GRAMNER CASE STUDY

SUBMISSION

BY:

Shubhi Upadhyay

Nikhil Srivastava

Lankireddy Sita rami reddy

Siddharth jain

Abstract

Business Objective:

- the company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.
- The company can utilize this knowledge for its portfolio and risk assessment.

Strategy: Plot charts to visually identify problems areas

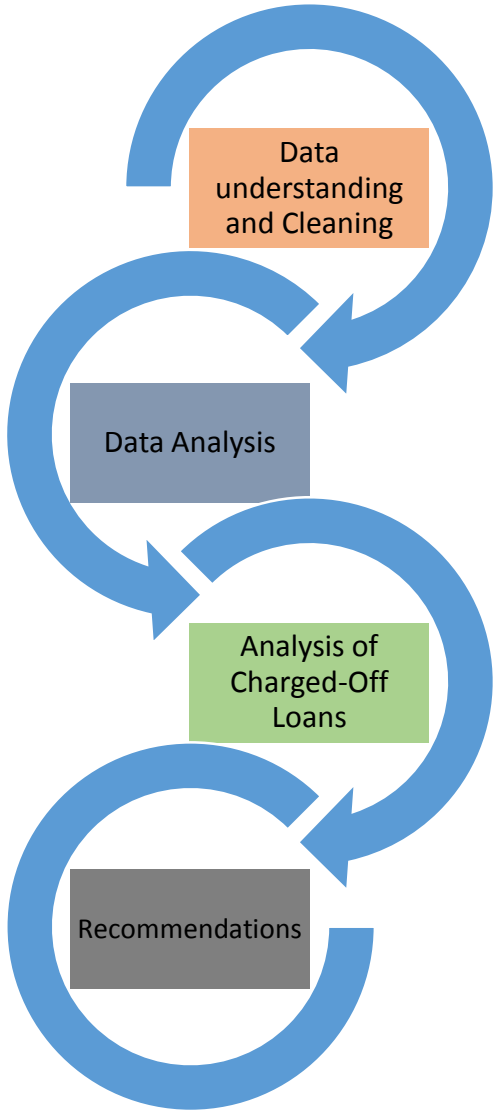
Constraints:

- Loans rejected by bank are not considered

Data Source:

- complete loan data for all loans issued through the time period 2007 to 2011.

Problem solving methodology



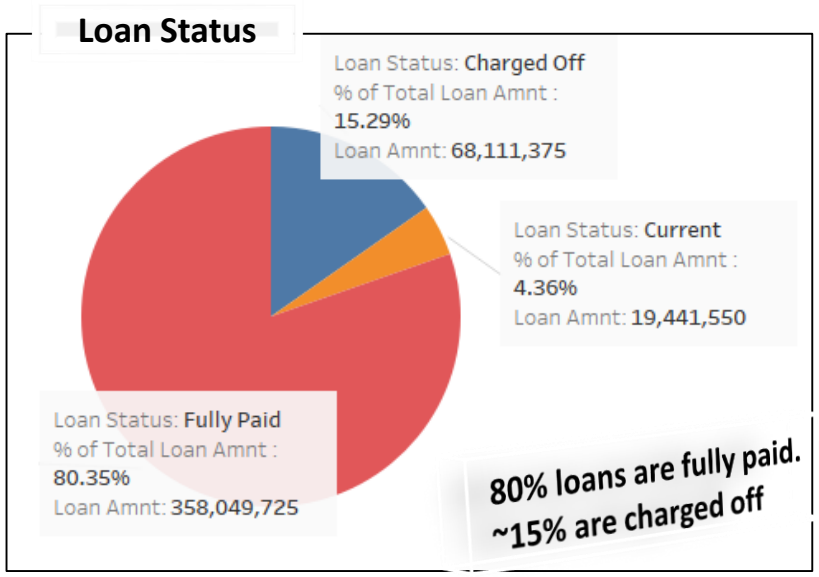
- AS-IS loading Loan in R
- Analysis of key columns of loan data.
- Remove unwanted columns having 0 and NA
- Check for Duplicate & null values
- Convert Issue Date to standard format

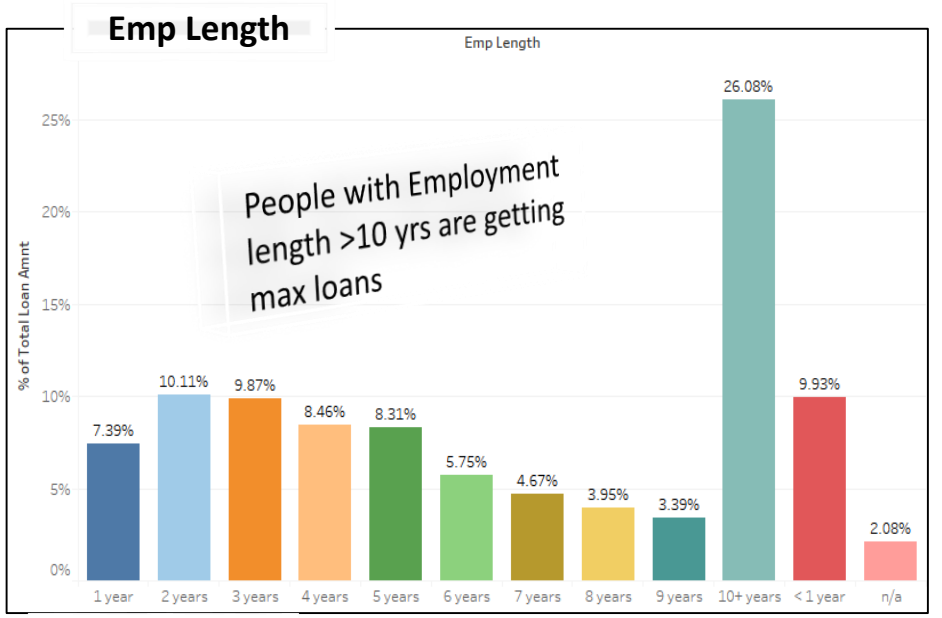
- Analysis of raw data using univariate & bivariate analysis.

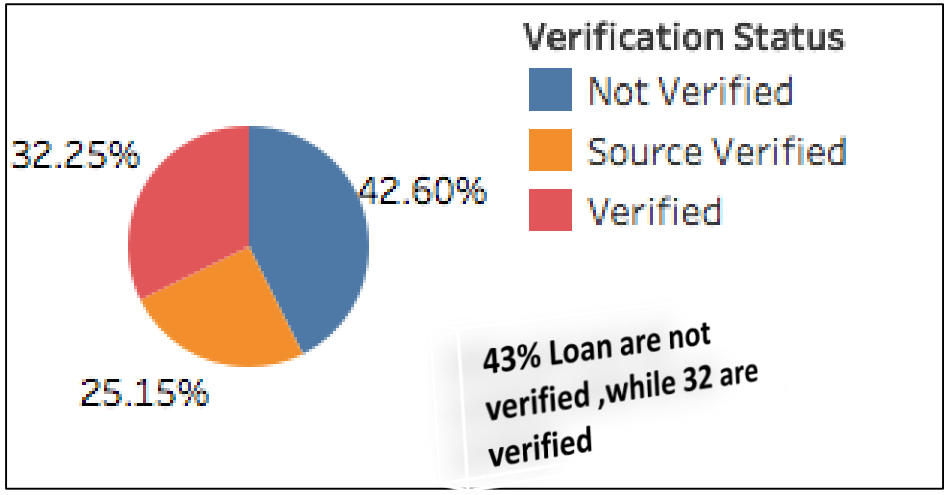
- Segmentation Analysis of Charged off Loans

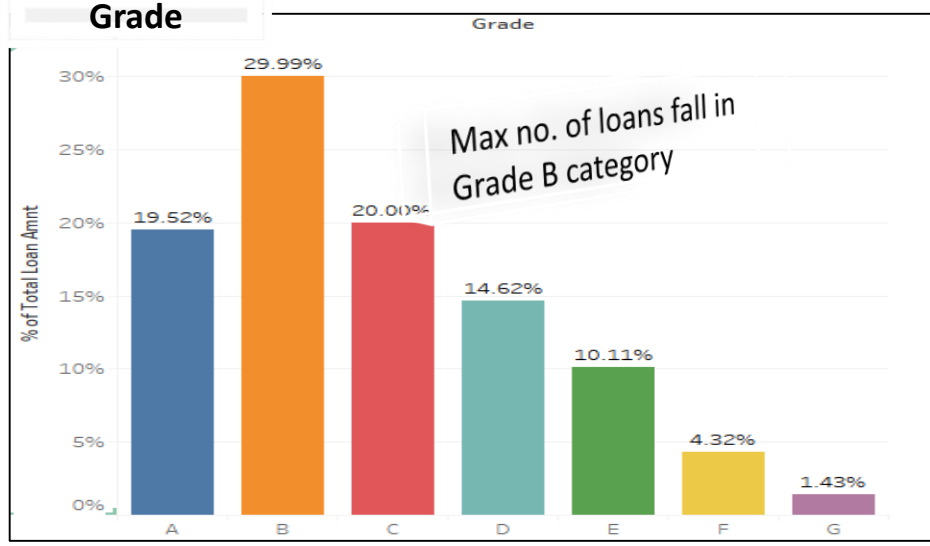
- Identify driving factors behind loan defaults

Univariate analysis on Key attributes

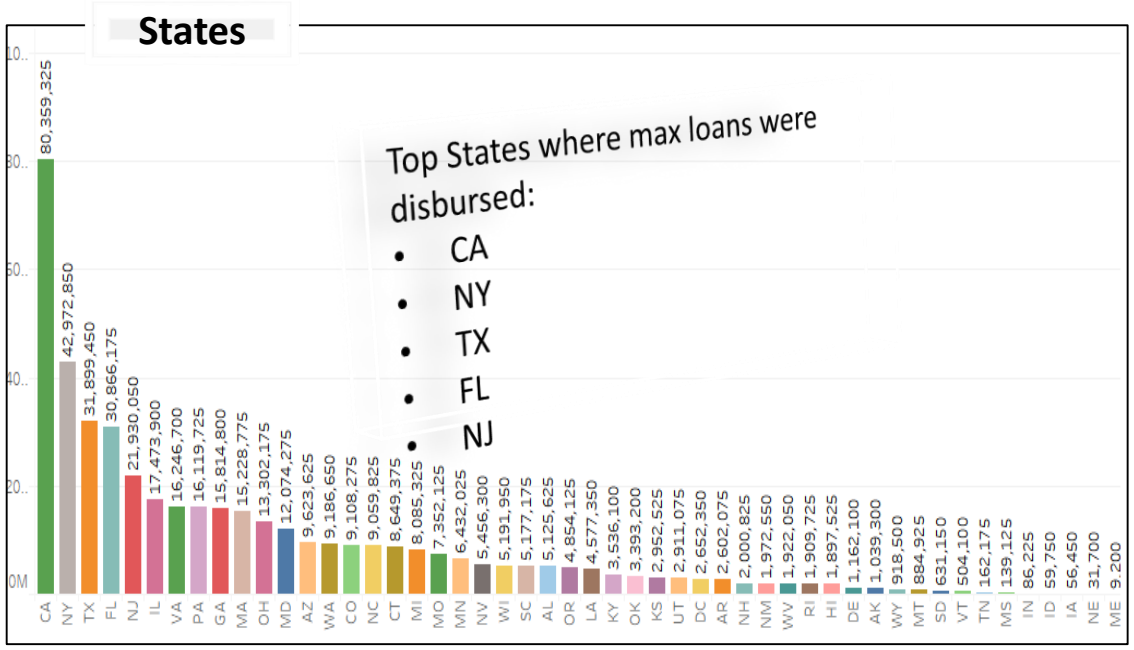


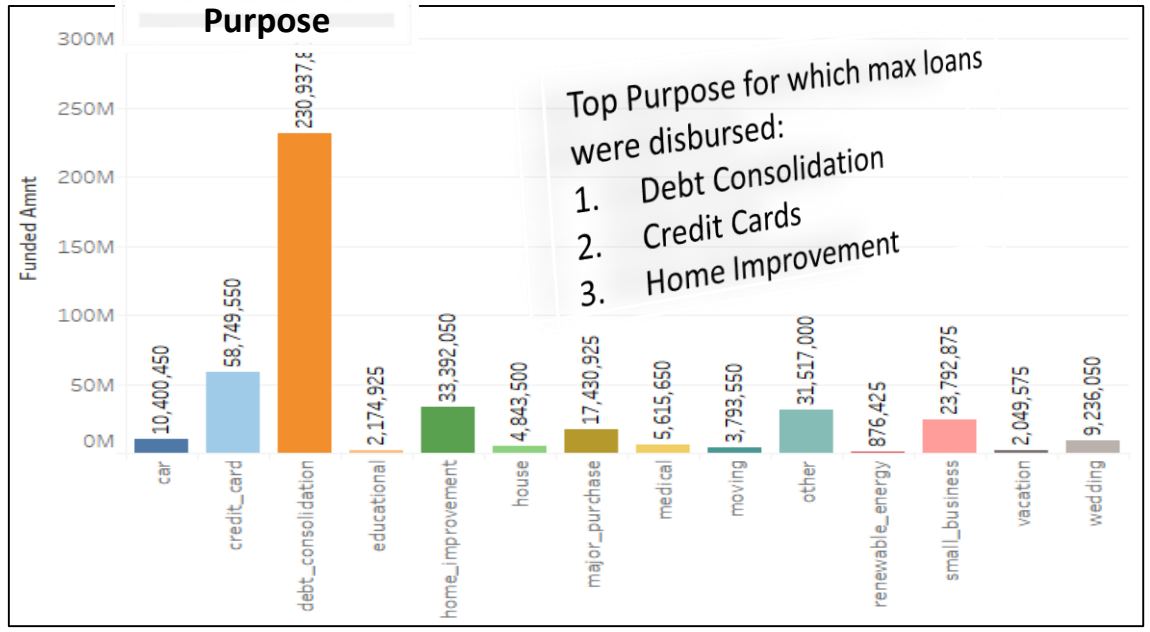


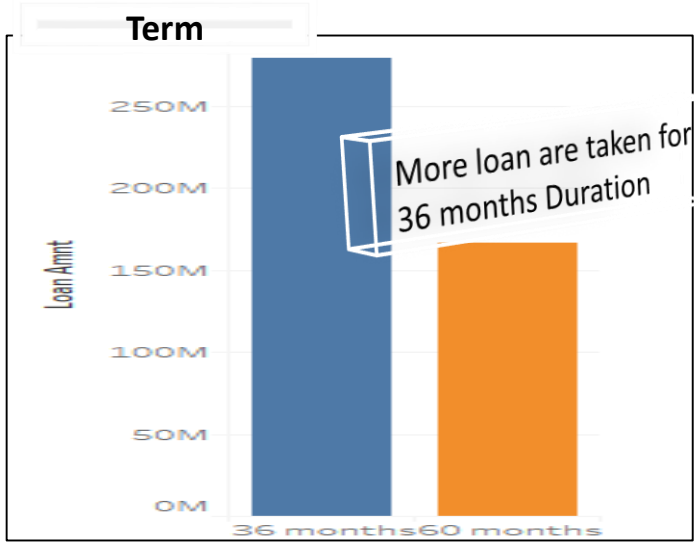


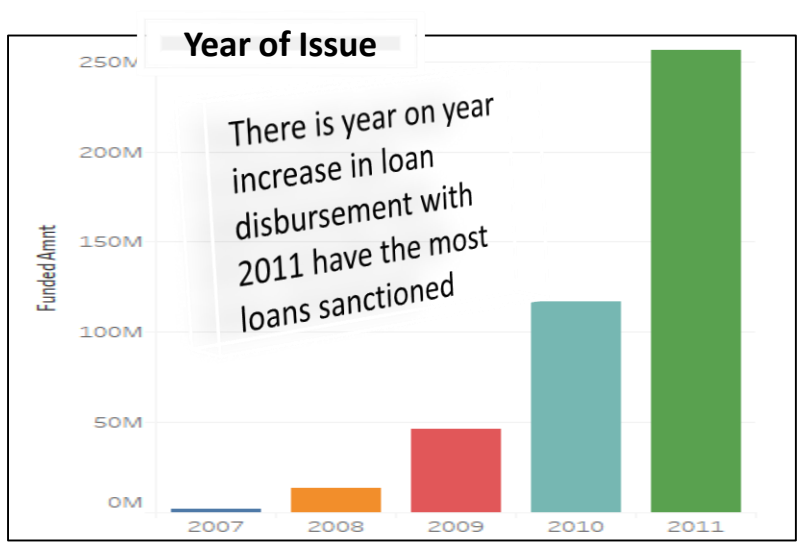


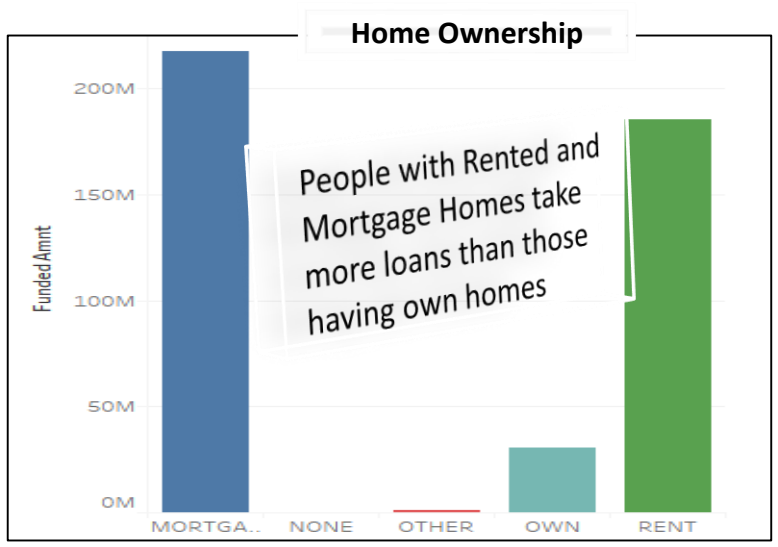
Univariate analysis on Key attributes (contd..)





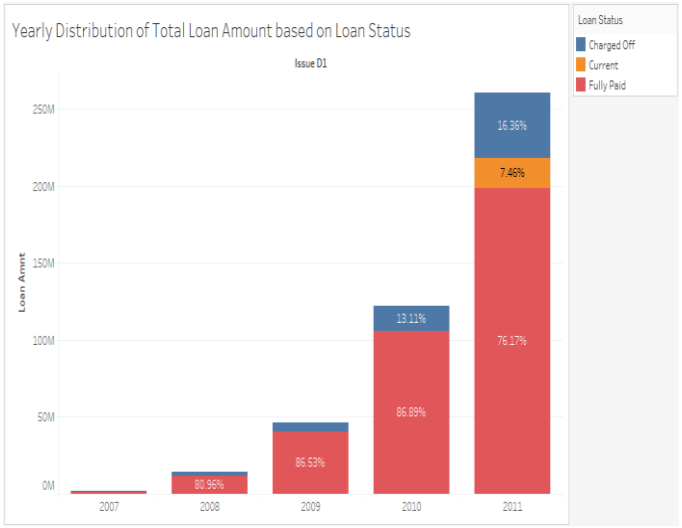


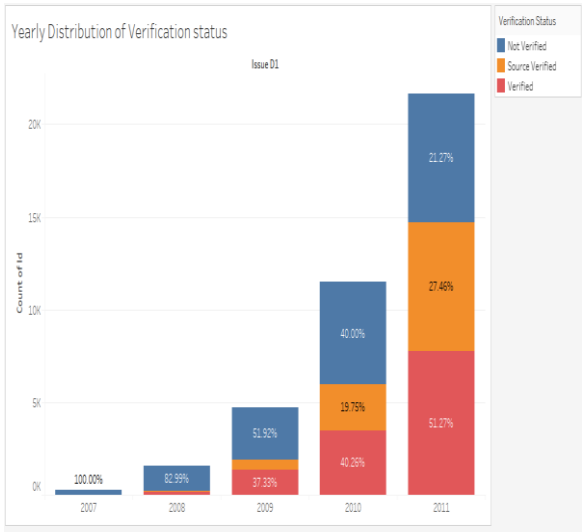


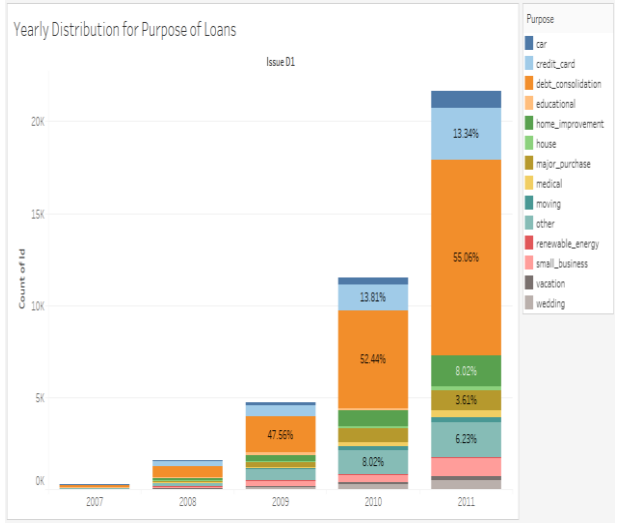


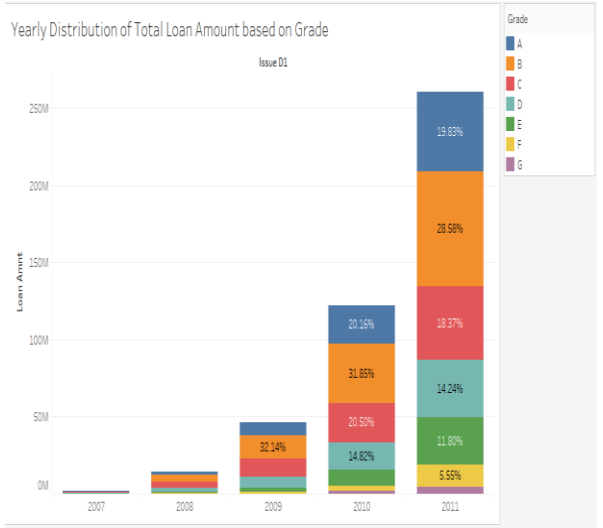
Bivariate analysis on Key attributes

Variable: Year in which Loans were sanctioned

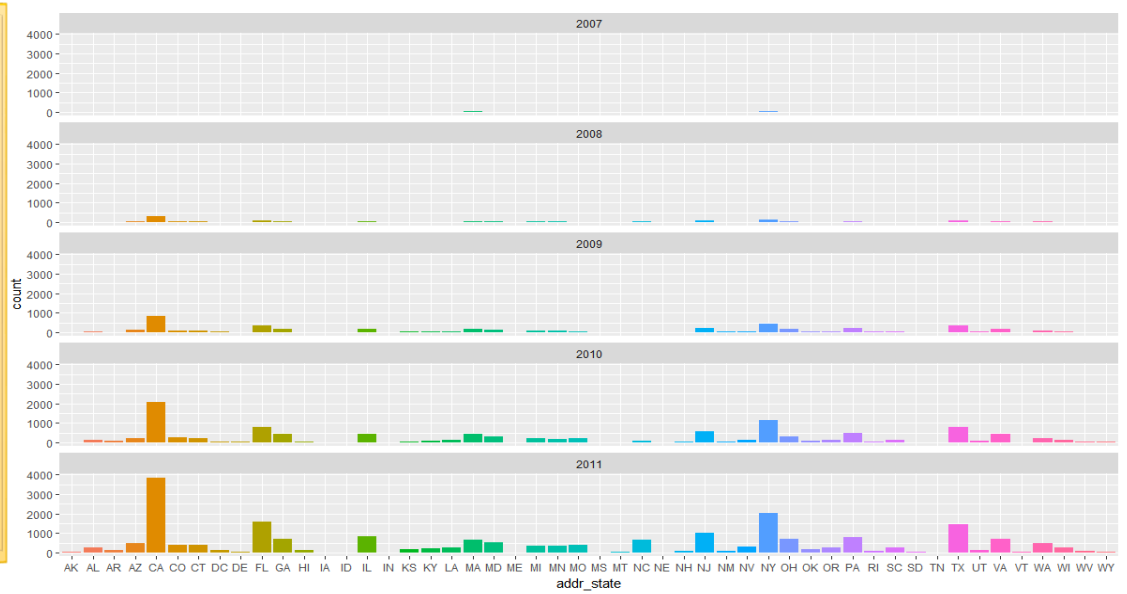




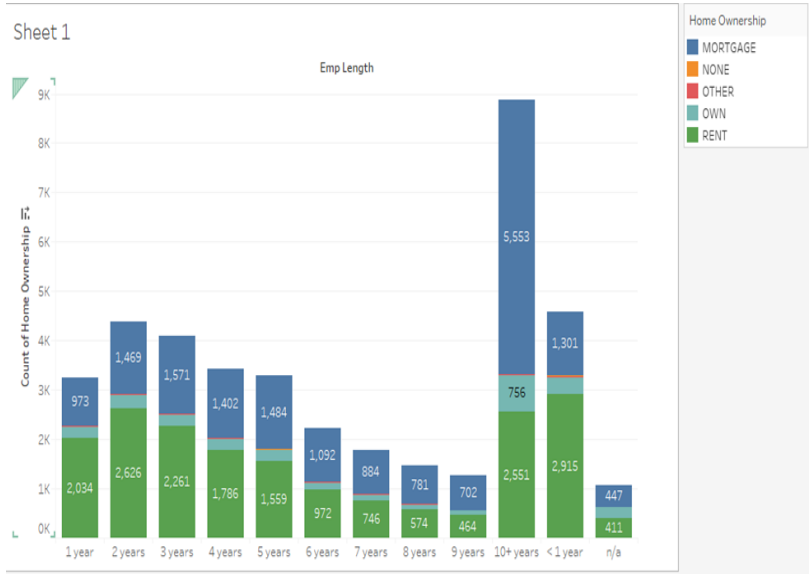


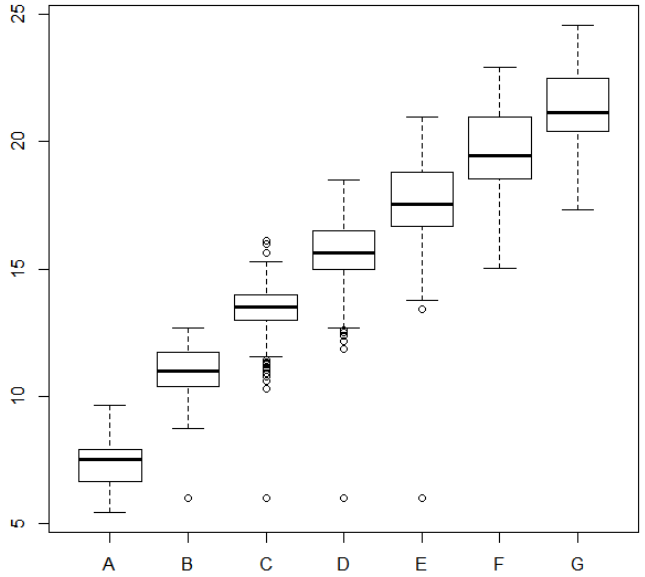


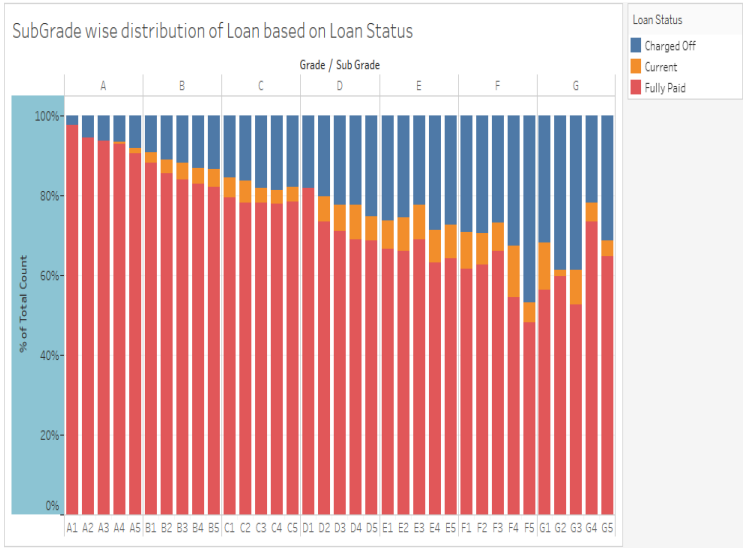
- Since most loans were disbursed in 2011, most bivariate analysis shows higher count/percentage in 2011.
- The trends identified in univariate analysis are also visible in bivariate analysis using issued year.
- **On Loan Status:**
  - A good percentage of loan are Fully Paid year on year.
  - The count and percentage of Charged-Off increases as we move from 2007 to 2011
- **On Verification Status:**
  - More loans are getting Verified year on year while not-verified %age is decreasing.
- **On Purpose:**
  - Top categories every year: Debt consolidation, credit card, home improvement
- **On Grade:**
  - Max loans sanctioned in Grade B year on year.
- **On State:**
  - Max loans sanctioned to States: CA, NY, TX, FL , NJ

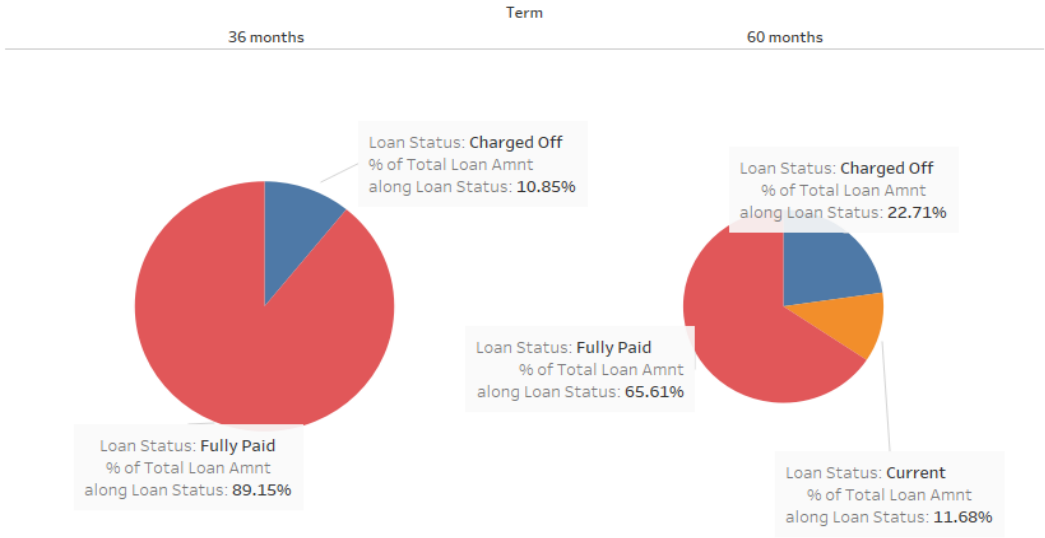


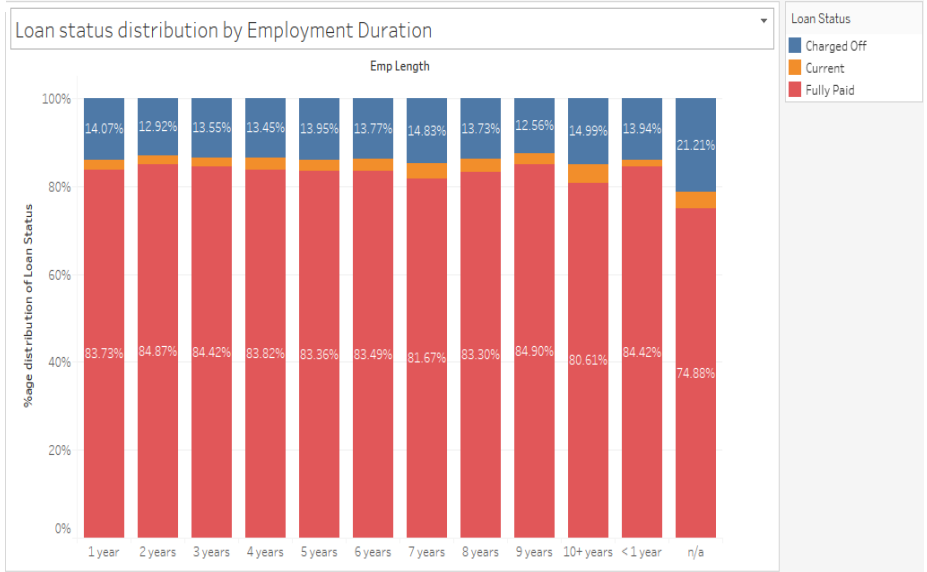
Bivariate analysis on Key attributes (contd..)



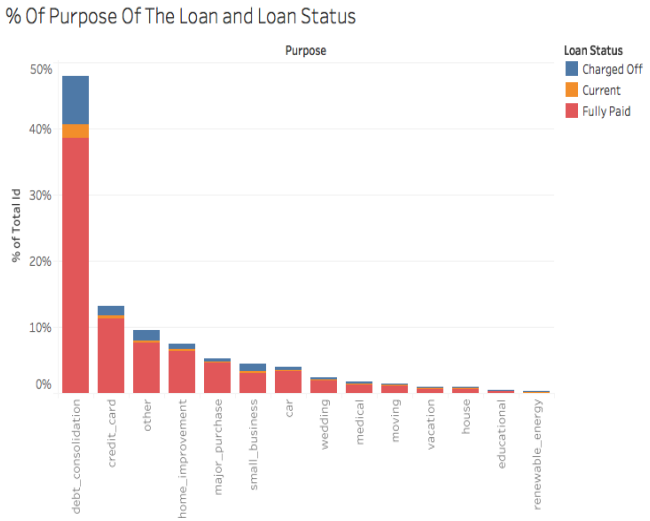


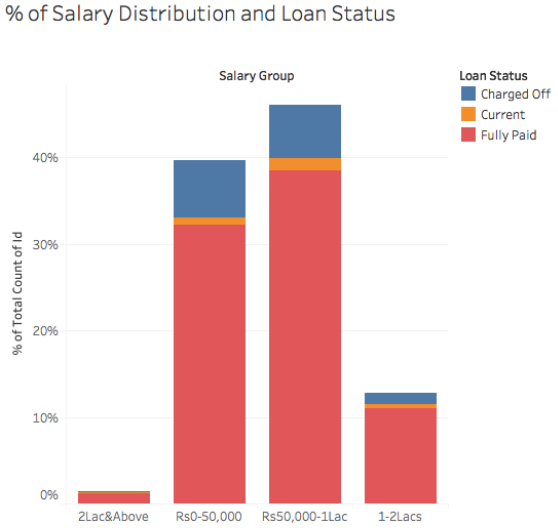


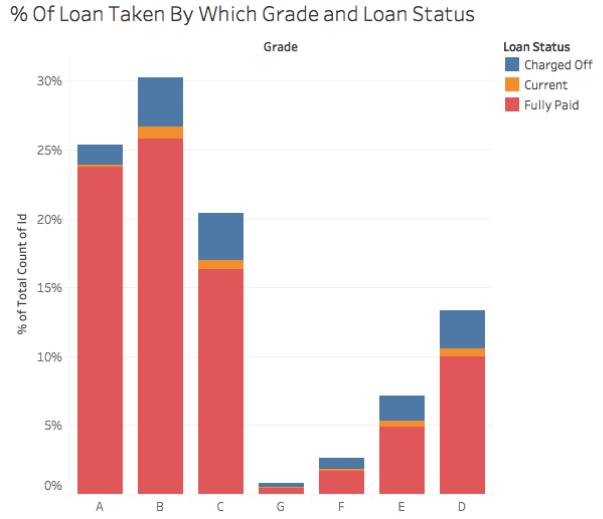




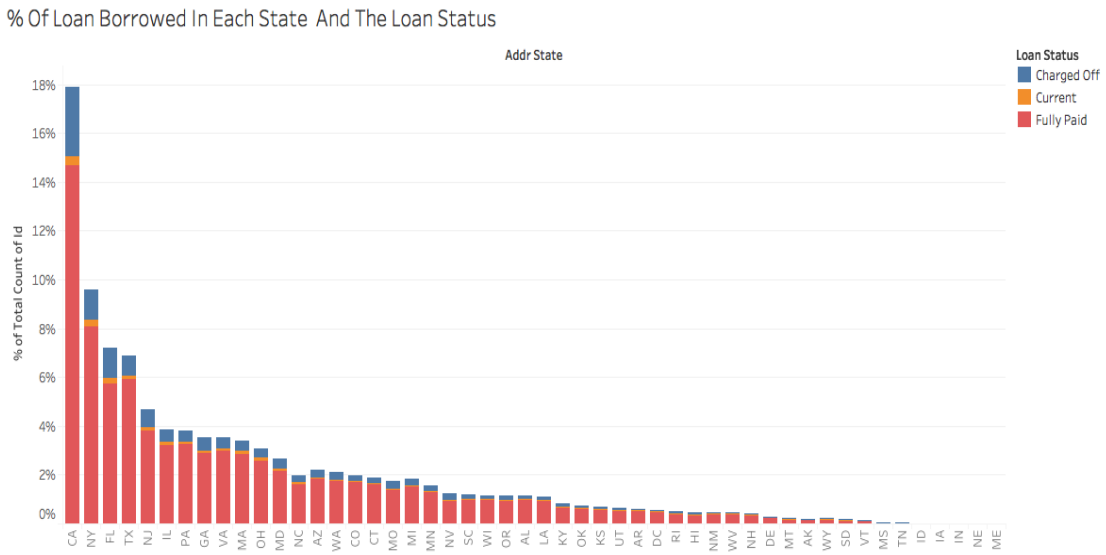
Bivariate analysis on Key attributes







- The top most 5 factors where the % of charged off customer is highest are mentioned below:
- The trends identified in bivariate analysis is also visible in segmented analysis.
- **On Term:**
  - A good percentage of loans which were charged off lie in 36 months term.
  - The same thing is visible in segmented analysis as well
- **On Purpose:**
  - Top categories : Debt consolidation, credit card, home improvement
- **On Salary:**
  - People with salary between 0 to 1 Lac have high chances of getting defaulted.
- **On Grade:**
  - Max loans sanctioned in Grade B.
- **On State:**
  - Max loans sanctioned to States: CA, NY, FL, TX , NJ



Bivariate analysis on Key attributes (contd..)

Salary Distribution & Loan Status

- People with salary between 0 – 50,000 are the one who take most loan and the one who get defaulted.

Interest Rate & Grade

- As the Grades increases, the interest rates increase.

Subgrade & Loan Status

- **Grade A and its subgrades are best categories** in terms of payments and has least count of loans that had been charged-off
- There is a increase in count of loans that were charged off from A to G (barring some exception)
- **Grade F and SubGrade F5 has the worst rate** with 47% loans marked as "Charged-Off"

Term & Loan Status

- 36 Months:- Most loan were defaulted in 36 months term.
- **The variation in % distribution is highest in 60 term but if we see overall variation 36 month tops the list.**

Employment length & Loan Status

- **80-85%** of the Total loans were **Fully Paid** across all employment durations.
- **12-15%** of total loans were **Charged-Off** across all employment durations.

Segmented Analysis on “Charged-off” loans

Problem Statement

- Understanding Charged off Loans across various variables.
- Identify the key variables that impact bad loans

Criteria

- Only loans that were marked as “Charged-Off” would be considered
- Amount of loan pending for such loans used for analysis

Assumption

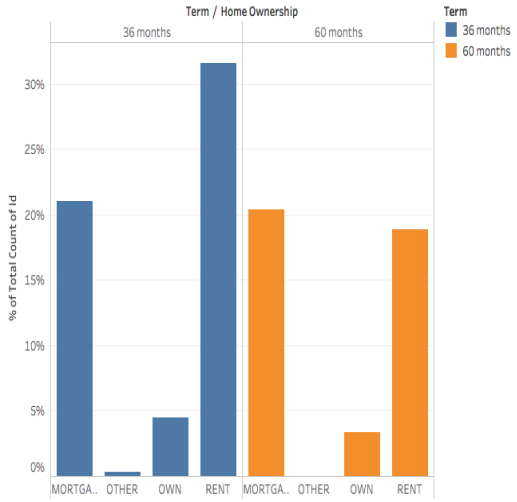
- Columns having NAs and 0 were removed

Conclusion

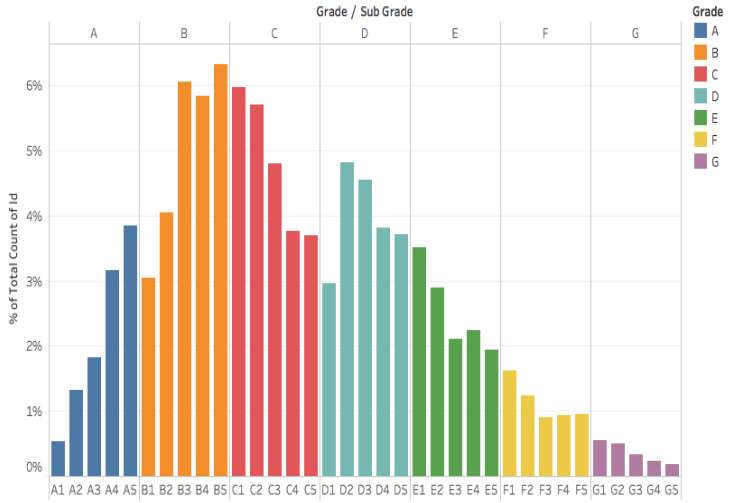
- Key Variables:
  - Purpose, State, Grade, Home ownership

Segmented Analysis on “Charged-off” loans

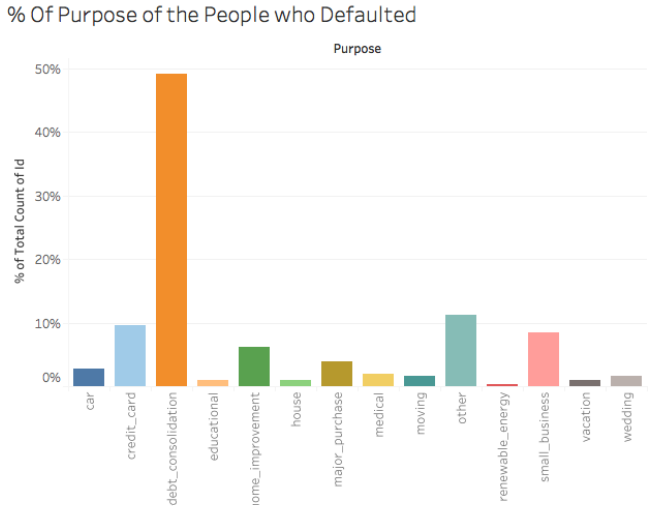
% Of Defaulted people in Each Term With Home Ownership



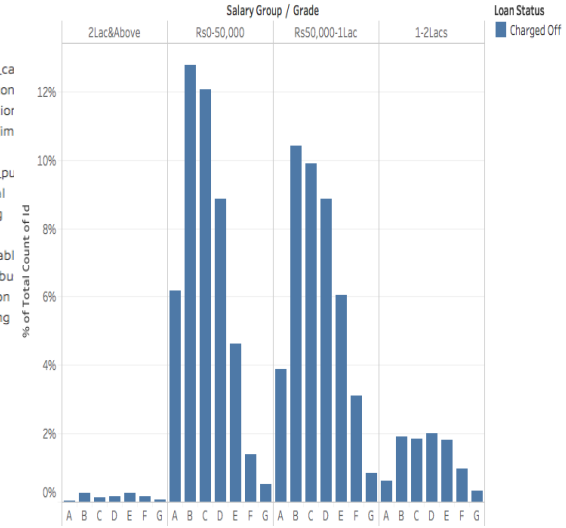
% Of Defaulted people in Each Grade



% Of Purpose of the People who Defaulted

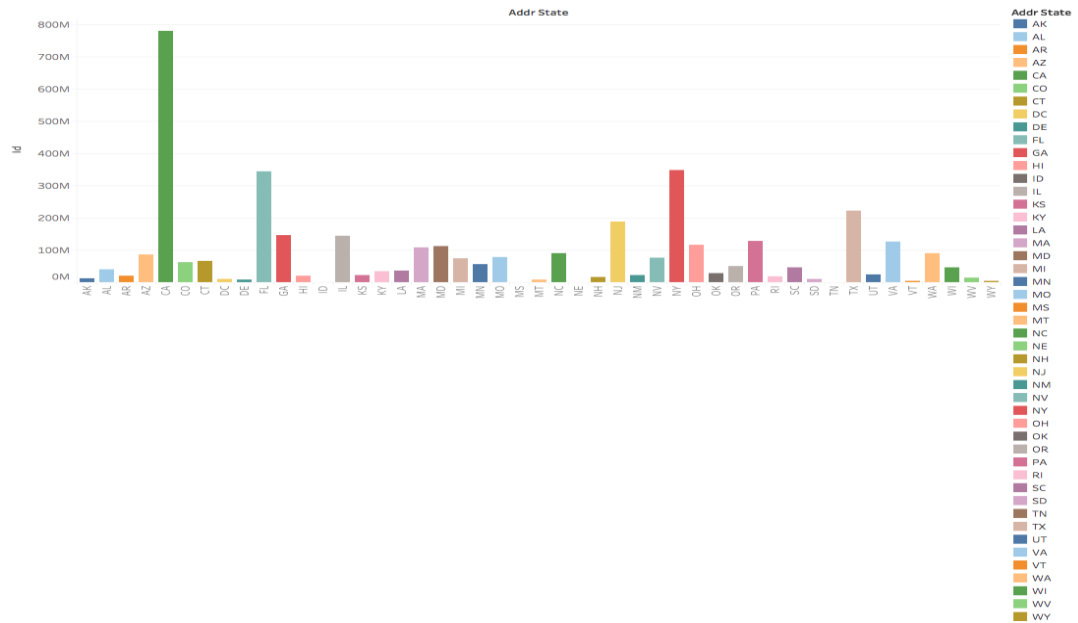


% Of People in each Salary Group for Charged off People

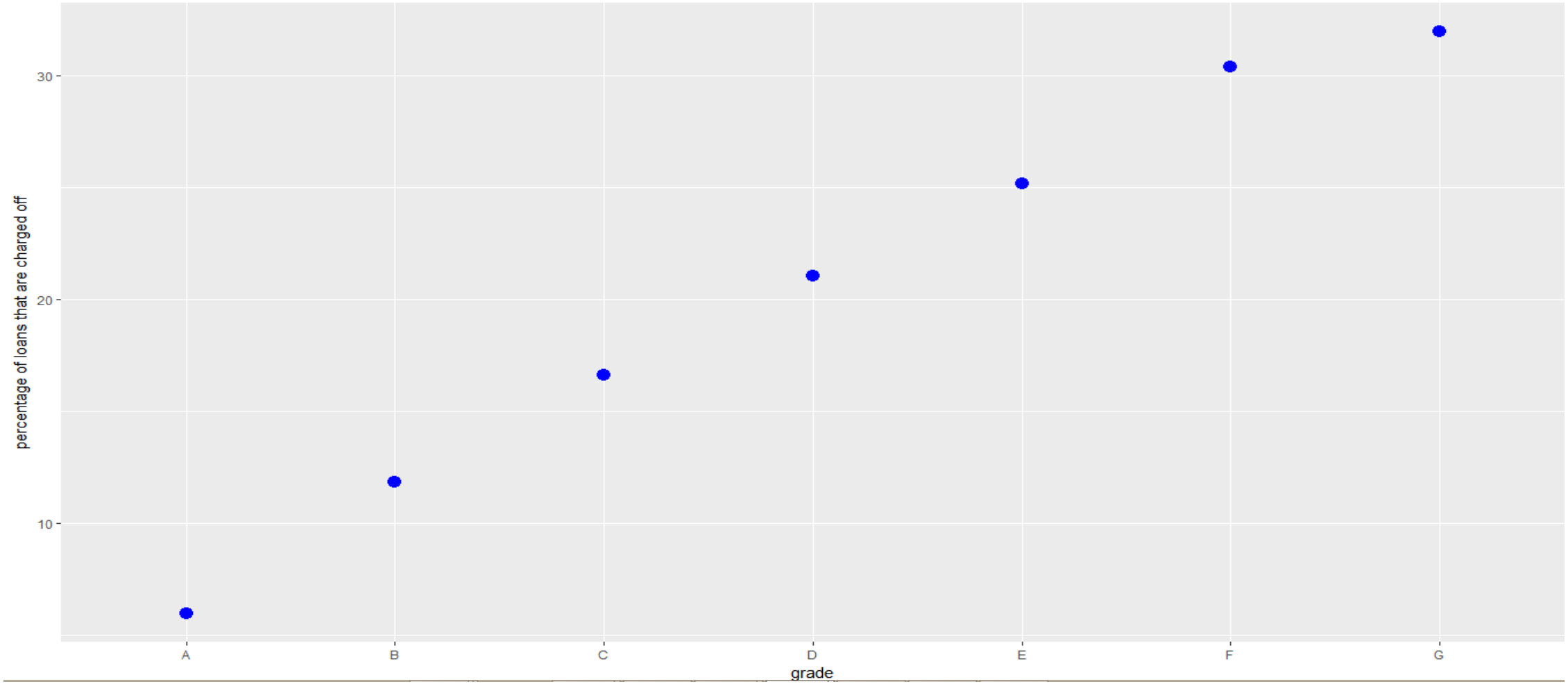


- If we do similar analysis only on people who charged off the loan, we can see that again the below given factors tops the list.
- **On Term:**
  - Again loan with 36 term top the list.
  - People who stay on rent or taken mortgage take most loan here
- **On Purpose:**
  - Top categories Debt consolidation, credit card, home improvement here also
- **On Salary:**
  - People with salary between 0 to 1 Lac have high chances of getting defaulted.
- **On Grade:**
  - Max loans sanctioned in Grade B with subgrade B5 , B3 and B4 topping the list.
- **On State:**
  - Max loans sanctioned to States: CA, NY, FL, TX , NJ

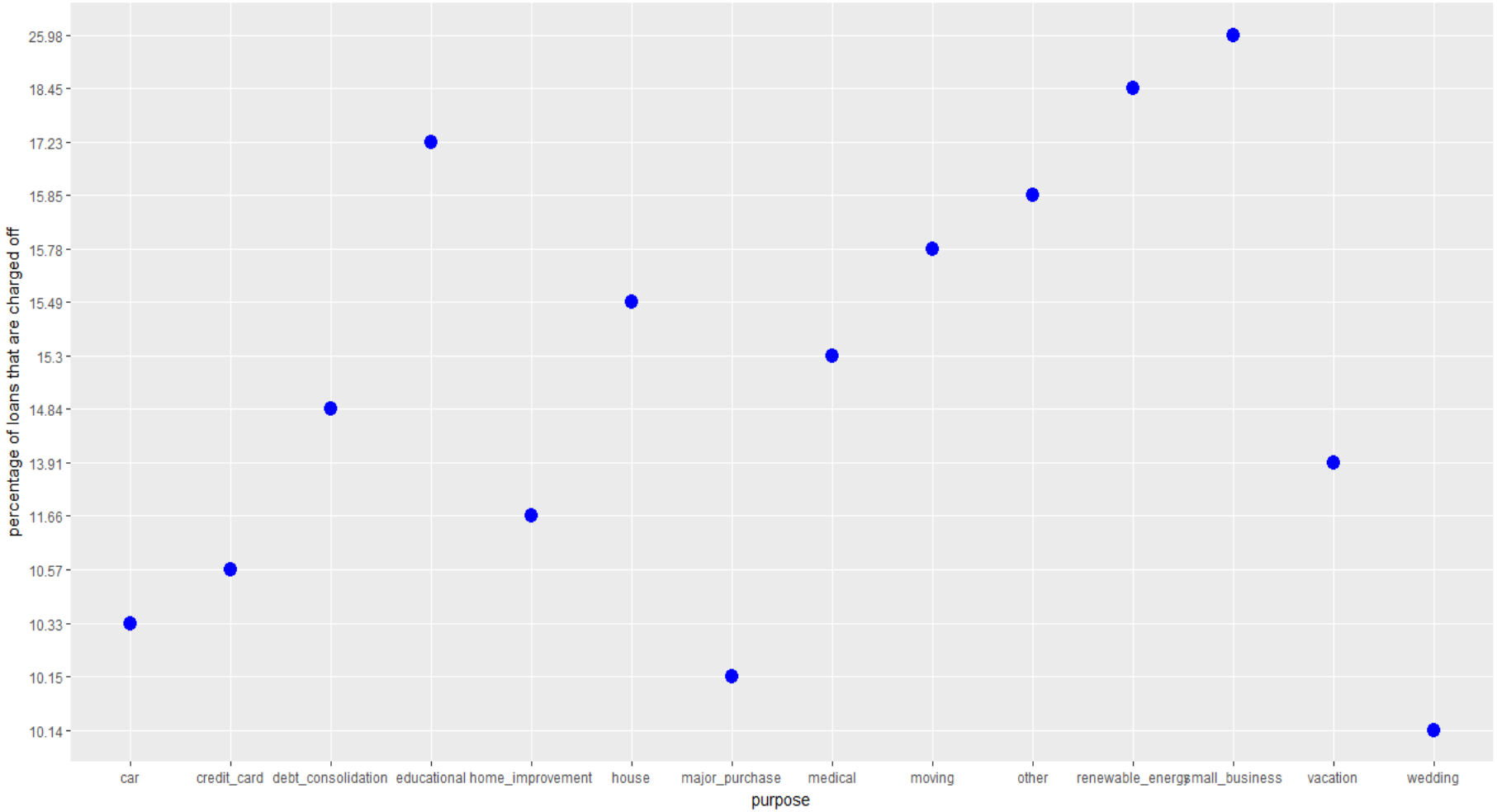
% Of People who Defaulted in each country



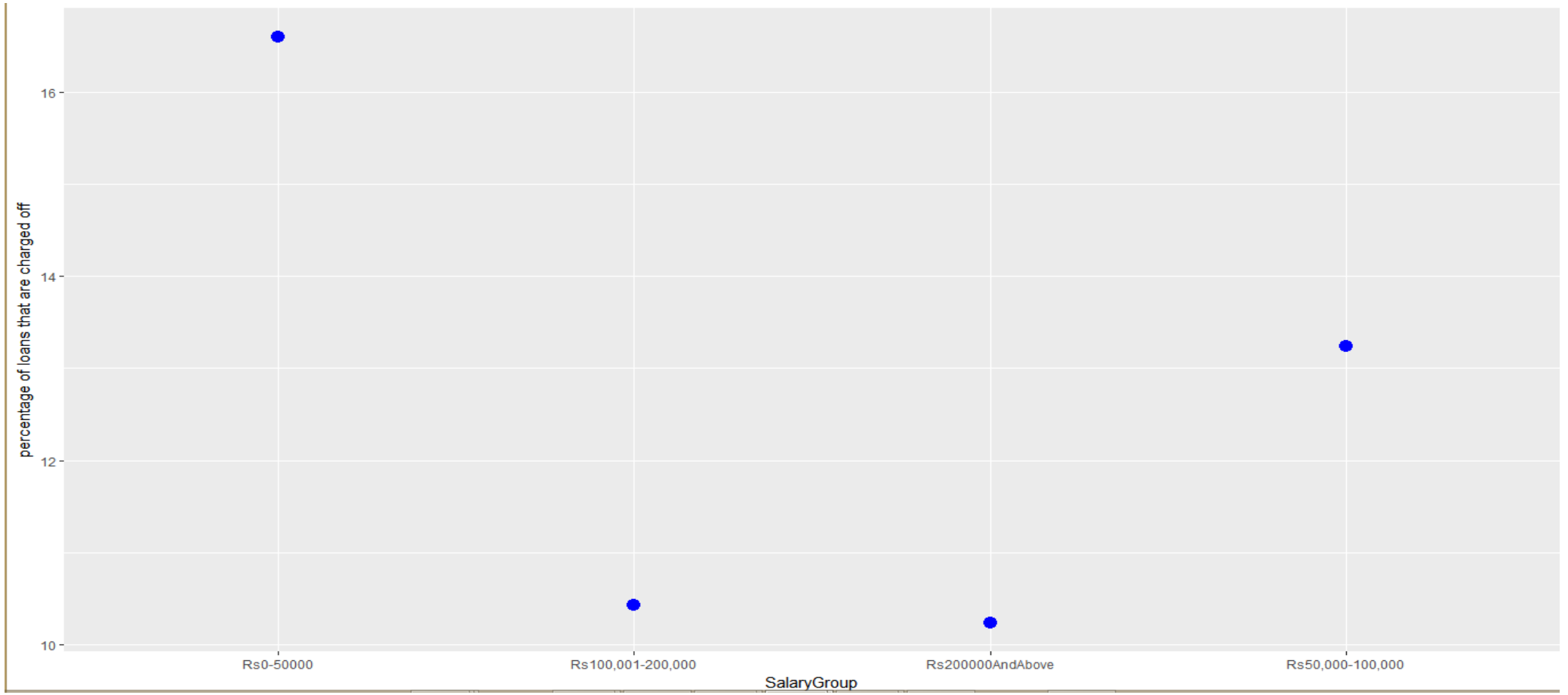
From the below plot we conclude that as grade moves from A to G the percentage of defaulters increases, so we conclude that as grade moves towards G the chances of defaulters increases.



From the below plot we conclude that there are more percentage of defaulters in small business fallowed by renewable energy. So there are more chances of defaulters in these groups



From the below plot we conclude that the percentage of defaulters is high in salary groups less than 50000 fallowed by salary group of 50000 – 100000.
So giving loans for these two groups will lead to more chances for defaulters.



Overall Analysis

- From the above analysis we can conclude that , the factors that will effect the loan status are:
    1. Grade
    2. Purpose
    3. Annual income
- As the grade increases interest rate increases.
- 80% of loans are fully paid.
- People with greater then 10+ years employee length are getting more loans.
- Maximum number of loans fall under grade B.
- CA,NY,TX,FL,NJ are the top 5 states where maximum loans are disbursed.
- Debit consolidation, credit cards, Home improvement are the top purpose for which loan is disbursed.

Overall Analysis

Year of Issue: Loan disbursement is increasing every year from 2007 to 2011.

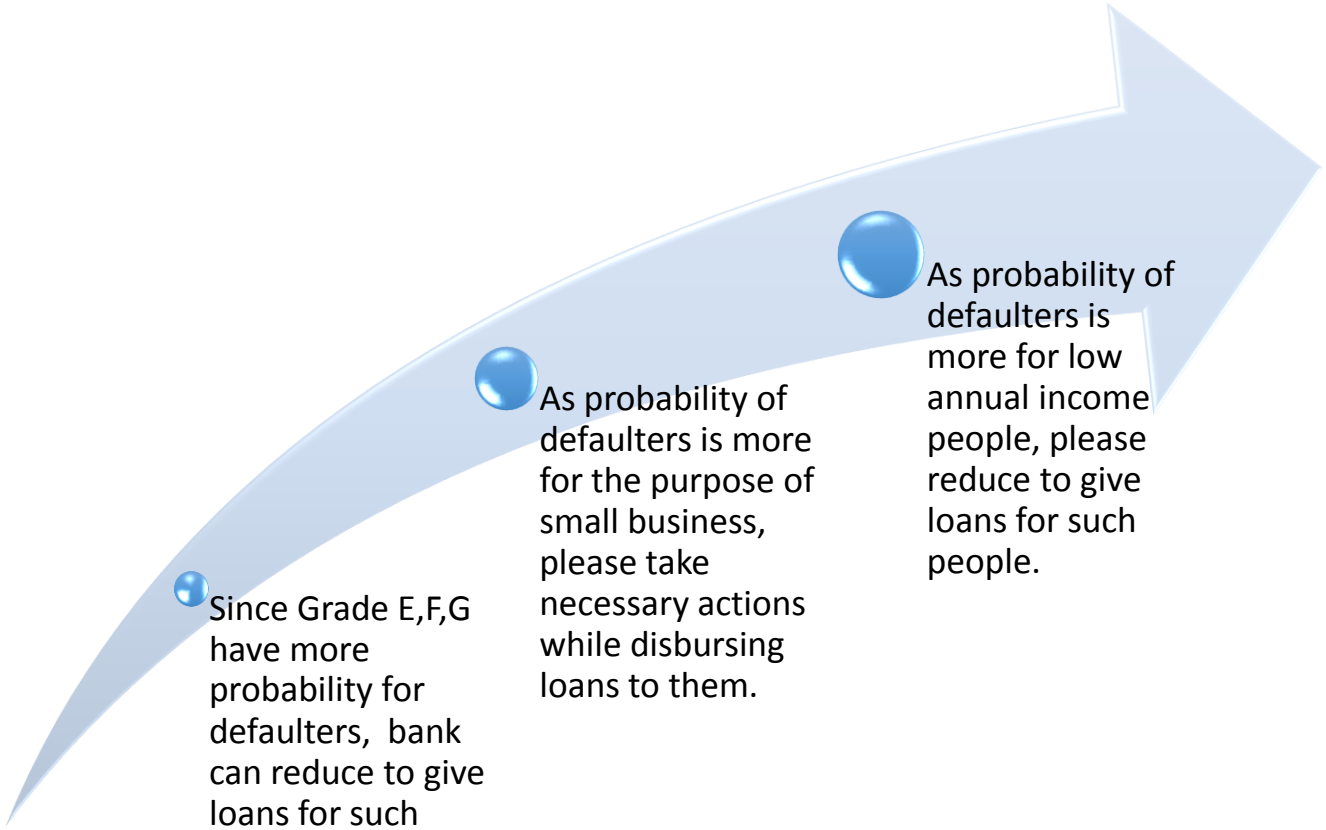
Home Ownership: People with rented and mortgage home are taking more loans.

Grade: As the grade increase from A to G the probability of defaulters increases.

Purpose: Regarding purpose of loan small business and renewable energy have more probability for defaulters.

Income: Annual income with less than 1lakh have more chances for default.

Recommendations



Thank You