Lending Club Case Study

Group Members:
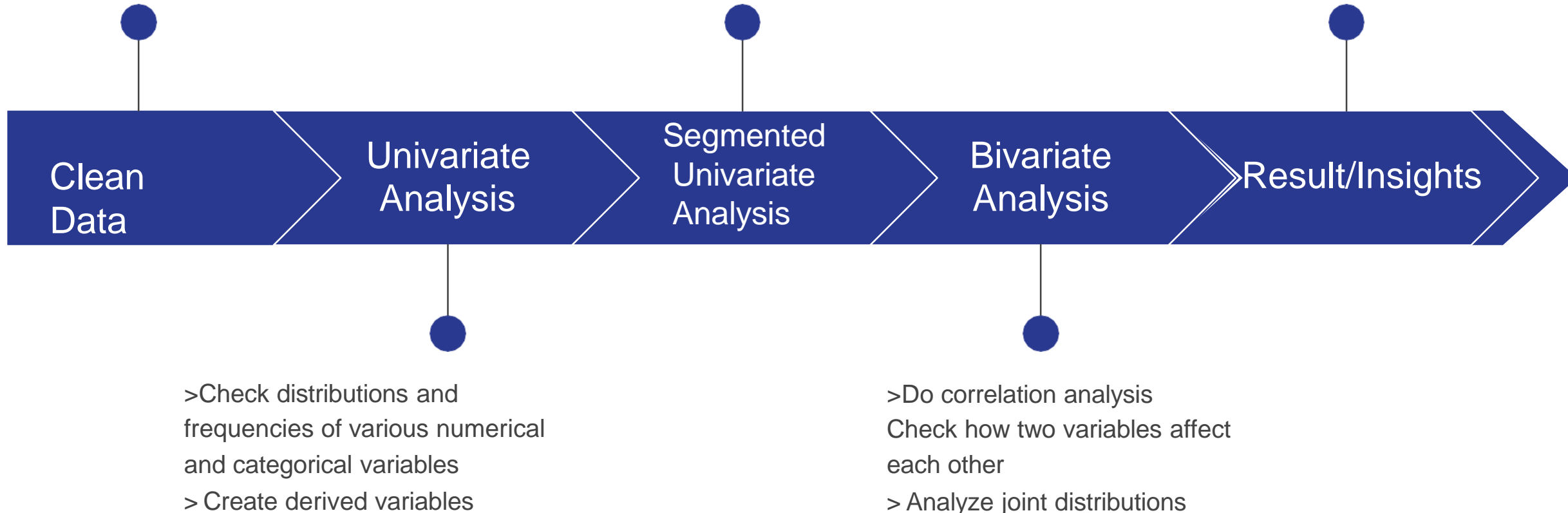*Shrawan Kumar Sahu*
Ganpati Prasad

Abstract

➤ 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.

➤ Based on the past loan applicants data, Company has to take decision for a loan approval and reduce two types of risks associated :

- 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

Problem Solving Methodology

>Drop columns with null values, all random values or single category value
>Convert values to proper int, float, date representations

>Analyze variables against segments of other variables
> Create derived variables

Publish insights and observations for each Feature Analysis

| Clean Data | Univariate Analysis | Segmented Univariate Analysis | Bivariate Analysis | Result/Insights |

>Check distributions and frequencies of various numerical and categorical variables
> Create derived variables

>Do correlation analysis Check how two variables affect each other
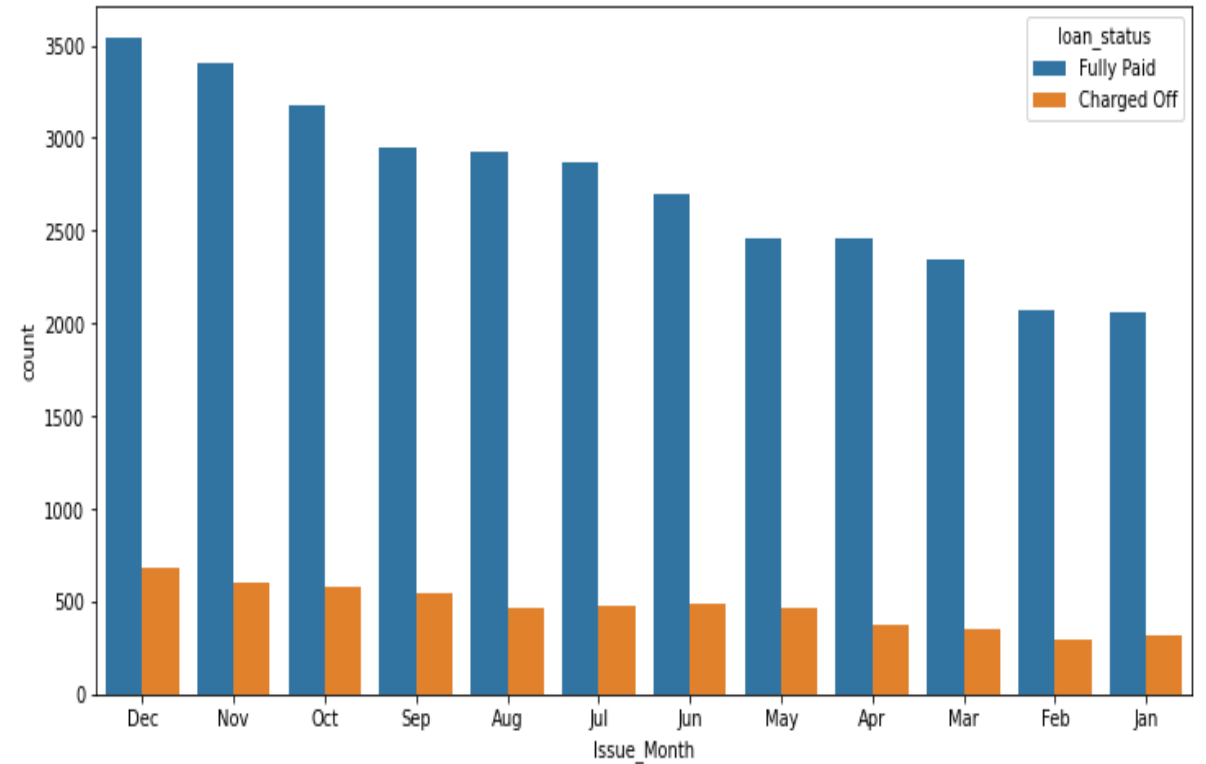> Analyze joint distributions

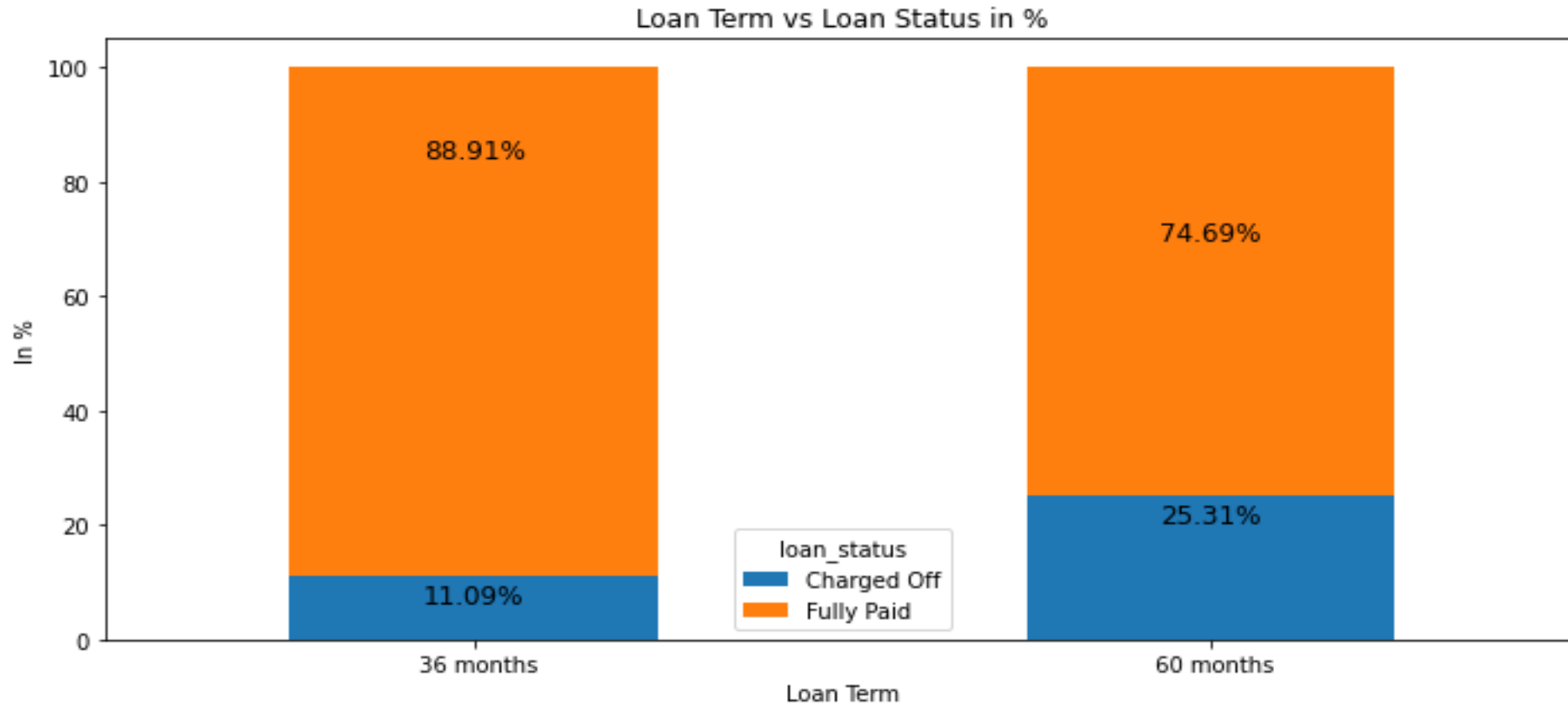# **Analysis**: Loan Application Yearly and Monthly

Lending club has really expanded year by year, the number of loan issued are doubled every year
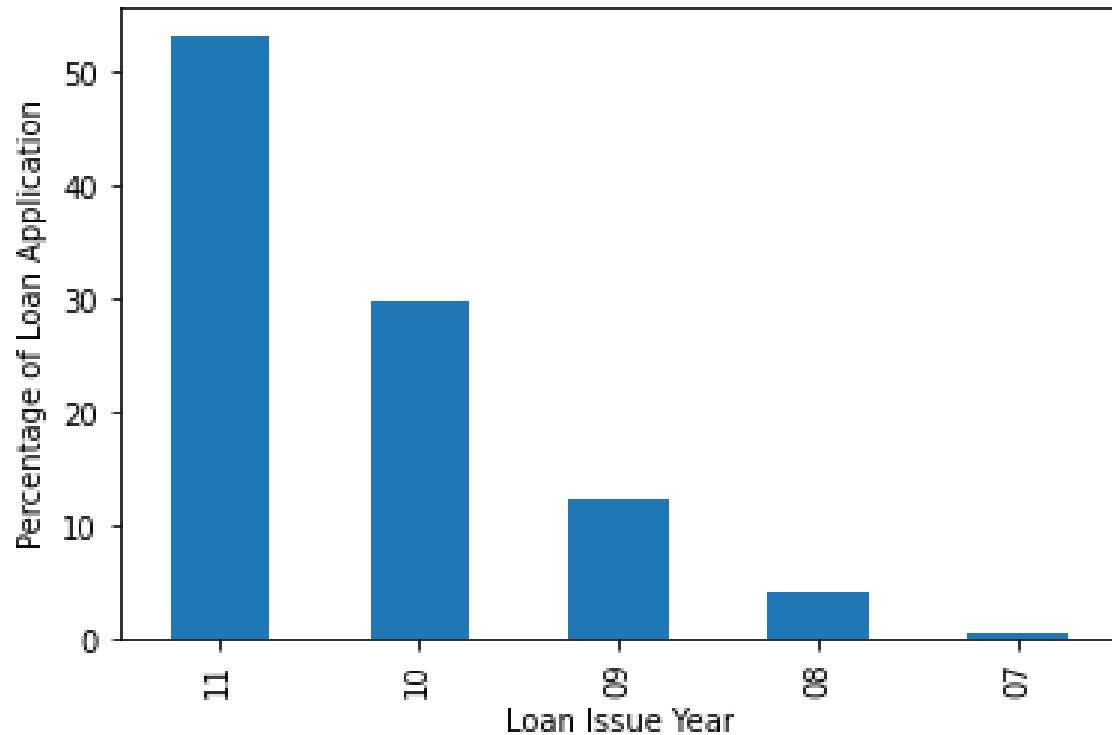
Most of the Applicant has been issued the Loan in the December Month

# **Analysis**: Loan Tenure (Term)

Loan Term vs Loan Status in %

❖ Borrowers with 3 Years tenure are most likely to Fully Pay where as Borrowers with 5 Years tenure are most likely to Charged off

Analysis: Interest Rate

There is presence of outliers, which indicate that our analysis should be based on median statistical measure.

The column int_rate is an influential feature in the approval of loan process as the analysis clearly shows that interest rate below 12% are most likely to Fully Pay where as interest rate above 12% are most like to Charged off.
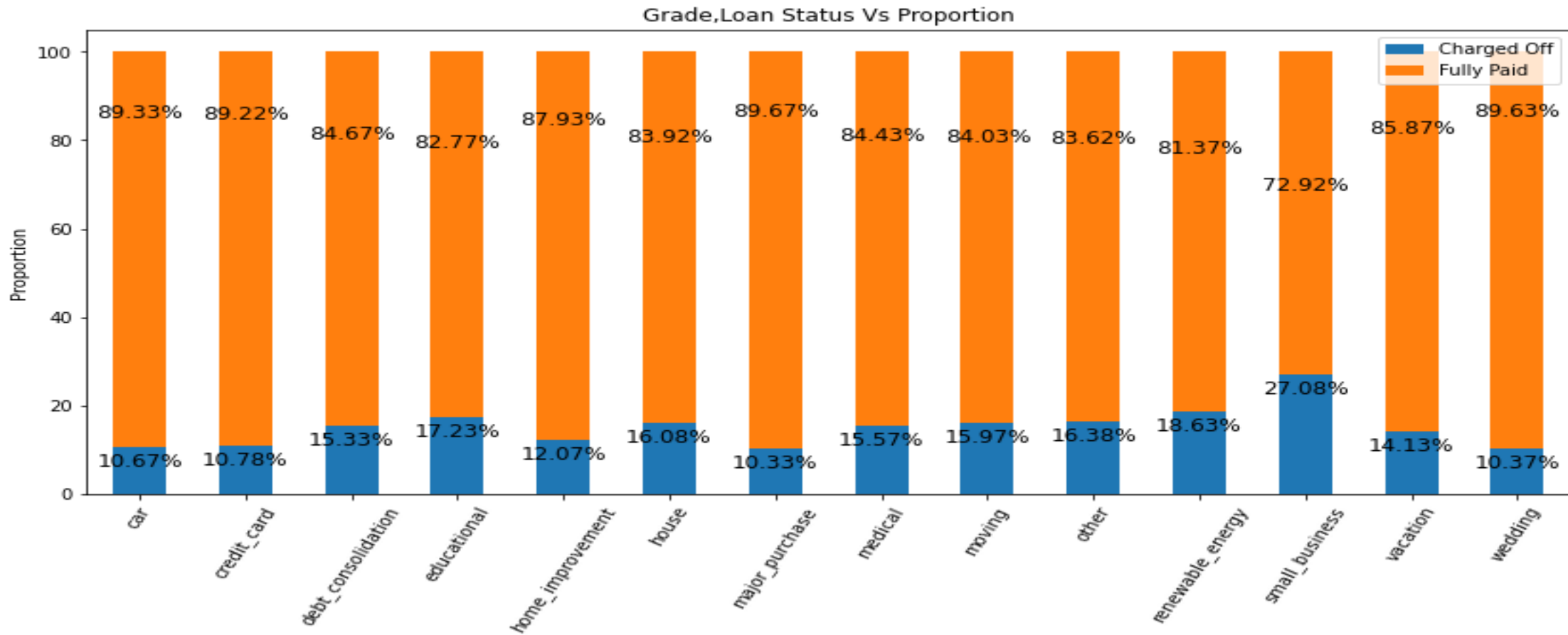
Analysis: Grade

Grade,Loan Status Vs Proportion
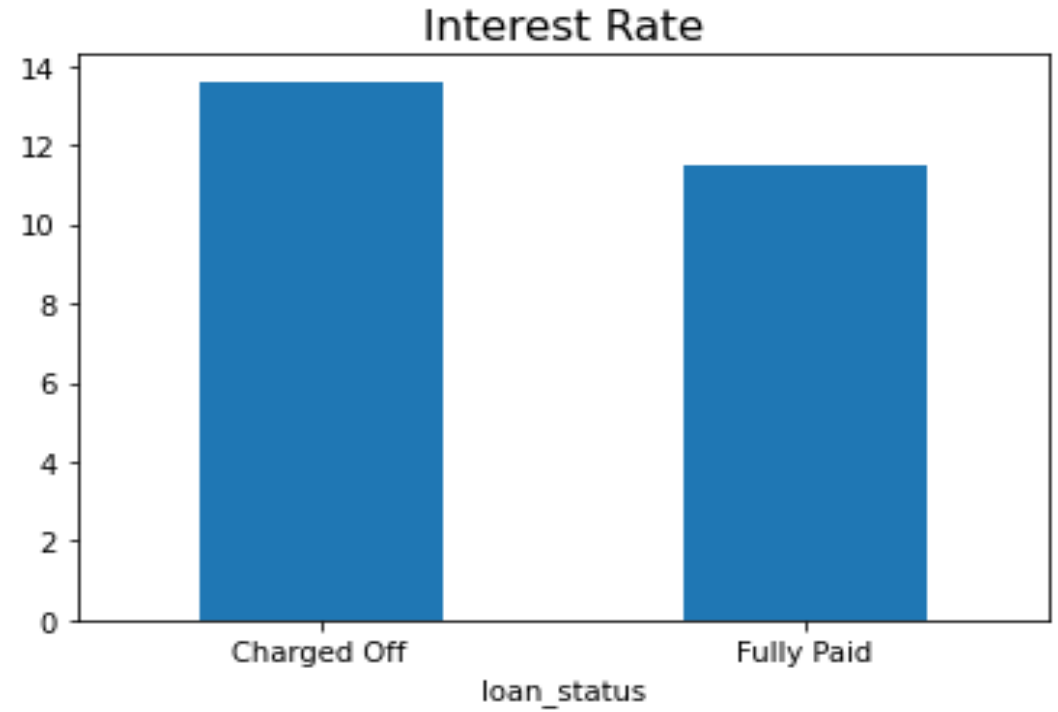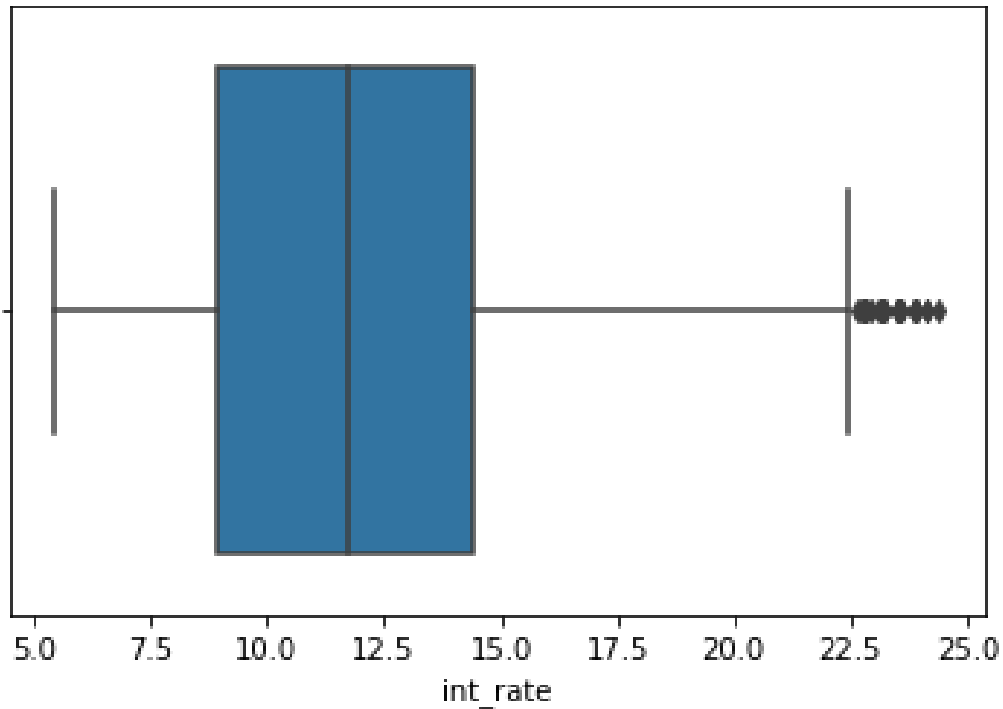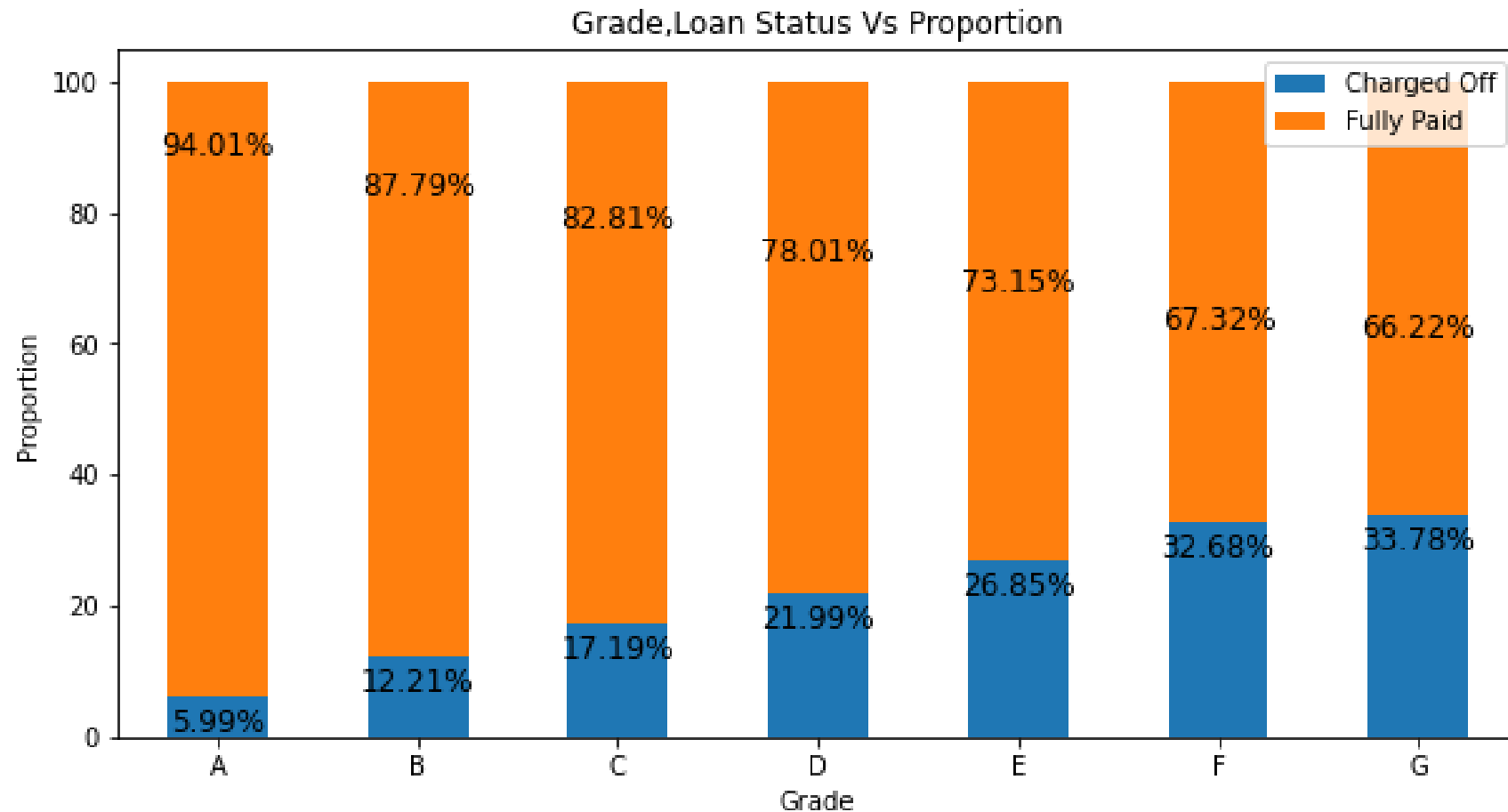
❖ Grade is having significant impact on Loan Approval Process. Higher the Grade are more likely to Fully Pay and lower the Grade are more likely to Charged Off (A>B>C>D>E>F>G).
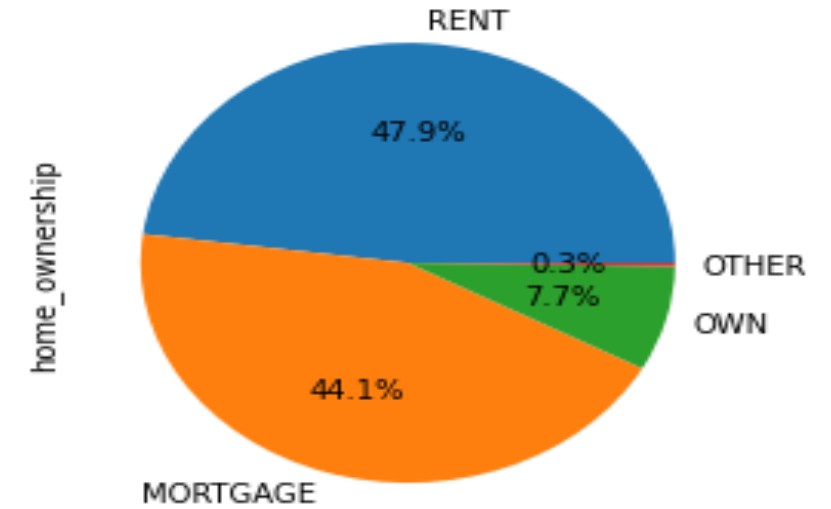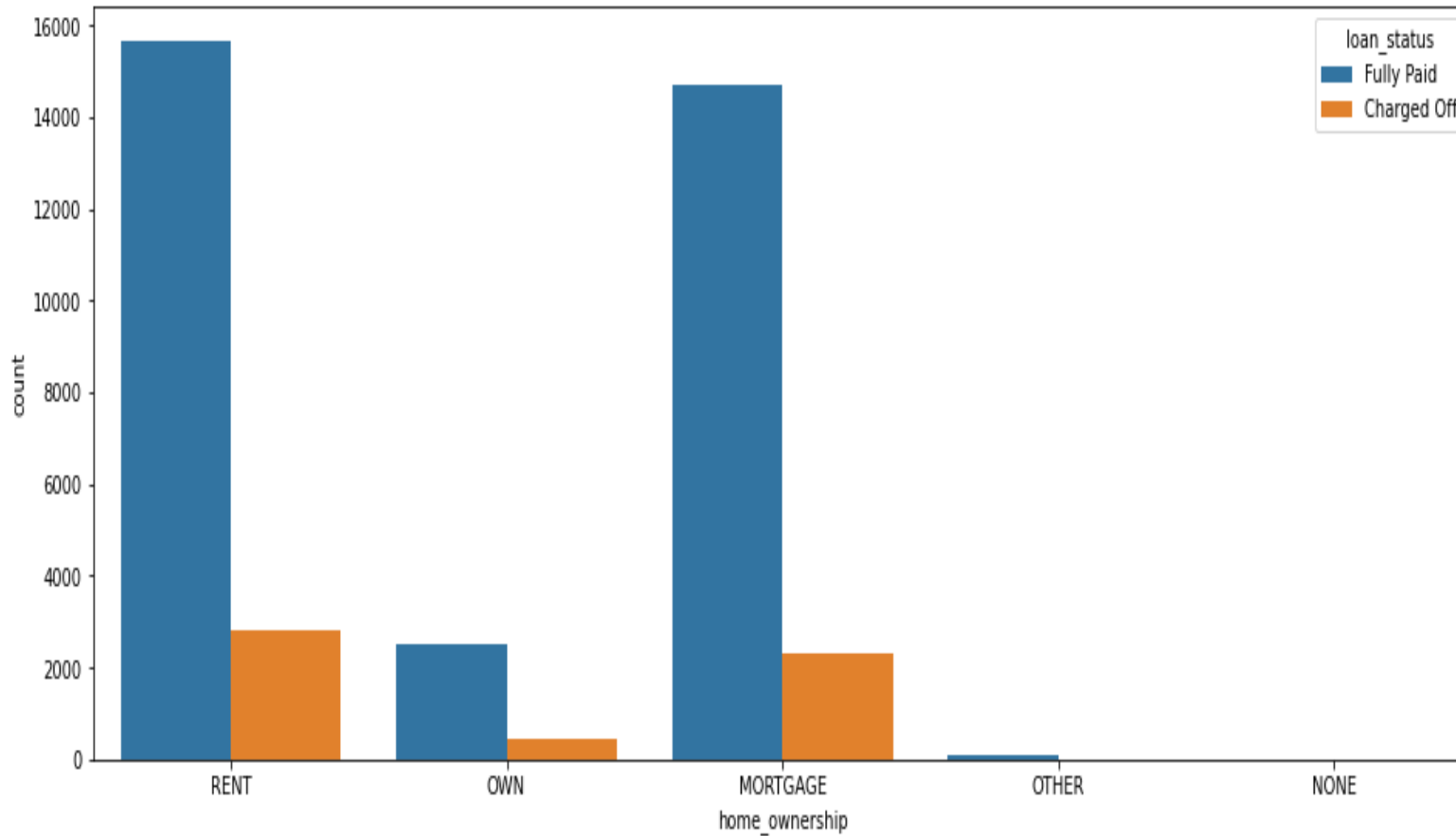
Analysis: Purpose

Grade,Loan Status Vs Proportion

❖ Purpose is important Feature as we can see borrower having purpose of major purchase are more likely to Fully Pay and purpose of small business are most like to Charged Off.

Analysis: Home Ownership

❖ The feature Home Ownership reflect an interesting insights, that most of the borrower who stays in Rent or Mortgage are the loan applicants and out of these two who stays in Rent is having higher percentage of Fully Paid.

Analysis: Debt to Income Ratio (DTI)

There is no outlier found in the boxplot of DTI Ratio so we can go for mean plot to interoperate average DTI.
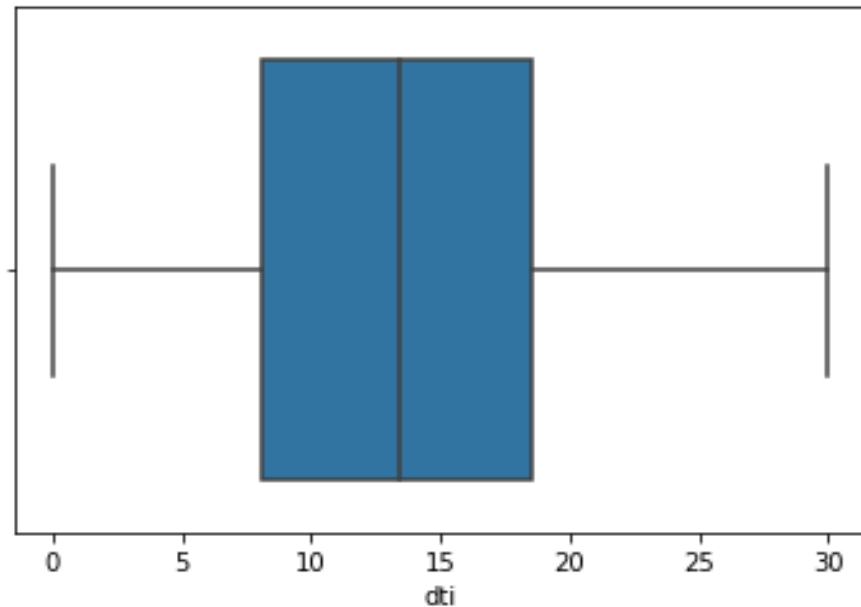
If the DTI Ratio is more than 13.27 then Borrowers are most likely to Default.

Debt to Income Ratio(DTI)
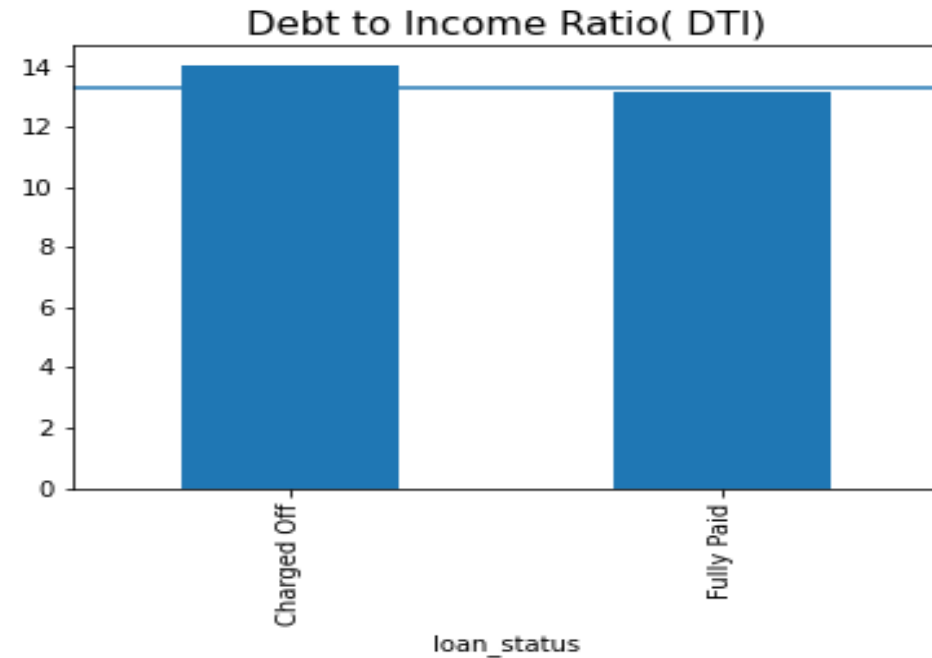
Conclusion

➢ Lending club should reduce the high interest loans for 60 months tenure, they are prone to loan default.

➢ Grades (A to G) are good metric for detecting defaulters. Lending club should examine more information from borrowers before issuing loans to low grade profile.

➢ If the DTI Ratio is more than 13.27 then Borrowers are most likely to Default.

➢ Small business loans are defaulted more. Lending club should try to reduce or verify properly before issuing the loans to them.

➢ It is really surprising to see that the Borrowers having a verified income are most like Charged Off. Hence Lending club should verify the profile very closely.

➢ Around 14.6% Borrower got defaulted whose average annual income is below 59K USD for the average loan amount of around 11K USD. Hence Lending Club should check such profile before approving Loan.