

baseball.R

ROG

2021-07-21

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3      v purrr  0.3.4
## v tibble  3.0.6      v dplyr  1.0.4
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(Hmisc)
```

```
## Warning: package 'Hmisc' was built under R version 4.0.5
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
##
```

```
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:dplyr':
```

```
##
```

```
##      src, summarize
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      format.pval, units
```

```
library(ggpubr)
```

```
## Warning: package 'ggpubr' was built under R version 4.0.5
```

```
library(Metrics)
```

```
## Warning: package 'Metrics' was built under R version 4.0.5
```

```
# Objectives
```

```
# This project aims to determine factors associated with game performance and  
# in particular whether tight defense is associated with increased game won, and  
# whether the conclusion that tight defenses increase game won is valid across  
# multiple time periods.
```

```
data <- read.csv('C:/Users/ROG/Desktop/baseball_teams.csv')  
head(data)
```

```
##   Year League Team Franchise Division Final_Standing Games_Played  X Games_Won  
## 1 1871  <NA>  BS1      BNA              3             31 NA         20  
## 2 1871  <NA>  CH1      CNA              2             28 NA         19  
## 3 1871  <NA>  CL1      CFC              8             29 NA         10  
## 4 1871  <NA>  FW1      KEK              7             19 NA          7  
## 5 1871  <NA>  NY2      NNA              5             33 NA         16  
## 6 1871  <NA>  PH1      PNA              1             28 NA         21  
##   Games_Lost X.1 X.2 League_Win World_Series Runs_Scored At_Bats Hits Doubles  
## 1          10 NA  NA      FALSE          NA         401   1372  426    70  
## 2           9 NA  NA      FALSE          NA         302   1196  323    52  
## 3          19 NA  NA      FALSE          NA         249   1186  328    35  
## 4          12 NA  NA      FALSE          NA         137    746  178    19  
## 5          17 NA  NA      FALSE          NA         302   1404  403    43  
## 6           7 NA  NA       TRUE          NA         376   1281  410    66  
##   Triples Home_Runs Walks Strike_Outs Stolen_Bases Caught_Stealing Hit_By_Pitch  
## 1        37         3   60          19          73           NA         NA  
## 2        21        10   60          22          69           NA         NA  
## 3        40         7   26          25          18           NA         NA  
## 4         8         2   33           9          16           NA         NA  
## 5        21         1   33          15          46           NA         NA  
## 6        27         9   46          23          56           NA         NA  
##   Sacrifice_Fly Runs_Against Earned_Runs Earned_Run_Average Complete_Games  
## 1              NA         303         109          3.55           22  
## 2              NA         241          77          2.76           25  
## 3              NA         341         116          4.11           23  
## 4              NA         243          97          5.17           19  
## 5              NA         313         121          3.72           32  
## 6              NA         266         137          4.95           27  
##   Shutout Saves Infield_Put_Outs Hits_Allowed Home_Run_Allowed Walks_Allowed  
## 1         1     3             828         367             2           42  
## 2         0     1             753         308             6           28  
## 3         0     0             762         346            13           53  
## 4         1     0             507         261             5           21  
## 5         1     0             879         373             7           42  
## 6         0     0             747         329             3           53  
##   Strikeouts_Allowed Errors Double_Plays Fielding_Percentage  
## 1                 23   225          NA          0.83  
## 2                 22   218          NA          0.82  
## 3                 34   223          NA          0.81  
## 4                 17   163          NA          0.80
```

```
## 5          22    227          NA          0.83
## 6          16    194          NA          0.84
##          Team_Name          Home_Ball_Park Attendance
## 1    Boston Red Stockings          South End Grounds I          NA
## 2    Chicago White Stockings          Union Base-Ball Grounds          NA
## 3    Cleveland Forest Citys National Association Grounds          NA
## 4      Fort Wayne Kekiongas          Hamilton Field          NA
## 5      New York Mutuals          Union Grounds (Brooklyn)          NA
## 6    Philadelphia Athletics          Jefferson Street Grounds          NA
```

```
df <- data[, c(1, 9, 10, 15:40)]
```

```
# Check missing point
apply(df, 2, function(x) sum(is.na(x)))
```

```
##          Year          Games_Won          Games_Lost          Runs_Scored
##          0          0          0          0
##          At_Bats          Hits          Doubles          Triples
##          0          0          0          0
##          Home_Runs          Walks          Strike-Outs          Stolen_Bases
##          0          0          120          144
##          Caught_Stealing          Hit_By_Pitch          Sacrifice_Fly          Runs_Against
##          859          2325          2325          0
##          Earned_Runs          Earned_Run_Average          Complete_Games          Shutout
##          0          0          0          0
##          Saves          Infield_Put-Outs          Hits_Allowed          Home_Run_Allowed
##          0          0          0          0
##          Walks_Allowed          Strikeouts_Allowed          Errors          Double_Plays
##          0          0          0          317
##          Fielding_Percentage
##          0
```

```
# Period 1 - before 1920
df_before_1920 <- filter(df, Year < 1920)

# Period 2 - 1920 to 1960
df_1920_1960 <- filter(df, between(Year, 1920, 1960))

# Period 3 - 1960 to 1990
df_1960_1990 <- filter(df, between(Year, 1960, 1990))

# Period 4 - 1990 to 2010
df_1990_2010 <- filter(df, between(Year, 1990, 2010))

# Perform an exploratory data analysis to understand the statistical
# properties and distribution shapes of 8 key variables.
# Perform this analysis for a single dataset that spans 1960 to 2010

# Calculate the descriptive statistics for the 8 variables
df_1960_2010 <- filter(df, between(Year, 1960, 2010))
key_variables <- df_1960_2010[, c('Runs_Scored', 'At_Bats', 'Hits', 'Walks',
```

```

                                'Runs_Against', 'Earned_Run_Average', 'Fielding_Percentage',
                                'Saves'])
summary(key_variables)

```

```

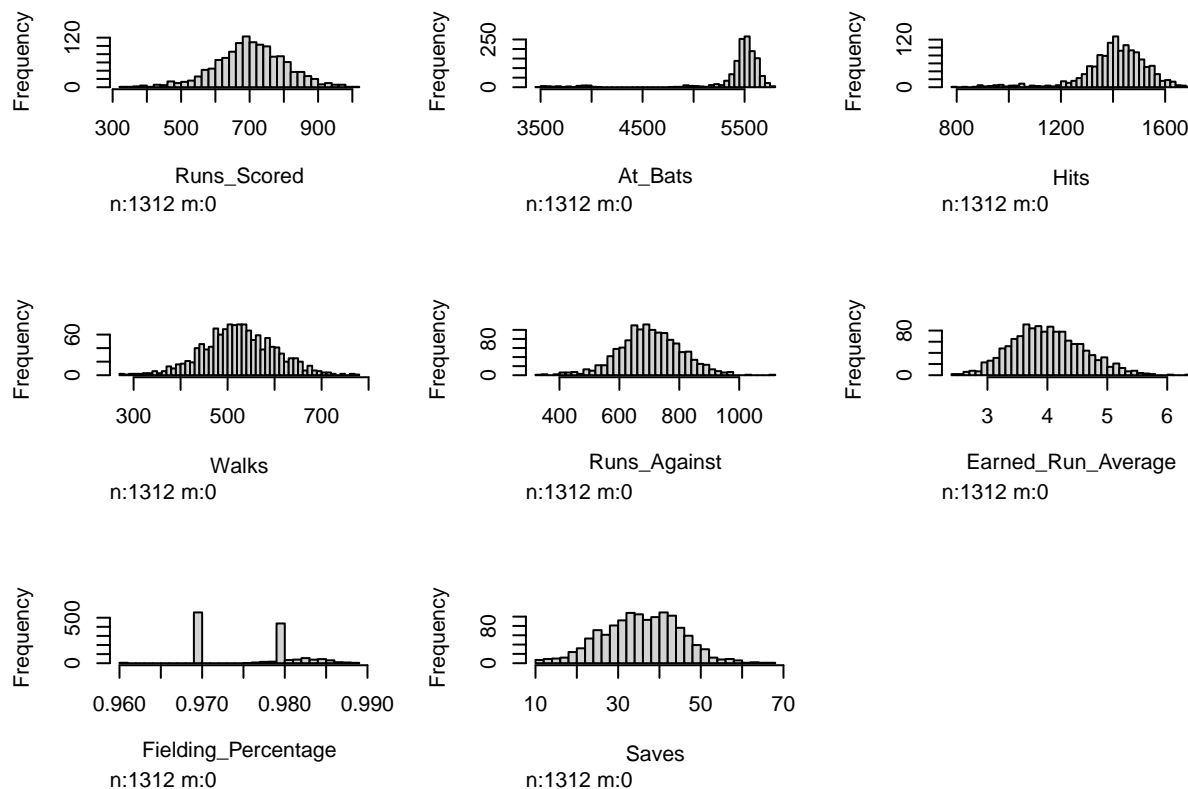
##   Runs_Scored      At_Bats      Hits      Walks
##   Min.       : 329.0    Min.       :3493    Min.       : 797    Min.       :275.0
##   1st Qu.: 641.0    1st Qu.:5441    1st Qu.:1357    1st Qu.:474.8
##   Median : 704.0    Median :5506    Median :1418    Median :523.0
##   Mean      : 703.4    Mean      :5428    Mean      :1409    Mean      :524.7
##   3rd Qu.: 771.0    3rd Qu.:5575    3rd Qu.:1489    3rd Qu.:575.2
##   Max.      :1009.0    Max.      :5781    Max.      :1684    Max.      :775.0
##   Runs_Against Earned_Run_Average Fielding_Percentage Saves
##   Min.       : 331.0    Min.       :2.450    Min.       :0.9600    Min.       :10.00
##   1st Qu.: 640.0    1st Qu.:3.580      1st Qu.:0.9700      1st Qu.:30.00
##   Median : 701.0    Median :3.970      Median :0.9800      Median :36.00
##   Mean      : 703.4    Mean      :4.007      Mean      :0.9764      Mean      :36.16
##   3rd Qu.: 770.2    3rd Qu.:4.390      3rd Qu.:0.9800      3rd Qu.:43.00
##   Max.      :1103.0    Max.      :6.380      Max.      :0.9890      Max.      :68.00

```

```

# histogram for each of the 8 variables
hist.data.frame(key_variables)

```

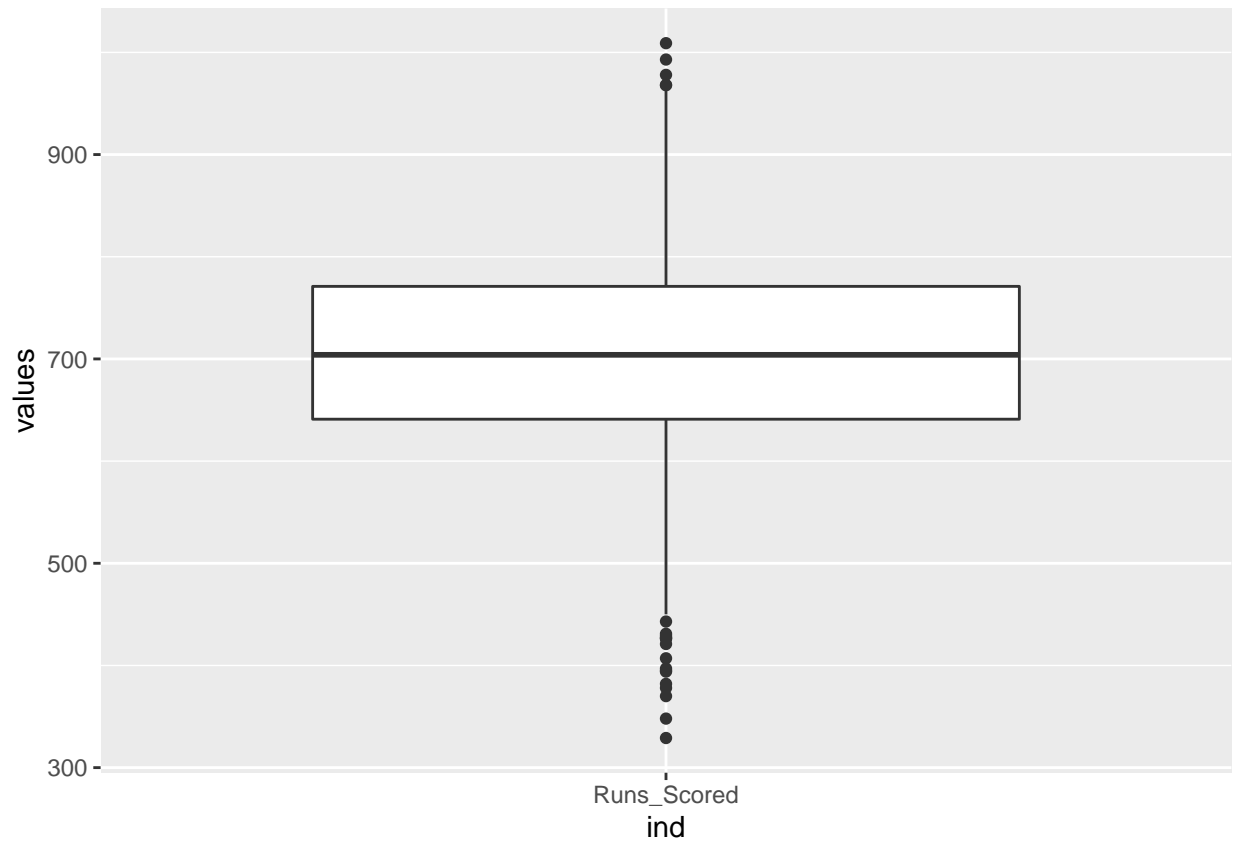


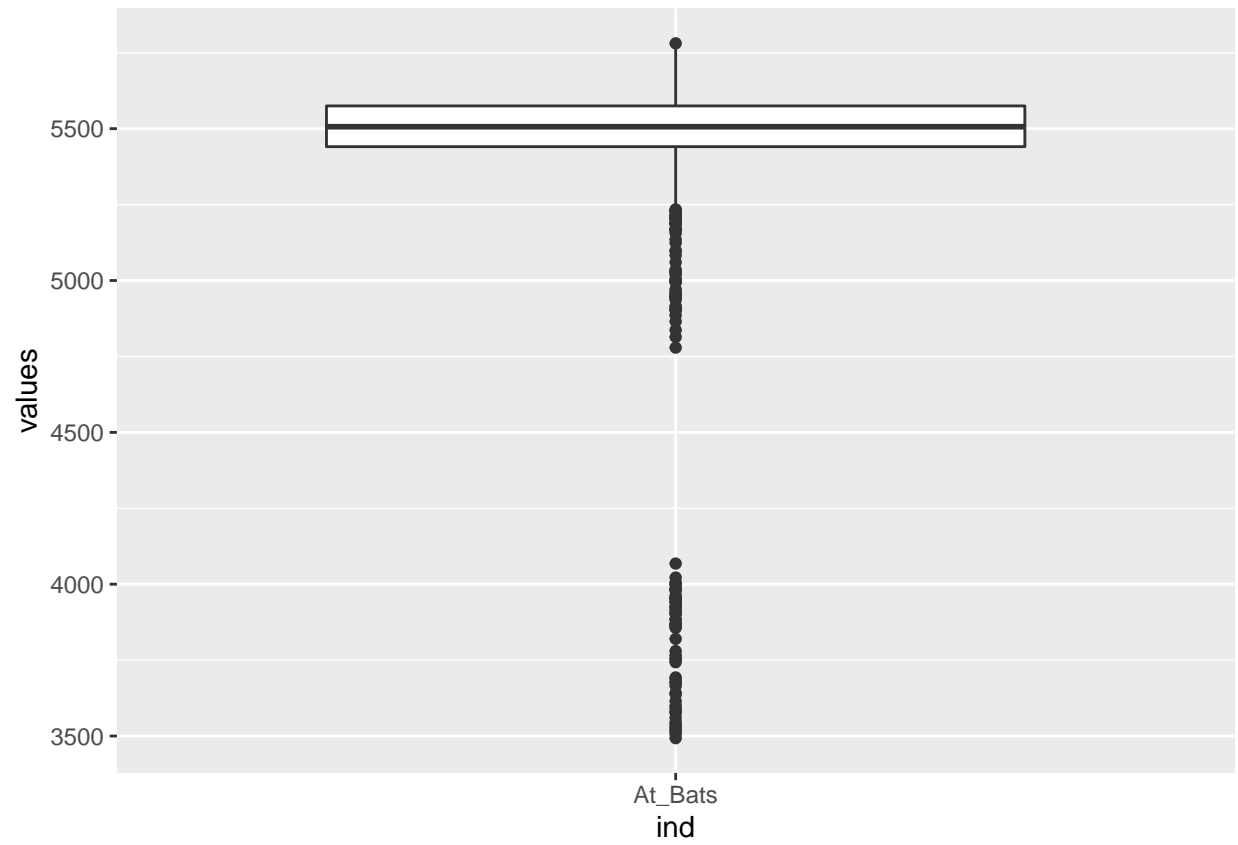
```

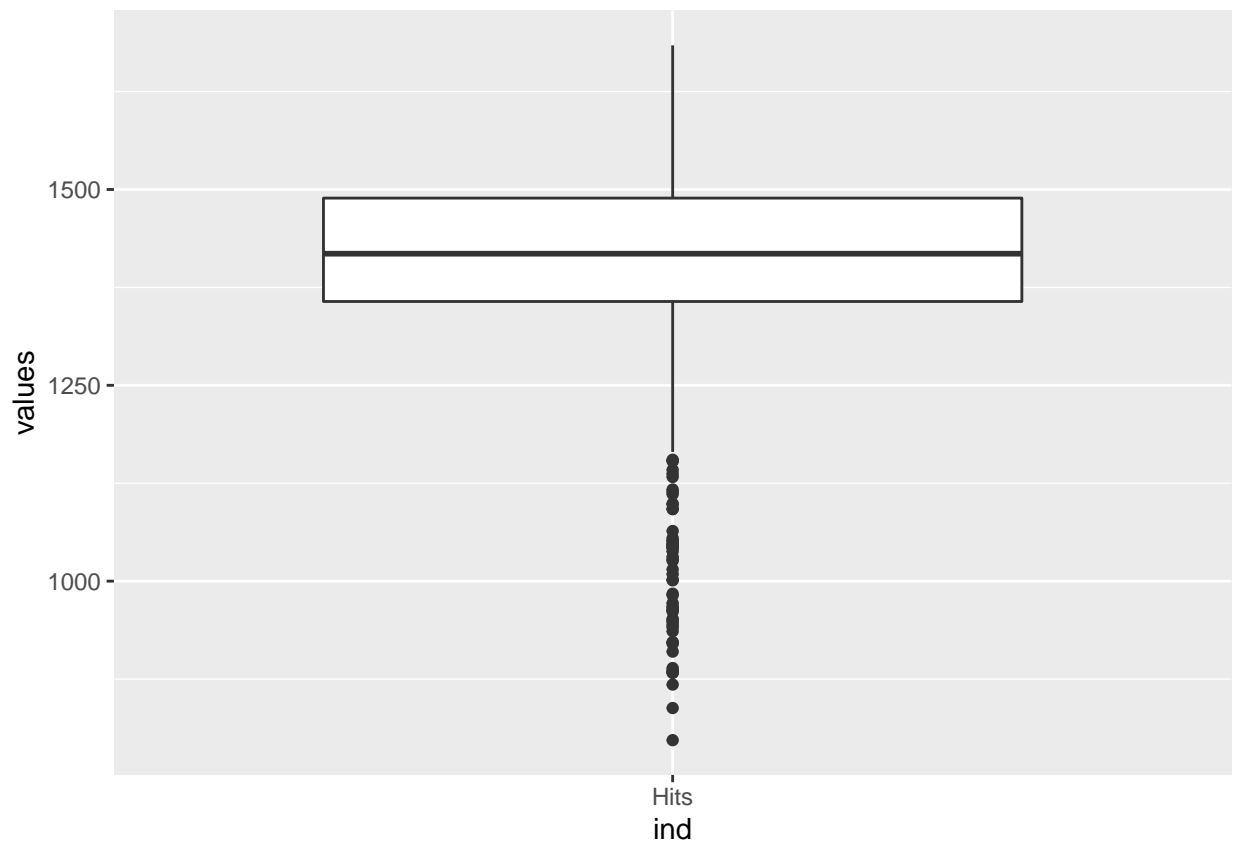
# boxplot for each of the 8 variables
for (i in 1:ncol(key_variables)){

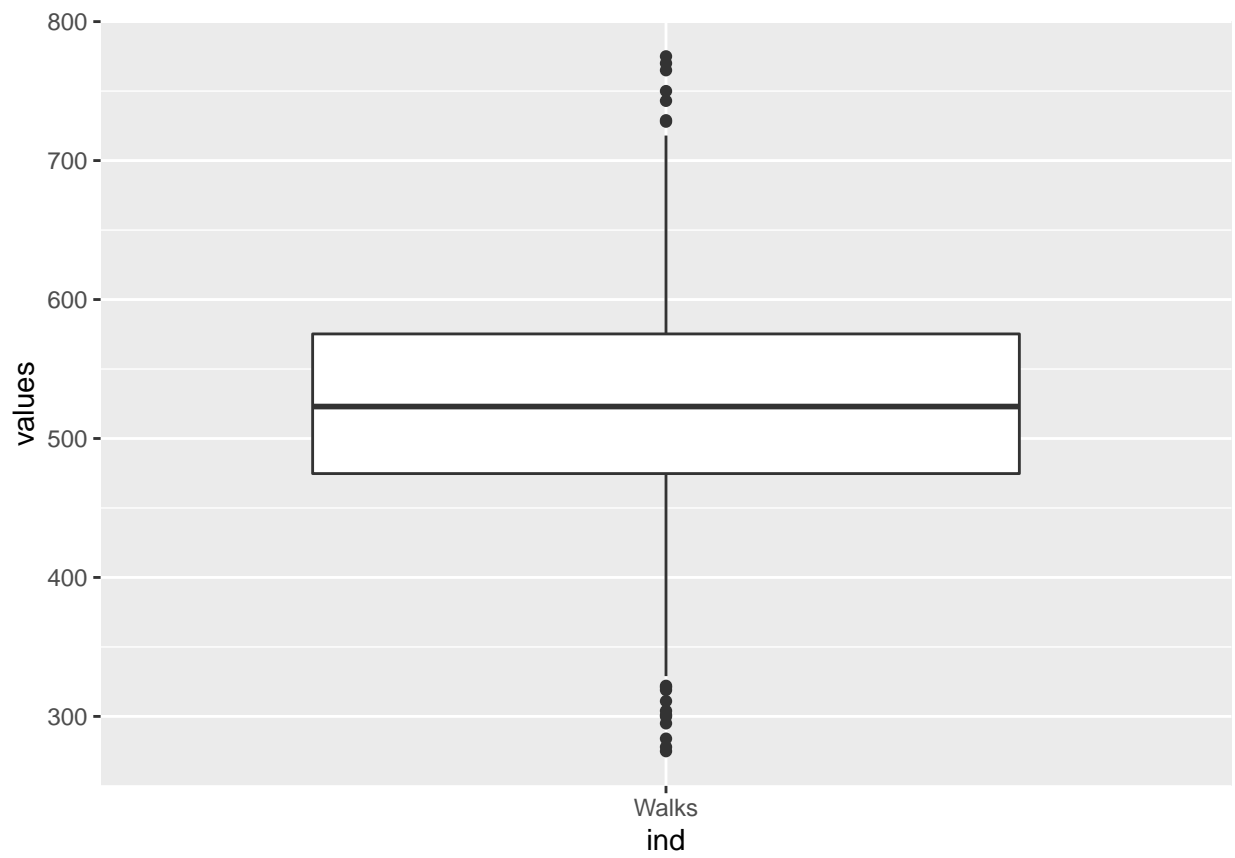
```

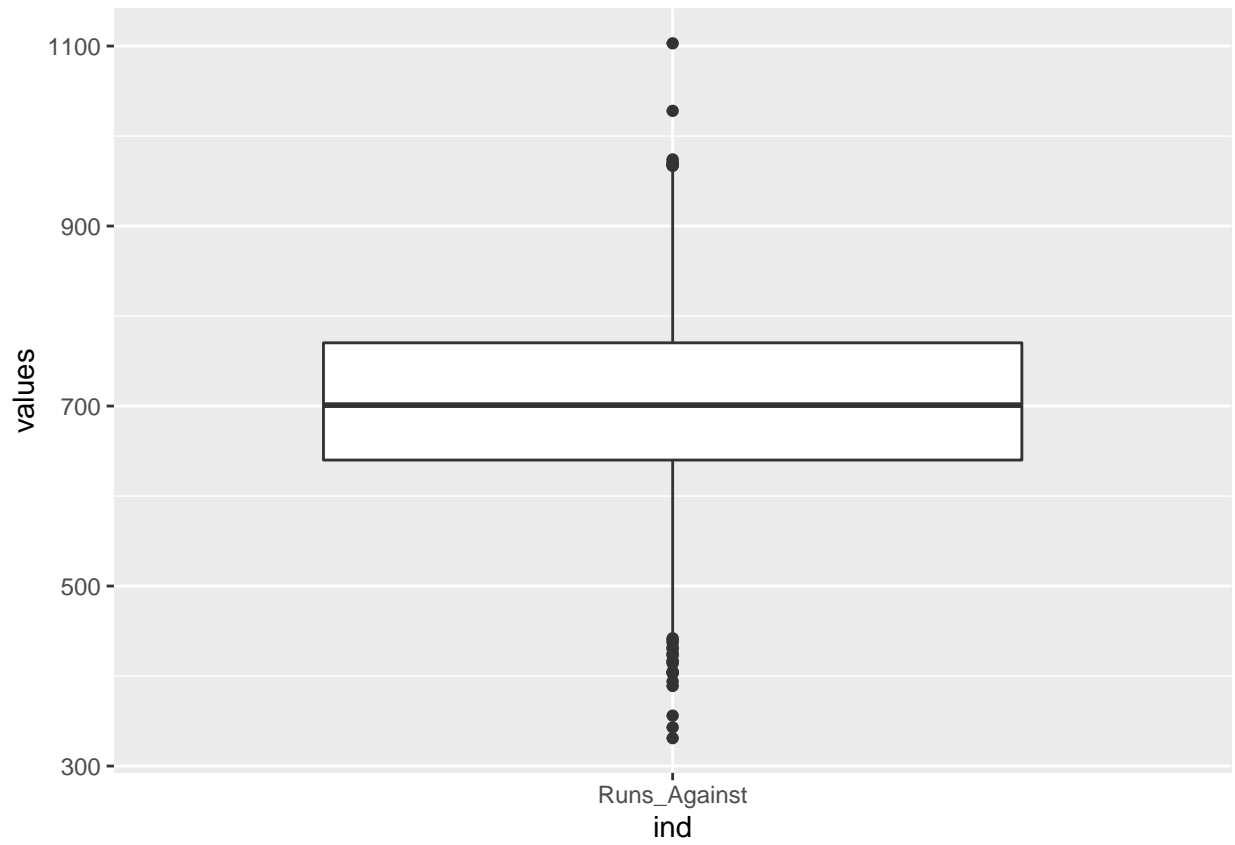
```
print(ggplot(stack(key_variables[i]), aes(x = ind, y = values)) +
      geom_boxplot())
}
```

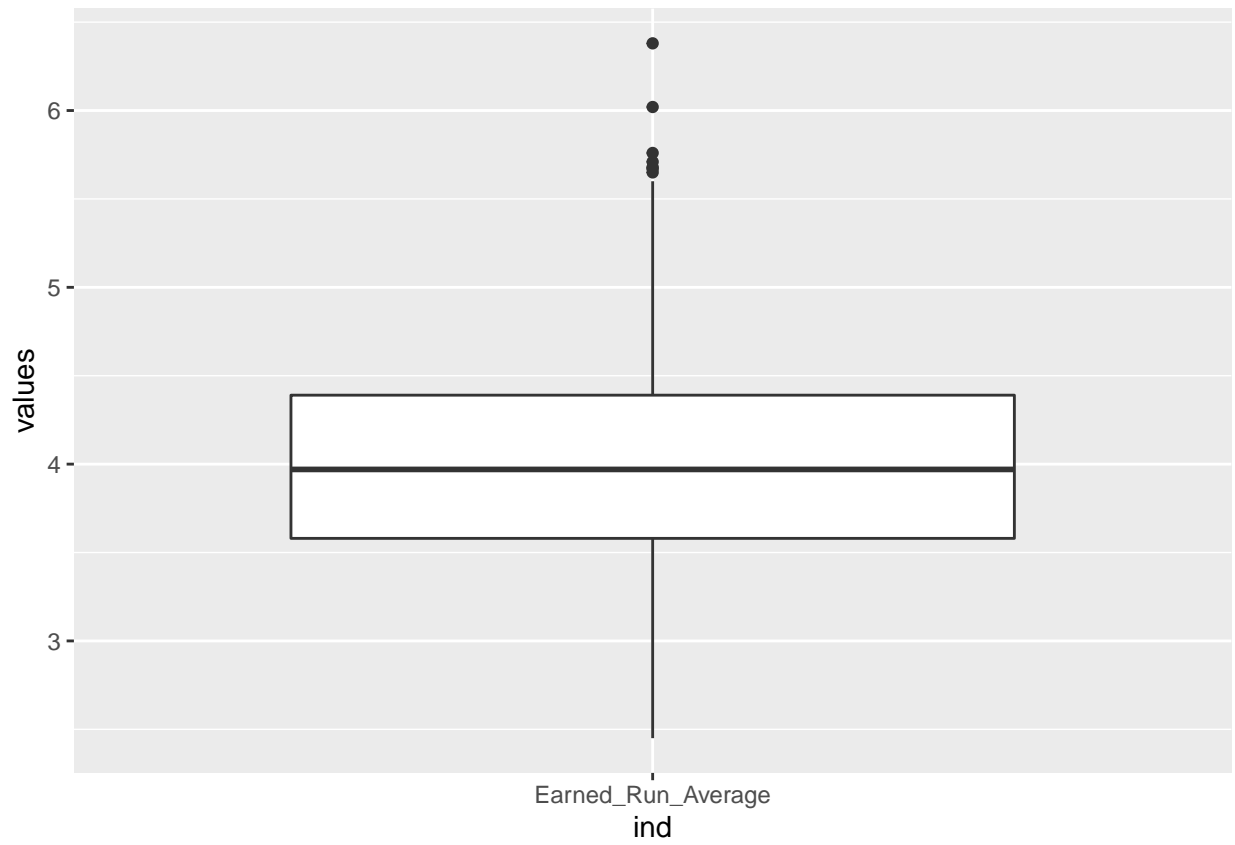


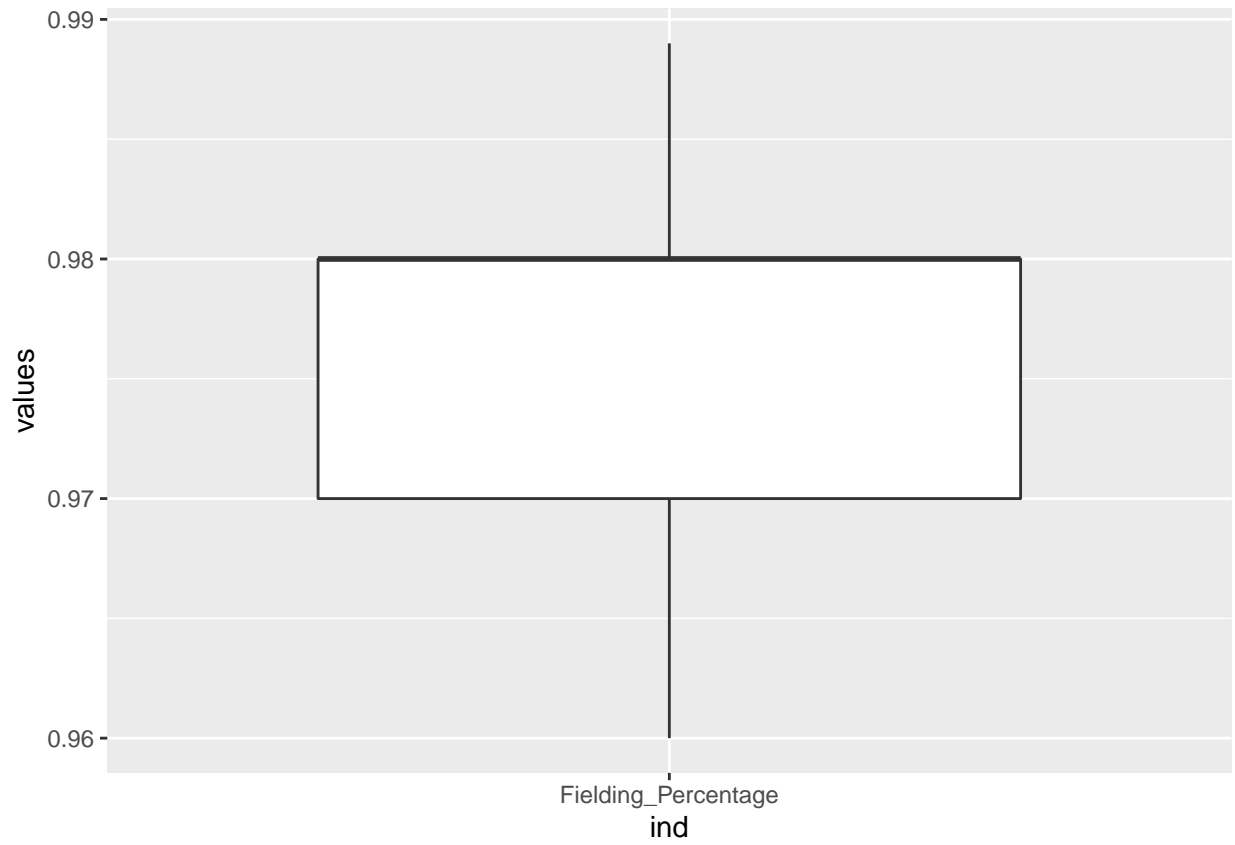


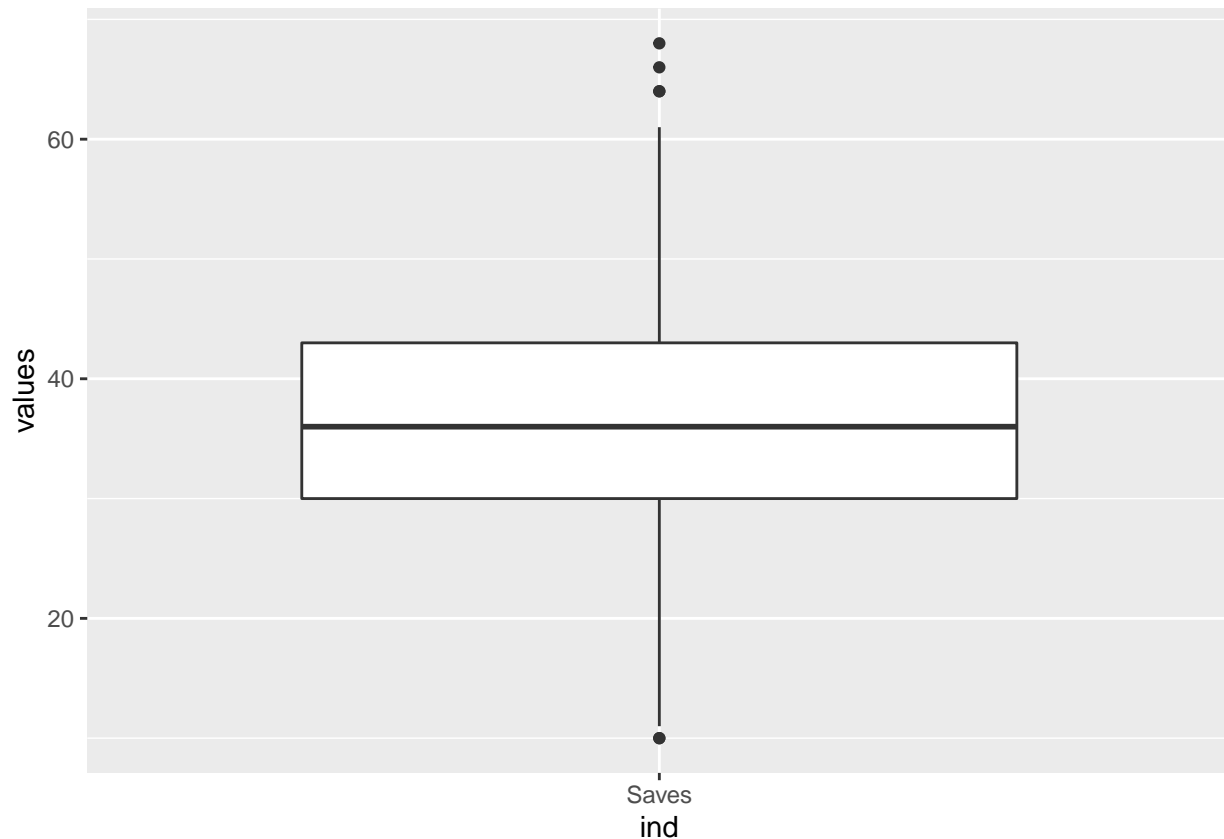












validity of the assumption that each of the 8 variables is normally distributed

*# if the p-value > 0.05 implying that the distribution of the data are not
significantly different from normal distribution*

```
test <- matrix(0, nrow = ncol(key_variables), ncol = 2)
for (i in 1:ncol(key_variables)) {
  test[i, 1] <- names(key_variables)[i]
  test[i, 2] <- shapiro.test(key_variables[,i])[[2]]
}
tibble('variables' = test[,1], 'p_value' = test[,2])
```

```
## # A tibble: 8 x 2
##   variables      p_value
##   <chr>         <chr>
## 1 Runs_Scored  0.000244025534546546
## 2 At_Bats      1.76036455340308e-51
## 3 Hits         7.03424660341483e-28
## 4 Walks        0.139395243265944
## 5 Runs_Against 0.00637230131975897
## 6 Earned_Run_Average 2.95774045541599e-05
## 7 Fielding_Percentage 8.02485485174246e-37
## 8 Saves        0.0053189326971369
```

*# Correlation analysis for the following pairs of fields for the complete data
spanning 1960 to 2010*

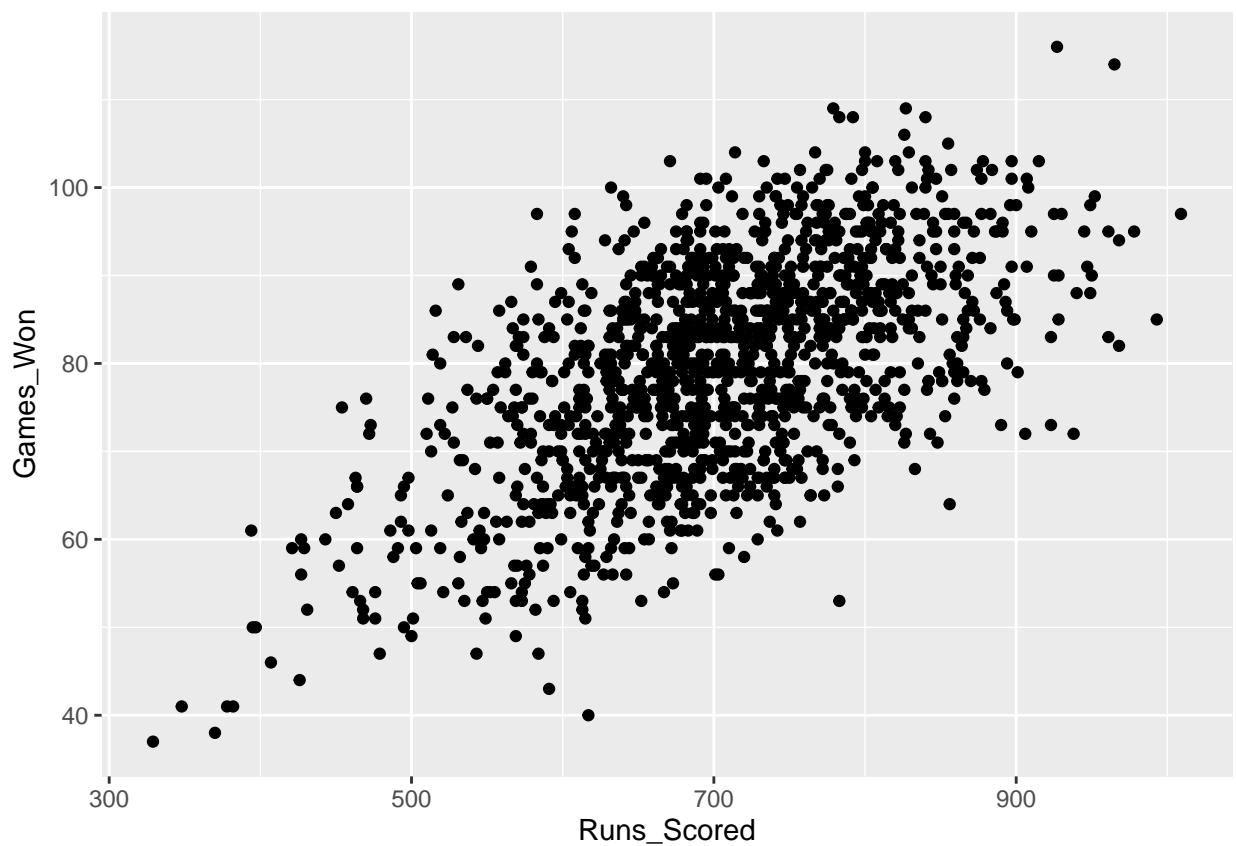
```

# Games Won vs Runs Scored, Games Lost vs Runs Scored
# Games Won vs Runs Against, Games Lost vs Runs Against
# Games Won vs (Runs Scored minus Runs Against), Games Lost vs (Runs Scored
# minus Runs Against)

df2 <- df_1960_2010 %>%
  select(Games_Won, Games_Lost, Runs_Scored, Runs_Against) %>%
  mutate('Runs_Scored-Runs_Against' = Runs_Scored - Runs_Against)

for (i in 3:ncol(df2)){
  for (j in 1:2) {
    print(ggplot(mapping = aes(x = unlist(df2[i]), y = unlist(df2[j]))) +
      geom_point() +
      xlab(names(df2)[i]) +
      ylab(names(df2)[j]))
    print(cor(df2[i], df2[j]))
  }
}

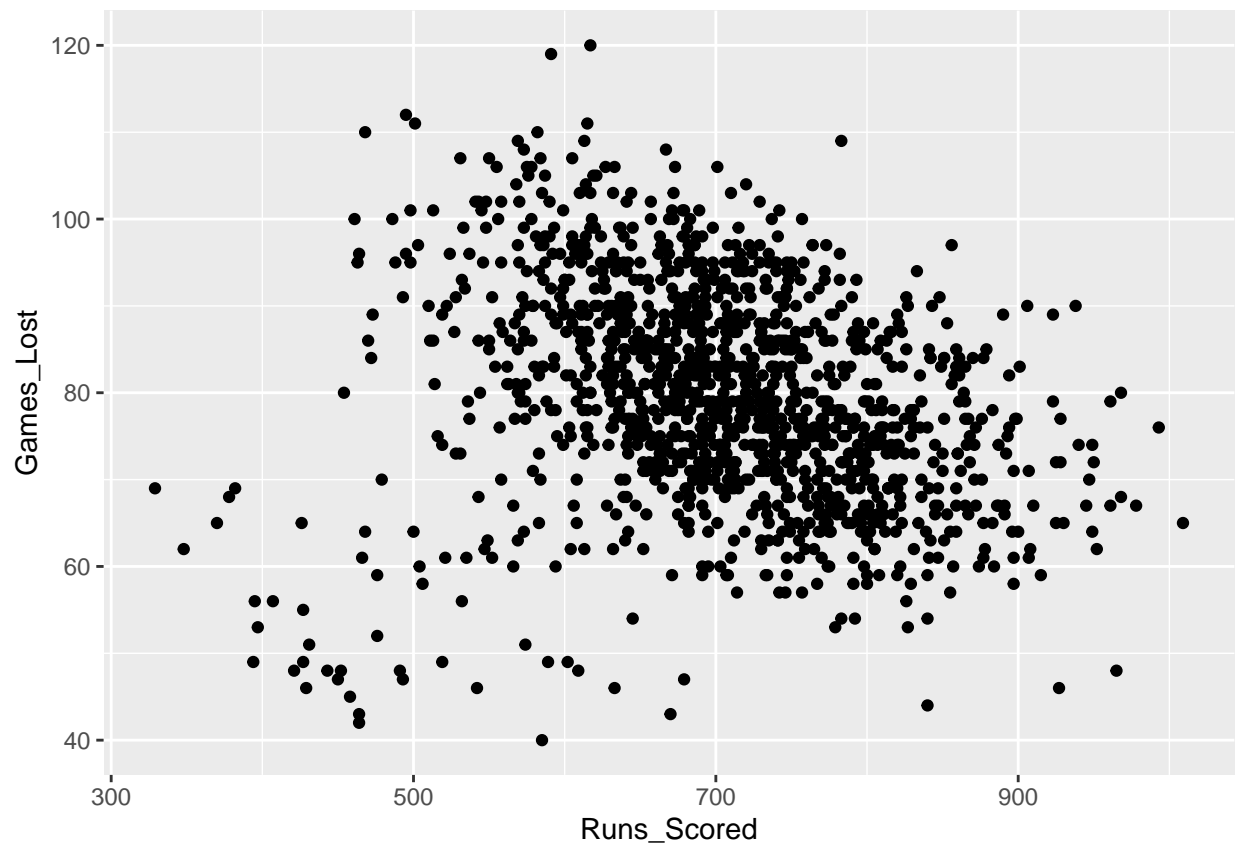
```



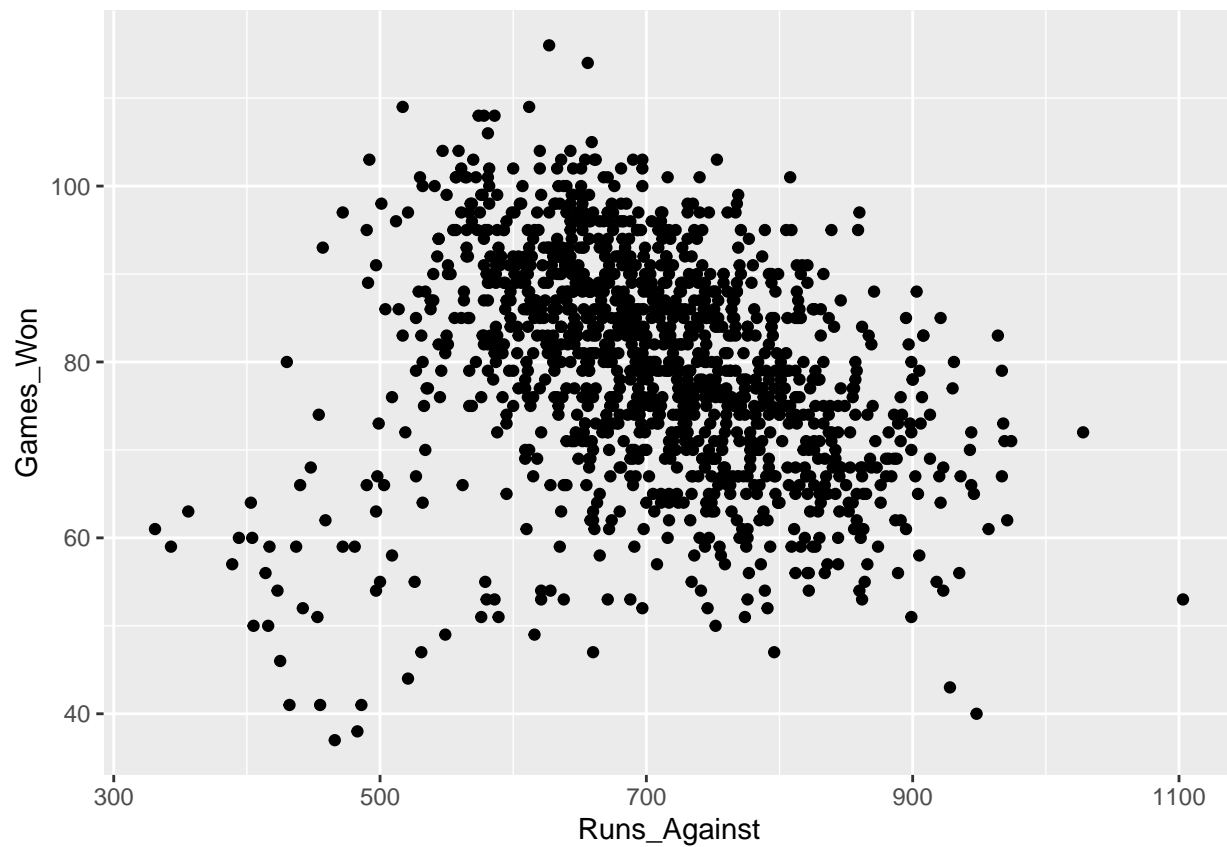
```

##           Games_Won
## Runs_Scored 0.5958278

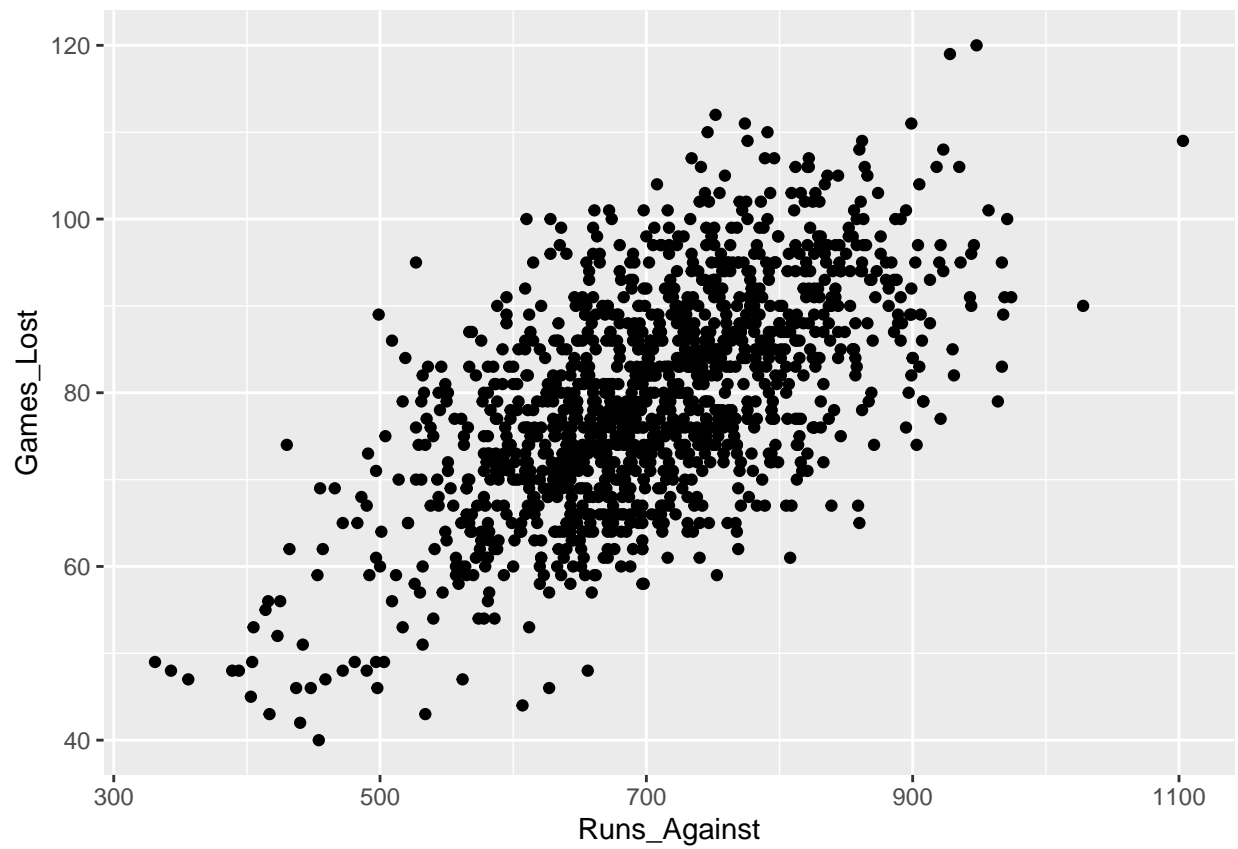
```



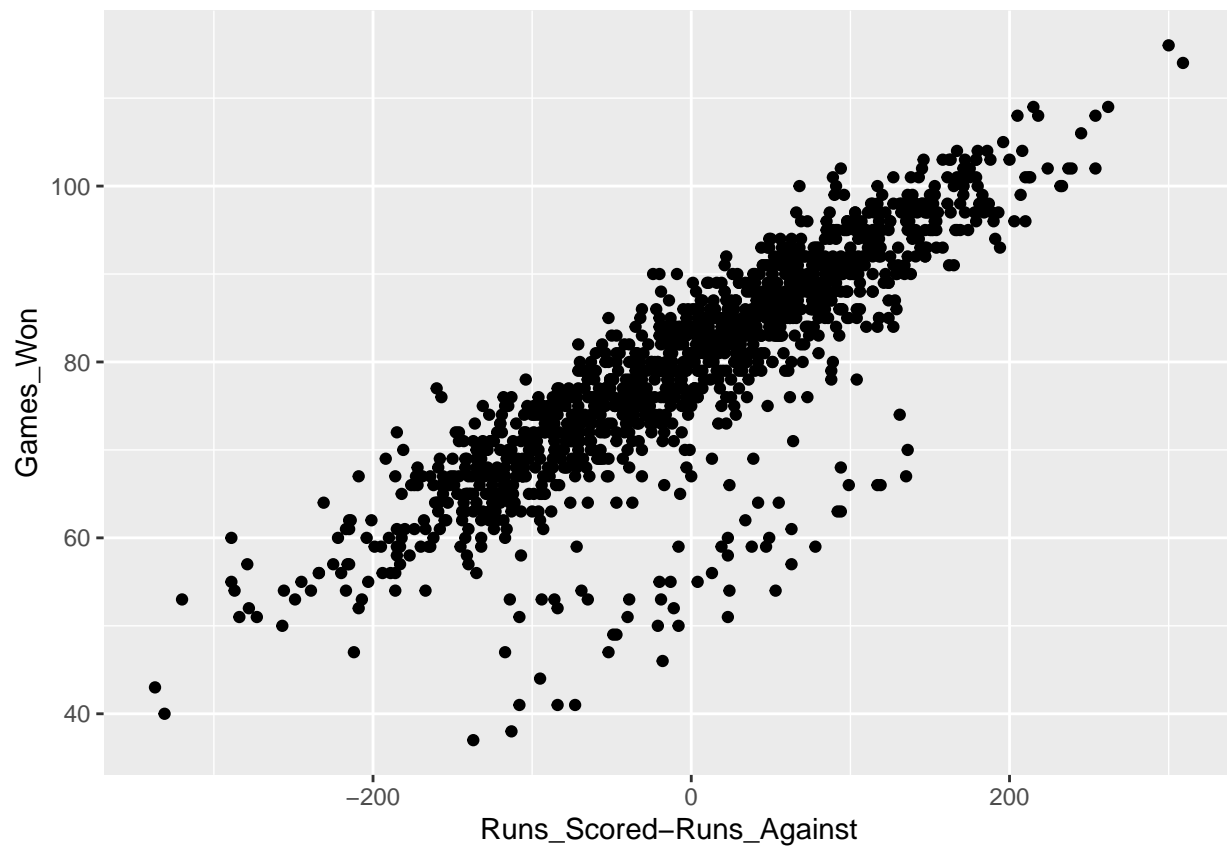
```
##           Games_Lost
## Runs_Scored -0.2267723
```



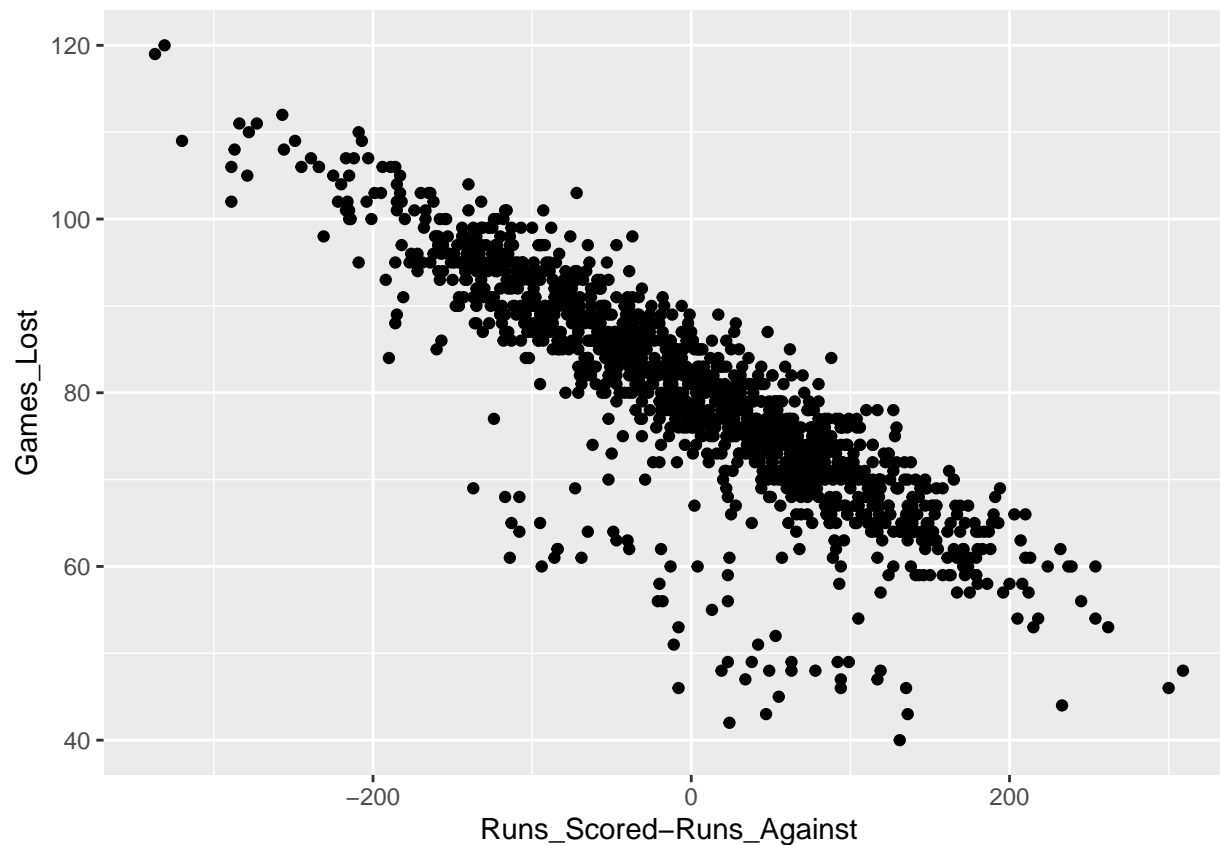
```
## Games_Won
## Runs_Against -0.2481575
```



```
##           Games_Lost
## Runs_Against 0.6111768
```

```
##                                     Games_Won
## Runs_Scored-Runs_Against 0.8504114
```



```
##                                     Games_Lost
## Runs_Scored-Runs_Against -0.8493582
```

*# Create a multiple linear regression model for each of the 4 time periods noted
above and select the independent variables that are good predictors of games won*

```
drop.na <- function(data) {
  dt <- data
  empty_frame <- data.frame(matrix(0, nrow(dt), ncol(dt)))
  colnames(empty_frame) <- names(dt)
  for (i in 1:ncol(dt)){
    if (colSums(is.na(dt[i])) < nrow(dt)) {
      empty_frame[i] <- dt[i]
    }
  }
  new_dt <- na.omit(empty_frame) %>%
    select_if(colSums(.) != 0)
  rownames(new_dt) <- NULL
  new_dt
}
```

Period 1 - before 1920

```
df_p1 <- drop.na(df_before_1920[-c(1, 3)])
```

Forward stepwise regression: starts with the intercept and adds new ones until the model cannot be improved

```
null_modelp1 <- lm(Games_Won ~ 1, data = df_p1)
full_modelp1 <- lm(Games_Won ~ ., data = df_p1)
```

```
step(null_modelp1, direction = "forward", scope = formula(full_modelp1))
```

Start: AIC=130.71

Games_Won ~ 1

	Df	Sum of Sq	RSS	AIC
## + Runs_Against	1	3269.6	1850.4	108.28
## + Earned_Run_Average	1	2885.9	2234.0	112.80
## + Earned_Runs	1	2791.3	2328.6	113.80
## + Runs_Scored	1	2737.3	2382.7	114.35
## + Fielding_Percentage	1	2306.0	2814.0	118.34
## + Saves	1	2274.6	2845.4	118.61
## + Hits_Allowed	1	2232.7	2887.2	118.96
## + Shutout	1	2081.8	3038.1	120.18
## + Errors	1	1952.5	3167.4	121.18
## + Strike_Outs	1	1687.8	3432.2	123.11
## + Walks_Allowed	1	1686.1	3433.8	123.12
## + Hits	1	1266.4	3853.5	125.89
## + Infield_Put_Outs	1	1184.8	3935.2	126.39
## + Complete_Games	1	1156.1	3963.9	126.57
## + Walks	1	938.1	4181.9	127.85
## <none>			5120.0	130.71
## + Doubles	1	362.7	4757.3	130.94
## + Double_Plays	1	334.9	4785.1	131.09
## + Triples	1	292.0	4827.9	131.30
## + At_Bats	1	259.8	4860.1	131.46
## + Stolen_Bases	1	246.4	4873.5	131.52
## + Home_Runs	1	239.0	4880.9	131.56
## + Home_Run_Allowed	1	209.0	4910.9	131.71
## + Caught_Stealing	1	203.3	4916.6	131.74
## + Strikeouts_Allowed	1	164.6	4955.4	131.92

Step: AIC=108.28

Games_Won ~ Runs_Against

	Df	Sum of Sq	RSS	AIC
## + Runs_Scored	1	1383.36	467.03	77.240
## + Hits	1	866.66	983.73	95.119
## + Strike_Outs	1	510.42	1339.97	102.536
## + Saves	1	509.29	1341.10	102.557
## + Walks	1	449.11	1401.28	103.610
## + Triples	1	232.77	1617.63	107.056
## + Doubles	1	220.58	1629.81	107.236
## <none>			1850.39	108.282
## + Home_Run_Allowed	1	143.10	1707.29	108.351
## + Infield_Put_Outs	1	106.04	1744.35	108.866
## + Walks_Allowed	1	65.74	1784.65	109.414
## + Stolen_Bases	1	46.26	1804.13	109.675
## + Earned_Runs	1	37.30	1813.09	109.794

```
## + Hits_Allowed      1      23.81 1826.58 109.972
## + Shutout           1      22.48 1827.91 109.989
## + Complete_Games    1      22.29 1828.11 109.992
## + Fielding_Percentage 1      21.41 1828.99 110.003
## + Home_Runs          1      17.64 1832.75 110.052
## + Earned_Run_Average 1      13.46 1836.94 110.107
## + Caught_Stealing    1      11.00 1839.40 110.139
## + At_Bats            1      10.06 1840.33 110.152
## + Errors             1       0.31 1850.08 110.278
## + Strikeouts_Allowed 1       0.21 1850.18 110.280
## + Double_Plays       1       0.05 1850.34 110.282
```

```
##
```

```
## Step: AIC=77.24
```

```
## Games_Won ~ Runs_Against + Runs_Scored
```

```
##
```

	Df	Sum of Sq	RSS	AIC
## + Errors	1	99.486	367.55	73.491
## + Doubles	1	85.518	381.52	74.386
## + Hits_Allowed	1	66.750	400.28	75.539
## + Infield_Put_Outs	1	66.288	400.75	75.567
## + Earned_Run_Average	1	65.281	401.75	75.627
## + Saves	1	60.478	406.56	75.912
## + Hits	1	59.231	407.80	75.986
## + Strike_Outs	1	51.657	415.38	76.427
## + Fielding_Percentage	1	48.697	418.34	76.598
## + Earned_Runs	1	39.268	427.77	77.133
## <none>			467.03	77.240
## + Shutout	1	34.467	432.57	77.400
## + At_Bats	1	28.463	438.57	77.731
## + Stolen_Bases	1	26.478	440.56	77.840
## + Walks_Allowed	1	18.224	448.81	78.285
## + Walks	1	15.101	451.93	78.452
## + Strikeouts_Allowed	1	11.603	455.43	78.637
## + Home_Run_Allowed	1	11.524	455.51	78.641
## + Complete_Games	1	4.099	462.93	79.029
## + Caught_Stealing	1	3.878	463.16	79.040
## + Double_Plays	1	3.853	463.18	79.042
## + Home_Runs	1	2.745	464.29	79.099
## + Triples	1	0.271	466.76	79.226

```
##
```

```
## Step: AIC=73.49
```

```
## Games_Won ~ Runs_Against + Runs_Scored + Errors
```

```
##
```

	Df	Sum of Sq	RSS	AIC
## + Strike_Outs	1	114.897	252.65	66.495
## + Stolen_Bases	1	85.443	282.10	69.141
## + Doubles	1	85.354	282.19	69.149
## + Caught_Stealing	1	42.583	324.96	72.536
## + Hits	1	35.479	332.07	73.055
## + Shutout	1	31.594	335.95	73.334
## <none>			367.55	73.491
## + At_Bats	1	29.076	338.47	73.513
## + Strikeouts_Allowed	1	27.340	340.21	73.636
## + Infield_Put_Outs	1	26.418	341.13	73.701

```

## + Saves          1      21.211 346.34 74.065
## + Hits_Allowed   1      20.114 347.43 74.140
## + Walks          1      19.995 347.55 74.149
## + Home_Run_Allowed 1       6.865 360.68 75.039
## + Double_Plays    1       5.927 361.62 75.101
## + Complete_Games  1       5.394 362.15 75.136
## + Walks_Allowed   1       3.961 363.59 75.231
## + Earned_Run_Average 1      3.300 364.25 75.275
## + Home_Runs       1       2.197 365.35 75.347
## + Fielding_Percentage 1      1.807 365.74 75.373
## + Triples         1       1.323 366.23 75.405
## + Earned_Runs     1       0.153 367.39 75.481
##
## Step:  AIC=66.49
## Games_Won ~ Runs_Against + Runs_Scored + Errors + Strike-Outs
##
##              Df Sum of Sq   RSS   AIC
## + Doubles      1    56.471 196.18 62.424
## + Strikeouts_Allowed 1    41.917 210.73 64.141
## + Complete_Games  1    32.353 220.30 65.206
## <none>                252.65 66.495
## + Stolen_Bases    1    18.948 233.70 66.624
## + Home_Run_Allowed 1    18.111 234.54 66.710
## + Infield_Put-Outs 1    17.722 234.93 66.749
## + Triples         1    11.976 240.68 67.329
## + Caught_Stealing  1     8.270 244.38 67.696
## + Double_Plays    1     8.006 244.64 67.722
## + Earned_Runs     1     5.798 246.85 67.938
## + Saves          1     5.293 247.36 67.987
## + At_Bats         1     3.755 248.90 68.136
## + Walks_Allowed   1     2.883 249.77 68.219
## + Hits_Allowed    1     2.607 250.04 68.246
## + Shutout         1     2.555 250.10 68.251
## + Earned_Run_Average 1     1.732 250.92 68.330
## + Hits            1     1.067 251.58 68.393
## + Walks           1     0.398 252.25 68.457
## + Fielding_Percentage 1     0.121 252.53 68.483
## + Home_Runs       1     0.001 252.65 68.495
##
## Step:  AIC=62.42
## Games_Won ~ Runs_Against + Runs_Scored + Errors + Strike-Outs +
##           Doubles
##
##              Df Sum of Sq   RSS   AIC
## + Strikeouts_Allowed 1    33.506 162.67 59.929
## + Complete_Games     1    27.983 168.20 60.730
## <none>                196.18 62.424
## + Home_Run_Allowed   1    13.450 182.73 62.719
## + Walks_Allowed      1    11.615 184.56 62.959
## + Earned_Runs        1     7.738 188.44 63.458
## + Infield_Put-Outs   1     7.073 189.11 63.542
## + Home_Runs          1     5.432 190.75 63.750
## + Triples            1     5.132 191.05 63.787
## + Earned_Run_Average 1     4.980 191.20 63.806

```

```

## + Double_Plays          1      4.585 191.59 63.856
## + Hits_Allowed          1      3.133 193.05 64.037
## + Saves                  1      2.367 193.81 64.132
## + Stolen_Bases          1      1.239 194.94 64.271
## + Walks                  1      0.811 195.37 64.324
## + Shutout                1      0.223 195.96 64.396
## + Caught_Stealing        1      0.119 196.06 64.409
## + At_Bats                1      0.037 196.14 64.419
## + Fielding_Percentage    1      0.020 196.16 64.421
## + Hits                   1      0.014 196.17 64.422
##
## Step:  AIC=59.93
## Games_Won ~ Runs_Against + Runs_Scored + Errors + Strike_Outs +
##      Doubles + Strikeouts_Allowed
##
##              Df Sum of Sq    RSS    AIC
## + Walks_Allowed      1   27.0759 135.60 57.559
## + Shutout             1   14.5177 148.16 59.685
## <none>                  162.67 59.929
## + Triples             1   12.4567 150.22 60.017
## + Complete_Games      1   11.0223 151.65 60.245
## + Earned_Runs          1   10.7548 151.92 60.287
## + Hits_Allowed         1   10.3340 152.34 60.353
## + Earned_Run_Average   1    9.4484 153.23 60.493
## + Saves                1    9.2265 153.45 60.527
## + Double_Plays         1    6.3907 156.28 60.967
## + Infield_Put_Outs     1    4.8059 157.87 61.209
## + Home_Run_Allowed     1    4.3548 158.32 61.277
## + Stolen_Bases         1    3.0471 159.63 61.475
## + Home_Runs            1    2.5379 160.14 61.551
## + Hits                 1    1.0426 161.63 61.774
## + Caught_Stealing      1    0.6840 161.99 61.828
## + Walks                1    0.2258 162.45 61.895
## + Fielding_Percentage  1    0.1894 162.49 61.901
## + At_Bats              1    0.0861 162.59 61.916
##
## Step:  AIC=57.56
## Games_Won ~ Runs_Against + Runs_Scored + Errors + Strike_Outs +
##      Doubles + Strikeouts_Allowed + Walks_Allowed
##
##              Df Sum of Sq    RSS    AIC
## + Home_Run_Allowed     1   14.7339 120.86 56.799
## + Triples              1   11.4117 124.19 57.450
## <none>                  135.60 57.559
## + Earned_Runs          1    6.6476 128.95 58.353
## + Earned_Run_Average   1    6.3129 129.28 58.415
## + Saves                1    6.1126 129.49 58.452
## + Shutout              1    5.8194 129.78 58.507
## + Complete_Games       1    4.7737 130.82 58.699
## + Stolen_Bases         1    4.6943 130.90 58.714
## + Infield_Put_Outs     1    2.1364 133.46 59.178
## + Hits_Allowed         1    1.3395 134.26 59.321
## + Hits                 1    1.2918 134.31 59.330
## + Double_Plays         1    1.0232 134.57 59.378

```

```

## + Home_Runs          1    0.4940 135.10 59.472
## + Walks              1    0.3127 135.29 59.504
## + Caught_Stealing    1    0.2440 135.35 59.516
## + At_Bats            1    0.1983 135.40 59.524
## + Fielding_Percentage 1    0.0878 135.51 59.544
##
## Step: AIC=56.8
## Games_Won ~ Runs_Against + Runs_Scored + Errors + Strike-Outs +
##      Doubles + Strikeouts_Allowed + Walks_Allowed + Home_Run_Allowed
##
##              Df Sum of Sq   RSS   AIC
## <none>                120.86 56.799
## + Saves              1    8.0047 112.86 57.154
## + Home_Runs          1    6.6322 114.23 57.444
## + Stolen_Bases       1    6.1119 114.75 57.553
## + Triples            1    2.9043 117.96 58.215
## + Shutout            1    2.8147 118.05 58.233
## + Infield_Put-Outs   1    1.8888 118.97 58.421
## + Complete_Games     1    1.8499 119.01 58.429
## + Hits_Allowed       1    1.7721 119.09 58.444
## + Hits               1    1.6218 119.24 58.475
## + At_Bats            1    1.0927 119.77 58.581
## + Caught_Stealing    1    0.8605 120.00 58.627
## + Walks              1    0.6093 120.25 58.677
## + Double_Plays       1    0.1850 120.68 58.762
## + Fielding_Percentage 1    0.1666 120.70 58.766
## + Earned_Run_Average 1    0.0267 120.84 58.793
## + Earned_Runs        1    0.0050 120.86 58.798
##
##
## Call:
## lm(formula = Games_Won ~ Runs_Against + Runs_Scored + Errors +
##      Strike-Outs + Doubles + Strikeouts_Allowed + Walks_Allowed +
##      Home_Run_Allowed, data = df_p1)
##
## Coefficients:
##      (Intercept)      Runs_Against      Runs_Scored           Errors
##      81.36756         -0.16765           0.10578         0.13734
##      Strike-Outs      Doubles  Strikeouts_Allowed      Walks_Allowed
##      -0.02958          0.07589          -0.02074          0.02187
##      Home_Run_Allowed
##      0.13804

fit1 <- lm(Games_Won ~ Runs_Against + Runs_Scored + Errors +
      Strike-Outs + Doubles + Strikeouts_Allowed + Walks_Allowed +
      Home_Run_Allowed, data = df_p1)

summary(fit1)

##
## Call:
## lm(formula = Games_Won ~ Runs_Against + Runs_Scored + Errors +
##      Strike-Outs + Doubles + Strikeouts_Allowed + Walks_Allowed +

```

```
## Home_Run_Allowed, data = df_p1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.0654 -0.9247 -0.0054  1.7052  4.8746
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   81.367559  14.364171   5.665 4.49e-05 ***
## Runs_Against  -0.167649   0.015890 -10.550 2.46e-08 ***
## Runs_Scored    0.105783   0.010050  10.526 2.54e-08 ***
## Errors         0.137341   0.031883   4.308 0.000622 ***
## Strike-Outs   -0.029577   0.008255  -3.583 0.002720 **
## Doubles        0.075888   0.027427   2.767 0.014389 *
## Strikeouts_Allowed -0.020736  0.009766  -2.123 0.050771 .
## Walks_Allowed  0.021870   0.010144   2.156 0.047728 *
## Home_Run_Allowed 0.138035   0.102079   1.352 0.196334
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.839 on 15 degrees of freedom
## Multiple R-squared:  0.9764, Adjusted R-squared:  0.9638
## F-statistic: 77.55 on 8 and 15 DF,  p-value: 8.327e-11
```

```
# Period 2 - 1920 to 1960
```

```
df_p2 <- drop.na(df_1920_1960[-c(1, 3)])
```

```
## Forward stepwise regression
```

```
null_modelp2 <- lm(Games_Won ~ 1, data = df_p2)
```

```
full_modelp2 <- lm(Games_Won ~ ., data = df_p2)
```

```
step(null_modelp2, direction = "forward", scope = formula(full_modelp2))
```

```
## Start: AIC=2423.73
```

```
## Games_Won ~ 1
```

```
##
##              Df Sum of Sq  RSS    AIC
## + Runs_Scored    1    38182 54165 2182.4
## + Runs_Against    1    37767 54581 2185.9
## + Earned_Run_Average 1    36414 55933 2197.1
## + Earned_Runs      1    35788 56559 2202.2
## + Shutout          1    23915 68432 2289.1
## + Hits_Allowed     1    22860 69487 2296.0
## + Saves            1    16685 75662 2334.9
## + Hits             1    14275 78072 2349.2
## + Walks            1    13940 78407 2351.1
## + Home_Runs        1    13124 79223 2355.8
## + Errors           1    10085 82263 2373.0
## + Infield_Put-Outs 1     8894 83453 2379.6
## + Fielding_Percentage 1     8757 83590 2380.3
## + Complete_Games   1     8573 83775 2381.3
## + Strikeouts_Allowed 1     7936 84411 2384.8
## + Walks_Allowed    1     5807 86540 2396.1
## + Triples          1     3904 88444 2406.0
```



```
## + Doubles          1      3676 88672 2407.2
## + At_Bats          1      2552 89795 2412.9
## + Stolen_Bases     1      2199 90149 2414.8
## + Home_Run_Allowed 1      1474 90874 2418.4
## <none>              92347 2423.7
## + Strike-Outs      1       325 92023 2424.1
## + Double_Plays      1       170 92178 2424.9
## + Caught_Stealing  1        34 92314 2425.6
##
```

```
## Step: AIC=2182.45
```

```
## Games_Won ~ Runs_Scored
```

```
##
##              Df Sum of Sq  RSS    AIC
## + Runs_Against      1    46327  7838 1303.0
## + Earned_Run_Average 1    43522 10643 1442.5
## + Earned_Runs        1    42898 11267 1468.4
## + Hits_Allowed       1    33983 20182 1734.3
## + Shutout            1    24612 29553 1908.2
## + Saves              1    12155 42010 2068.6
## + Errors             1    12134 42031 2068.8
## + Fielding_Percentage 1    10885 43280 2082.1
## + Infield_Put-Outs   1    10833 43332 2082.7
## + Strikeouts_Allowed 1    10207 43958 2089.2
## + Doubles            1     7450 46714 2117.0
## + Walks_Allowed      1     5977 48188 2131.1
## + Hits               1     3654 50511 2152.6
## + At_Bats            1     3488 50677 2154.1
## + Home_Runs          1     3182 50983 2156.8
## + Complete_Games     1     1107 53058 2175.0
## + Strike-Outs        1     1057 53108 2175.5
## + Walks              1      903 53262 2176.8
## + Triples            1      595 53570 2179.4
## + Home_Run_Allowed   1      441 53724 2180.7
## + Double_Plays        1      262 53903 2182.2
## <none>              54165 2182.4
## + Stolen_Bases       1        47 54118 2184.1
## + Caught_Stealing    1        18 54147 2184.3
##
```

```
## Step: AIC=1302.96
```

```
## Games_Won ~ Runs_Scored + Runs_Against
```

```
##
##              Df Sum of Sq  RSS    AIC
## + Saves            1    416.16 7421.5 1280.1
## + Infield_Put-Outs 1    383.12 7454.5 1282.1
## + Strikeouts_Allowed 1     73.68 7764.0 1300.7
## <none>              7837.7 1303.0
## + Earned_Run_Average 1     28.38 7809.3 1303.3
## + Double_Plays      1     26.54 7811.1 1303.4
## + Home_Run_Allowed   1     22.04 7815.6 1303.7
## + Fielding_Percentage 1     20.12 7817.5 1303.8
## + Shutout           1     16.51 7821.2 1304.0
## + Walks             1     15.27 7822.4 1304.1
## + Errors            1     11.68 7826.0 1304.3
## + Doubles           1     11.15 7826.5 1304.3
```

```

## + Home_Runs          1      10.60 7827.1 1304.3
## + Stolen_Bases       1       9.66 7828.0 1304.4
## + Walks_Allowed      1       2.86 7834.8 1304.8
## + Hits_Allowed       1       2.57 7835.1 1304.8
## + Caught_Stealing    1       1.21 7836.5 1304.9
## + Strike-Outs        1       1.18 7836.5 1304.9
## + Complete_Games     1       0.71 7837.0 1304.9
## + Hits               1       0.16 7837.5 1305.0
## + Triples            1       0.15 7837.5 1305.0
## + Earned_Runs        1       0.15 7837.5 1305.0
## + At_Bats            1       0.11 7837.5 1305.0
##
## Step:  AIC=1280.08
## Games_Won ~ Runs_Scored + Runs_Against + Saves
##
##              Df Sum of Sq   RSS   AIC
## + Complete_Games      1    689.30 6732.2 1237.6
## + Infield_Put-Outs    1    304.75 7116.7 1263.0
## + Earned_Run_Average  1    174.20 7247.3 1271.2
## + Strike-Outs         1    158.63 7262.9 1272.2
## + Caught_Stealing     1    124.46 7297.0 1274.4
## + Home_Runs           1     92.34 7329.2 1276.4
## + Walks_Allowed       1     90.19 7331.3 1276.5
## + Hits_Allowed        1     89.29 7332.2 1276.6
## + Triples             1     86.47 7335.0 1276.7
## + Hits                1     76.31 7345.2 1277.4
## + Home_Run_Allowed    1     76.05 7345.5 1277.4
## + Earned_Runs         1     71.64 7349.9 1277.7
## + Errors              1     59.36 7362.1 1278.4
## <none>                7421.5 1280.1
## + Shutout             1     22.78 7398.7 1280.7
## + Stolen_Bases        1     17.76 7403.7 1281.0
## + Strikeouts_Allowed  1     17.57 7403.9 1281.0
## + Fielding_Percentage 1     11.09 7410.4 1281.4
## + Doubles             1      4.77 7416.7 1281.8
## + Walks               1      4.16 7417.3 1281.8
## + At_Bats             1      1.73 7419.8 1282.0
## + Double_Plays        1      0.54 7421.0 1282.0
##
## Step:  AIC=1237.63
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games
##
##              Df Sum of Sq   RSS   AIC
## + Infield_Put-Outs    1   292.396 6439.8 1219.4
## + Strikeouts_Allowed  1    42.230 6690.0 1236.8
## + Walks_Allowed       1    32.507 6699.7 1237.4
## <none>                6732.2 1237.6
## + Fielding_Percentage 1    23.465 6708.7 1238.0
## + Earned_Run_Average  1    22.619 6709.6 1238.1
## + Home_Run_Allowed    1    14.734 6717.5 1238.6
## + Hits_Allowed        1    13.746 6718.5 1238.7
## + Hits                1    10.734 6721.5 1238.9
## + Caught_Stealing     1    10.447 6721.8 1238.9
## + Shutout             1      6.761 6725.4 1239.2

```

```

## + Double_Plays      1      6.421 6725.8 1239.2
## + Triples           1      4.428 6727.8 1239.3
## + Stolen_Bases      1      4.249 6728.0 1239.3
## + At_Bats           1      3.992 6728.2 1239.4
## + Home_Runs         1      2.604 6729.6 1239.5
## + Errors            1      2.532 6729.7 1239.5
## + Strike-Outs       1      2.510 6729.7 1239.5
## + Doubles           1      2.056 6730.1 1239.5
## + Earned_Runs       1      1.077 6731.1 1239.5
## + Walks             1      0.058 6732.1 1239.6
##
## Step:  AIC=1219.38
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games +
##      Infield_Put-Outs
##
##              Df Sum of Sq    RSS    AIC
## + At_Bats      1    163.443 6276.4 1209.7
## <none>                      6439.8 1219.4
## + Walks_Allowed  1     25.857 6413.9 1219.5
## + Stolen_Bases  1     18.257 6421.5 1220.1
## + Fielding_Percentage 1    15.845 6424.0 1220.3
## + Home_Run_Allowed  1    12.505 6427.3 1220.5
## + Strikeouts_Allowed 1    10.461 6429.3 1220.6
## + Double_Plays   1      7.795 6432.0 1220.8
## + Strike-Outs    1      6.825 6433.0 1220.9
## + Errors         1      6.116 6433.7 1221.0
## + Earned_Run_Average 1      3.760 6436.0 1221.1
## + Home_Runs      1      3.719 6436.1 1221.1
## + Shutout        1      2.670 6437.1 1221.2
## + Earned_Runs    1      2.253 6437.6 1221.2
## + Doubles        1      1.483 6438.3 1221.3
## + Triples        1      1.281 6438.5 1221.3
## + Hits           1      0.612 6439.2 1221.3
## + Caught_Stealing 1      0.358 6439.4 1221.3
## + Walks          1      0.225 6439.6 1221.4
## + Hits_Allowed   1      0.010 6439.8 1221.4
##
## Step:  AIC=1209.66
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games +
##      Infield_Put-Outs + At_Bats
##
##              Df Sum of Sq    RSS    AIC
## + Hits          1     43.965 6232.4 1208.5
## + Stolen_Bases  1     39.674 6236.7 1208.8
## <none>                      6276.4 1209.7
## + Walks_Allowed  1     23.665 6252.7 1209.9
## + Double_Plays   1     22.435 6253.9 1210.0
## + Walks          1     20.196 6256.2 1210.2
## + Fielding_Percentage 1    14.585 6261.8 1210.6
## + Strike-Outs    1     13.887 6262.5 1210.7
## + Strikeouts_Allowed 1    10.562 6265.8 1210.9
## + Caught_Stealing 1      8.183 6268.2 1211.1
## + Home_Run_Allowed  1      7.753 6268.6 1211.1
## + Earned_Run_Average 1      7.123 6269.2 1211.1

```

```

## + Errors          1      5.493 6270.9 1211.3
## + Triples         1      5.102 6271.3 1211.3
## + Earned_Runs     1      4.823 6271.5 1211.3
## + Doubles         1      3.467 6272.9 1211.4
## + Shutout         1      2.972 6273.4 1211.4
## + Home_Runs       1      1.857 6274.5 1211.5
## + Hits_Allowed    1      0.088 6276.3 1211.7
##
## Step:  AIC=1208.45
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games +
##      Infield_Put_Outs + At_Bats + Hits
##
##              Df Sum of Sq    RSS    AIC
## + Stolen_Bases      1      96.187 6136.2 1203.4
## + Caught_Stealing   1      77.000 6155.4 1204.8
## + Double_Plays      1      40.656 6191.7 1207.5
## + Strikeouts_Allowed 1      39.860 6192.5 1207.5
## + Home_Run_Allowed  1      38.412 6194.0 1207.6
## + Fielding_Percentage 1      35.425 6197.0 1207.8
## + Earned_Run_Average 1      34.466 6197.9 1207.9
## + Errors            1      33.051 6199.3 1208.0
## + Home_Runs         1      30.529 6201.9 1208.2
## + Earned_Runs       1      28.335 6204.1 1208.4
## <none>                6232.4 1208.5
## + Hits_Allowed      1      9.797 6222.6 1209.7
## + Shutout           1      5.614 6226.8 1210.0
## + Walks_Allowed     1      5.181 6227.2 1210.1
## + Triples           1      1.181 6231.2 1210.4
## + Walks             1      0.477 6231.9 1210.4
## + Strike_Outs       1      0.014 6232.4 1210.5
## + Doubles           1      0.011 6232.4 1210.5
##
## Step:  AIC=1203.36
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games +
##      Infield_Put_Outs + At_Bats + Hits + Stolen_Bases
##
##              Df Sum of Sq    RSS    AIC
## + Double_Plays      1      36.413 6099.8 1202.6
## <none>                6136.2 1203.4
## + Strikeouts_Allowed 1      25.042 6111.2 1203.5
## + Fielding_Percentage 1      21.405 6114.8 1203.8
## + Earned_Run_Average 1      21.178 6115.0 1203.8
## + Home_Run_Allowed  1      21.063 6115.1 1203.8
## + Earned_Runs       1      15.947 6120.3 1204.2
## + Errors            1      12.721 6123.5 1204.4
## + Walks_Allowed     1      12.657 6123.6 1204.4
## + Caught_Stealing   1      11.662 6124.5 1204.5
## + Home_Runs         1      7.135 6129.1 1204.8
## + Hits_Allowed      1      5.776 6130.4 1204.9
## + Walks             1      5.268 6130.9 1205.0
## + Triples           1      3.635 6132.6 1205.1
## + Shutout           1      3.518 6132.7 1205.1
## + Strike_Outs       1      0.913 6135.3 1205.3
## + Doubles           1      0.067 6136.1 1205.3

```

```

##
## Step: AIC=1202.64
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games +
##      Infield_Put_Outs + At_Bats + Hits + Stolen_Bases + Double_Plays
##
##      Df Sum of Sq    RSS    AIC
## + Strikeouts_Allowed  1    47.934 6051.9 1201.0
## + Walks_Allowed      1    37.002 6062.8 1201.9
## <none>                6099.8 1202.6
## + Home_Run_Allowed   1    21.410 6078.4 1203.0
## + Walks              1    14.199 6085.6 1203.6
## + Fielding_Percentage 1    12.724 6087.1 1203.7
## + Earned_Run_Average  1    11.701 6088.1 1203.8
## + Caught_Stealing     1    11.586 6088.2 1203.8
## + Home_Runs           1     9.291 6090.5 1204.0
## + Earned_Runs         1     7.827 6092.0 1204.1
## + Errors              1     5.852 6093.9 1204.2
## + Hits_Allowed       1     5.153 6094.6 1204.3
## + Triples            1     4.995 6094.8 1204.3
## + Shutout            1     2.711 6097.1 1204.4
## + Strike_Outs        1     0.060 6099.7 1204.6
## + Doubles            1     0.009 6099.8 1204.6
##
## Step: AIC=1201.05
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games +
##      Infield_Put_Outs + At_Bats + Hits + Stolen_Bases + Double_Plays +
##      Strikeouts_Allowed
##
##      Df Sum of Sq    RSS    AIC
## + Walks_Allowed      1    52.354 5999.5 1199.1
## <none>                6051.9 1201.0
## + Walks              1    14.216 6037.6 1202.0
## + Triples            1    13.733 6038.1 1202.0
## + Strike_Outs        1     8.187 6043.7 1202.4
## + Fielding_Percentage 1     7.947 6043.9 1202.5
## + Caught_Stealing     1     4.795 6047.1 1202.7
## + Earned_Run_Average  1     4.566 6047.3 1202.7
## + Home_Run_Allowed   1     2.448 6049.4 1202.9
## + Earned_Runs         1     2.197 6049.7 1202.9
## + Shutout            1     1.658 6050.2 1202.9
## + Errors              1     0.736 6051.1 1203.0
## + Doubles            1     0.378 6051.5 1203.0
## + Home_Runs           1     0.279 6051.6 1203.0
## + Hits_Allowed       1     0.009 6051.9 1203.0
##
## Step: AIC=1199.08
## Games_Won ~ Runs_Scored + Runs_Against + Saves + Complete_Games +
##      Infield_Put_Outs + At_Bats + Hits + Stolen_Bases + Double_Plays +
##      Strikeouts_Allowed + Walks_Allowed
##
##      Df Sum of Sq    RSS    AIC
## <none>                5999.5 1199.1
## + Hits_Allowed       1    24.2893 5975.2 1199.2
## + Strike_Outs        1    20.3796 5979.1 1199.5

```

```
## + Triples          1    13.3995 5986.1 1200.1
## + Fielding_Percentage 1    11.0138 5988.5 1200.2
## + Earned_Run_Average 1     9.8395 5989.7 1200.3
## + Caught_Stealing    1     8.3856 5991.1 1200.5
## + Earned_Runs        1     5.9193 5993.6 1200.6
## + Shutout           1     3.2306 5996.3 1200.8
## + Errors            1     2.1544 5997.4 1200.9
## + Walks             1     2.1237 5997.4 1200.9
## + Home_Runs          1     0.9994 5998.5 1201.0
## + Home_Run_Allowed   1     0.7924 5998.7 1201.0
## + Doubles           1     0.7431 5998.8 1201.0

##
## Call:
## lm(formula = Games_Won ~ Runs_Scored + Runs_Against + Saves +
##     Complete_Games + Infield_Put_Outs + At_Bats + Hits + Stolen_Bases +
##     Double_Plays + Strikeouts_Allowed + Walks_Allowed, data = df_p2)
##
## Coefficients:
##      (Intercept)      Runs_Scored      Runs_Against          Saves
##      0.956490         0.083433         -0.069990         0.317535
## Complete_Games  Infield_Put_Outs          At_Bats          Hits
##      0.144090         0.029437         -0.017327         0.013152
## Stolen_Bases    Double_Plays  Strikeouts_Allowed  Walks_Allowed
##      -0.017152         0.028312         0.004682        -0.005361
```

```
fit2 <- lm(Games_Won ~ Runs_Scored + Runs_Against + Saves +
            Complete_Games + Infield_Put_Outs + At_Bats + Hits + Stolen_Bases +
            Double_Plays + Strikeouts_Allowed + Walks_Allowed, data = df_p2)

summary(fit2)
```

```
##
## Call:
## lm(formula = Games_Won ~ Runs_Scored + Runs_Against + Saves +
##     Complete_Games + Infield_Put_Outs + At_Bats + Hits + Stolen_Bases +
##     Double_Plays + Strikeouts_Allowed + Walks_Allowed, data = df_p2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9118  -2.5259  -0.0583   2.4852  10.0460
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.956490   14.014307   0.068  0.94562
## Runs_Scored     0.083433    0.003461  24.108 < 2e-16 ***
## Runs_Against   -0.069990    0.003111 -22.498 < 2e-16 ***
## Saves          0.317535    0.042249   7.516 3.15e-13 ***
## Complete_Games 0.144090    0.020926   6.886 1.97e-11 ***
## Infield_Put_Outs 0.029437    0.004597   6.404 3.86e-10 ***
## At_Bats        -0.017327    0.003507  -4.941 1.11e-06 ***
## Hits           0.013152    0.004541   2.896 0.00396 **
## Stolen_Bases   -0.017152    0.006934  -2.474 0.01375 *
```

```
## Double_Plays      0.028312  0.010543  2.685  0.00751 **
## Strikeouts_Allowed 0.004682  0.002163  2.164  0.03099 *
## Walks_Allowed     -0.005361  0.002723  -1.968  0.04965 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.676 on 444 degrees of freedom
## Multiple R-squared:  0.935, Adjusted R-squared:  0.9334
## F-statistic: 580.9 on 11 and 444 DF,  p-value: < 2.2e-16
```

```
# Period 3 - 1960 to 1990
```

```
df_p3 <- drop.na(df_1960_1990[-c(1, 3)])
```

```
## Forward stepwise regression
```

```
null_modelp3 <- lm(Games_Won ~ 1, data = df_p3)
```

```
full_modelp3 <- lm(Games_Won ~ ., data = df_p3)
```

```
step(null_modelp3, direction = "forward", scope = formula(full_modelp3))
```

```
## Start:  AIC=3701.27
```

```
## Games_Won ~ 1
```

```
##
```

	Df	Sum of Sq	RSS	AIC
## + Runs_Scored	1	52361	63546	3264.5
## + Hits	1	38074	77832	3412.6
## + Saves	1	29214	86692	3491.3
## + Earned_Run_Average	1	26316	89590	3515.3
## + Infield_Put_Outs	1	24954	90952	3526.3
## + At_Bats	1	24040	91866	3533.6
## + Shutout	1	23606	92300	3537.0
## + Home_Runs	1	21388	94518	3554.4
## + Doubles	1	19278	96629	3570.5
## + Walks	1	18279	97627	3578.0
## + Strikeouts_Allowed	1	14014	101892	3609.2
## + Complete_Games	1	10428	105478	3634.4
## + Fielding_Percentage	1	7263	108643	3656.0
## + Runs_Against	1	7207	108699	3656.4
## + Earned_Runs	1	5987	109919	3664.6
## + Triples	1	5965	109941	3664.7
## + Stolen_Bases	1	4495	111412	3674.4
## + Walks_Allowed	1	3129	112777	3683.3
## + Double_Plays	1	2898	113008	3684.8
## + Errors	1	2546	113360	3687.1
## + Strike_Outs	1	2096	113810	3689.9
## <none>			115906	3701.3
## + Caught_Stealing	1	171	115736	3702.2
## + Hits_Allowed	1	122	115784	3702.5
## + Home_Run_Allowed	1	16	115890	3703.2

```
##
```

```
## Step:  AIC=3264.52
```

```
## Games_Won ~ Runs_Scored
```

```
##
```

	Df	Sum of Sq	RSS	AIC
## + Earned_Run_Average	1	50127	13418	2131.3

```

## + Earned_Runs      1      40039 23507 2540.6
## + Runs_Against     1      40010 23536 2541.5
## + Shutout          1      29204 34341 2817.3
## + Home_Run_Allowed 1      17221 46324 3035.8
## + Hits_Allowed     1      16290 47256 3050.3
## + Walks_Allowed    1      10867 52679 3129.6
## + Saves            1       8309 55236 3164.2
## + Complete_Games   1       6395 57150 3189.1
## + Strikeouts_Allowed 1      4826 58720 3208.9
## + Fielding_Percentage 1     3948 59598 3219.7
## + Errors           1     3590 59956 3224.1
## + Infield_Put_Outs 1     2242 61304 3240.3
## + Doubles          1       781 62765 3257.5
## + Stolen_Bases     1       764 62782 3257.7
## + Double_Plays     1       647 62899 3259.1
## + Home_Runs        1       424 63122 3261.6
## + At_Bats          1       291 63255 3263.2
## + Triples          1       182 63363 3264.4
## <none>              1     63546 3264.5
## + Strike_Outs      1       157 63389 3264.7
## + Walks            1        45 63501 3266.0
## + Hits             1         9 63537 3266.4
## + Caught_Stealing  1         8 63538 3266.4
##
## Step:  AIC=2131.26
## Games_Won ~ Runs_Scored + Earned_Run_Average
##
##              Df Sum of Sq  RSS    AIC
## + Saves      1    1510.98 11907 2046.0
## + Fielding_Percentage 1    1081.11 12337 2071.9
## + Errors      1     833.80 12584 2086.4
## + Shutout     1     471.27 12947 2107.2
## + Infield_Put_Outs 1      96.99 13321 2128.0
## + Strikeouts_Allowed 1      86.41 13332 2128.6
## + Earned_Runs  1      70.98 13347 2129.4
## + Runs_Against 1      54.61 13364 2130.3
## + Home_Run_Allowed 1      50.81 13367 2130.5
## + Triples     1      40.84 13377 2131.0
## <none>         1     13418 2131.3
## + Stolen_Bases 1      27.97 13390 2131.7
## + Walks        1      23.34 13395 2132.0
## + At_Bats      1      20.99 13397 2132.1
## + Double_Plays 1       8.27 13410 2132.8
## + Home_Runs    1       6.37 13412 2132.9
## + Hits         1       6.13 13412 2132.9
## + Strike_Outs  1       4.47 13414 2133.0
## + Doubles      1       3.22 13415 2133.1
## + Hits_Allowed 1       0.84 13417 2133.2
## + Complete_Games 1       0.74 13417 2133.2
## + Caught_Stealing 1       0.23 13418 2133.2
## + Walks_Allowed 1       0.12 13418 2133.3
##
## Step:  AIC=2046.05
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves

```



```

##
##              Df Sum of Sq    RSS    AIC
## + Complete_Games      1   1820.91 10086 1926.9
## + Shutout              1    794.62 11112 1997.6
## + Fielding_Percentage  1    751.49 11156 2000.5
## + Errors               1    510.61 11396 2016.1
## + Infield_Put_Outs     1     83.48 11824 2042.9
## + Earned_Runs          1     58.46 11849 2044.5
## + Walks                1     53.30 11854 2044.8
## + Home_Run_Allowed     1     48.16 11859 2045.1
## + Doubles              1     38.05 11869 2045.7
## <none>                  1    11907 2046.0
## + Caught_Stealing      1     29.48 11878 2046.2
## + Strikeouts_Allowed   1     28.94 11878 2046.3
## + Runs_Against         1     28.88 11878 2046.3
## + Triples              1     24.24 11883 2046.6
## + At_Bats              1     20.50 11887 2046.8
## + Double_Plays         1     13.15 11894 2047.2
## + Stolen_Bases         1     12.46 11895 2047.3
## + Strike_Outs          1     10.87 11896 2047.4
## + Hits                 1      8.19 11899 2047.5
## + Home_Runs            1      5.07 11902 2047.7
## + Hits_Allowed         1      2.43 11905 2047.9
## + Walks_Allowed        1      2.00 11905 2047.9
##
## Step:  AIC=1926.89
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games
##
##              Df Sum of Sq    RSS    AIC
## + Fielding_Percentage  1    797.27  9289.0 1868.8
## + Errors               1    672.45  9413.8 1878.5
## + Shutout              1    334.57  9751.7 1904.3
## + Runs_Against         1    131.67  9954.6 1919.3
## + Stolen_Bases         1     48.01 10038.2 1925.4
## + Walks                1     43.44 10042.8 1925.7
## + Triples              1     28.65 10057.6 1926.8
## <none>                  1    10086.2 1926.9
## + Strikeouts_Allowed   1     16.96 10069.3 1927.7
## + Hits_Allowed         1     16.73 10069.5 1927.7
## + Strike_Outs          1     13.59 10072.7 1927.9
## + Walks_Allowed        1     12.72 10073.5 1928.0
## + At_Bats              1      8.62 10077.6 1928.3
## + Doubles              1      8.01 10078.2 1928.3
## + Double_Plays         1      7.49 10078.8 1928.3
## + Home_Runs            1      4.10 10082.1 1928.6
## + Hits                 1      2.14 10084.1 1928.7
## + Infield_Put_Outs     1      1.32 10084.9 1928.8
## + Home_Run_Allowed     1      0.78 10085.5 1928.8
## + Caught_Stealing      1      0.77 10085.5 1928.8
## + Earned_Runs          1      0.00 10086.2 1928.9
##
## Step:  AIC=1868.78
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##      Fielding_Percentage

```

```

##
##           Df Sum of Sq    RSS    AIC
## + Shutout      1   305.360 8983.6 1846.4
## + Errors       1    72.260 9216.7 1865.1
## + Stolen_Bases  1    58.702 9230.3 1866.2
## + Walks        1    49.782 9239.2 1866.9
## + Strikeouts_Allowed 1    40.627 9248.3 1867.6
## <none>                9289.0 1868.8
## + Infield_Put_Outs 1    22.913 9266.1 1869.0
## + Home_Runs     1    19.992 9269.0 1869.2
## + Earned_Runs   1    11.482 9277.5 1869.9
## + Runs_Against  1     9.921 9279.1 1870.0
## + Caught_Stealing 1     9.503 9279.5 1870.0
## + Double_Plays  1     8.477 9280.5 1870.1
## + Home_Run_Allowed 1     5.541 9283.4 1870.3
## + Triples       1     4.858 9284.1 1870.4
## + Hits          1     1.393 9287.6 1870.7
## + At_Bats       1     0.821 9288.2 1870.7
## + Hits_Allowed  1     0.588 9288.4 1870.7
## + Strike_Outs   1     0.420 9288.6 1870.8
## + Walks_Allowed 1     0.388 9288.6 1870.8
## + Doubles       1     0.064 9288.9 1870.8
##
## Step:  AIC=1846.38
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##      Fielding_Percentage + Shutout
##
##           Df Sum of Sq    RSS    AIC
## + Errors      1    87.712 8895.9 1841.2
## + Walks       1    53.755 8929.9 1844.0
## + Stolen_Bases 1    52.846 8930.8 1844.1
## + Runs_Against 1    31.028 8952.6 1845.9
## <none>                8983.6 1846.4
## + Double_Plays 1    17.838 8965.8 1846.9
## + Home_Runs     1    12.574 8971.0 1847.4
## + Strikeouts_Allowed 1    11.862 8971.7 1847.4
## + Strike_Outs   1     8.752 8974.9 1847.7
## + Caught_Stealing 1     7.860 8975.8 1847.7
## + Triples       1     4.858 8978.8 1848.0
## + Home_Run_Allowed 1     4.478 8979.1 1848.0
## + At_Bats       1     4.359 8979.3 1848.0
## + Infield_Put_Outs 1     3.797 8979.8 1848.1
## + Doubles       1     3.022 8980.6 1848.1
## + Hits          1     2.637 8981.0 1848.2
## + Hits_Allowed  1     1.022 8982.6 1848.3
## + Earned_Runs   1     0.421 8983.2 1848.3
## + Walks_Allowed 1     0.194 8983.4 1848.4
##
## Step:  AIC=1841.22
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##      Fielding_Percentage + Shutout + Errors
##
##           Df Sum of Sq    RSS    AIC
## + Infield_Put_Outs 1    75.223 8820.7 1837.0

```

```

## + Walks          1      75.009 8820.9 1837.0
## + Stolen_Bases   1      59.342 8836.6 1838.3
## + Earned_Runs    1      50.790 8845.1 1839.0
## + Strikeouts_Allowed 1      28.275 8867.6 1840.9
## + Hits_Allowed   1      26.121 8869.8 1841.1
## <none>           1      8895.9 1841.2
## + Home_Runs      1      17.129 8878.8 1841.8
## + Caught_Stealing 1      15.455 8880.4 1842.0
## + At_Bats        1      13.824 8882.1 1842.1
## + Double_Plays   1       5.932 8890.0 1842.7
## + Walks_Allowed  1       4.036 8891.9 1842.9
## + Hits           1       3.360 8892.5 1842.9
## + Home_Run_Allowed 1       2.977 8892.9 1843.0
## + Doubles        1       2.706 8893.2 1843.0
## + Triples        1       1.974 8893.9 1843.1
## + Runs_Against   1       1.556 8894.3 1843.1
## + Strike-Outs    1       0.024 8895.9 1843.2
##
## Step:  AIC=1837.02
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##      Fielding_Percentage + Shutout + Errors + Infield_Put_Outs
##
##              Df Sum of Sq    RSS    AIC
## + At_Bats      1    376.40 8444.3 1807.2
## + Runs_Against 1    323.09 8497.6 1811.8
## + Earned_Runs   1     98.33 8722.3 1830.8
## + Stolen_Bases  1     49.98 8770.7 1834.9
## + Walks         1     48.48 8772.2 1835.0
## <none>          1    8820.7 1837.0
## + Double_Plays  1     23.10 8797.6 1837.1
## + Home_Run_Allowed 1     21.56 8799.1 1837.2
## + Strike-Outs   1     21.28 8799.4 1837.3
## + Home_Runs     1     14.24 8806.4 1837.8
## + Hits          1     11.64 8809.0 1838.1
## + Doubles       1     11.21 8809.5 1838.1
## + Caught_Stealing 1      8.05 8812.6 1838.3
## + Walks_Allowed 1      4.29 8816.4 1838.7
## + Hits_Allowed  1      3.19 8817.5 1838.8
## + Triples       1      2.61 8818.1 1838.8
## + Strikeouts_Allowed 1      2.02 8818.7 1838.8
##
## Step:  AIC=1807.19
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##      Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
##      At_Bats
##
##              Df Sum of Sq    RSS    AIC
## + Hits          1    397.36 8046.9 1774.0
## + Runs_Against  1    296.50 8147.8 1783.1
## + Earned_Runs    1     82.39 8361.9 1802.0
## + Strike-Outs    1     68.32 8376.0 1803.3
## + Home_Run_Allowed 1     57.68 8386.6 1804.2
## + Walks          1     43.69 8400.6 1805.4
## + Home_Runs      1     40.11 8404.2 1805.7

```

```

## <none>                                8444.3 1807.2
## + Walks_Allowed          1      23.01 8421.3 1807.2
## + Stolen_Bases           1      10.04 8434.2 1808.3
## + Doubles                 1       9.74 8434.5 1808.3
## + Triples                 1       7.91 8436.4 1808.5
## + Double_Plays           1       7.91 8436.4 1808.5
## + Caught_Stealing        1       5.14 8439.1 1808.7
## + Hits_Allowed           1       5.01 8439.3 1808.8
## + Strikeouts_Allowed     1       0.00 8444.3 1809.2
##
## Step: AIC=1774
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##   Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
##   At_Bats + Hits
##
##              Df Sum of Sq    RSS    AIC
## + Runs_Against      1   230.709 7816.2 1754.8
## + Caught_Stealing    1   227.099 7819.8 1755.1
## + Earned_Runs        1    78.961 7968.0 1768.8
## + Home_Runs          1    40.088 8006.8 1772.3
## + Stolen_Bases       1    23.679 8023.2 1773.8
## + Strikeouts_Allowed 1    23.390 8023.5 1773.9
## <none>                                8046.9 1774.0
## + Double_Plays       1    21.172 8025.7 1774.1
## + Walks              1    18.503 8028.4 1774.3
## + Walks_Allowed      1    18.186 8028.7 1774.3
## + Strike_Outs        1    14.113 8032.8 1774.7
## + Hits_Allowed       1     9.383 8037.5 1775.2
## + Doubles            1     6.058 8040.9 1775.5
## + Home_Run_Allowed   1     0.659 8046.3 1775.9
## + Triples            1     0.322 8046.6 1776.0
##
## Step: AIC=1754.76
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##   Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
##   At_Bats + Hits + Runs_Against
##
##              Df Sum of Sq    RSS    AIC
## + Caught_Stealing    1   209.253 7606.9 1737.0
## + Home_Runs          1    42.338 7773.9 1752.8
## + Strikeouts_Allowed 1    29.535 7786.7 1754.0
## + Double_Plays       1    27.375 7788.8 1754.2
## + Stolen_Bases       1    25.918 7790.3 1754.3
## <none>                                7816.2 1754.8
## + Earned_Runs        1    17.921 7798.3 1755.1
## + Strike_Outs        1    13.113 7803.1 1755.5
## + Hits_Allowed       1     8.210 7808.0 1756.0
## + Walks              1     8.036 7808.2 1756.0
## + Doubles            1     6.212 7810.0 1756.2
## + Walks_Allowed      1     5.420 7810.8 1756.3
## + Home_Run_Allowed   1     4.546 7811.7 1756.3
## + Triples            1     0.283 7815.9 1756.7
##
## Step: AIC=1736.95

```

```

## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##   Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
##   At_Bats + Hits + Runs_Against + Caught_Stealing
##
##
##      Df Sum of Sq    RSS    AIC
## + Double_Plays      1    36.627 7570.3 1735.4
## + Stolen_Bases      1    32.807 7574.1 1735.8
## + Strike_Outs       1    31.258 7575.7 1736.0
## + Walks             1    28.333 7578.6 1736.2
## + Home_Runs         1    21.871 7585.1 1736.8
## <none>                      7606.9 1737.0
## + Strikeouts_Allowed 1    18.045 7588.9 1737.2
## + Earned_Runs        1    15.724 7591.2 1737.4
## + Hits_Allowed       1     8.667 7598.3 1738.1
## + Walks_Allowed      1     6.379 7600.6 1738.3
## + Triples            1     2.191 7604.8 1738.7
## + Home_Run_Allowed   1     1.057 7605.9 1738.8
## + Doubles            1     0.138 7606.8 1738.9
##
## Step:  AIC=1735.43
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##   Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
##   At_Bats + Hits + Runs_Against + Caught_Stealing + Double_Plays
##
##
##      Df Sum of Sq    RSS    AIC
## + Strike_Outs      1   28.6309 7541.7 1734.7
## + Walks             1   25.4813 7544.8 1735.0
## + Stolen_Bases      1   23.1543 7547.2 1735.2
## <none>                      7570.3 1735.4
## + Home_Runs         1   19.5777 7550.7 1735.5
## + Earned_Runs        1   16.6367 7553.7 1735.8
## + Strikeouts_Allowed 1    4.6051 7565.7 1737.0
## + Hits_Allowed       1    3.3837 7566.9 1737.1
## + Walks_Allowed      1    2.8961 7567.4 1737.2
## + Triples            1    2.7085 7567.6 1737.2
## + Doubles            1    0.8205 7569.5 1737.3
## + Home_Run_Allowed   1    0.2829 7570.0 1737.4
##
## Step:  AIC=1734.66
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
##   Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
##   At_Bats + Hits + Runs_Against + Caught_Stealing + Double_Plays +
##   Strike_Outs
##
##
##      Df Sum of Sq    RSS    AIC
## + Walks             1   24.7409 7517.0 1734.3
## + Stolen_Bases      1   23.6532 7518.0 1734.4
## <none>                      7541.7 1734.7
## + Earned_Runs        1   16.8517 7524.8 1735.0
## + Home_Runs         1   12.1437 7529.5 1735.5
## + Walks_Allowed      1    4.4441 7537.2 1736.2
## + Hits_Allowed       1    1.3384 7540.4 1736.5
## + Triples            1    1.2032 7540.5 1736.5
## + Doubles            1    0.5432 7541.1 1736.6

```

```

## + Strikeouts_Allowed 1 0.3847 7541.3 1736.6
## + Home_Run_Allowed 1 0.0044 7541.7 1736.7
##
## Step: AIC=1734.27
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
## Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
## At_Bats + Hits + Runs_Against + Caught_Stealing + Double_Plays +
## Strike_Outs + Walks
##
## Df Sum of Sq RSS AIC
## + Stolen_Bases 1 29.1793 7487.8 1733.4
## <none> 7517.0 1734.3
## + Earned_Runs 1 14.4245 7502.5 1734.9
## + Home_Runs 1 5.0675 7511.9 1735.8
## + Walks_Allowed 1 2.2676 7514.7 1736.0
## + Hits_Allowed 1 0.9549 7516.0 1736.2
## + Doubles 1 0.8624 7516.1 1736.2
## + Home_Run_Allowed 1 0.8151 7516.1 1736.2
## + Strikeouts_Allowed 1 0.3513 7516.6 1736.2
## + Triples 1 0.2608 7516.7 1736.2
##
## Step: AIC=1733.43
## Games_Won ~ Runs_Scored + Earned_Run_Average + Saves + Complete_Games +
## Fielding_Percentage + Shutout + Errors + Infield_Put_Outs +
## At_Bats + Hits + Runs_Against + Caught_Stealing + Double_Plays +
## Strike_Outs + Walks + Stolen_Bases
##
## Df Sum of Sq RSS AIC
## <none> 7487.8 1733.4
## + Home_Runs 1 17.4191 7470.4 1733.7
## + Earned_Runs 1 14.2028 7473.6 1734.0
## + Strikeouts_Allowed 1 2.6140 7485.2 1735.2
## + Walks_Allowed 1 2.0987 7485.7 1735.2
## + Doubles 1 1.8371 7485.9 1735.2
## + Hits_Allowed 1 0.9245 7486.8 1735.3
## + Home_Run_Allowed 1 0.1328 7487.6 1735.4
## + Triples 1 0.0773 7487.7 1735.4
##
##
## Call:
## lm(formula = Games_Won ~ Runs_Scored + Earned_Run_Average + Saves +
## Complete_Games + Fielding_Percentage + Shutout + Errors +
## Infield_Put_Outs + At_Bats + Hits + Runs_Against + Caught_Stealing +
## Double_Plays + Strike_Outs + Walks + Stolen_Bases, data = df_p3)
##
## Coefficients:
## (Intercept) Runs_Scored Earned_Run_Average
## -1.001e+02 8.567e-02 -2.275e+00
## Saves Complete_Games Fielding_Percentage
## 2.955e-01 1.406e-01 1.105e+02
## Shutout Errors Infield_Put_Outs
## 1.703e-01 -1.241e-03 5.703e-02
## At_Bats Hits Runs_Against
## -4.673e-02 3.857e-02 -5.050e-02

```

```
##      Caught_Stealing      Double_Plays      Strike_Outs
##      -5.678e-02      -9.962e-03      2.472e-03
##      Walks      Stolen_Bases
##      -4.602e-03      6.869e-03
```

```
fit3 <- lm(Games_Won ~ Runs_Scored + Earned_Run_Average + Saves +
           Complete_Games + Fielding_Percentage + Shutout + Errors +
           Infield_Put_Outs + At_Bats + Hits + Runs_Against + Caught_Stealing +
           Double_Plays + Strike_Outs + Walks + Stolen_Bases, data = df_p3)

summary(fit3)
```

```
##
## Call:
## lm(formula = Games_Won ~ Runs_Scored + Earned_Run_Average + Saves +
##     Complete_Games + Fielding_Percentage + Shutout + Errors +
##     Infield_Put_Outs + At_Bats + Hits + Runs_Against + Caught_Stealing +
##     Double_Plays + Strike_Outs + Walks + Stolen_Bases, data = df_p3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.0646  -2.1949   0.0916   2.1205  10.3707
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.001e+02  3.867e+01  -2.588  0.00984 **
## Runs_Scored      8.567e-02  3.893e-03  22.003 < 2e-16 ***
## Earned_Run_Average -2.275e+00  1.982e+00  -1.148  0.25144
## Saves           2.955e-01  2.253e-02  13.117 < 2e-16 ***
## Complete_Games    1.406e-01  1.529e-02   9.193 < 2e-16 ***
## Fielding_Percentage 1.105e+02  3.844e+01   2.874  0.00417 **
## Shutout          1.703e-01  3.950e-02   4.311 1.86e-05 ***
## Errors          -1.241e-03  1.189e-02  -0.104  0.91688
## Infield_Put_Outs   5.703e-02  5.393e-03  10.576 < 2e-16 ***
## At_Bats          -4.673e-02  5.169e-03  -9.040 < 2e-16 ***
## Hits             3.857e-02  5.552e-03   6.947 8.40e-12 ***
## Runs_Against      -5.050e-02  1.196e-02  -4.222 2.74e-05 ***
## Caught_Stealing    -5.678e-02  1.179e-02  -4.814 1.81e-06 ***
## Double_Plays      -9.962e-03  6.995e-03  -1.424  0.15485
## Strike_Outs        2.472e-03  1.504e-03   1.643  0.10076
## Walks            -4.602e-03  2.711e-03  -1.698  0.09001 .
## Stolen_Bases       6.869e-03  4.121e-03   1.667  0.09598 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.241 on 713 degrees of freedom
## Multiple R-squared:  0.9354, Adjusted R-squared:  0.9339
## F-statistic: 645.2 on 16 and 713 DF, p-value: < 2.2e-16
```

```
# Period 4 - 1990 to 2010
df_p4 <- drop.na(df_1990_2010[-c(1, 3)])

## Forward stepwise regression
```

```

null_modelp4 <- lm(Games_Won ~ 1, data = df_p4)
full_modelp4 <- lm(Games_Won ~ . , data = df_p4)

step(null_modelp4, direction = "forward", scope = formula(full_modelp4))

```

```

## Start:  AIC=1619.18
## Games_Won ~ 1
##
##
##      Df Sum of Sq  RSS    AIC
## + Saves          1   19728.7 24611 1426.9
## + Runs_Against    1   19487.7 24852 1430.1
## + Earned_Run_Average 1   19333.2 25006 1432.2
## + Earned_Runs      1   18422.0 25917 1444.0
## + Runs_Scored      1   15400.5 28939 1480.4
## + Hits_Allowed     1   12364.0 31975 1513.3
## + Infield_Put_Outs 1   11292.4 33047 1524.2
## + Walks_Allowed    1   10037.6 34302 1536.5
## + Shutout          1   10004.9 34335 1536.8
## + Walks            1    8519.6 35820 1550.8
## + Home_Runs        1    7574.2 36765 1559.4
## + Fielding_Percentage 1    7284.2 37055 1562.0
## + Hits             1    7265.3 37074 1562.1
## + Errors           1    6992.7 37347 1564.5
## + Home_Run_Allowed 1    6720.4 37619 1566.9
## + Strikeouts_Allowed 1    6566.2 37773 1568.3
## + Sacrifice_Fly    1    4757.9 39581 1583.7
## + Doubles          1    2831.5 41508 1599.4
## + At_Bats          1    1825.3 42514 1607.3
## + Double_Plays     1    1533.7 42806 1609.6
## + Strike_Outs      1    1076.9 43263 1613.1
## + Complete_Games   1     935.7 43404 1614.1
## + Hit_By_Pitch     1     556.5 43783 1617.0
## + Stolen_Bases     1     422.2 43917 1618.0
## + Caught_Stealing  1     415.4 43924 1618.1
## <none>              44339 1619.2
## + Triples          1     213.0 44126 1619.6
##
## Step:  AIC=1426.91
## Games_Won ~ Saves
##
##
##      Df Sum of Sq  RSS    AIC
## + Runs_Scored      1   10503.0 14108 1245.3
## + Runs_Against     1    6196.4 18414 1333.2
## + Earned_Run_Average 1    6116.7 18494 1334.6
## + Walks            1    5848.4 18762 1339.4
## + Earned_Runs      1    5689.8 18921 1342.2
## + Hits             1    5440.4 19170 1346.5
## + Home_Runs        1    4096.4 20514 1368.8
## + Hits_Allowed     1    3614.4 20996 1376.5
## + Infield_Put_Outs 1    3421.0 21190 1379.5
## + Doubles          1    3072.1 21539 1384.9
## + Fielding_Percentage 1    3040.5 21570 1385.4
## + Shutout          1    2952.3 21658 1386.7

```



```

## + Errors          1    2913.8 21697 1387.3
## + Walks_Allowed   1    2676.2 21935 1390.9
## + Sacrifice_Fly    1    2653.1 21958 1391.3
## + Strikeouts_Allowed 1    2367.0 22244 1395.5
## + At_Bats          1    2306.3 22304 1396.4
## + Home_Run_Allowed 1    1890.2 22720 1402.5
## + Complete_Games   1    1465.0 23146 1408.7
## + Caught_Stealing  1     527.7 24083 1421.8
## + Strike-Outs       1     291.6 24319 1425.0
## + Hit_By_Pitch     1     156.2 24455 1426.8
## <none>              1    24611 1426.9
## + Double_Plays     1      57.0 24554 1428.1
## + Triples          1      50.7 24560 1428.2
## + Stolen_Bases     1      18.6 24592 1428.7
##
## Step:  AIC=1245.28
## Games_Won ~ Saves + Runs_Scored
##
##              Df Sum of Sq    RSS    AIC
## + Runs_Against    1   10964.5  3143.2  751.79
## + Earned_Run_Average 1   10697.1  3410.6  778.73
## + Earned_Runs      1   10630.2  3477.5  785.14
## + Hits_Allowed     1    7150.6  6957.1 1013.98
## + Shutout          1    4757.4  9350.3 1111.54
## + Home_Run_Allowed 1    4455.9  9651.8 1122.02
## + Strikeouts_Allowed 1    2774.8 11332.9 1175.00
## + Walks_Allowed    1    2398.4 11709.4 1185.79
## + Fielding_Percentage 1    2047.2 12060.5 1195.54
## + Infield_Put_Outs  1    2011.4 12096.3 1196.52
## + Errors           1    1977.6 12130.2 1197.44
## + Complete_Games   1     940.8 13166.9 1224.50
## + Walks             1     584.9 13522.8 1233.30
## + At_Bats           1     474.9 13632.8 1235.98
## + Triples           1     267.0 13840.7 1240.97
## + Double_Plays      1     228.4 13879.3 1241.89
## + Caught_Stealing   1     118.3 13989.4 1244.50
## + Hits              1     113.4 13994.4 1244.61
## + Home_Runs         1      90.4 14017.3 1245.15
## <none>              1    14107.7 1245.28
## + Strike-Outs       1      67.4 14040.3 1245.70
## + Sacrifice_Fly     1      55.2 14052.5 1245.98
## + Hit_By_Pitch      1      47.6 14060.1 1246.16
## + Stolen_Bases      1      10.9 14096.8 1247.02
## + Doubles           1       0.0 14107.7 1247.28
##
## Step:  AIC=751.79
## Games_Won ~ Saves + Runs_Scored + Runs_Against
##
##              Df Sum of Sq    RSS    AIC
## + Triples          1    108.610  3034.6  742.19
## + Shutout          1     80.243  3063.0  745.26
## + Complete_Games   1     69.088  3074.1  746.46
## + Strikeouts_Allowed 1     50.722  3092.5  748.42
## + At_Bats          1     43.552  3099.7  749.19

```

```

## + Infield_Put_Outs      1      38.822 3104.4 749.69
## + Walks_Allowed         1      37.723 3105.5 749.81
## + Walks                  1      28.756 3114.5 750.76
## <none>                   1      3143.2 751.79
## + Doubles                1      15.641 3127.6 752.14
## + Errors                 1      14.349 3128.9 752.28
## + Earned_Run_Average    1      13.680 3129.5 752.35
## + Double_Plays           1      12.308 3130.9 752.50
## + Home_Runs              1      12.001 3131.2 752.53
## + Caught_Stealing        1      11.012 3132.2 752.63
## + Stolen_Bases           1      10.353 3132.9 752.70
## + Fielding_Percentage    1       6.357 3136.9 753.12
## + Hits_Allowed           1       5.309 3137.9 753.23
## + Strike_Outs            1       3.728 3139.5 753.40
## + Home_Run_Allowed       1       3.635 3139.6 753.41
## + Hits                   1       2.012 3141.2 753.58
## + Sacrifice_Fly          1       1.674 3141.5 753.61
## + Earned_Runs            1       1.338 3141.9 753.65
## + Hit_By_Pitch           1       0.192 3143.0 753.77
##
## Step: AIC=742.19
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples
##
##              Df Sum of Sq    RSS    AIC
## + Shutout      1      84.912 2949.7 734.82
## + Complete_Games 1      68.793 2965.8 736.62
## + Strikeouts_Allowed 1     51.894 2982.7 738.49
## + At_Bats       1      40.310 2994.3 739.77
## + Infield_Put_Outs 1      36.848 2997.8 740.15
## + Walks_Allowed  1      29.993 3004.6 740.91
## + Caught_Stealing 1      25.656 3008.9 741.38
## + Stolen_Bases   1      25.525 3009.1 741.40
## <none>           1      3034.6 742.19
## + Errors         1      18.320 3016.3 742.19
## + Walks          1      16.655 3017.9 742.37
## + Doubles        1      14.305 3020.3 742.63
## + Sacrifice_Fly  1      11.243 3023.4 742.96
## + Fielding_Percentage 1       9.092 3025.5 743.20
## + Double_Plays    1       9.035 3025.6 743.20
## + Earned_Run_Average 1       8.004 3026.6 743.31
## + Strike_Outs     1       4.795 3029.8 743.66
## + Home_Run_Allowed 1       4.456 3030.2 743.70
## + Hits_Allowed    1       3.545 3031.1 743.80
## + Hit_By_Pitch    1       3.264 3031.3 743.83
## + Earned_Runs     1       0.112 3034.5 744.17
## + Home_Runs       1       0.025 3034.6 744.18
## + Hits            1       0.005 3034.6 744.19
##
## Step: AIC=734.82
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout
##
##              Df Sum of Sq    RSS    AIC
## + Complete_Games  1      49.822 2899.9 731.20
## + Strikeouts_Allowed 1      43.834 2905.9 731.88

```

```

## + At_Bats          1    40.493 2909.2 732.26
## + Infield_Put_Outs 1    38.349 2911.3 732.50
## + Caught_Stealing  1    33.506 2916.2 733.05
## + Walks_Allowed    1    32.613 2917.1 733.15
## + Stolen_Bases     1    25.893 2923.8 733.91
## + Walks            1    19.016 2930.7 734.69
## <none>              1    2949.7 734.82
## + Doubles          1    16.524 2933.2 734.97
## + Errors           1    14.190 2935.5 735.23
## + Double_Plays     1    11.628 2938.1 735.52
## + Sacrifice_Fly    1    10.258 2939.4 735.67
## + Earned_Run_Average 1     6.440 2943.3 736.10
## + Fielding_Percentage 1     6.357 2943.3 736.11
## + Hits_Allowed     1     4.853 2944.8 736.28
## + Strike_Outs      1     4.388 2945.3 736.33
## + Hit_By_Pitch     1     2.584 2947.1 736.53
## + Home_Runs        1     1.224 2948.5 736.68
## + Home_Run_Allowed 1     0.568 2949.1 736.76
## + Hits             1     0.004 2949.7 736.82
## + Earned_Runs      1     0.003 2949.7 736.82
##
## Step:  AIC=731.2
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games
##
##              Df Sum of Sq   RSS   AIC
## + Strikeouts_Allowed 1    60.323 2839.6 726.26
## + Infield_Put_Outs   1    43.118 2856.8 728.26
## + At_Bats            1    30.644 2869.2 729.69
## + Caught_Stealing    1    22.961 2876.9 730.58
## + Stolen_Bases       1    21.182 2878.7 730.78
## + Walks              1    21.073 2878.8 730.79
## + Walks_Allowed      1    19.839 2880.0 730.93
## + Errors             1    19.414 2880.5 730.98
## <none>               1    2899.9 731.20
## + Double_Plays       1    14.723 2885.2 731.52
## + Doubles            1    13.999 2885.9 731.60
## + Fielding_Percentage 1     9.625 2890.2 732.10
## + Earned_Run_Average 1     6.688 2893.2 732.44
## + Sacrifice_Fly      1     6.646 2893.2 732.44
## + Hits_Allowed       1     4.008 2895.9 732.74
## + Home_Run_Allowed   1     3.591 2896.3 732.79
## + Hit_By_Pitch       1     3.577 2896.3 732.79
## + Home_Runs          1     0.640 2899.2 733.13
## + Strike_Outs        1     0.527 2899.3 733.14
## + Hits               1     0.492 2899.4 733.14
## + Earned_Runs        1     0.026 2899.8 733.20
##
## Step:  AIC=726.26
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed
##
##              Df Sum of Sq   RSS   AIC
## + Walks_Allowed      1    39.735 2799.8 723.61

```

```

## + At_Bats          1    30.781 2808.8 724.67
## + Hits_Allowed     1    29.447 2810.1 724.82
## + Infield_Put_Outs 1    28.259 2811.3 724.96
## + Caught_Stealing   1    23.682 2815.9 725.50
## + Stolen_Bases      1    18.905 2820.7 726.06
## + Errors            1    18.525 2821.0 726.10
## <none>              2839.6 726.26
## + Doubles           1    14.008 2825.5 726.63
## + Walks             1    12.474 2827.1 726.81
## + Fielding_Percentage 1    11.469 2828.1 726.93
## + Sacrifice_Fly     1    11.293 2828.3 726.95
## + Strike_Outs       1     5.777 2833.8 727.59
## + Home_Run_Allowed  1     4.218 2835.3 727.77
## + Earned_Run_Average 1     2.890 2836.7 727.93
## + Hit_By_Pitch      1     2.259 2837.3 728.00
## + Hits              1     1.483 2838.1 728.09
## + Home_Runs         1     0.281 2839.3 728.23
## + Earned_Runs       1     0.253 2839.3 728.23
## + Double_Plays      1     0.236 2839.3 728.23
##
## Step: AIC=723.61
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed
##
##              Df Sum of Sq   RSS   AIC
## + At_Bats          1    35.503 2764.3 721.40
## + Infield_Put_Outs 1    28.237 2771.6 722.27
## + Caught_Stealing   1    21.623 2778.2 723.05
## + Stolen_Bases      1    18.734 2781.1 723.40
## + Errors            1    17.912 2781.9 723.49
## <none>              2799.8 723.61
## + Walks             1    16.696 2783.1 723.64
## + Doubles           1    15.934 2783.9 723.73
## + Fielding_Percentage 1    11.278 2788.5 724.28
## + Home_Run_Allowed  1    10.880 2788.9 724.33
## + Sacrifice_Fly     1     9.893 2789.9 724.44
## + Hits_Allowed     1     6.735 2793.1 724.82
## + Strike_Outs       1     5.739 2794.1 724.93
## + Hit_By_Pitch      1     3.640 2796.2 725.18
## + Earned_Run_Average 1     2.055 2797.8 725.37
## + Double_Plays      1     0.915 2798.9 725.50
## + Earned_Runs       1     0.612 2799.2 725.54
## + Hits              1     0.545 2799.3 725.55
## + Home_Runs         1     0.020 2799.8 725.61
##
## Step: AIC=721.4
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed + At_Bats
##
##              Df Sum of Sq   RSS   AIC
## + Infield_Put_Outs 1    106.920 2657.4 710.38
## + Hits              1     57.794 2706.5 716.43
## + Errors            1     27.453 2736.9 720.11
## + Fielding_Percentage 1     20.259 2744.1 720.97

```

```

## + Stolen_Bases      1      17.451 2746.9 721.31
## + Strike_Outs       1      17.446 2746.9 721.31
## <none>              2764.3 721.40
## + Caught_Stealing   1      13.038 2751.3 721.84
## + Hits_Allowed      1      12.179 2752.1 721.94
## + Home_Run_Allowed  1      11.624 2752.7 722.01
## + Sacrifice_Fly     1       8.966 2755.3 722.33
## + Hit_By_Pitch      1       6.409 2757.9 722.63
## + Doubles           1       6.395 2757.9 722.64
## + Earned_Run_Average 1       5.064 2759.2 722.79
## + Walks             1       3.623 2760.7 722.97
## + Double_Plays      1       2.472 2761.8 723.10
## + Home_Runs         1       1.928 2762.4 723.17
## + Earned_Runs       1       1.741 2762.6 723.19
##
## Step:  AIC=710.38
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed + At_Bats +
##      Infield_Put_Outs
##
##              Df Sum of Sq    RSS    AIC
## + Hits              1    200.511 2456.9 686.49
## + Strike_Outs       1     41.722 2615.7 707.16
## + Errors            1     31.683 2625.7 708.42
## + Home_Run_Allowed  1     21.921 2635.5 709.65
## + Fielding_Percentage 1     19.352 2638.0 709.97
## <none>              2657.4 710.38
## + Hit_By_Pitch      1     13.975 2643.4 710.64
## + Stolen_Bases      1     13.302 2644.1 710.73
## + Home_Runs         1     12.010 2645.4 710.89
## + Sacrifice_Fly     1      6.396 2651.0 711.59
## + Caught_Stealing   1      5.000 2652.4 711.76
## + Walks             1      4.957 2652.4 711.77
## + Earned_Runs       1      2.946 2654.4 712.02
## + Earned_Run_Average 1      2.924 2654.5 712.02
## + Doubles           1      2.669 2654.7 712.05
## + Hits_Allowed      1      2.360 2655.0 712.09
## + Double_Plays      1      0.182 2657.2 712.36
##
## Step:  AIC=686.49
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed + At_Bats +
##      Infield_Put_Outs + Hits
##
##              Df Sum of Sq    RSS    AIC
## + Home_Run_Allowed  1    22.8573 2434.0 685.41
## + Caught_Stealing   1    16.8543 2440.0 686.22
## + Errors            1    16.8218 2440.1 686.23
## <none>              2456.9 686.49
## + Home_Runs         1     9.8217 2447.1 687.17
## + Fielding_Percentage 1     9.3939 2447.5 687.23
## + Hit_By_Pitch      1     4.1908 2452.7 687.93
## + Sacrifice_Fly     1     3.8513 2453.0 687.98
## + Doubles           1     2.9928 2453.9 688.09

```

```

## + Stolen_Bases      1      1.0873 2455.8 688.35
## + Strike_Outs       1      0.8376 2456.1 688.38
## + Walks             1      0.4023 2456.5 688.44
## + Double_Plays      1      0.2216 2456.7 688.46
## + Hits_Allowed      1      0.2181 2456.7 688.46
## + Earned_Runs       1      0.1040 2456.8 688.48
## + Earned_Run_Average 1      0.0783 2456.8 688.48
##
## Step:  AIC=685.41
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed + At_Bats +
##      Infield_Put_Outs + Hits + Home_Run_Allowed
##
##              Df Sum of Sq    RSS    AIC
## + Errors      1    25.6830 2408.3 683.91
## + Fielding_Percentage 1    15.0219 2419.0 685.37
## <none>                        2434.0 685.41
## + Caught_Stealing    1    12.2845 2421.7 685.74
## + Home_Runs          1    11.9366 2422.1 685.79
## + Doubles            1     6.5224 2427.5 686.52
## + Hit_By_Pitch       1     3.8344 2430.2 686.89
## + Sacrifice_Fly      1     3.6330 2430.4 686.92
## + Stolen_Bases       1     3.0271 2431.0 687.00
## + Earned_Runs        1     1.9343 2432.1 687.15
## + Earned_Run_Average 1     1.7188 2432.3 687.18
## + Strike_Outs        1     0.5507 2433.5 687.33
## + Hits_Allowed       1     0.5220 2433.5 687.34
## + Double_Plays       1     0.1635 2433.9 687.39
## + Walks              1     0.0450 2434.0 687.40
##
## Step:  AIC=683.91
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed + At_Bats +
##      Infield_Put_Outs + Hits + Home_Run_Allowed + Errors
##
##              Df Sum of Sq    RSS    AIC
## + Fielding_Percentage 1     49.031 2359.3 679.12
## <none>                        2408.3 683.91
## + Home_Runs          1    12.929 2395.4 684.13
## + Earned_Run_Average 1     7.385 2401.0 684.89
## + Doubles            1     7.181 2401.2 684.92
## + Earned_Runs        1     6.893 2401.4 684.96
## + Caught_Stealing    1     6.482 2401.9 685.02
## + Sacrifice_Fly      1     4.233 2404.1 685.33
## + Hit_By_Pitch       1     3.941 2404.4 685.37
## + Stolen_Bases       1     3.416 2404.9 685.44
## + Hits_Allowed       1     2.994 2405.3 685.50
## + Strike_Outs        1     1.249 2407.1 685.74
## + Double_Plays       1     0.741 2407.6 685.81
## + Walks              1     0.012 2408.3 685.91
##
## Step:  AIC=679.12
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed + At_Bats +

```

```

##      Infield_Put_Outs + Hits + Home_Run_Allowed + Errors + Fielding_Percentage
##
##              Df Sum of Sq    RSS    AIC
## + Home_Runs      1   17.2765 2342.0 678.70
## <none>                      2359.3 679.12
## + Earned_Run_Average 1    9.1798 2350.1 679.83
## + Earned_Runs      1    8.9818 2350.3 679.86
## + Caught_Stealing   1    6.8546 2352.5 680.16
## + Sacrifice_Fly     1    6.5983 2352.7 680.20
## + Hit_By_Pitch      1    4.7356 2354.6 680.46
## + Doubles           1    3.8322 2355.5 680.58
## + Double_Plays      1    2.0118 2357.3 680.84
## + Stolen_Bases      1    1.2769 2358.0 680.94
## + Hits_Allowed      1    1.0301 2358.3 680.98
## + Strike-Outs       1    0.4132 2358.9 681.06
## + Walks             1    0.0176 2359.3 681.12
##
## Step:  AIC=678.7
## Games_Won ~ Saves + Runs_Scored + Runs_Against + Triples + Shutout +
##      Complete_Games + Strikeouts_Allowed + Walks_Allowed + At_Bats +
##      Infield_Put_Outs + Hits + Home_Run_Allowed + Errors + Fielding_Percentage +
##      Home_Runs
##
##              Df Sum of Sq    RSS    AIC
## <none>                      2342.0 678.70
## + Earned_Run_Average 1    9.3615 2332.7 679.37
## + Earned_Runs      1    9.1813 2332.8 679.40
## + Caught_Stealing   1    5.9140 2336.1 679.86
## + Hit_By_Pitch      1    5.1676 2336.9 679.97
## + Stolen_Bases      1    2.9138 2339.1 680.28
## + Sacrifice_Fly     1    2.5188 2339.5 680.34
## + Doubles           1    2.4708 2339.6 680.35
## + Double_Plays      1    1.6169 2340.4 680.47
## + Walks             1    1.3406 2340.7 680.51
## + Hits_Allowed      1    0.7988 2341.2 680.58
## + Strike-Outs       1    0.0200 2342.0 680.69
##
##
## Call:
## lm(formula = Games_Won ~ Saves + Runs_Scored + Runs_Against +
##      Triples + Shutout + Complete_Games + Strikeouts_Allowed +
##      Walks_Allowed + At_Bats + Infield_Put_Outs + Hits + Home_Run_Allowed +
##      Errors + Fielding_Percentage + Home_Runs, data = df_p4)
##
## Coefficients:
##      (Intercept)              Saves      Runs_Scored
##      1.108e+03      3.678e-01      7.879e-02
##      Runs_Against      Triples      Shutout
##      -5.653e-02      -6.822e-02      1.519e-01
##      Complete_Games  Strikeouts_Allowed  Walks_Allowed
##      1.239e-01      2.328e-03      -8.176e-03
##      At_Bats      Infield_Put_Outs      Hits
##      -3.440e-02      4.363e-02      3.042e-02
##      Home_Run_Allowed      Errors  Fielding_Percentage

```

```
##          -2.304e-02          -1.949e-01          -1.096e+03
##          Home_Runs
##          1.234e-02
```

```
fit4 <- lm(Games_Won ~ Saves + Runs_Scored + Runs_Against +
           Triples + Shutout + Complete_Games + Strikeouts_Allowed +
           Walks_Allowed + At_Bats + Infield_Put_Outs + Hits + Home_Run_Allowed +
           Errors + Fielding_Percentage + Home_Runs, data = df_p4)

summary(fit4)
```

```
##
## Call:
## lm(formula = Games_Won ~ Saves + Runs_Scored + Runs_Against +
##     Triples + Shutout + Complete_Games + Strikeouts_Allowed +
##     Walks_Allowed + At_Bats + Infield_Put_Outs + Hits + Home_Run_Allowed +
##     Errors + Fielding_Percentage + Home_Runs, data = df_p4)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.7050 -1.8150 -0.1751  1.7257  8.6409
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.108e+03  4.040e+02   2.741  0.006467 **
## Saves          3.678e-01  2.754e-02  13.352 < 2e-16 ***
## Runs_Scored    7.879e-02  4.912e-03  16.038 < 2e-16 ***
## Runs_Against  -5.653e-02  4.187e-03 -13.499 < 2e-16 ***
## Triples       -6.821e-02  1.833e-02  -3.721  0.000235 ***
## Shutout        1.519e-01  5.713e-02   2.659  0.008239 **
## Complete_Games 1.239e-01  4.906e-02   2.526  0.012041 *
## Strikeouts_Allowed 2.328e-03  1.840e-03   1.265  0.206668
## Walks_Allowed  -8.176e-03  2.859e-03  -2.859  0.004532 **
## At_Bats       -3.440e-02  5.145e-03  -6.685  1.06e-10 ***
## Infield_Put_Outs 4.363e-02  6.758e-03   6.456  4.09e-10 ***
## Hits          3.042e-02  6.185e-03   4.918  1.41e-06 ***
## Home_Run_Allowed -2.304e-02  9.143e-03  -2.520  0.012244 *
## Errors        -1.949e-01  6.621e-02  -2.944  0.003484 **
## Fielding_Percentage -1.096e+03  4.096e+02  -2.675  0.007860 **
## Home_Runs      1.234e-02  8.107e-03   1.522  0.129031
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.731 on 314 degrees of freedom
## Multiple R-squared:  0.9472, Adjusted R-squared:  0.9447
## F-statistic: 375.4 on 15 and 314 DF, p-value: < 2.2e-16
```

```
# From the summary, we can find
# The p-value indicates if the regression coefficient is significantly different
# from 0. If the p-value is smaller than 0.05, then we reject the null hypothesis
# that the regression coefficient is equal to 0 at 0.05 significance level.
# R^2 is a measure of the variance of y that is explained by the model.
# The higher the R^2 is, the "better" is the model. However, adding additional
```



```

# variables will always increase R^2 while this may not improve the model in the
# sense that it may not improve prediction for new data. The adjusted
# R^2 accounts for the number of variables in the model and is a better measure
# of the model's quality.
# F-statistic tells us whether the model is statistically significant or not.
# The null hypothesis is that all coefficients are zero and the alternative
# hypothesis is that not all coefficients are zero.

# Use the 4th regression model from 1990 to 2010 and forecast the number of games
# won for the New York Yankees and the Toronto Blue Jays using values for the
# independent variables for 2012 and 2015.

# New York Yankees
df_yk <- data %>%
  filter(Team_Name == 'New York Yankees', Year %in% c(2012, 2015))

df_yk2 <- df_yk %>%
  select(Saves, Runs_Scored, Runs_Against, Triples, Shutout, Complete_Games, Strikeouts_Allowed,
         Walks_Allowed, At_Bats, Infield_Put_Outs, Hits, Home_Run_Allowed,
         Errors, Fielding_Percentage, Home_Runs)

pred_yk <- predict(fit4, df_yk2)

cor(pred_yk, df_yk$Games_Won)

## [1] 1

# Low values of RMSE indicate good fit
rmse(as.double(df_yk$Games_Won), as.double(pred_yk))

## [1] 2.779608

# Toronto Blue Jays
df_bj <- data %>%
  filter(Team_Name == 'Toronto Blue Jays', Year %in% c(2012, 2015))

df_bj2 <- df_bj %>%
  select(Saves, Runs_Scored, Runs_Against, Triples, Shutout, Complete_Games, Strikeouts_Allowed,
         Walks_Allowed, At_Bats, Infield_Put_Outs, Hits, Home_Run_Allowed,
         Errors, Fielding_Percentage, Home_Runs)

pred_bj <- predict(fit4, df_bj2)

cor(pred_bj, df_bj$Games_Won)

## [1] 1

# Low values of RMSE indicate good fit
rmse(as.double(df_bj$Games_Won), as.double(pred_bj))

## [1] 5.060376

```