
# Lending club – Risk Analytics

**By Rohit Chaudhary and Giridhar Tarare**

# Our Understanding

**Consumer finance company ( Banking and financial )** which 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.

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

❑ Two type of decision will take for any loan application as "Loan Accepted " and "Loan Rejected"

❑ Identify the loan Accepted and load Rejected applications & criteria

| Loan Accepted | Loan Rejected |
|---|---|
| 1. **Fully paid:** Applicant has fully paid the loan (the principal and the interest rate)<br>2. **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'<br>3. **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 | The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset) |

# Business Objective and Scope

- Analysis & Identify these risky loan applicants profiles and loan application:
- Borrowers can easily access lower interest rate loans through a fast online interface.
- Company wants to understand the driving factors (or driver variables) behind loan default
- Advised to independently research a little about risk analytics (understanding the types of variables and their significance should be enough).

Scope : Risk Analytics

- Mention the problem statement and the analysis approach briefly
  - Understand the data and make data smart ( data cleaning , data standardized ,removed redundant columns, etc.)
- Explain the results of univariate, bivariate analysis etc. in business terms

- Include visualizations and summaries the most important results in the presentation

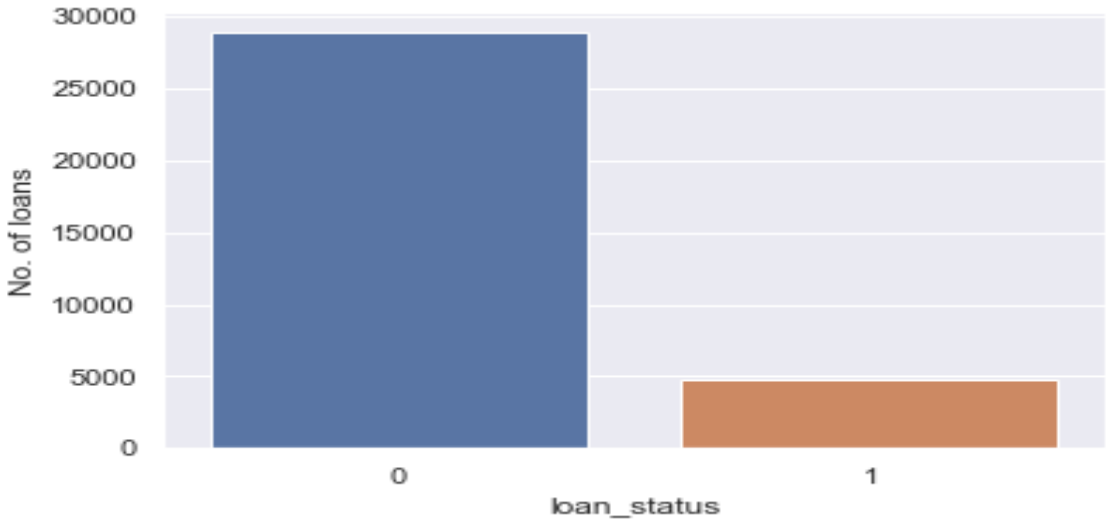
# Data Understanding and Data Cleaning and Manipulation

# Implementation Summary : Risk Analytics ( Loan Applications)

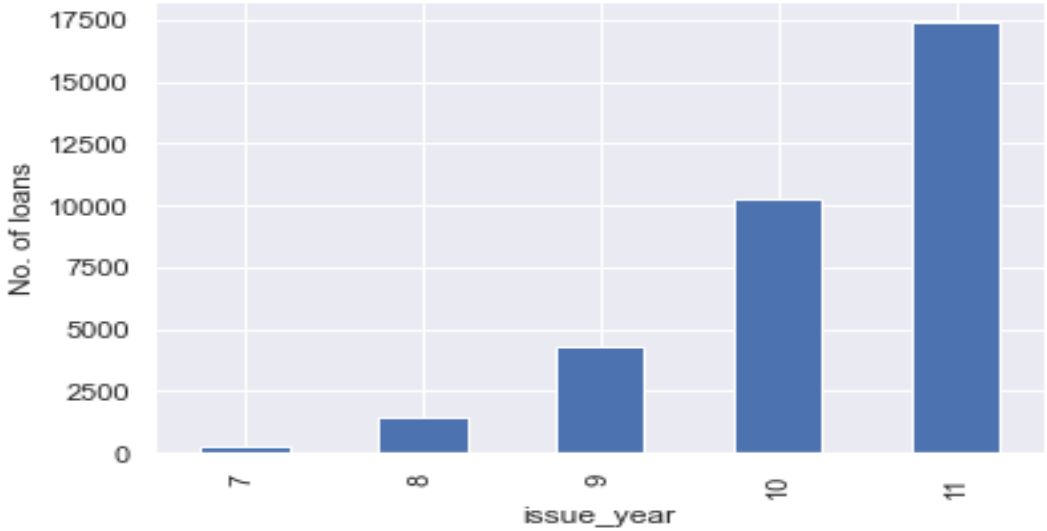
1. Data Load - Loan data and related dictionary data
2. Load data into Data frame and understand the data & Columns
3. Data Cleaning and standardized ( removed unwanted columns, rows, removed null values from columns, amount percentage , zip code etc.)
   - Check the number of null values in the columns
   - Dropping columns & rows with all the Null values
   - Deleting all the columns which are not required and having only unique value and all unique values like id, member_id, url etc.
   - Dropping rows with null values
   - Interest rate column and removing % and converting to float type for analysis
4. Validate the loan amount and Funded amount to identify any discrepancy in amounts for approval process.
5. Univariate Analysis for Risk Analytics
6. Bivariate Analysis for Risk Analytics
7. Multivariate Analysis for Risk Analytics

# Data Analysis

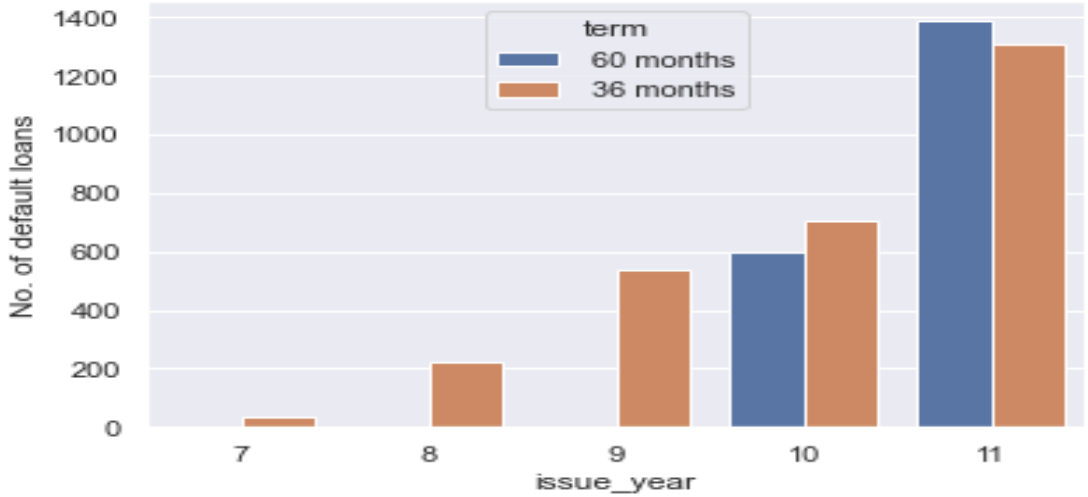
# Univariate Analysis




Loan defaulted by loan applicants [0 -28873 and 1- 4794 ]

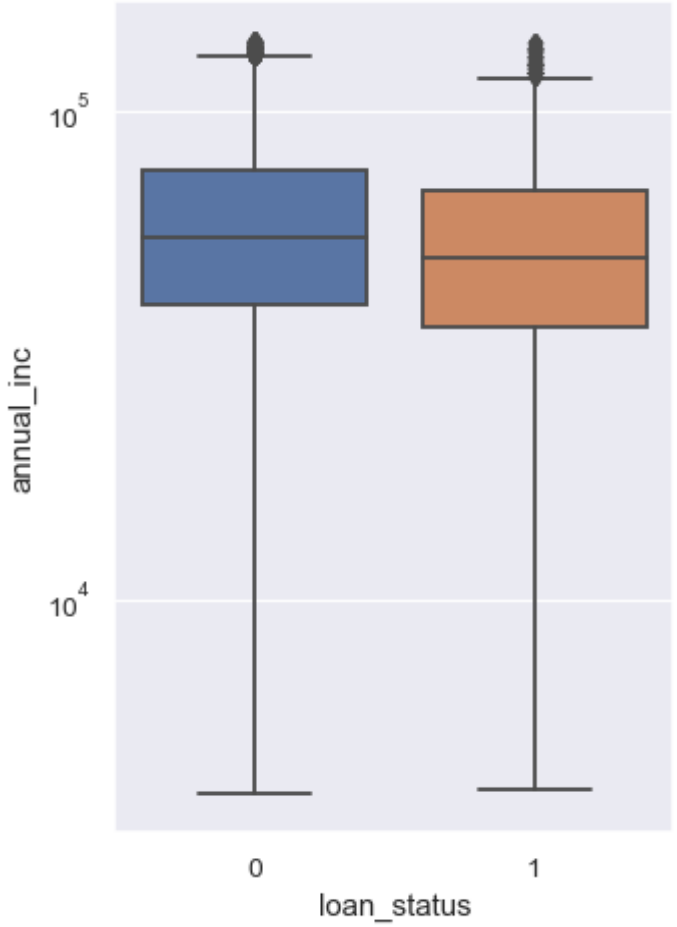



Distribution of loans over the year



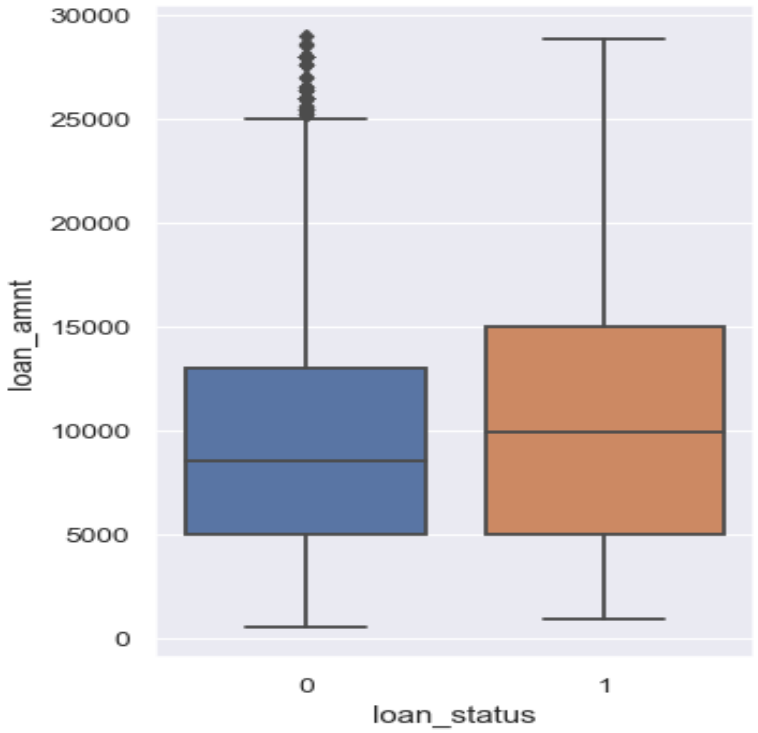

Distribution of loan compared to term of loan (60 & 30 months )

# Univariate Analysis continue…



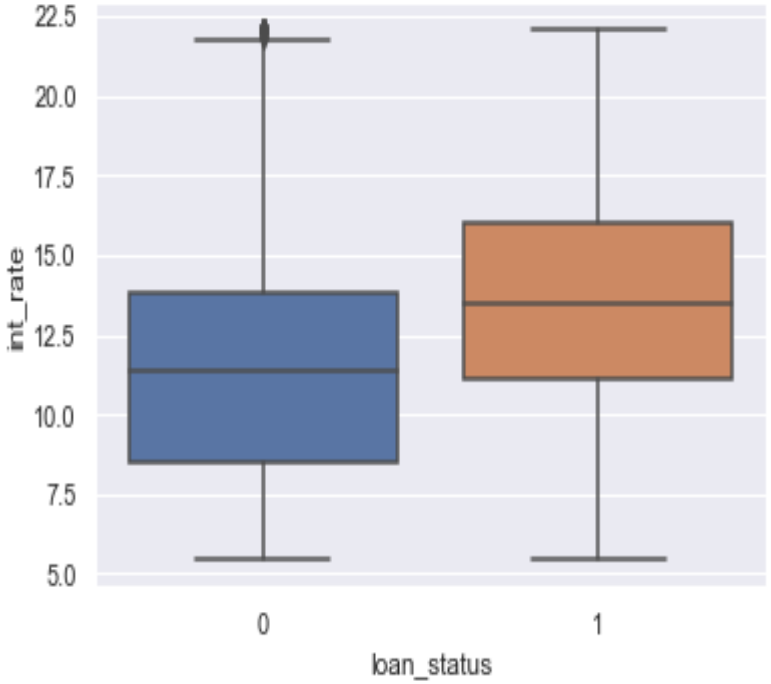

0 = Fully Paid
1= Charged Off

Annual income on log scale of loan applicants for loan paid off and default loan




0 = Fully Paid
1= Charged Off

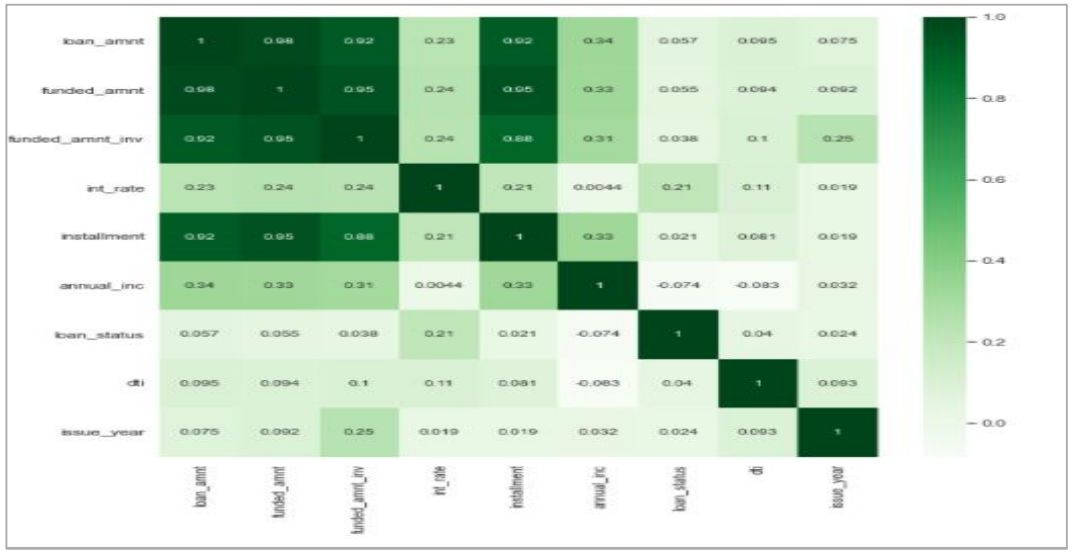
Loan amount and loan status



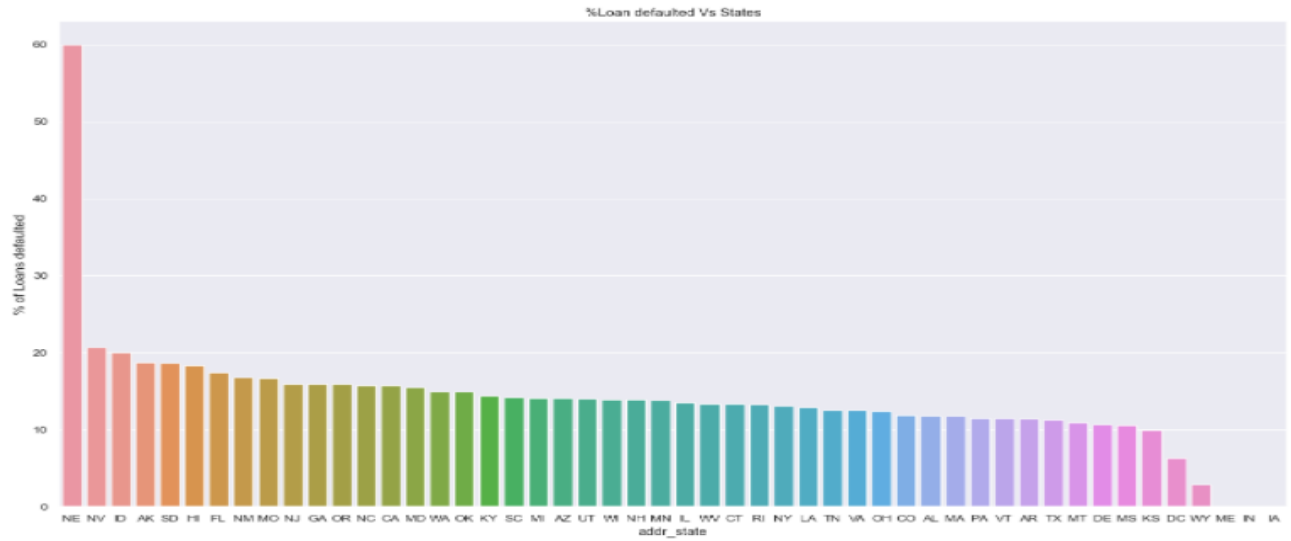

0 = Fully Paid
1= Charged Off

Interest rate will be low other then providers

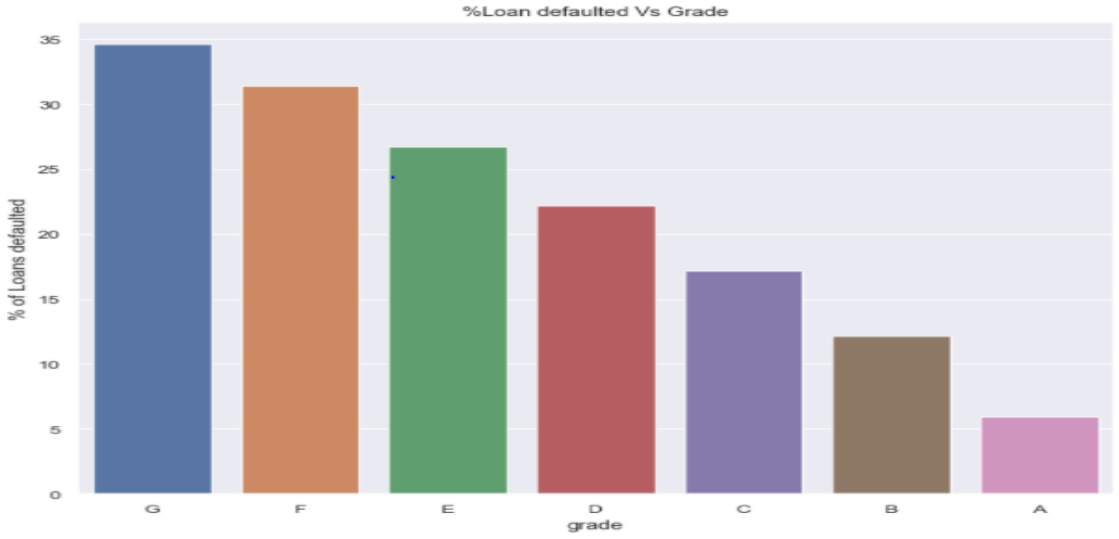
# Bivariate Analysis



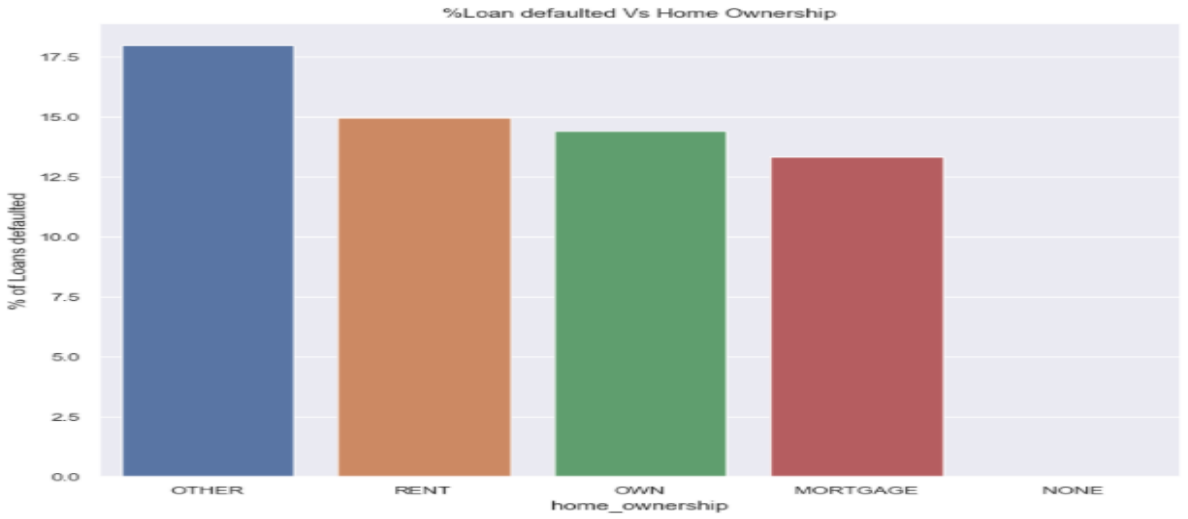

Funded amount and funded amount by investor has a strong correlation with each other and with loan amount, therefore theses two columns can be dropped




Seems like NE is having maximum loan defaults among all the states. Let us check no. of loans counts in NE

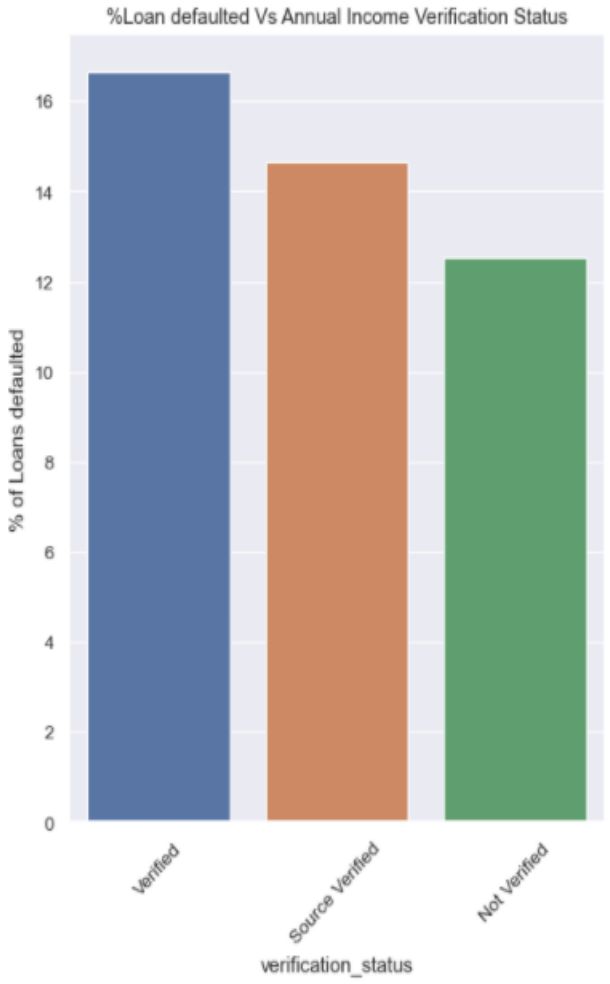



As apparent from above, grade G loans are having highest default rate

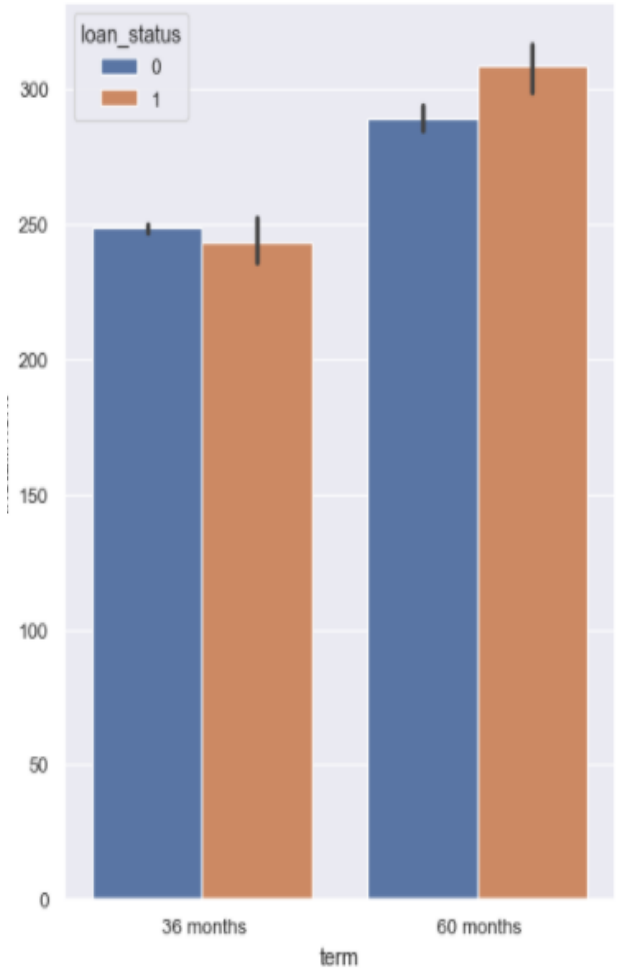



Loan defaulted vs home ownership

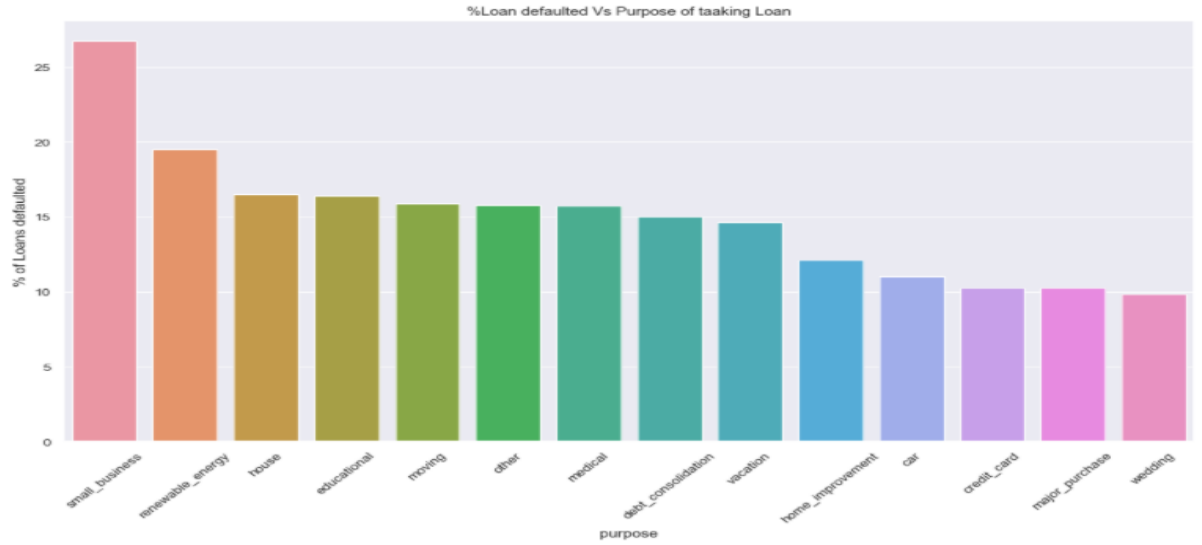
# Bivariate Analysis Continue..



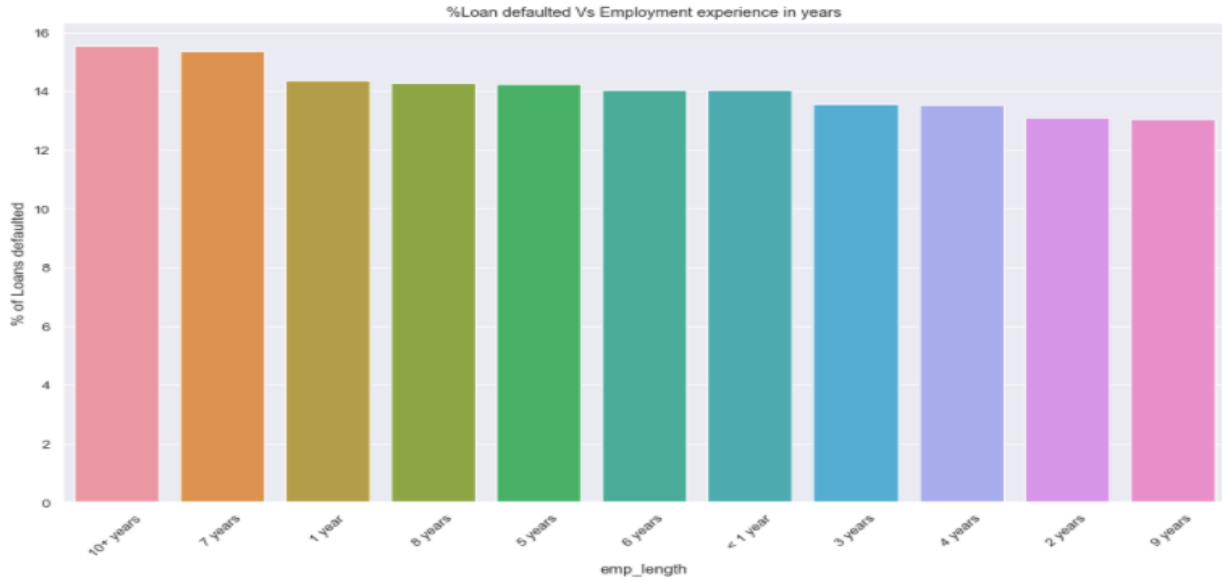

Verification status for all defaulted customer's




Loan Status of the months of 30 & 60 months

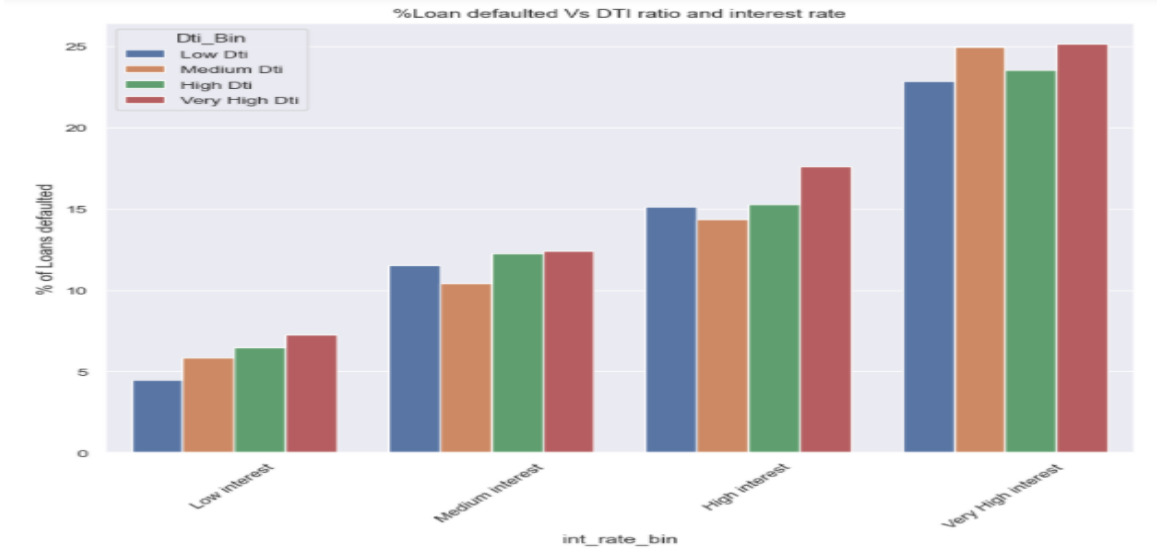



Loan defaulted vs Purpose of loan taking

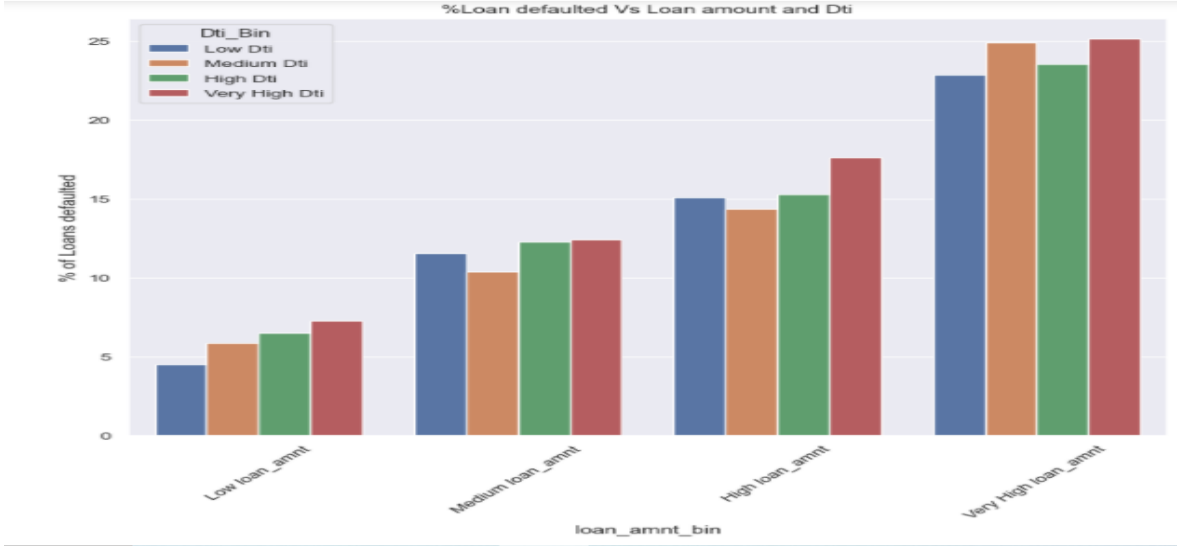



Employees with experience more that 10 years are more likely to default loans.

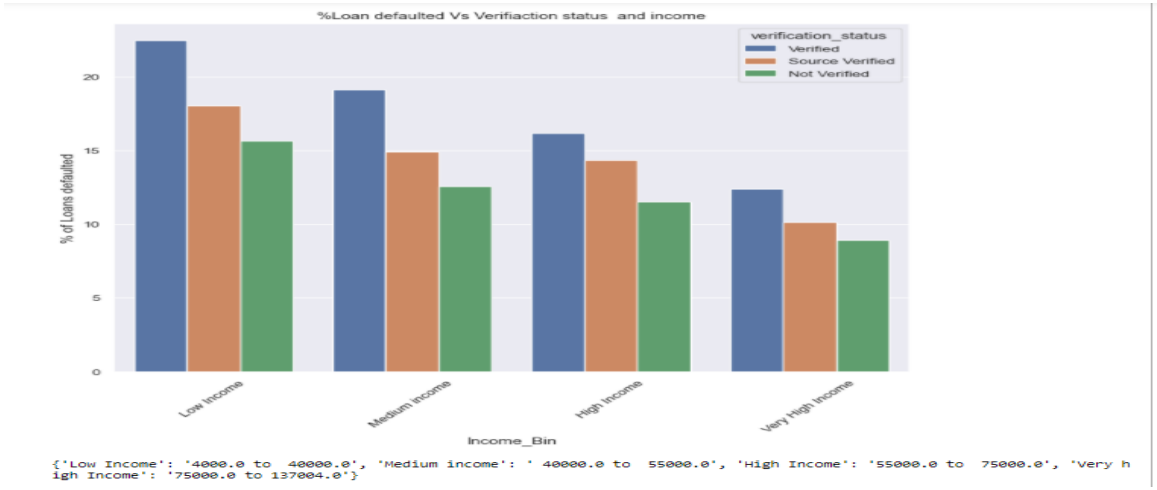
# Multivariate analysis



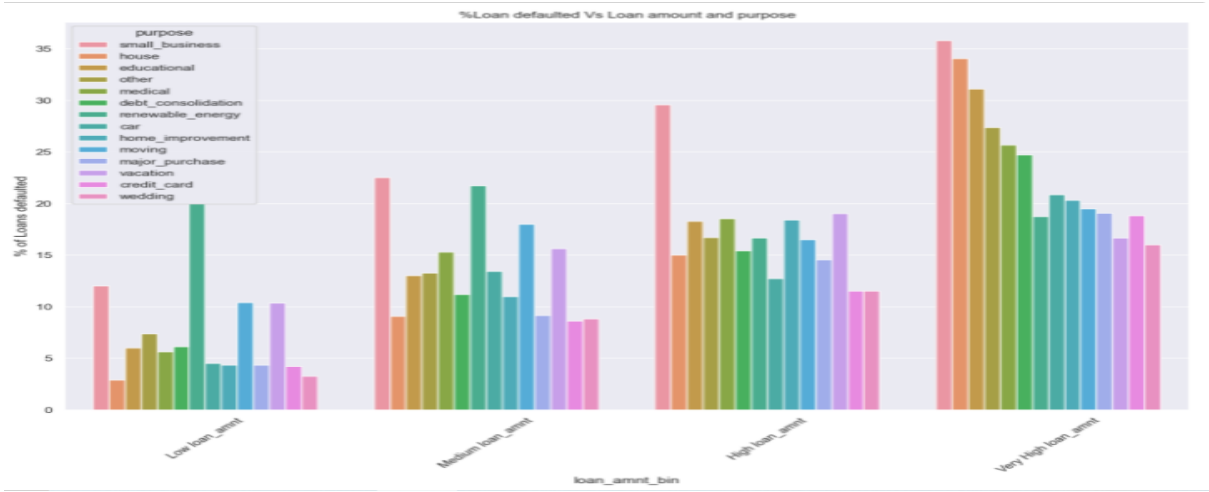

High interest rate, medium DTI is yielding almost same default rate as very high DTI




Loan defaulted vs Loan amount and DTI




Verification status : Variation is as per earlier variation brought out above. No abnormality observed.




Combined effect of Loan amount and purpose on default % - Loan amount bin vs Purpose

# Conclusion

1. No. of loans being offered are in an increasing trend from 2007 to 2011 and so does the loan default rate.
2. Default rate is much higher for 60 months term compared to 36 month term ever since its inception in 2010. In fact default rate is on decreasing trend for 36 months term may be due to increasing 60 month term loan. Therefore no. of 60 months term loan may be reduced.
3. It is observed that there is higher tendency of loan default when loan is being offered for purpose of small business, therefore no. of loans to small business may be reduced or offered after due diligence.
4. Less defaults are being observed for lower interest rate irrespective of dti, therefore, low interest rate loans may be given in more nos.. Also. Loans are not being offered for higher dti(> ~25%) and higher interest rate (> ~15%) may be due to high default risk associated with such loans.
5. For low and small loan amount default rates are higher for debt consolidation purpose of loan and also for very high loan amount, house and education loan are having max default rate after small business.
6. In case of very high interest rate, medium DTI is yielding almost same default rate as very high DTI.

# Recommendation

- Adjust the operating principles with market  and investment shifts.
- Reduce the risk of defaulter and proactively get applicant’s other data to validate for loan approve or reject.
- Increment the revenues of bank because of less defaulter and regular payment by customers.

# Thank you!