

# LENDING CLUB CASE STUDY

## Using EDA

Group Members:

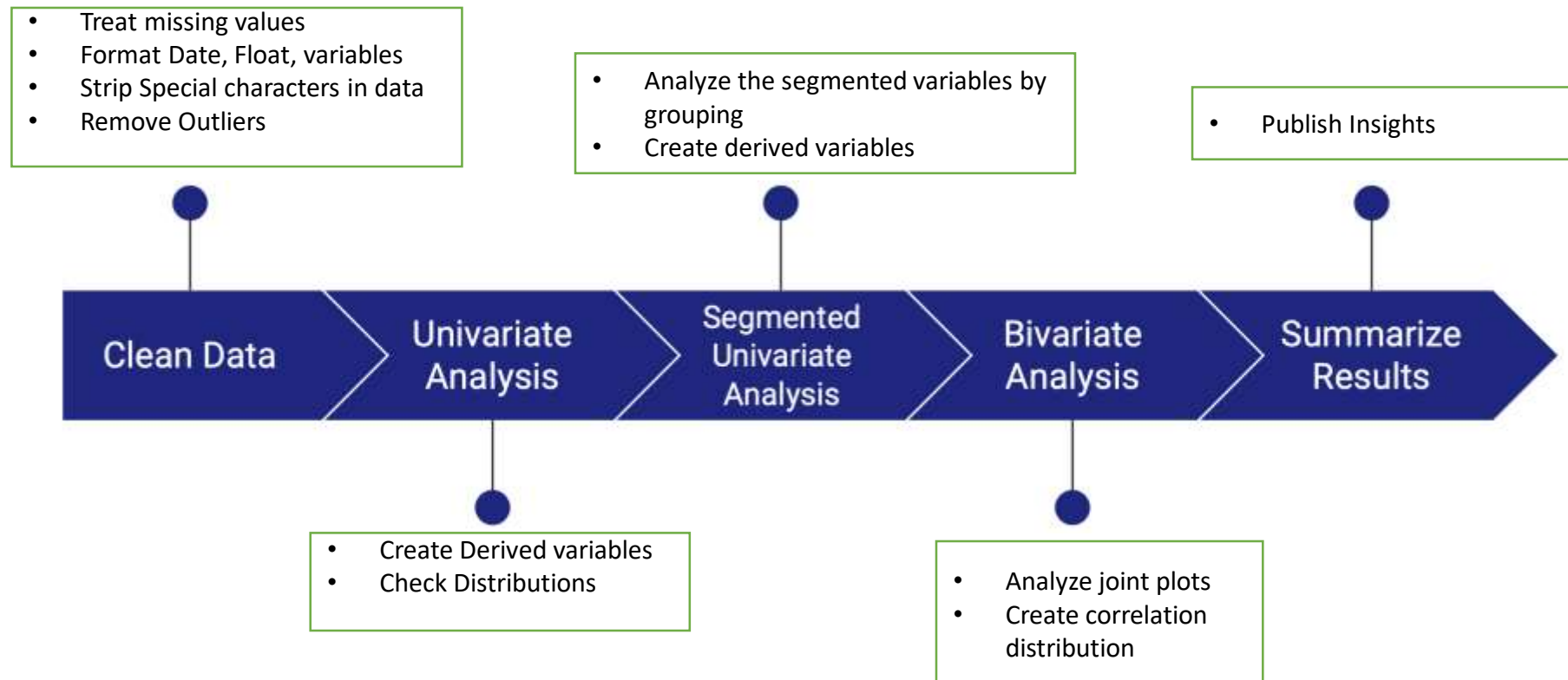
Vijay Anand Rajavel([vijay.onlyforonline@gmail.com](mailto:vijay.onlyforonline@gmail.com))

Ravneet Singh ([ravneetsingh1986@gmail.com](mailto:ravneetsingh1986@gmail.com))

## Business Objective

- 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.
- 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.
- 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.

# Problem Solving Methodology



# DATA UNDERSTANDING

## □ BUSINESS UNDERSTANDING:

- This includes the business study and constraints that are already defined in slide 2.
- Goals of data analysis are defined crisply.

## □ DATA UNDERSTANDING:

- This includes the study of data files which are “loan” & “data dictionary”.
- Steps:
  - Loading File : Read loan.csv and apply encoding techniques.
  - Check columns data types and Shape.
  - Check and get Numeric data for needed columns.
  - Check for missing values.
  - Identify the target column.

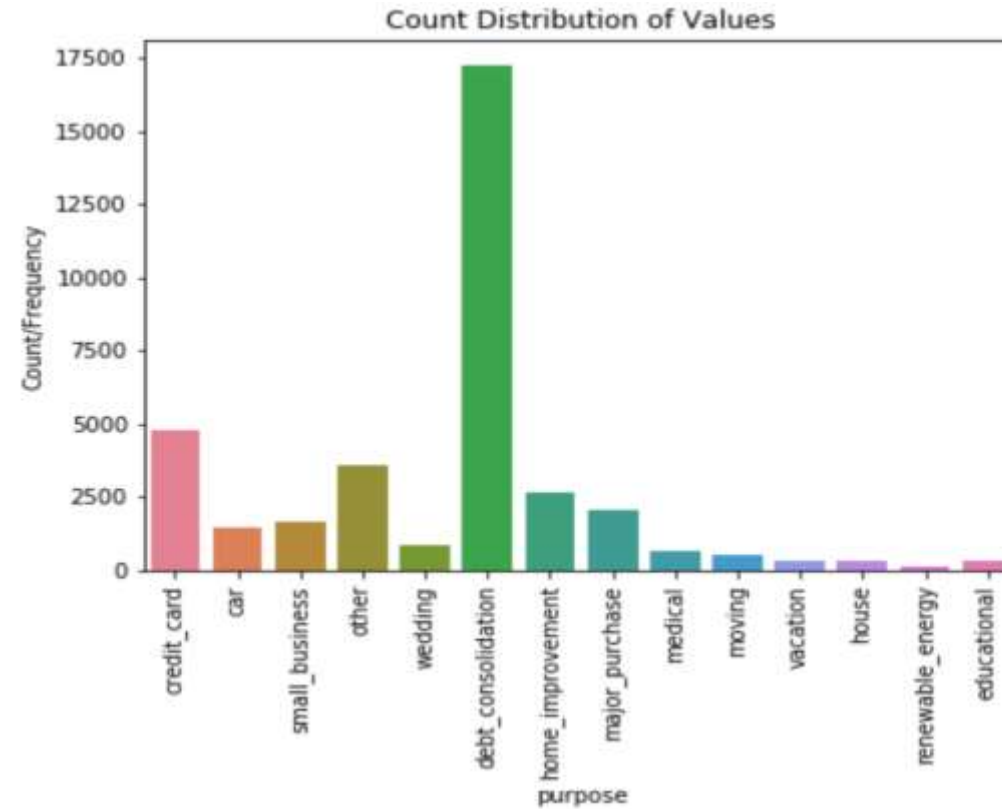
# UNIVARIATE ANALYSIS

## DATA PREPARATION/ CLEANING:

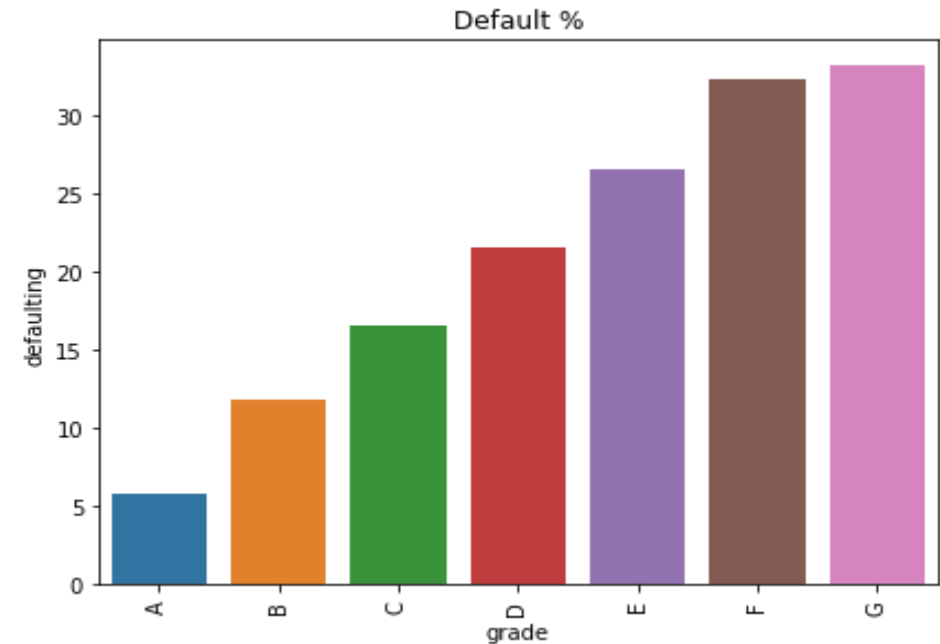
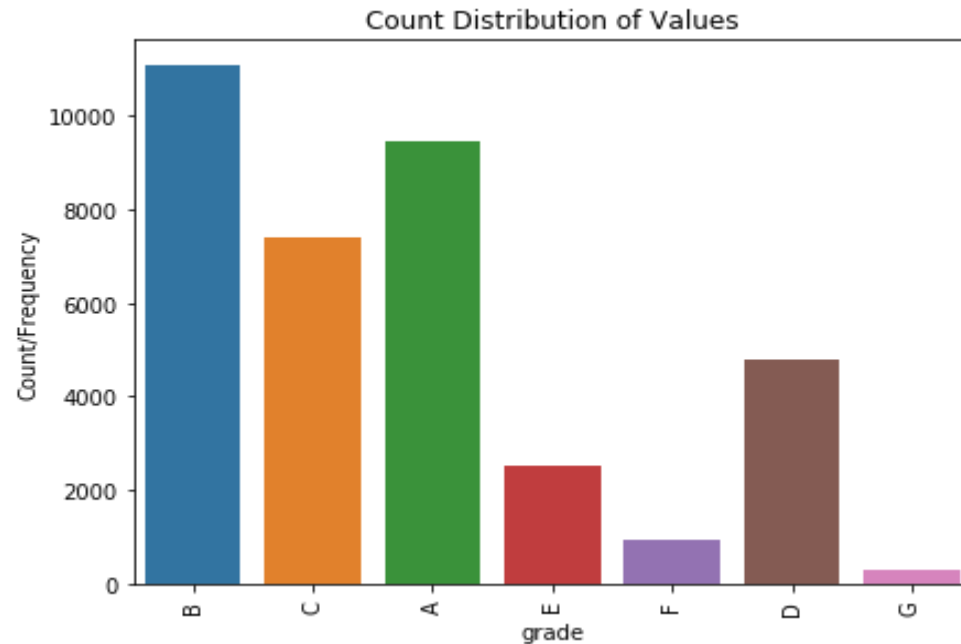
- We decided to remove the rows / columns that contains all NULL values.
- Further we remove the columns that contains 50% NULL values.
- Checking Outliers for numeric data and handling them.
- Converting object data type to appropriate data type.
- Considering the factors those support, applicants and before the loan is sanctioned characteristics such as 'addr\_state', 'grade', 'sub\_grade', ' emp\_length', so on., which are described in python file.

# Understanding the Loans

Maximum number of Loans are for debt consolidation followed by Credit cards

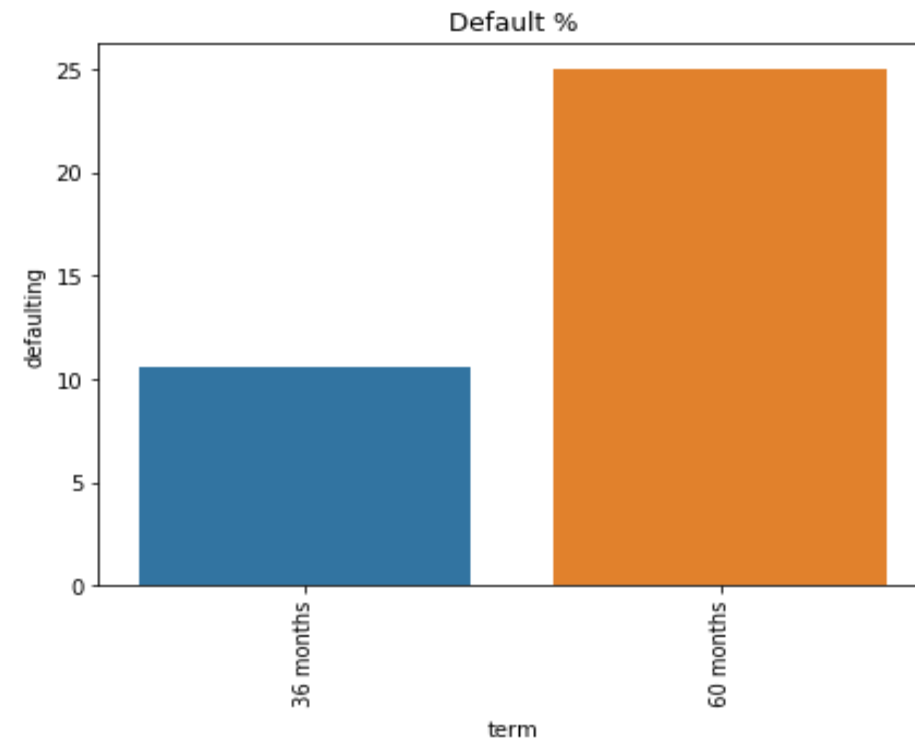
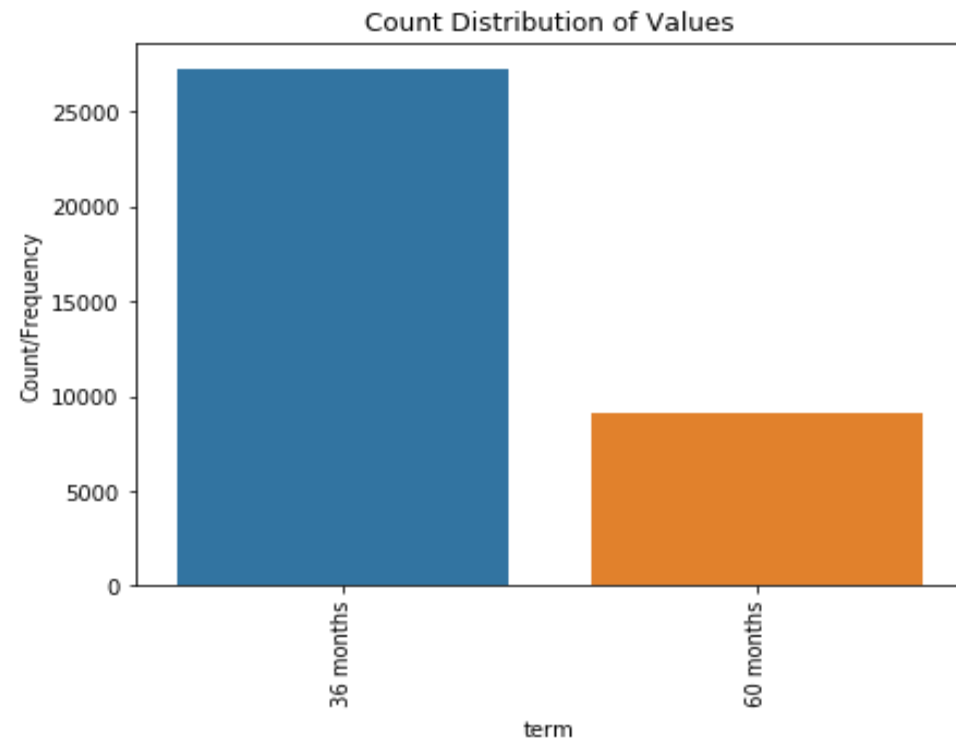


# Understanding the Loans (contd..)



- From above graph, it is considered that grades E, F, and G have more defaulters even though possess low range.

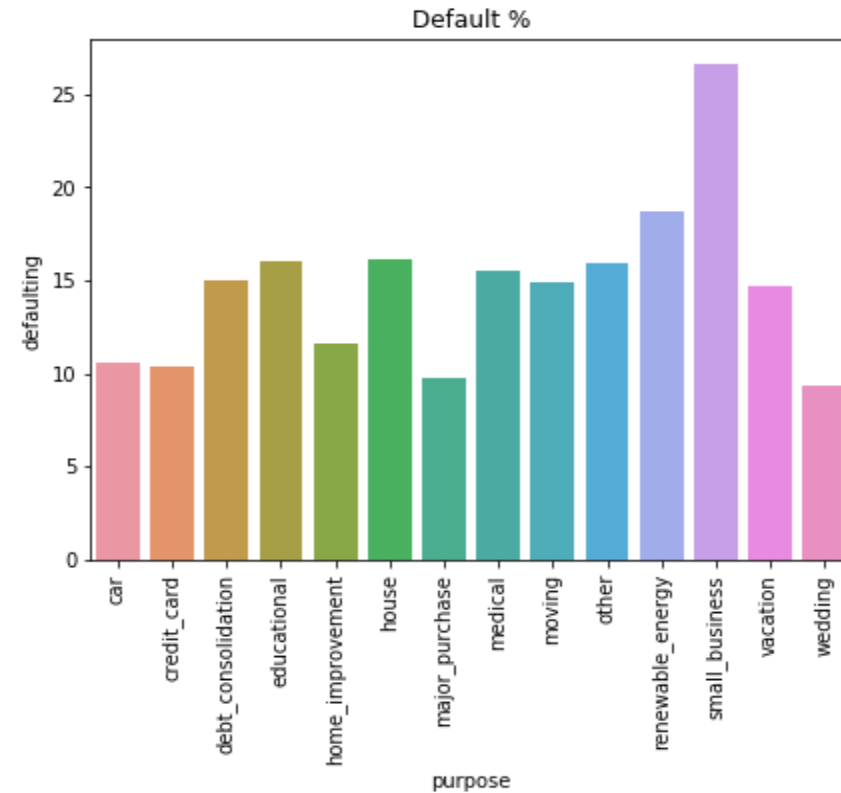
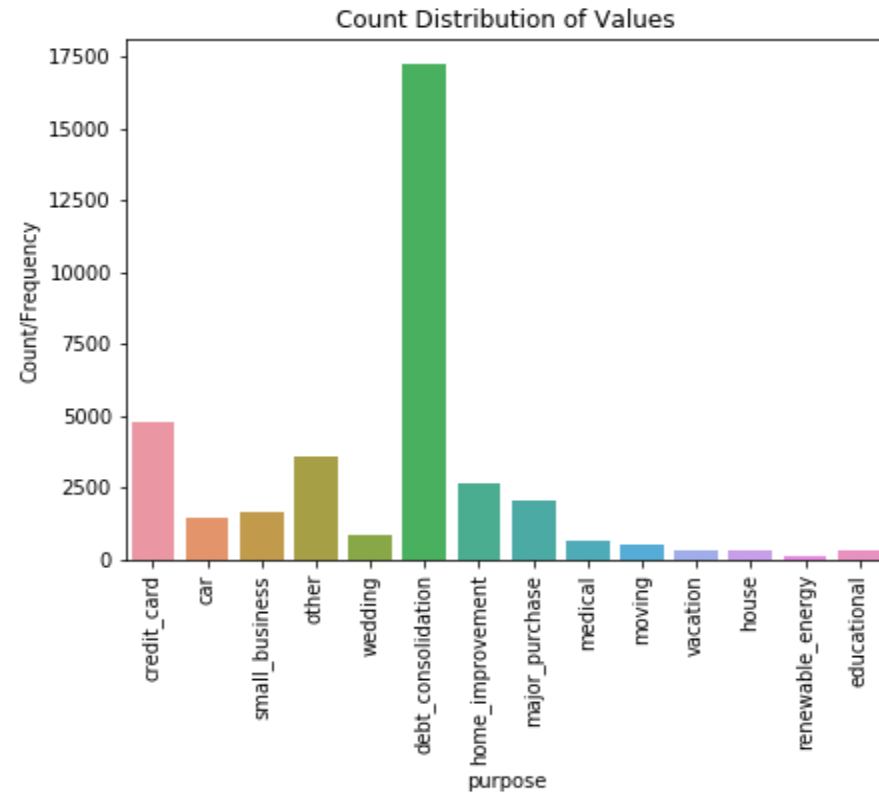
# Continues..



- From this graph, it is considered that even though the range in 'term' for 60 months is less, still the defaulters are more.

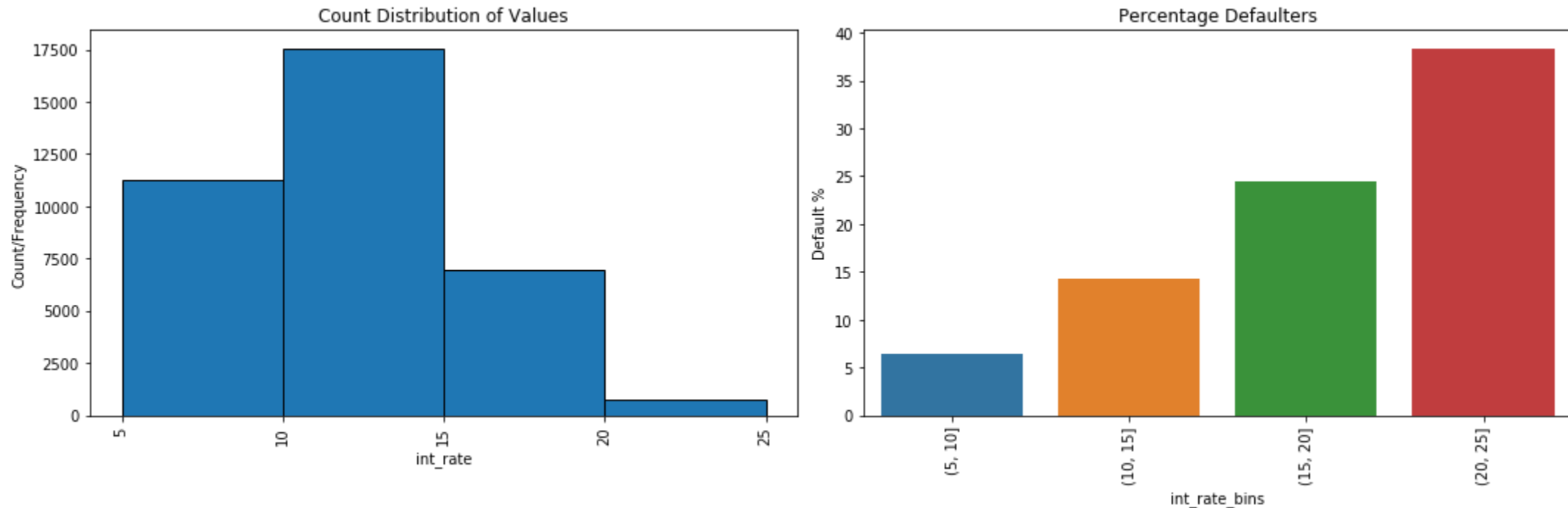


# Continues..



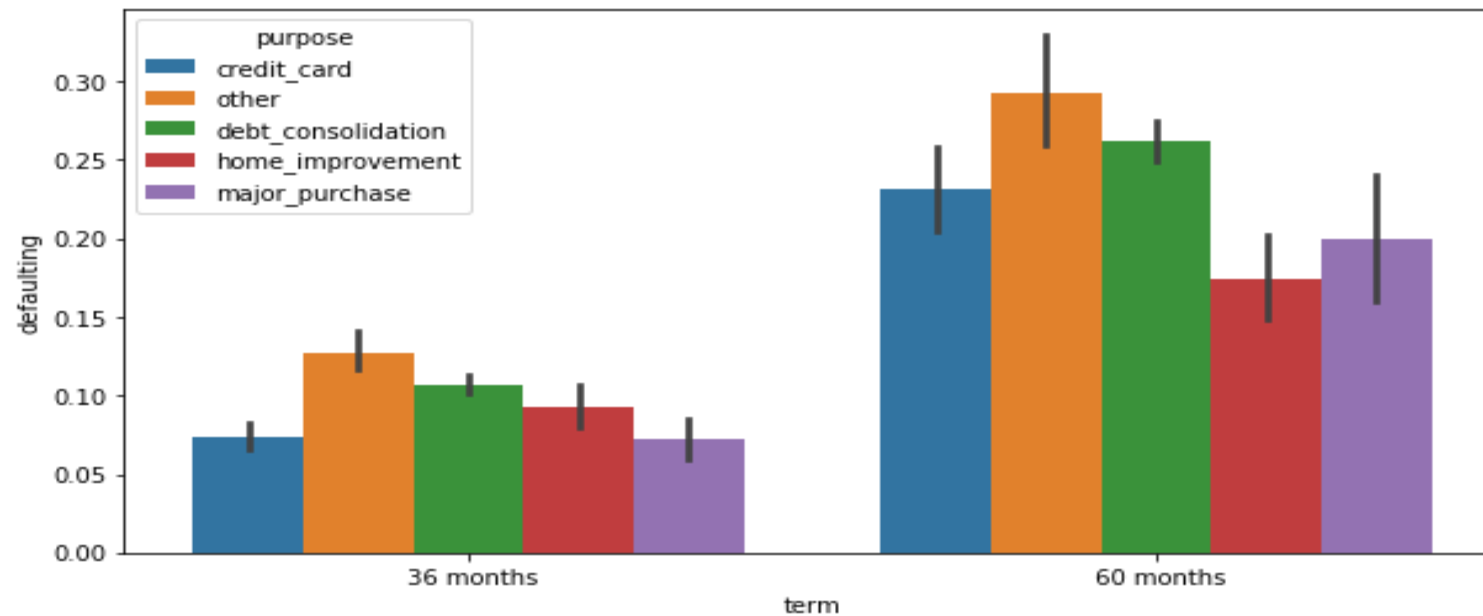
- Small\_business have more no. of defaulters even if their range is low as shown from the above graph.

# Continues...



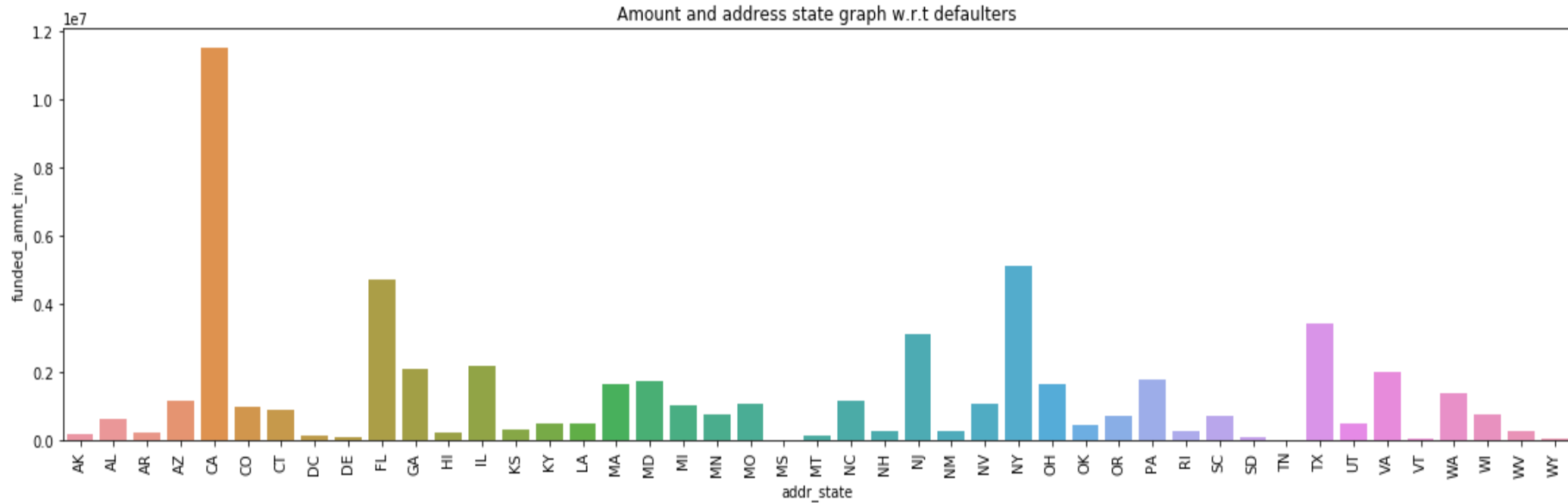
- Percentage of Defaults increases monotonically with higher interest rates. At rates above 20%, more than 35% of loans are Charged-Off.

# SEGMENTED ANALYSIS



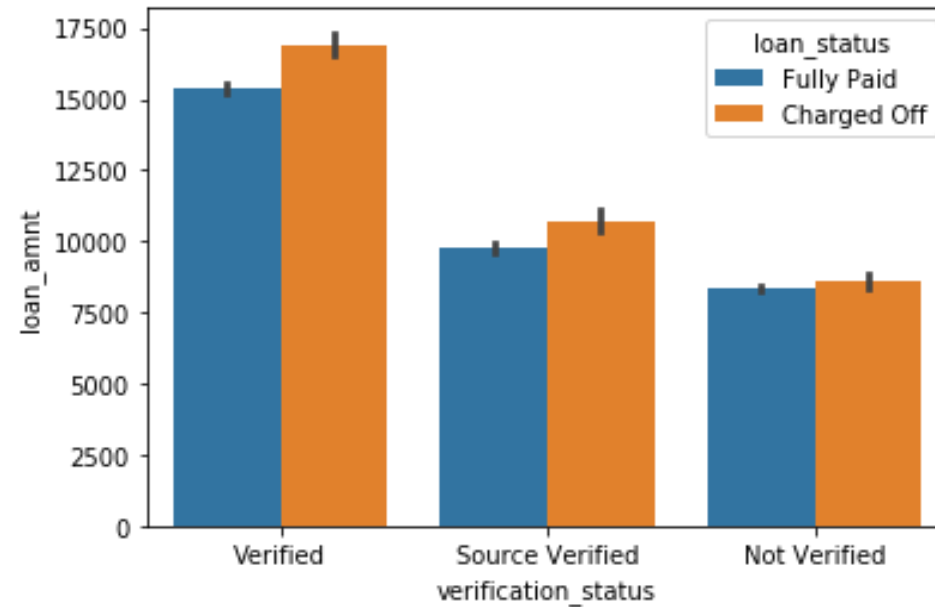
- Above graph shows defaulters grouping term and purpose which indicates ‘others’ is highest among all.

# BIVARIATE ANALYSIS



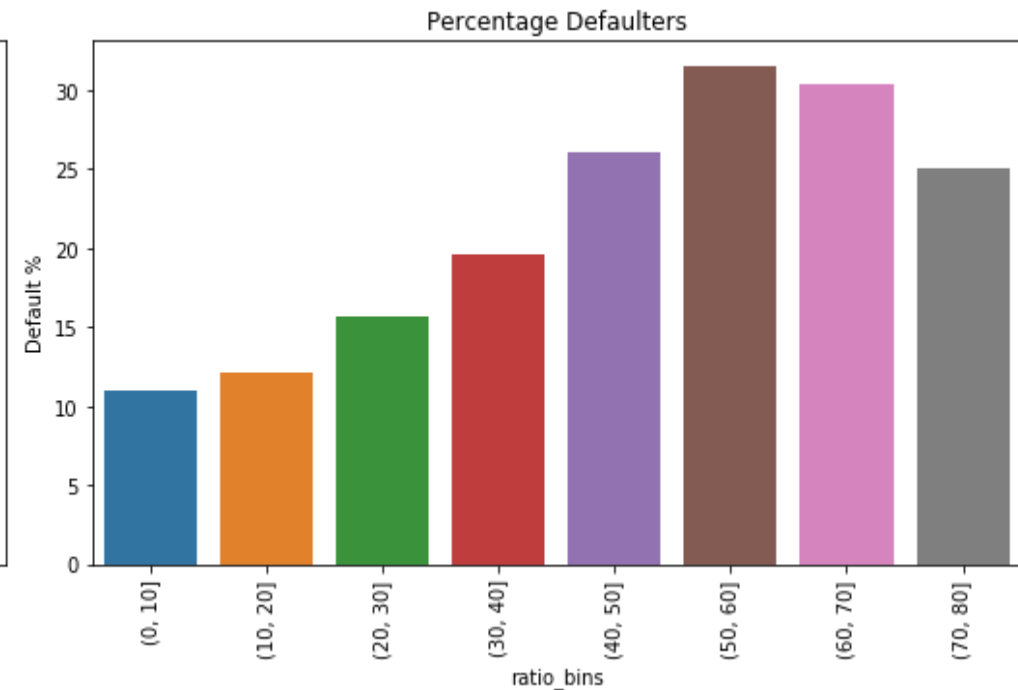
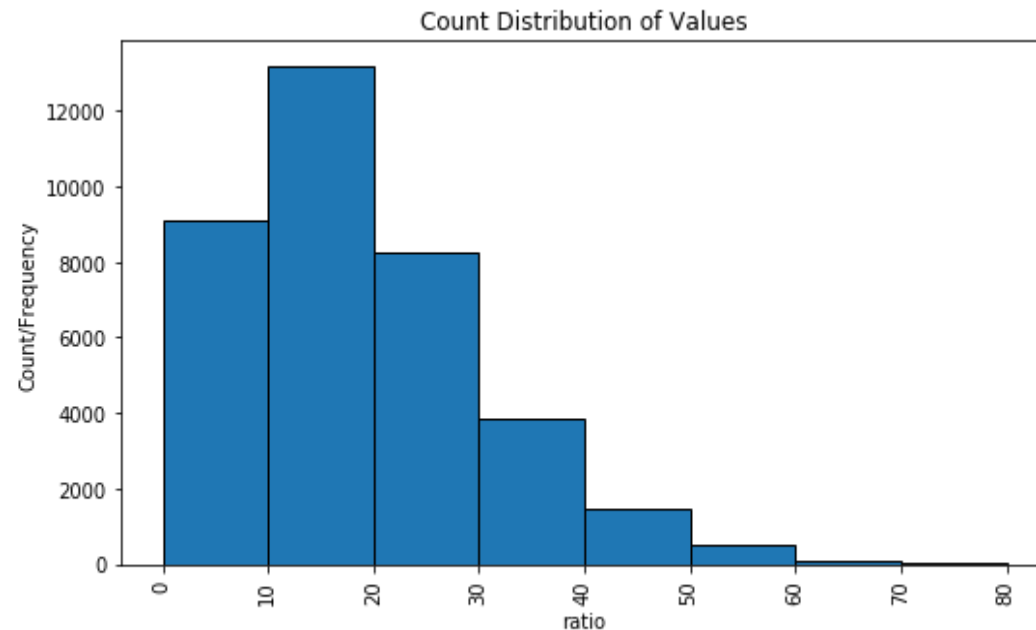
- Above graph depicts that the 'CA' has lost higher amount in terms of defaulters.

# Continues...



- We can observe from the above graph that the verified category have the more chances to default.

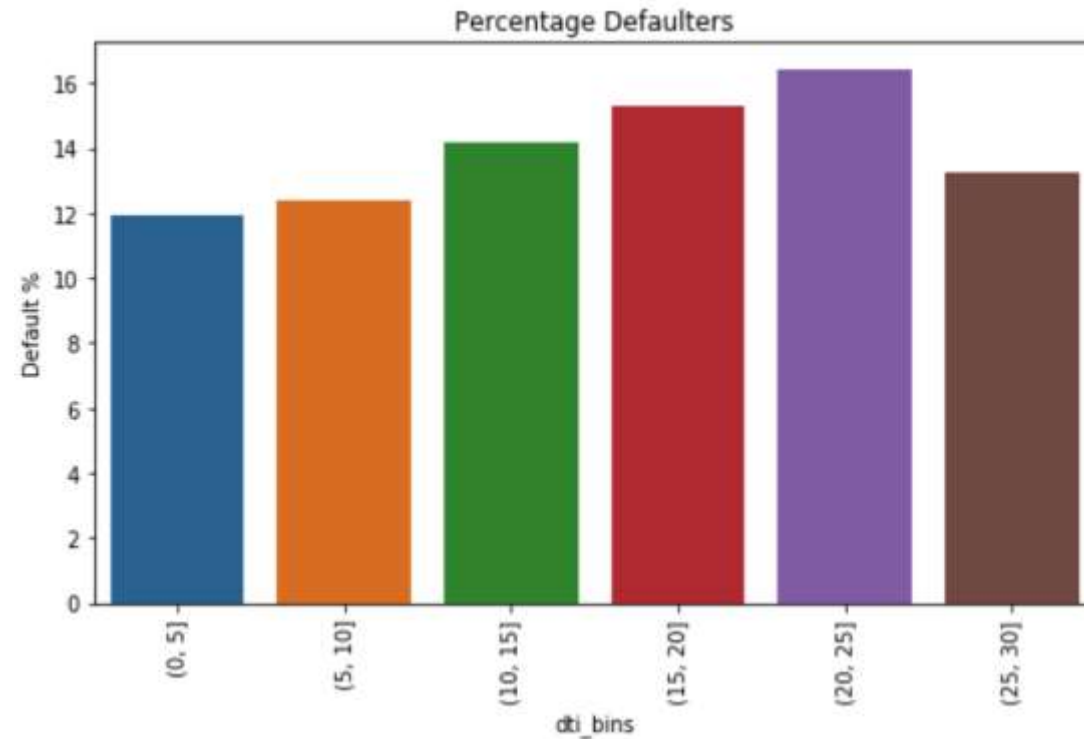
# DERIVED METRICS



- As long as loan amount is less than 30% of annual income, defaults are low.
- Loan amounts of 40% of annual income or higher see a high rate of default

# Defaults vs Debt to Income Ratio

When the dti is between 15 and 25, the defaulting is more

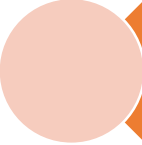





# INFERENCES

- From the analysis shown in previous slides, we get following observations which have more probability of defaulting(Charged Off):
  - Grade
  - Term
  - Address state
  - Business type such as small business
  - Interest rate
  - Verification status
  - Purpose
  - Ratio of loan amount to annual income.



# Recommendations

-  Stop—approving loans where Loan/income ratio is more than 40%
-  Reduce – the number of approvals where purpose is small business
-  Stop—approving loans to people with prior bad record. Or at least stop approving high-valued loans
-  Start—charging higher interest rates for loans with dti greater than 15

THANK YOU