

Peer 2 Peer Lending Data Analysis

Homer Kay

5/3/2018

```
#### DATA CLEANING ####
LD <- read.csv("~/Desktop/Data Mining/prosperLoanData.csv")
#Remove Listing Key, Listing Number, CreditGrade, Prosper Rating (Alpha)
LD <- LD[,c(-1,-2,-4,-15)]
str(LD)

## 'data.frame': 113937 obs. of 77 variables:
## $ ListingCreationDate : Factor w/ 113064 levels "2005-11-0
9 20:44:28.847000000",...: 14184 111894 6429 64760 85967 100310 72556 74019 97
834 97834 ...
## $ Term : int 36 36 36 36 36 60 36 36 36 36
...
## $ LoanStatus : Factor w/ 12 levels "Cancelled","C
hargedoff",...: 3 4 3 4 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 1 ...
## $ BorrowerAPR : num 0.165 0.12 0.283 0.125 0.246
...
## $ BorrowerRate : num 0.158 0.092 0.275 0.0974 0.20
85 ...
## $ LenderYield : num 0.138 0.082 0.24 0.0874 0.198
5 ...
## $ 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 ...
## $ 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 1 ...
```

```

## $ GroupKey : Factor w/ 707 levels "", "003433769
01312423168731",...: 1 1 335 1 1 1 1 1 1 1 ...
## $ DateCreditPulled : Factor w/ 112992 levels "2005-11-0
9 00:30:04.487000000",...: 14347 111883 6446 64724 85857 100382 72500 73937 97
888 97888 ...
## $ CreditScoreRangeLower : int 640 680 480 800 680 740 680 7
00 820 820 ...
## $ CreditScoreRangeUpper : int 659 699 499 819 699 759 699 7
19 839 839 ...
## $ 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 2 2 ...
## $ StatedMonthlyIncome : num 3083 6125 2083 2875 9583 ...
## $ LoanKey : Factor w/ 113066 levels "000036836
05746079487FF7",...: 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 10267
0 123257 88353 90051 121268 121268 ...
## $ LoanOriginalAmount                    : int  9425 10000 3001 10000 15000 1
5000 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 "0000339769
7413387CAF966",...: 11071 10302 33781 54939 19465 48037 60448 40951 26129 2612
9 ...
## $ 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 ...

```

summary(LD)

```

##                               ListingCreationDate      Term
## 2013-10-02 17:20:16.550000000:      6 Min.      :12.00
## 2013-08-28 20:31:41.107000000:      4 1st Qu.:36.00
## 2013-09-08 09:27:44.853000000:      4 Median :36.00
## 2013-12-06 05:43:13.830000000:      4 Mean   :40.83
## 2013-12-06 11:44:58.283000000:      4 3rd Qu.:36.00
## 2013-08-21 07:25:22.360000000:      3 Max.    :60.00
## (Other)                               :113912

```

```

##          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. ProsperScore ListingCategory..numeric.
## Min. :1.000  Min. : 1.00  Min. : 0.000
## 1st Qu.:3.000  1st Qu.: 4.00  1st Qu.: 1.000
## Median :4.000  Median : 6.00  Median : 1.000
## Mean :4.072  Mean : 5.95  Mean : 2.774
## 3rd Qu.:5.000  3rd Qu.: 8.00  3rd Qu.: 3.000
## Max. :7.000  Max. :11.00  Max. :20.000
## NA's :29084  NA's :29084
## BorrowerState Occupation EmploymentStatus
## CA :14717 Other :28617 Employed :67322
## TX : 6842 Professional :13628 Full-time :26355
## NY : 6729 Computer Programmer : 4478 Self-employed: 6134
## FL : 6720 Executive : 4311 Not available: 5347
## IL : 5921 Teacher : 3759 Other : 3806
## : 5515 Administrative Assistant: 3688 : 2255
## (Other):67493 (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
## ScoreExchangeAtTimeOfListing 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 Min. : 1
## 1st Qu.: 9.00 1st Qu.: 6.0 1st Qu.: 37332
## Median :14.00 Median : 21.0 Median : 68599
## Mean :16.27 Mean : 31.9 Mean : 69444
## 3rd Qu.:22.00 3rd Qu.: 65.0 3rd Qu.:101901
## Max. :44.00 Max. :100.0 Max. :136486
## 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 Min. : -2.35
## 16083364744933457E57FB9: 8 1st Qu.: 131.6 1st Qu.: 1005.76
## 3A2F3380477699707C81385: 8 Median : 217.7 Median : 2583.83
## 4D9C3403302047712AD0CDD: 8 Mean : 272.5 Mean : 4183.08
## 739C338135235294782AE75: 8 3rd Qu.: 371.6 3rd Qu.: 5548.40
## 7E1733653050264822FAA3D: 8 Max. :2251.5 Max. :40702.39
## (Other) :113888
## LP_CustomerPrincipalPayments LP_InterestandFees LP_ServiceFees
## Min. : 0.0 Min. : -2.35 Min. : -664.87
## 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
##

## There are 25 missing values in APR.
str(LD[,c("BorrowerAPR")])

## num [1:113937] 0.165 0.12 0.283 0.125 0.246 ...

summary(LD[,c("BorrowerAPR")])

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's
## 0.00653 0.15629 0.20976 0.21883 0.28381 0.51229 25

## Imputing the average fee based on APR and Rate variables
FEE <- LD[,c("BorrowerAPR")] - LD[,c("BorrowerRate")]
FEE.avg <- mean(FEE,na.rm = TRUE)
## Average is 0.026
## Replace Values with FEE.avg + BorrowerRate
## Did this in excel
LD2 <- read.csv("~/Desktop/Data Mining/prosperLoanData2.csv")
LD2.red <- LD2[,c(-48, -49, -50, -51, -52, -53, -54, -55)]
summary(LD2.red)

## ListingCreationDate Term LoanStatus
## 20:28.5: 12 Min. :12.00 Current :56576
## 20:16.5: 11 1st Qu.:36.00 Completed :19664
## 24:18.7: 11 Median :36.00 Chargedoff : 5336
## 04:49.3: 10 Mean :42.49 Defaulted : 1005
## 09:43.7: 10 3rd Qu.:60.00 Past Due (1-15 days) : 806

```



```

## 19:42.9: 10 Max. :60.00 Past Due (31-60 days): 363
## (Other):84789 (Other) : 1103
## ClosedDate BorrowerAPR BorrowerRate LenderYield
## :58848 Min. :0.04583 Min. :0.0400 Min. :0.0300
## 3/4/14 0:00 : 105 1st Qu.:0.16328 1st Qu.:0.1359 1st Qu.:0.1259
## 2/19/14 0:00 : 100 Median :0.21945 Median :0.1875 Median :0.1775
## 2/11/14 0:00 : 92 Mean :0.22666 Mean :0.1960 Mean :0.1860
## 10/30/12 0:00: 81 3rd Qu.:0.29254 3rd Qu.:0.2574 3rd Qu.:0.2474
## 2/26/13 0:00 : 78 Max. :0.42395 Max. :0.3600 Max. :0.3400
## (Other) :25549
## EstimatedEffectiveYield EstimatedLoss EstimatedReturn
## Min. :-0.1827 Min. :0.00490 Min. :-0.18270
## 1st Qu.: 0.1157 1st Qu.:0.04240 1st Qu.: 0.07408
## Median : 0.1615 Median :0.07240 Median : 0.09170
## Mean : 0.1687 Mean :0.08031 Mean : 0.09607
## 3rd Qu.: 0.2243 3rd Qu.:0.11200 3rd Qu.: 0.11660
## Max. : 0.3199 Max. :0.36600 Max. : 0.28370
##
## ProsperRating..numeric. ProsperScore ListingCategory..numeric.
## Min. :1.000 Min. : 1.00 Min. : 0.000
## 1st Qu.:3.000 1st Qu.: 4.00 1st Qu.: 1.000
## Median :4.000 Median : 6.00 Median : 1.000
## Mean :4.072 Mean : 5.95 Mean : 3.313
## 3rd Qu.:5.000 3rd Qu.: 8.00 3rd Qu.: 3.000
## Max. :7.000 Max. :11.00 Max. :20.000
##
## BorrowerState Occupation EmploymentStatus
## CA :10761 Other :21317 Employed :67310
## NY : 5844 Professional :10542 Full-time : 7927
## TX : 5634 Executive : 3468 Not employed : 649
## FL : 5406 Computer Programmer: 3236 Other : 3806
## IL : 4264 Teacher : 2888 Part-time : 256
## OH : 3373 Analyst : 2735 Retired : 367
## (Other):49571 (Other) :40667 Self-employed: 4538
## EmploymentStatusDuration IsBorrowerHomeowner CurrentlyInGroup
## Min. : 0.0 Mode :logical Mode :logical
## 1st Qu.: 30.0 FALSE:40005 FALSE:82607
## Median : 74.0 TRUE :44848 TRUE :2246
## Mean :103.1
## 3rd Qu.:148.0
## Max. :755.0
## NA's :19
##
## GroupKey DateCreditPulled
## :82827 11/4/13 14:12: 8
## 3D4D3366260257624AB272D: 310 12/23/13 9:38: 6
## 783C3371218786870A73D20: 208 9/25/13 15:33: 6
## 52EA3425051368132B80C96: 150 11/1/13 13:02: 5
## B0473364376920128370B13: 83 11/4/13 15:10: 5
## FEF83377364176536637E50: 82 12/9/13 10:12: 5
## (Other) : 1193 (Other) :84818

```

```

## CreditScoreRangeLower CreditScoreRangeUpper FirstRecordedCreditLine
## Min. :600.0 Min. :619.0 12/1/93 0:00: 148
## 1st Qu.:660.0 1st Qu.:679.0 11/1/94 0:00: 139
## Median :700.0 Median :719.0 11/1/95 0:00: 125
## Mean :699.4 Mean :718.4 3/1/96 0:00: 123
## 3rd Qu.:720.0 3rd Qu.:739.0 3/1/95 0:00: 121
## Max. :880.0 Max. :899.0 11/1/89 0:00: 118
## (Other) :84079
## CurrentCreditLines OpenCreditLines TotalCreditLinespast7years
## Min. : 0.00 Min. : 0.000 Min. : 2.00
## 1st Qu.: 7.00 1st Qu.: 6.000 1st Qu.: 18.00
## Median :10.00 Median : 9.000 Median : 26.00
## Mean :10.51 Mean : 9.529 Mean : 27.66
## 3rd Qu.:13.00 3rd Qu.:12.000 3rd Qu.: 35.00
## Max. :59.00 Max. :54.000 Max. :125.00
##
## OpenRevolvingAccounts OpenRevolvingMonthlyPayment InquiriesLast6Months
## Min. : 0.000 Min. : 0.0 Min. : 0.0000
## 1st Qu.: 4.000 1st Qu.: 156.0 1st Qu.: 0.0000
## Median : 7.000 Median : 311.0 Median : 0.0000
## Mean : 7.386 Mean : 430.7 Mean : 0.9646
## 3rd Qu.:10.000 3rd Qu.: 564.0 3rd Qu.: 1.0000
## Max. :50.000 Max. :13765.0 Max. :27.0000
##
## TotalInquiries CurrentDelinquencies AmountDelinquent
## Min. : 0.00 Min. : 0.0000 Min. : 0.0
## 1st Qu.: 2.00 1st Qu.: 0.0000 1st Qu.: 0.0
## Median : 3.00 Median : 0.0000 Median : 0.0
## Mean : 4.29 Mean : 0.3225 Mean : 950.8
## 3rd Qu.: 6.00 3rd Qu.: 0.0000 3rd Qu.: 0.0
## Max. :78.00 Max. :51.0000 Max. :463881.0
##
## DelinquenciesLast7Years PublicRecordsLast10Years
## Min. : 0.000 Min. : 0.0000
## 1st Qu.: 0.000 1st Qu.: 0.0000
## Median : 0.000 Median : 0.0000
## Mean : 3.659 Mean : 0.2851
## 3rd Qu.: 2.000 3rd Qu.: 0.0000
## Max. :99.000 Max. :38.0000
##
## PublicRecordsLast12Months RevolvingCreditBalance BankcardUtilization
## Min. : 0.000000 Min. : 0 Min. :0.0000
## 1st Qu.: 0.000000 1st Qu.: 3824 1st Qu.:0.3300
## Median : 0.000000 Median : 9323 Median :0.6000
## Mean : 0.009169 Mean : 17940 Mean :0.5642
## 3rd Qu.: 0.000000 3rd Qu.: 20339 3rd Qu.:0.8300
## Max. :20.000000 Max. :999165 Max. :2.5000
##
## AvailableBankcardCredit TotalTrades
## Min. : 0 Min. : 1.00

```

```

## 1st Qu.: 1148          1st Qu.: 16.00
## Median : 4574          Median : 23.00
## Mean   : 11401         Mean   : 23.93
## 3rd Qu.: 13910         3rd Qu.: 31.00
## Max.   :498374         Max.    :122.00
##
## TradesNeverDelinquent..percentage. TradesOpenedLast6Months
## Min.    :0.0800         Min.    : 0.0000
## 1st Qu.:0.8500         1st Qu.: 0.0000
## Median :0.9500         Median : 0.0000
## Mean    :0.9059         Mean    : 0.7298
## 3rd Qu.:1.0000         3rd Qu.: 1.0000
## Max.    :1.0000         Max.    :20.0000
##
## DebtToIncomeRatio      IncomeRange      IncomeVerifiable
## Min.    : 0.000        $0          : 45      Mode :logical
## 1st Qu.: 0.150        $1-24,999   : 4654    FALSE:7333
## Median : 0.220        $100,000+   :15205    TRUE :77520
## Mean    : 0.259        $25,000-49,999:24175
## 3rd Qu.: 0.320        $50,000-74,999:25627
## Max.    :10.010       $75,000-99,999:14498
## NA's    :7296         Not employed : 649
## StatedMonthlyIncome      LoanKey
## Min.    : 0          CB1B37030986463208432A1: 6
## 1st Qu.: 3434        2DEE3698211017519D7333F: 4
## Median : 5000        9F4B37043517554537C364C: 4
## Mean    : 5931        D895370150591392337ED6D: 4
## 3rd Qu.: 7083        E6FB37073953690388BC56D: 4
## Max.    :1750003      0D8F37036734373301ED419: 3
##                      (Other)          :84828
## LoanCurrentDaysDelinquent LoanFirstDefaultedCycleNumber
## Min.    : 0.00        Min.    : 1.00
## 1st Qu.: 0.00        1st Qu.: 9.00
## Median : 0.00        Median :13.00
## Mean    : 36.63       Mean    :14.47
## 3rd Qu.: 0.00        3rd Qu.:19.00
## Max.    :1593.00      Max.    :41.00
##                      NA's    :78609
## LoanMonthsSinceOrigination LoanNumber      LoanOriginalAmount
## Min.    : 0.00        Min.    : 38045    Min.    : 1000
## 1st Qu.: 4.00        1st Qu.: 60950    1st Qu.: 4000
## Median :11.00        Median : 87789    Median : 7500
## Mean    :16.02       Mean    : 86592    Mean    : 9083
## 3rd Qu.:25.00       3rd Qu.:109096    3rd Qu.:13500
## Max.    :56.00       Max.    :136486    Max.    :35000
##
## LoanOriginationDate LoanOriginationQuarter
## 1/22/14 0:00 : 491    Q4 2013:14450
## 11/13/13 0:00: 490    Q1 2014:12172
## 2/19/14 0:00 : 439    Q3 2013: 9180

```

```

## 10/16/13 0:00: 434 Q2 2013: 7099
## 1/28/14 0:00 : 339 Q3 2012: 5632
## 9/24/13 0:00 : 316 Q2 2012: 5061
## (Other) :82344 (Other):31259
## MemberKey MonthlyLoanPayment LP_CustomerPayments
## 720D3508651090808DC328F: 7 Min. : 0.0 Min. : -2.35
## C70934206057523078260C7: 7 1st Qu.: 157.3 1st Qu.: 799.39
## E4AF3422677498955FFA00E: 7 Median : 251.9 Median : 2214.76
## 18F6337949289842881D0A8: 6 Mean : 291.9 Mean : 3669.06
## 3D6B34225353312993B9700: 6 3rd Qu.: 388.4 3rd Qu.: 4869.99
## 43DB3366978035224D7D9E3: 6 Max. :2251.5 Max. :37369.16
## (Other) :84814
## LP_CustomerPrincipalPayments LP_InterestandFees LP_ServiceFees
## Min. : 0.0 Min. : -2.35 Min. : -589.95
## 1st Qu.: 393.8 1st Qu.: 254.06 1st Qu.: -72.30
## Median : 1250.0 Median : 674.39 Median : -34.90
## Mean : 2626.9 Mean : 1042.20 Mean : -54.64
## 3rd Qu.: 3414.3 3rd Qu.: 1436.06 3rd Qu.: -14.28
## Max. :35000.0 Max. :10572.78 Max. : 3.01
##
## LP_CollectionFees LP_GrossPrincipalLoss LP_NetPrincipalLoss
## Min. : -4865.080 Min. : -94.2 Min. : -504.4
## 1st Qu.: 0.000 1st Qu.: 0.0 1st Qu.: 0.0
## Median : 0.000 Median : 0.0 Median : 0.0
## Mean : -8.206 Mean : 376.1 Mean : 367.7
## 3rd Qu.: 0.000 3rd Qu.: 0.0 3rd Qu.: 0.0
## Max. : 0.000 Max. :25000.0 Max. :25000.0
##
## LP_NonPrincipalRecoverypayments PercentFunded Recommendations
## Min. : 0.000 Min. :0.7000 Min. : 0.00000
## 1st Qu.: 0.000 1st Qu.:1.0000 1st Qu.: 0.00000
## Median : 0.000 Median :1.0000 Median : 0.00000
## Mean : 7.647 Mean :0.9981 Mean : 0.01756
## 3rd Qu.: 0.000 3rd Qu.:1.0000 3rd Qu.: 0.00000
## Max. :7780.030 Max. :1.0125 Max. :19.00000
##
## InvestmentFromFriendsCount InvestmentFromFriendsAmount Investors
## Min. :0.000000 Min. : 0.000 Min. : 1.00
## 1st Qu.:0.000000 1st Qu.: 0.000 1st Qu.: 1.00
## Median :0.000000 Median : 0.000 Median : 32.00
## Mean :0.008049 Mean : 4.316 Mean : 68.26
## 3rd Qu.:0.000000 3rd Qu.: 0.000 3rd Qu.: 97.00
## Max. :9.000000 Max. :11000.000 Max. :1189.00
##
str(LD2.red)

## 'data.frame': 84853 obs. of 69 variables:
## $ ListingCreationDate : Factor w/ 32723 levels "00:00.0","0
0:00.1",...: 15307 1383 21068 14476 28832 26912 23782 23782 2160 641 ...

```

```

## $ Term : int 36 36 36 60 36 36 36 36 60 36
...
## $ LoanStatus : Factor w/ 11 levels "Chargedoff","Completed",...: 3 3 3 3 3 3 3 3 7 ...
## $ ClosedDate : Factor w/ 1453 levels "", "1/1/12 0:00",...: 1 1 1 1 1 1 1 1 1 ...
## $ BorrowerAPR : num 0.12 0.125 0.246 0.154 0.31 ..
.
## $ BorrowerRate : num 0.092 0.0974 0.2085 0.1314 0.2712 ...
## $ LenderYield : num 0.082 0.0874 0.1985 0.1214 0.2612 ...
## $ EstimatedEffectiveYield : num 0.0796 0.0849 0.1832 0.1157 0.2382 ...
## $ EstimatedLoss : num 0.0249 0.0249 0.0925 0.0449 0.1275 ...
## $ EstimatedReturn : num 0.0547 0.06 0.0907 0.0708 0.1107 ...
## $ ProsperRating..numeric. : int 6 6 3 5 2 4 7 7 4 5 ...
## $ ProsperScore : int 7 9 4 10 2 4 9 11 7 4 ...
## $ ListingCategory..numeric. : int 2 16 2 1 1 2 7 7 1 1 ...
## $ BorrowerState : Factor w/ 48 levels "AK","AL","AR",...: 6 11 22 30 16 5 14 14 20 2 ...
## $ Occupation : Factor w/ 68 levels "", "Accountant/CPA",...: 43 52 21 43 50 29 24 24 22 50 ...
## $ EmploymentStatus : Factor w/ 7 levels "Employed","Full-time",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ EmploymentStatusDuration : int 44 113 44 82 172 103 269 269 300 1 ...
## $ IsBorrowerHomeowner : logi FALSE TRUE TRUE TRUE FALSE FALSE ...
## $ CurrentlyInGroup : logi FALSE FALSE FALSE FALSE FALSE FALSE ...
## $ GroupKey : Factor w/ 221 levels "", "016833805323396548B2370",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ DateCreditPulled : Factor w/ 80238 levels "1/1/10 11:24",...: 37269 12138 74303 26000 45384 53854 27316 27316 46983 26175 ...
## $ CreditScoreRangeLower : int 680 800 680 740 680 700 820 820 640 680 ...
## $ CreditScoreRangeUpper : int 699 819 699 759 699 719 839 839 659 699 ...
## $ FirstRecordedCreditLine : Factor w/ 11096 levels "1/1/00 0:00",...: 4878 4332 4080 4586 10837 4250 5562 5562 2548 2788 ...
## $ CurrentCreditLines : int 14 5 19 21 10 6 17 17 2 9 ...
## $ OpenCreditLines : int 14 5 19 17 7 6 16 16 2 7 ...
## $ TotalCreditLinespast7years : int 29 29 49 49 20 10 32 32 56 29 ...
## $ OpenRevolvingAccounts : int 13 7 6 13 6 5 12 12 4 8 ...
## $ OpenRevolvingMonthlyPayment : int 389 115 220 1410 214 101 219 219 25 290 ...

```

```

## $ InquiriesLast6Months      : int  3 0 1 0 0 3 1 1 1 1 ...
## $ TotalInquiries            : int  5 1 9 2 0 16 6 6 2 4 ...
## $ CurrentDelinquencies      : int  0 4 0 0 0 0 0 0 1 0 ...
## $ AmountDelinquent          : int  0 10056 0 0 0 0 0 0 253 0 ...
## $ DelinquenciesLast7Years   : int  0 14 0 0 0 0 0 0 28 0 ...
## $ PublicRecordsLast10Years  : int  1 0 0 0 0 1 0 0 1 0 ...
## $ PublicRecordsLast12Months : int  0 0 0 0 0 0 0 0 0 0 ...
## $ RevolvingCreditBalance    : int  3989 1444 6193 62999 5812 1260
9906 9906 387 8624 ...
## $ BankcardUtilization       : num  0.21 0.04 0.81 0.39 0.72 0.13
0.11 0.11 0.51 0.7 ...
## $ AvailableBankcardCredit   : int  10266 30754 695 86509 1929 218
1 77696 77696 363 3626 ...
## $ TotalTrades               : int  29 26 39 47 16 10 29 29 47 20
...
## $ TradesNeverDelinquent..percentage.: num  1 0.76 0.95 1 0.68 0.8 1 1 0.7
2 1 ...
## $ TradesOpenedLast6Months   : int  2 0 2 0 0 0 1 1 0 0 ...
## $ DebtToIncomeRatio         : num  0.18 0.15 0.26 0.36 0.27 0.24
0.25 0.25 0.12 0.18 ...
## $ IncomeRange               : Factor w/ 7 levels "$0 ", "$1-24,999
",...: 5 4 3 3 4 4 4 6 4 ...
## $ IncomeVerifiable          : logi  TRUE TRUE TRUE TRUE TRUE TRUE
...
## $ StatedMonthlyIncome       : num  6125 2875 9583 8333 2083 ...
## $ LoanKey                   : Factor w/ 83962 levels "00003683605
746079487FF7",...: 51773 52467 52917 64201 67774 3949 648 648 40495 10635 ...
## $ LoanCurrentDaysDelinquent : int  0 0 0 0 0 0 0 0 0 3 ...
## $ LoanFirstDefaultedCycleNumber : int  NA NA NA NA NA NA NA NA NA NA
...
## $ LoanMonthsSinceOrigination : int  0 16 6 3 11 10 3 3 22 2 ...
## $ LoanNumber                : int  134815 77296 102670 123257 883
53 90051 121268 121268 65946 125045 ...
## $ LoanOriginalAmount        : int  10000 10000 15000 15000 3000 1
0000 10000 10000 13500 4000 ...
## $ LoanOriginationDate       : Factor w/ 1165 levels "1/10/11 0:00
",...: 577 209 1108 365 632 704 317 317 715 95 ...
## $ LoanOriginationQuarter    : Factor w/ 19 levels "Q1 2010", "Q1 2
011",...: 5 18 14 19 9 9 19 19 8 5 ...
## $ MemberKey                 : Factor w/ 68984 levels "00003529701
5484885C64F8",...: 7723 41723 14645 36439 45877 31041 19702 19702 19611 42972
...
## $ MonthlyLoanPayment        : num  319 321 564 342 123 ...
## $ LP_CustomerPayments       : num  0 5143 2820 679 1227 ...
## $ LP_CustomerPrincipalPayments : num  0 4091 1563 352 604 ...
## $ LP_InterestandFees        : num  0 1052 1257 327 622 ...
## $ LP_ServiceFees            : num  0 -108 -60.3 -25.3 -22.9 ...
## $ 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 1 158 20 1 1 1 1 1 19 1 ...
```

##Remove LoanFirstDefaultedCycle Variable (All Missing Values, same as Loan current days delinquent)

```
LD2.red <- LD2.red[,c(-49)]
```

```
summary(LD2.red)
```

```
## ListingCreationDate      Term      LoanStatus
## 20:28.5: 12      Min.   :12.00      Current      :56576
## 20:16.5: 11      1st Qu.:36.00      Completed    :19664
## 24:18.7: 11      Median :36.00      Chargedoff   : 5336
## 04:49.3: 10      Mean   :42.49      Defaulted    : 1005
## 09:43.7: 10      3rd Qu.:60.00      Past Due (1-15 days) : 806
## 19:42.9: 10      Max.   :60.00      Past Due (31-60 days): 363
## (Other):84789      (Other)      : 1103
##      ClosedDate      BorrowerAPR      BorrowerRate      LenderYield
##      :58848      Min.   :0.04583      Min.   :0.0400      Min.   :0.0300
## 3/4/14 0:00 : 105      1st Qu.:0.16328      1st Qu.:0.1359      1st Qu.:0.1259
## 2/19/14 0:00 : 100      Median :0.21945      Median :0.1875      Median :0.1775
## 2/11/14 0:00 : 92      Mean   :0.22666      Mean   :0.1960      Mean   :0.1860
## 10/30/12 0:00: 81      3rd Qu.:0.29254      3rd Qu.:0.2574      3rd Qu.:0.2474
## 2/26/13 0:00 : 78      Max.   :0.42395      Max.   :0.3600      Max.   :0.3400
## (Other)      :25549
## EstimatedEffectiveYield EstimatedLoss      EstimatedReturn
## Min.   :-0.1827      Min.   :0.00490      Min.   :-0.18270
## 1st Qu.: 0.1157      1st Qu.:0.04240      1st Qu.: 0.07408
## Median : 0.1615      Median :0.07240      Median : 0.09170
## Mean   : 0.1687      Mean   :0.08031      Mean   : 0.09607
## 3rd Qu.: 0.2243      3rd Qu.:0.11200      3rd Qu.: 0.11660
## Max.   : 0.3199      Max.   :0.36600      Max.   : 0.28370
##
## ProsperRating..numeric. ProsperScore      ListingCategory..numeric.
## Min.   :1.000      Min.   : 1.00      Min.   : 0.000
## 1st Qu.:3.000      1st Qu.: 4.00      1st Qu.: 1.000
## Median :4.000      Median : 6.00      Median : 1.000
## Mean   :4.072      Mean   : 5.95      Mean   : 3.313
## 3rd Qu.:5.000      3rd Qu.: 8.00      3rd Qu.: 3.000
## Max.   :7.000      Max.   :11.00      Max.   :20.000
##
## BorrowerState      Occupation      EmploymentStatus
## CA      :10761      Other      :21317      Employed      :67310
## NY      : 5844      Professional :10542      Full-time     : 7927
## TX      : 5634      Executive   : 3468      Not employed  : 649
## FL      : 5406      Computer Programmer: 3236      Other         : 3806
## IL      : 4264      Teacher     : 2888      Part-time     : 256
```

```

## OH      : 3373   Analyst      : 2735   Retired      : 367
## (Other):49571   (Other)      :40667   Self-employed: 4538
## EmploymentStatusDuration IsBorrowerHomeowner CurrentlyInGroup
## Min.    : 0.0      Mode :logical      Mode :logical
## 1st Qu.: 30.0      FALSE:40005      FALSE:82607
## Median  : 74.0      TRUE :44848      TRUE :2246
## Mean    :103.1
## 3rd Qu.:148.0
## Max.    :755.0
## NA's    :19
##
##              GroupKey          DateCreditPulled
##              :82827    11/4/13 14:12:      8
## 3D4D3366260257624AB272D: 310    12/23/13 9:38:      6
## 783C3371218786870A73D20: 208    9/25/13 15:33:      6
## 52EA3425051368132B80C96: 150    11/1/13 13:02:      5
## B0473364376920128370B13: 83     11/4/13 15:10:      5
## FEF83377364176536637E50: 82     12/9/13 10:12:      5
## (Other)      : 1193   (Other)      :84818
## CreditScoreRangeLower CreditScoreRangeUpper FirstRecordedCreditLine
## Min.    :600.0      Min.    :619.0      12/1/93 0:00: 148
## 1st Qu.:660.0      1st Qu.:679.0      11/1/94 0:00: 139
## Median  :700.0      Median  :719.0      11/1/95 0:00: 125
## Mean    :699.4      Mean    :718.4      3/1/96 0:00: 123
## 3rd Qu.:720.0      3rd Qu.:739.0      3/1/95 0:00: 121
## Max.    :880.0      Max.    :899.0      11/1/89 0:00: 118
## (Other)      :84079
## CurrentCreditLines OpenCreditLines TotalCreditLinespast7years
## Min.    : 0.00      Min.    : 0.000      Min.    : 2.00
## 1st Qu.: 7.00      1st Qu.: 6.000      1st Qu.: 18.00
## Median  :10.00      Median  : 9.000      Median  : 26.00
## Mean    :10.51      Mean    : 9.529      Mean    : 27.66
## 3rd Qu.:13.00      3rd Qu.:12.000      3rd Qu.: 35.00
## Max.    :59.00      Max.    :54.000      Max.    :125.00
##
## OpenRevolvingAccounts OpenRevolvingMonthlyPayment InquiriesLast6Months
## Min.    : 0.000      Min.    : 0.0      Min.    : 0.0000
## 1st Qu.: 4.000      1st Qu.: 156.0      1st Qu.: 0.0000
## Median  : 7.000      Median  : 311.0      Median  : 0.0000
## Mean    : 7.386      Mean    : 430.7      Mean    : 0.9646
## 3rd Qu.:10.000      3rd Qu.: 564.0      3rd Qu.: 1.0000
## Max.    :50.000      Max.    :13765.0      Max.    :27.0000
##
## TotalInquiries CurrentDelinquencies AmountDelinquent
## Min.    : 0.00      Min.    : 0.0000      Min.    : 0.0
## 1st Qu.: 2.00      1st Qu.: 0.0000      1st Qu.: 0.0
## Median  : 3.00      Median  : 0.0000      Median  : 0.0
## Mean    : 4.29      Mean    : 0.3225      Mean    : 950.8
## 3rd Qu.: 6.00      3rd Qu.: 0.0000      3rd Qu.: 0.0
## Max.    :78.00      Max.    :51.0000      Max.    :463881.0
##

```



```

## DelinquenciesLast7Years PublicRecordsLast10Years
## Min. : 0.000 Min. : 0.0000
## 1st Qu.: 0.000 1st Qu.: 0.0000
## Median : 0.000 Median : 0.0000
## Mean : 3.659 Mean : 0.2851
## 3rd Qu.: 2.000 3rd Qu.: 0.0000
## Max. :99.000 Max. :38.0000
##
## PublicRecordsLast12Months RevolvingCreditBalance BankcardUtilization
## Min. : 0.000000 Min. : 0 Min. :0.0000
## 1st Qu.: 0.000000 1st Qu.: 3824 1st Qu.:0.3300
## Median : 0.000000 Median : 9323 Median :0.6000
## Mean : 0.009169 Mean : 17940 Mean :0.5642
## 3rd Qu.: 0.000000 3rd Qu.: 20339 3rd Qu.:0.8300
## Max. :20.000000 Max. :999165 Max. :2.5000
##
## AvailableBankcardCredit TotalTrades
## Min. : 0 Min. : 1.00
## 1st Qu.: 1148 1st Qu.: 16.00
## Median : 4574 Median : 23.00
## Mean : 11401 Mean : 23.93
## 3rd Qu.: 13910 3rd Qu.: 31.00
## Max. :498374 Max. :122.00
##
## TradesNeverDelinquent..percentage. TradesOpenedLast6Months
## Min. :0.0800 Min. : 0.0000
## 1st Qu.:0.8500 1st Qu.: 0.0000
## Median :0.9500 Median : 0.0000
## Mean :0.9059 Mean : 0.7298
## 3rd Qu.:1.0000 3rd Qu.: 1.0000
## Max. :1.0000 Max. :20.0000
##
## DebtToIncomeRatio IncomeRange IncomeVerifiable
## Min. : 0.000 $0 : 45 Mode :logical
## 1st Qu.: 0.150 $1-24,999 : 4654 FALSE:7333
## Median : 0.220 $100,000+ :15205 TRUE :77520
## Mean : 0.259 $25,000-49,999:24175
## 3rd Qu.: 0.320 $50,000-74,999:25627
## Max. :10.010 $75,000-99,999:14498
## NA's :7296 Not employed : 649
## StatedMonthlyIncome LoanKey
## Min. : 0 CB1B37030986463208432A1: 6
## 1st Qu.: 3434 2DEE3698211017519D7333F: 4
## Median : 5000 9F4B37043517554537C364C: 4
## Mean : 5931 D895370150591392337ED6D: 4
## 3rd Qu.: 7083 E6FB37073953690388BC56D: 4
## Max. :1750003 0D8F37036734373301ED419: 3
## (Other) :84828
## LoanCurrentDaysDelinquent LoanMonthsSinceOrigination LoanNumber
## Min. : 0.00 Min. : 0.00 Min. : 38045

```

```

## 1st Qu.: 0.00          1st Qu.: 4.00          1st Qu.: 60950
## Median : 0.00          Median :11.00         Median : 87789
## Mean   : 36.63         Mean   :16.02         Mean   : 86592
## 3rd Qu.: 0.00          3rd Qu.:25.00        3rd Qu.:109096
## Max.   :1593.00        Max.   :56.00         Max.   :136486
##
## LoanOriginalAmount    LoanOriginationDate LoanOriginationQuarter
## Min.   : 1000          1/22/14 0:00 : 491    Q4 2013:14450
## 1st Qu.: 4000          11/13/13 0:00: 490    Q1 2014:12172
## Median : 7500          2/19/14 0:00 : 439    Q3 2013: 9180
## Mean   : 9083          10/16/13 0:00: 434    Q2 2013: 7099
## 3rd Qu.:13500          1/28/14 0:00 : 339    Q3 2012: 5632
## Max.   :35000          9/24/13 0:00 : 316    Q2 2012: 5061
##                               (Other) :82344    (Other):31259
##                               MemberKey MonthlyLoanPayment LP_CustomerPayments
## 720D3508651090808DC328F: 7 Min.   : 0.0    Min.   : -2.35
## C70934206057523078260C7: 7 1st Qu.: 157.3  1st Qu.: 799.39
## E4AF3422677498955FFA00E: 7 Median : 251.9  Median : 2214.76
## 18F6337949289842881D0A8: 6 Mean   : 291.9  Mean   : 3669.06
## 3D6B34225353312993B9700: 6 3rd Qu.: 388.4  3rd Qu.: 4869.99
## 43DB3366978035224D7D9E3: 6 Max.   :2251.5  Max.   :37369.16
## (Other)                  :84814
## LP_CustomerPrincipalPayments LP_InterestandFees LP_ServiceFees
## Min.   : 0.0    Min.   : -2.35  Min.   : -589.95
## 1st Qu.: 393.8  1st Qu.: 254.06 1st Qu.: -72.30
## Median : 1250.0 Median : 674.39 Median : -34.90
## Mean   : 2626.9 Mean   : 1042.20 Mean   : -54.64
## 3rd Qu.: 3414.3 3rd Qu.: 1436.06 3rd Qu.: -14.28
## Max.   :35000.0 Max.   :10572.78 Max.   : 3.01
##
## LP_CollectionFees    LP_GrossPrincipalLoss LP_NetPrincipalLoss
## Min.   : -4865.080  Min.   : -94.2    Min.   : -504.4
## 1st Qu.: 0.000      1st Qu.: 0.0      1st Qu.: 0.0
## Median : 0.000      Median : 0.0      Median : 0.0
## Mean   : -8.206     Mean   : 376.1     Mean   : 367.7
## 3rd Qu.: 0.000      3rd Qu.: 0.0      3rd Qu.: 0.0
## Max.   : 0.000      Max.   :25000.0    Max.   :25000.0
##
## LP_NonPrincipalRecoverypayments PercentFunded Recommendations
## Min.   : 0.000      Min.   :0.7000    Min.   : 0.00000
## 1st Qu.: 0.000      1st Qu.:1.0000    1st Qu.: 0.00000
## Median : 0.000      Median :1.0000    Median : 0.00000
## Mean   : 7.647      Mean   :0.9981     Mean   : 0.01756
## 3rd Qu.: 0.000      3rd Qu.:1.0000    3rd Qu.: 0.00000
## Max.   :7780.030    Max.   :1.0125     Max.   :19.00000
##
## InvestmentFromFriendsCount InvestmentFromFriendsAmount Investors
## Min.   :0.000000    Min.   : 0.000    Min.   : 1.00
## 1st Qu.:0.000000    1st Qu.: 0.000    1st Qu.: 1.00
## Median :0.000000    Median : 0.000    Median : 32.00

```

```
## Mean :0.008049      Mean : 4.316      Mean : 68.26
## 3rd Qu.:0.000000    3rd Qu.: 0.000      3rd Qu.: 97.00
## Max. :9.000000      Max. :11000.000     Max. :1189.00
##

## Impute mean of Debt/Income Ratio into 7000 missing values
## Tried imputing other methods such as LoanPayment/MonthlyIncome, however these values were often far off.
## However imputing 7000 rows is considerable. Keeping in mind for analysis.
LD2.red$DebtToIncomeRatio.imp <- ifelse(is.na(LD2.red$DebtToIncomeRatio), mean(LD2.red$DebtToIncomeRatio, na.rm=TRUE), LD2.red$DebtToIncomeRatio)
## Check it worked
summary(LD2.red$DebtToIncomeRatio.imp)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0000  0.1600  0.2400  0.2588  0.3100 10.0100

summary(LD2.red)

## ListingCreationDate      Term      LoanStatus
## 20:28.5: 12      Min. :12.00      Current      :56576
## 20:16.5: 11      1st Qu.:36.00      Completed      :19664
## 24:18.7: 11      Median :36.00      Chargedoff      : 5336
## 04:49.3: 10      Mean :42.49      Defaulted      : 1005
## 09:43.7: 10      3rd Qu.:60.00      Past Due (1-15 days) : 806
## 19:42.9: 10      Max. :60.00      Past Due (31-60 days): 363
## (Other):84789      (Other)      : 1103
##      ClosedDate      BorrowerAPR      BorrowerRate      LenderYield
##      :58848      Min. :0.04583      Min. :0.0400      Min. :0.0300
## 3/4/14 0:00 : 105      1st Qu.:0.16328      1st Qu.:0.1359      1st Qu.:0.1259
## 2/19/14 0:00 : 100      Median :0.21945      Median :0.1875      Median :0.1775
## 2/11/14 0:00 : 92      Mean :0.22666      Mean :0.1960      Mean :0.1860
## 10/30/12 0:00: 81      3rd Qu.:0.29254      3rd Qu.:0.2574      3rd Qu.:0.2474
## 2/26/13 0:00 : 78      Max. :0.42395      Max. :0.3600      Max. :0.3400
## (Other) :25549
## EstimatedEffectiveYield EstimatedLoss      EstimatedReturn
## Min. : -0.1827      Min. :0.00490      Min. : -0.18270
## 1st Qu.: 0.1157      1st Qu.:0.04240      1st Qu.: 0.07408
## Median : 0.1615      Median :0.07240      Median : 0.09170
## Mean : 0.1687      Mean :0.08031      Mean : 0.09607
## 3rd Qu.: 0.2243      3rd Qu.:0.11200      3rd Qu.: 0.11660
## Max. : 0.3199      Max. :0.36600      Max. : 0.28370
##
## ProsperRating..numeric. ProsperScore      ListingCategory..numeric.
## Min. :1.000      Min. : 1.00      Min. : 0.000
## 1st Qu.:3.000      1st Qu.: 4.00      1st Qu.: 1.000
## Median :4.000      Median : 6.00      Median : 1.000
## Mean :4.072      Mean : 5.95      Mean : 3.313
## 3rd Qu.:5.000      3rd Qu.: 8.00      3rd Qu.: 3.000
## Max. :7.000      Max. :11.00      Max. :20.000
##
```

```

## BorrowerState Occupation EmploymentStatus
## CA :10761 Other :21317 Employed :67310
## NY : 5844 Professional :10542 Full-time : 7927
## TX : 5634 Executive : 3468 Not employed : 649
## FL : 5406 Computer Programmer: 3236 Other : 3806
## IL : 4264 Teacher : 2888 Part-time : 256
## OH : 3373 Analyst : 2735 Retired : 367
## (Other):49571 (Other) :40667 Self-employed: 4538
## EmploymentStatusDuration IsBorrowerHomeowner CurrentlyInGroup
## Min. : 0.0 Mode :logical Mode :logical
## 1st Qu.: 30.0 FALSE:40005 FALSE:82607
## Median : 74.0 TRUE :44848 TRUE :2246
## Mean :103.1
## 3rd Qu.:148.0
## Max. :755.0
## NA's :19
## GroupKey DateCreditPulled
## :82827 11/4/13 14:12: 8
## 3D4D3366260257624AB272D: 310 12/23/13 9:38: 6
## 783C3371218786870A73D20: 208 9/25/13 15:33: 6
## 52EA3425051368132B80C96: 150 11/1/13 13:02: 5
## B0473364376920128370B13: 83 11/4/13 15:10: 5
## FEF83377364176536637E50: 82 12/9/13 10:12: 5
## (Other) : 1193 (Other) :84818
## CreditScoreRangeLower CreditScoreRangeUpper FirstRecordedCreditLine
## Min. :600.0 Min. :619.0 12/1/93 0:00: 148
## 1st Qu.:660.0 1st Qu.:679.0 11/1/94 0:00: 139
## Median :700.0 Median :719.0 11/1/95 0:00: 125
## Mean :699.4 Mean :718.4 3/1/96 0:00: 123
## 3rd Qu.:720.0 3rd Qu.:739.0 3/1/95 0:00: 121
## Max. :880.0 Max. :899.0 11/1/89 0:00: 118
## (Other) :84079
## CurrentCreditLines OpenCreditLines TotalCreditLinespast7years
## Min. : 0.00 Min. : 0.000 Min. : 2.00
## 1st Qu.: 7.00 1st Qu.: 6.000 1st Qu.: 18.00
## Median :10.00 Median : 9.000 Median : 26.00
## Mean :10.51 Mean : 9.529 Mean : 27.66
## 3rd Qu.:13.00 3rd Qu.:12.000 3rd Qu.: 35.00
## Max. :59.00 Max. :54.000 Max. :125.00
##
## OpenRevolvingAccounts OpenRevolvingMonthlyPayment InquiriesLast6Months
## Min. : 0.000 Min. : 0.0 Min. : 0.0000
## 1st Qu.: 4.000 1st Qu.: 156.0 1st Qu.: 0.0000
## Median : 7.000 Median : 311.0 Median : 0.0000
## Mean : 7.386 Mean : 430.7 Mean : 0.9646
## 3rd Qu.:10.000 3rd Qu.: 564.0 3rd Qu.: 1.0000
## Max. :50.000 Max. :13765.0 Max. :27.0000
##
## TotalInquiries CurrentDelinquencies AmountDelinquent
## Min. : 0.00 Min. : 0.0000 Min. : 0.0

```

```

## 1st Qu.: 2.00    1st Qu.: 0.0000    1st Qu.:    0.0
## Median : 3.00    Median : 0.0000    Median :    0.0
## Mean   : 4.29    Mean   : 0.3225    Mean   :   950.8
## 3rd Qu.: 6.00    3rd Qu.: 0.0000    3rd Qu.:    0.0
## Max.   :78.00    Max.   :51.0000    Max.   :463881.0
##
## DelinquenciesLast7Years PublicRecordsLast10Years
## Min.    : 0.000    Min.    : 0.0000
## 1st Qu.: 0.000    1st Qu.: 0.0000
## Median : 0.000    Median : 0.0000
## Mean    : 3.659    Mean    : 0.2851
## 3rd Qu.: 2.000    3rd Qu.: 0.0000
## Max.    :99.000    Max.    :38.0000
##
## PublicRecordsLast12Months RevolvingCreditBalance BankcardUtilization
## Min.    : 0.000000    Min.    :    0    Min.    :0.0000
## 1st Qu.: 0.000000    1st Qu.: 3824    1st Qu.:0.3300
## Median : 0.000000    Median : 9323    Median :0.6000
## Mean    : 0.009169    Mean    : 17940    Mean    :0.5642
## 3rd Qu.: 0.000000    3rd Qu.: 20339    3rd Qu.:0.8300
## Max.    :20.000000    Max.    :999165    Max.    :2.5000
##
## AvailableBankcardCredit TotalTrades
## Min.    :    0    Min.    : 1.00
## 1st Qu.: 1148    1st Qu.: 16.00
## Median : 4574    Median : 23.00
## Mean    : 11401    Mean    : 23.93
## 3rd Qu.: 13910    3rd Qu.: 31.00
## Max.    :498374    Max.    :122.00
##
## TradesNeverDelinquent..percentage. TradesOpenedLast6Months
## Min.    :0.0800    Min.    : 0.0000
## 1st Qu.:0.8500    1st Qu.: 0.0000
## Median :0.9500    Median : 0.0000
## Mean    :0.9059    Mean    : 0.7298
## 3rd Qu.:1.0000    3rd Qu.: 1.0000
## Max.    :1.0000    Max.    :20.0000
##
## DebtToIncomeRatio IncomeRange IncomeVerifiable
## Min.    : 0.000    $0 : 45 Mode :logical
## 1st Qu.: 0.150    $1-24,999 : 4654 FALSE:7333
## Median : 0.220    $100,000+ :15205 TRUE :77520
## Mean    : 0.259    $25,000-49,999:24175
## 3rd Qu.: 0.320    $50,000-74,999:25627
## Max.    :10.010    $75,000-99,999:14498
## NA's    :7296    Not employed : 649
## StatedMonthlyIncome LoanKey
## Min.    :    0    CB1B37030986463208432A1: 6
## 1st Qu.: 3434    2DEE3698211017519D7333F: 4
## Median : 5000    9F4B37043517554537C364C: 4

```

```

## Mean      : 5931      D895370150591392337ED6D: 4
## 3rd Qu.: 7083      E6FB37073953690388BC56D: 4
## Max.      :1750003    0D8F37036734373301ED419: 3
##
##              (Other)              :84828
## LoanCurrentDaysDelinquent LoanMonthsSinceOrigination LoanNumber
## Min.      : 0.00      Min.      : 0.00      Min.      : 38045
## 1st Qu.: 0.00      1st Qu.: 4.00      1st Qu.: 60950
## Median : 0.00      Median :11.00      Median : 87789
## Mean      : 36.63      Mean      :16.02      Mean      : 86592
## 3rd Qu.: 0.00      3rd Qu.:25.00      3rd Qu.:109096
## Max.      :1593.00      Max.      :56.00      Max.      :136486
##
## LoanOriginalAmount LoanOriginationDate LoanOriginationQuarter
## Min.      : 1000      1/22/14 0:00 : 491      Q4 2013:14450
## 1st Qu.: 4000      11/13/13 0:00: 490      Q1 2014:12172
## Median : 7500      2/19/14 0:00 : 439      Q3 2013: 9180
## Mean      : 9083      10/16/13 0:00: 434      Q2 2013: 7099
## 3rd Qu.:13500      1/28/14 0:00 : 339      Q3 2012: 5632
## Max.      :35000      9/24/13 0:00 : 316      Q2 2012: 5061
##
##              (Other)              :82344      (Other):31259
##              MemberKey      MonthlyLoanPayment LP_CustomerPayments
## 720D3508651090808DC328F: 7 Min.      : 0.0      Min.      : -2.35
## C70934206057523078260C7: 7 1st Qu.: 157.3      1st Qu.: 799.39
## E4AF3422677498955FFA00E: 7 Median : 251.9      Median : 2214.76
## 18F6337949289842881D0A8: 6 Mean      : 291.9      Mean      : 3669.06
## 3D6B34225353312993B9700: 6 3rd Qu.: 388.4      3rd Qu.: 4869.99
## 43DB3366978035224D7D9E3: 6 Max.      :2251.5      Max.      :37369.16
##
##              (Other)              :84814
## LP_CustomerPrincipalPayments LP_InterestandFees LP_ServiceFees
## Min.      : 0.0      Min.      : -2.35      Min.      : -589.95
## 1st Qu.: 393.8      1st Qu.: 254.06      1st Qu.: -72.30
## Median : 1250.0      Median : 674.39      Median : -34.90
## Mean      : 2626.9      Mean      :1042.20      Mean      : -54.64
## 3rd Qu.: 3414.3      3rd Qu.: 1436.06      3rd Qu.: -14.28
## Max.      :35000.0      Max.      :10572.78      Max.      : 3.01
##
## LP_CollectionFees LP_GrossPrincipalLoss LP_NetPrincipalLoss
## Min.      : -4865.080      Min.      : -94.2      Min.      : -504.4
## 1st Qu.: 0.000      1st Qu.: 0.0      1st Qu.: 0.0
## Median : 0.000      Median : 0.0      Median : 0.0
## Mean      : -8.206      Mean      : 376.1      Mean      : 367.7
## 3rd Qu.: 0.000      3rd Qu.: 0.0      3rd Qu.: 0.0
## Max.      : 0.000      Max.      :25000.0      Max.      :25000.0
##
## LP_NonPrincipalRecoverypayments PercentFunded Recommendations
## Min.      : 0.000      Min.      :0.7000      Min.      : 0.00000
## 1st Qu.: 0.000      1st Qu.:1.0000      1st Qu.: 0.00000
## Median : 0.000      Median :1.0000      Median : 0.00000
## Mean      : 7.647      Mean      :0.9981      Mean      : 0.01756
## 3rd Qu.: 0.000      3rd Qu.:1.0000      3rd Qu.: 0.00000

```

```
## Max. :7780.030 Max. :1.0125 Max. :19.00000
##
## InvestmentFromFriendsCount InvestmentFromFriendsAmount Investors
## Min. :0.000000 Min. : 0.000 Min. : 1.00
## 1st Qu.:0.000000 1st Qu.: 0.000 1st Qu.: 1.00
## Median :0.000000 Median : 0.000 Median : 32.00
## Mean :0.008049 Mean : 4.316 Mean : 68.26
## 3rd Qu.:0.000000 3rd Qu.: 0.000 3rd Qu.: 97.00
## Max. :9.000000 Max. :11000.000 Max. :1189.00
##
## DebtToIncomeRatio.imp
## Min. : 0.0000
## 1st Qu.: 0.1600
## Median : 0.2400
## Mean : 0.2588
## 3rd Qu.: 0.3100
## Max. :10.0100
##
```

```
str(LD2.red)
```

```
## 'data.frame': 84853 obs. of 69 variables:
## $ ListingCreationDate : Factor w/ 32723 levels "00:00.0","0
0:00.1",...: 15307 1383 21068 14476 28832 26912 23782 23782 2160 641 ...
## $ Term : int 36 36 36 60 36 36 36 36 60 36
...
## $ LoanStatus : Factor w/ 11 levels "Chargedoff","C
ompleted",...: 3 3 3 3 3 3 3 3 3 7 ...
## $ ClosedDate : Factor w/ 1453 levels "", "1/1/12 0:
00",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ BorrowerAPR : num 0.12 0.125 0.246 0.154 0.31 ..
.
## $ BorrowerRate : num 0.092 0.0974 0.2085 0.1314 0.2
712 ...
## $ LenderYield : num 0.082 0.0874 0.1985 0.1214 0.2
612 ...
## $ EstimatedEffectiveYield : num 0.0796 0.0849 0.1832 0.1157 0.
2382 ...
## $ EstimatedLoss : num 0.0249 0.0249 0.0925 0.0449 0.
1275 ...
## $ EstimatedReturn : num 0.0547 0.06 0.0907 0.0708 0.11
07 ...
## $ ProsperRating..numeric. : int 6 6 3 5 2 4 7 7 4 5 ...
## $ ProsperScore : int 7 9 4 10 2 4 9 11 7 4 ...
## $ ListingCategory..numeric. : int 2 16 2 1 1 2 7 7 1 1 ...
## $ BorrowerState : Factor w/ 48 levels "AK","AL","AR",
...: 6 11 22 30 16 5 14 14 20 2 ...
## $ Occupation : Factor w/ 68 levels "", "Accountant/
CPA",...: 43 52 21 43 50 29 24 24 22 50 ...
## $ EmploymentStatus : Factor w/ 7 levels "Employed","Full
```

```

-time",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ EmploymentStatusDuration          : int   44 113 44 82 172 103 269 269 3
00 1 ...
## $ IsBorrowerHomeowner                : logi   FALSE TRUE TRUE TRUE FALSE FA
LSE ...
## $ CurrentlyInGroup                    : logi   FALSE FALSE FALSE FALSE FALSE
FALSE ...
## $ GroupKey                            : Factor w/ 221 levels "", "0168338053
23396548B2370",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ DateCreditPulled                    : Factor w/ 80238 levels "1/1/10 11:2
4",...: 37269 12138 74303 26000 45384 53854 27316 27316 46983 26175 ...
## $ CreditScoreRangeLower               : int   680 800 680 740 680 700 820 82
0 640 680 ...
## $ CreditScoreRangeUpper               : int   699 819 699 759 699 719 839 83
9 659 699 ...
## $ FirstRecordedCreditLine             : Factor w/ 11096 levels "1/1/00 0:00
",...: 4878 4332 4080 4586 10837 4250 5562 5562 2548 2788 ...
## $ CurrentCreditLines                  : int   14 5 19 21 10 6 17 17 2 9 ...
## $ OpenCreditLines                     : int   14 5 19 17 7 6 16 16 2 7 ...
## $ TotalCreditLinespast7years          : int   29 29 49 49 20 10 32 32 56 29
...
## $ OpenRevolvingAccounts                : int   13 7 6 13 6 5 12 12 4 8 ...
## $ OpenRevolvingMonthlyPayment         : int   389 115 220 1410 214 101 219 2
19 25 290 ...
## $ InquiriesLast6Months                : int    3 0 1 0 0 3 1 1 1 1 ...
## $ TotalInquiries                      : int    5 1 9 2 0 16 6 6 2 4 ...
## $ CurrentDelinquencies                 : int    0 4 0 0 0 0 0 0 1 0 ...
## $ AmountDelinquent                    : int    0 10056 0 0 0 0 0 0 253 0 ...
## $ DelinquenciesLast7Years              : int    0 14 0 0 0 0 0 0 28 0 ...
## $ PublicRecordsLast10Years             : int    1 0 0 0 0 1 0 0 1 0 ...
## $ PublicRecordsLast12Months            : int    0 0 0 0 0 0 0 0 0 0 ...
## $ RevolvingCreditBalance               : int   3989 1444 6193 62999 5812 1260
9906 9906 387 8624 ...
## $ BankcardUtilization                  : num    0.21 0.04 0.81 0.39 0.72 0.13
0.11 0.11 0.51 0.7 ...
## $ AvailableBankcardCredit              : int   10266 30754 695 86509 1929 218
1 77696 77696 363 3626 ...
## $ TotalTrades                          : int    29 26 39 47 16 10 29 29 47 20
...
## $ TradesNeverDelinquent..percentage.: num    1 0.76 0.95 1 0.68 0.8 1 1 0.7
2 1 ...
## $ TradesOpenedLast6Months              : int    2 0 2 0 0 0 1 1 0 0 ...
## $ DebtToIncomeRatio                    : num    0.18 0.15 0.26 0.36 0.27 0.24
0.25 0.25 0.12 0.18 ...
## $ IncomeRange                          : Factor w/ 7 levels "$0 ", "$1-24,999
",...: 5 4 3 3 4 4 4 6 4 ...
## $ IncomeVerifiable                     : logi   TRUE TRUE TRUE TRUE TRUE TRUE
...
## $ StatedMonthlyIncome                  : num   6125 2875 9583 8333 2083 ...
## $ LoanKey                              : Factor w/ 83962 levels "00003683605

```



```

746079487FF7",...: 51773 52467 52917 64201 67774 3949 648 648 40495 10635 ...
## $ LoanCurrentDaysDelinquent      : int  0 0 0 0 0 0 0 0 0 3 ...
## $ LoanMonthsSinceOrigination      : int  0 16 6 3 11 10 3 3 22 2 ...
## $ LoanNumber                      : int  134815 77296 102670 123257 883
53 90051 121268 121268 65946 125045 ...
## $ LoanOriginalAmount              : int  10000 10000 15000 15000 3000 1
0000 10000 10000 13500 4000 ...
## $ LoanOriginationDate              : Factor w/ 1165 levels "1/10/11 0:00
",...: 577 209 1108 365 632 704 317 317 715 95 ...
## $ LoanOriginationQuarter          : Factor w/ 19 levels "Q1 2010","Q1 2
011",...: 5 18 14 19 9 9 19 19 8 5 ...
## $ MemberKey                      : Factor w/ 68984 levels "00003529701
5484885C64F8",...: 7723 41723 14645 36439 45877 31041 19702 19702 19611 42972
...
## $ MonthlyLoanPayment              : num  319 321 564 342 123 ...
## $ LP_CustomerPayments              : num  0 5143 2820 679 1227 ...
## $ LP_CustomerPrincipalPayments     : num  0 4091 1563 352 604 ...
## $ LP_InterestandFees               : num  0 1052 1257 327 622 ...
## $ LP_ServiceFees                   : num  0 -108 -60.3 -25.3 -22.9 ...
## $ 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  1 158 20 1 1 1 1 1 19 1 ...
## $ DebtToIncomeRatio.imp            : num  0.18 0.15 0.26 0.36 0.27 0.24
0.25 0.25 0.12 0.18 ...

```

#Remove original debt/income ratio column.

```
LD2.red <- LD2.red[,c(-43)]
```

```
## Imput EmploymentStatusDuration Missing Values with Median, Missing 19 Values
```

```
LD2.red$EmploymentStatusDuration.imp <- ifelse(is.na(LD2.red$EmploymentStatusDuration),
median(LD2.red$EmploymentStatusDuration, na.rm=TRUE), LD2.red$EmploymentStatusDuration)
```

```
#### Check work
```

```
summary(LD2.red$EmploymentStatusDuration.imp)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##         0      30      74     103     148     755
```

```
## Remove original employment duration
```

```
LD2.red <- LD2.red[,c(-17)]
```

```
summary(LD2.red)
```

```
## ListingCreationDate      Term      LoanStatus
## 20:28.5: 12      Min.    :12.00   Current      :56576
## 20:16.5: 11      1st Qu.:36.00   Completed    :19664
```

```

## 24:18.7: 11      Median :36.00   Chargedoff      : 5336
## 04:49.3: 10      Mean   :42.49   Defaulted       : 1005
## 09:43.7: 10      3rd Qu.:60.00   Past Due (1-15 days) : 806
## 19:42.9: 10      Max.    :60.00   Past Due (31-60 days): 363
## (Other):84789      (Other)      : 1103
##      ClosedDate      BorrowerAPR      BorrowerRate      LenderYield
##      :58848      Min.    :0.04583      Min.    :0.0400      Min.    :0.0300
## 3/4/14 0:00 : 105      1st Qu.:0.16328      1st Qu.:0.1359      1st Qu.:0.1259
## 2/19/14 0:00 : 100      Median :0.21945      Median :0.1875      Median :0.1775
## 2/11/14 0:00 : 92      Mean    :0.22666      Mean    :0.1960      Mean    :0.1860
## 10/30/12 0:00: 81      3rd Qu.:0.29254      3rd Qu.:0.2574      3rd Qu.:0.2474
## 2/26/13 0:00 : 78      Max.    :0.42395      Max.    :0.3600      Max.    :0.3400
## (Other)      :25549
## EstimatedEffectiveYield EstimatedLoss      EstimatedReturn
## Min.    :-0.1827      Min.    :0.00490      Min.    :-0.18270
## 1st Qu.: 0.1157      1st Qu.:0.04240      1st Qu.: 0.07408
## Median : 0.1615      Median :0.07240      Median : 0.09170
## Mean    : 0.1687      Mean    :0.08031      Mean    : 0.09607
## 3rd Qu.: 0.2243      3rd Qu.:0.11200      3rd Qu.: 0.11660
## Max.    : 0.3199      Max.    :0.36600      Max.    : 0.28370
##
## ProsperRating..numeric. ProsperScore      ListingCategory..numeric.
## Min.    :1.000      Min.    : 1.00      Min.    : 0.000
## 1st Qu.:3.000      1st Qu.: 4.00      1st Qu.: 1.000
## Median :4.000      Median : 6.00      Median : 1.000
## Mean    :4.072      Mean    : 5.95      Mean    : 3.313
## 3rd Qu.:5.000      3rd Qu.: 8.00      3rd Qu.: 3.000
## Max.    :7.000      Max.    :11.00      Max.    :20.000
##
## BorrowerState      Occupation      EmploymentStatus
## CA      :10761      Other      :21317      Employed      :67310
## NY      : 5844      Professional :10542      Full-time     : 7927
## TX      : 5634      Executive   : 3468      Not employed  : 649
## FL      : 5406      Computer Programmer: 3236      Other         : 3806
## IL      : 4264      Teacher     : 2888      Part-time     : 256
## OH      : 3373      Analyst     : 2735      Retired       : 367
## (Other):49571      (Other)     :40667      Self-employed: 4538
## IsBorrowerHomeowner CurrentlyInGroup      GroupKey
## Mode :logical      Mode :logical      :82827
## FALSE:40005      FALSE:82607      3D4D3366260257624AB272D: 310
## TRUE :44848      TRUE :2246      783C3371218786870A73D20: 208
##      :      :      52EA3425051368132B80C96: 150
##      :      :      B0473364376920128370B13: 83
##      :      :      FEF83377364176536637E50: 82
##      :      :      (Other)      : 1193
##      DateCreditPulled CreditScoreRangeLower CreditScoreRangeUpper
## 11/4/13 14:12: 8      Min.    :600.0      Min.    :619.0
## 12/23/13 9:38: 6      1st Qu.:660.0      1st Qu.:679.0
## 9/25/13 15:33: 6      Median :700.0      Median :719.0
## 11/1/13 13:02: 5      Mean    :699.4      Mean    :718.4

```

```

## 11/4/13 15:10:      5      3rd Qu.:720.0          3rd Qu.:739.0
## 12/9/13 10:12:      5      Max.      :880.0          Max.      :899.0
## (Other)           :84818
## FirstRecordedCreditLine CurrentCreditLines OpenCreditLines
## 12/1/93 0:00:  148      Min.      : 0.00      Min.      : 0.000
## 11/1/94 0:00:  139      1st Qu.: 7.00      1st Qu.: 6.000
## 11/1/95 0:00:  125      Median :10.00      Median : 9.000
## 3/1/96 0:00 :  123      Mean      :10.51      Mean      : 9.529
## 3/1/95 0:00 :  121      3rd Qu.:13.00      3rd Qu.:12.000
## 11/1/89 0:00:  118      Max.      :59.00      Max.      :54.000
## (Other)           :84079
## TotalCreditLinespast7years OpenRevolvingAccounts
## Min.      : 2.00      Min.      : 0.000
## 1st Qu.: 18.00      1st Qu.: 4.000
## Median : 26.00      Median : 7.000
## Mean      : 27.66      Mean      : 7.386
## 3rd Qu.: 35.00      3rd Qu.:10.000
## Max.      :125.00      Max.      :50.000
##
## OpenRevolvingMonthlyPayment InquiriesLast6Months TotalInquiries
## Min.      : 0.0      Min.      : 0.0000      Min.      : 0.00
## 1st Qu.: 156.0      1st Qu.: 0.0000      1st Qu.: 2.00
## Median : 311.0      Median : 0.0000      Median : 3.00
## Mean      : 430.7      Mean      : 0.9646      Mean      : 4.29
## 3rd Qu.: 564.0      3rd Qu.: 1.0000      3rd Qu.: 6.00
## Max.      :13765.0      Max.      :27.0000      Max.      :78.00
##
## 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      : 0.3225      Mean      : 950.8      Mean      : 3.659
## 3rd Qu.: 0.0000      3rd Qu.: 0.0      3rd Qu.: 2.000
## Max.      :51.0000      Max.      :463881.0      Max.      :99.000
##
## PublicRecordsLast10Years PublicRecordsLast12Months RevolvingCreditBalance
## Min.      : 0.0000      Min.      : 0.000000      Min.      : 0
## 1st Qu.: 0.0000      1st Qu.: 0.000000      1st Qu.: 3824
## Median : 0.0000      Median : 0.000000      Median : 9323
## Mean      : 0.2851      Mean      : 0.009169      Mean      : 17940
## 3rd Qu.: 0.0000      3rd Qu.: 0.000000      3rd Qu.: 20339
## Max.      :38.0000      Max.      :20.000000      Max.      :999165
##
## BankcardUtilization AvailableBankcardCredit TotalTrades
## Min.      :0.0000      Min.      : 0      Min.      : 1.00
## 1st Qu.:0.3300      1st Qu.: 1148      1st Qu.: 16.00
## Median :0.6000      Median : 4574      Median : 23.00
## Mean      :0.5642      Mean      : 11401      Mean      : 23.93
## 3rd Qu.:0.8300      3rd Qu.: 13910      3rd Qu.: 31.00
## Max.      :2.5000      Max.      :498374      Max.      :122.00

```

```

##
## TradesNeverDelinquent..percentage. TradesOpenedLast6Months
## Min. :0.0800 Min. : 0.0000
## 1st Qu.:0.8500 1st Qu.: 0.0000
## Median :0.9500 Median : 0.0000
## Mean :0.9059 Mean : 0.7298
## 3rd Qu.:1.0000 3rd Qu.: 1.0000
## Max. :1.0000 Max. :20.0000
##
## IncomeRange IncomeVerifiable StatedMonthlyIncome
## $0 : 45 Mode :logical Min. : 0
## $1-24,999 : 4654 FALSE:7333 1st Qu.: 3434
## $100,000+ :15205 TRUE :77520 Median : 5000
## $25,000-49,999:24175 Mean : 5931
## $50,000-74,999:25627 3rd Qu.: 7083
## $75,000-99,999:14498 Max. :1750003
## Not employed : 649
## LoanKey LoanCurrentDaysDelinquent
## CB1B37030986463208432A1: 6 Min. : 0.00
## 2DEE3698211017519D7333F: 4 1st Qu.: 0.00
## 9F4B37043517554537C364C: 4 Median : 0.00
## D895370150591392337ED6D: 4 Mean : 36.63
## E6FB37073953690388BC56D: 4 3rd Qu.: 0.00
## 0D8F37036734373301ED419: 3 Max. :1593.00
## (Other) :84828
## LoanMonthsSinceOrigination LoanNumber LoanOriginalAmount
## Min. : 0.00 Min. : 38045 Min. : 1000
## 1st Qu.: 4.00 1st Qu.: 60950 1st Qu.: 4000
## Median :11.00 Median : 87789 Median : 7500
## Mean :16.02 Mean : 86592 Mean : 9083
## 3rd Qu.:25.00 3rd Qu.:109096 3rd Qu.:13500
## Max. :56.00 Max. :136486 Max. :35000
##
## LoanOriginationDate LoanOriginationQuarter
## 1/22/14 0:00 : 491 Q4 2013:14450
## 11/13/13 0:00: 490 Q1 2014:12172
## 2/19/14 0:00 : 439 Q3 2013: 9180
## 10/16/13 0:00: 434 Q2 2013: 7099
## 1/28/14 0:00 : 339 Q3 2012: 5632
## 9/24/13 0:00 : 316 Q2 2012: 5061
## (Other) :82344 (Other):31259
## MemberKey MonthlyLoanPayment LP_CustomerPayments
## 720D3508651090808DC328F: 7 Min. : 0.0 Min. : -2.35
## C70934206057523078260C7: 7 1st Qu.: 157.3 1st Qu.: 799.39
## E4AF3422677498955FFA00E: 7 Median : 251.9 Median : 2214.76
## 18F6337949289842881D0A8: 6 Mean : 291.9 Mean : 3669.06
## 3D6B34225353312993B9700: 6 3rd Qu.: 388.4 3rd Qu.: 4869.99
## 43DB3366978035224D7D9E3: 6 Max. :2251.5 Max. :37369.16
## (Other) :84814
## LP_CustomerPrincipalPayments LP_InterestandFees LP_ServiceFees

```

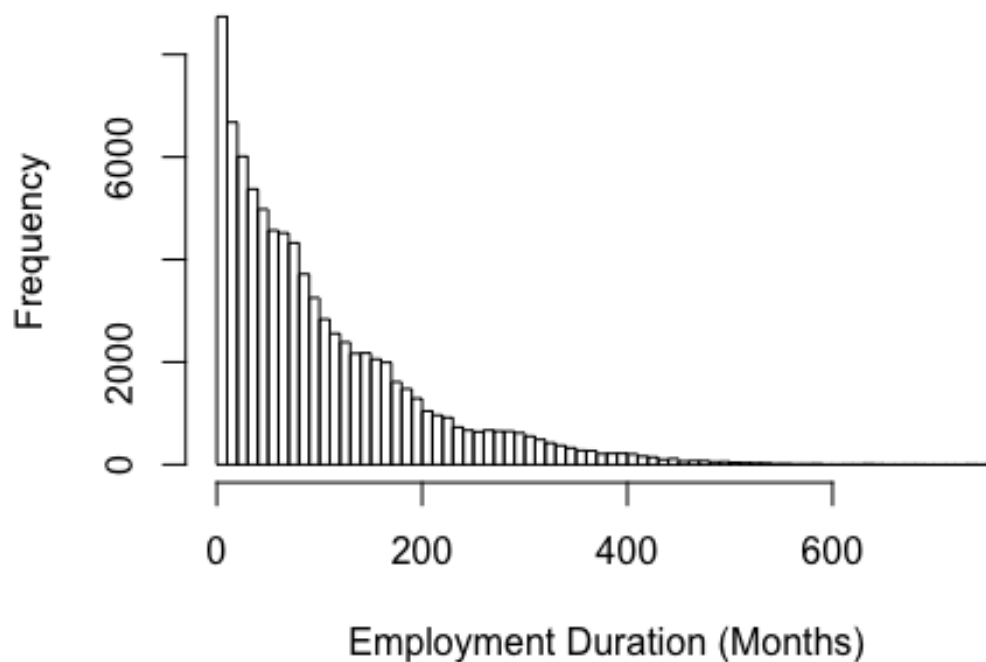
```

## Min. : 0.0 Min. : -2.35 Min. : -589.95
## 1st Qu.: 393.8 1st Qu.: 254.06 1st Qu.: -72.30
## Median : 1250.0 Median : 674.39 Median : -34.90
## Mean : 2626.9 Mean : 1042.20 Mean : -54.64
## 3rd Qu.: 3414.3 3rd Qu.: 1436.06 3rd Qu.: -14.28
## Max. : 35000.0 Max. : 10572.78 Max. : 3.01
##
## LP_CollectionFees LP_GrossPrincipalLoss LP_NetPrincipalLoss
## Min. : -4865.080 Min. : -94.2 Min. : -504.4
## 1st Qu.: 0.000 1st Qu.: 0.0 1st Qu.: 0.0
## Median : 0.000 Median : 0.0 Median : 0.0
## Mean : -8.206 Mean : 376.1 Mean : 367.7
## 3rd Qu.: 0.000 3rd Qu.: 0.0 3rd Qu.: 0.0
## Max. : 0.000 Max. : 25000.0 Max. : 25000.0
##
## LP_NonPrincipalRecoverypayments PercentFunded Recommendations
## Min. : 0.000 Min. : 0.7000 Min. : 0.00000
## 1st Qu.: 0.000 1st Qu.: 1.0000 1st Qu.: 0.00000
## Median : 0.000 Median : 1.0000 Median : 0.00000
## Mean : 7.647 Mean : 0.9981 Mean : 0.01756
## 3rd Qu.: 0.000 3rd Qu.: 1.0000 3rd Qu.: 0.00000
## Max. : 7780.030 Max. : 1.0125 Max. : 19.00000
##
## InvestmentFromFriendsCount InvestmentFromFriendsAmount Investors
## Min. : 0.000000 Min. : 0.000 Min. : 1.00
## 1st Qu.: 0.000000 1st Qu.: 0.000 1st Qu.: 1.00
## Median : 0.000000 Median : 0.000 Median : 32.00
## Mean : 0.008049 Mean : 4.316 Mean : 68.26
## 3rd Qu.: 0.000000 3rd Qu.: 0.000 3rd Qu.: 97.00
## Max. : 9.000000 Max. : 11000.000 Max. : 1189.00
##
## DebtToIncomeRatio.imp EmploymentStatusDuration.imp
## Min. : 0.0000 Min. : 0
## 1st Qu.: 0.1600 1st Qu.: 30
## Median : 0.2400 Median : 74
## Mean : 0.2588 Mean : 103
## 3rd Qu.: 0.3100 3rd Qu.: 148
## Max. : 10.0100 Max. : 755
##
## Removed or imputed all "NA"s to this point.
## 84,853 Observations are left. Ready for EDA

#### EXPLORATORY DATA ANALYSIS ####
hist(LD2.red$EmploymentStatusDuration.imp, xlab="Employment Duration (Months)",
     ,main = "Employment Duration Hist.", 100)

```

Employment Duration Hist.



```
summary(LD2.red$EmploymentStatus)
```

```
##      Employed      Full-time  Not employed      Other      Part-time
##      67310       7927         649         3806         256
##      Retired Self-employed
##      367         4538
```

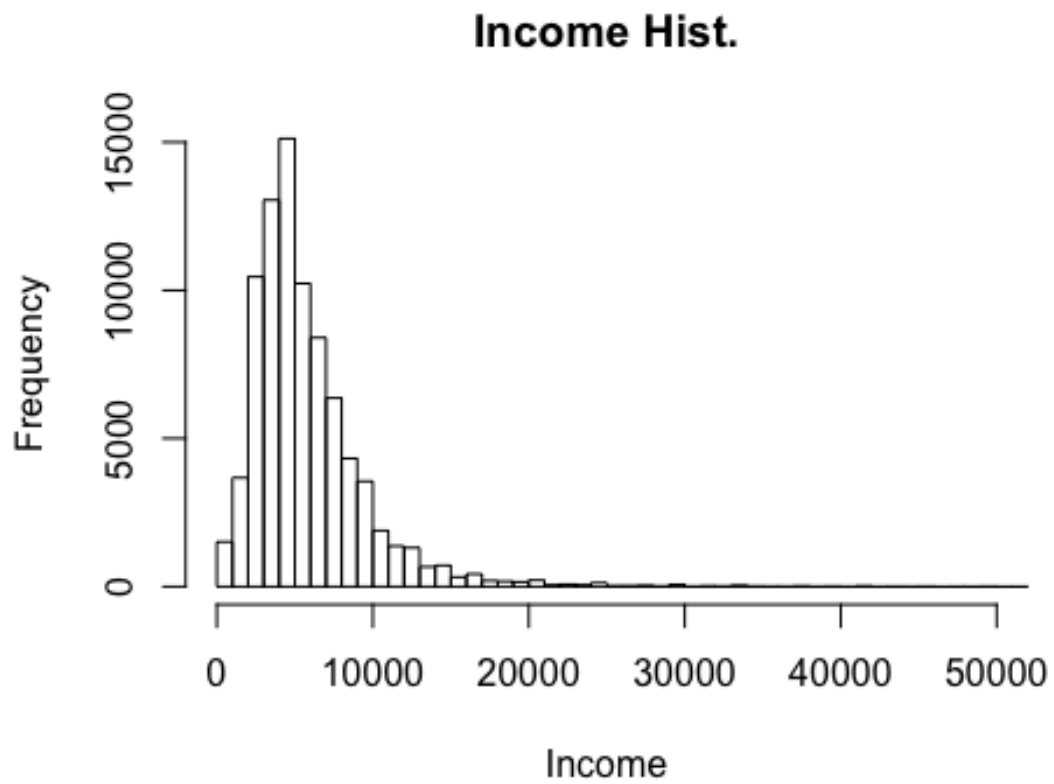
```
### DUMMY VARIABLES
```

```
## Employment Status: Employed, Full-time, Not employed, Other, Part-time, Re
tired, Self-employed
```

```
## To simplify making "Part-Time" a 0 (only 256 entries of this)
```

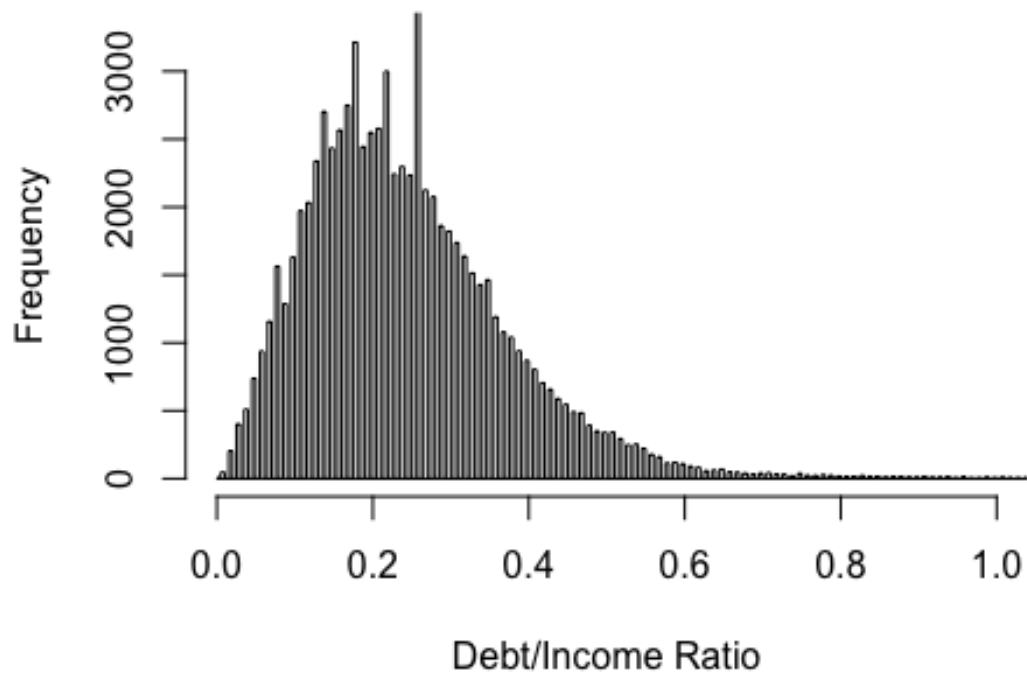
```
LD2.red$EmploymentStatus.dum <- ifelse(LD2.red$EmploymentStatus=="Employed",
1,
  ifelse(LD2.red$EmploymentStatus=="Full-time", 1,
    ifelse(LD2.red$EmploymentStatus=="Self-employed", 1, 0)
  )
)
```

```
hist(LD2.red$StatedMonthlyIncome, xlab = "Income", main = "Income Hist.", 200
0, xlim=c(0,50000))
```

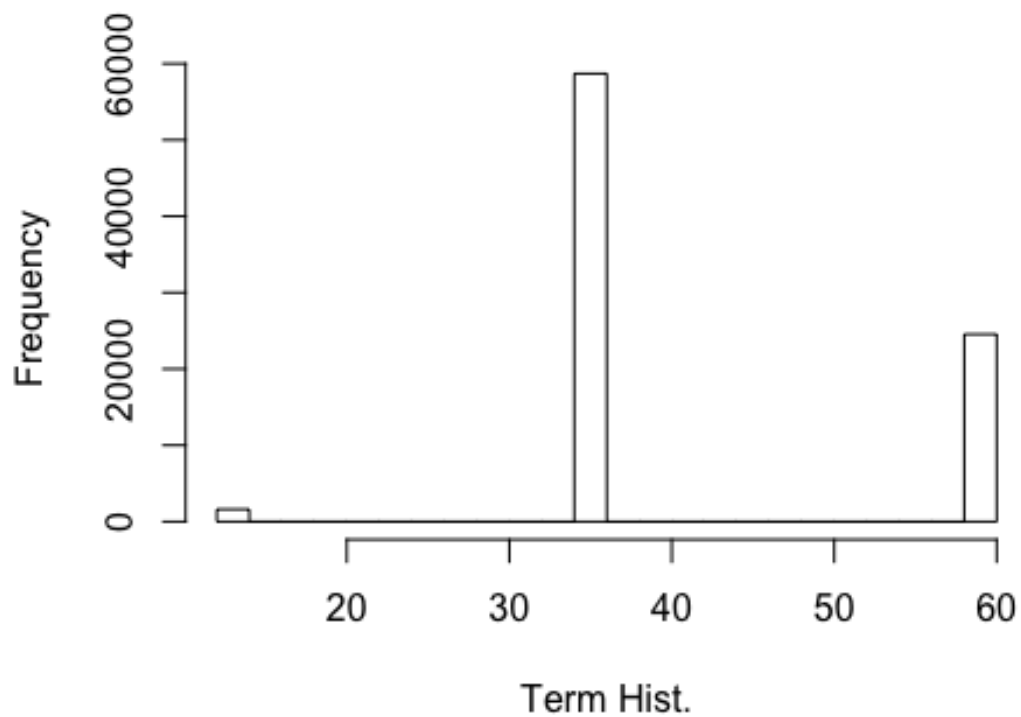


```
hist(LD2.red$DebtToIncomeRatio.imp, xlab="Debt/Income Ratio", main = "D/I Hist.", 2000, xlim=c(0,1), ylim=c(0,3300))
```

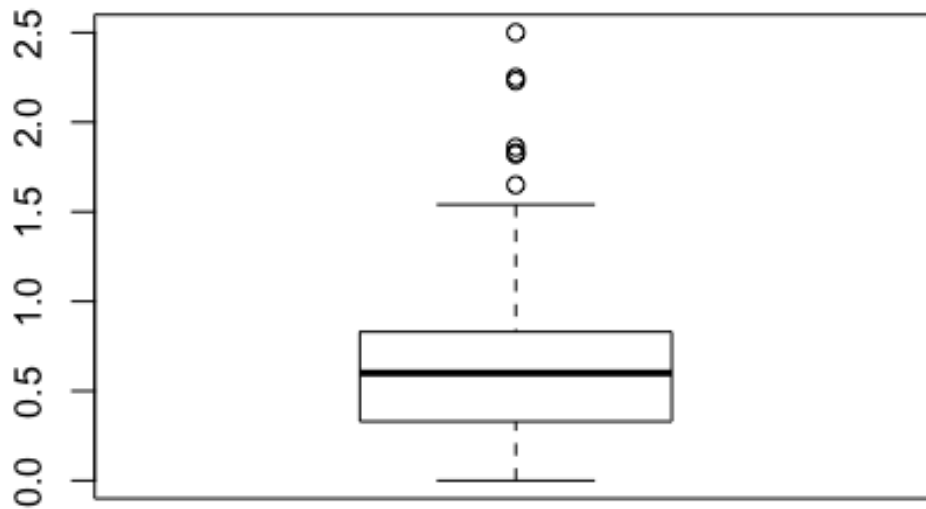
D/I Hist.



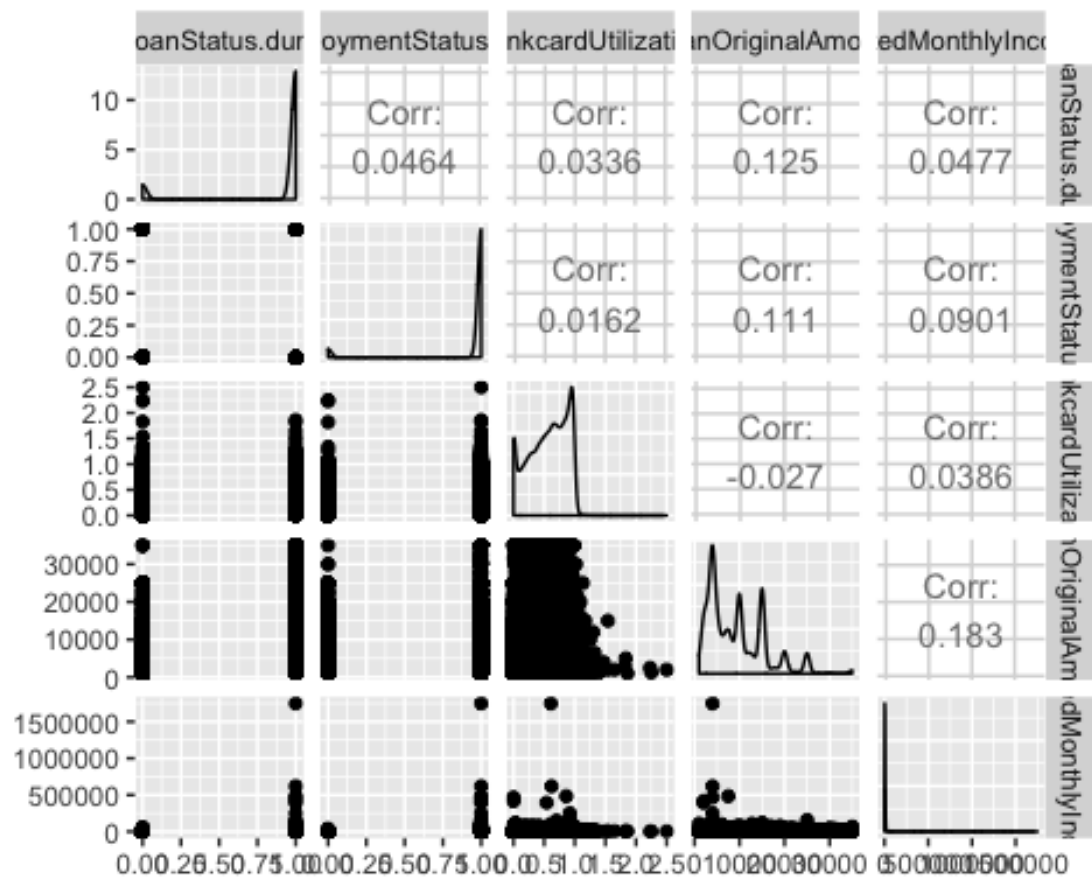
```
## Peak because of Imputed into 7000 data points.  
hist(LD2.red$Term, xlab="Term Hist.", main="")
```

```
## BOXPLOT Bank Card Utilization
boxplot(LD2.red$BankcardUtilization)
## Loan Status DUMMIES
LD2.red$LoanStatus.dum <- ifelse(LD2.red$LoanStatus=="Completed", 1,
  ifelse(LD2.red$LoanStatus=="Current", 1,
    ifelse(LD2.red$LoanStatus=="FinalPaymentInProgress", 1, 0)
  )
)
library(GGally)
```



```
graphs1 = subset(LD2.red, select=c("LoanStatus.dum", "EmploymentStatus.dum",  
"BankcardUtilization", "LoanOriginalAmount", "StatedMonthlyIncome"))  
ggpairs(graphs1, diag=list(continuous="density", discrete="bar"), axisLabels=  
"show")  
  
## Warning in check_and_set_ggpairs_defaults("diag", diag, continuous =  
## "densityDiag", : Changing diag$continuous from 'density' to 'densityDiag'  
  
## Warning in check_and_set_ggpairs_defaults("diag", diag, continuous =  
## "densityDiag", : Changing diag$discrete from 'bar' to 'barDiag'
```



```
cor(LD2.red$BorrowerRate, LD2.red$EstimatedEffectiveYield)
## [1] 0.8952825

## 89%, Rid of EstimatedEffectiveYield
cor(LD2.red$ProsperRating.numeric., LD2.red$ProsperScore)
## [1] 0.7052214

# 70% Correlated.
cor(LD2.red$CreditScoreRangeLower, LD2.red$CreditScoreRangeUpper)
## [1] 1

# 100% Correlated, Removed CreditScoreRangeUpper
cor(LD2.red$CurrentCreditLines, LD2.red$OpenCreditLines)
## [1] 0.9633239

# 96% Remove OpenCreditLines
cor(LD2.red$BankcardUtilization, LD2.red$DebtToIncomeRatio.imp)
## [1] 0.05473621
```

```

# 5%
LD.write <- LD2.red[1:100, ]
#write.csv(LD.write, file = "pLD_Check2.csv", row.names=FALSE)

##Reducing Dimensions:
## Removed all "Date" Variables
## Removed Highly Correlated Variables (As stated above)
## Removed admin variables (group key, member key)
LD3 <- LD2.red[,c(-1,-4,-19,-20,-22,-23,-25,-42,-44,-46,-51,-52,-53)]
LD3 <- LD3[,c(-16)]
#### Create Dummy for Homeowner
LD3$IsBorrowerHomeowner.dum <- ifelse(LD3$IsBorrowerHomeowner=="TRUE", 1, 0)

### TEST/TRAIN SETS
set.seed(123) #ensure we always get the same output
LD3.index <- LD3[order(runif(84853)), ]#randomized the observations
## 80/20 Split
train <- LD3.index[1:67882, ] #create training set
valid <- LD3.index[67883:84853, ] #create validation set

#### NUMERICAL PREDICTION FOR "PROSPER SCORE" ####
train.num <- train[,c(-2,-11,-12,-13,-14,-15,-34)]

train.num <- train.num[,c(-28,-32)]
### Got rid of categorical variables, some now replaced by dummies
str(train.num)

## 'data.frame':    67882 obs. of  48 variables:
## $ Term : int  36 60 60 36 36 36 60 36 36 60
...
## $ BorrowerAPR : num  0.288 0.265 0.204 0.232 0.242
...
## $ BorrowerRate : num  0.249 0.239 0.184 0.201 0.204
...
## $ LenderYield : num  0.239 0.229 0.174 0.191 0.195
...
## $ EstimatedEffectiveYield : num  0.219 0.21 0.172 0.122 0.18 ..
.
## $ EstimatedLoss : num  0.1125 0.1175 0.0595 0.065 0.0
899 ...
## $ EstimatedReturn : num  0.1065 0.0921 0.112 0.1216 0.0
898 ...
## $ ProsperRating..numeric. : int  3 3 5 4 4 5 4 4 7 2 ...
## $ ProsperScore : int  3 4 8 9 4 8 6 2 10 3 ...
## $ CreditScoreRangeLower : int  640 660 700 640 740 680 740 74
0 800 660 ...
## $ CurrentCreditLines : int  7 8 7 2 8 9 22 14 12 9 ...
## $ TotalCreditLinespast7years : int  29 36 8 34 14 14 44 27 20 22 .
..
## $ OpenRevolvingAccounts : int  2 5 2 4 5 5 18 4 9 7 ...

```

```

## $ OpenRevolvingMonthlyPayment      : int  60 213 127 40 442 285 866 84 1
55 192 ...
## $ InquiriesLast6Months              : int  0 0 0 0 1 1 0 5 1 6 ...
## $ TotalInquiries                    : int  6 3 0 4 4 2 0 8 2 11 ...
## $ CurrentDelinquencies              : int  0 1 0 0 0 0 0 0 0 0 ...
## $ AmountDelinquent                  : int  0 0 0 0 0 0 0 0 0 0 ...
## $ DelinquenciesLast7Years           : int  0 7 0 4 0 0 0 0 0 5 ...
## $ PublicRecordsLast10Years          : int  1 1 0 1 0 0 0 0 0 2 ...
## $ PublicRecordsLast12Months         : int  0 0 0 0 0 0 0 0 0 0 ...
## $ RevolvingCreditBalance            : int  961 4668 4982 901 17127 5916 2
7182 1008 316 6087 ...
## $ BankcardUtilization                : num  0.86 0.46 0.73 0.85 0.59 0.56
0.45 0.08 0 0.64 ...
## $ AvailableBankcardCredit           : int  108 3202 1443 149 11281 4586 2
5568 10971 43434 2280 ...
## $ TotalTrades                       : int  19 32 7 31 14 14 37 26 19 21 .
..
## $ TradesNeverDelinquent..percentage.: num  1 0.75 1 0.7 1 0.71 0.97 1 1 0
.9 ...
## $ TradesOpenedLast6Months           : int  0 0 0 1 1 1 6 5 2 3 ...
## $ LoanMonthsSinceOrigination         : int  8 2 25 51 4 29 26 27 5 5 ...
## $ LoanNumber                        : int  95175 128064 60408 39758 10973
4 54438 59506 58116 105814 109096 ...
## $ LoanOriginalAmount                : int  7000 4000 15000 1200 3000 4000
15000 5000 15000 10000 ...
## $ LP_CustomerPayments               : num  1941 115 9230 1337 335 ...
## $ LP_CustomerPrincipalPayments      : num  926.7 33.6 4425.2 1200 185.1 .
..
## $ LP_InterestandFees                : num  1014.7 81.3 4804.7 136.6 149.7
...
## $ LP_ServiceFees                    : num  -40.92 -3.4 -260.42 -6.8 -7.41
...
## $ 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  1 2 1 50 1 37 79 38 276 185 ..
.
## $ DebtToIncomeRatio.imp              : num  0.36 0.46 0.1 0.02 0.43 ...
## $ EmploymentStatusDuration.imp       : num  52 111 2 17 141 16 84 113 281
184 ...
## $ EmploymentStatus.dum               : num  1 1 1 1 1 1 0 1 1 1 ...
## $ LoanStatus.dum                     : num  1 1 1 1 1 1 1 1 1 1 ...
## $ IsBorrowerHomeowner.dum           : num  0 0 0 0 0 0 1 0 1 1 ...

```

`summary(train.num)`

##	Term	BorrowerAPR	BorrowerRate	LenderYield
##	Min. :12.00	Min. :0.04583	Min. :0.0400	Min. :0.0300
##	1st Qu.:36.00	1st Qu.:0.16324	1st Qu.:0.1359	1st Qu.:0.1259
##	Median :36.00	Median :0.21945	Median :0.1875	Median :0.1775
##	Mean :42.46	Mean :0.22668	Mean :0.1960	Mean :0.1860
##	3rd Qu.:60.00	3rd Qu.:0.29254	3rd Qu.:0.2574	3rd Qu.:0.2474
##	Max. :60.00	Max. :0.42395	Max. :0.3600	Max. :0.3400
##	EstimatedEffectiveYield	EstimatedLoss	EstimatedReturn	
##	Min. :-0.1827	Min. :0.00490	Min. :-0.18270	
##	1st Qu.: 0.1157	1st Qu.:0.04240	1st Qu.: 0.07408	
##	Median : 0.1615	Median :0.07240	Median : 0.09167	
##	Mean : 0.1687	Mean :0.08031	Mean : 0.09607	
##	3rd Qu.: 0.2243	3rd Qu.:0.11200	3rd Qu.: 0.11660	
##	Max. : 0.3199	Max. :0.36600	Max. : 0.26670	
##	ProsperRating..numeric.	ProsperScore	CreditScoreRangeLower	
##	Min. :1.000	Min. : 1.000	Min. :600.0	
##	1st Qu.:3.000	1st Qu.: 4.000	1st Qu.:660.0	
##	Median :4.000	Median : 6.000	Median :700.0	
##	Mean :4.073	Mean : 5.956	Mean :699.4	
##	3rd Qu.:5.000	3rd Qu.: 8.000	3rd Qu.:720.0	
##	Max. :7.000	Max. :11.000	Max. :880.0	
##	CurrentCreditLines	TotalCreditLinespast7years	OpenRevolvingAccounts	
##	Min. : 0.00	Min. : 2.00	Min. : 0.000	
##	1st Qu.: 7.00	1st Qu.: 18.00	1st Qu.: 4.000	
##	Median :10.00	Median : 26.00	Median : 7.000	
##	Mean :10.51	Mean : 27.63	Mean : 7.388	
##	3rd Qu.:13.00	3rd Qu.: 35.00	3rd Qu.:10.000	
##	Max. :56.00	Max. :125.00	Max. :50.000	
##	OpenRevolvingMonthlyPayment	InquiriesLast6Months	TotalInquiries	
##	Min. : 0.0	Min. : 0.0000	Min. : 0.00	
##	1st Qu.: 155.0	1st Qu.: 0.0000	1st Qu.: 2.00	
##	Median : 311.0	Median : 0.0000	Median : 3.00	
##	Mean : 430.8	Mean : 0.9652	Mean : 4.29	
##	3rd Qu.: 564.0	3rd Qu.: 1.0000	3rd Qu.: 6.00	
##	Max. :13765.0	Max. :22.0000	Max. :78.00	
##	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 : 0.3208	Mean : 950.6	Mean : 3.669	
##	3rd Qu.: 0.0000	3rd Qu.: 0.0	3rd Qu.: 2.000	
##	Max. :51.0000	Max. :327677.0	Max. :99.000	
##	PublicRecordsLast10Years	PublicRecordsLast12Months	RevolvingCreditBalance	
##	Min. : 0.0000	Min. : 0.000000	Min. : 0	
##	1st Qu.: 0.0000	1st Qu.: 0.000000	1st Qu.: 3825	
##	Median : 0.0000	Median : 0.000000	Median : 9308	
##	Mean : 0.2878	Mean : 0.009325	Mean : 17932	
##	3rd Qu.: 0.0000	3rd Qu.: 0.000000	3rd Qu.: 20310	
##	Max. :38.0000	Max. :20.000000	Max. :999165	
##	BankcardUtilization	AvailableBankcardCredit	TotalTrades	

```

## Min. :0.0000 Min. : 0 Min. : 1.00
## 1st Qu.:0.3300 1st Qu.: 1154 1st Qu.: 15.00
## Median :0.6000 Median : 4597 Median : 22.00
## Mean :0.5634 Mean : 11390 Mean : 23.91
## 3rd Qu.:0.8300 3rd Qu.: 13934 3rd Qu.: 31.00
## Max. :2.5000 Max. :498374 Max. :122.00
## TradesNeverDelinquent..percentage. TradesOpenedLast6Months
## Min. :0.0800 Min. : 0.0000
## 1st Qu.:0.8500 1st Qu.: 0.0000
## Median :0.9500 Median : 0.0000
## Mean :0.9059 Mean : 0.7298
## 3rd Qu.:1.0000 3rd Qu.: 1.0000
## Max. :1.0000 Max. :20.0000
## LoanMonthsSinceOrigination LoanNumber LoanOriginalAmount
## Min. : 0.00 Min. : 38045 Min. : 1000
## 1st Qu.: 4.00 1st Qu.: 60923 1st Qu.: 4000
## Median :11.00 Median : 87758 Median : 7500
## Mean :16.03 Mean : 86599 Mean : 9091
## 3rd Qu.:25.00 3rd Qu.:109199 3rd Qu.:13500
## Max. :56.00 Max. :136486 Max. :35000
## LP_CustomerPayments LP_CustomerPrincipalPayments LP_InterestandFees
## Min. : -2.35 Min. : 0.0 Min. : -2.35
## 1st Qu.: 797.52 1st Qu.: 392.3 1st Qu.: 252.42
## Median : 2219.91 Median : 1253.9 Median : 674.25
## Mean : 3677.10 Mean : 2636.9 Mean : 1040.21
## 3rd Qu.: 4882.64 3rd Qu.: 3428.9 3rd Qu.: 1431.98
## Max. :37369.16 Max. :35000.0 Max. :10572.78
## LP_ServiceFees LP_CollectionFees LP_GrossPrincipalLoss
## Min. : -589.95 Min. : -4865.08 Min. : -94.2
## 1st Qu.: -72.12 1st Qu.: 0.00 1st Qu.: 0.0
## Median : -34.76 Median : 0.00 Median : 0.0
## Mean : -54.51 Mean : -7.99 Mean : 372.5
## 3rd Qu.: -14.21 3rd Qu.: 0.00 3rd Qu.: 0.0
## Max. : 3.01 Max. : 0.00 Max. :25000.0
## LP_NetPrincipalLoss LP_NonPrincipalRecoverypayments PercentFunded
## Min. : -504.4 Min. : 0.00 Min. :0.7000
## 1st Qu.: 0.0 1st Qu.: 0.00 1st Qu.:1.0000
## Median : 0.0 Median : 0.00 Median :1.0000
## Mean : 364.4 Mean : 7.45 Mean :0.9981
## 3rd Qu.: 0.0 3rd Qu.: 0.00 3rd Qu.:1.0000
## Max. :25000.0 Max. :6296.96 Max. :1.0125
## Recommendations InvestmentFromFriendsCount InvestmentFromFriendsAmount
## Min. : 0.00000 Min. :0.000000 Min. : 0.000
## 1st Qu.: 0.00000 1st Qu.:0.000000 1st Qu.: 0.000
## Median : 0.00000 Median :0.000000 Median : 0.000
## Mean : 0.01728 Mean :0.008102 Mean : 4.202
## 3rd Qu.: 0.00000 3rd Qu.:0.000000 3rd Qu.: 0.000
## Max. :19.00000 Max. :9.000000 Max. :11000.000
## Investors DebtToIncomeRatio.imp EmploymentStatusDuration.imp
## Min. : 1.00 Min. : 0.0000 Min. : 0.0

```

```
## 1st Qu.: 1.00 1st Qu.: 0.1600 1st Qu.: 30.0
## Median : 32.00 Median : 0.2400 Median : 74.0
## Mean : 68.14 Mean : 0.2598 Mean : 102.9
## 3rd Qu.: 97.00 3rd Qu.: 0.3100 3rd Qu.: 147.0
## Max. : 1189.00 Max. : 10.0100 Max. : 755.0
## EmploymentStatus.dum LoanStatus.dum IsBorrowerHomeowner.dum
## Min. :0.0000 Min. :0.000 Min. :0.0000
## 1st Qu.:1.0000 1st Qu.:1.000 1st Qu.:0.0000
## Median :1.0000 Median :1.000 Median :1.0000
## Mean :0.9396 Mean :0.901 Mean :0.5273
## 3rd Qu.:1.0000 3rd Qu.:1.000 3rd Qu.:1.0000
## Max. :1.0000 Max. :1.000 Max. :1.0000
```

```
options(scipen=999)
```

```
### Linear Regression
```

```
train.lm <- lm(ProsperScore~., data=train.num)
```

```
library(psych)
```

```
summary(train.lm)
```

```
##
```

```
## Call:
```

```
## lm(formula = ProsperScore ~ ., data = train.num)
```

```
##
```

```
## Residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -6.0514 -0.9004  0.0083  0.8698  7.8796
```

```
##
```

```
## Coefficients: (1 not defined because of singularities)
```

	Estimate	Std. Error	t value
## (Intercept)	14.1681360437	0.4311692402	32.860
## Term	0.0012867989	0.0007189228	1.790
## BorrowerAPR	-26.0295817838	0.9001376779	-28.917
## BorrowerRate	19.2777874159	25.1581381522	0.766
## LenderYield	-93.1668655065	25.2505422090	-3.690
## EstimatedEffectiveYield	-0.9660775531	0.2471348203	-3.909
## EstimatedLoss	99.9150899235	2.0700664528	48.267
## EstimatedReturn	98.7616107740	1.9822420247	49.823
## ProsperRating..numeric.	0.4756579445	0.0168823716	28.175
## CreditScoreRangeLower	-0.0072210943	0.0001797389	-40.175
## CurrentCreditLines	-0.0124803933	0.0022387854	-5.575
## TotalCreditLinespast7years	-0.0015306850	0.0011923829	-1.284
## OpenRevolvingAccounts	0.0102629296	0.0024669803	4.160
## OpenRevolvingMonthlyPayment	-0.0000038511	0.0000235301	-0.164
## InquiriesLast6Months	-0.1814605978	0.0050420780	-35.989
## TotalInquiries	-0.0221831890	0.0017995132	-12.327
## CurrentDelinquencies	-0.0569806777	0.0055717684	-10.227
## AmountDelinquent	-0.0000015533	0.0000007932	-1.958
## DelinquenciesLast7Years	0.0028290541	0.0007221588	3.917
## PublicRecordsLast10Years	0.0527379955	0.0091278042	5.778
## PublicRecordsLast12Months	-0.0786294616	0.0425645582	-1.847

## RevolvingCreditBalance	0.0000004034	0.0000002662	1.515
## BankcardUtilization	-0.4280507822	0.0245854464	-17.411
## AvailableBankcardCredit	0.0000126907	0.0000003905	32.495
## TotalTrades	0.0002056766	0.0014378563	0.143
## TradesNeverDelinquent..percentage.	-0.2651071892	0.0604735667	-4.384
## TradesOpenedLast6Months	-0.1680466252	0.0060756467	-27.659
## LoanMonthsSinceOrigination	0.0167455901	0.0016266675	10.294
## LoanNumber	-0.0000166246	0.0000007824	-21.249
## LoanOriginalAmount	0.0000526293	0.0000014373	36.616
## LP_CustomerPayments	-0.0001361028	0.0000146221	-9.308
## LP_CustomerPrincipalPayments	0.0001007874	0.0000145888	6.909
## LP_InterestandFees	NA	NA	NA
## LP_ServiceFees	0.0021192651	0.0002966358	7.144
## LP_CollectionFees	0.0000366597	0.0001157428	0.317
## LP_GrossPrincipalLoss	-0.0000362794	0.0000365878	-0.992
## LP_NetPrincipalLoss	-0.0000278054	0.0000366558	-0.759
## LP_NonPrincipalRecoverypayments	-0.0001137466	0.0000634366	-1.793
## PercentFunded	-1.1869585998	0.2600222529	-4.565
## Recommendations	-0.0225243384	0.0340554552	-0.661
## InvestmentFromFriendsCount	-0.1424564153	0.0602370410	-2.365
## InvestmentFromFriendsAmount	-0.0000869777	0.0000495312	-1.756
## Investors	0.0006355812	0.0000748885	8.487
## DebtToIncomeRatio.imp	-0.2536257208	0.0174658817	-14.521
## EmploymentStatusDuration.imp	-0.0000867362	0.0000576969	-1.503
## EmploymentStatus.dum	0.0612082522	0.0229123795	2.671
## LoanStatus.dum	-0.0164432309	0.0255476376	-0.644
## IsBorrowerHomeowner.dum	0.0082660454	0.0124813708	0.662
##	Pr(> t)		
## (Intercept)	< 0.00000000000000002	***	
## Term	0.073475	.	
## BorrowerAPR	< 0.00000000000000002	***	
## BorrowerRate	0.443522		
## LenderYield	0.000225	***	
## EstimatedEffectiveYield	0.000092726191538	***	
## EstimatedLoss	< 0.00000000000000002	***	
## EstimatedReturn	< 0.00000000000000002	***	
## ProsperRating..numeric.	< 0.00000000000000002	***	
## CreditScoreRangeLower	< 0.00000000000000002	***	
## CurrentCreditLines	0.000000024900150	***	
## TotalCreditLinespast7years	0.199245		
## OpenRevolvingAccounts	0.000031847413990	***	
## OpenRevolvingMonthlyPayment	0.869993		
## InquiriesLast6Months	< 0.00000000000000002	***	
## TotalInquiries	< 0.00000000000000002	***	
## CurrentDelinquencies	< 0.00000000000000002	***	
## AmountDelinquent	0.050192	.	
## DelinquenciesLast7Years	0.000089561117253	***	
## PublicRecordsLast10Years	0.000000007604478	***	
## PublicRecordsLast12Months	0.064708	.	
## RevolvingCreditBalance	0.129757		

```

## BankcardUtilization          < 0.0000000000000002 ***
## AvailableBankcardCredit      < 0.0000000000000002 ***
## TotalTrades                  0.886256
## TradesNeverDelinquent..percentage. 0.000011677382533 ***
## TradesOpenedLast6Months      < 0.0000000000000002 ***
## LoanMonthsSinceOrigination    < 0.0000000000000002 ***
## LoanNumber                   < 0.0000000000000002 ***
## LoanOriginalAmount           < 0.0000000000000002 ***
## LP_CustomerPayments          < 0.0000000000000002 ***
## LP_CustomerPrincipalPayments 0.000000000004940 ***
## LP_InterestandFees           NA
## LP_ServiceFees               0.000000000000913 ***
## LP_CollectionFees            0.751446
## LP_GrossPrincipalLoss        0.321411
## LP_NetPrincipalLoss          0.448121
## LP_NonPrincipalRecoverypayments 0.072966 .
## PercentFunded                0.000005007640864 ***
## Recommendations              0.508357
## InvestmentFromFriendsCount    0.018036 *
## InvestmentFromFriendsAmount    0.079090 .
## Investors                    < 0.0000000000000002 ***
## DebtToIncomeRatio.imp        < 0.0000000000000002 ***
## EmploymentStatusDuration.imp  0.132764
## EmploymentStatus.dum         0.007555 **
## LoanStatus.dum               0.519817
## IsBorrowerHomeowner.dum      0.507800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.396 on 67835 degrees of freedom
## Multiple R-squared:  0.6547, Adjusted R-squared:  0.6544
## F-statistic: 2796 on 46 and 67835 DF, p-value: < 0.00000000000000022

#### Stepwise Regression
stepreg.train <- step(train.lm, direction="both")

## Start: AIC=45334.26
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##   EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##   ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##   TotalCreditLinespast7years + OpenRevolvingAccounts + OpenRevolvingMont
hlyPayment +
##   InquiriesLast6Months + TotalInquiries + CurrentDelinquencies +
##   AmountDelinquent + DelinquenciesLast7Years + PublicRecordsLast10Years
+
##   PublicRecordsLast12Months + RevolvingCreditBalance + BankcardUtilizati
on +
##   AvailableBankcardCredit + TotalTrades + TradesNeverDelinquent..percent
age. +
##   TradesOpenedLast6Months + LoanMonthsSinceOrigination + LoanNumber +

```

```

##      LoanOriginalAmount + LP_CustomerPayments + LP_CustomerPrincipalPayment
##      S +
##      LP_InterestandFees + LP_ServiceFees + LP_CollectionFees +
##      LP_GrossPrincipalLoss + LP_NetPrincipalLoss + LP_NonPrincipalRecoveryp
ayments +
##      PercentFunded + Recommendations + InvestmentFromFriendsCount +
##      InvestmentFromFriendsAmount + Investors + DebtToIncomeRatio.imp +
##      EmploymentStatusDuration.imp + EmploymentStatus.dum + LoanStatus.dum +
##      IsBorrowerHomeowner.dum
##
##
## Step:  AIC=45334.26
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##      EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##      ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##      TotalCreditLinespast7years + OpenRevolvingAccounts + OpenRevolvingMont
hlyPayment +
##      InquiriesLast6Months + TotalInquiries + CurrentDelinquencies +
##      AmountDelinquent + DelinquenciesLast7Years + PublicRecordsLast10Years
+
##      PublicRecordsLast12Months + RevolvingCreditBalance + BankcardUtilizati
on +
##      AvailableBankcardCredit + TotalTrades + TradesNeverDelinquent..percent
age. +
##      TradesOpenedLast6Months + LoanMonthsSinceOrigination + LoanNumber +
##      LoanOriginalAmount + LP_CustomerPayments + LP_CustomerPrincipalPayment
S +
##      LP_ServiceFees + LP_CollectionFees + LP_GrossPrincipalLoss +
##      LP_NetPrincipalLoss + LP_NonPrincipalRecoverypayments + PercentFunded
+
##      Recommendations + InvestmentFromFriendsCount + InvestmentFromFriendsAm
ount +
##      Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
##      EmploymentStatus.dum + LoanStatus.dum + IsBorrowerHomeowner.dum
##
##
##              Df Sum of Sq    RSS    AIC
## - TotalTrades          1      0.0 132188 45332
## - OpenRevolvingMonthlyPayment      1      0.1 132188 45332
## - LP_CollectionFees          1      0.2 132188 45332
## - LoanStatus.dum          1      0.8 132189 45333
## - Recommendations          1      0.9 132189 45333
## - IsBorrowerHomeowner.dum          1      0.9 132189 45333
## - LP_NetPrincipalLoss          1      1.1 132189 45333
## - BorrowerRate          1      1.1 132189 45333
## - LP_GrossPrincipalLoss          1      1.9 132190 45333
## - TotalCreditLinespast7years          1      3.2 132191 45334
## <none>                                132188 45334
## - EmploymentStatusDuration.imp          1      4.4 132192 45335
## - RevolvingCreditBalance          1      4.5 132193 45335
## - InvestmentFromFriendsAmount          1      6.0 132194 45335

```

```

## - Term 1 6.2 132194 45335
## - LP_NonPrincipalRecoverypayments 1 6.3 132194 45335
## - PublicRecordsLast12Months 1 6.6 132195 45336
## - AmountDelinquent 1 7.5 132196 45336
## - InvestmentFromFriendsCount 1 10.9 132199 45338
## - EmploymentStatus.dum 1 13.9 132202 45339
## - LenderYield 1 26.5 132215 45346
## - EstimatedEffectiveYield 1 29.8 132218 45348
## - DelinquenciesLast7Years 1 29.9 132218 45348
## - OpenRevolvingAccounts 1 33.7 132222 45350
## - TradesNeverDelinquent..percentage. 1 37.4 132225 45351
## - PercentFunded 1 40.6 132229 45353
## - CurrentCreditLines 1 60.6 132249 45363
## - PublicRecordsLast10Years 1 65.1 132253 45366
## - LP_CustomerPrincipalPayments 1 93.0 132281 45380
## - LP_ServiceFees 1 99.5 132287 45383
## - Investors 1 140.4 132328 45404
## - LP_CustomerPayments 1 168.8 132357 45419
## - CurrentDelinquencies 1 203.8 132392 45437
## - LoanMonthsSinceOrigination 1 206.5 132395 45438
## - TotalInquiries 1 296.1 132484 45484
## - DebtToIncomeRatio.imp 1 410.9 132599 45543
## - BankcardUtilization 1 590.7 132779 45635
## - LoanNumber 1 879.9 133068 45783
## - TradesOpenedLast6Months 1 1490.8 133679 46094
## - ProsperRating..numeric. 1 1546.9 133735 46122
## - BorrowerAPR 1 1629.5 133818 46164
## - AvailableBankcardCredit 1 2057.6 134246 46381
## - InquiriesLast6Months 1 2524.0 134712 46616
## - LoanOriginalAmount 1 2612.7 134801 46661
## - CreditScoreRangeLower 1 3145.3 135333 46929
## - EstimatedLoss 1 4539.7 136728 47624
## - EstimatedReturn 1 4837.3 137025 47772
##
## Step: AIC=45332.28
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
## EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
## ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
## TotalCreditLinespast7years + OpenRevolvingAccounts + OpenRevolvingMont
hlyPayment +
## InquiriesLast6Months + TotalInquiries + CurrentDelinquencies +
## AmountDelinquent + DelinquenciesLast7Years + PublicRecordsLast10Years
+
## PublicRecordsLast12Months + RevolvingCreditBalance + BankcardUtilizati
on +
## AvailableBankcardCredit + TradesNeverDelinquent..percentage. +
## TradesOpenedLast6Months + LoanMonthsSinceOrigination + LoanNumber +
## LoanOriginalAmount + LP_CustomerPayments + LP_CustomerPrincipalPayment
s +
## LP_ServiceFees + LP_CollectionFees + LP_GrossPrincipalLoss +

```

```

##      LP_NetPrincipalLoss + LP_NonPrincipalRecoverypayments + PercentFunded
+
##      Recommendations + InvestmentFromFriendsCount + InvestmentFromFriendsAm
ount +
##      Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
##      EmploymentStatus.dum + LoanStatus.dum + IsBorrowerHomeowner.dum
##
##                               Df Sum of Sq      RSS      AIC
## - OpenRevolvingMonthlyPayment      1      0.0 132188 45330
## - LP_CollectionFees                  1      0.2 132188 45330
## - LoanStatus.dum                     1      0.8 132189 45331
## - Recommendations                   1      0.8 132189 45331
## - IsBorrowerHomeowner.dum           1      0.9 132189 45331
## - LP_NetPrincipalLoss                1      1.1 132189 45331
## - BorrowerRate                      1      1.1 132189 45331
## - LP_GrossPrincipalLoss              1      1.9 132190 45331
## <none>                               132188 45332
## - EmploymentStatusDuration.imp       1      4.4 132192 45333
## - RevolvingCreditBalance             1      4.5 132193 45333
## - InvestmentFromFriendsAmount        1      6.0 132194 45333
## - Term                              1      6.2 132194 45333
## - LP_NonPrincipalRecoverypayments    1      6.3 132194 45333
## - PublicRecordsLast12Months          1      6.7 132195 45334
## - AmountDelinquent                  1      7.4 132196 45334
## + TotalTrades                        1      0.0 132188 45334
## - InvestmentFromFriendsCount         1     10.9 132199 45336
## - TotalCreditLinespast7years         1     11.6 132200 45336
## - EmploymentStatus.dum               1     13.9 132202 45337
## - LenderYield                       1     26.5 132215 45344
## - EstimatedEffectiveYield             1     29.8 132218 45346
## - DelinquenciesLast7Years            1     30.1 132218 45346
## - OpenRevolvingAccounts              1     34.5 132223 45348
## - TradesNeverDelinquent..percentage. 1     37.5 132226 45350
## - PercentFunded                     1     40.6 132229 45351
## - CurrentCreditLines                 1     60.5 132249 45361
## - PublicRecordsLast10Years           1     65.2 132253 45364
## - LP_CustomerPrincipalPayments       1     93.0 132281 45378
## - LP_ServiceFees                     1     99.5 132288 45381
## - Investors                         1    140.3 132328 45402
## - LP_CustomerPayments                1    168.8 132357 45417
## - CurrentDelinquencies               1    205.5 132394 45436
## - LoanMonthsSinceOrigination          1    206.5 132395 45436
## - TotalInquiries                     1    296.8 132485 45483
## - DebtToIncomeRatio.imp              1    410.9 132599 45541
## - BankcardUtilization                 1    591.4 132780 45633
## - LoanNumber                         1    880.4 133068 45781
## - TradesOpenedLast6Months            1   1492.6 133681 46092
## - ProsperRating..numeric.            1   1547.0 133735 46120
## - BorrowerAPR                       1   1629.6 133818 46162
## - AvailableBankcardCredit            1   2062.2 134250 46381

```

```

## - InquiriesLast6Months          1    2524.0 134712 46614
## - LoanOriginalAmount            1    2615.0 134803 46660
## - CreditScoreRangeLower         1    3158.8 135347 46933
## - EstimatedLoss                  1    4540.6 136729 47623
## - EstimatedReturn                1    4838.1 137026 47770
##
## Step: AIC=45330.31
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##   EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##   ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##   TotalCreditLinespast7years + OpenRevolvingAccounts + InquiriesLast6Months +
##   TotalInquiries + CurrentDelinquencies + AmountDelinquent +
##   DelinquenciesLast7Years + PublicRecordsLast10Years + PublicRecordsLast12Months +
##   RevolvingCreditBalance + BankcardUtilization + AvailableBankcardCredit +
##   TradesNeverDelinquent..percentage. + TradesOpenedLast6Months +
##   LoanMonthsSinceOrigination + LoanNumber + LoanOriginalAmount +
##   LP_CustomerPayments + LP_CustomerPrincipalPayments + LP_ServiceFees +
##   LP_CollectionFees + LP_GrossPrincipalLoss + LP_NetPrincipalLoss +
##   LP_NonPrincipalRecoverypayments + PercentFunded + Recommendations +
##   InvestmentFromFriendsCount + InvestmentFromFriendsAmount +
##   Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
##   EmploymentStatus.dum + LoanStatus.dum + IsBorrowerHomeowner.dum
##
##                                     Df Sum of Sq    RSS    AIC
## - LP_CollectionFees                1         0.2 132188 45328
## - LoanStatus.dum                   1         0.8 132189 45329
## - Recommendations                  1         0.8 132189 45329
## - IsBorrowerHomeowner.dum          1         0.9 132189 45329
## - LP_NetPrincipalLoss               1         1.1 132189 45329
## - BorrowerRate                     1         1.1 132189 45329
## - LP_GrossPrincipalLoss             1         1.9 132190 45329
## <none>                             132188 45330
## - EmploymentStatusDuration.imp      1         4.4 132193 45331
## - InvestmentFromFriendsAmount       1         6.0 132194 45331
## - Term                             1         6.2 132194 45332
## - LP_NonPrincipalRecoverypayments  1         6.3 132194 45332
## - PublicRecordsLast12Months         1         6.7 132195 45332
## - RevolvingCreditBalance            1         6.8 132195 45332
## - AmountDelinquent                 1         7.4 132196 45332
## + OpenRevolvingMonthlyPayment       1         0.0 132188 45332
## + TotalTrades                       1         0.0 132188 45332
## - InvestmentFromFriendsCount        1        10.9 132199 45334
## - TotalCreditLinespast7years       1        11.7 132200 45334
## - EmploymentStatus.dum              1        13.9 132202 45335
## - LenderYield                      1        26.5 132215 45342
## - EstimatedEffectiveYield           1        29.8 132218 45344
## - DelinquenciesLast7Years           1        30.2 132218 45344

```

```

## - OpenRevolvingAccounts          1      36.5 132225 45347
## - TradesNeverDelinquent..percentage. 1      37.9 132226 45348
## - PercentFunded                  1      40.6 132229 45349
## - CurrentCreditLines              1      60.8 132249 45359
## - PublicRecordsLast10Years        1      65.4 132254 45362
## - LP_CustomerPrincipalPayments    1      93.1 132281 45376
## - LP_ServiceFees                  1      99.5 132288 45379
## - Investors                       1     140.5 132329 45400
## - LP_CustomerPayments              1     168.9 132357 45415
## - CurrentDelinquencies             1     205.4 132394 45434
## - LoanMonthsSinceOrigination       1     206.5 132395 45434
## - TotalInquiries                  1     298.0 132486 45481
## - DebtToIncomeRatio.imp            1     412.2 132600 45540
## - BankcardUtilization              1     607.0 132795 45639
## - LoanNumber                      1     881.4 133070 45779
## - TradesOpenedLast6Months          1    1504.9 133693 46097
## - ProsperRating..numeric.          1    1547.1 133735 46118
## - BorrowerAPR                     1    1631.8 133820 46161
## - AvailableBankcardCredit          1    2067.1 134255 46382
## - InquiriesLast6Months             1    2525.3 134713 46613
## - LoanOriginalAmount               1    2628.9 134817 46665
## - CreditScoreRangeLower            1    3158.8 135347 46931
## - EstimatedLoss                    1    4542.3 136730 47622
## - EstimatedReturn                   1    4839.8 137028 47769
##
## Step:  AIC=45328.41
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##   EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##   ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##   TotalCreditLinespast7years + OpenRevolvingAccounts + InquiriesLast6Months +
##   TotalInquiries + CurrentDelinquencies + AmountDelinquent +
##   DelinquenciesLast7Years + PublicRecordsLast10Years + PublicRecordsLast12Months +
##   RevolvingCreditBalance + BankcardUtilization + AvailableBankcardCredit +
##   TradesNeverDelinquent..percentage. + TradesOpenedLast6Months +
##   LoanMonthsSinceOrigination + LoanNumber + LoanOriginalAmount +
##   LP_CustomerPayments + LP_CustomerPrincipalPayments + LP_ServiceFees +
##   LP_GrossPrincipalLoss + LP_NetPrincipalLoss + LP_NonPrincipalRecoverypayments +
##   PercentFunded + Recommendations + InvestmentFromFriendsCount +
##   InvestmentFromFriendsAmount + Investors + DebtToIncomeRatio.imp +
##   EmploymentStatusDuration.imp + EmploymentStatus.dum + LoanStatus.dum +
##   IsBorrowerHomeowner.dum
##
##           Df Sum of Sq    RSS    AIC
## - LoanStatus.dum      1      0.7 132189 45327
## - Recommendations     1      0.8 132189 45327
## - IsBorrowerHomeowner.dum 1      0.8 132189 45327

```

```

## - LP_NetPrincipalLoss          1          0.9 132189 45327
## - BorrowerRate                  1          1.1 132189 45327
## - LP_GrossPrincipalLoss         1          3.4 132192 45328
## <none>                          132188 45328
## - EmploymentStatusDuration.imp  1          4.4 132193 45329
## - InvestmentFromFriendsAmount    1          6.0 132194 45329
## - Term                           1          6.3 132195 45330
## - PublicRecordsLast12Months      1          6.7 132195 45330
## - RevolvingCreditBalance         1          6.8 132195 45330
## - AmountDelinquent              1          7.4 132196 45330
## + LP_CollectionFees              1          0.2 132188 45330
## + OpenRevolvingMonthlyPayment    1          0.0 132188 45330
## + TotalTrades                    1          0.0 132188 45330
## - LP_NonPrincipalRecoverypayments 1          9.8 132198 45331
## - InvestmentFromFriendsCount      1         10.9 132199 45332
## - TotalCreditLinespast7years     1         11.7 132200 45332
## - EmploymentStatus.dum           1         13.9 132202 45334
## - LenderYield                    1         26.5 132215 45340
## - EstimatedEffectiveYield         1         29.9 132218 45342
## - DelinquenciesLast7Years         1         30.2 132219 45342
## - OpenRevolvingAccounts           1         36.6 132225 45345
## - TradesNeverDelinquent..percentage. 1         37.8 132226 45346
## - PercentFunded                  1         40.6 132229 45347
## - CurrentCreditLines              1         60.7 132249 45358
## - PublicRecordsLast10Years        1         65.5 132254 45360
## - LP_CustomerPrincipalPayments    1         96.4 132285 45376
## - LP_ServiceFees                  1         99.6 132288 45378
## - Investors                       1        140.4 132329 45398
## - LP_CustomerPayments              1        174.5 132363 45416
## - CurrentDelinquencies             1        205.4 132394 45432
## - LoanMonthsSinceOrigination       1        206.4 132395 45432
## - TotalInquiries                  1        298.1 132486 45479
## - DebtToIncomeRatio.imp           1        412.1 132600 45538
## - BankcardUtilization              1        607.1 132795 45637
## - LoanNumber                      1        881.5 133070 45778
## - TradesOpenedLast6Months          1       1504.9 133693 46095
## - ProsperRating..numeric.          1       1547.4 133736 46116
## - BorrowerAPR                     1       1631.8 133820 46159
## - AvailableBankcardCredit          1       2066.9 134255 46380
## - InquiriesLast6Months             1       2525.1 134713 46611
## - LoanOriginalAmount               1       2629.1 134817 46663
## - CreditScoreRangeLower            1       3158.6 135347 46929
## - EstimatedLoss                    1       4544.9 136733 47621
## - EstimatedReturn                   1       4843.5 137032 47769
##
## Step:  AIC=45326.76
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##     EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##     ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##     TotalCreditLinespast7years + OpenRevolvingAccounts + InquiriesLast6Mon

```



```

ths +
##      TotalInquiries + CurrentDelinquencies + AmountDelinquent +
##      DelinquenciesLast7Years + PublicRecordsLast10Years + PublicRecordsLast
12Months +
##      RevolvingCreditBalance + BankcardUtilization + AvailableBankcardCredit
+
##      TradesNeverDelinquent..percentage. + TradesOpenedLast6Months +
##      LoanMonthsSinceOrigination + LoanNumber + LoanOriginalAmount +
##      LP_CustomerPayments + LP_CustomerPrincipalPayments + LP_ServiceFees +
##      LP_GrossPrincipalLoss + LP_NetPrincipalLoss + LP_NonPrincipalRecoveryp
ayments +
##      PercentFunded + Recommendations + InvestmentFromFriendsCount +
##      InvestmentFromFriendsAmount + Investors + DebtToIncomeRatio.imp +
##      EmploymentStatusDuration.imp + EmploymentStatus.dum + IsBorrowerHomeow
ner.dum
##
##
##          Df Sum of Sq    RSS    AIC
## - Recommendations          1      0.8 132190 45325
## - IsBorrowerHomeowner.dum    1      0.9 132190 45325
## - LP_NetPrincipalLoss        1      0.9 132190 45325
## - BorrowerRate              1      1.2 132190 45325
## - LP_GrossPrincipalLoss      1      3.2 132192 45326
## <none>                      132189 45327
## - EmploymentStatusDuration.imp 1      4.5 132193 45327
## - InvestmentFromFriendsAmount 1      6.0 132195 45328
## - Term                      1      6.1 132195 45328
## - PublicRecordsLast12Months  1      6.6 132196 45328
## - RevolvingCreditBalance     1      6.8 132196 45328
## + LoanStatus.dum            1      0.7 132188 45328
## - AmountDelinquent          1      7.5 132196 45329
## + LP_CollectionFees         1      0.1 132189 45329
## + OpenRevolvingMonthlyPayment 1      0.0 132189 45329
## + TotalTrades               1      0.0 132189 45329
## - LP_NonPrincipalRecoverypayments 1      9.6 132199 45330
## - InvestmentFromFriendsCount  1     10.9 132200 45330
## - TotalCreditLinespast7years  1     11.7 132201 45331
## - EmploymentStatus.dum      1     13.8 132203 45332
## - LenderYield               1     26.6 132216 45338
## - EstimatedEffectiveYield    1     29.7 132219 45340
## - DelinquenciesLast7Years    1     30.1 132219 45340
## - OpenRevolvingAccounts      1     36.5 132226 45344
## - TradesNeverDelinquent..percentage. 1     37.7 132227 45344
## - PercentFunded             1     40.7 132230 45346
## - CurrentCreditLines        1     60.9 132250 45356
## - PublicRecordsLast10Years   1     65.4 132254 45358
## - LP_CustomerPrincipalPayments 1     96.7 132286 45374
## - LP_ServiceFees            1     99.4 132288 45376
## - Investors                 1    140.8 132330 45397
## - LP_CustomerPayments       1    175.6 132365 45415
## - CurrentDelinquencies      1    205.0 132394 45430

```

```

## - LoanMonthsSinceOrigination      1      206.1 132395 45431
## - TotalInquiries                   1      298.1 132487 45478
## - DebtToIncomeRatio.imp            1      411.8 132601 45536
## - BankcardUtilization              1      608.9 132798 45637
## - LoanNumber                       1      889.1 133078 45780
## - TradesOpenedLast6Months          1     1504.2 133693 46093
## - ProsperRating..numeric.          1     1547.6 133737 46115
## - BorrowerAPR                      1     1635.3 133824 46159
## - AvailableBankcardCredit          1     2067.7 134257 46378
## - InquiriesLast6Months             1     2524.4 134713 46609
## - LoanOriginalAmount               1     2636.4 134825 46665
## - CreditScoreRangeLower            1     3163.1 135352 46930
## - EstimatedLoss                     1     4559.3 136748 47627
## - EstimatedReturn                   1     4856.7 137046 47774
##
## Step:  AIC=45325.2
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##   EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##   ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##   TotalCreditLinespast7years + OpenRevolvingAccounts + InquiriesLast6Months +
##   TotalInquiries + CurrentDelinquencies + AmountDelinquent +
##   DelinquenciesLast7Years + PublicRecordsLast10Years + PublicRecordsLast12Months +
##   RevolvingCreditBalance + BankcardUtilization + AvailableBankcardCredit +
##   TradesNeverDelinquent..percentage. + TradesOpenedLast6Months +
##   LoanMonthsSinceOrigination + LoanNumber + LoanOriginalAmount +
##   LP_CustomerPayments + LP_CustomerPrincipalPayments + LP_ServiceFees +
##   LP_GrossPrincipalLoss + LP_NetPrincipalLoss + LP_NonPrincipalRecoverypayments +
##   PercentFunded + InvestmentFromFriendsCount + InvestmentFromFriendsAmount +
##   Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
##   EmploymentStatus.dum + IsBorrowerHomeowner.dum
##
##                                     Df Sum of Sq    RSS    AIC
## - IsBorrowerHomeowner.dum          1         0.9 132191 45324
## - LP_NetPrincipalLoss               1         0.9 132191 45324
## - BorrowerRate                     1         1.2 132191 45324
## - LP_GrossPrincipalLoss             1         3.2 132193 45325
## <none>                             132190 45325
## - EmploymentStatusDuration.imp      1         4.5 132194 45325
## - InvestmentFromFriendsAmount        1         5.7 132196 45326
## - Term                              1         6.1 132196 45326
## - PublicRecordsLast12Months          1         6.6 132196 45327
## - RevolvingCreditBalance            1         6.9 132197 45327
## + Recommendations                   1         0.8 132189 45327
## + LoanStatus.dum                    1         0.7 132189 45327
## - AmountDelinquent                  1         7.5 132197 45327

```

```

## + LP_CollectionFees          1      0.1 132190 45327
## + OpenRevolvingMonthlyPayment 1      0.0 132190 45327
## + TotalTrades                 1      0.0 132190 45327
## - LP_NonPrincipalRecoverypayments 1      9.6 132199 45328
## - TotalCreditLinespast7years 1     11.8 132202 45329
## - EmploymentStatus.dum       1     13.7 132204 45330
## - InvestmentFromFriendsCount  1     18.6 132208 45333
## - LenderYield                 1     26.6 132216 45337
## - EstimatedEffectiveYield     1     29.4 132219 45338
## - DelinquenciesLast7Years    1     30.2 132220 45339
## - OpenRevolvingAccounts      1     36.6 132226 45342
## - TradesNeverDelinquent..percentage. 1    37.6 132227 45343
## - PercentFunded              1     40.9 132231 45344
## - CurrentCreditLines         1     61.0 132251 45355
## - PublicRecordsLast10Years   1     65.4 132255 45357
## - LP_CustomerPrincipalPayments 1     97.0 132287 45373
## - LP_ServiceFees             1     99.1 132289 45374
## - Investors                   1    140.8 132331 45395
## - LP_CustomerPayments        1    176.1 132366 45414
## - CurrentDelinquencies       1    205.1 132395 45428
## - LoanMonthsSinceOrigination  1    206.6 132396 45429
## - TotalInquiries             1    298.5 132488 45476
## - DebtToIncomeRatio.imp      1   412.0 132602 45534
## - BankcardUtilization        1   609.1 132799 45635
## - LoanNumber                 1   888.3 133078 45778
## - TradesOpenedLast6Months    1  1503.8 133694 46091
## - ProsperRating..numeric.    1  1549.8 133740 46114
## - BorrowerAPR                1  1634.9 133825 46158
## - AvailableBankcardCredit    1  2066.9 134257 46376
## - InquiriesLast6Months       1  2524.8 134715 46608
## - LoanOriginalAmount         1  2636.6 134826 46664
## - CreditScoreRangeLower      1  3163.7 135354 46929
## - EstimatedLoss              1  4559.2 136749 47625
## - EstimatedReturn            1  4856.3 137046 47772
##
## Step:  AIC=45323.64
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##   EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##   ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##   TotalCreditLinespast7years + OpenRevolvingAccounts + InquiriesLast6Months +
##   TotalInquiries + CurrentDelinquencies + AmountDelinquent +
##   DelinquenciesLast7Years + PublicRecordsLast10Years + PublicRecordsLast12Months +
##   RevolvingCreditBalance + BankcardUtilization + AvailableBankcardCredit +
##   TradesNeverDelinquent..percentage. + TradesOpenedLast6Months +
##   LoanMonthsSinceOrigination + LoanNumber + LoanOriginalAmount +
##   LP_CustomerPayments + LP_CustomerPrincipalPayments + LP_ServiceFees +
##   LP_GrossPrincipalLoss + LP_NetPrincipalLoss + LP_NonPrincipalRecoveryp

```

```

payments +
##      PercentFunded + InvestmentFromFriendsCount + InvestmentFromFriendsAmount +
##      Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
##      EmploymentStatus.dum
##
##              Df Sum of Sq    RSS    AIC
## - LP_NetPrincipalLoss      1      0.9 132192 45322
## - BorrowerRate              1      1.2 132192 45322
## - LP_GrossPrincipalLoss     1      3.2 132194 45323
## <none>                      132191 45324
## - EmploymentStatusDuration.imp 1      4.1 132195 45324
## - InvestmentFromFriendsAmount  1      5.6 132196 45325
## - Term                      1      6.1 132197 45325
## - PublicRecordsLast12Months  1      6.6 132197 45325
## + IsBorrowerHomeowner.dum    1      0.9 132190 45325
## + Recommendations            1      0.8 132190 45325
## + LoanStatus.dum             1      0.7 132190 45325
## - AmountDelinquent           1      7.3 132198 45325
## - RevolvingCreditBalance     1      7.4 132198 45325
## + LP_CollectionFees          1      0.1 132191 45326
## + TotalTrades                1      0.0 132191 45326
## + OpenRevolvingMonthlyPayment 1      0.0 132191 45326
## - LP_NonPrincipalRecoverypayments 1      9.6 132200 45327
## - TotalCreditLinespast7years  1     11.3 132202 45327
## - EmploymentStatus.dum       1     13.8 132204 45329
## - InvestmentFromFriendsCount  1     18.7 132209 45331
## - LenderYield                1     26.6 132217 45335
## - EstimatedEffectiveYield     1     29.3 132220 45337
## - DelinquenciesLast7Years     1     30.0 132221 45337
## - OpenRevolvingAccounts       1     35.8 132226 45340
## - TradesNeverDelinquent..percentage. 1     38.5 132229 45341
## - PercentFunded              1     40.8 132232 45343
## - CurrentCreditLines          1     60.3 132251 45353
## - PublicRecordsLast10Years    1     65.1 132256 45355
## - LP_CustomerPrincipalPayments 1     97.0 132288 45371
## - LP_ServiceFees              1     99.2 132290 45373
## - Investors                   1    140.9 132332 45394
## - LP_CustomerPayments         1    176.1 132367 45412
## - CurrentDelinquencies        1    204.3 132395 45426
## - LoanMonthsSinceOrigination  1    207.0 132398 45428
## - TotalInquiries              1    298.1 132489 45475
## - DebtToIncomeRatio.imp       1    414.1 132605 45534
## - BankcardUtilization         1    624.0 132815 45641
## - LoanNumber                  1    888.1 133079 45776
## - TradesOpenedLast6Months     1   1503.1 133694 46089
## - ProsperRating..numeric.     1   1549.4 133740 46113
## - BorrowerAPR                 1   1637.0 133828 46157
## - AvailableBankcardCredit     1   2066.1 134257 46374
## - InquiriesLast6Months        1  2526.0 134717 46607

```

```

## - LoanOriginalAmount          1    2649.4 134840 46669
## - CreditScoreRangeLower       1    3432.6 135623 47062
## - EstimatedLoss                1    4564.4 136755 47626
## - EstimatedReturn              1    4860.6 137051 47773
##
## Step: AIC=45322.12
## ProsperScore ~ Term + BorrowerAPR + BorrowerRate + LenderYield +
##   EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##   ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##   TotalCreditLinespast7years + OpenRevolvingAccounts + InquiriesLast6Mon
ths +
##   TotalInquiries + CurrentDelinquencies + AmountDelinquent +
##   DelinquenciesLast7Years + PublicRecordsLast10Years + PublicRecordsLast
12Months +
##   RevolvingCreditBalance + BankcardUtilization + AvailableBankcardCredit
+
##   TradesNeverDelinquent..percentage. + TradesOpenedLast6Months +
##   LoanMonthsSinceOrigination + LoanNumber + LoanOriginalAmount +
##   LP_CustomerPayments + LP_CustomerPrincipalPayments + LP_ServiceFees +
##   LP_GrossPrincipalLoss + LP_NonPrincipalRecoverypayments +
##   PercentFunded + InvestmentFromFriendsCount + InvestmentFromFriendsAmou
nt +
##   Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
##   EmploymentStatus.dum
##
##                               Df Sum of Sq    RSS    AIC
## - BorrowerRate                1         1.2 132193 45321
## <none>                        132192 45322
## - EmploymentStatusDuration.imp 1         4.1 132196 45322
## - InvestmentFromFriendsAmount  1         5.6 132197 45323
## - Term                         1         6.1 132198 45323
## - PublicRecordsLast12Months    1         6.6 132198 45323
## + LP_NetPrincipalLoss          1         0.9 132191 45324
## + IsBorrowerHomeowner.dum      1         0.9 132191 45324
## + Recommendations              1         0.8 132191 45324
## + LoanStatus.dum               1         0.7 132191 45324
## - AmountDelinquent             1         7.3 132199 45324
## - RevolvingCreditBalance        1         7.4 132199 45324
## + LP_CollectionFees            1         0.1 132192 45324
## + TotalTrades                  1         0.1 132192 45324
## + OpenRevolvingMonthlyPayment  1         0.0 132192 45324
## - LP_NonPrincipalRecoverypayments 1         9.2 132201 45325
## - TotalCreditLinespast7years   1        11.4 132203 45326
## - EmploymentStatus.dum         1        13.8 132205 45327
## - InvestmentFromFriendsCount    1        18.7 132210 45330
## - LenderYield                  1        26.7 132218 45334
## - EstimatedEffectiveYield       1        29.3 132221 45335
## - DelinquenciesLast7Years       1        30.0 132222 45336
## - OpenRevolvingAccounts         1        35.7 132227 45338
## - TradesNeverDelinquent..percentage. 1        38.5 132230 45340

```

```

## - PercentFunded          1      40.7 132232 45341
## - CurrentCreditLines     1      60.2 132252 45351
## - PublicRecordsLast10Years 1      65.0 132257 45353
## - LP_CustomerPrincipalPayments 1      97.8 132289 45370
## - LP_ServiceFees         1      98.5 132290 45371
## - Investors              1     140.8 132332 45392
## - LP_CustomerPayments    1     177.2 132369 45411
## - CurrentDelinquencies   1     204.3 132396 45425
## - LoanMonthsSinceOrigination 1     207.4 132399 45427
## - TotalInquiries         1     298.0 132490 45473
## - DebtToIncomeRatio.imp   1     414.7 132606 45533
## - LP_GrossPrincipalLoss   1     618.4 132810 45637
## - BankcardUtilization     1     624.0 132816 45640
## - LoanNumber              1     887.3 133079 45774
## - TradesOpenedLast6Months 1    1503.1 133695 46088
## - ProsperRating..numeric. 1    1549.1 133741 46111
## - BorrowerAPR             1    1636.3 133828 46155
## - AvailableBankcardCredit 1    2067.0 134259 46373
## - InquiriesLast6Months    1    2525.9 134718 46605
## - LoanOriginalAmount       1    2648.5 134840 46667
## - CreditScoreRangeLower    1    3432.3 135624 47060
## - EstimatedLoss            1    4564.8 136756 47625
## - EstimatedReturn          1    4861.0 137053 47771
##
## Step: AIC=45320.72
## ProsperScore ~ Term + BorrowerAPR + LenderYield + EstimatedEffectiveYield
+
## EstimatedLoss + EstimatedReturn + ProsperRating..numeric. +
## CreditScoreRangeLower + CurrentCreditLines + TotalCreditLinespast7years +
## OpenRevolvingAccounts + InquiriesLast6Months + TotalInquiries +
## CurrentDelinquencies + AmountDelinquent + DelinquenciesLast7Years +
## PublicRecordsLast10Years + PublicRecordsLast12Months + RevolvingCredit
Balance +
## BankcardUtilization + AvailableBankcardCredit + TradesNeverDelinquent.
.percentage. +
## TradesOpenedLast6Months + LoanMonthsSinceOrigination + LoanNumber +
## LoanOriginalAmount + LP_CustomerPayments + LP_CustomerPrincipalPayments +
## LP_ServiceFees + LP_GrossPrincipalLoss + LP_NonPrincipalRecoverypayments +
## PercentFunded + InvestmentFromFriendsCount + InvestmentFromFriendsAmount +
## Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
## EmploymentStatus.dum
##
## Df Sum of Sq RSS AIC
## <none> 132193 45321
## - EmploymentStatusDuration.imp 1 4.1 132197 45321
## - InvestmentFromFriendsAmount 1 5.6 132198 45322

```

## - Term	1	6.3	132199	45322
## - PublicRecordsLast12Months	1	6.6	132199	45322
## + BorrowerRate	1	1.2	132192	45322
## + LP_NetPrincipalLoss	1	0.9	132192	45322
## + IsBorrowerHomeowner.dum	1	0.9	132192	45322
## + Recommendations	1	0.8	132192	45322
## + LoanStatus.dum	1	0.7	132192	45322
## - AmountDelinquent	1	7.3	132200	45322
## - RevolvingCreditBalance	1	7.4	132200	45323
## + LP_CollectionFees	1	0.1	132193	45323
## + TotalTrades	1	0.1	132193	45323
## + OpenRevolvingMonthlyPayment	1	0.0	132193	45323
## - LP_NonPrincipalRecoverypayments	1	9.2	132202	45323
## - TotalCreditLinespast7years	1	11.4	132204	45325
## - EmploymentStatus.dum	1	13.7	132207	45326
## - InvestmentFromFriendsCount	1	18.8	132212	45328
## - EstimatedEffectiveYield	1	29.3	132222	45334
## - DelinquenciesLast7Years	1	29.9	132223	45334
## - OpenRevolvingAccounts	1	35.8	132229	45337
## - TradesNeverDelinquent..percentage.	1	38.6	132231	45339
## - PercentFunded	1	40.6	132233	45340
## - CurrentCreditLines	1	60.2	132253	45350
## - PublicRecordsLast10Years	1	65.0	132258	45352
## - LP_CustomerPrincipalPayments	1	97.7	132291	45369
## - LP_ServiceFees	1	98.6	132291	45369
## - Investors	1	140.9	132334	45391
## - LP_CustomerPayments	1	177.2	132370	45410
## - CurrentDelinquencies	1	204.3	132397	45424
## - LoanMonthsSinceOrigination	1	209.5	132402	45426
## - TotalInquiries	1	297.7	132491	45471
## - DebtToIncomeRatio.imp	1	414.7	132608	45531
## - LP_GrossPrincipalLoss	1	618.5	132811	45636
## - BankcardUtilization	1	623.9	132817	45638
## - LoanNumber	1	886.2	133079	45772
## - TradesOpenedLast6Months	1	1502.6	133695	46086
## - ProsperRating..numeric.	1	1551.1	133744	46111
## - BorrowerAPR	1	1635.5	133828	46153
## - AvailableBankcardCredit	1	2067.2	134260	46372
## - InquiriesLast6Months	1	2526.1	134719	46604
## - LoanOriginalAmount	1	2648.2	134841	46665
## - LenderYield	1	2915.7	135108	46800
## - CreditScoreRangeLower	1	3431.9	135625	47059
## - EstimatedLoss	1	4564.4	136757	47623
## - EstimatedReturn	1	4860.7	137053	47770

```
## 5 Minutes atleast
summary(stepreg.train)
```

```
##
```

```
## Call:
```

```

## lm(formula = ProsperScore ~ Term + BorrowerAPR + LenderYield +
##     EstimatedEffectiveYield + EstimatedLoss + EstimatedReturn +
##     ProsperRating..numeric. + CreditScoreRangeLower + CurrentCreditLines +
##     TotalCreditLinespast7years + OpenRevolvingAccounts + InquiriesLast6Mon
ths +
##     TotalInquiries + CurrentDelinquencies + AmountDelinquent +
##     DelinquenciesLast7Years + PublicRecordsLast10Years + PublicRecordsLast
12Months +
##     RevolvingCreditBalance + BankcardUtilization + AvailableBankcardCredit
+
##     TradesNeverDelinquent..percentage. + TradesOpenedLast6Months +
##     LoanMonthsSinceOrigination + LoanNumber + LoanOriginalAmount +
##     LP_CustomerPayments + LP_CustomerPrincipalPayments + LP_ServiceFees +
##     LP_GrossPrincipalLoss + LP_NonPrincipalRecoverypayments +
##     PercentFunded + InvestmentFromFriendsCount + InvestmentFromFriendsAmou
nt +
##     Investors + DebtToIncomeRatio.imp + EmploymentStatusDuration.imp +
##     EmploymentStatus.dum, data = train.num)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.0548 -0.9012  0.0083  0.8696  7.8824
##
## Coefficients:
##              Estimate      Std. Error t value
## (Intercept)    14.3128266684    0.3533235754   40.509
## Term              0.0012900899    0.0007177426    1.797
## BorrowerAPR    -26.0229160007    0.8982224551  -28.972
## LenderYield    -73.9344884279    1.9113054059  -38.683
## EstimatedEffectiveYield -0.9571329696    0.2468723976   -3.877
## EstimatedLoss    99.9788201942    2.0657066811   48.399
## EstimatedReturn   98.8123503543    1.9784014127   49.946
## ProsperRating..numeric.  0.4760940224    0.0168741882   28.214
## CreditScoreRangeLower -0.0071810911    0.0001711103  -41.968
## CurrentCreditLines -0.0122556154    0.0022041034   -5.560
## TotalCreditLinespast7years -0.0013652769    0.0005641519   -2.420
## OpenRevolvingAccounts  0.0099887693    0.0023308836    4.285
## InquiriesLast6Months -0.1814781682    0.0050402524  -36.006
## TotalInquiries    -0.0220454872    0.0017833910  -12.362
## CurrentDelinquencies -0.0567376869    0.0055404499  -10.241
## AmountDelinquent -0.0000015298    0.0000007917   -1.932
## DelinquenciesLast7Years  0.0028248897    0.0007205663    3.920
## PublicRecordsLast10Years  0.0526385055    0.0091135229    5.776
## PublicRecordsLast12Months -0.0782912323    0.0425560049   -1.840
## RevolvingCreditBalance  0.0000003902    0.0000002001    1.950
## BankcardUtilization -0.4258251731    0.0237965396  -17.894
## AvailableBankcardCredit  0.0000126854    0.0000003895   32.572
## TradesNeverDelinquent..percentage. -0.2677862459    0.0601810792   -4.450
## TradesOpenedLast6Months -0.1677484368    0.0060406543  -27.770
## LoanMonthsSinceOrigination  0.0168269447    0.0016228391   10.369

```


## LoanNumber	-0.0000166183	0.0000007792	-21.326
## LoanOriginalAmount	0.0000526860	0.0000014291	36.866
## LP_CustomerPayments	-0.0001376449	0.0000144345	-9.536
## LP_CustomerPrincipalPayments	0.0001020924	0.0000144149	7.082
## LP_ServiceFees	0.0020967246	0.0002947813	7.113
## LP_GrossPrincipalLoss	-0.0000623858	0.0000035017	-17.816
## LP_NonPrincipalRecoverypayments	-0.0001028159	0.0000474041	-2.169
## PercentFunded	-1.1870649079	0.2599432836	-4.567
## InvestmentFromFriendsCount	-0.1628221776	0.0524415770	-3.105
## InvestmentFromFriendsAmount	-0.0000839632	0.0000493627	-1.701
## Investors	0.0006366196	0.0000748569	8.504
## DebtToIncomeRatio.imp	-0.2542829274	0.0174295991	-14.589
## EmploymentStatusDuration.imp	-0.0000823767	0.0000568239	-1.450
## EmploymentStatus.dum	0.0607859262	0.0228965403	2.655
##	Pr(> t)		
## (Intercept)	< 0.0000000000000002	***	
## Term	0.072272	.	
## BorrowerAPR	< 0.0000000000000002	***	
## LenderYield	< 0.0000000000000002	***	
## EstimatedEffectiveYield	0.000106	***	
## EstimatedLoss	< 0.0000000000000002	***	
## EstimatedReturn	< 0.0000000000000002	***	
## ProsperRating..numeric.	< 0.0000000000000002	***	
## CreditScoreRangeLower	< 0.0000000000000002	***	
## CurrentCreditLines	0.00000002702242	***	
## TotalCreditLinespast7years	0.015521	*	
## OpenRevolvingAccounts	0.00001826615014	***	
## InquiriesLast6Months	< 0.0000000000000002	***	
## TotalInquiries	< 0.0000000000000002	***	
## CurrentDelinquencies	< 0.0000000000000002	***	
## AmountDelinquent	0.053331	.	
## DelinquenciesLast7Years	0.00008849830791	***	
## PublicRecordsLast10Years	0.00000000768905	***	
## PublicRecordsLast12Months	0.065813	.	
## RevolvingCreditBalance	0.051220	.	
## BankcardUtilization	< 0.0000000000000002	***	
## AvailableBankcardCredit	< 0.0000000000000002	***	
## TradesNeverDelinquent..percentage.	0.00000861369576	***	
## TradesOpenedLast6Months	< 0.0000000000000002	***	
## LoanMonthsSinceOrigination	< 0.0000000000000002	***	
## LoanNumber	< 0.0000000000000002	***	
## LoanOriginalAmount	< 0.0000000000000002	***	
## LP_CustomerPayments	< 0.0000000000000002	***	
## LP_CustomerPrincipalPayments	0.00000000000143	***	
## LP_ServiceFees	0.00000000000115	***	
## LP_GrossPrincipalLoss	< 0.0000000000000002	***	
## LP_NonPrincipalRecoverypayments	0.030092	*	
## PercentFunded	0.00000496496436	***	
## InvestmentFromFriendsCount	0.001905	**	
## InvestmentFromFriendsAmount	0.088958	.	

```

## Investors < 0.0000000000000002 ***
## DebtToIncomeRatio.imp < 0.0000000000000002 ***
## EmploymentStatusDuration.imp 0.147152
## EmploymentStatus.dum 0.007937 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.396 on 67843 degrees of freedom
## Multiple R-squared: 0.6547, Adjusted R-squared: 0.6545
## F-statistic: 3385 on 38 and 67843 DF, p-value: < 0.0000000000000022

## Reducing Less Significant Variables
train.num.red <- train.num[,c(-5,-11,-13,-18,-21,-35,-45,-48,-49)]
train.lm.2 <- lm(ProsperScore~.,data=train.num.red)
## Try 2
summary(train.lm.2)

##
## Call:
## lm(formula = ProsperScore ~ ., data = train.num.red)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.0423 -0.9003  0.0085  0.8708  7.8900
##
## Coefficients: (1 not defined because of singularities)
##              Estimate      Std. Error t value
## (Intercept)  13.8434776090    0.4221842702   32.790
## Term         0.0015370886    0.0007155689    2.148
## BorrowerAPR  -25.8538726608    0.8992291977  -28.751
## BorrowerRate  19.5668117842   25.1660103253    0.778
## LenderYield  -93.5208971599   25.2587107052   -3.703
## EstimatedLoss  99.6330640287    2.0650582365   48.247
## EstimatedReturn 97.6113888425    1.9548490456   49.933
## ProsperRating..numeric.  0.4931825287    0.0164066551   30.060
## CreditScoreRangeLower  -0.0072386667    0.0001704169  -42.476
## TotalCreditLinespast7years -0.0023039542    0.0011733723   -1.964
## OpenRevolvingMonthlyPayment -0.0000057408    0.0000215844   -0.266
## InquiriesLast6Months  -0.1820837987    0.0050321951  -36.184
## TotalInquiries  -0.0225407465    0.0017830113  -12.642
## CurrentDelinquencies  -0.0593086097    0.0052535640  -11.289
## DelinquenciesLast7Years  0.0030309496    0.0007070723    4.287
## PublicRecordsLast10Years  0.0536130900    0.0087283461    6.142
## RevolvingCreditBalance  0.0000003652    0.0000002619    1.394
## BankcardUtilization  -0.4314016337    0.0240758996  -17.918
## AvailableBankcardCredit  0.0000127227    0.0000003815   33.349
## TotalTrades  -0.0003320567    0.0013893527   -0.239
## TradesNeverDelinquent..percentage. -0.2957621967    0.0595601965   -4.966
## TradesOpenedLast6Months  -0.1720840679    0.0059723923  -28.813
## LoanMonthsSinceOrigination  0.0200503504    0.0013897115   14.428

```

## LoanNumber	-0.0000155725	0.0000007392	-21.066
## LoanOriginalAmount	0.0000531239	0.0000014296	37.161
## LP_CustomerPayments	-0.0001393271	0.0000144563	-9.638
## LP_CustomerPrincipalPayments	0.0001035592	0.0000144268	7.178
## LP_InterestandFees	NA	NA	NA
## LP_ServiceFees	0.0020890030	0.0002952360	7.076
## LP_GrossPrincipalLoss	-0.0000433248	0.0000318071	-1.362
## LP_NetPrincipalLoss	-0.0000210713	0.0000316796	-0.665
## LP_NonPrincipalRecoverypayments	-0.0001246122	0.0000552254	-2.256
## PercentFunded	-1.1206126156	0.2594272391	-4.320
## Recommendations	-0.0208159705	0.0340520445	-0.611
## InvestmentFromFriendsCount	-0.1435928075	0.0602516531	-2.383
## InvestmentFromFriendsAmount	-0.0000813595	0.0000495318	-1.643
## Investors	0.0006456043	0.0000748815	8.622
## DebtToIncomeRatio.imp	-0.2607282305	0.0174032498	-14.982
## EmploymentStatus.dum	0.0499015491	0.0228347968	2.185
## LoanStatus.dum	-0.0141825380	0.0251428929	-0.564
##	Pr(> t)		
## (Intercept)	< 0.00000000000000002	***	
## Term	0.031712	*	
## BorrowerAPR	< 0.00000000000000002	***	
## BorrowerRate	0.436861		
## LenderYield	0.000214	***	
## EstimatedLoss	< 0.00000000000000002	***	
## EstimatedReturn	< 0.00000000000000002	***	
## ProsperRating..numeric.	< 0.00000000000000002	***	
## CreditScoreRangeLower	< 0.00000000000000002	***	
## TotalCreditLinespast7years	0.049588	*	
## OpenRevolvingMonthlyPayment	0.790263		
## InquiriesLast6Months	< 0.00000000000000002	***	
## TotalInquiries	< 0.00000000000000002	***	
## CurrentDelinquencies	< 0.00000000000000002	***	
## DelinquenciesLast7Years	0.000018166311866	***	
## PublicRecordsLast10Years	0.000000000817279	***	
## RevolvingCreditBalance	0.163249		
## BankcardUtilization	< 0.00000000000000002	***	
## AvailableBankcardCredit	< 0.00000000000000002	***	
## TotalTrades	0.811106		
## TradesNeverDelinquent..percentage.	0.000000685949900	***	
## TradesOpenedLast6Months	< 0.00000000000000002	***	
## LoanMonthsSinceOrigination	< 0.00000000000000002	***	
## LoanNumber	< 0.00000000000000002	***	
## LoanOriginalAmount	< 0.00000000000000002	***	
## LP_CustomerPayments	< 0.00000000000000002	***	
## LP_CustomerPrincipalPayments	0.0000000000000713	***	
## LP_InterestandFees	NA		
## LP_ServiceFees	0.0000000000001501	***	
## LP_GrossPrincipalLoss	0.173167		
## LP_NetPrincipalLoss	0.505965		
## LP_NonPrincipalRecoverypayments	0.024047	*	

```

## PercentFunded          0.000015655940395 ***
## Recommendations        0.541004
## InvestmentFromFriendsCount 0.017165 *
## InvestmentFromFriendsAmount 0.100476
## Investors              < 0.0000000000000002 ***
## DebtToIncomeRatio.imp  < 0.0000000000000002 ***
## EmploymentStatus.dum   0.028868 *
## LoanStatus.dum         0.572703
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.396 on 67843 degrees of freedom
## Multiple R-squared:  0.6544, Adjusted R-squared:  0.6542
## F-statistic: 3381 on 38 and 67843 DF, p-value: < 0.00000000000000022

## Reduced 2 (Super Reduced!)
train.num.red.2 <- train.num.red[,c(1,7,8,9,14,18,19,23,40)]
train.lm.3 <- lm(ProsperScore~., data=train.num.red.2)
summary(train.lm.3)

##
## Call:
## lm(formula = ProsperScore ~ ., data = train.num.red.2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.0915 -1.0676  0.0109  0.9461  7.8102
##
## Coefficients:
##              Estimate      Std. Error t value
## (Intercept)    5.0824960242    0.1147876727  44.277
## Term           0.0111211358    0.0005122535  21.710
## ProsperRating..numeric.  1.1248054384    0.0043346417 259.492
## CreditScoreRangeLower  -0.0072338373    0.0001660285 -43.570
## CurrentDelinquencies  -0.0353084702    0.0052997900  -6.662
## BankcardUtilization  -0.2403749247    0.0222440629 -10.806
## AvailableBankcardCredit  0.0000103855    0.0000003662  28.358
## LoanMonthsSinceOrigination 0.0562653247    0.0004530881 124.182
## LoanStatus.dum       0.0071797628    0.0202173785   0.355
##
##              Pr(>|t|)
## (Intercept) < 0.0000000000000002 ***
## Term        < 0.0000000000000002 ***
## ProsperRating..numeric. < 0.0000000000000002 ***
## CreditScoreRangeLower  < 0.0000000000000002 ***
## CurrentDelinquencies   0.0000000000272 ***
## BankcardUtilization    < 0.0000000000000002 ***
## AvailableBankcardCredit < 0.0000000000000002 ***
## LoanMonthsSinceOrigination < 0.0000000000000002 ***
## LoanStatus.dum         0.722
## ---

```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.499 on 67873 degrees of freedom
## Multiple R-squared:  0.6018, Adjusted R-squared:  0.6018
## F-statistic: 1.282e+04 on 8 and 67873 DF,  p-value: < 0.000000000000000022

## Exhaustive Search
library(leaps)
exh.search <- regsubsets(ProsperScore~., data = train.num.red, nbest=1, nvmax
=dim(train.num.red)[2], method="exhaustive")

## Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax,
## force.in = force.in, : 1 linear dependencies found

## Reordering variables and trying again:

sum <- summary(exh.search)
sum$which

##      (Intercept)  Term BorrowerAPR BorrowerRate LenderYield EstimatedLoss
## 1      TRUE FALSE      FALSE      FALSE      FALSE      FALSE
## 2      TRUE FALSE      FALSE      FALSE      FALSE      FALSE
## 3      TRUE FALSE      FALSE      FALSE      FALSE      FALSE
## 4      TRUE FALSE      FALSE      FALSE      FALSE      FALSE
## 5      TRUE FALSE      FALSE      FALSE      FALSE      FALSE
## 6      TRUE FALSE      FALSE      FALSE      FALSE      FALSE
## 7      TRUE FALSE      FALSE      FALSE      FALSE      FALSE
## 8      TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 9      TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 10     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 11     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 12     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 13     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 14     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 15     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 16     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 17     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 18     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 19     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 20     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 21     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 22     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 23     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 24     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 25     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 26     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 27     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 28     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 29     TRUE FALSE      TRUE      FALSE      TRUE      TRUE
## 30     TRUE  TRUE      TRUE      FALSE      TRUE      TRUE
## 31     TRUE  TRUE      TRUE      FALSE      TRUE      TRUE

```

## 32	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE
## 33	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## 34	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## 35	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
##	EstimatedReturn ProsperRating..numeric. CreditScoreRangeLower					
## 1	FALSE			TRUE		FALSE
## 2	FALSE			TRUE		FALSE
## 3	FALSE			TRUE		FALSE
## 4	FALSE			TRUE		TRUE
## 5	FALSE			TRUE		TRUE
## 6	FALSE			TRUE		TRUE
## 7	FALSE			TRUE		TRUE
## 8	TRUE			FALSE		TRUE
## 9	TRUE			TRUE		TRUE
## 10	TRUE			TRUE		TRUE
## 11	TRUE			TRUE		TRUE
## 12	TRUE			TRUE		TRUE
## 13	TRUE			TRUE		TRUE
## 14	TRUE			TRUE		TRUE
## 15	TRUE			TRUE		TRUE
## 16	TRUE			TRUE		TRUE
## 17	TRUE			TRUE		TRUE
## 18	TRUE			TRUE		TRUE
## 19	TRUE			TRUE		TRUE
## 20	TRUE			TRUE		TRUE
## 21	TRUE			TRUE		TRUE
## 22	TRUE			TRUE		TRUE
## 23	TRUE			TRUE		TRUE
## 24	TRUE			TRUE		TRUE
## 25	TRUE			TRUE		TRUE
## 26	TRUE			TRUE		TRUE
## 27	TRUE			TRUE		TRUE
## 28	TRUE			TRUE		TRUE
## 29	TRUE			TRUE		TRUE
## 30	TRUE			TRUE		TRUE
## 31	TRUE			TRUE		TRUE
## 32	TRUE			TRUE		TRUE
## 33	TRUE			TRUE		TRUE
## 34	TRUE			TRUE		TRUE
## 35	TRUE			TRUE		TRUE
## 36	TRUE			TRUE		TRUE
## 37	TRUE			TRUE		TRUE
## 38	TRUE			TRUE		TRUE
##	TotalCreditLinespast7years OpenRevolvingMonthlyPayment					
## 1			FALSE			FALSE
## 2			FALSE			FALSE
## 3			FALSE			FALSE

## 4		FALSE	FALSE
## 5		FALSE	FALSE
## 6		FALSE	FALSE
## 7		FALSE	FALSE
## 8		FALSE	FALSE
## 9		FALSE	FALSE
## 10		FALSE	FALSE
## 11		FALSE	FALSE
## 12		FALSE	FALSE
## 13		FALSE	FALSE
## 14		FALSE	FALSE
## 15		FALSE	FALSE
## 16		FALSE	FALSE
## 17		FALSE	FALSE
## 18		FALSE	FALSE
## 19		FALSE	FALSE
## 20		FALSE	FALSE
## 21		FALSE	FALSE
## 22		FALSE	FALSE
## 23		FALSE	FALSE
## 24		TRUE	FALSE
## 25		TRUE	FALSE
## 26		TRUE	FALSE
## 27		TRUE	FALSE
## 28		TRUE	FALSE
## 29		TRUE	FALSE
## 30		TRUE	FALSE
## 31		TRUE	FALSE
## 32		TRUE	FALSE
## 33		TRUE	FALSE
## 34		TRUE	FALSE
## 35		TRUE	FALSE
## 36		TRUE	FALSE
## 37		TRUE	TRUE
## 38		TRUE	TRUE
##	InquiriesLast6Months	TotalInquiries	CurrentDelinquencies
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	TRUE	FALSE	FALSE
## 4	TRUE	FALSE	FALSE
## 5	TRUE	FALSE	FALSE
## 6	TRUE	FALSE	FALSE
## 7	TRUE	FALSE	FALSE
## 8	TRUE	FALSE	FALSE
## 9	TRUE	FALSE	FALSE
## 10	TRUE	FALSE	FALSE
## 11	TRUE	FALSE	FALSE
## 12	TRUE	FALSE	FALSE
## 13	TRUE	FALSE	FALSE
## 14	TRUE	FALSE	FALSE

## 15	TRUE	FALSE	FALSE
## 16	TRUE	FALSE	FALSE
## 17	TRUE	FALSE	FALSE
## 18	TRUE	TRUE	FALSE
## 19	TRUE	TRUE	TRUE
## 20	TRUE	TRUE	TRUE
## 21	TRUE	TRUE	TRUE
## 22	TRUE	TRUE	TRUE
## 23	TRUE	TRUE	TRUE
## 24	TRUE	TRUE	TRUE
## 25	TRUE	TRUE	TRUE
## 26	TRUE	TRUE	TRUE
## 27	TRUE	TRUE	TRUE
## 28	TRUE	TRUE	TRUE
## 29	TRUE	TRUE	TRUE
## 30	TRUE	TRUE	TRUE
## 31	TRUE	TRUE	TRUE
## 32	TRUE	TRUE	TRUE
## 33	TRUE	TRUE	TRUE
## 34	TRUE	TRUE	TRUE
## 35	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE
##	DelinquenciesLast7Years	PublicRecordsLast10Years	RevolvingCreditBalance
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	FALSE	FALSE
## 6	FALSE	FALSE	FALSE
## 7	FALSE	FALSE	FALSE
## 8	FALSE	FALSE	FALSE
## 9	FALSE	FALSE	FALSE
## 10	FALSE	FALSE	FALSE
## 11	FALSE	FALSE	FALSE
## 12	FALSE	FALSE	FALSE
## 13	FALSE	FALSE	FALSE
## 14	FALSE	FALSE	FALSE
## 15	FALSE	FALSE	FALSE
## 16	FALSE	FALSE	FALSE
## 17	FALSE	FALSE	FALSE
## 18	FALSE	FALSE	FALSE
## 19	FALSE	FALSE	FALSE
## 20	FALSE	FALSE	FALSE
## 21	FALSE	FALSE	FALSE
## 22	FALSE	FALSE	FALSE
## 23	FALSE	TRUE	FALSE
## 24	FALSE	TRUE	FALSE
## 25	FALSE	TRUE	FALSE

## 26	TRUE	TRUE	FALSE
## 27	TRUE	TRUE	FALSE
## 28	TRUE	TRUE	FALSE
## 29	TRUE	TRUE	FALSE
## 30	TRUE	TRUE	FALSE
## 31	TRUE	TRUE	FALSE
## 32	TRUE	TRUE	TRUE
## 33	TRUE	TRUE	TRUE
## 34	TRUE	TRUE	TRUE
## 35	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE
##	BankcardUtilization	AvailableBankcardCredit	TotalTrades
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	TRUE	FALSE
## 6	FALSE	TRUE	FALSE
## 7	FALSE	TRUE	FALSE
## 8	FALSE	TRUE	FALSE
## 9	FALSE	TRUE	FALSE
## 10	FALSE	TRUE	FALSE
## 11	TRUE	TRUE	FALSE
## 12	FALSE	TRUE	FALSE
## 13	TRUE	TRUE	FALSE
## 14	TRUE	TRUE	FALSE
## 15	TRUE	TRUE	FALSE
## 16	TRUE	TRUE	FALSE
## 17	TRUE	TRUE	FALSE
## 18	TRUE	TRUE	FALSE
## 19	TRUE	TRUE	FALSE
## 20	TRUE	TRUE	FALSE
## 21	TRUE	TRUE	FALSE
## 22	TRUE	TRUE	FALSE
## 23	TRUE	TRUE	FALSE
## 24	TRUE	TRUE	FALSE
## 25	TRUE	TRUE	FALSE
## 26	TRUE	TRUE	FALSE
## 27	TRUE	TRUE	FALSE
## 28	TRUE	TRUE	FALSE
## 29	TRUE	TRUE	FALSE
## 30	TRUE	TRUE	FALSE
## 31	TRUE	TRUE	FALSE
## 32	TRUE	TRUE	FALSE
## 33	TRUE	TRUE	FALSE
## 34	TRUE	TRUE	FALSE
## 35	TRUE	TRUE	FALSE
## 36	TRUE	TRUE	FALSE

## 37	TRUE	TRUE	FALSE
## 38	TRUE	TRUE	TRUE
##	TradesNeverDelinquent..percentage.	TradesOpenedLast6Months	
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	FALSE	FALSE
## 6	FALSE	TRUE	TRUE
## 7	FALSE	TRUE	TRUE
## 8	FALSE	FALSE	FALSE
## 9	FALSE	FALSE	FALSE
## 10	FALSE	TRUE	TRUE
## 11	FALSE	TRUE	TRUE
## 12	FALSE	TRUE	TRUE
## 13	FALSE	TRUE	TRUE
## 14	FALSE	TRUE	TRUE
## 15	FALSE	TRUE	TRUE
## 16	FALSE	TRUE	TRUE
## 17	FALSE	TRUE	TRUE
## 18	FALSE	TRUE	TRUE
## 19	FALSE	TRUE	TRUE
## 20	TRUE	TRUE	TRUE
## 21	TRUE	TRUE	TRUE
## 22	TRUE	TRUE	TRUE
## 23	TRUE	TRUE	TRUE
## 24	TRUE	TRUE	TRUE
## 25	TRUE	TRUE	TRUE
## 26	TRUE	TRUE	TRUE
## 27	TRUE	TRUE	TRUE
## 28	TRUE	TRUE	TRUE
## 29	TRUE	TRUE	TRUE
## 30	TRUE	TRUE	TRUE
## 31	TRUE	TRUE	TRUE
## 32	TRUE	TRUE	TRUE
## 33	TRUE	TRUE	TRUE
## 34	TRUE	TRUE	TRUE
## 35	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE
##	LoanMonthsSinceOrigination	LoanNumber	LoanOriginalAmount
## 1	FALSE	FALSE	FALSE
## 2	TRUE	FALSE	FALSE
## 3	TRUE	FALSE	FALSE
## 4	TRUE	FALSE	FALSE
## 5	TRUE	FALSE	FALSE
## 6	TRUE	FALSE	FALSE
## 7	TRUE	FALSE	TRUE
## 8	TRUE	FALSE	FALSE

## 9	TRUE	FALSE	FALSE
## 10	TRUE	FALSE	FALSE
## 11	TRUE	FALSE	FALSE
## 12	TRUE	FALSE	TRUE
## 13	TRUE	FALSE	TRUE
## 14	TRUE	FALSE	TRUE
## 15	TRUE	TRUE	TRUE
## 16	FALSE	TRUE	TRUE
## 17	TRUE	TRUE	TRUE
## 18	TRUE	TRUE	TRUE
## 19	TRUE	TRUE	TRUE
## 20	TRUE	TRUE	TRUE
## 21	TRUE	TRUE	TRUE
## 22	TRUE	TRUE	TRUE
## 23	TRUE	TRUE	TRUE
## 24	TRUE	TRUE	TRUE
## 25	TRUE	TRUE	TRUE
## 26	TRUE	TRUE	TRUE
## 27	TRUE	TRUE	TRUE
## 28	TRUE	TRUE	TRUE
## 29	TRUE	TRUE	TRUE
## 30	TRUE	TRUE	TRUE
## 31	TRUE	TRUE	TRUE
## 32	TRUE	TRUE	TRUE
## 33	TRUE	TRUE	TRUE
## 34	TRUE	TRUE	TRUE
## 35	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE

##	LP_CustomerPayments	LP_CustomerPrincipalPayments	LP_InterestandFees
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	FALSE	FALSE
## 6	FALSE	FALSE	FALSE
## 7	FALSE	FALSE	FALSE
## 8	FALSE	FALSE	FALSE
## 9	FALSE	FALSE	FALSE
## 10	FALSE	FALSE	FALSE
## 11	FALSE	FALSE	FALSE
## 12	FALSE	FALSE	TRUE
## 13	FALSE	FALSE	TRUE
## 14	FALSE	FALSE	TRUE
## 15	FALSE	FALSE	TRUE
## 16	TRUE	FALSE	TRUE
## 17	TRUE	FALSE	TRUE
## 18	TRUE	FALSE	TRUE
## 19	TRUE	FALSE	TRUE

## 20	TRUE	TRUE	FALSE
## 21	FALSE	TRUE	TRUE
## 22	TRUE	FALSE	TRUE
## 23	TRUE	FALSE	TRUE
## 24	TRUE	TRUE	FALSE
## 25	FALSE	TRUE	TRUE
## 26	FALSE	TRUE	TRUE
## 27	TRUE	FALSE	TRUE
## 28	TRUE	TRUE	FALSE
## 29	TRUE	TRUE	FALSE
## 30	FALSE	TRUE	TRUE
## 31	FALSE	TRUE	TRUE
## 32	FALSE	TRUE	TRUE
## 33	FALSE	TRUE	TRUE
## 34	FALSE	TRUE	TRUE
## 35	FALSE	TRUE	TRUE
## 36	FALSE	TRUE	TRUE
## 37	TRUE	TRUE	FALSE
## 38	TRUE	TRUE	FALSE
##	LP_ServiceFees	LP_GrossPrincipalLoss	LP_NetPrincipalLoss
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	FALSE	FALSE
## 6	FALSE	FALSE	FALSE
## 7	FALSE	FALSE	FALSE
## 8	FALSE	FALSE	FALSE
## 9	FALSE	FALSE	FALSE
## 10	FALSE	FALSE	FALSE
## 11	FALSE	FALSE	FALSE
## 12	FALSE	FALSE	FALSE
## 13	FALSE	FALSE	FALSE
## 14	FALSE	FALSE	FALSE
## 15	FALSE	FALSE	FALSE
## 16	FALSE	TRUE	FALSE
## 17	FALSE	TRUE	FALSE
## 18	FALSE	TRUE	FALSE
## 19	FALSE	TRUE	FALSE
## 20	FALSE	TRUE	FALSE
## 21	FALSE	TRUE	FALSE
## 22	TRUE	TRUE	FALSE
## 23	TRUE	TRUE	FALSE
## 24	TRUE	TRUE	FALSE
## 25	TRUE	TRUE	FALSE
## 26	TRUE	TRUE	FALSE
## 27	TRUE	TRUE	FALSE
## 28	TRUE	TRUE	FALSE
## 29	TRUE	TRUE	FALSE
## 30	TRUE	TRUE	FALSE

## 31	TRUE	TRUE	FALSE
## 32	TRUE	TRUE	FALSE
## 33	TRUE	TRUE	FALSE
## 34	TRUE	TRUE	TRUE
## 35	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE
##	LP_NonPrincipalRecoverypayments	PercentFunded	Recommendations
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	FALSE	FALSE
## 6	FALSE	FALSE	FALSE
## 7	FALSE	FALSE	FALSE
## 8	FALSE	FALSE	FALSE
## 9	FALSE	FALSE	FALSE
## 10	FALSE	FALSE	FALSE
## 11	FALSE	FALSE	FALSE
## 12	FALSE	FALSE	FALSE
## 13	FALSE	FALSE	FALSE
## 14	FALSE	FALSE	FALSE
## 15	FALSE	FALSE	FALSE
## 16	FALSE	FALSE	FALSE
## 17	FALSE	FALSE	FALSE
## 18	FALSE	FALSE	FALSE
## 19	FALSE	FALSE	FALSE
## 20	FALSE	FALSE	FALSE
## 21	FALSE	FALSE	FALSE
## 22	FALSE	FALSE	FALSE
## 23	FALSE	FALSE	FALSE
## 24	FALSE	FALSE	FALSE
## 25	FALSE	TRUE	FALSE
## 26	FALSE	TRUE	FALSE
## 27	FALSE	TRUE	FALSE
## 28	TRUE	TRUE	FALSE
## 29	TRUE	TRUE	FALSE
## 30	TRUE	TRUE	FALSE
## 31	TRUE	TRUE	FALSE
## 32	TRUE	TRUE	FALSE
## 33	TRUE	TRUE	FALSE
## 34	TRUE	TRUE	FALSE
## 35	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE
##	InvestmentFromFriendsCount	InvestmentFromFriendsAmount	Investors
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE

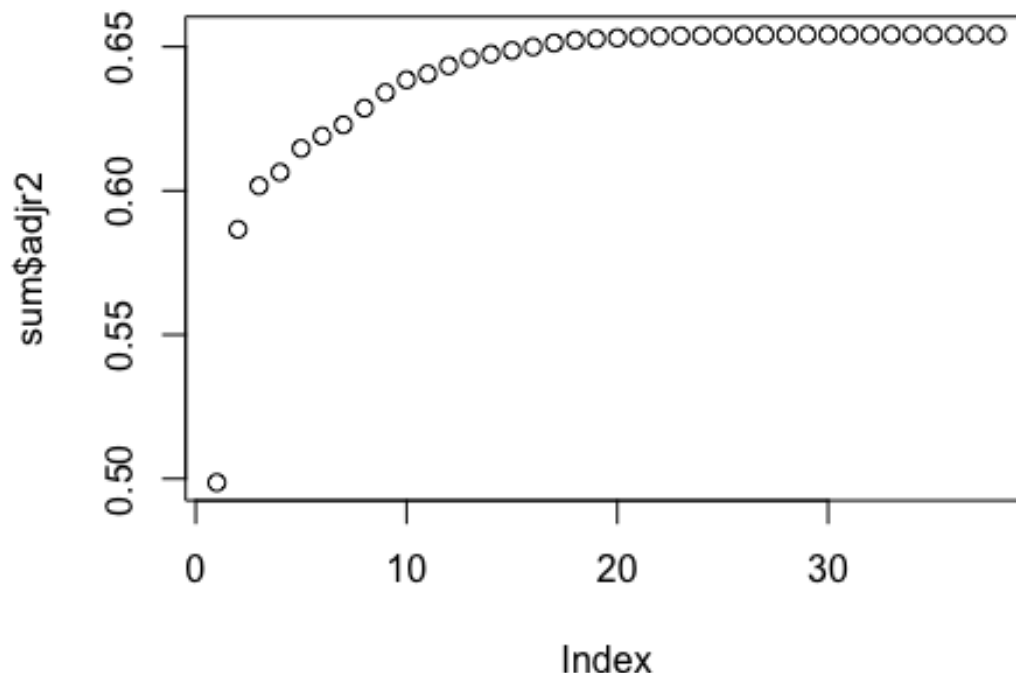
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	FALSE	FALSE
## 6	FALSE	FALSE	FALSE
## 7	FALSE	FALSE	FALSE
## 8	FALSE	FALSE	FALSE
## 9	FALSE	FALSE	FALSE
## 10	FALSE	FALSE	FALSE
## 11	FALSE	FALSE	FALSE
## 12	FALSE	FALSE	FALSE
## 13	FALSE	FALSE	FALSE
## 14	FALSE	FALSE	FALSE
## 15	FALSE	FALSE	FALSE
## 16	FALSE	FALSE	FALSE
## 17	FALSE	FALSE	FALSE
## 18	FALSE	FALSE	FALSE
## 19	FALSE	FALSE	FALSE
## 20	FALSE	FALSE	FALSE
## 21	FALSE	FALSE	TRUE
## 22	FALSE	FALSE	TRUE
## 23	FALSE	FALSE	TRUE
## 24	FALSE	FALSE	TRUE
## 25	FALSE	FALSE	TRUE
## 26	FALSE	FALSE	TRUE
## 27	TRUE	FALSE	TRUE
## 28	TRUE	FALSE	TRUE
## 29	TRUE	FALSE	TRUE
## 30	TRUE	FALSE	TRUE
## 31	TRUE	TRUE	TRUE
## 32	TRUE	TRUE	TRUE
## 33	TRUE	TRUE	TRUE
## 34	TRUE	TRUE	TRUE
## 35	TRUE	TRUE	TRUE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE
##	DebtToIncomeRatio.imp	EmploymentStatus.dum	LoanStatus.dum
## 1	FALSE	FALSE	FALSE
## 2	FALSE	FALSE	FALSE
## 3	FALSE	FALSE	FALSE
## 4	FALSE	FALSE	FALSE
## 5	FALSE	FALSE	FALSE
## 6	FALSE	FALSE	FALSE
## 7	FALSE	FALSE	FALSE
## 8	FALSE	FALSE	FALSE
## 9	FALSE	FALSE	FALSE
## 10	FALSE	FALSE	FALSE
## 11	FALSE	FALSE	FALSE
## 12	FALSE	FALSE	FALSE
## 13	FALSE	FALSE	FALSE

## 14	TRUE	FALSE	FALSE
## 15	TRUE	FALSE	FALSE
## 16	TRUE	FALSE	FALSE
## 17	TRUE	FALSE	FALSE
## 18	TRUE	FALSE	FALSE
## 19	TRUE	FALSE	FALSE
## 20	TRUE	FALSE	FALSE
## 21	TRUE	FALSE	FALSE
## 22	TRUE	FALSE	FALSE
## 23	TRUE	FALSE	FALSE
## 24	TRUE	FALSE	FALSE
## 25	TRUE	FALSE	FALSE
## 26	TRUE	FALSE	FALSE
## 27	TRUE	FALSE	FALSE
## 28	TRUE	FALSE	FALSE
## 29	TRUE	TRUE	FALSE
## 30	TRUE	TRUE	FALSE
## 31	TRUE	TRUE	FALSE
## 32	TRUE	TRUE	FALSE
## 33	TRUE	TRUE	FALSE
## 34	TRUE	TRUE	FALSE
## 35	TRUE	TRUE	FALSE
## 36	TRUE	TRUE	TRUE
## 37	TRUE	TRUE	TRUE
## 38	TRUE	TRUE	TRUE

```
sum$adjr2
```

```
## [1] 0.4986608 0.5865845 0.6017042 0.6063979 0.6147008 0.6190039 0.6229020
## [8] 0.6286804 0.6341050 0.6385178 0.6406092 0.6434122 0.6460152 0.6474665
## [15] 0.6486897 0.6499689 0.6513038 0.6522772 0.6527228 0.6530685 0.6533502
## [22] 0.6535770 0.6537805 0.6538959 0.6539892 0.6540780 0.6541571 0.6541777
## [29] 0.6541977 0.6542154 0.6542232 0.6542310 0.6542290 0.6542262 0.6542230
## [36] 0.6542196 0.6542149 0.6542101
```

```
plot(sum$adjr2)
```



```

### Model 10 seems to be the most important model
## Model 10 includes: BorrowerAPR, LenderYield, Estimated Loss, Estimated Return, Prosper Rating Num, Credit Score Range Lower, INquiries last 6 months, Available Bank Card Credit, Trades opened last 6 months, LoanMonthsSinceOrigination
### ANSWER!

##Compare Predictive Accuracy
## Stepwise
library(forecast)
pred_v_stepreg <- predict(stepreg.train, valid)
accuracy(pred_v_stepreg, valid$ProsperScore)

##               ME      RMSE      MAE      MPE      MAPE
## Test set -0.02484609 1.404636 1.095558 -10.26007 25.91077

## Exhaustive Model
exh.lm <- lm(ProsperScore~BorrowerAPR+LenderYield+EstimatedLoss+EstimatedReturn+ProsperRating.numeric.+CreditScoreRangeLower+InquiriesLast6Months+AvailableBankcardCredit+TradesOpenedLast6Months+LoanMonthsSinceOrigination, data=train)
pred_v_exh <- predict(exh.lm, valid)
accuracy(pred_v_exh, valid$ProsperScore)

```



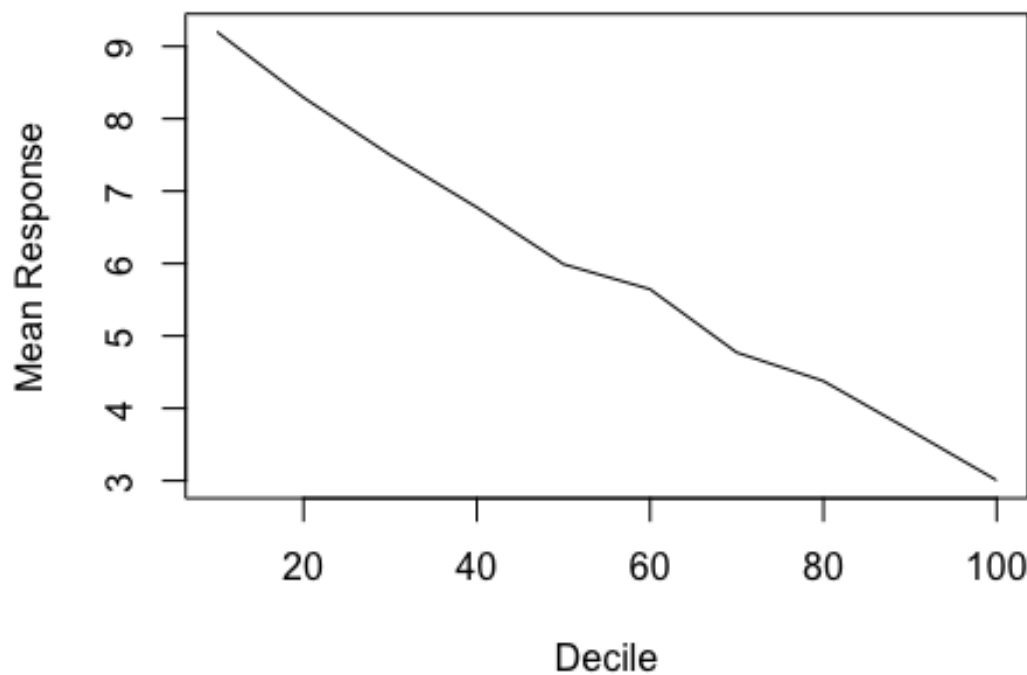
```
##               ME      RMSE      MAE      MPE      MAPE
## Test set -0.02918859 1.439297 1.124845 -10.8617 26.60153

pred_v_train.3 <- predict(train.lm.3,valid)
accuracy(pred_v_train.3,valid$ProsperScore)

##               ME      RMSE      MAE      MPE      MAPE
## Test set -0.0277709 1.510647 1.201994 -12.29348 28.77092

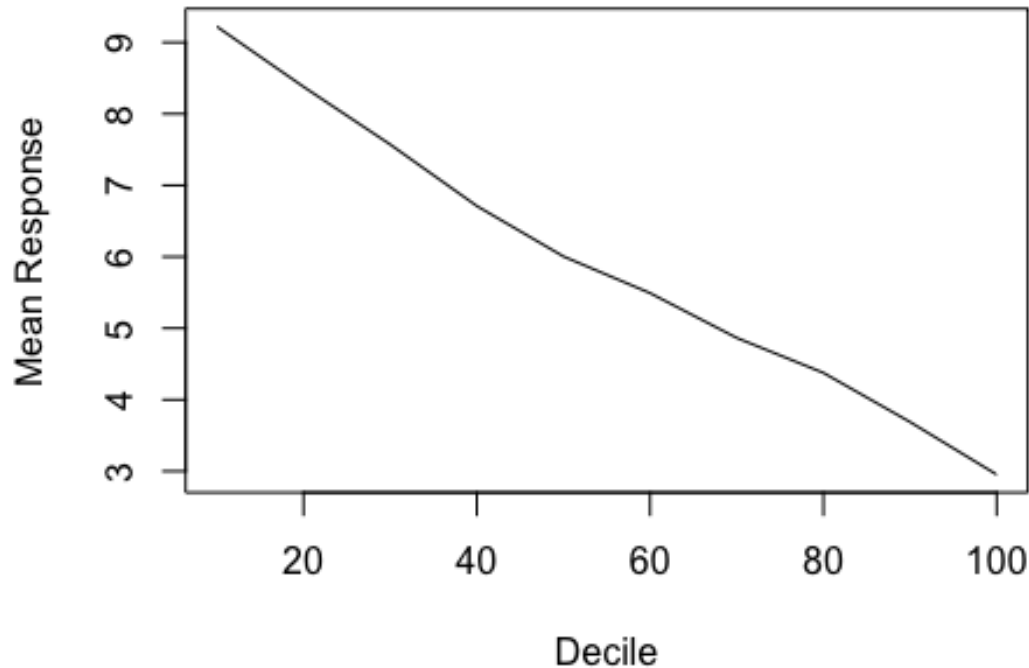
##Lift Charts - Exh
library(gains)
gain.num <- gains(valid$ProsperScore, pred_v_exh, groups=10)
plot(gain.num$depth, gain.num$mean.resp, xlab= "Decile", ylab = "Mean Response",
     main = "(Exhaustive) Decile vs. Mean Response",type="l")
```

(Exhaustive) Decile vs. Mean Response



```
##Lift Charts - Step
gain.num2 <- gains(valid$ProsperScore, pred_v_stepreg, groups=10)
plot(gain.num2$depth, gain.num2$mean.resp, xlab= "Decile", ylab = "Mean Response",
     main = "(Stepwise) Decile vs. Mean Response",type="l")
```

(Stepwise) Decile vs. Mean Response



LASSO REGRESSION

```
##set.seed(123)
##trin<-sample(1:nrow(train.num),) #80% train
#tst<-(-trin) #20% test

#library(glmnet)
#x <- model.matrix(ProsperScore~., trin)[, -9]
#head(x,5)
#y<- trin$ProsperScore
#set.seed(123)
#lasso.model<-cv.glmnet(x[trin,], y[trin], alpha=1)
#default is 10 fold cv
#alpha=1 is the option to do lasso regression
#lasso.pred<-predict(lasso.model, newx=x[tst,])
#lasso.pred.df <- as.data.frame(lasso.pred)
#lasso.pred.df$actual.rating <-cereals.complete[test,15]#column 15 is rating

#MSE<- mean((lasso.pred.df$`1` - lasso.pred.df$actual.rating)^2)
#RMSE<-sqrt(MSE)

#print(MSE)
#print(RMSE)
```

```

##### CLASSIFICATION PREDICTIONS #####
#KNN
## Scaling with Z-Standardization
#train.z <- as.data.frame(scale(train))

#train.num.2 <- train[,c(-2,-11,-12,-13,-14,-15,-34,-56,-57,-58)]
#train.num.z <- as.data.frame(scale(train.num.2))
#train.binary <- train[,c(55,57)]
#train.knn.total <- cbind(train.num.z, train.binary)

#library(class)
#train.knn.predictors <- train[,-57]

#train.knn.target <- valid[,57]

#valid.num <- valid[,c(-2,-11,-12,-13,-14,-15,-34,-56,-57,-58)]
#valid.num.z <- as.data.frame(scale(valid.num))
#valid.binary <- valid[,c(56,58)]
#valid.knn.predictors <- cbind(valid.num.z, valid.binary)

#valid.knn.target <- valid[,57]

#set.seed(123)
#preds <- knn(train=train.knn.predictors, test = valid.knn.predictors,
#            cl=train.knn.target, k=1, prob=TRUE)

#CONF_MATRIX<-table(preds,valid.knn.target)
#CONF_MATRIX

##### Logistic Regression #####
### Logit Model All Variables
train.log <- train[,c(3,5,7,8,9,10,16,21,30,33,37,56,57)]
LOGIT_MODEL <- glm(LoanStatus.dum~., family=binomial(), data=train.log)
summary(LOGIT_MODEL)

##
## Call:
## glm(formula = LoanStatus.dum ~ ., family = binomial(), data = train.log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.2413   0.1690   0.2824   0.4590   1.4379
##
## Coefficients:
##              Estimate      Std. Error z value
## (Intercept)  1.6206318641    0.4164469481    3.892
## BorrowerAPR  0.0524912708    1.6341860230    0.032
## LenderYield  9.9998701822    4.3202647515    2.315

```

```

## EstimatedLoss          -18.0998089762    4.6813610394   -3.866
## EstimatedReturn        -22.7398022847    4.4202889971   -5.144
## ProsperRating..numeric. -0.0492669694    0.0396531450   -1.242
## ProsperScore           0.0714106234    0.0100697475    7.092
## CreditScoreRangeLower   0.0002475912    0.0003641760    0.680
## InquiriesLast6Months   -0.0254202076    0.0091713746   -2.772
## AvailableBankcardCredit -0.0000035302    0.0000009141   -3.862
## TradesOpenedLast6Months -0.0718578694    0.0131978318   -5.445
## LoanNumber             0.0000297785    0.0000009827   30.303
## IsBorrowerHomeowner.dum 0.1708386039    0.0287277905    5.947
##                               Pr(>|z|)
## (Intercept)             0.00009959822236 ***
## BorrowerAPR              0.974376
## LenderYield              0.020632 *
## EstimatedLoss            0.000110 ***
## EstimatedReturn          0.00000026835531 ***
## ProsperRating..numeric.  0.214071
## ProsperScore             0.00000000000133 ***
## CreditScoreRangeLower    0.496589
## InquiriesLast6Months     0.005577 **
## AvailableBankcardCredit  0.000112 ***
## TradesOpenedLast6Months  0.00000005190084 ***
## LoanNumber               < 0.0000000000000002 ***
## IsBorrowerHomeowner.dum  0.00000000273425 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 43821  on 67881  degrees of freedom
## Residual deviance: 36264  on 67869  degrees of freedom
## AIC: 36290
##
## Number of Fisher Scoring iterations: 6

library(Discriminer)
logLik(LOGIT_MODEL)

## 'log Lik.' -18132.03 (df=13)

pred.LR<-predict(LOGIT_MODEL,valid)
LOGITS <- data.frame(pred.LR)
ODDS <- exp(LOGITS)
PROBABILITIES <- ODDS/(ODDS+1)
library(psych)
describe(PROBABILITIES)

##      vars      n mean  sd median trimmed  mad min max range  skew
## pred.LR      1 16971  0.9 0.11   0.94   0.92 0.06 0.3   1   0.7 -1.57
##      kurtosis se
## pred.LR      1.93 0

```

```

SC_PROB <- data.frame(fitted(LOGIT_MODEL))
PREDICTIONS <- ifelse(PROBABILITIES>.50,1,0)
invisible(data.frame(PREDICTIONS))
table(PREDICTIONS,valid$LoanStatus.dum)

##
## PREDICTIONS      0      1
##           0    20    22
##           1  1671 15258

##ROC Curve Reduced Model
library(ROCR)

## Loading required package: gplots

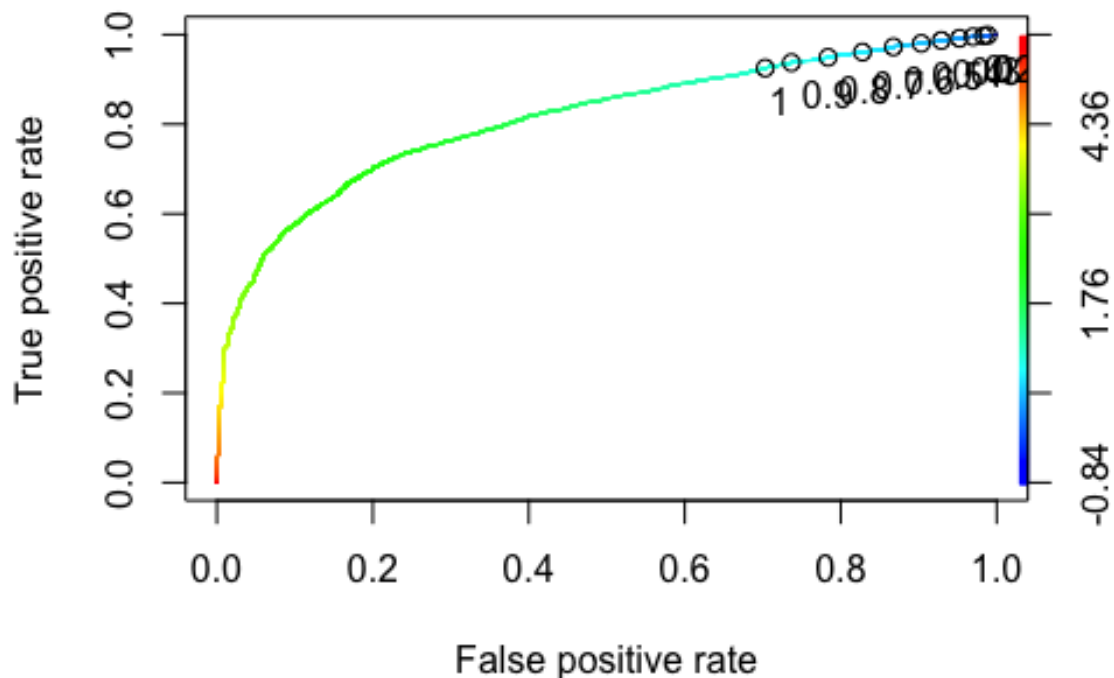
##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##      lowess

pred_logit <- prediction(pred_LR, valid$LoanStatus.dum)
perf_logit <- performance(pred_logit, "tpr", "fpr")
plot(perf_logit, colorize=TRUE, print.cutoffs.at=seq(0,1,by=0.1), text.adj=c(
-0.2,1.7),
      main = "ROC Curve for Logistic Regression Model")

```

ROC Curve for Logistic Regression Model



```
#### Logit Model 2, Reduced
### Rid of: BorrowerAPR, Prosper Rating, Credit SCore Range Lower
train.log <- train[,c(5,7,8,10,21,30,33,37,56,57)]
LOGIT_MODEL <- glm(LoanStatus.dum~., family=binomial(), data=train.log)
summary(LOGIT_MODEL)

##
## Call:
## glm(formula = LoanStatus.dum ~ ., family = binomial(), data = train.log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.2500   0.1692   0.2822   0.4589   1.4394
##
## Coefficients:
##              Estimate      Std. Error z value
## (Intercept)    1.4824637979    0.1543142088    9.607
## LenderYield    11.8420991986    4.0785804907    2.903
## EstimatedLoss  -18.8180278065    4.5521812065   -4.134
## EstimatedReturn -24.1580259683    4.1832642571   -5.775
## ProsperScore    0.0687693177    0.0098645694    6.971
## InquiriesLast6Months -0.0256676121    0.0091639484   -2.801
## AvailableBankcardCredit -0.0000034751    0.0000008733   -3.979
```

```

## TradesOpenedLast6Months -0.0719394789 0.0131906656 -5.454
## LoanNumber 0.0000296473 0.0000009654 30.710
## IsBorrowerHomeowner.dum 0.1735619042 0.0278847529 6.224
##
## Pr(>|z|)
## (Intercept) < 0.0000000000000002 ***
## LenderYield 0.00369 **
## EstimatedLoss 0.00003567385599 ***
## EstimatedReturn 0.00000000769886 ***
## ProsperScore 0.00000000000314 ***
## InquiriesLast6Months 0.00510 **
## AvailableBankcardCredit 0.00006913805815 ***
## TradesOpenedLast6Months 0.00000004929988 ***
## LoanNumber < 0.0000000000000002 ***
## IsBorrowerHomeowner.dum 0.0000000048384 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 43821 on 67881 degrees of freedom
## Residual deviance: 36266 on 67872 degrees of freedom
## AIC: 36286
##
## Number of Fisher Scoring iterations: 6

library(Discriminer)
logLik(LOGIT_MODEL)

## 'log Lik.' -18133.06 (df=10)

pred.LR<-predict(LOGIT_MODEL,valid)
LOGITS <- data.frame(pred.LR)
ODDS <- exp(LOGITS)
PROBABILITIES <- ODDS/(ODDS+1)
library(psych)
describe(PROBABILITIES)

## vars n mean sd median trimmed mad min max range skew
## pred.LR 1 16971 0.9 0.11 0.94 0.92 0.06 0.31 1 0.69 -1.55
## kurtosis se
## pred.LR 1.82 0

SC_PROB <- data.frame(fitted(LOGIT_MODEL))
PREDICTIONS2 <- ifelse(PROBABILITIES>.50,1,0)
invisible(data.frame(PREDICTIONS2))
table(PREDICTIONS2,valid$LoanStatus.dum)

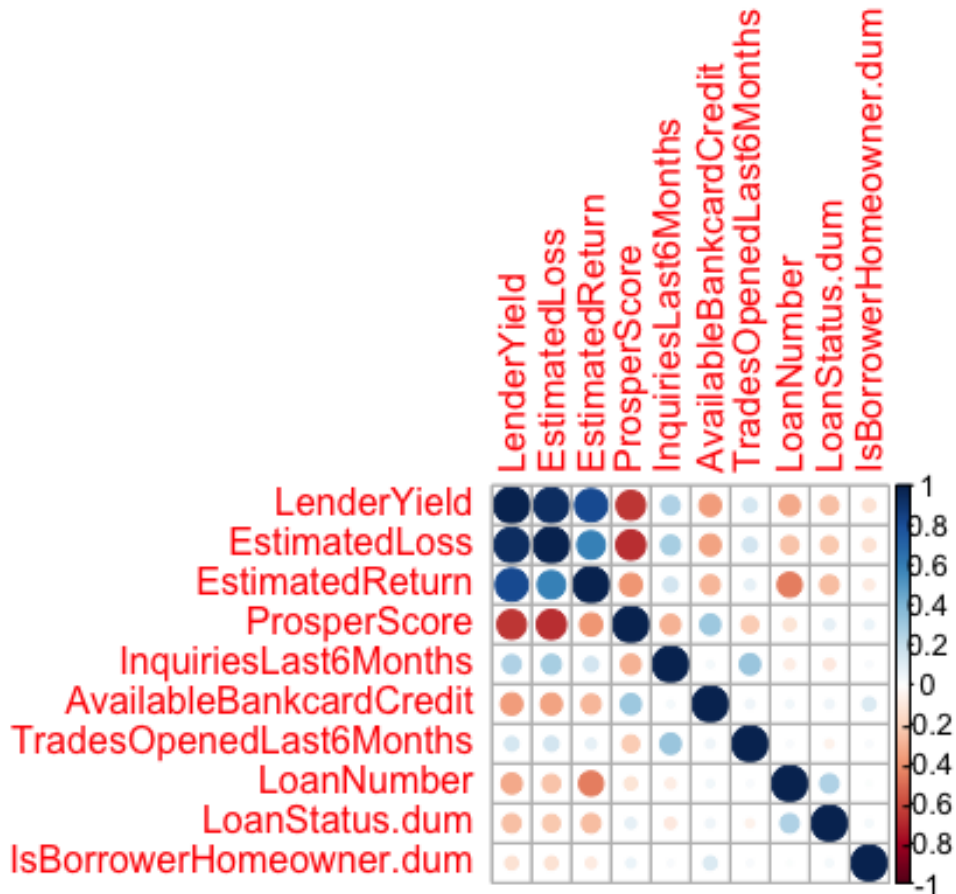
##
## PREDICTIONS2 0 1
## 0 16 21
## 1 1675 15259

```

```
#### Linear Discriminant Analysis
library(Discriminer)
## Running LDA gave error, check for Multicollinearity
library(corrplot)

## corrplot 0.84 loaded

corrs <- cor(train.log)
corrs.matrix <- as.matrix(corrs)
corrplot(corrs.matrix)
```



```
### Multicollinearity exists: get rid of Estimated Loss, Estimated Return
##train.lda <- train.log[,c(-2,-3)]
##train.lda <- train.lda[,c(-2)]
##LDA_MODEL <- linDA((train.lda),train.lda$LoanStatus.dum)
##summary(LDA_MODEL)
```

```
#####Try Naive Bayes #####
library(e1071)
##Output must be factor
train.nb <- train[,c(5,7,8,10,21,30,33,37,56,57)]
train.nb$LoanStatus.dum <- as.factor(train.nb$LoanStatus.dum)
```



```

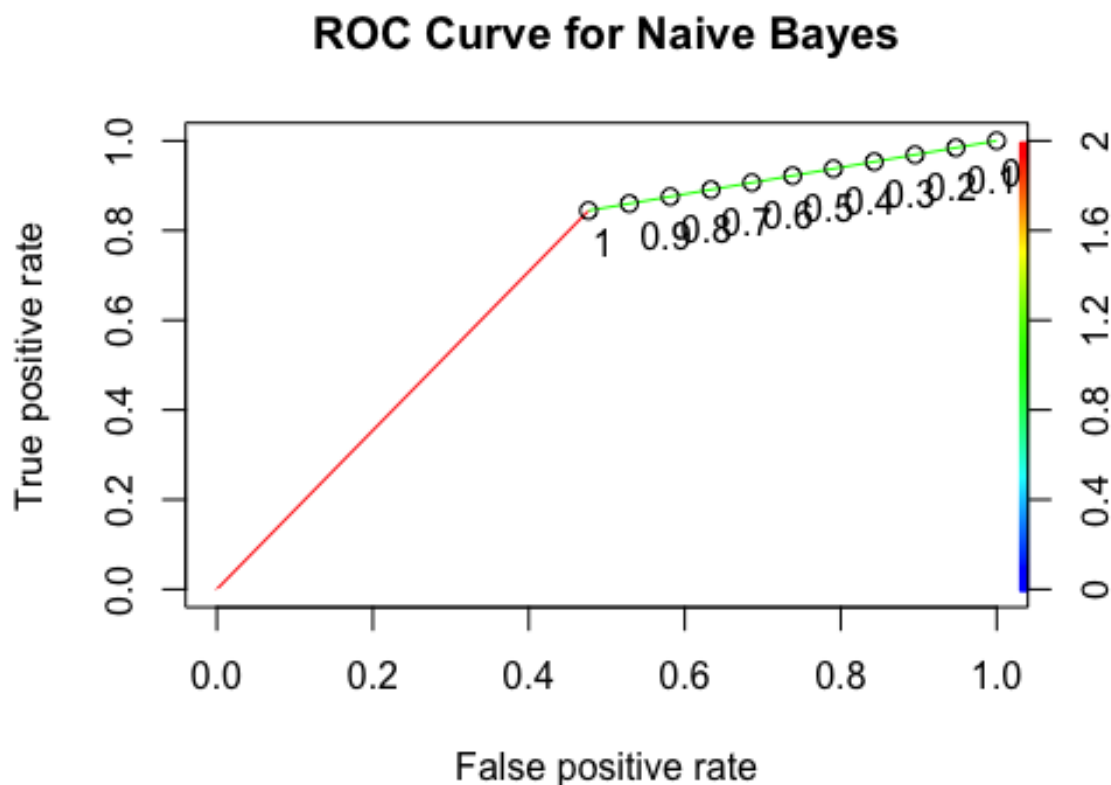
Status_classifier <- naiveBayes(LoanStatus.dum~., data = train.nb)
Status.pred <- predict(Status_classifier, valid[,c(-56)])

CONF_MATRIX.nb <- table(Status.pred,valid$LoanStatus.dum)

PROB.nb <- ifelse(Status.pred == 1, 1, 0)
library(ROCR)
pred.nb <- prediction(PROB.nb, valid[,56])
perf.nb <- performance(pred.nb, "tpr", "fpr")

plot(perf.nb, colorize=TRUE, print.cutoffs.at=seq(0,1,by=0.1), text.adj=c(-0.
2,1.7),
      main = "ROC Curve for Naive Bayes")

```



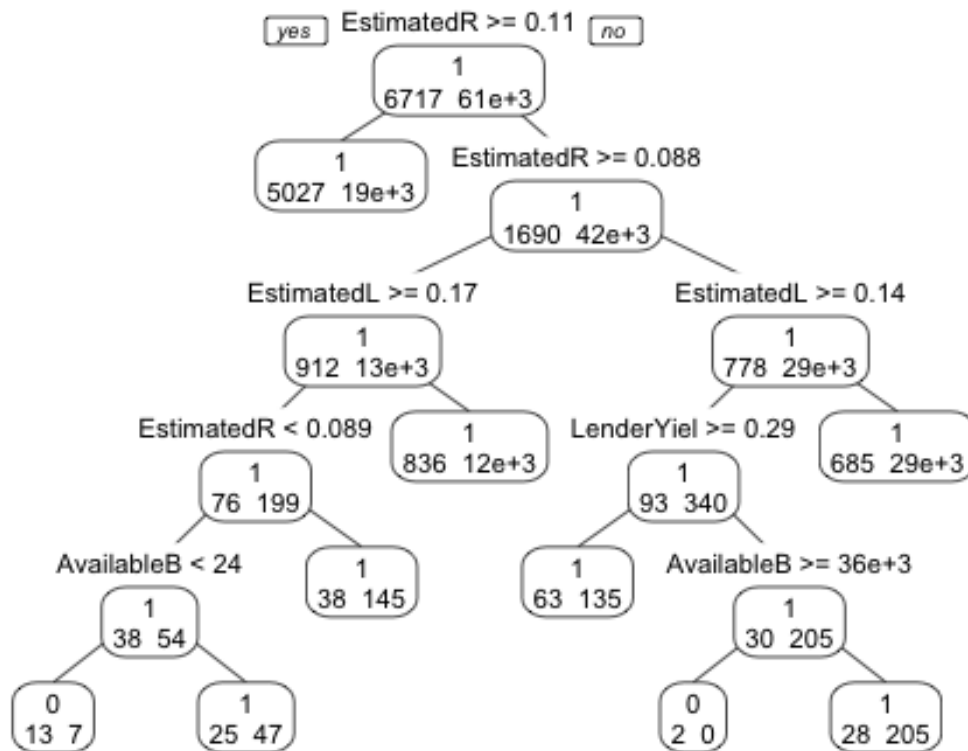
```

##### Decision Tree Model
library(rpart)
library(rpart.plot)
#Decision Tree Models
train.nb.1 <- train.nb[,c(-8)]
data_rpart_1 <- rpart(train.nb.1$LoanStatus.dum~., data=train.nb.1,method="class",
parms=list(split="information"), control=rpart.control(minsplit = 1, max

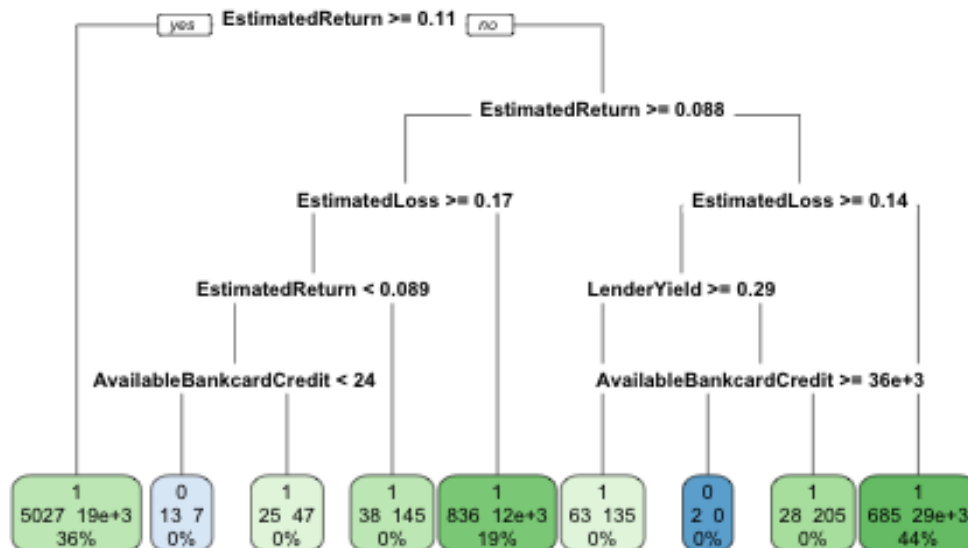
```

```
depth = 5, cp=0.00001))
```

```
prp(data_rpart_1, type=1, extra=1, split.font=1, varlen = -10)
```



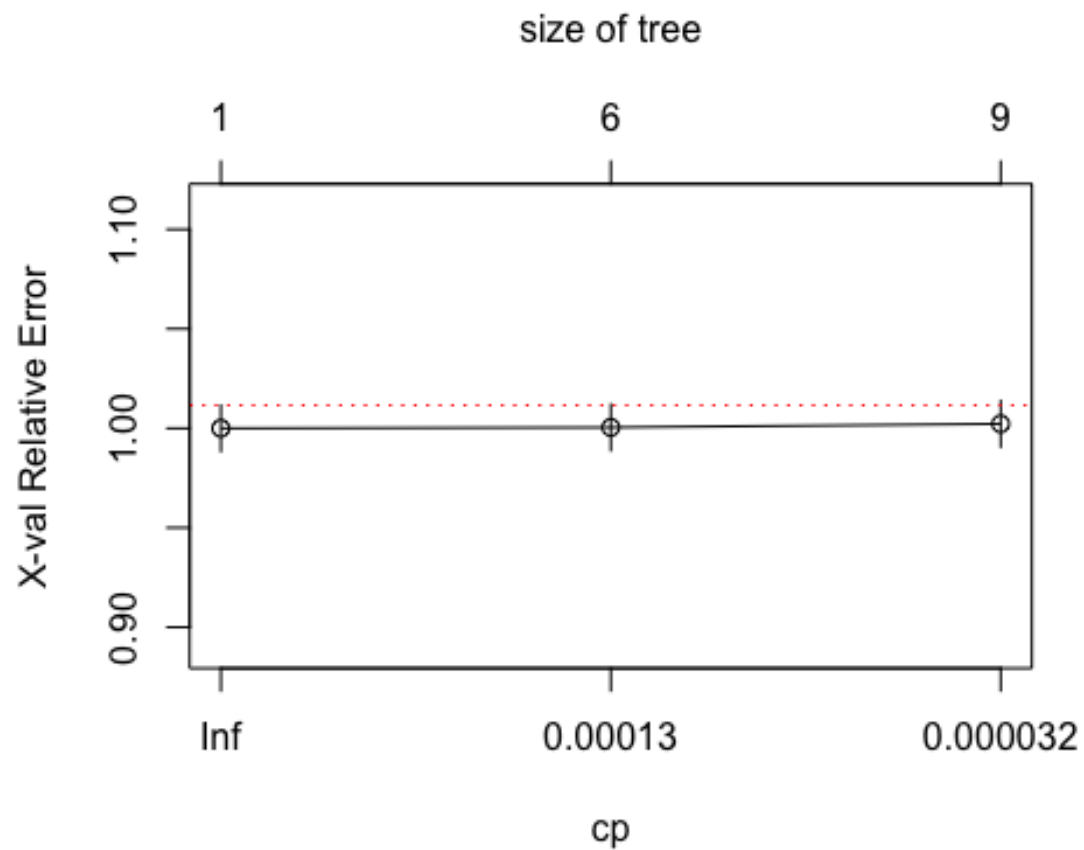
```
rpart.plot(data_rpart_1, type=0, extra=101)
```



```
##CP Values Reporting
cptable<-printcp(data_rpart_1)

##
## Classification tree:
## rpart(formula = train.nb.1$LoanStatus.dum ~ ., data = train.nb.1,
##       method = "class", parms = list(split = "information"), control = rpart
##       .control(minsplit = 1,
##       maxdepth = 5, cp = 0.00001))
##
## Variables actually used in tree construction:
## [1] AvailableBankcardCredit EstimatedLoss           EstimatedReturn
## [4] LenderYield
##
## Root node error: 6717/67882 = 0.098951
##
## n= 67882
##
##      CP nsplit rel error xerror   xstd
## 1 0.000178651      0  1.00000 1.0000 0.011582
## 2 0.000099251      5  0.99911 1.0004 0.011584
## 3 0.000010000      8  0.99881 1.0022 0.011594
```

```
plotcp(data_rpart_1, minline=TRUE, col="red")
```



```
rpart_pred_conf1 <- predict(data_rpart_1, valid, type="class")
CONF_MX1 <- table(rpart_pred_conf1, valid$LoanStatus.dum)
CONF_MX1
```

```
##
## rpart_pred_conf1      0      1
##           0      1      2
##           1 1690 15278
```