Multiple Regression Analysis

```
library(tidyverse)

library(knitr)

library(kableExtra)

library(treemap)

library(ggthemes)

library(highcharter)

library(summarytools)

library(corrplot)

library(formattable)

library(ggcorrplot)



# Loading the packages

options(warn = -1)

packags <- c("tidyverse", "knitr", "kableExtra","ggthemes", "treemap", "highcharter", "summarytools",
"ggcorrplot", "knitr","formattable")

purrr::walk(packags, library, character.only = T, quietly = T)


data <-read.csv("C:/Users/nitis/OneDrive/Desktop/Subject Semester 2/MVA/Project/Suicide Rate
Analysis.csv")


data

#data= Suicide.Rate.Analysis

names(data)[1]= "country"

data


# Data summary

# We are using str() & head() function to inspect and have a brief overwiew of the dataset.
```

```
str(data)

summary(data)

######################################################################################
###############################################


#Multivariate Analysis


#correlation between pairs of variables

data[,sapply(data, is.numeric)] %>%     cor(use = "complete.obs")

fit1 <- lm(suicides_no~sex+population+suicides.100k.pop+country+HDI.for.year+generation, data=data)

summary(fit1)


# country has most p-value among variables so we can remove it

fit2 <- lm(suicides_no~sex+population+suicides.100k.pop+HDI.for.year+generation, data=data)

anova(fit1, fit2)


#The results of anova shows that p-value>0.05 so we can accept null hypothesis(country is not
important)

coefficients(fit2)


library(GGally)

ggpairs(data=data, title="Data")

confint(fit2,level=0.95)

anova(fit2)

vcov(fit2)

cov2cor(vcov(fit2))

temp <- influence.measures(fit2)

temp

View(temp)
```

```
#diagnostic plots

plot(fit2)



# Assessing Outliers

library(car)

outlierTest(fit2)

qqPlot(fit2, main="QQ Plot")

leveragePlots(fit2) # leverage plots

plot(fit2)


# Influence Plot

library(mvinfluence)


influencePlot(fit2, id.method="identify", main="Influence Plot", sub="Circle size is proportial to Cook's Distance" )

# Normality of Residuals

# qq plot for studentized resid

qqPlot(fit2, main="QQ Plot")

# distribution of studentized residuals

library(MASS)

sresid <- studres(fit2)

hist(sresid, freq=FALSE,

     main="Distribution of Studentized Residuals")

xfit<-seq(min(sresid),max(sresid),length=40)

yfit<-dnorm(xfit)

lines(xfit, yfit)
```

```
#Non-constant Error Variance

# Evaluate homoscedasticity

# non-constant error variance test

ncvTest(fit2)

# plot studentized residuals vs. fitted values

spreadLevelPlot(fit2)

#Multi-collinearity

# Evaluate Collinearity

vif(fit2) # variance inflation factors

sqrt(vif(fit2)) > 2 # problem?

#Nonlinearity

# component + residual plot

crPlots(fit2)

# Ceres plots

#ceresPlots(fit2)


#Non-independence of Errors

#Test Autocorrelated Errors

durbinWatsonTest(fit2)


# Global test of model assumptions

library(gvlma)

gvmodel <- gvlma(fit2)

summary(gvmodel)

fit2

summary(fit2)

fit3 <- fit2

fit4 <- lm(suicides_no~sex+population+suicides.100k.pop+HDI.for.year+generation, data=data)

fit4
```

```
# compare models
anova(fit3, fit4)
step <- stepAIC(fit2, direction="both")
step$anova # display results

library(leaps)
leaps<-regsubsets(suicides_no~sex+population+suicides.100k.pop+HDI.for.year+generation, data=data, nbest=10)
# view results
summary(leaps)

# plot a table of models showing variables in each model.
# models are ordered by the selection statistic.
plot(leaps)
plot(leaps,scale="r2")
#subsets(leaps, statistic="rsq")

# All Subsets Regression
plot(leaps,scale="bic")
summary(leaps)
?regsubsets
summary(leaps)
View(leaps)
leaps
coef(leaps,1:5)

# Calculate Relative Importance for Each Predictor
library(relaimpo)
```

calc.relimp(fit2 ,type=c("lmg","last","first","pratt"),rela=TRUE)

Bootstrap Measures of Relative Importance (1000 samples)

boot <- boot.relimp(fit2, b = 1000, type = c("lmg", "last", "first", "pratt"), rank = TRUE, diff = TRUE, rela = TRUE)

booteval.relimp(boot) # print result

plot(booteval.relimp(boot,sort=TRUE)) # plot result

#https://rpubs.com/davoodastaraky/mtRegression

summary(fit2)

Output

```
> library(tidyverse)
> library(knitr)
> library(kableExtra)
> library(treemap)
> library(ggthemes)
> library(highcharter)
> library(summarytools)
> library(corrplot)
> library(formattable)
> library(ggcorrplot)
> # Loading the packages
> options(warn = -1)
> packags <- c("tidyverse", "knitr", "kableExtra","ggthemes", "treemap", "highcharter", "summarytools", "ggcorrplot", "knitr","formattable")
> purrr::walk(packags, library, character.only = T, quietly = T)
>
> data <-read.csv("C:/Users/nitis/OneDrive/Desktop/Subject Semester 2/MVA/Project/Suicide Rate Analysis.csv")
>
> data
   ï..country year    sex         age suicides_no population suicides.100k.pop country.year HDI.for.year gdp_for_year.... gdp_per_capita....       generation
1     Albania 1987   male 15-24 years          21     312900              6.71  Albania1987           NA    2,156,624,900                796     Generation X
2     Albania 1987   male 35-54 years          16     308000              5.19  Albania1987           NA    2,156,624,900                796           Silent
3     Albania 1987 female 15-24 years          14     289700              4.83  Albania1987           NA    2,156,624,900                796     Generation X
4     Albania 1987   male   75+ years           1      21800              4.59  Albania1987           NA    2,156,624,900                796 G.I. Generation
5     Albania 1987   male 25-34 years           9     274300              3.28  Albania1987           NA    2,156,624,900                796          Boomers
6     Albania 1987 female   75+ years           1      35600              2.81  Albania1987           NA    2,156,624,900                796 G.I. Generation
7     Albania 1987 female 35-54 years           6     278800              2.15  Albania1987           NA    2,156,624,900                796           Silent
8     Albania 1987 female 25-34 years           4     257200              1.56  Albania1987           NA    2,156,624,900                796          Boomers
9     Albania 1987   male 55-74 years           1     137500              0.73  Albania1987           NA    2,156,624,900                796 G.I. Generation
10    Albania 1987 female  5-14 years           0     311000              0.00  Albania1987           NA    2,156,624,900                796     Generation X
11    Albania 1987 female 55-74 years           0     144600              0.00  Albania1987           NA    2,156,624,900                796 G.I. Generation
12    Albania 1987   male  5-14 years           0     338200              0.00  Albania1987           NA    2,156,624,900                796     Generation X
13    Albania 1988 female   75+ years           2      36400              5.49  Albania1988           NA    2,126,000,000                769 G.I. Generation
14    Albania 1988   male 15-24 years          17     319200              5.33  Albania1988           NA    2,126,000,000                769     Generation X
15    Albania 1988   male   75+ years           1      22300              4.48  Albania1988           NA    2,126,000,000                769 G.I. Generation
16    Albania 1988   male 35-54 years          14     314100              4.46  Albania1988           NA    2,126,000,000                769           Silent
17    Albania 1988   male 55-74 years           4     140200              2.85  Albania1988           NA    2,126,000,000                769 G.I. Generation
18    Albania 1988 female 15-24 years           8     295600              2.71  Albania1988           NA    2,126,000,000                769     Generation X
19    Albania 1988 female 55-74 years           3     147500              2.03  Albania1988           NA    2,126,000,000                769 G.I. Generation
20    Albania 1988 female 25-34 years           5     262400              1.91  Albania1988           NA    2,126,000,000                769          Boomers
21    Albania 1988   male 25-34 years           5     279900              1.79  Albania1988           NA    2,126,000,000                769          Boomers
22    Albania 1988 female 35-54 years           4     284500              1.41  Albania1988           NA    2,126,000,000                769           Silent
23    Albania 1988 female  5-14 years           0     317200              0.00  Albania1988           NA    2,126,000,000                769     Generation X
24    Albania 1988   male  5-14 years           0     345000              0.00  Albania1988           NA    2,126,000,000                769     Generation X
25    Albania 1989   male   75+ years           2      22500              8.89  Albania1989           NA    2,335,124,988                833 G.I. Generation
26    Albania 1989   male 25-34 years          18     283600              6.35  Albania1989           NA    2,335,124,988                833          Boomers
27    Albania 1989   male 35-54 years          15     318400              4.71  Albania1989           NA    2,335,124,988                833           Silent
28    Albania 1989   male 55-74 years           6     142100              4.22  Albania1989           NA    2,335,124,988                833 G.I. Generation
29    Albania 1989   male 15-24 years          12     323500              3.71  Albania1989           NA    2,335,124,988                833     Generation X
30    Albania 1989 female 35-54 years           7     288600              2.43  Albania1989           NA    2,335,124,988                833           Silent
31    Albania 1989 female 15-24 years           5     299900              1.67  Albania1989           NA    2,335,124,988                833     Generation X
32    Albania 1989 female 25-34 years           2     266300              0.75  Albania1989           NA    2,335,124,988                833          Boomers
33    Albania 1989 female 55-74 years           1     149600              0.67  Albania1989           NA    2,335,124,988                833 G.I. Generation
```

```
> #data= Suicide.Rate.Analysis
> names(data)[1]= "country"
> data
   country year    sex         age suicides_no population suicides.100k.pop country.year HDI.for.year gdp_for_year.... gdp_per_capita....       generation
1  Albania 1987   male 15-24 years          21     312900              6.71  Albania1987           NA    2,156,624,900                796     Generation X
2  Albania 1987   male 35-54 years          16     308000              5.19  Albania1987           NA    2,156,624,900                796           Silent
3  Albania 1987 female 15-24 years          14     289700              4.83  Albania1987           NA    2,156,624,900                796     Generation X
4  Albania 1987   male   75+ years           1      21800              4.59  Albania1987           NA    2,156,624,900                796 G.I. Generation
5  Albania 1987   male 25-34 years           9     274300              3.28  Albania1987           NA    2,156,624,900                796          Boomers
6  Albania 1987 female   75+ years           1      35600              2.81  Albania1987           NA    2,156,624,900                796 G.I. Generation
7  Albania 1987 female 35-54 years           6     278800              2.15  Albania1987           NA    2,156,624,900                796           Silent
8  Albania 1987 female 25-34 years           4     257200              1.56  Albania1987           NA    2,156,624,900                796          Boomers
9  Albania 1987   male 55-74 years           1     137500              0.73  Albania1987           NA    2,156,624,900                796 G.I. Generation
10 Albania 1987 female  5-14 years           0     311000              0.00  Albania1987           NA    2,156,624,900                796     Generation X
11 Albania 1987 female 55-74 years           0     144600              0.00  Albania1987           NA    2,156,624,900                796 G.I. Generation
12 Albania 1987   male  5-14 years           0     338200              0.00  Albania1987           NA    2,156,624,900                796     Generation X
13 Albania 1988 female   75+ years           2      36400              5.49  Albania1988           NA    2,126,000,000                769 G.I. Generation
14 Albania 1988   male 15-24 years          17     319200              5.33  Albania1988           NA    2,126,000,000                769     Generation X
15 Albania 1988   male   75+ years           1      22300              4.48  Albania1988           NA    2,126,000,000                769 G.I. Generation
16 Albania 1988   male 35-54 years          14     314100              4.46  Albania1988           NA    2,126,000,000                769           Silent
17 Albania 1988   male 55-74 years           4     140200              2.85  Albania1988           NA    2,126,000,000                769 G.I. Generation
18 Albania 1988 female 15-24 years           8     295600              2.71  Albania1988           NA    2,126,000,000                769     Generation X
19 Albania 1988 female 55-74 years           3     147500              2.03  Albania1988           NA    2,126,000,000                769 G.I. Generation
20 Albania 1988 female 25-34 years           5     262400              1.91  Albania1988           NA    2,126,000,000                769          Boomers
21 Albania 1988   male 25-34 years           5     279900              1.79  Albania1988           NA    2,126,000,000                769          Boomers
22 Albania 1988 female 35-54 years           4     284500              1.41  Albania1988           NA    2,126,000,000                769           Silent
23 Albania 1988 female  5-14 years           0     317200              0.00  Albania1988           NA    2,126,000,000                769     Generation X
24 Albania 1988   male  5-14 years           0     345000              0.00  Albania1988           NA    2,126,000,000                769     Generation X
25 Albania 1989   male   75+ years           2      22500              8.89  Albania1989           NA    2,335,124,988                833 G.I. Generation
26 Albania 1989   male 25-34 years          18     283600              6.35  Albania1989           NA    2,335,124,988                833          Boomers
27 Albania 1989   male 35-54 years          15     318400              4.71  Albania1989           NA    2,335,124,988                833           Silent
28 Albania 1989   male 55-74 years           6     142100              4.22  Albania1989           NA    2,335,124,988                833 G.I. Generation
29 Albania 1989   male 15-24 years          12     323500              3.71  Albania1989           NA    2,335,124,988                833     Generation X
30 Albania 1989 female 35-54 years           7     288600              2.43  Albania1989           NA    2,335,124,988                833           Silent
31 Albania 1989 female 15-24 years           5     299900              1.67  Albania1989           NA    2,335,124,988                833     Generation X
32 Albania 1989 female 25-34 years           2     266300              0.75  Albania1989           NA    2,335,124,988                833          Boomers
33 Albania 1989 female 55-74 years           1     149600              0.67  Albania1989           NA    2,335,124,988                833 G.I. Generation
```

```
>
> # Data summary
> # We are using str() & head() function to inspect and have a brief overwiew of the dataset.
> str(data)
'data.frame':	27820 obs. of  12 variables:
 $ country           : Factor w/ 101 levels "Albania","Antigua and Barbuda",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ year              : int  1987 1987 1987 1987 1987 1987 1987 1987 1987 1987 ...
 $ sex               : Factor w/ 2 levels "female","male": 2 2 1 2 2 1 1 1 2 1 ...
 $ age               : Factor w/ 6 levels "15-24 years",..: 1 3 1 6 2 6 3 2 5 4 ...
 $ suicides_no       : int  21 16 14 1 9 1 6 4 1 0 ...
 $ population        : int  312900 308000 289700 21800 274300 35600 278800 257200 137500 311000 ...
 $ suicides.100k.pop : num  6.71 5.19 4.83 4.59 3.28 2.81 2.15 1.56 0.73 0 ...
 $ country.year      : Factor w/ 2321 levels "Albania1987",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ HDI.for.year      : num  NA NA NA NA NA NA NA NA NA NA ...
 $ gdp_for_year....  : Factor w/ 2321 levels "1,002,219,052,968",..: 727 727 727 727 727 727 727 727 727 727 ...
 $ gdp_per_capita....: int  796 796 796 796 796 796 796 796 796 796 ...
 $ generation        : Factor w/ 6 levels "Boomers","G.I. Generation",..: 3 6 3 2 1 2 6 1 2 3 ...
> summary(data)
       country          year          sex                 age         suicides_no        population       suicides.100k.pop       country.year     HDI.for.year  
 Austria    :  382   Min.   :1985   female:13910   15-24 years:4642   Min.   :    0.0   Min.   :     278   Min.   :  0.00    Albania1987:   12   Min.   :0.483  
 Iceland    :  382   1st Qu.:1995   male  :13910   25-34 years:4642   1st Qu.:    3.0   1st Qu.:   97498   1st Qu.:  0.92    Albania1988:   12   1st Qu.:0.711  
 Mauritius  :  382   Median :2002                  35-54 years:4642   Median :   25.0   Median :  430150   Median :  5.99    Albania1989:   12   Median :0.775  
 Netherlands:  382   Mean   :2001                  5-14 years :4610   Mean   :  242.6   Mean   : 1844794   Mean   : 12.82    Albania1992:   12   Mean   :0.774  
 Argentina  :  372   3rd Qu.:2008                  55-74 years:4642   3rd Qu.:  131.0   3rd Qu.: 1486143   3rd Qu.: 16.62    Albania1993:   12   3rd Qu.:0.851  
 Belgium    :  372   Max.   :2016                  75+ years  :4642   Max.   :22338.0   Max.   :43805214   Max.   :224.97    Albania1994:   12   Max.   :0.944  
 (Other)    :25548                                                                                                           (Other)    :27748   NA's   :19096  
          gdp_for_year.... gdp_per_capita....           generation  
 1,002,219,052,968:   12   Min.   :   251     Boomers        :4990  
 1,011,797,457,139:   12   1st Qu.:  3447     G.I. Generation:2744  
 1,016,418,229    :   12   Median :  9372     Generation X   :6408  
 1,018,847,043,277:   12   Mean   : 16866     Generation Z   :1470  
 1,022,191,296    :   12   3rd Qu.: 24874     Millenials     :5844  
 1,023,196,003,075:   12   Max.   :126352     Silent         :6364  
 (Other)          :27748                                          
> 
```

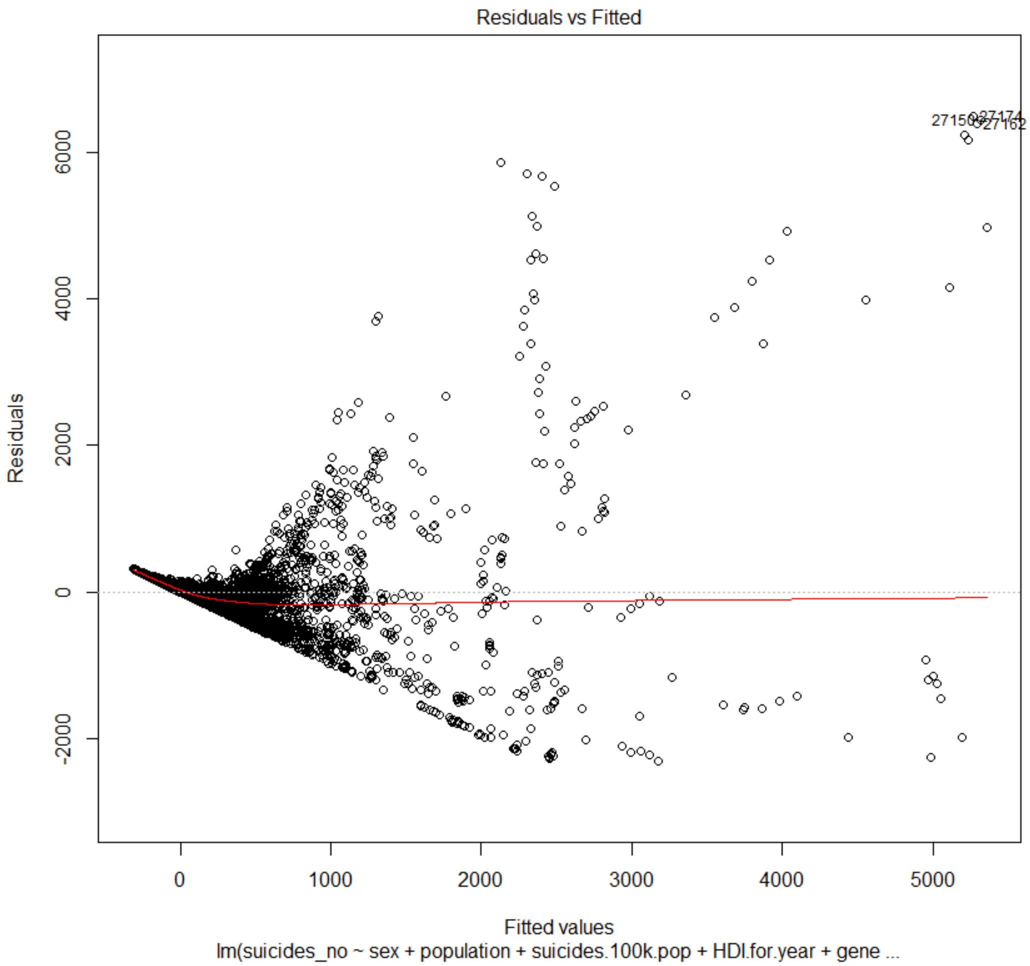
RStudio

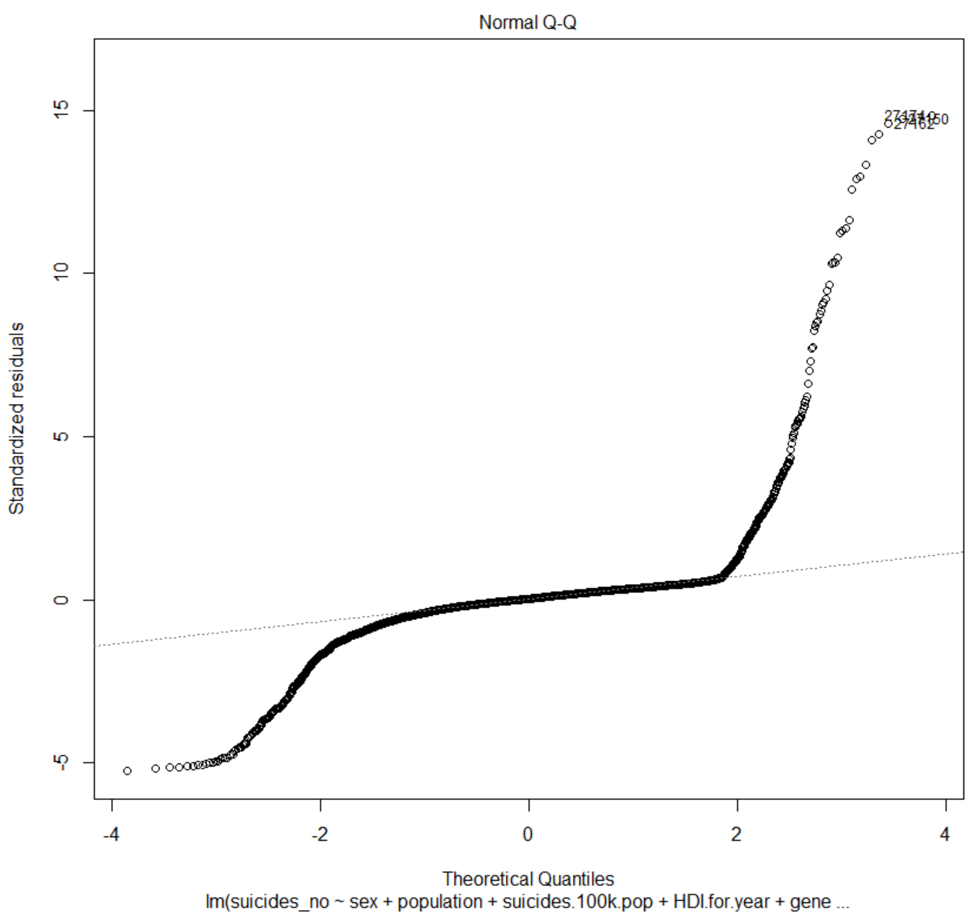
```
> ##############################################################################################
>
> #Multivariate Analysis
>
> #correlation between pairs of variables
> data[,sapply(data, is.numeric)] %>%  cor(use = "complete.obs")
                          year suicides_no  population suicides.100k.pop HDI.for.year gdp_per_capita....
year                1.00000000 -0.02122209 -0.01193929      -0.071105083   0.37820770        0.307208408
suicides_no        -0.02122209  1.00000000  0.69895204       0.240515931   0.15569802        0.109844385
population         -0.01193929  0.69895204  1.00000000      -0.018162496   0.10967510        0.079469017
suicides.100k.pop  -0.07110508  0.24051593 -0.01816250       1.000000000   0.08379932        0.007847636
HDI.for.year        0.37820770  0.15569802  0.10967510       0.083799321   1.00000000        0.774158774
gdp_per_capita....  0.30720841  0.10984438  0.07946902       0.007847636   0.77415877        1.000000000
> fit1 <- lm(suicides_no~sex+population+suicides.100k.pop+country+HDI.for.year+generation, data=data)
> summary(fit1)

Call:
lm(formula = suicides_no ~ sex + population + suicides.100k.pop +
    country + HDI.for.year + generation, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-2960.2   -89.1     6.9    94.9  5792.9

Coefficients:
                                   Estimate Std. Error t value Pr(>|t|)
(Intercept)                      -2.899e+02  8.477e+01  -3.420 0.000629 ***
sexmale                           1.341e+02  9.820e+00  13.657  < 2e-16 ***
population                        1.442e-04  2.457e-06  58.690  < 2e-16 ***
suicides.100k.pop                 6.231e+00  3.574e-01  17.434  < 2e-16 ***
countryAntigua and Barbuda        3.311e+01  4.774e+01   0.694 0.487955
countryArgentina                 -2.311e+02  3.920e+01  -5.894 3.90e-09 ***
countryArmenia                    7.138e+00  4.835e+01   0.148 0.882624
countryAustralia                 -1.211e+02  5.606e+01  -2.160 0.030764 *
countryAustria                   -1.031e+02  4.963e+01  -2.078 0.037739 *
countryAzerbaijan                -2.881e+01  8.534e+01  -0.338 0.735664
countryBahamas                    1.738e+01  5.503e+01   0.316 0.752173
countryBahrain                    2.825e+00  5.071e+01   0.056 0.955566
countryBarbados                   8.234e+00  4.749e+01   0.173 0.862362
countryBelarus                   -4.003e+01  5.923e+01  -0.676 0.499172
countryBelgium                   -7.981e+01  5.024e+01  -1.588 0.112230
```

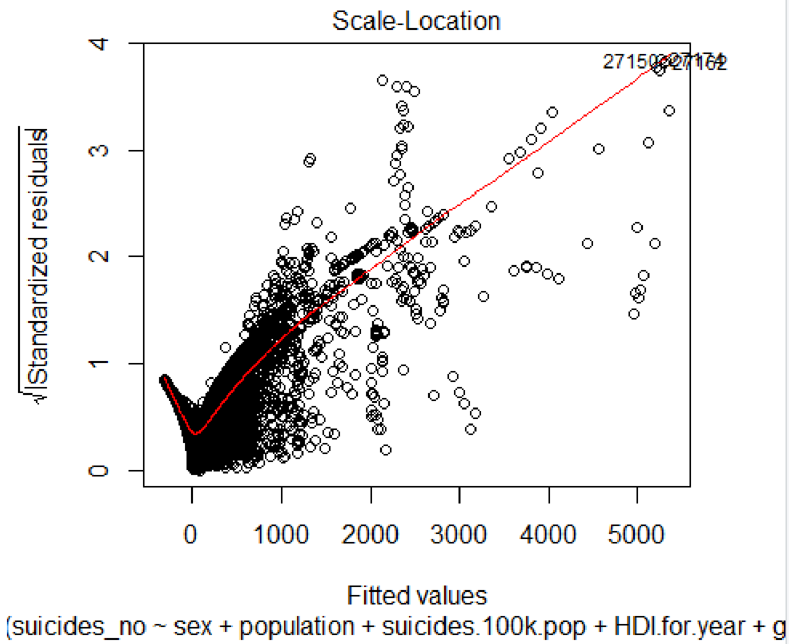
```
> # country has most p-value among variables so we can remove it
> fit2 <- lm(suicides_no~sex+population+suicides.100k.pop+HDI.for.year+generation, data=data)
> anova(fit1, fit2)
Analysis of Variance Table

Model 1: suicides_no ~ sex + population + suicides.100k.pop + country +
    HDI.for.year + generation
Model 2: suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year +
    generation
  Res.Df        RSS  Df  Sum of Sq      F    Pr(>F)
1   8625 1358781213
2   8714 1691320567 -89 -332539354 23.717 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> #The results of anova shows that p-value>0.05 so we can accept null hypothesis(country is not important)
> coefficients(fit2)
           (Intercept)                sexmale             population      suicides.100k.pop           HDI.for.year generationG.I. Generation
         -5.041471e+02           1.159584e+02           1.190084e-04           7.422114e+00           5.233851e+02          -1.982477e+01
  generationGeneration X   generationGeneration Z      generationMillenials       generationSilent
         -6.248608e+01          -1.604875e+02          -1.053919e+02          -2.491644e+01
> library(GGally)
> ggpairs(data=data, title="Data")
```

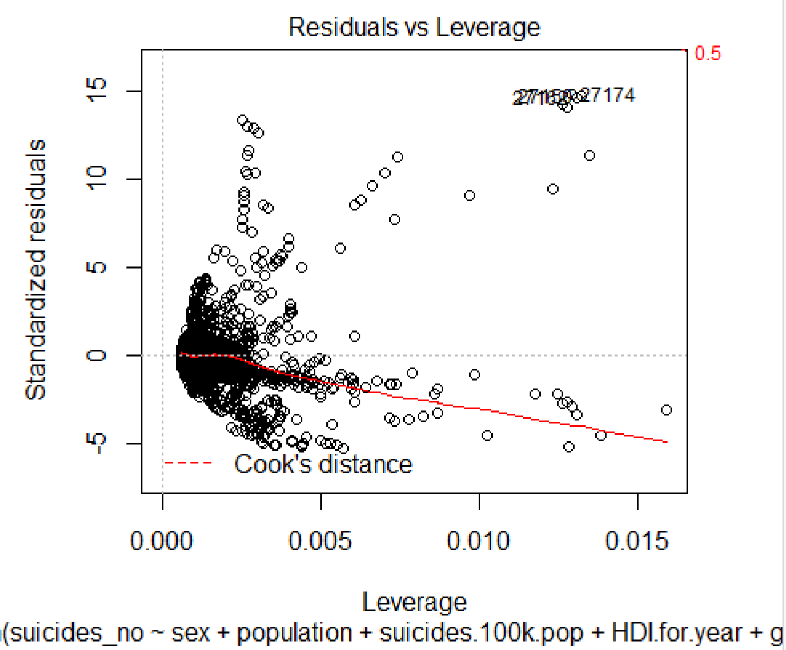
```
> anova(fit2)
Analysis of Variance Table

Response: suicides_no
                   Df     Sum Sq    Mean Sq  F value    Pr(>F)    
sex                 1   95724216   95724216  493.189 < 2.2e-16 ***
population          1 1909009468 1909009468 9835.574 < 2.2e-16 ***
suicides.100k.pop   1  163575485  163575485  842.771 < 2.2e-16 ***
HDI.for.year        1   14193068   14193068   73.125 < 2.2e-16 ***
generation          5   17598770    3519754   18.134 < 2.2e-16 ***
Residuals        8714 1691320567     194092                       
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> vcov(fit2)
                              (Intercept)       sexmale    population suicides.100k.pop  HDI.for.year generationG.I. Generation generationGeneration X
(Intercept)                  1.784016e+03 -6.477050e+01  1.895694e-06      1.485671e+00 -2.112734e+03             -2.608925e+02          -1.542904e+02
sexmale                     -6.477050e+01  1.094504e+02  1.213477e-08     -1.488957e+00  3.981020e+01              1.723180e+01          -4.350319e+00
population                   1.895694e-06  1.213477e-08  1.505555e-12      6.883728e-09 -7.203180e-06              1.770986e-06          -4.274332e-08
suicides.100k.pop            1.485671e+00 -1.488957e+00  6.883728e-09      1.084081e-01 -2.935028e+00             -1.245082e+00           3.164127e-01
HDI.for.year                -2.112734e+03  3.981020e+01 -7.203180e-06     -2.935028e+00  2.791859e+03              1.506770e+02           3.734515e+00
generationG.I. Generation   -2.608925e+02  1.723180e+01  1.770986e-06     -1.245082e+00  1.506770e+02              4.676345e+02           1.461533e+02
generationGeneration X      -1.542904e+02 -4.350319e+00 -4.274332e-08      3.164127e-01  3.734515e+00              1.461533e+02           2.431489e+02
generationGeneration Z      -6.317742e+01 -2.092217e+01  1.427175e-06      1.530176e+00 -1.301466e+02              1.294948e+02           1.533426e+02
generationMillenials        -9.946374e+01 -1.099794e+01  1.139179e-06      8.063592e-01 -7.556768e+01              1.390971e+02           1.513918e+02
generationSilent            -1.405598e+02  4.516185e+00  1.633340e-06     -3.202792e-01 -1.351606e+01              1.543343e+02           1.482331e+02
                            generationGeneration Z generationMillenials generationSilent
(Intercept)                          -6.317742e+01        -9.946374e+01    -1.405598e+02
sexmale                              -2.092217e+01        -1.099794e+01     4.516185e+00
population                            1.427175e-06         1.139179e-06     1.633340e-06
suicides.100k.pop                     1.530176e+00         8.063592e-01    -3.202792e-01
HDI.for.year                         -1.301466e+02        -7.556768e+01    -1.351606e+01
generationG.I. Generation             1.294948e+02         1.390971e+02     1.543343e+02
generationGeneration X                1.533426e+02         1.513918e+02     1.482331e+02
generationGeneration Z                4.160249e+02         1.631435e+02     1.466882e+02
generationMillenials                  1.631435e+02         2.506438e+02     1.483778e+02
generationSilent                      1.466882e+02         1.483778e+02     2.598525e+02
> cov2cor(vcov(fit2))
                            (Intercept)       sexmale    population suicides.100k.pop HDI.for.year generationG.I. Generation generationGeneration X
(Intercept)                  1.00000000 -0.1465781667  0.0365780501        0.10682984 -0.946670239               -0.28563341           -0.234262782
sexmale                     -0.14657817  1.0000000000  0.0009453109       -0.43225763  0.072017667                0.07616735           -0.026667136
population                   0.03657805  0.0009453109  1.0000000000        0.01703902 -0.111103878                0.06674419           -0.002234003
suicides.100k.pop            0.10682984 -0.4322576314  0.0170390213        1.00000000 -0.168707613               -0.17486938            0.061629259
HDI.for.year                -0.94667024  0.0720176672 -0.1111038775       -0.16870761  1.000000000                0.13187026            0.004532641
generationG.I. Generation   -0.28563341  0.0761673459  0.0667441857       -0.17486938  0.131870261                1.00000000            0.433430267
generationGeneration X      -0.23426278 -0.0266671364 -0.0022340034        0.06162926  0.004532641                0.43343027            1.000000000
generationGeneration Z      -0.07333358 -0.0980479446  0.0570255175        0.22785102 -0.120760910                0.29358876            0.482133315
generationMillenials        -0.14874315 -0.0664009692  0.0586428806        0.15469253 -0.090336010                0.40629029            0.613250433
generationSilent            -0.20644216  0.0267793296  0.0825781024       -0.06034405 -0.015868647                0.44273698            0.589719495
                            generationGeneration Z generationMillenials generationSilent
(Intercept)                            -0.07333358          -0.14874315      -0.20644216
sexmale                                -0.09804794          -0.06640097       0.02677933
population                              0.05702552           0.05864288       0.08257810
suicides.100k.pop                       0.22785102           0.15469253      -0.06034405
HDI.for.year                           -0.12076091          -0.09033601      -0.01586865
generationG.I. Generation               0.29358876           0.40629029       0.44273698
generationGeneration X                  0.48213332           0.61325043       0.58971950
generationGeneration Z                  1.00000000           0.50522131       0.44614084
```

```
> temp <- influence.measures(fit2)
> temp
Influence measures of
	 lm(formula = suicides_no ~ sex + population + suicides.100k.pop +      HDI.for.year + generation, data = data) :

      dfb.1_  dfb.sxml  dfb.pplt  dfb.s.10  dfb.HDI.  dfb.gG.G  dfb.gnGX  dfb.gnGZ  dfb.gnrM  dfb.gnrS     dffit cov.r   cook.d      hat inf
37  8.15e-04  4.76e-04 -1.23e-04 -3.34e-04 -5.74e-04 -5.63e-04 -7.78e-04 -6.03e-04 -7.56e-04 -7.02e-04  1.27e-03     1 1.60e-07 0.001353    
38  2.96e-03  2.48e-03 -7.46e-04 -1.46e-03 -3.17e-03 -1.96e-04  2.46e-03 -9.20e-05 -5.43e-05  8.98e-05  6.12e-03     1 3.74e-06 0.001008    
39  2.11e-03  1.95e-03 -1.40e-04 -1.60e-03 -2.25e-03 -1.83e-05 -1.36e-04 -1.68e-04 -1.00e-04  2.12e-03  4.55e-03     1 2.07e-06 0.001207    
40  1.43e-03  8.48e-04 -2.33e-04 -6.16e-04 -1.00e-03 -9.85e-04 -1.37e-03 -1.07e-03 -1.33e-03 -1.23e-03  2.23e-03     1 4.99e-07 0.001365    
41  8.65e-03 -4.85e-03 -1.76e-03  8.86e-04 -8.33e-03 -1.24e-03  6.16e-03  7.75e-04  5.45e-04 -5.40e-05  1.42e-02     1 2.03e-05 0.000955    
42  9.01e-03 -3.10e-03 -1.28e-03 -3.09e-04 -6.05e-03 -5.98e-03 -7.37e-03 -5.19e-03 -6.89e-03 -7.00e-03  1.19e-02     1 1.41e-05 0.001263    
43  8.86e-03 -2.96e-03 -1.21e-03 -5.21e-04 -5.93e-03 -5.86e-03 -7.28e-03 -5.16e-03 -6.82e-03 -6.88e-03  1.17e-02     1 1.37e-05 0.001263    
44  8.25e-03 -3.42e-03 -4.72e-04 -1.62e-03 -7.91e-03 -6.90e-04 -2.24e-04  2.63e-04  2.23e-04  6.67e-03  1.40e-02     1 1.95e-05 0.001073    
45  1.23e-02 -6.39e-03 -8.85e-04  2.01e-03 -1.21e-02 -1.77e-03 -5.34e-05  1.35e-03  8.64e-03  1.21e-05  1.86e-02     1 3.46e-05 0.001062    
46  5.87e-03 -2.32e-03 -5.54e-04 -4.87e-03 -5.21e-03  1.43e-02 -3.91e-04 -6.49e-04 -4.11e-04  3.62e-04  2.01e-02     1 4.05e-05 0.001937    
47  5.42e-03  3.77e-03 -4.37e-04 -2.03e-03 -5.85e-03 -3.88e-04 -2.18e-04  1.45e-05  4.08e-03  1.99e-04  9.93e-03     1 9.87e-06 0.001094    
48  1.85e-03  3.70e-03 -2.43e-04 -4.14e-03 -1.91e-03  6.55e-03 -2.94e-04 -7.42e-04 -4.90e-04  2.82e-04  9.41e-03     1 8.85e-06 0.002261    
49  2.27e-03  1.72e-03 -5.79e-04 -6.89e-04 -2.45e-03 -2.21e-04  1.92e-03  2.29e-05  2.12e-05  4.07e-05  4.58e-03     1 2.10e-06 0.000960    
50  1.77e-03  1.63e-03 -1.20e-04 -1.28e-03 -1.89e-03 -2.64e-05 -1.11e-04 -1.27e-04 -7.45e-05  1.80e-03  3.83e-03     1 1.47e-06 0.001175    
51  1.18e-03  2.20e-03 -1.54e-04 -2.27e-03 -1.24e-03  4.13e-03 -1.64e-04 -3.91e-04 -2.58e-04  1.56e-04  5.85e-03     1 3.42e-06 0.002134    
52  1.11e-03  6.54e-04 -1.87e-04 -4.51e-04 -7.79e-04 -7.82e-04 -1.08e-03 -8.35e-04 -1.05e-03 -9.75e-04  1.75e-03     1 3.05e-07 0.001338    
53  8.28e-03 -4.85e-03 -1.73e-03  1.21e-03 -7.98e-03 -1.25e-03  6.00e-03  8.17e-04  5.71e-04 -7.83e-05  1.38e-02     1 1.89e-05 0.000947    
54  8.23e-04  4.97e-04 -1.27e-04 -3.62e-04 -5.74e-04 -5.74e-04 -8.00e-04 -6.25e-04 -7.80e-04 -7.20e-04  1.30e-03     1 1.69e-07 0.001351    
55  8.53e-03 -3.07e-03 -1.24e-03 -3.60e-05 -5.71e-03 -5.75e-03 -7.02e-03 -4.91e-03 -6.55e-03 -6.70e-03  1.13e-02     1 1.27e-05 0.001249    
56  8.27e-03 -2.93e-03 -1.14e-03 -1.64e-04 -5.53e-03 -5.55e-03 -6.82e-03 -4.79e-03 -6.37e-03 -6.50e-03  1.09e-02     1 1.20e-05 0.001248    
57  7.90e-03 -3.39e-03 -4.67e-04 -1.42e-03 -7.57e-03 -6.81e-04 -2.06e-04  2.78e-04  2.31e-04  6.47e-03  1.35e-02     1 1.81e-05 0.001057    
58  1.20e-02 -6.34e-03 -8.88e-04  2.07e-03 -1.18e-02 -1.74e-03 -4.43e-05  1.34e-03  8.54e-03  1.79e-06  1.83e-02     1 3.33e-05 0.001048    
59  5.20e-03  3.67e-03 -4.34e-04 -1.94e-03 -5.62e-03 -3.75e-04 -2.09e-04  1.51e-05  3.98e-03  1.90e-04  9.63e-03     1 9.27e-06 0.001077    
60  5.66e-03 -2.29e-03 -5.67e-04 -4.86e-03 -5.00e-03  1.42e-02 -3.87e-04 -6.64e-04 -4.22e-04  3.57e-04  1.99e-02     1 3.97e-05 0.001929    
61  4.86e-04  8.74e-04 -6.74e-05 -8.13e-04 -5.21e-04  1.74e-03 -6.00e-05 -1.35e-04 -8.86e-05  5.63e-05  2.44e-03     1 5.93e-07 0.002022    
62  1.15e-03  1.05e-03 -8.37e-05 -7.46e-04 -1.24e-03 -3.20e-05 -6.66e-05 -6.40e-05 -3.60e-05  1.21e-03  2.52e-03     1 6.33e-07 0.001116    
63  4.47e-03 -2.42e-03 -4.63e-04 -2.59e-03 -4.00e-03  1.13e-02 -2.27e-04 -2.55e-04 -1.53e-04  2.01e-04  1.57e-02     1 2.48e-05 0.001847    
64  1.26e-03  7.90e-04 -2.24e-04 -5.88e-04 -8.60e-04 -8.97e-04 -1.26e-03 -9.92e-04 -1.24e-03 -1.14e-03  2.03e-03     1 4.14e-07 0.001334    
```

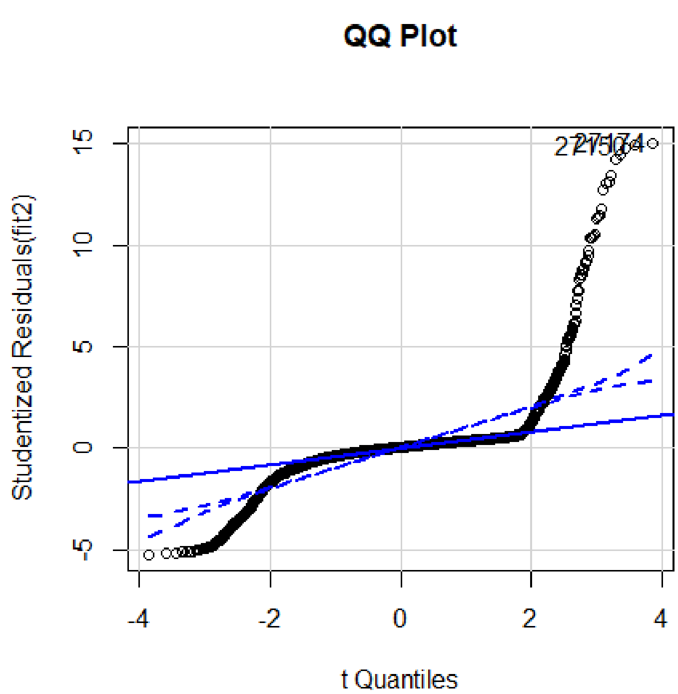

Residuals vs Fitted
Residuals
Fitted values
lm(suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year + gene ...
27150
27174
27162
6000
4000
2000
0
-2000
0
1000
2000
3000
4000
5000


Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year + gene ...
15
10
5
0
-5
-4
-2
0
2
4

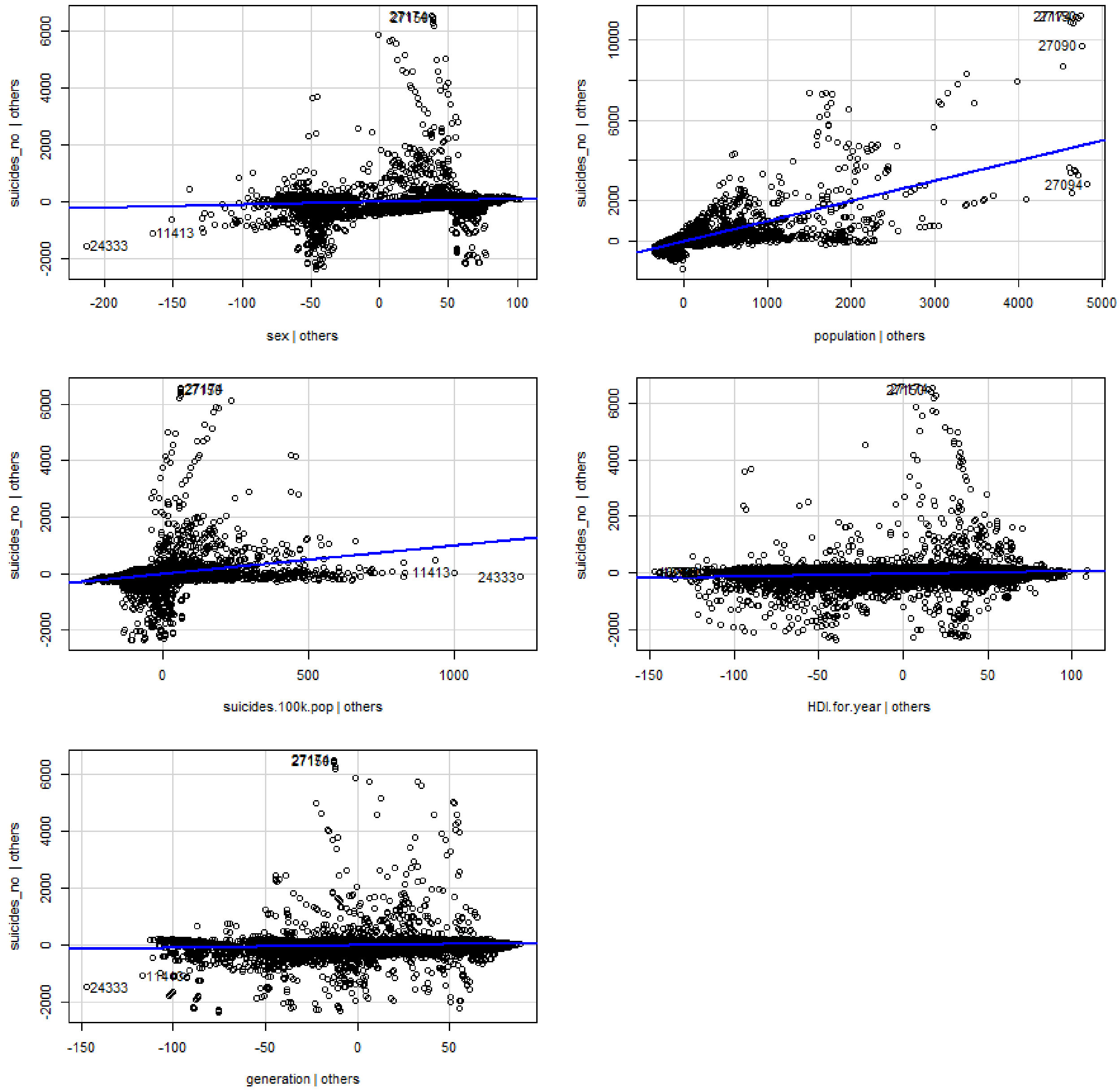

Scale-Location
√|Standardized residuals|
Fitted values
(suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year + g
0
1
2
3
4
0
1000
2000
3000
4000
5000

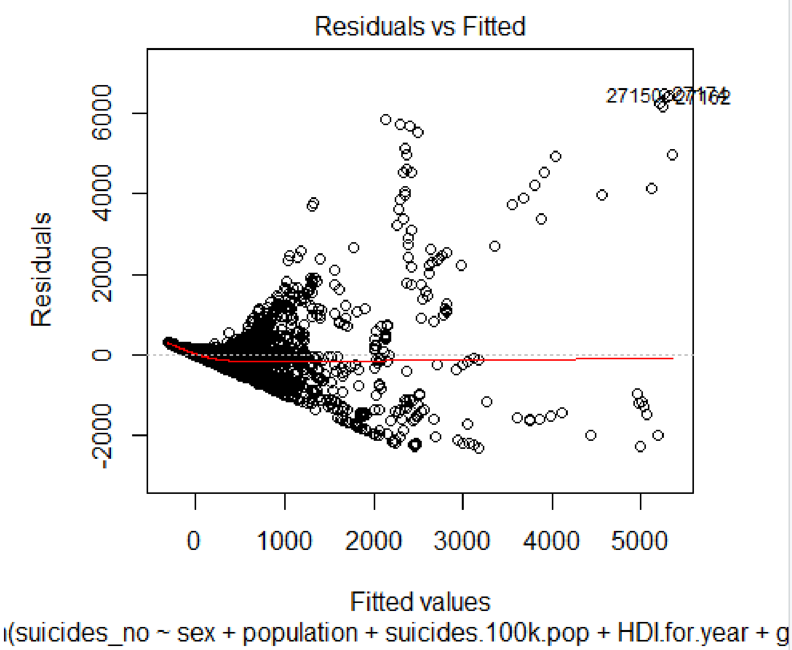

Residuals vs Leverage
Standardized residuals
Leverage
(suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year + g
Cook's distance
0.5
27174
-5
0
5
10
15
0.000
0.005
0.010
0.015

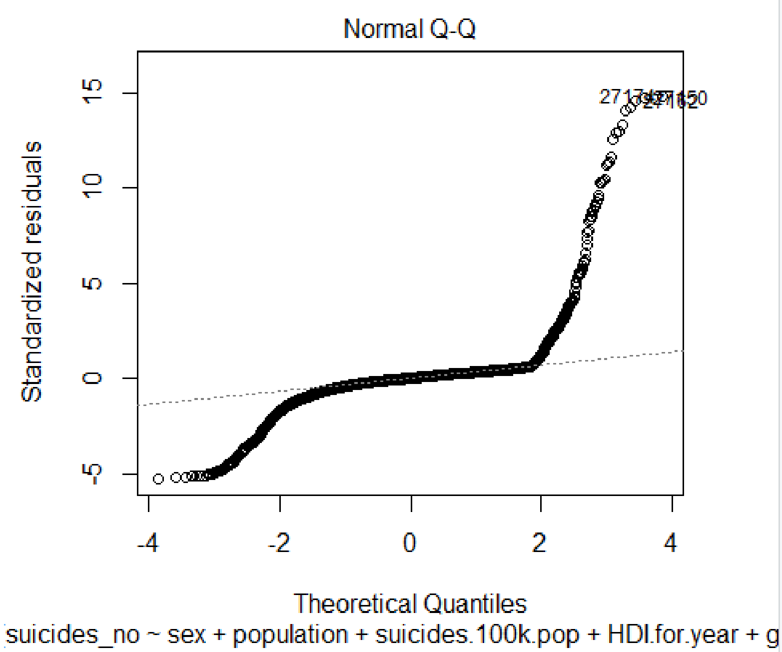
```
> # Assessing Outliers
> library(car)
> outlierTest(fit2)
      rstudent unadjusted p-value Bonferroni p
27174 15.02399         2.1797e-50   1.9015e-46
27150 14.91131         1.1366e-49   9.9161e-46
27162 14.77019         8.8470e-49   7.7182e-45
27199 14.42745         1.1962e-46   1.0436e-42
27187 14.23244         1.8586e-45   1.6214e-41
13545 13.45473         7.3760e-41   6.4348e-37
13605 13.09217         8.4708e-39   7.3899e-35
13606 13.01356         2.3309e-38   2.0335e-34
13547 12.69927         1.2589e-36   1.0983e-32
13665 11.74345         1.3198e-31   1.1514e-27
> qqPlot(fit2, main="QQ Plot")
[1] 27150 27174
```



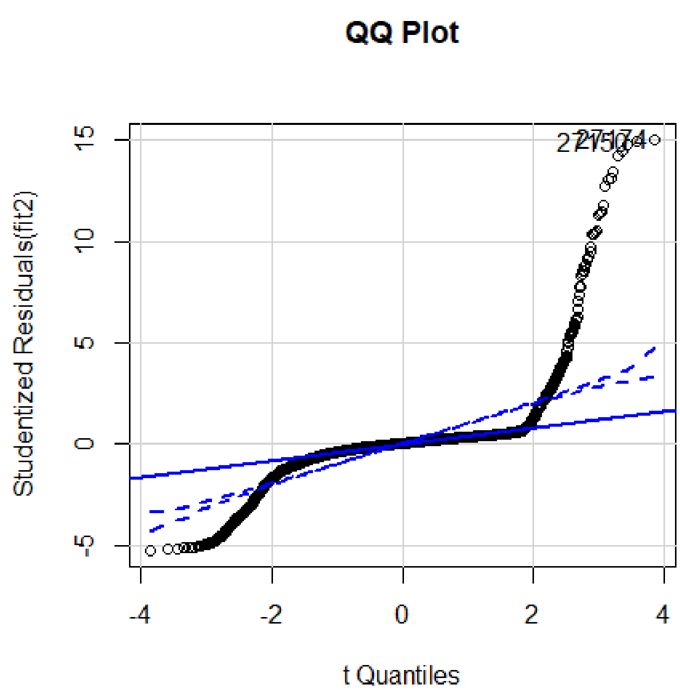
Leverage Plots



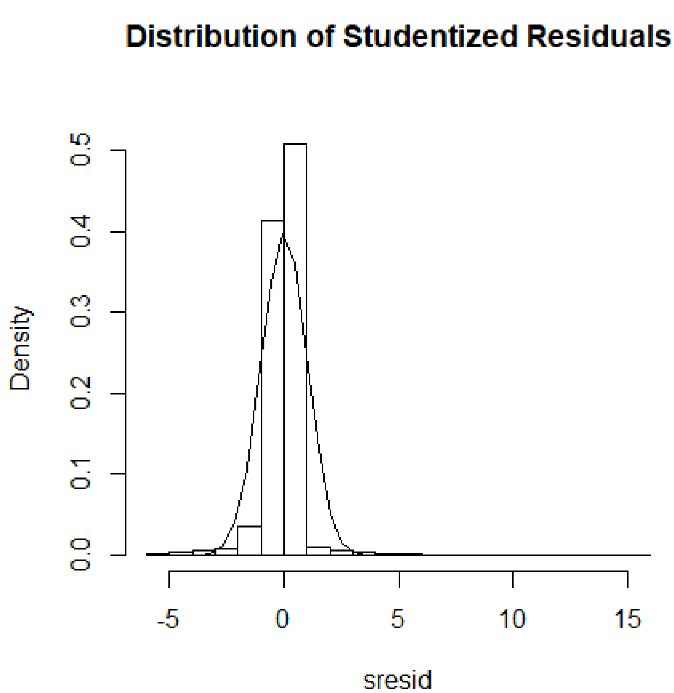

Residuals vs Fitted
Residuals
6000
4000
2000
0
-2000
0
1000
2000
3000
4000
5000
Fitted values
ı(suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year + g


Normal Q-Q
Standardized residuals
15
10
5
0
-5
-4
-2
0
2
4
Theoretical Quantiles
suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year + g

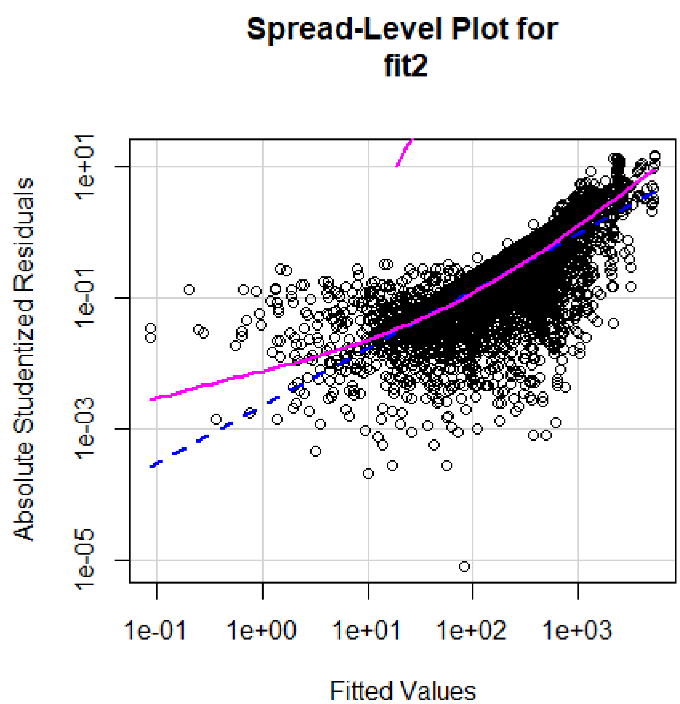
```
> # Assessing Outliers
> library(car)
> outlierTest(fit2)
      rstudent unadjusted p-value Bonferroni p
27174 15.02399         2.1797e-50   1.9015e-46
27150 14.91131         1.1366e-49   9.9161e-46
27162 14.77019         8.8470e-49   7.7182e-45
27199 14.42745         1.1962e-46   1.0436e-42
27187 14.23244         1.8586e-45   1.6214e-41
13545 13.45473         7.3760e-41   6.4348e-37
13605 13.09217         8.4708e-39   7.3899e-35
13606 13.01356         2.3309e-38   2.0335e-34
13547 12.69927         1.2589e-36   1.0983e-32
13665 11.74345         1.3198e-31   1.1514e-27
> qqPlot(fit2, main="QQ Plot")
[1] 27150 27174
> leveragePlots(fit2) # leverage plots
> plot(fit2)
Hit <Return> to see next plot:
Hit <Return> to see next plot:
Hit <Return> to see next plot: # Influence Plot
Hit <Return> to see next plot: library(mvinfluence)
> influencePlot(fit2, id.method="identify", main="Influence Plot", sub="Circle size is proportial to Cook's Distance" )
        StudRes        Hat      CookD
24333 -3.063060 0.01590841 0.01515251
27094 -4.551611 0.01384053 0.02901044
27150 14.911314 0.01319190 0.28987617
27174 15.023990 0.01290747 0.28773731
> # Normality of Residuals
> # qq plot for studentized resid
> qqPlot(fit2, main="QQ Plot")
[1] 27150 27174
> # distribution of studentized residuals
> library(MASS)
```

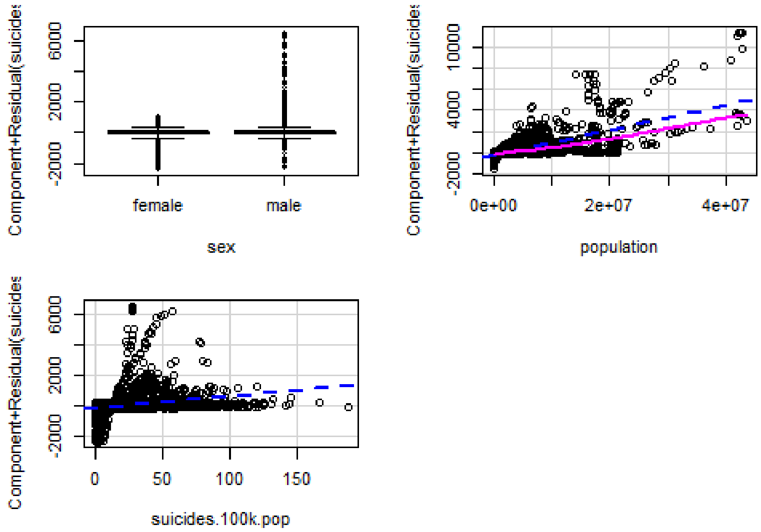


```
> library(MASS)
> # distribution of studentized residuals
> library(MASS)
> sresid <- studres(fit2)
> hist(sresid, freq=FALSE,
+      main="Distribution of Studentized Residuals")
> xfit<-seq(min(sresid),max(sresid),length=40)
> yfit<-dnorm(xfit)
> lines(xfit, yfit)
> #Non-constant Error Variance
> # Evaluate homoscedasticity
> # non-constant error variance test
> ncvTest(fit2)
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 84068.74, Df = 1, p = < 2.22e-16
> # plot studentized residuals vs. fitted values
> spreadLevelPlot(fit2)

Suggested power transformation:  0.1210214
> #Multi-collinearity
> # Evaluate Collinearity
> vif(fit2) # variance inflation factors
                     GVIF Df GVIF^(1/(2*Df))
sex              1.229885  1        1.109002
population       1.029200  1        1.014495
suicides.100k.pop 1.430225 1        1.195920
HDI.for.year     1.083056  1        1.040700
generation       1.256773  5        1.023118
>
```

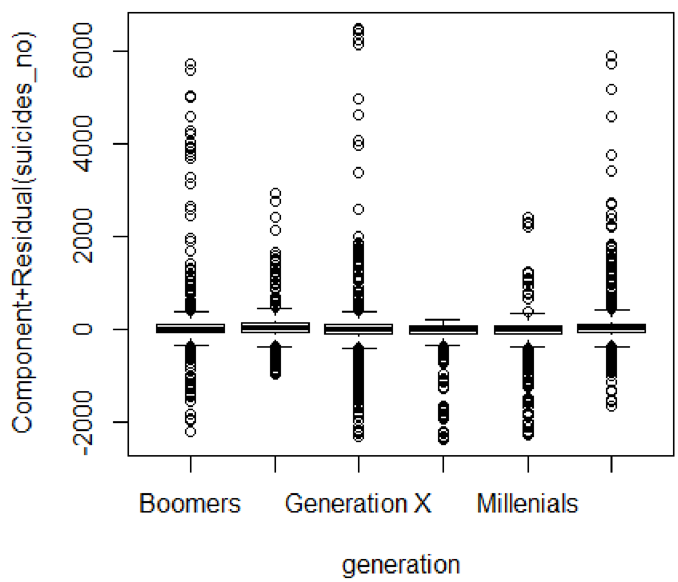


```
> # plot studentized residuals vs. fitted values
> spreadLevelPlot(fit2)

Suggested power transformation:  0.1210214
> #Multi-collinearity
> # Evaluate Collinearity
> vif(fit2) # variance inflation factors
                    GVIF Df GVIF^(1/(2*Df))
sex             1.229885  1        1.109002
population      1.029200  1        1.014495
suicides.100k.pop 1.430225  1      1.195920
HDI.for.year    1.083056  1        1.040700
generation      1.256773  5        1.023118
> sqrt(vif(fit2)) > 2 # problem?
                   GVIF    Df GVIF^(1/(2*Df))
sex               FALSE FALSE           FALSE
population        FALSE FALSE           FALSE
suicides.100k.pop FALSE FALSE           FALSE
HDI.for.year      FALSE FALSE           FALSE
generation        FALSE  TRUE           FALSE
> #Nonlinearity
> # component + residual plot
> crPlots(fit2)
```

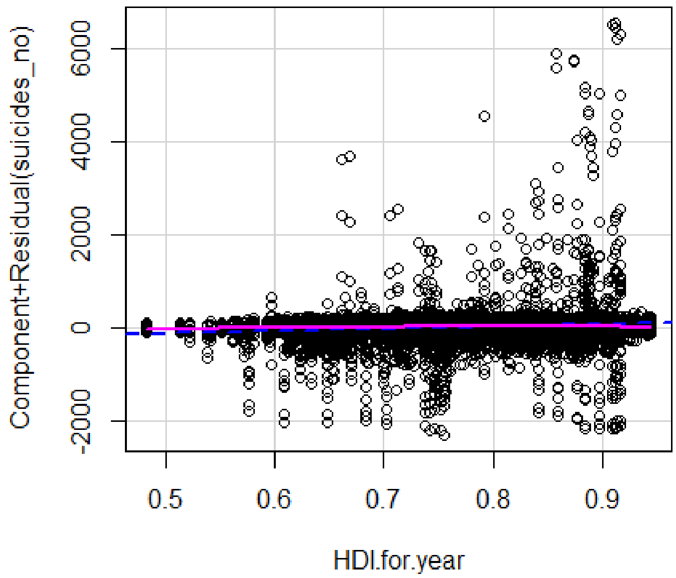




Component+Residual(suicides
-2000
2000
6000
female
male
sex
10000
4000
0e+00
2e+07
4e+07
population
0
50
100
150
suicides.100k.pop



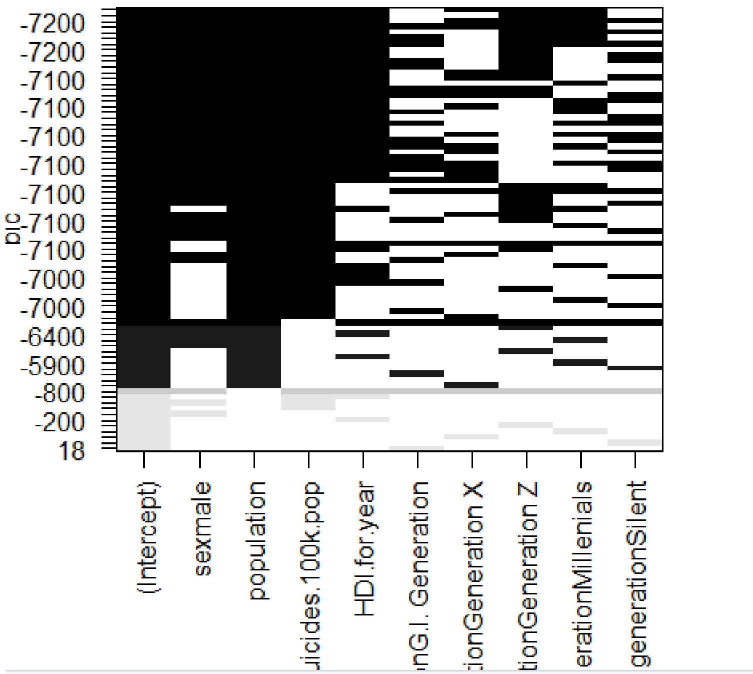
Component+Residual(suicides_no)
-2000
0
2000
4000
6000
Boomers
Generation X
Millenials
generation



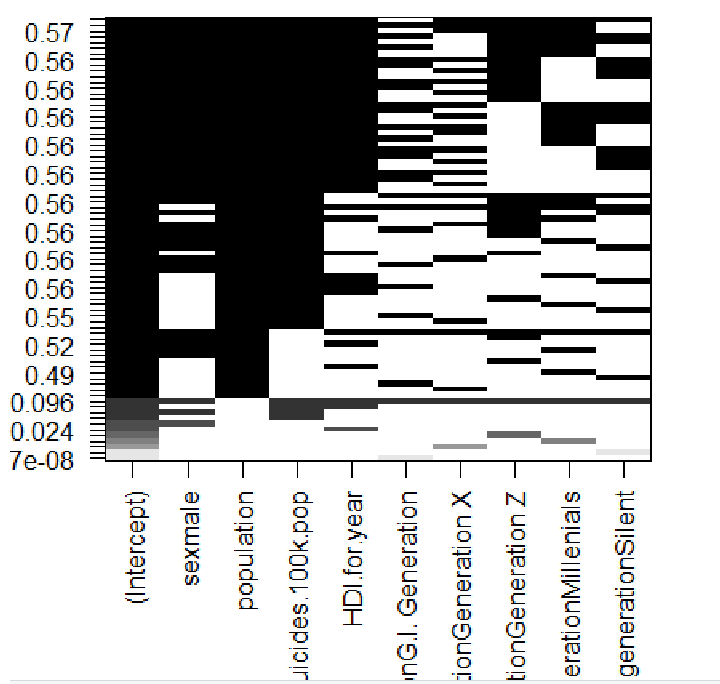

```
> #Nonlinearity
> # component + residual plot
> crPlots(fit2)
> #Non-independence of Errors
> #Test Autocorrelated Errors
> durbinWatsonTest(fit2)
 lag Autocorrelation D-W Statistic p-value
   1       0.6153124      0.769375       0
 Alternative hypothesis: rho != 0
> # Global test of model assumptions
> library(gvlma)
> gvmodel <- gvlma(fit2)
> summary(gvmodel)

Call:
lm(formula = suicides_no ~ sex + population + suicides.100k.pop +
    HDI.for.year + generation, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-2309.4   -94.5    12.3   111.6  6492.9

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)               -5.041e+02  4.224e+01 -11.936  < 2e-16 ***
sexmale                    1.160e+02  1.046e+01  11.084  < 2e-16 ***
population                 1.190e-04  1.227e-06  96.991  < 2e-16 ***
suicides.100k.pop          7.422e+00  3.293e-01  22.542  < 2e-16 ***
HDI.for.year               5.234e+02  5.284e+01   9.905  < 2e-16 ***
generationG.I. Generation -1.982e+01  2.162e+01  -0.917    0.359
generationGeneration X    -6.249e+01  1.559e+01  -4.007 6.19e-05 ***
generationGeneration Z    -1.605e+02  2.040e+01  -7.868 4.02e-15 ***
generationMillenials      -1.054e+02  1.583e+01  -6.657 2.96e-11 ***
generationSilent          -2.492e+01  1.612e+01  -1.546    0.122
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 440.6 on 8714 degrees of freedom
  (19096 observations deleted due to missingness)
Multiple R-squared:  0.5654,    Adjusted R-squared:  0.5649
F-statistic:  1259 on 9 and 8714 DF,  p-value: < 2.2e-16


ASSESSMENT OF THE LINEAR MODEL ASSUMPTIONS
USING THE GLOBAL TEST ON 4 DEGREES-OF-FREEDOM:
Level of Significance =  0.05

Call:
 gvlma(x = fit2)

                  Value p-value                   Decision
Global Stat     1517082       0 Assumptions NOT satisfied!
Skewness          33514       0 Assumptions NOT satisfied!
Kurtosis        1480357       0 Assumptions NOT satisfied!
```

```
> fit3 <- fit2
> fit4 <- lm(suicides_no~sex+population+suicides.100k.pop+HDI.for.year+generation, data=data)
> fit4

Call:
lm(formula = suicides_no ~ sex + population + suicides.100k.pop +
    HDI.for.year + generation, data = data)

Coefficients:
              (Intercept)                    sexmale                 population          suicides.100k.pop              HDI.for.year  generationG.I. Generation
               -5.041e+02                  1.160e+02                  1.190e-04                  7.422e+00                 5.234e+02                 -1.982e+01
   generationGeneration X     generationGeneration Z       generationMillenials           generationSilent
               -6.249e+01                 -1.605e+02                 -1.054e+02                 -2.492e+01

> # compare models
> anova(fit3, fit4)
Analysis of Variance Table

Model 1: suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year +
    generation
Model 2: suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year +
    generation
  Res.Df        RSS Df Sum of Sq F Pr(>F)
1   8714 1691320567
2   8714 1691320567  0         0
> step <- stepAIC(fit2, direction="both")
Start:  AIC=106234.2
suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year +
    generation

                    Df  Sum of Sq        RSS    AIC
<none>                           1691320567 106234
- generation         5   17598770 1708919336 106315
- HDI.for.year       1   19043976 1710364542 106330
- sex                1   23844925 1715165492 106354
- suicides.100k.pop  1   98628429 1789948996 106727
- population         1 1825857113 3517177680 112619
> step$anova # display results
Stepwise Model Path
Analysis of Deviance Table

Initial Model:
suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year +
    generation

Final Model:
suicides_no ~ sex + population + suicides.100k.pop + HDI.for.year +
    generation


  Step Df Deviance Resid. Df Resid. Dev      AIC
1                       8714 1691320567 106234.2
```

```
> library(leaps)
> leaps<-regsubsets(suicides_no~sex+population+suicides.100k.pop+HDI.for.year+generation, data=data, nbest=10)
> # view results
> summary(leaps)
Subset selection object
Call: regsubsets.formula(suicides_no ~ sex + population + suicides.100k.pop +
    HDI.for.year + generation, data = data, nbest = 10)
9 Variables  (and intercept)
                          Forced in Forced out
sexmale                       FALSE      FALSE
population                    FALSE      FALSE
suicides.100k.pop             FALSE      FALSE
HDI.for.year                  FALSE      FALSE
generationG.I. Generation     FALSE      FALSE
generationGeneration X        FALSE      FALSE
generationGeneration Z        FALSE      FALSE
generationMillenials          FALSE      FALSE
generationSilent              FALSE      FALSE
10 subsets of each size up to 8
Selection Algorithm: exhaustive
           sexmale population suicides.100k.pop HDI.for.year generationG.I. Generation generationGeneration X generationGeneration Z generationMillenials
1  ( 1 )  " "     "*"        " "               " "          " "                       " "                    " "                    " "
1  ( 2 )  " "     " "        "*"               " "          " "                       " "                    " "                    " "
1  ( 3 )  "*"     " "        " "               " "          " "                       " "                    " "                    " "
1  ( 4 )  " "     " "        " "               "*"          " "                       " "                    " "                    " "
1  ( 5 )  " "     " "        " "               " "          " "                       " "                    "*"                    " "
1  ( 6 )  " "     " "        " "               " "          " "                       " "                    " "                    "*"
1  ( 7 )  " "     " "        " "               " "          " "                       "*"                    " "                    " "
1  ( 8 )  " "     " "        " "               " "          " "                       " "                    " "                    " "
1  ( 9 )  " "     " "        " "               " "          "*"                       " "                    " "                    " "
2  ( 1 )  " "     "*"        "*"               " "          " "                       " "                    " "                    " "
2  ( 2 )  "*"     "*"        " "               " "          " "                       " "                    " "                    " "
2  ( 3 )  " "     "*"        " "               " "          " "                       " "                    "*"                    " "
2  ( 4 )  " "     "*"        " "               "*"          " "                       " "                    " "                    " "
2  ( 5 )  " "     "*"        " "               " "          " "                       " "                    " "                    "*"
2  ( 6 )  " "     "*"        " "               " "          " "                       " "                    " "                    " "
2  ( 7 )  " "     "*"        " "               " "          "*"                       " "                    " "                    " "
2  ( 8 )  " "     "*"        " "               " "          " "                       "*"                    " "                    " "
2  ( 9 )  " "     " "        "*"               "*"          " "                       " "                    " "                    " "
2  ( 10 ) "*"     " "        "*"               " "          " "                       " "                    " "                    " "
3  ( 1 )  "*"     "*"        "*"               " "          " "                       " "                    " "                    " "
3  ( 2 )  " "     "*"        "*"               "*"          " "                       " "                    " "                    " "
3  ( 3 )  " "     "*"        "*"               " "          " "                       " "                    "*"                    " "
3  ( 4 )  " "     "*"        "*"               " "          " "                       " "                    " "                    "*"
```



bic
-7200
-7200
-7100
-7100
-7100
-7100
-7100
-7100
-7100
-7000
-7000
-6400
-5900
-800
-200
18
(Intercept)
sexmale
population
uicides.100k.pop
HDI.for.year
nG.I.. Generation
tionGeneration X
tionGeneration Z
erationMillenials
generationSilent



0.57
0.56
0.56
0.56
0.56
0.56
0.56
0.56
0.56
0.56
0.55
0.52
0.49
0.096
0.024
7e-08
(Intercept)
sexmale
population
uicides.100k.pop
HDI.for.year
nG.I.. Generation
tionGeneration X
tionGeneration Z
erationMillenials
generationSilent

```
> leaps
Subset selection object
Call: regsubsets.formula(suicides_no ~ sex + population + suicides.100k.pop +
    HDI.for.year + generation, data = data, nbest = 10)
9 Variables  (and intercept)
                          Forced in Forced out
sexmale                       FALSE      FALSE
population                    FALSE      FALSE
suicides.100k.pop             FALSE      FALSE
HDI.for.year                  FALSE      FALSE
generationG.I. Generation     FALSE      FALSE
generationGeneration X        FALSE      FALSE
generationGeneration Z        FALSE      FALSE
generationMillenials          FALSE      FALSE
generationSilent              FALSE      FALSE
10 subsets of each size up to 8
Selection Algorithm: exhaustive
> coef(leaps,1:5)
[[1]]
  (Intercept)    population
-1.687177e+01  1.197004e-04

[[2]]
      (Intercept) suicides.100k.pop
        90.408288          9.376102

[[3]]
(Intercept)     sexmale
   95.44223   209.49954

[[4]]
 (Intercept) HDI.for.year
   -666.1087    1119.3203

[[5]]
           (Intercept) generationGeneration Z
              219.5040              -209.5501

>
```

```
> # Calculate Relative Importance for Each Predictor
> library(relaimpo)
```

```
Metrics betasq, pratt, genizi and car do not work with groups of factors.
> booteval.relimp(boot) # print result
Response variable: suicides_no
Total response variance: 446110.5
Analysis based on 8724 observations

5 Regressors:
sex population suicides.100k.pop HDI.for.year generation
Proportion of variance explained by model: 56.14%
Metrics are normalized to sum to 100% (rela=TRUE).

Relative importance metrics:

                         lmg         last       first       pratt
sex              0.026800311 0.0094290412 0.04089900 0.021400353
population       0.856730352 0.9080579283 0.81225777 0.865225316
suicides.100k.pop 0.085237232 0.0740025629 0.09618045 0.092486439
HDI.for.year     0.024903955 0.0075123784 0.04030559 0.017646912
generation       0.006328149 0.0009980892 0.01035719 0.003240979

Average coefficients for different model sizes:

                          1X           2Xs           3Xs           4Xs           5Xs
sex               2.094995e+02  1.832834e+02  1.575896e+02  1.312000e+02  1.023150e+02
population        1.197004e-04  1.194973e-04  1.193189e-04  1.191607e-04  1.190088e-04
suicides.100k.pop 9.376102e+00  9.073520e+00  8.796008e+00  8.566505e+00  8.415146e+00
HDI.for.year      1.119320e+03  9.652920e+02  8.039735e+02  6.349106e+02  4.574103e+02
generation       -3.089841e+01 -2.626573e+01 -2.110123e+01 -1.537090e+01 -9.024404e+00


 Confidence interval information ( 1000 bootstrap replicates, bty= perc ):
Relative Contributions with confidence intervals:

                                       Lower  Upper
                        percentage 0.95 0.95   0.95
sex.lmg                 0.0268     __CD_ 0.0230 0.0315
population.lmg          0.8567     A____ 0.8303 0.8779
suicides.100k.pop.lmg   0.0852     _B___ 0.0683 0.1093
HDI.for.year.lmg        0.0249     __CD_ 0.0211 0.0294
generation.lmg          0.0063     ____E 0.0033 0.0105

sex.last                0.0094     __CD_ 0.0069 0.0125
population.last         0.9081     A____ 0.8846 0.9252
suicides.100k.pop.last  0.0740     _B___ 0.0586 0.0969
HDI.for.year.last       0.0075     __CD_ 0.0050 0.0107
generation.last         0.0010     ____E 0.0001 0.0030

sex.first               0.0409     __CD_ 0.0341 0.0483
population.first        0.8123     A____ 0.7813 0.8357
suicides.100k.pop.first 0.0962     _B___ 0.0781 0.1214
HDI.for.year.first      0.0403     __CD_ 0.0346 0.0471
generation.first        0.0104     ____E 0.0053 0.0167

sex.pratt               0.0214     __CD_ 0.0177 0.0257
population.pratt        0.8652     A____ 0.8372 0.8866
```

```
Letters indicate the ranks covered by bootstrap CIs.
(Rank bootstrap confidence intervals always obtained by percentile method)
CAUTION: Bootstrap confidence intervals can be somewhat liberal.


 Differences between Relative Contributions:

                                                      Lower   Upper
                                     difference 0.95 0.95    0.95
sex-population.lmg                   -0.8299     *   -0.8542 -0.7997
sex-suicides.100k.pop.lmg            -0.0584     *   -0.0803 -0.0435
sex-HDI.for.year.lmg                  0.0019         -0.0034  0.0073
sex-generation.lmg                    0.0205     *    0.0149  0.0260
population-suicides.100k.pop.lmg      0.7715     *    0.7214  0.8086
population-HDI.for.year.lmg           0.8318     *    0.8049  0.8544
population-generation.lmg             0.8504     *    0.8228  0.8726
suicides.100k.pop-HDI.for.year.lmg    0.0603     *    0.0433  0.0837
suicides.100k.pop-generation.lmg      0.0789     *    0.0613  0.1028
HDI.for.year-generation.lmg           0.0186     *    0.0133  0.0239

sex-population.last                  -0.8986     *   -0.9166 -0.8752
sex-suicides.100k.pop.last           -0.0646     *   -0.0879 -0.0485
sex-HDI.for.year.last                 0.0019         -0.0014  0.0052
sex-generation.last                   0.0084     *    0.0054  0.0116
population-suicides.100k.pop.last     0.8341     *    0.7878  0.8669
population-HDI.for.year.last          0.9005     *    0.8763  0.9193
population-generation.last            0.9071     *    0.8833  0.9240
suicides.100k.pop-HDI.for.year.last   0.0665     *    0.0513  0.0888
suicides.100k.pop-generation.last     0.0730     *    0.0573  0.0960
HDI.for.year-generation.last          0.0065     *    0.0038  0.0095

sex-population.first                 -0.7714     *   -0.8008 -0.7353
sex-suicides.100k.pop.first          -0.0553     *   -0.0761 -0.0396
sex-HDI.for.year.first                0.0006         -0.0095  0.0101
sex-generation.first                  0.0305     *    0.0210  0.0402
population-suicides.100k.pop.first    0.7161     *    0.6619  0.7574
population-HDI.for.year.first         0.7720     *    0.7416  0.7975
population-generation.first           0.8019     *    0.7689  0.8276
suicides.100k.pop-HDI.for.year.first  0.0559     *    0.0368  0.0803
suicides.100k.pop-generation.first    0.0858     *    0.0667  0.1108
HDI.for.year-generation.first         0.0299     *    0.0211  0.0385

sex-population.pratt                 -0.8438     *   -0.8672 -0.8138
sex-suicides.100k.pop.pratt          -0.0711     *   -0.0973 -0.0532
sex-HDI.for.year.pratt                0.0038         -0.0008  0.0081
sex-generation.pratt                  0.0182     *    0.0136  0.0228
population-suicides.100k.pop.pratt    0.7727     *    0.7192  0.8114
population-HDI.for.year.pratt         0.8476     *    0.8181  0.8711
population-generation.pratt           0.8620     *    0.8331  0.8839
suicides.100k.pop-HDI.for.year.pratt  0.0748     *    0.0573  0.1006
suicides.100k.pop-generation.pratt    0.0892     *    0.0706  0.1151
```

```
* indicates that CI for difference does not include 0.
CAUTION: Bootstrap confidence intervals can be somewhat liberal.
> plot(booteval.relimp(boot,sort=TRUE)) # plot result
> #https://rpubs.com/davoodastaraky/mtRegression
> summary(fit2)

Call:
lm(formula = suicides_no ~ sex + population + suicides.100k.pop +
    HDI.for.year + generation, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-2309.4   -94.5    12.3   111.6  6492.9

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)               -5.041e+02  4.224e+01 -11.936  < 2e-16 ***
sexmale                    1.160e+02  1.046e+01  11.084  < 2e-16 ***
population                 1.190e-04  1.227e-06  96.991  < 2e-16 ***
suicides.100k.pop          7.422e+00  3.293e-01  22.542  < 2e-16 ***
HDI.for.year               5.234e+02  5.284e+01   9.905  < 2e-16 ***
generationG.I. Generation -1.982e+01  2.162e+01  -0.917    0.359
generationGeneration X    -6.249e+01  1.559e+01  -4.007 6.19e-05 ***
generationGeneration Z    -1.605e+02  2.040e+01  -7.868 4.02e-15 ***
generationMillenials      -1.054e+02  1.583e+01  -6.657 2.96e-11 ***
generationSilent          -2.492e+01  1.612e+01  -1.546    0.122
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 440.6 on 8714 degrees of freedom
  (19096 observations deleted due to missingness)
Multiple R-squared:  0.5654,	Adjusted R-squared:  0.5649
F-statistic:  1259 on 9 and 8714 DF,  p-value: < 2.2e-16
```