PREDICCION - CASO NBA

Nombre: Mayra Goicochea Neyra

Se tiene la información de los jugadores de la NBA para desarrollar un modelo predictivo que se ajuste adecuadamente y permita estimar el salario de otros jugadores. Los datos que se entregaron en la muestra son los siguientes:

- Nombre del Jugador
- Salario
- Draft Number de la NBA
- Edad
- Equipo
- Goals
- Minutos de juego
- Clasificación de eficiencia del jugador
- Porcentaje de disparo
- Tasa de intento de 3 puntos
- Tasa de intento de tiro libre
- Porcentaje de rebote ofensivo
- Porcentaje de rebote defensivo
- Porcentaje de rebote total
- Porcentaje de asistencia
- Porcentaje de robo
- Porcentaje de bloqueo
- Porcentaje de rotación
- Porcentaje de uso
- Acciones ofensivas ganadoras
- Acciones de victorias defensivas
- Acciones Ganadoras
- Acciones Ganadoras por 48 minutos
- Box Plus/Minus
- Valor sobre el jugador de reemplazo

Primero se carga la muestra entregada en un dataframe y se renombran las columnas. Adicionalmente se realiza un resumen de las características de la información.

```r
setwd("C:/Users/Goicochea/Desktop/CUNEF/Cursos/Prediccion/Tarea")
library(readr)
mNba=read.csv("./nba.csv")
colnames(mNba)<-c("Player","Salary","NBA_Country","NBA_DraftNumber","Age","Tm
","G","MP","PER","TS","TPAr","FTr","ORB","DRB","TRB","AST","STL","BLK","TOV",
"USG","OWS","DWS","WS","WS48","OBPM","DBPM","BPM","VORP")
summary(mNba)
```

```
##                Player        Salary            NBA_Country  NBA_DraftNumber
##  Kay Felder     :  3   Min.   :   46080   USA      :374   Min.   : 1.00
##  Aaron Brooks   :  1   1st Qu.: 1471382   Canada   : 12   1st Qu.:11.00
##  Aaron Gordon   :  1   Median : 3202217   France   :  9   Median :25.00
##  Aaron Harrison :  1   Mean   : 6636507   Australia:  8   Mean   :29.45
##  Abdel Nader    :  1   3rd Qu.:10000000   Spain    :  7   3rd Qu.:47.00
##  Al-Farouq Aminu:  1   Max.   :34682550   Croatia  :  6   Max.   :62.00
##  (Other)        :477                      (Other)  : 69
##       Age             Tm            G               MP
##  Min.   :19.00   TOT    : 55   Min.   : 1.00   Min.   :   1
##  1st Qu.:23.00   DAL    : 18   1st Qu.:29.00   1st Qu.: 381
##  Median :26.00   MEM    : 17   Median :59.00   Median :1134
##  Mean   :26.26   UTA    : 17   Mean   :50.17   Mean   :1154
##  3rd Qu.:29.00   ATL    : 16   3rd Qu.:71.00   3rd Qu.:1819
##  Max.   :41.00   GSW    : 16   Max.   :79.00   Max.   :2898
##                  (Other):346
##       PER               TS              TPAr              FTr
##  Min.   :-41.10   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:  9.80   1st Qu.:0.5055   1st Qu.:0.1670   1st Qu.:0.1550
##  Median : 13.20   Median :0.5450   Median :0.3460   Median :0.2310
##  Mean   : 13.26   Mean   :0.5354   Mean   :0.3374   Mean   :0.2634
##  3rd Qu.: 16.50   3rd Qu.:0.5825   3rd Qu.:0.4810   3rd Qu.:0.3195
##  Max.   :134.10   Max.   :1.5000   Max.   :1.0000   Max.   :5.3330
##                   NA's   :2        NA's   :2        NA's   :2
##       ORB              DRB             TRB               AST
##  Min.   : 0.000   Min.   : 0.00   Min.   : 0.000   Min.   : 0.00
##  1st Qu.: 1.800   1st Qu.:10.20   1st Qu.: 6.200   1st Qu.: 6.90
##  Median : 3.200   Median :14.00   Median : 8.700   Median : 9.90
##  Mean   : 4.874   Mean   :14.95   Mean   : 9.908   Mean   :12.95
##  3rd Qu.: 7.000   3rd Qu.:18.80   3rd Qu.:13.300   3rd Qu.:17.60
##  Max.   :35.900   Max.   :37.60   Max.   :26.500   Max.   :49.40
##
##       STL              BLK              TOV             USG
##  Min.   : 0.000   Min.   : 0.000   Min.   : 0.00   Min.   : 0.0
##  1st Qu.: 1.000   1st Qu.: 0.600   1st Qu.: 9.90   1st Qu.:15.0
##  Median : 1.500   Median : 1.200   Median :12.50   Median :17.9
##  Mean   : 1.529   Mean   : 1.713   Mean   :13.14   Mean   :18.9
##  3rd Qu.: 1.900   3rd Qu.: 2.200   3rd Qu.:15.75   3rd Qu.:22.2
##  Max.   :12.500   Max.   :13.400   Max.   :66.70   Max.   :45.1
##                                    NA's   :2
##       OWS              DWS              WS               WS48
##  Min.   :-2.300   Min.   :0.000   Min.   :-1.200   Min.   :-1.06300
##  1st Qu.: 0.000   1st Qu.:0.300   1st Qu.: 0.300   1st Qu.: 0.04000
##  Median : 0.800   Median :1.000   Median : 1.800   Median : 0.08300
##  Mean   : 1.275   Mean   :1.176   Mean   : 2.455   Mean   : 0.07996
##  3rd Qu.: 2.000   3rd Qu.:1.800   3rd Qu.: 3.600   3rd Qu.: 0.12300
##  Max.   :11.400   Max.   :5.600   Max.   :15.000   Max.   : 2.71300
##
##       OBPM              DBPM               BPM              VORP
##  Min.   :-36.500   Min.   :-14.3000   Min.   :-49.20   Min.   :-1.3000
```

```
##  1st Qu.: -2.700   1st Qu.: -1.7000   1st Qu.: -3.60   1st Qu.:-0.1000
##  Median : -1.100   Median : -0.4000   Median : -1.30   Median : 0.1000
##  Mean   : -1.271   Mean   : -0.4895   Mean   : -1.76   Mean   : 0.5988
##  3rd Qu.:  0.400   3rd Qu.:  1.0000   3rd Qu.:  0.50   3rd Qu.: 0.9000
##  Max.   : 68.700   Max.   :  6.8000   Max.   : 54.40   Max.   : 8.6000
##
```

```r
cat("NA values:",sum(is.na(mNba)),"\n")
```

```
## NA values: 8
```

Se observa que se tiene valores ausentes en 4 variables (TS, TPAr, FTr, TOV). Es así que se procede a ajustar estos datos con el método de imputación por hotdeck.

```r
library(rminer)
library(ggplot2)
library(kknn)

mNba<-imputation("hotdeck",mNba,"TOV")
mNba<-imputation("hotdeck",mNba,"FTr")
mNba<-imputation("hotdeck",mNba,"TPAr")
mNba<-imputation("hotdeck",mNba,"TS")
cat("NA values:",sum(is.na(mNba)),"\n")
```

```
## NA values: 0
```

Ahora si se puede proceder a proponer un modelo inicial para revisar el comportamiento de las variables explicativas. Se quita la variable Nombre del Jugador al no ser representativa para la predicción del salario del jugador.

```r
set.seed(5)
reg0<-lm(Salary~ .-Player, data=mNba)
summary(reg0)
```

```
##
## Call:
## lm(formula = Salary ~ . - Player, data = mNba)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -13937029  -2698206   -109317   2245370  21052085
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -11118373    7045737  -1.578 0.115376
## NBA_CountryAustralia           5503787    4443729   1.239 0.216262
## NBA_CountryAustria             1768353    6881150   0.257 0.797327
## NBA_CountryBahamas             1979037    6775885   0.292 0.770389
## NBA_CountryBosnia              9043800    6734228   1.343 0.180072
## NBA_CountryBosnia & Herz...   -4118619    6781790  -0.607 0.544004
## NBA_CountryBrazil              5334108    4710812   1.132 0.258205
```

```
## NBA_CountryCameroon         -1933569    5094380  -0.380 0.704488
## NBA_CountryCanada            4732952    4348266   1.088 0.277066
## NBA_CountryChina             5788010    6880690   0.841 0.400757
## NBA_CountryCroatia           4243887    4601185   0.922 0.356922
## NBA_CountryCzech Republic     489464    6739524   0.073 0.942141
## NBA_CountryDemocratic Re...  12803137    5504149   2.326 0.020530 *
## NBA_CountryDemocratic Re_    11298539    6754240   1.673 0.095173 .
## NBA_CountryDominican Rep...   9980364    5667310   1.761 0.079021 .
## NBA_CountryEgypt             8348979    6704412   1.245 0.213776
## NBA_CountryFinland           2838650    6761942   0.420 0.674866
## NBA_CountryFrance           10724624    4346296   2.468 0.014037 *
## NBA_CountryGeorgia           -628547    6688654  -0.094 0.925180
## NBA_CountryGermany           4537502    4665772   0.973 0.331405
## NBA_CountryGreece            4936396    5679698   0.869 0.385315
## NBA_CountryHaiti             4182049    6768228   0.618 0.537009
## NBA_CountryIsrael           -3159854    6745788  -0.468 0.639749
## NBA_CountryItaly             7942978    5479074   1.450 0.147954
## NBA_CountryLatvia            2331715    5456657   0.427 0.669387
## NBA_CountryLithuania         2770665    5231523   0.530 0.596686
## NBA_CountryMali              5319703    6742346   0.789 0.430596
## NBA_CountryMontenegro        5861284    5461354   1.073 0.283837
## NBA_CountryNew Zealand      10458577    6871350   1.522 0.128812
## NBA_CountryPoland            4331339    6685954   0.648 0.517481
## NBA_CountryPuerto Rico       2032492    5632614   0.361 0.718413
## NBA_CountryRussia           17126880    6721623   2.548 0.011219 *
## NBA_CountrySenegal          15023253    6678138   2.250 0.025035 *
## NBA_CountrySerbia            2565033    4670837   0.549 0.583213
## NBA_CountrySlovenia          8877503    6694404   1.326 0.185586
## NBA_CountrySouth Sudan       8254505    6758798   1.221 0.222717
## NBA_CountrySpain             5882300    4426354   1.329 0.184655
## NBA_CountrySweden            5125055    6623076   0.774 0.439511
## NBA_CountrySwitzerland      -2510959    5549329  -0.452 0.651177
## NBA_CountryTunisia           2937976    6942567   0.423 0.672397
## NBA_CountryTurkey            7065243    4630307   1.526 0.127859
## NBA_CountryUkraine           4851390    5621518   0.863 0.388670
## NBA_CountryUnited Kingdo...  10592050    5628525   1.882 0.060606 .
## NBA_CountryUSA               4788161    3992509   1.199 0.231150
## NBA_DraftNumber               -53128      14245  -3.730 0.000221 ***
## Age                           521276      63127   8.258 2.38e-15 ***
## TmBOS                       -2606730    2318012  -1.125 0.261475
## TmBRK                       -1268772    2063663  -0.615 0.539038
## TmCHI                       -1246480    2032977  -0.613 0.540150
## TmCHO                        1117152    1998311   0.559 0.576452
## TmCLE                         992786    2235172   0.444 0.657171
## TmDAL                       -1232036    2026773  -0.608 0.543623
## TmDEN                        -584764    2020372  -0.289 0.772404
## TmDET                        -272871    2068016  -0.132 0.895094
## TmGSW                        -873496    2287135  -0.382 0.702732
## TmHOU                       -2629122    2254287  -1.166 0.244221
## TmIND                       -1670309    1928034  -0.866 0.386847
```

```
## TmLAC                         -811984    2076243  -0.391 0.695951    
## TmLAL                        -2051129    1960717  -1.046 0.296162    
## TmMEM                          603789    1947587   0.310 0.756714    
## TmMIA                        -2289325    2027555  -1.129 0.259552    
## TmMIL                        -1405189    2112522  -0.665 0.506337    
## TmMIN                        -1113465    2241854  -0.497 0.619703    
## TmNOP                        -2923811    2076601  -1.408 0.159940    
## TmNYK                         -667518    2019870  -0.330 0.741219    
## TmOKC                         1365863    2142694   0.637 0.524208    
## TmORL                         -653276    1955431  -0.334 0.738498    
## TmPHI                        -3053356    2113545  -1.445 0.149362    
## TmPHO                         -522219    2023570  -0.258 0.796491    
## TmPOR                          460921    2050509   0.225 0.822266    
## TmSAC                          320103    2387836   0.134 0.893428    
## TmSAS                        -1176606    2107712  -0.558 0.577004    
## TmTOR                          309230    2247408   0.138 0.890633    
## TmTOT                        -1092147    1545969  -0.706 0.480334    
## TmUTA                        -3141828    2189324  -1.435 0.152075    
## TmWAS                          650039    2100331   0.309 0.757113    
## G                             -183665      27717  -6.626 1.16e-10 ***
## MP                               6217       1188   5.236 2.70e-07 ***
## PER                           -367094     316269  -1.161 0.246479    
## TS                            2639979    5681082   0.465 0.642410    
## TPAr                          -943127    3275156  -0.288 0.773528    
## FTr                           -157482     949274  -0.166 0.868324    
## ORB                          -2453115    1622682  -1.512 0.131410    
## DRB                          -2341000    1590038  -1.472 0.141755    
## TRB                           4914763    3214898   1.529 0.127144    
## AST                             28391      52304   0.543 0.587569    
## STL                           -268636     527255  -0.509 0.610694    
## BLK                            -82277     427477  -0.192 0.847474    
## TOV                            -80496      60322  -1.334 0.182843    
## USG                            239451     132996   1.800 0.072570 .  
## OWS                            589633    4943953   0.119 0.905129    
## DWS                            463804    4950202   0.094 0.925401    
## WS                              30652    4960038   0.006 0.995072    
## WS48                         14460422   15621188   0.926 0.355182    
## OBPM                          5177096    5071120   1.021 0.307941    
## DBPM                          5465060    4976386   1.098 0.272801    
## BPM                          -5106311    4987602  -1.024 0.306569    
## VORP                           385770     671561   0.574 0.566005    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 5102000 on 387 degrees of freedom
## Multiple R-squared:  0.6191, Adjusted R-squared:  0.5236 
## F-statistic: 6.485 on 97 and 387 DF,  p-value: < 2.2e-16
```

De igual forma se observa en el modelo inicial que el Pais NBA crea muchas variables dicotómicas no representativas para el modelo predictivo.

```r
reg01<-lm(Salary~ .-Player-NBA_Country, data=mNba)
summary(reg01)

##
## Call:
## lm(formula = Salary ~ . - Player - NBA_Country, data = mNba)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -14484389  -2854452   -462414   2278016  20268616
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -3182044    5647408  -0.563    0.573
## NBA_DraftNumber    -58801      13729  -4.283 2.28e-05 ***
## Age                523086      60103   8.703  < 2e-16 ***
## TmBOS             -239702    2167241  -0.111    0.912
## TmBRK              146052    2007044   0.073    0.942
## TmCHI            -1107402    1995228  -0.555    0.579
## TmCHO             1705799    1984095   0.860    0.390
## TmCLE             1364470    2186632   0.624    0.533
## TmDAL            -1040207    1907324  -0.545    0.586
## TmDEN             -868811    1982807  -0.438    0.661
## TmDET              555796    2033667   0.273    0.785
## TmGSW            -1177308    2243811  -0.525    0.600
## TmHOU            -2631863    2145811  -1.227    0.221
## TmIND            -1368884    1901746  -0.720    0.472
## TmLAC             -471163    2038053  -0.231    0.817
## TmLAL            -1309049    1929693  -0.678    0.498
## TmMEM              752755    1915301   0.393    0.694
## TmMIA            -1040623    1981150  -0.525    0.600
## TmMIL             -618190    1980194  -0.312    0.755
## TmMIN             -668629    2157623  -0.310    0.757
## TmNOP            -2260707    2027801  -1.115    0.266
## TmNYK             -319902    1976024  -0.162    0.871
## TmOKC             2202998    2070762   1.064    0.288
## TmORL              534233    1900132   0.281    0.779
## TmPHI            -1860281    2068608  -0.899    0.369
## TmPHO             -749465    1989777  -0.377    0.707
## TmPOR              508763    2015899   0.252    0.801
## TmSAC             -366029    2211337  -0.166    0.869
## TmSAS              186188    2051736   0.091    0.928
## TmTOR             1079678    2084792   0.518    0.605
## TmTOT             -410766    1519778  -0.270    0.787
## TmUTA            -1829949    2078919  -0.880    0.379
## TmWAS             1143758    1995871   0.573    0.567
## G                 -177714      26930  -6.599 1.22e-10 ***
## MP                   6444       1168   5.518 5.94e-08 ***
## PER               -437090     303987  -1.438    0.151
## TS                -505566    5403322  -0.094    0.925
```

```
## TPAr              -3009344    3157457  -0.953    0.341
## FTr                -278599     927382  -0.300    0.764
## ORB               -2049701    1504207  -1.363    0.174
## DRB               -1883011    1477321  -1.275    0.203
## TRB                4060271    2981367   1.362    0.174
## AST                   1666      49623   0.034    0.973
## STL                -144332     515871  -0.280    0.780
## BLK                 101903     410170   0.248    0.804
## TOV                 -13410      55317  -0.242    0.809
## USG                 199437     123813   1.611    0.108
## OWS                 235107    4734949   0.050    0.960
## DWS                -317630    4752555  -0.067    0.947
## WS                  327266    4742039   0.069    0.945
## WS48               9417015   14677428   0.642    0.521
## OBPM               3247012    4953991   0.655    0.513
## DBPM               2991924    4843036   0.618    0.537
## BPM               -2809800    4867403  -0.577    0.564
## VORP                544481     656051   0.830    0.407
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5162000 on 430 degrees of freedom
## Multiple R-squared:  0.5668, Adjusted R-squared:  0.5124
## F-statistic: 10.42 on 54 and 430 DF,  p-value: < 2.2e-16
```

Ahora con el modelo inicial se ejecutarán algunas pruebas para revisar el comportamiento de las variables.

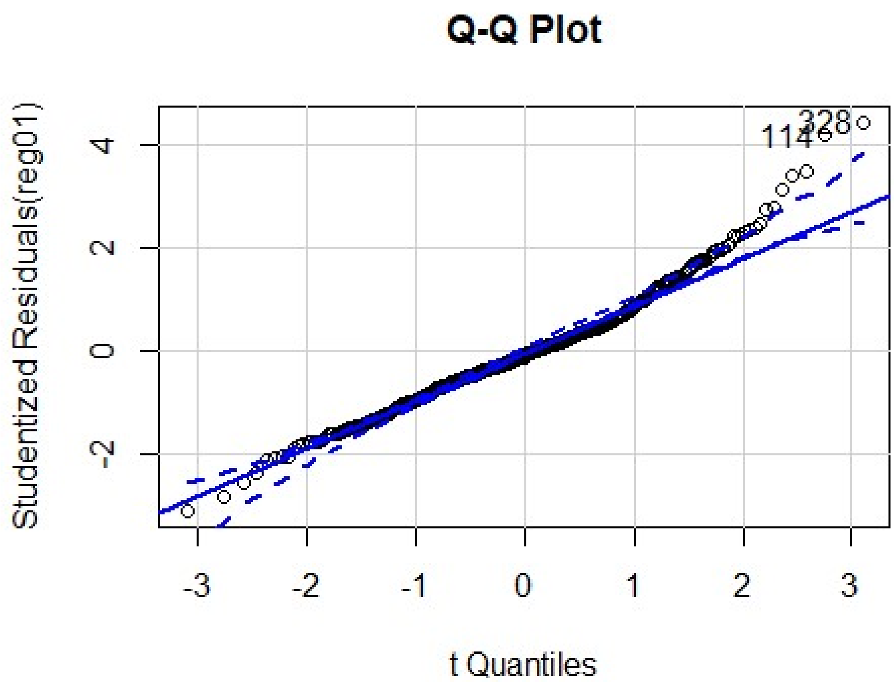
ANALISIS DE LAS VARIABLES EXPLICATIVAS

1. Pruebas de Normalidad

a. QQPLOT

Contrastamos gráficamente la distribución de densidad de la muestra con respecto a la distribución normal. La línea continua se encuentra dentro del intervalo de confianza en el centro, pero en los extremos se aleja. No se puede afirmar que el comportamiento de las variables siga una distribución normal.

```r
library(carData)
library(car)

qqPlot(reg01, labels=row.names(mNbaF), id.method="identify", simulate=TRUE, m
ain="Q-Q Plot")
```

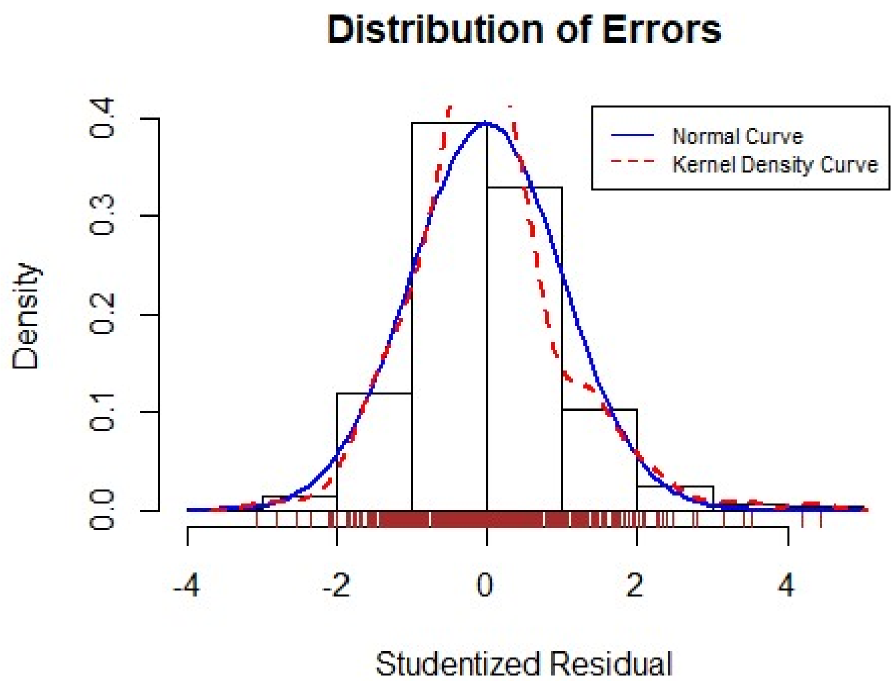



```
## [1] 114 328
```

b. Histograma + densidad + normal + rug

Esta técnica nos permite calcular el error residual del análisis de normalidad de la muestra.

```
residplot <- function(fit, nbreaks=10) {
  z <- rstudent(fit)
  hist(z, breaks=nbreaks, freq=FALSE,
       xlab="Studentized Residual",
       main="Distribution of Errors")
  rug(jitter(z), col="brown")
  curve(dnorm(x, mean=mean(z), sd=sd(z)),
        add=TRUE, col="blue", lwd=2)
  lines(density(z)$x, density(z)$y,
        col="red", lwd=2, lty=2)
  legend("topright",
         legend = c( "Normal Curve", "Kernel Density Curve"),
         lty=1:2, col=c("blue","red"), cex=.7)
}
residplot(reg01)
```



El histograma nos grafica de forma discreta como los datos se agrupan en los rangos de valores. La línea descontinua roja representa la muestra y la línea azul, la distribución normal. Como se observa, ambas curvas difieren tanto en mediana como en el rango de 0 a 2.

c. Jarque-Bera

```
vResid=resid(reg01)
library(fBasics)

## Loading required package: timeDate

## Loading required package: timeSeries

##
## Attaching package: 'fBasics'

## The following object is masked from 'package:car':
##
##     densityPlot

library(akima)
jbTest(vResid)

## Warning in interpp.old(x, y, z, xo, yo, ncp = 0, extrap = FALSE, duplicate
## = duplicate, : interpp.old() is deprecated, future versions will only
## provide interpp()

## Warning in interpp.old(x, y, z, xo, yo, ncp = 0, extrap = FALSE, duplicate
```

```
## = duplicate, : interpp.old() is deprecated, future versions will only
## provide interpp()

##
## Title:
##  Jarque - Bera Normality Test
##
## Test Results:
##   PARAMETER:
##     Sample Size: 485
##   STATISTIC:
##     LM: 78.787
##     ALM: 81.42
##   P VALUE:
##     ALM p-value: < 2.2e-16
##     Asymptotic: < 2.2e-16
##
## Description:
##  Thu Oct 10 01:09:38 2019 by user: Goicochea
```

Mediante la prueba de Jarque-Vera, se puede indicar que siendo p-value menor que 0.05 no se puede aceptar la hipótesis nula (Los datos se comportan en forma de una distribución normal).

d. Shapiro - Wilk

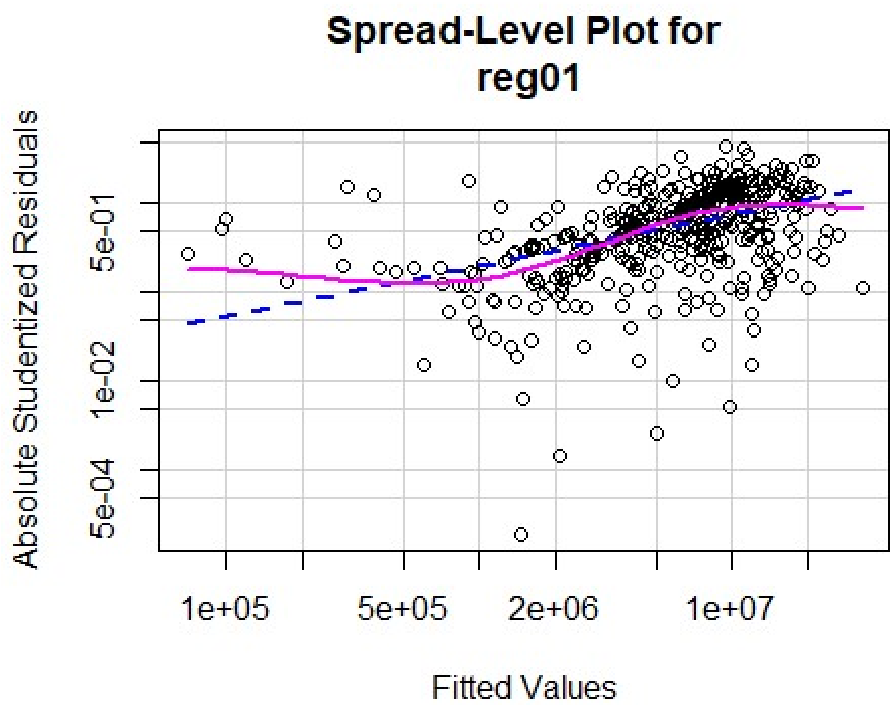
```
shapiro.test(vResid)

##
##  Shapiro-Wilk normality test
##
## data:  vResid
## W = 0.97319, p-value = 9.25e-08
```

De igual manera, la prueba de Shapiro nos indica que no se puede aceptar la hipótesis de que los datos se comportan en forma de una distribución normal.

2. Pruebas de Homocedasticidad

Se realiza el test de Breusch–Pagan para determinar si la varianza entre las variables es contaste.

```
ncvTest(reg01)

## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 81.80624, Df = 1, p = < 2.22e-16

spreadLevelPlot(reg01)

## Warning in spreadLevelPlot.lm(reg01):
## 45 negative fitted values removed
```




```
##
## Suggested power transformation:  0.430272
```

Siendo el valor Chisquare mayor a 0.05, se acepta que las varianzas son constantes.

3. Validación Global

```
library(gvlma)
gvmodel <- gvlma(reg01)
summary(gvmodel)
```

```
##
## Call:
## lm(formula = Salary ~ . - Player - NBA_Country, data = mNba)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -14484389  -2854452   -462414   2278016  20268616
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -3182044    5647408  -0.563    0.573
## NBA_DraftNumber     -58801      13729  -4.283 2.28e-05 ***
## Age                 523086      60103   8.703  < 2e-16 ***
## TmBOS              -239702    2167241  -0.111    0.912
## TmBRK               146052    2007044   0.073    0.942
## TmCHI             -1107402    1995228  -0.555    0.579
```

```
## TmCHO         1705799    1984095   0.860    0.390
## TmCLE         1364470    2186632   0.624    0.533
## TmDAL        -1040207    1907324  -0.545    0.586
## TmDEN         -868811    1982807  -0.438    0.661
## TmDET          555796    2033667   0.273    0.785
## TmGSW        -1177308    2243811  -0.525    0.600
## TmHOU        -2631863    2145811  -1.227    0.221
## TmIND        -1368884    1901746  -0.720    0.472
## TmLAC         -471163    2038053  -0.231    0.817
## TmLAL        -1309049    1929693  -0.678    0.498
## TmMEM          752755    1915301   0.393    0.694
## TmMIA        -1040623    1981150  -0.525    0.600
## TmMIL         -618190    1980194  -0.312    0.755
## TmMIN         -668629    2157623  -0.310    0.757
## TmNOP        -2260707    2027801  -1.115    0.266
## TmNYK         -319902    1976024  -0.162    0.871
## TmOKC         2202998    2070762   1.064    0.288
## TmORL          534233    1900132   0.281    0.779
## TmPHI        -1860281    2068608  -0.899    0.369
## TmPHO         -749465    1989777  -0.377    0.707
## TmPOR          508763    2015899   0.252    0.801
## TmSAC         -366029    2211337  -0.166    0.869
## TmSAS          186188    2051736   0.091    0.928
## TmTOR         1079678    2084792   0.518    0.605
## TmTOT         -410766    1519778  -0.270    0.787
## TmUTA        -1829949    2078919  -0.880    0.379
## TmWAS         1143758    1995871   0.573    0.567
## G             -177714      26930  -6.599 1.22e-10 ***
## MP               6444       1168   5.518 5.94e-08 ***
## PER           -437090     303987  -1.438    0.151
## TS            -505566    5403322  -0.094    0.925
## TPAr         -3009344    3157457  -0.953    0.341
## FTr           -278599     927382  -0.300    0.764
## ORB          -2049701    1504207  -1.363    0.174
## DRB          -1883011    1477321  -1.275    0.203
## TRB           4060271    2981367   1.362    0.174
## AST              1666      49623   0.034    0.973
## STL           -144332     515871  -0.280    0.780
## BLK            101903     410170   0.248    0.804
## TOV            -13410      55317  -0.242    0.809
## USG            199437     123813   1.611    0.108
## OWS            235107    4734949   0.050    0.960
## DWS           -317630    4752555  -0.067    0.947
## WS             327266    4742039   0.069    0.945
## WS48          9417015   14677428   0.642    0.521
## OBPM          3247012    4953991   0.655    0.513
## DBPM          2991924    4843036   0.618    0.537
## BPM          -2809800    4867403  -0.577    0.564
## VORP           544481     656051   0.830    0.407
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5162000 on 430 degrees of freedom
## Multiple R-squared:  0.5668, Adjusted R-squared:  0.5124
## F-statistic: 10.42 on 54 and 430 DF,  p-value: < 2.2e-16
##
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance =  0.05
##
## Call:
##  gvlma(x = reg01)
##
##                        Value   p-value                   Decision
## Global Stat        116.7865 0.000e+00 Assumptions NOT satisfied!
## Skewness            32.6716 1.091e-08 Assumptions NOT satisfied!
## Kurtosis            46.1157 1.115e-11 Assumptions NOT satisfied!
## Link Function       37.8539 7.625e-10 Assumptions NOT satisfied!
## Heteroscedasticity   0.1453 7.031e-01    Assumptions acceptable.
```

No se puede generar una validación general debido a que existen variables correlacionadas. Con el supuesto de que las variables relacionadas son Tm (Equipo), formulamos un nuevo modelo:

```
reg01<-lm(Salary~ .-Player-NBA_Country-Tm, data=mNba)
library(gvlma)
gvmodel <- gvlma(reg01)
summary(gvmodel)
```

```
##
## Call:
## lm(formula = Salary ~ . - Player - NBA_Country - Tm, data = mNba)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -15335041  -2876838   -393827   2084667  21561492
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -2818136    4496008  -0.627   0.5311
## NBA_DraftNumber     -60715      12702  -4.780 2.36e-06 ***
## Age                 516828      56274   9.184  < 2e-16 ***
## G                  -154068      24922  -6.182 1.40e-09 ***
## MP                    5694       1071   5.316 1.65e-07 ***
## PER                -370383     275829  -1.343   0.1800
## TS                -1508637    4700490  -0.321   0.7484
## TPAr              -3369918    2335268  -1.443   0.1497
## FTr                -178939     882668  -0.203   0.8394
## ORB               -1046505     900662  -1.162   0.2459
```

```
## DRB              -847463     891446  -0.951   0.3423
## TRB              1999401    1785035   1.120   0.2633
## AST               -14675      44899  -0.327   0.7439
## STL              -165292     411913  -0.401   0.6884
## BLK               118256     317038   0.373   0.7093
## TOV                -1117      49640  -0.022   0.9821
## USG               176582     101254   1.744   0.0818 .
## OWS             -1258098    4484355  -0.281   0.7792
## DWS             -1743380    4509571  -0.387   0.6992
## WS               1806640    4489961   0.402   0.6876
## WS48             2832104   11353750   0.249   0.8031
## OBPM             1599356    4694011   0.341   0.7335
## DBPM             1172219    4615294   0.254   0.7996
## BPM             -1037091    4634823  -0.224   0.8230
## VORP              636862     632152   1.007   0.3142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5093000 on 460 degrees of freedom
## Multiple R-squared:  0.5488, Adjusted R-squared:  0.5253
## F-statistic: 23.32 on 24 and 460 DF,  p-value: < 2.2e-16
##
##
## ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
## USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
## Level of Significance =  0.05
##
## Call:
##  gvlma(x = reg01)
##
##                        Value   p-value                   Decision
## Global Stat         117.9189 0.000e+00 Assumptions NOT satisfied!
## Skewness             32.5233 1.178e-08 Assumptions NOT satisfied!
## Kurtosis             52.5263 4.245e-13 Assumptions NOT satisfied!
## Link Function        32.2142 1.381e-08 Assumptions NOT satisfied!
## Heteroscedasticity    0.6551 4.183e-01    Assumptions acceptable.
```

Efectivamente esas eran las variables correlacionadas. Vemos que este nuevo modelo pasa la prueba de homocedastacidad.

4. Multicolinealidad

Se realiza las pruebas en los siguientes casos: - Todas las variables excepto el nombre del jugador. - Todas las variables excepto el nombre del jugador y Pais.

```
reg0<-lm(Salary~ .-Player, data=mNba)
vif(reg0)
```

```
##                         GVIF Df GVIF^(1/(2*Df))
## NBA_Country       277.925201 43        1.067623
## NBA_DraftNumber     1.683751  1        1.297594
```

```
## Age                  1.352283  1         1.162877
## Tm                1443.249985 30         1.128901
## G                    8.837655  1         2.972819
## MP                  17.257801  1         4.154251
## PER                143.007203  1        11.958562
## TS                   7.737028  1         2.781551
## TPAr                10.227368  1         3.198026
## FTr                  1.449060  1         1.203769
## ORB               1028.126234  1        32.064408
## DRB               2203.934346  1        46.946079
## TRB               4720.512367  1        68.705985
## AST                  4.223251  1         2.055055
## STL                  5.061083  1         2.249685
## BLK                  9.632054  1         3.103555
## TOV                  2.638265  1         1.624274
## USG                 11.604962  1         3.406606
## OWS               1608.596014  1        40.107306
## DWS                487.629346  1        22.082331
## WS                3269.651436  1        57.180866
## WS48               120.524534  1        10.978367
## OBPM             12078.587180  1       109.902626
## DBPM              2628.457905  1        51.268488
## BPM              14823.628103  1       121.752323
## VORP                13.010092  1         3.606950
```

```
sqrt(vif(reg0)) > 2
```

```
##                   GVIF    Df GVIF^(1/(2*Df))
## NBA_Country       TRUE  TRUE           FALSE
## NBA_DraftNumber  FALSE FALSE           FALSE
## Age              FALSE FALSE           FALSE
## Tm                TRUE  TRUE           FALSE
## G                 TRUE FALSE           FALSE
## MP                TRUE FALSE            TRUE
## PER               TRUE FALSE            TRUE
## TS                TRUE FALSE           FALSE
## TPAr              TRUE FALSE           FALSE
## FTr              FALSE FALSE           FALSE
## ORB               TRUE FALSE            TRUE
## DRB               TRUE FALSE            TRUE
## TRB               TRUE FALSE            TRUE
## AST               TRUE FALSE           FALSE
## STL               TRUE FALSE           FALSE
## BLK               TRUE FALSE           FALSE
## TOV              FALSE FALSE           FALSE
## USG               TRUE FALSE           FALSE
## OWS               TRUE FALSE            TRUE
## DWS               TRUE FALSE            TRUE
## WS                TRUE FALSE            TRUE
## WS48              TRUE FALSE            TRUE
```

```
## OBPM              TRUE  FALSE           TRUE
## DBPM              TRUE  FALSE           TRUE
## BPM               TRUE  FALSE           TRUE
## VORP              TRUE  FALSE          FALSE
```

```
reg0<-lm(Salary~ .-Player-NBA_Country, data=mNba)
vif(reg0)
```

```
##                          GVIF Df GVIF^(1/(2*Df))
## NBA_DraftNumber      1.527738  1        1.236017
## Age                  1.197508  1        1.094307
## Tm                  58.570344 30        1.070191
## G                    8.150091  1        2.854836
## MP                  16.306705  1        4.038156
## PER                129.061946  1       11.360543
## TS                   6.837203  1        2.614805
## TPAr                 9.285807  1        3.047262
## FTr                  1.351030  1        1.162338
## ORB                863.056504  1       29.377823
## DRB               1858.565759  1       43.111086
## TRB               3965.796611  1       62.974571
## AST                  3.713531  1        1.927052
## STL                  4.732913  1        2.175526
## BLK                  8.662929  1        2.943285
## TOV                  2.167379  1        1.472202
## USG                  9.825270  1        3.134529
## OWS               1441.364418  1       37.965305
## DWS                439.079279  1       20.954219
## WS                2919.486581  1       54.032274
## WS48               103.942224  1       10.195206
## OBPM             11260.649278  1      106.116206
## DBPM              2431.940999  1       49.314714
## BPM              13791.459094  1      117.437043
## VORP                12.129139  1        3.482691
```

```
sqrt(vif(reg0)) > 2
```

```
##                  GVIF    Df GVIF^(1/(2*Df))
## NBA_DraftNumber FALSE FALSE           FALSE
## Age             FALSE FALSE           FALSE
## Tm               TRUE  TRUE           FALSE
## G                TRUE FALSE           FALSE
## MP               TRUE FALSE            TRUE
## PER              TRUE FALSE            TRUE
## TS               TRUE FALSE           FALSE
## TPAr             TRUE FALSE           FALSE
## FTr             FALSE FALSE           FALSE
## ORB              TRUE FALSE            TRUE
## DRB              TRUE FALSE            TRUE
## TRB              TRUE FALSE            TRUE
## AST             FALSE FALSE           FALSE
```

```
## STL                 TRUE  FALSE            FALSE
## BLK                 TRUE  FALSE            FALSE
## TOV                FALSE  FALSE            FALSE
## USG                 TRUE  FALSE            FALSE
## OWS                 TRUE  FALSE             TRUE
## DWS                 TRUE  FALSE             TRUE
## WS                  TRUE  FALSE             TRUE
## WS48                TRUE  FALSE             TRUE
## OBPM                TRUE  FALSE             TRUE
## DBPM                TRUE  FALSE             TRUE
## BPM                 TRUE  FALSE             TRUE
## VORP                TRUE  FALSE            FALSE
```

No se puede calcular el error debido a que las variables están correlacionadas. Estos casos responden a multicolinealidad.

La siguiente prueba, quitando las variables el nombre del jugador, Pais y Equipo.

```
reg01<-lm(Salary~ .-Player-NBA_Country-Tm, data=mNba)
vif(reg01)

## NBA_DraftNumber             Age               G              MP
##        1.343368        1.078361        7.169812       14.087502
##             PER              TS            TPAr             FTr
##      109.154233        5.315156        5.217843        1.257232
##             ORB             DRB             TRB             AST
##      317.849153      695.170363     1460.380308        3.122943
##             STL             BLK             TOV             USG
##        3.099772        5.316598        1.792825        6.750052
##             OWS             DWS              WS            WS48
##     1328.055358      406.099148     2688.650674       63.891606
##            OBPM            DBPM             BPM            VORP
##    10385.187478     2268.764386    12845.617580       11.568339

sqrt(vif(reg01)) > 2

## NBA_DraftNumber             Age               G              MP
##           FALSE           FALSE            TRUE            TRUE
##             PER              TS            TPAr             FTr
##            TRUE            TRUE            TRUE           FALSE
##             ORB             DRB             TRB             AST
##            TRUE            TRUE            TRUE           FALSE
##             STL             BLK             TOV             USG
##           FALSE            TRUE           FALSE            TRUE
##             OWS             DWS              WS            WS48
##            TRUE            TRUE            TRUE            TRUE
##            OBPM            DBPM             BPM            VORP
##            TRUE            TRUE            TRUE            TRUE
```

Observaciones Anómalas

1. *Atípicos*

Para identificar los valores atípicos de la muestra se realiza una prueba Bonferroni p-values.

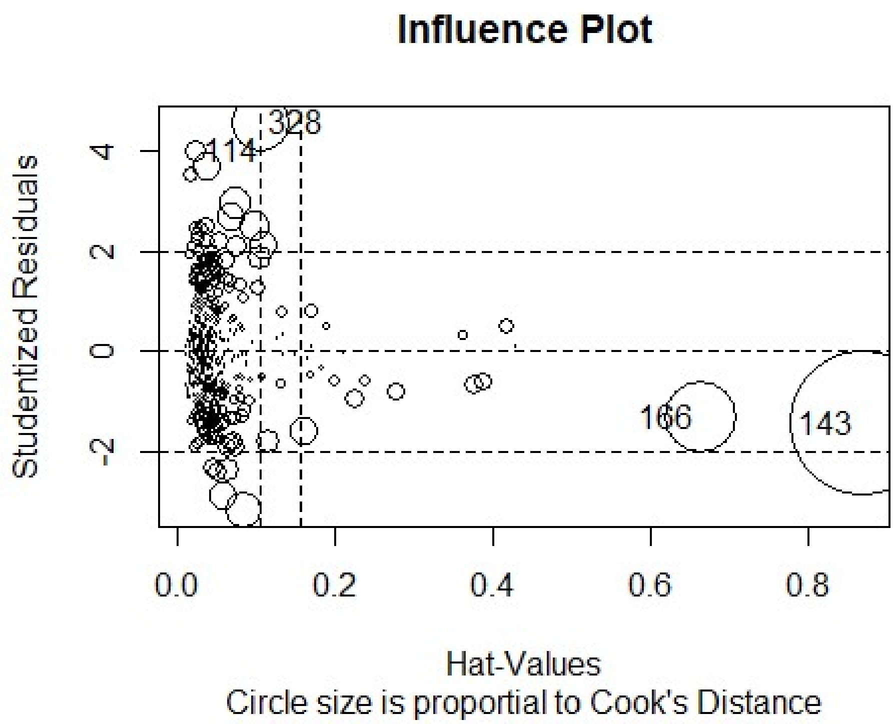
```
outlierTest(reg01)

##     rstudent unadjusted p-value Bonferroni p
## 328 4.565045         6.4214e-06    0.0031144
## 114 4.010267         7.0785e-05    0.0343310
```

Según el test, se identifica las filas 114 y 328 como casos atípicos.

2. *Influyentes*

Mediante el grafico de influencia se puede identificar los valores influyentes:

```
influencePlot(reg01, id.method="identify", main="Influence Plot",
              sub="Circle size is proportial to Cook's Distance" )

## Warning in plot.window(...): "id.method" is not a graphical parameter

## Warning in plot.xy(xy, type, ...): "id.method" is not a graphical paramete
r

## Warning in axis(side = side, at = at, labels = labels, ...): "id.method" i
s
## not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "id.method" i
s
## not a graphical parameter

## Warning in box(...): "id.method" is not a graphical parameter

## Warning in title(...): "id.method" is not a graphical parameter

## Warning in plot.xy(xy.coords(x, y), type = type, ...): "id.method" is not
a
## graphical parameter
```




```
##        StudRes        Hat      CookD
## 114  4.010267 0.02084044 0.01325712
## 143 -1.455868 0.86934235 0.56273522
## 166 -1.321064 0.66534057 0.13856248
## 328  4.565045 0.10299671 0.09175742
```

Los valores influyentes son las filas 114,143,166,328. Se procede a eliminar estos casos, porque si no historiarían el modelo predictivo.

```
mNba<-mNba[-c(114,143,166,328),]
```

SELECCION DE VARIABLES

A continuación, se ejecutan las técnicas de selección Stepwise y Cross Validation. Cada algoritmo generará el mejor modelo y mediante el menor valor de AIC alcanzado se elegirá el modelo predictivo resultante. Primero se iniciará con un modelo con todas las variables a excepción del nombre del jugador.

```
library(MASS)
library (leaps)

reg0<-lm(Salary~ .-Player, data=mNba)
```

1. *Stepwise*

```
#1. Forward Stepwise
stepAIC(reg0, direction="both")
```

```
## Start:  AIC=14896.49
## Salary ~ (Player + NBA_Country + NBA_DraftNumber + Age + Tm +
##     G + MP + PER + TS + TPAr + FTr + ORB + DRB + TRB + AST +
##     STL + BLK + TOV + USG + OWS + DWS + WS + WS48 + OBPM + DBPM +
##     BPM + VORP) - Player
##
##                   Df  Sum of Sq        RSS   AIC
## - Tm              30 5.3997e+14 9.5599e+15 14864
## - NBA_Country     43 1.4082e+15 1.0428e+16 14880
## - DWS              1 6.9653e+11 9.0206e+15 14894
## - OWS              1 1.0303e+12 9.0209e+15 14894
## - TPAr             1 1.1747e+12 9.0211e+15 14894
## - TS               1 1.1754e+12 9.0211e+15 14894
## - VORP             1 1.7658e+12 9.0217e+15 14895
## - WS               1 2.8502e+12 9.0227e+15 14895
## - AST              1 3.9274e+12 9.0238e+15 14895
## - BLK              1 6.6924e+12 9.0266e+15 14895
## - FTr              1 1.0614e+13 9.0305e+15 14895
## - TOV              1 1.2776e+13 9.0327e+15 14895
## - STL              1 1.3806e+13 9.0337e+15 14895
## - ORB              1 2.8557e+13 9.0485e+15 14896
## - OBPM             1 2.9584e+13 9.0495e+15 14896
## - BPM              1 3.0868e+13 9.0508e+15 14896
## - DRB              1 3.2210e+13 9.0521e+15 14896
## - TRB              1 3.2664e+13 9.0526e+15 14896
## - DBPM             1 3.7381e+13 9.0573e+15 14896
## <none>                          9.0199e+15 14896
## - WS48             1 4.2551e+13 9.0625e+15 14897
## - PER              1 4.7422e+13 9.0673e+15 14897
## - USG              1 1.2616e+14 9.1461e+15 14901
## - NBA_DraftNumber  1 3.6069e+14 9.3806e+15 14913
## - MP               1 5.7716e+14 9.5971e+15 14924
## - G                1 9.6884e+14 9.9887e+15 14944
## - Age              1 1.5730e+15 1.0593e+16 14972
##
## Step:  AIC=14864.45
## Salary ~ NBA_Country + NBA_DraftNumber + Age + G + MP + PER +
##     TS + TPAr + FTr + ORB + DRB + TRB + AST + STL + BLK + TOV +
##     USG + OWS + DWS + WS + WS48 + OBPM + DBPM + BPM + VORP
##
##                   Df  Sum of Sq        RSS   AIC
## - NBA_Country     43 1.4055e+15 1.0965e+16 14844
## - WS48             1 3.3291e+10 9.5599e+15 14862
## - AST              1 7.1581e+11 9.5606e+15 14862
## - TOV              1 1.0561e+12 9.5609e+15 14862
## - STL              1 1.0722e+12 9.5609e+15 14862
## - OWS              1 1.7410e+12 9.5616e+15 14862
## - DWS              1 2.4860e+12 9.5624e+15 14863
## - WS               1 4.0162e+12 9.5639e+15 14863
## - VORP             1 6.8381e+12 9.5667e+15 14863
```

```
## - BLK               1 7.3348e+12 9.5672e+15 14863
## - TS                1 9.6944e+12 9.5696e+15 14863
## - DRB               1 1.1755e+13 9.5716e+15 14863
## - TRB               1 1.7183e+13 9.5770e+15 14863
## - BPM               1 1.7207e+13 9.5771e+15 14863
## - ORB               1 1.7939e+13 9.5778e+15 14863
## - DBPM              1 1.8709e+13 9.5786e+15 14863
## - OBPM              1 2.2870e+13 9.5827e+15 14864
## - FTr               1 2.5467e+13 9.5853e+15 14864
## - TPAr              1 2.6011e+13 9.5859e+15 14864
## - PER               1 3.2343e+13 9.5922e+15 14864
## <none>                          9.5599e+15 14864
## - USG               1 5.7472e+13 9.6173e+15 14865
## - NBA_DraftNumber   1 4.0098e+14 9.9608e+15 14882
## - MP                1 5.5042e+14 1.0110e+16 14889
## + Tm               30 5.3997e+14 9.0199e+15 14896
## - G                 1 7.9704e+14 1.0357e+16 14901
## - Age               1 1.7892e+15 1.1349e+16 14945
##
## Step:  AIC=14844.43
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + DRB + TRB + AST + STL + BLK + TOV + USG + OWS + DWS +
##     WS + WS48 + OBPM + DBPM + BPM + VORP
##
##                    Df  Sum of Sq        RSS   AIC
## - WS48              1 1.5298e+12 1.0967e+16 14842
## - BPM               1 2.1304e+12 1.0967e+16 14842
## - STL               1 2.4496e+12 1.0968e+16 14842
## - DBPM              1 2.5774e+12 1.0968e+16 14842
## - OWS               1 3.1705e+12 1.0969e+16 14843
## - DWS               1 4.7383e+12 1.0970e+16 14843
## - OBPM              1 5.0098e+12 1.0970e+16 14843
## - BLK               1 5.6891e+12 1.0971e+16 14843
## - WS                1 5.9420e+12 1.0971e+16 14843
## - AST               1 8.9755e+12 1.0974e+16 14843
## - DRB               1 1.0540e+13 1.0976e+16 14843
## - TOV               1 1.1162e+13 1.0977e+16 14843
## - VORP              1 1.2742e+13 1.0978e+16 14843
## - TRB               1 1.5978e+13 1.0981e+16 14843
## - FTr               1 1.6476e+13 1.0982e+16 14843
## - ORB               1 1.6527e+13 1.0982e+16 14843
## - TS                1 2.6689e+13 1.0992e+16 14844
## <none>                          1.0965e+16 14844
## - PER               1 5.0075e+13 1.1015e+16 14845
## - TPAr              1 5.3457e+13 1.1019e+16 14845
## - USG               1 7.5964e+13 1.1041e+16 14846
## + NBA_Country      43 1.4055e+15 9.5599e+15 14864
## - NBA_DraftNumber   1 5.3907e+14 1.1504e+16 14866
## - MP                1 6.0742e+14 1.1573e+16 14868
## - G                 1 8.3438e+14 1.1800e+16 14878
```

```
## + Tm               30 5.3725e+14 1.0428e+16 14880
## - Age               1 2.0428e+15 1.3008e+16 14925
##
## Step:  AIC=14842.5
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + DRB + TRB + AST + STL + BLK + TOV + USG + OWS + DWS +
##     WS + OBPM + DBPM + BPM + VORP
##
##                   Df  Sum of Sq        RSS   AIC
## - BPM              1 2.0546e+12 1.0969e+16 14841
## - DBPM             1 2.5299e+12 1.0969e+16 14841
## - OWS              1 3.2497e+12 1.0970e+16 14841
## - STL              1 3.7918e+12 1.0971e+16 14841
## - DWS              1 4.4964e+12 1.0971e+16 14841
## - BLK              1 4.9284e+12 1.0972e+16 14841
## - OBPM             1 5.0589e+12 1.0972e+16 14841
## - WS               1 6.1169e+12 1.0973e+16 14841
## - TOV              1 9.9578e+12 1.0977e+16 14841
## - DRB              1 1.1528e+13 1.0978e+16 14841
## - VORP             1 1.1561e+13 1.0978e+16 14841
## - AST              1 1.1578e+13 1.0978e+16 14841
## - TRB              1 1.7156e+13 1.0984e+16 14841
## - ORB              1 1.8175e+13 1.0985e+16 14841
## - FTr              1 2.3660e+13 1.0991e+16 14842
## - TS               1 2.7154e+13 1.0994e+16 14842
## <none>                          1.0967e+16 14842
## - TPAr             1 5.4058e+13 1.1021e+16 14843
## - PER              1 6.8771e+13 1.1036e+16 14844
## + WS48             1 1.5298e+12 1.0965e+16 14844
## - USG              1 1.4716e+14 1.1114e+16 14847
## + NBA_Country     43 1.4070e+15 9.5599e+15 14862
## - NBA_DraftNumber  1 5.3767e+14 1.1505e+16 14864
## - MP               1 7.4561e+14 1.1713e+16 14872
## - G                1 8.3519e+14 1.1802e+16 14876
## + Tm              30 4.9917e+14 1.0468e+16 14880
## - Age              1 2.0500e+15 1.3017e+16 14923
##
## Step:  AIC=14840.59
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + DRB + TRB + AST + STL + BLK + TOV + USG + OWS + DWS +
##     WS + OBPM + DBPM + VORP
##
##                   Df  Sum of Sq        RSS   AIC
## - OWS              1 3.5227e+12 1.0972e+16 14839
## - STL              1 4.2630e+12 1.0973e+16 14839
## - BLK              1 4.4549e+12 1.0973e+16 14839
## - DWS              1 4.7876e+12 1.0974e+16 14839
## - DBPM             1 5.2300e+12 1.0974e+16 14839
## - WS               1 6.4673e+12 1.0975e+16 14839
## - TOV              1 9.4275e+12 1.0978e+16 14839
```

```
## - AST              1 1.1324e+13 1.0980e+16 14839
## - VORP             1 1.1748e+13 1.0981e+16 14839
## - DRB              1 1.2170e+13 1.0981e+16 14839
## - TRB              1 1.7838e+13 1.0987e+16 14839
## - ORB              1 1.8842e+13 1.0988e+16 14839
## - FTr              1 2.3253e+13 1.0992e+16 14840
## - TS               1 2.6559e+13 1.0996e+16 14840
## <none>                          1.0969e+16 14841
## - TPAr             1 5.2533e+13 1.1021e+16 14841
## - PER              1 6.6973e+13 1.1036e+16 14842
## - OBPM             1 8.8979e+13 1.1058e+16 14842
## + BPM              1 2.0546e+12 1.0967e+16 14842
## + WS48             1 1.4540e+12 1.0967e+16 14842
## - USG              1 1.4690e+14 1.1116e+16 14845
## + NBA_Country     43 1.3919e+15 9.5771e+15 14861
## - NBA_DraftNumber  1 5.4094e+14 1.1510e+16 14862
## - MP               1 7.4506e+14 1.1714e+16 14870
## - G                1 8.3454e+14 1.1803e+16 14874
## + Tm              30 4.9293e+14 1.0476e+16 14878
## - Age              1 2.0592e+15 1.3028e+16 14921
##
## Step:  AIC=14838.74
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + DRB + TRB + AST + STL + BLK + TOV + USG + DWS + WS +
##     OBPM + DBPM + VORP
##
##                   Df  Sum of Sq        RSS   AIC
## - STL              1 3.5698e+12 1.0976e+16 14837
## - BLK              1 4.8800e+12 1.0977e+16 14837
## - DBPM             1 4.8965e+12 1.0977e+16 14837
## - DWS              1 5.4193e+12 1.0978e+16 14837
## - TOV              1 1.0004e+13 1.0982e+16 14837
## - DRB              1 1.2156e+13 1.0985e+16 14837
## - VORP             1 1.2260e+13 1.0985e+16 14837
## - AST              1 1.2396e+13 1.0985e+16 14837
## - TRB              1 1.7865e+13 1.0990e+16 14838
## - ORB              1 1.8906e+13 1.0991e+16 14838
## - FTr              1 2.3522e+13 1.0996e+16 14838
## - TS               1 2.6585e+13 1.0999e+16 14838
## <none>                          1.0972e+16 14839
## - TPAr             1 5.2720e+13 1.1025e+16 14839
## - PER              1 6.8197e+13 1.1041e+16 14840
## - WS               1 7.1930e+13 1.1044e+16 14840
## + OWS              1 3.5227e+12 1.0969e+16 14841
## + BPM              1 2.3276e+12 1.0970e+16 14841
## + WS48             1 1.5297e+12 1.0971e+16 14841
## - OBPM             1 9.0218e+13 1.1063e+16 14841
## - USG              1 1.4877e+14 1.1121e+16 14843
## + NBA_Country     43 1.3933e+15 9.5792e+15 14859
## - NBA_DraftNumber  1 5.5063e+14 1.1523e+16 14860
```

```
## - MP               1 7.4186e+14 1.1714e+16 14868
## - G                1 8.3124e+14 1.1804e+16 14872
## + Tm              30 4.9518e+14 1.0477e+16 14876
## - Age              1 2.0567e+15 1.3029e+16 14919
##
## Step:  AIC=14836.9
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + DRB + TRB + AST + BLK + TOV + USG + DWS + WS + OBPM +
##     DBPM + VORP
##
##                   Df  Sum of Sq        RSS   AIC
## - DBPM             1 1.5020e+12 1.0978e+16 14835
## - DWS              1 3.8687e+12 1.0980e+16 14835
## - VORP             1 1.0829e+13 1.0987e+16 14835
## - DRB              1 1.1341e+13 1.0987e+16 14835
## - AST              1 1.1983e+13 1.0988e+16 14835
## - TOV              1 1.3730e+13 1.0990e+16 14836
## - BLK              1 1.7025e+13 1.0993e+16 14836
## - TRB              1 1.7624e+13 1.0994e+16 14836
## - ORB              1 1.8909e+13 1.0995e+16 14836
## - FTr              1 2.4990e+13 1.1001e+16 14836
## - TS               1 2.8005e+13 1.1004e+16 14836
## <none>                          1.0976e+16 14837
## - TPAr             1 7.4193e+13 1.1050e+16 14838
## - WS               1 7.4919e+13 1.1051e+16 14838
## + STL              1 3.5698e+12 1.0972e+16 14839
## + OWS              1 2.8296e+12 1.0973e+16 14839
## + WS48             1 2.8155e+12 1.0973e+16 14839
## + BPM              1 2.7538e+12 1.0973e+16 14839
## - PER              1 1.1299e+14 1.1089e+16 14840
## - OBPM             1 1.3810e+14 1.1114e+16 14841
## - USG              1 1.4521e+14 1.1121e+16 14841
## + NBA_Country     43 1.3951e+15 9.5809e+15 14858
## - NBA_DraftNumber  1 5.4957e+14 1.1526e+16 14858
## - MP               1 7.3832e+14 1.1714e+16 14866
## - G                1 8.2915e+14 1.1805e+16 14870
## + Tm              30 4.9339e+14 1.0483e+16 14875
## - Age              1 2.0537e+15 1.3030e+16 14917
##
## Step:  AIC=14834.96
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + DRB + TRB + AST + BLK + TOV + USG + DWS + WS + OBPM +
##     VORP
##
##                   Df  Sum of Sq        RSS   AIC
## - DWS              1 2.5432e+12 1.0980e+16 14833
## - DRB              1 1.1222e+13 1.0989e+16 14834
## - AST              1 1.2337e+13 1.0990e+16 14834
## - VORP             1 1.3183e+13 1.0991e+16 14834
## - TOV              1 1.6375e+13 1.0994e+16 14834
```

```
## - TRB               1 1.7745e+13 1.0995e+16 14834
## - ORB               1 1.9346e+13 1.0997e+16 14834
## - FTr               1 2.3801e+13 1.1001e+16 14834
## - BLK               1 2.8073e+13 1.1006e+16 14834
## - TS                1 3.7322e+13 1.1015e+16 14835
## <none>                          1.0978e+16 14835
## - WS                1 7.4737e+13 1.1052e+16 14836
## + OWS               1 3.0103e+12 1.0975e+16 14837
## + WS48              1 2.4666e+12 1.0975e+16 14837
## + DBPM              1 1.5020e+12 1.0976e+16 14837
## + BPM               1 1.3077e+12 1.0976e+16 14837
## + STL               1 1.7529e+11 1.0977e+16 14837
## - TPAr              1 9.8188e+13 1.1076e+16 14837
## - PER               1 1.1826e+14 1.1096e+16 14838
## - USG               1 1.6166e+14 1.1139e+16 14840
## - OBPM              1 1.6370e+14 1.1141e+16 14840
## + NBA_Country      43 1.3906e+15 9.5869e+15 14856
## - NBA_DraftNumber   1 5.4852e+14 1.1526e+16 14856
## - MP                1 7.4481e+14 1.1722e+16 14864
## - G                 1 8.3423e+14 1.1812e+16 14868
## + Tm               30 4.9122e+14 1.0486e+16 14873
## - Age               1 2.0553e+15 1.3033e+16 14916
##
## Step:  AIC=14833.08
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + DRB + TRB + AST + BLK + TOV + USG + WS + OBPM + VORP
##
##                    Df  Sum of Sq        RSS   AIC
## - DRB               1 9.7386e+12 1.0990e+16 14832
## - VORP              1 1.1157e+13 1.0991e+16 14832
## - AST               1 1.2431e+13 1.0993e+16 14832
## - TRB               1 1.5910e+13 1.0996e+16 14832
## - TOV               1 1.6031e+13 1.0996e+16 14832
## - ORB               1 1.7434e+13 1.0998e+16 14832
## - FTr               1 2.3462e+13 1.1004e+16 14832
## - BLK               1 2.7535e+13 1.1008e+16 14832
## - TS                1 3.7169e+13 1.1017e+16 14833
## <none>                          1.0980e+16 14833
## - WS                1 7.3401e+13 1.1053e+16 14834
## + DWS               1 2.5432e+12 1.0978e+16 14835
## + OWS               1 1.8586e+12 1.0978e+16 14835
## + STL               1 4.7400e+11 1.0980e+16 14835
## + WS48              1 4.2551e+11 1.0980e+16 14835
## + DBPM              1 1.7651e+11 1.0980e+16 14835
## + BPM               1 1.2306e+11 1.0980e+16 14835
## - TPAr              1 1.0169e+14 1.1082e+16 14836
## - PER               1 1.2566e+14 1.1106e+16 14836
## - USG               1 1.6886e+14 1.1149e+16 14838
## - OBPM              1 1.7576e+14 1.1156e+16 14839
## + NBA_Country      43 1.3930e+15 9.5871e+15 14854
```

```
## - NBA_DraftNumber  1 5.4603e+14 1.1526e+16 14854
## - MP               1 8.2602e+14 1.1806e+16 14866
## - G                1 8.6223e+14 1.1842e+16 14867
## + Tm              30 4.9307e+14 1.0487e+16 14871
## - Age              1 2.0660e+15 1.3046e+16 14914
##
## Step:  AIC=14831.5
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + TRB + AST + BLK + TOV + USG + WS + OBPM + VORP
##
##                   Df  Sum of Sq        RSS   AIC
## - VORP             1 9.6558e+12 1.0999e+16 14830
## - AST              1 1.2206e+13 1.1002e+16 14830
## - TOV              1 1.7171e+13 1.1007e+16 14830
## - FTr              1 2.2295e+13 1.1012e+16 14830
## - BLK              1 3.1464e+13 1.1021e+16 14831
## - TS               1 3.5029e+13 1.1025e+16 14831
## <none>                          1.0990e+16 14832
## + DRB              1 9.7386e+12 1.0980e+16 14833
## - WS               1 8.8156e+13 1.1078e+16 14833
## - ORB              1 8.8766e+13 1.1079e+16 14833
## + WS48             1 1.2812e+12 1.0989e+16 14833
## + DWS              1 1.0593e+12 1.0989e+16 14834
## + OWS              1 6.4613e+11 1.0989e+16 14834
## + DBPM             1 3.7646e+11 1.0989e+16 14834
## + BPM              1 2.8678e+11 1.0990e+16 14834
## + STL              1 2.3101e+11 1.0990e+16 14834
## - TPAr             1 1.0225e+14 1.1092e+16 14834
## - PER              1 1.3424e+14 1.1124e+16 14835
## - USG              1 1.7616e+14 1.1166e+16 14837
## - OBPM             1 1.8230e+14 1.1172e+16 14837
## - TRB              1 1.9655e+14 1.1186e+16 14838
## + NBA_Country     43 1.3883e+15 9.6016e+15 14852
## - NBA_DraftNumber  1 5.5089e+14 1.1541e+16 14853
## - MP               1 8.1804e+14 1.1808e+16 14864
## - G                1 8.6087e+14 1.1851e+16 14866
## + Tm              30 4.7711e+14 1.0513e+16 14870
## - Age              1 2.0639e+15 1.3054e+16 14912
##
## Step:  AIC=14829.92
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + TRB + AST + BLK + TOV + USG + WS + OBPM
##
##                   Df  Sum of Sq        RSS   AIC
## - AST              1 8.3860e+12 1.1008e+16 14828
## - FTr              1 2.2035e+13 1.1022e+16 14829
## - TOV              1 2.2585e+13 1.1022e+16 14829
## - TS               1 4.1772e+13 1.1041e+16 14830
## <none>                          1.0999e+16 14830
## - BLK              1 4.9589e+13 1.1049e+16 14830
```

```
## + VORP             1 9.6558e+12 1.0990e+16 14832
## + DRB              1 8.2373e+12 1.0991e+16 14832
## + DBPM             1 2.2402e+12 1.0997e+16 14832
## + BPM              1 2.0164e+12 1.0997e+16 14832
## + WS48             1 6.2055e+11 1.0999e+16 14832
## + STL              1 1.5982e+11 1.0999e+16 14832
## + DWS              1 7.8336e+10 1.0999e+16 14832
## + OWS              1 1.8121e+09 1.0999e+16 14832
## - ORB              1 9.5693e+13 1.1095e+16 14832
## - TPAr             1 1.0599e+14 1.1105e+16 14832
## - PER              1 1.7158e+14 1.1171e+16 14835
## - USG              1 1.9367e+14 1.1193e+16 14836
## - TRB              1 2.2667e+14 1.1226e+16 14838
## - OBPM             1 2.3106e+14 1.1231e+16 14838
## + NBA_Country     43 1.3887e+15 9.6108e+15 14851
## - NBA_DraftNumber  1 5.4563e+14 1.1545e+16 14851
## - WS               1 5.5457e+14 1.1554e+16 14852
## - MP               1 8.1309e+14 1.1813e+16 14862
## + Tm              30 4.8118e+14 1.0518e+16 14868
## - G                1 1.0474e+15 1.2047e+16 14872
## - Age              1 2.0557e+15 1.3055e+16 14910
##
## Step:  AIC=14828.29
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + TRB + BLK + TOV + USG + WS + OBPM
##
##                   Df  Sum of Sq        RSS   AIC
## - TOV              1 1.4391e+13 1.1022e+16 14827
## - FTr              1 2.7429e+13 1.1035e+16 14828
## - TS               1 3.3442e+13 1.1041e+16 14828
## <none>                          1.1008e+16 14828
## - BLK              1 5.2776e+13 1.1061e+16 14829
## + AST              1 8.3860e+12 1.0999e+16 14830
## + DRB              1 8.3443e+12 1.1000e+16 14830
## + VORP             1 5.8358e+12 1.1002e+16 14830
## - ORB              1 8.7538e+13 1.1095e+16 14830
## + WS48             1 1.9893e+12 1.1006e+16 14830
## + DBPM             1 1.9564e+12 1.1006e+16 14830
## + BPM              1 1.7579e+12 1.1006e+16 14830
## + DWS              1 2.0704e+11 1.1008e+16 14830
## + STL              1 1.2954e+11 1.1008e+16 14830
## + OWS              1 3.6926e+10 1.1008e+16 14830
## - TPAr             1 9.7872e+13 1.1106e+16 14830
## - PER              1 1.6921e+14 1.1177e+16 14834
## - USG              1 1.9411e+14 1.1202e+16 14835
## - OBPM             1 2.2732e+14 1.1235e+16 14836
## - TRB              1 2.3045e+14 1.1238e+16 14836
## + NBA_Country     43 1.3970e+15 9.6109e+15 14849
## - WS               1 5.4904e+14 1.1557e+16 14850
## - NBA_DraftNumber  1 5.6019e+14 1.1568e+16 14850
```

```
## - MP                1 8.4888e+14 1.1857e+16 14862
## + Tm               30 4.8249e+14 1.0525e+16 14867
## - G                 1 1.0775e+15 1.2085e+16 14871
## - Age               1 2.0475e+15 1.3055e+16 14908
##
## Step:  AIC=14826.92
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TS + TPAr + FTr +
##     ORB + TRB + BLK + USG + WS + OBPM
##
##                   Df  Sum of Sq        RSS   AIC
## - TS               1 2.4129e+13 1.1046e+16 14826
## - FTr              1 3.4217e+13 1.1056e+16 14826
## <none>                          1.1022e+16 14827
## - BLK              1 5.1906e+13 1.1074e+16 14827
## + TOV              1 1.4391e+13 1.1008e+16 14828
## + VORP             1 1.2245e+13 1.1010e+16 14828
## + DRB              1 9.0508e+12 1.1013e+16 14828
## + DBPM             1 6.1485e+12 1.1016e+16 14829
## + BPM              1 5.8268e+12 1.1016e+16 14829
## - ORB              1 8.9439e+13 1.1112e+16 14829
## + STL              1 3.2874e+11 1.1022e+16 14829
## + AST              1 1.9250e+11 1.1022e+16 14829
## + OWS              1 8.2417e+10 1.1022e+16 14829
## + WS48             1 5.9220e+10 1.1022e+16 14829
## + DWS              1 6.4346e+08 1.1022e+16 14829
## - TPAr             1 1.0854e+14 1.1131e+16 14830
## - PER              1 1.7101e+14 1.1193e+16 14832
## - USG              1 1.8722e+14 1.1209e+16 14833
## - OBPM             1 2.1767e+14 1.1240e+16 14834
## - TRB              1 2.2931e+14 1.1252e+16 14835
## + NBA_Country     43 1.4114e+15 9.6109e+15 14847
## - WS               1 5.4637e+14 1.1569e+16 14848
## - NBA_DraftNumber  1 5.5112e+14 1.1573e+16 14848
## - MP               1 8.9903e+14 1.1921e+16 14863
## + Tm              30 4.8653e+14 1.0536e+16 14865
## - G                1 1.1583e+15 1.2181e+16 14873
## - Age              1 2.0905e+15 1.3113e+16 14908
##
## Step:  AIC=14825.97
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + FTr +
##     ORB + TRB + BLK + USG + WS + OBPM
##
##                   Df  Sum of Sq        RSS   AIC
## - FTr              1 2.7270e+13 1.1074e+16 14825
## - BLK              1 4.5593e+13 1.1092e+16 14826
## <none>                          1.1046e+16 14826
## + TS               1 2.4129e+13 1.1022e+16 14827
## + VORP             1 1.8974e+13 1.1027e+16 14827
## + DBPM             1 1.5698e+13 1.1031e+16 14827
## + BPM              1 1.5237e+13 1.1031e+16 14827
```

```
## + DRB               1 6.2772e+12 1.1040e+16 14828
## + TOV               1 5.0782e+12 1.1041e+16 14828
## + STL               1 3.6911e+12 1.1043e+16 14828
## + AST               1 1.2320e+12 1.1045e+16 14828
## + OWS               1 2.0390e+11 1.1046e+16 14828
## + DWS               1 4.1004e+10 1.1046e+16 14828
## + WS48              1 3.2684e+09 1.1046e+16 14828
## - ORB               1 9.6147e+13 1.1143e+16 14828
## - TPAr              1 9.9329e+13 1.1146e+16 14828
## - PER               1 1.7649e+14 1.1223e+16 14832
## - OBPM              1 1.9498e+14 1.1241e+16 14832
## - USG               1 1.9598e+14 1.1242e+16 14832
## - TRB               1 2.3128e+14 1.1278e+16 14834
## + NBA_Country      43 1.4078e+15 9.6386e+15 14846
## - NBA_DraftNumber  1 5.7223e+14 1.1619e+16 14848
## - WS                1 5.7453e+14 1.1621e+16 14848
## - MP                1 9.4256e+14 1.1989e+16 14863
## + Tm               30 4.7363e+14 1.0573e+16 14865
## - G                 1 1.2567e+15 1.2303e+16 14876
## - Age               1 2.0683e+15 1.3115e+16 14906
##
## Step:  AIC=14825.16
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB +
##     TRB + BLK + USG + WS + OBPM
##
##                    Df  Sum of Sq        RSS   AIC
## - BLK               1 4.3100e+13 1.1117e+16 14825
## <none>                           1.1074e+16 14825
## + FTr               1 2.7270e+13 1.1046e+16 14826
## + VORP              1 1.7990e+13 1.1056e+16 14826
## + TS                1 1.7182e+13 1.1056e+16 14826
## + DBPM              1 1.1212e+13 1.1062e+16 14827
## + BPM               1 1.0872e+13 1.1063e+16 14827
## + TOV               1 9.9161e+12 1.1064e+16 14827
## + DRB               1 5.7360e+12 1.1068e+16 14827
## + WS48              1 3.2485e+12 1.1070e+16 14827
## + STL               1 1.0035e+12 1.1073e+16 14827
## + AST               1 3.1474e+11 1.1073e+16 14827
## + OWS               1 2.9991e+11 1.1073e+16 14827
## + DWS               1 8.3452e+10 1.1074e+16 14827
## - ORB               1 1.0029e+14 1.1174e+16 14828
## - TPAr              1 1.1768e+14 1.1191e+16 14828
## - PER               1 1.6452e+14 1.1238e+16 14830
## - USG               1 1.8932e+14 1.1263e+16 14831
## - OBPM              1 1.8933e+14 1.1263e+16 14831
## - TRB               1 2.2738e+14 1.1301e+16 14833
## + NBA_Country      43 1.4148e+15 9.6588e+15 14845
## - NBA_DraftNumber  1 5.5310e+14 1.1627e+16 14847
## - WS                1 6.1131e+14 1.1685e+16 14849
## - MP                1 9.3036e+14 1.2004e+16 14862
```

```
## + Tm              30 4.7356e+14 1.0600e+16 14864
## - G                1 1.2865e+15 1.2360e+16 14876
## - Age              1 2.0480e+15 1.3122e+16 14905
##
## Step:  AIC=14825.03
## Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB +
##     TRB + USG + WS + OBPM
##
##                   Df  Sum of Sq        RSS   AIC
## <none>                           1.1117e+16 14825
## + BLK              1 4.3100e+13 1.1074e+16 14825
## + DBPM             1 3.5607e+13 1.1081e+16 14826
## + BPM              1 3.5118e+13 1.1082e+16 14826
## + VORP             1 3.4610e+13 1.1082e+16 14826
## + FTr              1 2.4777e+13 1.1092e+16 14826
## - ORB              1 7.6130e+13 1.1193e+16 14826
## + TS               1 1.2369e+13 1.1104e+16 14826
## + TOV              1 1.0155e+13 1.1107e+16 14827
## + DRB              1 9.0678e+12 1.1108e+16 14827
## - TPAr             1 8.4890e+13 1.1202e+16 14827
## + WS48             1 2.6031e+12 1.1114e+16 14827
## + OWS              1 1.9057e+12 1.1115e+16 14827
## + DWS              1 1.2057e+12 1.1116e+16 14827
## + STL              1 1.0520e+11 1.1117e+16 14827
## + AST              1 1.8870e+09 1.1117e+16 14827
## - PER              1 1.2217e+14 1.1239e+16 14828
## - OBPM             1 1.4724e+14 1.1264e+16 14829
## - USG              1 1.4871e+14 1.1265e+16 14829
## - TRB              1 2.0511e+14 1.1322e+16 14832
## + NBA_Country     43 1.4115e+15 9.7053e+15 14846
## - NBA_DraftNumber  1 5.8649e+14 1.1703e+16 14848
## - WS               1 6.3630e+14 1.1753e+16 14850
## - MP               1 9.0936e+14 1.2026e+16 14861
## + Tm              30 4.8614e+14 1.0631e+16 14864
## - G                1 1.2443e+15 1.2361e+16 14874
## - Age              1 2.0455e+15 1.3162e+16 14904

##
## Call:
## lm(formula = Salary ~ NBA_DraftNumber + Age + G + MP + PER +
##     TPAr + ORB + TRB + USG + WS + OBPM, data = mNba)
##
## Coefficients:
##     (Intercept)  NBA_DraftNumber              Age                G
##        -3518568           -59427           491764          -150949
##              MP              PER             TPAr              ORB
##            5065          -321391         -3126988          -161178
##             TRB              USG               WS             OBPM
##          297419           132197           834067           584764
```

Este método selecciono el modelo Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB + TRB + USG + WS + OBPM, obteniendo un AIC de 14825.03.

2. *Cross Validation*

   Se validará con las técnicas del Cross validation la eficiencia de los 4 modelos que se han formulado. Caso 1 (Inicial): Salary~.-Player-NBA_Country-Tm Caso 2 (Summary): Salary~ NBA_DraftNumber+Age+G+MP Caso 3 (Colinealidad): Salary~ NBA_DraftNumber + Age + FTr + AST + STL + TOV Caso 4 (generado por Stepwise): Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB + TRB + USG + WS + OBPM

a. Validation Set

```
library(ISLR)
set.seed(400)
mNbaF<-mNba[,c(2,4:5,7:28)] #Se eliminan las columnas Player, NBA_Country y T
m por que afectan a las pruebas (problema de multicolinealidad)
numData=nrow(mNbaF)
train=sample(numData ,numData/2)

#Caso 1: reg0= Salary~.-Player-NBA_Country-Tm
regres.train =lm(Salary~.,mNbaF ,subset =train )
attach(mNbaF)
mean((Salary-predict(regres.train ,Auto))[-train ]^2)

## Warning: 'newdata' had 392 rows but variables found have 481 rows

## [1] 2.85034e+13

#2.Caso (Summary): Salary~ NBA_DraftNumber+Age+G+MP
regres.train2 =lm(Salary~ NBA_DraftNumber+Age+G+MP,mNbaF ,subset =train )
mean((Salary-predict(regres.train2 ,Auto))[-train ]^2)

## Warning: 'newdata' had 392 rows but variables found have 481 rows

## [1] 2.936511e+13

#3.Caso (Colinealidad): Salary~ NBA_DraftNumber + Age + FTr + AST + STL + TOV
regres.train3 =lm(Salary~ NBA_DraftNumber + Age + FTr + AST + STL + TOV,mNbaF
,subset =train )
mean((Salary-predict(regres.train3 ,Auto))[-train ]^2)

## Warning: 'newdata' had 392 rows but variables found have 481 rows

## [1] 4.06478e+13

#4.Caso (Stepwise): Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + OR
B + TRB + USG + WS + OBPM
regres.train4 =lm(Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB
+ TRB + USG + WS + OBPM,mNbaF ,subset =train )
mean((Salary-predict(regres.train4 ,Auto))[-train ]^2)
```

```
## Warning: 'newdata' had 392 rows but variables found have 481 rows

## [1] 2.607355e+13
```

Se encontró que el modelo de Stepwise tiene el menor margen de error.

b. Leave One Out Cross Validation

```r
library (boot)

##
## Attaching package: 'boot'

## The following object is masked from 'package:car':
##
##     logit

#1.Caso: reg0= Salary~.-Player-NBA_Country-Tm
glm.fit1=glm(Salary~.,mNbaF ,family = gaussian())
cv.err =cv.glm(mNbaF,glm.fit1)
cv.err$delta[1]

## [1] 2.545618e+13

#2.Caso (Summary): Salary~ NBA_DraftNumber+Age+G+MP
glm.fit2=glm(Salary~ NBA_DraftNumber+Age+G+MP,mNbaF ,family = gaussian())
cv.err =cv.glm(mNbaF,glm.fit2)
cv.err$delta[1]

## [1] 2.706614e+13

#3.Caso (Colinealidad): Salary~ NBA_DraftNumber + Age + FTr + AST + STL + TOV
glm.fit3=glm(Salary~ NBA_DraftNumber + Age + FTr + AST + STL + TOV,mNbaF ,fam
ily = gaussian())
cv.err =cv.glm(mNbaF,glm.fit3)
cv.err$delta[1]

## [1] 3.710536e+13

#4.Caso (Stepwise): Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + OR
B + TRB + USG + WS + OBPM
glm.fit4=glm(Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB + TRB
+ USG + WS + OBPM,mNbaF ,family = gaussian())
cv.err =cv.glm(mNbaF,glm.fit4)
cv.err$delta[1]

## [1] 2.429988e+13
```

Este método tambien selecciono el modelo generado por el Stepwise.

c. K Fold Cross Validation

```r
library (boot)
#1.Caso: reg0= Salary~.-Player-NBA_Country-Tm
```

```r
glm.fit1=glm(Salary~.,mNbaF ,family = gaussian())
cv.err =cv.glm(mNbaF,glm.fit1,K=10)
cv.err$delta[1] #raw cross-validation estimate of prediction error

## [1] 2.54324e+13

#2.Caso (Summary): Salary~ NBA_DraftNumber+Age+G+MP
glm.fit2=glm(Salary~ NBA_DraftNumber+Age+G+MP,mNbaF ,family = gaussian())
cv.err =cv.glm(mNbaF,glm.fit2,K=10)
cv.err$delta[1] #raw cross-validation estimate of prediction error

## [1] 2.69156e+13

#3.Caso (Colinealidad): Salary~ NBA_DraftNumber + Age + FTr + AST + STL + TOV
glm.fit3=glm(Salary~ NBA_DraftNumber + Age + FTr + AST + STL + TOV,mNbaF ,fam
ily = gaussian())
cv.err =cv.glm(mNbaF,glm.fit3,K=10)
cv.err$delta[1]  #raw cross-validation estimate of prediction error

## [1] 3.7387e+13

#4.Caso (Stepwise): Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + OR
B + TRB + USG + WS + OBPM
glm.fit4=glm(Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB + TRB
+ USG + WS + OBPM,mNbaF ,family = gaussian())
cv.err =cv.glm(mNbaF,glm.fit4)
cv.err$delta[1] #raw cross-validation estimate of prediction error

## [1] 2.429988e+13
```

De igual manera, el stepwise es el mejor bajo la revisión de los tres métodos.

Salary ~ NBA_DraftNumber + Age + G + MP + PER + TPAr + ORB + TRB + USG + WS + OBPM