Loan Risky Defaulters

EDA Analysis

Problem Statement

- Using the loan data set received, EDA is done to understand the risky loan applicants
- This approach will help the company reject such applicants and thus save on credit loss

Analysis

- This approach applies EDA techniques using python
- The dataset is cleaned and handled with imputed values for certain cases
- Analysis is performed on load status, public record and delinquent record.
- Few insights are drawn based on last payment date
- Overall loan analysis grant is showcased

Univariate analysis – Loan amount

- Total number of applicants around 37k.
- With IQR starting from 5k to 15k, the mean around 10k.

Bivariate analysis

- Around 93% loan applicants had fully paid loan status history
- Few outlier data having loan amount more than 30k had mostly Fully paid loan status applicants
- Home ownership data showed most of the loan applicants were mortgage owners
- Last payment date analysis show stark rise of loan amount in mid 2008.
- Higher delinquent, inquiry for more than 3 times and public record bankruptcies for more than 1 were rejected

Conclusion

- Overall 6.27% of credit loss was saved
- Below plot shows ranges for loan status, loan amount, applied and load status.

```
In [169]: sns.heatmap(res)
          plt.show()
```