LENDING CLUB CASE STUDY

Submitted By:

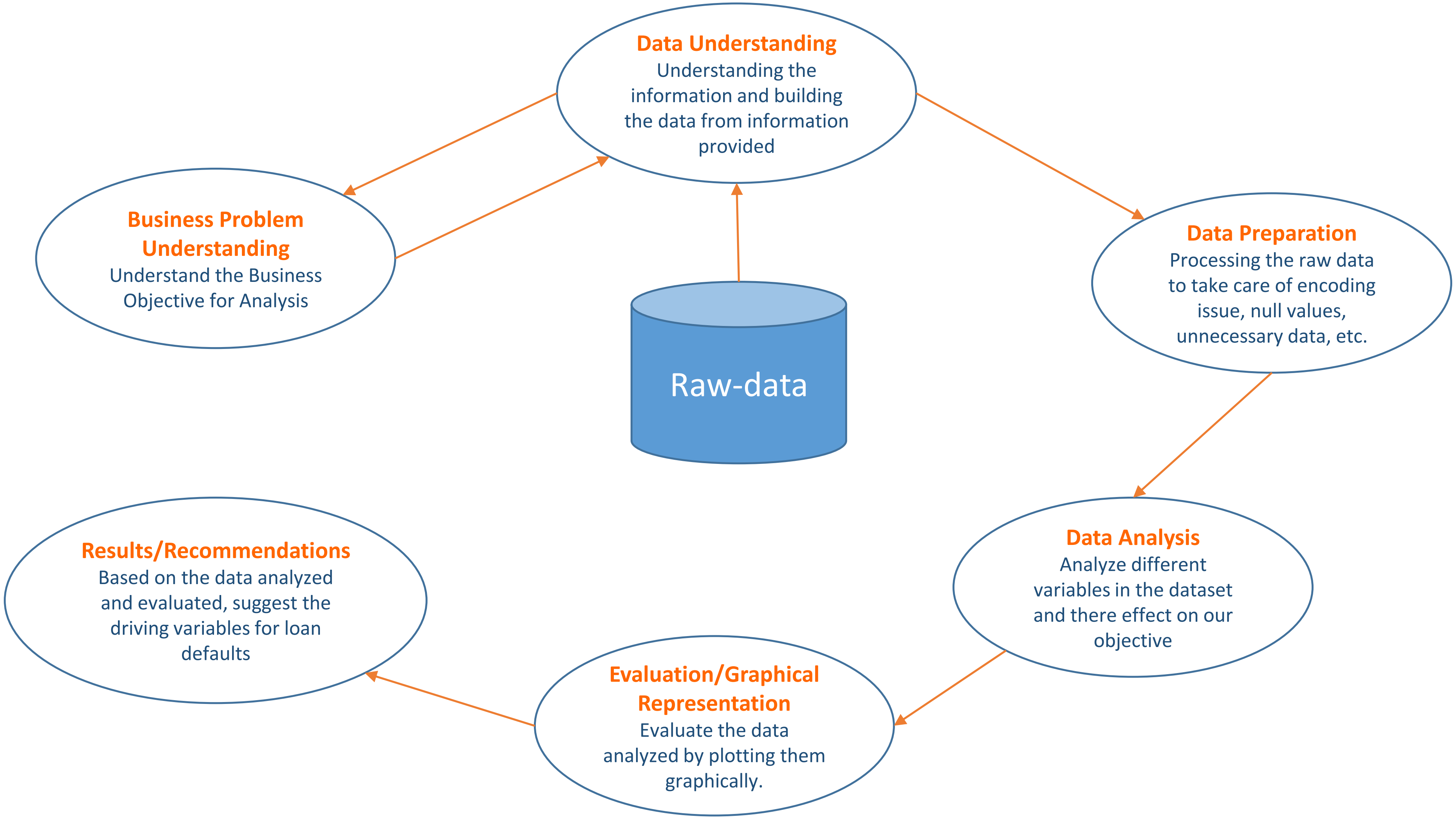
Mudassar Majgaonkar (Group Facilitator)

Nitin Gupta

Main Objective of the Case Study

- Aim of this case study is to help identify the consumer finance company the risky loan applicants or defaulters.
- By identifying these risky loan applicants, the company can reduce lending such loans and thereby cutting down the amount of credit loss.
- The objective of this case study is to find the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators leading to loan default. The company can utilize this knowledge for its portfolio and risk assessment.

Problem solving methodology



Note: The problem solving methodology is based on known CRISP-DM framework

Data Analysis

Data Pre-processing

- Calculate the percentage of null data in the dataset
- Remove the null data above 50% threshold.
- Take care of rest of the null values in dataset.

Variable Analysis

- Analyze the variables and categorize them Quantitative (Numeric) variables and Categorical Variables
- Normalize all the variables to standard format.
- Filter out variables which are of little or no importance in business context.

Univariate Analysis

- Analyze the distributions and count/frequencies of the single variables in the dataset.
- Analyze the segments of data for these variables (**Segmented Univariate Analysis**)
- Create Derived Variables if any (Extract month and year from timestamp variables)

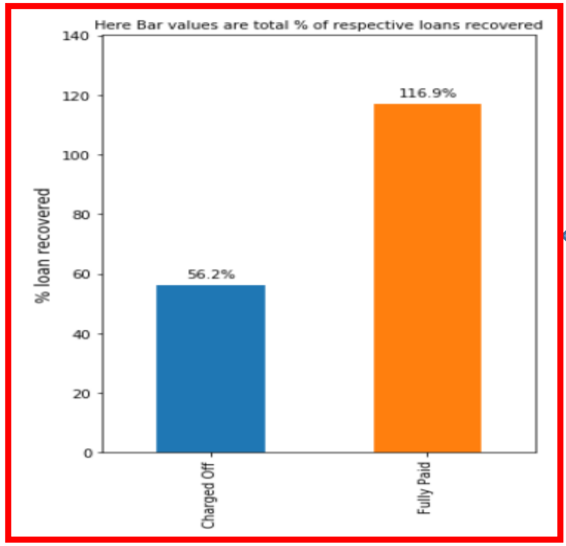
Bivariate Analysis

- Analyze the distributions and correlation between 2 variables at a time.
- Create the correlation matrix and analyze the dependent and independent variables.
- Analyze these related variables using joint distributions.

Derived Metric Analysis

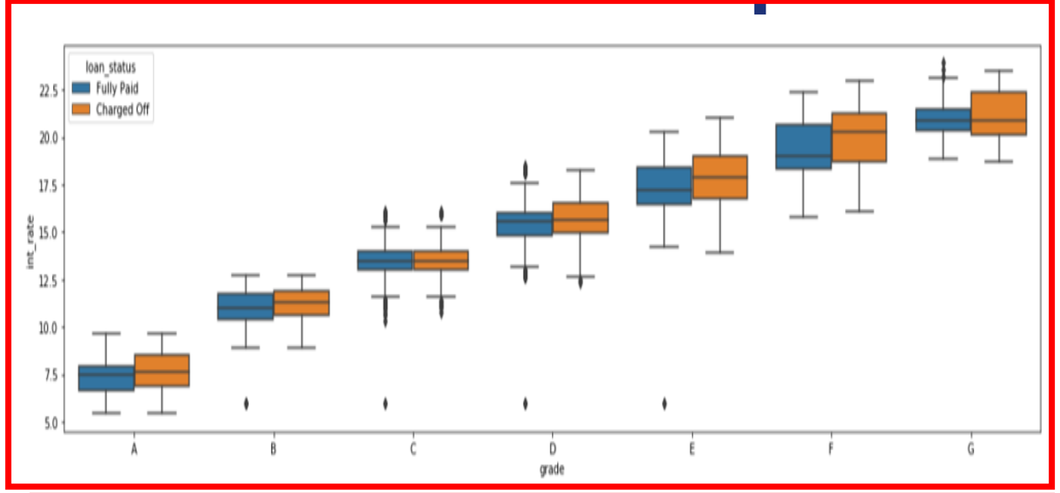
- Analyze the derived metrics from the earlier context.
- The derived metrics can be type driven, data driven or business context driven.
- Analyze these metrics according to the business objective.

Understanding the Loan Information and Characteristics

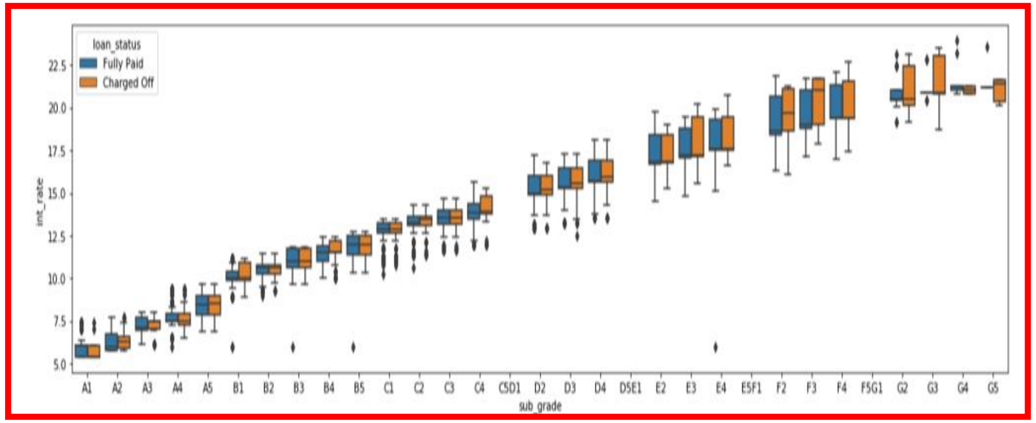


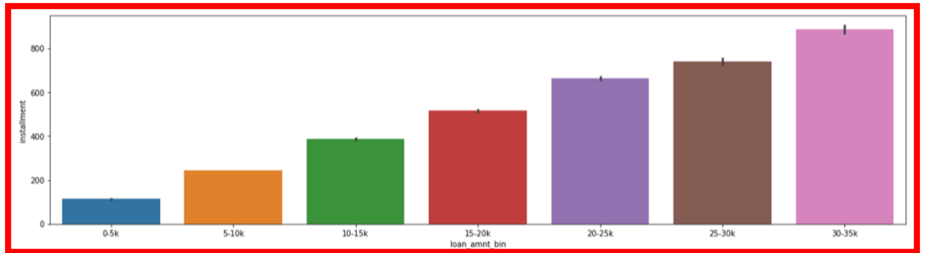
Loan Status
Fully Paid or
Charged Off

Loan Grade(A to G)
Interest rate
increases with
increase in grade



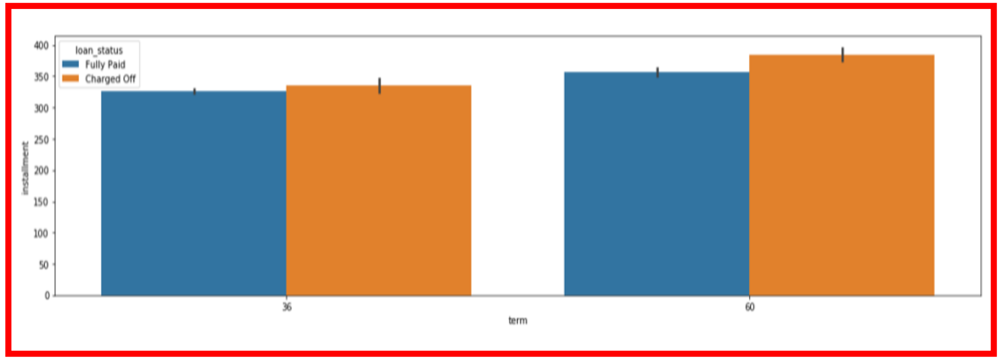
Loan Sub-Grade(A1 to
G5)
Interest rate increases
with increase in grade





Loan Amount
and Installment
-
Installment
Increases with
increase in Loan
Amount

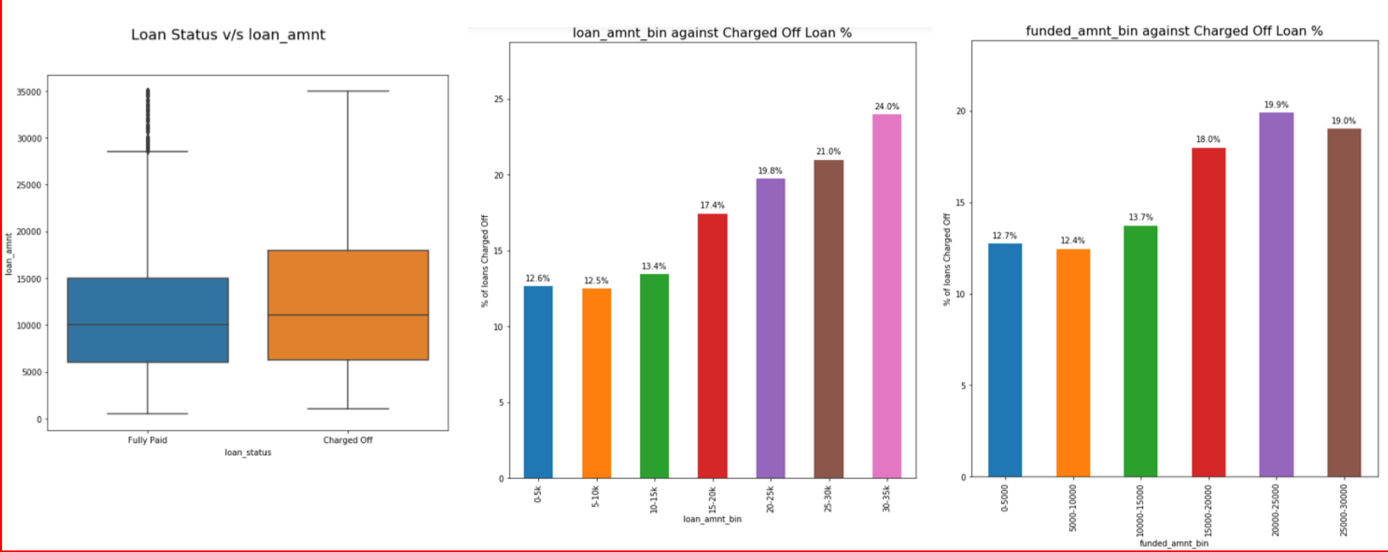
Loan Term
increases with increase in
loan amount and
installment



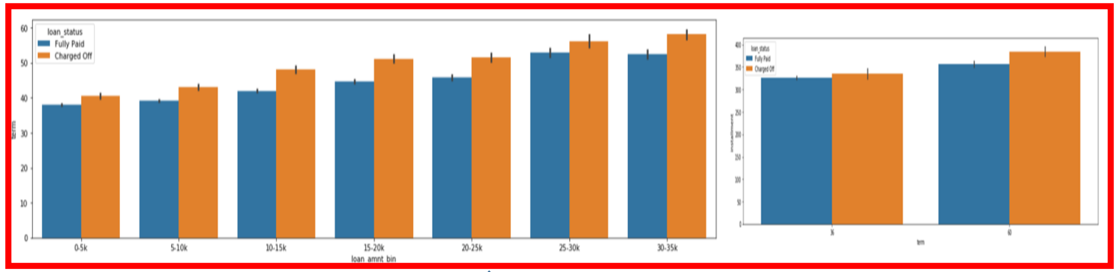
Defaults Based on Loan Characteristics

	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv
id	1	0.992632	0.187379	0.19715	0.276367
member_id	0.992632	1	0.18888	0.197977	0.287395
loan_amnt	0.187379	0.18888	1	0.978546	0.941765
funded_amnt	0.19715	0.197977	0.978546	1	0.963888
funded_amnt_inv	0.276367	0.287395	0.941765	0.963888	1

Loan Amount, Funded Amount and Funded Amount Investor are positively correlated. So analysis for one variable will hold true for others too

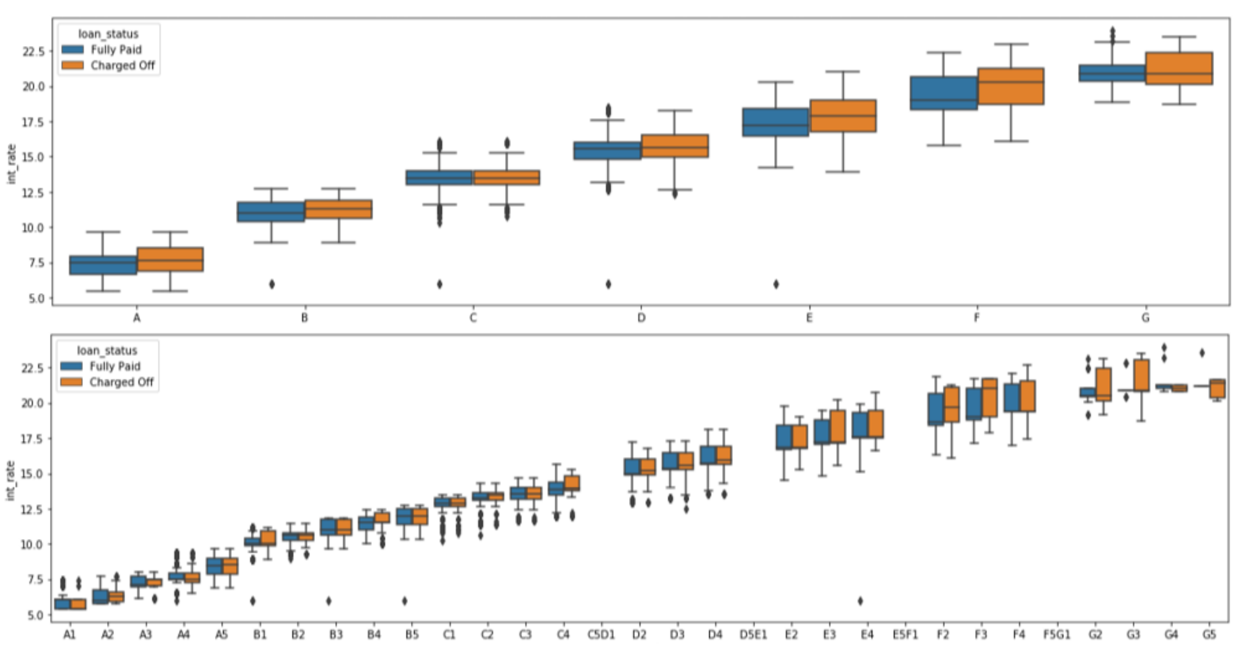


Probability of Loan Default increases with the increase in Loan Amount, Funded Amount or Funded Amount from Investor

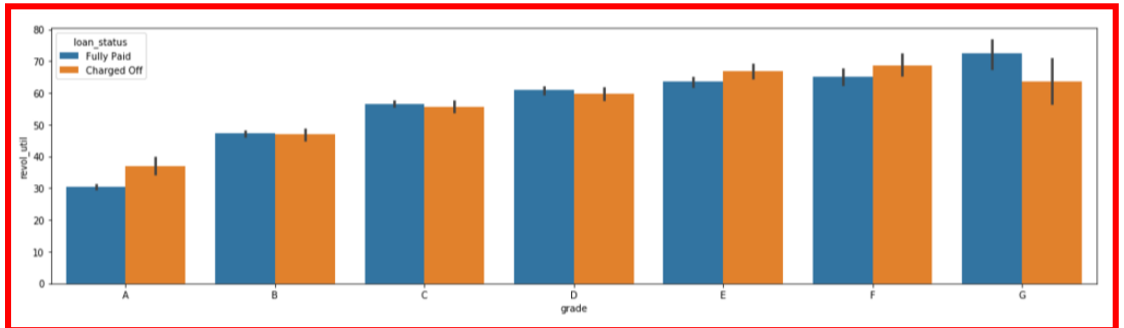


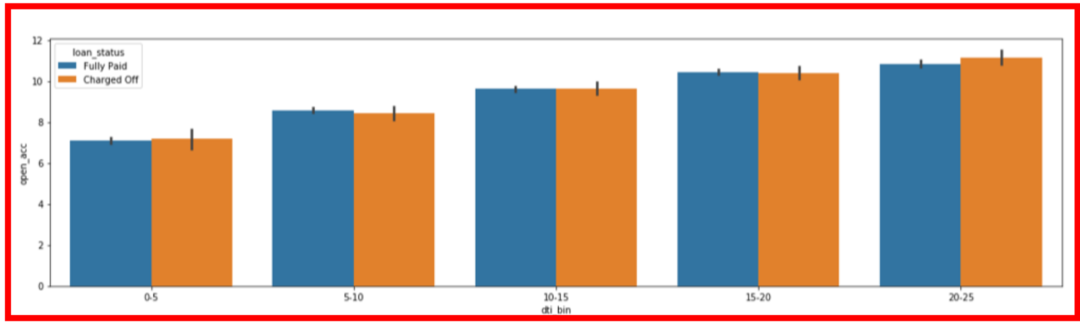
Higher the loan amount, higher is the interest, longer is the term for loan fulfillment and higher is the risk for loan default

Higher the Loan Grade/Sub Grade, higher is the interest rate and higher are the chances for loan default



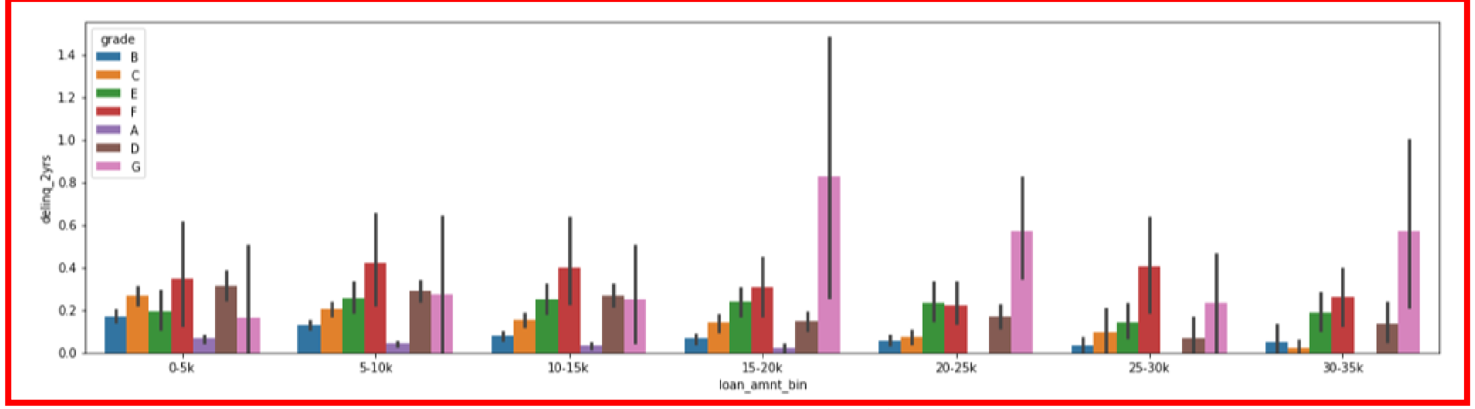
Defaults Based on Loan Characteristics – contd…



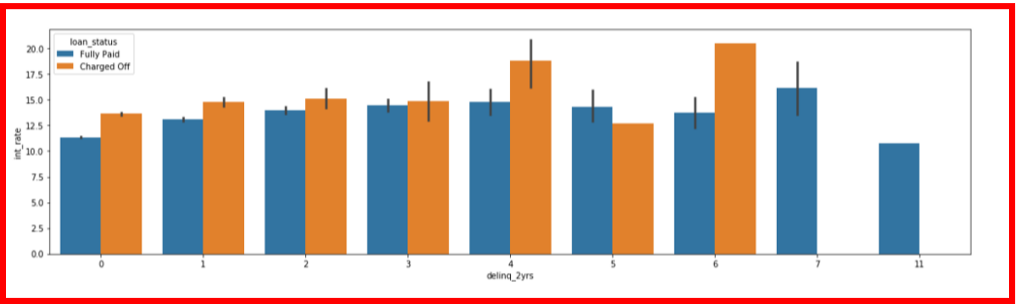


Higher the dti higher are the values for open_acc. And thereby higher is the risk for loan default

Higher the revol_util, higher is the risk for loan default. (higher loan amounts are associated with higher interest rates and with higher grades making them more prone to be defaulted)

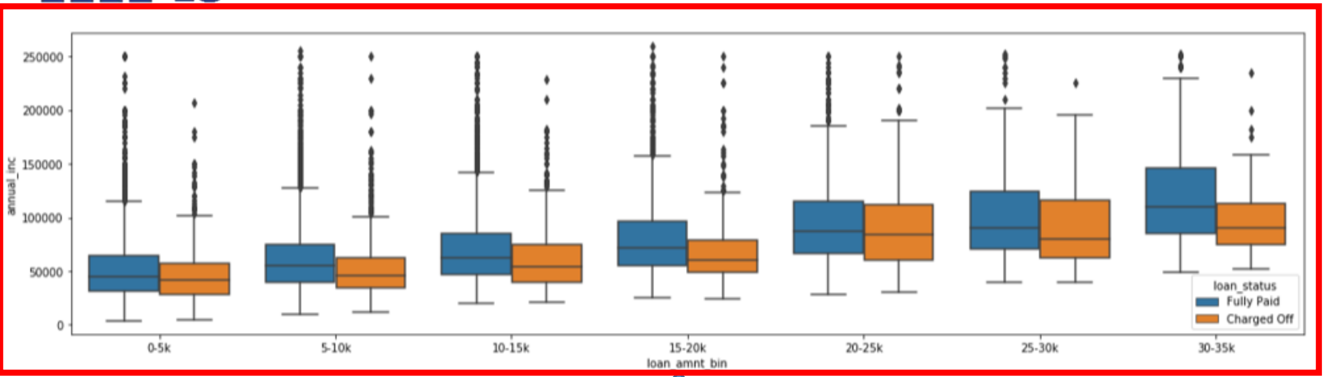


The interest rate offered increases with the number of delinquency of the borrower. With the increase in interest rate the applicant is more probable for default.

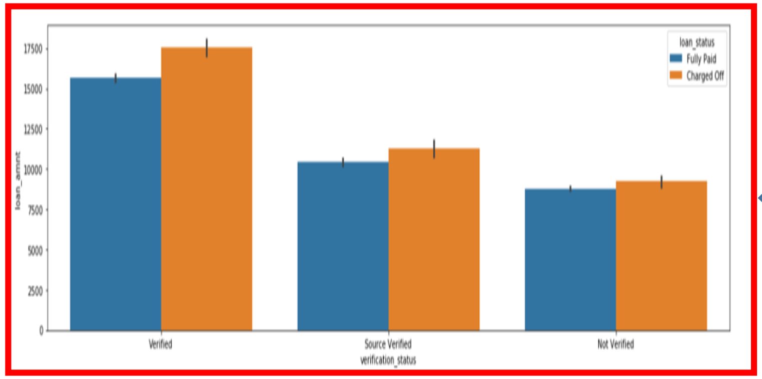


Not many loans receive investment with higher number of delinquencies. Even though the loan amount requested is low, these loans are not invested much in. These loans are thereby considered risky and more prone to default.

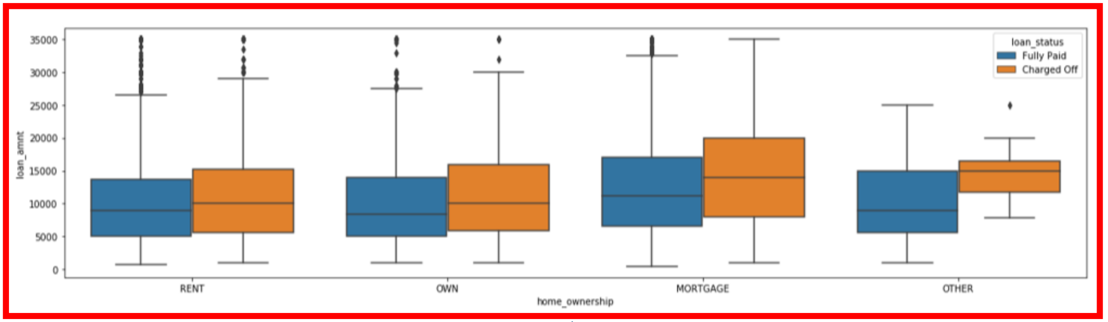
Defaults Based on Loan Applicant Behavioural Characteristics



Applicants with high annual income have applied for low loan amount. **As the annual_inc increases for low loan_amnt, the chances of loan default/charged off decreases. This assumption may not hold true for the applicants with high annual_inc who have applied for higher loan_amnts.**

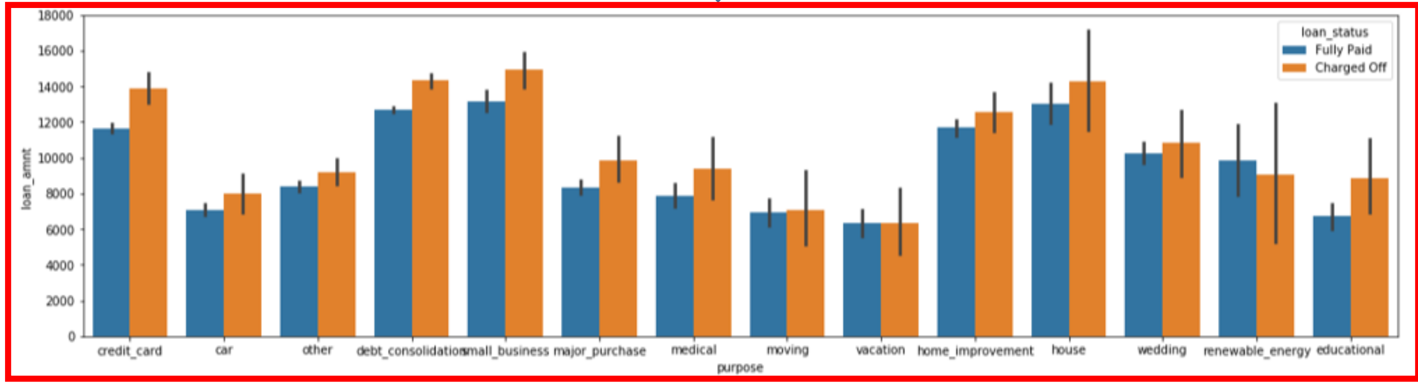


Higher loan amounts are most often verified Its not the verification status per se that's responsible for the default but the higher loan amount is. Higher loan amount are most often verified by the company and because of higher loan amount they are most likely to default/charged off.



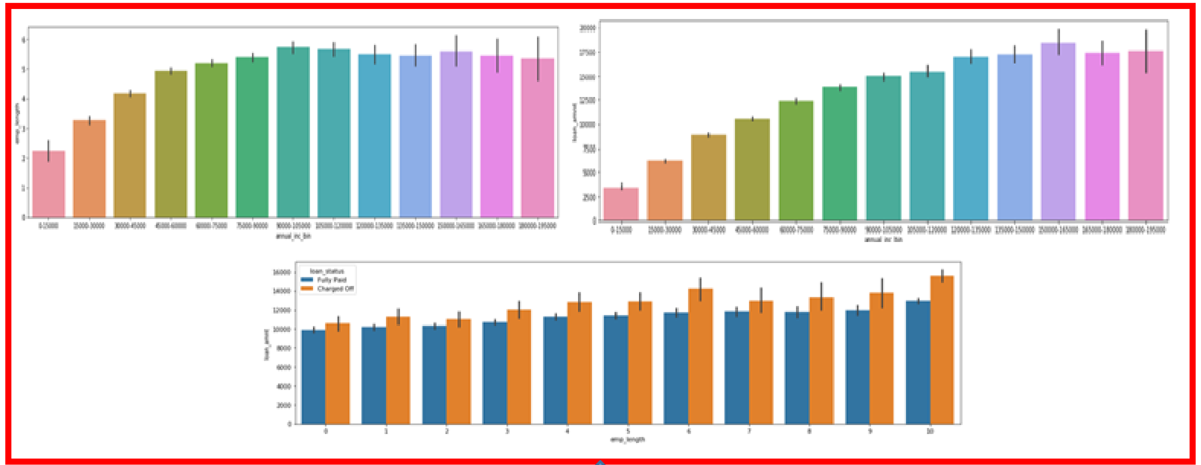
Maximum loan amount were requested on mortgage, rent and own. Mortgage seems to be more prone to loan default, followed by rent and then own in that order.

Maximum loans and loan amount is applied for small_business, credit_card, debt_consolidation and house. Small_business, credit card and debt_consolidation are more likely to default on their loans.

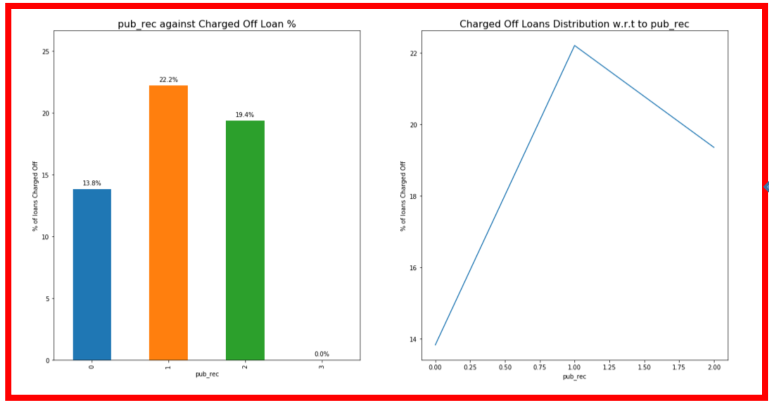


Defaults Based on Loan Applicant Behavioural Characteristics contd..

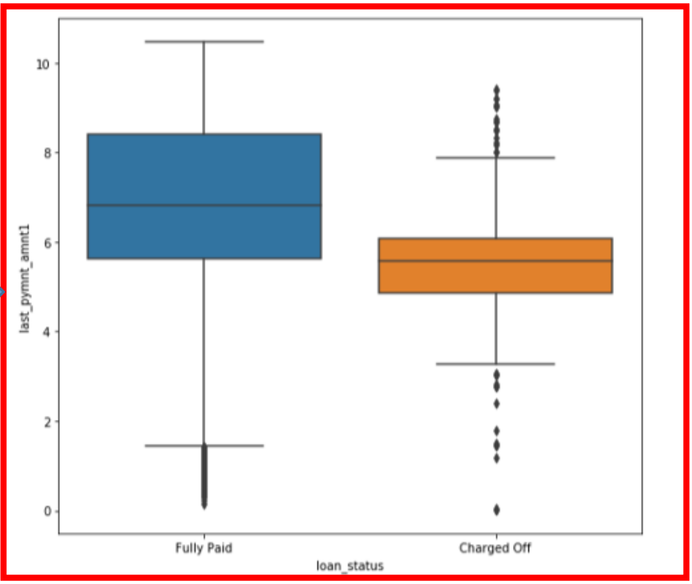
UpGrad



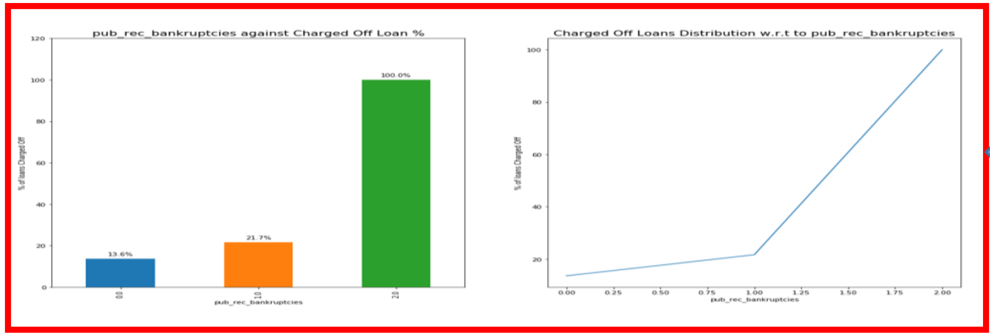
As the employment length increases so does the chances for loan default increases. This has not much to do with the employment length per se and more to do with the loan amount applied for. With more employment length comes more annual income and with more annual income comes more loan amount and with that more risk of loan default.



Even 1 derogatory public record the chances of defaulting loan increases significantly. So having at least 1 derogatory public record is strong indication of defaulting on loan



The amount of last payment goes on decreasing the loan applicant is probably going bankrupt and hence unable to pay the amount with higher chances of loan default



Charged off rate significantly increases from 1 indicating the applicants having at least 1 public record of bankruptcies record are more likely to default on the loan.

Recommendations Based On Analysis

Refrain from approving higher amount loans for applicants with lower annual income.

Verify the source of income for higher amount loans.

Refrain from approving loans to applicants with prior bad public record.

Approve loans to applicants with longer employment term and annual income.

Reduce approvals for loan Applicants with mortgage since they are highly likely to default.

Reduce approving loans for small businesses and debt consolidation. These are highly likely to default.