

# LENDING CLUB CASE STUDY

## SUBMISSION

### **SUBMITTED BY**

ARUN V S

ABHINAV ABHIPRIYA

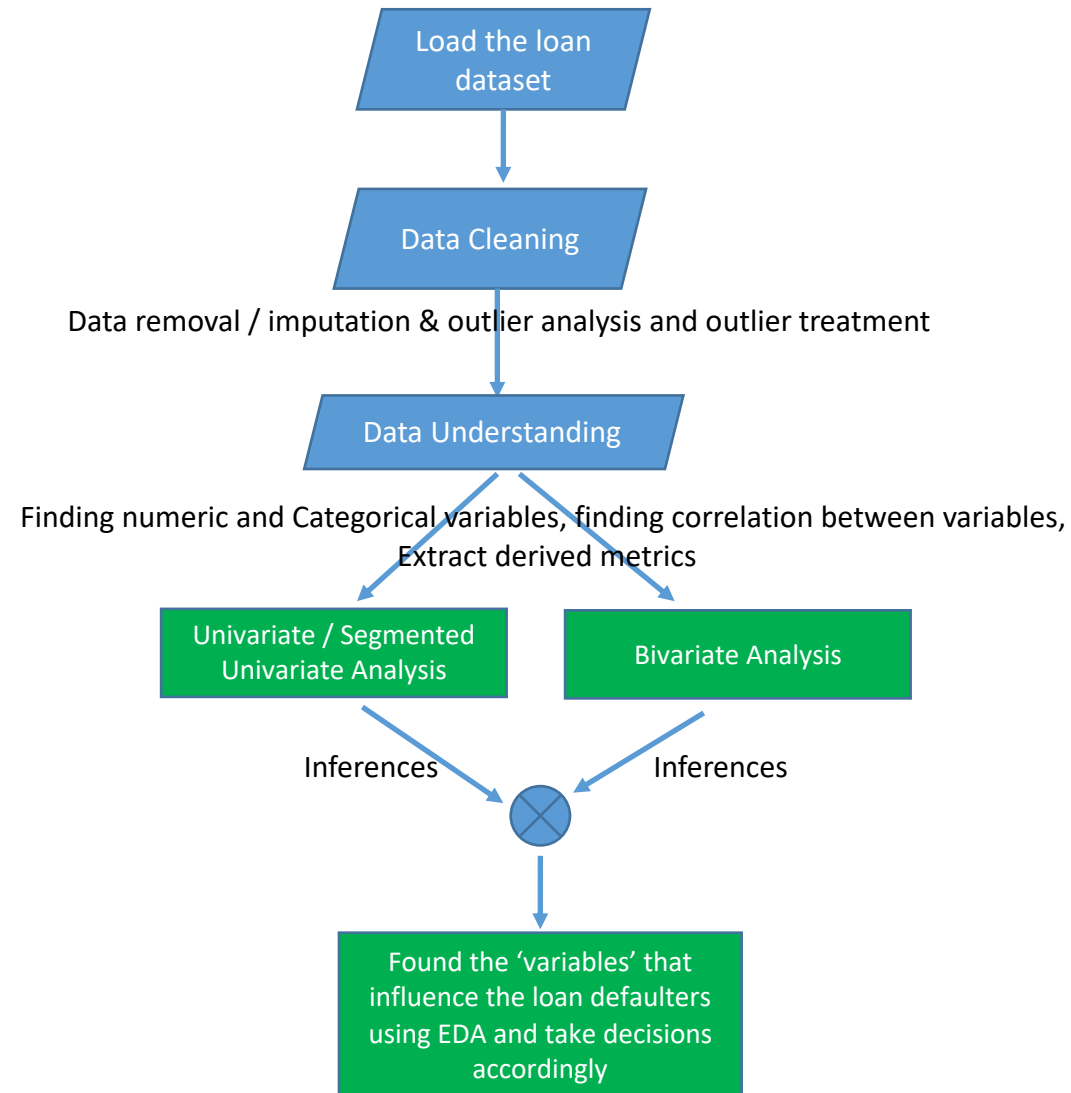
ARUN SAPPANI PRAKASAM

MANJUNATHA GOPALAIAH

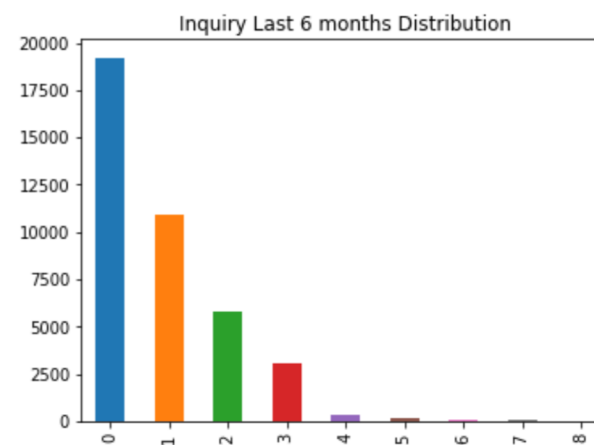
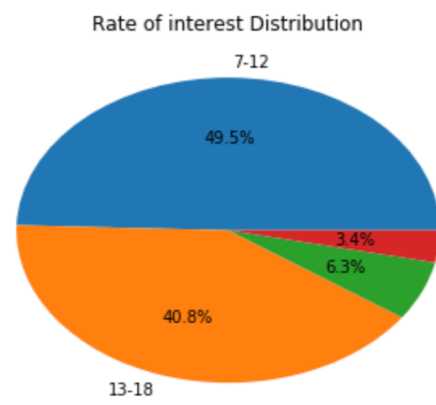
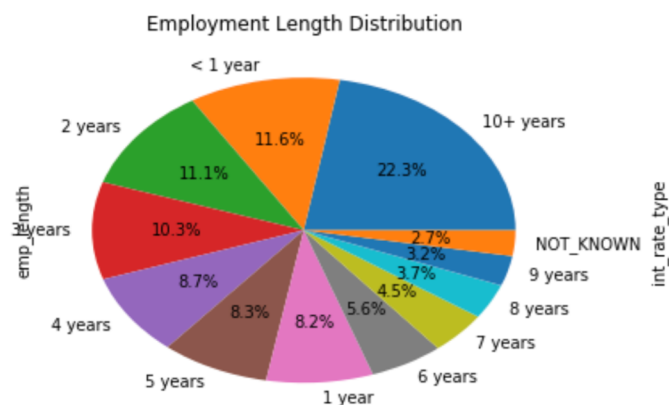
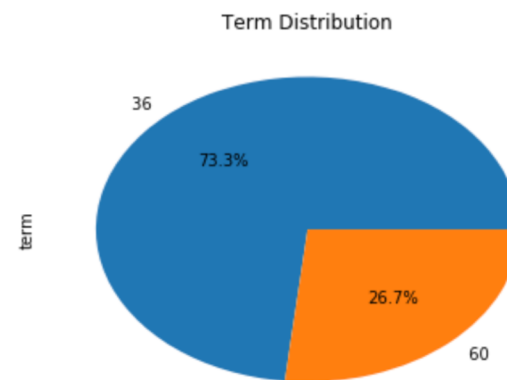
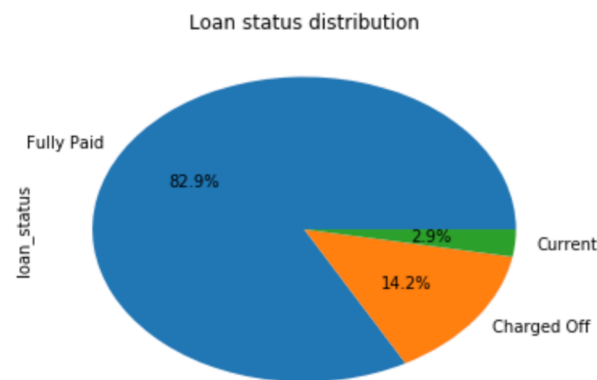
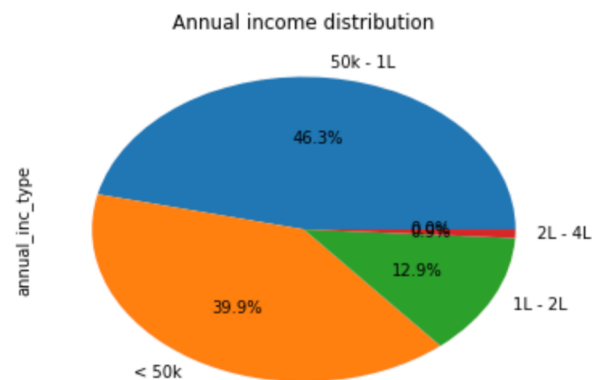
## Abstract

- Our team work for a **consumer finance company** which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval 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. he/she is likely to default, then approving the loan may lead to a **financial loss** for the company
- Assuming working in a consumer finance company, we need to analyze the data using various strategies like Exploratory Data Analysis and take decisions by finding the persons who are likely to pay loan and those who do not.

# Problem Solving Methodology



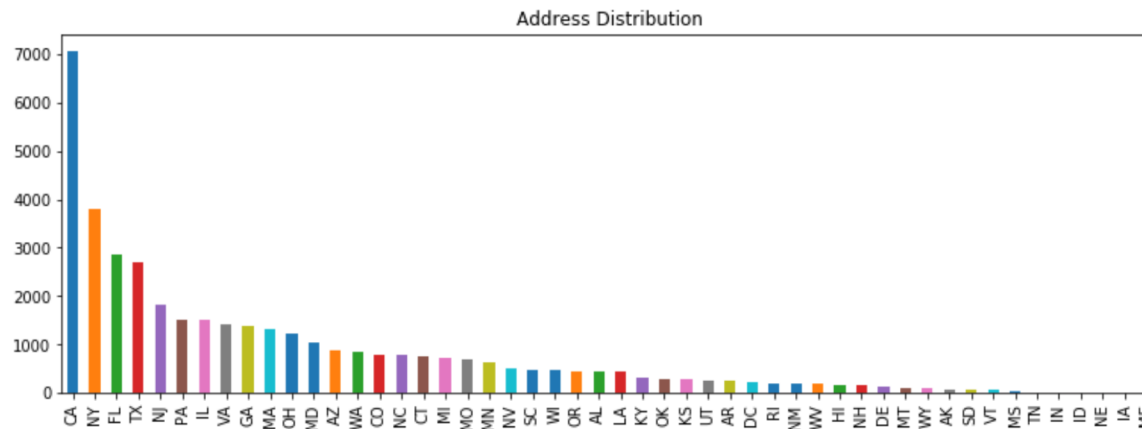
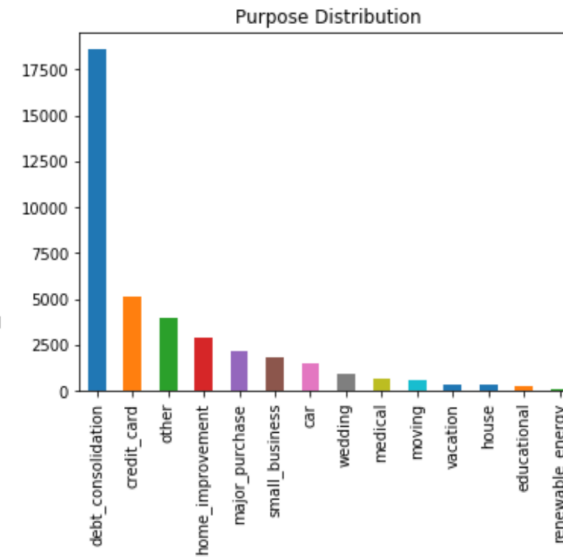
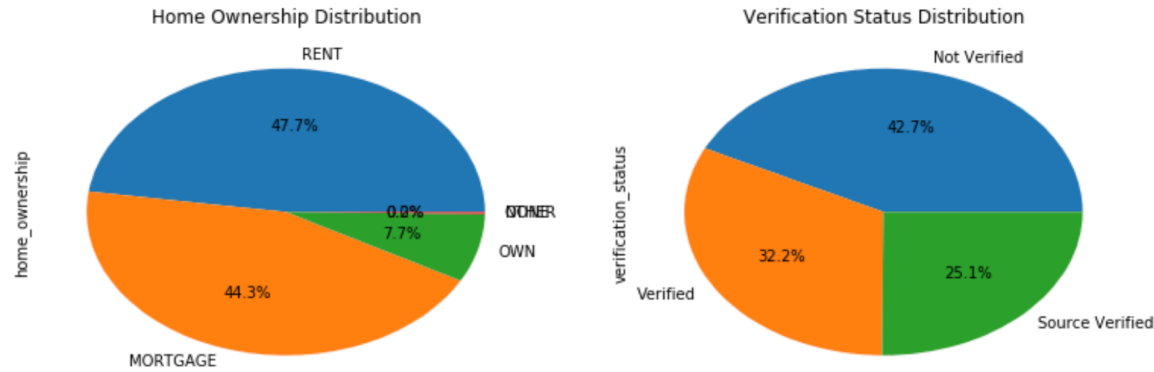
# Data Understanding



## Inferences

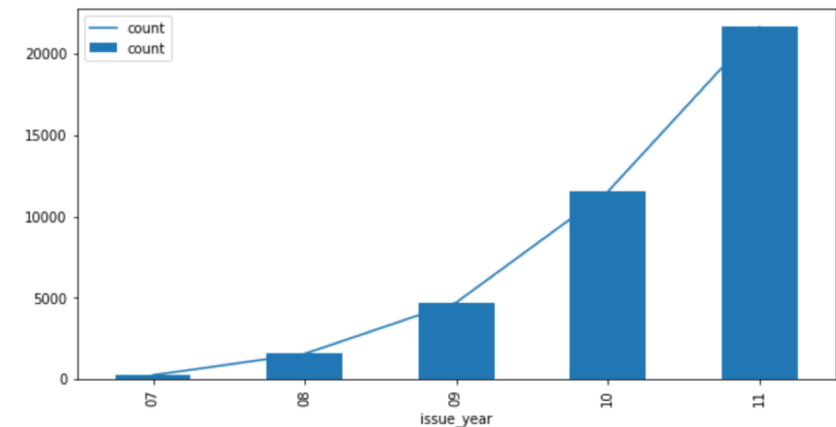
- 14.2% persons are labelled as 'Charged Off'
- Most people opted for 36 years as their loan term
- People with Annual income below 2 lakhs have taken more loans
- People with Employment Length > 10 years have taken more loan
- People have 7%-12% of rate of interest have opted more loans
- Most people who opted for loan didnt make any enquiries or only less number of enquiries.

# Data Understanding (Contd.)



## Inferences

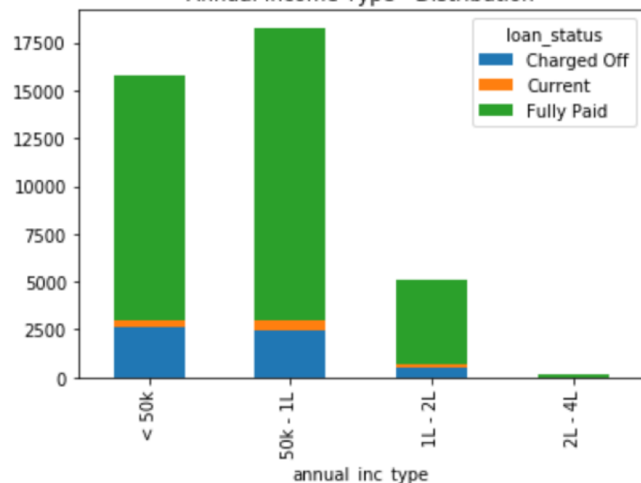
- People who has RENT or MORTGAGE takes loan
- Maximum loans were provided without verification
- People get loans more for the category 'debt\_consolidation'
- People in 'CA' have taken more loans.
- There is a steady increase in the issuance of loan as the year increases



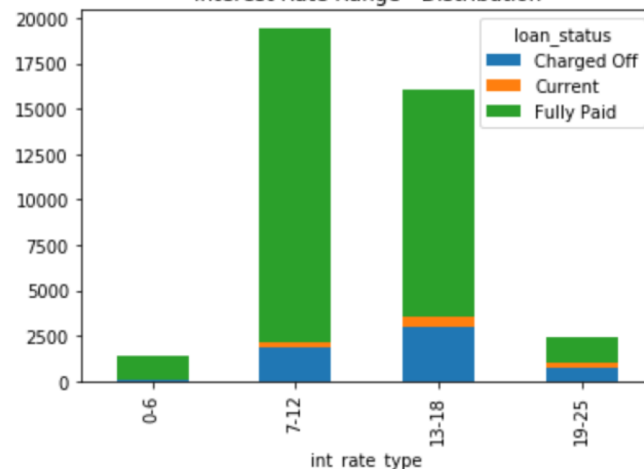


# Univariate Analysis

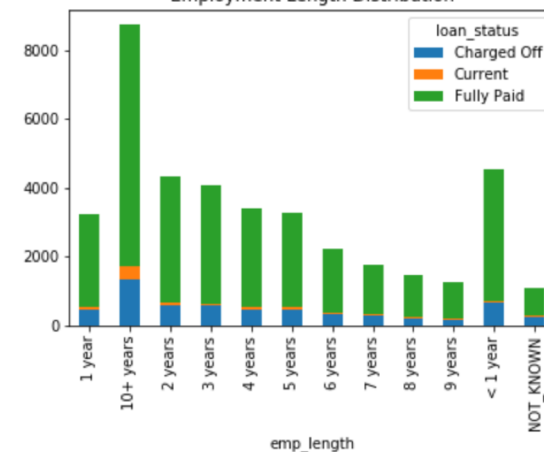
Annual Income Type - Distribution



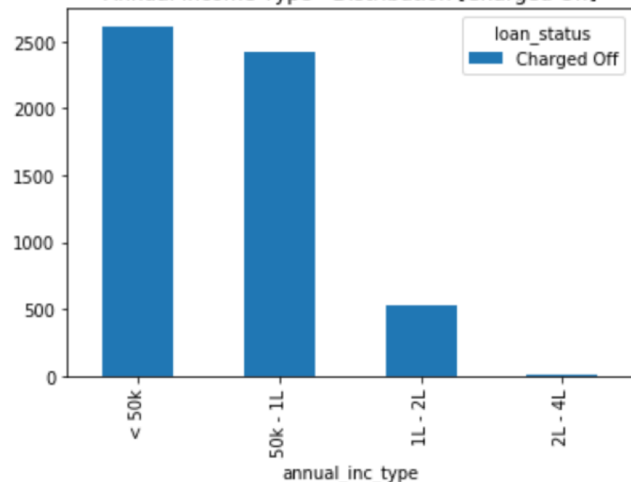
Interest Rate Range - Distribution



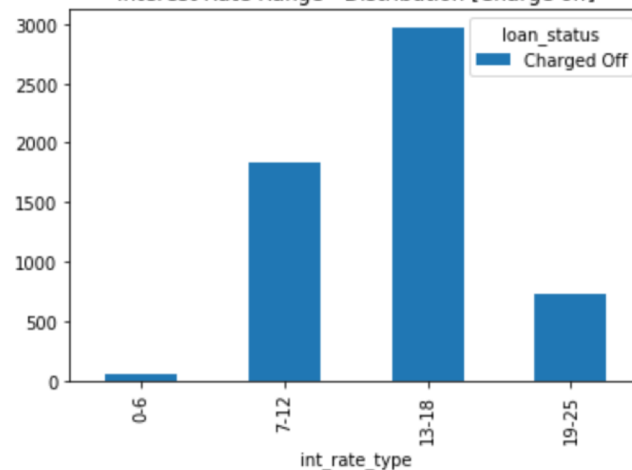
Employment Length Distribution



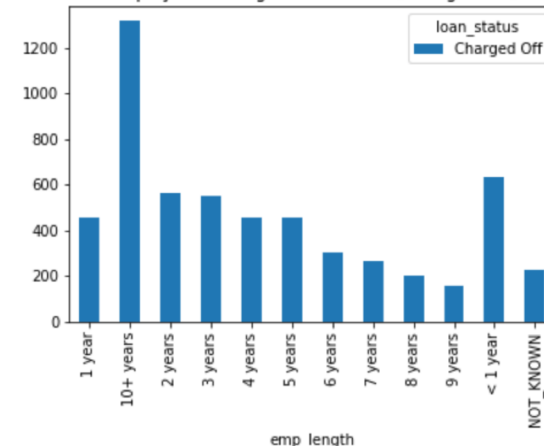
Annual Income Type - Distribution [Charged Off]



Interest Rate Range - Distribution [Charge off]



Employment Length Distribution [Charged Off]

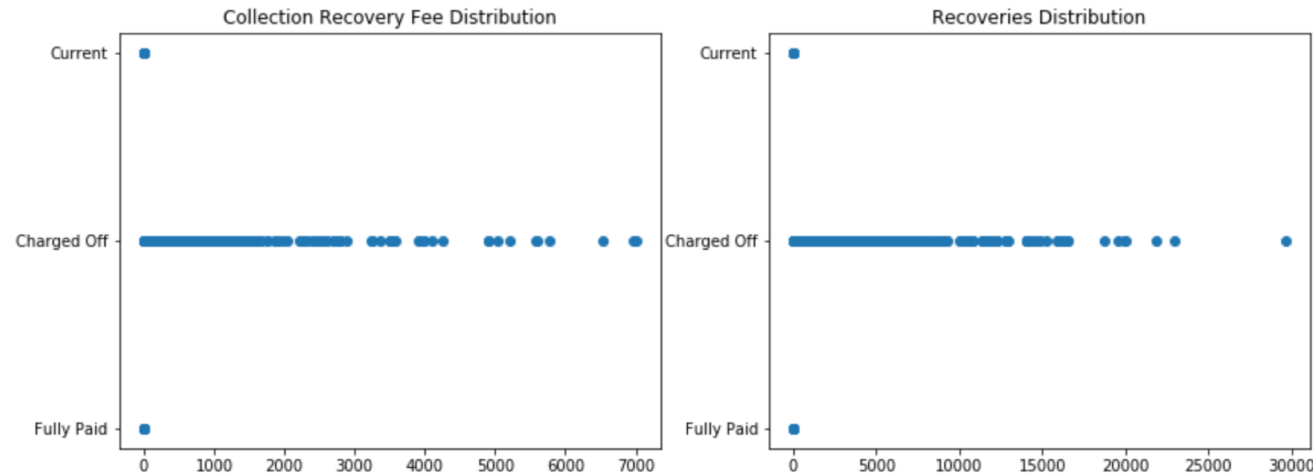


**Annual Income Type :** It is clear that People who have 50k-1L of annual income have opted more loans. But interestingly, people who have annual income <50k have more loan defaulters than people who earns 50k-1L(majority).

**Interest Range :** It is clear that People have 7%-12% of rate of interest have opted more loans. But interestingly, people with 13%-18% of rate of interest have more defaulters (i.e Charge off ratio is more in the range of 13-18%)

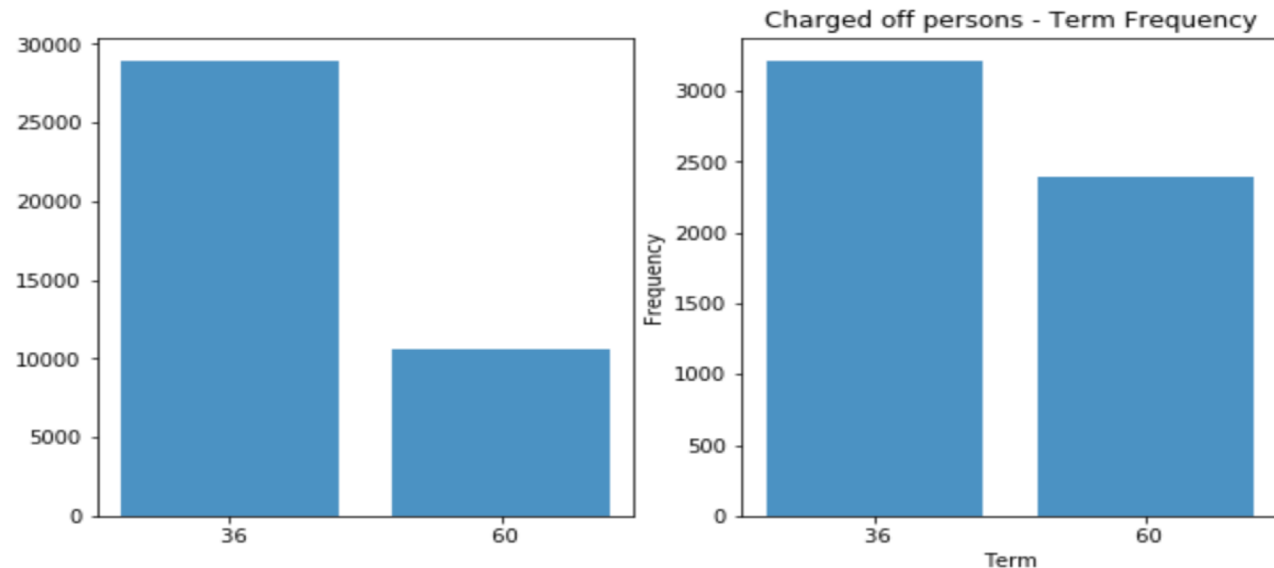
- **Employment Length:** It is evident from the graph that, probability of the frequency of the people who charged off are those who are less than a year experienced
- **With respect to the ratio of "ChargeOff/Total", person whose employment <1 year tends to be a defaulter**

# Univariate Analysis (Contd.)



## Collection Recovery Fee & Recoveries :

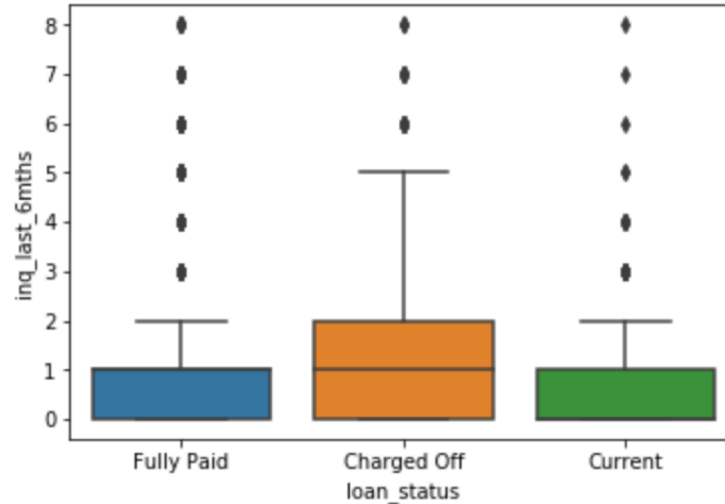
- It is very obvious that the Charged off persons will be charged with Collection Recovery Fee and Recoveries.
- So, it is evident that these 2 columns strongly contribute to the defaulters



## Term:

- From the graph - 'Charged off persons - Term Frequency', In both 36 months and 60 months, persons who are labelled as 'Charged Off' are available.
- On comparing the term ratio (Number of persons who opt for loan in 36 months to 60 months) for the entire dataset with the charged off dataset is 2.739 and 1.344. **Both the ratio are not almost equal and that gives us the interesting pattern.**
- Visually we can see from two graphs that, at 60 months, there is a raise in the second graph. It implies, defaulters are available in both terms.
- **But for the term 60 months, more defaulters are found even though the entire dataset has less number of persons who opted for 60 months**

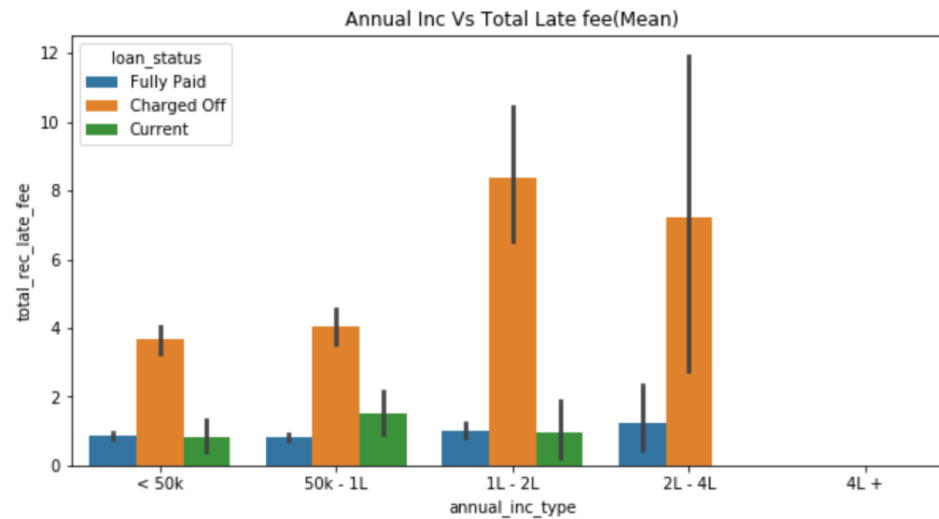
## Univariate Analysis (Contd.)



### Inquiry Last 6 months:

- It is evident from the graph that, frequency of the people who enquired 2 times or more in the last 6 months, tend to default more.

## Bivariate Analysis

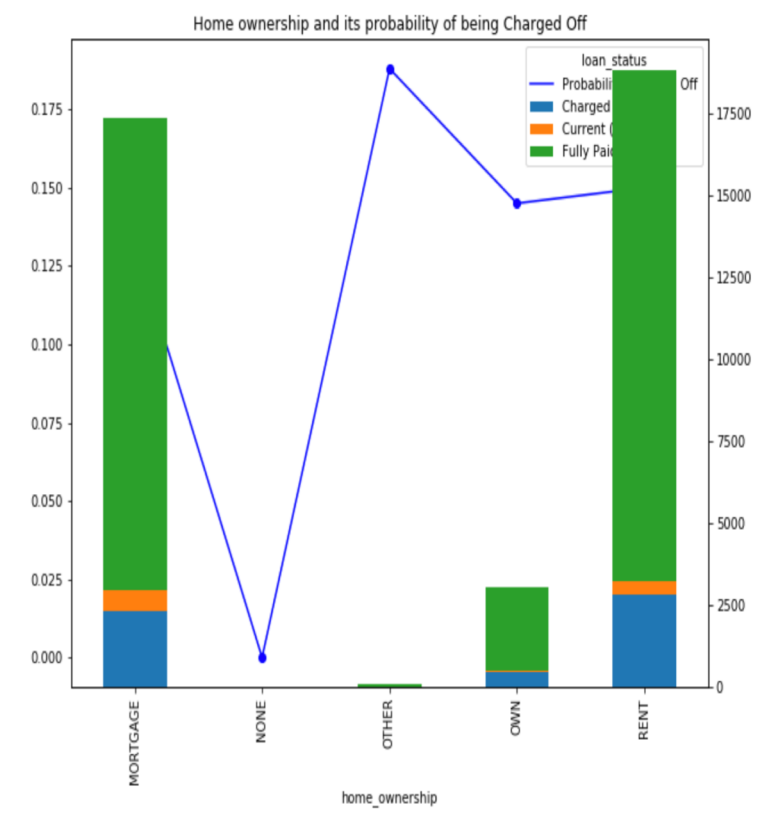
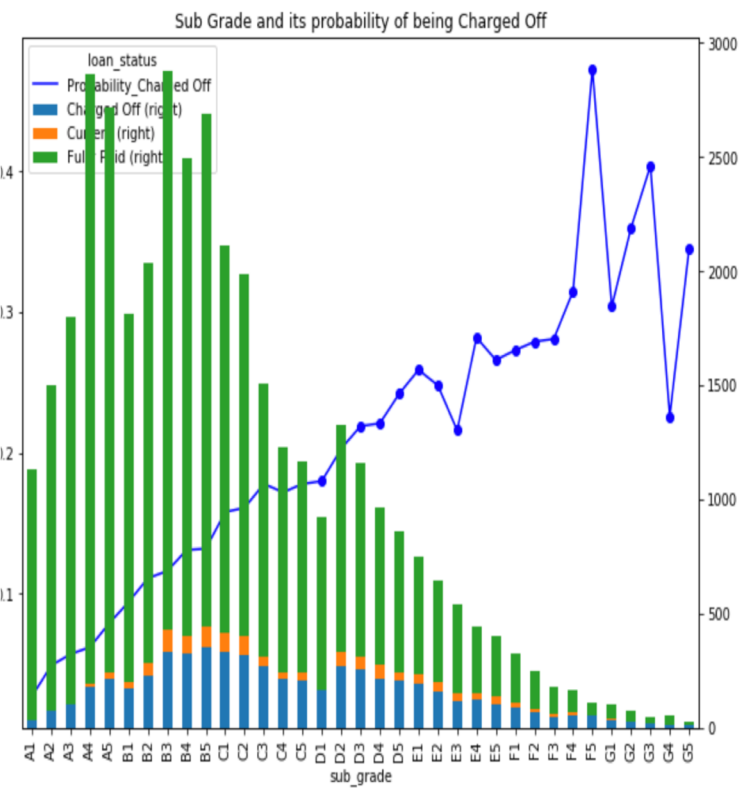
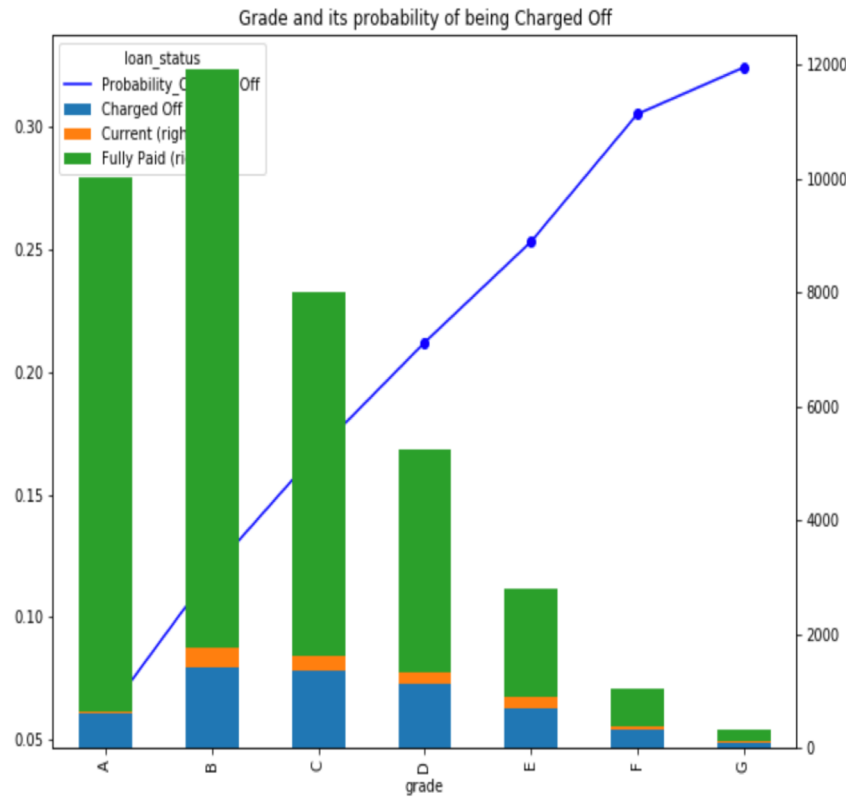


### Annual Income – Total Rec Late Fee:

- From the graph on 'Annual Inc Vs Total Late fee (Mean)', it is known that people whose income falls between 1 Lakh - 4 Lakh and those who avail late fee tends to be a defaulter more



# Bivariate Analysis (Contd.)



## Grade / Sub Grade :

- As the Grade increases, the probability of charge off is also increased. At Grades 'G' and 'F', we can observe more defaulters.

## Home Ownership with its Probability:

- People who stated 'OTHER' as their Home ownership tend to charge off more.
- Next to 'OTHER', people who stated 'RENT' as their ownership tend to charge off more

## Conclusions

- Based on the insights arrived by EDA on Loan data set, one should be cautious to the people to lend loan who fall under the multiple category below mentioned as they might tend to 'Charge Off'
  - Grades 'G' or 'F'
  - Annual Income < 50000
  - Rate of Interest between 8% to 13%
  - Employment Length < 1 year
  - Home Ownership – 'Other' or 'Rent'
  - Term – 60 months
  - Annual income between 1Lakh and 4 Lakh with Late fee
  - More collection recovery fee and collections