```
> #HW11
> if (FALSE)
+ {"
+ Use Ex_Cross_Validation.R to perform a 10-fold cross validation to choose between the model
+ that includes weight and abdom with the model that includes weight, abdom, and thigh.Choose the
+ model by RMSE, by Rsquared, and by MAE criteria. Are the choices consistent? Submit the output (no graphics needed)
+ and a short paragraph into Canvas summarizing the results.
+ "}
> ##------------------------------------------------------------------------------##
> library(faraway)
> library(caret)
> library(olsrr)
>
> #model one
> one <- lm(brozek ~ abdom + weight, data=fat)
> summary(one)

Call:
lm(formula = brozek ~ abdom + weight, data = fat)

Residuals:
     Min       1Q   Median       3Q      Max
-10.8307  -2.9773   0.0237   2.9397   9.7679

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -41.34812    2.41299 -17.136  < 2e-16 ***
abdom         0.91514    0.05254  17.419  < 2e-16 ***
weight       -0.13645    0.01928  -7.079 1.47e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.127 on 249 degrees of freedom
Multiple R-squared:  0.7187,    Adjusted R-squared:  0.7165
F-statistic: 318.1 on 2 and 249 DF,  p-value: < 2.2e-16

>
> #model two
> two <- lm(brozek ~ abdom + weight + thigh, data=fat)
> summary(two)

Call:
lm(formula = brozek ~ abdom + weight + thigh, data = fat)

Residuals:
    Min      1Q  Median      3Q     Max
-10.675  -2.973  -0.047   3.031   9.442

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) -48.03851    3.98692 -12.049  < 2e-16 ***
abdom         0.91743    0.05219  17.578  < 2e-16 ***
weight       -0.16963    0.02482  -6.834 6.37e-11 ***
thigh         0.20896    0.09952   2.100   0.0368 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.099 on 248 degrees of freedom
Multiple R-squared:  0.7236,    Adjusted R-squared:  0.7203
F-statistic: 216.5 on 3 and 248 DF,  p-value: < 2.2e-16

>
> #Start the cross-validation to choose between the model that includes weight and abdom
> # with the model that includes weight, abdom, and thigh
>
> # Define training control
> set.seed(13245)
> #asks for 10-fold validation with 5 repeats
> train.control <- trainControl(method = "repeatedcv",repeats=5, number = 10)
> ##the choices in this function are either 10 or 25
>
> # Train the model
> model_one <- train(brozek ~ abdom + weight,data = fat, method = "lm",
+ trControl = train.control)
> # Summarize the results
> print(model_one)
Linear Regression

252 samples
  2 predictor

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 5 times)
Summary of sample sizes: 226, 227, 227, 227, 227, 226, ...
Resampling results:

  RMSE      Rsquared   MAE
  4.138209  0.7182581  3.385683

Tuning parameter 'intercept' was held constant at a value of TRUE
> ##
>
> # Define training control
> set.seed(14235)
> #asks for 10-fold validation with 5 repeats
> train.control <- trainControl(method = "repeatedcv",repeats=5, number = 10)
> # Train the model
> model_two <- train(brozek ~ abdom + weight + thigh,data = fat, method = "lm",
+ trControl = train.control)
> # Summarize the results
```

```
> print(model_two)
Linear Regression

252 samples
  3 predictor

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 5 times)
Summary of sample sizes: 226, 227, 225, 227, 228, 226, ...
Resampling results:

  RMSE      Rsquared   MAE
  4.120946  0.7260265  3.386503

Tuning parameter 'intercept' was held constant at a value of TRUE
> ##
>
> summary(one)

Call:
lm(formula = brozek ~ abdom + weight, data = fat)

Residuals:
     Min       1Q   Median       3Q      Max
-10.8307  -2.9773   0.0237   2.9397   9.7679

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -41.34812    2.41299 -17.136  < 2e-16 ***
abdom         0.91514    0.05254  17.419  < 2e-16 ***
weight       -0.13645    0.01928  -7.079 1.47e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.127 on 249 degrees of freedom
Multiple R-squared:  0.7187,    Adjusted R-squared:  0.7165
F-statistic: 318.1 on 2 and 249 DF,  p-value: < 2.2e-16

> summary(two)

Call:
lm(formula = brozek ~ abdom + weight + thigh, data = fat)

Residuals:
    Min      1Q  Median      3Q     Max
-10.675  -2.973  -0.047   3.031   9.442

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -48.03851    3.98692 -12.049  < 2e-16 ***
```

```
abdom         0.91743    0.05219  17.578  < 2e-16 ***
weight       -0.16963    0.02482  -6.834 6.37e-11 ***
thigh         0.20896    0.09952   2.100   0.0368 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.099 on 248 degrees of freedom
Multiple R-squared:  0.7236,    Adjusted R-squared:  0.7203
F-statistic: 216.5 on 3 and 248 DF,  p-value: < 2.2e-16

>
> ####End of 10-fold Cross Validation####
>
```