Sir Parashurambhau College, Pune 30
(AUTONOMOUS)
Department of Statistics
ST35257: Statistics Practical Paper III
Experiment No. 04
Title: Multiple Linear Regression Model II by using R Software
(Solution)

Name: KASTURI SHASHIKANT SAKPAL

Roll No.: 9151

```
> #Q.1
> setwd("C:\\Users\\kasturi\\Downloads")
> GS=read.table("gasoline.txt")
> View(GS)
> dim(GS)
[1] 32 12
> GS1=na.omit(GS)
> dim(GS1)
[1] 30 12
> View(GS1)
> #a)
> #MLRM
> names(GS1)
 [1] "y"   "x1"  "x2"  "x3"  "x4"  "x5"  "x6"  "x7"  "x8"  "x9"  "x10"
[12] "x11"
> ?I
> LM1=lm(y~x1+x2+x3+x4+x5+x6+x7+x8+x9+x10+I(x11), data=GS1)
> LM1

Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 +
    x10 + I(x11), data = GS1)

Coefficients:
(Intercept)           x1            x2            x3            x4
  17.339838    -0.075588     -0.069163      0.115117      1.494737
         x5           x6            x7            x8            x9
   5.843495     0.317583     -3.205390      0.180811     -0.397945
        x10         I(x11)
  -0.005115      0.638483


> #b)
```

```
> summary(LM1)

Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 +
    x10 + I(x11), data = GS1)

Residuals:
    Min      1Q  Median      3Q     Max
-5.3441 -1.6711 -0.4486  1.4906  5.2508

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 17.339838  30.355375   0.571   0.5749
x1          -0.075588   0.056347  -1.341   0.1964
x2          -0.069163   0.087791  -0.788   0.4411
x3           0.115117   0.088113   1.306   0.2078
x4           1.494737   3.101464   0.482   0.6357
x5           5.843495   3.148438   1.856   0.0799 .
x6           0.317583   1.288967   0.246   0.8082
x7          -3.205390   3.109185  -1.031   0.3162
x8           0.180811   0.130301   1.388   0.1822
x9          -0.397945   0.323456  -1.230   0.2344
x10         -0.005115   0.005896  -0.868   0.3971
I(x11)       0.638483   3.021680   0.211   0.8350
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.227 on 18 degrees of freedom
Multiple R-squared:  0.8355, Adjusted R-squared:  0.7349
F-statistic:  8.31 on 11 and 18 DF,  p-value: 5.231e-05

>#X11 is not contributing significantly in the regression model

> #c)
> par(mfrow=c(2,2))
> plot(LM1)
```
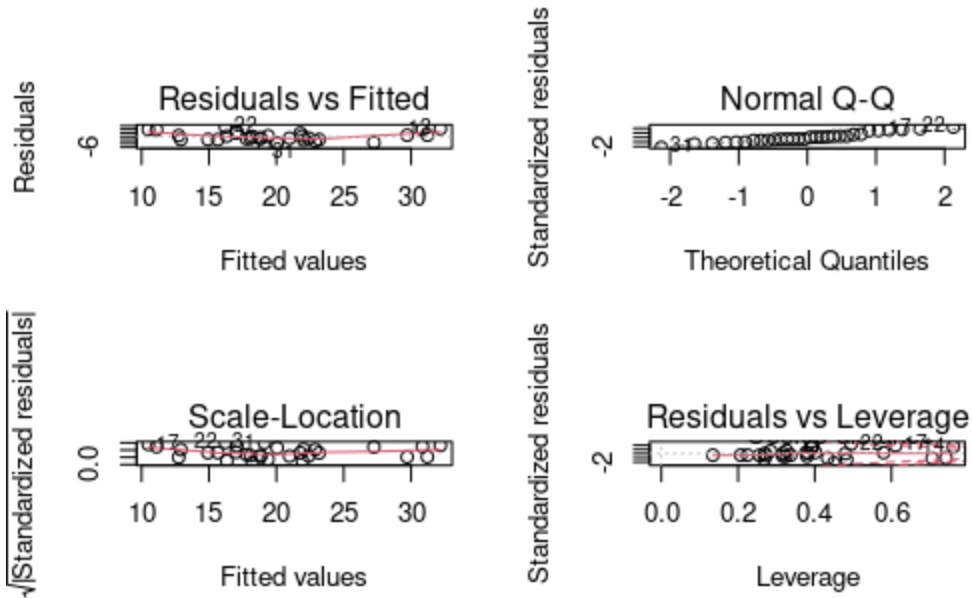
```
> #d)
> cor(GS1)
             y          x1          x2          x3          x4          x5
y    1.0000000 -0.8721701 -0.7968304 -0.8495915  0.42237247  0.6347500
x1  -0.8721701  1.0000000  0.9408473  0.9891628 -0.34697246 -0.6720903
x2  -0.7968304  0.9408473  1.0000000  0.9643592 -0.28989951 -0.5509642
x3  -0.8495915  0.9891628  0.9643592  1.0000000 -0.32599915 -0.6728661
x4   0.4223725 -0.3469725 -0.2898995 -0.3259992  1.00000000  0.4137808
x5   0.6347500 -0.6720903 -0.5509642 -0.6728661  0.41378081  1.0000000
x6  -0.4718055  0.6427984  0.7614190  0.6531263  0.03748643 -0.2195283
x7   0.7077682 -0.7719151 -0.6259445 -0.7461800  0.55823570  0.8717662
x8  -0.7528208  0.8623681  0.8027387  0.8641224 -0.30415026 -0.5613315
x9  -0.7629952  0.7974811  0.7105117  0.7881284 -0.37817358 -0.4534470
x10 -0.8528801  0.9515520  0.8878810  0.9434871 -0.35845879 -0.5798617
x11 -0.7212809  0.8244446  0.7086735  0.8012765 -0.44054570 -0.7546650
             x6          x7          x8          x9         x10         x11
y   -0.47180548  0.7077682 -0.7528208 -0.7629952 -0.8528801 -0.7212809
x1   0.64279836 -0.7719151  0.8623681  0.7974811  0.9515520  0.8244446
x2   0.76141897 -0.6259445  0.8027387  0.7105117  0.8878810  0.7086735
x3   0.65312630 -0.7461800  0.8641224  0.7881284  0.9434871  0.8012765
x4   0.03748643  0.5582357 -0.3041503 -0.3781736 -0.3584588 -0.4405457
x5  -0.21952829  0.8717662 -0.5613315 -0.4534470 -0.5798617 -0.7546650
x6   1.00000000 -0.2756386  0.4220680  0.3003862  0.5203669  0.3954893
x7  -0.27563863  1.0000000 -0.6552065 -0.6551300 -0.7058126 -0.8506963
x8   0.42206800 -0.6552065  1.0000000  0.8831512  0.9554541  0.6824919
x9   0.30038618 -0.6551300  0.8831512  1.0000000  0.8994711  0.6326677
x10  0.52036693 -0.7058126  0.9554541  0.8994711  1.0000000  0.7530353
x11  0.39548928 -0.8506963  0.6824919  0.6326677  0.7530353  1.0000000
> #There is high correlation between regressor variables. For example; X2
```

```
> #Which lead to   #### highly fitted model
> #e)
> library(MASS)
> install.packages("CARS")
Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
(as 'lib' is unspecified)
also installing the dependencies 'Rcpp', 'SparseM', 'MatrixModels',
'cubature', 'quadprog', 'quantreg', 'np'

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/Rcpp_1.0.8.tar.gz'
Content type 'application/x-gzip' length 4234256 bytes (4.0 MB)
==================================================
downloaded 4.0 MB

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/SparseM_1.81.tar.g
z'
Content type 'application/x-gzip' length 1099716 bytes (1.0 MB)
==================================================
downloaded 1.0 MB

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/MatrixModels_0.5-0
.tar.gz'
Content type 'application/x-gzip' length 442837 bytes (432 KB)
==================================================
downloaded 432 KB

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/cubature_2.0.4.2.t
ar.gz'
Content type 'application/x-gzip' length 3042928 bytes (2.9 MB)
==================================================
downloaded 2.9 MB

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/quadprog_1.5-8.tar
.gz'
Content type 'application/x-gzip' length 41922 bytes (40 KB)
==================================================
downloaded 40 KB

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/quantreg_5.87.tar.
gz'
Content type 'application/x-gzip' length 1648112 bytes (1.6 MB)
```

```
==================================================
downloaded 1.6 MB

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/np_0.60-11.tar.gz'
Content type 'application/x-gzip' length 3131848 bytes (3.0 MB)
==================================================
downloaded 3.0 MB

trying URL
'http://rspm/default/__linux__/focal/latest/src/contrib/CARS_0.2.2.tar.gz'
Content type 'application/x-gzip' length 27866 bytes (27 KB)
==================================================
downloaded 27 KB

* installing *binary* package 'Rcpp' ...
* DONE (Rcpp)
* installing *binary* package 'SparseM' ...
* DONE (SparseM)
* installing *binary* package 'MatrixModels' ...
* DONE (MatrixModels)
* installing *binary* package 'quadprog' ...
* DONE (quadprog)
* installing *binary* package 'cubature' ...
* DONE (cubature)
* installing *binary* package 'quantreg' ...
* DONE (quantreg)
* installing *binary* package 'np' ...
* DONE (np)
* installing *binary* package 'CARS' ...
* DONE (CARS)

The downloaded source packages are in
        '/tmp/RtmpWjw9Uf/downloaded_packages'

> library(CARS)
> #OR
> library(faraway)
> vif(LM1)
        x1          x2          x3          x4          x5          x6
119.487804   42.800811 149.234409    2.060036    7.729187    5.324730
        x7          x8          x9         x10      I(x11)
 11.761341   20.917632    9.397108   85.744344    5.145052
> #f)
> library(MASS)
> NL=lm(y~1, data=GS1)
> NL
```

```
Call:
lm(formula = y ~ 1, data = GS1)

Coefficients:
(Intercept)
      20.04

> FS=stepAIC(object=NL,scope=list(lower=NL, upper=LM1),
direction="forward")
Start:  AIC=111.1
y ~ 1

          Df Sum of Sq      RSS      AIC
+ x1       1     866.50   272.61   70.205
+ x10      1     828.59   310.52   74.111
+ x3       1     822.21   316.89   74.721
+ x2       1     723.26   415.84   82.873
+ x9       1     663.14   475.96   86.924
+ x8       1     645.58   493.53   88.012
+ I(x11)   1     592.62   546.49   91.070
+ x7       1     570.62   568.49   92.253
+ x5       1     458.95   680.15   97.634
+ x6       1     253.57   885.54  105.550
+ x4       1     203.21   935.89  107.209
<none>                   1139.11  111.104

Step:  AIC=70.21
y ~ x1

          Df Sum of Sq     RSS      AIC
+ x4       1    18.5716  254.04   70.089
<none>                   272.61   70.205
+ x6       1    15.3154  257.29   70.471
+ x9       1    14.2388  258.37   70.596
+ x3       1     9.1052  263.50   71.186
+ x10      1     6.3538  266.26   71.498
+ x2       1     5.5959  267.01   71.583
+ x5       1     4.9016  267.71   71.661
+ x7       1     3.3600  269.25   71.833
+ I(x11)   1     0.0176  272.59   72.203
+ x8       1     0.0021  272.61   72.205

Step:  AIC=70.09
y ~ x1 + x4

          Df Sum of Sq     RSS      AIC
```

```
<none>                  254.04 70.089
+ x9       1    9.3032 244.74 70.969
+ x6       1    6.3629 247.68 71.328
+ x3       1    6.2419 247.80 71.342
+ x10      1    4.4437 249.59 71.559
+ x2       1    3.5435 250.50 71.667
+ I(x11)   1    1.3744 252.66 71.926
+ x5       1    1.2821 252.76 71.937
+ x7       1    0.0927 253.94 72.078
+ x8       1    0.0000 254.04 72.089
> #g)
> BS=stepAIC(LM1,scope=list(lower=NL, upper=LM1), direction="backward")
Start:  AIC=78.96
y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10 + I(x11)


          Df Sum of Sq    RSS    AIC
- I(x11)   1     0.465 187.87 77.036
- x6       1     0.632 188.03 77.063
- x4       1     2.418 189.82 77.346
- x2       1     6.462 193.86 77.979
- x10      1     7.836 195.24 78.190
- x7       1    11.065 198.47 78.683
<none>                  187.40 78.962
- x9       1    15.758 203.16 79.384
- x3       1    17.770 205.17 79.679
- x1       1    18.736 206.14 79.820
- x8       1    20.047 207.45 80.011
- x5       1    35.864 223.26 82.215


Step:  AIC=77.04
y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10

        Df Sum of Sq    RSS    AIC
- x6     1     0.536 188.40 75.121
- x4     1     2.363 190.23 75.411
- x2     1     6.642 194.51 76.078
- x10    1     7.985 195.85 76.285
<none>                187.87 77.036
- x7     1    14.124 201.99 77.211
- x9     1    16.914 204.78 77.622
- x3     1    17.815 205.68 77.754
- x1     1    18.280 206.15 77.822
- x8     1    20.301 208.17 78.114
- x5     1    36.370 224.24 80.345


Step:  AIC=75.12
y ~ x1 + x2 + x3 + x4 + x5 + x7 + x8 + x9 + x10
```

```
        Df Sum of Sq    RSS    AIC
- x4    1       3.451 191.85 73.666
- x2    1       6.932 195.33 74.205
- x10   1       9.351 197.75 74.574
<none>              188.40 75.121
- x7    1      14.473 202.87 75.342
- x3    1      17.802 206.20 75.830
- x9    1      18.146 206.55 75.880
- x1    1      18.780 207.18 75.972
- x8    1      21.244 209.65 76.326
- x5    1      39.332 227.73 78.809

Step:  AIC=73.67
y ~ x1 + x2 + x3 + x5 + x7 + x8 + x9 + x10

        Df Sum of Sq    RSS    AIC
- x2    1      10.780 202.63 73.306
- x7    1      11.113 202.97 73.355
<none>              191.85 73.666
- x10   1      14.988 206.84 73.923
- x1    1      16.602 208.46 74.156
- x9    1      18.072 209.92 74.366
- x3    1      21.314 213.17 74.826
- x8    1      28.835 220.69 75.867
- x5    1      40.323 232.18 77.389

Step:  AIC=73.31
y ~ x1 + x3 + x5 + x7 + x8 + x9 + x10

        Df Sum of Sq    RSS    AIC
- x7    1      10.457 213.09 72.815
- x3    1      10.595 213.23 72.835
- x1    1      11.998 214.63 73.032
- x9    1      12.643 215.28 73.122
- x10   1      13.887 216.52 73.295
<none>              202.63 73.306
- x8    1      27.665 230.30 75.145
- x5    1      30.191 232.82 75.472

Step:  AIC=72.82
y ~ x1 + x3 + x5 + x8 + x9 + x10

        Df Sum of Sq    RSS    AIC
- x3    1      4.8720 217.96 71.494
- x9    1      5.2049 218.29 71.539
- x1    1      5.3212 218.41 71.555
```

```
<none>                  213.09 72.815
- x10    1    18.3677 231.46 73.296
- x5     1    23.3458 236.44 73.934
- x8     1    26.0316 239.12 74.273


Step:  AIC=71.49
y ~ x1 + x5 + x8 + x9 + x10

        Df Sum of Sq    RSS    AIC
- x1     1     0.765 218.73 69.599
- x9     1     5.863 223.82 70.290
<none>                 217.96 71.494
- x10    1    20.291 238.25 72.164
- x5     1    23.020 240.98 72.506
- x8     1    31.634 249.59 73.559


Step:  AIC=69.6
y ~ x5 + x8 + x9 + x10

        Df Sum of Sq    RSS    AIC
- x9     1     5.097 223.82 68.290
<none>                 218.73 69.599
- x5     1    40.404 259.13 72.684
- x8     1    57.407 276.13 74.591
- x10    1   135.105 353.83 82.029


Step:  AIC=68.29
y ~ x5 + x8 + x10

        Df Sum of Sq    RSS    AIC
<none>                 223.82 68.290
- x5     1    36.314 260.14 70.800
- x8     1    52.960 276.78 72.661
- x10    1   194.838 418.66 85.076
```
> #h)
> stepAIC(LM1,scope=list(lower=NL, upper=LM1), direction="both") #stepwise
regression model
```
Start:  AIC=78.96
y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10 + I(x11)

          Df Sum of Sq    RSS    AIC
- I(x11)   1     0.465 187.87 77.036
- x6       1     0.632 188.03 77.063
- x4       1     2.418 189.82 77.346
- x2       1     6.462 193.86 77.979
- x10      1     7.836 195.24 78.190
- x7       1    11.065 198.47 78.683
```

```
<none>                    187.40 78.962
- x9       1    15.758 203.16 79.384
- x3       1    17.770 205.17 79.679
- x1       1    18.736 206.14 79.820
- x8       1    20.047 207.45 80.011
- x5       1    35.864 223.26 82.215


Step:  AIC=77.04
y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10

          Df Sum of Sq    RSS    AIC
- x6       1     0.536 188.40 75.121
- x4       1     2.363 190.23 75.411
- x2       1     6.642 194.51 76.078
- x10      1     7.985 195.85 76.285
<none>                    187.87 77.036
- x7       1    14.124 201.99 77.211
- x9       1    16.914 204.78 77.622
- x3       1    17.815 205.68 77.754
- x1       1    18.280 206.15 77.822
- x8       1    20.301 208.17 78.114
+ I(x11)   1     0.465 187.40 78.962
- x5       1    36.370 224.24 80.345


Step:  AIC=75.12
y ~ x1 + x2 + x3 + x4 + x5 + x7 + x8 + x9 + x10

          Df Sum of Sq    RSS    AIC
- x4       1     3.451 191.85 73.666
- x2       1     6.932 195.33 74.205
- x10      1     9.351 197.75 74.574
<none>                    188.40 75.121
- x7       1    14.473 202.87 75.342
- x3       1    17.802 206.20 75.830
- x9       1    18.146 206.55 75.880
- x1       1    18.780 207.18 75.972
- x8       1    21.244 209.65 76.326
+ x6       1     0.536 187.87 77.036
+ I(x11)   1     0.368 188.03 77.063
- x5       1    39.332 227.73 78.809


Step:  AIC=73.67
y ~ x1 + x2 + x3 + x5 + x7 + x8 + x9 + x10

          Df Sum of Sq    RSS    AIC
- x2       1    10.780 202.63 73.306
- x7       1    11.113 202.97 73.355
```

```
<none>                    191.85 73.666
- x10      1    14.988 206.84 73.923
- x1       1    16.602 208.46 74.156
- x9       1    18.072 209.92 74.366
- x3       1    21.314 213.17 74.826
+ x4       1     3.451 188.40 75.121
+ x6       1     1.625 190.23 75.411
+ I(x11)   1     0.244 191.61 75.628
- x8       1    28.835 220.69 75.867
- x5       1    40.323 232.18 77.389


Step:  AIC=73.31
y ~ x1 + x3 + x5 + x7 + x8 + x9 + x10


          Df Sum of Sq     RSS     AIC
- x7       1    10.4572 213.09 72.815
- x3       1    10.5951 213.23 72.835
- x1       1    11.9982 214.63 73.032
- x9       1    12.6431 215.28 73.122
- x10      1    13.8874 216.52 73.295
<none>                    202.63 73.306
+ x2       1    10.7797 191.85 73.666
+ x4       1     7.2987 195.33 74.205
- x8       1    27.6646 230.30 75.145
+ I(x11)   1     0.7404 201.89 75.196
+ x6       1     0.3586 202.27 75.253
- x5       1    30.1906 232.82 75.472


Step:  AIC=72.82
y ~ x1 + x3 + x5 + x8 + x9 + x10


          Df Sum of Sq     RSS     AIC
- x3       1     4.8720 217.96 71.494
- x9       1     5.2049 218.29 71.539
- x1       1     5.3212 218.41 71.555
<none>                    213.09 72.815
- x10      1    18.3677 231.46 73.296
+ x7       1    10.4572 202.63 73.306
+ x2       1    10.1239 202.97 73.355
- x5       1    23.3458 236.44 73.934
+ I(x11)   1     3.9621 209.13 74.252
- x8       1    26.0316 239.12 74.273
+ x4       1     1.3059 211.78 74.631
+ x6       1     0.6112 212.48 74.729


Step:  AIC=71.49
y ~ x1 + x5 + x8 + x9 + x10
```

```
          Df Sum of Sq    RSS    AIC
- x1       1      0.765 218.73 69.599
- x9       1      5.863 223.82 70.290
<none>                   217.96 71.494
- x10      1     20.291 238.25 72.164
- x5       1     23.020 240.98 72.506
+ x3       1      4.872 213.09 72.815
+ x7       1      4.734 213.23 72.835
+ I(x11)   1      2.115 215.85 73.201
+ x4       1      1.891 216.07 73.232
+ x2       1      0.120 217.84 73.477
+ x6       1      0.053 217.91 73.486
- x8       1     31.634 249.59 73.559

Step:  AIC=69.6
y ~ x5 + x8 + x9 + x10

          Df Sum of Sq    RSS    AIC
- x9       1      5.097 223.82 68.290
<none>                   218.73 69.599
+ x7       1      3.713 215.01 71.085
+ I(x11)   1      1.354 217.37 71.413
+ x4       1      1.326 217.40 71.416
+ x1       1      0.765 217.96 71.494
+ x2       1      0.684 218.04 71.505
+ x6       1      0.547 218.18 71.524
+ x3       1      0.316 218.41 71.555
- x5       1     40.404 259.13 72.684
- x8       1     57.407 276.13 74.591
- x10      1    135.105 353.83 82.029

Step:  AIC=68.29
y ~ x5 + x8 + x10

          Df Sum of Sq    RSS    AIC
<none>                   223.82 68.290
+ x9       1      5.097 218.73 69.599
+ x4       1      2.730 221.09 69.922
+ x3       1      1.642 222.18 70.069
+ I(x11)   1      1.369 222.46 70.106
+ x7       1      0.610 223.21 70.208
+ x6       1      0.137 223.69 70.272
+ x2       1      0.017 223.81 70.288
+ x1       1      0.000 223.82 70.290
- x5       1     36.314 260.14 70.800
- x8       1     52.960 276.78 72.661
```

```
- x10     1   194.838 418.66 85.076

Call:
lm(formula = y ~ x5 + x8 + x10, data = GS1)

Coefficients:
(Intercept)              x5              x8             x10
   4.590404     2.597240      0.217814    -0.009485
```

```
> #k)
> vif(FS)
      x1         x4
1.136867 1.136867
> vif(BS)
       x5          x8         x10
 1.507958 11.490675 11.856751
```

```
> ##Q2)
> #a)
> NFL <- read.table("NationalFootballLeague.txt",header = TRUE)
> View(NFL)
>
>
> LM2 <- lm(y~x1+x2+x3+x4+x5+x6+x7+x8+x9, data = NFL)
> LM2

Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9,
    data = NFL)

Coefficients:
(Intercept)           x1           x2           x3           x4
 -7.292e+00    8.124e-04    3.631e-03    1.222e-01    3.189e-02
         x5           x6           x7           x8           x9
  1.511e-05    1.590e-03    1.544e-01   -3.895e-03   -1.791e-03


>
> #b)
> summary(LM2)

Call:
lm(formula = y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9,
    data = NFL)

Residuals:
    Min       1Q   Median       3Q      Max
-3.0408  -0.6802  -0.1131   0.9835   2.9785

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -7.292e+00  1.281e+01  -0.569 0.576312
x1           8.124e-04  2.006e-03   0.405 0.690329
x2           3.631e-03  8.410e-04   4.318 0.000414 ***
x3           1.222e-01  2.590e-01   0.472 0.642750
x4           3.189e-02  4.160e-02   0.767 0.453289
x5           1.511e-05  4.684e-02   0.000 0.999746
x6           1.590e-03  3.248e-03   0.490 0.630338
x7           1.544e-01  1.521e-01   1.015 0.323547
x8          -3.895e-03  2.052e-03  -1.898 0.073793 .
x9          -1.791e-03  1.417e-03  -1.264 0.222490
---
```
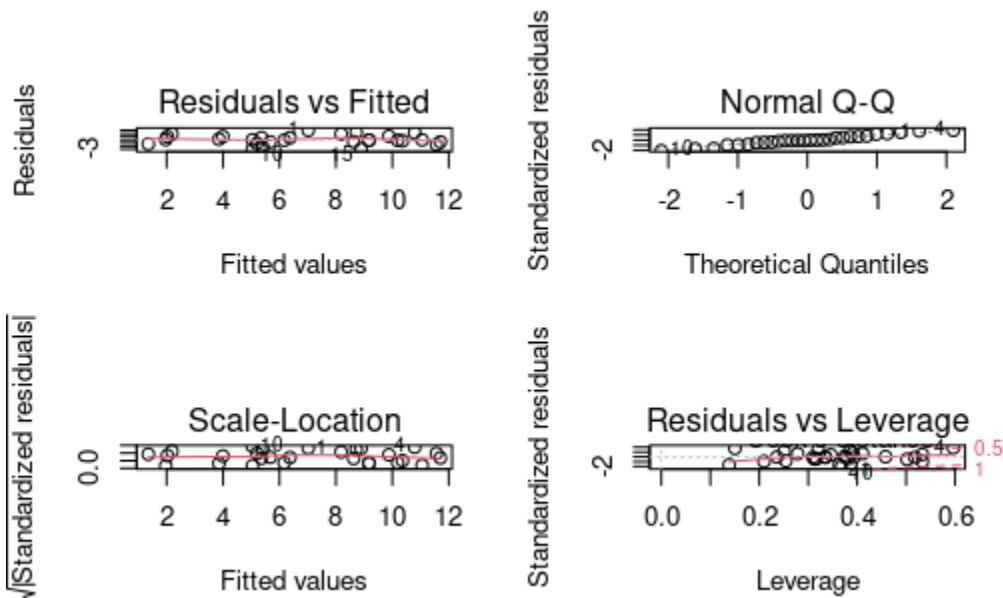
```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.83 on 18 degrees of freedom
Multiple R-squared:  0.8156,     Adjusted R-squared:  0.7234
F-statistic: 8.846 on 9 and 18 DF,  p-value: 5.303e-05


>
>
> #c)
> par(mfrow = c(2,2))
> plot(LM2)
```



```
>
> #d)
> cor(NFL)
            y            x1          x2           x3          x4
y   1.00000000  0.59323604  0.48273470 -0.080812472  0.25847477
x1  0.59323604  1.00000000 -0.03674736  0.212471227  0.07029904
x2  0.48273470 -0.03674736  1.00000000 -0.068815157  0.30151583
x3 -0.08081247  0.21247123 -0.06881516  1.000000000 -0.41309561
x4  0.25847477  0.07029904  0.30151583 -0.413095614  1.00000000
x5  0.51320624  0.59998017  0.13499515  0.115098074  0.14902865
x6  0.22403447  0.25297272 -0.19283713 -0.003115748 -0.12818435
x7  0.54534104  0.83728269 -0.19691540  0.162511469 -0.10100316
x8 -0.73802730 -0.65854627 -0.05104783  0.290438108 -0.16402353
x9 -0.30374811 -0.11055739  0.14598149  0.088195595  0.05913611
            x5          x6          x7           x8          x9
y   0.51320624  0.224034472  0.5453410 -0.73802730 -0.30374811
```

```
x1   0.59998017   0.252972716   0.8372827  -0.65854627  -0.11055739
x2   0.13499515  -0.192837129  -0.1969154  -0.05104783   0.14598149
x3   0.11509807  -0.003115748   0.1625115   0.29043811   0.08819559
x4   0.14902865  -0.128184348  -0.1010032  -0.16402353   0.05913611
x5   1.00000000   0.258915336   0.6095632  -0.47004608  -0.09028906
x6   0.25891534   1.000000000   0.3670779  -0.35249327  -0.17275608
x7   0.60956318   0.367077900   1.0000000  -0.68504573  -0.20331784
x8  -0.47004608  -0.352493271  -0.6850457   1.00000000   0.41746519
x9  -0.09028906  -0.172756078  -0.2033178   0.41746519   1.00000000
>
> #e) variance inflation factor(VIF)
> library(faraway)
> library(CARS)
> vif(LM2)                #variance inflation factor(VIF)
      x1         x2         x3         x4         x5         x6         x7
x8
4.827645 1.420161 2.126597 1.566107 1.924035 1.275979 5.414572
4.535643
      x9
1.423390
>
> #f)forward selection algorithm
> library(MASS)
>
> NL <- lm(y~1, data = NFL)        #Null model depends on all
regressor variables
> forward_sel <- stepAIC(NL, scope = list(lower = NL, upper =
LM2),direction = "forward")
Start:  AIC=70.81
y ~ 1

       Df Sum of Sq    RSS    AIC
+ x8    1   178.092 148.87 50.785
+ x1    1   115.068 211.90 60.669
+ x7    1    97.238 229.73 62.931
+ x5    1    86.116 240.85 64.255
+ x2    1    76.193 250.77 65.385
+ x9    1    30.167 296.80 70.104
<none>              326.96 70.814
+ x4    1    21.844 305.12 70.878
+ x6    1    16.411 310.55 71.372
+ x3    1     2.135 324.83 72.631

Step:  AIC=50.78
```

```
y ~ x8

        Df Sum of Sq      RSS     AIC
+ x2     1     64.934   83.938  36.741
+ x5     1     11.607  137.265  50.512
<none>                 148.872  50.785
+ x1     1      6.636  142.236  51.508
+ x3     1      6.368  142.504  51.561
+ x4     1      6.345  142.527  51.565
+ x7     1      0.974  147.898  52.601
+ x6     1      0.487  148.385  52.693
+ x9     1      0.008  148.864  52.783


Step:  AIC=36.74
y ~ x8 + x2

        Df Sum of Sq     RSS     AIC
+ x7     1    14.0682  69.870  33.604
+ x1     1    11.1905  72.748  34.734
+ x3     1     8.9010  75.037  35.602
+ x5     1     5.8147  78.124  36.730
<none>                83.938  36.741
+ x9     1     2.0256  81.913  38.057
+ x6     1     1.3216  82.617  38.296
+ x4     1     0.0161  83.922  38.735


Step:  AIC=33.6
y ~ x8 + x2 + x7

        Df Sum of Sq      RSS     AIC
+ x9     1     4.8657  65.004  33.583
<none>                 69.870  33.604
+ x3     1     1.3873  68.483  35.043
+ x4     1     0.9792  68.891  35.209
+ x1     1     0.9022  68.968  35.240
+ x6     1     0.4879  69.382  35.408
+ x5     1     0.2987  69.571  35.484


Step:  AIC=33.58
y ~ x8 + x2 + x7 + x9

        Df Sum of Sq     RSS     AIC
<none>                65.004  33.583
+ x1     1    1.86452  63.140  34.768
```

```
+ x4    1    1.74260 63.262 34.822
+ x3    1    0.70148 64.303 35.279
+ x6    1    0.45071 64.554 35.388
+ x5    1    0.32667 64.678 35.442
>
> #g)backward elimination
> backward_eli <- stepAIC(LM2, scope = list(lower = NL, upper =
LM2),direction = "backward")
Start:  AIC=41.48
y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9

       Df Sum of Sq     RSS    AIC
- x5    1     0.000  60.293 39.476
- x1    1     0.549  60.842 39.730
- x3    1     0.746  61.039 39.821
- x6    1     0.803  61.096 39.847
- x4    1     1.968  62.261 40.376
- x7    1     3.451  63.744 41.035
<none>              60.293 41.476
- x9    1     5.348  65.642 41.856
- x8    1    12.072  72.365 44.587
- x2    1    62.448 122.741 59.380

Step:  AIC=39.48
y ~ x1 + x2 + x3 + x4 + x6 + x7 + x8 + x9

       Df Sum of Sq     RSS    AIC
- x1    1     0.553  60.846 37.732
- x3    1     0.750  61.043 37.822
- x6    1     0.818  61.111 37.854
- x4    1     2.053  62.346 38.414
- x7    1     3.859  64.152 39.213
<none>              60.293 39.476
- x9    1     5.351  65.644 39.857
- x8    1    12.086  72.379 42.592
- x2    1    66.979 127.272 58.395

Step:  AIC=37.73
y ~ x2 + x3 + x4 + x6 + x7 + x8 + x9

       Df Sum of Sq     RSS    AIC
- x6    1     0.690  61.536 36.048
- x3    1     1.715  62.561 36.510
- x4    1     3.051  63.897 37.102
```

```
<none>                 60.846 37.732
- x9    1      4.852   65.698 37.880
- x7    1      8.961   69.807 39.579
- x8    1     16.599   77.445 42.486
- x2    1     67.010  127.856 56.524

Step:  AIC=36.05
y ~ x2 + x3 + x4 + x7 + x8 + x9

        Df Sum of Sq      RSS     AIC
- x3    1      1.726   63.262 34.822
- x4    1      2.767   64.303 35.279
<none>                 61.536 36.048
- x9    1      4.831   66.367 36.164
- x7    1      9.390   70.926 38.024
- x8    1     18.314   79.851 41.343
- x2    1     66.447  127.984 54.552

Step:  AIC=34.82
y ~ x2 + x4 + x7 + x8 + x9

        Df Sum of Sq      RSS     AIC
- x4    1      1.743   65.004 33.583
<none>                 63.262 34.822
- x9    1      5.629   68.891 35.209
- x8    1     17.701   80.962 39.730
- x7    1     18.583   81.845 40.033
- x2    1     75.598  138.860 54.835

Step:  AIC=33.58
y ~ x2 + x7 + x8 + x9

        Df Sum of Sq      RSS     AIC
<none>                 65.004 33.583
- x9    1      4.866   69.870 33.604
- x7    1     16.908   81.913 38.057
- x8    1     23.299   88.303 40.160
- x2    1     82.892  147.897 54.601
>
> #h)stepwise
> stepAIC(LM2, direction = "both")
Start:  AIC=41.48
y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9
```

```
        Df Sum of Sq     RSS     AIC
- x5    1     0.000   60.293  39.476
- x1    1     0.549   60.842  39.730
- x3    1     0.746   61.039  39.821
- x6    1     0.803   61.096  39.847
- x4    1     1.968   62.261  40.376
- x7    1     3.451   63.744  41.035
<none>                60.293  41.476
- x9    1     5.348   65.642  41.856
- x8    1    12.072   72.365  44.587
- x2    1    62.448  122.741  59.380

Step:  AIC=39.48
y ~ x1 + x2 + x3 + x4 + x6 + x7 + x8 + x9

        Df Sum of Sq     RSS     AIC
- x1    1     0.553   60.846  37.732
- x3    1     0.750   61.043  37.822
- x6    1     0.818   61.111  37.854
- x4    1     2.053   62.346  38.414
- x7    1     3.859   64.152  39.213
<none>                60.293  39.476
- x9    1     5.351   65.644  39.857
+ x5    1     0.000   60.293  41.476
- x8    1    12.086   72.379  42.592
- x2    1    66.979  127.272  58.395

Step:  AIC=37.73
y ~ x2 + x3 + x4 + x6 + x7 + x8 + x9

        Df Sum of Sq     RSS     AIC
- x6    1     0.690   61.536  36.048
- x3    1     1.715   62.561  36.510
- x4    1     3.051   63.897  37.102
<none>                60.846  37.732
- x9    1     4.852   65.698  37.880
+ x1    1     0.553   60.293  39.476
- x7    1     8.961   69.807  39.579
+ x5    1     0.004   60.842  39.730
- x8    1    16.599   77.445  42.486
- x2    1    67.010  127.856  56.524

Step:  AIC=36.05
y ~ x2 + x3 + x4 + x7 + x8 + x9
```

```
        Df Sum of Sq      RSS     AIC
- x3     1      1.726   63.262 34.822
- x4     1      2.767   64.303 35.279
<none>                  61.536 36.048
- x9     1      4.831   66.367 36.164
+ x6     1      0.690   60.846 37.732
+ x1     1      0.425   61.111 37.854
- x7     1      9.390   70.926 38.024
+ x5     1      0.027   61.509 38.036
- x8     1     18.314   79.851 41.343
- x2     1     66.447  127.984 54.552

Step:  AIC=34.82
y ~ x2 + x4 + x7 + x8 + x9


        Df Sum of Sq      RSS     AIC
- x4     1      1.743   65.004 33.583
<none>                  63.262 34.822
- x9     1      5.629   68.891 35.209
+ x3     1      1.726   61.536 36.048
+ x1     1      1.321   61.941 36.231
+ x6     1      0.700   62.561 36.510
+ x5     1      0.101   63.160 36.777
- x8     1     17.701   80.962 39.730
- x7     1     18.583   81.845 40.033
- x2     1     75.598  138.860 54.835

Step:  AIC=33.58
y ~ x2 + x7 + x8 + x9


        Df Sum of Sq      RSS     AIC
<none>                  65.004 33.583
- x9     1      4.866   69.870 33.604
+ x1     1      1.865   63.140 34.768
+ x4     1      1.743   63.262 34.822
+ x3     1      0.701   64.303 35.279
+ x6     1      0.451   64.554 35.388
+ x5     1      0.327   64.678 35.442
- x7     1     16.908   81.913 38.057
- x8     1     23.299   88.303 40.160
- x2     1     82.892  147.897 54.601

Call:
```

```
lm(formula = y ~ x2 + x7 + x8 + x9, data = NFL)

Coefficients:
(Intercept)            x2            x7            x8            x9
  -1.821703      0.003819      0.216894     -0.004015     -0.001635

> #k)
>
> vif(forward_sel)                #forward selection vif
      x8        x2        x7        x9
2.496630 1.183190 2.182090 1.304339
> vif(backward_eli)              #backward elimination vif
      x2        x7        x8        x9
1.183190 2.182090 2.496630 1.304339



 >
```