**DATA ANALYTICS ASSIGNMENT: 4**

Problem statement: The pseudo code of the algorithm you used to produce your samples and split them in training and test samples.

Solution:

Step1:

We import libraries: Mas, Lattice, Caret, gglpot

Step 2:

We use mu and sigma to store the values of mean vector and covariance matrix respectively.

Step 3:

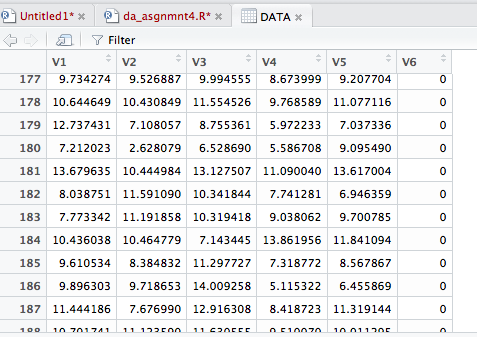
We use a function like the following:

X1 <- MASS:: mvrnorm (n = 500, mu, Sigma,tol = 1e-06, empirical = FALSE)

It generates 500 random numbers and store them to a matrix X1.

Step 4:

We add a new column for target to X1 matrix and assign it value either 0 or 1.



Step 5:

We add both the two matrices generated to a new matrix named DATA.

STEP 6:

Now, we split the data in the ratio 80:20 and store it as train and test respectively.

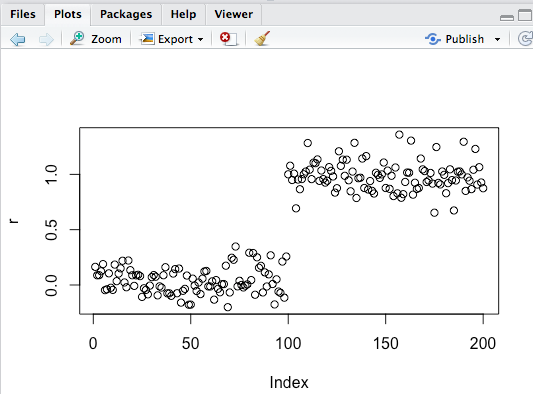
Step 7:

We need to create data frames for storing matrix data.

We fit our model using the training data, and apply it on the test data.

Step 8:

Now we find the accuracy using linear regression.



Step 9:

The classification technique of choice used: Confusion Matrix

Step 10:

The accuracy attained using both the models is 100%.

