LENDING CLUB CASE STUDY

ARCHANA TANEJA

MAHALAXMI BHAT

THE PROBLEM

What is the problem?
- When a new borrower request for a loan we need to have some parameters to decide if they will result in a default or not.

Who has this problem?
- Lending Club and Investors who approve the loans.

Why should this problem be solved?
- We want to minimize the loss which happens when a loan is charged off and also on the same time not miss on potential ones which results in the profit.

How will I know this problem has been solved?
- If by using the parameters suggested in this ppt investors are able to better filter out defaulters.

STEPS FOLLOWED

- Data Understanding and Data Cleaning.

- Finding suitable Imputation Method to deal with null values.

- Univariate Analysis

- Segmented Univariate Analysis

- Bivariate Analysis

- Analyzing Correlation matrix

- Conclusion Drawn

IMPUTATION METHOD

- Public record of bankruptcies is the only column which has few null values.

- We can't say that Public record of bankruptcies has dependency on any column because of which we have null values in it. It belongs to MCAR(missing completely at random) category.

- We can use Median for imputation as we see that the data is numeric and highly skewed.
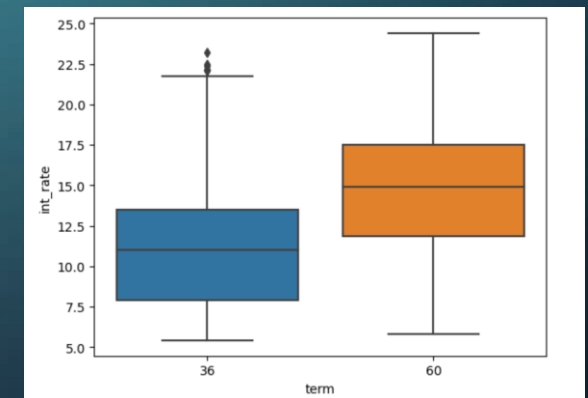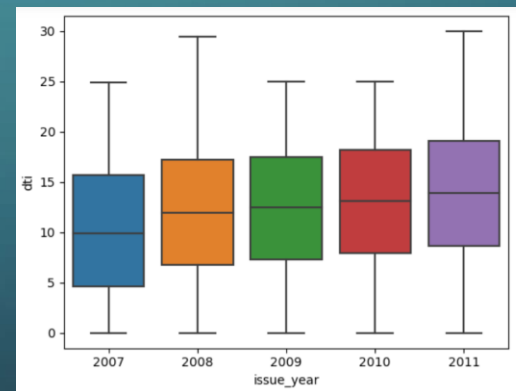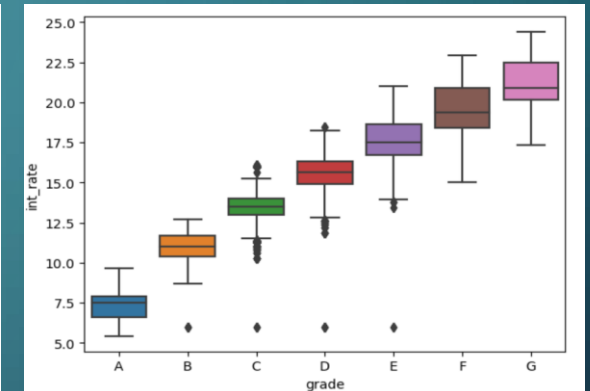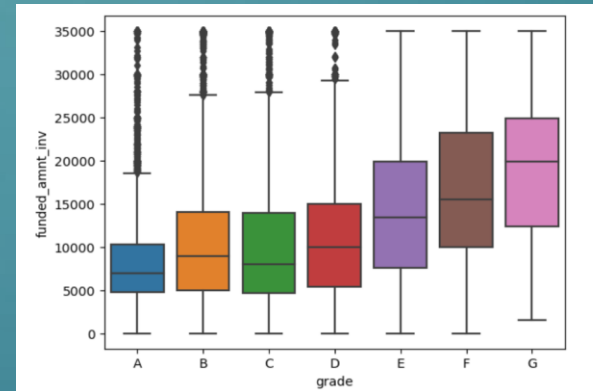
UNIVARIATE ANALYSIS

- Maximum loan is provided for grade B followed by grade A and C.

- More loan is taken for the amount between 5000 and 15000.

- people who are owning the home have taken less loans. Home ownership with "RENT" and "MORTAGE" have taken more loan.
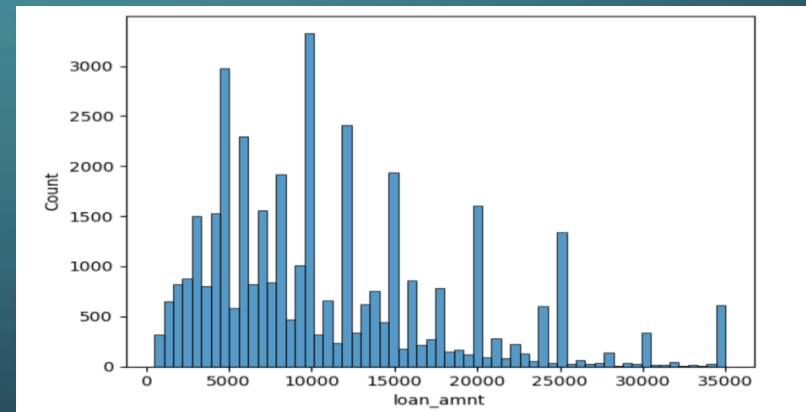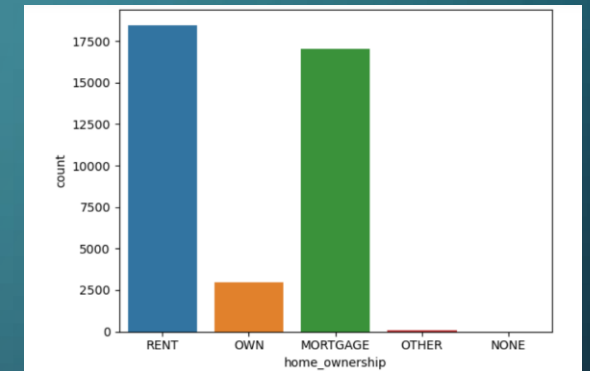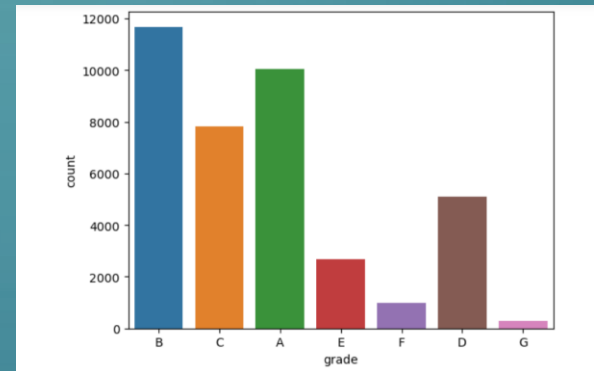
SEGMENTED UNIVARIATE ANALYSIS

- We see that there is a significant increase in the size of IQR for Grade E,F,G. This could mean that these people tend to request for wide range of funded amount.

- We see DTI(Debt to Income ratio) has increased over time for people requesting for loan. This is an indication of people relying more on loans now than in earlier times.

- Interest rates are more when the Grade is E,F,G or loan is acquired for longer tenure.

BIVARIATE ANALYSIS

- Highest percentage of charge offs are seen in case of small business. It shows this is the risky one and proper analysis should be done about the borrower before approving loan for small business purpose.

- We observe that borrowers with longer tenure to payback loan tend to default more.
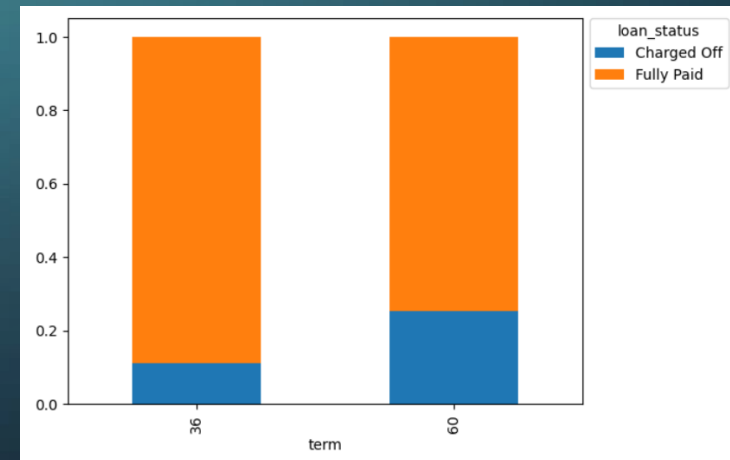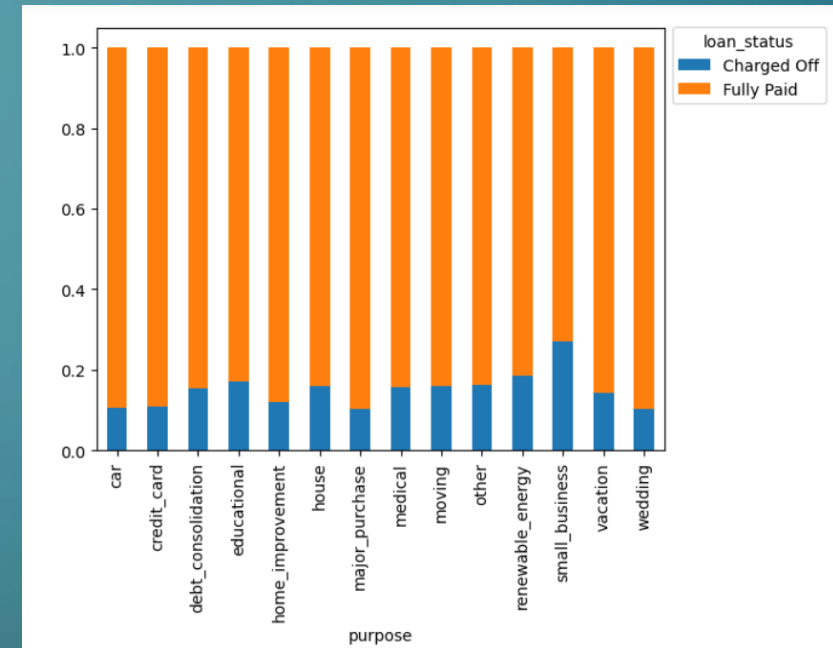
BIVARIATE ANALYSIS

- We observe that people who were verified are likely to default more. Reason could be that some of the time verification is done after the loan is passed or even for testing and analyzing purpose hence its not strongly affecting if the borrower would default. We observe that borrowers with longer tenure to payback loan tend to default more.

- Here we see that people who are unemployed or haven't completed 1 year of employment yet are more likely to default. Hence approver should be cautious when approving their loan request.

- Seeing the scatterplot we can assume that people who have less annual income request for less amount of loan. Also majority of loans are ranging from 5000 to 15000
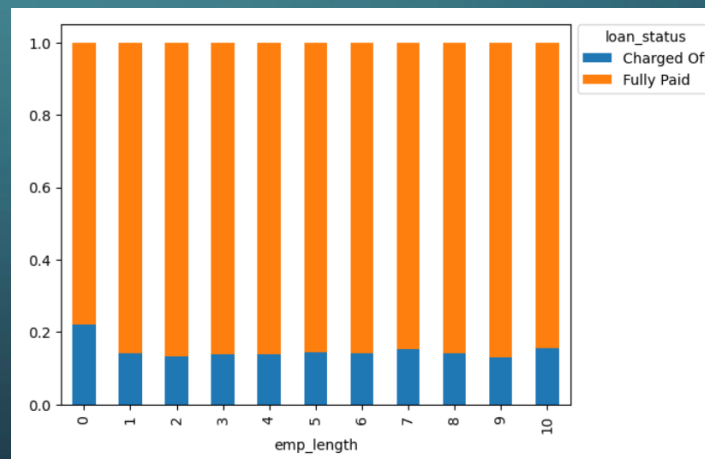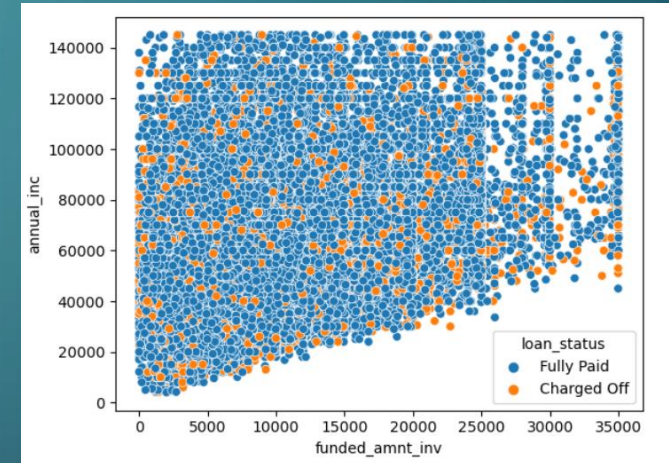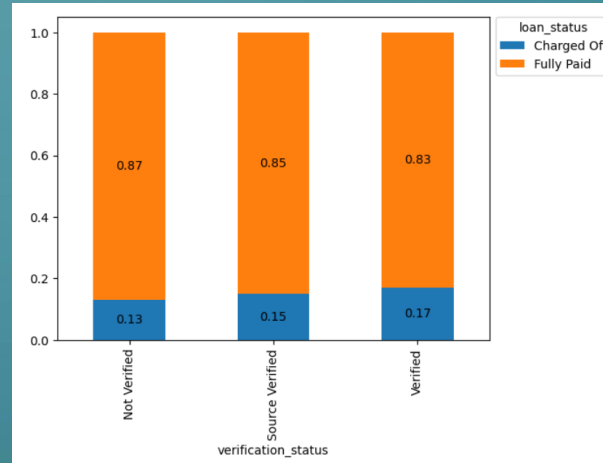
BIVARIATE ANALYSIS
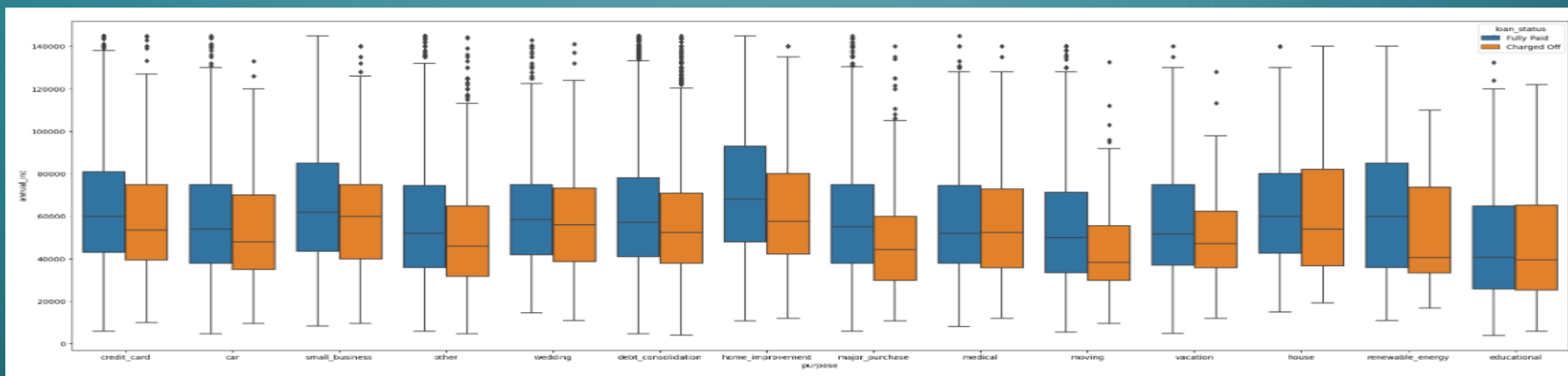
- We see loans taken for home improvement, moving, house and renewable energy are most likely to be fully paid. Whereas loans taken for education, debt consolidation and credit card are more likely to default.

- People taking loan for home improvement, small business, house and renewable energy also seem to earn well. One more thing to note is for home improvement, house and renewable energy defaulters median is quite below fully paid ones. But in case of small business even if you earn well there is a 50-50 chance you might default.

- We see people who earn quite high are placed in grade G, but when we look at boxplot we get to know people who default are having less annual income compared to there counterpart.

- And all who requested for loan having income more then 94000.0 (75% of charged off grade G borrower) have Fully Paid their loan.

CONCLUSION

The following are the parameters to consider to minimize Charge Off count:

- Grade E,F,G are more likely to default but in Grade G are the people who request for large fund amount and more than 50% of them Fully Pay it back. Hence rejecting their application is not a solution but we can have high interest on it.

- Borrowers from Grade G who requested for loan having annual income more than 94000.0 have Fully Paid their loan. So we might give them better offers considering their credit score and background verification.

- Highest percentage of charge offs are seen in case of small business. It shows this is the risky one and proper analysis should be done about the borrower before approving loan for small business purpose.

- Here we see that people who are unemployed or haven't completed 1 year of employment yet are more likely to default. Hence approver should be cautious when approving their loan request.

- We see loans taken for home improvement, moving, house and renewable energy are most likely to be fully paid.

- Loans with longer term have high probability of resulting in a charged off, hence give it on high interest rate.