Prosper Loans

December 24, 2022

1 Prosper loan Data Analysis

1.1 by Joseph Stanley

1.2 Investigation Overview

In this investigation I wanted to focus mainly on the apparent lending patterns of the company and how these may be related to peculiarities discovered in borrower profiles especially how the company's Lending Amount is particularly varied across various borrower profiles.

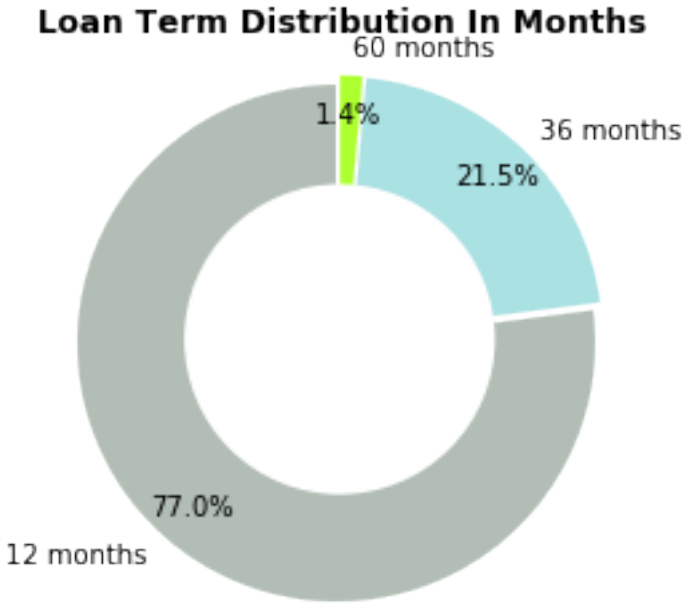
1.3 Dataset Overview

This data set contains 113,937 loan listings by a loan company Prosper with 81 variables on each loan which are uniquely identified by a listing Key and are stored on Udacity's server.
These loan variables include: * The loan amount. * The borrower's APR (Total annual percentsge rate). * The status of the loan. * The borrower's stated monthy income and many others available

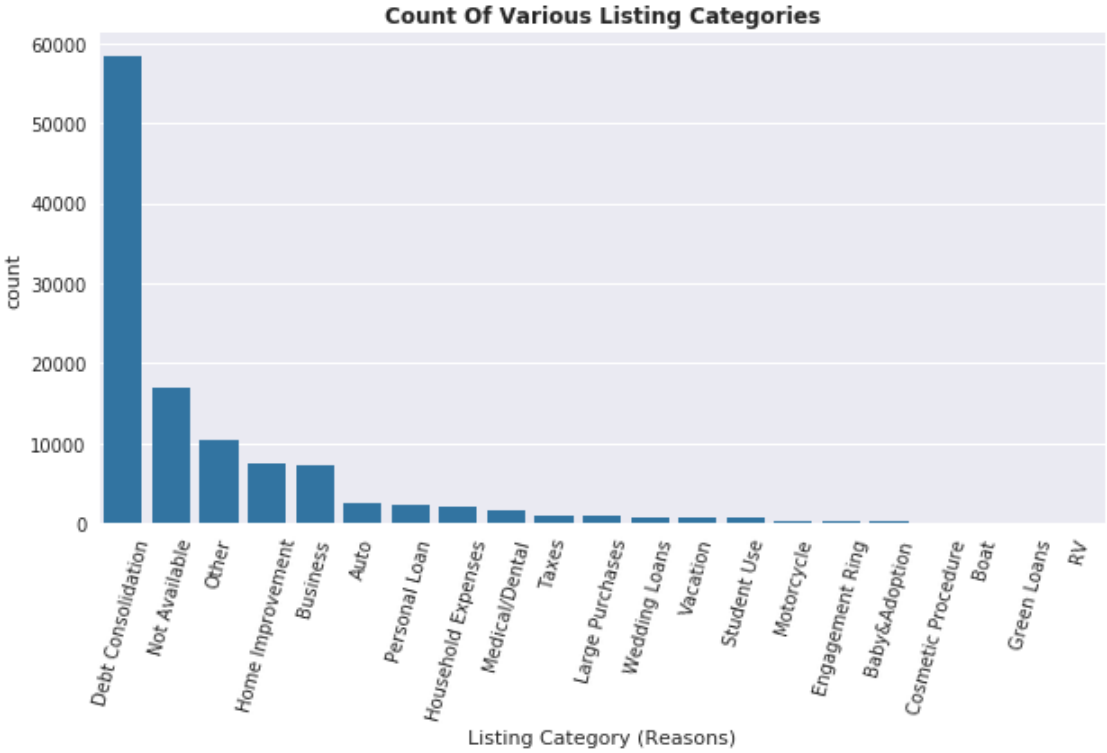
1.4 Basic Loan Characteristics

Majority of the loans were currently being serviced or completed and were listed on 36 and 60 month Terms. A very significant portion of these loans were also taken for debt consolidation i.e to offset other loans.

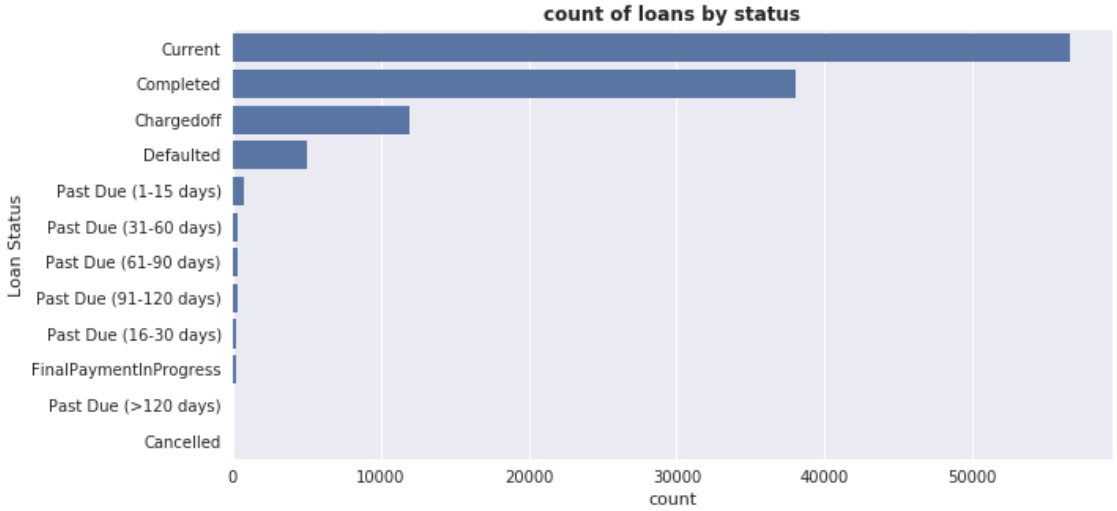
In [4]:


Loan Term Distribution In Months
60 months
1.4%
36 months
21.5%
77.0%
12 months

In [5]:


Count Of Various Listing Categories
60000
50000
40000
30000
20000
10000
0
count
Debt Consolidation
Not Available
Other
Home Improvement
Business
Auto
Personal Loan
Household Expenses
Medical/Dental
Taxes
Large Purchases
Wedding Loans
Vacation
Student Use
Motorcycle
Engagement Ring
Baby&Adoption
Cosmetic Procedure
Boat
Green Loans
RV
Listing Category (Reasons)

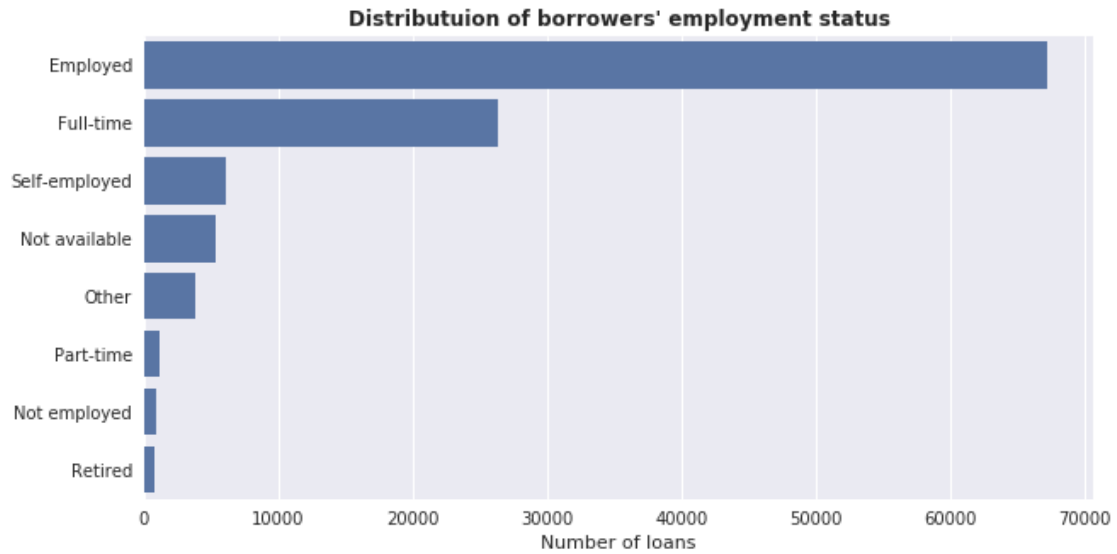
In [6]:



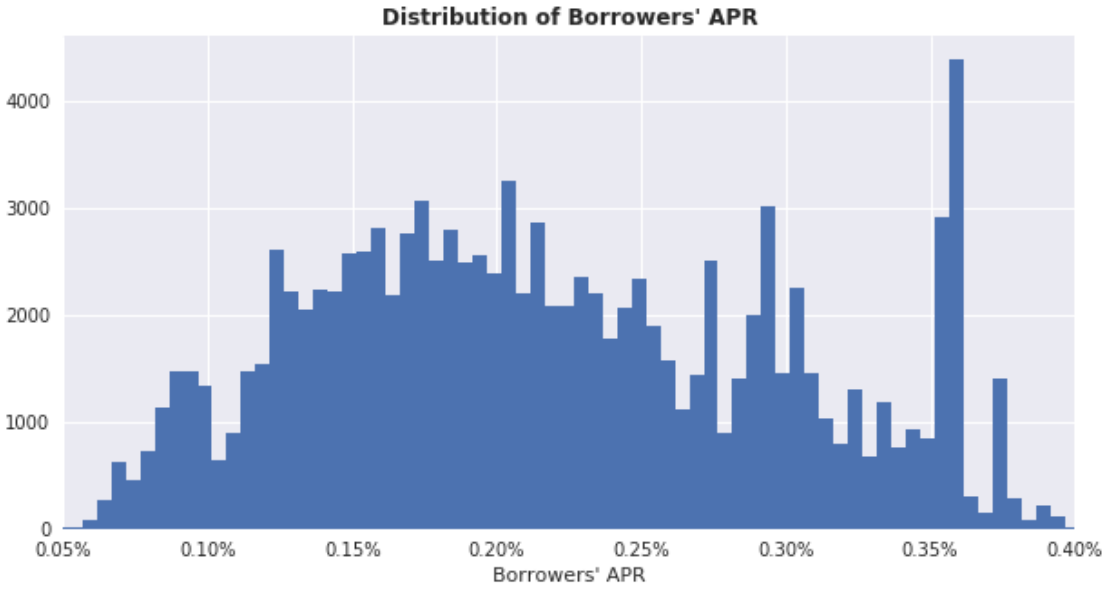
1.5 Borrowers' profiles

The borrowers were mostly employed persons with only a tiny fraction of them being unemployed and retired. The borrowers APR which is a percentage representation of the sum total ionterest a borrower was made to pay per loan on a histogram has a unimodal distribution that is skewed to the right with a very sharp spike at around 0.36%.

In [7]:



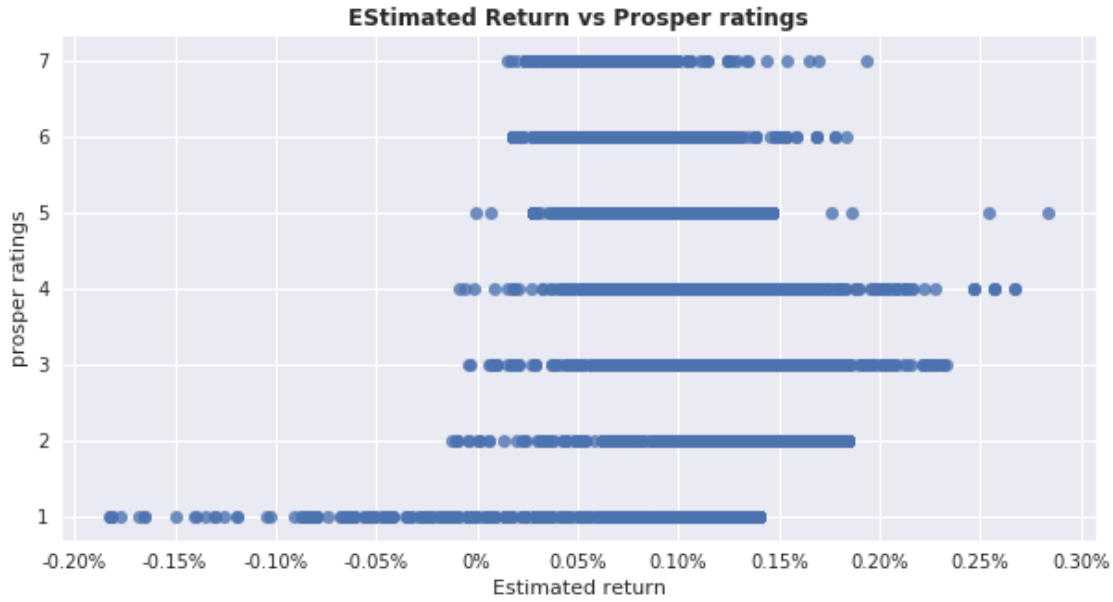
In [8]:



1.6 Prosper ratings vs positive and negative Estimated Return Loans.

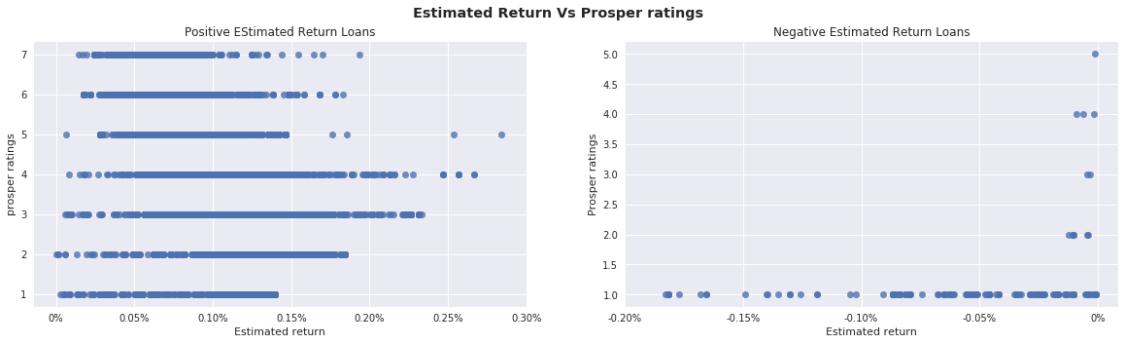
The company's risk grading of borrowers which is the Prosper rating when plotted against the company's estimated returns on the loans would first reveal that some loan have negative Estimated Returns of up to 0.2%. Representing these two different categories of Estimated returns would go further to reveal that loans with positive Estimaeed returns had prosper ratings spread from 1 - 7 representing bith high risk and low risk borrowers. Whereas, the loans with negative estimated returns were acquired by high risk borrowers with prosper rating of 1.

In [9]:



In [10]:

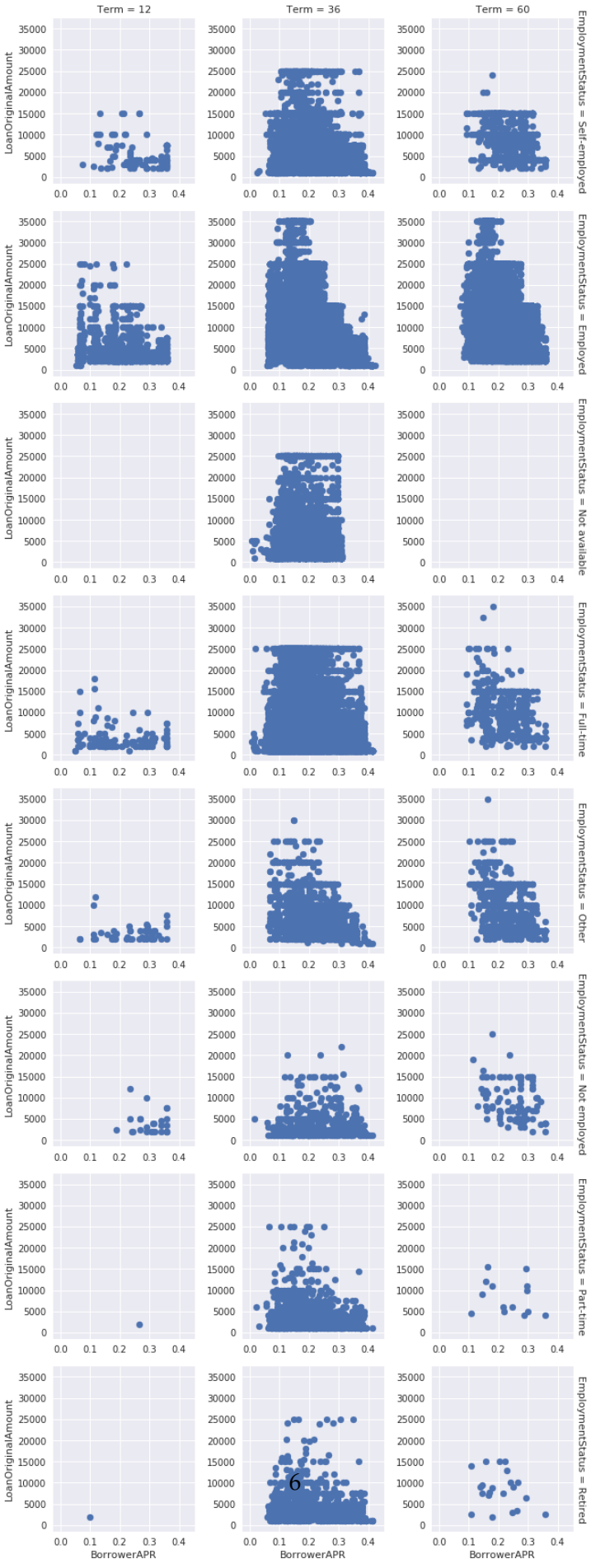
In [11]:



1.7 Loan Amount and Loan Term by Borrower APR and Employment Status.

A facet grid of these two Loan variables and two borrower characteristics would reveal a pattern that almost all the borrowers regardless of Employment status often times preferred taking longer term loans with most of the loans falling in the 36 month term category. There was also a maximum confinement on 36 months Term loans at around 25,000USD for almost all borrower employment status except borrowers in the Employed category who had dense points above that amount.

In [12]:


Term = 12
Term = 36
Term = 60
EmploymentStatus = Self-employed
EmploymentStatus = Employed
EmploymentStatus = Not available
EmploymentStatus = Full-time
EmploymentStatus = Other
EmploymentStatus = Not employed
EmploymentStatus = Part-time
EmploymentStatus = Retired
LoanOriginalAmount
BorrowerAPR