Regression Analysis of Mortality Rate

By Sailee Rumao,
Under Dr Linlin Chen,
ROCHESTER INSTITUTE OF TECHNOLOGY.
Course-741| MS,APP-STATS|
03 DECEMBER,2017

Table of Contents:

Abstract:

The focus of this study is to build a model to predict the death rate based on the variables under consideration using Multiple Linear Regression Analysis.

In this study, different methods of Regression Analysis for variable selection are used to build various Regression models. These models are compared to choose the best model based upon the statistical measures R-Square, PRESS, MSE, and AIC. The best model chosen by this process is further validated by evaluating the residual plots and examining if the predicted values lie in the prediction interval.

The analysis is concluded with its findings and Recommendations.

The statistical software used for this analysis is R-software.

Understanding:

In literal terms, Mortality means death. Death of one person can be called Mortality. But, by definition Mortality rate is death of a number of people in a region over a period of time.

Study of Mortality is important to understand the health and well-being of population,causes of death , social security, population forecasting and evaluation of medical measures and is widely allocated for Public Health expenditures & health care financing, planning and development of the nation.

The objectives of this analysis are as follows:

Objectives:

1. To fit the model and check for significance of the model
2. Test significance of overall model and individual variables
3. Build confidence intervals
4. Build prediction interval.
5. Check for multicollinearity
6. Check for model adequacy by Residual Analysis and other tests.
7. Conduct Influence points analysis.
8. Variable selection and model building
9. Validation for regression.

The analysis helps us understand the impact of unit change in any of the predictor variables on the Response variable. This can be used to take effective measures on the variables under study.

Data Understanding:

Data source:

The data for this study is taken from the department of scientific computing of the Florida state University, United States. However, the original source of data is as follows:

1.Richard Gunst, Robert Mason,
Regression Analysis and Its Applications: a data-oriented approach

2.Gary McDonald, Richard Schwing
Instabilities of regression estimates relating air pollution to mortality.

3. Helmut Spaeth,
Mathematical Algorithms for Linear Regression,
Academic Press.

The data is consistent over all variables and there are no MISSING values.

The data consists of 60 variables from 14 variables which are as follows:

- A1: the average annual precipitation;
- A2: the average January temperature;
- A3: the average July temperature;
- A4: the size of the population older than 65;
- A5: the number of members per household;
- A6: the number of years of schooling for persons over 22;
- A7: the number of households with fully equipped kitchens;
- A8: the population per square mile;
- A9: the number of office workers;
- A10: the number of families with an income less than $3000;
- A11: the hydrocarbon pollution index;
- A12: the nitric oxide pollution index;
- A13: the sulfur dioxide pollution index;
- A14: the degree of atmospheric moisture.
- B: the death rate.

- The model is as follows:

B = A1*X1 + A2*X2 + A3*X3 + A4*X4 + A5*X5 + A6*X6 + A7*X7 + A8*X8 + A9*X9 + A10*X10 + A11*X11 + A12*X12 + A13*X13 + A14*X14 +A15*X15

Data type:

The data is a data frame with its variables of integer or numeric type.

```
> str(data1)
'data.frame':	60 obs. of  15 variables:
 $ A1 : int  36 35 44 47 43 53 43 45 36 36 ...
 $ A2 : int  27 23 29 45 35 45 30 30 24 27 ...
 $ A3 : int  71 72 74 79 77 80 74 73 70 72 ...
 $ A4 : num  8.1 11.1 10.4 6.5 7.6 7.7 10.9 9.3 9 9.5 ...
 $ A5 : num  3.34 3.14 3.21 3.41 3.44 3.45 3.23 3.29 3.31 3.36 ...
 $ A6 : num  11.4 11 9.8 11.1 9.6 10.2 12.1 10.6 10.5 10.7 ...
 $ A7 : num  81.5 78.8 81.6 77.5 84.6 66.8 83.9 86 83.2 79.3 ...
 $ A8 : int  3243 4281 4260 3125 6441 3325 4679 2140 6582 4213 ...
 $ A9 : num  42.6 50.7 39.4 50.2 43.7 43.1 49.2 40.4 42.5 41 ...
 $ A10: num  11.7 14.4 12.4 20.6 14.3 25.5 11.3 10.5 12.6 13.2 ...
 $ A11: int  21 8 6 18 43 30 21 6 18 12 ...
 $ A12: int  15 10 6 8 38 32 32 4 12 7 ...
 $ A13: int  59 39 33 24 206 72 62 4 37 20 ...
 $ A14: int  59 57 54 56 55 54 56 56 61 59 ...
 $ B  : num  922 998 962 982 1071 ...
```

The descriptive statistics of the data are:

```
> summary(data)
       I               A1              A2              A3       
 Min.   : 1.00   Min.   :10.00   Min.   :12.00   Min.   :63.00  
 1st Qu.:15.75   1st Qu.:32.75   1st Qu.:27.00   1st Qu.:72.00  
 Median :30.50   Median :38.00   Median :31.50   Median :74.00  
 Mean   :30.50   Mean   :37.37   Mean   :34.82   Mean   :74.60  
 3rd Qu.:45.25   3rd Qu.:43.25   3rd Qu.:40.00   3rd Qu.:77.25  
 Max.   :60.00   Max.   :60.00   Max.   :83.00   Max.   :85.00  
       A4               A5              A6              A7       
 Min.   : 5.600   Min.   :2.920   Min.   : 9.00   Min.   :66.80  
 1st Qu.: 7.675   1st Qu.:3.210   1st Qu.:10.40   1st Qu.:78.38  
 Median : 9.000   Median :3.265   Median :11.05   Median :81.15  
 Mean   : 8.798   Mean   :3.263   Mean   :10.97   Mean   :80.91  
 3rd Qu.: 9.700   3rd Qu.:3.360   3rd Qu.:11.50   3rd Qu.:83.60  
 Max.   :11.800   Max.   :3.530   Max.   :12.30   Max.   :90.70  
       A8             A9             A10             A11        
 Min.   :1441   Min.   :33.80   Min.   : 9.40   Min.   :  1.00  
 1st Qu.:3104   1st Qu.:43.25   1st Qu.:12.00   1st Qu.:  7.00  
 Median :3567   Median :45.50   Median :13.20   Median : 14.50  
 Mean   :3876   Mean   :46.07   Mean   :14.37   Mean   : 37.85  
 3rd Qu.:4520   3rd Qu.:49.52   3rd Qu.:15.15   3rd Qu.: 30.25  
 Max.   :9699   Max.   :59.70   Max.   :26.40   Max.   :648.00  
      A12              A13              A14              B         
 Min.   :  1.00   Min.   :  1.00   Min.   :38.00   Min.   : 790.7  
 1st Qu.:  4.00   1st Qu.: 11.00   1st Qu.:54.75   1st Qu.: 898.4  
 Median :  9.00   Median : 30.00   Median :57.00   Median : 943.7  
 Mean   : 22.52   Mean   : 53.77   Mean   :57.53   Mean   : 940.4  
 3rd Qu.: 23.75   3rd Qu.: 69.00   3rd Qu.:60.00   3rd Qu.: 983.2  
 Max.   :319.00   Max.   :278.00   Max.   :73.00   Max.   :1113.2  
```

For Regression Analysis , the variables under study should have a relationship between them.

Checking this primary requirement by plotting scatterplot of the data.

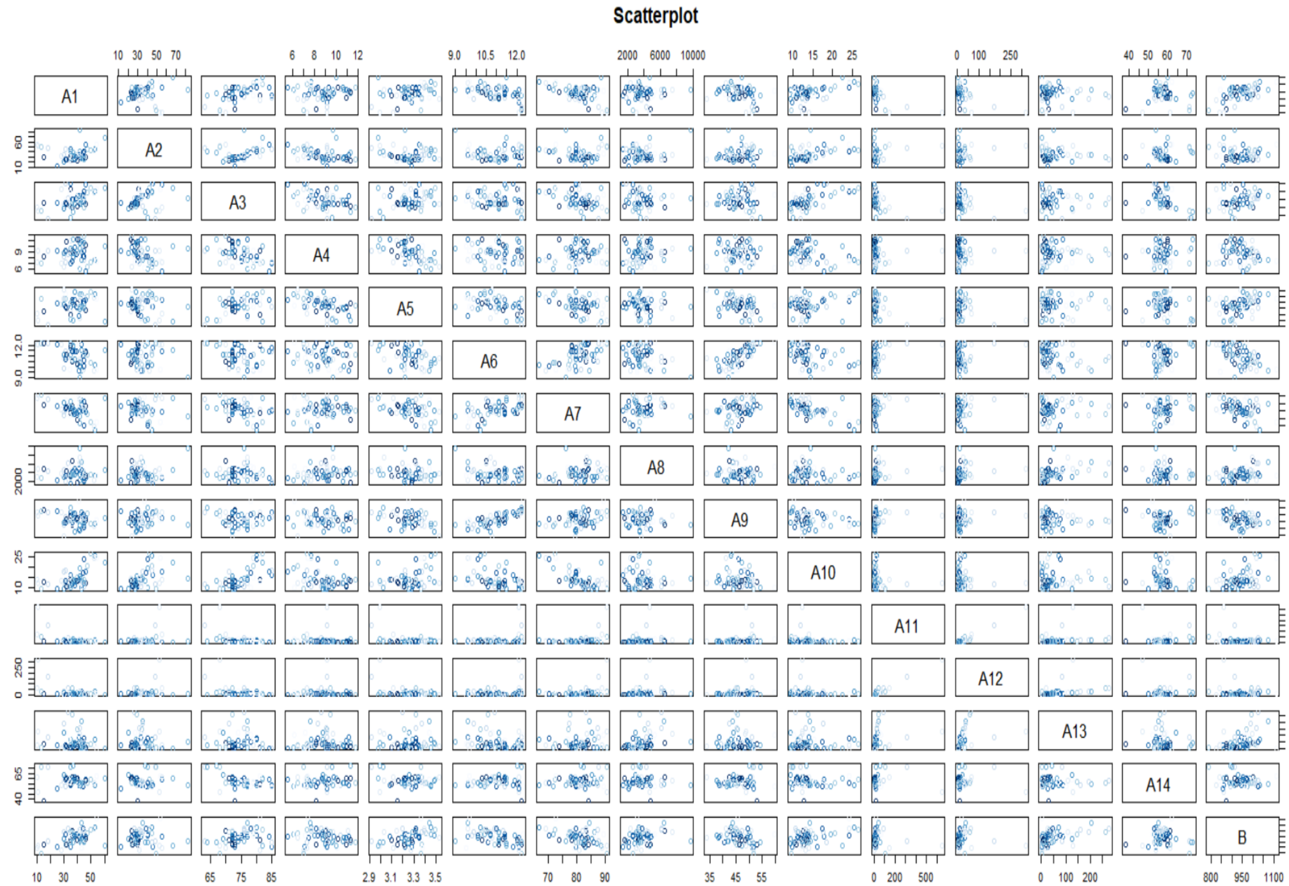


Fig (1.a)

The graph shows that all variable except the variables A11 & A12 show linear relationship.

For clearer picture, visualization of the scatterplot for these variables and comparing them with the other two variables showing linear relationship,

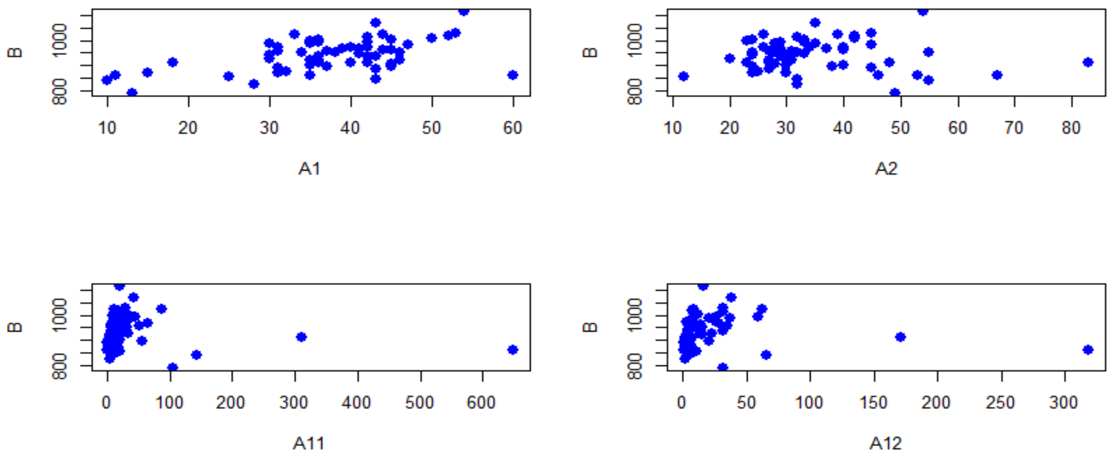


Fig (1.b)

ANALYSIS OF THE FULL MODEL.

A.1 Fitting the model and Hypothesis Testing:

The fitted Model obtained is ,

B =1818.708479+ 3.35*A1-2.95466*A2-2.600*A3-27.28804*A4-113.48783*A5- 25.96596*A6-1.14122*A7+ 0.01391*A8+ 1.76738*A9+ 4.21249*A10 - 0.94825*A11+ 2.18534*A12-0.01936*A13+ 1.26610*A14.

Hypothesis Testing:

```
> summary(data_regression)

Call:
lm(formula = B ~ A1 + A2 + A3 + A4 + A5 + A6 + A7 + A8 + A9 +
    A10 + A11 + A12 + A13 + A14, data = data1)

Residuals:
    Min      1Q  Median      3Q     Max
-81.529 -13.121  -2.039  13.875  78.748

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.819e+03  4.395e+02   4.138 0.000152 ***
A1           3.350e+00  7.521e-01   4.455 5.51e-05 ***
A2          -2.954e+00  6.570e-01  -4.496 4.83e-05 ***
A3          -2.601e+00  1.894e+00  -1.373 0.176471
A4          -2.729e+01  6.503e+00  -4.196 0.000126 ***
A5          -1.135e+02  6.638e+01  -1.710 0.094238 .
A6          -2.597e+01  1.199e+01  -2.166 0.035623 *
A7          -1.141e+00  1.570e+00  -0.727 0.471121
A8           1.391e-02  4.151e-03   3.351 0.001638 **
A9           1.767e+00  1.589e+00   1.112 0.271852
A10          4.212e+00  2.279e+00   1.849 0.071078 .
A11         -9.482e-01  4.839e-01  -1.960 0.056233 .
A12          2.185e+00  9.929e-01   2.201 0.032907 *
A13         -1.936e-02  1.522e-01  -0.127 0.899360
A14          1.266e+00  1.089e+00   1.162 0.251285
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 34.62 on 45 degrees of freedom
Multiple R-squared:  0.7638,    Adjusted R-squared:  0.6903
F-statistic: 10.39 on 14 and 45 DF,  p-value: 7.25e-10
```

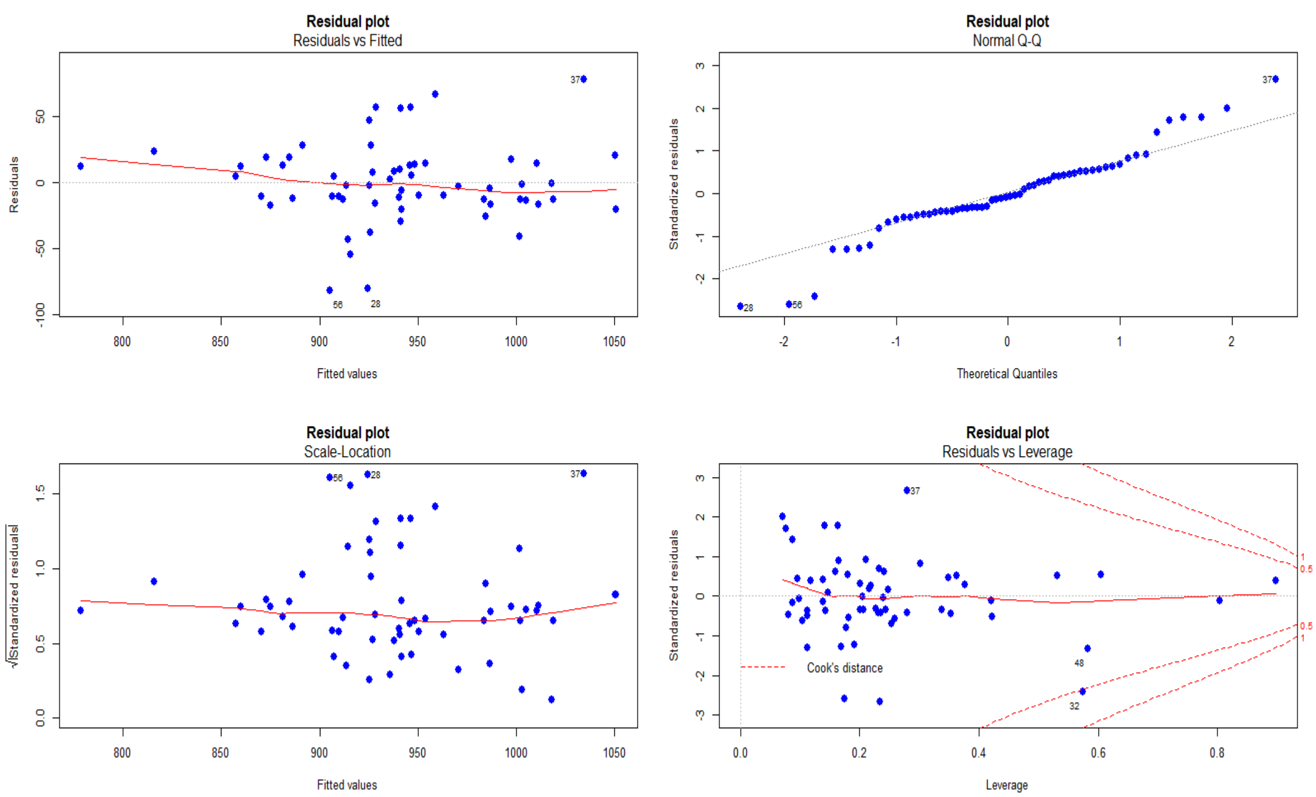
1. From the above results, it is clear that the overall model is significant as the p-value is very small.
2. For the individual coefficients, we see that the variables A3,A5,A7,A9,A10,A11,A13,A14 turn out to be insignificant.

A.2 Residual Analysis.

It is important that the Multiple linear Regression Model satisfies the following assumptions of regressions:

1. The residuals are independent.(No auto-correlation).
2. The residuals are normally distributed
3. Residuals show Homoscadesticity
4. There is no multicollinearity between the variables

These assumptions are checked by studying the residual plots.



Fig(A.2.1)

Deductions from the Residual Analysis:

1. The dotted line in the residuals vs fitted values plot is the line of our fitted model. The points above it are positive residuals and the ones below are the negative residuals visualized blue in colour. This plot shows that the residuals are random.
2. The normal Q-Q plot shows that most observations are on the straight line following the normality assumption except a few observations at the top and bottom ends.
3. The scale-location plot is a horizontal straight line displaying Homoscadesticity throughout.
4. The dotted lines in the Residuals vs leverage plot are the cook's distances. We see a few observations in this area indicating that these might be influential points. For our further analysis, the inclusion of these influential points needs to be justified or they have to be removed from the data to obtain better results.

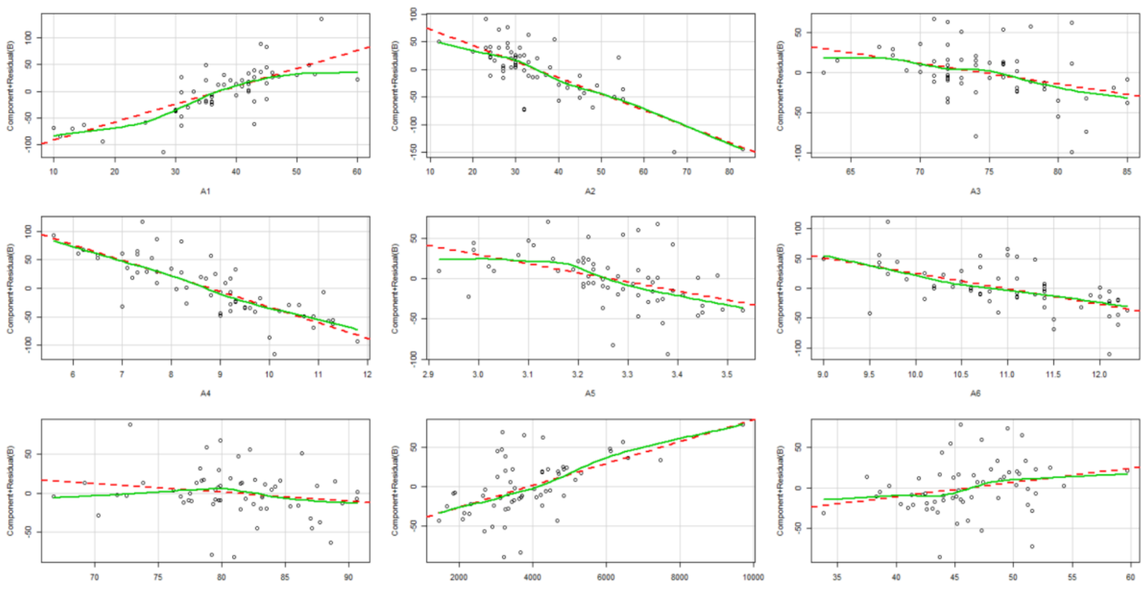
In addition to the scale-location plot the assumption of constant variance is verified by the breusch pagan test.

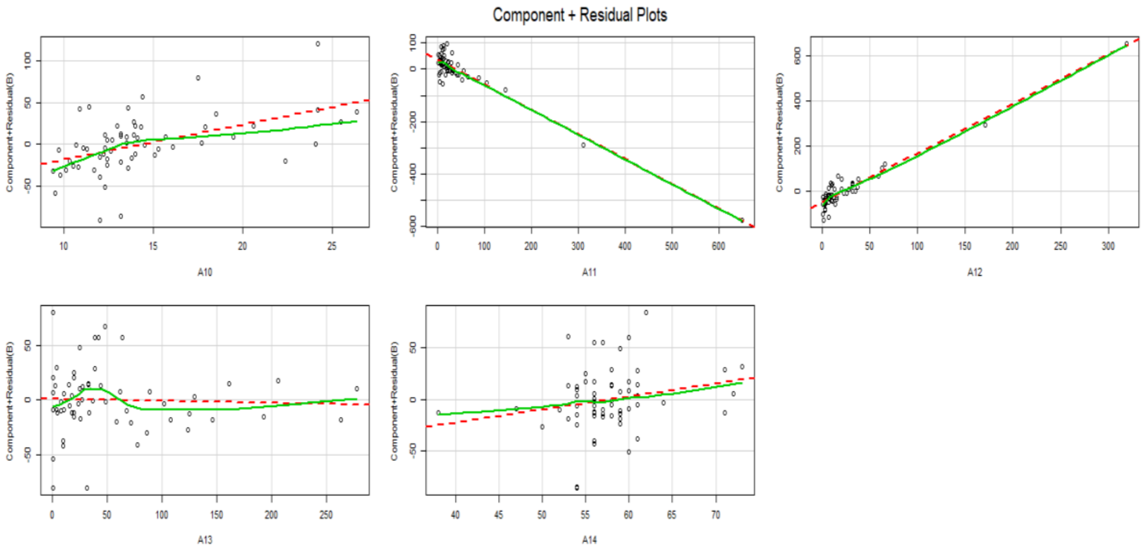
```
> ncvTest(data_regression) #test for non-constant variance
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 0.1635263    Df = 1     p = 0.6859305
```

A.3 Evaluation of Non-linearity between predictor and dependent variable.

The analysis of non-linearity is done using the component-residual plots for each variable.These plots show the relationship between the predictor and the dependent variable.

We can see that there is no significant difference in the residual line and component line for any of the variables. This implies that the predictors and the dependent variable have a linear relationship. We can also see a perfect overlap for a few variables.



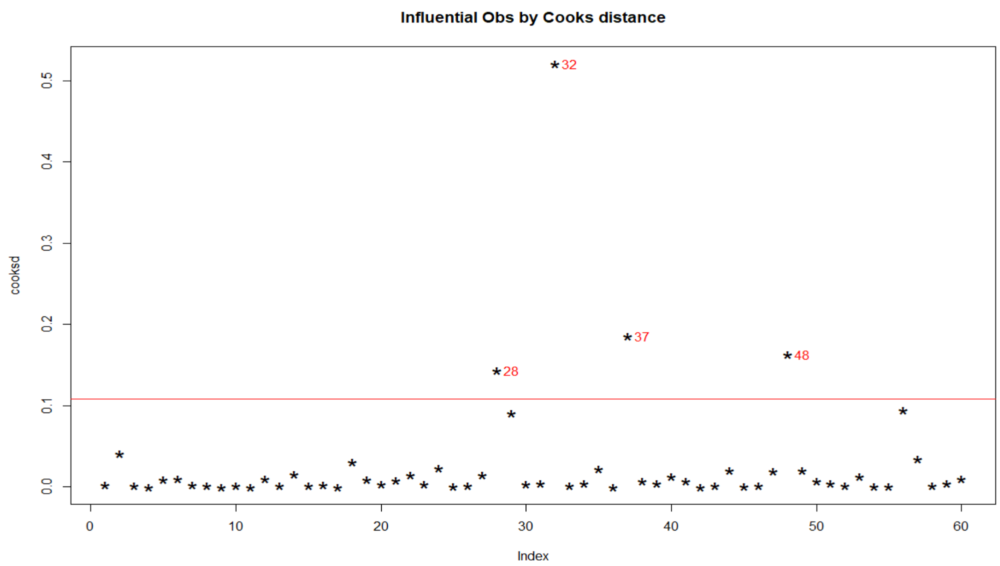


A.4 Study of Multicollinearity

```
> vif(data_regression)
       A1        A2        A3        A4        A5        A6        A7
 2.775642  3.047173  4.011431  4.465599  3.968329  5.053283  3.208267
       A8        A9       A10       A11       A12       A13       A14
 1.793299  2.659816  4.423172 97.493809 104.287199 4.583739  1.741655
```

Above, we see that the vif values for the variables A11 and A12 are very high. This show high multicollinearity between the two variables.we need to remove either of the two variables to satisfy the assumption of no multicollinearity between the dependent variables.

A.5 cook's Distance



The graph above indicates presence of four influential observations that might have affected the model significantly. The presence of these observations is not justified and thus there are removed in further analysis.

Variable Selection and Comparison of Models.

Methodology:

Three methods were used to select important variables in the model. The forward selection method resulted into the same original full model.Thus, the full model represented this model for the comparison of models. The other two models were obtained by backward selection method and stepwise variable selection method.

The three models were compared based on the statistical measures R-square,Adjusted R-square,Press residual,AIC and Mean squared error.

This information is tabulated as follows:

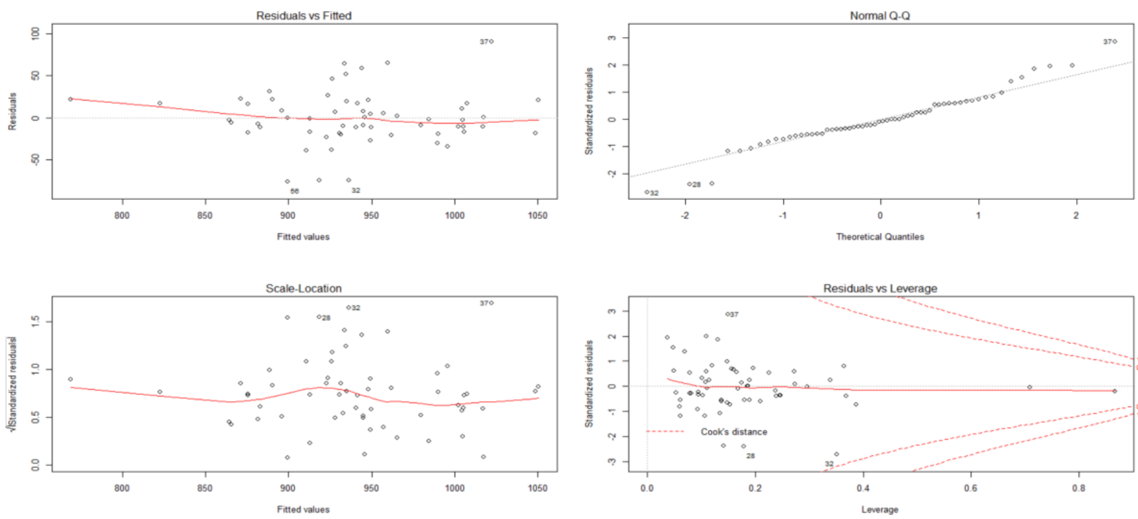
	R_square	Adj_R_sqaure	AIC	PRESS	MSE
full_model	0.7638	0.6903	438.0720	100216.90	34.62
step-wise	0.6680	0.6160	446.4826	108513.20	38.55
backward_selection	0.7462	0.6944	434.3680	83534.77	34.99

Based on this study, the model obtained by backward selection method was chosen to be the best model of the three models obtained

Analaysis of the selected Model

The Model by Backward selection Method is:
B=A1*x1+A2*x2+A3*x3+A4*x4+A5*x5+A6*x6+A8*x8+A10*x10+A11*x11+A12*x12+error

C.1 Residual Analysis



Deductions from the Residual plots:

The Residual analysis clearly show that all the assumptions of linear regression are satisfied.
However, we can still see a few influential/leverage points in the residual vs leverage plot which lie in the region of cook's distance

c.2 Detection and remedial measures for Influential points.

Using influential.measures in R programming , the influential points for this model were detected and removed for a better unbiased model.

sc.3 Study of Multicollinearity

```
> vif(backward.lm)
       A1        A2        A3        A4        A5        A6        A8 
 2.597147  2.492156  2.608558  3.841331  3.580952  2.535692  1.585898 
      A10       A11       A12 
 2.968115 45.262039 41.088311 
```

This model shows high multicollinearity between A11 and A12.
Remedial measure for multicollinearty in this situation was to remove either of the two variables. The model without A11 has higher R-Square value than the model without A12. Thus , the model without A12 was chosen.

After removing the A11 and the influential points,

The Model becomes:

$$B=A1*x1+A2*x2+A3*x3+A4*x4+A5*x5+A6*x6+A8*x8+A10*x10+A12*x12+error$$

C.4 Testing overall Significance and Hypothesis testing for individual variables.

1. The overall model is significant.
2. Except A3 and A5 all variables are significant

```
> summary(a.lm)

Call:
lm(formula = B ~ A1 + A2 + A3 + A4 + A5 + A6 + A8 + A10 + A12,
    data = a)

Residuals:
    Min      1Q  Median      3Q     Max
-61.055 -13.882  -2.647  11.128  67.221

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.581e+03  3.051e+02   5.184 5.85e-06 ***
A1           4.139e+00  6.355e-01   6.513 7.25e-08 ***
A2          -2.713e+00  4.820e-01  -5.628 1.36e-06 ***
A3          -1.802e+00  1.312e+00  -1.373 0.177150
A4          -1.943e+01  5.048e+00  -3.848 0.000399 ***
A5          -9.424e+01  5.318e+01  -1.772 0.083622 .
A6          -1.953e+01  7.233e+00  -2.701 0.009930 **
A8           1.261e-02  3.334e-03   3.783 0.000484 ***
A10          3.951e+00  1.647e+00   2.399 0.020971 *
A12          1.203e+00  3.078e-01   3.910 0.000331 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 27.19 on 42 degrees of freedom
Multiple R-squared:  0.802,     Adjusted R-squared:  0.7595
F-statistic:  18.9 on 9 and 42 DF,  p-value: 4.25e-12
```

c.5 Comparison of selected and other models

Compared to the other model, the selected model has highest R-square value.
Also, AIC ,PRESS and MSE significantly reduced.

R_square	Adj_R_Square	AIC	PRESS	MSE
0.802	0.7597	352.4095	43294.52	27.19

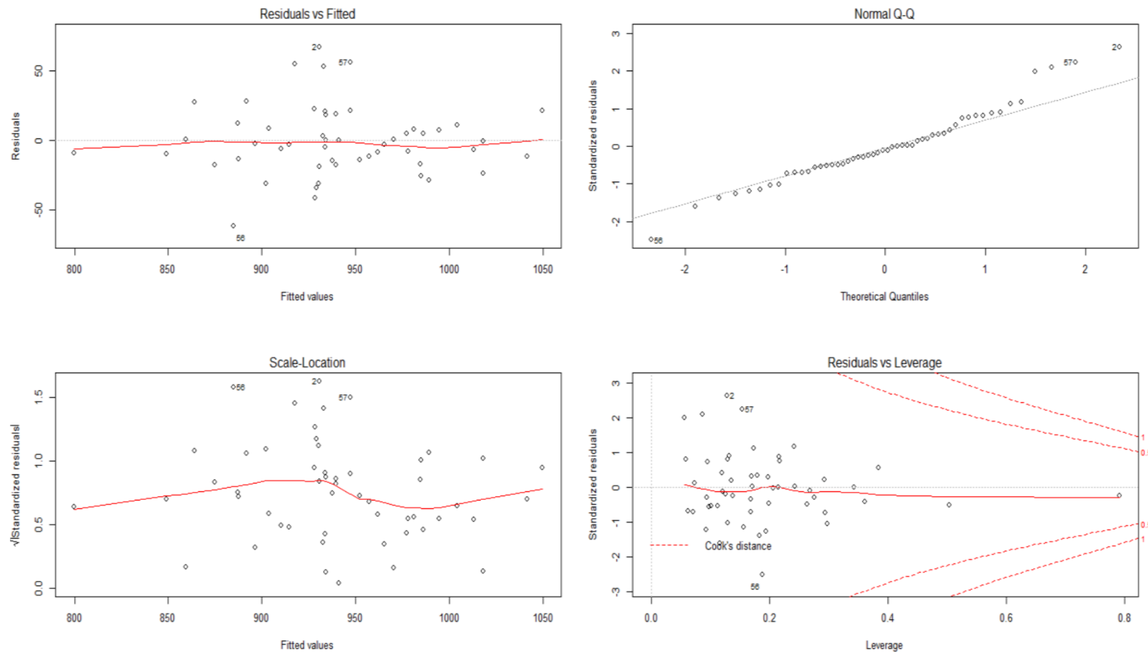
	R_square	Adj_R_sqaure	AIC	PRESS	MSE
full_model	0.7638	0.6903	438.0720	100216.90	34.62
step-wise	0.6680	0.6160	446.4826	108513.20	38.55
backward_selection	0.7462	0.6944	434.3680	83534.77	34.99

c.7 Checking the multicollinearity of the selected model.

The vif values for all the variables are less than 5 verifying that there is no multicollinearity between the variables.

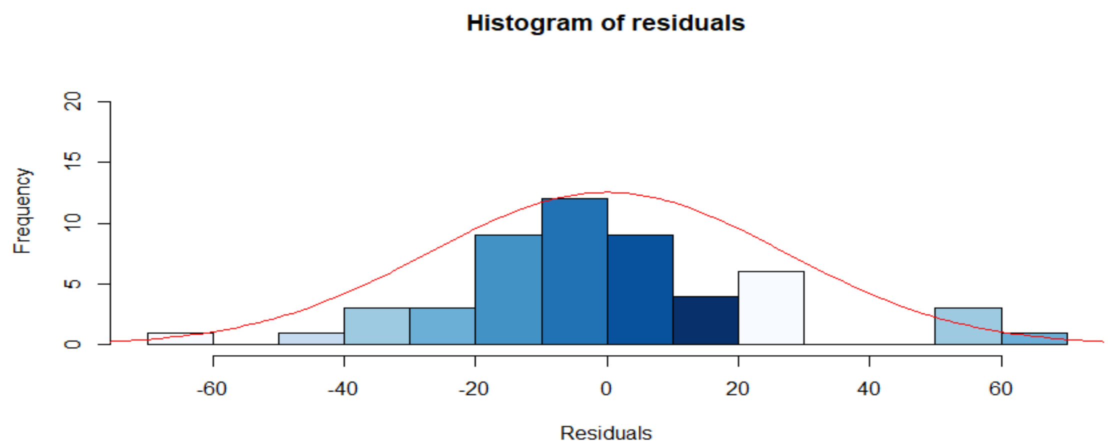
```
> vif(a.lm)
      A1       A2       A3       A4       A5       A6       A8      A10
1.867316 1.988473 2.445840 4.147898 2.872253 2.388618 1.746328 2.996818
     A12
1.379458
```

c.8 Residual Analysis of the modified selected model.



The residual plot shows that the chosen modified model obtained by the backward selection method satisfies the assumptions of regression.

Normality of Residual



The histogram clearly displays the normality of the residuals in addition to the Normal probability plot in the Residual analysis.

Test for Autocorrelation of residuals

The durbin Watson test clarifies that the residuals are independently distributed.

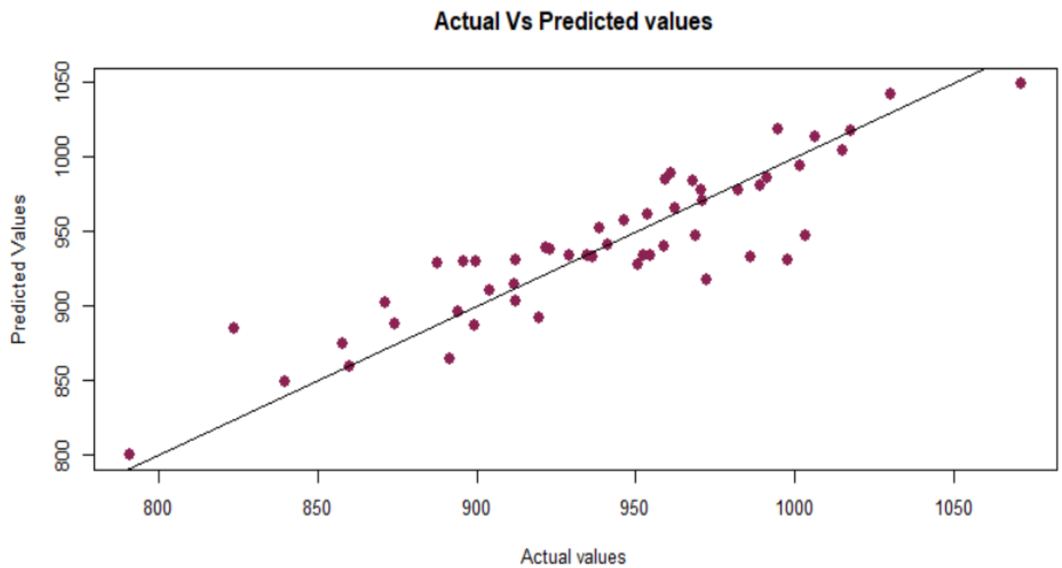
```
> durbinWatsonTest(a.lm)
 lag Autocorrelation D-W Statistic p-value
   1      -0.1009946      2.177965   0.578
 Alternative hypothesis: rho != 0
```

Validation of Analysis:

The confidence interval of each variable is as follows.

```
> confint(a.lm)
                    2.5 %        97.5 %
(Intercept)  9.657195e+02 2196.97838781
A1           2.856739e+00    5.42182720
A2          -3.685156e+00   -1.73986852
A3          -4.450289e+00    0.84713685
A4          -2.961273e+01   -9.23729525
A5          -2.015523e+02   13.07718508
A6          -3.413008e+01   -4.93789505
A8           5.885605e-03    0.01934353
A10          6.269250e-01    7.27487079
A12          5.821979e-01    1.82438010
```

To validate the data, the predicted values were plotted against the actual values for the model obtained. The graph of the same is as below.



The graph clearly shows a linear relationship indicating that the predicated values obtained are very close to the actual values of the model.
Thus, the model is obtained is verified and is best suitable for futher predictions.

Conclusion and Recommendation:

```
> coef(a.lm)
  (Intercept)           A1           A2           A3           A4 
1581.34896823   4.13928328  -2.71251244  -1.80157632 -19.42501145 
           A5           A6           A8          A10          A12 
 -94.23754909 -19.53398614   0.01261457   3.95089788   1.20328899 
```

Conclusion:

We know that for a unit increase in a variable there is a significant change in the response variable equivalent to the coefficient of the variable.

From the coefficients obtained for the model , we can understand that

1. For a unit increase in A4: the size of the population older than 65 there is a decrease of 19.42 units in the death rate. In my understanding the reason for this might be the development in the field of medicine that has resulted in people having longer lives in comparison to the old days.
2. Similarly, for an unit increase in A5: the number of members per household there is a significant decrease of 94.237 units in the death rate. This can be partly answered with the fact that more people support the living of a family. But it remains questionable if the members of household are unable to support the living of the household.(children,handicap,unemployed etc.)
3. The change caused by A6: the number of years of schooling for persons over 22; Can be explained easily as more education implies employment,ethical and hygienic habits etc leading to lesser deaths.
4. For A10: the number of families with an income less than $3000 the death rate variable increases by a factor of 3. The reason could be insufficiency of food and resources and unavailability of health care facilities.
5. Under hypothetical condition that the coefficients of all variables would be zero the death rate would be high (1581.34897).
   Other variables can also be interpreted in a similar way.

Recommendations:

- The backward selection Model is best suited for Prediction of Mortality rate.
- More observations can be used for better and reliable result
- As this study is based on time and area accuracy, the data might need to be updated for any future studies.

References:

1. Data link: http://people.sc.fsu.edu/~jburkardt/datasets/regression/x28.txt
2. Data link: http://people.sc.fsu.edu/~jburkardt/datasets/regression/x11.txt
3. http://docs.statwing.com/interpreting-residual-plots-to-improve-your-regression/#transform
4. https://rstudio-pubs-static.s3.amazonaws.com/155304_cc51f448116744069664b35e7762999f.html
5. http://michael.hahsler.net/SMU/EMIS7331/R/regression.html
6. http://analyticspro.org/2016/03/02/r-tutorial-multiple-linear-regression/
7. http://www.data-mania.com/blog/a-5-step-checklist-for-multiple-linear-regression/
8. https://www.princeton.edu/~otorres/NiceOutputR.pdf
9. https://www.statmethods.net/stats/rdiagnostics.html
10. https://www.r-bloggers.com/how-to-detect-heteroscedasticity-and-rectify-it/
11. http://r-statistics.co/Outlier-Treatment-With-R.html
12. https://stats.stackexchange.com/questions/40884/f-test-for-lack-of-fit-using-r

Appendix:

Codes used:

```
data<-read.table(file.choose(),header = T)
data
plot(B~A1+A2+A3,data=data)
plot(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14+A15,data=data)
plot(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14+A15,data=data)
REG<-
lm(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14+A15,data=data)
REG
summary(REG)
anova(REG)
plot(REG)
data<-read.table(file.choose(),header = T)
data
data<-read.table(file.choose(),header = T)
data<-read.table(file.choose(),header = T)
data
plot(data,col="blue")
plot(data,col="blues.9")
plot(data,col=blues.9)
plot(data,col=blue.9)
plot(data,col="red")
data
plot(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14+A15,COL="blue")
plot(data,col=blues9)
plot(data,col=blues9,main="Scatterplot")
summary(data)
data.frame(summary(data))
plot(B~A1A12+A13,data=data)
plot(B~A1+A12+A13,data=data)
plot(B~A1+A12+A13,data=data,col=blues9)
plot(B~A1+A2+A12+A13,data=data,col=blues9)
par(mfrow=c(2,2))
plot(B~A1+A2+A12+A13,data=data,col=blues9)
plot(B~A1+A2+A12+A13,data=data,col="red")
plot(B~A1+A2+A12+A13,data=data,col="red",pch=20)
par(mfrow=c(2,2))
plot(B~A1+A2+A12+A13,data=data,col="red",pch=20,cex=2)
install.packages("stargazer")
library(stargazer)
```

```
setwd("D:/APP STATS/741/Project")
stargazer(data, type = "text", title="Descriptive statistics", digits=1, out="table1.txt")
summary(data)
data_regression<-
lm(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14+A15,data = data)
data_regression
data_regression<-lm(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A14+A15,data
= data)
data_regression
summary(data_regression1)
data_regression<-
lm(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14+A15,data = data)
data_regression
data_regression1<-lm(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A14+A15,data
= data)
data_regression1
summary(data_regression)
summary(data_regression1)
coef(data_regression)
coef(data_regression1)
vif(data_regression)
install.packages(leaps)
install.packages("leaps")
library(leaps)
vif(data_regression)
install.packages("cars")
install.packages("CARS")
library(CARS)
vif(data_regression)
install.packages("car")
library(car)
vif(data_regression)
vif(data_regression1)
influence.measures(data_regression)
```

```
plot(B~A1+A2+A12+A13,data=data1,col="blue",pch=20,cex=2)
par(mfrow=c(2,2))
plot(B~A1+A2+A11+A12,data=data1,col="blue",pch=20,cex=2)
install.packages("stargazer")
library(stargazer)
stargazer(data, type = "text", title="Descriptive statistics", digits=1, out="table1.txt")
data_regression<-lm(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14,data = data1)
data_regression
par(mfrow=c(2,2))
plot(data_regression,main = "Residual plots")
plot(data_regression,main = "Residual plots",pch=20,cex=2.5,col="blue")
plot(data_regression,main = "Residual plots",pch=20,cex=2,col="blue")
plot(data_regression,main = "Residual plots",pch=20,cex=1.5,col="blue")
plot(data_regression,main = "Residual plots",pch=20,cex=2,col="blue")
plot(data_regression,main = "Residual plot",pch=20,cex=2,col="blue")
summary(data_regression)
Anova(data_regression)
anova(data_regression)
vif(data1)
library(car)
vif(data1)
vcov(data1)
a<-as.numeric(data1)
vif(a)
a<-as.numeric(data1)
vif(a)
install.packages("leaps")
library(leaps)
cp<-plot(14,cp)
cp<-plot(14,Cp)
?leaps
vif(data_regression)
install.packages("gvlma")
library(gvlma)
gvmodel<-gvlma(data_regression)
gvmodel
crPlots(data_regression)
```

```
ncv.test(data_regression)
ncvTest(data_regression)
?ncvTest
ceresPlots(data_regression)
?ceresPlots
?crPlots
av.plots(data_regression)
influencePlot(data_regression,id.method="identify", main="Influence Plot")
?influence.measures
influence.measures(data_regression)
data.fwd<-regsubsets(B~.,data = data1,method = "forward")
summary(data.fwd)
data.bwd<-regsubsets(B~.,data = data1,method = "backward")
summary(data.bwd)
data.seq<-regsubsets(B~.,data = data1,method = "seqrep")
summary(data.seq)
null<-lm(B~1,data=data1)
step(null,scope = list(upper=data_regression,lower=null),direction = "forward")
step(null,scope = list(upper=data_regression,lower=null),direction = "backward")
step(data_regression,direction = "backward",trace = 0)
step(null,scope = list(upper=data_regression,lower=null),direction = "both")
step(null,scope = list(upper=data_regression,lower=null),direction = "forward")
forward.lm<-lm( B ~ A6 + A13 + A1 + A4 + A2 + A8 + A10 + A3 + A5,
data = data1)
plot(forward.lm)
summary(forward.lm)
plot(forward.lm)
plots(leaps,scale="bic")
plot(leaps,scale="bic")
influence.measures(forward.lm)
summary(forward.lm)
summary(data_regression)
influence.measures(forward.lm)
a<-data1[-c(12,32,37,40,48,49,59),]
forward.lm<-lm( B ~ A6 + A13 + A1 + A4 + A2 + A8 + A10 + A3 + A5,
data = a)
summary(forward.lm)
influence.measures(forward.lm)
```

```
vcov(as.numeric(data1))
cov(data1)
vif(forward.lm)
install.packages("MPV")
library(MPV)
PRES
PRESS(data_regression)
PRESS(forward.lm)
confint(data.bwd)
confint(data1)
predict(data_regression,interval = "confidence")
yhat<-data_regression$fit
r<-residua(data_regression)
r<-residuals(data_regression)
plot(r,yhat)
boxplot(data1)
boxplot(data1$A1)
boxplot(data1$A2)
?mallows cp
?mallows
install.packages("CombMSC")
library(CombMSC)
Cp(data_regression)
Cp
?`CombMSC-package`
?CombMSC
malow<-subsets(data_regression,statistic="cp")
boxplot(B~A1+A2+A3+A4+A5+A6,data = data1)
>boxplot(B~A1,data = data1)
>boxplot(B~A1)
mod<-lm(B~.,data = data1)
cooksd<-cooks.distance(mod)
plot(cooksd, pch="*", cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
abline(h = 4*mean(cooksd, na.rm=T), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
plot(cooksd, pch=22, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
```

```
plot(cooksd, pch=20, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=18, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=17, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=15, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=10, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=8, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=3, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=6, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
?pch
plot(cooksd, pch=6, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
plot(cooksd, pch=8, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
?pch
plot(cooksd, pch=18, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
abline(h = 4*mean(cooksd, na.rm=T), col="red")
> text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
durbinWatsonTest(data_regression)
library(gvlma)
gvmodel<-gvlma(data_regression)
gvmodel
step(null,scope = list(upper=data_regression,lower=null),direction = "both")
step(null,scope = list(upper=data_regression,lower=null),direction = "forward")
step(null,scope = list(upper=data_regression,lower=null),direction = "forward")
step(null,scope = list(upper=data_regression,lower=null),direction = "forward")
step(data_regression,direction = "backward",trace = 0)
extractAIC(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11,data=data1))
extractAIC(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data=data1))
extractAIC(lm(B~A1+A2+A3+A4+A5+A6+A8+A10,data=data1))
step.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10,data=data1))
backward.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data=data1))
summary(step.lm)
summary(backward.lm)
plot(backward.lm)
cooks.distance(backward.lm)
plot(cooksd, pch=18, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
cooks.distance(backward.lm)
abline(h = 4*mean(cooksd, na.rm=T), col="red")
```

```
cooks.distance(backward.lm)
plot(cooksd, pch=18, cex=2, main="Influential Obs by Cooks distance") # plot cook's distance
cooks.distance(backward.lm)
abline(h = 4*mean(cooksd, na.rm=T), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
new_data<-data1[-c(28,32,37,48),]
backward.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data=new_data))
summary(backward.lm)
plot(backward.lm)
vif(backward.lm)
library(car)
cov(data1)
vif(backward.lm)
backward.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11,data=new_data))
vif(backward.lm)
summary(backward.lm)
plot(backward.lm)
backward.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=new_data))
vif(backward.lm)
summary(backward.lm)
fit<-fitted(backward.lm)
plot(a$B,fit)
full.lm<-lm(B~A1+A2+A3+A4+A5+A6+A7+A8+A9+A10+A11+A12+A13+A14,data = data1)
summary(full.lm)
extractAIC(full.lm)
step.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10,data = data1)
summary(step.lm)
extractAIC(step.lm)
backward.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data = data1)
summary(backward.lm)
extractAIC(backward.lm)
library(MPV)
PRESS(backward.lm)
PRESS(step.lm)
PRESS(full.lm)
R-SQAURE<-C(0.7638,0.668,0.7462)
R_SQUARE<-C(0.7638,0.668,0.7462)
a<-C(0.7638,0.668,0.7462)
a>-C(0.7638,0.668,0.7462)
```

```
a<-C(0.7638,0.668,0.7462)
a>-C(0.7638,0.668,0.7462)
extractAIC(full.lm)
a<-c(1,2,3)
r_square<-c(0.7638,0.668,0.7462)
adj_r_sqaure<-c(0.6903,0.616,0.6944)
AIC<-c(438.072,446.4826,434.368)
PRESS<-c(100216.9,108513.2,83534.77)
MSE<-c(34.62,38.55,34.99)
data.frame(r_square,adj_r_sqaure,AIC,PRESS,MSE)
row.names(c("full_model","step-wise","backward_selection"))
table<-data.frame(r_square,adj_r_sqaure,AIC,PRESS,MSE)
row.names(table)<-(c("full_model","step-wise","backward_selection"))
table
table<-data.frame(R_square,a\Adj_R_sqaure,AIC,PRESS,MSE)
table<-data.frame(R_square,Adj_R_sqaure,AIC,PRESS,MSE)
R_square<-c(0.7638,0.668,0.7462)
Adj_R_sqaure<-c(0.6903,0.616,0.6944)
table<-data.frame(R_square,Adj_R_sqaure,AIC,PRESS,MSE)
table
row.names(table)<-(c("full_model","step-wise","backward_selection"))
table
Adj_R_sqaure<-c(0.6903,0.616,0.6944)
backward.lm
summary(backward.lm)
par(mfrow=c(2,2))
plot(backward.lm)
vif(backward.lm)
backward.lm.1<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11,data=data1))
backward.lm.2<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=data1))
summary(backward.lm.1)
backward.lm.2
summary(backward.lm.2)
vif(backward.lm.1)
vif(backward.lm.2)
influence.measures(backward.lm.2)
```

```
vif(backward.lm.1)
vif(backward.lm.2)
influence.measures(backward.lm.2)
backward.data.1<-data1[-c(29,32,37,48,49,50,59),]
summary(backward.data.1)
backward.data.1.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=backward.data.1))
summary(backward.data.1.lm)
backward.data.2.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11,data=backward.data.1))
summary(backward.data.2.lm)
influence.measures(backward.lm.1)
cooksd_back<-cooks.distance(backward.data.2.lm)
plot(cooksd, pch=18, cex=2, main="Influential Obs by Cooks distance")
abline(h = 4*mean(cooksd, na.rm=T), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
backward.data.2.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=backward.data.2))
backward.data.2<-data1[-c(29,32,37,48,49,50,59),]
summary(backward.data.2)
backward.data.2.lm<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=backward.data.2))
summary(backward.data.2.lm)
cooksd_back<-cooks.distance(backward.data.2.lm)
plot(cooksd, pch=18, cex=2, main="Influential Obs by Cooks distance")
abline(h = 4*mean(cooksd, na.rm=T), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
outiers_removed<-backward.data.2[-c(32,28,37,48),]
backward.data.2.lm_removed<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=outiers_removed))
summary(backward.data.2.lm_removed
summary(backward.data.2.lm_removed)
PRESS(backward.data.2.lm_removed)
extractAIC(backward.data.2.lm_removed)
plot(backward.data.2.lm_removed)
plot(backward.data.2.lm_removed)
cooksd_REMOVED<-cooks.distance(backward.data.2.lm_removed)
plot(cooksd_REMOVED, pch=18, cex=2, main="Influential Obs by Cooks distance")
abline(h = 4*mean(cooksd, na.rm=T), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
plot(cooksd_back, pch=18, cex=2, main="Influential Obs by Cooks distance")
abline(h = 4*mean(cooksd, na.rm=T), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
```

```
PRESS(backward.lm)
backward.lm.1<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11,data=data1))
backward.lm.2<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=data1))
summary(backward.lm.1)
summary(backward.lm.2)
wo_influential_points<-data1[-c(28,32,37,48),]
backward.lm.ifp<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data = wo_influential_points)
summary(backward.lm.ifp)
influence.measures(backward.lm.ifp)
vif(backward.lm.ifp)
library(car)
library(MPV)
vif(backward.lm.ifp)
backward.lm.A12<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data = wo_influential_points)
summary(backward.lm.A12)
vif(backward.lm.A12)
influence.measures(backward.lm.A12)
bckwrd<-data1[-c(28,32,37,48),]
bckwrd.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10++A11+A12,data=bckwrd)
summary(bckwrd.lm)
influence.measures(backward.lm)
bckwrd.1<-data1[-c(7,29,32,37,40,48,49,50,55,59),]
bckward.1.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data = bckward.1)
bckward.1.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data = bckwrd.1)
summary(bckward.1.lm)
vif(bckward.1.lm)
bckward.A11.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data = bckwrd.1)
summary(bckward.A11.lm)
bckward.A12.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11,data = bckwrd.1)
summary(bckward.A12.lm)
library(MPV)
library(car)
library(leaps)
```

```
res<-cooks_removed.lm$residuals
plot(res)
res
plot(res)
hist(res)
cooksd_back<-cooks.distance(backward.data.2.lm)
plot(cooksd, pch=18, cex=2, main="Influential Obs by Cooks distance")
abline(h = 4*mean(cooksd, na.rm=T), col="red")
text(x=1:length(cooksd)+1, y=cooksd, labels=ifelse(cooksd>4*mean(cooksd, na.rm=T),names(cooksd),""), col="red")
summary(backward.lm)
influence.measures(backward.lm)
bckward<-data1[-c(7,29,32,37,40,48,49,50,55,59),]
bckwrd.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11+A12,data=bckward)
summary(bckwrd.lm)
vif(bckwrd.lm)
bckwrd.lm.1<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=bckward)
summary(bckward.lm.1)
summary(bckwrd.lm.1)
influence.measures(bckwrd.lm.1)
a<-bckwrd[-c(12,18,28,41),]
a.lm<-lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A12,data=a)
summary(a)
summary(a.lm)
PRESS(a.lm)
extractAIC(a.lm)
vif(a.lm)
cooks.distance(a.lm)
cooks.A11<-cooks.distance(a.lm)
plot(cooks.a, pch=18, cex=2, main="Influential Obs by Cooks distance")
plot(cooks.A11, pch=18, cex=2, main="Influential Obs by Cooks distance")
abline(h = 4*mean(cooks.A11, na.rm=T), col="red")
text(x=1:length(cooks.A11)+1, y=cooks.A11, labels=ifelse(cooks.A11>4*mean(cooks.A11, na.rm=T),names(cooks.A11),""), col="red")
Anova(a.lm)
```

```
durbinWatsonTest(a.lm)
?durbinWatsonTest
confint(a.lm)
predict(a.lm)
predict.lm(a.lm)
temp_var<-predict(a.lm,interval = "prediction")
new_df<-cbind(bckwrd$B,temp_var)
new_df<-cbind(bckward,new_df)
new_df<-cbind(bckward,temp_var)
fit<-fitted.values(a.lm)
plot(a$B,fit,pch=20,cex=2,col=blues9)
plot(a$B,fit,pch=20,cex=2,col=blues9,xlab = "Actual values",ylab = "Predicted Values",main = "Actual Vs Predicted values")
plot(a$B,fit,pch=20,cex=2,col="violetred4",xlab = "Actual values",ylab = "Predicted Values",main = "Actual Vs Predicted values")
abline(a.lm)
lines(a$B,fit)
plot(a$B,fit,pch=20,cex=2,col="violetred4",xlab = "Actual values",ylab = "Predicted Values",main = "Actual Vs Predicted values")
lines(fit)
lines(a$B)
lines(a$B,)
lines(,fit)
lines(1,fit)
abline(a.lm)
lines(a.lm)
abline(a=0,b=1)
pred<-predict(a.lm)
pred
plot(a$B,pred,pch=20,cex=2,col="violetred4",xlab = "Actual values",ylab = "Predicted Values",main = "Actual Vs Predicted values")
abline(a=0,b=1)
backward.lm.1<-(lm(B~A1+A2+A3+A4+A5+A6+A8+A10+A11,data=data1))
vif(backward.lm)
summary(a.lm)
summary(a.lm)
PRESS(a.lm)
Press<-PRESS(a.lm)
AIC<-extractAIC(a.lm)
MSE<-27.19
r_square<-0.802
Adj_R_sqaure<-0.7595
```

Data:

```
 1   36   27   71    8.1   3.34   11.4   81.5   3243    8.8   42.6   11.7    21
15    59   59    921.870

 2   35   23   72   11.1   3.14   11.0   78.8   4281    3.6   50.7   14.4     8
10    39   57    997.875

 3   44   29   74   10.4   3.21    9.8   81.6   4260    0.8   39.4   12.4     6
6    33   54    962.354

 4   47   45   79    6.5   3.41   11.1   77.5   3125   27.1   50.2   20.6    18
8    24   56    982.291

 5   43   35   77    7.6   3.44    9.6   84.6   6441   24.4   43.7   14.3    43
38   206   55   1071.289

 6   53   45   80    7.7   3.45   10.2   66.8   3325   38.5   43.1   25.5    30
32    72   54   1030.380

 7   43   30   74   10.9   3.23   12.1   83.9   4679    3.5   49.2   11.3    21
32    62   56    934.700

 8   45   30   73    9.3   3.29   10.6   86.0   2140    5.3   40.4   10.5     6
4     4   56    899.529

 9   36   24   70    9.0   3.31   10.5   83.2   6582    8.1   42.5   12.6    18
12    37   61   1001.902

10   36   27   72    9.5   3.36   10.7   79.3   4213    6.7   41.0   13.2    12
7    20   59    912.347

11   52   42   79    7.7   3.39    9.6   69.2   2302   22.2   41.3   24.2    18
8    27   56   1017.613
```

12	33	26	76	8.6	3.20	10.9	83.4	6122	16.3	44.9	10.7	88	63	278	58	1024.885
13	40	34	77	9.2	3.21	10.2	77.0	4101	13.0	45.7	15.1	26	26	146	57	970.467
14	35	28	71	8.8	3.29	11.1	86.3	3042	14.7	44.6	11.4	31	21	64	60	985.950
15	37	31	75	8.0	3.26	11.9	78.4	4259	13.1	49.6	13.9	23	9	15	58	958.839
16	35	46	85	7.1	3.22	11.8	79.9	1441	14.8	51.2	16.1	1	1	1	54	860.101
17	36	30	75	7.5	3.35	11.4	81.9	4029	12.4	44.0	12.0	6	4	16	58	936.234
18	15	30	73	8.2	3.15	12.2	84.2	4824	4.7	53.1	12.7	17	8	28	38	871.766
19	31	27	74	7.2	3.44	10.8	87.0	4834	15.8	43.5	13.6	52	35	124	59	959.221
20	30	24	72	6.5	3.53	10.8	79.5	3694	13.1	33.8	12.4	11	4	11	61	941.181
21	31	45	85	7.3	3.22	11.4	80.7	1844	11.5	48.1	18.5	1	1	1	53	891.708
22	31	24	72	9.0	3.37	10.9	82.8	3226	5.1	45.2	12.3	5	3	10	61	871.338
23	42	40	77	6.1	3.45	10.4	71.8	2269	22.7	41.4	19.5	8	3	5	53	971.122
24	43	27	72	9.0	3.25	11.5	87.1	2909	7.2	51.6	9.5	7	3	10	56	887.466
25	46	55	84	5.6	3.35	11.4	79.7	2647	21.0	46.9	17.9	6	5	1	59	952.529
26	39	29	76	8.7	3.23	11.4	78.6	4412	15.6	46.6	13.2	13	7	33	60	968.665
27	35	31	81	9.2	3.10	12.0	78.3	3262	12.6	48.6	13.9	7	4	4	55	919.729
28	43	32	74	10.1	3.38	9.5	79.2	3214	2.9	43.7	12.0	11	7	32	54	844.053
29	11	53	68	9.2	2.99	12.1	90.6	4700	7.8	48.9	12.3	648	319	130	47	861.833
30	30	35	71	8.3	3.37	9.9	77.4	4474	13.1	42.6	17.7	38	37	193	57	989.265
31	50	42	82	7.3	3.49	10.4	72.5	3497	36.7	43.3	26.4	15	10	34	59	1006.490
32	60	67	82	10.0	2.98	11.5	88.6	4657	13.6	47.3	22.4	3	1	1	60	861.439
33	30	20	69	8.8	3.26	11.1	85.4	2934	5.8	44.0	9.4	33	23	125	64	929.150

34	25	12	73	9.2	3.28	12.1	83.1	2095	2.0	51.9	9.8	20	11	26	50	857.622
35	45	40	80	8.3	3.32	10.1	70.3	2682	21.0	46.1	24.1	17	14	78	56	961.009
36	46	30	72	10.2	3.16	11.3	83.2	3327	8.8	45.3	12.2	4	3	8	58	923.234
37	54	54	81	7.4	3.36	9.7	72.8	3172	31.4	45.5	24.2	20	17	1	62	1113.156
38	42	33	77	9.7	3.03	10.7	83.5	7462	11.3	48.7	12.4	41	26	108	58	994.648
39	42	32	76	9.1	3.32	10.5	87.5	6092	17.5	45.3	13.2	29	32	161	54	1015.023
40	36	29	72	9.5	3.32	10.6	77.6	3437	8.1	45.5	13.8	45	59	263	56	991.290
41	37	38	67	11.3	2.99	12.0	81.5	3387	3.6	50.3	13.5	56	21	44	73	893.991
42	42	29	72	10.7	3.19	10.1	79.5	3508	2.2	38.3	15.7	6	4	18	56	938.500
43	41	33	77	11.2	3.08	9.6	79.9	4843	2.7	38.6	14.1	11	11	89	54	946.185
44	44	39	78	8.2	3.32	11.0	79.9	3768	28.6	49.5	17.5	12	9	48	53	1025.502
45	32	25	72	10.9	3.21	11.1	82.5	4355	5.0	46.4	10.8	7	4	18	60	874.281
46	34	32	79	9.3	3.23	9.7	76.8	5160	17.2	45.1	15.3	31	15	68	57	953.560
47	10	55	70	7.3	3.11	12.1	88.9	3033	5.9	51.0	14.0	144	66	20	61	839.709
48	18	48	63	9.2	2.92	12.2	87.7	4253	13.7	51.2	12.0	311	171	86	71	911.701
49	13	49	68	7.0	3.36	12.2	90.7	2702	3.0	51.9	9.7	105	32	3	71	790.733
50	35	40	64	9.6	3.02	12.2	82.5	3626	5.7	54.3	10.1	20	7	20	72	899.264
51	45	28	74	10.6	3.21	11.1	82.6	1883	3.4	41.9	12.3	5	4	20	56	904.155
52	38	24	72	9.8	3.34	11.4	78.0	4923	3.8	50.5	11.1	8	5	25	61	950.672
53	31	26	73	9.3	3.22	10.7	81.3	3249	9.5	43.9	13.6	11	7	25	59	972.464
54	40	23	71	11.3	3.28	10.3	73.8	1671	2.5	47.4	13.5	5	2	11	60	912.202
55	41	37	78	6.2	3.25	12.3	89.5	5308	25.9	59.7	10.3	65	28	102	52	967.803

56	28	32	81	7.0	3.27	12.1	81.0	3665	7.5	51.6	13.2	4	2	1	54	823.764
57	45	33	76	7.7	3.39	11.3	82.2	3152	12.1	47.3	10.9	14	11	42	56	1003.502
58	45	24	70	11.8	3.25	11.1	79.8	3678	1.0	44.8	14.0	7	3	8	56	895.696
59	42	83	76	9.7	3.22	9.0	76.2	9699	4.8	42.2	14.5	8	8	49	54	911.817
60	38	28	72	8.9	3.48	10.7	79.8	3451	11.7	37.5	13.0	14	13	39	58	954.442