**Use Case Summary**

**Objective**

Customer attrition or churn, is a *critical phenomenon* in the banking industry that refers to the rate at which customers leave or discontinue their relationship with a particular bank.

The objective of this project is to analyse bank customer data to identify key factors influencing the customer churn and help bank to develop retention strategies.

**Preprocessing and Modelling**

Below are some of the data preprocessing steps performed on the Customer Churn dataset:

After importing the CSV data as a Pandas DataFrame, I performed data cleaning. I checked for missing values and duplications. The dataset is very clean with no missing values and duplications.

Next, I carried out feature engineering. I created a correlation matrix to check for correlation with target variable with features. The correlation matrix indicates that most features have a weak correlation with the target variable, *Exited*. However, *Age* and *Balance* show a moderate correlation with the target, suggesting they may influence customer churn.

I then performed exploratory data analysis (EDA) to examine the distribution of key features including Age, Balance and Creditscore.

I developed a Logistic Regression model. The model achieved an accuracy of **81%** on the training data and **81%** on the test data.

I did validation using the metrics and confusion matrix.

After that I developed an Artificial Neural Network model, The model achieved an accuracy of 79**%** on the training data and **79 %** on the test data.

The Logistic Regression and ANN algorithms predict churn with an accuracy of **81% and 79%** respectively. However, these are not sufficient for effective customer retention as it is very sensitive in the banking sector. Therefore, I aim to develop advanced algorithms to build a model with higher accuracy.