



Lending Club Case Study:

Assessment By: Divakar & Vivek

OBJECTIVE

Identify the **driving factors (or driver variables)** behind loan default

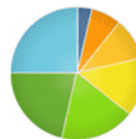
What is Lending Club?

Lending Club is a marketplace for personal loans that matches borrowers who are seeking a loan with investors looking to lend money and make a return.

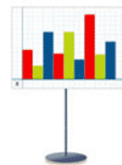
How Lending Club Works



Borrowers apply for loans.
Investors open an account.



Borrowers get funded.
Investors build a portfolio.

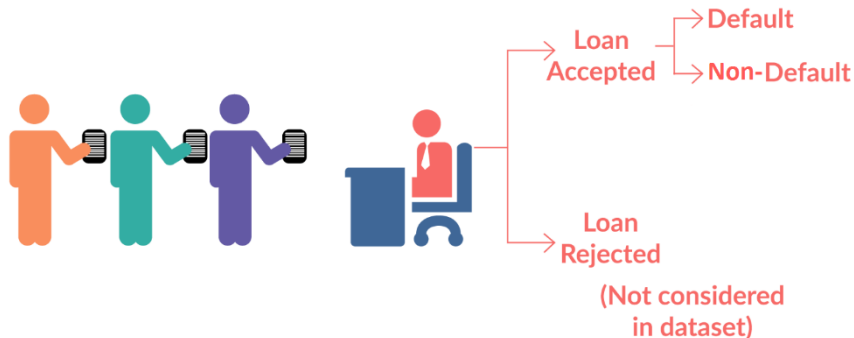


Borrowers repay automatically.
Investors earn & reinvest.

When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

- 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, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company
- If one can identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.
- Identification of such applicant's using EDA is the aim of this case study

LOAN DATASET



Fully paid: Applicant has fully paid the loan (the principal and the interest rate)

Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.

Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan

Lending Club: Dataset Critical Parameters

Parameters	Description	Parameters	Description
loan_amnt :	Amount applied by potential borrowers	annual_inc:	yearly income of the loan applicant
funded_amnt:	Amount recommended/approved by Lending Club	int_rate:	rate of interest of the availed loan
funded_amnt_inv :	Amount funded by investors	emp_length:	work experience of the loan applicant
dti:	Debt to Income ratio	verification_status:	income verification status of the loan applicant
grade:	loan classification based on interest rate	home_ownership:	loan applicant home ownership status
purpose:	purpose of loan such as car, house, vacation..etc	issue_d:	date of issue of loan
installment:	monthly EMI amount	term:	loan repayment EMI number of months
loan_status:	status of the loan Fully Paid/Charged Off/Current		

- Import Raw Loan dataset ("loan.csv") to → "loands" variable.
- Remove null value rows (referring "emp_length" column) from the dataset to create → "loands2".
- Drop rows based on outliers in "int_rate", "installment" & "loan_amnt" in loands2 dataset
- Drop less important columns from "loands2".
- Extract → "month" & "year" column from "issue_d".
- Convert → "int_rate" to float after removing % symbol
- Generate → "exp_exper" column ("Beginner" / "Junior" / "Senior" / "Vsenior") from the "emp_length"
- Calculate → "monthly_inc" from the "annual_inc"
- Calculate → "netcash_monthly" from the "monthly_inc", "dti" & "installment".
- Drop → "Current" category rows based on loan status("Charged Off", "Fully Paid" and "Current") to form → "loanNcur" for better correlations

Calculated Parameter	Description
year:	year in which loan was applied. Extracted from “issue_d”
month:	month in which loan was applied. Extracted from “issue_d”
monthly_inc:	monthly income of the loan applicant. Calculated by “annual_inc”/12
netcash_monthly (NCM):	Net cash available with loan applicant after repaying debt monthly in percentage. netcash_monthly= $(\text{monthly_inc} - \text{installment} - (\text{dti} * \text{monthly_inc}) / 100) * 100 / \text{monthly_inc}$
emp_exper :	loan applicant work experience category. Calculated from “emp_length”. emp_length (0 – 2yrs) :Beginner emp_length (3 – 5yrs) :Junior emp_length (6 – 9yrs) :Senior emp_length (>=10yrs) :VSenior

Lending Club: Total Charged Off Loans

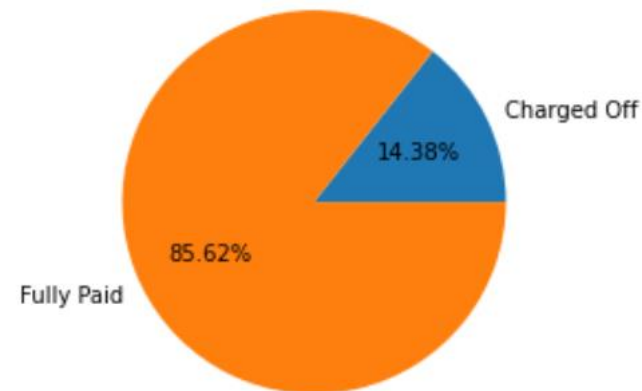
Total Loans from the dataset: 39717

“Fully Paid” Loans: 32950

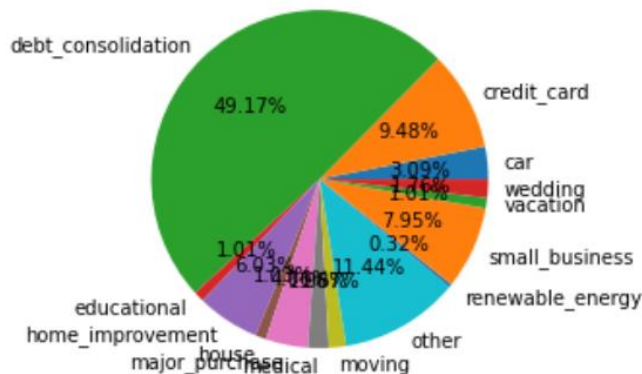
“Charged Off” Loans: 5627

“Current” Loans: 1140

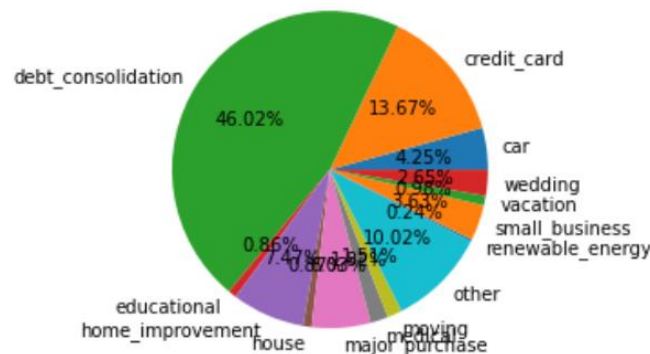
Charge Off Loans



Charge Off -Loan

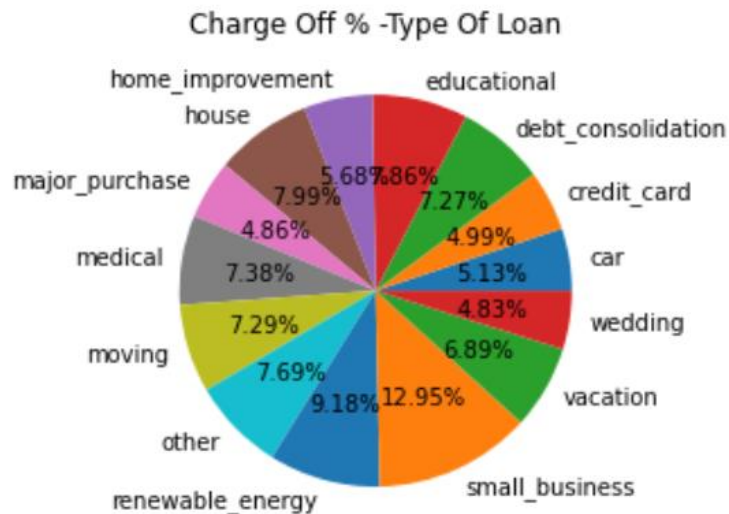


Fully Paid-Loan



Looking at the defaulter's distribution based on their loan's availed

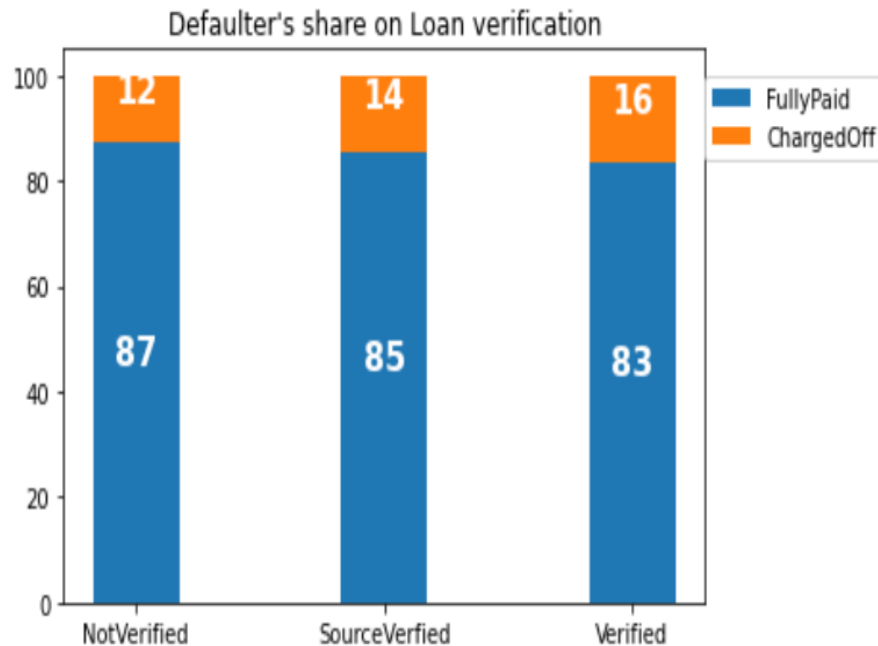
- In terms of %
 - “Small_business” type of loan contributes the highest % of defaulters
 - Followed by “renewable_energy”
 - House, educational, other, medical, moving, debt_consolidation, & vacation type of loans show similar % defaulters
- In terms of volume,
 - “debt_consolidation” loan contributes to nearly 49.2% of defaulters,
 - followed by “home_improvement” and “other” loans



Lending Club: Defaulter %- On Loan verification status upGrad

Loans applicant's income are usually verified, and status is categorized as Verified, Not Verified and Source Verified

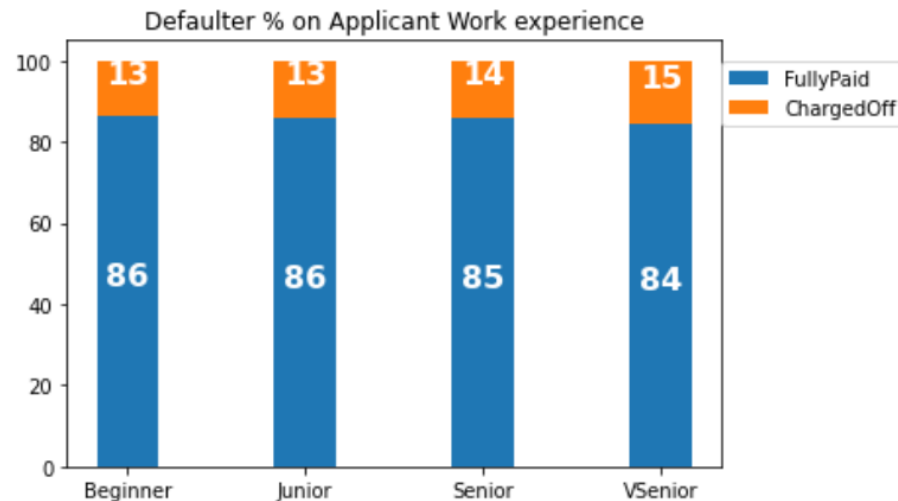
- According to the income verification status, when we plot the defaulter's distribution, we can't see much difference between "Verified" & "Not Verified" loans.
- The trend shows slightly reverse belief. Instead of "NotVerified" , "Verified" loans has slightly high defaulters.



Lending Club: Defaulter %- On Loan Applicant Experience upGrad

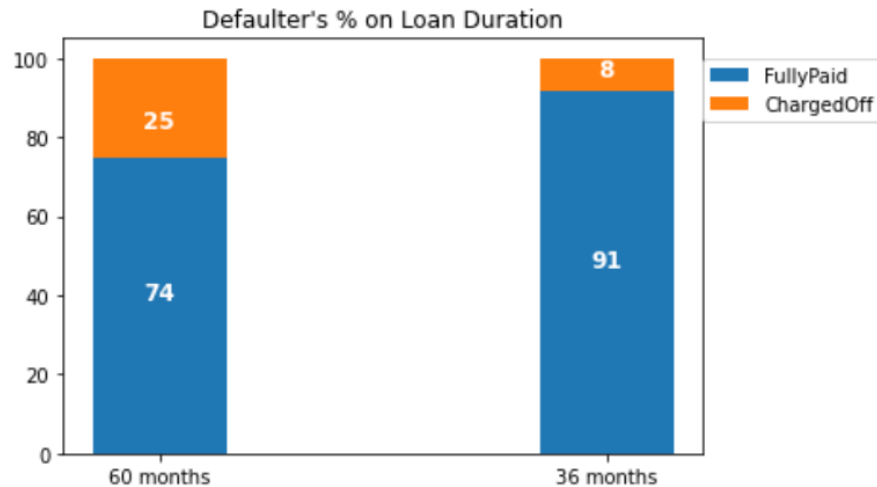
Based on applicant's employment length they are grouped as Beginners Junior, Senior and VSenior. When we plot the defaulter's % distribution based on work experience

- We don't see much difference in the applicant defaulting behavior based on their experience.
- We expected a lower defaulting % with increasing experience, but the results are slightly reverse



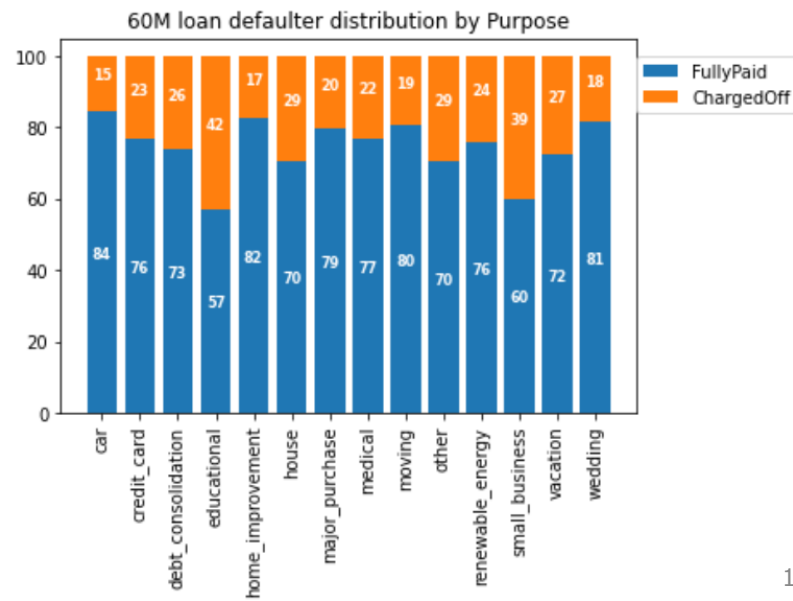
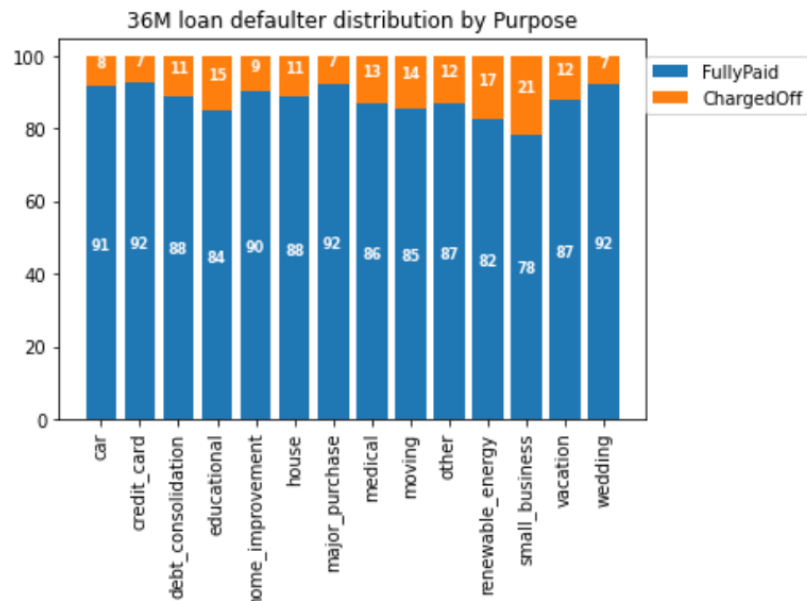
The applicant can choose between 60-month term loan or 36-month term loan. When we plot the defaulter's % distribution based on the loan term

- We see 300% more defaulters in 60 months loan when compared to 36 months loan.



When we perform bivariate analysis on loan defaulter w.r.t loan term and loan purpose, we can understand the following

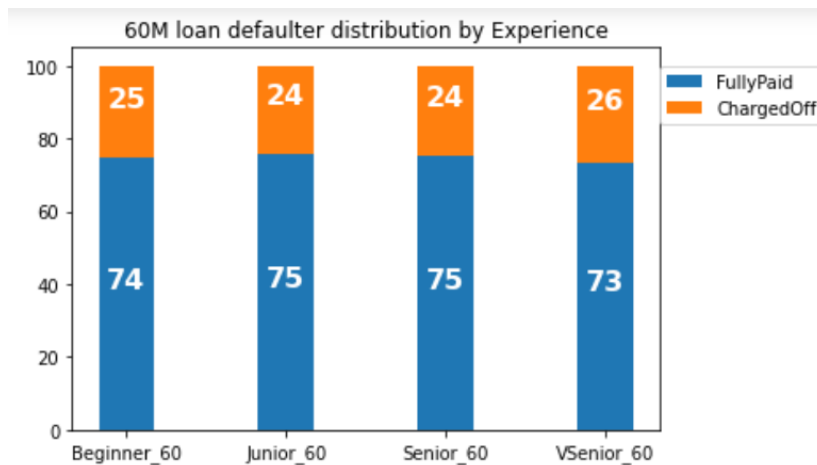
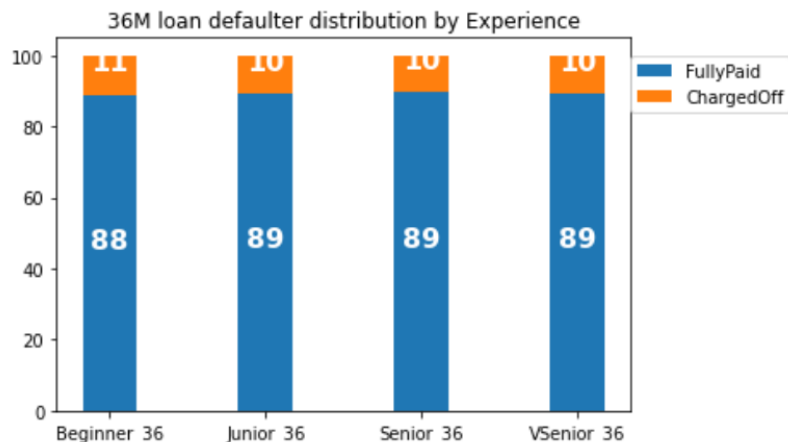
- On an average loan defaulters in 60-month term is twice/thrice that of 36-month loan irrespective of the purpose of the loan
- Educational and Small business type of loan has the highest % of defaulter



Lending Club: Defaulter %- On Loan Duration & Experience

When we perform bivariate analysis on loan defaulter w.r.t loan term and applicant experience, we can understand the following

- On an average loan defaulters in 60-month term is twice/thrice that of 36-month loan irrespective of the loan applicant experience

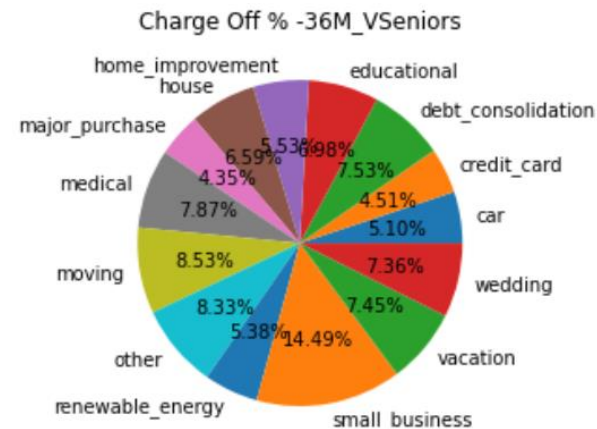
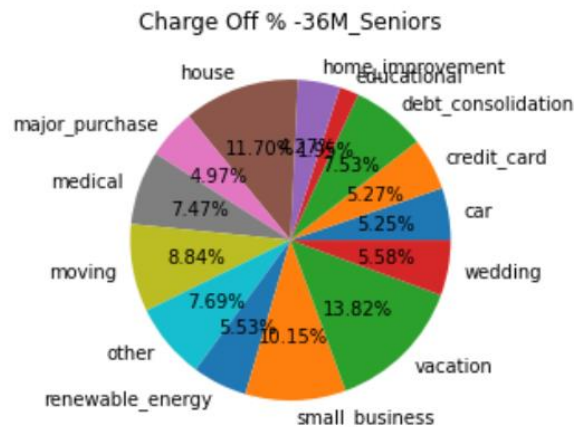
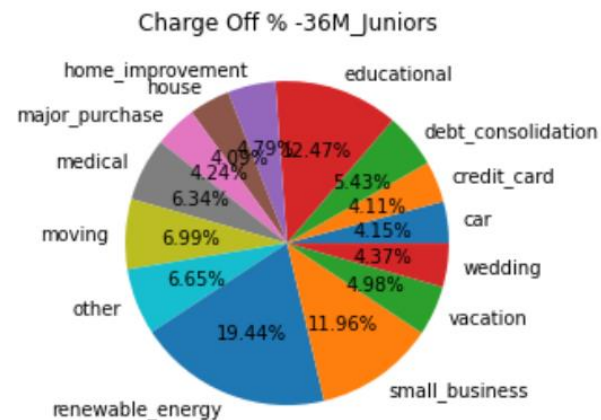
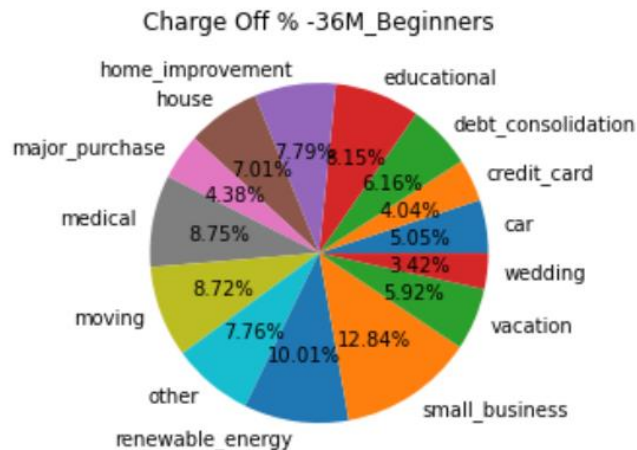


Multivariate Analysis on 36M-Loan Defaulter %

- Loan Term, Experience & Purpose

36 month loan defaulter's

- “Renewable Energy” loan contributes max % of defaulters in Beginners and Juniors
- Defaults in “Small_business” loans are highest in all experience
- Seniors seems to default more in housing loan & and vacation loan.

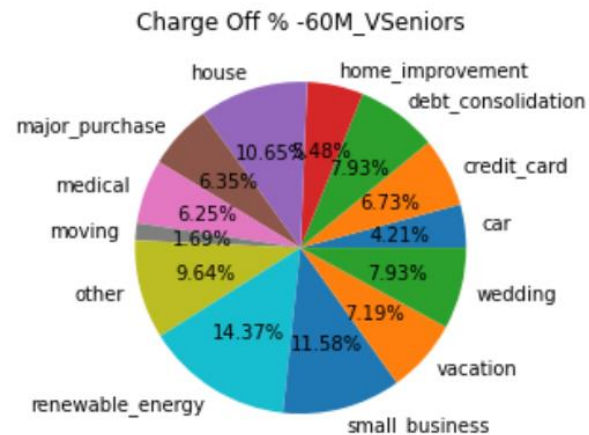
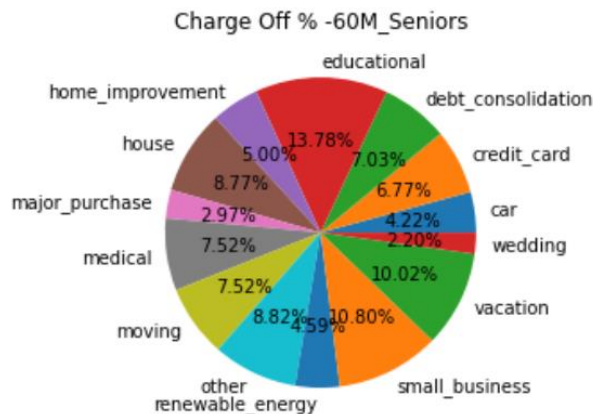
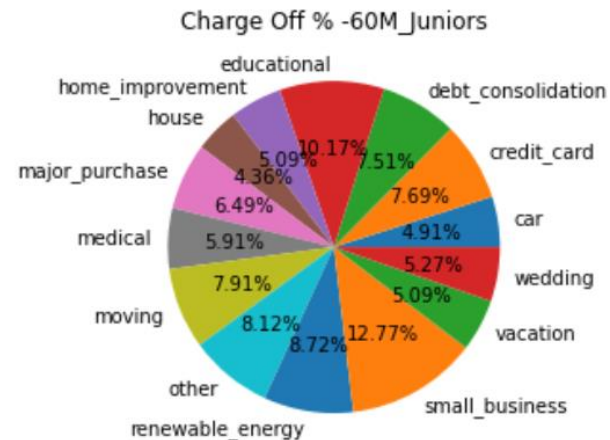
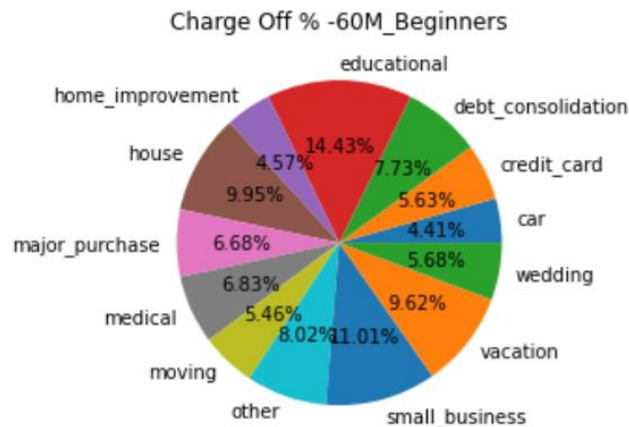


Multivariate Analysis on 60M-Loan Defaulter %

- Loan Term, Experience & Purpose

60- month loan defaulter's

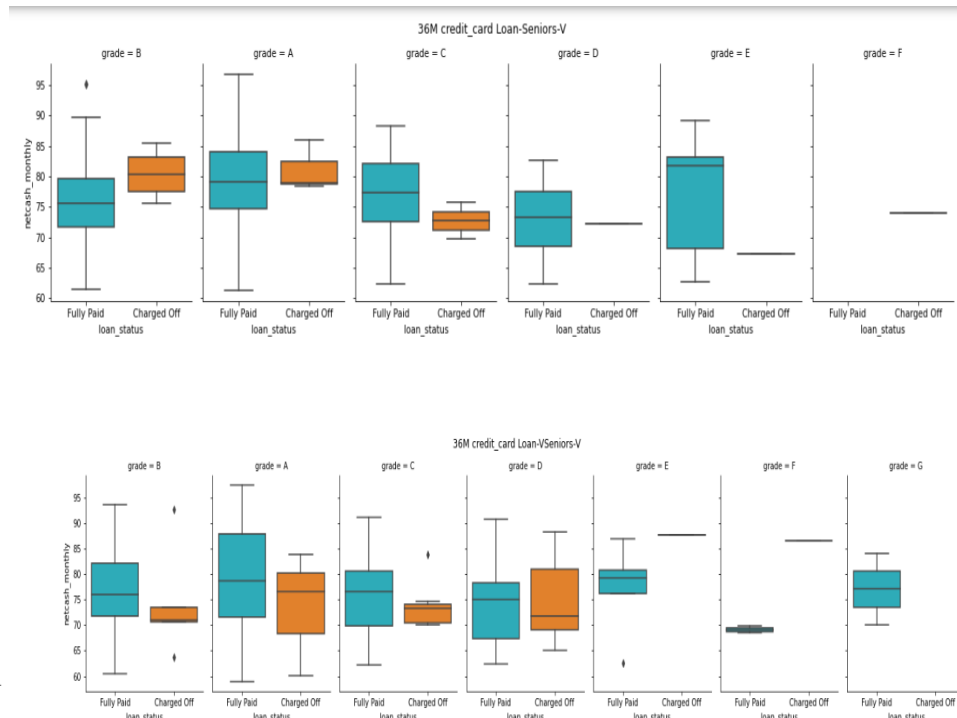
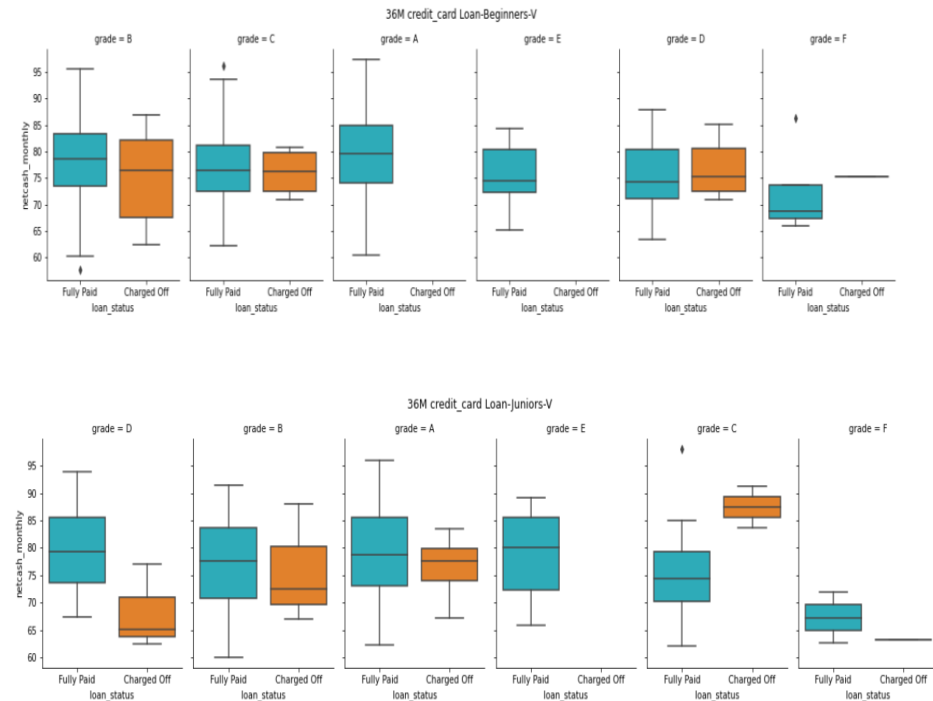
- “Educational” loan contributes max % of defaulters in Beginners, Juniors and seniors
- Defaults in “Small_business” loans are highest in all experience
- Seniors seems to default more in housing loan & and vacation loan.
- Vseniors seems to default more on renewable energy and house



Multivariate Analysis: 36M-Credit Card Loan

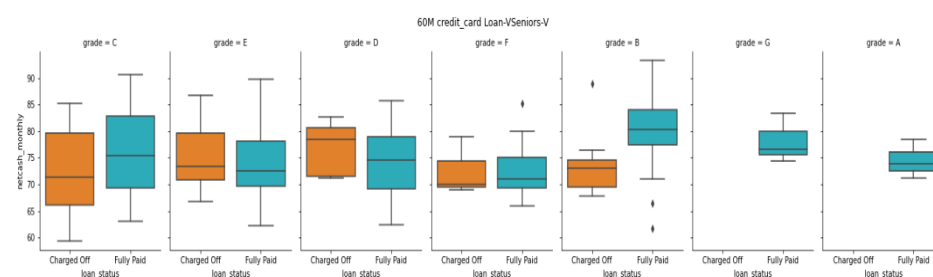
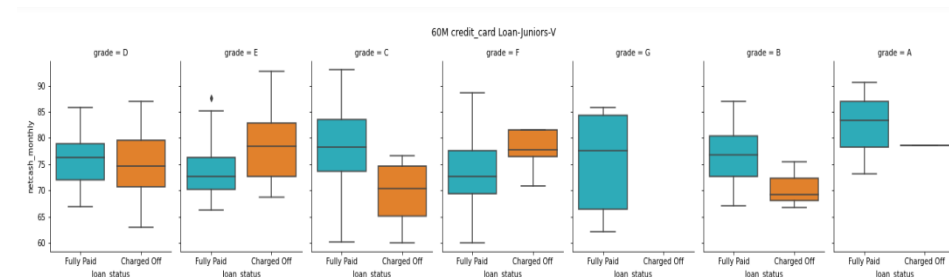
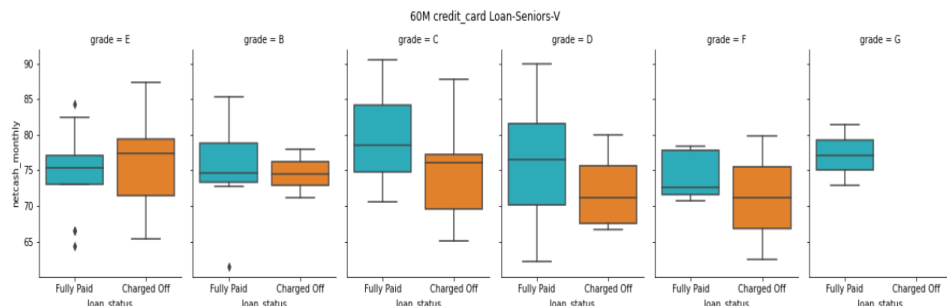
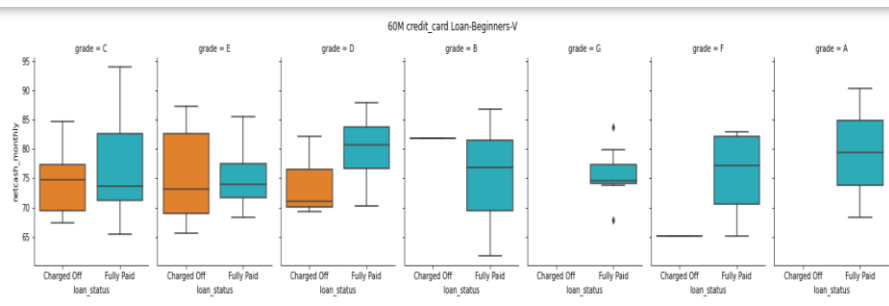
High possibility of default by:

- Applicants availing “A5”, “B” , “C”, “D” grade loan with NCM <75%
- Not-Verified loan applicants



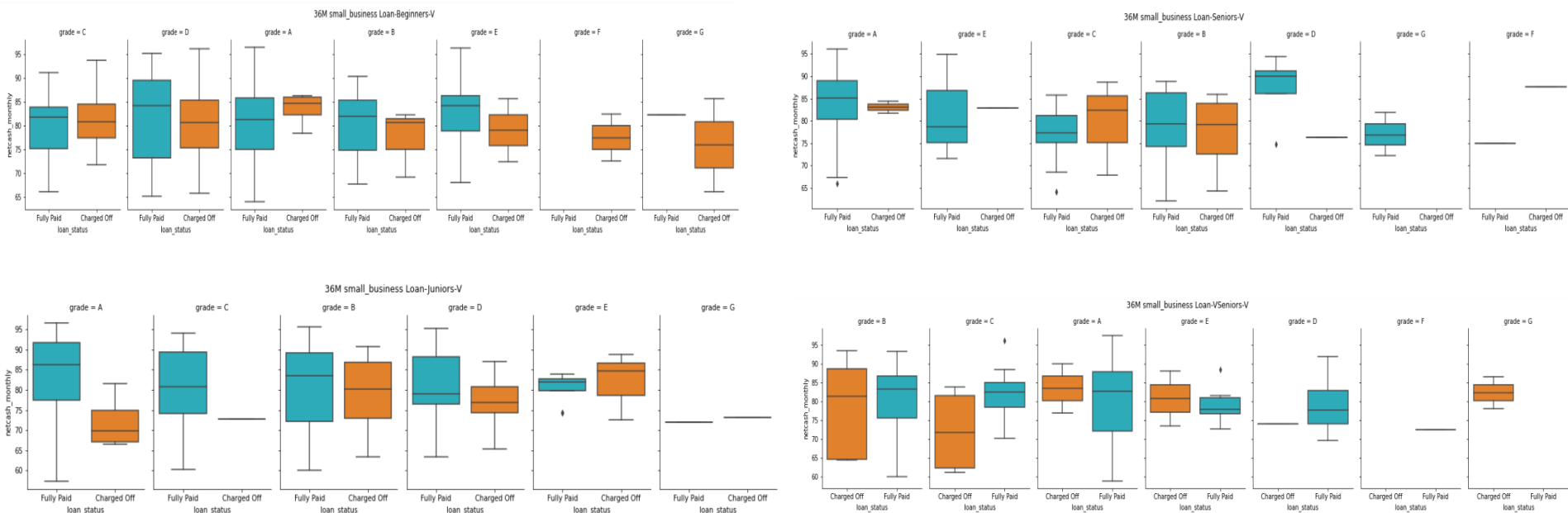
High possibility of default by:

- Applicants availing “B4”, “C”, “D”, “E” grade loan with NCM <78%



High possibility of default by:

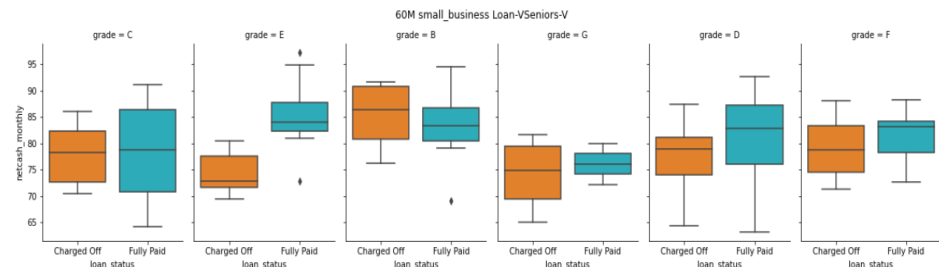
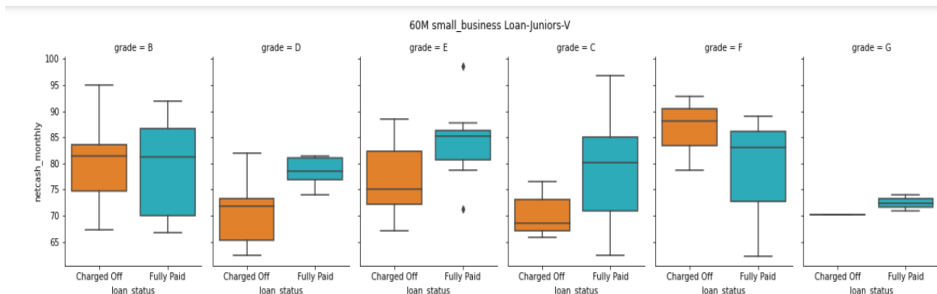
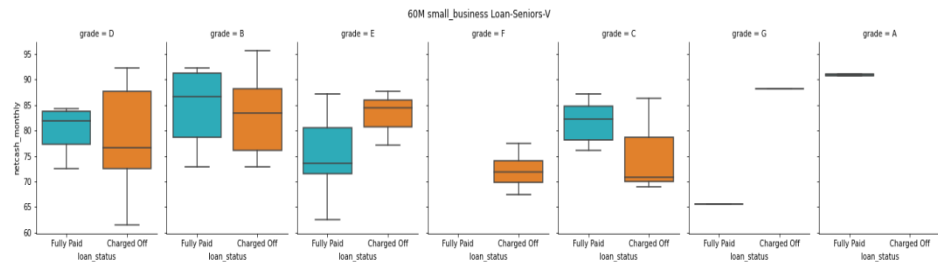
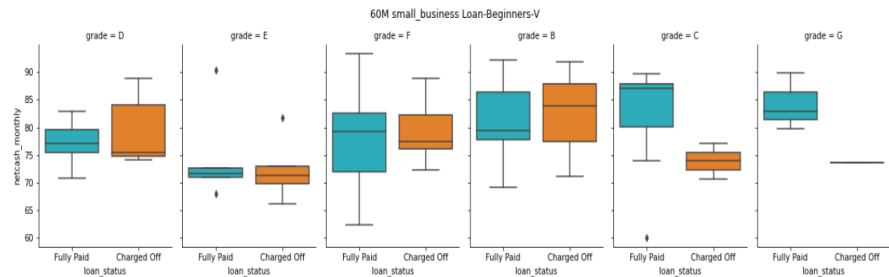
- Applicants availing “C”, “D” grade loan who posses NCM <80%
- Applicants availing “B” grade loan with NCM <78%



Multivariate Analysis: 60M-SmallBusiness Loan

High possibility of default:

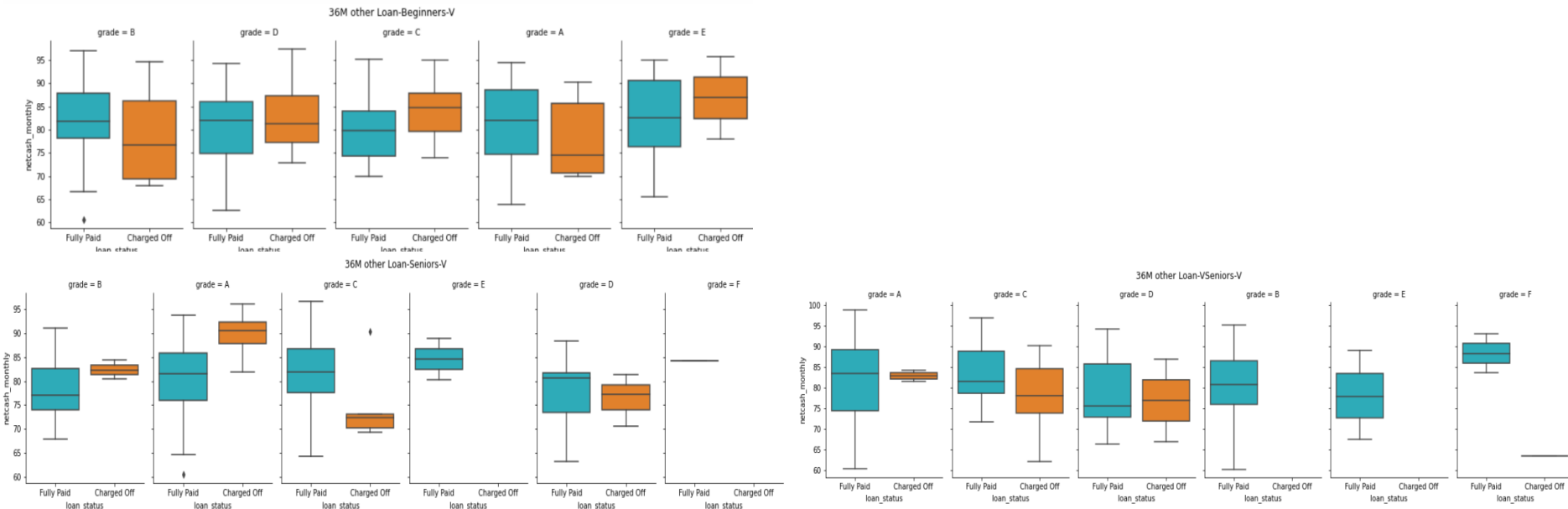
- More % of defaulters compared to same 36M loan
- Applicants availing "B", "F" grade loan who posses NCM <82%
- Applicants availing "C", "G" grade loan who posses NCM <76%
- Applicants availing "D", "E" grade loan who posses NCM <80%



Multivariate Analysis: 36M-Other Loan

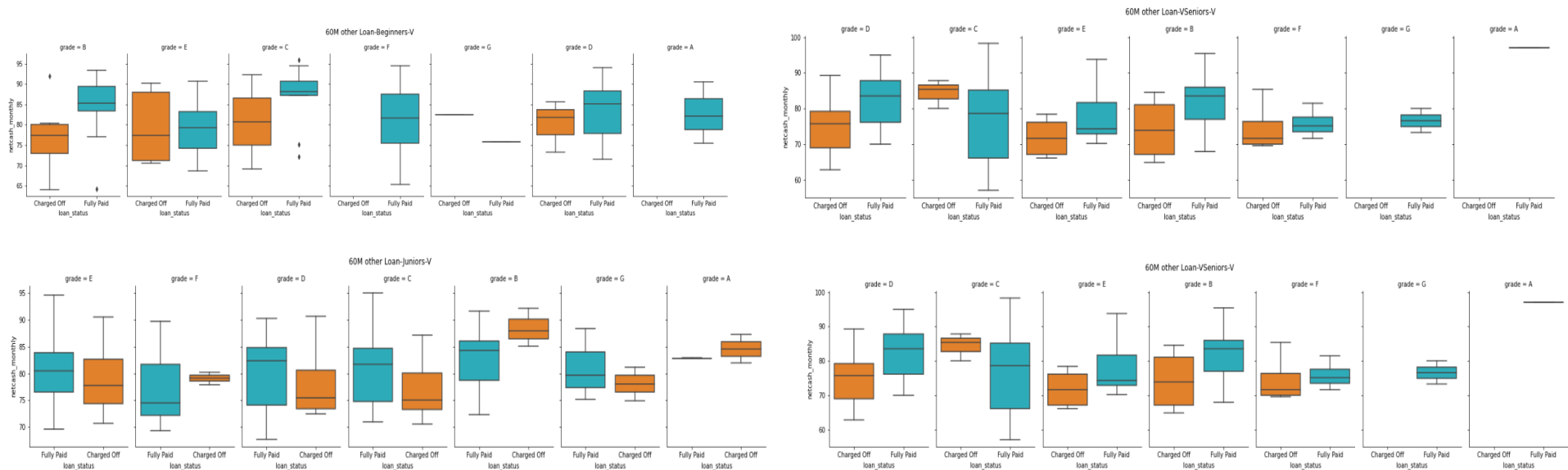
High possibility of default:

- More % of defaulters in Not Verified Category
- Applicant availing “B” grade loan who posses NCM <76%
- Applicant availing “A” grade loan who posses NCM <80%
- Applicants availing “C”, “D” grade loan who posses NCM <78%
- Applicant availing “E” grade loan who posses NCM <85%



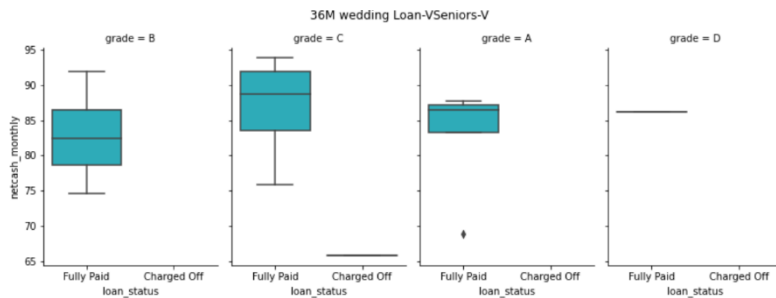
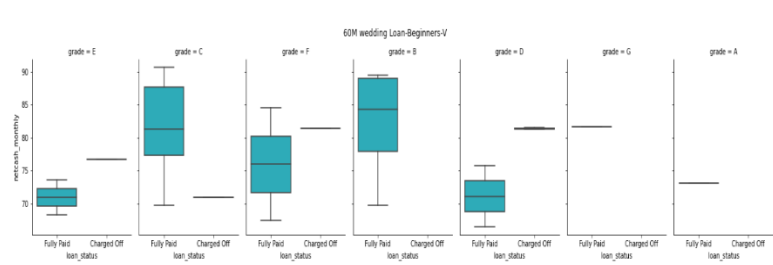
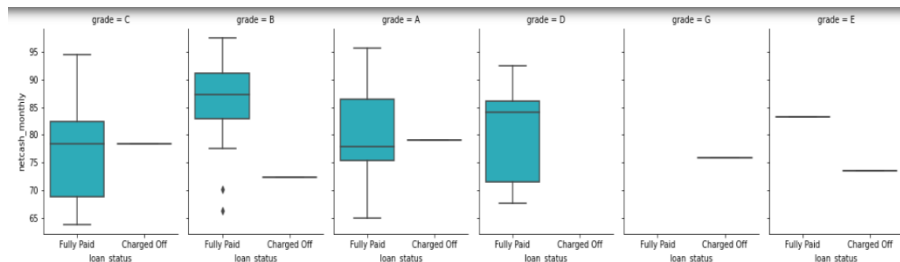
High possibility of default:

- Applicants availing “B” grade loan who posses NCM <76%
- Applicants availing “C” grade loan who posses NCM <84%
- Applicants availing “D”, “E” grade loan who posses NCM <81% ,82%



Multivariate Analysis: 36/60M-Wedding Loan

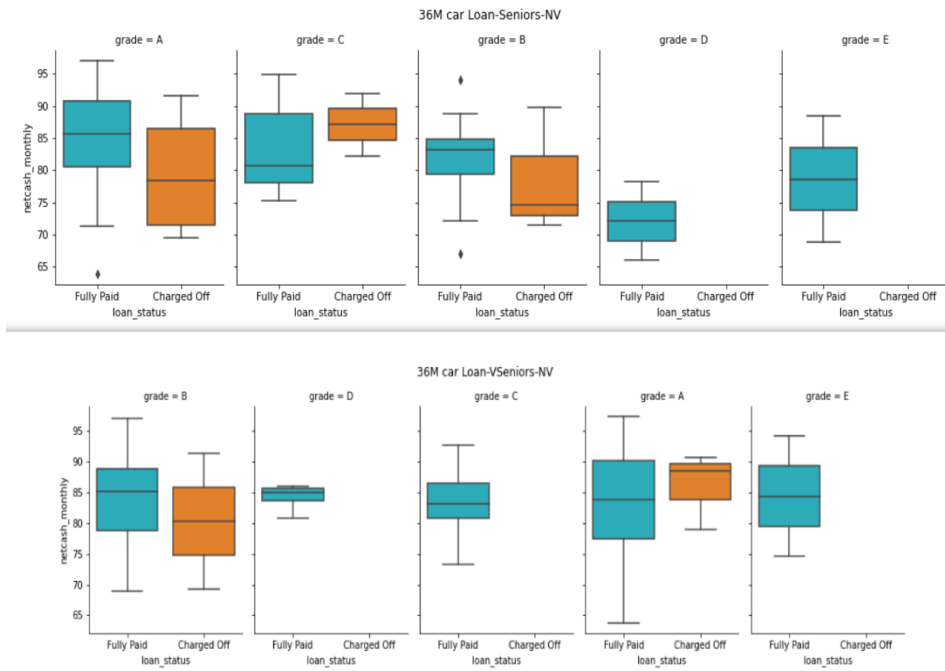
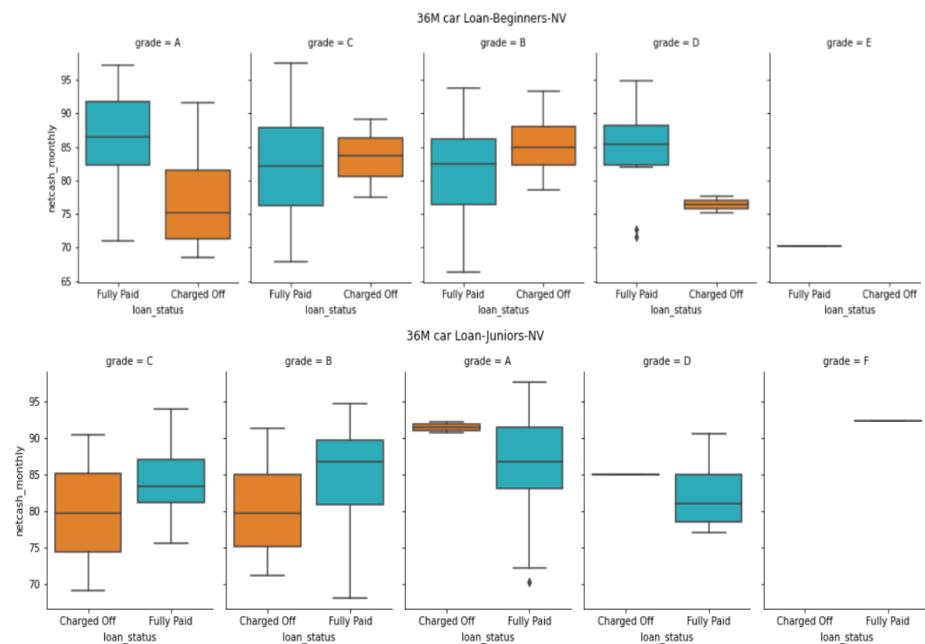
- Overall default % in wedding loan is less compared to other loans
- Applicants availing “B”, ”D” grade loan who possess NCM <88% most likely to default in the Beginners or Junior category



Multivariate Analysis: 36M- Car Loan

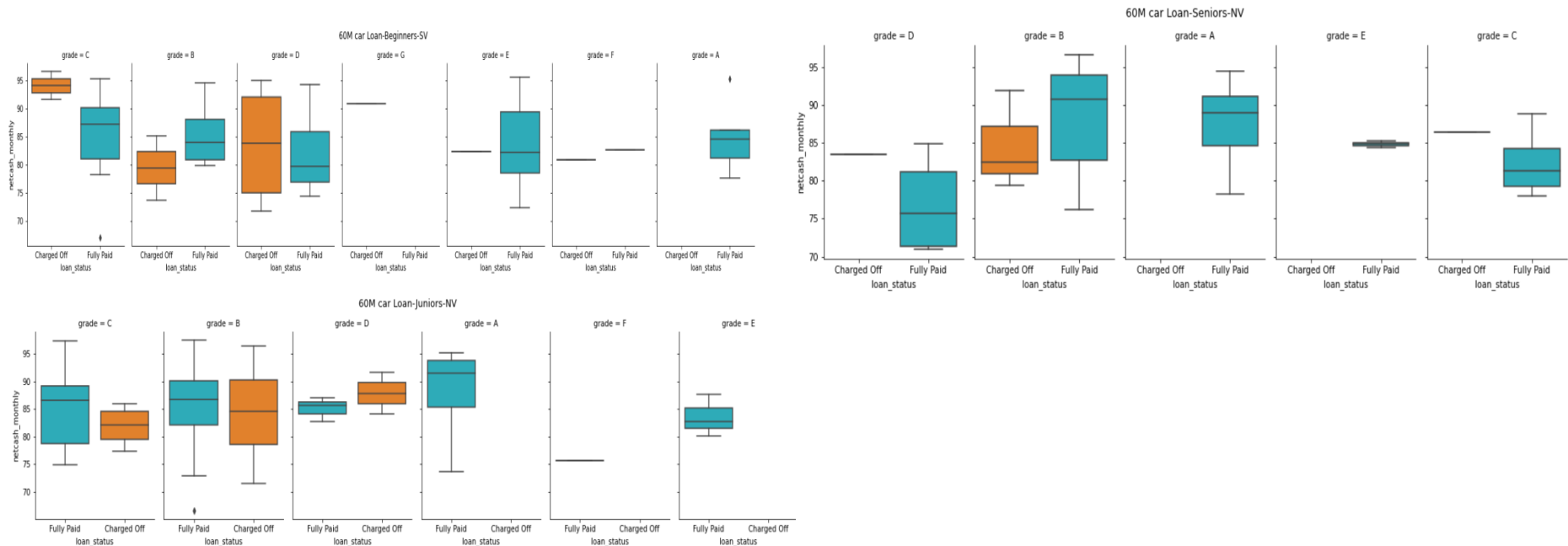
High possibility of default:

- Most of the defaults are in Not-Verified category
- Applicants availing “A”, “B” grade loan who posses NCM <78%
- Applicants availing “C” grade loan who posses NCM <84%



High possibility of default:

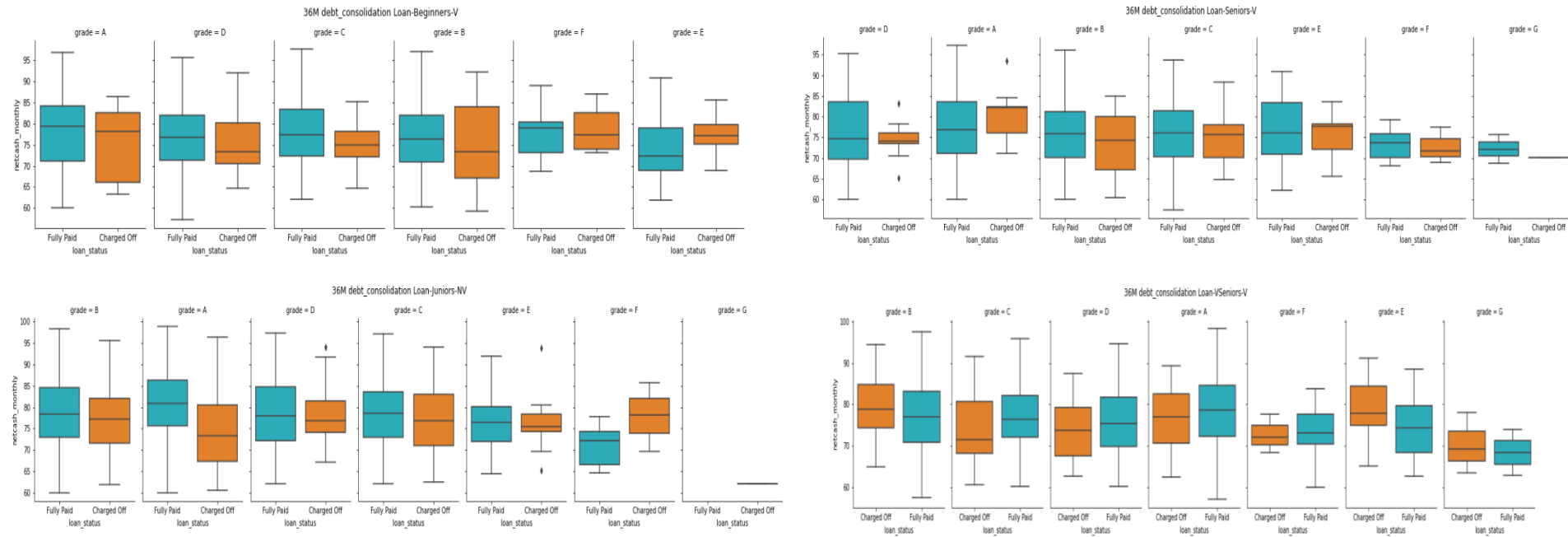
- Most of the defaults are in Not-Verified category
- Applicants availing “B”, “C”, “D” grade loan who posses NCM <82%, 80%, 86%



Multivariate Analysis: 36M- Debit Consolidation Loan

High possibility of default by:

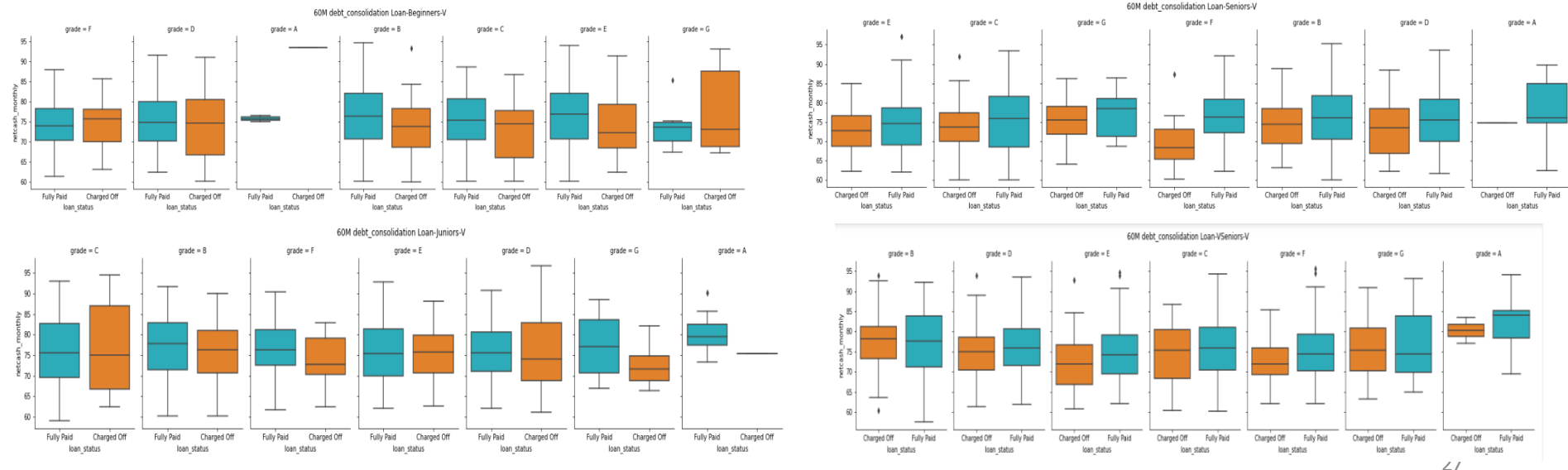
- Most availed loan with highest vol% of defaulter in Verified and Not Verified category
- Applicants availing “A”, “D” grade loan who posses NCM <75%
- Applicants availing “B”, “C” grade loan who posses NCM <73%
- Applicants availing “E” grade loan who posses NCM < 78%



Multivariate Analysis: 60M- Debit Consolidation Loan

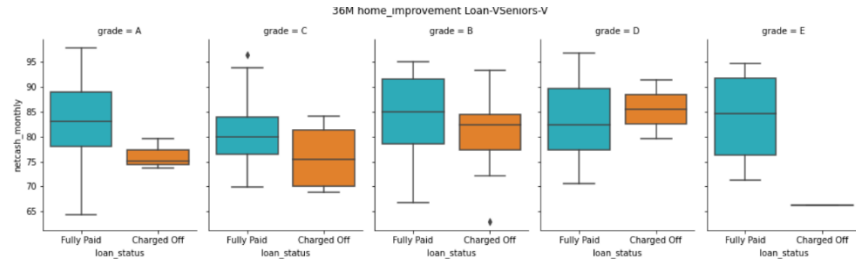
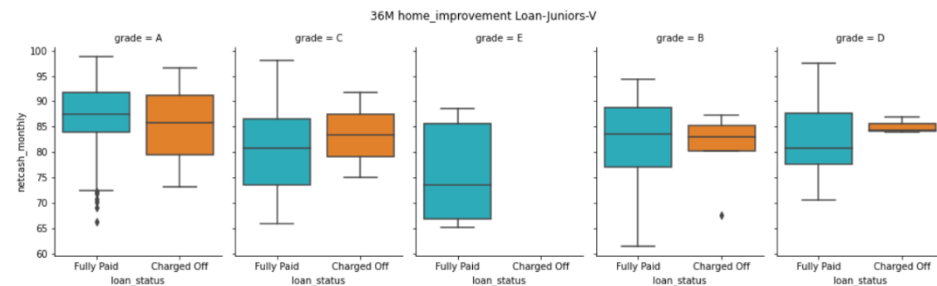
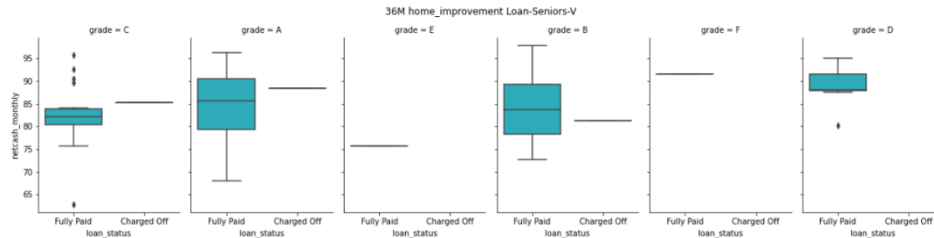
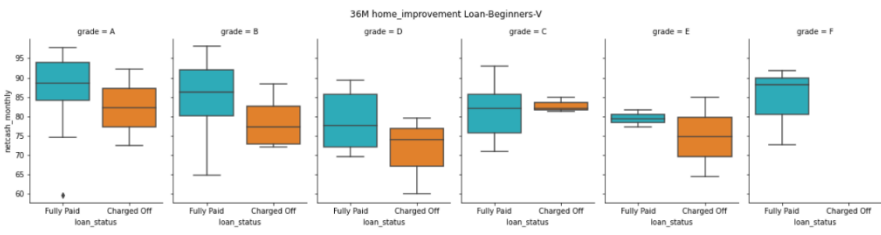
High possibility of default by:

- Most availed loan with high vol% of defaulters in Verified and Not Verified Category
- Applicants availing “B”, “D” grade loan who posses NCM <73%
- Applicants availing “A”, “E”, “F”, “G” grade loan who posses NCM <75%
- Applicants availing “C” grade loan who posses NCM <74%



High possibility of default by:

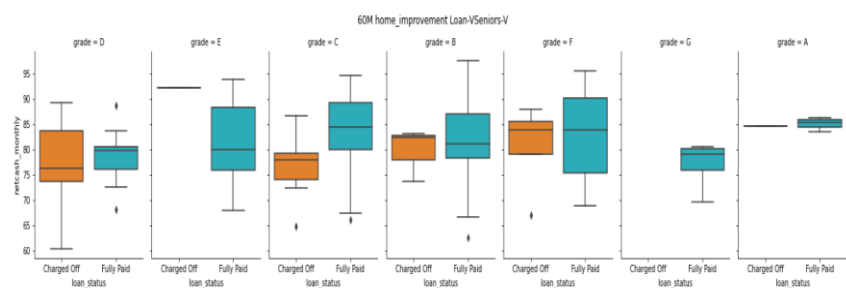
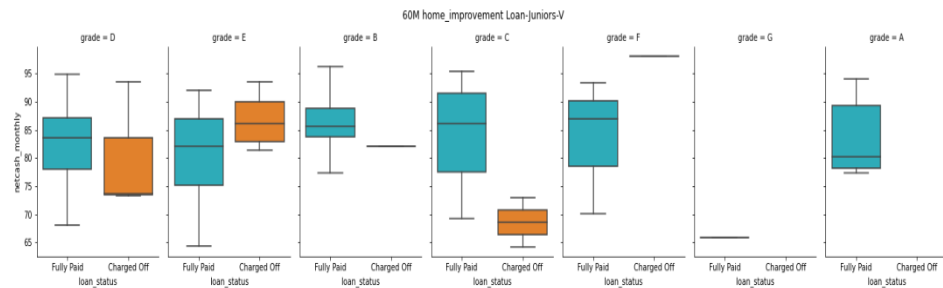
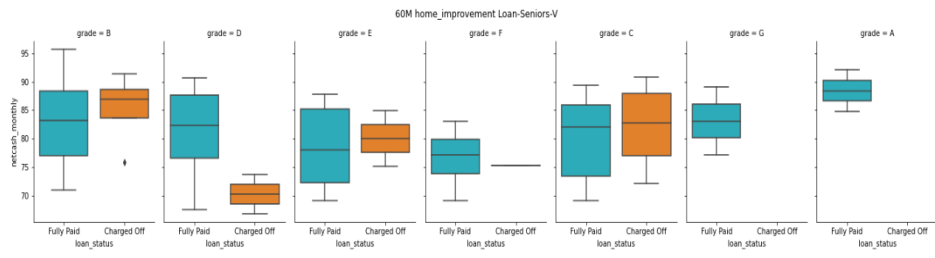
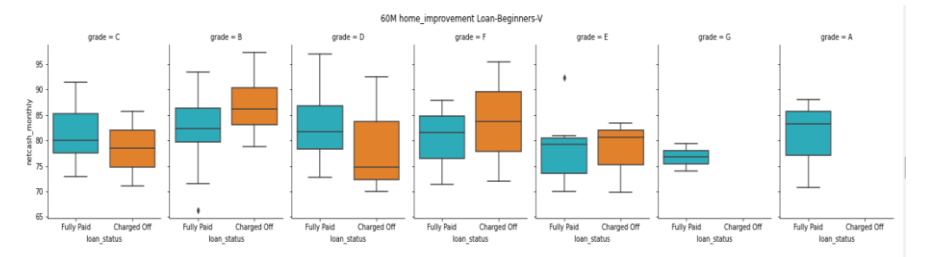
- More % of defaulters in Not Verified category compared to Verified category
- Applicants availing “B”, “C” grade loan who posses NCM <80%, <78%
- More % of defaulters in “B” & “C” loan compared to “A” / “D” loan



Multivariate Analysis: 60M- Home Improvement Loan

High possibility of default by:

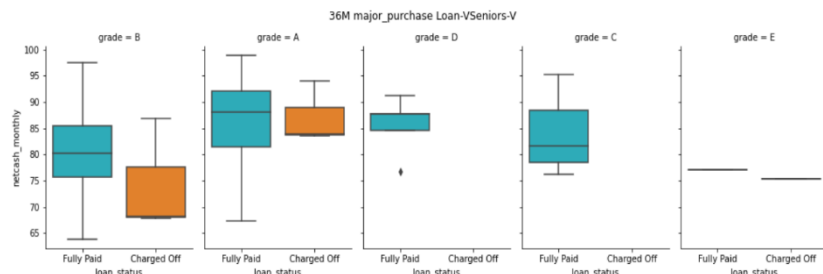
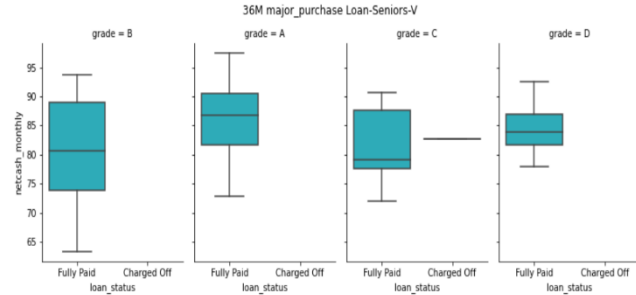
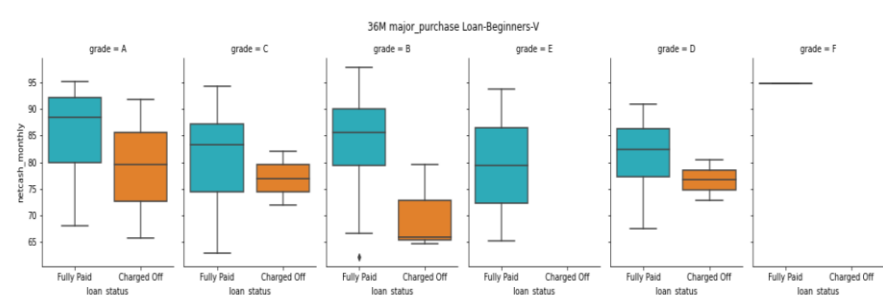
- Compared to 36M loan the defaulter % is less
- Applicants availing “B” grade loan who posses NCM <84%
- Applicants availing “C” grade loan who posses NCM <81%
- Applicants availing “D” grade loan who posses NCM <85%



Multivariate Analysis: 36M-Major Purchase Loan

High possibility of default by:

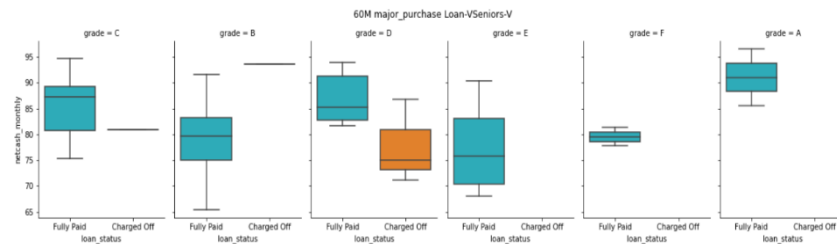
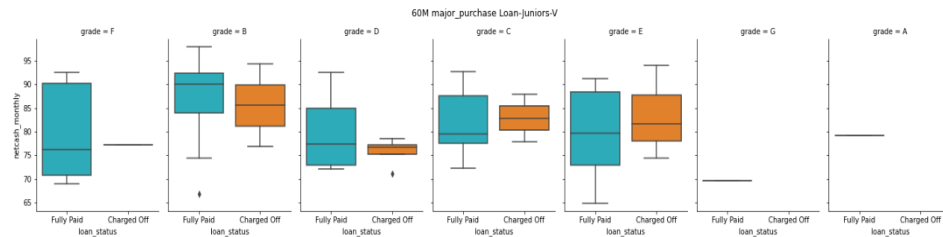
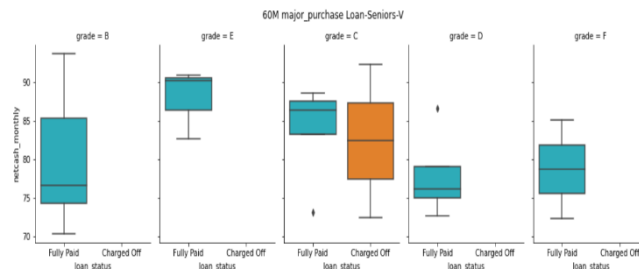
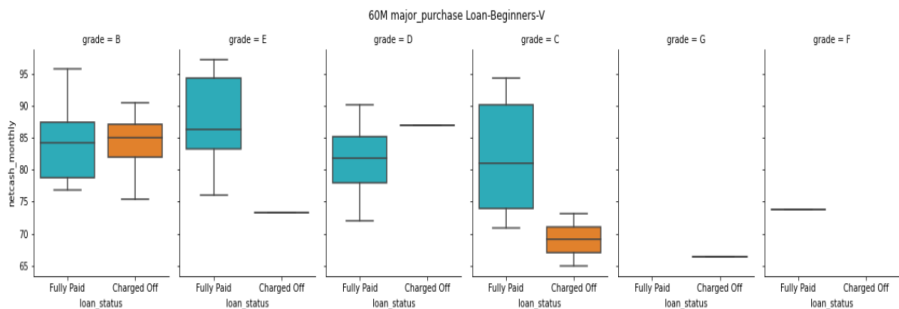
- % default by Juniors is less compared to others
- Applicants availing “A”, “C” grade loan who posses NCM <82%
- Applicants availing “B” grade loan who posses NCM <80%
- Applicants availing “D” loan who posses NCM <90%



Multivariate Analysis: 60M-Major Purchase Loan

High possibility of default by:

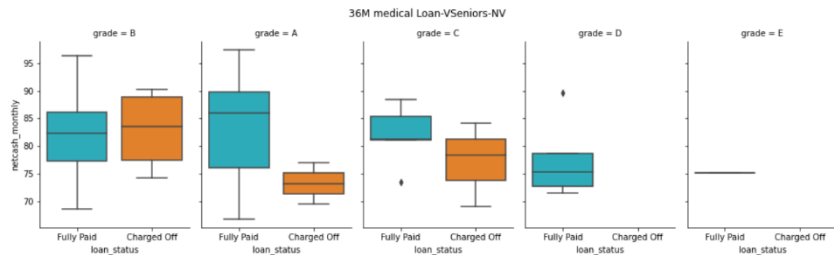
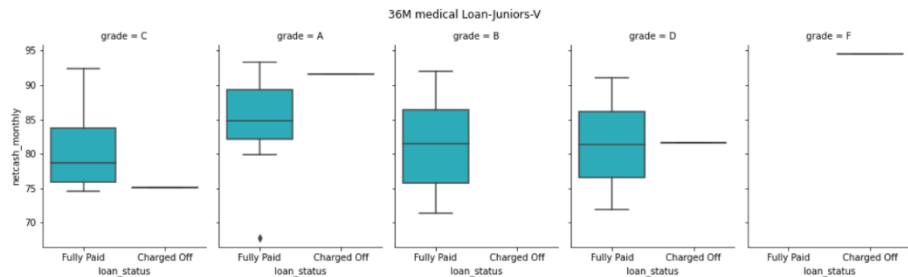
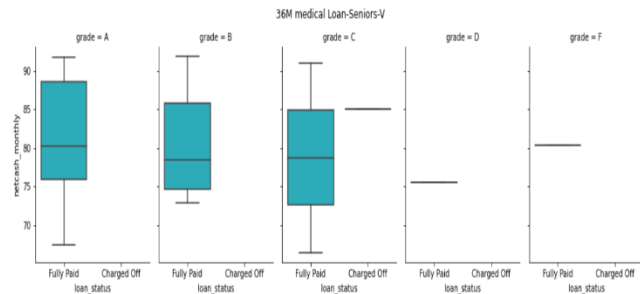
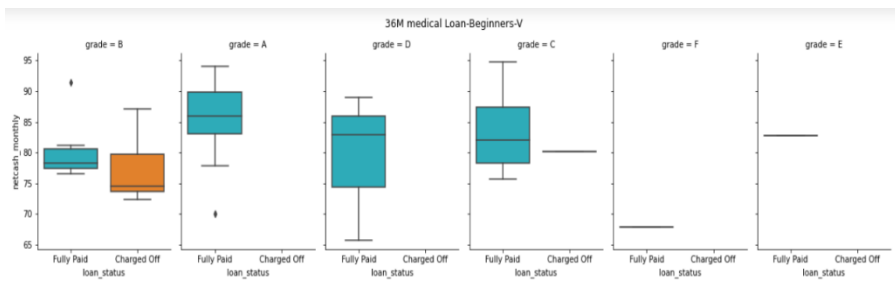
- % defaulters in Seniors, Vseniors is less compared to Beginners & Juniors
- Applicants availing “C” grade loan who posses NCM <82%
- Applicants availing “D” grade loan who posses NCM <83%
- More % of defaulters in “B” & “C” loans compared to other loans



Multivariate Analysis: 36M-Medical Loan

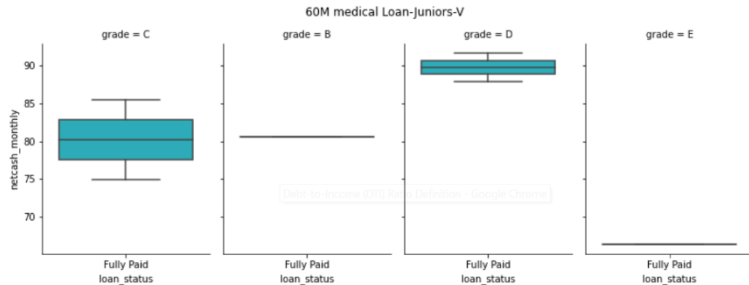
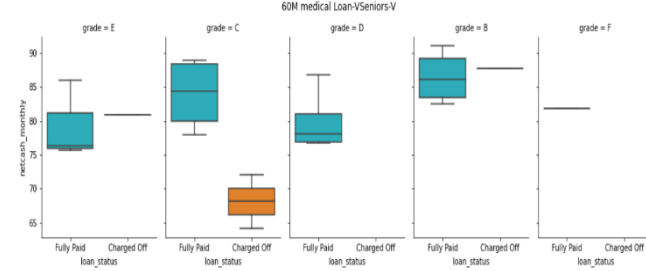
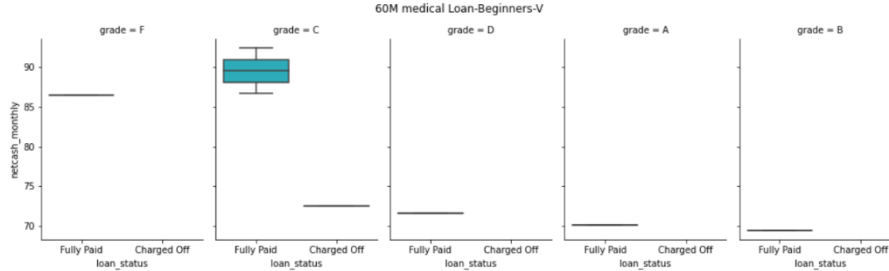
High possibility of default:

- Beginners availing “B” grade loan who posses NCM <78%
- Verified juniors, seniors and vseniors are not likely to default
- Most of the defaulters are in Not-Verified category



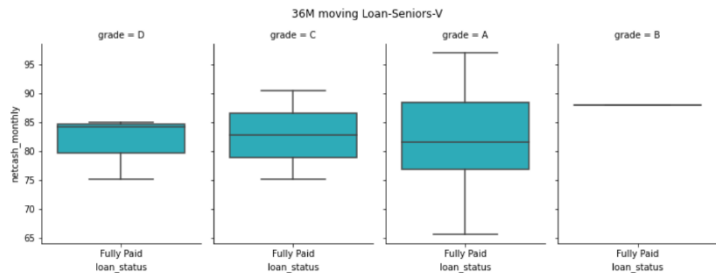
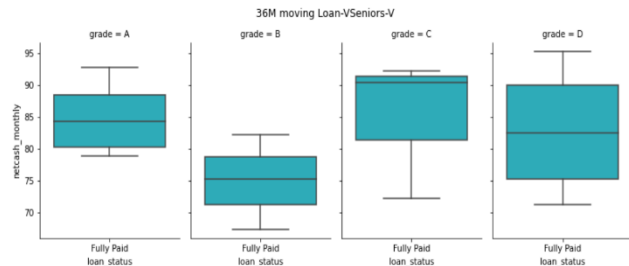
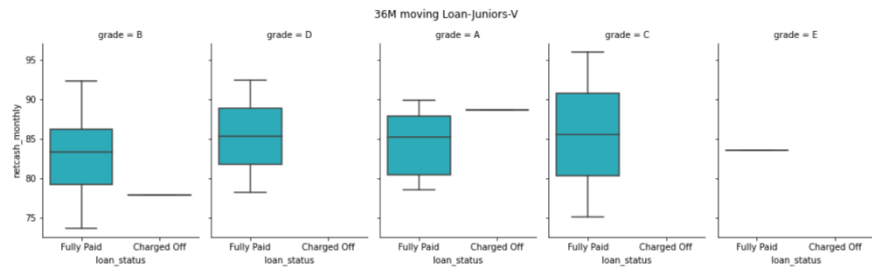
High possibility of default:

- Very less % of defaulters overall
- Only applicants in Not-Verified category are defaulting



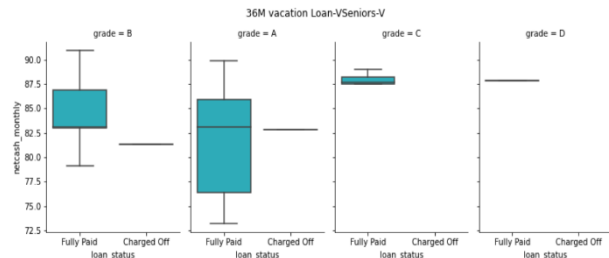
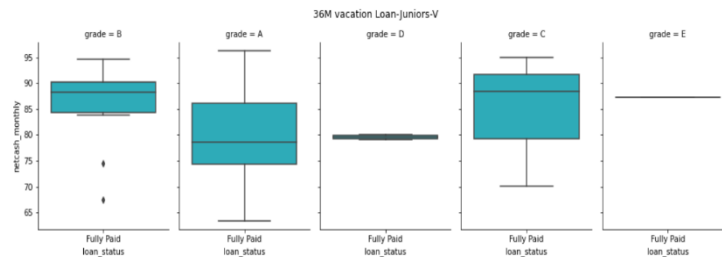
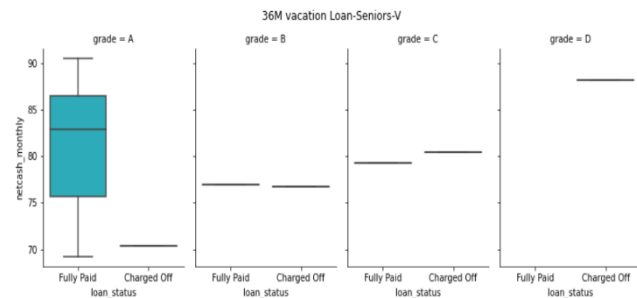
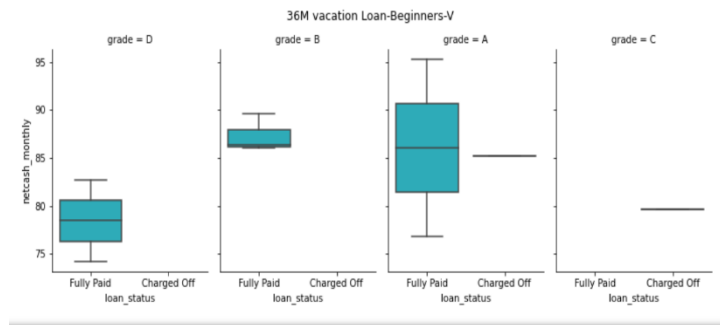
High possibility of default:

- Most of the Verified loan applicants are not likely to default
- Most of the defaulters in moving loan are in Not-Verified category
- Very less % of defaulters compared to other loan



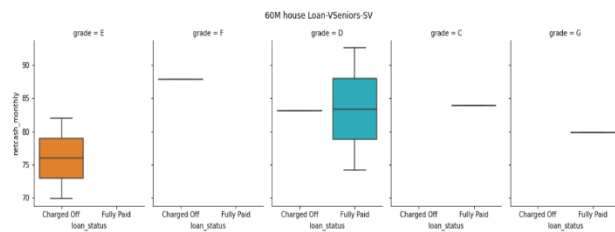
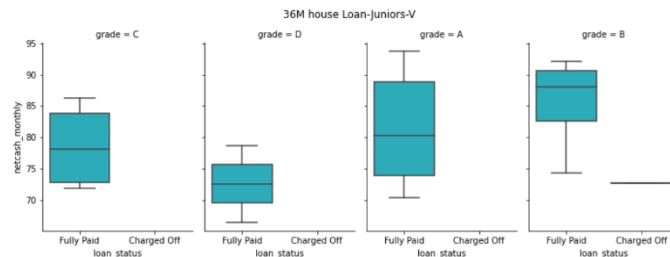
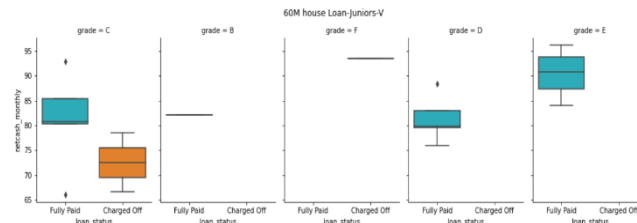
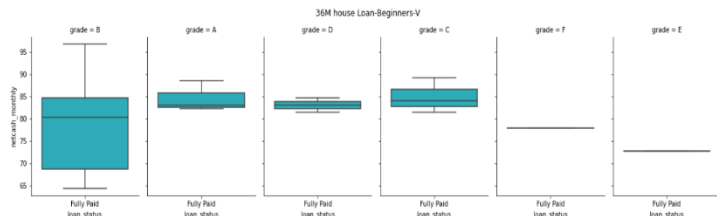
High possibility of default:

- Most of the Verified loan applicants are not likely to default
- Very less % of defaulters compared to other loans
- Most of the defaulters in vacation loan are in Not-Verified category



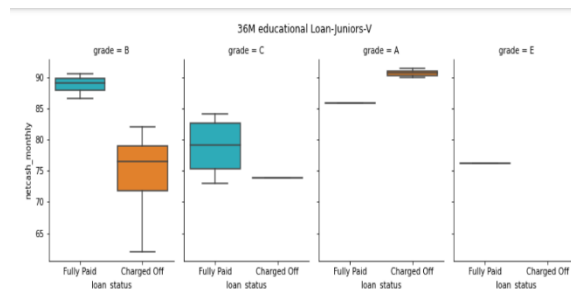
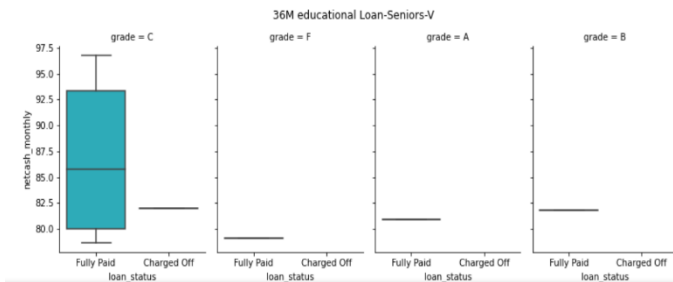
High possibility of default by:

- Most of the Verified loan applicants are not likely to default
- Very less % of defaulters compared to other loan
- Most of the defaulters in Not verified category



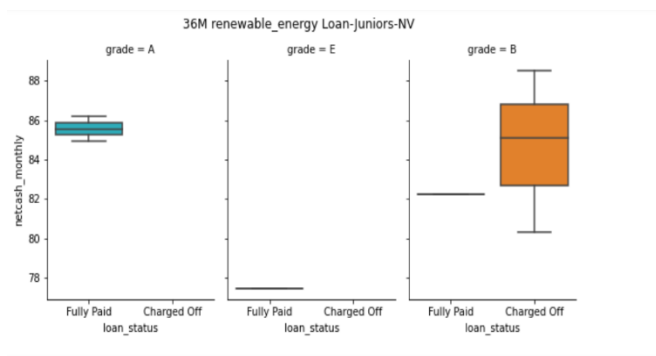
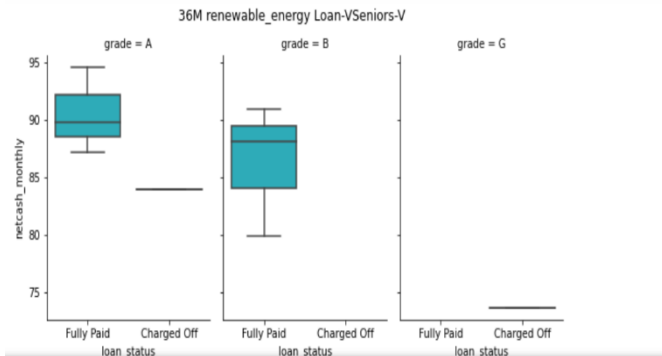
High possibility of default:

- Most of the loan applicants are not likely to default
- Very less % of defaulters compared to other loans
- Likely defaulters in Not Verified category



High possibility of default:

- Most of the verified loan applicants are not likely to default
- Most likely defaulters of RE loans are in Not Verified category
- Very less % of applicants and defaulters



upGrad



Thank You!