CAPSTONE PROJECT

ABG MOTORS IN INDIA

By : Akshat Chauhan

PROBLEM STATEMENT :

- This Project is based on a car company ‘ABG Motors’, a fictional Japan-based car motors company .
- ABG Motors believes that the Indian market is very similar to Japanese, in which the company currently operates.
- Before entering the new market, the company wants to be sure that the whole process will be profitable for them.

Hence, we are given the task to check for the following conditions that must be fulfilled in the Indian market for the company to enter:

- Sale of a minimum of 10,000 Cars over the sample data in one year .

T1: JUSTIFICATION DURING MODEL DEVELOPMENT

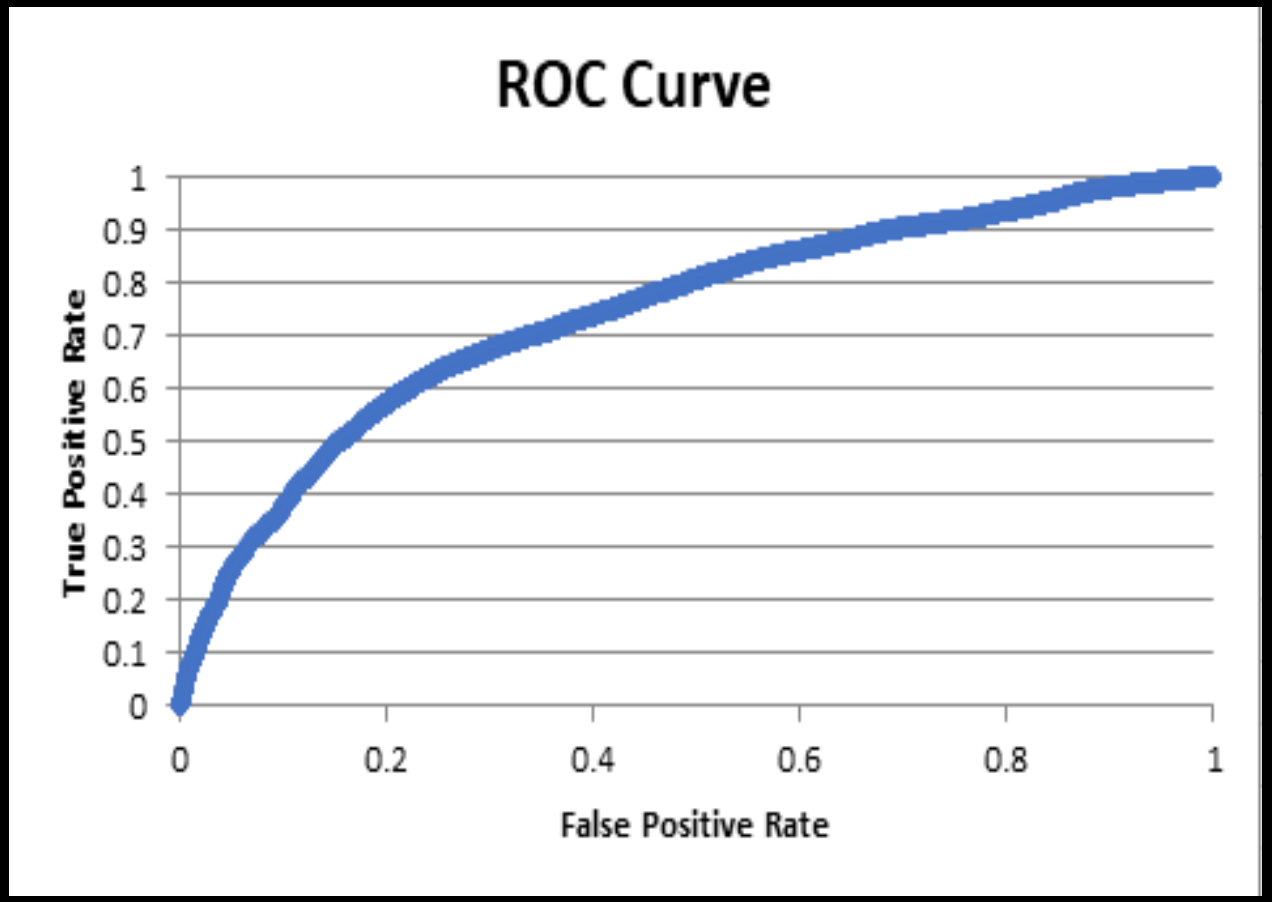
- Gender was classified into binary data as male (1) and female(0). The annual income was converted into INR for matching the situation of Indian currency
- From the Chinese customer data it is clear that purchase decision depends on 4 factors. Customer age, Gender, Car Age, and annual income. The following changes were made on the raw data

Days	Segment
<200	1
200-360	2
360-500	3
>500	4

- Then after dataset is converted into training and test set with 70:30 Rule and Logistic Regression Modelling Is Performed On both of them.
- And then ROC Curve, Beta Values and Conversion Matrix (Including Accuracy, Sensitivity, Precision etc.)Is Computed from train and test data through K-means Clustering

Classification Table			
	Obs Suc	Obs Fail	Total
Pred Suc	15035	9609	24644
Pred Fail	998	2357	3355
Total	16033	11966	27999
Accuracy	0.937753	0.196975	0.621165
Cutoff	0.3		
AUC	0.741296		

Covariance matrix				
0.004793	-4.9E-05	-0.0005	-3.4E-09	-0.00064
-4.9E-05	1.3E-06	-3.7E-08	-3.1E-11	-1.1E-06
-0.0005	-3.7E-08	0.000714	2.43E-10	2.43E-05
-3.4E-09	-3.1E-11	2.43E-10	2.04E-14	2.91E-10
-0.00064	-1.1E-06	2.43E-05	2.91E-10	0.000273



Cut-off	Accuracy
0.3	0.93
0.4	0.89
0.5	0.74

Cutoff	0.3
AUC	0.741296

Covariance matrix

0.004793	-4.9E-05	-0.0005	-3.4E-09	-0.00064
-4.9E-05	1.3E-06	-3.7E-08	-3.1E-11	-1.1E-06
-0.0005	-3.7E-08	0.000714	2.43E-10	2.43E-05
-3.4E-09	-3.1E-11	2.43E-10	2.04E-14	2.91E-10
-0.00064	-1.1E-06	2.43E-05	2.91E-10	0.000273

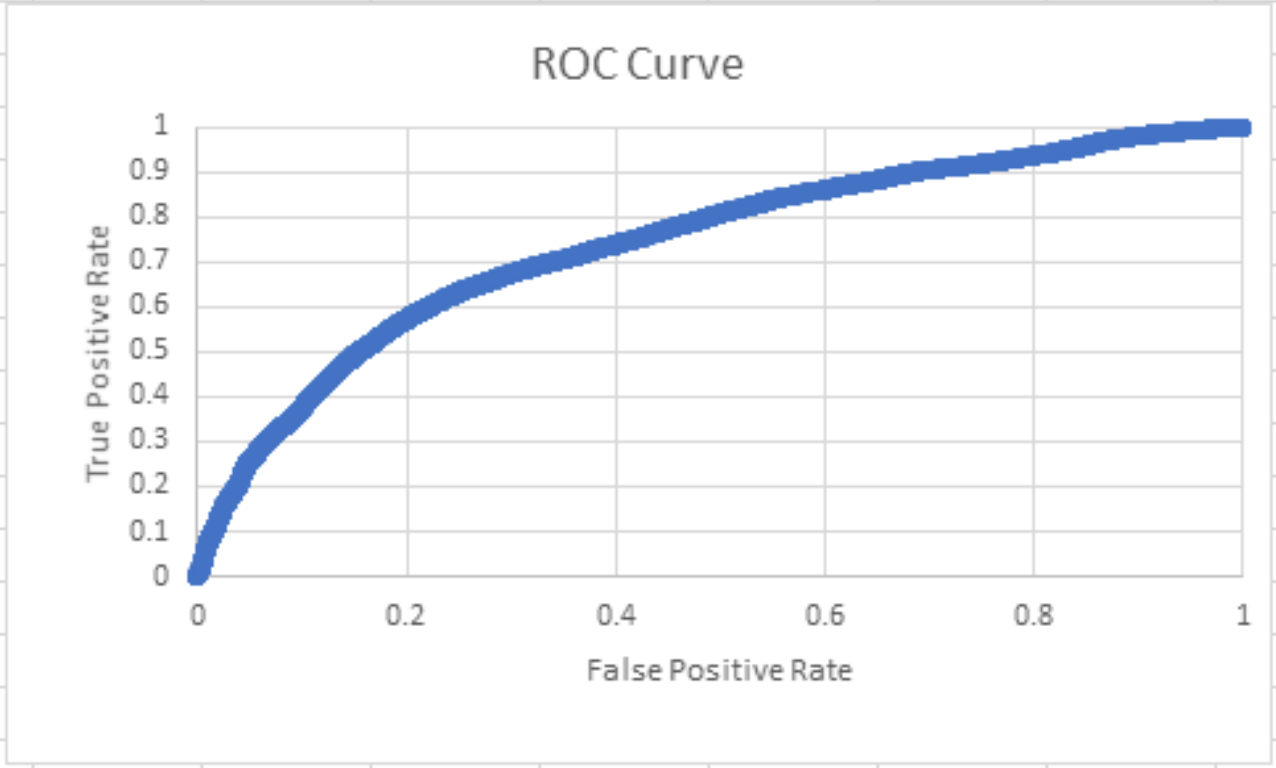
Converge

7.89E-13
4.22E-11
4.75E-13
2.24E-07
2.06E-12

Classification Table

	Obs Suc	Obs Fail	Total
Pred Suc	14290	7967	22257
Pred Fail	1743	3999	5742
Total	16033	11966	27999
Accuracy	0.891287	0.334197	0.653202

Cutoff	0.4
AUC	0.741296



T1: PIVOT TABLES

Testing Data

	Column Labels		
	0	1	Grand Total
Sum of AGE_CAR_CATEGORY	1475	27335	28810

Training Data

	Column Labels		
	0	1	Grand Total
Sum of AGE_CAR_CATEGORY	23961	43065	67026

	Column Labels		
	0	1	Grand Total
Sum of PURCHASE	443	6555	6998

CLASSIFICATION MODEL BASED ON JAPANESE DATASET

- The logistic regression was made on the formatted data by considering the factors like age, gender, income, car life and purchase.
- The coefficients B0 to B4 and ROC Curve of training Data are computed as follows:

	coeff
intercept	-2.421454
CURR_AGE	-0.011216
GENDER	0.226001
ANN_INCO	4.08E-06
AGE_CAR_(	0.973804

Metric	Value
Accuracy	62.7
Precision	61.9
Recall	93.6
F1-Score	0.2

		Predicted	
		1	0
Actual	1	6555	443
	0	4033	970

T1: COUNT OF POTENTIAL CUSTOMERS IN INDIA BASED ON MODEL

The data set is formatted such that gender is converted into a binomial model and the phone age is calculated by considering the purchase date as 1st July 2019.

The phone life is converted into 4 categories as follow:

The probability is computed based on coefficients (B0 to B4) obtained from the Japanese Dataset.

Days	Segment
<200	1
200-360	2
360-500	3
>500	4

	Column Lab		
	0	1	Grand Total
Count of Predic	4	20997	21001

	Column Lab				
	1	2	3	4	Grand Total
Sum of CAR_AG	5544	11406	16842	16560	50352

	Column Labels	
	1	Grand Total
Sum of CAR_AGE_CA	117481	117481

	Column Labels				
	1	2	3	4	Grand Total
Count of Predicted	12952	13291	13077	9679	48999

	coeff
intercept	-2.42145357
CURR_AGE	-0.01121648
GENDER	0.22600065
ANN_INCOMI	4.0783E-06
AGE_CAR_CA	0.97380418

JUSTIFICATION FOR FINAL RESULTS

Japan

	Column Labels		
	0	1	Grand Total
Sum of PURCHASE	443	6555	6998

India

	Column Labels				
	1	2	3	4	Grand Total
Count of Predicted	12952	13291	13077	9679	48999

JUSTIFICATION FOR FINAL RESULTS

- From the Table Below we can conclude that Customer Age Segment 3 ,count of purchase is 13077 which is way over our required result that is minimum 10000 cars, in INDIA as compared to other segments so ABG Motors company must target Middle aged customers and so, yes they can enter in INDIAN market as there is no loss for XYZ Mobile Company.