Lending Club Case Study

About the Company

This company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface. It faces significant financial losses due to credit loss from 'risky' loan applicants who default on their loans.

Key Objectives / Problem Statement:

- **Identify Risky Loan Applicants**: Determine the driving factors / driver variables of loan default by borrowers, i.e. the variables which are strong indicators of default.
- **Reduce Credit Loss by making use of driving factors**: Minimize the financial losses due to defaults, by identifying and potentially rejecting high-risk loan applications based on different driver variables which are strong indicators of default.

Analysis Approach:

- Performed Exploratory Data Analysis (EDA) to investigate the dataset and identify patterns or trends associated with loan defaults.
- While performing EDA, data cleaning activities have been performed and some major steps taken were: removal of columns with all null values, removal of columns not contributing in analysis and removal of 'loan status' as current and keeping only the Fully Paid and Charged off loan status.
- As per the problem statement, analysis is more focused on loan default as count rather in monetary terms. Few places (as needed in the analysis), monetary term has also been considered.
- Examined company's historical record / report trend on loan and derived outcomes based on Univariate and Bivariate analysis.
- Identified all key variables that are strong predictors of loan default.
- Develop a comprehensive understanding of the borrower profiles and loan characteristics that contribute to higher default rates.
- Made extensive use of python and it's libraries like pandas, matplotlib and seaborn.

Lending Club Case Study

Conclusion:

After analysing multiple driver variables using EDA, there are several observations and conclusions drawn (as listed below) which leads to higher loan default rates among loan borrowers.

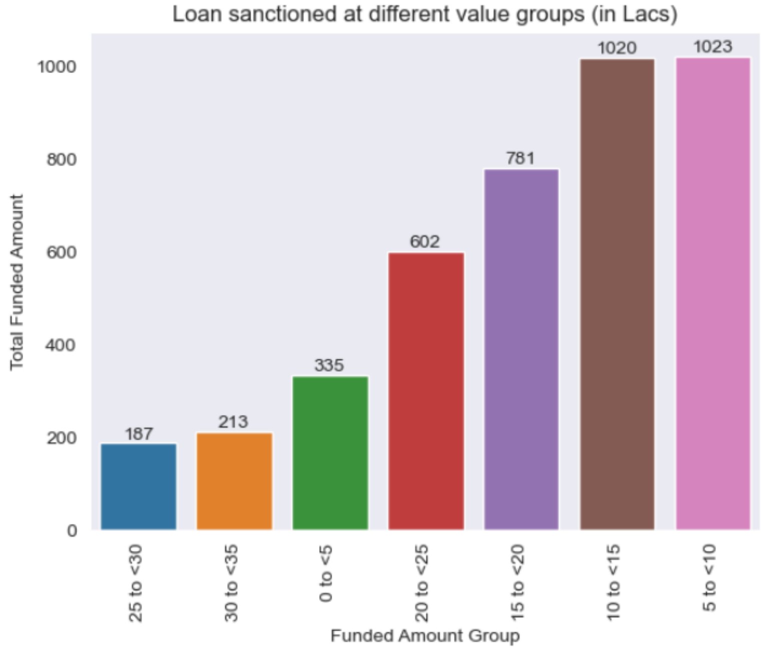
Key Outcomes from Univariate & Bivariate Analysis:

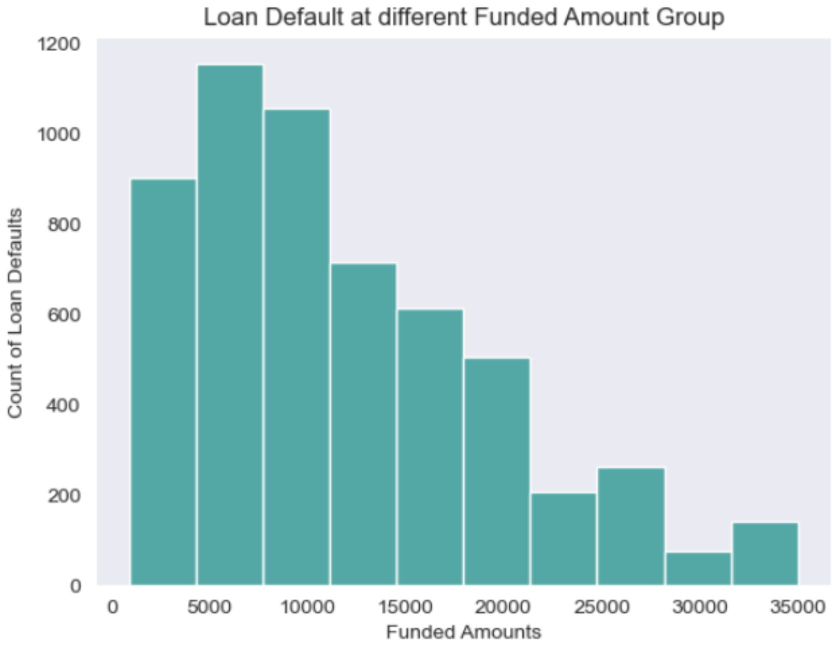
1. **Sanctioned Loan Amount**: Loan sanction volume being higher at lower disbursement values and thus, loan default volume is higher at same level. Bivariate analysis leads to the conclusion that – Special due-diligence is needed for loan sanction in the disbursement bucket of “20K to 25K” and “30K to 35K” due to higher loan default percentage.
2. **Loan Tenure**: Loan with higher tenure (60 months) has the higher loan default percentage (+25%) based on total loan sanctioned at same tenure.
3. **Interest Rates**: Loan default increases with increase in intertest rate and highest loan default is between “20% to 23%” interest rate. Interest rate should be maintained at market standards / mean level. Issuing risky loans at higher interest rate would make the loan more risky.
4. **Loan Grades**: Loan grades assigned by the LC seems to be working fine with higher loan defaults at lower grade & sub-grade level. Loan sanctions to lower grades should be minimized and more scrutinized.
5. **Employee Experience**: There isn’t much variance in loan default at different employee service length and loan default percentage at each employee service length ranges between +12% to +15%. Loan disbursed to employee with service length of 7 and 10+ years need better scrutiny since loan default percentage at these length are +15%.
6. **Loan Purpose:** Loan disbursed for the purpose of “Small-Business” need better scrutiny since it has the highest loan default percentage of +27%. Also, since major loan business volume comes from loan purpose of “Debt-Consolidation”, loan for this purpose need to be better scrutinized.
7. **Geographical Factors**: Loans sanctioned to borrowers in the states of NV and NE show higher default rates.NV is more relevant since there are very few borrowers in state of NE with 60% loan default rate.
8. **Seasonal / Periodical factor**: Maximum loan default happens in the quarter-end months (month 3) and year end (i.e. Dec). This could be because, sales team grant more number of loans to applicants at liberal terms to achieve their sales targets.
9. **Borrower Income:** Source of income for low annual income groups should be scrutinized to minimize loan defaults. Also, employees with instalment of 20% to 25% of annual income need better scrutiny since the loan default is highest at this level.

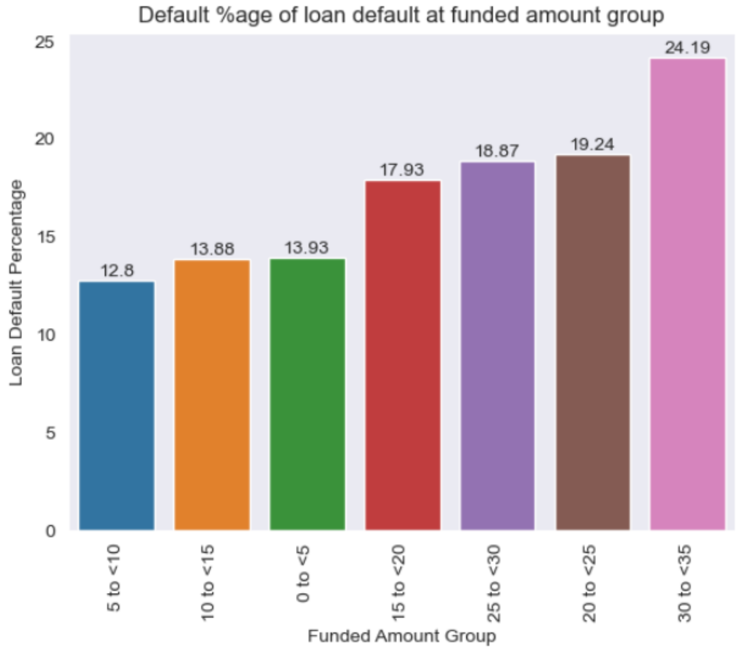
Overall Conclusion:

The analysis highlights that loan defaults are influenced by a combination of factors related to multiple loan driver variables. By focusing on these key indicators, the company can further process its risk assessment and lending strategies to mitigate the risk of loan defaults, thereby reducing financial losses. Specifically, loans with lower amounts, specific interest rate ranges, particular borrower grades, term and specific purposes or geographical locations require more stringent evaluation criteria to manage and minimize default risks effectively.

Sanctioned Loan Amount



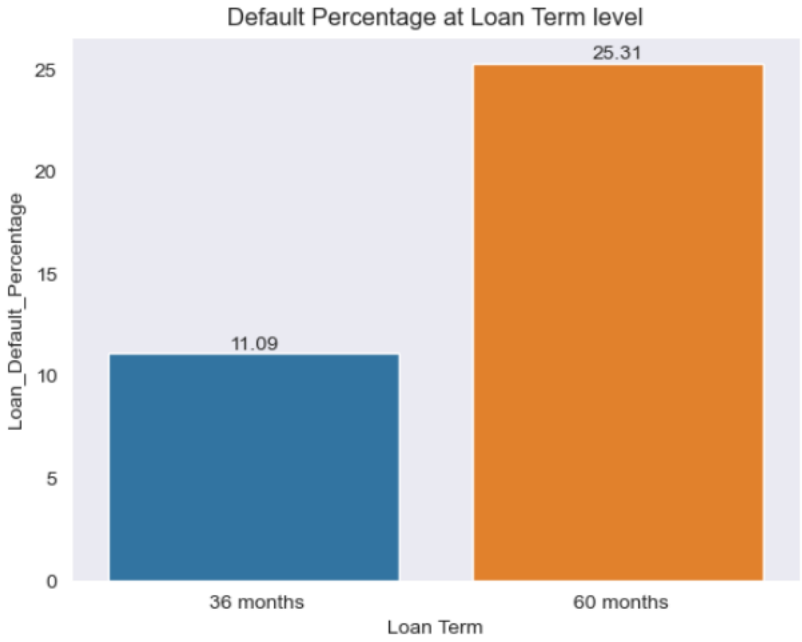




Observations & Key Call-Outs:

- Majority loan sanctions happens at lower loan disbursement amount, primarily between 5K to 15K.
- Since maximum loan sanctions volume sits at small disbursement values, loan default cases are higher at these disbursement values.
- Loan sanction amount group of “20K to 25K” to be better scrutinized since loan default percentage for this group is +19% with total loan sanction/funded amount of 602 lakhs.
- Extra due-diligence need to be provided for Loan sanction amount group of “30K to 35K” since it has the highest loan default percentage at +24% though the loan disbursement volume is low (213 lakhs).

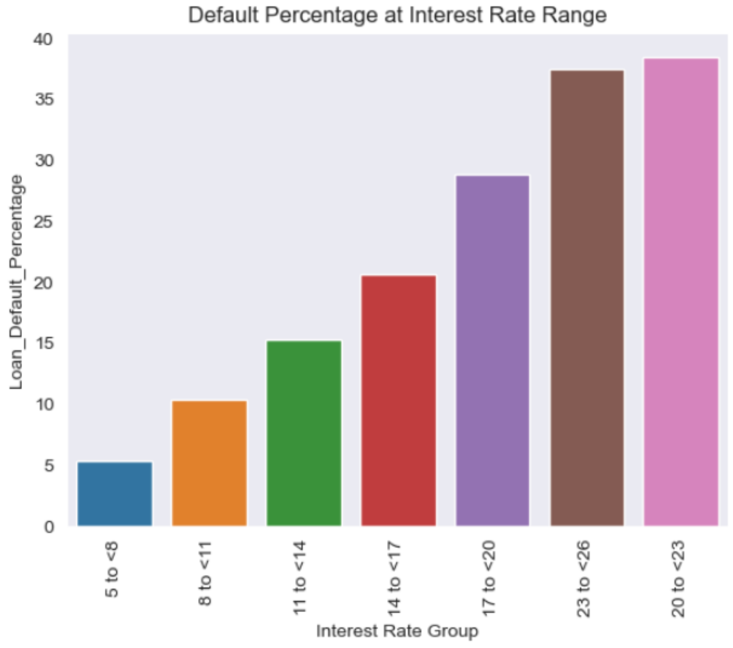
Loan Tenure

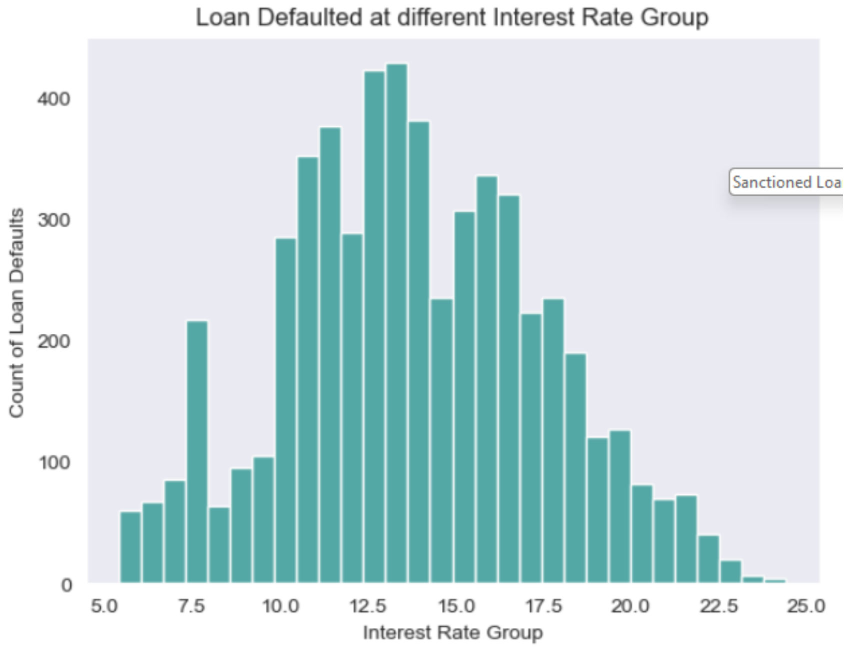


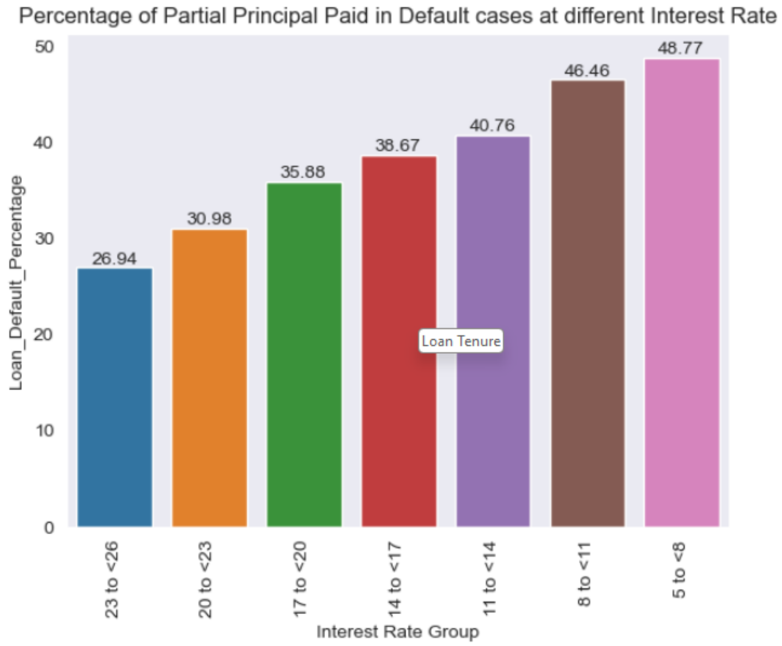
Observations & Key Call-Outs:

- Maximum loan sanction amount and loan default volume/count sits at loan tenure of 36 months.
- Higher loan default amount and Percentage of loan default based on total loan sanctioned at each term level is at term level of 60 months.
- Loan with higher tenure need extra scrutiny and could be limited to smaller term.

Loan Lending Interest Rate



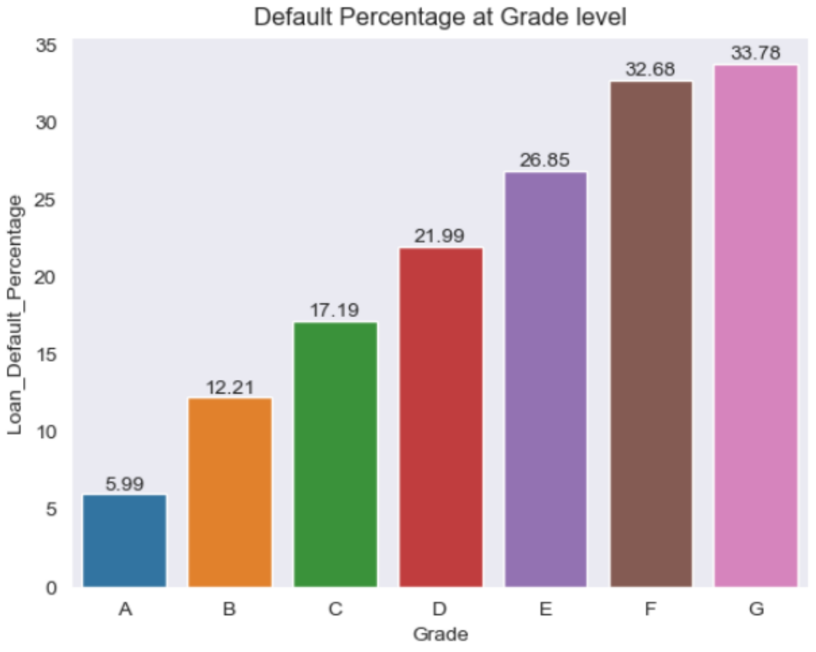


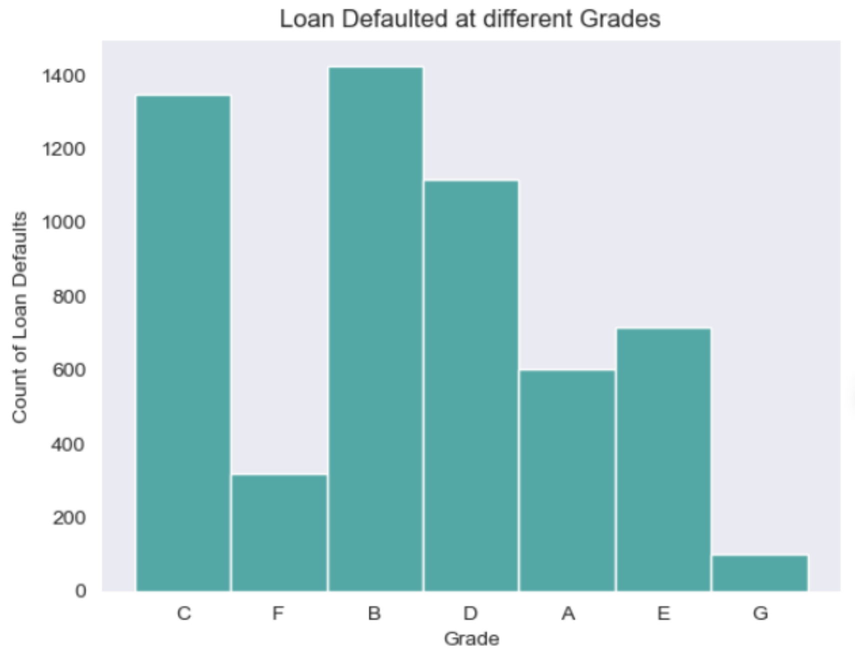


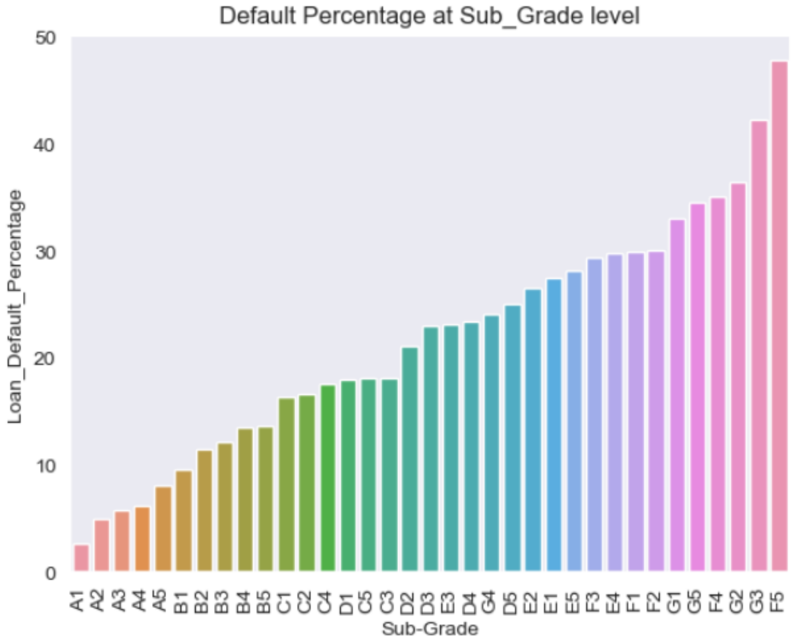
Observations & Key Call-Outs:

- Higher volume / count of loan default is between 12% to 14% interest rate .
- Loan default increases with increase in intertest rate and highest loan default is between "20% to 23%" interest rate.
- Partial recovery (principal + recoveries) is low at higher interest rate group which indicates earlier loan defaults at higher interest rate.
- Interest rate should be maintained at market standards / mean level. Issuing risky loans at higher interest rate would again make the loan more risky.

Grade & Sub-Grades



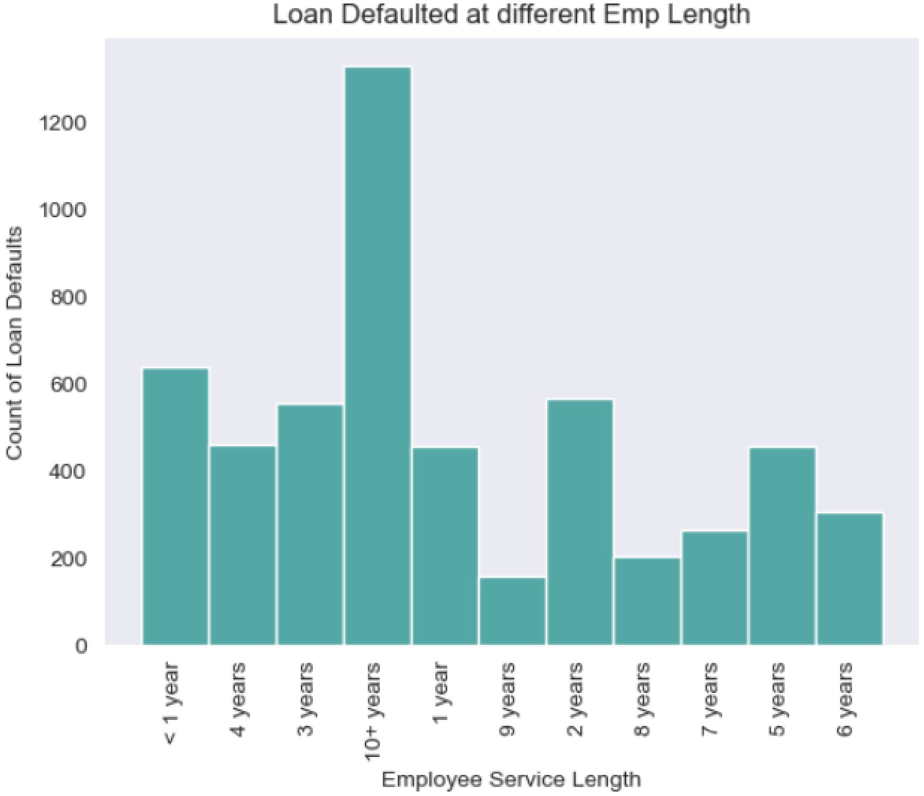


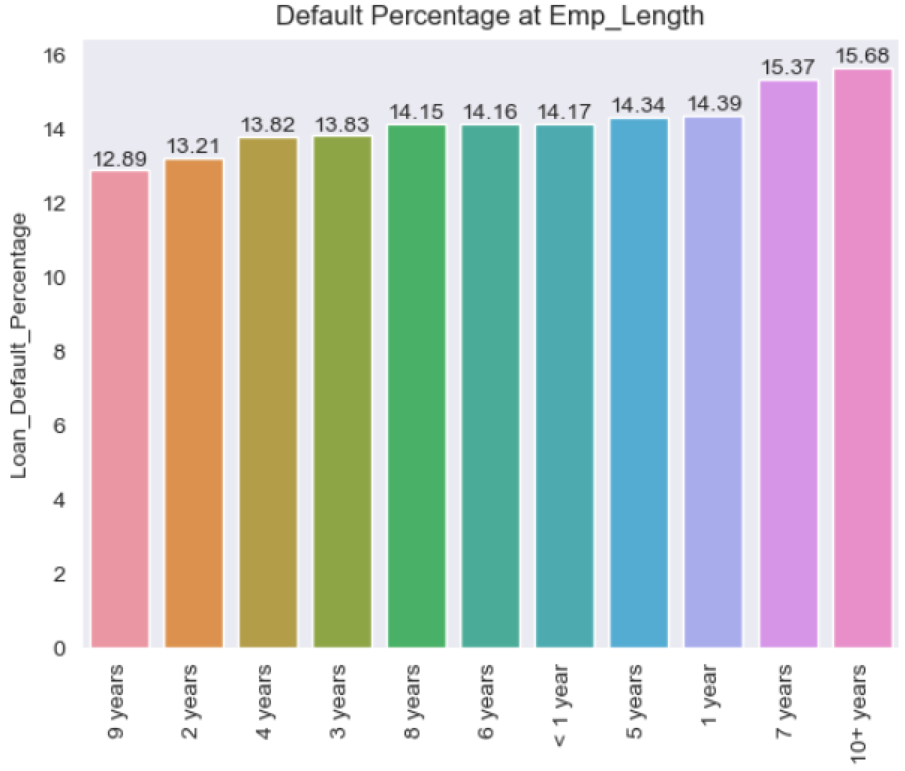


Observations & Key Call-Outs:

- Higher volume / count of loan default is at grades B, C and D.
- Loan default increases with decreasing grades with most loan defaults happening at grade E, F and G.
- Similarly, Loan default increases with decreasing sub-grades with most loan defaults happening at grade F and G
- Loan sanctions to lower grades should be minimized and more scrutinized.

Employee Experience (Years)

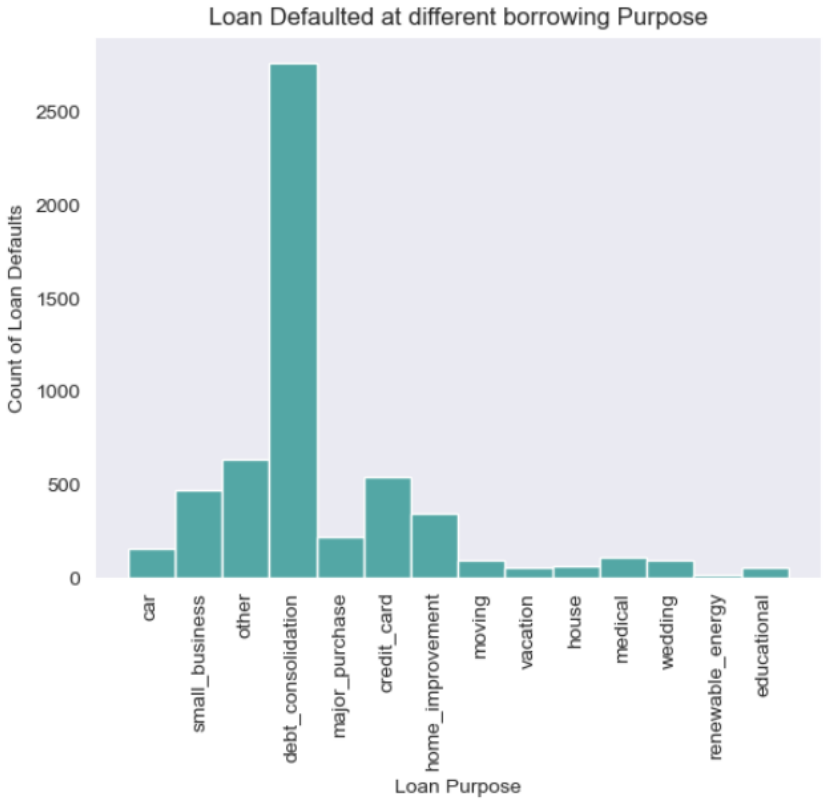


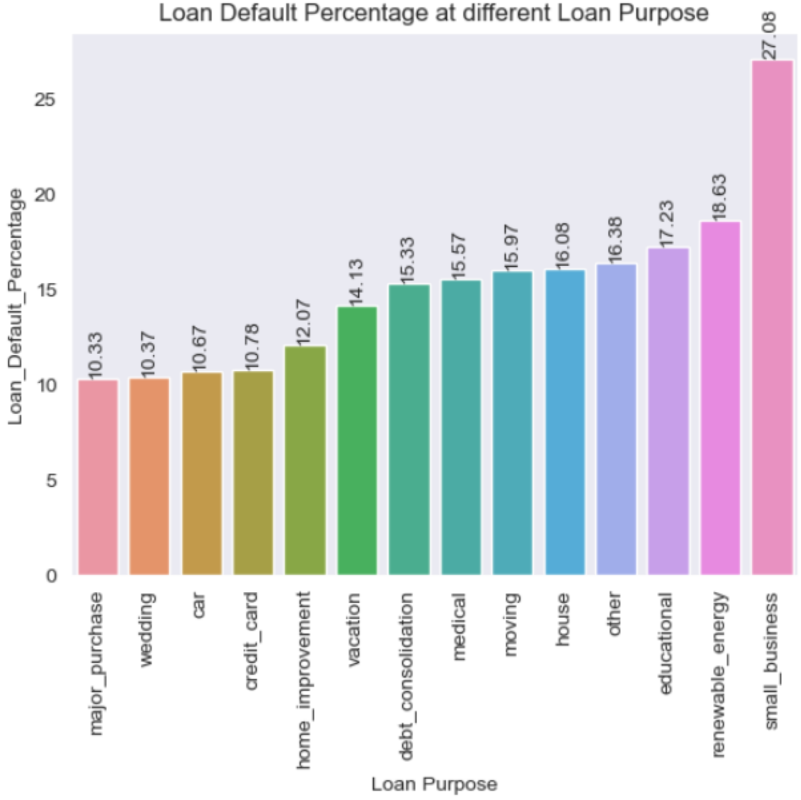


Observations & Key Call-Outs:

- Maximum loan default volume/count sits at employee service length of 10+ years.
- There isn't much variance in loan default at different employee service length and loan default percentage at each employee service length ranges between +12% to +15%.
- Loan disbursed to employee with service length of 7 and 10+ years need better scrutiny since loan default percentage at these length are +15%.

Purpose of Loan

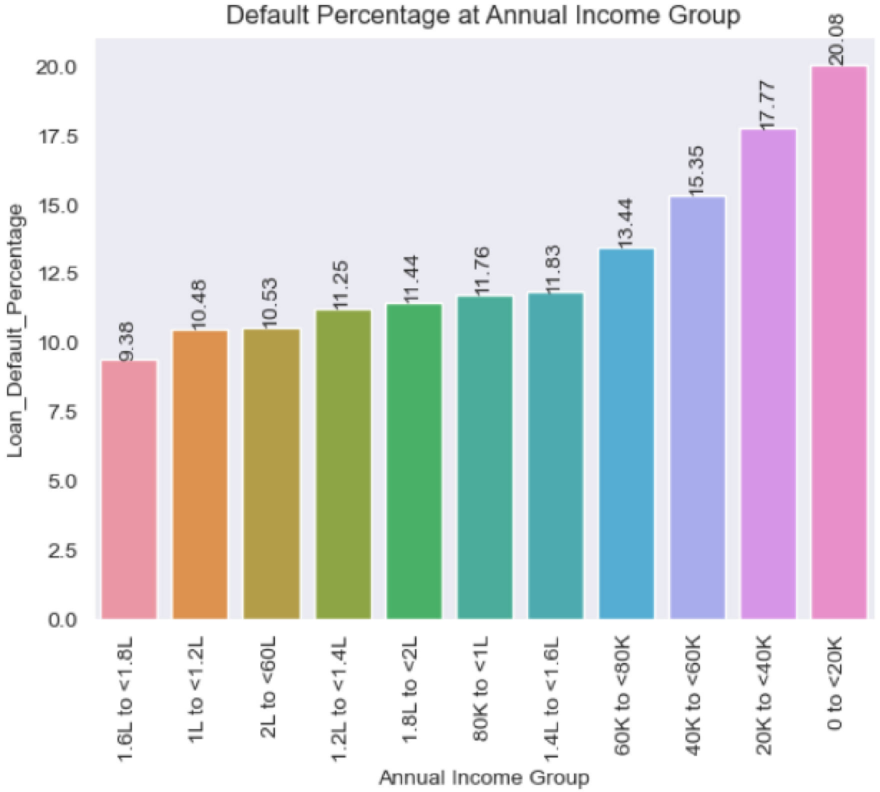


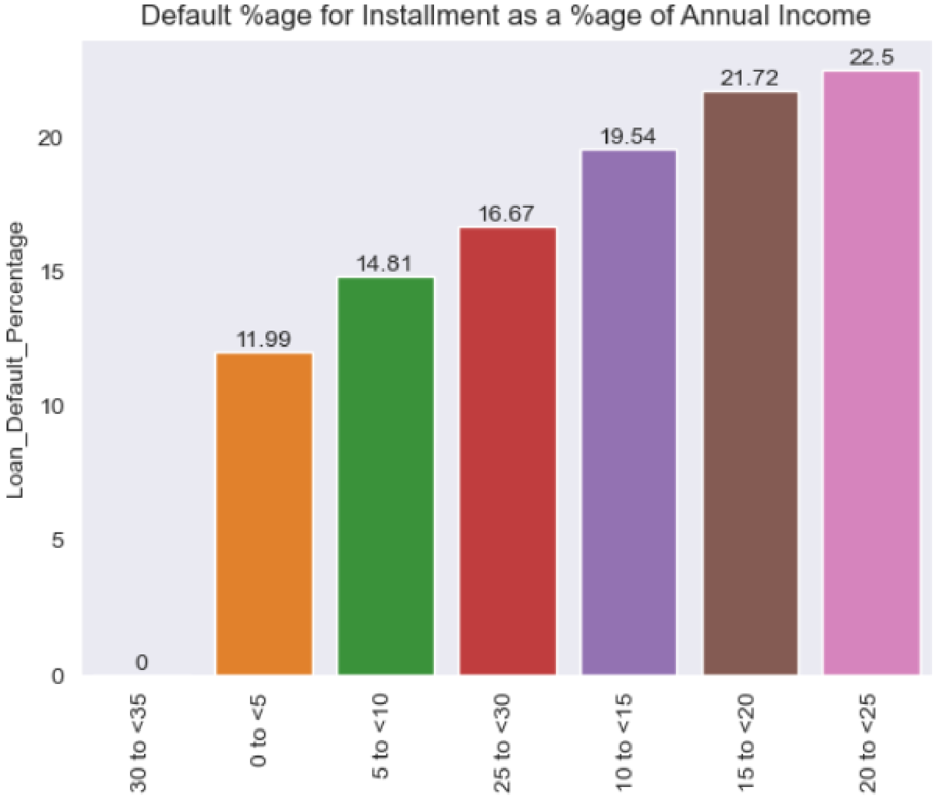


Observations & Key Call-Outs:

- Maximum loan disbursement and loan default volume is for loan purpose – 'Debt-Consolidation'.
- +50% of loan business is obtained from loan purpose Debt-Consolidation, followed by Credit-Card, Home-Improvement and Small-Business.
- Loan disbursed for the purpose of "Small-Business" need better scrutiny since it has the highest loan default percentage of +27%. Also, since major loan business volume comes from loan purpose of "Debt-Consolidation", loan for this purpose need to be better scrutinized.

Installments and Annual Income

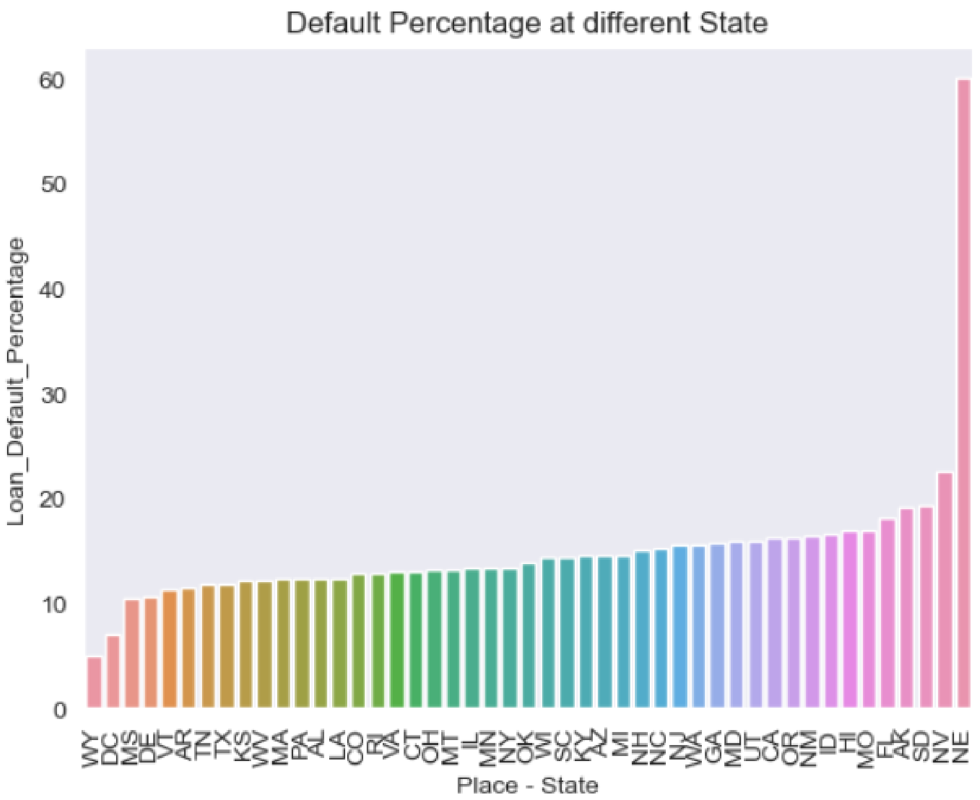


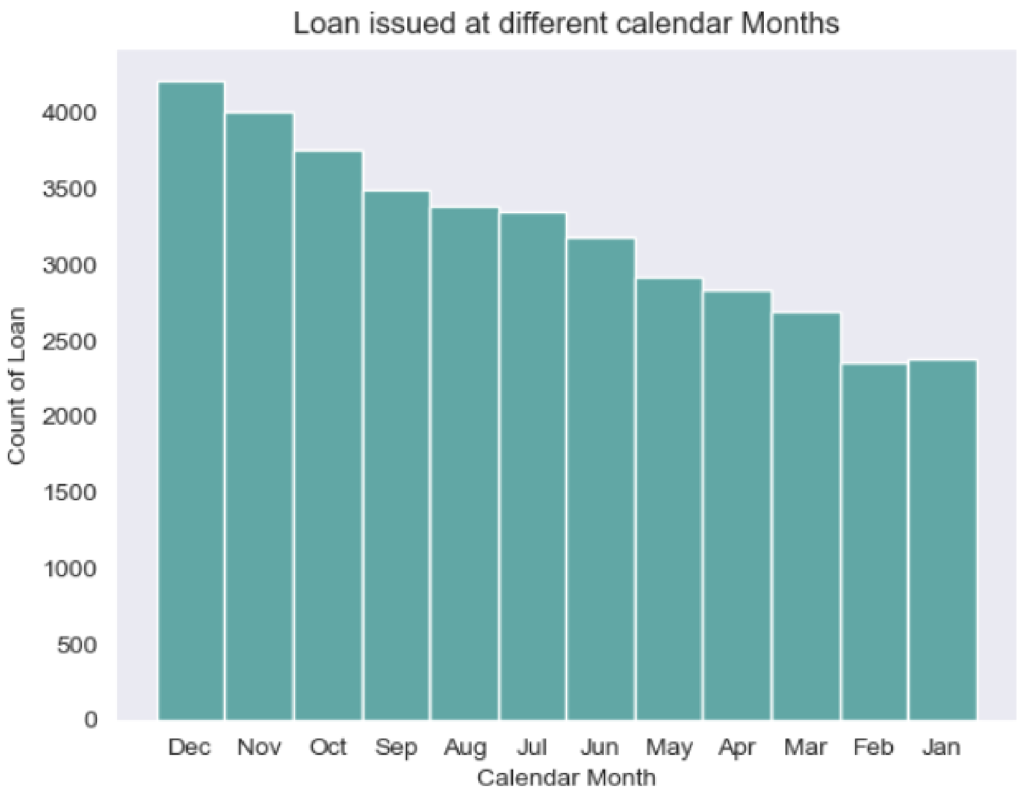


Observations & Key Call-Outs:

- Maximum loan default is at lower income group (up-to 80K)
- Lower monthly instalment as percentage of employee annual income results in lower loan default and it increases with increase in monthly instalment as percentage of employee annual income.
- Source of income for low annual income groups should be scrutinized to minimize loan defaults. Also, employees with instalment of 20% to 25% of annual income need better scrutiny since the loan default is highest at this level.

Others – Place (State) and Period





Observations & Key Call-Outs:

- Borrower living in the state of NV and NE defaults maximum. NV is more relevant since there are very few borrowers in state of NE with 60% of them have defaulted in loan payment.
- Maximum loan default happens in the quarter-end months (month 3) and year end (i.e. Dec). This could be because, sales team grant more number of loans to applicants at liberal terms to achieve their sales targets.