Module 6: Case Study 1 – Multiple Logistic Regression

Problem Statement:

Sam's next exam is on '**Multiple Logistic Regression**'. Questions would be asked on the basis of what you've learnt in the respective module

Tasks to be performed:

1. Build a multiple logistic regression model:
   a. Start off by dividing the data-set into 'train' & 'test' sets in 65:35 ratio, with the splitcriteria being determined by 'gender' column
   b. Build a logistic regression model on the train set where the dependent variable is 'gender' & the independent variables are 'Dependents', 'InternetService' & 'Contract' & store the result in 'log_mod_multi'
   c. Predict the values on top of the test set & store the result in 'result_log_multi'
   d. Have a look at the range of 'result_log_multi' & build a confusion matrix where the threshold of predicted values is greater than '0.49'
   e. Calculate the accuracy of the model from the confusion matrix

2. Build second logistic regression model on the same 'train' & 'test' sets
   a. In this case dependent variable is 'gender' & the independent variables are 'tenure', 'MonthlyCharges' & 'PaymentMethod
   b. Predict the values on top of the test set & store the result in 'result_log_multi2'
   c. Have a look at the range of 'result_log_multi2' & build a confusion matrix where the threshold of predicted values is greater than 0.49
   d. Calculate the accuracy of the model from the confusion matrix