

Home Work Assignment - 01 - Test/ Evaluation Data

Critical Thinking Group 5

Contents

Data Exploration	2
Summary / Descriptives / Correlation	2
Distribution and Correlation	3
Data Preparation	16
Additional Variables	19

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

url <- "https://raw.githubusercontent.com/kishkp/data621-ctg5/master/HW1/moneyball-evaluation-data.csv"
moneyball<- read.csv(url)
head(moneyball)
```

```
##   INDEX TEAM_BATTING_H TEAM_BATTING_2B TEAM_BATTING_3B TEAM_BATTING_HR
## 1     9          1209          170          33          83
## 2    10          1221          151          29          88
## 3    14          1395          183          29          93
## 4    47          1539          309          29         159
## 5    60          1445          203          68           5
## 6    63          1431          236          53          10
##   TEAM_BATTING_BB TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS
## 1             447             1080             62             50
## 2             516             929              54             39
## 3             509             816              59             47
## 4             486             914             148             57
## 5              95             416              NA             NA
## 6             215             377              NA             NA
##   TEAM_BATTING_HBP TEAM_PITCHING_H TEAM_PITCHING_HR TEAM_PITCHING_BB
## 1              NA             1209             83             447
## 2              NA             1221             88             516
## 3              NA             1395             93             509
## 4              42             1539             159             486
## 5              NA             3902              14             257
## 6              NA             2793              20             420
##   TEAM_PITCHING_SO TEAM_FIELDING_E TEAM_FIELDING_DP
```

```
## 1      1080      140      156
## 2      929      135      164
## 3      816      156      153
## 4      914      124      154
## 5     1123      616      130
## 6      736      572      105
```

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

```
## TEAM_BATTING_H TEAM_BATTING_2B TEAM_BATTING_3B TEAM_BATTING_HR
## Min. : 819 Min. : 44.0 Min. : 14.00 Min. : 0.00
## 1st Qu.:1387 1st Qu.:210.0 1st Qu.: 35.00 1st Qu.: 44.50
## Median :1455 Median :239.0 Median : 52.00 Median :101.00
## Mean :1469 Mean :241.3 Mean : 55.91 Mean : 95.63
## 3rd Qu.:1548 3rd Qu.:278.5 3rd Qu.: 72.00 3rd Qu.:135.50
## Max. :2170 Max. :376.0 Max. :155.00 Max. :242.00
##
## TEAM_BATTING_BB TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS
## Min. : 15.0 Min. : 0.0 Min. : 0.0 Min. : 0.00
## 1st Qu.:436.5 1st Qu.: 545.0 1st Qu.: 59.0 1st Qu.: 38.00
## Median :509.0 Median : 686.0 Median : 92.0 Median : 49.50
## Mean :499.0 Mean : 709.3 Mean :123.7 Mean : 52.32
## 3rd Qu.:565.5 3rd Qu.: 912.0 3rd Qu.:151.8 3rd Qu.: 63.00
## Max. :792.0 Max. :1268.0 Max. :580.0 Max. :154.00
## NA's :18 NA's :13 NA's :87
## TEAM_BATTING_HBP TEAM_PITCHING_H TEAM_PITCHING_HR TEAM_PITCHING_BB
## Min. :42.00 Min. : 1155 Min. : 0.0 Min. : 136.0
## 1st Qu.:53.50 1st Qu.: 1426 1st Qu.: 52.0 1st Qu.: 471.0
## Median :62.00 Median : 1515 Median :104.0 Median : 526.0
## Mean :62.37 Mean : 1813 Mean :102.1 Mean : 552.4
## 3rd Qu.:67.50 3rd Qu.: 1681 3rd Qu.:142.5 3rd Qu.: 606.5
## Max. :96.00 Max. :22768 Max. :336.0 Max. :2008.0
## NA's :240
## TEAM_PITCHING_SO TEAM_FIELDING_E TEAM_FIELDING_DP
## Min. : 0.0 Min. : 73.0 Min. : 69.0
## 1st Qu.: 613.0 1st Qu.: 131.0 1st Qu.:131.0
## Median : 745.0 Median : 163.0 Median :148.0
## Mean : 799.7 Mean : 249.7 Mean :146.1
## 3rd Qu.: 938.0 3rd Qu.: 252.0 3rd Qu.:164.0
## Max. :9963.0 Max. :1568.0 Max. :204.0
## NA's :18 NA's :31
```

Data Exploration

Summary / Descriptives / Correlation

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

```
Variable<- rownames(ds_stats)

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

Distribution and Correlation

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

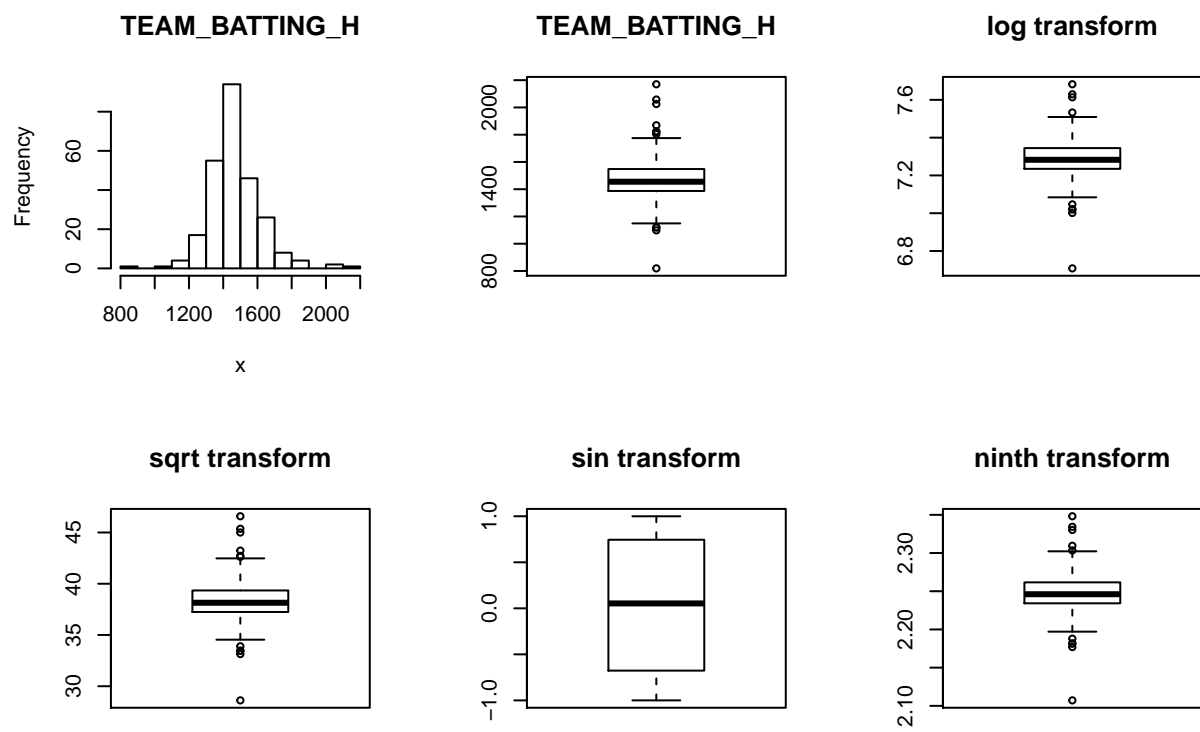
  par(mfrow=c(2,3))

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

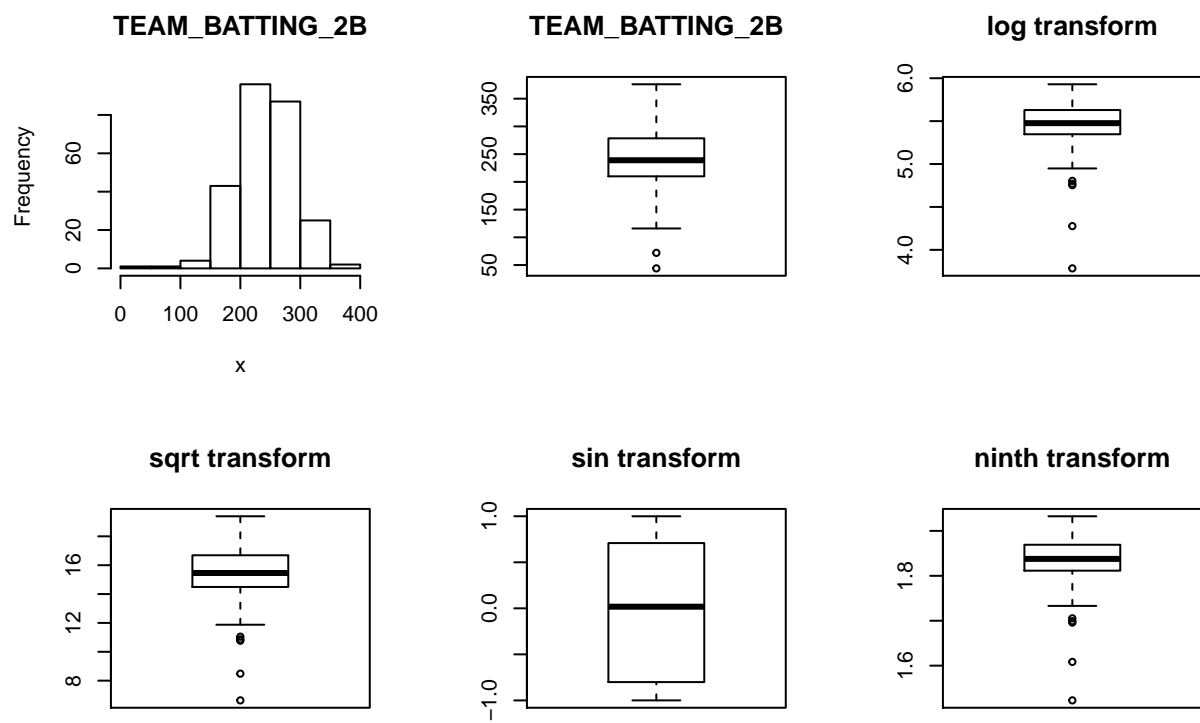
  hist(x,main=xlabel)
  boxplot(x,main=xlabel)

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

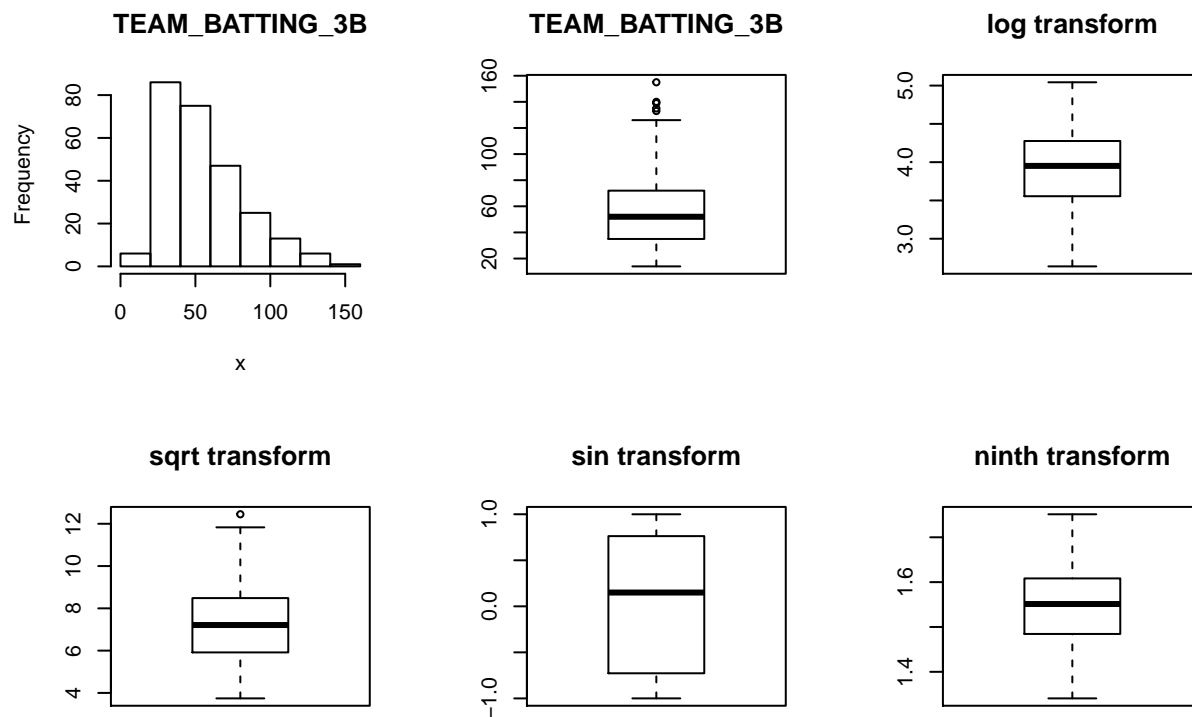
show_charts(moneyball2$TEAM_BATTING_H)
```



```
show_charts(moneyball12$TEAM_BATTING_2B)
```

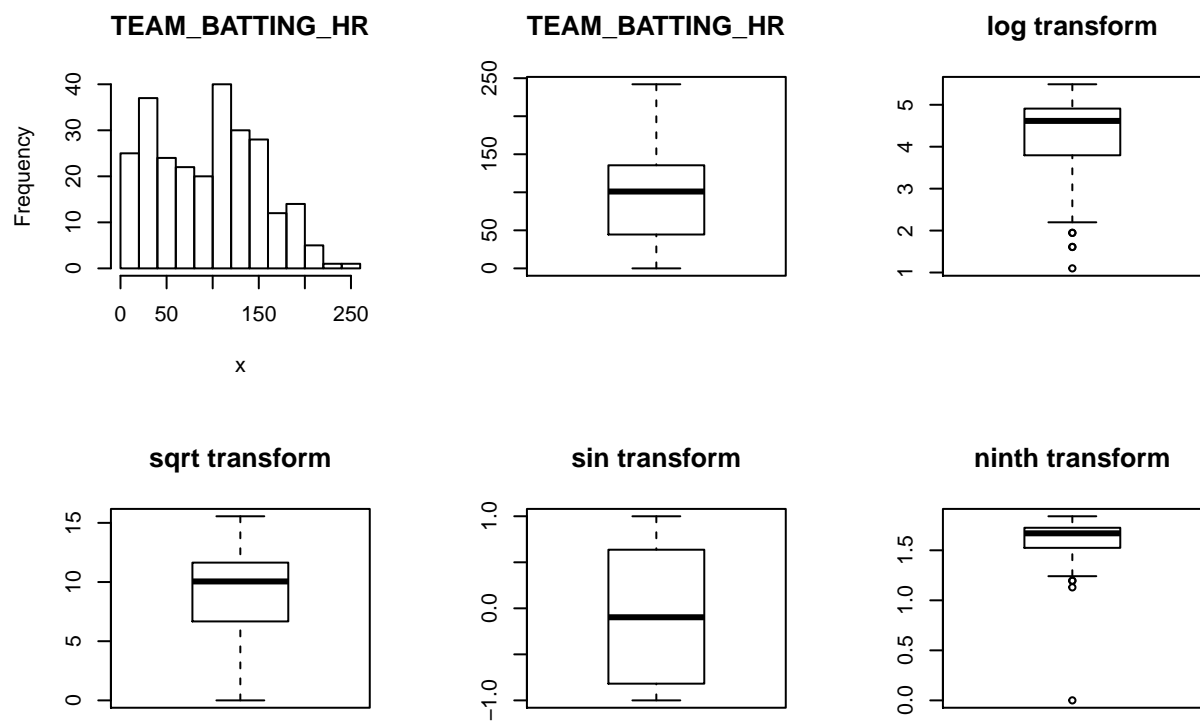


```
show_charts(moneyball12$TEAM_BATTING_3B)
```

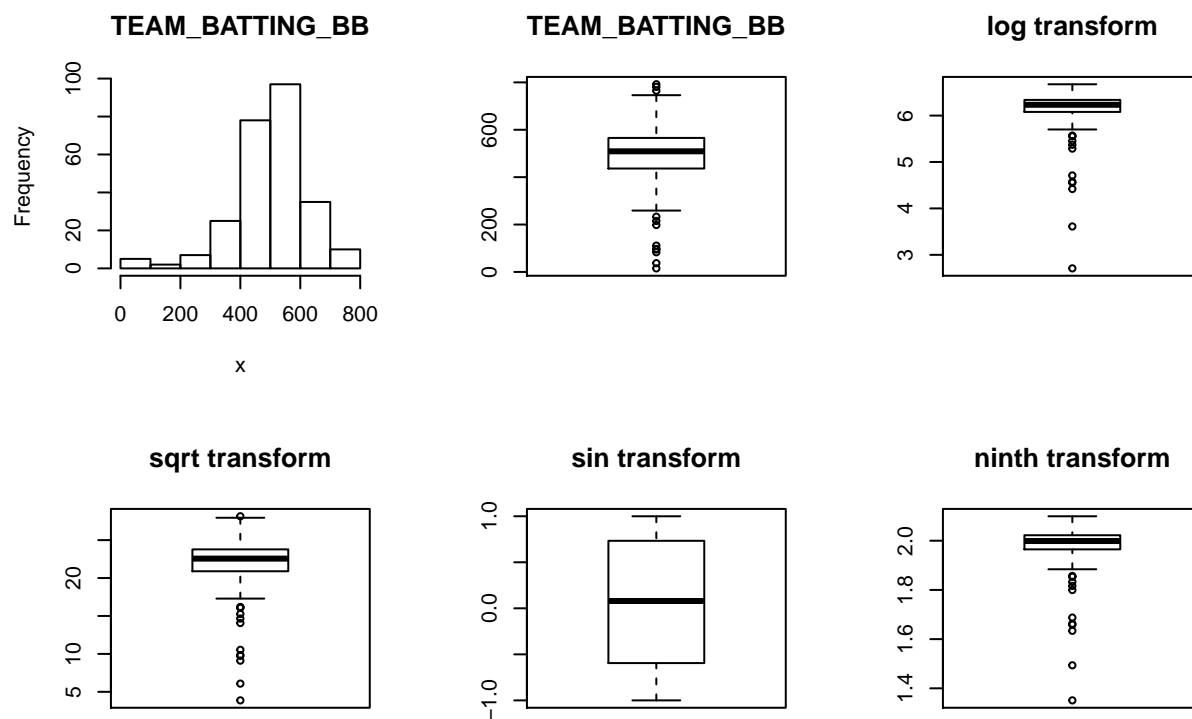


```
show_charts(moneyball2$TEAM_BATTING_HR)
```

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

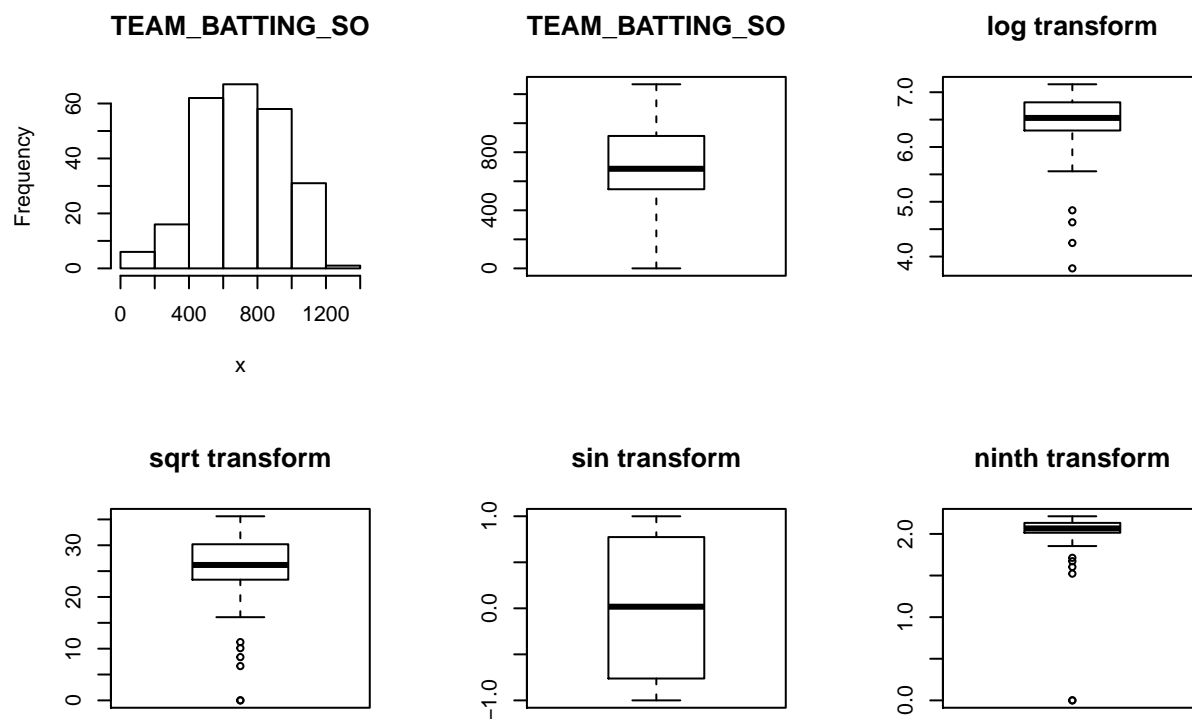


```
show_charts(moneyball12$TEAM_BATTING_BB)
```



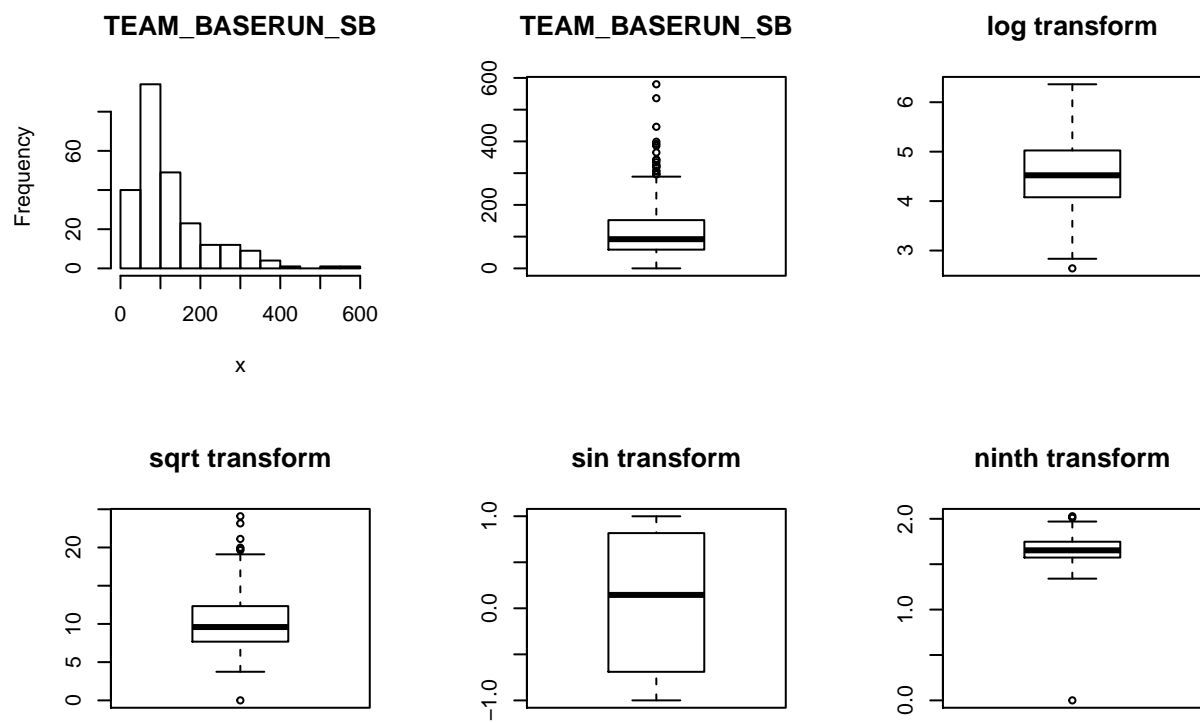
```
show_charts(moneyball2$TEAM_BATTING_SO)
```

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

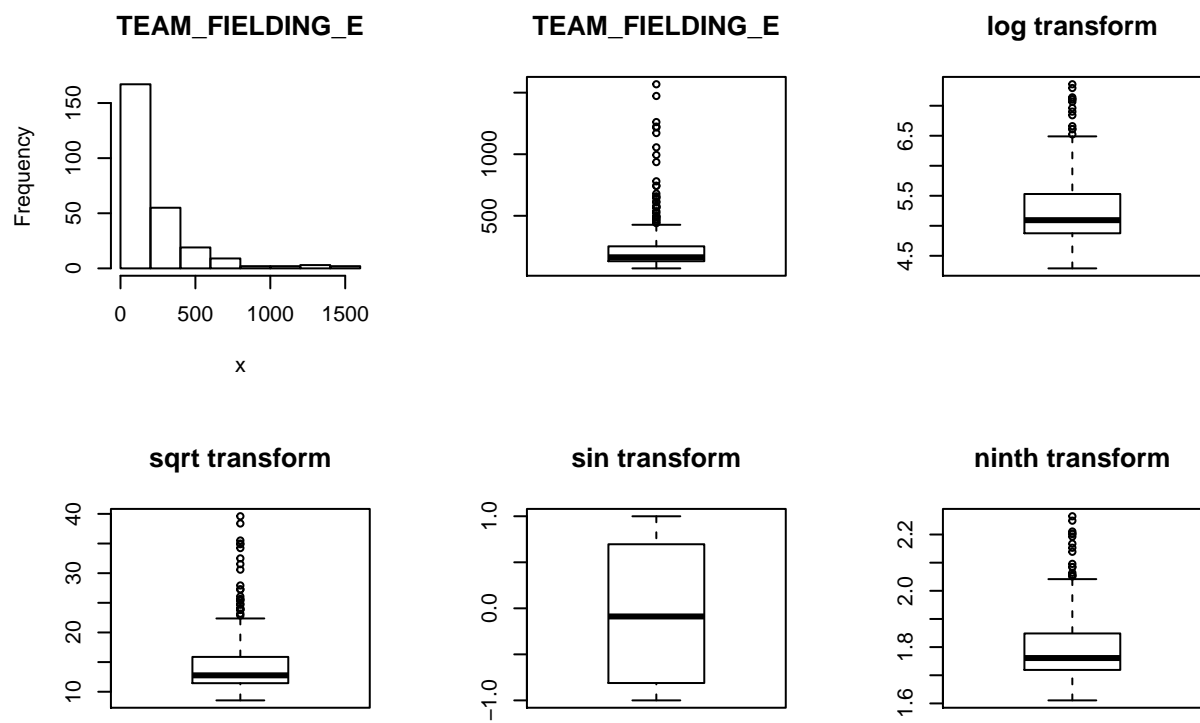



```
show_charts(moneyball2$TEAM_BASERUN_SB)
```

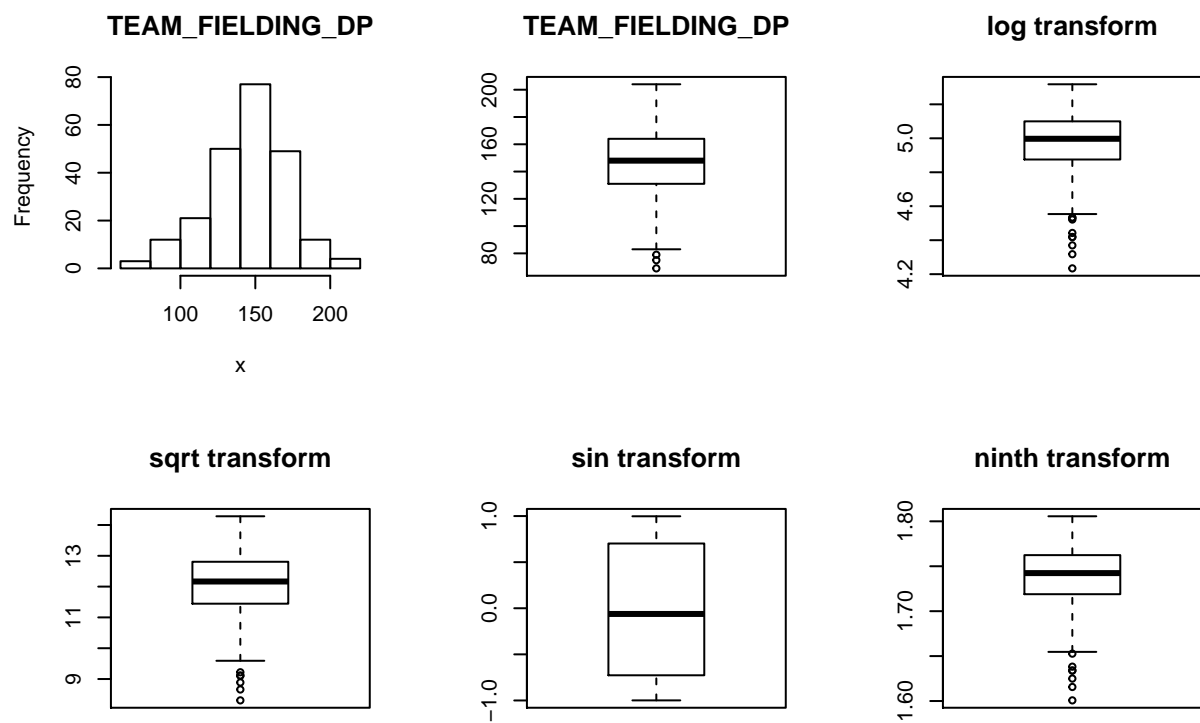
```
## Warning in bplt(at[i], wid = width[i], stats = z$stats[, i], out = z$out[z
## $group == : Outlier (-Inf) in boxplot 1 is not drawn
```



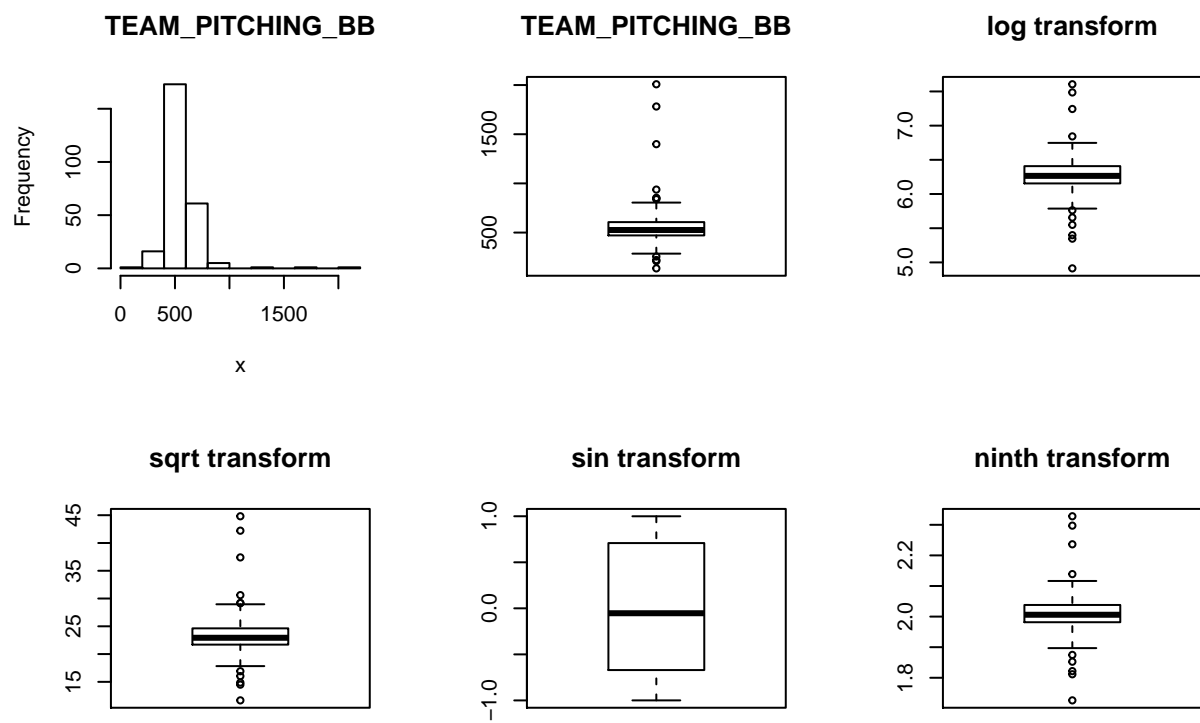
```
show_charts(moneyball12$TEAM_FIELDING_E)
```



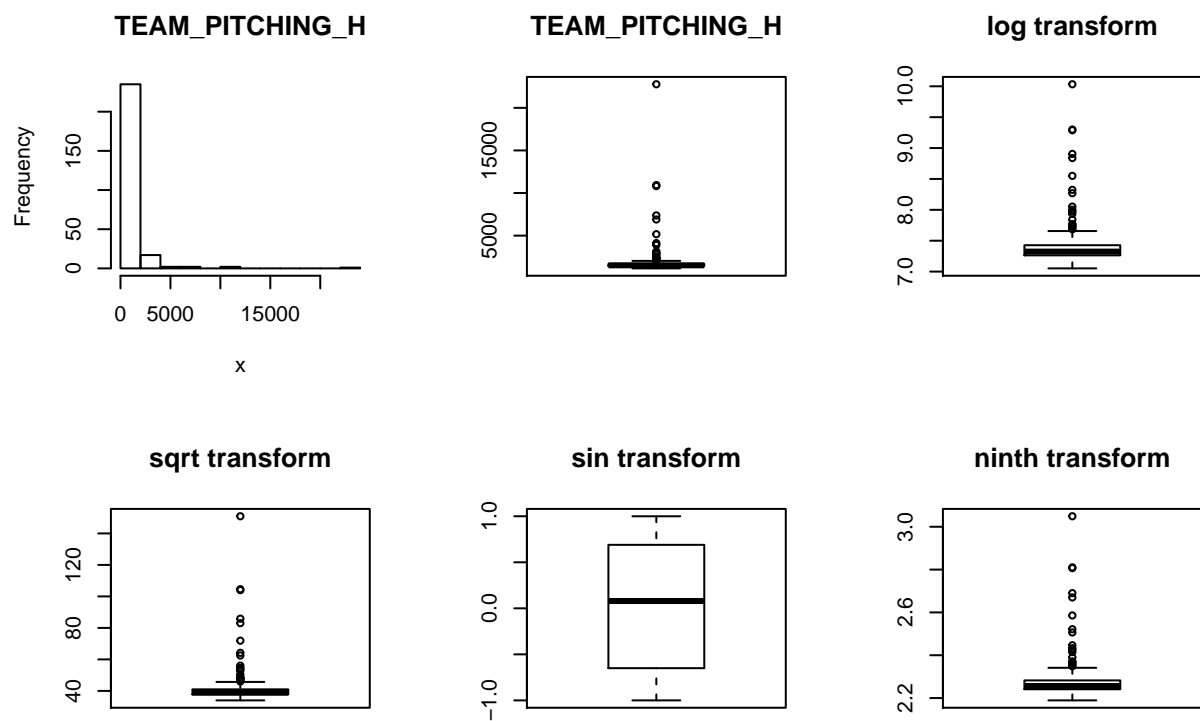
```
show_charts(moneyball12$TEAM_FIELDING_DP)
```



```
show_charts(moneyball12$TEAM_PITCHING_BB)
```

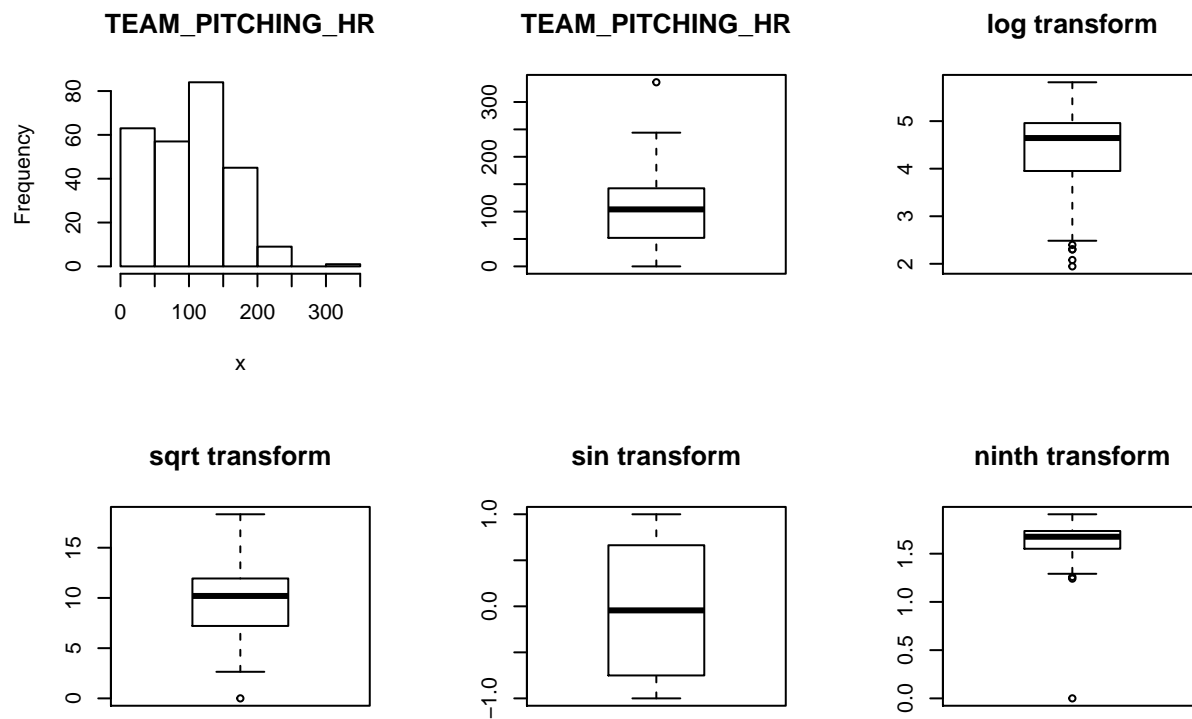


```
show_charts(moneyball12$TEAM_PITCHING_H)
```



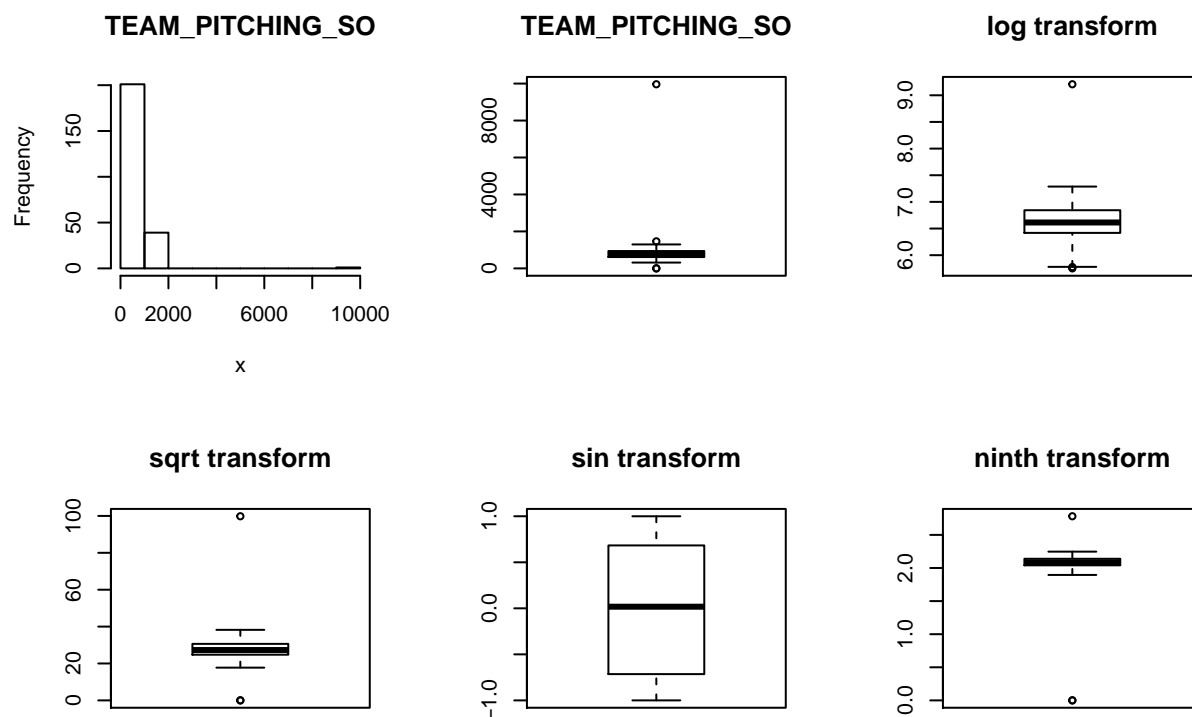
```
show_charts(moneyball12$TEAM_PITCHING_HR)
```

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



```
show_charts(moneyball2$TEAM_PITCHING_S0)
```

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



Data Preparation

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

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

Outliers

For outliers, we will create 2 sets of variables.

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

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

```
TEAM_BATTING_H_NEW TEAM_BATTING_2B_NEW TEAM_BATTING_3B_NEW TEAM_BATTING_BB_NEW
TEAM_BASERUN_SB_NEW TEAM_FIELDING_E_NEW TEAM_FIELDING_DP_NEW
TEAM_PITCHING_BB_NEW TEAM_PITCHING_H_NEW TEAM_PITCHING_HR_NEW
TEAM_PITCHING_SO_NEW
```

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

```
treat_outliers <- function(x) {
  qnt <- quantile(x, probs=c(.25, .75), na.rm = T)
```



```

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

return(x)
}

moneyball12$TEAM_BATTING_H_NEW <- treat_outliers(moneyball12$TEAM_BATTING_H)
moneyball12$TEAM_BATTING_2B_NEW <- treat_outliers(moneyball12$TEAM_BATTING_2B)
moneyball12$TEAM_BATTING_3B_NEW <- treat_outliers(moneyball12$TEAM_BATTING_3B)
moneyball12$TEAM_BATTING_BB_NEW <- treat_outliers(moneyball12$TEAM_BATTING_BB)
moneyball12$TEAM_BASERUN_SB_NEW <- treat_outliers(moneyball12$TEAM_BASERUN_SB)
moneyball12$TEAM_FIELDING_E_NEW <- treat_outliers(moneyball12$TEAM_FIELDING_E)
moneyball12$TEAM_FIELDING_DP_NEW <- treat_outliers(moneyball12$TEAM_FIELDING_DP)
moneyball12$TEAM_PITCHING_BB_NEW <- treat_outliers(moneyball12$TEAM_PITCHING_BB)
moneyball12$TEAM_PITCHING_H_NEW <- treat_outliers(moneyball12$TEAM_PITCHING_H)
moneyball12$TEAM_PITCHING_HR_NEW <- treat_outliers(moneyball12$TEAM_PITCHING_HR)
moneyball12$TEAM_PITCHING_SO_NEW <- treat_outliers(moneyball12$TEAM_PITCHING_SO)

```

Lets see how the new variables look in boxplots.

```

par(mfrow=c(3,4))

boxplot(moneyball12$TEAM_BATTING_H_NEW,main="TEAM_BATTING_H_NEW")
boxplot(moneyball12$TEAM_BATTING_2B_NEW,main="TEAM_BATTING_2B_NEW")
boxplot(moneyball12$TEAM_BATTING_3B_NEW,main="TEAM_BATTING_3B_NEW")
boxplot(moneyball12$TEAM_BATTING_BB_NEW,main="TEAM_BATTING_BB_NEW")
boxplot(moneyball12$TEAM_BASERUN_SB_NEW,main="TEAM_BASERUN_SB_NEW")
boxplot(moneyball12$TEAM_FIELDING_E_NEW,main="TEAM_FIELDING_E_NEW")
boxplot(moneyball12$TEAM_FIELDING_DP_NEW,main="TEAM_FIELDING_DP_NEW")
boxplot(moneyball12$TEAM_PITCHING_BB_NEW,main="TEAM_PITCHING_BB_NEW")
boxplot(moneyball12$TEAM_PITCHING_H_NEW,main="TEAM_PITCHING_H_NEW")
boxplot(moneyball12$TEAM_PITCHING_HR_NEW,main="TEAM_PITCHING_HR_NEW")
boxplot(moneyball12$TEAM_PITCHING_SO_NEW,main="TEAM_PITCHING_SO_NEW")

```

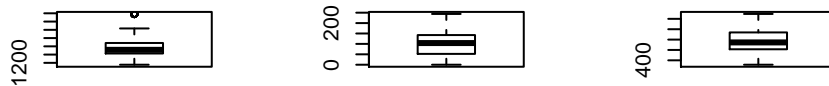
TEAM_BATTING_H_NE TEAM_BATTING_2B_NI TEAM_BATTING_3B_NI TEAM_BATTING_BB_NI



TEAM_BASERUN_SB_N TEAM_FIELDING_E_NE TEAM_FIELDING_DP_N TEAM_PITCHING_BB_N



TEAM_PITCHING_H_NI TEAM_PITCHING_HR_N TEAM_PITCHING_SO_N



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

TEAM_BATTING_H_SIN TEAM_BATTING_2B_SIN TEAM_BATTING_3B_SIN TEAM_BATTING_BB_SIN
TEAM_BASERUN_SB_SIN TEAM_FIELDING_E_SIN TEAM_FIELDING_DP_SIN TEAM_PITCHING_BB_SIN
TEAM_PITCHING_H_SIN TEAM_PITCHING_HR_SIN TEAM_PITCHING_SO_SIN

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

Missing Values

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

TEAM_BATTING_SO_NEW

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

```
TEAM_PITCHING_SO_NEW TEAM_BASERUN_SB_NEW TEAM_FIELDING_DP_NEW
```

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

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

Additional Variables

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

Missing Flags

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

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

Ratios

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

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

Calculated Variables

Finally, we create some calculated variables as below:

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

```
moneyball2$TEAM_BATTING_EB <- moneyball2$TEAM_BATTING_2B + moneyball2$TEAM_BATTING_3B + moneyball2$TEAM_BATTING_HR
moneyball2$TEAM_BATTING_1B <- moneyball2$TEAM_BATTING_H - moneyball2$TEAM_BATTING_EB
summary(moneyball2)
```

```
## TEAM_BATTING_H TEAM_BATTING_2B TEAM_BATTING_3B TEAM_BATTING_HR
## Min. : 819 Min. : 44.0 Min. : 14.00 Min. : 0.00
## 1st Qu.:1387 1st Qu.:210.0 1st Qu.: 35.00 1st Qu.: 44.50
## Median :1455 Median :239.0 Median : 52.00 Median :101.00
## Mean :1469 Mean :241.3 Mean : 55.91 Mean : 95.63
## 3rd Qu.:1548 3rd Qu.:278.5 3rd Qu.: 72.00 3rd Qu.:135.50
```

```

## Max. :2170 Max. :376.0 Max. :155.00 Max. :242.00
##
## TEAM_BATTING_BB TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS
## Min. : 15.0 Min. : 0.0 Min. : 0.0 Min. : 0.00
## 1st Qu.:436.5 1st Qu.: 545.0 1st Qu.: 59.0 1st Qu.: 38.00
## Median :509.0 Median : 686.0 Median : 92.0 Median : 49.50
## Mean :499.0 Mean : 709.3 Mean :123.7 Mean : 52.32
## 3rd Qu.:565.5 3rd Qu.: 912.0 3rd Qu.:151.8 3rd Qu.: 63.00
## Max. :792.0 Max. :1268.0 Max. :580.0 Max. :154.00
## NA's :18 NA's :13 NA's :87
## TEAM_BATTING_HBP TEAM_PITCHING_H TEAM_PITCHING_HR TEAM_PITCHING_BB
## Min. :42.00 Min. : 1155 Min. : 0.0 Min. : 136.0
## 1st Qu.:53.50 1st Qu.: 1426 1st Qu.: 52.0 1st Qu.: 471.0
## Median :62.00 Median : 1515 Median :104.0 Median : 526.0
## Mean :62.37 Mean : 1813 Mean :102.1 Mean : 552.4
## 3rd Qu.:67.50 3rd Qu.: 1681 3rd Qu.:142.5 3rd Qu.: 606.5
## Max. :96.00 Max. :22768 Max. :336.0 Max. :2008.0
## NA's :240
## TEAM_PITCHING_SO TEAM_FIELDING_E TEAM_FIELDING_DP TEAM_BATTING_H_NEW
## Min. : 0.0 Min. : 73.0 Min. : 69.0 Min. :1149
## 1st Qu.: 613.0 1st Qu.: 131.0 1st Qu.:131.0 1st Qu.:1387
## Median : 745.0 Median : 163.0 Median :148.0 Median :1455
## Mean : 799.7 Mean : 249.7 Mean :146.1 Mean :1467
## 3rd Qu.: 938.0 3rd Qu.: 252.0 3rd Qu.:164.0 3rd Qu.:1548
## Max. :9963.0 Max. :1568.0 Max. :204.0 Max. :1775
## NA's :18 NA's :31
## TEAM_BATTING_2B_NEW TEAM_BATTING_3B_NEW TEAM_BATTING_BB_NEW
## Min. :116.0 Min. : 14.00 Min. :259.0
## 1st Qu.:210.0 1st Qu.: 35.00 1st Qu.:436.5
## Median :239.0 Median : 52.00 Median :509.0
## Mean :242.1 Mean : 55.26 Mean :503.8
## 3rd Qu.:278.5 3rd Qu.: 72.00 3rd Qu.:565.5
## Max. :376.0 Max. :126.00 Max. :746.0
##
## TEAM_BASERUN_SB_NEW TEAM_FIELDING_E_NEW TEAM_FIELDING_DP_NEW
## Min. : 0.0 Min. : 73.0 Min. : 83.0
## 1st Qu.: 60.5 1st Qu.:131.0 1st Qu.:134.5
## Median : 96.0 Median :163.0 Median :146.4
## Mean :120.2 Mean :238.8 Mean :146.4
## 3rd Qu.:149.0 3rd Qu.:252.0 3rd Qu.:160.5
## Max. :319.0 Max. :660.2 Max. :204.0
##
## TEAM_PITCHING_BB_NEW TEAM_PITCHING_H_NEW TEAM_PITCHING_HR_NEW
## Min. :286.0 Min. :1155 Min. : 0.0
## 1st Qu.:471.0 1st Qu.:1426 1st Qu.: 52.0
## Median :526.0 Median :1515 Median :104.0
## Mean :542.3 Mean :1606 Mean :101.6
## 3rd Qu.:606.5 3rd Qu.:1681 3rd Qu.:142.5
## Max. :805.0 Max. :2385 Max. :244.0
##
## TEAM_PITCHING_SO_NEW TEAM_BATTING_H_SIN TEAM_BATTING_2B_SIN
## Min. : 315.0 Min. : -0.99999 Min. : -0.99976
## 1st Qu.: 622.5 1st Qu.: -0.67675 1st Qu.: -0.80113
## Median : 764.8 Median : 0.05320 Median : 0.01770

```

```

## Mean      : 764.8      Mean      : 0.02619      Mean      :-0.01888
## 3rd Qu.: 927.5      3rd Qu.: 0.74501      3rd Qu.: 0.70868
## Max.      :1295.0      Max.      : 0.99991      Max.      : 0.99991
##
## TEAM_BATTING_3B_SIN TEAM_BATTING_BB_SIN TEAM_BASERUN_SB_SIN
## Min.      :-0.9998      Min.      :-0.99999      Min.      :-0.99975
## 1st Qu.: -0.7271      1st Qu.: -0.59492      1st Qu.: -0.68318
## Median : 0.1499      Median : 0.07961      Median : 0.14546
## Mean      : 0.0324      Mean      : 0.06589      Mean      : 0.06904
## 3rd Qu.: 0.7626      3rd Qu.: 0.73317      3rd Qu.: 0.81676
## Max.      : 0.9999      Max.      : 0.99976      Max.      : 0.99952
##                                     NA's      :13
## TEAM_FIELDING_E_SIN TEAM_FIELDING_DP_SIN TEAM_PITCHING_BB_SIN
## Min.      :-0.99999      Min.      :-0.999207      Min.      :-0.999522
## 1st Qu.: -0.81160      1st Qu.: -0.727143      1st Qu.: -0.670218
## Median : -0.08834      Median : -0.061920      Median : -0.053114
## Mean      :-0.05169      Mean      :-0.008299      Mean      :-0.003834
## 3rd Qu.: 0.69608      3rd Qu.: 0.699225      3rd Qu.: 0.708642
## Max.      : 0.99991      Max.      : 0.997799      Max.      : 0.999757
##                                     NA's      :31
## TEAM_PITCHING_H_SIN TEAM_PITCHING_HR_SIN TEAM_PITCHING_SO_SIN
## Min.      :-0.99999      Min.      :-0.99999      Min.      :-0.99999
## 1st Qu.: -0.65017      1st Qu.: -0.75099      1st Qu.: -0.71483
## Median : 0.07940      Median : -0.04424      Median : 0.01761
## Mean      : 0.01651      Mean      :-0.03292      Mean      :-0.01170
## 3rd Qu.: 0.68966      3rd Qu.: 0.66361      3rd Qu.: 0.68320
## Max.      : 0.99991      Max.      : 0.99991      Max.      : 0.99991
##                                     NA's      :18
## TEAM_BATTING_SO_NEW TEAM_BATTING_HBP_Missing TEAM_BASERUN_CS_Missing
## Min.      : 0.0      Min.      :0.00000      Min.      :0.0000
## 1st Qu.: 565.0      1st Qu.:0.00000      1st Qu.:0.0000
## Median : 709.3      Median :0.00000      Median :1.0000
## Mean      : 709.3      Mean      :0.07336      Mean      :0.6641
## 3rd Qu.: 904.5      3rd Qu.:0.00000      3rd Qu.:1.0000
## Max.      :1268.0      Max.      :1.00000      Max.      :1.0000
##
## Hits_R      Walks_R      HomeRuns_R      Strikeout_R
## Min.      :0.0679      Min.      :0.06787      Min.      :0.1111      Min.      :0.0679
## 1st Qu.:0.9382      1st Qu.:0.93871      1st Qu.:0.9358      1st Qu.:0.9388
## Median :0.9506      Median :0.95067      Median :0.9516      Median :0.9508
## Mean      :0.9168      Mean      :0.91681      Mean      :0.9199      Mean      :0.9204
## 3rd Qu.:1.0000      3rd Qu.:1.00000      3rd Qu.:1.0000      3rd Qu.:1.0000
## Max.      :1.0187      Max.      :1.01786      Max.      :1.0204      Max.      :1.0189
##                                     NA's      :1      NA's      :20
## TEAM_BATTING_EB TEAM_BATTING_1B
## Min.      : 73.0      Min.      : 657.0
## 1st Qu.:339.5      1st Qu.: 990.5
## Median :397.0      Median :1059.0
## Mean      :392.9      Mean      :1076.5
## 3rd Qu.:450.0      3rd Qu.:1134.0
## Max.      :615.0      Max.      :1846.0
##

```