STATS 202 homework 1

Siping Wang 006405652
July 2019

1 Chapter 2, Exercise 2

(a)

This scenario is a regression problem, and we are most interested in inference, because we are caring about the relationship between the CEO salary and the three factors. In this scenario, n = 500 and p = 3.

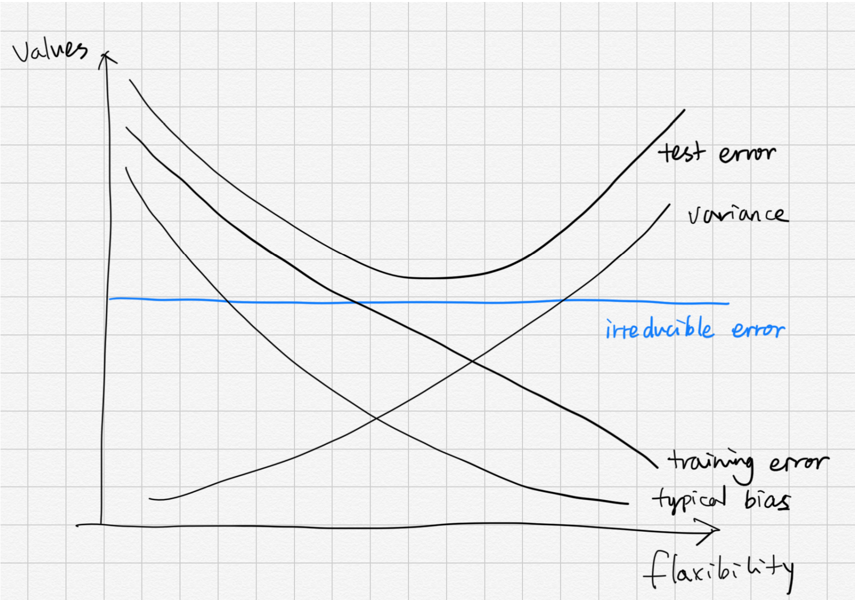
(b)

This scenario is a classification problem, and we are most interested in prediction, because we are want to know whether the product would be a success or failure in the future. In this scenario, n = 20 and p = 13.

(c)

This scenario is a regression problem, and we are most interested in prediction, because we want to predict the change in US dollar in relation to the weekly changes in the world stock markets. In this scenario, n = 52 and p = 3.

2 Chapter 2, Exercise 3



- Squared bias refers to the error that is introduced by approximating a real-life problem by a much simpler model, so it is high when the flexibility is low because there are fewer assumptions made about the shape of the fit, and it goes down as the model becoming more complex.

- Variance will increase with the flexibility increases because changing data points will have more effect on the parameter estimates.
- Training error will decrease with more flexibility because an overfit model will always produce lower error on the training data.
- Test error is U-shaped because when a f curve yields a small training error but a large test error we are actually overfitting the data, and thus the test error will increase.
- Irreducible error is always the same regardless of model fit.

3 Chapter 2, Exercise 7

(a)

- Obs.1: $\sqrt{0^2+3^2+0^2}=3$.
- Obs.2: $\sqrt{2^2+0^2+0^2}=2$.
- Obs.3: $\sqrt{0^2+1^2+3^2}=\sqrt{10}$.
- Obs.4: $\sqrt{0^2+1^2+2^2}=\sqrt{5}$.
- Obs.5: $\sqrt{(-1)^2+0^2+1^2}=\sqrt{2}$.
- Obs.6: $\sqrt{1^2+1^2+1^2}=\sqrt{3}$.

(b)

When K = 1, the closest neighbour is Obs.5, which is green. So our prediction is green.

(c)

When K = 3, the closest neighbours are Obs.5, Obs.6 and Obs.2, which are green, red, red. So our prediction is red.

(d)

We would expect K to be small to be able to capture more of the non-linear decision boundary.

4 Chapter 10, Exercise 1

(a)

$$\begin{aligned}\frac{1}{|C_k|}\sum_{i,i'\in C_k}\sum_{j=1}^{p}(x_{ij}-xi'j)^2 &= \sum_i\sum_j x_{ij}^2-2\sum_i\sum_j x_{ij}\bar{x}_{kj}+\sum_i\sum_j xij^2 \\ &= 2\sum_i\sum_j x_{ij}^2-2|C_k|\sum_j\bar{x}_{kj}^2\end{aligned}$$

and also

$$
\begin{aligned}
2\sum_{i\in C_k}\sum_{j=1}^{p}(x_{ij}-\bar{x}_{kj})^2 &= 2\sum_i\sum_j x_{ij}^2 - 4\sum_i\sum_j x_{ij}\bar{x}_{kj} + 2\sum_i\sum_j \bar{x}_{kj}^2 \\
&= 2\sum_i\sum_j x_{ij}^2 - 2|C_k|\sum_j \bar{x}_{kj}^2
\end{aligned}
$$

So we can prove (10.12).

(b)

The K-means clustering algorithm decreases the objective at each iteration because in this algorithm, observations are re-assigned to their closest cluster, thus minizing the Euclidean distance.

5 Chapter 10, Exercise 2

(a)

```
d = as.dist(matrix(c(0, 0.3, 0.4, 0.7,
                     0.3, 0, 0.5, 0.8,
                     0.4, 0.5, 0.0, 0.45,
                     0.7, 0.8, 0.45, 0.0), nrow = 4))
plot(hclust(d, method = "complete"))
```

Following the complete linkage algorithm,

- fusion1: (1, 2)
- fusion2: (3, 4)
- fusion3: ((1, 2), (3, 4))

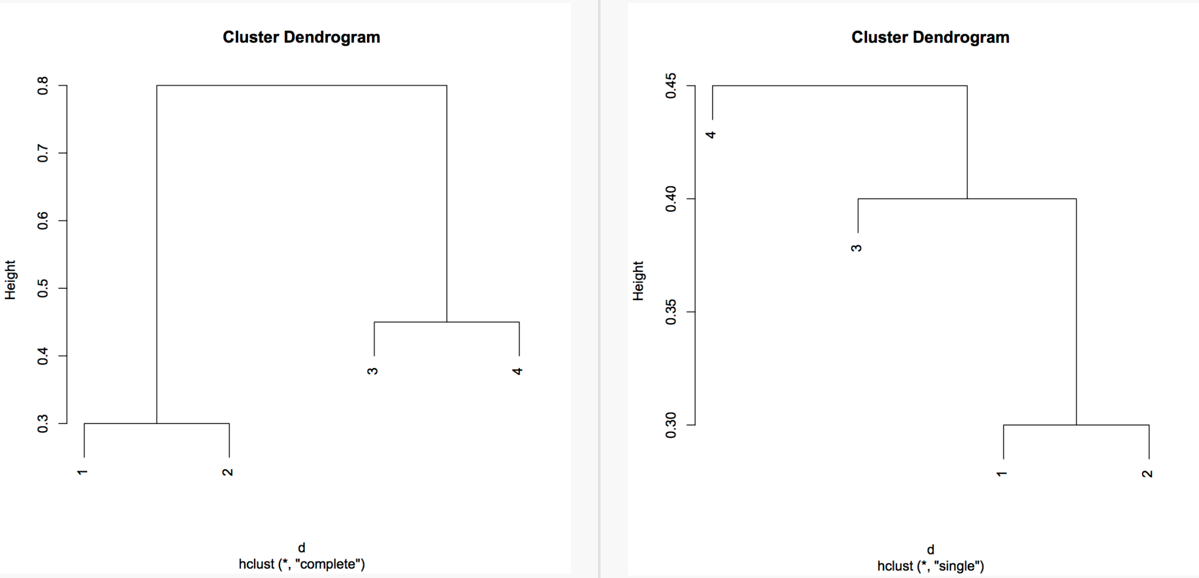
(b)

```
plot(hclust(d, method = "single"))
```

Following the single linkage algorithm,

- fusion1: (1, 2)
- fusion2: ((1, 2), 3)
- fusion3: (((1, 2), 3), 4)

Output of (a) and (b):



(a)	(b)

(c)

Clusters (1, 2) and (3, 4).

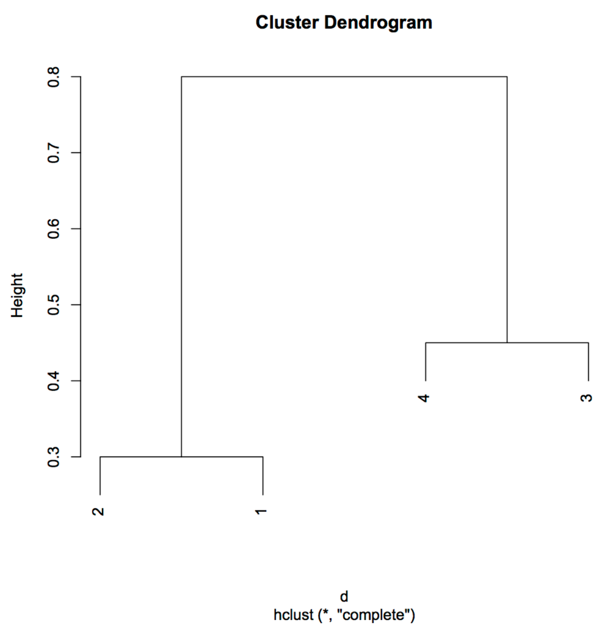
(d)

Clusters ((1, 2), 3) and (4).

(e)

```
plot(hclust(d, method = "complete"), labels = c(2,1,4,3))
```

Output:



6 Chapter 10, Exercise 4

(a)

Using the single linkage algorithm, when {1, 2, 3} and {4, 5} fuse, it means that the minimum of the distances beween any pairs of the two clusters is the smallest among the minimum of the distances beween any pairs of any other two clusters, i. e.,

$$height_1 = minD_{(1,2,3),(4,5)} = min\{minD_{others}\}$$

Using the complete linkage algorithm, when {1, 2, 3} and {4, 5} fuse, it means that the maximum of the distances beween any pairs of the two clusters is the smallest among the maximum of the distances beween any pairs of any other two clusters, i.e.,

$$height_2 = maxD_{(1,2,3),(4,5)} = min\{maxD_{others}\}$$

Clearly we can conclude that $height_1 \geq height_2$.

But if $height_1 = height_2$, it means that the distances between any pairs of the two clusters is the same, i.e.,

$$d = D_{1,4} = D_{1,5} = D_{2,4} = D_{2,5} = D_{3,4} = D_{3,5}$$

which means that Obs.1, Obs.2 and Obs.3 must be different dots at the intersection of two circles with Obs.4, Obs.5 as the center and d as the radius. However, two different circles have at most two intersections, so it is impossible.

Thus, $height_1 > height_2$ is true.

(b)

Similiar to (a), we can conclude that

$$height_1 = maxD_{(5),(6)} \geq height_2 = minD_{(5),(6)}.$$

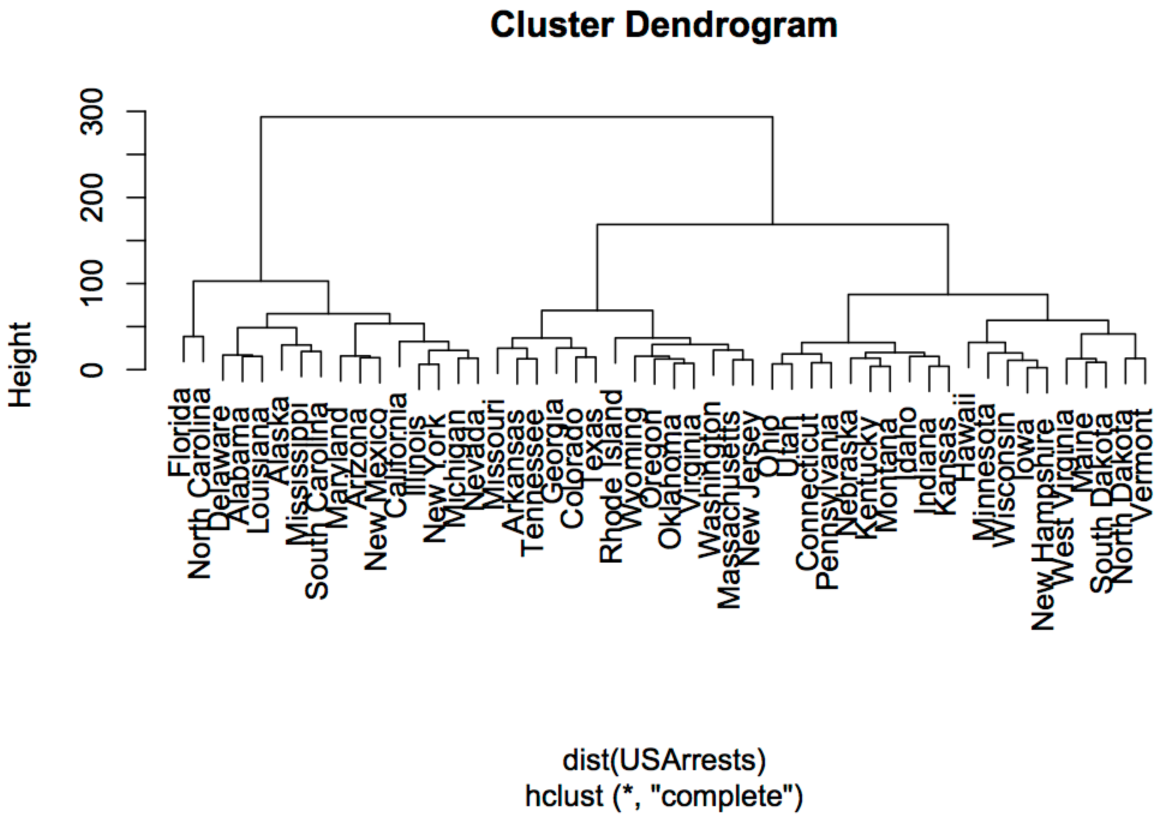
But this time, there is not enough information to tell if $height_1 > height_2$.

7 Chapter 10, Exercise 9

(a)

```
hclust.out=hclust(dist(USArrests),method='complete')
plot(hclust.out)
```

Output:



(b)

```
cutree(hclust.out, 3)
```

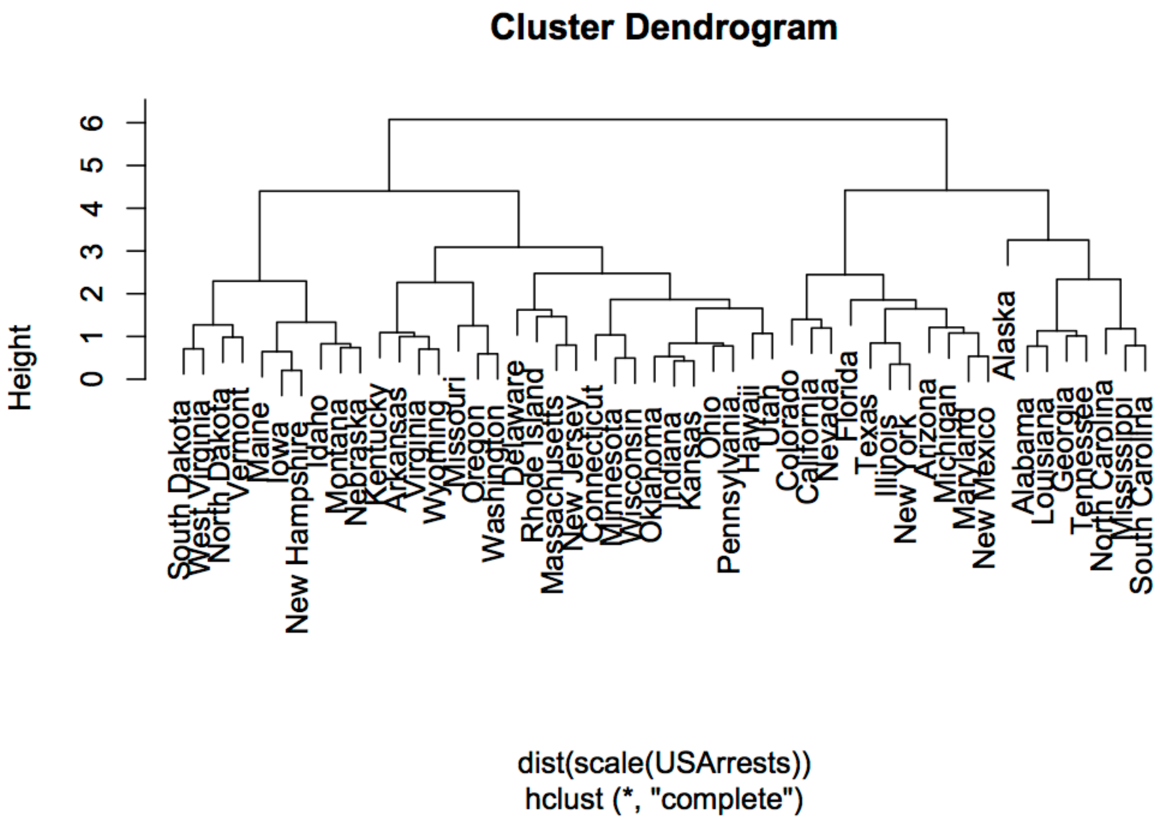
Output:

```
       Alabama         Alaska        Arizona       Arkansas     California
             1              1              1              2              1
      Colorado    Connecticut       Delaware        Florida        Georgia
             2              3              1              1              2
        Hawaii          Idaho       Illinois        Indiana           Iowa
             3              3              1              3              3
        Kansas       Kentucky      Louisiana          Maine       Maryland
             3              3              1              3              1
 Massachusetts       Michigan      Minnesota    Mississippi       Missouri
             2              1              3              1              2
       Montana       Nebraska         Nevada  New Hampshire     New Jersey
             3              3              1              3              2
    New Mexico       New York North Carolina   North Dakota           Ohio
             1              1              1              3              3
      Oklahoma         Oregon   Pennsylvania   Rhode Island South Carolina
             2              2              3              2              1
  South Dakota      Tennessee          Texas           Utah        Vermont
             3              2              2              3              3
      Virginia     Washington  West Virginia      Wisconsin        Wyoming
             2              2              3              3              2
```

(c)

```
hclust.out=hclust(dist(scale(USArrests)),method='complete')
plot(hclust.out)
```

Output:



(d)

```
cutree(hclust.out, 3)
```

Output:

```
       Alabama         Alaska        Arizona       Arkansas     California
             1              1              2              3              2
      Colorado    Connecticut       Delaware        Florida        Georgia
             2              3              3              2              1
        Hawaii          Idaho       Illinois        Indiana           Iowa
             3              3              2              3              3
        Kansas       Kentucky      Louisiana          Maine       Maryland
             3              3              1              3              2
 Massachusetts       Michigan      Minnesota    Mississippi       Missouri
             3              2              3              1              3
       Montana       Nebraska         Nevada  New Hampshire     New Jersey
             3              3              2              3              3
    New Mexico       New York North Carolina   North Dakota           Ohio
             2              2              1              3              3
      Oklahoma         Oregon   Pennsylvania   Rhode Island South Carolina
             3              3              3              3              1
  South Dakota      Tennessee          Texas           Utah        Vermont
             3              1              2              3              3
      Virginia     Washington  West Virginia      Wisconsin        Wyoming
             3              3              3              3              3
```

According to the outputs from (a) and (c), we can find that scaling can significantly reduces the range and spread of the height of the tree. Also by doing the cutting same as (b), the results are slightly different.

In my opinion, the variables should be scaled before the inter-obsercation dissimilarities are computed. Because Murder, Assault and Rape all have units of per 100,000 people while UrbanPop is the percentage of the state population that lives in urban areas. Therefore, by scaling the data, the units of UrbanPop would have an equal contribution to the hierarchical clustering algorithm as the other variables.

8 Chapter 3, Exercise 4

(a)

Since the information is not enough, it is difficult to tell which training RSS is exactly lower than the other. However, as the true relationship between X and Y is linear, we may expect the least squares line to be close to the true regression line, and consequently the RSS for the linear regression may be lower than that for the cubic regression.

(b)

We may also expect the RSS for the linear regression may be lower than that for the cubic regression, since the true relationship between X and Y is linear, the overfit from training of the cubic regression would have more error than the linear regression.

(c)

The cubic regression fit should produce a better RSS on the training set because it can adjust for the non-linearity.

(d)

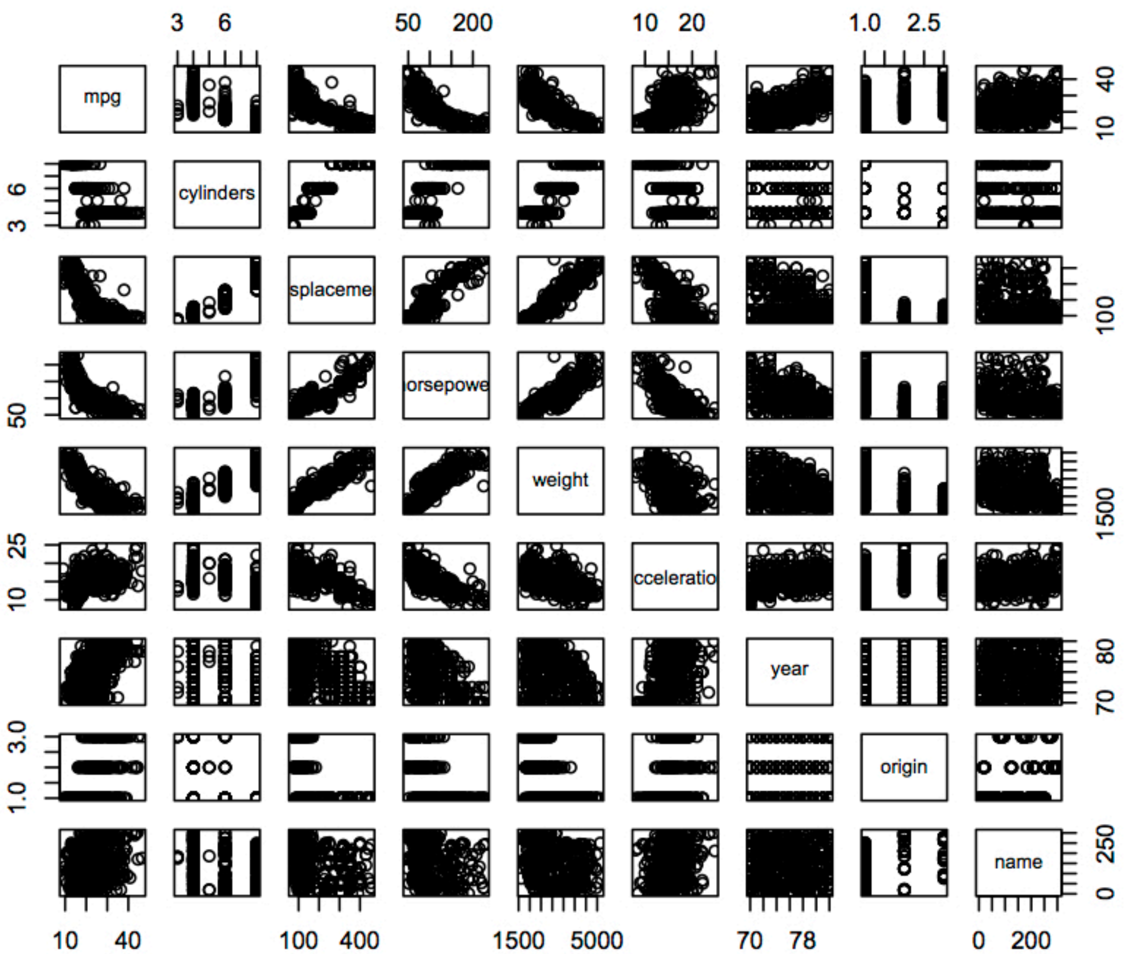
There is not enough information to tell which RSS is lower. If the true relationship between X and Y is closer to linear than cubic, the linear regression test RSS could be lower than the cubic regression test RSS. Otherwise, the cubic regression test RSS could be lower than that of the other.

9 Chapter 3, Exercise 9

(a)

```
require(ISLR)
data(Auto)
pairs(Auto)
```

Output:



(b)

```
cor(subset(Auto, select=-name))
```

Output:

```
                    mpg  cylinders displacement horsepower     weight
mpg           1.0000000 -0.7776175   -0.8051269 -0.7784268 -0.8322442
cylinders    -0.7776175  1.0000000    0.9508233  0.8429834  0.8975273
displacement -0.8051269  0.9508233    1.0000000  0.8972570  0.9329944
horsepower   -0.7784268  0.8429834    0.8972570  1.0000000  0.8645377
weight       -0.8322442  0.8975273    0.9329944  0.8645377  1.0000000
acceleration  0.4233285 -0.5046834   -0.5438005 -0.6891955 -0.4168392
year          0.5805410 -0.3456474   -0.3698552 -0.4163615 -0.3091199
origin        0.5652088 -0.5689316   -0.6145351 -0.4551715 -0.5850054
             acceleration       year     origin
mpg             0.4233285  0.5805410  0.5652088
cylinders      -0.5046834 -0.3456474 -0.5689316
displacement   -0.5438005 -0.3698552 -0.6145351
horsepower     -0.6891955 -0.4163615 -0.4551715
weight         -0.4168392 -0.3091199 -0.5850054
acceleration    1.0000000  0.2903161  0.2127458
year            0.2903161  1.0000000  0.1815277
origin          0.2127458  0.1815277  1.0000000
```

(c)

```
fit.lm <- lm(mpg~.-name, data=Auto)
summary(fit.lm)
```

Output:

```
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
```

Comment:

i) There is a relationship between the predictors and the response.

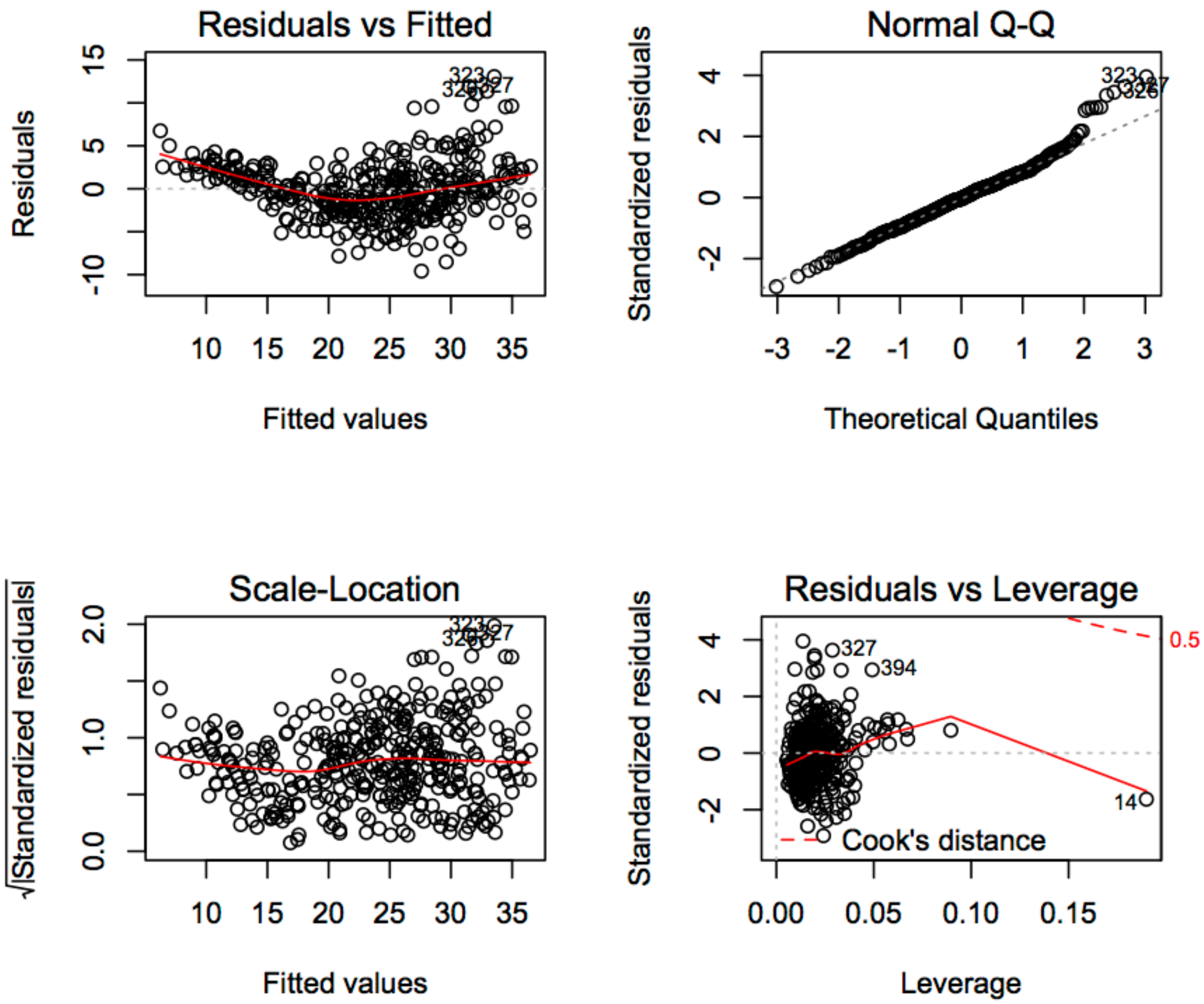
ii) The 3 predictors, weight, year and origin, appear to have a statistically significant relationship to the response. The predictor displacement also appears to have a statistically relationship to the response.

iii) The coeffidient for the year variable, 0.750773, suggests that the average effect of an increase of 1 year is an increase of 0.7507727 in mpg.

(d)

```
par(mfrow=c(2,2))
plot(fit.lm)
```

Output:



Comment:

The residual plots suggest that there are a few outliers (higher than 2 or lower than -2) and one high leverage point.

(e)

```
# try 3 interactions
fit.lm1 <- lm(mpg~displacement+weight+year*origin, data=Auto)
fit.lm2 <- lm(mpg~displacement+origin+year*weight, data=Auto)
fit.lm3 <- lm(mpg~cylinders*displacement+displacement*weight, data=Auto)
summary(fit.lm1)
summary(fit.lm2)
summary(fit.lm3)
```

Output:

```
# fit.lm1
Call:
lm(formula = mpg ~ displacement + weight + year * origin, data = Auto)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.7541 -1.8722 -0.0936  1.6900 12.4650 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)   7.927e+00  8.873e+00   0.893 0.372229    
displacement  1.551e-03  4.859e-03   0.319 0.749735    
weight       -6.394e-03  5.526e-04 -11.571  < 2e-16 ***
year          4.313e-01  1.130e-01   3.818 0.000157 ***
origin       -1.449e+01  4.707e+00  -3.079 0.002225 ** 
year:origin   2.023e-01  6.047e-02   3.345 0.000904 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.303 on 386 degrees of freedom
Multiple R-squared:  0.8232,	Adjusted R-squared:  0.8209 
F-statistic: 359.5 on 5 and 386 DF,  p-value: < 2.2e-16
```

```
# fit.lm2
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
```

```
# fit.lm3
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
```

Comment:

Both the first 2 interactions appear to be statistically significant. For the last interaction, we can find out that the interaction between cylinders and displacement is not statistically significant, while the interaction between weight and displacement is.

(f)

```
# try 3 interactions
fit.lm4 <- lm(mpg~displacement+log(weight)+year+origin, data=Auto)
fit.lm5 <- lm(mpg~displacement+weight+sqrt(year)+origin, data=Auto)
fit.lm6 <- lm(mpg~displacement+I(weight^2)+year+origin, data=Auto)
summary(fit.lm4)
summary(fit.lm5)
summary(fit.lm6)
```

Output:

```
# fit.lm4
Call:
lm(formula = mpg ~ displacement + log(weight) + year + origin,
    data = Auto)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.7136 -1.9214  0.0447  1.5790 12.9864 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)  131.274483  11.082986  11.845  < 2e-16 ***
displacement   0.007711   0.004052   1.903 0.057810 .  
log(weight)  -21.584745   1.451851 -14.867  < 2e-16 ***
year           0.804835   0.046532  17.296  < 2e-16 ***
origin         0.836143   0.250485   3.338 0.000925 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.113 on 387 degrees of freedom
Multiple R-squared:  0.8425,	Adjusted R-squared:  0.8409 
F-statistic: 517.7 on 4 and 387 DF,  p-value: < 2.2e-16
```

```
# fit.lm5
Call:
lm(formula = mpg ~ displacement + weight + sqrt(year) + origin,
    data = Auto)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.8339 -2.1130 -0.0335  1.7946 13.2206 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)  -7.669e+01  7.744e+00  -9.903  < 2e-16 ***
displacement  5.699e-03  4.782e-03   1.192    0.234    
weight       -6.595e-03  5.586e-04 -11.807  < 2e-16 ***
sqrt(year)    1.340e+01  8.703e-01  15.392  < 2e-16 ***
origin        1.226e+00  2.676e-01   4.583 6.19e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.354 on 387 degrees of freedom
Multiple R-squared:  0.8173,	Adjusted R-squared:  0.8154 
F-statistic: 432.7 on 4 and 387 DF,  p-value: < 2.2e-16
```

```
# fit.lm6
Call:
lm(formula = mpg ~ displacement + I(weight^2) + year + origin,
    data = Auto)

Residuals:
     Min       1Q   Median       3Q      Max
-10.0988  -2.2549  -0.1057   1.8704  13.4702

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.609e+01  4.349e+00  -5.999 4.56e-09 ***
displacement -9.114e-03  5.118e-03  -1.781   0.0757 .
I(weight^2)  -7.068e-07  9.075e-08  -7.789 6.28e-14 ***
year          7.336e-01  5.380e-02  13.635  < 2e-16 ***
origin        1.488e+00  2.900e-01   5.132 4.56e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.628 on 387 degrees of freedom
Multiple R-squared:  0.7861,  Adjusted R-squared:  0.7839
F-statistic: 355.7 on 4 and 387 DF,  p-value: < 2.2e-16
```

Comment:

All the 3 transformations of the variables are statistically significant.

10 Chapter 3, Exercise 14

(a)

```
set.seed(1)
x1 <- runif(100)
x2 <- 0.5 * x1 + rnorm(100)/10
y <- 2 + 2 * x1 + 0.3 * x2 + rnorm(100)
```

The form of the linear model is: $Y = 2 + 2X_1 + 0.3X_2 + \epsilon$.

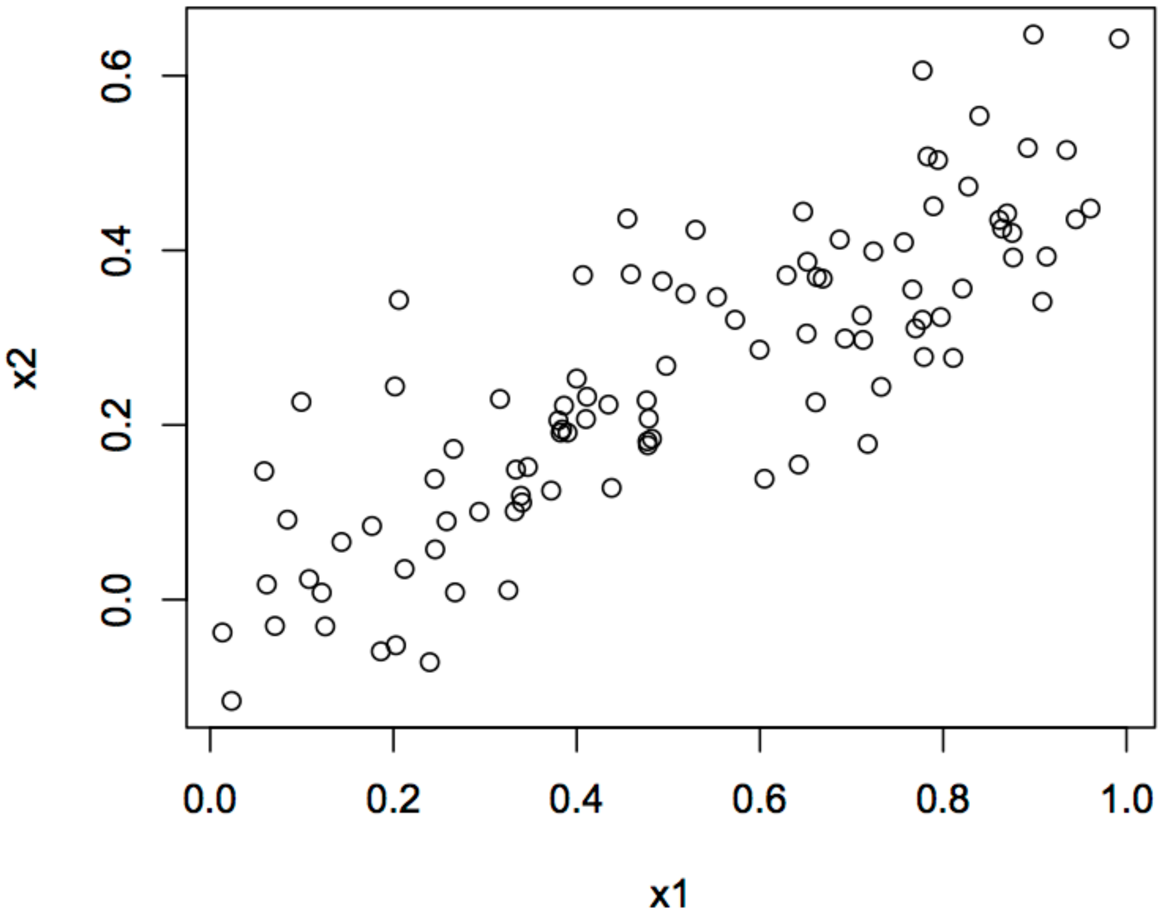
The regression coefficients are respectively 2, 2, 0.3.

(b)

```
cor(x1, x2)
plot(x1, x2)
```

Output:

```
[1] 0.8351212
```



(c)

```
fit <- lm(y ~ x1 + x2)
summary(fit)
```

Output:

```
Call:
lm(formula = y ~ x1 + x2)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.8311 -0.7273 -0.0537  0.6338  2.3359 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.1305     0.2319   9.188 7.61e-15 ***
x1            1.4396     0.7212   1.996   0.0487 *  
x2            1.0097     1.1337   0.891   0.3754    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.056 on 97 degrees of freedom
Multiple R-squared:  0.2088,  Adjusted R-squared:  0.1925 
F-statistic:  12.8 on 2 and 97 DF,  p-value: 1.164e-05
```

Results:

- $\hat{\beta}_0$ is 2.1305, $\hat{\beta}_1$ is 1.4396, $\hat{\beta}_2$ is 1.0097.
- As the p-value is less than 0.05 we may reject the null hypothesis $H_0 : \beta_1 = 0$.
- However, as the p-value is higher than 0.05, we cannot reject $H_0 : \beta_2 = 0$.

(d)

```
fit <- lm(y ~ x1)
summary(fit)
```

Output:

```
Call:
lm(formula = y ~ x1)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.89495 -0.66874 -0.07785  0.59221  2.45560 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.1124     0.2307   9.155 8.27e-15 ***
x1            1.9759     0.3963   4.986 2.66e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.055 on 98 degrees of freedom
Multiple R-squared:  0.2024,	Adjusted R-squared:  0.1942 
F-statistic: 24.86 on 1 and 98 DF,  p-value: 2.661e-06
```

Comment:

- As p-value is very low, we may reject the null hypothesis $H_0 : \beta_1 = 0$.

(e)

```
fit <- lm(y ~ x2)
summary(fit)
```

Output:

```
Call:
lm(formula = y ~ x2)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.62687 -0.75156 -0.03598  0.72383  2.44890 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.3899     0.1949   12.26  < 2e-16 ***
x2            2.8996     0.6330    4.58 1.37e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.072 on 98 degrees of freedom
Multiple R-squared:  0.1763,	Adjusted R-squared:  0.1679 
F-statistic: 20.98 on 1 and 98 DF,  p-value: 1.366e-05
```

Commant:

- As p-value is very low, we may reject the null hypothesis $H_0 : \beta_2 = 0$.

(f)

The results obtained in (c) - (e) don't contradict each other. The results indicate that without the presence of other parameters, both β_1 and β_2 are statistically significant. However, with the presence of β_1. β_2 is no longer significant.

(g)

```
x1 <- c(x1, 0.1)
x2 <- c(x2, 0.8)
y <- c(y, 6)
fit1 <- lm(y ~ x1 + x2)
fit2 <- lm(y ~ x1)
fit3 <- lm(y ~ x2)
summary(fit1)
plot(fit1)
summary(fit2)
plot(fit2)
summary(fit3)
plot(fit3)
```

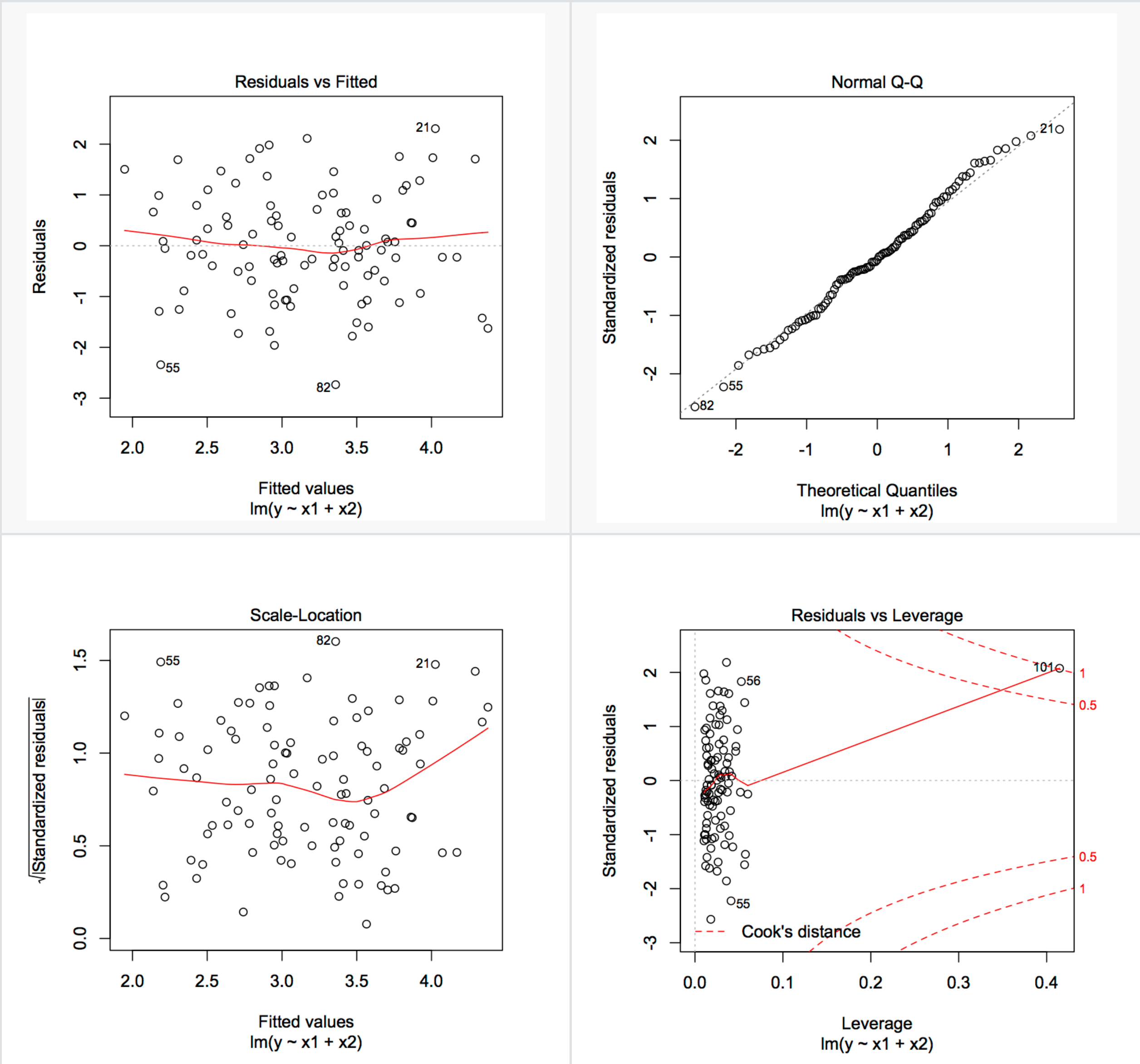
Output:

```
# fit1
Call:
lm(formula = y ~ x1 + x2)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.73348 -0.69318 -0.05263  0.66385  2.30619 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.2267     0.2314   9.624 7.91e-16 ***
x1            0.5394     0.5922   0.911  0.36458    
x2            2.5146     0.8977   2.801  0.00614 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.075 on 98 degrees of freedom
Multiple R-squared:  0.2188,  Adjusted R-squared:  0.2029 
F-statistic: 13.72 on 2 and 98 DF,  p-value: 5.564e-06
```

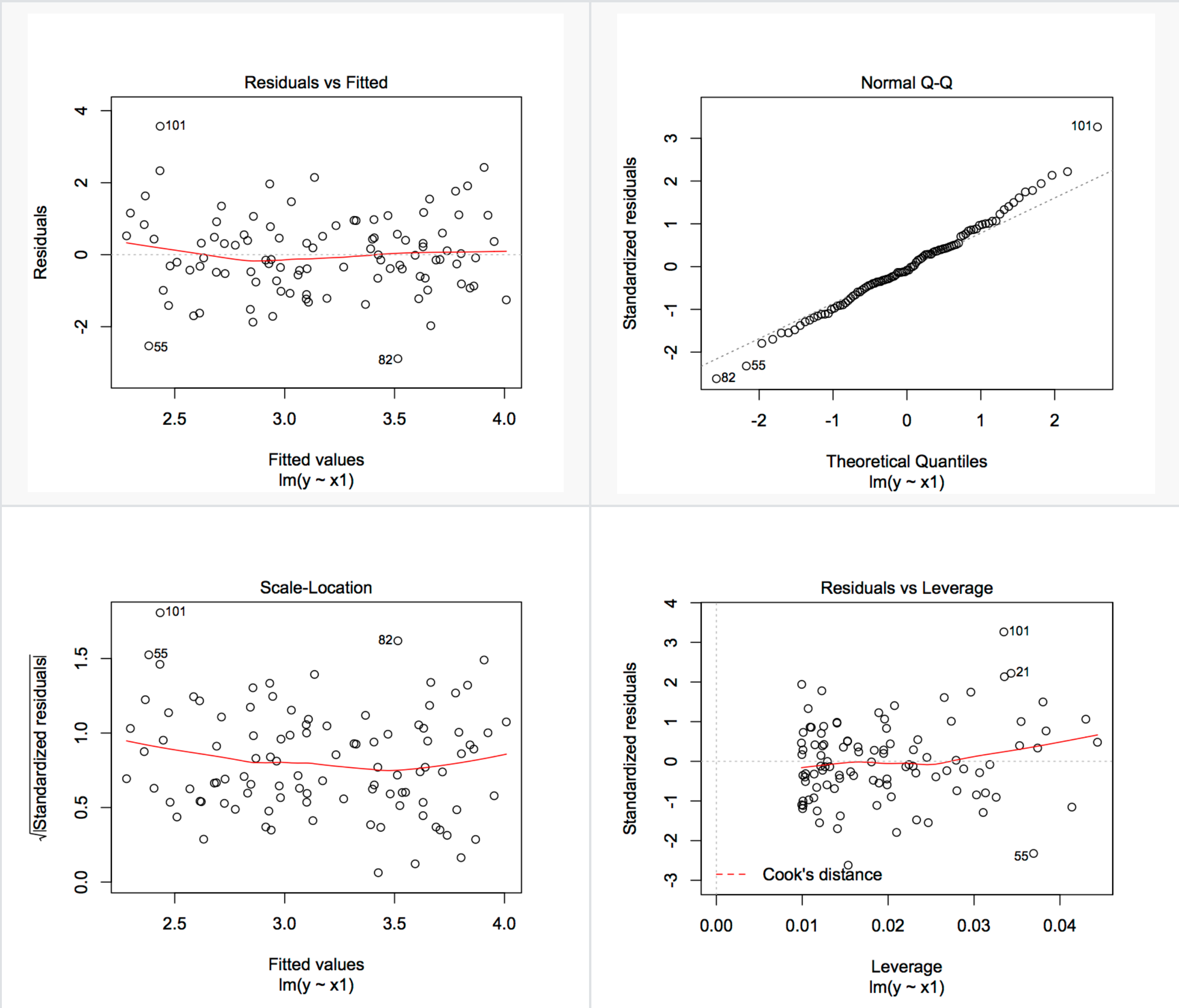



```
# fit2
Call:
lm(formula = y ~ x1)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.8897 -0.6556 -0.0909  0.5682  3.5665 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.2569     0.2390   9.445 1.78e-15 ***
x1            1.7657     0.4124   4.282 4.29e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.111 on 99 degrees of freedom
Multiple R-squared:  0.1562,	Adjusted R-squared:  0.1477 
F-statistic: 18.33 on 1 and 99 DF,  p-value: 4.295e-05
```

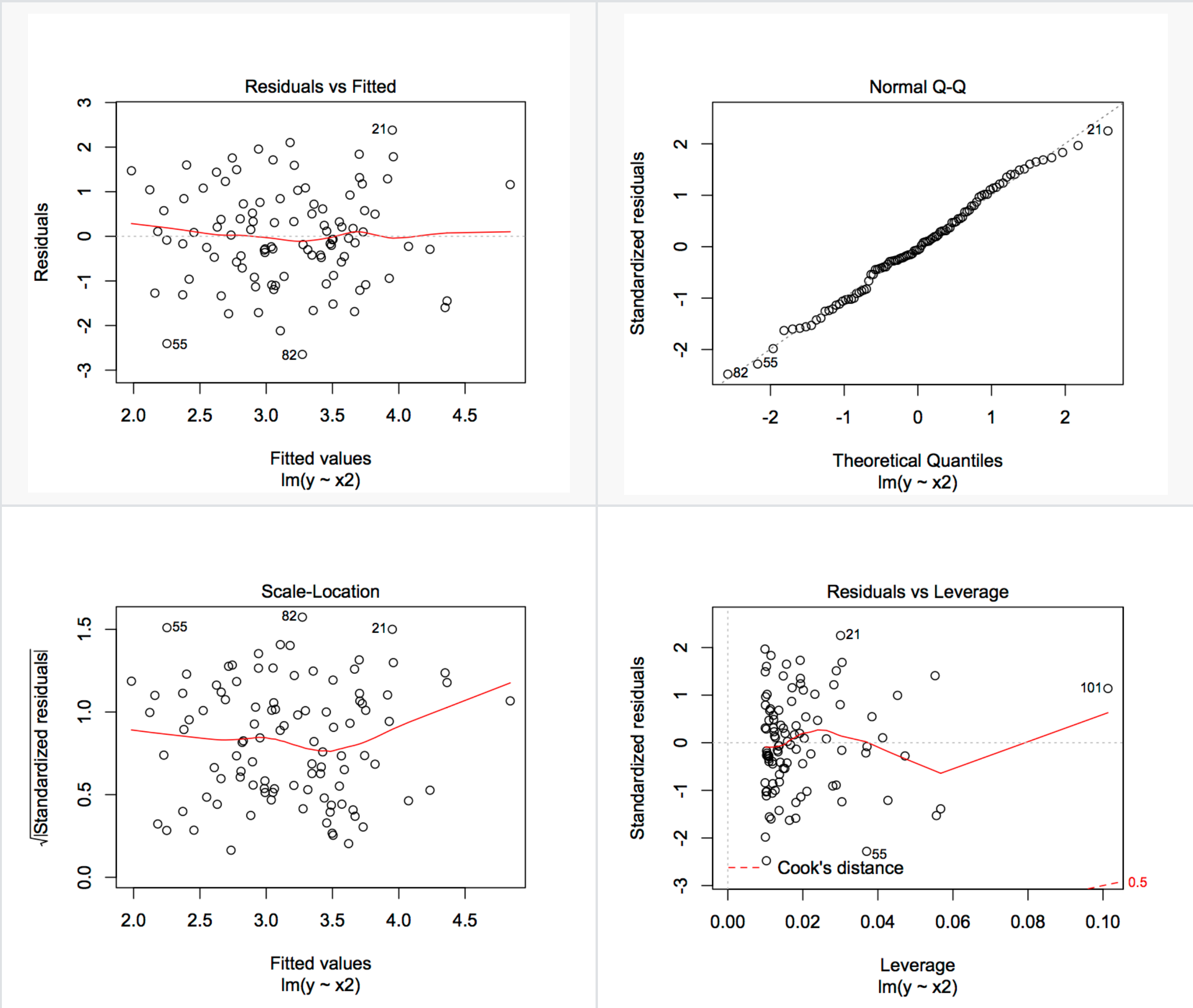



```
# fit3
Call:
lm(formula = y ~ x2)

Residuals:
     Min       1Q   Median       3Q      Max
-2.64729 -0.71021 -0.06899  0.72699  2.38074

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.3451     0.1912  12.264  < 2e-16 ***
x2            3.1190     0.6040   5.164 1.25e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.074 on 99 degrees of freedom
Multiple R-squared:  0.2122,	Adjusted R-squared:  0.2042
F-statistic: 26.66 on 1 and 99 DF,  p-value: 1.253e-06
```



- Two predictors model: the last point is a high-leverage point.
- x1 model: the last point is an outlier.
- x2 model: the last point is a high leverage point.