GRAMENER CASE STUDY

Presented by:

Subarna Saha

OBJECTIVE

- **Business Understanding**

A consumer finance company 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.

- **Business Objective:**

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 utilize this knowledge for its portfolio and risk assessment.

- **Business Decision:**

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 instalments, 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 instalments 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)
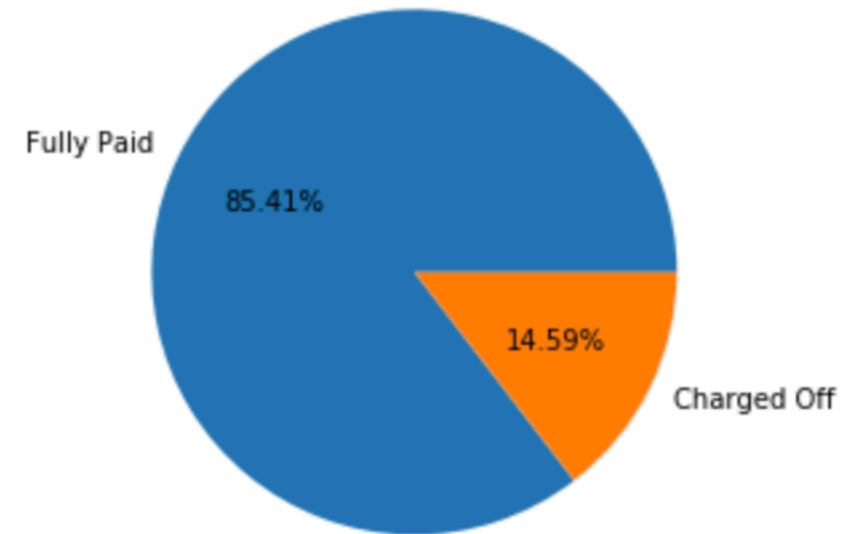
DATA CLEANING AND FORMATTING

- We observe that we have 39717 rows and 111 columns.

- We first discard columns that have more than 50 % null value.

- Drop columns that do not add an value to analysis of loan defaulters.

- We then format those columns to correct their data type and clean any extra characters added to data.

- We then end up with 27 columns on which we will proceed our analysis.

- We then remove the rows having loan_status="Current" , as it cannot give us any information whether the person will be a defaulter or not as its still on-going and we cannot know the final outcome. Therefore, we will be working with only the "Fully Paid" and "Charged off" loan_status.

DATA ANALYSIS

- From the data we see that during the time period 2007 to 2011 the company had 14.59% of loans being charged off and 85.41% Fully paid.
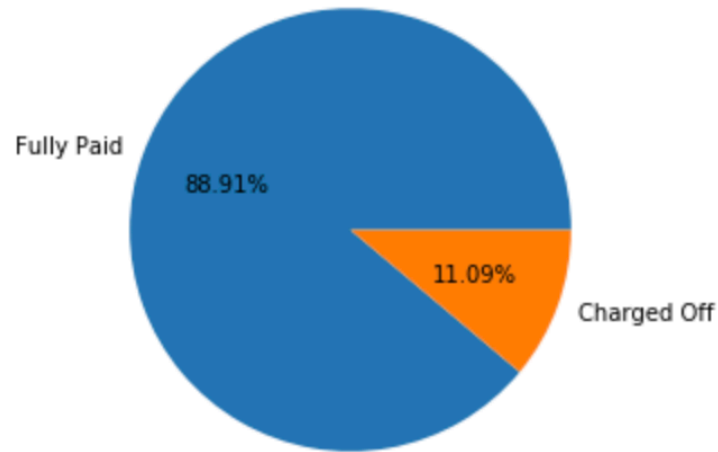
- This 14.59% was a loss to the company.

percentage of each loan status

Fully Paid
85.41%

14.59%

Charged Off

- We have loan terms of 36 months and 60 months.

- Now we compare the percentage of charged off loans for each "term". We see that when term is 60 months we see a higher percent of loan being charged off and a loss to the lending company.

- Also we see that applicants with a higher loan amount choses a 60 month term.
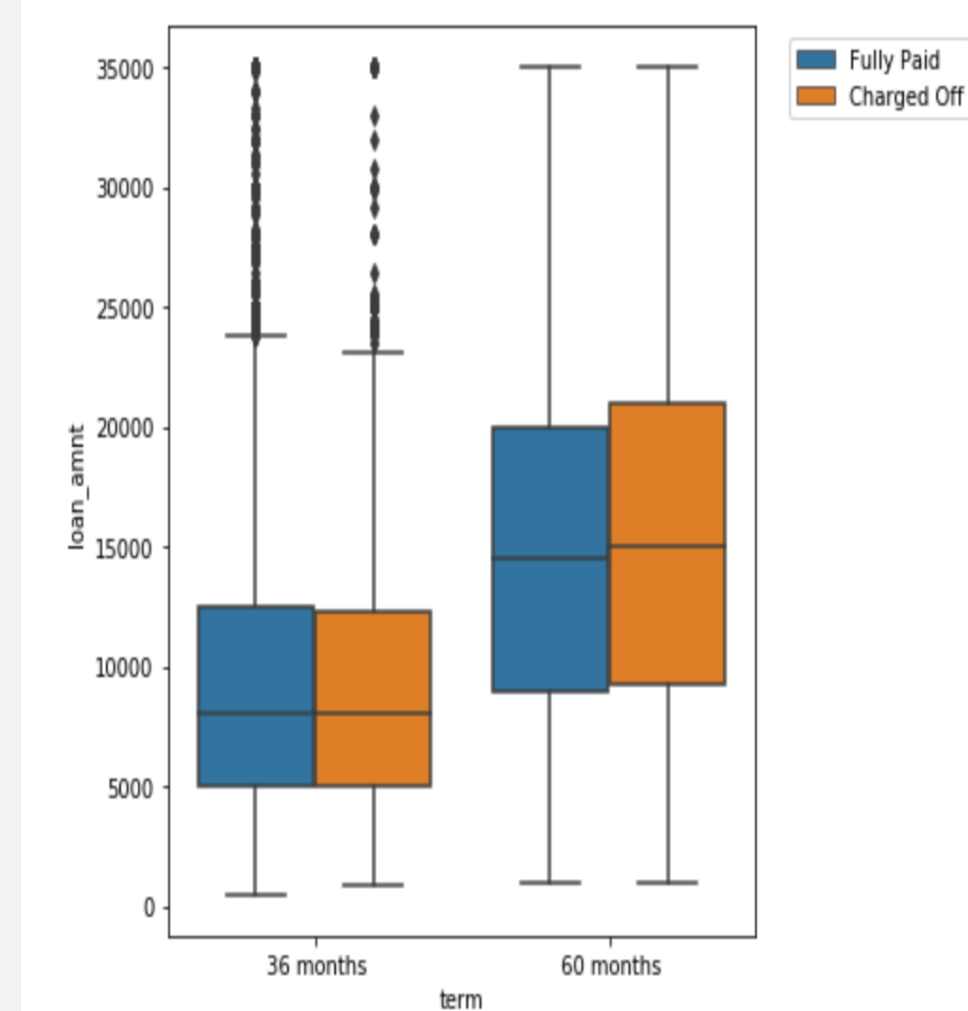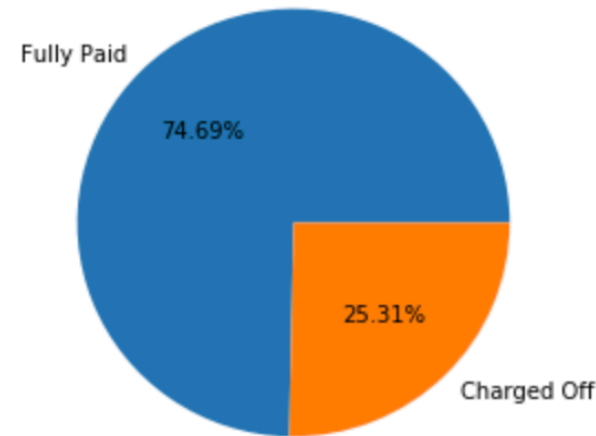
DATA ANALYSIS: GRADE AND SUBGRADE

- We the analyze with grade and subgrade for a loan.

- We observe that most of the loans charged off belonged to Grade B and subgrade B5
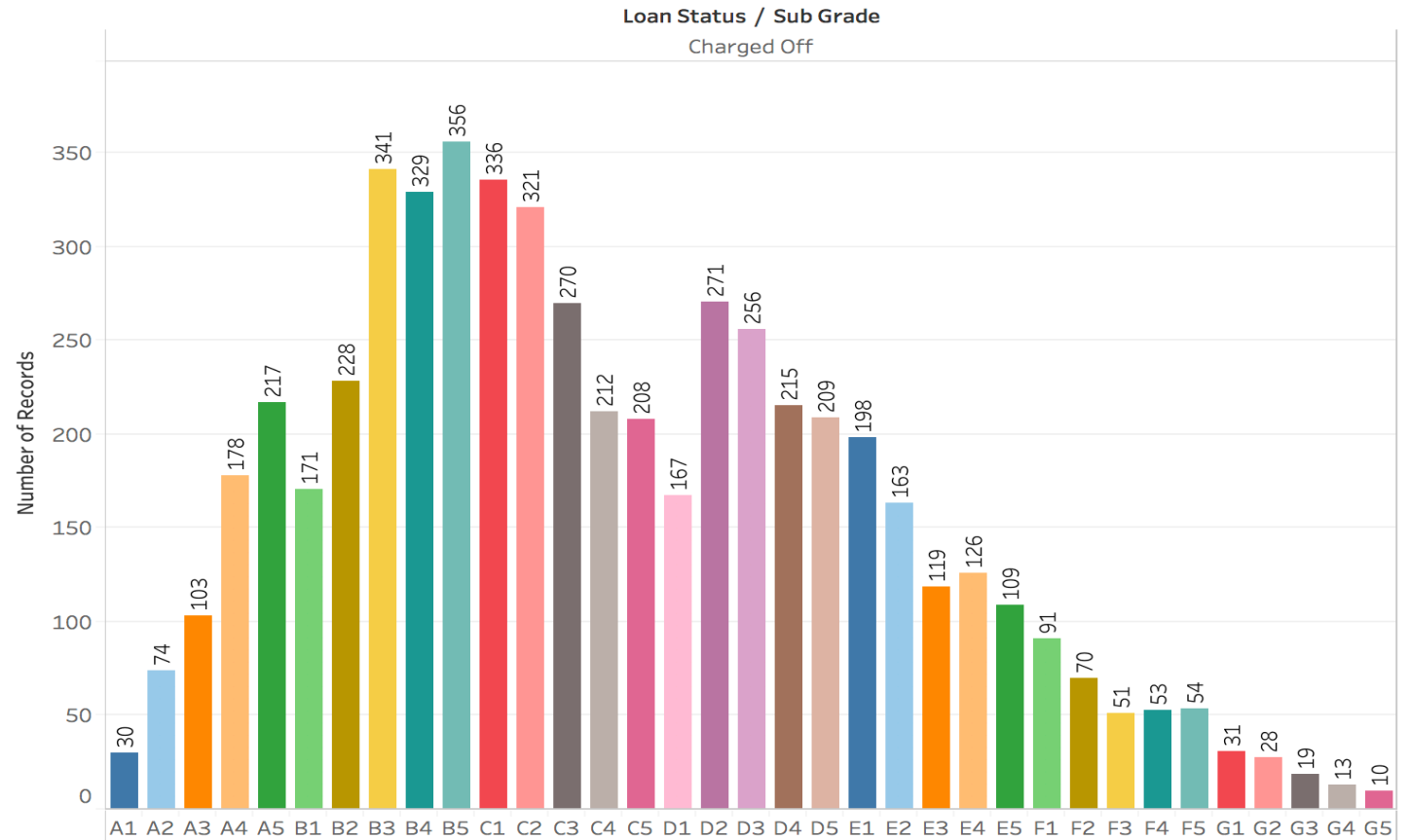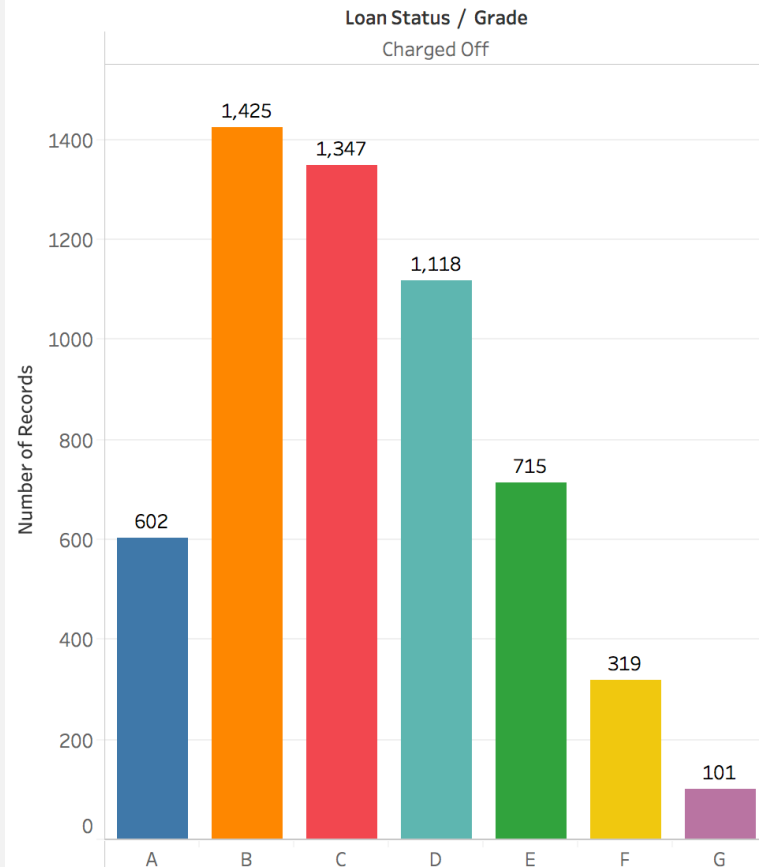
DATA ANALYSIS- HOME OWNERSHIP AND VERIFICATION STATUS

- We observe that charged off loans tend to have home ownership to be as rent or mortgage

- We observe that "Not verified" applicants tend to default most.

- We observe that not verified applicants mostly stay in rented accommodation.
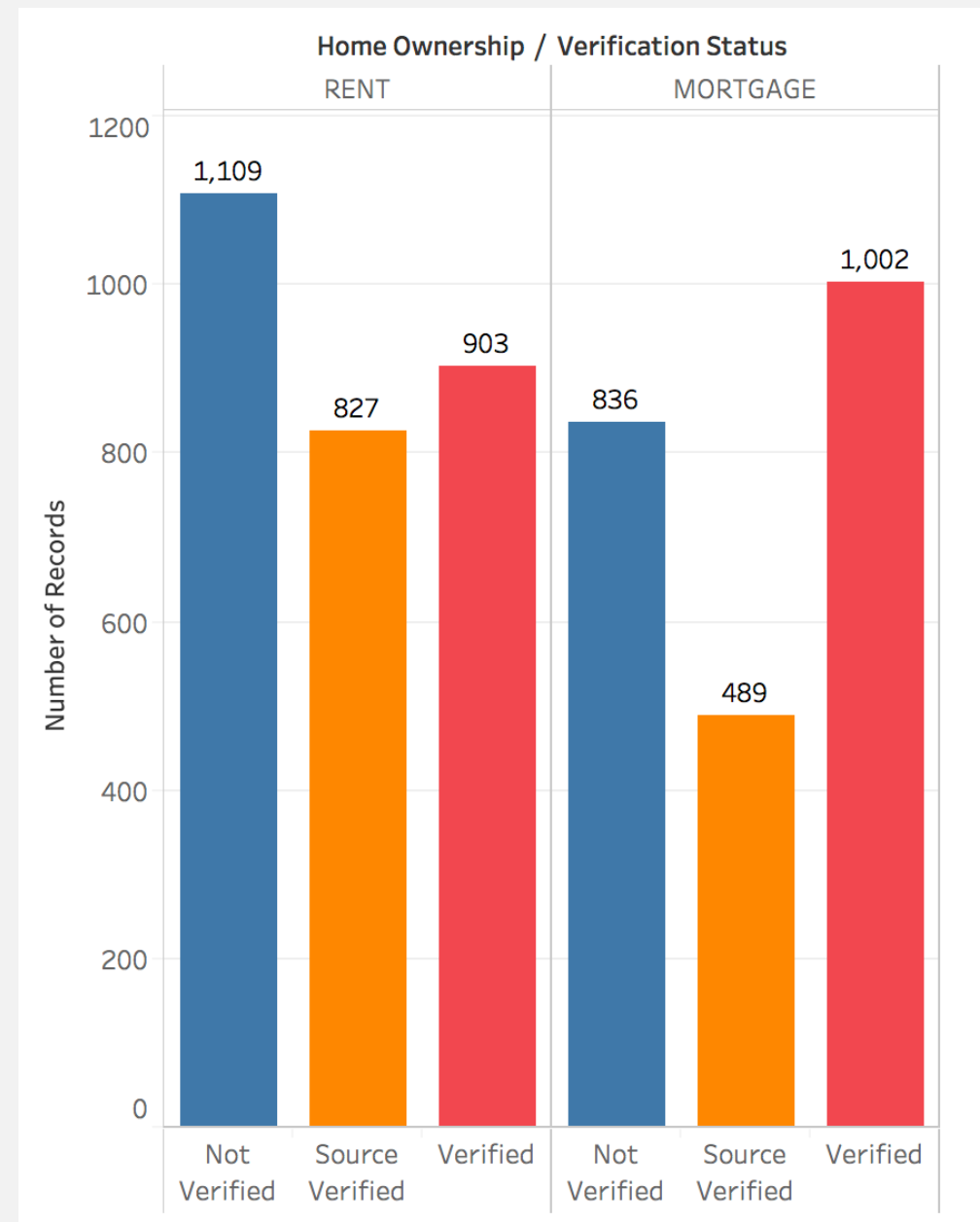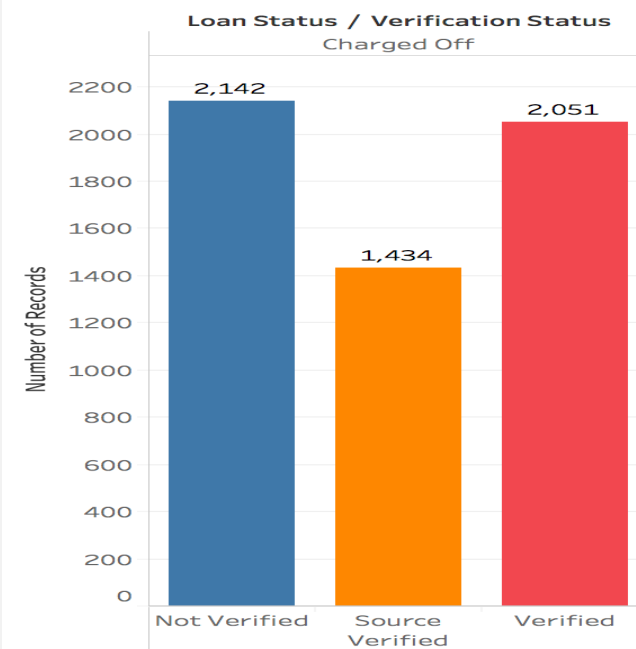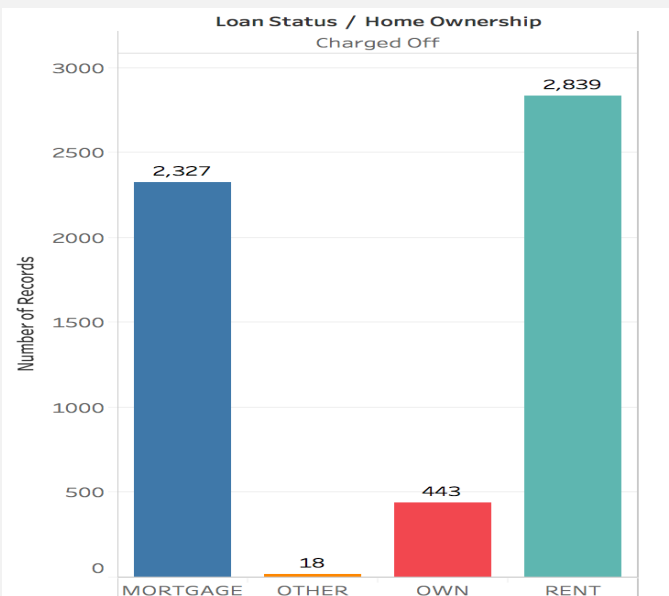
DATA ANALYSIS-PURPOSE OF LOAN

- We observe that purpose stated as "debt consolidation" tend to default most.
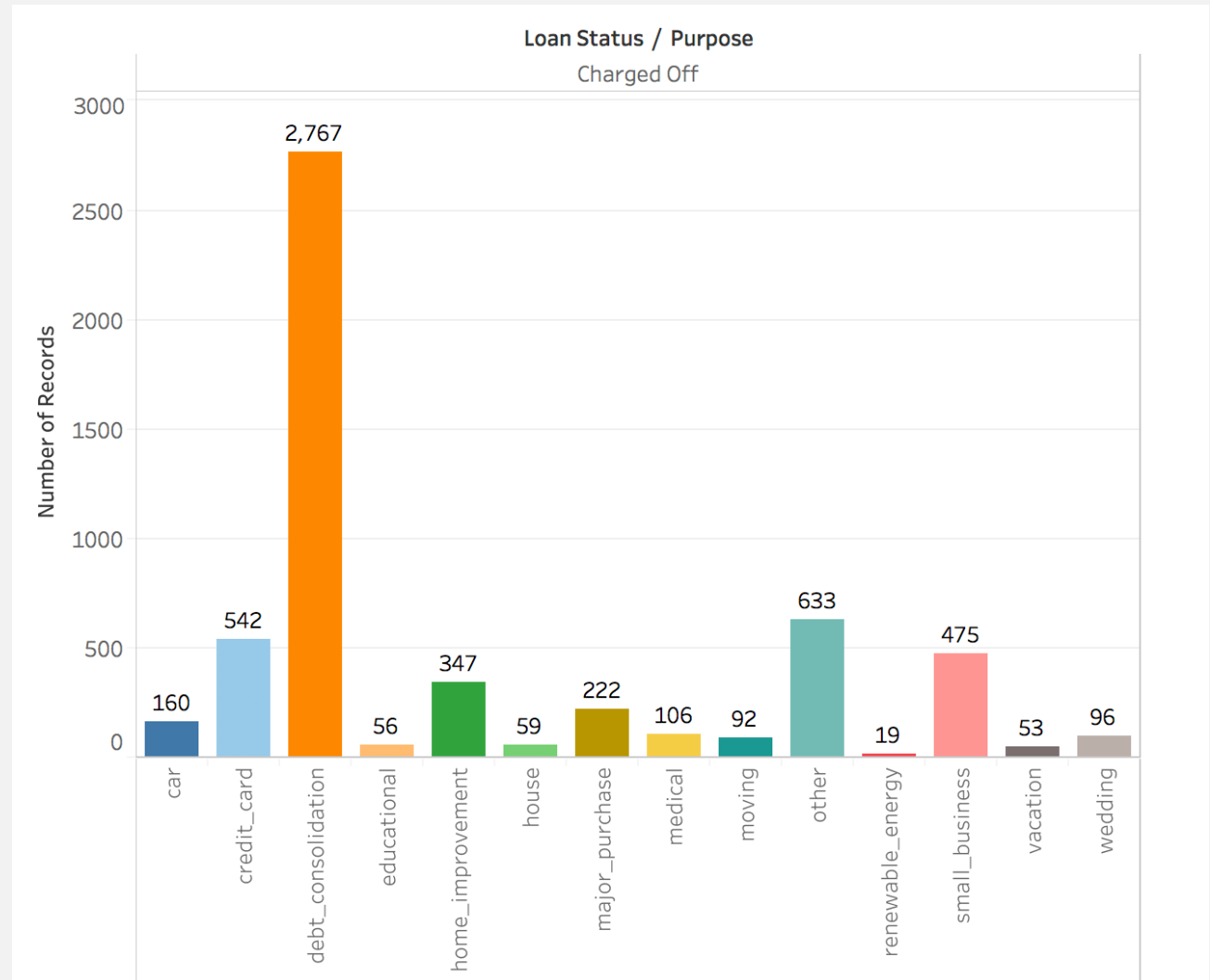
DATA ANALYSIS- ANNUAL INCOME VS LOAN AMOUNT

- Applicants with low income tend to default more.

- Applicants that belong to the high income category took higher loan amounts than people from low and medium income categories. We can observe that applicants who were charged off had taken loans of amount higher than the average amount taken by applicants who fully paid their loan.
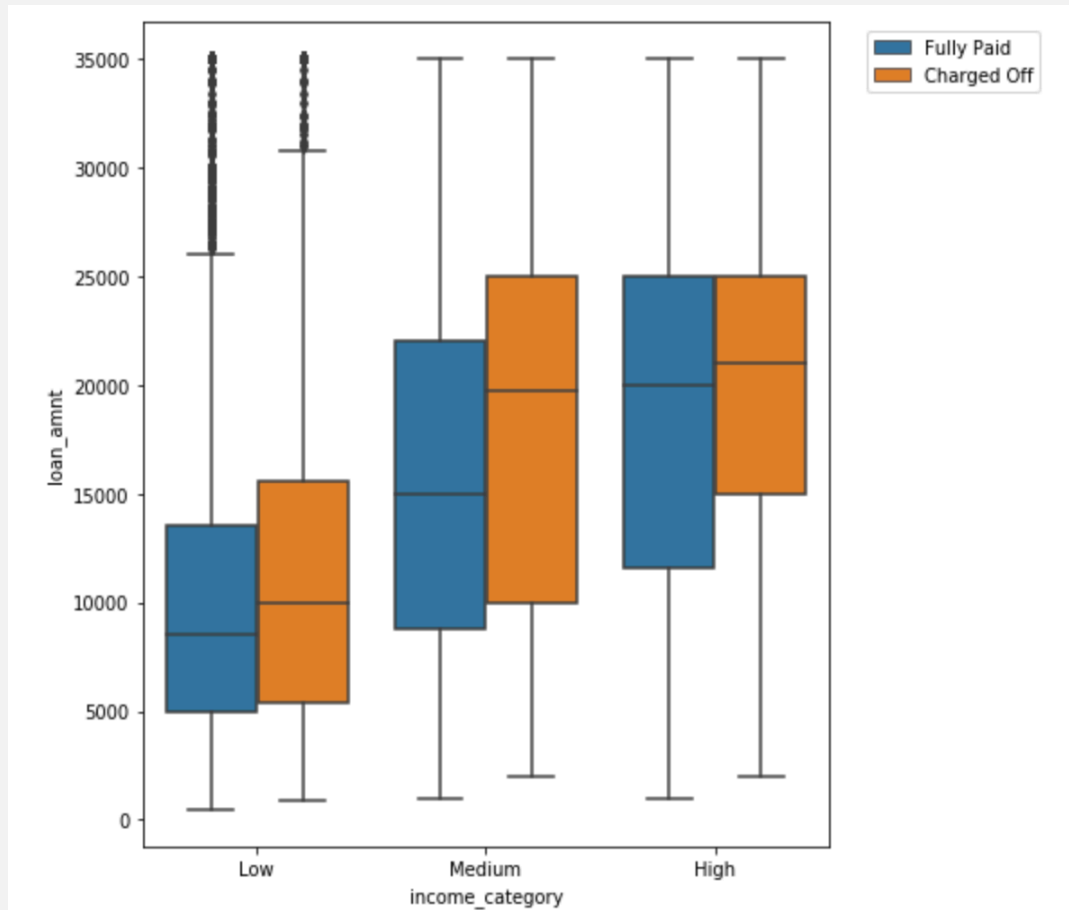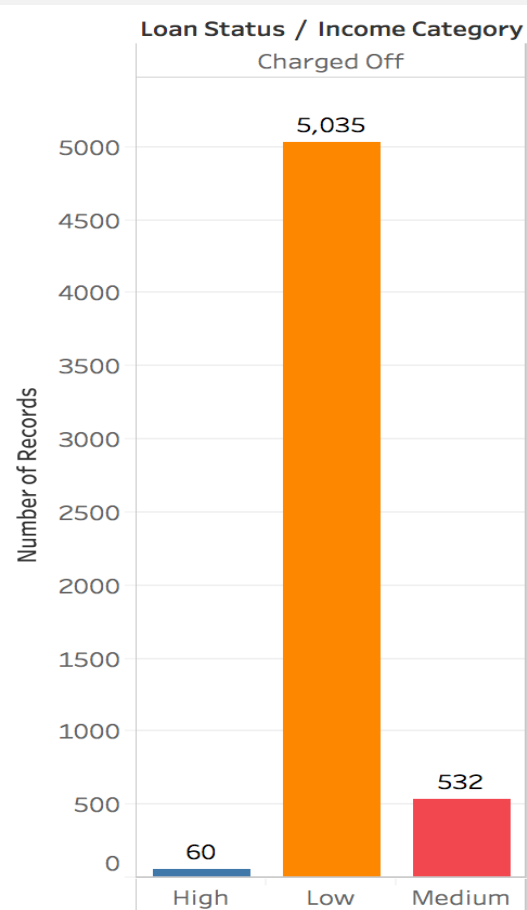
Assuming the income given are in USD,

We create the below income brackets:

Low income category: <=100,000 USD.
Medium income category: >100,000 USD & <=200,000 USD
High income category: >200,000 USD

DATA ANALYSIS- EMPLOYMENT LENGTH

- Applicants with High and Medium annual incomes have a longer employment length than people with lower incomes.

- Applicants with more than 10 yrs experience take higher amounts of loan and more probable to default.

- We also see that applicants who fall under high income category but have an average of less number of experience can default.
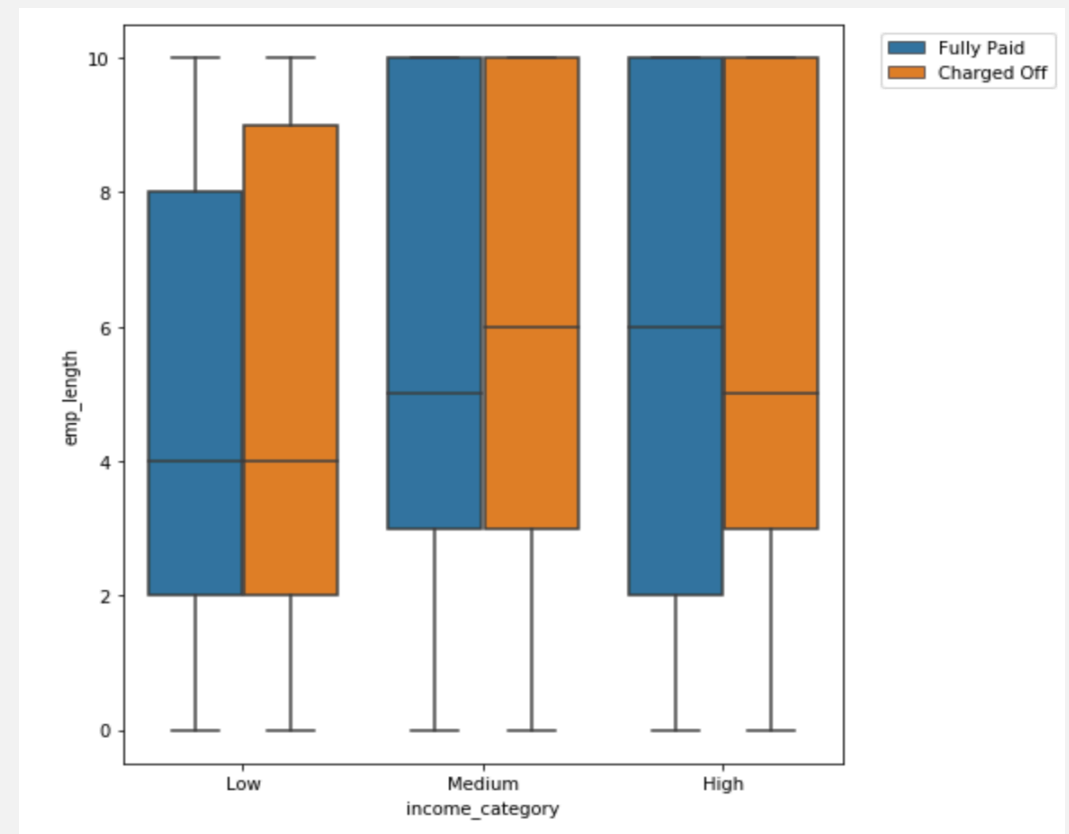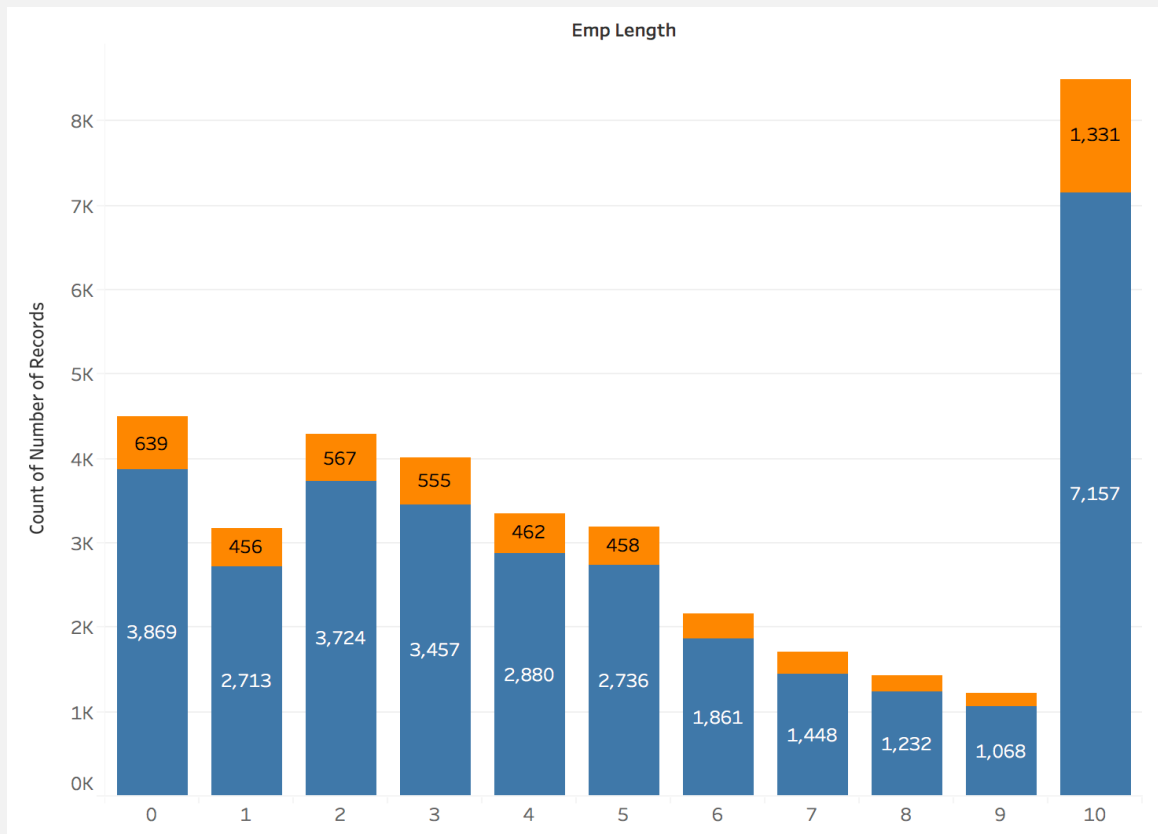
DATA ANALYSIS- DTI & REVOLVING UTILIZATION

Revolving utilization is a percentage value which tells how much percent of balance revolving credit forms to the total credit limit/total credit taken. Therefore we see that that applicants that have defaulted all had a higher revolving utilization

Also, we observe that applicants with higher dti tend to default more.
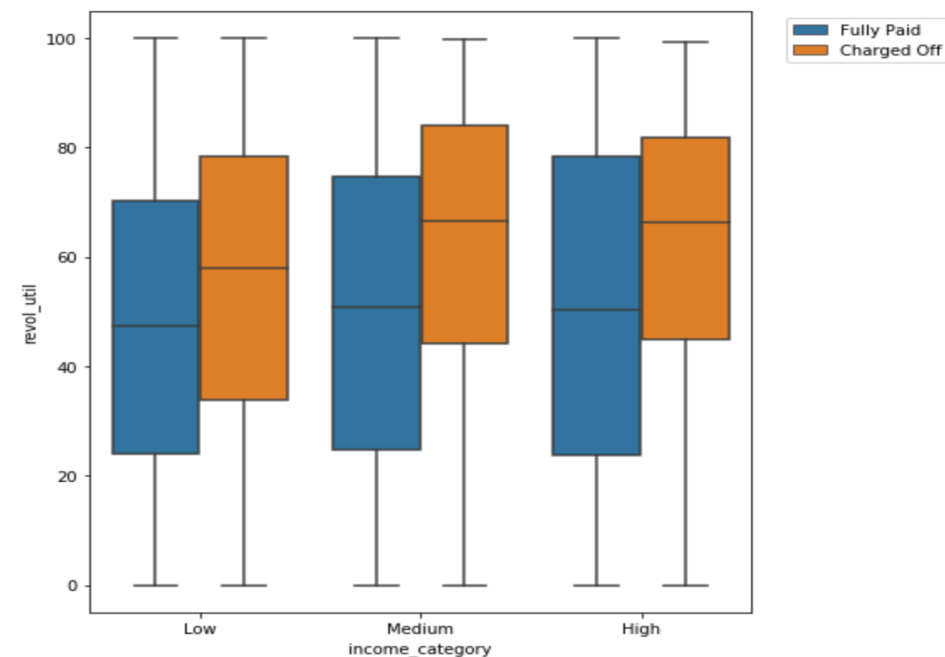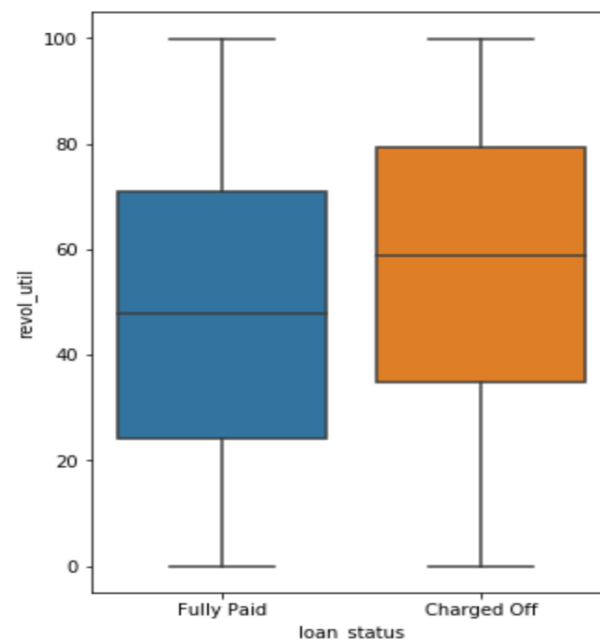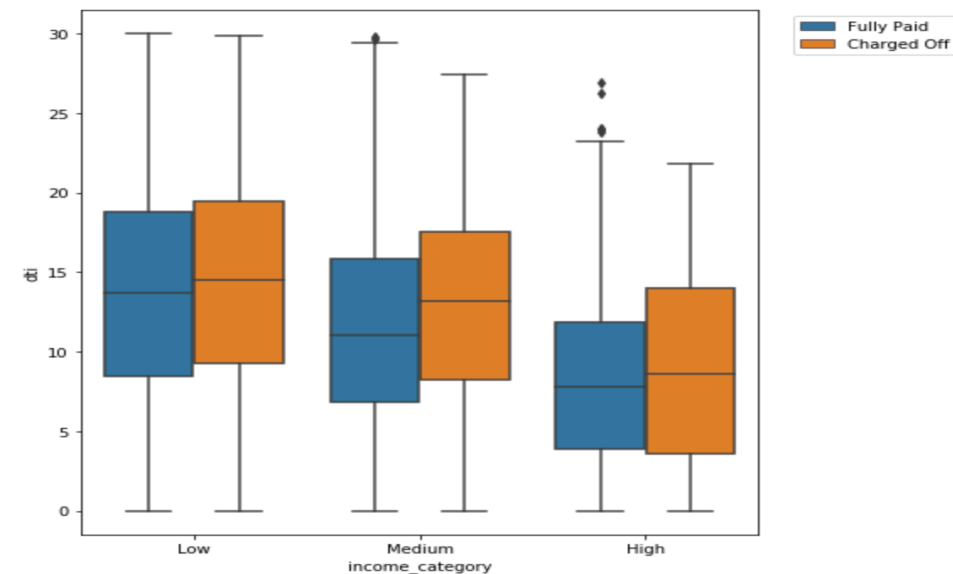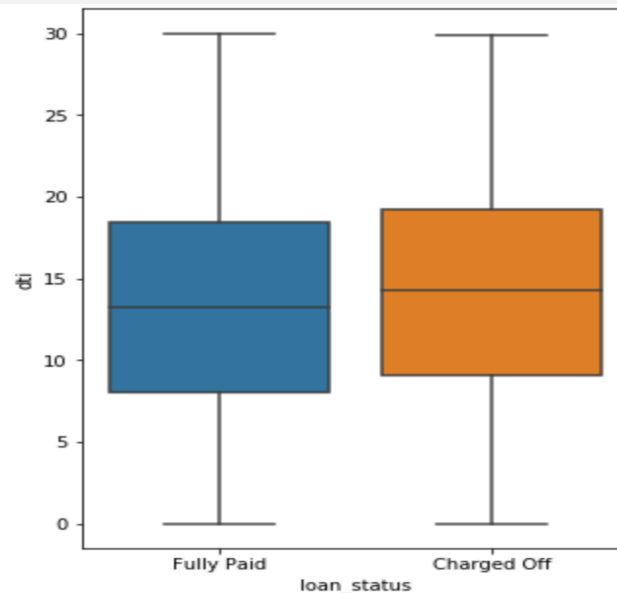
CONCLUSION

- Verification process should be made more stringent.

- Applicants with rented accommodation need to verified further.

- Applicants dti score and revolving utilization need to be checked and loans defaulters tend to have a higher score.

- Debt consolidation purpose of loan tend to default more.

- Low income applicants have higher possibility to default.

- Loans with a 60 month tenure tend to get defaulted more.