The effect of pollution on women's health in the U.S., China, India, Brazil, and Russia

Mandy Hooks, Natalia Jaffee, Amrita Sood

Spring 2021

```
##
## Call:
## lm(formula = SDG_Data.short$GDP ~ SDG_Data.short$CO2)
##
## Coefficients:
##        (Intercept)  SDG_Data.short$CO2
##               3952                1964
```

```
##
## Call:
## lm(formula = SDG_Data.short$GDP ~ SDG_Data.short$CO2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -23385.1  -5850.0     18.3   9416.4  21473.5
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           3952.0     6466.0   0.611   0.5535
## SDG_Data.short$CO2    1963.6      641.1   3.063   0.0108 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13920 on 11 degrees of freedom
## Multiple R-squared:  0.4602, Adjusted R-squared:  0.4112
## F-statistic: 9.379 on 1 and 11 DF,  p-value: 0.0108
```

```
##
## Call:
## lm(formula = SDG_Data.short$GDP ~ SDG_Data.short$PM2.5)
##
## Coefficients:
```

Table 1: Summary Statistics of Variables in Countries

Variable	Mean	Median	Standard Deviation
Access to clean fuels and technologies for cooking (% of population)	79.9320	95.9900	30.600
CO2 emissions (metric tons per capita)	8.0916	7.3985	6.265
GDP per capita (current US$)	19840.0000	13574.0000	18130.000
Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70, female (%)	14.0920	13.8000	5.522
PM2.5 air pollution, mean annual exposure (micrograms per cubic meter)	34.3370	16.6400	33.200
Renewable electricity output (% of total electricity output)	27.0110	18.1950	21.290

```
##          (Intercept)  SDG_Data.short$PM2.5
##              28776.8                -260.3
```

```
##
## Call:
## lm(formula = SDG_Data.short$GDP ~ SDG_Data.short$PM2.5)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -21114 -11180  -5338  15701  30144
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)           28776.8     6785.0   4.241  0.00139 **
## SDG_Data.short$PM2.5   -260.3      144.8  -1.798  0.09968 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16650 on 11 degrees of freedom
## Multiple R-squared:  0.2271, Adjusted R-squared:  0.1568
## F-statistic: 3.232 on 1 and 11 DF,  p-value: 0.09968
```

```
##
## Call:
## lm(formula = SDG_Data.short$PM2.5 ~ SDG_Data.short$Mortality)
##
## Coefficients:
##              (Intercept)  SDG_Data.short$Mortality
##                  -10.317                     3.169
```

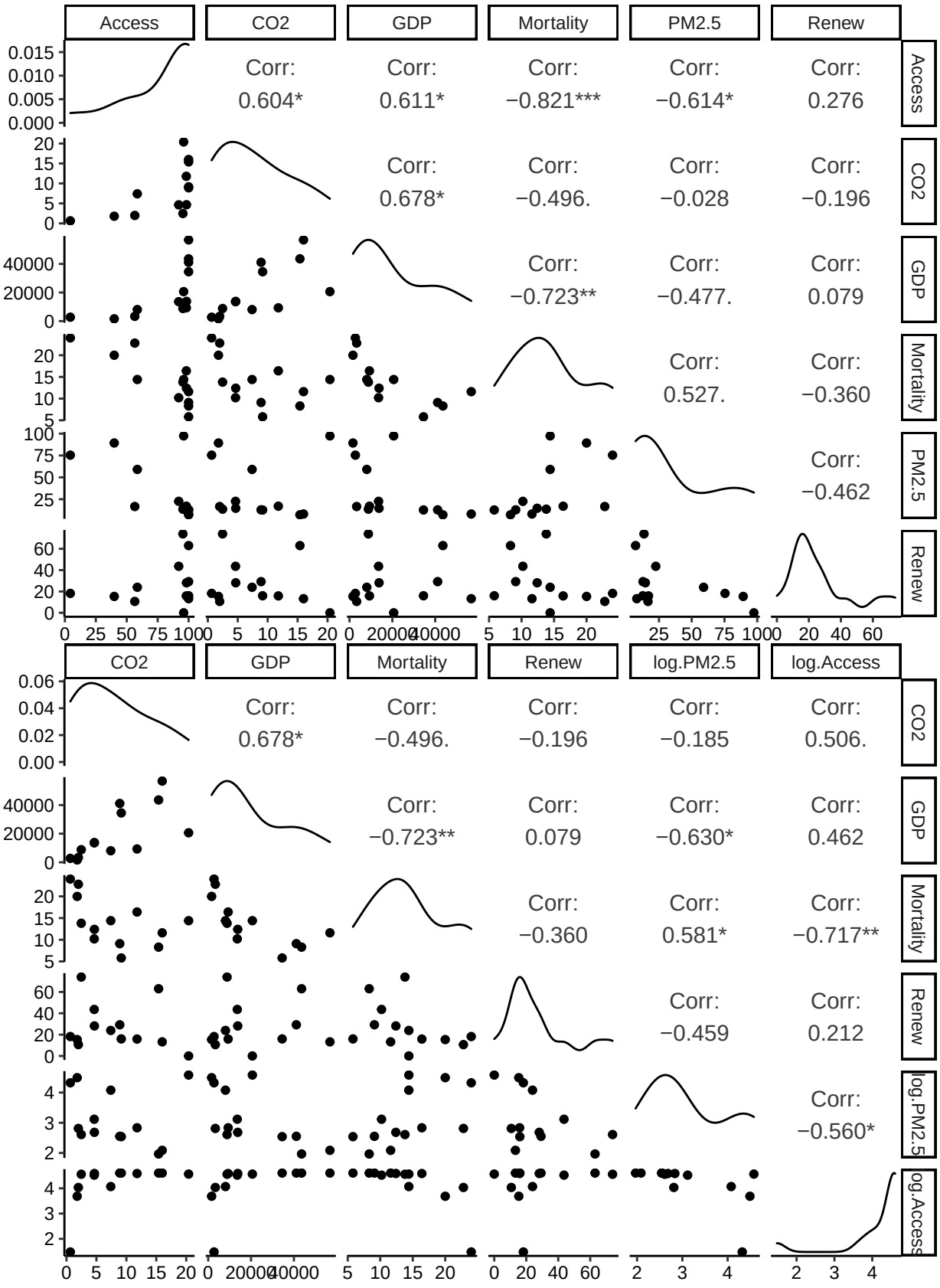
```
##
## Call:
## lm(formula = SDG_Data.short$PM2.5 ~ SDG_Data.short$Mortality)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -45.289 -18.376  -5.692   9.667  62.120
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               -10.317     23.201  -0.445   0.6652
## SDG_Data.short$Mortality    3.169      1.541   2.057   0.0642 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 29.47 on 11 degrees of freedom
## Multiple R-squared:  0.2777, Adjusted R-squared:  0.2121
## F-statistic:  4.23 on 1 and 11 DF,  p-value: 0.06424
```

```
##
## Call:
## lm(formula = SDG_Data.short$CO2 ~ SDG_Data.short$Renew)
##
## Coefficients:
##          (Intercept)  SDG_Data.short$Renew
##               9.6502               -0.0577
```

```
##
## Call:
## lm(formula = SDG_Data.short$CO2 ~ SDG_Data.short$Renew)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.960 -3.362 -0.871  3.050 10.752
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           9.65017    2.94801   3.273  0.00742 **
## SDG_Data.short$Renew -0.05770    0.08701  -0.663  0.52088
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.417 on 11 degrees of freedom
## Multiple R-squared:  0.03844,    Adjusted R-squared:  -0.04897
## F-statistic: 0.4398 on 1 and 11 DF,  p-value: 0.5209
```

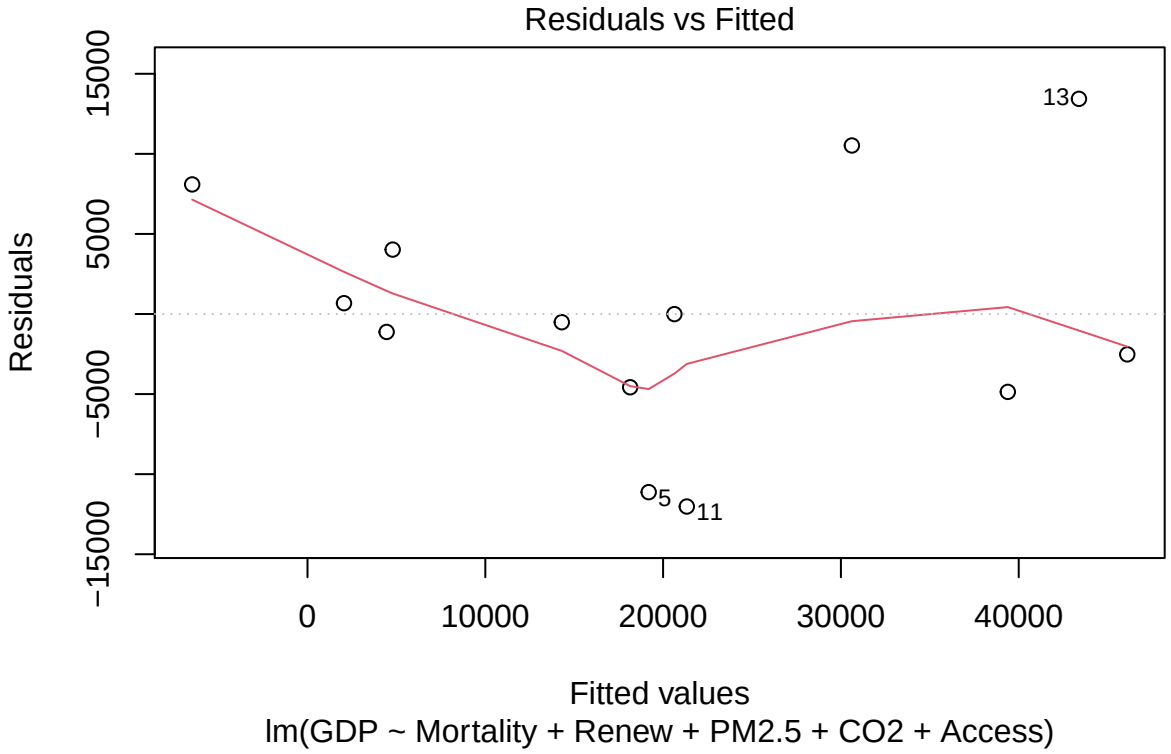
```
##
## Call:
## lm(formula = SDG_Data.short$Mortality ~ SDG_Data.short$Renew)
##
## Coefficients:
##          (Intercept)  SDG_Data.short$Renew
##              16.61404              -0.09336
```

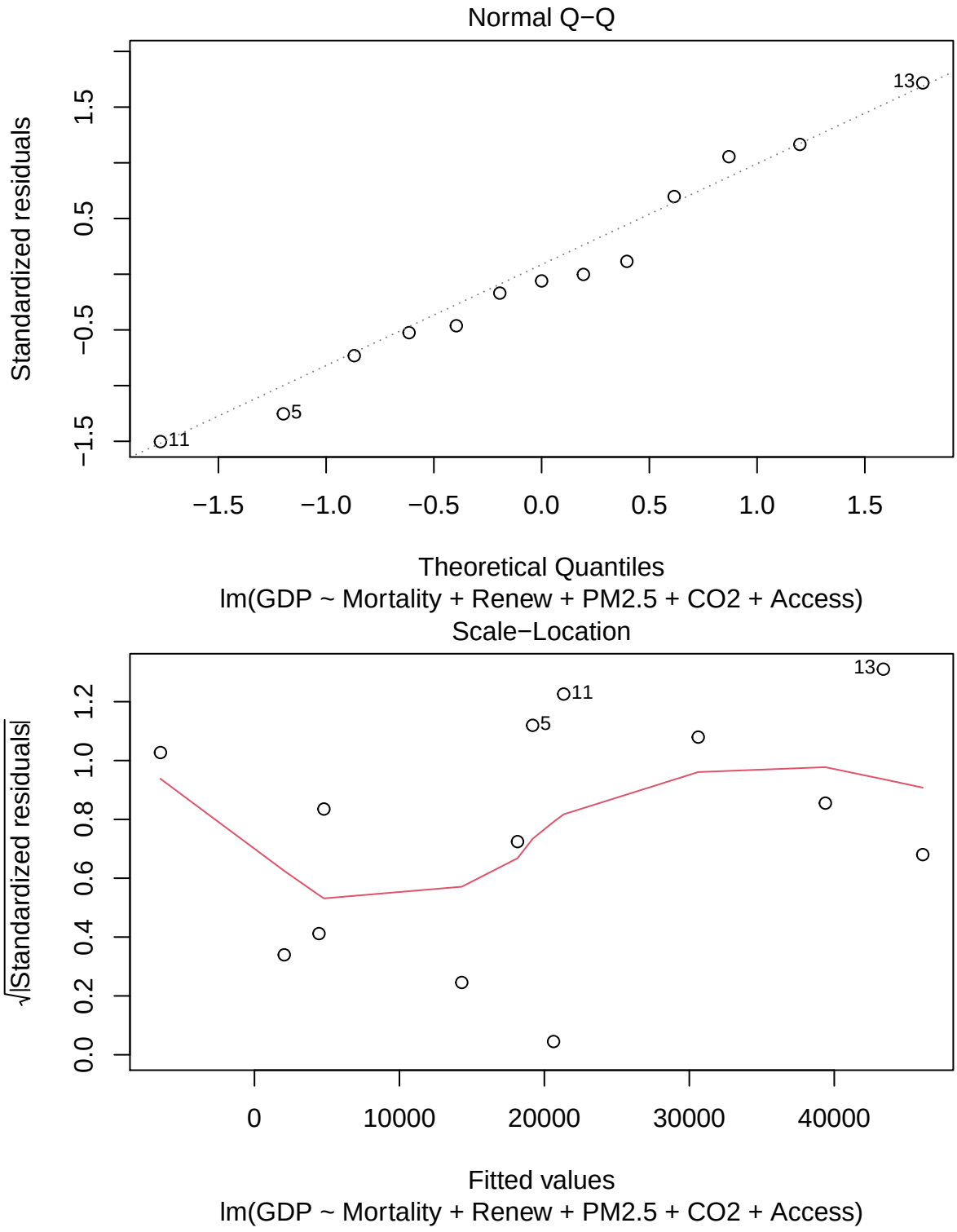
```
##
## Call:
## lm(formula = SDG_Data.short$Mortality ~ SDG_Data.short$Renew)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.322 -2.431 -1.587  4.091  9.085
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           16.61404    2.47225   6.720 3.29e-05 ***
## SDG_Data.short$Renew -0.09336    0.07297  -1.279    0.227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.381 on 11 degrees of freedom
## Multiple R-squared:  0.1295, Adjusted R-squared:  0.05041
## F-statistic: 1.637 on 1 and 11 DF,  p-value: 0.2271
```

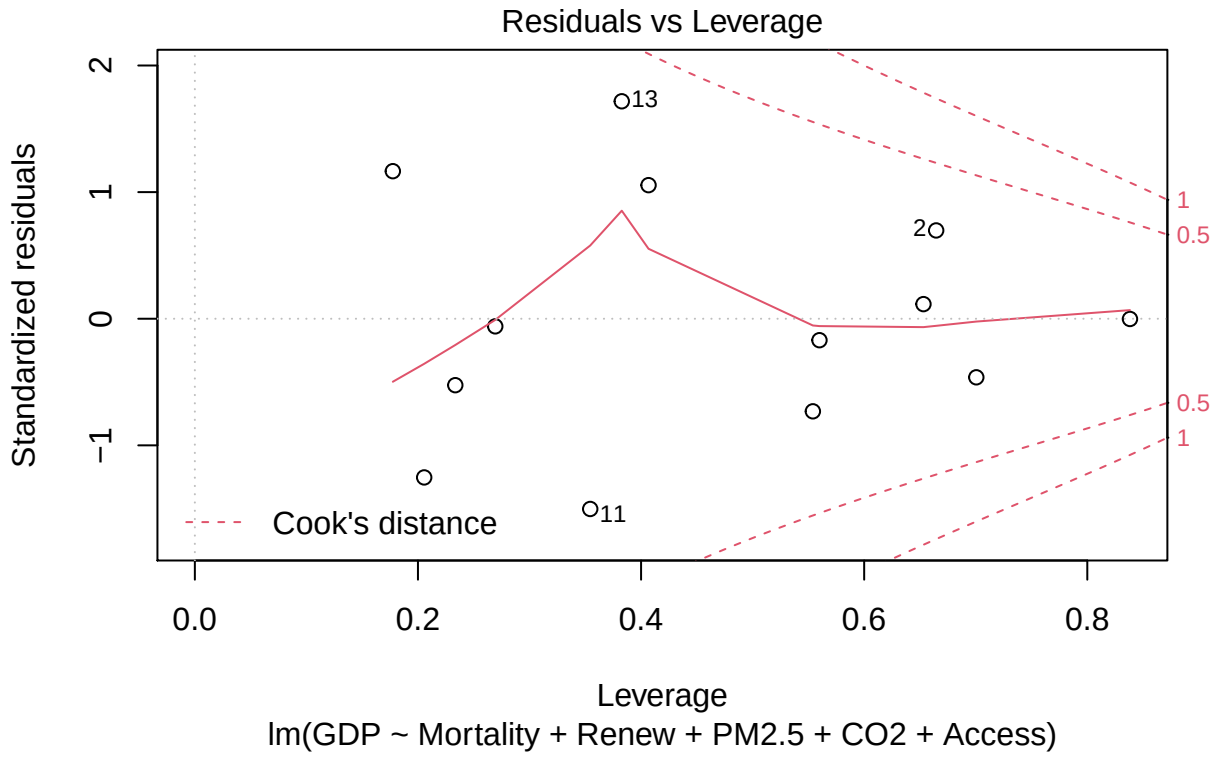

Access
CO2
GDP
Mortality
PM2.5
Renew
Corr: 0.604*
Corr: 0.611*
Corr: −0.821***
Corr: −0.614*
Corr: 0.276
Corr: 0.678*
Corr: −0.496.
Corr: −0.028
Corr: −0.196
Corr: −0.723**
Corr: −0.477.
Corr: 0.079
Corr: 0.527.
Corr: −0.360
Corr: −0.462
CO2
GDP
Mortality
Renew
log.PM2.5
log.Access
Corr: 0.678*
Corr: −0.496.
Corr: −0.196
Corr: −0.185
Corr: 0.506.
Corr: −0.723**
Corr: 0.079
Corr: −0.630*
Corr: 0.462
Corr: −0.360
Corr: 0.581*
Corr: −0.717**
Corr: −0.459
Corr: 0.212
Corr: −0.560*

##

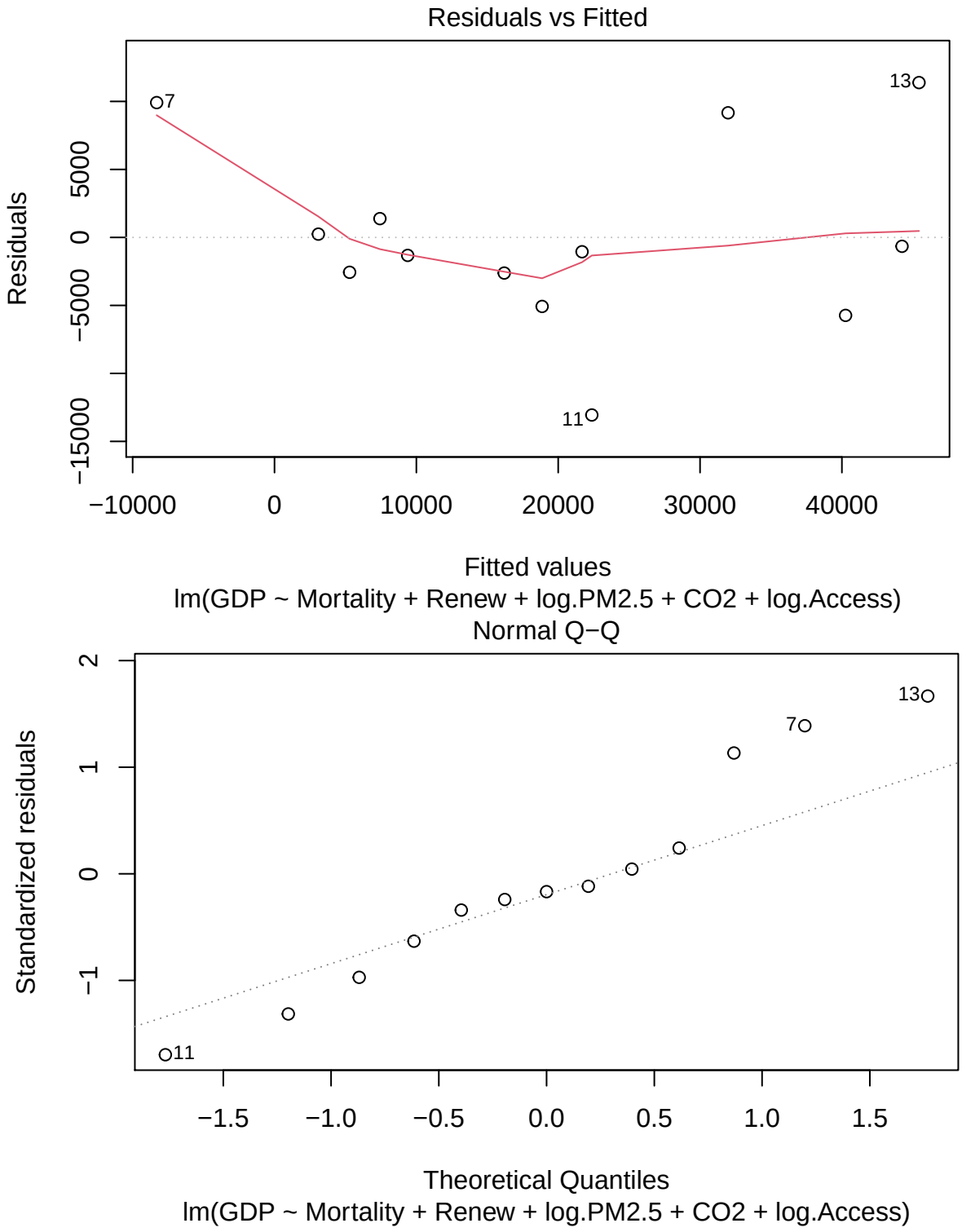
```
## Call:
## lm(formula = GDP ~ Mortality + Renew + PM2.5 + CO2 + Access,
##     data = SDG_Data.short)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12020.6  -4574.9   -513.6   4021.8  13434.4
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 80952.90   28902.36   2.801   0.0265 *
## Mortality   -2193.66     947.64  -2.315   0.0538 .
## Renew         -70.21     168.19  -0.417   0.6889
## PM2.5        -324.42     132.65  -2.446   0.0444 *
## CO2          2198.56     735.93   2.987   0.0203 *
## Access       -437.28     218.60  -2.000   0.0856 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9957 on 7 degrees of freedom
## Multiple R-squared:  0.8241, Adjusted R-squared:  0.6985
## F-statistic: 6.561 on 5 and 7 DF,  p-value: 0.0142
```

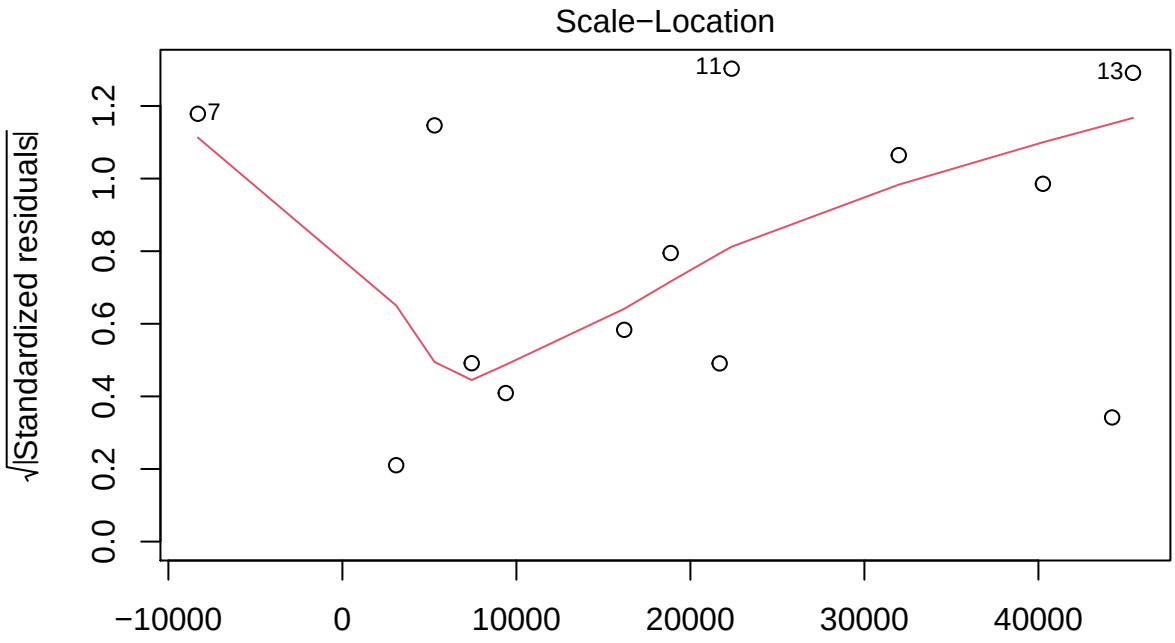


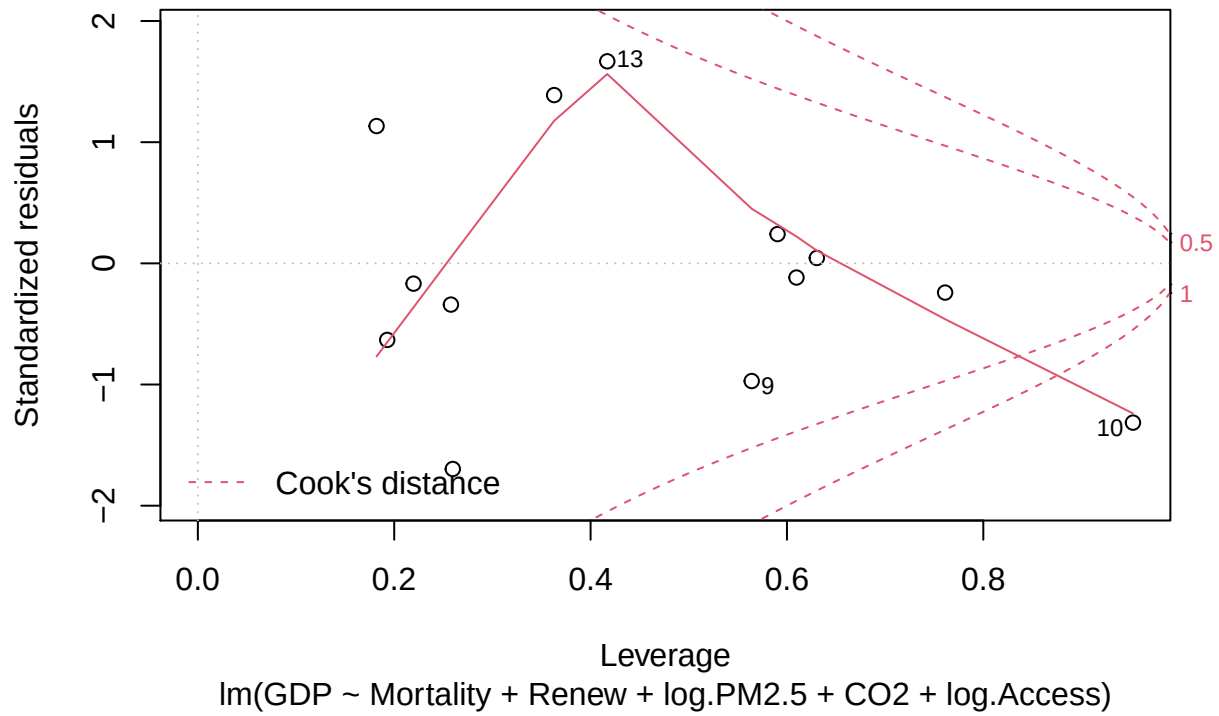

Normal Q−Q
Standardized residuals
1.5
0.5
−0.5
−1.5
13
5
11
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Theoretical Quantiles
lm(GDP ~ Mortality + Renew + PM2.5 + CO2 + Access)
Scale−Location
√|Standardized residuals|
1.2
1.0
0.8
0.6
0.4
0.2
0.0
13
11
5
0
10000
20000
30000
40000
Fitted values
lm(GDP ~ Mortality + Renew + PM2.5 + CO2 + Access)




```
##
## Call:
## lm(formula = GDP ~ Mortality + Renew + log.PM2.5 + CO2 + log.Access,
##     data = SDG_Data.short)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -13058  -2622  -1053   1382  11385
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 111240.5    31958.5   3.481   0.0103 *
## Mortality    -1760.3      778.4  -2.261   0.0582 .
## Renew         -149.9      153.8  -0.975   0.3620
## log.PM2.5   -11084.7     3844.7  -2.883   0.0235 *
## CO2           1455.6      553.1   2.632   0.0338 *
## log.Access   -9454.0     4625.6  -2.044   0.0803 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8942 on 7 degrees of freedom
## Multiple R-squared:  0.8582, Adjusted R-squared:  0.7568
## F-statistic:  8.47 on 5 and 7 DF,  p-value: 0.007003
```

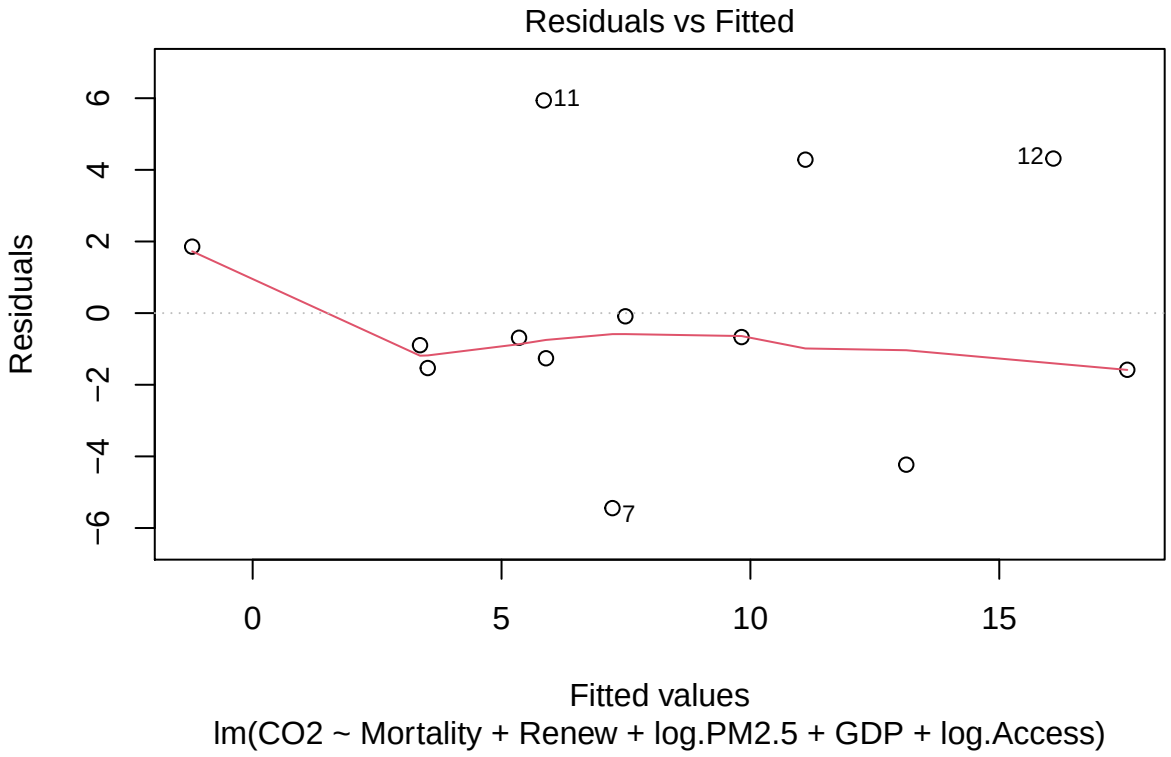

Residuals vs Fitted
Residuals
Fitted values
lm(GDP ~ Mortality + Renew + log.PM2.5 + CO2 + log.Access)
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(GDP ~ Mortality + Renew + log.PM2.5 + CO2 + log.Access)

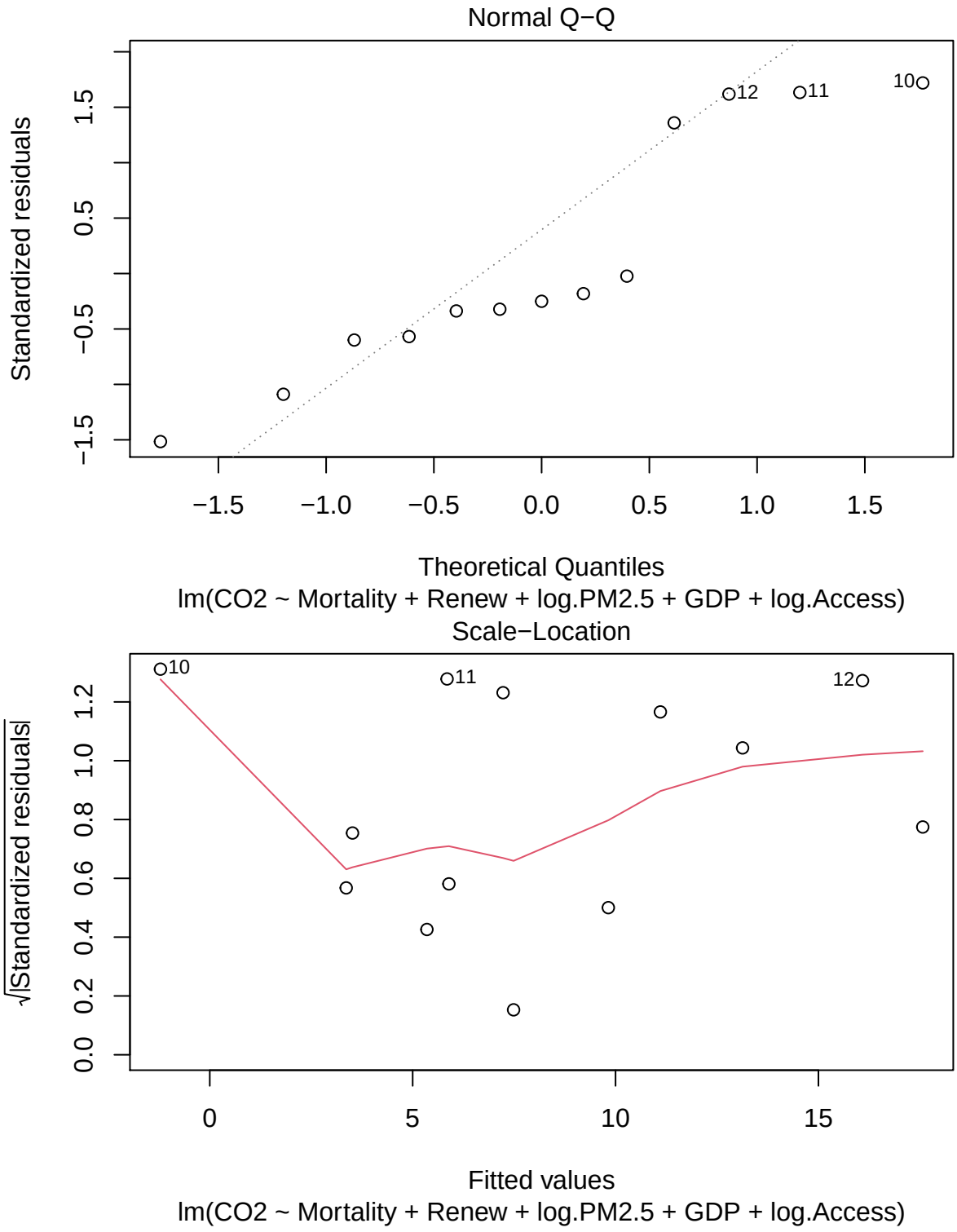


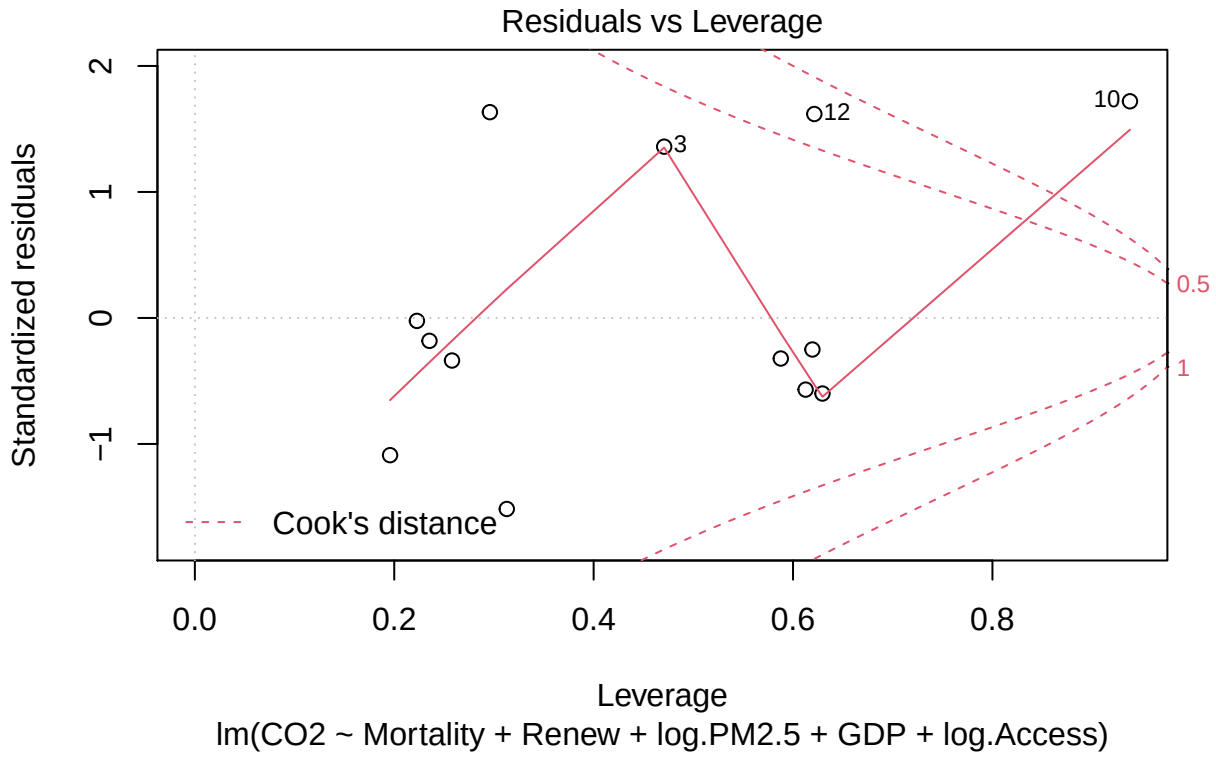



```
##
## Call:
## lm(formula = CO2 ~ Mortality + Renew + log.PM2.5 + GDP + log.Access,
##     data = SDG_Data.short)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -5.4440 -1.5334 -0.6877  1.8555  5.9373
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.376e+01  2.218e+01  -1.522   0.1717
## Mortality    3.336e-01  4.798e-01   0.695   0.5093
## Renew       -7.241e-03  7.934e-02  -0.091   0.9298
## log.PM2.5    4.024e+00  2.297e+00   1.752   0.1233
## GDP          3.417e-04  1.298e-04   2.632   0.0338 *
## log.Access   4.288e+00  2.322e+00   1.846   0.1073
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.332 on 7 degrees of freedom
## Multiple R-squared:  0.7211, Adjusted R-squared:  0.5219
## F-statistic:  3.62 on 5 and 7 DF,  p-value: 0.06171
```

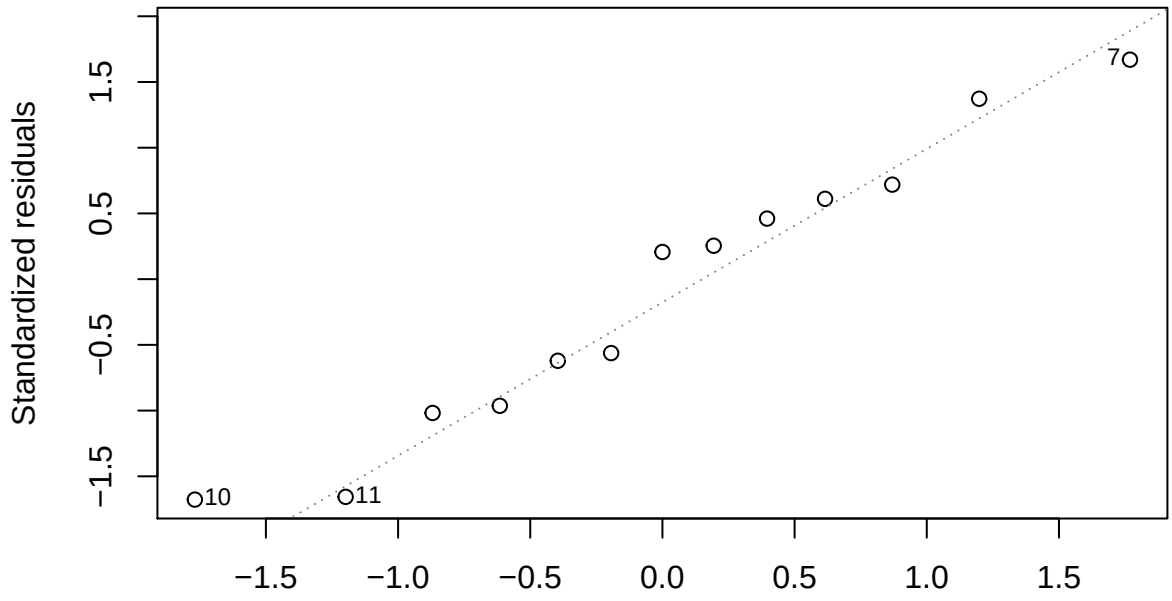


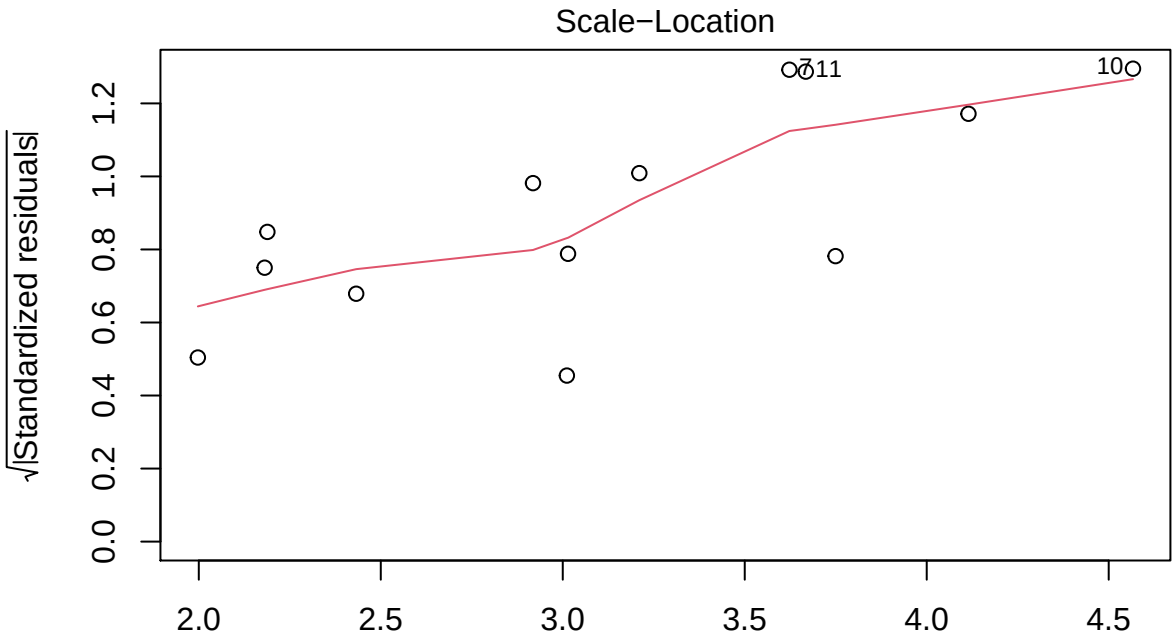

Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(CO2 ~ Mortality + Renew + log.PM2.5 + GDP + log.Access)
Scale−Location
√|Standardized residuals|
Fitted values
lm(CO2 ~ Mortality + Renew + log.PM2.5 + GDP + log.Access)

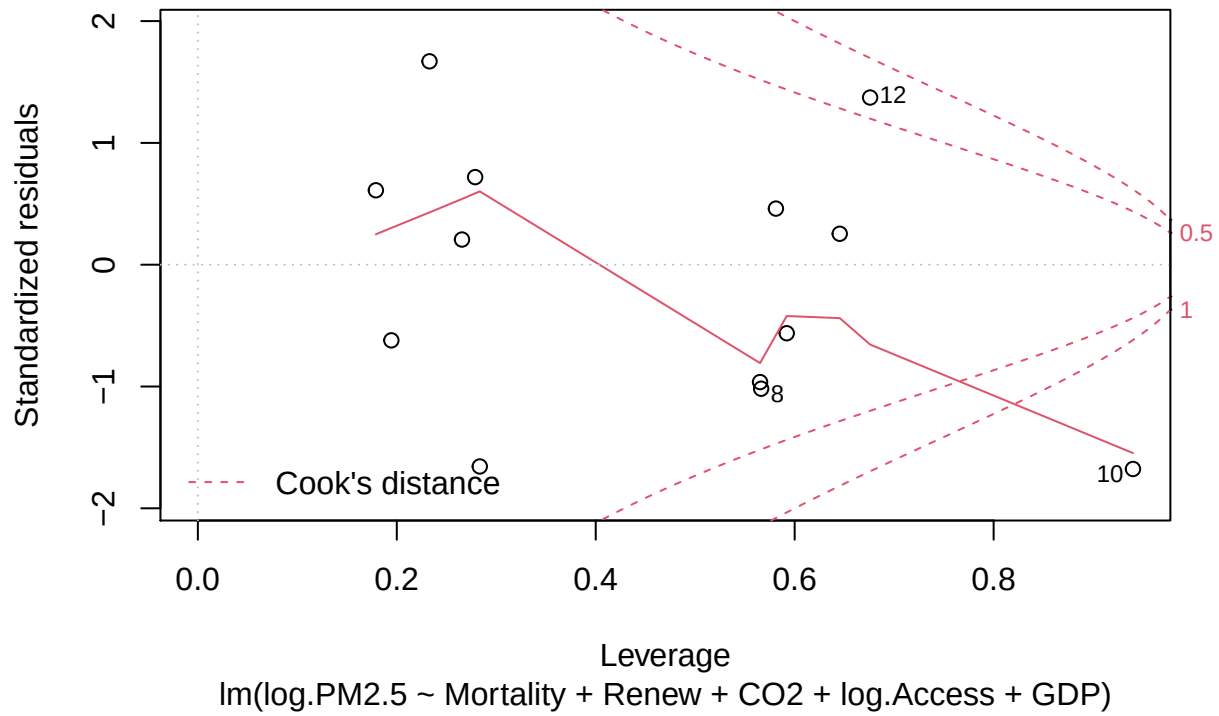



```
##
## Call:
## lm(formula = log.PM2.5 ~ Mortality + Renew + CO2 + log.Access +
##     GDP, data = SDG_Data.short)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.83337 -0.33133  0.08997  0.32913  0.86930
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.394e+00  2.124e+00   3.481   0.0103 *
## Mortality   -6.584e-02  6.335e-02  -1.039   0.3332
## Renew       -1.279e-02  9.761e-03  -1.310   0.2316
## CO2          7.575e-02  4.324e-02   1.752   0.1233
## log.Access  -6.280e-01  3.076e-01  -2.042   0.0805 .
## GDP         -4.897e-05  1.699e-05  -2.883   0.0235 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5944 on 7 degrees of freedom
## Multiple R-squared:  0.7541, Adjusted R-squared:  0.5785
## F-statistic: 4.294 on 5 and 7 DF,  p-value: 0.04165
```


Residuals vs Fitted
Residuals
1.0
0.5
0.0
−0.5
−1.0
7
12
11
2.0
2.5
3.0
3.5
4.0
4.5
Fitted values
lm(log.PM2.5 ~ Mortality + Renew + CO2 + log.Access + GDP)

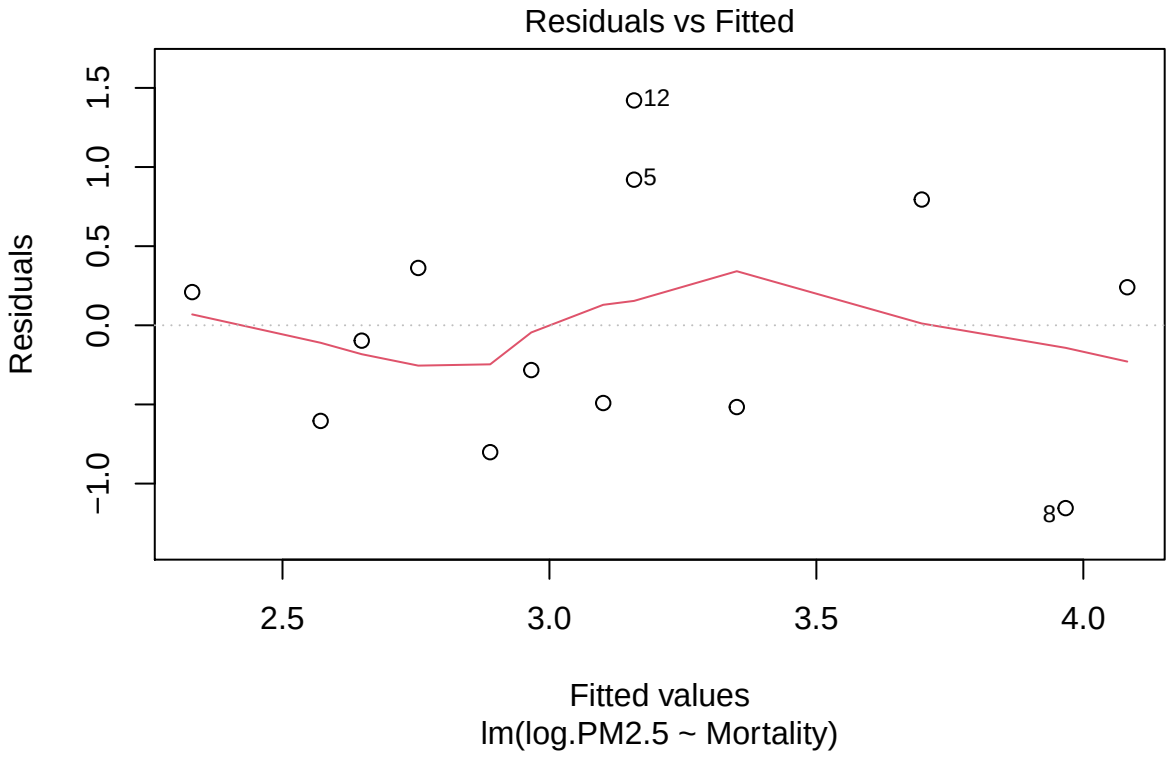

Normal Q−Q
Standardized residuals
1.5
0.5
−0.5
−1.5
7
10
11
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Theoretical Quantiles
lm(log.PM2.5 ~ Mortality + Renew + CO2 + log.Access + GDP)

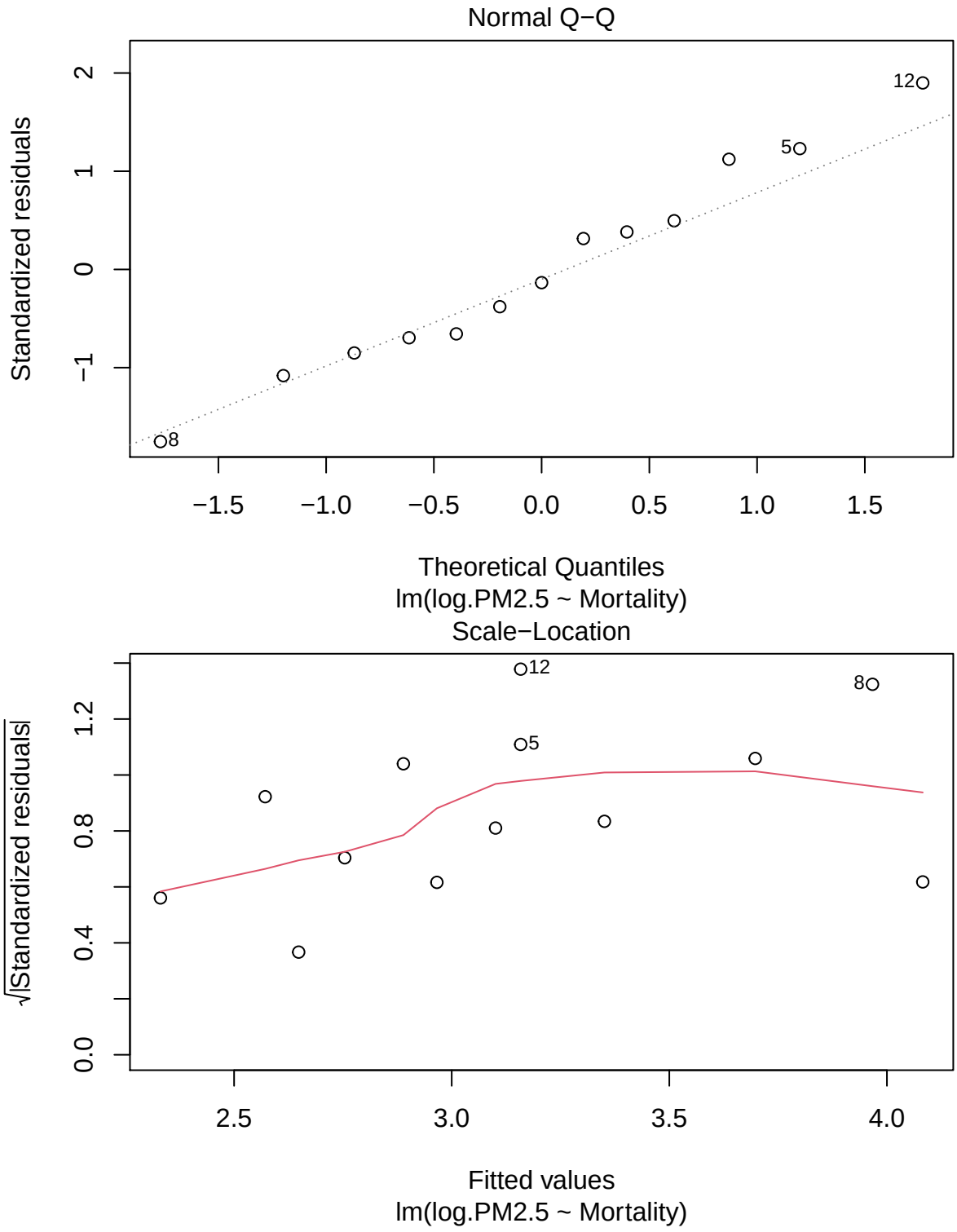


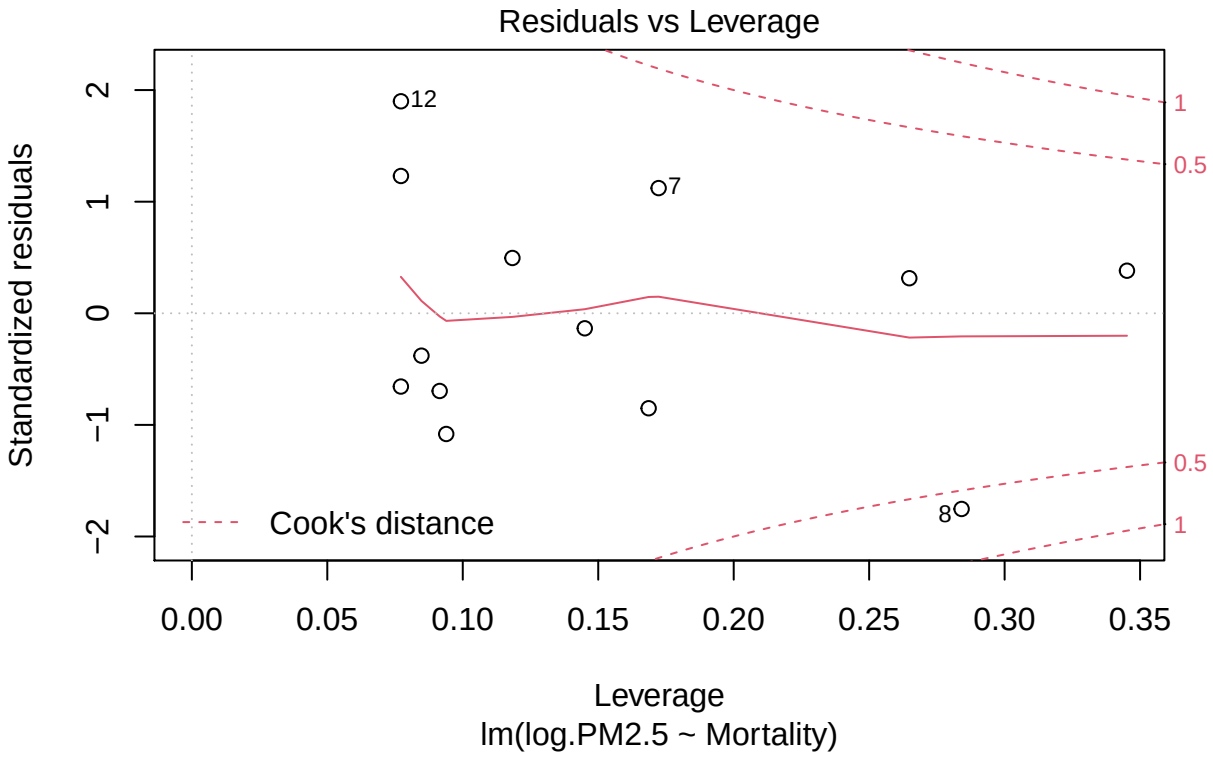


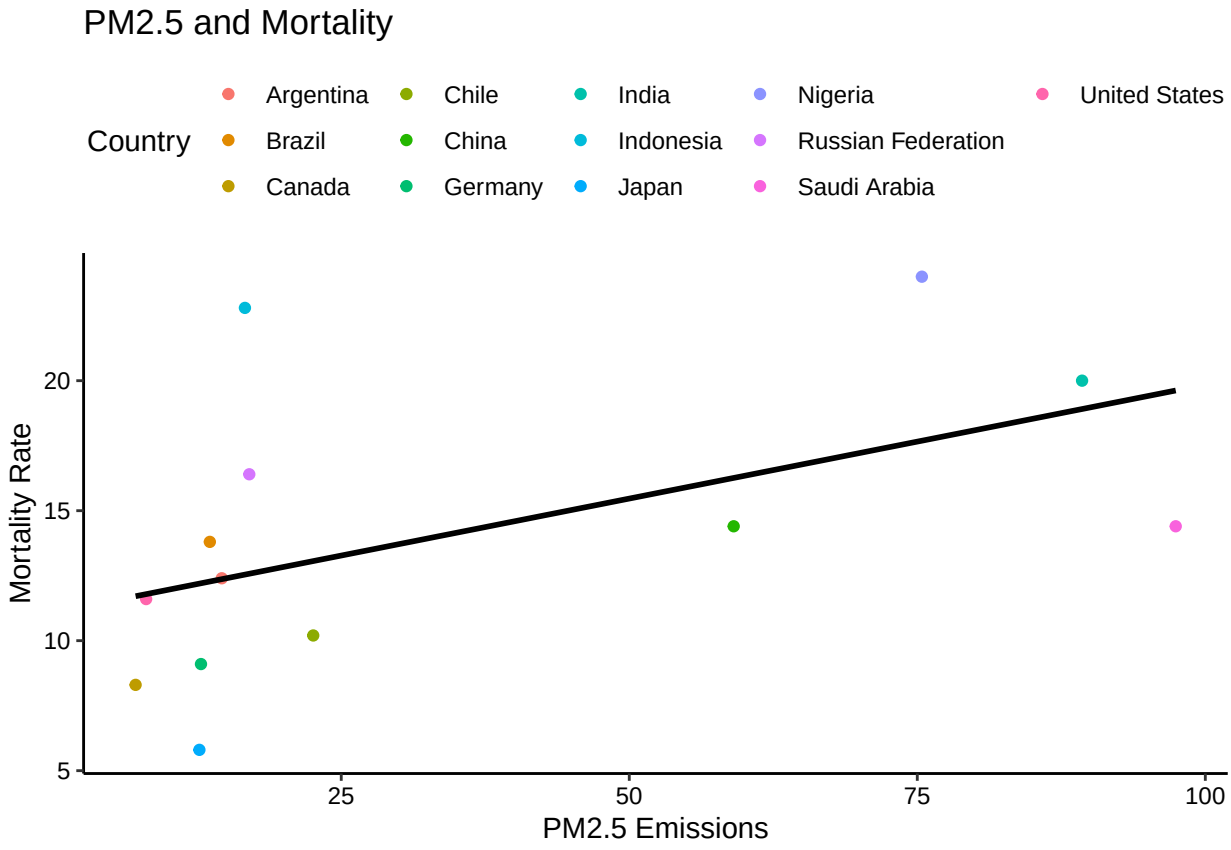

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Mortality    1  3.389   3.389   5.591 0.0375 *
## Residuals   11  6.669   0.606
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

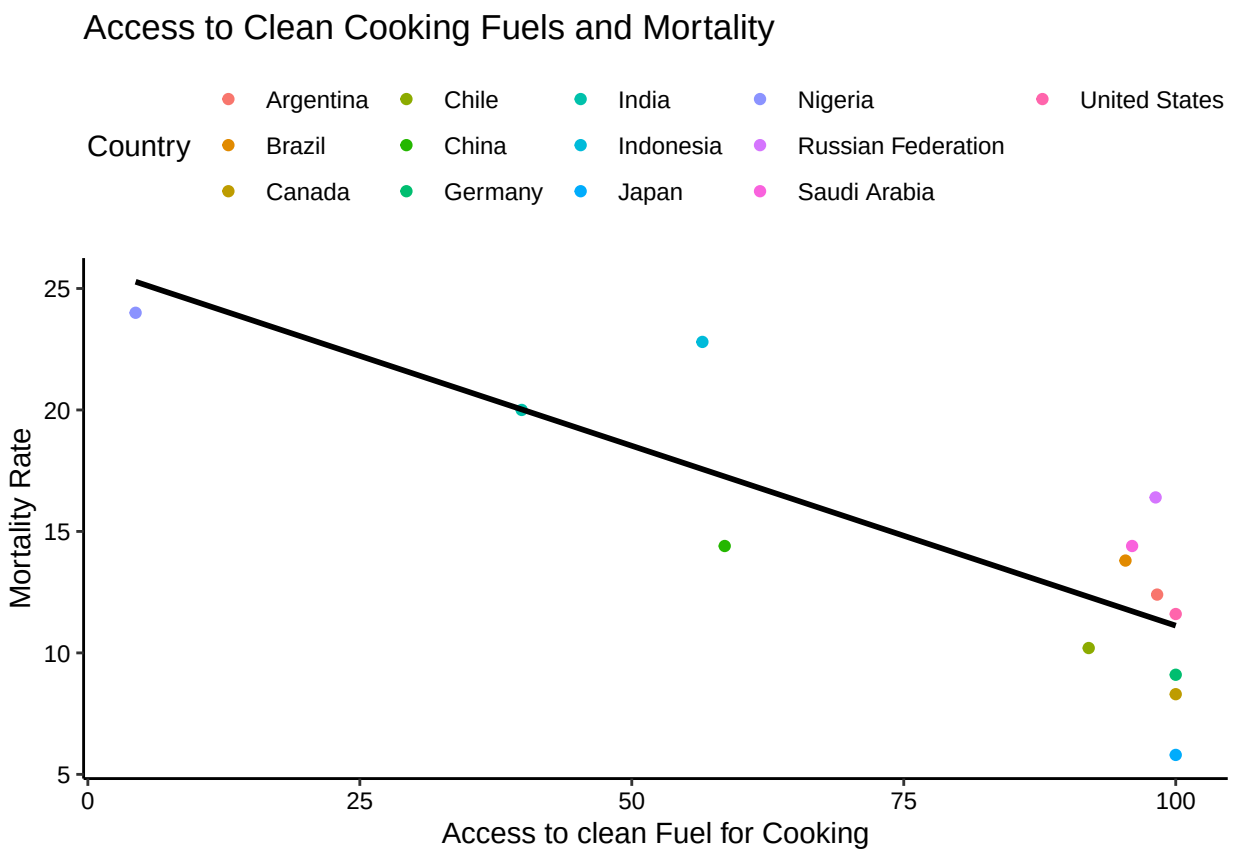
```
## Call:
## lm(formula = log.PM2.5 ~ Mortality, data = SDG_Data.short)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.15503 -0.51660 -0.09692  0.36235  1.42069
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.77264    0.61288   2.892   0.0146 *
## Mortality    0.09624    0.04070   2.364   0.0375 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7786 on 11 degrees of freedom
## Multiple R-squared:  0.337,  Adjusted R-squared:  0.2767
## F-statistic: 5.591 on 1 and 11 DF,  p-value: 0.03752
```

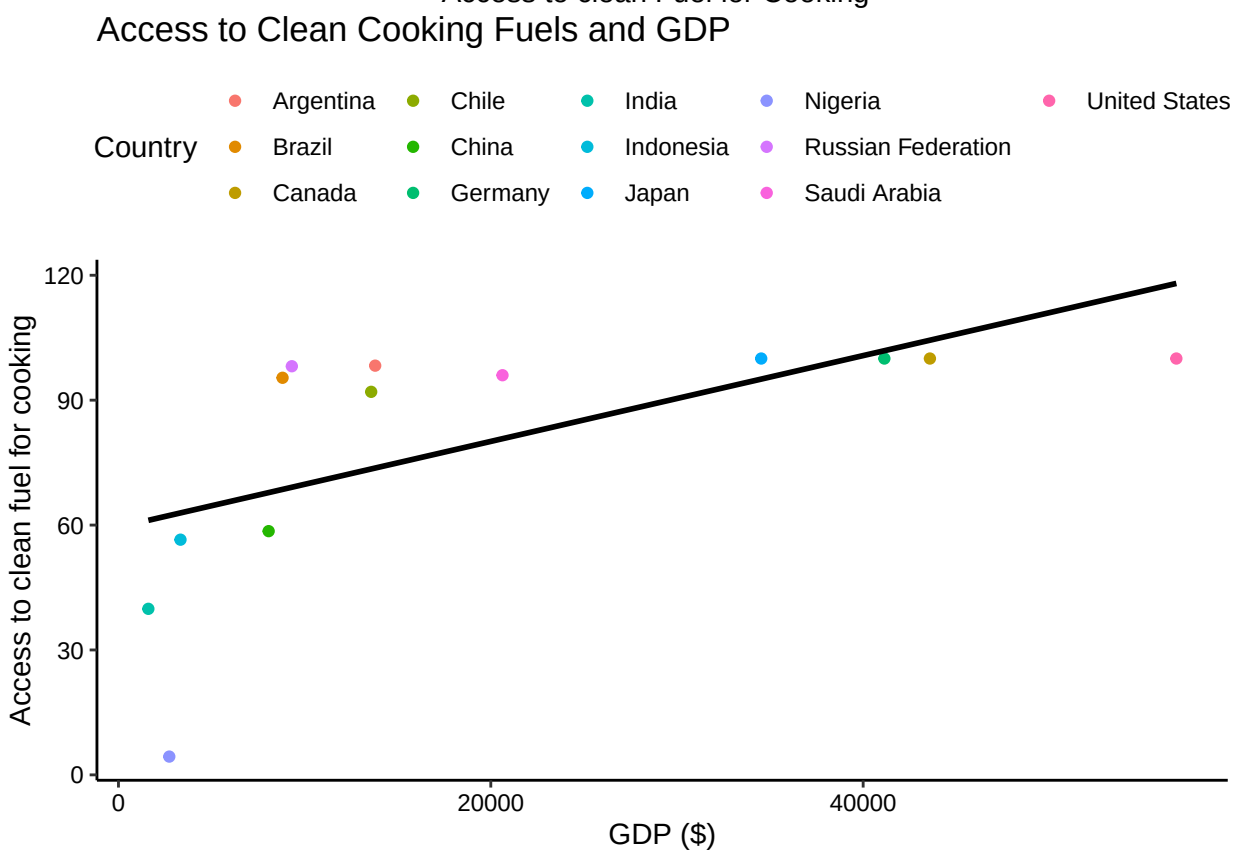



Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(log.PM2.5 ~ Mortality)
Scale−Location
√|Standardized residuals|
Fitted values
lm(log.PM2.5 ~ Mortality)


Residuals vs Leverage
Standardized residuals
Leverage
lm(log.PM2.5 ~ Mortality)
Cook's distance
12
7
8
1
0.5
0.5
1
2
1
0
−1
−2
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35


PM2.5 and Mortality
Country
Argentina
Brazil
Canada
Chile
China
Germany
India
Indonesia
Japan
Nigeria
Russian Federation
Saudi Arabia
United States
Mortality Rate
PM2.5 Emissions
5
10
15
20
25
50
75
100


Access to Clean Cooking Fuels and Mortality
Country
Argentina
Brazil
Canada
Chile
China
Germany
India
Indonesia
Japan
Nigeria
Russian Federation
Saudi Arabia
United States
Mortality Rate
25
20
15
10
5
0
25
50
75
100
Access to clean Fuel for Cooking


Access to Clean Cooking Fuels and GDP
Country
Argentina
Brazil
Canada
Chile
China
Germany
India
Indonesia
Japan
Nigeria
Russian Federation
Saudi Arabia
United States
Access to clean fuel for cooking
120
90
60
30
0
0
20000
40000
GDP ($)