# Random Forest Exercise 1

## Part A: Interpreting the Software Output

Run the Rscript RF Heart1.R.

The Random Forest confusion matrix on Heart.csv data was presented:

```
Call:
 randomForest(formula = AHD ~ ., data = heart.df, importance = T,      na.action = na.omit)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 3

        OOB estimate of  error rate: 18.18%
Confusion matrix:
     No Yes class.error
No  136  24   0.1500000
Yes  30 107   0.2189781
```

An internet search shows that some researchers thought the rows of the matrix represent actuals while other researchers thought the columns of the matrix represent actuals.

1. Verify if the rows represent actuals or columns represent actuals.
2. Explain the meaning of the class.error.

## Part B: Checking the Default Parameter Settings.

3. Execute Random Forest with different settings on Heart.csv to predict AHD and save the respective OOB error in a table. How does the default settings for B and RSF size perform in terms of OOB error? Recommend a way to use/set the B and RSF size in your future work.
   - B = 25, RSF size = 1
   - B = 25, RSF size = floor(sqrt(M)) = 3
   - B = 25, RSF size = M = 13
   - B = 100, RSF size = 1
   - B = 100, RSF size = floor(sqrt(M)) = 3
   - B = 100, RSF size = M = 13
   - B = 500, RSF size = 1
   - B = 500, RSF size = floor(sqrt(M)) = 3
   - B = 500, RSF size = M = 13