Capstone_Final_Report

This dataset contains information on borrowers of sanctioned loans. The problem presented here is predicting whet

A few questions I initially asked were as follows:

Is there imbalance in the data?

Do all borrowers share common features associated with defaulting?

What loans are being sanctioned for borrowers?

Do borrowers with payment issues have red flags exclusive to them?

What does the socioeconomic status of all borrowers look like? What do the distributions of the income ranges lool

What age groups are applying for loans?

With these questions as my foundation, I began by dropping irrelevant columns, changing date formats, and imputin ationships between borrower features. There were absolutely distinct fingerprints for borrowers with and without pa ry school, were young or middle-aged, and belonged to the middle - lower class, among other features.

The bulk of applicants were middle to lower class and sought out cash loans. A lot of borrowers were in a relationshi n't provide specifics of the loan, (what the loan was needed for) the total amount of credit loaned out was a feature. t these loans were used for bigger purchases such as mortgages or vehicles.

Further Research:

It would be interesting to measure discrimination bias, if any, in a dataset like this. Analysis of this data showed that eir probability to default or not. Instead of building a simple classification model that either predicts if a borrower wi ability of defaulting instead of a yes/no classification. A fluid scale like this would open up important dialogue betwee titution would offer a payment plan that is structured around the borrower's current situation or they'd offer lenienc

Recommendations:

Incentivize borrowers to opt into revolving loans instead of cash loans. This will provide a small safety net incase bor

Cap the amount of credit offered to higher risk individuals to 500K. Higher risk individuals in this case refers to worki

Provide leniency/assisstance to borrowers with lower socio-economic standing. Low income women make up a lot o redit limit would reduce defaultee rates, particularly among low income women.