Project Summary for Customer Churn Analysis

Introduction:

This project is one of the few capstone projects done at Masterschool and involves customer churn analysis for credit card services of a bank. The project goal is to analyse the provided data to understand the main reasons for the churning of customers as well as provide recommendations for how the bank can mitigate that by proactively implementing the recommendations to keep the customers from churning.

The provided data include.

BankChurners.csv file; which contains basic information about each client’s relationship with the bank, basic\_client\_info.csv file which contains some basic demographic info and enriched\_churn\_data.csv file which contains some enriched data about each client.

To begin the task, it was a great idea to have a feel of what the data entails by writing some SQL queries to get some specific info from the dataset. Some SQL queries written answered questions such as: How many clients does the bank have and are above the age of 50? What’s the distribution (in %) between male and female clients? Defining a new variable age\_group based on categories given. Out of the male clients, who holds the “blue” card, how many (in %) hold the income category 40K - 60K? and many more. There was a need to evaluate the dataset to identify quality issues with the data and correct them before further analysis. Python came in handy for this task’s phase as a programmatical assessment of data quality and tidiness issues in a dataset is the best approach using python.