

```
## Start:  AIC=901.99
## TARGET_WINS ~ TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + TEAM_BATTING_HBP_Missing +
##     TEAM_BASERUN_CS_Missing + Hits_R + Walks_R + HomeRuns_R +
##     Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##
## Step:  AIC=901.99
## TARGET_WINS ~ TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + TEAM_BATTING_HBP_Missing +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R + TEAM_BATTING_EB +
##     TEAM_BATTING_1B
##
##
## Step:  AIC=901.99
## TARGET_WINS ~ TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - Hits_R                 1       0.3 17788 899.99  0.0025  0.960256
## - TEAM_BATTING_2B_SIN    1       0.6 17788 899.99  0.0058  0.939598
## - TEAM_PITCHING_BB_SIN   1       4.5 17792 900.04  0.0441  0.833924
```

```
## - TEAM_BATTING_H_SIN    1       9.0 17796 900.08  0.0871  0.768229
## - TEAM_PITCHING_H_SIN   1      14.3 17802 900.14  0.1395  0.709202
## - TEAM_FIELDING_E_SIN   1      21.7 17809 900.22  0.2114  0.646255
## - TEAM_BATTING_3B_SIN   1      27.2 17815 900.28  0.2645  0.607716
## - Strikeout_R           1      37.5 17825 900.39  0.3646  0.546730
## - TEAM_PITCHING_SO_SIN  1      50.2 17838 900.53  0.4883  0.485643
## - TEAM_BATTING_BB_SIN   1     105.3 17893 901.12  1.0241  0.312953
## <none>                              17787 901.99
## - TEAM_PITCHING_HR_SIN  1     244.0 18031 902.59  2.3732  0.125257
## - Walks_R               1     383.0 18170 904.06  3.7249  0.055242 .
## - HomeRuns_R            1     398.2 18186 904.22  3.8726  0.050679 .
## - TEAM_BASERUN_SB_SIN   1     402.1 18189 904.26  3.9106  0.049570 *
## - TEAM_FIELDING_DP_SIN  1     459.1 18246 904.86  4.4654  0.036023 *
## - TEAM_BATTING_1B       1    1121.9 18909 911.67 10.9115  0.001161 **
## - TEAM_BATTING_EB       1    3473.5 21261 934.06 33.7836 2.902e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=899.99
## TARGET_WINS ~ TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Walks_R + HomeRuns_R +
##     Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_2B_SIN    1       0.5 17788 898.00  0.0050  0.943480
## - TEAM_PITCHING_BB_SIN   1       5.8 17793 898.05  0.0571  0.811476
## - TEAM_BATTING_H_SIN     1       9.0 17797 898.09  0.0884  0.766539
## - TEAM_PITCHING_H_SIN    1      14.9 17802 898.15  0.1454  0.703465
## - TEAM_FIELDING_E_SIN    1      21.7 17809 898.22  0.2121  0.645705
## - TEAM_BATTING_3B_SIN    1      27.1 17815 898.28  0.2651  0.607306
## - TEAM_PITCHING_SO_SIN   1      51.2 17839 898.54  0.5012  0.479915
## - TEAM_BATTING_BB_SIN    1     121.3 17909 899.29  1.1866  0.277529
## - Strikeout_R            1     126.7 17914 899.35  1.2391  0.267187
## <none>                              17788 899.99
## - TEAM_PITCHING_HR_SIN   1     244.1 18032 900.59  2.3883  0.124066
## - TEAM_BASERUN_SB_SIN    1     402.1 18190 902.26  3.9336  0.048904 *
## - HomeRuns_R             1     409.7 18197 902.34  4.0078  0.046844 *
## - Walks_R                1     463.8 18251 902.91  4.5371  0.034574 *
## - TEAM_FIELDING_DP_SIN   1     466.1 18254 902.93  4.5593  0.034139 *
## - TEAM_BATTING_1B        1    1124.2 18912 909.70 10.9974  0.001111 **
## - TEAM_BATTING_EB        1    3490.7 21278 932.21 34.1460 2.463e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=898
## TARGET_WINS ~ TEAM_BATTING_H_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Walks_R + HomeRuns_R + Strikeout_R +
##     TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
```

```
## - TEAM_PITCHING_BB_SIN  1       5.9 17794 896.06  0.0577  0.810441
## - TEAM_BATTING_H_SIN    1       8.9 17797 896.09  0.0877  0.767457
## - TEAM_PITCHING_H_SIN   1      14.8 17803 896.15  0.1455  0.703331
## - TEAM_FIELDING_E_SIN   1      21.7 17810 896.23  0.2131  0.644929
## - TEAM_BATTING_3B_SIN   1      26.7 17815 896.28  0.2622  0.609243
## - TEAM_PITCHING_SO_SIN  1      52.8 17841 896.56  0.5199  0.471853
## - TEAM_BATTING_BB_SIN   1     122.3 17910 897.30  1.2032  0.274189
## - Strikeout_R           1     126.9 17915 897.35  1.2487  0.265337
## <none>                              17788 898.00
## - TEAM_PITCHING_HR_SIN  1     243.7 18032 898.60  2.3979  0.123303
## - HomeRuns_R            1     410.1 18198 900.35  4.0344  0.046119 *
## - TEAM_BASERUN_SB_SIN   1     413.3 18201 900.38  4.0656  0.045293 *
## - Walks_R               1     464.6 18253 900.92  4.5704  0.033917 *
## - TEAM_FIELDING_DP_SIN  1     470.7 18259 900.98  4.6304  0.032781 *
## - TEAM_BATTING_1B       1    1136.9 18925 907.83 11.1852  0.001009 **
## - TEAM_BATTING_EB       1    3495.5 21284 930.26 34.3892 2.201e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=896.06
## TARGET_WINS ~ TEAM_BATTING_H_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Walks_R + HomeRuns_R + Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_H_SIN     1      12.8 17807 894.20  0.1261 0.7229122
## - TEAM_PITCHING_H_SIN    1      19.7 17814 894.27  0.1948 0.6594623
## - TEAM_FIELDING_E_SIN    1      21.7 17816 894.29  0.2142 0.6440472
## - TEAM_BATTING_3B_SIN    1      28.9 17823 894.37  0.2862 0.5933715
## - TEAM_PITCHING_SO_SIN   1      49.2 17843 894.59  0.4868 0.4862673
## - Strikeout_R            1     121.4 17915 895.36  1.2008 0.2746606
## <none>                              17794 896.06
## - TEAM_PITCHING_HR_SIN   1     250.7 18045 896.73  2.4802 0.1170865
## - TEAM_BATTING_BB_SIN    1     326.1 18120 897.53  3.2256 0.0742118 .
## - TEAM_BASERUN_SB_SIN    1     442.8 18237 898.75  4.3796 0.0378045 *
## - HomeRuns_R             1     449.3 18243 898.82  4.4436 0.0364474 *
## - Walks_R                1     458.7 18253 898.92  4.5370 0.0345591 *
## - TEAM_FIELDING_DP_SIN   1     466.7 18261 899.00  4.6160 0.0330422 *
## - TEAM_BATTING_1B        1    1134.4 18928 905.86 11.2204 0.0009898 ***
## - TEAM_BATTING_EB        1    3493.7 21288 928.30 34.5565 2.032e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=894.2
## TARGET_WINS ~ TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Walks_R + HomeRuns_R +
##     Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_PITCHING_H_SIN    1       7.0 17814 892.27  0.0697  0.792043
## - TEAM_FIELDING_E_SIN    1      23.9 17831 892.45  0.2372  0.626873
## - TEAM_BATTING_3B_SIN    1      28.1 17835 892.50  0.2792  0.597870
```

```
## - TEAM_PITCHING_SO_SIN  1      52.5 17859 892.76  0.5217  0.471055
## - Strikeout_R           1     131.9 17939 893.61  1.3106  0.253831
## <none>                              17807 894.20
## - TEAM_PITCHING_HR_SIN  1     271.4 18078 895.09  2.6978  0.102262
## - TEAM_BATTING_BB_SIN   1     333.0 18140 895.73  3.3100  0.070550 .
## - TEAM_BASERUN_SB_SIN   1     462.5 18269 897.09  4.5969  0.033394 *
## - TEAM_FIELDING_DP_SIN  1     469.2 18276 897.16  4.6643  0.032140 *
## - HomeRuns_R            1     481.2 18288 897.29  4.7830  0.030054 *
## - Walks_R               1     507.4 18314 897.56  5.0438  0.025951 *
## - TEAM_BATTING_1B       1    1138.6 18945 904.03 11.3182  0.000941 ***
## - TEAM_BATTING_EB       1    3481.0 21288 926.30 34.6016 1.978e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=892.27
## TARGET_WINS ~ TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Walks_R + HomeRuns_R + Strikeout_R +
##     TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_FIELDING_E_SIN    1      25.6 17839 890.55  0.2557 0.6137323
## - TEAM_BATTING_3B_SIN    1      28.9 17843 890.58  0.2891 0.5914530
## - TEAM_PITCHING_SO_SIN   1      53.1 17867 890.84  0.5309 0.4671655
## - Strikeout_R            1     129.7 17944 891.66  1.2962 0.2564317
## <none>                              17814 892.27
## - TEAM_PITCHING_HR_SIN   1     268.9 18083 893.13  2.6871 0.1029289
## - TEAM_BATTING_BB_SIN    1     340.2 18154 893.88  3.3991 0.0668946 .
## - TEAM_BASERUN_SB_SIN    1     463.7 18277 895.18  4.6334 0.0327021 *
## - TEAM_FIELDING_DP_SIN   1     470.7 18284 895.25  4.7034 0.0314295 *
## - HomeRuns_R             1     492.6 18306 895.48  4.9226 0.0277726 *
## - Walks_R                1     506.8 18321 895.63  5.0643 0.0256477 *
## - TEAM_BATTING_1B        1    1136.9 18951 902.09 11.3598 0.0009204 ***
## - TEAM_BATTING_EB        1    3514.3 21328 924.66 35.1156 1.573e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=890.55
## TARGET_WINS ~ TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Walks_R + HomeRuns_R + Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_3B_SIN    1      24.7 17864 888.81  0.2475 0.6194230
## - TEAM_PITCHING_SO_SIN   1      50.7 17890 889.09  0.5088 0.4766042
## - Strikeout_R            1     136.0 17975 890.00  1.3647 0.2442863
## <none>                              17839 890.55
## - TEAM_PITCHING_HR_SIN   1     284.6 18124 891.57  2.8555 0.0928043 .
## - TEAM_BATTING_BB_SIN    1     338.8 18178 892.14  3.3996 0.0668640 .
## - TEAM_FIELDING_DP_SIN   1     457.7 18297 893.38  4.5924 0.0334646 *
## - TEAM_BASERUN_SB_SIN    1     463.0 18302 893.44  4.6461 0.0324598 *
## - HomeRuns_R             1     505.7 18345 893.88  5.0743 0.0254968 *
## - Walks_R                1     526.3 18366 894.10  5.2806 0.0227203 *
## - TEAM_BATTING_1B        1    1211.8 19051 901.10 12.1595 0.0006145 ***
```

```
## - TEAM_BATTING_EB       1    3494.2 21334 922.71 35.0612 1.598e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=888.81
## TARGET_WINS ~ TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Walks_R + HomeRuns_R +
##     Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_PITCHING_SO_SIN  1      54.1 17918 887.39  0.5451 0.4612907
## - Strikeout_R           1     129.9 17994 888.19  1.3085 0.2541899
## <none>                              17864 888.81
## - TEAM_PITCHING_HR_SIN  1     299.8 18164 889.99  3.0213 0.0838877 .
## - TEAM_BATTING_BB_SIN   1     357.6 18222 890.60  3.6037 0.0592529 .
## - TEAM_BASERUN_SB_SIN   1     438.8 18303 891.44  4.4219 0.0368715 *
## - HomeRuns_R            1     501.5 18365 892.10  5.0527 0.0258014 *
## - Walks_R               1     511.6 18376 892.20  5.1546 0.0243702 *
## - TEAM_FIELDING_DP_SIN  1     572.9 18437 892.84  5.7731 0.0172895 *
## - TEAM_BATTING_1B       1    1216.1 19080 899.39 12.2540 0.0005853 ***
## - TEAM_BATTING_EB       1    3476.9 21341 920.78 35.0341 1.604e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=887.39
## TARGET_WINS ~ TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_HR_SIN + Walks_R + HomeRuns_R + Strikeout_R +
##     TEAM_BATTING_EB + TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - Strikeout_R           1     126.3 18044 886.73  1.2758  0.260176
## <none>                              17918 887.39
## - TEAM_PITCHING_HR_SIN  1     335.7 18254 888.93  3.3907  0.067200 .
## - TEAM_BATTING_BB_SIN   1     366.5 18285 889.25  3.7024  0.055902 .
## - TEAM_BASERUN_SB_SIN   1     435.0 18353 889.97  4.3940  0.037456 *
## - HomeRuns_R            1     494.2 18412 890.58  4.9922  0.026686 *
## - Walks_R               1     497.7 18416 890.62  5.0271  0.026168 *
## - TEAM_FIELDING_DP_SIN  1     556.4 18475 891.23  5.6201  0.018806 *
## - TEAM_BATTING_1B       1    1210.9 19129 897.88 12.2323  0.000591 ***
## - TEAM_BATTING_EB       1    3571.5 21490 920.10 36.0775 1.017e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=886.73
## TARGET_WINS ~ TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_HR_SIN + Walks_R + HomeRuns_R + TEAM_BATTING_EB +
##     TEAM_BATTING_1B
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## <none>                              18044 886.73
## - TEAM_PITCHING_HR_SIN  1     288.0 18332 887.75  2.9052 0.0900020 .
## - TEAM_BATTING_BB_SIN   1     364.3 18409 888.55  3.6746 0.0568132 .
## - TEAM_BASERUN_SB_SIN   1     453.0 18497 889.46  4.5695 0.0338803 *
## - TEAM_FIELDING_DP_SIN  1     535.8 18580 890.32  5.4040 0.0211945 *
```

```
## - HomeRuns_R            1     801.4 18846 893.03  8.0827 0.0049799 **
## - Walks_R               1     866.8 18911 893.69  8.7426 0.0035208 **
## - TEAM_BATTING_1B       1    1246.5 19291 897.49 12.5721 0.0004978 ***
## - TEAM_BATTING_EB       1    3589.5 21634 919.38 36.2050  9.55e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coefficients(step3)
```

```
##          (Intercept)  TEAM_BATTING_BB_SIN  TEAM_BASERUN_SB_SIN
##         4.073260e+02        -2.007580e+00        -2.201608e+00
## TEAM_FIELDING_DP_SIN TEAM_PITCHING_HR_SIN              Walks_R
##         2.397104e+00        -1.778182e+00        -5.393058e+03
##           HomeRuns_R      TEAM_BATTING_EB      TEAM_BATTING_1B
##         4.971968e+03         1.018484e-01         4.418862e-02
```

Model4: This is the fourth and final model.

```
model4<-lm(TARGET_WINS~., na.omit(moneyball2))

summary(model4)
```

```
##
## Call:
## lm(formula = TARGET_WINS ~ ., data = na.omit(moneyball2))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -22.5999  -4.8981  -0.1927   5.0706  25.6716
##
## Coefficients: (14 not defined because of singularities)
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)              1.580e+04  1.633e+04   0.967   0.3348
## TEAM_BATTING_H           1.419e+01  1.160e+01   1.224   0.2229
## TEAM_BATTING_2B          6.909e-02  1.226e-01   0.564   0.5738
## TEAM_BATTING_3B         -6.702e-02  8.034e-02  -0.834   0.4055
## TEAM_BATTING_HR         -3.138e+01  2.208e+01  -1.422   0.1571
## TEAM_BATTING_BB          1.486e+01  8.533e+00   1.742   0.0835 .
## TEAM_BATTING_SO         -7.531e+00  3.956e+00  -1.904   0.0588 .
## TEAM_BASERUN_SB          2.898e-02  3.006e-02   0.964   0.3365
## TEAM_BASERUN_CS         -2.575e-02  7.291e-02  -0.353   0.7244
## TEAM_BATTING_HBP         8.849e-02  5.054e-02   1.751   0.0819 .
## TEAM_PITCHING_H         -1.418e+01  1.160e+01  -1.223   0.2233
## TEAM_PITCHING_HR         3.147e+01  2.207e+01   1.426   0.1560
## TEAM_PITCHING_BB        -1.484e+01  8.538e+00  -1.738   0.0841 .
## TEAM_PITCHING_SO         7.497e+00  3.954e+00   1.896   0.0598 .
## TEAM_FIELDING_E         -1.900e-01  4.329e-02  -4.388 2.09e-05 ***
## TEAM_FIELDING_DP        -9.830e-02  3.818e-02  -2.574   0.0110 *
## TEAM_BATTING_H_NEW              NA         NA      NA       NA
## TEAM_BATTING_2B_NEW     -4.696e-02  1.246e-01  -0.377   0.7069
## TEAM_BATTING_3B_NEW             NA         NA      NA       NA
## TEAM_BATTING_BB_NEW      3.109e-02  8.693e-02   0.358   0.7211
```

```
## TEAM_BASERUN_SB_NEW              NA         NA      NA       NA
## TEAM_FIELDING_E_NEW              NA         NA      NA       NA
## TEAM_FIELDING_DP_NEW             NA         NA      NA       NA
## TEAM_PITCHING_BB_NEW             NA         NA      NA       NA
## TEAM_PITCHING_H_NEW              NA         NA      NA       NA
## TEAM_PITCHING_HR_NEW             NA         NA      NA       NA
## TEAM_PITCHING_SO_NEW             NA         NA      NA       NA
## TEAM_BATTING_H_SIN       -8.180e-01  1.886e+00  -0.434   0.6651
## TEAM_BATTING_2B_SIN      -6.811e-01  9.278e-01  -0.734   0.4640
## TEAM_BATTING_3B_SIN      -4.102e-01  9.855e-01  -0.416   0.6778
## TEAM_BATTING_BB_SIN      -1.008e+00  1.983e+00  -0.508   0.6118
## TEAM_BASERUN_SB_SIN      -2.301e+00  9.340e-01  -2.464   0.0148 *
## TEAM_FIELDING_E_SIN      -4.924e-01  9.278e-01  -0.531   0.5964
## TEAM_FIELDING_DP_SIN      1.766e+00  9.543e-01   1.851   0.0661 .
## TEAM_PITCHING_BB_SIN     -7.478e-02  1.943e+00  -0.038   0.9694
## TEAM_PITCHING_H_SIN       1.078e+00  1.869e+00   0.577   0.5648
## TEAM_PITCHING_HR_SIN     -9.515e-01  9.362e-01  -1.016   0.3110
## TEAM_PITCHING_SO_SIN     -9.082e-01  9.605e-01  -0.946   0.3458
## TEAM_BATTING_SO_NEW              NA         NA      NA       NA
## TEAM_BATTING_HBP_Missing         NA         NA      NA       NA
## TEAM_BASERUN_CS_Missing          NA         NA      NA       NA
## Hits_R                   -2.362e+04  2.053e+04  -1.150   0.2518
## Walks_R                  -1.804e+04  9.127e+03  -1.977   0.0498 *
## HomeRuns_R                1.188e+04  6.110e+03   1.944   0.0537 .
## Strikeout_R               1.405e+04  8.910e+03   1.577   0.1168
## TEAM_BATTING_EB                  NA         NA      NA       NA
## TEAM_BATTING_1B                  NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.272 on 158 degrees of freedom
## Multiple R-squared:  0.6123, Adjusted R-squared:  0.5338
## F-statistic: 7.798 on 32 and 158 DF,  p-value: < 2.2e-16
```

```
# coefficients(model1) # model coefficients
# confint(model1, level=0.95) # CIs for model parameters
# fitted(model1) # predicted values
# residuals(model1) # residuals
# anova(model1) # anova table
# vcov(model1) # covariance matrix for model parameters
# influence(model1) # regression diagnostics
```

Lets now step thru this model and retain only those variables that have the most impact.

```
step4 <- step(model4,direction="backward",test="F")
```

```
## Start:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
```

```
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + TEAM_BATTING_SO_NEW + TEAM_BATTING_HBP_Missing +
##     TEAM_BASERUN_CS_Missing + Hits_R + Walks_R + HomeRuns_R +
##     Strikeout_R + TEAM_BATTING_EB + TEAM_BATTING_1B
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + TEAM_BATTING_SO_NEW + TEAM_BATTING_HBP_Missing +
##     TEAM_BASERUN_CS_Missing + Hits_R + Walks_R + HomeRuns_R +
##     Strikeout_R + TEAM_BATTING_EB
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + TEAM_BATTING_SO_NEW + TEAM_BATTING_HBP_Missing +
##     TEAM_BASERUN_CS_Missing + Hits_R + Walks_R + HomeRuns_R +
##     Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
```

```
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + TEAM_BATTING_SO_NEW + TEAM_BATTING_HBP_Missing +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + TEAM_BATTING_SO_NEW + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_PITCHING_SO_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_PITCHING_HR_NEW + TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN +
##     TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##
```

```
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_PITCHING_H_NEW +
##     TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_PITCHING_BB_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_FIELDING_DP_NEW + TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN +
##     TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_FIELDING_E_NEW +
##     TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
```

```
##     HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BASERUN_SB_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_3B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN +
##     TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_H_NEW + TEAM_BATTING_2B_NEW + TEAM_BATTING_BB_NEW +
##     TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R
##
##
## Step:  AIC=836.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_2B_NEW + TEAM_BATTING_BB_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_BB_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
```

```
## - TEAM_PITCHING_BB_SIN  1      0.10 10812 834.89  0.0015   0.96935
## - TEAM_BASERUN_CS       1      8.54 10820 835.04  0.1248   0.72440
## - TEAM_BATTING_BB_NEW   1      8.75 10820 835.05  0.1279   0.72107
## - TEAM_BATTING_2B_NEW   1      9.71 10821 835.06  0.1419   0.70686
## - TEAM_BATTING_3B_SIN   1     11.86 10823 835.10  0.1733   0.67780
## - TEAM_BATTING_H_SIN    1     12.87 10824 835.12  0.1881   0.66508
## - TEAM_BATTING_BB_SIN   1     17.69 10829 835.20  0.2586   0.61182
## - TEAM_FIELDING_E_SIN   1     19.27 10831 835.23  0.2816   0.59639
## - TEAM_BATTING_2B       1     21.74 10833 835.27  0.3177   0.57382
## - TEAM_PITCHING_H_SIN   1     22.78 10834 835.29  0.3329   0.56479
## - TEAM_BATTING_2B_SIN   1     36.88 10848 835.54  0.5390   0.46395
## - TEAM_BATTING_3B       1     47.61 10859 835.73  0.6958   0.40546
## - TEAM_PITCHING_SO_SIN  1     61.18 10872 835.97  0.8941   0.34581
## - TEAM_BASERUN_SB       1     63.58 10875 836.01  0.9292   0.33655
## - TEAM_PITCHING_HR_SIN  1     70.67 10882 836.14  1.0329   0.31104
## - Hits_R                1     90.52 10902 836.48  1.3229   0.25181
## - TEAM_PITCHING_H       1    102.30 10914 836.69  1.4950   0.22327
## - TEAM_BATTING_H        1    102.48 10914 836.69  1.4976   0.22286
## <none>                             10811 836.89
## - TEAM_BATTING_HR       1    138.28 10950 837.32  2.0208   0.15712
## - TEAM_PITCHING_HR      1    139.05 10950 837.33  2.0322   0.15597
## - Strikeout_R           1    170.21 10982 837.87  2.4875   0.11675
## - TEAM_PITCHING_BB      1    206.80 11018 838.51  3.0223   0.08408 .
## - TEAM_BATTING_BB       1    207.64 11019 838.52  3.0346   0.08345 .
## - TEAM_BATTING_HBP      1    209.81 11021 838.56  3.0662   0.08188 .
## - TEAM_FIELDING_DP_SIN  1    234.37 11046 838.99  3.4251   0.06608 .
## - TEAM_PITCHING_SO      1    246.00 11057 839.19  3.5951   0.05978 .
## - TEAM_BATTING_SO       1    247.98 11059 839.22  3.6240   0.05877 .
## - HomeRuns_R            1    258.61 11070 839.41  3.7793   0.05367 .
## - Walks_R               1    267.37 11079 839.56  3.9074   0.04981 *
## - TEAM_BASERUN_SB_SIN   1    415.38 11227 842.09  6.0704   0.01482 *
## - TEAM_FIELDING_DP      1    453.43 11265 842.74  6.6265   0.01097 *
## - TEAM_FIELDING_E       1   1317.23 12129 856.85 19.2503 2.087e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=834.89
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BASERUN_CS + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_2B_NEW + TEAM_BATTING_BB_NEW + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BASERUN_CS        1      8.69 10820 833.05  0.1278   0.72121
## - TEAM_BATTING_BB_NEW    1      8.75 10820 833.05  0.1287   0.72022
## - TEAM_BATTING_2B_NEW    1      9.73 10821 833.06  0.1430   0.70579
## - TEAM_BATTING_3B_SIN    1     12.09 10824 833.11  0.1778   0.67384
## - TEAM_BATTING_H_SIN     1     13.49 10825 833.13  0.1985   0.65658
## - TEAM_FIELDING_E_SIN    1     19.27 10831 833.23  0.2834   0.59524
```

```
## - TEAM_BATTING_2B       1     21.77 10833 833.28  0.3201   0.57232
## - TEAM_PITCHING_H_SIN   1     24.12 10836 833.32  0.3547   0.55234
## - TEAM_BATTING_2B_SIN   1     37.34 10849 833.55  0.5492   0.45974
## - TEAM_BATTING_3B       1     47.53 10859 833.73  0.6990   0.40436
## - TEAM_PITCHING_SO_SIN  1     61.90 10873 833.98  0.9103   0.34148
## - TEAM_BASERUN_SB       1     64.32 10876 834.03  0.9460   0.33222
## - TEAM_PITCHING_HR_SIN  1     70.63 10882 834.14  1.0387   0.30967
## - Hits_R                1     90.49 10902 834.48  1.3308   0.25040
## - TEAM_BATTING_BB_SIN   1     93.14 10905 834.53  1.3697   0.24361
## - TEAM_PITCHING_H       1    105.40 10917 834.75  1.5501   0.21495
## - TEAM_BATTING_H        1    105.59 10917 834.75  1.5529   0.21454
## <none>                             10812 834.89
## - TEAM_BATTING_HR       1    147.25 10959 835.48  2.1655   0.14311
## - TEAM_PITCHING_HR      1    148.08 10960 835.49  2.1778   0.14200
## - Strikeout_R           1    177.34 10989 836.00  2.6081   0.10830
## - TEAM_BATTING_HBP      1    209.78 11021 836.56  3.0852   0.08093 .
## - TEAM_PITCHING_BB      1    213.07 11024 836.62  3.1336   0.07861 .
## - TEAM_BATTING_BB       1    213.95 11025 836.64  3.1465   0.07800 .
## - TEAM_FIELDING_DP_SIN  1    236.68 11048 837.03  3.4808   0.06393 .
## - TEAM_PITCHING_SO      1    253.26 11065 837.32  3.7247   0.05539 .
## - TEAM_BATTING_SO       1    255.30 11067 837.35  3.7546   0.05443 .
## - Walks_R               1    271.09 11082 837.62  3.9868   0.04757 *
## - HomeRuns_R            1    276.55 11088 837.72  4.0671   0.04541 *
## - TEAM_BASERUN_SB_SIN   1    423.04 11234 840.22  6.2215   0.01364 *
## - TEAM_FIELDING_DP      1    458.64 11270 840.83  6.7450   0.01028 *
## - TEAM_FIELDING_E       1   1321.53 12133 854.92 19.4353 1.908e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=833.05
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_2B_NEW +
##     TEAM_BATTING_BB_NEW + TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_3B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_2B_NEW    1     10.02 10830 831.22  0.1482  0.700769
## - TEAM_BATTING_3B_SIN    1     10.43 10831 831.23  0.1543  0.694983
## - TEAM_BATTING_BB_NEW    1     10.83 10831 831.24  0.1601  0.689615
## - TEAM_BATTING_H_SIN     1     13.37 10834 831.28  0.1976  0.657236
## - TEAM_FIELDING_E_SIN    1     17.36 10838 831.35  0.2566  0.613136
## - TEAM_BATTING_2B        1     22.08 10842 831.44  0.3264  0.568568
## - TEAM_PITCHING_H_SIN    1     24.02 10844 831.47  0.3552  0.552019
## - TEAM_BATTING_2B_SIN    1     34.66 10855 831.66  0.5125  0.475094
## - TEAM_BATTING_3B        1     54.21 10874 832.00  0.8016  0.371958
## - TEAM_BASERUN_SB        1     59.81 10880 832.10  0.8845  0.348399
## - TEAM_PITCHING_SO_SIN   1     64.80 10885 832.19  0.9582  0.329123
## - TEAM_PITCHING_HR_SIN   1     74.92 10895 832.36  1.1078  0.294141
## - Hits_R                 1     92.53 10913 832.67  1.3683  0.243846
```

```
## - TEAM_BATTING_BB_SIN   1     93.61 10914 832.69  1.3843  0.241120
## - TEAM_PITCHING_H       1    108.15 10928 832.95  1.5992  0.207856
## - TEAM_BATTING_H        1    108.35 10928 832.95  1.6022  0.207432
## <none>                             10820 833.05
## - TEAM_BATTING_HR       1    146.52 10967 833.62  2.1667  0.142995
## - TEAM_PITCHING_HR      1    147.36 10968 833.63  2.1791  0.141865
## - Strikeout_R           1    177.50 10998 834.15  2.6248  0.107176
## - TEAM_PITCHING_BB      1    208.49 11029 834.69  3.0830  0.081027 .
## - TEAM_BATTING_BB       1    209.30 11029 834.71  3.0950  0.080446 .
## - TEAM_BATTING_HBP      1    210.23 11030 834.72  3.1087  0.079782 .
## - TEAM_PITCHING_SO      1    248.86 11069 835.39  3.6800  0.056851 .
## - TEAM_FIELDING_DP_SIN  1    250.28 11070 835.41  3.7010  0.056155 .
## - TEAM_BATTING_SO       1    250.87 11071 835.42  3.7096  0.055872 .
## - Walks_R               1    269.10 11089 835.74  3.9792  0.047764 *
## - HomeRuns_R            1    273.62 11094 835.82  4.0461  0.045951 *
## - TEAM_BASERUN_SB_SIN   1    416.87 11237 838.27  6.1644  0.014066 *
## - TEAM_FIELDING_DP      1    462.18 11282 839.04  6.8343  0.009796 **
## - TEAM_FIELDING_E       1   1464.88 12285 855.30 21.6615 6.786e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=831.22
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_NEW +
##     TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_3B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_3B_SIN   1      9.18 10839 829.38  0.1365   0.71228
## - TEAM_BATTING_BB_NEW   1     10.54 10841 829.41  0.1567   0.69273
## - TEAM_BATTING_H_SIN    1     13.15 10843 829.45  0.1955   0.65900
## - TEAM_FIELDING_E_SIN   1     19.56 10850 829.57  0.2908   0.59048
## - TEAM_PITCHING_H_SIN   1     23.46 10854 829.64  0.3488   0.55562
## - TEAM_BATTING_2B_SIN   1     33.46 10864 829.81  0.4974   0.48165
## - TEAM_BATTING_2B       1     41.33 10872 829.95  0.6145   0.43426
## - TEAM_BATTING_3B       1     52.93 10883 830.15  0.7869   0.37637
## - TEAM_BASERUN_SB       1     58.84 10889 830.26  0.8748   0.35104
## - TEAM_PITCHING_SO_SIN  1     60.56 10891 830.29  0.9003   0.34411
## - TEAM_PITCHING_HR_SIN  1     75.13 10905 830.54  1.1169   0.29216
## - Hits_R                1     96.19 10926 830.91  1.4299   0.23354
## - TEAM_BATTING_BB_SIN   1     96.86 10927 830.92  1.4399   0.23191
## - TEAM_PITCHING_H       1    112.62 10943 831.20  1.6742   0.19755
## - TEAM_BATTING_H        1    112.84 10943 831.20  1.6774   0.19712
## <none>                             10830 831.22
## - TEAM_BATTING_HR       1    148.85 10979 831.83  2.2128   0.13882
## - TEAM_PITCHING_HR      1    149.70 10980 831.85  2.2254   0.13772
## - Strikeout_R           1    177.01 11007 832.32  2.6315   0.10672
## - TEAM_BATTING_HBP      1    206.11 11036 832.82  3.0640   0.08195 .
## - TEAM_PITCHING_BB      1    207.31 11038 832.84  3.0819   0.08107 .
## - TEAM_BATTING_BB       1    208.14 11038 832.86  3.0941   0.08048 .
```

```
## - TEAM_FIELDING_DP_SIN  1    241.99 11072 833.44  3.5974   0.05966 .
## - TEAM_PITCHING_SO      1    244.66 11075 833.49  3.6371   0.05829 .
## - TEAM_BATTING_SO       1    246.65 11077 833.52  3.6667   0.05728 .
## - Walks_R               1    268.17 11098 833.89  3.9865   0.04755 *
## - HomeRuns_R            1    274.36 11104 834.00  4.0786   0.04509 *
## - TEAM_BASERUN_SB_SIN   1    424.86 11255 836.57  6.3160   0.01295 *
## - TEAM_FIELDING_DP      1    493.47 11324 837.73  7.3359   0.00749 **
## - TEAM_FIELDING_E       1   1459.27 12289 853.37 21.6933 6.659e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=829.38
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_NEW +
##     TEAM_BATTING_H_SIN + TEAM_BATTING_2B_SIN + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_BB_NEW    1     11.14 10850 827.58  0.1665  0.683772
## - TEAM_BATTING_H_SIN     1     13.56 10853 827.62  0.2027  0.653129
## - TEAM_FIELDING_E_SIN    1     17.42 10857 827.69  0.2604  0.610531
## - TEAM_PITCHING_H_SIN    1     22.90 10862 827.79  0.3423  0.559315
## - TEAM_BATTING_2B_SIN    1     31.90 10871 827.95  0.4768  0.490862
## - TEAM_BATTING_2B        1     40.90 10880 828.10  0.6112  0.435464
## - TEAM_BATTING_3B        1     52.37 10892 828.31  0.7827  0.377623
## - TEAM_BASERUN_SB        1     59.05 10898 828.42  0.8825  0.348906
## - TEAM_PITCHING_SO_SIN   1     63.00 10902 828.49  0.9416  0.333309
## - TEAM_PITCHING_HR_SIN   1     71.49 10911 828.64  1.0684  0.302845
## - TEAM_BATTING_BB_SIN    1     93.08 10932 829.02  1.3911  0.239953
## - Hits_R                 1     98.84 10938 829.12  1.4773  0.225971
## <none>                               10839 829.38
## - TEAM_PITCHING_H        1    115.34 10955 829.41  1.7238  0.191063
## - TEAM_BATTING_H         1    115.56 10955 829.41  1.7271  0.190637
## - TEAM_BATTING_HR        1    143.42 10983 829.90  2.1435  0.145117
## - TEAM_PITCHING_HR       1    144.25 10984 829.91  2.1560  0.143957
## - Strikeout_R            1    186.38 11026 830.64  2.7855  0.097053 .
## - TEAM_BATTING_HBP       1    198.10 11037 830.84  2.9608  0.087216 .
## - TEAM_PITCHING_BB       1    209.47 11049 831.04  3.1307  0.078714 .
## - TEAM_BATTING_BB        1    210.26 11050 831.05  3.1424  0.078161 .
## - TEAM_FIELDING_DP_SIN   1    233.94 11073 831.46  3.4964  0.063306 .
## - TEAM_PITCHING_SO       1    250.07 11089 831.74  3.7374  0.054951 .
## - TEAM_BATTING_SO        1    252.04 11091 831.78  3.7669  0.054013 .
## - HomeRuns_R             1    275.33 11115 832.18  4.1150  0.044142 *
## - Walks_R                1    282.22 11122 832.29  4.2179  0.041608 *
## - TEAM_BASERUN_SB_SIN    1    469.85 11309 835.49  7.0221  0.008849 **
## - TEAM_FIELDING_DP       1    493.10 11332 835.88  7.3697  0.007352 **
## - TEAM_FIELDING_E        1   1500.66 12340 852.15 22.4283 4.735e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Step:  AIC=827.58
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_H_SIN +
##     TEAM_BATTING_2B_SIN + TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN +
##     TEAM_FIELDING_E_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_H_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_H_SIN     1     13.09 10864 825.81  0.1966  0.658031
## - TEAM_FIELDING_E_SIN    1     14.27 10865 825.83  0.2144  0.643995
## - TEAM_PITCHING_H_SIN    1     23.42 10874 825.99  0.3519  0.553879
## - TEAM_BATTING_2B_SIN    1     28.30 10879 826.08  0.4251  0.515309
## - TEAM_BATTING_2B        1     40.28 10891 826.29  0.6052  0.437747
## - TEAM_BATTING_3B        1     50.79 10901 826.47  0.7630  0.383679
## - TEAM_BASERUN_SB        1     53.25 10904 826.52  0.8000  0.372422
## - TEAM_PITCHING_SO_SIN   1     58.05 10908 826.60  0.8720  0.351773
## - TEAM_PITCHING_HR_SIN   1     68.80 10919 826.79  1.0335  0.310837
## - Hits_R                 1     97.94 10948 827.30  1.4713  0.226898
## - TEAM_BATTING_BB_SIN    1     99.20 10950 827.32  1.4902  0.223943
## <none>                               10850 827.58
## - TEAM_PITCHING_H        1    114.40 10965 827.58  1.7186  0.191718
## - TEAM_BATTING_H         1    114.63 10965 827.59  1.7220  0.191281
## - TEAM_BATTING_HR        1    142.96 10993 828.08  2.1476  0.144717
## - TEAM_PITCHING_HR       1    143.79 10994 828.10  2.1601  0.143561
## - Strikeout_R            1    184.31 11035 828.80  2.7688  0.098041 .
## - TEAM_PITCHING_BB       1    202.71 11053 829.12  3.0452  0.082861 .
## - TEAM_BATTING_BB        1    204.11 11055 829.14  3.0662  0.081820 .
## - TEAM_BATTING_HBP       1    204.27 11055 829.14  3.0686  0.081699 .
## - TEAM_FIELDING_DP_SIN   1    229.37 11080 829.58  3.4457  0.065222 .
## - TEAM_PITCHING_SO       1    246.35 11097 829.87  3.7008  0.056128 .
## - TEAM_BATTING_SO        1    248.31 11099 829.90  3.7303  0.055170 .
## - HomeRuns_R             1    271.89 11122 830.31  4.0844  0.044916 *
## - Walks_R                1    276.49 11127 830.39  4.1535  0.043166 *
## - TEAM_BASERUN_SB_SIN    1    479.07 11330 833.83  7.1968  0.008056 **
## - TEAM_FIELDING_DP       1    519.87 11370 834.52  7.8096  0.005821 **
## - TEAM_FIELDING_E        1   1490.78 12341 850.17 22.3951 4.787e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=825.81
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_H_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value   Pr(>F)
## - TEAM_PITCHING_H_SIN    1     11.67 10875 824.02  0.1762 0.675206
## - TEAM_FIELDING_E_SIN    1     12.45 10876 824.03  0.1879 0.665202
```

```
## - TEAM_BATTING_2B_SIN   1     29.75 10893 824.33  0.4491 0.503715
## - TEAM_BATTING_2B       1     45.00 10909 824.60  0.6793 0.411015
## - TEAM_PITCHING_SO_SIN  1     53.03 10917 824.74  0.8006 0.372233
## - TEAM_BATTING_3B       1     55.57 10919 824.79  0.8389 0.361063
## - TEAM_BASERUN_SB       1     67.85 10931 825.00  1.0243 0.312983
## - TEAM_PITCHING_HR_SIN  1     82.74 10946 825.26  1.2490 0.265369
## - TEAM_BATTING_BB_SIN   1    100.37 10964 825.57  1.5152 0.220115
## <none>                             10864 825.81
## - Hits_R                1    118.31 10982 825.88  1.7861 0.183258
## - TEAM_PITCHING_H       1    120.37 10984 825.92  1.8172 0.179507
## - TEAM_BATTING_H        1    120.60 10984 825.92  1.8206 0.179100
## - TEAM_BATTING_HR       1    141.71 11005 826.29  2.1392 0.145487
## - TEAM_PITCHING_HR      1    142.53 11006 826.30  2.1517 0.144324
## - TEAM_BATTING_HBP      1    201.59 11065 827.32  3.0433 0.082946 .
## - Strikeout_R           1    218.56 11082 827.62  3.2994 0.071129 .
## - TEAM_PITCHING_BB      1    233.12 11097 827.87  3.5193 0.062436 .
## - TEAM_PITCHING_SO      1    233.41 11097 827.87  3.5236 0.062276 .
## - TEAM_BATTING_BB       1    234.64 11098 827.89  3.5422 0.061599 .
## - TEAM_FIELDING_DP_SIN  1    235.03 11099 827.90  3.5481 0.061384 .
## - TEAM_BATTING_SO       1    235.36 11099 827.91  3.5531 0.061204 .
## - HomeRuns_R            1    302.05 11166 829.05  4.5599 0.034214 *
## - Walks_R               1    314.83 11178 829.27  4.7527 0.030677 *
## - TEAM_BASERUN_SB_SIN   1    507.97 11372 832.54  7.6685 0.006267 **
## - TEAM_FIELDING_DP      1    515.61 11379 832.67  7.7838 0.005897 **
## - TEAM_FIELDING_E       1   1489.54 12353 848.35 22.4866 4.57e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=824.02
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_E_SIN +
##     TEAM_FIELDING_DP_SIN + TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_FIELDING_E_SIN   1     10.89 10886 822.21  0.1652  0.684914
## - TEAM_BATTING_2B_SIN   1     28.78 10904 822.52  0.4367  0.509649
## - TEAM_BATTING_3B       1     48.34 10924 822.86  0.7334  0.393011
## - TEAM_BATTING_2B       1     50.44 10926 822.90  0.7652  0.382971
## - TEAM_PITCHING_SO_SIN  1     52.35 10928 822.93  0.7942  0.374123
## - TEAM_BASERUN_SB       1     67.25 10942 823.19  1.0203  0.313931
## - TEAM_PITCHING_HR_SIN  1     81.58 10957 823.44  1.2377  0.267531
## - TEAM_BATTING_BB_SIN   1    106.43 10982 823.88  1.6147  0.205618
## <none>                             10875 824.02
## - TEAM_PITCHING_H       1    127.05 11002 824.23  1.9277  0.166885
## - TEAM_BATTING_H        1    127.28 11002 824.24  1.9311  0.166514
## - Hits_R                1    130.64 11006 824.30  1.9820  0.161057
## - TEAM_BATTING_HR       1    143.74 11019 824.52  2.1809  0.141641
## - TEAM_PITCHING_HR      1    144.58 11020 824.54  2.1936  0.140489
## - TEAM_BATTING_HBP      1    199.23 11074 825.48  3.0228  0.083967 .
## - Strikeout_R           1    237.11 11112 826.14  3.5974  0.059616 .
```

```
## - TEAM_PITCHING_SO      1    237.83 11113 826.15  3.6084  0.059231 .
## - TEAM_BATTING_SO       1    239.80 11115 826.18  3.6383  0.058201 .
## - TEAM_FIELDING_DP_SIN  1    240.06 11115 826.19  3.6422  0.058070 .
## - TEAM_PITCHING_BB      1    245.95 11121 826.29  3.7316  0.055106 .
## - TEAM_BATTING_BB       1    247.50 11123 826.31  3.7551  0.054354 .
## - HomeRuns_R            1    319.33 11195 827.54  4.8450  0.029115 *
## - Walks_R               1    326.28 11202 827.66  4.9503  0.027442 *
## - TEAM_BASERUN_SB_SIN   1    513.91 11389 830.84  7.7971  0.005851 **
## - TEAM_FIELDING_DP      1    520.12 11395 830.94  7.8913  0.005568 **
## - TEAM_FIELDING_E       1   1559.67 12435 847.61 23.6634 2.661e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=822.21
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_2B_SIN +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_2B_SIN    1     27.70 10914 820.69  0.4224  0.516636
## - TEAM_BATTING_3B        1     46.05 10932 821.01  0.7022  0.403239
## - TEAM_PITCHING_SO_SIN   1     50.50 10937 821.09  0.7701  0.381470
## - TEAM_BATTING_2B        1     57.31 10943 821.21  0.8739  0.351228
## - TEAM_BASERUN_SB        1     65.08 10951 821.35  0.9924  0.320608
## - TEAM_PITCHING_HR_SIN   1     79.17 10965 821.59  1.2073  0.273456
## - TEAM_BATTING_BB_SIN    1    108.77 10995 822.11  1.6587  0.199576
## <none>                              10886 822.21
## - TEAM_PITCHING_H        1    122.00 11008 822.34  1.8603  0.174434
## - TEAM_BATTING_H         1    122.21 11008 822.34  1.8636  0.174061
## - Hits_R                 1    125.66 11012 822.40  1.9161  0.168141
## - TEAM_BATTING_HR        1    139.64 11026 822.64  2.1293  0.146395
## - TEAM_PITCHING_HR       1    140.48 11027 822.66  2.1421  0.145195
## - TEAM_BATTING_HBP       1    201.43 11088 823.71  3.0715  0.081520 .
## - Strikeout_R            1    230.70 11117 824.21  3.5179  0.062467 .
## - TEAM_PITCHING_SO       1    231.98 11118 824.24  3.5374  0.061750 .
## - TEAM_BATTING_SO        1    233.93 11120 824.27  3.5671  0.060678 .
## - TEAM_PITCHING_BB       1    241.52 11128 824.40  3.6829  0.056688 .
## - TEAM_BATTING_BB        1    243.05 11129 824.43  3.7062  0.055919 .
## - TEAM_FIELDING_DP_SIN   1    255.04 11141 824.63  3.8890  0.050266 .
## - HomeRuns_R             1    311.80 11198 825.60  4.7546  0.030628 *
## - Walks_R                1    319.44 11206 825.73  4.8711  0.028682 *
## - TEAM_FIELDING_DP       1    509.92 11396 828.95  7.7756  0.005914 **
## - TEAM_BASERUN_SB_SIN    1    512.07 11398 828.99  7.8085  0.005813 **
## - TEAM_FIELDING_E        1   1567.55 12454 845.90 23.9031 2.375e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=820.69
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_3B +
##     TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB +
```

```
##     TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_3B        1     41.59 10955 819.42  0.6364  0.426135
## - TEAM_BATTING_2B        1     53.07 10967 819.62  0.8120  0.368816
## - TEAM_PITCHING_SO_SIN   1     58.81 10973 819.72  0.8999  0.344184
## - TEAM_BASERUN_SB        1     59.58 10973 819.73  0.9116  0.341064
## - TEAM_PITCHING_HR_SIN   1     86.06 11000 820.19  1.3168  0.252803
## - TEAM_BATTING_BB_SIN    1    104.00 11018 820.50  1.5913  0.208894
## <none>                              10914 820.69
## - TEAM_PITCHING_H        1    115.02 11029 820.70  1.7600  0.186436
## - TEAM_BATTING_H         1    115.24 11029 820.70  1.7633  0.186020
## - Hits_R                 1    118.23 11032 820.75  1.8091  0.180441
## - TEAM_BATTING_HR        1    143.37 11057 821.19  2.1938  0.140451
## - TEAM_PITCHING_HR       1    144.20 11058 821.20  2.2065  0.139315
## - TEAM_BATTING_HBP       1    218.35 11132 822.48  3.3411  0.069352 .
## - Strikeout_R            1    221.23 11135 822.53  3.3852  0.067558 .
## - TEAM_PITCHING_SO       1    223.17 11137 822.56  3.4148  0.066382 .
## - TEAM_BATTING_SO        1    225.08 11139 822.59  3.4441  0.065241 .
## - TEAM_FIELDING_DP_SIN   1    251.40 11165 823.04  3.8468  0.051504 .
## - TEAM_PITCHING_BB       1    251.81 11166 823.05  3.8531  0.051315 .
## - TEAM_BATTING_BB        1    253.38 11167 823.08  3.8771  0.050605 .
## - HomeRuns_R             1    311.49 11225 824.07  4.7663  0.030418 *
## - Walks_R                1    322.12 11236 824.25  4.9290  0.027755 *
## - TEAM_BASERUN_SB_SIN    1    487.01 11401 827.03  7.4521  0.007016 **
## - TEAM_FIELDING_DP       1    490.56 11404 827.09  7.5064  0.006816 **
## - TEAM_FIELDING_E        1   1566.51 12480 844.31 23.9702 2.293e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=819.42
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_HR +
##     TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BASERUN_SB + TEAM_BATTING_HBP +
##     TEAM_PITCHING_H + TEAM_PITCHING_HR + TEAM_PITCHING_BB + TEAM_PITCHING_SO +
##     TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BASERUN_SB        1     47.68 11003 818.25  0.7312  0.393723
## - TEAM_BATTING_2B        1     57.59 11013 818.42  0.8831  0.348709
## - TEAM_PITCHING_SO_SIN   1     60.17 11016 818.47  0.9227  0.338145
## - TEAM_PITCHING_HR_SIN   1     92.17 11048 819.02  1.4134  0.236172
## <none>                              10955 819.42
## - TEAM_PITCHING_H        1    116.63 11072 819.44  1.7884  0.182924
## - TEAM_BATTING_H         1    116.80 11072 819.45  1.7911  0.182604
## - TEAM_BATTING_BB_SIN    1    120.03 11076 819.50  1.8406  0.176704
## - Hits_R                 1    128.94 11084 819.65  1.9772  0.161532
## - TEAM_BATTING_HR        1    155.66 11111 820.11  2.3870  0.124226
## - TEAM_PITCHING_HR       1    156.58 11112 820.13  2.4011  0.123129
## - TEAM_PITCHING_SO       1    247.88 11203 821.69  3.8012  0.052879 .
```

```
## - TEAM_BATTING_SO       1    249.94 11205 821.73  3.8328  0.051916 .
## - TEAM_BATTING_HBP      1    252.71 11208 821.78  3.8752  0.050649 .
## - TEAM_PITCHING_BB      1    254.76 11210 821.81  3.9067  0.049730 *
## - TEAM_BATTING_BB       1    256.37 11212 821.84  3.9313  0.049025 *
## - Strikeout_R           1    265.62 11221 822.00  4.0733  0.045158 *
## - TEAM_FIELDING_DP_SIN  1    271.73 11227 822.10  4.1670  0.042783 *
## - Walks_R               1    317.91 11273 822.88  4.8750  0.028602 *
## - HomeRuns_R            1    318.81 11274 822.90  4.8888  0.028381 *
## - TEAM_BASERUN_SB_SIN   1    481.28 11437 825.63  7.3804  0.007284 **
## - TEAM_FIELDING_DP      1    523.65 11479 826.34  8.0301  0.005165 **
## - TEAM_FIELDING_E       1   1578.65 12534 843.13 24.2084 2.048e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=818.25
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B + TEAM_BATTING_HR +
##     TEAM_BATTING_BB + TEAM_BATTING_SO + TEAM_BATTING_HBP + TEAM_PITCHING_H +
##     TEAM_PITCHING_HR + TEAM_PITCHING_BB + TEAM_PITCHING_SO +
##     TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN + TEAM_PITCHING_HR_SIN +
##     TEAM_PITCHING_SO_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                        Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_BATTING_2B       1     32.55 11036 816.81  0.4999  0.480502
## - TEAM_PITCHING_SO_SIN  1     73.27 11076 817.52  1.1253  0.290288
## - TEAM_PITCHING_HR_SIN  1     89.30 11092 817.79  1.3716  0.243179
## - TEAM_PITCHING_H       1    102.93 11106 818.03  1.5809  0.210365
## - TEAM_BATTING_H        1    103.13 11106 818.03  1.5840  0.209920
## <none>                             11003 818.25
## - Hits_R                1    117.16 11120 818.27  1.7995  0.181574
## - TEAM_BATTING_HR       1    152.47 11156 818.88  2.3419  0.127807
## - TEAM_PITCHING_HR      1    153.33 11156 818.89  2.3551  0.126745
## - TEAM_BATTING_BB_SIN   1    157.01 11160 818.96  2.4116  0.122308
## - TEAM_PITCHING_BB      1    237.85 11241 820.33  3.6533  0.057651 .
## - TEAM_BATTING_BB       1    239.42 11242 820.36  3.6773  0.056844 .
## - TEAM_BATTING_HBP      1    245.84 11249 820.47  3.7759  0.053657 .
## - TEAM_PITCHING_SO      1    259.86 11263 820.71  3.9913  0.047341 *
## - TEAM_BATTING_SO       1    261.94 11265 820.74  4.0233  0.046472 *
## - TEAM_FIELDING_DP_SIN  1    279.27 11282 821.04  4.2894  0.039872 *
## - Strikeout_R           1    279.76 11283 821.04  4.2969  0.039699 *
## - HomeRuns_R            1    292.57 11296 821.26  4.4937  0.035479 *
## - Walks_R               1    297.06 11300 821.34  4.5627  0.034115 *
## - TEAM_BASERUN_SB_SIN   1    492.49 11496 824.61  7.5643  0.006602 **
## - TEAM_FIELDING_DP      1    589.73 11593 826.22  9.0579  0.003016 **
## - TEAM_FIELDING_E       1   1554.91 12558 841.50 23.8823 2.365e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=816.81
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HR + TEAM_BATTING_BB +
##     TEAM_BATTING_SO + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_HR_SIN + TEAM_PITCHING_SO_SIN + Hits_R + Walks_R +
```

```
##     HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_PITCHING_SO_SIN   1     78.91 11115 816.17  1.2155  0.271799
## - TEAM_PITCHING_HR_SIN   1     79.51 11115 816.18  1.2248  0.269983
## - TEAM_PITCHING_H        1     94.40 11130 816.44  1.4542  0.229530
## - TEAM_BATTING_H         1     94.65 11130 816.44  1.4580  0.228929
## - Hits_R                 1    104.92 11141 816.62  1.6162  0.205363
## <none>                               11036 816.81
## - TEAM_BATTING_HR        1    140.66 11176 817.23  2.1668  0.142871
## - TEAM_PITCHING_HR       1    141.48 11177 817.25  2.1795  0.141714
## - TEAM_BATTING_BB_SIN    1    174.75 11210 817.81  2.6920  0.102703
## - TEAM_PITCHING_BB       1    238.21 11274 818.89  3.6695  0.057095 .
## - TEAM_BATTING_BB        1    239.79 11276 818.92  3.6939  0.056283 .
## - TEAM_BATTING_HBP       1    258.51 11294 819.24  3.9822  0.047580 *
## - TEAM_PITCHING_SO       1    269.40 11305 819.42  4.1499  0.043187 *
## - Strikeout_R            1    269.71 11305 819.43  4.1548  0.043066 *
## - TEAM_BATTING_SO        1    271.46 11307 819.45  4.1818  0.042403 *
## - HomeRuns_R             1    281.54 11317 819.62  4.3369  0.038791 *
## - TEAM_FIELDING_DP_SIN   1    285.83 11322 819.70  4.4031  0.037352 *
## - Walks_R                1    302.46 11338 819.98  4.6592  0.032290 *
## - TEAM_BASERUN_SB_SIN    1    478.12 11514 822.91  7.3652  0.007335 **
## - TEAM_FIELDING_DP       1    590.30 11626 824.77  9.0933  0.002959 **
## - TEAM_FIELDING_E        1   1596.05 12632 840.61 24.5866 1.709e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=816.17
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HR + TEAM_BATTING_BB +
##     TEAM_BATTING_SO + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN +
##     TEAM_PITCHING_HR_SIN + Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_PITCHING_HR_SIN   1     65.61 11180 815.30  1.0095  0.316445
## - TEAM_PITCHING_H        1     98.65 11213 815.86  1.5178  0.219645
## - TEAM_BATTING_H         1     98.91 11214 815.87  1.5218  0.219045
## - Hits_R                 1    108.34 11223 816.03  1.6669  0.198420
## <none>                               11115 816.17
## - TEAM_BATTING_HR        1    148.05 11263 816.70  2.2777  0.133090
## - TEAM_PITCHING_HR       1    148.86 11263 816.72  2.2903  0.132032
## - TEAM_BATTING_BB_SIN    1    171.45 11286 817.10  2.6378  0.106187
## - TEAM_BATTING_HBP       1    229.10 11344 818.07  3.5247  0.062164 .
## - TEAM_PITCHING_BB       1    232.56 11347 818.13  3.5781  0.060238 .
## - TEAM_BATTING_BB        1    234.12 11349 818.16  3.6020  0.059393 .
## - TEAM_PITCHING_SO       1    261.16 11376 818.61  4.0180  0.046596 *
## - TEAM_BATTING_SO        1    263.19 11378 818.64  4.0492  0.045763 *
## - Strikeout_R            1    266.19 11381 818.69  4.0953  0.044559 *
## - HomeRuns_R             1    288.13 11403 819.06  4.4329  0.036714 *
## - TEAM_FIELDING_DP_SIN   1    294.50 11409 819.17  4.5310  0.034719 *
## - Walks_R                1    302.47 11417 819.30  4.6535  0.032386 *
## - TEAM_BASERUN_SB_SIN    1    483.45 11598 822.31  7.4380  0.007052 **
## - TEAM_FIELDING_DP       1    558.83 11673 823.54  8.5977  0.003827 **
```

```
## - TEAM_FIELDING_E       1   1524.96 12640 838.73 23.4618 2.839e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=815.3
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HR + TEAM_BATTING_BB +
##     TEAM_BATTING_SO + TEAM_BATTING_HBP + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP +
##     TEAM_BATTING_BB_SIN + TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN +
##     Hits_R + Walks_R + HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value    Pr(>F)
## - TEAM_PITCHING_H        1    104.99 11285 815.08  1.6152  0.205485
## - TEAM_BATTING_H         1    105.24 11285 815.09  1.6190  0.204944
## - Hits_R                 1    114.62 11295 815.25  1.7633  0.185970
## <none>                               11180 815.30
## - TEAM_BATTING_HR        1    150.04 11330 815.84  2.3082  0.130528
## - TEAM_PITCHING_HR       1    150.84 11331 815.86  2.3207  0.129503
## - TEAM_BATTING_BB_SIN    1    166.75 11347 816.13  2.5653  0.111066
## - TEAM_PITCHING_BB       1    246.40 11427 817.46  3.7907  0.053166 .
## - TEAM_BATTING_BB        1    248.02 11428 817.49  3.8156  0.052401 .
## - TEAM_BATTING_HBP       1    253.66 11434 817.58  3.9024  0.049818 *
## - TEAM_FIELDING_DP_SIN   1    272.09 11452 817.89  4.1859  0.042285 *
## - Strikeout_R            1    275.53 11456 817.95  4.2388  0.041017 *
## - Walks_R                1    309.69 11490 818.52  4.7644  0.030409 *
## - HomeRuns_R             1    313.55 11494 818.58  4.8238  0.029407 *
## - TEAM_PITCHING_SO       1    315.61 11496 818.62  4.8554  0.028888 *
## - TEAM_BATTING_SO        1    317.95 11498 818.65  4.8915  0.028307 *
## - TEAM_BASERUN_SB_SIN    1    502.94 11683 821.70  7.7373  0.006013 **
## - TEAM_FIELDING_DP       1    630.08 11810 823.77  9.6933  0.002167 **
## - TEAM_FIELDING_E        1   1496.72 12677 837.30 23.0261 3.451e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=815.08
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HR + TEAM_BATTING_BB +
##     TEAM_BATTING_SO + TEAM_BATTING_HBP + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN + Hits_R + Walks_R +
##     HomeRuns_R + Strikeout_R
##
##                         Df Sum of Sq   RSS    AIC F value   Pr(>F)
## - Hits_R                 1     10.51 11296 813.26  0.1611 0.688656
## - TEAM_BATTING_HR        1     45.46 11331 813.85  0.6969 0.404974
## - TEAM_PITCHING_HR       1     46.21 11331 813.86  0.7084 0.401122
## <none>                               11285 815.08
## - TEAM_PITCHING_BB       1    143.99 11429 815.51  2.2074 0.139170
## - TEAM_BATTING_BB        1    145.50 11431 815.53  2.2306 0.137127
## - TEAM_BATTING_H         1    173.73 11459 816.00  2.6632 0.104513
## - TEAM_BATTING_BB_SIN    1    186.83 11472 816.22  2.8641 0.092374 .
## - Strikeout_R            1    191.48 11477 816.30  2.9354 0.088450 .
## - Walks_R                1    205.54 11491 816.53  3.1509 0.077645 .
## - TEAM_PITCHING_SO       1    239.37 11524 817.09  3.6694 0.057068 .
## - TEAM_BATTING_SO        1    241.56 11527 817.13  3.7031 0.055953 .
```

```
## - TEAM_BATTING_HBP      1    257.19 11542 817.39  3.9427 0.048655 *
## - TEAM_FIELDING_DP_SIN  1    306.87 11592 818.21  4.7043 0.031453 *
## - HomeRuns_R            1    321.36 11606 818.45  4.9264 0.027750 *
## - TEAM_BASERUN_SB_SIN   1    513.49 11799 821.58  7.8717 0.005597 **
## - TEAM_FIELDING_DP      1    592.53 11878 822.86  9.0834 0.002967 **
## - TEAM_FIELDING_E       1   1434.91 12720 835.94 21.9970 5.52e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=813.26
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HR + TEAM_BATTING_BB +
##     TEAM_BATTING_SO + TEAM_BATTING_HBP + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN + Walks_R + HomeRuns_R +
##     Strikeout_R
##
##                        Df Sum of Sq   RSS    AIC F value   Pr(>F)
## - TEAM_BATTING_HR       1     45.33 11341 812.03  0.6982 0.404530
## - TEAM_PITCHING_HR      1     46.08 11342 812.04  0.7098 0.400677
## <none>                              11296 813.26
## - TEAM_PITCHING_BB      1    133.53 11429 813.51  2.0569 0.153315
## - TEAM_BATTING_BB       1    135.03 11431 813.53  2.0800 0.151033
## - TEAM_BATTING_H        1    178.10 11474 814.25  2.7434 0.099458 .
## - TEAM_BATTING_BB_SIN   1    181.81 11478 814.31  2.8006 0.096023 .
## - Walks_R               1    214.12 11510 814.85  3.2983 0.071072 .
## - TEAM_PITCHING_SO      1    240.83 11536 815.29  3.7098 0.055725 .
## - TEAM_BATTING_SO       1    243.04 11539 815.33  3.7438 0.054627 .
## - Strikeout_R           1    249.22 11545 815.43  3.8390 0.051671 .
## - TEAM_BATTING_HBP      1    254.16 11550 815.51  3.9151 0.049431 *
## - TEAM_FIELDING_DP_SIN  1    298.25 11594 816.24  4.5942 0.033468 *
## - HomeRuns_R            1    313.69 11609 816.49  4.8321 0.029255 *
## - TEAM_BASERUN_SB_SIN   1    515.55 11811 819.79  7.9416 0.005391 **
## - TEAM_FIELDING_DP      1    600.38 11896 821.15  9.2484 0.002722 **
## - TEAM_FIELDING_E       1   1431.36 12727 834.05 22.0488 5.37e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=812.03
## TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_BB + TEAM_BATTING_SO +
##     TEAM_BATTING_HBP + TEAM_PITCHING_HR + TEAM_PITCHING_BB +
##     TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP + TEAM_BATTING_BB_SIN +
##     TEAM_BASERUN_SB_SIN + TEAM_FIELDING_DP_SIN + Walks_R + HomeRuns_R +
##     Strikeout_R
##
##                        Df Sum of Sq   RSS    AIC F value    Pr(>F)
## <none>                              11341 812.03
## - TEAM_PITCHING_BB      1    122.66 11464 812.08  1.8927 0.1706516
## - TEAM_BATTING_BB       1    124.09 11465 812.10  1.9149 0.1681859
## - TEAM_BATTING_BB_SIN   1    192.94 11534 813.25  2.9772 0.0862118 .
## - TEAM_BATTING_H        1    193.12 11534 813.25  2.9800 0.0860649 .
## - TEAM_PITCHING_SO      1    233.73 11575 813.92  3.6066 0.0591955 .
## - TEAM_BATTING_SO       1    235.91 11577 813.96  3.6402 0.0580378 .
## - TEAM_BATTING_HBP      1    263.66 11605 814.42  4.0684 0.0452199 *
## - HomeRuns_R            1    268.40 11609 814.49  4.1416 0.0433498 *
```

```
## - Walks_R               1    310.05 11651 815.18  4.7843 0.0300457 *
## - Strikeout_R           1    314.04 11655 815.24  4.8459 0.0290201 *
## - TEAM_FIELDING_DP_SIN  1    320.27 11661 815.35  4.9421 0.0274913 *
## - TEAM_BASERUN_SB_SIN   1    533.18 11874 818.80  8.2274 0.0046341 **
## - TEAM_FIELDING_DP      1    622.14 11963 820.23  9.6001 0.0022673 **
## - TEAM_PITCHING_HR      1    788.55 12130 822.87 12.1680 0.0006149 ***
## - TEAM_FIELDING_E       1   1449.87 12791 833.00 22.3725 4.607e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coefficients(step4)
```

```
##          (Intercept)       TEAM_BATTING_H      TEAM_BATTING_BB
##        -1.818043e+03         1.733817e-02         8.603592e+00
##      TEAM_BATTING_SO     TEAM_BATTING_HBP     TEAM_PITCHING_HR
##        -6.535137e+00         9.492188e-02         8.312811e-02
##     TEAM_PITCHING_BB     TEAM_PITCHING_SO      TEAM_FIELDING_E
##        -8.550658e+00         6.500379e+00        -1.810601e-01
##     TEAM_FIELDING_DP  TEAM_BATTING_BB_SIN  TEAM_BASERUN_SB_SIN
##        -1.069442e-01        -1.473778e+00        -2.441250e+00
## TEAM_FIELDING_DP_SIN              Walks_R           HomeRuns_R
##         1.896564e+00        -1.474774e+04         4.845243e+03
##          Strikeout_R
##         1.179872e+04
```

Discuss the coefficients in the models, do they make sense? For example, if a team hits a lot of Home Runs, it would be reasonably expected that such a team would win more games. However, if the coefficient is negative (suggesting that the team would lose more games), then that needs to be discussed. Are you keeping the model even though it is counter intuitive? Why? The boss needs to know.

Select Models

Decide on the criteria for selecting the best multiple linear regression model. Will you select a model with slightly worse performance if it makes more sense or is more parsimonious? Discuss why you selected your model. For the multiple linear regression model, will you use a metric such as Adjusted R2, RMSE, etc.? Be sure to explain how you can make inferences from the model, discuss multi-collinearity issues (if any), and discuss other relevant model output. Using the training data set, evaluate the multiple linear regression model based on (a) mean squared error, (b) R2, (c) F-statistic, and (d) residual plots. Make predictions using the evaluation data set.

Model One with original data

```
library(car)
```

```
## Warning: package 'car' was built under R version 3.2.4
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:psych':
##
##     logit
```

```
mod1<- lm(TARGET_WINS ~
     TEAM_BATTING_H +
     TEAM_BATTING_2B +
     TEAM_BATTING_3B +
     TEAM_BATTING_HR +
     TEAM_BATTING_BB +
     TEAM_BATTING_HBP +
     TEAM_BATTING_SO +
     TEAM_BASERUN_SB +
     TEAM_BASERUN_CS +
     TEAM_FIELDING_E +
     TEAM_FIELDING_DP +
     TEAM_PITCHING_BB +
     TEAM_PITCHING_H +
     TEAM_PITCHING_HR +
     TEAM_PITCHING_SO, moneyball2
   )

summary(mod1)
```

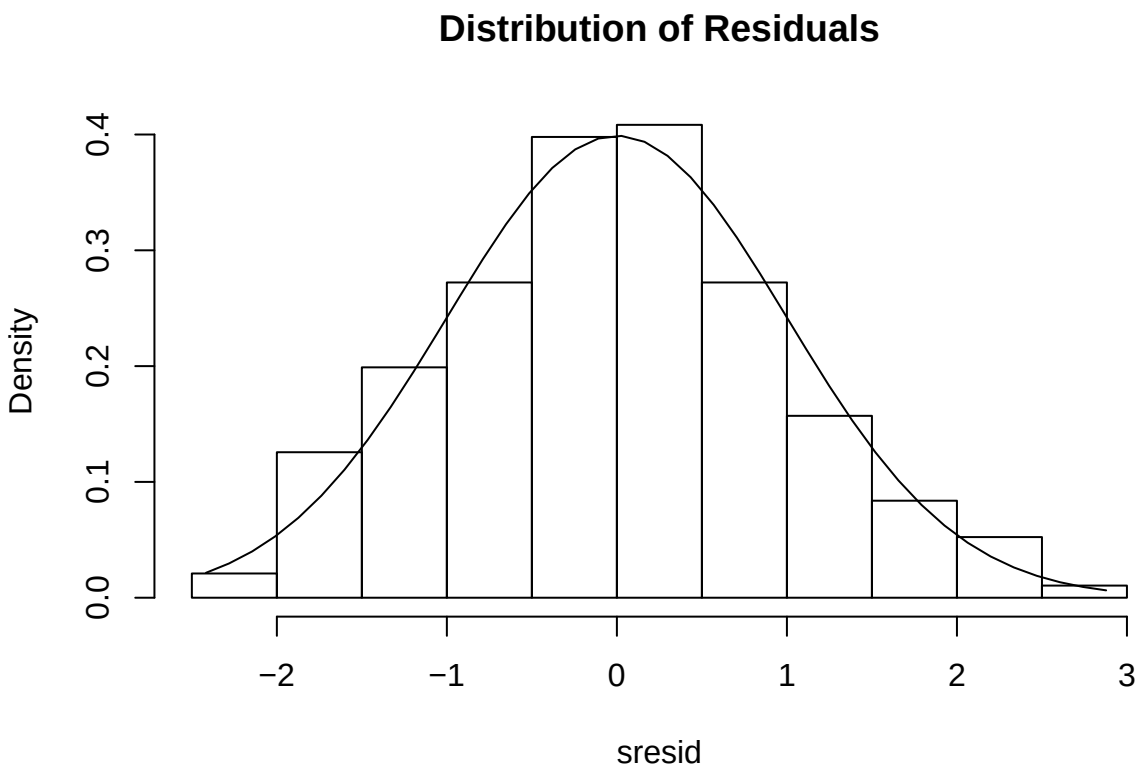
```
##
## Call:
## lm(formula = TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B +
##     TEAM_BATTING_3B + TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_BATTING_HBP +
##     TEAM_BATTING_SO + TEAM_BASERUN_SB + TEAM_BASERUN_CS + TEAM_FIELDING_E +
##     TEAM_FIELDING_DP + TEAM_PITCHING_BB + TEAM_PITCHING_H + TEAM_PITCHING_HR +
##     TEAM_PITCHING_SO, data = moneyball2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -19.8708  -5.6564  -0.0599   5.2545  22.9274
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      60.28826   19.67842   3.064  0.00253 **
## TEAM_BATTING_H    1.91348    2.76139   0.693  0.48927
## TEAM_BATTING_2B   0.02639    0.03029   0.871  0.38484
## TEAM_BATTING_3B  -0.10118    0.07751  -1.305  0.19348
## TEAM_BATTING_HR  -4.84371   10.50851  -0.461  0.64542
## TEAM_BATTING_BB  -4.45969    3.63624  -1.226  0.22167
## TEAM_BATTING_HBP  0.08247    0.04960   1.663  0.09815 .
## TEAM_BATTING_SO   0.34196    2.59876   0.132  0.89546
## TEAM_BASERUN_SB   0.03304    0.02867   1.152  0.25071
## TEAM_BASERUN_CS  -0.01104    0.07143  -0.155  0.87730
## TEAM_FIELDING_E  -0.17204    0.04140  -4.155 5.08e-05 ***
## TEAM_FIELDING_DP -0.10819    0.03654  -2.961  0.00349 **
## TEAM_PITCHING_BB  4.51089    3.63372   1.241  0.21612
## TEAM_PITCHING_H  -1.89096    2.76095  -0.685  0.49432
## TEAM_PITCHING_HR  4.93043   10.50664   0.469  0.63946
## TEAM_PITCHING_SO -0.37364    2.59705  -0.144  0.88577
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.467 on 175 degrees of freedom
##   (2085 observations deleted due to missingness)
## Multiple R-squared:  0.5501, Adjusted R-squared:  0.5116
## F-statistic: 14.27 on 15 and 175 DF,  p-value: < 2.2e-16
```

```
#library(faraway)
#sumary(mod1)
```

Normality check of Residuals

```
#  First let plot residuals to see if they look like a normal distribution:

library(MASS)
sresid <- studres(mod1)
hist(sresid, freq=FALSE,
     main="Distribution of Residuals")
xfit<-seq(min(sresid),max(sresid),length=40)
yfit<-dnorm(xfit)
lines(xfit, yfit)
```



The residuals are normally distributed, this indicates That the mean of the difference between our predictions

and the actual values is close to 0 which is good for our analysis.

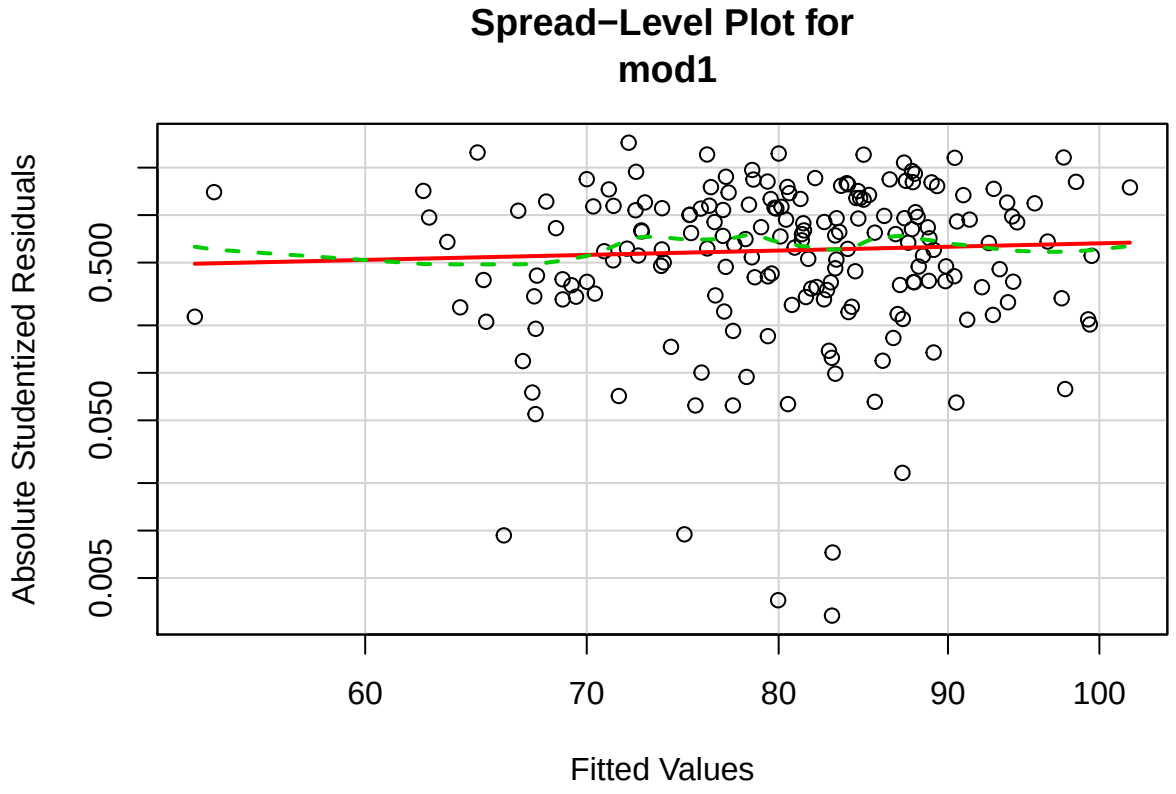
Also, it's unlikely that no relationship exists between TEAM_FIELDING_E and TARGET_WINS.

homoscedasticity check or non-constant error variance test

```
# Evaluate homoscedasticity
ncvTest(mod1)
```

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 0.03994848    Df = 1     p = 0.8415813
```

```
# plot studentized residuals vs. fitted values
spreadLevelPlot(mod1)
```




```
##
## Suggested power transformation:  0.5267026
```

The test confirms the non-constant error variance test. It also has a p-value higher than a significance level of 0.05.

Therefore we can accept the null hypothesis that the variance of the residuals is constant and infer that heteroscedasticity is not present.

Collinearity Check

```
# Evaluate Collinearity
vif(mod1) # variance inflation factors
```

```
##   TEAM_BATTING_H  TEAM_BATTING_2B  TEAM_BATTING_3B  TEAM_BATTING_HR
##     1.171824e+05     1.685623e+00     1.302198e+00     3.074804e+05
##  TEAM_BATTING_BB TEAM_BATTING_HBP  TEAM_BATTING_SO  TEAM_BASERUN_SB
##     1.962853e+05     1.096334e+00     1.941752e+05     1.950069e+00
##  TEAM_BASERUN_CS  TEAM_FIELDING_E TEAM_FIELDING_DP TEAM_PITCHING_BB
##     1.914415e+00     1.256819e+00     1.097611e+00     1.964039e+05
##  TEAM_PITCHING_H TEAM_PITCHING_HR TEAM_PITCHING_SO
##     1.160417e+05     3.069624e+05     1.946316e+05
```

```
Collinearity<- sqrt(vif(mod1)) > 3 # 3 problem?
data.frame(Collinearity)
```

```
##                  Collinearity
## TEAM_BATTING_H           TRUE
## TEAM_BATTING_2B         FALSE
## TEAM_BATTING_3B         FALSE
## TEAM_BATTING_HR          TRUE
## TEAM_BATTING_BB          TRUE
## TEAM_BATTING_HBP        FALSE
## TEAM_BATTING_SO          TRUE
## TEAM_BASERUN_SB         FALSE
## TEAM_BASERUN_CS         FALSE
## TEAM_FIELDING_E         FALSE
## TEAM_FIELDING_DP        FALSE
## TEAM_PITCHING_BB         TRUE
## TEAM_PITCHING_H          TRUE
## TEAM_PITCHING_HR         TRUE
## TEAM_PITCHING_SO         TRUE
```

Test for Autocorrelated Errors

durbinWatsonTest(mod1)

```
durbinWatsonTest(mod1)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1       0.2128921      1.567453       0
##  Alternative hypothesis: rho != 0
```

goodness of fit of your model

using R-squared and adjusted R-squared, our model is about 55% predicts the TARGET_WINS