**APPROACH**

* A customer data of 3333 customers with 21 variables was given.
* The data is divided into two parts train data and test data
* Train data is used to construct regression model and test data is used to test the accuracy of the model
* First model was constructed using all the variables. the model took each state of the column as a single variable
* In model2 column state was removed
* Yet there are lot of insignificant variables in the data
* Step function was used to identify the significant variables and model 3 was formed
* As the train data contained unequal number of 1s and 0s
* Model 4 was constructed by removing imbalance
* Accuracy measures and performance trade off was carried out using model 4

**Results**

* outcome4

0 1

1. 451 117
2. 27 72

* Accuracy = 0.7841079
* Sensitivity = 0.7841079
* Specificity = 0.7272727
* AUC value = 0.8211703
* ROC CURVE

