Prosper Loan Data Analysis

By Sarah Pedigo

Introduction

This dataset is from Prosper loan company who specializes in personal loans. A personal loan is a loan for anything that is not a house, car, or college payment. They can be used for weddings, paying off credit card debt, repairing your house, funding a trip, and many other things.

Necessary definitions

ProsperRating- (numeric) The Prosper Rating assigned at the time the listing was created: 0 - N/A, 1 - HR, 2 - E, 3 - D, 4 - C, 5 - B, 6 - A, 7 - AA. Applicable for loans originated after July 2009.

ProsperScore- A custom risk score built using historical Prosper data. The score ranges from 1-10, with 10 being the best, or lowest risk score. Applicable for loans originated after July 2009.

DebtToIncomeRatio- The debt to income ratio of the borrower at the time the credit profile was pulled. This value is Null if the debt to income ratio is not available. This value is capped at (10.01) (any debt to income ratio larger than 1000% will be returned as 1001%).

Main feature explored

Question: What relationship does Prosper score have with the other variables? What factors affect your score? What does your score affect?

Steps taken

1. loaded the CSV file into a dataframe and made a copy of it
2. dropped columns I did not plan on using so that the chart was more readable
3. defined my variables
4. created a dataframe where entries are grouped by Prosper score
5. created a histogram depicting frequencies of Prosper score
6. explored rest of variables univariately
7. created a bar chart comparing Prosper scores and monthly loan payment
8. created scatter plot to compare Prosper score vs monthly income, created a bar plot when the scatter plot did not show any patterns

Steps taken, cont.

9. created scatter plot using Prosper score and monthly income, then a bar chart when the scatter plot did not show any patterns

10. created scatter plot using borrower APR and Prosper score

11. explored relationships between other variables

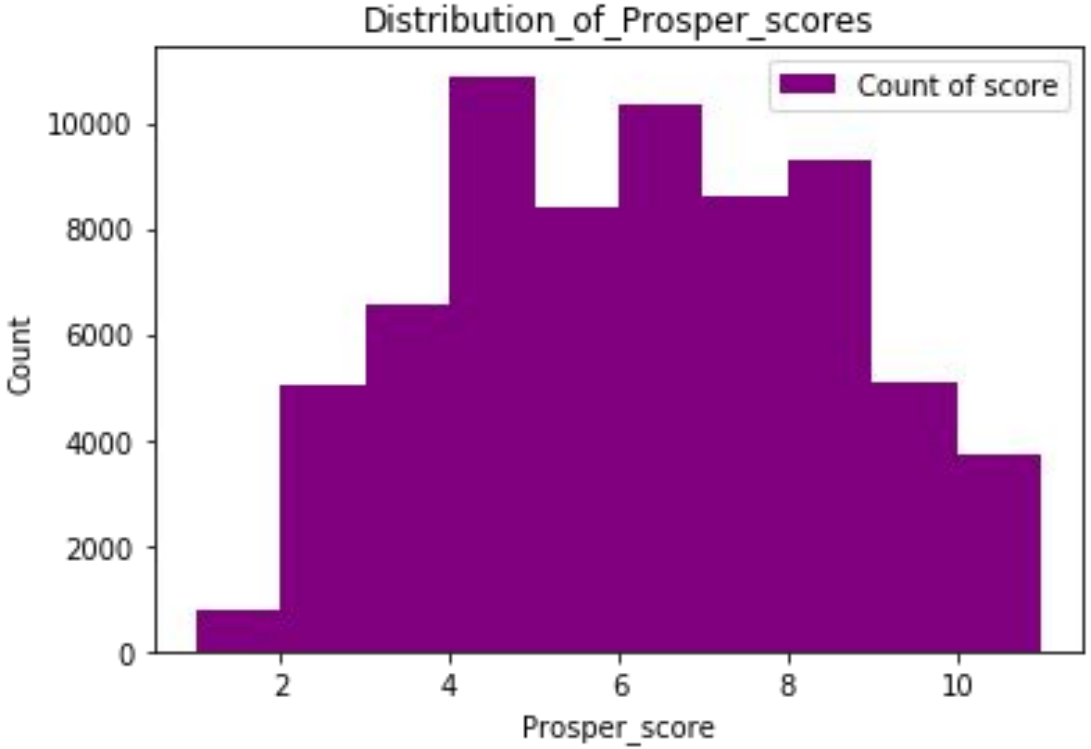
12. created a correlation matrix for several of the chosen variables

13. created a scatter matrix with several of the chosen variables

14. summarized findings

Analysis

Distribution of Prosper scores



Comments on distribution of Prosper scores

- My first visualization is a pie chart that depicts the distribution of Prosper scores.
- The majority of scores fall within 4-8.
- 4 is the most frequent value.
- The purpose of this chart is to get familiar with the main variable I will be using.

Prosper score vs monthly loan payment

Comments on Prosper score vs monthly loan payment graph

- The purpose of this chart is to examine the monthly loan payment for each Prosper score and visualize trends.
- I predicted that the monthly payment would increase with decreasing Prosper scores.
- As we can see on the previous graph, it actually does the opposite!
- My hypothesis is that this decrease is due to decrease in overall credit balance available to them.

Prosper score vs monthly income

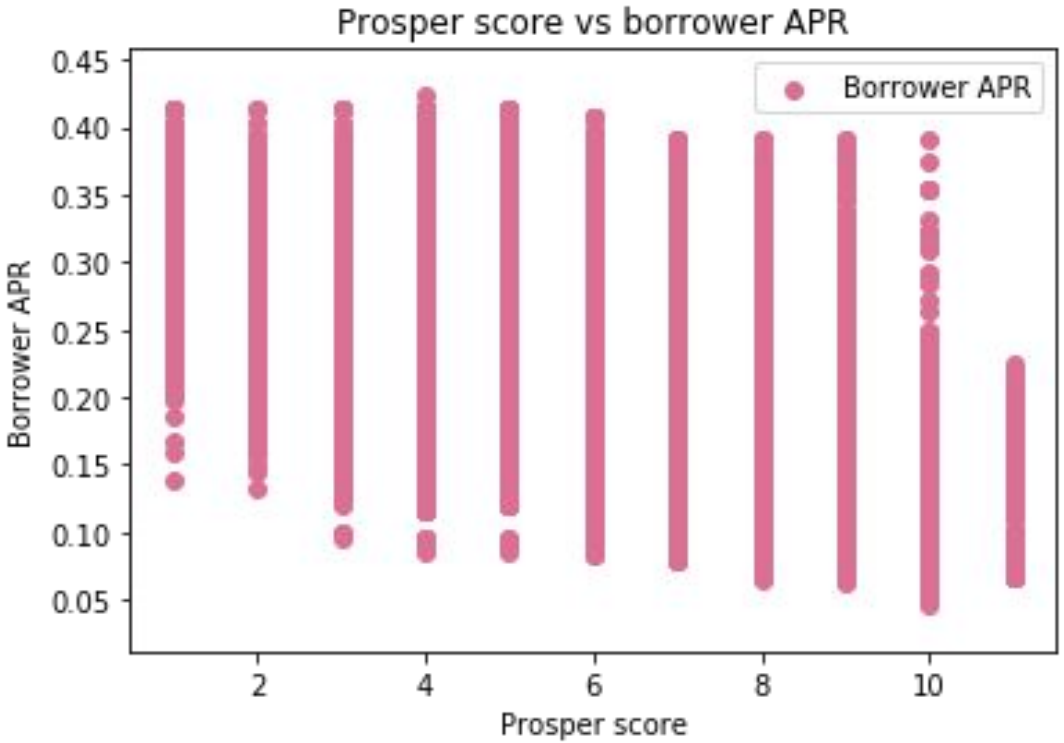
Comments on Monthly income vs Prosper score, cont.

- The purpose of this chart is to show the overall trend for the relationship between monthly income and Prosper score
- In general, the higher the score the higher the income.
- Monthly income ranges by approximately $2,000.
- A Prosper score of 1 has a mean income that is slightly higher than those that follow it, but there are far fewer incomes to average out with its low frequency.

Being a homeowner vs Prosper score

- Being a homeowner has little to no effect on Prosper score.
- Most common score in both homeowner and non-homeowner group is 4.

Prosper Score vs borrower APR



Comments on Prosper score vs borrower APR

- The purpose of this chart is to view the relationship between Prosper score and borrower APR.
- APR generally increases as Prosper score decreases.
- Range of 0.2 (20%!)
- Those with low Prosper scores can have a lower APR but those with high Prosper scores do not have high APRs.

Revolving credit balance vs Prosper score

- Balance increases as Prosper score increases.
- This is likely due to higher credit limits for those with higher scores

Notable findings

1. Among borrowers, there are almost the same amount of homeowners as non-homeowners.
2. The majority of Prosper scores fall between 4 and 8 (distribution is slightly bell shaped)
3. APRs tend to stay in the .15-.25 range, but go close to 0 and up to .45.
4. With loan status, users have 5 times more current ones than defaulted!
5. An estimated 80-90% of monthly loan payments fall below $500 with most falling below the $300-350 range.
6. Lower Prosper score tends to predict a lower monthly payment.
7. Income increases gradually with Prosper score.

Key insights/conclusions

1. Being a homeowner has no conclusive effect on a client's Prosper score.
2. Prosper scores have some effect on the APR a client receives.
3. Income can affect a client's given score.
4. Monthly loan payment is calculated by the client's earned income.

References

Prosper Loan Data - Variable Definitions. (n.d.). Retrieved December 15, 2019, from https://docs.google.com/spreadsheets/d/1gDyi_L4UvIrLTEC6Wri5nbaMmkGmLQBk-Yx3z0XDEtI/edit#gid=0.