

# Lending Club Case Study

---

Animesh Pandey

# Objective

---

- The aim of the case study is to implement Exploratory Data Analysis concepts to a real life dataset and scenario to bring about the needed results.
- To get the basic idea of how real life business scenario looks like.
- To create multiple visualization graphs and charts to enhance the understanding of data.
- To use analysis prowess to analyze the pattern of an action which can be used for better results in future.

# Business Objective

---

The business objective is to take a decision which are

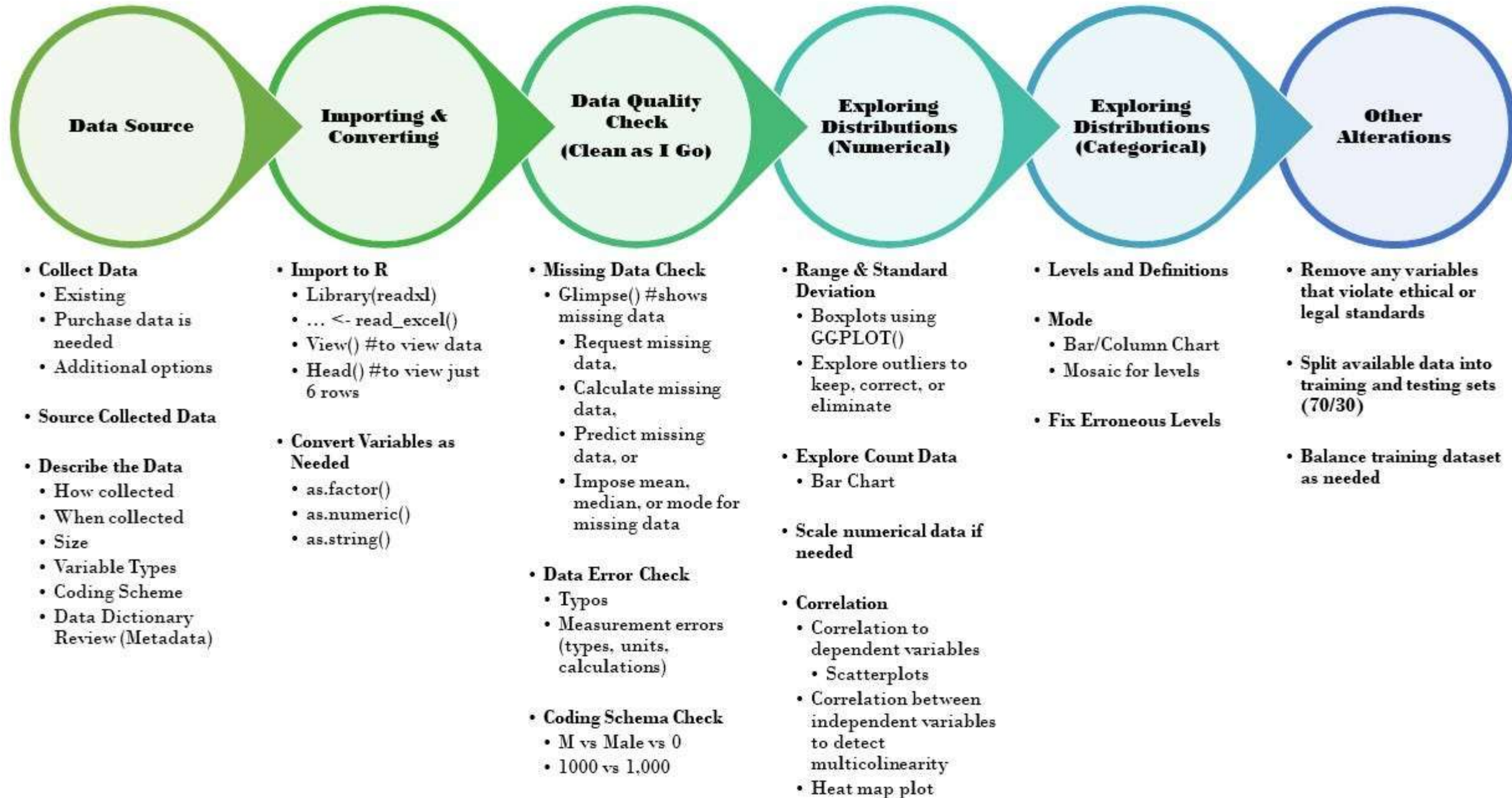
If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company

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

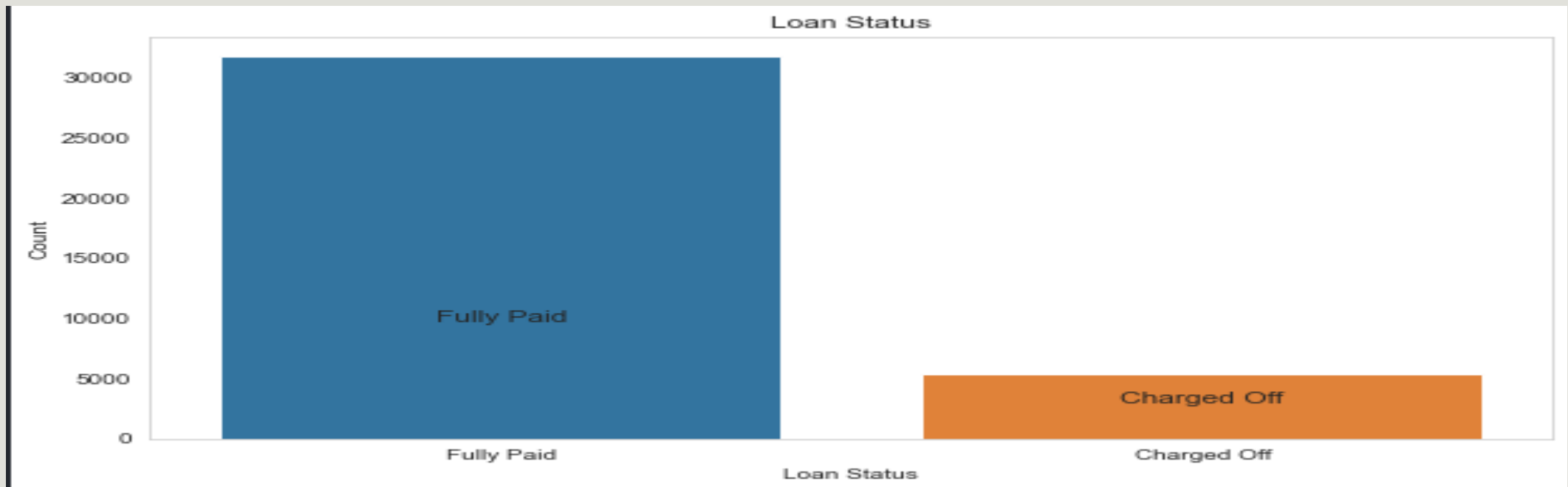
## **Dataset Details:**

The given dataset contains information about past loan applicants and whether they 'defaulted' or not. The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. past loan applicants and whether they 'defaulted' or not. Data has details regarding approved loan not the rejected ones. It has 3 status of loan which is Fully Paid, Current and Charged-Off.

# Data Understanding & Preparation Flow

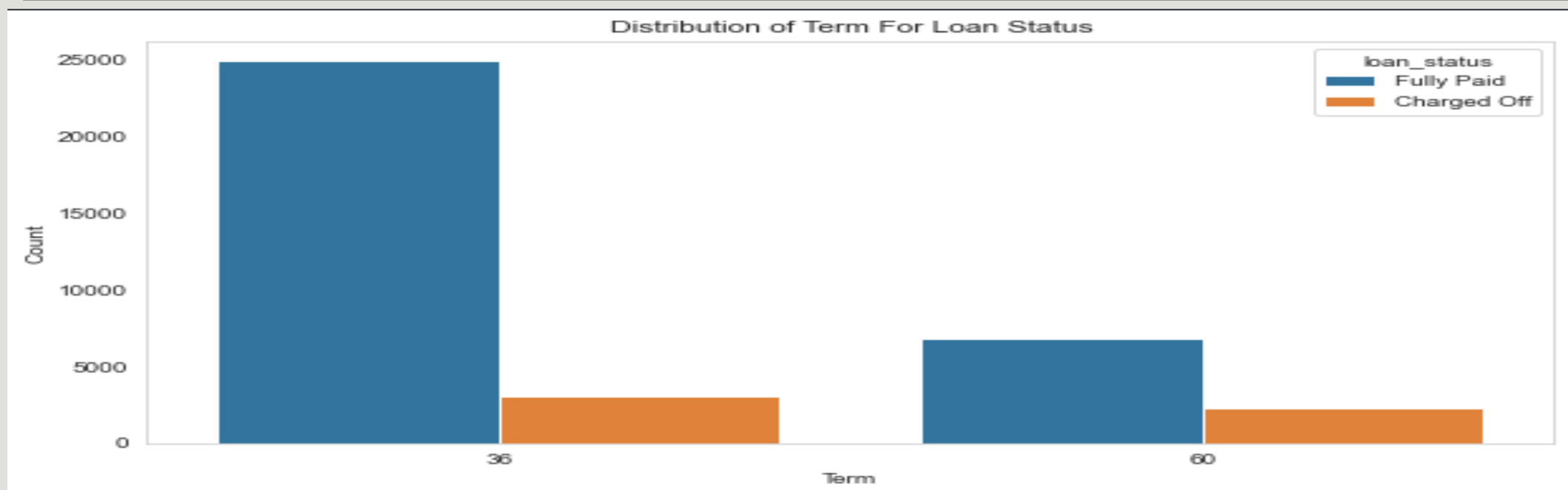


# Loan Status



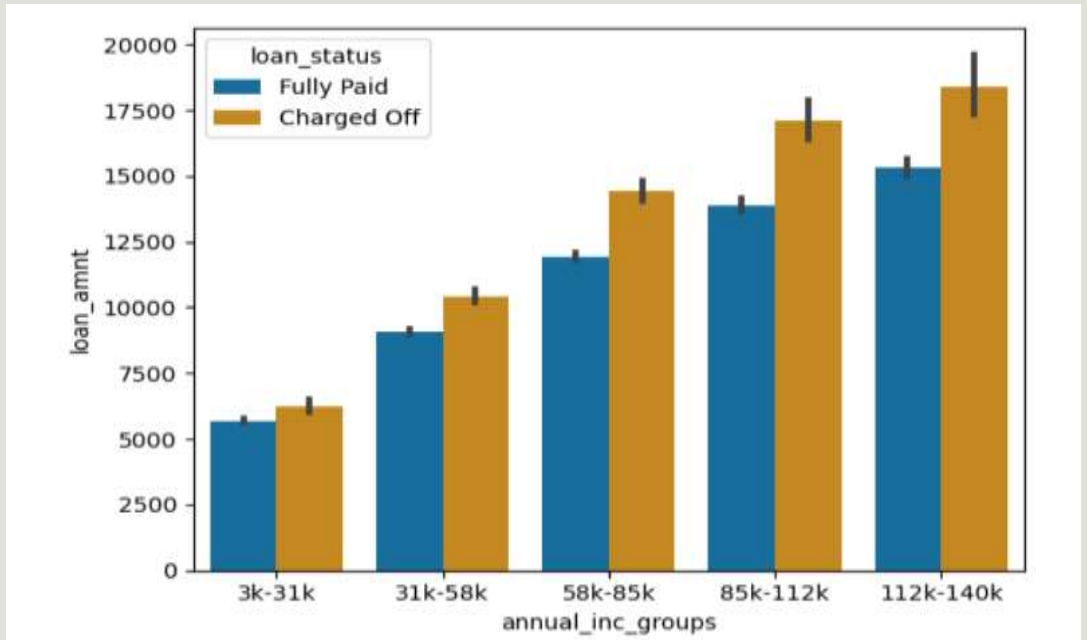
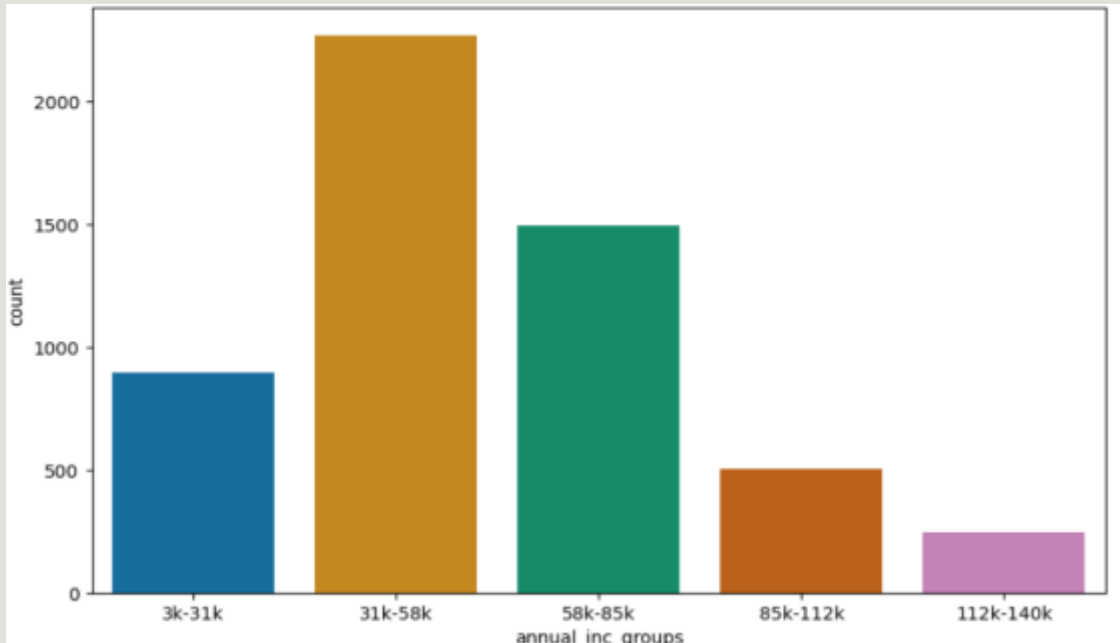
The number of charged off loan is way smaller compared to the fully paid count which means the loan defaulters are less.

# Loan Term differentiation



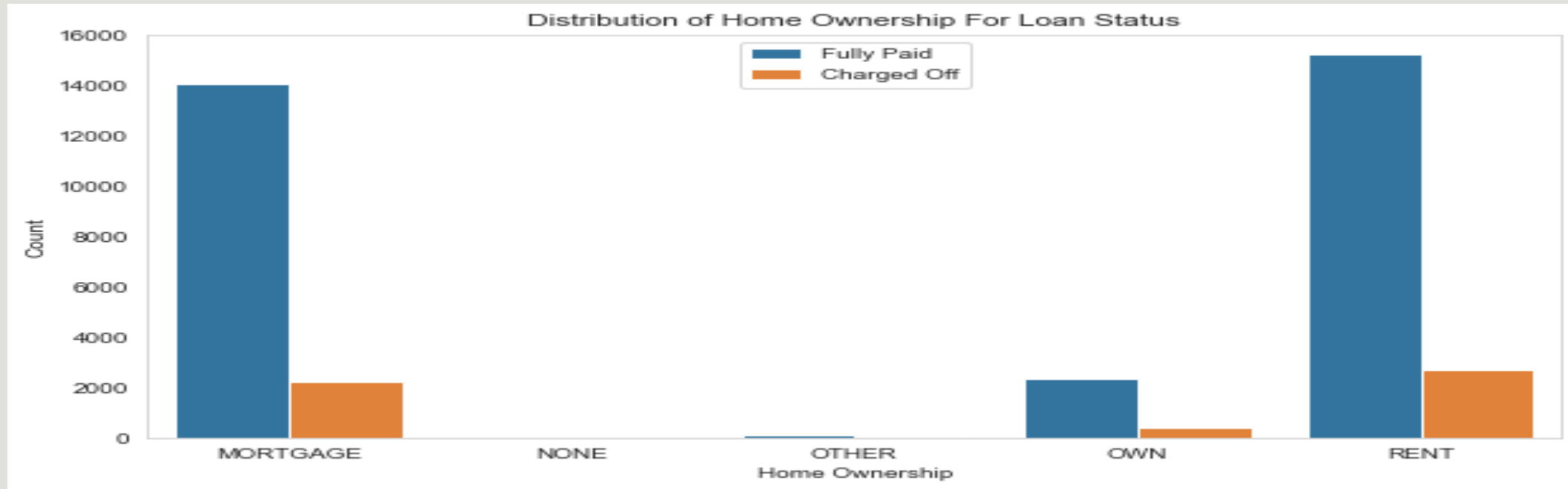
There are 2 terms for loan, 60 months and 36 months and as the above graph suggests, 36 months loan term has more count than that of 60 months. The loan defaulters in 36 months term is significantly lower than 60 months loan term.

# Loan as per Annual Income Groups



- As the above graph suggests, the income group of 31K to 58K are taking loans in larger count than any employee group.
- The second graph describes that people with larger salaries take loan of larger amount and the fully paid vs charged off count ratio is almost similar for all income groups but it decreases a bit as the annual salary and the loan amount decreases.

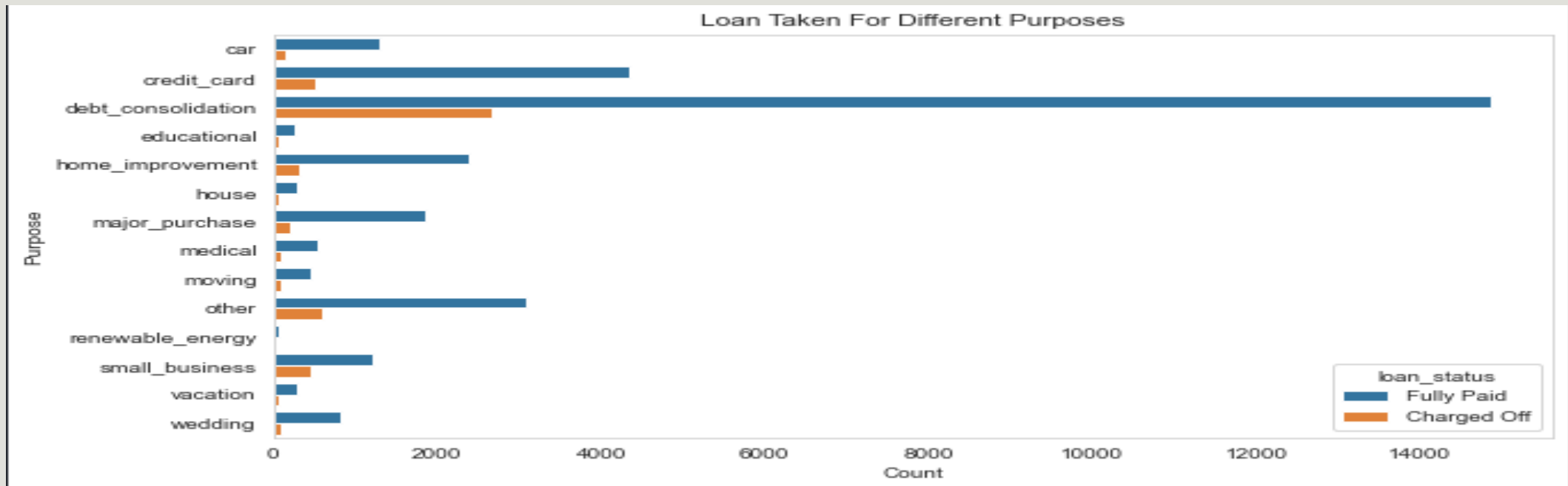
# Home Ownership versus Loan



As the graph suggests people with no home ownership i.e. with Mortgage and Rent has most of the loan and charged off count.



# Purpose of loan versus loan status



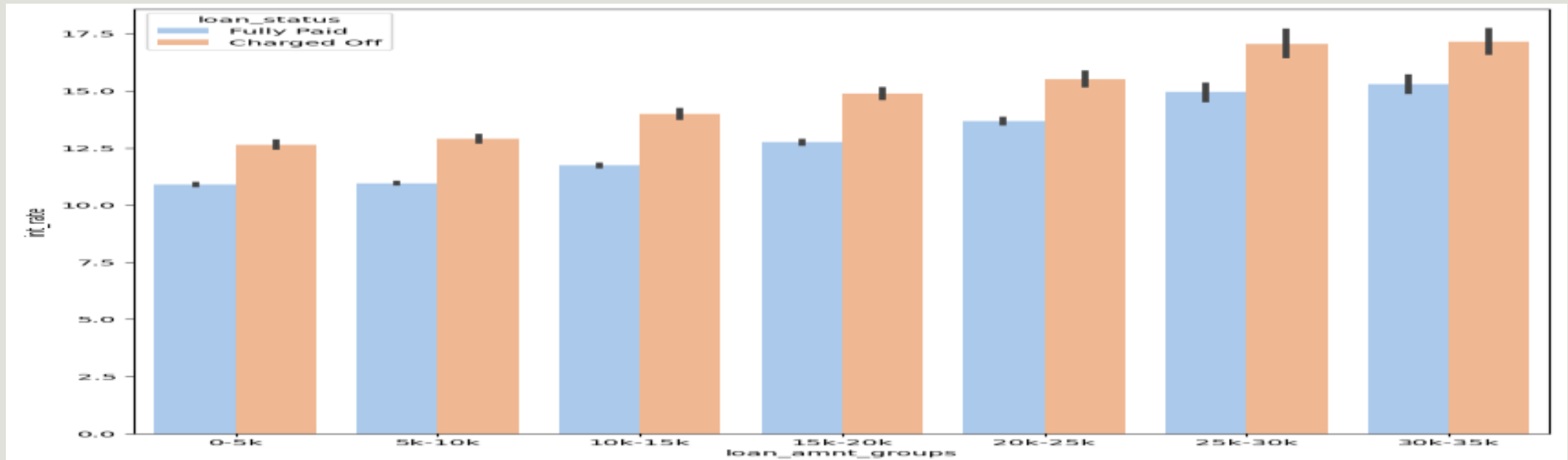
The debt consolidation has highest fully paid loan but also has the highest charged off count too.

# Loan Trend over years



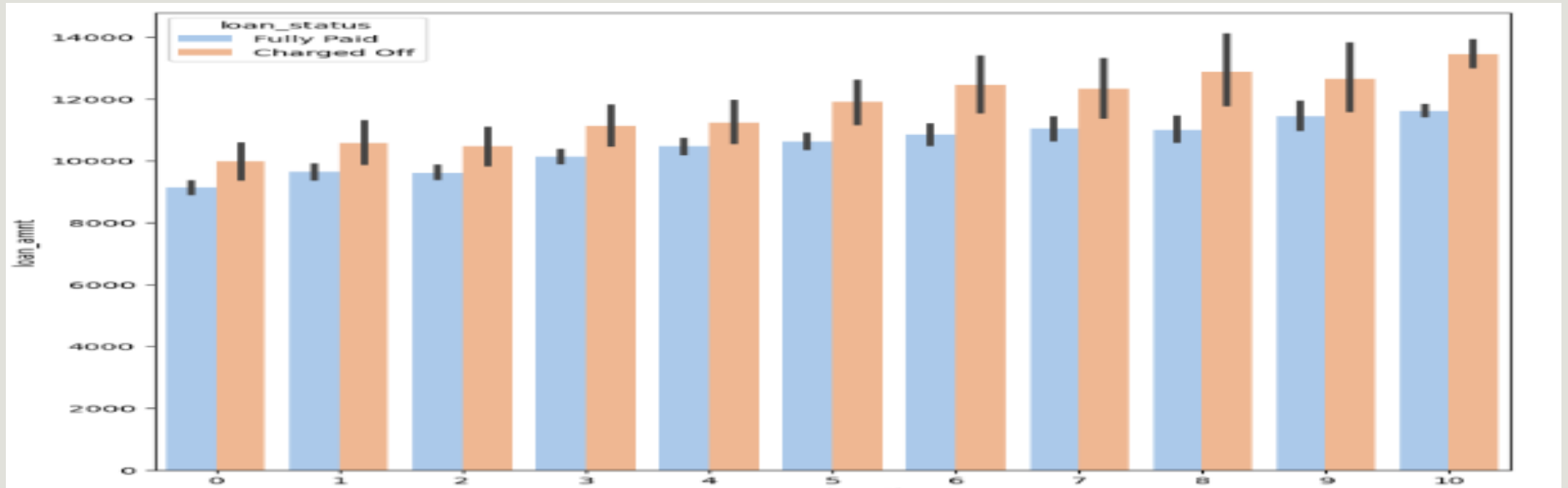
The loan over years have increased almost exponentially and charged off has also increased with time.

# Loan Amount versus Interest Rate



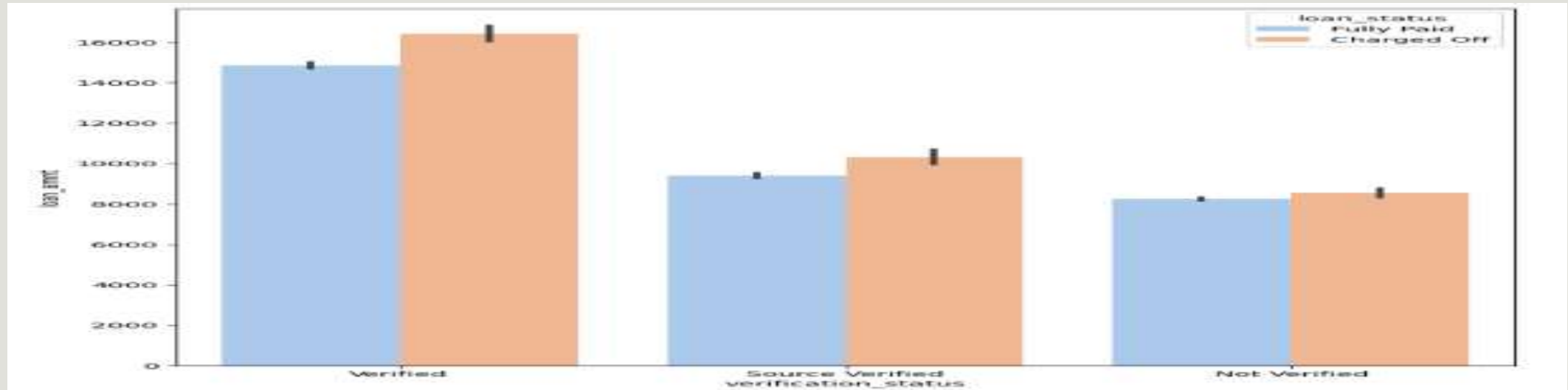
As the above graph represents, the loan taken with different interest groups does not have much impact on fully paid or charged off counts.

# Employment Length versus loan status



Employees with higher employment length are suppose to get higher loans and charged off count also increases.

# Verification Status versus Loan



People who do not have their status verified have lesser loan amounts sanctioned to them.

# Conclusions

---

## Recommendations

Major Driving factor which can be used to predict the chance of defaulting and avoiding Credit Loss:

1. DTI
2. Grades
3. Verification Status
4. Annual income
5. Pub\_rec\_bankruptcies

Other considerations for 'defaults' :

1. Borrowers not from large urban cities like California, new york, texas, florida etc.
2. Borrowers having annual income in the range 50000-100000.
3. Borrowers having Public Recorded Bankruptcy.
4. Borrowers with least grades like E,F,G which indicates high risk.
5. Borrowers with very high Debt to Income value.
6. Borrowers with working experience 10+ years.