

Lending Club Case Study

SHALEEN KUMAR TRIPATHI



Agenda

- ❑ Problem Statement
- ❑ Analysis Approach
- ❑ Analysis Observations
- ❑ Conclusion

Problem Statement

A famous online loan providing company want to minimize its lending risk i.e. 'defaulters'. Therefore, they wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

Analysis Approach

I have used Exploratory Data Analysis approach for identifying strong indicator for default cases, I have performed following major activities under this approach -

- ❖ Data cleaning

- Missing values
- Fixing Row and Columns
- Standardising Values
- Invalid Values

- ❖ Univariate Analysis

- ❖ Bivariate Analysis

Univariate Analysis

Observations

- Most annual incomes are between 50k to 70k
- Most ratio values are between 5 to 23
- Most cases are either fresh employee or very senior employee
- Most funding is between 1k to 10k
- Most investor funding is between 1k to 10k
- Most case are of 1 time installment default

Univariate Analysis

Observations

- Most installment are from 80 to 350
- Mostly offered interest rate is between 10 to 15%
- Most last payment amount in less than 800
- Most loan amount is between 1k to 10k
- Most applicant has 3 to 11 open account
- Most loans are issued to A & B person

Univariate Analysis

Observations

- Most of loans are unsecured
- 90% loans are fully paid
- Most loans are issued for Business purpose
- Most loans are not verified category
- Most people are applying for loan between Aug-Jan month
- Most loans are issued between Jul-Dec month

Univariate Analysis

Observations

- Most people are re-applying for loan with 6 months
- Most loans are issued between 2010 and 2011

Bivariate Analysis

Observation

- Most loan have been issued to approx. 50k annual income person
- DTI is high for B, C, D grade person
- Loan Issue cases for Grade A people is less than Grade G
- Grade B and D has more open accounts
- Grade G person are taking frequent loans

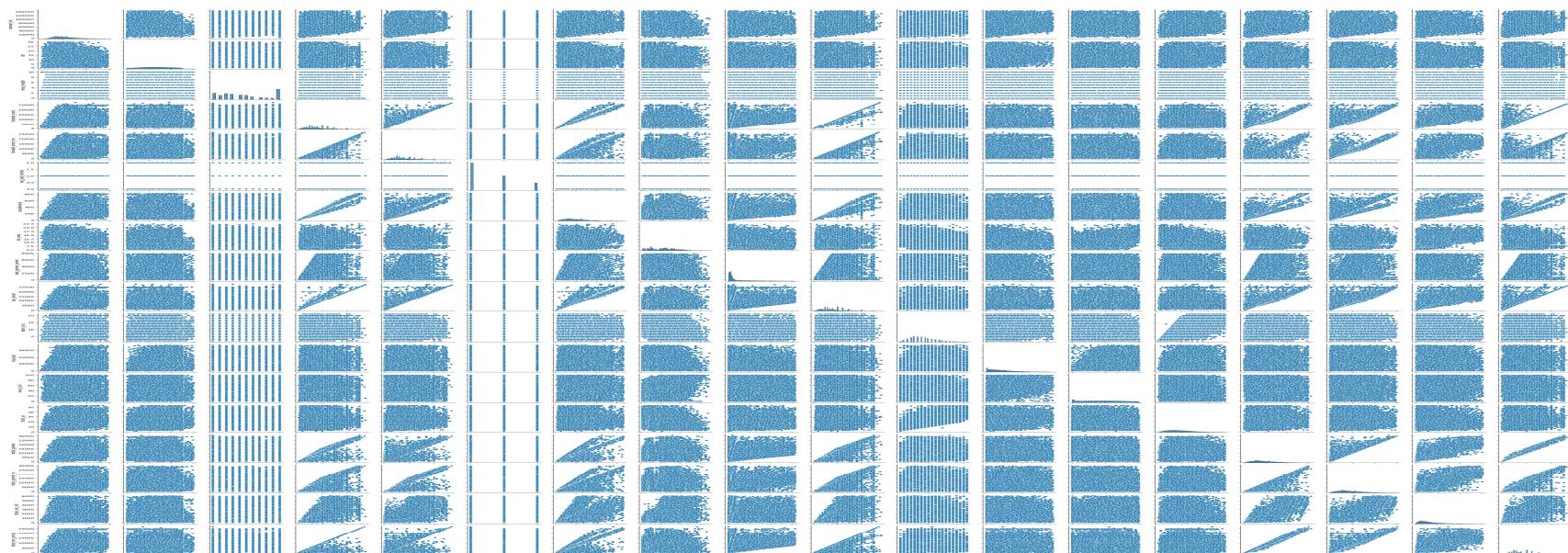
Bivariate Analysis

Observation

- DTI is high for not verified loans
- Funded Amount is higher for Charged Off case in comparison of Fully Paid
- DTI is high for Rented person

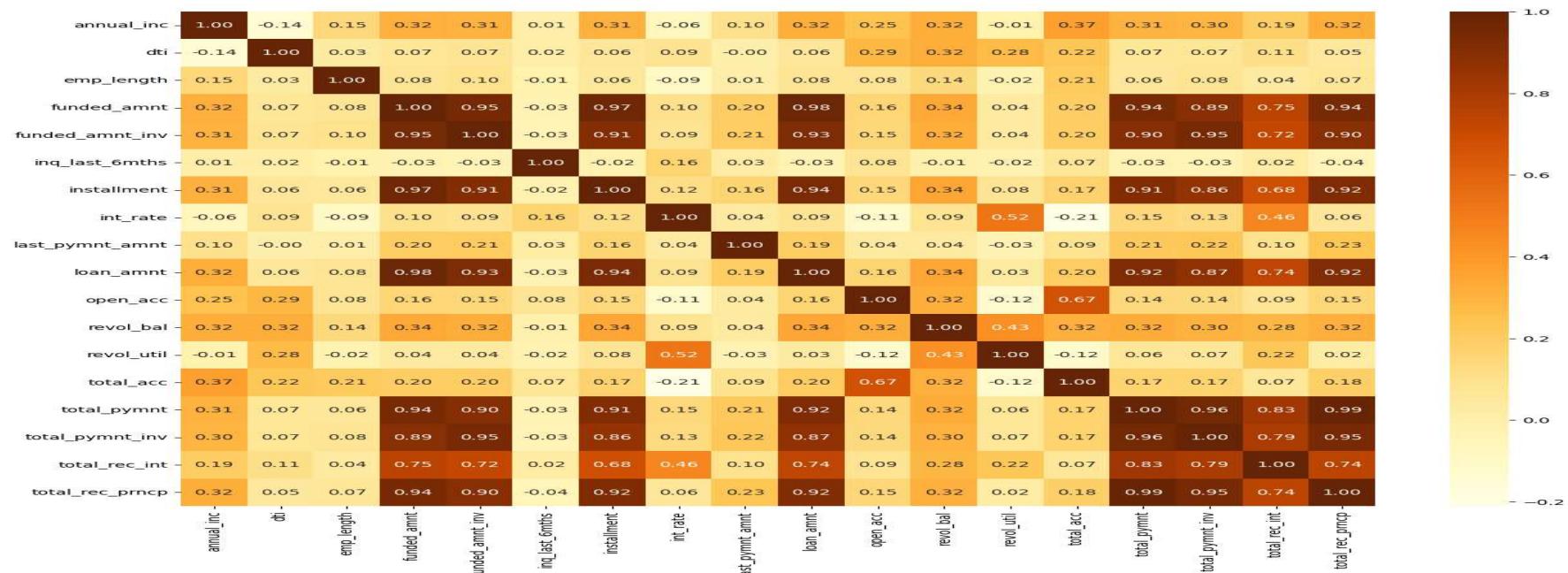
Bivariate Analysis

Pair Plot Graph with Quantitative variables



Bivariate Analysis

Heat Map with Quantitative variables

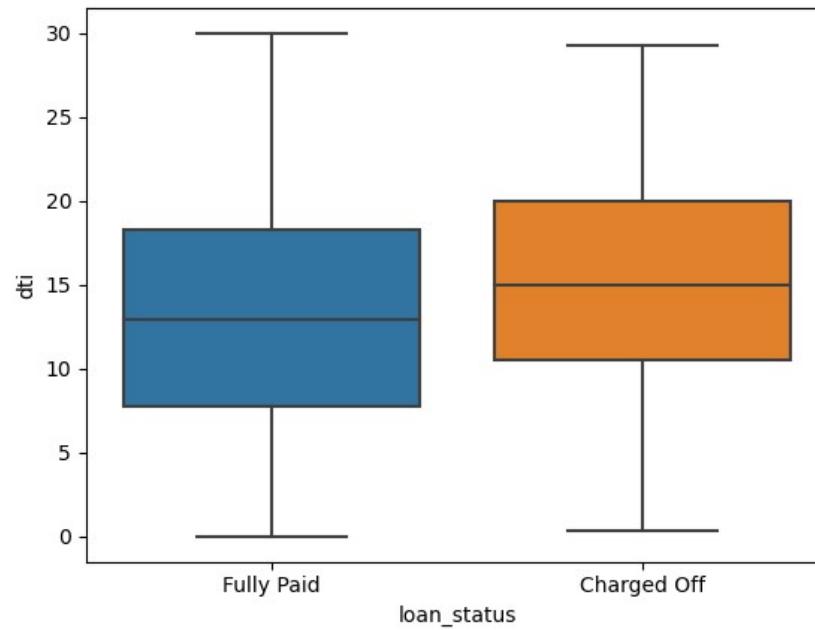


Conclusion

- ❑ DTI is a very strong variable which shows higher chance of default of loan payment.
- ❑ EDA shows that person, who have taken loan more than then annual income, has strong chances of default.
- ❑ High DTI for non-verified loan cases are also strong indication for default.
- ❑ Low grade borrower, who has higher DTI, has more risk of default.
- ❑ Lower grade borrower who has frequently re-financing loan, has higher chance of default.

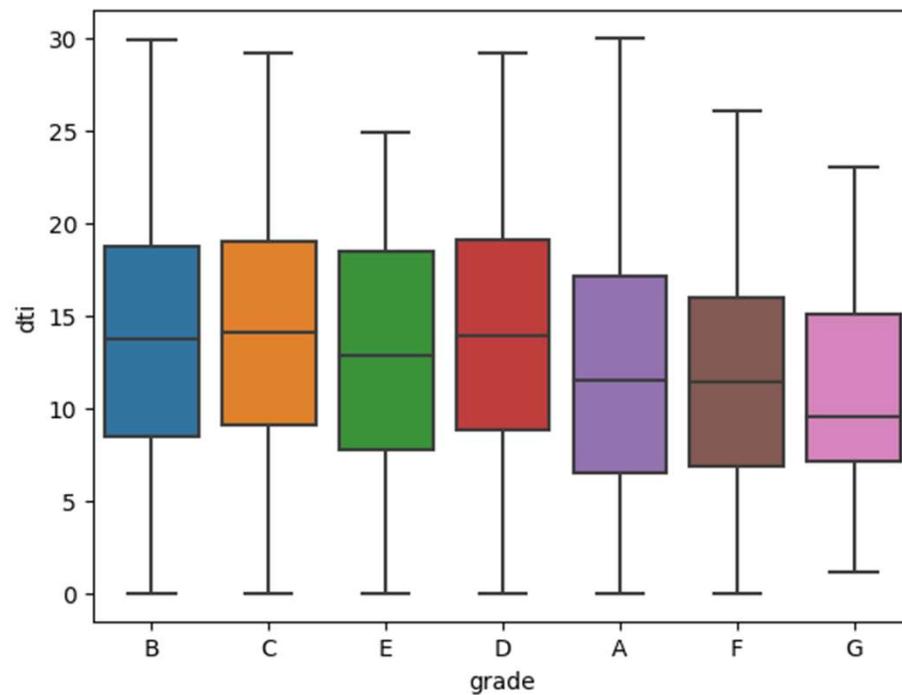
Conclusion

DTI Vs Loan Status



Conclusion

DTI Vs Grade

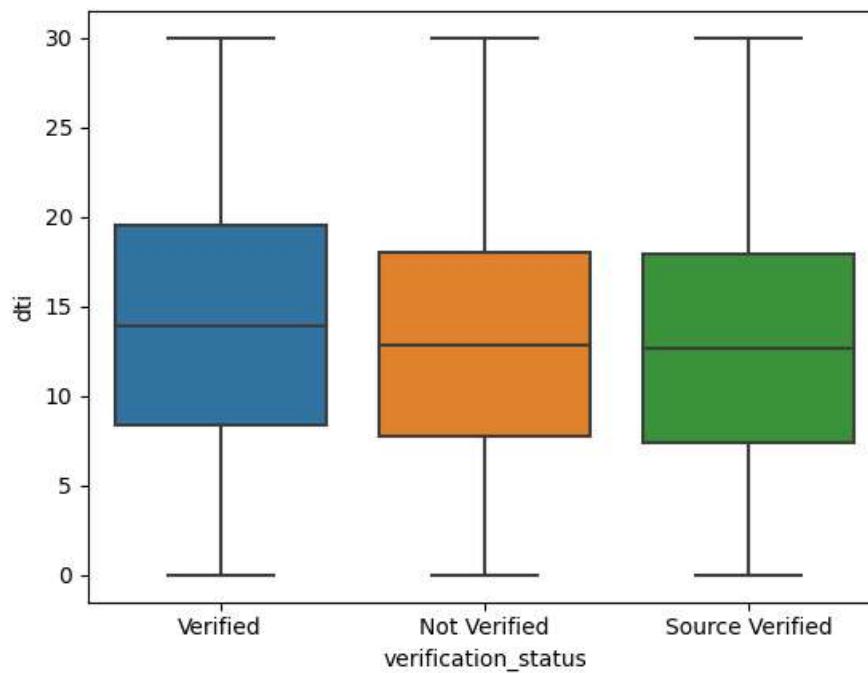


Conclusion

DTI Vs

Loan Verification

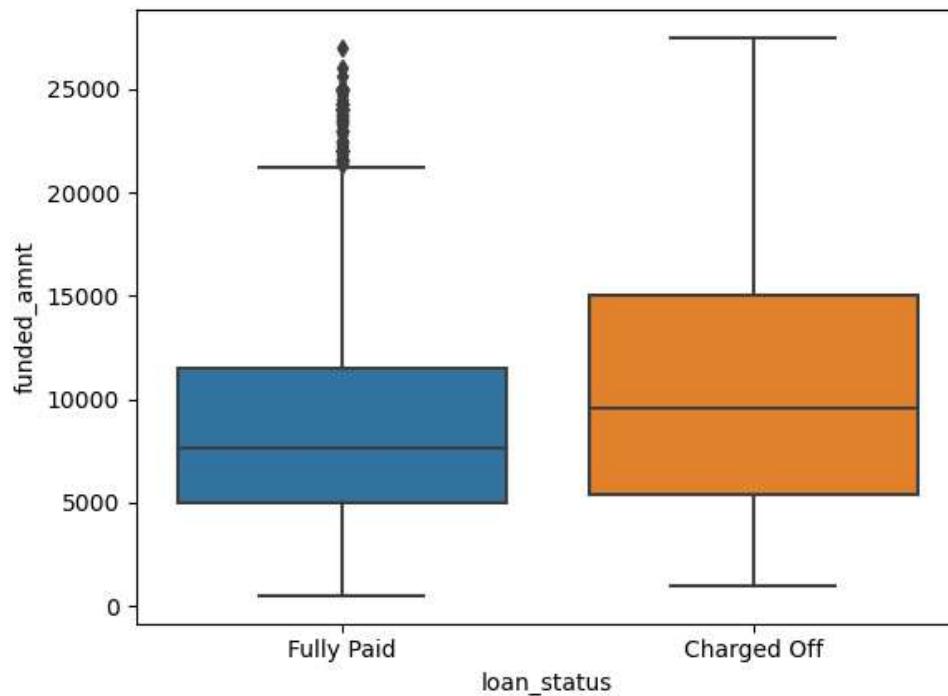
Status



Conclusion

Funded Amount

Vs Grade

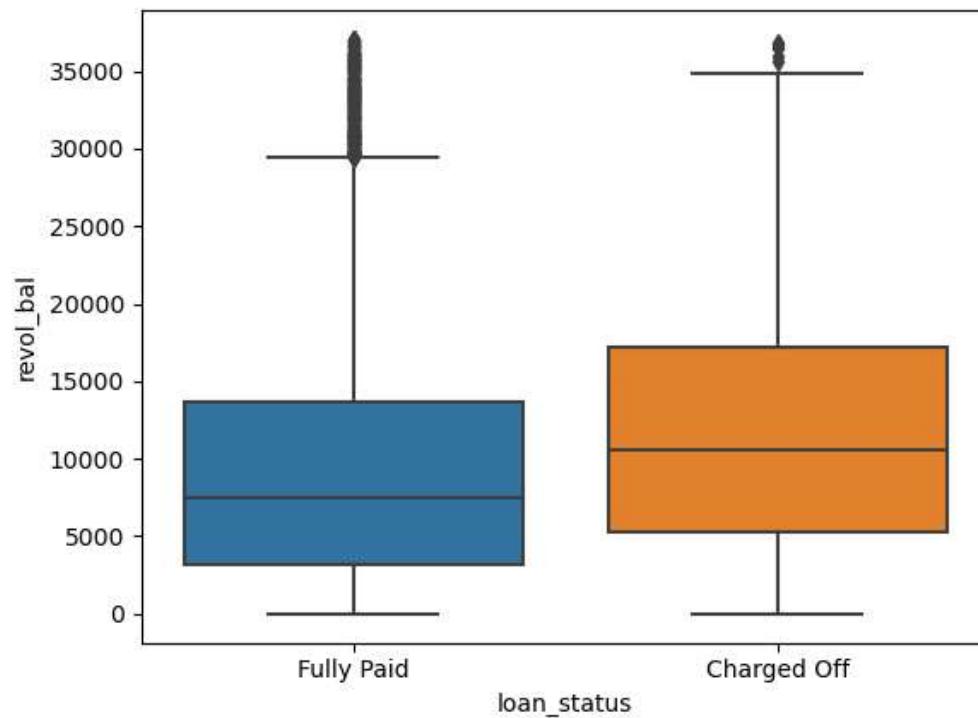


Conclusion

Total credit

revolving balance

Vs Loan Status



Conclusion

Total Accounts

Vs Loan Status

