Loan Default Prediction Project | Exploratory Data Analysis (Milestone 3)

Executive Summary Report

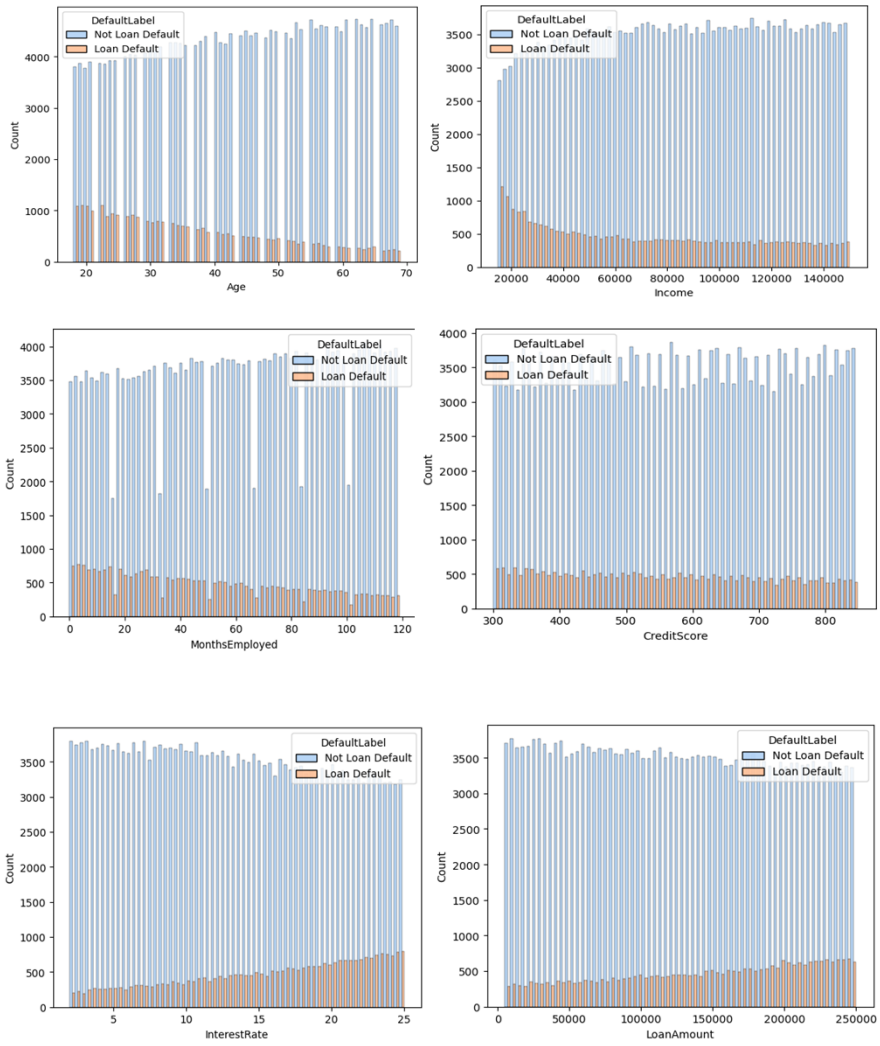
Prepared by: PhD Aleksandar Osmanli

Project Overview

I'm currently developing a data analytics project aimed at decreasing overall loan default among borrowers by predicting which client attributes most contribute their loan default. Thorough exploratory data analysis (EDA) enables financial institution to make better decisions about how to proactively target clients likely to loan default, thereby improving loan payment. **This report offers details and key insights from Milestone 3, which impact the future development of the overall project.**

Milestone 3 - EDA results

- **The 'Default' target variable is of object type with only two possible values: loan defaulted (1) or not (0).**
- **The loan default rate among clients is 12/88 which makes the loan default class highly imbalanced.**
- **The loan default rate decrease as the client's age, the client's income, the client's number of months employed, and the client's credit score increase.**
- **The loan default rate increases as the loan interest rate and the loan amount increase.**
- **The above 6 basic features bring the most predictive power to loan default.**
- **I identified that the 'Education' variiable has blank characters and single quotas among its possible values. In order to avoid future problems, I dropped the blank characters and single quotas before I continue with modeling.**



Next steps

➔ **Through Feature Engineering reveal combine features from the basic predictive features which will discover trends and patterns of client's loan default.**

➔ **Deeply explore their impact on the loan default.**