23rdOctober 2024

Lending Club Case Study

Submitted By :

Prashant Saxena

Pratham Desai

Table of Content

Executive Summary

Problem Statement

- The Consumer Finance Company specializes in lending various types of loans to urban customers
- When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile
  - If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
  - If the applicant is not likely to repay the loan, , then approving the loan may lead to a financial loss for the company
- The company needs a framework to evaluate the loan applications, to improve the profitability of the business

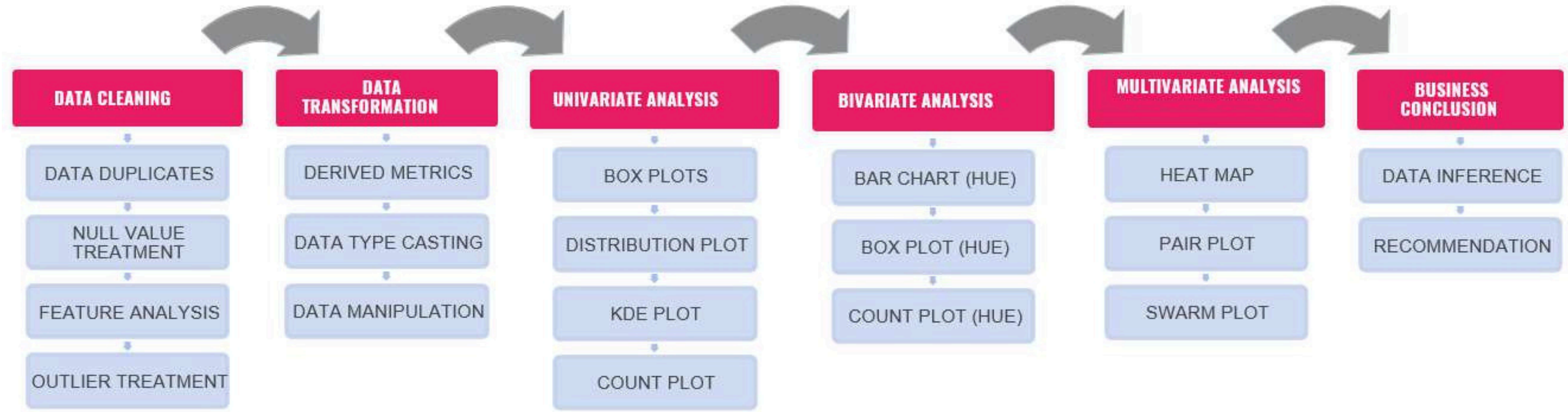
Objective of the Analysis

- To identify the key drivers impacting the default behaviour of borrowers
  - Also understand the interrelationship between variables that impact default behaviour
- To identify the key drivers impacting the recoveries from defaulted loans
- To prioritize the drivers in terms of their impact, and suggest ways to assess loan application to maximize business profitability

Recommendations for the Business

- Caution must be exercised in granting lower grade loans, especially graded 'D' and below, which experience more than 20% default rate
  - Additional measures such as physical verification, income documentation, etc should be taken for such loans
- Loans graded 'D' and below should be charged even higher interest rates to compensate for substantially higher chances of default
- High income group borrowers (typically above $90k) should be incentivized to apply for more loans due to better risk-adjusted returns, given the low default rate
- Debt to Income (DTI) ratio is a primary indicator of default that must be considered and verified in granting loans
  - Higher DTI increases the chances of default
- If an applicant has made multiple inquiries in the last 6 months, loan is likely to have some defect due to which it was probably rejected
  - Such loans should go through an extra layer of due diligence
- Loans with large installment sizes (typically higher than 200 USD) should be regularly monitored, and if required restructuring plans should be considered to reduce incidences of default due to budget related burden on borrowers
- Lower rates of interest should be offered for 3 yr loans in view of lower defaults in this bucket

Methodology



- Data Cleaning
  - Outlier treatment - Identify and treat outliers by calculating IQR (Q3-Q1), then remove data points outside 1.5 * IQR from Q1 and Q3
  - Data conversion – '%' symbol is removed from variables such as 'int_rate' and 'revol_util' columns to treat them as float
- To understand the default behaviour, only closed loans are analyzed - 'Charged Off' and 'Fully Paid' loans
  - 'Current' loans are excluded because their default status cannot be determined yet.
- Derived metrics such as default rate and recovery rate are calculated to understand their relationship with other variables
- Default rate has been calculated as 'Charged Off' loans as percentage of total closed loans
- Default rate has been calculated in terms of both:
  - Count of loans defaulted
  - Amount of loans defaulted
- Recovery Rate has been calculated as 'recoveries' as percentage of defaulted amount
  - Defaulted amount is considered as 'loan_amnt' minus 'total_rec_prncp', which denotes the principal amount that has been charged off.
- Univariate, Bivariate and Multivariate analysis is performed to explain relationship among variables
  - Interdependency of variables is also analyzed avoid any misleading conclusion regarding any variable driving default rate
- Correlation and regression techniques have been applied to establish relationship among variables

Key Findings –Univariate, Bivariate and Multivariate Analysis

Grade	•Lower grade loans are more likely to make defaults •Low grade borrowers tend to apply for long duration loans •Interest rates are higher for lower grade loans
Term	•Default rates are noticeably higher in 5 yr term bucket compared to 3 yr bucket •Higher default rate for 5 yr term loans is primarily driven by higher composition of lower grade loans in this bucket •5 yr term bucket comprises of ~49% loans (by count) graded below 'C', as compared to only ~15% in 3 yr bucket.
Interest Rate	•Higher default rate is observed in high interest rateloans •Lower grade loans are usually offered higher interest rates to compensate for higher credit risk
Annual Income	•High income borrowers certainly have lower defaults •Default rates are consistently lower as the income of borrowers increase
Debt to Income (DTI) ratio	•Borrowers with higher DTI tend to have higher default rates •DTI is a leading indicator of default that represents the ability of borrower to repay the loan installments
Revolving Line Utilization Rate	•Default rates are higher for borrowers with higher revolving line utilization rate •Borrowers tend to utilize the revolving line in time of crisis, or to repay the debt
Inquiry in last 6 months	•As a general trend, default rates are higher for borrowers with higher inquiries in last 6 months •Multiple inquiries also suggest that loans for such borrowers must have been rejected due to some defect by multiple lenders, which also indicate higher chances of default
Purpose	•Small business loans have significantly high default rates •Disproportionally large amount of loans are made for purpose of debt consolidation •Consumer loan such as purchases, car, credit card, wedding, etc have relatively lower default rates •Such loans are typically optional in nature and tend to be taken by borrowers with expected cashflows in future to repay the loan
Installment	•Beyond a certain threshold of ~200 USD, chances of default tend to increase with the installment size •Larger installments increases the burden on borrower, and may be subject to default in case of any adverse condition on borrower income

Conclusion

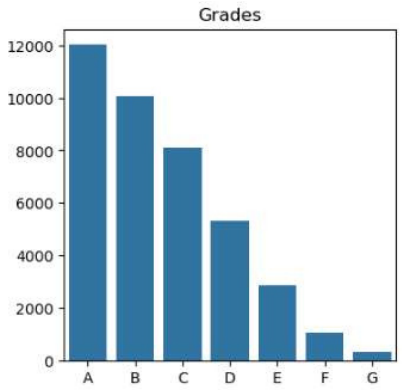
- Caution must be exercised in grantinglowergradeloans,especially graded ‘D’ and below, which experience more than 20% default rate
  - Additional measures such as physical verification, income documentation, etc should be taken for such loans
- Loansgraded ‘D’ and below should be charged even higher interest rates to compensate for substantially higher chances of default
- Borrowers in grades E, F, and G with an average loan amount over 15,000 are noteworthy.
- High income group borrowers (typically above $90k) should be incentivized to apply for more loans due to better risk-adjusted returns, given the low default rate
- DebttoIncome(DTI)ratio is a primary indicator of default that must be considered and verified in granting loans
  - HigherDTIincreasesthechances of default
- If an applicant has made multiple inquiries in the last 6 months, loan is likely to have some defect due to which it was probably rejected
  - Such loans should go through an extra layer of due diligence
- Loans with large installment sizes (typically higher than 200 USD) should be regularly monitored, and if required restructuring plans should be considered to reduce incidences of default due to budget related burden on borrowers
- Borrowers with high interest rates and over 10 years of employment are at risk.
- Loans for purposes such as debt consolidation, credit card payments, or small businesses with amounts greater than 12,000 are also notable.
- Lower ratesofinterestshouldbeofferedfor3yrloansinviewoflowerdefaultsinthisbucket

References

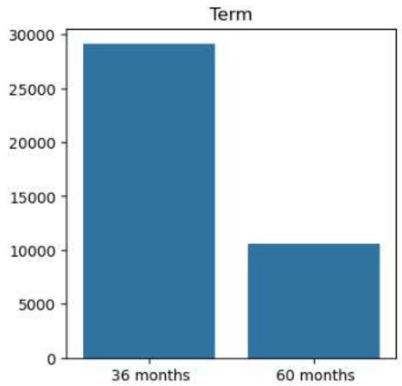
- Python Documentation
  Official Python documentation:
  Link: https://docs.python.org/3/
- Pandas Documentation
  For data manipulation and analysis:
  Link: https://pandas.pydata.org/pandas-docs/stable/
- NumPy Documentation
  For numerical operations and arrays:
  Link: https://numpy.org/doc/stable/
- Seaborn Documentation
  For statistical data visualization:
  Link: https://seaborn.pydata.org/
- Matplotlib Documentation
  For creating plots and visualizations:
  Link: https://matplotlib.org/stable/contents.html
- Kaggle
  For datasets and example notebooks:
  Link: https://www.kaggle.com/
- Stack Overflow
  For troubleshooting and solutions to common coding issues:
  https://stackoverflow.com/
  Link:
- Online Tutorials and Guides
  Real Python: https://realpython.com/
  Towards Data Science: https://towardsdatascience.com/

Appendix

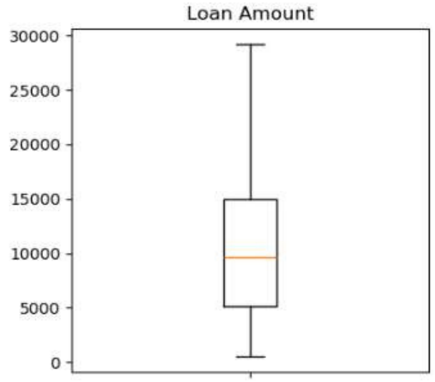
Additional Insights –Univariate Analysis



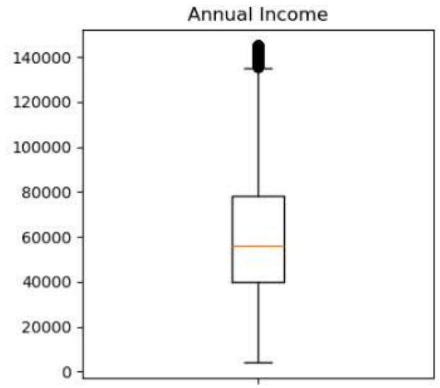
Company tends to grant more loans to high quality borrowers



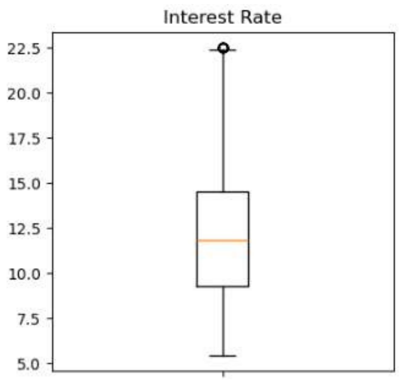
Shorter duration loans are more popular than longer duration



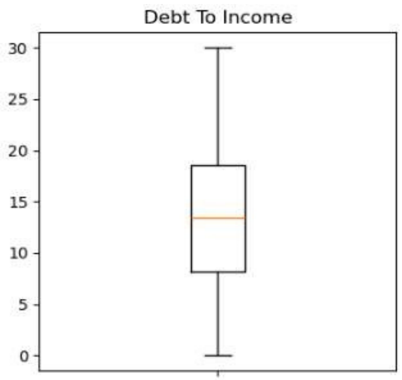
Majority of borrowers tend to choose loan amounts ranging from 5kto15k



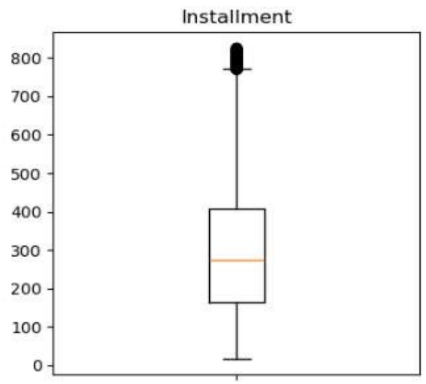
Annual income of borrowers is majorly concentrated between $40k to $80k range



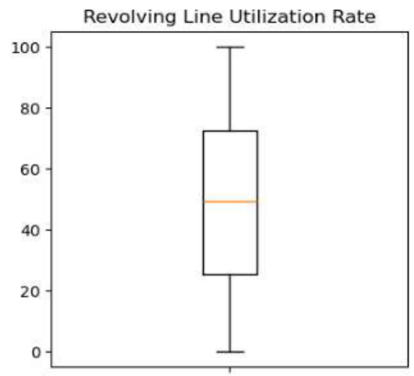
Nearly all the borrowers receive interest rate between 9 to 14%



Average Debt to Income rat io of borrowers is ~14%



Most of the borrowers have installments ranging from $190 to $400

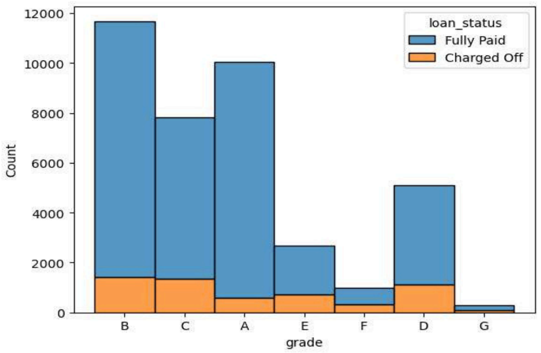


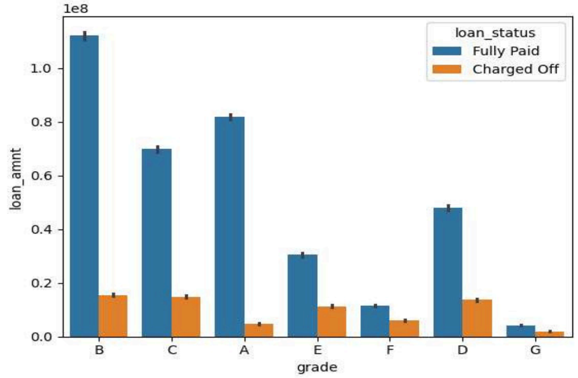
Borrowers on average utilize 50% of their credit line

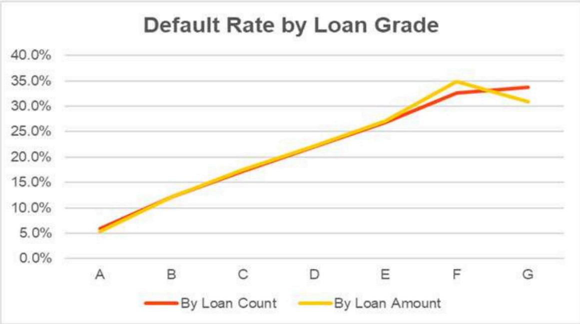
Additional Insights –Understanding Default Behaviour

Loan Grade

- It can be safely concluded that lower grade loans are more likely to make defaults

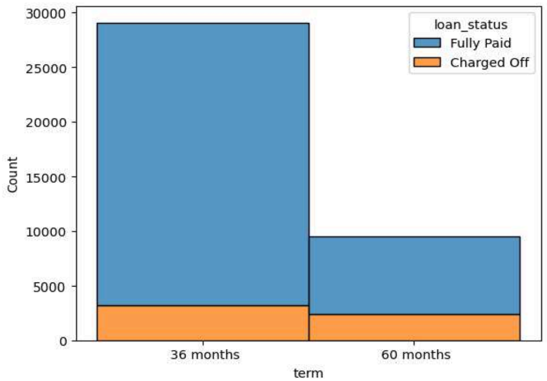


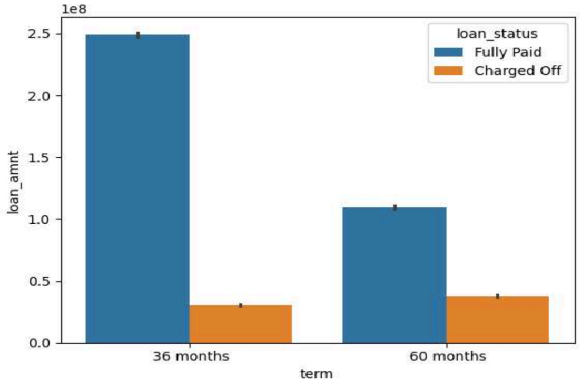


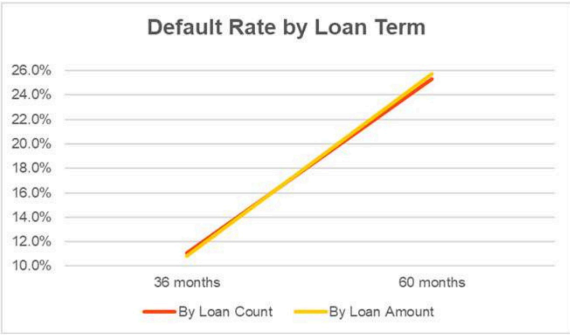


Loan Term

- Default rates are noticeably higher in 5 yr term bucket, but it is inconclusive without understanding the composition of termbuckets







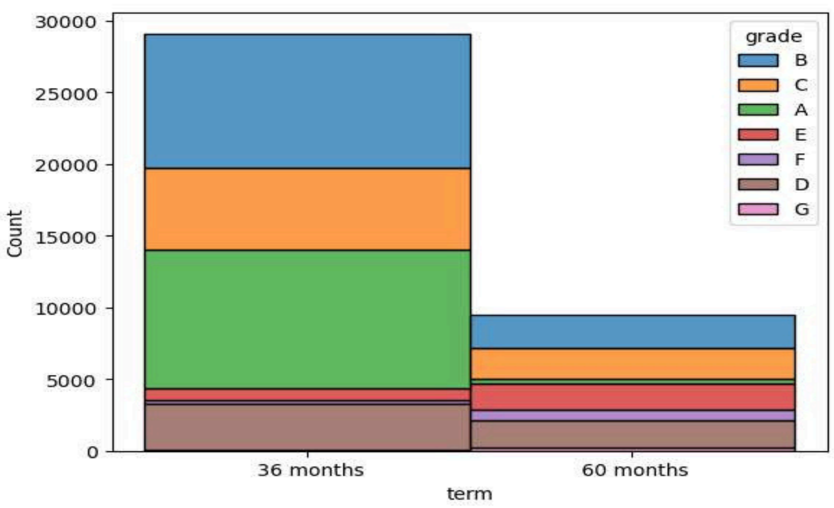
Additional Insights –Understanding Default Behaviour

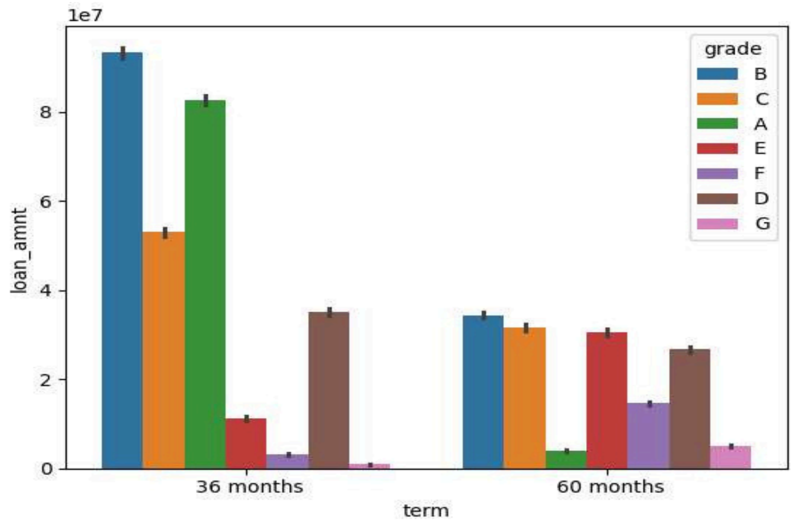
Relationship between Loan Term and Grade

- Higher default rate for 5 yr term loans is primarily driven by higher composition of lower grade loans in this bucket

- Low grade borrowers tend to apply for long duration loans
- 5 yr term bucket comprises of ~49% loans (by count) graded below 'C', as comparedtoonly ~15%in 3yrbucket.
- It clearly shows that the higher default rate in 5 yr term bucket is driven by higher composition of lower grade loans in the bucket.

		Default Rate	
Term	Gra	By Loan Count	By Loan Amount
36 months	de	33. 2%	29. 6%
	A B C	32. 1%	33. 4%
	D E F G	19. 8%	19. 0%
	A B C	11. 1%	12. 6%
	D E F G	3.0%	4.0%
		0.7%	1.1%
		0.2%	0.3%
60 months		4.2%	2.7%
		24. 7%	23. 4%
		22. 0%	21. 5%
		19. 6%	18. 1%
		18. 9%	20. 8%
		8.1%	9.9%
		2.6%	3.5%

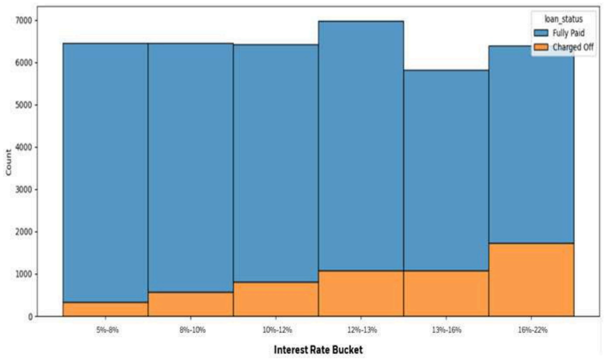


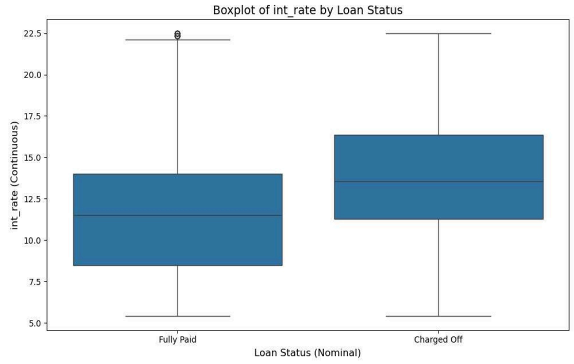


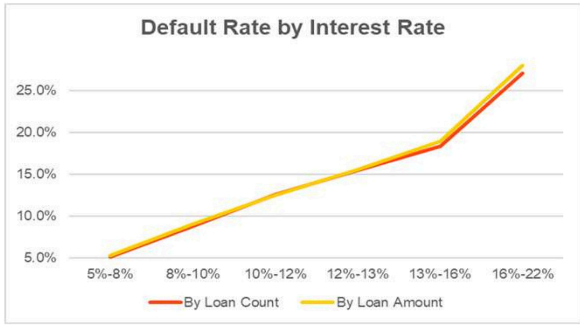
Additional Insights –Understanding Default Behaviour

Interest Rate*

- Higher default rate is observed in high interest rateloans

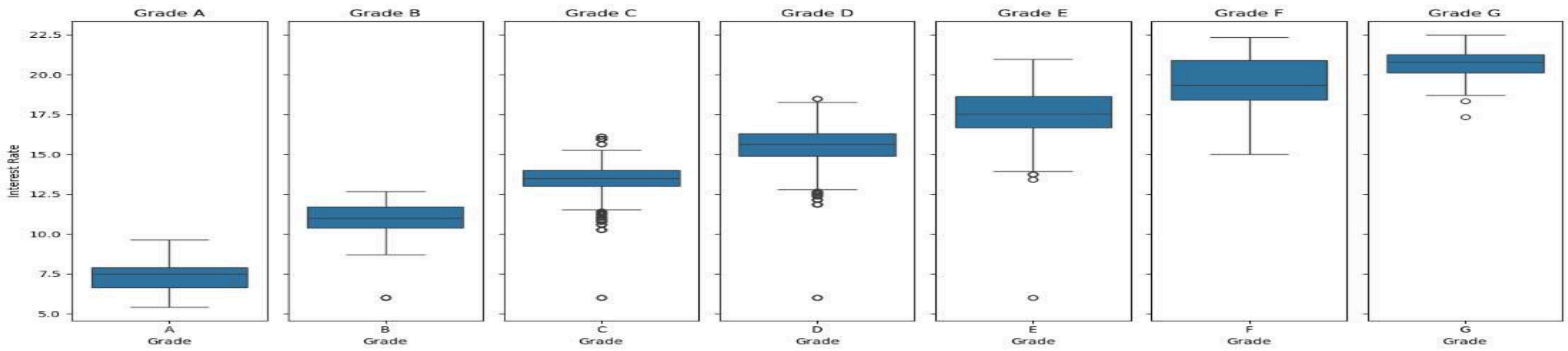






Relationship between Grade and Interest Rate*

- Lower grade loans are usually offered higher interest rates to compensate for higher credit risk

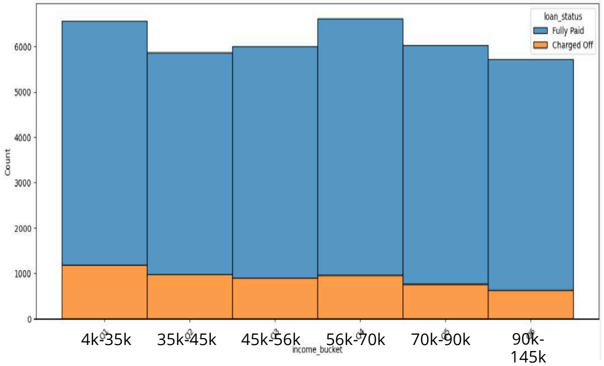


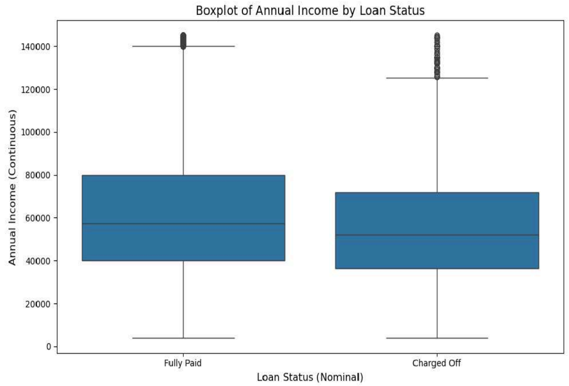
*Analysis has been done after removing the outliers from interest rates. Buckets have been created based on 6 equal quantilestoensure equal representation in each bucket.

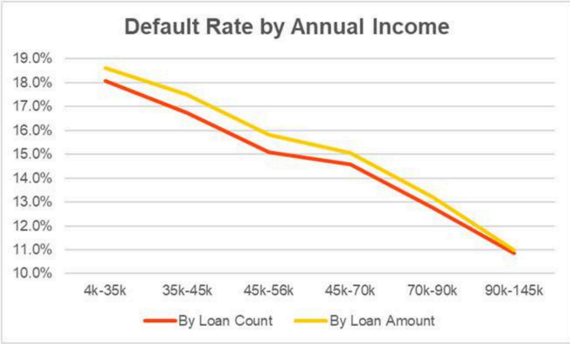
Additional Insights –Understanding Default Behaviour

Annual Income*

- High income borrowers certainly have lower defaults

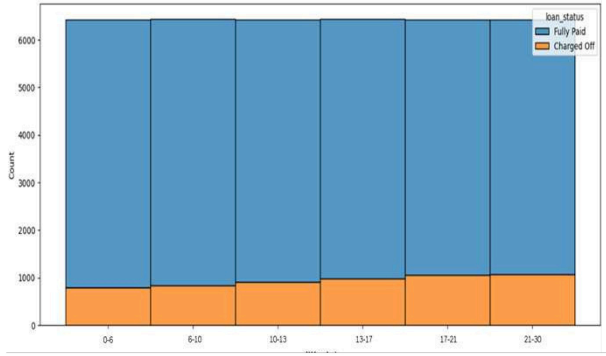


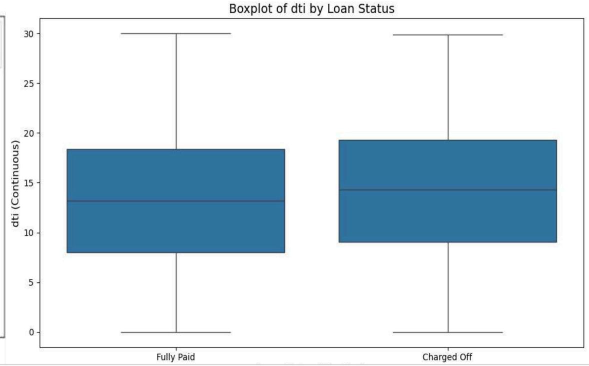


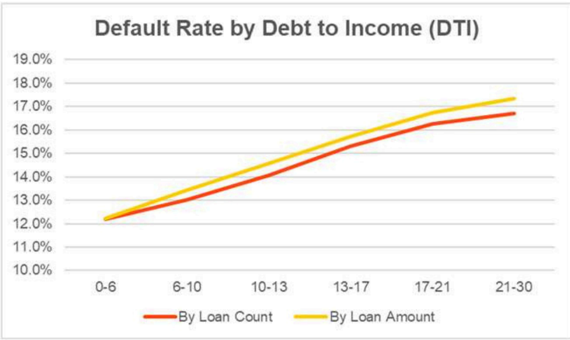


Debt to Income Ratio* (DTI)

- Borrowers with higher DTI tend to have higher default rates







Interdependence of Annual Income* and DTI*

- The correlation between annual income and dti is only -0.075, i.e.very low to establish a relationship between these two variables

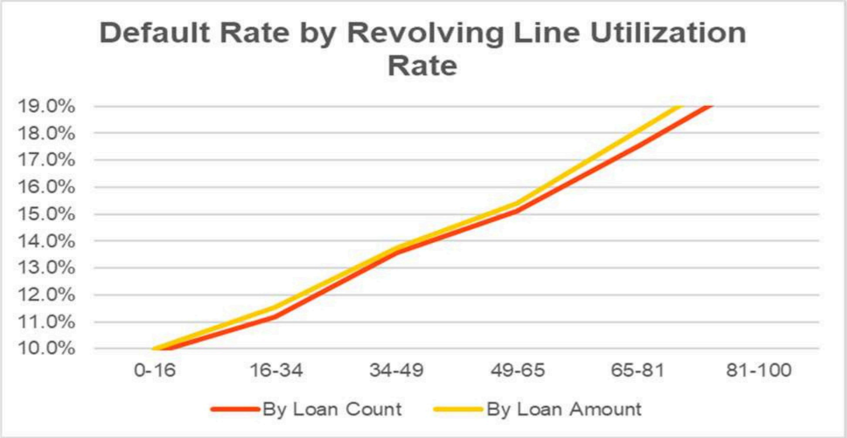
*Analysis has been done after removing the outliers from annual income and dti. Buckets have been created based on 6 equal quantiles to ensure equal representation in each bucket.

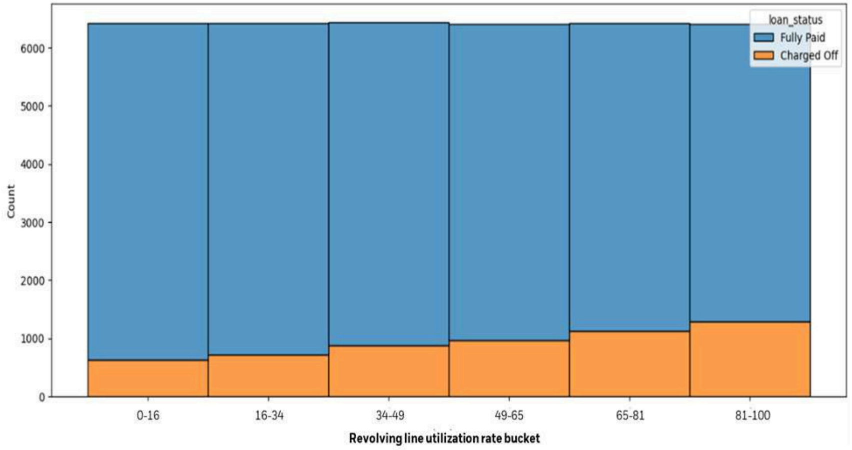
Additional Insights –Understanding Default Behaviour

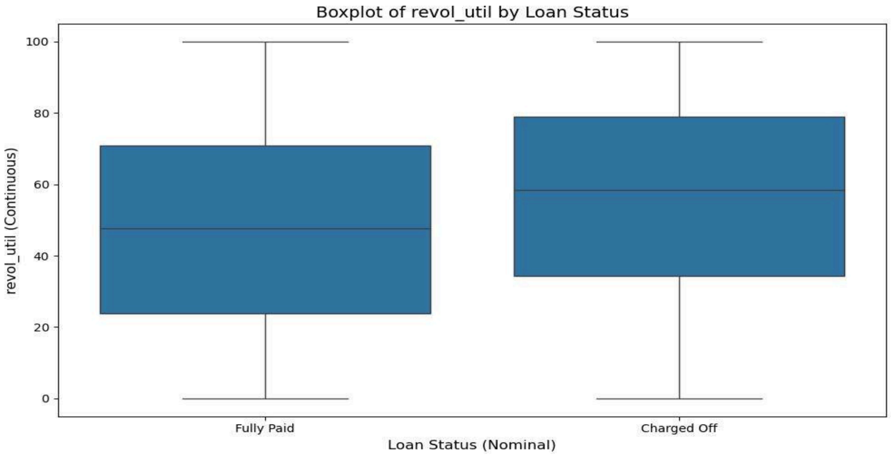
Revolving Line Utilization Rate*

- Default rates are higher for borrowers with higher revolving line utilization rate

- Borrowers may have revolving lines available for their disposal, and they tend to utilize these lines to pay off their debts in times of crisis or before making any default High utilizations rate is a leading indicator of likelihood of default Default
- rates are consistently higher for borrowers with higher revolving line utilization
- rate.





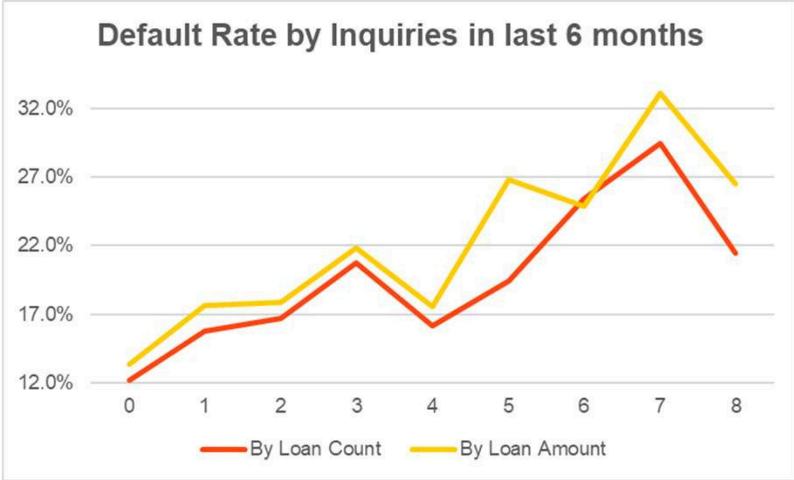


*Analysis has been done after removing the outliers from revol_util. Buckets have been created based on 6 equal quantiles to ensure equal representation in each bucket.

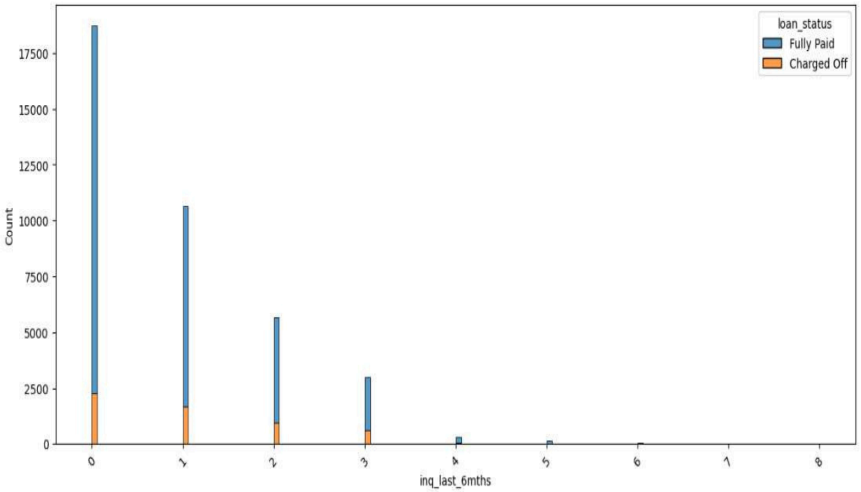
Additional Insights –Understanding Default Behaviour

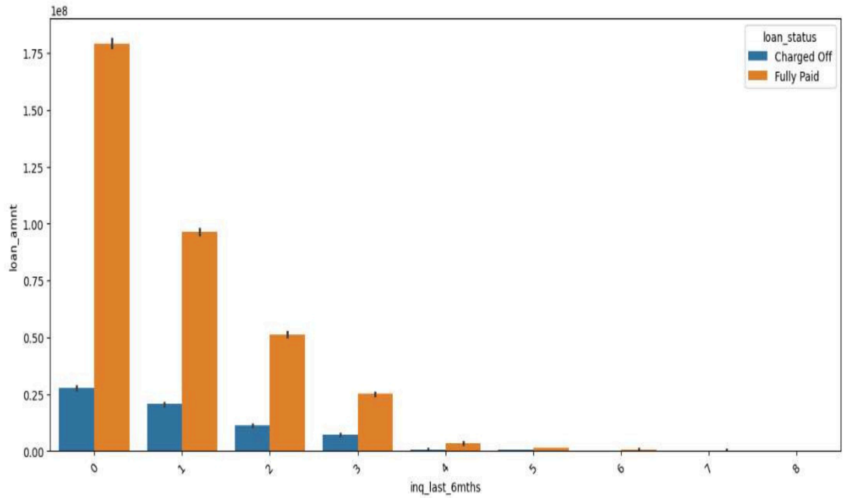
Inquiry in last 6 months

- As a general trend, default rates are higher for borrowers with higher inquiries in last 6 months



- Usuallyborrowers with lower credit quality make multiple inquiries to get favourable terms on the loan
- This is evidentfromthefindingthatlowgradeborrowershavebeenoffered higherinterestrates
- Multiple inquiries also suggest that loans for such borrowers must have been rejected due to some defect by multiple lenders, which also indicate higher chances of default





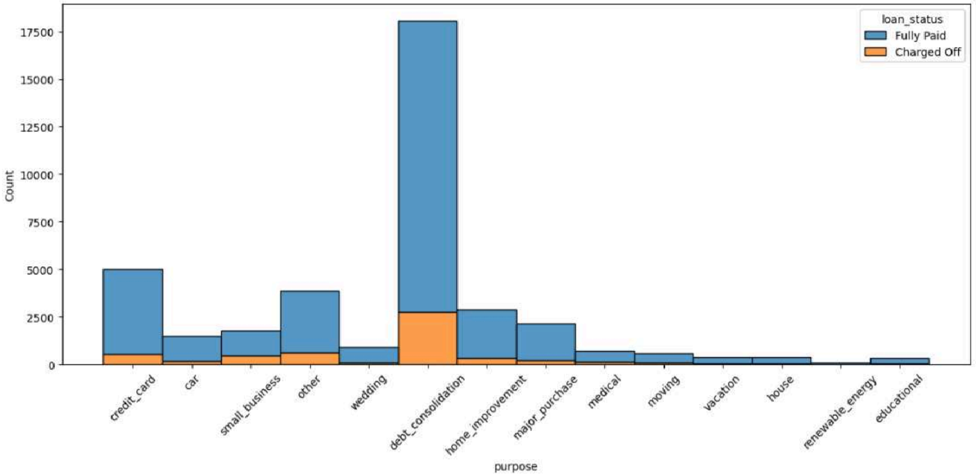
Additional Insights –Understanding Default Behaviour

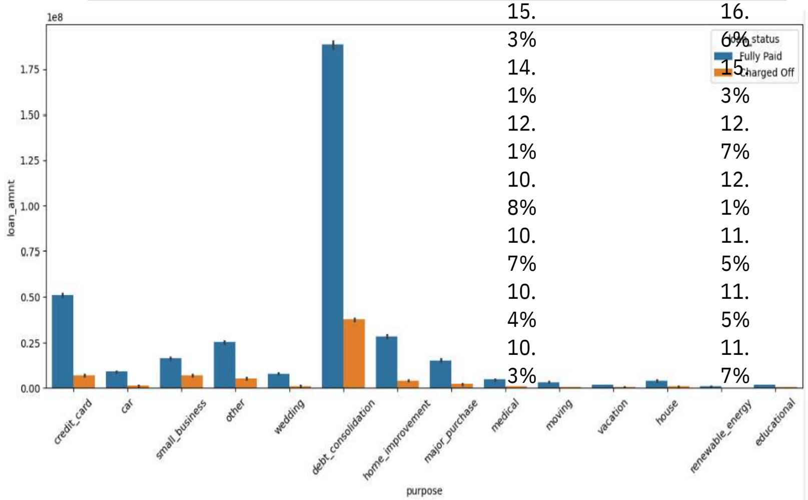
Purpose of Loan

- Small business loans have significantly high default rates
- Disproportionally large amount of loans are made for purpose of debt consolidation

- Substantially large amount of loans are made for the purpose of debt consolidation
  - Suchloans havemoderatelevelsof defaultrates
  - Suchborrowerstendtorepaytheloan fromfurtherconsolidation
- Significantly high default rates (~27% by loan count) are observed in loans for the purpose of small business
- Consumer loan such as purchases, car, credit card, wedding, etc have relatively lower defaultrates
  - Such loans are typically optional in nature and tend to be taken by borrowers withexpectedcashflows infuturetorepaytheloan

	Default Rate	
Purpose	By Loan Count	By Loan Amount
s mal l_bus ines s	27.	29.
renewable_energy	1%	6%
educ at ional ot her	18.	18.
hous e moving medi	6%	4%
c al debt _c ons ol i	17.	19.
dat i on vacation	2%	9%
home_improvement	16.	17.
c redi t _c ard car	4%	1%
wedding maj or_purc	16.	15.
has e	1%	9%
	16.	14.
	0%	4%
	15.	16.
	6%	4%
	15.	16.
	3%	6%
	14.	15
	1%	3%
	12.	12.
	1%	7%
	10.	12.
	8%	1%
	10.	11.
	7%	5%
	10.	11.
	4%	5%
	10.	11.
	3%	7%





Additional Insights –Understanding Default Behaviour

Installment

- Beyond a certain threshold of ~200 USD, chances of default tend to increase with the installment size

- Higher installment sizes may result in increased burden on borrowers
- Suchloansarelikelytodefaultincase of loss of employment or any adverse condition impacting the income of the borrower
- Default rates tend to increase when installment increases beyond a threshold level