Module_2 applied Que

#8

In [23]:
```
#Module_2 Applied
#solution to que 8
library(MASS)
library(ISLR)
```

In [95]:
```
install.packages("ISLR")
```

```
Updating HTML index of packages in '.Library'
Making 'packages.html' ... done
```

In [25]:
```
Auto = read.csv("/Users/priyanka/desktop/Auto.csv", header = T, na.strings =
Auto = na.omit(Auto)
dim(Auto)
summary(Auto)
```

1. 392
2. 9

```
      mpg          cylinders      displacement     horsepower        weight    
 Min.   : 9.00   Min.   :3.000   Min.   : 68.0   Min.   : 46.0   Min.   :1613  
 1st Qu.:17.00   1st Qu.:4.000   1st Qu.:105.0   1st Qu.: 75.0   1st Qu.:2225  
 Median :22.75   Median :4.000   Median :151.0   Median : 93.5   Median :2804  
 Mean   :23.45   Mean   :5.472   Mean   :194.4   Mean   :104.5   Mean   :2978  
 3rd Qu.:29.00   3rd Qu.:8.000   3rd Qu.:275.8   3rd Qu.:126.0   3rd Qu.:3615  
 Max.   :46.60   Max.   :8.000   Max.   :455.0   Max.   :230.0   Max.   :5140  
                                                                               
  acceleration        year           origin                      name    
 Min.   : 8.00   Min.   :70.00   Min.   :1.000   amc matador       :  5  
 1st Qu.:13.78   1st Qu.:73.00   1st Qu.:1.000   ford pinto        :  5  
 Median :15.50   Median :76.00   Median :1.000   toyota corolla    :  5  
 Mean   :15.54   Mean   :75.98   Mean   :1.577   amc gremlin       :  4  
 3rd Qu.:17.02   3rd Qu.:79.00   3rd Qu.:2.000   amc hornet        :  4  
 Max.   :24.80   Max.   :82.00   Max.   :3.000   chevrolet chevette:  4  
                                                 (Other)           :365  
```

8(a)

In [26]:
```
#8(a)
data(Auto)
lm.fit = lm(mpg~horsepower,data=Auto)
summary(lm.fit)
```

```
Call:
lm(formula = mpg ~ horsepower, data = Auto)

Residuals:
     Min       1Q   Median       3Q      Max 
-13.5710  -3.2592  -0.3435   2.7630  16.9240 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 39.935861   0.717499   55.66   <2e-16 ***
horsepower  -0.157845   0.006446  -24.49   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.906 on 390 degrees of freedom
Multiple R-squared:  0.6059,	Adjusted R-squared:  0.6049 
F-statistic: 599.7 on 1 and 390 DF,  p-value: < 2.2e-16
```

In [28]:

```
#8(a)
#(i) yes, there is a relationship which can be calculated by
#testing null-hypothesis
#F-stat is large and p-value is close to zero, so we reject null-hypothesis.
#there is a significant relation b/w horsepower and mpg is significant

#(ii) as the p-value is close to 0, so relationship between
#predictor and response is strong

#(iii) the relationship b/w mpg and horsepower is negetive. the more horsepow
# mpg fuel efficiency of a vehicle will have.

#(iv)
predict(lm.fit, data.frame(horsepower=c(98)), interval="confidence", level=0.
```

fit	lwr	upr
24.46708	23.97308	24.96108

In [29]:

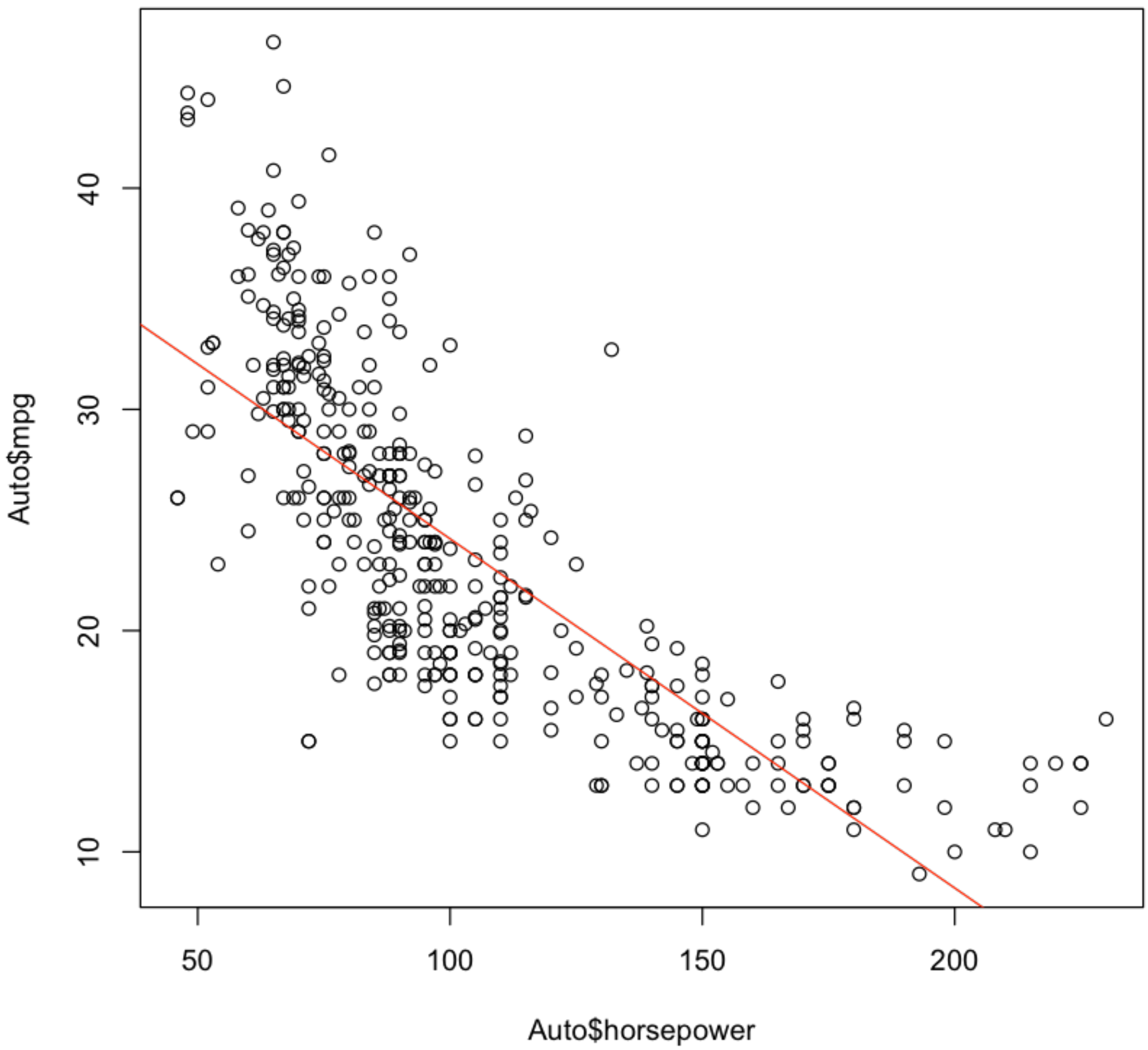
```
predict(lm.fit, data.frame(horsepower=c(98)), interval="prediction", level=0.
```

fit	lwr	upr
24.46708	14.8094	34.12476

#8(b)

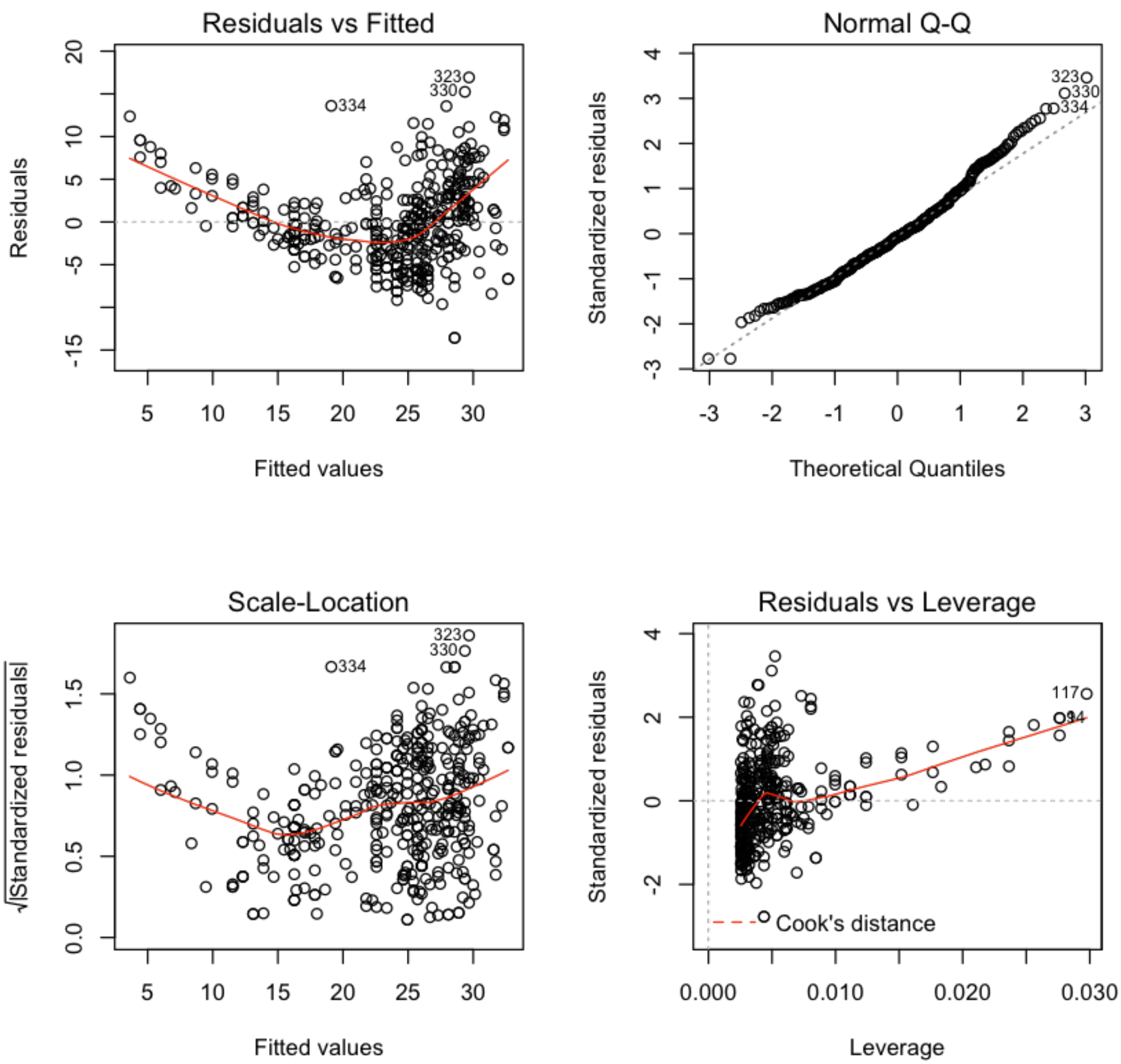
In [17]:

```
#8(b)
plot(Auto$horsepower, Auto$mpg)
#abline(lm.fit)
abline(lm.fit,col='red')
```



#8(c)

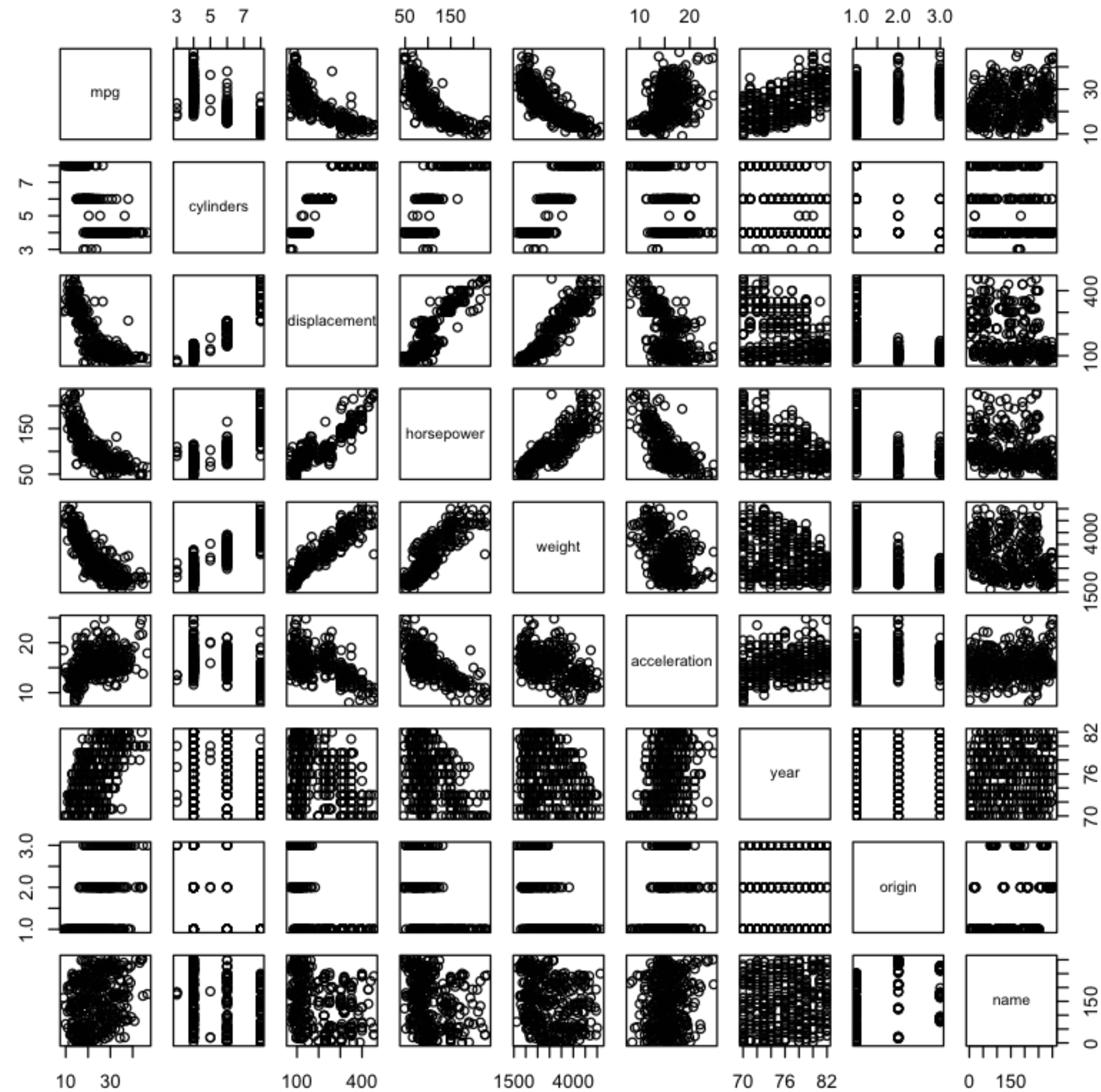
In [31]:
```
#8(c)
par(mfrow=c(2,2))
plot(lm.fit)

# residuals vs fitted value shows the non-linear relation.
```



#9(a)

In [19]:

```
#solution to que. 9(a)
pairs(Auto)
```



#9(b)

In [20]:
```
#9(b)
cor(subset(Auto, select=-name))
```

	mpg	cylinders	displacement	horsepower	weight	acceleration
mpg	1.0000000	-0.7776175	-0.8051269	-0.7784268	-0.8322442	0.4233285
cylinders	-0.7776175	1.0000000	0.9508233	0.8429834	0.8975273	-0.5046834
displacement	-0.8051269	0.9508233	1.0000000	0.8972570	0.9329944	-0.5438005
horsepower	-0.7784268	0.8429834	0.8972570	1.0000000	0.8645377	-0.6891955
weight	-0.8322442	0.8975273	0.9329944	0.8645377	1.0000000	-0.4168392
acceleration	0.4233285	-0.5046834	-0.5438005	-0.6891955	-0.4168392	1.0000000
year	0.5805410	-0.3456474	-0.3698552	-0.4163615	-0.3091199	0.2903161
origin	0.5652088	-0.5689316	-0.6145351	-0.4551715	-0.5850054	0.2127458

In [21]:

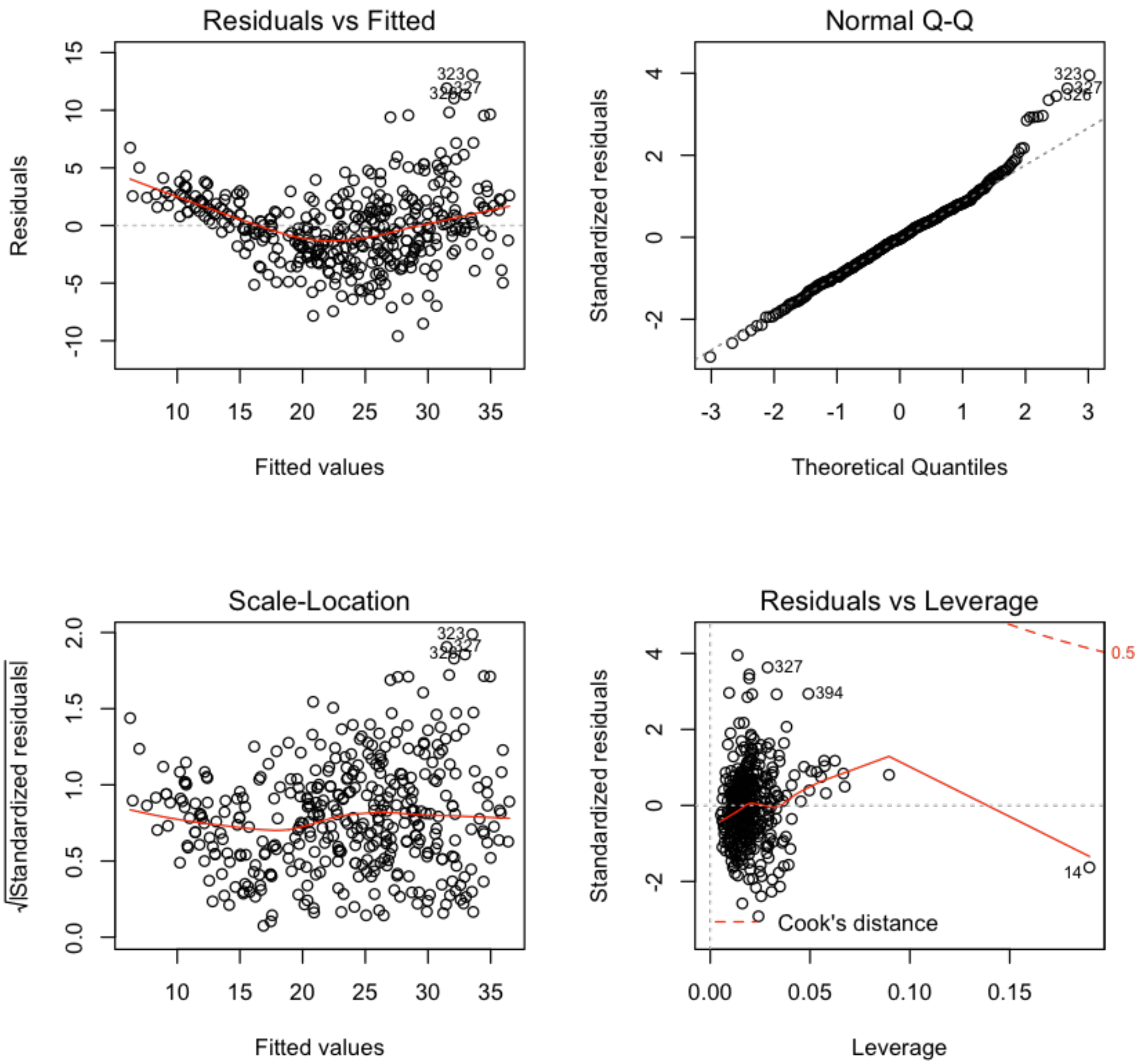
```
#9(c)
lm.fit_1 = lm(mpg~.-name, data=Auto)
summary(lm.fit_1)
```

```
Call:
lm(formula = mpg ~ . - name, data = Auto)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.5903 -2.1565 -0.1169  1.8690 13.0604 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -17.218435   4.644294  -3.707  0.00024 ***
cylinders     -0.493376   0.323282  -1.526  0.12780    
displacement   0.019896   0.007515   2.647  0.00844 ** 
horsepower    -0.016951   0.013787  -1.230  0.21963    
weight        -0.006474   0.000652  -9.929  < 2e-16 ***
acceleration   0.080576   0.098845   0.815  0.41548    
year           0.750773   0.050973  14.729  < 2e-16 ***
origin         1.426141   0.278136   5.127 4.67e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.328 on 384 degrees of freedom
Multiple R-squared:  0.8215,	Adjusted R-squared:  0.8182 
F-statistic: 252.4 on 7 and 384 DF,  p-value: < 2.2e-16
```

In [ ]:

```
#9 explanation
#(i):yes there is a relation between predictor and resp. F-stat>1 and
# p-value is smaller

#(ii):from p-values of predictors:disp, weight,year and orgin have significan
# relation.

#(iii)reg coef for year is 0.7507- means with every year cars are becoming
#more fuel efficient
```

In [23]:
```
#9(d)
par(mfrow=c(2,2))
plot(lm.fit_1)
```



In [24]:
```
#residuals are non linear fit
#between 30-35 mpg, there are some high residuals (323,327,328)
#from leverabge plot, 14 has the high leverage
```

In [25]:
```
#9(e)
lm.fit_2 = lm(mpg~cylinders*displacement+displacement*weight, data=Auto)
summary(lm.fit_2)
```

```
Call:
lm(formula = mpg ~ cylinders * displacement + displacement *
    weight, data = Auto)

Residuals:
     Min       1Q   Median       3Q      Max
-13.2934  -2.5184  -0.3476   1.8399  17.7723

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              5.262e+01  2.237e+00  23.519  < 2e-16 ***
cylinders                7.606e-01  7.669e-01   0.992    0.322
displacement            -7.351e-02  1.669e-02  -4.403 1.38e-05 ***
weight                  -9.888e-03  1.329e-03  -7.438 6.69e-13 ***
cylinders:displacement -2.986e-03  3.426e-03  -0.872    0.384
displacement:weight      2.128e-05  5.002e-06   4.254 2.64e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.103 on 386 degrees of freedom
Multiple R-squared:  0.7272,    Adjusted R-squared:  0.7237
F-statistic: 205.8 on 5 and 386 DF,  p-value: < 2.2e-16
```

In [38]:

```
lm.fit_3 = lm(mpg~displacement+origin+year*weight, data=Auto)
summary(lm.fit_3)
```

```
Call:
lm(formula = mpg ~ displacement + origin + year * weight, data = Auto)

Residuals:
    Min      1Q  Median      3Q     Max
-8.9402 -1.8736 -0.0966  1.5924 12.2125

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.076e+02  1.290e+01  -8.339 1.34e-15 ***
displacement -4.020e-04  4.558e-03  -0.088 0.929767
origin        9.116e-01  2.547e-01   3.579 0.000388 ***
year          1.962e+00  1.716e-01  11.436  < 2e-16 ***
weight        2.605e-02  4.552e-03   5.722 2.12e-08 ***
year:weight  -4.305e-04  5.967e-05  -7.214 2.89e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.145 on 386 degrees of freedom
Multiple R-squared:  0.8397,    Adjusted R-squared:  0.8376
F-statistic: 404.4 on 5 and 386 DF,  p-value: < 2.2e-16
```

In [27]:

```
# lm.fit_2,from the p-values we can see the interaction between
# cylinders and disp not significant, disp-weight has significant interaction
# lm.fit_3, from year-weight has signif interaction
#interaction b/w cylinders and horsepower is signficant
#horsepower and weight not significant
```

In [37]:
```
lm.fit_3 = lm(mpg~cylinders*horsepower+horsepower*weight+displacement*weight+
summary(lm.fit_3)

#interactions b/w weight and year
```

```
Call:
lm(formula = mpg ~ cylinders * horsepower + horsepower * weight +
    displacement * weight + origin + year * weight, data = Auto)

Residuals:
    Min      1Q  Median      3Q     Max
-8.8816 -1.5616 -0.0744  1.2508 12.2585

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)          -5.224e+01  1.464e+01  -3.569 0.000404 ***
cylinders            -1.940e+00  8.024e-01  -2.417 0.016115 *
horsepower           -1.549e-01  4.342e-02  -3.567 0.000407 ***
weight                1.266e-02  5.237e-03   2.417 0.016103 *
displacement         -4.167e-02  1.883e-02  -2.212 0.027541 *
origin                6.262e-01  2.565e-01   2.441 0.015097 *
year                  1.467e+00  1.788e-01   8.203 3.64e-15 ***
cylinders:horsepower  2.025e-02  7.222e-03   2.804 0.005301 **
horsepower:weight    -7.173e-06  1.590e-05  -0.451 0.652212
weight:displacement   1.258e-05  5.089e-06   2.473 0.013852 *
weight:year          -2.580e-04  6.381e-05  -4.044 6.36e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.841 on 381 degrees of freedom
Multiple R-squared:  0.8709,    Adjusted R-squared:  0.8675
F-statistic:   257 on 10 and 381 DF,  p-value: < 2.2e-16
```

In [49]:
```
#9(f)
lm.fit_4 = lm(mpg~cylinders*horsepower+sqrt(horsepower)*sqrt(weight)+displace
#lm.fit_4 = lm(mpg~log(acceleration)+(horsepower^2)+(year^2)+sqrt(weight)+(or
summary(lm.fit_4)

#we did no find any significant improvement
```

```
Call:
lm(formula = mpg ~ cylinders * horsepower + sqrt(horsepower) *
    sqrt(weight) + displacement * weight + origin + year * weight,
    data = Auto)

Residuals:
    Min      1Q  Median      3Q     Max
-9.0352 -1.5352  0.0034  1.3336 12.3590

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                   -7.793e+01  2.895e+01  -2.692  0.00742 **
cylinders                     -8.149e-01  1.037e+00  -0.786  0.43248
horsepower                     1.391e-01  1.568e-01   0.887  0.37549
sqrt(horsepower)               7.192e-01  2.471e+00   0.291  0.77116
sqrt(weight)                   7.517e-01  9.403e-01   0.799  0.42453
displacement                  -7.354e-02  2.754e-02  -2.670  0.00791 **
weight                         1.560e-02  1.150e-02   1.356  0.17580
origin                         5.128e-01  2.576e-01   1.990  0.04726 *
year                           1.516e+00  1.774e-01   8.547 3.13e-16 ***
cylinders:horsepower           1.279e-02  9.115e-03   1.404  0.16125
sqrt(horsepower):sqrt(weight) -1.159e-01  5.790e-02  -2.002  0.04599 *
displacement:weight            2.054e-05  7.386e-06   2.781  0.00569 **
weight:year                   -2.748e-04  6.319e-05  -4.350 1.76e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.828 on 379 degrees of freedom
Multiple R-squared:  0.8728,    Adjusted R-squared:  0.8687
F-statistic: 216.7 on 12 and 379 DF,  p-value: < 2.2e-16
```

In [47]:

```
#10
data(Carseats)
summary(Carseats)
```

```
     Sales          CompPrice       Income        Advertising
 Min.   : 0.000   Min.   : 77   Min.   : 21.00   Min.   : 0.000
 1st Qu.: 5.390   1st Qu.:115   1st Qu.: 42.75   1st Qu.: 0.000
 Median : 7.490   Median :125   Median : 69.00   Median : 5.000
 Mean   : 7.496   Mean   :125   Mean   : 68.66   Mean   : 6.635
 3rd Qu.: 9.320   3rd Qu.:135   3rd Qu.: 91.00   3rd Qu.:12.000
 Max.   :16.270   Max.   :175   Max.   :120.00   Max.   :29.000
   Population        Price        ShelveLoc        Age          Education
 Min.   : 10.0   Min.   : 24.0   Bad   : 96   Min.   :25.00   Min.   :10.0
 1st Qu.:139.0   1st Qu.:100.0   Good  : 85   1st Qu.:39.75   1st Qu.:12.0
 Median :272.0   Median :117.0   Medium:219   Median :54.50   Median :14.0
 Mean   :264.8   Mean   :115.8                Mean   :53.32   Mean   :13.9
 3rd Qu.:398.5   3rd Qu.:131.0                3rd Qu.:66.00   3rd Qu.:16.0
 Max.   :509.0   Max.   :191.0                Max.   :80.00   Max.   :18.0
 Urban       US
 No :118   No :142
 Yes:282   Yes:258
```

In [50]:
```
attach(Carseats)
#attach() fun is used to access the variables present in the data
#framework without calling the dataframe
library(MASS)
library(ISLR)
```

In [55]:
```
?Carseats
```

In [33]:
```
#10(a)
lm.fit = lm(Sales~Price+Urban+US)
summary(lm.fit)
```

```
Call:
lm(formula = Sales ~ Price + Urban + US)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.9206 -1.6220 -0.0564  1.5786  7.0581 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 13.043469   0.651012  20.036  < 2e-16 ***
Price       -0.054459   0.005242 -10.389  < 2e-16 ***
UrbanYes    -0.021916   0.271650  -0.081    0.936    
USYes        1.200573   0.259042   4.635 4.86e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.472 on 396 degrees of freedom
Multiple R-squared:  0.2393,	Adjusted R-squared:  0.2335 
F-statistic: 41.52 on 3 and 396 DF,  p-value: < 2.2e-16
```

In [34]:
```
#10(b) explanation
#price- significant - sales droped by 54 for each $1000 increase
#UrbanYes- sales are lower for urban locations - not significant
#USYes- sales are higher in USlocations - significant relation

#10(c)
#sales= 13.0434-0.5445*Price-0.02191*(UrbanYes)+1.2005*USYes

#10(d)
#Null-hypothisis for predictors- Price and USYes can reject.
```

In [35]:
```
#10(e)
lm.fit_n = lm(Sales ~ Price + US)
summary(lm.fit_n)
```

```
Call:
lm(formula = Sales ~ Price + US)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.9269 -1.6286 -0.0574  1.5766  7.0515 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 13.03079    0.63098  20.652  < 2e-16 ***
Price       -0.05448    0.00523 -10.416  < 2e-16 ***
USYes        1.19964    0.25846   4.641 4.71e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.469 on 397 degrees of freedom
Multiple R-squared:  0.2393,	Adjusted R-squared:  0.2354 
F-statistic: 62.43 on 2 and 397 DF,  p-value: < 2.2e-16
```

In [ ]:

```
#10(f)
#In case of lm.fit_n, data fit better than lm.fit. Although values of RMS and
#R-square are almost equal but lm.fit_n has slightly less RMS than
#lm.fit and no of predictors are also less than lm.fit
```

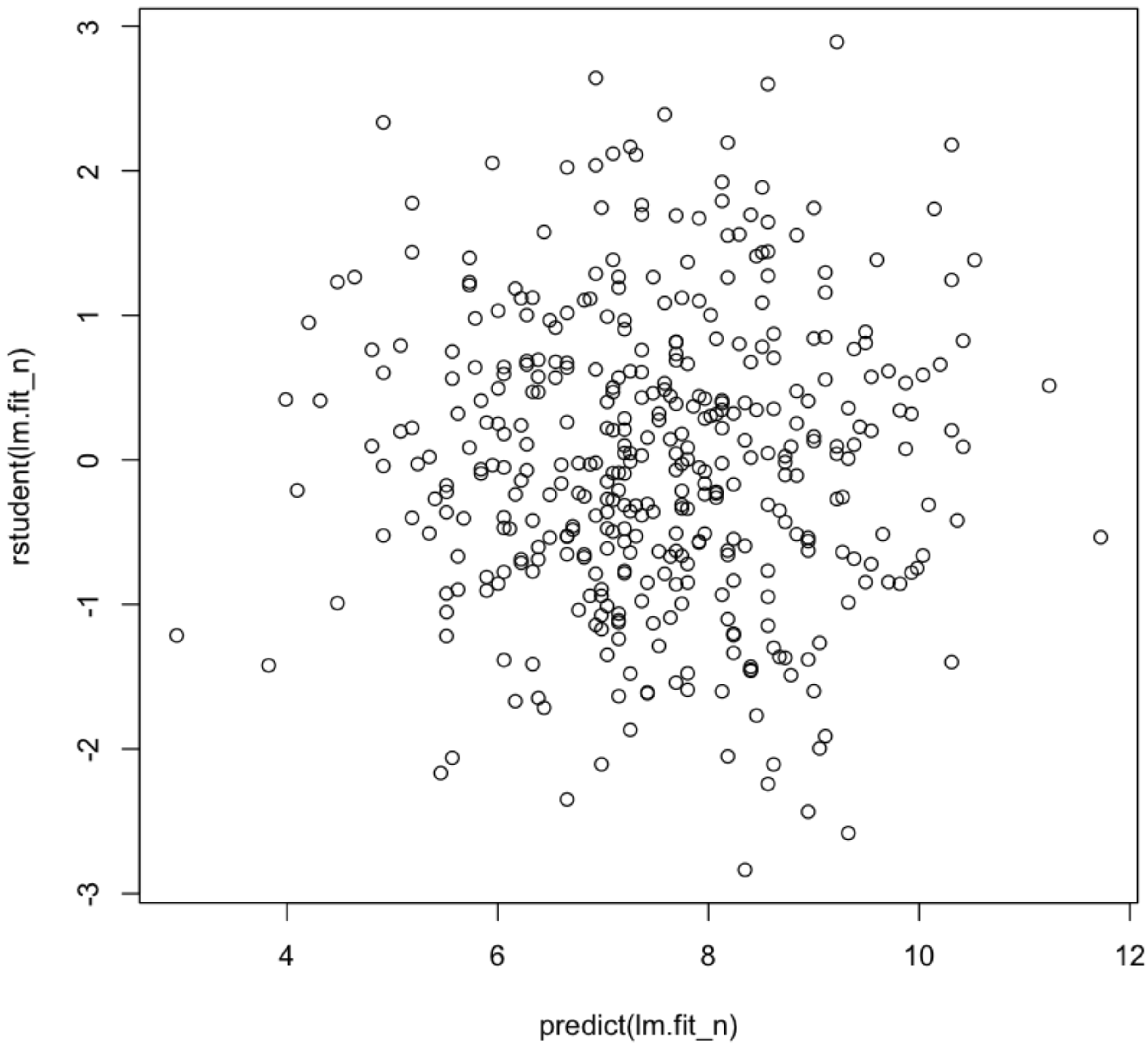
In [36]:

```
#10(g)
#model from 10(e), 95% confidence intervals for the coefficient(s).

confint(lm.fit_n)
```

	2.5 %	97.5 %
(Intercept)	11.79032020	14.27126531
Price	-0.06475984	-0.04419543
USYes	0.69151957	1.70776632

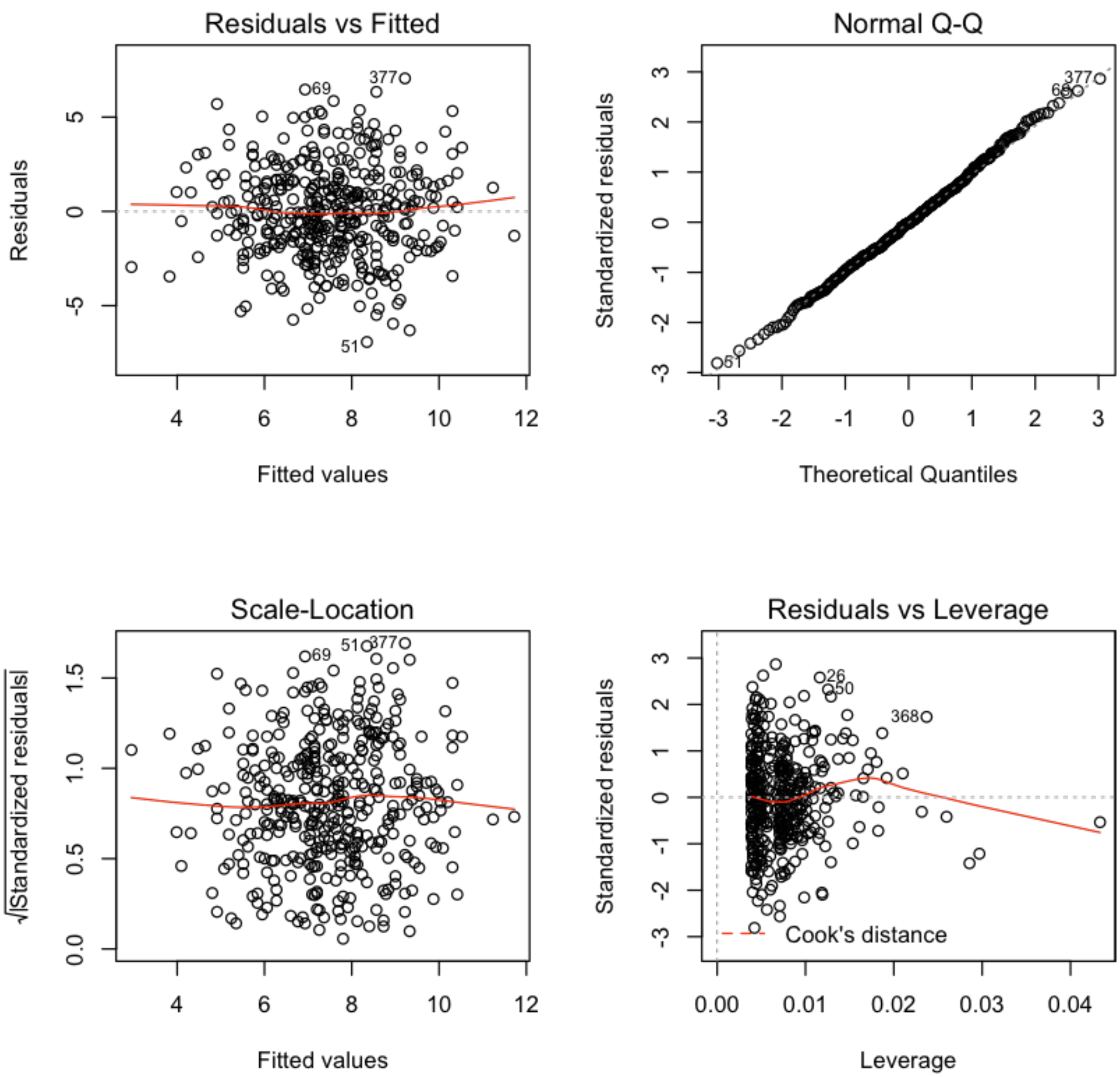
In [37]:

```
#10(h)   //using10(e) model
plot(predict(lm.fit_n), rstudent(lm.fit_n))

#no evidence fo outliners
```



In [38]:

```
#10(h)
par(mfrow=c(2,2))
plot(lm.fit_n)

#from Leverage plot, few value exceeds appox at 0.07 so this plots show
#the high leverge observations.
```



In [2]:
```
#15
library(MASS)
library(ISLR)
```

In [3]:
```
data(Boston)
summary(Boston)
```

```
      crim                zn             indus            chas        
 Min.   : 0.00632   Min.   :  0.00   Min.   : 0.46   Min.   :0.00000  
 1st Qu.: 0.08204   1st Qu.:  0.00   1st Qu.: 5.19   1st Qu.:0.00000  
 Median : 0.25651   Median :  0.00   Median : 9.69   Median :0.00000  
 Mean   : 3.61352   Mean   : 11.36   Mean   :11.14   Mean   :0.06917  
 3rd Qu.: 3.67708   3rd Qu.: 12.50   3rd Qu.:18.10   3rd Qu.:0.00000  
 Max.   :88.97620   Max.   :100.00   Max.   :27.74   Max.   :1.00000  
      nox               rm             age              dis        
 Min.   :0.3850   Min.   :3.561   Min.   :  2.90   Min.   : 1.130  
 1st Qu.:0.4490   1st Qu.:5.886   1st Qu.: 45.02   1st Qu.: 2.100  
 Median :0.5380   Median :6.208   Median : 77.50   Median : 3.207  
 Mean   :0.5547   Mean   :6.285   Mean   : 68.57   Mean   : 3.795  
 3rd Qu.:0.6240   3rd Qu.:6.623   3rd Qu.: 94.08   3rd Qu.: 5.188  
 Max.   :0.8710   Max.   :8.780   Max.   :100.00   Max.   :12.127  
      rad              tax           ptratio          black       
 Min.   : 1.000   Min.   :187.0   Min.   :12.60   Min.   :  0.32  
 1st Qu.: 4.000   1st Qu.:279.0   1st Qu.:17.40   1st Qu.:375.38  
 Median : 5.000   Median :330.0   Median :19.05   Median :391.44  
 Mean   : 9.549   Mean   :408.2   Mean   :18.46   Mean   :356.67  
 3rd Qu.:24.000   3rd Qu.:666.0   3rd Qu.:20.20   3rd Qu.:396.23  
 Max.   :24.000   Max.   :711.0   Max.   :22.00   Max.   :396.90  
     lstat            medv      
 Min.   : 1.73   Min.   : 5.00  
 1st Qu.: 6.95   1st Qu.:17.02  
 Median :11.36   Median :21.20  
 Mean   :12.65   Mean   :22.53  
 3rd Qu.:16.95   3rd Qu.:25.00  
 Max.   :37.97   Max.   :50.00  
```

In [56]:
```
?Boston
```

In [4]:
```
names(Boston)
```

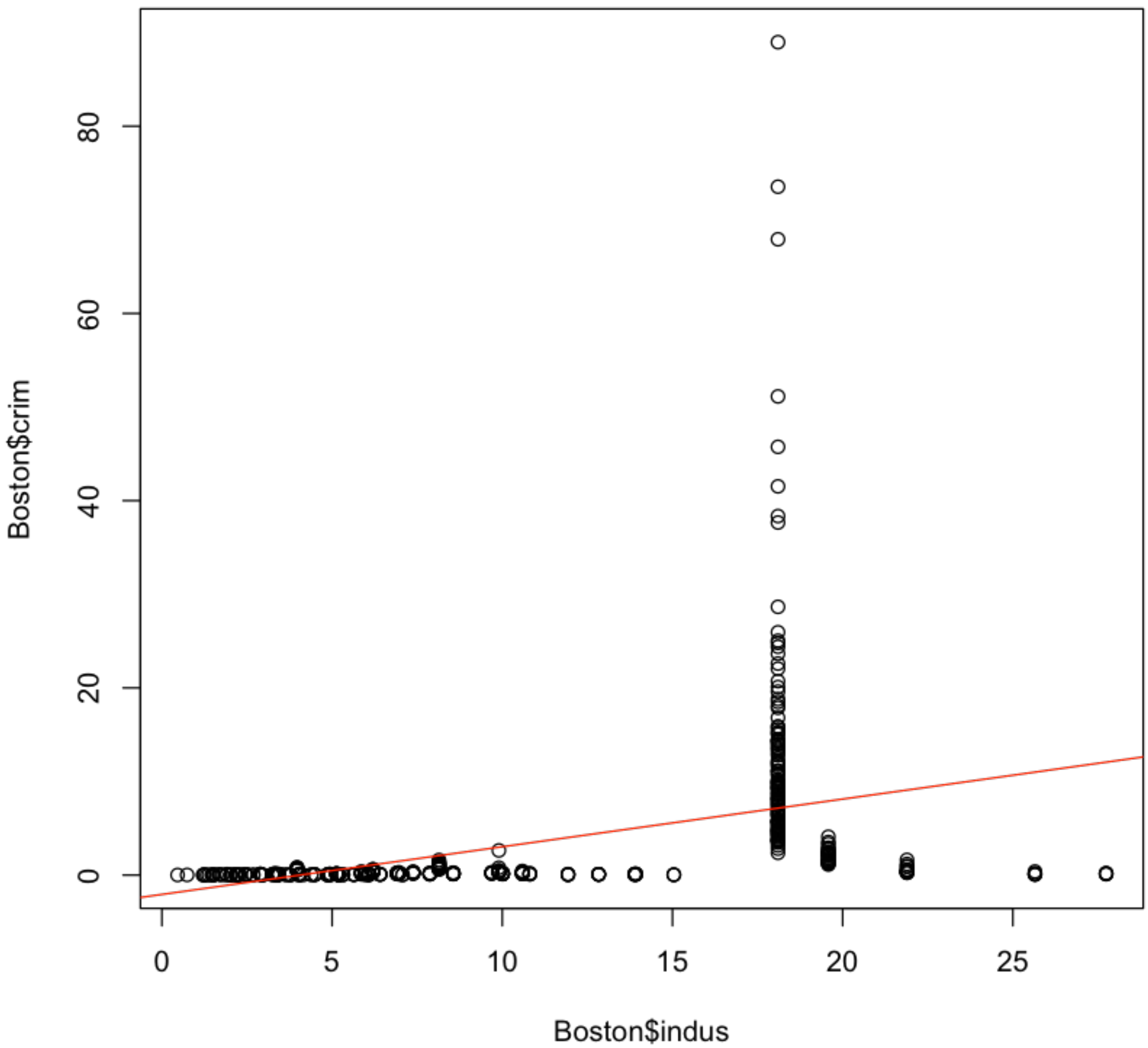
1. 'crim'
2. 'zn'
3. 'indus'
4. 'chas'
5. 'nox'
6. 'rm'
7. 'age'
8. 'dis'
9. 'rad'
10. 'tax'
11. 'ptratio'
12. 'black'
13. 'lstat'
14. 'medv'

In [5]:
```
#15(a)
attach(Boston)
lm.zn = lm(crim~zn)
summary(lm.zn)
```

```
Call:
lm(formula = crim ~ zn)

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
```

In [59]:
```
lm.indus = lm(crim~indus)
summary(lm.indus)
plot(Boston$indus, Boston$crim)
abline(lm.indus, col="red")
```

```
Call:
lm(formula = crim ~ indus)

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
```

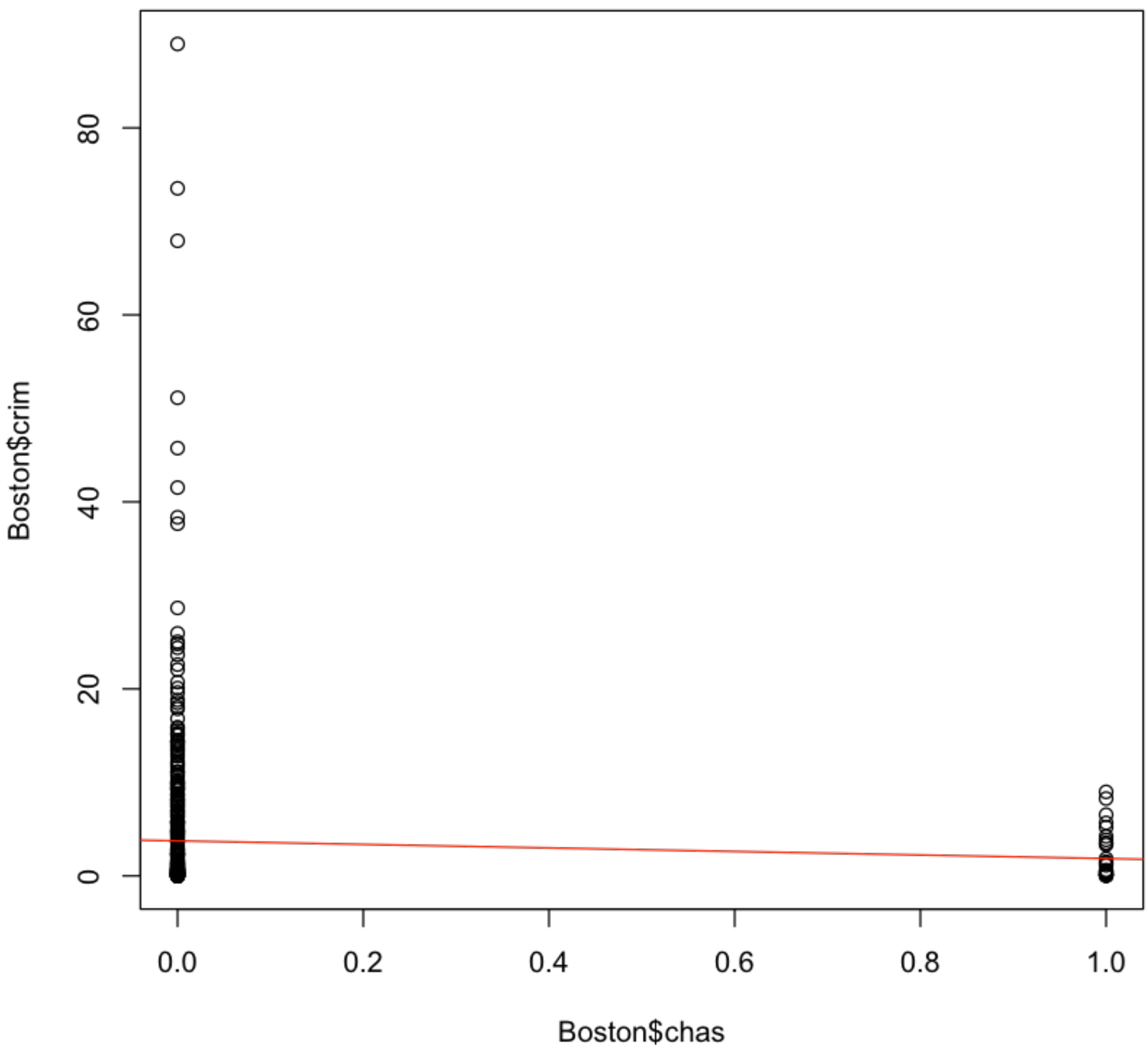


In [60]:

```
lm.chas = lm(crim~chas)
summary(lm.chas)
plot(Boston$chas, Boston$crim)
abline(lm.chas, col="red")
```

```
Call:
lm(formula = crim ~ chas)

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
```

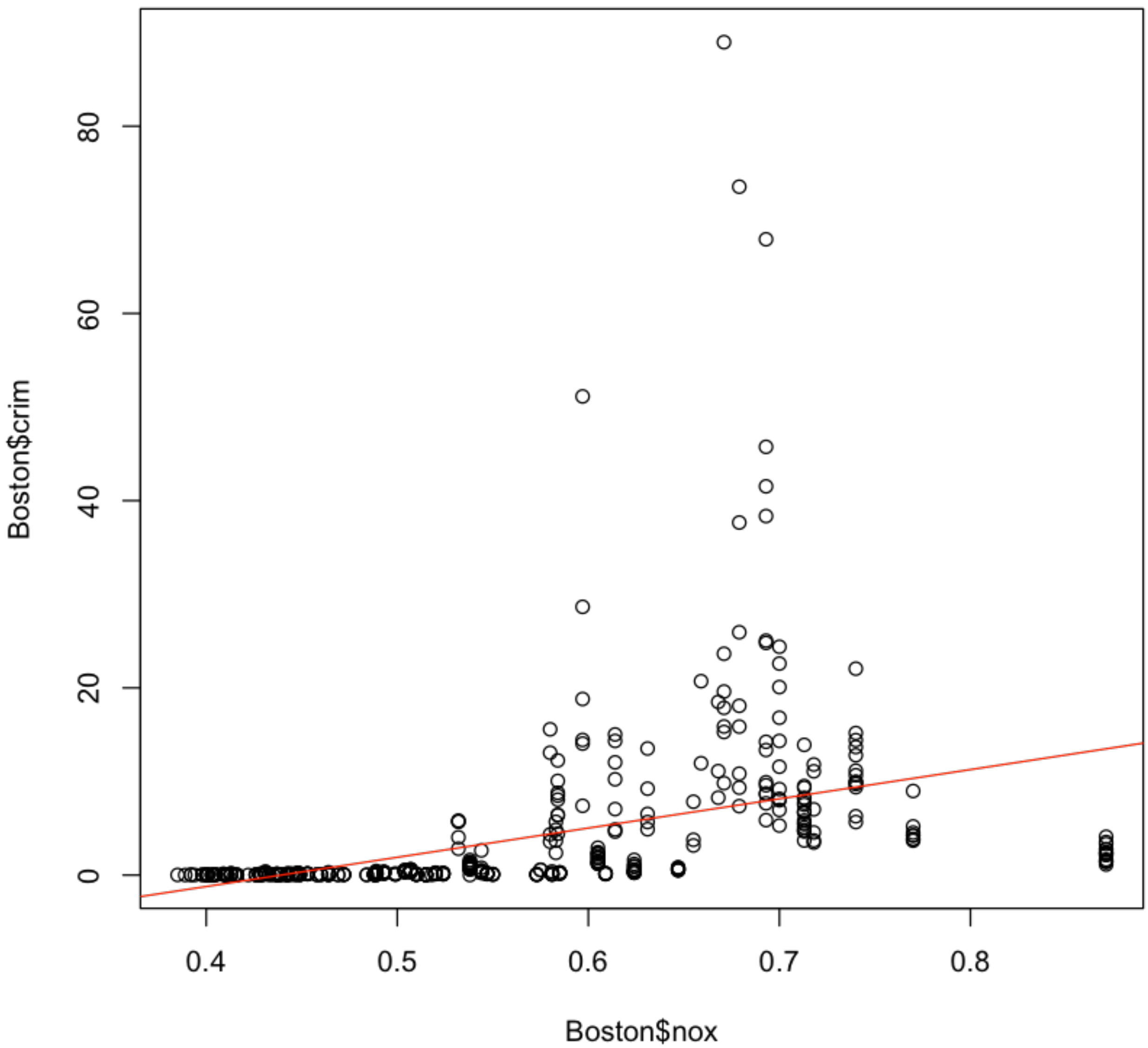


In [61]:

```
lm.nox = lm(crim~nox)
summary(lm.nox)
plot(Boston$nox, Boston$crim)
abline(lm.nox, col="red")
```

```
Call:
```

```
lm(formula = crim ~ nox)

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
```

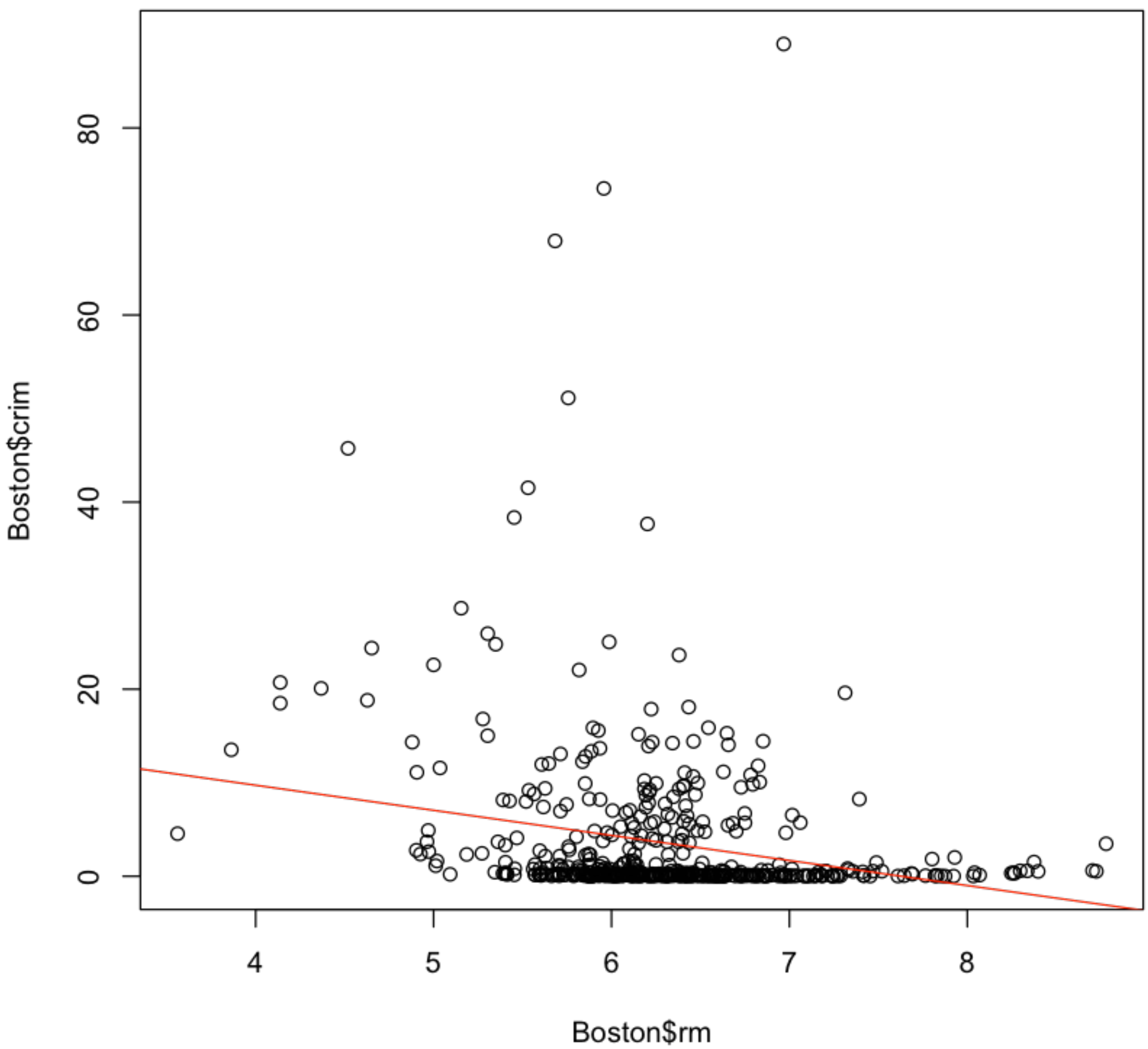


In [62]:

```
lm.rm = lm(crim~rm)
summary(lm.rm)
plot(Boston$rm, Boston$crim)
abline(lm.rm, col="red")
```

```
Call:
lm(formula = crim ~ rm)

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
```

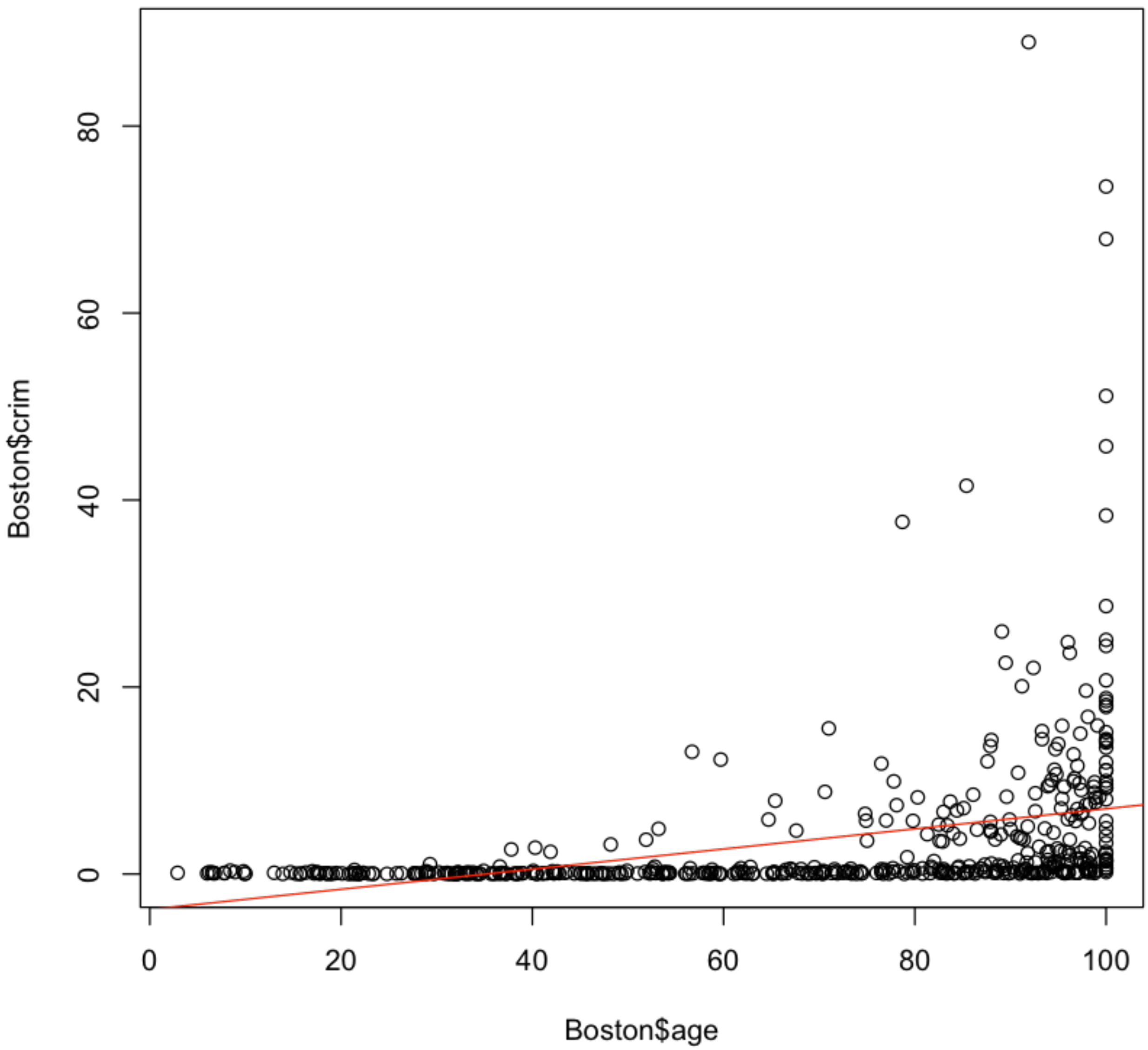


In [63]:

```
lm.age = lm(crim~age)
summary(lm.age)
plot(Boston$age, Boston$crim)
abline(lm.age, col="red")
```

```
Call:
```

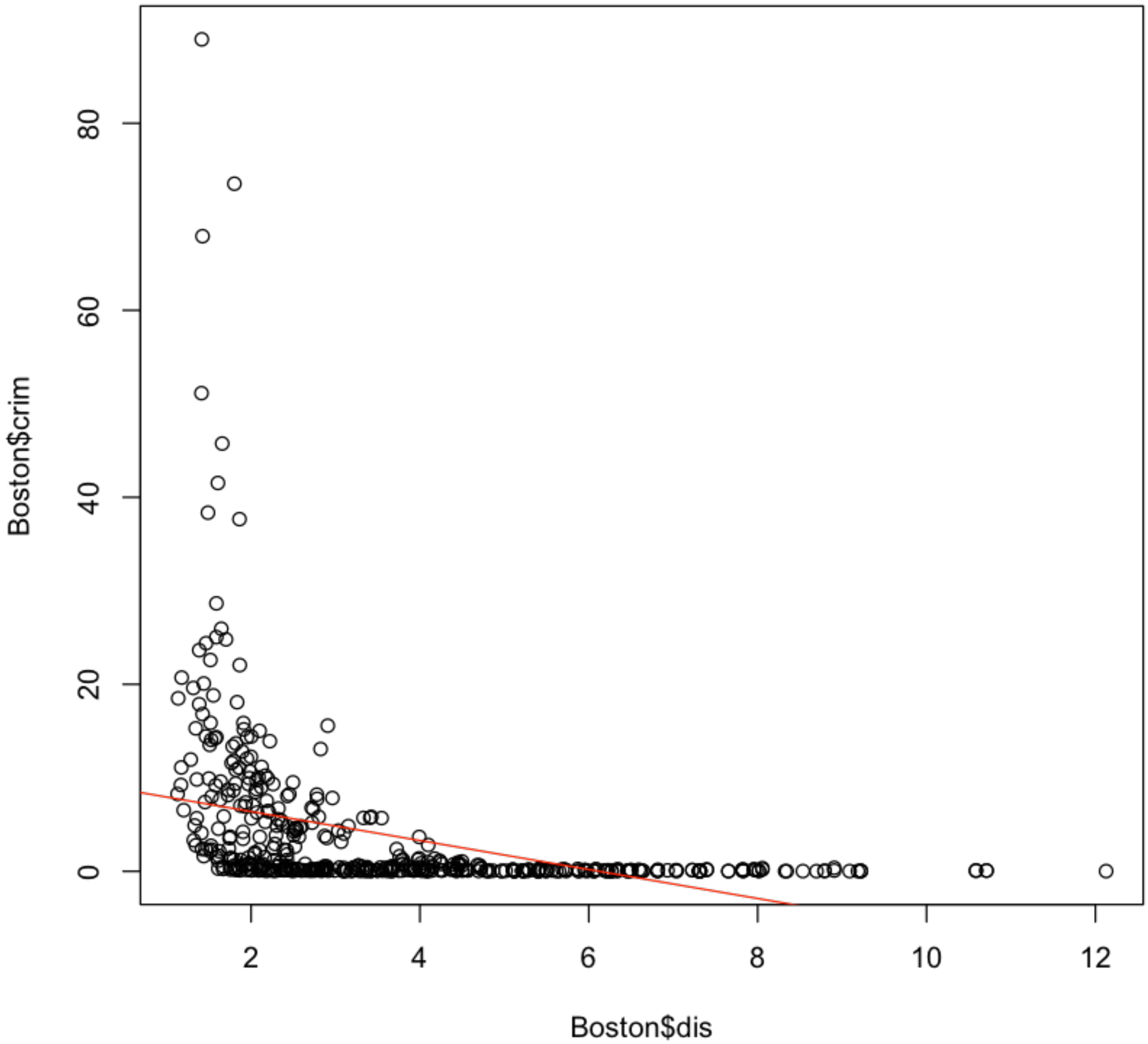
```
lm(formula = crim ~ age)

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
```



In [64]:
```
lm.dis = lm(crim~dis)
summary(lm.dis)
plot(Boston$dis, Boston$crim)
abline(lm.dis, col="red")
```

```
Call:
lm(formula = crim ~ dis)

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
```

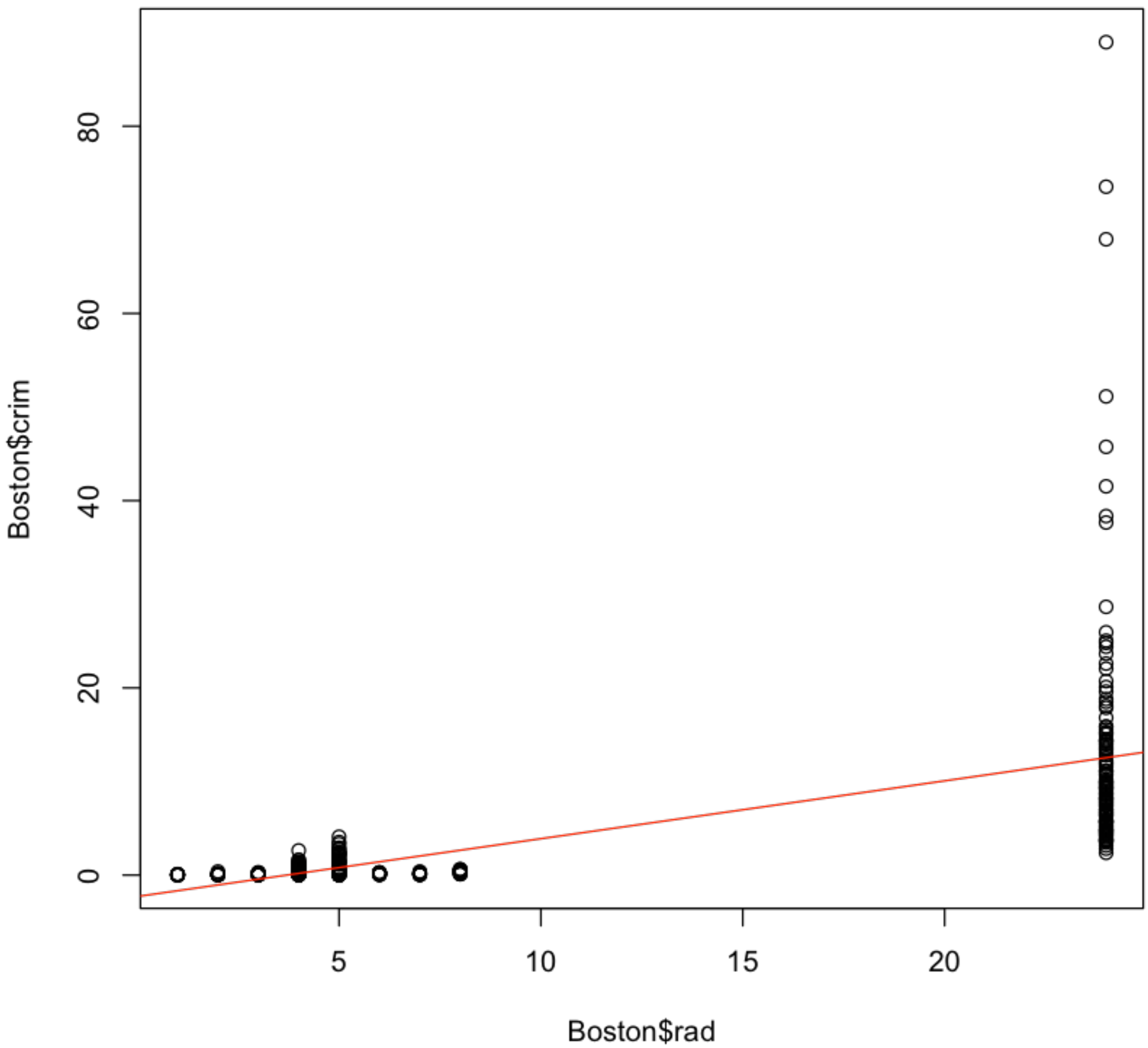


In [65]:

```
lm.rad = lm(crim~rad)
summary(lm.rad)
plot(Boston$rad, Boston$crim)
abline(lm.rad, col="red")
```

```
Call:
```

```
lm(formula = crim ~ rad)

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
```

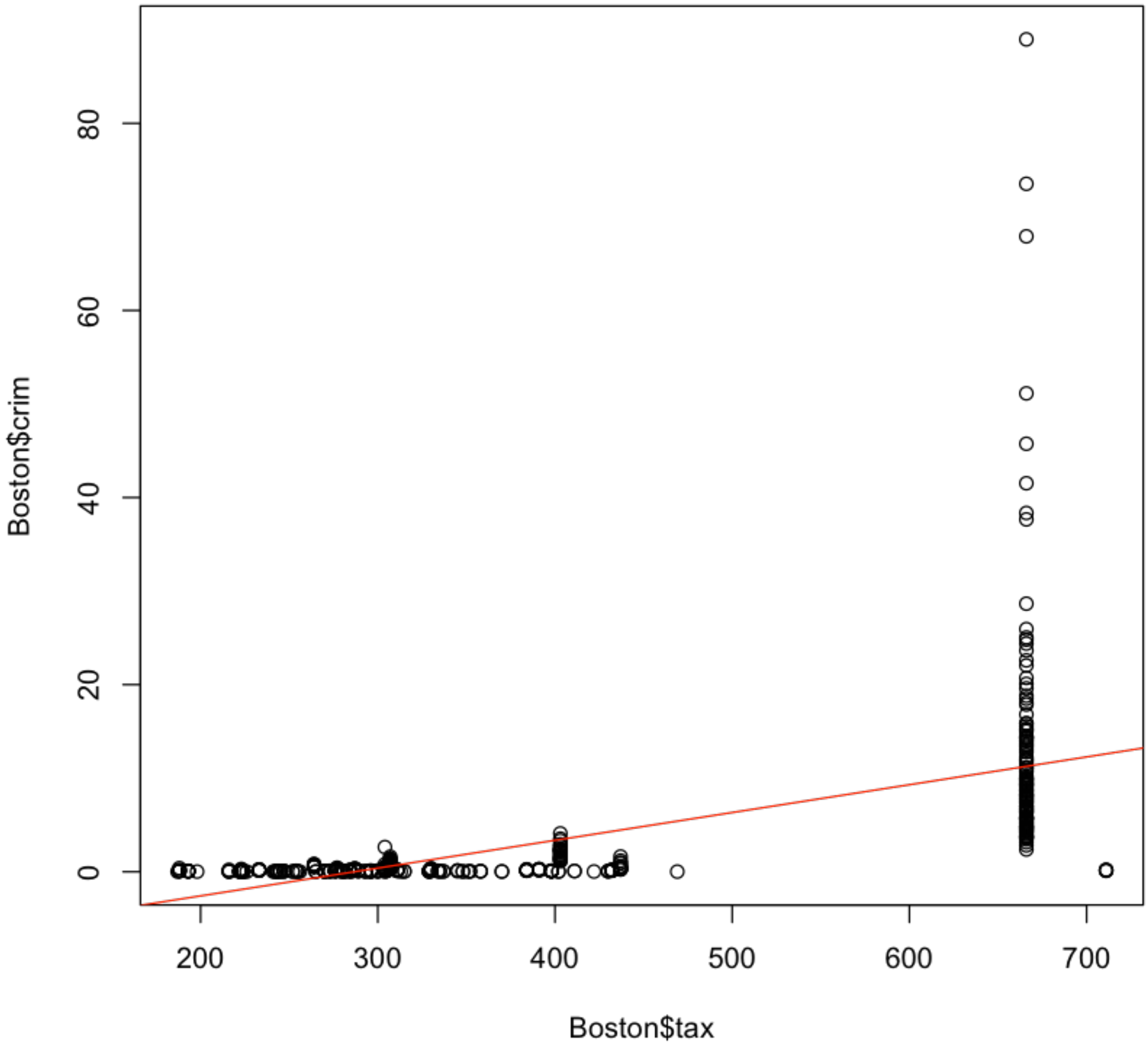


In [66]:

```
lm.tax = lm(crim~tax)
summary(lm.tax)
plot(Boston$tax, Boston$crim)
abline(lm.tax, col="red")
```

```
Call:
lm(formula = crim ~ tax)

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
```

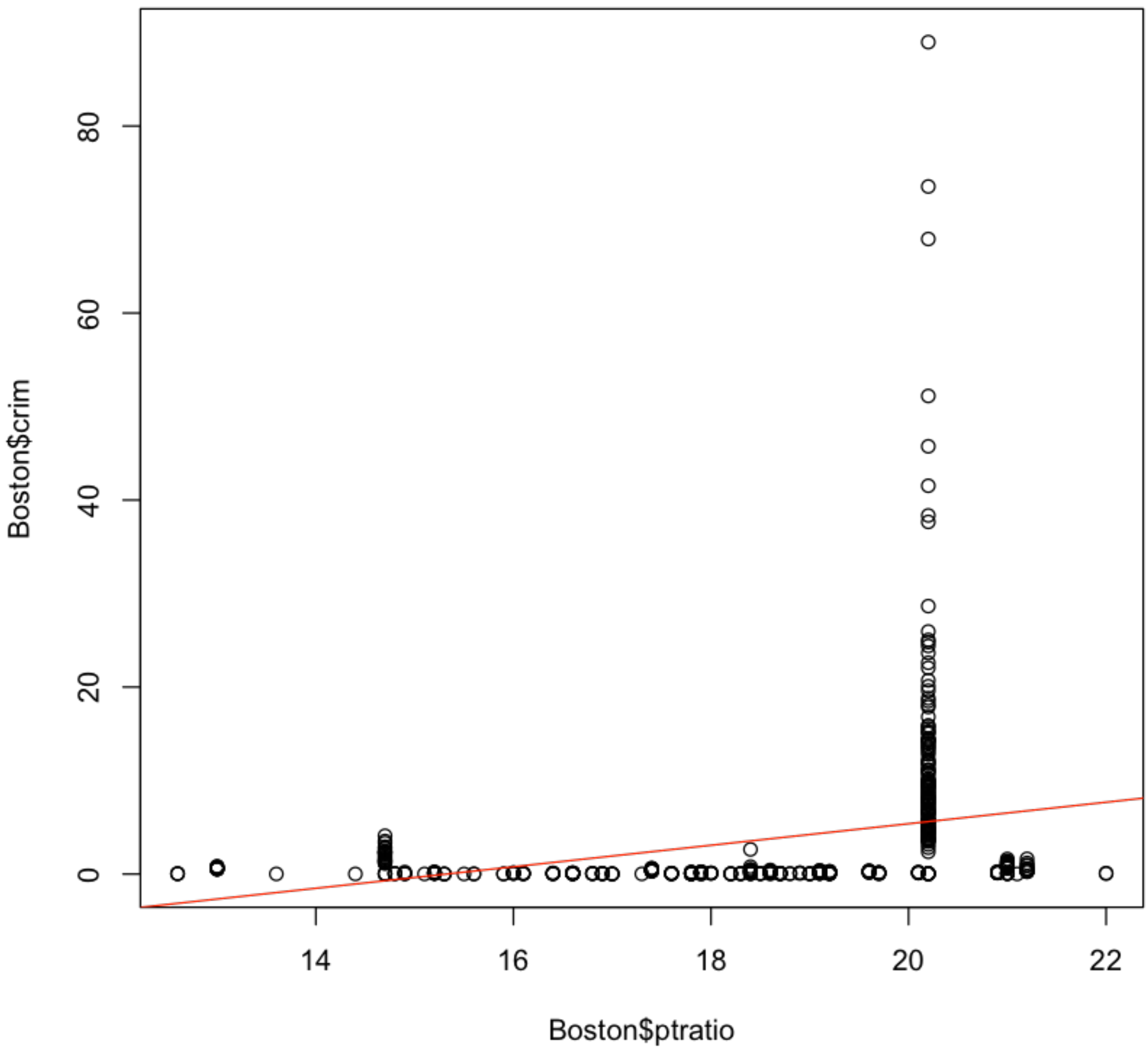


In [67]:

```
lm.ptratio = lm(crim~ptratio)
summary(lm.ptratio)
plot(Boston$ptratio, Boston$crim)
abline(lm.ptratio, col="red")
```

```
Call:
```

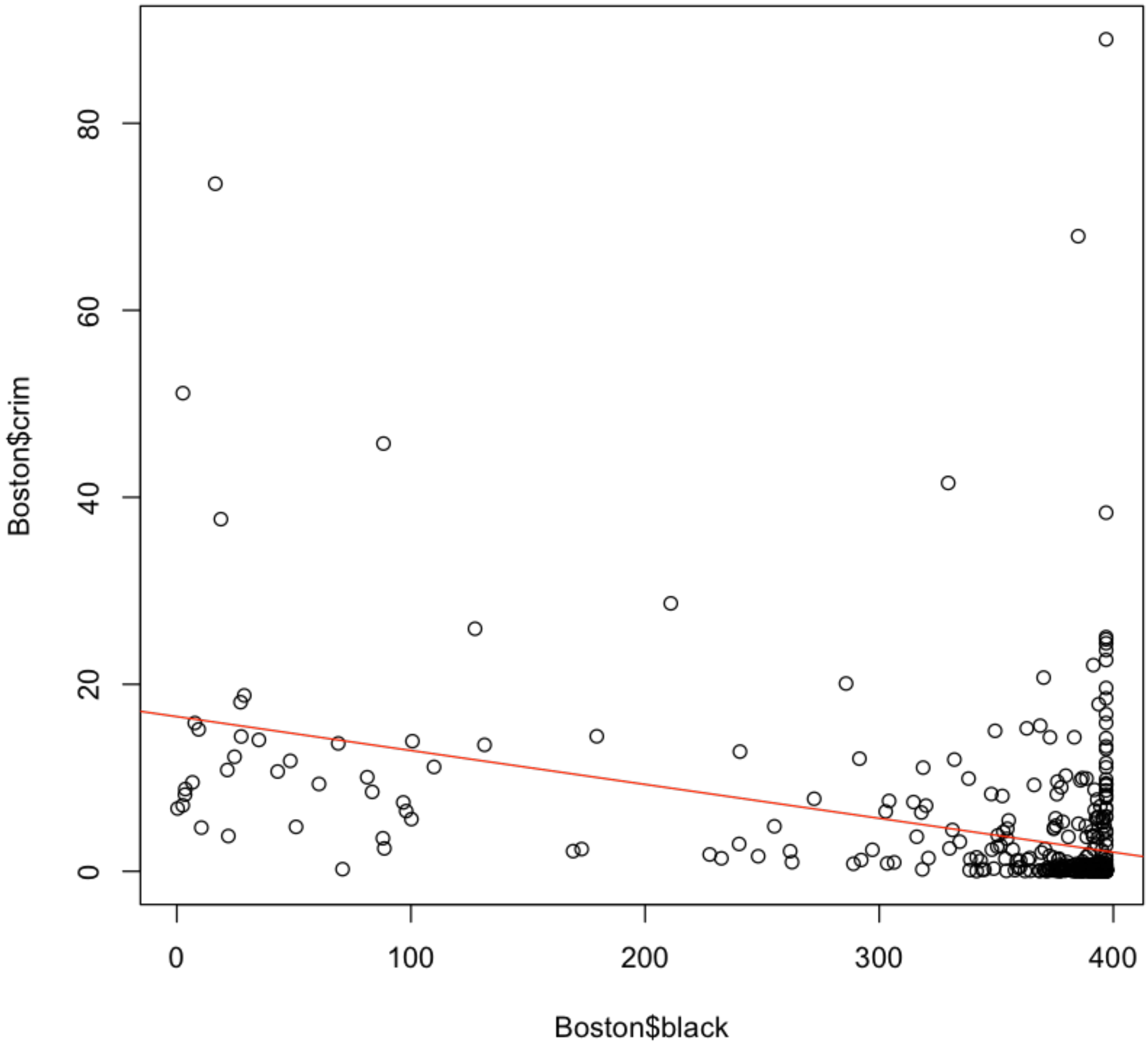
```
lm(formula = crim ~ ptratio)

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
```



In [68]:
```
lm.black = lm(crim~black)
summary(lm.black)
plot(Boston$black, Boston$crim)
abline(lm.black, col="red")
```

```
Call:
lm(formula = crim ~ black)

Residuals:
    Min      1Q  Median      3Q     Max 
-13.756  -2.299  -2.095  -1.296  86.822 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
```

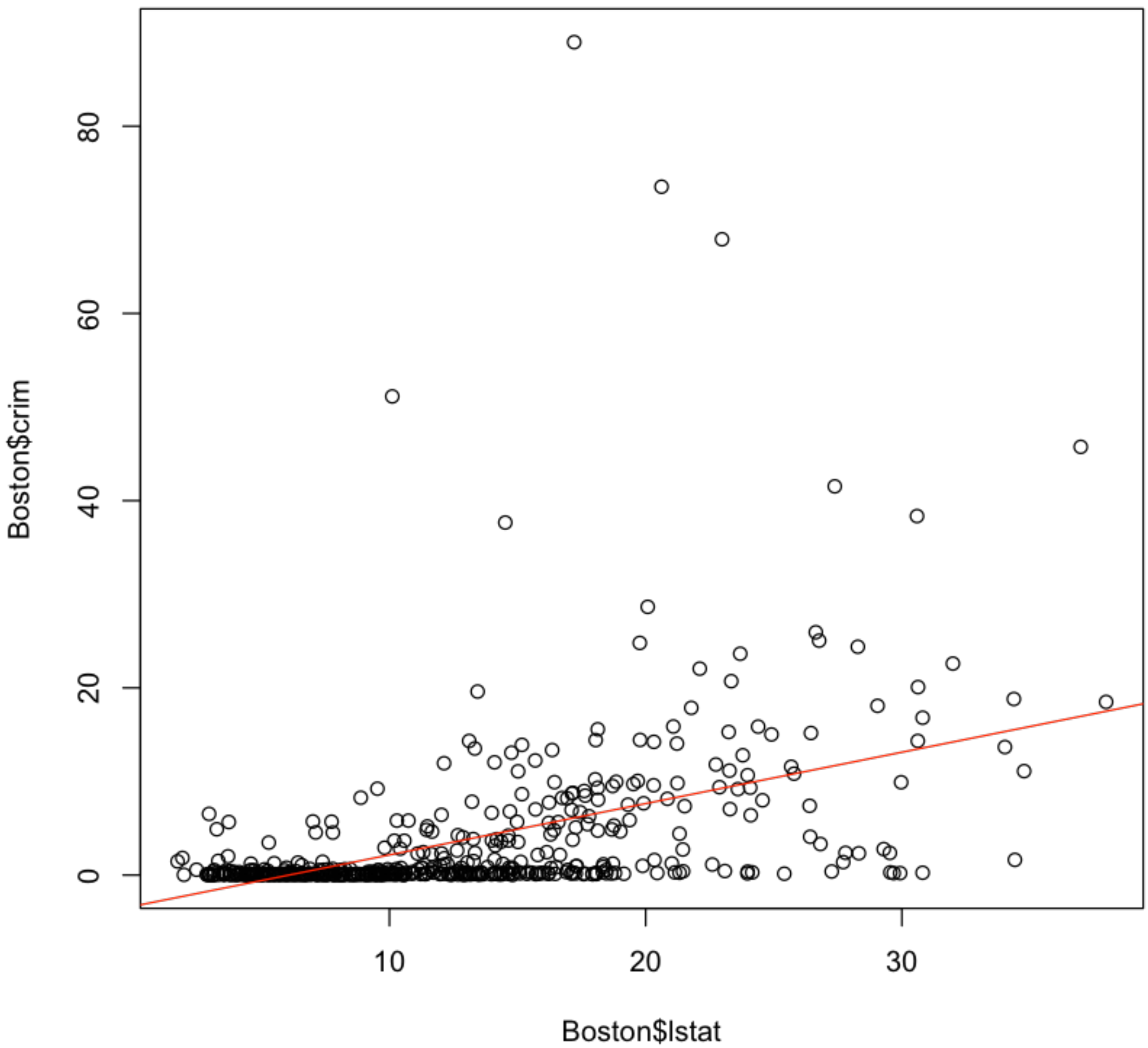
```
(Intercept) 16.553529   1.425903  11.609   <2e-16 ***
black       -0.036280   0.003873  -9.367   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

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
```



In [69]:
```
lm.lstat = lm(crim~lstat)
summary(lm.lstat)
plot(Boston$lstat, Boston$crim)
abline(lm.lstat, col="red")
```

```
Call:
```

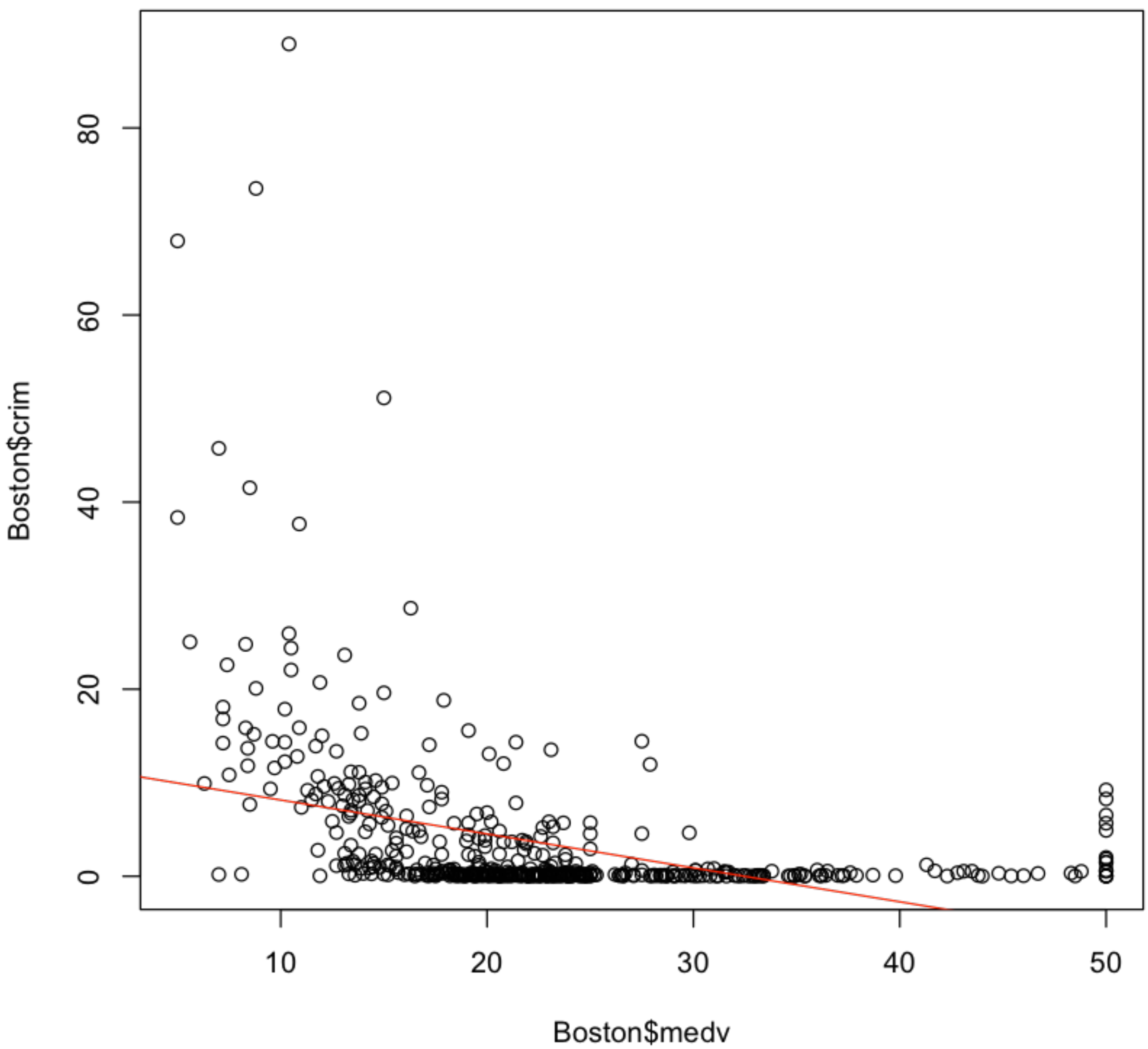
```
lm(formula = crim ~ lstat)

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
```



In [70]:
```
lm.medv = lm(crim~medv)
summary(lm.medv)
plot(Boston$medv, Boston$crim)
abline(lm.medv, col="red")

#Each predictor has statistically significant association
#with response except chas
```

```
Call:
lm(formula = crim ~ medv)

Residuals:
   Min     1Q Median     3Q    Max
-9.071 -4.022 -2.343  1.298 80.957
```

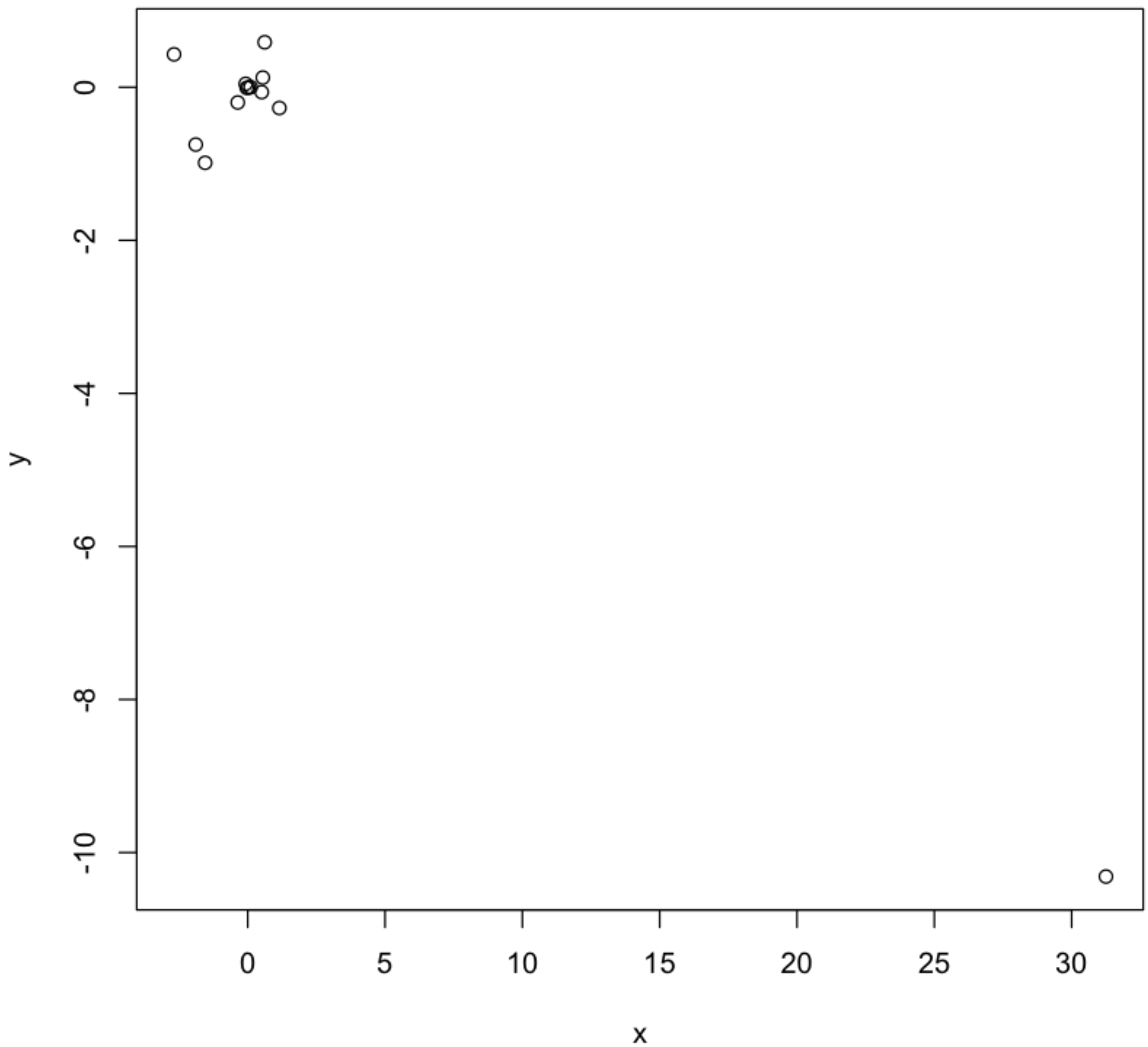
```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.79654    0.93419   12.63   <2e-16 ***
medv        -0.36316    0.03839   -9.46   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

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
```



In [18]:
```
#15(b)
lm.all = lm(crim~., data=Boston)
summary(lm.all)

# for zn, dis, rad, black, medv, we can reject null-hyp
```

```
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
```

In [94]:
```
#15(c)

x = c(coefficients(lm.zn)[2],
      coefficients(lm.indus)[2],
      coefficients(lm.chas)[2],
      coefficients(lm.nox)[2],
      coefficients(lm.rm)[2],
      coefficients(lm.age)[2],
      coefficients(lm.dis)[2],
      coefficients(lm.rad)[2],
      coefficients(lm.tax)[2],
      coefficients(lm.ptratio)[2],
      coefficients(lm.black)[2],
      coefficients(lm.lstat)[2],
      coefficients(lm.medv)[2])
y = coefficients(lm.all)[2:14]
plot(x, y)
#results vary between single variable and multi-variable linear regression va
#value in multi-variable regression than single variable regression, due to s
```



15 (d)

In [76]:

```
lm.zn_nl = lm(crim~zn + I(zn^2) + I(zn^3), data=Boston)
summary(lm.zn_nl)
```

```
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
```

In [77]:

```
lm.indus_nl = lm(crim~indus + I(indus^2) + I(indus^3), data=Boston)
summary(lm.indus_nl)
```

```
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
```

In [78]:

```
lm.chas_nl = lm(crim~chas + I(chas^2) + I(chas^3), data=Boston)
summary(lm.chas_nl)
```

```
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
```

In [79]:

```
lm.nox_nl = lm(crim~nox + I(nox^2) + I(nox^3), data=Boston)
summary(lm.nox_nl)
```

```
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
```

In [80]:

```
lm.rm_nl = lm(crim~rm + I(rm^2) + I(rm^3), data=Boston)
summary(lm.rm_nl)
```

```
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
```

In [81]:

```
lm.age_nl = lm(crim~age + I(age^2) + I(age^3), data=Boston)
summary(lm.age_nl)
```

```
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
```

In [82]:

```
lm.dis_nl = lm(crim~dis + I(dis^2) + I(dis^3), data=Boston)
summary(lm.dis_nl)
```

```
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
```

In [83]:

```
lm.rad_nl = lm(crim~rad + I(rad^2) + I(rad^3), data=Boston)
summary(lm.rad_nl)
```

```
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
```

In [84]:

```
lm.tax_nl = lm(crim~tax + I(tax^2) + I(tax^3), data=Boston)
summary(lm.tax_nl)
```

```
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
```

In [85]:
```
lm.ptratio_nl = lm(crim~ptratio + I(ptratio^2) + I(ptratio^3), data=Boston)
summary(lm.ptratio_nl)
```

```
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
```

In [75]:
```
lm.lstat_nl = lm(crim~lstat + I(lstat^2) + I(lstat^3), data=Boston)
summary(lm.lstat_nl)
```

```
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
```

In [86]:
```
lm.black_nl = lm(crim~black + I(black^2) + I(black^3), data=Boston)
summary(lm.black_nl)
```

```
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
```

In [ ]:
```
lm.lstat_nl = lm(crim~lstat + I(lstat^2) + I(lstat^3), data=Boston)
summary(lm.lstat_nl)
```

In [87]:
```
lm.medv_nl = lm(crim~medv + I(medv^2) + I(medv^3), data=Boston)
summary(lm.medv_nl)
```

```
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
```

In [ ]:

```
# 15(d) observations:
# medv, ptratio, dis, age, nox, indus have non-linear association with the re
```