Task 1 - 'Faraway' package in R contains 'fat' dataset that has more than 13 variables. We are running the best subset regression with exhaustive search taking 'body fat' as the response variable and considering 'BIC' criteria to select the subset size.

```
library(faraway)
library(leaps)

head(fat)

##    brozek siri density age weight height adipos  free neck chest abdom
hip
## 1   12.6 12.3  1.0708  23 154.25  67.75   23.7 134.9 36.2  93.1  85.2
94.5
## 2    6.9  6.1  1.0853  22 173.25  72.25   23.4 161.3 38.5  93.6  83.0
98.7
## 3   24.6 25.3  1.0414  22 154.00  66.25   24.7 116.0 34.0  95.8  87.9
99.2
## 4   10.9 10.4  1.0751  26 184.75  72.25   24.9 164.7 37.4 101.8  86.4
101.2
## 5   27.8 28.7  1.0340  24 184.25  71.25   25.6 133.1 34.4  97.3 100.0
101.9
## 6   20.6 20.9  1.0502  24 210.25  74.75   26.5 167.0 39.0 104.5  94.4
107.8
##    thigh knee ankle biceps forearm wrist
## 1  59.0 37.3  21.9   32.0    27.4  17.1
## 2  58.7 37.3  23.4   30.5    28.9  18.2
## 3  59.6 38.9  24.0   28.8    25.2  16.6
## 4  60.1 37.3  22.8   32.4    29.4  18.2
## 5  63.2 42.2  24.0   32.2    27.7  17.7 ##
6  66.0 42.0  25.6   35.7    30.6  18.8


model_1 = regsubsets(siri ~
age+weight+height+neck+chest+abdom+hip+thigh+knee+ankle+biceps+forearm+wrist,
data=fat, nvmax=12)

result_1 = summary(model_1)

result_1

## Subset selection object
## Call: regsubsets.formula(siri ~ age + weight + height + neck + chest +
##     abdom + hip + thigh + knee + ankle + biceps + forearm + wrist,
```

```
##    data = fat, nvmax = 12)
## 13 Variables  (and intercept)
##         Forced in Forced out
## age         FALSE      FALSE
## weight      FALSE      FALSE
## height      FALSE      FALSE
## neck        FALSE      FALSE
## chest       FALSE      FALSE
## abdom       FALSE      FALSE
## hip         FALSE      FALSE
## thigh       FALSE      FALSE
## knee        FALSE      FALSE
## ankle       FALSE      FALSE
## biceps      FALSE      FALSE
## forearm     FALSE      FALSE
## wrist       FALSE      FALSE
```

```
## 1 subsets of each size up to 12
## Selection Algorithm: exhaustive
##           age weight height neck chest abdom hip thigh knee ankle biceps
## 1  ( 1 )  " " " "    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 2  ( 1 )  " " "*"    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 3  ( 1 )  " " "*"    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 4  ( 1 )  " " "*"    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 5  ( 1 )  " " "*"    " "    "*"  " "   "*"   " " " "   " "  " "   " "
## 6  ( 1 )  "*" "*"    " "    " "  " "   "*"   " " "*"   " "  " "   " "
## 7  ( 1 )  "*" "*"    " "    "*"  " "   "*"   " " "*"   " "  " "   " "
## 8  ( 1 )  "*" "*"    " "    "*"  " "   "*"   "*" "*"   " "  " "   " "
## 9  ( 1 )  "*" "*"    " "    "*"  " "   "*"   "*" "*"   " "  " "   "*"
## 10 ( 1 )  "*" "*"    " "    "*"  " "   "*"   "*" "*"   " "  "*"   "*"
## 11 ( 1 )  "*" "*"    "*"    "*"  " "   "*"   "*" "*"   " "  "*"   "*"    ##
12 ( 1 ) "*" "*"    "*"    "*"  "*"   "*"   "*" "*"   " "  "*"   "*"    ##
forearm wrist
## 1  ( 1 )  " "     " "
## 2  ( 1 )  " "     " "
## 3  ( 1 )  " "     "*"
## 4  ( 1 )  "*"     "*"
## 5  ( 1 )  "*"     "*"
## 6  ( 1 )  "*"     "*"
## 7  ( 1 )  "*"     "*"
## 8  ( 1 )  "*"     "*"
## 9  ( 1 )  "*"     "*"
## 10 ( 1 )  "*"     "*"
```

```
## 11  ( 1 ) "*"      "*"    ##
12  ( 1 ) "*"      "*"
result_1$bic
```

```
## [1] -262.0435 -303.1197 -305.7338 -307.0259 -304.2743 -301.6966 -299.6035
## [8] -296.1315 -291.7763 -287.0028 -281.9683 -276.5036
```

```
which.min(result_1$bic)
```

```
## [1] 4
```

Using the BIC criteria, we should go for the model with 4 variables

```
which(result_1$which[which.min(result_1$bic), ])
```

```
## (Intercept)        weight        abdom      forearm        wrist
##            1             3            7           13           14
```

```
coef(model_1, which.min(result_1$bic))
```

```
## (Intercept)        weight        abdom      forearm        wrist
## -34.8540743   -0.1356315    0.9957513    0.4729284   -1.5055620
```

Task 2 – Considering the same dataset, we are running the best subset regression via forward stepwise selection and considering 'BIC' criteria to select the model.

```
model_2 = regsubsets(siri ~
age+weight+height+neck+chest+abdom+hip+thigh+knee+ankle+biceps+forearm+wrist,
data=fat, nvmax=12, method = "forward")

result_2 = summary(model_2)

result_2


## Subset selection object
## Call: regsubsets.formula(siri ~ age + weight + height + neck + chest +
##     abdom + hip + thigh + knee + ankle + biceps + forearm + wrist,
##     data = fat, nvmax = 12, method = "forward")
## 13 Variables  (and intercept)
##           Forced in Forced out
## age           FALSE      FALSE
## weight        FALSE      FALSE
## height        FALSE      FALSE
## neck          FALSE      FALSE
## chest         FALSE      FALSE
## abdom         FALSE      FALSE
## hip           FALSE      FALSE
## thigh         FALSE      FALSE
## knee          FALSE      FALSE
## ankle         FALSE      FALSE
## biceps        FALSE      FALSE
## forearm       FALSE      FALSE
## wrist         FALSE      FALSE
## 1 subsets of each size up to 12
## Selection Algorithm: forward
```

```
##           age weight height neck chest abdom hip thigh knee ankle biceps
## 1  ( 1 )  " " " "    " "    " " " "   "*"   " " " "   " "  " "   " "
## 2  ( 1 )  " " "*"    " "    " " " "   "*"   " " " "   " "  " "   " "
## 3  ( 1 )  " " "*"    " "    " " " "   "*"   " " " "   " "  " "   " "
## 4  ( 1 )  " " "*"    " "    " " " "   "*"   " " " "   " "  " "   " "
## 5  ( 1 )  " " "*"    " "    "*" " "   "*"   " " " "   " "  " "   " "
## 6  ( 1 )  "*" "*"    " "    "*" " "   "*"   " " " "   " "  " "   " "
## 7  ( 1 )  "*" "*"    " "    "*" " "   "*"   " " "*"   " "  " "   " "
## 8  ( 1 )  "*" "*"    " "    "*" " "   "*"   "*" "*"   " "  " "   " "
## 9  ( 1 )  "*" "*"    " "    "*" " "   "*"   "*" "*"   " "  " "   "*"
## 10 ( 1 )  "*" "*"    " "    "*" " "   "*"   "*" "*"   " "  "*"   "*"
## 11 ( 1 )  "*" "*"    "*"    "*" " "   "*"   "*" "*"   " "  "*"   "*"    ##
12 ( 1 ) "*" "*"    "*"    "*" "*"   "*"   "*" "*"   " "  "*"   "*"    ##
forearm wrist
## 1  ( 1 )  " "    " "
## 2  ( 1 )  " "    " "
## 3  ( 1 )  " "    "*"
## 4  ( 1 )  "*"    "*"
## 5  ( 1 )  "*"    "*"
## 6  ( 1 )  "*"    "*"
## 7  ( 1 )  "*"    "*"
## 8  ( 1 )  "*"    "*"
## 9  ( 1 )  "*"    "*"
## 10 ( 1 ) "*"    "*"
## 11 ( 1 ) "*"    "*"    ##
12 ( 1 ) "*"    "*"
```

result_2$bic

```
##  [1] -262.0435 -303.1197 -305.7338 -307.0259 -304.2743 -301.3805 -299.6035
##  [8] -296.1315 -291.7763 -287.0028 -281.9683 -276.5036
```

which.min(result_2$bic)

```
## [1] 4
```

which(result_2$which[which.min(result_2$bic),])

```
## (Intercept)      weight       abdom     forearm       wrist
##           1           3           7          13          14
```

```
coef(model_2, which.min(result_2$bic))

## (Intercept)       weight        abdom      forearm        wrist
## -34.8540743   -0.1356315    0.9957513    0.4729284   -1.5055620
```

Task 3 – Considering the same dataset, we are running the best subset regression via backward stepwise selection and considering 'BIC' criteria to select the model.

```
model_3 = regsubsets(siri ~
age+weight+height+neck+chest+abdom+hip+thigh+knee+ankle+biceps+forearm+wrist,
data=fat, nvmax=12, method = "backward")

result_3 = summary(model_3)

result_3


## Subset selection object
## Call: regsubsets.formula(siri ~ age + weight + height + neck + chest +
##      abdom + hip + thigh + knee + ankle + biceps + forearm + wrist,
##      data = fat, nvmax = 12, method = "backward")
## 13 Variables  (and intercept)
##          Forced in Forced out
## age          FALSE      FALSE
## weight       FALSE      FALSE
## height       FALSE      FALSE
## neck         FALSE      FALSE
## chest        FALSE      FALSE
## abdom        FALSE      FALSE
## hip          FALSE      FALSE
## thigh        FALSE      FALSE
## knee         FALSE      FALSE
## ankle        FALSE      FALSE
## biceps       FALSE      FALSE
## forearm      FALSE      FALSE
## wrist        FALSE      FALSE
```

```
## 1 subsets of each size up to 12
## Selection Algorithm: backward
##           age weight height neck chest abdom hip thigh knee ankle biceps
## 1  ( 1 )  " " " "    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 2  ( 1 )  " " "*"    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 3  ( 1 )  " " "*"    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 4  ( 1 )  " " "*"    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 5  ( 1 )  "*" "*"    " "    " "  " "   "*"   " " " "   " "  " "   " "
## 6  ( 1 )  "*" "*"    " "    " "  " "   "*"   " " "*"   " "  " "   " "
## 7  ( 1 )  "*" "*"    " "    "*"  " "   "*"   " " "*"   " "  " "   " "
## 8  ( 1 )  "*" "*"    " "    "*"  " "   "*"   "*" "*"   " "  " "   " "
## 9  ( 1 )  "*" "*"    " "    "*"  " "   "*"   "*" "*"   " "  " "   "*"
## 10 ( 1 )  "*" "*"    " "    "*"  " "   "*"   "*" "*"   " "  "*"   "*"
## 11 ( 1 )  "*" "*"    "*"    "*"  " "   "*"   "*" "*"   " "  "*"   "*"
## 12 ( 1 )  "*" "*"    "*"    "*"  "*"   "*"   "*" "*"   " "  "*"   "*"
##           forearm wrist
## 1  ( 1 )  " "     " "
## 2  ( 1 )  " "     " "
```

```
## 3  ( 1 )  " "     "*"
## 4  ( 1 )  "*"     "*"
## 5  ( 1 )  "*"     "*"
## 6  ( 1 )  "*"     "*"
## 7  ( 1 )  "*"     "*"
## 8  ( 1 )  "*"     "*"
## 9  ( 1 )  "*"     "*"
## 10 ( 1 )  "*"     "*"
## 11 ( 1 )  "*"     "*"   ##
12 ( 1 ) "*"     "*"
```

```
result_3$bic
```

```
## [1] -262.0435 -303.1197 -305.7338 -307.0259 -303.5804 -301.6966 -299.6035
## [8] -296.1315 -291.7763 -287.0028 -281.9683 -276.5036
```

```
which.min(result_3$bic)
```

```
## [1] 4
```

```
which(result_3$which[which.min(result_3$bic),])
```

```
## (Intercept)      weight       abdom     forearm       wrist
##           1           3           7          13          14
```

```
coef(model_3, which.min(result_3$bic))

## (Intercept)      weight       abdom      forearm       wrist
## -34.8540743  -0.1356315   0.9957513   0.4729284  -1.5055620
```