sta674ex1f21

Melissa Pittard

3/15/2021

Problem 1

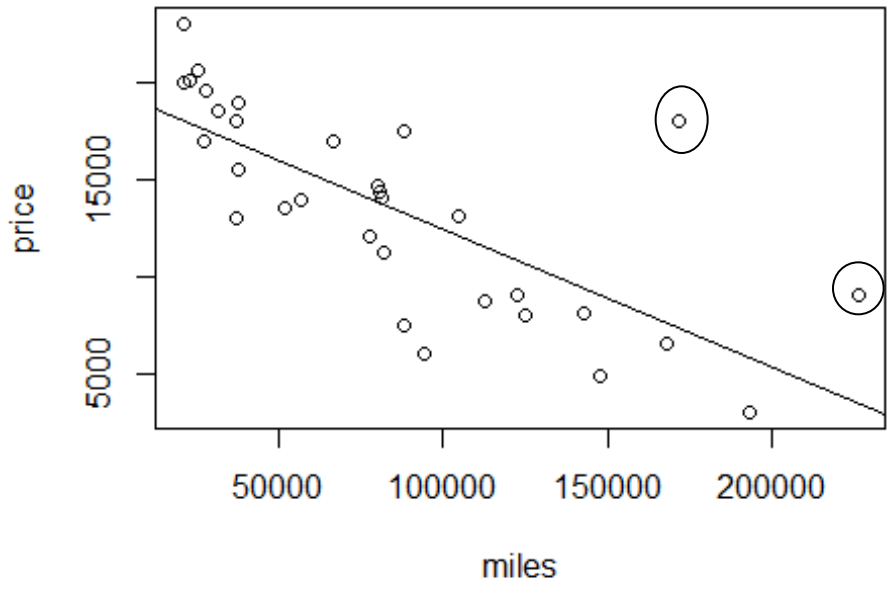
```
setwd("C:/Users/Melissa/Dropbox (Personal)/STA 674/exam 1")
prius = read.csv("priuspricesO.csv", header = T)
print(prius)
```

```
##    year  miles price age
## 1  2010 226331  8990  11
## 2  2017  66462 16900   4
## 3  2018  25152 20588   3
## 4  2015  81792 11198   6
## 5  2017  27825 19516   4
## 6  2010 122354  8995  11
## 7  2016 171953 17995   5
## 8  2017  80756 14385   4
## 9  2002 193000  2995  19
## 10 2018  26950 16957   3
## 11 2014  36705 12990   7
## 12 2006  94397  5995  15
## 13 2009 147598  4888  12
## 14 2017  80209 14633   4
## 15 2008 125065  7998  13
## 16 2015  56409 13950   6
## 17 2017  37450 18955   4
## 18 2017  87764 17464   4
## 19 2015 112460  8699   6
## 20 2018  21053 22995   3
## 21 2017 104397 13064   4
## 22 2016  36882 17981   5
## 23 2014 142903  8125   7
## 24 2017  31484 18480   4
## 25 2011 168139  6500  10
## 26 2019  22680 20095   2
## 27 2016  81016 13990   5
## 28 2019  20681 19985   2
## 29 2015  51825 13500   6
## 30 2007  88171  7499  14
## 31 2015  77809 11995   6
## 32 2017  37815 15500   4
```

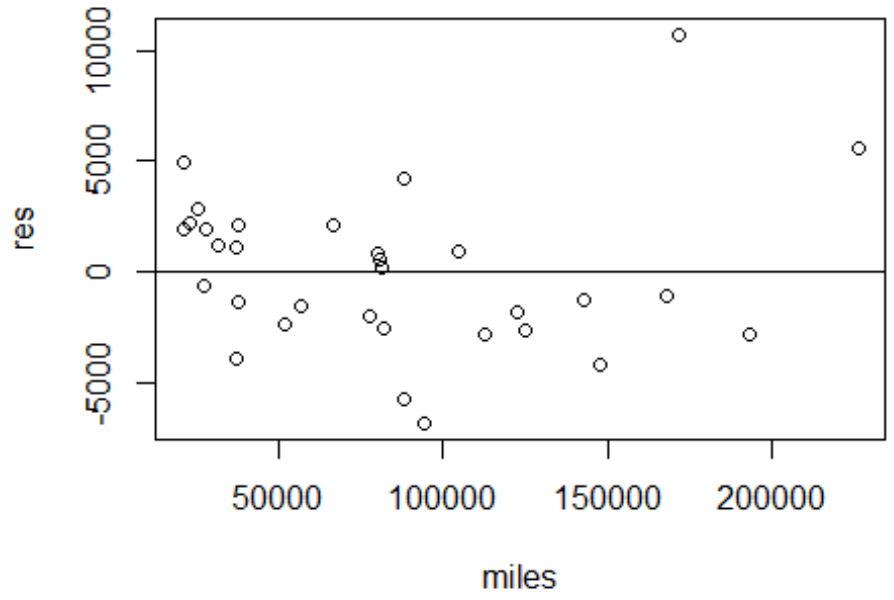
```
prius.lm = lm(price ~ miles, data=prius)
summary(prius.lm)
```

```
##
## Call:
## lm(formula = price ~ miles, data = prius)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6816.5 -2380.9  -211.1  2002.7 10696.9
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.952e+04  1.163e+03  16.791  < 2e-16 ***
## miles       -7.109e-02  1.163e-02  -6.111 1.02e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3572 on 30 degrees of freedom
## Multiple R-squared:  0.5545, Adjusted R-squared:  0.5397
## F-statistic: 37.35 on 1 and 30 DF,  p-value: 1.022e-06
```

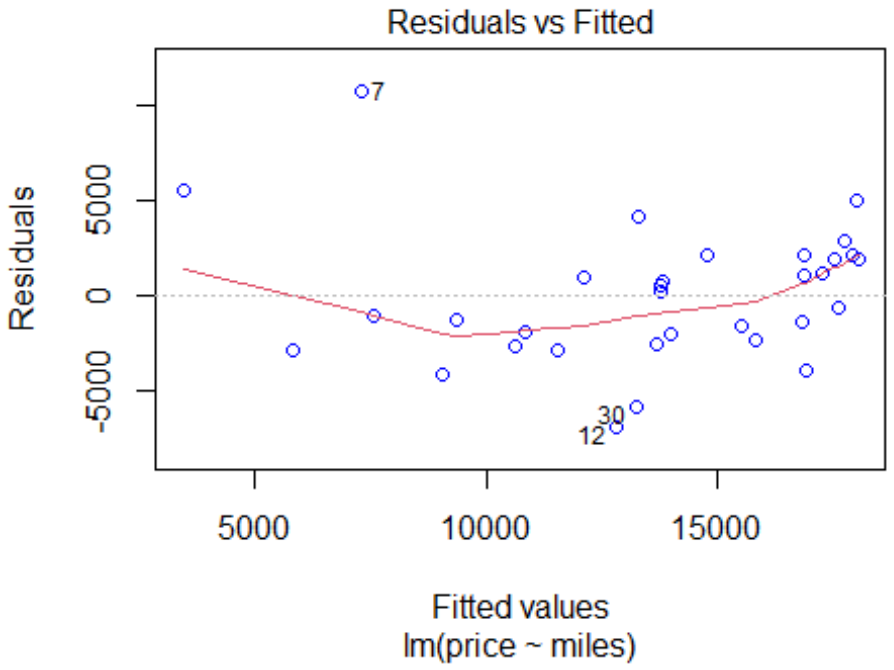
```
plot(price ~ miles, data=prius)
abline(prius.lm)
```



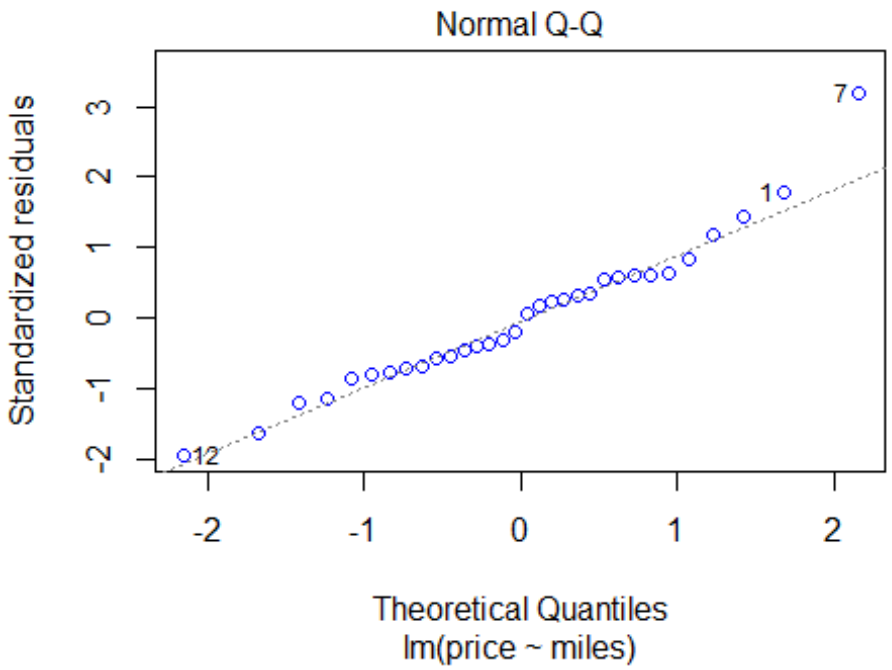

```
res = residuals(prius.lm)
plot(res ~ miles, data=prius)
abline(h=0)
```



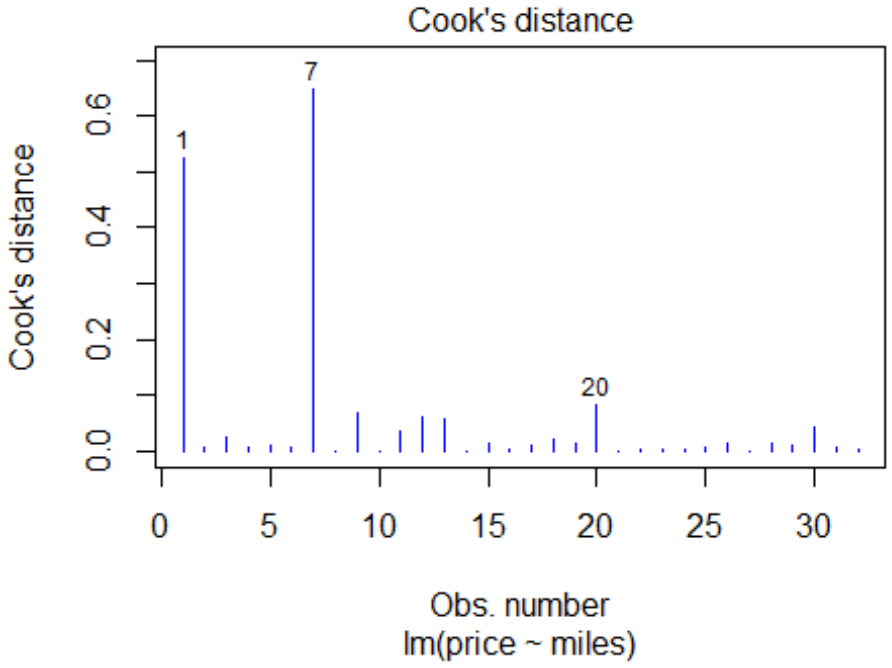

```
plot(prius.lm, which=1, col=c("blue"))
```




```
plot(prius.lm, which=2, col=c("blue"))
```

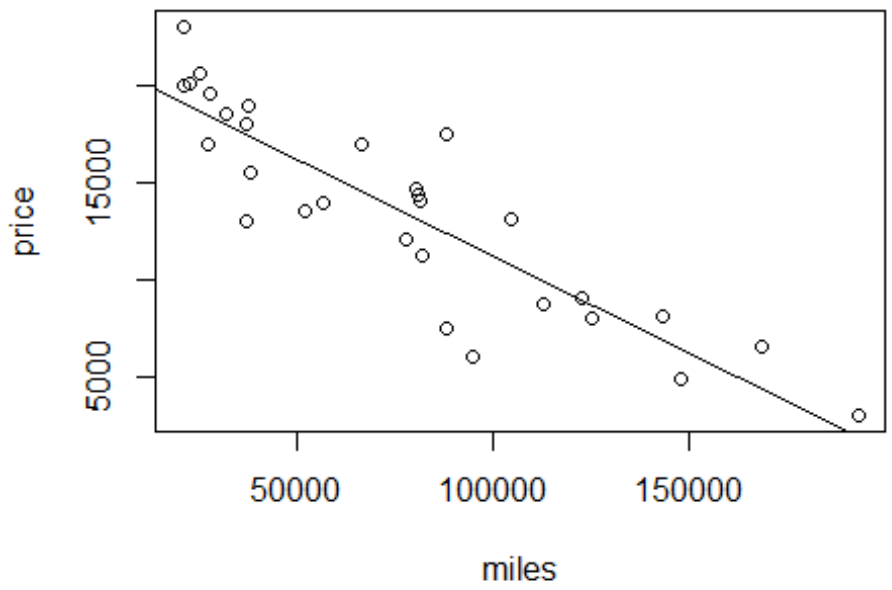



```r
plot(prius.lm, which=4, col=c("blue"))
```

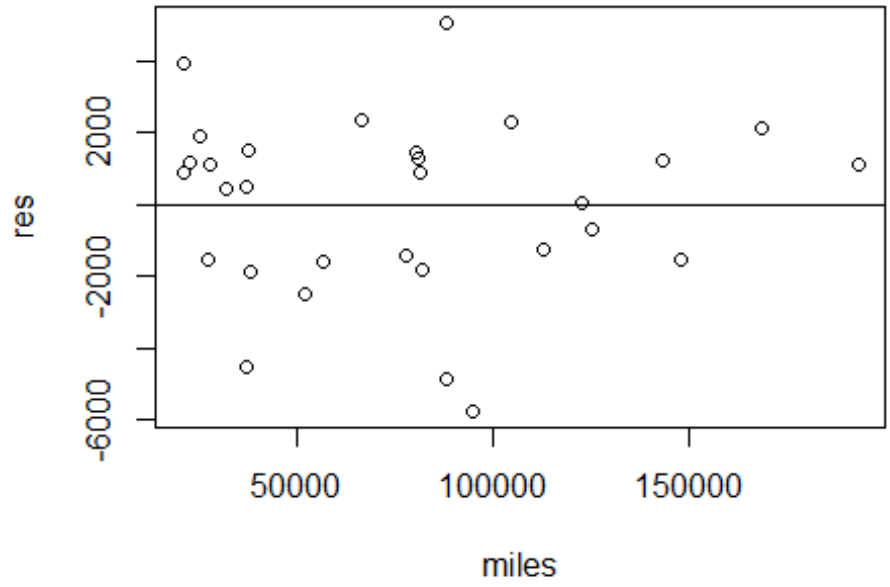


Problem 2

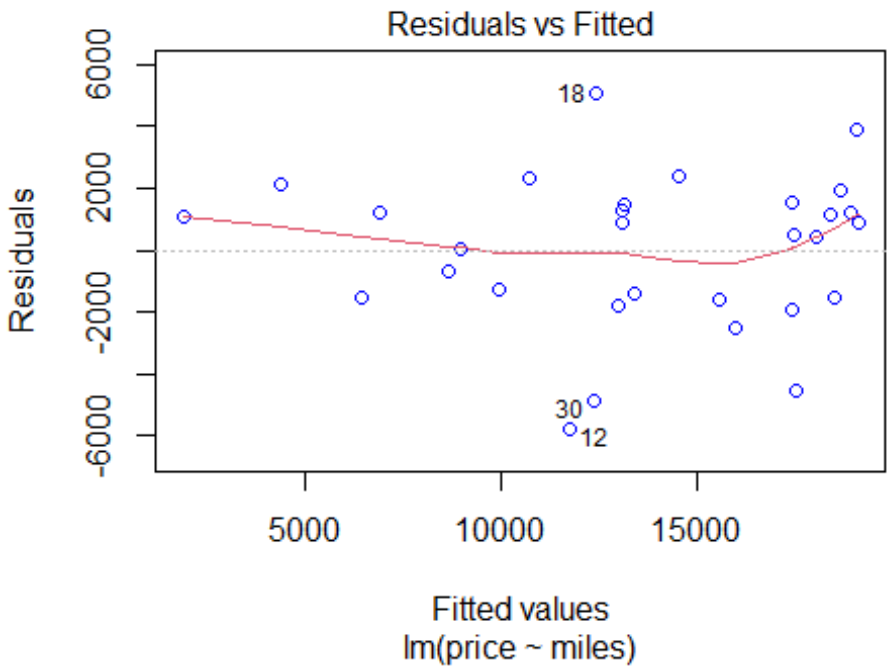
```
prius.new = prius[-c(1,7), ]
prius.new.lm = lm(price ~ miles, data=prius.new)
summary(prius.new.lm)

##
## Call:
## lm(formula = price ~ miles, data = prius.new)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5751.6 -1539.8   681.4  1420.7  5054.7
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.118e+04  8.788e+02   24.10  < 2e-16 ***
## miles       -9.990e-02  9.849e-03  -10.14 7.02e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2501 on 28 degrees of freedom
## Multiple R-squared:  0.7861, Adjusted R-squared:  0.7784
## F-statistic: 102.9 on 1 and 28 DF,  p-value: 7.019e-11

plot(price ~ miles, data = prius.new)
abline(prius.new.lm)
```



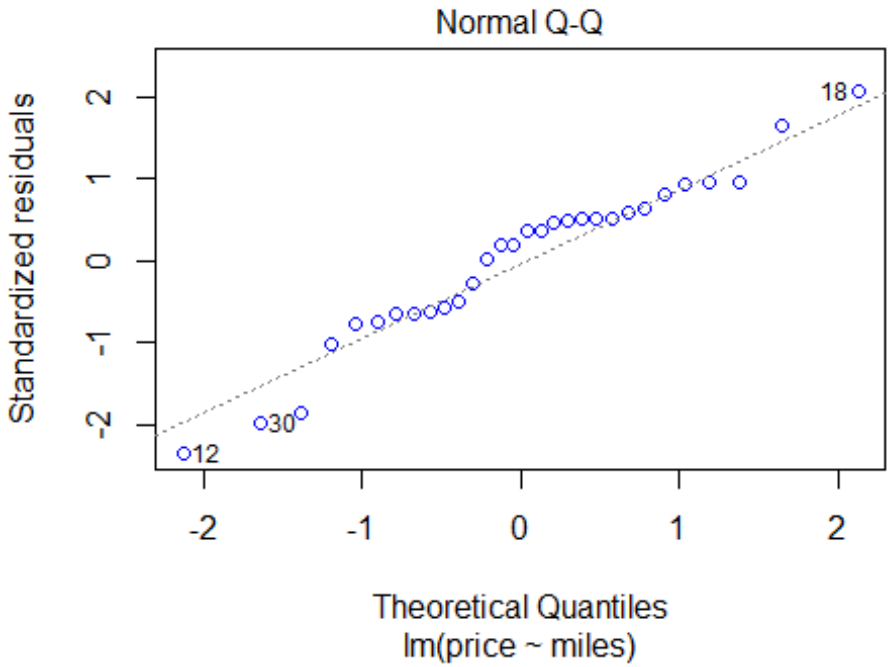
```
res = residuals(prius.new.lm)
plot(res ~ miles, data=prius.new)
abline(h=0)
```



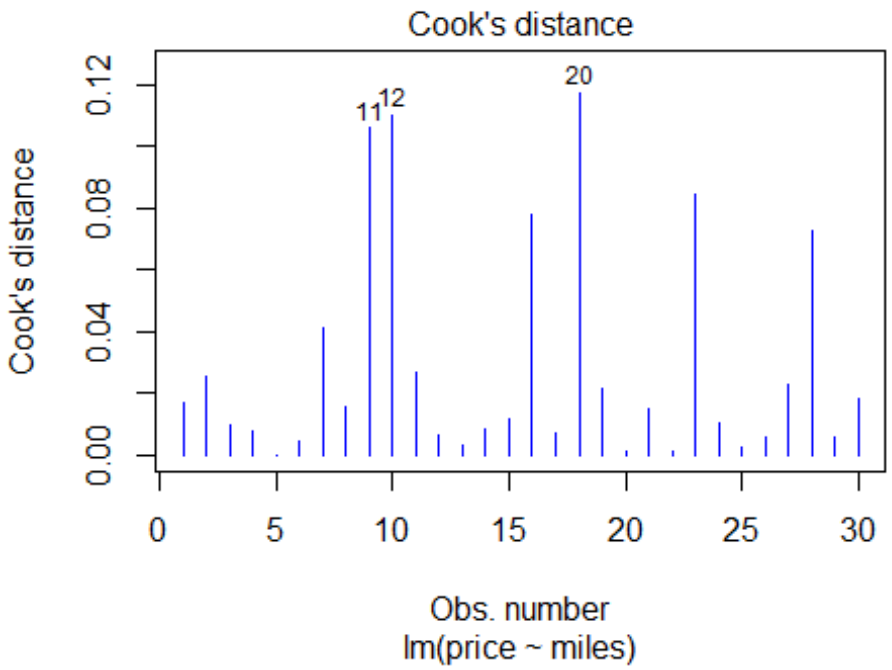

```
plot(prius.new.lm, which=1, col=c("blue"))
```




```
plot(prius.new.lm, which=2, col=c("blue"))
```

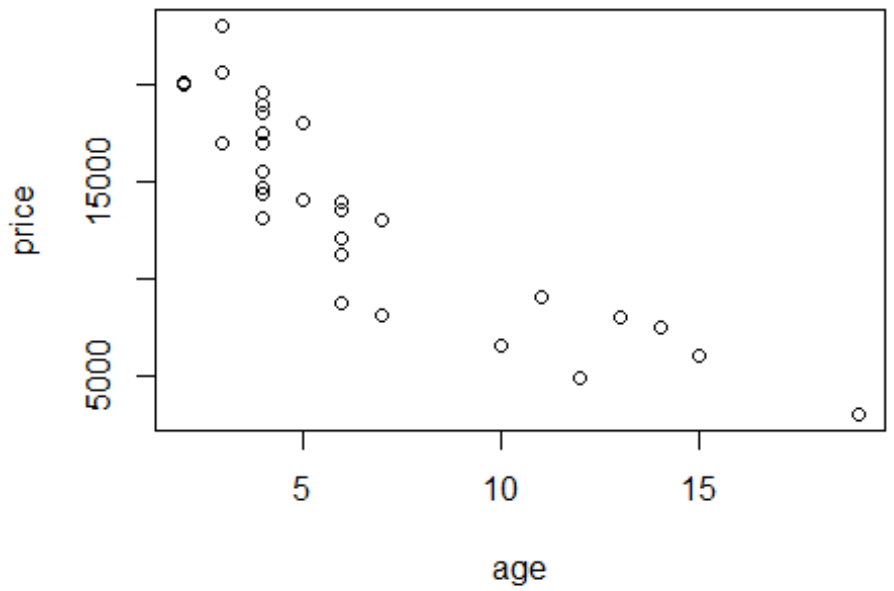



```
plot(prius.new.lm, which=4, col=c("blue"))
```

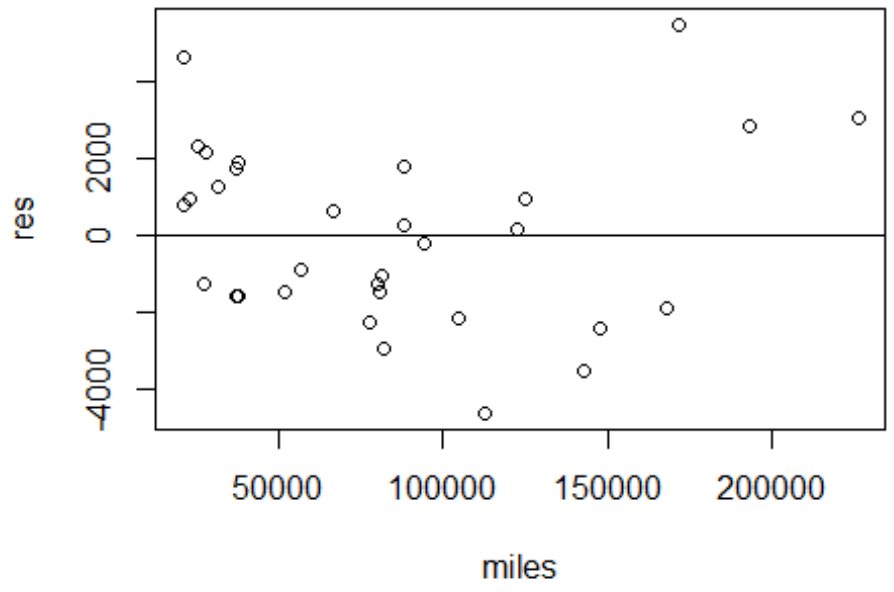


Problem 3

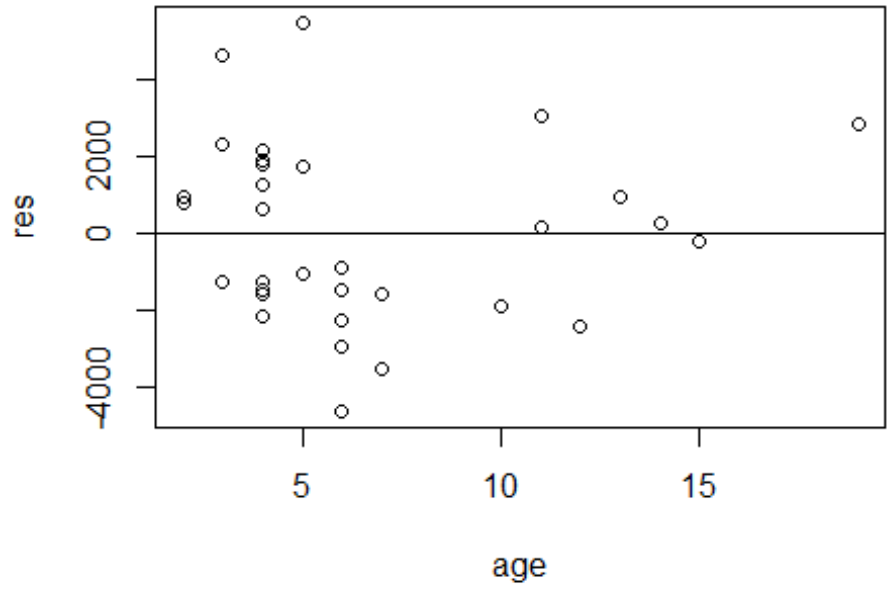
```
priusfull.lm = lm(price ~ age + miles, data=prius)
summary(priusfull.lm)

##
## Call:
## lm(formula = price ~ age + miles, data = prius)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4625.2 -1569.4   -34.9  1742.0  5465.0
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.148e+04  8.504e+02  25.264  < 2e-16 ***
## age         -8.438e+02  1.395e+02  -6.051 1.39e-06 ***
## miles       -2.753e-02  1.066e-02  -2.582   0.0151 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2415 on 29 degrees of freedom
## Multiple R-squared:  0.8031, Adjusted R-squared:  0.7895
## F-statistic: 59.15 on 2 and 29 DF,  p-value: 5.835e-11

plot(price ~ age, data = prius)
```



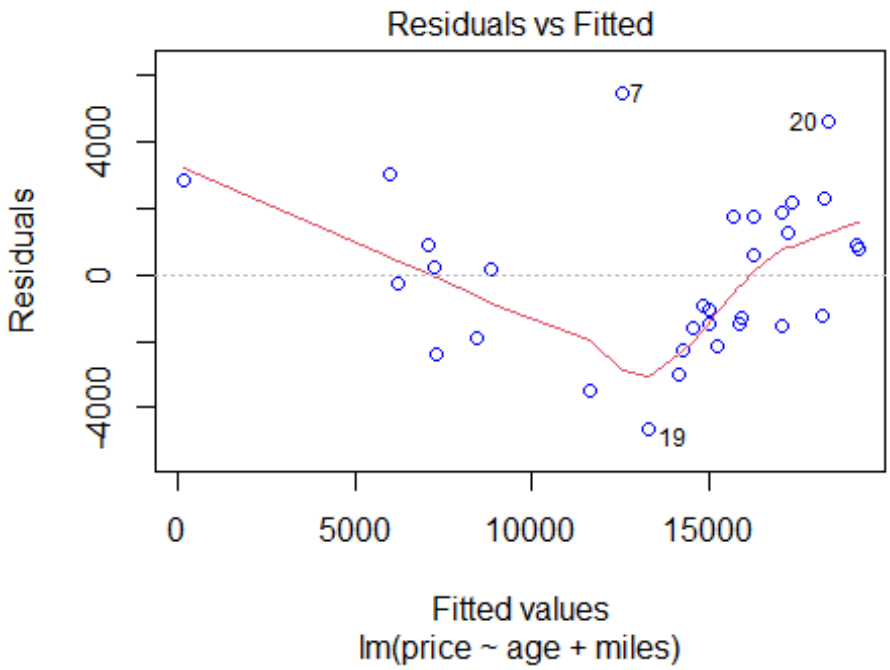

```
res = residuals(priusfull.lm)
plot(res ~ miles, data=prius)
abline(h=0)
```



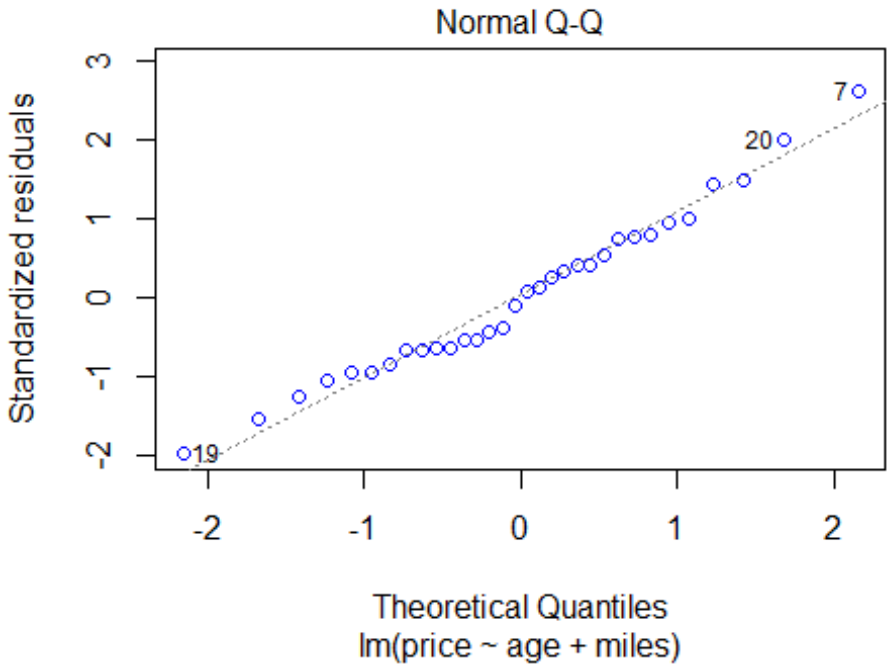
```r
plot(res ~ age, data = prius)
abline(h=0)
```



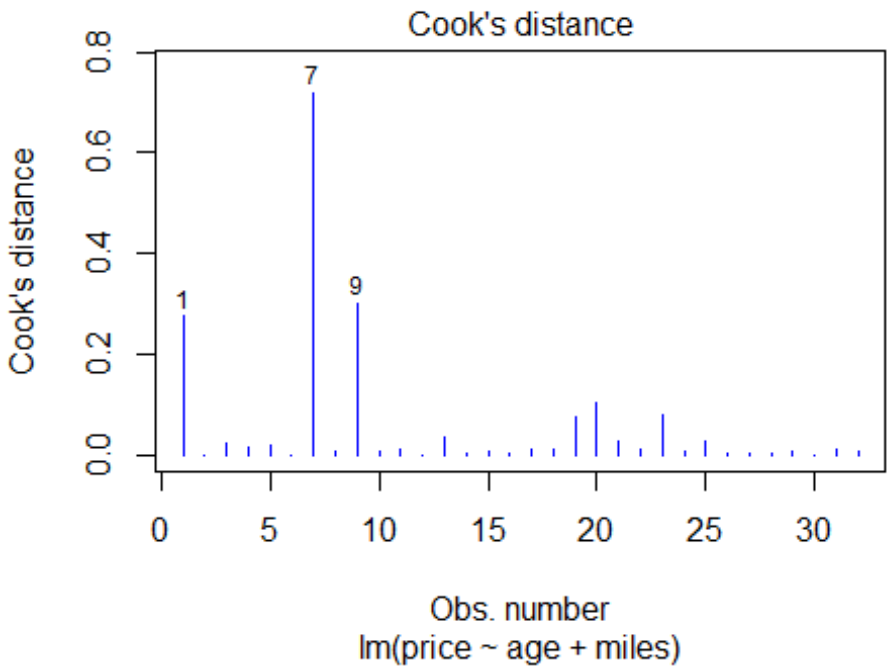

```r
plot(priusfull.lm, which=1, col=c("blue"))
```




```
plot(priusfull.lm, which=2, col=c("blue"))
```




```
plot(priusfull.lm, which=4, col=c("blue"))
```


Cook's distance
Cook's distance
0.0
0.2
0.4
0.6
0.8
1
7
9
0
5
10
15
20
25
30
Obs. number
lm(price ~ age + miles)