LOAN EVALUATION

BIG IDEA

How might we use ML to identify patterns and predictors to more efficiently predict a loan applicants credit worthiness, overcoming the information asymmetry within peer to peer (P2P) loan markets.

THE PROBLEM

Investors face many challenges in predicting the likelihood of repayment due to the lack of borrower information in P2P markets

IMPACT

- Improved consumer confidence and trust
- Eliminate Grey metrics
- Risk mitigation
- Operational Efficiency

SAVINGS

- Global personal loans forecasted to be 620 Billion in 2032
- A 5% increase in accuracy for an average loan amount of $10,000 over 100 loans results in $50,000 in savings.

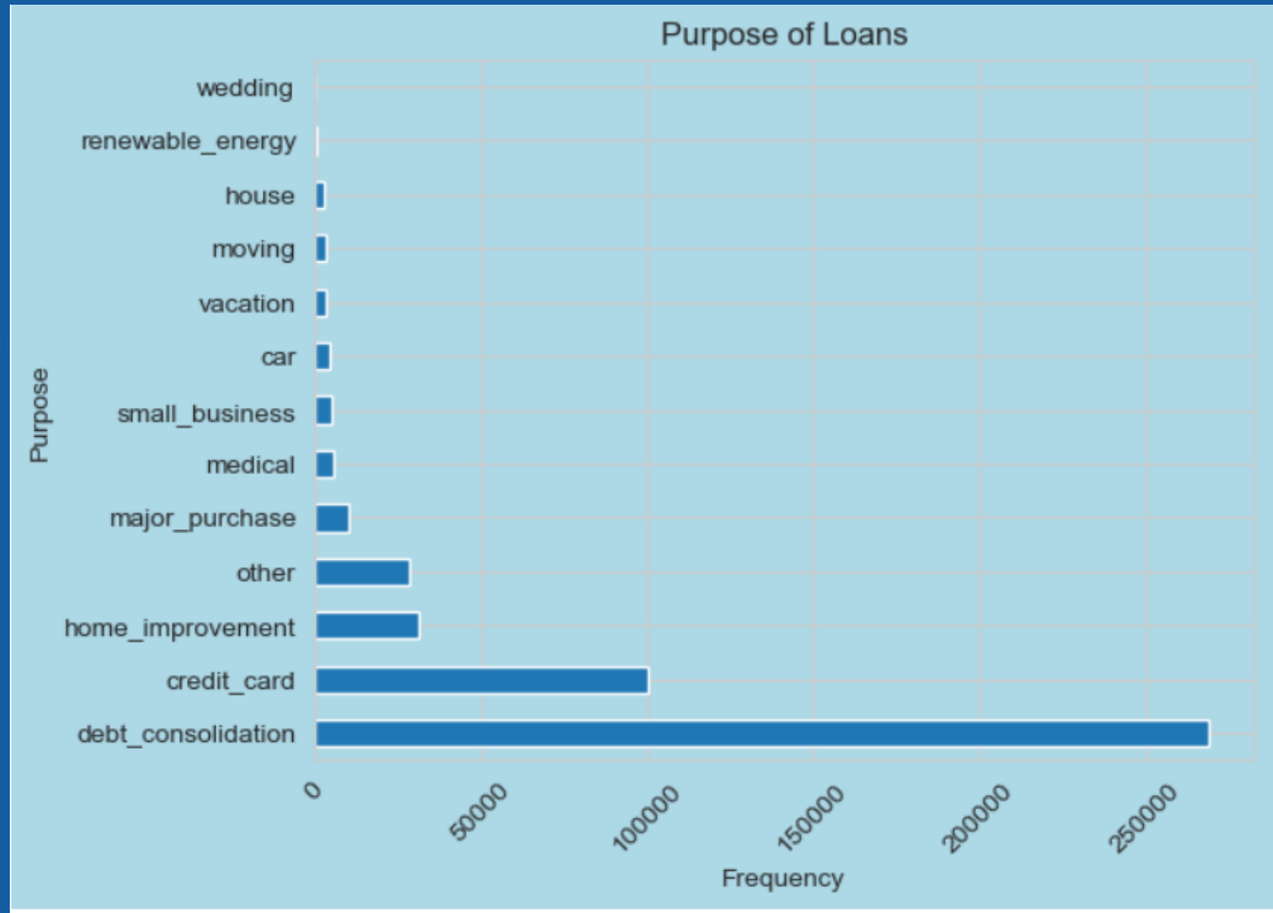
LendingClub

THE DATASET

Real-World Data-set from a US based P2P lending platform, spanning 11 years

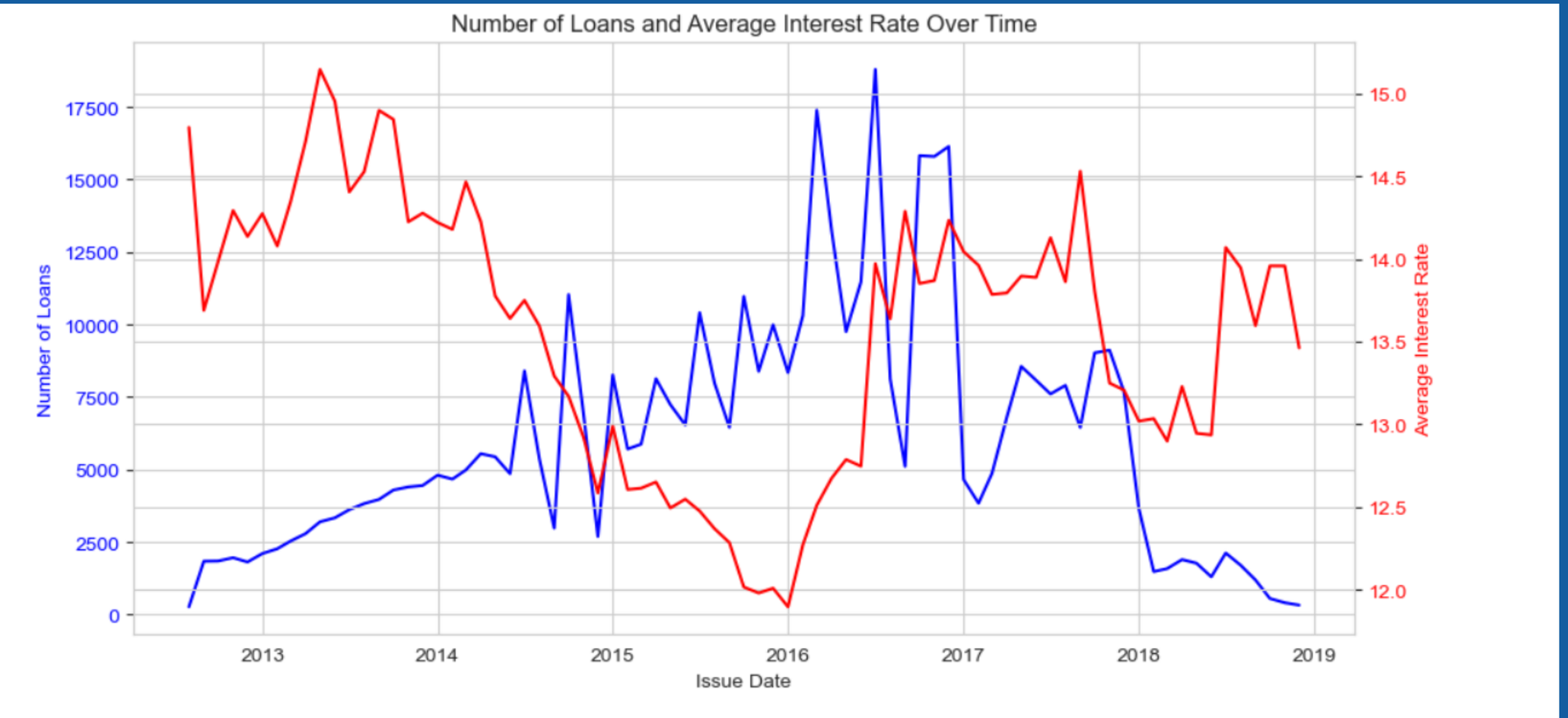
PreProcessing Steps

- Data Cleaning
- Finding which features leak information
- debt to income metric
- avoid PCA for now



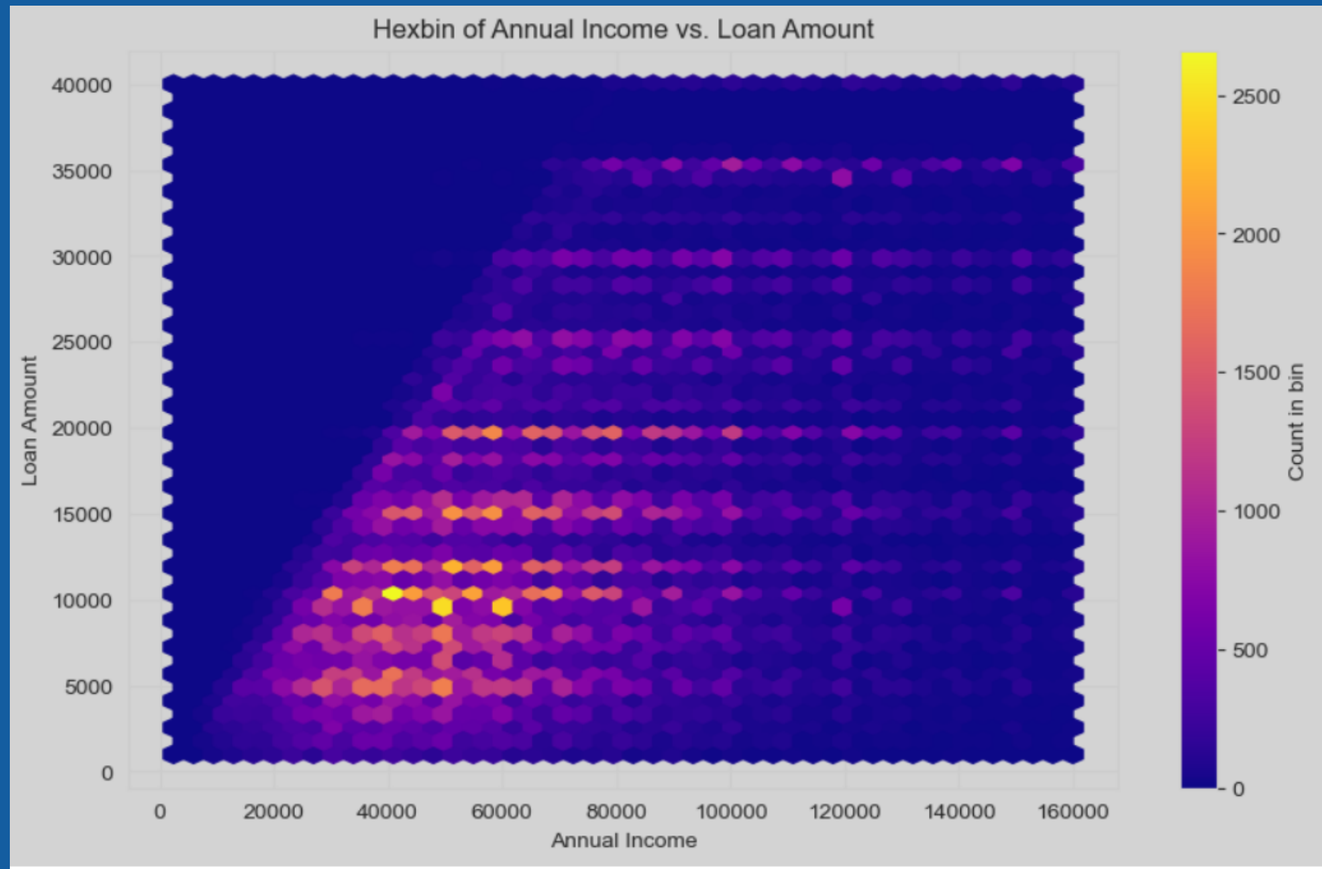
FINDING

- Sensitive to economic data
- Correlation



BASELINE MODEL

- Logistic Regression
- train accuracy of 66%
- test accuracy 65%
- recall of 66%



NEXT STEPS

Optimize the Log Reg model

Expand on Feature Engineering

Fit more complex model