



# Exploratory Data Analysis- Lending Club Case Study

Group Members:

1 - Subhrajit Pyne

2 - Utkarsh Srivastava

## Company

Lending Club 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.

## Context

Lending Club wants to understand the driving factors behind **loan default**, i.e. the driver variables which are strong indicators of default.

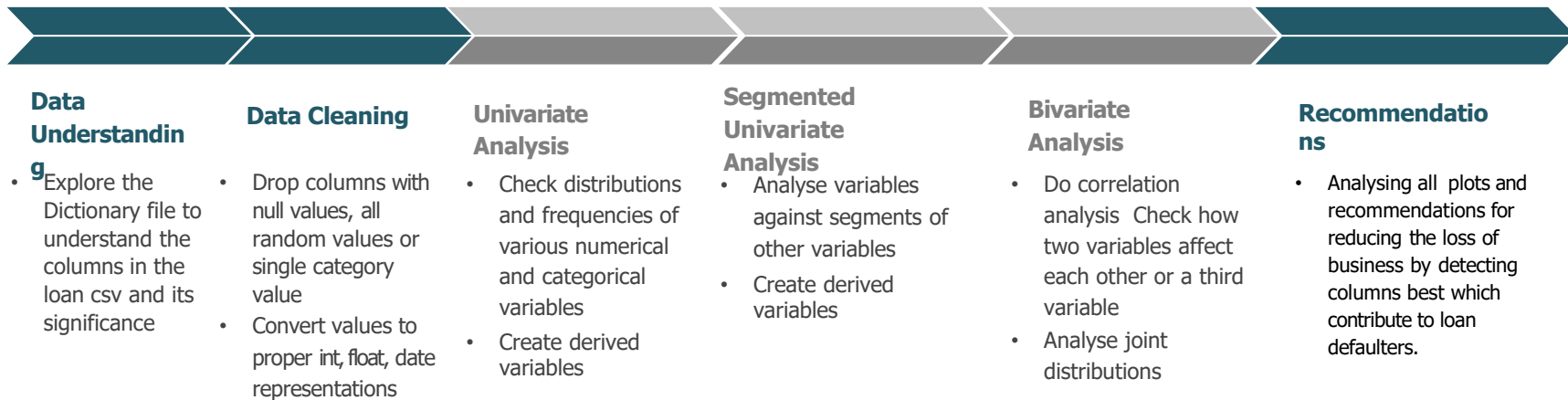
The company can utilise this knowledge for its portfolio and risk assessment.

## Problem statement

Identify patterns how **consumer attributes** and **loan attributes** influence the tendency of default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (too risky applicants) at a higher interest rate, etc.

# Analysis Approach

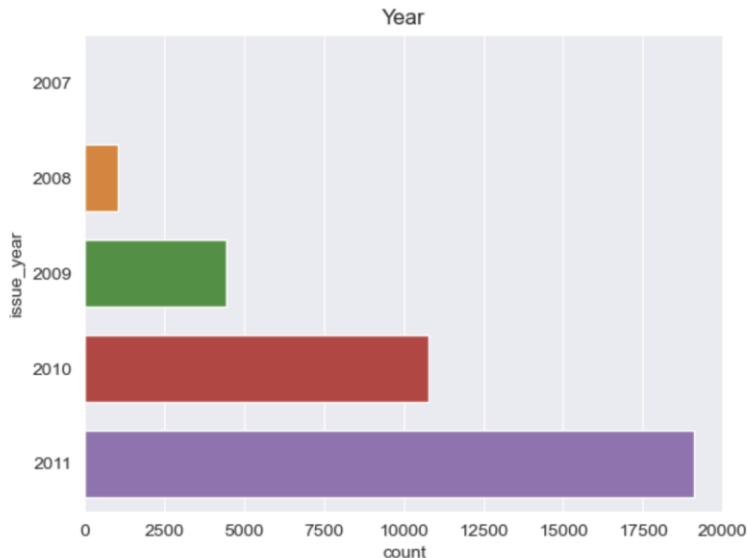
## Data Analysis



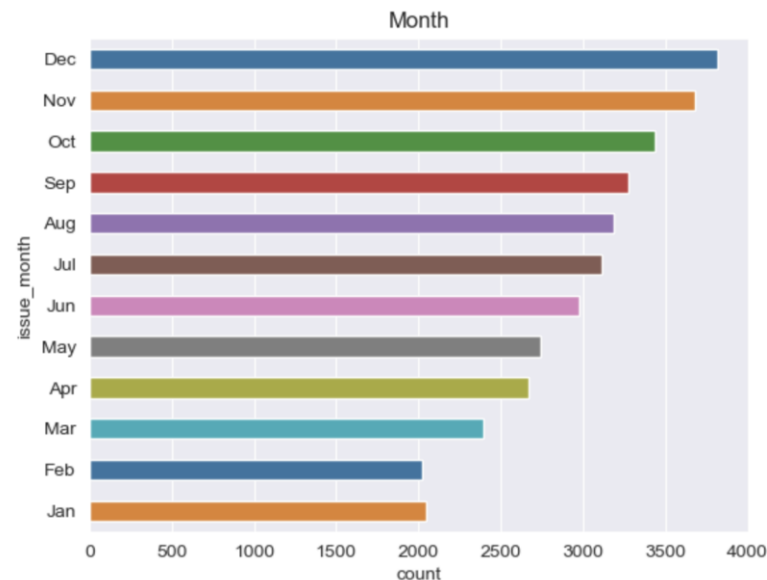
# Variable Types

- Consumer – (Annual income, Employee Length, Employee Title)
- Loan information & characteristics - (Loan Amount, Loan Status, Loan Grade/Sub Grade, Amount investment, Interest rate)
- Customer behavior – (delinquency year-2, earliest, purpose, recoveries, earliest cred line, application type, revolving balance)

# Analysis

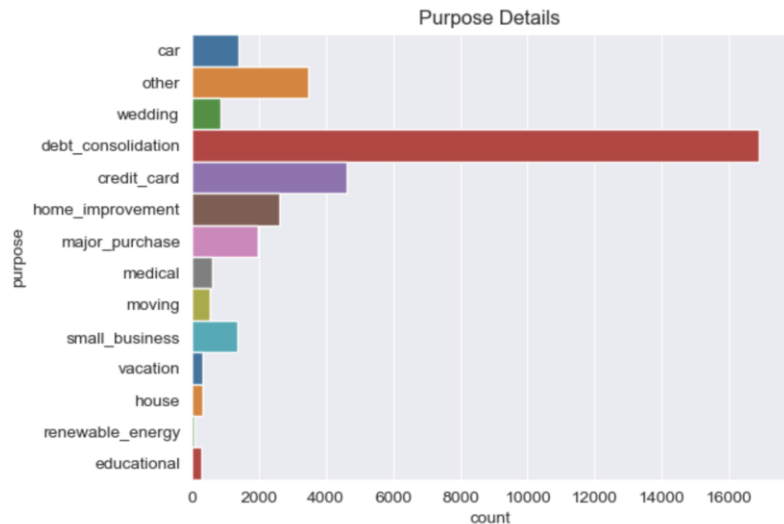


**Loan Issue Year:** Loan issuing has increasing year by year, the number of loan issued has doubled every year. In 2011, number of applicants are more

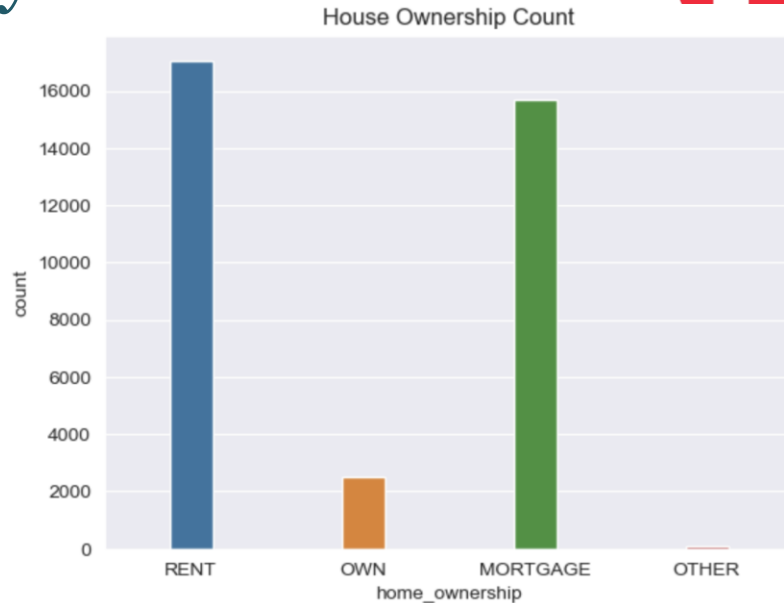


**Loan Issue Month:** Loan issuing month is increasing from January to December. In last quarter, most loans has been issued. This could be because of new year resolution.

# Analysis

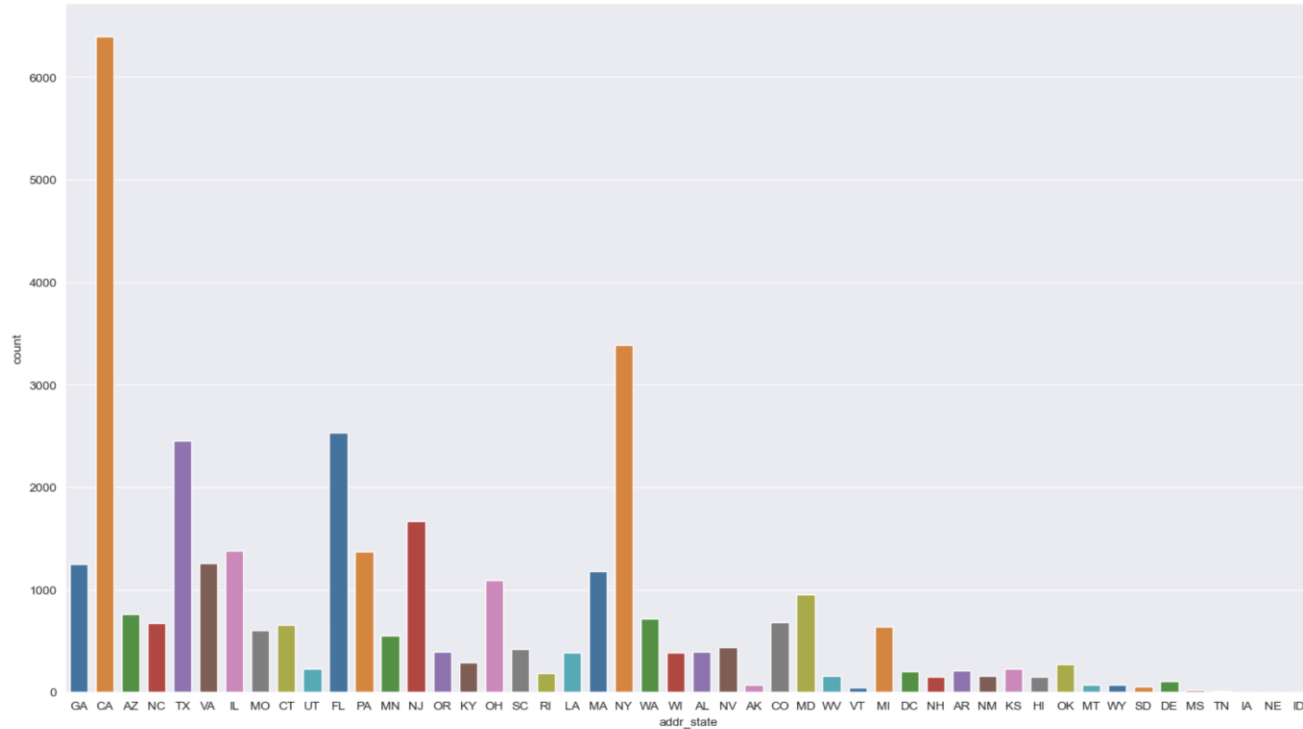


**Purpose:** Most loan are taken for debt consolidation, credit card and others.



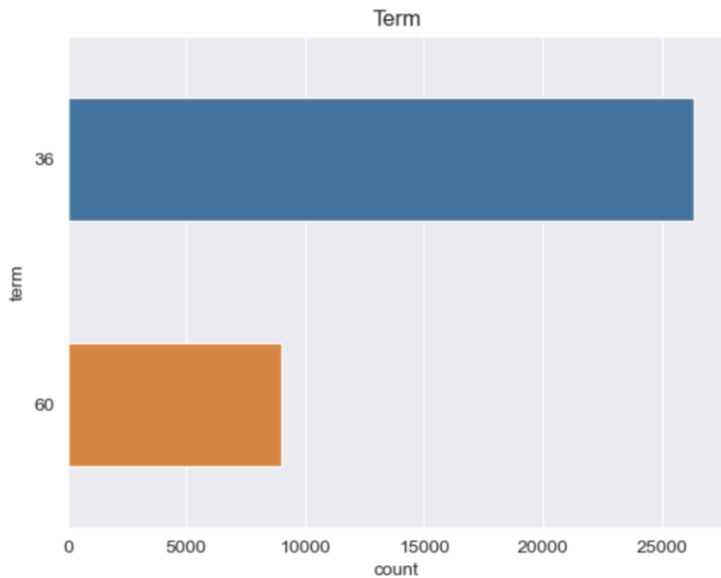
**House Ownership:** Most loan applicants live in rented house or house which is already under mortgage. These applicants have higher chance of defaulting

# Analysis

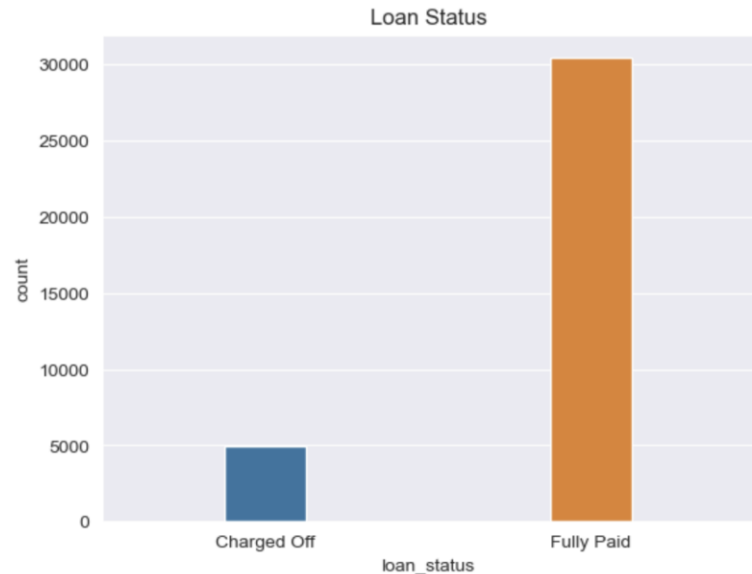


**State Address:** For cities like CA, NY, FL and TX state has the greatest number of loan applicants so chances of getting defaulted is more than other states.

# Analysis



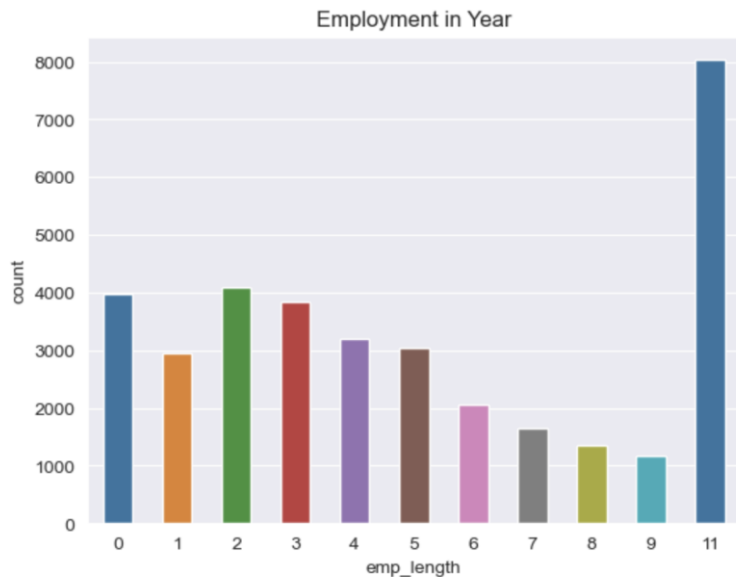
**Loan Term:** There are only two loan terms 36 and 60 months. Most of the borrowers applied for 36 months term



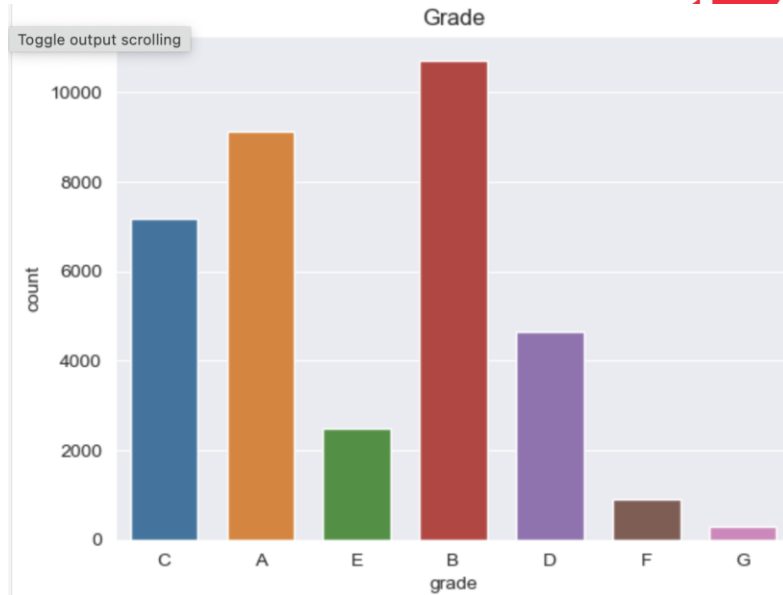
**Loan Status:** Fully paid borrowers are around 85% and charged off borrowers are around 15%



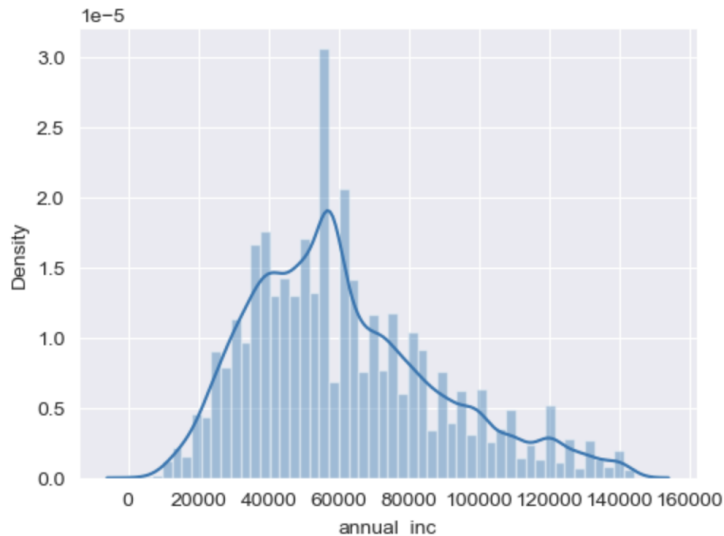
# Analysis



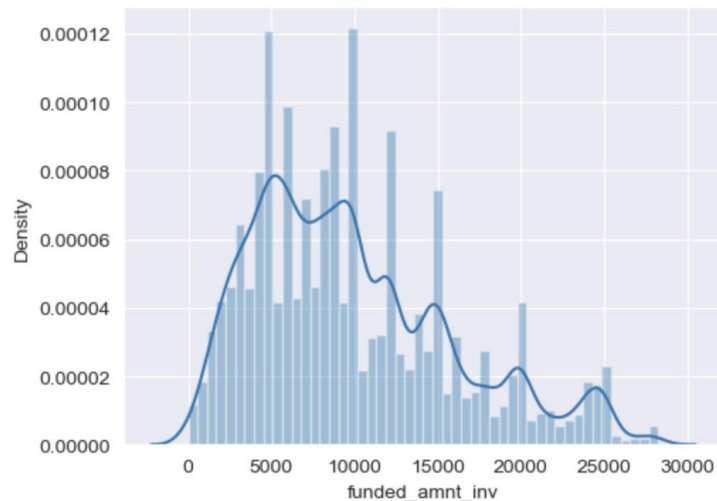
**Employment Length:** Majority of borrowers have 10+ years of experience and has highest number of defaulted loan.



**Loan Grade:** Majority of borrowers are from grade B.

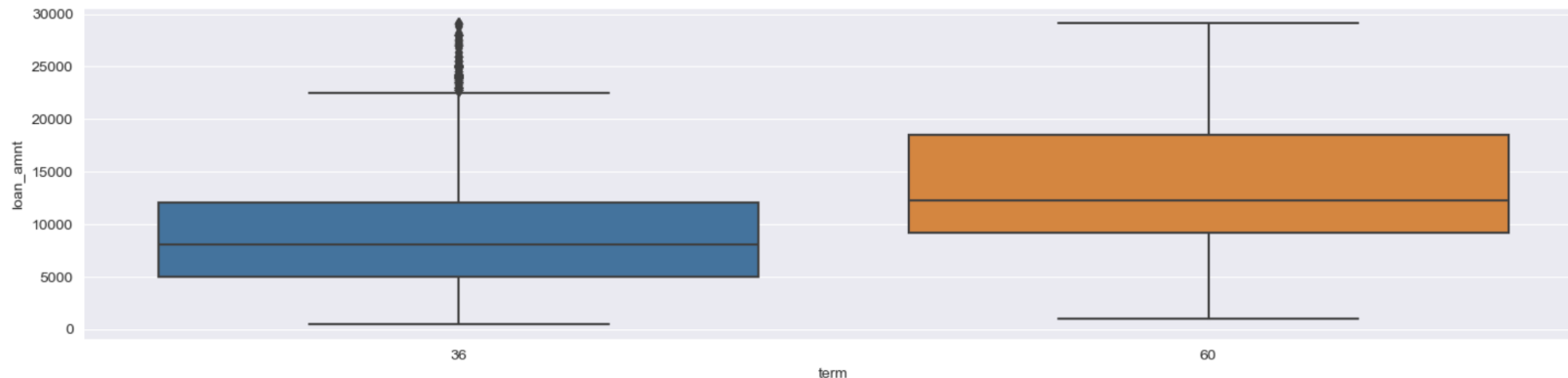


**Annual Income :** Majority borrowers have low annual income compared to rest and income lower than 57k has higher chance of defaulting.



**Approved Loan:** Approved loan amounts are varring between 5k to 14k (Q1 and Q3)

# Bivariate Analysis

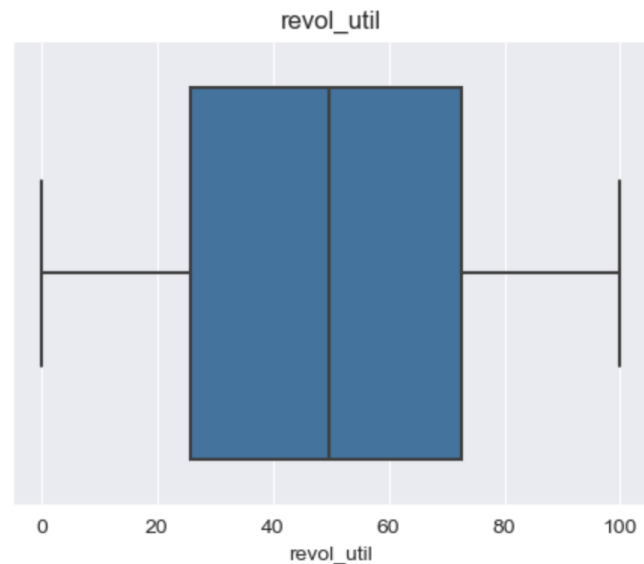
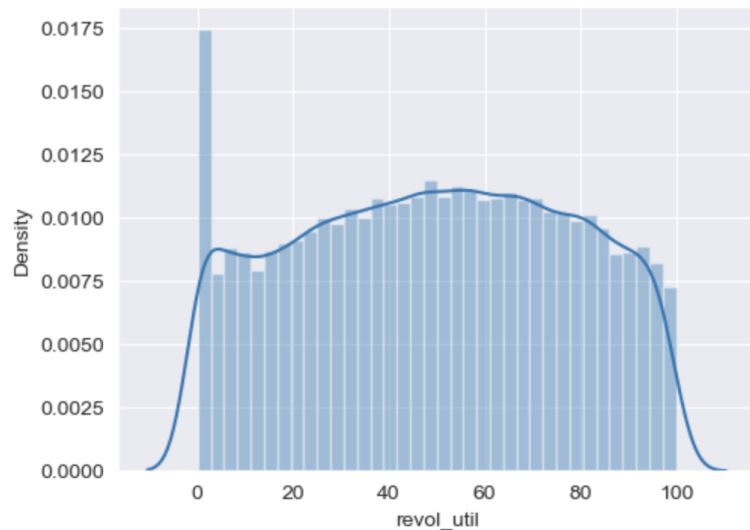


## Loan Amount :-

36 Month tenure has more fully paid customers and has a greater number of defaulters but when we go by ratio, 60 month has more defaulter than 36 months tenure

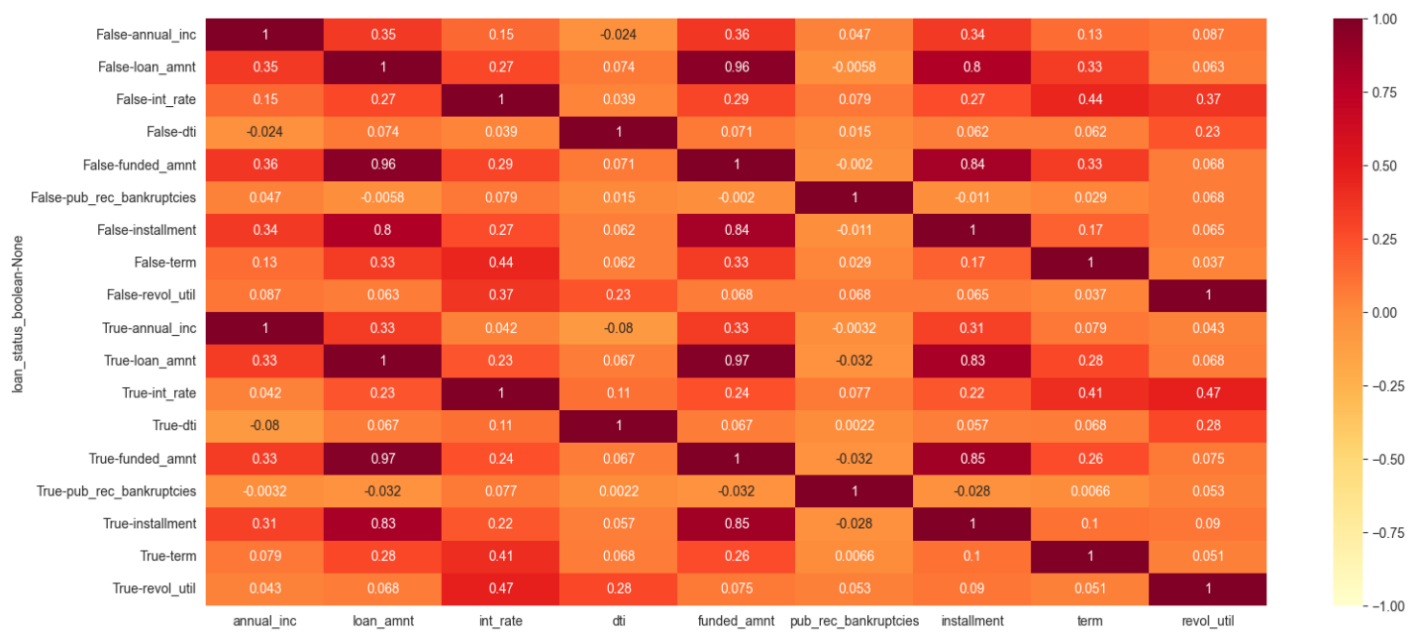
**Analysis:-** Thus, most of the 60 months tenure choosers having more chance to default

# Bivariate Analysis



**Outcome-** Maximum loan applicants are between 25 to 72.

# Bivariate Analysis



**Outcome:** More-darker the point with False, more the risk

# Final Summary

Upon analysing on all the points, We have derived some points which Lending Club can follow to minimize bad loans:

- Giving loan to the applicants having annual income higher than 40K and low interest rate (below 15%) may minimize the bad loan.
- Smaller term (36 months) can minimize the bad loans.
- Loan grading has a significant impact on loans.
- Loan can be provided who has own house.
- Purpose like small business and another loan repayment has most numbers of defaulters.
- Public bankruptcy record must be 0 or 1. Higher the record value, higher the risk of bad loan.
- Address Cities has lower impact on loan status as well.
- High dti(15+) leads to more bad loans thus loan funding should be made on annual income.
- Month and year has lower impact.
- Credit Utilization Ratio (revol\_util) has a big significance. High utilization leads to bad loan. Thus, Lending Club should provide loans to lower utilization rate applicants.