621 MoneyBall

Brandon Chung, Jiaxin Zheng, Andreina Arias, and Stephanie Chiang

2025-09-28

Introduction

In this homework assignment we will explore, analyze and model a data set containing 2276 professional baseball team records from the years 1871 to 2006. Our objective is to build a multiple linear regression model on the given training data to predict the number of wins for each team in the test data.

Table 1: Variable Definitions and Theoretical Effects on Wins

Variable_Name	Definition	Theoretical_Effect
INDEX	Identification variable (do not use)	None
TARGET_WINS	Number of wins	—
TEAM_BATTING_H	Base hits (1B, 2B, 3B, HR)	Positive
TEAM_BATTING_2B	Doubles	Positive
TEAM_BATTING_3B	Triples	Positive
TEAM_BATTING_HR	Homeruns	Positive
TEAM_BATTING_BB	Walks	Positive
TEAM_BATTING_HBP	Hit by pitch	Positive
TEAM_BATTING_SO	Strikeouts	Negative
TEAM_BASERUN_SB	Stolen bases	Positive
TEAM_BASERUN_CS	Caught stealing	Negative
TEAM_FIELDING_E	Errors	Negative
TEAM_FIELDING_DP	Double plays	Positive
TEAM_PITCHING_BB	Walks allowed	Negative
TEAM_PITCHING_H	Hits allowed	Negative
TEAM_PITCHING_HR	Homeruns allowed	Negative
TEAM_PITCHING_SO	Strikeouts by pitchers	Positive

Data Exploration

Data Summary

The moneyball training data set contains 16 variables, excluding the index, and 2,276 observations. Each observational unit represents a single team's statistics for that year's performance. There are 15 predictor variables which are counts of various actions in baseball such as base hits, home runs, strikeouts, stolen bases, caught stealing, hits allows and more.

As seen below in our numerical summary, the data contains NA values in certain variables (TEAM_BATTING_SO, TEAM_BASERUN_SB, TEAM_BASERUN_CS, TEAM_BATTING_HBP, TEAM_PITCHING_SO, and TEAM_FIELDING_DP). These NA values will be addressed in the data preparation. In addition,

TEAM_BATTING_HBP contains a large amount of NAs at a count of 2085. There is also certain variables with max values that deviate significantly from the interquartile ranges such as TEAM_PITCHING_H and TEAM_PITCHING_SO.

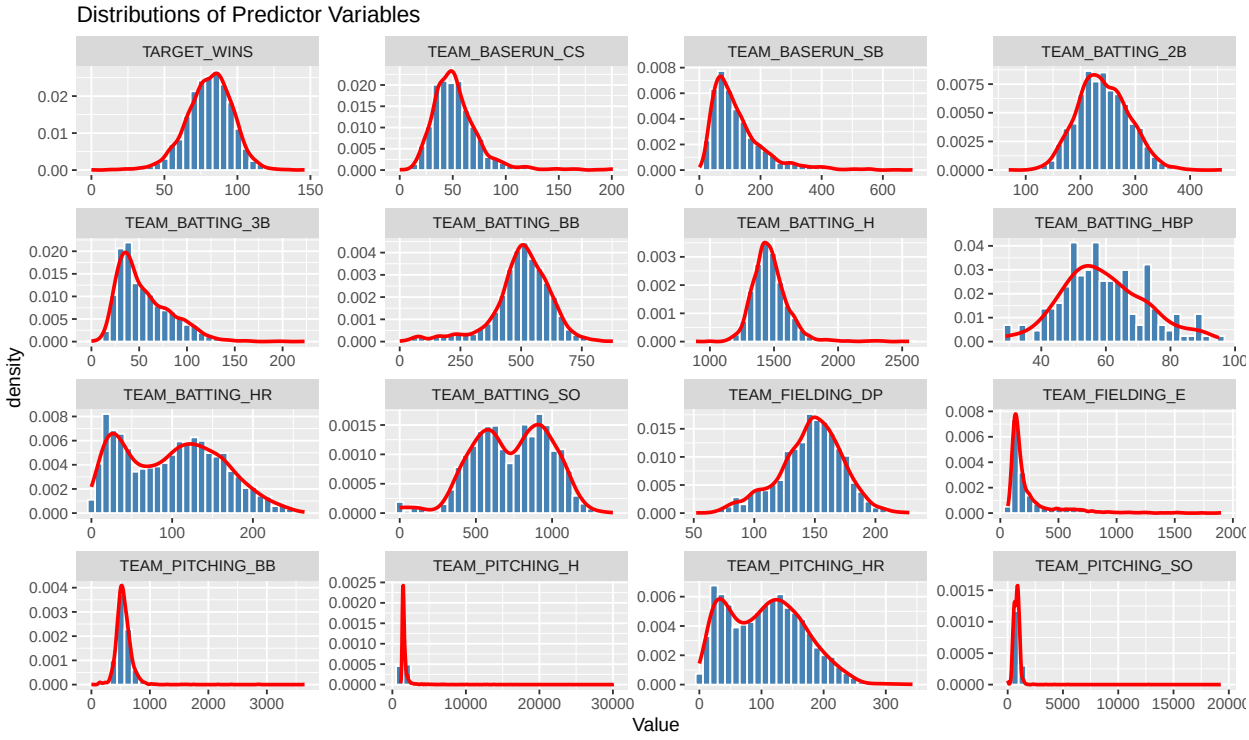
```
## Rows: 2,276
## Columns: 16
## $ TARGET_WINS      <int> 39, 70, 86, 70, 82, 75, 80, 85, 86, 76, 78, 68, 72, 7~
## $ TEAM_BATTING_H   <int> 1445, 1339, 1377, 1387, 1297, 1279, 1244, 1273, 1391,~
## $ TEAM_BATTING_2B  <int> 194, 219, 232, 209, 186, 200, 179, 171, 197, 213, 179~
## $ TEAM_BATTING_3B  <int> 39, 22, 35, 38, 27, 36, 54, 37, 40, 18, 27, 31, 41, 2~
## $ TEAM_BATTING_HR  <int> 13, 190, 137, 96, 102, 92, 122, 115, 114, 96, 82, 95,~
## $ TEAM_BATTING_BB  <int> 143, 685, 602, 451, 472, 443, 525, 456, 447, 441, 374~
## $ TEAM_BATTING_SO  <int> 842, 1075, 917, 922, 920, 973, 1062, 1027, 922, 827, ~
## $ TEAM_BASERUN_SB  <int> NA, 37, 46, 43, 49, 107, 80, 40, 69, 72, 60, 119, 221~
## $ TEAM_BASERUN_CS  <int> NA, 28, 27, 30, 39, 59, 54, 36, 27, 34, 39, 79, 109, ~
## $ TEAM_BATTING_HBP <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ TEAM_PITCHING_H  <int> 9364, 1347, 1377, 1396, 1297, 1279, 1244, 1281, 1391,~
## $ TEAM_PITCHING_HR <int> 84, 191, 137, 97, 102, 92, 122, 116, 114, 96, 86, 95,~
## $ TEAM_PITCHING_BB <int> 927, 689, 602, 454, 472, 443, 525, 459, 447, 441, 391~
## $ TEAM_PITCHING_SO <int> 5456, 1082, 917, 928, 920, 973, 1062, 1033, 922, 827,~
## $ TEAM_FIELDING_E  <int> 1011, 193, 175, 164, 138, 123, 136, 112, 127, 131, 11~
## $ TEAM_FIELDING_DP <int> NA, 155, 153, 156, 168, 149, 186, 136, 169, 159, 141,~
```

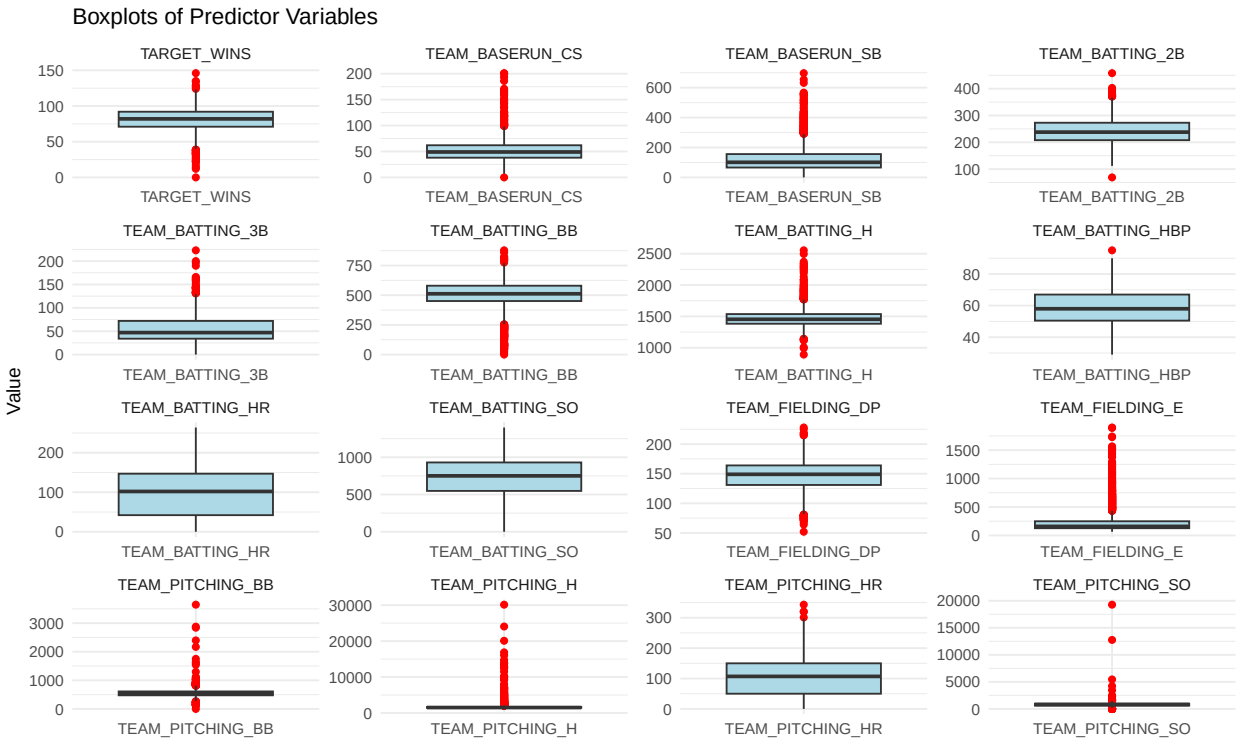
```
##      TARGET_WINS   TEAM_BATTING_H  TEAM_BATTING_2B  TEAM_BATTING_3B
##                0                0                0                0
##  TEAM_BATTING_HR  TEAM_BATTING_BB  TEAM_BATTING_SO  TEAM_BASERUN_SB
##                0                0              102              131
##  TEAM_BASERUN_CS TEAM_BATTING_HBP  TEAM_PITCHING_H TEAM_PITCHING_HR
##              772             2085                0                0
## TEAM_PITCHING_BB TEAM_PITCHING_SO  TEAM_FIELDING_E TEAM_FIELDING_DP
##                0              102                0              286
```

```
##   TARGET_WINS     TEAM_BATTING_H TEAM_BATTING_2B TEAM_BATTING_3B
##  Min.   :  0.00   Min.   : 891   Min.   : 69.0   Min.   :  0.00
##  1st Qu.: 71.00   1st Qu.:1383   1st Qu.:208.0   1st Qu.: 34.00
##  Median : 82.00   Median :1454   Median :238.0   Median : 47.00
##  Mean   : 80.79   Mean   :1469   Mean   :241.2   Mean   : 55.25
##  3rd Qu.: 92.00   3rd Qu.:1537   3rd Qu.:273.0   3rd Qu.: 72.00
##  Max.   :146.00   Max.   :2554   Max.   :458.0   Max.   :223.00
##
##  TEAM_BATTING_HR  TEAM_BATTING_BB TEAM_BATTING_SO  TEAM_BASERUN_SB
##  Min.   :  0.00   Min.   :  0.0   Min.   :   0.0   Min.   :  0.0
##  1st Qu.: 42.00   1st Qu.:451.0   1st Qu.: 548.0   1st Qu.: 66.0
##  Median :102.00   Median :512.0   Median : 750.0   Median :101.0
##  Mean   : 99.61   Mean   :501.6   Mean   : 735.6   Mean   :124.8
##  3rd Qu.:147.00   3rd Qu.:580.0   3rd Qu.: 930.0   3rd Qu.:156.0
##  Max.   :264.00   Max.   :878.0   Max.   :1399.0   Max.   :697.0
##                                   NA's   :102      NA's   :131
##  TEAM_BASERUN_CS TEAM_BATTING_HBP TEAM_PITCHING_H TEAM_PITCHING_HR
##  Min.   :  0.0   Min.   :29.00    Min.   : 1137   Min.   :  0.0
##  1st Qu.: 38.0   1st Qu.:50.50    1st Qu.: 1419   1st Qu.: 50.0
##  Median : 49.0   Median :58.00    Median : 1518   Median :107.0
##  Mean   : 52.8   Mean   :59.36    Mean   : 1779   Mean   :105.7
```

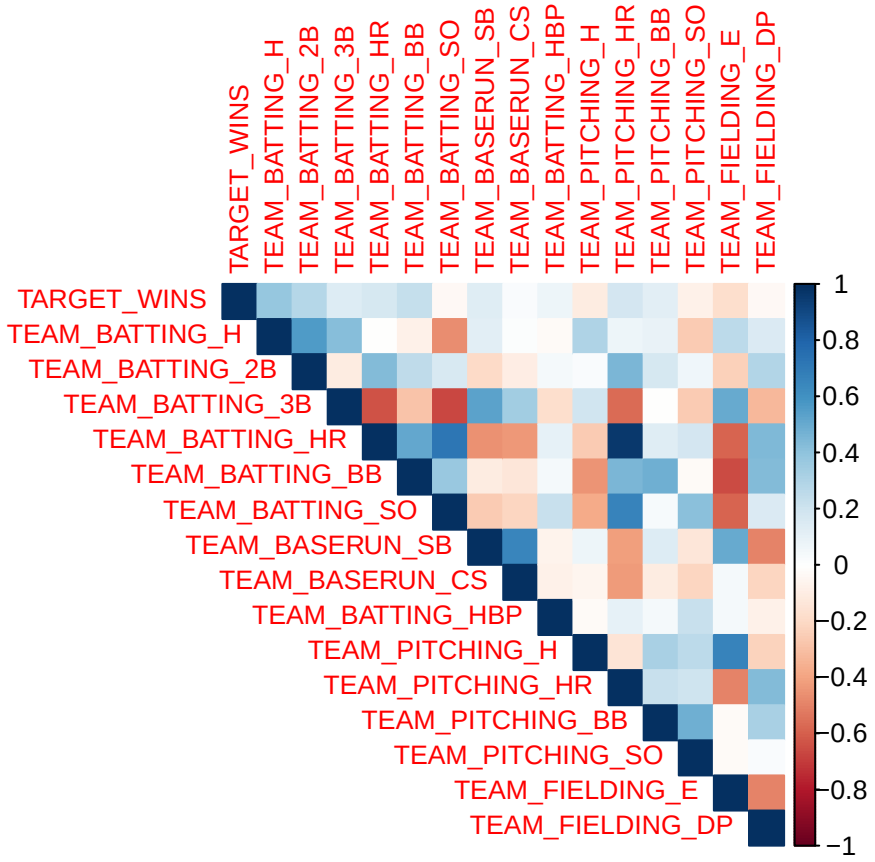
```
##  3rd Qu.: 62.0   3rd Qu.:67.00    3rd Qu.: 1682   3rd Qu.:150.0
##  Max.   :201.0   Max.   :95.00    Max.   :30132   Max.   :343.0
##  NA's   :772     NA's   :2085
##  TEAM_PITCHING_BB TEAM_PITCHING_SO  TEAM_FIELDING_E  TEAM_FIELDING_DP
##  Min.   :   0.0   Min.   :    0.0   Min.   :  65.0   Min.   : 52.0
##  1st Qu.: 476.0   1st Qu.:  615.0   1st Qu.: 127.0   1st Qu.:131.0
##  Median : 536.5   Median :  813.5   Median : 159.0   Median :149.0
##  Mean   : 553.0   Mean   :  817.7   Mean   : 246.5   Mean   :146.4
##  3rd Qu.: 611.0   3rd Qu.:  968.0   3rd Qu.: 249.2   3rd Qu.:164.0
##  Max.   :3645.0   Max.   :19278.0   Max.   :1898.0   Max.   :228.0
##                   NA's   :102                        NA's   :286
```

Data Visualizations





The histogram and box plots above provide a better understanding of the distribution of our predictor variables. Most variables have a relatively normal distribution where others show strong left and right side skewing. The box plots also clue us into possible data entry errors as may be the case for TEAM_PITCHING_SO.

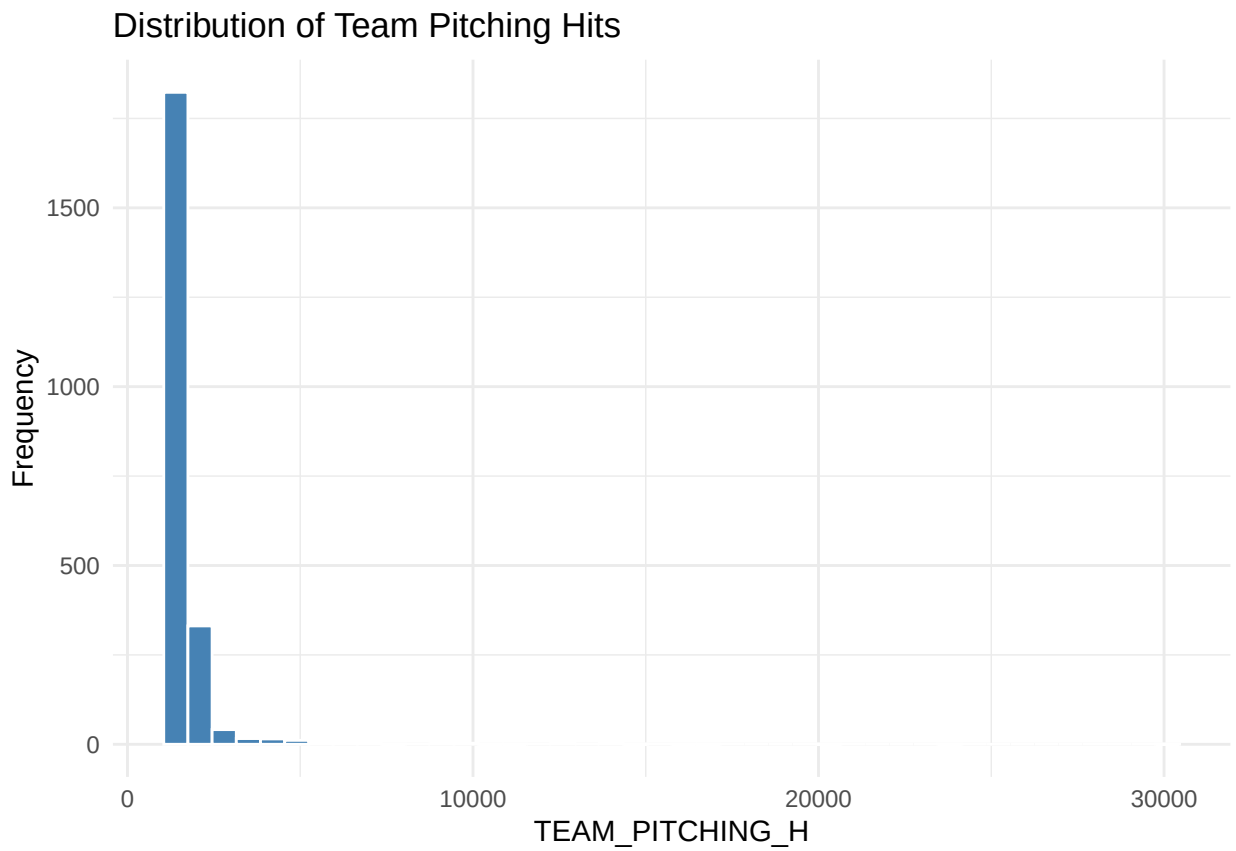


The correlation heat map helps us to see the relationship of variables against the target variable and other predictors. Correlations are mostly what was expected based on the theoretical effect given in the introduction with some exceptions. An example of this can be seen with TEAM_BASERUN_CS where the correlation is slightly positive (0.02240407) when the theoretical effect is to have a negative impact on wins.

Diving deeper into the outliers for the TEAM_PITCHING_SO (pitchers striking out the opposing team's hitter) variable we can see that the record for these teams also are paired with a 0 TEAM_PITCHING_HR (home runs allowed by the pitchers), and so it stand to reason that these outliers are not data errors.

```
##   TARGET_WINS TEAM_BATTING_H TEAM_BATTING_2B TEAM_BATTING_3B TEAM_BATTING_HR
## 1          41            992             263              20               0
## 2         108           1188             338               0               0
##   TEAM_BATTING_BB TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS
## 1             142             952              NA              NA
## 2             270             945              NA              NA
##   TEAM_BATTING_HBP TEAM_PITCHING_H TEAM_PITCHING_HR TEAM_PITCHING_BB
## 1               NA           20088                0             2876
## 2               NA           16038                0             3645
##   TEAM_PITCHING_SO TEAM_FIELDING_E TEAM_FIELDING_DP
## 1            19278             952              NA
## 2            12758             716              NA
```

For the outliers in TEAM_PITCHING_H (hits allowed by pitchers) our distribution shows us that the outliers are likely not data errors either. There are infrequent but other recorded values between our outliers and the IQR of our variable. Our outliers in this variable are plausible real recorded values that happen to fall far on our distribution's right sided tail.



Data Preparation

The variable TEAM_BATTING_HBP which represents a batter being hit by a pitch was removed as the influence is a factor outside of the batter's controls and it's not a repeatable skill. The variable also contained 2,085 NA values out of the total of 2,276 observations.

```
## 'data.frame':    2276 obs. of  15 variables:
##  $ TARGET_WINS     : int  39 70 86 70 82 75 80 85 86 76 ...
##  $ TEAM_BATTING_H  : int  1445 1339 1377 1387 1297 1279 1244 1273 1391 1271 ...
##  $ TEAM_BATTING_2B : int  194 219 232 209 186 200 179 171 197 213 ...
##  $ TEAM_BATTING_3B : int  39 22 35 38 27 36 54 37 40 18 ...
##  $ TEAM_BATTING_HR : int  13 190 137 96 102 92 122 115 114 96 ...
##  $ TEAM_BATTING_BB : int  143 685 602 451 472 443 525 456 447 441 ...
##  $ TEAM_BATTING_SO : int  842 1075 917 922 920 973 1062 1027 922 827 ...
##  $ TEAM_BASERUN_SB : int  NA 37 46 43 49 107 80 40 69 72 ...
##  $ TEAM_BASERUN_CS : int  NA 28 27 30 39 59 54 36 27 34 ...
##  $ TEAM_PITCHING_H : int  9364 1347 1377 1396 1297 1279 1244 1281 1391 1271 ...
##  $ TEAM_PITCHING_HR: int  84 191 137 97 102 92 122 116 114 96 ...
##  $ TEAM_PITCHING_BB: int  927 689 602 454 472 443 525 459 447 441 ...
##  $ TEAM_PITCHING_SO: int  5456 1082 917 928 920 973 1062 1033 922 827 ...
##  $ TEAM_FIELDING_E : int  1011 193 175 164 138 123 136 112 127 131 ...
##  $ TEAM_FIELDING_DP: int  NA 155 153 156 168 149 186 136 169 159 ...
```

Near zero variance variables are variables with observed values that barely change across observations.

Because of this they contribute little to analysis and introduce unnecessary complexity along with multicollinearity risk. No variables were found to be near zero variance as seen below.

```
##                  freqRatio percentUnique zeroVar   nzv
## TARGET_WINS       1.014493      4.745167   FALSE FALSE
## TEAM_BATTING_H    1.333333     25.000000   FALSE FALSE
## TEAM_BATTING_2B   1.000000     10.544815   FALSE FALSE
## TEAM_BATTING_3B   1.074074      6.326889   FALSE FALSE
## TEAM_BATTING_HR   1.038462     10.676626   FALSE FALSE
## TEAM_BATTING_BB   1.058824     23.418278   FALSE FALSE
## TEAM_BATTING_SO   2.000000     36.115993   FALSE FALSE
## TEAM_BASERUN_SB   1.000000     15.289982   FALSE FALSE
## TEAM_BASERUN_CS   1.125000      5.623902   FALSE FALSE
## TEAM_BATTING_HBP  1.000000      2.416520   FALSE FALSE
## TEAM_PITCHING_H   1.083333     37.038664   FALSE FALSE
## TEAM_PITCHING_HR  1.038462     11.247803   FALSE FALSE
## TEAM_PITCHING_BB  1.066667     23.506151   FALSE FALSE
## TEAM_PITCHING_SO  2.222222     36.159930   FALSE FALSE
## TEAM_FIELDING_E   1.037037     24.121265   FALSE FALSE
## TEAM_FIELDING_DP  1.000000      6.326889   FALSE FALSE
```

For data imputation we looked at the columns with missing values and we used a multiple imputation approach to impute the missing data using MICE predictive mean matching method those columns, as the variable had a high percentage of missing data and the missingness is not completely at random. MICE imputation would be best to preserve the varianc and the correlation structure of the data, a simple imputation wouldn't be ideal as it it would impute with a fixed number leading to an underestimates variance, the simple imputation would also be bias ignoring the relationships between the variables.

```
##      TARGET_WINS   TEAM_BATTING_H  TEAM_BATTING_2B  TEAM_BATTING_3B
##         0.000000         0.000000         0.000000         0.000000
##  TEAM_BATTING_HR  TEAM_BATTING_BB  TEAM_BATTING_SO  TEAM_BASERUN_SB
##         0.000000         0.000000         4.481547         5.755712
##  TEAM_BASERUN_CS  TEAM_PITCHING_H TEAM_PITCHING_HR TEAM_PITCHING_BB
##        33.919156         0.000000         0.000000         0.000000
## TEAM_PITCHING_SO  TEAM_FIELDING_E TEAM_FIELDING_DP
##         4.481547         0.000000        12.565905
```

Used multiple imputation to impute the missing data using MICE predictive mean matching method.

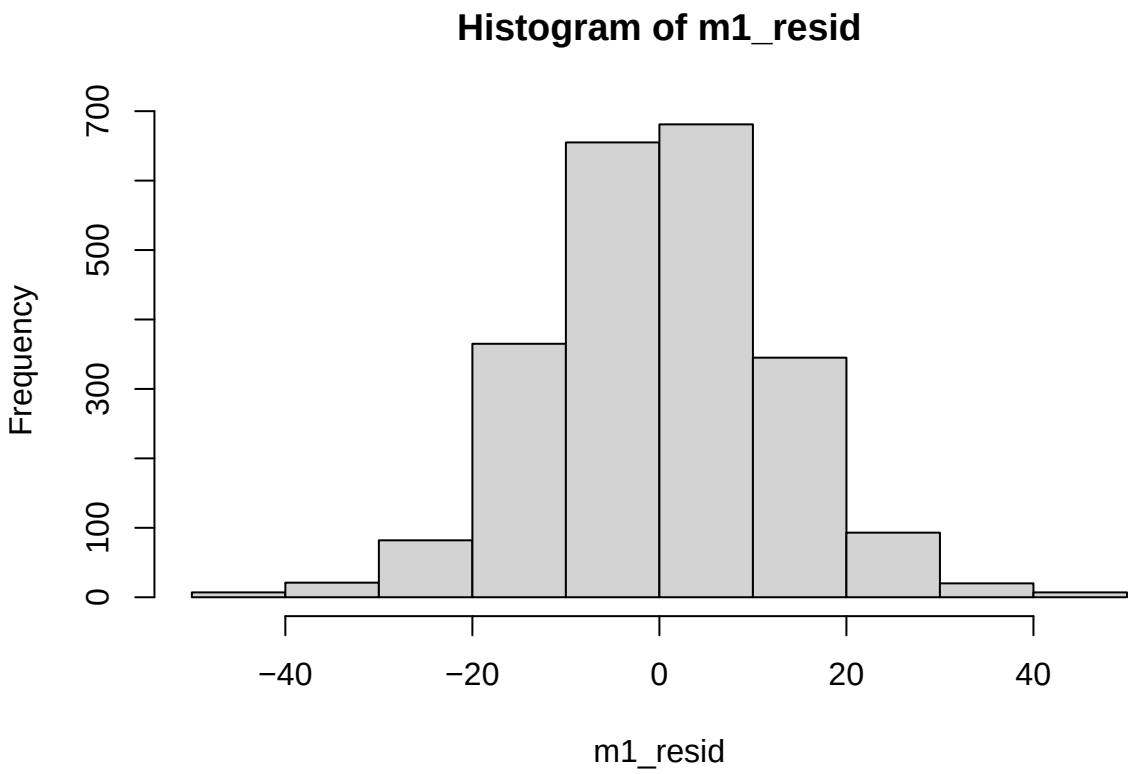
Multiple Linear Regression Models

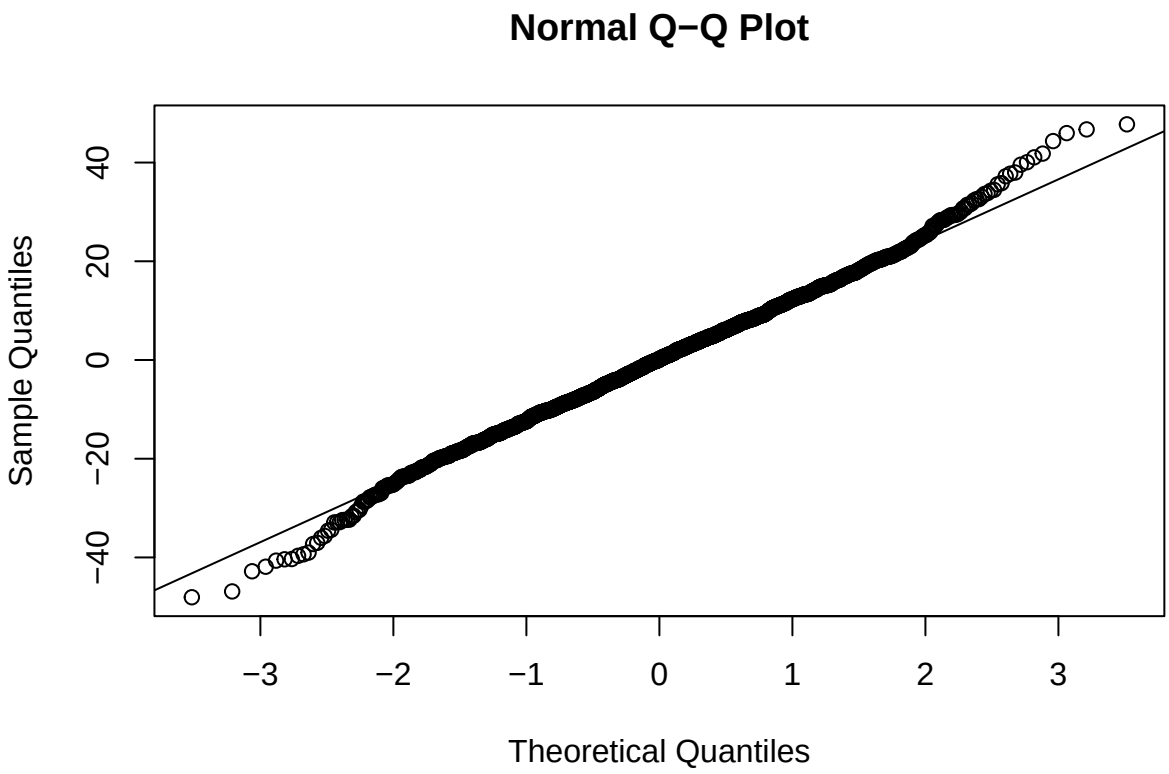
Model 1: All Features

For the first model we choose to include all the predictive variables. This will allow us to see which features have significant influence on our TARGET_WINS dependent variable.

```
##
## Call:
## lm(formula = 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_PITCHING_H + TEAM_PITCHING_HR +
```

```
##     TEAM_PITCHING_BB + TEAM_PITCHING_SO + TEAM_FIELDING_E + TEAM_FIELDING_DP,
##     data = Training_imp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -48.066  -8.413   0.173   8.114  47.738
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)      33.6652346  5.1731357   6.508 9.37e-11 ***
## TEAM_BATTING_H    0.0431257  0.0035895  12.014  < 2e-16 ***
## TEAM_BATTING_2B  -0.0199054  0.0088954  -2.238 0.025337 *
## TEAM_BATTING_3B   0.0412403  0.0164442   2.508 0.012215 *
## TEAM_BATTING_HR   0.0576471  0.0265424   2.172 0.029968 *
## TEAM_BATTING_BB   0.0130473  0.0056243   2.320 0.020440 *
## TEAM_BATTING_SO  -0.0150600  0.0024780  -6.077 1.43e-09 ***
## TEAM_BASERUN_SB   0.0494468  0.0054066   9.146  < 2e-16 ***
## TEAM_BASERUN_CS   0.0020950  0.0110596   0.189 0.849777
## TEAM_PITCHING_H   0.0013758  0.0003859   3.566 0.000371 ***
## TEAM_PITCHING_HR  0.0236405  0.0235842   1.002 0.316263
## TEAM_PITCHING_BB -0.0036554  0.0040041  -0.913 0.361385
## TEAM_PITCHING_SO  0.0015600  0.0008943   1.744 0.081220 .
## TEAM_FIELDING_E  -0.0415048  0.0027079 -15.327  < 2e-16 ***
## TEAM_FIELDING_DP -0.1119556  0.0124114  -9.020  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.66 on 2261 degrees of freedom
## Multiple R-squared:  0.358,  Adjusted R-squared:  0.354
## F-statistic: 90.06 on 14 and 2261 DF,  p-value: < 2.2e-16
```


Histogram of m1_resid
Frequency
0
100
300
500
700
−40
−20
0
20
40
m1_resid

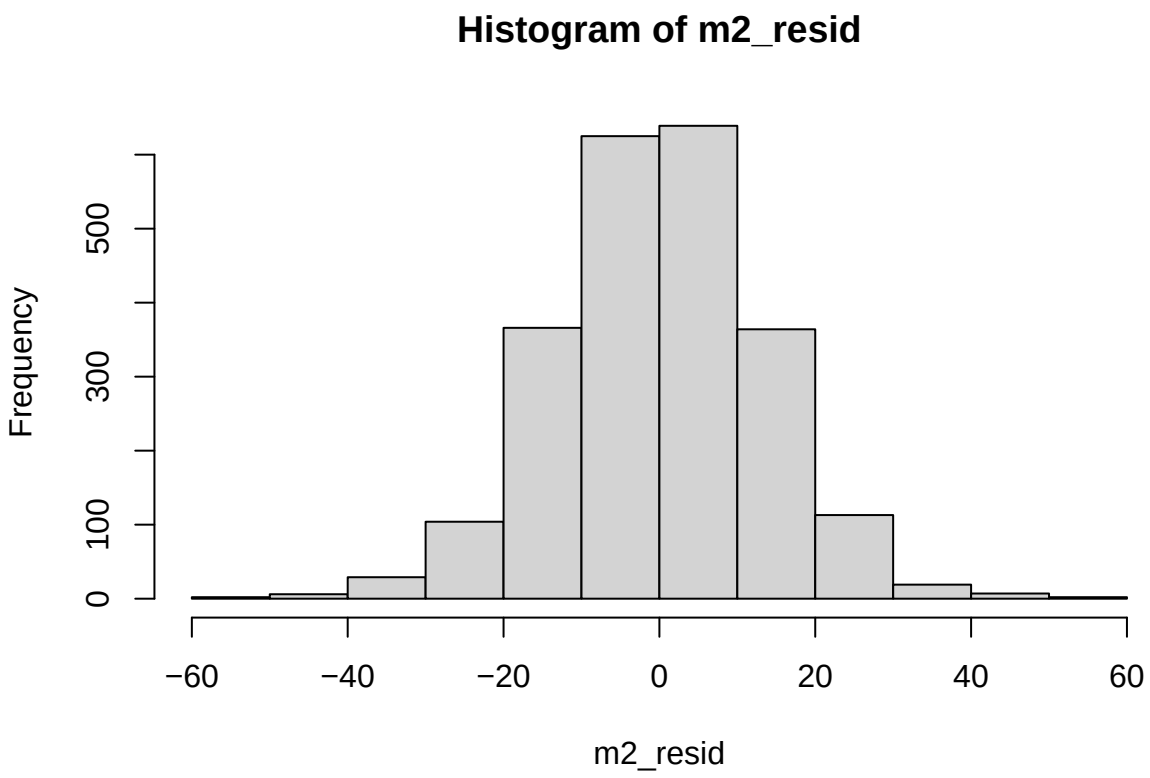


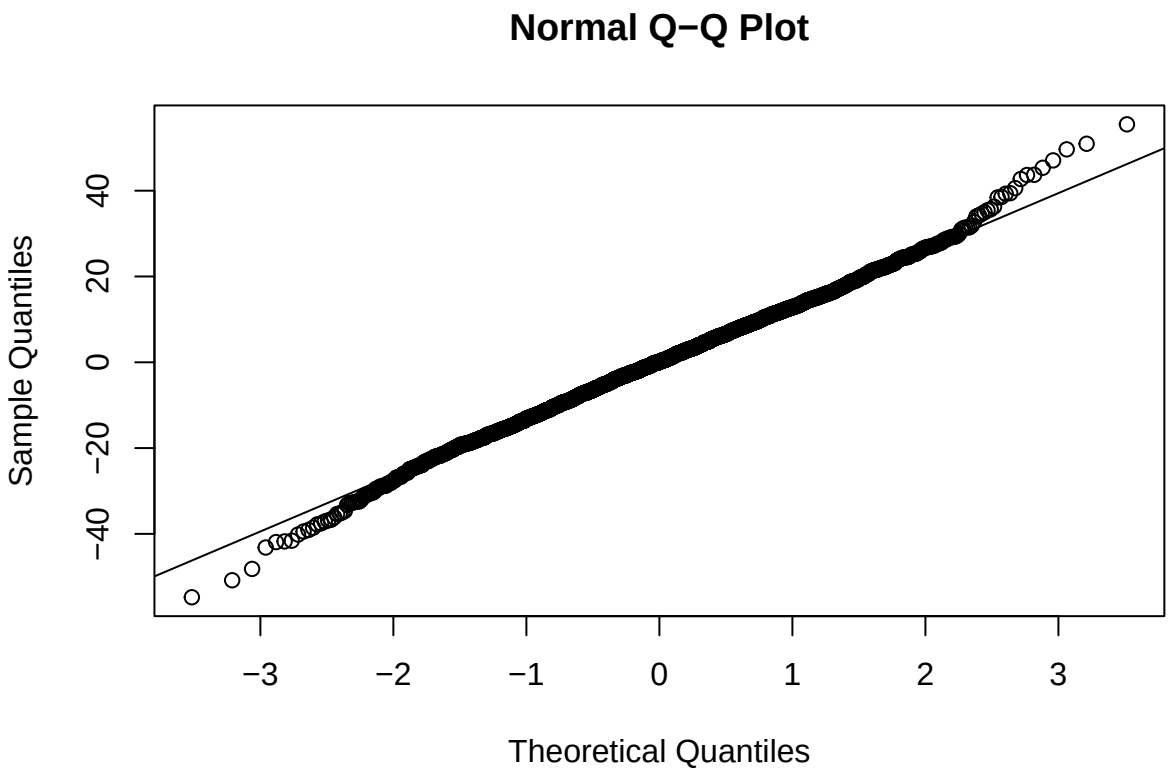
Model 2:

For the second model we narrowed down the variable selection based on our findings that TEAM_PITCHING_HR has high multicollinearity with TEAM_BATTING_HR, therefore we removed TEAM_PITCHING_HR. In addition, we removed TEAM_BATTING_SO, TEAM_BASERUN_SB, TEAM_BASERUN_CS, TEAM_PITCHING_SO, TEAM_FIELDING_DP for missing values. Our thoughts here is that by removing these variables our model is more reliable due to removal of imputed values and reduced model complexity.

```
##
## Call:
## lm(formula = TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_2B +
##     TEAM_BATTING_3B + TEAM_BATTING_HR + TEAM_BATTING_BB + TEAM_PITCHING_H +
##     TEAM_PITCHING_BB + TEAM_FIELDING_E, data = Training_imp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -54.776  -8.875   0.097   8.860  55.466
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       7.290e+00  3.443e+00   2.117 0.034376 *
## TEAM_BATTING_H    4.848e-02  3.207e-03  15.118  < 2e-16 ***
## TEAM_BATTING_2B  -2.582e-02  9.057e-03  -2.851 0.004400 **
## TEAM_BATTING_3B   1.011e-01  1.665e-02   6.072 1.48e-09 ***
```

```
## TEAM_BATTING_HR   3.672e-02  7.749e-03   4.739 2.28e-06 ***
## TEAM_BATTING_BB  -7.926e-05  4.585e-03  -0.017 0.986208
## TEAM_PITCHING_H  -1.312e-03  3.683e-04  -3.561 0.000377 ***
## TEAM_PITCHING_BB  1.036e-02  2.802e-03   3.695 0.000225 ***
## TEAM_FIELDING_E  -1.664e-02  2.368e-03  -7.025 2.81e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.48 on 2267 degrees of freedom
## Multiple R-squared:   0.27,  Adjusted R-squared:  0.2675
## F-statistic: 104.8 on 8 and 2267 DF,  p-value: < 2.2e-16
```



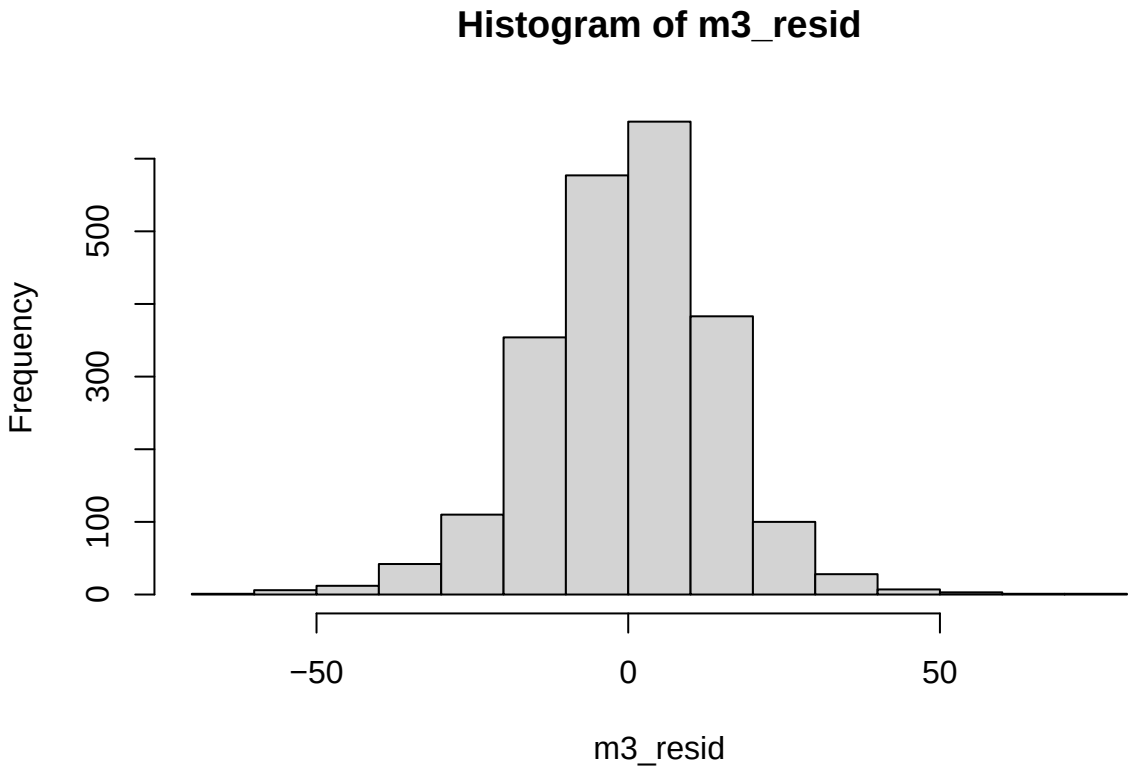


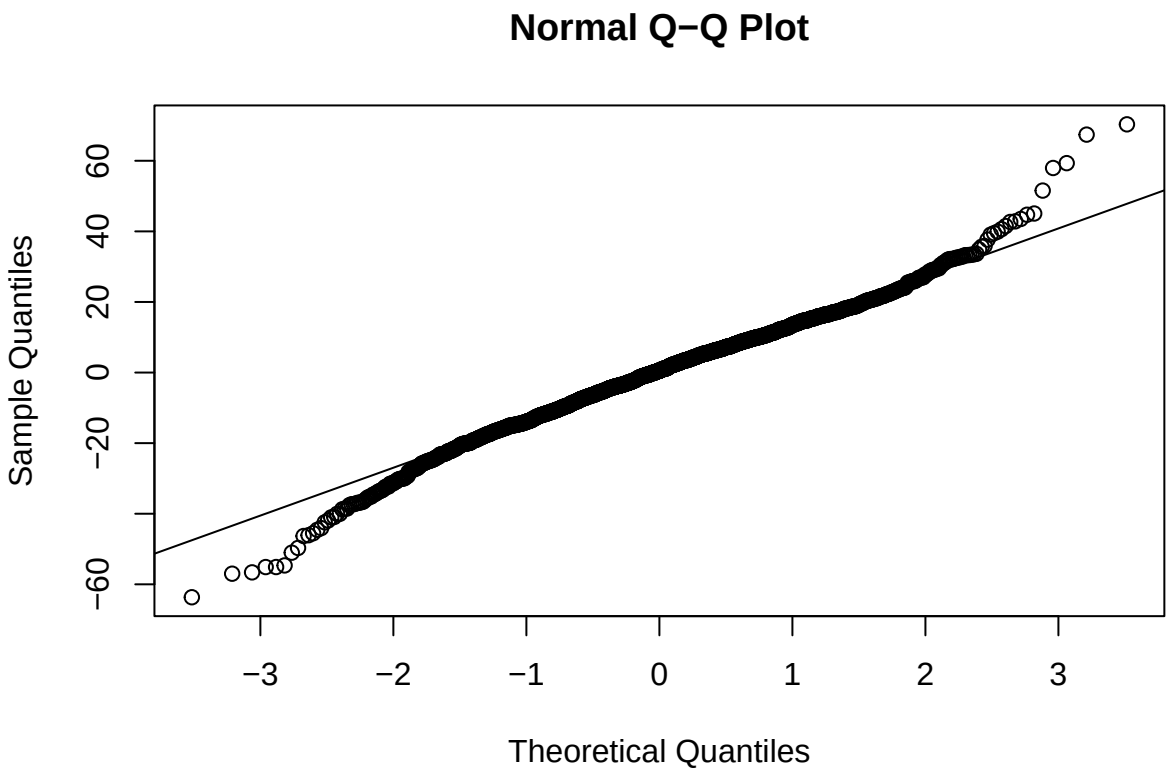
Model 3:

For our third model our group utilized the backward selection process where we removed the lowest p-value variables noted from model 1 and 2. Included in this model were only variables with p-values greater than 0.05.

```
##
## Call:
## lm(formula = TARGET_WINS ~ TEAM_BATTING_SO + TEAM_BASERUN_CS +
##     TEAM_PITCHING_HR + TEAM_PITCHING_BB + TEAM_BATTING_BB, data = Training_imp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -63.659  -8.994   0.549   9.297  70.322
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      63.658983   1.850740  34.397  < 2e-16 ***
## TEAM_BATTING_SO  -0.021016   0.001696 -12.388  < 2e-16 ***
## TEAM_BASERUN_CS   0.083583   0.007696  10.860  < 2e-16 ***
## TEAM_PITCHING_HR  0.116163   0.007706  15.075  < 2e-16 ***
## TEAM_PITCHING_BB -0.009051   0.002172  -4.166 3.21e-05 ***
## TEAM_BATTING_BB   0.037613   0.003223  11.669  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 14.4 on 2270 degrees of freedom
## Multiple R-squared:  0.1657, Adjusted R-squared:  0.1638
## F-statistic: 90.14 on 5 and 2270 DF,  p-value: < 2.2e-16
```





Select Models:

While Model 1 has higher multicollinearity in certain predictors, our analysis identified Model 1 as the strongest regression model. It achieved the lowest residual error (12.66) and the highest adjusted R^2 (0.354), making it the most accurate and reliable predictor of team wins. Model 1's residuals show a normal distribution and a normal looking Q-Q plot.

Model 1 shows that for a baseball team to increase their amount of wins for the season they should focus on increasing their batting home runs and stolen bases. TEAM_BATTING_HR has the greatest positive impact at a coefficient of 0.05764 and TEAM_BASERUN_SB has the second greatest positive impact with a coefficient of 0.04945. Conversely, minimizing fielding errors (TEAM_FIELDING_E) as this variable has the largest negative impact on wins with a coefficient of -0.041504.

The variable TEAM_BATTING_HR is noted to be highly correlated with TEAM_PITCHING_HR, however both of these variables have large theoretical impact to the probability of winning. Hitting a home run or allowing a home run directly influences the game's score and therefore our group decided to keep these variables.

```
## # A tibble: 3 x 4
##   Model     RSE Adj.R2 F.Statistic
##   <chr>   <dbl>  <dbl>       <dbl>
## 1 Model 1  12.7  0.354        90.1
## 2 Model 2  13.5  0.267       105.
## 3 Model 3  14.4  0.164        90.1
```

Model 1 variables VIF

```
##   TEAM_BATTING_H  TEAM_BATTING_2B  TEAM_BATTING_3B  TEAM_BATTING_HR
##         3.823342         2.460052         2.995896        36.657149
##  TEAM_BATTING_BB  TEAM_BATTING_SO  TEAM_BASERUN_SB  TEAM_BASERUN_CS
##         6.756380         5.274069         4.349937         4.373084
##  TEAM_PITCHING_H TEAM_PITCHING_HR TEAM_PITCHING_BB TEAM_PITCHING_SO
##         4.182680        29.664612         6.297724         3.336076
##  TEAM_FIELDING_E TEAM_FIELDING_DP
##         5.399699         1.872039
```

Model 2 variables VIF

```
##   TEAM_BATTING_H  TEAM_BATTING_2B  TEAM_BATTING_3B  TEAM_BATTING_HR
##         2.691190         2.248967         2.707698         2.755238
##  TEAM_BATTING_BB  TEAM_PITCHING_H TEAM_PITCHING_BB  TEAM_FIELDING_E
##         3.958646         3.361075         2.720094         3.642208
```

Model 3 variables VIF

```
##  TEAM_BATTING_SO  TEAM_BASERUN_CS TEAM_PITCHING_HR TEAM_PITCHING_BB
##         1.909613         1.635908         2.446552         1.432176
##  TEAM_BATTING_BB
##         1.714261
```

Utilizing our model 1 below we can see our predicted TARGET_WINS for the evaluation data.

```
##   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20
##  62  64  74  87  59  77  86  76  70  73  69  83  82  83  86  78  74  78  72  91
##  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36  37  38  39  40
##  81  84  80  72  83  88  49  74  83  74  90  86  81  83  79  86  75  90  87  91
##  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
##  83  91  27 110  97  98 101  76  68  80  77  86  76  74  75  79  93  76  65  81
##  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80
##  88  73  88  86  85 109  73  79  79  88  85  71  78  90  80  84  82  84  73  78
##  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
##  85  87  97  74  84  82  84  84  90  90  83  84  74  88  86  85  88 101  86  87
## 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
##  79  74  84  84  78  63  58  77  86  57  85  87  95  92  81  78  86  81  73  78
## 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
##  99  69  70  68  68  88  90  77  93  91  85  80  80  85  88  72  74  78  89  81
## 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
##  64  74  90  72  71  71  77  79  79  83  82  80  53  69  76  71  90  68  91  76
## 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
## 103 107  94 103  97  90  82  83  74  81  90  89  80  94  83  73  78  70  74  79
## 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
##  90  89  86  85  85  99  88  65  64 115  71  84  77  77  79  68  78  84  79  83
## 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220
##  74  79  72  92  82  83  77  77  83  73 105  90  81  65  68  83  78  95  78  78
## 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
##  78  74  81  73  71  75  82  78  81  84  82  93  79  89  80  75  82  77  89  72
## 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259
##  88  86  83  82  61  88  81  85  73  84  80  63  96  34  69  77  83  85  77
```