Prosper Distressed Loan Analysis

July 28, 2018

1 Stories

- First version of the story
- Final version of the story
- Updated Final version of the story

2 Summary

Prosper is one of the leading platforms that provide peer-to-peer lending. Their loans from 2006 to 2014 are available. We used the data to attempt to discover what contribute most to the distressed scenario, when loan went bad. We found that the charged off status is a good measure of the distress - whether people suffer economically or not. We then looked at the available data and characteristics of the loan such as borrower's credit rating, employement status, homeowneship status, and the state of residence, as well as the term of the loan. We found each of these factors might or might not affect the load charged off rate. If they do, the difference appears to be small.

We found that the most dominate factor aaffecting the increase of charged off rate is time. In the bad year of 2008 (world-wide economic recession), and 2010 (EU recession and US slowed down), everyone suffered. Some categories might do a bit better relative to others, but can never be better than in good years.

3 Design

Prosper is one of the leading platforms that provide peer-to-peer lending. The platform has two unique characteristics that separate it from the traditional lending industry. First is that the online nature of the business. Second is they use novel way of assessing creditworthiness of borrowers. Although the platforms might not have predetermined market segment to target to, these two unique characteristics might attract certainly type of borrowers more than others.

We observed the status of the loan and identified that charged off is a better way to analyze the distressed status of the loan as there is always certain amount of charged off over time. We also calculate the charged off rate, which is the percentage of the loan in charged off status over the total amount of loans. This is essential as Prosper is a fast growing business with number of loan increase expotentially over the years. If we just use the number of record of charged off, they will always increas over time and cannot show the distressed status we want to investigate.

Then we look at various attributes to attempt to determine are there primary factors that cause the distress. We looked at:

- Time
- Term of the loan: 12, 36, and 60 months
- Borrower is a Homeonwer or not
- Borrower's employment status
- Borrower's credit rating
- Borrower's resident state

Proper did some methodology changes in Q2 of 2009. For example, before that, Credit grade was gather and used. But Credit Grade was abandoned after that, instead, they used a proprietary Prosper Rating. Employment Status went through some changes as at Q2 of 2010 as well. Before that, Full Time is the majority of status. After that, the majority transitioned to Employed. While Full Time status still exist, they eventually disappeared after a few years.

3.1 Chart Selection

- We used stacked bar chart to look at the complete number of records of various loan status. THis identified what type of loan status is more signicicant and relevant to our study - which is charged off and default.
- Now that there are only two status, we use linec chart to show the difference between charged off and default.
- For **Term, Homeownership, Employment Status, Credit Grade, ProsperRating**, we used stacked bar chart of both Rate, and Number of Records. The number of records is to emphasis the wide variation of the number in each group, hence the rate will be a better measure. So each picture will contain both rate and number of records in two separate charts. Highlandsight this decision make the **rate** more difficult to identify the difference because the difference is small.
- We use a map to show state-wide charged off, a obvious choice. Also added a filter to be able to show individual years.

4 Feedback

Rebecca Barnes provided great feedbacks for the first version of the story on a stack post.

5 Redesign based on feedback

The feedbacks are awesome and give very to the point suggestions. Based on the feedback, we adjusted the story as below:

- **Removed the number of records as a chart and use it as a tooltip.** As explained in Design section, putting the number of records is to emphasis the important that rate is a better metric instead of number of records. But using a chart to tell the same idea over and over just distract reader more. It's better to hide this information as tooltip and let the main idea come out.
- **Use line chart for comparison.** The stacked bar chart is a way to consistency when showing both rate and number of records. Given that we already removed the number of record chart, a natural way for comparing value like line chart is better fit.

- **Edited color of employment status.** Given that we want to highlight employed and full-time status, and there all too many status that full color chart will distract reader, we changed all to gray scale for all else and only use color on Employed and Full-time. We also added a highlight section in case reader want to drip into each employment status.
- **Put Credit Grade and ProsperRating into the same chart and use color and size respectively.** Prosper uses two value for the score about credit as they have different systems before and after 2009. However, this is really one concept. Putting them on two different charts might give reader the idea that they are different. So we put them on one chart and use color and size respectively to represent difference in the grades. Because Credit Grade has all null value after 2009, and ProsperRating has all null value before 2009, the final chart looks surprising readable.
- **Added a time series chart for the state**, in addition to map to hightlight the difference in time.
- **Enarlged caption boxes** to be able to put more words and described the story more clearly.

6 Resources

- Prosper's definition of default
- Prosper's definition of charged off