Lending CLub Case Study

Driving factors behind loan default by Saurabh Dubey

Problem Overview

Company

A Consumer finance company is into various types of loans to urban customers. The company is trying to reach a decision whether to give a loan to a certain individual or not, thus reducing two types of risks:

• If the person is likely to repay the loan, then not approving the loan will result in loss of business

• If the person is not likely to repay the loan, then approving the loan will result in financial loss

Context

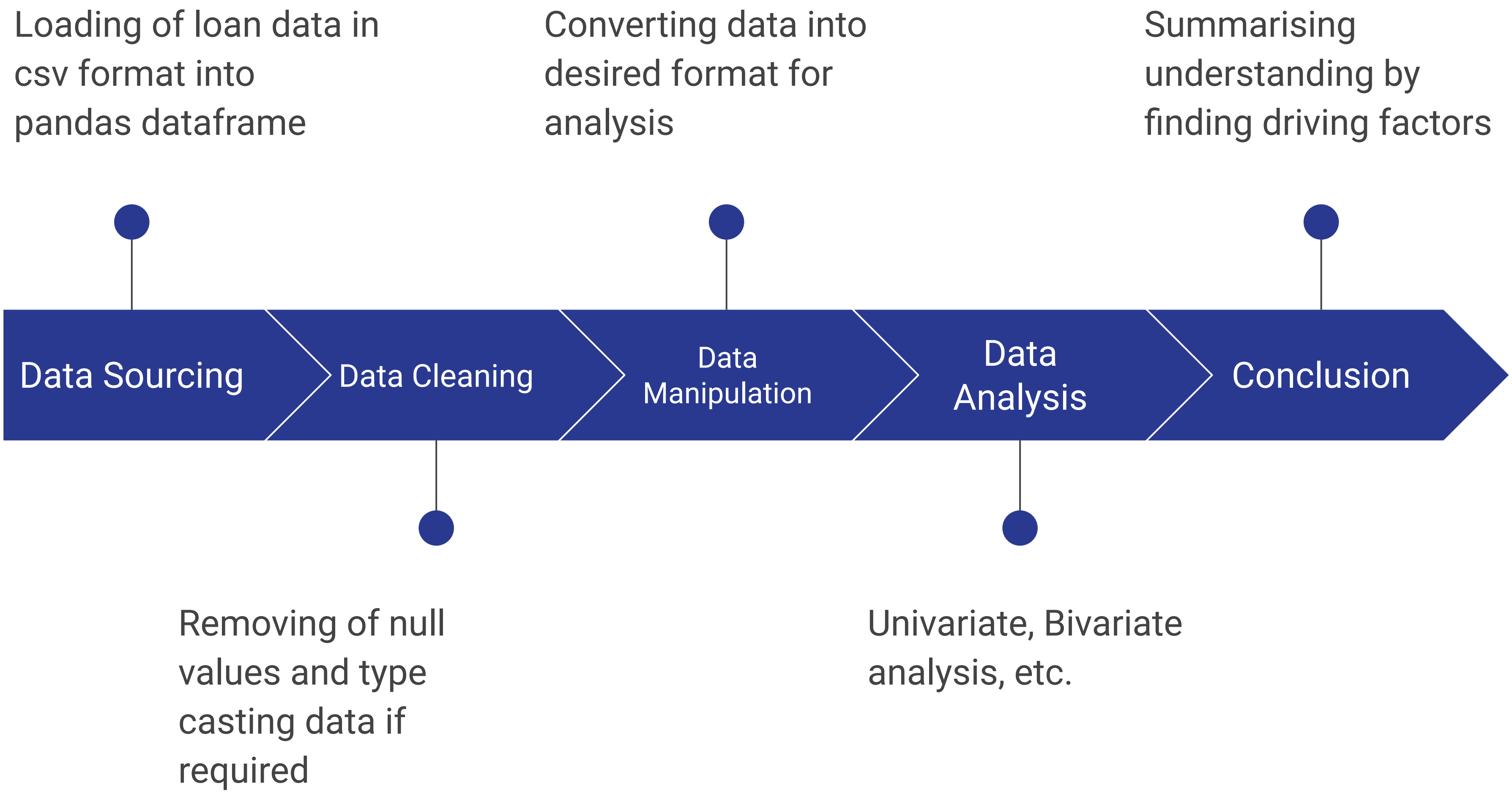
A company can either choose to accept or reject the loan application. While accepting the loan application, there are various scenarios.

- The customer has **fully paid** the loan
- There is a current loan **on going**
- **Charged off**: the customer has missed or delayed some payments for an existing loan

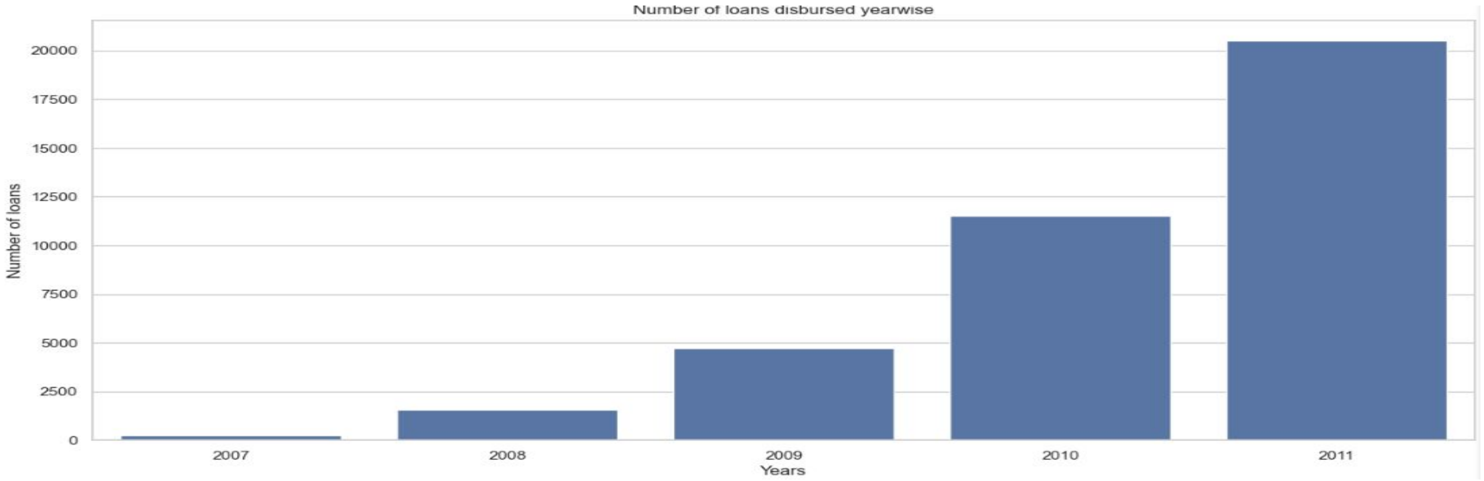
Problem statement

By performing Exploratory data analysis(EDA), we try to understand/find out which parameters about the customer such as age, monthly income, loan amount, etc. acting as an driving factors for defaulters.

Ultimately helping company decide customer should be given a loan.

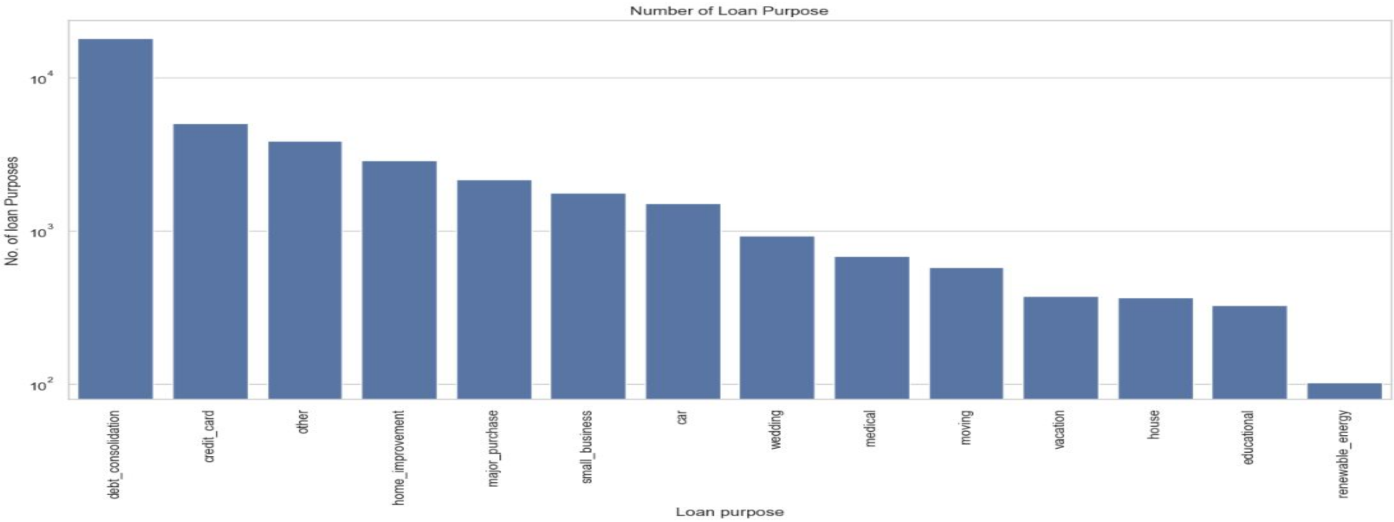

Loading of loan data in csv format into pandas dataframe
Converting data into desired format for analysis
Summarising understanding by finding driving factors
Data Sourcing
Data Cleaning
Data Manipulation
Data Analysis
Conclusion
Removing of null values and type casting data if required
Univariate, Bivariate analysis, etc.

Growth in Loan disbursed



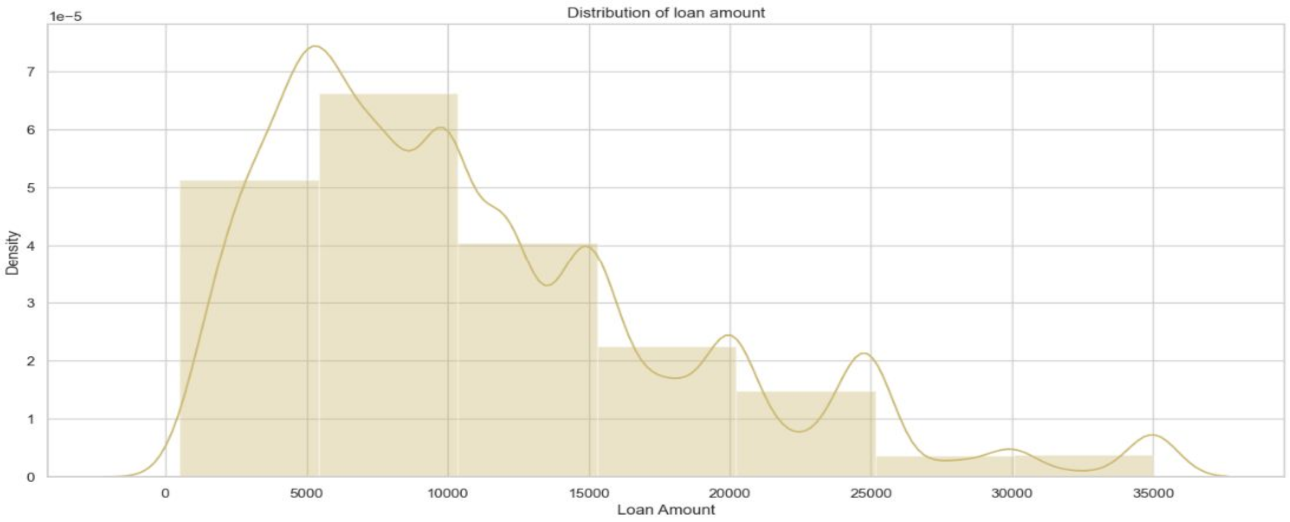
- There is substantial growth between 2007 - 2011
- Doubling almost every year approximately

Purpose of Loan



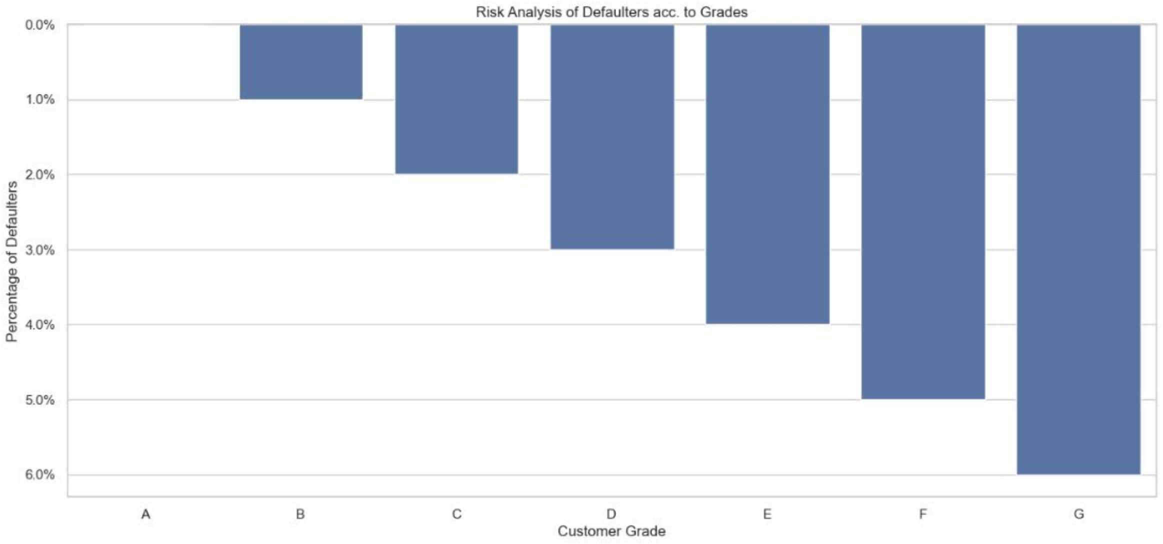
It is clearly visible from above plot that top 7 purpose of large number of applicants taking out loan is Debt consolidation , Credit card, Other , Home improvement, Major purchase, Small business & Car.

Loan amount distribution



The distribution of loan amount i.e the amount of loan applied can be visualized from the above plot . It has been observed that huge number of applicants apply loan in the range of 5-10k.

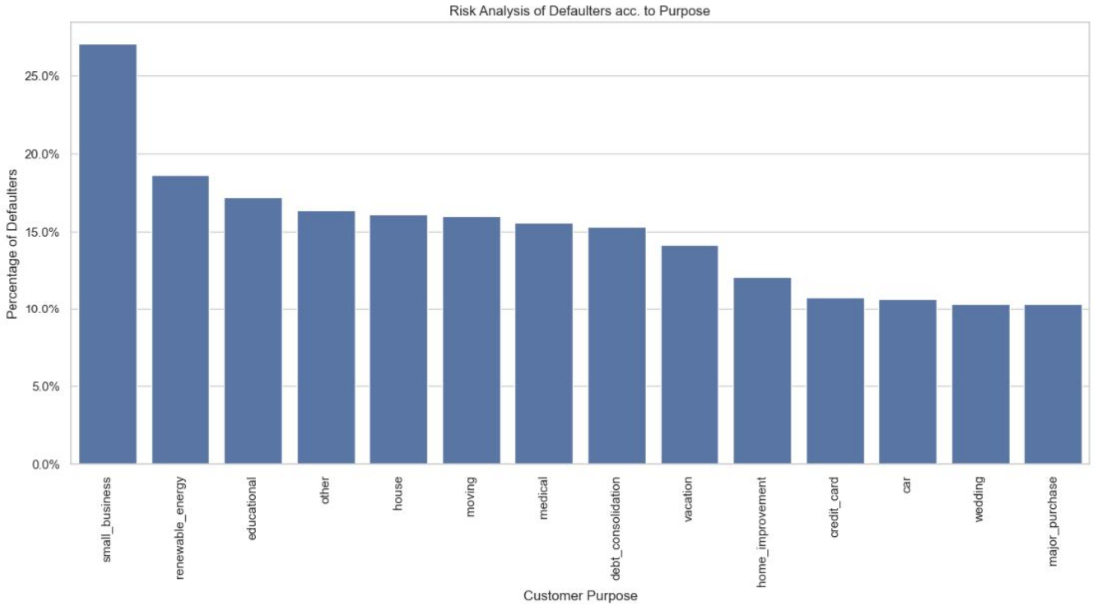
Percentage of loan defaulters according to grade



grade	
A	5.99
B	12.21
C	17.19
D	21.99
E	26.85
F	32.68
G	33.78

The percentage of Defaulters is the highest in Grade G group

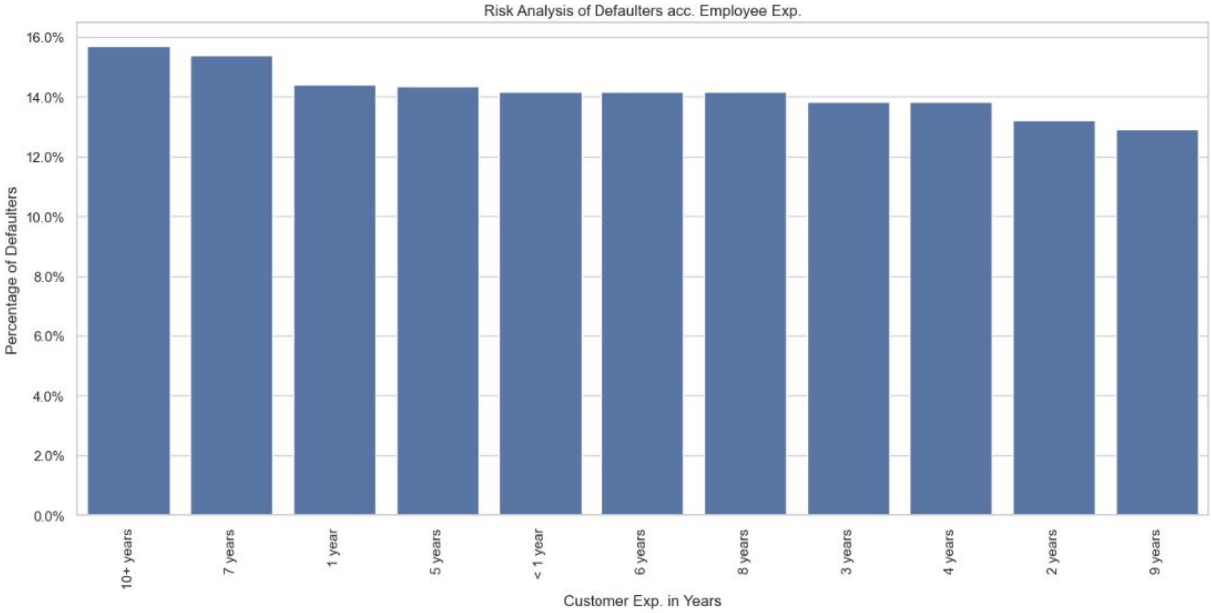
Percentage of loan defaulters according to loan purpose



purpose	
small_business	27.08
renewable_energy	18.63
educational	17.23
other	16.38
house	16.08
moving	15.97
medical	15.57
debt_consolidation	15.33
vacation	14.13
home_improvement	12.07
credit_card	10.78
car	10.67
wedding	10.37
major_purchase	10.33

The percentage of risk of Defaulters is the largest in the Loan Purpose – Small business followed by Renewable energy, Educationa’ and so on.

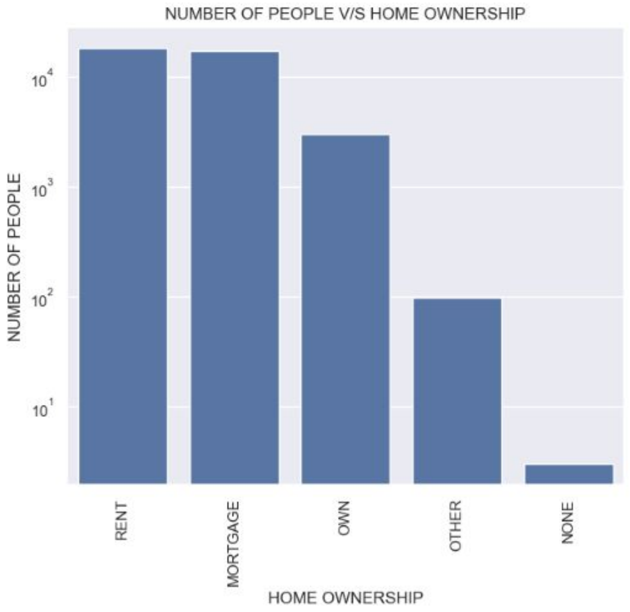
Percentage of loan defaulters according to Employee experience

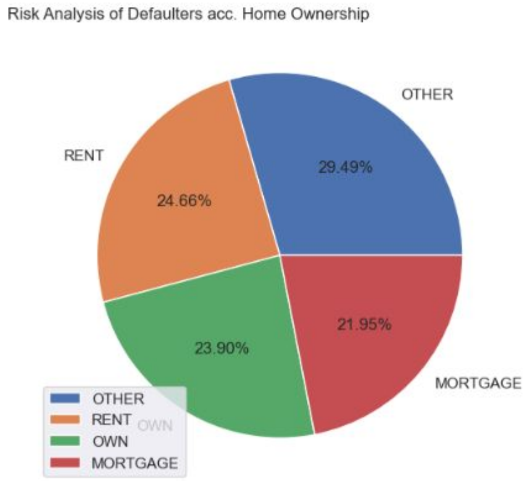


emp_length	
10+ years	15.68
7 years	15.37
1 year	14.39
5 years	14.34
< 1 year	14.17
6 years	14.16
8 years	14.15
3 years	13.83
4 years	13.82
2 years	13.21
9 years	12.89

Customers having 10+ years of experience have defaulted highest while least is 9 years of experience.

Percentage of loan defaulters according to home ownership

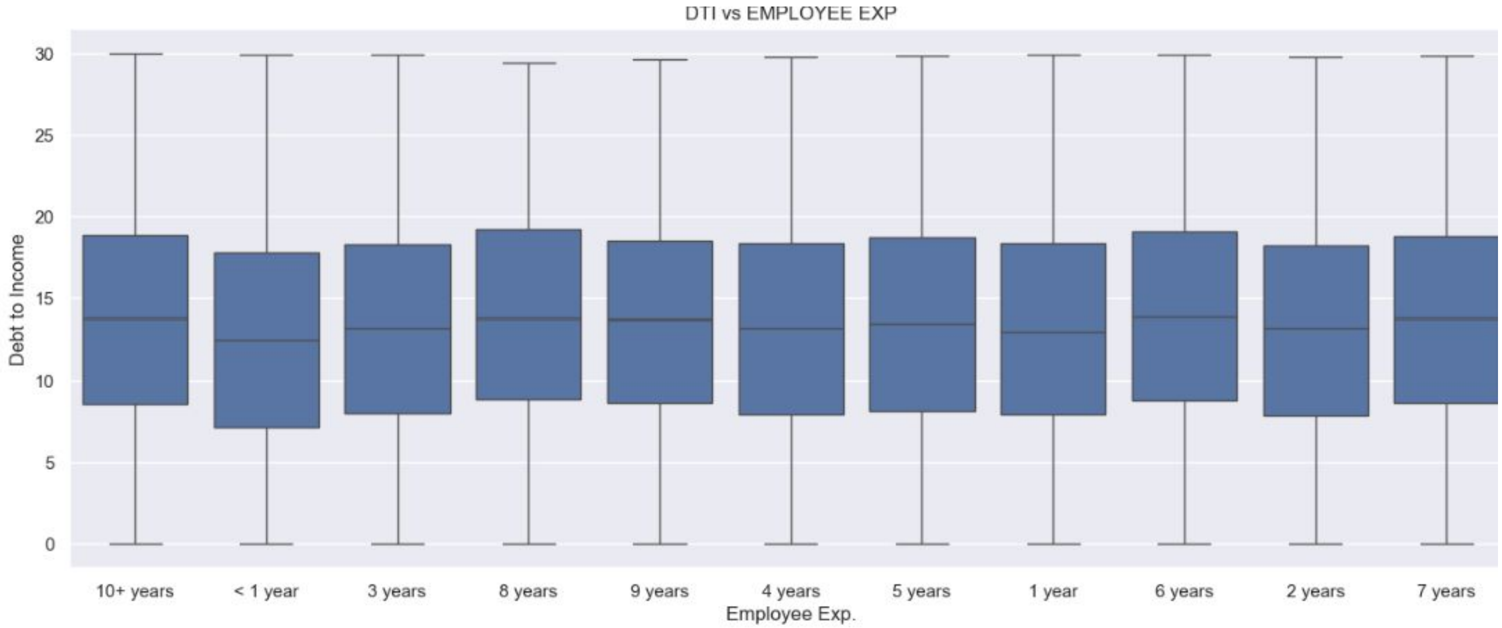




home_ownership	
OTHER	18.37
RENT	15.36
OWN	14.89
MORTGAGE	13.67

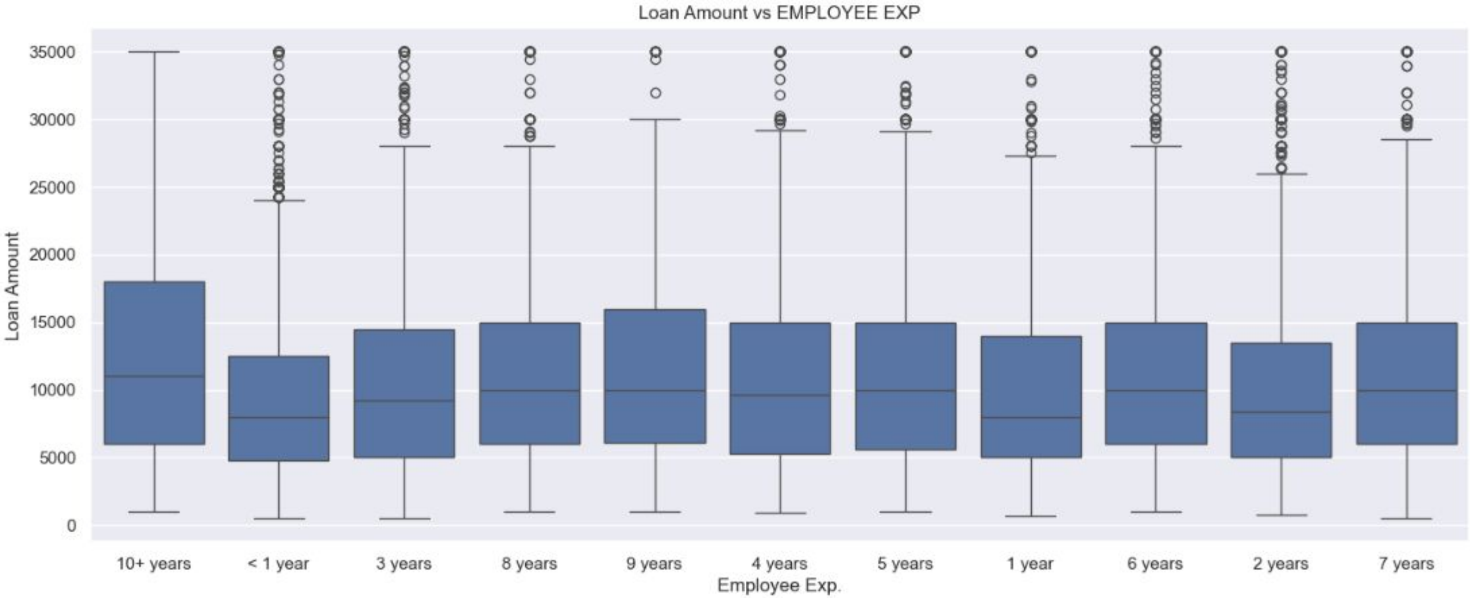
1. The number of people who have home ownership - “Rent” is highest(with very slim margin when compared to Mortgage)
2. The highest percentage of defaulter under home ownership is “Other” followed by “Rent” and “Own” and “Mortgage”

Debt to Income Vs Employee experience



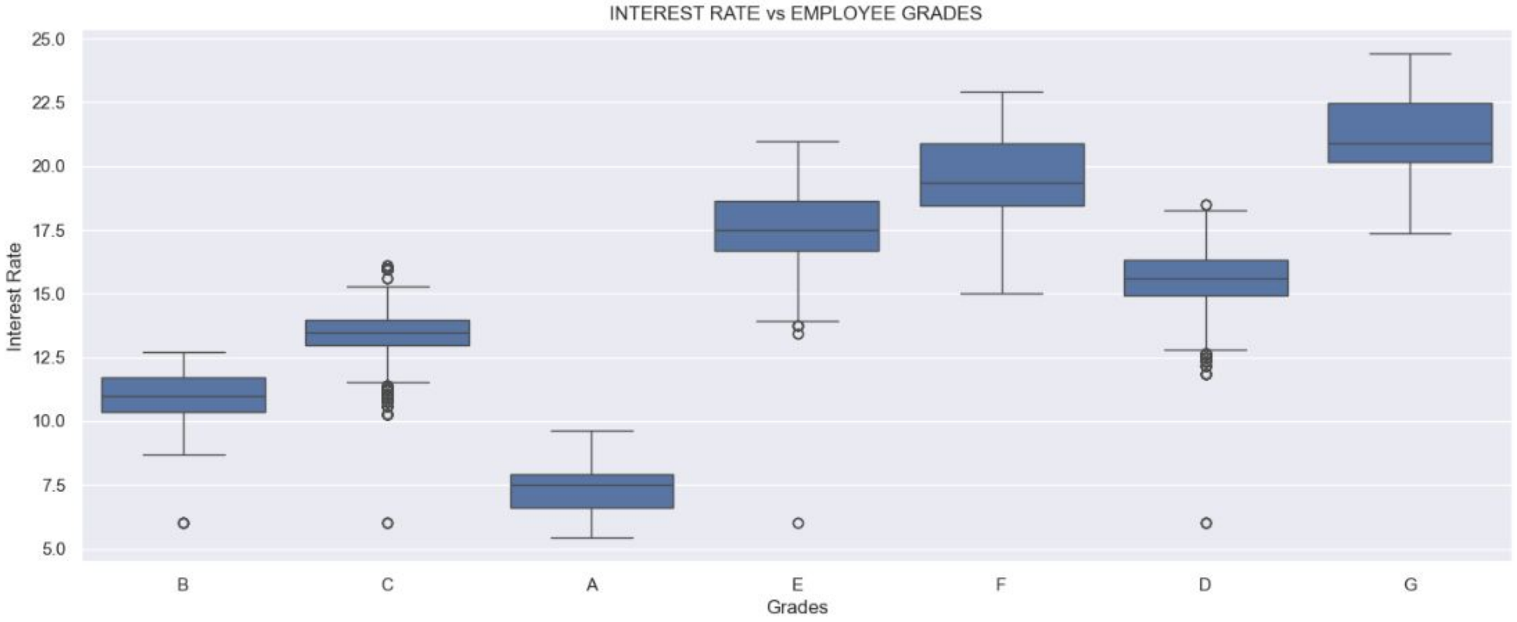
By doing bivariate analysis it can been observed that 10+ years of employee exp. is having debt to income in maximum range, for all others year of employee exp. are more or less in same range.

Loan Amount Vs Employee experience



By doing bivariate analysis it can been observed that employee with exp. 10+ years is taking more amount of loan and on the reverse side those who have less than 1 year is taking less amount of loan.

Interest rates Vs Employee grades



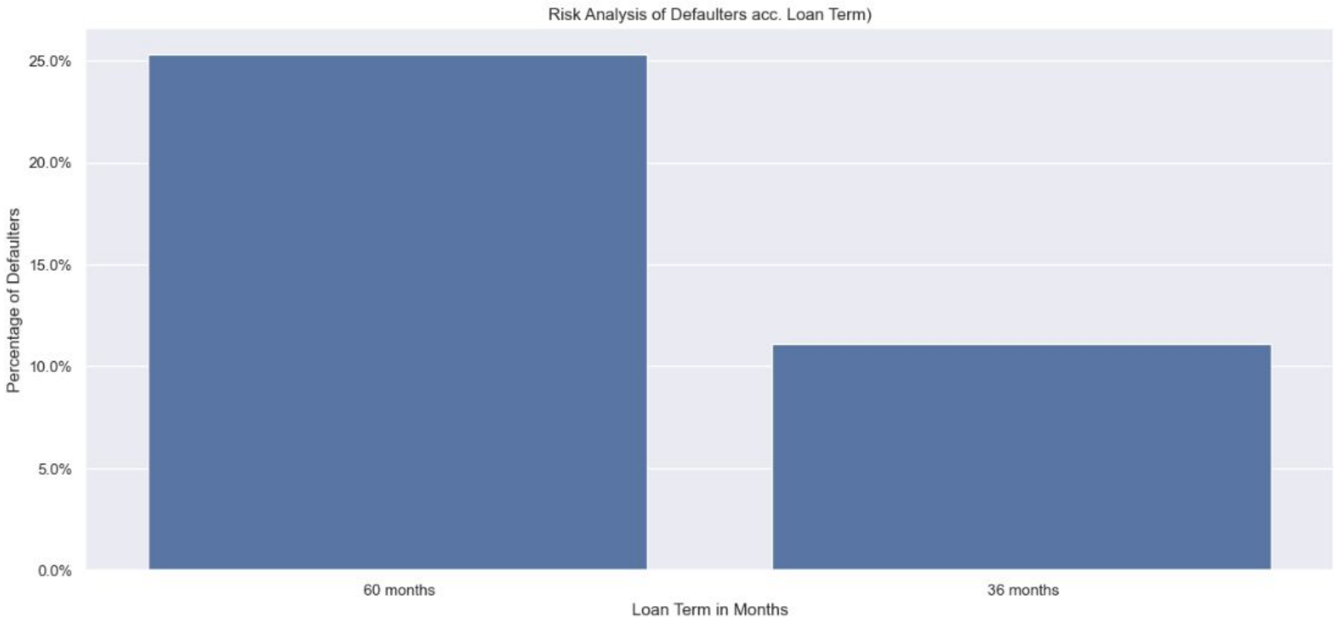
From the above plot it has been visualize that – “A” is having minimum interest rate, whereas “G” is having higher interest rate.

Percentage of loan defaulters according to address of a state

addr_state	
NE	60.00
NV	22.55
SD	19.35
AK	19.23
FL	18.12
MO	17.01
HI	16.87
ID	16.67
NM	16.39
OR	16.32

The percentage of Defaulters is maximum from NE (state) which is 60%

Percentage of loan defaulters according to term



	term
60 months	25.31
36 months	11.09

The percentage of Defaulters is found in 60 months as compared to 36 months – Loan Term

Conclusion

1. Debt Consolidation is the most common loan purpose, but Small Business Loans have the highest default risk, with Renewable Energy and Educational Loans also showing significant, though lower, risk. Understanding these varying risks can help in better managing defaults

2. Most loans are between 5000 and 10000 dollars. Borrowers with 10+ years of experience apply for larger loans and have higher debt-to-income ratios, while less experienced borrowers typically request smaller amounts. This suggests a correlation between borrower experience and loan size

3. Grade G borrowers have the highest default rates, while interest rates rise with loan grades from Grade A to Grade G. Lower-grade loans are riskier, so adjusting policies based on loan grades can improve risk management

4. Borrowers with 'Other' home ownership status have the highest default rates, followed by renters. Homeowners (Own and Mortgage) generally have lower default rates. Analyzing home ownership status helps assess borrower stability and default risk

5. 60-month loans have higher default rates than 36-month loans, with longer terms linked to increased risk. Shorter-term loans generally perform better, suggesting that re-evaluating loan term policies could help reduce defaults