Roger(Siwei Zhang) hw5
20335254

8.a

i. Because p-value is very close to zero, so there is a relationship between the predictor and the response.
ii. They have a strong relationship because R-squared is 0.6 and RSE is 4.9 which is low.
iii. The relationship is negative. The expected change in mpg for 1 unit change in horsepower is -0.16.
iv. associated 24.46

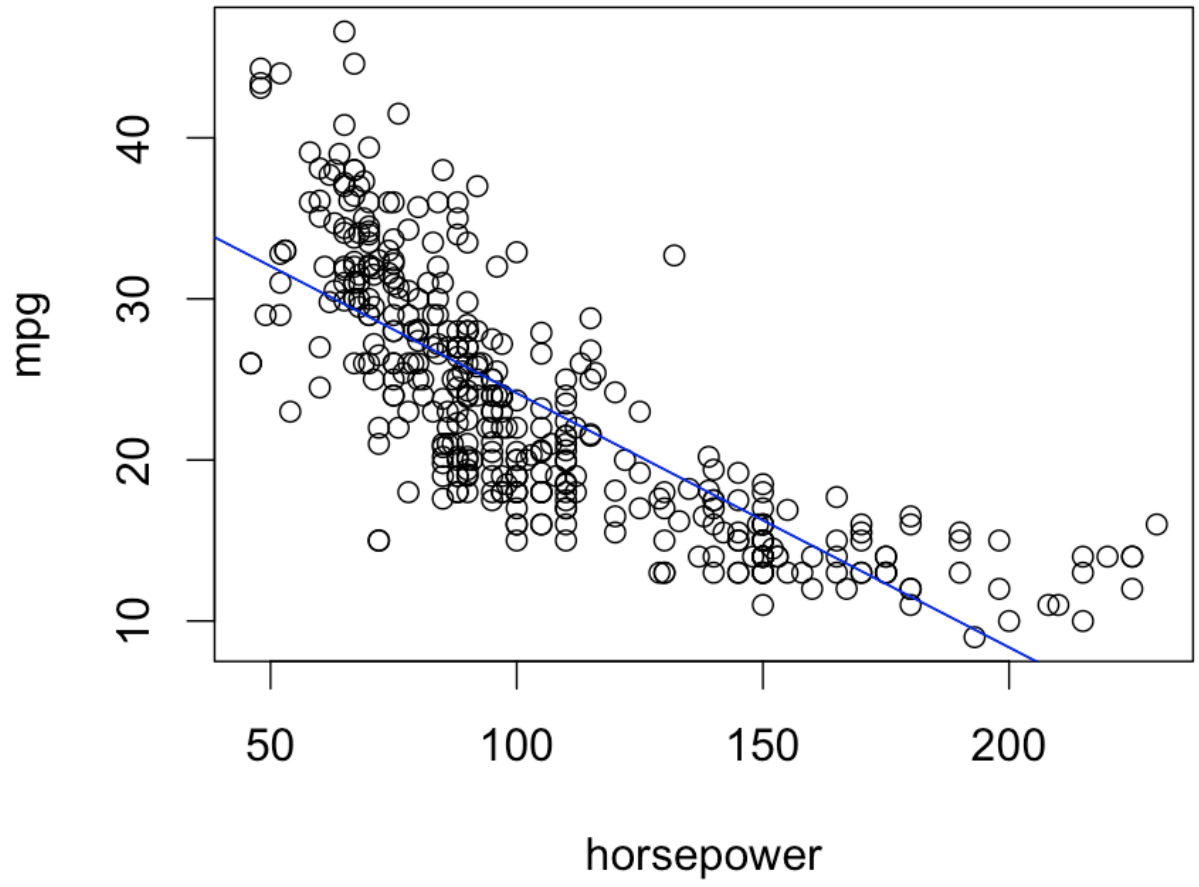
confidence

fit	lwr	upr
24.46708	23.97308	24.96108

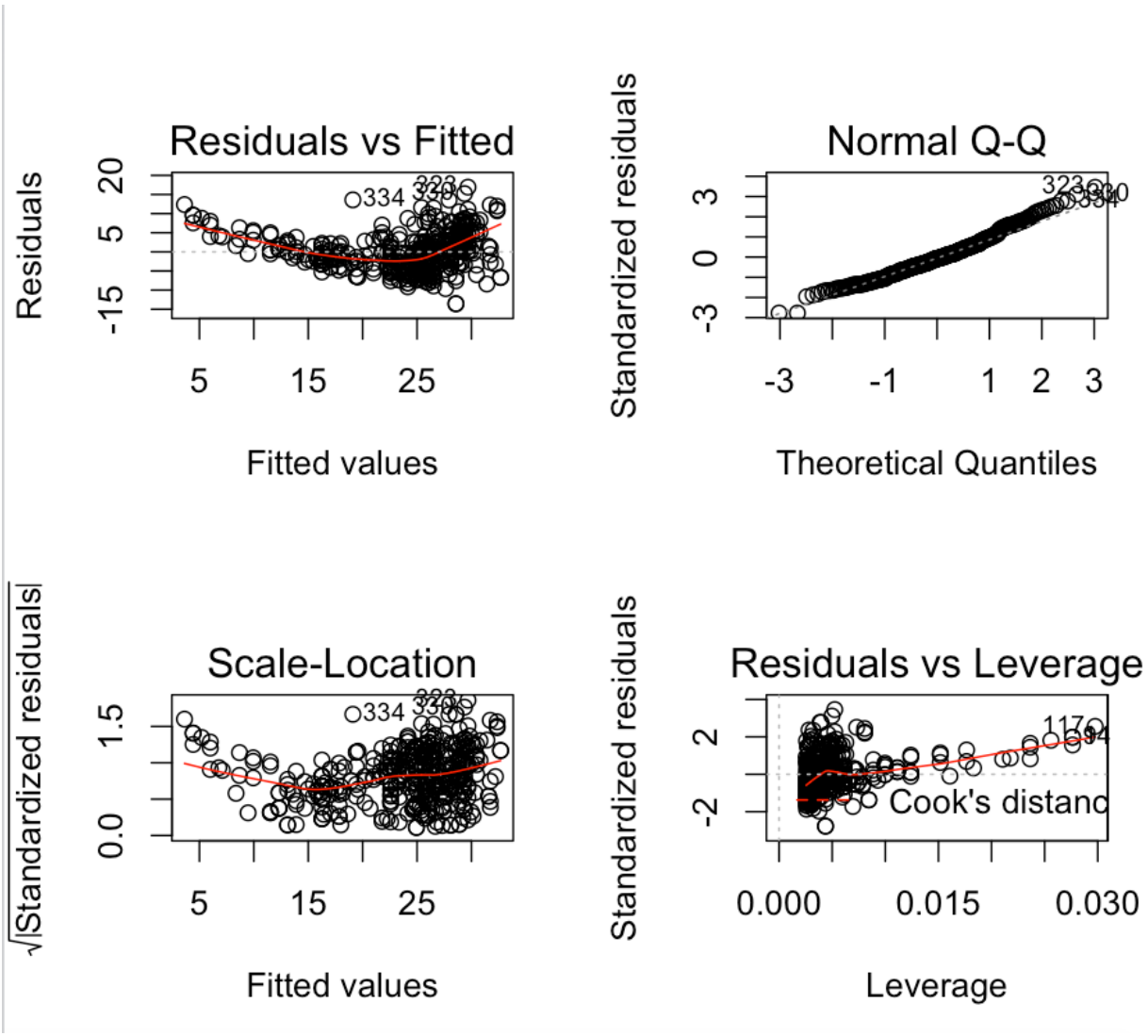
prediction

fit	lwr	upr
24.46708	14.8094	34.12476

b

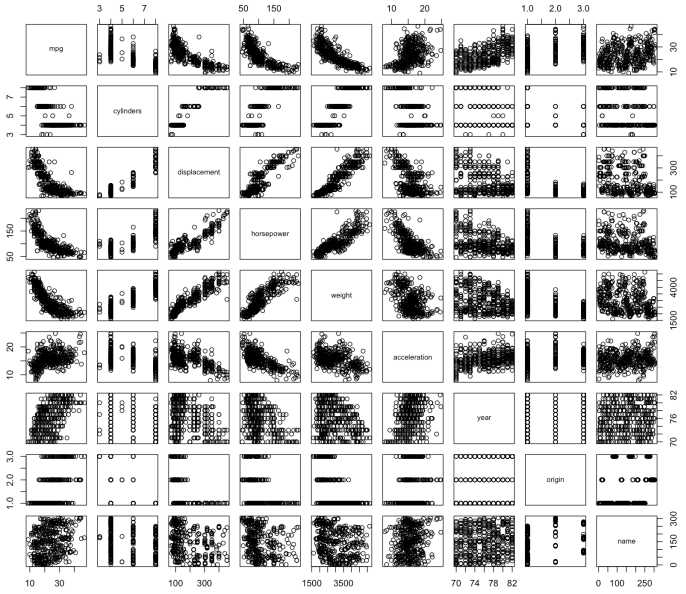


c



The model is not good fit the data and several points are high leverage.

9.a



b

```
                     mpg  cylinders displacement horsepower     weight
mpg           1.0000000 -0.7762599   -0.8044430         NA -0.8317389
cylinders    -0.7762599  1.0000000    0.9509199         NA  0.8970169
displacement -0.8044430  0.9509199    1.0000000         NA  0.9331044
horsepower           NA         NA           NA          1         NA
weight       -0.8317389  0.8970169    0.9331044         NA  1.0000000
acceleration  0.4222974 -0.5040606   -0.5441618         NA -0.4195023
year          0.5814695 -0.3467172   -0.3698041         NA -0.3079004
origin        0.5636979 -0.5649716   -0.6106643         NA -0.5812652
              acceleration       year     origin
mpg              0.4222974  0.5814695  0.5636979
cylinders       -0.5040606 -0.3467172 -0.5649716
displacement    -0.5441618 -0.3698041 -0.6106643
horsepower              NA         NA         NA
weight          -0.4195023 -0.3079004 -0.5812652
acceleration     1.0000000  0.2829009  0.2100836
year             0.2829009  1.0000000  0.1843141
origin           0.2100836  0.1843141  1.0000000
```

c

i. From the p-value, we can see that there is a relationship between the predictors and the response.
ii. displacement, weight, year and origin
iii. For every 1 unit year increase, the mpg increases by 0.75.

d

From the residuals and fitted plot we can see that the model is not very good fit to data. There are high leverage points.

e

i. All p-values changed.
ii. cylinders:displacement and cylinders:horsepower have no significant relationship.
iii. weight:year and weight cylinders still have significant relationship.

f

sqrt(displacement) and I(horsepower^2) have significant relationship, and log(cylinders) have no significant relationship.

10.a

```
fit<-lm(Sales~Price+Urban+US,data=Carseats)
summary(fit)
```

b

The price coefficient shows that 1 units increase in price, increases sales by -0.05.
And the US and Urban are qualitative.
The US coefficient shows that if US is yes, increases sales by 1200.
The Urban coefficient shows that if Urban is yes, decreases sales by 21.

c

Sales = intercept + price*xi1 + urban*xi2 + us*xi3

d
Price and US.

e
```
fit<-lm(Sales~Price+US,data=Carseats)
summary(fit)
```

f
Both models are not very good fit the data. And model e is slightly better with bigger F value and smaller RSE.

g

```
> confint(fit)
                  2.5 %      97.5 %
(Intercept) 11.79032020 14.27126531
Price       -0.06475984 -0.04419543
USYes        0.69151957  1.70776632
```

h
There are some outliers or high leverage observations in the model from (e).