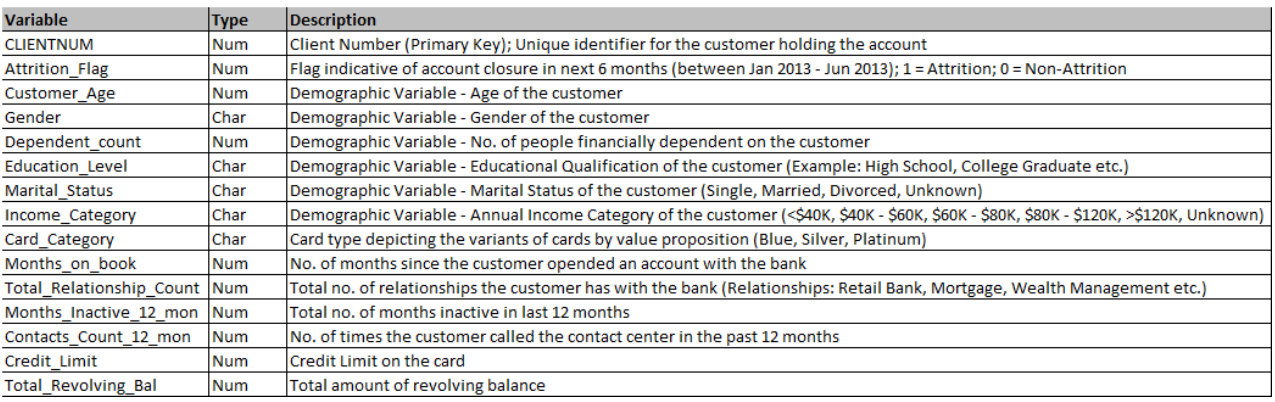
**Credit Card Customer Attrition: Application of Logistic Regression**

**Project Description –**

### Goal and Theory

The goal in this task is to develop a ML model that can determine the probability of attrition of an account holder within the next six months by the use of Logistic Regression. We are given with the banking credit attrition dataset containing over 10000 data. It includes the following information as shown –



**Approach –**

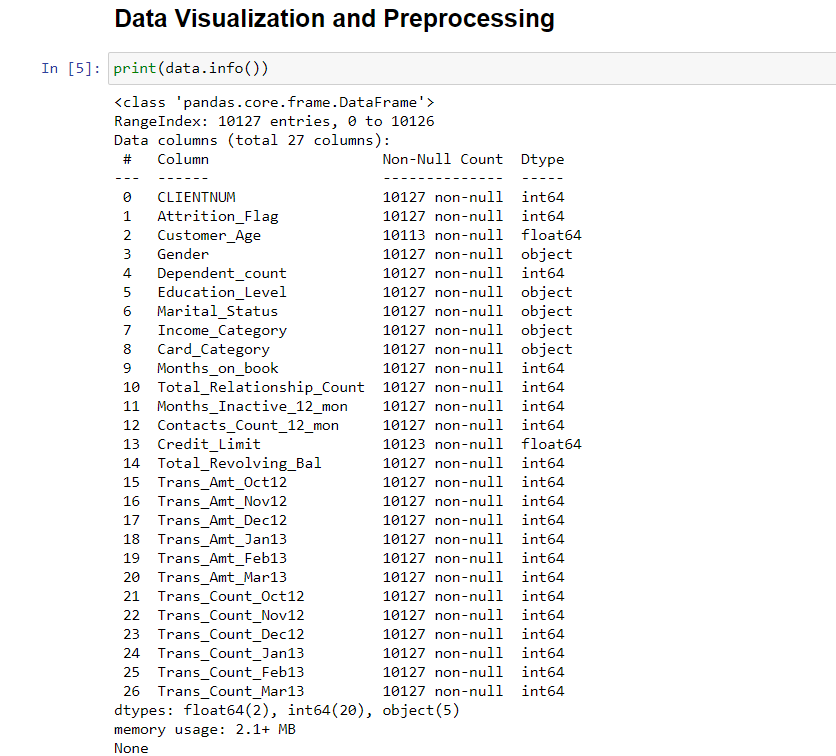
**The steps followed for this include:**

1. **Importing libraries, packages**
2. **Importing the given data in the form of csv files**
3. **Data preprocessing and analysis**
4. **Fitting of the pre processed data in an ML model**
5. **Predicting results using the test dataset**

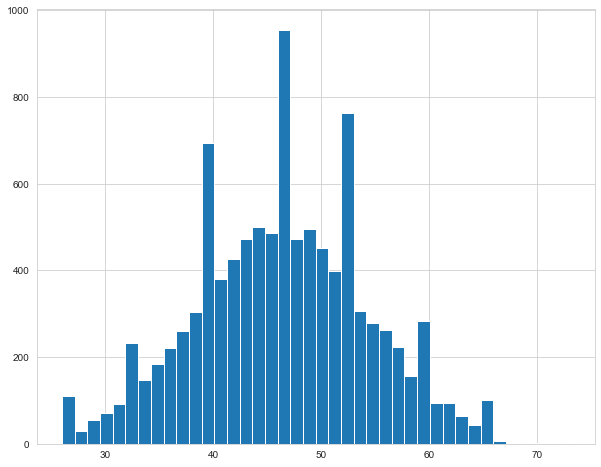
**Based on this, the following plots are obtained**

1. **Data Table**

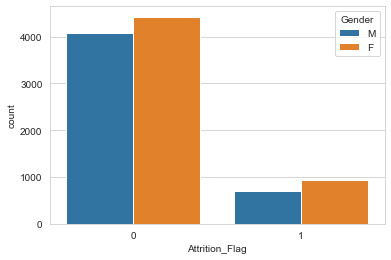
**Two new features are engineering using the past six months data of transaction amount and transaction count, to help reduce the data spread and simply data analysis.**



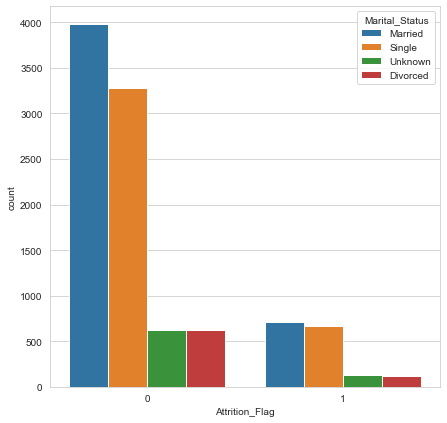
1. **Age Distribution**



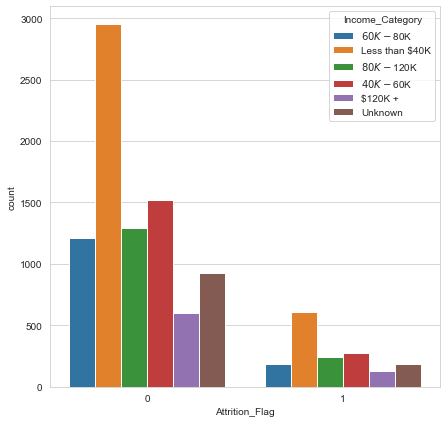
1. **Gender Distribution**



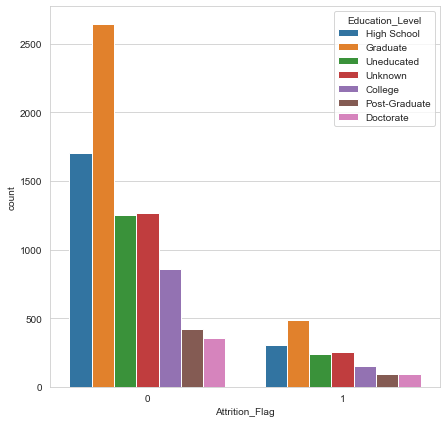
1. **Marital Status**



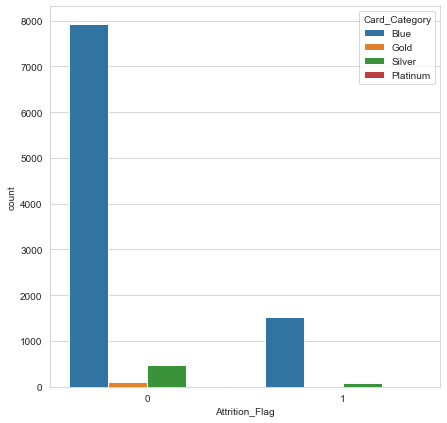
1. Income Category



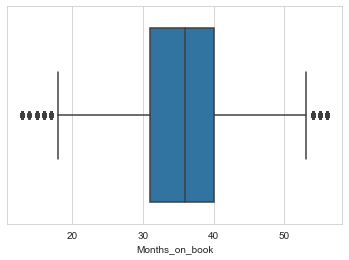
1. Education Level



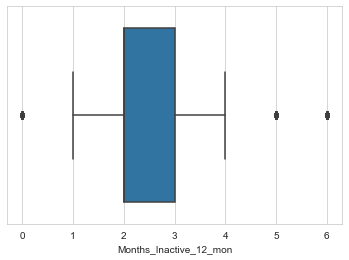
1. Card Category



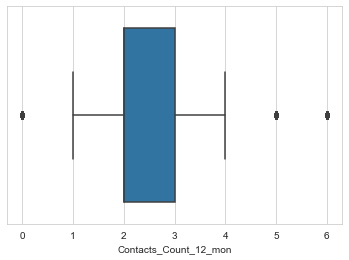
1. Months on Book



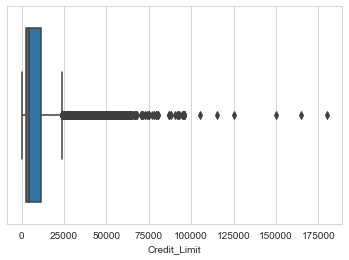
1. Inactive Months



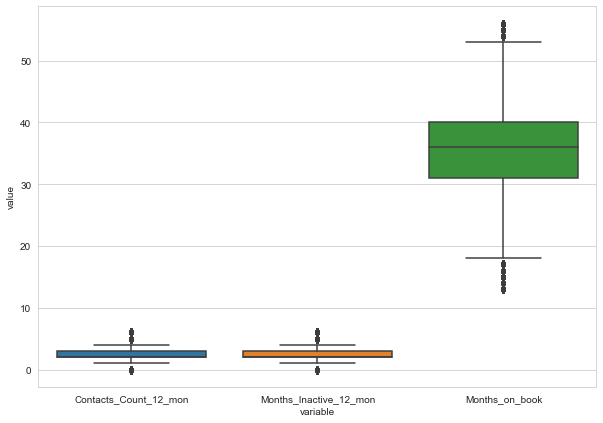
1. Contacts



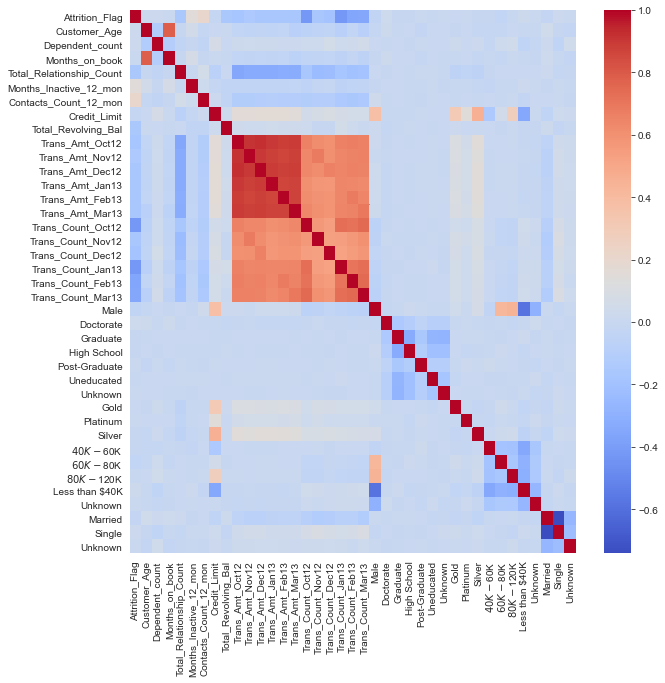
1. Credit Limit



1. Combined Plot

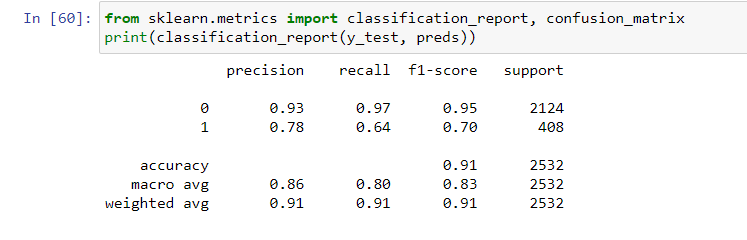


1. Heat Map



**Results –**

**Confusion Matrix scores**



**The accuracy shown in terms of f1-score is 91%**

**Conclusion –**

**The data was preprocessed successfully and all different plots from the dataset were made for better understanding. The data was then fit into the logistic regression model, since this is a classification problem. An accuracy of 91% was obtained for prediction on the basis of training done with the available dataset.**