Bank Loan Case Study

BY SYED SARFARAZ AHMED

Description

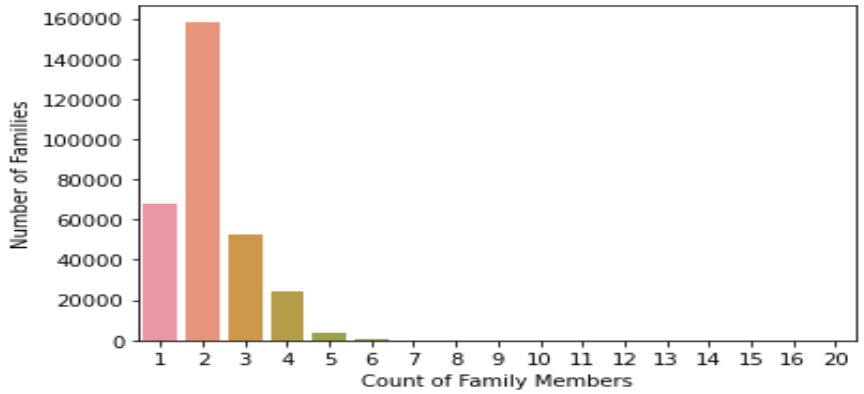
- This case study aims to give you an idea of applying EDA in a real business scenario. In this case study, apart from applying the techniques that you have learnt in the EDA module, you will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers.
- Aim of this project is identify the customers who are capable of repaying loan.

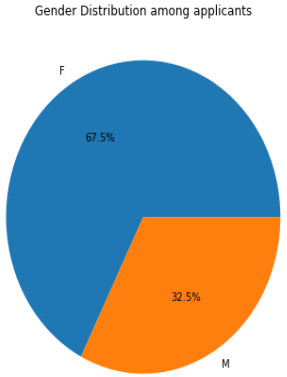
Approach

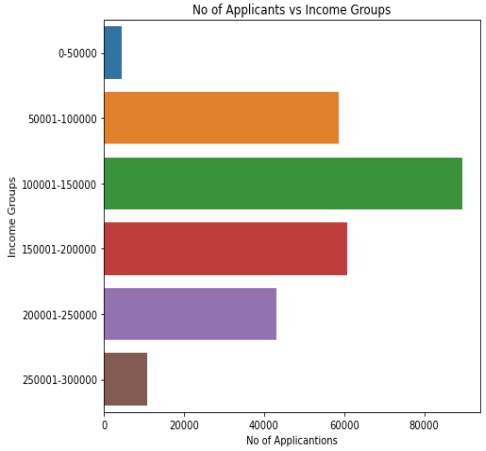
- Firstly import the dataset to your preferred editor to perform analysis (I've used Jupyter notebook as my editor and used python programming language along with it)
- After the above the second major step is to design an efficient procedure which will answer all the asked/given questions about the dataset. Also Jupyter Notebook along with Python Programming Language can be used to create and plot precise and fine Graph/Charts/Visuals for Data.
- Here, first we had to handle the data(i.e: clean, replace inappropriate values and preprocess the data) then after modification of dataset we had to perform necessary methods to get the desired output/result.

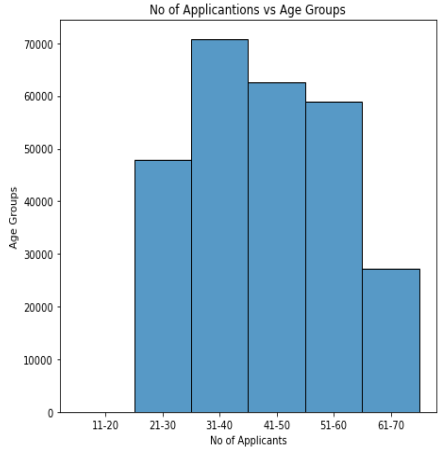
Approach

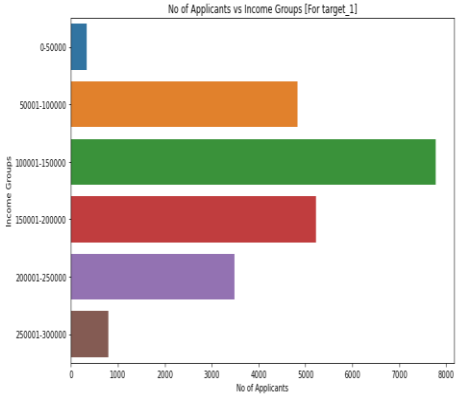
- Below are some of the Charts/Visuals created using Jupyter Notebook with Python Programming for the given Dataset.

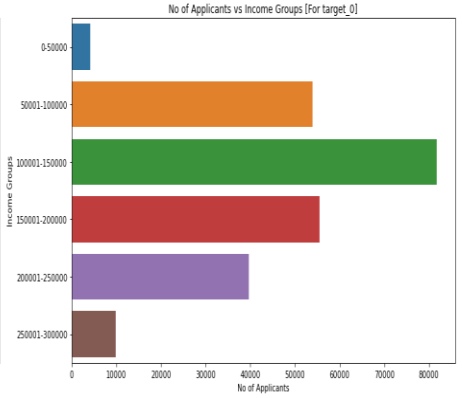


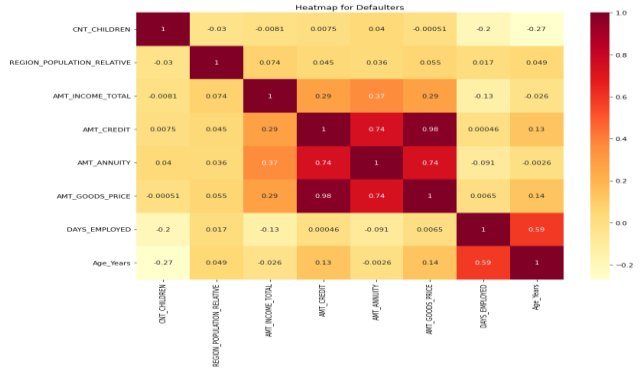


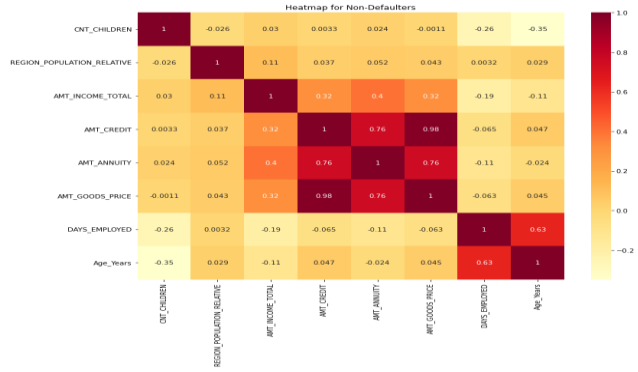












Tech-Stack

- MS PowerPoint - Documentation
- MS Excel - For Getting logic and Viewing Dataset.
- Jupyter Notebook - For loading database and writing the procedure to solve the given problem.
- Python Programming - For writing the procedure ,creating charts/plots.

- Jupyter Notebook Version - 6.4.5
- Python Version - 3.8.2

Insights

There were a number of insights gained while analysis the given dataset some of the insights from the analysis are,

- The percentage of females is higher compared to males in loan applications.
- Higher loan applications have come from applicants with income group 100001-150000.
- People in age group 31-40 have applied the highest number of loans.
- Proportion of females is higher than males in both defaulters and non-defaulters.
- Proportion of both defaulters and non-defaulters not having a car is higher than those who have it.

Result and Drive Link

This project helped me in understanding the use of python programming language and its function to perform data analysis and also helped in discovering various plotting methods and features and draw out insights from those plots.

Drive Link :

https://drive.google.com/drive/folders/1q_pzENxP8P_lL2mhCgJBEJI140eeQv48?usp=sharing