

VariableSelection-HittersData

SP

28/10/2019

```
library(car)
```

```
## Loading required package: carData
```

```
library(MASS)
```

```
library(LEAP)
```

```
## Warning: package 'LEAP' was built under R version 3.6.1
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:MASS':
```

```
##
```

```
##      select
```

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

```
##
```

```
##      recode
```

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

```
##
```

```
##      filter, lag
```

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

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(leaps)
```

```
## Warning: package 'leaps' was built under R version 3.6.1
```

```
library(ISLR)
```

```
## Warning: package 'ISLR' was built under R version 3.6.1
```

```
ds<-read.csv("Hitters.csv")
```

```
head(ds)
```

```
##      AtBat Hits HmRun Runs RBI Walks Years CatBat CHit CHmRun CRun CRBI
## 1      293   66     1  30  29   14     1    293   66     1  30  29
## 2      315  950     7  24  38   39    14   3449  835     69 321 414
## 3      479  130    18  66  72   76     3   1624  457     63 224 266
## 4      496  141    20  65  78   37    11   5628 1575    225 828 838
## 5      321   87    10  39  42   30     2    396  101     12  48  46
## 6      594  169     4  74  51   35    11   4408 1133     19 501 336
##      CWalks League Division PutOuts Assists Errors Salary NewLeague
## 1         14      A        E     446     33     20      NA        A
## 2        375      N        W     632     43     10   475.0        N
## 3        263      A        W     880     82     14   480.0        A
## 4        354      N        E     200     11      3   500.0        N
## 5         33      N        E     805     40      4    91.5        N
## 6        194      A        W     282    421     25   750.0        A
```

```
dim(ds)
```

```
## [1] 322 20
```

```
names(ds)
```

```
## [1] "AtBat"      "Hits"       "HmRun"      "Runs"       "RBI"
## [6] "Walks"      "Years"      "CatBat"     "CHit"       "CHmRun"
## [11] "CRun"       "CRBI"       "CWalks"     "League"     "Division"
## [16] "PutOuts"    "Assists"    "Errors"     "Salary"     "NewLeague"
```

```
summary(ds)
```

```
##      AtBat      Hits      HmRun      Runs
## Min.   : 16.0   Min.   : 1.0   Min.   : 0.00   Min.   : 0.00
## 1st Qu.:255.2   1st Qu.: 64.0   1st Qu.: 4.00   1st Qu.: 30.25
## Median :379.5   Median : 96.0   Median : 8.00   Median : 48.00
## Mean   :380.9   Mean   :103.7   Mean   :10.77   Mean   : 50.91
## 3rd Qu.:512.0   3rd Qu.:137.8   3rd Qu.:16.00   3rd Qu.: 69.00
## Max.   :687.0   Max.   :950.0   Max.   :40.00   Max.   :130.00
##
##      RBI      Walks      Years      CatBat
## Min.   : 0.00   Min.   : 0.00   Min.   : 1.000   Min.   : 19.0
## 1st Qu.: 28.00   1st Qu.: 22.00   1st Qu.: 4.000   1st Qu.: 816.8
## Median : 44.00   Median : 35.00   Median : 6.000   Median : 1928.0
## Mean    : 48.03   Mean    : 38.74   Mean    : 7.444   Mean    : 2648.7
## 3rd Qu.: 64.75   3rd Qu.: 53.00   3rd Qu.:11.000   3rd Qu.: 3924.2
## Max.    :121.00   Max.    :105.00   Max.    :24.000   Max.    :14053.0
##
##      CHit      CHmRun      CRun      CRBI
## Min.   : 4.0   Min.   : 0.00   Min.   : 1.0   Min.   : 0.00
## 1st Qu.:209.0   1st Qu.: 14.00   1st Qu.:100.2   1st Qu.: 88.75
## Median : 508.0   Median : 37.50   Median : 247.0   Median : 220.50
## Mean    : 717.6   Mean    : 69.49   Mean    : 358.8   Mean    : 330.12
## 3rd Qu.:1059.2   3rd Qu.: 90.00   3rd Qu.: 526.2   3rd Qu.: 426.25
## Max.    :4256.0   Max.    :548.00   Max.    :2165.0   Max.    :1659.00
##
```

```
##           CWalks           League Division      PutOuts           Assists
## Min.      : 0.00   A:175   E:157   Min.      : 0.0   Min.      : 0.0
## 1st Qu.: 67.25   N:147   W:165   1st Qu.: 109.2   1st Qu.: 7.0
## Median : 170.50                        Median : 212.0   Median : 39.5
## Mean      : 260.24                        Mean      : 288.9   Mean      :106.9
## 3rd Qu.: 339.25                        3rd Qu.: 325.0   3rd Qu.:166.0
## Max.      :1566.00                        Max.      :1378.0   Max.      :492.0
##
##           Errors           Salary           NewLeague
## Min.      : 0.00   Min.      : 67.5   A:176
## 1st Qu.: 3.00   1st Qu.: 190.0   N:146
## Median : 6.00   Median : 425.0
## Mean      : 8.04   Mean      : 535.9
## 3rd Qu.:11.00   3rd Qu.: 750.0
## Max.      :32.00   Max.      :2460.0
##
##           NA's      :59
```

```
ds=na.omit(ds) # removing the NA's of the Salary
dim(ds)
```

```
## [1] 263 20
```

```
model1=lm(Salary~., data=ds)
summary(model1)
```

```
##
## Call:
## lm(formula = Salary ~ ., data = ds)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -869.56 -193.32  -22.83   134.60  1873.28
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  145.22575   92.75984    1.566  0.118741
## AtBat        -0.64967    0.46111   -1.409  0.160141
## Hits          0.39145    0.37498    1.044  0.297562
## HmRun        -1.83802    5.96208   -0.308  0.758129
## Runs          2.94786    2.46373    1.197  0.232666
## RBI           1.22712    2.53702    0.484  0.629046
## Walks         4.47276    1.76643    2.532  0.011970 *
## Years        -1.80218   12.71370   -0.142  0.887394
## CatBat       -0.34012    0.12566   -2.707  0.007278 **
## CHit          1.15335    0.59559    1.936  0.053970 .
## CHmRun        0.89954    1.60684    0.560  0.576116
## CRun          0.61973    0.70857    0.875  0.382649
## CRBI          0.37370    0.68990    0.542  0.588539
## CWalks       -0.55144    0.32187   -1.713  0.087947 .
## LeagueN       78.50054   80.50542    0.975  0.330482
## DivisionW    -122.89677   41.10013   -2.990  0.003075 **
## PutOuts       0.27614    0.07902    3.495  0.000564 ***
## Assists       0.45110    0.22449    2.009  0.045592 *
```

```
## Errors          -5.56545    4.42373   -1.258 0.209567
## NewLeagueN     -43.21387    80.21913   -0.539 0.590589
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 321.3 on 243 degrees of freedom
## Multiple R-squared:  0.5296, Adjusted R-squared:  0.4929
## F-statistic: 14.4 on 19 and 243 DF,  p-value: < 2.2e-16
```

```
#cor(ds)
```

```
AIC(model1)
```

```
## [1] 3803.763
```

```
BIC(model1)
```

```
## [1] 3878.779
```

Let us choose the 7 variables

```
reg_all=regsubsets(Salary~., data=ds, nvmax=7)
```

```
summary(reg_all)
```

```
## Subset selection object
## Call: regsubsets.formula(Salary ~ ., data = ds, nvmax = 7)
## 19 Variables (and intercept)
##              Forced in Forced out
## AtBat          FALSE      FALSE
## Hits           FALSE      FALSE
## HmRun          FALSE      FALSE
## Runs           FALSE      FALSE
## RBI            FALSE      FALSE
## Walks          FALSE      FALSE
## Years          FALSE      FALSE
## CatBat         FALSE      FALSE
## CHit           FALSE      FALSE
## CHmRun         FALSE      FALSE
## CRun           FALSE      FALSE
## CRBI           FALSE      FALSE
## CWalks         FALSE      FALSE
## LeagueN        FALSE      FALSE
## DivisionW      FALSE      FALSE
## PutOuts        FALSE      FALSE
## Assists        FALSE      FALSE
## Errors         FALSE      FALSE
## NewLeagueN     FALSE      FALSE
## 1 subsets of each size up to 7
```

```
## Selection Algorithm: exhaustive
##           AtBat Hits HmRun Runs RBI Walks Years CatBat CHit CHmRun CRun
## 1 ( 1 ) " " " " " " " " " " " " " " " " " " " "
## 2 ( 1 ) " " " " " " "*" " " " " " " " " " " " "
## 3 ( 1 ) " " " " " " "*" " " " " " " " " " " " "
## 4 ( 1 ) " " " " " " "*" " " " " " " " " " " " "
## 5 ( 1 ) " " " " " " " " " " "*" " " "*" "*" " " " "
## 6 ( 1 ) " " " " " " " " " " "*" " " "*" "*" "*" " " "
## 7 ( 1 ) " " " " " " " " " " "*" " " "*" "*" "*" " " "
##           CRBI CWalks LeagueN DivisionW PutOuts Assists Errors NewLeagueN
## 1 ( 1 ) "*" " " " " " " " " " " " " " "
## 2 ( 1 ) "*" " " " " " " " " " " " " " "
## 3 ( 1 ) "*" " " " " " " "*" " " " " " "
## 4 ( 1 ) "*" " " " " "*" "*" " " " " " "
## 5 ( 1 ) " " " " " " "*" "*" " " " " " "
## 6 ( 1 ) " " " " " " "*" "*" " " " " " "
## 7 ( 1 ) " " "*" " " " "*" "*" " " " " " "
```

```
coef(reg_all, 7)
```

```
## (Intercept)      Walks      CatBat      CHit      CHmRun
## 119.5168836    5.6070682   -0.3563967    1.5622172    1.5675278
##      CWalks    DivisionW      PutOuts
##   -0.4027386 -138.4992434    0.2464922
```

functions available in regsubset and summary

```
names(reg_all)
```

```
## [1] "np"      "nrbar"   "d"       "rbar"    "thetab"
## [6] "first"   "last"    "vorder"  "tol"     "rss"
## [11] "bound"   "nvmax"   "ress"    "ir"      "nbest"
## [16] "lopt"    "il"      "ier"     "xnames"  "method"
## [21] "force.in" "force.out" "sserr"   "intercept" "lindep"
## [26] "nullrss" "nn"      "call"
```

```
reg_all_summary = summary(reg_all)
```

```
names(reg_all_summary)
```

```
## [1] "which" "rsq"    "rss"    "adjr2" "cp"     "bic"    "outmat" "obj"
```

Forward Selection with nvmax number of regressors

```
forward_reg = regsubsets(Salary~., data=ds, nvmax= 19,method="forward")
summary(forward_reg)
```

```
## Subset selection object
## Call: regsubsets.formula(Salary ~ ., data = ds, nvmax = 19, method = "forward")
## 19 Variables (and intercept)
##           Forced in Forced out
## AtBat      FALSE      FALSE
## Hits       FALSE      FALSE
## HmRun      FALSE      FALSE
## Runs       FALSE      FALSE
## RBI        FALSE      FALSE
## Walks      FALSE      FALSE
## Years      FALSE      FALSE
## CatBat     FALSE      FALSE
## CHit       FALSE      FALSE
## CHmRun     FALSE      FALSE
## CRun       FALSE      FALSE
## CRBI       FALSE      FALSE
## CWalks     FALSE      FALSE
## LeagueN    FALSE      FALSE
## DivisionW  FALSE      FALSE
## PutOuts    FALSE      FALSE
## Assists    FALSE      FALSE
## Errors     FALSE      FALSE
## NewLeagueN FALSE      FALSE
## 1 subsets of each size up to 19
## Selection Algorithm: forward
##           AtBat Hits HmRun Runs RBI Walks Years CatBat CHit CHmRun CRun
## 1 ( 1 ) " " " " " " " " " " " " " " " " " "
## 2 ( 1 ) " " " " " " "*" " " " " " " " " " "
## 3 ( 1 ) " " " " " " "*" " " " " " " " " " "
## 4 ( 1 ) " " " " " " "*" " " " " " " " " " "
## 5 ( 1 ) " " " " " " "*" " " "*" " " " " " " " "
## 6 ( 1 ) " " " " " " "*" " " "*" " " " " " " " "
## 7 ( 1 ) " " " " " " "*" " " "*" " " " " " " "*"
## 8 ( 1 ) " " " " " " "*" " " "*" " " "*" " " " " "*"
## 9 ( 1 ) " " " " " " "*" " " "*" " " "*" "*" " " " "*"
## 10 ( 1 ) " " " " " " "*" " " "*" " " "*" "*" " " " "*"
## 11 ( 1 ) "*" " " " " "*" " " "*" " " "*" "*" " " " "*"
## 12 ( 1 ) "*" " " " " "*" " " "*" " " "*" "*" " " " "*"
## 13 ( 1 ) "*" "*" " " "*" " " "*" " " "*" "*" " " " "*"
## 14 ( 1 ) "*" "*" " " "*" " " "*" " " "*" "*" " " " "*"
## 15 ( 1 ) "*" "*" " " "*" " " "*" " " "*" "*" " " " "*"
## 16 ( 1 ) "*" "*" " " "*" " " "*" " " "*" "*" "*" " " "*"
## 17 ( 1 ) "*" "*" " " "*" "*" "*" " " "*" "*" "*" " " "*"
## 18 ( 1 ) "*" "*" "*" "*" "*" "*" " " "*" "*" "*" " " "*"
## 19 ( 1 ) "*" "*" "*" "*" "*" "*" "*" "*" "*" "*" " " "*"
##           CRBI CWalks LeagueN DivisionW PutOuts Assists Errors NewLeagueN
## 1 ( 1 ) "*" " " " " " " " " " " " "
## 2 ( 1 ) "*" " " " " " " " " " " " "
## 3 ( 1 ) "*" " " " " " " "*" " " " " "
## 4 ( 1 ) "*" " " " " "*" "*" " " " " "
## 5 ( 1 ) "*" " " " " "*" "*" " " " " "
## 6 ( 1 ) "*" "*" " " "*" "*" " " " " "
## 7 ( 1 ) "*" "*" " " "*" "*" " " " " "
## 8 ( 1 ) "*" "*" " " "*" "*" " " " " "
```

```
## 9 ( 1 ) "*" "*" " " "*" "*" " " " " " "
## 10 ( 1 ) "*" "*" " " "*" "*" "*" " " " "
## 11 ( 1 ) "*" "*" " " "*" "*" "*" " " " "
## 12 ( 1 ) "*" "*" " " "*" "*" "*" "*" " "
## 13 ( 1 ) "*" "*" " " "*" "*" "*" "*" " "
## 14 ( 1 ) "*" "*" "*" "*" "*" "*" "*" " "
## 15 ( 1 ) "*" "*" "*" "*" "*" "*" "*" "*"
## 16 ( 1 ) "*" "*" "*" "*" "*" "*" "*" "*"
## 17 ( 1 ) "*" "*" "*" "*" "*" "*" "*" "*"
## 18 ( 1 ) "*" "*" "*" "*" "*" "*" "*" "*"
## 19 ( 1 ) "*" "*" "*" "*" "*" "*" "*" "*"

```

```
names(forward_reg)
```

```
## [1] "np"      "nrbar"    "d"        "rbar"     "thetab"
## [6] "first"   "last"     "vorder"   "tol"      "rss"
## [11] "bound"   "nvmax"    "ress"     "ir"       "nbest"
## [16] "lopt"    "il"       "ier"      "xnames"   "method"
## [21] "force.in" "force.out" "sserr"    "intercept" "lindep"
## [26] "nullrss" "nn"       "call"

```

Backward Elimination with nvmax number of regressors

```
backward_reg = regsubsets(Salary~., data=ds, nvmax= 19, method="backward")
summary(backward_reg)
```

```
## Subset selection object
## Call: regsubsets.formula(Salary ~ ., data = ds, nvmax = 19, method = "backward")
## 19 Variables (and intercept)
##           Forced in Forced out
## AtBat      FALSE      FALSE
## Hits        FALSE      FALSE
## HmRun        FALSE      FALSE
## Runs         FALSE      FALSE
## RBI          FALSE      FALSE
## Walks        FALSE      FALSE
## Years        FALSE      FALSE
## CatBat       FALSE      FALSE
## CHit         FALSE      FALSE
## CHmRun        FALSE      FALSE
## CRun          FALSE      FALSE
## CRBI          FALSE      FALSE
## CWalks        FALSE      FALSE
## LeagueN       FALSE      FALSE
## DivisionW     FALSE      FALSE
## PutOuts       FALSE      FALSE
## Assists       FALSE      FALSE
## Errors        FALSE      FALSE
## NewLeagueN    FALSE      FALSE
## 1 subsets of each size up to 19
## Selection Algorithm: backward

```

```

##           AtBat Hits HmRun Runs RBI Walks Years CatBat CHit CHmRun CRun
## 1 ( 1 ) " " " " " " " " " " " " " " " " " " " " " "
## 2 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 3 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 4 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 5 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 6 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 7 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 8 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 9 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 10 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 11 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 12 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 13 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 14 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 15 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 16 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 17 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 18 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 19 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
##           CRBI CWalks LeagueN DivisionW PutOuts Assists Errors NewLeagueN
## 1 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 2 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 3 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 4 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 5 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 6 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 7 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 8 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 9 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 10 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 11 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 12 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 13 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 14 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 15 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 16 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 17 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 18 ( 1 ) " " " " " " " " " " " " " " " " " " " " "
## 19 ( 1 ) " " " " " " " " " " " " " " " " " " " " "

```

```
names(backward_reg)
```

```

## [1] "np"          "nrbar"       "d"           "rbar"        "thetab"
## [6] "first"       "last"        "vorder"      "tol"         "rss"
## [11] "bound"       "nvmax"       "ress"        "ir"          "nbest"
## [16] "lopt"        "il"          "ier"         "xnames"      "method"
## [21] "force.in"    "force.out"   "sserr"       "intercept"   "lindep"
## [26] "nullrss"     "nn"          "call"

```


Variable Selection using nvmax=5

```
reg_all_5 = regsubsets(Salary~., data=ds, nvmax=5)
#reg_all_5
```

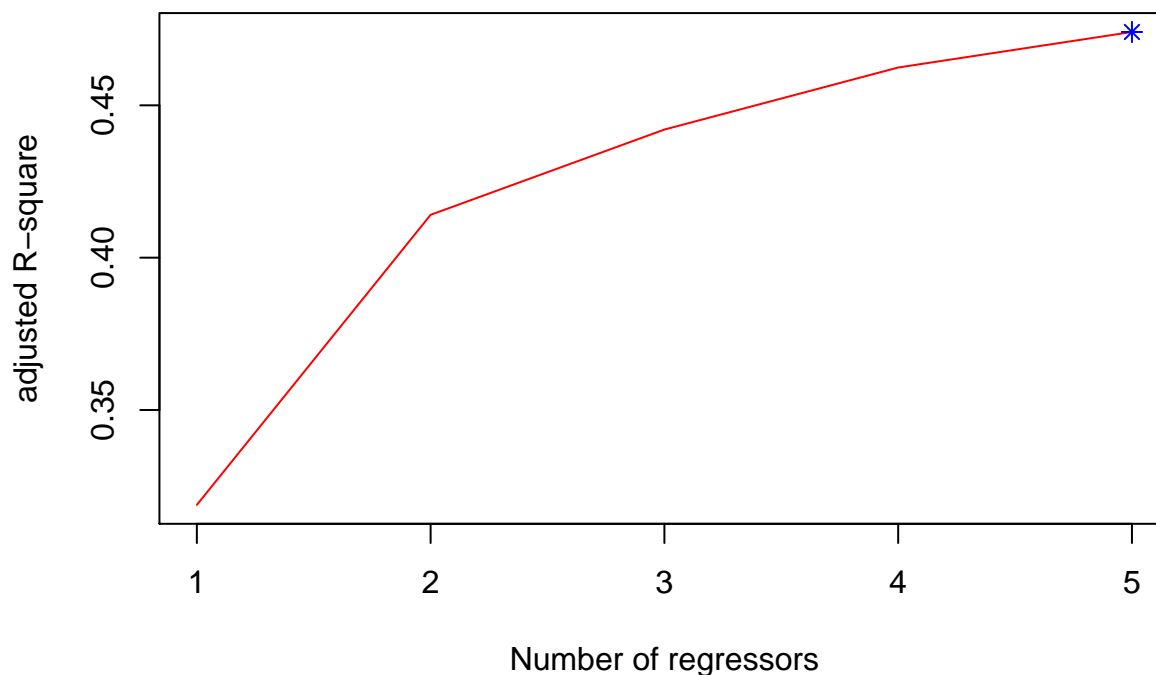
```
summary_5 = summary(reg_all_5)
summary_5
```

```
## Subset selection object
## Call: regsubsets.formula(Salary ~ ., data = ds, nvmax = 5)
## 19 Variables (and intercept)
##           Forced in Forced out
## AtBat      FALSE      FALSE
## Hits       FALSE      FALSE
## HmRun       FALSE      FALSE
## Runs        FALSE      FALSE
## RBI         FALSE      FALSE
## Walks       FALSE      FALSE
## Years       FALSE      FALSE
## CatBat      FALSE      FALSE
## CHit        FALSE      FALSE
## CHmRun      FALSE      FALSE
## CRun        FALSE      FALSE
## CRBI        FALSE      FALSE
## CWalks      FALSE      FALSE
## LeagueN     FALSE      FALSE
## DivisionW   FALSE      FALSE
## PutOuts     FALSE      FALSE
## Assists     FALSE      FALSE
## Errors      FALSE      FALSE
## NewLeagueN  FALSE      FALSE
## 1 subsets of each size up to 5
## Selection Algorithm: exhaustive
##           AtBat Hits HmRun Runs RBI Walks Years CatBat CHit CHmRun CRun
## 1 ( 1 ) " " " " " " " " " " " " " " " " " "
## 2 ( 1 ) " " " " " " "*" " " " " " " " " " "
## 3 ( 1 ) " " " " " " "*" " " " " " " " " " "
## 4 ( 1 ) " " " " " " "*" " " " " " " " " " "
## 5 ( 1 ) " " " " " " " " " " "*" " " "*" " " "
##           CRBI CWalks LeagueN DivisionW PutOuts Assists Errors NewLeagueN
## 1 ( 1 ) "*" " " " " " " " " " " " " " "
## 2 ( 1 ) "*" " " " " " " " " " " " " " "
## 3 ( 1 ) "*" " " " " " " "*" " " " " " "
## 4 ( 1 ) "*" " " " " "*" "*" " " " " " "
## 5 ( 1 ) " " " " " " "*" "*" " " " " " "
```

```
bestsubset_5=which.max(summary_5$adjr2)
bestsubset_5
```

```
## [1] 5
```

```
plot(summary_5$adjr2, xlab="Number of regressors", ylab="adjusted R-square", type = "l", col="red")
points(bestsubset_5,summary_5$adjr2[bestsubset_5],col="blue",pch = 8)
```



From the above curve we can notice that, it is not yet reached the optimum number of regressors

Variable Selection using nvmax=19

```
reg_all_19=regsubsets(Salary~., data=ds, nvmax=19)
```

```
summary(reg_all_19)
```

```
## Subset selection object
## Call: regsubsets.formula(Salary ~ ., data = ds, nvmax = 19)
## 19 Variables (and intercept)
##           Forced in Forced out
## AtBat      FALSE      FALSE
## Hits       FALSE      FALSE
## HmRun       FALSE      FALSE
## Runs       FALSE      FALSE
## RBI        FALSE      FALSE
## Walks      FALSE      FALSE
## Years      FALSE      FALSE
## CatBat     FALSE      FALSE
```

```

## CHit          FALSE      FALSE
## CHmRun        FALSE      FALSE
## CRun          FALSE      FALSE
## CRBI          FALSE      FALSE
## CWalks        FALSE      FALSE
## LeagueN       FALSE      FALSE
## DivisionW     FALSE      FALSE
## PutOuts       FALSE      FALSE
## Assists       FALSE      FALSE
## Errors        FALSE      FALSE
## NewLeagueN    FALSE      FALSE
## 1 subsets of each size up to 19
## Selection Algorithm: exhaustive
##           AtBat Hits HmRun Runs RBI Walks Years CatBat CHit CHmRun CRun
## 1 ( 1 ) " " " " " " " " " " " " " " " " " "
## 2 ( 1 ) " " " " " " "*" " " " " " " " " " " " "
## 3 ( 1 ) " " " " " " "*" " " " " " " " " " " " "
## 4 ( 1 ) " " " " " " "*" " " " " " " " " " " " "
## 5 ( 1 ) " " " " " " " " " " "*" " " "*" "*" " " " "
## 6 ( 1 ) " " " " " " " " " " "*" " " "*" "*" "*" " " "
## 7 ( 1 ) " " " " " " " " " " "*" " " "*" "*" "*" " " "
## 8 ( 1 ) " " " " " " " " " " "*" " " "*" "*" "*" " " "
## 9 ( 1 ) " " " " " " " " " " "*" " " "*" " " "*"
## 10 ( 1 ) " " " " " " " " " " "*" " " "*" " " "*"
## 11 ( 1 ) "*" " " " " "*" " " "*" " " "*" "*" "*" " " "
## 12 ( 1 ) "*" "*" " " "*" " " "*" " " "*" "*" "*" " " "
## 13 ( 1 ) "*" "*" " " "*" " " "*" " " "*" "*" "*" " " "
## 14 ( 1 ) "*" "*" " " "*" " " "*" " " "*" " " "*"
## 15 ( 1 ) "*" "*" " " "*" " " "*" " " "*" " " "*"
## 16 ( 1 ) "*" "*" " " "*" " " "*" " " "*" "*" "*"
## 17 ( 1 ) "*" "*" " " "*" "*" "*" " " "*" "*" "*"
## 18 ( 1 ) "*" "*" "*" "*" "*" "*" " " "*" "*" "*"
## 19 ( 1 ) "*" "*" "*" "*" "*" "*" "*" "*" "*" "*"
##           CRBI CWalks LeagueN DivisionW PutOuts Assists Errors NewLeagueN
## 1 ( 1 ) "*" " " " " " " " " " " " "
## 2 ( 1 ) "*" " " " " " " " " " " " "
## 3 ( 1 ) "*" " " " " " " "*" " " " " "
## 4 ( 1 ) "*" " " " " "*" "*" " " " " "
## 5 ( 1 ) " " " " " " "*" "*" " " " "
## 6 ( 1 ) " " " " " " "*" "*" " " " "
## 7 ( 1 ) " " "*" " " " "*" "*" " " " "
## 8 ( 1 ) " " "*" " " " "*" "*" "*" " " "
## 9 ( 1 ) "*" "*" " " " " "*" "*" "*" " " "
## 10 ( 1 ) "*" "*" " " " " "*" "*" "*" " " "
## 11 ( 1 ) " " "*" " " " "*" "*" "*" " " "
## 12 ( 1 ) " " "*" " " " "*" "*" "*" " " "
## 13 ( 1 ) " " "*" "*" " " "*" "*" "*" " " "
## 14 ( 1 ) "*" "*" "*" " " " "*" "*" "*" " " "
## 15 ( 1 ) "*" "*" "*" " " " "*" "*" "*" "*"
## 16 ( 1 ) "*" "*" "*" " " " "*" "*" "*" "*"
## 17 ( 1 ) "*" "*" "*" " " " "*" "*" "*" "*"
## 18 ( 1 ) "*" "*" "*" " " " "*" "*" "*" "*"
## 19 ( 1 ) "*" "*" "*" " " " "*" "*" "*" "*"

```

```
summary_19 = summary(reg_all_19)
summary_19$adjr2
```

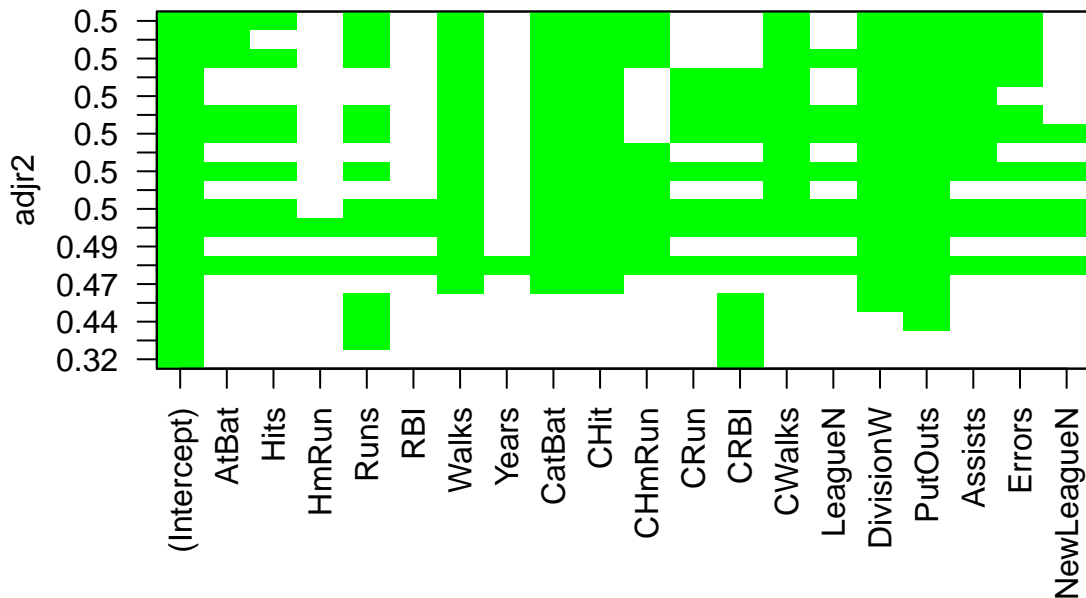
```
## [1] 0.3188503 0.4140632 0.4420427 0.4624180 0.4740707 0.4933275 0.4981264
## [8] 0.4999857 0.5015259 0.5019771 0.5020751 0.5021206 0.5019859 0.5014591
## [15] 0.4999994 0.4985190 0.4967643 0.4948937 0.4928570
```

```
bestsubset_19=which.max(summary_19$adjr2)
bestsubset_19
```

```
## [1] 12
```

Plots for variable selection

```
plot(reg_all_19, scale="adjr2", col="green")
```



```
plot(summary_19$adjr2, xlab="Number of regressors", ylab="adjusted R-square", type="l", col="red")
points(bestsubset_19, summary_19$adjr2[bestsubset_19], col="blue", pch=8)
```

