

# Lending Club Case Study

Mahendra Singh

# Agenda

- Problem Statement
- Data Load and Cleanup
- Univariate Analysis
- Bivariate Analysis
- Conclusion

# Problem Statement

- Objective is to identify patterns which indicate if a person is likely to default or not, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate.

# Data Load and Cleanup

- Data cleanup is a process to remove unwanted data or noise in the data and convert the data into desired format so that data analysis can be done on the dataset

Below are the steps we have followed for data cleaning:

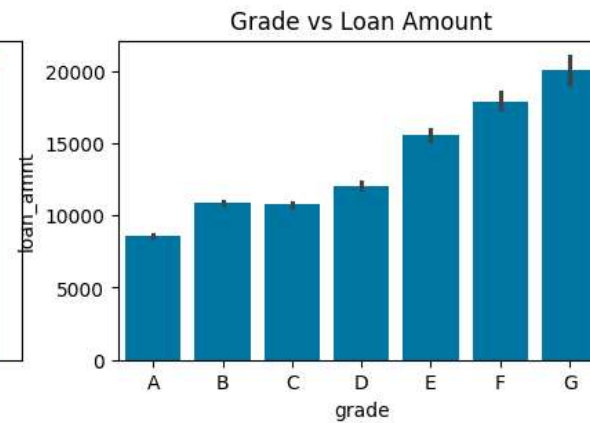
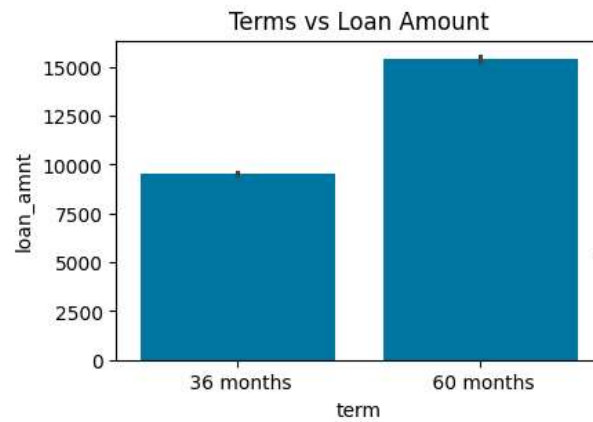
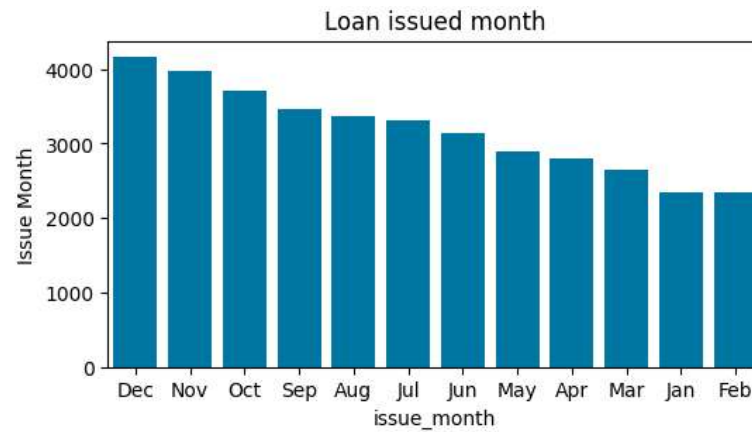
- Loaded loan.csv file with 39717 rows and 111 columns
- Removed all the empty columns
- Removed all the columns that are empty more than 70%
- 27 columns remained after data cleaning
- Converted the column values to the expected format and data types
- Validated duplicate rows bases on id column
- Created new columns for further analysis

# Univariate Analysis

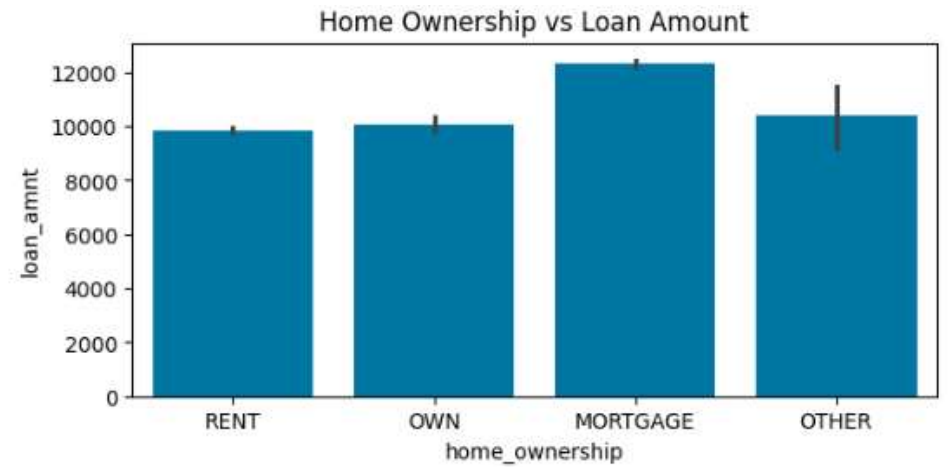
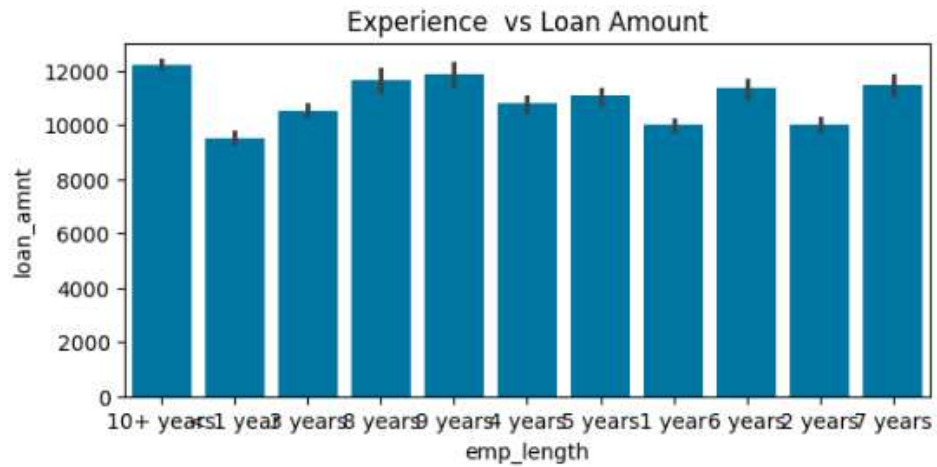
Done the analysis for important features as follows:

- Loan issued month
- Terms vs Loan Amount
- Grade vs Loan Amount
- Home Ownership vs Loan Amount
- Experience vs Loan Amount
- Verification Status vs Loan Amount
- Loan Status vs Loan Amount
- Grade vs Int Rate
- Verification Status vs Int Rate

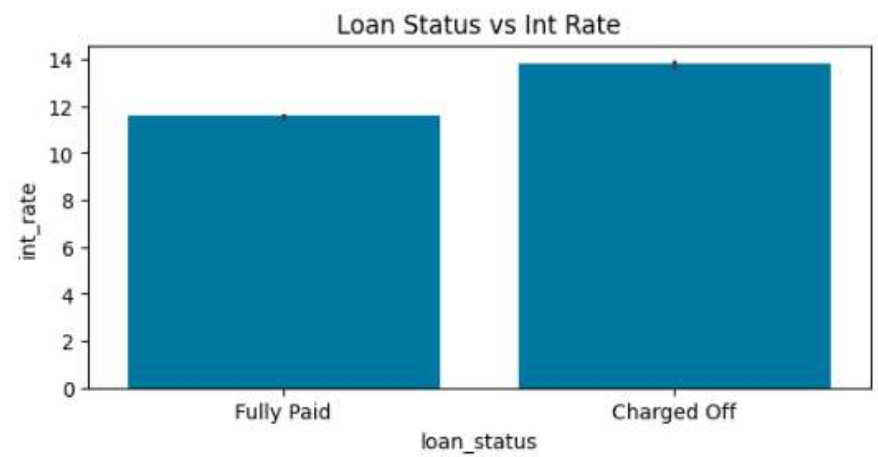
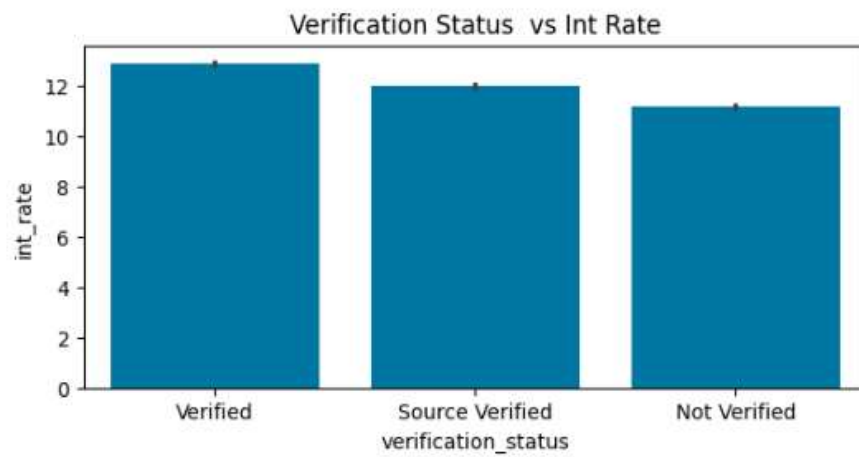
# Univariate Analysis



# Univariate Analysis



# Univariate Analysis



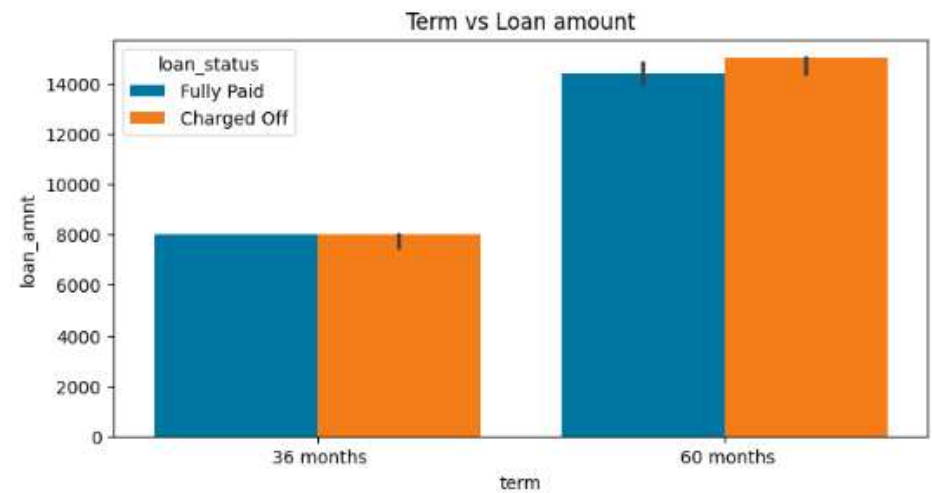
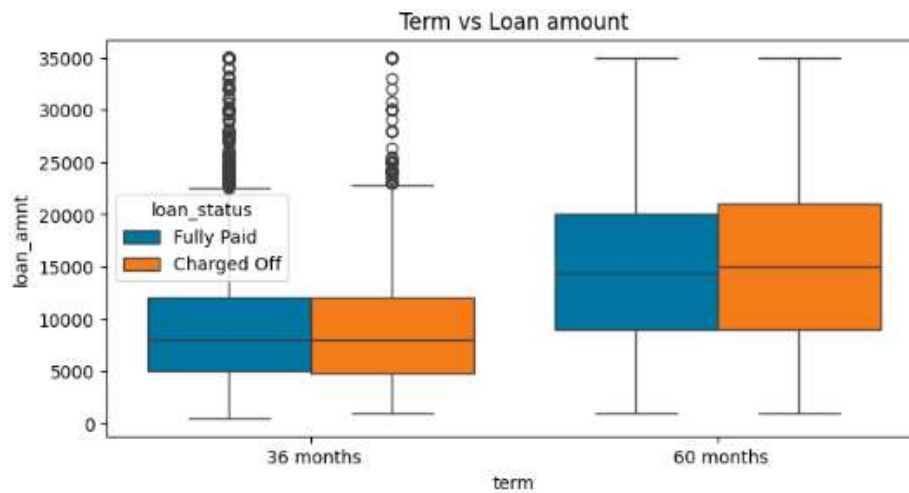


# Bivariate Analysis

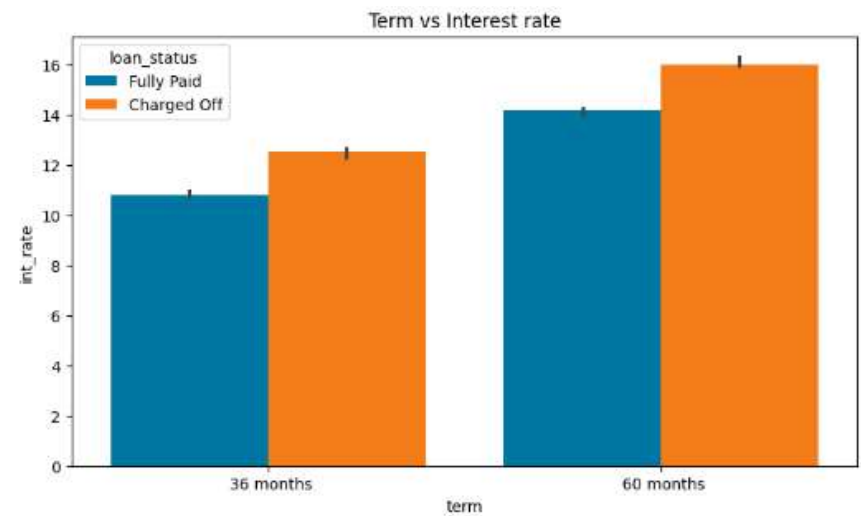
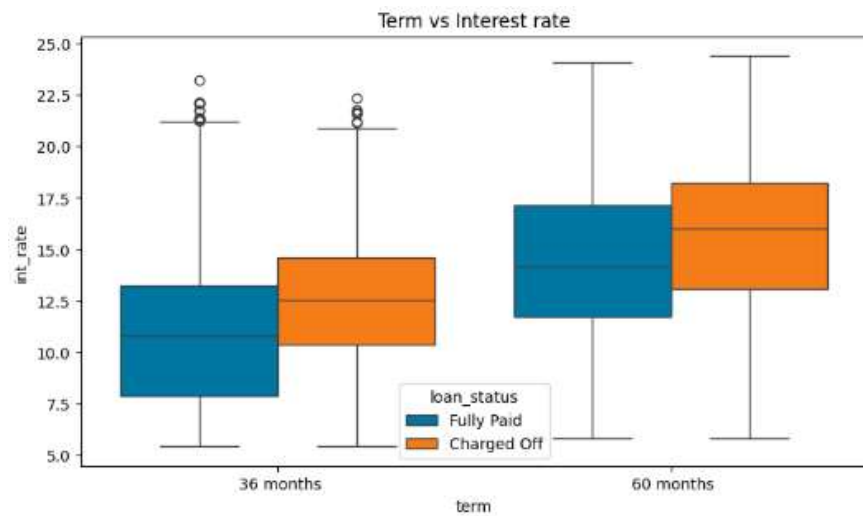
Done the analysis for important features as follows:

- Term vs Loan Amount
- Term vs Interest Amount
- Comparison of DTI Based On Grade For Loan Status
- Public Record Bankruptcies Vs Annual Income
- Annual Income vs Grade

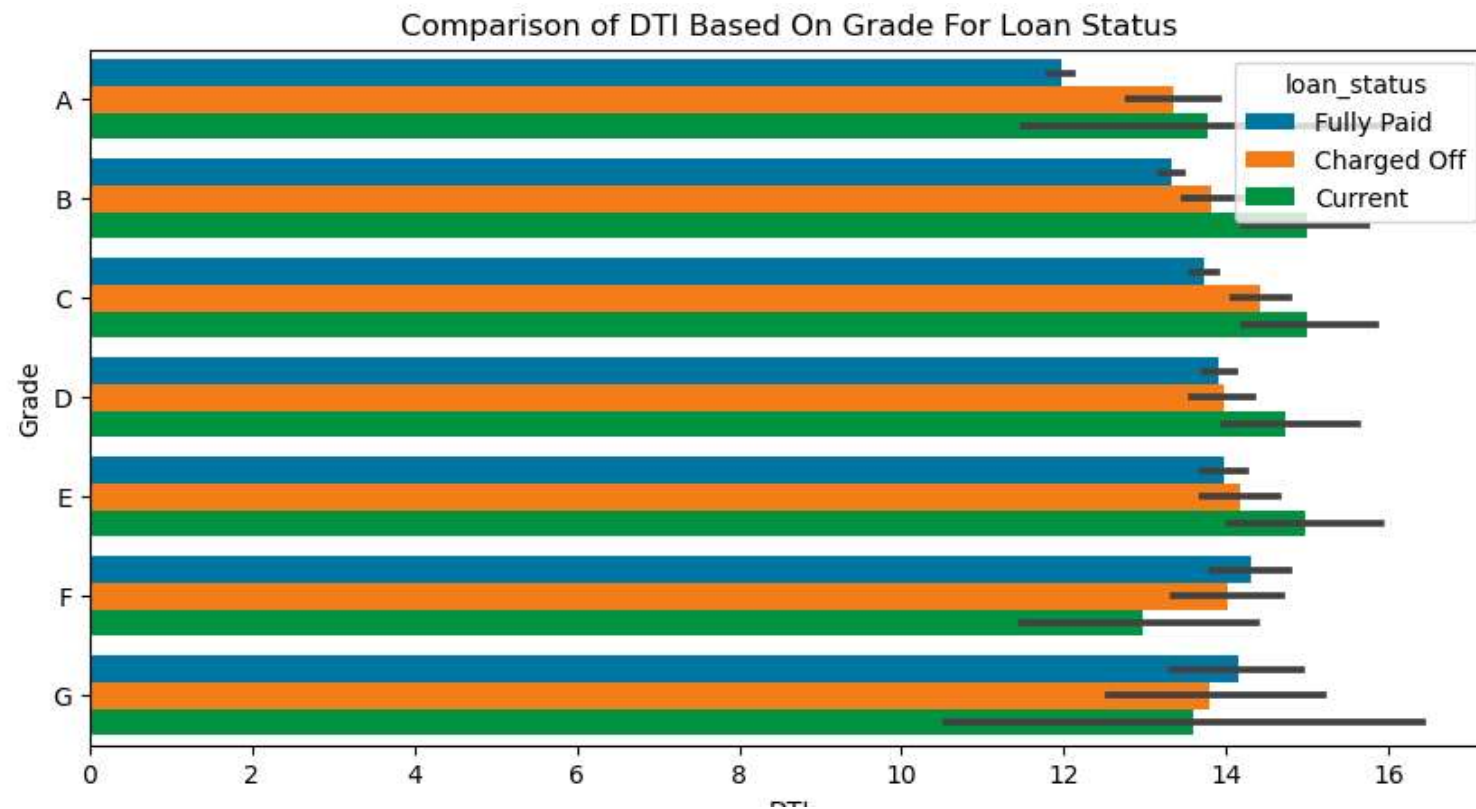
# Bivariate Analysis



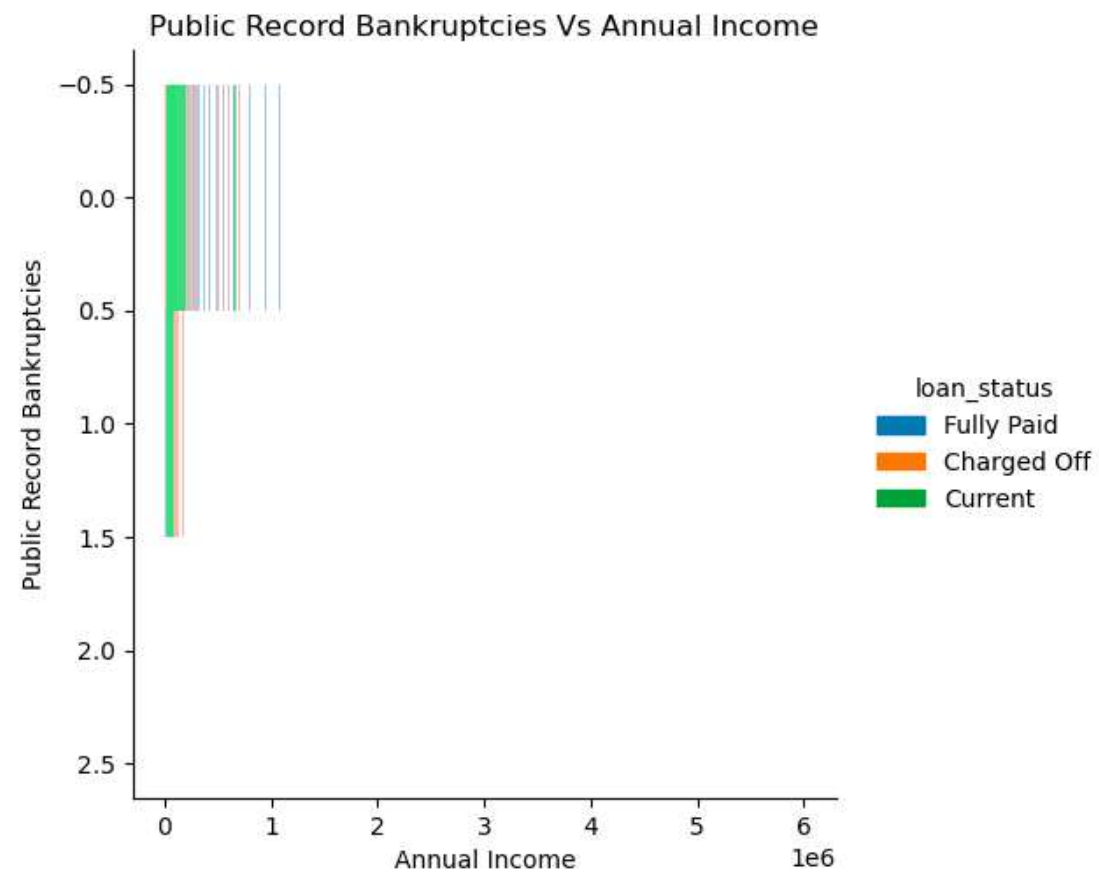
# Bivariate Analysis



# Bivariate Analysis



# Bivariate Analysis



# Conclusion

- Higher the interest rate higher charge off ratio
- Higher the annual income the loan amount is slightly higher
- The number of defaulted loan is 7 times less than the number of fully paid loan
- The majority of loan has a term of 36 months compared to 60 months.
- The interest rate is more crowded around 5-10 and 10-15 with a drop near 10
- Loan taken for small business purpose had to repay the loan with more interest rate as compared to other
- A large amount of loans are with grade 'A' and 'B' compared to rest showing most loans are high grade loan
- The Grade represent risk factor thus we can say interest rate increases with the risk
- The Grade A which is lowest risk also has lowest DTI ratio which we can say that higher grade has lower rate of default
- The borrowers are mostly having no record of Public Recorded Bankruptcy and are safe choice for loan issue
- Borrowers with least grades like E,F,G which indicates high risk

Thank You