****

**Loan Defaulter Prediction**

**Presented By:**

Apoorva Prasad, Kumari Juhi, Manishi Kumari,

Prachi Jain, Pranav Handa

**Mentored By:**

Jennifer Eigo

University of Connecticut – OPIM 5604

Predictive Modeling

CONTENTS

[Executive Summary……………………………………….………………………………………3](#_Toc12244)

[Introduction to Problem Statement……………………………………………………………… 4](#_Toc12245)

[Methodology………………………………………………………………………………………..4](#_Toc12249)

[Sample……………………………………………………………………………………………...4](#_Toc12250)

[Explore……………………………………………………………………………………………...4](#_Toc12251)

Modify…………………………………………………………………………………………........[9](#_Toc12252)

[Model….…………………………………………………………………………………………..10](#_Toc12253)

Bootstrap.………………………………………………………..………………………………..10

Boosted Tree……………………………………………………………………………………..10

Logistic Regression………………………………..…………………………………………….10

[Result……………………………………………………………………………………………...11](#_Toc12254)

[Conclusions & Recommendations……………………………………………………………..12](#_Toc12257)

[References………………………………………………………………………………………..13](#_Toc12258)

[Appendix 1………………………………………………………………………………………..19](#_Toc12259)

Appendix 2………………………………………………………………………………………..21

**EXECUTIVE SUMMARY:**

Lending Club offers a peer-to-peer lending platform which is transparent and customer-friendly alternative to the traditional banking system. It provides unsecured loan at lower rate of interest to borrowers with an opportunity for lenders to choose their borrowers based on credit history and other information provided. This paper discusses the methodology and approach to devise a classification model to identify potential loan defaulters and assess the features of risky loans. This would help in making key decisions by both the investors and borrowers.

An initial exploration of several features revealed some key facts about the nature of borrowers and loan profile. Lending Club potentially determines interest rates based on grade which is calculated using credit history and other factors. Grades through “A” to “G” have increasing interest rate with increased loan defaulters. Most of the borrowers belong to economically major cities like California, NY, Florida and Texas, accounting for 38.2% of total loan borrowers which aligns with the fact that these are high GDP states. Further analysis shows that most of the borrowers have taken loan for credit card payment and debt consolidation. There are very few borrowers who have taken loan for a major purchase, home improvement, etc.

After performing data treatment for missing values, three different models - Logistic Regression, Bootstrap algorithm and Boosted Trees algorithm, were built to identify the defaulters. Two different sampling techniques – random sampling and stratified sampling were used to improve model performance. Out of these models, stratified sampling on Logistic regression yielded the best result with a total accuracy of 64.96%, precision of 37.31%, sensitivity of 68.29% and Receiver Operating Characteristics (ROC) of 72.22%.

Key features based on logistic regression model are loan term (term), sub grade (Sub\_grade), employee tenure (emp\_length), state (addr\_state), loan amount (loan\_amnt), number of currently active revolving trades (Num\_actv\_rev\_tl) and type of home ownership (home\_ownership).

**INTRODUCTION TO PROBLEM STATEMENT:**

Lending Club is a peer-to-peer lending company which brings borrowers and private investors together. It enables borrowers to take unsecured personal loans between $1,000 to $40,000. The club uses risk assessment model to grade borrower’s case based on credit history, credit score and desired loan amount. This grading helps investors to choose borrowers they would like to fund and determine the payable interest as well as repayment plan.

The purpose of this paper is to examine, analyze and evaluate the key metrics involved to predict loan defaulters by building a classification model. The dataset includes detailed information for every loan issued by Lending Club from 2015 to 2018, including a borrower’s annual income, zip codes, revolving balances, and purpose of loan. Among a total of 145 variables, 121 are numerical and 24 are categorical in nature. The initial dataset has 0.4 million records for each year. Based on historical records of the borrowers and loan details, the main motive is to find out the probability of a borrower to default.

**METHODOLOGY:**

SEMMA Approach:

Sample:

The current loan dataset consists of loan records from 2015-2018 (~1.7M records). Loan status (dependent variable) has 7 categories - Charged Off, Current, Default, Fully Paid, In Grace Period, Late (16-30 days) and Late (31-120 days). We have considered Late (16-30 days), Late (31-120 days), Default and Charged Off as defaulted loans and Fully Paid as a desirable loan. Current and In Grace Period loans have been removed as these are ongoing loans whose final status is not known as of now. Hence, the current number of records in dataset reduces to ~0.8MM after removing loans with Current and In Grace Period records. Base ratio of defaulters, non-defaulters in the data is 74:26 which makes the data a little skewed.

Explore:

Missing Data Analysis:

The Missing Value Report of all the variables shows 49 variables having over 50% missing records. These 49 variables were removed from the data. Further, 9 more variables were having 35% (0.3 MM) missing records on which imputation would have led to highly impure data and removal of records would have led to significant loss of data.

Hence, these 9 variables were also removed. Referring the data dictionary helped to find variables like employee title, zip code, loan funding, start date, etc. to be irrelevant. Similarly, variables like last payment date (last\_payment\_d), payment received to date for total amount funded (total\_pymnt), payment received to date for portion of total amount funded (total\_pymnt\_inv), last payment amount received (last\_pymnt\_amnt), interest received to date (total\_rec\_int) etc. were found to be insignificant. Such variables have been dropped from the data-set. Variables like total\_pymnt, total\_rec\_late\_fee, etc are relevant only after loan has been issued, but the prediction of a defaulted loan must be made before the loan is issued. Hence, these variables were also dropped from the data.

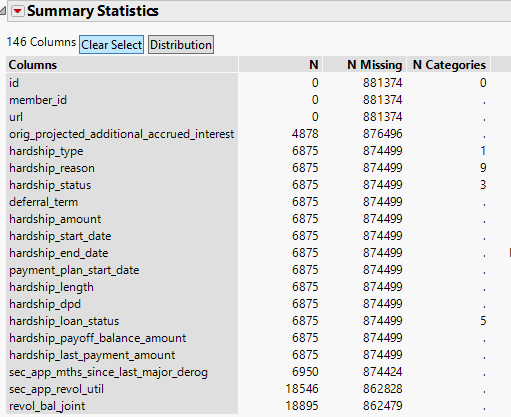


Fig.a: Sample variables with more than 50% missing data

Correlation Matrix:

Correlation matrix (Figure 1: Correlation Matrix) between remaining 81 numerical variables was studied to remove highly correlated variables from data. Threshold for highly correlated variables was kept at ±0.75. For example, instalment was highly correlated with loan\_amnt and hence, was dropped. Similarly, a total of 41 numerical variables were dropped in this manner.

*Variable Exploration by Loan Status:*

There were several interesting information and patterns that were identified by analysing different categorical and continuous variables:

*Grade vs Loan Status:*

Grades from A to G are given to each loan application based on the borrower’s credit score and other customer KPIs. The grades get worse while moving from A to G. As observed from the above graph, proportion of loan defaulters increases across grades from A to G. This is expected as borrowers with a bad credit score are more likely to default as compared to those with good grades.

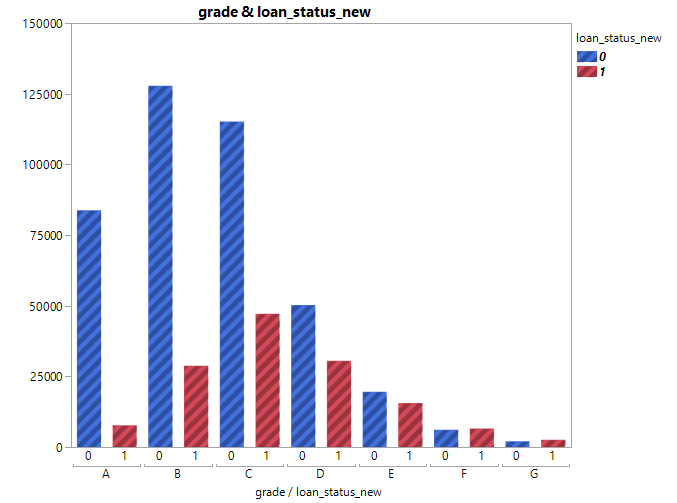


Fig.b: Distribution of Grade by Loan Status

*Home Ownership:*

Most of the borrowers are living in rented houses or have mortgages. Graph shows that borrowers with home on rent or mortgage are likely to default more as compared to those who own a house.

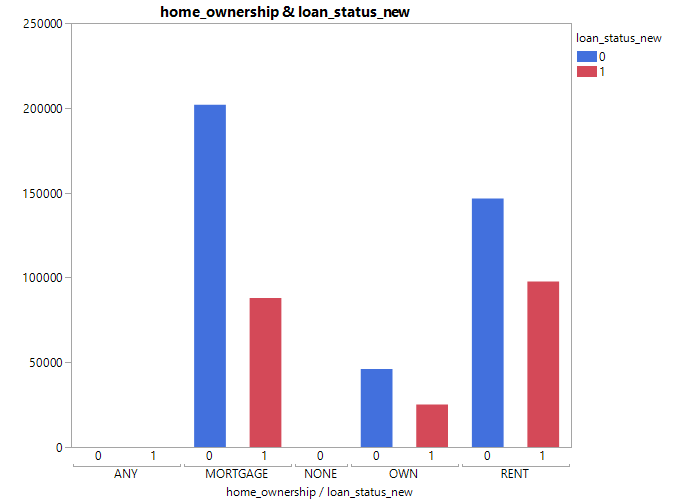


Fig.c: Distribution of Home Ownership by Loan Status

*Purpose:*

As per the above graph, borrowers with purpose of taking loan to consolidate debt are more likely to default than borrowers who took loan for other purposes.

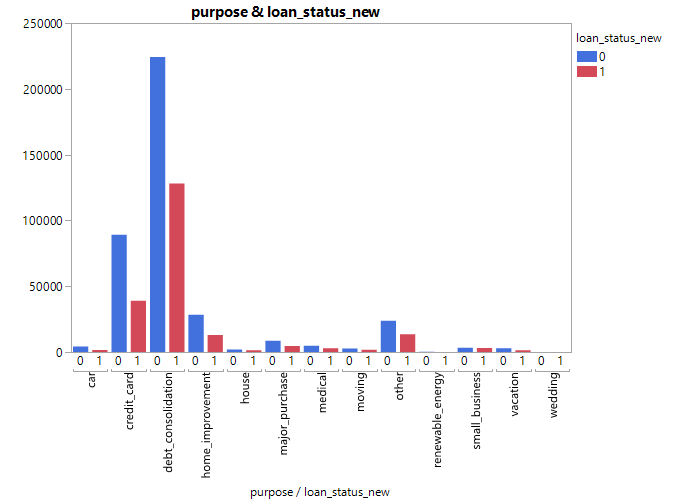


Fig.d: Distribution of Purpose by Loan Status

*Address State:*

38.2% of the total borrowers are from California, Texas, New York, and Florida. This observation aligns with the higher economic output and GDP of these states.

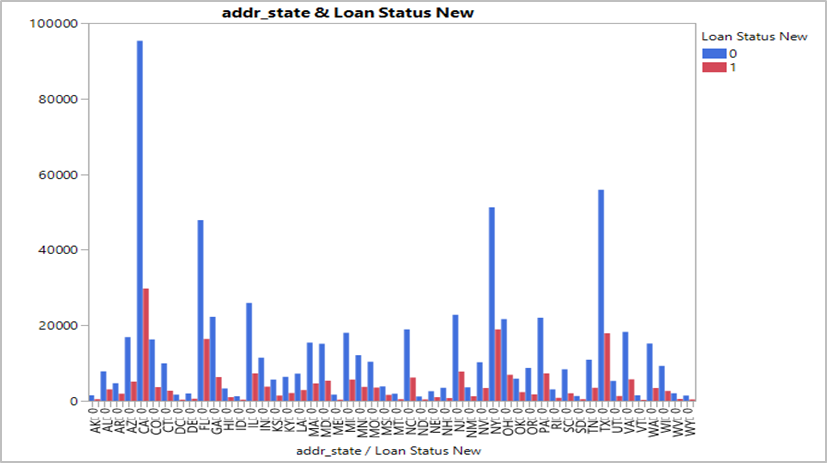


Fig.e: Distribution of Address State by Loan Status

*Debt Settlement:*

The observation from the graph aligns with the fact that borrowers working with any debt settlement company are likely to default more. From the graph, 3% of the total borrowers working with a debt settlement company are found to be 100% defaulters.

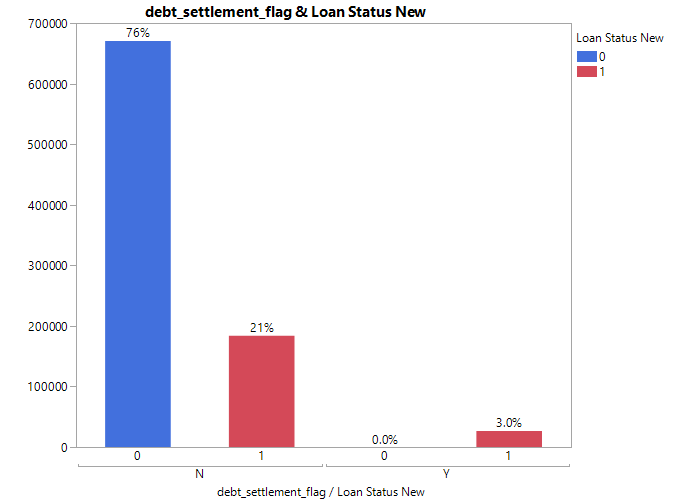


Fig.f: Distribution of Debt Settlement with Loan Status

*Hardship flag:*

Lending Club offers hardship plans to their borrowers by which the EMI payment starts only after 3 months of loan disbursement. It is interesting to observe that all ~0.1% (Figure 2: Distribution of Hardship Flag with Loan Status) of the total borrowers who took a hardship plan between 2015-2018 defaulted in loan repayment. This variable has been dropped from analysis due to biased observations.

*Verification Status:*

Lending Club captures information whether source of income of their borrowers is verified by their system or not. However, it is interesting to observe that borrowers whose source of income has been verified have higher proportion of defaulters as compared to borrowers whose source of income is not verified.

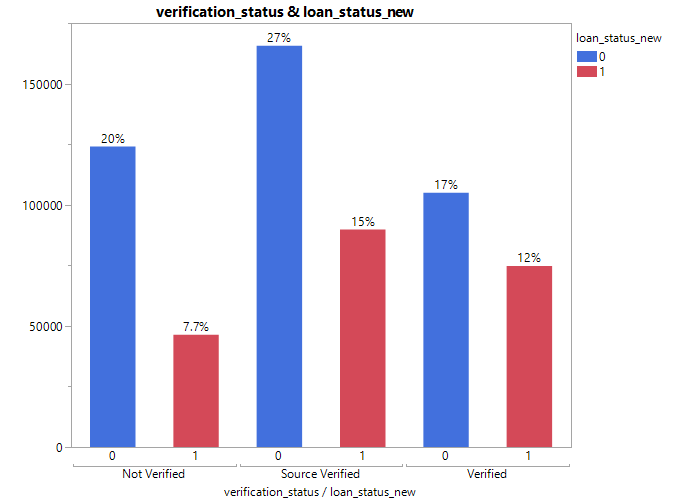


Fig.g: Distribution of Verification Status by Loan Status

*Interest Rate vs Grade vs Loan Status:*

Loans are classified by grade (A through G) and subgrade (1 through 5). These groups correspond to given interest rates, which are determined by the perceived risk of the borrowers. The graph shows expected default distribution for each grade and interest rate which aligns with the fact that bad loans come with higher interest rates.

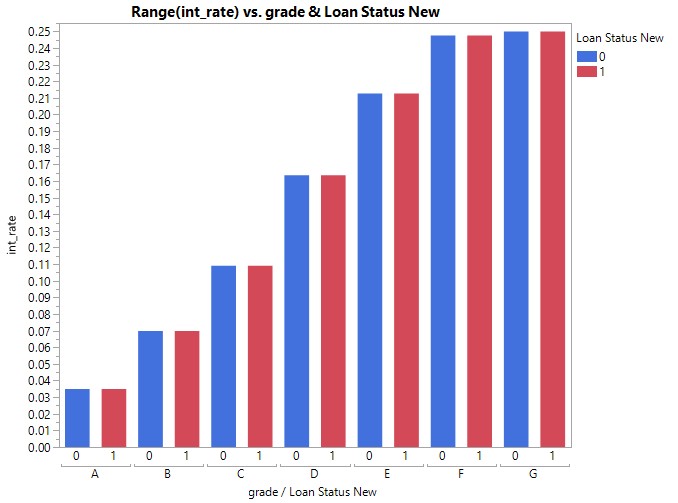


Fig.h: Distribution of Interest Rate by Grade and Loan Status

Modify:

All the irrelevant variables along with those having too many missing values have been dropped from data. Further removal of numerical variables has been done by multivariate analysis based on correlation matrix.

*Imputation:*

From the missing value report created and analysing missing data pattern, 41 columns were finalized for further analysis. After dropping variables with more than 35% (~0.3MM records) missing values, there were few variables which had maximum 2.8% of missing records. Those variables were imputed using normal imputation in JMP.

*Feature Engineering:*

Principal Component Analysis (PCA): It is a feature extraction technique used for dimensionality reduction. Input variables are combined to create new variables in such a way that least contributing variables are dropped while still retaining most of the variance in data.

On performing PCA, out of 31 numerical variables, 24 were selected for the model, retaining ~94% of information in data.



Fig.i: Eigen values of Principal Components

Model:

To enhance model accuracy, it should be validated through out-of-sample testing. Data has been divided into in-sample data (data used to develop the model – training - 60%, validation - 20%) and out-of-sample (model testing data – 20%). There were four different classification models built to predict loan defaulters -

1. Logistic Regression: Logistic Regression predicts a discrete outcome (yes or no) from a set of categorical/ continuous variables. Based on model results for random and stratified data, best model was built on stratified data having base distribution of 1’s and 0’s as 74:26. (Figure 3: Logistic Regression Model Summary (Random)).
2. Bootstrap Tree: This technique uses sampling with replacement method to build each independent tree. This reduces the chance of over-fitting data and is robust to outliers in data. (Figure 4: Bootstrap Model Summary for (Stratified and Random))
3. Boosted Tree: Similar to bootstrap algorithm, this technique samples data with replacement but each successive tree is created to reduce error in predecessor tree. (Figure 5: Boosted Tree Model Summary (Stratified and Random))
4. Logistic Regression using PCAs: Upon using the selected 24 principal components along with remaining categorical variables in a logistic model, comparable results were obtained with reduced model complexity. (Figure 6: Logistic Regression Using Principal Component Analysis (Stratified and Random))

**RESULTS:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| S.No. | Data Sample | Method | Overall Accuracy | Precision | ROC | Sensitivity |
| 1. | Random | Bootstrap | 76.37% | 64.20% | 70.83% | 2.93% |
| 2. | Random | Boosted | 76.56% | 62.26% | 71.15% | 4.14% |
| 3. | Random | Logistic | 76.94% | 57.23% | 72.28% | 14.19% |
| 4. | Random | PCA - Logistic | 76.74% | 57.12% | 71.63% | 12.77% |
| 5. | Stratified | Bootstrap | 61.36% | 34.77% | 70.27% | 70.24% |
| 6. | Stratified | Boosted | 63.70% | 36.18% | 71.23% | 68.08% |
| 7. | Stratified | Logistic | 64.96% | 37.31% | 72.22% | 68.29% |
| 8. | Stratified | PCA - Logistic | 24.50% | 24.05% | 71.75% | 99.95% |

Table 1: Model results for test data

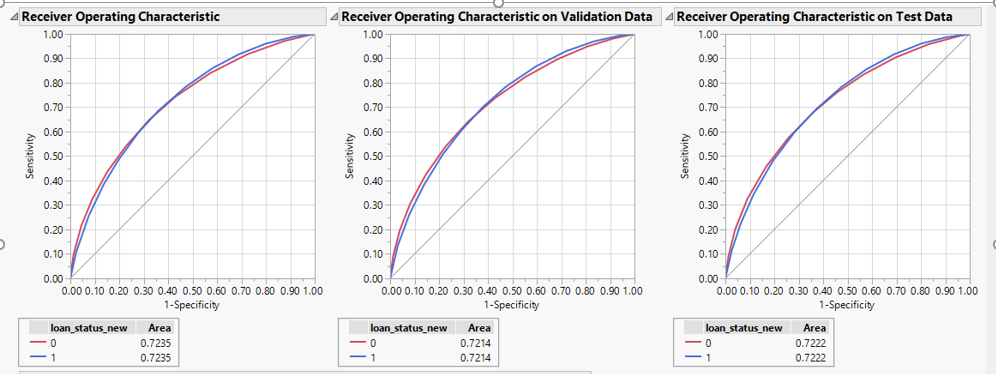


Fig.j: ROC curves for train, validation and test data for Logistic Regression model

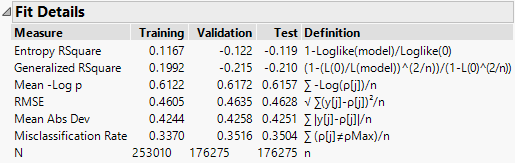


Fig.k: Performance metrics for train, validation and test data for Logistic Regression model (Stratified Sample Data)

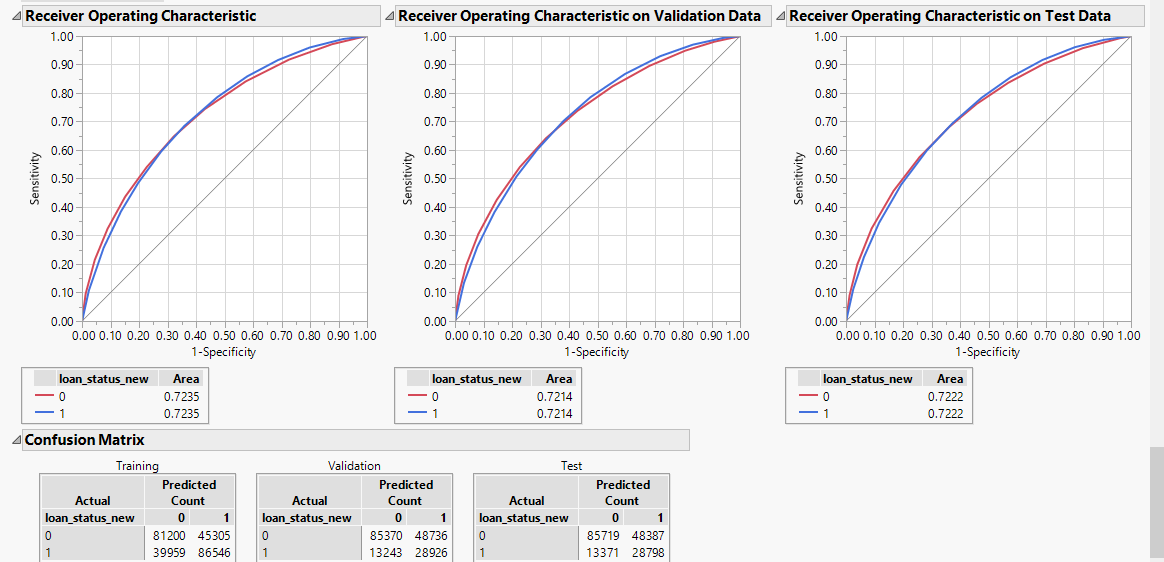


Fig.l: Confusion matrix for train, validation and test data for Logistic Regression model

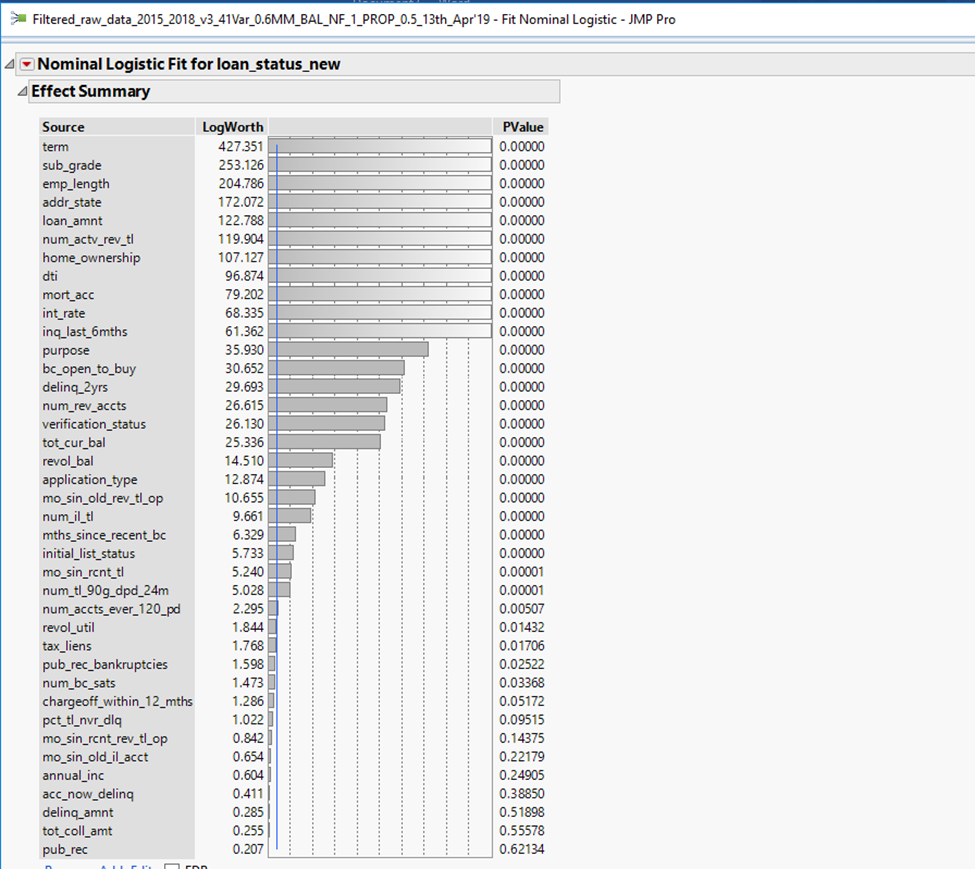


Fig.m: Log-worth of important variables with p-values from Logistic Regression model

On analysing and comparing every model, logistic regression model with stratified data sampling outperformed every other model based on basic model performance KPIs like overall, accuracy, precision, sensitivity and ROC. There were several key features which contributed significantly in capturing variance in data and predicting defaulters – grades based on the credit score (sub-grade), number of payments in months (term) (Figure 7: Distribution of Term by Loan Status), employee tenure (emp\_length), state (addr\_state), loan amount (loan\_amnt), number of currently active revolving trades (Num\_actv\_rev\_tl) (Figure 8: Distribution of Number of Active Revolving Trade with Loan Status), type of home ownership (home\_ownership), ratio of monthly debt to income (dti) (Figure 9: Mean Debt to Income Ratio by Loan Status), number of inquiries in past 6 months (inq\_last\_6mths) (Figure 10: Mean Number of Inquiries in Last 6 Months by Loan Status), loan taken against purpose (purpose) etc. After analysing results of all the models created, logistic regression with stratified sample yielded the best result with total accuracy of ~65%, precision of ~37% and ROC of ~72%.

**CONCLUSIONS AND RECOMMENDATIONS:**

The main goal of this project was to build a classification model to predict potential loan defaulters and identify the features that would help in analysing risky loans. This would help both borrowers and investors to make better informed decisions. Based on the key features identified by logistic regression model, investors should closely scrutinize the borrower’s profile like grade assigned by Lending Club (preferably between A to C) in conjunction with the state they belong to (high GDP states), in order to choose their borrowers. In addition to these features, lenders should also analyse their type of home ownership (preferably who own a house) and number of inquiries made against them in the past 6 months. As observed, borrowers with more than six inquiries (on an average) have higher rate of defaulting as compared to those with lesser inquiries. Borrowers with lower debt-to-income ratio (<20.7 on an average) should be preferred by lenders. All these attributes about the loan profile would help lenders identify riskier loans and take key decisions.

**REFERENCES:**

1. <https://www.lendingclub.com/info/download-data.action>
2. Jennifer Eigo Spring 2019, Class Slides
3. <https://www.lendingclub.com/fileDownload.action?file=javelin.pdf&type=press>

**APPENDIX 1:**

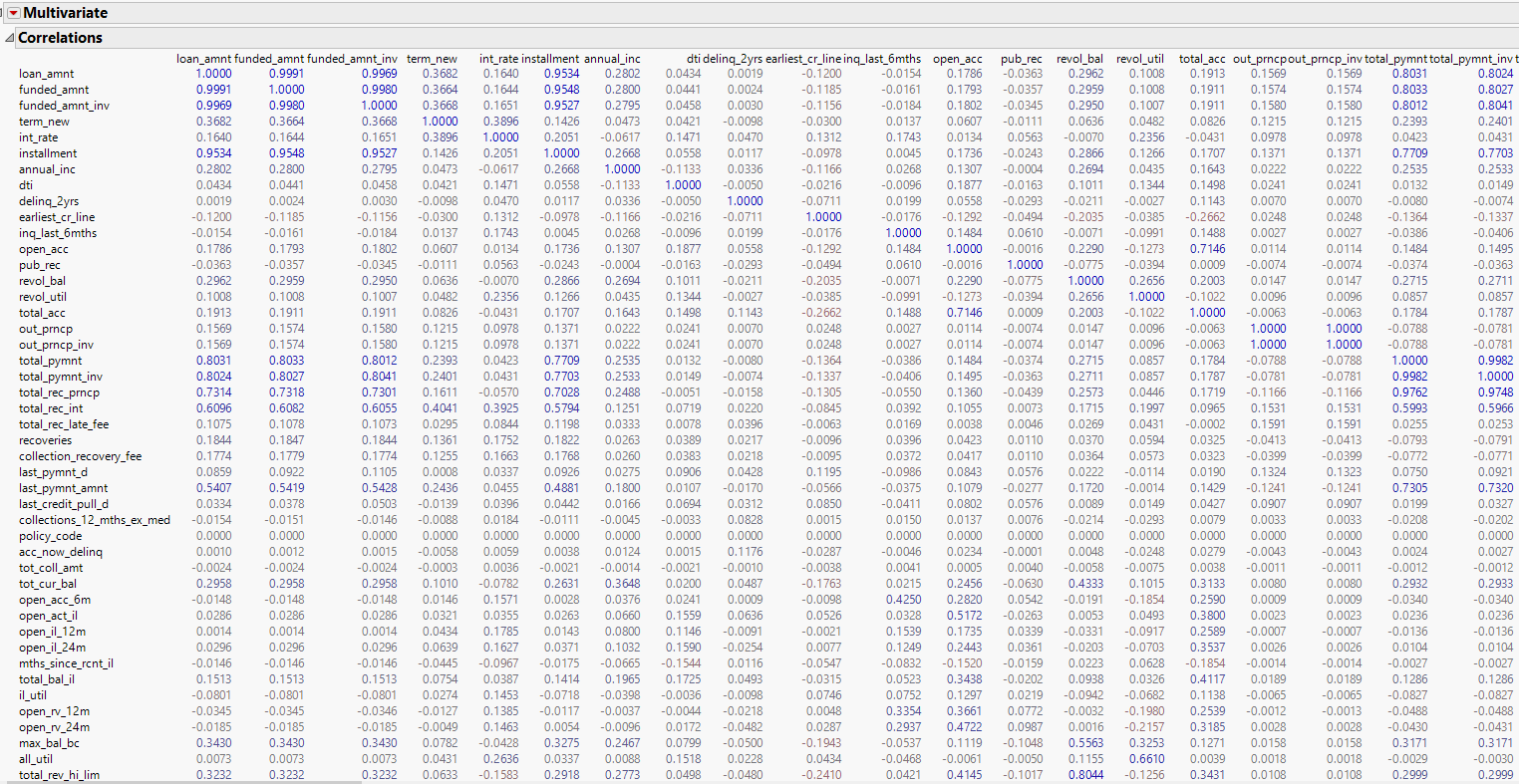


Figure Correlation Matrix

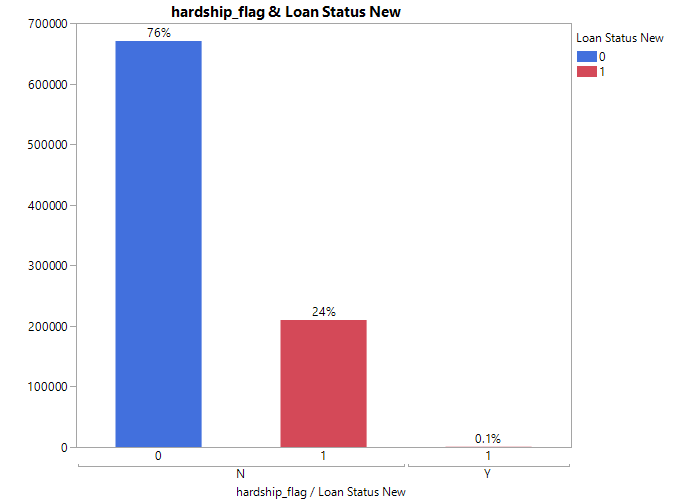
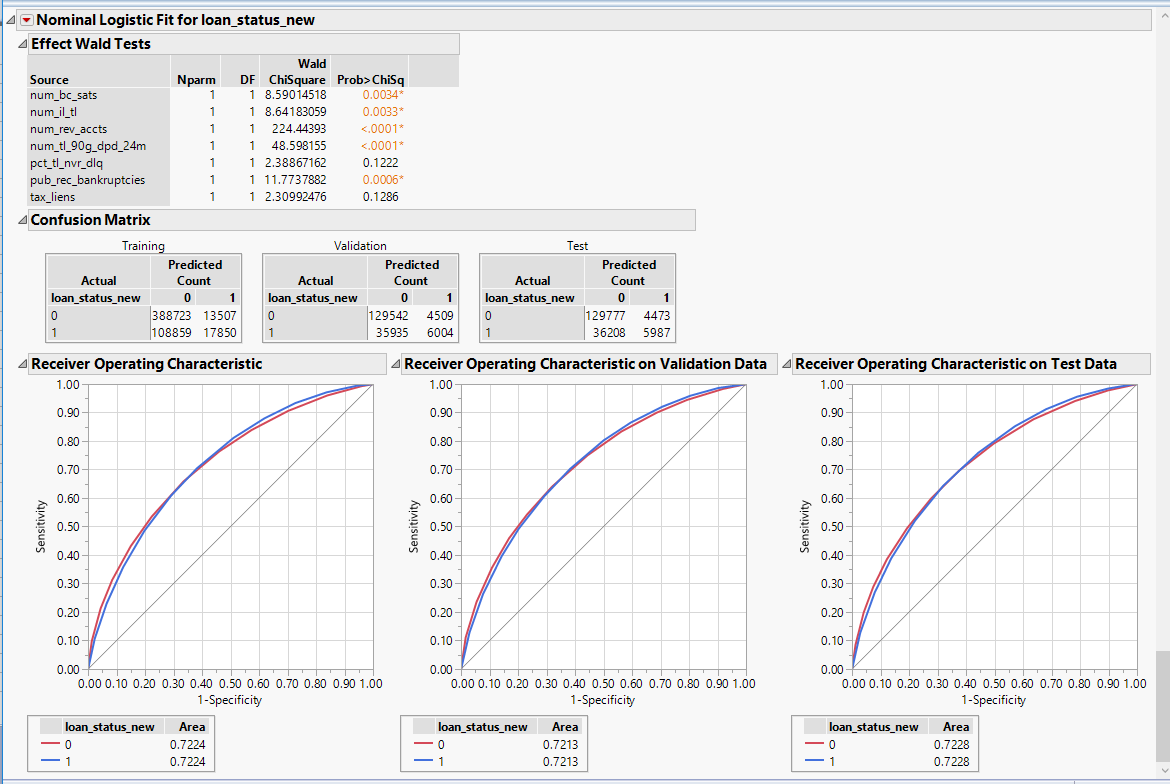


Figure : Distribution of Hardship Flag with Loan Status



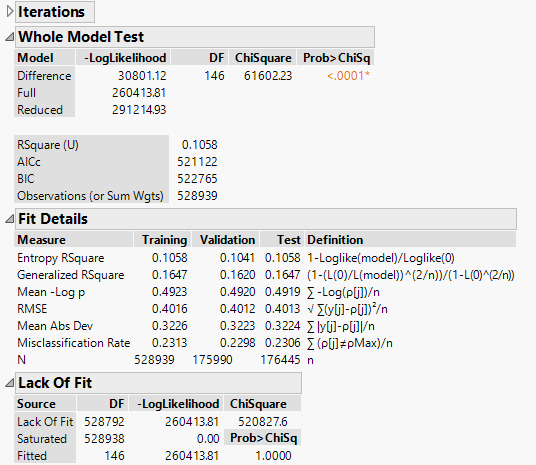
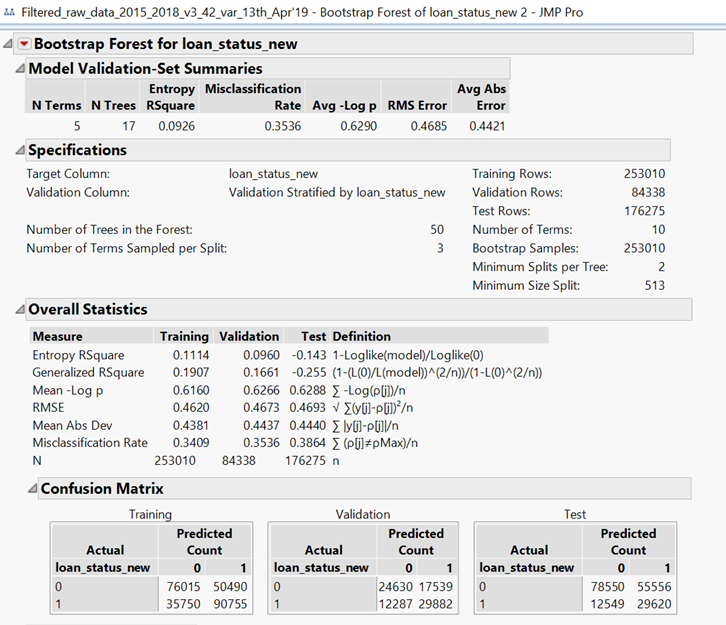
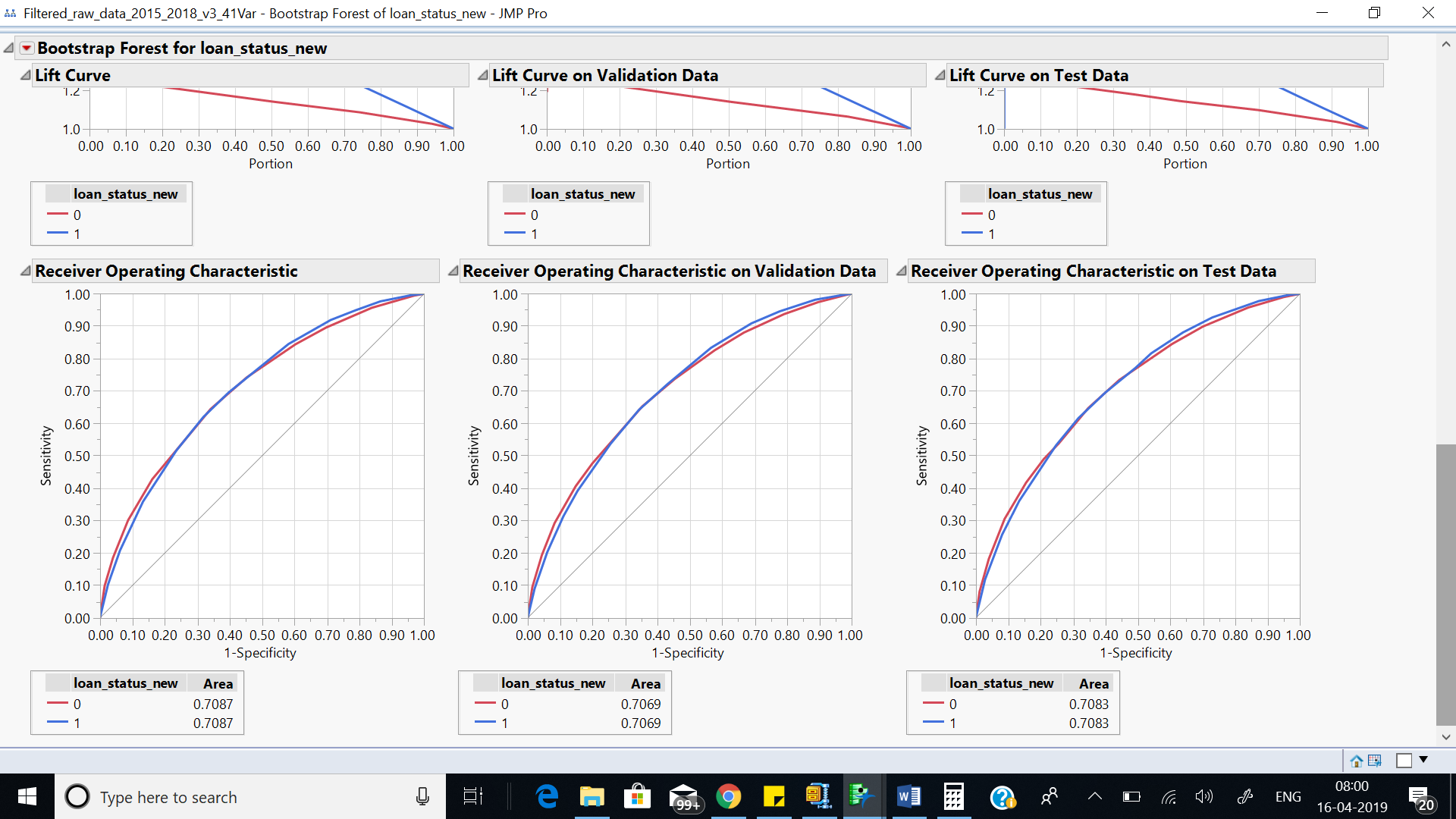
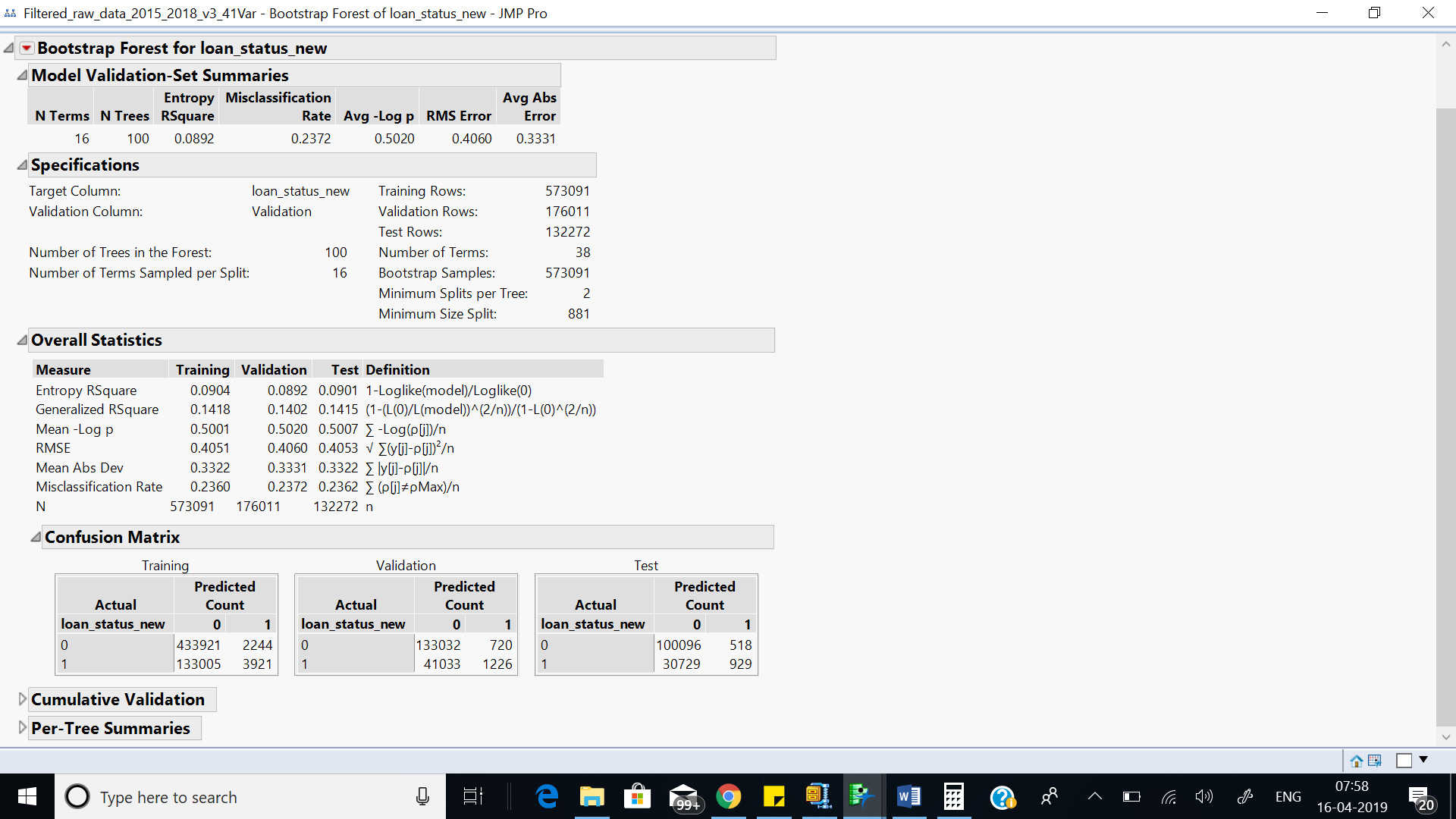


Figure : Logistic Regression Model Summary (Random)







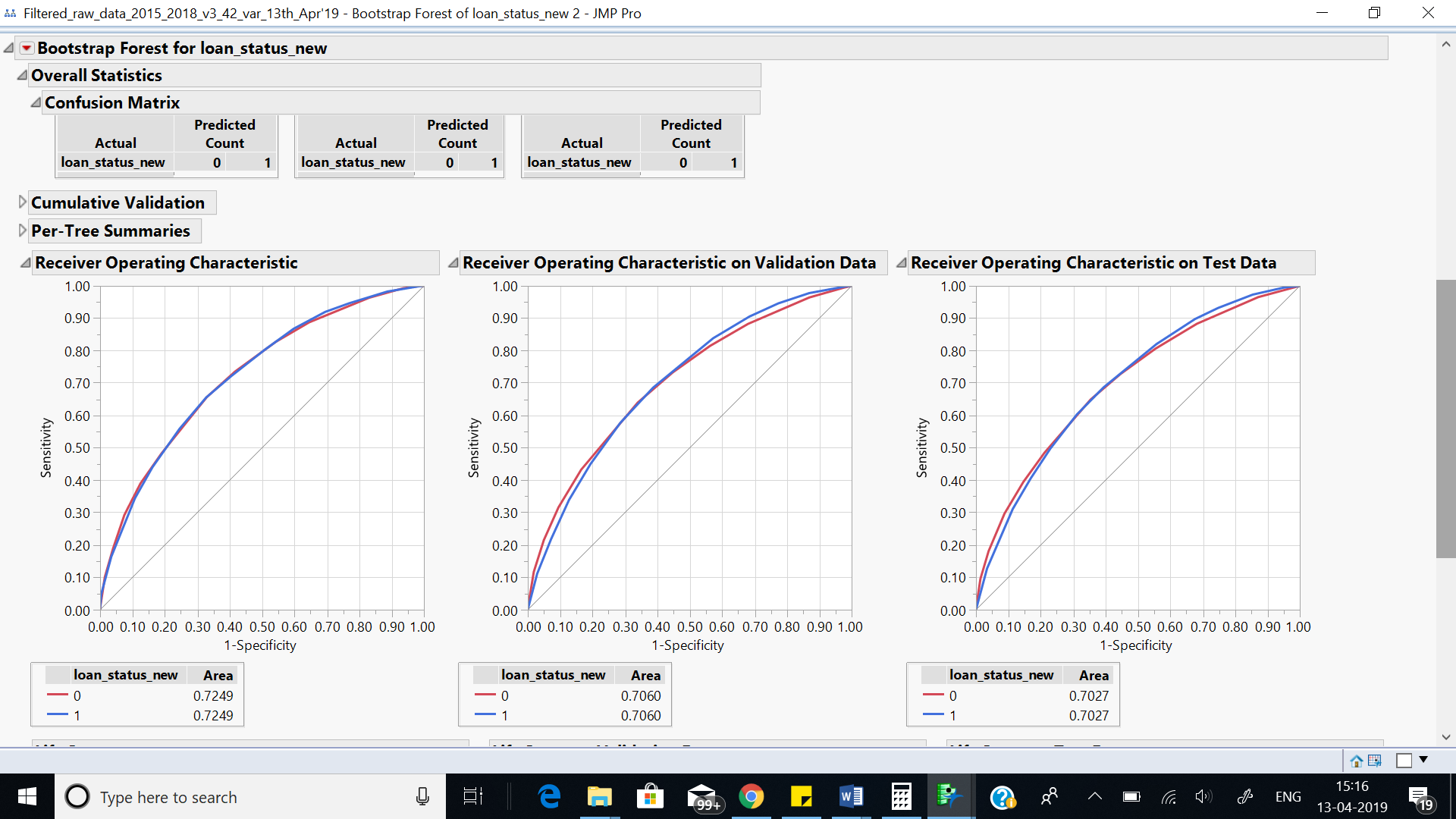
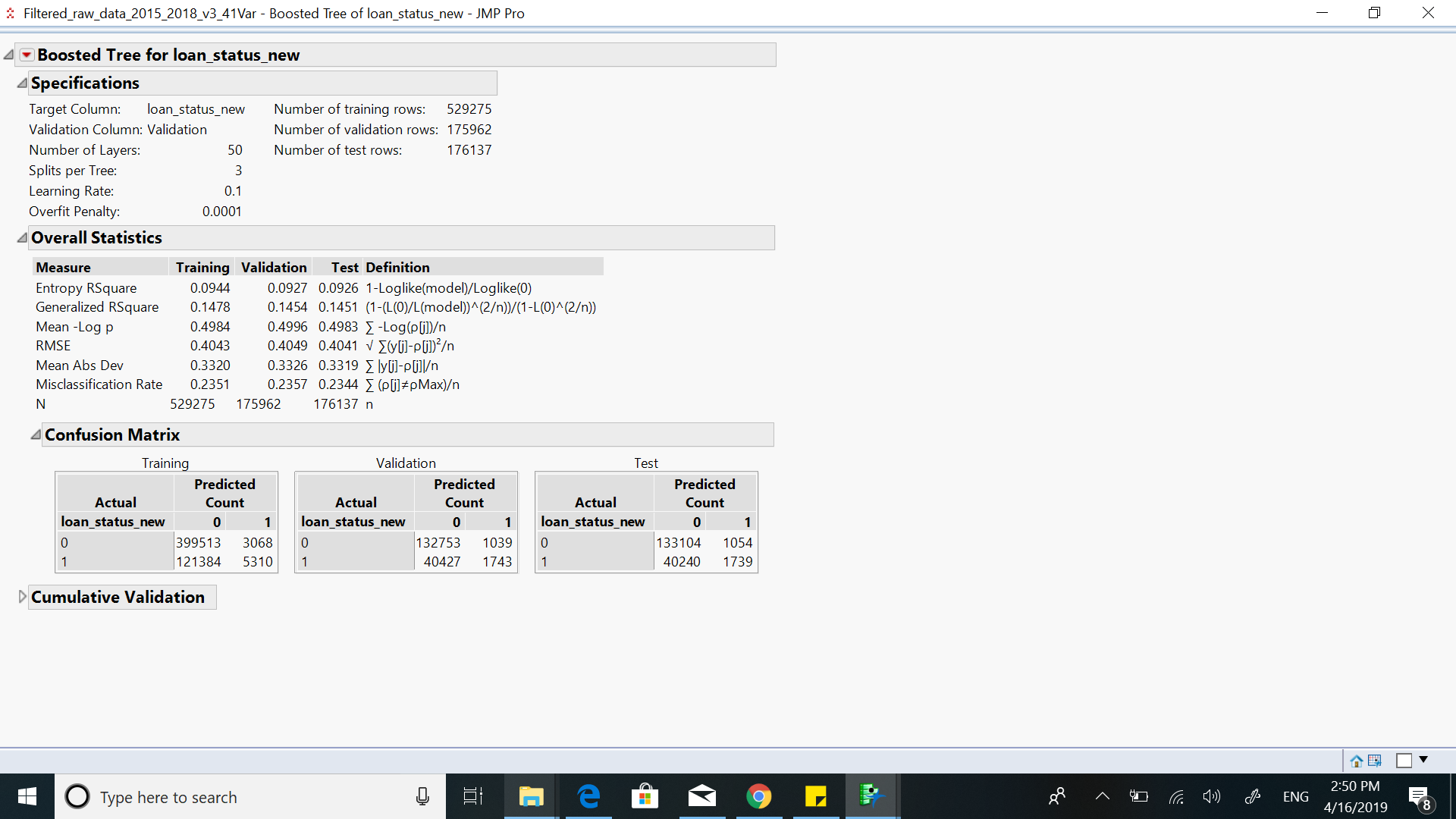
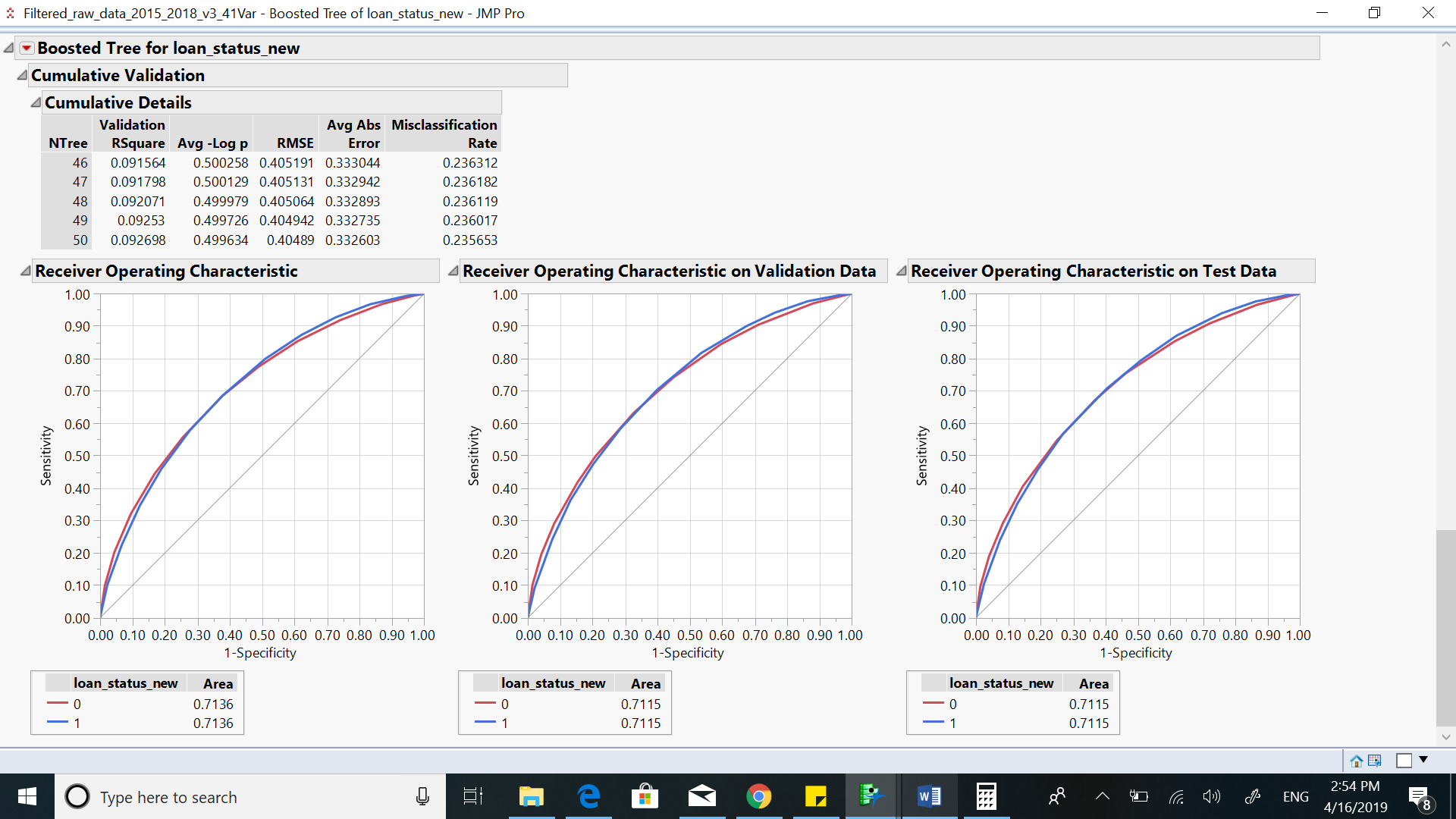
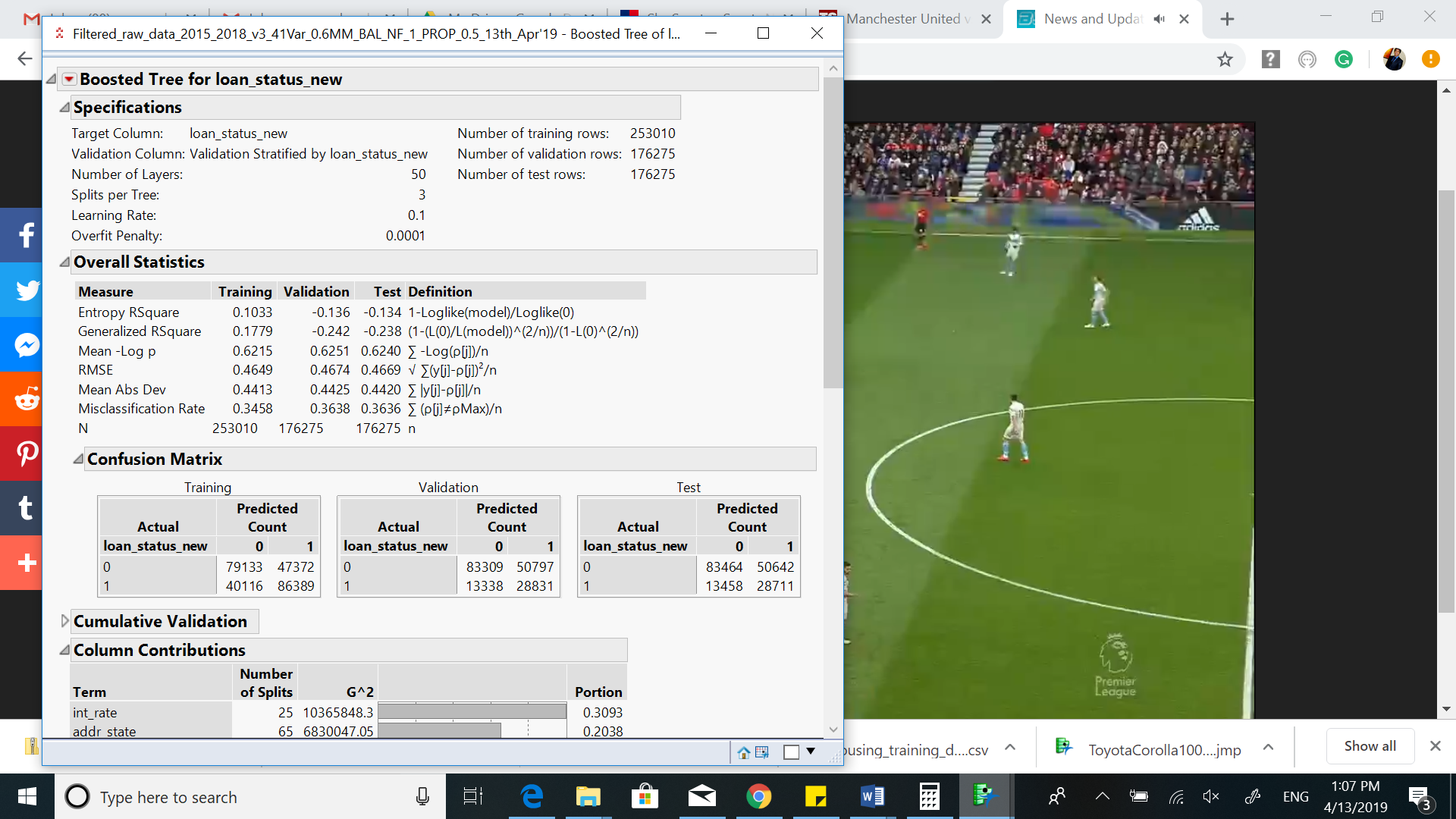


Figure : Bootstrap Model Summary for (Stratified and Random)







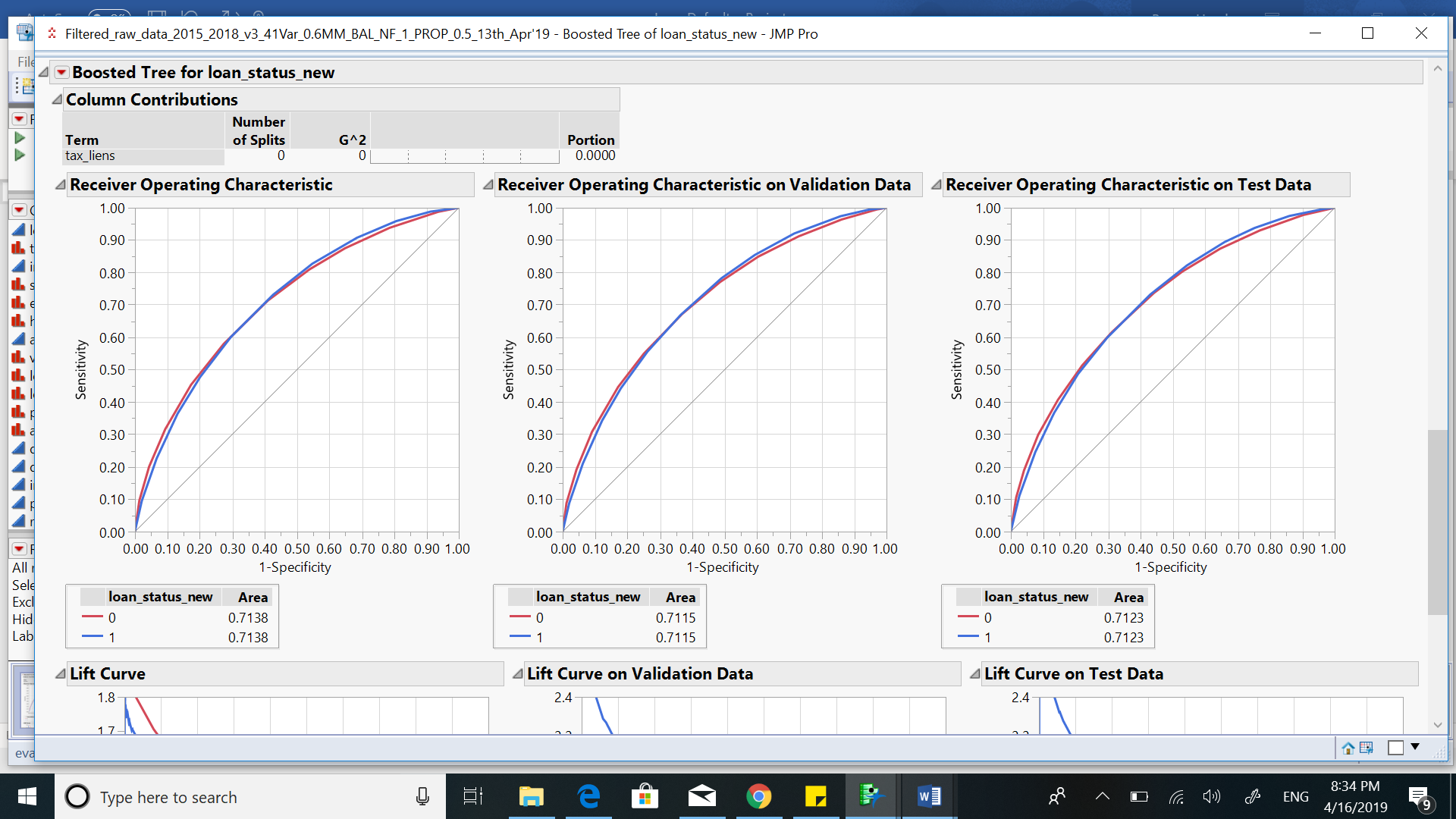
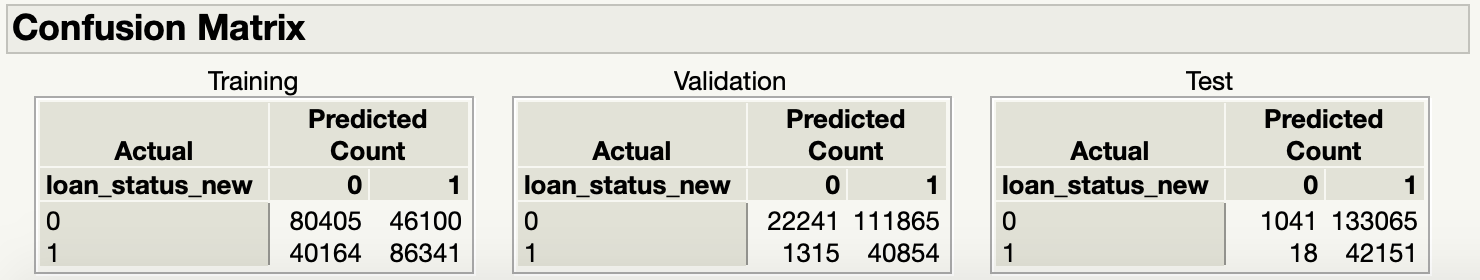
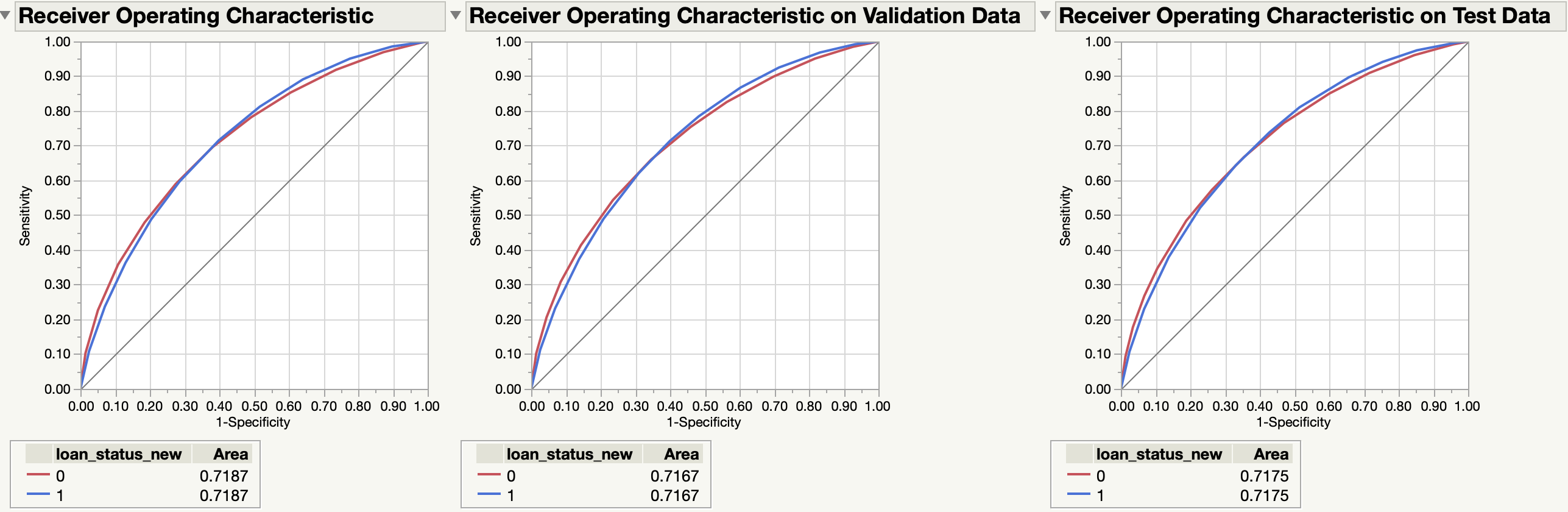
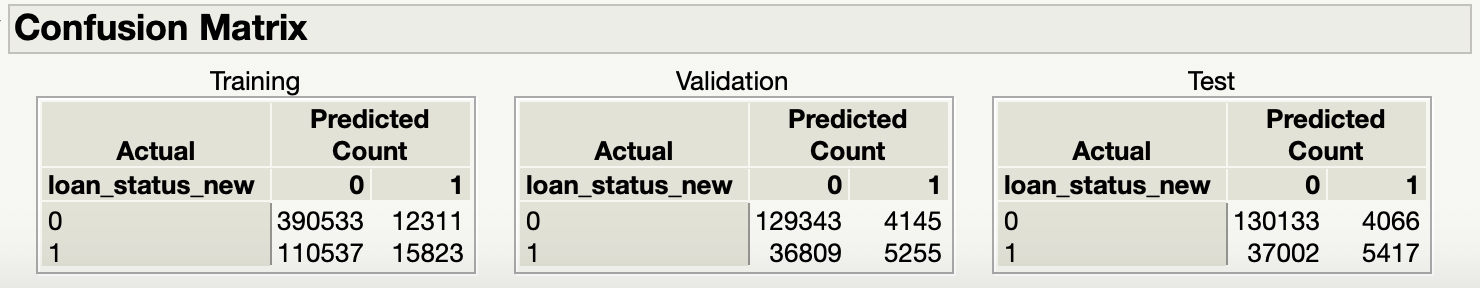


Figure : Boosted Tree Model Summary (Stratified and Random)







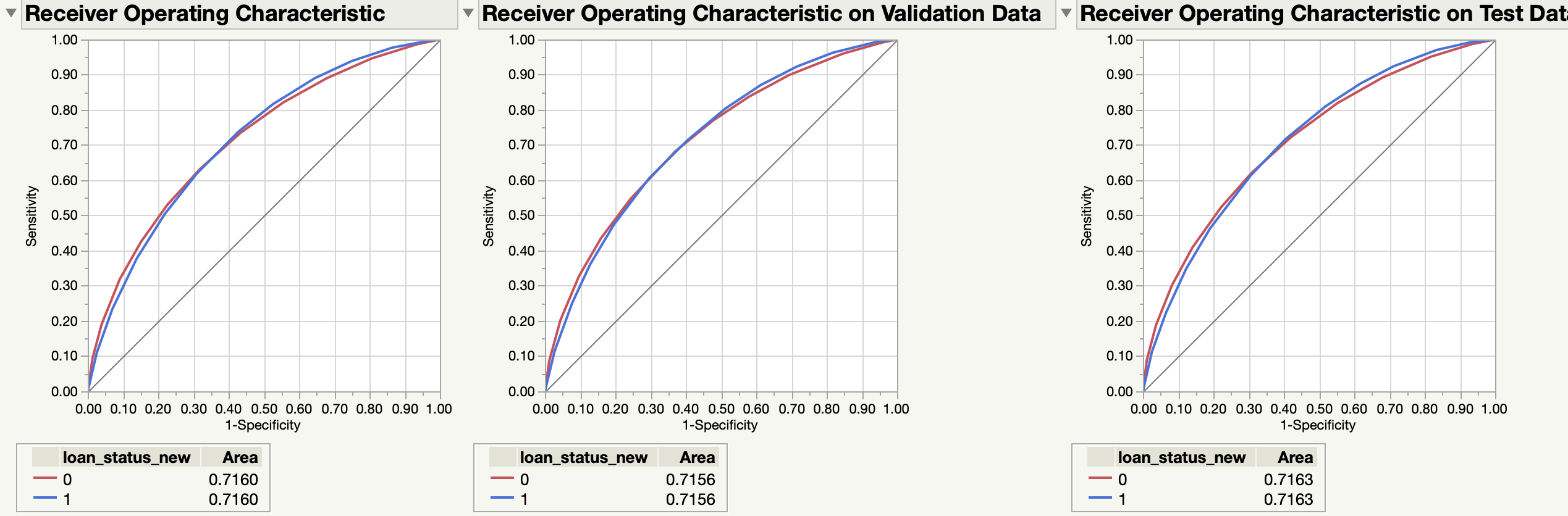


Figure : Logistic Regression Using Principal Component Analysis (Stratified and Random)

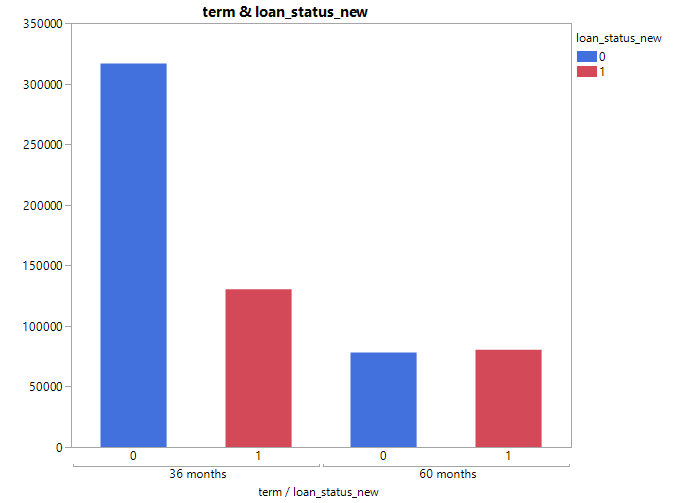


Figure : Distribution of Term by Loan Status

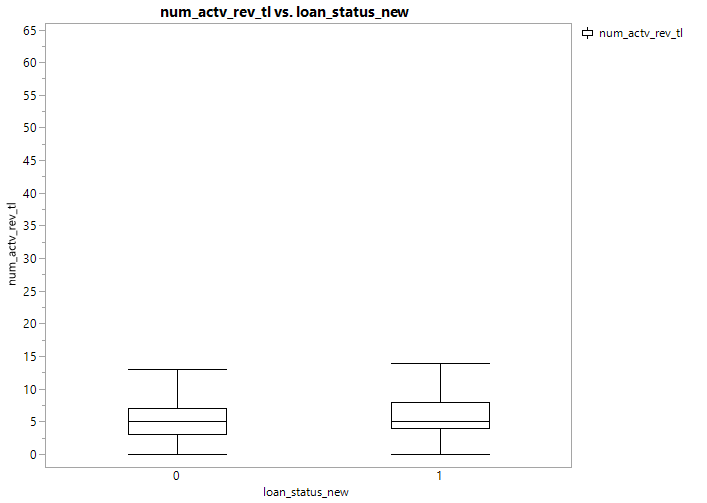


Figure : Distribution of Number of Active Revolving Trade with Loan Status

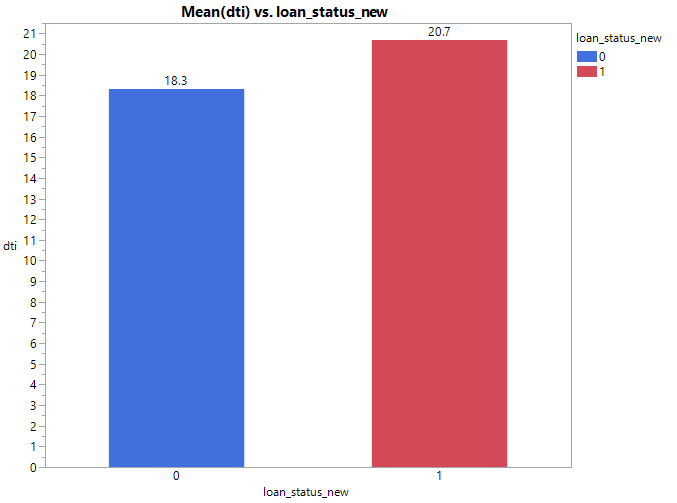


Figure : Mean Debt to Income Ratio by Loan Status

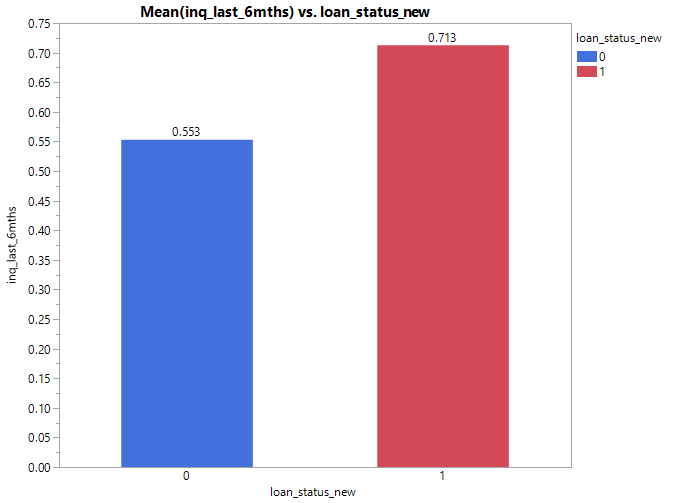


Figure : Mean Number of Inquiries in Last 6 Months by Loan Status

**APPENDIX 2:**

**Data Dictionary (145 variables)**

|  |  |
| --- | --- |
| Columns | Description |
| acc\_now\_delinq | The number of accounts on which the borrower is now delinquent. |
| acc\_open\_past\_24mths | Number of trades opened in past 24 months. |
| addr\_state | The state provided by the borrower in the loan application |
| all\_util | Balance to credit limit on all trades |
| annual\_inc | The self-reported annual income provided by the borrower during registration. |
| annual\_inc\_joint | The combined self-reported annual income provided by the co-borrowers during registration |
| application\_type | Indicates whether the loan is an individual application or a joint application with two co-borrowers |
| avg\_cur\_bal | Average current balance of all accounts |
| bc\_open\_to\_buy | Total open to buy on revolving bankcards. |
| bc\_util | Ratio of total current balance to high credit/credit limit for all bankcard accounts. |
| chargeoff\_within\_12\_mths | Number of charge-offs within 12 months |
| collection\_recovery\_fee | post charge off collection fee |
| collections\_12\_mths\_ex\_med | Number of collections in 12 months excluding medical collections |
| delinq\_2yrs | The number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years |
| delinq\_amnt | The past-due amount owed for the accounts on which the borrower is now delinquent. |
| desc | Loan description provided by the borrower |
| dti | A ratio calculated using the borrower’s total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower’s self-reported monthly income. |
| dti\_joint | A ratio calculated using the co-borrowers' total monthly payments on the total debt obligations, excluding mortgages and the requested LC loan, divided by the co-borrowers' combined self-reported monthly income |
| earliest\_cr\_line | The month the borrower's earliest reported credit line was opened |
| emp\_length | Employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years. |
| emp\_title | The job title supplied by the Borrower when applying for the loan.\* |
| fico\_range\_high | The upper boundary range the borrower’s FICO at loan origination belongs to. |
| fico\_range\_low | The lower boundary range the borrower’s FICO at loan origination belongs to. |
| funded\_amnt | The total amount committed to that loan at that point in time. |
| funded\_amnt\_inv | The total amount committed by investors for that loan at that point in time. |
| grade | LC assigned loan grade |
| home\_ownership | The home ownership status provided by the borrower during registration or obtained from the credit report. Our values are: RENT, OWN, MORTGAGE, OTHER |
| id | A unique LC assigned ID for the loan listing. |
| il\_util | Ratio of total current balance to high credit/credit limit on all install acct |
| initial\_list\_status | The initial listing status of the loan. Possible values are – W, F |
| inq\_fi | Number of personal finance inquiries |
| inq\_last\_12m | Number of credit inquiries in past 12 months |
| inq\_last\_6mths | The number of inquiries in past 6 months (excluding auto and mortgage inquiries) |
| installment | The monthly payment owed by the borrower if the loan originates. |
| int\_rate | Interest Rate on the loan |
| issue\_d | The month which the loan was funded |
| last\_credit\_pull\_d | The most recent month LC pulled credit for this loan |
| last\_fico\_range\_high | The upper boundary range the borrower’s last FICO pulled belongs to. |
| last\_fico\_range\_low | The lower boundary range the borrower’s last FICO pulled belongs to. |
| last\_pymnt\_amnt | Last total payment amount received |
| last\_pymnt\_d | Last month payment was received |
| loan\_amnt | The listed amount of the loan applied for by the borrower. If at some point in time, the credit department reduces the loan amount, then it will be reflected in this value. |
| loan\_status | Current status of the loan |
| max\_bal\_bc | Maximum current balance owed on all revolving accounts |
| member\_id | A unique LC assigned Id for the borrower member. |
| mo\_sin\_old\_il\_acct | Months since oldest bank installment account opened |
| mo\_sin\_old\_rev\_tl\_op | Months since oldest revolving account opened |
| mo\_sin\_rcnt\_rev\_tl\_op | Months since most recent revolving account opened |
| mo\_sin\_rcnt\_tl | Months since most recent account opened |
| mort\_acc | Number of mortgage accounts. |
| mths\_since\_last\_delinq | The number of months since the borrower's last delinquency. |
| mths\_since\_last\_major\_derog | Months since most recent 90-day or worse rating |
| mths\_since\_last\_record | The number of months since the last public record. |
| mths\_since\_rcnt\_il | Months since most recent installment accounts opened |
| mths\_since\_recent\_bc | Months since most recent bankcard account opened. |
| mths\_since\_recent\_bc\_dlq | Months since most recent bankcard delinquency |
| mths\_since\_recent\_inq | Months since most recent inquiry. |
| mths\_since\_recent\_revol\_delinq | Months since most recent revolving delinquency. |
| next\_pymnt\_d | Next scheduled payment date |
| num\_accts\_ever\_120\_pd | Number of accounts ever 120 or more days past due |
| num\_actv\_bc\_tl | Number of currently active bankcard accounts |
| num\_actv\_rev\_tl | Number of currently active revolving trades |
| num\_bc\_sats | Number of satisfactory bankcard accounts |
| num\_bc\_tl | Number of bankcard accounts |
| num\_il\_tl | Number of installment accounts |
| num\_op\_rev\_tl | Number of open revolving accounts |
| num\_rev\_accts | Number of revolving accounts |
| num\_rev\_tl\_bal\_gt\_0 | Number of revolving trades with balance >0 |
| num\_sats | Number of satisfactory accounts |
| num\_tl\_120dpd\_2m | Number of accounts currently 120 days past due (updated in past 2 months) |
| num\_tl\_30dpd | Number of accounts currently 30 days past due (updated in past 2 months) |
| num\_tl\_90g\_dpd\_24m | Number of accounts 90 or more days past due in last 24 months |
| num\_tl\_op\_past\_12m | Number of accounts opened in past 12 months |
| open\_acc | The number of open credit lines in the borrower's credit file. |
| open\_acc\_6m | Number of open trades in last 6 months |
| open\_il\_12m | Number of installment accounts opened in past 12 months |
| open\_il\_24m | Number of installment accounts opened in past 24 months |
| open\_act\_il | Number of currently active installment trades |
| open\_rv\_12m | Number of revolving trades opened in past 12 months |
| open\_rv\_24m | Number of revolving trades opened in past 24 months |
| out\_prncp | Remaining outstanding principal for total amount funded |
| out\_prncp\_inv | Remaining outstanding principal for portion of total amount funded by investors |
| pct\_tl\_nvr\_dlq | Percent of trades never delinquent |
| percent\_bc\_gt\_75 | Percentage of all bankcard accounts > 75% of limit. |
| policy\_code | publicly available policy\_code=1 new products not publicly available policy\_code=2 |
| pub\_rec | Number of derogatory public records |
| pub\_rec\_bankruptcies | Number of public record bankruptcies |
| purpose | A category provided by the borrower for the loan request. |
| pymnt\_plan | Indicates if a payment plan has been put in place for the loan |
| recoveries | post charge off gross recovery |
| revol\_bal | Total credit revolving balance |
| revol\_util | Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit. |
| sub\_grade | LC assigned loan subgrade |
| tax\_liens | Number of tax liens |
| term | The number of payments on the loan. Values are in months and can be either 36 or 60. |
| title | The loan title provided by the borrower |
| tot\_coll\_amt | Total collection amounts ever owed |
| tot\_cur\_bal | Total current balance of all accounts |
| tot\_hi\_cred\_lim | Total high credit/credit limit |
| total\_acc | The total number of credit lines currently in the borrower's credit file |
| total\_bal\_ex\_mort | Total credit balance excluding mortgage |
| total\_bal\_il | Total current balance of all installment accounts |
| total\_bc\_limit | Total bankcard high credit/credit limit |
| total\_cu\_tl | Number of finance trades |
| total\_il\_high\_credit\_limit | Total installment high credit/credit limit |
| total\_pymnt | Payments received to date for total amount funded |
| total\_pymnt\_inv | Payments received to date for portion of total amount funded by investors |
| total\_rec\_int | Interest received to date |
| total\_rec\_late\_fee | Late fees received to date |
| total\_rec\_prncp | Principal received to date |
| total\_rev\_hi\_lim | Total revolving high credit/credit limit |
| url | URL for the LC page with listing data. |
| verification\_status | Indicates if income was verified by LC, not verified, or if the income source was verified |
| verified\_status\_joint | Indicates if the co-borrowers' joint income was verified by LC, not verified, or if the income source was verified |
| zip\_code | The first 3 numbers of the zip code provided by the borrower in the loan application. |
| revol\_bal\_joint | Sum of revolving credit balance of the co-borrowers, net of duplicate balances |
| sec\_app\_earliest\_cr\_line | Earliest credit line at time of application for the secondary applicant |
| sec\_app\_inq\_last\_6mths | Credit inquiries in the last 6 months at time of application for the secondary applicant |
| sec\_app\_mort\_acc | Number of mortgage accounts at time of application for the secondary applicant |
| sec\_app\_open\_acc | Number of open trades at time of application for the secondary applicant |
| sec\_app\_revol\_util | Ratio of total current balance to high credit/credit limit for all revolving accounts |
| sec\_app\_open\_act\_il | Number of currently active installment trades at time of application for the secondary applicant |
| sec\_app\_num\_rev\_accts | Number of revolving accounts at time of application for the secondary applicant |
| sec\_app\_chargeoff\_within\_12\_mths | Number of charge-offs within last 12 months at time of application for the secondary applicant |
| sec\_app\_collections\_12\_mths\_ex\_med | Number of collections within last 12 months excluding medical collections at time of application for the secondary applicant |
| sec\_app\_mths\_since\_last\_major\_derog | Months since most recent 90-day or worse rating at time of application for the secondary applicant |
| hardship\_flag | Flags whether or not the borrower is on a hardship plan |
| hardship\_type | Describes the hardship plan offering |
| hardship\_reason | Describes the reason the hardship plan was offered |
| hardship\_status | Describes if the hardship plan is active, pending, canceled, completed, or broken |
| deferral\_term | Amount of months that the borrower is expected to pay less than the contractual monthly payment amount due to a hardship plan |
| hardship\_amount | The interest payment that the borrower has committed to make each month while they are on a hardship plan |
| hardship\_start\_date | The start date of the hardship plan period |
| hardship\_end\_date | The end date of the hardship plan period |
| payment\_plan\_start\_date | The day the first hardship plan payment is due. For example, if a borrower has a hardship plan period of 3 months, the start date is the start of the three-month period in which the borrower is allowed to make interest-only payments. |
| hardship\_length | The number of months the borrower will make smaller payments than normally obligated due to a hardship plan |
| hardship\_dpd | Account days past due as of the hardship plan start date |
| hardship\_loan\_status | Loan Status as of the hardship plan start date |
| orig\_projected\_additional\_accrued\_interest | The original projected additional interest amount that will accrue for the given hardship payment plan as of the Hardship Start Date. This field will be null if the borrower has broken their hardship payment plan. |
| hardship\_payoff\_balance\_amount | The payoff balance amount as of the hardship plan start date |
| hardship\_last\_payment\_amount | The last payment amount as of the hardship plan start date |
| disbursement\_method | The method by which the borrower receives their loan. Possible values are: CASH, DIRECT\_PAY |
| debt\_settlement\_flag | Flags whether or not the borrower, who has charged-off, is working with a debt-settlement company. |