CARLSON SCHOOL OF MANAGEMENT

Lending Club

Project for Introduction to Statistics

Team:

- Chuqiao Chen
- Nithya Prasad Veeramacheneni
- Shishir Shirbhate
- Sophie Zheng

Contents

1. Introduction

1.1. Background of P2P Lending & Lending Club

1.2. Why we choose this project?

1.3. Questions of Interest

We have organized our questions in the increasing manner of complexity. Following are our questions of interest in the increasing order of complexity:

1. What are the characteristics of people who invest in Lending club and what are their requirements?
2. Which factors affect the interest rates of loans and to what extent?
3. What recommendations can we provide based on our understanding?

2. Analysis

2.1. Overview of data and sampling method

2.2. Understanding the customer base and their requirements

2.3. How different variables affect Interest Rate?

Before we understand how different variables affect the interest rate, let us take a step back to understand our process of how we reduced from 115 variables to 10 variables of interest:

1. Removed multiple variables which did not have available data in it
2. Removed variables which were just acting as a proxy for interest rates (E.g. Grade, Sub-grade)
3. Removed variables which did not make business sense. E.g. Installments is one of the variables which can only be identified after interest rate is decided. And since we did not have panel data we couldn't lag installments. Hence removed installments.

Below is the list of 10 variables which were considered:

1. Loan Amount
2. Credit used / credit limit
3. Monthly debt / monthly income
4. Total balance of all accounts
5. Total high credit limit
6. Home ownership
7. Total bank card credit limit
8. Term of loan
9. No. of bank card accounts used > 75% of limit / Total bank card accounts
10. Purpose of loan

We ran three models on the above variables:

#	Predictors	K	R-square	S	highest p-value	Corresponding variable	Comments
1	All 10	10	0.5108	3.113	0.565966	home ownership	

2	(home ownership, total current balance)	8	0.5017	3.141	0.013379	purpose-house	total current balance is highly collinear with total high credit limit
3	(purpose)	7	0.4355	3.341	2.04E-08	dti	removed purpose

Finally we choose Model no. 2 as our final model as 50% variance in interest rate can be explained by this model.

From model 2 we understand how are the following variables affect the interest rates by 1%:

1. Loan amount is directly proportional to interest rates:
   Around $14.6k to $19.2 increase in loan amount will increase the interest rates by 1%

2. Total Credit limit of customers is inversely proportional to interest rate:
   If credit limit decreases by $174.5k to $199.5k the interest rate increases by 1%

3. Total Bank Card credit limit is inversely proportional to interest rate:
   If total bank card credit limit decreases by $22.4k to $26.k then the interest rate increases by 1%

Additionally analyzing the factors which are in the form of ratios:

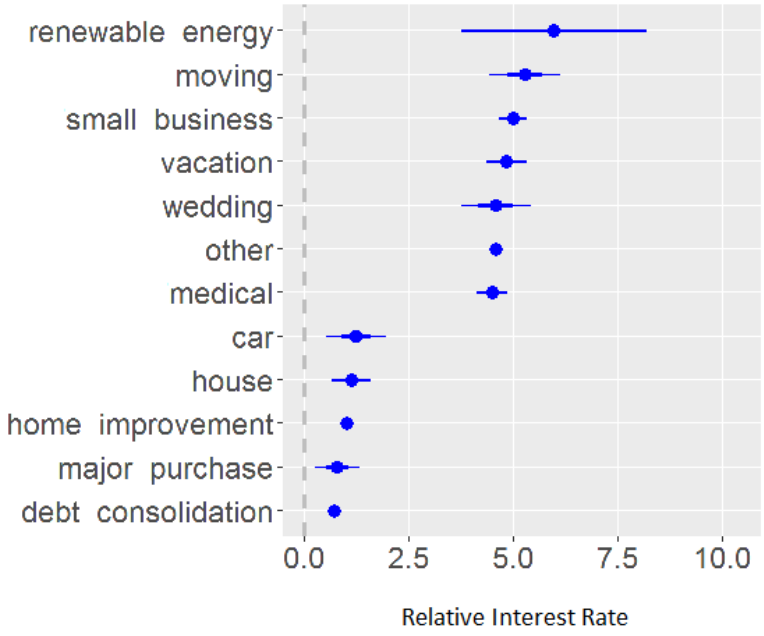
1. Credit card used / credit limit:
   As the credit used / limit increases to 26% to 36% then the interest rate increases by 1%

2. Monthly debt / monthly Income:
   As the ratio of monthly debt to income ratio increases by 27% to 41% the interest rate increases by 1%

3. No. of bank cards used more than 75% of limit / total bank card accounts:
   As the above ratio increases by 41.7% to 49.5% the interest rate increases by 1 %

Analyzing the interest rates for various purposes for which it is taken:
The below chart show how the interest rates vary among each other for different purposes of loans.
For e.g. Interest rate for moving is approximately 5% higher than interest rates for credit cards:

Note – The purpose for reference is credit card. All the other interest rates for purposes are in comparison with interest rate of credit card (Assuming all the other factors are kept fixed).



The interest rates for loans for 60 months duration are typically higher by 4.29% to 4.56% as compared to loans for 36 months duration keeping all other factors constant.

3. Conclusion

4. References

5. Appendix

5.1. Final model details:

```
##
## Call:
## lm(formula = int_rate ~ loan_amnt + revol_util + dti + tot_hi_cred_lim +
##     total_bc_limit + term + percent_bc_gt_75 + purposeF, data = Loan)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.5088  -2.2441  -0.1893   2.0942  14.9374
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 9.783e+00  1.392e-01  70.299  < 2e-16 ***
## loan_amnt                   6.031e-05  4.964e-06  12.150  < 2e-16 ***
## revol_util                  3.448e+00  2.017e-01  17.096  < 2e-16 ***
## dti                         2.936e-02  4.268e-03   6.879 6.39e-12 ***
## tot_hi_cred_lim            -5.369e-06  2.187e-07 -24.545  < 2e-16 ***
## total_bc_limit             -4.125e-05  1.971e-06 -20.928  < 2e-16 ***
## term 60 months              4.432e+00  8.110e-02  54.650  < 2e-16 ***
## percent_bc_gt_75            2.237e-02  1.329e-03  16.831  < 2e-16 ***
## purposeFcar                 1.234e+00  3.568e-01   3.459 0.000544 ***
## purposeFdebt_consolidation  7.116e-01  7.748e-02   9.184  < 2e-16 ***
## purposeFhome_improvement    1.019e+00  1.514e-01   6.727 1.83e-11 ***
## purposeFhouse               1.126e+00  4.550e-01   2.474 0.013379 *
## purposeFmajor_purchase      7.832e-01  2.634e-01   2.973 0.002952 **
## purposeFmedical             4.496e+00  3.649e-01  12.323  < 2e-16 ***
## purposeFmoving              5.284e+00  4.258e-01  12.411  < 2e-16 ***
## purposeFother               4.583e+00  1.631e-01  28.105  < 2e-16 ***
## purposeFrenewable_energy    5.972e+00  2.222e+00   2.687 0.007212 **
## purposeFsmall_business      4.998e+00  3.232e-01  15.466  < 2e-16 ***
## purposeFvacation            4.843e+00  4.745e-01  10.207  < 2e-16 ***
```

```
## purposeFwedding             4.592e+00  4.185e-01  10.974  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.141 on 9980 degrees of freedom
## Multiple R-squared:  0.5017, Adjusted R-squared:  0.5007
## F-statistic: 528.8 on 19 and 9980 DF,  p-value: < 2.2e-16
```

Variables:

- revol_util - Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit.
- dti - A ratio calculated using the borrower's total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower's self-reported monthly income.
- tot_hi_cred_lim - Total high credit/credit limit
- total_bc_limit - Total bankcard high credit/credit limit
- percent_bc_gt_75 - Percentage of all bankcard accounts > 75% of limit.

5.2. Confidence intervals for all variables at 90% level of significance:

```
##                                    5 %          95 %
## (Intercept)                9.554265e+00  1.001212e+01
## loan_amnt                  5.214751e-05  6.847997e-05
## revol_util                 3.116407e+00  3.780005e+00
## dti                        2.233988e-02  3.638256e-02
## tot_hi_cred_lim           -5.728854e-06 -5.009197e-06
## total_bc_limit            -4.449356e-05 -3.800851e-05
## term 60 months             4.298460e+00  4.565267e+00
## percent_bc_gt_75           2.017982e-02  2.455187e-02
## purposeFcar                6.474585e-01  1.821486e+00
## purposeFdebt_consolidation 5.841308e-01  8.390385e-01
## purposeFhome_improvement   7.694837e-01  1.267671e+00
## purposeFhouse              3.771988e-01  1.874239e+00
## purposeFmajor_purchase     3.499006e-01  1.216486e+00
## purposeFmedical            3.896189e+00  5.096651e+00
```

```
## purposeFmoving              4.583866e+00  5.984725e+00
## purposeFother               4.315117e+00  4.851663e+00
## purposeFrenewable_energy    2.316585e+00  9.628044e+00
## purposeFsmall_business      4.466285e+00  5.529477e+00
## purposeFvacation            4.062466e+00  5.623469e+00
## purposeFwedding             3.903860e+00  5.280666e+00
```