Problem Statement

For any business to flourish, customer satisfaction is the key. A customer doesn't want to continue his relationship with the business if he is not satisfied. The task of identifying these unsatisfied customers is challenging because in general they do not voice out their dissatisfaction but calmly move away!!

Coral bank is in business for quite some time and have seen many customers moving away. The bank wants to take your help in identifying the dissatisfied customers early in their relationship so that they can intervene and take proactive steps to make them happy and retain them.

In this context, you'll work with the data provided by the bank to solve the problem. As the information is financial related, to maintain the confidentiality, the feature names and the values provided are masked. It is up to you how you would want to treat these features.

Data Set

You are provided with two csv files- "data_train.csv" and "data_test.csv". The "data_train.csv" has the target variable (whether the customer is satisfied or not) and the data_test.csv doesn't have a target.

Evaluations

The metric we are interested for this problem is AUC.

Step1: **Visualizations**

Since this forms an important aspect in data science problems, we would evaluate your visualizations. To make it a bit simpler, here are the following questions for which you may develop visualizations. Also, please report if you have observed any insights from the data.

1. Is there any relationship between Var15 and the Target
2. Do you find any relationship between the Num_Var4 anf the Target
3. Understanding the distribution of mortgage. Is it normal. If this variable is somehow transformed (say log transformed) would it be closer to normal distribution. What is the most frequent value of it.

    Report your observations

Step2: **Benchmark AUC**

You are required to submit your predictions on data test through the shiny app provided and check the AUC obtained. The benchmark AUC is 70%. Make sure submission file schema is same as sample provided. Submit your best predictions and relevant code and report through piazza by 7:00 PM