

# Lending Club Case Study

Presented By -

Ashutosh Pandey & Rajendra Singh Rao

Batch - AI / ML - C36

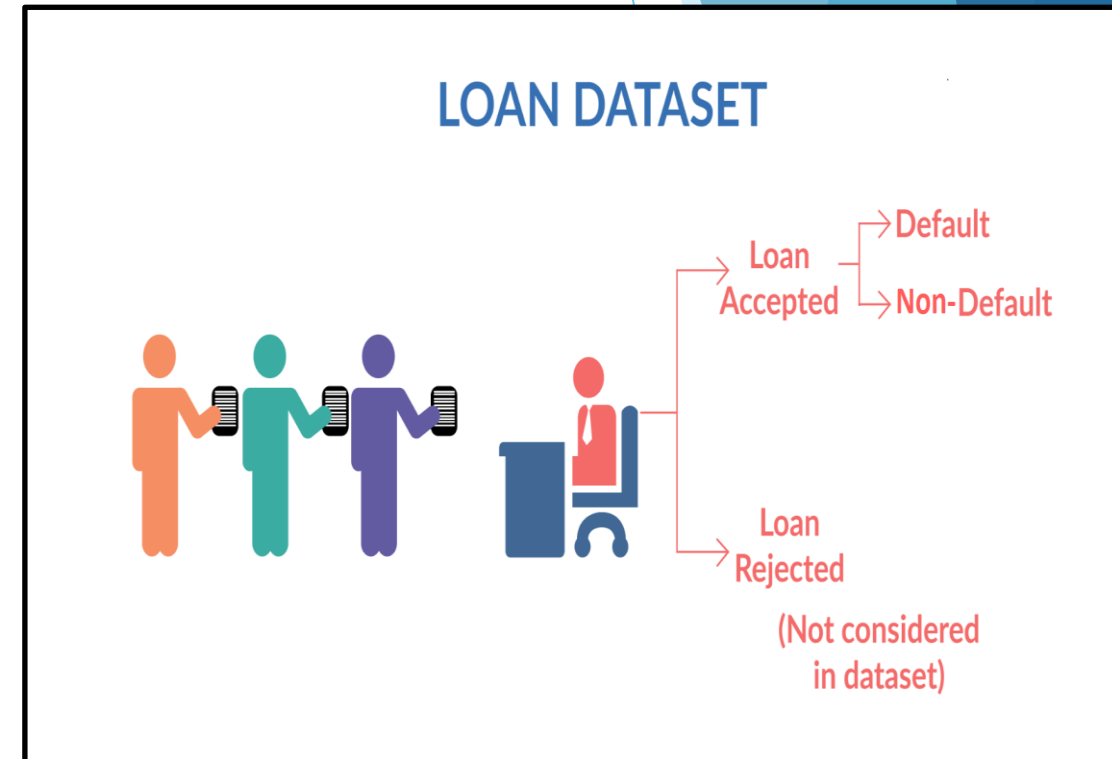
08- Feb- 2022

# Business Understanding

- ▶ This case study is about a **consumer finance company** which specialises 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

# Business Understanding

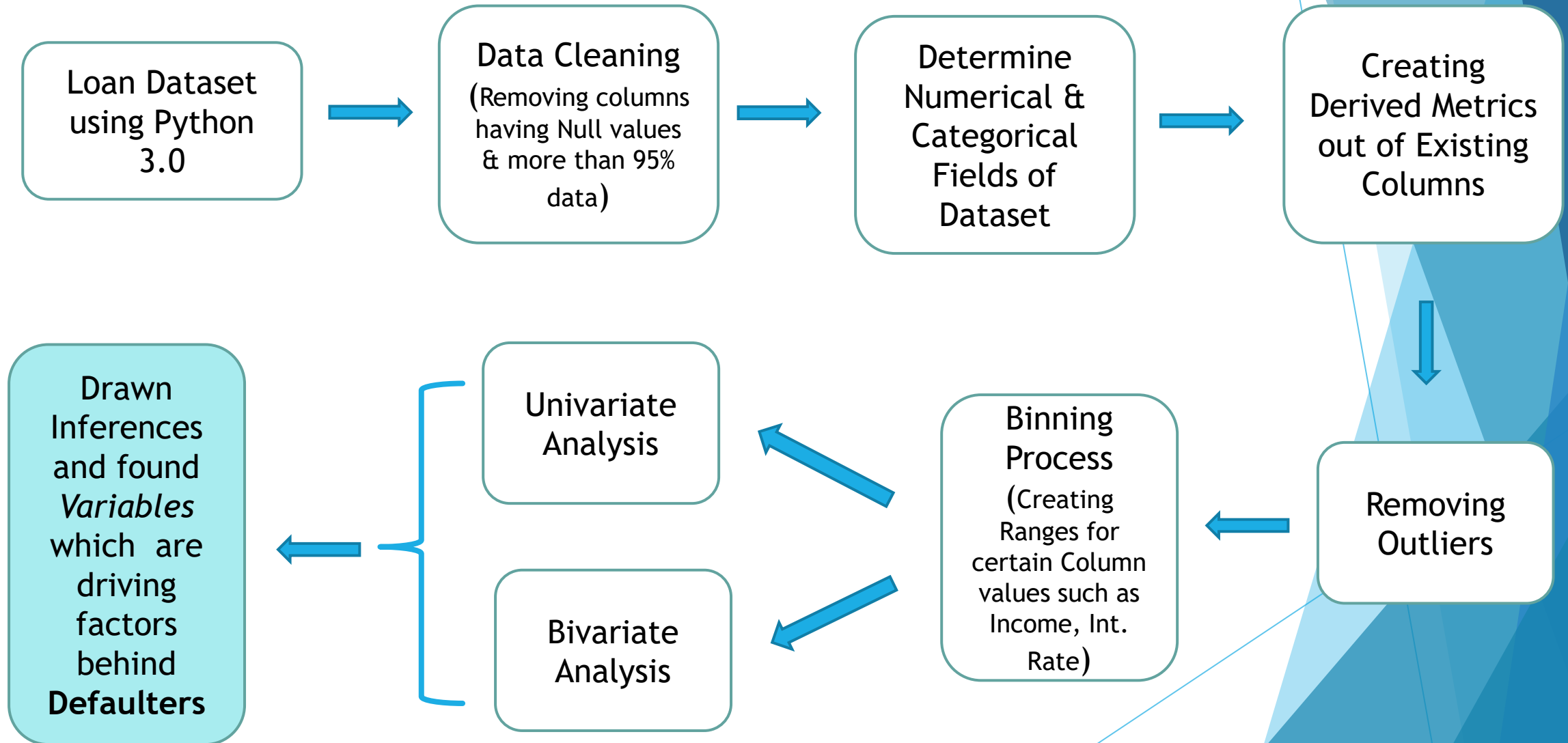
- ▶ In this case study, we have used Exploratory Data Analysis (EDA) to understand how **consumer attributes** and **loan attributes influence** the tendency of default.
- ▶ When a person applies for a loan, there are **two types of decisions** that could be taken by the company:
- ▶ **Loan accepted**: If the company approves the loan, there are 3 possible scenarios described below:
  - ▶ **Fully paid**: Applicant has fully paid the loan (the principal and the interest rate)
  - ▶ **Current**: Applicant is in the process of paying the installments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
  - ▶ **Charged-off**: Applicant has not paid the installments in due time for a long period of time, i.e. he/she has **defaulted** on the loan
- ▶ **Loan rejected**: The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset)



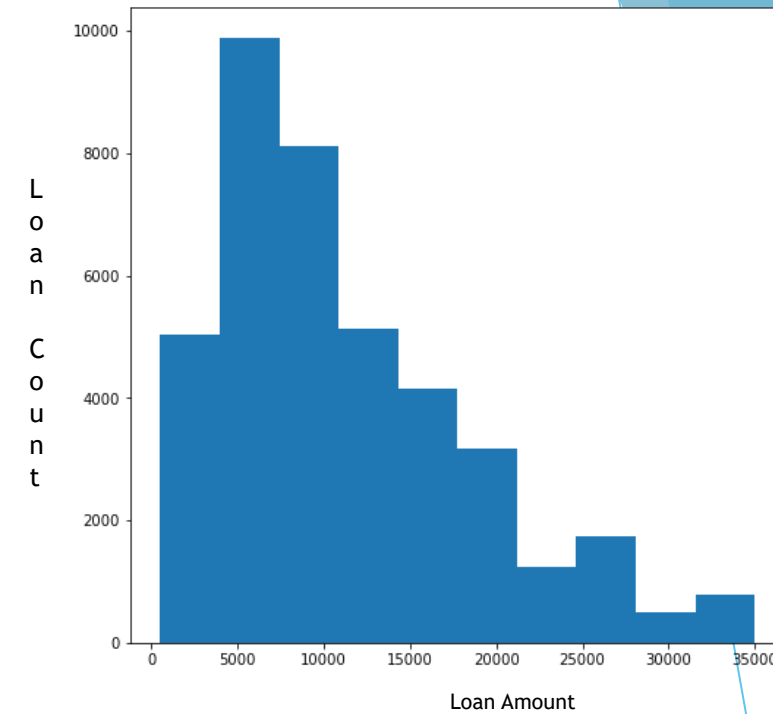
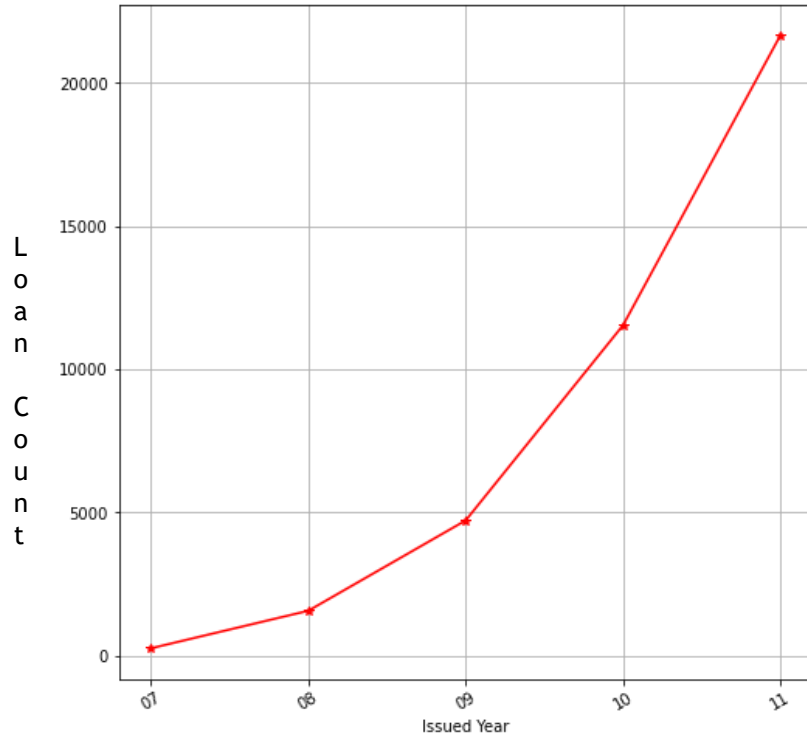
# Business Objectives

- ▶ Like most other lending companies, lending loans to 'risky' applicants is the largest source of financial loss (called credit loss).
- ▶ Borrowers who **default** cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.
- ▶ Identification of such risky loan applicants using EDA is the aim of this case study then such loans can be reduced thereby cutting down the amount of credit loss.
- ▶ 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 utilise this knowledge for its portfolio and risk assessment.

# Methodology Used

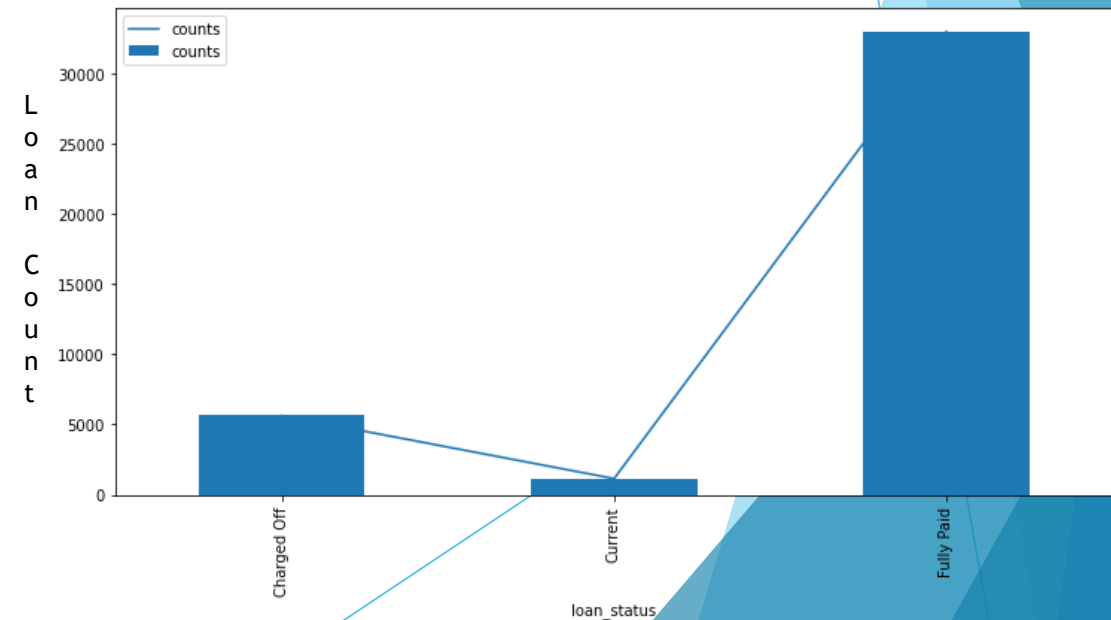


# Data Understanding



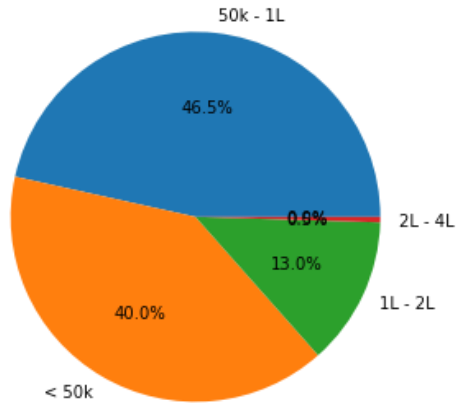
## Observations:

- ▶ Loans issued per year has been an increasing trend.
- ▶ Fully paid loans are the majority of the given dataset. After which Charged off cases and then Current loan cases.
- ▶ Loan amount between 5,000 - 10,000 is the majority.

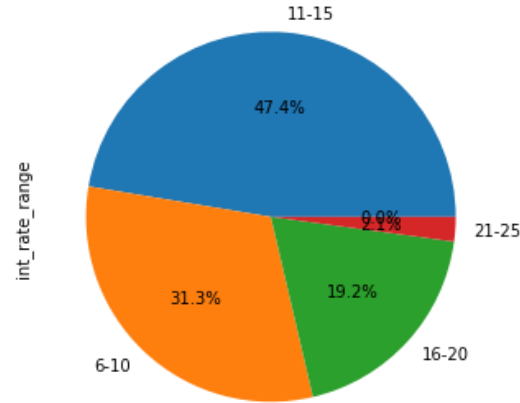


# Data Understanding

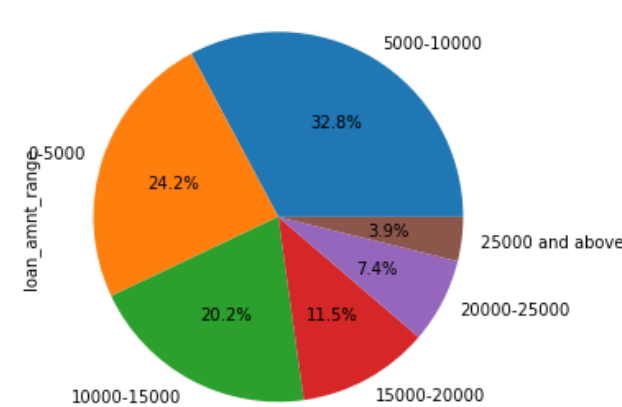
Annual income distribution



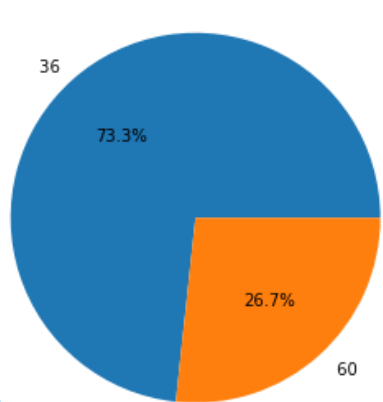
Rate of interest Distribution



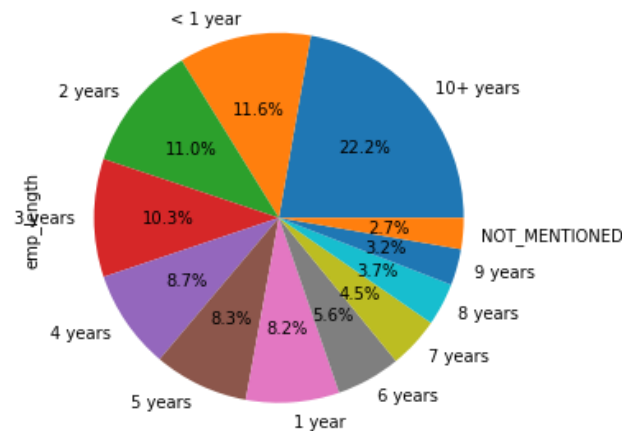
Loan Amount Distribution



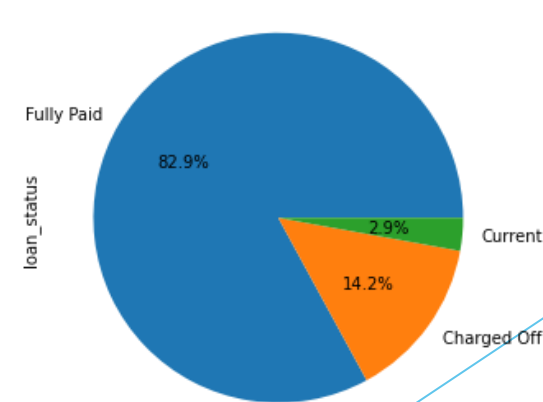
Term Distribution



Employment Length Distribution



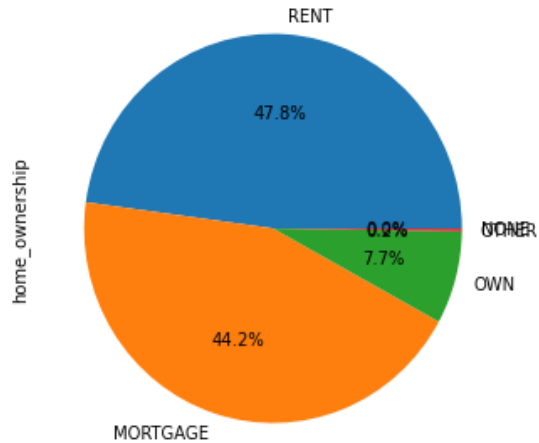
Loan Status Distribution



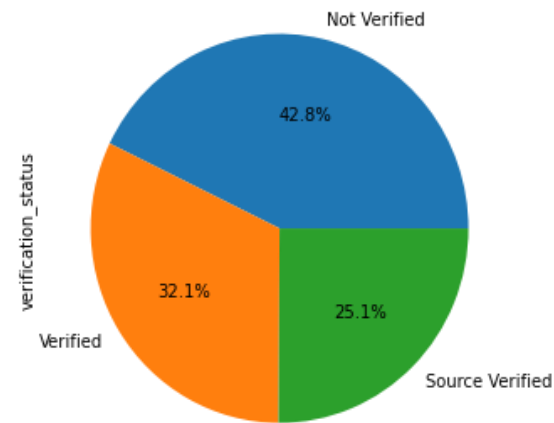
- ▶ **Observations:**
- ▶ After removing Outliers and binning the dataset into different ranges, below inferences are drawn:
- ▶ Maximum people of the dataset who have taken loan are of income under 2 Lacs i.e. about 86.5%
- ▶ 11-15 % is the interest rate offered by bank to majority of loans.
- ▶ 5,000 - 10,000 is the loan amount range offered in majority by bank.
- ▶ Most People about 73.3% have taken 36 months as term of loan.
- ▶ People with 10+ experience are in majority of total people who have opted for loan.
- ▶ In the dataset, about 83% loans are fully paid, 14% are charged-off and 3% are currently active.

# Data Understanding

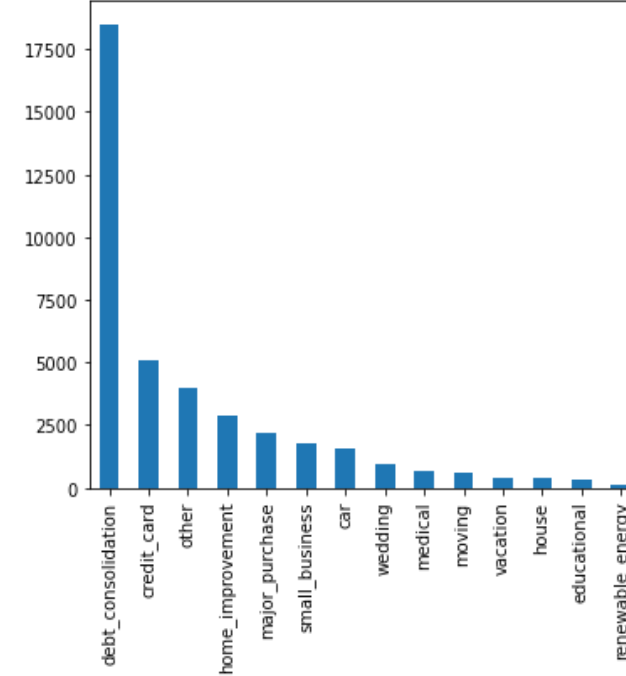
Home Ownership Distribution



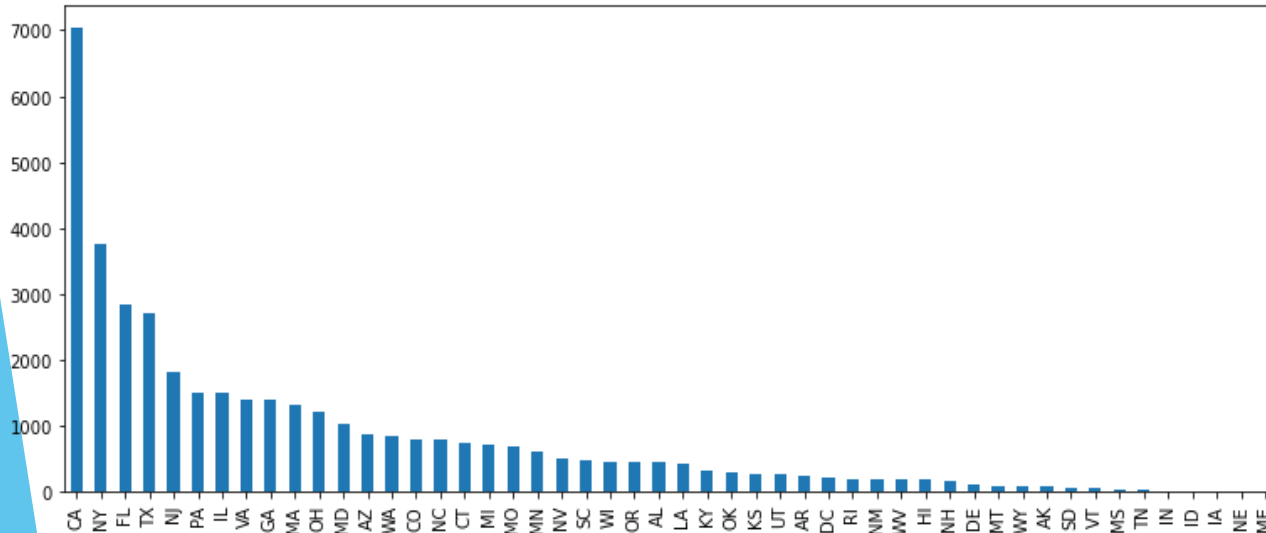
Verification Status Distribution



Purpose Distribution



Address Distribution

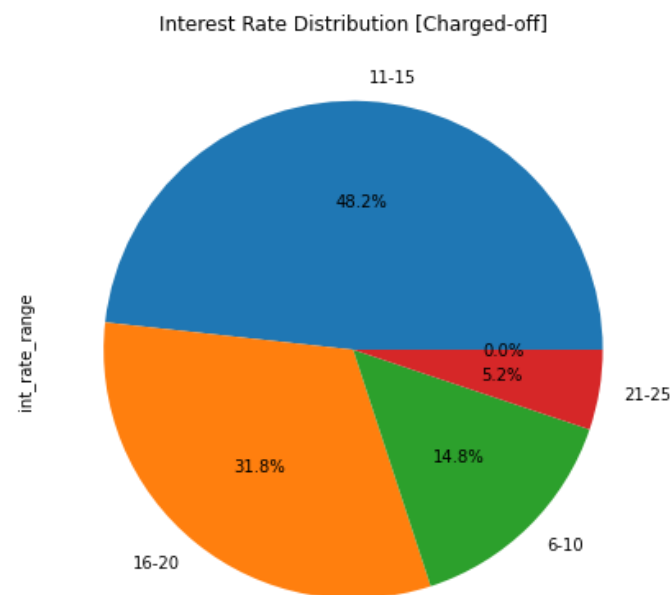
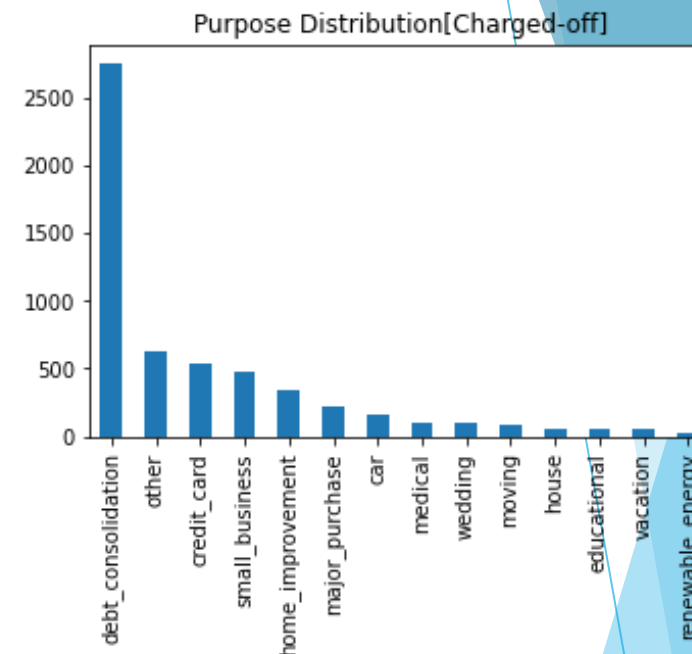
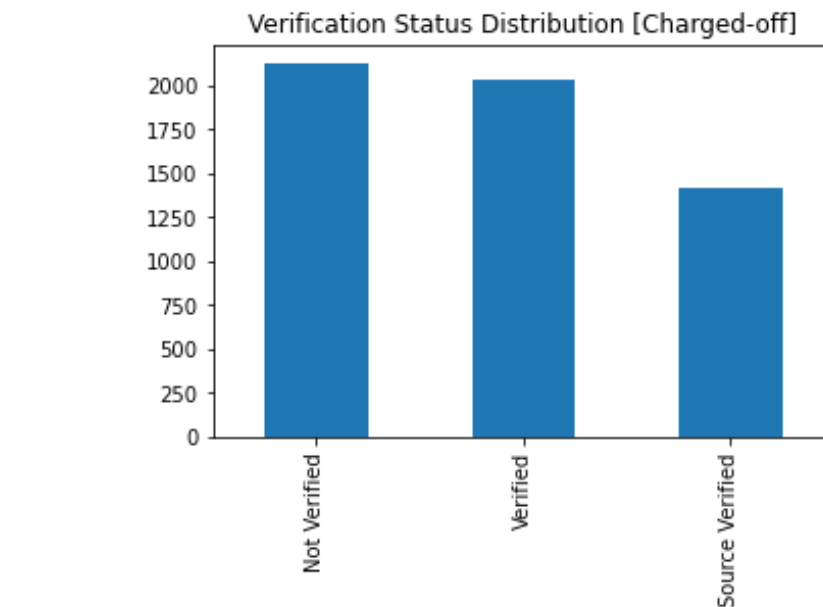
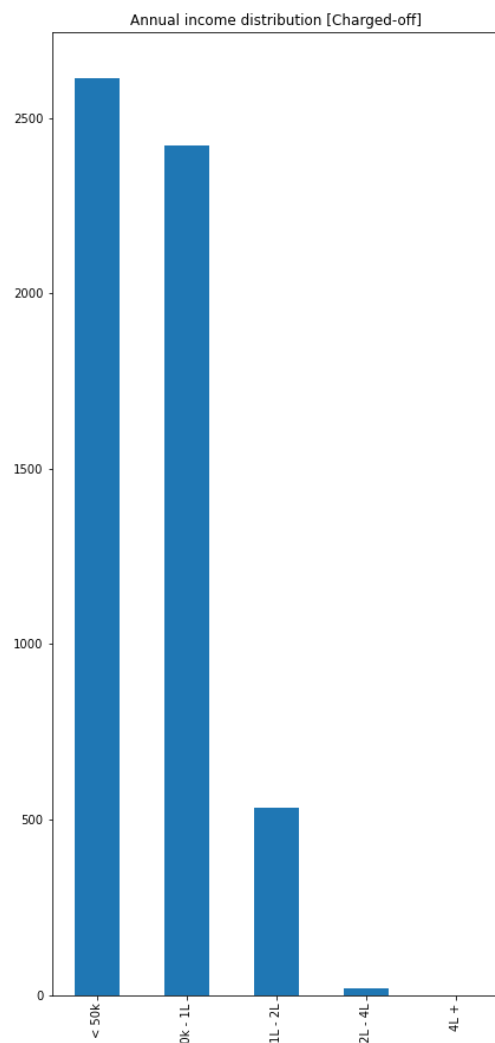


## Observation:

- ▶ People whose home ownership is either RENT or Mortgage tend to have taken more loans.
- ▶ Verification status shows that maximum loans have been given without verification i.e. close to 43%
- ▶ Maximum number are taken for debt consolidation.
- ▶ People of address 'CA' have taken most loans.



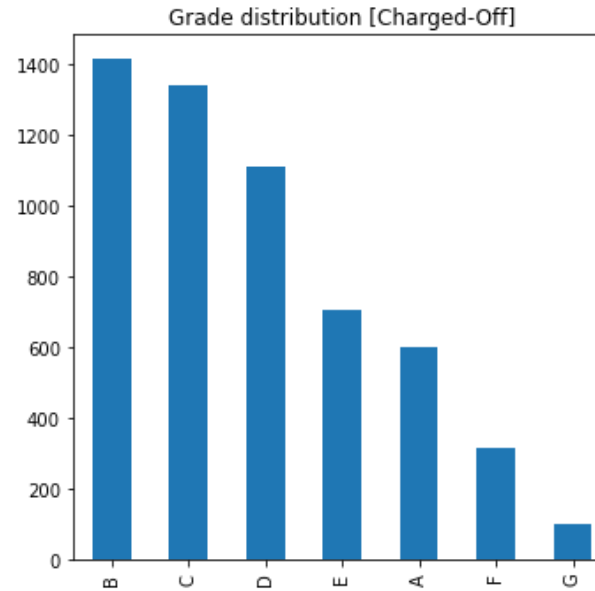
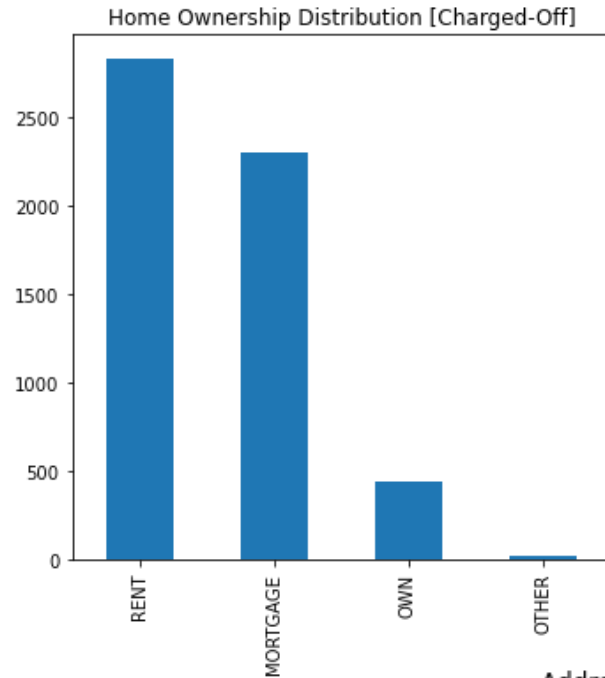
# Univariate Analysis



## Observation:

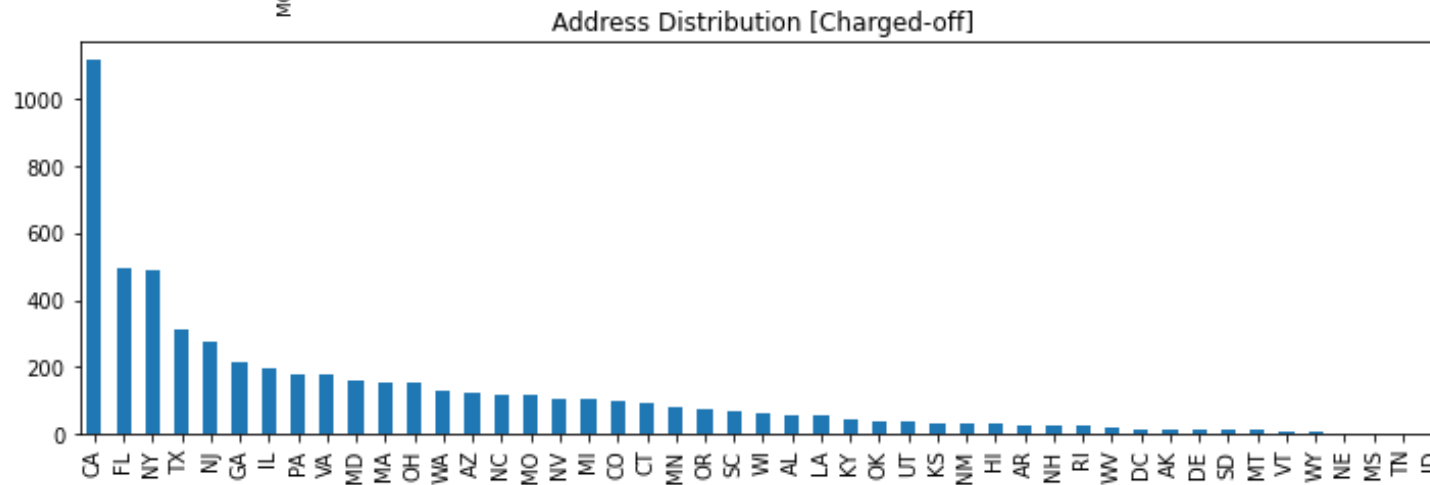
- ▶ People having income less than 50,000 are majority (about 48%) who are turning into **Default / Charged-off** loans.
- ▶ Loans with interest rate between 11-15% have mostly turned up into Default. They consist of nearly 48% of total default loan population.
- ▶ People whose ID is not verified tend to **Default** more.
- ▶ People who have shown loan purpose as **Debt Consolidation** have greater number in default loans.

# Univariate Analysis

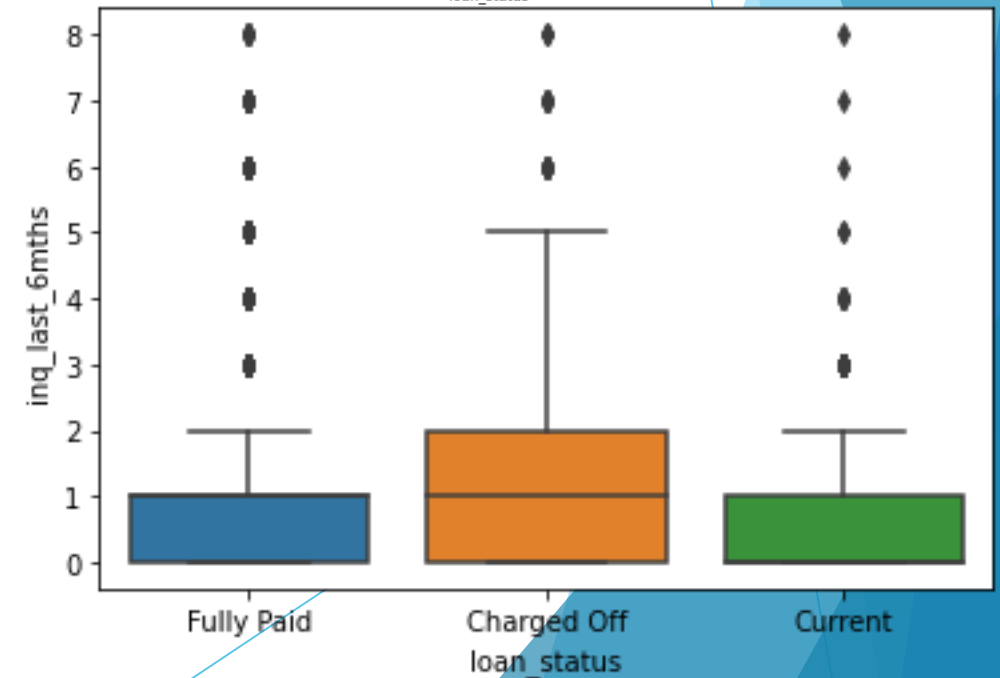
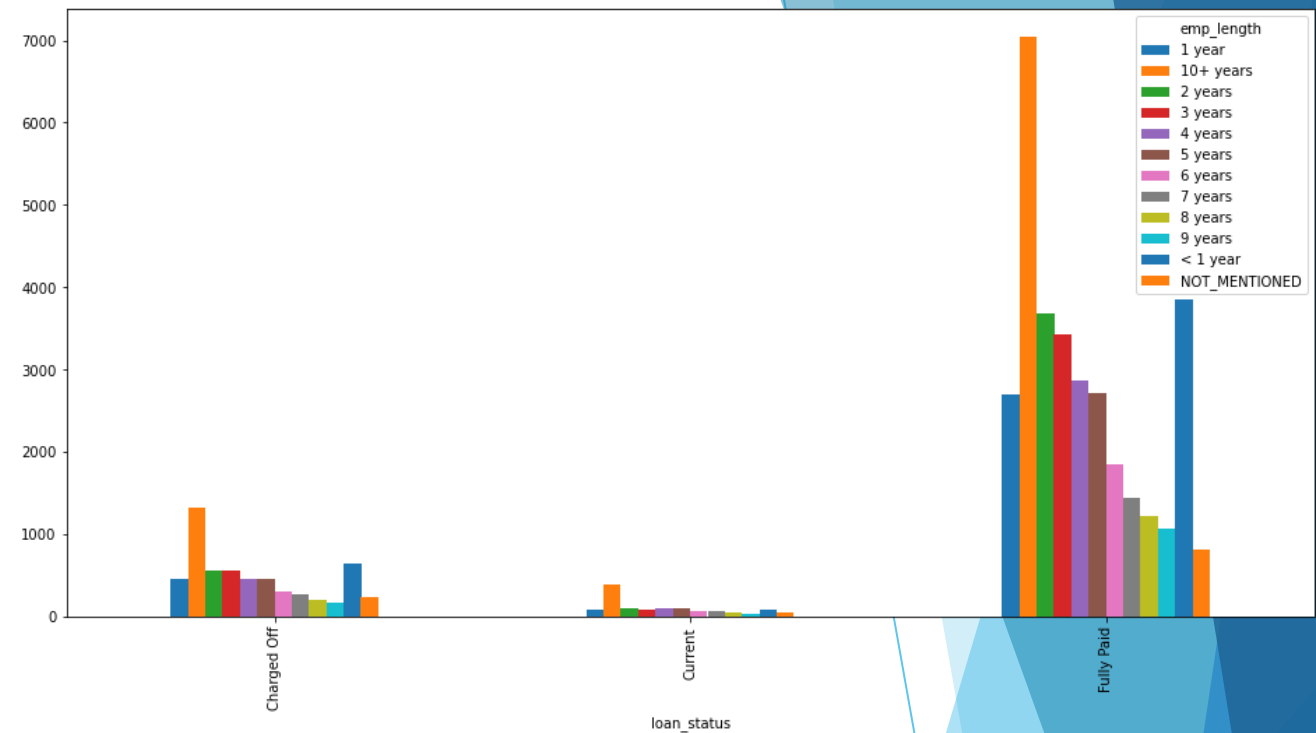
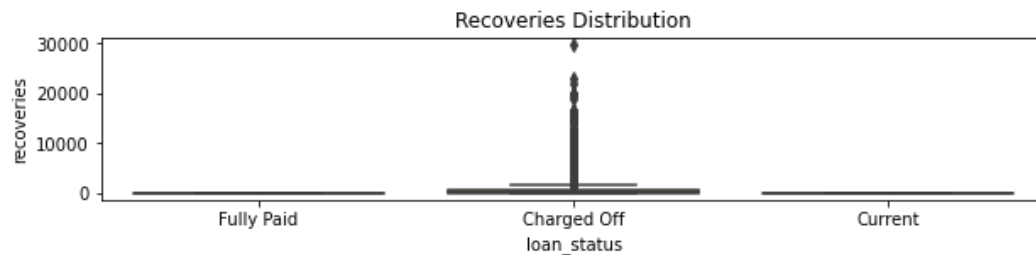
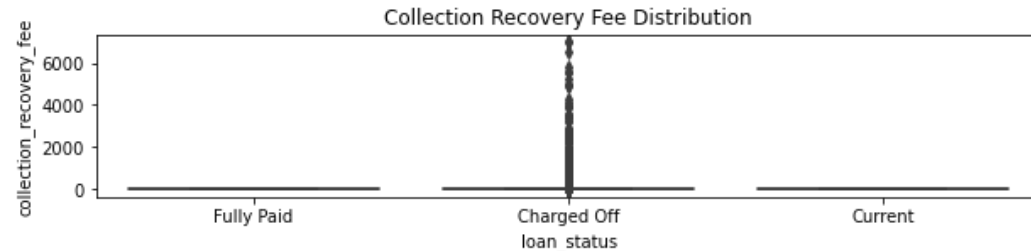


## Observation:

- ▶ People who have Home Ownership as "RENT" have majority in Charged-off loans.
- ▶ People with **grade B,C and D** have defaulted more in comparison to other grades.
- ▶ Most people from **CA address** have being defaulters.



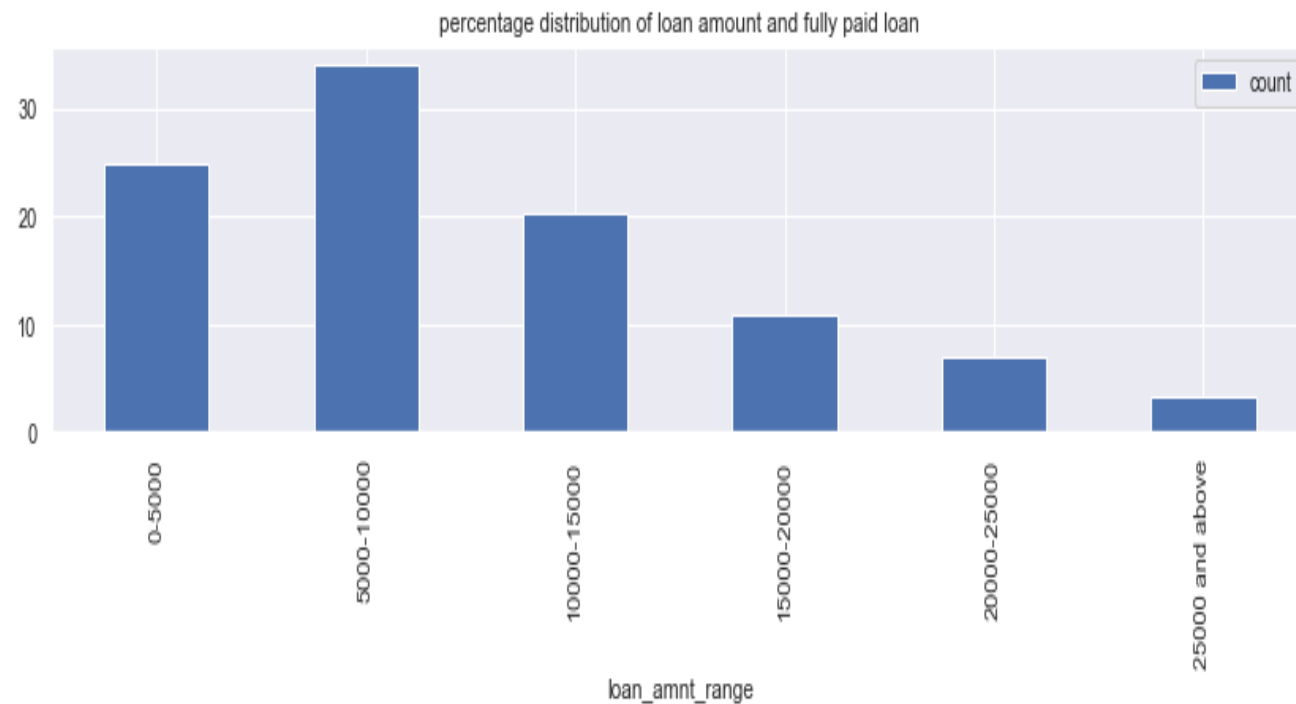
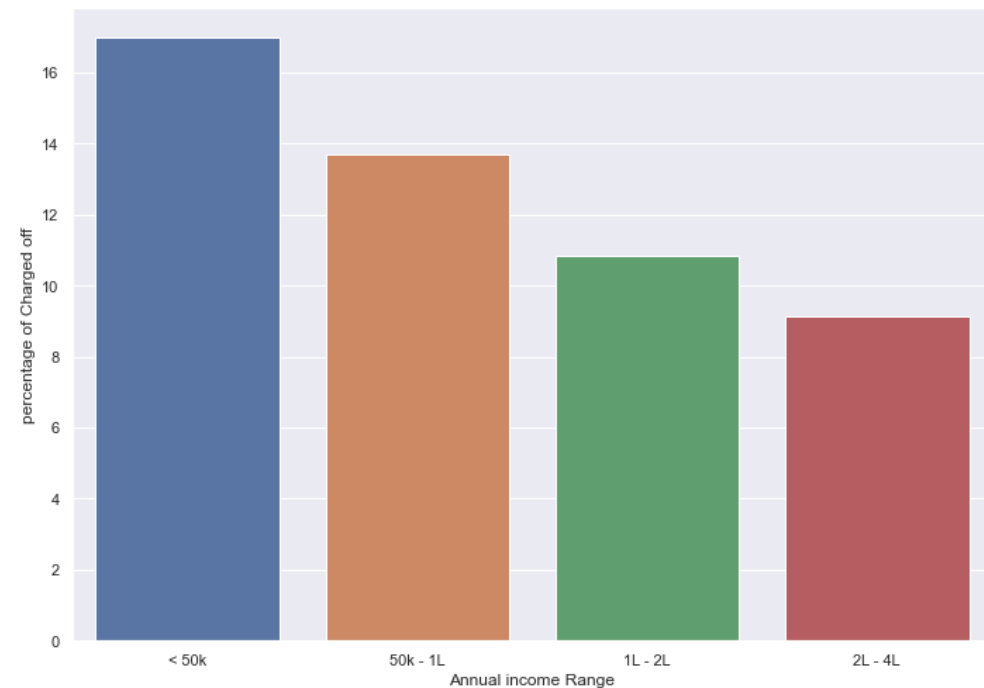
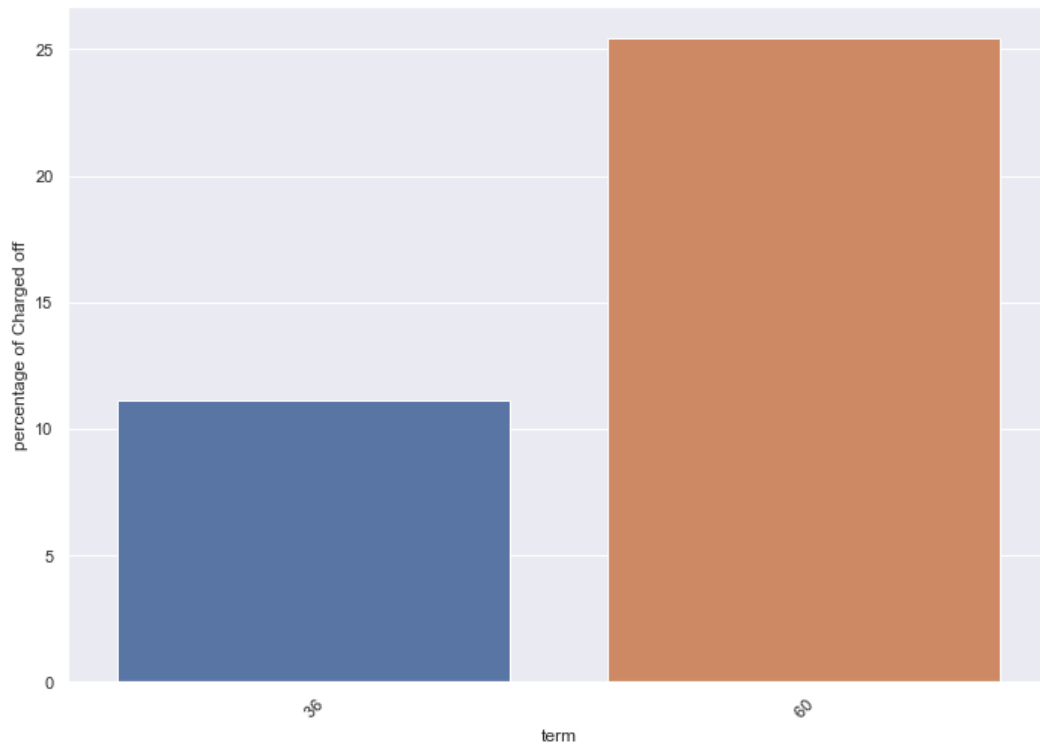
# Bivariate Analysis



## ► Observation:

- It can be inferred that Collection Recovery Fee and Recoveries Distribution has been highest for Charged-off loans.
- It can be clearly seen that people who have **inquired more than two times** in last 6 months have **defaulted** more in comparison to Fully Paid and Current loans. Most people from CA address have being defaulters.

# Bivariate Analysis



- **Observation:**
- People opted for **60 Months** have **defaulted more** in comparison to people who have taken 36 months term.
  - As **annual income increases**, **percentage of charged off loan decreases**.
  - It has be seen clearly as **loan amount increase** **number of loan taken also decrease** in both **Fully paid** and **Charged off** loan. Probably it is because small amount loan is easily taken but bigger amount loan is less required.

# Conclusions

- ▶ Based on study done by EDA on the given dataset, many useful insights have come into notice which can be utilized by Lending Company to reduce Default loans in future. The trend of people who are getting defaulted can be of one or many below categories mentioned as follows -
  - ▶ People with Annual Income less than 50,000/-
  - ▶ ID status - Not Verified
  - ▶ Employment length less than 1 year
  - ▶ Loans given with interest range between 11-15 %
  - ▶ Home Ownership - Rent or Mortgage
  - ▶ Term of loan - 60 Months
  - ▶ More collection recovery fees
  - ▶ Grades B, C and D
  - ▶ Purpose Distribution - Debt Consolidation