

## Problem Statement:

You work for a **consumer finance company** that specialises in providing various types of loans to urban customers. When the company receives a loan application, it has to decide whether to approve or reject it based on the applicant's profile. Two **types of risks** are associated with the bank's decision:

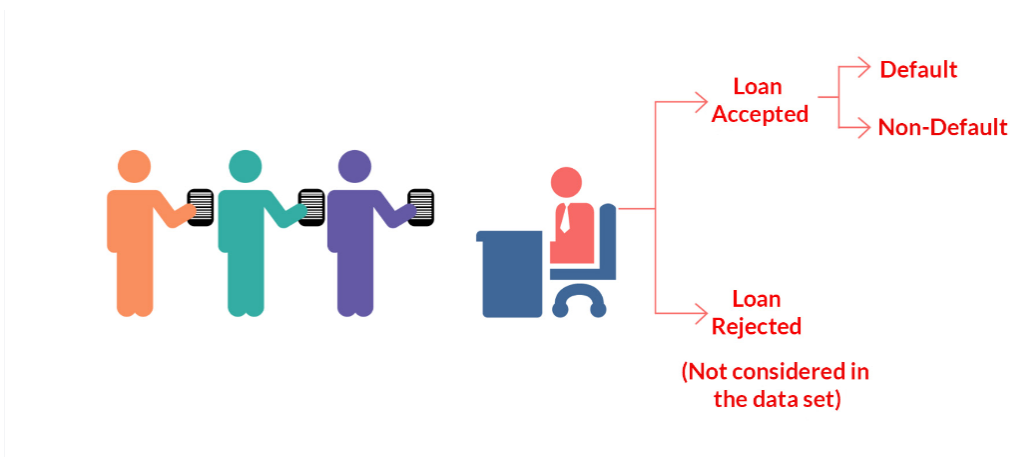
If the applicant is **likely to repay the loan**, then not approving the loan results in a **loss of business** to the company.

If the applicant is **not likely to repay the loan**, i.e., they are likely to default, then approving the loan may lead to **financial loss** for the company.

The data given below contains information about past loan applicants and whether they 'defaulted'. The aim is to identify patterns indicating that a person is likely to default, which may be used to deny the loan, reduce the loan amount, lend (to risky applicants) at a higher interest rate, etc.

In this case study, you will use EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of defaulting.

The following image depicts the decisions that could be undertaken by the firm.



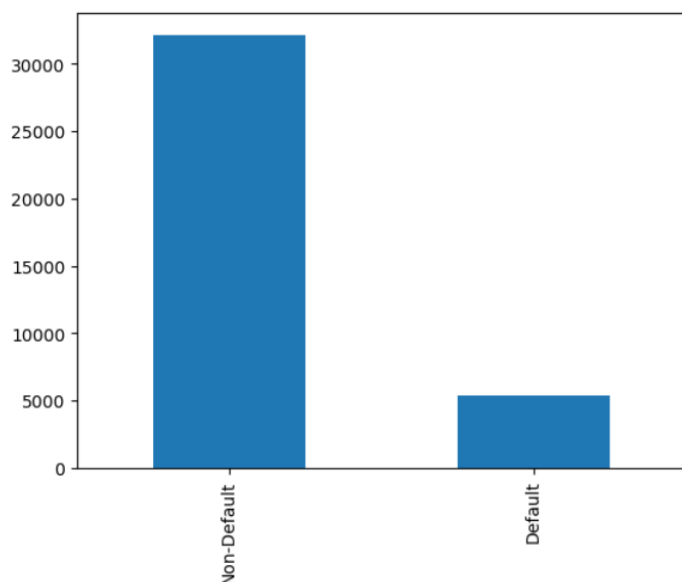
## Data Understanding :

Customer Data
1. Emp Length (How many years customer is working)
2. Annual Income
3. Employee Title

Loan Information
1. Loan Amount
2. Funded Amount
3. Term (number of months)
4. Interest Rate
5. Loan Status
6. Loan Grade

Customer Behaviour data
1. Earliest credit line
2. Recoveries
3. Application Type
4. Loan Purpose

## Calculating the overall default rate:



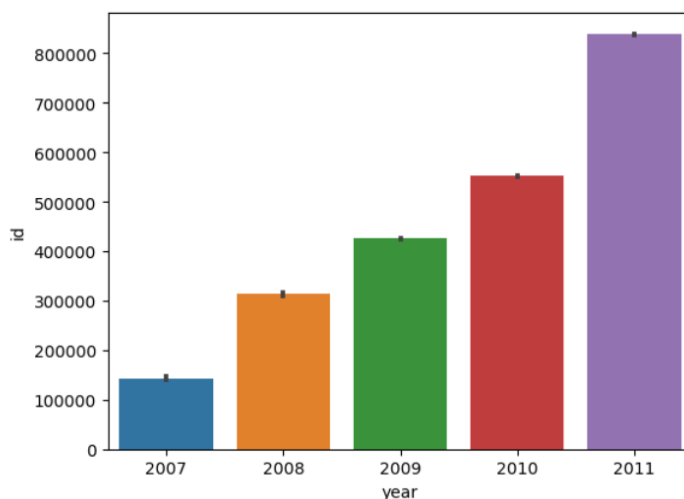
It can be clearly seen that there are less % of default rate in the entire dataset.

The default rate is 14%.

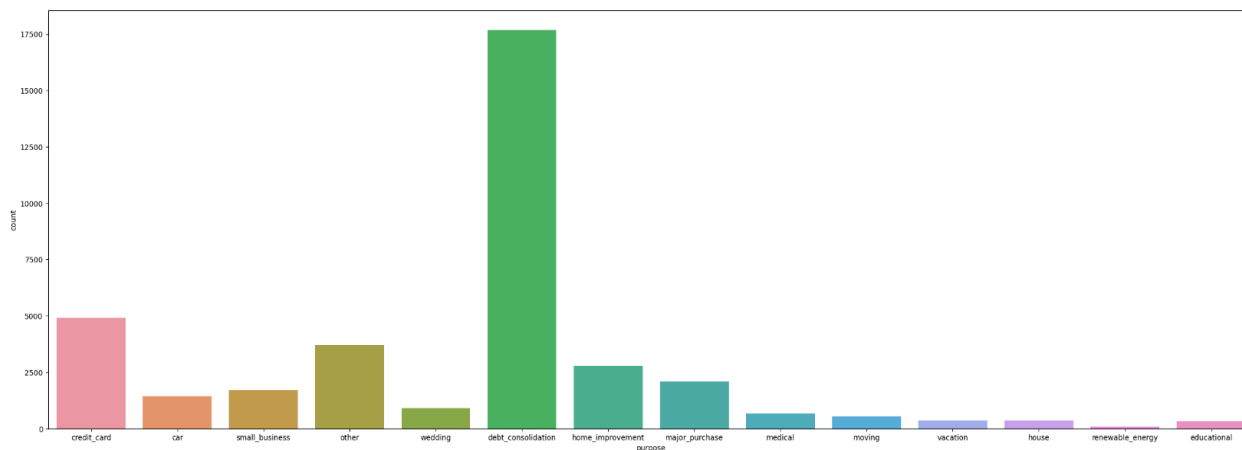
## Time Frame of Data:

Increase in number of applicants by year.

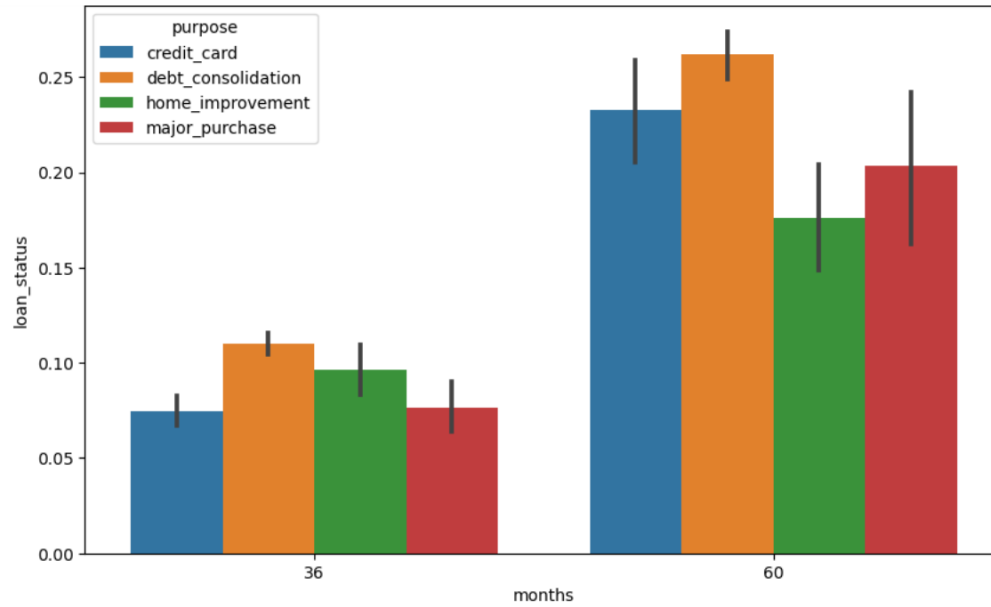
2011 has highest number of applicants for loan.



## Loan Purpose Distribution : Debt Consolidation Loan is the most popular



**Default rates with respect to purpose of loan and term period – Both 36 months and 60 months has debt consolidation as highest default rate.**



**Default rates with respect to grade – As Grade increases default rate increase A1 with low default rate A2 and goes on .**

