Supplement for Lecture 17: Techniques for Choosing Predictors

Load Data

```
data("BodyFat") # Load Data

bf = BodyFat

head(bf)
```

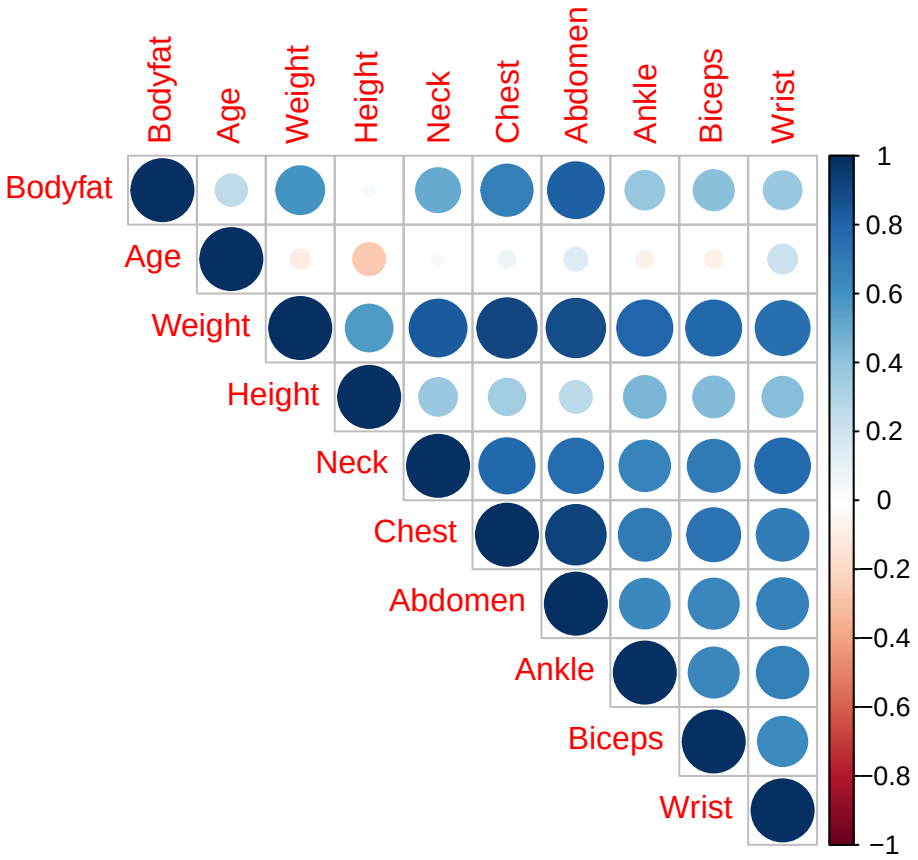
```
##   Bodyfat Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist
## 1    32.3  41 247.25  73.50 42.1 117.0   115.6  26.3   37.3  19.7
## 2    22.5  31 177.25  71.50 36.2 101.1    92.4  24.6   30.1  18.2
## 3    22.0  42 156.25  69.00 35.5  97.8    86.0  24.0   31.2  17.4
## 4    12.3  23 154.25  67.75 36.2  93.1    85.2  21.9   32.0  17.1
## 5    20.5  46 177.00  70.00 37.2  99.7    95.6  22.5   29.1  17.7
## 6    22.6  54 198.00  72.00 39.9 107.6   100.0  22.0   35.9  18.9
```

Check for Multicollinearity

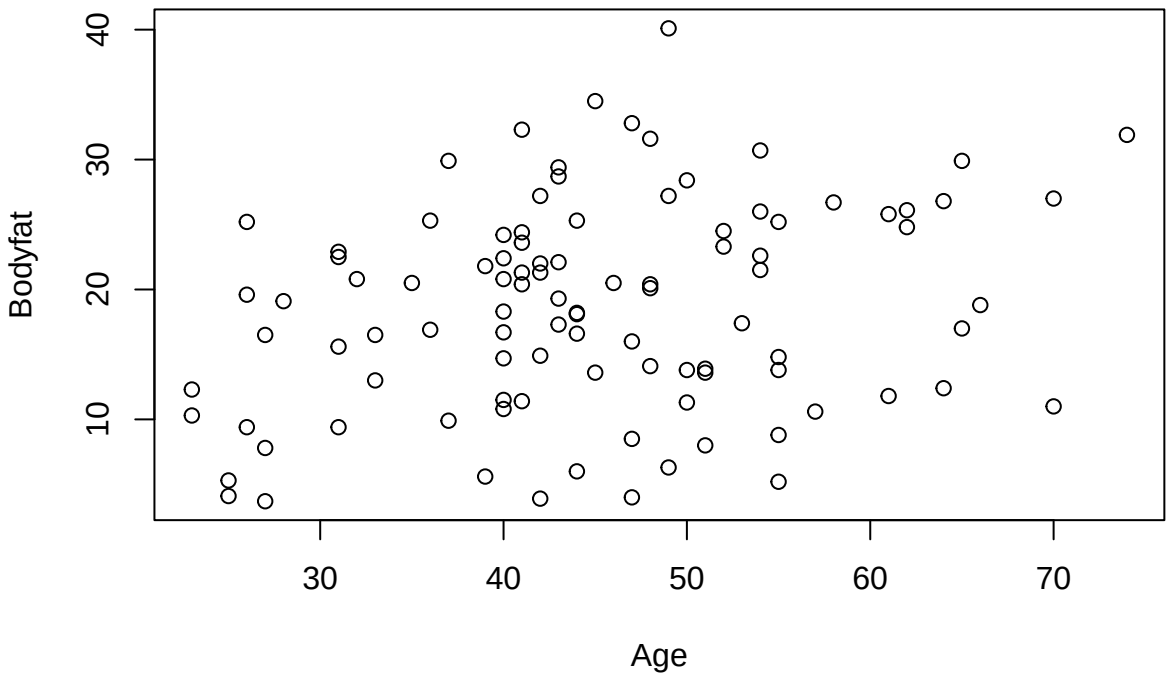
```
# Correlation Matrix from Base R
round(cor(bf),4)
```

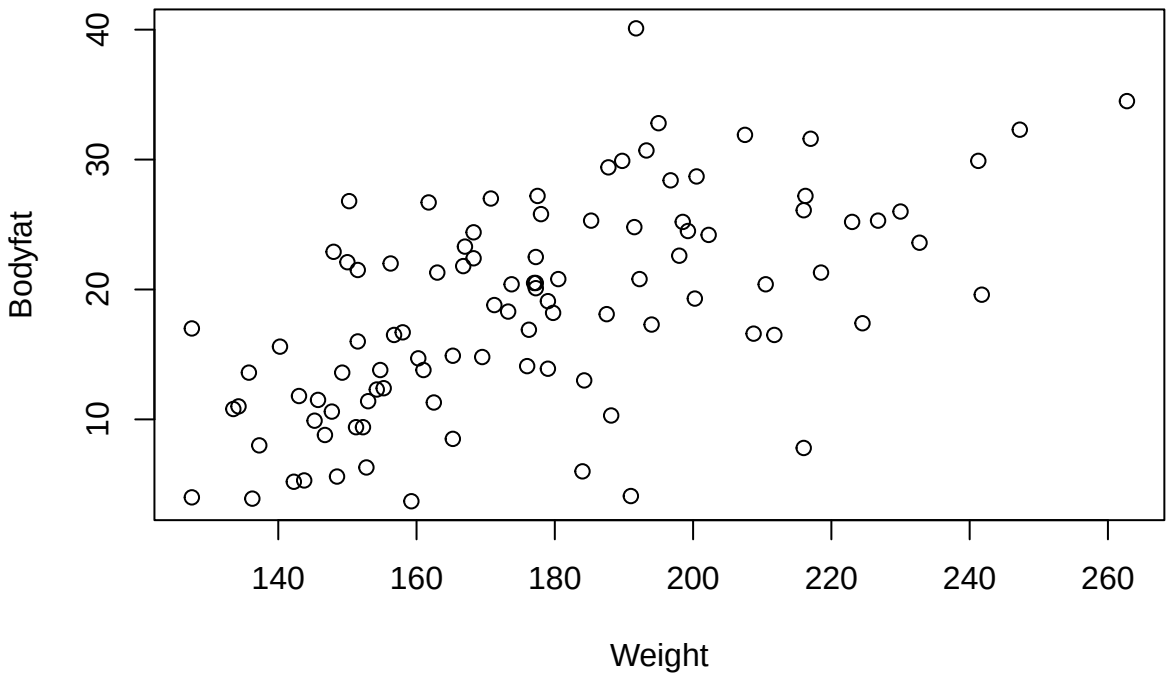
```
##         Bodyfat     Age  Weight  Height   Neck  Chest Abdomen   Ankle  Biceps
## Bodyfat  1.0000  0.2558  0.5960  0.0322 0.5044 0.6823  0.8116  0.3879  0.4181
## Age      0.2558  1.0000 -0.0979 -0.2687 0.0385 0.0729  0.1413 -0.0735 -0.0794
## Weight   0.5960 -0.0979  1.0000  0.5684 0.8390 0.9160  0.8883  0.7904  0.7863
## Height   0.0322 -0.2687  0.5684  1.0000 0.3704 0.3453  0.2654  0.4550  0.4355
## Neck     0.5044  0.0385  0.8390  0.3704 1.0000 0.7898  0.7691  0.6641  0.7011
## Chest    0.6823  0.0729  0.9160  0.3453 0.7898 1.0000  0.9227  0.7001  0.7301
## Abdomen  0.8116  0.1413  0.8883  0.2654 0.7691 0.9227  1.0000  0.6432  0.6521
## Ankle    0.3879 -0.0735  0.7904  0.4550 0.6641 0.7001  0.6432  1.0000  0.6508
## Biceps   0.4181 -0.0794  0.7863  0.4355 0.7011 0.7301  0.6521  0.6508  1.0000
## Wrist    0.3731  0.2163  0.7551  0.4209 0.7773 0.6958  0.6721  0.6883  0.6305
##           Wrist
## Bodyfat  0.3731
## Age      0.2163
## Weight   0.7551
## Height   0.4209
## Neck     0.7773
## Chest    0.6958
## Abdomen  0.6721
## Ankle    0.6883
## Biceps   0.6305
## Wrist    1.0000
```

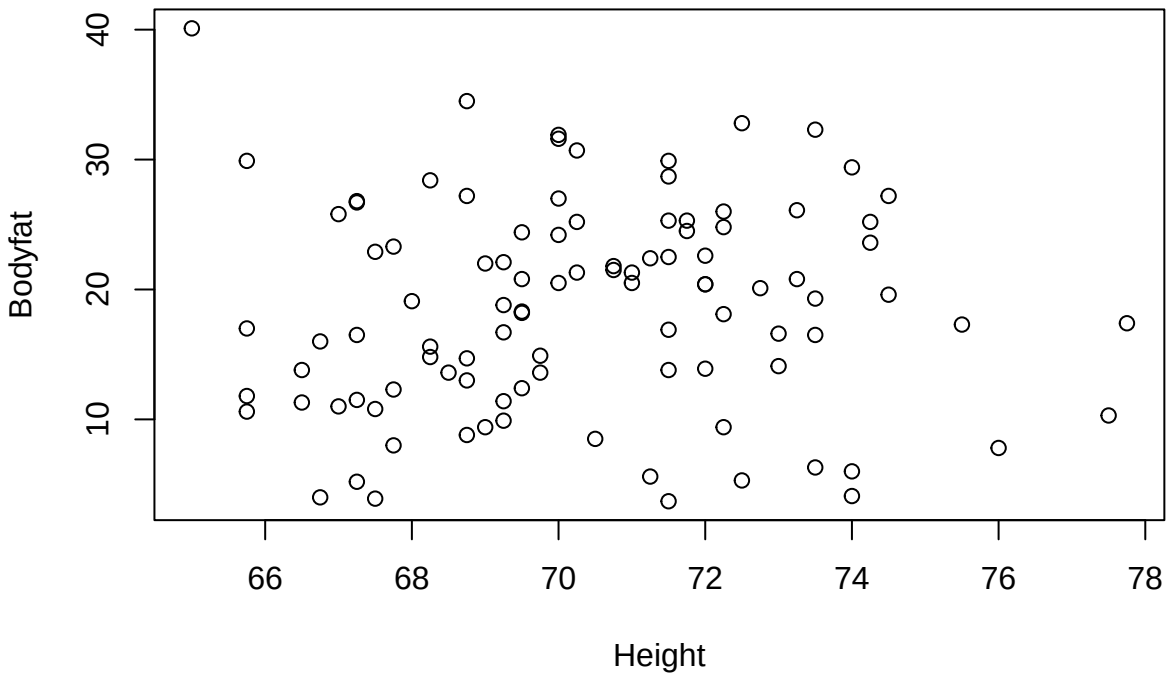
```
# Tile Plot of Correlation Matrix (Correlogram) from corrplot package
corrplot(cor(bf), type="upper")
```

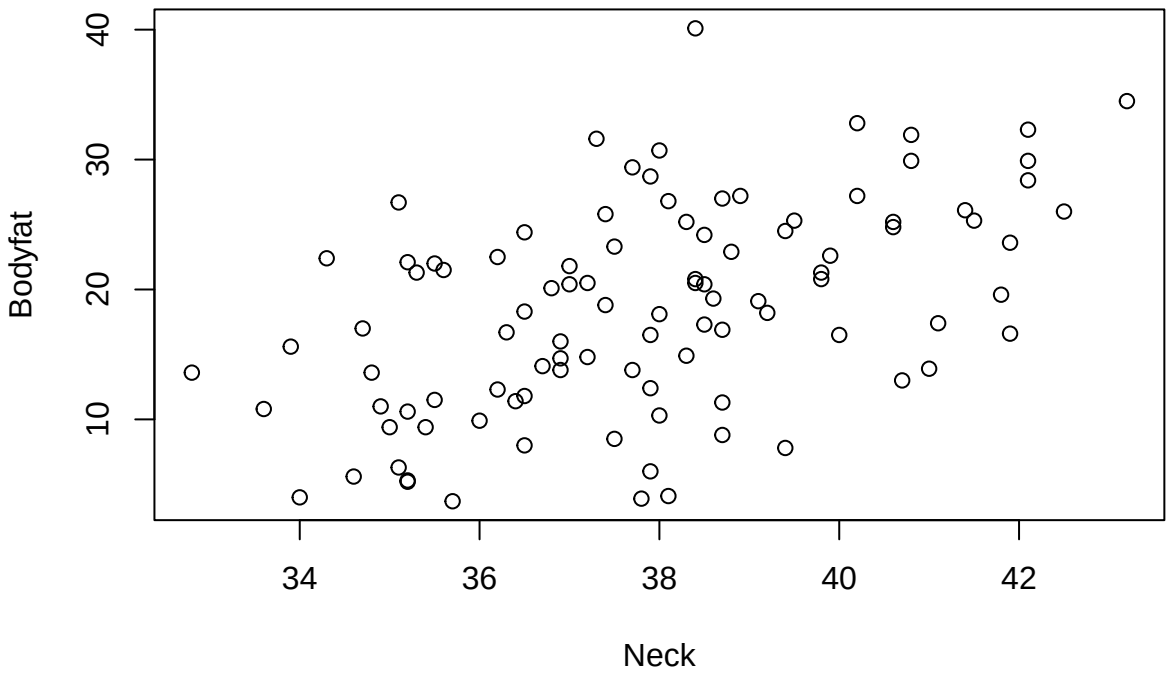



```
#Scatterplots of Bodyfat Variable with Each Other Predictor
plot(Bodyfat~., data=bf)
```

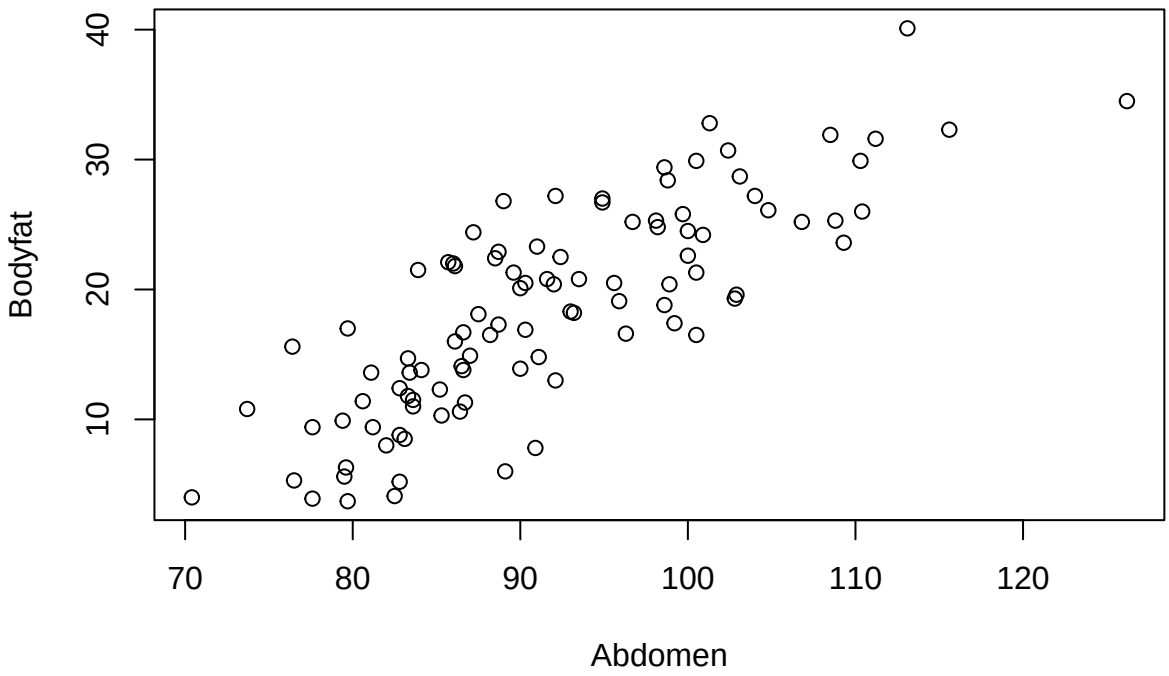

Bodyfat
10
20
30
40
30
40
50
60
70
Age

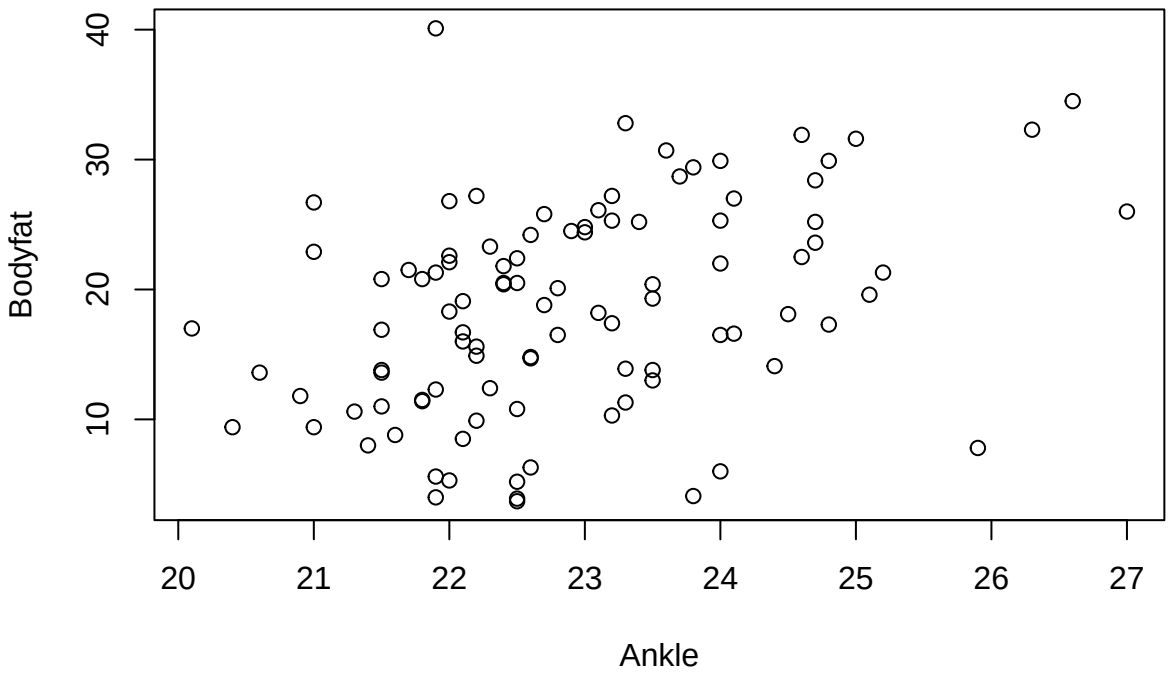

Bodyfat
10
20
30
40
140
160
180
200
220
240
260
Weight

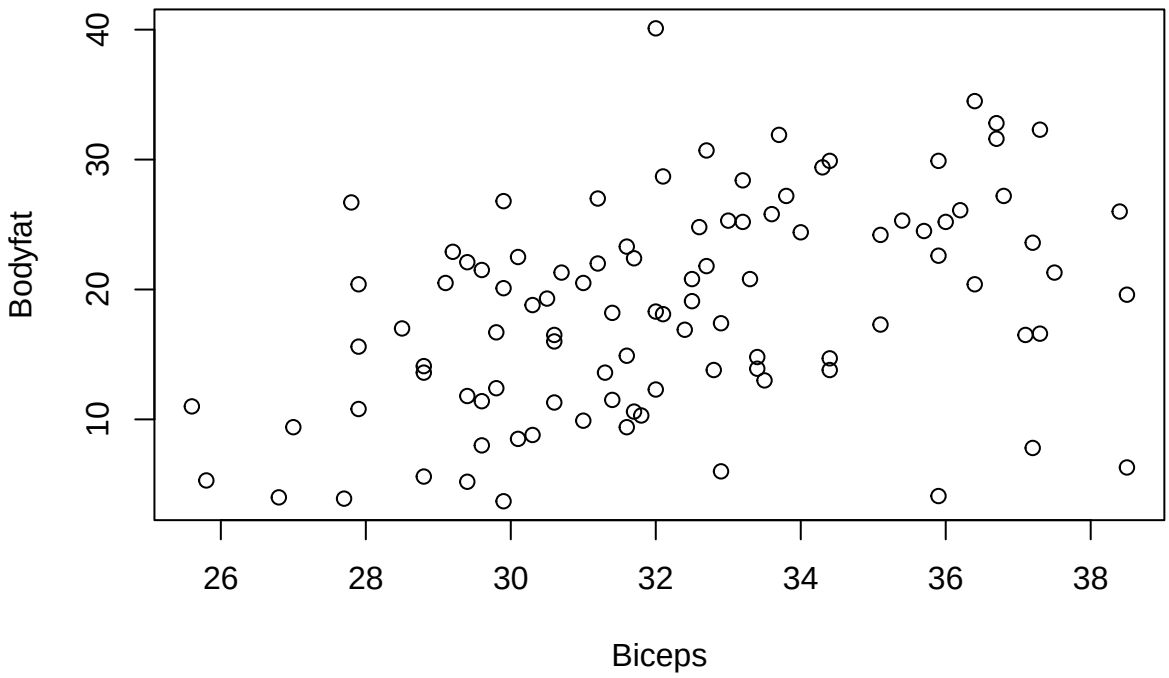

Bodyfat
10
20
30
40
66
68
70
72
74
76
78
Height

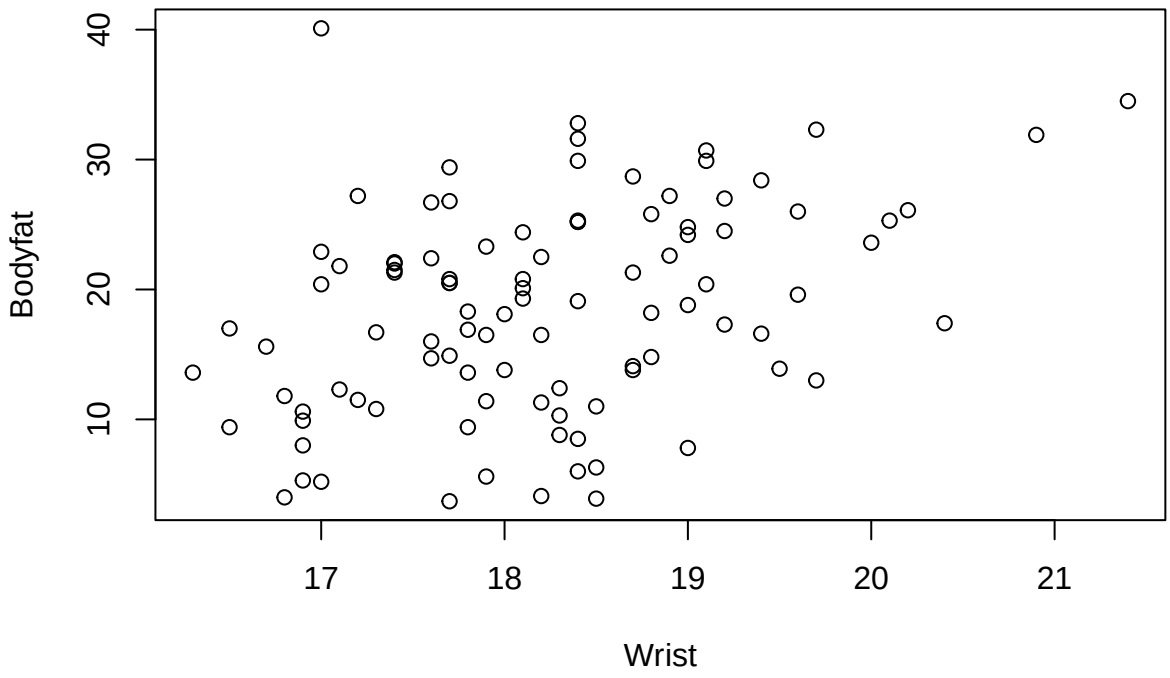

Bodyfat
10
20
30
40
34
36
38
40
42
Neck


Bodyfat
10
20
30
40
90
100
110
120
130
Chest

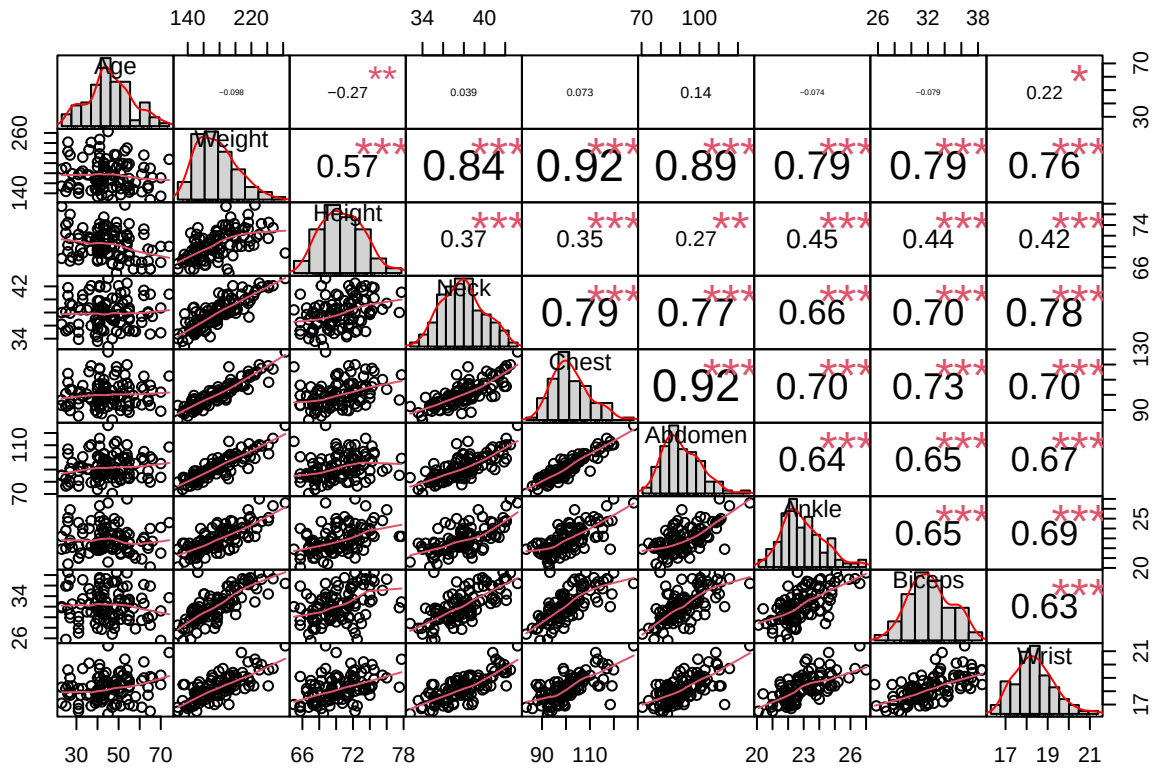

Bodyfat
10
20
30
40
70
80
90
100
110
120
Abdomen


Bodyfat
10
20
30
40
20
21
22
23
24
25
26
27
Ankle


Bodyfat
10
20
30
40
26
28
30
32
34
36
38
Biceps




```
# Cool Visual from PerformanceAnalytics package
chart.Correlation(bf[,-1], histogram=TRUE)
```



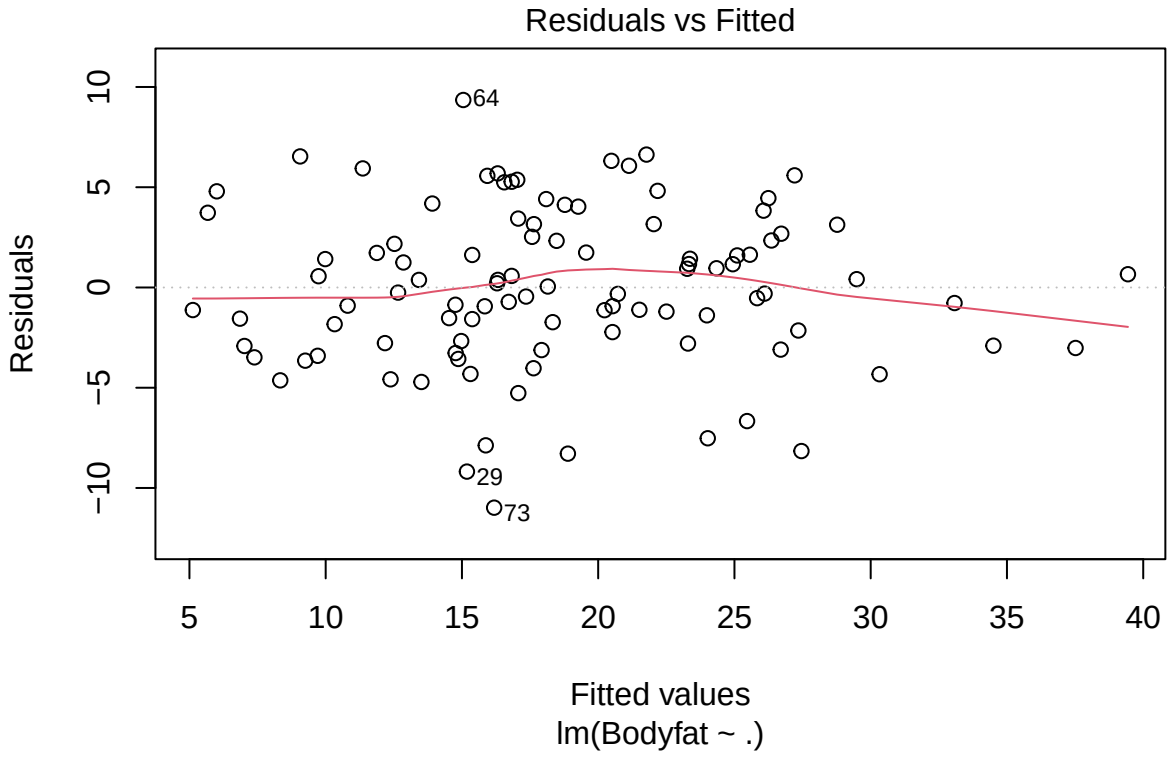
Variance Inflation Factor

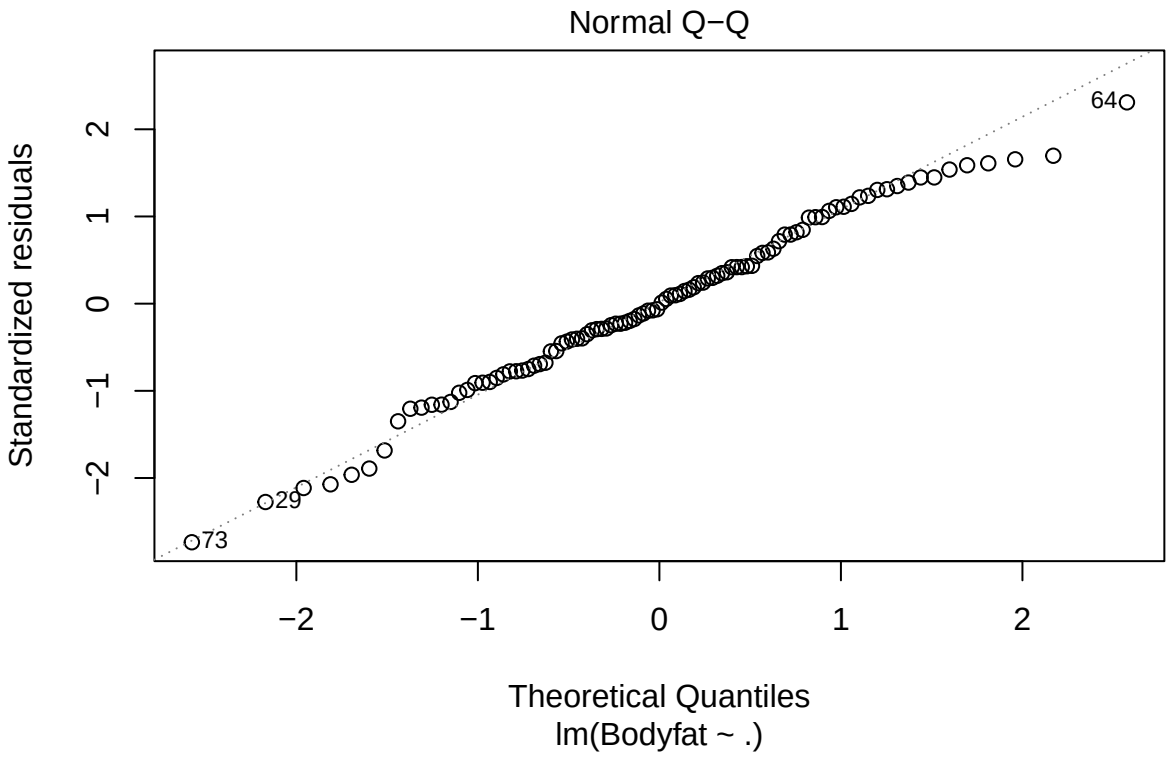
```
mod.full = lm(Bodyfat~.,data=bf)
summary(mod.full)
```

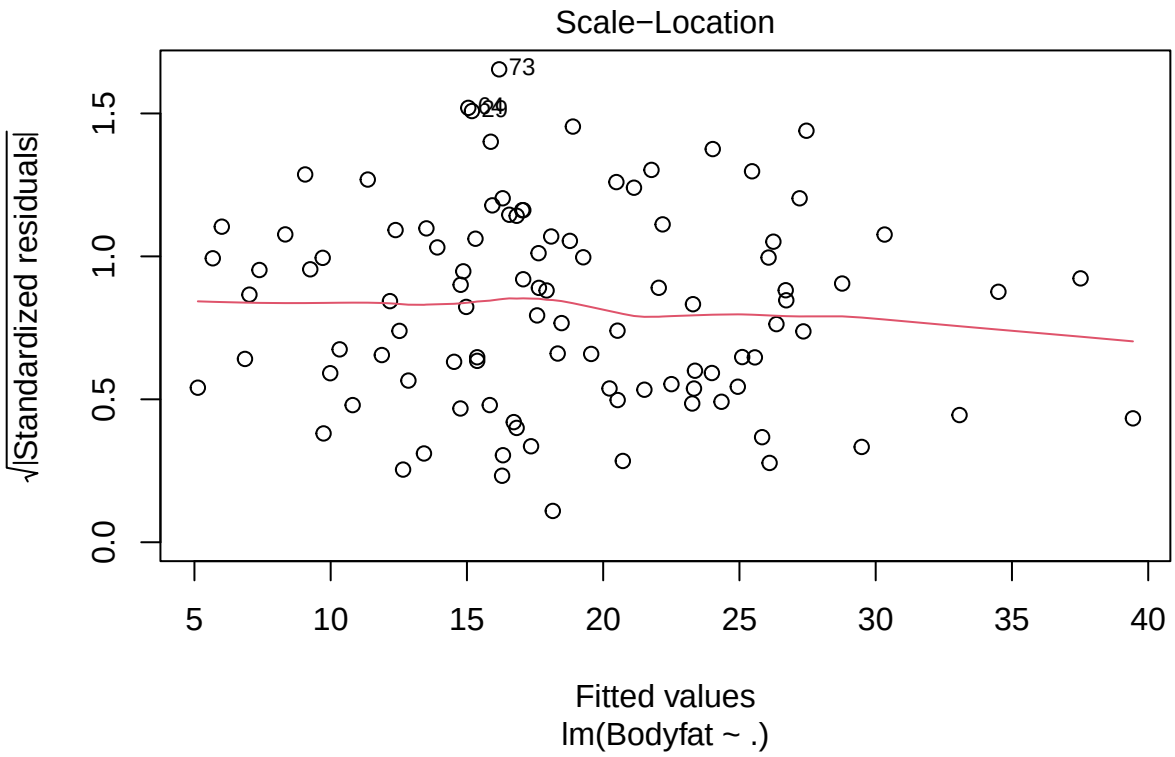
```
##
## Call:
## lm(formula = Bodyfat ~ ., data = bf)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9819  -2.7798  -0.1037   2.7939   9.3505
##
## Coefficients:
##               Estimate Std. Error t value       Pr(>|t|)
## (Intercept) -23.664200  29.459443  -0.803        0.42393
## Age           0.083779   0.050655   1.654        0.10163
## Weight       -0.083322   0.084706  -0.984        0.32792
## Height        0.035932   0.265770   0.135        0.89276
## Neck          0.001123   0.380079   0.003        0.99765
## Chest        -0.138742   0.160900  -0.862        0.39082
## Abdomen       1.032741   0.145857   7.080 0.000000000306 ***
## Ankle         0.225943   0.541725   0.417        0.67761
## Biceps        0.148276   0.229519   0.646        0.51990
## Wrist        -2.203399   0.812923  -2.710        0.00805 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

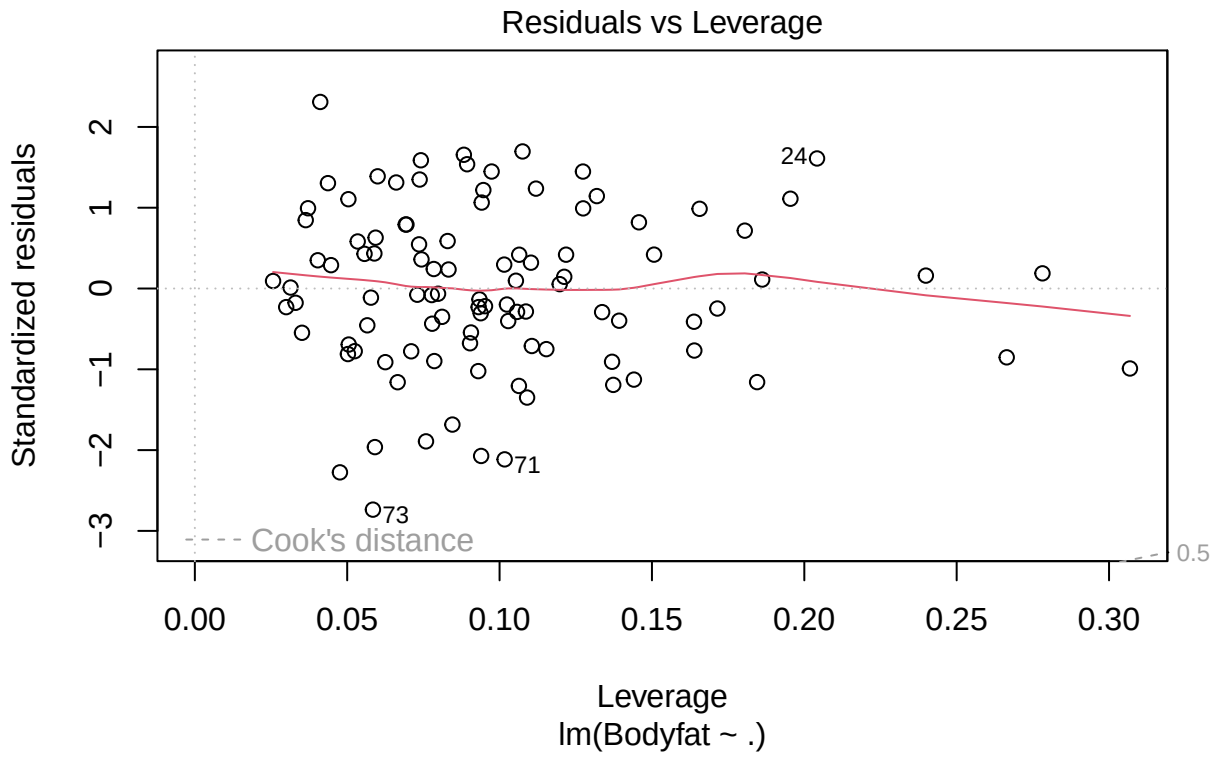
```
##
## Residual standard error: 4.136 on 90 degrees of freedom
## Multiple R-squared:  0.7575, Adjusted R-squared:  0.7332
## F-statistic: 31.23 on 9 and 90 DF,  p-value: < 0.00000000000000022
```

```
plot(mod.full)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Bodyfat ~ .)
64
29
73
−2
−1
0
1
2


Scale−Location
73
√|Standardized residuals|
0.0
0.5
1.0
1.5
5
10
15
20
25
30
35
40
Fitted values
lm(Bodyfat ~ .)




```
vif(mod.full) # From car package
```

```
##       Age    Weight    Height      Neck     Chest   Abdomen     Ankle    Biceps
##  1.936408 36.164456  3.013450  4.435176 10.869437 12.966704  3.079738  2.864922
##     Wrist
##  3.820191
```

```
mod.noWeight = lm(Bodyfat~.-Weight,data=bf)
vif(mod.noWeight)
```

```
##      Age   Height     Neck    Chest  Abdomen    Ankle   Biceps    Wrist
## 1.423017 1.535216 3.996662 8.987621 7.398178 2.634682 2.647918 3.678592
```

```
mod.noWeightChest = lm(Bodyfat~.-Weight-Chest,data=bf)
vif(mod.noWeightChest)
```

```
##      Age   Height     Neck  Abdomen    Ankle   Biceps    Wrist
## 1.422846 1.530002 3.942897 2.936500 2.564204 2.447678 3.678552
```

Fit All Subsets

The `regsubsets()` function fits all subset models up to a maximum number of variables. Notice the `nvmax` argument. The asterisk indicates which variables are included in the best model for each possible choice for k.

```
all = regsubsets(Bodyfat~., data = bf) #From leaps package
summary(all)
```

```
## Subset selection object
```

```
## Call: regsubsets.formula(Bodyfat ~ ., data = bf)
## 9 Variables  (and intercept)
##         Forced in Forced out
## Age          FALSE      FALSE
## Weight       FALSE      FALSE
## Height       FALSE      FALSE
## Neck         FALSE      FALSE
## Chest        FALSE      FALSE
## Abdomen      FALSE      FALSE
## Ankle        FALSE      FALSE
## Biceps       FALSE      FALSE
## Wrist        FALSE      FALSE
## 1 subsets of each size up to 8
## Selection Algorithm: exhaustive
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist
## 1  ( 1 ) " " " "    " "    " "  " "   "*"     " "   " "    " "
## 2  ( 1 ) " " "*"    " "    " "  " "   "*"     " "   " "    " "
## 3  ( 1 ) " " "*"    " "    " "  " "   "*"     " "   " "    "*"
## 4  ( 1 ) "*" "*"    " "    " "  " "   "*"     " "   " "    "*"
## 5  ( 1 ) "*" "*"    " "    " "  "*"   "*"     " "   " "    "*"
## 6  ( 1 ) "*" "*"    " "    " "  "*"   "*"     " "   "*"    "*"
## 7  ( 1 ) "*" "*"    " "    " "  "*"   "*"     "*"   "*"    "*"
## 8  ( 1 ) "*" "*"    "*"    " "  "*"   "*"     "*"   "*"    "*"
```

```
all2 = regsubsets(Bodyfat~., data = bf,nvmax=9) #From leaps package
summary(all2)
```

```
## Subset selection object
## Call: regsubsets.formula(Bodyfat ~ ., data = bf, nvmax = 9)
## 9 Variables  (and intercept)
##         Forced in Forced out
## Age          FALSE      FALSE
## Weight       FALSE      FALSE
## Height       FALSE      FALSE
## Neck         FALSE      FALSE
## Chest        FALSE      FALSE
## Abdomen      FALSE      FALSE
## Ankle        FALSE      FALSE
## Biceps       FALSE      FALSE
## Wrist        FALSE      FALSE
## 1 subsets of each size up to 9
## Selection Algorithm: exhaustive
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist
## 1  ( 1 ) " " " "    " "    " "  " "   "*"     " "   " "    " "
## 2  ( 1 ) " " "*"    " "    " "  " "   "*"     " "   " "    " "
## 3  ( 1 ) " " "*"    " "    " "  " "   "*"     " "   " "    "*"
## 4  ( 1 ) "*" "*"    " "    " "  " "   "*"     " "   " "    "*"
## 5  ( 1 ) "*" "*"    " "    " "  "*"   "*"     " "   " "    "*"
## 6  ( 1 ) "*" "*"    " "    " "  "*"   "*"     " "   "*"    "*"
## 7  ( 1 ) "*" "*"    " "    " "  "*"   "*"     "*"   "*"    "*"
## 8  ( 1 ) "*" "*"    "*"    " "  "*"   "*"     "*"   "*"    "*"
## 9  ( 1 ) "*" "*"    "*"    "*"  "*"   "*"     "*"   "*"    "*"
```

Now we identify the "best" model based off the criteria R-Squared, adjusted R-Squared, and Mallow's Cp. We can use the `ShowSubsets()` function created by Dr. McLean.

```
out2 = ShowSubsets(all2)
out2
```

```
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist   Rsq adjRsq
## 1  ( 1 )                                   *                     65.87  65.52
## 2  ( 1 )          *                        *                     73.28  72.73
## 3  ( 1 )          *                        *                   * 74.71  73.92
## 4  ( 1 )   *      *                        *                   * 75.40  74.37
## 5  ( 1 )   *      *                 *      *                   * 75.59  74.29
## 6  ( 1 )   *      *                 *      *             *     * 75.70  74.13
## 7  ( 1 )   *      *                 *      *     *       *     * 75.74  73.90
## 8  ( 1 )   *      *      *          *      *     *       *     * 75.75  73.62
## 9  ( 1 )   *      *      *    *     *      *     *       *     * 75.75  73.32
##             Cp
## 1  ( 1 ) 30.66
## 2  ( 1 )  5.17
## 3  ( 1 )  1.83
## 4  ( 1 )  1.27
## 5  ( 1 )  2.60
## 6  ( 1 )  4.18
## 7  ( 1 )  6.02
## 8  ( 1 )  8.00
## 9  ( 1 ) 10.00
```

```
# Best Model According to R-Squared
out2[which.max(out2$Rsq),]
```

```
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist   Rsq adjRsq
## 8  ( 1 )   *      *      *          *       *     *      *     * 75.75  73.62
##          Cp
## 8  ( 1 )  8
```

```
# Best Model According to Adjusted R-Squared
out2[which.max(out2$adjRsq),]
```

```
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist  Rsq adjRsq
## 4  ( 1 )   *      *                        *                   * 75.4  74.37
##            Cp
## 4  ( 1 ) 1.27
```

```
# Best Model According to Mallows Cp
out2[which.min(out2$Cp),]
```

```
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist  Rsq adjRsq
## 4  ( 1 )   *      *                        *                   * 75.4  74.37
##            Cp
## 4  ( 1 ) 1.27
```

We can also tell the function using `nbest` the number of top models for each choice of k that we want to see in the output. We can also calculate the BIC for each of the models and identify the best model according to BIC.

```
all3 = regsubsets(Bodyfat~., data = BodyFat, nvmax=9, nbest=2)
out3 = ShowSubsets(all3)
out3
```

```
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist   Rsq adjRsq
## 1  ( 1 )                                   *                     65.87  65.52
```

```
## 1  ( 2 )                          *                          46.55  46.00
## 2  ( 1 )          *                          *                73.28  72.73
## 2  ( 2 )                                     *               * 71.29  70.69
## 3  ( 1 )          *                          *               * 74.71  73.92
## 3  ( 2 )   *                                 *               * 74.66  73.86
## 4  ( 1 )   *      *                          *               * 75.40  74.37
## 4  ( 2 )   *                          *      *               * 75.28  74.24
## 5  ( 1 )   *      *                   *      *               * 75.59  74.29
## 5  ( 2 )   *      *                          *          *    * 75.47  74.17
## 6  ( 1 )   *      *                   *      *          *    * 75.70  74.13
## 6  ( 2 )   *      *                   *      *     *         * 75.63  74.06
## 7  ( 1 )   *      *                   *      *     *     *   * 75.74  73.90
## 7  ( 2 )   *      *      *            *      *           *   * 75.70  73.85
## 8  ( 1 )   *      *      *            *      *     *     *   * 75.75  73.62
## 8  ( 2 )   *      *           *      *      *     *     *   * 75.74  73.61
## 9  ( 1 )   *      *      *    *      *      *     *     *   * 75.75  73.32
##              Cp
## 1  ( 1 )  30.66
## 1  ( 2 ) 102.35
## 2  ( 1 )   5.17
## 2  ( 2 )  12.56
## 3  ( 1 )   1.83
## 3  ( 2 )   2.05
## 4  ( 1 )   1.27
## 4  ( 2 )   1.72
## 5  ( 1 )   2.60
## 5  ( 2 )   3.03
## 6  ( 1 )   4.18
## 6  ( 2 )   4.42
## 7  ( 1 )   6.02
## 7  ( 2 )   6.18
## 8  ( 1 )   8.00
## 8  ( 2 )   8.02
## 9  ( 1 )  10.00
```

```
#Get BIC for each of the models
summ.all3 = summary(all3)
summ.all3$bic
```

```
##  [1]  -98.28408  -53.42961 -118.14668 -110.96042 -119.07503 -118.83843
##  [7] -117.23164 -116.74943 -113.37315 -112.89961 -109.22590 -108.95840
## [13] -104.80096 -104.62649 -100.21719 -100.19689  -95.61203
```

```
#Find Best Model According to BIC
out3[which.min(summ.all3$bic),]
```

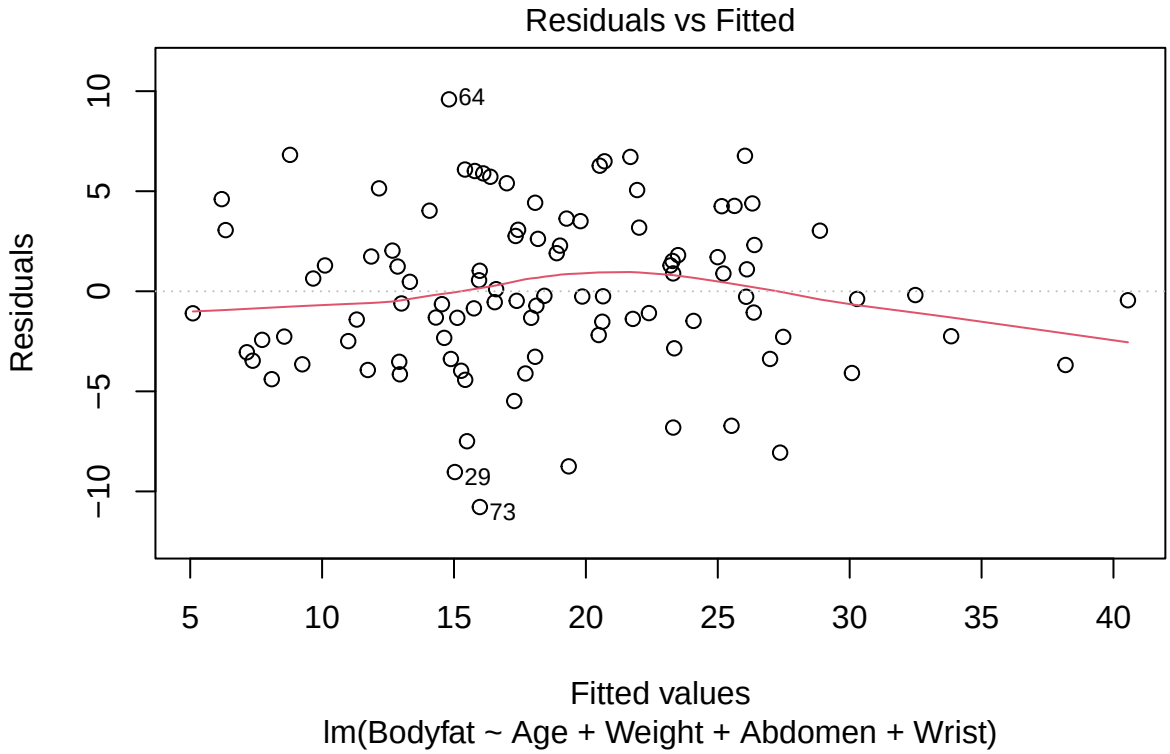
```
##          Age Weight Height Neck Chest Abdomen Ankle Biceps Wrist   Rsq adjRsq
## 3  ( 1 )          *                          *                 * 74.71  73.92
##            Cp
## 3  ( 1 ) 1.83
```

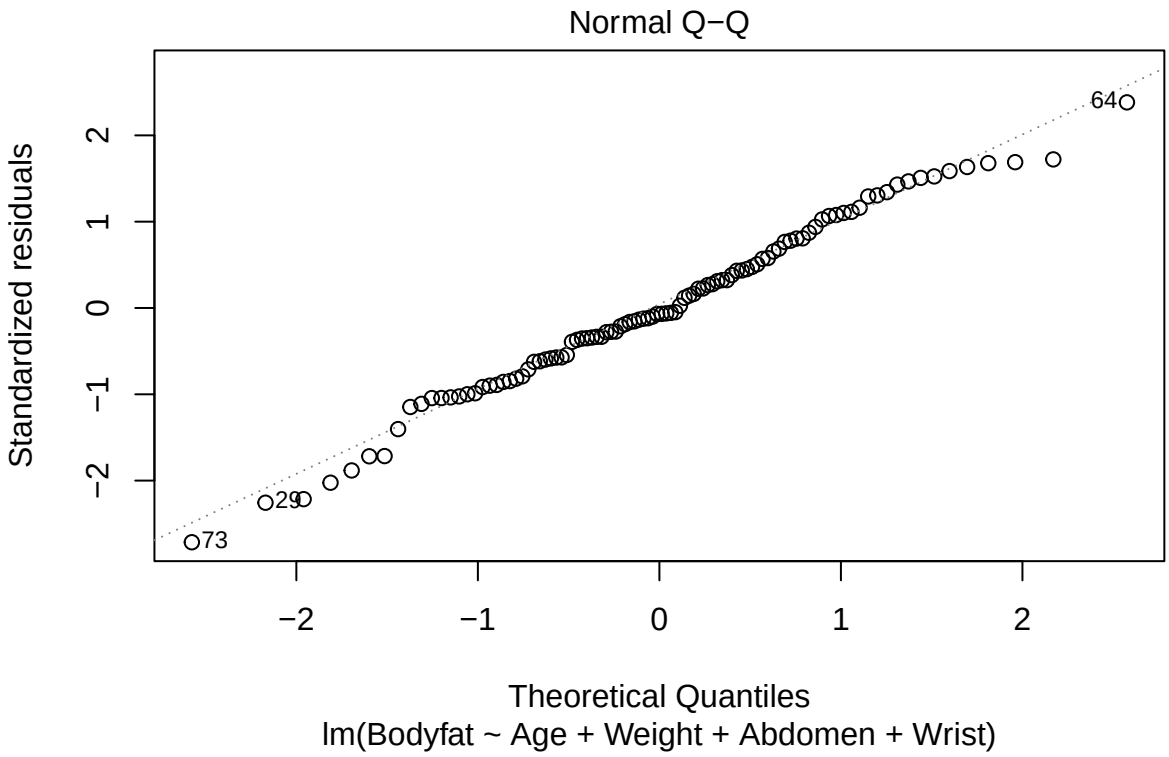
```
#Calculated adjusted R-squared by hand
1-(1-0.7471)*((100-1)/(100-3-1)) #NOtice these equals the adjusted R-Squared in the BIC Model
```

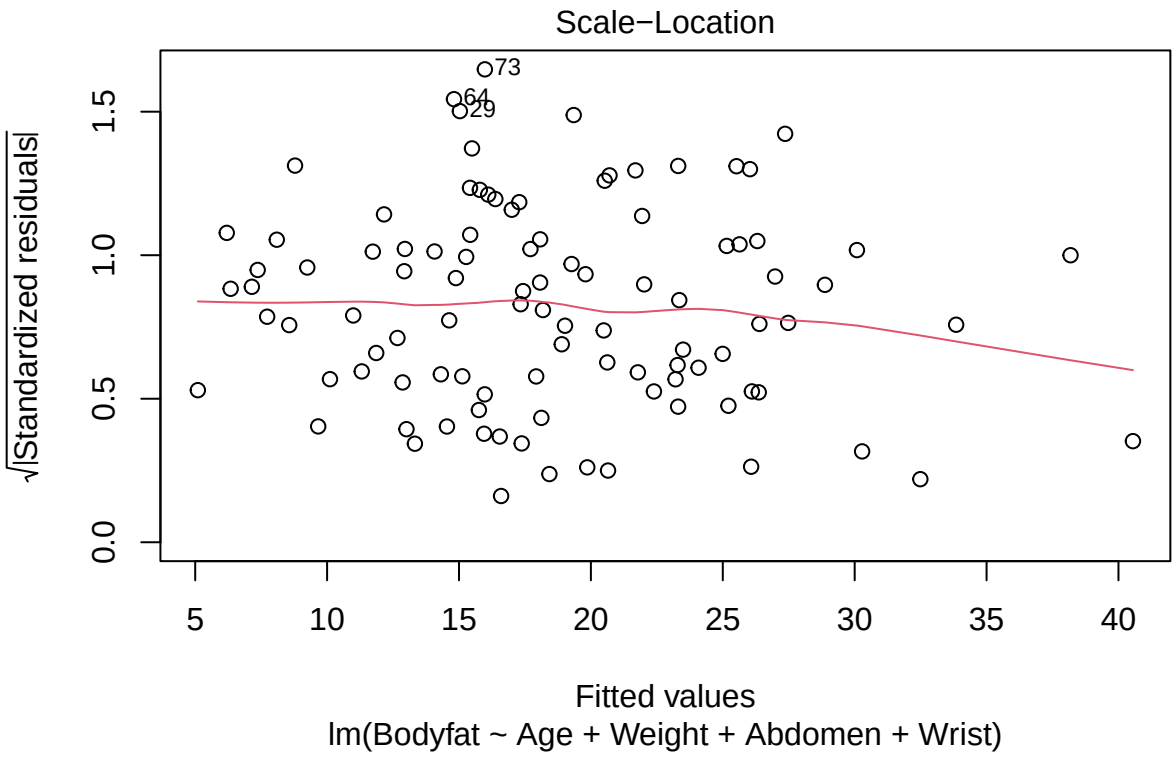
```
## [1] 0.7391969
```

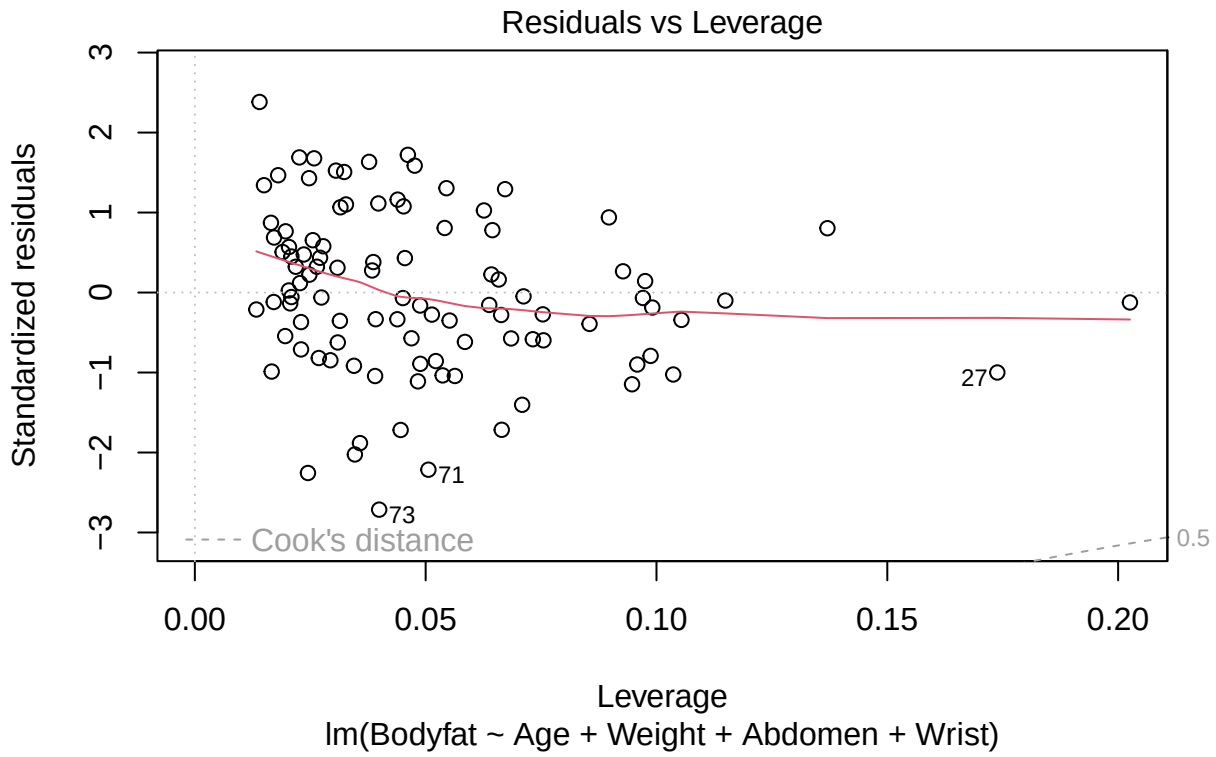
Now we fit our "best" models according to adjusted R-squared/Mallow's Cp and BIC.

```
mod.rsqmallow = lm(Bodyfat ~ Age + Weight + Abdomen + Wrist, data=bf)
plot(mod.rsqmallow)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Bodyfat ~ Age + Weight + Abdomen + Wrist)
64
29
73
−2
−1
0
1
2

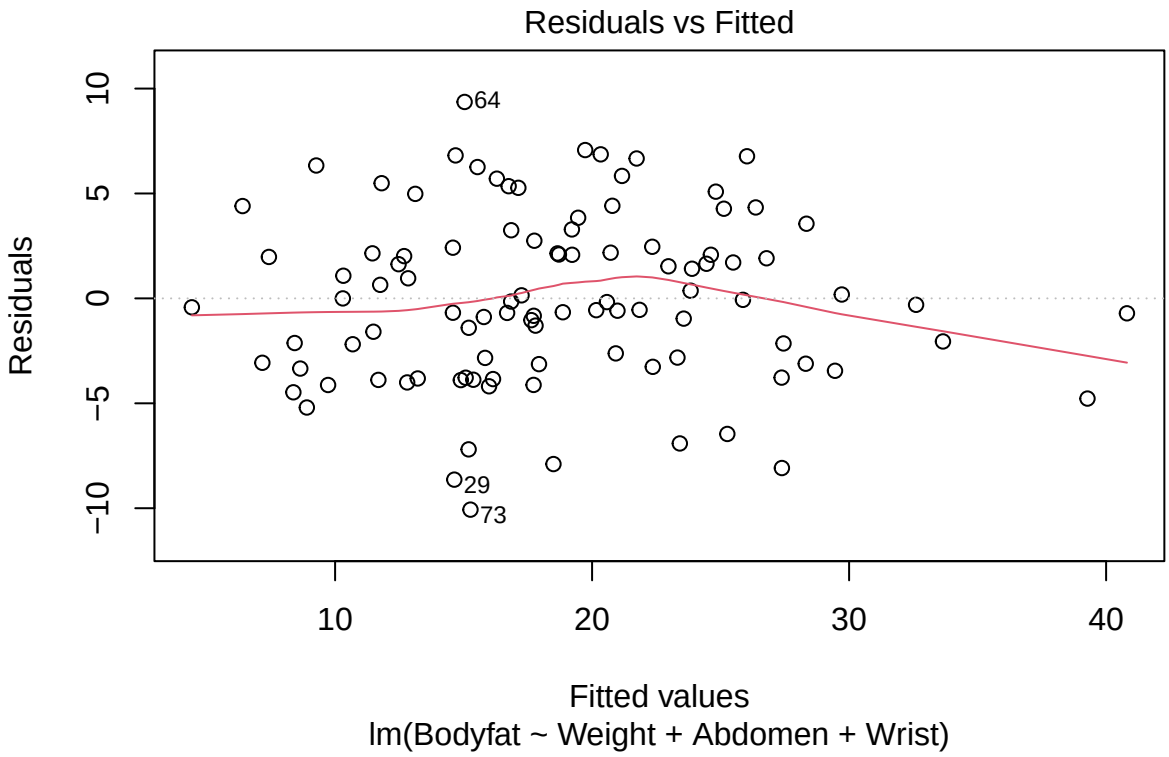

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
73
64
29
5
10
15
20
25
30
35
40
Fitted values
lm(Bodyfat ~ Age + Weight + Abdomen + Wrist)

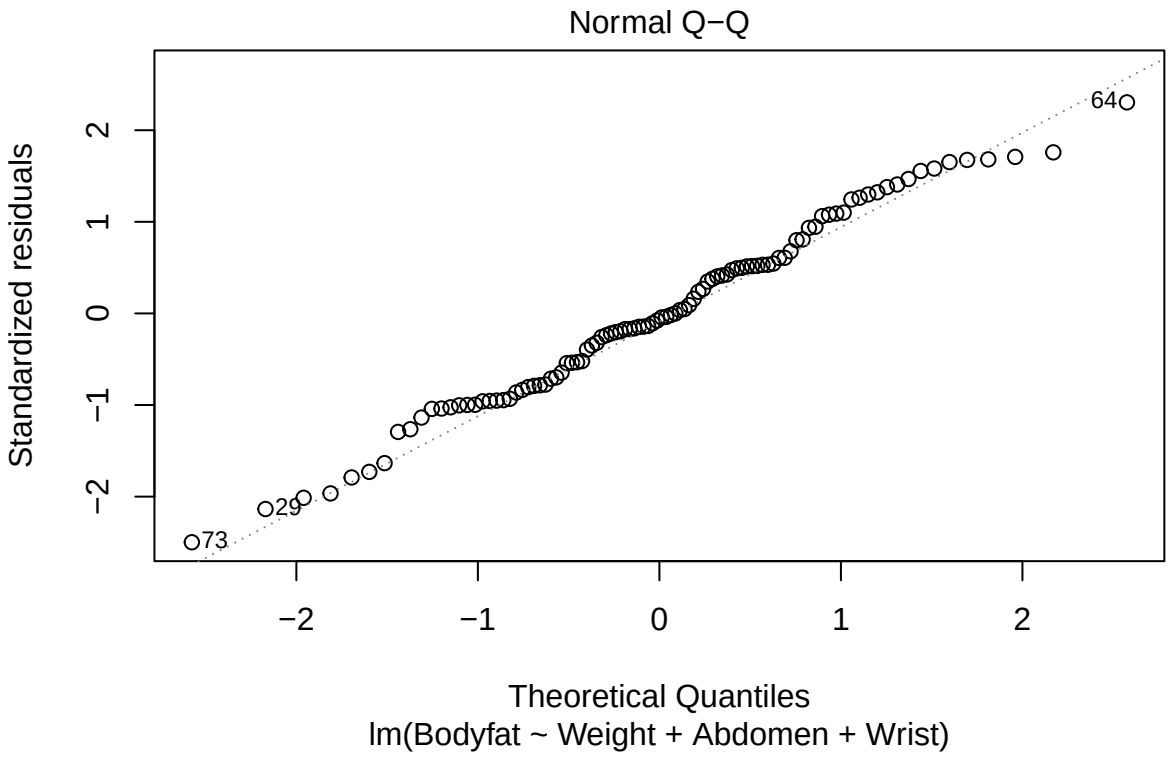


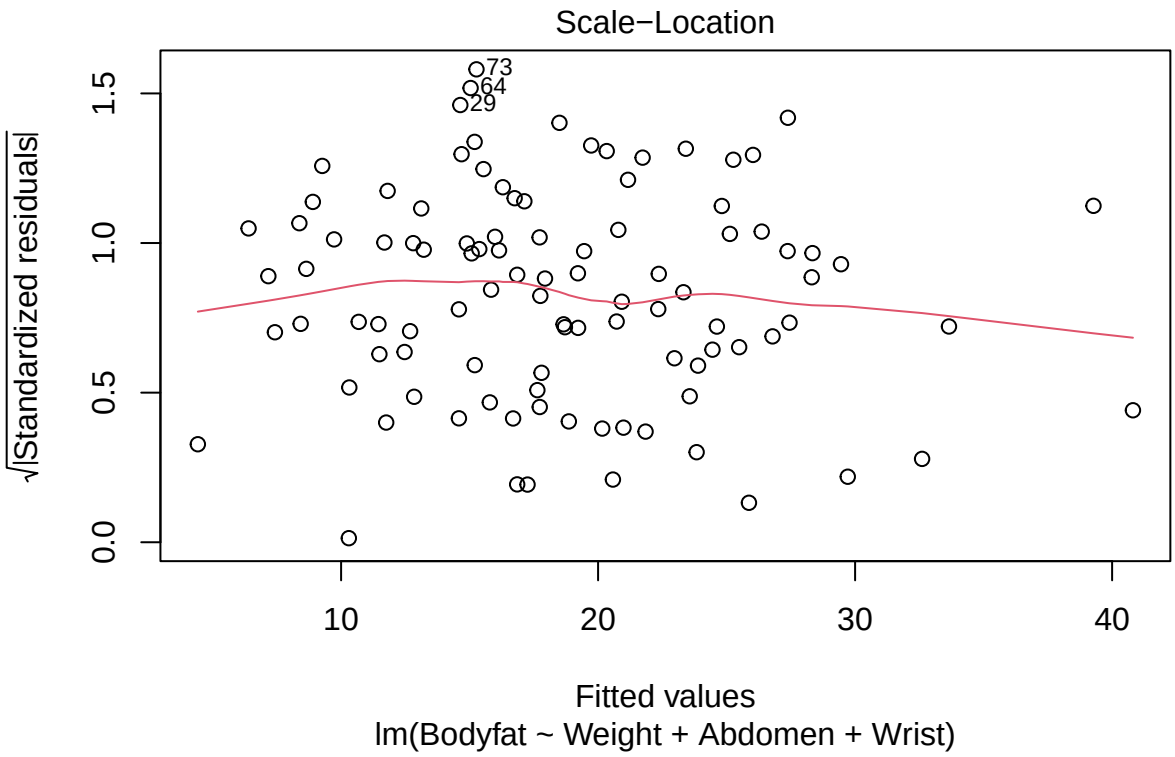

```
vif(mod.rsqmallow)
```

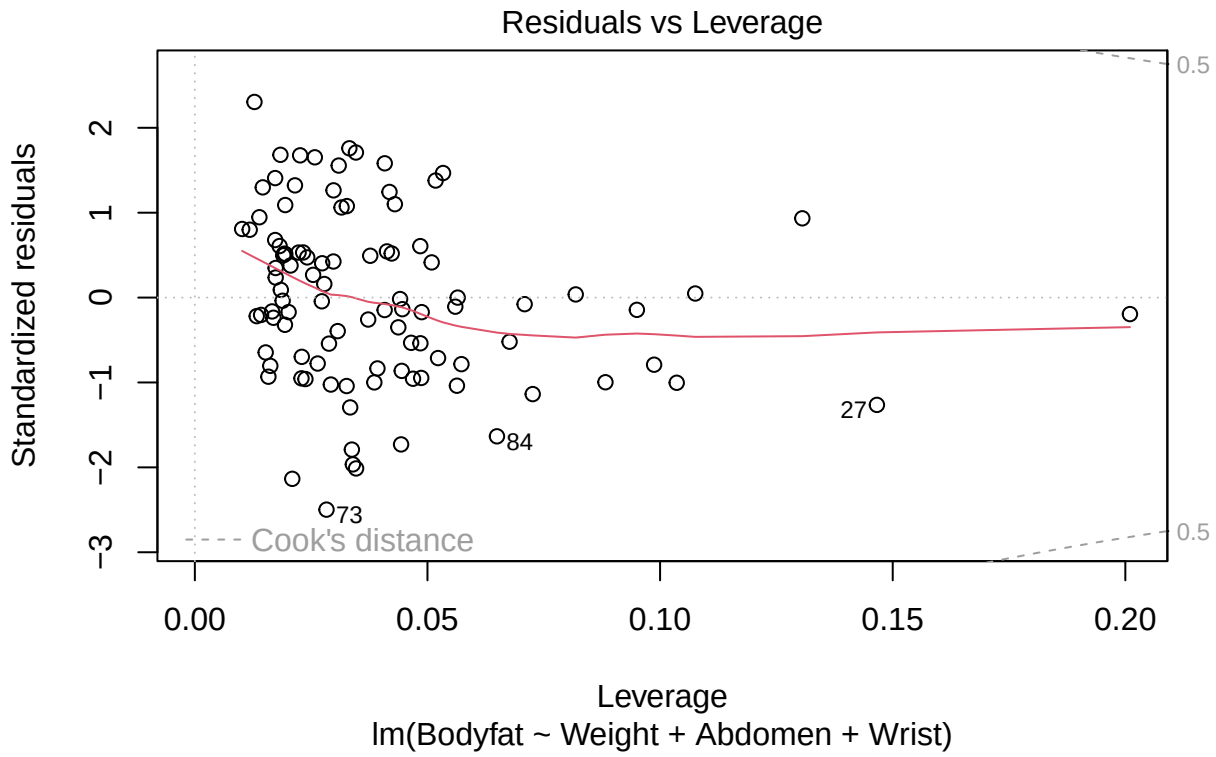
```
##       Age    Weight   Abdomen     Wrist
##  1.821182 10.501749  6.860258  3.149493
```

```
mod.bic = lm(Bodyfat ~ Weight + Abdomen + Wrist,data=bf)
plot(mod.bic)
```


Residuals vs Fitted
Residuals
10
5
0
−5
−10
64
29
73
10
20
30
40
Fitted values
lm(Bodyfat ~ Weight + Abdomen + Wrist)


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Bodyfat ~ Weight + Abdomen + Wrist)
64
29
73
−2
−1
0
1
2


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
73
64
29
10
20
30
40
Fitted values
lm(Bodyfat ~ Weight + Abdomen + Wrist)




```
vif(mod.bic)
```

```
##   Weight  Abdomen    Wrist
## 6.049705 4.741816 2.326765
```

Backwards, Forwards, and Stepwise Algorithms

Built-in `step()` function doesn't use p-values to determine what variables to remove or keep. It uses the AIC measurement which is similar to BIC. Let's first look at backward's elimination.

```
Full = lm(Bodyfat~.,data=bf)
back.out = step(Full, direction="backward")
```

```
## Start:  AIC=293.39
## Bodyfat ~ Age + Weight + Height + Neck + Chest + Abdomen + Ankle +
##     Biceps + Wrist
##
##           Df Sum of Sq    RSS    AIC
## - Neck     1      0.00 1539.2 291.39
## - Height   1      0.31 1539.5 291.41
## - Ankle    1      2.98 1542.2 291.58
## - Biceps   1      7.14 1546.4 291.85
## - Chest    1     12.72 1551.9 292.21
## - Weight   1     16.55 1555.8 292.46
## <none>                 1539.2 293.39
## - Age      1     46.78 1586.0 294.38
## - Wrist    1    125.65 1664.9 299.23
```

```
## - Abdomen  1    857.41 2396.6 335.67
##
## Step:  AIC=291.39
## Bodyfat ~ Age + Weight + Height + Chest + Abdomen + Ankle + Biceps +
##     Wrist
##
##           Df Sum of Sq    RSS    AIC
## - Height   1      0.33 1539.6 289.41
## - Ankle    1      3.02 1542.2 289.58
## - Biceps   1      7.19 1546.4 289.85
## - Chest    1     12.73 1552.0 290.21
## - Weight   1     18.33 1557.6 290.57
## <none>                 1539.2 291.39
## - Age      1     46.82 1586.0 292.38
## - Wrist    1    144.48 1683.7 298.36
## - Abdomen  1    862.49 2401.7 333.88
##
## Step:  AIC=289.41
## Bodyfat ~ Age + Weight + Chest + Abdomen + Ankle + Biceps + Wrist
##
##           Df Sum of Sq    RSS    AIC
## - Ankle    1      2.78 1542.3 287.59
## - Biceps   1      6.91 1546.5 287.86
## - Chest    1     14.74 1554.3 288.36
## - Weight   1     27.96 1567.5 289.21
## <none>                 1539.6 289.41
## - Age      1     49.53 1589.1 290.57
## - Wrist    1    145.64 1685.2 296.45
## - Abdomen  1   1153.22 2692.8 343.32
##
## Step:  AIC=287.59
## Bodyfat ~ Age + Weight + Chest + Abdomen + Biceps + Wrist
##
##           Df Sum of Sq    RSS    AIC
## - Biceps   1      7.08 1549.4 286.05
## - Chest    1     14.43 1556.8 286.52
## - Weight   1     25.22 1567.5 287.21
## <none>                 1542.3 287.59
## - Age      1     49.72 1592.1 288.76
## - Wrist    1    143.37 1685.7 294.48
## - Abdomen  1   1168.38 2710.7 341.98
##
## Step:  AIC=286.05
## Bodyfat ~ Age + Weight + Chest + Abdomen + Wrist
##
##           Df Sum of Sq    RSS    AIC
## - Chest    1     11.61 1561.0 284.79
## - Weight   1     19.16 1568.6 285.27
## <none>                 1549.4 286.05
## - Age      1     50.55 1600.0 287.26
## - Wrist    1    139.74 1689.2 292.68
## - Abdomen  1   1186.64 2736.1 340.91
##
## Step:  AIC=284.79
```

```
## Bodyfat ~ Age + Weight + Abdomen + Wrist
##
##           Df Sum of Sq    RSS    AIC
## <none>                  1561.0 284.79
## - Age      1     43.71 1604.7 285.55
## - Weight   1     47.51 1608.5 285.79
## - Wrist    1    134.38 1695.4 291.05
## - Abdomen  1   1373.33 2934.3 345.91
```

```
summary(back.out)
```

```
##
## Call:
## lm(formula = Bodyfat ~ Age + Weight + Abdomen + Wrist, data = bf)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.7803  -2.4434  -0.2681   2.8294   9.5902
##
## Coefficients:
##              Estimate Std. Error t value           Pr(>|t|)
## (Intercept) -21.06107   10.52814  -2.000            0.04831 *
## Age           0.07854    0.04815   1.631            0.10620
## Weight       -0.07608    0.04474  -1.700            0.09231 .
## Abdomen       0.95069    0.10399   9.142 0.0000000000000113 ***
## Wrist        -2.06898    0.72350  -2.860            0.00521 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.054 on 95 degrees of freedom
## Multiple R-squared:  0.754,  Adjusted R-squared:  0.7437
## F-statistic: 72.81 on 4 and 95 DF,  p-value: < 0.00000000000000022
```

```
vif(back.out)
```

```
##       Age    Weight   Abdomen     Wrist
##  1.821182 10.501749  6.860258  3.149493
```

Now let's look at forward selection. In this case, we need to start by initiating the empty model and telling the `step()` function to consider all models up to possibly the full model.

```
Empty = lm(Bodyfat~1,data=bf)
forward.out = step(Empty,scope=list(upper=Full),direction="forward",trace=F)
summary(forward.out)
```

```
##
## Call:
## lm(formula = Bodyfat ~ Abdomen + Weight + Wrist + Age, data = bf)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.7803  -2.4434  -0.2681   2.8294   9.5902
##
## Coefficients:
##              Estimate Std. Error t value           Pr(>|t|)
## (Intercept) -21.06107   10.52814  -2.000            0.04831 *
## Abdomen       0.95069    0.10399   9.142 0.0000000000000113 ***
```

```
## Weight       -0.07608    0.04474  -1.700            0.09231 .
## Wrist        -2.06898    0.72350  -2.860            0.00521 **
## Age           0.07854    0.04815   1.631            0.10620
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.054 on 95 degrees of freedom
## Multiple R-squared:  0.754,  Adjusted R-squared:  0.7437
## F-statistic: 72.81 on 4 and 95 DF,  p-value: < 0.00000000000000022
```

```
vif(forward.out)
```

```
##   Abdomen    Weight     Wrist       Age
##  6.860258 10.501749  3.149493  1.821182
```

Now, we specify the `direction="both"` to conduct stepwise regression where variables can be both added and removed.

```
Empty = lm(Bodyfat~1,data=bf)
step.out = step(Empty,scope=list(upper=Full),direction="both",trace=F)
summary(step.out)
```

```
##
## Call:
## lm(formula = Bodyfat ~ Abdomen + Weight + Wrist + Age, data = bf)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.7803  -2.4434  -0.2681   2.8294   9.5902
##
## Coefficients:
##              Estimate Std. Error t value           Pr(>|t|)
## (Intercept) -21.06107   10.52814  -2.000            0.04831 *
## Abdomen       0.95069    0.10399   9.142 0.0000000000000113 ***
## Weight       -0.07608    0.04474  -1.700            0.09231 .
## Wrist        -2.06898    0.72350  -2.860            0.00521 **
## Age           0.07854    0.04815   1.631            0.10620
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.054 on 95 degrees of freedom
## Multiple R-squared:  0.754,  Adjusted R-squared:  0.7437
## F-statistic: 72.81 on 4 and 95 DF,  p-value: < 0.00000000000000022
```

```
vif(step.out)
```

```
##   Abdomen    Weight     Wrist       Age
##  6.860258 10.501749  3.149493  1.821182
```

All Models Summarized

```
mod.full
```

```
##
## Call:
## lm(formula = Bodyfat ~ ., data = bf)
```

```
##
## Coefficients:
## (Intercept)          Age       Weight       Height         Neck        Chest
##  -23.664200     0.083779    -0.083322     0.035932     0.001123    -0.138742
##     Abdomen        Ankle       Biceps        Wrist
##    1.032741     0.225943     0.148276    -2.203399
```

```
mod.noWeight
```

```
##
## Call:
## lm(formula = Bodyfat ~ . - Weight, data = bf)
##
## Coefficients:
## (Intercept)          Age       Height         Neck        Chest      Abdomen
##     2.45399      0.10944     -0.14717     -0.11644     -0.20460      0.93872
##       Ankle       Biceps        Wrist
##     0.02337      0.08614     -2.35735
```

```
mod.noWeightChest
```

```
##
## Call:
## lm(formula = Bodyfat ~ . - Weight - Chest, data = bf)
##
## Coefficients:
## (Intercept)          Age       Height         Neck      Abdomen        Ankle
##    1.419294     0.110102    -0.162629    -0.174955     0.819076    -0.091225
##      Biceps        Wrist
##    0.001288    -2.353703
```

```
mod.rsqmallow
```

```
##
## Call:
## lm(formula = Bodyfat ~ Age + Weight + Abdomen + Wrist, data = bf)
##
## Coefficients:
## (Intercept)          Age       Weight      Abdomen        Wrist
##   -21.06107      0.07854     -0.07608      0.95069     -2.06898
```

```
mod.bic
```

```
##
## Call:
## lm(formula = Bodyfat ~ Weight + Abdomen + Wrist, data = bf)
##
## Coefficients:
## (Intercept)       Weight      Abdomen        Wrist
##    -28.7531      -0.1236       1.0449      -1.4659
```

```
back.out
```

```
##
## Call:
## lm(formula = Bodyfat ~ Age + Weight + Abdomen + Wrist, data = bf)
##
## Coefficients:
```

```
## (Intercept)          Age       Weight      Abdomen        Wrist
##   -21.06107      0.07854     -0.07608      0.95069     -2.06898
```

```
forward.out
```

```
##
## Call:
## lm(formula = Bodyfat ~ Abdomen + Weight + Wrist + Age, data = bf)
##
## Coefficients:
## (Intercept)      Abdomen       Weight        Wrist          Age
##   -21.06107      0.95069     -0.07608     -2.06898      0.07854
```

```
step.out
```

```
##
## Call:
## lm(formula = Bodyfat ~ Abdomen + Weight + Wrist + Age, data = bf)
##
## Coefficients:
## (Intercept)      Abdomen       Weight        Wrist          Age
##   -21.06107      0.95069     -0.07608     -2.06898      0.07854
```