

HW3

Reading data

```
library('MASS')  
data(Boston)
```

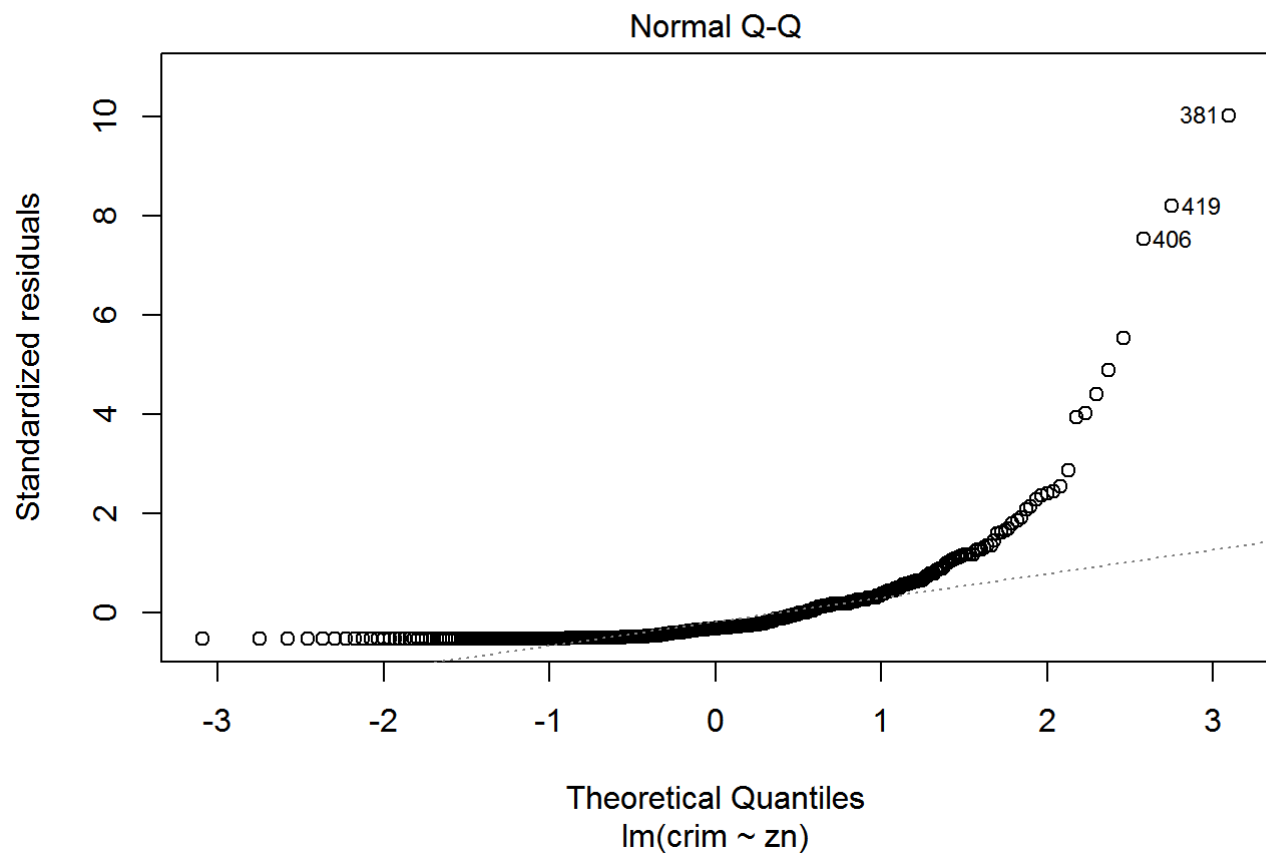
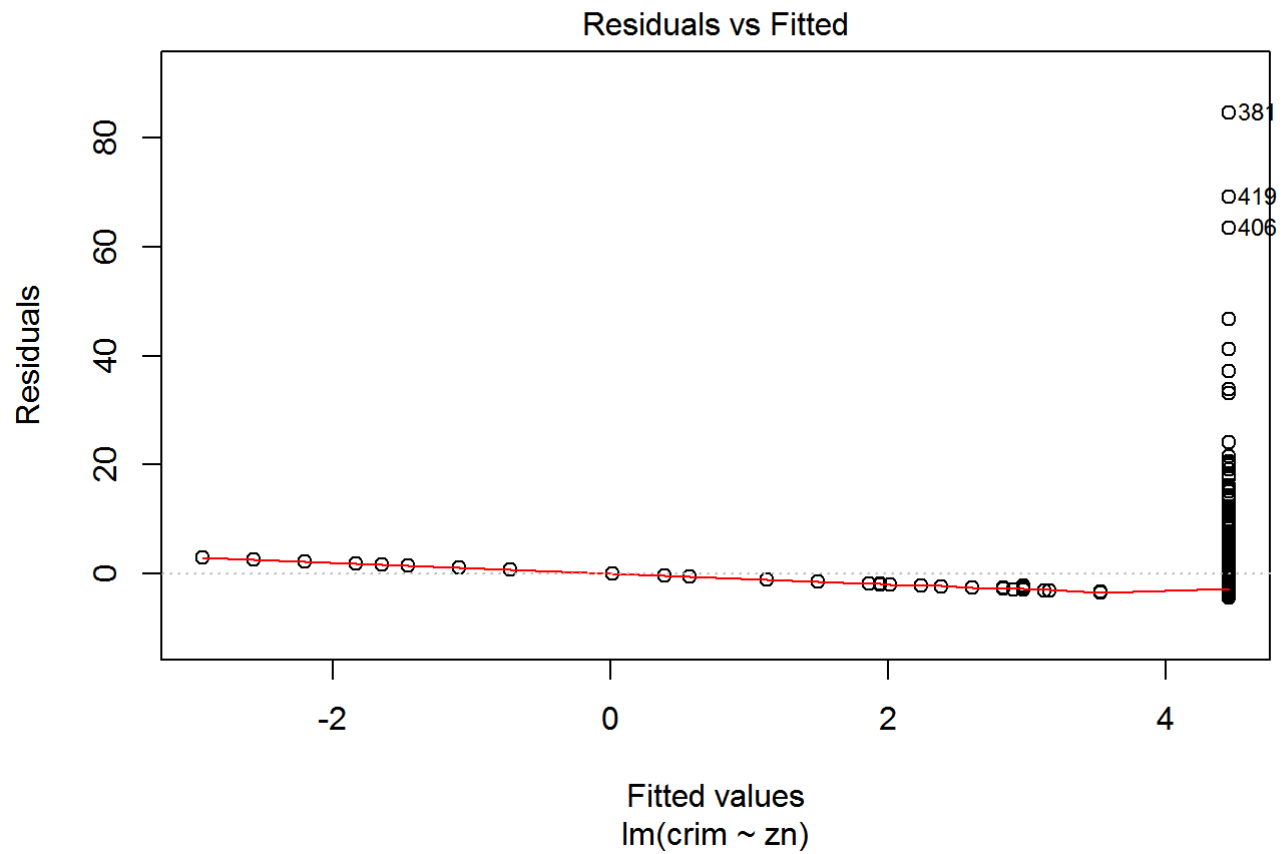
For each predictor, fit a simple linear regression model

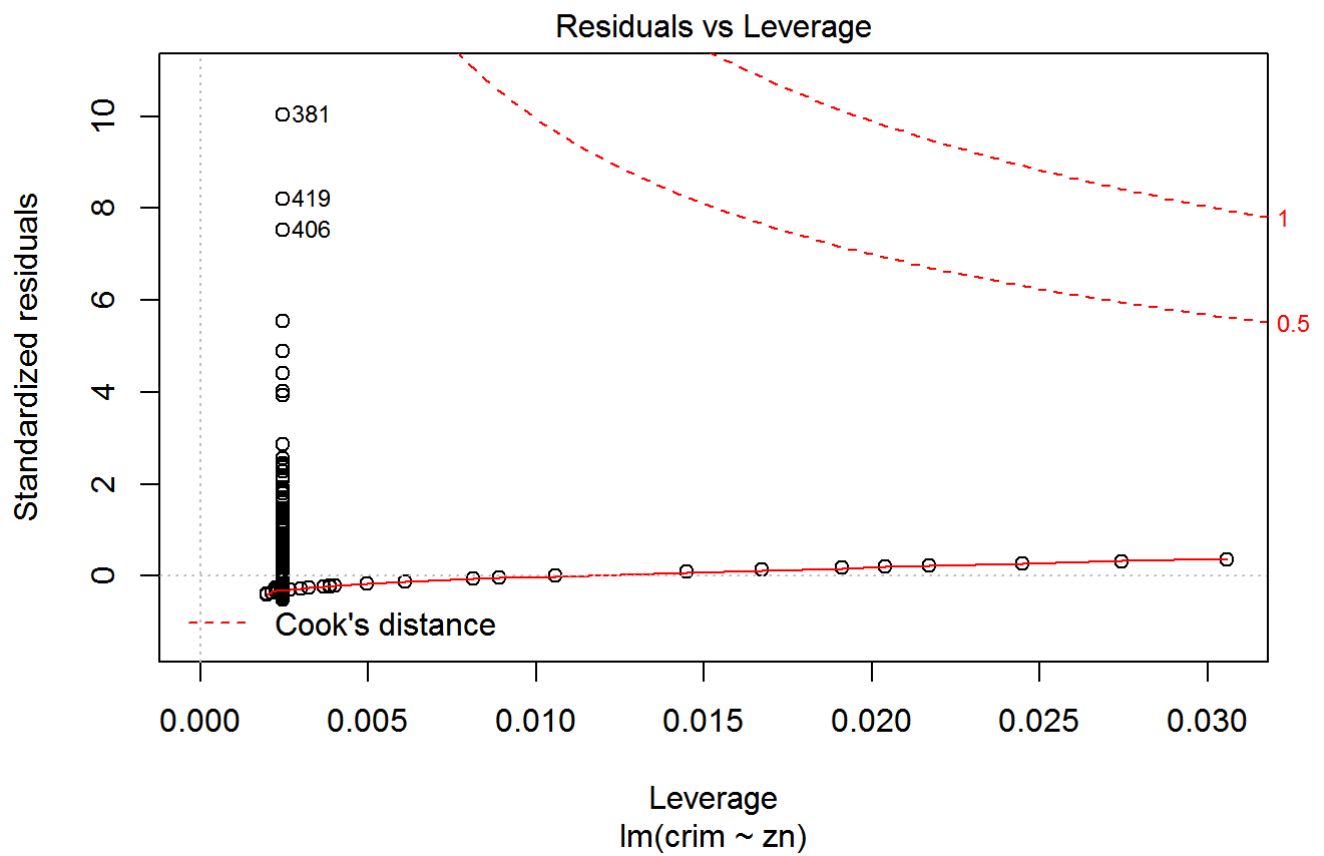
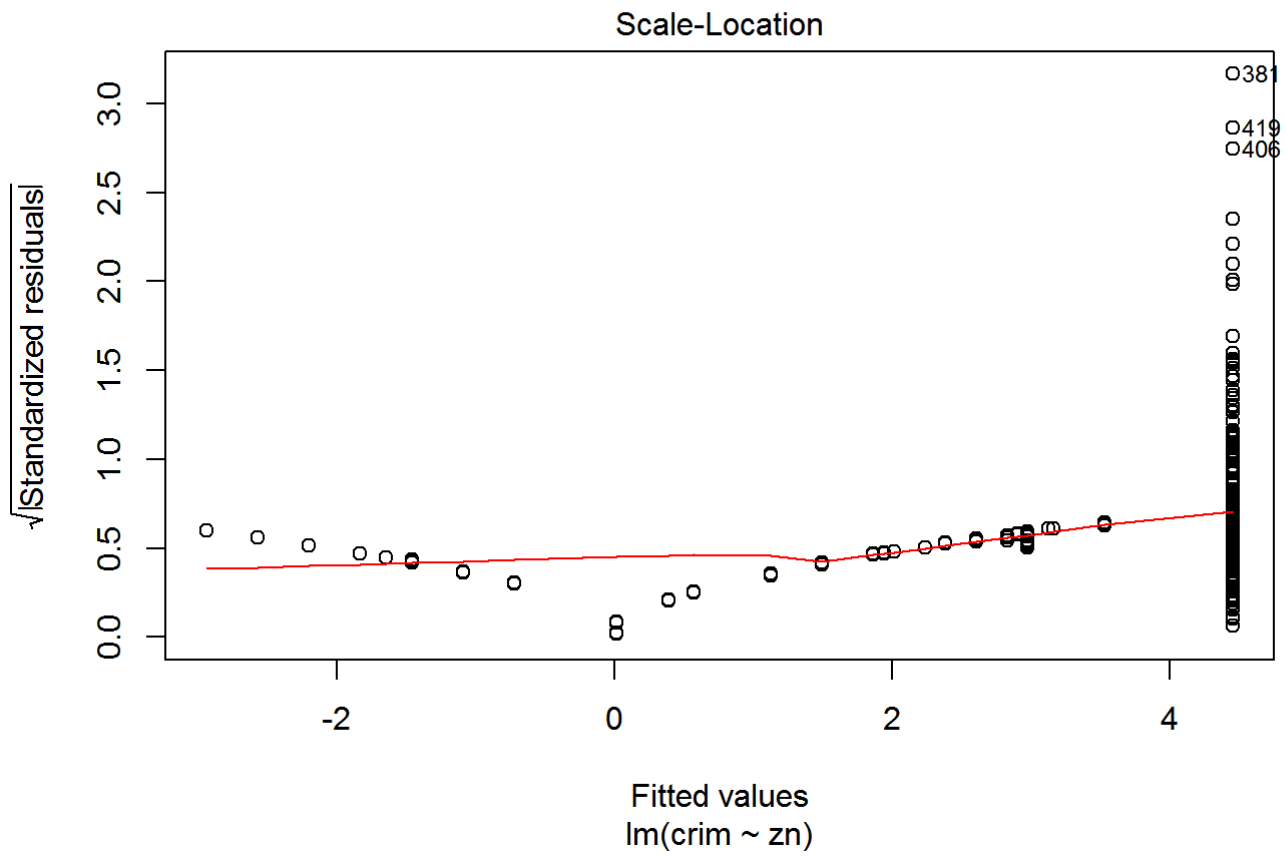
For zn

```
m1_zn<-lm(crim~zn,data = Boston)  
summary(m1_zn)
```

```
##  
## Call:  
## lm(formula = crim ~ zn, data = Boston)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -4.429 -4.222 -2.620  1.250  84.523   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)  4.45369    0.41722  10.675  < 2e-16 ***  
## zn          -0.07393    0.01609  -4.594 5.51e-06 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 8.435 on 504 degrees of freedom  
## Multiple R-squared:  0.04019,    Adjusted R-squared:  0.03828   
## F-statistic: 21.1 on 1 and 504 DF,  p-value: 5.506e-06
```

```
plot(m1_zn)
```



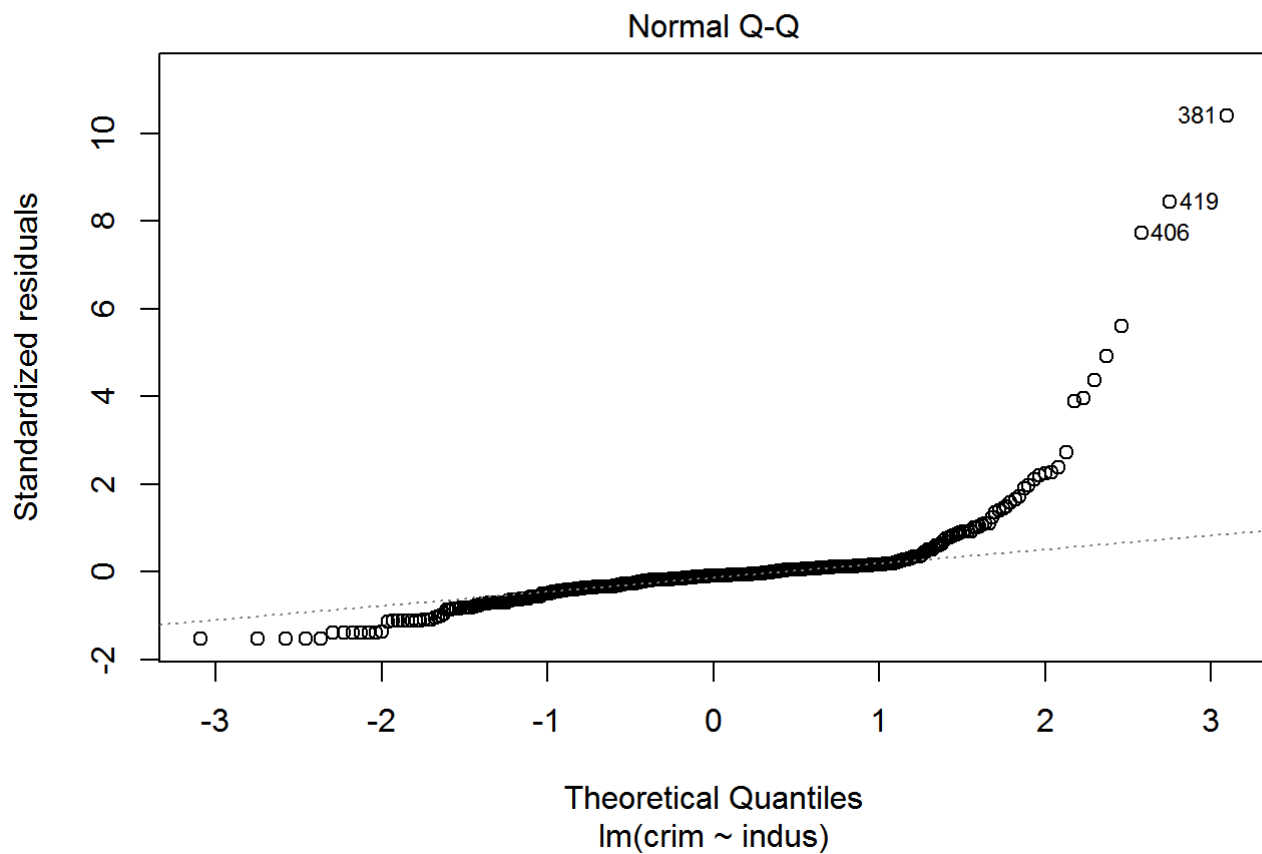
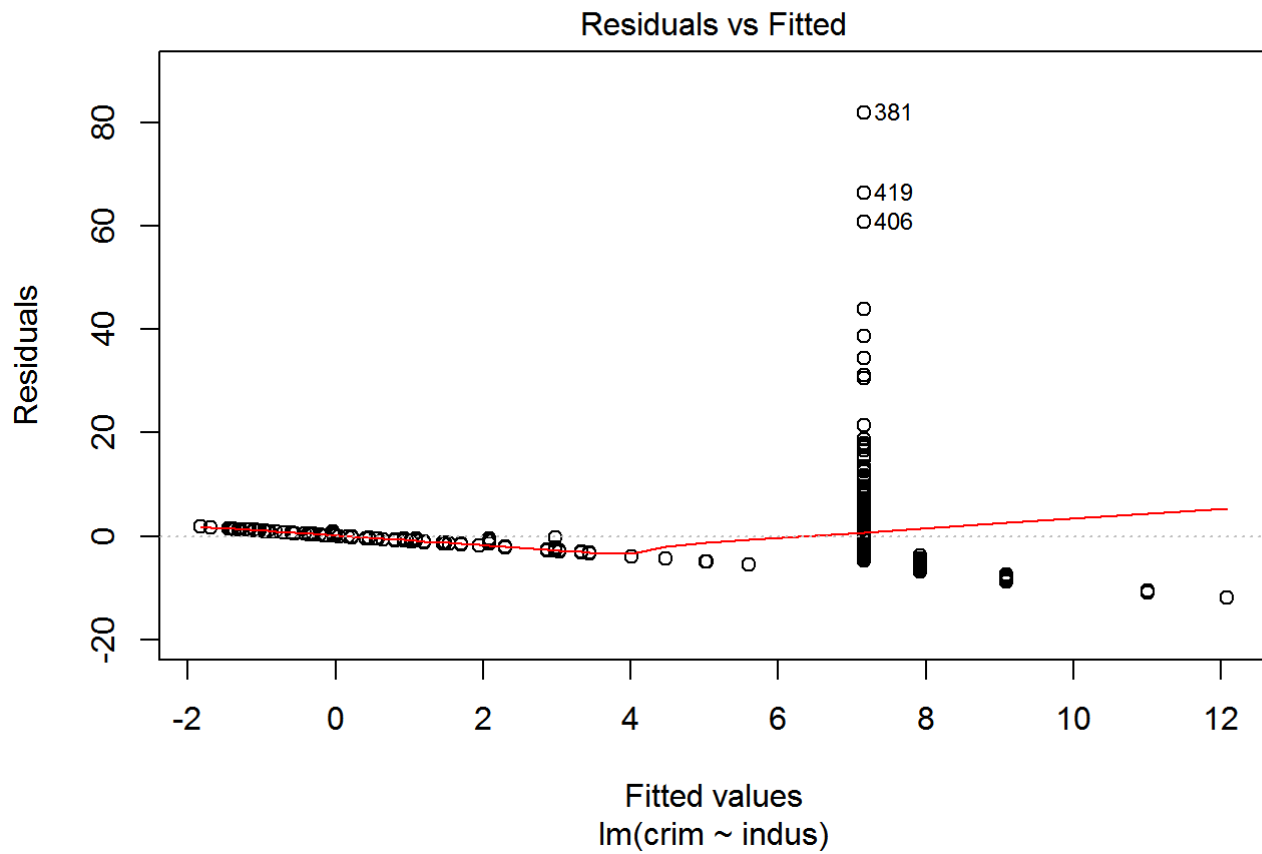
For

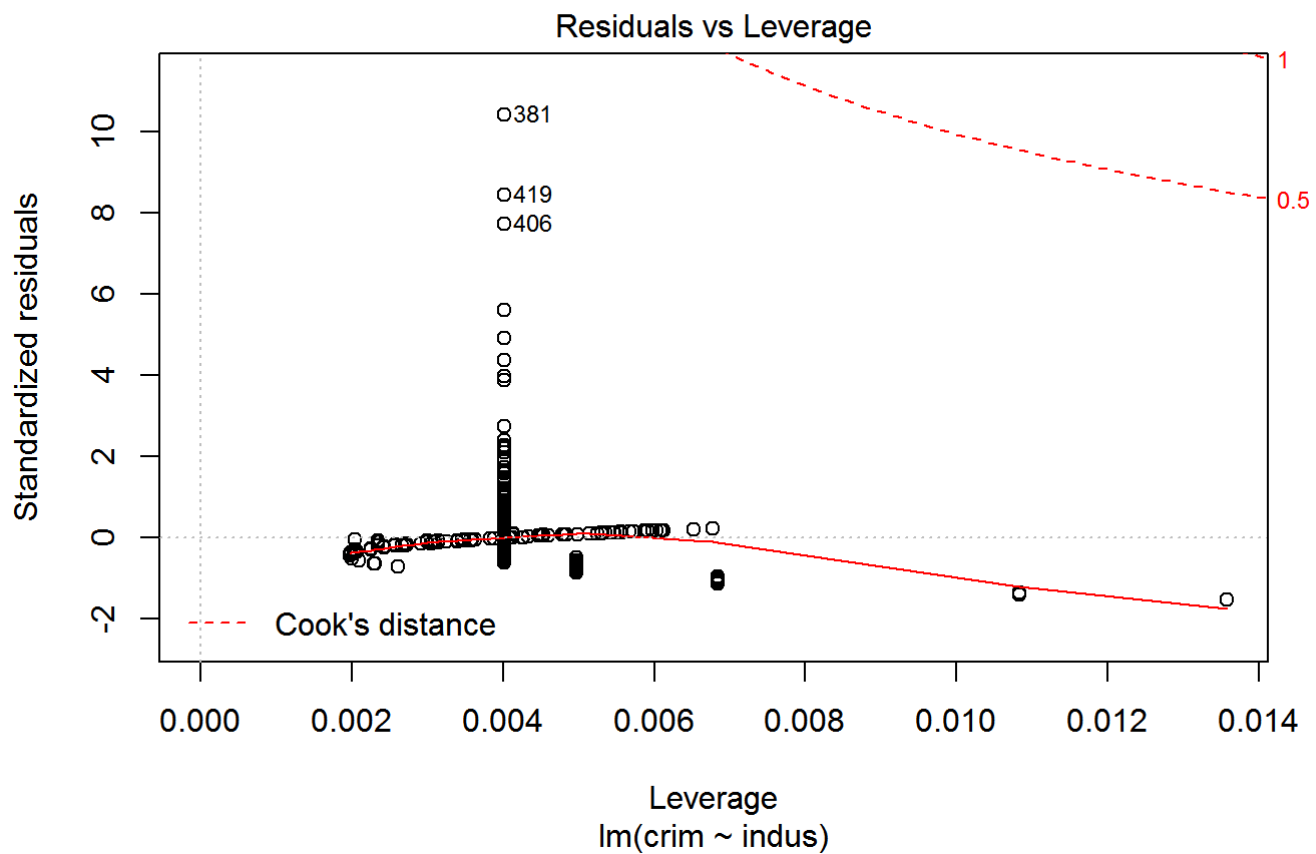
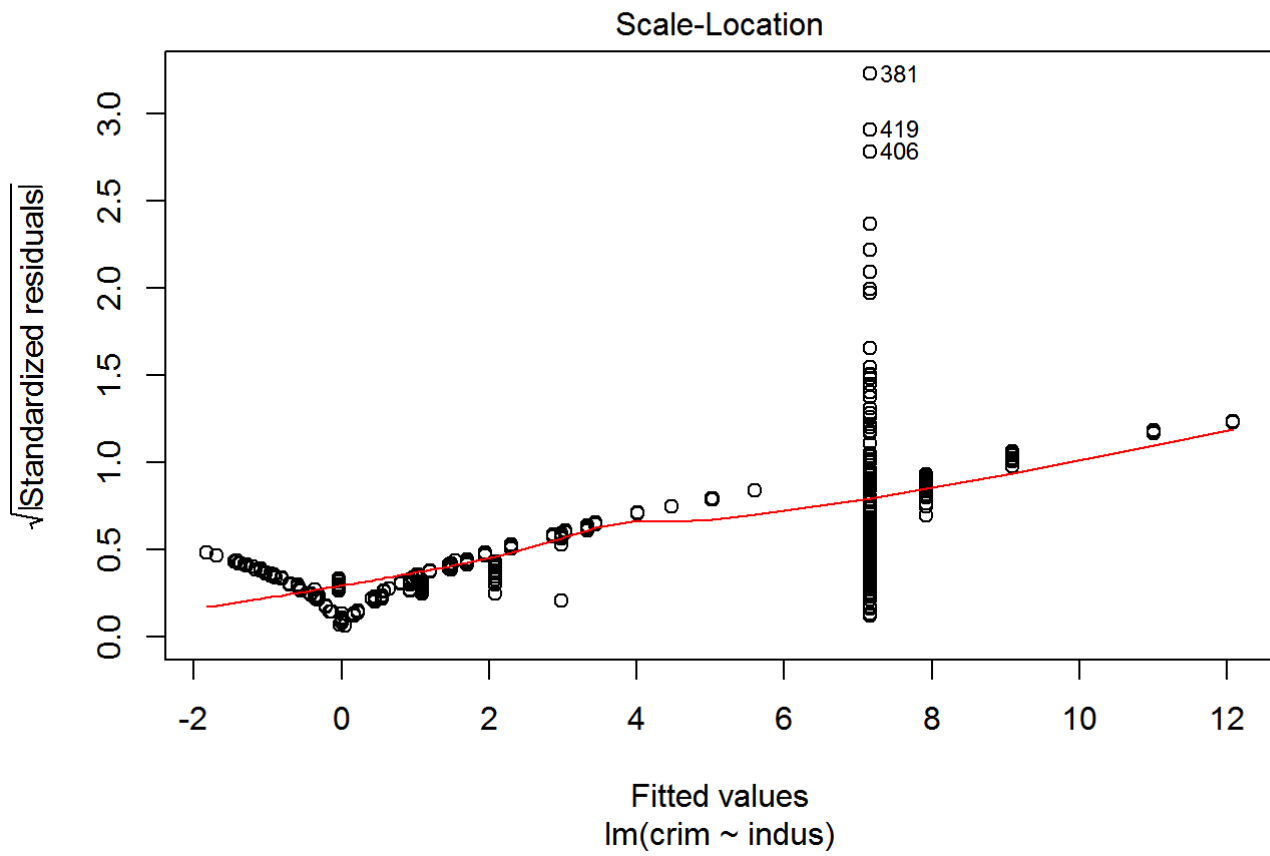
indus

```
m2_indus<-lm(crim~indus,data = Boston)
summary(m2_indus)
```

```
##
## Call:
## lm(formula = crim ~ indus, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.972  -2.698  -0.736   0.712  81.813
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.06374    0.66723  -3.093  0.00209 **
## indus        0.50978    0.05102   9.991 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.866 on 504 degrees of freedom
## Multiple R-squared:  0.1653, Adjusted R-squared:  0.1637
## F-statistic: 99.82 on 1 and 504 DF,  p-value: < 2.2e-16
```

```
plot(m2_indus)
```





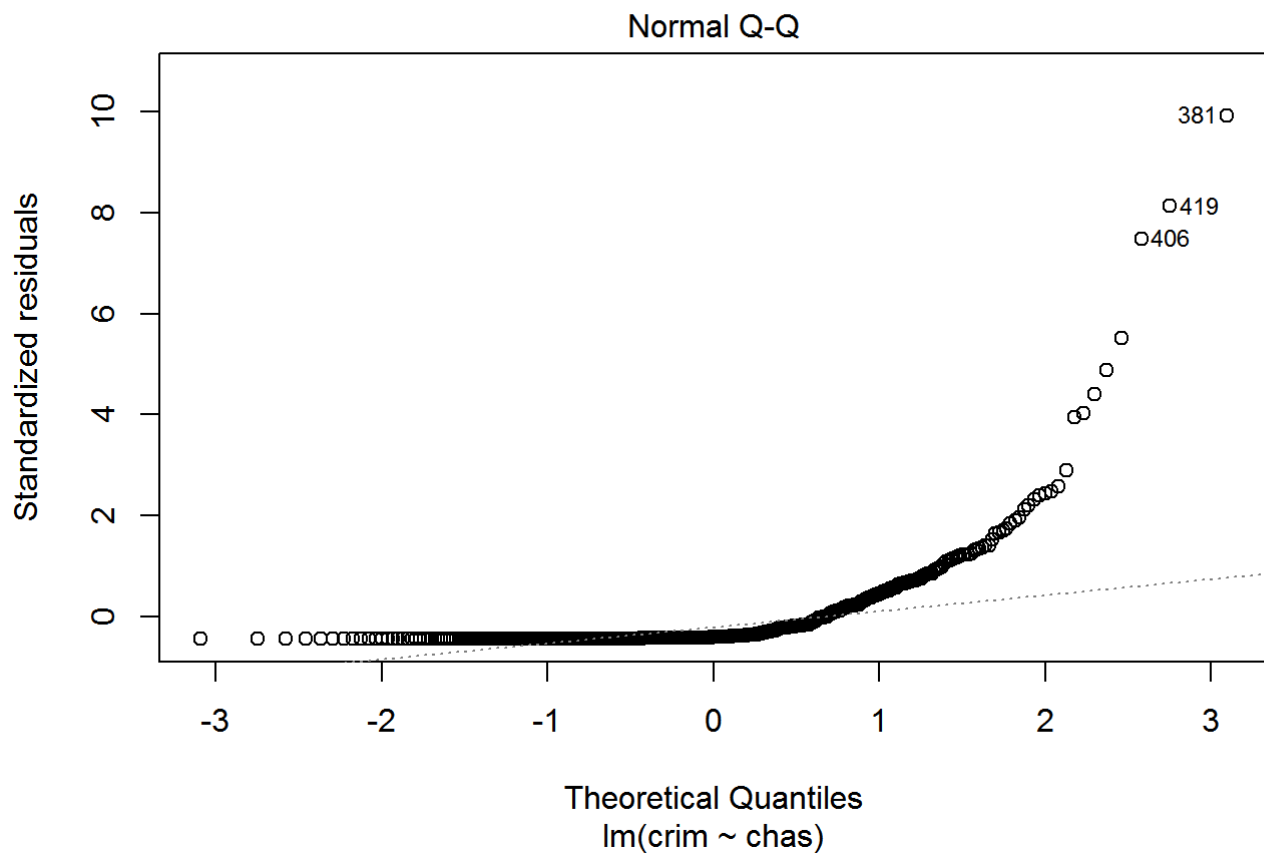
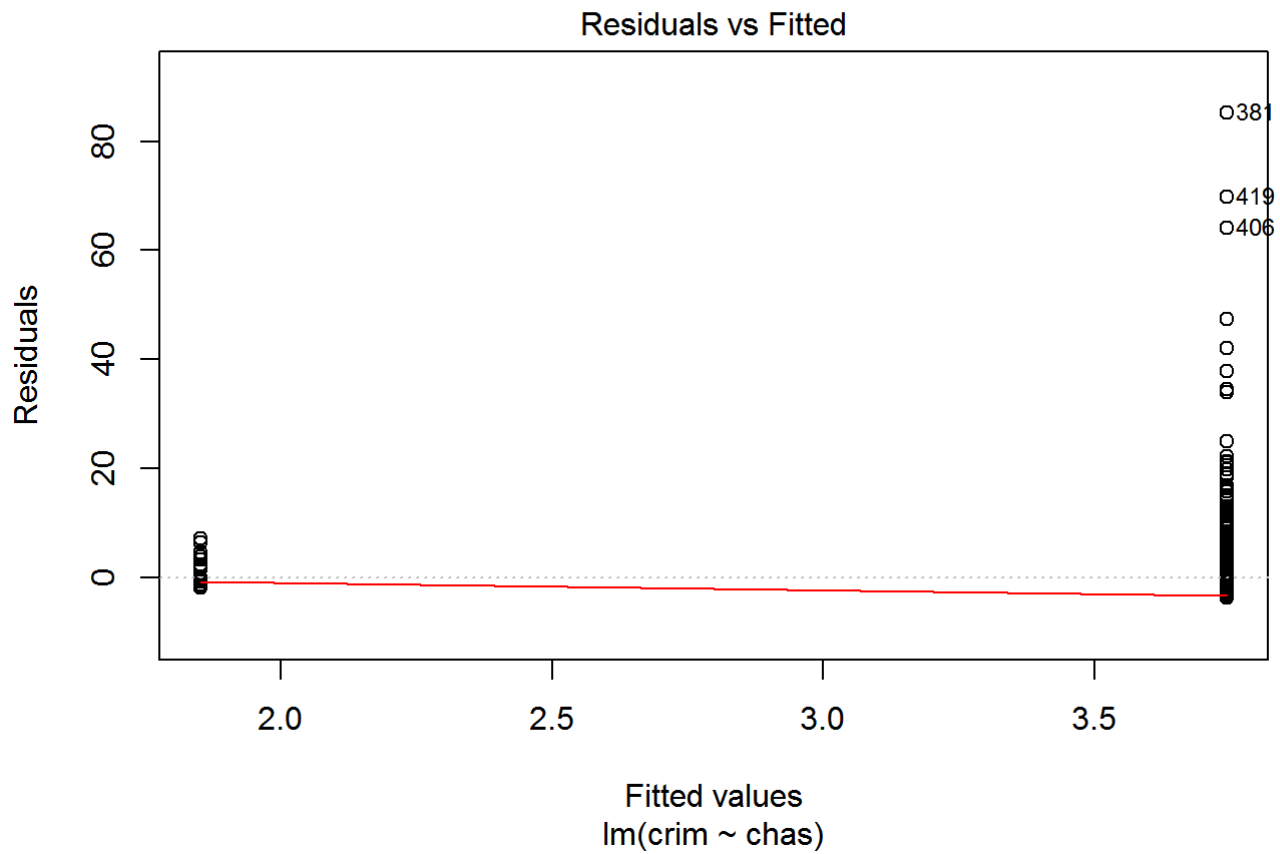
For

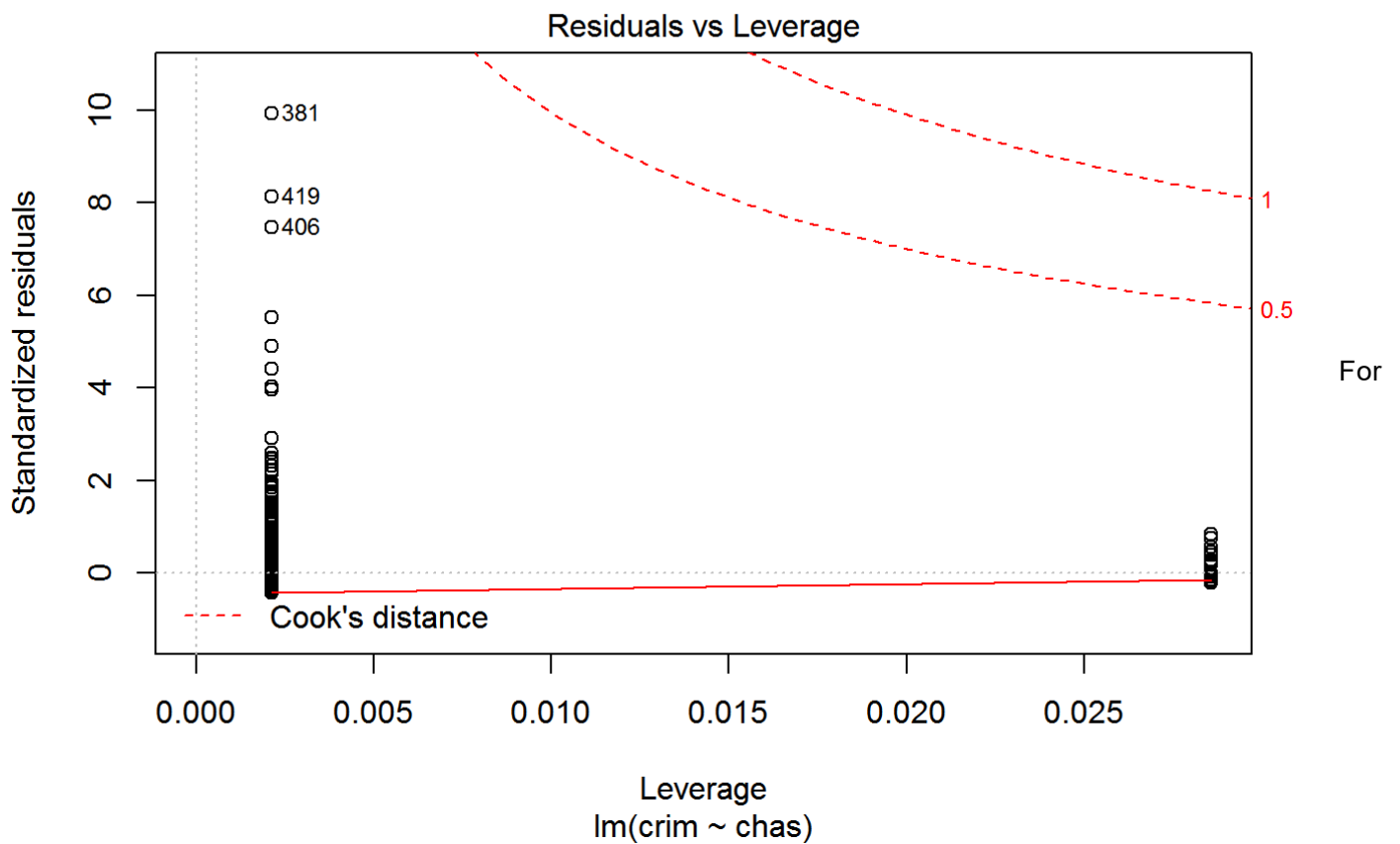
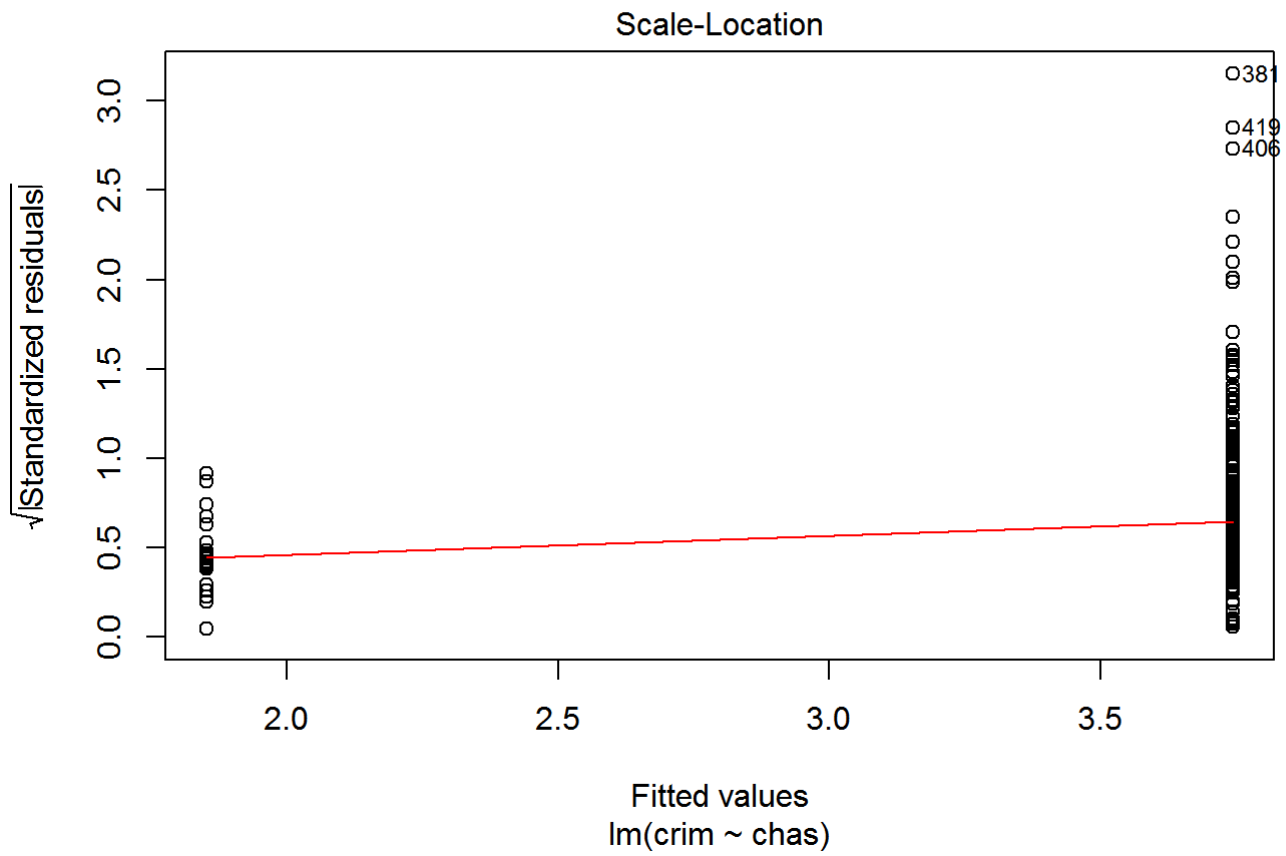
chas

```
m3_chas<-lm(crim~chas,data = Boston)
summary(m3_chas)
```

```
##
## Call:
## lm(formula = crim ~ chas, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.738 -3.661 -3.435  0.018 85.232
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.7444     0.3961   9.453  <2e-16 ***
## chas         -1.8928     1.5061  -1.257   0.209
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.597 on 504 degrees of freedom
## Multiple R-squared:  0.003124,    Adjusted R-squared:  0.001146
## F-statistic: 1.579 on 1 and 504 DF,  p-value: 0.2094
```

```
plot(m3_chas)
```

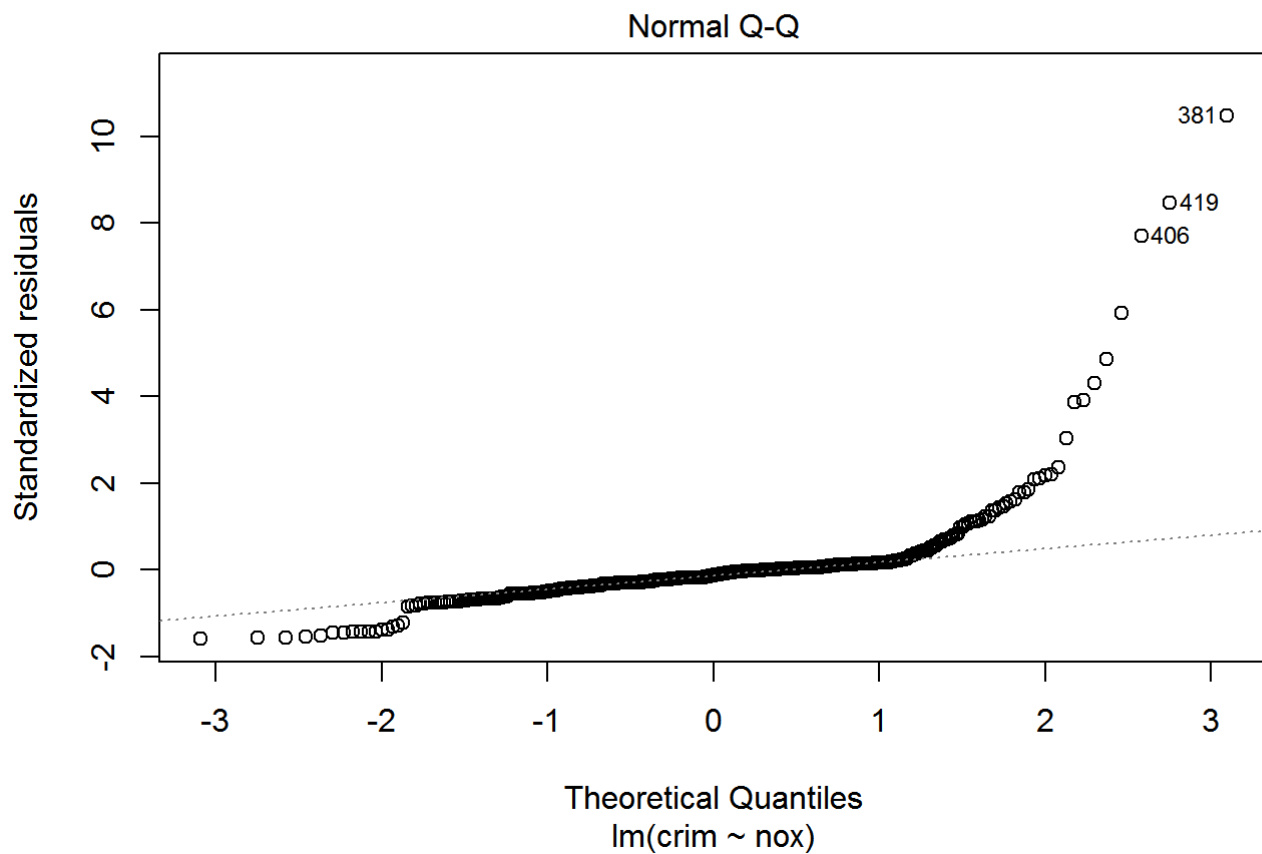
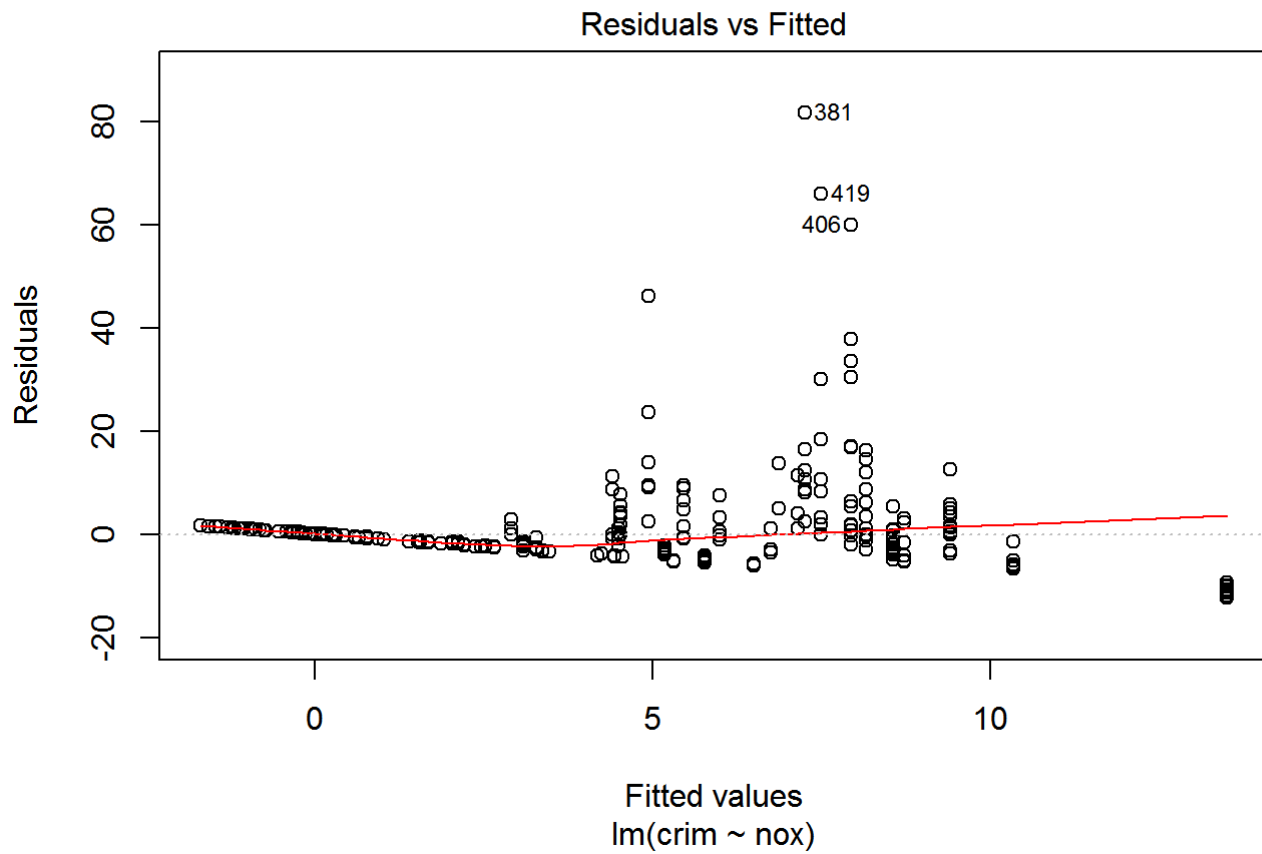


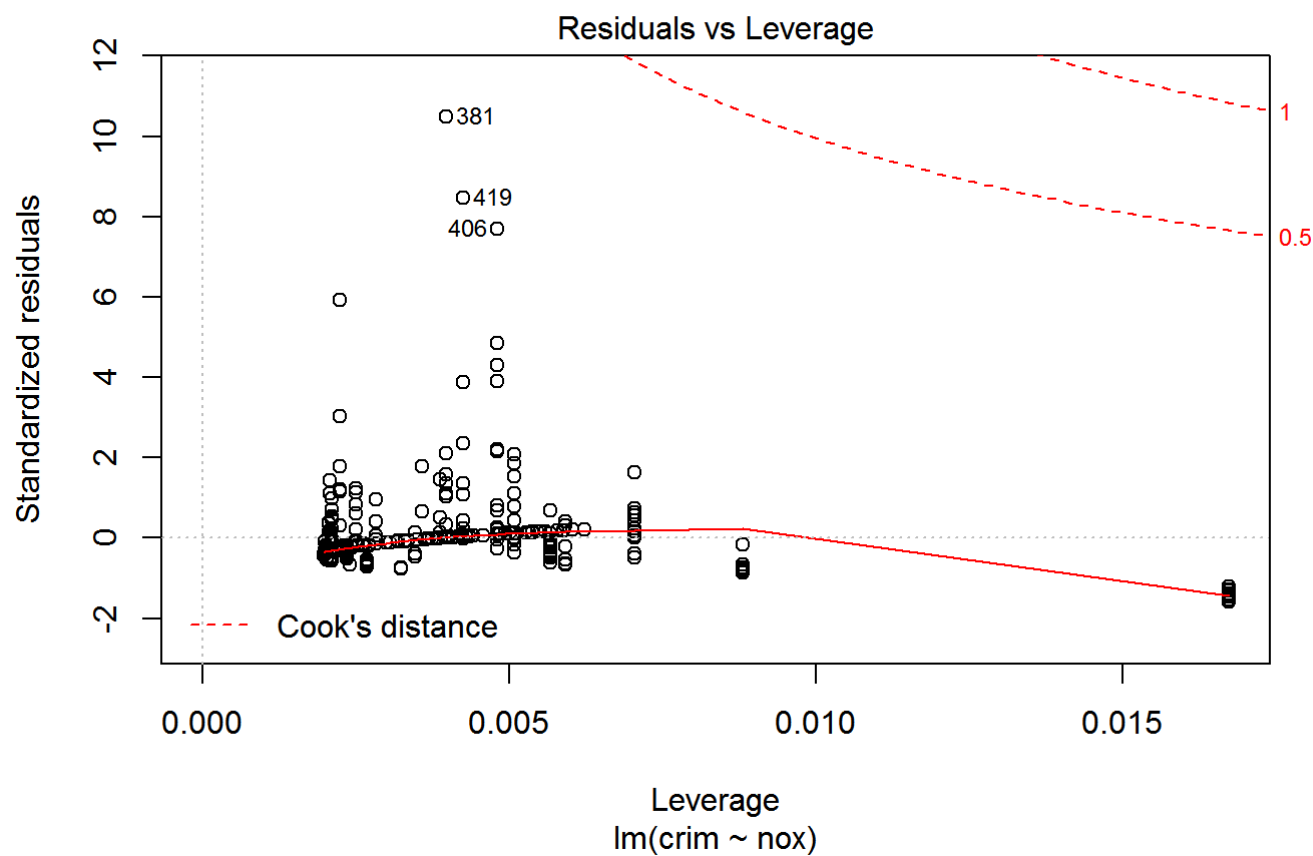
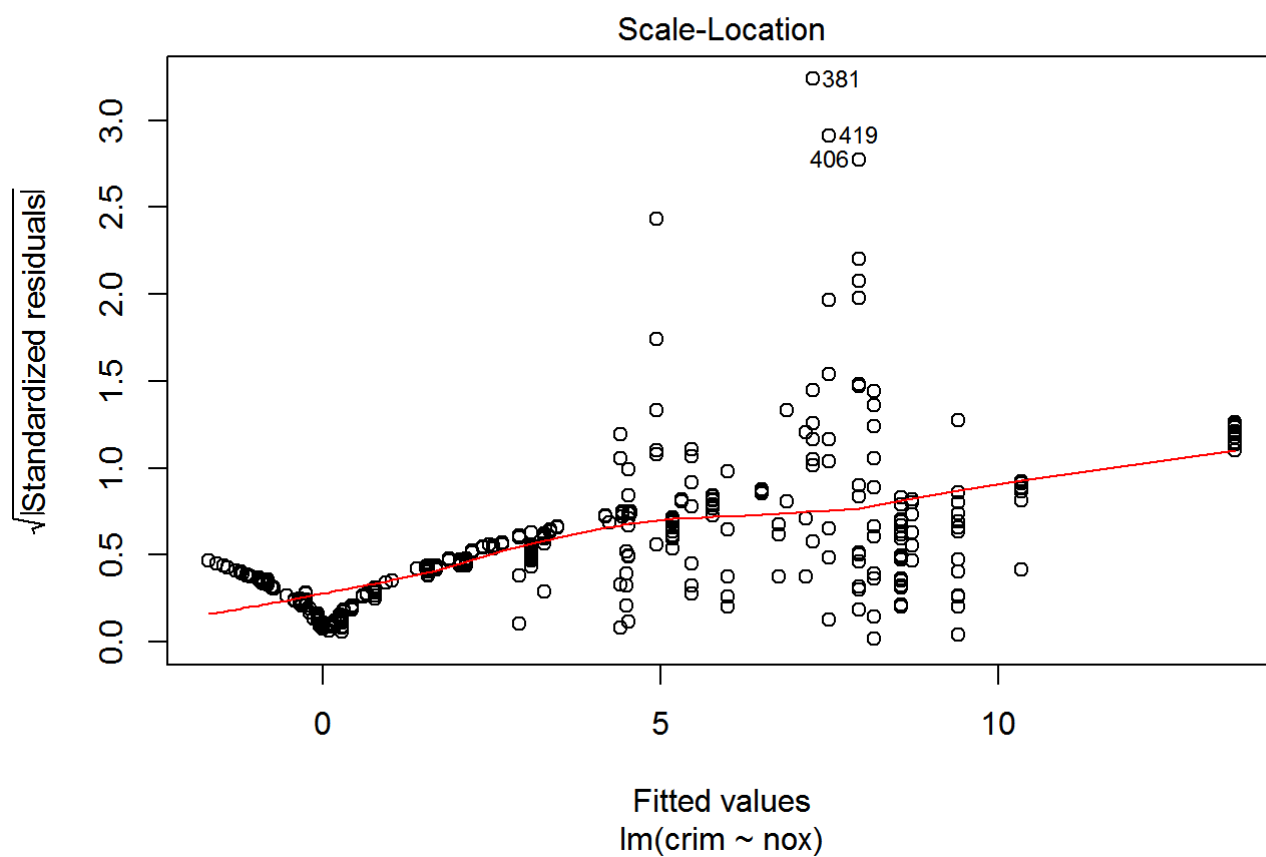
nox

```
m4_nox<-lm(crim~nox,data = Boston)
summary(m4_nox)
```

```
##
## Call:
## lm(formula = crim ~ nox, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.371  -2.738  -0.974   0.559   81.728
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -13.720      1.699   -8.073 5.08e-15 ***
## nox           31.249      2.999  10.419 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.81 on 504 degrees of freedom
## Multiple R-squared:  0.1772, Adjusted R-squared:  0.1756
## F-statistic: 108.6 on 1 and 504 DF,  p-value: < 2.2e-16
```

```
plot(m4_nox)
```





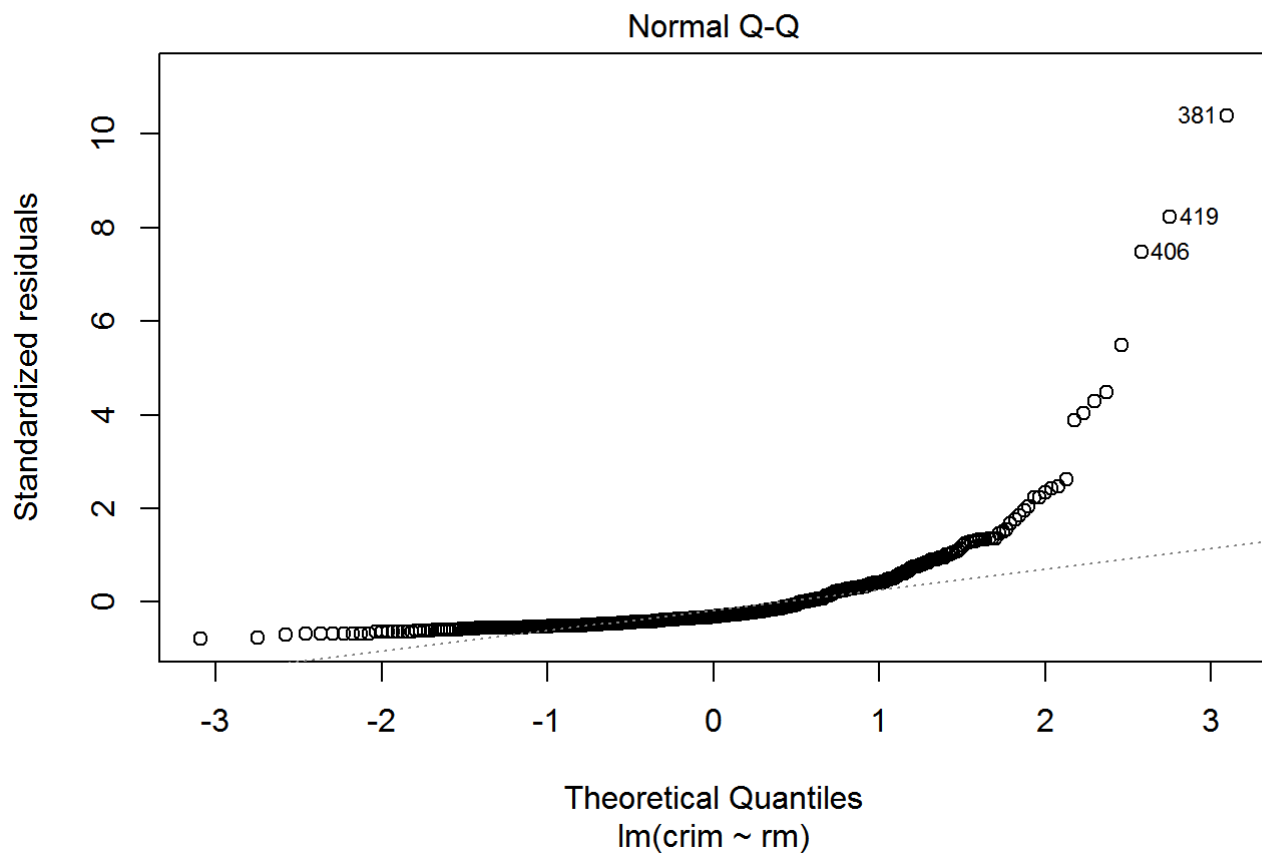
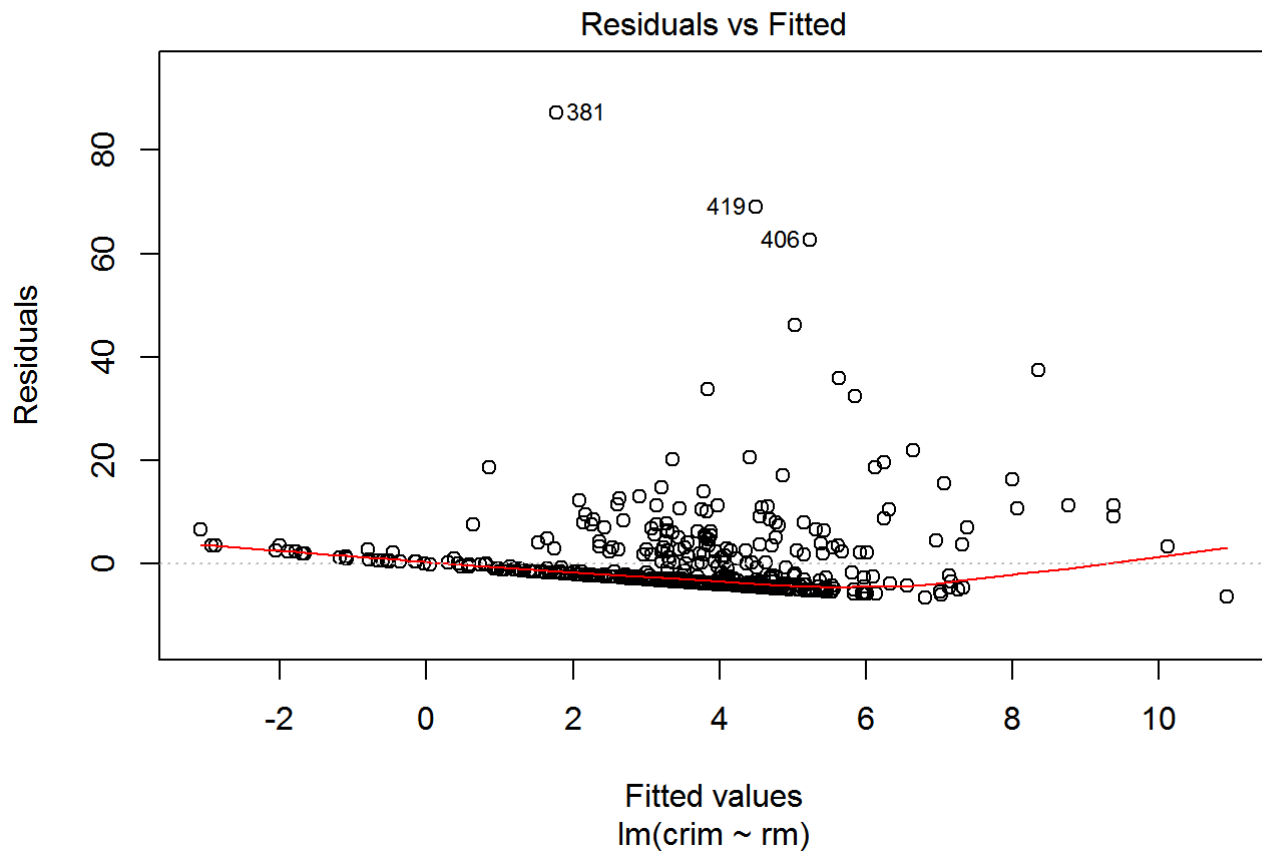
For

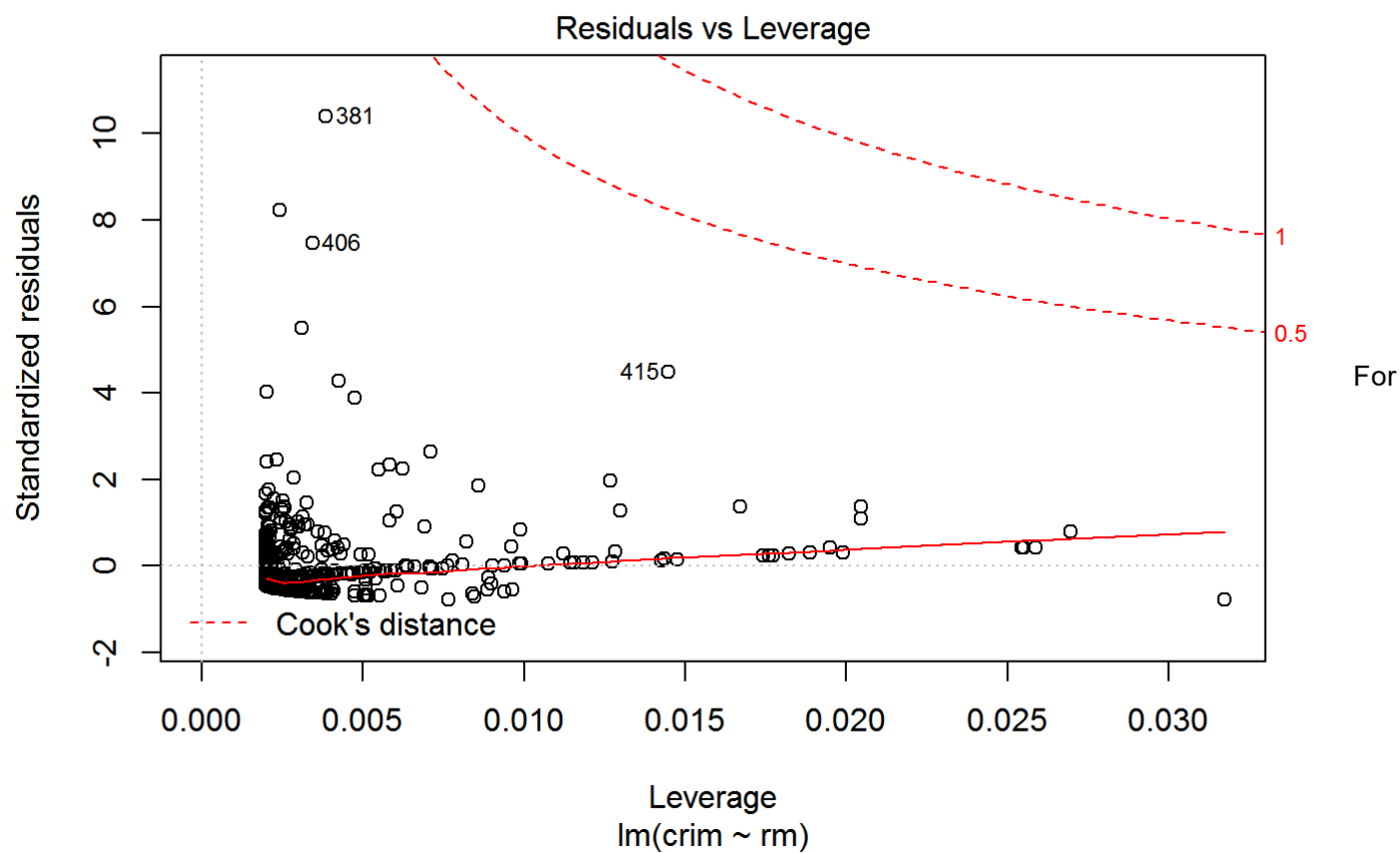
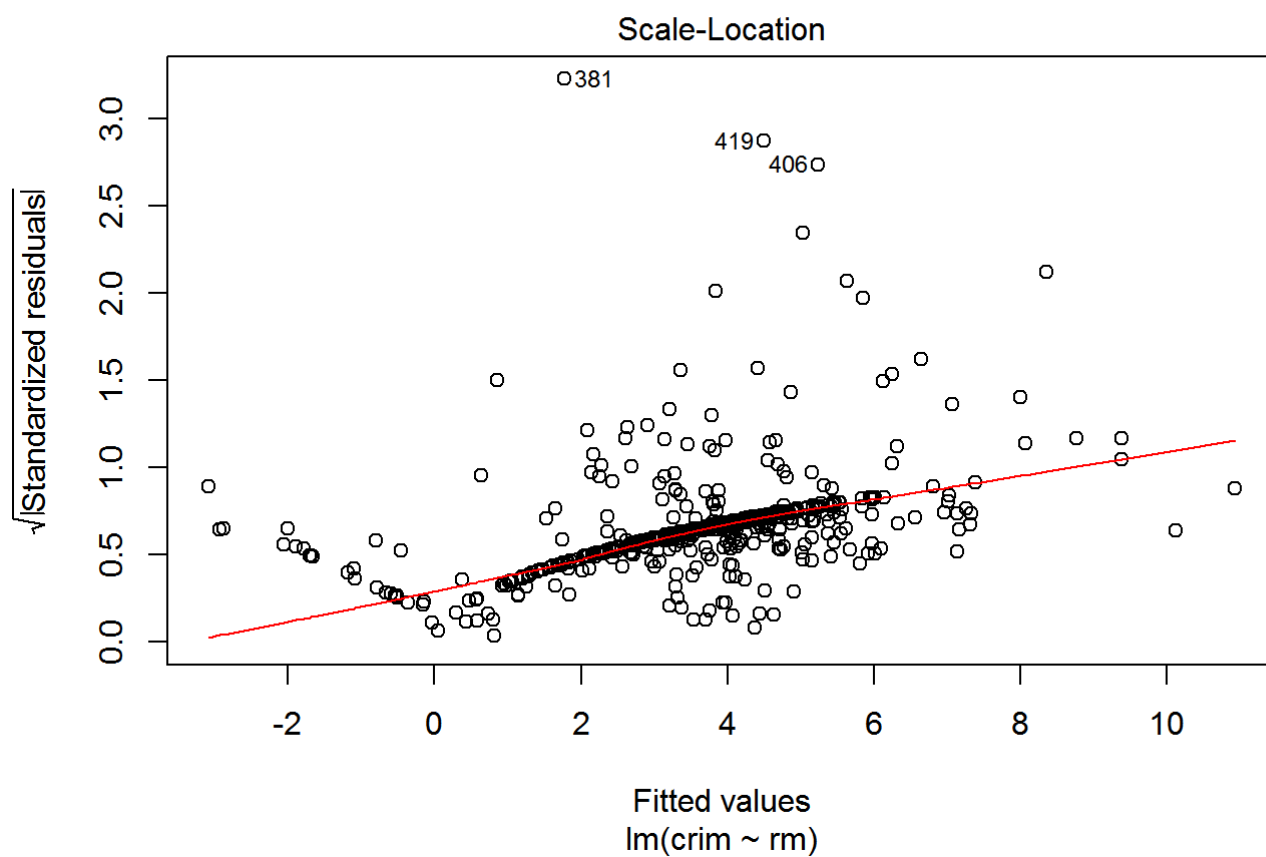
rm

```
m5_rm<-lm(crim~rm,data = Boston)
summary(m5_rm)
```

```
##
## Call:
## lm(formula = crim ~ rm, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.604  -3.952  -2.654   0.989  87.197
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   20.482     3.365   6.088 2.27e-09 ***
## rm           -2.684     0.532  -5.045 6.35e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.401 on 504 degrees of freedom
## Multiple R-squared:  0.04807,    Adjusted R-squared:  0.04618
## F-statistic: 25.45 on 1 and 504 DF,  p-value: 6.347e-07
```

```
plot(m5_rm)
```



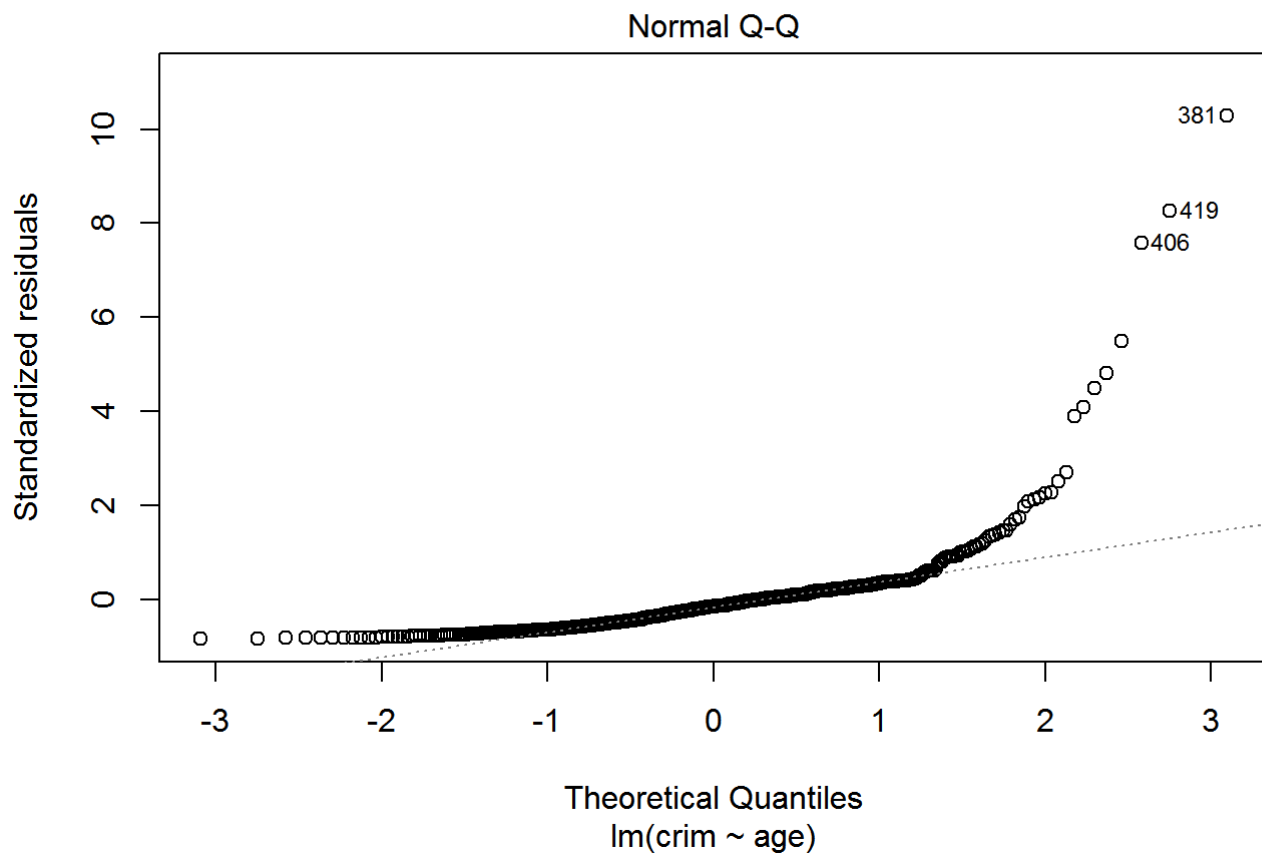
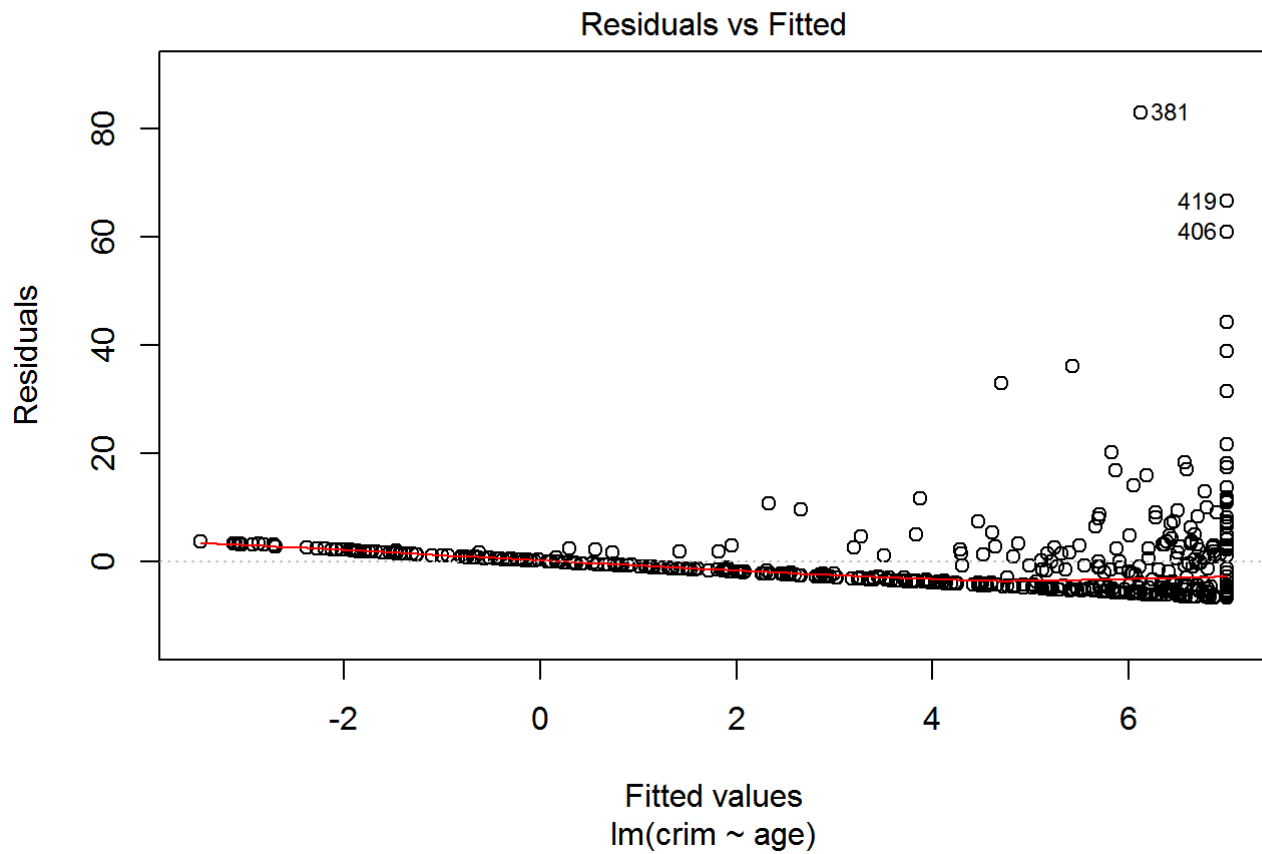


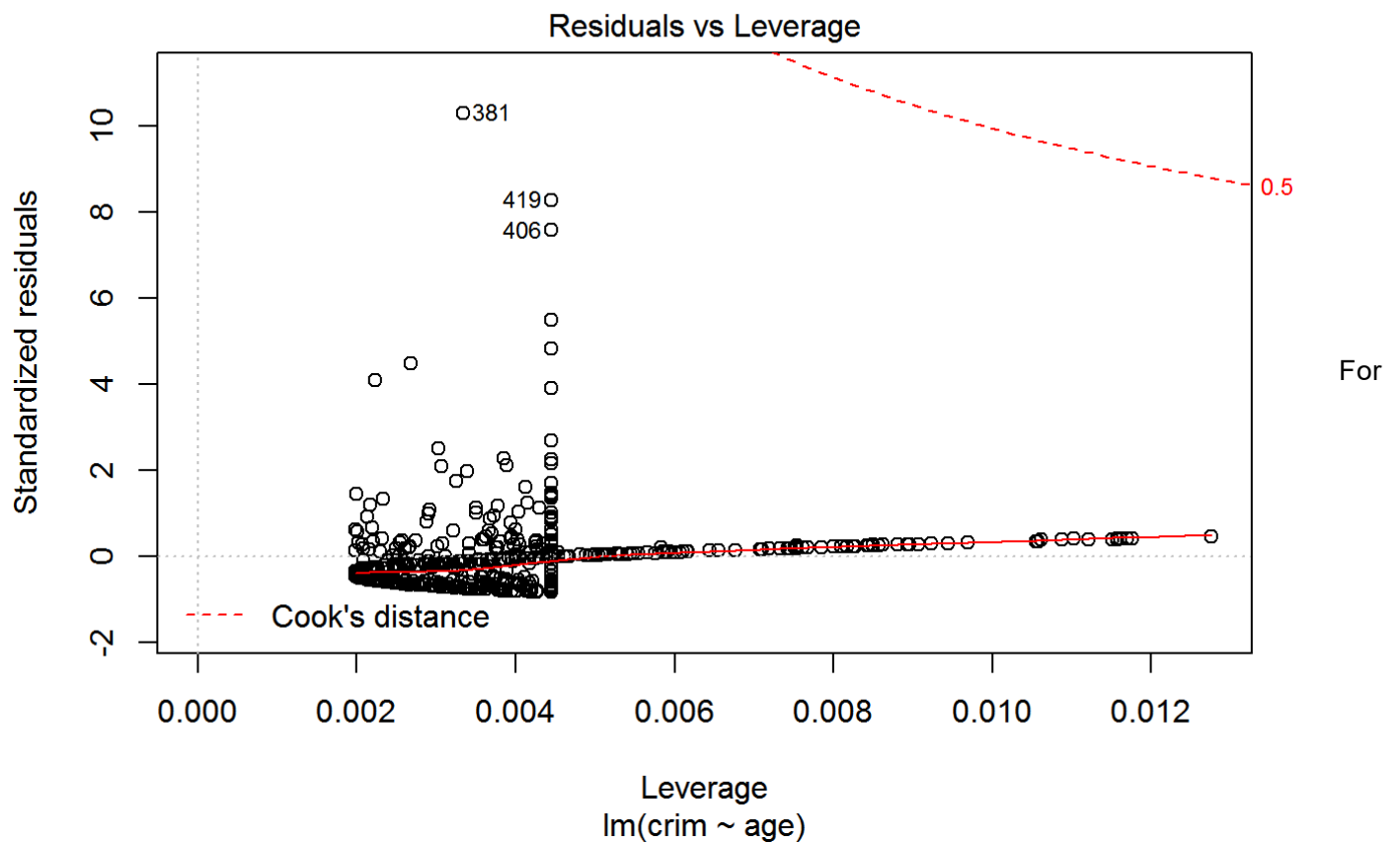
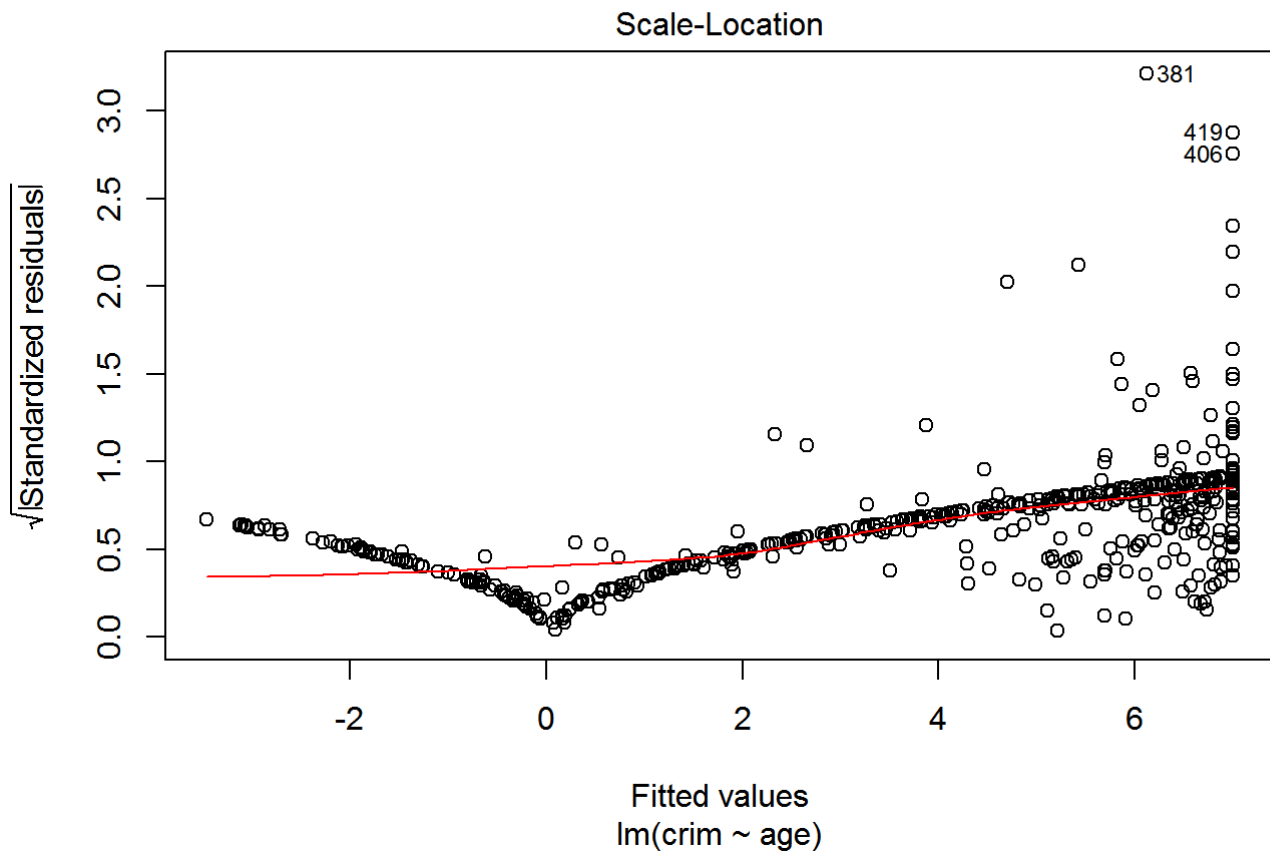
age


```
m6_age<-lm(crim~age,data = Boston)
summary(m6_age)
```

```
##
## Call:
## lm(formula = crim ~ age, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.789  -4.257  -1.230   1.527  82.849
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.77791    0.94398  -4.002 7.22e-05 ***
## age          0.10779    0.01274   8.463 2.85e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.057 on 504 degrees of freedom
## Multiple R-squared:  0.1244, Adjusted R-squared:  0.1227
## F-statistic: 71.62 on 1 and 504 DF,  p-value: 2.855e-16
```

```
plot(m6_age)
```



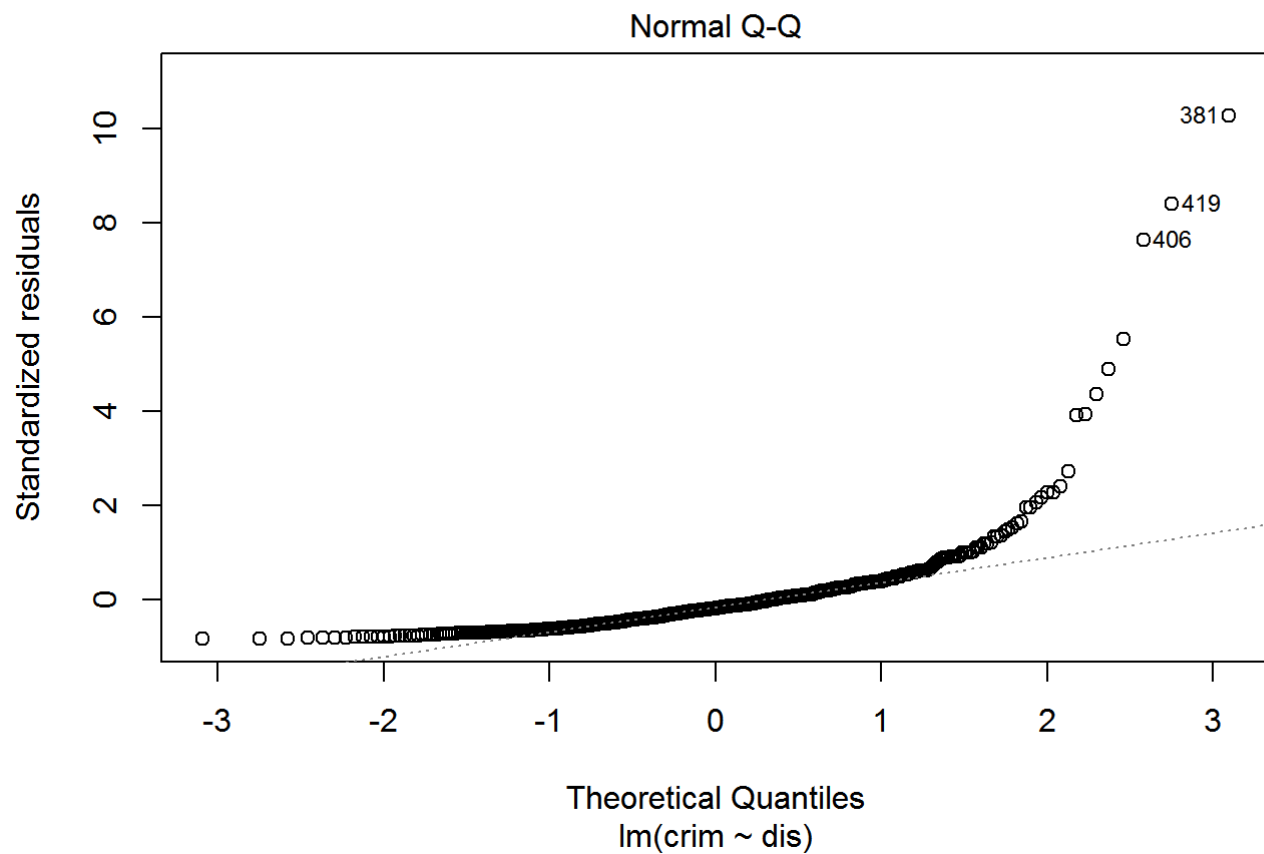
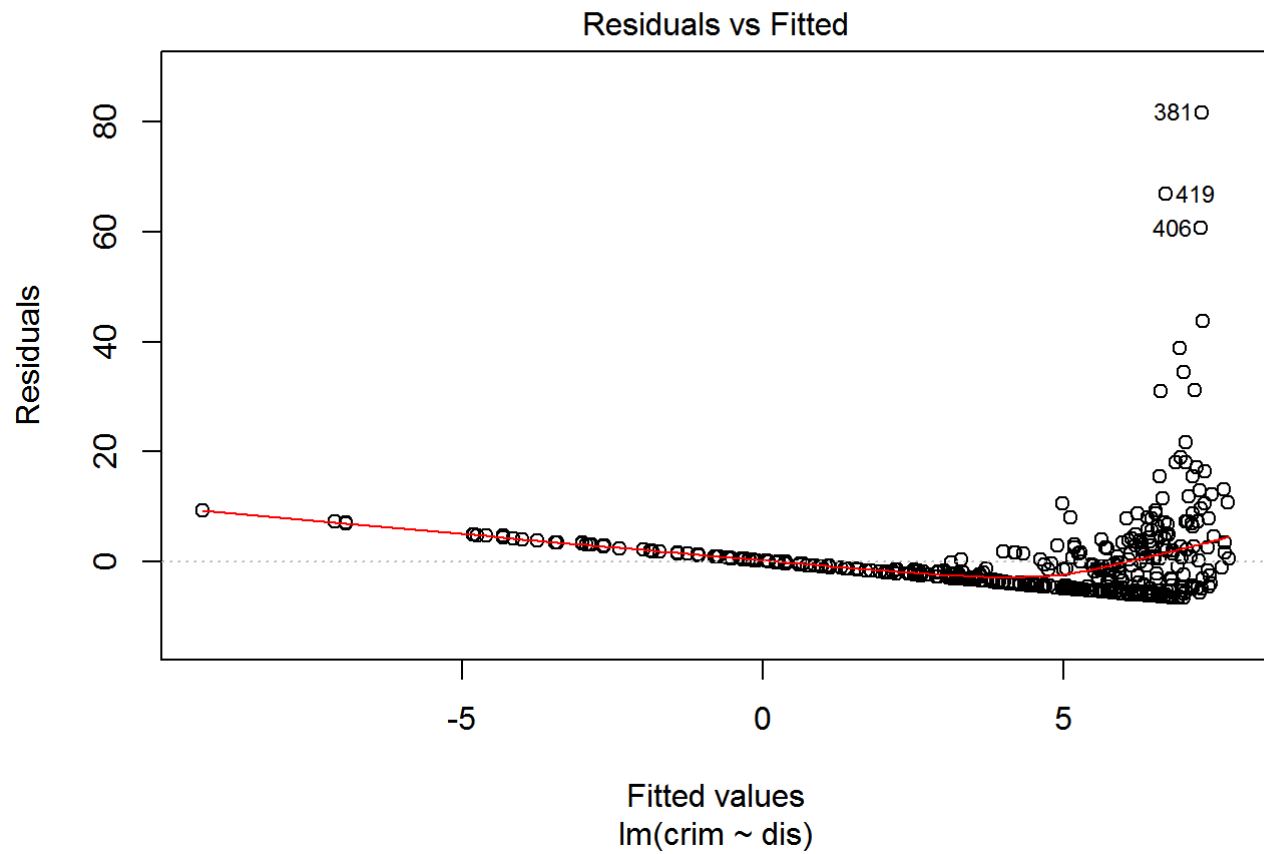


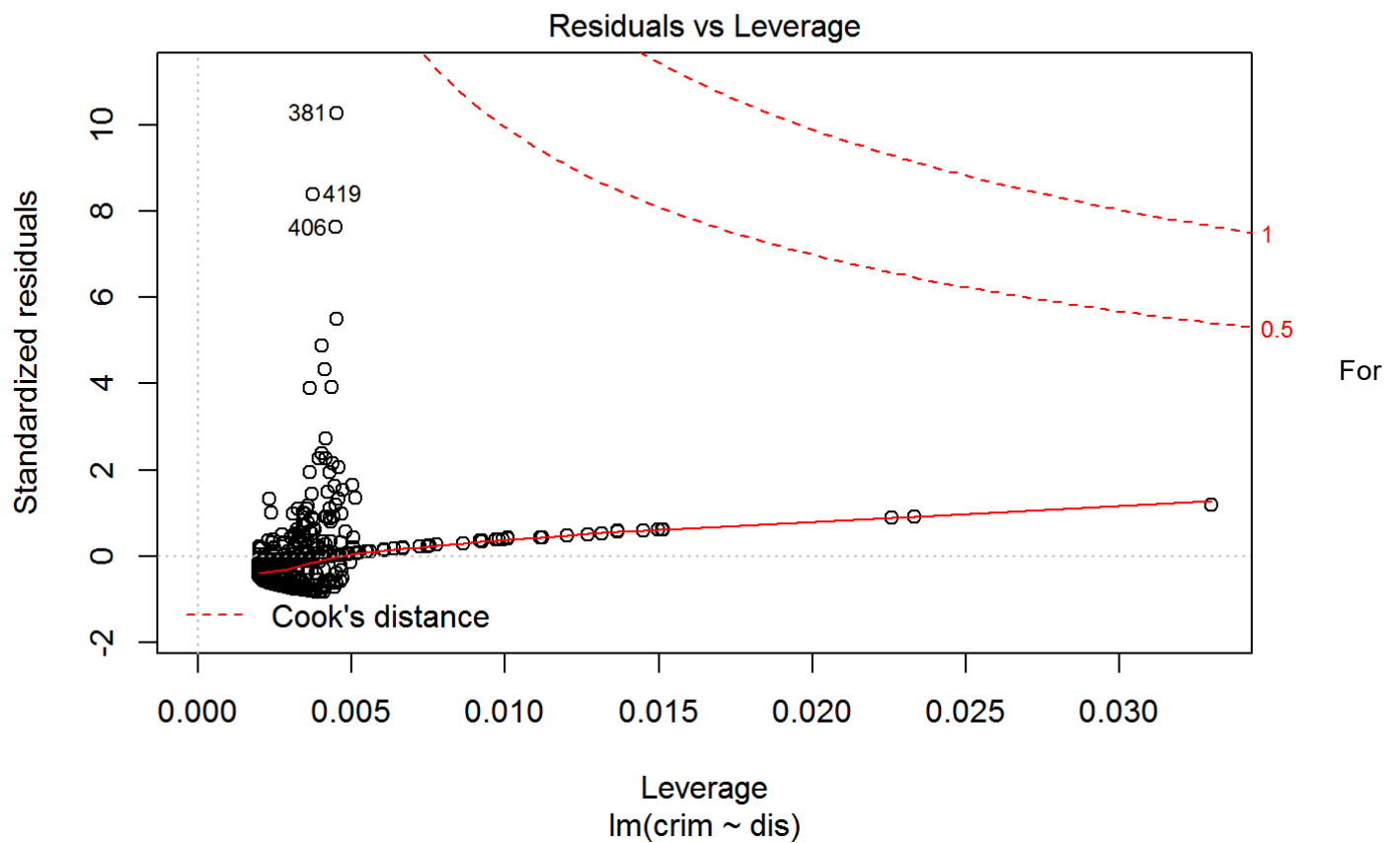
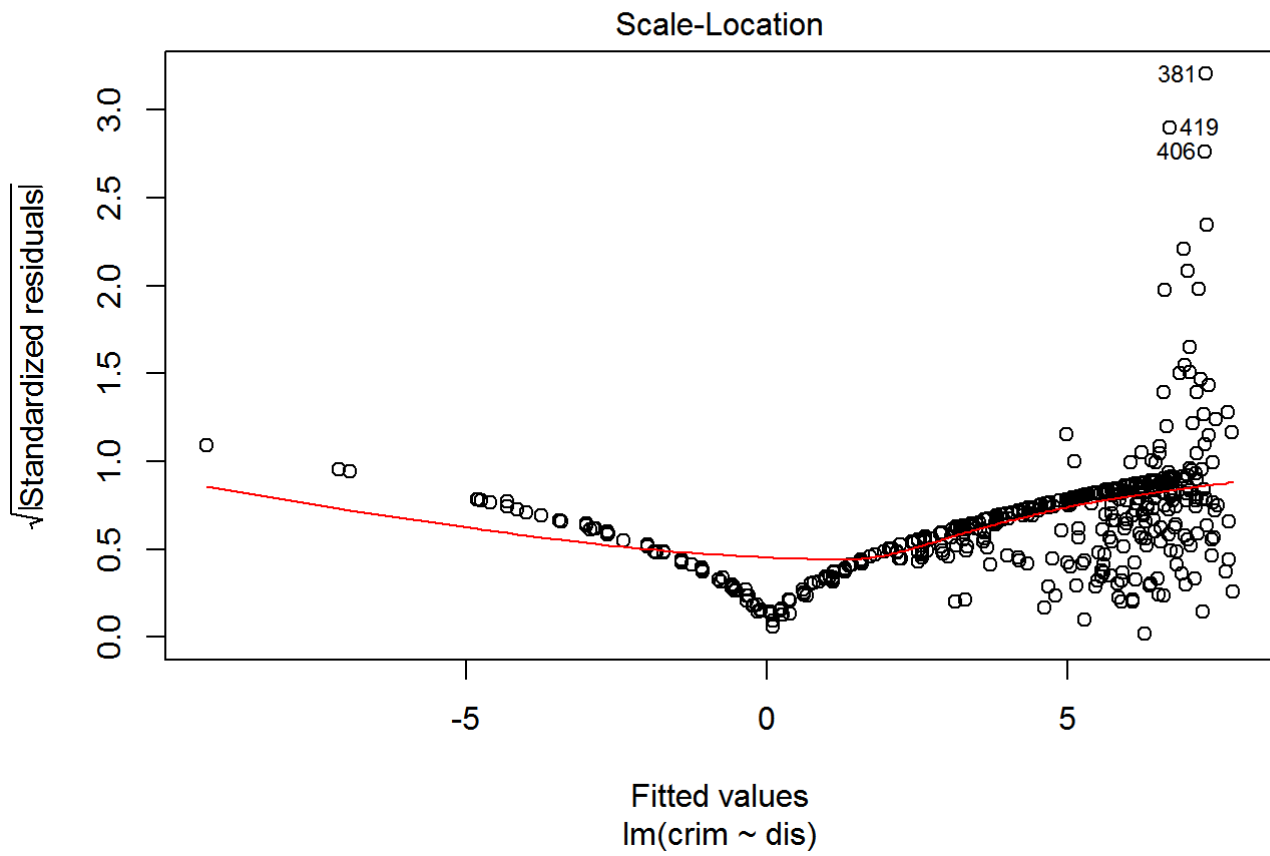
dis

```
m7_dis<-lm(crim~dis,data = Boston)
summary(m7_dis)
```

```
##
## Call:
## lm(formula = crim ~ dis, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.708  -4.134  -1.527   1.516  81.674
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.4993     0.7304   13.006  <2e-16 ***
## dis          -1.5509     0.1683   -9.213  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.965 on 504 degrees of freedom
## Multiple R-squared:  0.1441, Adjusted R-squared:  0.1425
## F-statistic: 84.89 on 1 and 504 DF,  p-value: < 2.2e-16
```

```
plot(m7_dis)
```



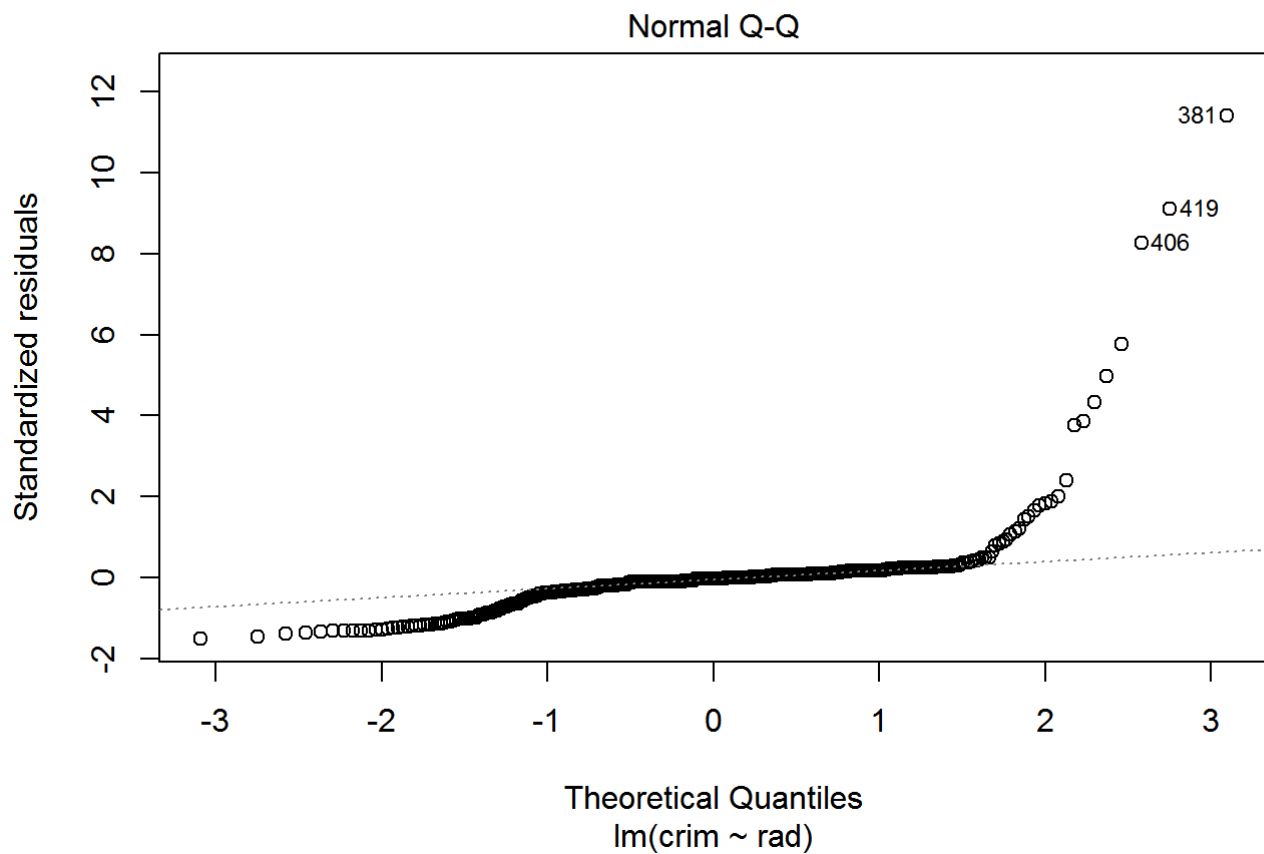
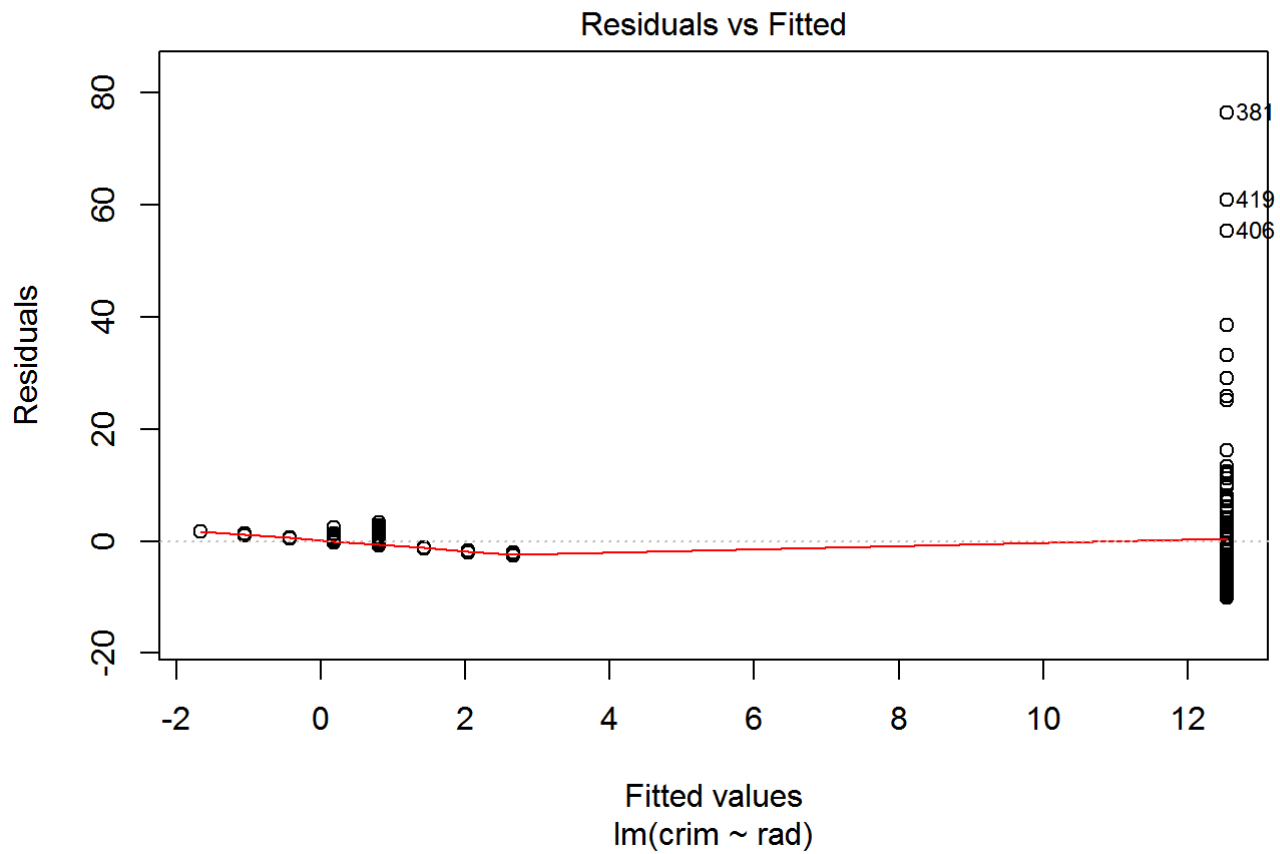


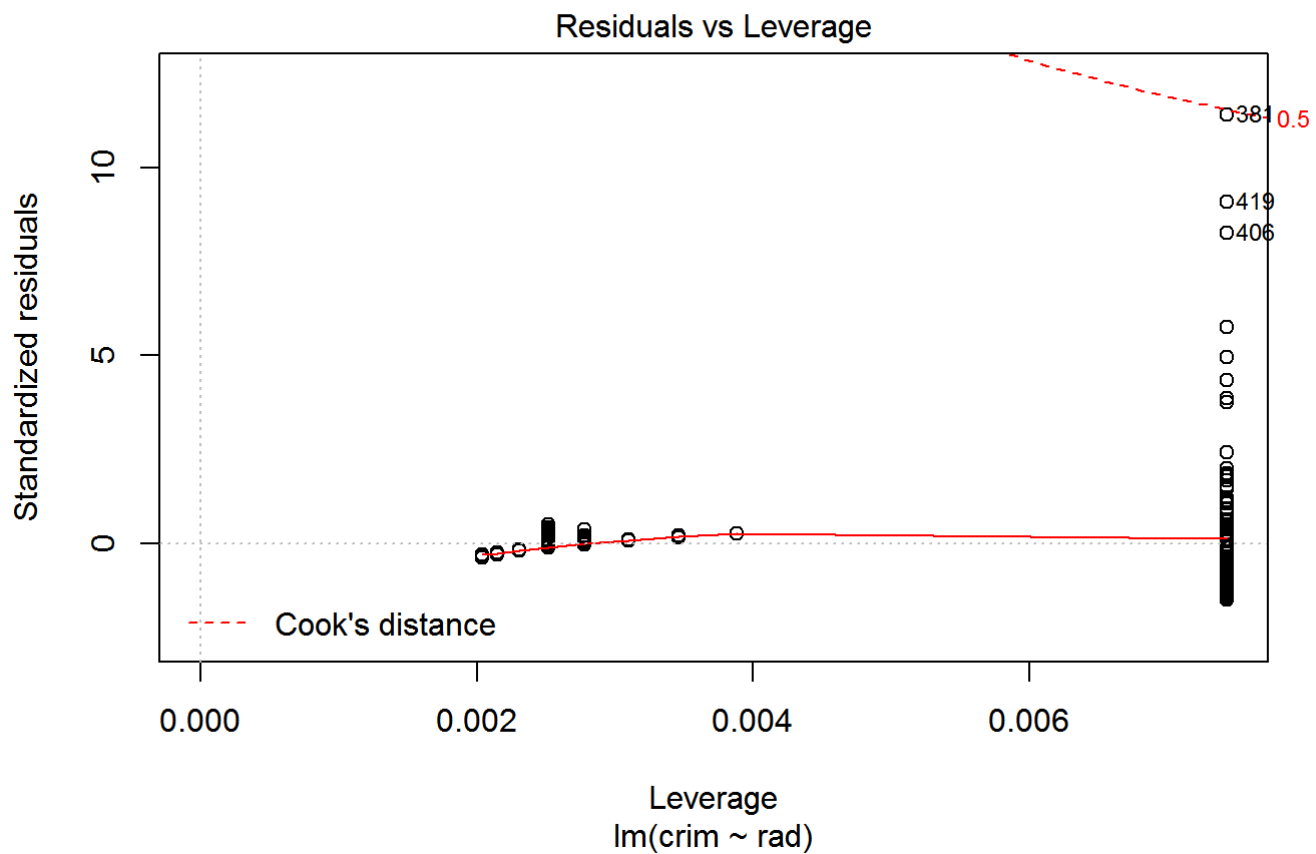
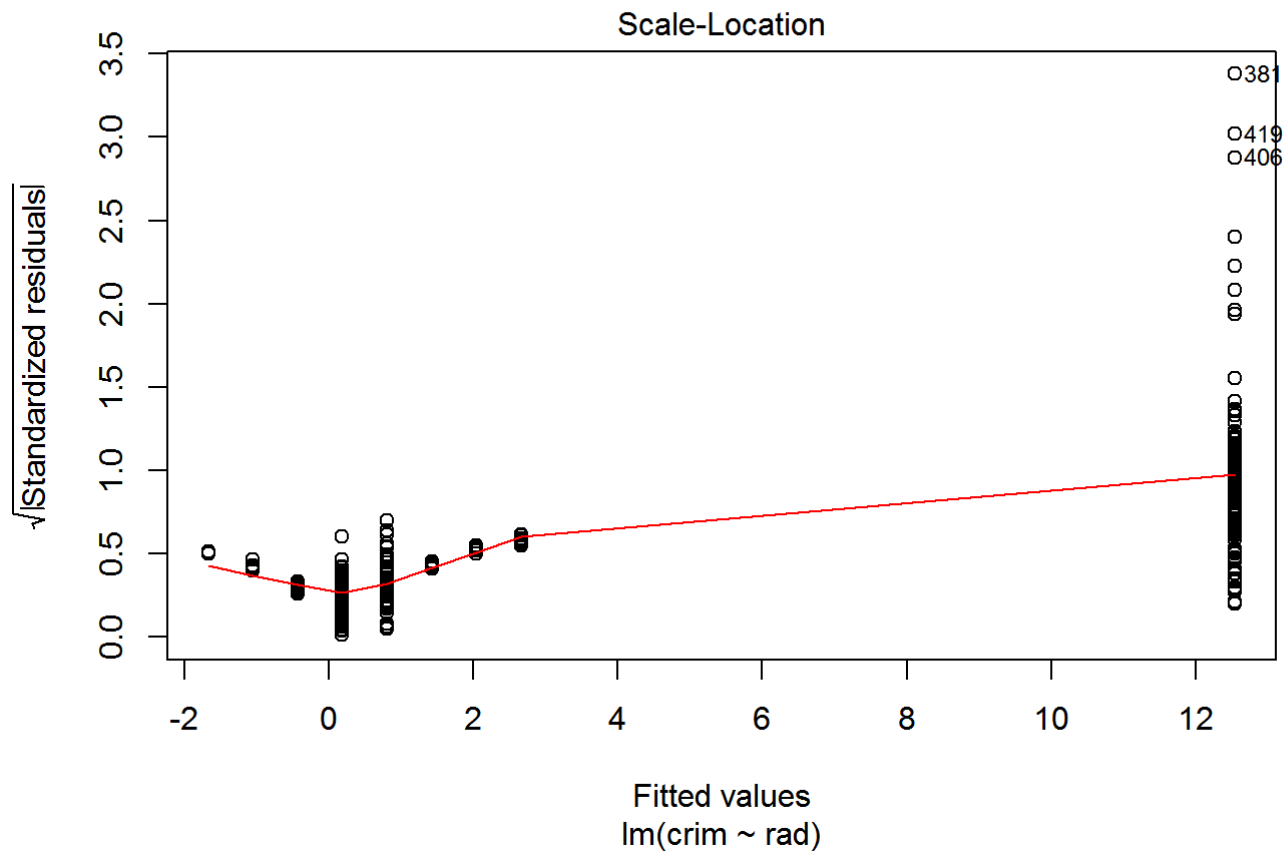
rad

```
m8_rad<-lm(crim~rad,data = Boston)
summary(m8_rad)
```

```
##
## Call:
## lm(formula = crim ~ rad, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.164  -1.381  -0.141   0.660   76.433
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.28716    0.44348  -5.157 3.61e-07 ***
## rad          0.61791    0.03433  17.998 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.718 on 504 degrees of freedom
## Multiple R-squared:  0.3913, Adjusted R-squared:  0.39
## F-statistic: 323.9 on 1 and 504 DF,  p-value: < 2.2e-16
```

```
plot(m8_rad)
```





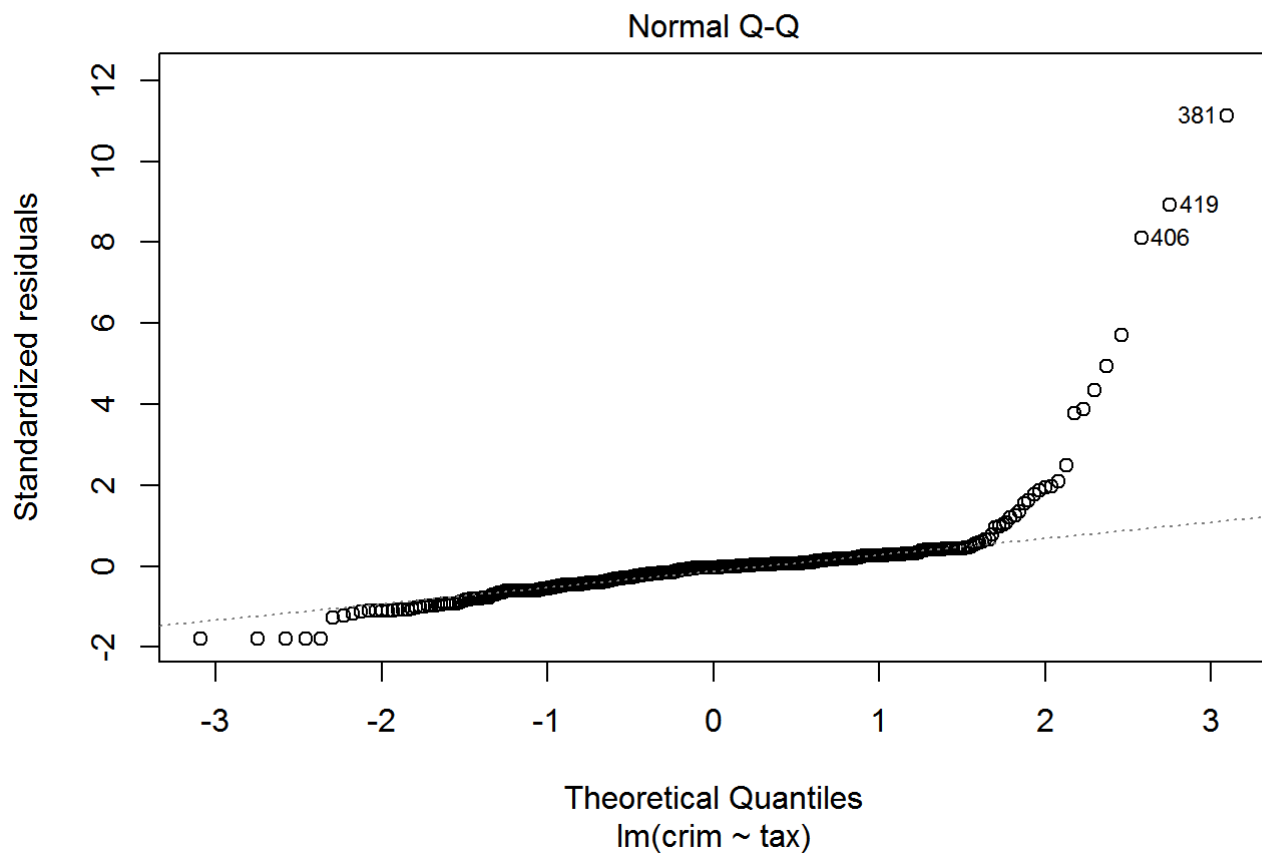
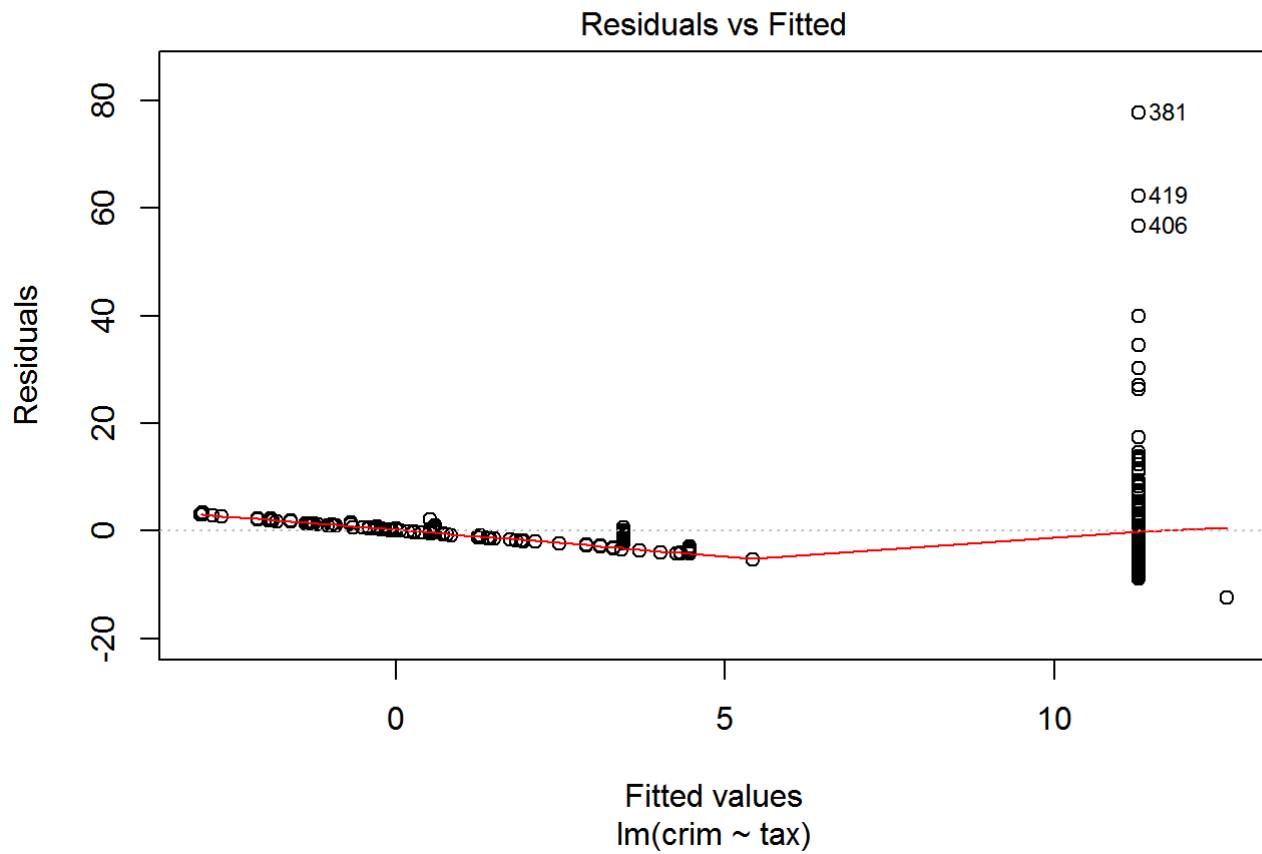
For

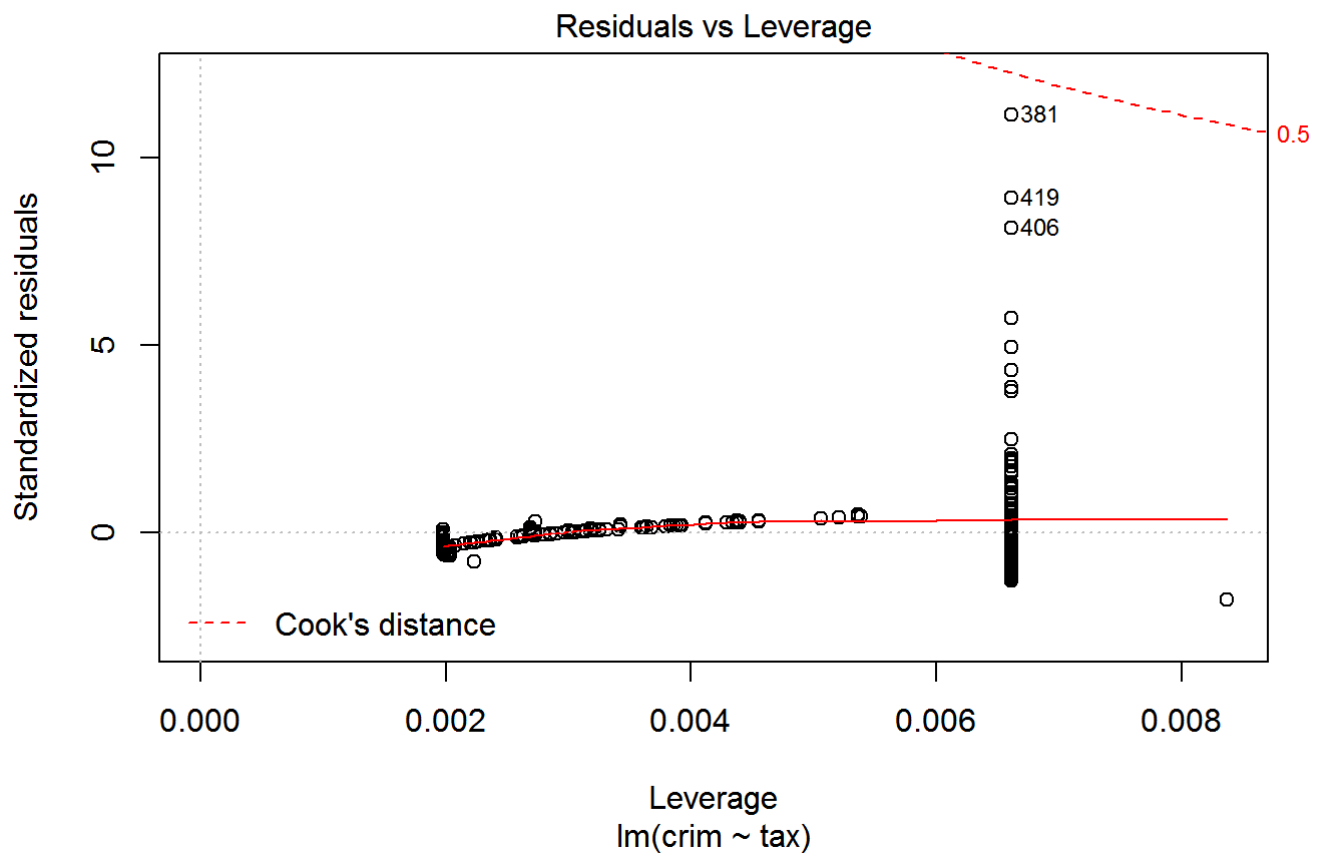
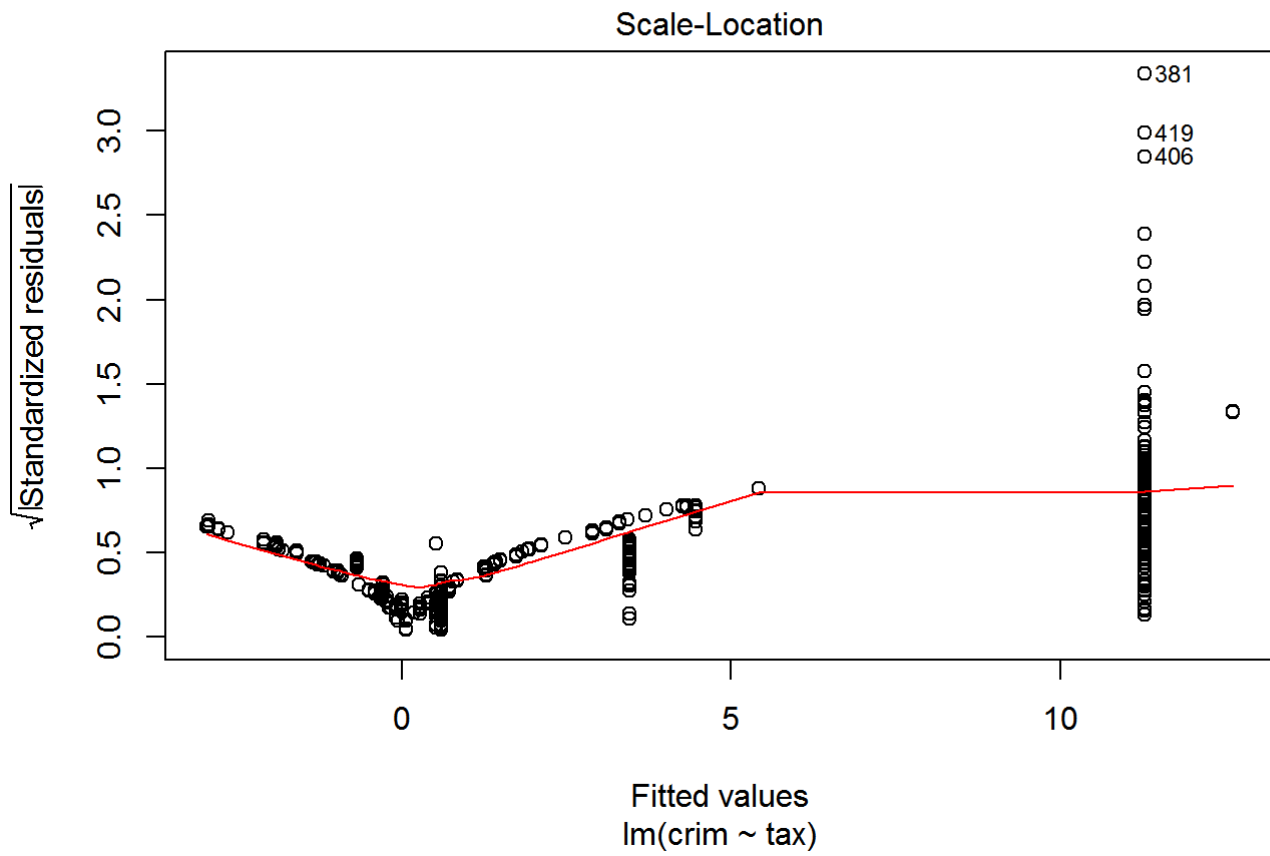
tax

```
m9_tax<-lm(crim~tax,data = Boston)
summary(m9_tax)
```

```
##
## Call:
## lm(formula = crim ~ tax, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.513  -2.738  -0.194   1.065  77.696
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -8.528369   0.815809  -10.45  <2e-16 ***
## tax          0.029742   0.001847   16.10  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.997 on 504 degrees of freedom
## Multiple R-squared:  0.3396, Adjusted R-squared:  0.3383
## F-statistic: 259.2 on 1 and 504 DF,  p-value: < 2.2e-16
```

```
plot(m9_tax)
```





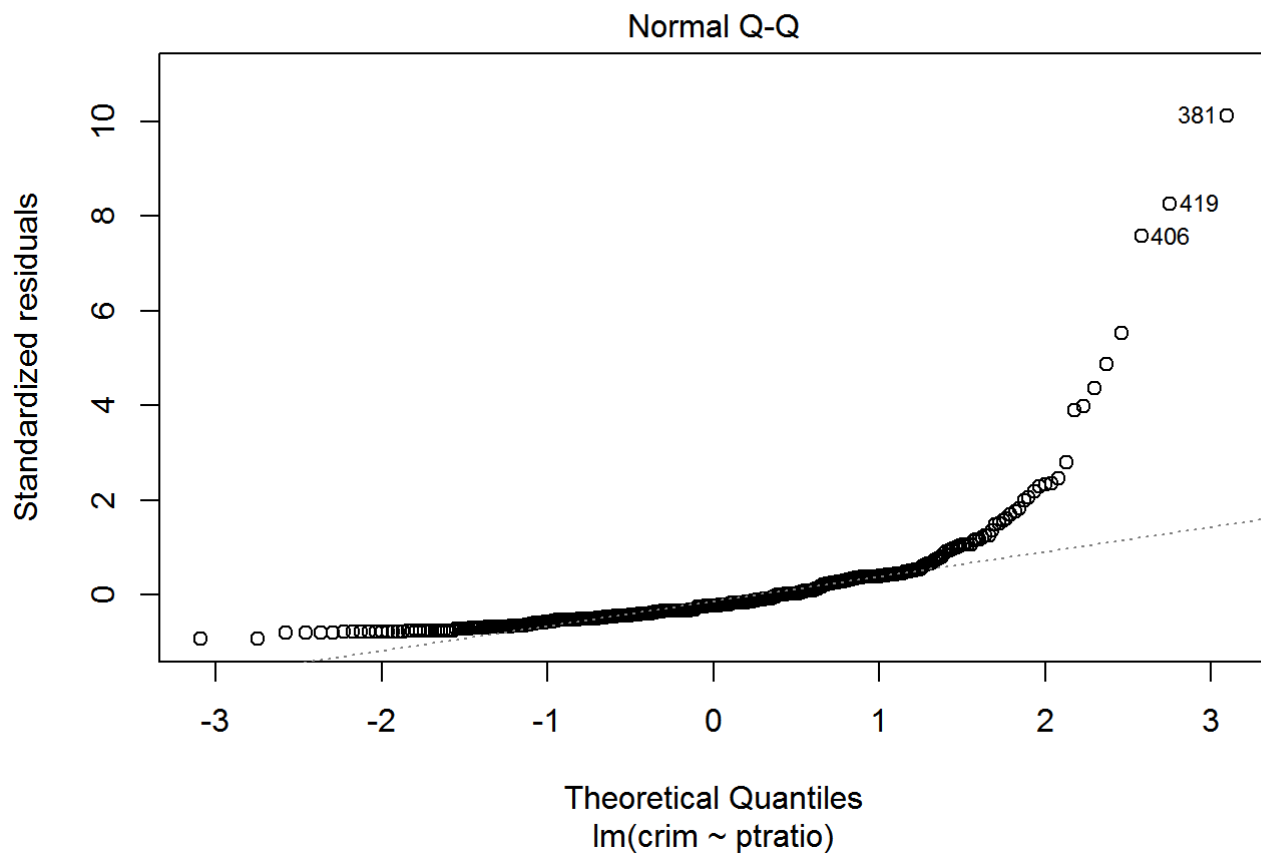
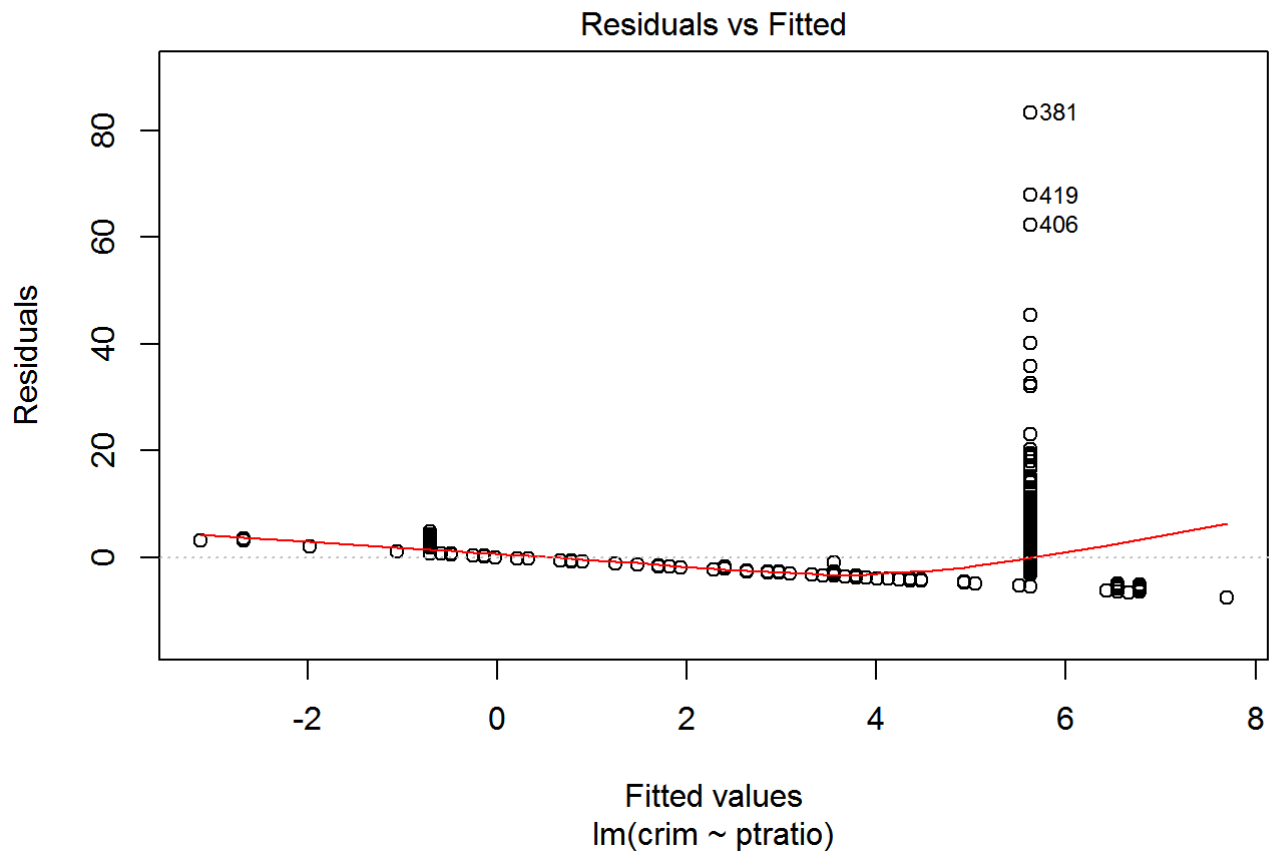
For

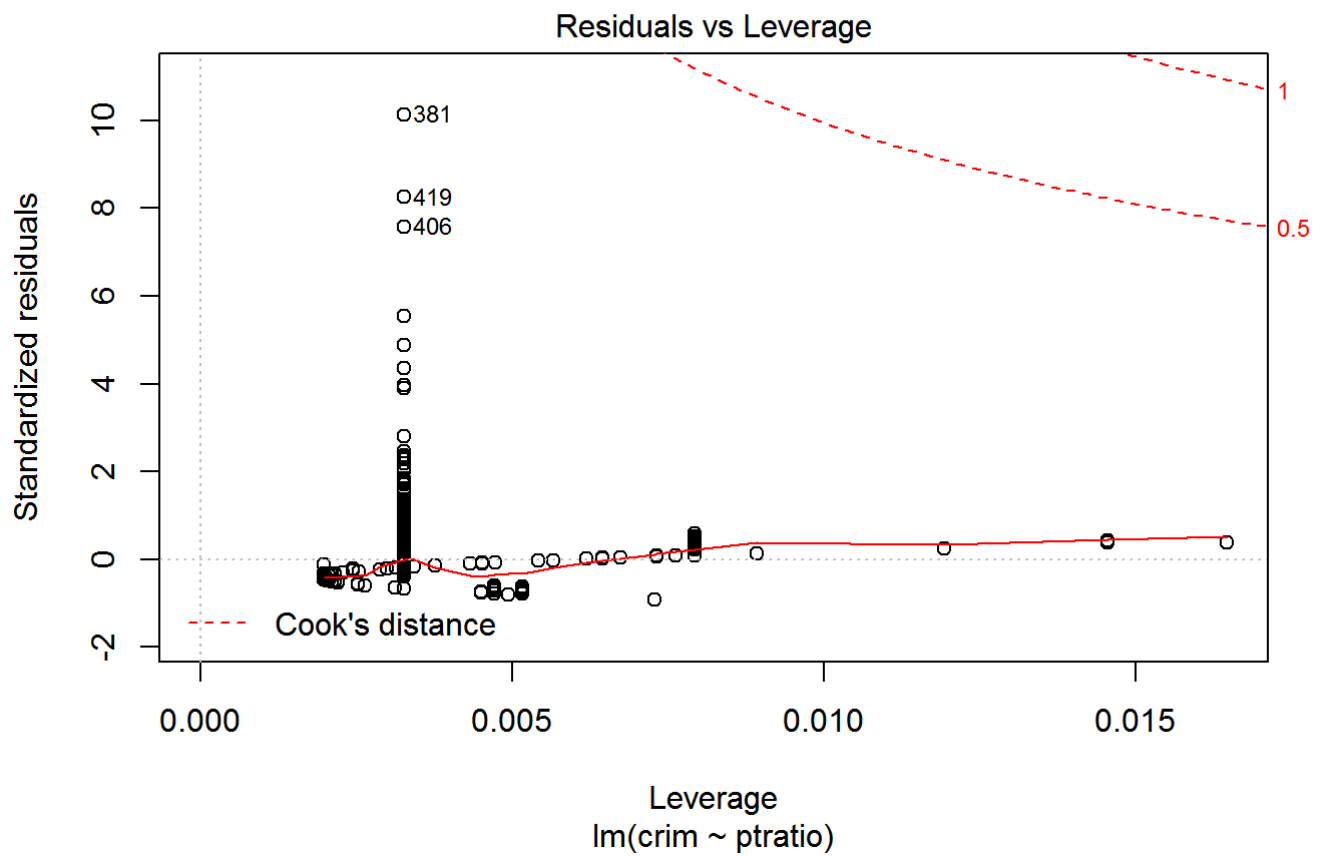
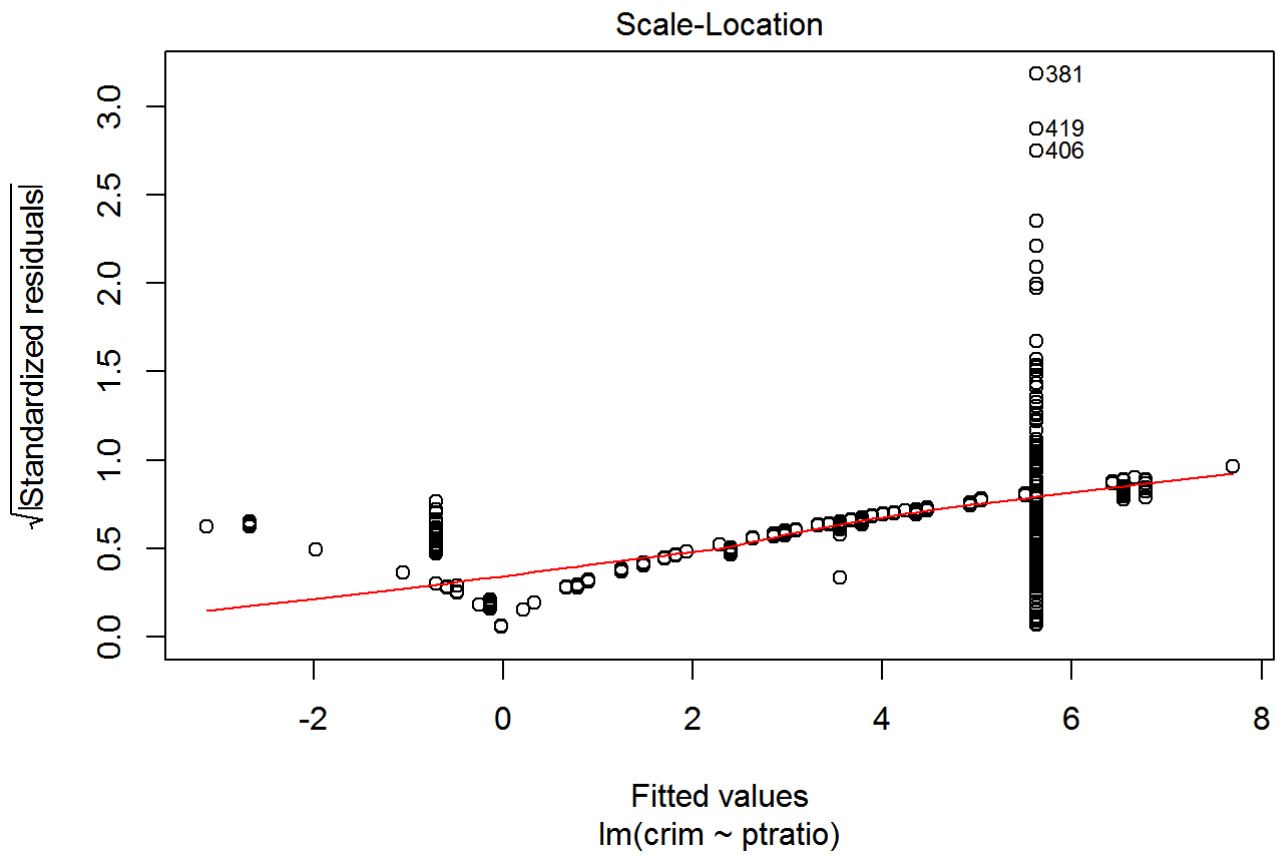
ptratio

```
m10_ptratio<-lm(crim~ptratio,data = Boston)
summary(m10_ptratio)
```

```
##
## Call:
## lm(formula = crim ~ ptratio, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.654  -3.985  -1.912   1.825  83.353
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -17.6469     3.1473  -5.607 3.40e-08 ***
## ptratio      1.1520     0.1694   6.801 2.94e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.24 on 504 degrees of freedom
## Multiple R-squared:  0.08407,    Adjusted R-squared:  0.08225
## F-statistic: 46.26 on 1 and 504 DF,  p-value: 2.943e-11
```

```
plot(m10_ptratio)
```





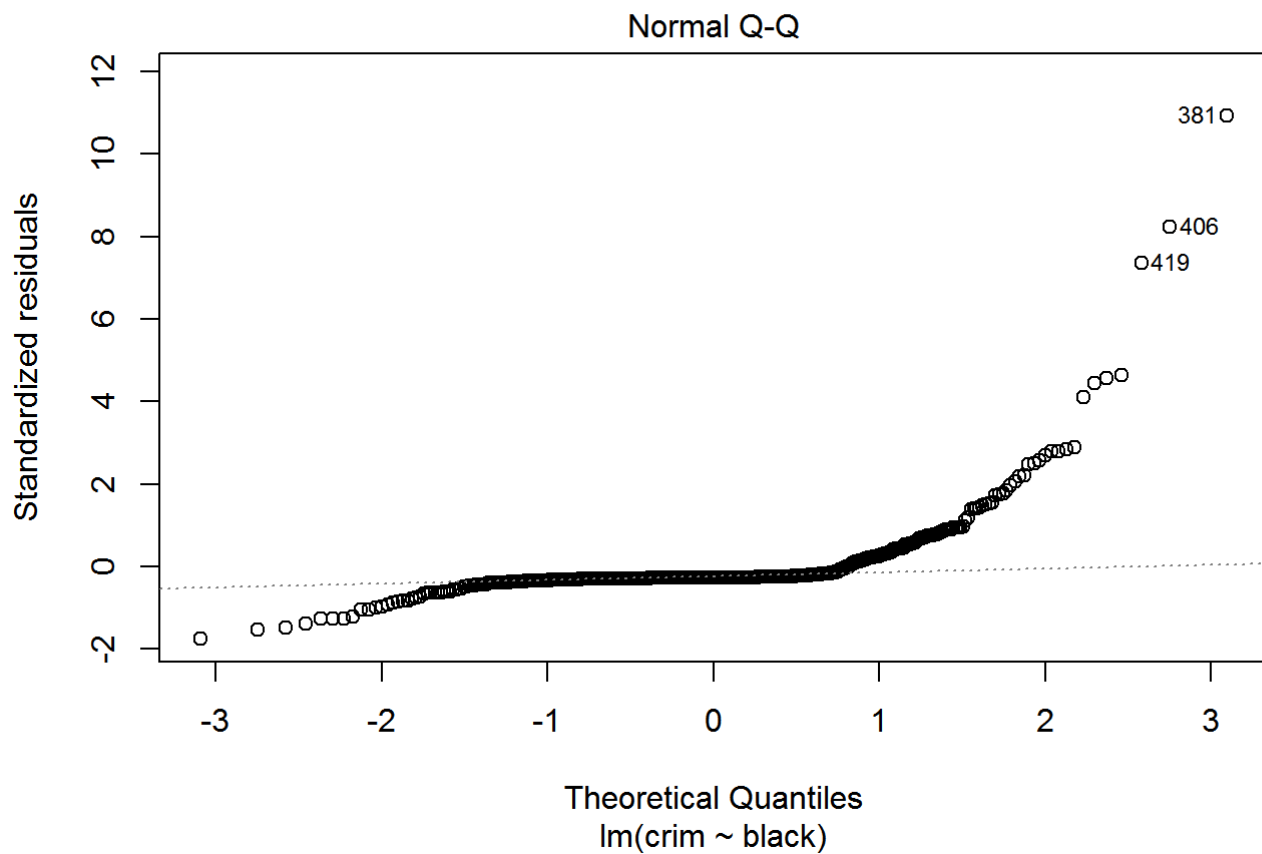
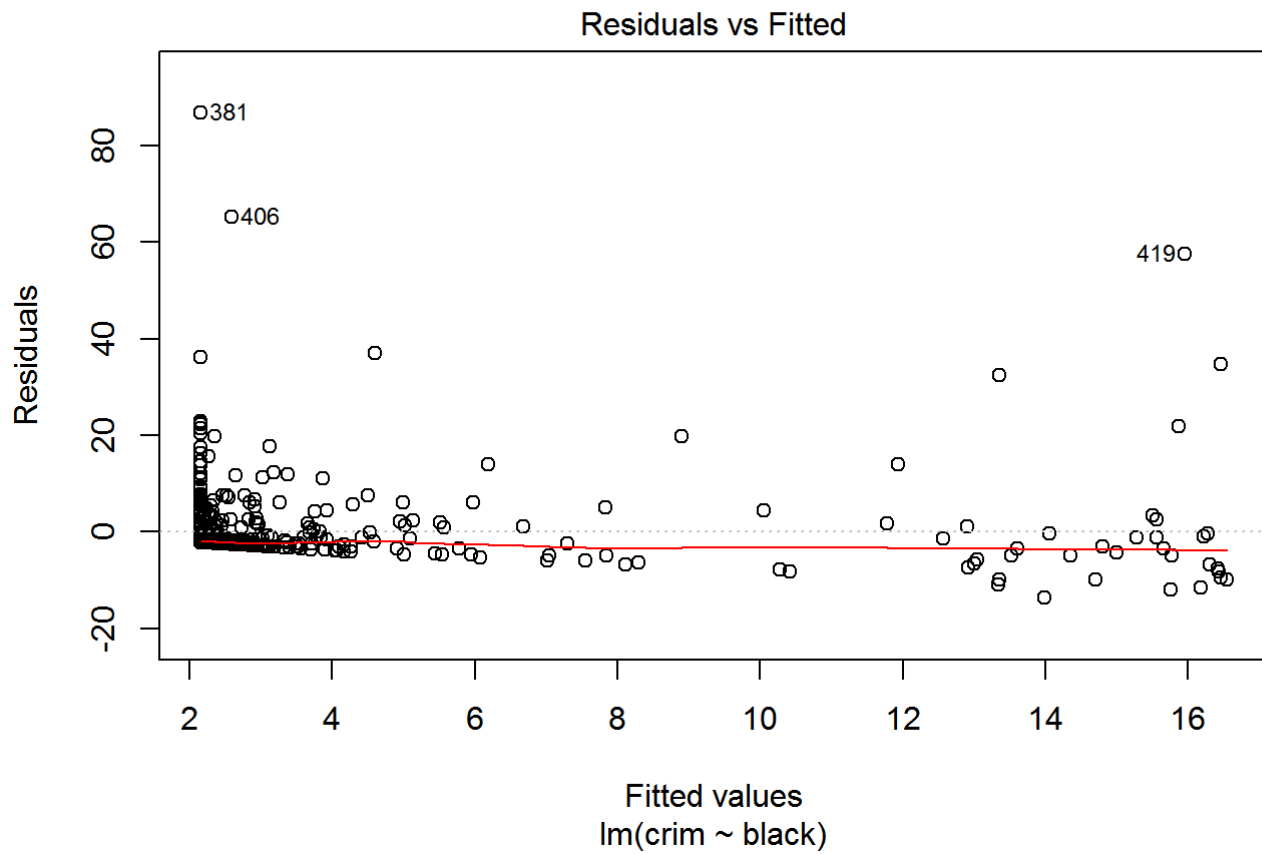
For

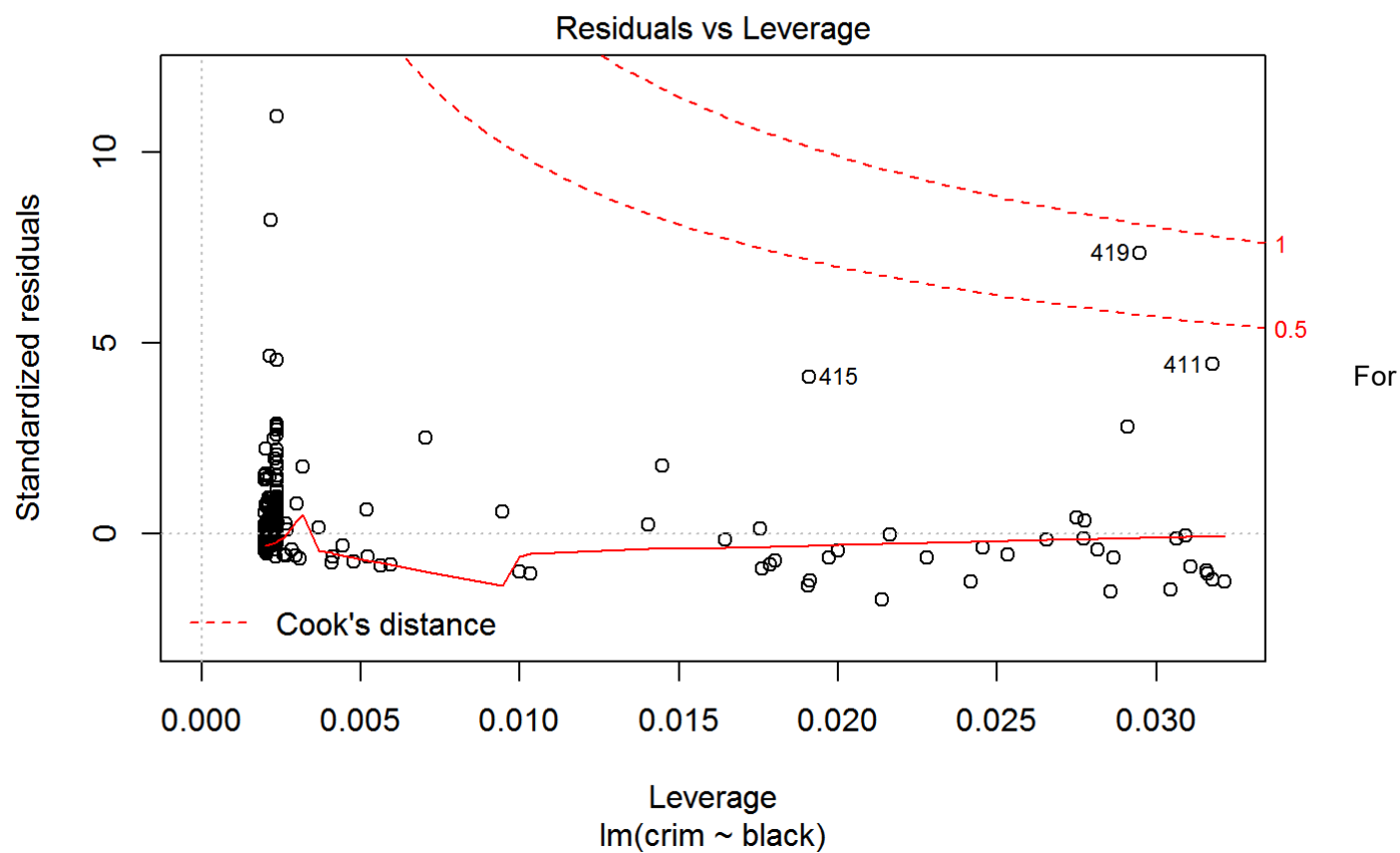
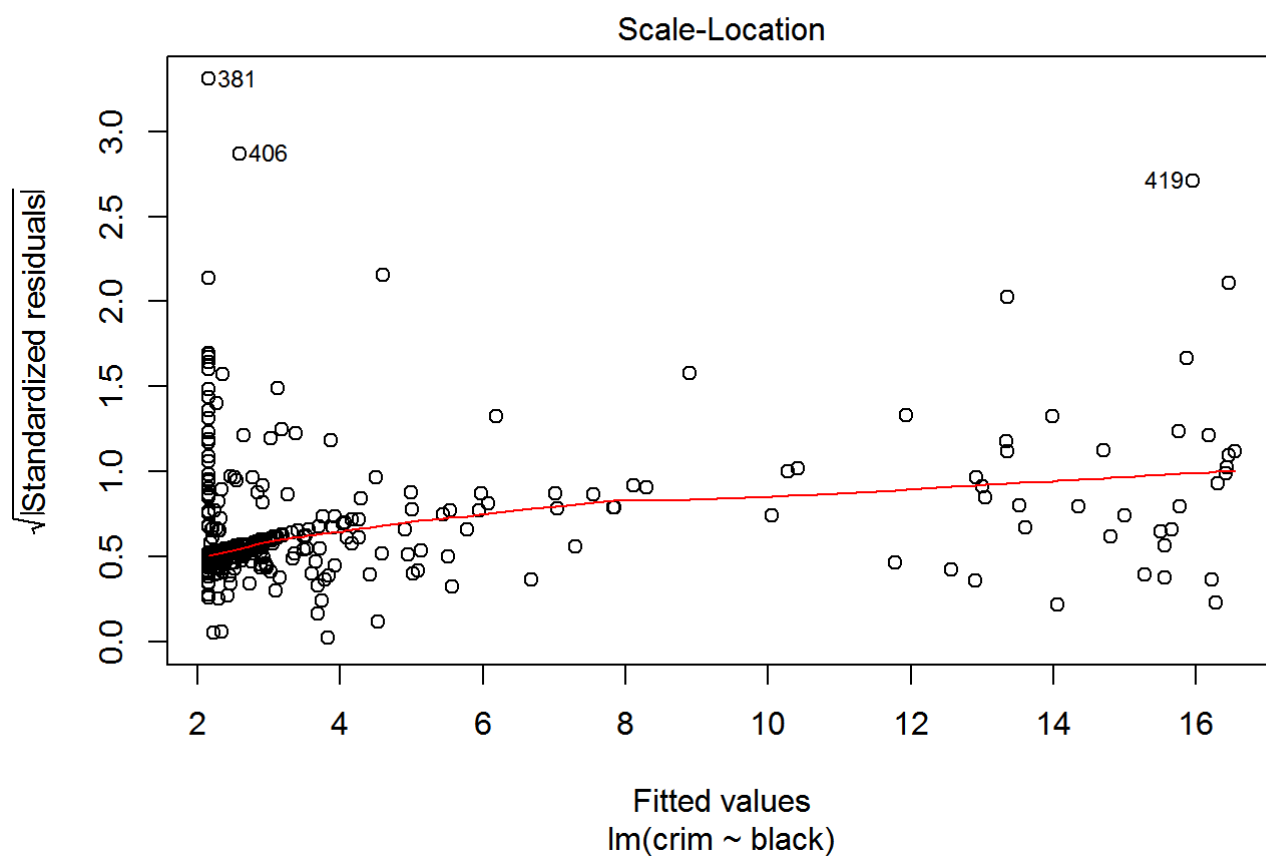
black

```
m11_black<-lm(crim~black,data = Boston)
summary(m11_black)
```

```
##
## Call:
## lm(formula = crim ~ black, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.756  -2.299  -2.095  -1.296   86.822
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 16.553529   1.425903  11.609  <2e-16 ***
## black       -0.036280   0.003873  -9.367  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.946 on 504 degrees of freedom
## Multiple R-squared:  0.1483, Adjusted R-squared:  0.1466
## F-statistic: 87.74 on 1 and 504 DF,  p-value: < 2.2e-16
```

```
plot(m11_black)
```

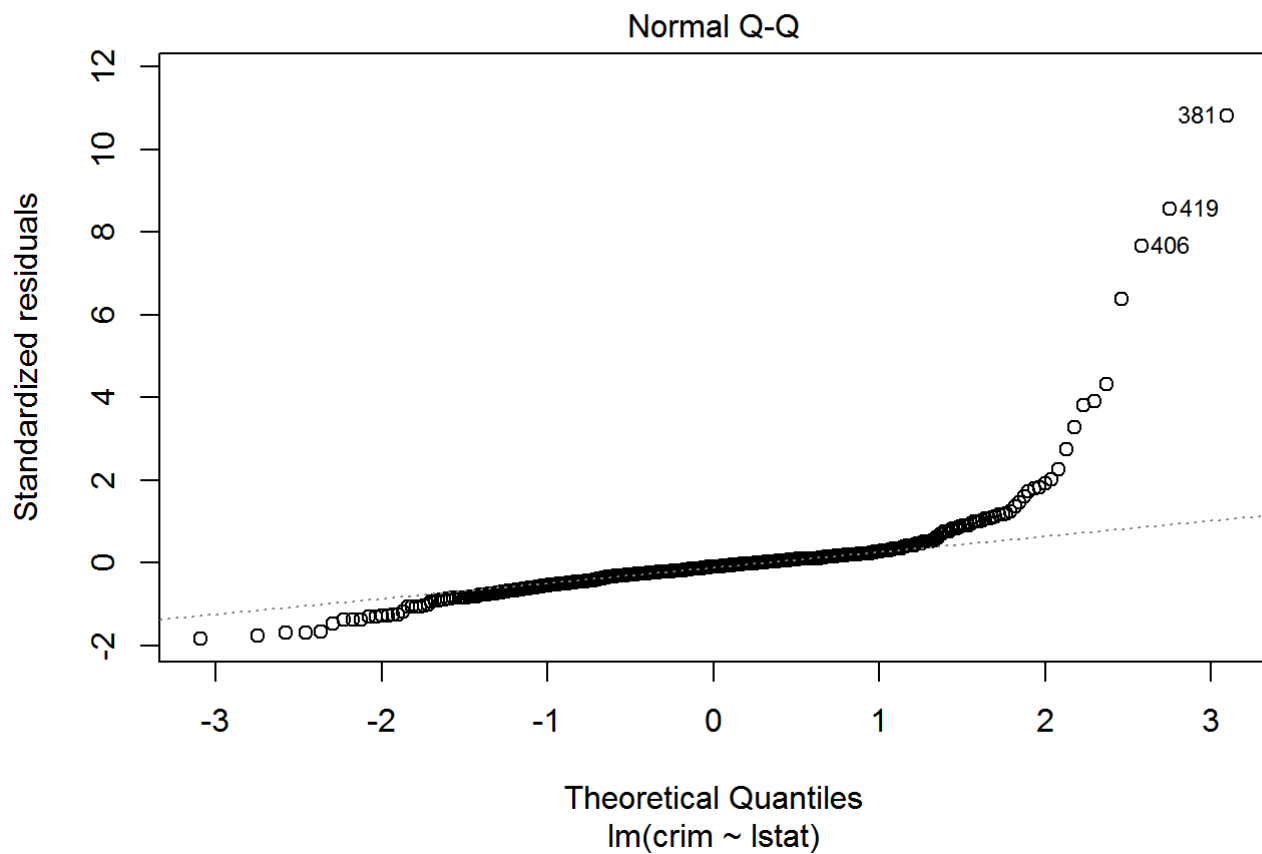
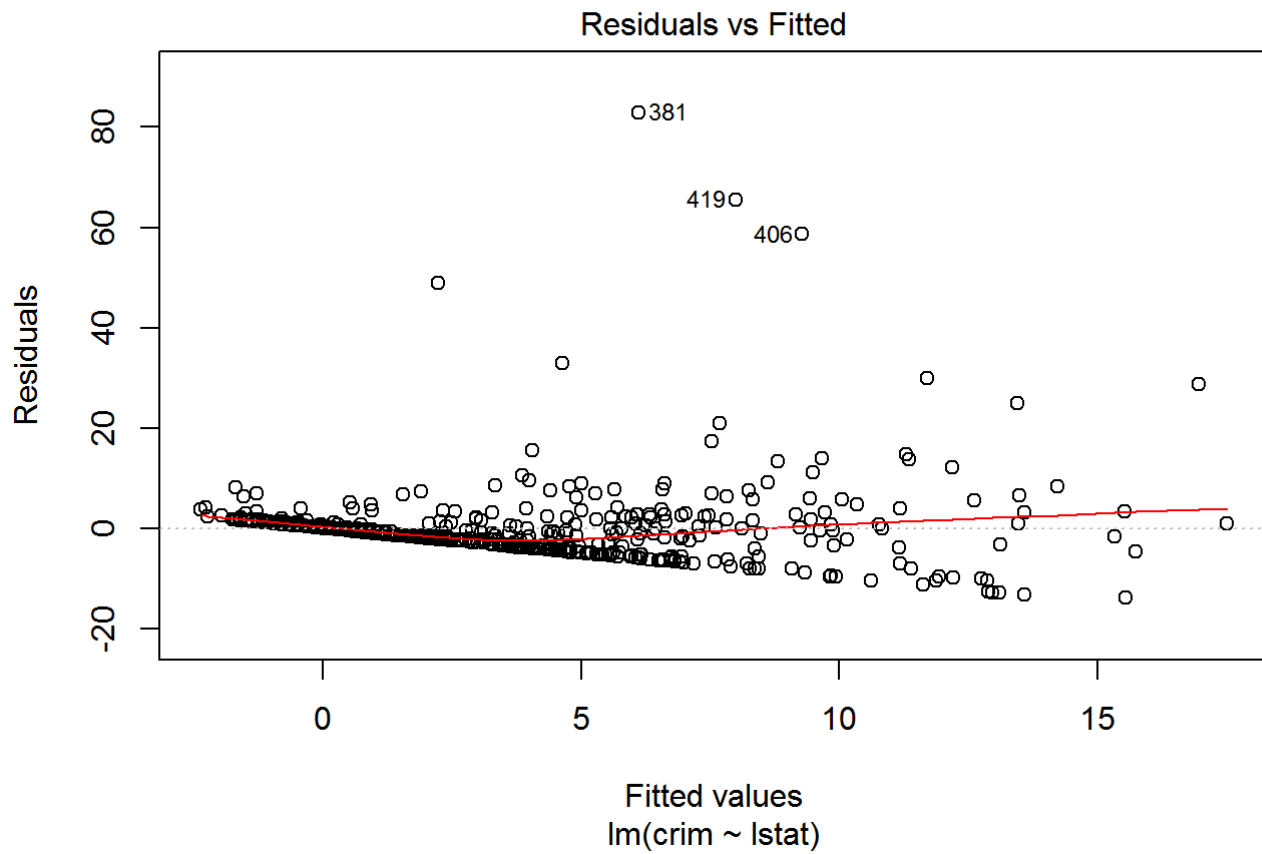


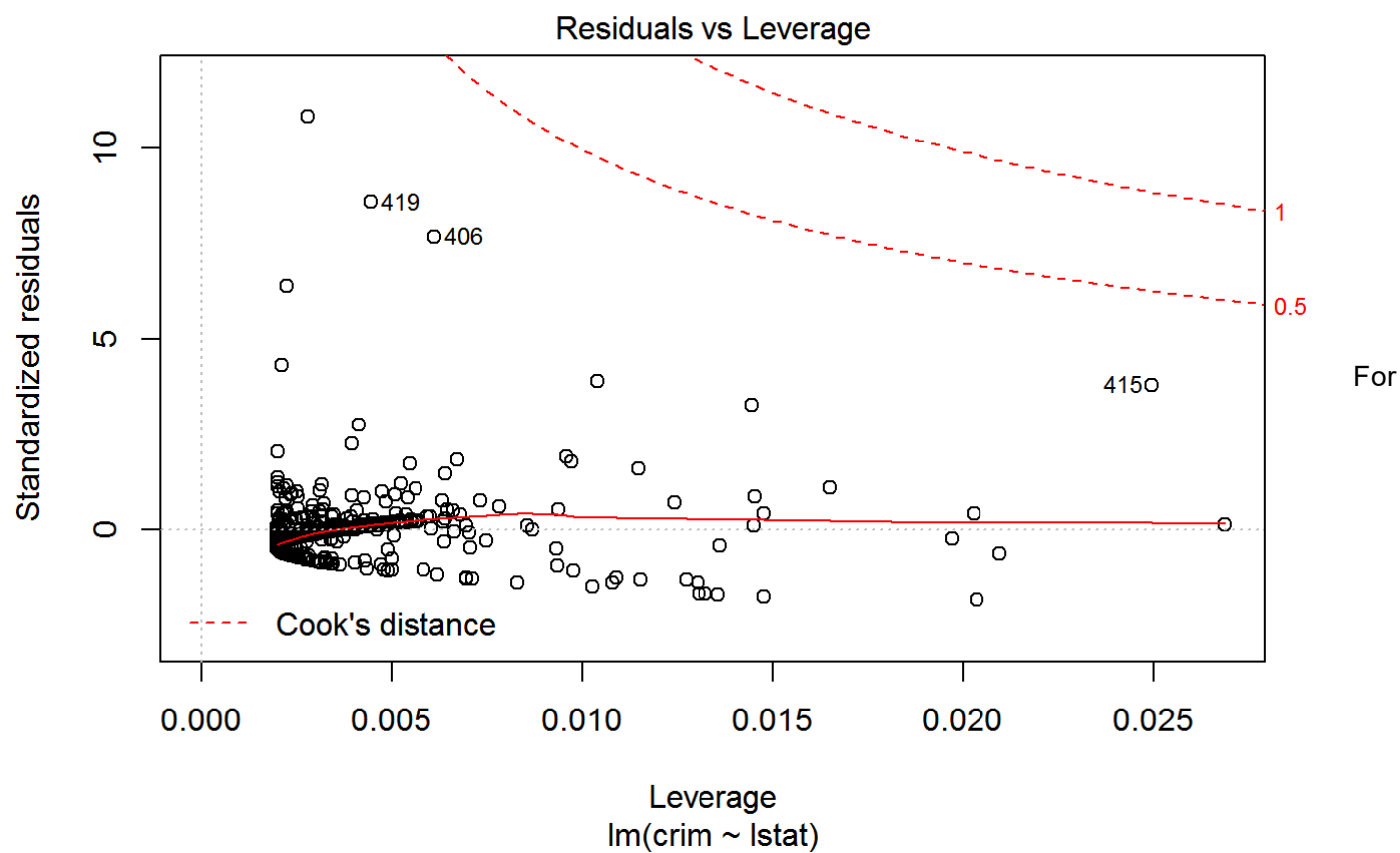
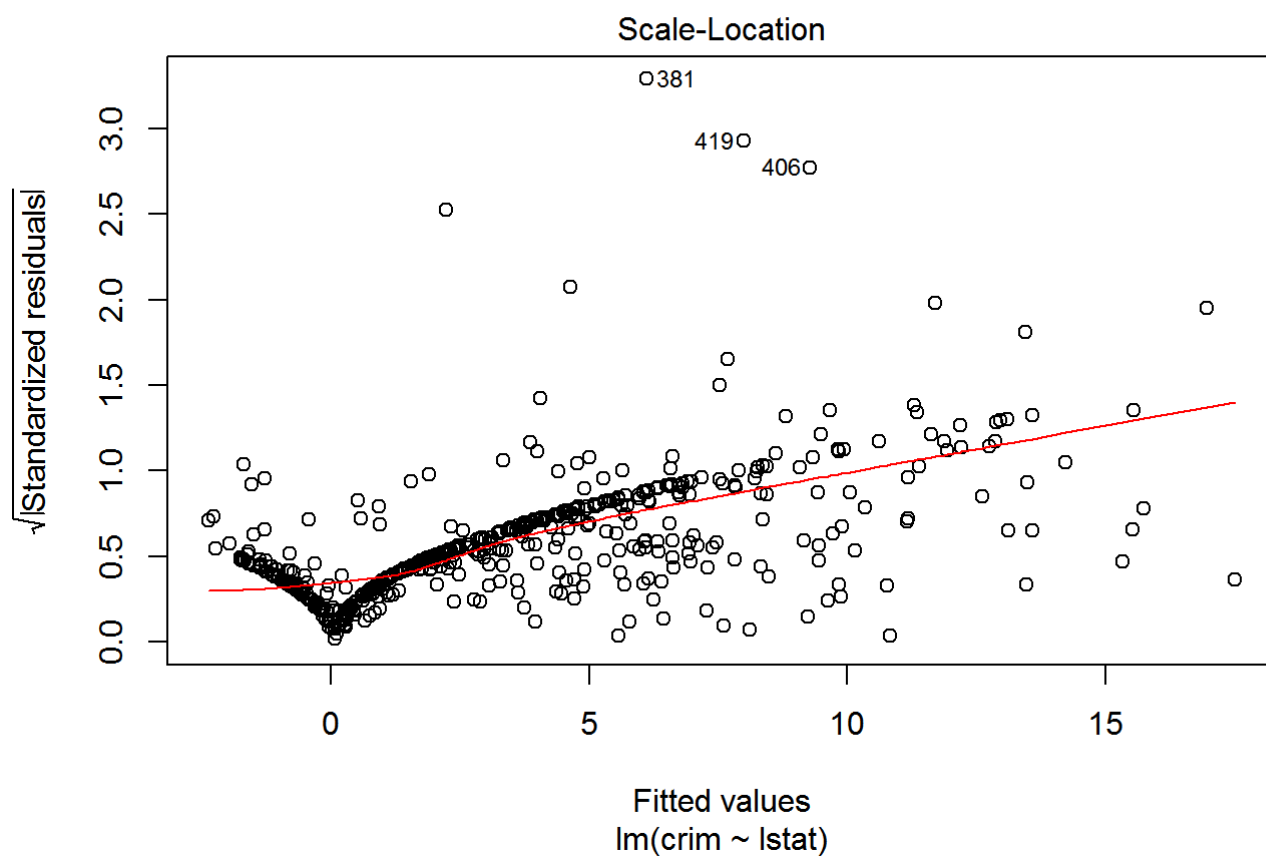
Istat

```
m12_lstat<-lm(crim~lstat,data = Boston)
summary(m12_lstat)
```

```
##
## Call:
## lm(formula = crim ~ lstat, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.925  -2.822  -0.664   1.079   82.862
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.33054    0.69376  -4.801 2.09e-06 ***
## lstat        0.54880    0.04776  11.491 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.664 on 504 degrees of freedom
## Multiple R-squared:  0.2076, Adjusted R-squared:  0.206
## F-statistic: 132 on 1 and 504 DF, p-value: < 2.2e-16
```

```
plot(m12_lstat)
```



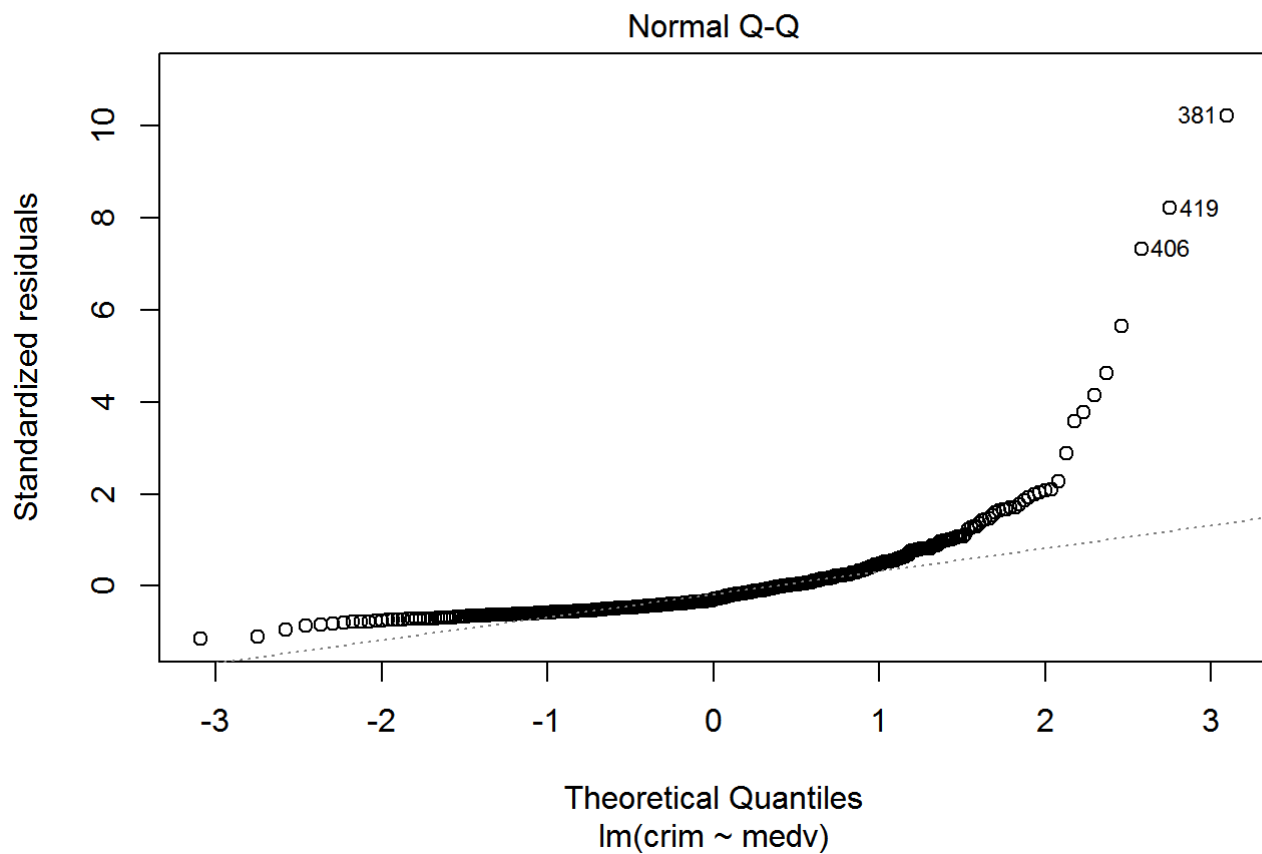
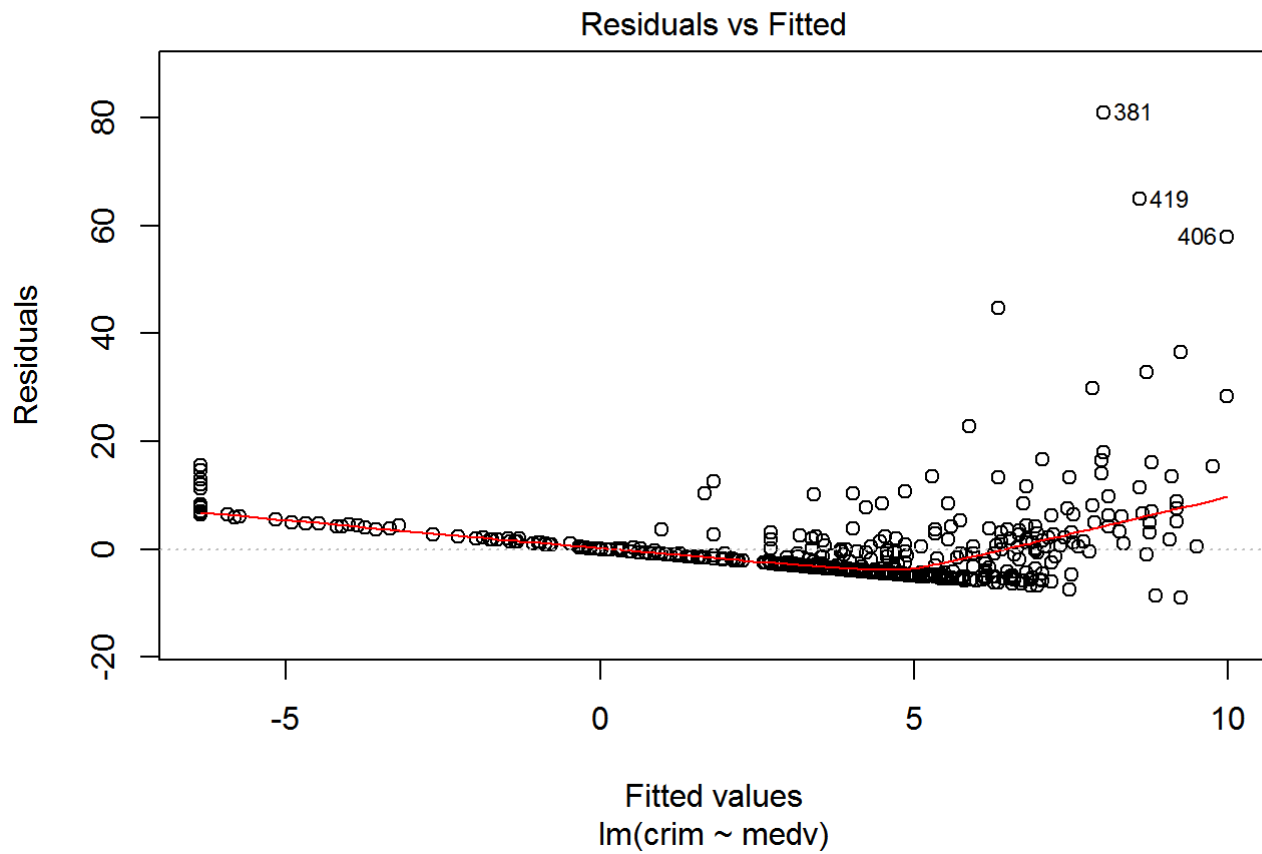


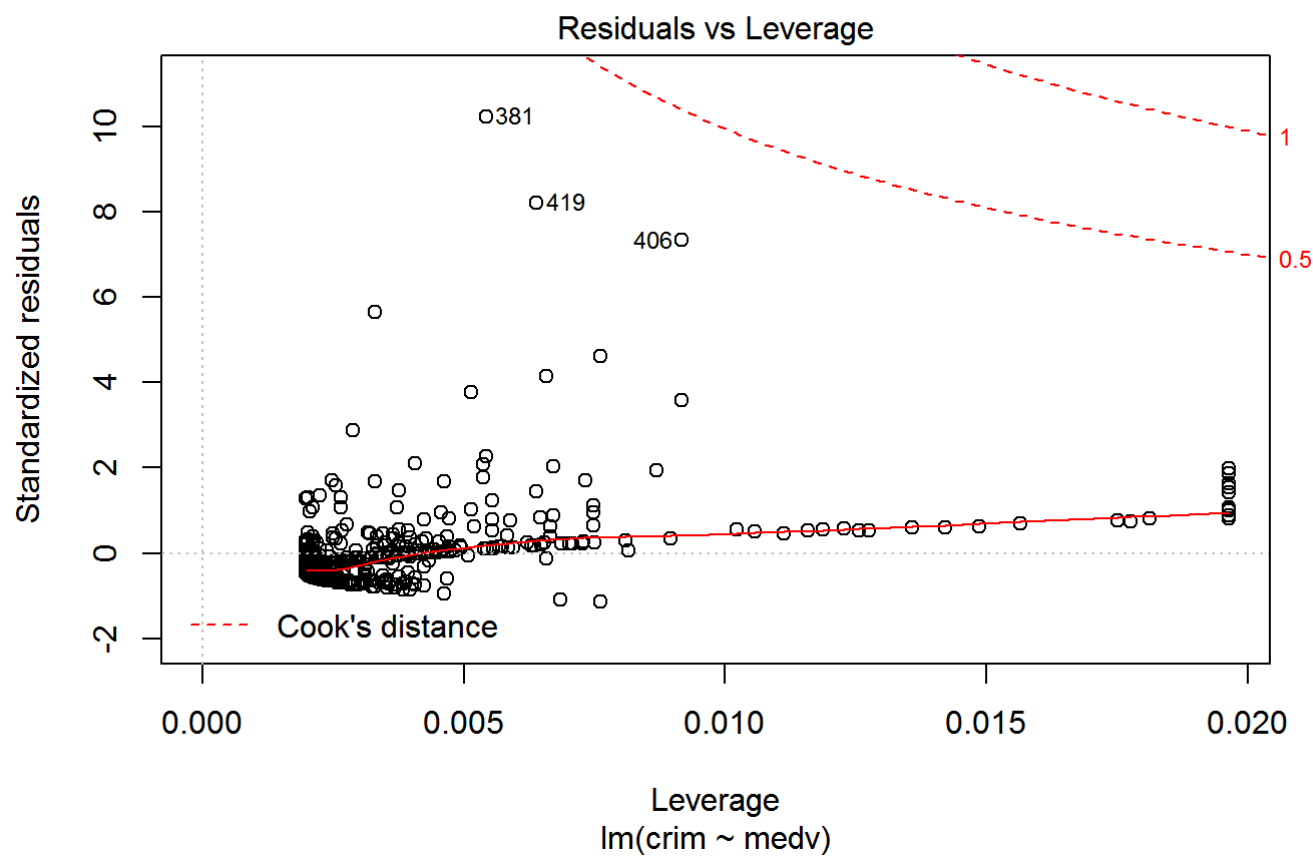
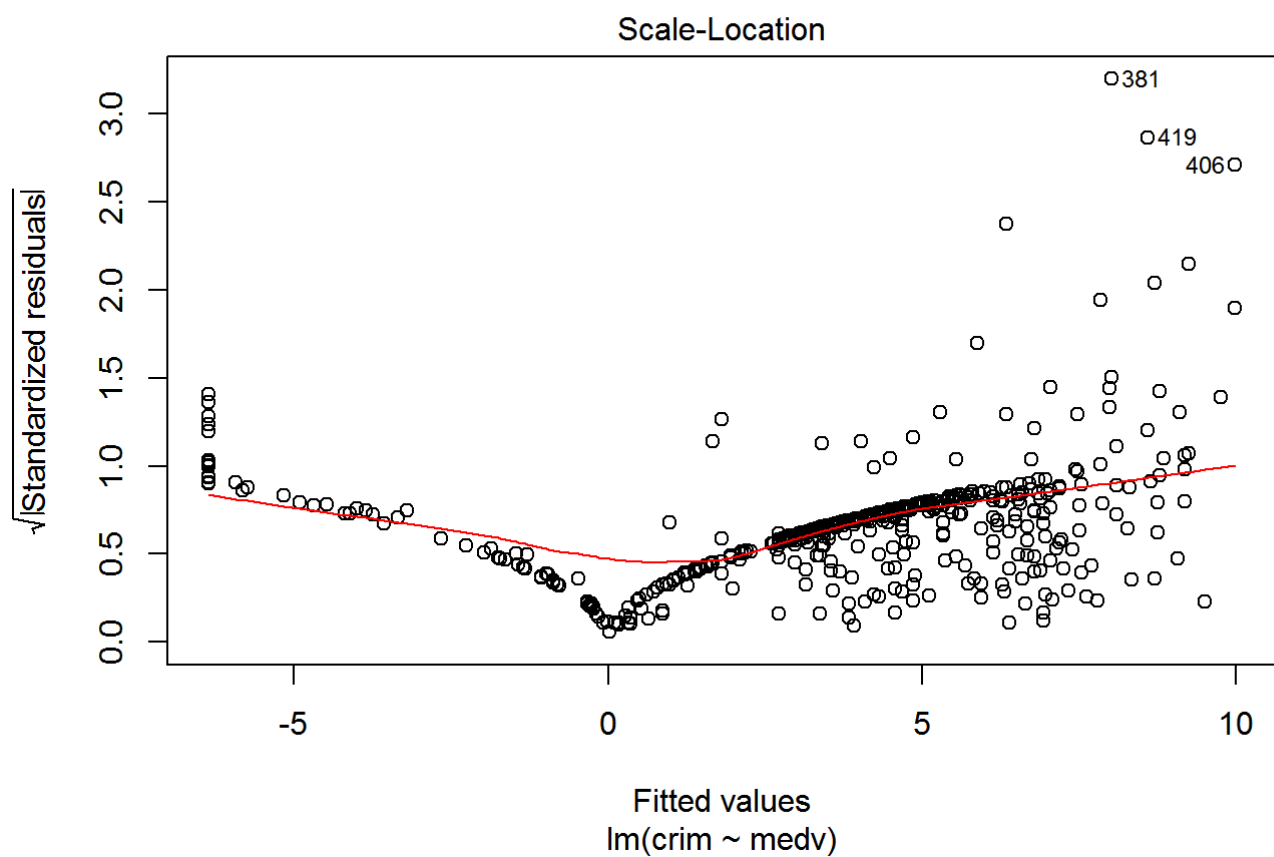
medv

```
m13_medv<-lm(crim~medv,data = Boston)
summary(m13_medv)
```

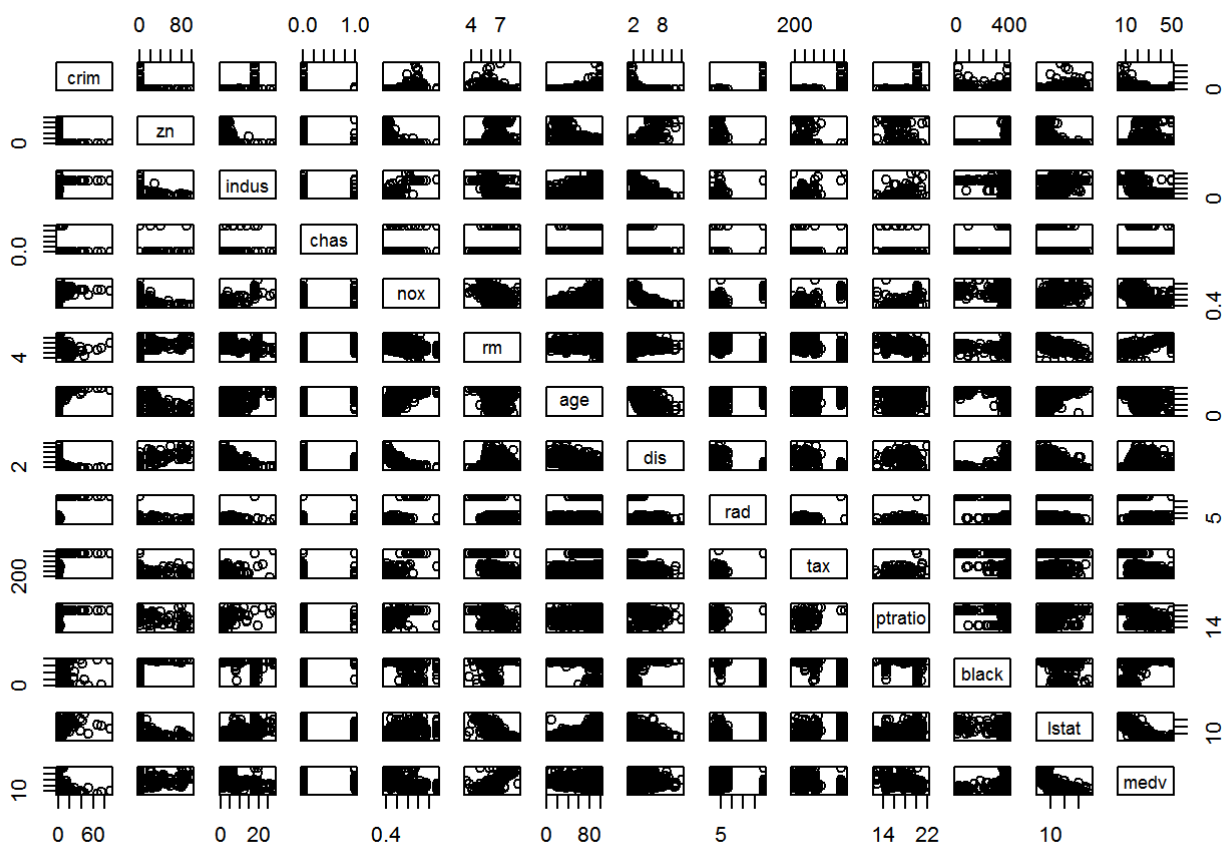
```
##
## Call:
## lm(formula = crim ~ medv, data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.071  -4.022  -2.343   1.298  80.957
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.79654    0.93419   12.63  <2e-16 ***
## medv        -0.36316    0.03839   -9.46  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.934 on 504 degrees of freedom
## Multiple R-squared:  0.1508, Adjusted R-squared:  0.1491
## F-statistic: 89.49 on 1 and 504 DF,  p-value: < 2.2e-16
```

```
plot(m13_medv)
```






```
plot(Boston)
```



Conclusion

In all the models apart from the one with chas variable(m3_chas) there is a statistically significant association between the predictor and the response.

Multiple regression using all variables

```
m14_all<-lm(crim~.,data=Boston)
summary(m14_all)
```

```
##
## Call:
## lm(formula = crim ~ ., data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.924 -2.120 -0.353  1.019 75.051
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.033228   7.234903   2.354 0.018949 *
## zn           0.044855   0.018734   2.394 0.017025 *
## indus        -0.063855   0.083407  -0.766 0.444294
## chas         -0.749134   1.180147  -0.635 0.525867
## nox          -10.313535   5.275536  -1.955 0.051152 .
## rm           0.430131   0.612830   0.702 0.483089
## age          0.001452   0.017925   0.081 0.935488
## dis          -0.987176   0.281817  -3.503 0.000502 ***
## rad           0.588209   0.088049   6.680 6.46e-11 ***
## tax          -0.003780   0.005156  -0.733 0.463793
## ptratio      -0.271081   0.186450  -1.454 0.146611
## black        -0.007538   0.003673  -2.052 0.040702 *
## lstat         0.126211   0.075725   1.667 0.096208 .
## medv         -0.198887   0.060516  -3.287 0.001087 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.439 on 492 degrees of freedom
## Multiple R-squared:  0.454, Adjusted R-squared:  0.4396
## F-statistic: 31.47 on 13 and 492 DF,  p-value: < 2.2e-16
```

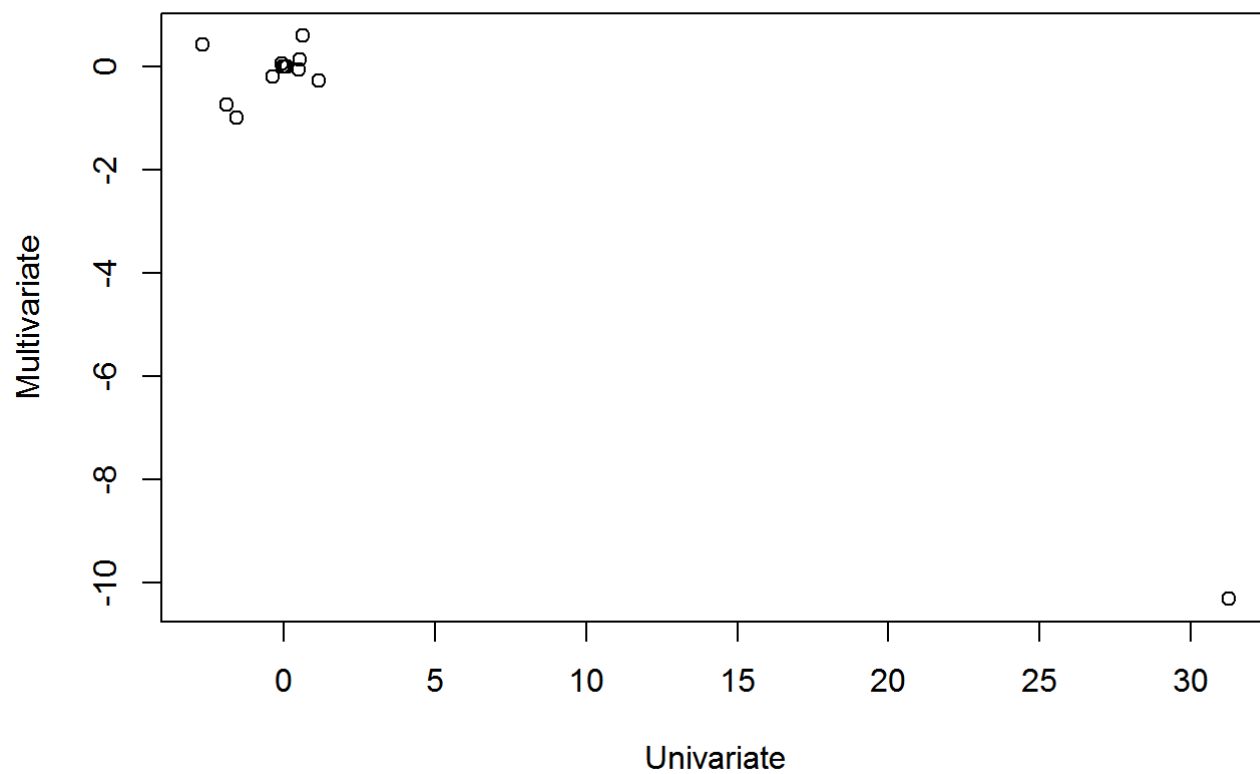
We can reject null hypothesis ($H_0 : \beta_j = 0$) for dis and rad to be statistically significant at .001, medv to be statistically significant at .01, intercept,zn and black to be statistically significant at .05 and nox and age to be statistically significant at .1

Result comparsion and plot

It appeared that all variables except chas were individually statistically significant but when all of them are used together many of the variables including chas were found not to be statistically significant. Also the R squared in all the models was very low.

```
x<-c(m1_zn$coefficients[-1],m2_indus$coefficients[-1],m3_chas$coefficients[-1],m4_nox$coefficients[-1],m5_rm$coefficients[-1],m6_age$coefficients[-1],m7_dis$coefficients[-1],m8_rad$coefficients[-1],m9_tax$coefficients[-1],m10_ptratio$coefficients[-1],m11_black$coefficients[-1],m12_lstat$coefficients[-1],m13_medv$coefficients[-1])
y<-m14_all$coefficients[-1]
plot(x,y,xlab = "Univariate",ylab = "Multivariate",main = "Uni vs Multi")
```

Uni vs Multi



Non-linear associations

For zn

```
m15_znq<-lm(crim~zn+I(zn^2)+I(zn^3),data = Boston)
summary(m15_znq)
```

```
##
## Call:
## lm(formula = crim ~ zn + I(zn^2) + I(zn^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.821 -4.614 -1.294  0.473  84.130
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.846e+00  4.330e-01  11.192  < 2e-16 ***
## zn          -3.322e-01  1.098e-01  -3.025  0.00261 **
## I(zn^2)       6.483e-03  3.861e-03   1.679  0.09375 .
## I(zn^3)      -3.776e-05  3.139e-05  -1.203  0.22954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.372 on 502 degrees of freedom
## Multiple R-squared:  0.05824,    Adjusted R-squared:  0.05261
## F-statistic: 10.35 on 3 and 502 DF,  p-value: 1.281e-06
```

For indus

```
m16_indus<-lm(crim~indus+I(indus^2)+I(indus^3),data = Boston)
summary(m16_indus)
```

```
##
## Call:
## lm(formula = crim ~ indus + I(indus^2) + I(indus^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.278 -2.514  0.054  0.764  79.713
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.6625683  1.5739833   2.327  0.0204 *
## indus        -1.9652129  0.4819901  -4.077 5.30e-05 ***
## I(indus^2)    0.2519373  0.0393221   6.407 3.42e-10 ***
## I(indus^3)   -0.0069760  0.0009567  -7.292 1.20e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.423 on 502 degrees of freedom
## Multiple R-squared:  0.2597, Adjusted R-squared:  0.2552
## F-statistic: 58.69 on 3 and 502 DF,  p-value: < 2.2e-16
```

For chas

```
m17_chas<-lm(crim~chas+I(chas^2)+I(chas^3),data = Boston)
summary(m17_chas)
```

```
##
## Call:
## lm(formula = crim ~ chas + I(chas^2) + I(chas^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.738 -3.661 -3.435  0.018 85.232
##
## Coefficients: (2 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.7444     0.3961   9.453  <2e-16 ***
## chas          -1.8928     1.5061  -1.257   0.209
## I(chas^2)         NA          NA      NA      NA
## I(chas^3)         NA          NA      NA      NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.597 on 504 degrees of freedom
## Multiple R-squared:  0.003124, Adjusted R-squared:  0.001146
## F-statistic: 1.579 on 1 and 504 DF, p-value: 0.2094
```

For nox

```
m18_nox<-lm(crim~nox+I(nox^2)+I(nox^3),data = Boston)
summary(m18_nox)
```

```
##
## Call:
## lm(formula = crim ~ nox + I(nox^2) + I(nox^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.110 -2.068 -0.255  0.739 78.302
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   233.09      33.64   6.928 1.31e-11 ***
## nox          -1279.37     170.40  -7.508 2.76e-13 ***
## I(nox^2)       2248.54     279.90   8.033 6.81e-15 ***
## I(nox^3)      -1245.70     149.28  -8.345 6.96e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.234 on 502 degrees of freedom
## Multiple R-squared:  0.297, Adjusted R-squared:  0.2928
## F-statistic: 70.69 on 3 and 502 DF, p-value: < 2.2e-16
```

For rm

```
m19_rm<-lm(crim~rm+I(rm^2)+I(rm^3),data = Boston)
summary(m19_rm)
```

```
##
## Call:
## lm(formula = crim ~ rm + I(rm^2) + I(rm^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.485  -3.468  -2.221  -0.015   87.219
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  112.6246    64.5172   1.746   0.0815 .
## rm          -39.1501    31.3115  -1.250   0.2118
## I(rm^2)       4.5509     5.0099   0.908   0.3641
## I(rm^3)      -0.1745     0.2637  -0.662   0.5086
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.33 on 502 degrees of freedom
## Multiple R-squared:  0.06779, Adjusted R-squared:  0.06222
## F-statistic: 12.17 on 3 and 502 DF, p-value: 1.067e-07
```

For age

```
m20_age<-lm(crim~age+I(age^2)+I(age^3),data = Boston)
summary(m20_age)
```

```
##
## Call:
## lm(formula = crim ~ age + I(age^2) + I(age^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
##  -9.762  -2.673  -0.516   0.019  82.842
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.549e+00  2.769e+00  -0.920   0.35780
## age          2.737e-01  1.864e-01   1.468   0.14266
## I(age^2)     -7.230e-03  3.637e-03  -1.988   0.04738 *
## I(age^3)      5.745e-05  2.109e-05   2.724   0.00668 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.84 on 502 degrees of freedom
## Multiple R-squared:  0.1742, Adjusted R-squared:  0.1693
## F-statistic: 35.31 on 3 and 502 DF, p-value: < 2.2e-16
```

For dis

```
m21_dis<-lm(crim~dis+I(dis^2)+I(dis^3),data = Boston)
summary(m21_dis)
```

```
##
## Call:
## lm(formula = crim ~ dis + I(dis^2) + I(dis^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.757  -2.588   0.031   1.267  76.378
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  30.0476     2.4459   12.285 < 2e-16 ***
## dis         -15.5543     1.7360   -8.960 < 2e-16 ***
## I(dis^2)      2.4521     0.3464    7.078 4.94e-12 ***
## I(dis^3)     -0.1186     0.0204   -5.814 1.09e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.331 on 502 degrees of freedom
## Multiple R-squared:  0.2778, Adjusted R-squared:  0.2735
## F-statistic: 64.37 on 3 and 502 DF,  p-value: < 2.2e-16
```

For rad

```
m22_rad<-lm(crim~rad+I(rad^2)+I(rad^3),data = Boston)
summary(m22_rad)
```

```
##
## Call:
## lm(formula = crim ~ rad + I(rad^2) + I(rad^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.381  -0.412  -0.269   0.179  76.217
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.605545     2.050108  -0.295   0.768
## rad          0.512736     1.043597   0.491   0.623
## I(rad^2)     -0.075177     0.148543  -0.506   0.613
## I(rad^3)      0.003209     0.004564   0.703   0.482
##
## Residual standard error: 6.682 on 502 degrees of freedom
## Multiple R-squared:  0.4, Adjusted R-squared:  0.3965
## F-statistic: 111.6 on 3 and 502 DF,  p-value: < 2.2e-16
```

For tax

```
m23_tax<-lm(crim~tax+I(tax^2)+I(tax^3),data = Boston)
summary(m23_tax)
```

```
##
## Call:
## lm(formula = crim ~ tax + I(tax^2) + I(tax^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.273  -1.389   0.046   0.536  76.950
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.918e+01  1.180e+01   1.626   0.105
## tax          -1.533e-01  9.568e-02  -1.602   0.110
## I(tax^2)      3.608e-04  2.425e-04   1.488   0.137
## I(tax^3)     -2.204e-07  1.889e-07  -1.167   0.244
##
## Residual standard error: 6.854 on 502 degrees of freedom
## Multiple R-squared:  0.3689, Adjusted R-squared:  0.3651
## F-statistic: 97.8 on 3 and 502 DF,  p-value: < 2.2e-16
```

For ptratio

```
m24_ptratio<-lm(crim~ptratio+I(ptratio^2)+I(ptratio^3),data = Boston)
summary(m24_ptratio)
```

```
##
## Call:
## lm(formula = crim ~ ptratio + I(ptratio^2) + I(ptratio^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.833  -4.146  -1.655   1.408  82.697
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  477.18405  156.79498   3.043  0.00246 **
## ptratio      -82.36054   27.64394  -2.979  0.00303 **
## I(ptratio^2)   4.63535    1.60832   2.882  0.00412 **
## I(ptratio^3)  -0.08476    0.03090  -2.743  0.00630 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.122 on 502 degrees of freedom
## Multiple R-squared:  0.1138, Adjusted R-squared:  0.1085
## F-statistic: 21.48 on 3 and 502 DF,  p-value: 4.171e-13
```

For black

```
m24_black<-lm(crim~black+I(black^2)+I(black^3),data = Boston)
summary(m24_black)
```



```
##
## Call:
## lm(formula = crim ~ black + I(black^2) + I(black^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.096  -2.343  -2.128  -1.439   86.790
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.826e+01  2.305e+00   7.924  1.5e-14 ***
## black        -8.356e-02  5.633e-02  -1.483   0.139
## I(black^2)    2.137e-04  2.984e-04   0.716   0.474
## I(black^3)   -2.652e-07  4.364e-07  -0.608   0.544
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.955 on 502 degrees of freedom
## Multiple R-squared:  0.1498, Adjusted R-squared:  0.1448
## F-statistic: 29.49 on 3 and 502 DF,  p-value: < 2.2e-16
```

For lstat

```
m25_lstat<-lm(crim~lstat+I(lstat^2)+I(lstat^3),data = Boston)
summary(m25_lstat)
```

```
##
## Call:
## lm(formula = crim ~ lstat + I(lstat^2) + I(lstat^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.234  -2.151  -0.486   0.066   83.353
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.2009656  2.0286452   0.592   0.5541
## lstat        -0.4490656  0.4648911  -0.966   0.3345
## I(lstat^2)    0.0557794  0.0301156   1.852   0.0646 .
## I(lstat^3)   -0.0008574  0.0005652  -1.517   0.1299
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.629 on 502 degrees of freedom
## Multiple R-squared:  0.2179, Adjusted R-squared:  0.2133
## F-statistic: 46.63 on 3 and 502 DF,  p-value: < 2.2e-16
```

For medv

```
m26_medv<-lm(crim~medv+I(medv^2)+I(medv^3),data = Boston)
summary(m26_medv)
```

```
##
## Call:
## lm(formula = crim ~ medv + I(medv^2) + I(medv^3), data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -24.427  -1.976  -0.437   0.439  73.655
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 53.1655381  3.3563105  15.840  < 2e-16 ***
## medv        -5.0948305  0.4338321 -11.744  < 2e-16 ***
## I(medv^2)    0.1554965  0.0171904   9.046  < 2e-16 ***
## I(medv^3)   -0.0014901  0.0002038  -7.312 1.05e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.569 on 502 degrees of freedom
## Multiple R-squared:  0.4202, Adjusted R-squared:  0.4167
## F-statistic: 121.3 on 3 and 502 DF,  p-value: < 2.2e-16
```

Conclusion:

There is evidence of non-linear association for zn as its squared term is statistically significant at .1, for indus as its squared and cubed terms are statistically significant at .001, for nox as its squared and cubed terms are statistically significant at .001, for age as its squared term is statistically significant at .05 and its cubed term is statistically significant at .01, for dis as its squared and cubed terms are statistically significant at .001, for pratio as its squared and cubed terms are statistically significant at .01, for lstat as its squared term is statistically significant at .1 and for nox as its squared and cubed terms are statistically significant at .001