PROSPER LOAN DATA EXPLORATION REPORT

By Maximus Ikengwa

Introduction

Prosper Loan company is is a lending platform located in USA. They provide personal loans to borrowers with good credit history. To qualify applicants, the company uses a proprietary rating system that considers data points like employment status, income, and debt-to-income ratio. Borrowers are then assigned a Prosper score, which investors use to decide on loans to fund.
This investigation explores the motivations behind loans taken through Prosper while identifying how prosper ratings may influence loan favorability.

Dataset Overview

The dataset contains 113,937 loans with 81 features which includes but not limited to the following; LoanOriginalAmount, BorrowerAPR, StatedMonthlyIncome, Term, ProsperRating (Alpha), EmploymentStatus and many others.it also has more than 50% of the columns containing numerical data. 871 data points were removed due to inconsistency and non-relevance to the main focus of my investigations

The following steps were taken throughout this exploration exercise;

Data Cleaning: I went through the dataset and removed some columns that were not needed for my explorations. I also removed over 8oo duplicated data points.

Data Exploration; Here I tried to view different columns on my interest separately, this helped me to understand how there are distributed along the data frame. Some of the data I looked at in the entire process are as follows;

BorrowerAPR: The distribution of Borrowers'APR seems multimodal. we have peaks at 0.09,0.8,0.22,0.3 and 0.35 areas, then a very sharp peac peak at 0.37 area. The least of the APR comes after 0.43.

ProsperRating (Alpha): The distribution of Prosper Ratings is a one mode distribution, with the most common rating belonging to the central grade represented (**C**), has the highest ratings, followed by (**B**) and then (**A**). (**AA**) are the least common.

Loan Original Amount : The very large spikes in frequency are at 10k, 15k, 20k, 25k and 35k. There are also small spikes at 8k,9k,11k,12k,13k,14k etc. It means that most of the loans are multiples of 1k.

Summary of the Exploration In terms of borrower motivations, we found surprising results. Rather than take loans to start businesses or purchase assets, the largest population seems to take large amount of Loans t finance weddings, child adoptions, boat acquisitions, and the purchase of engagement rings; These are things that ordinarily doesn't have any return on investment, hence money spent on them cannot be recovered. This explains the reason, we have Many people taking Loans just to finance their debts. Debt consolidation also accounts for the highest loan amounts collected from the platform on average. While all these may point to possibly self-indulgent reasons, it can also be the lifestyle of the people living in that Region

We also measured loan favorability using the annual percentage rate attached to a loan (the Borrower APR). we found out that the success of most Loan application is influenced positively by high and verifiable incomes, homeownership, low debt to income ratio, and the presence of a current means of employment. Also Borrower APR is negatively correlated with the loan original amount, loan term, and prosper rating(these are only detailed information needed while filling out an application form not an asset that might be considered as collateral).

On further exploration, another surprising interaction was discovered. There seemed to be a dichotomy in the interaction between borrower APR and prosper ratings. Between the lower ratings of HR to B, borrower APR and prosper ratings were negatively correlated. This interaction turns positive within the high prosper rating group (B to AA). We attributed this to the possible influence of lurking variables, such as the loan term, and borrowing power of high-income earners who are usually rated higher on the prosper scale. High income earners seemed to borrow more when long-term loans are involved, increasing their debt to income ratio. Hence, an increase in APR might be a great way to disincentivize 'overborrowing'. On the other hand, decreasing APR by term might be a great way to encourage low-rated borrowers to take long-term loans; giving them enough time to repay their loan.