

Assignment1

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Load Data

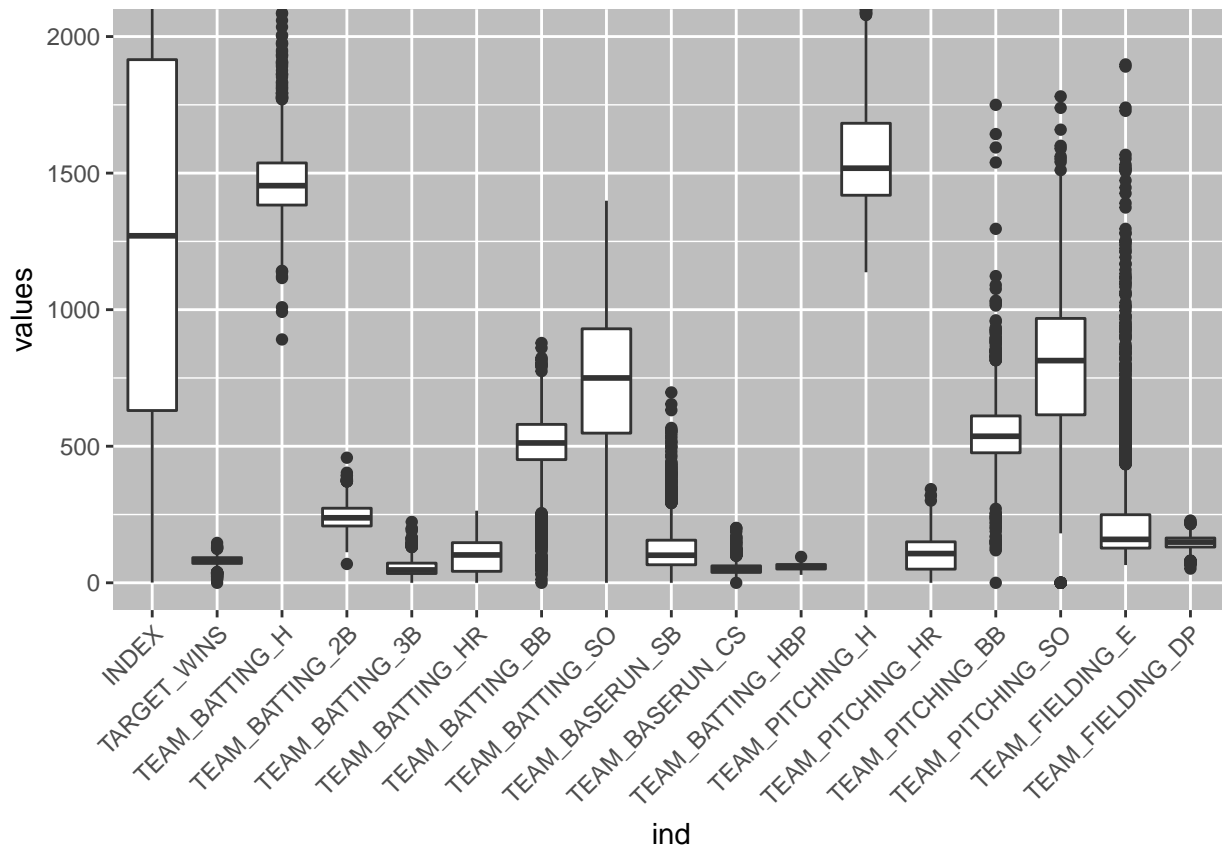
```
file1 <- 'C:/Users/andre/Downloads/moneyball-evaluation-data.csv'
eval<-read.csv(file1)
file2 <- 'C:/Users/andre/Downloads/moneyball-training-data.csv'
train<-read.csv(file2)
summary(train)
```

```
##      INDEX      TARGET_WINS      TEAM_BATTING_H TEAM_BATTING_2B
##  Min.   : 1.0    Min.   : 0.00    Min.   : 891    Min.   : 69.0
## 1st Qu.: 630.8  1st Qu.: 71.00    1st Qu.:1383   1st Qu.:208.0
## Median :1270.5  Median : 82.00    Median :1454   Median :238.0
## Mean   :1268.5  Mean   : 80.79    Mean   :1469   Mean   :241.2
## 3rd Qu.:1915.5  3rd Qu.: 92.00    3rd Qu.:1537   3rd Qu.:273.0
## Max.   :2535.0  Max.   :146.00    Max.   :2554   Max.   :458.0
##
## TEAM_BATTING_3B TEAM_BATTING_HR TEAM_BATTING_BB TEAM_BATTING_SO
##  Min.   : 0.00    Min.   : 0.00    Min.   : 0.0    Min.   : 0.0
## 1st Qu.: 34.00    1st Qu.: 42.00    1st Qu.:451.0   1st Qu.: 548.0
## Median : 47.00    Median :102.00    Median :512.0   Median : 750.0
## Mean   : 55.25    Mean   : 99.61    Mean   :501.6   Mean   : 735.6
## 3rd Qu.: 72.00    3rd Qu.:147.00    3rd Qu.:580.0   3rd Qu.: 930.0
## Max.   :223.00    Max.   :264.00    Max.   :878.0   Max.   :1399.0
##                      NA's   :102
## TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_BATTING_HBP TEAM_PITCHING_H
##  Min.   : 0.0    Min.   : 0.0    Min.   :29.00    Min.   : 1137
## 1st Qu.: 66.0    1st Qu.: 38.0    1st Qu.:50.50    1st Qu.: 1419
## Median :101.0    Median : 49.0    Median :58.00    Median : 1518
## Mean   :124.8    Mean   : 52.8    Mean   :59.36    Mean   : 1779
## 3rd Qu.:156.0    3rd Qu.: 62.0    3rd Qu.:67.00    3rd Qu.: 1682
## Max.   :697.0    Max.   :201.0    Max.   :95.00    Max.   :30132
## NA's   :131     NA's   :772     NA's   :2085
## TEAM_PITCHING_HR TEAM_PITCHING_BB TEAM_PITCHING_SO TEAM_FIELDING_E
##  Min.   : 0.0    Min.   : 0.0    Min.   : 0.0    Min.   : 65.0
## 1st Qu.: 50.0    1st Qu.: 476.0   1st Qu.: 615.0   1st Qu.: 127.0
## Median :107.0    Median : 536.5   Median : 813.5   Median : 159.0
## Mean   :105.7    Mean   : 553.0   Mean   : 817.7   Mean   : 246.5
## 3rd Qu.:150.0    3rd Qu.: 611.0   3rd Qu.: 968.0   3rd Qu.: 249.2
## Max.   :343.0    Max.   :3645.0   Max.   :19278.0   Max.   :1898.0
##                      NA's   :102
## TEAM_FIELDING_DP
```

```
## Min.   : 52.0
## 1st Qu.:131.0
## Median :149.0
## Mean   :146.4
## 3rd Qu.:164.0
## Max.   :228.0
## NA's   :286
```

Find Outlier

```
## Warning: Removed 3478 rows containing non-finite values (stat_boxplot).
```



```
## Remove TEAM_BATTING_HBP since it has mostly missing values
```

```
train <- train[, -1 ]
train <- train[, -10]
```

Find correlations

```
cor(drop_na(train))
```

```
##
## TARGET_WINS      TARGET_WINS TEAM_BATTING_H TEAM_BATTING_2B
## TEAM_BATTING_H    0.35910266  1.000000000    0.67597673
## TEAM_BATTING_2B   0.19377063   0.675976730    1.00000000
```

##	TEAM_BATTING_3B	0.08273171	0.424243387	0.04159537
##	TEAM_BATTING_HR	0.28346449	0.121533598	0.31393818
##	TEAM_BATTING_BB	0.34829184	0.169265952	0.19769347
##	TEAM_BATTING_SO	-0.06289963	-0.365563214	0.07894699
##	TEAM_BASERUN_SB	0.12035129	0.006327368	0.06540808
##	TEAM_BASERUN_CS	-0.01119995	0.014056697	-0.11760390
##	TEAM_PITCHING_H	0.21610375	0.687898063	0.45457406
##	TEAM_PITCHING_HR	0.27872264	0.136743031	0.31787719
##	TEAM_PITCHING_BB	0.29493756	0.188208391	0.19232998
##	TEAM_PITCHING_SO	-0.06673654	-0.330034165	0.09116245
##	TEAM_FIELDING_E	-0.25450741	0.121279902	-0.19416232
##	TEAM_FIELDING_DP	-0.04949709	0.148619532	0.04710052
##	TEAM_BATTING_3B	TEAM_BATTING_HR	TEAM_BATTING_BB	
##	TARGET_WINS	0.08273171	0.283464494	0.34829184
##	TEAM_BATTING_H	0.42424339	0.121533598	0.16926595
##	TEAM_BATTING_2B	0.04159537	0.313938180	0.19769347
##	TEAM_BATTING_3B	1.00000000	-0.554984363	-0.09728038
##	TEAM_BATTING_HR	-0.55498436	1.000000000	0.28987751
##	TEAM_BATTING_BB	-0.09728038	0.289877505	1.00000000
##	TEAM_BATTING_SO	-0.69034724	0.640159102	0.02384741
##	TEAM_BASERUN_SB	0.04553554	-0.138348758	-0.07810816
##	TEAM_BASERUN_CS	0.36934182	-0.470715823	-0.20628472
##	TEAM_PITCHING_H	0.35643177	-0.002382248	0.12825248
##	TEAM_PITCHING_HR	-0.52736563	0.971652181	0.29599870
##	TEAM_PITCHING_BB	-0.02376740	0.200030244	0.87467545
##	TEAM_PITCHING_SO	-0.63929275	0.597391699	0.03414065
##	TEAM_FIELDING_E	0.64109245	-0.621227175	-0.16115860
##	TEAM_FIELDING_DP	0.00887003	0.070219016	0.16343997
##	TEAM_BATTING_SO	TEAM_BASERUN_SB	TEAM_BASERUN_CS	
##	TARGET_WINS	-0.06289963	0.120351295	-0.01119995
##	TEAM_BATTING_H	-0.36556321	0.006327368	0.01405670
##	TEAM_BATTING_2B	0.07894699	0.065408077	-0.11760390
##	TEAM_BATTING_3B	-0.69034724	0.045535545	0.36934182
##	TEAM_BATTING_HR	0.64015910	-0.138348758	-0.47071582
##	TEAM_BATTING_BB	0.02384741	-0.078108165	-0.20628472
##	TEAM_BATTING_SO	1.00000000	0.126866866	-0.26423776
##	TEAM_BASERUN_SB	0.12686687	1.000000000	0.65233884
##	TEAM_BASERUN_CS	-0.26423776	0.652338836	1.00000000
##	TEAM_PITCHING_H	-0.34335935	0.021181111	0.06909025
##	TEAM_PITCHING_HR	0.60500635	-0.130403756	-0.45085011
##	TEAM_PITCHING_BB	-0.05535063	-0.059437263	-0.14133608
##	TEAM_PITCHING_SO	0.93221165	0.134325496	-0.23168177
##	TEAM_FIELDING_E	-0.62668120	-0.011939775	0.42695850
##	TEAM_FIELDING_DP	-0.13647576	-0.255985407	-0.21424801
##	TEAM_PITCHING_H	TEAM_PITCHING_HR	TEAM_PITCHING_BB	
##	TARGET_WINS	0.216103749	0.27872264	0.29493756
##	TEAM_BATTING_H	0.687898063	0.13674303	0.18820839
##	TEAM_BATTING_2B	0.454574060	0.31787719	0.19232998
##	TEAM_BATTING_3B	0.356431774	-0.52736563	-0.02376740
##	TEAM_BATTING_HR	-0.002382248	0.97165218	0.20003024
##	TEAM_BATTING_BB	0.128252478	0.29599870	0.87467545
##	TEAM_BATTING_SO	-0.343359353	0.60500635	-0.05535063
##	TEAM_BASERUN_SB	0.021181111	-0.13040376	-0.05943726
##	TEAM_BASERUN_CS	0.069090246	-0.45085011	-0.14133608

```
## TEAM_PITCHING_H      1.000000000      0.17271371      0.48745563
## TEAM_PITCHING_HR     0.172713710      1.00000000      0.31601752
## TEAM_PITCHING_BB     0.487455634      0.31601752      1.00000000
## TEAM_PITCHING_SO    -0.056621268      0.64681064      0.12892998
## TEAM_FIELDING_E      0.144892622     -0.60020742     -0.09659142
## TEAM_FIELDING_DP     0.122916134      0.07928388      0.16263604
##
##      TEAM_PITCHING_SO TEAM_FIELDING_E TEAM_FIELDING_DP
## TARGET_WINS          -0.06673654     -0.25450741     -0.04949709
## TEAM_BATTING_H        -0.33003416      0.12127990      0.14861953
## TEAM_BATTING_2B        0.09116245     -0.19416232      0.04710052
## TEAM_BATTING_3B       -0.63929275      0.64109245      0.00887003
## TEAM_BATTING_HR        0.59739170     -0.62122718      0.07021902
## TEAM_BATTING_BB        0.03414065     -0.16115860      0.16343997
## TEAM_BATTING_SO        0.93221165     -0.62668120     -0.13647576
## TEAM_BASERUN_SB        0.13432550     -0.01193978     -0.25598541
## TEAM_BASERUN_CS       -0.23168177      0.42695850     -0.21424801
## TEAM_PITCHING_H       -0.05662127      0.14489262      0.12291613
## TEAM_PITCHING_HR       0.64681064     -0.60020742      0.07928388
## TEAM_PITCHING_BB       0.12892998     -0.09659142      0.16263604
## TEAM_PITCHING_SO      1.000000000     -0.58691832     -0.11885440
## TEAM_FIELDING_E       -0.58691832      1.00000000     -0.07870923
## TEAM_FIELDING_DP      -0.11885440     -0.07870923      1.00000000
```

TEAM_PITCHING_HR/BB and TEAM_BATTING_HR/BB are highly correlated, so we can remove one of them.

```
train <- train[, -11]
train<-train[, -11]
```

```
refin_data <- mice(train, m=5, maxit = 5, method = 'pmm')
```

```
##
## iter imp variable
## 1 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
```

```
## 5 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP

refin_data <- complete(refin_data)
```

m1

```
model1 <- lm(TARGET_WINS ~ ., refin_data)
summary(model1)

##
## Call:
## lm(formula = TARGET_WINS ~ ., data = refin_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -51.294  -8.288   0.171   8.147  47.733
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    35.8904382    5.0364434     7.126 1.38e-12 ***
## TEAM_BATTING_H     0.0426396    0.0035568    11.988 < 2e-16 ***
## TEAM_BATTING_2B   -0.0196205    0.0088679    -2.213 0.027029 *
## TEAM_BATTING_3B     0.0316733    0.0164089     1.930 0.053700 .
## TEAM_BATTING_HR     0.0857039    0.0093855     9.132 < 2e-16 ***
## TEAM_BATTING_BB     0.0073853    0.0030547     2.418 0.015699 *
## TEAM_BATTING_SO   -0.0159297    0.0023813    -6.690 2.81e-11 ***
## TEAM_BASERUN_SB     0.0542472    0.0052481    10.337 < 2e-16 ***
## TEAM_BASERUN_CS   -0.0075826    0.0104039    -0.729 0.466188
## TEAM_PITCHING_H     0.0011855    0.0003413     3.473 0.000524 ***
## TEAM_PITCHING_SO    0.0009161    0.0006714     1.364 0.172593
## TEAM_FIELDING_E    -0.0399249    0.0026255   -15.207 < 2e-16 ***
## TEAM_FIELDING_DP  -0.1071835    0.0128911    -8.315 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.6 on 2263 degrees of freedom
## Multiple R-squared:  0.3637, Adjusted R-squared:  0.3604
## F-statistic: 107.8 on 12 and 2263 DF, p-value: < 2.2e-16
```

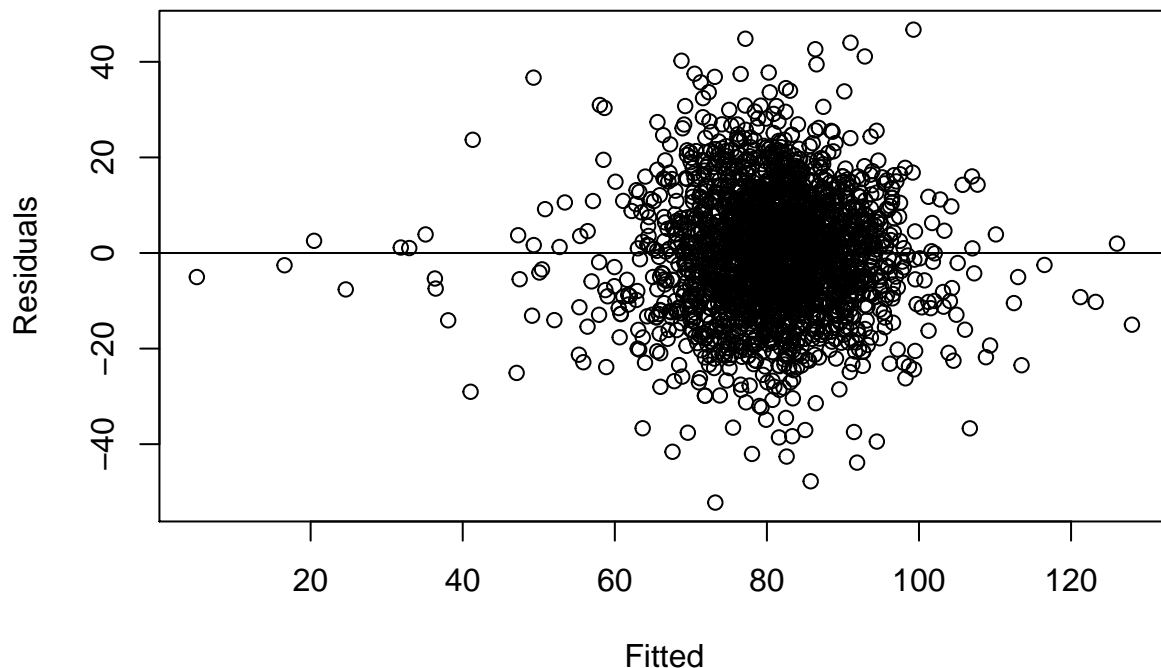
m2

```
model2 <- lm(TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HR + TEAM_BATTING_SO + TEAM_BASERUN_SB + TEAM_PITCHING_H + TEAM_FIELDING_E + TEAM_FIELDING_DP, data = refin_data)
summary(model2)

##
## Call:
## lm(formula = TARGET_WINS ~ TEAM_BATTING_H + TEAM_BATTING_HR +
##      TEAM_BATTING_SO + TEAM_BASERUN_SB + TEAM_PITCHING_H + TEAM_FIELDING_E +
##      TEAM_FIELDING_DP, data = refin_data)
##
```

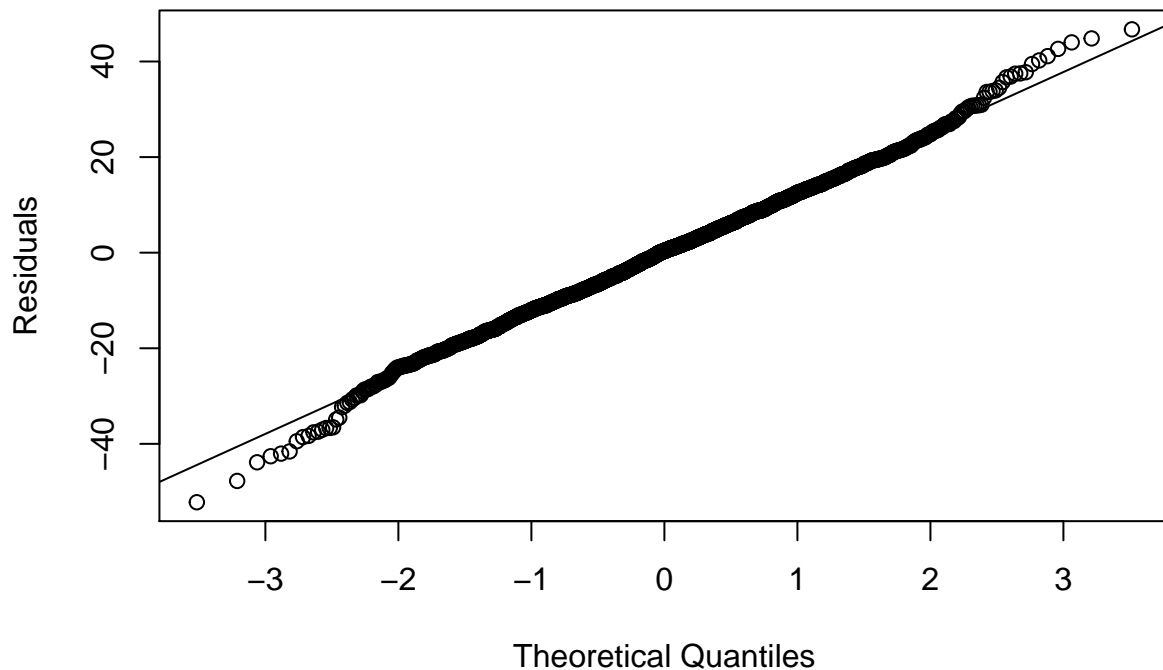
```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -52.217  -8.574   0.357   8.449  46.727
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  42.9700167  4.3626536   9.850 < 2e-16 ***
## TEAM_BATTING_H  0.0393555  0.0024842  15.842 < 2e-16 ***
## TEAM_BATTING_HR  0.0859065  0.0085881  10.003 < 2e-16 ***
## TEAM_BATTING_SO -0.0176053  0.0021032  -8.371 < 2e-16 ***
## TEAM_BASERUN_SB  0.0570343  0.0041535  13.732 < 2e-16 ***
## TEAM_PITCHING_H  0.0012536  0.0002898   4.325 1.59e-05 ***
## TEAM_FIELDING_E -0.0420590  0.0023456 -17.931 < 2e-16 ***
## TEAM_FIELDING_DP -0.1080635  0.0126006  -8.576 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.63 on 2268 degrees of freedom
## Multiple R-squared:  0.359, Adjusted R-squared:  0.357
## F-statistic: 181.4 on 7 and 2268 DF, p-value: < 2.2e-16
```

```
plot(fitted(model2),residuals(model2),xlab = "Fitted", ylab = "Residuals")
abline(h=0)
```



```
par(mfrow = c(1,1))
qqnorm(residuals(model2),ylab = "Residuals")
qqline(residuals(model2))
```

Normal Q-Q Plot



```
## transfrom data
```

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 3.4.4
```

```
library(e1071)
```

```
## Warning: package 'e1071' was built under R version 3.4.4
```

```
t = preProcess(refin_data,  
               c("BoxCox", "center", "scale"))  
refin_data = data.frame(  
  t = predict(t, refin_data))
```

m3

```
model3 <- lm(t.TARGET_WINS ~ t.TEAM_BATTING_H + t.TEAM_BATTING_HR + t.TEAM_BATTING_SO + t.TEAM_BASERUN_SB +  
             t.TEAM_PITCHING_H + t.TEAM_FIELDING_E + t.TEAM_FIELDING_DP, data = refin_data)  
summary(model3)
```

```
##
```

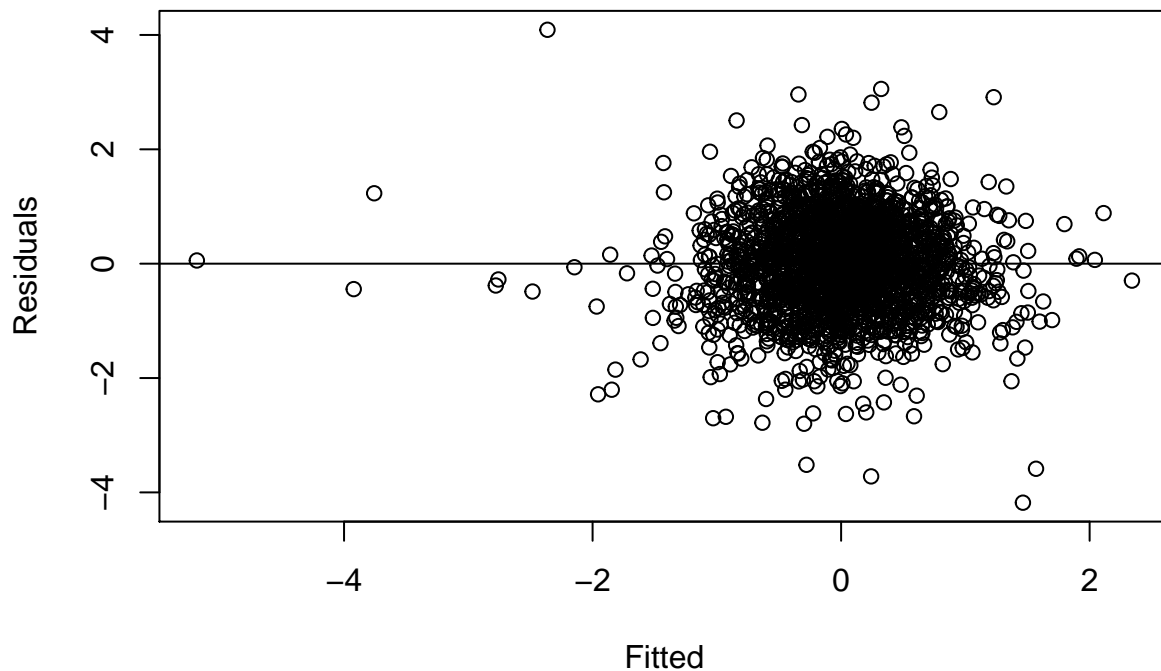
```
## Call:
```

```
## lm(formula = t.TARGET_WINS ~ t.TEAM_BATTING_H + t.TEAM_BATTING_HR +  
##     t.TEAM_BATTING_SO + t.TEAM_BASERUN_SB + t.TEAM_PITCHING_H +  
##     t.TEAM_FIELDING_E + t.TEAM_FIELDING_DP, data = refin_data)  
##
```

```
## Residuals:
```

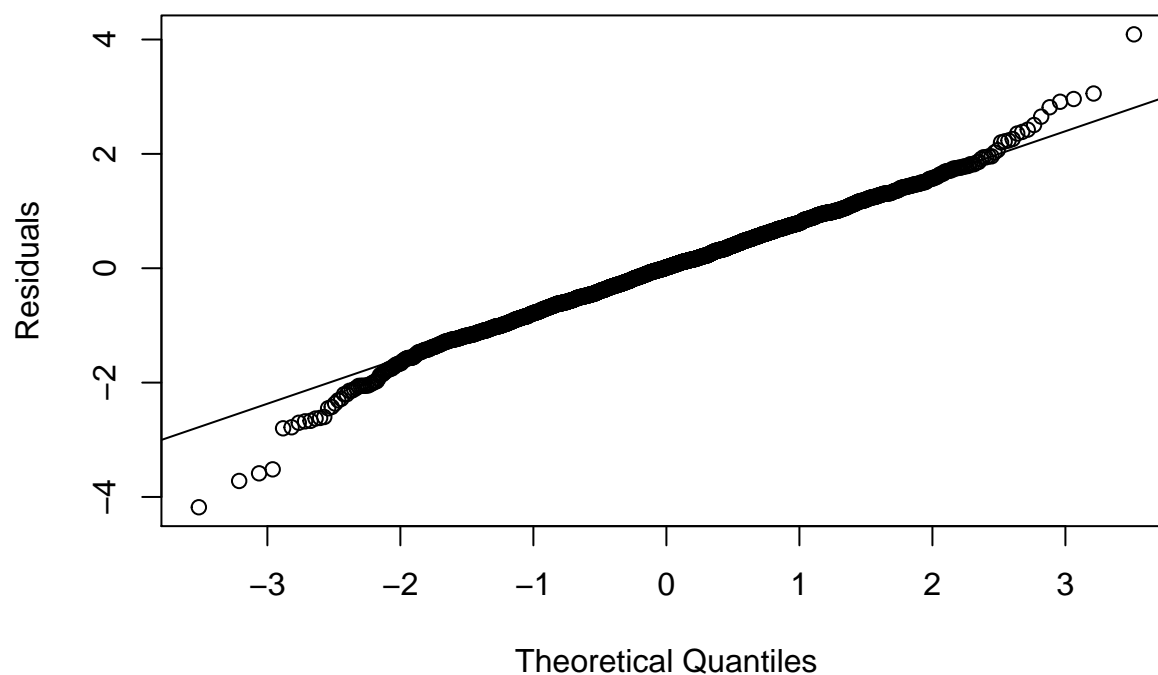
```
##      Min      1Q  Median      3Q      Max
## -4.1797 -0.5224  0.0154  0.5492  4.0903
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.657e-11  1.733e-02   0.000      1
## t.TEAM_BATTING_H  4.264e-01  2.706e-02  15.758 < 2e-16 ***
## t.TEAM_BATTING_HR  2.389e-01  3.647e-02   6.551 7.07e-11 ***
## t.TEAM_BATTING_SO -3.503e-01  3.562e-02  -9.836 < 2e-16 ***
## t.TEAM_BASERUN_SB  3.207e-01  2.625e-02  12.214 < 2e-16 ***
## t.TEAM_PITCHING_H -2.050e-01  2.968e-02  -6.906 6.45e-12 ***
## t.TEAM_FIELDING_E -4.434e-01  3.655e-02 -12.129 < 2e-16 ***
## t.TEAM_FIELDING_DP -1.950e-01  2.346e-02  -8.310 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8267 on 2268 degrees of freedom
## Multiple R-squared:  0.3187, Adjusted R-squared:  0.3166
## F-statistic: 151.6 on 7 and 2268 DF,  p-value: < 2.2e-16
```

```
plot(fitted(model3),residuals(model3),xlab = "Fitted", ylab = "Residuals")
abline(h=0)
```



```
par(mfrow = c(1,1))
qqnorm(residuals(model3),ylab = "Residuals")
qqline(residuals(model3))
```


Normal Q-Q Plot



```
eval <- eval[, -1 ]
eval <- eval[, -9]
eval<-eval[, -10]
eval<-eval[, -10]
```

```
refin_val <- mice(eval, m=5, maxit = 5, method = 'pmm')
```

```
##
## iter imp variable
## 1 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 1 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 2 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 3 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
```

```
## 4 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 4 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 1 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 2 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 3 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 4 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
## 5 5 TEAM_BATTING_SO TEAM_BASERUN_SB TEAM_BASERUN_CS TEAM_PITCHING_SO TEAM_FIELDING_DP
```

```
refin_val <- complete(refin_val)
```

```
eval_data <- predict(model2, newdata = refin_val, interval="prediction")
summary(eval_data)
```

```
##      fit      lwr      upr
## Min.   : 16.22   Min.   :-8.941   Min.   : 41.38
## 1st Qu.: 75.13   1st Qu.:50.342   1st Qu.: 99.92
## Median : 81.19   Median :56.373   Median :106.00
## Mean   : 80.56   Mean   :55.742   Mean   :105.37
## 3rd Qu.: 86.52   3rd Qu.:61.728   3rd Qu.:111.32
## Max.   :115.30   Max.   :90.417   Max.   :140.19
```

```
eval_data
```

```
##      fit      lwr      upr
## 1  60.97285 36.154285 85.79141
## 2  63.43760 38.635429 88.23977
## 3  73.51314 48.729361 98.29691
## 4  89.61923 64.833314 114.40514
## 5  61.58643 36.761305 86.41156
## 6  68.67862 43.867437 93.48979
## 7  83.73024 58.915693 108.54479
## 8  77.14634 52.357480 101.93521
## 9  68.83209 44.033289 93.63090
## 10 74.07499 49.285295 98.86469
## 11 69.50261 44.694597 94.31062
## 12 83.30555 58.504931 108.10617
## 13 81.44405 56.634095 106.25400
## 14 84.49484 59.686074 109.30360
## 15 86.88368 62.064998 111.70237
## 16 78.85485 54.060593 103.64911
## 17 73.77952 48.987070 98.57197
## 18 78.61458 53.825631 103.40352
## 19 68.74554 43.917000 93.57408
## 20 92.28935 67.479436 117.09927
## 21 81.28553 56.491090 106.07997
## 22 84.21901 59.427791 109.01023
## 23 77.88608 53.102320 102.66984
## 24 72.73186 47.937566 97.52615
## 25 84.64860 59.866639 109.43056
## 26 90.54215 65.756712 115.32759
## 27 54.85635 29.867634 79.84508
## 28 76.17981 51.387948 100.97167
## 29 82.87747 58.076388 107.67855
## 30 76.44911 51.642905 101.25531
## 31 89.28983 64.499899 114.07975
## 32 84.94880 60.160018 109.73758
```

## 33	80.01289	55.228373	104.79740
## 34	80.86956	56.079937	105.65919
## 35	78.12198	53.337445	102.90652
## 36	86.37383	61.554484	111.19318
## 37	75.45278	50.671004	100.23456
## 38	89.27492	64.473226	114.07662
## 39	85.37468	60.587410	110.16195
## 40	90.90944	66.096859	115.72202
## 41	84.69753	59.909240	109.48582
## 42	91.02530	66.224317	115.82629
## 43	36.29716	11.003657	61.59066
## 44	104.03089	79.134914	128.92687
## 45	97.79480	72.957605	122.63200
## 46	92.21338	67.404664	117.02210
## 47	98.78221	73.967620	123.59681
## 48	76.81005	52.017074	101.60302
## 49	68.11645	43.320140	92.91275
## 50	80.20280	55.407071	104.99852
## 51	78.75877	53.965717	103.55182
## 52	87.28834	62.493930	112.08275
## 53	74.88005	50.091039	99.66907
## 54	73.89821	49.111834	98.68458
## 55	75.28982	50.506542	100.07310
## 56	79.45513	54.674393	104.23586
## 57	91.47595	66.664719	116.28718
## 58	77.10176	52.296578	101.90695
## 59	62.70534	37.884349	87.52634
## 60	81.01344	56.217197	105.80969
## 61	88.61671	63.818031	113.41538
## 62	70.45841	45.642884	95.27394
## 63	88.40913	63.621359	113.19690
## 64	87.96412	63.148747	112.77949
## 65	87.31556	62.515313	112.11580
## 66	104.30730	79.399942	129.21465
## 67	72.22715	47.436905	97.01740
## 68	78.76804	53.965505	103.57057
## 69	80.72255	55.905730	105.53938
## 70	85.10999	60.307430	109.91255
## 71	82.72707	57.915418	107.53872
## 72	71.98849	47.157166	96.81982
## 73	77.59582	52.775805	102.41583
## 74	89.91676	65.078038	114.75548
## 75	81.18776	56.373217	106.00231
## 76	82.94998	58.152439	107.74753
## 77	80.84865	56.064785	105.63251
## 78	84.63057	59.841893	109.41924
## 79	74.02457	49.224217	98.82493
## 80	78.08146	53.286614	102.87630
## 81	87.23304	62.434338	112.03175
## 82	87.82774	63.033840	112.62164
## 83	97.12731	72.313774	121.94085
## 84	74.48265	49.665042	99.30025
## 85	82.25889	57.471991	107.04578
## 86	83.51109	58.716389	108.30580

## 87	84.49422	59.698530	109.28991
## 88	83.51145	58.732130	108.29078
## 89	90.03725	65.244166	114.83033
## 90	91.18429	66.386237	115.98234
## 91	80.14521	55.350287	104.94014
## 92	84.06950	59.002865	109.13613
## 93	74.13646	49.335364	98.93755
## 94	85.26167	60.448218	110.07512
## 95	89.04273	64.234993	113.85047
## 96	86.49910	61.687463	111.31074
## 97	84.84887	60.054082	109.64366
## 98	98.69389	73.848305	123.53947
## 99	83.99347	59.189934	108.79700
## 100	85.22964	60.416514	110.04277
## 101	78.95298	54.154408	103.75155
## 102	74.89261	50.093955	99.69126
## 103	83.94581	59.159003	108.73262
## 104	83.02619	58.223678	107.82870
## 105	77.87008	53.060561	102.67959
## 106	70.68295	45.867718	95.49818
## 107	57.77746	32.890931	82.66398
## 108	78.34319	53.541562	103.14482
## 109	86.78630	61.992470	111.58014
## 110	57.56474	32.689286	82.44019
## 111	85.33855	60.549669	110.12743
## 112	89.18259	64.382979	113.98220
## 113	95.09320	70.300902	119.88549
## 114	93.17635	68.389816	117.96289
## 115	82.41319	57.633954	107.19243
## 116	79.43268	54.644287	104.22108
## 117	85.72630	60.919444	110.53316
## 118	82.22768	57.446105	107.00926
## 119	74.87928	50.084557	99.67399
## 120	80.48918	55.660013	105.31834
## 121	93.24793	68.429487	118.06637
## 122	68.61983	43.810864	93.42879
## 123	67.52108	42.714252	92.32791
## 124	63.94556	39.087666	88.80345
## 125	66.92089	42.120098	91.72167
## 126	89.03973	64.236770	113.84269
## 127	89.32975	64.512277	114.14722
## 128	76.81046	52.020721	101.60021
## 129	92.97442	68.167232	117.78161
## 130	91.40945	66.605529	116.21337
## 131	85.00845	60.214833	109.80206
## 132	81.09936	56.309332	105.88938
## 133	81.44764	56.659640	106.23565
## 134	83.46855	58.666381	108.27071
## 135	87.14517	62.348594	111.94175
## 136	77.92244	53.099267	102.74562
## 137	74.03815	49.250633	98.82567
## 138	78.71935	53.937630	103.50107
## 139	91.35966	66.546845	116.17248
## 140	81.57631	56.795464	106.35717

## 141	64.47481	39.657619	89.29200
## 142	72.58740	47.780828	97.39398
## 143	89.64711	64.840423	114.45381
## 144	71.27260	46.476007	96.06920
## 145	70.94211	46.146918	95.73730
## 146	71.02519	46.235307	95.81507
## 147	76.42568	51.642275	101.20908
## 148	78.77578	53.988560	103.56300
## 149	79.28292	54.475367	104.09048
## 150	84.18369	59.399150	108.96823
## 151	82.83016	58.037381	107.62294
## 152	81.38184	56.584846	106.17884
## 153	51.58034	25.294461	77.86621
## 154	69.55554	44.761347	94.34973
## 155	75.73121	50.940192	100.52224
## 156	71.01286	46.209620	95.81610
## 157	90.74730	65.944116	115.55049
## 158	84.32134	59.385352	109.25732
## 159	88.30726	63.505041	113.10948
## 160	72.79448	47.999286	97.58968
## 161	99.40638	74.591161	124.22160
## 162	104.39146	79.573042	129.20988
## 163	92.36437	67.558480	117.17027
## 164	101.12231	76.299245	125.94538
## 165	94.81318	69.997073	119.62929
## 166	87.85043	63.040623	112.66023
## 167	80.06505	55.267709	104.86239
## 168	83.42704	58.612639	108.24143
## 169	74.96804	50.177476	99.75861
## 170	82.35322	57.571035	107.13541
## 171	86.11549	61.306150	110.92483
## 172	88.20970	63.414485	113.00491
## 173	80.12929	55.338103	104.92047
## 174	93.94512	69.134932	118.75532
## 175	83.40247	58.615166	108.18976
## 176	71.78859	46.993047	96.58414
## 177	76.59964	51.808031	101.39126
## 178	70.08482	45.279759	94.88988
## 179	73.28195	48.496000	98.06789
## 180	78.19407	53.411718	102.97643
## 181	91.48034	66.625145	116.33554
## 182	88.45507	63.647471	113.26266
## 183	86.55032	61.767855	111.33278
## 184	84.49026	59.690137	109.29038
## 185	85.10514	60.074397	110.13588
## 186	101.00380	76.088909	125.91869
## 187	88.21806	63.405505	113.03061
## 188	55.72995	30.814000	80.64589
## 189	68.06919	43.246795	92.89159
## 190	115.30304	90.416866	140.18921
## 191	69.98206	45.182769	94.78135
## 192	81.74419	56.951427	106.53696
## 193	78.80420	54.023274	103.58513
## 194	79.11760	54.325513	103.90969

```

## 195 81.32178 56.521648 106.12190
## 196 69.25723 44.458483 94.05599
## 197 79.07213 54.287175 103.85709
## 198 82.54158 57.732486 107.35068
## 199 78.03132 53.242709 102.81993
## 200 80.84830 56.058879 105.63772
## 201 71.84958 47.050522 96.64864
## 202 77.31938 52.526555 102.11220
## 203 70.69329 45.888950 95.49764
## 204 91.70518 66.910991 116.49937
## 205 82.80306 58.023941 107.58219
## 206 81.78088 56.992399 106.56937
## 207 77.21449 52.414727 102.01424
## 208 77.56355 52.769093 102.35800
## 209 83.24922 58.457890 108.04055
## 210 72.73594 47.928391 97.54350
## 211 102.96897 78.134000 127.80395
## 212 89.43868 64.636499 114.24087
## 213 78.64835 53.866221 103.43047
## 214 64.81584 40.012130 89.61955
## 215 67.22291 42.419120 92.02670
## 216 81.75250 56.963727 106.54127
## 217 77.54043 52.738876 102.34198
## 218 95.80783 71.003274 120.61239
## 219 78.66951 53.888720 103.45030
## 220 77.71847 52.932729 102.50421
## 221 77.51432 52.718755 102.30989
## 222 75.93201 51.127540 100.73647
## 223 81.48131 56.690687 106.27194
## 224 73.41762 48.609567 98.22567
## 225 72.02379 46.951295 97.09628
## 226 74.25784 49.473693 99.04200
## 227 82.27626 57.494074 107.05844
## 228 77.90784 53.107154 102.70852
## 229 80.80524 56.020167 105.59031
## 230 77.43289 52.596037 102.26975
## 231 80.81225 55.994144 105.63036
## 232 92.56427 67.763153 117.36538
## 233 77.10201 52.296959 101.90706
## 234 89.84769 65.042617 114.65276
## 235 79.13558 54.353680 103.91749
## 236 74.14809 49.359973 98.93621
## 237 81.80917 57.013458 106.60488
## 238 77.13036 52.347828 101.91289
## 239 85.77376 60.971112 110.57640
## 240 71.61542 46.823962 96.40688
## 241 87.68913 62.902100 112.47617
## 242 86.58148 61.785822 111.37714
## 243 84.01641 59.222533 108.81029
## 244 83.21416 58.422729 108.00558
## 245 61.10136 36.281252 85.92147
## 246 89.96287 65.172664 114.75308
## 247 82.22010 57.440406 106.99979
## 248 86.28001 61.491988 111.06804

```

##	249	73.89574	49.104961	98.68651
##	250	85.86242	61.051063	110.67378
##	251	80.46371	55.668164	105.25925
##	252	62.85112	37.989687	87.71255
##	253	96.40464	71.585917	121.22337
##	254	16.21773	-8.940658	41.37612
##	255	69.27864	44.489050	94.06824
##	256	75.32631	50.541139	100.11149
##	257	84.38305	59.585272	109.18083
##	258	86.32412	61.529542	111.11870
##	259	79.00680	54.216843	103.79676