Predictive Analysis for Loan Delay and Default as well as the Prospective Regions and Fields for the Kiva Lenders and Borrowers

By

Sanzida Parvin

A Capstone Project Paper Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

In

Data Science

University of Wisconsin-La Crosse

La Crosse, Wisconsin

May 2020

ABSTRACT

According to World bank, around 10 percentage or 735 million of world population lived under extreme poverty line (estimated data on 2015). Expense is less than $1.90 per day. Their estimation on recent pandemic COVID-19 will increase the number by 0.3 to 0.7 percentage, to around 9 percent in 2020. The number of populations living under $3.20 per day will increase by 0.3 to 1.7 percentage, to 23 percent or higher. Also, the people living under $5.50 per day will increase by 0.4 to 1.9 percentage, to 42 percent or higher. These huge numbers are considered as under poverty line population. Their main sources of income are agriculture, day laborer, livestock, small retail shop etc. Most of them do not have any bank access for their transections or not qualified for the standard bank loan. Microfinance opened a new window for these types of poverty line's small business owners and/or entrepreneurs. Because of microfinancing they can think out of box and can raise their livelihood out of poverty line.

Kiva.org is a non-profit micro financial organization, worked in more than 80 countries of the world to help those underserved communities financially and raised their living standard. It was founded in October 2005, inspired by the work of Nobel Laureate Dr. Mohammad Yunus, considered as the pioneer of modern microfinance.

One of the biggest challenges for this type of microfinance organization is to find out potentially good or bad loans. From the borrower's side, probable reasons for time delay of getting the loan or completely not getting the loan are another concern. The main purpose of this project is to predict the probable good/bad loans as well as the reasons behind the loan delay for Kiva.org and the borrowers. The dataset was collected from Kaggle competition site provided by Kiva organization. The model accuracy was measured using logistic regression method.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Background

To help the poor people raise their living standard the non-profit Kiva.org started their mission in October 2005. Until now they are working in 80 plus countries with more than 1.5 million people to fund over 2 million borrowers all over the world. Inspired by the work of Dr. Mohammad Yunus, pioneer of micro finance and founder of Grameen Bank, Kiva began their journey. With the help of internet and crowdfunding, anyone through Kiva can participate in the program to help the borrower start or grow a business, go to school, access clean energy or realize their potential. The lower limit of the lending is as little as $25. Kiva has their own field partners through which they distribute the loan to the borrowers with zero interest rate. Field Partners are micro-finance organizations around the world "responsible for screening borrowers, posting loan requests to Kiva website, disbursing loans and collecting repayments, and otherwise administering Kiva loans". The repayment rate from the borrowers of Kiva is 96.4 percent with bullet, irregular, monthly or weekly repayment interval methods. When a borrower applies for loan, the Kiva field partner sanction it after a minimal investigation. After that the loan published on the Kiva site and started collecting money by lender/s to reimburse the field partners. A borrower can apply for the loan in two ways, one is by the field partner's an another is to Kiva site directly. Similarly, at a specific repayment interval the borrower repays the loan to the lender/s through the field partner or directly. The repayment goes directly to the lender's account no matter how the repayment was made. The lenders can use the repayments to fund new loans, donate or withdraw the money. The operating cost of Kiva covers mostly from the donation of the lenders instead of loan. The rest expenses are covered through grants, donations from foundations and supporters and Field Partner’s service fees. At a glance of how Kiva works,

Figure 1*:* Flow chart of how kiva works

During the project proposal time, when I found the Kiva loans datasets through an open source platform it attracted my eyes. Moreover, after seeing the great repayment rate, I was interested to know about the rate of finally sanctioning the loans and any loans default that can be addressed to solve the problem/s to increase the percentage of funded loans. I choose the predictive analysis method to find out the prospective bad loans by measuring the accuracy level of the selected model.

Objective

The main objective of this project was to fit a model that can predict the loans default due to time delay and/or partially or fully unfunded loans with an accepted level of accuracy. After that, based on this prediction the prospective regions and sectors of loan can be determined or find out. Being a full-time student, I choose this project that was not related to any employer or direct client. But this was a client-based project and the prospective client of this project would be the Kiva organization itself and the lenders and borrowers of it. To measure the accuracy level of the model, the total cleaned dataset was divided into two parts, train set and test set. The train set was used to feed the model and the test set was used to measure the accuracy level.

DATA SOURCES

The relevant information and data were collected from Kaggle, an online community of data scientists and machine learning practitioners and an open source of original data sets and the Kiva Data Snapshots. Kiva opened their loan data sets from 2014 to 2017 to make a competition for their own objectives. The data set that was used in this capstone has 671205 observations of 20 different variables. The descriptions of the variables are,

a. id - Unique ID for loan
b. funded_amount - The amount disbursed by Kiva to the field agent(USD)
c. loan_amount - The amount disbursed by the field agent to the borrower(USD)
d. activity - More granular category
e. sector - High level category
f. use - Exact usage of loan amount
g. country_code - ISO country code of country in which loan was disbursed
h. country - Full country name of country in which loan was disbursed
i. region - Full region name within the country
j. currency - The currency in which the loan was disbursed
k. partner_id - ID of partner organization
l. posted_time - The time at which the loan is posted on Kiva by the field agent

m. disbursed_time - The time at which the loan is disbursed by the field agent to the borrower
n. funded_time - The time at which the loan posted to Kiva gets funded by lenders completely
o. term_in_months - The duration for which the loan was disbursed in months
p. lender_count - The total number of lenders that contributed to this loan
q. tags
r. borrower_genders - Comma separated M,F letters, where each instance represents a single male/female in the group
s. repayment_interval
t. date - Date at which the data was posted in the dataset

Additionally, some other resources and the Kiva website were used to better understand about Kiva, micro financing and what and how they work.

METHODOLOGY

Data Processing

The main time consuming and tedious part of data science project is cleaning and processing the data for analysis. Without processing, the result may bias or can give false impression about the final outcomes. It is obvious to have some unusual or missing values inside a big data set due to human error or some other types of data errors. So, data cleaning and processing is very important for a good and reliable analysis.

The Kiva_loans dataset had some missing values under different variables. The percentage range of those missing values varied from 0.001% to 25.6% by variables. After cleaning and manipulating the Kiva_loans dataset, around 3% of total observations were lost, which was acceptable. The table below will show the variable wise missing values and their percentage,

Table 1: List of missing values

	Variables	Missing_values	Percentage
1	id	0	0.000
2	funded_amount	0	0.000
3	loan_amount	0	0.000
4	activity	0	0.000
5	sector	0	0.000
6	use	4228	0.630
7	country_code	8	0.001
8	country	0	0.000
9	region	56800	8.460
10	currency	0	0.000
11	partner_id	13507	2.010
12	posted_time	0	0.000
13	disbursed_time	2396	0.360
14	funded_time	48331	7.200
15	term_in_months	0	0.000
16	lender_count	0	0.000
17	tags	171416	25.540
18	borrower_genders	4221	0.630
19	repayment_interval	0	0.000
20	date	0	0.000

From the table, it can be seen that the maximum missing values were under variable tag, around 25.5%, which was probably used for social media purposes and had very less to no effect on loans default. Therefore, that variable was removed completely and saved around 25% of total data loss. Other observations with one or more missing fields were removed from the dataset . After dealing with missing values, some redundant, highly correlated and comparatively less important variables were removed for a smooth analysis. The borrower_genders column had more than 10 thousand different combinations of male and female borrowers. Which could affect the analysis and increased the time of analysis for couple of days. Some feature engineering was done on that column and converted it to 3 different types of borrower_genders, 'male'(only male/s), 'female'(only female/s) and 'female,male'(both male/s and female/s) using Microsoft Excel. Because, more that 10thousands unique values were not possible to identified and converted through R functions. All date columns posted_time, disbursed_time, funded_time and date had 4years × 365days × n-borrower observations. Which was almost impossible to process in a home machine like laptop or desktop. Those variables were also removed from the dataset after extracting the required information of time delay and putting them in new columns. After that, activity, use and country_code columns were removed due to redundancy. Finally, the cleaned dataset has 653896 observations with 13 variables.

a. id - Unique ID for loan
b. funded_amount - The amount disbursed by Kiva to the field agent(USD)
c. loan_amount - The amount disbursed by the field agent to the borrower(USD)
d. sector - High level category
e. country - Full country name of country in which loan was disbursed
f. partner_id - ID of partner organization
g. term_in_months - The duration for which the loan was disbursed in months
h. lender_count - The total number of lenders that contributed to this loan

i. borrower_genders - Comma separated M,F letters, where each instance represents a single male/female in the group
j. repayment_interval
k. loan_status - Whether the loan is defaulted or not
l. fund_status - If the loan was fully funded or partially funded or unfunded
m. expiration - If the funded time has expired or not

The loan_status of the borrowers was defined from the definition of Kiva's loan default. According to Kiva, the loan amount has to be fully distributed within 30 days. Otherwise the loan will be defaulted. Based on that definition the fund_status had three states, 'Partially Funded' (loans that were funded partially), 'Not Funded' (loans that were not funded at all), 'Time Expired' (loans that were exceeded 30 days after posting). Finally, the good and bad loan_status was defined from the fund_status. The loans that were partially funded or unfunded or expired defined as bad loans.

Statistical Techniques and Outcomes

I had spent a decent amount of time determining the method that will suite the analysis and give an acceptable accuracy rate. I had tried several techniques and finally ended up using logistic regression model with categorical response variable. The cleaned and modified dataset then divided into two groups, training data - randomly selected 80% of sample observations and test data - randomly selected 20% of sample observations. The logistic regression model was run on training dataset and then predict the model accuracy level using the test dataset. In regression model only those variables were used which were significant for the response variable, also did not have high collinearity with other variables. The overall accuracy level of the model found 87.28%, which was a good number. And percentage of correctly predicted bad loans was 92.8%. The outcomes of the statistical calculations (predicted probability and accuracy level) using R programming language is shown below.

```
> predprob = predict(loan_model, newdata = test_data, type="response")
> pred_goodbad = cut(predprob, breaks = c(-Inf, 0.5, Inf), labels=c("Bad", "Good"))
> contin_table1 = table(test_data$loan_status, pred_goodbad)
> addmargins(contin_table1)
      pred_goodbad
          Bad   Good    Sum
  bad    9864  15864  25728
  good    765 104287 105052
  Sum   10629 120151 130780
> proportion1 = sum(diag(contin_table1)) / sum(contin_table1)
> print(paste('Overall Accuracy(Correctly predicted outcomes):',round((proportion1*100),2)))
[1] "Overall Accuracy(Correctly predicted outcomes): 87.28"
> TP = contin_table1[1,1]
> FN = contin_table1[2,1]
> TN = contin_table1[2,2]
> FP = contin_table1[1,2]
> true_bad = TP/(TP+FN)
> print(paste('Correctly predicted bad loans: ', round((true_bad*100),2)))
[1] "Correctly predicted bad loans:  92.8"
```

The estimated coefficients, standard errors and average effect on odds from the logistic regression model presented in Appendix A.

DISCUSSION

Technical Limitations

It is obvious that most of the analysis in Data Science project has some limitations. And to know about those limitations are very important before interpreting the final results. So that the interpretation of the analysis would be more understandable. The first and foremost limitation of this project was the lack of direct communication with any personnel related to Kiva. Any possible communication could help to visualize and understand the in-depth concept and scenario of the Kiva_loans dataset including their terms and conditions. Another limitation was not having the proper and real time conditions of measuring the loans default which were applied in reality during sanctioning the loans, rather than the definition on Kiva site. Because, after exploring the dataset it was founded that, a lot of expired(after 30 days of posting) loans were disbursed as well as funded on the Kiva site. Same thing happened in partially or unfunded loans. To do the analysis in this project the default loans were defined with the general definition from Kiva site. Due to the limitation of home machine some variables needed to remove from the original dataset, such as all ‘dates’ and ‘use’ columns.

Situational Limitations

The devastated COVID-19 pandemic was outbreak in USA just after few weeks of the Spring-2020 semester has begun. Living in New Jersey, the second most affected zone in USA, destroyed the mental condition to concentrate on the project work. Moreover, the daily life was so much impacted that the situation didn’t allow me to do some of my extra plans to make this project more in-depth. But I am glad that I was able to finish the main objectives of this capstone project despite of all the unwanted situations.

Results and Interpretation

The overall accuracy level of the model was 87.28% with 92.8% of correctly predicted bad loans, which was very satisfactory. The summery of the logistic regression model given in the appendix, shown that the variables funded_amount, loan_amount, country and repayment_interval were statistically significant with great p-value. The standard error rate for all variables were also very low. Overall the model as a whole fits significantly better than an empty model. From the lists of top frequent countries and sectors with fully funded loans, (see APPENDIX A) it can be seen that, Philippines has the highest number of fully funded loans following with Kenya, Cambodia, El Salvador Pakistan etc. Similarly, Agriculture is the most frequent sector of getting fully funded loans following with Food, Retail, Services etc.

CONCLUSION AND RECOMMENDATIONS

Micro financing has opened a new window to the under-poverty line population around the world. During this project the main concern was to find out a way that can help the nonprofit organizations like Kiva and others to minimize the odds and maximize the welfare for both lenders and borrowers.

The future recommendation of this project will be to fit some other statistical models and compare the accuracy levels in order to get the best one. At the same time, collect some other similar organization's dataset and apply the model to compare the performance between them. Also, if possible and applicable then try to find out the strategies of operation of that best organization and modify Kiva's strategies to help the lenders and borrowers at its level best.

REFERENCES

The World Bank [*Official site*]. Retrieved from https://www.worldbank.org/en/topic/poverty/overview

Pocchiari, Martina (2016). Predicting a successful loan in prosocial lending: the role of partner accreditation and storytelling on Kiva.org. *ResearchGate*

Data Science for Good: Kiva Crowdfunding. *Kaggle*. Retrieved from https://www.kaggle.com/kiva/data-science-for-good-kiva-crowdfunding

SRK, (2018). Simple Exploration Notebook – Kiva. *Kaggle*. Retrieved from https://www.kaggle.com/sudalairajkumar/simple-exploration-notebook-kiva

Kiva.org [*Official site*]. Retrieved from https://www.kiva.org/

Rana, Muhammad Sohail Jafar. (2016). Micro-Finance and Its Impacts on People and Societies. *ResearchGate*

Q&A: Expiring loans, Credit Limits, and the evolution of Kiva. *Kiva*. Retrieved from https://pages.kiva.org/blog/qa-expiring-loans-credit-limits-and-the-evolution-of-kiva

APPENDIX A

Summary of logistic regression model,

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-4.0706   0.2214   0.3218   0.5723   8.4904

Coefficients:
                                              Estimate Std. Error z value Pr(>|z|)
(Intercept)                                  2.487e+01  6.838e+04   0.000    1.000
id                                           3.141e-07  2.390e-08  13.145   <2e-16 ***
funded_amount                                5.769e-01  8.588e-01   0.672    0.502
loan_amount                                 -5.768e-01  8.588e-01  -0.672    0.502
countryAlbania                              -2.238e+01  6.838e+04   0.000    1.000
countryArmenia                              -2.307e+01  6.838e+04   0.000    1.000
countryAzerbaijan                           -2.290e+01  6.838e+04   0.000    1.000
countryBelize                               -1.805e+01  6.838e+04   0.000    1.000
countryBenin                                -2.376e+01  6.838e+04   0.000    1.000
countryBhutan                                1.855e+00  1.048e+05   0.000    1.000
countryBolivia                              -2.299e+01  6.838e+04   0.000    1.000
countryBrazil                               -2.113e+01  6.838e+04   0.000    1.000
countryBurkina Faso                         -2.180e+01  6.838e+04   0.000    1.000
countryBurundi                              -2.264e+01  6.838e+04   0.000    1.000
countryCambodia                             -2.191e+01  6.838e+04   0.000    1.000
countryCameroon                             -2.292e+01  6.838e+04   0.000    1.000
countryChile                                 3.170e+01  2.536e+07   0.000    1.000
countryChina                                 2.200e-01  6.874e+04   0.000    1.000
countryColombia                             -2.359e+01  6.838e+04   0.000    1.000
countryCongo                                 2.026e-01  6.874e+04   0.000    1.000
countryCosta Rica                           -2.215e+01  6.838e+04   0.000    1.000
countryCote D'Ivoire                         1.921e+01  1.048e+05   0.000    1.000
countryDominican Republic                   -1.814e+01  6.838e+04   0.000    1.000
countryEcuador                              -2.222e+01  6.838e+04   0.000    1.000
countryEgypt                                -2.145e+01  6.838e+04   0.000    1.000
countryEl Salvador                          -2.386e+01  6.838e+04   0.000    1.000
countryGeorgia                              -2.352e+01  6.838e+04   0.000    1.000
countryGhana                                -2.247e+01  6.838e+04   0.000    1.000
countryGuatemala                            -2.255e+01  6.838e+04   0.000    1.000
countryHaiti                                -2.067e+01  6.838e+04   0.000    1.000
countryHonduras                             -2.332e+01  6.838e+04   0.000    1.000
countryIndia                                -2.145e+01  6.838e+04   0.000    1.000
countryIndonesia                            -2.281e+01  6.838e+04   0.000    1.000
countryIraq                                 -2.006e+01  6.838e+04   0.000    1.000
countryIsrael                               -1.771e+01  6.838e+04   0.000    1.000
countryJordan                               -2.238e+01  6.838e+04   0.000    1.000
countryKenya                                -2.331e+01  6.838e+04   0.000    1.000
countryKosovo                               -2.256e+01  6.838e+04   0.000    1.000
countryKyrgyzstan                           -2.332e+01  6.838e+04   0.000    1.000
countryLao People's Democratic Republic     -1.871e+01  6.838e+04   0.000    1.000
countryLebanon                              -2.235e+01  6.838e+04   0.000    1.000
countryLesotho                              -1.972e+01  6.838e+04   0.000    1.000
countryLiberia                              -1.983e+01  6.838e+04   0.000    1.000
countryMadagascar                           -2.026e+01  6.838e+04   0.000    1.000
countryMalawi                               -2.116e+01  6.838e+04   0.000    1.000
countryMali                                 -2.301e+01  6.838e+04   0.000    1.000
countryMauritania                            5.613e+00  1.048e+05   0.000    1.000
countryMexico                               -2.202e+01  6.838e+04   0.000    1.000
countryMoldova                              -2.118e+01  6.838e+04   0.000    1.000
```

Summary of logistic regression model continued,

```
countryMongolia                            -2.107e+01  6.838e+04   0.000    1.000    
countryMozambique                          -2.303e+01  6.838e+04   0.000    1.000    
countryMyanmar (Burma)                     -2.252e+01  6.838e+04   0.000    1.000    
countryNamibia                              3.853e-01  7.367e+04   0.000    1.000    
countryNepal                               -1.771e+01  6.838e+04   0.000    1.000    
countryNicaragua                           -2.336e+01  6.838e+04   0.000    1.000    
countryNigeria                             -2.266e+01  6.838e+04   0.000    1.000    
countryPakistan                            -2.384e+01  6.838e+04   0.000    1.000    
countryPalestine                           -2.160e+01  6.838e+04   0.000    1.000    
countryPanama                              -1.554e+01  6.838e+04   0.000    1.000    
countryParaguay                            -2.213e+01  6.838e+04   0.000    1.000    
countryPeru                                -2.226e+01  6.838e+04   0.000    1.000    
countryPhilippines                         -2.267e+01  6.838e+04   0.000    1.000    
countryRwanda                              -2.273e+01  6.838e+04   0.000    1.000    
countrySaint Vincent and the Grenadines     9.399e-01  7.173e+04   0.000    1.000    
countrySamoa                               -2.372e+01  6.838e+04   0.000    1.000    
countrySenegal                             -2.281e+01  6.838e+04   0.000    1.000    
countrySierra Leone                        -2.319e+01  6.838e+04   0.000    1.000    
countrySolomon Islands                     -2.146e+01  6.838e+04   0.000    1.000    
countrySomalia                             -1.945e+01  6.838e+04   0.000    1.000    
countrySouth Africa                        -1.719e+01  6.838e+04   0.000    1.000    
countrySouth Sudan                         -2.083e+01  6.838e+04   0.000    1.000    
countrySuriname                            -2.181e+01  6.838e+04   0.000    1.000    
countryTajikistan                          -2.326e+01  6.838e+04   0.000    1.000    
countryTanzania                            -2.232e+01  6.838e+04   0.000    1.000    
countryThailand                            -1.999e-01  6.867e+04   0.000    1.000    
countryThe Democratic Republic of the Congo -2.259e+01  6.838e+04   0.000    1.000    
countryTimor-Leste                         -2.253e+01  6.838e+04   0.000    1.000    
countryTogo                                -2.174e+01  6.838e+04   0.000    1.000    
countryTurkey                              -2.019e+01  6.838e+04   0.000    1.000    
countryUganda                              -2.295e+01  6.838e+04   0.000    1.000    
countryUkraine                             -2.102e+01  6.838e+04   0.000    1.000    
countryUnited States                       -2.180e+01  6.838e+04   0.000    1.000    
countryVanuatu                             -4.735e+01  7.897e+04  -0.001    1.000    
countryVietnam                             -2.257e+01  6.838e+04   0.000    1.000    
countryYemen                               -2.231e+01  6.838e+04   0.000    1.000    
countryZambia                              -2.197e+01  6.838e+04   0.000    1.000    
countryZimbabwe                            -2.185e+01  6.838e+04   0.000    1.000    
partner_id                                  1.329e-03  7.400e-05  17.960   <2e-16 ***
term_in_months                             -2.573e-02  5.958e-04 -43.179   <2e-16 ***
lender_count                               -1.535e-02  3.969e-04 -38.680   <2e-16 ***
borrower_gendersfemale,male                -3.976e-01  2.213e-02 -17.963   <2e-16 ***
borrower_gendersmale                       -7.373e-01  1.096e-02 -67.263   <2e-16 ***
repayment_intervalirregular                 7.216e-01  2.106e-02  34.254   <2e-16 ***
repayment_intervalmonthly                   1.996e-02  1.720e-02   1.161    0.246    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 516343  on 523115  degrees of freedom
Residual deviance: 332194  on 523022  degrees of freedom
AIC: 332382

Number of Fisher Scoring iterations: 22
```

Table 2: Country wise funding status of loans (from high to low)

	country	fund_status	Freq
60	Philippines	fully_funded	157607
35	Kenya	fully_funded	70379
12	Cambodia	fully_funded	33464
23	El Salvador	fully_funded	32723
55	Pakistan	fully_funded	24920
59	Peru	fully_funded	21618
79	Uganda	fully_funded	18250
16	Colombia	fully_funded	18033
72	Tajikistan	fully_funded	17529
21	Ecuador	fully_funded	12890
58	Paraguay	fully_funded	11580
30	India	fully_funded	10911
53	Nicaragua	fully_funded	10499
83	Vietnam	fully_funded	9880
54	Nigeria	fully_funded	9137
39	Lebanon	fully_funded	8094
8	Bolivia	fully_funded	7561
56	Palestine	fully_funded	7172
197	El Salvador	partially_funded	6927
3	Armenia	fully_funded	6807
27	Guatemala	fully_funded	6803

Table 3: Sector wise funding status of loans (from high to low)

	sector	fund_status	Freq
1	Agriculture	fully_funded	166501
7	Food	fully_funded	128579
12	Retail	fully_funded	114213
13	Services	fully_funded	41249
11	Personal Use	fully_funded	34923
5	Education	fully_funded	30235
3	Clothing	fully_funded	29754
9	Housing	fully_funded	29726
14	Transportation	fully_funded	13795
31	Agriculture	partially_funded	13138
2	Arts	fully_funded	11857
42	Retail	partially_funded	9662
8	Health	fully_funded	8609
37	Food	partially_funded	7383
10	Manufacturing	fully_funded	6162
4	Construction	fully_funded	5936
39	Housing	partially_funded	3871
43	Services	partially_funded	3300
33	Clothing	partially_funded	2784
44	Transportation	partially_funded	1601
41	Personal Use	partially_funded	1358

APPENDIX B

Data Exploration

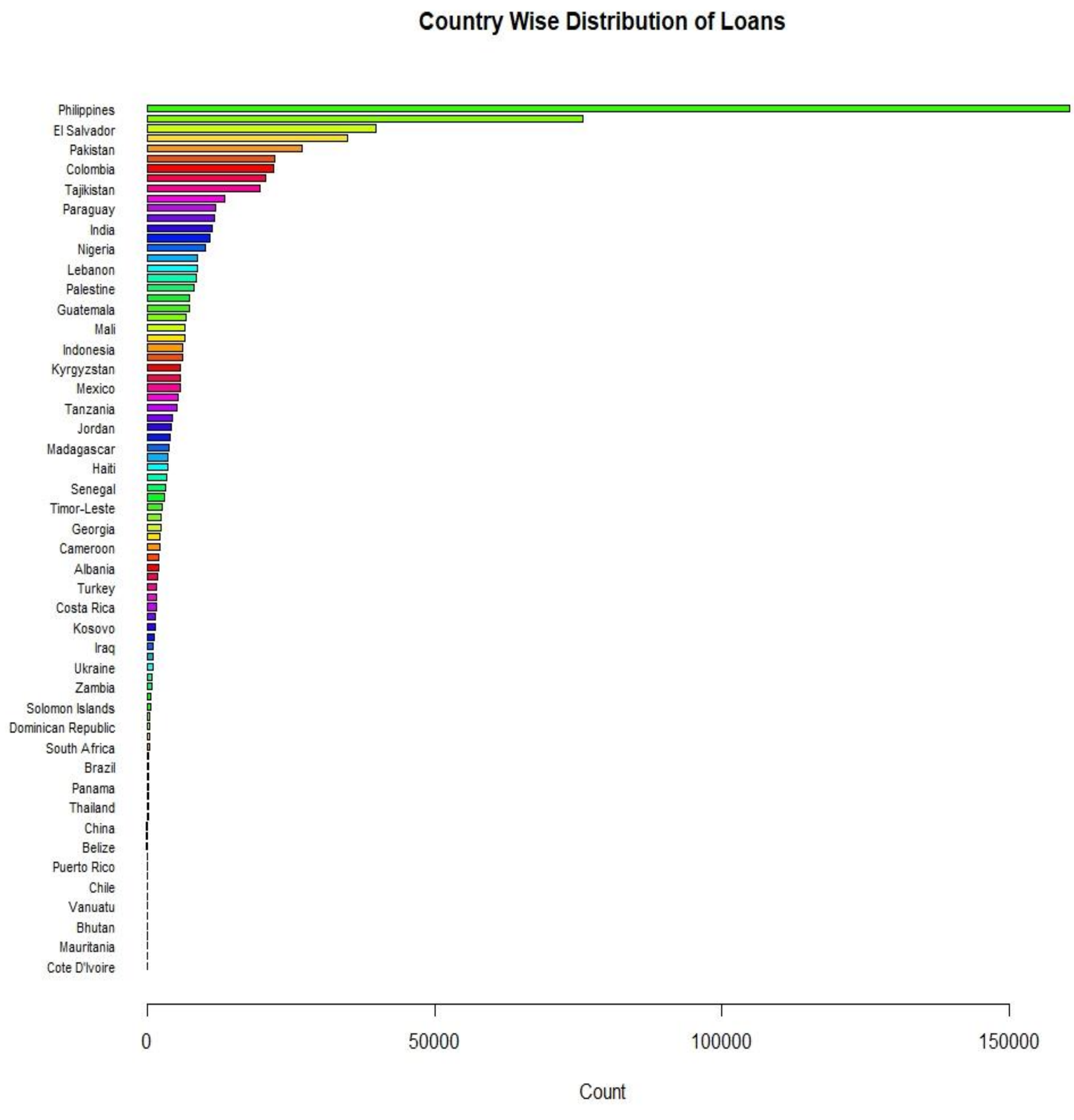

Figure 2: Country wise distribution of loans

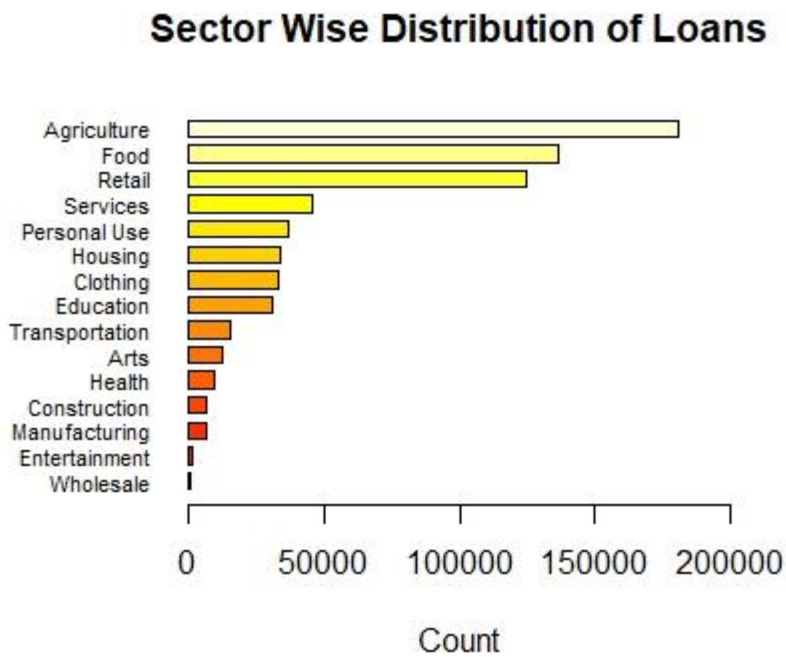

Figure 3: Sector wise distribution of loans

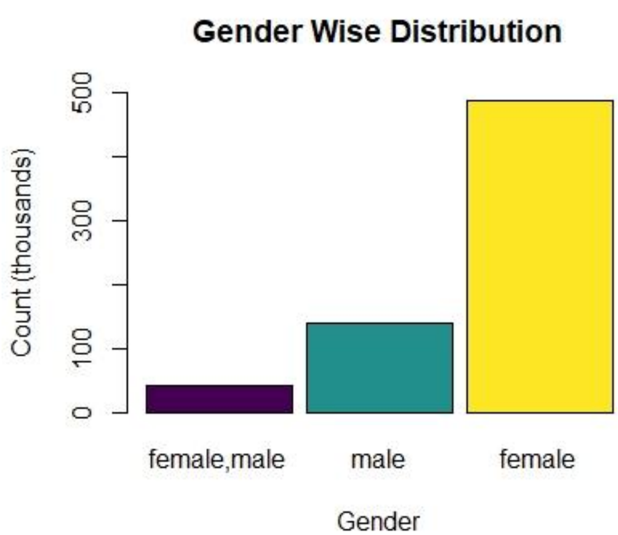

Figure 4: Gender wise distribution of loans

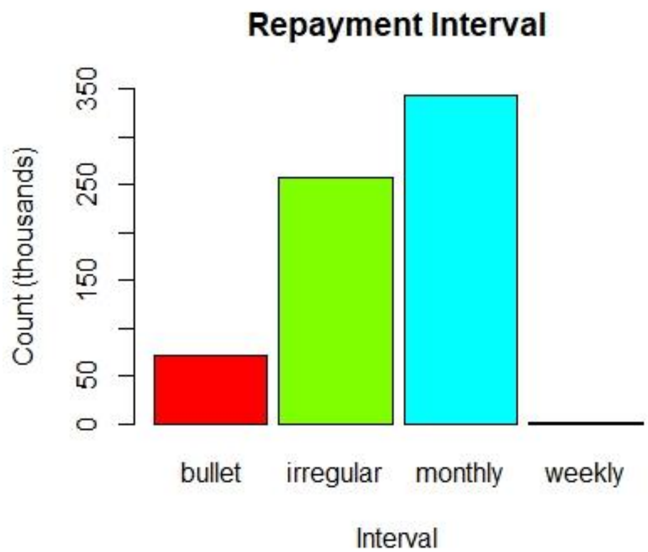

Figure 5: Frequency of repayment interval

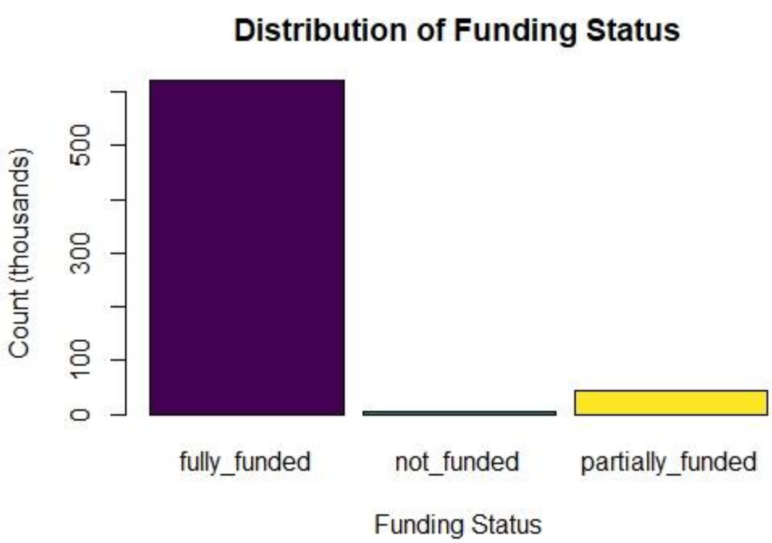

Figure 6: Frequency of funding status

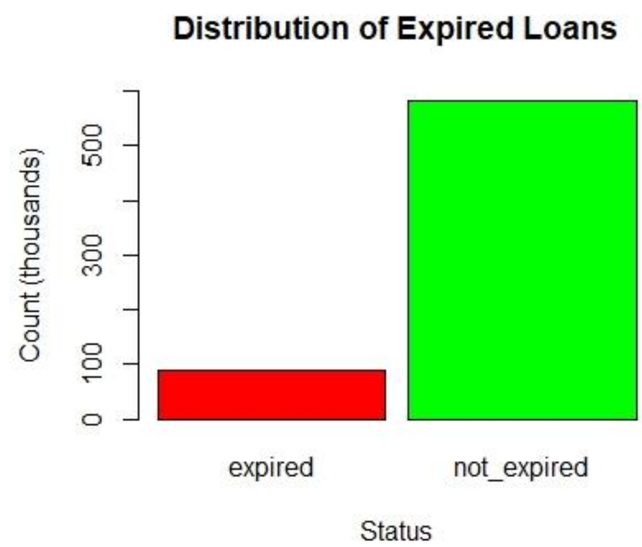

Figure 7: Frequency of expired loan status

Table 4: Frequency and percentage of expired loan status

	Default.Category	Total.Number	Percentage
1	Partially Funded	48328	7.20
2	Not Funded	3383	0.50
3	Time Expired	88912	13.25

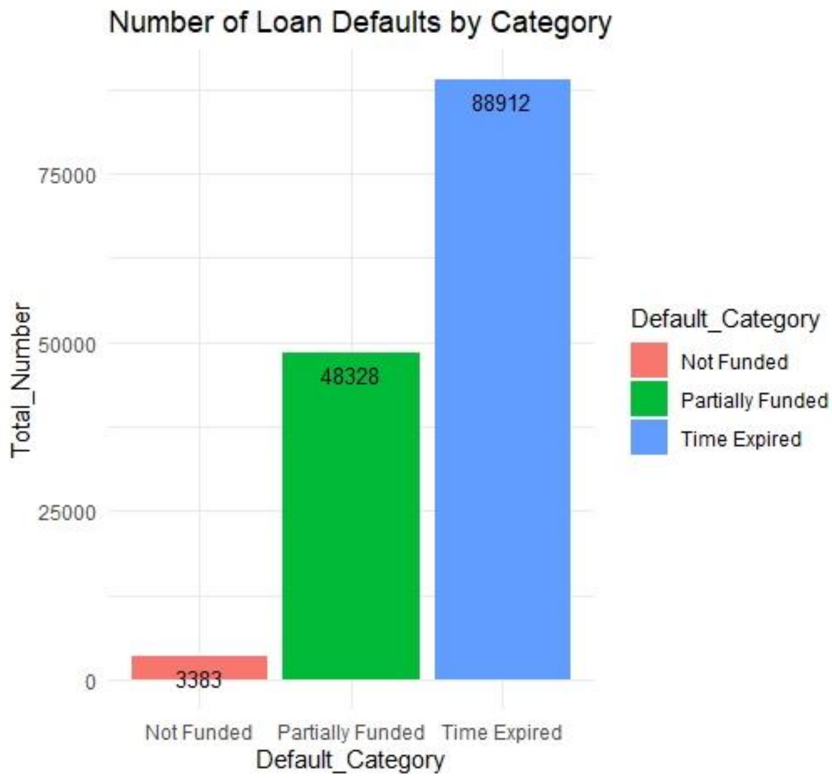

Figure 8: Distribution of default loans by category