



BUILDING A LENDING BOT

*Project McNulty Group 4
Will Cosby, Carl van Haren, Derek Updegraff, Naoya Kanai*





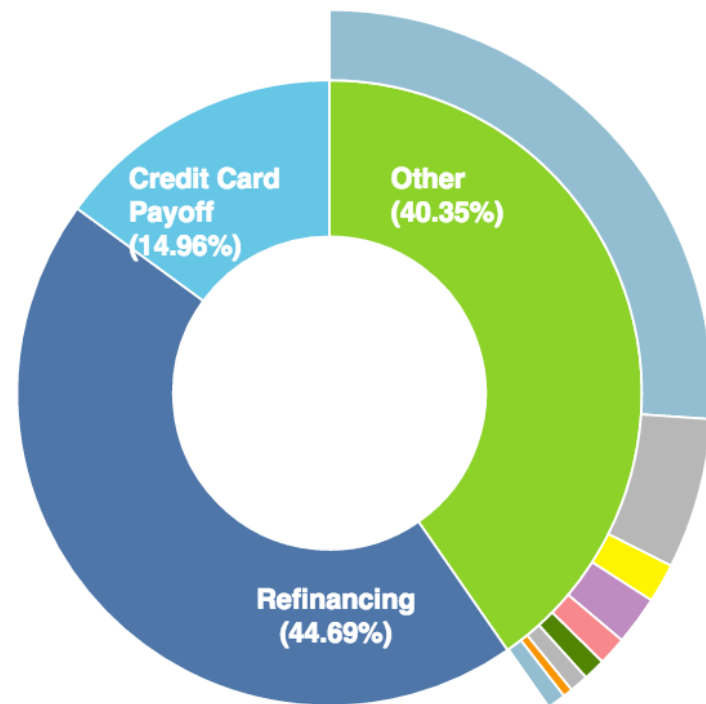
34.79	-3.76	-3.92	92.23	-4.40	-1.38	127.40	-1.10
63.71	-0.83	-1.87	16.82	-1.82	-1.61	134.08	-1.10
4.35	-1.22	-1.90	29.41	-4.07	-1.39	138.08	-1.10
12.77	-0.83	-0.89	29.73	-1.40	-1.67	172.12	-1.10
16.82	-1.77	-3.53	82.16	-2.34	-1.21	187.40	-1.10
29.41	-7.98	-7.98	27.82	-4.92	-1.07	62.12	-1.10
29.73	-6.61	-6.61	34.34	-1.58	-1.39	142.40	-1.10
82.16	-2.99	-2.99	11.35	-1.38	-1.11	128.40	-1.10
27.82	-0.87	-4.67	37.18	-1.46	-1.07	183.40	-1.10
34.34	-1.42	-2.54	2.22	-1.66	-1.43	164.40	-1.10
11.35	-2.54	-0.92	4.00	-4.00	-1.07	150.40	-1.10
37.18	-1.50	-5.12	54.35	-4.73	-1.43	144.40	-1.10
2.22	-1.94	-5.12	12.10	-4.40	-1.40	140.40	-1.10



**WAKE UP
STAND UP
SPEAK UP
RISE UP
TAKE THE POWER BACK !**



WHAT IS LENDING CLUB?



- P2P consumer lending platform (NYSE: LC)
- Provides credit for debt refinancing, credit card payoffs, auto, education, etc.
- Over \$20B funded to date
- Investors can browse and fund loans at 36- and 60-month terms
- Average interest rate of 12.76%

Browse Loans

Summary Portfolio Builder Browse Loans Alert Transfer Trading Account Automated Investing									
Available: \$36.06		Showing Loans 1 - 15 of 463							
Add Funds		Add to Order							
Build a Portfolio									
Per Loan: \$25									
Filter Loans									
Loan Term									
<input checked="" type="checkbox"/> 36-month									
<input checked="" type="checkbox"/> 60-month									
Public Records									
<input type="checkbox"/> Exclude Loans with Public Records									
Location State									
Earliest CREDIT line									
Funding Progress									
		<input type="checkbox"/> Investment	Rate	Term	FICO®	Amount	Purpose	% Funded	Amount / Time Left
		<input type="checkbox"/> \$0	B 3 9.99%	60	685-689	\$12,000	Loan Refinancing & Consolidation	81%	\$2,175 10 days
		<input type="checkbox"/> \$0	B 3 9.99%	60	725-729	\$24,000	Loan Refinancing & Consolidation	83%	\$3,875 10 days
		<input type="checkbox"/> \$0	B 3 9.99%	60	720-724	\$16,600	Loan Refinancing & Consolidation	93%	\$1,100 11 days
		<input type="checkbox"/> \$0	B 2 9.17%	60	680-684	\$15,000	Other	62%	\$5,600 10 days
		<input type="checkbox"/> \$0	A 5 7.89%	60	790-794	\$25,600	Major Purchase	68%	\$8,075 10 days
		<input type="checkbox"/> \$0	B 5 11.53%	36	675-679	\$11,000	Credit Card Payoff	94%	\$650 12 days

A SILICON VALLEY FAIRY TALE...

Why you shouldn't panic about the Lending Club scandal

MARKETS

LendingClub CEO Fired Over Faulty Loans

Renaud Laplanche's ouster is likely to ir

FOUNDERS 40

After Firing Its CEO, Lending Club Is Facing a Crisis

Here's what you need to know about the latest developments in the scandal involving the peer-to-peer lend

BY OSCAR WILLIAMS-GRUT *Business Insider*

LendingClub Rises on Signs of Rebound Even as Scandal Widens

by How LendingClub's Scandal Echoes the Financial Crisis

by Lauren Silva Laughlin

MAY 10, 2016, 2:21 PM EDT

August 9, 2016 12:57 am

Lending Club results reveal pain of governance scandal

Ben McLannahan in New York

JUL 22, 2016 @ 02:41 PM

2,348 VIEWS

The Litt

Lending Club Scandal Provokes Major UK 'Peer-To-Peer' Investigation



Roger Aitken, CONTRIBUTOR

I write about financial markets, exchanges, IT and trading technology.

FULL BIO ✓

DATA OVERVIEW

.....

member_id	loan_amnt	funded_amn	funded_amn	term	int_rate
d	n	n	n	s	s
1296599	5000	5000	4975	36 months	10.6
1314167	2500	2500	2500	60 months	15.2
1313524	2400	2400	2400	36 months	15.9
1277178	10000	10000	10000	36 months	13.4
1311748	3000	3000	3000	60 months	12.6
1311441	5000	5000	5000	36 months	7.9
1304742	7000	7000	7000	60 months	15.9
1288686	3000	3000	3000	36 months	18.6
1306957	5600	5600	5600	60 months	21.2
1306721	5375	5375	5350	60 months	12.6
1305201	6500	6500	6500	60 months	14.6
1305008	12000	12000	12000	36 months	12.6
1298717	9000	9000	9000	36 months	13.4
1304956	3000	3000	3000	36 months	9.9
1303503	10000	10000	10000	36 months	10.6
1304871	1000	1000	1000	36 months	16.2
1299699	10000	10000	10000	36 months	15.2
1304884	3600	3600	3600	36 months	6.0

- Public data on all LC funded loans available via website & Kaggle
- Income
- Demographics
- Past credit events
- Lending Club outcomes

SAMPLE LENDING CLUB LOAN

Title: “HITEK EQUIPMENT”

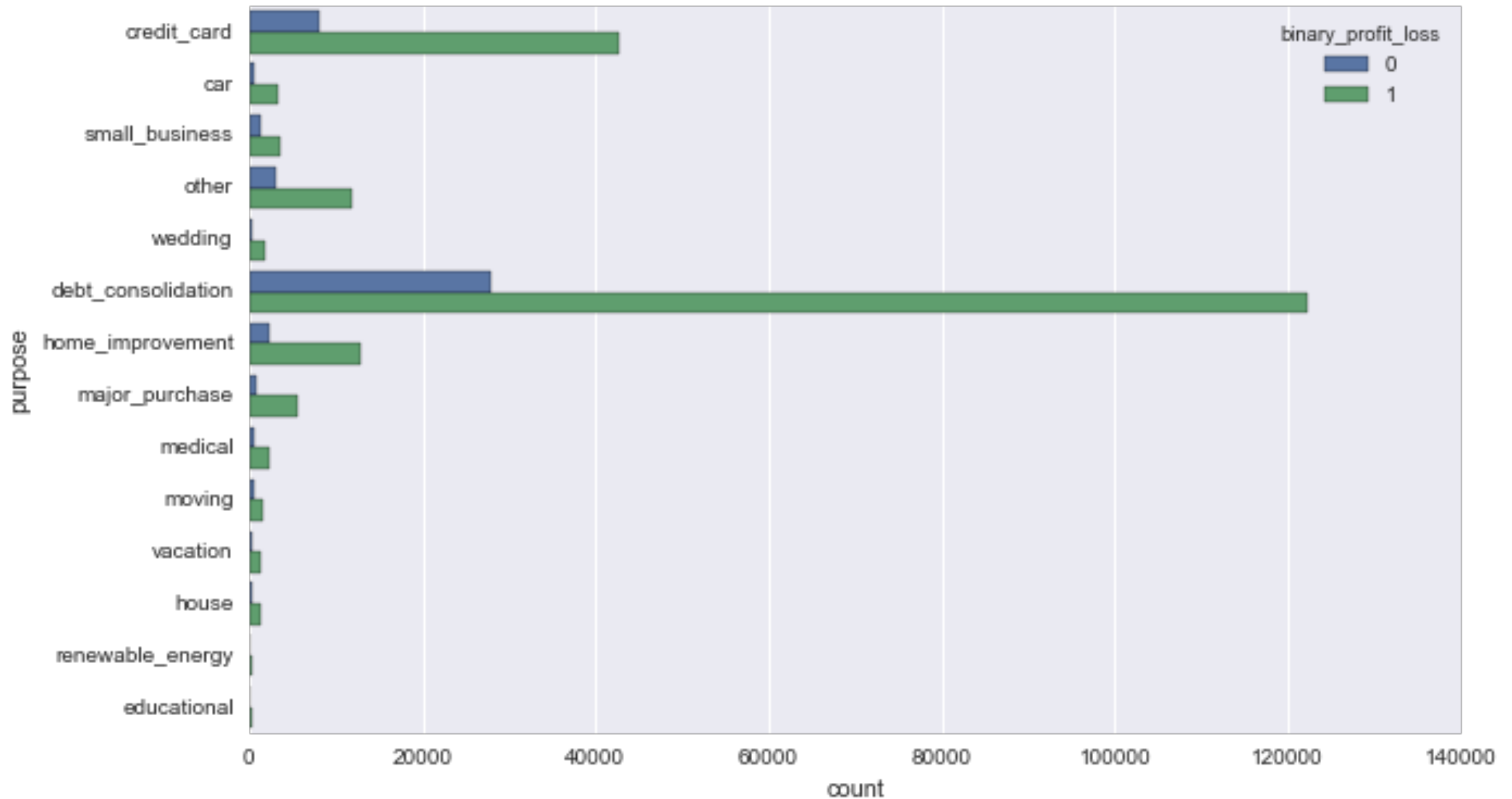
Description:

“YOU ARE WITH ME IN THE MOST IMPORTANT STEP IN MY WAY , SHOP AN EQUIPMENT WHICH FULL FILE MY HOPE IN A NEW LIFE , THANKS. MBM ENTERPRISES OF SOUTH FLORIDA HAVE A FRANCHISE WITH FAST FIX JEWELRY AND WATCH REPPAIR , THEM OWNER 5 LOCATION IN SOUTH FLORIDA -MIAMI , BOCA RATON , OTHER -THERE ARE AT INTERNATIONAL MALL , DADELAND MALL , OTHER , KIOSK AND STORES THAT MAKE SERVICES IN THE REPPAIR WATCHES AND JEWELRY , I WORK IN THIS BUFF LIKE A SALES PERSON AND TECHNICIAN ASSISTANT .
 Borrower added on 06/14/10 > WITH THIS EXPERIENCE , I SAW THE NEED THAT HAVE ALL THE HANDCRAFTER , JEWELLER , WATCHMAKER , CELLULAR PHONE REPAIR TO HAVE TOOLS AND BENCH IN THE LOCATE OF COMMERCE , THERE OWN KIOSK WORKING IN SHOPSTORES , MALLS , STREETS , MY FIRM DESIGN AND BUIL THIS , ERGONOMIC , USEFUL AND ECONOMIC POWERED WITH SOLAR ENERGY AND BEEN ABLE TO LOCATE .”

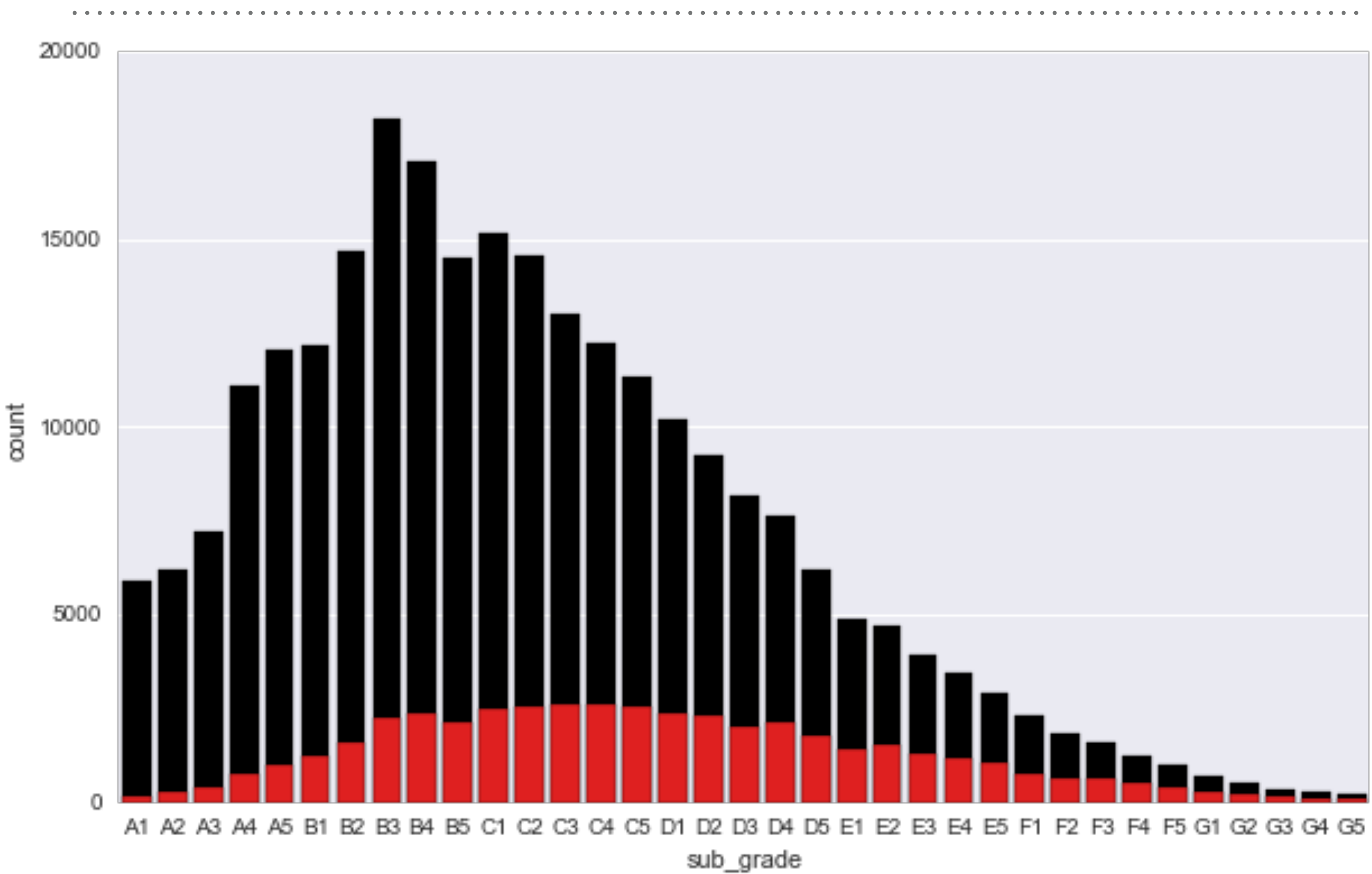
Loan Amount: \$3000.00

Total Payments Received: \$567.60

HOW DO BORROWERS USE LOAN FUNDS?



MODERATE-RISK B & C GRADE LOANS ARE MOST COMMON





MODEL OBJECTIVE

.....

- Classify/predict good & bad loans
- “Good” loan = profitable for the investor, taking into account current interest rates etc.
- “Bad” loan = investor loses money (poor return, default, etc.)



FEATURES & CHARACTERISTICS

.....

➤ Dataset

- $X = 43$ features
- $y = [0, 1]$ (Loan profitability)
- $n = 256,824$

➤ Modeling

- Random forest classifier
- Grid search to tune for number of estimators / max features per node
- Test for overfitting

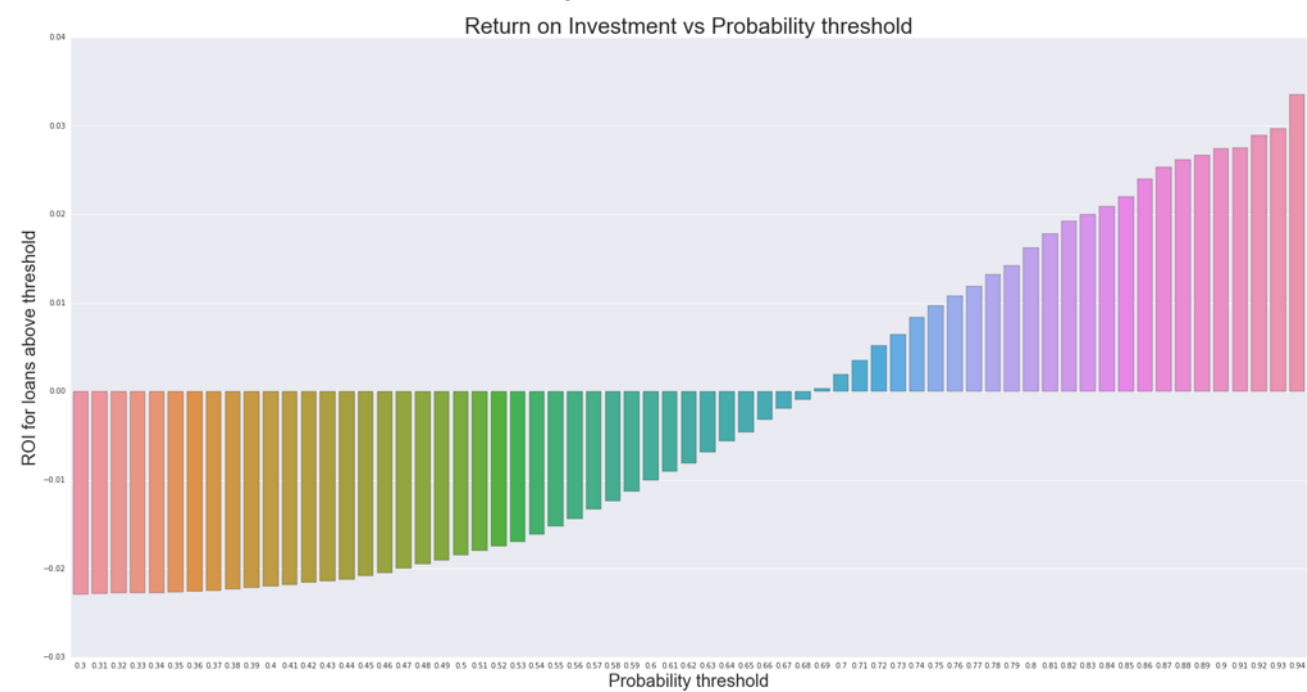


RESULTS

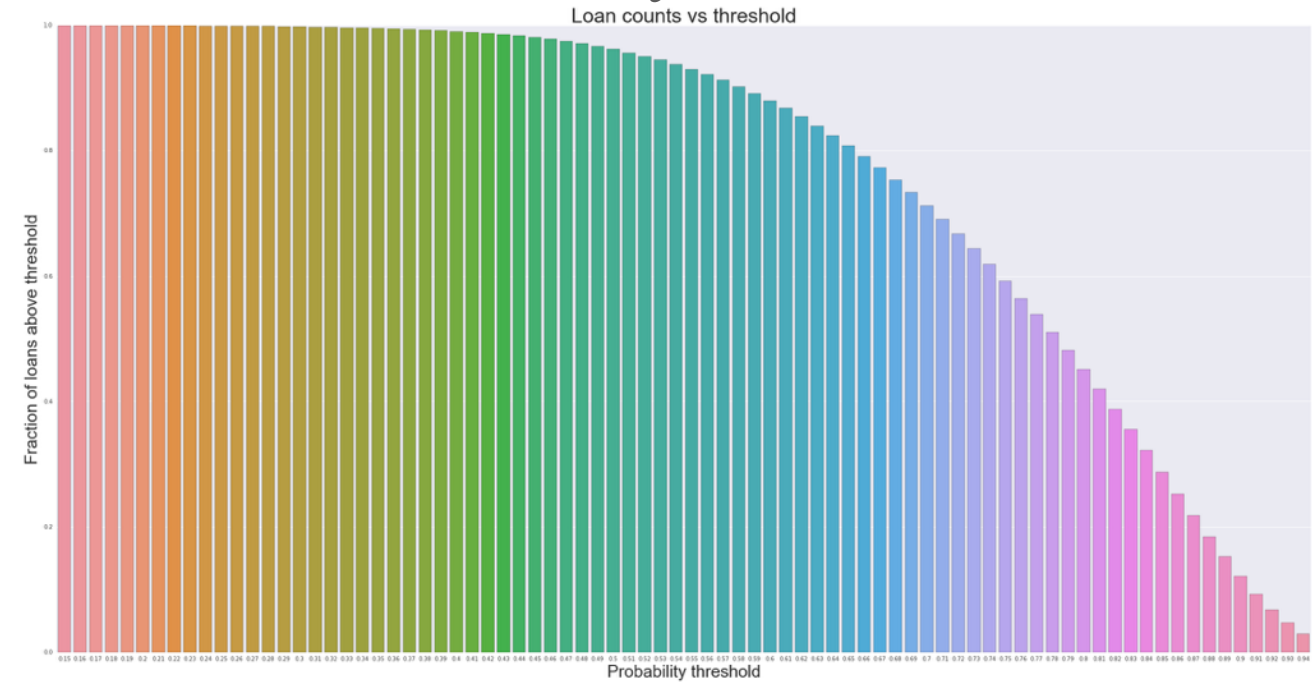
- Model very effective at predicting good loans
 - Annual ROI 5% better with model than without (2.7% gain vs 2.3% loss)
 - ROI increases monotonically with model confidence
 - Model predictions better than simply choosing a loan grade

MODEL RESULTS

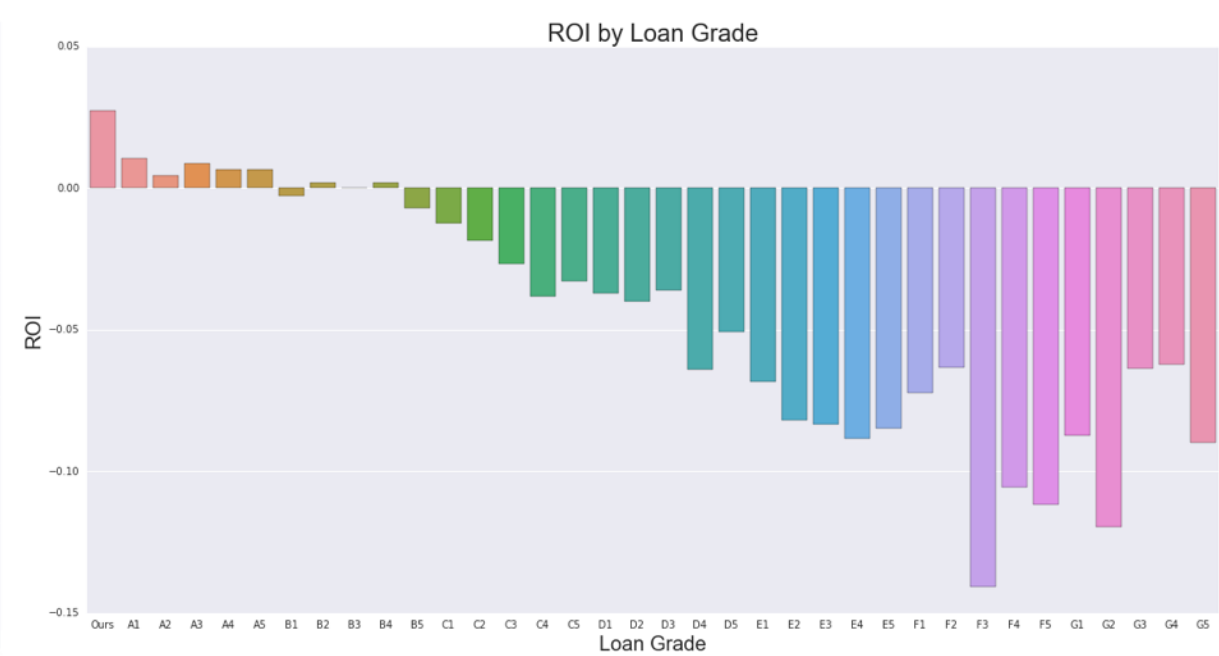
ROI by threshold



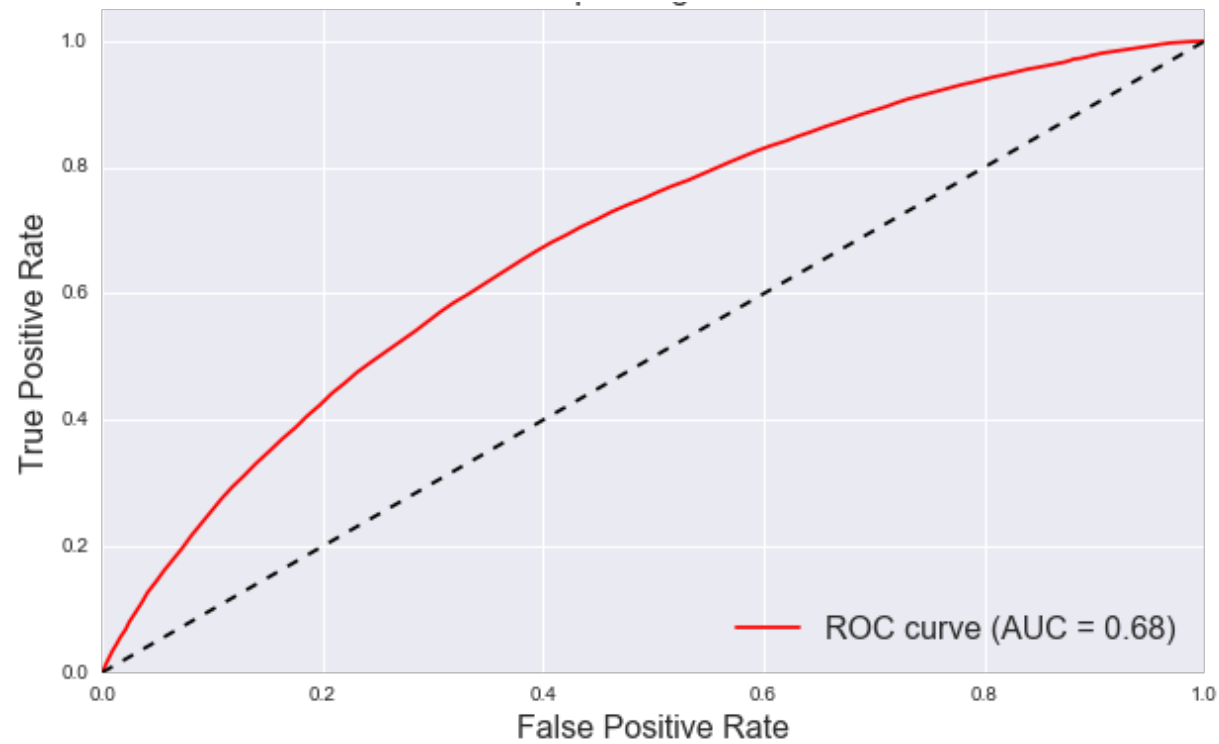
Fraction by threshold



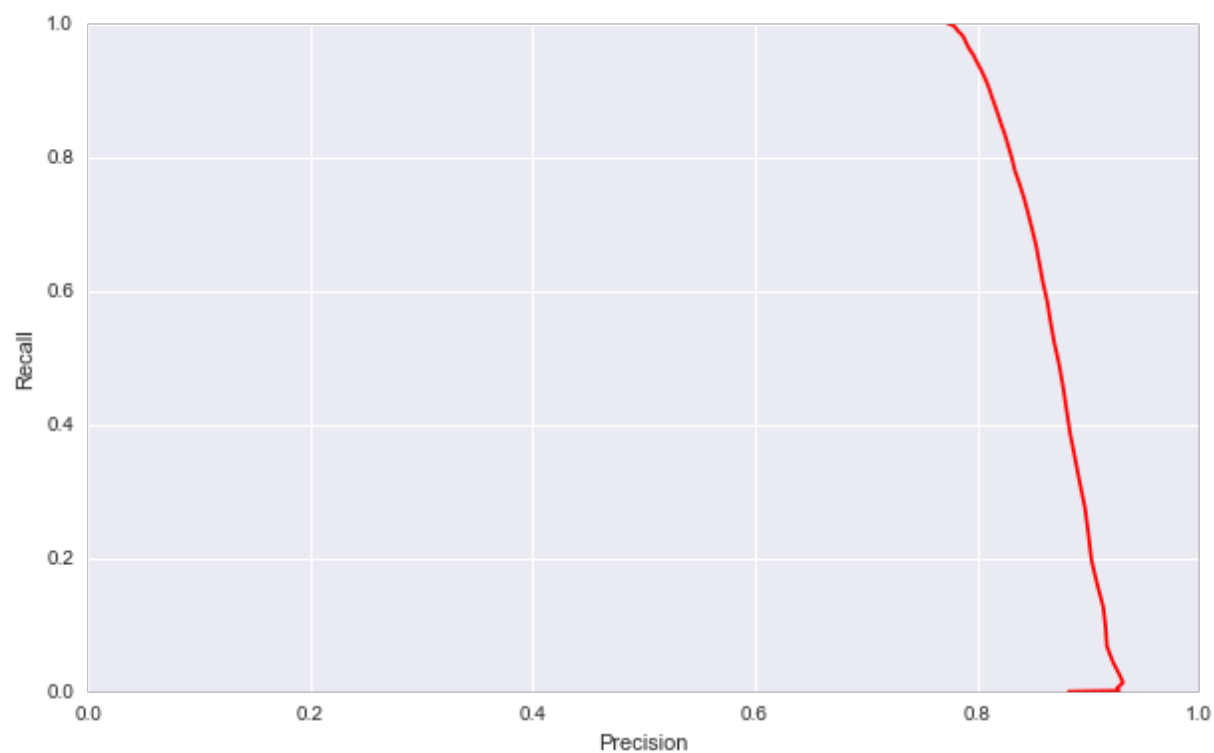
ROI by loan grade



ROC curve



PR curve



MODEL INSIGHTS

.....

- As retail investors, minimizing loss is important
- Generally prioritize precision in model tuning
- Reduce false positives (investing in bad loans) while false negatives (missing out on good loans) more acceptable

WHAT FACTORS PREDICT LOAN QUALITY?

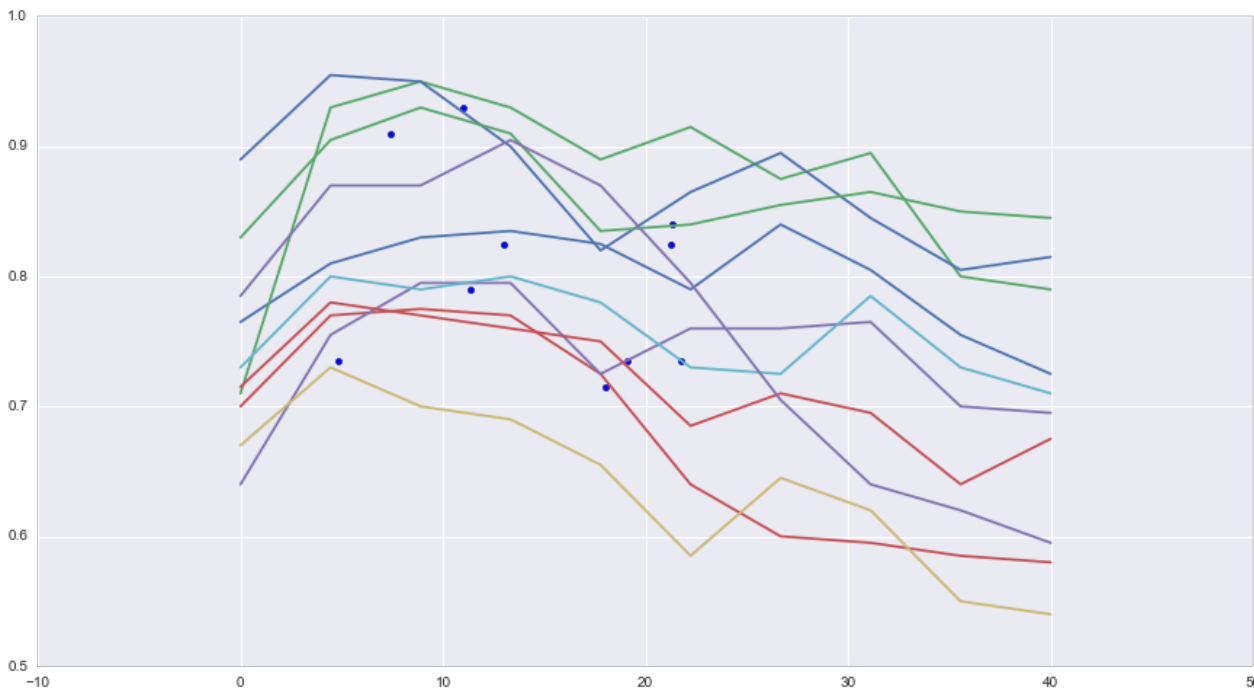
.....

<i>Feature</i>	<i>Relative importance</i>
dti	0.105739
zip3_mean_income	0.096541
annual_inc	0.094343
funded_amnt	0.087083
sub_grade_num	0.086094
pre_loan_credit_hist_length	0.076211
issue_d_month	0.067340
delinq	0.050987
emp_length_num	0.050815
issue_d_year	0.037504
derog	0.026922
desc_length	0.026219
title_length	0.026095
term_num	0.021847
record	0.018437

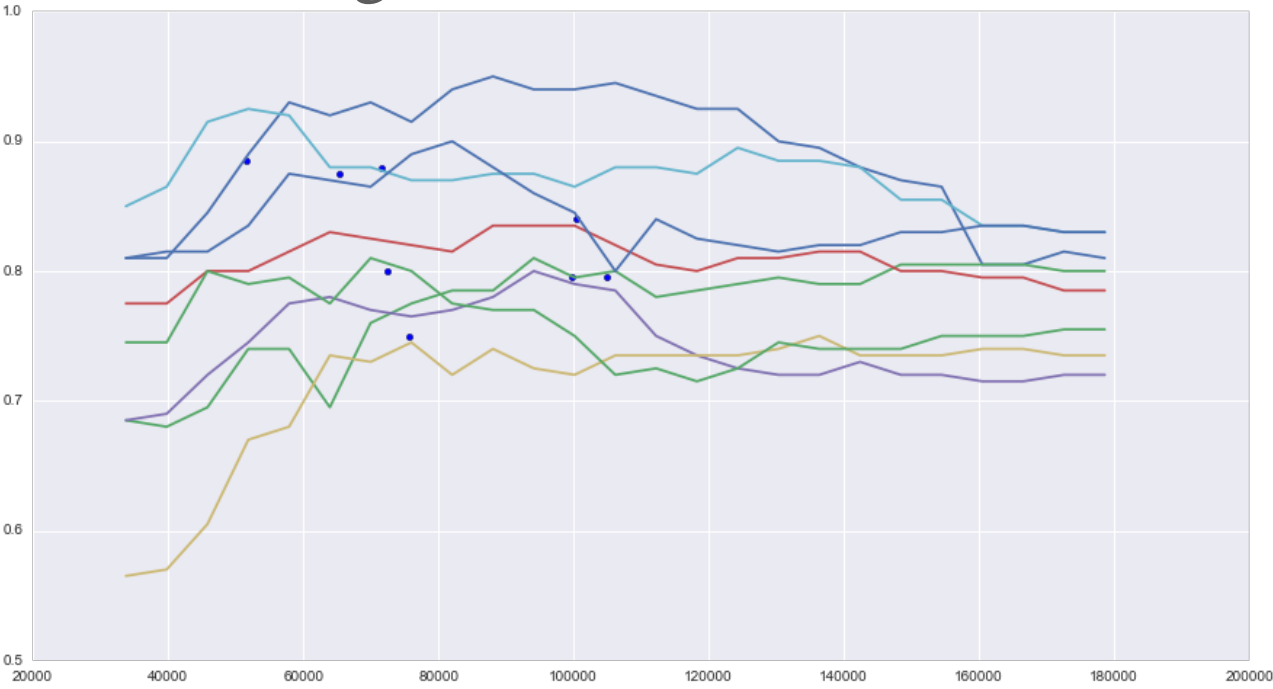
DIRECTIONAL INFLUENCE OF TOP FEATURES BY IMPORTANCE

.....

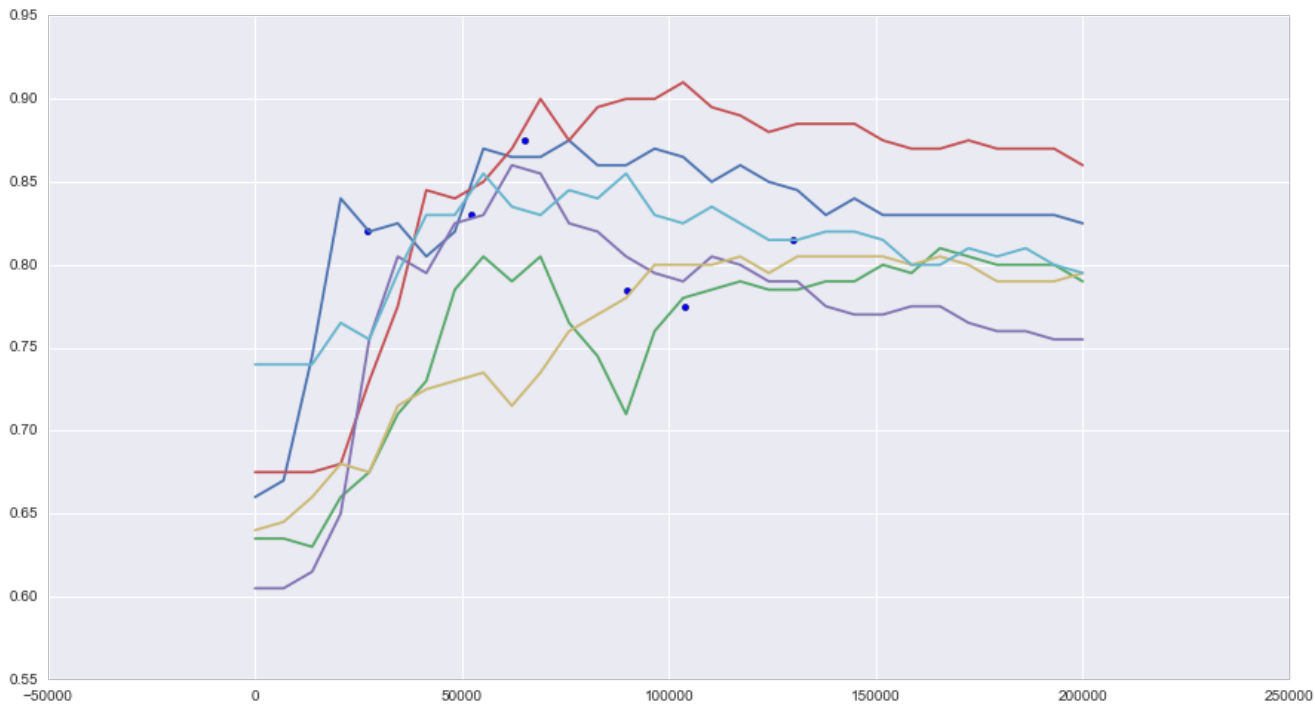
Debt-to-income ratio



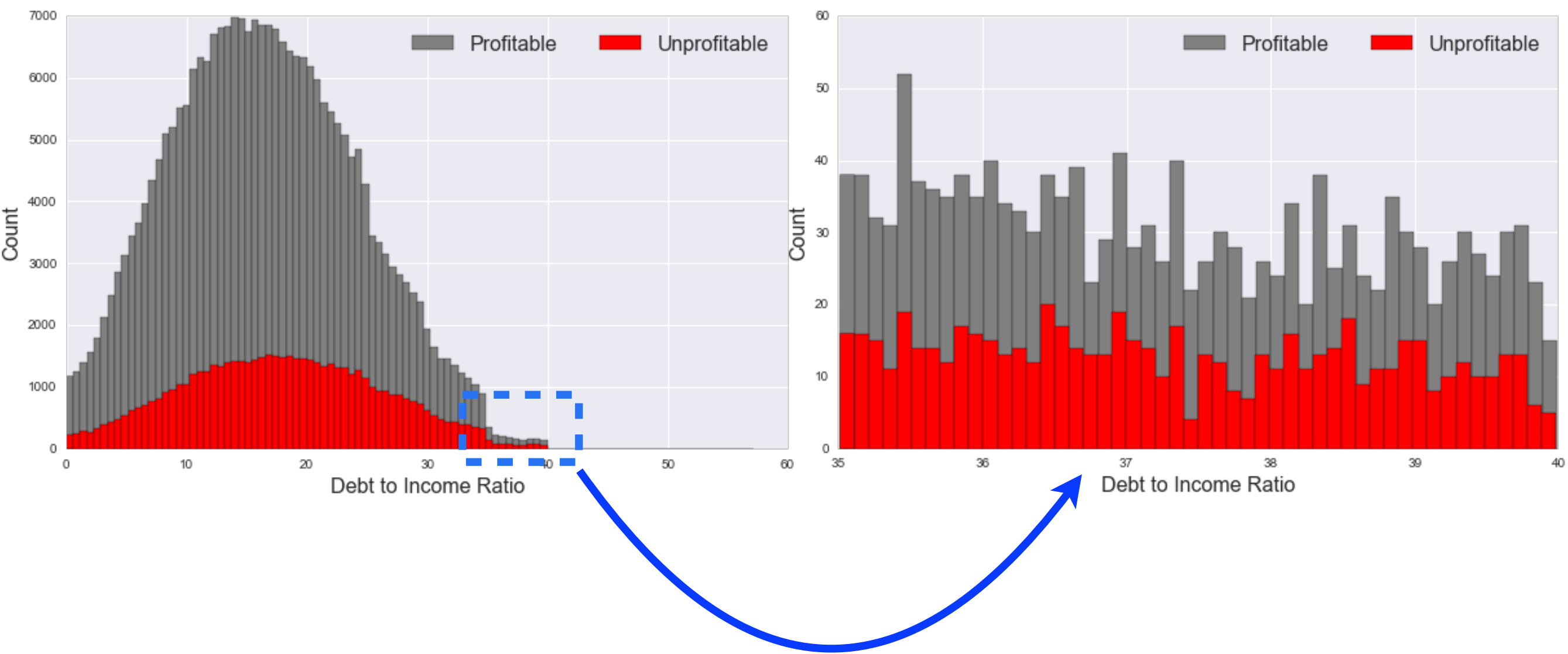
Avg income in ZIP3 area



Annual income



MODERATE DEBT-TO-INCOME LEVELS ASSOCIATED WITH HIGH RETURNS





KEY TAKEAWAYS

.....

- Defining positive/negative outcomes is a non-trivial problem
- Once classes are defined, reliable prediction not overly difficult
- Unusual features can have outsized predictive power
- D3 is a mystery



FURTHER EXPLORATIONS

.....

- Modelling and feature engineering
- Clarify and refine opportunity cost of money and value of payments
- Split train/test data by year
- Make web app real-time-ier
- Map viz by zip codes
- Master d3, challenge Mike Bostock to epic bubble based showdown