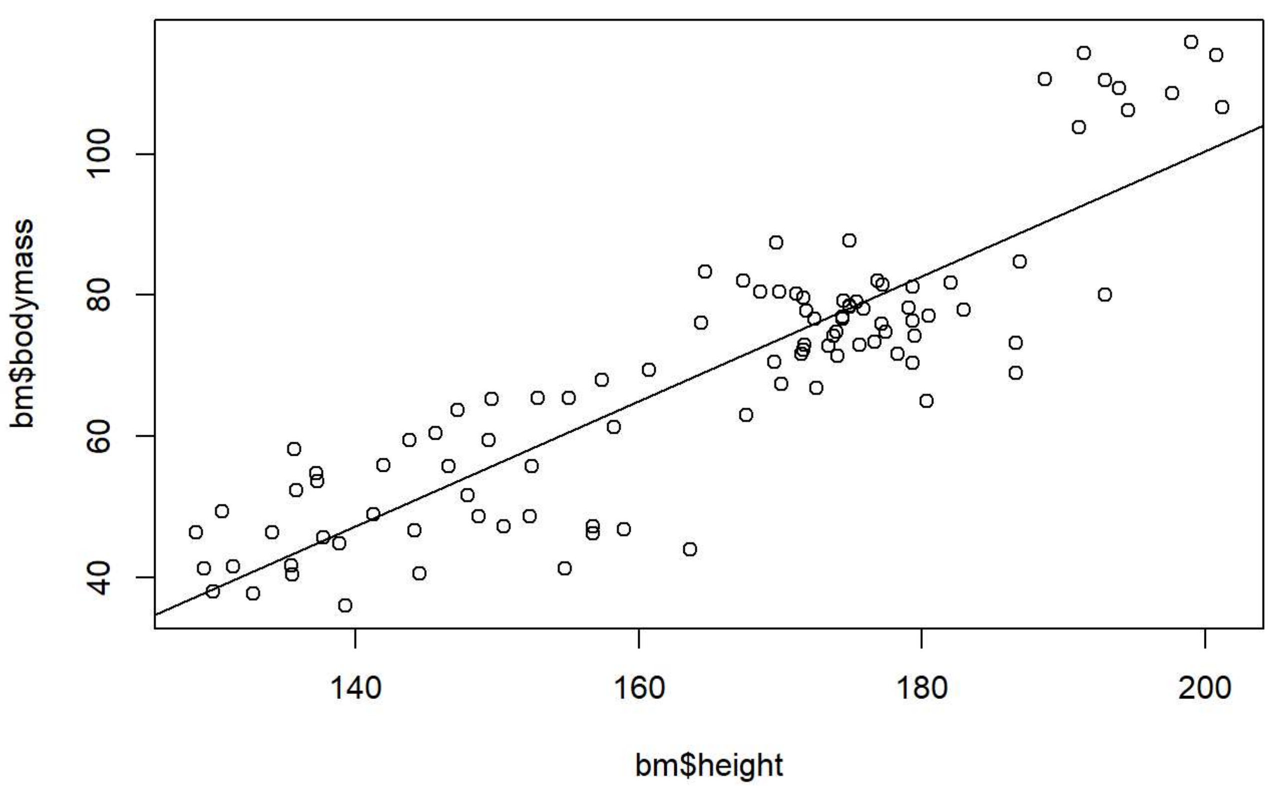
Stats 110 Homework 4

Owen Lin

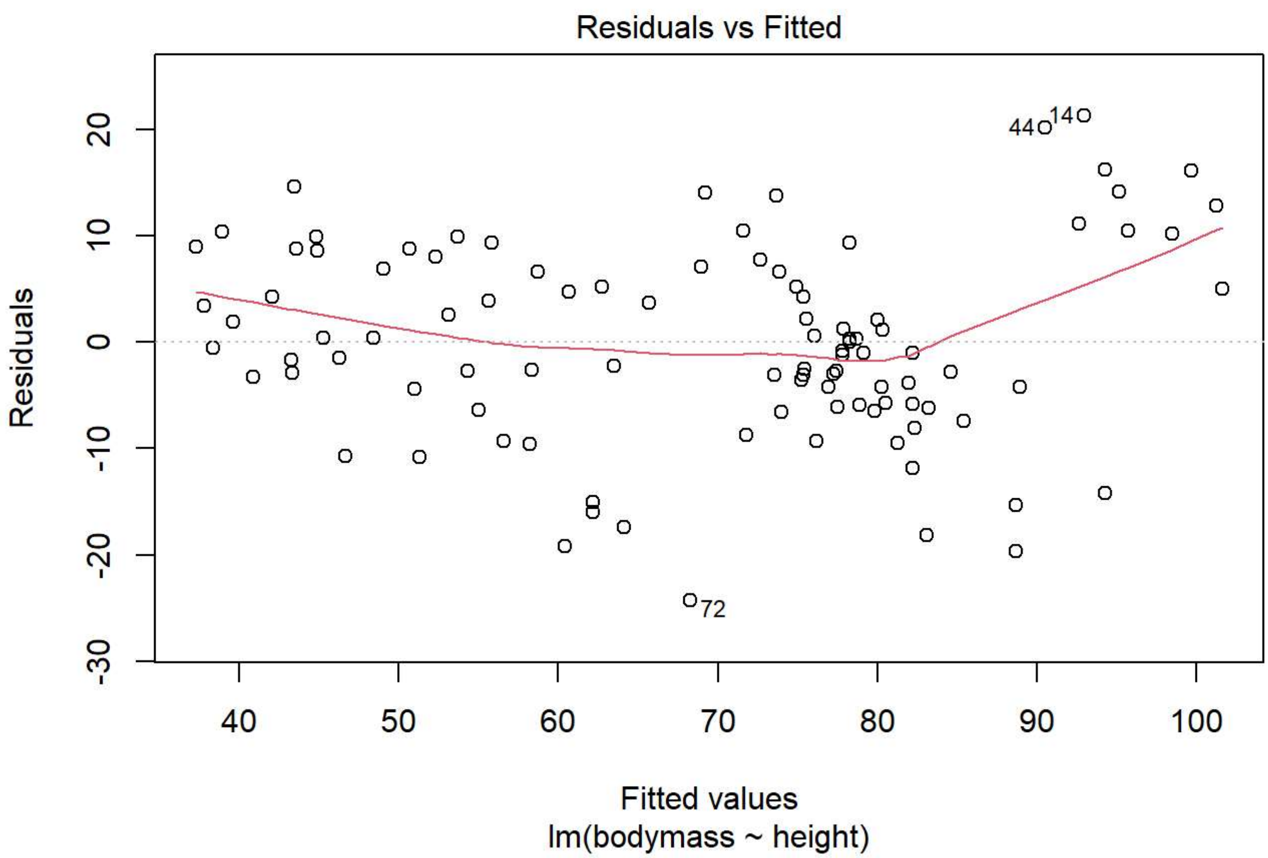
11/16/2022

1. a. This observation will have a small influence on the estimate $\hat{\beta}_1$. The main effect will be on the intercept estimate: the regression line will shift accordingly rather than a change in slope estimate.
   b. This observation will have no influence on the estimate of $\hat{\beta}_1$, because even though it is an outlier in X, it still locates around on the regression line.
   c. This observation will have a large influence on both estimates. The regression line needs to be flatten/steepen in order to minimize the square of residual from the outlier at the cost of all other observations.
2. a. It seems pretty linear other than the top right cluster.
   b. The adjusted r-squared is 0.773.
   c. See output.

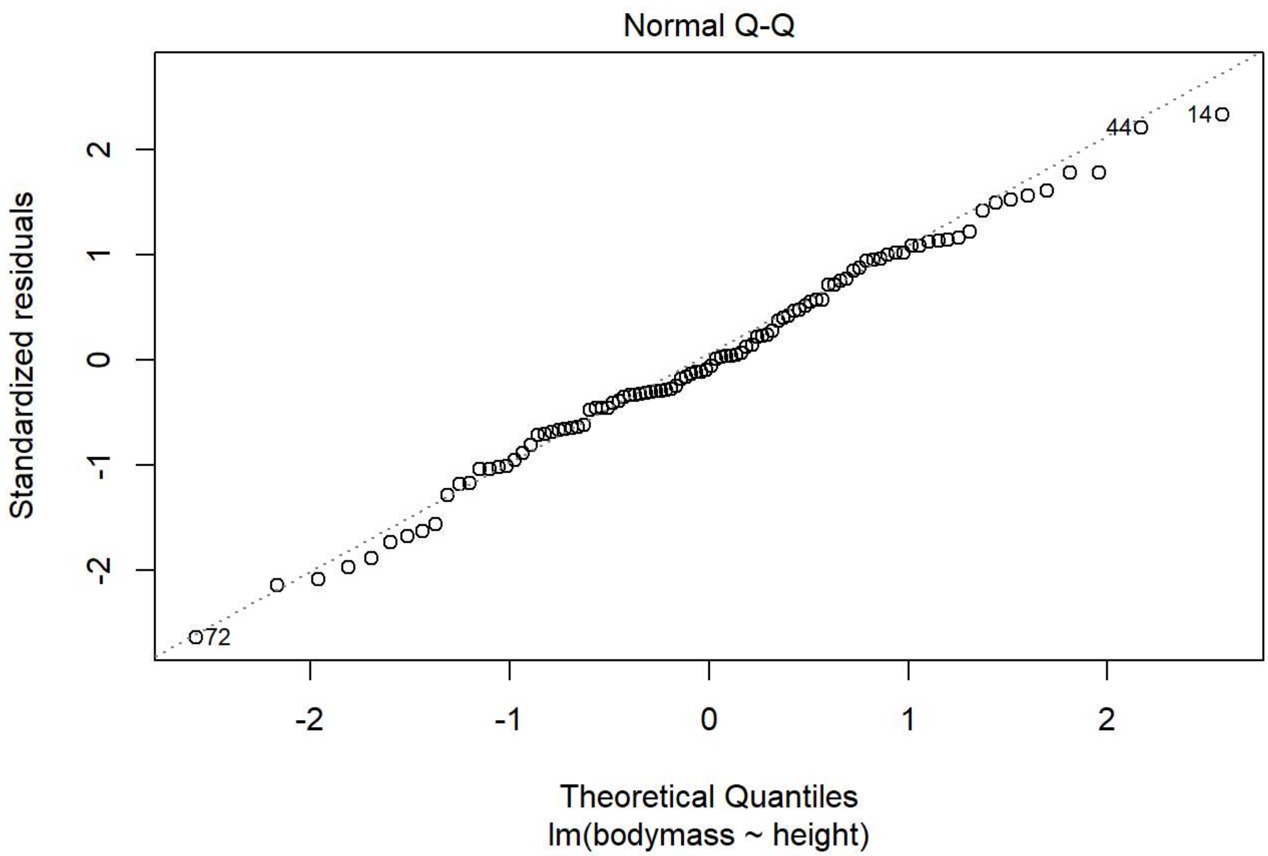
```
bm <- read.csv("bodymass.csv", fill = TRUE, header = TRUE)
plot(bm$height, bm$bodymass)
model1 <- lm(bodymass ~ height, data = bm)
abline(model1)
```



```
summary(model1)$adj.r.squared
plot(model1, 1)
```




```
plot(model1, 2)
```




```
## [1] 0.7730191
```

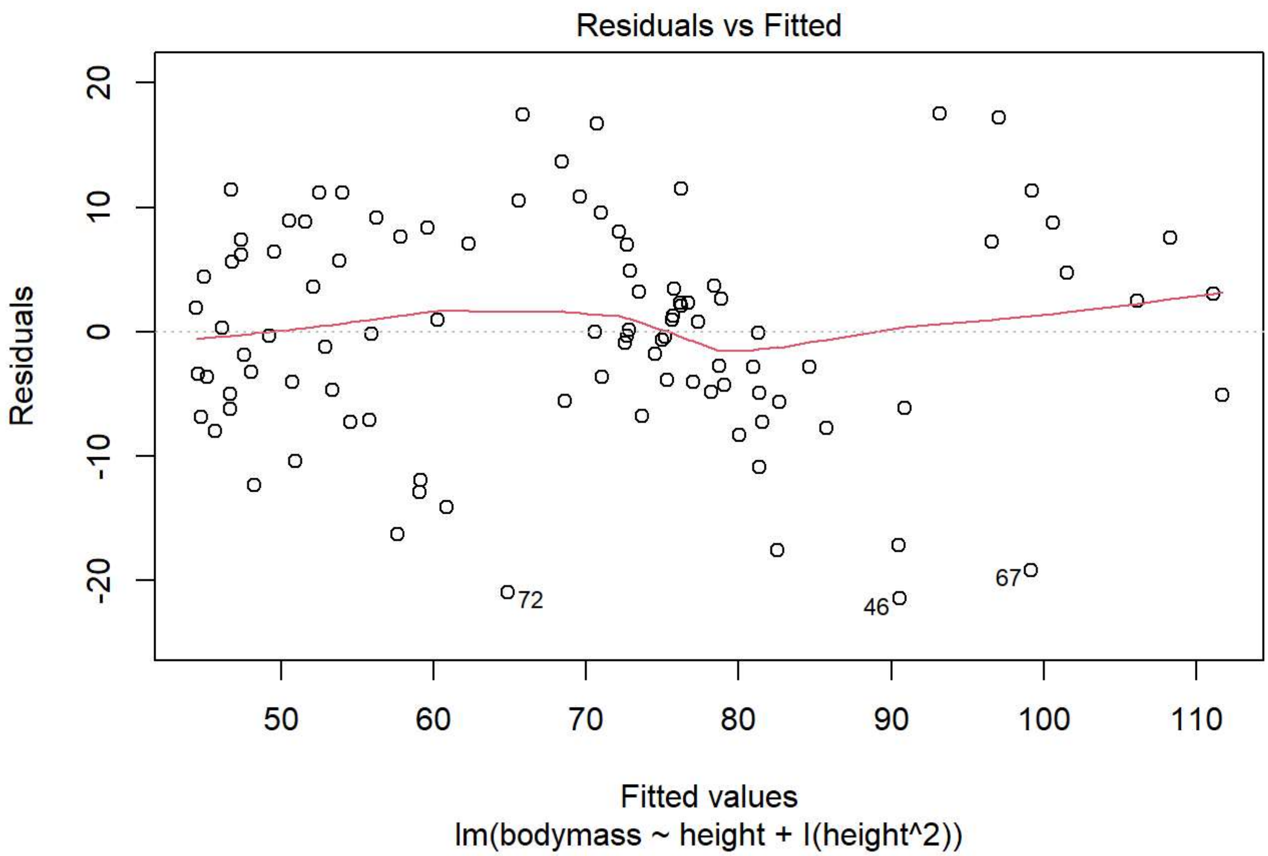
d. Linearity assumptions is invalidated by the graph while normality assumption holds for the linear regression model.
e. See output

```
model2 <- lm(bodymass ~ height + I(height^2), data = bm)
summary(model2)
```

```
##
## Call:
## lm(formula = bodymass ~ height + I(height^2), data = bm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.4796  -5.0268  -0.0875   6.5396  17.5271
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 160.727653  61.146026   2.629 0.009969 **
## height       -2.075533   0.758131  -2.738 0.007361 **
## I(height^2)   0.009104   0.002326   3.914 0.000169 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.668 on 97 degrees of freedom
## Multiple R-squared:  0.806,  Adjusted R-squared:  0.802
## F-statistic: 201.4 on 2 and 97 DF,  p-value: < 2.2e-16
```

f. H0: $\beta_2 = 0$
   Ha: $\beta \neq 0$
   Test-statistic: 3.914 ~ t(97+2+1)
   p-value: 0.000169 < 0.05
   Conclusion: reject the null at 95% significant level and conclude that we need quadratic term in the model.
g. The adjusted r-squared is 0.802 > 0.773, there is an improvement.
h. The line is flatten comparing to the model without quadratic term, but there is still some deviation away from 0.

```
plot(model2, 1)
```




```
summary(model1)
```

```
##
## Call:
## lm(formula = bodymass ~ height, data = bm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -24.3327  -5.8686  -0.6653   6.9646  21.3385
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -76.81867    7.99913  -9.603 8.79e-16 ***
## height        0.88673    0.04822  18.389  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.28 on 98 degrees of freedom
## Multiple R-squared:  0.7753, Adjusted R-squared:  0.773
## F-statistic: 338.2 on 1 and 98 DF,  p-value: < 2.2e-16
```

i. H0: $\beta_1 = \beta_2 = 0$
Test-statistic: 201.4 ~ F(2, 97)

p-value: < 2.2e-16
Conclusion: reject the null at 95% significant level and conclude that height has an influence on bodymass.

j. The vif of height and height^2 is the same. That means regress height^2 on height or regress height on height^2 result in the same R-squared. (In fact, if there's only 2 covariates in the regression equation, the vif will always be the same. We can prove it by the definition R = corr^2 and corr = cov(x,y)/(var(x)*var(y))).

```
vif(model2)
```

```
##      height I(height^2)
##    283.3114    283.3114
```

k. If a subject who is 170cm tall increases height by 1cm, we expect the bodymass to increase by $-2.07 + (171^2 - 170^2) * 0.0091 = 1.0331$ kg.
if the increased height is 5cm, we expect the body mass to increase by $-2.07 * 5 + (175^2 - 170^2) * 0.0091 = 5.3475$ kg.

l. The 64th observation has the highest leverage, the large x-value causes it to have a high leverage (with the highest x-value).

```
bm$leverage <- hatvalues(model2)
head(bm[order(bm$leverage, decreasing=TRUE),])
```

```
##    subject  bodymass   height   leverage
## 64      64 106.60295 201.2101 0.13450811
## 4        4 114.13046 200.8275 0.12913279
## 74      74 115.85467 199.0427 0.10633186
## 84      84 108.61465 197.6424 0.09088713
## 75      75  46.35658 128.7124 0.08884905
## 5        5  41.19526 129.2606 0.08374079
```

```
summary(bm)
```

```
##     subject          bodymass          height         leverage
##  Min.   :  1.00   Min.   : 35.94   Min.   :128.7   Min.   :0.01611
##  1st Qu.: 25.75   1st Qu.: 53.26   1st Qu.:148.5   1st Qu.:0.01689
##  Median : 50.50   Median : 71.70   Median :170.6   Median :0.01942
##  Mean   : 50.50   Mean   : 69.29   Mean   :164.8   Mean   :0.03000
##  3rd Qu.: 75.25   3rd Qu.: 79.07   3rd Qu.:177.3   3rd Qu.:0.03114
##  Max.   :100.00   Max.   :115.85   Max.   :201.2   Max.   :0.13451
```

3. a. cor(male, female) = -1
cor(RtFoot, LeftFoot) = 0.9438
cor(HeadCirc, RtFoot) = 0.4754
cor(HeadCirc, LeftFoot) = 0.4666
cor(HeadCirc, Male) = 0.4894

```
physical <- read.table("PhysicalData.txt", fill = TRUE, header = TRUE)
round(cor(physical), 4)
```

```
##            Height LeftArm   RtArm LeftFoot  RtFoot LeftHand  RtHand HeadCirc
## Height     1.0000  0.7481  0.6949   0.8190  0.8094   0.3690  0.4150   0.4236
## LeftArm    0.7481  1.0000  0.8850   0.6029  0.5930   0.3345  0.3558   0.3044
## RtArm      0.6949  0.8850  1.0000   0.5614  0.6201   0.2877  0.3203   0.3496
## LeftFoot   0.8190  0.6029  0.5614   1.0000  0.9438   0.3105  0.3741   0.4666
## RtFoot     0.8094  0.5930  0.6201   0.9438  1.0000   0.2845  0.3760   0.4754
## LeftHand   0.3690  0.3345  0.2877   0.3105  0.2845   1.0000  0.9353   0.0413
## RtHand     0.4150  0.3558  0.3203   0.3741  0.3760   0.9353  1.0000   0.0927
## HeadCirc   0.4236  0.3044  0.3496   0.4666  0.4754   0.0413  0.0927   1.0000
## nose       0.2843  0.3199  0.2975   0.3040  0.2870   0.1744  0.2170   0.1367
## Female    -0.7110 -0.5687 -0.5108  -0.7729 -0.7174  -0.4950 -0.5309  -0.4894
## Male       0.7110  0.5687  0.5108   0.7729  0.7174   0.4950  0.5309   0.4894
##              nose  Female    Male
## Height     0.2843 -0.7110  0.7110
## LeftArm    0.3199 -0.5687  0.5687
## RtArm      0.2975 -0.5108  0.5108
## LeftFoot   0.3040 -0.7729  0.7729
## RtFoot     0.2870 -0.7174  0.7174
## LeftHand   0.1744 -0.4950  0.4950
## RtHand     0.2170 -0.5309  0.5309
## HeadCirc   0.1367 -0.4894  0.4894
## nose       1.0000 -0.3844  0.3844
## Female    -0.3844  1.0000 -1.0000
## Male       0.3844 -1.0000  1.0000
```

b. The model fail to give a estimate to female because male and female are perfectly multicolinear.

```
model3b <- lm(HeadCirc ~ Male + Female, data = physical)
summary(model3b)
```

```
##
## Call:
## lm(formula = HeadCirc ~ Male + Female, data = physical)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0760 -1.2313  0.1133  1.0187  5.1133
##
## Coefficients: (1 not defined because of singularities)
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  55.8867     0.3613 154.680  < 2e-16 ***
## Male          2.1893     0.5359   4.085  0.00015 ***
## Female            NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.979 on 53 degrees of freedom
## Multiple R-squared:  0.2395, Adjusted R-squared:  0.2251
## F-statistic: 16.69 on 1 and 53 DF,  p-value: 0.0001497
```

c. $HeadCirc = 50.621 + 1.367 * Male + 0.219 * RtFoot$. The adjusted R-squared is 0.2433

```
model3c <- lm(HeadCirc ~ Male + RtFoot, data = physical)
summary(model3c)
```

```
##
## Call:
## lm(formula = HeadCirc ~ Male + RtFoot, data = physical)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0198 -1.3304  0.2419  1.1636  5.1148
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  50.6211     3.5097  14.423   <2e-16 ***
## Male          1.3670     0.7601   1.798   0.0779 .
## RtFoot        0.2193     0.1454   1.508   0.1376
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.956 on 52 degrees of freedom
## Multiple R-squared:  0.2714, Adjusted R-squared:  0.2433
## F-statistic: 9.683 on 2 and 52 DF,  p-value: 0.0002663
```

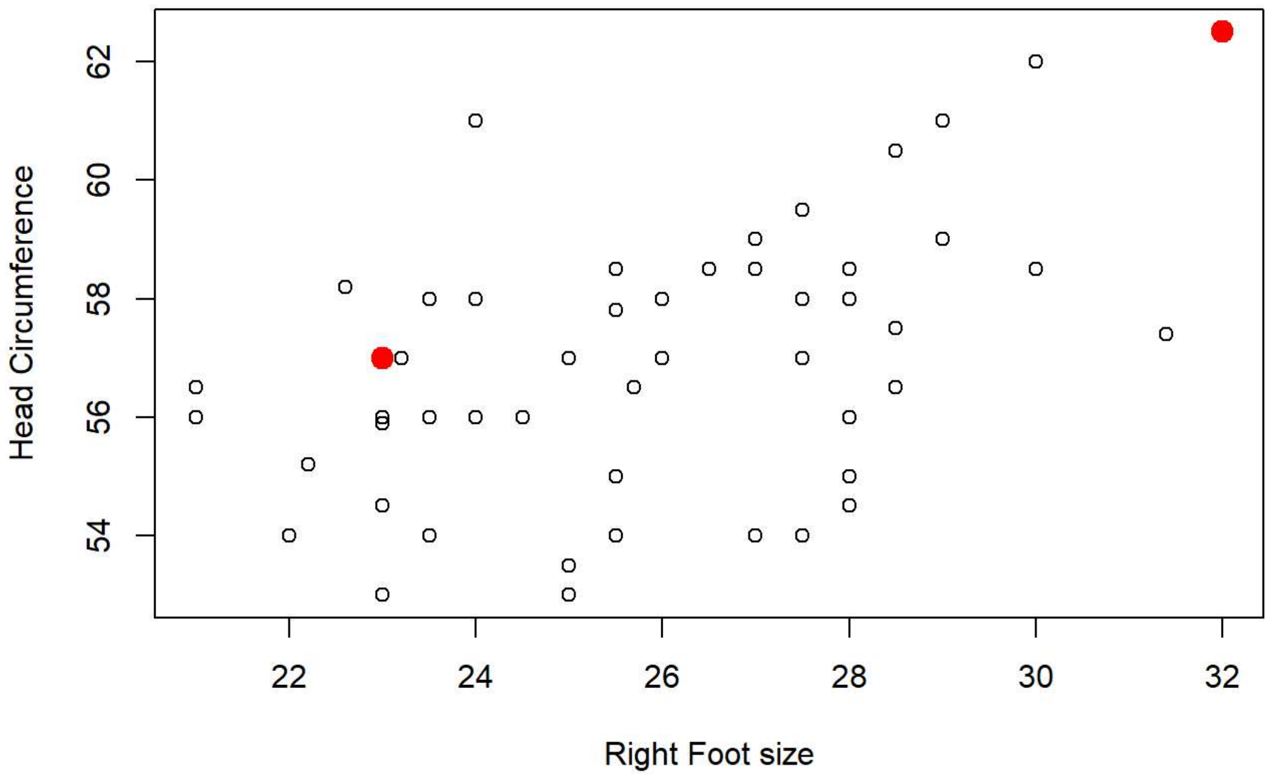
d. Observation 5 and 3 have the two highest leverage values. One has the highest RtFoot value and one has a very low RtFoot value.

```
physical$leverage <- hatvalues(model3c)
head(physical[order(physical$leverage, decreasing = TRUE),])
summary(physical)
```

```
##    Height LeftArm RtArm LeftFoot RtFoot LeftHand RtHand HeadCirc nose Female
## 5      65    25.0  25.0     23.5   23.0     9.50   9.40     57.0  4.4      0
## 3      75    27.0  27.5     31.0   32.0     3.75   3.75     62.5  5.0      0
## 43     71    25.5  28.0     25.5   28.0     7.50   8.50     58.0  5.0      1
## 53     71    25.5  28.0     25.5   28.0     7.50   7.50     54.5  5.0      1
## 25     73    30.0  29.0     23.5   24.0     9.00   9.50     58.0  6.0      0
## 9      73    28.0  28.4     30.6   31.4     8.50   8.90     57.4  6.4      0
##    Male  leverage
## 5     1 0.1651079
## 3     1 0.1396212
## 43    0 0.1215340
## 53    0 0.1215340
## 25    1 0.1180283
## 9     1 0.1134442
##      Height         LeftArm          RtArm          LeftFoot    
##  Min.   :61.00   Min.   :22.00   Min.   :21.50   Min.   :21.00  
##  1st Qu.:64.75   1st Qu.:23.75   1st Qu.:24.00   1st Qu.:23.50  
##  Median :68.00   Median :25.50   Median :25.50   Median :25.50  
##  Mean   :68.31   Mean   :25.47   Mean   :25.48   Mean   :25.59  
##  3rd Qu.:71.75   3rd Qu.:26.75   3rd Qu.:27.00   3rd Qu.:27.70  
##  Max.   :79.00   Max.   :31.00   Max.   :30.50   Max.   :31.00  
##      RtFoot         LeftHand          RtHand          HeadCirc    
##  Min.   :21.00   Min.   : 3.750   Min.   : 3.750   Min.   :53.00  
##  1st Qu.:23.35   1st Qu.: 7.500   1st Qu.: 7.450   1st Qu.:55.55  
##  Median :25.50   Median : 8.000   Median : 8.200   Median :57.00  
##  Mean   :25.71   Mean   : 8.237   Mean   : 8.228   Mean   :56.88  
##  3rd Qu.:27.75   3rd Qu.: 9.050   3rd Qu.: 9.000   3rd Qu.:58.35  
##  Max.   :32.00   Max.   :11.500   Max.   :11.500   Max.   :62.50  
##       nose           Female            Male           leverage      
##  Min.   :4.000   Min.   :0.0000   Min.   :0.0000   Min.   :0.03333  
##  1st Qu.:4.500   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.03894  
##  Median :5.000   Median :1.0000   Median :0.0000   Median :0.04306  
##  Mean   :4.947   Mean   :0.5455   Mean   :0.4545   Mean   :0.05455  
##  3rd Qu.:5.050   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.05335  
##  Max.   :6.500   Max.   :1.0000   Max.   :1.0000   Max.   :0.16511  
```

e. See Output

```
plot(physical$RtFoot, physical$HeadCirc, xlab="Right Foot size" , ylab="Head Circumference")
points(head(physical[order(physical$leverage, decreasing=TRUE),])[1,"RtFoot"], head(physical[ord
er(physical$leverage, decreasing=TRUE),])[1,"HeadCirc"], col="red", cex=1.5 , pch=19)
points(head(physical[order(physical$leverage, decreasing=TRUE),])[2,"RtFoot"], head(physical[ord
er(physical$leverage, decreasing=TRUE),])[2,"HeadCirc"], col="red", cex=1.5 , pch=19)
```



f. $HeadCirc = 50.621 - 0.1132 * LeftFoot + 0.3057 * RtFoot + 1.478 * Male$. The adjusted R-squared is 0.23.

```
model3f <- lm(HeadCirc ~ LeftFoot + RtFoot + Male, data = physical)
summary(model3f)
```

```
##
## Call:
## lm(formula = HeadCirc ~ LeftFoot + RtFoot + Male, data = physical)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9001 -1.3365  0.3252  1.1663  5.1327
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  51.2471     4.0556  12.636   <2e-16 ***
## LeftFoot     -0.1132     0.3576  -0.316   0.7530
## RtFoot        0.3057     0.3098   0.987   0.3284
## Male          1.4780     0.8432   1.753   0.0857 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.973 on 51 degrees of freedom
## Multiple R-squared:  0.2728, Adjusted R-squared:   0.23
## F-statistic: 6.377 on 3 and 51 DF,  p-value: 0.0009392
```

g. No, adding RtFoot to the model only adds minimal amount to SSR (3.065)

```
anova(model3f)
1.973^2 * (55-4)
```

```
## Analysis of Variance Table
##
## Response: HeadCirc
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## LeftFoot   1  59.429  59.429 15.2710 0.0002754 ***
## RtFoot     1   3.065   3.065  0.7875 0.3790132
## Male       1  11.956  11.956  3.0722 0.0856515 .
## Residuals 51 198.472   3.892
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## [1] 198.5292
```

h. SSE: $1.973^2 * (55-4) = 198.5292$
MSE: $SSE/n = 198.5292/55 = 3.6096$

i. Both LeftFoot and RtFoot have a high VIF, indicating that we should drop one of them to avoid collinearity.

```
vif(model3f)
```

```
##  LeftFoot    RtFoot      Male
## 11.072450  9.186436  2.491458
```

j. Observation 43 and 53 have the two highest leverage values.Both have a large RtFoot value of 28.

```
physical$leverage3f <- hatvalues(model3f)
head(physical[order(physical$leverage3f, decreasing = TRUE),])
summary(physical)
```

```
##    Height LeftArm RtArm LeftFoot RtFoot LeftHand RtHand HeadCirc nose Female
## 43     71    25.5  28.0     25.5     28     7.50   8.50     58.0  5.0      1
## 53     71    25.5  28.0     25.5     28     7.50   7.50     54.5  5.0      1
## 25     73    30.0  29.0     23.5     24     9.00   9.50     58.0  6.0      0
## 5      65    25.0  25.0     23.5     23     9.50   9.40     57.0  4.4      0
## 13     71    28.0  27.0     29.0     27     8.00   8.00     59.0  6.0      0
## 3      75    27.0  27.5     31.0     32     3.75   3.75     62.5  5.0      0
##    Male   leverage leverage3f
## 43    0 0.12153404  0.1853391
## 53    0 0.12153404  0.1853391
## 25    1 0.11802829  0.1755097
## 5     1 0.16510787  0.1753975
## 13    1 0.04316115  0.1604163
## 3     1 0.13962118  0.1398018
##      Height         LeftArm          RtArm          LeftFoot
##  Min.   :61.00   Min.   :22.00   Min.   :21.50   Min.   :21.00
##  1st Qu.:64.75   1st Qu.:23.75   1st Qu.:24.00   1st Qu.:23.50
##  Median :68.00   Median :25.50   Median :25.50   Median :25.50
##  Mean   :68.31   Mean   :25.47   Mean   :25.48   Mean   :25.59
##  3rd Qu.:71.75   3rd Qu.:26.75   3rd Qu.:27.00   3rd Qu.:27.70
##  Max.   :79.00   Max.   :31.00   Max.   :30.50   Max.   :31.00
##      RtFoot         LeftHand          RtHand          HeadCirc
##  Min.   :21.00   Min.   : 3.750   Min.   : 3.750   Min.   :53.00
##  1st Qu.:23.35   1st Qu.: 7.500   1st Qu.: 7.450   1st Qu.:55.55
##  Median :25.50   Median : 8.000   Median : 8.200   Median :57.00
##  Mean   :25.71   Mean   : 8.237   Mean   : 8.228   Mean   :56.88
##  3rd Qu.:27.75   3rd Qu.: 9.050   3rd Qu.: 9.000   3rd Qu.:58.35
##  Max.   :32.00   Max.   :11.500   Max.   :11.500   Max.   :62.50
##       nose           Female            Male           leverage
##  Min.   :4.000   Min.   :0.0000   Min.   :0.0000   Min.   :0.03333
##  1st Qu.:4.500   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.03894
##  Median :5.000   Median :1.0000   Median :0.0000   Median :0.04306
##  Mean   :4.947   Mean   :0.5455   Mean   :0.4545   Mean   :0.05455
##  3rd Qu.:5.050   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.05335
##  Max.   :6.500   Max.   :1.0000   Max.   :1.0000   Max.   :0.16511
##    leverage3f
##  Min.   :0.03416
##  1st Qu.:0.04490
##  Median :0.06190
##  Mean   :0.07273
##  3rd Qu.:0.07916
##  Max.   :0.18534
```

k. p-value: 0.3193
   Conclusion: fail to reject the null, adding foot size to the model didn't improve the prediction of HeadCirc.

```
model3f_reduce <- lm(HeadCirc ~ Male, data = physical)
anova(model3f_reduce, model3f)
```

```
## Analysis of Variance Table
##
## Model 1: HeadCirc ~ Male
## Model 2: HeadCirc ~ LeftFoot + RtFoot + Male
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     53 207.56
## 2     51 198.47  2    9.0878 1.1676 0.3193
```

l. $HeadCirc = \beta_0 + \beta_1 * Male + \beta_2 * LeftHand$. The adjusted R-squared is 0.2658.

```
full <- lm(HeadCirc~.-Female, data=physical)
forward = step(lm(HeadCirc~1, data=physical), scope=list(upper=full), direction="forward")
model3l <- lm(HeadCirc ~ Male + LeftHand, data = physical)
summary(model3l)$adj.r.squared
```

```
## Start:  AIC=90.1
## HeadCirc ~ 1
##
##              Df Sum of Sq    RSS    AIC
## + Male        1    65.362 207.56 77.045
## + RtFoot      1    61.692 211.23 78.009
## + LeftFoot    1    59.429 213.49 78.595
## + Height      1    48.966 223.96 81.226
## + RtArm       1    33.355 239.57 84.932
## + LeftArm     1    25.281 247.64 86.756
## + leverage    1    14.410 258.51 89.119
## <none>                    272.92 90.102
## + nose        1     5.100 267.82 91.064
## + leverage3f  1     3.103 269.82 91.473
## + RtHand      1     2.347 270.57 91.627
## + LeftHand    1     0.466 272.46 92.008
##
## Step:  AIC=77.04
## HeadCirc ~ Male
##
##              Df Sum of Sq    RSS    AIC
## + LeftHand    1   14.5913 192.97 75.036
## + RtHand      1   10.6065 196.95 76.160
## + RtFoot      1    8.6981 198.86 76.690
## <none>                    207.56 77.045
## + LeftFoot    1    5.2989 202.26 77.623
## + leverage    1    4.7546 202.81 77.770
## + RtArm       1    3.6658 203.90 78.065
## + Height      1    3.1559 204.40 78.202
## + nose        1    0.8471 206.71 78.820
## + LeftArm     1    0.2740 207.29 78.972
## + leverage3f  1    0.1168 207.44 79.014
##
## Step:  AIC=75.04
## HeadCirc ~ Male + LeftHand
##
##              Df Sum of Sq    RSS    AIC
## <none>                    192.97 75.036
## + RtFoot      1    6.3574 186.61 75.193
## + RtArm       1    4.3905 188.58 75.770
## + Height      1    3.5484 189.42 76.015
## + LeftFoot    1    3.3061 189.66 76.085
## + leverage    1    1.5787 191.39 76.584
## + nose        1    0.9927 191.98 76.752
## + LeftArm     1    0.6545 192.31 76.849
## + RtHand      1    0.3247 192.64 76.943
## + leverage3f  1    0.0197 192.95 77.030
## [1] 0.2657574
```

m. 2-covariates: $HeadCirc = \beta_0 + \beta_1 * LeftHand + \beta_2 * Male$
   3-covariates: $HeadCirc = \beta_0 + \beta_1 * RtFoot + \beta_2 * LeftHand + \beta_3 * Male$

```
subsets <- regsubsets(HeadCirc~.-Female, data = physical)
summary(subsets)
```

```
## Subset selection object
## Call: regsubsets.formula(HeadCirc ~ . - Female, data = physical)
## 11 Variables  (and intercept)
##            Forced in Forced out
## Height         FALSE      FALSE
## LeftArm        FALSE      FALSE
## RtArm          FALSE      FALSE
## LeftFoot       FALSE      FALSE
## RtFoot         FALSE      FALSE
## LeftHand       FALSE      FALSE
## RtHand         FALSE      FALSE
## nose           FALSE      FALSE
## Male           FALSE      FALSE
## leverage       FALSE      FALSE
## leverage3f     FALSE      FALSE
## 1 subsets of each size up to 8
## Selection Algorithm: exhaustive
##          Height LeftArm RtArm LeftFoot RtFoot LeftHand RtHand nose Male
## 1  ( 1 ) " "    " "     " "   " "      " "    " "      " "    " "  "*"
## 2  ( 1 ) " "    " "     " "   " "      " "    "*"      " "    " "  "*"
## 3  ( 1 ) " "    " "     " "   " "      "*"    "*"      " "    " "  "*"
## 4  ( 1 ) " "    " "     "*"   " "      "*"    "*"      " "    " "  "*"
## 5  ( 1 ) " "    " "     " "   " "      "*"    "*"      " "    " "  "*"
## 6  ( 1 ) " "    " "     "*"   "*"      " "    "*"      " "    " "  "*"
## 7  ( 1 ) "*"    "*"     "*"   " "      " "    "*"      " "    " "  "*"
## 8  ( 1 ) "*"    "*"     "*"   "*"      " "    "*"      " "    " "  "*"
##          leverage leverage3f
## 1  ( 1 ) " "      " "
## 2  ( 1 ) " "      " "
## 3  ( 1 ) " "      " "
## 4  ( 1 ) " "      " "
## 5  ( 1 ) "*"      "*"
## 6  ( 1 ) "*"      "*"
## 7  ( 1 ) "*"      "*"
## 8  ( 1 ) "*"      "*"
```

n. $HeadCirc = 46.047 + 0.148 * LeftFoot + 0.274 * RtFoot$
VIF is the same for both, and they are pretty high at 9.156.

```
model3n <- lm(HeadCirc ~ LeftFoot + RtFoot, data = physical)
model3n$coef
vif(model3n)
```

```
## (Intercept)    LeftFoot      RtFoot
##  46.0470850   0.1476146   0.2744700
## LeftFoot   RtFoot
## 9.156082 9.156082
```

o. The estimated coefficient of the slope is 0.42, more than doubled comparing to model3n. Here's the logic: as LeftFoot increases, RtFoot is likely to increases as well. From model3n, we would expect HeadCirc to increase from effect of both LeftFoot and RtFoot. But sincet Foot is no long included in model3o, LeftFoot as the only regressor needs to take both effect into account (effect of RtFoot on HeadCirc is unseen in model3o).

```
model3o <- lm(HeadCirc ~ LeftFoot, data = physical)
model3o$coef
```

```
## (Intercept)    LeftFoot
##  46.1325423   0.4200129
```