

Important factors that influence loan's interest rate

Introduction

Various potential factors attribute the loan interest rate. The first one that may come to our mind is the loader's reliability. Generally speaking, the more reliability a loaner has, the more that the loan issuer will trust him and as a result, the lower interest rate. And as a matter of fact, FICO[1] score is a commonly used indication of loaner's reliability. However, taking into account only the FICO score is not quite practical when deciding the loan interest rate, as there is a variety of other factors that are attributable.

In this study, we performed primarily observational data analysis on a date set of peer-to-peer loan records, of each which record contains the information of the loaner and the final interest rate decided by the lender. We made the following works:

1. we analyzed why FICO score is the most important factor among all the factors on interest rate.
2. we found out the other four important factors on loan interest rate, which are 1), requested loan amount, 2), amount funded by the investors, 3), loan time length, 4), lender's inquiry count.
3. we studied the association between interest rate and each of the above five factors respectively, under different levels of FICO score,

Method

Data collection

The data set used in this study consists of a sample of 2,500 peer-to-peer loans issued through the Lending Club[2]. The interest rate of these loans is determined by the Lending Club on the basis of characteristics of the person asking for the loan such as their employment history, credit history, and creditworthiness scores.

Exploratory data analysis

Exploratory analysis was performed by examining the collected data table. In this step, the following work have been done:

1. gaining a rough concept of what is data set is like, for example the data type for each column and what characteristics each column tries to capture.
2. performing necessary data type conversion, such as the interest rate from string type ("0.12%") to numeric type (0.0012)
3. identifying missing values in the data set

Statistical modeling

In order to find out the important factors on interest rate, we performed individual studies on how loan interest rate varies with a specific factor. Here we made different processing on qualitative and quantitative variables:

1. If the factor being studied is qualitative, we plotted the box whisker plot of the interest rate under each possible value of the factor.
2. For quantitative factors, we cut it into a certain number of value ranges for better visualization and subsequently performed the same operation as the qualitative one.

In certain cases, if the correlation between interest rate and one factor is not quite obvious, we linearly fitted the data in order to see a clearer trend.

In addition, as FICO score is the primary influencing factor on interest rate, we studied under different FICO score, how interest rate varies with other factors respectively.

Reproducibility

To reproduce the exact result, the scripts in the script must be executed. Scripts used in this study is contained in the *script/final* folder.

Results

FICO score as the primary factor

As FICO score is quantitative, we cut all the scores into five ranges(644~684, 684~699, 699~719, 719~744 and 744~834), each of which contain nearly 500 individual scores. After grouping interest rates based on FICO score into 5 groups, the median interest rate for each group is

FICO Score	[644,684)	[684,699)	[699,719)	[719,744)	[744,834]
Median of Interest Rate	0.1629	0.1411	0.1269	0.1025	0.0790

As we can see a clear drop of interest rate as the loaner FICO score rises.

Requested loan amount and amount funded by the investors

As we observed that the correlation between the two factors is nearly 0.97, which is very high, we can treat them as one factor for the sake of simplicity, thus we *only* take amount requested into consideration. As the amount is quantitative, we divided into 4 ranges, 1000~ 6075, 6075~10050, 10050~17200 and 17200~35000. Similar to the above method, we have the following data:

Amount Requested	[1000, 6075)	[6075,10050)	[10050,17200)	[17200,35000]
Median of Interest Rate	0.1197	0.1235	0.1311	0.1531

As we can see a clear rise of interest rate as the loaner's requested amount rises. Our interpretation is as the amount to be lent becomes larger, it is more risky for the lender to issue this loan because he will lose more in case the loaner doesn't pay back. To reduce the risk or to some degree, relieve the lender, the lender rises the interest rate.

Taking the effect of FICO score into account, we found out as the loaner becomes more reliable, the increasing rate of interests rate will be smaller. Based on the previous interpretation, it is easy to conclude as the loaner becomes more reliable, it is less risky for the lender to issue the loan, thus lower interest rate.

Loan length

Loan length indicates the length(in months) of the loan

Loan length' case is quite similar to the previous, as:

1. interest rate increases as the loan length becomes greater
2. the loaner's FICO score has a positive impact on the interest rate. As the FICO score increases, the rate of interest rate increasing will be smaller.

Also, we have similar interpretations for the result:

1. As things change constantly, the longer the loan is, the more unpredictability there will be, and thus the more uncertainty or doubt the lender will have, which makes the interest rate rises.
2. If the loaner is reliable, there is be less chance that he will not pay back the money and the lender will be more confident about this loan, thus lower interest rate.

Times the loaner is inquired on his/her creditworthiness[3]

Subjectively thinking, a loaner, who is seldom inquired on his creditworthy because people trust him and think there is no need to check his trustfulness, is more likely to receive a low-interest- rate loan, than a loaner who is inquired more often because he is less reliable.

As a matter of fact, this way of thinking is reasonable as we compared interest rates issued to loaners who are less inquired on creditworthiness and those who receive more. In our data set, those who has not been inquired in the last six months have a median interest rate of 12%, while for those who have been

inquired for at least once in the last 6 months, the median value is approximately 14%, 2 percent higher.

Taking into account the FICO score, there is a similar pattern here as in both the loan length and amount requested case. As the loaner becomes more reliable, the lender's uncertainty rising rate will be slower.

One thing to note here, there is a substantial difference on interest rate between those who has never be inquired and those who have. We interpret this phenomenon as the lenders think it is a important sign for creditworthiness if the loaner has never be inquired.

Conclusions

Our analysis suggests that besides FICO score, which indicates loaners' creditworthiness, three factors on the loaner and the loan play an important role in determining the interest rate issued. The first two are requested loan amount and amount funded by the investors. And according to our finding, they are closely correlated. The third one is loan time length. We interpret this as, as the loan duration grows, there will be more unpredictability, so the lender rises the interest rate in hope of lowering the risk. Lender's inquiry times on the loaner is the last factor. Our speculation suggests this factor maybe an indicator of the creditworthiness of the loaners.

As FICO score is the primary influencing factor, we studied its influence on the above four factors and surprisingly find out a common pattern shared by them. As the loaner becomes more reliable, the increasing rate of interest rate(as one factor poses more impact), actually decreases.

While our analysis is an interesting first step it is based on a limited sample of 2500 loan records. A larger collection of representative earthquakes may be more appropriate for understanding the relationship between earthquake magnitude and depth.

References

1. FICO score, http://en.wikipedia.org/wiki/Credit_score_in_the_United_States, Accessed on 2013.2.18
2. Lending Club, <https://www.lendingclub.com/home.action> , Accessed on 2013.2.16
3. Credit Inquiry, <http://www.myfico.com/crediteducation/creditinquiries.aspx>, Accessed on 2013.2.18
4. R documentation, <http://www.r-project.org/other-docs.html>, Accessed on 2013.2.18