

Lending Club Case study

Analysis and Recommendation

Group member:

- Abhijeet Joshi

Lending Club

Business

- The company 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.

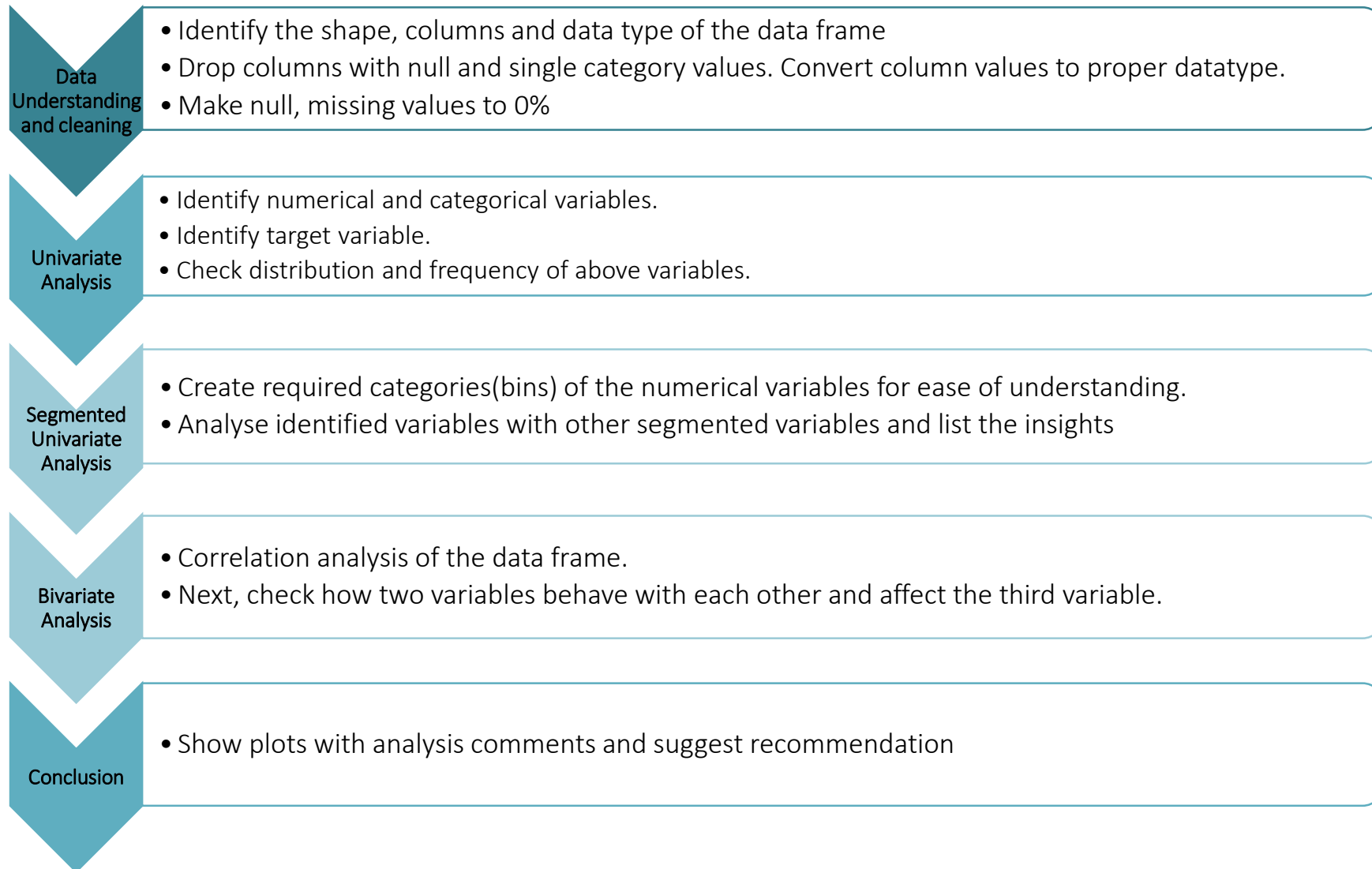
Objectives

- To reduce credit loss by defaulters.
- Identify the driving factor which are indicators of defaulters
- Utilizing previous knowledge to the identify driving factors.

Problem statement

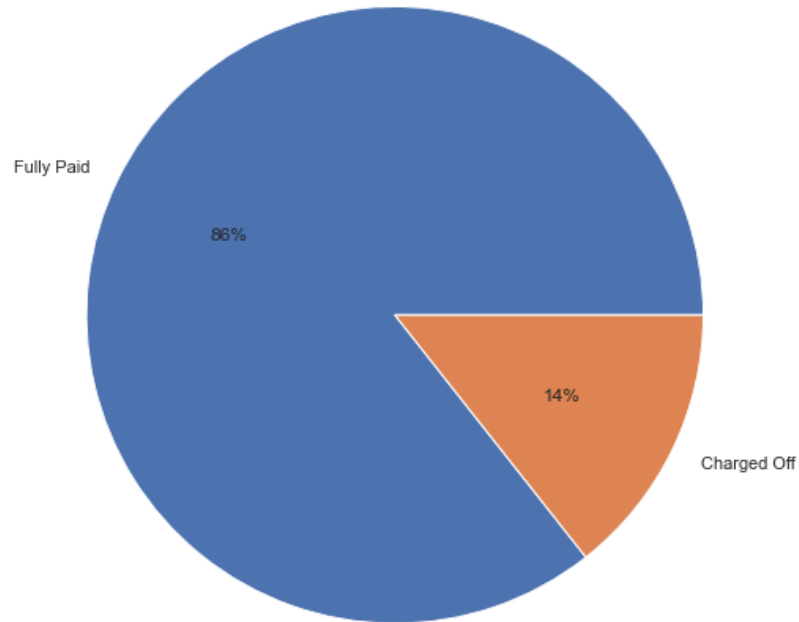
- Perform EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of default

Steps taken for analysis



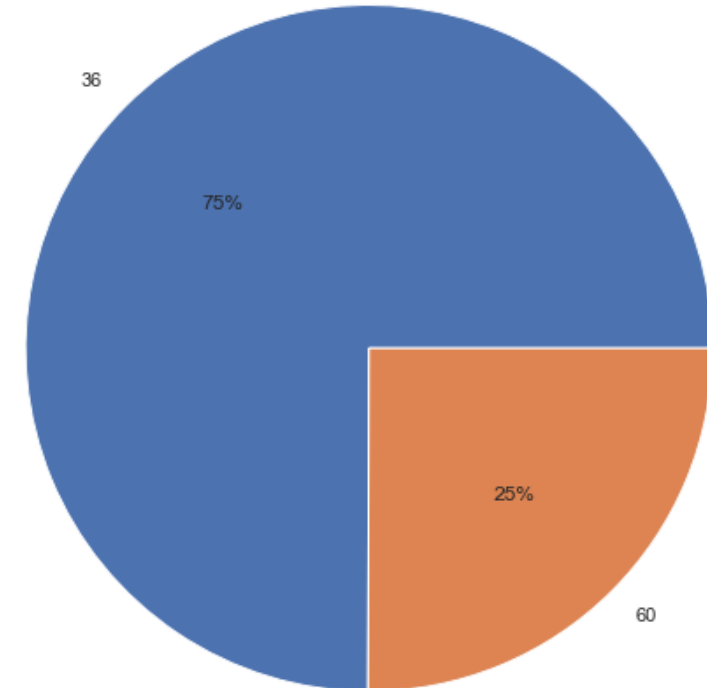
Lending analysis

Percentage of loans fully paid and charged off



- ❑ The pie chart shows that 14% of the total loans have been charged-off.
- ❑ It accounts for a loss of 15.67% of the total funded

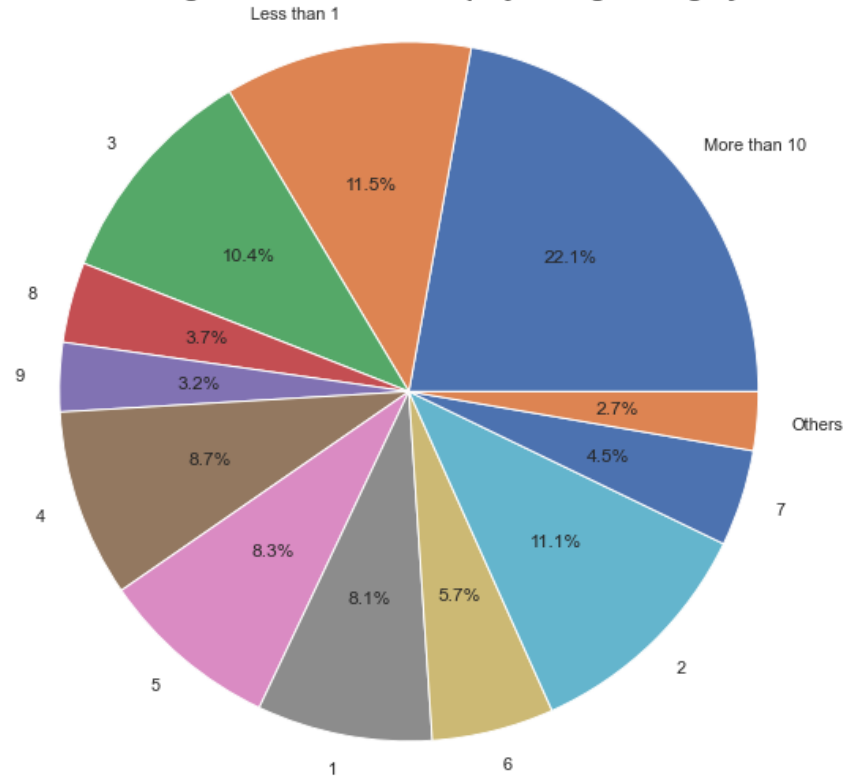
Percentage of loan terms



- ❑ The above pie chart shows that 75% of the total loans have a term of 36 months and 25% with 60 months term

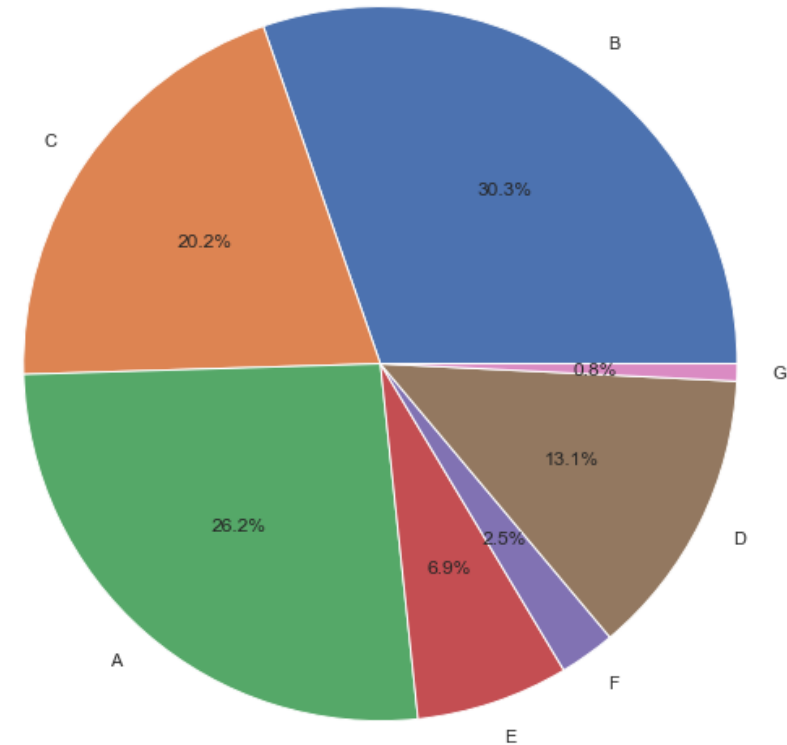
Lending analysis

Percentage of loans for each employee length category



- ❑ The pie chart shows that 22.1% of the total loans were given to consumers with 10+ years experience followed by 11.5% of total loans to 2 year experience consumers.
- ❑ Consumers with 1, 3, 4, 5 years of experience has more or less same distribution of loans.

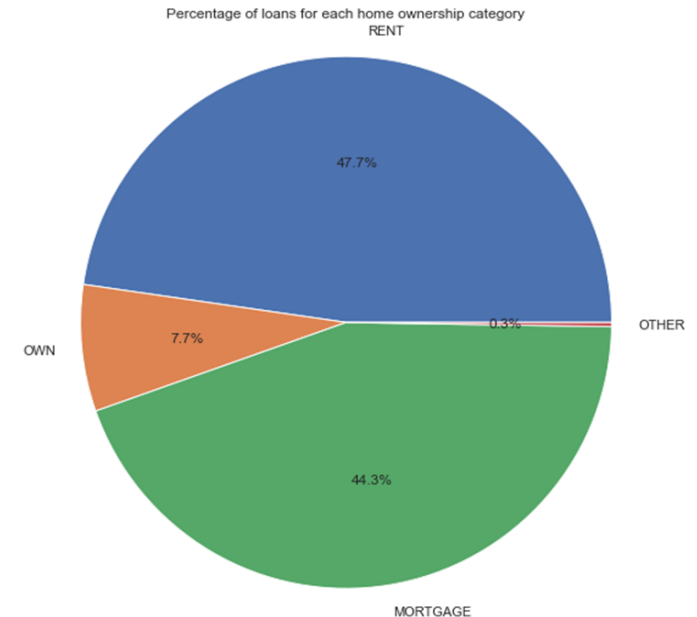
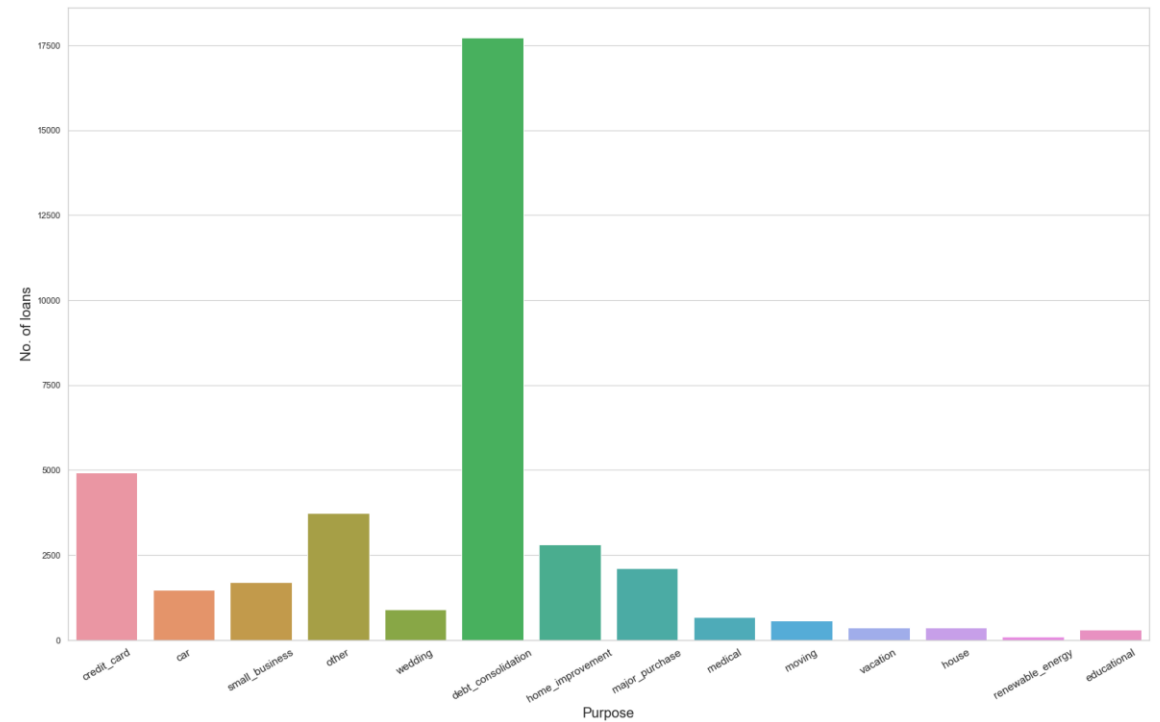
Percentage of loans for each grade



- ❑ The above pie chart shows that maximum % of loans were provided to grade B consumers i.e. 30.3% , followed by grade A consumers (26.2%).
- ❑ Note that grade A => low risk consumers, grade D => high risk. This shows that quality of providing loans is quite good.

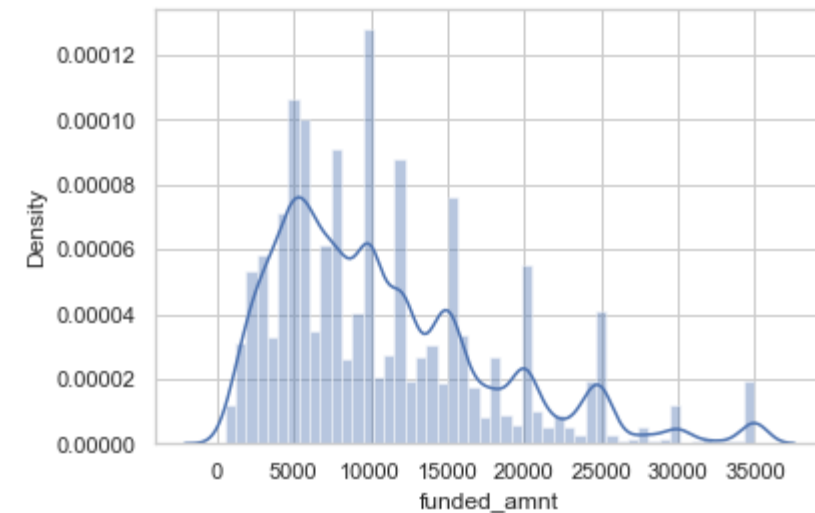
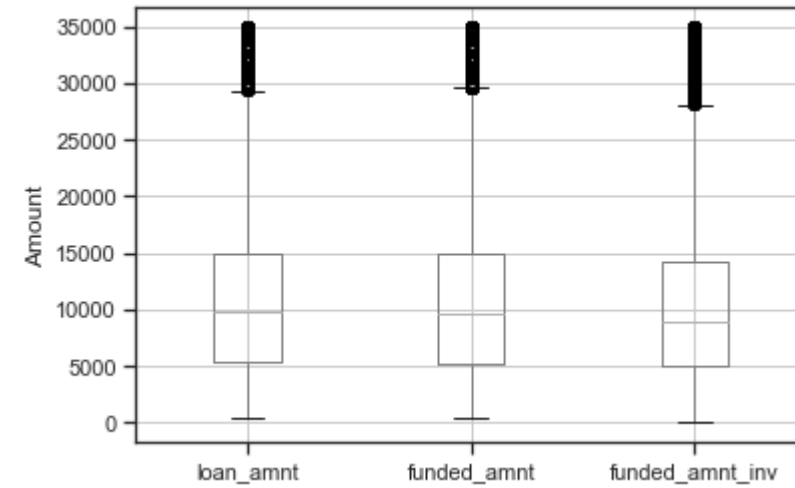
Lending analysis

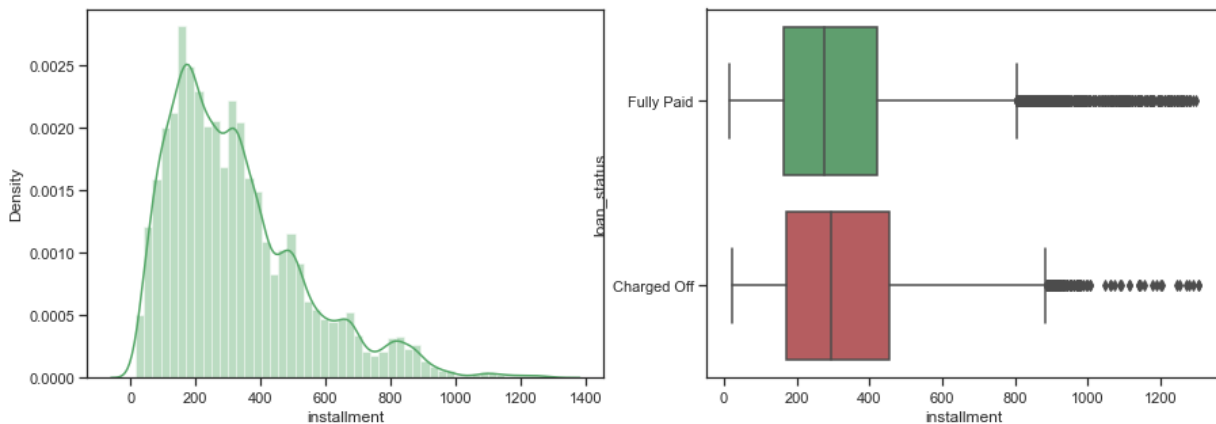
- ❑ The bar graph shows that most of the loans were taken for debt consolidation followed by credit card payment.
- ❑ The pie chart shows that 47.7% of the total loans were given to consumers with home ownership status as RENT followed by 44.3% of total loans to MORTGAGE
- ❑ OWN Home ownership status has very less % of loans.



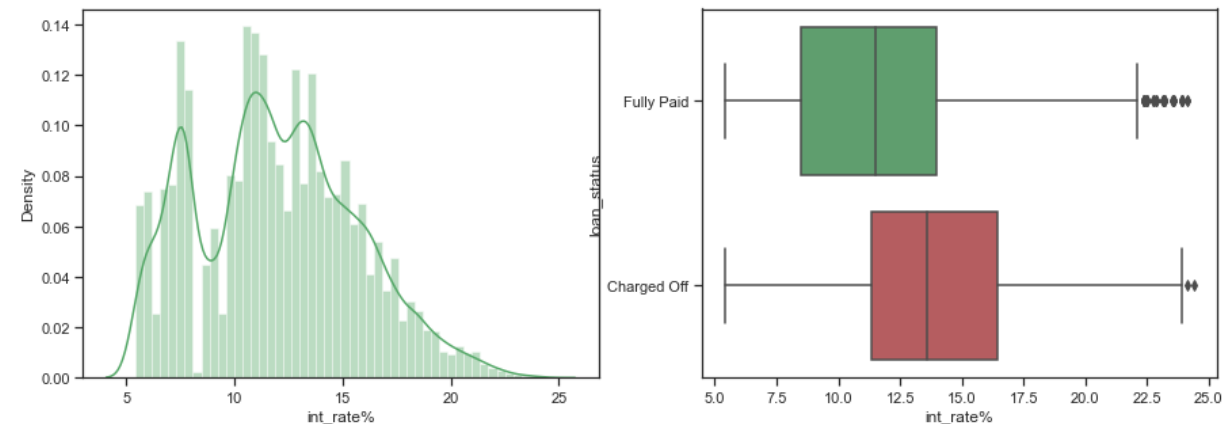
Lending analysis

- ❑ The boxplot shows that loan amount, funded amount and invested funds are all in the same range
- ❑ Funded amount is almost equal to loan amount.
- ❑ This implies, in most of the cases, the customers demand of required loan amount is fulfilled by Lending club.
- ❑ By the distribution plot we can see that most of loan amount is round 5000-10000 range.
- ❑ We can also observe that there are few high loan amount(outliers) but the density of high loan amount is less.



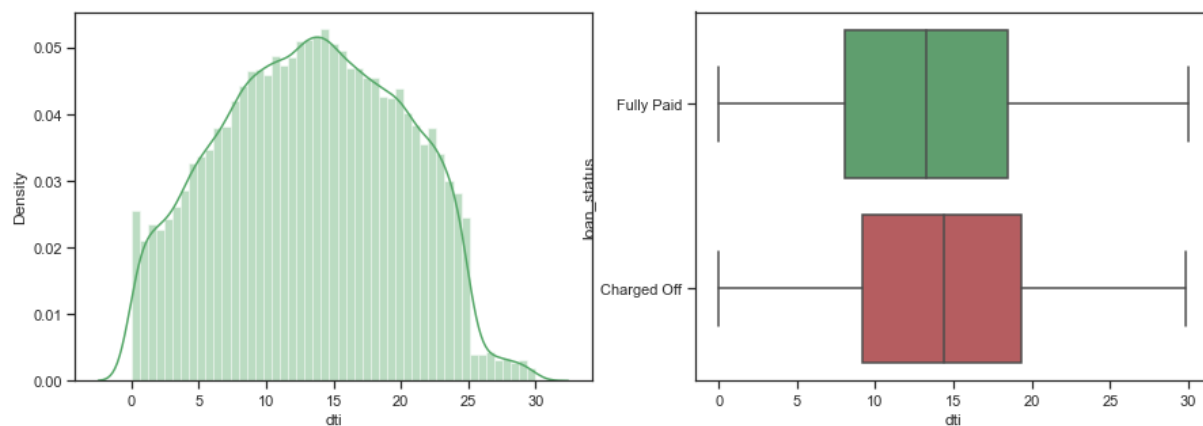


- The installments for most of the loans vary around 200-400 installment range.
- From box plot, we can see that the installments taken by customers is higher in charged-off loan compared to fully paid loans
- There are many higher installment loans(outliers) but the density is less

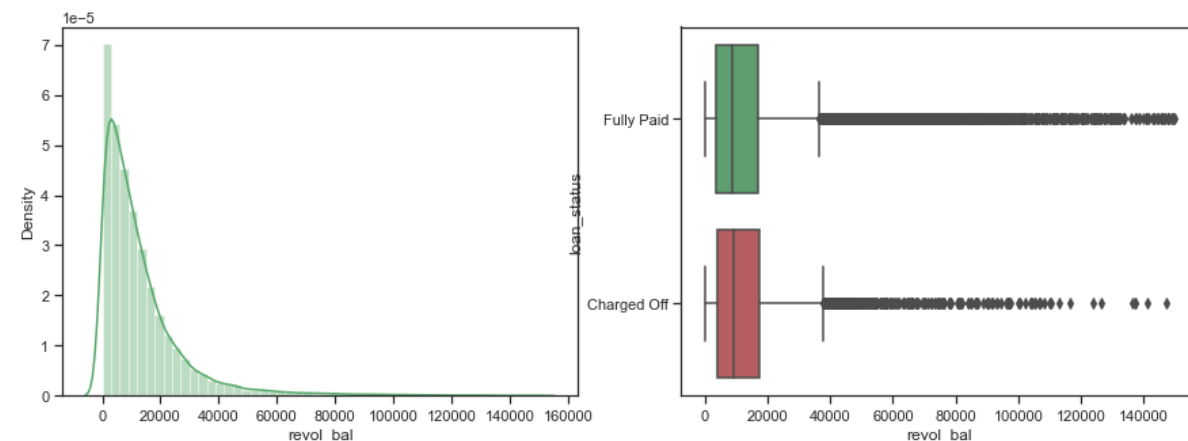


- The interest rates for most of the loans vary around 10-15 %
- From box plot, we can see that the fully paid loans have lower interest rates compared to charged of rates.
- Charged-off loan interest rate range: 11%~16%
- Fully paid loan interest rate range: 8%~13.5%

- The debt to income ratio for most of the loans vary around to 22.
- From box plot, we can see that the dti of customers is higher in charged-off loan compared to fully paid loans



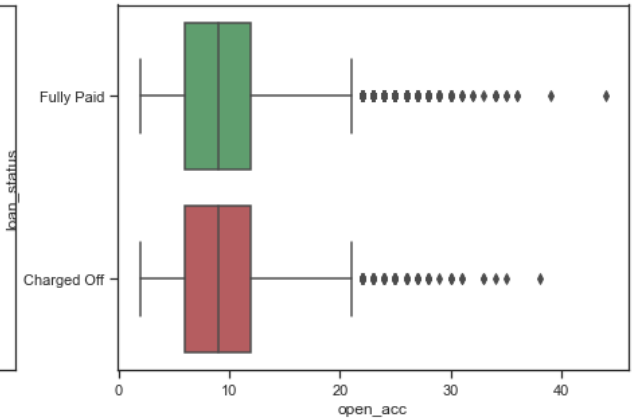
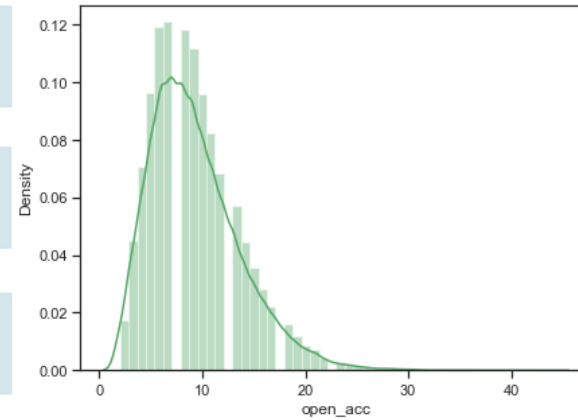
- The revolving balance of most of the loans vary around \$1000-\$18000
- From box plot, we can see that there are many customers with higher revolving



The number of existing loan accounts for most of the customers vary around 6 to 13 account range.

From box plot, we can see that there are few customers with higher open accounts(outliers).

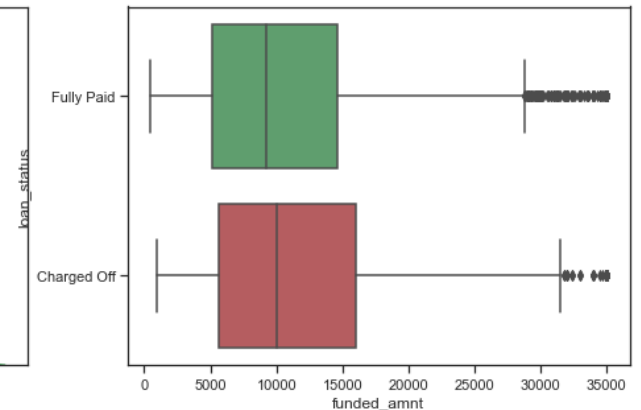
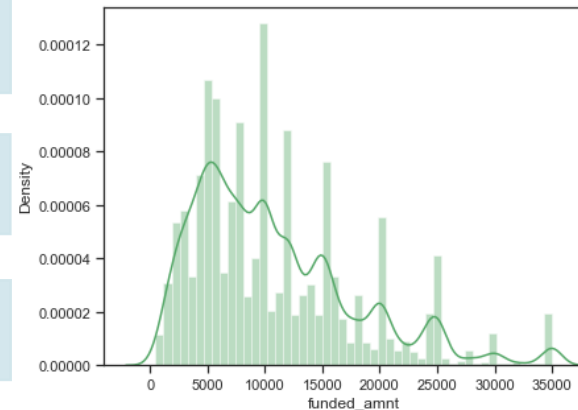
However observation is that open accounts for customers who fully paid and for charged-off customers are almost same.



The funded amount for most of the customers vary around \$4000 to \$13000.

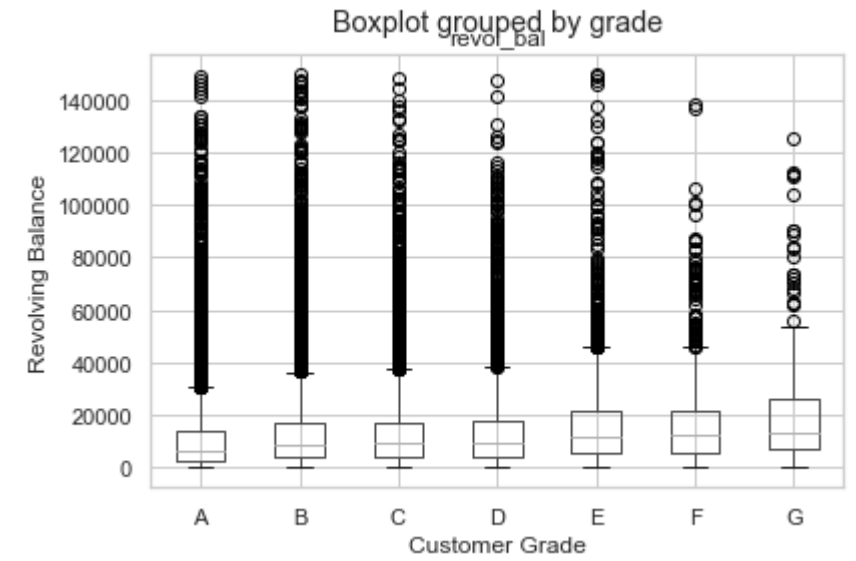
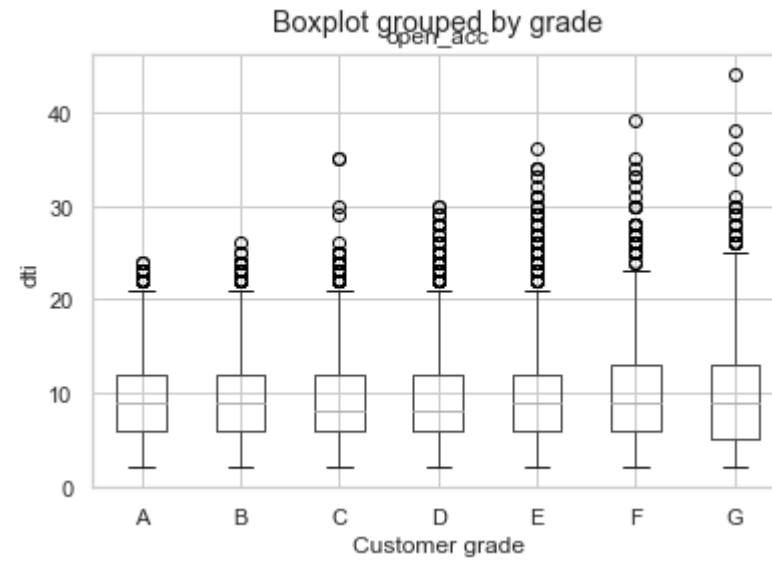
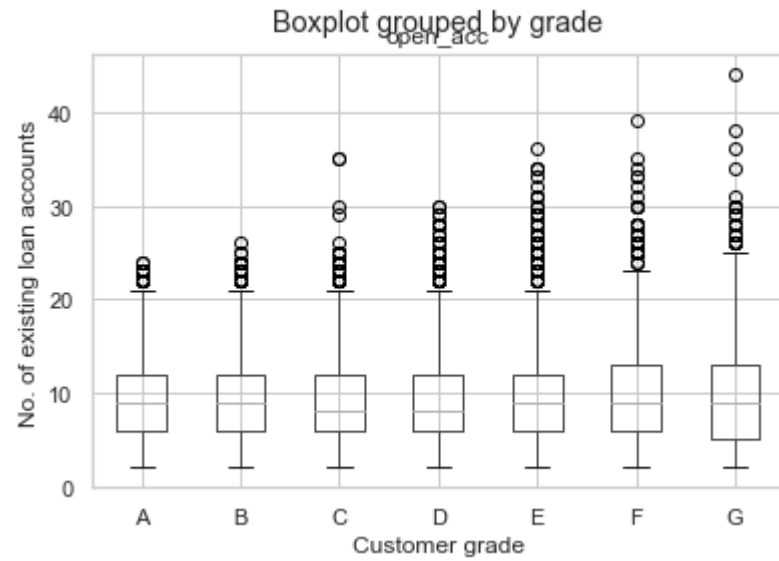
From box plot, we can see charged-off loans were funded more than the fully paid loans. This implies some faults in verification of customer financial background.

There are many outliers but the density of outliers is less.



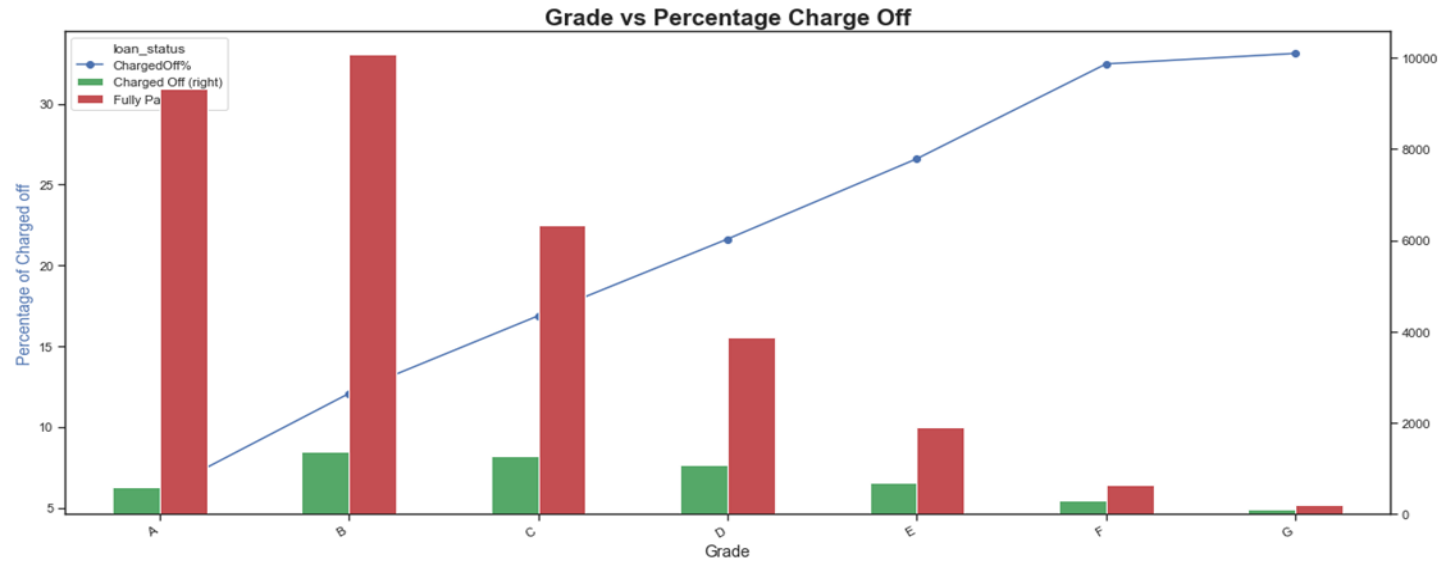
Next, lets check how customer attributes are affecting the loan attributes.

Analysis based on Customer grade

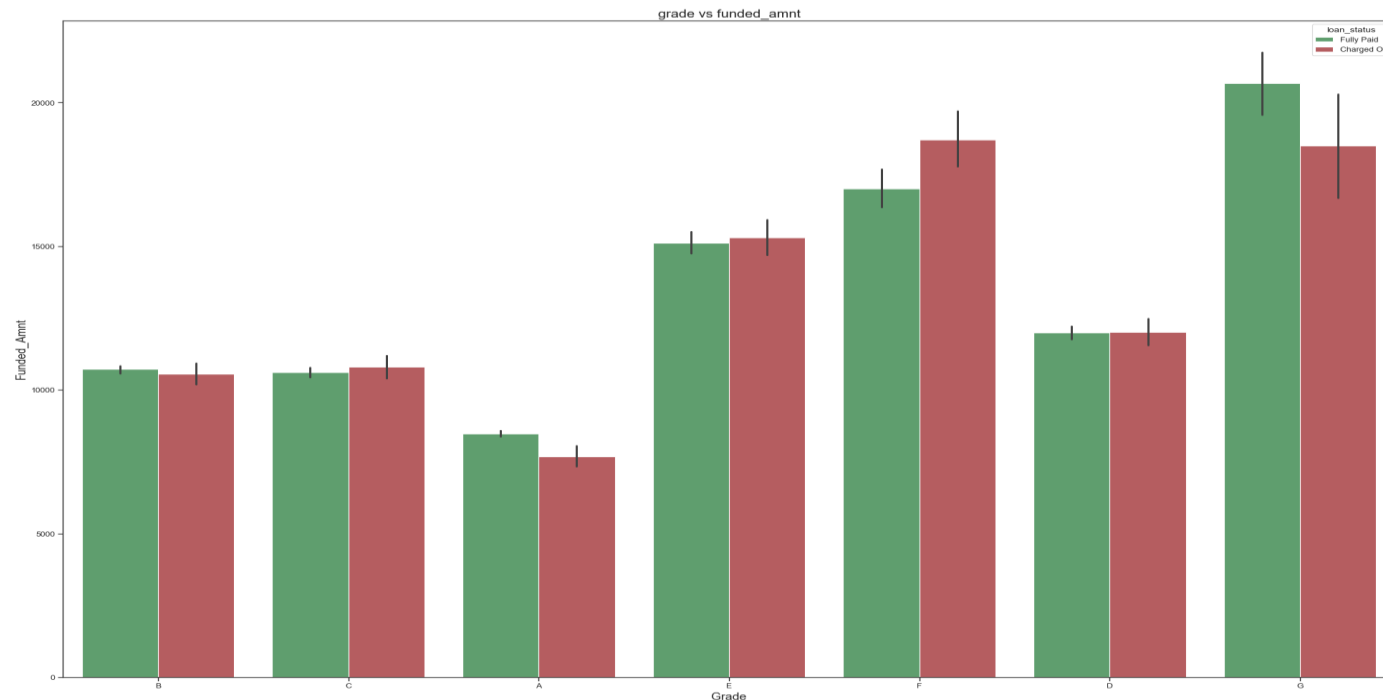


- From the customer grade vs open account box plot, we can see that grade F and grade G customers are having higher open accounts. Grade C and Grade D have lower open accounts.
 - From the customer grade vs dti box plot, we can see that grade F and grade G customers are having higher debt to income ratio.
 - From the customer grade vs dti box plot, we can see that grade E, grade F and grade G customers are having higher revolving balance.
- Lower grade customers(E, F, G) are risky customers(tend to be charged-off)

Analysis based on Customer grade

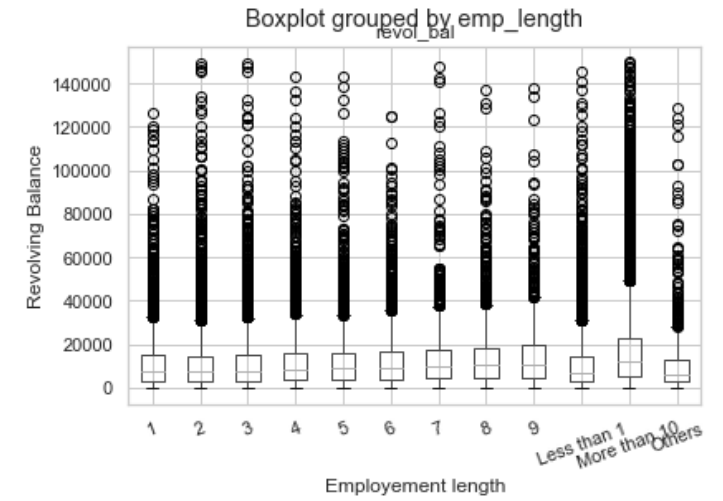
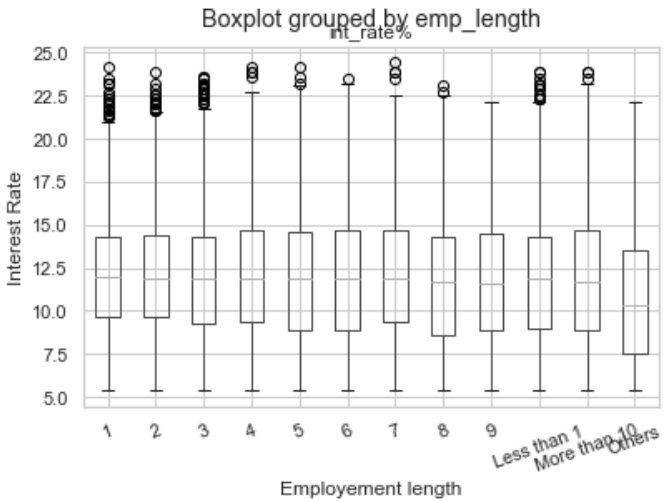


The 1st bar plots here confirm the analysis before. The plot shows that the charge-off percentage keeps on increasing as the grade of customers decrease



The 2nd bar plots shows that the funded amount is highest in the low graded customers. Grade E, F, and G have highest ratio of funded amount.

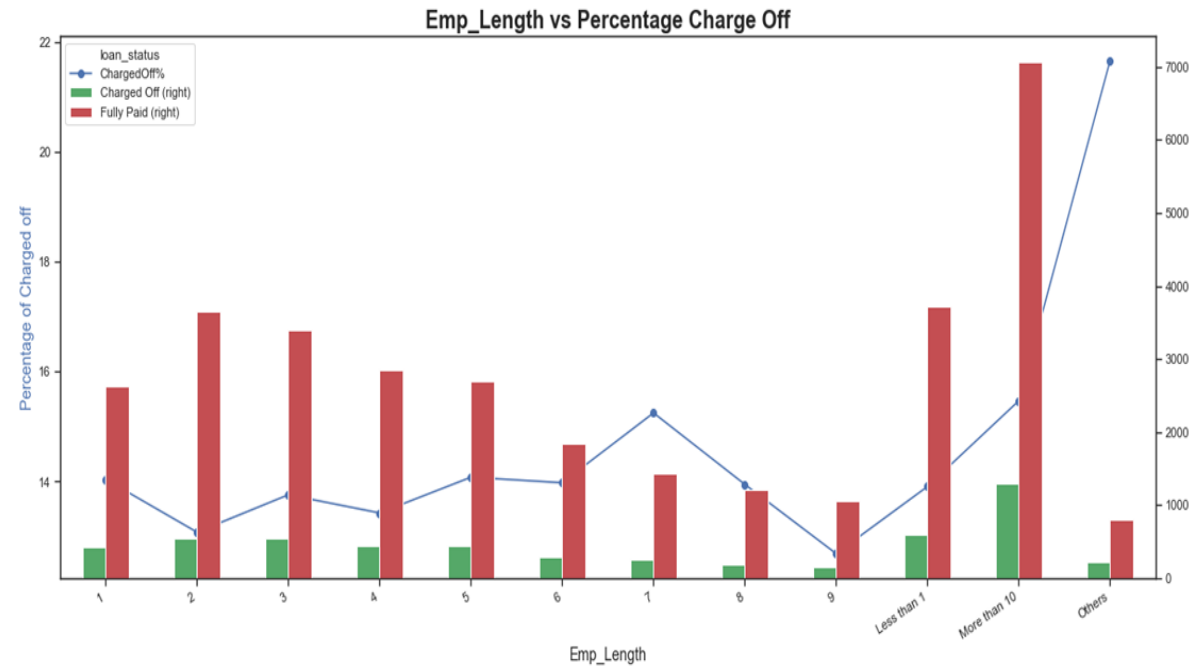
Analysis based on customer's Employment length



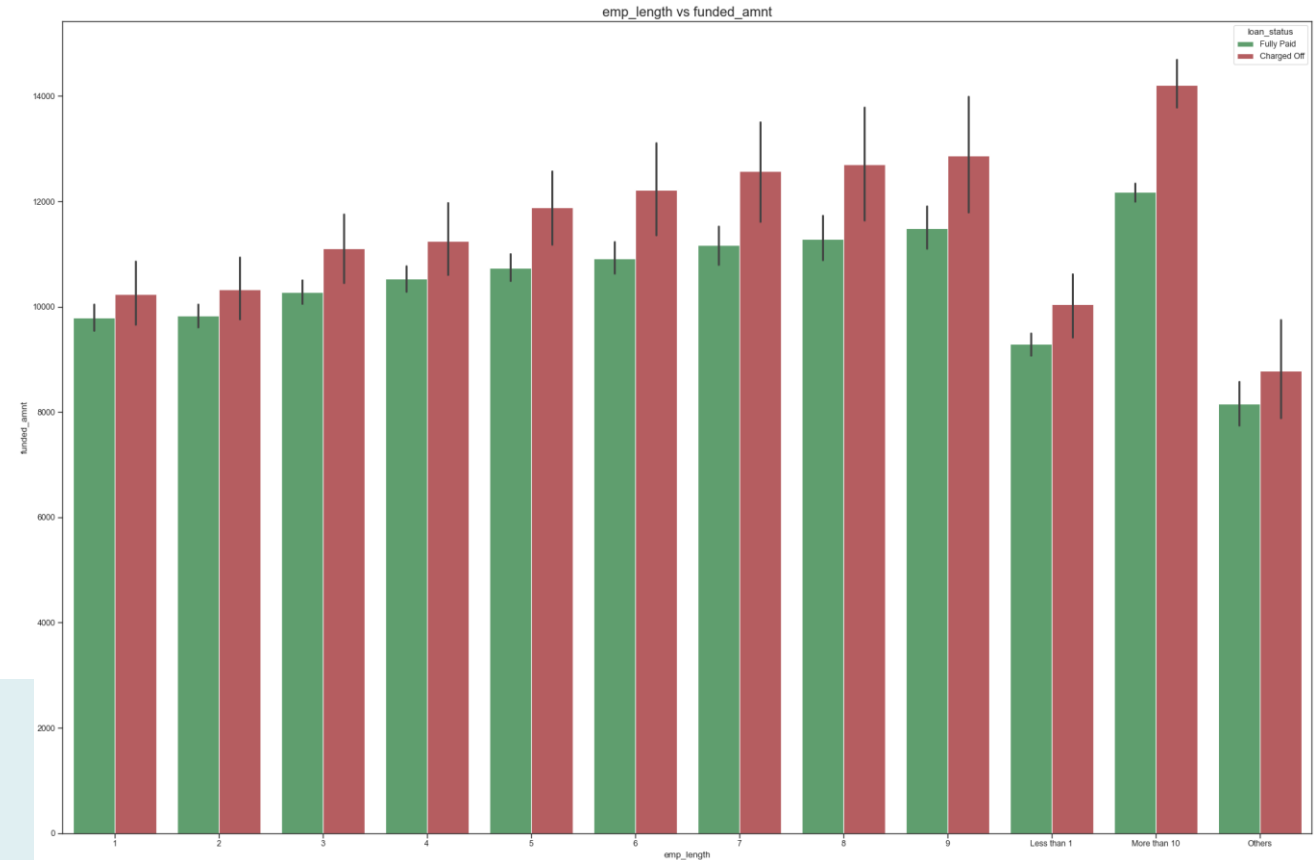
- From the customer grade vs open account box plot, we can see that grade F and grade G customers are having higher open accounts. Grade C and Grade D have lower open accounts.
- From the customer grade vs dti box plot, we can see that grade F and grade G customers are having higher debt to income ratio.
- From the customer grade vs revolving balance box plot, we can see that grade E, grade F and grade G customers are having higher revolving balance.

→ Lower grade customers(E, F, G) are risky customers(tend to be charged-off)

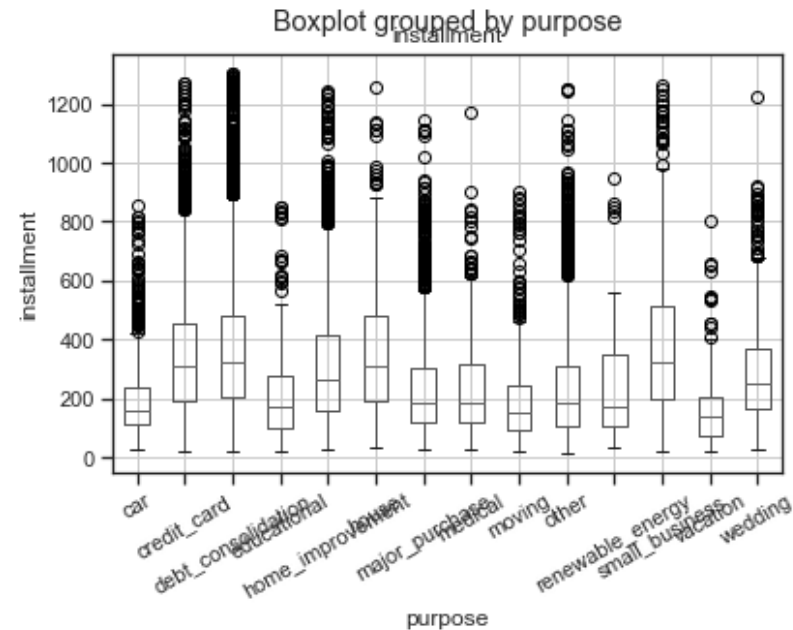
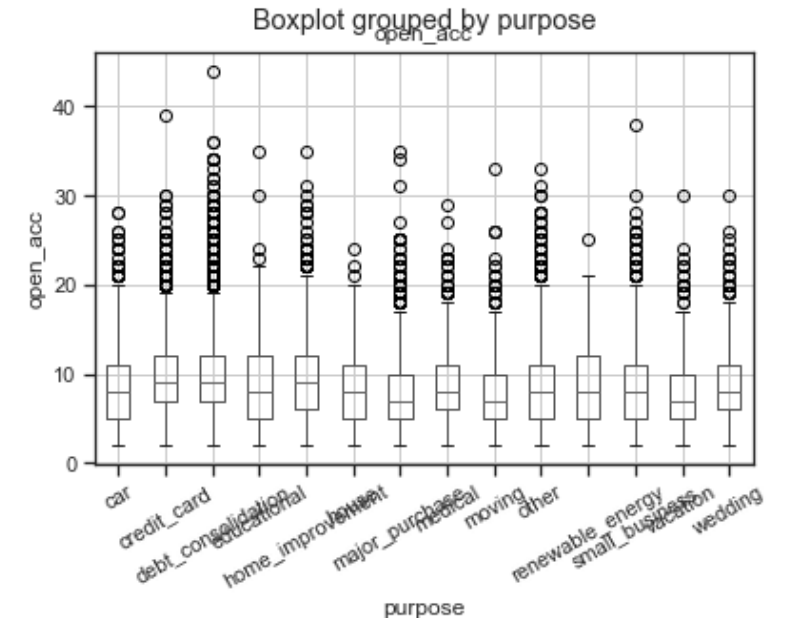
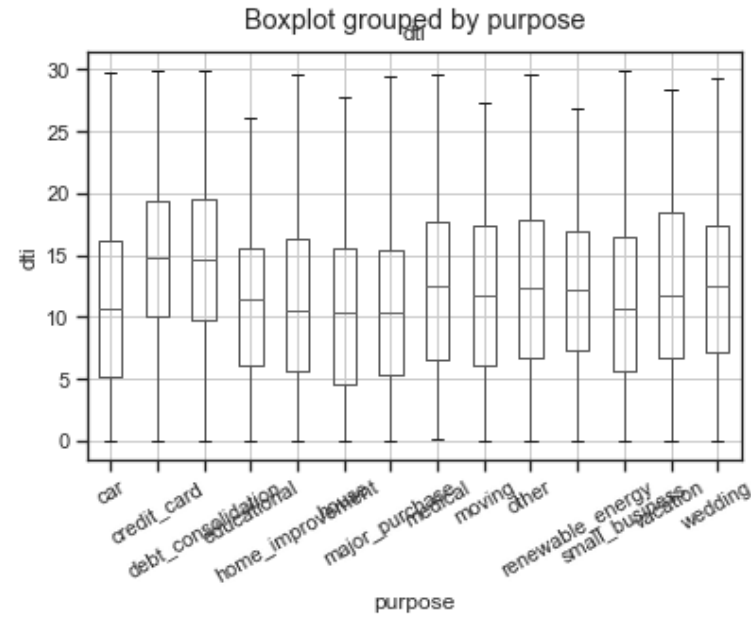
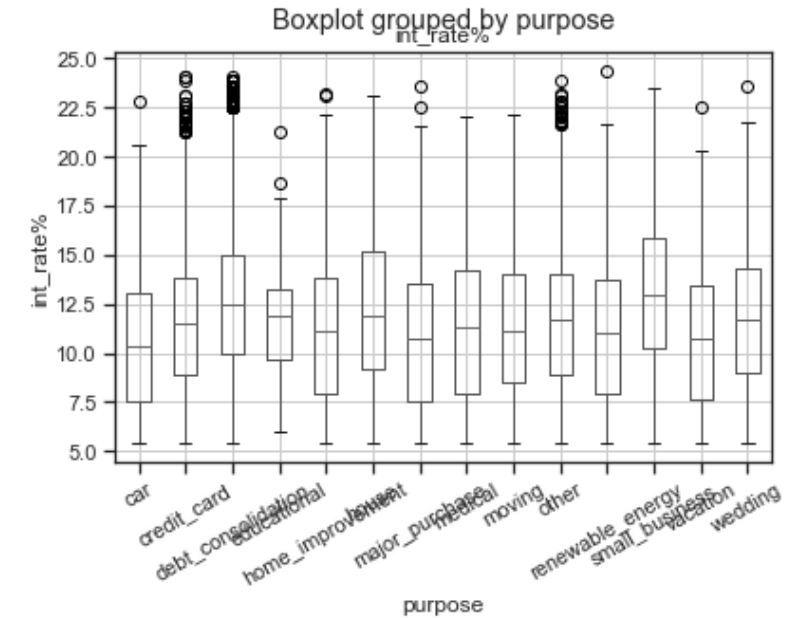
Analysis based on customer's Employment length



- The charge-off percentage is higher in 10+ year bar as most of the loans were given with this range.
- However, the charged-off loans decrease with increasing employment years till 9 years.
- Most of the fully paid loans are in 10 and 10+ year of customer employment.

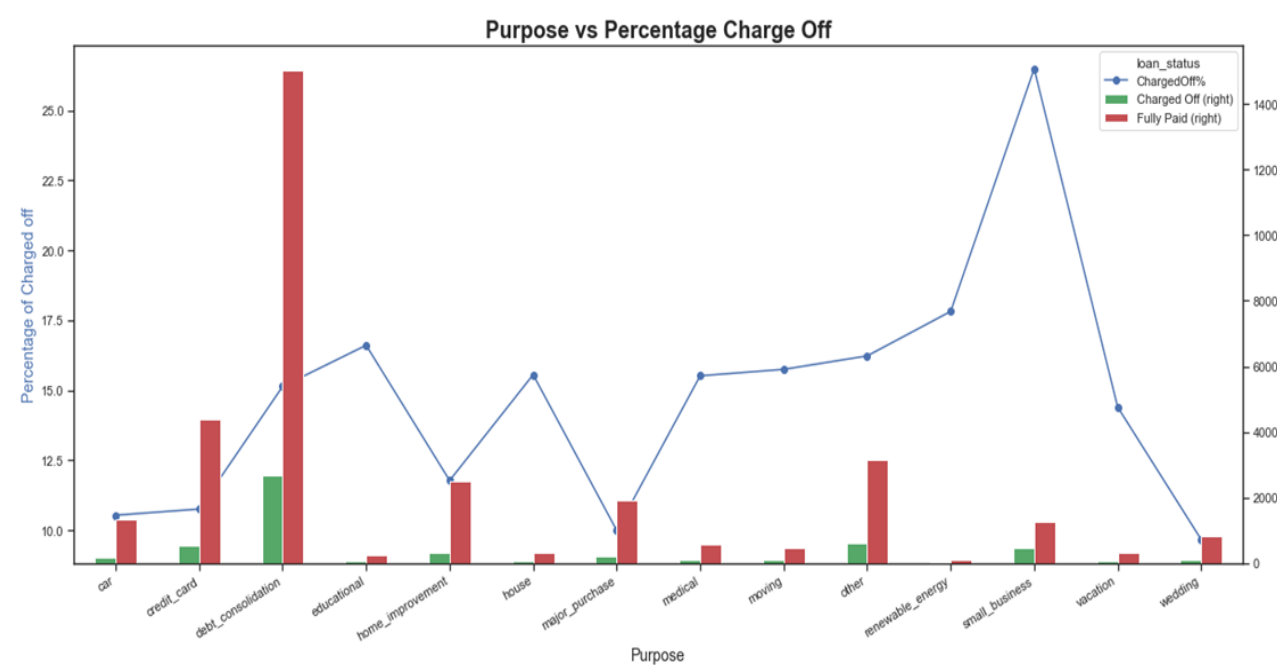


Analysis based on customer's loan purpose



- From the customer grade vs open account box plot, we can see that grade F and grade G customers are having higher open accounts. Grade C and Grade D have lower open accounts.
- From the customer grade vs dti box plot, we can see that grade F and grade G customers are having higher debt to income ratio.
- From the customer grade vs dti box plot, we can see that grade E, grade F and grade G customers are having higher revolving balance.

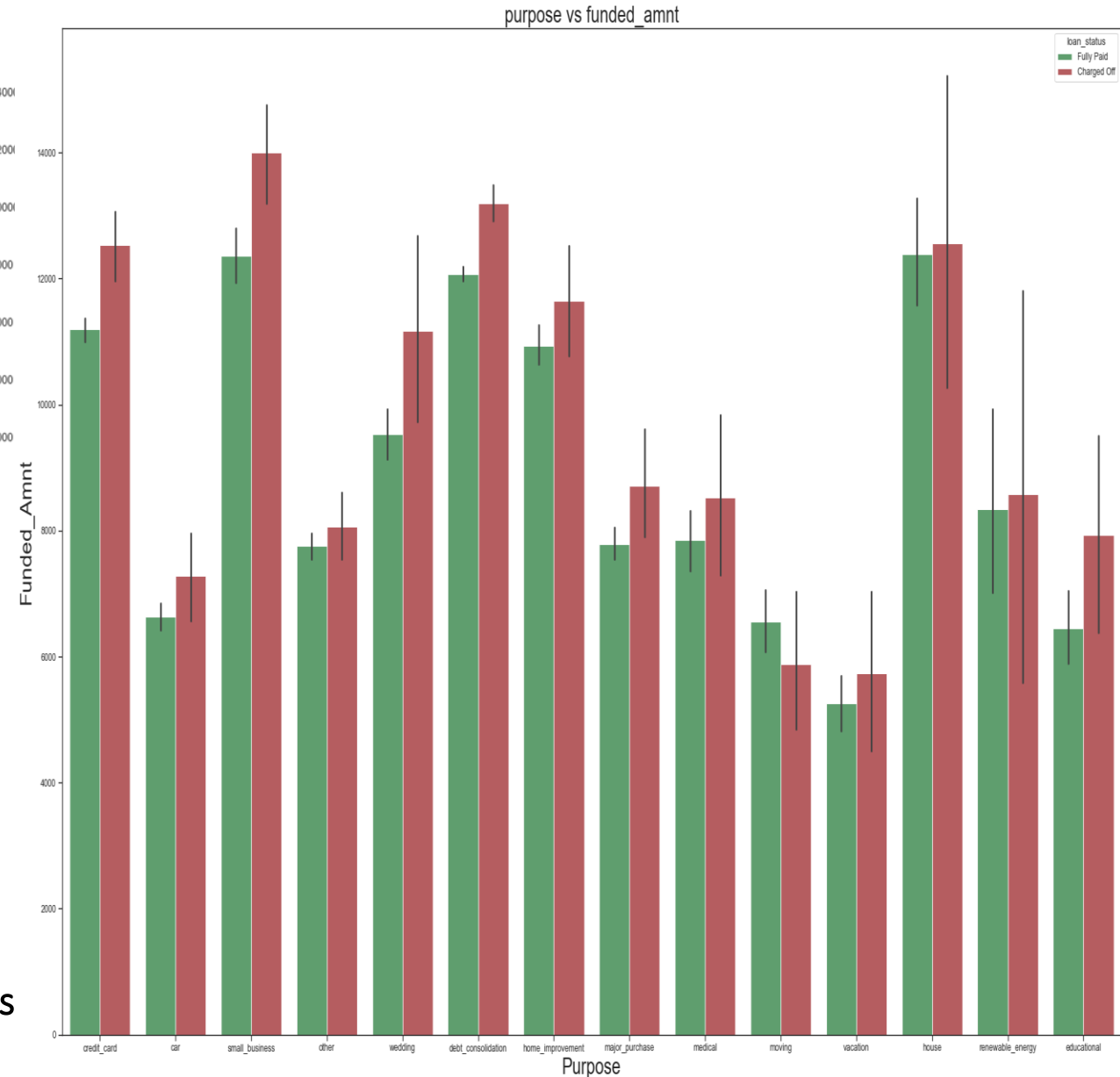
Analysis based on customer's loan purpose



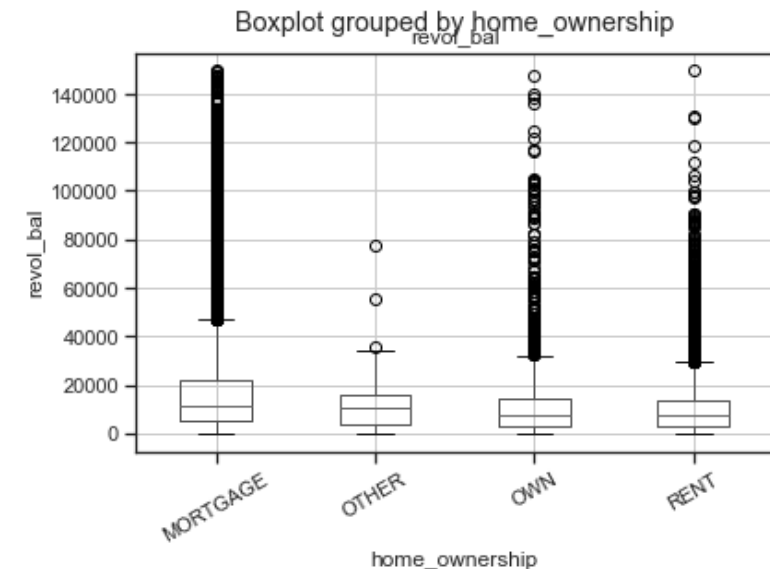
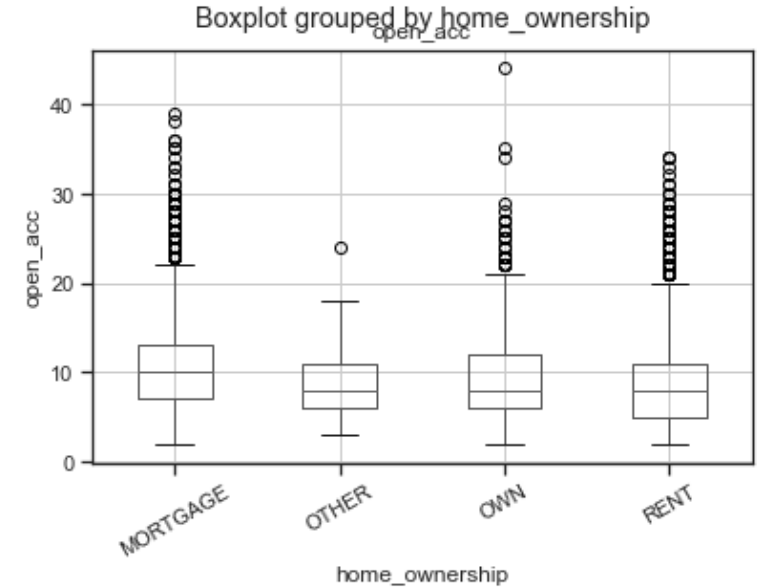
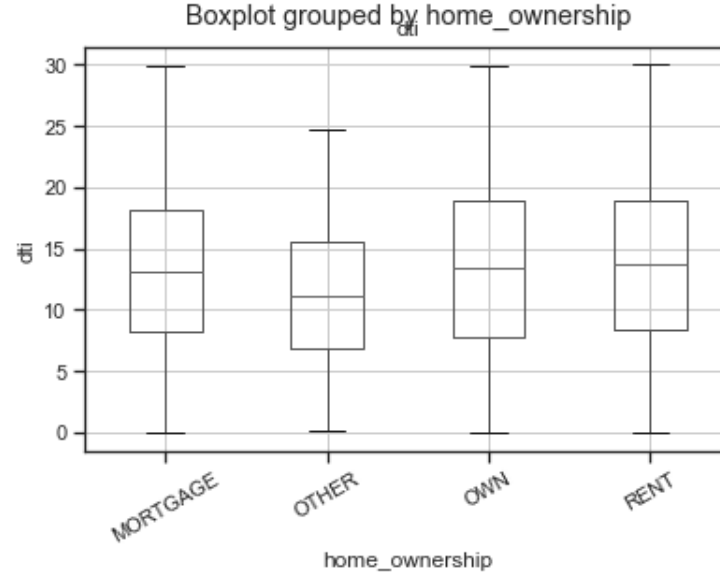
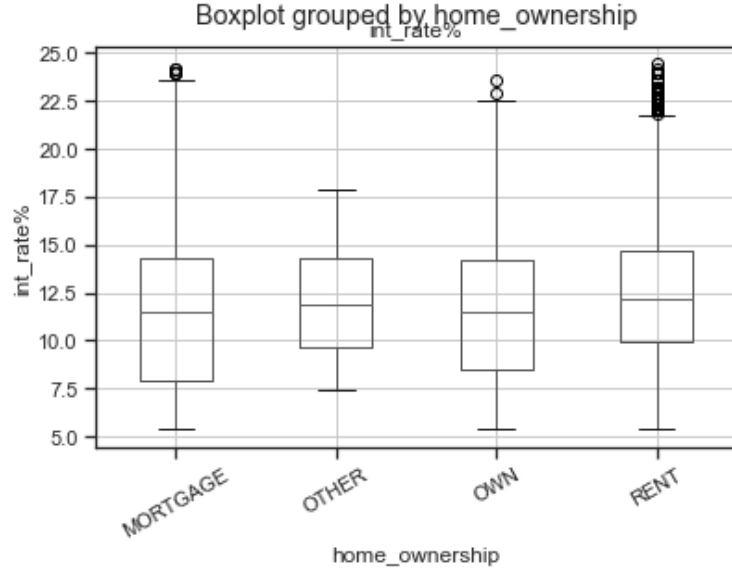
From bar plot, we observe that customer's purpose for 'small business', 'debt_consolidation' and 'credit card' have most of the loan.

However, purpose with *small_business* has highest percentage of charged-off.

Therefore, lower grade customers(E, F, G) and *small_business* purpose customers are of high risk.

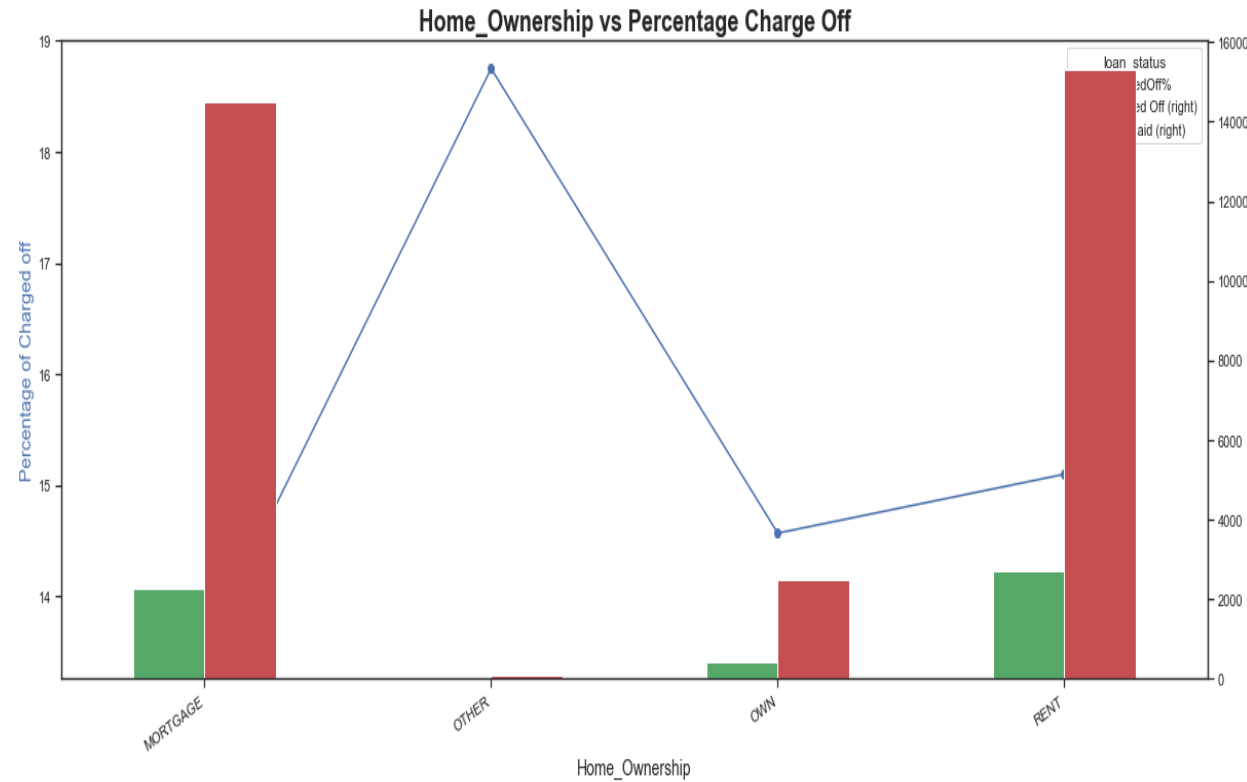


Analysis based on customer's home ownership

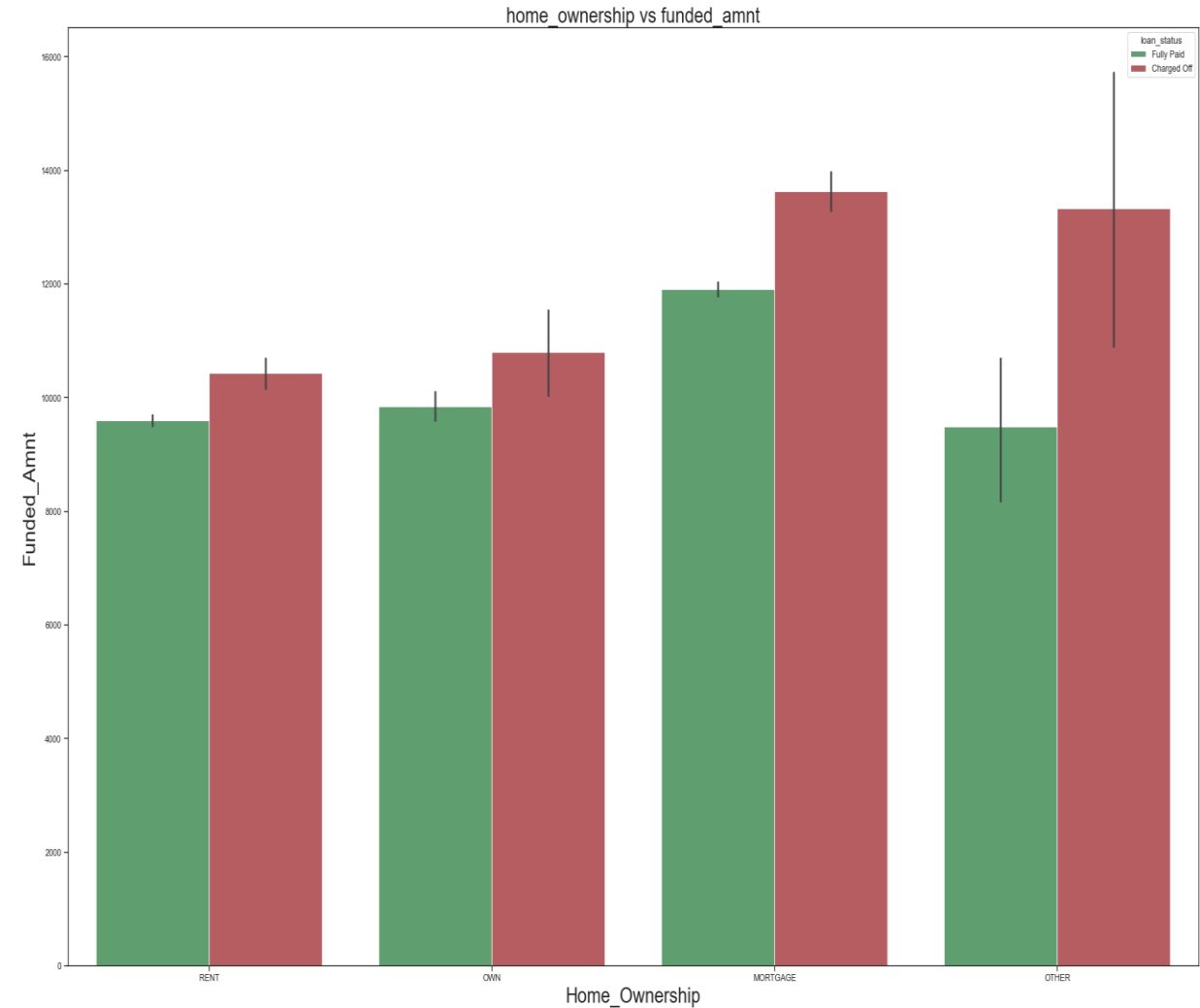


- From the home_ownership vs int_rate% box plot, we can see that ownership with RENT has higher interest rate median compared to others.
- From the home_ownership vs dti box plot, we can see that ownership with RENT and OWN have higher dti ratio median.
- From the home_ownership vs open_acc and home_ownership vs revol_bal box plot, MORTGAGE has higher open accounts and revolving balancing respectively.

Analysis based on customer's home ownership

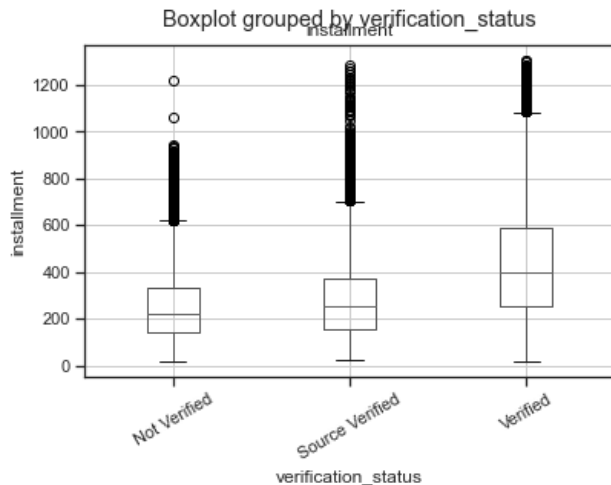
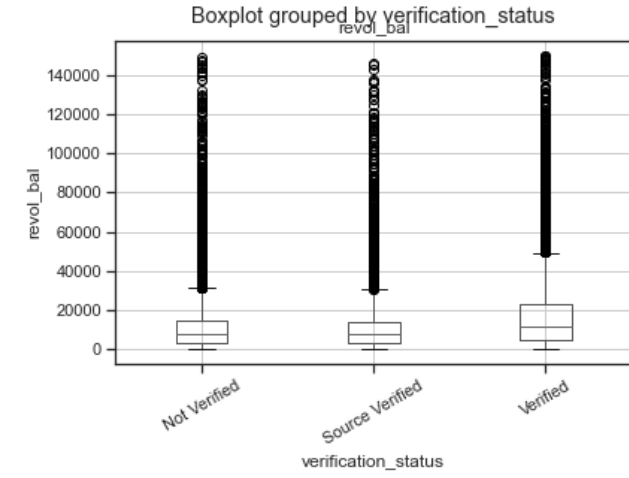
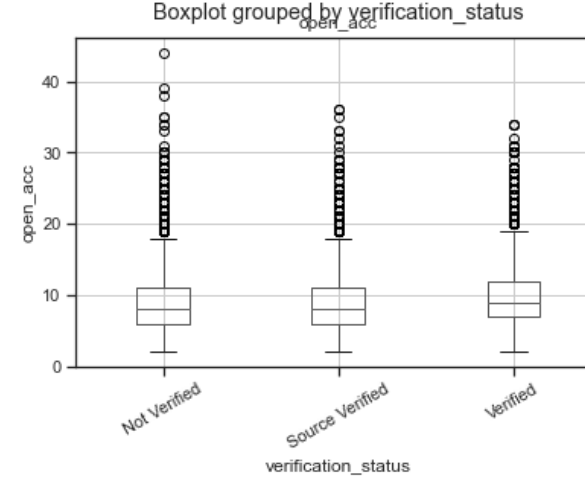
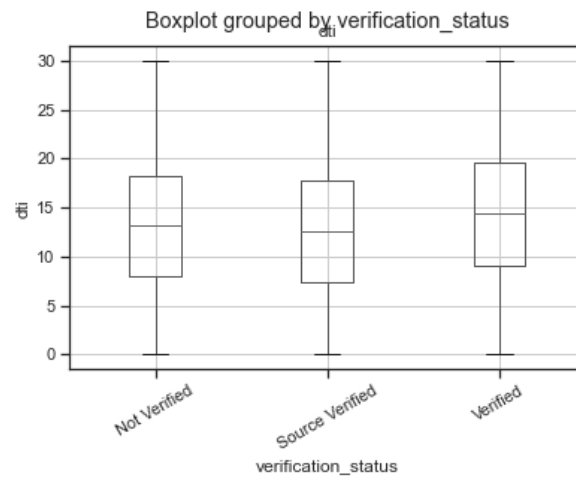
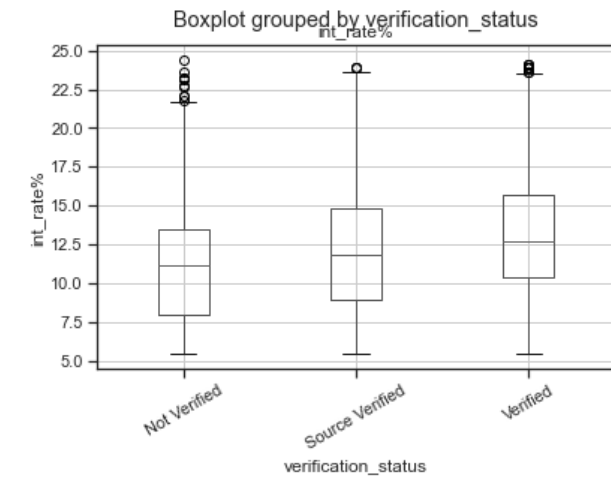


From bar plot, we observe that customer's home ownership of RENT has higher charged-off percentage.



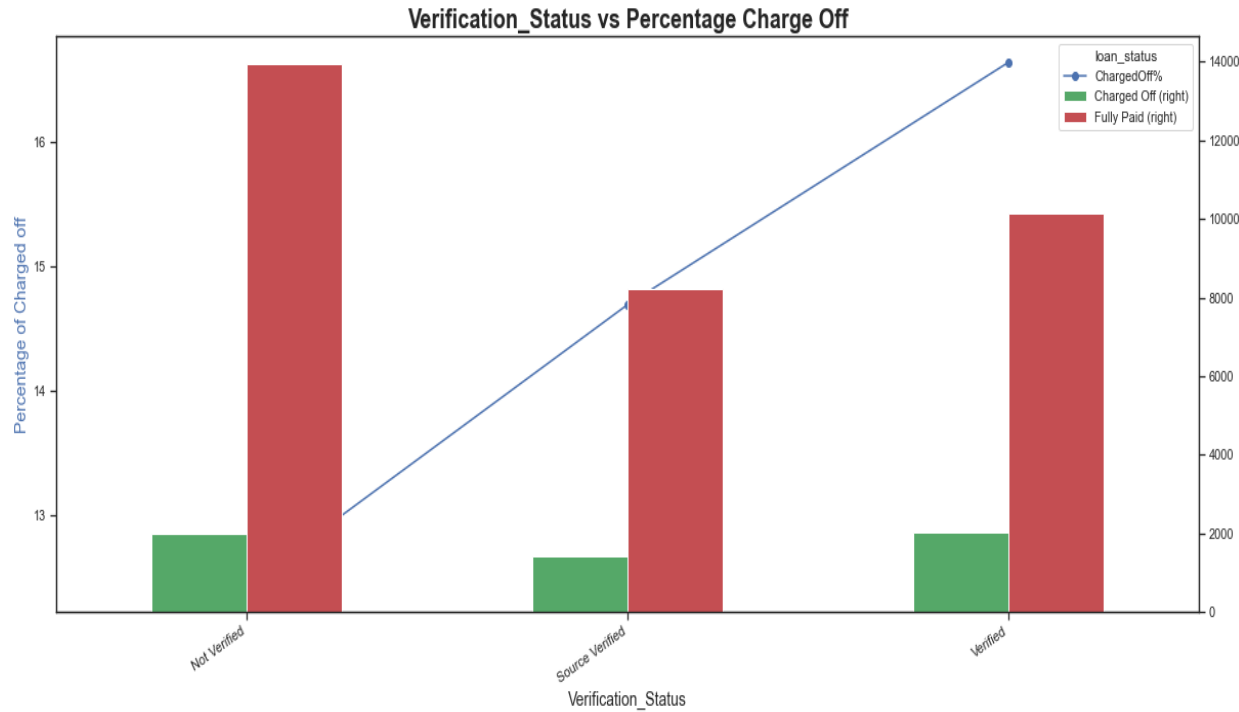
Therefore, customers with RENT home ownership tend to be high risk.

Analysis based on customer's verification status

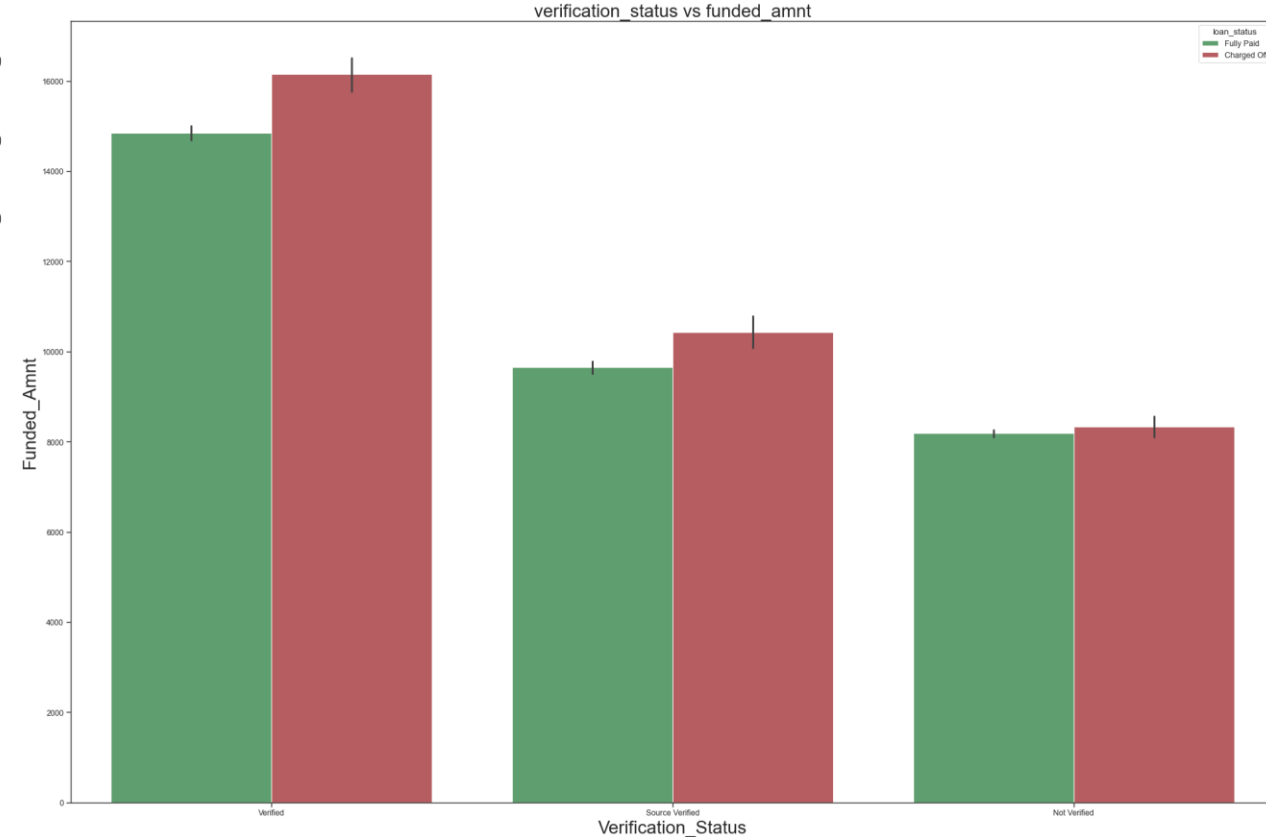


- From the verification status vs int_rate% box plot, we can see that verified customers have higher interest rate median compared to others.
- From the verification status vs dti box plot, we can see that verified customers have higher dti ratio median compared to others.
- From the verification status vs open_acc and verification status vs revol_bal box plot, we can see that verified customers have higher open account median and higher revolving balance median compared to others.

Analysis based on customer's verification status

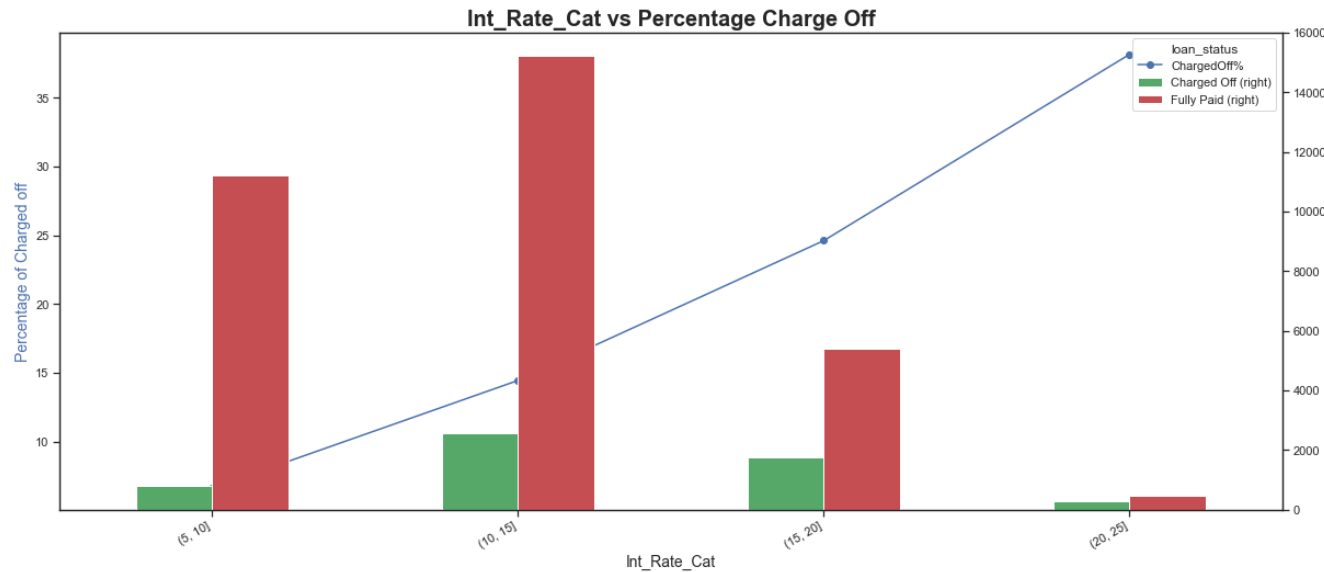


- From bar plot, we observe that verified customers have been funded and charged off funds are also higher.
- Also we can observe that verified customers have highest charged-off percentage.



Therefore, VERIFIED customers tend to be of high risk. This also implies, verification of customer financial background is not done correctly.

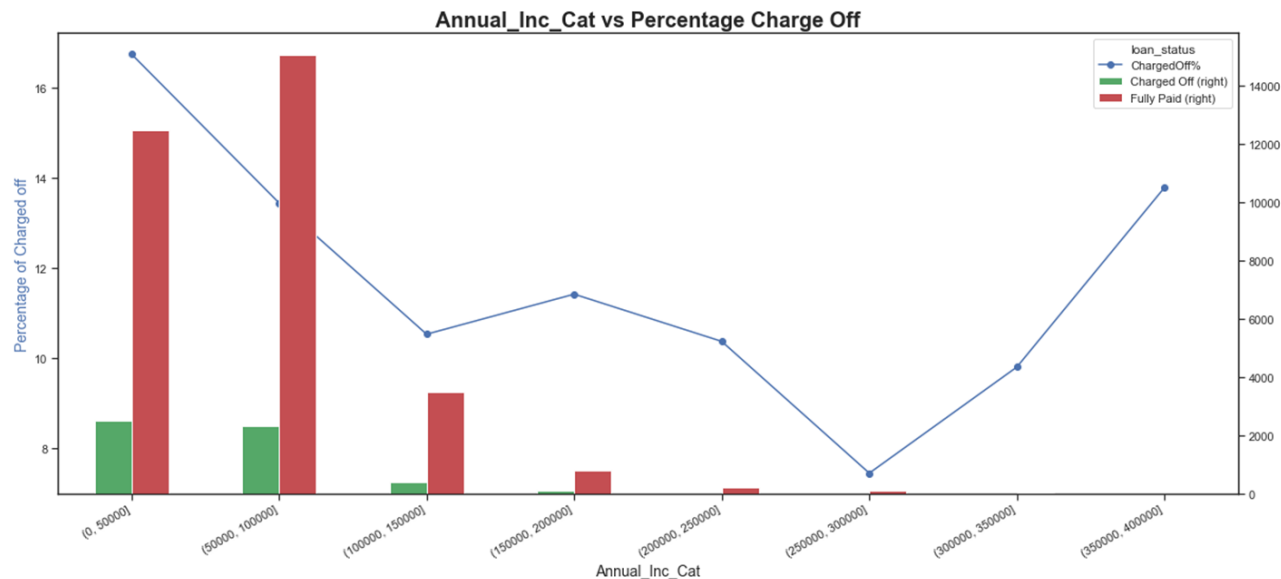
Analysis



Based on Interest rate provided

Most of the loans are in 5%-15% interest.

Also the percentage of charged off increases with increasing interest rate.

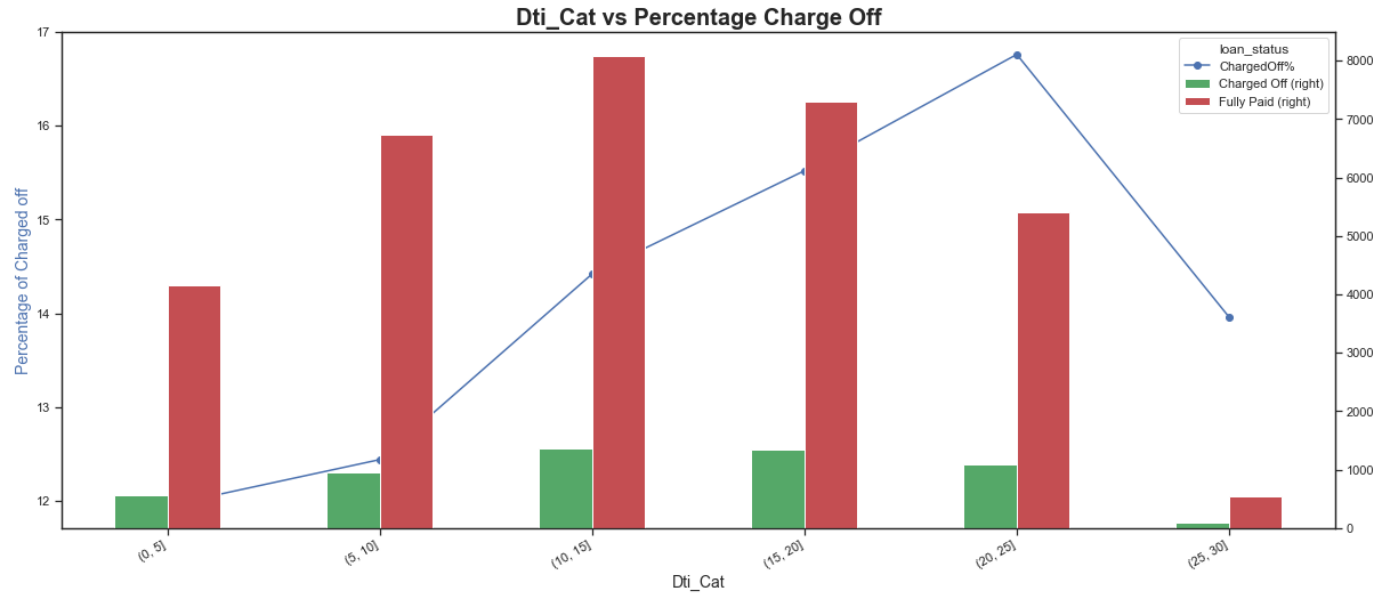


Based on Annual income

Most of the customers have annual income in the range \$50000 to \$100000

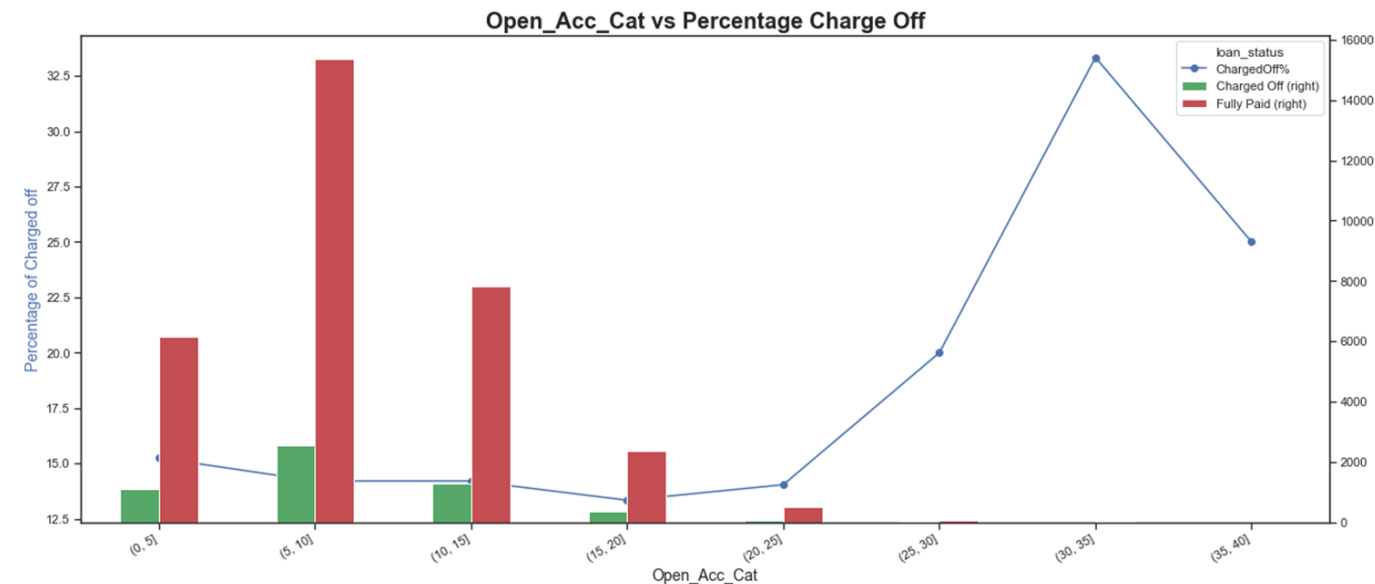
Also the percentage of charged off decrease with increasing annual income.

Analysis based on



Based on customer's Debt to Income ratio

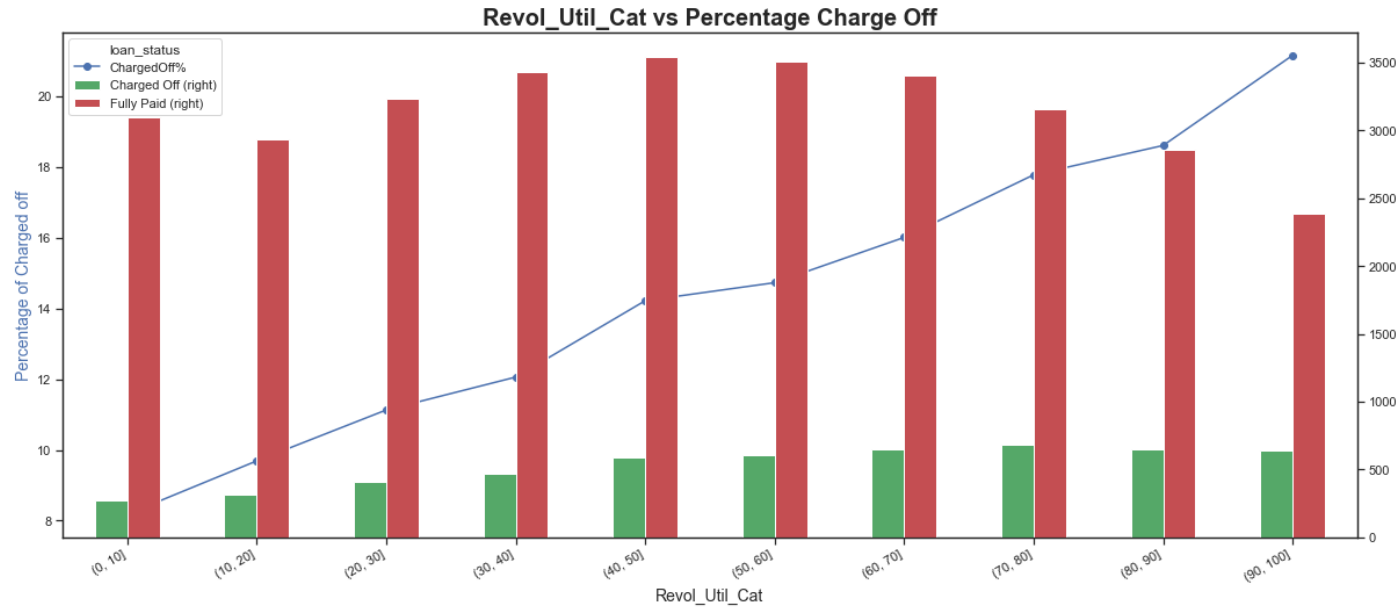
The percentage of charged off increases with increasing dti ratio of customer.



Based on existing loan accounts

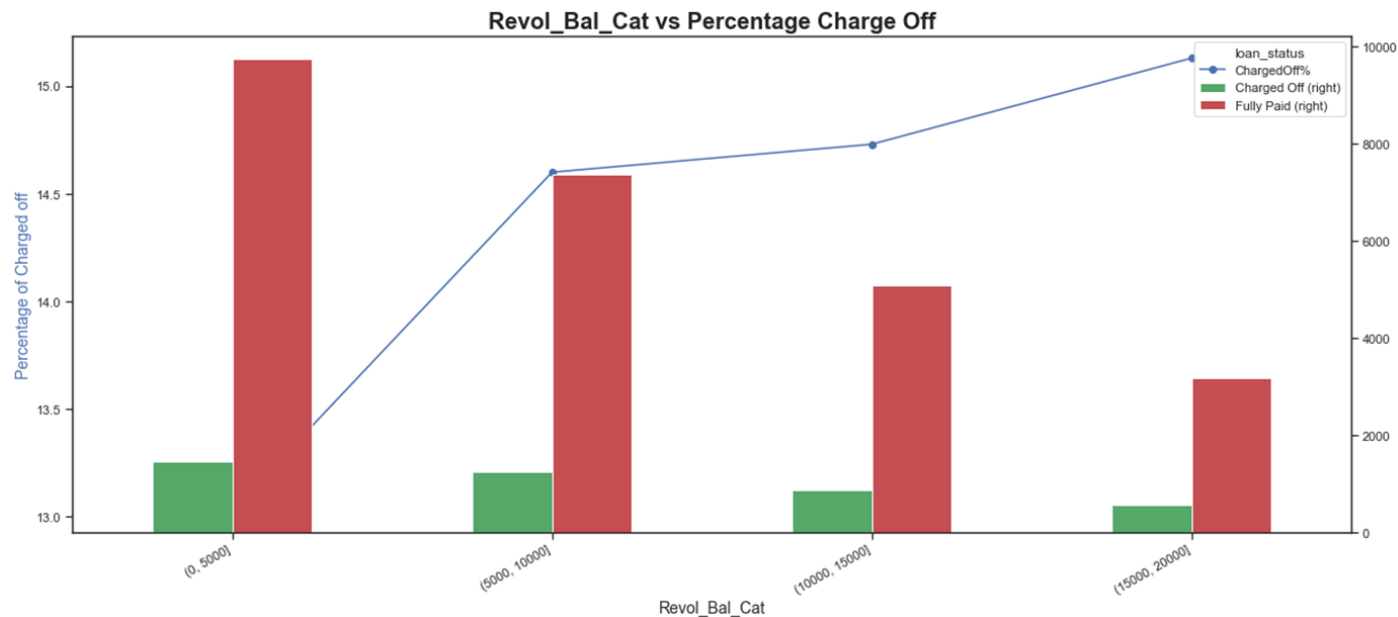
The customers with open accounts till 20, tend to have less charge-off percentage, but as the open account increases about 25, there is a higher percentage of customer going default.

Analysis:



Based on revolving Utility rate

The percentage of charged-off keeps on increasing as the revolving utility rate of loan customers increases.



Based on revolving balance

The percentage of charged-off increases as the customer's revolving balancing increases.

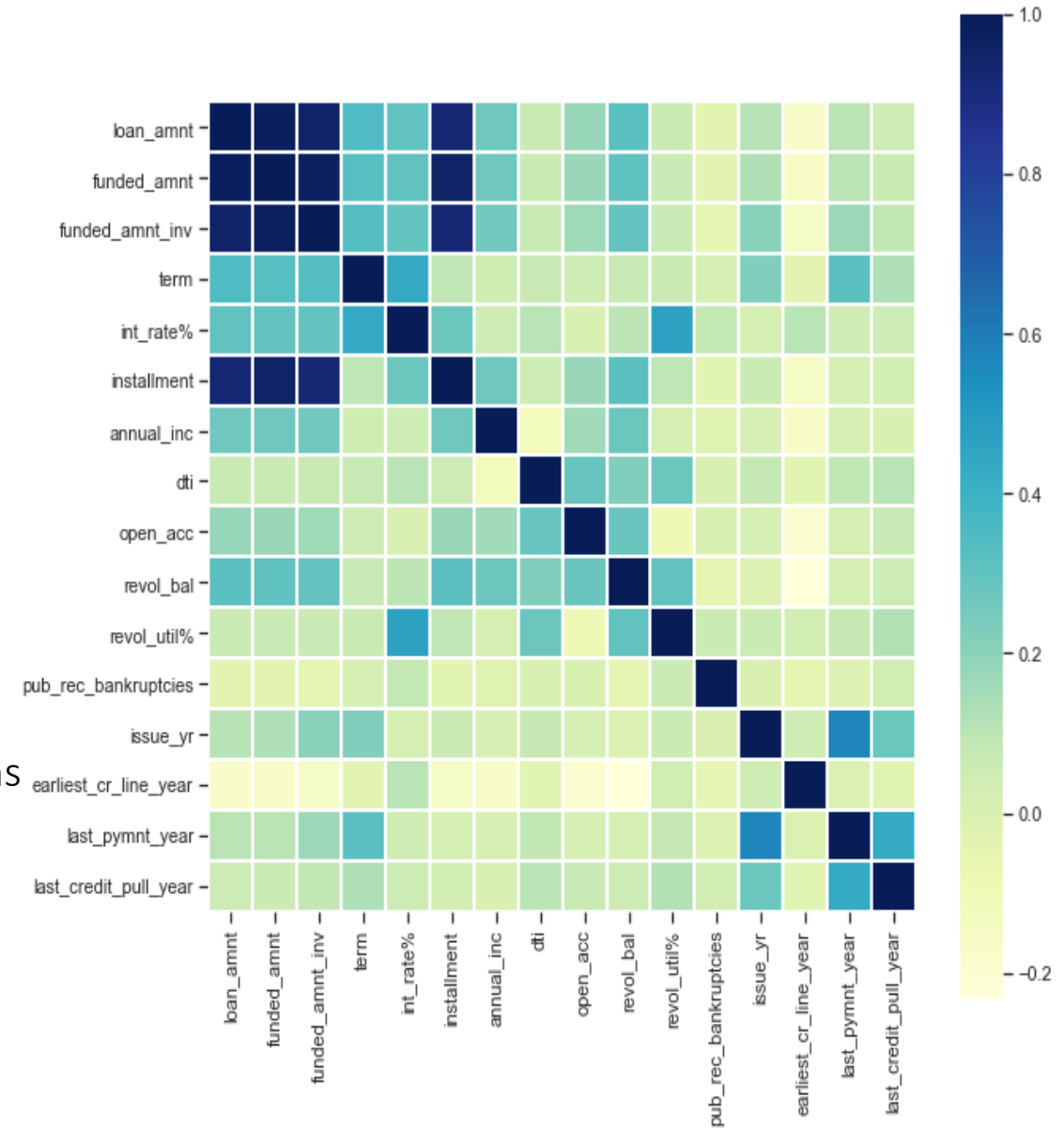
Insights from correlation map

- ❖ Correlation of Interest rate and revolving utilization rate is high.
- ❖ As the interest rate increases, the charge-off percentage also increase linearly

loan_status	Charged Off	Fully Paid	Total	ChargedOff%	FullyPaid%
int_rate_cat					
(5, 10]	796	11224	12020	6.62	93.38
(10, 15]	2574	15241	17815	14.45	85.55
(15, 20]	1763	5399	7162	24.62	75.38
(20, 25]	292	473	765	38.17	61.83

- ❖ Correlation of Term and interest rate is quite high.
- ❖ The charge-off percentage is higher with loan term of 60 months

loan_status	Charged Off	Fully Paid	Total	ChargedOff%	FullyPaid%
term					
36	3045	25256	28301	10.76	89.24
60	2380	7081	9461	25.16	74.84



Recommendations

1. Reduce providing loans for small business which are at higher risk of default.
2. Reduce providing the loan with 60 month tenure which are at higher risk of default.
3. Reduce giving loans to lower graded customers(E, F, G).
4. Stop providing loans with higher interest rate which are at higher risk of default.
5. In order to stop loans with high interest rate, stop providing loans with higher dti ratio. Higher dti -> Higher tendency of default.
6. Reduce providing loans with higher customer revolving balance.
7. Increase giving loans with higher customer annual income.
8. Start having stricter verification protocols. The above steps might help in achieving stricter verification protocols.

Thank you!