

# Exploring Prosper Loan (P2P) by Min Lu

## Introduction:

Prosper is a P2P marketplace providing person-to-person lending utilizing a listing and bidding process to get competitive rates for loans.

In this project, I will use R and data analyst skills to explore relationships between multiple variables.

## Explore Dataset

```
## 'data.frame': 113937 obs. of 81 variables:
## $ ListingKey : Factor w/ 113066 levels
"00003546482094282EF90E5",...: 7180 7193 6647 6669 6686 6689 6699 6706 6687
6687 ...
## $ ListingNumber : int 193129 1209647 81716 658116
909464 1074836 750899 768193 1023355 1023355 ...
## $ ListingCreationDate : Factor w/ 113064 levels "2005-11-
09 20:44:28.847000000",...: 14184 111894 6429 64760 85967 100310 72556 74019
97834 97834 ...
## $ CreditGrade : Factor w/ 9 levels
"", "A", "AA", "B",...: 5 1 8 1 1 1 1 1 1 ...
## $ Term : int 36 36 36 36 36 60 36 36 36
36 ...
## $ LoanStatus : Factor w/ 12 levels
"Cancelled", "Chargedoff",...: 3 4 3 4 4 4 4 4 4 ...
## $ ClosedDate : Factor w/ 2803 levels "", "2005-11-
25 00:00:00",...: 1138 1 1263 1 1 1 1 1 1 ...
## $ BorrowerAPR : num 0.165 0.12 0.283 0.125
0.246 ...
## $ BorrowerRate : num 0.158 0.092 0.275 0.0974
0.2085 ...
## $ LenderYield : num 0.138 0.082 0.24 0.0874
0.1985 ...
## $ EstimatedEffectiveYield : num NA 0.0796 NA 0.0849
0.1832 ...
## $ EstimatedLoss : num NA 0.0249 NA 0.0249
0.0925 ...
## $ EstimatedReturn : num NA 0.0547 NA 0.06 0.0907 ...
## $ ProsperRating..numeric. : int NA 6 NA 6 3 5 2 4 7 7 ...
## $ ProsperRating..Alpha. : Factor w/ 8 levels
"", "A", "AA", "B",...: 1 2 1 2 6 4 7 5 3 3 ...
## $ ProsperScore : num NA 7 NA 9 4 10 2 4 9 11 ...
## $ ListingCategory..numeric. : int 0 2 0 16 2 1 1 2 7 7 ...
## $ BorrowerState : Factor w/ 52 levels
"", "AK", "AL", "AR",...: 7 7 12 12 25 34 18 6 16 16 ...
## $ Occupation : Factor w/ 68 levels
"", "Accountant/CPA",...: 37 43 37 52 21 43 50 29 24 24 ...
```

```

## $ EmploymentStatus : Factor w/ 9 levels
"", "Employed", ...: 9 2 4 2 2 2 2 2 2 ...
## $ EmploymentStatusDuration : int 2 44 NA 113 44 82 172 103 269
269 ...
## $ IsBorrowerHomeowner : Factor w/ 2 levels "False", "True":
2 1 1 2 2 2 1 1 2 2 ...
## $ CurrentlyInGroup : Factor w/ 2 levels "False", "True":
2 1 2 1 1 1 1 1 1 ...
## $ GroupKey : Factor w/ 707 levels
"", "00343376901312423168731", ...: 1 1 335 1 1 1 1 1 1 ...
## $ DateCreditPulled : Factor w/ 112992 levels "2005-11-
09 00:30:04.487000000", ...: 14347 111883 6446 64724 85857 100382 72500 73937
97888 97888 ...
## $ CreditScoreRangeLower : int 640 680 480 800 680 740 680
700 820 820 ...
## $ CreditScoreRangeUpper : int 659 699 499 819 699 759 699
719 839 839 ...
## $ FirstRecordedCreditLine : Factor w/ 11586 levels "", "1947-
08-24 00:00:00", ...: 8639 6617 8927 2247 9498 497 8265 7685 5543 5543 ...
## $ CurrentCreditLines : int 5 14 NA 5 19 21 10 6 17
17 ...
## $ OpenCreditLines : int 4 14 NA 5 19 17 7 6 16 16 ...
## $ TotalCreditLinespast7years : int 12 29 3 29 49 49 20 10 32
32 ...
## $ OpenRevolvingAccounts : int 1 13 0 7 6 13 6 5 12 12 ...
## $ OpenRevolvingMonthlyPayment : num 24 389 0 115 220 1410 214 101
219 219 ...
## $ InquiriesLast6Months : int 3 3 0 0 1 0 0 3 1 1 ...
## $ TotalInquiries : num 3 5 1 1 9 2 0 16 6 6 ...
## $ CurrentDelinquencies : int 2 0 1 4 0 0 0 0 0 0 ...
## $ AmountDelinquent : num 472 0 NA 10056 0 ...
## $ DelinquenciesLast7Years : int 4 0 0 14 0 0 0 0 0 0 ...
## $ PublicRecordsLast10Years : int 0 1 0 0 0 0 0 1 0 0 ...
## $ PublicRecordsLast12Months : int 0 0 NA 0 0 0 0 0 0 0 ...
## $ RevolvingCreditBalance : num 0 3989 NA 1444 6193 ...
## $ BankcardUtilization : num 0 0.21 NA 0.04 0.81 0.39 0.72
0.13 0.11 0.11 ...
## $ AvailableBankcardCredit : num 1500 10266 NA 30754 695 ...
## $ TotalTrades : num 11 29 NA 26 39 47 16 10 29
29 ...
## $ TradesNeverDelinquent..percentage. : num 0.81 1 NA 0.76 0.95 1 0.68
0.8 1 1 ...
## $ TradesOpenedLast6Months : num 0 2 NA 0 2 0 0 0 1 1 ...
## $ DebtToIncomeRatio : num 0.17 0.18 0.06 0.15 0.26 0.36
0.27 0.24 0.25 0.25 ...
## $ IncomeRange : Factor w/ 8 levels "$0", "$1-
24,999", ...: 4 5 7 4 3 3 4 4 4 4 ...
## $ IncomeVerifiable : Factor w/ 2 levels "False", "True":
2 2 2 2 2 2 2 2 ...
## $ StatedMonthlyIncome : num 3083 6125 2083 2875 9583 ...
## $ LoanKey : Factor w/ 113066 levels
"00003683605746079487FF7", ...: 100337 69837 46303 70776 71387 86505 91250 5425
908 908 ...
## $ TotalProsperLoans : int NA NA NA NA 1 NA NA NA NA
NA ...
## $ TotalProsperPaymentsBilled : int NA NA NA NA 11 NA NA NA NA
NA ...

```

```

## $ OnTimeProsperPayments          : int  NA NA NA NA 11 NA NA NA NA
NA ...
## $ ProsperPaymentsLessThanOneMonthLate: int  NA NA NA NA 0 NA NA NA NA
NA ...
## $ ProsperPaymentsOneMonthPlusLate   : int  NA NA NA NA 0 NA NA NA NA
NA ...
## $ ProsperPrincipalBorrowed          : num  NA NA NA NA 11000 NA NA NA NA
NA ...
## $ ProsperPrincipalOutstanding        : num  NA NA NA NA 9948 ...
## $ ScorexChangeAtTimeOfListing        : int  NA NA NA NA NA NA NA NA NA
NA ...
## $ LoanCurrentDaysDelinquent          : int  0 0 0 0 0 0 0 0 0 0 ...
## $ LoanFirstDefaultedCycleNumber      : int  NA NA NA NA NA NA NA NA NA
NA ...
## $ LoanMonthsSinceOrigination         : int  78 0 86 16 6 3 11 10 3 3 ...
## $ LoanNumber                        : int  19141 134815 6466 77296
102670 123257 88353 90051 121268 121268 ...
## $ LoanOriginalAmount                : int  9425 10000 3001 10000 15000
15000 3000 10000 10000 10000 ...
## $ LoanOriginationDate                : Factor w/ 1873 levels "2005-11-15
00:00:00",...: 426 1866 260 1535 1757 1821 1649 1666 1813 1813 ...
## $ LoanOriginationQuarter             : Factor w/ 33 levels "Q1 2006","Q1
2007",...: 18 8 2 32 24 33 16 16 33 33 ...
## $ MemberKey                         : Factor w/ 90831 levels
"00003397697413387CAF966",...: 11071 10302 33781 54939 19465 48037 60448 40951
26129 26129 ...
## $ MonthlyLoanPayment                : num  330 319 123 321 564 ...
## $ LP_CustomerPayments                : num  11396 0 4187 5143 2820 ...
## $ LP_CustomerPrincipalPayments       : num  9425 0 3001 4091 1563 ...
## $ LP_InterestandFees                 : num  1971 0 1186 1052 1257 ...
## $ LP_ServiceFees                     : num  -133.2 0 -24.2 -108 -60.3 ...
## $ LP_CollectionFees                  : num  0 0 0 0 0 0 0 0 0 0 ...
## $ LP_GrossPrincipalLoss               : num  0 0 0 0 0 0 0 0 0 0 ...
## $ LP_NetPrincipalLoss                 : num  0 0 0 0 0 0 0 0 0 0 ...
## $ LP_NonPrincipalRecoverypayments     : num  0 0 0 0 0 0 0 0 0 0 ...
## $ PercentFunded                      : num  1 1 1 1 1 1 1 1 1 1 ...
## $ Recommendations                    : int  0 0 0 0 0 0 0 0 0 0 ...
## $ InvestmentFromFriendsCount          : int  0 0 0 0 0 0 0 0 0 0 ...
## $ InvestmentFromFriendsAmount         : num  0 0 0 0 0 0 0 0 0 0 ...
## $ Investors                          : int  258 1 41 158 20 1 1 1 1 1 ...
## [1] "ListingKey"
## [2] "ListingNumber"
## [3] "ListingCreationDate"
## [4] "CreditGrade"
## [5] "Term"
## [6] "LoanStatus"
## [7] "ClosedDate"
## [8] "BorrowerAPR"
## [9] "BorrowerRate"
## [10] "LenderYield"
## [11] "EstimatedEffectiveYield"
## [12] "EstimatedLoss"
## [13] "EstimatedReturn"
## [14] "ProsperRating..numeric."
## [15] "ProsperRating..Alpha."
## [16] "ProsperScore"
## [17] "ListingCategory..numeric."

```

```
## [18] "BorrowerState"
## [19] "Occupation"
## [20] "EmploymentStatus"
## [21] "EmploymentStatusDuration"
## [22] "IsBorrowerHomeowner"
## [23] "CurrentlyInGroup"
## [24] "GroupKey"
## [25] "DateCreditPulled"
## [26] "CreditScoreRangeLower"
## [27] "CreditScoreRangeUpper"
## [28] "FirstRecordedCreditLine"
## [29] "CurrentCreditLines"
## [30] "OpenCreditLines"
## [31] "TotalCreditLinespast7years"
## [32] "OpenRevolvingAccounts"
## [33] "OpenRevolvingMonthlyPayment"
## [34] "InquiriesLast6Months"
## [35] "TotalInquiries"
## [36] "CurrentDelinquencies"
## [37] "AmountDelinquent"
## [38] "DelinquenciesLast7Years"
## [39] "PublicRecordsLast10Years"
## [40] "PublicRecordsLast12Months"
## [41] "RevolvingCreditBalance"
## [42] "BankcardUtilization"
## [43] "AvailableBankcardCredit"
## [44] "TotalTrades"
## [45] "TradesNeverDelinquent..percentage."
## [46] "TradesOpenedLast6Months"
## [47] "DebtToIncomeRatio"
## [48] "IncomeRange"
## [49] "IncomeVerifiable"
## [50] "StatedMonthlyIncome"
## [51] "LoanKey"
## [52] "TotalProsperLoans"
## [53] "TotalProsperPaymentsBilled"
## [54] "OnTimeProsperPayments"
## [55] "ProsperPaymentsLessThanOneMonthLate"
## [56] "ProsperPaymentsOneMonthPlusLate"
## [57] "ProsperPrincipalBorrowed"
## [58] "ProsperPrincipalOutstanding"
## [59] "ScorexChangeAtTimeOfListing"
## [60] "LoanCurrentDaysDelinquent"
## [61] "LoanFirstDefaultedCycleNumber"
## [62] "LoanMonthsSinceOrigination"
## [63] "LoanNumber"
## [64] "LoanOriginalAmount"
## [65] "LoanOriginationDate"
## [66] "LoanOriginationQuarter"
## [67] "MemberKey"
## [68] "MonthlyLoanPayment"
## [69] "LP_CustomerPayments"
## [70] "LP_CustomerPrincipalPayments"
## [71] "LP_InterestandFees"
## [72] "LP_ServiceFees"
## [73] "LP_CollectionFees"
## [74] "LP_GrossPrincipalLoss"
```

```

## [75] "LP_NetPrincipalLoss"
## [76] "LP_NonPrincipalRecoverypayments"
## [77] "PercentFunded"
## [78] "Recommendations"
## [79] "InvestmentFromFriendsCount"
## [80] "InvestmentFromFriendsAmount"
## [81] "Investors"
## [1] 113937
## [1] 81
##
##           ListingKey      ListingNumber
## 17A93590655669644DB4C06:      6      Min.      :      4
## 349D3587495831350F0F648:      4      1st Qu.: 400919
## 47C1359638497431975670B:      4      Median : 600554
## 8474358854651984137201C:      4      Mean   : 627886
## DE8535960513435199406CE:      4      3rd Qu.: 892634
## 04C13599434217079754AEE:      3      Max.    :1255725
## (Other)                :113912
##
##           ListingCreationDate  CreditGrade      Term
## 2013-10-02 17:20:16.550000000:      6           :84984      Min.    :12.00
## 2013-08-28 20:31:41.107000000:      4      C           : 5649      1st Qu.:36.00
## 2013-09-08 09:27:44.853000000:      4      D           : 5153      Median :36.00
## 2013-12-06 05:43:13.830000000:      4      B           : 4389      Mean   :40.83
## 2013-12-06 11:44:58.283000000:      4      AA          : 3509      3rd Qu.:36.00
## 2013-08-21 07:25:22.360000000:      3      HR          : 3508      Max.    :60.00
## (Other)                :113912      (Other): 6745
##
##           LoanStatus      ClosedDate
## Current           :56576           :58848
## Completed         :38074      2014-03-04 00:00:00: 105
## Chargedoff        :11992      2014-02-19 00:00:00: 100
## Defaulted         : 5018      2014-02-11 00:00:00: 92
## Past Due (1-15 days) : 806      2012-10-30 00:00:00: 81
## Past Due (31-60 days): 363      2013-02-26 00:00:00: 78
## (Other)           : 1108      (Other)           :54633
##
## BorrowerAPR      BorrowerRate      LenderYield
## Min.    :0.00653      Min.    :0.0000      Min.    :-0.0100
## 1st Qu.:0.15629      1st Qu.:0.1340      1st Qu.: 0.1242
## Median :0.20976      Median :0.1840      Median : 0.1730
## Mean   :0.21883      Mean   :0.1928      Mean   : 0.1827
## 3rd Qu.:0.28381      3rd Qu.:0.2500      3rd Qu.: 0.2400
## Max.    :0.51229      Max.    :0.4975      Max.    : 0.4925
## NA's    :25
## EstimatedEffectiveYield EstimatedLoss      EstimatedReturn
## Min.    :-0.183      Min.    :0.005      Min.    :-0.183
## 1st Qu.: 0.116      1st Qu.:0.042      1st Qu.: 0.074
## Median : 0.162      Median :0.072      Median : 0.092
## Mean   : 0.169      Mean   :0.080      Mean   : 0.096
## 3rd Qu.: 0.224      3rd Qu.:0.112      3rd Qu.: 0.117
## Max.    : 0.320      Max.    :0.366      Max.    : 0.284
## NA's    :29084      NA's    :29084      NA's    :29084
## ProsperRating..numeric. ProsperRating..Alpha. ProsperScore
## Min.    :1.000      :29084      Min.    : 1.00
## 1st Qu.:3.000      C      :18345      1st Qu.: 4.00
## Median :4.000      B      :15581      Median : 6.00
## Mean   :4.072      A      :14551      Mean   : 5.95
## 3rd Qu.:5.000      D      :14274      3rd Qu.: 8.00
## Max.    :7.000      E      : 9795      Max.    :11.00
## NA's    :29084      (Other):12307      NA's    :29084

```

```

## ListingCategory..numeric. BorrowerState
## Min. : 0.000 CA :14717
## 1st Qu.: 1.000 TX : 6842
## Median : 1.000 NY : 6729
## Mean : 2.774 FL : 6720
## 3rd Qu.: 3.000 IL : 5921
## Max. :20.000 : 5515
## (Other):67493
## Occupation EmploymentStatus
## Other :28617 Employed :67322
## Professional :13628 Full-time :26355
## Computer Programmer : 4478 Self-employed: 6134
## Executive : 4311 Not available: 5347
## Teacher : 3759 Other : 3806
## Administrative Assistant: 3688 : 2255
## (Other) :55456 (Other) : 2718
## EmploymentStatusDuration IsBorrowerHomeowner CurrentlyInGroup
## Min. : 0.00 False:56459 False:101218
## 1st Qu.: 26.00 True :57478 True : 12719
## Median : 67.00
## Mean : 96.07
## 3rd Qu.:137.00
## Max. :755.00
## NA's :7625
## GroupKey DateCreditPulled
## :100596 2013-12-23 09:38:12: 6
## 783C3371218786870A73D20: 1140 2013-11-21 09:09:41: 4
## 3D4D3366260257624AB272D: 916 2013-12-06 05:43:16: 4
## 6A3B336601725506917317E: 698 2014-01-14 20:17:49: 4
## FEF83377364176536637E50: 611 2014-02-09 12:14:41: 4
## C9643379247860156A00EC0: 342 2013-09-27 22:04:54: 3
## (Other) : 9634 (Other) :113912
## CreditScoreRangeLower CreditScoreRangeUpper
## Min. : 0.0 Min. : 19.0
## 1st Qu.:660.0 1st Qu.:679.0
## Median :680.0 Median :699.0
## Mean :685.6 Mean :704.6
## 3rd Qu.:720.0 3rd Qu.:739.0
## Max. :880.0 Max. :899.0
## NA's :591 NA's :591
## FirstRecordedCreditLine CurrentCreditLines OpenCreditLines
## : 697 Min. : 0.00 Min. : 0.00
## 1993-12-01 00:00:00: 185 1st Qu.: 7.00 1st Qu.: 6.00
## 1994-11-01 00:00:00: 178 Median :10.00 Median : 9.00
## 1995-11-01 00:00:00: 168 Mean :10.32 Mean : 9.26
## 1990-04-01 00:00:00: 161 3rd Qu.:13.00 3rd Qu.:12.00
## 1995-03-01 00:00:00: 159 Max. :59.00 Max. :54.00
## (Other) :112389 NA's :7604 NA's :7604
## TotalCreditLinespast7years OpenRevolvingAccounts
## Min. : 2.00 Min. : 0.00
## 1st Qu.: 17.00 1st Qu.: 4.00
## Median : 25.00 Median : 6.00
## Mean : 26.75 Mean : 6.97
## 3rd Qu.: 35.00 3rd Qu.: 9.00
## Max. :136.00 Max. :51.00
## NA's :697
## OpenRevolvingMonthlyPayment InquiriesLast6Months TotalInquiries

```

|    |                                    |   |         |                           |         |          |                         |       |         |  |  |
|----|------------------------------------|---|---------|---------------------------|---------|----------|-------------------------|-------|---------|--|--|
| ## | Min.                               | : | 0.0     | Min.                      | :       | 0.000    | Min.                    | :     | 0.000   |  |  |
| ## | 1st Qu.:                           |   | 114.0   | 1st Qu.:                  |         | 0.000    | 1st Qu.:                |       | 2.000   |  |  |
| ## | Median :                           |   | 271.0   | Median :                  |         | 1.000    | Median :                |       | 4.000   |  |  |
| ## | Mean :                             |   | 398.3   | Mean :                    |         | 1.435    | Mean :                  |       | 5.584   |  |  |
| ## | 3rd Qu.:                           |   | 525.0   | 3rd Qu.:                  |         | 2.000    | 3rd Qu.:                |       | 7.000   |  |  |
| ## | Max.                               | : | 14985.0 | Max.                      | :       | 105.000  | Max.                    | :     | 379.000 |  |  |
| ## |                                    |   |         | NA's                      | :       | 697      | NA's                    | :     | 1159    |  |  |
| ## | CurrentDelinquencies               |   |         | AmountDelinquent          |         |          | DelinquenciesLast7Years |       |         |  |  |
| ## | Min.                               | : | 0.0000  | Min.                      | :       | 0.0      | Min.                    | :     | 0.000   |  |  |
| ## | 1st Qu.:                           |   | 0.0000  | 1st Qu.:                  |         | 0.0      | 1st Qu.:                |       | 0.000   |  |  |
| ## | Median :                           |   | 0.0000  | Median :                  |         | 0.0      | Median :                |       | 0.000   |  |  |
| ## | Mean :                             |   | 0.5921  | Mean :                    |         | 984.5    | Mean :                  |       | 4.155   |  |  |
| ## | 3rd Qu.:                           |   | 0.0000  | 3rd Qu.:                  |         | 0.0      | 3rd Qu.:                |       | 3.000   |  |  |
| ## | Max.                               | : | 83.0000 | Max.                      | :       | 463881.0 | Max.                    | :     | 99.000  |  |  |
| ## | NA's                               | : | 697     | NA's                      | :       | 7622     | NA's                    | :     | 990     |  |  |
| ## | PublicRecordsLast10Years           |   |         | PublicRecordsLast12Months |         |          | RevolvingCreditBalance  |       |         |  |  |
| ## | Min.                               | : | 0.0000  | Min.                      | :       | 0.000    | Min.                    | :     | 0       |  |  |
| ## | 1st Qu.:                           |   | 0.0000  | 1st Qu.:                  |         | 0.000    | 1st Qu.:                |       | 3121    |  |  |
| ## | Median :                           |   | 0.0000  | Median :                  |         | 0.000    | Median :                |       | 8549    |  |  |
| ## | Mean :                             |   | 0.3126  | Mean :                    |         | 0.015    | Mean :                  |       | 17599   |  |  |
| ## | 3rd Qu.:                           |   | 0.0000  | 3rd Qu.:                  |         | 0.000    | 3rd Qu.:                |       | 19521   |  |  |
| ## | Max.                               | : | 38.0000 | Max.                      | :       | 20.000   | Max.                    | :     | 1435667 |  |  |
| ## | NA's                               | : | 697     | NA's                      | :       | 7604     | NA's                    | :     | 7604    |  |  |
| ## | BankcardUtilization                |   |         | AvailableBankcardCredit   |         |          | TotalTrades             |       |         |  |  |
| ## | Min.                               | : | 0.000   | Min.                      | :       | 0        | Min.                    | :     | 0.00    |  |  |
| ## | 1st Qu.:                           |   | 0.310   | 1st Qu.:                  |         | 880      | 1st Qu.:                |       | 15.00   |  |  |
| ## | Median :                           |   | 0.600   | Median :                  |         | 4100     | Median :                |       | 22.00   |  |  |
| ## | Mean :                             |   | 0.561   | Mean :                    |         | 11210    | Mean :                  |       | 23.23   |  |  |
| ## | 3rd Qu.:                           |   | 0.840   | 3rd Qu.:                  |         | 13180    | 3rd Qu.:                |       | 30.00   |  |  |
| ## | Max.                               | : | 5.950   | Max.                      | :       | 646285   | Max.                    | :     | 126.00  |  |  |
| ## | NA's                               | : | 7604    | NA's                      | :       | 7544     | NA's                    | :     | 7544    |  |  |
| ## | TradesNeverDelinquent..percentage. |   |         | TradesOpenedLast6Months   |         |          |                         |       |         |  |  |
| ## | Min.                               | : | 0.000   | Min.                      | :       | 0.000    |                         |       |         |  |  |
| ## | 1st Qu.:                           |   | 0.820   | 1st Qu.:                  |         | 0.000    |                         |       |         |  |  |
| ## | Median :                           |   | 0.940   | Median :                  |         | 0.000    |                         |       |         |  |  |
| ## | Mean :                             |   | 0.886   | Mean :                    |         | 0.802    |                         |       |         |  |  |
| ## | 3rd Qu.:                           |   | 1.000   | 3rd Qu.:                  |         | 1.000    |                         |       |         |  |  |
| ## | Max.                               | : | 1.000   | Max.                      | :       | 20.000   |                         |       |         |  |  |
| ## | NA's                               | : | 7544    | NA's                      | :       | 7544     |                         |       |         |  |  |
| ## | DebtToIncomeRatio                  |   |         | IncomeRange               |         |          | IncomeVerifiable        |       |         |  |  |
| ## | Min.                               | : | 0.000   | \$25,000-49,999:          | 32192   | False:   | 8669                    |       |         |  |  |
| ## | 1st Qu.:                           |   | 0.140   | \$50,000-74,999:          | 31050   | True :   | 105268                  |       |         |  |  |
| ## | Median :                           |   | 0.220   | \$100,000+                | :17337  |          |                         |       |         |  |  |
| ## | Mean :                             |   | 0.276   | \$75,000-99,999:          | 16916   |          |                         |       |         |  |  |
| ## | 3rd Qu.:                           |   | 0.320   | Not displayed :           | 7741    |          |                         |       |         |  |  |
| ## | Max.                               | : | 10.010  | \$1-24,999                | : 7274  |          |                         |       |         |  |  |
| ## | NA's                               | : | 8554    | (Other)                   | : 1427  |          |                         |       |         |  |  |
| ## | StatedMonthlyIncome                |   |         | LoanKey                   |         |          | TotalProsperLoans       |       |         |  |  |
| ## | Min.                               | : | 0       | CB1B37030986463208432A1:  | 6       | Min.     | :                       | 0.00  |         |  |  |
| ## | 1st Qu.:                           |   | 3200    | 2DEE3698211017519D7333F:  | 4       | 1st Qu.: |                         | 1.00  |         |  |  |
| ## | Median :                           |   | 4667    | 9F4B37043517554537C364C:  | 4       | Median : |                         | 1.00  |         |  |  |
| ## | Mean :                             |   | 5608    | D895370150591392337ED6D:  | 4       | Mean :   |                         | 1.42  |         |  |  |
| ## | 3rd Qu.:                           |   | 6825    | E6FB37073953690388BC56D:  | 4       | 3rd Qu.: |                         | 2.00  |         |  |  |
| ## | Max.                               | : | 1750003 | 0D8F37036734373301ED419:  | 3       | Max.     | :                       | 8.00  |         |  |  |
| ## |                                    |   |         | (Other)                   | :113912 | NA's     | :                       | 91852 |         |  |  |
| ## | TotalProsperPaymentsBilled         |   |         | OnTimeProsperPayments     |         |          |                         |       |         |  |  |
| ## | Min.                               | : | 0.00    | Min.                      | :       | 0.00     |                         |       |         |  |  |

|    |   |         |                      |                   |
|----|---|---------|----------------------|-------------------|
| ## | 1st Qu.:  | 9.00    | 1st Qu.:             | 9.00              |
| ## | Median :  | 16.00   | Median :             | 15.00             |
| ## | Mean :  | 22.93   | Mean :               | 22.27             |
| ## | 3rd Qu.:  | 33.00   | 3rd Qu.:             | 32.00             |
| ## | Max. :  | 141.00  | Max. :               | 141.00            |
| ## | NA's :  | 91852   | NA's :               | 91852             |
| ## | ProsperPaymentsLessThanOneMonthLate ProsperPaymentsOneMonthPlusLate |         |                      |                   |
| ## | Min. :  | 0.00    | Min. :               | 0.00              |
| ## | 1st Qu.:  | 0.00    | 1st Qu.:             | 0.00              |
| ## | Median :  | 0.00    | Median :             | 0.00              |
| ## | Mean :  | 0.61    | Mean :               | 0.05              |
| ## | 3rd Qu.:  | 0.00    | 3rd Qu.:             | 0.00              |
| ## | Max. :  | 42.00   | Max. :               | 21.00             |
| ## | NA's :  | 91852   | NA's :               | 91852             |
| ## | ProsperPrincipalBorrowed ProsperPrincipalOutstanding                |         |                      |                   |
| ## | Min. :  | 0       | Min. :               | 0                 |
| ## | 1st Qu.:  | 3500    | 1st Qu.:             | 0                 |
| ## | Median :  | 6000    | Median :             | 1627              |
| ## | Mean :  | 8472    | Mean :               | 2930              |
| ## | 3rd Qu.:  | 11000   | 3rd Qu.:             | 4127              |
| ## | Max. :  | 72499   | Max. :               | 23451             |
| ## | NA's :  | 91852   | NA's :               | 91852             |
| ## | ScorexChangeAtTimeOfListing LoanCurrentDaysDelinquent               |         |                      |                   |
| ## | Min. :  | -209.00 | Min. :               | 0.0               |
| ## | 1st Qu.:  | -35.00  | 1st Qu.:             | 0.0               |
| ## | Median :  | -3.00   | Median :             | 0.0               |
| ## | Mean :  | -3.22   | Mean :               | 152.8             |
| ## | 3rd Qu.:  | 25.00   | 3rd Qu.:             | 0.0               |
| ## | Max. :  | 286.00  | Max. :               | 2704.0            |
| ## | NA's :  | 95009   |                      |                   |
| ## | LoanFirstDefaultedCycleNumber LoanMonthsSinceOrigination LoanNumber |         |                      |                   |
| ## | Min. :  | 0.00    | Min. :               | 0.0               |
| ## | 1st Qu.:  | 9.00    | 1st Qu.:             | 6.0               |
| ## | Median :  | 14.00   | Median :             | 21.0              |
| ## | Mean :  | 16.27   | Mean :               | 31.9              |
| ## | 3rd Qu.:  | 22.00   | 3rd Qu.:             | 65.0              |
| ## | Max. :  | 44.00   | Max. :               | 100.0             |
| ## | NA's :  | 96985   |                      |                   |
| ## | LoanOriginalAmount LoanOriginationDate LoanOriginationQuarter       |         |                      |                   |
| ## | Min. :  | 1000    | 2014-01-22 00:00:00: | 491 Q4 2013:14450 |
| ## | 1st Qu.:  | 4000    | 2013-11-13 00:00:00: | 490 Q1 2014:12172 |
| ## | Median :  | 6500    | 2014-02-19 00:00:00: | 439 Q3 2013: 9180 |
| ## | Mean :  | 8337    | 2013-10-16 00:00:00: | 434 Q2 2013: 7099 |
| ## | 3rd Qu.:  | 12000   | 2014-01-28 00:00:00: | 339 Q3 2012: 5632 |
| ## | Max. :  | 35000   | 2013-09-24 00:00:00: | 316 Q2 2012: 5061 |
| ## |   | (Other) | :111428              | (Other):60343     |
| ## | MemberKey MonthlyLoanPayment LP_CustomerPayments                    |         |                      |                   |
| ## | 63CA34120866140639431C9:  | 9       | Min. :               | 0.0               |
| ## | 16083364744933457E57FB9:  | 8       | 1st Qu.:             | 131.6             |
| ## | 3A2F3380477699707C81385:  | 8       | Median :             | 217.7             |
| ## | 4D9C3403302047712AD0CDD:  | 8       | Mean :               | 272.5             |
| ## | 739C338135235294782AE75:  | 8       | 3rd Qu.:             | 371.6             |
| ## | 7E1733653050264822FAA3D:  | 8       | Max. :               | 2251.5            |
| ## | (Other)   | :113888 |                      |                   |
| ## | LP_CustomerPrincipalPayments LP_InterestandFees LP_ServiceFees      |         |                      |                   |
| ## | Min. :  | 0.0     | Min. :               | -2.35             |
| ## | 1st Qu.:  | 500.9   | 1st Qu.:             | 274.87            |
|    |   |         | 1st Qu.:             | -73.18            |



```

## Median : 1587.5           Median : 700.84   Median : -34.44
## Mean   : 3105.5           Mean    : 1077.54  Mean    : -54.73
## 3rd Qu.: 4000.0           3rd Qu.: 1458.54  3rd Qu.: -13.92
## Max.   :35000.0           Max.    :15617.03  Max.    : 32.06
##
## LP_CollectionFees LP_GrossPrincipalLoss LP_NetPrincipalLoss
## Min.   :-9274.75   Min.    : -94.2    Min.    : -954.5
## 1st Qu.: 0.00     1st Qu.: 0.0      1st Qu.: 0.0
## Median : 0.00     Median : 0.0      Median : 0.0
## Mean   : -14.24   Mean    : 700.4    Mean    : 681.4
## 3rd Qu.: 0.00     3rd Qu.: 0.0      3rd Qu.: 0.0
## Max.   : 0.00     Max.    :25000.0   Max.    :25000.0
##
## LP_NonPrincipalRecoverypayments PercentFunded Recommendations
## Min.   : 0.00     Min.    :0.7000   Min.    : 0.00000
## 1st Qu.: 0.00     1st Qu.:1.0000   1st Qu.: 0.00000
## Median : 0.00     Median :1.0000   Median : 0.00000
## Mean   : 25.14    Mean    :0.9986   Mean    : 0.04803
## 3rd Qu.: 0.00     3rd Qu.:1.0000   3rd Qu.: 0.00000
## Max.   :21117.90  Max.    :1.0125   Max.    :39.00000
##
## InvestmentFromFriendsCount InvestmentFromFriendsAmount Investors
## Min.   : 0.00000   Min.    : 0.00    Min.    : 1.00
## 1st Qu.: 0.00000   1st Qu.: 0.00    1st Qu.: 2.00
## Median : 0.00000   Median : 0.00    Median : 44.00
## Mean   : 0.02346   Mean    : 16.55   Mean    : 80.48
## 3rd Qu.: 0.00000   3rd Qu.: 0.00    3rd Qu.: 115.00
## Max.   :33.00000   Max.    :25000.00 Max.    :1189.00
##

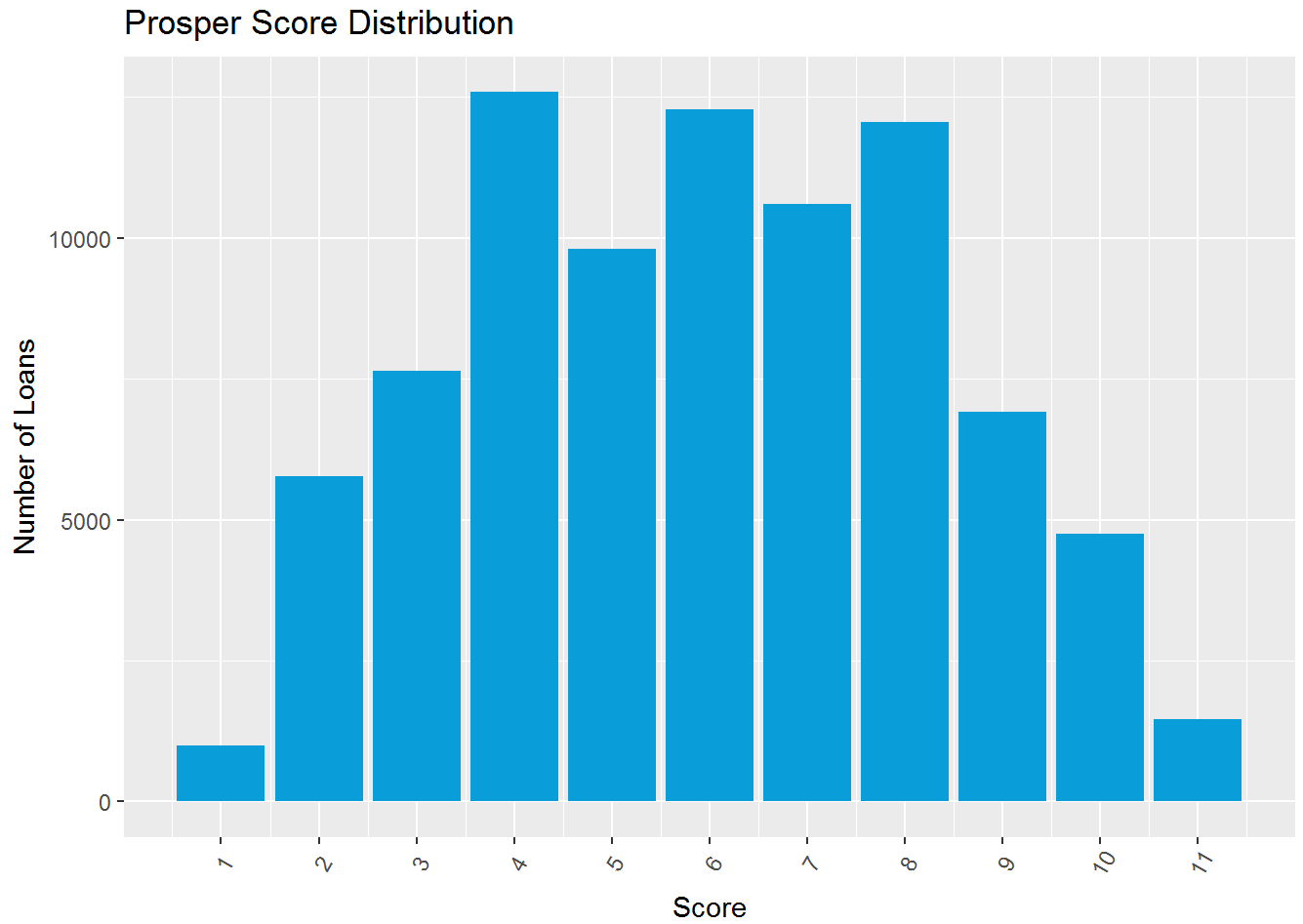
```

## Univariate Plots Section

```

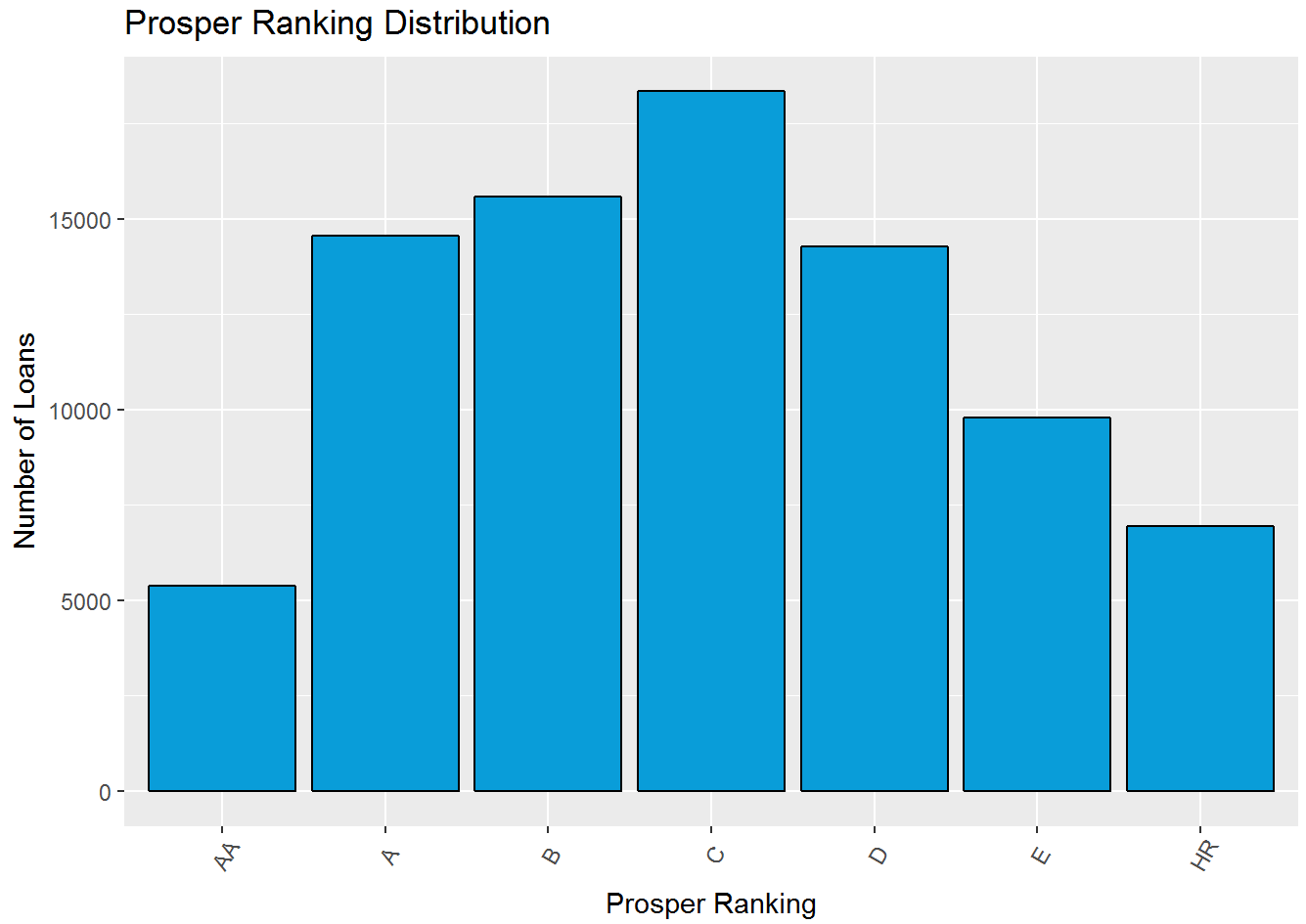
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's
##      1.00   4.00   6.00   5.95   8.00   11.00  29084
## Warning: Removed 29084 rows containing non-finite values (stat_count).

```



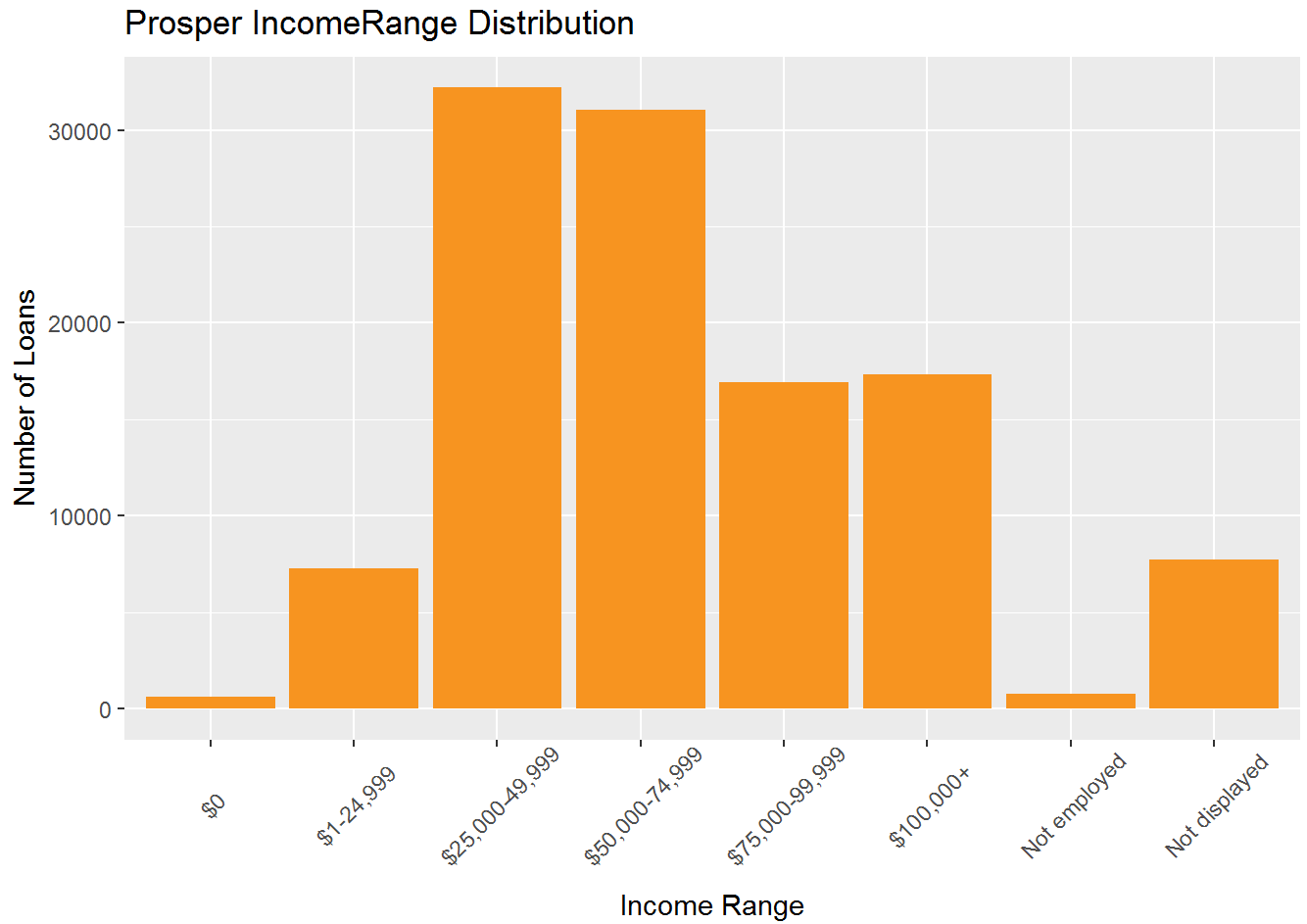
Here we can see a normal distribution chart if we did not count NA. It looks like most loans are under 4-8 score ranges.

|    |       |       |      |       |       |       |      |      |
|----|-------|-------|------|-------|-------|-------|------|------|
| ## | A     | AA    | B    | C     | D     | E     | HR   |      |
| ## | 29084 | 14551 | 5372 | 15581 | 18345 | 14274 | 9795 | 6935 |

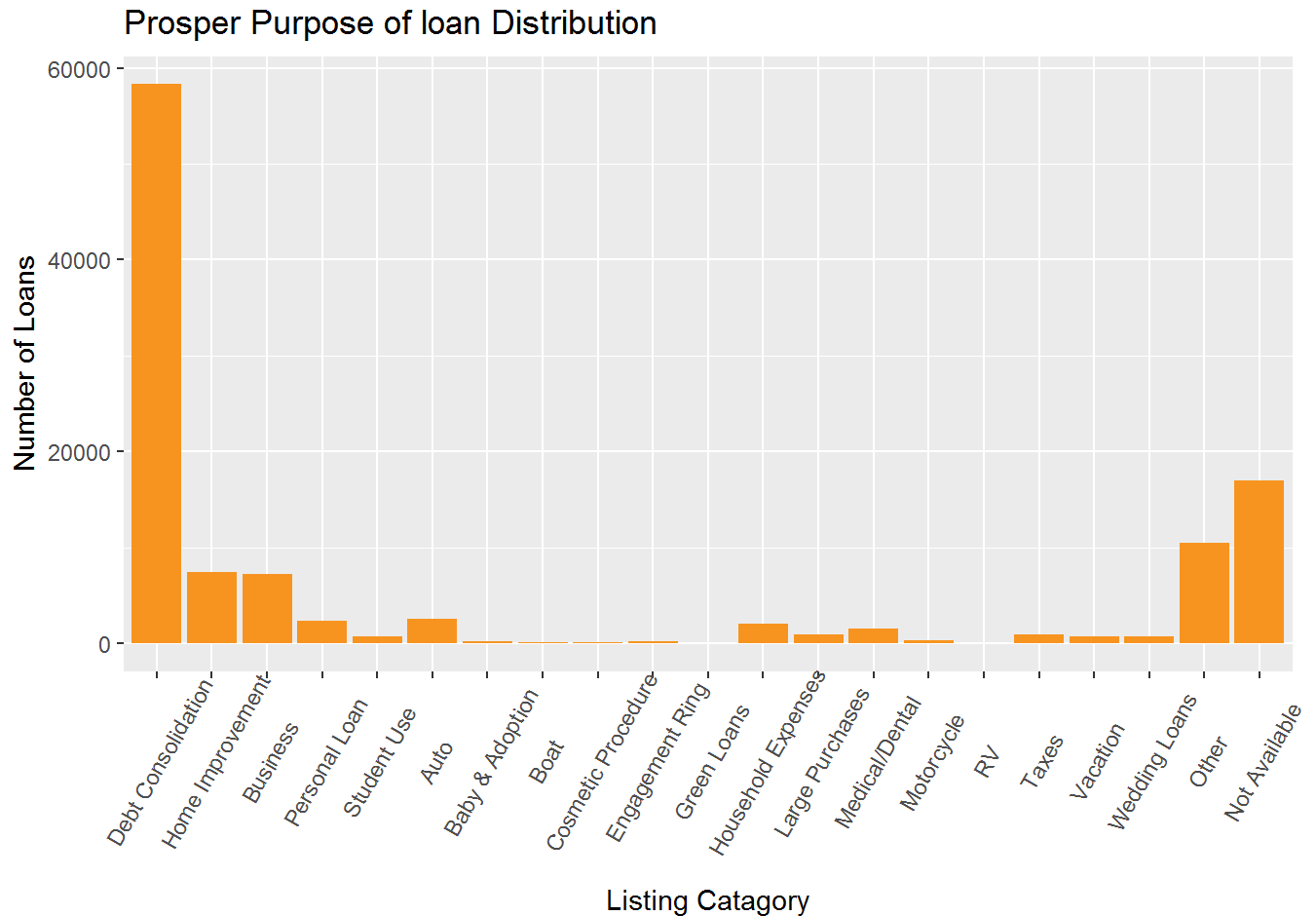


From the chart above, the majority of loan are being rated as “A”, “B”, “c” and “D”.

|    |            |              |                 |                 |                 |
|----|------------|--------------|-----------------|-----------------|-----------------|
| ## | \$0        | \$1-24,999   | \$25,000-49,999 | \$50,000-74,999 | \$75,000-99,999 |
| ## | 621        | 7274         | 32192           | 31050           | 16916           |
| ## | \$100,000+ | Not employed | Not displayed   |                 |                 |
| ## | 17337      | 806          | 7741            |                 |                 |



People from income range 25,000-49,999 and 50,000-74,999 are the main borrower. However, people with income under 25000 are not major borrower. So, probably Prosper.com has some income limits on borrowers or Prosper.com main target are future home-owners or car-owners, but not for student loan.

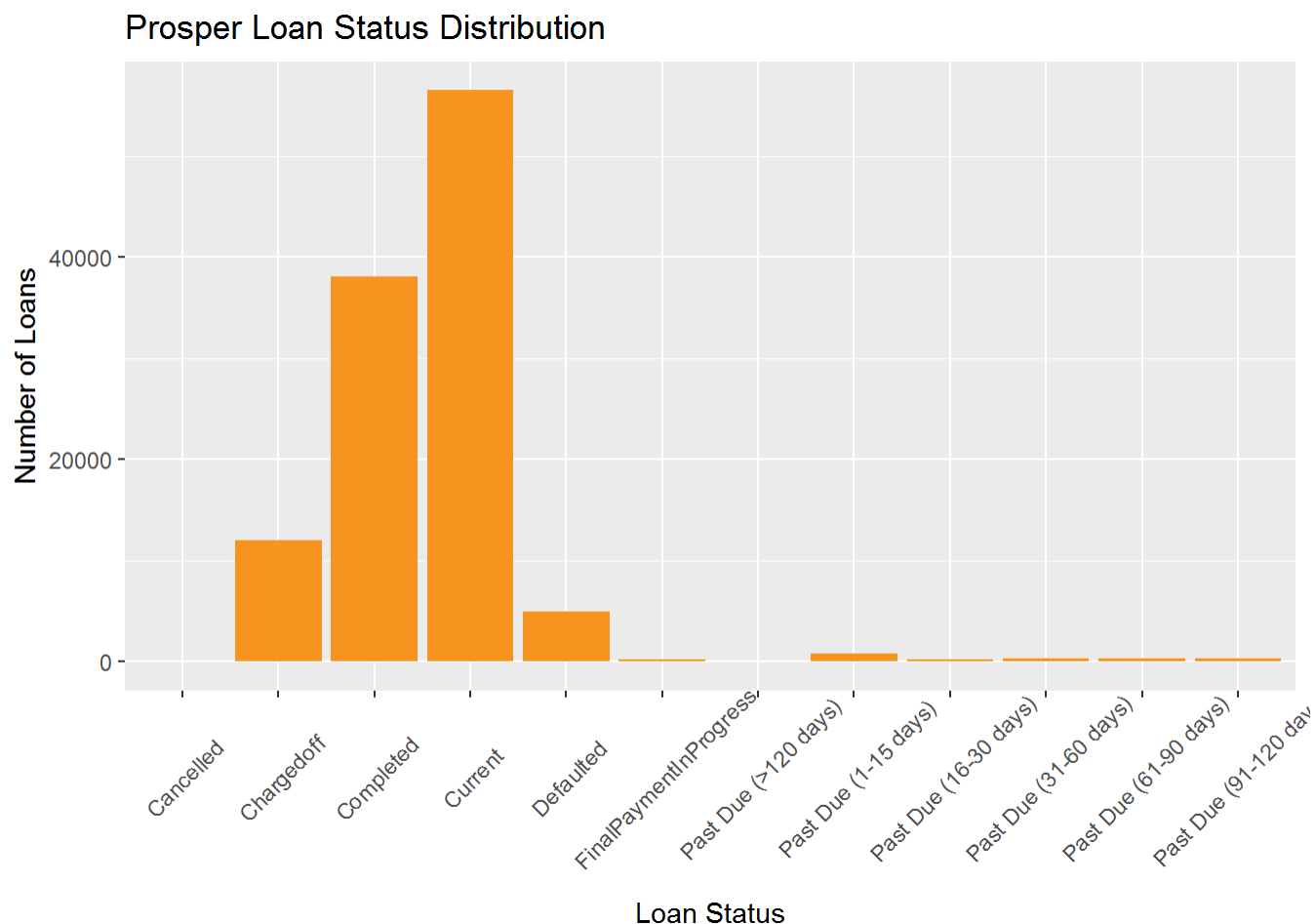


|                       |                  |                    |
|-----------------------|------------------|--------------------|
| ## Debt Consolidation | Home Improvement | Business           |
| ## 58308              | 7433             | 7189               |
| ## Personal Loan      | Student Use      | Auto               |
| ## 2395               | 756              | 2572               |
| ## Baby & Adoption    | Boat             | Cosmetic Procedure |
| ## 199                | 85               | 91                 |
| ## Engagement Ring    | Green Loans      | Household Expenses |
| ## 217                | 59               | 1996               |
| ## Large Purchases    | Medical/Dental   | Motorcycle         |
| ## 876                | 1522             | 304                |
| ## RV                 | Taxes            | Vacation           |
| ## 52                 | 885              | 768                |
| ## Wedding Loans      | Other            | Not Available      |
| ## 771                | 10494            | 16965              |

From this chart, we could see most people want money for debit consolidation, home improvement and business. Debit consolidation has a surprising amount which almost 50% of total number of loans. These data shows that most borrower already have their first job and is looking to pay their debt.

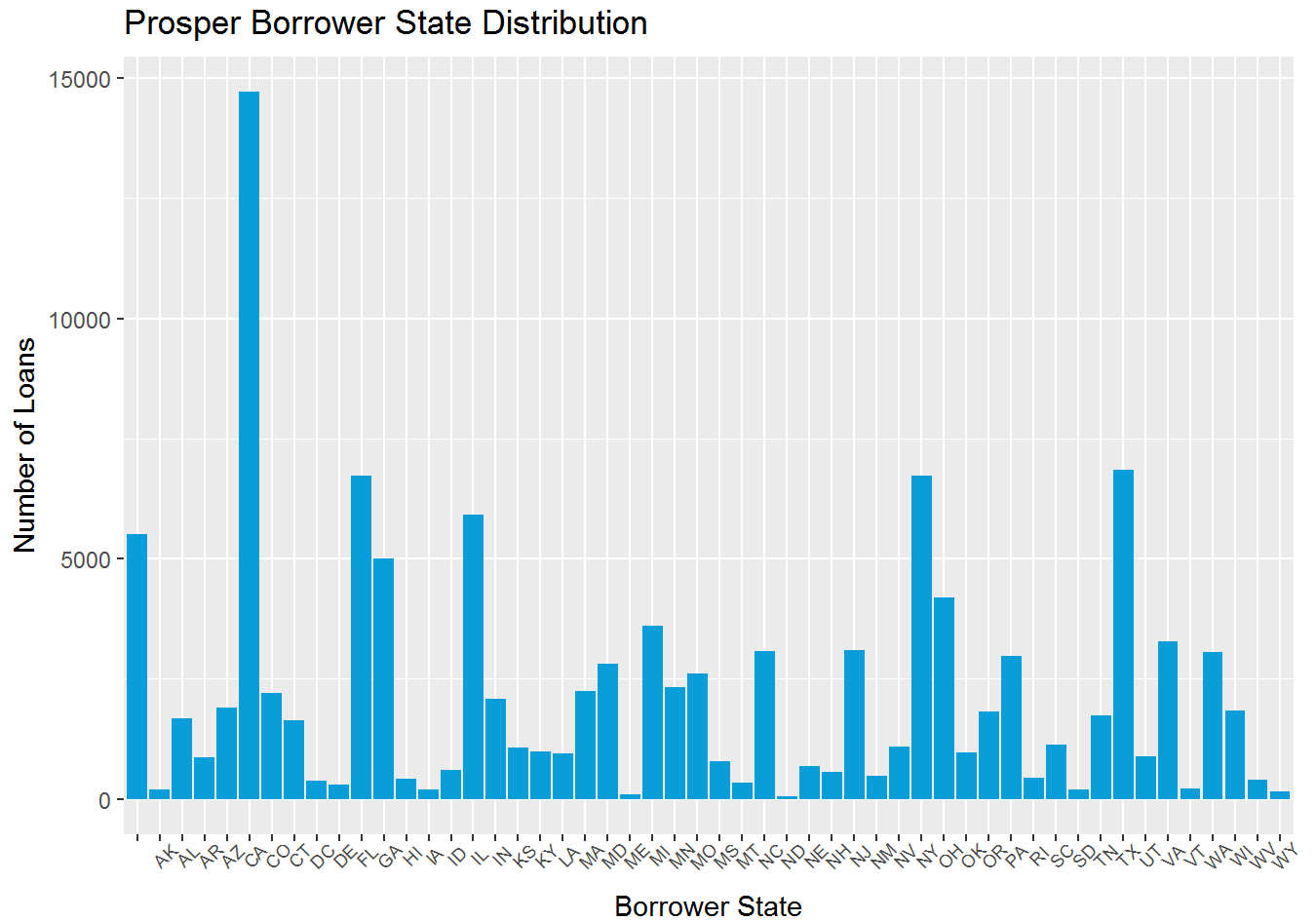
|              |            |           |
|--------------|------------|-----------|
| ## Cancelled | Chargedoff | Completed |
| ## 5         | 11992      | 38074     |

|    |                       |                       |                        |
|----|-----------------------|-----------------------|------------------------|
| ## | Current               | Defaulted             | FinalPaymentInProgress |
| ## | 56576                 | 5018                  | 205                    |
| ## | Past Due (>120 days)  | Past Due (1-15 days)  | Past Due (16-30 days)  |
| ## | 16                    | 806                   | 265                    |
| ## | Past Due (31-60 days) | Past Due (61-90 days) | Past Due (91-120 days) |
| ## | 363                   | 313                   | 304                    |



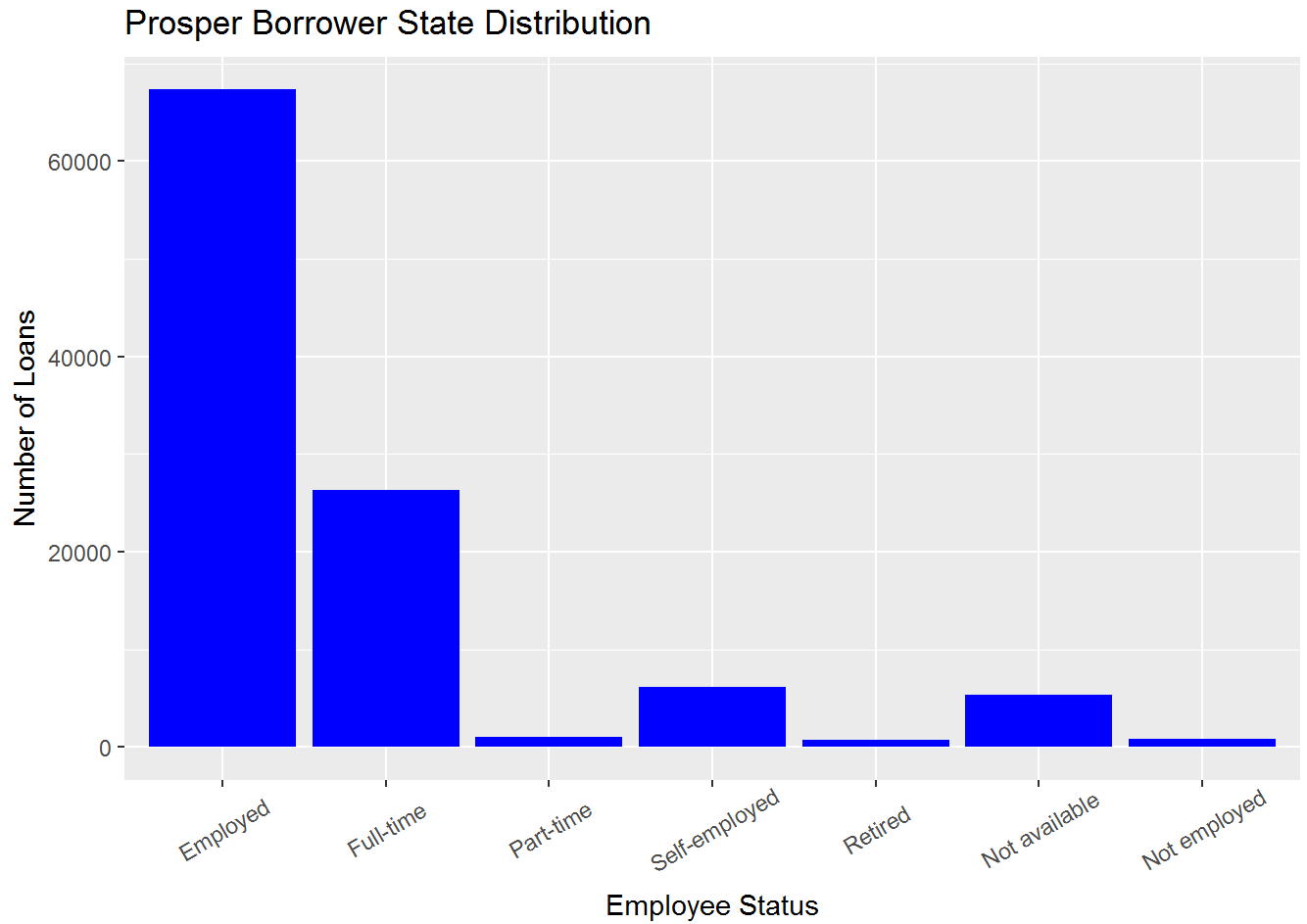
Here we can see that more than 50% of loans are under active status and almost 38% are complete. However, more than 10% loans are defaulted or past due, which probably will reduce the company revenue.

|    |      |      |      |      |      |       |      |      |      |      |      |
|----|------|------|------|------|------|-------|------|------|------|------|------|
| ## | AK   | AL   | AR   | AZ   | CA   | CO    | CT   | DC   | DE   | FL   | GA   |
| ## | 5515 | 200  | 1679 | 855  | 1901 | 14717 | 2210 | 1627 | 382  | 300  | 6720 |
| ## | HI   | IA   | ID   | IL   | IN   | KS    | KY   | LA   | MA   | MD   | ME   |
| ## | 409  | 186  | 599  | 5921 | 2078 | 1062  | 983  | 954  | 2242 | 2821 | 101  |
| ## | MN   | MO   | MS   | MT   | NC   | ND    | NE   | NH   | NJ   | NM   | NV   |
| ## | 2318 | 2615 | 787  | 330  | 3084 | 52    | 674  | 551  | 3097 | 472  | 1090 |
| ## | OH   | OK   | OR   | PA   | RI   | SC    | SD   | TN   | TX   | UT   | VA   |
| ## | 4197 | 971  | 1817 | 2972 | 435  | 1122  | 189  | 1737 | 6842 | 877  | 3278 |
| ## | WA   | WI   | WV   | WY   |      |       |      |      |      |      |      |
| ## | 3048 | 1842 | 391  | 150  |      |       |      |      |      |      |      |



From this chart, we could see the top borrower is from “CA”, which is almost 15,000 borrowers. Other popular states are “FL”, “IL”, “NY” and “TX”.

|    |               |              |           |               |         |
|----|---------------|--------------|-----------|---------------|---------|
| ## | Employed      | Full-time    | Part-time | Self-employed | Retired |
| ## | 67322         | 26355        | 1088      | 6134          | 795     |
| ## | Not available | Not employed | others    |               | NA's    |
| ## | 5347          | 835          | 0         | 0             | 6061    |

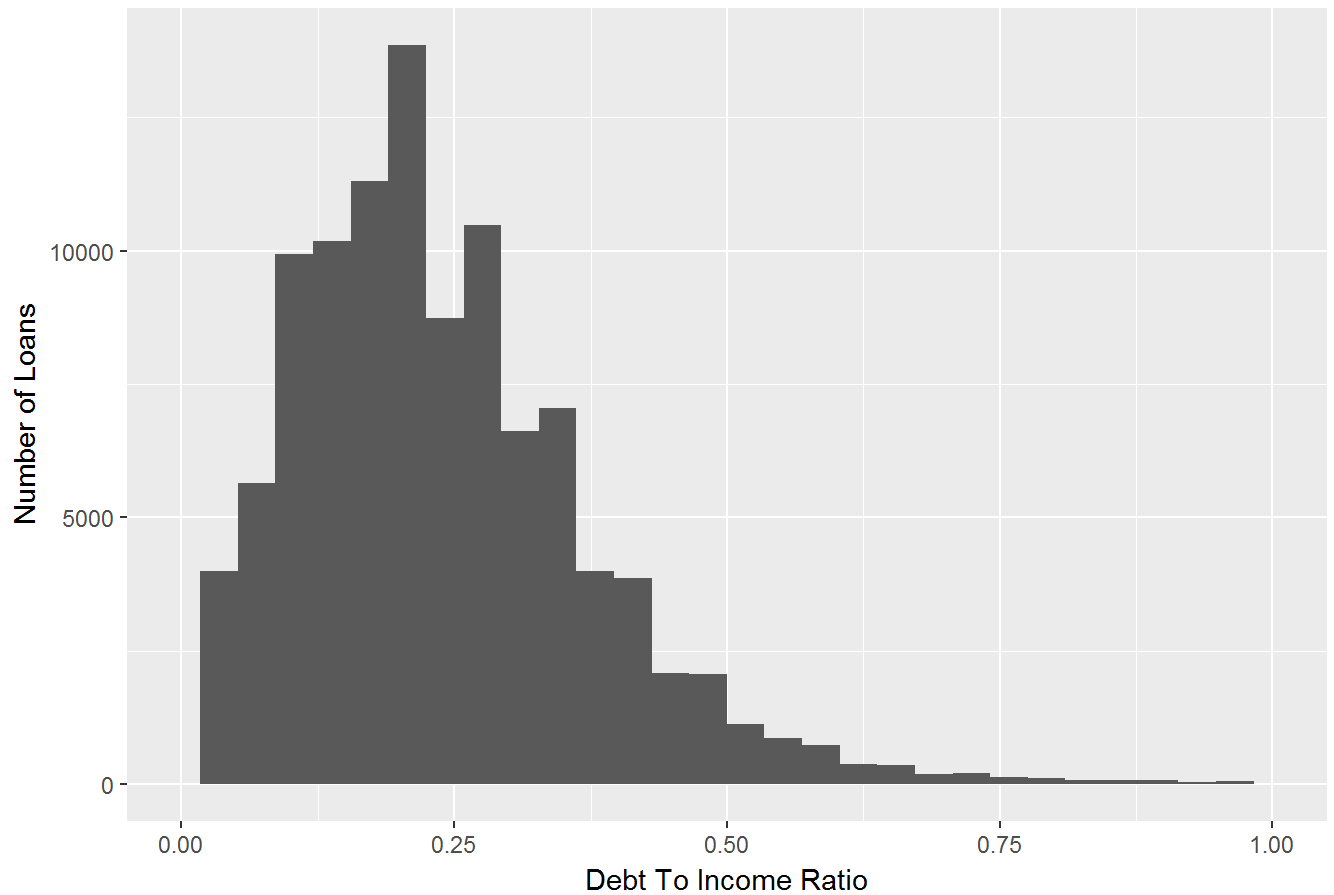


Most borrowers are employed or had a full-time job. Surprisingly, there're few part-time job borrowers. Probabaly because part-time job is not stable or do not have enough income to pay back the loan.

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
##    0.000   0.140   0.220   0.276   0.320  10.010    8554
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 9353 rows containing non-finite values (stat_bin).
```



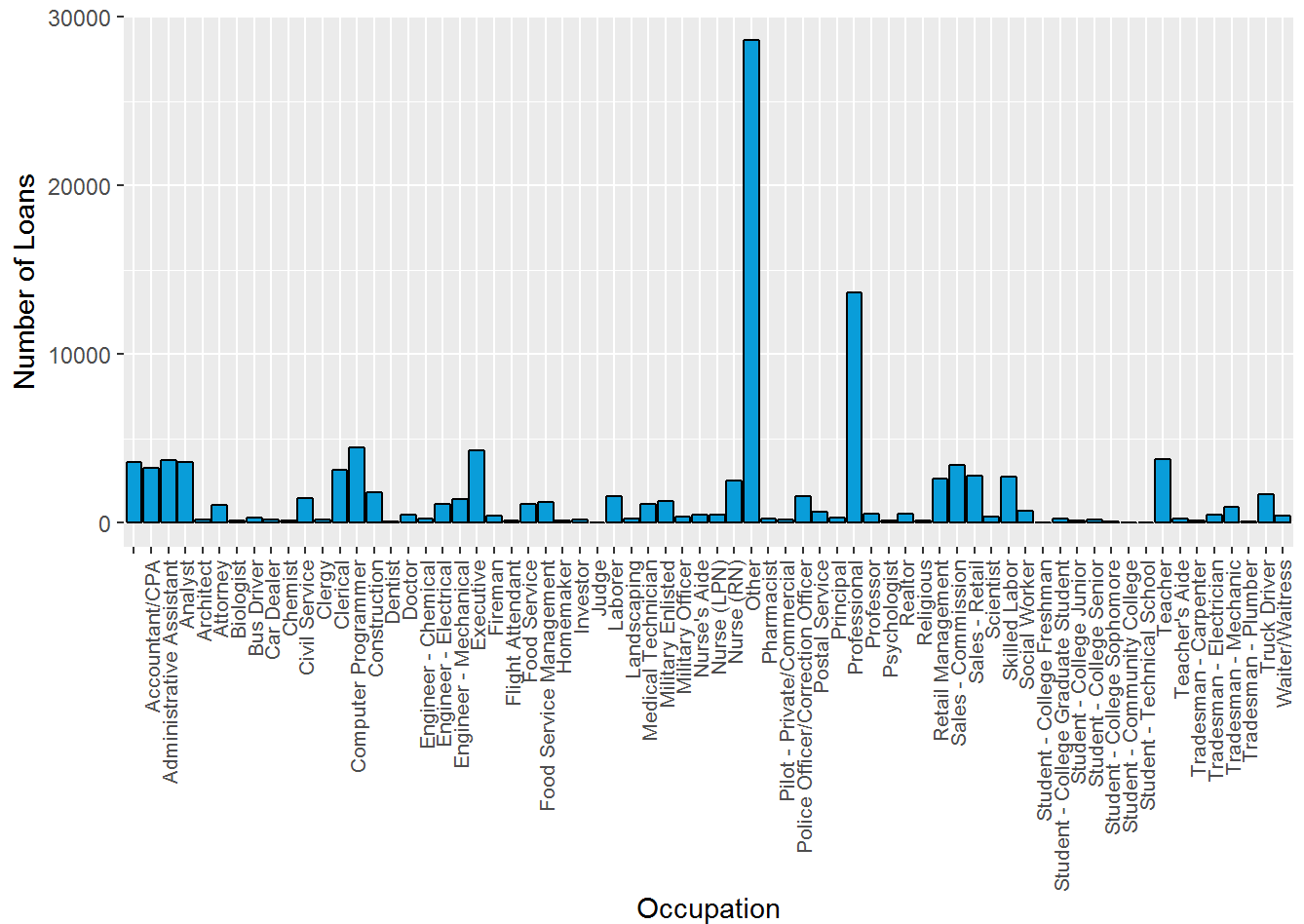
Prosper Debt to Income ratio Distribution



Most borrower's debt to income ratio is between 0-0.6. The peak is around 0.23. However, there're still some borrower's ratio is above 0.75.

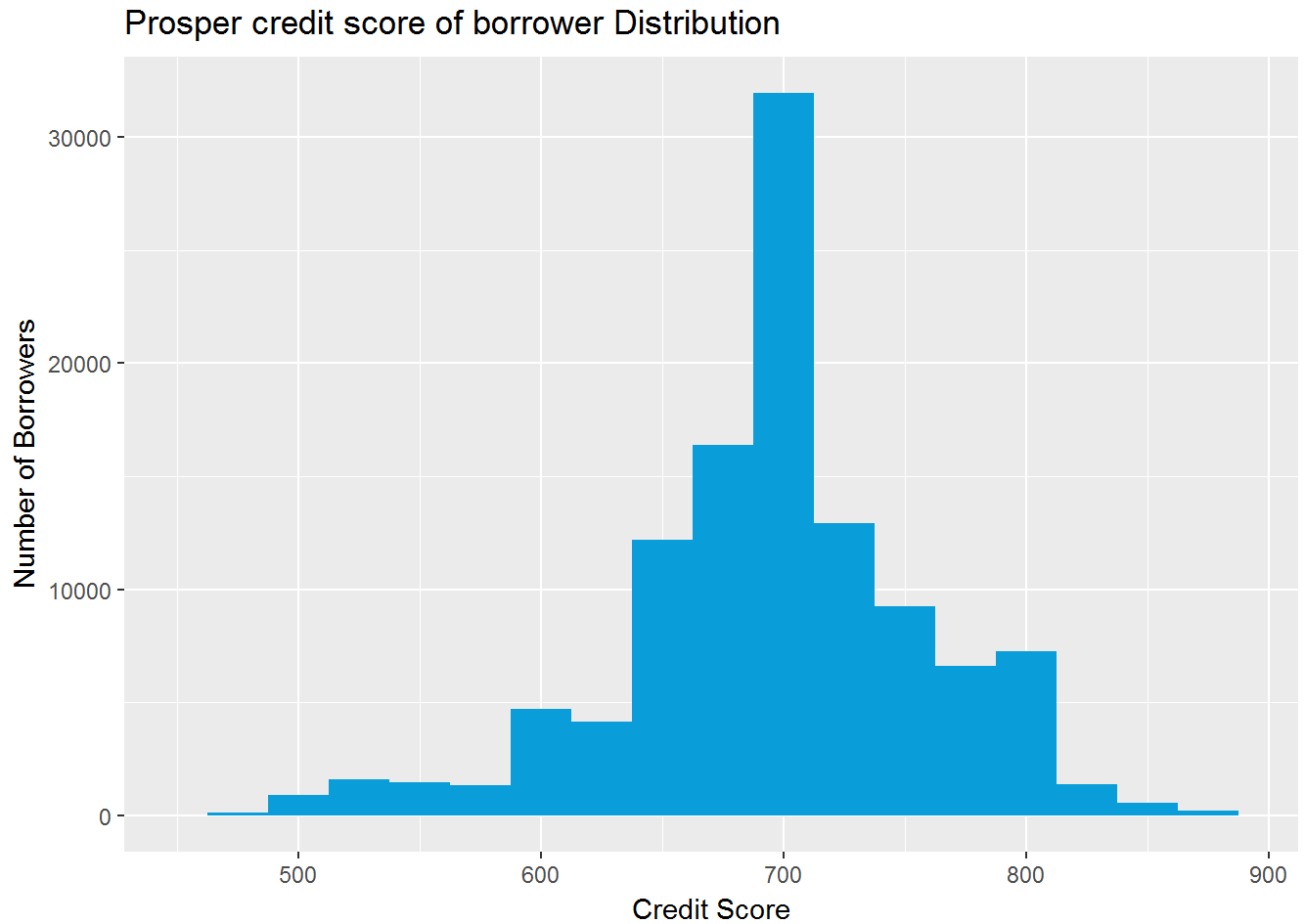
|    |                          |                       |
|----|--------------------------|-----------------------|
| ## |                          | Accountant/CPA        |
| ## | 3588                     | 3233                  |
| ## | Administrative Assistant | Analyst               |
| ## | 3688                     | 3602                  |
| ## | Architect                | Attorney              |
| ## | 213                      | 1046                  |
| ## | Biologist                | Bus Driver            |
| ## | 125                      | 316                   |
| ## | Car Dealer               | Chemist               |
| ## | 180                      | 145                   |
| ## | Civil Service            | Clergy                |
| ## | 1457                     | 196                   |
| ## | Clerical                 | Computer Programmer   |
| ## | 3164                     | 4478                  |
| ## | Construction             | Dentist               |
| ## | 1790                     | 68                    |
| ## | Doctor                   | Engineer - Chemical   |
| ## | 494                      | 225                   |
| ## | Engineer - Electrical    | Engineer - Mechanical |
| ## | 1125                     | 1406                  |

|    |                                    |                                   |
|----|------------------------------------|-----------------------------------|
| ## | Executive                          | Fireman                           |
| ## | 4311                               | 422                               |
| ## | Flight Attendant                   | Food Service                      |
| ## | 123                                | 1123                              |
| ## | Food Service Management            | Homemaker                         |
| ## | 1239                               | 120                               |
| ## | Investor                           | Judge                             |
| ## | 214                                | 22                                |
| ## | Laborer                            | Landscaping                       |
| ## | 1595                               | 236                               |
| ## | Medical Technician                 | Military Enlisted                 |
| ## | 1117                               | 1272                              |
| ## | Military Officer                   | Nurse's Aide                      |
| ## | 346                                | 491                               |
| ## | Nurse (LPN)                        | Nurse (RN)                        |
| ## | 492                                | 2489                              |
| ## | Other                              | Pharmacist                        |
| ## | 28617                              | 257                               |
| ## | Pilot - Private/Commercial         | Police Officer/Correction Officer |
| ## | 199                                | 1578                              |
| ## | Postal Service                     | Principal                         |
| ## | 627                                | 312                               |
| ## | Professional                       | Professor                         |
| ## | 13628                              | 557                               |
| ## | Psychologist                       | Realtor                           |
| ## | 145                                | 543                               |
| ## | Religious                          | Retail Management                 |
| ## | 124                                | 2602                              |
| ## | Sales - Commission                 | Sales - Retail                    |
| ## | 3446                               | 2797                              |
| ## | Scientist                          | Skilled Labor                     |
| ## | 372                                | 2746                              |
| ## | Social Worker                      | Student - College Freshman        |
| ## | 741                                | 41                                |
| ## | Student - College Graduate Student | Student - College Junior          |
| ## | 245                                | 112                               |
| ## | Student - College Senior           | Student - College Sophomore       |
| ## | 188                                | 69                                |
| ## | Student - Community College        | Student - Technical School        |
| ## | 28                                 | 16                                |
| ## | Teacher                            | Teacher's Aide                    |
| ## | 3759                               | 276                               |
| ## | Tradesman - Carpenter              | Tradesman - Electrician           |
| ## | 120                                | 477                               |
| ## | Tradesman - Mechanic               | Tradesman - Plumber               |
| ## | 951                                | 102                               |
| ## | Truck Driver                       | Waiter/Waitress                   |
| ## | 1675                               | 436                               |



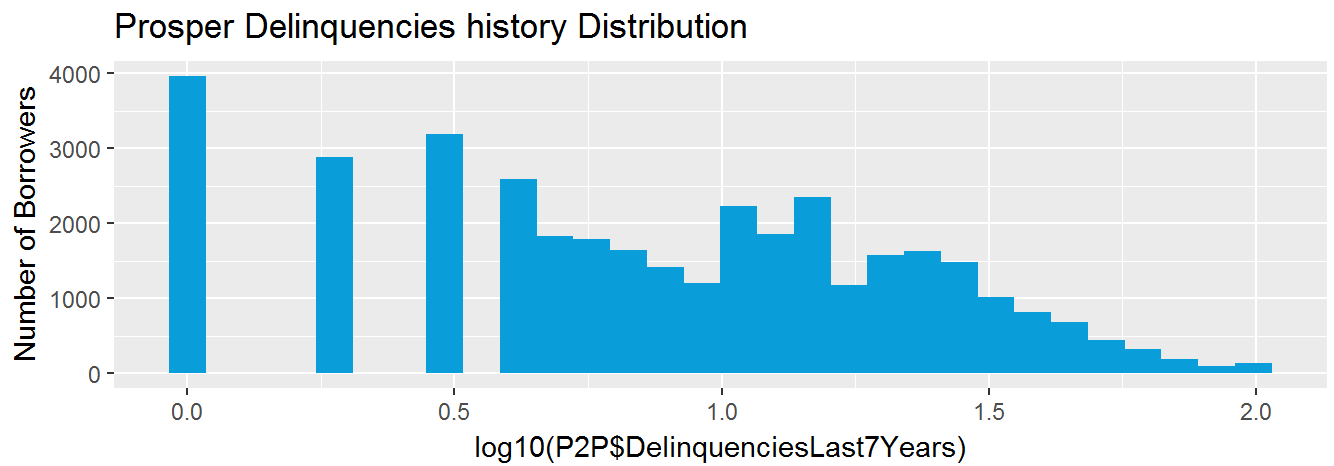
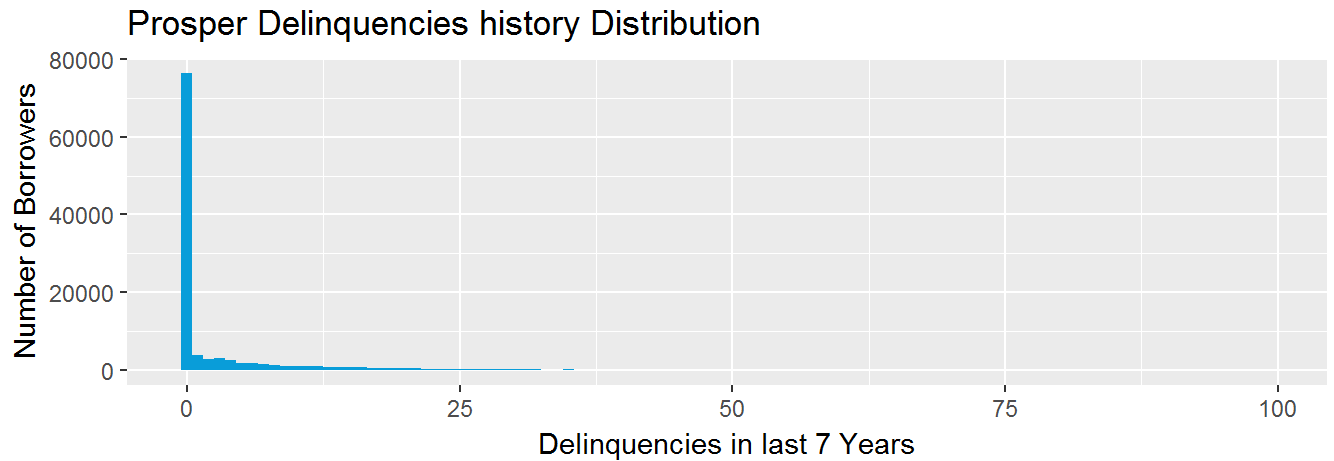
Here we can see borrowers are from all over of different industry. There are also many people choose “Other” or “professional” not sepcifying their occupation.

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
##      9.5   669.5   689.5   695.1  729.5   889.5     591
## Warning: Removed 766 rows containing non-finite values (stat_bin).
## Warning: Removed 1 rows containing missing values (geom_bar).
```



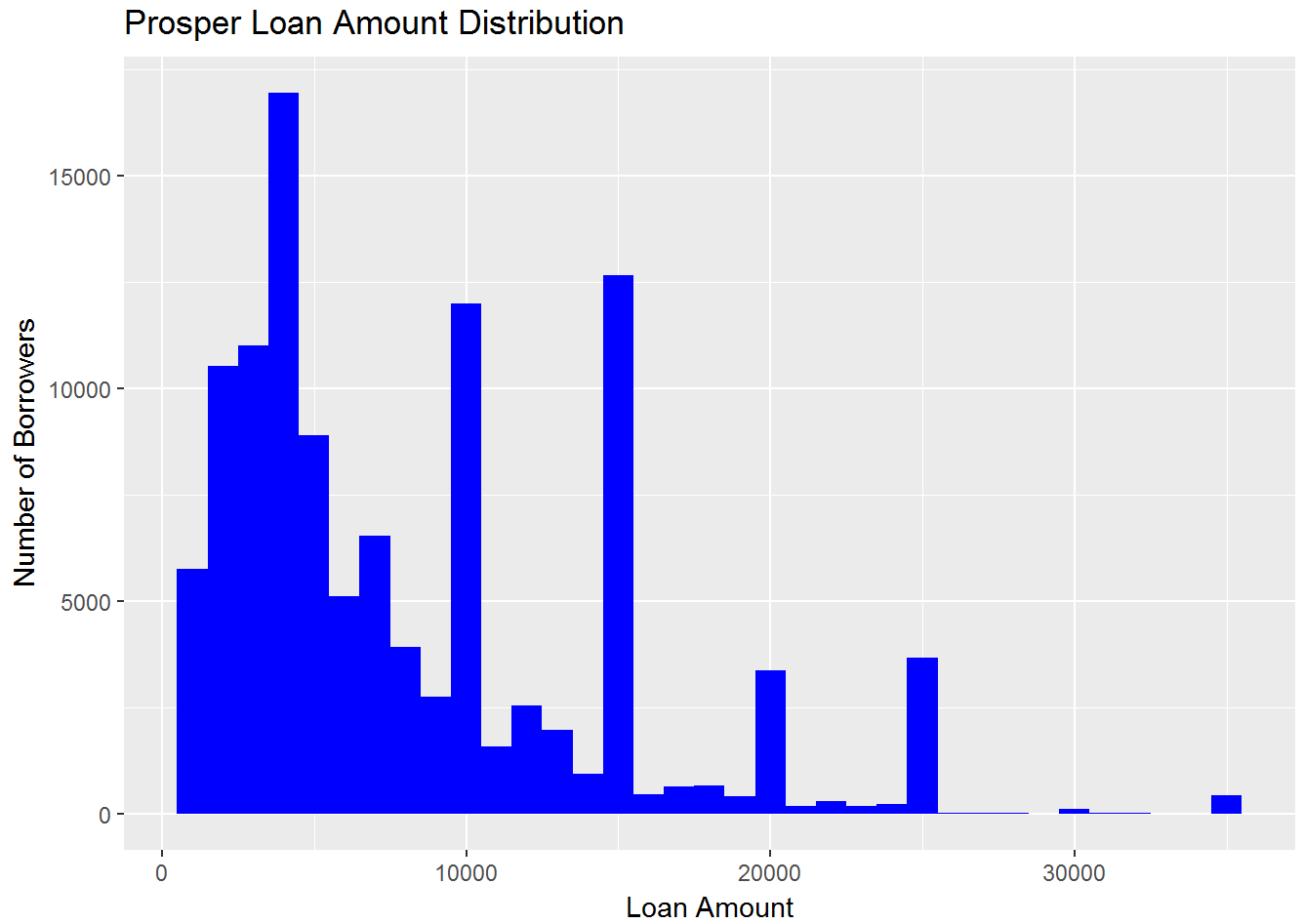
Here we could see a normal distribution chart. The main average of credit score from borrowers are around 700. Most borrowers are from 600-800.

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's
##    0.000  0.000   0.000   4.155  3.000  99.000   990
## Warning: Removed 990 rows containing non-finite values (stat_bin).
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 77429 rows containing non-finite values (stat_bin).
```



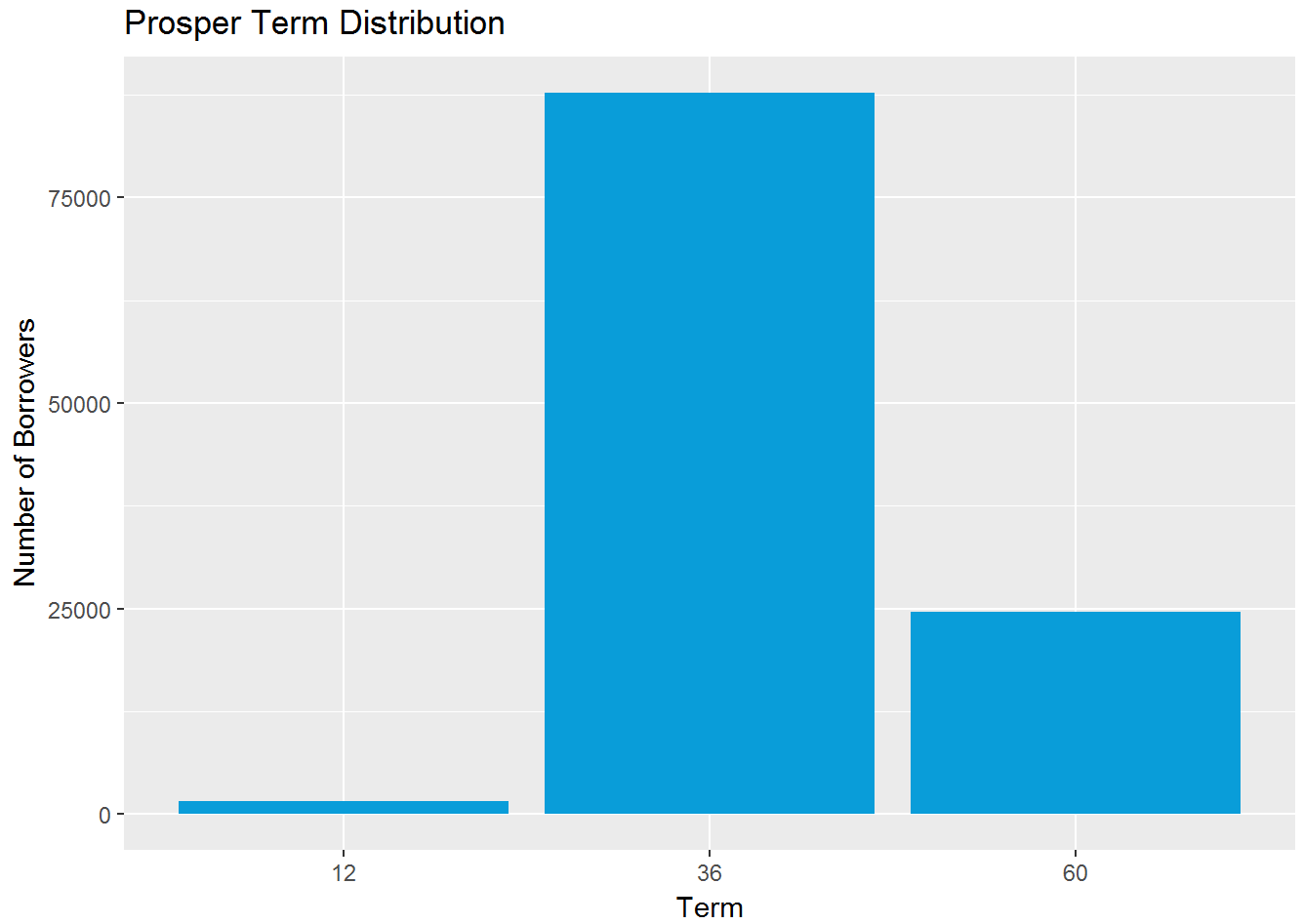
Here we could see a positive skewed chart and a log transformation chart. Most people do not have delinquency record. And people with few delinquency record can still get the loan.

| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max.  |
|----|------|---------|--------|------|---------|-------|
| ## | 1000 | 4000    | 6500   | 8337 | 12000   | 35000 |



Here we can see main loan amount are \$4000,\$10000 and \$15000. And minimum loan amount is \$1000.

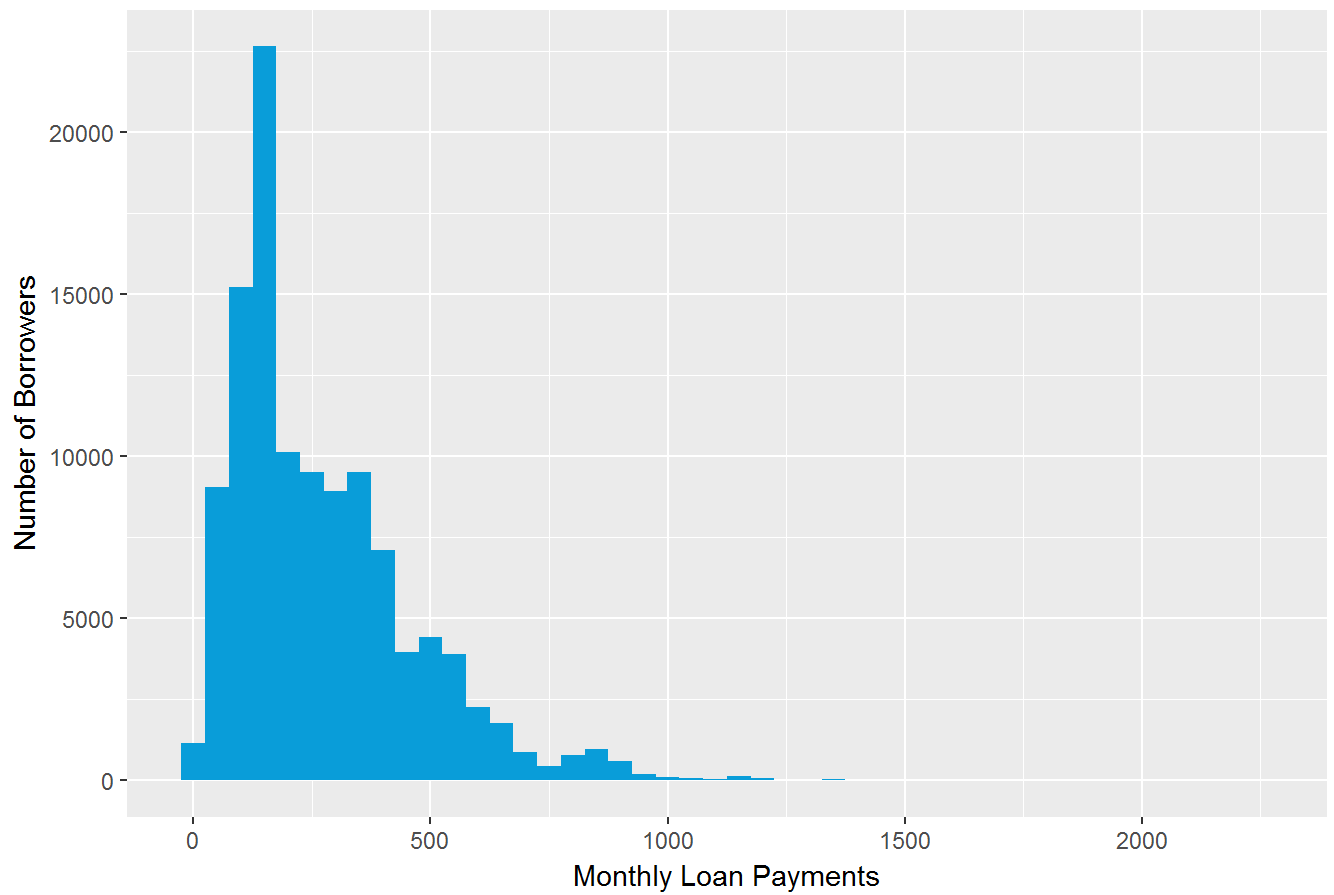
|    |       |         |        |       |         |       |
|----|-------|---------|--------|-------|---------|-------|
| ## | Min.  | 1st Qu. | Median | Mean  | 3rd Qu. | Max.  |
| ## | 12.00 | 36.00   | 36.00  | 40.83 | 36.00   | 60.00 |



Based on chart and summary, we could see most people choose 36 months, then 60 months and last 12 months. 36 months loan is more than tирple of 60 months loan.

|    |      |         |        |       |         |        |
|----|------|---------|--------|-------|---------|--------|
| ## | Min. | 1st Qu. | Median | Mean  | 3rd Qu. | Max.   |
| ## | 0.0  | 131.6   | 217.7  | 272.5 | 371.6   | 2252.0 |

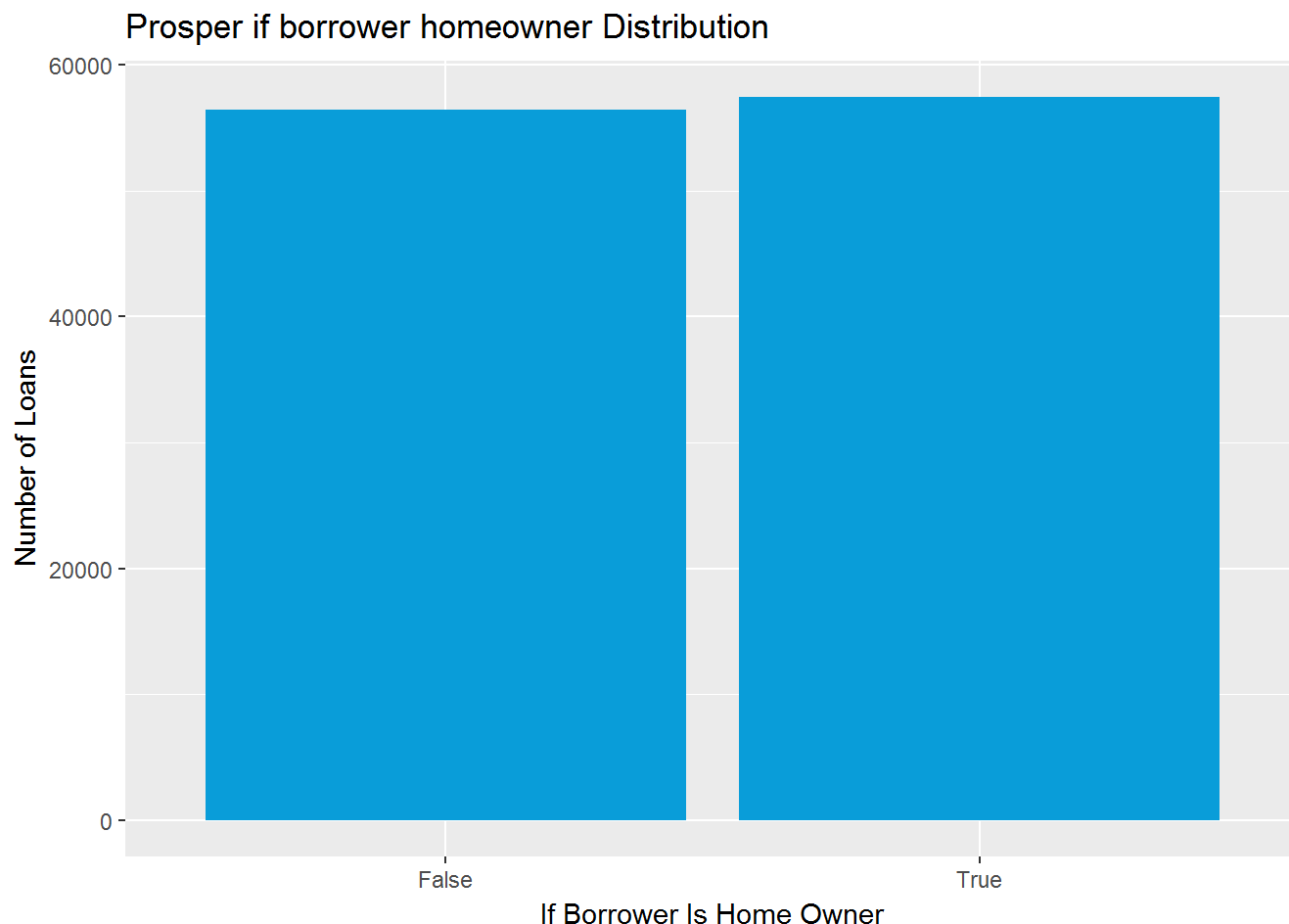
Prosper Monthly Loan Payments Distribution



From this chart, we could see most borrower's monthly loan payments are between \$0-\$500. Mean and median monthly loan payment is \$272.50 and \$217.70.

```
## False True
## 56459 57478
```





Here we could see almost half of borrower are homeowners and other half are not.

## Univariate Analysis

### What is the structure of your dataset?

The dataset has 87 variables and 113937 observations. And the loans cover the period from 2005 to 2014. The dataset contains classes int, numeric, data and factor.

### What is/are the main feature(s) of interest in your dataset?

My main features of interest are: 1. Borrower rate. I am interested to know the relationship between borrower rate and other variables relationship. What factors will mostly affect borrower rate? Like prosper score, income range, loan amount etc. 2. Number of investors. I am interested to know which loans are more attractive for investing. What are the main characteristics of those loans?

## What other features in the dataset do you think will help support your investigation into your feature(s) of interest?

Other features could be credit score, debt to income ratio, prosper rating, occupation and employment status.

## Did you create any new variables from existing variables in the dataset?

No, I did not

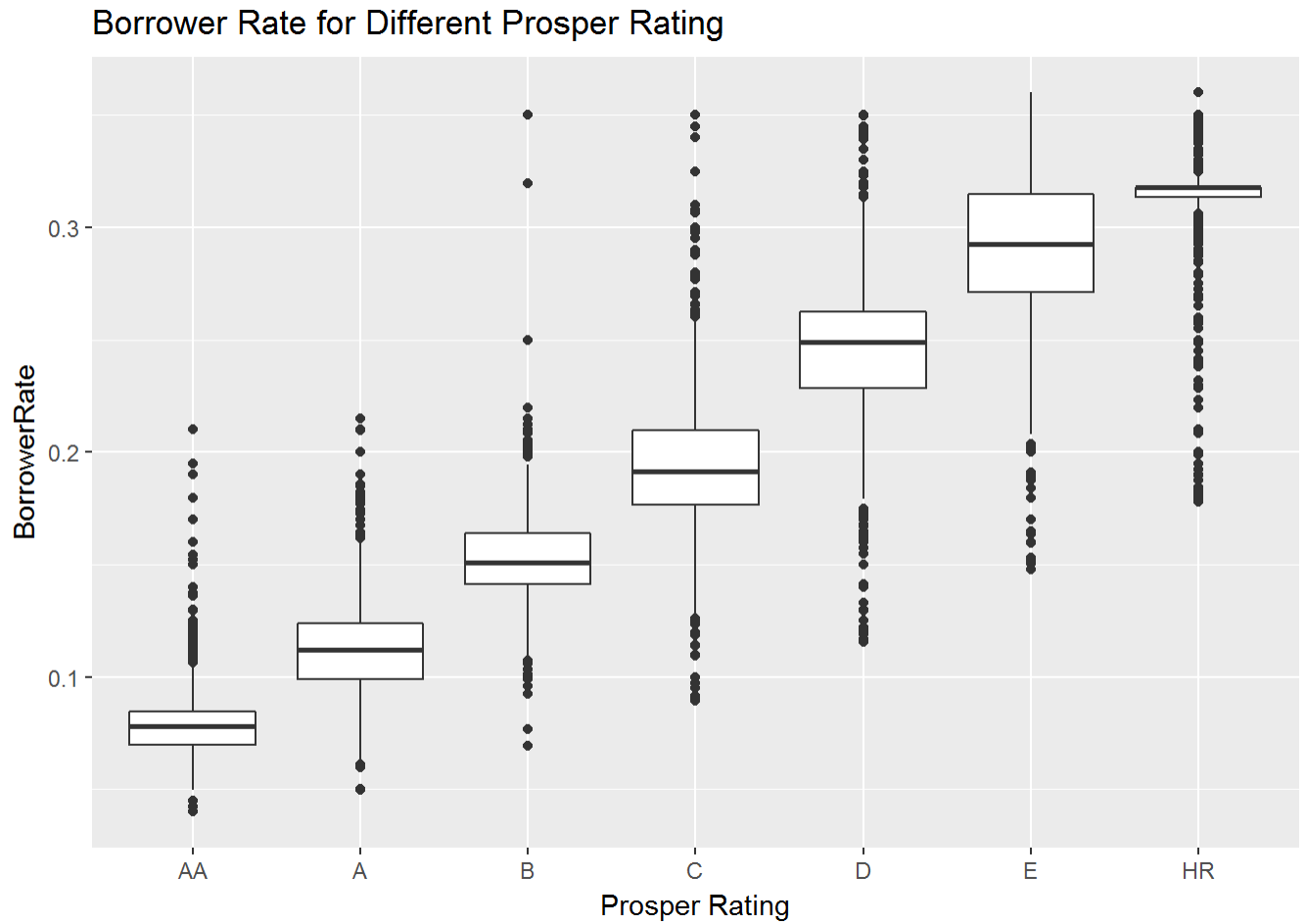
## Of the features you investigated, were there any unusual distributions? Did you perform any operations on the data to tidy, adjust, or change the form of the data? If so, why did you do this?

During my investigation, I find some unusual distributions. When I do income range and job research, I find income lower than \$25000 and part-time job borrowers are not main borrowers. I'm wondering if prosper set up any income limits for borrowers. I did not perform any operations on the data.

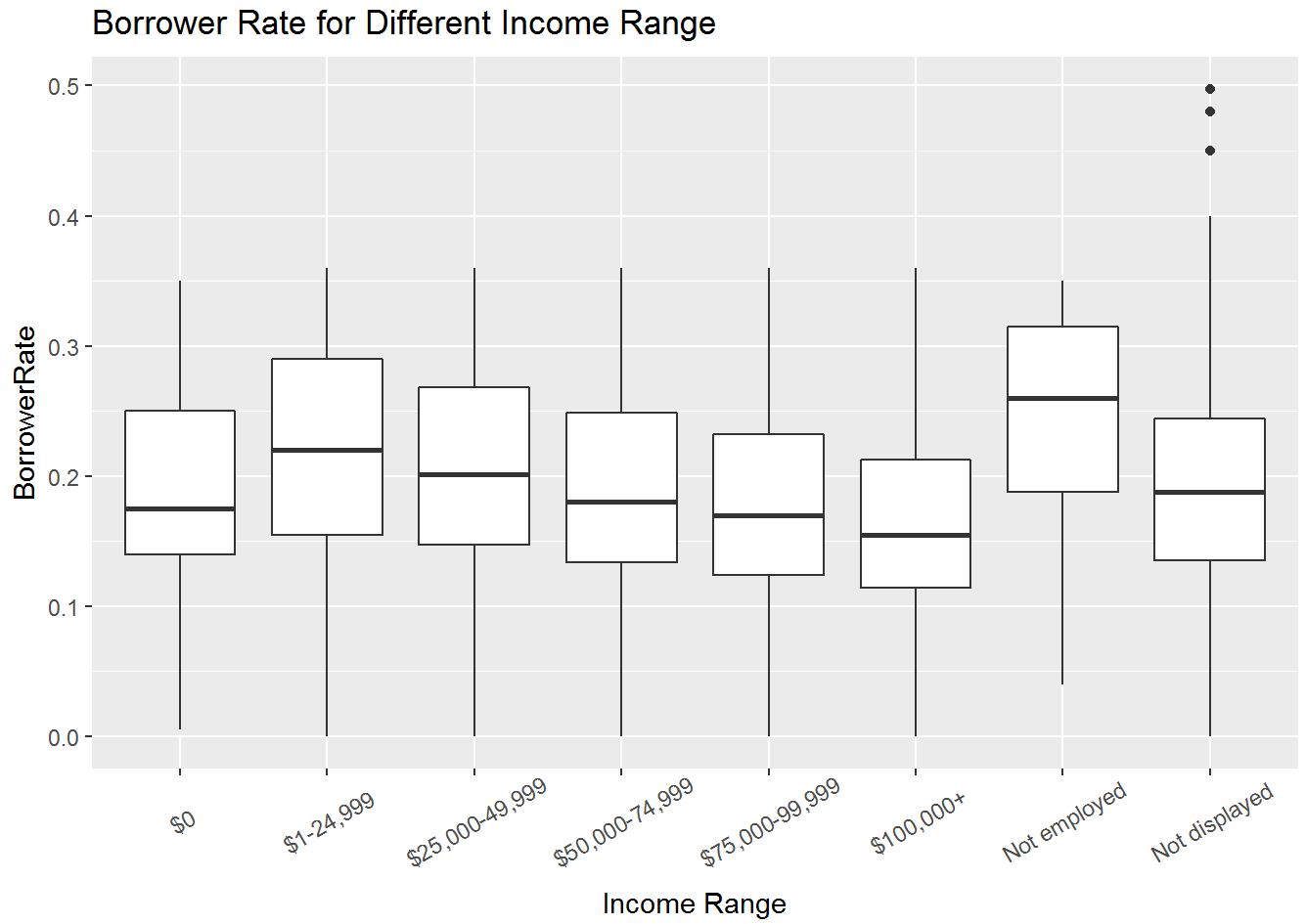
## Bivariate Plots Section

```
## P2P$ProsperRating..Alpha.: AA
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.04000 0.06990 0.07790 0.07912 0.08450 0.21000
## -----
## P2P$ProsperRating..Alpha.: A
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0498 0.0990 0.1119 0.1129 0.1239 0.2150
## -----
## P2P$ProsperRating..Alpha.: B
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0693 0.1414 0.1509 0.1545 0.1639 0.3500
## -----
## P2P$ProsperRating..Alpha.: C
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0895 0.1765 0.1914 0.1944 0.2099 0.3500
## -----
## P2P$ProsperRating..Alpha.: D
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.1157 0.2287 0.2492 0.2464 0.2625 0.3500
## -----
## P2P$ProsperRating..Alpha.: E
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

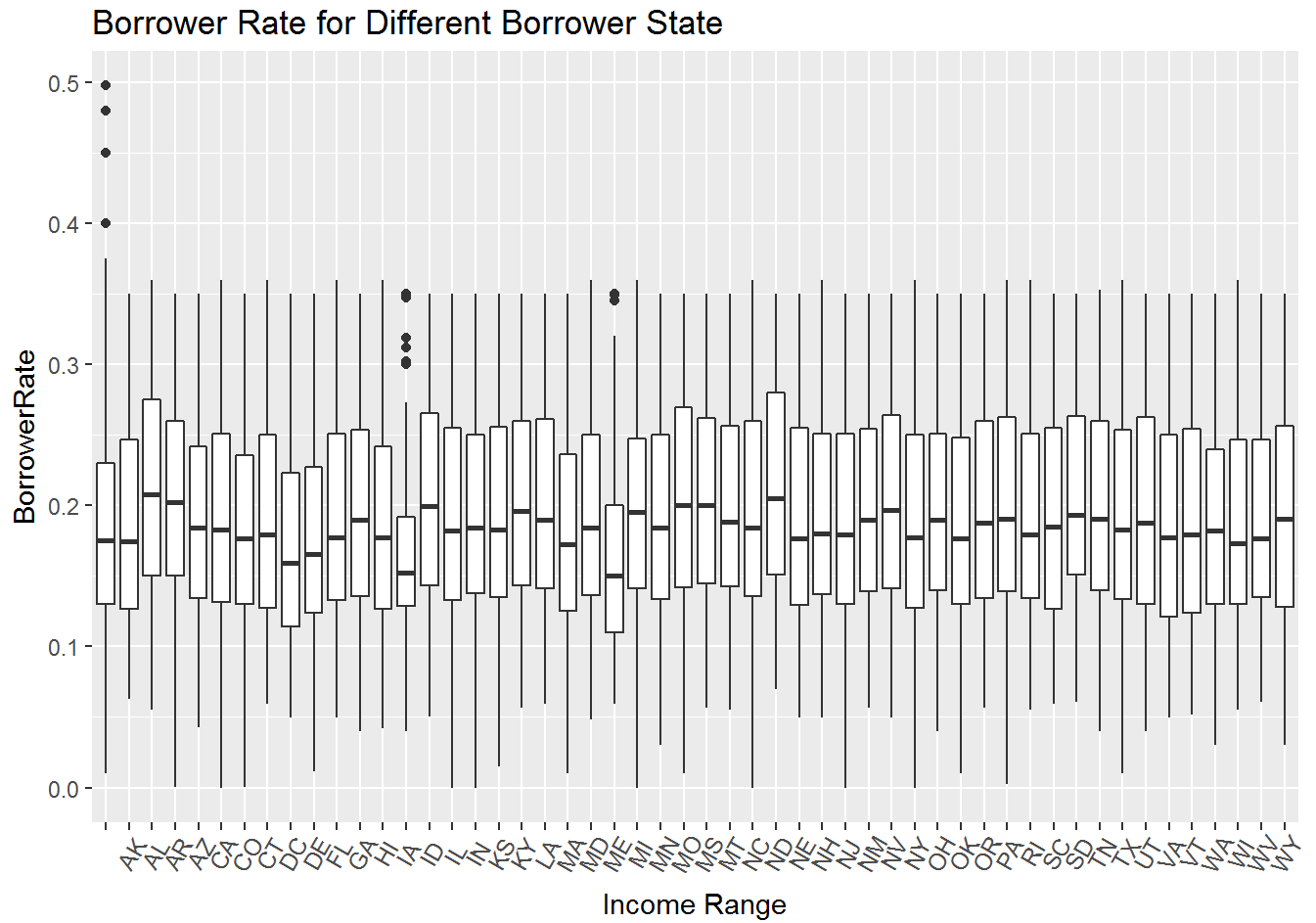
```
## 0.1479 0.2712 0.2925 0.2933 0.3149 0.3600
## -----
## P2P$ProsperRating..Alpha.: HR
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.1779 0.3134 0.3177 0.3173 0.3177 0.3600
## -----
## P2P$ProsperRating..Alpha.:
## NULL
```



From this chart, we could see the reversed relationship between borrower rate and prosper rating. The higher prosper rating, the lower borrower rate.

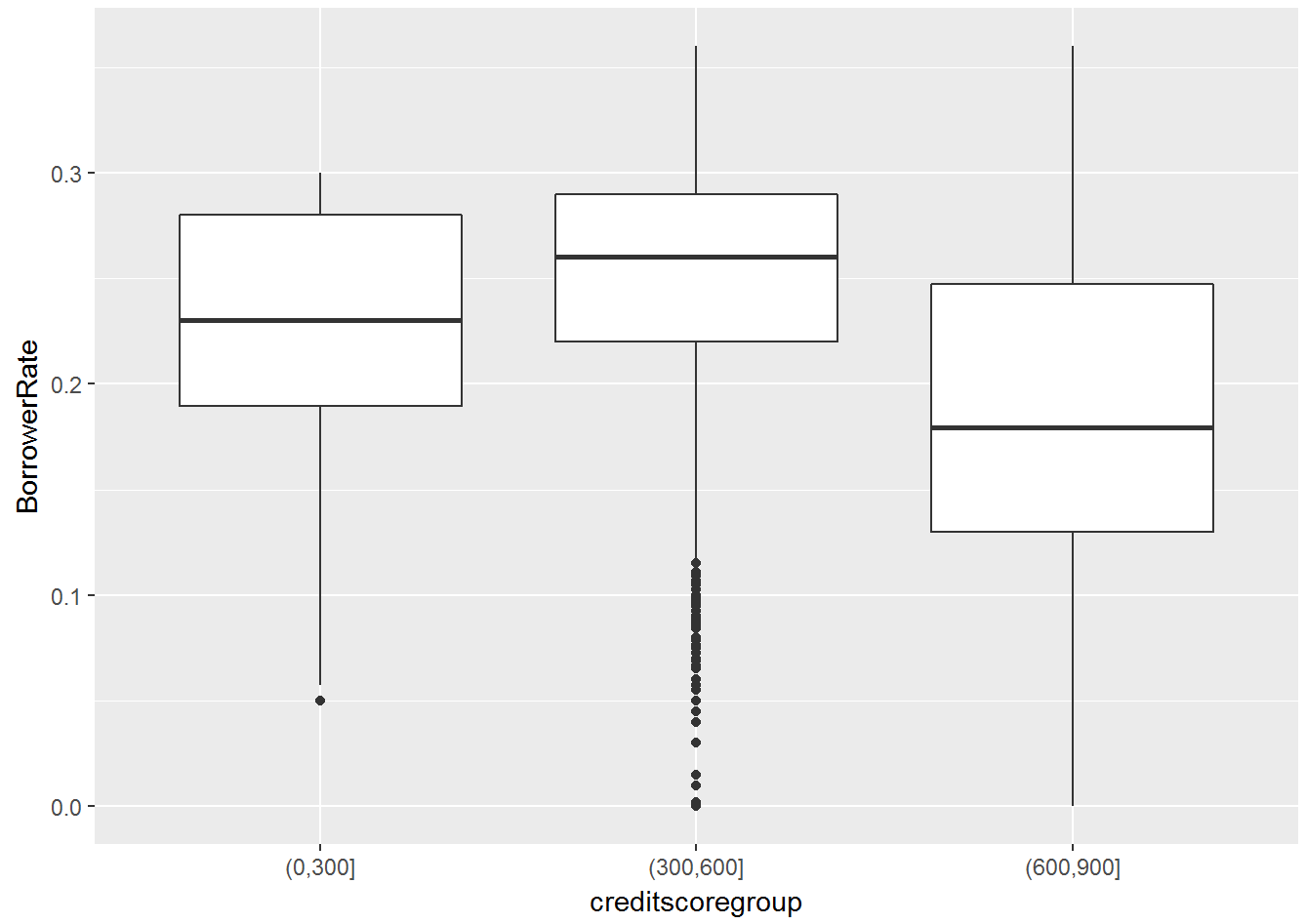


It looks like borrower with higher income will get lower borrower rate.



From this chart, we could see borrower rate has not a strong relationship with borrower state.

```
##
##      (0, 300]  (300, 600]  (600, 900]
##           133      6632    106581
```

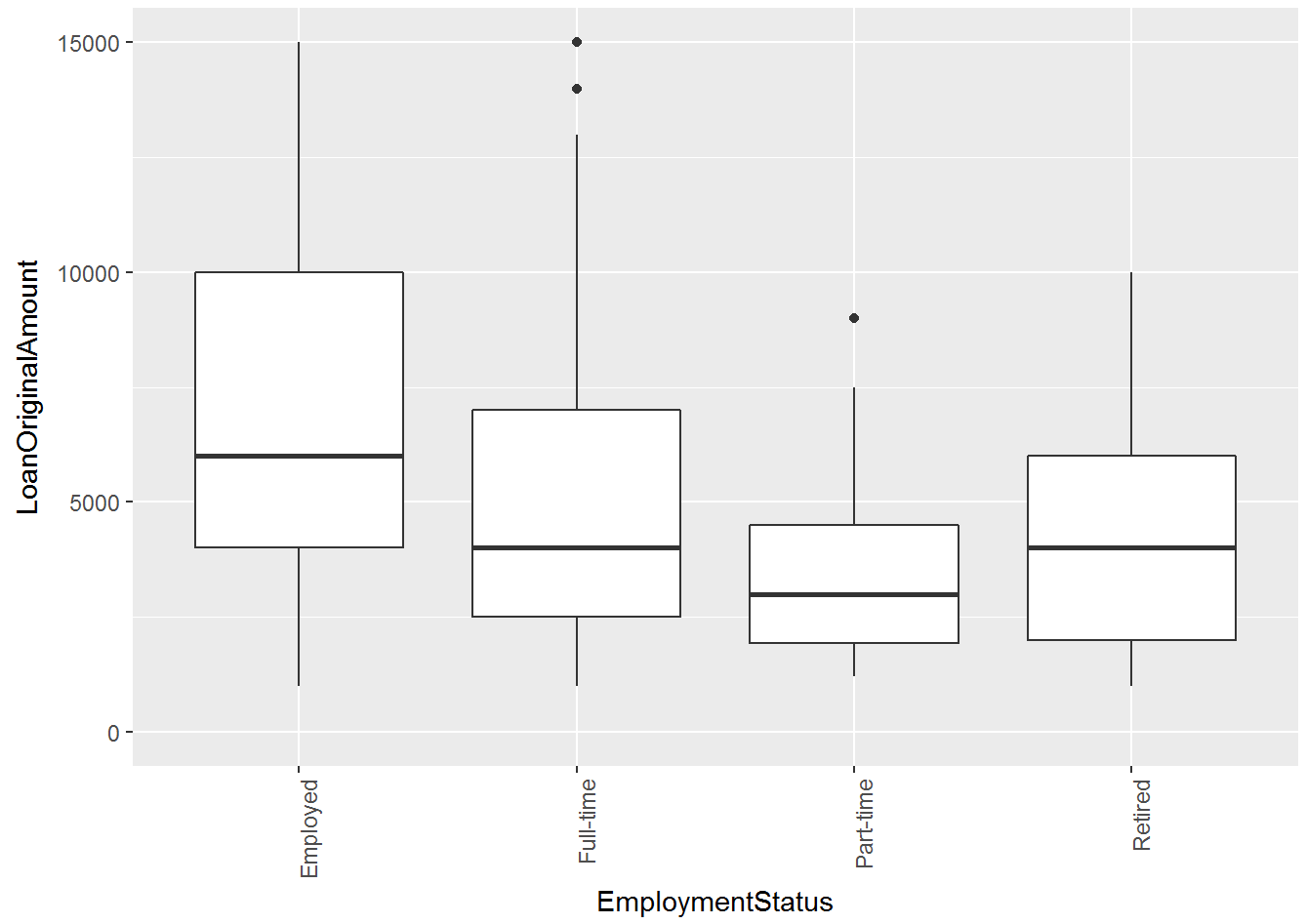


## Warning: Removed 981 rows containing missing values (geom\_point).



Here we use two charts to explore the relationship between borrower credit score and borrower rate. From first chart, we could see the relationship between different credit score group and borrower rate. Borrower with 600-900 credit score has lowest borrower rate, which average is around 0.18. From second chart, we could see a scatter plot chart with borrower's credit score 500 or above. We could clearly see that the higher credit score, the lower borrower rate.

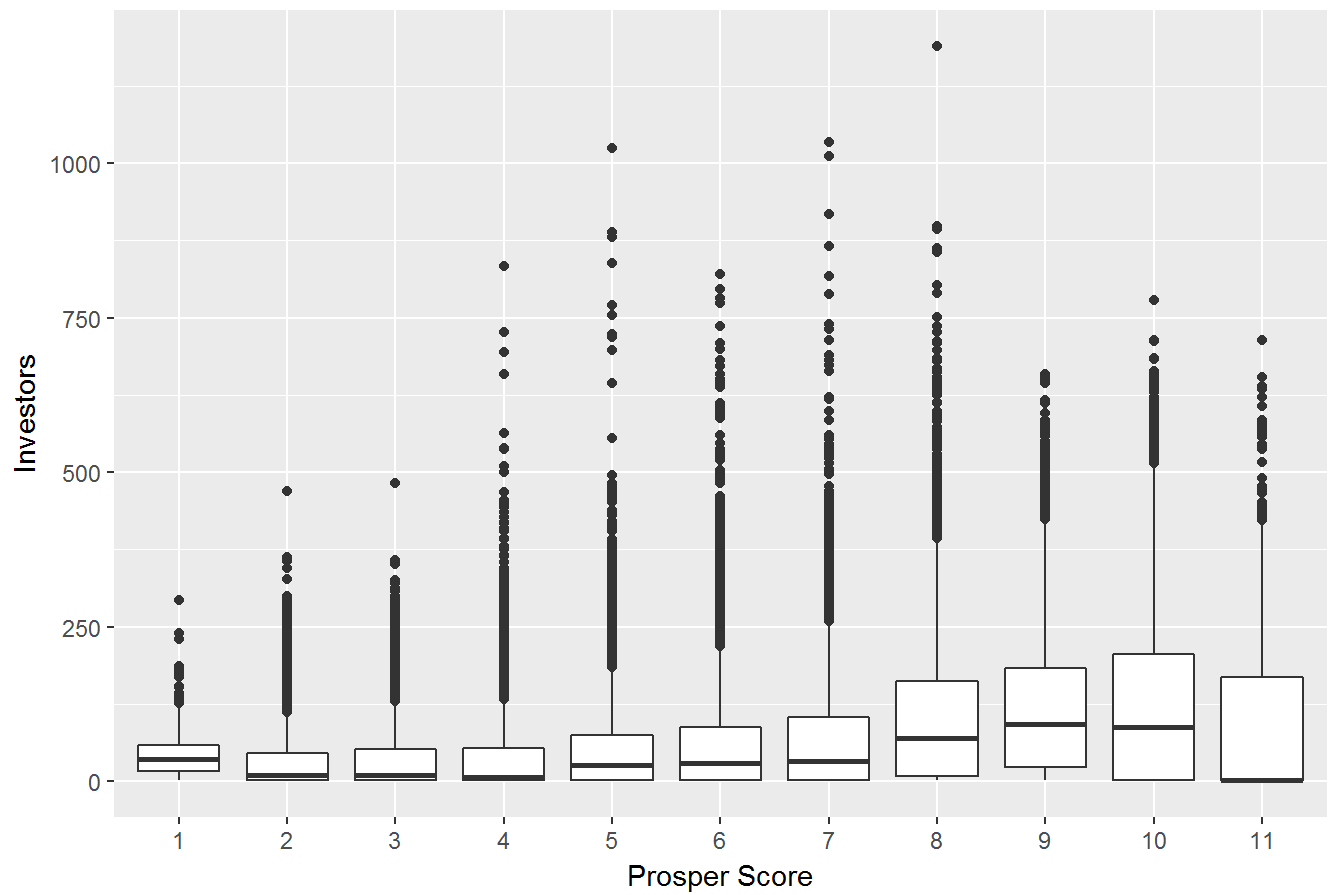
```
## Warning: Removed 22 rows containing non-finite values (stat_boxplot).
```



Employed and retired borrower seems to have higher loan amount than part-time.

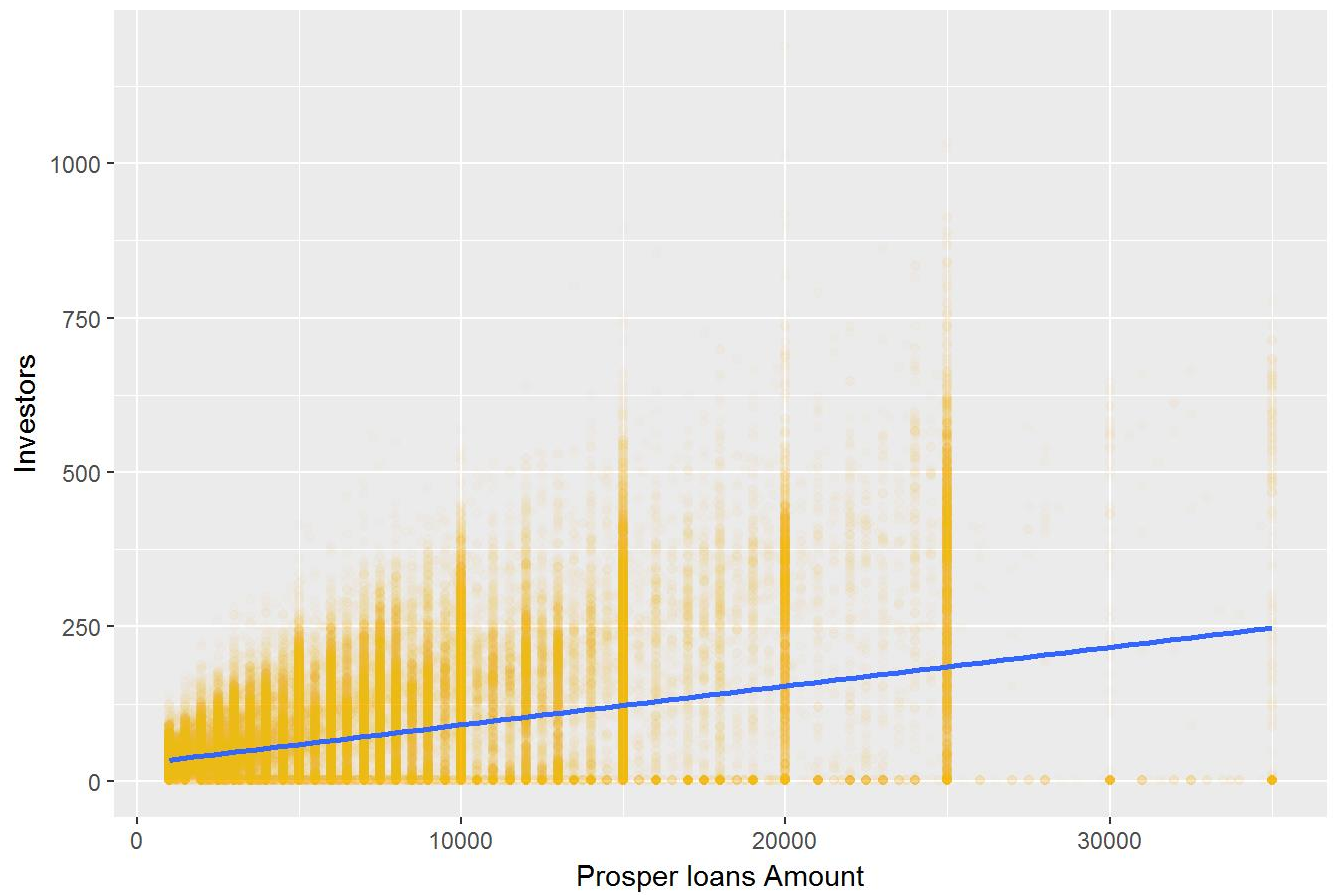


Prosper Score vs Number of Investors



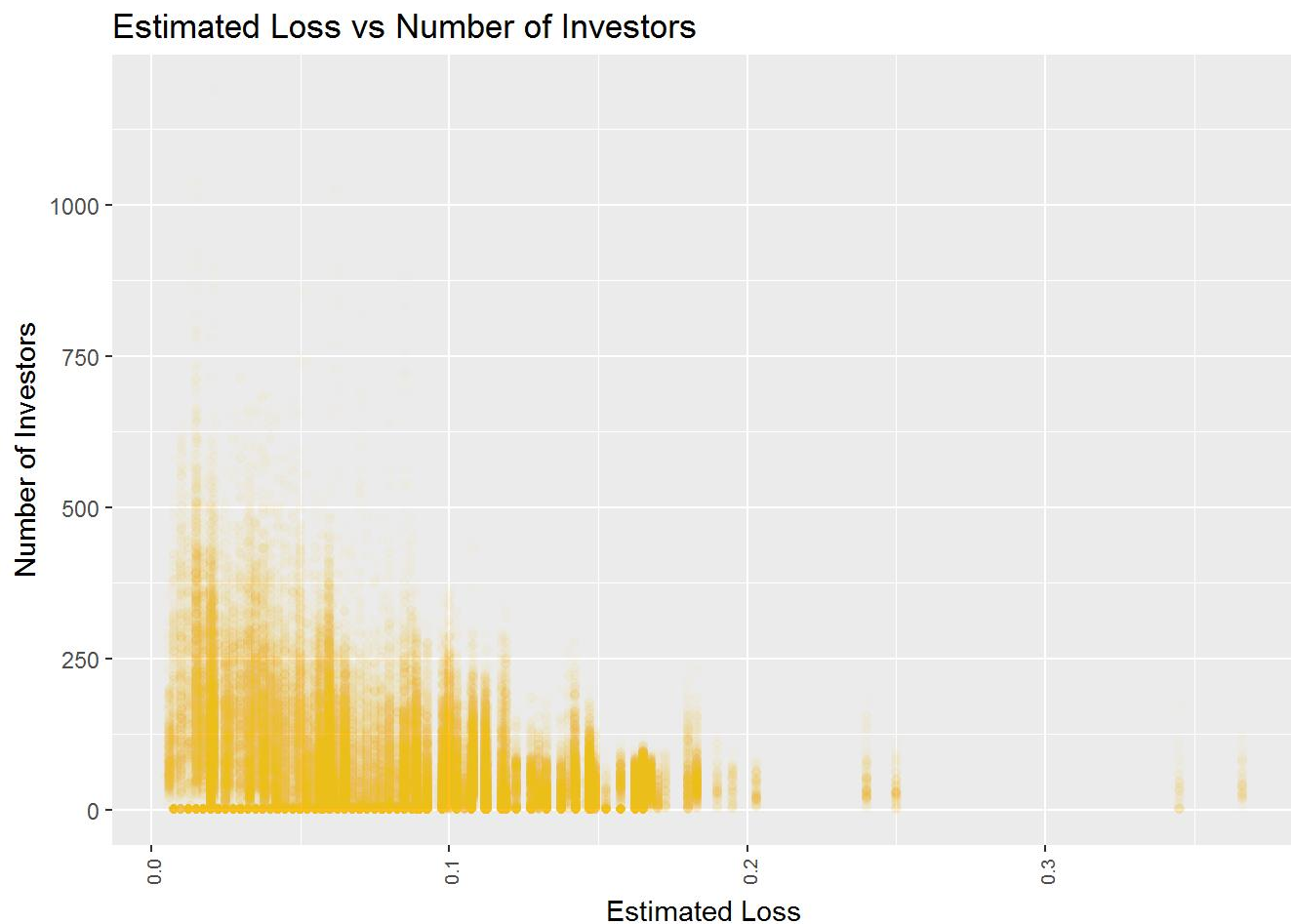
From the chart, we could see the higher the Prosper Score, the more investors. The peak is prosper score equal to 10.

Prosper loans vs Number of Investors



Here we could see that the higher loans amount, the more investors.

```
## Warning: Removed 29084 rows containing missing values (geom_point).
```



We can find the lower the estimated loss, the more investors.

## Bivariate Analysis

**Talk about some of the relationships you observed in this part of the investigation. How did the feature(s) of interest vary with other features in the dataset?**

1. Borrow Rate: There's reversed relationship between borrower rate and prosper rating. The higher prosper rating, the lower borrower rate. And borrower with higher income will get lower borrower rate. The higher credit score, the lower borrower rate. Employed and retired borrower seems to have higher loan amount than part-time. 2. Investors: We can find the higher the Prosper Score, the more investors. The peak is prosper score equal to 10. We also could see that the higher loans amount, the more investors. We can find the lower the estimated loss, the more investors.

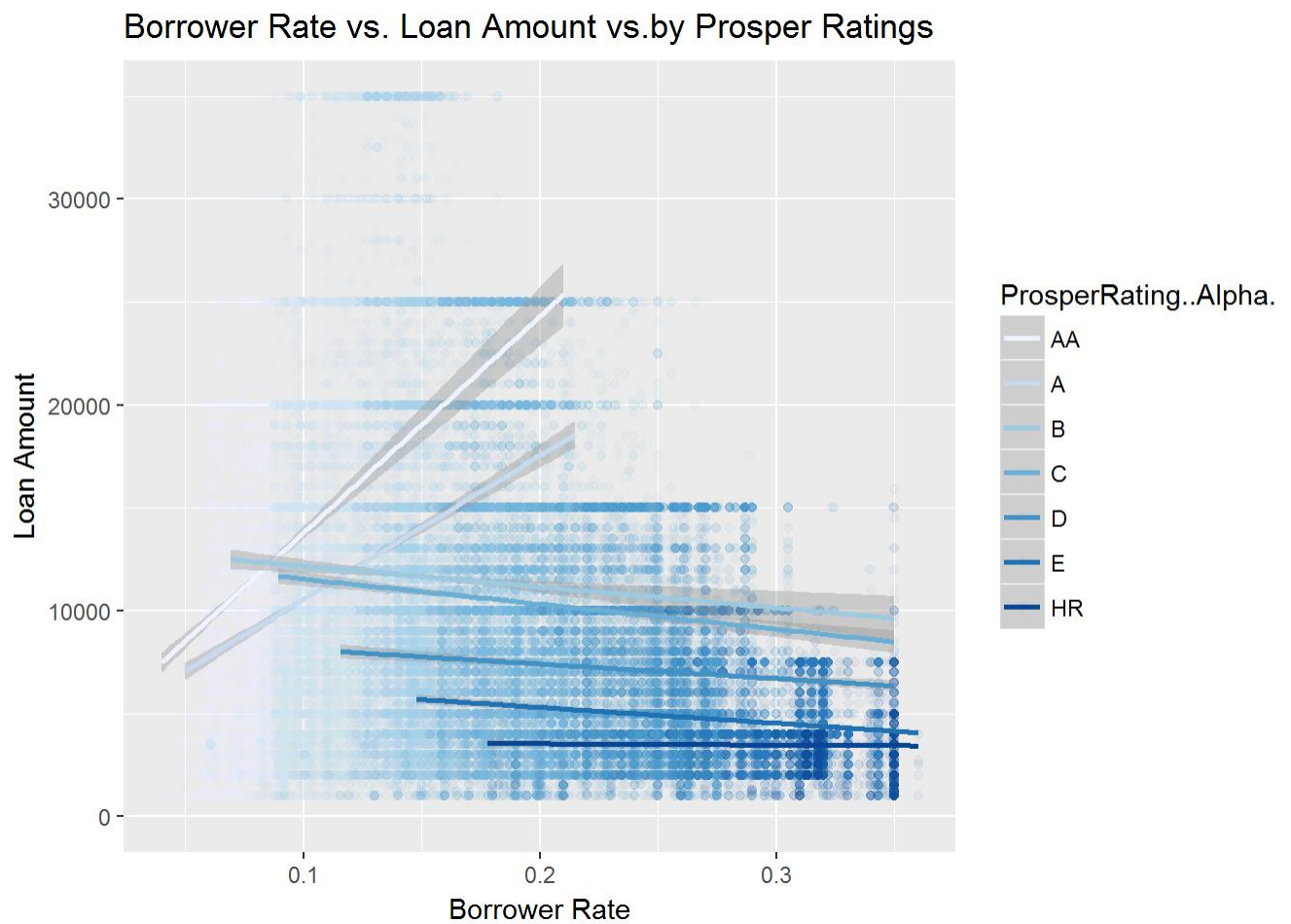
**Did you observe any interesting relationships between the other features (not the main feature(s) of interest)?**

No

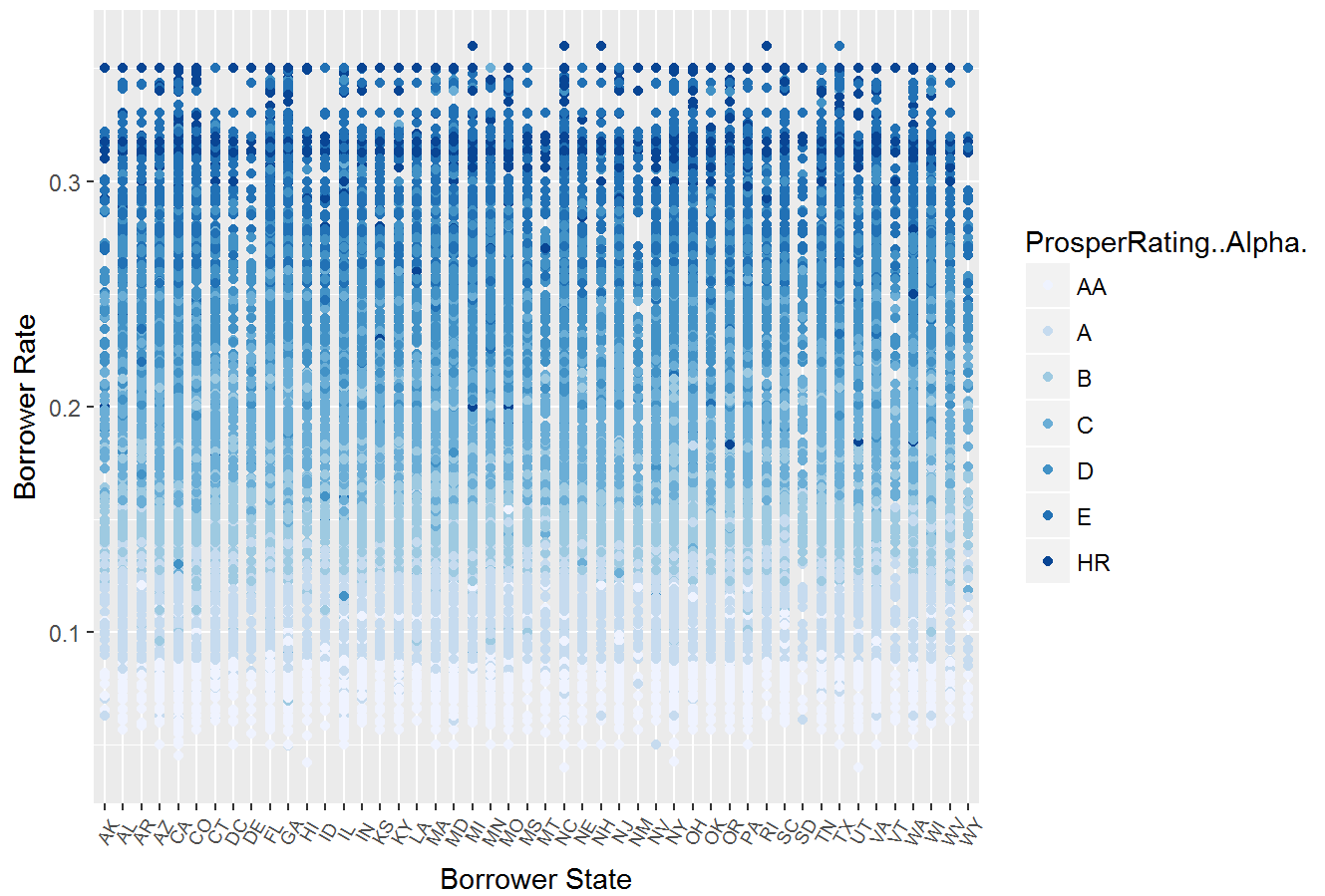
**What was the strongest relationship you found?**

The reverse relationship between borrower rate and prosper rating.

## Multivariate Plots Section

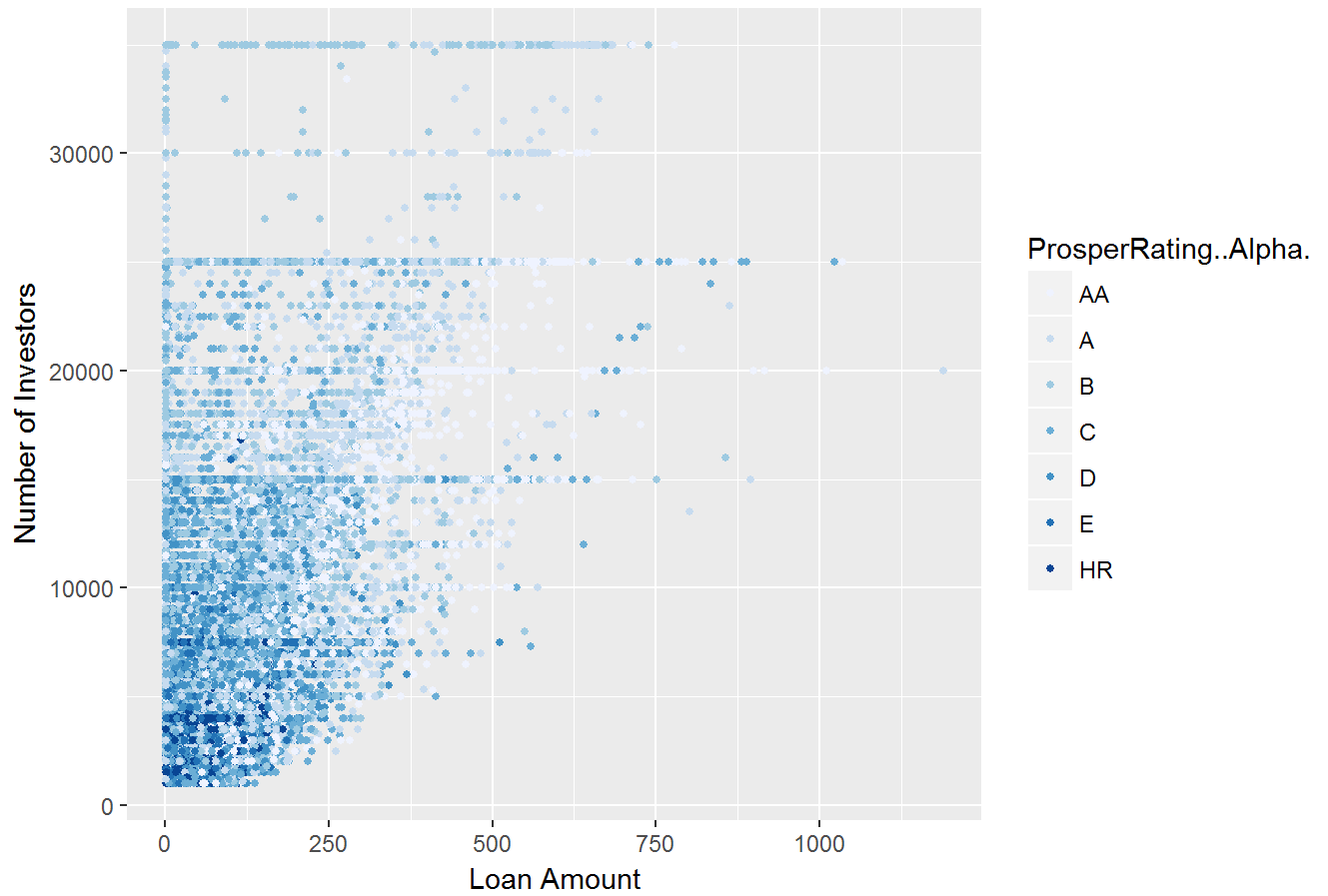


Borrower State vs. Borrower Rate vs.by Prosper Ratings



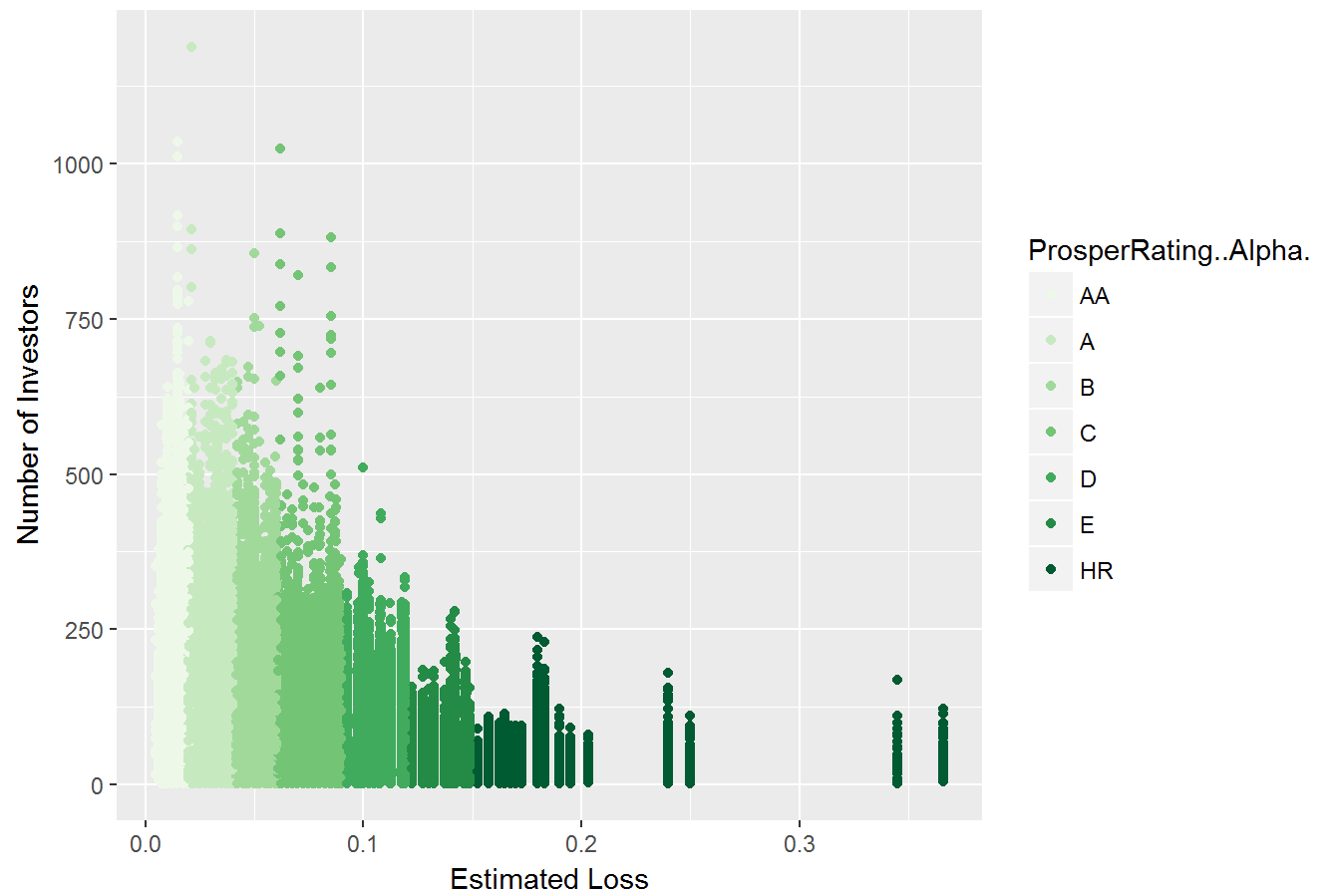
Each state is almost same, but still the lower the prosper rating, the higher the borrower rate.

Number of Investors vs. Loan Amount vs. by Prosper Ratings



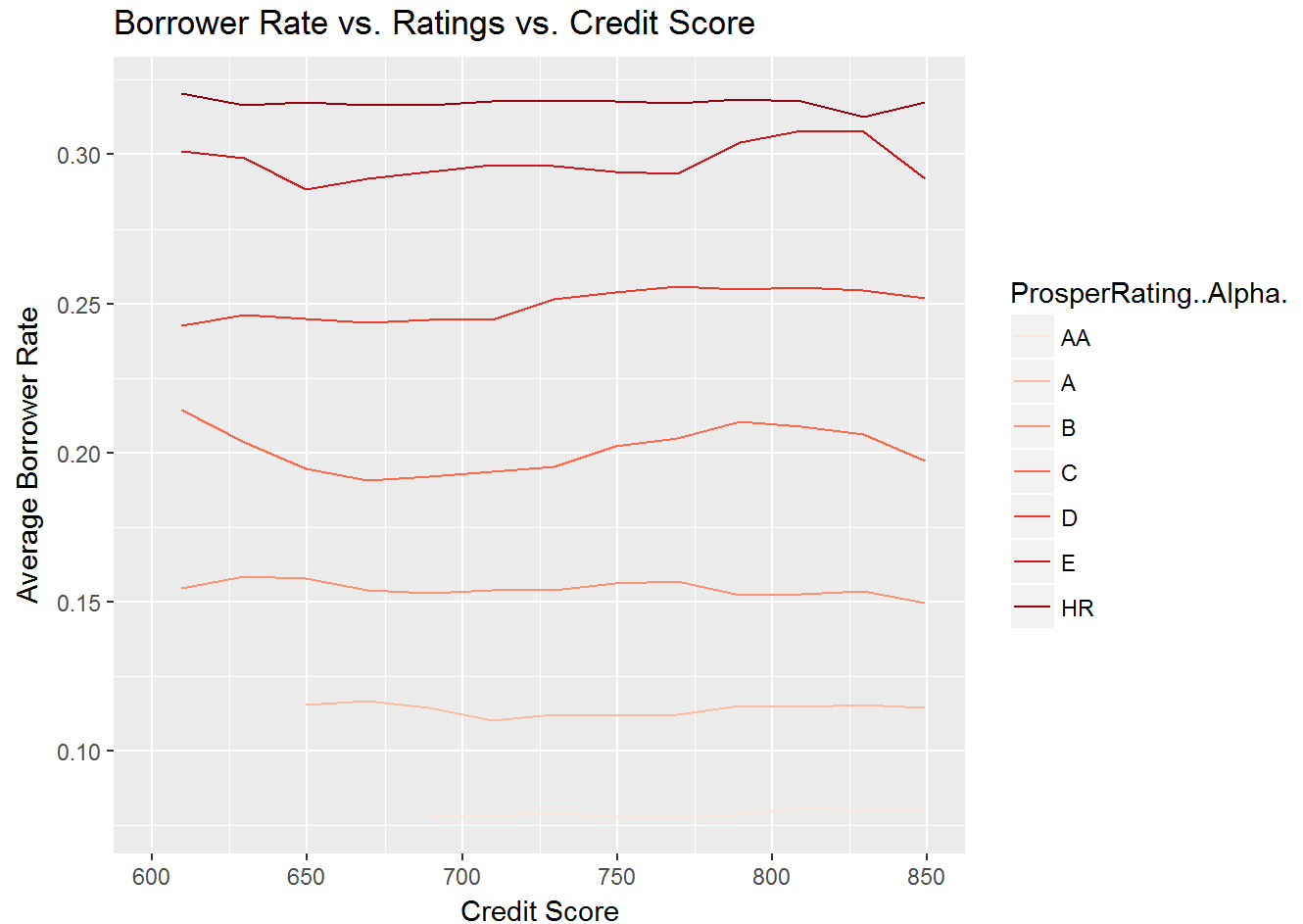
From this chart, we could see the relationship between Investors, loan amount and prosper ratings. The higher the prosper ratings, there will be more investors. However, most investors are interested in loan amount less than \$30,000. \$10,000 to \$20,000 are most popular investor's choice no matter what rating.

Number of Investors vs. Estimated Loss vs. by Prosper Ratings



The higher the prosper rating, the lower the estimated loss, the more the number of investors.

```
## Warning: Removed 139 rows containing non-finite values (stat_summary).
```



This chart summarizes the relationship between borrower rate, rating, and borrower credit score. We can see that loans with high ratings usually require higher credit scores.

## Multivariate Analysis

**Talk about some of the relationships you observed in this part of the investigation. Were there features that strengthened each other in terms of looking at your feature(s) of interest?**

Prosper rating and investors are strongly correlated. Here we can see that Prosper rating, investors, loan amount, and estimated loss are affecting each other.

**Were there any interesting or surprising interactions between features?**



There were interesting interactions between prosper rating, loan amount and number of investors. It looks like up to \$25,000, higher loan amount means more investors, but when it comes above \$25,000, there were only few investors.

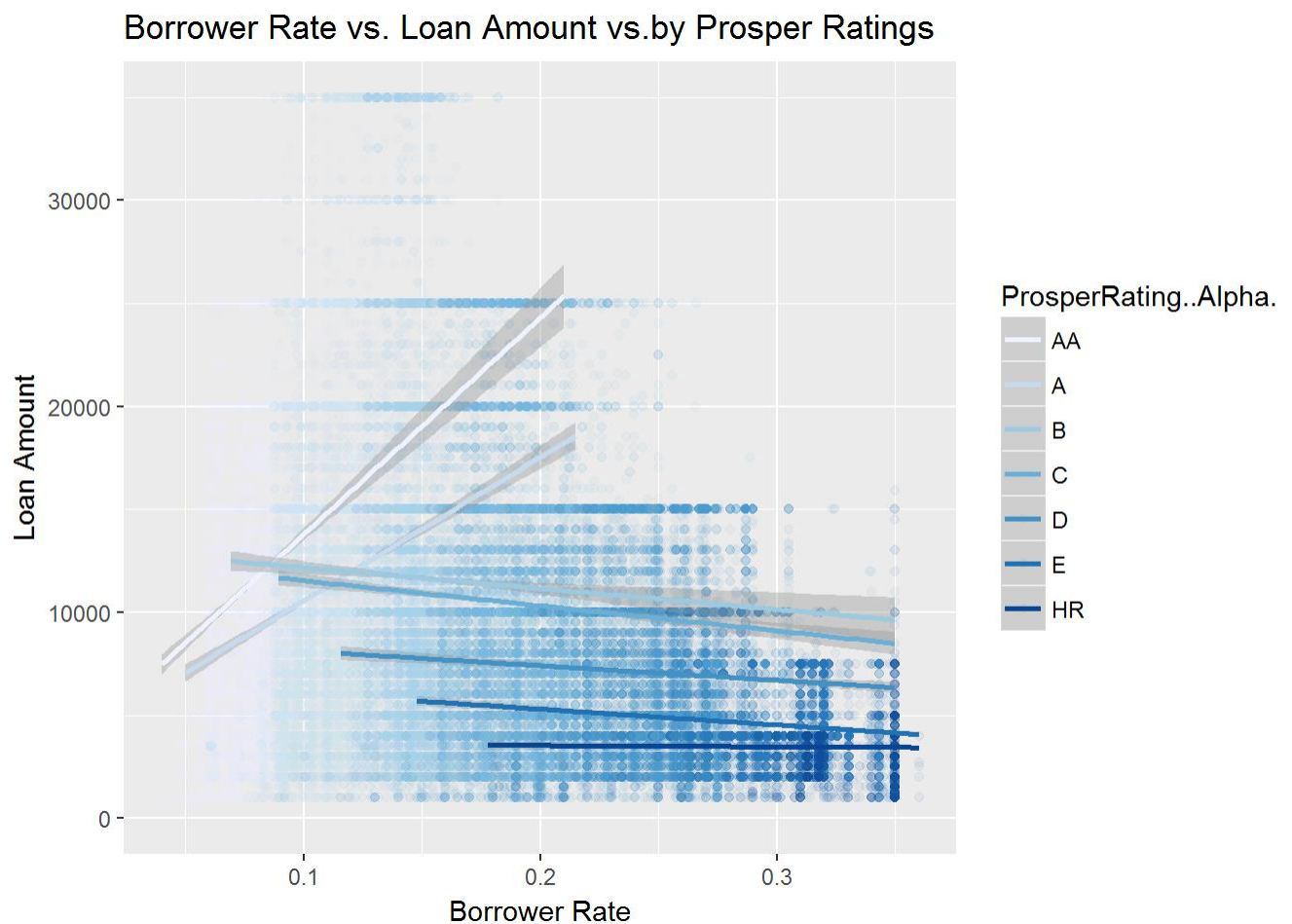
**OPTIONAL: Did you create any models with your dataset?  
Discuss the strengths and limitations of your model.**

No

---

## Final Plots and Summary

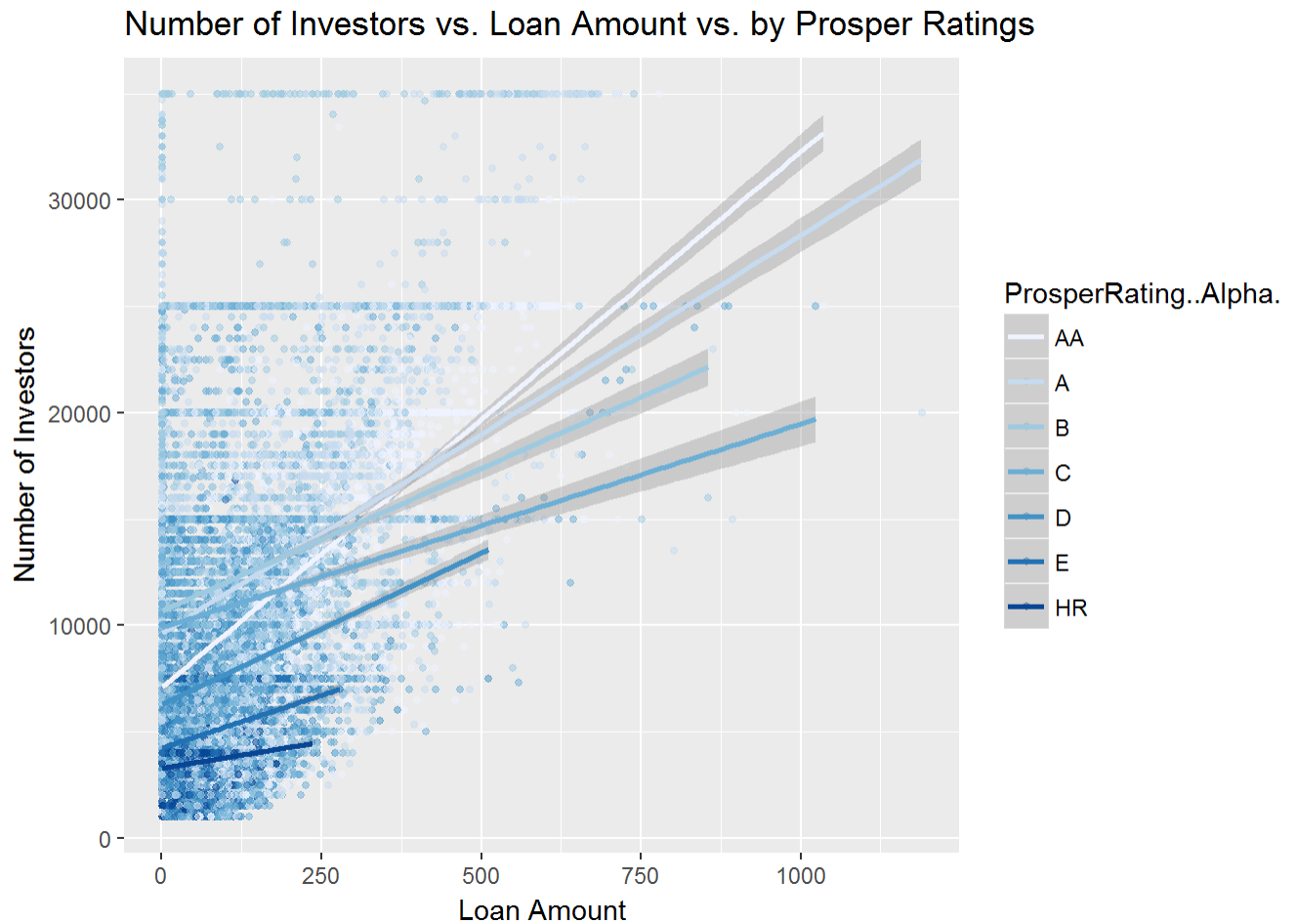
### Plot One



### Description One

This chart basically summarize the relationship between borrower rate, loan amount and prosper ratings. We can clearly see that high rating loans usually got low borrower rate. And some large loan amount (more than \$25,000) are only available for high rating loans. And we also can see loans with higher borrower rate and lower rating usually has small loan amount.

## Plot Two

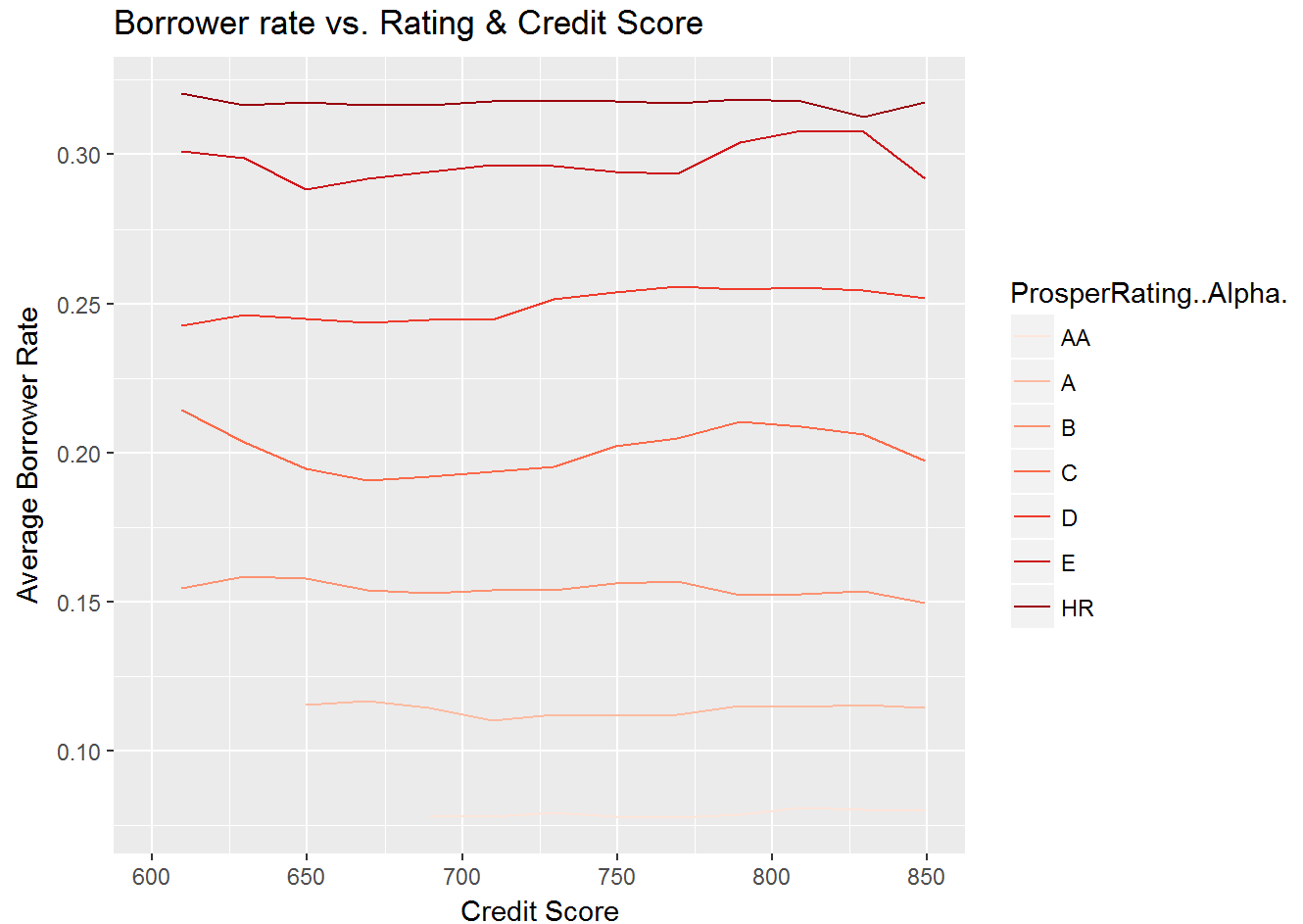


## Description Two

This chart summarize the relationship between number of investors, loan amount and prosper ratings. We can see most investors are interested in investing loan amount less than \$25,000. For loans above \$25,000, investors are Only interested in high rating loans, like “A” or “AA”. When investors was interested on low rating loans, they only invest loans with smaller loan amount (less than \$15,000). Probaly they are worried about risk with large loan amount.

## Plot Three

```
## Warning: Removed 139 rows containing non-finite values (stat_summary).
```



## Description Three

This chart summarizes the relationship between borrower rate, rating, and borrower credit score. We can see that loans with high ratings usually require higher credit scores. "AA" rating loans require borrowers to have almost a 700 credit score in order to qualify to apply. Other rating loans' minimum credit scores start from 600. We also can see that higher rating loans are always associated with lower borrower rates.

## Reflection

I used to be an Investment Consultant for more than 6 years, so I had some general ideas about loans and investments. So I chose this database as my research topic. After I reviewed the dataset, I was interested to find the relationship between loan rating, borrower rate, and number of investors. I wanted to learn more about P2P companies and see what factors will affect investors to make their decisions.

I was new to computer system. So at the beginning, I was struggling with ggplot2 syntax. After lots of practise, reserching, reading and learning on Google, I gradually become comfortable with ggplot2 sytax. Before I did this project, I did not learn deeply about P2P company like Prosper.com. But, this topic was quite interesting to me and related to my previous job. I also checked Prosper.com to see the meaning of each column and how they operate their business. I also took time to read database and wanted to find most related factors.

Then I imported the databased and used the ggplot2 to see the correlation with each factors and variables. After I did two variable comprison and three variable comprison, I finlly find the relationship between borrower rate, number of investors, ratings, creidt socre and loan amount. I think these are important factors which will most affect the profit of the loan.

After I finished my project, I got some future ideas about this database. When I do the quick reserch, I realized that there was a global financial crisis in 2008, but it was not obvious on these database. So if I will continue to do future work on this dataset, I will split the timeframe between 2004-2009 and 2009-2014 to see if there was any financial pressure or effects on number of borrowers, borrower rate and number of investors.