**Group\_Submission**

QWE\_Case Study

Use random seed 123 to split the data into training and test data (split Ratio =0.8)

Q.1 Is Wall's belief about the dependence of churn rates on customer age supported by data? To get some intuition, try visualizing the dependence.

#Q.2 Build the appropriate logistic regression model.

# a) Is the model overall significant?

#b) What factors are significant in the model?

#c) Assess the performance of the model by measuring AUC.

#d) Find the threshold probability for the problem in order to maximize accuracy

Q.3 Build a model with bagging approach (5000 trees) and comment on the accuracy of the model.

Q.4 Build a random forest (5000 trees) and comment on the accuracy of the model.