

Report I - SF2930 Regression Analysis

Chen Hanmo, 990904-T072 Shen Linying, 981208-T066

9 March, 2019 (Revised on 15 March,2019)

Contents

1	Introduction	2
1.1	Background	2
1.2	Goals	2
2	Model development	2
2.1	The full model	2
2.2	Residual Analysis	3
2.2.1	Standardized Residuals	3
2.2.2	Normal Probability Plot	3
2.2.3	Residuals against fitted values and regressors	4
2.2.4	Partial Residual Plots	5
2.3	Diagnostics of Leverage and Influence	5
2.3.1	Leverage Plots	5
2.3.2	Residuals V.S. Leverage	5
2.3.3	Cook's distance	5
2.4	Variable Transformation	7
2.5	Multicollinearity diagnostics and treatments	9
2.5.1	Variance Inflation factors	9
2.5.2	Eigensystem Analysis	9
2.5.3	Ridge Regression	9
2.5.4	Principal Component Regression	11
2.6	Variable Selection	12
2.6.1	All possible regression	12
2.6.2	Stepwise Regression	12
2.6.3	Validation of variable selection	12
2.7	Bootstrapping	13
3	Results	13
3.1	Our Models	13
3.2	Assessment	13
4	Conclusion	13

Project 1

Scenario I: Body fat assessment

1 Introduction

1.1 Background

The World Health organization (WHO) reported that obesity is a major risk factor for a number of chronic diseases, including diabetes, cardiovascular diseases and cancer. Obesity is defined as "the disease in which excess of body fat has accumulated to such extend that health may be adversely affected". Once being considered as a problem only for high income countries, obesity is now rise in low- and middle-income countries. An important issue for medical purposes is thus is to reliably identify people with the fat excess.

1.2 Goals

As a major risk for many chronic diseases, obesity can be influenced by many factors including one's age, height and other body indexes.

In the dataset from <http://lib.stat.cmu.edu/datasets/bodyfat>, it uses *body fat mass* (BFM) instead of the *body mass index*(BMI) as the measure of body fatness and other 13 indexs are included as factors that may influence the fatness.

In this project, we are trying to develop a regression model based on the body dataset and use statistical techniques to modify and validate it.

2 Model development

We establish and modify our model following the flow chart[1] presented in Montgomery's book. Although our model-building process covers many other aspects, we will mainly focus on the **residual analysis** and **multicollinearity diagnostics** in the following illustration.

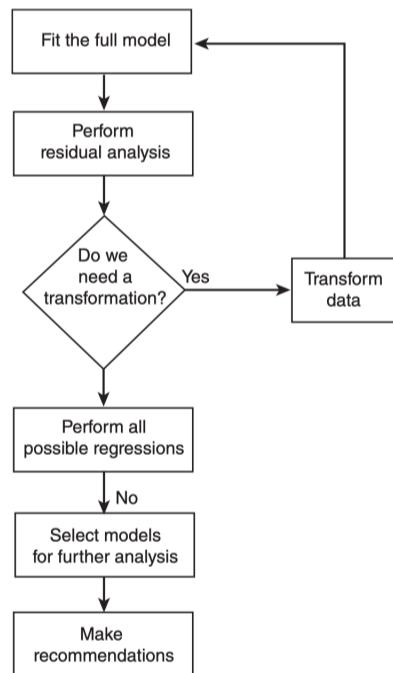


Figure 1: the model-building process

2.1 The full model

Firstly, we try to establish the full model with all the columns in the data frame. And the summary of the model is presented as [2].

```

Call:
lm(formula = density ~ age + weight + height + neck + chest +
    abdomen + hip + thigh + knee + ankle + biceps + forearm +
    wrist, data = mendata)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.0225107 -0.0071735  0.0002816  0.0064878  0.0254670 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.156e+00  5.061e-02 22.846 < 2e-16 ***
age        -1.320e-04  7.392e-05 -1.785  0.07550 .  
weight      2.378e-04  1.408e-04  1.689  0.09254 .  
height     -2.594e-05  4.083e-04 -0.064  0.94939  
neck        1.072e-03  5.371e-04  1.995  0.04720 *  
chest       1.169e-05  2.360e-04  0.050  0.96056  
abdomen    -2.200e-03  2.072e-04 -10.618 < 2e-16 *** 
hip         5.268e-04  3.336e-04  1.579  0.11569  
thigh      -6.343e-04  3.336e-04 -1.901  0.05849 .  
knee        -3.418e-05  5.640e-04 -0.061  0.95172  
ankle      -4.449e-04  5.107e-04 -0.871  0.38459  
biceps     -4.274e-04  3.942e-04 -1.084  0.27940  
forearm    -1.040e-03  4.527e-04 -2.298  0.02245 *  
wrist       3.651e-03  1.227e-03  2.976  0.00322 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.009781 on 234 degrees of freedom
Multiple R-squared:  0.7451,   Adjusted R-squared:  0.731 
F-statistic: 52.63 on 13 and 234 DF,  p-value: < 2.2e-16

```

Figure 2: Summary of the full model

The original full model is

$$\begin{aligned}
\text{density} = & 1.156121 - 1.319699 \times 10^{-4} * \text{age} + 2.377541 \times 10^{-4} * \text{weight} \\
& - 2.594465 \times 10^{-5} * \text{height} + 1.071540 \times 10^{-3} * \text{neck} + 1.168530 \times 10^{-5} * \text{chest} \\
& - 2.199601 \times 10^{-3} * \text{abdomen} + 5.267858 \times 10^{-4} * \text{hip} - 6.342697 \times 10^{-4} * \text{thigh} \\
& - 3.418355 \times 10^{-5} * \text{knee} - 4.448680 \times 10^{-4} * \text{ankle} - 4.273484 \times 10^{-4} * \text{biceps} \\
& - 1.040264 \times 10^{-3} * \text{forearm} + 3.651081 \times 10^{-3} * \text{wrist}
\end{aligned} \tag{1}$$

Based on that result, it can be seen that the full model is not so satisfying. For example, the **R-square** and the **p-value** of many coefficients is not appropriate for a validated model and further modification is needed.

2.2 Residual Analysis

Residual analysis is essential in our model-building to check model adequacy.

2.2.1 Standardized Residuals

First, we re-scale the residuals to the standardized values

$$d_i = \frac{e_i}{\sqrt{MS_{res}}}$$

and following [3] is the histogram of the standardized residuals.

According to the histogram, the standardized residuals all fall into the interval of $[-3, 3]$, but do not perfectly correspond with the normal distribution, which needs further inspection(as we did in the next section).

2.2.2 Normal Probability Plot

To check the normality assumption we constucted the normal probability plot of the residuals as follow [4]

As is seen in the graph, there is small departures from the normality. And the heavy-tailed error distributions indicates that there may be outliers that “pull” the least-squares fit too much in their direction. We may consider robust regression methods in the future processing.

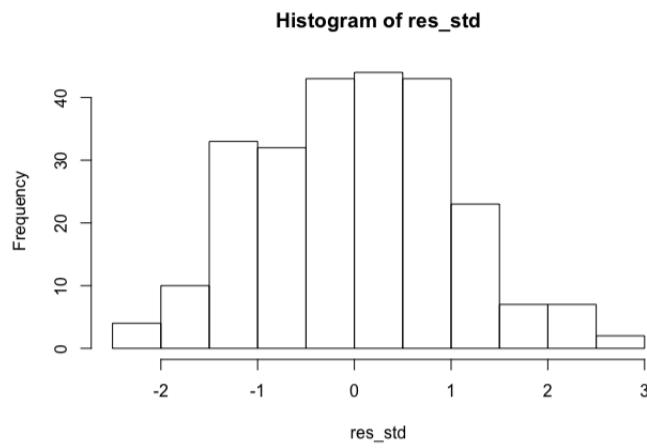


Figure 3: Histogram of the Standardized Residuals

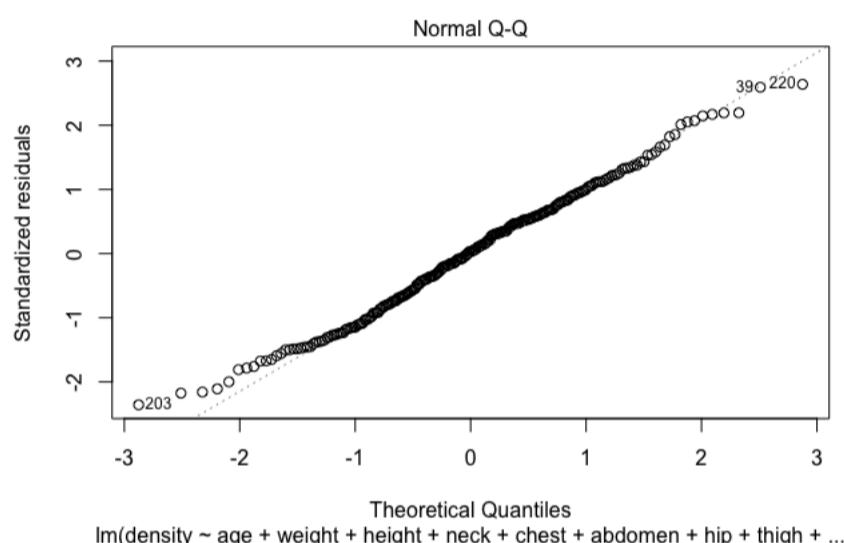


Figure 4: Normal Q-Q Plot

2.2.3 Residuals against fitted values and regressors

Another way of checking model adequacy is to plot the residuals (here we use *externally studentized residuals*) against the fitted values of the model and results are as follows[5].

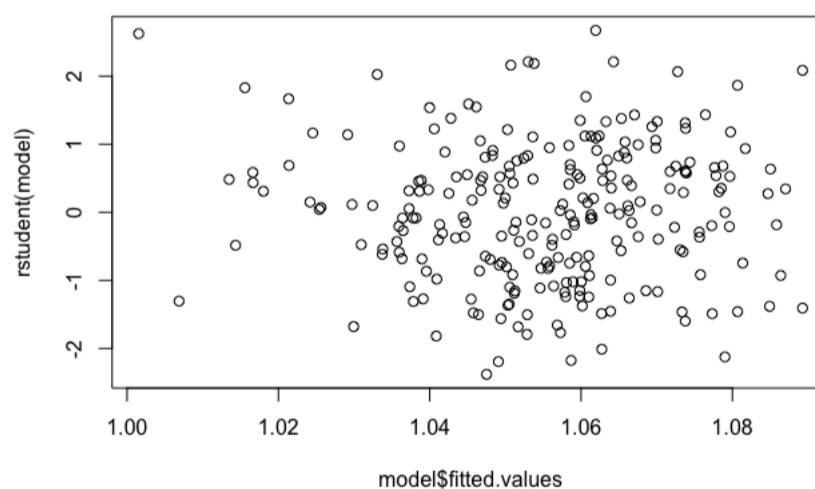


Figure 5: Residuals against fitted values

The residuals can be contained in a horizontal band, which means there are no obvious model defects.

Similarly, we can construct the 13 graphs of residuals against 13 regressors separately and the result [6][7] is also similar with the previous one.

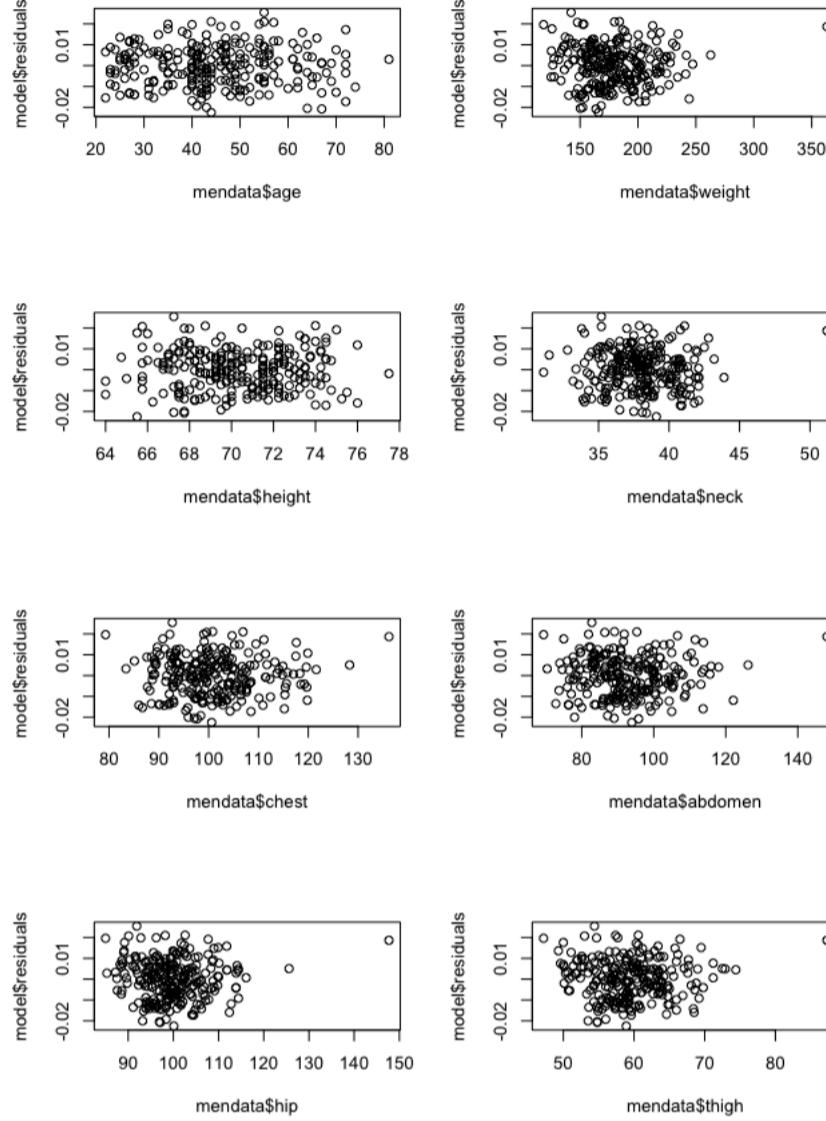


Figure 6: Residuals against regressors

2.2.4 Partial Residual Plots

Additionally, we can construct Added-Variable Plots to see the marginal effect on the target of each regressor.

As is shown in the figure [8], some regressors such as *weight, neck, hip, forearm, abdomen, wrist* have obvious effect in this model while others don't such as *age and knee*.

2.3 Diagnostics of Leverage and Influence

2.3.1 Leverage Plots

Using `leveragePlots` in the `car` package, the leverage plots of each regressor are as follows [9].

2.3.2 Residuals V.S. Leverage

We can also construct the plot of Residuals V.S. Leverage [10].

2.3.3 Cook's distance

What's more, Cook's distance is an important indicator of leverage and influence.

As is shown in graph [11,12], because the maximum Cook's distance is just around 0.4(a suggested cut-off is 1 because $F_{0.5,p,n-p}$ is approximately equal to 1 when n is large) so there is no need to rule out any point.

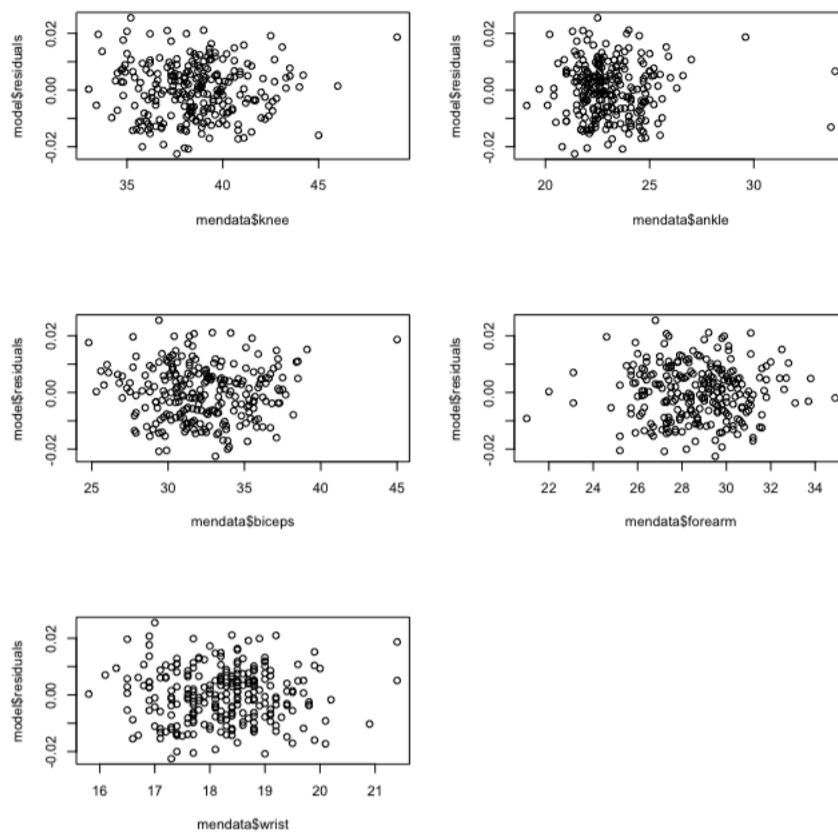


Figure 7: Residuals against regressors

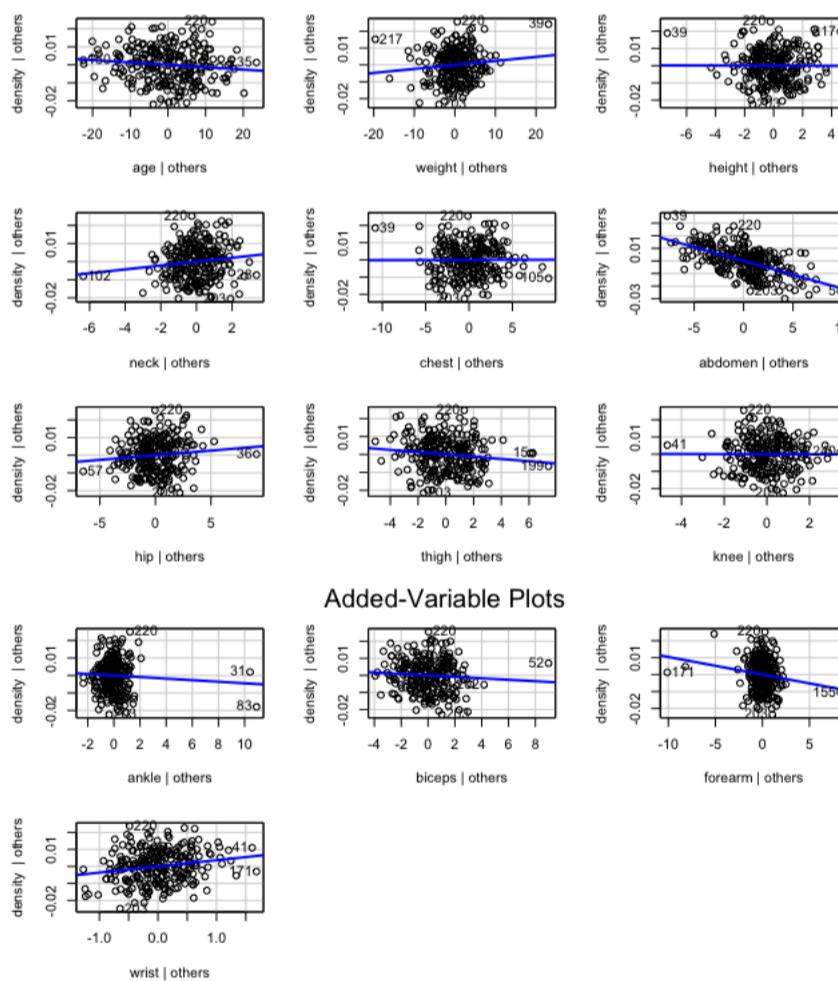


Figure 8: Residuals against regressors

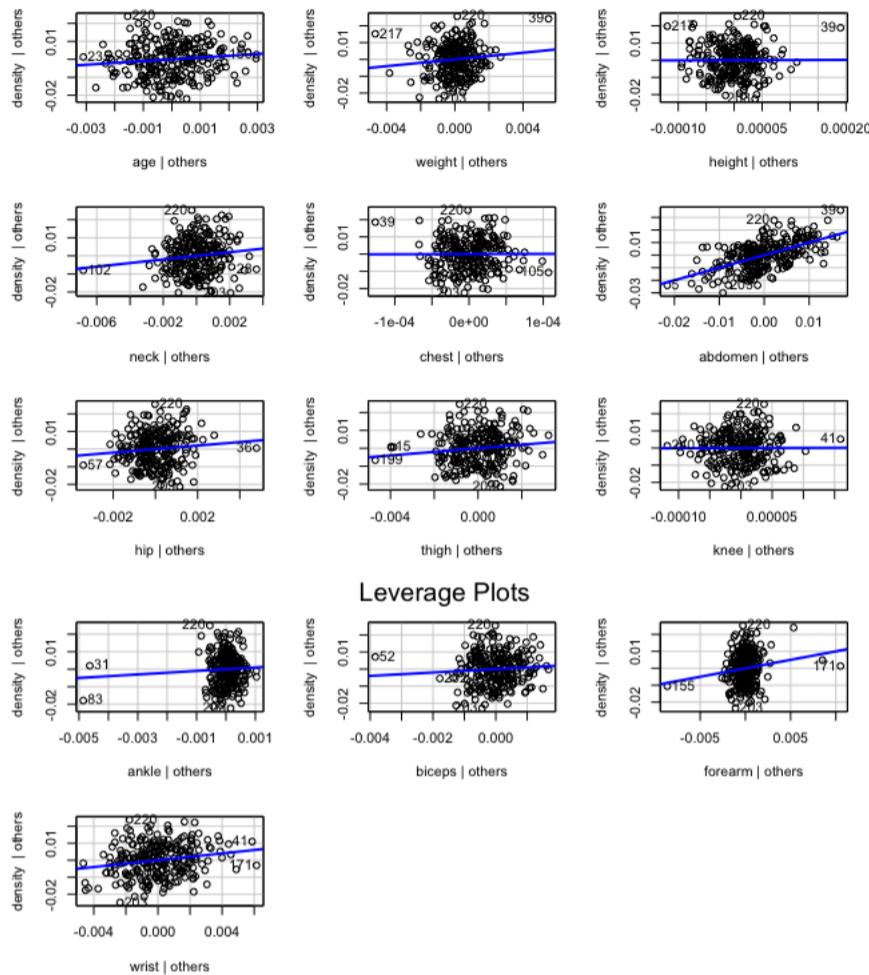


Figure 9: Leverage of regressors

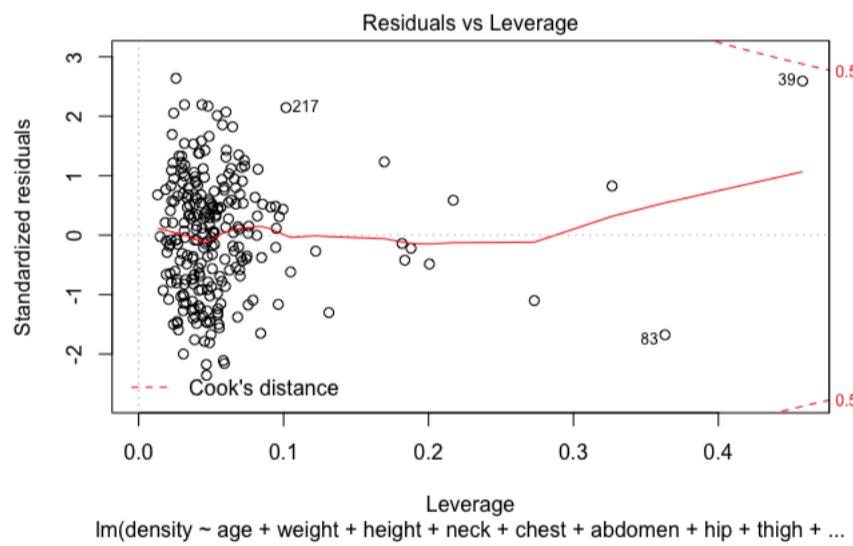


Figure 10: Residuals V.S. Leverage

2.4 Variable Transformation

Then we will consider possible variable transformation to improve the model.

According to the externally studentized residuals against fitted values plot [5], there is no obvious heteroscedasticity between fitted values. Also we can perform `ncvtest()` to test heteroscedasticity and the result is as follows [13].

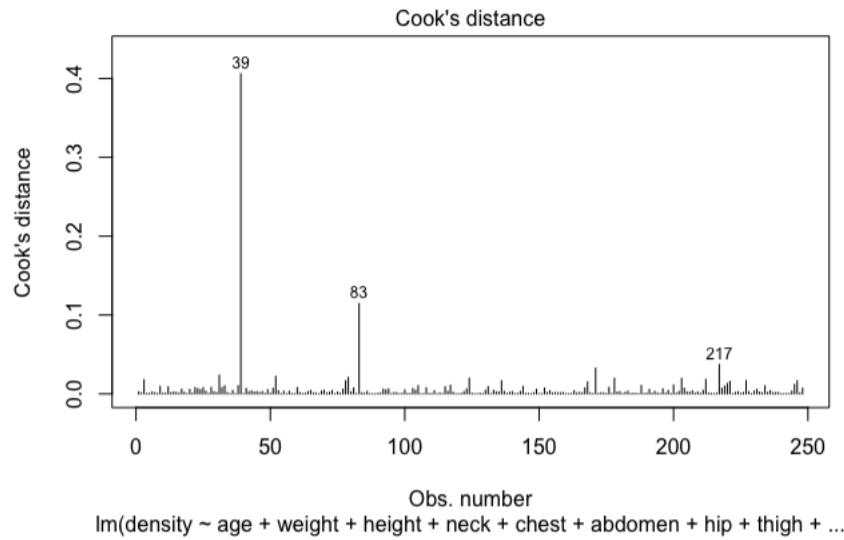


Figure 11: Cook's Distance

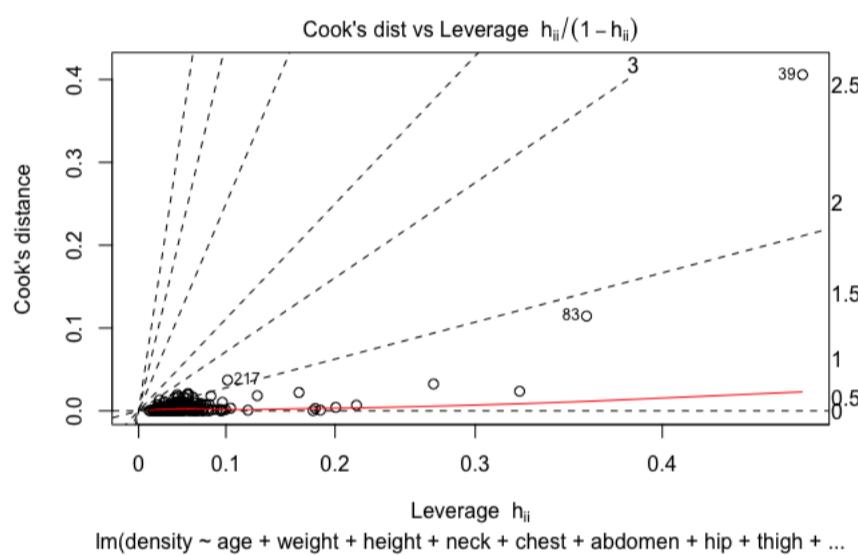


Figure 12: Cook's Distance vs Leverage

```
> ncvTest(model)
Non-constant Variance Score Test
Variance formula: ~ fitted.values
Chisquare = 0.1116408, Df = 1, p = 0.73828
```

Figure 13: Non-Constant Variance Test

However, we can also perform Box-Cox Transform to see if there is any possible power transforms. We use the `boxCox()` function from the `car` package and got a graph of log-likelihood of the parameter λ [14].

As the graph illustrates, the constant transform of the parameter 1 is within the confidence interval of 95% so we can continue safely with the original model.

$$\begin{aligned}
\text{density} = & 1.156121 - 1.319699 \times 10^{-4} * \text{age} + 2.377541 \times 10^{-4} * \text{weight} \\
& - 2.594465 \times 10^{-5} * \text{height} + 1.071540 \times 10^{-3} * \text{neck} + 1.168530 \times 10^{-5} * \text{chest} \\
& - 2.199601 \times 10^{-3} * \text{abdomen} + 5.267858 \times 10^{-4} * \text{hip} - 6.342697 \times 10^{-4} * \text{thigh} \\
& - 3.418355 \times 10^{-5} * \text{knee} - 4.448680 \times 10^{-4} * \text{ankle} - 4.273484 \times 10^{-4} * \text{biceps} \\
& - 1.040264 \times 10^{-3} * \text{forearm} + 3.651081 \times 10^{-3} * \text{wrist}
\end{aligned} \tag{2}$$

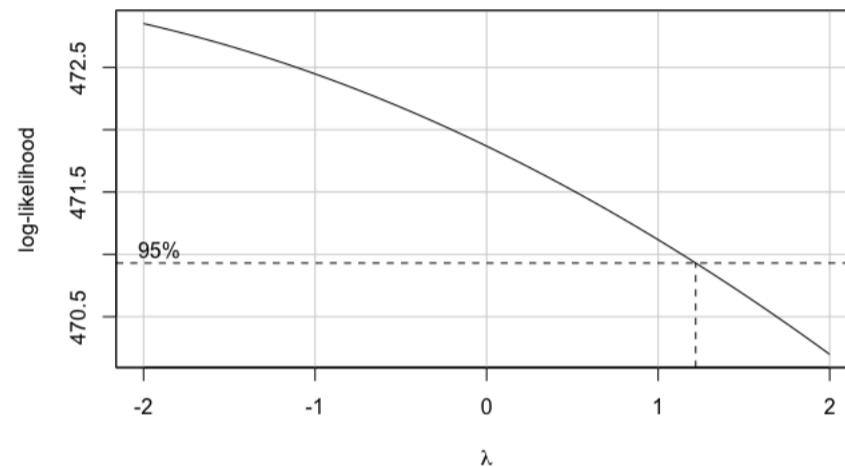


Figure 14: Log-likelihood of λ

2.5 Multicollinearity diagnostics and treatments

2.5.1 Variance Inflation factors

One way to check multicollinearity is to calculate VIFs.

$$VIF_j = C_{jj} = (1 - R_j^2)^{-1}$$

The results are as follows[15].

```

> vif(model)
      age    weight    height    neck    chest
      2.256000 43.944746  2.865731  4.391047 10.165371
abdomen      hip    thigh    knee    ankle
      12.881638 14.546865  7.815291  4.744625  1.952864
biceps    forearm    wrist
      3.683412  2.172323   3.354584

```

Figure 15: VIF

Because the maximum value is 43.94 so there exists multicollinearity.

2.5.2 Eigensystem Analysis

Then we try another way of multicollinearity diagnostics is **Eigensystem Analysis** by calculating condition indices[16]. Since many values of condition indices exceed 100, there exists multicollinearity.

2.5.3 Ridge Regression

To address the multicollinearity, we can use ridge regression and we use the `linearRidge()` function from the `ridge` package.

cond.index	
1	1.00000
2	13.72002
3	26.26194
4	59.90731
5	66.43640
6	81.47706
7	88.54941
8	96.13695
9	111.28599
10	146.64233
11	148.03658
12	155.09323
13	193.44492
14	389.25888

Figure 16: Condition Indices

```
Call:
linearRidge(formula = density^4 ~ age + weight + height + neck +
+ chest + abdomen + hip + thigh + knee + ankle + biceps +
forearm + wrist, data = mendata)

Coefficients:
            Estimate Scaled estimate Std. Error (scaled) t value (scaled) Pr(>|t| )
(Intercept) 0.6381081          NA             NA             NA             NA
age          0.0005514        0.1095838        0.0396031      2.767  0.00566 ** 
weight       -0.0002217       -0.1021407       0.0932969      1.095  0.27361  
height       -0.0015318       -0.0621158       0.0393504      1.579  0.11444  
neck         -0.0031983       -0.1220432       0.0545168      2.239  0.02518 *  
chest         0.0003103        0.0409896        0.0696246      0.589  0.55605  
abdomen       0.0052509        0.8898084        0.0770179     11.553 < 2e-16 *** 
hip           -0.0008015       -0.0896181       0.0776183      1.155  0.24825  
thigh         0.0013329        0.1092517        0.0677525      1.613  0.10685  
knee          -0.0001007       -0.0038054       0.0576420      0.066  0.94736  
ankle         0.0007217        0.0193147        0.0391247      0.494  0.62154  
biceps        0.0006767        0.0322248        0.0515354      0.625  0.53178  
forearm       0.0028941        0.0921612        0.0415145      2.220  0.02642 *  
wrist         -0.0120498       -0.1759642       0.0498843      3.527  0.00042 *** 

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Ridge parameter: 0.02135394, chosen automatically, computed using 6 PCs

Degrees of freedom: model 11.39 , variance 10.25 , residual 12.54
```

Figure 17: Ridge Regression Model

The result is as follows [17].

And we get the scaled ridge regression model [3]

$$\begin{aligned}
\text{density} = & 1.119775 - 1.918480 \times 10^{-4} * \text{age} + 8.416402 \times 10^{-5} * \text{weight} \\
& + 3.845388 \times 10^{-4} * \text{height} + 1.083985 \times 10^{-3} * \text{neck} - 1.0832994 \times 10^{-4} * \text{chest} \\
& - 1.698716 \times 10^{-3} * \text{abdomen} + 3.273120 \times 10^{-4} * \text{hip} - 5.345895 \times 10^{-4} * \text{thigh} \\
& - 4.764966 \times 10^{-5} * \text{knee} - 2.003772 \times 10^{-4} * \text{ankle} - 2.589640 \times 10^{-4} * \text{biceps} \\
& - 9.426395 \times 10^{-4} * \text{forearm} + 3.845387 \times 10^{-3} * \text{wrist}
\end{aligned} \tag{3}$$

2.5.4 Principal Component Regression

We use the `pcr()` function in the `pls` package to conduct our PCR(Principal Component Regression).

```

Data:      X dimension: 248 13
          Y dimension: 248 1
Fit method: svdpc
Number of components considered: 13

VALIDATION: RMSEP
Cross-validated using 10 random segments.
      (Intercept) 1 comps 2 comps 3 comps 4 comps 5 comps 6 comps 7 comps 8 comps 9 comps
CV          0.0189 0.01495 0.01311 0.01229 0.01219 0.01240 0.01165 0.01146 0.01136 0.01114
adjCV       0.0189 0.01495 0.01309 0.01227 0.01215 0.01237 0.01161 0.01143 0.01133 0.01113

      12 comps 13 comps
CV          0.01037 0.01030
adjCV       0.01032 0.01025

TRAINING: % variance explained
      1 comps 2 comps 3 comps 4 comps 5 comps 6 comps 7 comps 8 comps 9 comps 10 comps
X           62.80   73.90   80.48   85.73   89.98   92.44   94.53   96.41   97.84   98.87
density     37.85   53.56   58.88   61.18   61.26   65.59   66.73   67.17   67.33   67.56

      13 comps
X           100.00
density     74.51

```

Figure 18: Principal Component Regression Model

To show how much each component contributes, we can use `validationplot()` function[19].

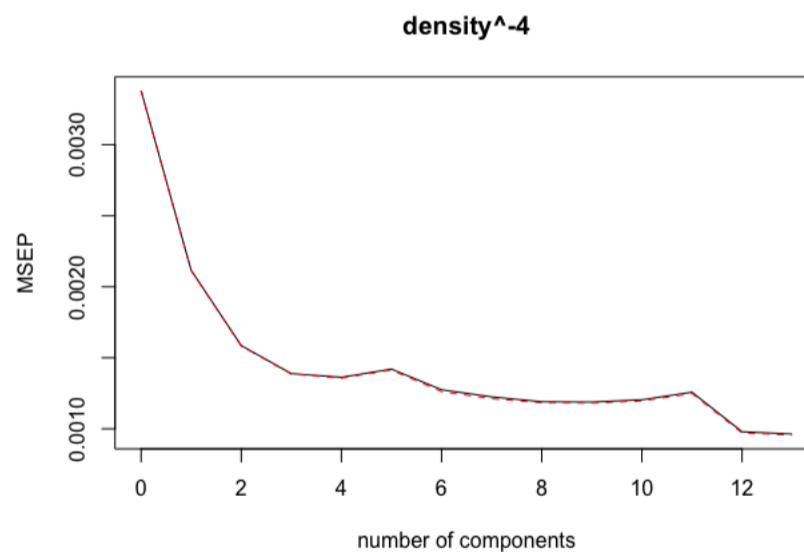


Figure 19: MSEP

And we get the scaled Principal Component Regression Model [4] using 6 components.

$$\begin{aligned} \text{density} = & 0.01187 - 0.0047059537 * \text{age} - 0.0017569642 * \text{weight} \\ & + 0.0037269895 * \text{height} + 0.0026024632 * \text{neck} - 0.0037694967 * \text{chest} \\ & - 0.0052066460 * \text{abdomen} - 0.0032809032 * \text{hip} - 0.0025914377 * \text{thigh} \\ & - 0.0036500052 * \text{knee} + 0.0012599117 * \text{ankle} - 0.0005902094 * \text{biceps} \\ & - 0.0012072402 * \text{forearm} + 0.0040723312 * \text{wrist} \end{aligned} \quad (4)$$

2.6 Variable Selection

2.6.1 All possible regression

First, we try to perform all possible regression for variable selection by using the `ols_all_possible()` function in the `olsrr` package.

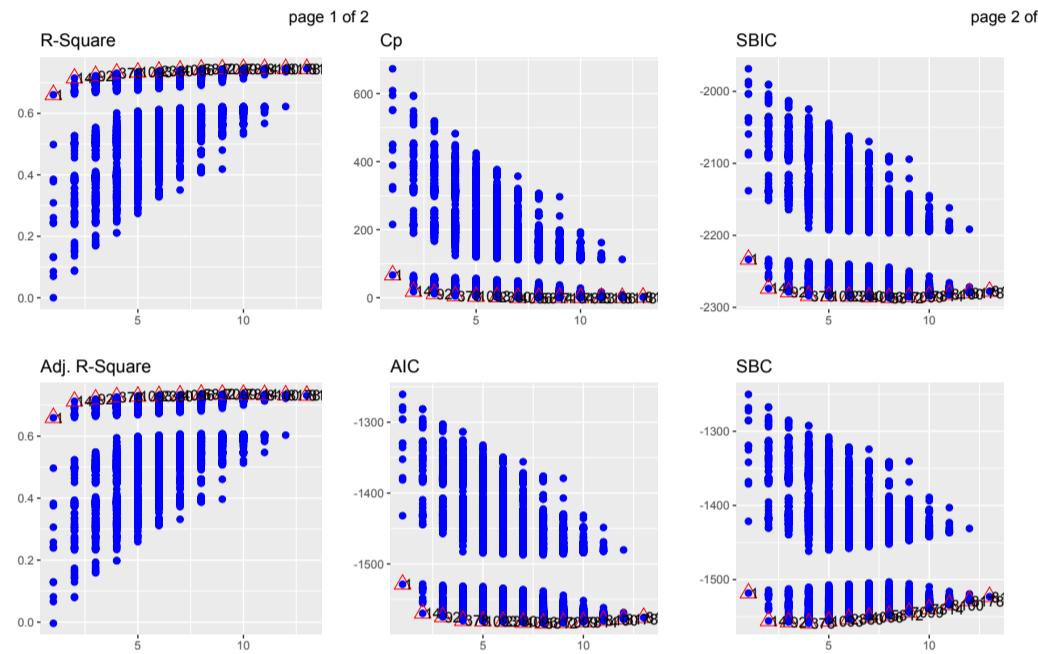


Figure 20: All Possible Regression Results

Although it took a quite long time to calculate all the $2^{13} = 8192$ combinations of variables, we can compare all possible models based on the result.

Measured by the adjusted R-square, akaike information criteria, and bayesian information criteria, we finally choose 8 variables, *age, weight, neck, abdomen, hip, thigh, forearm, wrist*

Using them as predictors, we got our reduced model(called RM1)[5].

$$\begin{aligned} \text{density} = & 1.1458191899 - 0.0001351300 * \text{age} + 0.0001893056 * \text{weight} + 0.0010947786 * \text{neck} - 0.0021502245 * \text{abdomen} \\ & + 0.0005607852 * \text{hip} - 0.0007324054 * \text{thigh} - 0.0011777856 * \text{forearm} + 0.0033211813 * \text{wrist} \end{aligned} \quad (5)$$

2.6.2 Stepwise Regression

For our model, all possible regression is feasible, however, for more variables involved, all possible regression would be impossible so we still need other methods, such as Stepwise Regression and forward/backward regression.

As for Stepwise Regression, we can use the `ols_step_forward_p()` function in the `olsrr` package.

As the output shows, we can just use 4 predictors *abdomen, weight, wrist and forearm* to construct another reduced model(called RM2)[6].

$$\begin{aligned} \text{density} = & 1.1791672 + 0.0003043 * \text{weight} - 0.0022308 * \text{abdomen} \\ & - 0.0011152 * \text{forearm} + 0.0033099 * \text{wrist} \end{aligned} \quad (6)$$

2.6.3 Validation of variable selection

We can use stepwise regression for variable selection.The modified model is as follows [21].

First, we perform log-likelihood ratio test between our full model and two reduced models. To do this, we use the `lrtest()` function from the `lmtest` package.

According to the test, there is no significant difference between the full model and reduced model 1, but there is significant difference between the reduced model 1 and reduced model 2. So finally we adopted the RM1[5].

Having selected the reduced model, we can also use 10-fold cross validation method to validate the variable selection. Here we use the `caret` package (the reference link is <https://machinelearningmastery.com/how-to-estimate-model-accuracy-in-r-using-the-caret-package/>)

So we can safely choose to use 8 predictors to build the final model.

2.7 Bootstrapping

The last step is to use bootstrapping to check the reduced model [5] by using the `Boot()` function in `car` package and then we get the final model [24].

Also we can get the confidence intervals of coefficients by bootstrapping[25].

3 Results

3.1 Our Models

1. First we establish the original full model by simple linear regression .
2. Then after residual analysis and leverage and influence diagnostic we try to use Box-Cox Transform to build a new model however it appears there is no need for variable transformation.
3. Through multicollinearity diagnostics, we try to address the multicollinearity with two different approaches, Ridge Regression and Principal Component Regression(PCR).
4. Based on previous models, we then consider our selection of predictors. We first use all possible regression and stepwise regression to eliminate some predictors and construct the reduced model. Then we use log-likelihood ratio test to test the full model and reduced model, and perform cross validation to further validate the reduced model.
5. The last step is to assess last model with bootstrapping, through this, we get the estimated prediction standard error and confidence intervals of coefficients, thus achieving the final model.

3.2 Assessment

- The first model[1] is easiest to build , however, is neither validated nor modified.
- After model adequacy checking, we get the second model[2], although we didn't perform any transformation, but it improves a lot because of our analysis. But the model[2] still have problems such as multicollinearity.
- To handle this, we construct the 3rd[3] and 4th model[4], these two models tend to reach perfection and works well but still need to be simplified.
- By two means we get 2 reduced model[5,6] and the test result shows that the full model is too complicated and the reduced model[6] is too simple so we choose model[5] as our final model.
- Finally we use bootstrapping for a small modification to complete the final model. However, it is still far from perfect, one possible way to improve it is to maybe add some non-linear components to the regression model.

4 Conclusion

In this project, we try to establish a useful model for bodyfat. It is easy to raise a model, but it is even easier to overthrow it. So the most difficult part is how to modify the model to be more sturdy.

Through all of those struggle, we finally manage to set up our model. It is not so beautiful, viewed from a mathematical prospective, but it is built and tested on the basis of practical data.

It is hard to interpret what the equation really means (maybe it doesn't have any meaning at all) but it can be of practical use even without any practical meaning.

$$\begin{aligned} \text{density} = & 1.1458191899 - 0.0001351300 * \text{age} + 0.0001893056 * \text{weight} + 0.0010947786 * \text{neck} - 0.0021502245 * \text{abdomen} \\ & + 0.0005607852 * \text{hip} - 0.0007324054 * \text{thigh} - 0.0011777856 * \text{forearm} + 0.0033211813 * \text{wrist} \end{aligned}$$

The conclusion sounds boring, but it takes much work to come to it, maybe that is why it is more like empirical rather than theoretical work.

```

> ols_step_both_p(model)
Stepwise Selection Method
-----
Candidate Terms:
1. age
2. weight
3. height
4. neck
5. chest
6. abdomen
7. hip
8. thigh
9. knee
10. ankle
11. biceps
12. forearm
13. wrist

We are selecting variables based on p value...

Variables Entered/Removed:
✓ abdomen
✓ weight
✓ wrist
✓ forearm

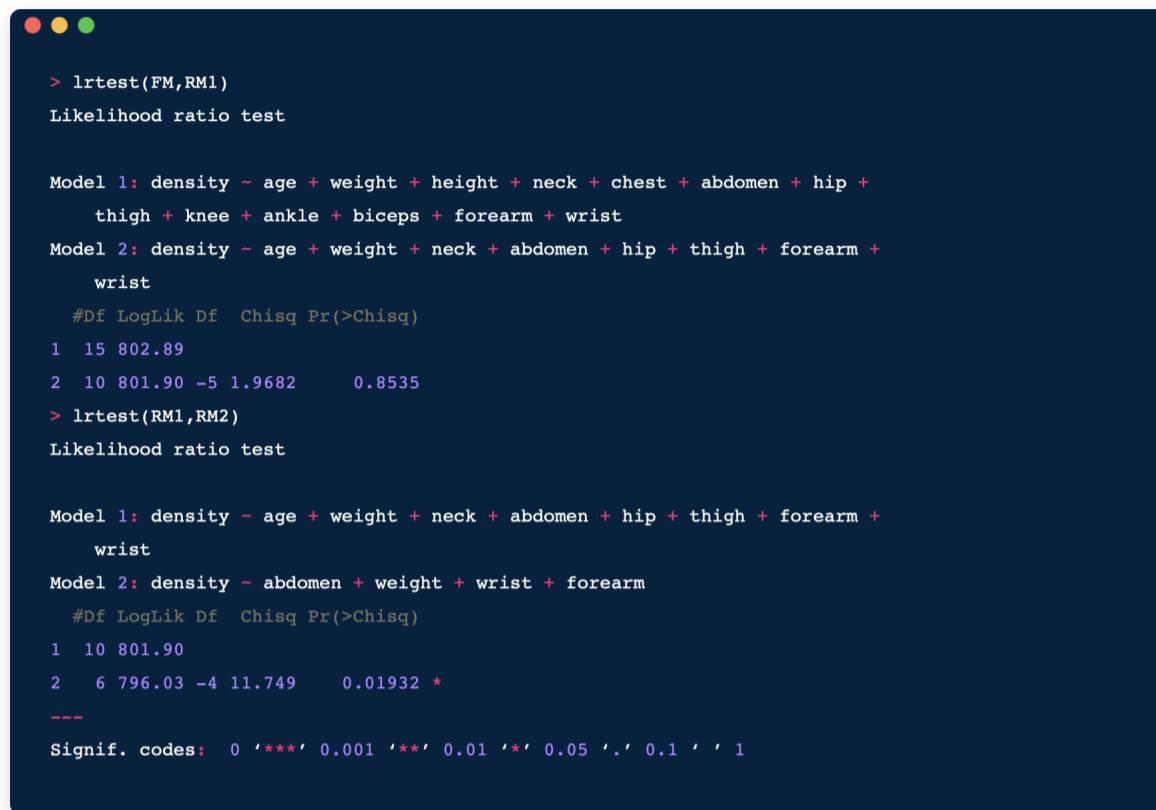
No more variables to be added/removed.

Final Model Output
-----
Model Summary
-----
R          0.855      RMSE       0.010
R-Squared  0.731      Coef. Var  0.935
Adj. R-Squared 0.726      MSE        0.000
Pred R-Squared 0.716      MAE        0.008
-----
RMSE: Root Mean Square Error
MSE: Mean Square Error
MAE: Mean Absolute Error

ANOVA
-----
Sum of
Squares   DF   Mean Square   F      Sig.
-----
Regression 0.064     4      0.016    164.791  0.0000
Residual   0.024    243      0.000
Total      0.088    247
-----
Parameter Estimates
-----
model   Beta   Std. Error   Std. Beta   t      Sig.   lower   upper
-----
(Intercept) 1.179     0.017           71.042  0.000   1.146   1.212
abdomen   -0.002     0.000      -1.276  -17.401  0.000  -0.002  -0.002
weight     0.000     0.000       0.473   5.382  0.000   0.000   0.000
wrist      0.003     0.001       0.163   3.228  0.001   0.001   0.005
forearm   -0.001     0.000      -0.120  -2.697  0.007  -0.002  0.000
-----
Stepwise Selection Summary
-----
Step  Variable  Added/Removed  R-Square  Adj. R-Square  C(p)  AIC  RMSE
-----
1  abdomen  addition  0.661  0.659  67.5630  -1528.7740  0.0110
2  weight   addition  0.715  0.712  20.0300  -1569.7134  0.0101
3  wrist    addition  0.723  0.719  14.7090  -1574.7409  0.0100
4  forearm  addition  0.731  0.726  9.3080  -1580.0541  0.0099

```

Figure 21: Stepwise Regression



```

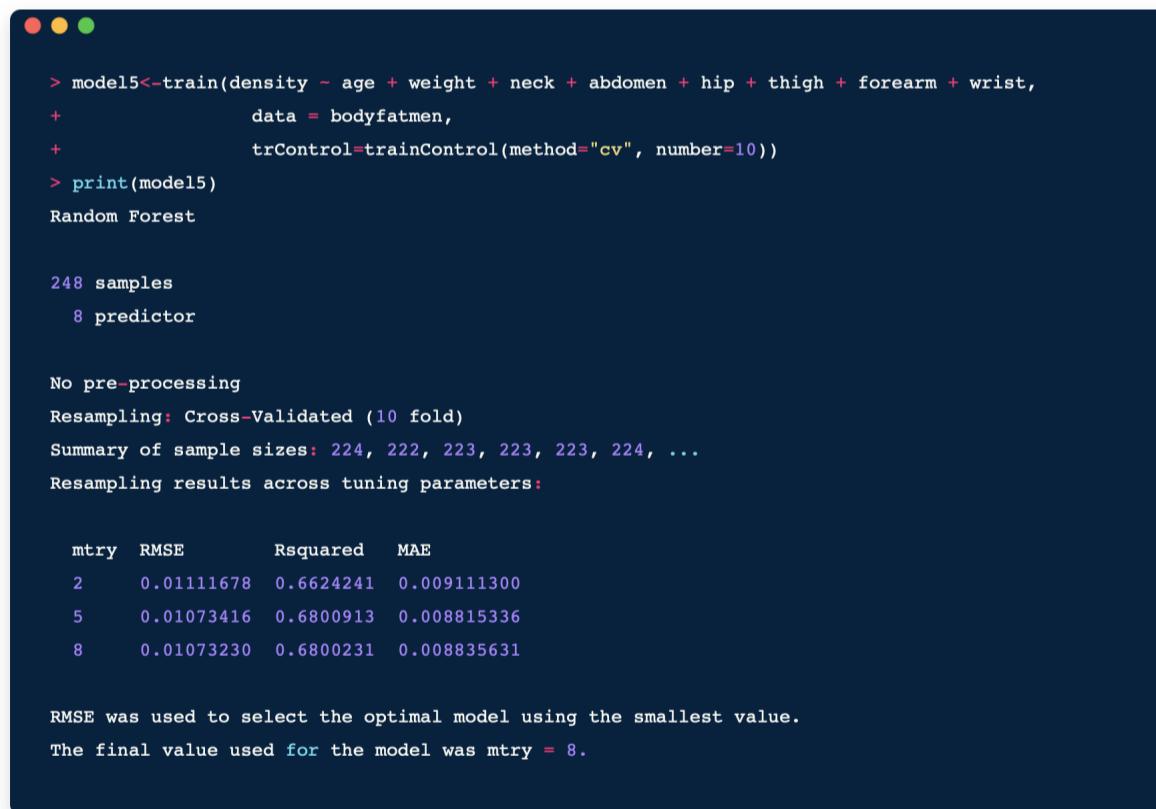
> lrtest(FM,RM1)
Likelihood ratio test

Model 1: density ~ age + weight + height + neck + chest + abdomen + hip +
    thigh + knee + ankle + biceps + forearm + wrist
Model 2: density ~ age + weight + neck + abdomen + hip + thigh + forearm +
    wrist
#Df LogLik Df  Chisq Pr(>Chisq)
1  15 802.89
2  10 801.90 -5 1.9682     0.8535
> lrtest(RM1,RM2)
Likelihood ratio test

Model 1: density ~ age + weight + neck + abdomen + hip + thigh + forearm +
    wrist
Model 2: density ~ abdomen + weight + wrist + forearm
#Df LogLik Df  Chisq Pr(>Chisq)
1  10 801.90
2   6 796.03 -4 11.749    0.01932 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figure 22: log-likelihood ratio test



```

> model5<-train(density ~ age + weight + neck + abdomen + hip + thigh + forearm + wrist,
+                   data = bodyfatmen,
+                   trControl=trainControl(method="cv", number=10))
> print(model5)
Random Forest

248 samples
  8 predictor

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 224, 222, 223, 223, 223, 224, ...
Resampling results across tuning parameters:

  mtry   RMSE      Rsquared      MAE
  2      0.01111678  0.6624241  0.009111300
  5      0.01073416  0.6800913  0.008815336
  8      0.01073230  0.6800231  0.008835631

RMSE was used to select the optimal model using the smallest value.
The final value used for the model was mtry = 8.

```

Figure 23: Cross Validation

```
Number of bootstrap replications R = 999
      original   bootBias   bootSE   bootMed
(Intercept) 1.14581919 -4.5038e-04 2.4870e-02 1.14608046
age         -0.00013513 -9.1925e-06 6.3262e-05 -0.00014448
weight       0.00018931 -1.1972e-05 9.2544e-05 0.00017796
neck          0.00109478 -1.6424e-05 5.0433e-04 0.00108179
abdomen      -0.00215022 2.3222e-05 1.6362e-04 -0.00212449
hip            0.00056079 -4.4153e-06 2.8674e-04 0.00055453
thigh         -0.00073241 -1.1254e-05 2.6666e-04 -0.00073458
forearm      -0.00117779 2.4068e-05 4.6791e-04 -0.00115281
wrist          0.00332118 1.0380e-04 1.1563e-03 0.00339349
```

Figure 24: Bootstrapping

```
Bootstrap bca confidence intervals

      2.5 %      97.5 %
(Intercept) 1.094415e+00 1.192978e+00
age        -2.476209e-04 -5.997014e-06
weight       1.954450e-05  3.847767e-04
neck          1.439075e-04  2.077178e-03
abdomen      -2.466253e-03 -1.857902e-03
hip            1.757445e-05  1.144802e-03
thigh         -1.286049e-03 -2.343162e-04
forearm      -2.214744e-03 -3.570913e-04
wrist          9.313450e-04  5.690133e-03
```

Figure 25: Confidence Intervals of Coefficients