Lending Club Case Study

Team Faciliator:Amreeta Sharma

Team Member:Muskaan Sharma

Problem Statement

- The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.

Analysis approach

❑ While studying this Case Study we have followed the systematic approach for Exploratory Data Analysis (EDA)

- Loading Dataset
- Checking for null and duplicate values
- Removing Null values from all the columns which are having null values.
- Exploratory Data Analysis
  - Checking Outliers
  - Insights from the data

Univariate Analysis

- In univariate analysis, there is only one dependable variable.

- In our example we have used histogram for getting insights of annual income . As shown in fig 1.1, most of the applicants count have an annual income between 0 to 50,000.
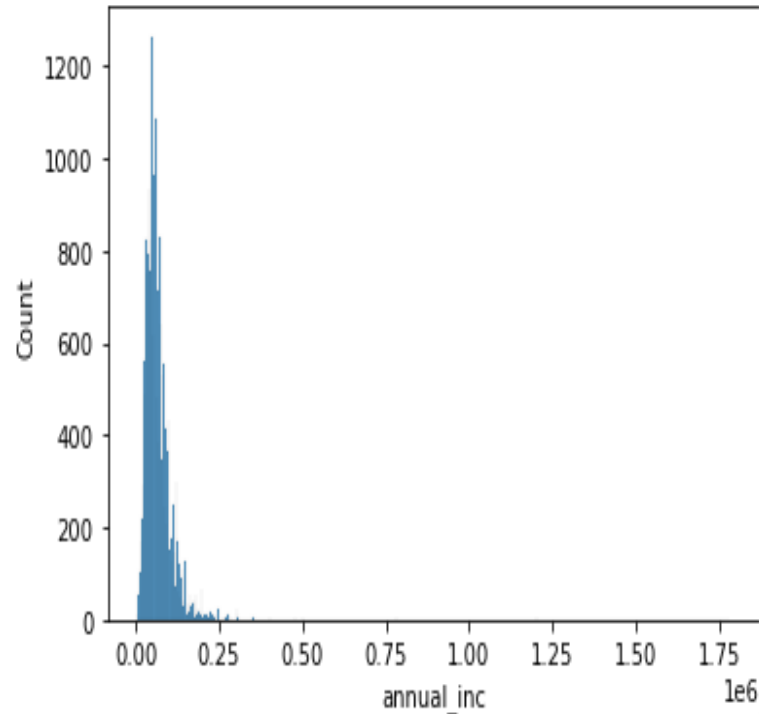

Fig 1.1

- In our second example of univariate analysis, we are plotting the histogram for loan_amnt.

- As shown in fig 1.2, Most of the loan amount is between 5000 to 10000.

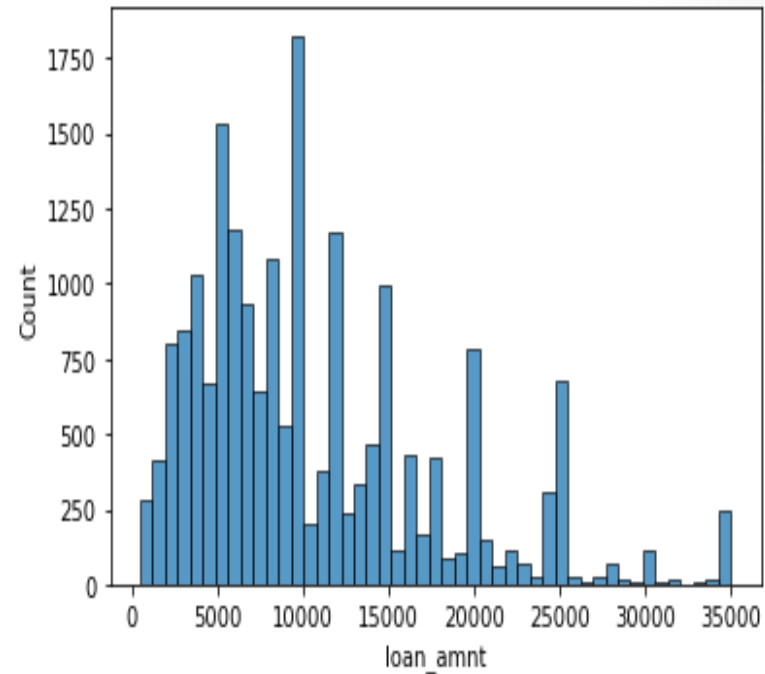- We can also find most of the values between 10000 to 15000.

Fig 1.2

Bivariate Analysis

- With bivariate analysis, plotting the graph makes it easier to check if two variables are related to each other.

- Probability of fully paid loan amount is higher between 5000 to 15000.

- Probability of charged-off loan amount is higher between 5000 to 18000.

- Probability of Current loan amount is higher between 9000 to 23000.

- Avg amount is 10000 and the density is higher between 5000 to 10000.

- For charged off avg amount is 10000 and high density is between 5000 to 9000.

- For Current avg would be 12000 and the density is higher density is between 5000 to 15000.
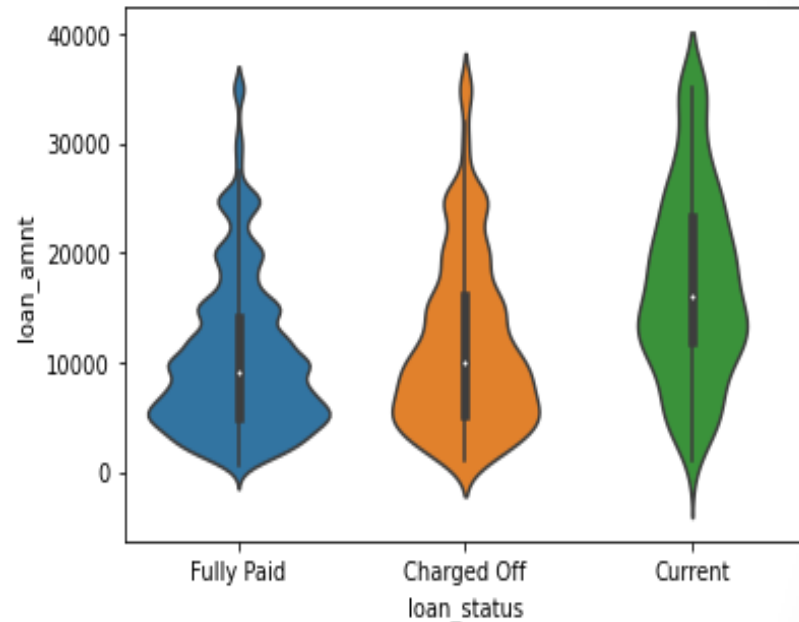

Fig 2.1

- Here, DTI is the ratio of annual income and the amount deducted from monthly salary for debt_payment.

- Now in fig 2.2, we are trying to compare the dti

  and loan_status.

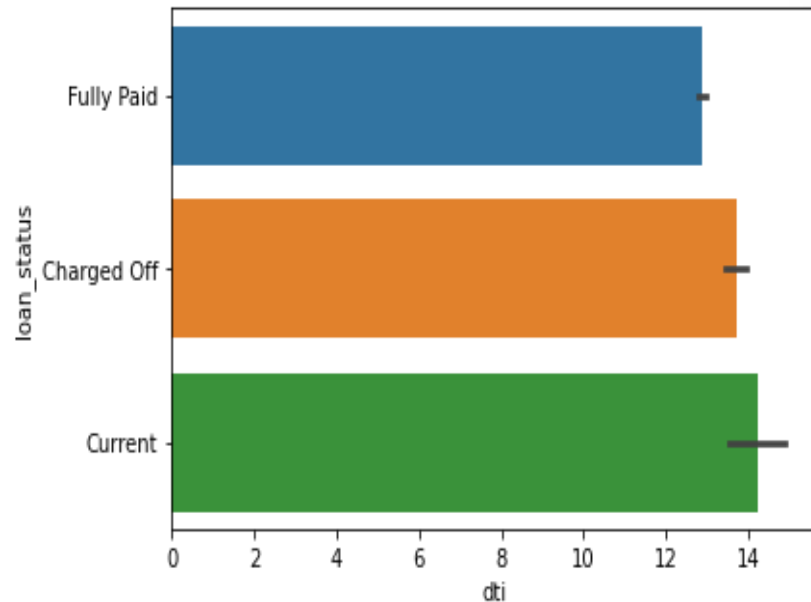- In comparison with fully paid and charged-off applicants, Current applicant are tend to pay more dti amount.

Fig 2.2

- In fig 2.3, below are the given insights which we got with our analysis

- Avg installments paid by fully paid applicant is approx 300

- Avg installments paid by Charge off applicant is approx 290

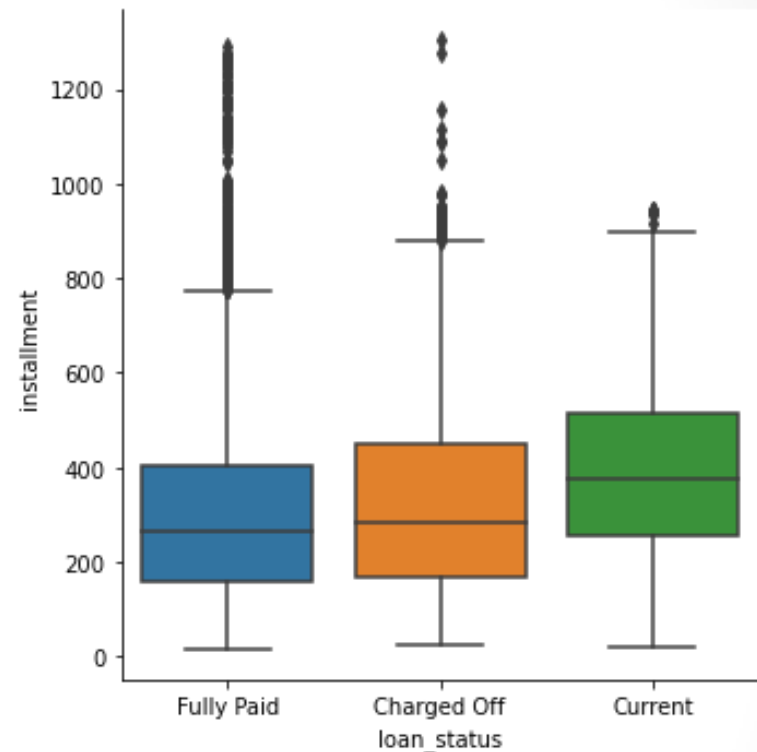- Avg installments paid by Charge off applicant is approx 390

Fig 2.3

- Here we are using the Stripplot, and got insight that most of the principal amount is received from applicant who have fully paid on time.
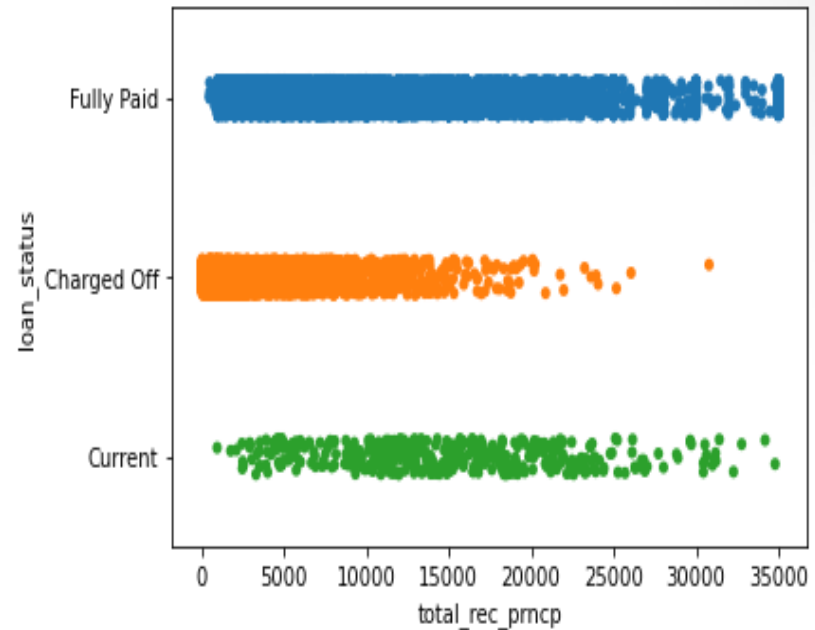

Fig 2.4

Conclusion

From Our Data Analysis some of the insights we got are as follows:-

- Majority of people are taking loan for debt consolidation purpose.

- There is a constant relation of revol_bal and the loan_amnt fully paid are able to pay the amount sucessfully.

- The most of the fully paid applicants are not verified.

- There are no such variables affecting the loan_status

- average annual income for each term is same .

- average dti for each term is same

Thank You!!