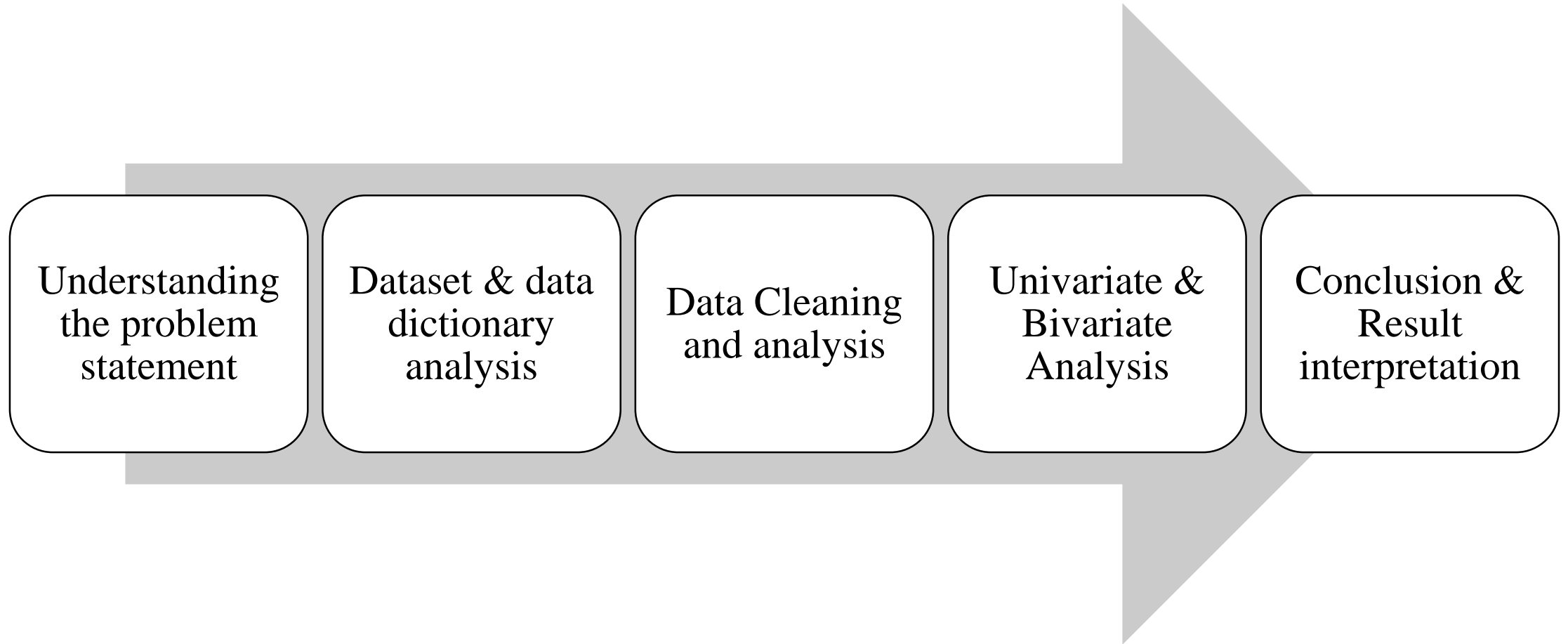# Lending Club Case Study

Group Facilitator – Pratik Advani

Member – Neeraj Keswani

Objectives

- To understand the driving parameters/factors behind loan default.

- To identify patterns which indicate if a person is likely to default.

- To help the company save credit loss.

Approach:

| Understanding the problem statement | Dataset & data dictionary analysis | Data Cleaning and analysis | Univariate & Bivariate Analysis | Conclusion & Result interpretation |

Data Cleaning and Analysis

- Read the dataset and remove the columns that will not help us in achieving the objective.
- Calculate the null values of each column & drop the columns that have all the null values.
- Identify unique values in each column & remove those columns that have only 1 unique value.
- Extract numbers for emp_length column & convert the datatype to numeric.
- NA values in emp_length column are assigned with mode value of the column.
- Derive new columns "issue_d_month", "issue_d_year" from "issue_d" column.
- Removed the rows where loan status is "current" because this status won't help us in providing any insights as we cannot assume if the current loan is going to be defaulted or not.
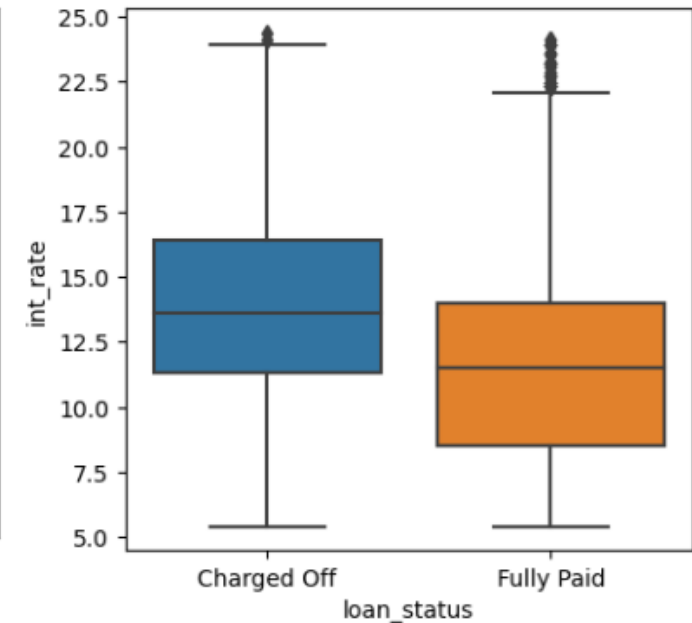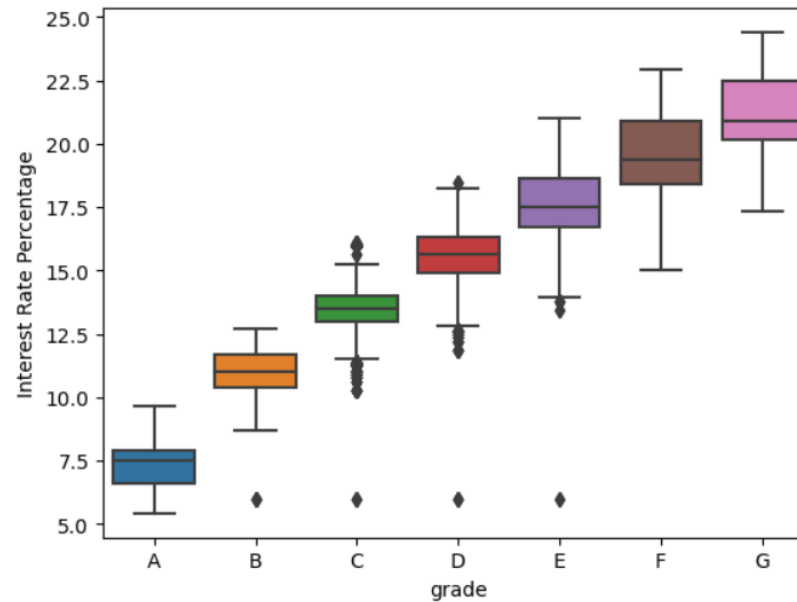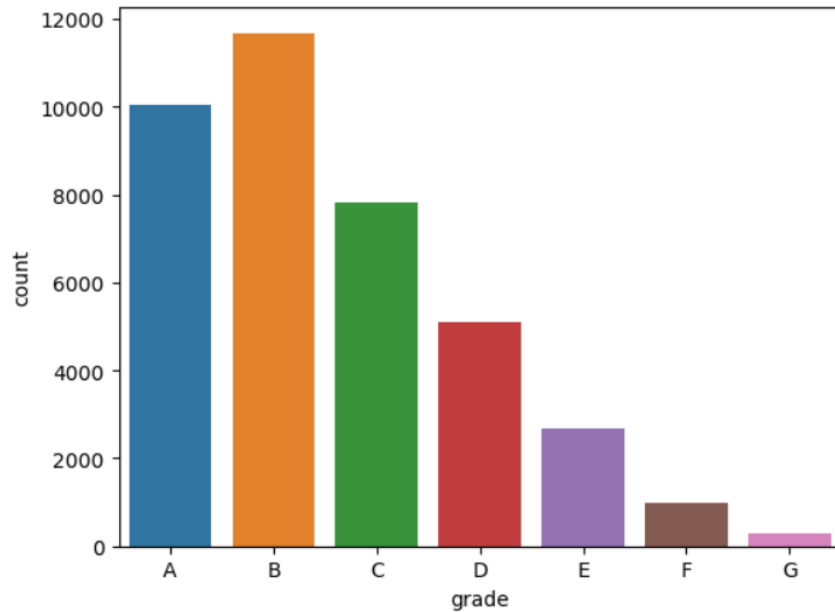- Divide the cleansed data into Categorical or numerical columns.

Loan Status Analysis
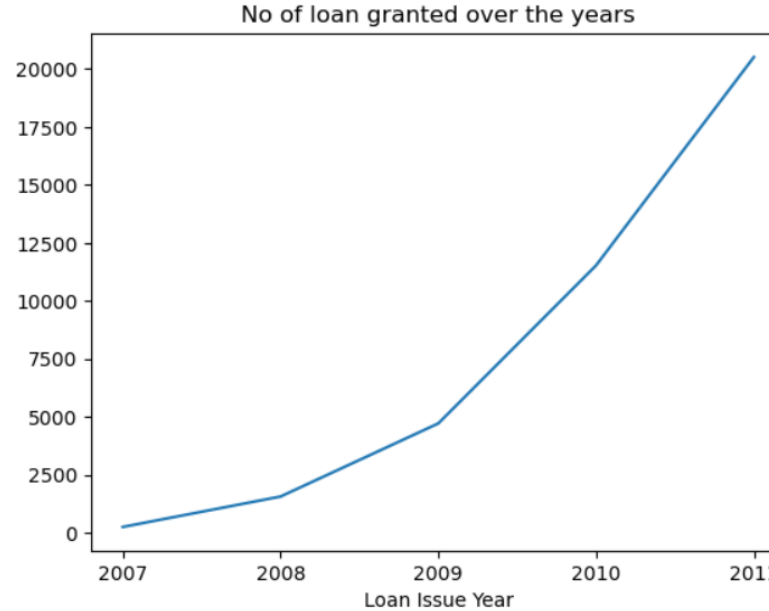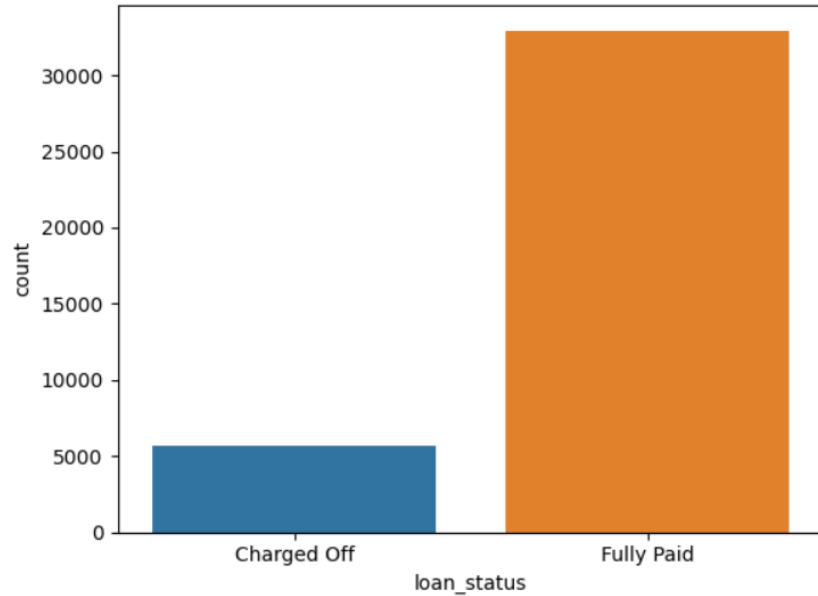
- Observations:
  1. Most of the loans are fully paid
  2. About 14% of loans have been defaulted
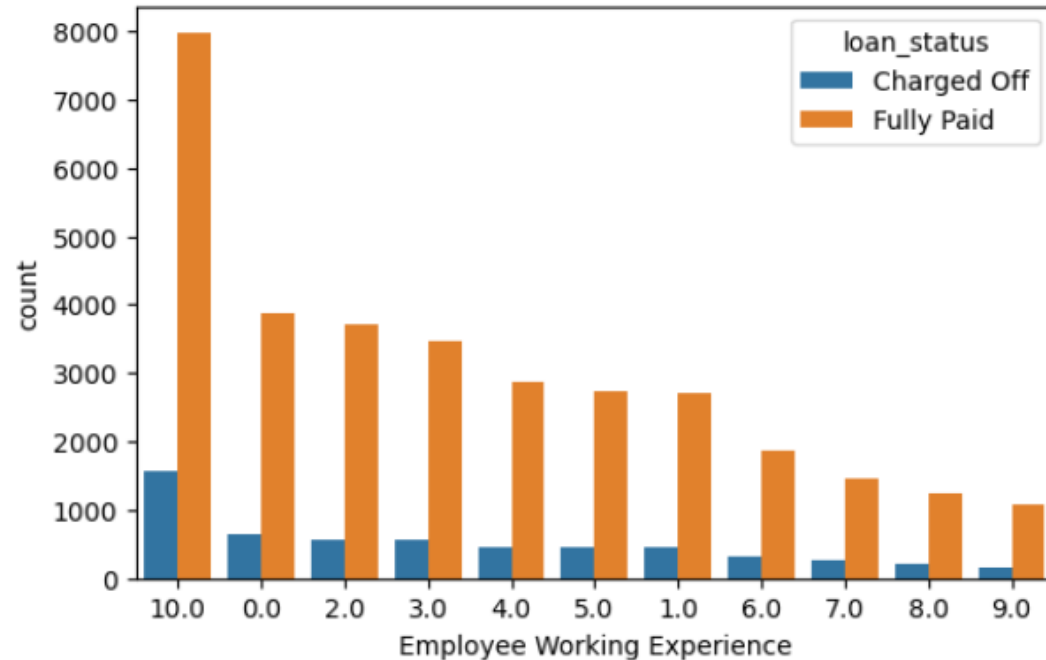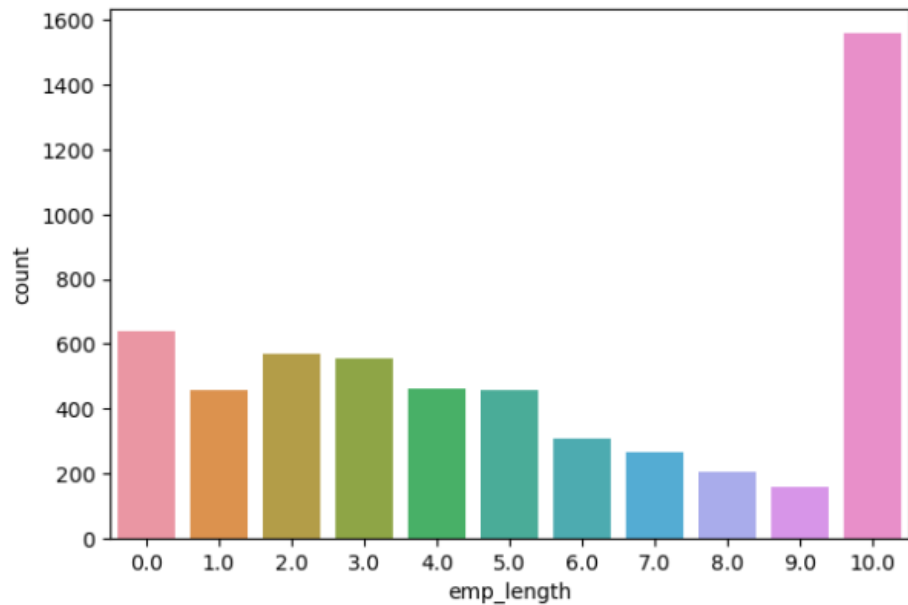 3. Most of the loans were granted in the year 2008 to 2011

Loan Grade Analysis

- Observations:
    1. Most of the loans have grade A&B, thereby we can infer most of the loans are highly graded loans.
    2. High Quality Loans have low interest rates.
    3. Which depicts, loans with higher interest rates have a higher tendency to default

# Loan Applicants Work Experience Analysis

- Observations:
  1. Most of the people who have applied for the loan have 10/10+ years of employment tenure
  2. Tendency to default is also the highest among the people with the 10/10+ years of Experience. Lending Club can do a thorough check of such people before granting loans

Loan Defaulters by State Analysis

- Observations: Applicants from CA - California have highest tendency to default followed by FL –Florida, and NY – New York.(Assuming the data is of US region).
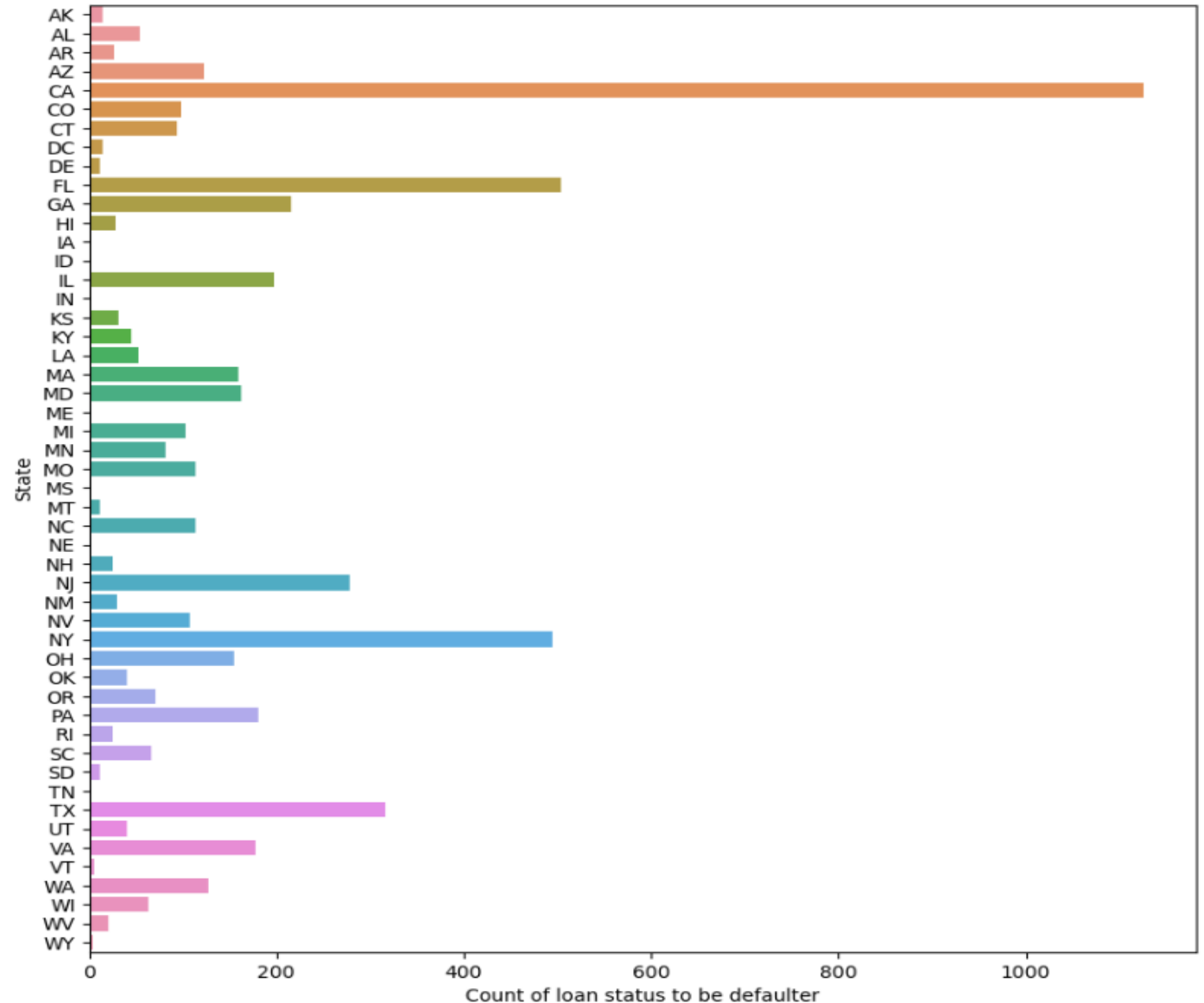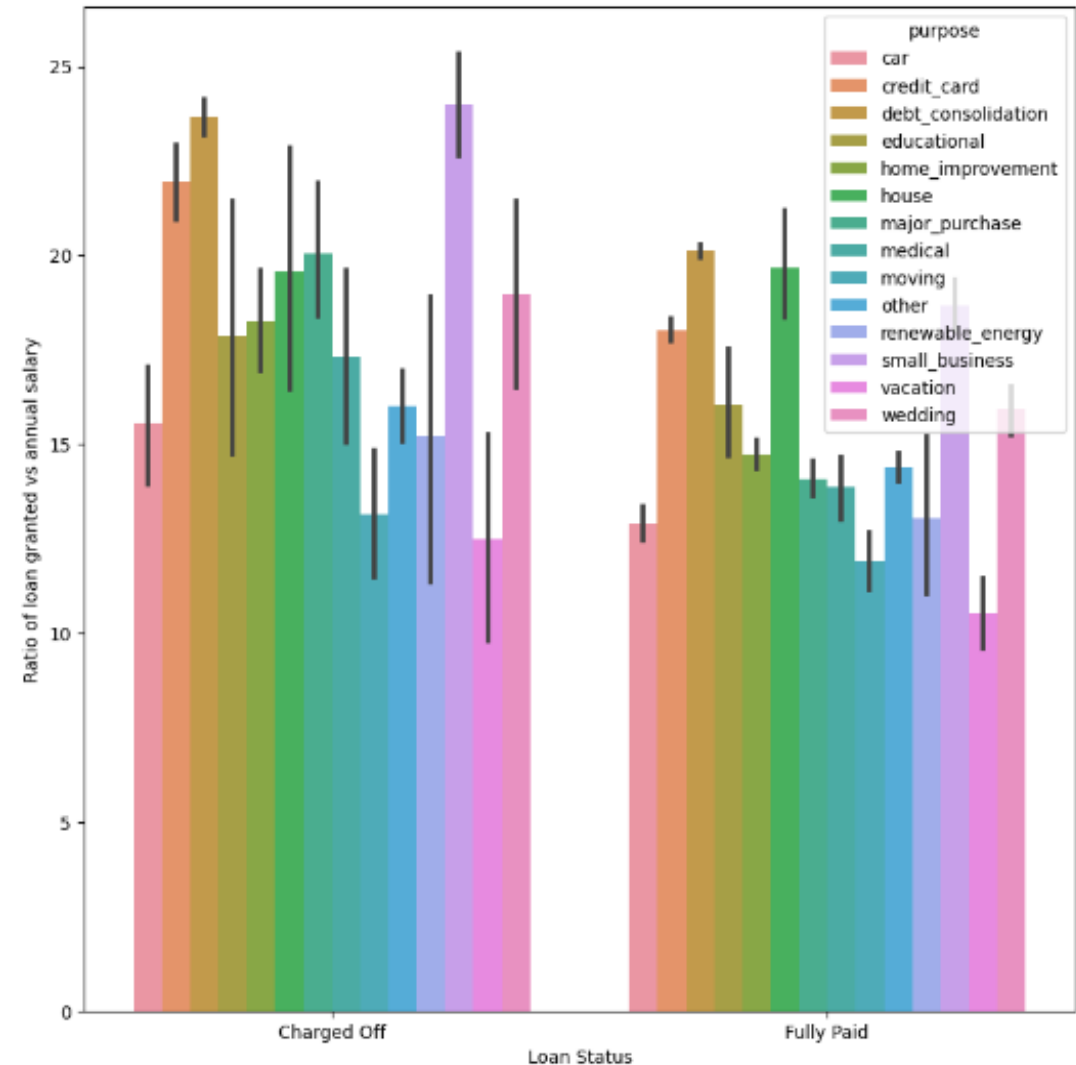
Purpose for Applying Loans Analysis
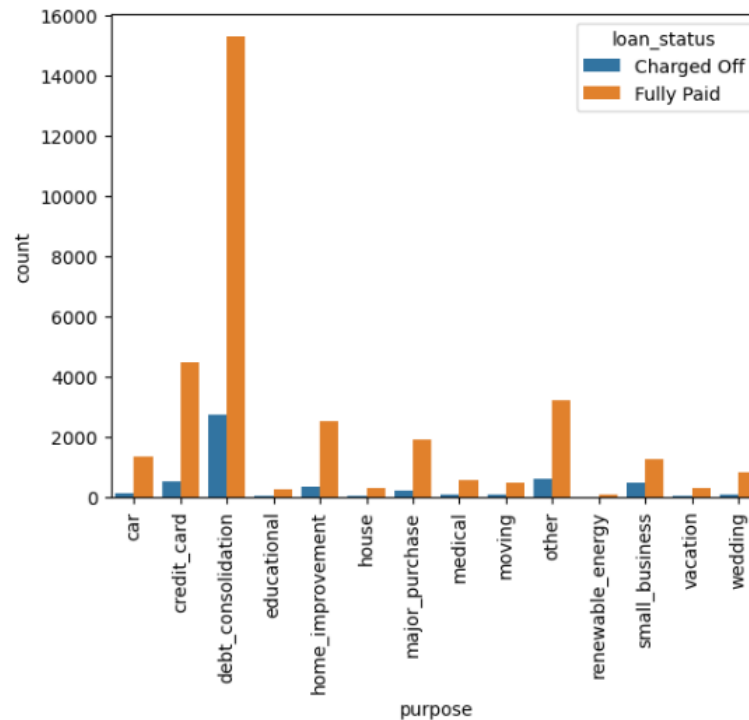
- Observations:
  1. Most of loans borrowed were for the purpose of debt_consolidation and at the same time most defaulters also have taken the loan for same purpose.
  2. The tendency to default has been seen more in the borrowers who have more annual income.

Loan Status vs Loan Tenure

- Observations:
    1. Most of the loans are borrowed for 36 Months
    2. Most of the borrowers who have taken a loan of 36 months have a higher tendency to default
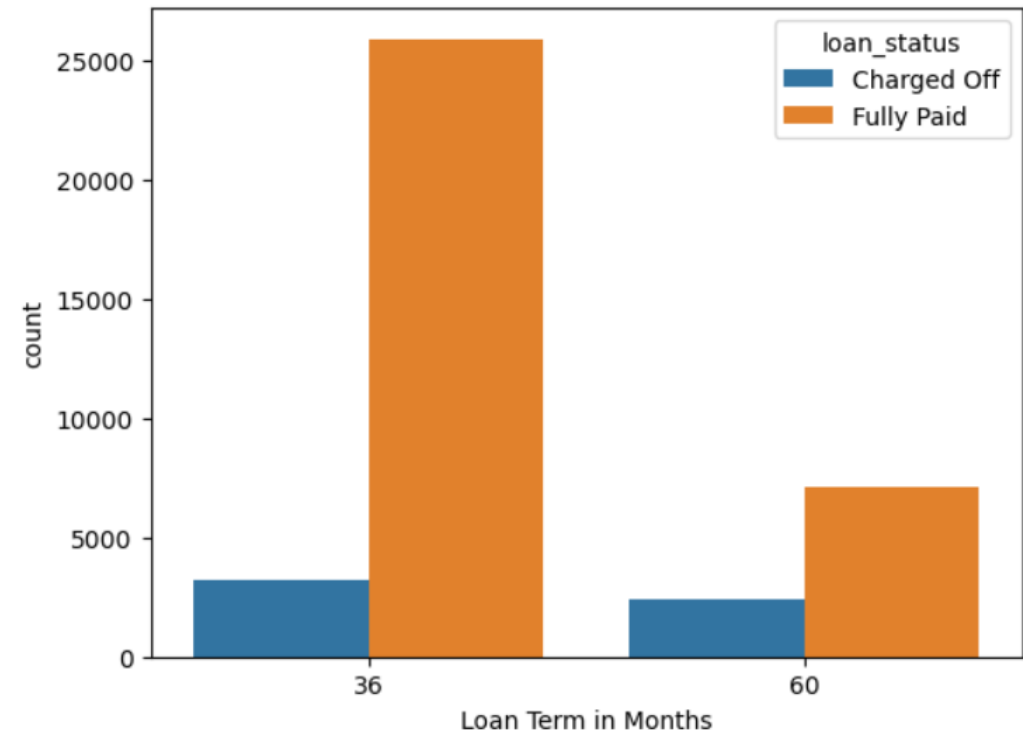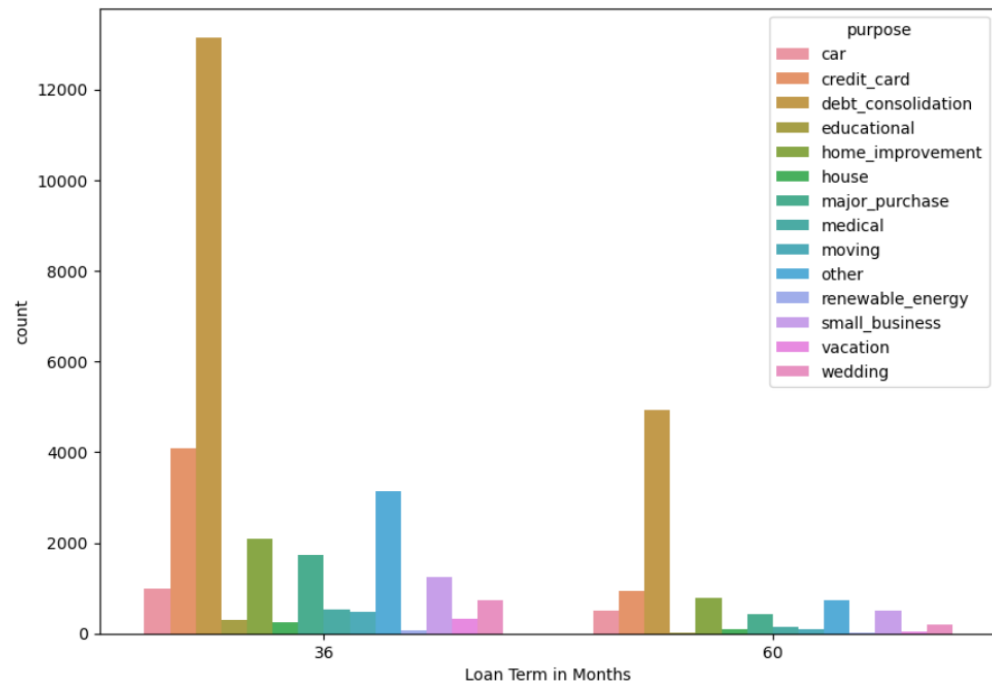
Home Ownership vs Loan Status

- Observations:
    1. Most of the loan borrowers are living on rent or on mortgage homes.
    2. Borrowers who are living on mortgage or rented homes have a very high tendency to default than who are owning their homes. So Company should thorough checks for such people before granting loans.
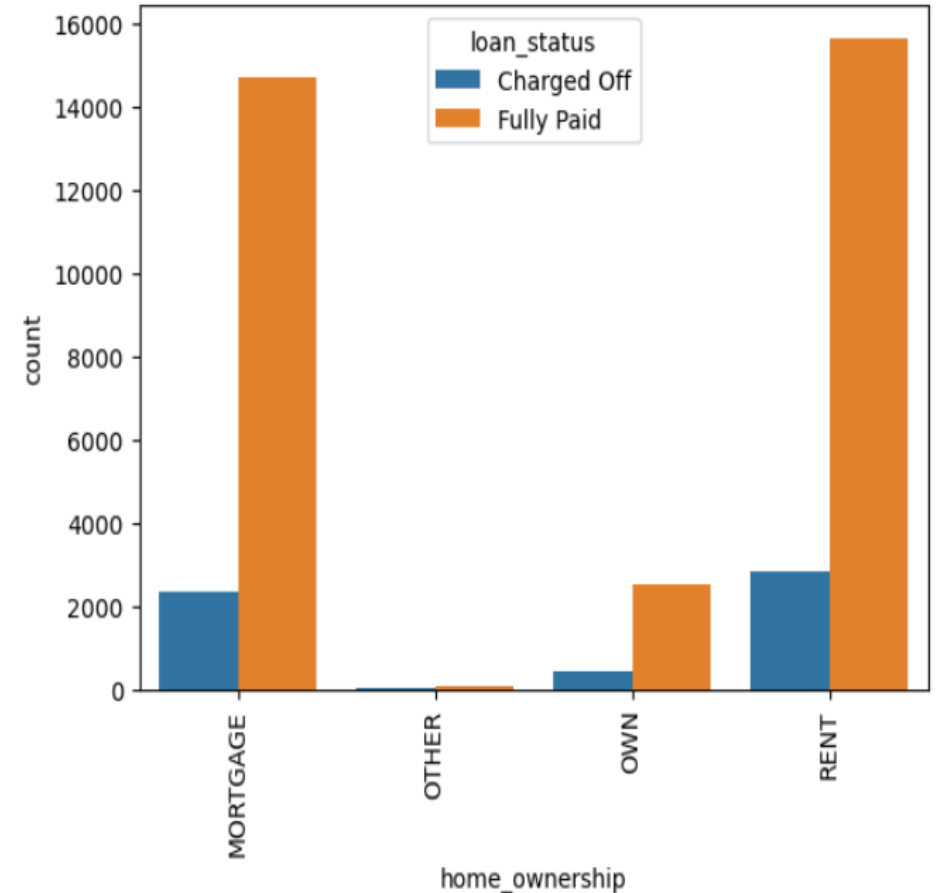
Loan issue time vs Loan Status
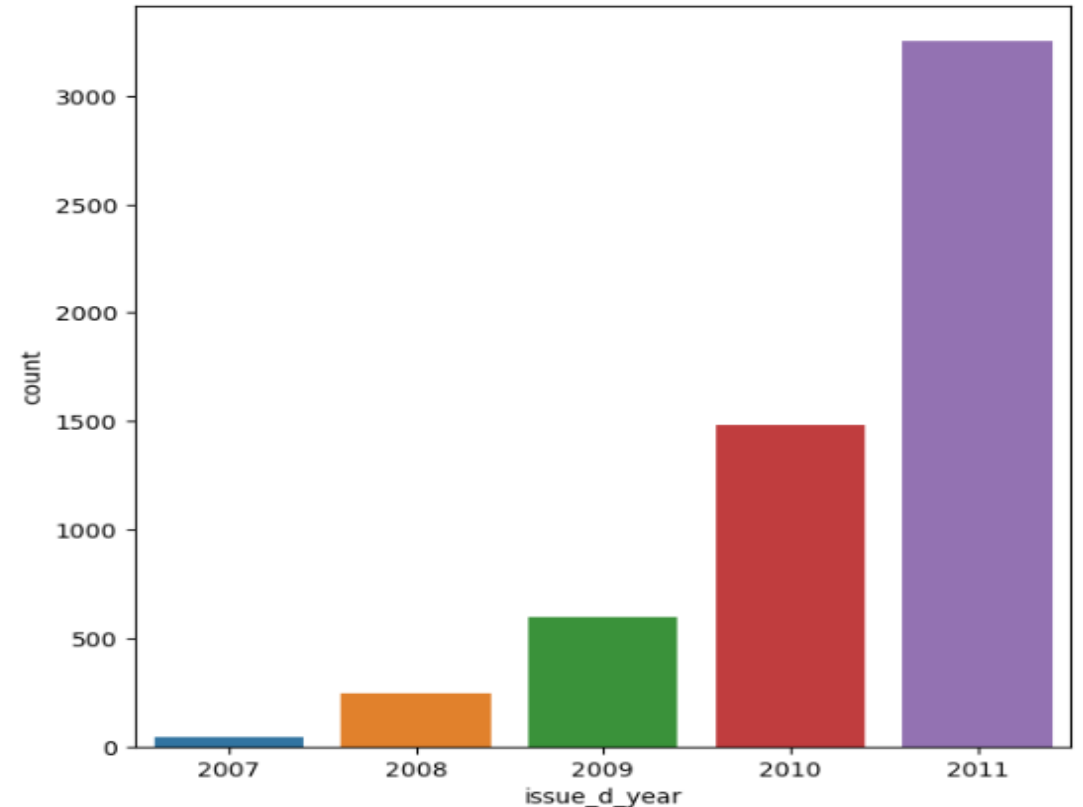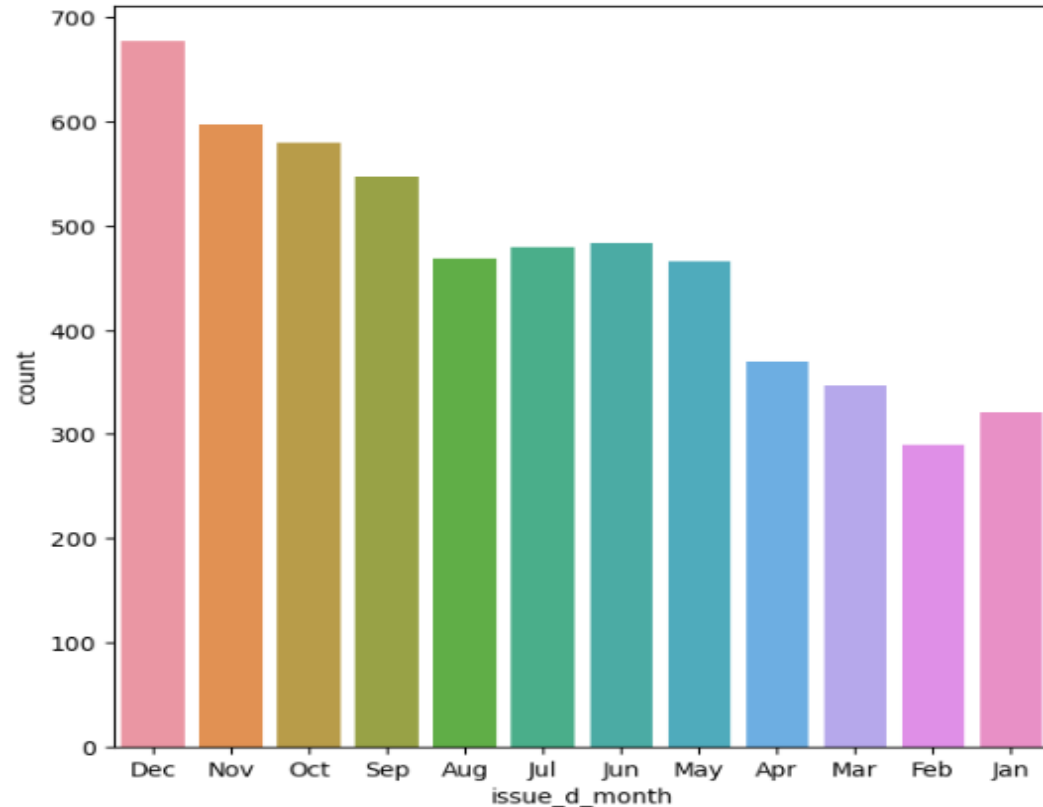

- Most no of loans were issued in 2011 and especially in the ending months of that year the number of loans steadily rose.

- Most no of defaulters were of the people whose loan was issued in the same time period. This could indicate a financial crisis resulting in more no of loans and more defaulters.

Conclusion

- **Final Observations:**

  1. Low Grade Loans have a high tendency to default. Grading system is working as expected.
  2. Borrowers who have been granted these loans at higher interest rate have a higher tendency to default.
  3. Loan Borrowers from state of CA have highest tendency to default , followed by FL and NY. Thus,for Loan applications from these states the company should scrutinize background thoroughly before granting loans.
  4. When the borrowers are taking a loan for the purpose of debt consolidation they are highly likely to default.
  5. People who do not own their homes have a very high tendency to default and loans must be granted to such population very carefully.
  6. People who are getting loans only for 36 months are defaulting the most so we should thorough checks here as well.

- **One Interesting Factor:** We figured out that the number of loans have increased steadily throughout the years and it seems most loans were taken in the year 2011 in the ending months and this was also the time period in which most loans were defaulted. It could indicated that this happened due to some sort of financial crisis.

Driving Factors leading to loan defaults

Factor	Description(as per Data Dictionary)
Purpose	A category provided by the borrower for the loan request.
addr_state	The state provided by the borrower in the loan application
grade	LC assigned loan grade
home_ownership	The home ownership status provided by the borrower during registration. Our values are: RENT, OWN, MORTGAGE, OTHER.
Term	The number of payments on the loan. Values are in months and can be either 36 or 60.
int_rate	The number of payments on the loan. Values are in months and can be either 36 or 60.