Lending Club Case Study

Group Member:
Varun Karkera

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.
- The objective of analysis is to use the information about past loan applicants and find whether they ‘defaulted’ or not.

Problem solving methodology

Data Analysis

Data Analysis

Data Analysis

Data Cleaning

Removing the null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.

Data Understanding

Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses

Univariate Analysis

Analyzing each column, plotting the distributions of each column.

Segmented Univariate Analysis

Analyzing the continuous data columns with respect to the categorical column

Bivariate Analysis

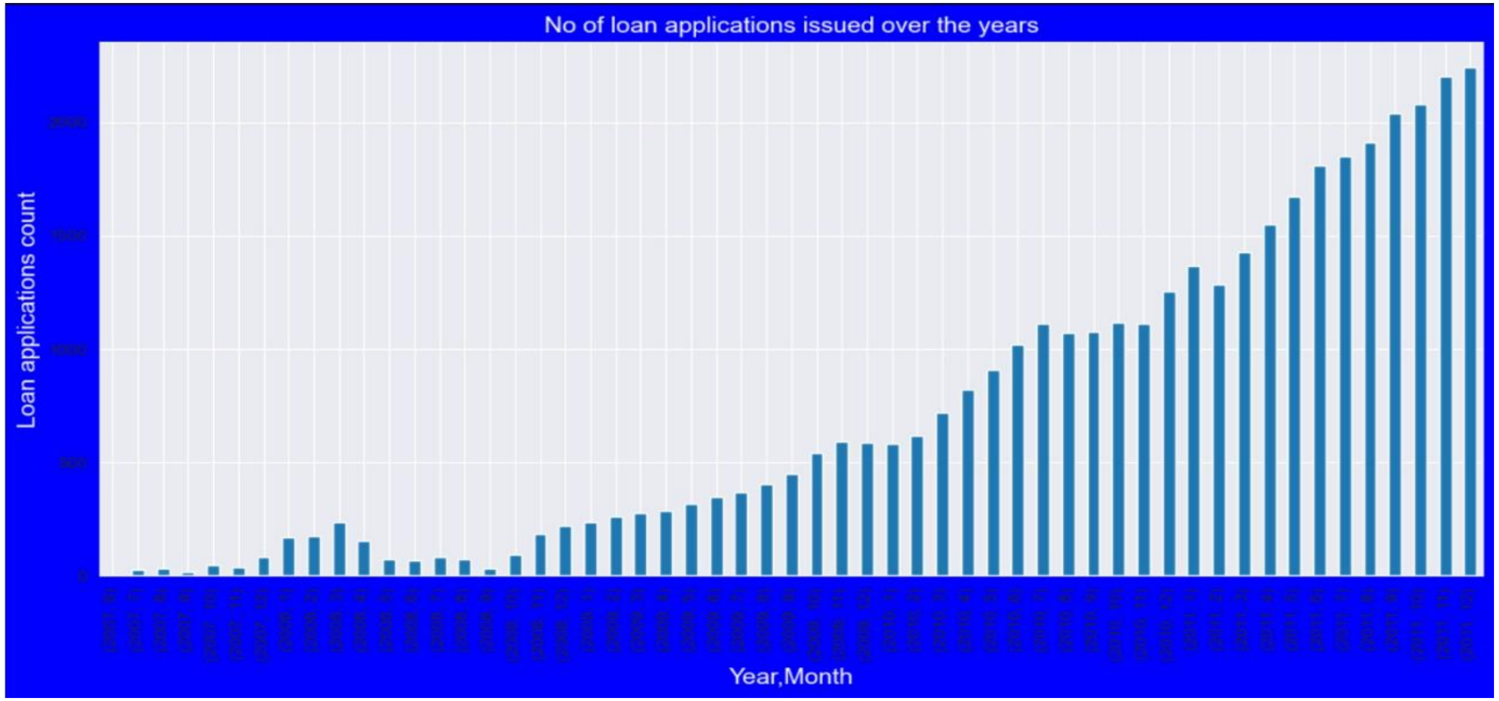
Analyzing the two-variable behavior like term and loan status with respect to loan amount.

Recommendations

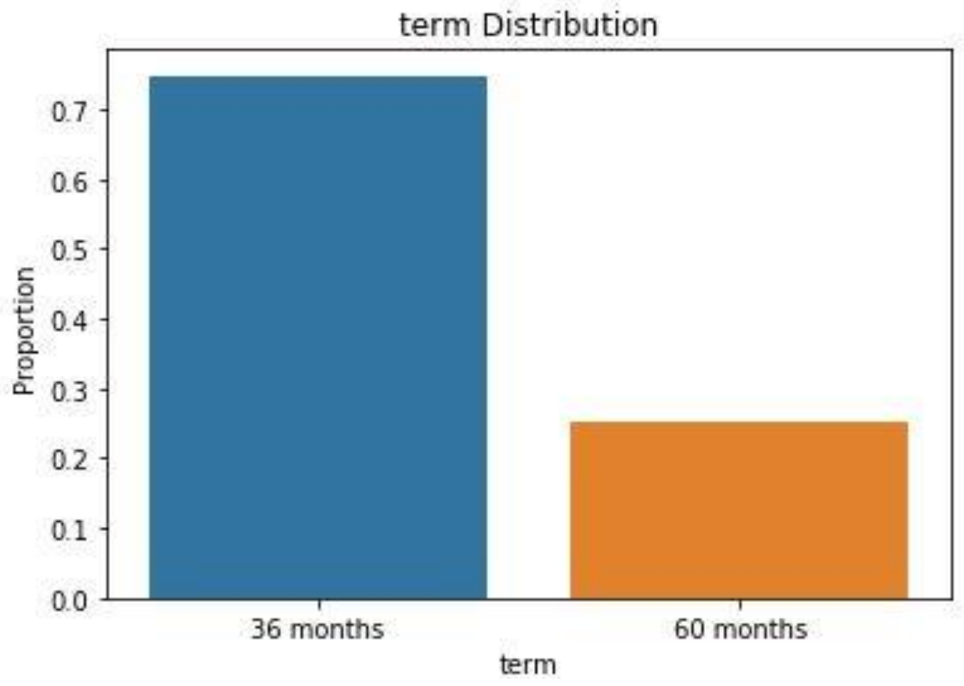
Analyzing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

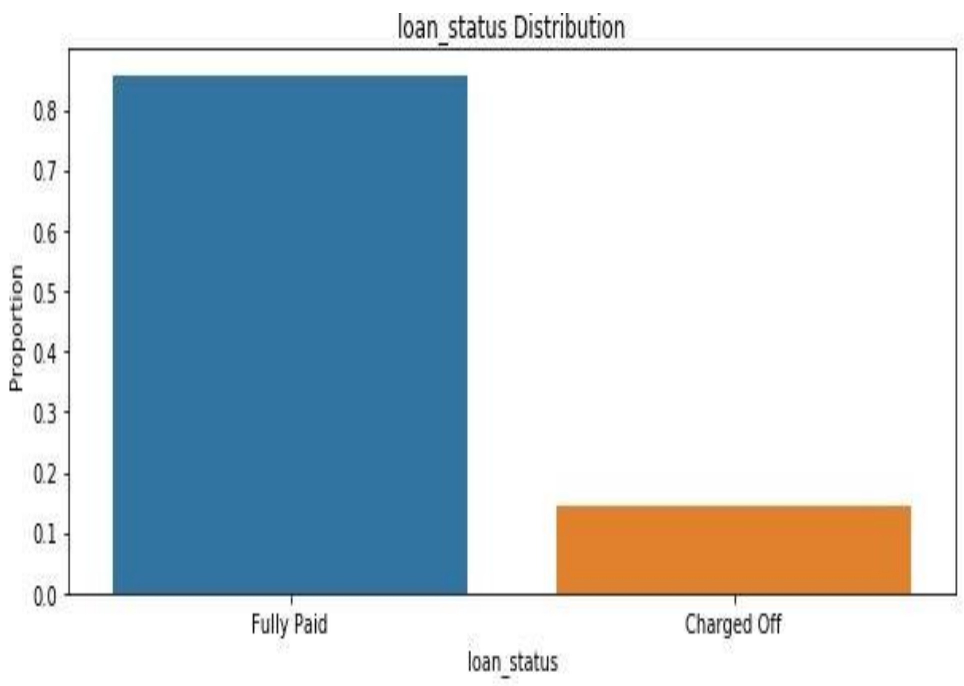
Analysis

- Lending club has really expanded year by year, the number of loans issued are doubled every year.
- Also, the issued month of loans is also increasing from January to December. In the final quarter of year there are more loans issued this could be because of vacation and Christmas.

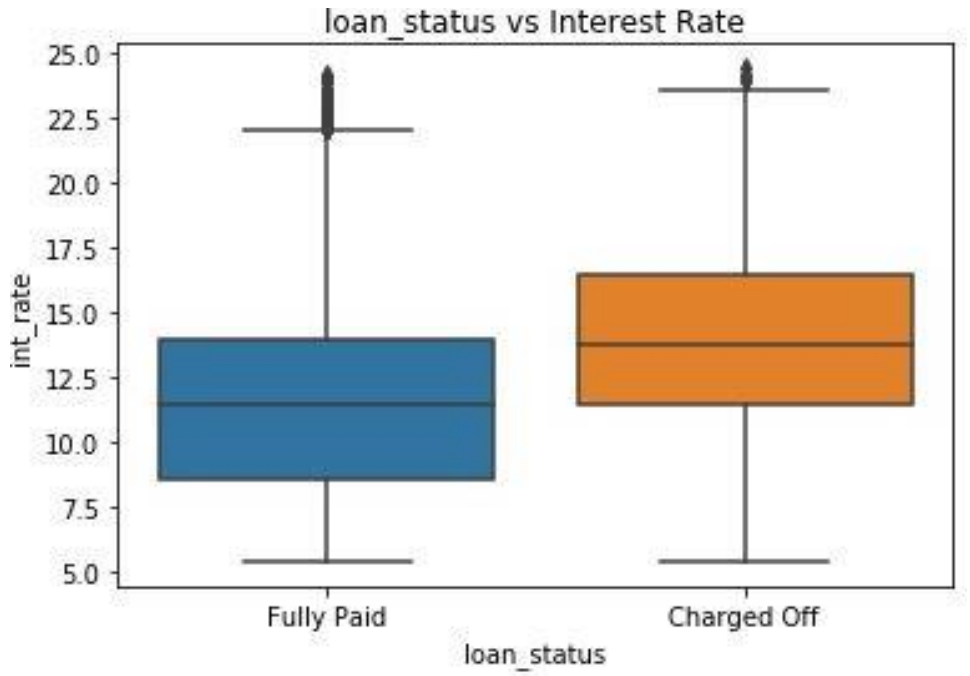


Analysis





- There are only two loan terms 36 and 60 months. Around 75% borrowers took loans with 36 months term.
- The charged off borrowers are around 15% and fully paid is around 85% in the given data set.

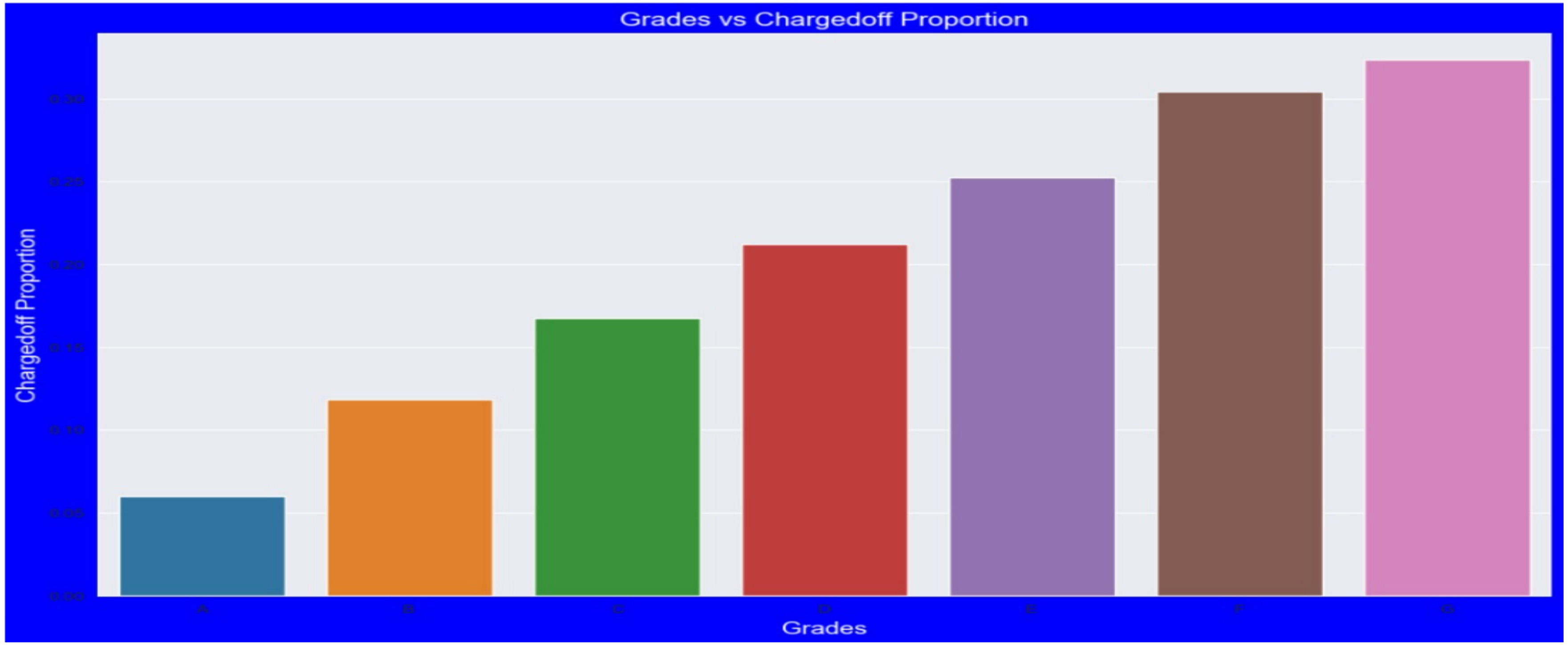


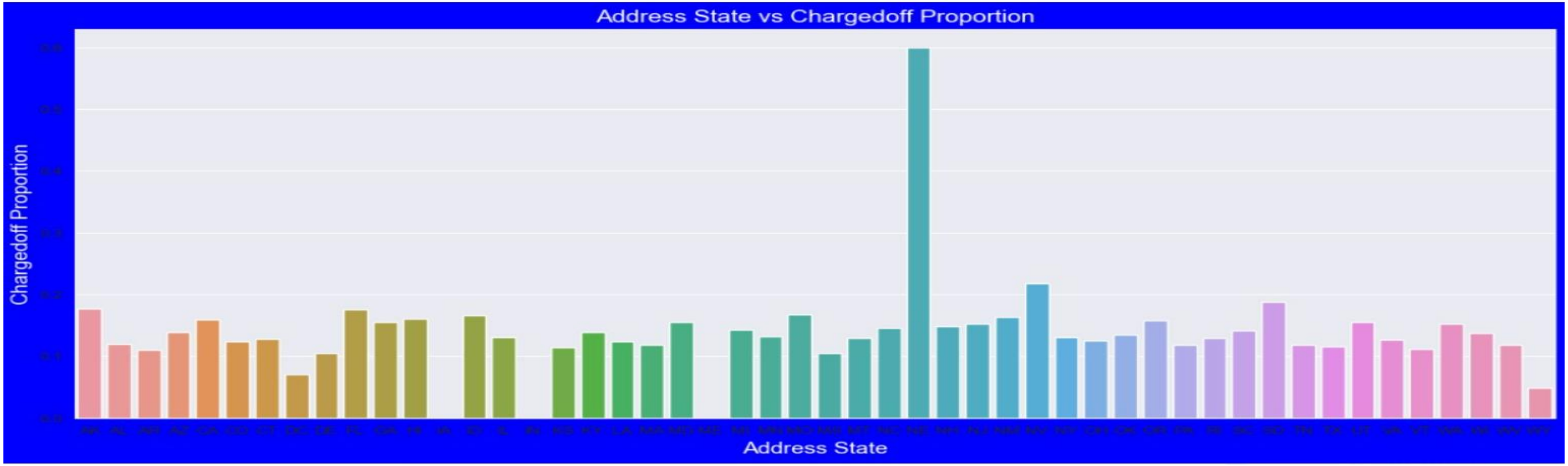
When the loan interest rate is high there is high chance of loan getting defaulted.

Analysis

The default rate is high in 60 months tenure because most people took high loan amount with high interest rate in it and they faced difficulties in returning the sum to bank.

- Grades are very good category to tell the borrower probability of defaulting the loan.
- The Lower grades (E, F, G) have higher chances of defaulting the loan than Higher ones(A,B)
- Also, the Lower grades are getting loans for higher interest rates which might be the cause for loan default.

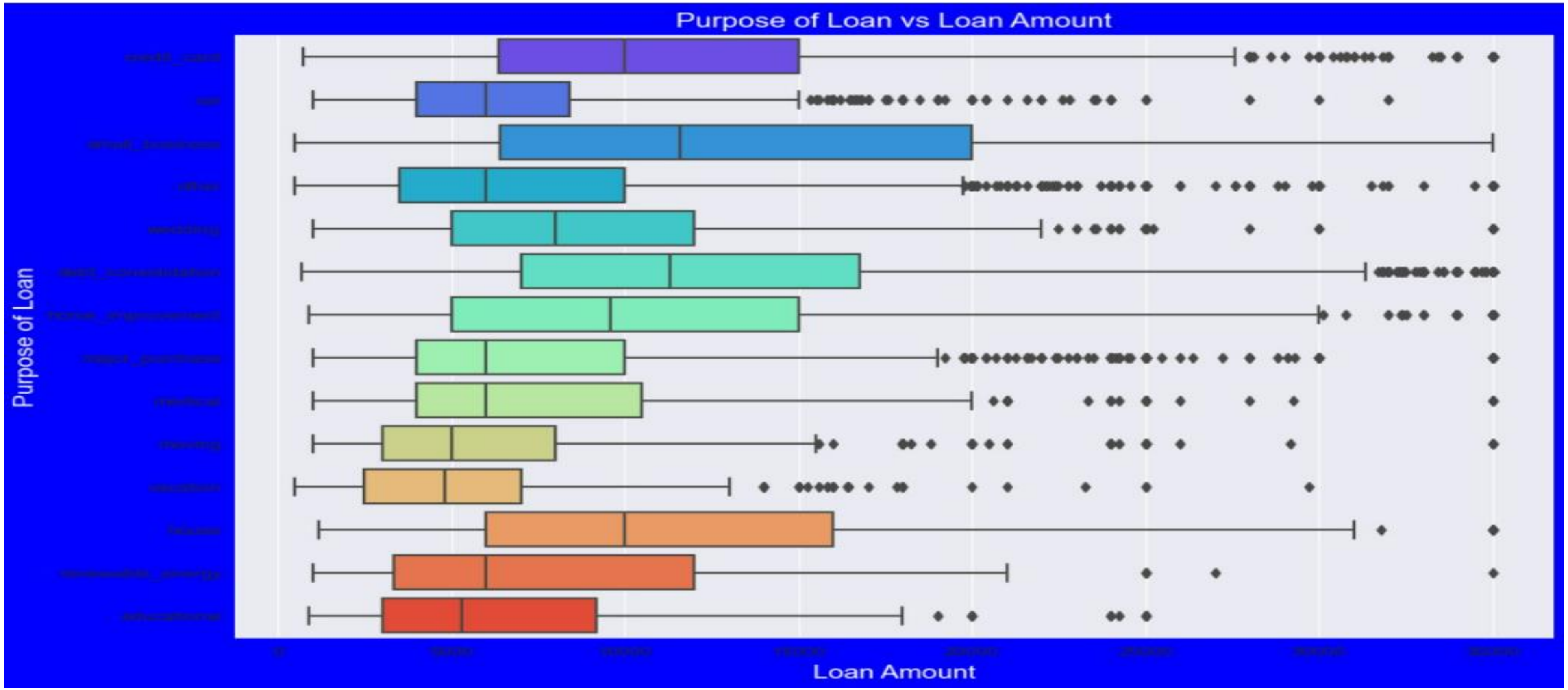


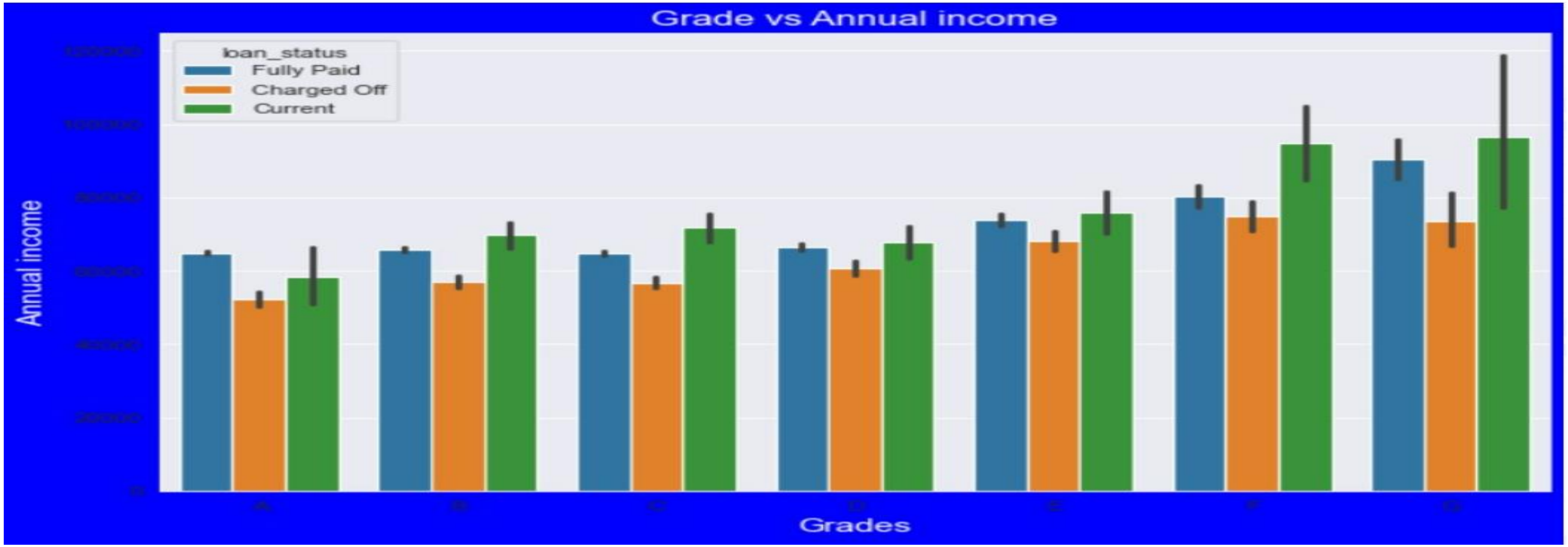


State NE has the high chances of charged off but the number of applications are too low to make any decisions.

NV, CA and FL states show good number of charged offs in good number of applications.

Borrowers who took loans for small business purpose have defaulted more.





From this we can conclude that the ones getting ‘charged off’ have lower annual incomes than the ones who ‘paid fully’ for each and every grade(i.e.at the same interest range)

Conclusions

- Lending club should reduce the high interest loans for 60 months tenure, they are prone to loan default.
- Grades are good metric for detecting defaulters. Lending club should examine more information from borrowers before issuing loans to Low grade (G to A).
- Lending Club should control their number of loan issues to borrowers who are from CA, FL and NY to make profits.
- Small business loans are defaulted more. Lending club should stop/reduce issuing the loans to them.