# Case Study

You have been approached by the consumer lending department to help them automate their underwriting process. As part of the exercise, they have gathered loan-level performance data indicating whether or not a given credit defaulted. This is captured in the *y* variable where:

* *1: default* and
* *0: non-default*.

Also included are 6 other variables the department believes will be predictive of default (*X1 – X6)*.

Propose a model-based solution for the department. As part of the proposal, please justify your decision using the following as guidelines:

1. How well should the department expect your model to identify defaults? ***– Accuracy, Confusion matrix, ROC***
2. How confident are you the model will work well in implementation (new data)? ***Test Accuracy,***
3. Briefly explain the effect of each variable in your final model on credit risk. – ***Marginal Effects, Odds Ratio, Coefficients.***
4. Please provide the probabilities of default (probability of class = 1) in a *csv* file as a new column labeled PD.

Finally, document your analysis.

Please feel free to reach out to [andungu@tcfbank.com](mailto:andungu@tcfbank.com) with any questions regarding this exercise.