# Loan Payback Prediction - P2P Lending

By Gareth Loh

#### What is P2P Lending?

- Cuts out financial institutions as the middle man
- Enables individuals to obtain loans from other individuals
- Easier for borrowers to access loans
- Lenders able to generate more interest income than conventional savings accounts

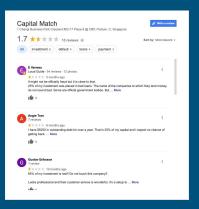
#### Problem Statement

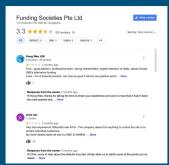
Liquid funds for instruments to earn returns

P2P Lending

Help investors assess and predict if the loan will be paid back

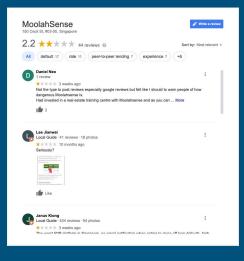
#### So who is this for...?





Investors P2P
Lending
Sites





#### Data Processing

The Lending Club data from Kaggle

Remove data unknown to investors

Impute FICO median

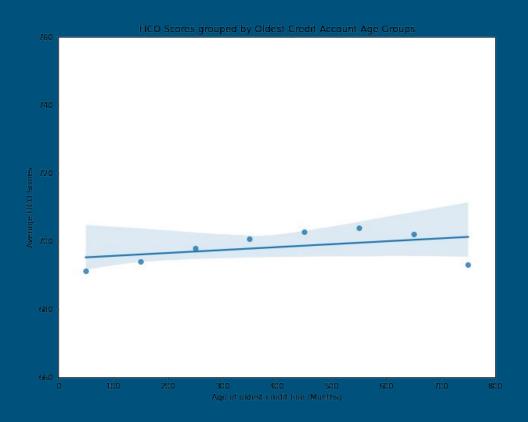
Convert employment length

Consolidate income and DTI columns

Update null values

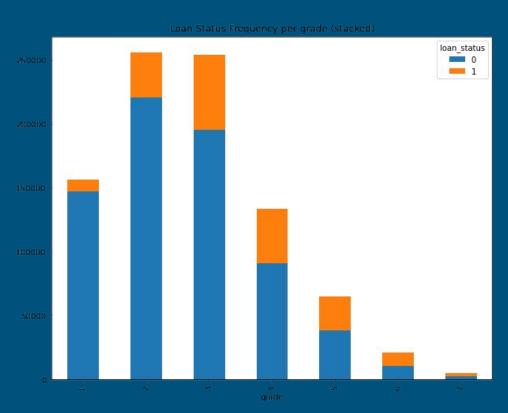
# **Exploratory Data Analysis**

## FICO Scores by Credit Account Age

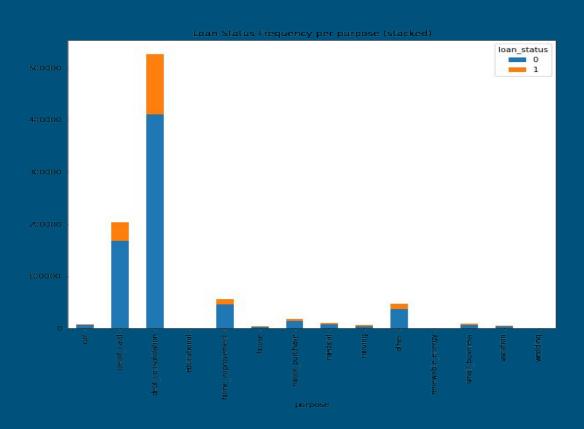


- FICO scores factor in the age of the oldest credit account
- Median FICO score increases with the age

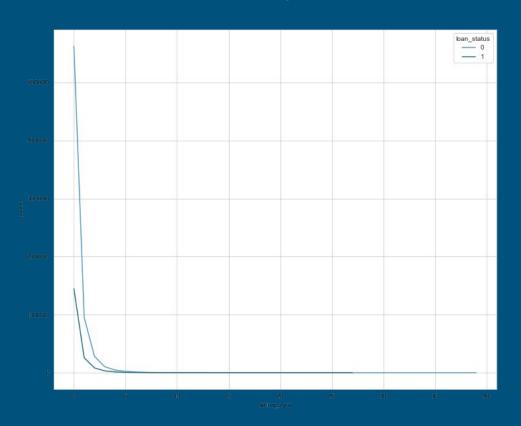
## Lending Club Grades to Loan Payback



## Loan Purpose to Loan Payback



## Delinquencies in the last 2 years to Loan Payback



## Modeling

## Modeling - Classifiers

Logistic Regression

Stochastic Gradient Descent

Random Forest

**XG Boost** 

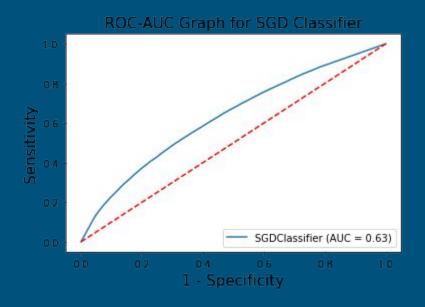
#### Models Used

#### **Models**

- Logistic regression
- Stochastic Gradient Descent
- Random forest
- XG Boost

## Modeling - Results





#### Conclusions and Recommendations

- Tends to predict loans that will be paid off
- Lower accuracy on loans that might be defaulted

#### Recommendations

- Take more current data to train the model
- Better feature engineering