Project_1_Linear_Regression.R

```
setwd("C:/Sourav/Technology/Trainings/Data Analytics with R, Excel &
Tableau/R_Directory/Files/")

library(readr)

FE2010 <- read_csv("FE2010.csv")

## Parsed with column specification:
## cols(
##   EngDispl = col_double(),
##   NumCyl = col_integer(),
##   FE = col_double(),
##   NumGears = col_integer(),
##   TransLockup = col_integer(),
##   TransCreeperGear = col_integer(),
##   IntakeValvePerCyl = col_integer(),
##   ExhaustValvesPerCyl = col_integer(),
##   VarValveTiming = col_integer(),
##   VarValveLift = col_integer()
## )

View(FE2010) str(FE2010)

## Classes 'tbl_df', 'tbl' and 'data.frame':    1107 obs. of  10 variables:
##  $ EngDispl           : num  4.7 4.7 4.2 4.2 5.2 5.2 2 6 3 3 ...
##  $ NumCyl             : int  8 8 8 8 10 10 4 12 6 6 ...
##  $ FE                 : num  28 25.6 26.8 25 24.8 ...
##  $ NumGears           : int  6 6 6 6 6 6 6 6 6 6 ...
##  $ TransLockup        : int  1 1 1 1 0 0 0 0 1 0 ...
##  $ TransCreeperGear   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ IntakeValvePerCyl  : int  2 2 2 2 2 2 2 2 2 2 ...
##  $ ExhaustValvesPerCyl: int  2 2 2 2 2 2 2 2 2 2 ...
##  $ VarValveTiming     : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ VarValveLift       : int  0 0 0 0 0 0 0 0 1 1 ...
##  - attr(*, "spec")=List of 2
##   ..$ cols   :List of 10
##   .. ..$ EngDispl           : list()
##   .. .. ..- attr(*, "class")= chr  "collector_double" "collector"
##   .. ..$ NumCyl             : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ FE                 : list()
##   .. .. ..- attr(*, "class")= chr  "collector_double" "collector"
##   .. ..$ NumGears           : list()
```

```
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ TransLockup        : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ TransCreeperGear   : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ IntakeValvePerCyl  : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ ExhaustValvesPerCyl: list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ VarValveTiming     : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   .. ..$ VarValveLift       : list()
##   .. .. ..- attr(*, "class")= chr  "collector_integer" "collector"
##   ..$ default: list()
##   .. ..- attr(*, "class")= chr  "collector_guess" "collector"
##   ..- attr(*, "class")= chr "col_spec"
```

```
summary(FE2010)
```

```
##     EngDispl         NumCyl             FE            NumGears
##  Min.   :1.000   Min.   : 2.000   Min.   :17.50   Min.   :1.000
##  1st Qu.:2.400   1st Qu.: 4.000   1st Qu.:29.09   1st Qu.:5.000
##  Median :3.500   Median : 6.000   Median :34.51   Median :6.000
##  Mean   :3.507   Mean   : 5.971   Mean   :34.71   Mean   :5.268
##  3rd Qu.:4.300   3rd Qu.: 8.000   3rd Qu.:39.20   3rd Qu.:6.000
##  Max.   :8.400   Max.   :16.000   Max.   :69.64   Max.   :8.000
##   TransLockup     TransCreeperGear   IntakeValvePerCyl ExhaustValvesPerCyl
##  Min.   :0.0000   Min.   :0.00000   Min.   :0.000     Min.   :0.000
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:2.000     1st Qu.:2.000
##  Median :1.0000   Median :0.00000   Median :2.000     Median :2.000
##  Mean   :0.6802   Mean   :0.04878   Mean   :1.862     Mean   :1.837
##  3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:2.000     3rd Qu.:2.000
##  Max.   :1.0000   Max.   :1.00000   Max.   :3.000     Max.   :2.000
##  VarValveTiming    VarValveLift
##  Min.   :0.0000   Min.   :0.0000
##  1st Qu.:1.0000   1st Qu.:0.0000
##  Median :1.0000   Median :0.0000
##  Mean   :0.8229   Mean   :0.1671
##  3rd Qu.:1.0000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.0000
```

```
library(ggplot2)
library(reshape2)
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
library(e1071)
library(caret)
```

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

```r
library(rpart)
library(C50)
library(party)
```

```
## Loading required package: grid
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: modeltools
```

```
## Loading required package: stats4
```

```
## Loading required package: strucchange
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: sandwich
```

```r
#library(partykit)
library(randomForest)
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin
```

```r
library(ROCR)
```

```
## Loading required package: gplots
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
##
##     lowess
```

```r
library(dplyr)
```

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

```
## The following object is masked from 'package:randomForest':
##
##     combine

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(car)
```

```
## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##     recode

## The following object is masked from 'package:modeltools':
##
##     Predict
```

```
mydata<-FE2010
head(mydata)
```

```
## # A tibble: 6 x 10
##   EngDispl NumCyl    FE NumGears TransLockup TransCreeperGear
##      <dbl>  <int> <dbl>    <int>       <int>             <int>
## 1      4.7      8  28.0        6           1                 0
## 2      4.7      8  25.6        6           1                 0
## 3      4.2      8  26.8        6           1                 0
## 4      4.2      8  25.0        6           1                 0
## 5      5.2     10  24.8        6           0                 0
## 6      5.2     10  23.9        6           0                 0
## # ... with 4 more variables: IntakeValvePerCyl <int>,
## #   ExhaustValvesPerCyl <int>, VarValveTiming <int>, VarValveLift <int>
```

```
summary(mydata)
```

```
##     EngDispl         NumCyl             FE           NumGears
##  Min.   :1.000   Min.   : 2.000   Min.   :17.50   Min.   :1.000
##  1st Qu.:2.400   1st Qu.: 4.000   1st Qu.:29.09   1st Qu.:5.000
##  Median :3.500   Median : 6.000   Median :34.51   Median :6.000
##  Mean   :3.507   Mean   : 5.971   Mean   :34.71   Mean   :5.268
##  3rd Qu.:4.300   3rd Qu.: 8.000   3rd Qu.:39.20   3rd Qu.:6.000
##  Max.   :8.400   Max.   :16.000   Max.   :69.64   Max.   :8.000
##   TransLockup     TransCreeperGear   IntakeValvePerCyl ExhaustValvesPerCyl
```

```
##  Min.   :0.0000   Min.   :0.00000   Min.   :0.000     Min.   :0.000
##  1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:2.000     1st Qu.:2.000
##  Median :1.0000   Median :0.00000   Median :2.000     Median :2.000
##  Mean   :0.6802   Mean   :0.04878   Mean   :1.862     Mean   :1.837
##  3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:2.000     3rd Qu.:2.000
##  Max.   :1.0000   Max.   :1.00000   Max.   :3.000     Max.   :2.000
##  VarValveTiming    VarValveLift
##  Min.   :0.0000   Min.   :0.0000
##  1st Qu.:1.0000   1st Qu.:0.0000
##  Median :1.0000   Median :0.0000
##  Mean   :0.8229   Mean   :0.1671
##  3rd Qu.:1.0000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.0000
```

```
View(mydata)
sapply(mydata, sd)
```

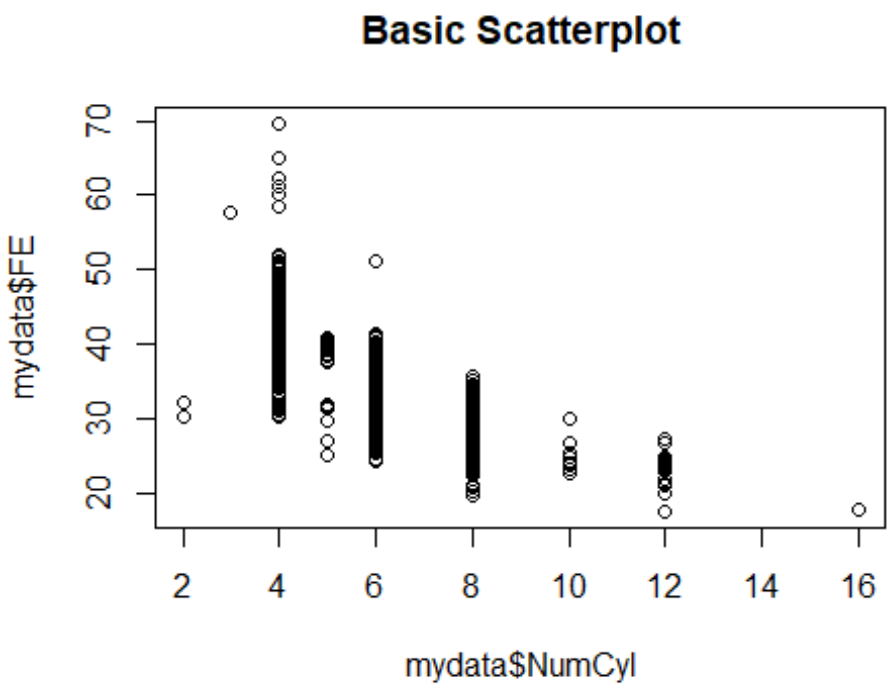
```
##            EngDispl              NumCyl                  FE
##           1.3059051           1.9005745           7.4980326
##            NumGears         TransLockup    TransCreeperGear
##           1.3966238           0.4666032           0.2155062
##   IntakeValvePerCyl ExhaustValvesPerCyl      VarValveTiming
##           0.3530462           0.3740349           0.3818879
##        VarValveLift
##           0.3732501
```

```
cormatrix <- round(cor(mydata), digits = 2 )
cormatrix
```

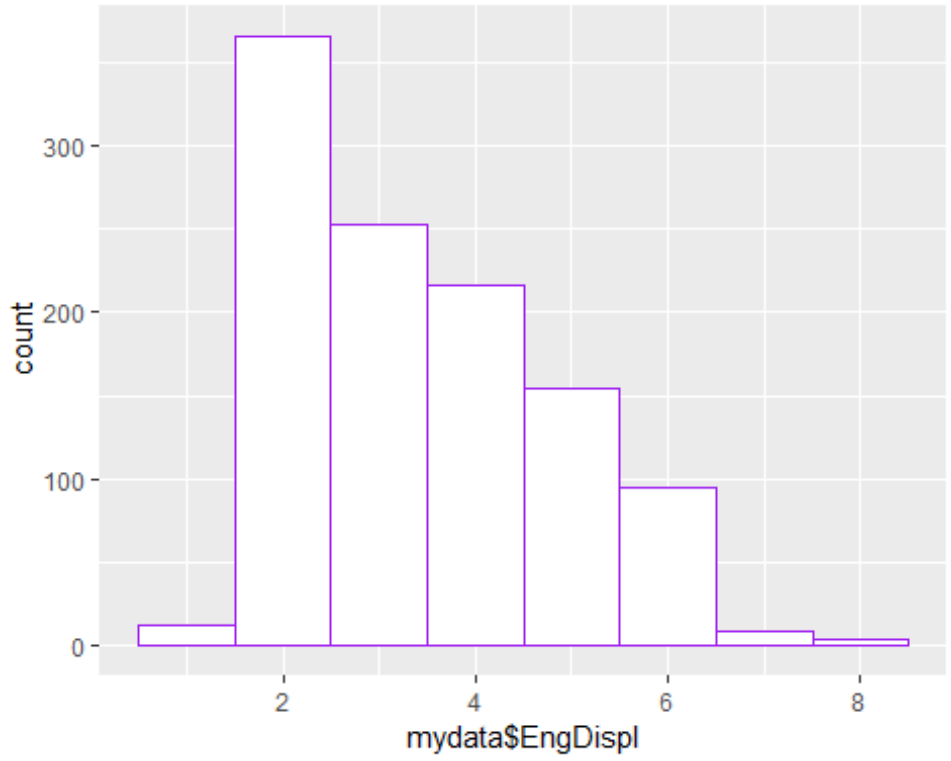
```
##                     EngDispl NumCyl    FE NumGears TransLockup
## EngDispl                1.00    0.91 -0.79     0.21        0.23
## NumCyl                  0.91    1.00 -0.74     0.29        0.21
## FE                     -0.79   -0.74  1.00    -0.21       -0.27
## NumGears                0.21    0.29 -0.21     1.00        0.00
## TransLockup             0.23    0.21 -0.27     0.00        1.00
## TransCreeperGear        0.03    0.03 -0.07     0.04        0.09
## IntakeValvePerCyl      -0.42   -0.25  0.28     0.18       -0.13
## ExhaustValvesPerCyl    -0.48   -0.34  0.34     0.15       -0.16
## VarValveTiming         -0.07    0.01  0.12     0.09       -0.09
## VarValveLift           -0.09   -0.06  0.10     0.13       -0.10
##                     TransCreeperGear IntakeValvePerCyl ExhaustValvesPerCyl
## EngDispl                        0.03             -0.42               -0.48
## NumCyl                          0.03             -0.25               -0.34
## FE                             -0.07              0.28                0.34
## NumGears                        0.04              0.18                0.15
## TransLockup                     0.09             -0.13               -0.16
## TransCreeperGear                1.00             -0.08               -0.17
## IntakeValvePerCyl              -0.08              1.00                0.91
## ExhaustValvesPerCyl            -0.17              0.91                1.00
## VarValveTiming                 -0.24              0.24                0.28
## VarValveLift                   -0.10              0.15                0.18
```

```
##                       VarValveTiming VarValveLift
## EngDispl                       -0.07        -0.09
## NumCyl                          0.01        -0.06
## FE                              0.12         0.10
## NumGears                        0.09         0.13
## TransLockup                    -0.09        -0.10
## TransCreeperGear               -0.24        -0.10
## IntakeValvePerCyl               0.24         0.15
## ExhaustValvesPerCyl             0.28         0.18
## VarValveTiming                  1.00         0.06
## VarValveLift                    0.06         1.00
```

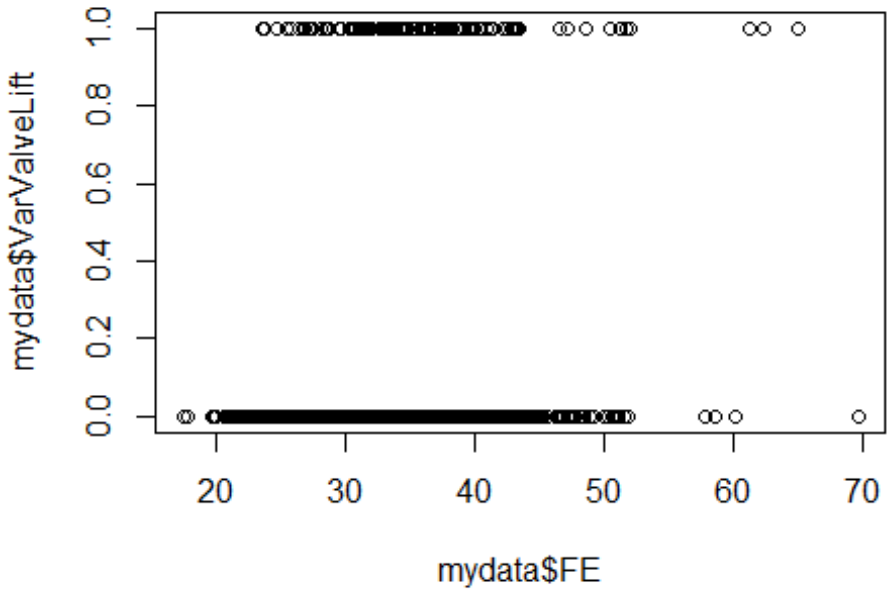
```
plot.new()
plot(mydata$FE ~mydata$NumCyl)
title('Basic Scatterplot')
```



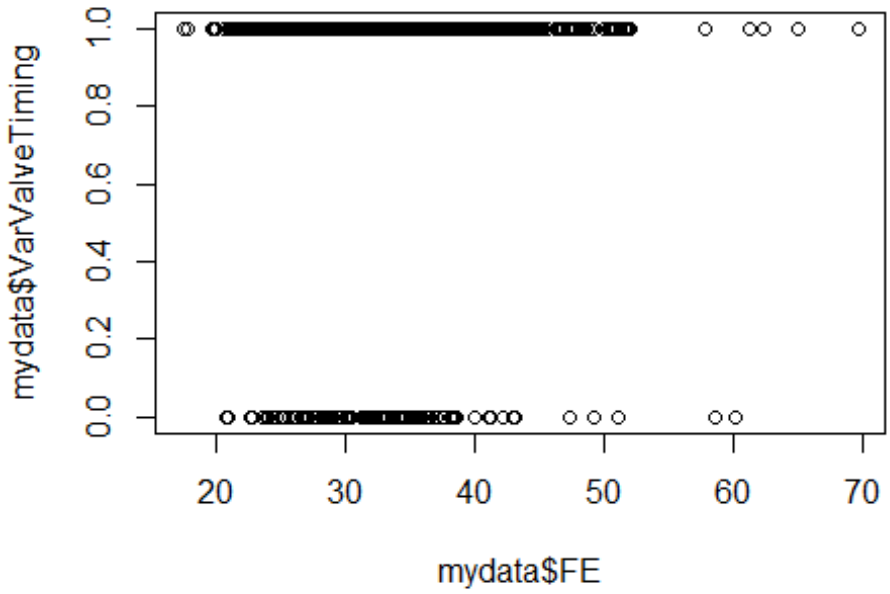

```
ggplot(mydata, aes(x=mydata$EngDispl)) + geom_histogram(binwidth = 1, fill =
"white", color = "purple")
```



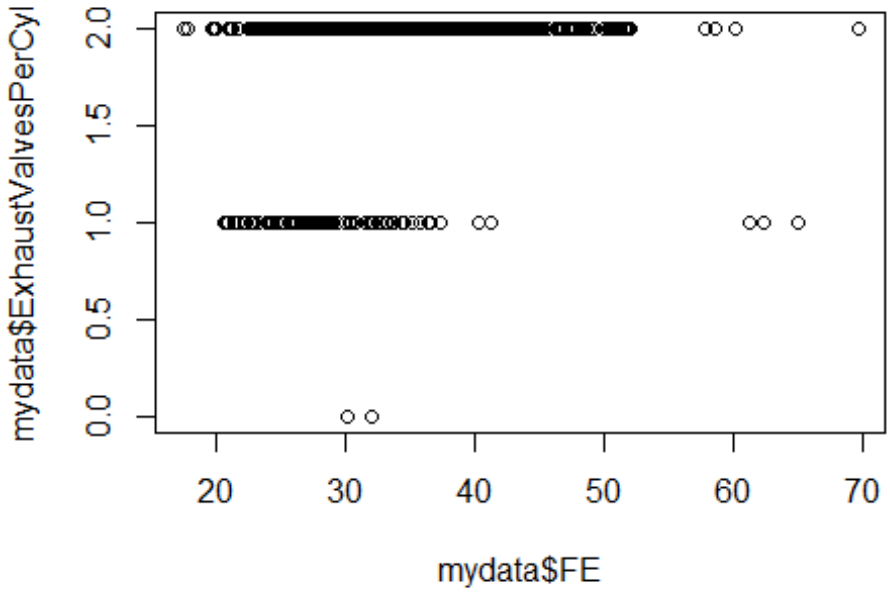

```
plot(mydata$FE,mydata$VarValveLift)
```



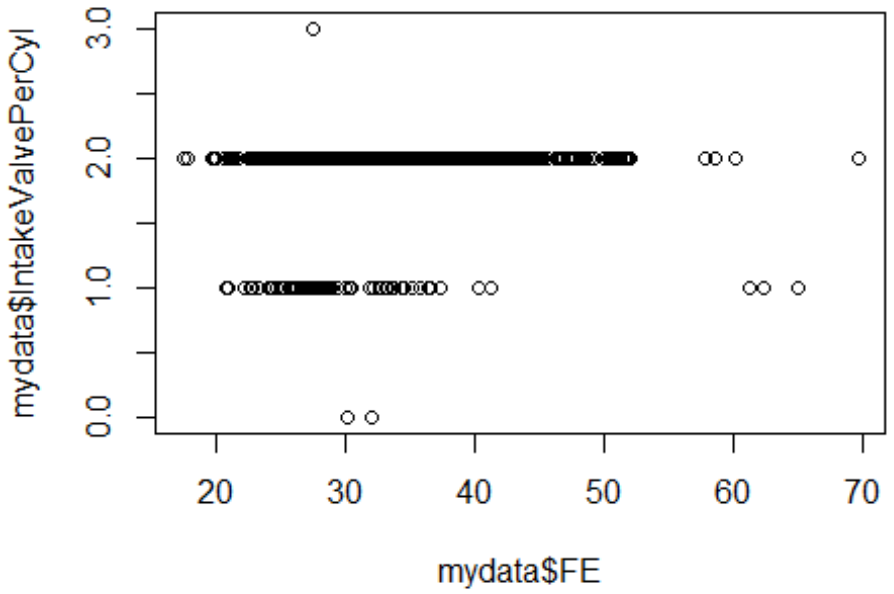

```
plot(mydata$FE,mydata$VarValveTiming)
```



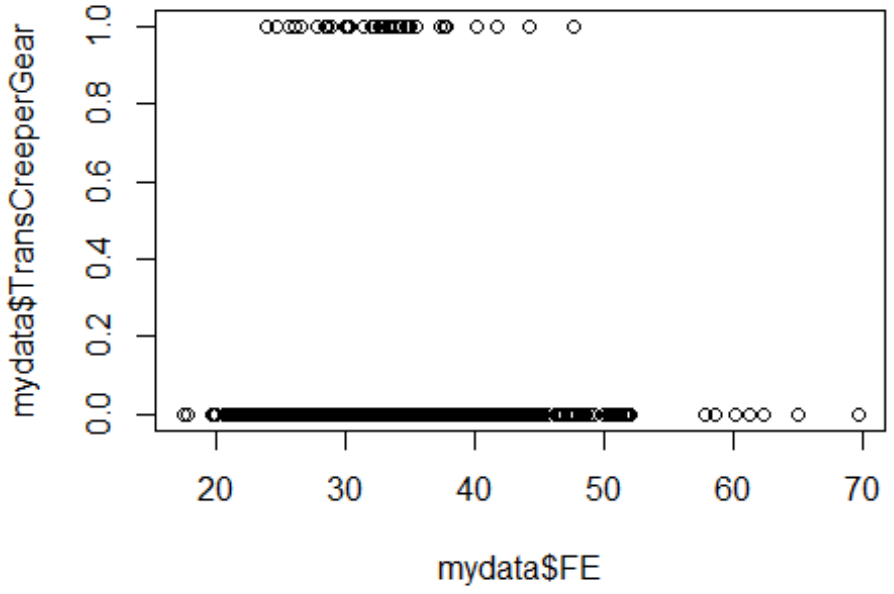

```
plot(mydata$FE,mydata$ExhaustValvesPerCyl)
```



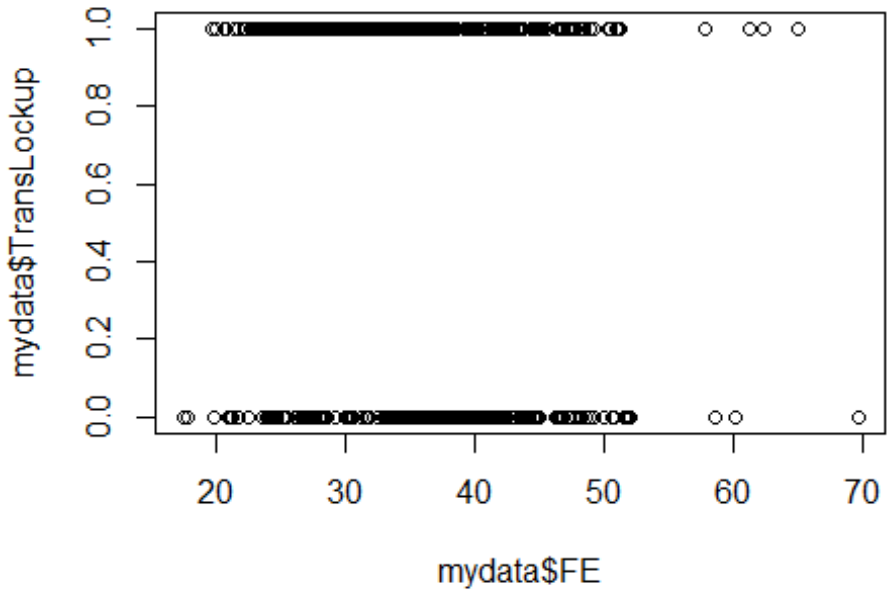

```
plot(mydata$FE,mydata$IntakeValvePerCyl)
```



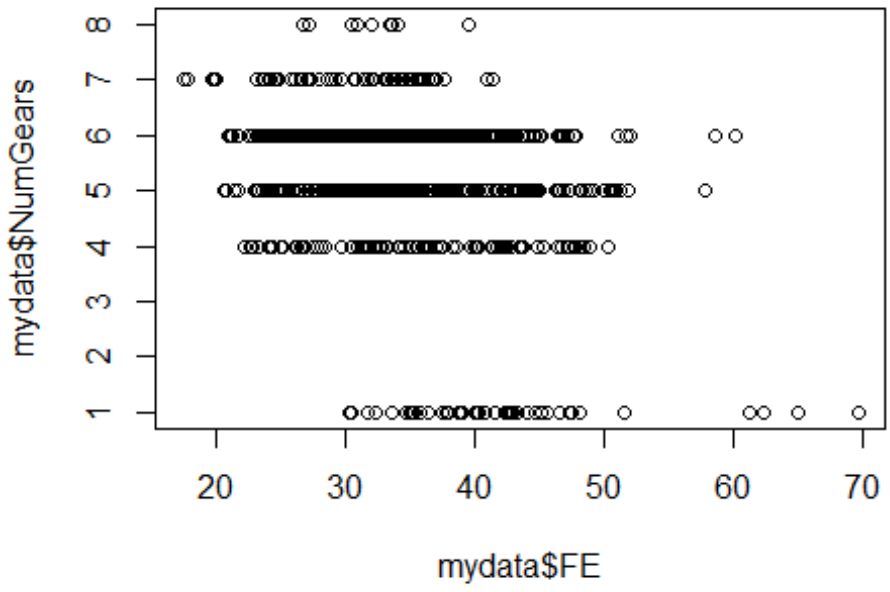

```
plot(mydata$FE,mydata$TransCreeperGear)
```



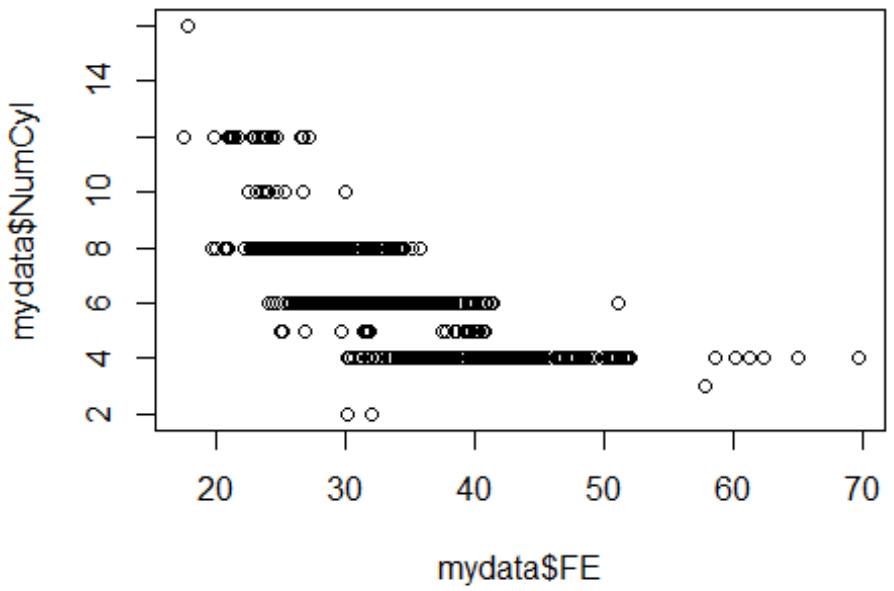

```
plot(mydata$FE,mydata$TransLockup)
```



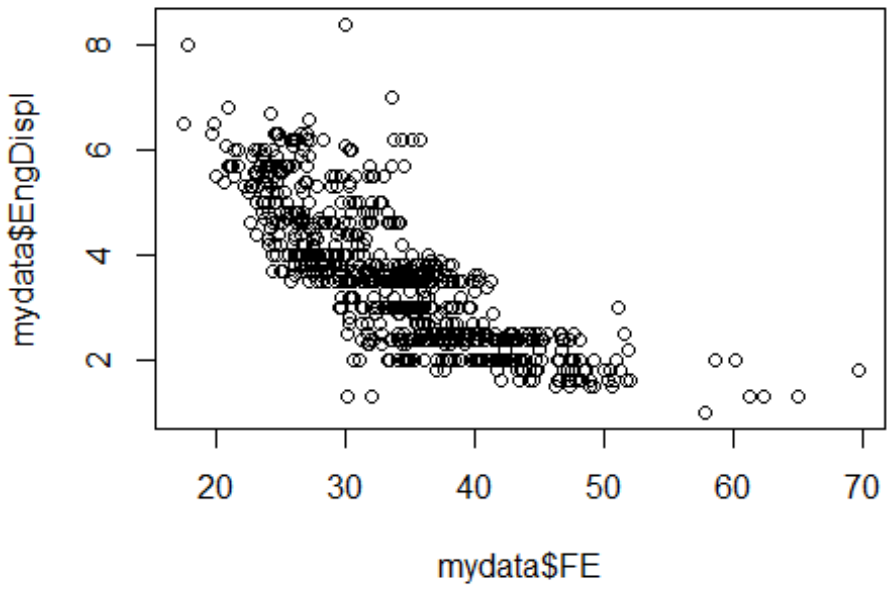

```r
plot(mydata$FE,mydata$NumGears)
```




```r
plot(mydata$FE,mydata$NumCyl)
```




```
plot(mydata$FE,mydata$EngDispl)
```




```
cor(mydata$FE,mydata$EngDispl)
```

```
## [1] -0.7873938
```

```
cor(mydata$FE,mydata$VarValveLift)
```

```
## [1] 0.09621127
```

```
cor(mydata$FE,mydata$VarValveTiming)
```

```
## [1] 0.1249528
```

```
cor(mydata$FE,mydata$ExhaustValvesPerCyl)
```

```
## [1] 0.3356529
```

```
cor(mydata$FE,mydata$IntakeValvePerCyl)
```

```
## [1] 0.280344
```

```
cor(mydata$FE,mydata$TransCreeperGear)
```

```
## [1] -0.06962168
```

```
cor(mydata$FE,mydata$TransLockup)
```

```
## [1] -0.2719389
```

```
cor(mydata$FE,mydata$NumGears)
```

```
## [1] -0.2112849
```

```
cor(mydata$FE,mydata$NumCyl)
```

```
## [1] -0.740218
```

```
mod=lm(mydata$FE~mydata$EngDispl)
mod
```

```
##
## Call:
## lm(formula = mydata$FE ~ mydata$EngDispl)
##
## Coefficients:
##     (Intercept)  mydata$EngDispl
##          50.563           -4.521
```

```
summary(mod)
```

```
##
## Call:
## lm(formula = mydata$FE ~ mydata$EngDispl)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -14.486  -3.192  -0.365   2.671  27.215
```

```
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      50.5632     0.3985  126.89   <2e-16 ***
## mydata$EngDispl  -4.5209     0.1065  -42.46   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.624 on 1105 degrees of freedom
## Multiple R-squared:   0.62,  Adjusted R-squared:  0.6196
## F-statistic:  1803 on 1 and 1105 DF,  p-value: < 2.2e-16
```

```
predict(mod)
```

```
##        1        2        3        4        5        6        7        8
## 29.31486 29.31486 31.57533 31.57533 27.05440 27.05440 41.52137 23.43765
##        9       10       11       12       13       14       15       16
## 37.00044 37.00044 37.00044 37.00044 14.39580 22.53347 22.53347 22.53347
##       17       18       19       20       21       22       23       24
## 18.91672 12.58742 12.58742 30.21905 24.79393 24.79393 27.05440 27.05440
##       25       26       27       28       29       30       31       32
## 27.05440 27.05440 21.17719 21.17719 21.17719 21.17719 21.17719 42.42556
##       33       34       35       36       37       38       39       40
## 42.42556 41.52137 41.52137 41.52137 25.69812 37.00044 34.73998 34.73998
##       41       42       43       44       45       46       47       48
## 34.73998 25.69812 46.04230 46.04230 33.83579 33.83579 33.83579 33.83579
##       49       50       51       52       53       54       55       56
## 41.52137 41.52137 39.71300 39.71300 33.38370 33.38370 37.45254 37.45254
##       57       58       59       60       61       62       63       64
## 35.19207 35.19207 37.45254 37.45254 35.19207 35.19207 41.52137 41.52137
##       65       66       67       68       69       70       71       72
## 39.71300 39.71300 31.57533 23.88975 23.88975 23.88975 23.88975 31.12323
##       73       74       75       76       77       78       79       80
## 27.95858 27.95858 27.95858 31.12323 34.73998 43.32974 43.32974 43.32974
##       81       82       83       84       85       86       87       88
## 43.32974 43.32974 43.32974 43.32974 43.32974 43.32974 43.32974 39.71300
##       89       90       91       92       93       94       95       96
## 33.38370 34.28788 34.28788 34.28788 34.28788 34.28788 34.28788 34.28788
##       97       98       99      100      101      102      103      104
## 34.28788 33.38370 33.38370 33.38370 33.38370 33.38370 33.38370 34.28788
##      105      106      107      108      109      110      111      112
## 34.28788 33.38370 33.38370 33.38370 33.38370 33.38370 33.38370 33.38370
##      113      114      115      116      117      118      119      120
## 33.38370 33.38370 33.38370 39.26091 23.88975 41.52137 41.52137 41.52137
##      121      122      123      124      125      126      127      128
## 41.52137 36.09626 31.57533 31.57533 37.00044 41.52137 23.43765 37.00044
##      129      130      131      132      133      134      135      136
## 37.00044 37.00044 37.00044 37.00044 37.00044 37.00044 37.00044 37.00044
##      137      138      139      140      141      142      143      144
## 37.00044 37.00044 37.00044 37.00044 37.00044 37.00044 37.00044 37.00044
```

```
##      145      146      147      148      149      150      151      152
## 37.00044 37.00044 37.00044 28.86277 28.86277 28.86277 28.86277 32.47951
##      153      154      155      156      157      158      159      160
## 32.47951 32.47951 32.47951 27.95858 27.95858 27.95858 27.95858 43.32974
##      161      162      163      164      165      166      167      168
## 43.32974 40.61719 40.61719 41.52137 40.61719 32.47951 32.47951 29.76696
##      169      170      171      172      173      174      175      176
## 29.76696 26.15021 42.42556 42.42556 42.42556 41.52137 41.52137 41.52137
##      177      178      179      180      181      182      183      184
## 33.38370 33.38370 33.83579 33.83579 33.83579 33.83579 33.83579 39.26091
##      185      186      187      188      189      190      191      192
## 39.26091 39.26091 34.73998 27.95858 31.57533 29.31486 29.31486 44.68602
##      193      194      195      196      197      198      199      200
## 44.68602 34.73998 25.69812 43.32974 43.32974 43.32974 43.32974 43.32974
##      201      202      203      204      205      206      207      208
## 39.71300 39.71300 33.38370 33.38370 39.26091 39.26091 34.73998 34.73998
##      209      210      211      212      213      214      215      216
## 33.38370 40.61719 40.61719 40.61719 40.61719 40.61719 29.76696 29.76696
##      217      218      219      220      221      222      223      224
## 41.52137 41.52137 43.32974 43.32974 39.71300 39.71300 42.42556 42.42556
##      225      226      227      228      229      230      231      232
## 43.78184 43.78184 41.52137 41.52137 39.26091 39.26091 39.26091 39.26091
##      233      234      235      236      237      238      239      240
## 39.71300 39.71300 34.73998 41.52137 41.52137 41.52137 37.00044 37.00044
##      241      242      243      244      245      246      247      248
## 19.82091 19.82091 23.43765 37.00044 37.00044 37.00044 37.00044 37.00044
##      249      250      251      252      253      254      255      256
## 37.00044 37.00044 37.00044 37.00044 32.47951 32.47951 43.32974 43.32974
##      257      258      259      260      261      262      263      264
## 34.28788 34.28788 22.53347 22.53347 40.61719 40.61719 40.61719 39.71300
##      265      266      267      268      269      270      271      272
## 38.35672 34.73998 34.73998 24.79393 24.79393 22.98556 22.98556 41.52137
##      273      274      275      276      277      278      279      280
## 41.52137 39.71300 39.71300 34.73998 34.73998 44.68602 44.68602 44.68602
##      281      282      283      284      285      286      287      288
## 43.32974 43.32974 43.32974 41.52137 41.52137 41.52137 39.71300 39.71300
##      289      290      291      292      293      294      295      296
## 43.32974 43.32974 34.73998 39.71300 41.52137 41.52137 39.26091 39.26091
##      297      298      299      300      301      302      303      304
## 37.00044 37.00044 34.73998 37.00044 34.73998 34.73998 22.08138 25.69812
##      305      306      307      308      309      310      311      312
## 25.69812 22.08138 23.43765 25.69812 22.08138 41.52137 41.52137 41.52137
##      313      314      315      316      317      318      319      320
## 39.71300 39.71300 41.52137 41.52137 43.32974 43.32974 39.71300 39.71300
##      321      322      323      324      325      326      327      328
## 39.71300 39.71300 34.73998 34.73998 34.28788 34.28788 20.27300 20.27300
##      329      330      331      332      333      334      335      336
## 41.52137 41.52137 41.52137 41.52137 41.52137 39.26091 39.26091 39.26091
##      337      338      339      340      341      342      343      344
## 39.26091 39.71300 39.71300 39.71300 39.71300 39.71300 39.71300 41.52137
```

```
##      345      346      347      348      349      350      351      352
## 41.52137 41.52137 41.52137 42.42556 42.42556 39.71300 39.71300 41.52137
##      353      354      355      356      357      358      359      360
## 41.52137 34.28788 34.28788 41.52137 41.52137 39.26091 39.26091 41.52137
##      361      362      363      364      365      366      367      368
## 41.52137 41.52137 41.52137 41.52137 41.52137 39.26091 39.26091 39.71300
##      369      370      371      372      373      374      375      376
## 39.71300 39.26091 39.26091 39.26091 39.26091 39.71300 39.71300 39.26091
##      377      378      379      380      381      382      383      384
## 39.26091 33.83579 34.73998 33.83579 33.83579 36.09626 37.00044 31.57533
##      385      386      387      388      389      390      391      392
## 31.57533 27.05440 23.43765 37.00044 37.00044 37.00044 37.00044 37.00044
##      393      394      395      396      397      398      399      400
## 37.00044 37.00044 37.00044 28.86277 28.86277 27.95858 27.95858 39.71300
##      401      402      403      404      405      406      407      408
## 37.00044 34.28788 37.00044 37.00044 37.00044 34.28788 34.28788 22.53347
##      409      410      411      412      413      414      415      416
## 22.53347 37.00044 34.28788 34.28788 29.76696 34.28788 29.76696 39.71300
##      417      418      419      420      421      422      423      424
## 39.71300 39.71300 39.71300 34.73998 34.73998 34.28788 39.71300 39.71300
##      425      426      427      428      429      430      431      432
## 38.35672 34.73998 39.71300 38.35672 34.73998 24.79393 24.79393 34.73998
##      433      434      435      436      437      438      439      440
## 37.00044 39.26091 39.26091 37.00044 34.73998 39.26091 39.26091 39.26091
##      441      442      443      444      445      446      447      448
## 41.52137 41.52137 41.52137 33.83579 33.83579 33.83579 31.57533 27.95858
##      449      450      451      452      453      454      455      456
## 27.95858 39.71300 39.71300 38.35672 34.73998 34.73998 34.73998 29.76696
##      457      458      459      460      461      462      463      464
## 29.76696 29.76696 29.76696 29.76696 27.95858 37.00044 39.26091 39.26091
##      465      466      467      468      469      470      471      472
## 37.00044 39.26091 39.26091 39.26091 34.73998 34.73998 39.26091 39.26091
##      473      474      475      476      477      478      479      480
## 33.83579 40.16509 34.73998 34.73998 25.69812 25.69812 22.08138 39.71300
##      481      482      483      484      485      486      487      488
## 39.26091 34.73998 34.73998 39.26091 34.73998 41.52137 41.52137 39.26091
##      489      490      491      492      493      494      495      496
## 39.26091 43.32974 43.32974 42.42556 42.42556 42.42556 20.27300 37.90463
##      497      498      499      500      501      502      503      504
## 39.71300 39.71300 34.28788 39.26091 39.26091 39.26091 34.28788 39.26091
##      505      506      507      508      509      510      511      512
## 39.26091 34.73998 39.71300 42.42556 41.52137 37.00044 30.67114 36.09626
##      513      514      515      516      517      518      519      520
## 31.57533 37.00044 30.67114 30.67114 30.67114 30.67114 30.67114 30.67114
##      521      522      523      524      525      526      527      528
## 23.43765 32.93161 32.93161 29.76696 29.76696 29.76696 29.76696 29.76696
##      529      530      531      532      533      534      535      536
## 34.73998 34.73998 32.93161 34.73998 24.79393 38.35672 34.73998 24.79393
##      537      538      539      540      541      542      543      544
## 22.98556 38.35672 34.73998 24.79393 22.98556 34.73998 24.79393 29.76696
```

```
##      545      546      547      548      549      550      551      552
## 34.73998 34.73998 34.73998 34.73998 39.71300 39.71300 34.73998 35.64416
##      553      554      555      556      557      558      559      560
## 33.38370 33.38370 29.76696 39.71300 39.71300 35.64416 34.73998 34.73998
##      561      562      563      564      565      566      567      568
## 30.21905 30.21905 27.95858 27.95858 27.95858 29.76696 34.73998 34.73998
##      569      570      571      572      573      574      575      576
## 34.73998 29.76696 31.57533 29.31486 25.69812 23.43765 23.43765 25.69812
##      577      578      579      580      581      582      583      584
## 25.69812 25.69812 22.08138 23.43765 34.73998 28.86277 28.86277 28.86277
##      585      586      587      588      589      590      591      592
## 20.72510 20.27300 34.73998 41.52137 41.52137 41.52137 41.52137 41.52137
##      593      594      595      596      597      598      599      600
## 37.00044 37.00044 37.00044 37.00044 37.00044 34.28788 37.00044 34.28788
##      601      602      603      604      605      606      607      608
## 41.52137 41.52137 39.71300 39.71300 43.78184 43.78184 43.78184 41.52137
##      609      610      611      612      613      614      615      616
## 41.52137 34.73998 34.73998 43.32974 43.32974 41.52137 41.52137 41.52137
##      617      618      619      620      621      622      623      624
## 39.71300 39.71300 42.42556 42.42556 42.42556 42.42556 39.71300 39.71300
##      625      626      627      628      629      630      631      632
## 39.71300 41.52137 41.52137 41.52137 41.52137 41.52137 39.26091 39.26091
##      633      634      635      636      637      638      639      640
## 39.26091 39.26091 39.26091 41.52137 41.52137 41.52137 41.52137 42.42556
##      641      642      643      644      645      646      647      648
## 42.42556 39.71300 39.71300 39.71300 39.71300 39.71300 41.52137 41.52137
##      649      650      651      652      653      654      655      656
## 39.26091 39.26091 39.26091 39.26091 39.71300 39.26091 39.26091 37.00044
##      657      658      659      660      661      662      663      664
## 37.00044 37.00044 39.71300 38.35672 41.52137 36.09626 37.45254 37.45254
##      665      666      667      668      669      670      671      672
## 33.83579 26.60230 33.83579 37.45254 37.45254 33.83579 26.60230 40.16509
##      673      674      675      676      677      678      679      680
## 40.16509 32.47951 32.47951 37.45254 37.45254 33.83579 26.60230 33.83579
##      681      682      683      684      685      686      687      688
## 37.45254 37.45254 33.83579 26.60230 39.26091 39.26091 32.47951 32.47951
##      689      690      691      692      693      694      695      696
## 39.26091 39.26091 32.47951 38.35672 38.35672 32.47951 32.47951 37.45254
##      697      698      699      700      701      702      703      704
## 37.45254 33.83579 26.60230 33.83579 33.83579 26.60230 32.47951 32.47951
##      705      706      707      708      709      710      711      712
## 37.45254 37.45254 33.83579 26.60230 33.83579 33.83579 26.60230 32.47951
##      713      714      715      716      717      718      719      720
## 32.47951 32.47951 38.35672 32.47951 32.47951 31.12323 28.86277 26.60230
##      721      722      723      724      725      726      727      728
## 22.53347 23.43765 26.60230 33.83579 29.31486 33.83579 29.31486 24.79393
##      729      730      731      732      733      734      735      736
## 32.47951 29.76696 26.15021 29.76696 29.76696 31.12323 28.86277 26.60230
##      737      738      739      740      741      742      743      744
## 22.53347 23.43765 26.60230 25.24603 25.24603 32.47951 29.76696 24.79393
```

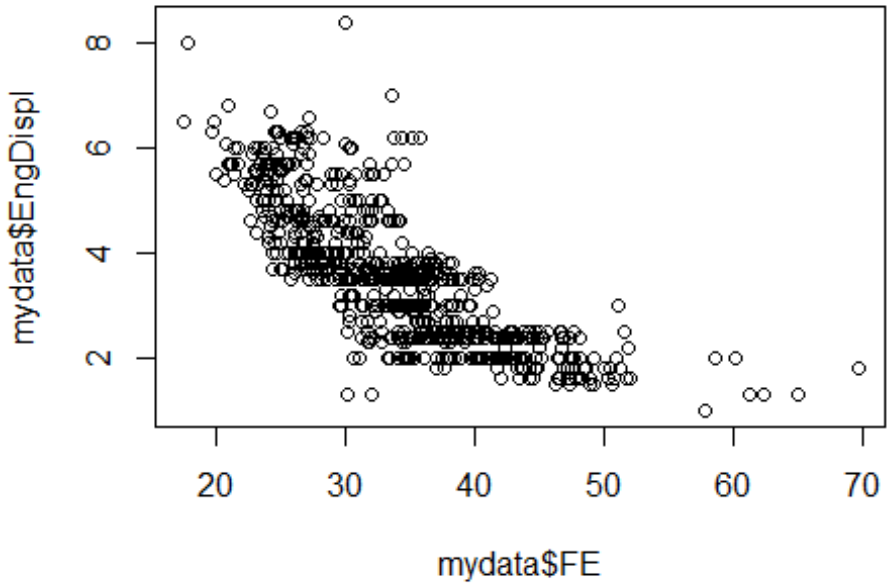
```
##      745      746      747      748      749      750      751      752
## 31.12323 28.86277 26.60230 22.53347 23.43765 33.83579 29.31486 29.31486
##      753      754      755      756      757      758      759      760
## 24.79393 32.47951 29.76696 26.15021 29.76696 29.76696 31.12323 28.86277
##      761      762      763      764      765      766      767      768
## 26.60230 22.53347 23.43765 22.53347 34.73998 33.83579 33.83579 26.60230
##      769      770      771      772      773      774      775      776
## 25.24603 25.24603 29.76696 24.79393 24.79393 31.12323 26.60230 26.60230
##      777      778      779      780      781      782      783      784
## 26.60230 26.60230 31.12323 26.60230 26.60230 26.60230 26.60230 26.60230
##      785      786      787      788      789      790      791      792
## 26.60230 26.60230 26.60230 41.52137 35.64416 33.38370 32.47951 35.64416
##      793      794      795      796      797      798      799      800
## 33.38370 32.47951 34.73998 34.73998 33.38370 33.38370 40.16509 40.16509
##      801      802      803      804      805      806      807      808
## 34.73998 33.38370 32.47951 34.73998 40.16509 34.28788 22.53347 23.43765
##      809      810      811      812      813      814      815      816
## 22.53347 37.00044 26.60230 22.53347 26.60230 22.53347 26.60230 23.43765
##      817      818      819      820      821      822      823      824
## 39.71300 37.00044 41.52137 41.52137 40.61719 40.61719 39.71300 39.71300
##      825      826      827      828      829      830      831      832
## 40.61719 40.61719 39.71300 39.71300 34.28788 39.71300 39.71300 34.73998
##      833      834      835      836      837      838      839      840
## 33.83579 32.47951 34.73998 39.26091 39.26091 37.00044 39.26091 26.15021
##      841      842      843      844      845      846      847      848
## 32.47951 29.76696 34.73998 34.28788 26.60230 22.53347 23.43765 26.60230
##      849      850      851      852      853      854      855      856
## 22.53347 22.53347 39.71300 37.00044 34.73998 39.71300 39.71300 39.71300
##      857      858      859      860      861      862      863      864
## 34.73998 39.71300 39.71300 34.73998 39.71300 39.71300 33.38370 34.73998
##      865      866      867      868      869      870      871      872
## 25.24603 33.83579 24.79393 41.52137 41.52137 39.71300 39.71300 33.83579
##      873      874      875      876      877      878      879      880
## 24.79393 33.83579 41.52137 41.52137 39.71300 39.71300 33.38370 33.38370
##      881      882      883      884      885      886      887      888
## 29.76696 41.52137 41.52137 38.35672 34.73998 34.73998 34.73998 34.73998
##      889      890      891      892      893      894      895      896
## 26.15021 40.16509 39.26091 33.83579 39.26091 39.26091 37.00044 39.26091
##      897      898      899      900      901      902      903      904
## 34.73998 34.73998 39.26091 37.00044 39.26091 32.47951 29.76696 39.71300
##      905      906      907      908      909      910      911      912
## 37.00044 33.38370 25.24603 25.24603 34.73998 32.47951 32.47951 39.26091
##      913      914      915      916      917      918      919      920
## 32.47951 32.47951 34.28788 39.71300 34.28788 34.28788 39.71300 39.71300
##      921      922      923      924      925      926      927      928
## 39.71300 36.09626 38.35672 32.47951 32.47951 38.35672 34.73998 39.26091
##      929      930      931      932      933      934      935      936
## 34.73998 29.76696 24.79393 38.35672 34.73998 41.52137 41.52137 36.09626
##      937      938      939      940      941      942      943      944
## 36.09626 34.73998 40.16509 33.83579 36.09626 37.00044 34.28788 31.57533
```

```
##      945      946      947      948      949      950      951      952
## 30.67114 37.00044 37.00044 37.00044 37.00044 28.86277 30.67114 37.00044
##      953      954      955      956      957      958      959      960
## 30.67114 30.67114 34.28788 22.53347 37.90463 37.00044 39.71300 37.00044
##      961      962      963      964      965      966      967      968
## 26.60230 23.43765 34.28788 34.73998 33.83579 32.47951 34.73998 39.26091
##      969      970      971      972      973      974      975      976
## 37.00044 39.26091 26.15021 32.47951 29.76696 34.73998 34.73998 34.28788
##      977      978      979      980      981      982      983      984
## 26.60230 23.43765 22.53347 39.71300 37.00044 34.73998 39.71300 39.71300
##      985      986      987      988      989      990      991      992
## 39.71300 34.73998 33.83579 33.83579 26.60230 39.71300 34.73998 39.71300
##      993      994      995      996      997      998      999     1000
## 39.71300 33.38370 34.73998 27.95858 25.24603 33.83579 24.79393 39.71300
##     1001     1002     1003     1004     1005     1006     1007     1008
## 39.71300 33.83579 24.79393 22.98556 33.83579 39.71300 39.71300 39.71300
##     1009     1010     1011     1012     1013     1014     1015     1016
## 33.38370 33.38370 33.38370 29.76696 41.52137 38.35672 36.09626 27.95858
##     1017     1018     1019     1020     1021     1022     1023     1024
## 27.95858 27.95858 27.95858 27.95858 29.76696 24.79393 34.73998 34.73998
##     1025     1026     1027     1028     1029     1030     1031     1032
## 34.73998 34.73998 34.73998 40.16509 33.83579 39.26091 37.00044 39.26091
##     1033     1034     1035     1036     1037     1038     1039     1040
## 26.15021 25.69812 37.00044 29.31486 25.69812 34.73998 34.73998 37.00044
##     1041     1042     1043     1044     1045     1046     1047     1048
## 25.69812 22.08138 34.73998 34.73998 37.00044 39.26091 37.00044 39.26091
##     1049     1050     1051     1052     1053     1054     1055     1056
## 32.47951 29.76696 39.71300 37.00044 33.38370 25.24603 25.24603 34.73998
##     1057     1058     1059     1060     1061     1062     1063     1064
## 32.47951 25.24603 39.26091 32.47951 32.47951 34.28788 34.28788 28.86277
##     1065     1066     1067     1068     1069     1070     1071     1072
## 28.86277 28.86277 28.86277 28.86277 28.86277 28.86277 34.28788 34.73998
##     1073     1074     1075     1076     1077     1078     1079     1080
## 34.28788 34.28788 39.26091 39.26091 39.26091 39.26091 39.26091 34.28788
##     1081     1082     1083     1084     1085     1086     1087     1088
## 34.28788 39.71300 39.71300 36.09626 32.47951 32.47951 32.47951 32.47951
##     1089     1090     1091     1092     1093     1094     1095     1096
## 34.73998 35.64416 24.79393 39.26091 34.73998 29.76696 24.79393 24.79393
##     1097     1098     1099     1100     1101     1102     1103     1104
## 38.35672 34.73998 41.52137 37.00044 34.28788 37.00044 36.09626 37.00044
##     1105     1106     1107
## 36.09626 36.09626 30.67114
```

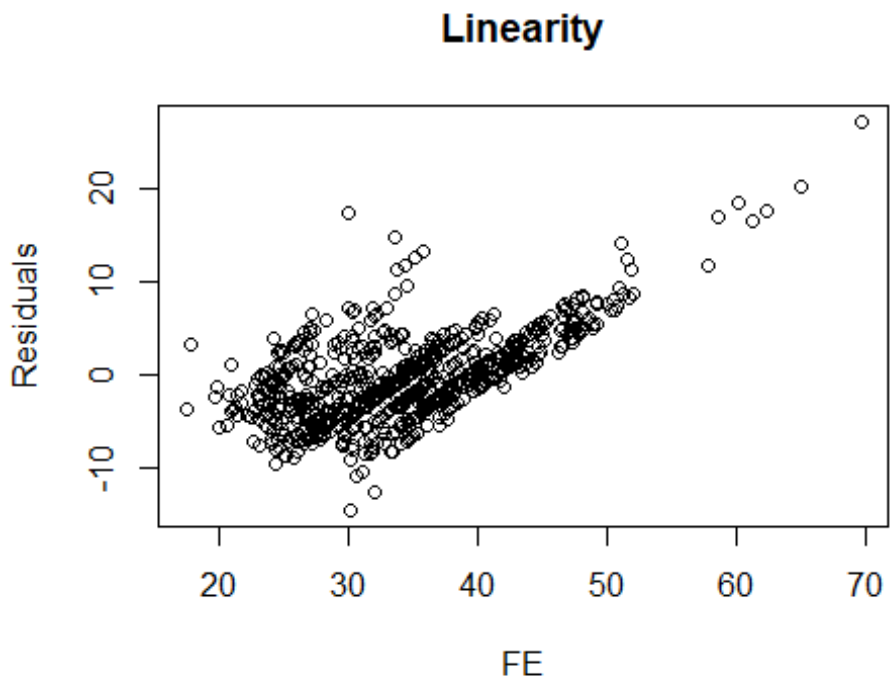
```
pred=predict(mod)
mydata$predicted=NA
mydata$predicted=pred
mydata$error=mod$residuals
library(car)
dwt(mod)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1         0.55367       0.891743       0
##  Alternative hypothesis: rho != 0
```

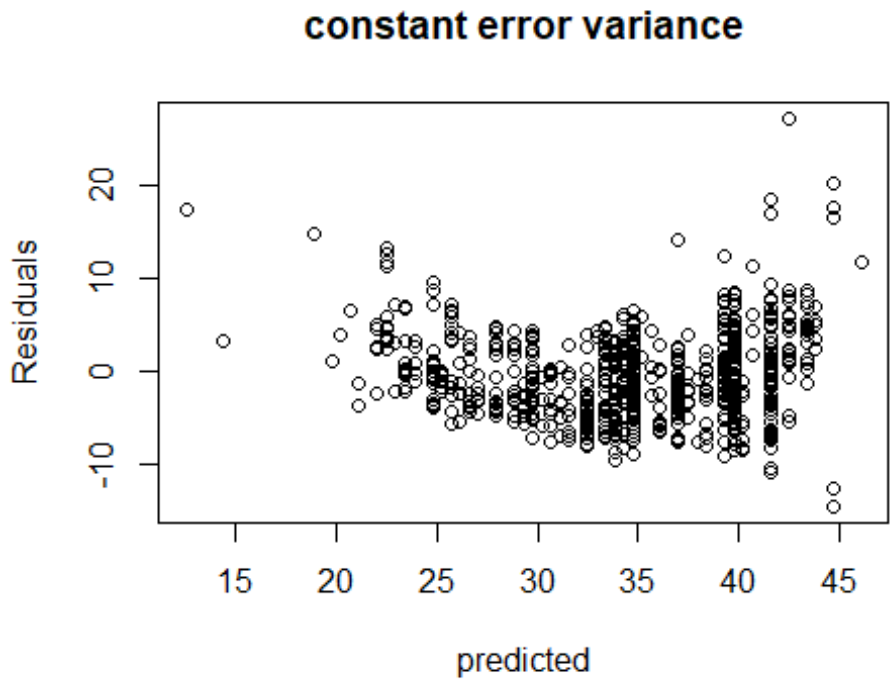
```
plot(mydata$FE,mydata$EngDispl,abline(lm(mydata$FE~mydata$EngDispl),col="oran
ge"))
```



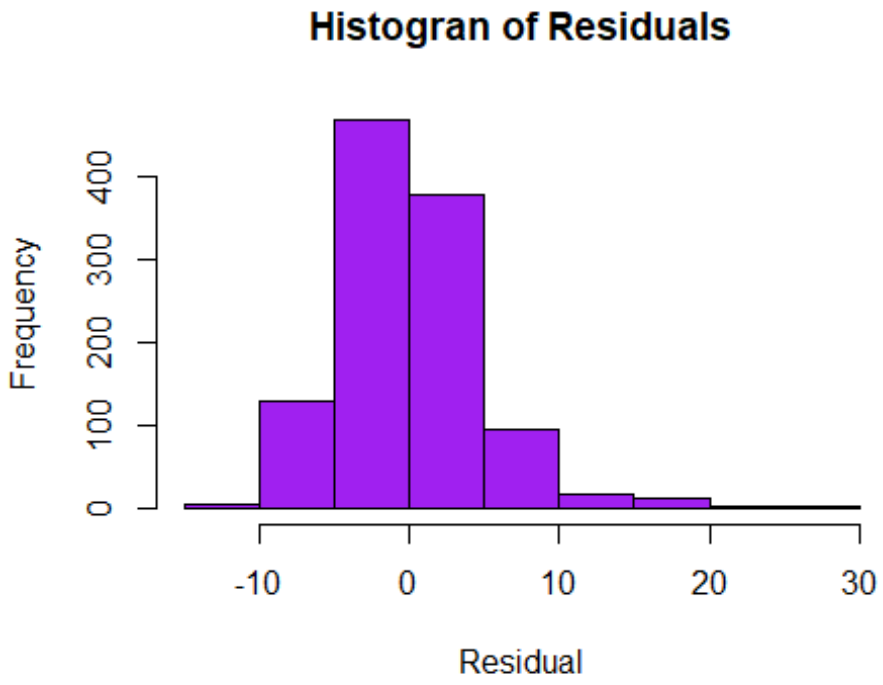

```
#Assumption1 Linearity
plot(mydata$FE,mydata$error,xlab="FE",ylab="Residuals",main="Linearity")
```




```r
#Assumption 2 constant error variance
plot(mydata$predicted,mydata$error, xlab="predicted",
ylab="Residuals",main="constant error variance")
```



```
#Assumption 3 constant error variance
#plot(mydata$observation.no,mydata$error,
xlab="observation.no",ylab="Residuals",main="independence of error")
#Assumption 4: Normality
hist(mydata$error, xlab="Residual", main="Histogran of
Residuals",col="purple")
```




```
#Save all newly inserted variables like predicted, error along with original
variables in a new file
head(mydata)
```

```
## # A tibble: 6 x 12
##   EngDispl NumCyl    FE NumGears TransLockup TransCreeperGear
##      <dbl>  <int> <dbl>    <int>       <int>             <int>
## 1      4.7      8  28.0        6           1                 0
## 2      4.7      8  25.6        6           1                 0
## 3      4.2      8  26.8        6           1                 0
## 4      4.2      8  25.0        6           1                 0
## 5      5.2     10  24.8        6           0                 0
## 6      5.2     10  23.9        6           0                 0
## # ... with 6 more variables: IntakeValvePerCyl <int>,
## #   ExhaustValvesPerCyl <int>, VarValveTiming <int>, VarValveLift <int>,
## #   predicted <dbl>, error <dbl>
```

```
FE2010new<-mydata
head(FE2010new)
```

```
## # A tibble: 6 x 12
##   EngDispl NumCyl    FE NumGears TransLockup TransCreeperGear
##      <dbl>  <int> <dbl>    <int>       <int>             <int>
## 1      4.7      8  28.0        6           1                 0
## 2      4.7      8  25.6        6           1                 0
## 3      4.2      8  26.8        6           1                 0
## 4      4.2      8  25.0        6           1                 0
## 5      5.2     10  24.8        6           0                 0
## 6      5.2     10  23.9        6           0                 0
## # ... with 6 more variables: IntakeValvePerCyl <int>,
## #   ExhaustValvesPerCyl <int>, VarValveTiming <int>, VarValveLift <int>,
## #   predicted <dbl>, error <dbl>
```

```
write.csv(FE2010new,"C:/Sourav/Technology/Trainings/Data
Analytics with R, Excel &
Tableau/R_Directory/Files/FE2010new.csv")
names(FE2010)
```

```
##  [1] "EngDispl"            "NumCyl"              "FE"
##  [4] "NumGears"            "TransLockup"         "TransCreeperGear"
##  [7] "IntakeValvePerCyl"   "ExhaustValvesPerCyl" "VarValveTiming"
## [10] "VarValveLift"
```

```
fit<-
lm(FE~EngDispl+NumCyl+NumGears+TransLockup+TransCreeperGear+IntakeValvePerCyl
+ExhaustValvesPerCyl+VarValveTiming+VarValveLift,data=FE2010)
fit
```

```
##
## Call:
## lm(formula = FE ~ EngDispl + NumCyl + NumGears + TransLockup +
##     TransCreeperGear + IntakeValvePerCyl + ExhaustValvesPerCyl +
##     VarValveTiming + VarValveLift, data = FE2010)
##
## Coefficients:
##         (Intercept)             EngDispl               NumCyl
##             54.3472              -3.8610              -0.4888
##            NumGears           TransLockup     TransCreeperGear
##             -0.1725              -1.4450              -0.9138
##   IntakeValvePerCyl   ExhaustValvesPerCyl        VarValveTiming
##             -0.3737              -1.1105               1.6870
##        VarValveLift
##              0.6235
```

```
summary(fit)
```

```
##
## Call:
## lm(formula = FE ~ EngDispl + NumCyl + NumGears + TransLockup +
##     TransCreeperGear + IntakeValvePerCyl + ExhaustValvesPerCyl +
##     VarValveTiming + VarValveLift, data = FE2010)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.1153  -2.7142  -0.3535   2.4191  25.6521
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          54.3472     1.0973  49.530  < 2e-16 ***
## EngDispl             -3.8610     0.2805 -13.765  < 2e-16 ***
## NumCyl               -0.4888     0.1845  -2.649  0.00819 **
## NumGears             -0.1725     0.1065  -1.620  0.10555
## TransLockup          -1.4450     0.3000  -4.817 1.66e-06 ***
## TransCreeperGear     -0.9138     0.6681  -1.368  0.17167
## IntakeValvePerCyl    -0.3737     0.9892  -0.378  0.70566
## ExhaustValvesPerCyl  -1.1105     0.9598  -1.157  0.24752
## VarValveTiming        1.6870     0.3796   4.444 9.71e-06 ***
## VarValveLift          0.6235     0.3719   1.676  0.09393 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.489 on 1097 degrees of freedom
## Multiple R-squared:  0.6445, Adjusted R-squared:  0.6415
## F-statistic: 220.9 on 9 and 1097 DF,  p-value: < 2.2e-16

vif(fit)

##            EngDispl              NumCyl            NumGears
##            7.363137            6.750388            1.214238
##         TransLockup    TransCreeperGear   IntakeValvePerCyl
##            1.075253            1.137623            6.693985
## ExhaustValvesPerCyl      VarValveTiming        VarValveLift
##            7.073284            1.153276            1.057688

vif(fit)>5

##            EngDispl              NumCyl            NumGears
##                TRUE                TRUE               FALSE
##         TransLockup    TransCreeperGear   IntakeValvePerCyl
##               FALSE               FALSE                TRUE
## ExhaustValvesPerCyl      VarValveTiming        VarValveLift
##                TRUE               FALSE               FALSE
```