```r
## Coffee example  (Coffee Quality Institute, 2018) continued
coffee <- read.csv("coffee_arabica.csv")

mfull <- lm(Flavor~ factor(Processing.Method) + Aroma + Aftertaste +
    Body + Acidity + Balance + Sweetness + Uniformity + Moisture, dat=coffee)
summary(mfull)$adj.r.squared
```

```
## [1] 0.8073297
```

```r
AIC(mfull)
```

```
## [1] -1087.524
```

```r
BIC(mfull)
```

```
## [1] -1027.282
```

```r
library(leaps)
all_regs <- regsubsets(Flavor ~ ., data = coffee, nvmax = 10, nbest = 2^10,
                       really.big = TRUE)
all_regs_summ <- summary(all_regs)
# all_regs_summ$which
# all_regs_summ$adjr2
# all_regs_summ$bic

# Organize results according to number of variables in model
p <- 10
k <- c(rep(1, choose(p,1)),
        rep(2, choose(p,2)),
        rep(3, choose(p,3)),
        rep(4, choose(p,4)),
        rep(5, choose(p,5)),
        rep(6, choose(p,6)),
        rep(7, choose(p,7)),
        rep(8, choose(p,8)),
        rep(9, choose(p,9)),
        rep(10, choose(p,10)))
boxplot(all_regs_summ$adjr2 ~ k, xlab = "Number of predictors", ylab =
        expression(R[adj]^2), ylim = c(0,1))
abline(h = c(0,1), lty = 2, col = "red")
```
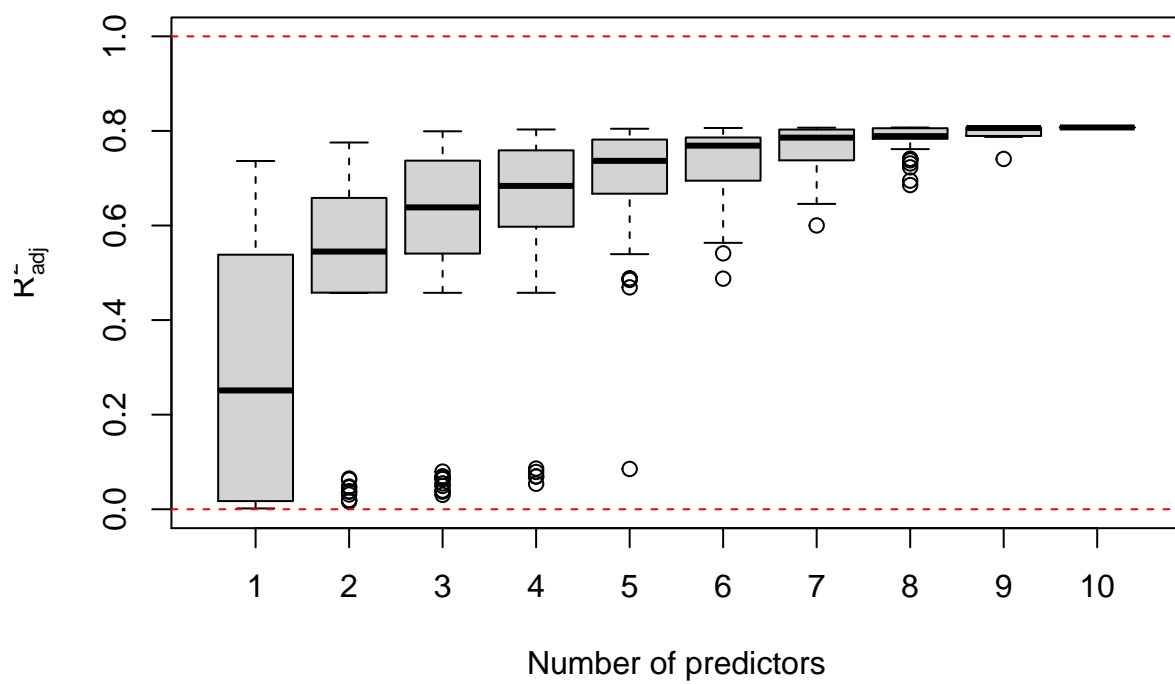
```r
boxplot(all_regs_summ$bic ~ k, xlab = "Number of predictors", ylab = "BIC")
```

Number of predictors
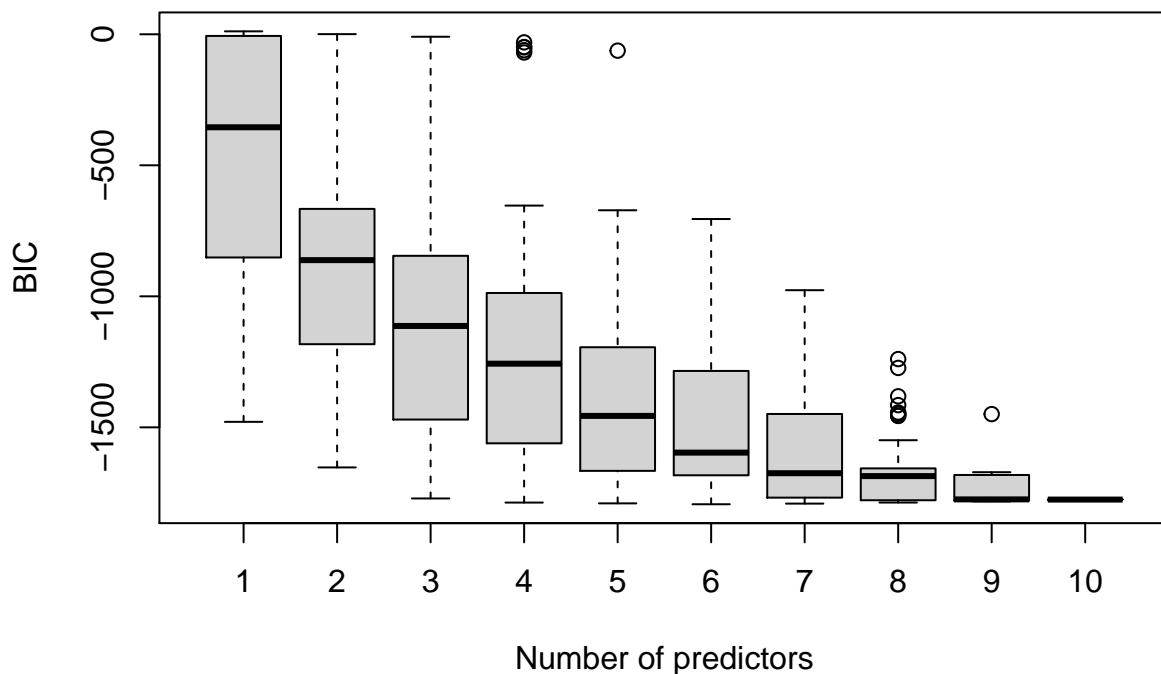
```r
max(all_regs_summ$adjr2)
```

```
## [1] 0.8075027
```

```r
bestR2adj <- which.max(all_regs_summ$adjr2)
min(all_regs_summ$bic)
```

```
## [1] -1793.389
```

```r
bestBIC <- which.min(all_regs_summ$bic)

# Find out which predictors in those models
all_regs_summ$which[bestR2adj,]
```

```
##                                     (Intercept)
##                                            TRUE
## Processing.MethodSemi-washed / Semi-pulped
##                                           FALSE
##               Processing.MethodWashed / Wet
##                                            TRUE
##                                           Aroma
##                                            TRUE
##                                      Aftertaste
##                                            TRUE
##                                            Body
##                                            TRUE
##                                          Acidity
##                                            TRUE
```

```
##                               Balance
##                                  TRUE
##                             Sweetness
##                                  TRUE
##                            Uniformity
##                                  TRUE
##                              Moisture
##                                  TRUE
```

```
all_regs_summ$which[bestBIC,]
```

```
##                           (Intercept)
##                                  TRUE
## Processing.MethodSemi-washed / Semi-pulped
##                                 FALSE
##          Processing.MethodWashed / Wet
##                                  TRUE
##                                 Aroma
##                                  TRUE
##                             Aftertaste
##                                  TRUE
##                                  Body
##                                  TRUE
##                                Acidity
##                                  TRUE
##                                Balance
##                                 FALSE
##                             Sweetness
##                                  TRUE
##                            Uniformity
##                                 FALSE
##                              Moisture
##                                 FALSE
```

```
coffee$wet <- ifelse(coffee$Processing.Method == 'Washed / Wet', 1,
                 0) # 1 = wet, 0 otherwise
coffee$semi <- ifelse(coffee$Processing.Method == 'Semi-washed / Semi-pulped',
                 1, 0) # 1 = semi/dry, 0 otherwise
coffee$Processing.Method <- NULL

m_bestr2adj <- lm(Flavor~ wet + Aroma + Aftertaste +
         Body + Acidity + Balance + Sweetness + Uniformity + Moisture,
         dat=coffee)
summary(m_bestr2adj)
```

```
##
## Call:
## lm(formula = Flavor ~ wet + Aroma + Aftertaste + Body + Acidity +
##      Balance + Sweetness + Uniformity + Moisture, data = coffee)
##
## Residuals:
##       Min       1Q   Median       3Q      Max
## -0.68587 -0.08469  0.00080  0.08923  0.63660
##
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.728709   0.168439  -4.326 1.65e-05 ***
## wet         -0.032797   0.010197  -3.216  0.00134 **
## Aroma        0.220278   0.020434  10.780  < 2e-16 ***
## Aftertaste   0.468749   0.023901  19.612  < 2e-16 ***
## Body         0.096194   0.024308   3.957 8.06e-05 ***
## Acidity      0.216754   0.021185  10.232  < 2e-16 ***
## Balance      0.046793   0.022547   2.075  0.03819 *
## Sweetness    0.025480   0.010136   2.514  0.01209 *
## Uniformity   0.016291   0.009798   1.663  0.09665 .
## Moisture     0.168439   0.102033   1.651  0.09906 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1479 on 1109 degrees of freedom
## Multiple R-squared:  0.8091, Adjusted R-squared:  0.8075
## F-statistic: 522.1 on 9 and 1109 DF,  p-value: < 2.2e-16
```

```r
AIC(m_bestr2adj)
```

```
## [1] -1089.52
```

```r
BIC(m_bestr2adj)
```

```
## [1] -1034.298
```

```r
m_bestBIC <- lm(Flavor~ wet + Aroma + Aftertaste +
                 Body + Acidity + Sweetness , dat=coffee)
summary(m_bestBIC)
```

```
##
## Call:
## lm(formula = Flavor ~ wet + Aroma + Aftertaste + Body + Acidity +
##     Sweetness, data = coffee)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.65627 -0.08781  0.00032  0.08529  0.63010
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.609003   0.159910  -3.808 0.000148 ***
## wet         -0.032852   0.010198  -3.221 0.001313 **
## Aroma        0.225969   0.020378  11.089  < 2e-16 ***
## Aftertaste   0.490988   0.021938  22.381  < 2e-16 ***
## Body         0.103438   0.022926   4.512 7.11e-06 ***
## Acidity      0.225638   0.020994  10.748  < 2e-16 ***
## Sweetness    0.033445   0.009582   3.491 0.000501 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1484 on 1112 degrees of freedom
## Multiple R-squared:  0.8073, Adjusted R-squared:  0.8063
## F-statistic: 776.4 on 6 and 1112 DF,  p-value: < 2.2e-16
```

```r
AIC(m_bestBIC)
```

```
## [1] -1085.26
```

```r
BIC(m_bestBIC)
```

```
## [1] -1045.098
```

```r
# Let's also try stepwise methods
library(MASS)

# Full model and empty model with just intercept
full <- lm(Flavor ~ ., data = coffee)
empty <- lm(Flavor ~ 1, data = coffee)

# default stepAIC uses AIC criterion
stepAIC(object = empty, scope = list(upper = full, lower = empty), direction
        = "forward")
```

```
## Start:  AIC=-2432.31
## Flavor ~ 1
##
##               Df Sum of Sq     RSS      AIC
## + Aftertaste   1    93.607  33.465 -3923.3
## + Acidity      1    69.294  57.778 -3312.3
## + Aroma        1    68.457  58.615 -3296.1
## + Balance      1    68.173  58.899 -3290.7
## + Body         1    58.232  68.840 -3116.2
## + Uniformity   1     5.778 121.294 -2482.4
## + wet          1     2.313 124.759 -2450.9
## + Sweetness    1     2.300 124.772 -2450.8
## + Moisture     1     2.239 124.833 -2450.2
## + semi         1     0.331 126.741 -2433.2
## <none>                     127.072 -2432.3
##
## Step:  AIC=-3923.33
## Flavor ~ Aftertaste
##
##               Df Sum of Sq     RSS      AIC
## + Acidity      1    4.9955  28.470 -4102.2
## + Aroma        1    4.9082  28.557 -4098.8
## + Body         1    2.2551  31.210 -3999.4
## + Balance      1    1.7369  31.729 -3981.0
## + Sweetness    1    0.1384  33.327 -3926.0
## + Uniformity   1    0.1143  33.351 -3925.2
## + wet          1    0.0871  33.378 -3924.2
## <none>                      33.465 -3923.3
## + Moisture     1    0.0382  33.427 -3922.6
## + semi         1    0.0179  33.448 -3921.9
##
## Step:  AIC=-4102.23
## Flavor ~ Aftertaste + Acidity
##
##               Df Sum of Sq     RSS      AIC
## + Aroma        1   3.02166  25.448 -4225.8
```

```
## + Body         1   0.89556 27.575 -4136.0
## + Balance      1   0.65424 27.816 -4126.2
## + wet          1   0.22561 28.244 -4109.1
## + Sweetness    1   0.17094 28.299 -4107.0
## + Uniformity   1   0.11428 28.356 -4104.7
## <none>                    28.470 -4102.2
## + semi         1   0.04453 28.425 -4102.0
## + Moisture     1   0.01991 28.450 -4101.0
##
## Step:  AIC=-4225.78
## Flavor ~ Aftertaste + Acidity + Aroma
##
##              Df Sum of Sq    RSS      AIC
## + Body        1   0.50898 24.939 -4246.4
## + Balance     1   0.32565 25.123 -4238.2
## + wet         1   0.26887 25.180 -4235.7
## + Sweetness   1   0.19006 25.258 -4232.2
## + Uniformity  1   0.11405 25.334 -4228.8
## <none>                   25.448 -4225.8
## + semi        1   0.04166 25.407 -4225.6
## + Moisture    1   0.01953 25.429 -4224.6
##
## Step:  AIC=-4246.39
## Flavor ~ Aftertaste + Acidity + Aroma + Body
##
##              Df Sum of Sq    RSS      AIC
## + Sweetness   1  0.223266 24.716 -4254.5
## + wet         1  0.183460 24.756 -4252.6
## + Uniformity  1  0.172293 24.767 -4252.1
## + Balance     1  0.132479 24.807 -4250.3
## + Moisture    1  0.058517 24.881 -4247.0
## <none>                   24.939 -4246.4
## + semi        1  0.040446 24.899 -4246.2
##
## Step:  AIC=-4254.45
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness
##
##              Df Sum of Sq    RSS      AIC
## + wet         1  0.228511 24.488 -4262.8
## + Balance     1  0.118556 24.598 -4257.8
## + Uniformity  1  0.075546 24.641 -4255.9
## <none>                   24.716 -4254.5
## + Moisture    1  0.038620 24.677 -4254.2
## + semi        1  0.037654 24.678 -4254.2
##
## Step:  AIC=-4262.84
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness + wet
##
##              Df Sum of Sq    RSS      AIC
## + Balance     1  0.099946 24.388 -4265.4
## + Uniformity  1  0.084340 24.403 -4264.7
## + Moisture    1  0.046681 24.441 -4263.0
## <none>                   24.488 -4262.8
## + semi        1  0.000241 24.487 -4260.9
```

```
##
## Step:  AIC=-4265.42
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness + wet +
##     Balance
##
##             Df Sum of Sq    RSS     AIC
## + Uniformity  1  0.063931 24.324 -4266.4
## + Moisture    1  0.063069 24.325 -4266.3
## <none>                    24.388 -4265.4
## + semi        1  0.000236 24.387 -4263.4
##
## Step:  AIC=-4266.36
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness + wet +
##     Balance + Uniformity
##
##            Df Sum of Sq    RSS     AIC
## + Moisture  1  0.059626 24.264 -4267.1
## <none>                  24.324 -4266.4
## + semi      1  0.000150 24.324 -4264.4
##
## Step:  AIC=-4267.1
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness + wet +
##     Balance + Uniformity + Moisture
##
##        Df  Sum of Sq    RSS     AIC
## <none>              24.264 -4267.1
## + semi  1 8.7985e-05 24.264 -4265.1
##
##
## Call:
## lm(formula = Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness +
##     wet + Balance + Uniformity + Moisture, data = coffee)
##
## Coefficients:
## (Intercept)    Aftertaste      Acidity        Aroma         Body    Sweetness
##    -0.72871       0.46875      0.21675      0.22028      0.09619      0.02548
##         wet       Balance   Uniformity     Moisture
##    -0.03280       0.04679      0.01629      0.16844
```

```r
# Let's get stepAIC to use BIC by specifying the penalty k = log(n)
# Forward
stepAIC(object = empty, scope = list(upper = full, lower = empty), direction
        = "forward", k = log(nrow(coffee)))
```

```
## Start:  AIC=-2427.29
## Flavor ~ 1
##
##               Df Sum of Sq     RSS     AIC
## + Aftertaste   1    93.607  33.465 -3913.3
## + Acidity      1    69.294  57.778 -3302.2
## + Aroma        1    68.457  58.615 -3286.1
## + Balance      1    68.173  58.899 -3280.7
## + Body         1    58.232  68.840 -3106.2
## + Uniformity   1     5.778 121.294 -2472.3
## + wet          1     2.313 124.759 -2440.8
```

```
## + Sweetness    1      2.300 124.772 -2440.7
## + Moisture     1      2.239 124.833 -2440.2
## <none>                       127.072 -2427.3
## + semi         1      0.331 126.741 -2423.2
##
## Step:  AIC=-3913.29
## Flavor ~ Aftertaste
##
##             Df Sum of Sq    RSS      AIC
## + Acidity    1     4.9955 28.470 -4087.2
## + Aroma      1     4.9082 28.557 -4083.7
## + Body       1     2.2551 31.210 -3984.3
## + Balance    1     1.7369 31.729 -3965.9
## <none>                    33.465 -3913.3
## + Sweetness  1     0.1384 33.327 -3910.9
## + Uniformity 1     0.1143 33.351 -3910.1
## + wet        1     0.0871 33.378 -3909.2
## + Moisture   1     0.0382 33.427 -3907.5
## + semi       1     0.0179 33.448 -3906.9
##
## Step:  AIC=-4087.17
## Flavor ~ Aftertaste + Acidity
##
##             Df Sum of Sq    RSS      AIC
## + Aroma      1    3.02166 25.448 -4205.7
## + Body       1    0.89556 27.575 -4115.9
## + Balance    1    0.65424 27.816 -4106.2
## + wet        1    0.22561 28.244 -4089.0
## <none>                    28.470 -4087.2
## + Sweetness  1    0.17094 28.299 -4086.9
## + Uniformity 1    0.11428 28.356 -4084.6
## + semi       1    0.04453 28.425 -4081.9
## + Moisture   1    0.01991 28.450 -4080.9
##
## Step:  AIC=-4205.7
## Flavor ~ Aftertaste + Acidity + Aroma
##
##             Df Sum of Sq    RSS      AIC
## + Body       1    0.50898 24.939 -4221.3
## + Balance    1    0.32565 25.123 -4213.1
## + wet        1    0.26887 25.180 -4210.6
## + Sweetness  1    0.19006 25.258 -4207.1
## <none>                    25.448 -4205.7
## + Uniformity 1    0.11405 25.334 -4203.7
## + semi       1    0.04166 25.407 -4200.5
## + Moisture   1    0.01953 25.429 -4199.5
##
## Step:  AIC=-4221.29
## Flavor ~ Aftertaste + Acidity + Aroma + Body
##
##             Df Sum of Sq    RSS      AIC
## + Sweetness  1   0.223266 24.716 -4224.3
## + wet        1   0.183460 24.756 -4222.5
## + Uniformity 1   0.172293 24.767 -4222.0
```

```
## <none>                        24.939 -4221.3
## + Balance    1  0.132479 24.807 -4220.2
## + Moisture   1  0.058517 24.881 -4216.9
## + semi       1  0.040446 24.899 -4216.1
##
## Step:  AIC=-4224.33
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness
##
##              Df Sum of Sq    RSS     AIC
## + wet         1  0.228511 24.488 -4227.7
## <none>                     24.716 -4224.3
## + Balance     1  0.118556 24.598 -4222.7
## + Uniformity  1  0.075546 24.641 -4220.7
## + Moisture    1  0.038620 24.677 -4219.1
## + semi        1  0.037654 24.678 -4219.0
##
## Step:  AIC=-4227.7
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness + wet
##
##              Df Sum of Sq    RSS     AIC
## <none>                     24.488 -4227.7
## + Balance     1  0.099946 24.388 -4225.3
## + Uniformity  1  0.084340 24.403 -4224.5
## + Moisture    1  0.046681 24.441 -4222.8
## + semi        1  0.000241 24.487 -4220.7
##
## Call:
## lm(formula = Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness +
##     wet, data = coffee)
##
## Coefficients:
## (Intercept)    Aftertaste       Acidity         Aroma          Body     Sweetness
##    -0.60900       0.49099       0.22564       0.22597       0.10344       0.03345
##         wet
##    -0.03285
```

```r
m_f <- stepAIC(object = empty, scope = list(upper = full, lower = empty),
               direction = "forward", trace = 0, k = log(nrow(coffee)))
summary(m_f)
```

```
##
## Call:
## lm(formula = Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness +
##     wet, data = coffee)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.65627 -0.08781  0.00032  0.08529  0.63010
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.609003   0.159910  -3.808 0.000148 ***
## Aftertaste   0.490988   0.021938  22.381  < 2e-16 ***
## Acidity      0.225638   0.020994  10.748  < 2e-16 ***
```

```
## Aroma          0.225969    0.020378   11.089  < 2e-16 ***
## Body            0.103438    0.022926    4.512 7.11e-06 ***
## Sweetness       0.033445    0.009582    3.491 0.000501 ***
## wet            -0.032852    0.010198   -3.221 0.001313 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1484 on 1112 degrees of freedom
## Multiple R-squared:  0.8073, Adjusted R-squared:  0.8063
## F-statistic: 776.4 on 6 and 1112 DF,  p-value: < 2.2e-16
```

```
# Backward
stepAIC(object = full, scope = list(upper = full, lower = empty),
        direction = "backward", k = log(nrow(coffee)))
```

```
## Start:  AIC=-4209.89
## Flavor ~ Aroma + Aftertaste + Body + Acidity + Balance + Sweetness +
##     Uniformity + Moisture + wet + semi
##
##               Df Sum of Sq    RSS      AIC
## - semi         1    0.0001 24.264 -4216.9
## - Moisture     1    0.0596 24.324 -4214.2
## - Uniformity   1    0.0605 24.325 -4214.1
## - Balance      1    0.0943 24.358 -4212.6
## - Sweetness    1    0.1383 24.402 -4210.5
## <none>                     24.264 -4209.9
## - wet          1    0.1970 24.461 -4207.9
## - Body         1    0.3418 24.606 -4201.3
## - Acidity      1    2.2904 26.554 -4116.0
## - Aroma        1    2.5422 26.806 -4105.4
## - Aftertaste   1    8.4155 32.679 -3883.7
##
## Step:  AIC=-4216.9
## Flavor ~ Aroma + Aftertaste + Body + Acidity + Balance + Sweetness +
##     Uniformity + Moisture + wet
##
##               Df Sum of Sq    RSS      AIC
## - Moisture     1    0.0596 24.324 -4221.2
## - Uniformity   1    0.0605 24.325 -4221.1
## - Balance      1    0.0942 24.358 -4219.6
## - Sweetness    1    0.1383 24.402 -4217.6
## <none>                     24.264 -4216.9
## - wet          1    0.2263 24.490 -4213.5
## - Body         1    0.3426 24.607 -4208.2
## - Acidity      1    2.2905 26.555 -4123.0
## - Aroma        1    2.5426 26.807 -4112.4
## - Aftertaste   1    8.4155 32.680 -3890.7
##
## Step:  AIC=-4221.18
## Flavor ~ Aroma + Aftertaste + Body + Acidity + Balance + Sweetness +
##     Uniformity + wet
##
##               Df Sum of Sq    RSS      AIC
## - Uniformity   1    0.0639 24.388 -4225.3
## - Balance      1    0.0795 24.403 -4224.5
```

```
## <none>                      24.324 -4221.2
## - Sweetness   1    0.1553 24.479 -4221.1
## - wet         1    0.2189 24.543 -4218.2
## - Body        1    0.3205 24.644 -4213.5
## - Acidity     1    2.3500 26.674 -4125.0
## - Aroma       1    2.5685 26.892 -4115.9
## - Aftertaste  1    8.3791 32.703 -3897.0
##
## Step:  AIC=-4225.26
## Flavor ~ Aroma + Aftertaste + Body + Acidity + Balance + Sweetness +
##     wet
##
##              Df Sum of Sq    RSS     AIC
## - Balance    1    0.0999 24.488 -4227.7
## <none>                    24.388 -4225.3
## - wet        1    0.2099 24.598 -4222.7
## - Sweetness  1    0.2522 24.640 -4220.8
## - Body       1    0.2905 24.678 -4219.0
## - Acidity    1    2.3543 26.742 -4129.2
## - Aroma      1    2.5711 26.959 -4120.1
## - Aftertaste 1    8.5578 32.945 -3895.7
##
## Step:  AIC=-4227.7
## Flavor ~ Aroma + Aftertaste + Body + Acidity + Sweetness + wet
##
##              Df Sum of Sq    RSS     AIC
## <none>                    24.488 -4227.7
## - wet        1    0.2285 24.716 -4224.3
## - Sweetness  1    0.2683 24.756 -4222.5
## - Body       1    0.4483 24.936 -4214.4
## - Acidity    1    2.5437 27.031 -4124.1
## - Aroma      1    2.7077 27.195 -4117.4
## - Aftertaste 1   11.0308 35.518 -3818.6

##
## Call:
## lm(formula = Flavor ~ Aroma + Aftertaste + Body + Acidity + Sweetness +
##     wet, data = coffee)
##
## Coefficients:
## (Intercept)       Aroma    Aftertaste          Body      Acidity     Sweetness
##    -0.60900     0.22597       0.49099       0.10344      0.22564       0.03345
##         wet
##    -0.03285
```

```r
m_b <- stepAIC(object = full, scope = list(upper = full, lower = empty),
               direction = "backward", trace = 0, k = log(nrow(coffee)))
summary(m_b)
```

```
##
## Call:
## lm(formula = Flavor ~ Aroma + Aftertaste + Body + Acidity + Sweetness +
##     wet, data = coffee)
##
## Residuals:
```

```
##       Min      1Q   Median      3Q      Max
## -0.65627 -0.08781  0.00032  0.08529  0.63010
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.609003   0.159910  -3.808 0.000148 ***
## Aroma        0.225969   0.020378  11.089  < 2e-16 ***
## Aftertaste   0.490988   0.021938  22.381  < 2e-16 ***
## Body         0.103438   0.022926   4.512 7.11e-06 ***
## Acidity      0.225638   0.020994  10.748  < 2e-16 ***
## Sweetness    0.033445   0.009582   3.491 0.000501 ***
## wet         -0.032852   0.010198  -3.221 0.001313 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1484 on 1112 degrees of freedom
## Multiple R-squared:  0.8073, Adjusted R-squared:  0.8063
## F-statistic: 776.4 on 6 and 1112 DF,  p-value: < 2.2e-16
```

```r
# Forward-backward
stepAIC(object = empty, scope = list(upper = full, lower = empty),
        direction = "both", k = log(nrow(coffee)))
```

```
## Start:  AIC=-2427.29
## Flavor ~ 1
##
##               Df Sum of Sq     RSS      AIC
## + Aftertaste   1    93.607  33.465 -3913.3
## + Acidity      1    69.294  57.778 -3302.2
## + Aroma        1    68.457  58.615 -3286.1
## + Balance      1    68.173  58.899 -3280.7
## + Body         1    58.232  68.840 -3106.2
## + Uniformity   1     5.778 121.294 -2472.3
## + wet          1     2.313 124.759 -2440.8
## + Sweetness    1     2.300 124.772 -2440.7
## + Moisture     1     2.239 124.833 -2440.2
## <none>                     127.072 -2427.3
## + semi         1     0.331 126.741 -2423.2
##
## Step:  AIC=-3913.29
## Flavor ~ Aftertaste
##
##               Df Sum of Sq     RSS      AIC
## + Acidity      1     4.995  28.470 -4087.2
## + Aroma        1     4.908  28.557 -4083.7
## + Body         1     2.255  31.210 -3984.3
## + Balance      1     1.737  31.729 -3965.9
## <none>                      33.465 -3913.3
## + Sweetness    1     0.138  33.327 -3910.9
## + Uniformity   1     0.114  33.351 -3910.1
## + wet          1     0.087  33.378 -3909.2
## + Moisture     1     0.038  33.427 -3907.5
## + semi         1     0.018  33.448 -3906.9
## - Aftertaste   1    93.607 127.072 -2427.3
##
```

```
## Step:  AIC=-4087.17
## Flavor ~ Aftertaste + Acidity
##
##               Df Sum of Sq    RSS     AIC
## + Aroma        1    3.0217 25.448 -4205.7
## + Body         1    0.8956 27.574 -4115.9
## + Balance      1    0.6542 27.816 -4106.2
## + wet          1    0.2256 28.244 -4089.0
## <none>                      28.470 -4087.2
## + Sweetness    1    0.1709 28.299 -4086.9
## + Uniformity   1    0.1143 28.356 -4084.6
## + semi         1    0.0445 28.425 -4081.9
## + Moisture     1    0.0199 28.450 -4080.9
## - Acidity      1    4.9955 33.465 -3913.3
## - Aftertaste   1   29.3075 57.778 -3302.2
##
## Step:  AIC=-4205.7
## Flavor ~ Aftertaste + Acidity + Aroma
##
##               Df Sum of Sq    RSS     AIC
## + Body         1    0.5090 24.939 -4221.3
## + Balance      1    0.3257 25.123 -4213.1
## + wet          1    0.2689 25.179 -4210.6
## + Sweetness    1    0.1901 25.258 -4207.1
## <none>                      25.448 -4205.7
## + Uniformity   1    0.1141 25.334 -4203.7
## + semi         1    0.0417 25.407 -4200.5
## + Moisture     1    0.0195 25.429 -4199.5
## - Aroma        1    3.0217 28.470 -4087.2
## - Acidity      1    3.1089 28.557 -4083.7
## - Aftertaste   1   15.5890 41.037 -3678.0
##
## Step:  AIC=-4221.29
## Flavor ~ Aftertaste + Acidity + Aroma + Body
##
##               Df Sum of Sq    RSS     AIC
## + Sweetness    1    0.2233 24.716 -4224.3
## + wet          1    0.1835 24.756 -4222.5
## + Uniformity   1    0.1723 24.767 -4222.0
## <none>                      24.939 -4221.3
## + Balance      1    0.1325 24.807 -4220.2
## + Moisture     1    0.0585 24.881 -4216.9
## + semi         1    0.0404 24.899 -4216.1
## - Body         1    0.5090 25.448 -4205.7
## - Acidity      1    2.4119 27.351 -4125.0
## - Aroma        1    2.6351 27.574 -4115.9
## - Aftertaste   1   11.9225 36.862 -3791.1
##
## Step:  AIC=-4224.33
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness
##
##               Df Sum of Sq    RSS     AIC
## + wet          1    0.2285 24.488 -4227.7
## <none>                      24.716 -4224.3
```

```
## + Balance       1     0.1186 24.598 -4222.7
## - Sweetness      1     0.2233 24.939 -4221.3
## + Uniformity     1     0.0755 24.641 -4220.7
## + Moisture       1     0.0386 24.677 -4219.1
## + semi           1     0.0377 24.678 -4219.0
## - Body           1     0.5422 25.258 -4207.1
## - Acidity        1     2.4163 27.132 -4127.0
## - Aroma          1     2.6440 27.360 -4117.6
## - Aftertaste     1    11.4557 36.172 -3805.2
##
## Step:  AIC=-4227.7
## Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness + wet
##
##               Df Sum of Sq    RSS     AIC
## <none>                     24.488 -4227.7
## + Balance       1     0.0999 24.388 -4225.3
## + Uniformity    1     0.0843 24.403 -4224.5
## - wet           1     0.2285 24.716 -4224.3
## + Moisture      1     0.0467 24.441 -4222.8
## - Sweetness     1     0.2683 24.756 -4222.5
## + semi          1     0.0002 24.487 -4220.7
## - Body          1     0.4483 24.936 -4214.4
## - Acidity       1     2.5437 27.031 -4124.1
## - Aroma         1     2.7077 27.195 -4117.4
## - Aftertaste    1    11.0308 35.518 -3818.6
##
##
## Call:
## lm(formula = Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness +
##     wet, data = coffee)
##
## Coefficients:
## (Intercept)    Aftertaste      Acidity        Aroma         Body     Sweetness
##    -0.60900       0.49099      0.22564      0.22597      0.10344       0.03345
##         wet
##    -0.03285
```

```r
m_h <- stepAIC(object = empty, scope = list(upper = full, lower = empty),
               direction = "both", trace = 0, k = log(nrow(coffee)))
summary(m_h)
```

```
##
## Call:
## lm(formula = Flavor ~ Aftertaste + Acidity + Aroma + Body + Sweetness +
##     wet, data = coffee)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.65627 -0.08781  0.00032  0.08529  0.63010
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.609003   0.159910  -3.808 0.000148 ***
## Aftertaste   0.490988   0.021938  22.381  < 2e-16 ***
## Acidity      0.225638   0.020994  10.748  < 2e-16 ***
```

```
## Aroma         0.225969   0.020378  11.089  < 2e-16 ***
## Body          0.103438   0.022926   4.512 7.11e-06 ***
## Sweetness      0.033445   0.009582   3.491 0.000501 ***
## wet           -0.032852   0.010198  -3.221 0.001313 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1484 on 1112 degrees of freedom
## Multiple R-squared:  0.8073, Adjusted R-squared:  0.8063
## F-statistic: 776.4 on 6 and 1112 DF,  p-value: < 2.2e-16
```

```
# 10 variables is still a fairly small problem:  in this example
# all 3 approaches identify the same BIC-based model as the exhaustive search.
```