Predicting Loans with Lending Club
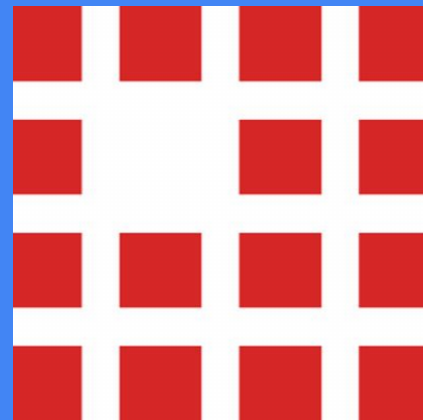
Travis, Catherine, Will, and Kaushik

LendingClub

-World's largest peer-to-peer lending platform

-Borrowers input:
- Desired Loan Amount
- Current Debt/Income
- Length at Job
- Etc.

-Public, de-identified datasets!

Monthly Payments

Borrowers

Investors [2]

Invest

The Problem: Will You Be Approved?

-Will you be approved for your loan?

-If so, how high will your interest be?

Check Your Rate

Get a custom rate for your **$2,000** loan in **1 click**

First Name

Last Name

Street Address

The Data: Predicting Loan Approval

- Two datasets with over three million loan records combined:
  - One accepted
  - One rejected

Common Features				
Amount Requested	Debt-to-Income Ratio	State	Employment Length	Reason for Loan

The Data: Predicting Loan Grade

e.g. Interest Rate

- "Accepted" dataset contains over 100 features
- Over 400,000 loan records

- Goal is to predict loan grade among those accepted

- Grades A-F indicate level of risk and are strongly related to the interest rate offered
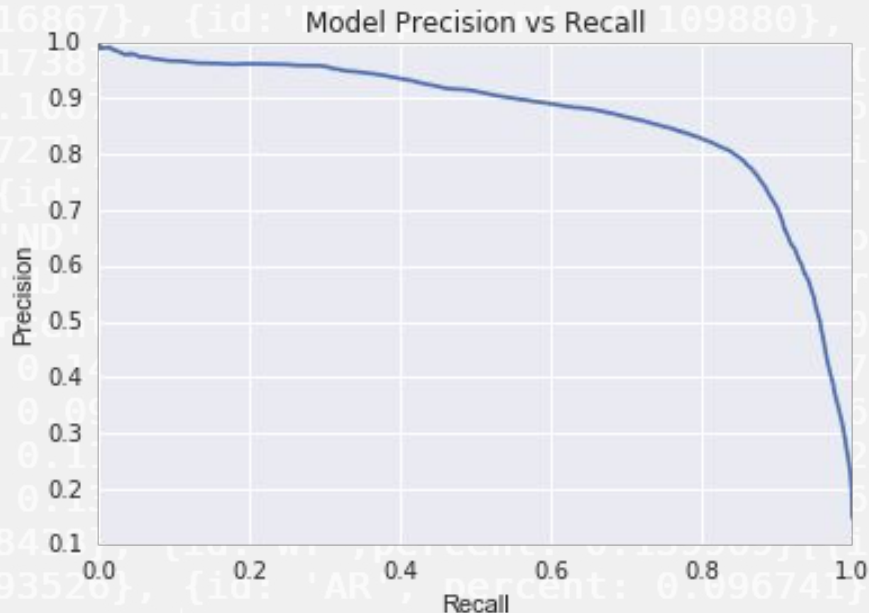
The Models
Accepted/Rejected Predictor

- Extreme gradient boosting

- Precision: 82.62%

- Recall: 80.65%

- 94.94% Accuracy

Model Precision vs Recall

(y-axis: Precision, x-axis: Recall)

The Models

Grade Predictor

- Grade shows a 97% correlation with interest rate

- Used extreme gradient boosting, as with Accept/Reject model

- 56.7% Accuracy when choosing between the 5 grades (random guessing yields ~20% accuracy)

'CT',percent: 0.153628}, {id:'DC',percent: 0.149413}, {id: 'DE ',percent: 0.114722}
percent: 0.116850}, {id:'GA',percent: 0.116867}, {id:'HI',percent: 0.109880}, {
percent: 0.100871}, {id:'LA', percent: 0.100741}, {id:'MA', percent: 0.144365},
percent 0.144065}, {id:'ME',percent: 0.133727}, {id:'MI',percent: 0.128808}, {id:
 0.176245}, {id:'MO', percent: 0.118098} {id:'MS' percent: 0.084202}, {id:'MT',
136899}, {id:'NE',pe
129760}, {id:'NM',percent: 0.150974}, {id:'NJ',percent: 0.148437}, {id:'NM',perce
, {id:'NV',percent: 0.144028}, {id:'NY',percent: 0.133222}, {id:'OH',percent: 0.1
'OK',percent: 0.109462}, {id:'OR',percent: 0.146298}, {id:'PA',percent: 0.125274}
'RI',percent: 0.139231}, {id:'SC',percent: 0.097424}, {id:'SD',percent: 0.120665}
'TN',pe   0.131029}
'VA',percent: 0.132941}, {id:'VI',percent: 0.134176}, {id:'WA',percent: 0.148364}
percent: 0.139969}[{
percent: 0.126727}, {id: 'AL',percent: 0.093526}, {id: 'AR', percent: 0.096741},
percent: 0.141241}, {id:'CA',percent: 0.141041}, {id:'CO',percent: 0.155460},
'CT',percent: 0.153628}, {id:'DC',percent: 0.149413}, {id:'DE ',percent: 0.11472

Takeaways and Future Directions

-More dynamic visualization based on user inputs

- Over time
- By state
- By loan type

-More Accurate Loan Grade Prediction