NATIONAL UNIVERSITY OF SINGAPORE

DAO2702 DECISION ANALYTICS USING SPREADSHEET

DAO2702 Final Project Report

Session: TA15

Team Name: wed4pm

Prepared For: Xiong Peng

Ammar Bin Hussein Bagharib	A0218111X
Celine Siah Wen Hui	A0223445L
Cho Hui Wen	A0204805M
Guo Shiliang	A0189172W
Yee Xian Hao, Samuel	A0218207L

Content Page

1 Introduction

Established in 2006, LendingClub was an American peer-to-peer (P2P) lending company that links borrowers to investors. Due to its affordability and accessibility, P2P lending platforms have seen a rise in credit demand over the past few years. With the rise in credit demand came a rise in borrowers defaulting on their loans. As LendingClub profits by charging both borrowers and investors a fee, the high defaulting rate would discourage investors from using the platform, leading to a loss in profits.

Our objective is to maximise LendingClub profits by decreasing losses via measures to prevent default among high default risk borrowers, and to increase profits by encouraging low default risk borrowers to take up more loans

1.1 Methodology

The original dataset contains a total of 396030 rows and 27 variable columns. We did an analysis of the major categories deemed useful for our project against smaller variables to identify significant trends. Our target variable is "loan_status", which represents whether the customer is able to pay off his debt. This variable is one of the company's key parameters which directly affects profits. The variable had 2 categories, 'Fully Paid' and 'Charged off', representing loans that were paid off and loans that were defaulted respectively. We analysed significant variables in borrowers' profiles to examine trends between the variable and frequency of default.

1.2 Data cleaning

We kept 10 numerical and 6 categorical variables that we assumed to be more significant in influencing likelihood of default and removed the remaining 11 variables. Rows with null values were removed. For 3 variables, namely "dti", "annual_inc" and "total_acc", rows with outliers that were more than 1.5 times of the interquartile range were removed due to the extreme anomalies observed. The cleaned dataset contained 370705 rows and 16 variables which made up 93.61% of the original data.

2 Data Analysis

Using a Pie Chart, the dataset was split into 2 categories, loans with a 60 month and 36 month term (Figure 1). We then further examined the percentage of charged off and fully paid loans for each category (Figure 2). Results show that longer term loans have double the proportion of loans being charged off. This may be attributed to the fact that longer term loans are usually preferred by borrowers with poor credit ratings due to lower monthly payments. Our team has chosen to focus our analysis on the 60 month loans as there is great room for improvement in terms of lowering charged off rates.

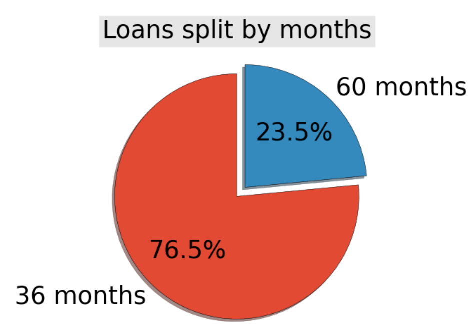

Figure 1: Pie Chart of Loans Split by Term

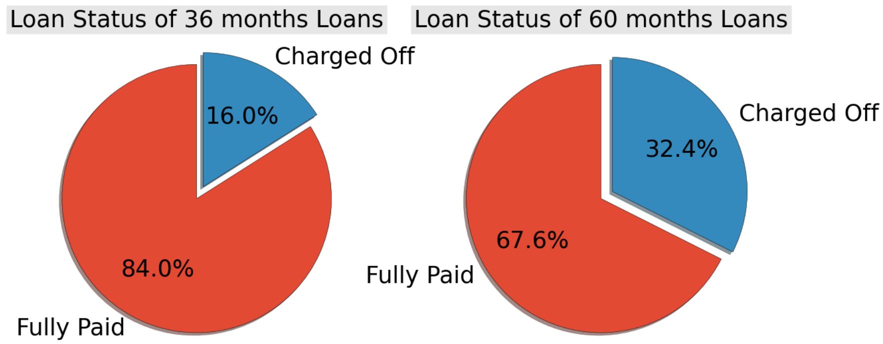

Figure 2: Pie Chart of Charged Off vs Fully Paid Split by Term

2.1 Debt to Income Ratio (DTI)

There is a positive relationship between debt to income ratio and frequency of charged off loans. When debt to income increases, frequency of charged off loans increases steeply (Figure 3). This is expected because higher DTI ratio indicates that the borrower is holding more debt relative to his income level and is less likely to meet his current debt obligations with LendingClub considering his level of income. Furthermore, borrowers with high DTI ratios may have taken up loans to clear their existing debt obligations. Another possible reason would be that they've consolidated their various debts into a consolidated loan under LendingClub, but they ended up finding it even more difficult to pay such a loan due to their level of annual income.

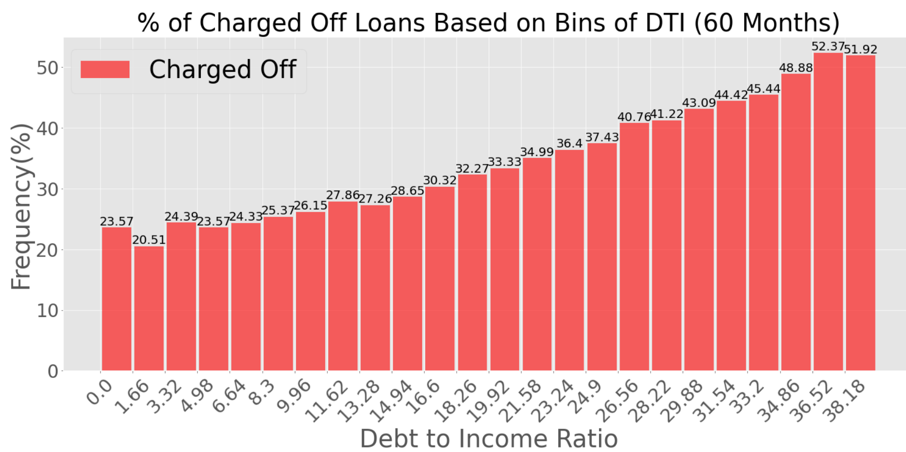

Figure 3: Bar Chart of DTI against Frequency of Charged Off Loans (%)

2.2 Open Accounts

The numerical variable "open_acc" refers to the number of open credit lines on the borrower's credit file that are unpaid or in the midst of repayment. We can see an increasing trend of charged off loans as the number of open accounts increase (Figure 4).

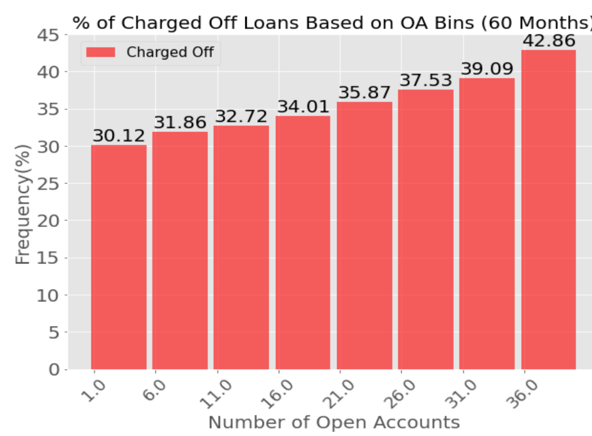

Figure 4: Bar Chart of Number of Open Accounts Against Frequency of Charged Off Loans (%)

Borrowers with more open accounts would typically find it more difficult to fulfil all of their repayment obligations. Alternatively, it may be because these borrowers did not have money to begin with, hence requiring many open accounts to fund multiple obligations. The rate of increase is consistent with the increase in open accounts. However, we've noticed that the increase in open accounts from 31-35 bracket to >= 36 bracket saw the greatest increase in frequency of charged off loans by 3.77% to 42.86%, compared to other bins.

2.3 Annual Income

In general, there is a negative relationship between the percentage of charged off loans incurred by borrowers and their respective annual incomes (Figure 5). The percentage of charged off 60 months loans increases with a decrease in annual income. This may be because borrowers with higher annual income have higher disposable income, hence putting them in a better position to repay their loans.

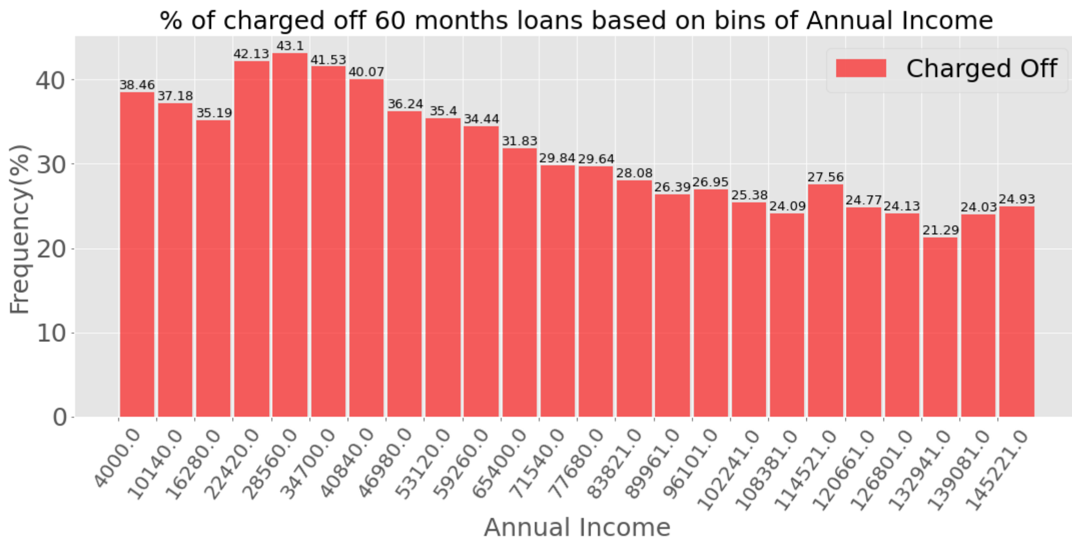

Figure 5: Bar Chart of Annual Income against Frequency of Charged Off Loans

In addition, the purpose of taking on the loan may differ between borrowers with higher and lower annual incomes. We assume borrowers with lower income to take up loans out of necessity, e.g. to pay off hospital bills. Thus, we'll explore this group further below.

2.4 Purpose

The categorical variable 'purpose' contains 14 categories showing the purpose of the loan. As seen in Figure 6, debt consolidation constituted the main purpose for the large majority of LendingClub loans at 59.52%, followed by credit card loans at 20.98%.

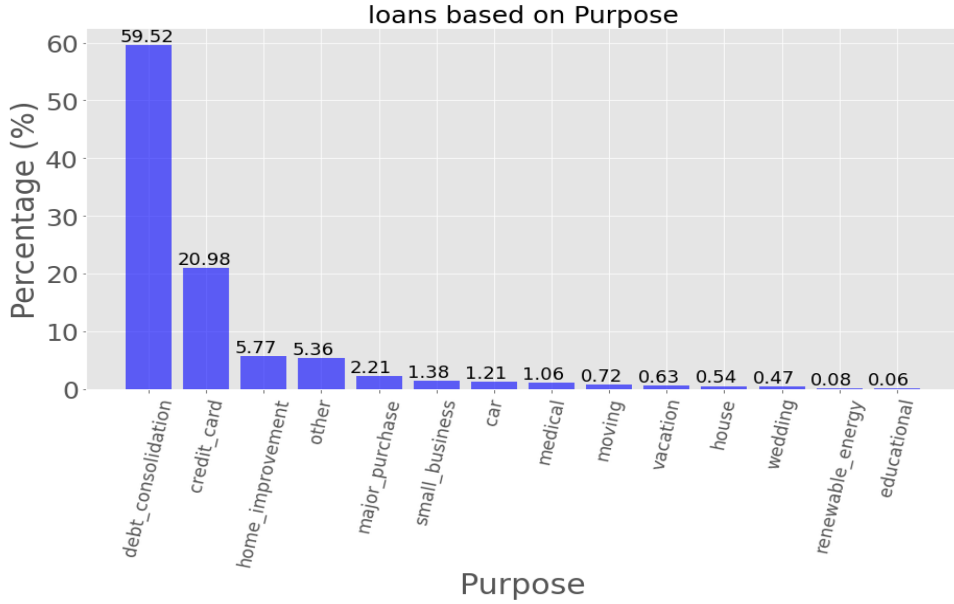

Figure 6: Bar Chart of Purpose against Percentage of Loans for each purpose

For our analysis, we further grouped up the rest of the purposes into five broad categories:
Events = [‘wedding’, ‘vacation’, ‘other’]
Necessity = [‘medical’, ‘educational’]
Home_related = [‘home_improvement’, ‘house’, ‘moving’, ‘car’]
Business_related = [‘renewable_energy’, ‘major_purchase’, ‘small_business’]
Home = [‘house’, ‘moving’, ‘home_improvement’, ‘car’]

This is done to increase the number of records to facilitate more accurate and significant analyses, as the other purposes only occupied a small percentage of the data set. We’ve decided to narrow our analysis to the Necessity category for 60 month loans. We’ve chosen as such due to our initial assumption from our previous analysis on charged off loans against annual income bins (refer to 2.3). For the rest of the categories, we’ve attached our graphs in Appendix A.

Interestingly, the ‘Necessity’ category observed the greatest number of anomalies, with 100% default rate by borrowers from the 6-12k and 17.9-23.8k income brackets (Figure 7). Furthermore, borrowers from the 11.9k-17.9k income brackets have a 50% default rate. This suggests that borrowers from these 3 income brackets who take on 60 month loans for medical and educational purposes have an extremely high chance of defaulting. One possible explanation for this could be that these lower income group borrowers may take on loans out of necessity, such as to pay off hospital bills even if they are unable to repay them.

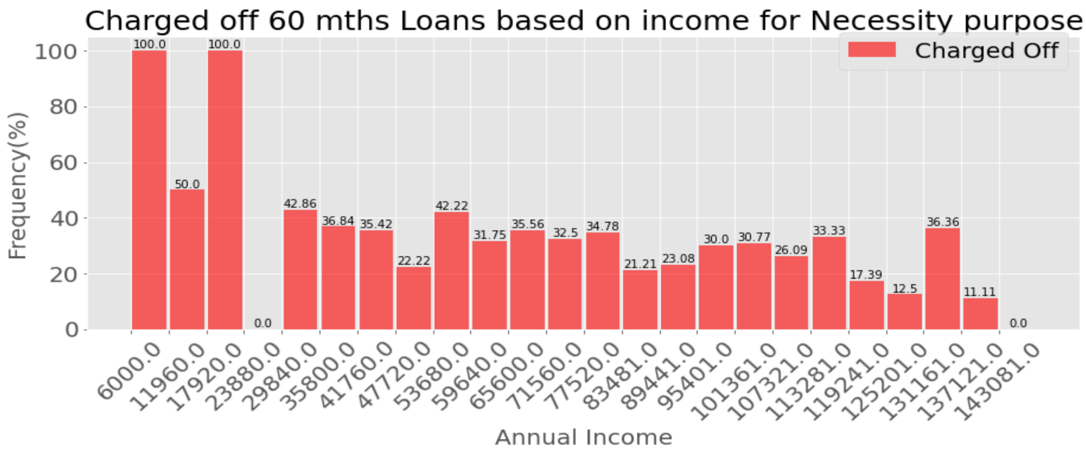

Figure 7: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for Necessity purpose

2.5 Home Ownership

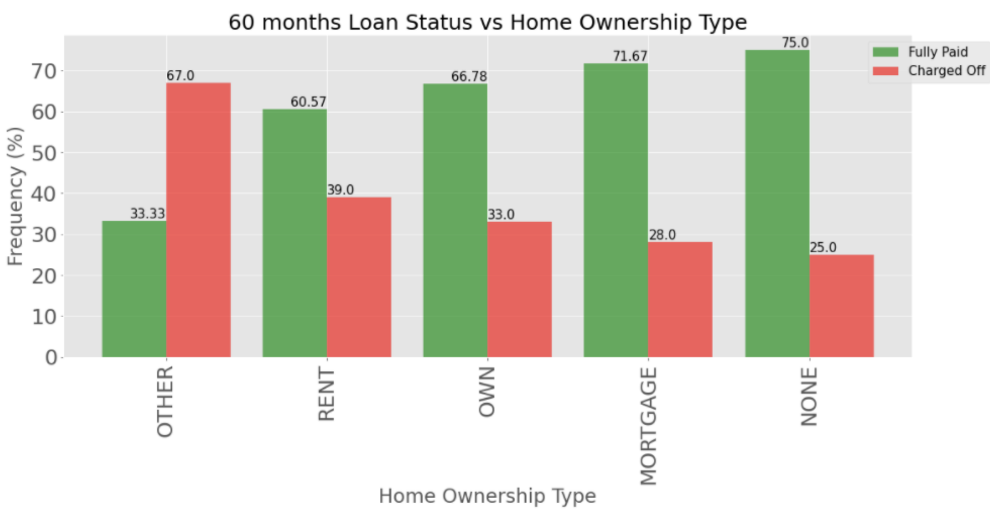

Figure 8: Bar Chart of Home Ownership Types against Frequency of Charged Off

Borrowers fall under 5 home ownership categories: ‘Other’, ‘Rent’, ‘Own’, ‘Mortgage’ and ‘None’. Figure 8 shows that borrowers that fall under the ‘Other’ category incur the highest rate of charged-off loans at 67%. The frequency of charged off loans decreases across the ‘Rent’, ‘Own’, ‘Mortgage’ and ‘None’ categories from 39% to 25%. It can be seen that borrowers under the ‘Rent’ category have the second highest frequency of charged off loans at 39% (Figure 8). This may be because the extra obligation to pay rent eats into their disposable income and hence their ability

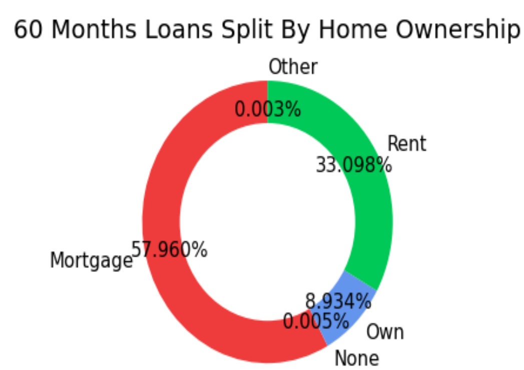

Figure 9: Pie Chart Showing Distribution of 60 Months Loan

to pay off their loans. We also noticed that there are very few ‘Other’ and ‘None’ home ownership types (Figure 9). Furthermore, it is not clear as to what the ‘None’ category refers to. Thus, we focused on the ‘Rent’, ‘Mortgage’ and ‘Own’ categories. We further narrowed on our analysis on ‘Rent’ home ownership types since we assume that these groups of people would be less capable in meeting fully paying their 60 months loans as they are financially less stable. For the ‘mortgage’ and ‘own’ groups, we’ve added our graphs in Appendix B.

It can also be seen that as annual income increases, there is a general decrease in frequency of charged off loans among borrowers who rent (Figure 10). Excluding the income brackets of 12k - 24k, the income brackets between 6k - 42k has a relatively high default rate of > 46%.

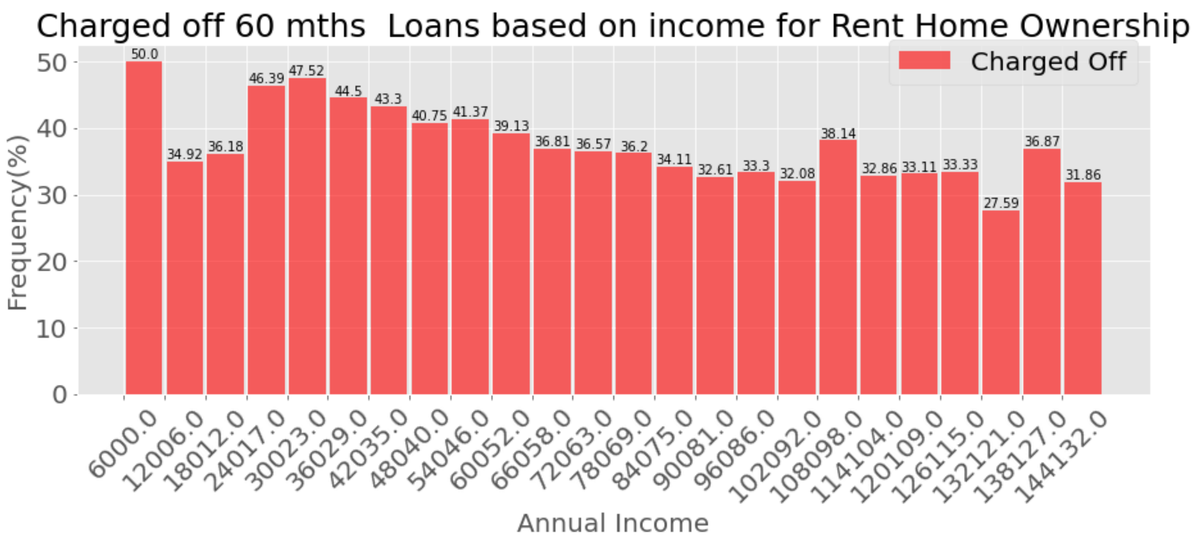

Figure 10: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for ‘Rent’ Home Ownership

2.6 Subgrades

This variable refers to the credit rating given according to LendingClub’s evaluation of the borrower’s credit and default risk. As subgrades are given independently of loan term, we analysed all loans regardless of term, against the frequency of charged off loans. Naturally, we assume the charge off rates to increase with a drop in subgrades. Thus, we’ve narrowed down our analysis for subgroups E1 onwards to G5, assuming that these subgrades signal some of the more risky borrowers. Based on our results, we observed that the proportion of charged off loans increases as the subgrade level drops from E1 to G5. (Figure 11). However, interestingly, as the subgrade falls, the percentage of fully paid and charged off loans begins to equalise. For grades ‘G3’ and ‘G5’, the percentage of charged off loans exceeds fully paid loans. For a graph of all the subgrades against charge off rates, we’ve added it under Appendix C.

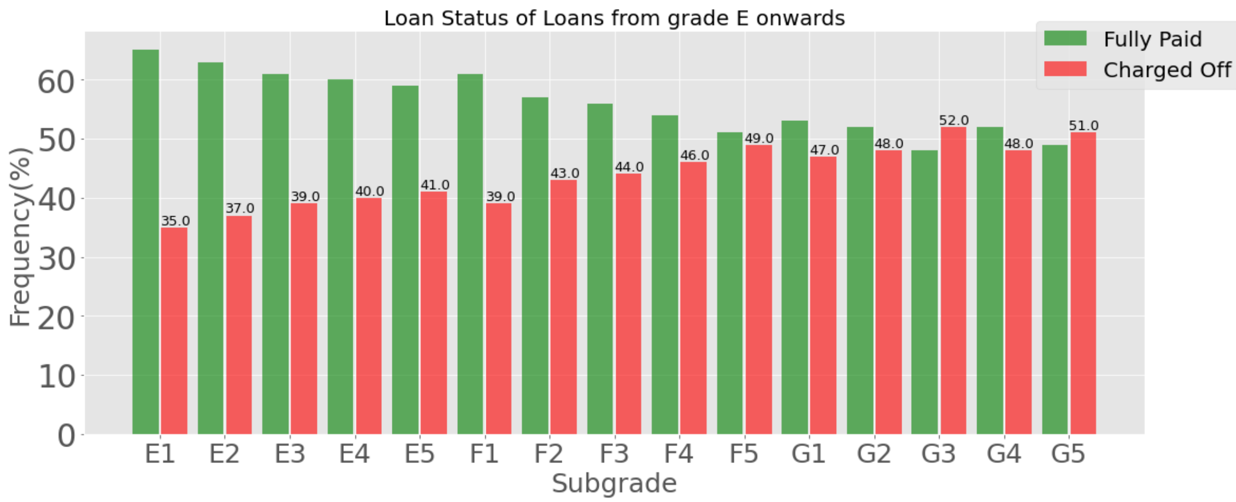

Figure 11: Bar Chart of Frequency of Charged Off and Fully Paid Loans subgrade E1 onwards

3 Recommendations

The frequency of charged off loans should remain sufficiently low to attract potential investors who are concerned about the risk of their loans being defaulted. Given LendingClub's bottomline as a loan platform, it earns revenue via loans handled, and it goes against its business model to severely restrict borrowers. However, if the number of loans being charged off goes too high, it risks hurting its reputation as an alternative investment option for lenders on its platform. Hence, our recommendations will mostly revolve around identifying and changing rules for people who are statistically more and less likely to repay their loans, and label them as high and low risk.

3.1 Recommendations For Customers with High Default Risk

We've decided upon a 45% charge off rate as the frequency threshold for 3 different factors: 'DTI', 'Rent Home Ownership' and 'Subgrades'. Upon analysis of the data, we found that for several factors in the data, there is a clear increasing trend of charged off proportion in its corresponding markers.

3.1.1 45% charged-off threshold for 'DTI', 'Rent' and 'Subgrades'

We decided on a 45% charged off frequency to show investors that a majority of loans are eventually fully paid for, and imposed the rule to distinct categories we identified that had a clear upward trend of increased charged off frequency for its corresponding data. We narrowed our observations to 3 variables: DTI, 'Rent' Home ownership and 'Subgrades'. From analyses done above (based on the 45% default rate threshold), we recommend LendingClub to limit the number of loans and to enhance the screening methods employed against these groups of borrowers who would take up 60 months loans in the future: With references to our data analysis based on the 45% charged off rate in Appendix D, we recommend the customers with:

1. Debt to income ratios above 33.2

2. Annual incomes of \$18,012 <= x <= \$30,000 and x <= \$12,006, where x represents annual income, whose home ownership types are listed as 'Rent'.
3. Subgrade F4 to G5

Upon limiting loans given out to these groups, Lending Club can also factor in their existing number of open accounts when choosing whom from amongst these high risk groups to give out loans to. Referencing from the findings in 2.2, and figure 4, LendingClub should minimise loans given out to those with the least number of open accounts as an added safety precaution.

3.1.2 Installment reliefs to borrowers who took up loans for 'Necessity' Purposes

With reference to figure 7, we also recommend LendingClub to introduce a debt moratorium to borrowers who took up 60 month loans for 'Necessity' reasons with an annual income of less than \$23,800. As these groups have extraordinarily high charged off rates, LendingClub can look into changing the debt repayment procedure for such borrowers, with focus on debt consolidation in order to return lenders their principle.

Such groups may find it difficult to fully pay the loans should we raise their interest rates. Thus, we recommend them to be entitled to installments reliefs of a minute percentage of up to 5 times i.e, for every 12 months worth of installments that they successfully pay off. These loan reliefs given by LendingClub are aimed to incentivise these groups to meet their debt obligations with LendingClub in a timely manner to enjoy such discounts.

3.2 Recommendations For Customers with LowDefault Risk

3.2.1 Reduced fees for responsible borrowers who have previously paid off a loan

As a peer to peer lending site, it benefits from the full repayment of loans by the borrowers as it gives better return to lenders and are more assured to put in more money. Thus, we recommend them to consider incentivising borrowers who have fully completed their loans to be eligible for a reduced origination fee (platform fee) for their next loan. This seeks to encourage responsible borrowers who have previously fully paid off their loans to continue borrowing from the platform.

3.2.2 Increased loan limit for low risk existing borrowers

We also recommend LendingClub to consider raising the loan limit for borrowers who have been responsible in paying off their loans within the categories of our analyses done above. This can be done by allowing such borrowers to borrow more than \$40,000 (the current maximum loan amount). For starters, to avoid unprecedented high risk borrowers from applying for such loans, LendingClub should only extend this option to their existing and reliable users who are of lower risk across most of the variables we analysed aboved. E.g. A prime customer would be a customer with a low number of open accounts and DTI ratio, and one who belongs to the higher annual income brackets, and who originally has paid off their existing 60 months loans. LendingClub should eventually publicise these metrics to incentivise larger groups of potential borrowers who fulfill such criteria in taking up loans under them in the long run.

4 Appendix

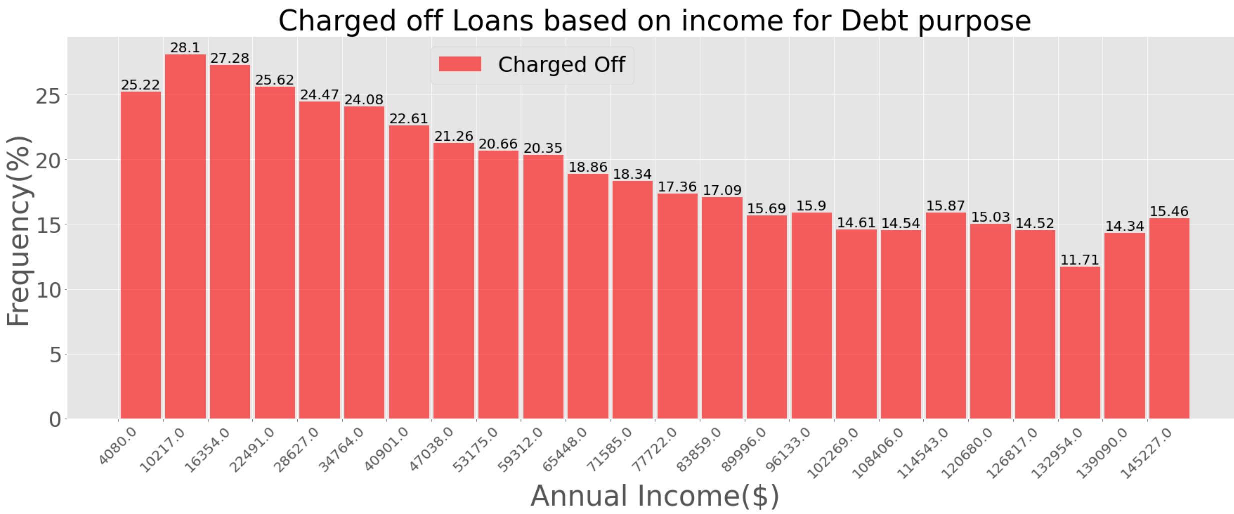

Figure A.1: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for Debt purpose

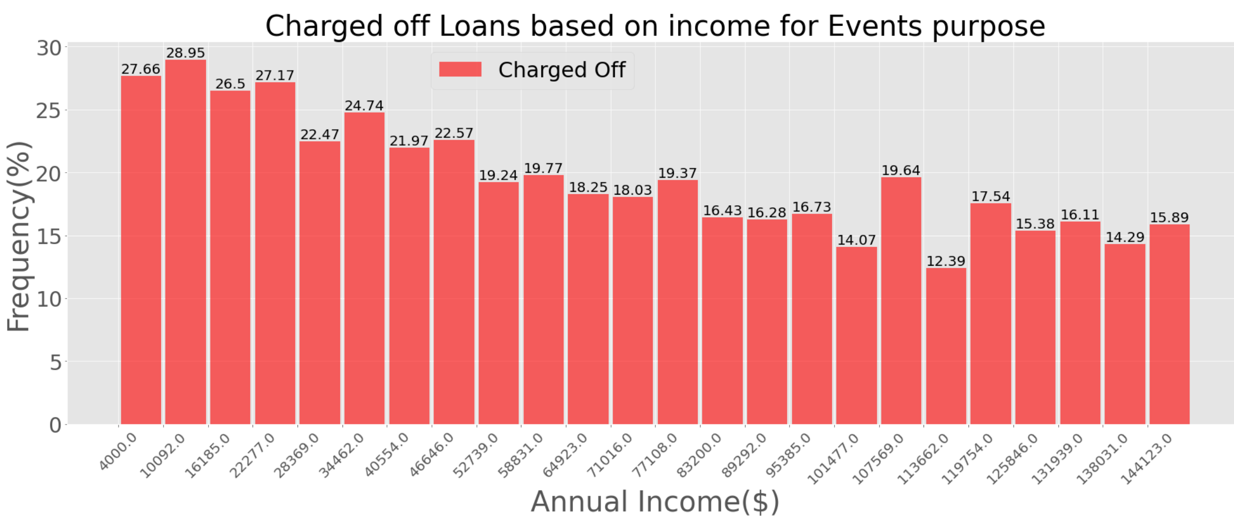

Figure A.2: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for Events purpose

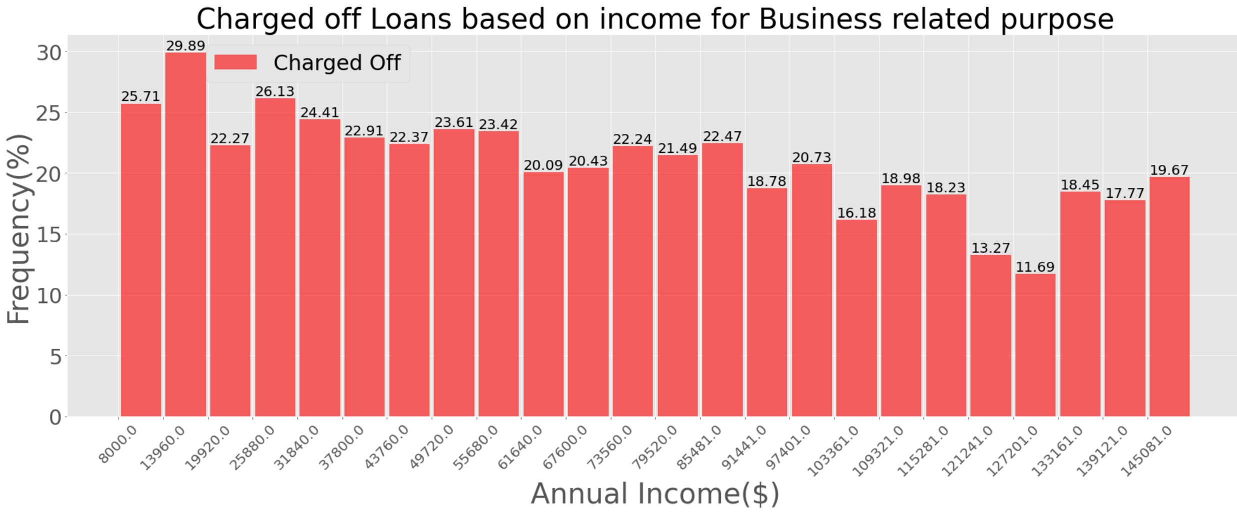

Figure A.3: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for Business-related purpose

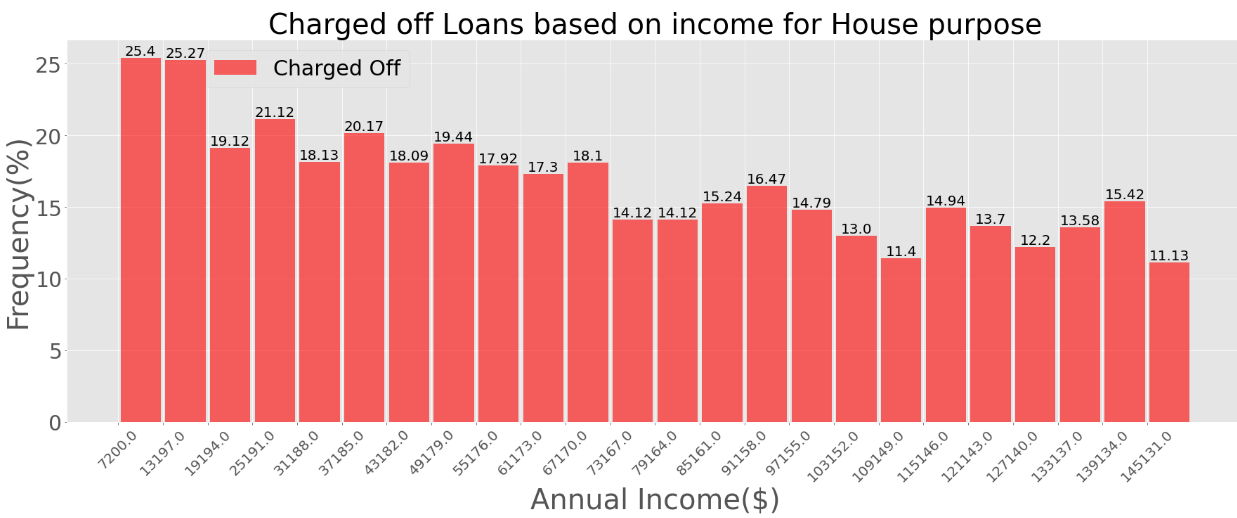

Figure A.4: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for House purpose

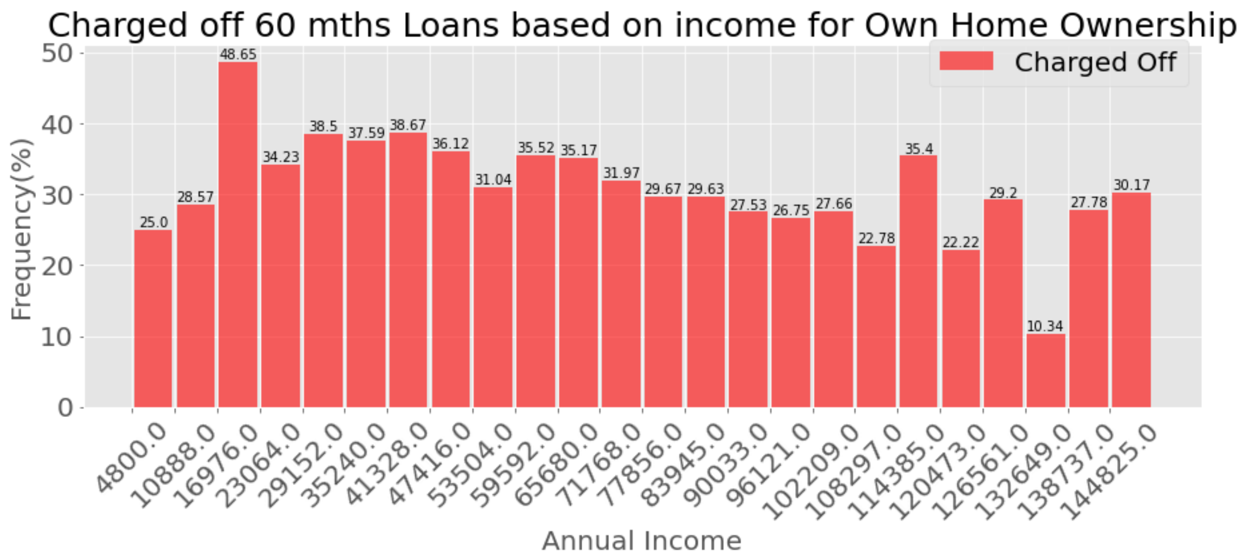

Figure B.1: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for Own Home Ownership

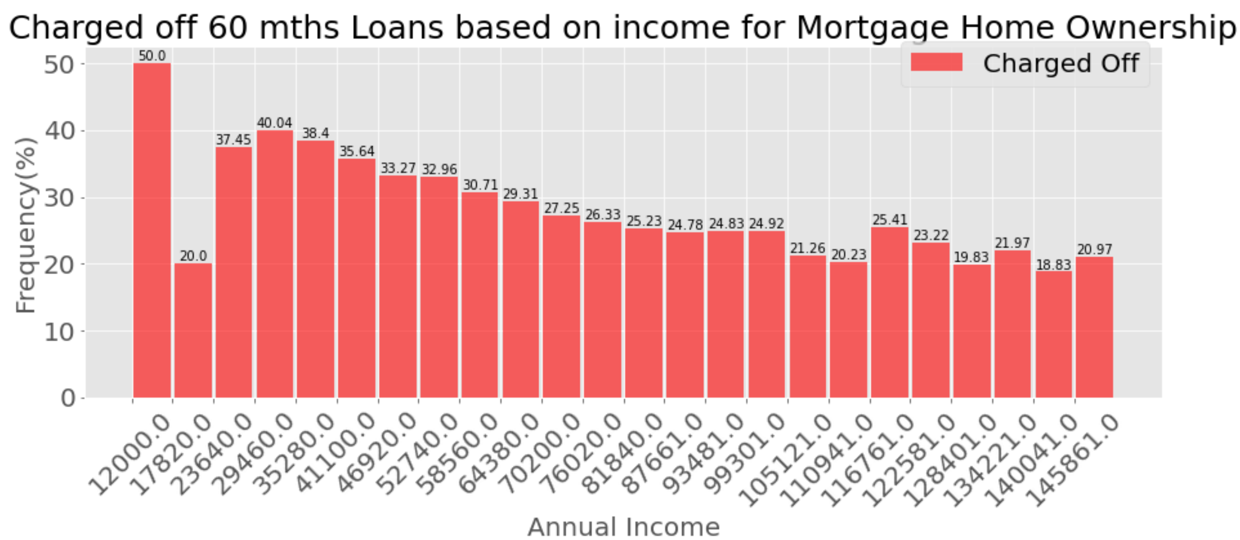

Figure B.2: Bar Chart of Annual Income bins against Frequency of Charged Off 60 month Loans for Mortgage Home Ownership

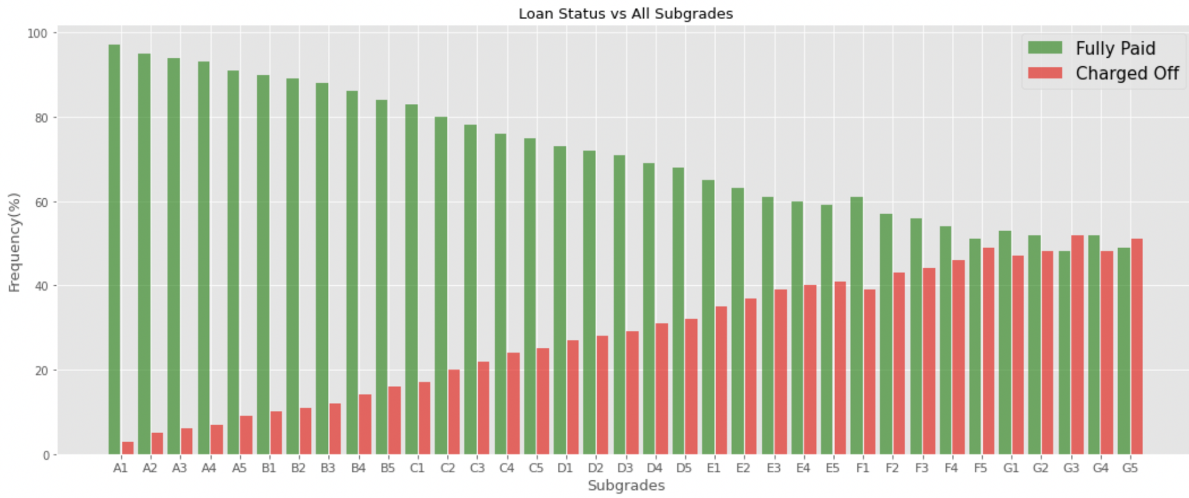

Figure C.1: Bar Chart of frequency of charged off loans against all subgrades

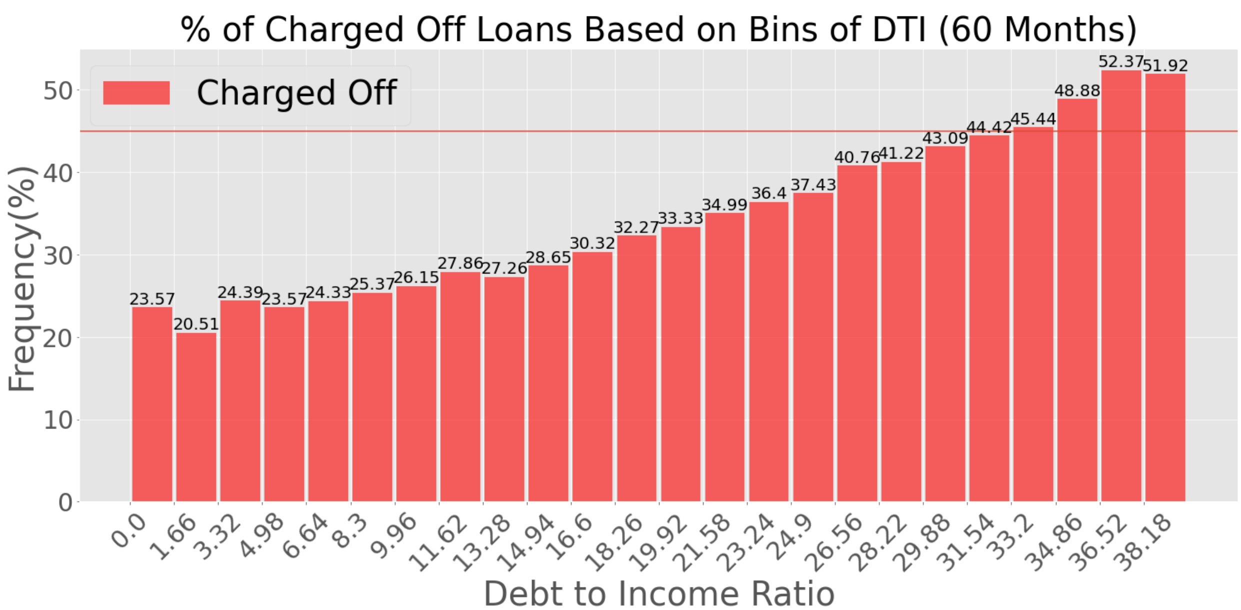

Figure D.1: Bar Chart Showing Cut-Off for DTI above 33.2

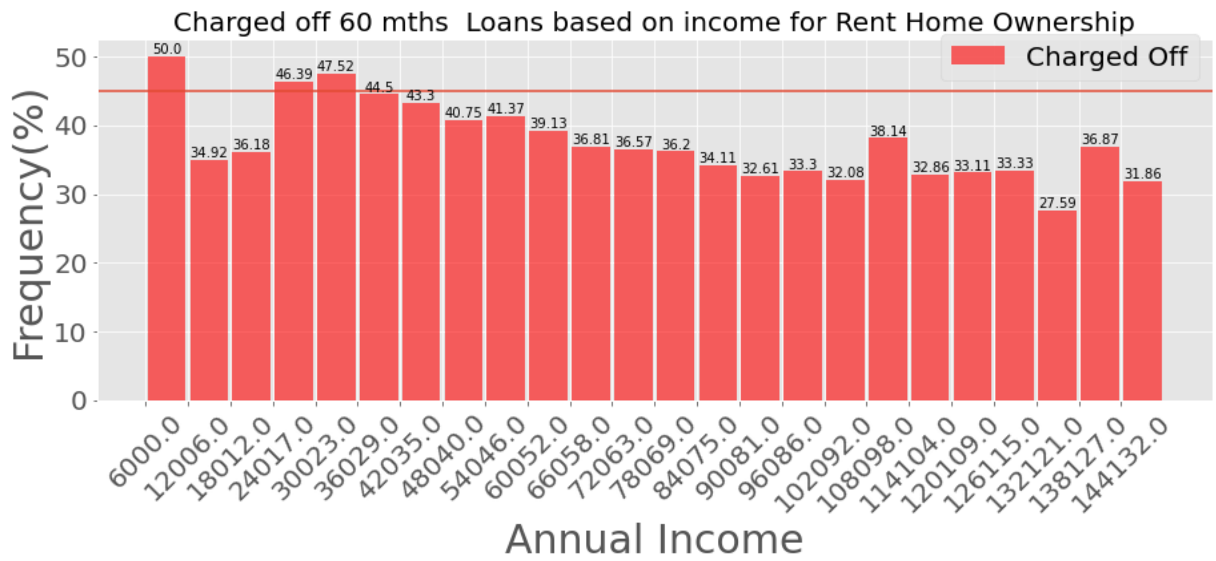

Figure D.2: Bar Chart Showing Cut-Off for ‘Rent’ Borrowers with Annual Income < $30023

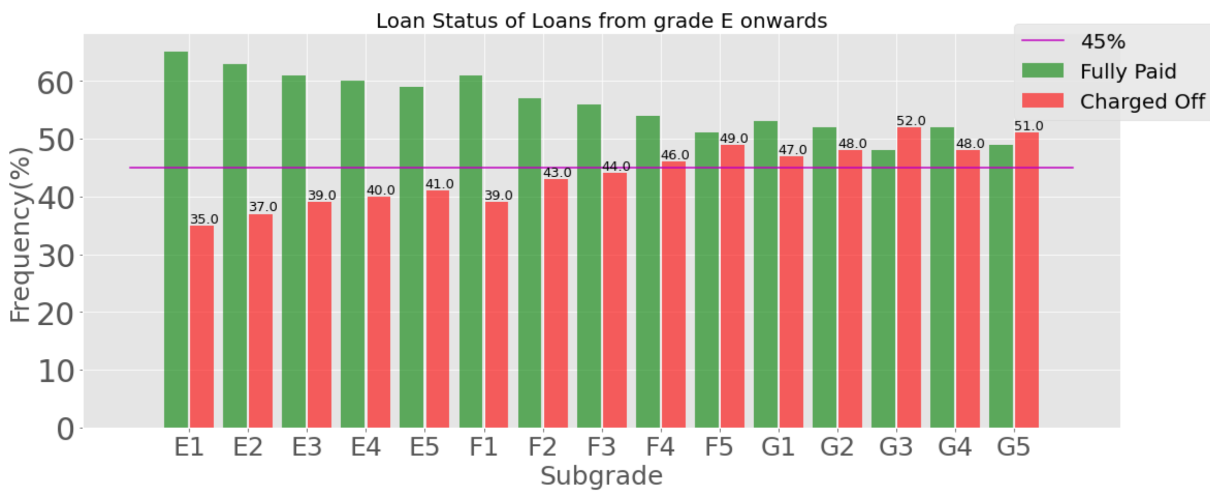

Figure D.3: Bar Chart Showing Cut-Off for Subgrades F4 to G5