ASSIGNMENT 4

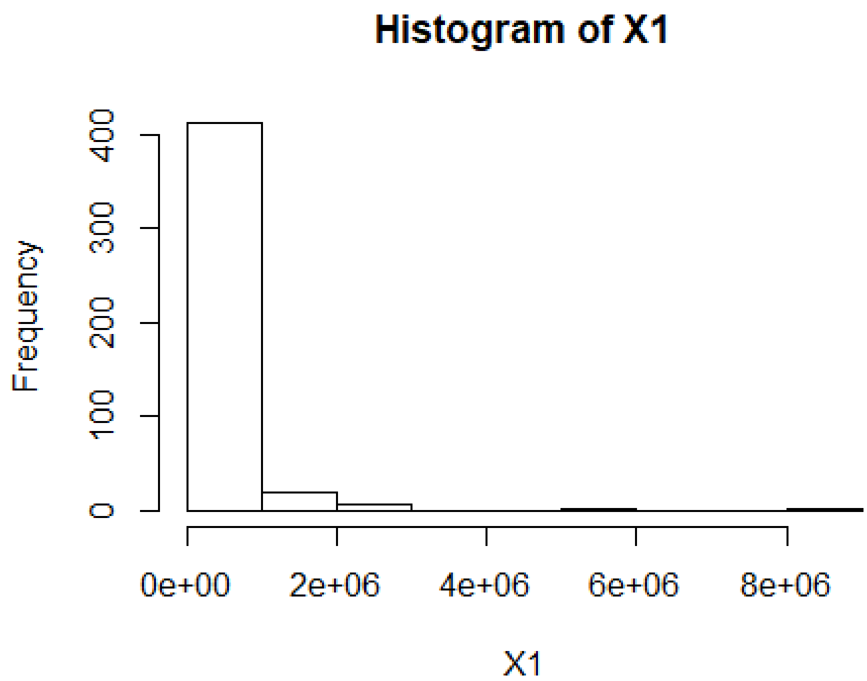
Kushal Kapadia

1213714799

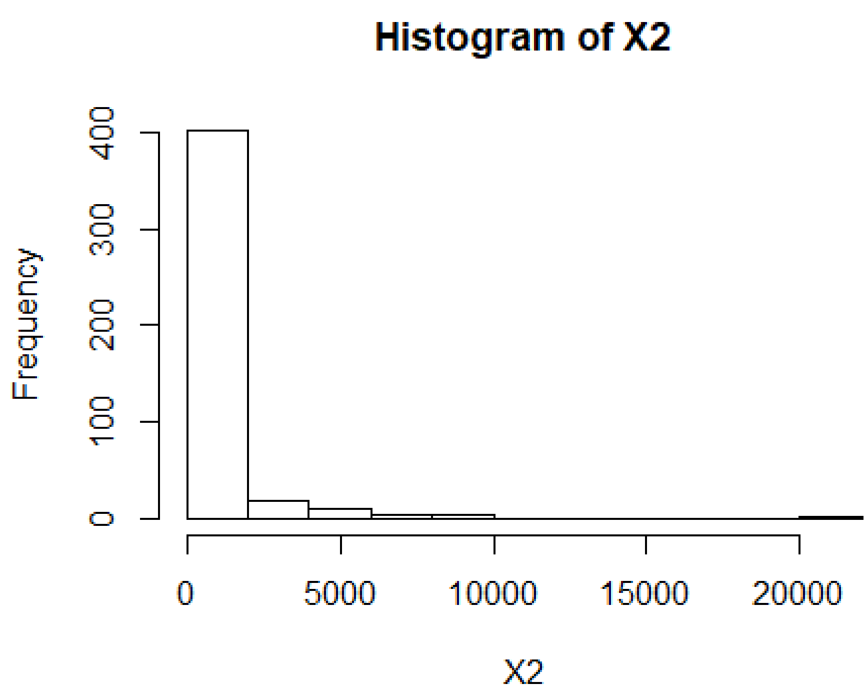
Part 1

(a)

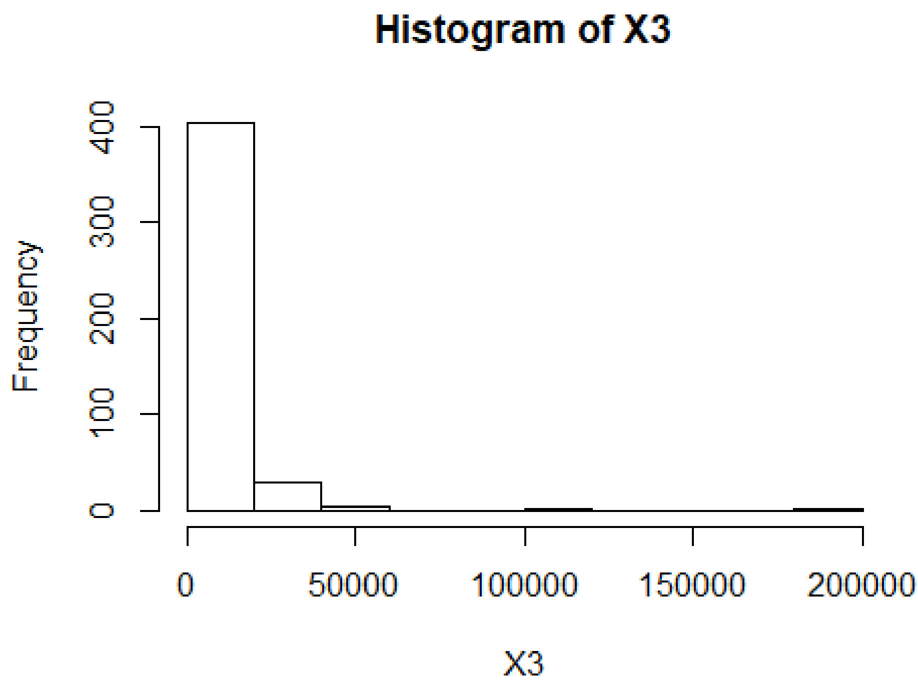
```r
cdi = read.table("D:\\ASU Stuff\\SEM-1\\STP 530\\CDI_data.txt")
Y = cdi$V8
# Model-I
X1 = cdi$V5
X2 = cdi$V4
X3 = cdi$V16

# Model-II
x1 = cdi$V5 / cdi$V4
x2 = cdi$V7
x3 = cdi$V16

hist(X1)
```



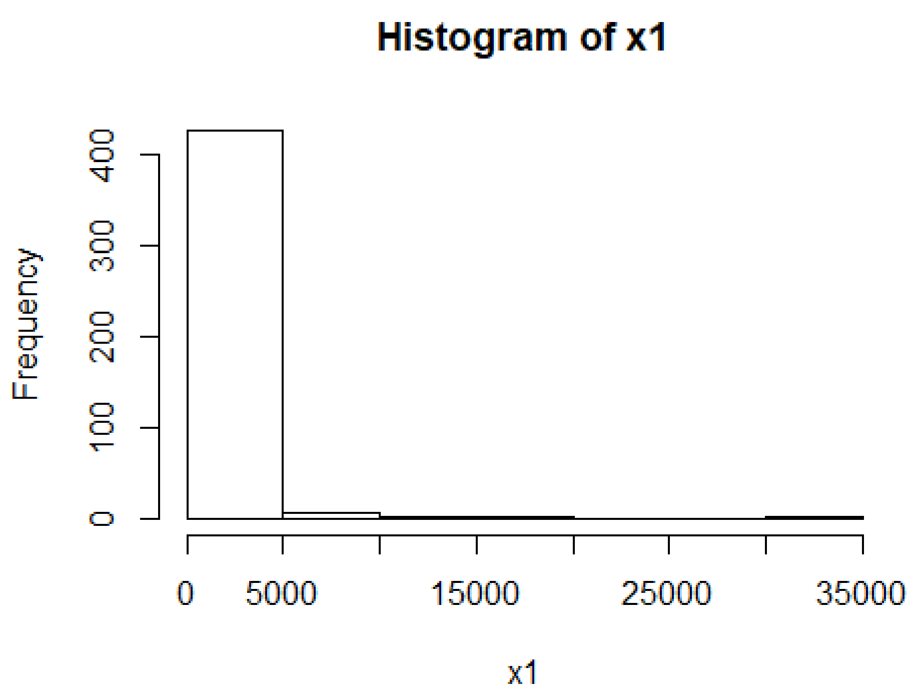

```r
hist(X2)
```



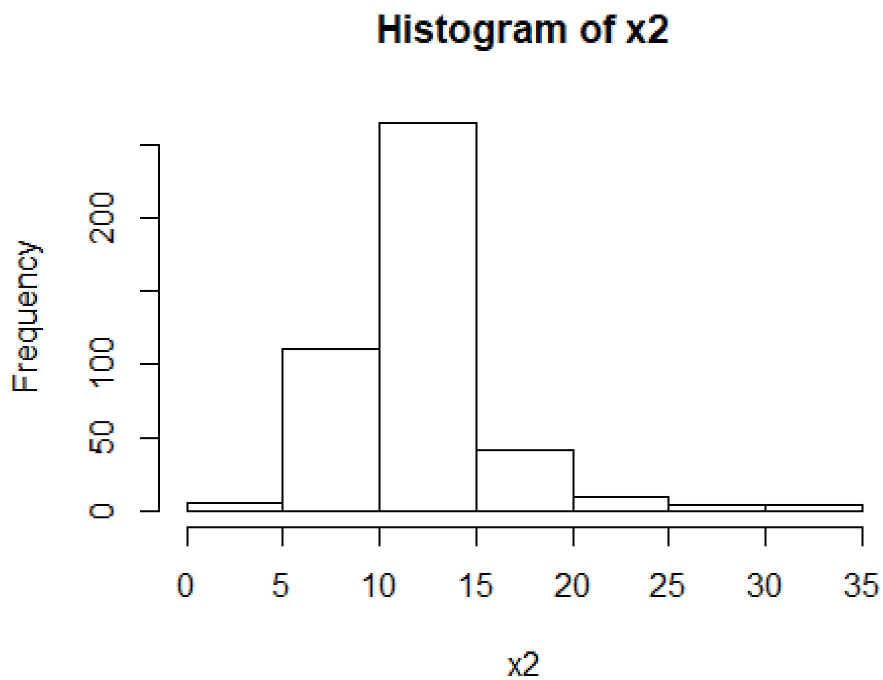

```r
hist(X3)
```



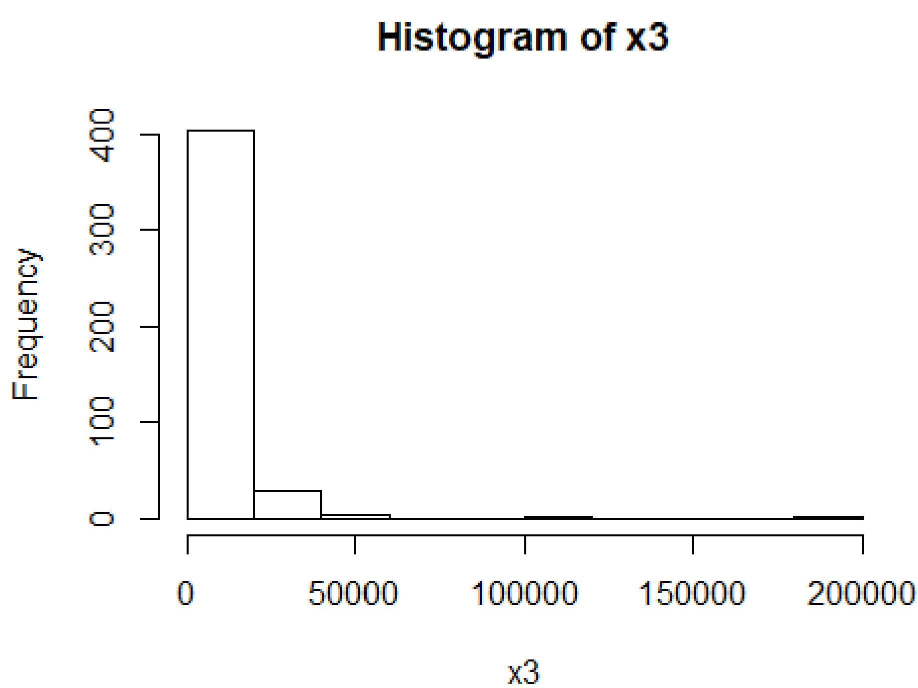

```
hist(x1)
```




```
hist(x2)
```


Histogram of x2
Frequency
0
50
100
200
0
5
10
15
20
25
30
35
x2

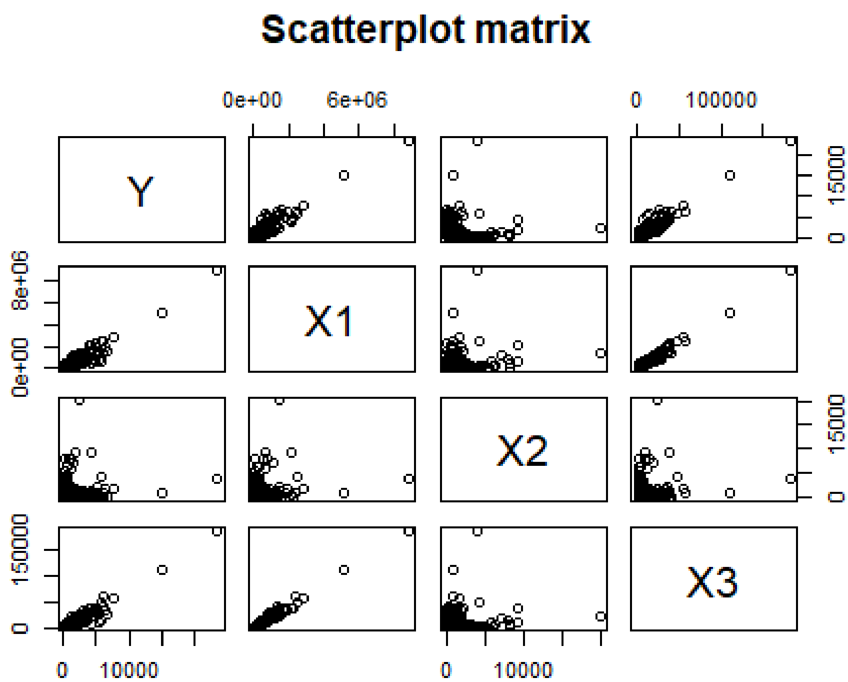
```
hist(x3)
```


Histogram of x3
Frequency
0
100
200
300
400
0
50000
100000
150000
200000
x3

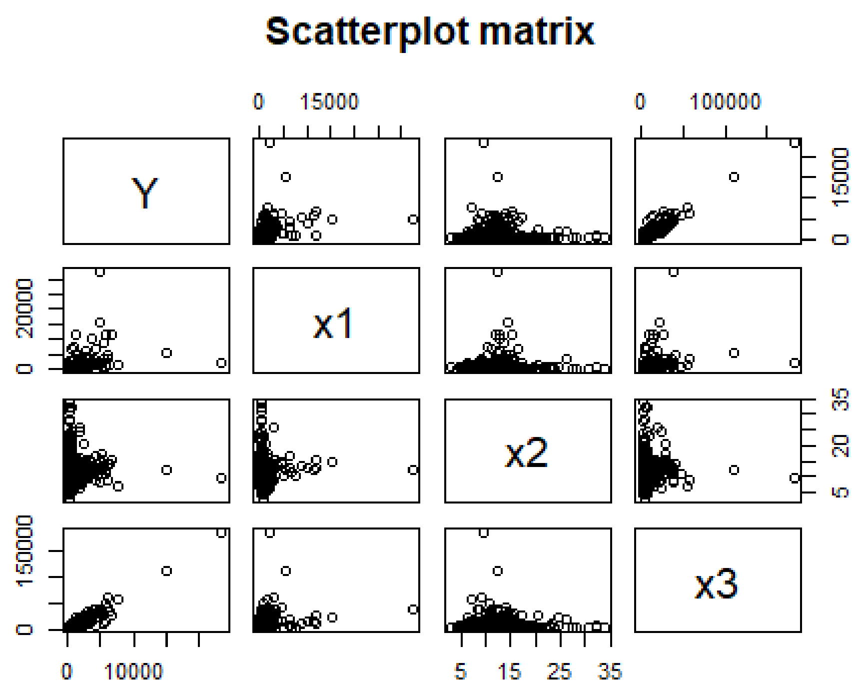
From the plots, we can see that the histogram plots for model-I are right-skewed but on the other hand, for model II, the histogram plot of percentage of population greater than 64 years old is approximately normal while the two histogram plots for model II are right-skewed.

(b)

```r
# Scatterplot matrix for model I
pairs(~Y+X1+X2+X3, data = cdi, main = "Scatterplot matrix")
```




```r
# Scatterplot matrix for model II
pairs(~Y+x1+x2+x3, data = cdi, main = "Scatterplot matrix")
```




```
# Correlation matrix for model I
C1 = cbind(X1,X2,X3)
cor(C1)

##           X1        X2        X3
## X1 1.0000000 0.1730834 0.9867476
## X2 0.1730834 1.0000000 0.1270743
## X3 0.9867476 0.1270743 1.0000000

# Correlation matrix for model II
C2 = cbind(x1,x2,x3)
cor(C2)

##            x1          x2          x3
## x1 1.00000000  0.02918445  0.31620475
## x2 0.02918445  1.00000000 -0.02273315
## x3 0.31620475 -0.02273315  1.00000000
```

From the scatterplots, we can actually determine the relations between response variables and predictor variables and also between predictor variables and predictor variables. From the scatterplot of modelI, we can see that the number of active physicians vary linearly with total population and total personal income while the same has no specific relation relation with land area. Clustering is in the starting smaller parts. From the scatterplot of modelII, we can observe that the number of active physicians vary linearly only with the total personal income and there is no specific relation between no of active physicians with population density and no of people greater than 64 years in age. Most of the data are clustered in the starting segments. From the correlation matrix in modelI, we can see that total population and total personal income are strongly correlated and from the correlation matrix in modelII, we can see that there is a weak correlation between predictor variables.

(c)

```
#Fitting the first order regression model for model I
model1 = lm(Y~X1+X2+X3)
model1

##
## Call:
## lm(formula = Y ~ X1 + X2 + X3)
##
## Coefficients:
## (Intercept)           X1           X2           X3
##  -1.332e+01    8.366e-04   -6.552e-02    9.413e-02

#Fitting the first order regression model for model I
model2 = lm(Y~x1+x2+x3)
model2

##
## Call:
## lm(formula = Y ~ x1 + x2 + x3)
##
## Coefficients:
## (Intercept)           x1           x2           x3
##  -170.57422      0.09616      6.33984      0.12657
```

Thus, the estimated regression function for modelI is -13.362+0.0008366X_1-0.065523X_2+0.09413X_3 and the estimated regression function for modelII is -170.574+0.09616X_1+6.33984X_2+0.1265X_3.

(d)

```
# Model I
yhat1 = model1$fitted.values
rsq1 = sum((yhat1 - mean(Y))^2) / sum((Y - mean(Y))^2)
rsq1

## [1] 0.9026432

# Model II
yhat2 = model2$fitted.values
rsq2 = sum((yhat2 - mean(Y))^2) / sum((Y - mean(Y))^2)
rsq2

## [1] 0.9117491
```

Thus, the value of R^2 for model I and mdoel II are 0.9026 and 0.9117 respectively. And, yes, model II explained more variance in the dependent variable than model I and thus is more preferrable.

(e)

```
#Residuals for model I
e1 = Y - yhat1
e1

##            1            2            3            4            5 
##  -800.6780229  515.4374974  144.3080224 -502.0168730 -1426.2901812 
##            6            7            8            9           10 
##  -684.7626779 -443.0294891 -1360.9624485 1544.2440120  -423.8873542 
##           11           12           13           14           15 
##  2841.0124665  797.5951318 -605.4050127  414.6857947  1679.0300557 
##           16           17           18           19           20 
##   723.2396545 1622.3368033 -1030.7944467 1263.2121638  -506.3677568 
##           21           22           23           24           25 
## -1143.7386407  434.8731762 -1001.9535363 -1305.4463657  441.8063320 
##           26           27           28           29           30 
##  -186.2090800  588.2759550 -1855.6061839  350.6304732  196.1630218 
##           31           32           33           34           35 
##   336.3725716 1141.9644241  801.1912369 -925.0844361  205.4556199 
##           36           37           38           39           40 
##  -793.2556524 -299.2196182   71.4885924 -740.8141266  426.9275152 
##           41           42           43           44           45 
##    60.0217210 -1186.0838861  860.8489737 -781.9474635  871.5151300 
##           46           47           48           49           50 
##  -431.6550305  385.0961473 1903.8340727 -419.6122684 3650.8200223 
##           51           52           53           54           55 
##  -655.2217850  495.9573369 2227.3283527 -1263.4377467  442.7440609 
##           56           57           58           59           60 
##    91.5568522 -786.5672710  897.8965969 -267.2108316  -79.8414346 
##           61           62           63           64           65 
##  -465.6868525 -769.9678548 -606.3505759    5.3918349  944.7007441 
##           66           67           68           69           70 
##   455.2265167 3689.0318671  930.9682227 -448.9951934 1479.1473259 
##           71           72           73           74           75 
##   108.8388711  721.5758503 1835.1289457  478.7505792  400.7389421 
##           76           77           78           79           80 
##  -263.1820444 -345.8902146 1068.2615798 -145.1832675 -117.8814126 
##           81           82           83           84           85 
##  -544.7041792 -371.6126273 -278.3762666 -469.0039165  -38.3316199 
##           86           87           88           89           90 
##  -669.2002360 -627.2121311  -75.3158165 -274.1808170  974.0406700 
##           91           92           93           94           95 
##  -684.6233972 -126.6496926  -81.1059375 -222.4729783 1333.9692908 
##           96           97           98           99          100 
##  -251.2116079  200.8083967 -330.2374586  480.1594247 -375.9442506 
##          101          102          103          104          105 
##   349.5659208 1523.9228936 -550.1349789  264.4093086 -243.1129170 
##          106          107          108          109          110 
##  -459.6841264  173.6895574 -641.3736247 -217.5227987 -618.4775675 
##          111          112          113          114          115 
```

```
##  -612.4199864  -649.7198506  -294.7709547  -617.7076351  -348.5233303
##           116           117           118           119           120
##  -455.6045167  -411.6319604   520.1648979  -210.0878816   -61.3127181
##           121           122           123           124           125
##  -354.2444476  -259.6082578  3198.1664897  -401.1169036  -275.0648033
##           126           127           128           129           130
##  -508.1050801  -174.1711953  -215.0313805   -73.6231269   114.4954054
##           131           132           133           134           135
##  -534.3227798  -275.4193874  -172.1815385  -302.4793104    -3.4355447
##           136           137           138           139           140
##  -246.1245702   693.8356105   121.8785117  -564.6835970  -203.2438194
##           141           142           143           144           145
##    46.3223146  -642.3343151   681.3994000  -224.2648514  -289.5409441
##           146           147           148           149           150
##  -374.8212608  -267.5297402  -337.7525386   200.4194523  -300.8438591
##           151           152           153           154           155
##  -115.5506886   -48.7050949   -91.3644550  -393.1528965    -5.4373046
##           156           157           158           159           160
##  -431.4565062  -284.6756249  -178.0483265  -184.8613557   738.7921737
##           161           162           163           164           165
##   481.4197543  -351.4935231    -5.9468959  -112.7121860  -334.0813513
##           166           167           168           169           170
##   362.0083707    57.6770814  1404.3525151   153.8842120    79.7732328
##           171           172           173           174           175
##  -234.0658439  -159.0458246  -202.9317523  -120.2768303  -525.8789591
##           176           177           178           179           180
##  -216.5844605  -399.5984799  -298.7427051   495.0518032    64.9965096
##           181           182           183           184           185
##   -93.3537165    14.7285480  -424.3608842   -69.8540619  -234.7090152
##           186           187           188           189           190
##   -21.8019881   495.7617156  -113.0950041  -157.7463001  -156.5388501
##           191           192           193           194           195
##   307.6385278   544.9971436   -81.1220552   -69.1058734  -294.5020970
##           196           197           198           199           200
##   379.6151725  -136.4227511  -146.6771792  -307.6181045   107.7607221
##           201           202           203           204           205
##  -135.1412583   -71.2652817  -161.8927356  -264.9403712   262.0592436
##           206           207           208           209           210
##    42.7735533  -360.0701259  -196.9680319   -68.8092991  -258.1315756
##           211           212           213           214           215
##   -85.0939048  -275.2369374   681.5943021  -128.8980226  -266.1611190
##           216           217           218           219           220
##    92.6071242   217.3997942  -244.0388034   -60.6552118   444.4187210
##           221           222           223           224           225
##   120.1343082    -8.3183828  -366.0826834  -318.5310799   -79.7123015
##           226           227           228           229           230
##  -331.9162901   -18.4756015  -113.2684895  -151.8920610   -34.7091015
##           231           232           233           234           235
##    48.0618934   131.8285623   -86.6264248  -185.8321847   373.7662387
##           236           237           238           239           240
```

```
##   -92.8438525   -64.3447676  -164.5099069    11.4256757  -224.0971655 
##           241           242           243           244           245 
##   192.3178877  -190.2330205   623.1681594   -41.3671982    94.5649535 
##           246           247           248           249           250 
##  -194.8643518  -268.1608735    82.3849295  -189.5749215     5.7356403 
##           251           252           253           254           255 
##  -182.5105535  -317.4092821   166.3595023  -238.7433607  -222.2480291 
##           256           257           258           259           260 
##    37.3931921  -215.7465891  1494.9832334   824.1678911  -191.8810737 
##           261           262           263           264           265 
##   -24.2049886   -59.0279654   191.7696107  -236.7131683  -180.8769311 
##           266           267           268           269           270 
##   100.4101631  -117.2769223  -206.3493050    39.7034913   -38.3038030 
##           271           272           273           274           275 
##  -282.4303624     0.4433911   -15.7989288   -46.1028720  -221.9515841 
##           276           277           278           279           280 
##    12.9228637   -86.0744913    20.5454633    -7.4282722  -164.8539812 
##           281           282           283           284           285 
##     0.8514891   166.1216326  -131.8136270   -53.6656203  -151.0586070 
##           286           287           288           289           290 
##    69.3508887  -125.3625819   -43.4861311  -169.4909038  -118.4221944 
##           291           292           293           294           295 
##  -246.6580824  -144.0020720  -134.9786045   -65.5447599  -181.9770908 
##           296           297           298           299           300 
##  -121.5938245    47.1494074    59.0544155  -179.0223852  -230.2103795 
##           301           302           303           304           305 
##   -98.6171383    94.4764316  -101.5244237  -160.7050142    14.3339393 
##           306           307           308           309           310 
##  -123.5796269  -142.5747750   180.1153054   167.0328529    27.0129825 
##           311           312           313           314           315 
##   121.6108928  -254.0542522  -157.5630236  -124.7414660   -56.3062154 
##           316           317           318           319           320 
##    -2.6018789   -57.1755356   -81.5077977    84.4097351   -67.6263557 
##           321           322           323           324           325 
##  -187.9214299    16.7628750   -58.2167056   -74.7589153    38.6159831 
##           326           327           328           329           330 
##   -61.5369438   -66.3842856  -137.9579561  -105.0506707    69.4464902 
##           331           332           333           334           335 
##  -107.7151984  -173.6131393   -86.0711844   175.2404678  -117.9326775 
##           336           337           338           339           340 
##  -196.6094676   116.2247614   139.6472961   387.1146924    54.9251236 
##           341           342           343           344           345 
##  -137.8465102  -229.9414697  -123.4057365   -26.0361986   -30.0131404 
##           346           347           348           349           350 
##   -76.7081910  -110.3710143  -169.3200482  -146.2496081  -145.1958686 
##           351           352           353           354           355 
##    55.8218922   -41.9320193  -158.2875076  -107.8649196   -61.1882704 
##           356           357           358           359           360 
##   122.8046301   -20.0028693  -159.2743165  -173.9534706   161.4552682 
##           361           362           363           364           365 
```

```
##    -2.1255604  -154.0298539   -25.2025759  -117.1186250   -86.9336301
##           366           367           368           369           370
##   -66.7120151   -93.3540241   169.0956104    28.9057173   -21.5124763
##           371           372           373           374           375
##   -20.8220514    -3.2213422   173.8517775   -39.4457349    36.0484935
##           376           377           378           379           380
##    50.6128948   -85.3375051   -74.4521503   -74.4783961   203.9907062
##           381           382           383           384           385
##   -61.2793944    27.2766017  -208.4643187    12.8470308   -11.6499804
##           386           387           388           389           390
##    19.9469381   -72.8105368   168.6762385    31.3884229   -99.8710424
##           391           392           393           394           395
##   -10.9194535   528.5063685    93.6815688  -123.5463615   -21.7556583
##           396           397           398           399           400
##   241.5601861   -73.9834687   -21.9982666   -30.9561949    78.6870683
##           401           402           403           404           405
##   -95.1488613   -96.3401386   -54.9137295   128.0422224   -26.4159808
##           406           407           408           409           410
##   -58.1248140   -90.7417731   -93.2855912   300.6861488  -171.6407136
##           411           412           413           414           415
##   -58.9915395   430.7587087   -55.9117764   -59.9500228   287.5854315
##           416           417           418           419           420
##   -72.0060052  -102.7889942  1575.4436830   -20.1896239   -31.9045732
##           421           422           423           424           425
##    -5.9739923   -36.9017809   -56.2983565   -56.5174791   -79.2994529
##           426           427           428           429           430
##  -102.5020062  -113.2637389  -103.0180993   228.3865345   -54.9073365
##           431           432           433           434           435
##  -173.5040223   -90.2713861   -31.3179387  -108.4031062  -158.0389082
##           436           437           438           439           440
##   -74.4021914   -99.3070734   -72.9820103    22.4794857   -65.2826263
```

```
#Residuals for model II
e2 = Y - yhat2
e2
```

```
##             1             2             3             4             5
##   258.8136730   686.2731418   560.2895393  -243.6790818 -1563.5533974
##             6             7             8             9            10
## -3055.7527132  -456.6957506 -1083.1459821  1890.9742682  -290.7005108
##            11            12            13            14            15
##  2230.2902607   775.5032344  -670.2285884  -372.9514039  1626.5445331
##            16            17            18            19            20
##   605.9275512  1667.0466356 -1075.5368926   634.5899287  -480.0564398
##            21            22            23            24            25
## -1156.7320030   751.7296586 -1089.4840866 -1123.1532029   315.1063742
##            26            27            28            29            30
##   -31.4063560   529.1097567 -1951.9848839   489.1888544   296.5805238
##            31            32            33            34            35
##   239.4881474   771.4414698   789.2961874 -1122.0573886   139.2056186
```

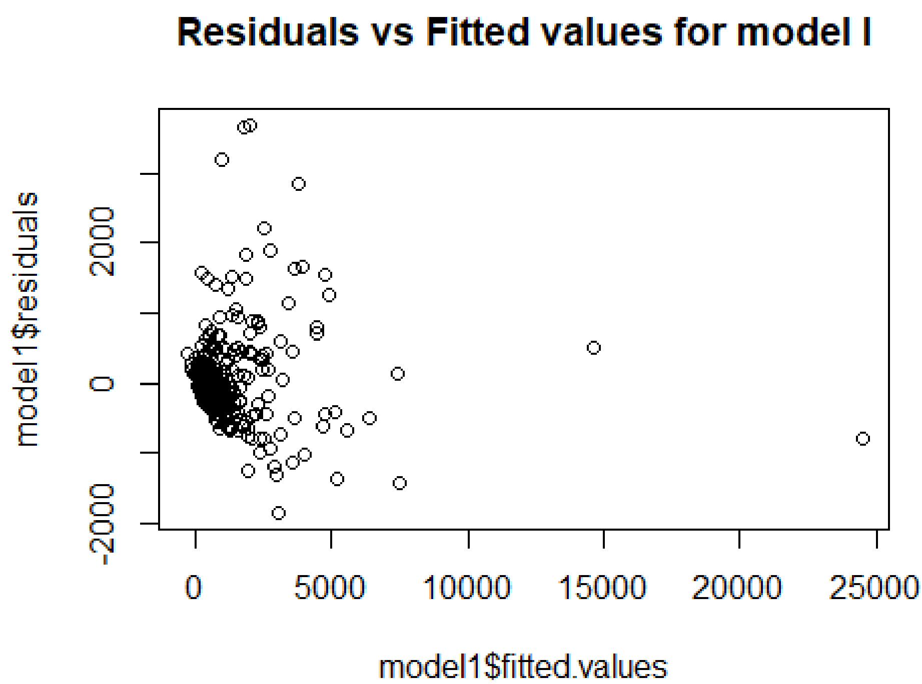
```
##           36           37           38           39           40 
##  -996.0720039  -263.3441870   247.4948888 -1011.8107856   560.9150965 
##           41           42           43           44           45 
##  -406.5250371 -1379.8431809   851.4282243  -841.2566943   906.9166399 
##           46           47           48           49           50 
##  -601.2816866  -108.2368674  1711.6753351  -683.8898611  3045.8079345 
##           51           52           53           54           55 
##  -570.5658780   701.0808218   744.1169218 -1246.0129018   469.5677249 
##           56           57           58           59           60 
##   153.4941528  -816.2395548   734.7404159  -134.4755644    52.7611820 
##           61           62           63           64           65 
##  -591.7698585  -799.7972654  -569.5674282  -100.6098087   635.4931441 
##           66           67           68           69           70 
##   479.7317243  2740.6200137  1021.7694013  -601.9818952  1483.0162391 
##           71           72           73           74           75 
##   190.2901097   592.8575609   993.6571643   588.1249203   445.2622154 
##           76           77           78           79           80 
##    -6.3861268  -216.1226481  1106.3318244   -15.6841336   -52.7865711 
##           81           82           83           84           85 
##  -640.2366694 -1298.0426795  -479.6926405  -494.2465624  -307.7362145 
##           86           87           88           89           90 
##  -649.4843285  -689.9216580   -32.0621508  -223.8658703  1031.2086335 
##           91           92           93           94           95 
##  -623.2277594   -58.4612439  -137.8919302  -130.4046498  1281.4959815 
##           96           97           98           99          100 
##  -664.0985018    57.6023608  -206.8428384   583.8098795  -275.2119119 
##          101          102          103          104          105 
##   412.7124973  1344.9676097  -501.9796641   305.8630425  -191.6493221 
##          106          107          108          109          110 
##  -575.6696651   231.7713225  -616.8810818  -209.8850750  -548.9256570 
##          111          112          113          114          115 
##  -574.6988812  -678.9905843  -183.6155374  -600.7308039  -273.6421229 
##          116          117          118          119          120 
##  -391.5745212  -516.9475129   545.7642894  -170.3076078     4.6704746 
##          121          122          123          124          125 
##  -301.4842673  -194.6711887  2729.7141107  -375.9429238  -220.6349264 
##          126          127          128          129          130 
##  -455.5277930  -181.7291285   -37.9358261    12.8206940   226.7003407 
##          131          132          133          134          135 
##  -530.6129405  -245.4171547   -78.3225670  -250.6728966   -69.9375101 
##          136          137          138          139          140 
##  -173.1895286   751.9459308   170.6754744  -476.6392337  -263.0710054 
##          141          142          143          144          145 
##    36.9810314  -553.7003959   757.4483110  -175.0911662  -183.9047374 
##          146          147          148          149          150 
##  -326.6817281  -236.1094062  -305.0880547   271.9874162  -317.7591885 
##          151          152          153          154          155 
##   -77.0427035  -129.7597410   -18.0822672  -345.2643807  -126.8007793 
##          156          157          158          159          160 
##  -367.8154847  -223.1505925  -144.4870656  -121.8675151   834.5766378 
```

```
##         161         162         163         164         165
##   502.9337431 -264.0075312   99.7813147  -99.7587925 -278.2987933
##         166         167         168         169         170
##   453.3822449  113.1364632 1456.8582217   16.3401919  172.9629804
##         171         172         173         174         175
##  -258.6295126 -158.5245472 -325.7373601  -51.5055619 -451.2890833
##         176         177         178         179         180
##   -67.2224964 -290.9631247 -214.1492111  529.5133110   -8.1713431
##         181         182         183         184         185
##    -0.5547246   56.9364880 -356.3173448   21.8988860 -157.6743399
##         186         187         188         189         190
##    67.1108826  197.2396105   25.5506853 -125.9461887 -109.3971964
##         191         192         193         194         195
##    17.0305328  625.1995823  -24.5376582  -33.6772942 -242.7511841
##         196         197         198         199         200
##   443.8351941 -102.5291619  -91.9612861 -256.1508271  134.8039160
##         201         202         203         204         205
##  -199.6381031  -22.4305852  -89.6210582 -182.9648531  308.1471186
##         206         207         208         209         210
##   -42.2367450 -306.3976116 -162.0237401  -25.9860768 -196.7931257
##         211         212         213         214         215
##   -14.0539674 -176.3506137  728.7008950  -65.7083448 -263.2510345
##         216         217         218         219         220
##   161.9953299  304.5210401 -191.8774727  -42.1636072  513.3192722
##         221         222         223         224         225
##    25.5054663   47.9742596 -274.6227443 -300.5386534  -15.0937996
##         226         227         228         229         230
##  -243.0442639   21.7445551  -49.1838083 -195.2666565   28.2348263
##         231         232         233         234         235
##   108.9634949  166.2139309  -36.4217189 -157.7807854   79.5823696
##         236         237         238         239         240
##   -49.2807144  -86.7386093  -99.6545867  103.7598545 -171.6826282
##         241         242         243         244         245
##   283.0050153 -118.7636616  700.1326709   10.6023448  -52.4927243
##         246         247         248         249         250
##   -79.9197714 -196.7311998  102.3369369 -227.7664273  -22.9167994
##         251         252         253         254         255
##  -137.8764963 -265.0916807  208.8154879 -169.7317410 -157.5747155
##         256         257         258         259         260
##    30.2978976 -170.2276094 1543.9222522  902.5663296 -133.7975222
##         261         262         263         264         265
##    30.4381107  -26.9226695  220.7194182 -172.9000757 -130.2245693
##         266         267         268         269         270
##   137.2400683   -4.7317165 -213.3749541  108.7495819  -29.4227561
##         271         272         273         274         275
##  -190.8968515 -572.4810717 -104.6794230   15.8063211 -148.3512125
##         276         277         278         279         280
##   101.1885345  -61.8228427   36.5716327   68.5462624 -126.2694564
##         281         282         283         284         285
##    31.5363991  189.1596473  -75.9359947    4.8243688 -161.9975889
```

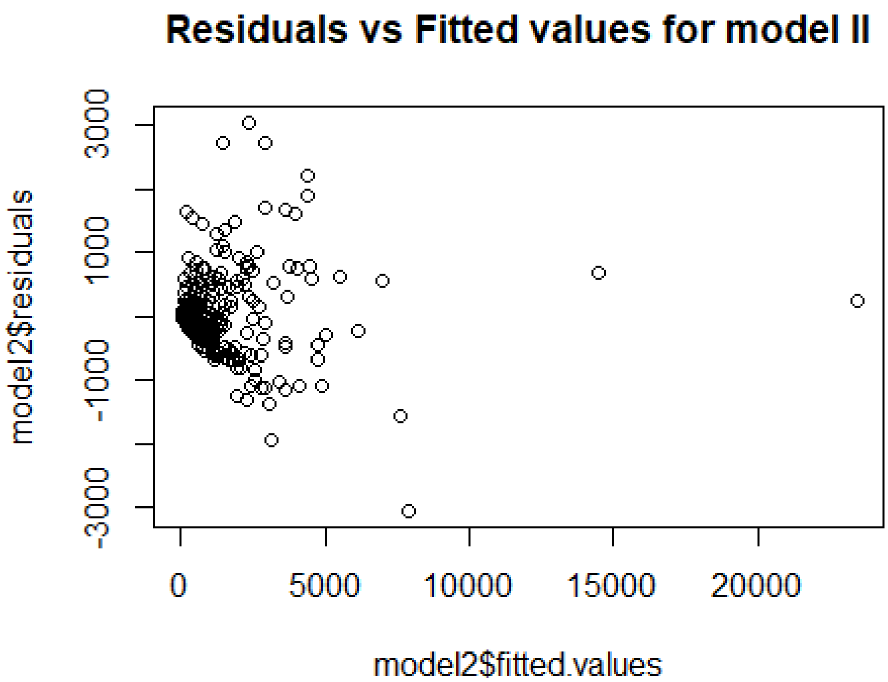
```
##          286          287          288          289          290 
##  111.9468605  -57.0564620   33.0993735 -100.2799493  -70.2426895 
##          291          292          293          294          295 
## -181.6890249 -134.2218557 -188.9123133 -191.7899199 -138.3974585 
##          296          297          298          299          300 
##  -39.0653006   81.6555173   98.4374753 -118.1256279 -143.1024159 
##          301          302          303          304          305 
##  -80.4642732  135.9846326   34.6266818  -99.3444466    8.5712400 
##          306          307          308          309          310 
##  -84.4164833  -90.8658435   41.5743538   73.5762290   87.1651878 
##          311          312          313          314          315 
##   37.1162002 -175.9085998 -105.8298250  -95.6838251    2.8001937 
##          316          317          318          319          320 
##   66.6262284  -15.0981635  -12.7160277  137.3141729  -47.2073887 
##          321          322          323          324          325 
## -105.2187899   95.6009340  -27.5149638  -50.6865279   83.0708318 
##          326          327          328          329          330 
##    1.5590967  -12.6901690  -65.3118089  -52.5733726   70.0313836 
##          331          332          333          334          335 
##  -55.0483062 -103.5025038  -34.5845173   87.1953230 -224.4152020 
##          336          337          338          339          340 
## -127.7910595   72.2994206    9.8296847  459.2248619   87.5469579 
##          341          342          343          344          345 
##  -72.3345686 -181.5244937  -70.8858546   55.2367816    7.7855807 
##          346          347          348          349          350 
##  -45.4468325  -37.8895055 -110.0923912  -87.7429860  -95.6877535 
##          351          352          353          354          355 
##   29.7018976   18.4407918  -91.3131097  -50.0633140  -71.2364737 
##          356          357          358          359          360 
##  167.9724253   38.2713666 -115.3020389 -122.0152464  107.1440964 
##          361          362          363          364          365 
##   48.5290346  -86.4241137   83.1881109  -59.1280022  -49.2637874 
##          366          367          368          369          370 
##  -38.1618682  -58.5546785   15.0416009   69.6033179    5.6480208 
##          371          372          373          374          375 
##   40.0883182   43.3071528   53.1276045   29.6697583   72.6420398 
##          376          377          378          379          380 
##   60.5406572   -0.7289711  -15.0627696  -37.0601739  -25.1927242 
##          381          382          383          384          385 
##    4.3855978   62.4192133 -142.5720933   16.6777555   69.1789896 
##          386          387          388          389          390 
##   80.2515486  -30.3027707  101.8774485   84.5961540  -34.9795087 
##          391          392          393          394          395 
##   -4.7450275  603.2706996   41.7493601  -70.8379912   45.7036242 
##          396          397          398          399          400 
## -401.4953482    0.2408944 -105.9335559    3.5653732   79.0323440 
##          401          402          403          404          405 
##  -52.3241972  -48.0240394   -2.2834495  -15.4706488   64.6437361 
##          406          407          408          409          410 
##   -4.8650241  -36.0602433  -52.4574781  371.3375822 -142.4890365 
```

```
##           411          412          413          414          415
##   -51.6612442  -54.1479393   -6.3620448  -61.6029302   43.7966233
##           416          417          418          419          420
##   -19.8551108  -49.2297990 1629.0761342   26.6511316   48.6915438
##           421          422          423          424          425
##    46.9220277   30.6681366   -7.4602092   23.8945205  -12.1050690
##           426          427          428          429          430
##   -64.8594516  -48.7513661  -32.7884736  233.9699767   30.8886907
##           431          432          433          434          435
##   -71.2551251  -39.4192139   18.1526001  -47.9130275  -72.0451376
##           436          437          438          439          440
##  -124.4791644 -174.6343004   22.1129299   47.5723066   -6.5861750
```

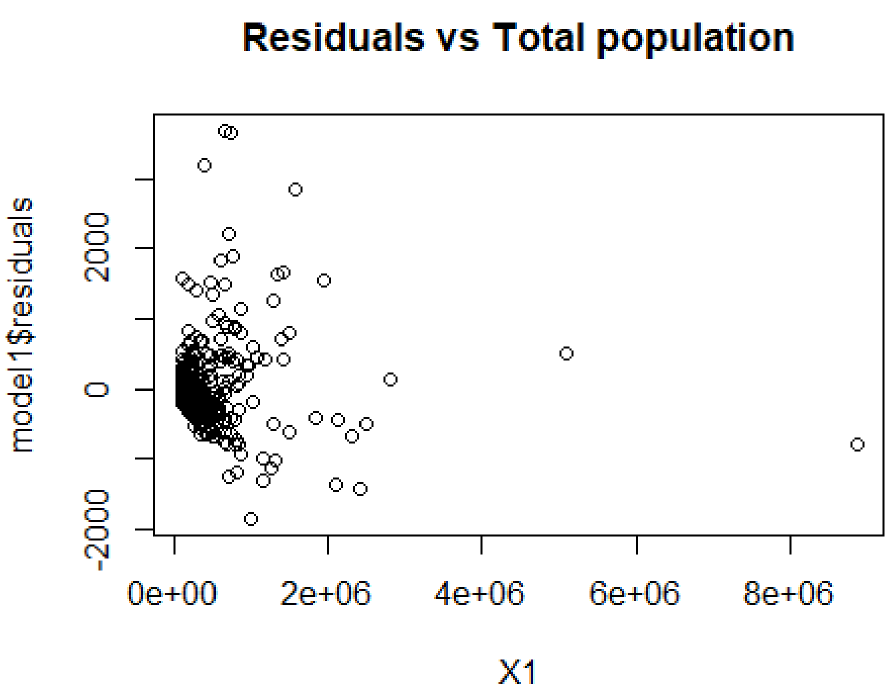
```
#Plots of residuals against fitted Y values
plot(model1$residuals ~ model1$fitted.values, main = "Residuals vs Fitted
values for model I")
```



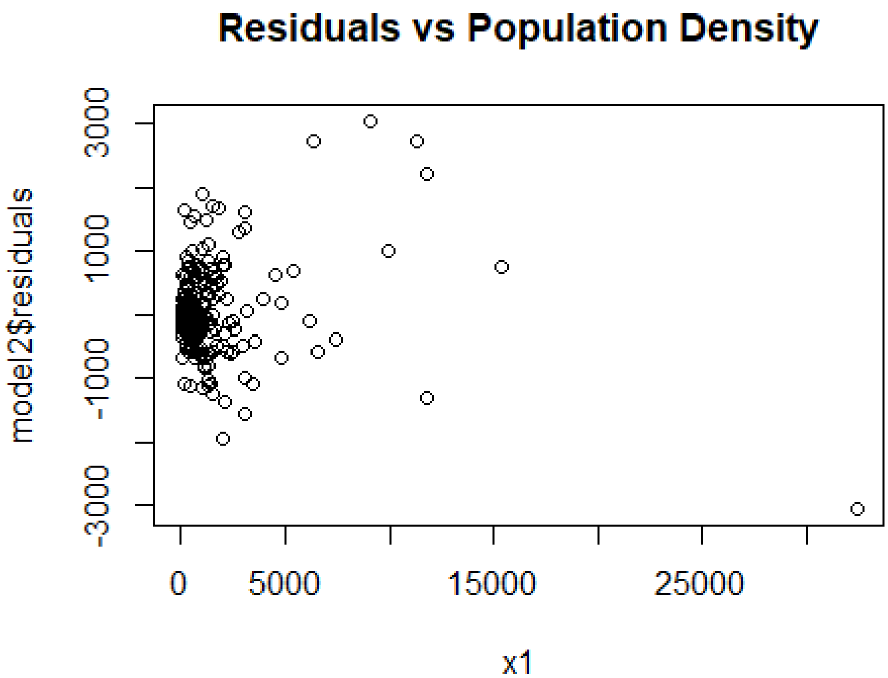

```
plot(model2$residuals ~ model2$fitted.values, main = "Residuals vs Fitted
values for model II")
```



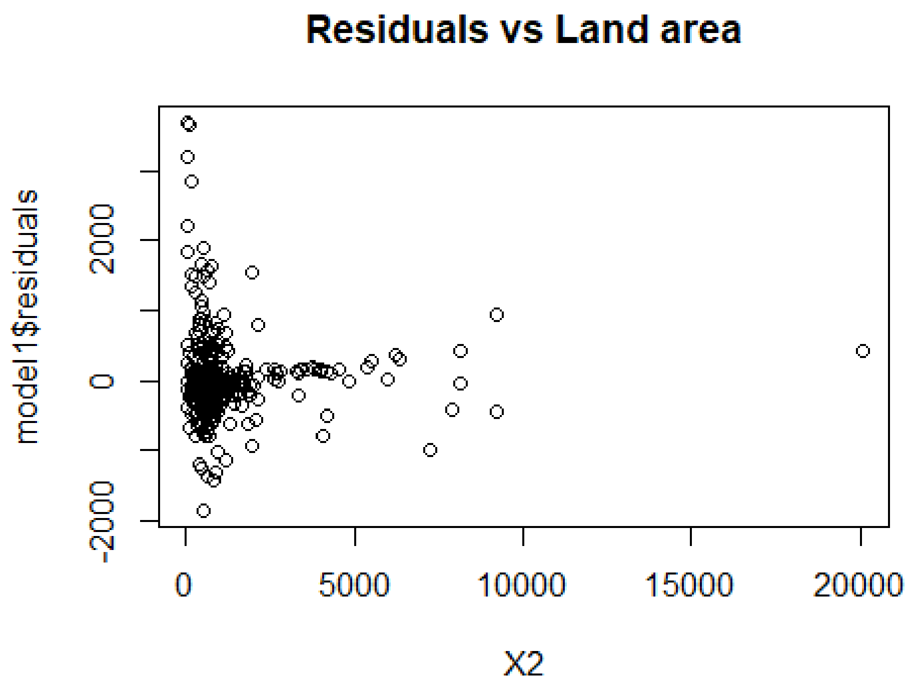

```r
#Plots of residuals against predictor variables X1 and x1
plot(model1$residuals~X1, main = "Residuals vs Total population")
```



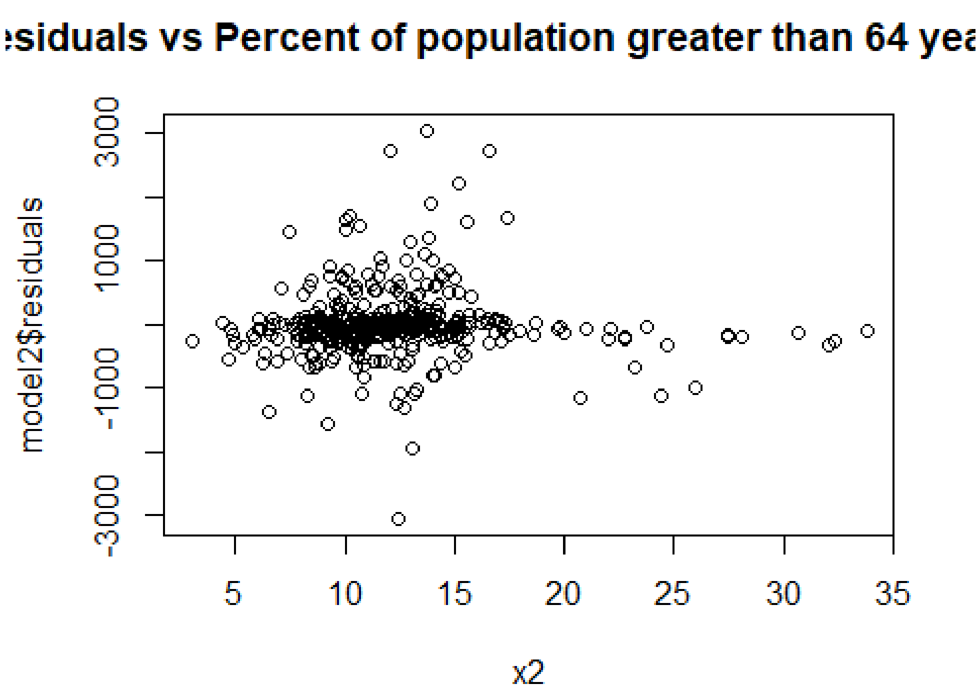
```r
plot(model2$residuals~x1, main = "Residuals vs Population Density")
```



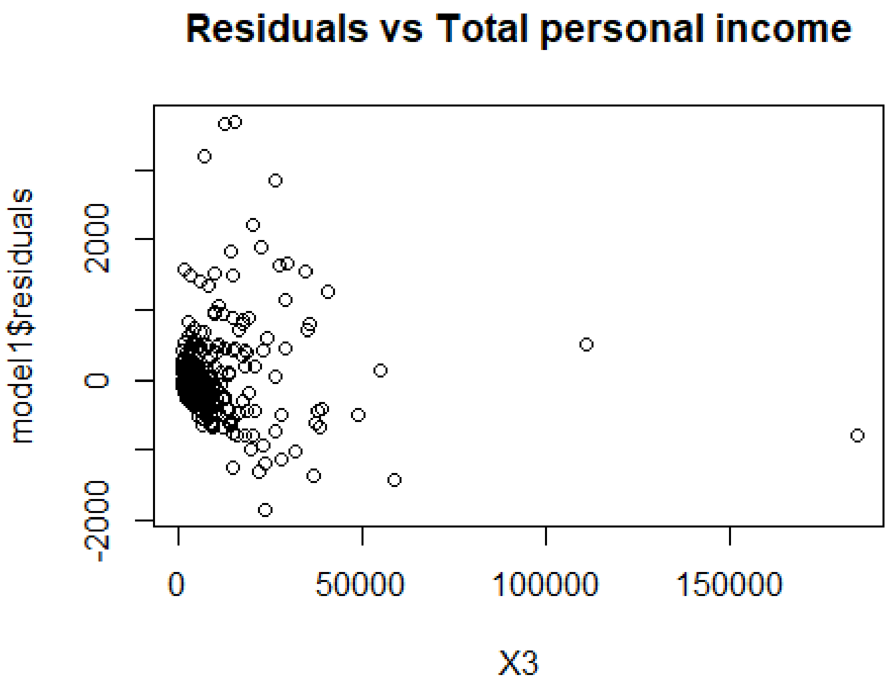

```r
#Plots of residuals against predictor variables X2 and x2
plot(model1$residuals~X2, main = "Residuals vs Land area")
```



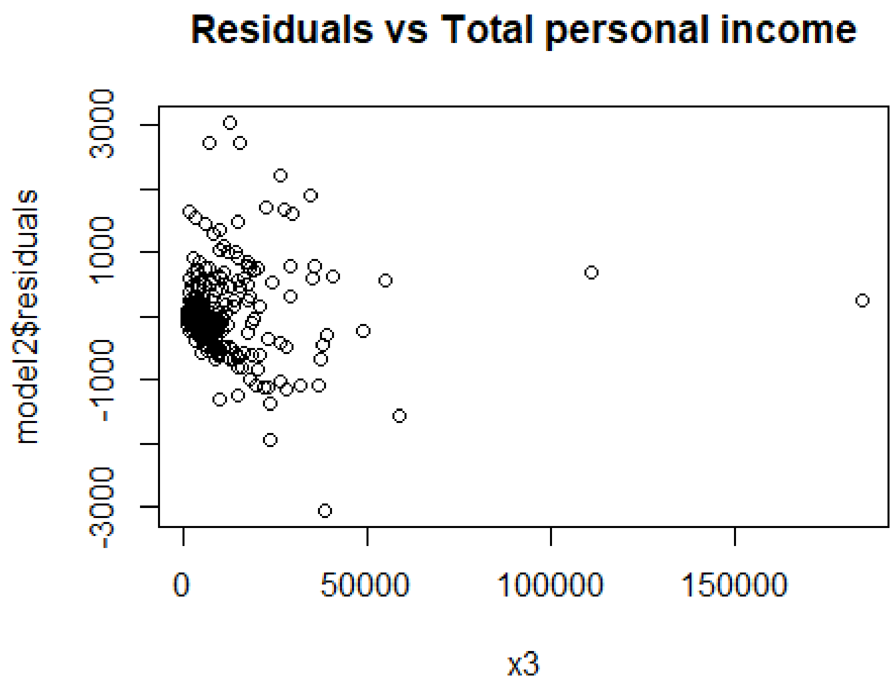

```r
plot(model2$residuals~x2, main = "Residuals vs Percent of population greater than 64 years old")
```



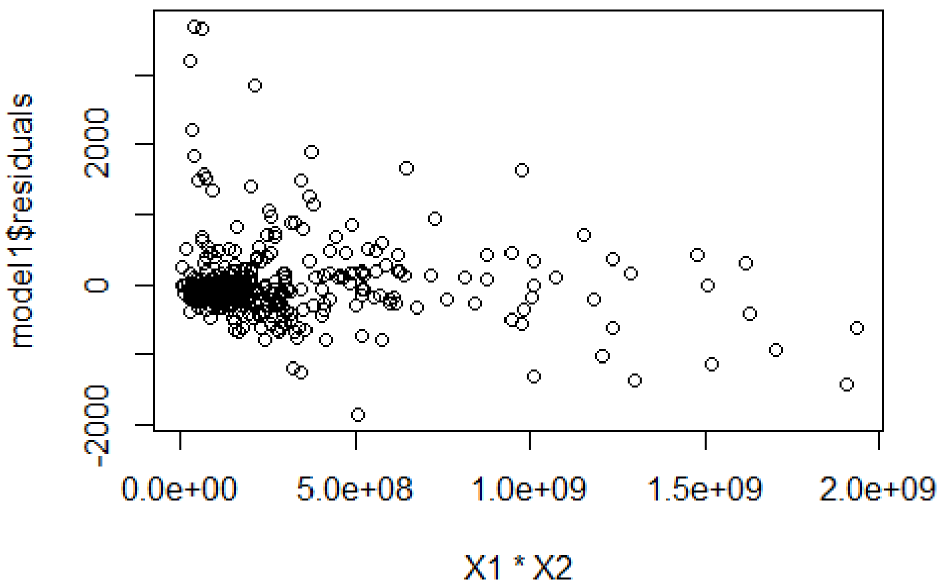
```r
#Plots of residuals against predictor variables X3 and x3
plot(model1$residuals~X3, main = "Residuals vs Total personal income")
```



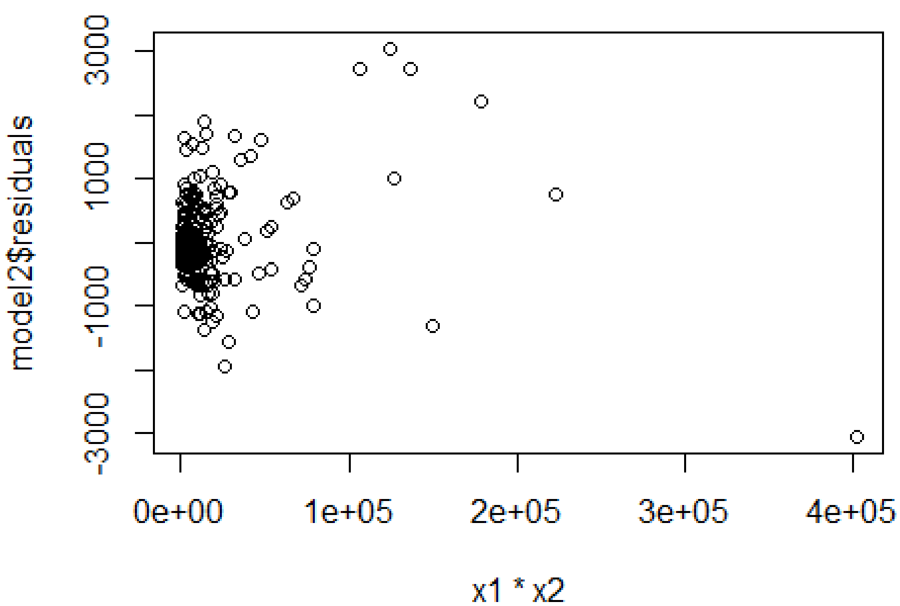

```r
plot(model2$residuals~x3, main = "Residuals vs Total personal income")
```



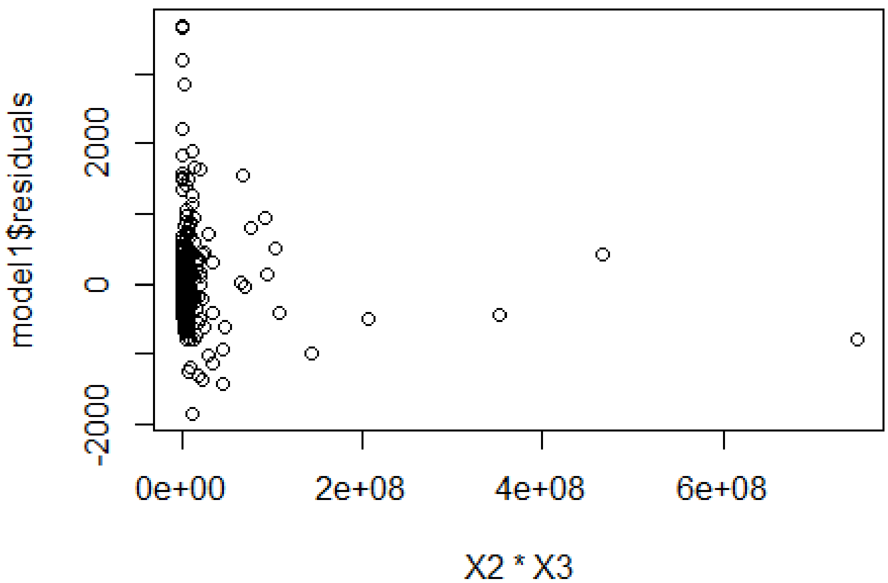

```
#Plots of residuals against predictor variables X1+X2 and x1+x2
plot(X1*X2,model1$residuals)

## Warning in X1 * X2: NAs produced by integer overflow
```



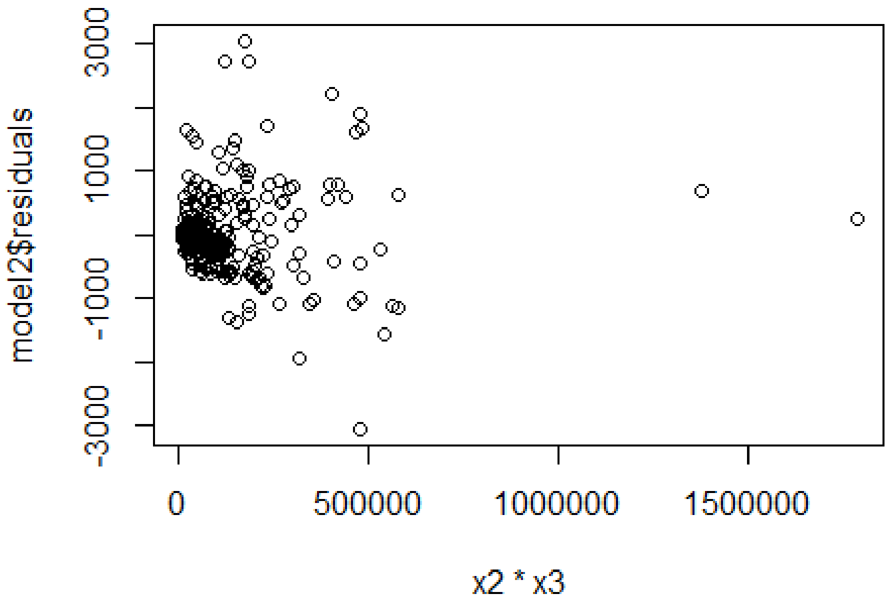

```
plot(x1*x2,model2$residuals)
```



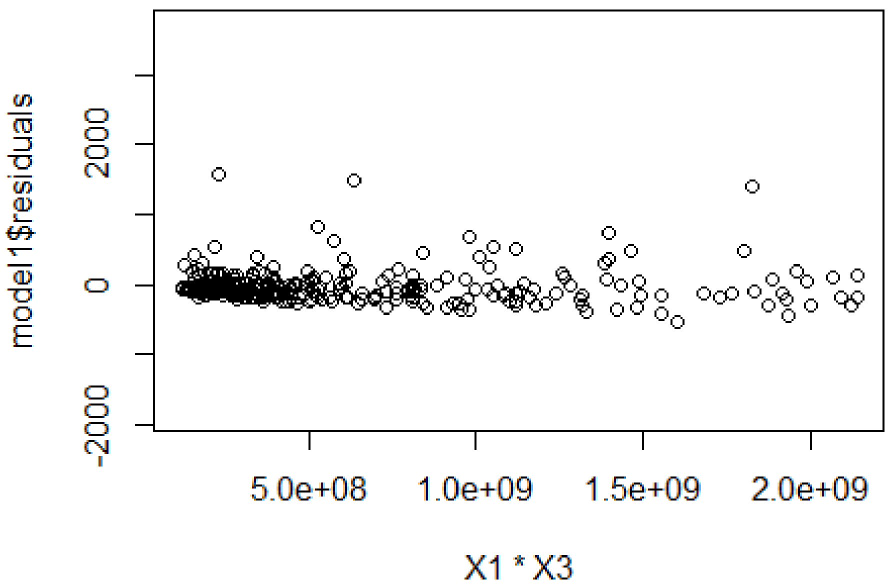
```r
#Plots of residuals against predictor variables X2+X3 and x2+x3
plot(X2*X3,model1$residuals)
```



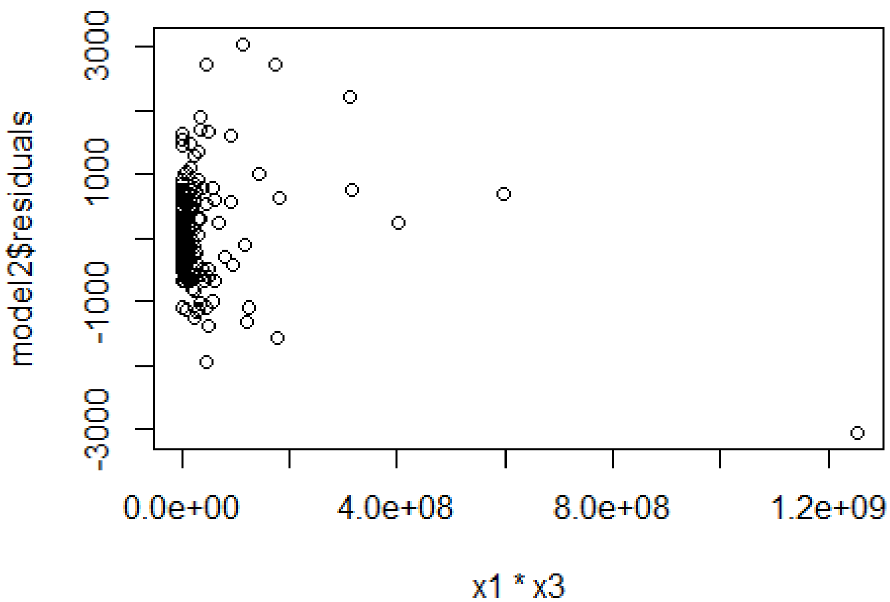

```r
plot(x2*x3,model2$residuals)
```



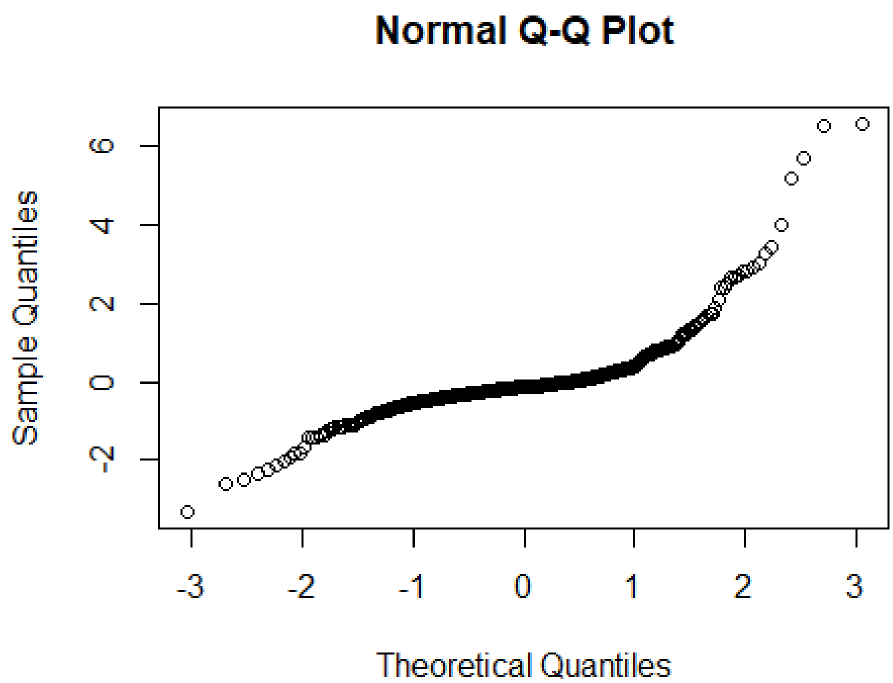

```
#Plots of residuals against predictor variables X1+X3 and x1+x3
plot(X1*X3,model1$residuals)

## Warning in X1 * X3: NAs produced by integer overflow
```



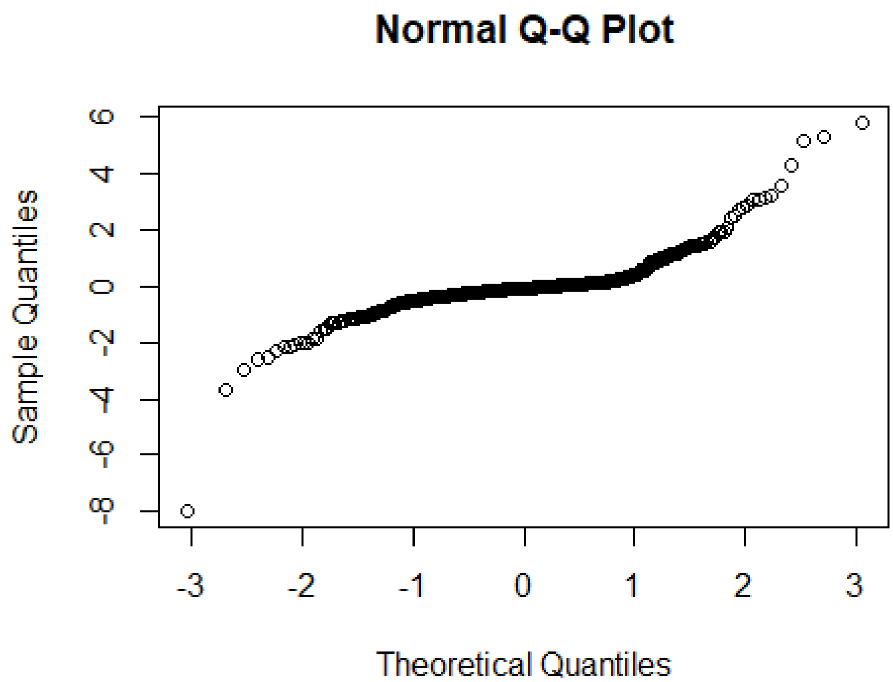

```
plot(x1*x3,model2$residuals)
```



```r
#Normal probability plot
#ModelI
plot1 = rstandard(model1)
qqnorm(plot1)
```




```r
#ModelII
plot2 = rstandard(model2)
qqnorm(plot2)
```



If we observe modelI, and if we specifically see the residuals plots against fitted values, there is no specific trend and there are some outliers as well. Variance and residuals are also not constant. From the plots of residuals vs predictors and two factor interactions, it can be seen that we can actually improve the model by including the variable X1*X2 and X1*X3 in our model. From the current predictor variables, the linear model doesn't seem to be appropriate.

If we observe modelII, and if we specifically see the residuals plots against fitted values, the variance for error term and residuals is not constant and there are some outliers as well. From the plots of residuals vs predictors and two factor interaction, the model seems to be appropriate with two factor interaction. Alos, total personal income varies constantly and shows normal distribution and there is no need to involve other variables. from the current predictors, the model is appropriate.

From the normal probability plot, we can see that the plot for both the models is increasing.

Overall, the model II is more appropriate as compared to model I.

Part 2

(a)

```
cdi = read.table("D:\\ASU Stuff\\SEM-1\\STP 530\\CDI_data.txt")

model1 = subset(cdi,cdi[,17]==1)
model2 = subset(cdi,cdi[,17]==2)
model3 = subset(cdi,cdi[,17]==3)
model4 = subset(cdi,cdi[,17]==4)

#Model1

Y1 = model1$V10
X1a = model1$V5 / model1$V4
X1b = model1$V15
X1c = model1$V11

mod1 = lm(Y1~X1a+X1b+X1c)
mod1

##
## Call:
## lm(formula = Y1 ~ X1a + X1b + X1c)
##
## Coefficients:
## (Intercept)          X1a          X1b          X1c
##  -64466.231       17.383       -1.406     1182.577

#Model2

Y2 = model2$V10
X2a = model2$V5 / model2$V4
X2b = model2$V15
X2c = model2$V11

mod2 = lm(Y2~X2a+X2b+X2c)
mod2

##
## Call:
## lm(formula = Y2 ~ X2a + X2b + X2c)
##
## Coefficients:
## (Intercept)          X2a          X2b          X2c
##  -4163.2673      33.6193       0.1024      -2.7616
```

```r
#Model3

Y3 = model3$V10
X3a = model3$V5 / model3$V4
X3b = model3$V15
X3c = model3$V11

mod3 = lm(Y3~X3a+X3b+X3c)
mod3

##
## Call:
## lm(formula = Y3 ~ X3a + X3b + X3c)
##
## Coefficients:
## (Intercept)          X3a          X3b          X3c
##   38862.667        5.537        1.957     -670.884

#Model4

Y4 = model4$V10
X4a = model4$V5 / model4$V4
X4b = model4$V15
X4c = model4$V11

mod4 = lm(Y4~X4a+X4b+X4c)
mod4

##
## Call:
## lm(formula = Y4 ~ X4a + X4b + X4c)
##
## Coefficients:
## (Intercept)          X4a          X4b          X4c
##  129323.415        5.717        4.342    -2159.920
```

Thus the estimated regression fucntion for model1 is $Y = -64466.231 + 17.383X_1 - 1.406X_2 + 1182.577X_3$, model2 is $Y = -4163.2673 + 33.6193X_1 + 0.1024X_2 - 2.7616X_3$, model3 is $Y = 38862.667 + 5.537X_1 + 1.957X_2 - 670.884X_3$ and model4 is $Y = 129323.415 + 5717X_1 + 4.342X_2 - 2159.920X_3$.

part (b)

The B_0 values are different for all the four models. It can be seen that B_1 values are similar for 3 and 4. B_2 and B_3 are different for all the four models.

part (c)

```r
summary(mod1)
```

```
##
## Call:
## lm(formula = Y1 ~ X1a + X1b + X1c)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -149353   -3776     339    5091  130462
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.447e+04  4.256e+04  -1.515   0.1330
## X1a          1.738e+01  8.336e-01  20.854   <2e-16 ***
## X1b         -1.406e+00  7.606e-01  -1.849   0.0674 .
## X1c          1.183e+03  6.411e+02   1.845   0.0681 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28060 on 99 degrees of freedom
## Multiple R-squared:  0.8352, Adjusted R-squared:  0.8302
## F-statistic: 167.2 on 3 and 99 DF,  p-value: < 2.2e-16
```

```
summary(mod2)
```

```
##
## Call:
## lm(formula = Y2 ~ X2a + X2b + X2c)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -148517   -4535   -1011    4321  257651
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4163.2673 49958.0969  -0.083    0.934
## X2a            33.6193     3.8553   8.720 4.79e-14 ***
## X2b             0.1024     1.6708   0.061    0.951
## X2c            -2.7616   809.2473  -0.003    0.997
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 32980 on 104 degrees of freedom
## Multiple R-squared:  0.5285, Adjusted R-squared:  0.5149
## F-statistic: 38.86 on 3 and 104 DF,  p-value: < 2.2e-16
```

```
summary(mod3)
```

```
##
## Call:
## lm(formula = Y3 ~ X3a + X3b + X3c)
##
## Residuals:
```

```
##    Min     1Q Median     3Q    Max
## -75116 -15583  -9111   3217 217684
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 38862.667  30662.721   1.267   0.2070
## X3a             5.537      2.377   2.329   0.0212 *
## X3b             1.957      1.096   1.787   0.0761 .
## X3c          -670.884    511.472  -1.312   0.1917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 36970 on 148 degrees of freedom
## Multiple R-squared:  0.09251,    Adjusted R-squared:  0.07411
## F-statistic: 5.029 on 3 and 148 DF,  p-value: 0.002392
```

```
summary(mod4)
```

```
##
## Call:
## lm(formula = Y4 ~ X4a + X4b + X4c)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -101554  -25633  -14428    3976  608081
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 129323.415  95555.937   1.353    0.180
## X4a              5.717      5.774   0.990    0.325
## X4b              4.342      2.731   1.590    0.116
## X4c          -2159.920   1353.579  -1.596    0.115
##
## Residual standard error: 81820 on 73 degrees of freedom
## Multiple R-squared:  0.08665,    Adjusted R-squared:  0.04912
## F-statistic: 2.309 on 3 and 73 DF,  p-value: 0.08351
```

Model1: MSE: 7.874 $R^2 = SSR/SSTO = 0.8352$

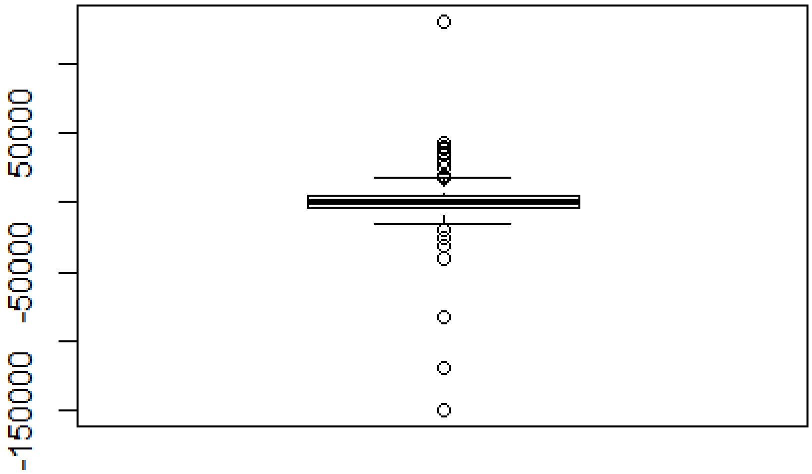
Model2: MSE: 1.08*10^9 $R^2 = 0.5280$

Model3: MSE: 1.365*10^9 $R^2 = 0.092$

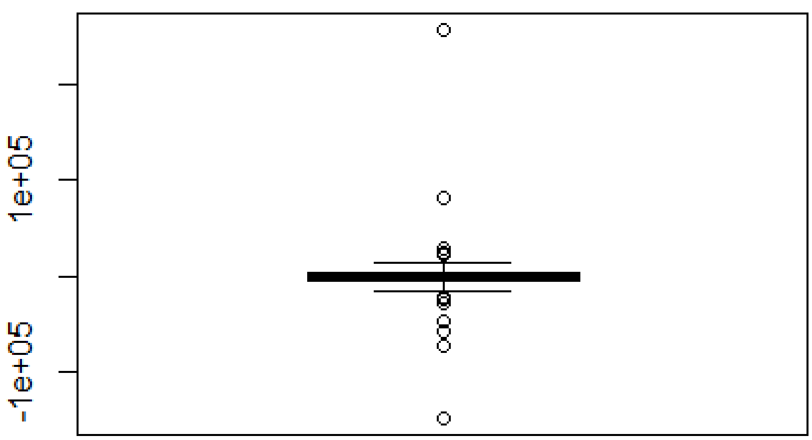
Model4: MSE: 6.69*10^9 $R^2 = 0.0867$

part (d)

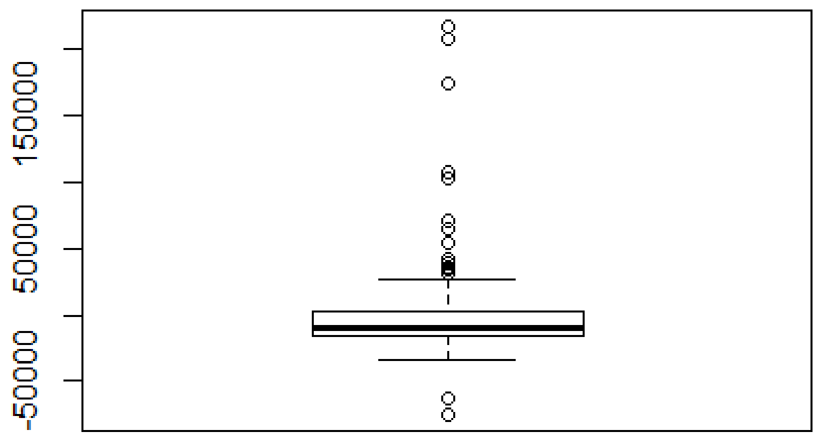
```
boxplot(resid(mod1))
```



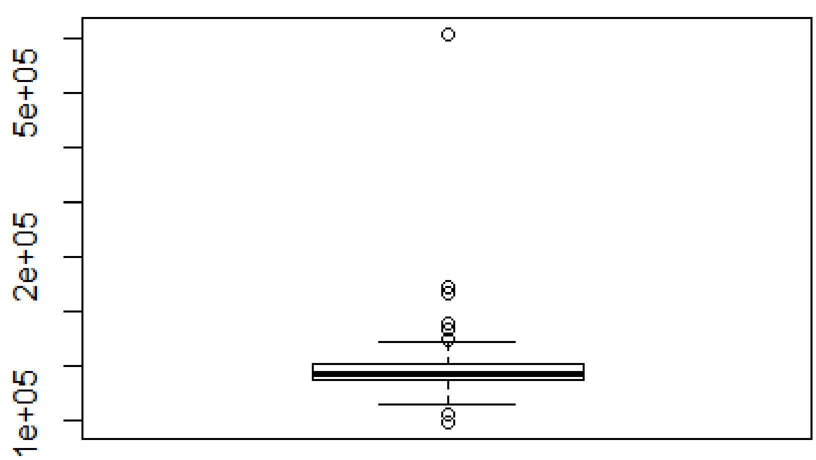

```r
boxplot(resid(mod2))
```




```r
boxplot(resid(mod3))
```




```
boxplot(resid(mod4))
```



Model1: In this model, we can see that there are a few outliers above as well as in the below region.

Model2: Majority of the outliers lie in the below region.

Model3: Majority of the outliers lie in the above region.

Model4: Majority of the otliers lie in the above region and one is specifically distant from each other.

Part-3 (10 points) :-

a) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 \log_{10} X_{i2} + \beta_3 X_{i1}^2 + \varepsilon_i$

By comparing this model with the general linear regression model, we can say easily say that it is not a general linear regression model.

But, it can be expressed in the form of (6.7) by making two transformations as follows:

$$X_{i3} = \log_{10} X_{i2} \quad \text{and} \quad X_{i4} = X_{i1}^2$$

$\Rightarrow Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i3} + \beta_3 X_{i4} + \varepsilon_i$

b) $Y_i = \varepsilon_i \exp(\beta_0 + \beta X_{i1} + \beta_2 X_{i2}^2)$

It is not a general linear regression model and also, it cannot be expressed in the form of (6.7) by a suitable transformation.

c) $Y_i = \log_{10}(\beta_1 X_{i1}) + \beta_2 X_{i2} + \varepsilon_i$

It is not a general linear regression model and also, it cannot be expressed in the form of (6.7) by a suitable transformation.

d) $Y_i = \beta_0 \exp(\beta_1 x_{i1}) + \varepsilon_i$

It is not a general linear regression model and also, it cannot be expressed in the form of (6.7) by a suitable transformation.

e) $Y_i = [1 + \exp(\beta_0 + \beta_1 x_{i1} + \varepsilon_i)]^{-1}$

It is not a general linear regression model and it can be expressed in the form of (6.7) by suitable transformations.

$Y_i' = \log_e\left(\frac{1}{Y_i} - 1\right)$

$\Rightarrow Y_i' = \beta_0 + \beta_1 x_{i1} + \varepsilon_i$

Scanned by CamScanner

Part-4 (10 points) :-

Multiple regression model :-

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i2} + \varepsilon_i \qquad i=1,\dots,n$$

$$\varepsilon_i \sim n(0, \sigma^2)$$

By transformation $X_{i3} = X_{i1}^2$, we can re-write the above equation as:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i3} + \beta_3 X_{i2} + \varepsilon_i$$

(a) least squares criterion :-

$$Q = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{i1} - \beta_2 X_{i3} - \beta_3 X_{i2})^2$$

The least square estimators are those values of $\beta_0, \beta_1, \beta_2, \beta_3$ that minimizes Q.

~~Least~~ least squares Normal Equations :-

Here,

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \qquad X = \begin{bmatrix} 1 & X_{11} & X_{13} & X_{12} \\ 1 & X_{21} & X_{23} & X_{22} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n3} & X_{n2} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

It can be represented in matrix terms as:-

$$\underset{n\times 1}{Y} = \underset{n\times 4}{X}\,\underset{4\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Now, if we differentiate Q w.r.t $\beta_0, \beta_1, \beta_2$ and β_3,

$$X'Xb = X'Y$$

$$\therefore \quad b = (X'X)^{-1}(X'Y)$$

Also, $X'X\,b = X'Y$ can be expressed in matrix as :-

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{13} & x_{23} & \cdots & x_{n3} \\ x_{12} & x_{22} & \cdots & x_{n2} \end{bmatrix} \begin{bmatrix} 1 & x_{11} & x_{13} & x_{12} \\ 1 & x_{21} & x_{23} & x_{22} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n3} & x_{n2} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{13} & x_{23} & \cdots & x_{n3} \\ x_{12} & x_{22} & \cdots & x_{n2} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

$$\therefore \begin{bmatrix} n & \Sigma_{i1} & \Sigma_{i3} & \Sigma x_{i2} \\ \Sigma x_{i1} & \Sigma x_{i1}^2 & \Sigma x_{i1}x_{i3} & \Sigma x_{i1}x_{i2} \\ \Sigma x_{i3} & \Sigma x_{i3}x_{i1} & \Sigma x_{i3}^2 & \Sigma x_{i3}x_{i2} \\ \Sigma x_{i2} & \Sigma x_{i2}x_{i1} & \Sigma x_{i2}x_{i3} & \Sigma x_{i2}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma x_{i1} Y_i \\ \Sigma x_{i3} Y_i \\ \Sigma x_{i2} Y_i \end{bmatrix}$$

Thus, least squares normal equations are:

$$\Sigma Y_i = nb_0 + b_1\Sigma_{i1} + b_2\Sigma_{i3} + b_3\Sigma x_{i2}$$

$$\Sigma x_{i1}Y_i = b_0\Sigma x_{i1} + b_1\Sigma x_{i1}^2 + b_2\Sigma x_{i1}x_{i3} + b_3\Sigma x_{i1}x_{i2}$$

$$\Sigma x_{i3}Y_i = b_0\Sigma x_{i3} + b_1\Sigma x_{i3}x_{i1} + b_2\Sigma x_{i3}^2 + b_3\Sigma x_{i3}x_{i2}$$

$$\Sigma x_{i2}Y_i = b_0\Sigma x_{i2} + b_1\Sigma x_{i2}x_{i1} + b_2\Sigma x_{i2}x_{i3} + b_3\Sigma x_{i2}^2$$

(b) Likelihood Function :-

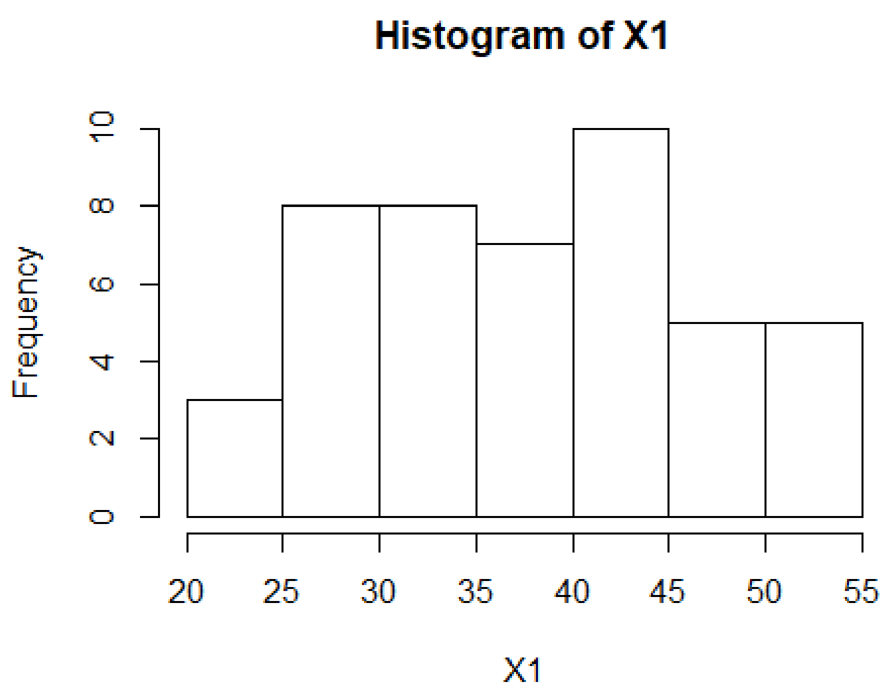
$$L(\beta_0, \beta_1, \beta_2, \beta_3, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{i1} - \beta_2 X_{i3} - \beta_3 X_{i2})^2 \right]$$

Here, maximum likelihood estimators will be the same as the least squares estimators because when the errors are normally distributed like in the normal regression model, the least squares estimators are optimal and fortunately, the maximum likelihood estimator also gives the optimal value when error is distributed normally.

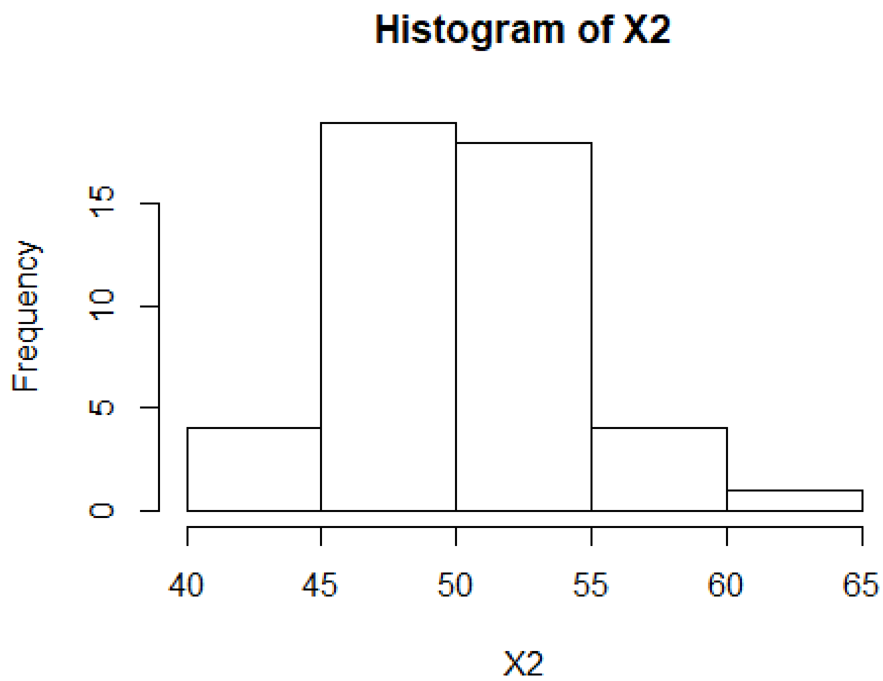
Part 5

(a)

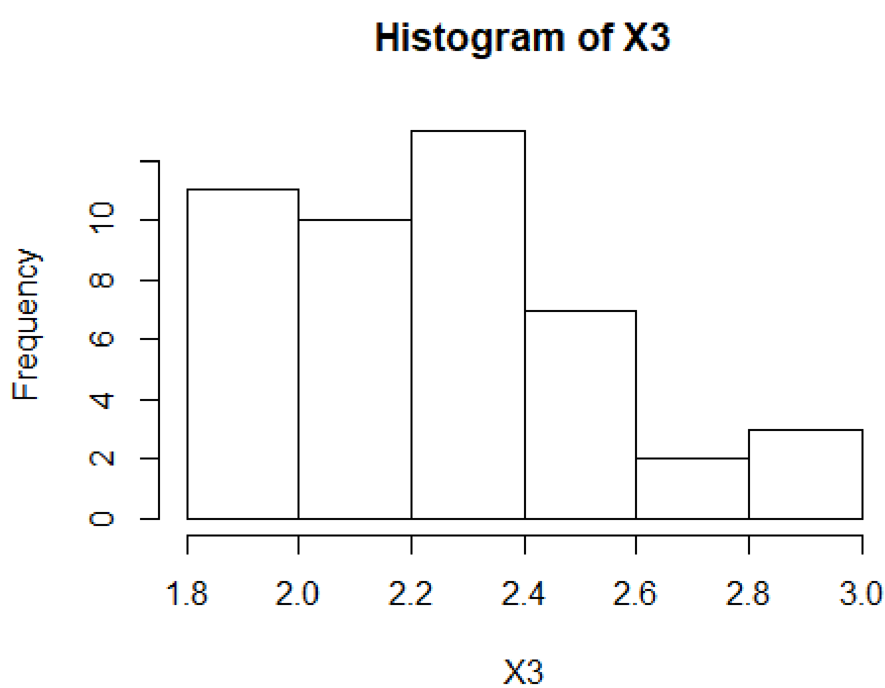
```
patsat = read.table("D:\\ASU Stuff\\SEM-1\\STP
530\\Patient_satisfaction.txt")
Y = patsat$V1
X1 = patsat$V2
X2 = patsat$V3
X3 = patsat$V4

#histogram of Patient's age
hist(X1)
```




```
#histogram of severity of illness
hist(X2)
```



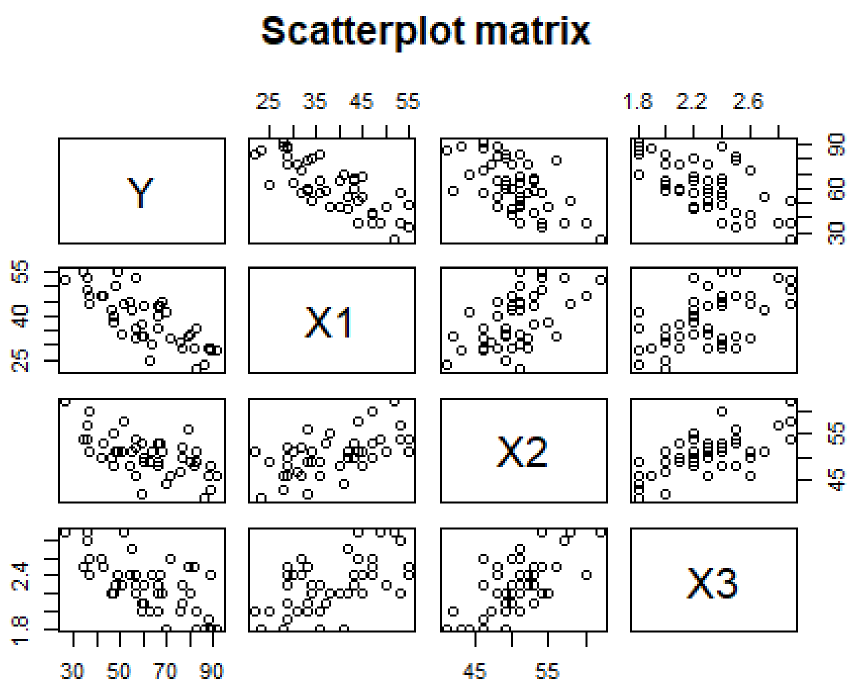

```r
#histogram of anxiety level
hist(X3)
```



From the histogram plots, no noteworthy features are observed for predictor variables X1 and X3, but the plot of variable X2 is approximately normal.

(b)

```
#Scatterplot matrix
pairs(~Y+X1+X2+X3, data=patsat, main="Scatterplot matrix")
```




```
#Correlation matrix
C1 = cbind(X1,X2,X3)
cor(C1)

##           X1        X2        X3
## X1 1.0000000 0.5679505 0.5696775
## X2 0.5679505 1.0000000 0.6705287
## X3 0.5696775 0.6705287 1.0000000
```

From the scatterplot matrix, we can see that there is a linear relation between response variable and all the predictor variables i.e. X1, X2 and X3. They are infact linearly decreasing. Also, the data is more clustered together in X1 as compared to X2 and X3. From the correlation matrix, X2 and X3 are strongly correlated than correlation of X1 and X2 and also correlation of X1 and X3.

(c)

```
model = lm(Y~X1+X2+X3)
model
```

```
##
## Call:
## lm(formula = Y ~ X1 + X2 + X3)
##
## Coefficients:
## (Intercept)           X1           X2           X3
##     158.491       -1.142       -0.442      -13.470
```

```
b2 = model$coefficients[3]
b2
```

```
##         X2
## -0.4420043
```

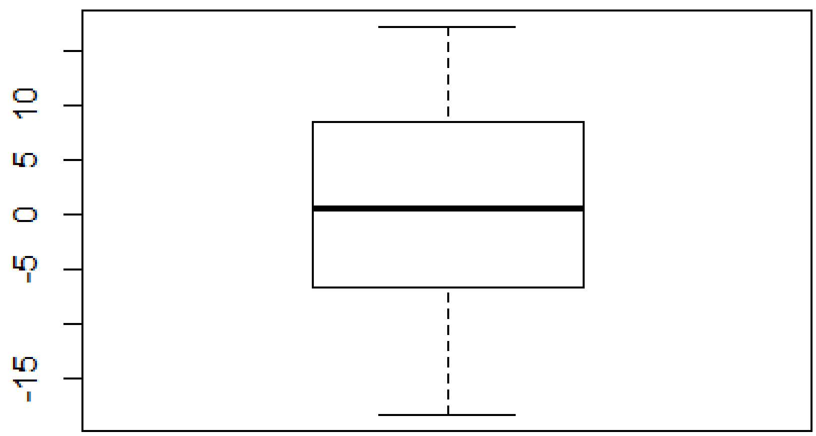
The estimated regression function is $Y = 158.491 - 1.142X_1 - 0.442X_2 - 13.470X_3$. the Value of b2 seems to be very less in this model.

(d)

```
e = resid(model)
e
```

```
##           1           2           3           4           5           6
##   0.1129334  -9.0796538   4.0237858   2.0093153   5.7263570  -3.6205678
##           7           8           9          10          11          12
## -12.8089820   0.4258777  -6.6596981   2.0030477  17.1600881  13.3526753
##          13          14          15          16          17          18
## -14.1654081 -15.1528562  12.5167654  -2.7946900  16.6095859   8.5409980
##          19          20          21          22          23          24
## -10.8725092   8.1680089   5.5810888   8.4393900   3.6796462  -3.8657107
##          25          26          27          28          29          30
##  -4.7338610  -4.1589620 -18.3524203   5.3949478  -9.6470593   3.3681039
##          31          32          33          34          35          36
## -16.3135553  11.5112774   0.6132423 -14.9762142   0.9248761  11.6161190
##          37          38          39          40          41          42
##  11.5071044  -5.3722872  -8.9868475  -5.7128575  11.0056590  -0.8932473
##          43          44          45          46
## -13.6956888  13.0578578  -5.5380448  10.0523698
```

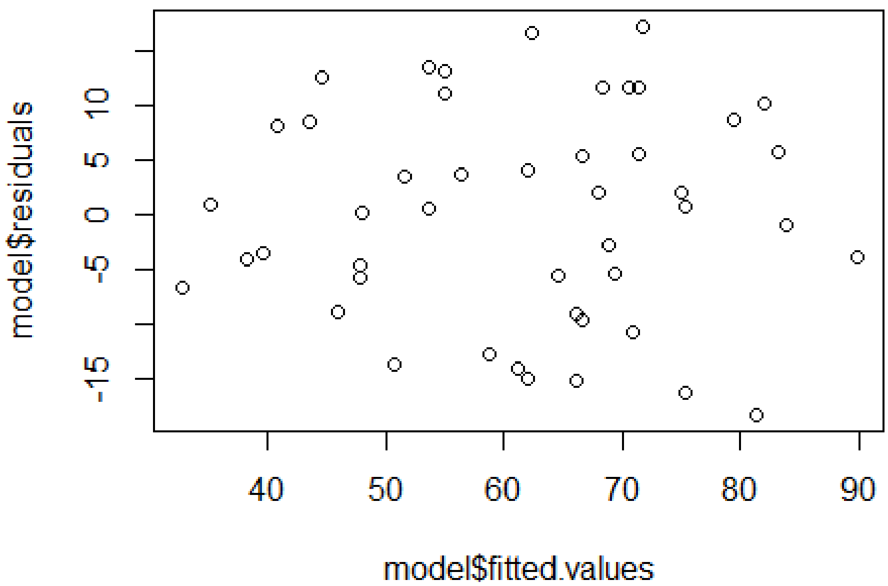
```
#Boxplot of residuals
boxplot(resid(model))
```



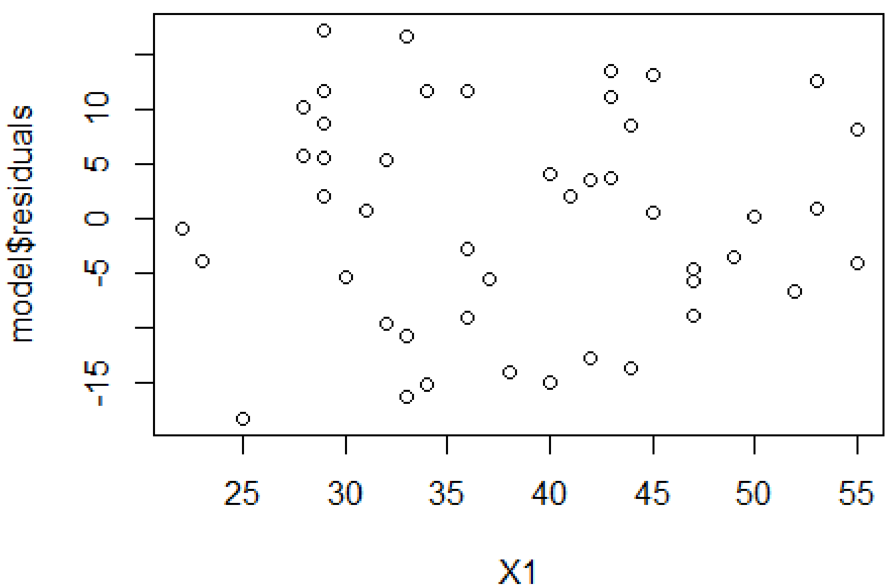
From the box-plot diagram, it is very evident that there are no outliers.

(e)

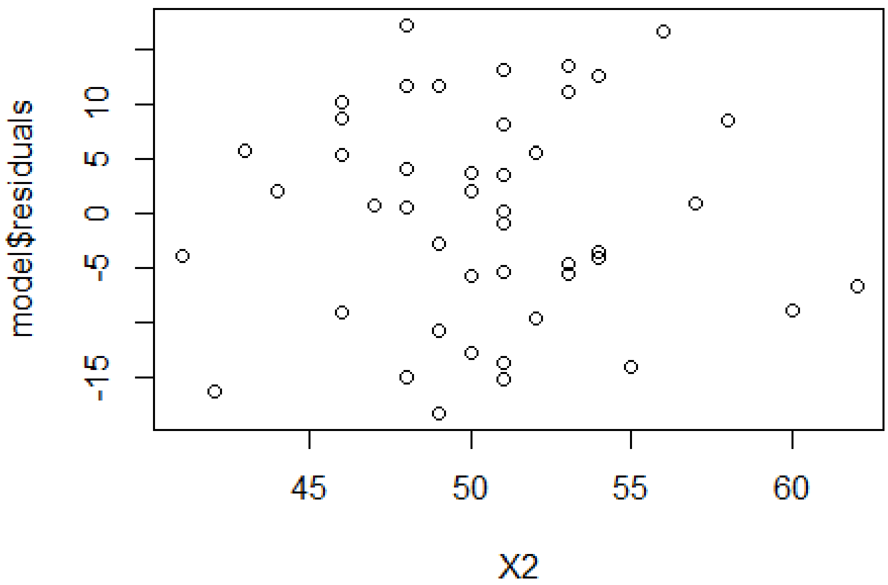
```
#Residuals vs fitted values
plot(model$residuals~model$fitted.values)
```



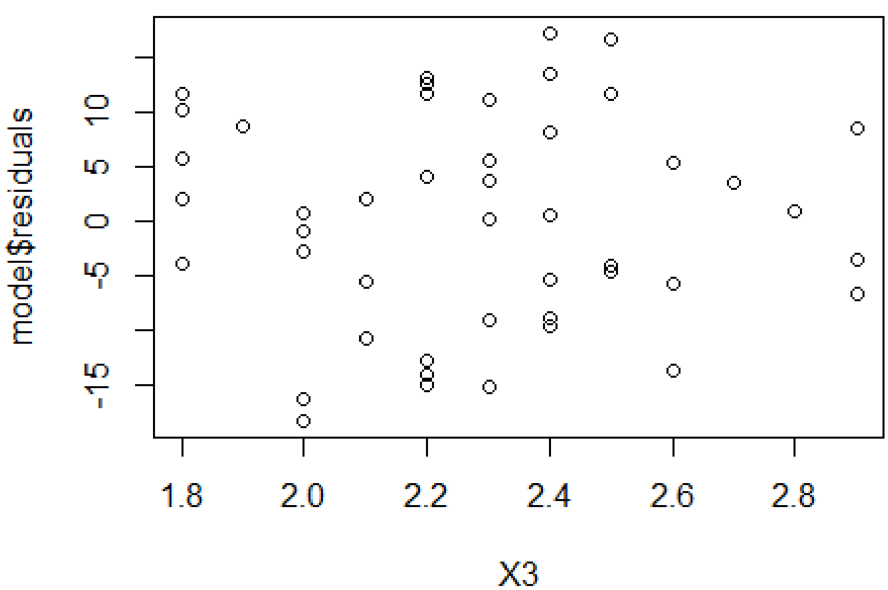

```
#Residuals vs predictor variables
plot(model$residuals~X1) #with predictor variable X1
```



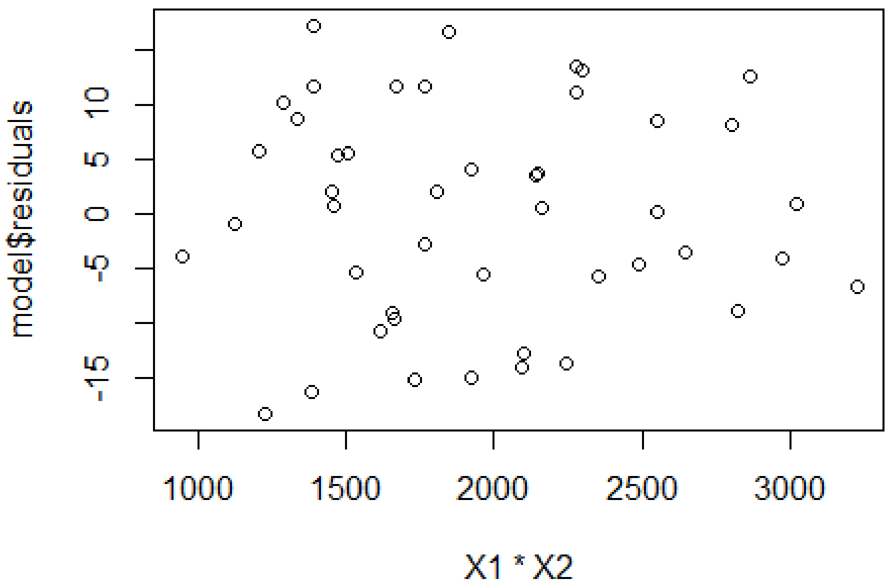
```r
plot(model$residuals~X2) #with predictor variable X2
```



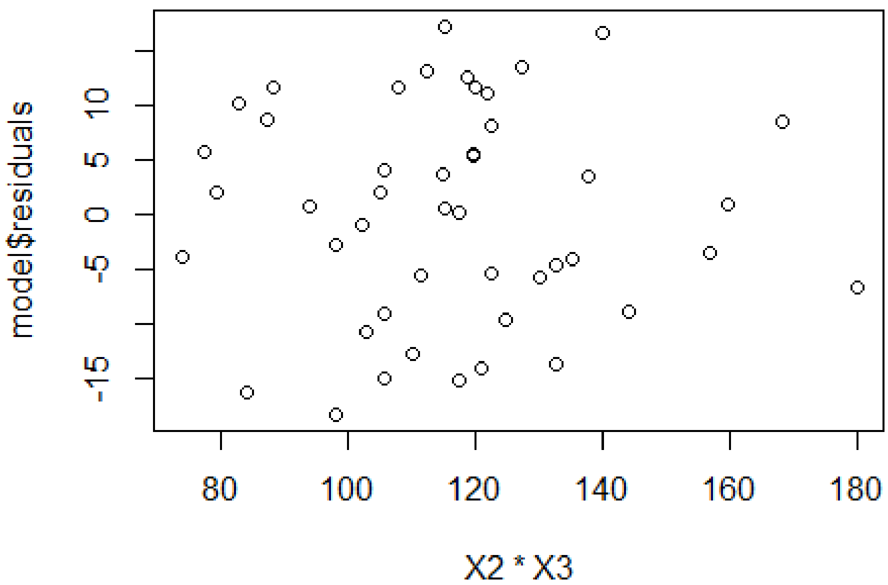

```r
plot(model$residuals~X3) #with predictor variable X3
```



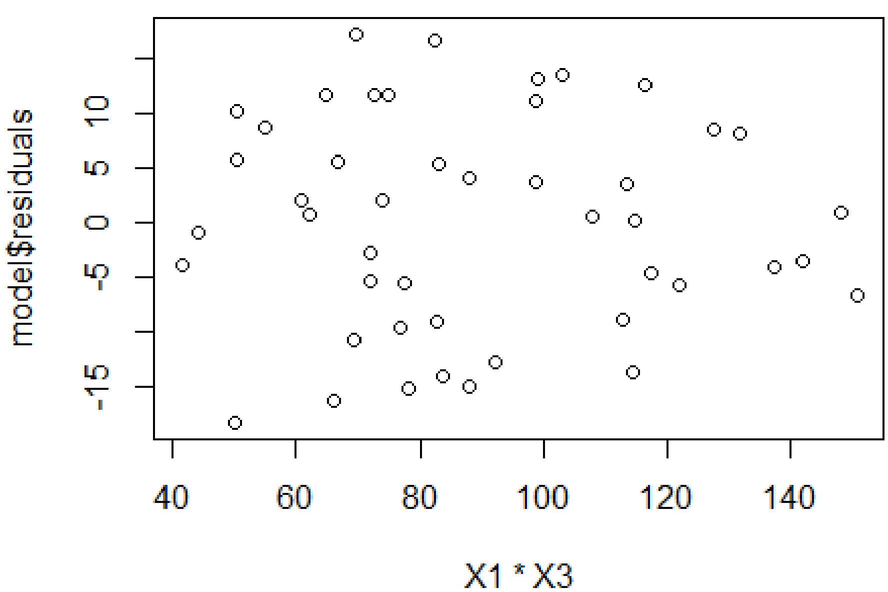
```r
#Residuals vs two factor interaction
plot(X1*X2, model$residuals)
```



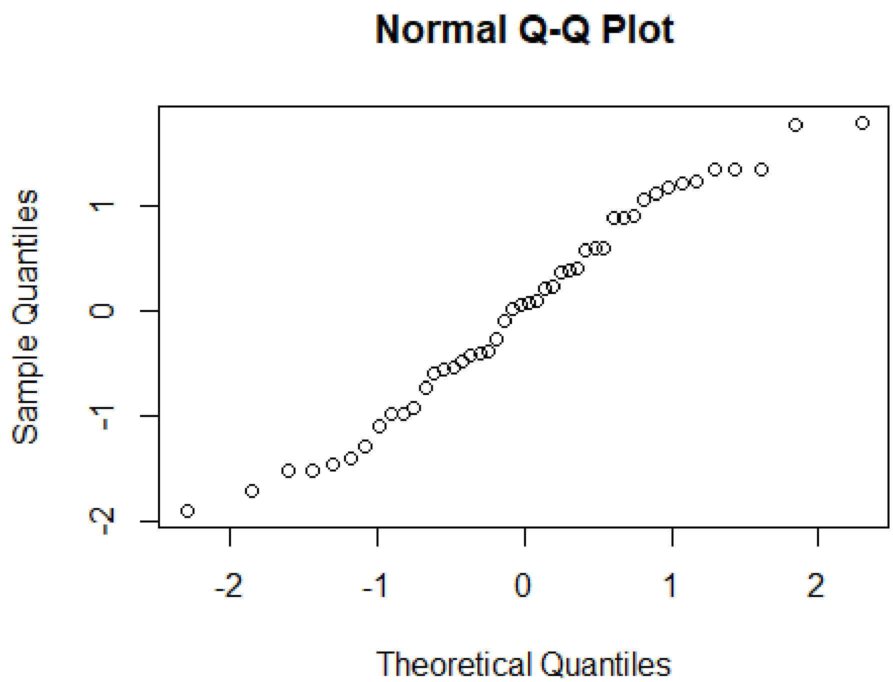

```r
plot(X2*X3, model$residuals)
```




```r
plot(X1*X3, model$residuals)
```



```
#Normal probability plot
plot1 = rstandard(model)
qqnorm(plot1)
```



From the plots of residuals vs fitted values, it clear that the error term of residuals and variance is constant which in turn gives that the linear model is appropriate. The same is the case with the plots of residuals vs each of the predictor variables and residuals vs two factor interaction.

Also, from the normal probability plot, it can be seen that as the theoretical quantiles increase, so do the sample quantiles.

(f)

When replicate observations are not available, an appropriate lack of fit test can be conducted if there are cases that have similar X_h vectors. In this case, we cannot do the formal test for lack of fit.

(g)

H_0:the variance is constant; H_1:The variance is not constant. SSR = 21,355.5, SSE = 4248.8, $X_B P^2 = (SSR/2)/(4248.8/46)^2 = 1.2516$. If $X_B^2 P \leq 11.34499$, conclude error variance constant, otherwise not constant. Hence, $X_B P^2 \leq X^2(0.99,3)$, therefore, we conclude H_0.

Part 6

```
patsat = read.table("D:\\ASU Stuff\\SEM-1\\STP
530\\Patient_satisfaction.txt")
Y = patsat$V1
X1 = patsat$V2
X2 = patsat$V3
X3 = patsat$V4

mod = lm(Y~X1+X2+X3)
mod

##
## Call:
## lm(formula = Y ~ X1 + X2 + X3)
##
## Coefficients:
## (Intercept)           X1           X2           X3
##     158.491       -1.142       -0.442      -13.470

anova(mod)

## Analysis of Variance Table
##
## Response: Y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## X1         1 8275.4  8275.4 81.8026 2.059e-11 ***
## X2         1  480.9   480.9  4.7539   0.03489 *
## X3         1  364.2   364.2  3.5997   0.06468 .
## Residuals 42 4248.8   101.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(mod)

##
## Call:
## lm(formula = Y ~ X1 + X2 + X3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.3524  -6.4230   0.5196   8.3715  17.1601
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 158.4913    18.1259   8.744 5.26e-11 ***
## X1           -1.1416     0.2148  -5.315 3.81e-06 ***
## X2           -0.4420     0.4920  -0.898   0.3741
## X3          -13.4702     7.0997  -1.897   0.0647 .
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.06 on 42 degrees of freedom
## Multiple R-squared:  0.6822, Adjusted R-squared:  0.6595
## F-statistic: 30.05 on 3 and 42 DF,  p-value: 1.542e-10

qf(0.9,3,42)

## [1] 2.219059

B = qt(1-0.1/6,42)

beta0 = mod$coefficients[1]
beta1 = mod$coefficients[2]
beta2 = mod$coefficients[3]
beta3 = mod$coefficients[4]


sbeta1 = 0.2148
sbeta2 = 0.4920
sbeta3 = 70.997

beta1a = beta1-B*sbeta1
beta0aa = beta1+B*sbeta1
beta0a = beta2-B*sbeta2
beta0aa = beta2+B*sbeta2
beta0a = beta3-B*sbeta3
beta0aa = beta3+B*sbeta3

yh = beta0 + beta1*35 + beta2*45 + beta3*2.2

dat = data.frame(X1=35,X2=45,X3=2.2)

predict(mod,dat,interval="confidence",level=0.90)

##        fit      lwr      upr
## 1 69.01029 64.52854 73.49204

predict(mod,dat,interval="prediction",level=0.90)

##        fit      lwr      upr
## 1 69.01029 51.50965 86.51092
```

Part - 6

a) $H_0: \beta_1 = \beta_2 = \beta_3 = 0$

H_a: Not all β_K are 0.

MSR = 3040.2, MSE = 101.2

$$F^* = \frac{MSR}{MSE} = \frac{3040.2}{101.2} = 30.05$$

If $F^* \leq F(1-\alpha; P-1, n-P)$, conclude H_0.

$F^* > F(1-\alpha; P-1, n-P)$, conclude H_a.

$F(1-0.1, 3, 42) = 2.0837$

$\Rightarrow F^* > F(1-\alpha; P-1; n-P)$

Thus, we conclude H_a.

P-value = 0.4878

b) $B = t\left(1-\frac{\alpha}{2g}, n-P\right)$

$= t\left(1-\frac{0.1}{6}, 42\right)$

$= 2.20036$

$s\{b_1\} = 0.2148$

$s\{b_2\} = 0.4420$

$s\{b_3\} = 7.0997$

$-1.6140 \leq \beta_1 \leq -0.6790$

$\cancel{-0.4394}$ $-1.5240 \leq \hat{\beta}_2 \leq 0.6400$

$-29.0858 \leq \hat{\beta}_3 \leq 2.145$

c) Coefficient of Multiple determination:

$$R^2 = \frac{SSR}{SSTO} = \frac{9120.50}{13.3694} = 0.8259$$

R^2 measures the proportionate reduction of total variation of Y w.r.t predictor variables. Thus, it means there is a 82.60% reduction of total variation of Y.

d) $\hat{Y}_h \pm t(1-\frac{\alpha}{2}, n-p)\, s\{\hat{Y}_h\}$

$\hat{Y}_h = 158.50 - 1.140 X_{1h} - 0.440 X_{2h} - 13.470 X_{3h}$

$X_{1h} = 35, \quad X_{2h} = 45, \quad X_{3h} = 2.2$

$\hat{Y}_h = 69.010$

e) Prediction interval:

$$\hat{Y}_h \pm t(1-\alpha/2, n-p)\, s\{pred\}$$

$51.509 \leq \hat{Y}_{h(new)} \leq 86.510$