Code

20122057_OLS_ASSUMPTION_PROOF

Hide

```
df= read.csv("D:/dwld/dat_df.csv",header = T)
View(df)
attach(df)
```

```
The following objects are masked from df (pos = 3):

    X, X0, X1, X10, X100, X101, X102, X103,
    X104, X105, X106, X107, X108, X109, X11,
    X110, X111, X112, X113, X114, X115,
    X116, X117, X118, X119, X12, X120, X121,
    X122, X123, X124, X125, X126, X127,
    X128, X129, X13, X130, X131, X132, X133,
    X134, X135, X136, X137, X138, X139, X14,
    X140, X141, X142, X143, X144, X145,
    X146, X147, X148, X149, X15, X150, X151,
    X152, X153, X154, X155, X156, X157,
    X158, X159, X16, X160, X161, X162, X163,
    X164, X165, X166, X167, X168, X169, X17,
    X170, X171, X172, X173, X174, X175,
    X176, X177, X178, X179, X18, X180, X181,
    X182, X183, X184, X185, X186, X187,
    X188, X189, X19, X190, X191, X192, X193,
    X194, X195, X196, X197, X198, X199, X2,
    X20, X200, X201, X202, X203, X204, X205,
    X206, X207, X208, X209, X21, X210, X211,
    X212, X213, X214, X215, X216, X217,
    X218, X219, X22, X220, X221, X222, X223,
    X224, X225, X226, X227, X228, X229, X23,
    X230, X231, X232, X233, X234, X235,
    X236, X237, X238, X239, X24, X240, X241,
    X242, X243, X244, X245, X246, X247,
    X248, X249, X25, X250, X251, X252, X253,
    X254, X255, X256, X257, X258, X259, X26,
    X260, X261, X262, X263, X264, X265,
    X266, X267, X268, X269, X27, X270, X271,
    X272, X273, X28, X29, X3, X30, X31, X32,
    X33, X34, X35, X36, X37, X38, X39, X4,
    X40, X41, X42, X43, X44, X45, X46, X47,
    X48, X49, X5, X50, X51, X52, X53, X54,
    X55, X56, X57, X58, X59, X6, X60, X61,
    X62, X63, X64, X65, X66, X67, X68, X69,
    X7, X70, X71, X72, X73, X74, X75, X76,
    X77, X78, X79, X8, X80, X81, X82, X83,
    X84, X85, X86, X87, X88, X89, X9, X90,
    X91, X92, X93, X94, X95, X96, X97, X98,
    X99
```

Hide

```
model1=lm(df$X273~.,df)
summary(model1)
```

```
Call:
lm(formula = df$X273 ~ ., data = df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.80100 -0.05491  0.00108  0.05963  2.15895

Coefficients: (43 not defined because of singularities)
              Estimate Std. Error t value
(Intercept)  3.709e-01  3.338e-03 111.141
X            6.873e-09  1.025e-08   0.670
X0           2.401e-02  1.876e-03  12.799
X1          -1.451e-01  6.249e-03 -23.217
X2          -1.395e-01  1.042e-02 -13.383
X3           4.156e-01  1.437e-02  28.924
X4           1.993e-01  3.426e-03  58.167
X5          -1.414e-03  9.076e-04  -1.558
X6           1.063e-02  3.647e-02   0.292
X7          -5.807e-01  7.149e-02  -8.122
X8          -4.105e-01  3.666e-02 -11.199
X9          -1.974e-01  1.507e-02 -13.103
X10         -2.499e-01  1.874e-02 -13.336
X11         -2.262e-01  1.877e-02 -12.051
X12         -2.488e-01  1.794e-02 -13.868
X13         -4.500e-01  2.799e-02 -16.080
X14                 NA         NA      NA
X15          6.823e-01  8.639e-02   7.898
X16                 NA         NA      NA
X17         -3.438e-03  5.994e-03  -0.574
X18          4.046e-02  7.770e-03   5.207
X19          1.373e+00  6.668e-02  20.586
X20         -4.907e-01  8.882e-02  -5.525
X21          1.693e-01  4.254e-03  39.805
X22         -6.673e-02  5.076e-03 -13.147
X23          1.017e+00  5.158e-02  19.718
X24          9.359e-03  2.407e-03   3.889
X25          1.491e-02  1.051e-03  14.179
X26          2.974e-02  1.499e-03  19.843
X27          1.175e-01  3.763e-02   3.122
X28         -6.088e-02  1.099e-02  -5.542
X29          5.917e-02  2.319e-02   2.552
X30          2.599e-01  3.016e-02   8.619
X31          5.709e-02  2.180e-02   2.619
X32          4.758e-02  2.606e-02   1.826
X33         -1.602e-01  3.493e-02  -4.585
X34         -3.601e-02  4.591e-02  -0.784
X35         -3.492e-01  1.574e-02 -22.188
X36         -4.930e-03  5.482e-03  -0.899
X37          5.763e-02  1.309e-02   4.404
X38          1.618e-02  1.564e-02   1.035
X39         -8.958e-02  2.045e-02  -4.379
X40         -1.307e+00  2.064e-01  -6.331
X41          8.993e-01  1.368e-01   6.572
X42          2.268e-02  5.587e-02   0.406
X43          7.177e-01  2.223e-02  32.280
X44          3.517e-01  1.339e-02  26.261
```

```
X45          7.997e-02  7.147e-03  11.189
X46          3.522e-02  9.590e-03   3.672
X47          3.150e-03  7.758e-03   0.406
X48         -1.585e-02  6.351e-03  -2.496
X49         -6.148e-02  6.352e-03  -9.679
X50         -6.610e-02  6.253e-03 -10.570
X51         -7.999e-02  6.692e-03 -11.953
X52         -1.109e-01  7.350e-03 -15.084
X53         -3.595e-01  8.998e-03 -39.959
X54         -5.278e-01  1.663e-02 -31.733
X55         -2.523e+00  8.223e-02 -30.680
X56          4.470e-02  4.681e-02   0.955
X57         -5.690e-02  2.464e-02  -2.309
X58         -6.520e-02  3.132e-02  -2.082
X59         -1.633e-01  2.949e-02  -5.537
X60         -4.750e-02  4.389e-02  -1.082
X61          2.120e-01  8.649e-02   2.451
X62          4.355e-01  8.859e-02   4.916
X63                 NA         NA      NA
X64          2.984e-02  2.773e-02   1.076
X65          8.677e-02  3.031e-02   2.863
X66          4.067e-02  3.759e-02   1.082
X67          1.824e-02  6.224e-02   0.293
X68         -1.349e-01  8.696e-02  -1.551
X69          1.722e+00  4.692e-02  36.695
X70         -1.820e-01  6.166e-02  -2.951
X71         -9.246e-03  2.755e-02  -0.336
X72         -1.552e-01  6.764e-02  -2.294
X73         -2.049e-02  1.695e-02  -1.209
X74         -6.379e-02  2.072e-02  -3.079
X75          4.060e-02  4.651e-02   0.873
X76         -6.397e-02  3.294e-02  -1.942
X77          8.849e-03  6.901e-02   0.128
X78         -2.910e-02  2.750e-03 -10.581
X79          7.870e-05  4.071e-03   0.019
X80          9.433e-03  6.961e-03   1.355
X81         -1.360e-02  5.363e-03  -2.536
X82         -7.228e-04  3.613e-03  -0.200
X83         -6.923e-03  3.394e-03  -2.039
X84          4.846e-03  3.241e-03   1.495
X85         -2.530e-02  1.542e-03 -16.410
X86          1.439e-02  2.834e-03   5.080
X87         -2.806e-02  2.258e-03 -12.430
X88          2.016e-02  3.142e-03   6.416
X89         -2.155e-03  1.135e-02  -0.190
X90                 NA         NA      NA
X91                 NA         NA      NA
X92                 NA         NA      NA
X93          1.359e-01  3.821e-02   3.557
X94         -2.066e-01  1.036e-01  -1.994
X95         -5.697e-02  4.623e-02  -1.232
X96         -1.085e-01  3.295e-02  -3.295
X97         -1.182e-01  6.361e-02  -1.858
X98         -1.139e-01  5.569e-02  -2.045
X99                 NA         NA      NA
X100                NA         NA      NA
X101        -1.762e-01  2.332e-02  -7.554
X102        -1.218e-03  1.131e-02  -0.108
```

X103	4.523e-02	2.139e-02	2.115
X104	4.152e-02	3.003e-02	1.383
X105	NA	NA	NA
X106	5.443e-01	1.238e-01	4.396
X107	-1.401e-01	3.583e-02	-3.911
X108	1.183e-01	6.117e-02	1.934
X109	4.483e-03	6.296e-02	0.071
X110	-1.534e-01	9.573e-02	-1.602
X111	2.171e-01	7.258e-02	2.991
X112	-1.208e-01	9.894e-02	-1.221
X113	1.238e-01	7.509e-02	1.648
X114	1.623e-01	8.518e-02	1.905
X115	-1.386e-01	1.101e-01	-1.259
X116	-5.579e-01	1.726e-01	-3.232
X117	-1.342e-01	4.434e-02	-3.026
X118	-7.807e-02	5.347e-02	-1.460
X119	-3.258e-03	5.220e-02	-0.062
X120	-3.951e-02	7.436e-02	-0.531
X121	1.124e-01	6.268e-02	1.793
X122	NA	NA	NA
X123	NA	NA	NA
X124	NA	NA	NA
X125	NA	NA	NA
X126	NA	NA	NA
X127	2.006e-01	8.280e-02	2.423
X128	8.645e-02	7.684e-02	1.125
X129	-1.121e-01	8.328e-02	-1.346
X130	-7.593e-02	1.296e-01	-0.586
X131	2.057e-01	9.912e-02	2.076
X132	NA	NA	NA
X133	NA	NA	NA
X134	NA	NA	NA
X135	NA	NA	NA
X136	NA	NA	NA
X137	NA	NA	NA
X138	NA	NA	NA
X139	NA	NA	NA
X140	NA	NA	NA
X141	NA	NA	NA
X142	-1.545e-01	2.124e-03	-72.766
X143	8.743e-02	2.817e-02	3.104
X144	-1.305e-02	3.264e-02	-0.400
X145	-2.742e-02	6.043e-02	-0.454
X146	4.143e-02	8.971e-02	0.462
X147	2.294e-02	5.797e-02	0.396
X148	-2.736e-02	6.206e-03	-4.409
X149	6.276e-02	2.581e-02	2.432
X150	-5.537e-02	4.523e-02	-1.224
X151	-8.345e-02	6.696e-02	-1.246
X152	-7.170e-03	9.410e-02	-0.076
X153	-5.093e-02	6.348e-02	-0.802
X154	-5.437e-02	2.760e-03	-19.699
X155	-1.606e-03	2.286e-02	-0.070
X156	-1.087e-01	6.183e-02	-1.757
X157	-9.889e-02	6.006e-02	-1.646
X158	1.977e-01	1.007e-01	1.964
X159	-7.588e-02	6.559e-02	-1.157
X160	2.690e-01	2.986e-02	9.009

```
X161        -3.356e-01  3.519e-02  -9.536
X162         6.507e-02  3.851e-02   1.690
X163        -9.206e-02  3.540e-02  -2.601
X164        -4.926e-02  3.278e-02  -1.503
X165        -3.200e-02  3.389e-02  -0.944
X166        -7.527e-02  2.815e-02  -2.674
X167         1.452e-01  3.066e-02   4.738
X168        -8.657e-02  2.857e-02  -3.030
X169         8.685e-03  3.755e-02   0.231
X170        -1.739e-02  3.620e-02  -0.480
X171        -1.584e-02  3.543e-02  -0.447
X172        -6.414e-03  2.620e-03  -2.448
X173         1.568e-01  2.642e-02   5.935
X174        -4.560e-02  1.132e-02  -4.027
X175        -8.223e-02  1.685e-02  -4.881
X176        -6.434e-02  2.483e-02  -2.592
X177                NA         NA      NA
X178        -2.920e-02  4.526e-02  -0.645
X179         8.106e-02  5.221e-02   1.552
X180        -1.368e-01  6.562e-02  -2.085
X181        -5.037e-03  5.733e-02  -0.088
X182         2.466e-02  5.089e-02   0.485
X183        -5.676e-02  3.793e-02  -1.497
X184         2.331e-01  8.940e-02   2.608
X185        -1.846e-01  5.406e-02  -3.414
X186        -7.938e-02  4.773e-02  -1.663
X187        -2.284e-01  7.477e-02  -3.055
X188         1.022e-01  8.256e-02   1.238
X189         7.999e-02  6.262e-02   1.277
X190         8.789e-02  4.883e-02   1.800
X191        -9.427e-02  4.119e-02  -2.288
X192         3.950e-02  5.386e-02   0.733
X193        -2.111e-02  3.875e-02  -0.545
X194                NA         NA      NA
X195                NA         NA      NA
X196                NA         NA      NA
X197                NA         NA      NA
            Pr(>|t|)
(Intercept)  < 2e-16 ***
X           0.502717
X0           < 2e-16 ***
X1           < 2e-16 ***
X2           < 2e-16 ***
X3           < 2e-16 ***
X4           < 2e-16 ***
X5          0.119197
X6          0.770585
X7          4.64e-16 ***
X8           < 2e-16 ***
X9           < 2e-16 ***
X10          < 2e-16 ***
X11          < 2e-16 ***
X12          < 2e-16 ***
X13          < 2e-16 ***
X14               NA
X15         2.87e-15 ***
X16               NA
X17         0.566270
```

```
X18         1.92e-07 ***
X19          < 2e-16 ***
X20         3.31e-08 ***
X21          < 2e-16 ***
X22          < 2e-16 ***
X23          < 2e-16 ***
X24         0.000101 ***
X25          < 2e-16 ***
X26          < 2e-16 ***
X27         0.001799 **
X28         3.01e-08 ***
X29         0.010725 *
X30          < 2e-16 ***
X31         0.008816 **
X32         0.067918 .
X33         4.54e-06 ***
X34         0.432841
X35          < 2e-16 ***
X36         0.368514
X37         1.06e-05 ***
X38         0.300905
X39         1.19e-05 ***
X40         2.45e-10 ***
X41         5.00e-11 ***
X42         0.684788
X43          < 2e-16 ***
X44          < 2e-16 ***
X45          < 2e-16 ***
X46         0.000240 ***
X47         0.684777
X48         0.012574 *
X49          < 2e-16 ***
X50          < 2e-16 ***
X51          < 2e-16 ***
X52          < 2e-16 ***
X53          < 2e-16 ***
X54          < 2e-16 ***
X55          < 2e-16 ***
X56         0.339649
X57         0.020940 *
X58         0.037339 *
X59         3.08e-08 ***
X60         0.279102
X61         0.014247 *
X62         8.86e-07 ***
X63               NA
X64         0.281916
X65         0.004198 **
X66         0.279280
X67         0.769467
X68         0.120908
X69          < 2e-16 ***
X70         0.003164 **
X71         0.737149
X72         0.021788 *
X73         0.226697
X74         0.002076 **
X75         0.382728
```

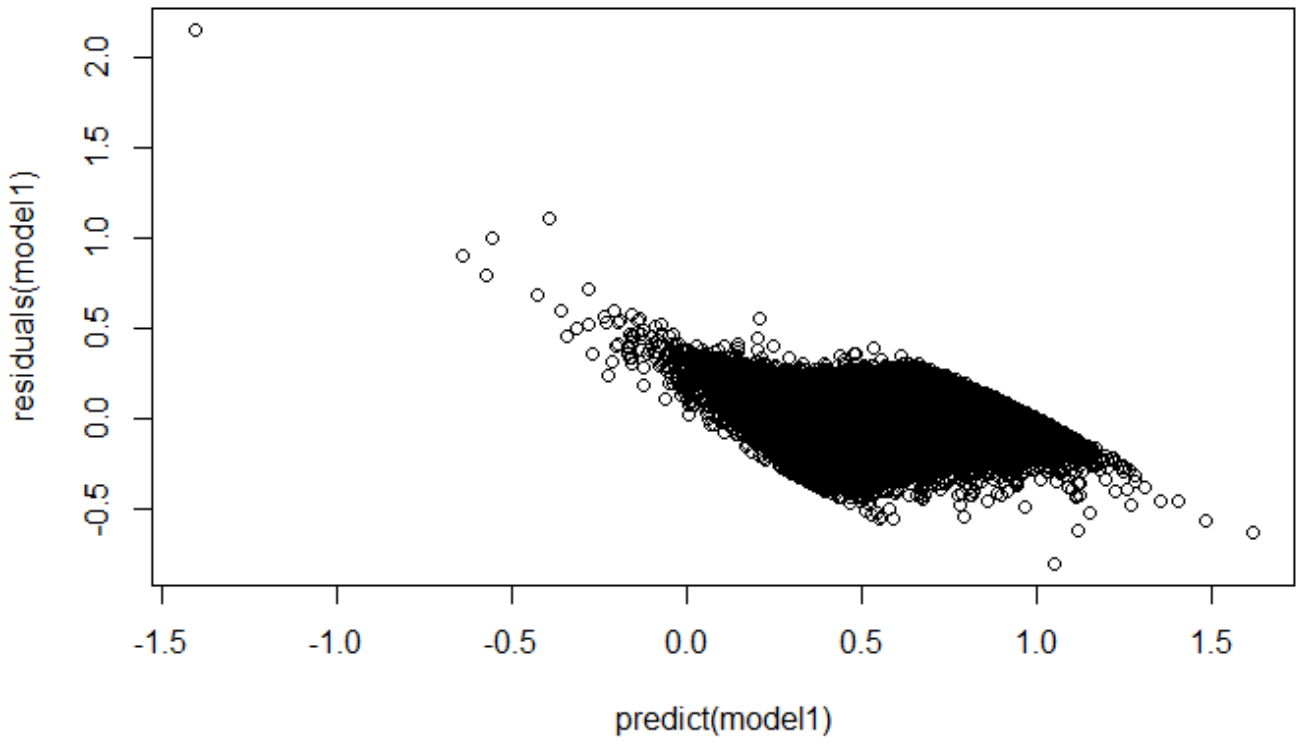
```
X76         0.052164 .
X77         0.897966
X78          < 2e-16 ***
X79         0.984577
X80         0.175386
X81         0.011229 *
X82         0.841431
X83         0.041407 *
X84         0.134881
X85          < 2e-16 ***
X86         3.79e-07 ***
X87          < 2e-16 ***
X88         1.41e-10 ***
X89         0.849349
X90               NA
X91               NA
X92               NA
X93         0.000375 ***
X94         0.046115 *
X95         0.217857
X96         0.000986 ***
X97         0.063212 .
X98         0.040879 *
X99               NA
X100              NA
X101        4.25e-14 ***
X102        0.914236
X103        0.034451 *
X104        0.166821
X105              NA
X106        1.11e-05 ***
X107        9.21e-05 ***
X108        0.053063 .
X109        0.943231
X110        0.109183
X111        0.002783 **
X112        0.222215
X113        0.099357 .
X114        0.056755 .
X115        0.208104
X116        0.001228 **
X117        0.002479 **
X118        0.144295
X119        0.950234
X120        0.595198
X121        0.072945 .
X122              NA
X123              NA
X124              NA
X125              NA
X126              NA
X127        0.015385 *
X128        0.260540
X129        0.178301
X130        0.557893
X131        0.037934 *
X132              NA
X133              NA
```

```
X134              NA
X135              NA
X136              NA
X137              NA
X138              NA
X139              NA
X140              NA
X141              NA
X142         < 2e-16 ***
X143        0.001911 **
X144        0.689398
X145        0.649952
X146        0.644254
X147        0.692325
X148        1.04e-05 ***
X149        0.015014 *
X150        0.220950
X151        0.212660
X152        0.939264
X153        0.422380
X154         < 2e-16 ***
X155        0.944004
X156        0.078843 .
X157        0.099672 .
X158        0.049508 *
X159        0.247328
X160         < 2e-16 ***
X161         < 2e-16 ***
X162        0.091118 .
X163        0.009304 **
X164        0.132900
X165        0.345029
X166        0.007494 **
X167        2.17e-06 ***
X168        0.002447 **
X169        0.817104
X170        0.631017
X171        0.654759
X172        0.014379 *
X173        2.96e-09 ***
X174        5.66e-05 ***
X175        1.06e-06 ***
X176        0.009557 **
X177              NA
X178        0.518789
X179        0.120579
X180        0.037099 *
X181        0.929987
X182        0.627966
X183        0.134510
X184        0.009109 **
X185        0.000641 ***
X186        0.096305 .
X187        0.002250 **
X188        0.215791
X189        0.201500
X190        0.071859 .
X191        0.022112 *
```

```
X192        0.463273
X193        0.585840
X194              NA
X195              NA
X196              NA
X197              NA
 [ reached getOption("max.print") -- omitted 75 rows ]
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.09352 on 99768 degrees of freedom
Multiple R-squared:  0.8193,    Adjusted R-squared:  0.8189
F-statistic:  1959 on 231 and 99768 DF,  p-value: < 2.2e-16
```

Hide

```
plot(predict(model1),residuals(model1))
```



Hide

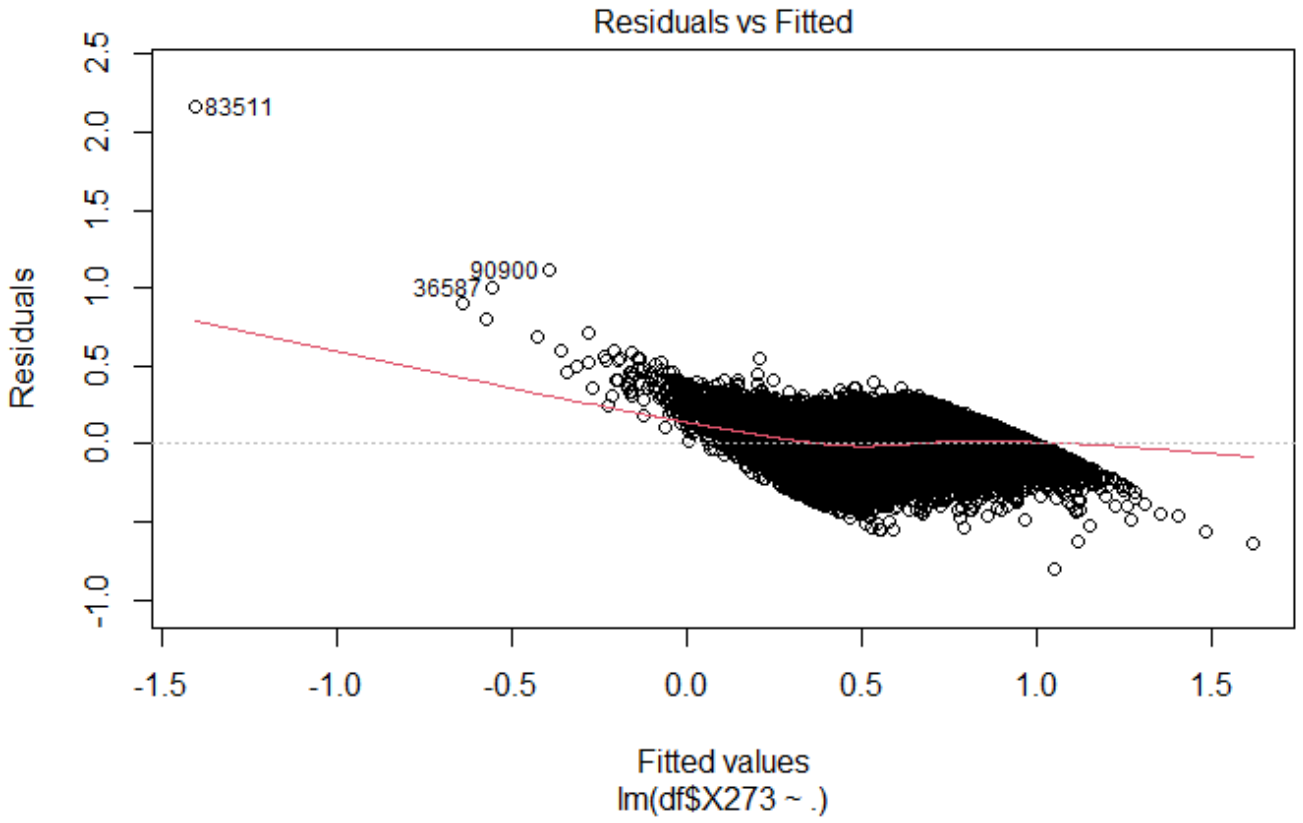
```
mean(residuals(model1))
```

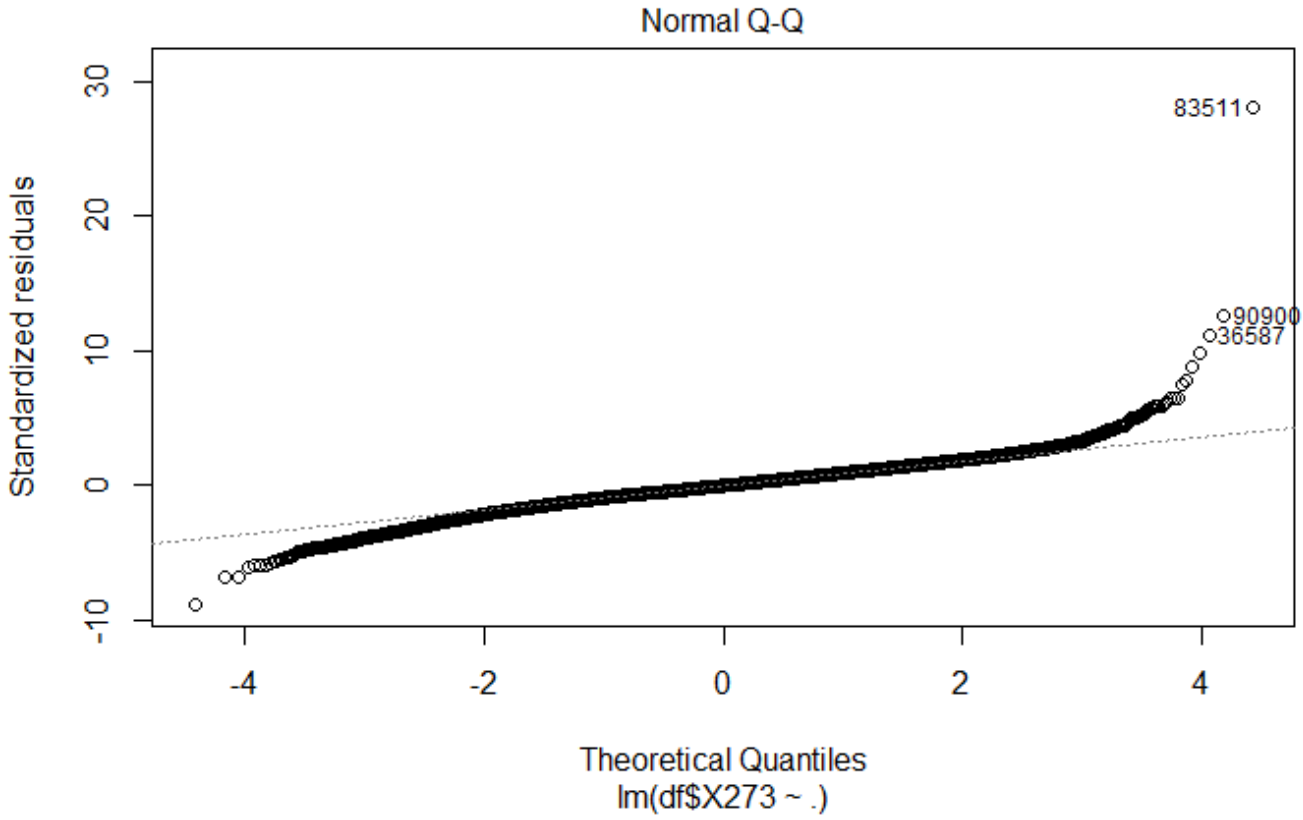
```
[1] -8.393522e-18
```

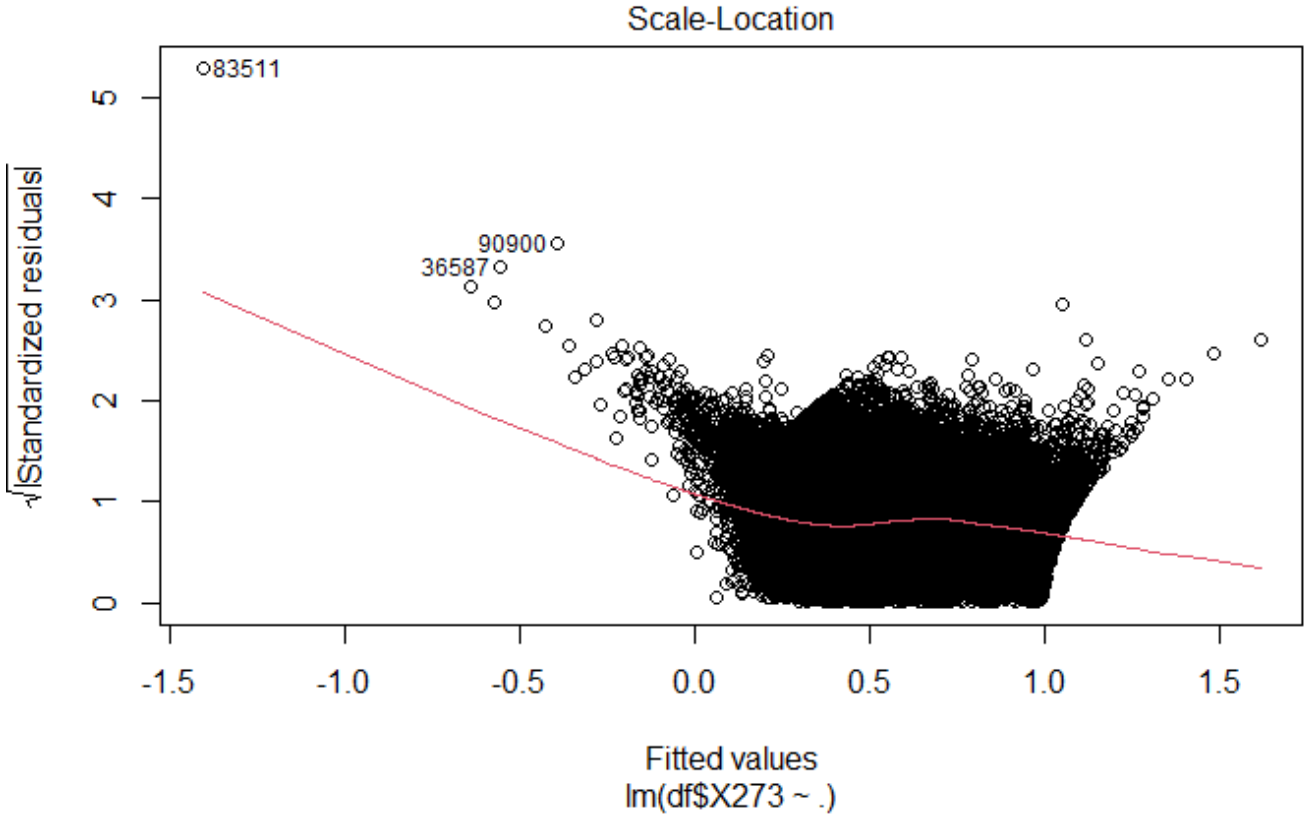
Diagonostic Plot for testing assumptions

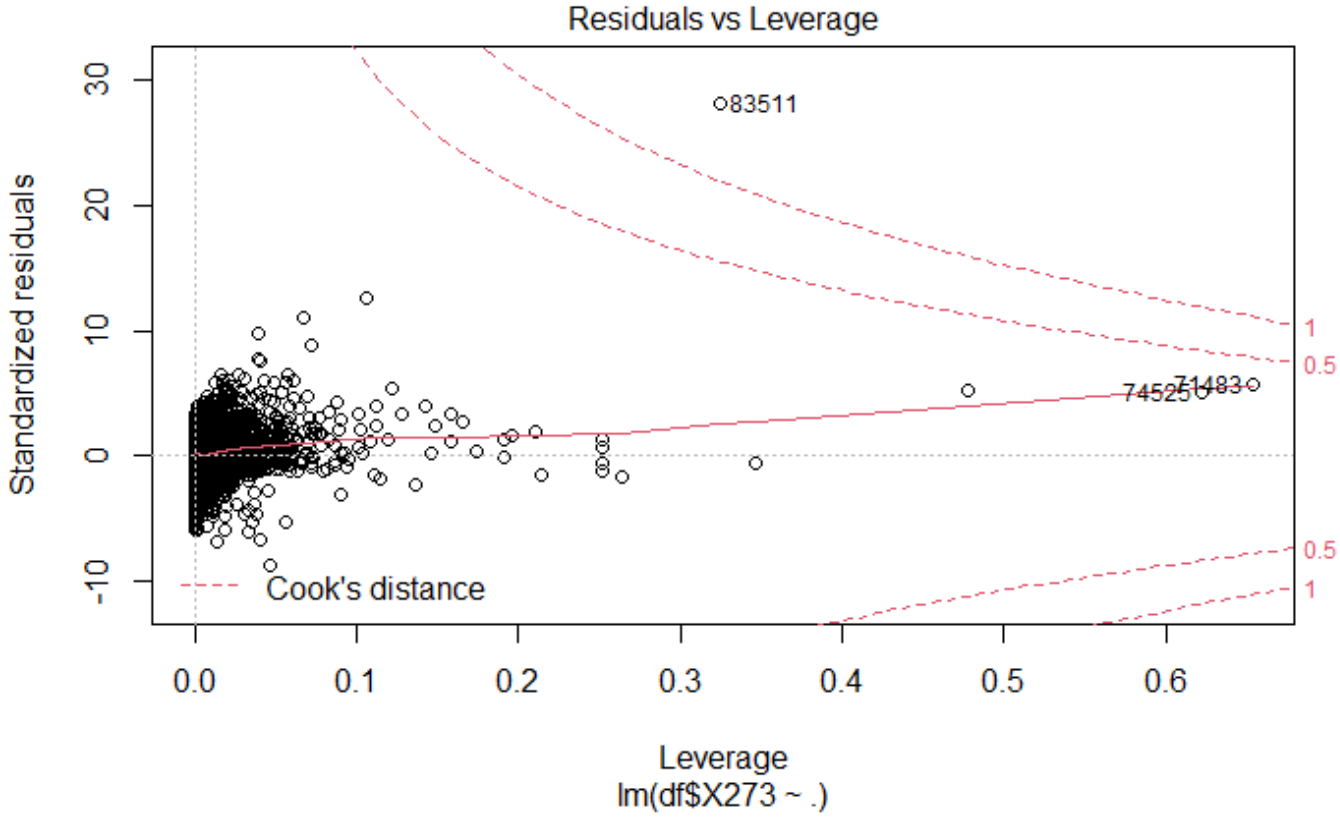
Hide

```
f=predict(model1)

plot(model1)
```


Residuals vs Fitted
Residuals
Fitted values
lm(df$X273 ~ .)
83511
90900
36587
2.5
2.0
1.5
1.0
0.5
0.0
-1.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5


Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(df$X273 ~ .)
83511
90900
36587
30
20
10
0
-10
-4
-2
0
2
4





A1. The linear regression model is “linear in parameters.” (parameters are alpha and beta values)

```
 From fig1(below) the red line near to the dense cluster is flat indicating linearity in para
meters
```

A2. There is a random sampling of observations.

A3. The residual mean should be zero

```
In R file i showed mean of residuals is zero that is assumption A3 is also proved.
```

A5. Spherical errors: There is homoscedasticity and no autocorrelation

```
The below figure  proves my A5 assumptions since error varies with constant variance with res
pect to variablesnjkkhyug
```

A6: Optional Assumption: Error terms should be normally distributed

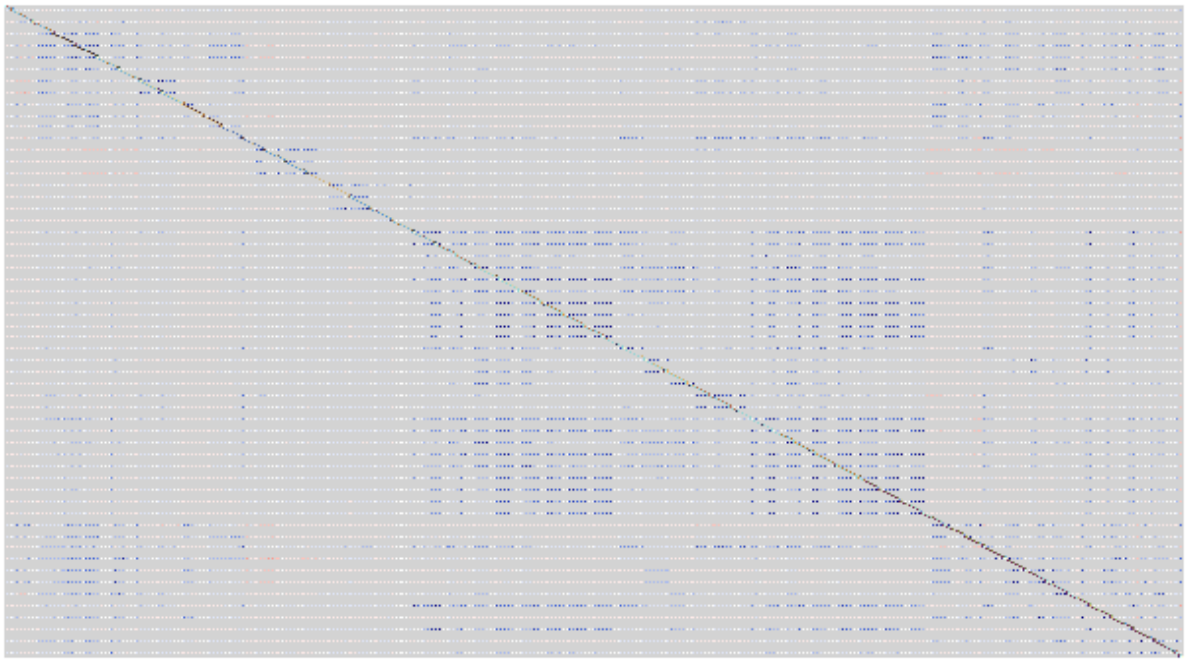
```
The above plot is a Q-Q- plot or Quantile - Quantile chart, X axis as theorectical X and Y ax
is as standardized residuals.

if standardized error is linear w.r.t my Theorectical value then we can say my residuals are
normally distributed.

Therefore from the above figure i can prove my A6 assumption( linearity is shown by dotted bl
ack line).
```

A4. There is no multi-collinearity (or perfect collinearity).

Hide

```
corrgram::corrgram(cor(df))
```

```
the standard deviation is zero
```



all the blue dots except the last column is showing collinearity between independent variables, thereisnt a serious problem or dark red box formation in correlation plot. we can go ahead with our assumption to be true.

ALL OUR OLS ASSUMPTIONS ARE PROVED