Lending Club Case Study

Identifying patterns for risky loans

The problem

Company

Lending Club is an online marketplace for lenders and borrowers.

The platform facilitates investors to lend money to borrowers who apply for loans.

Context

Several of these loans are unpaid and are written off by the company.

These are loss making loans and need to be reduced/avoided.

Problem statement

Identify patterns in the past loans to check what could be causing these borrowers to become defaulters.

Approach

Exploratory Data Analysis

Data to be analysed with respect to columns and their relationships to derive key insights

New metrics may be derived using existing data to find more insights

Some key insights could be derived post analysis of the existing loan data.

Approach steps:

- Data understanding
- Data cleanup
- Data analysis
  - Univariate
  - Bivariate
- Insights and Recommendations

About the data

Meta data

Description	the complete loan data for all loans issued
Source	Upgrad
Format	csv
Number of rows	39717
Each row is	Details of a particular loan issued
Sampling method	loans issued through the time period 2007 to 2011.

Looking into bad loans

Findings from the bad loans

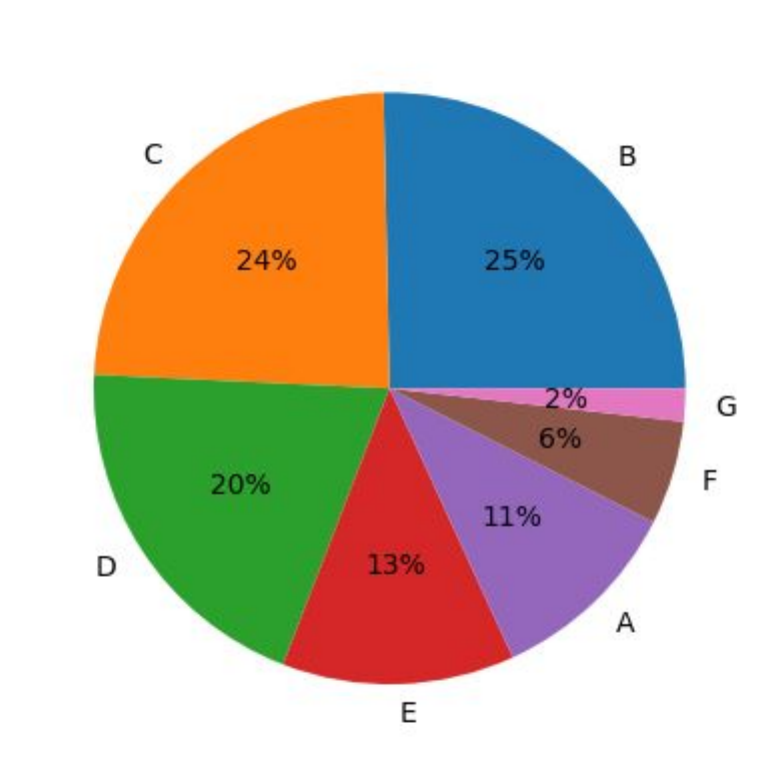
- most charged off asked loans are around 10000

- most charged off loans have an interest rate of around 10 to 15%, while interest rates start from 5.4% and go all the way up to 24%

- majority of income distribution is around 30000 to 60000

Grade Insights

- majority of charged off loans fall under B, C, D grades

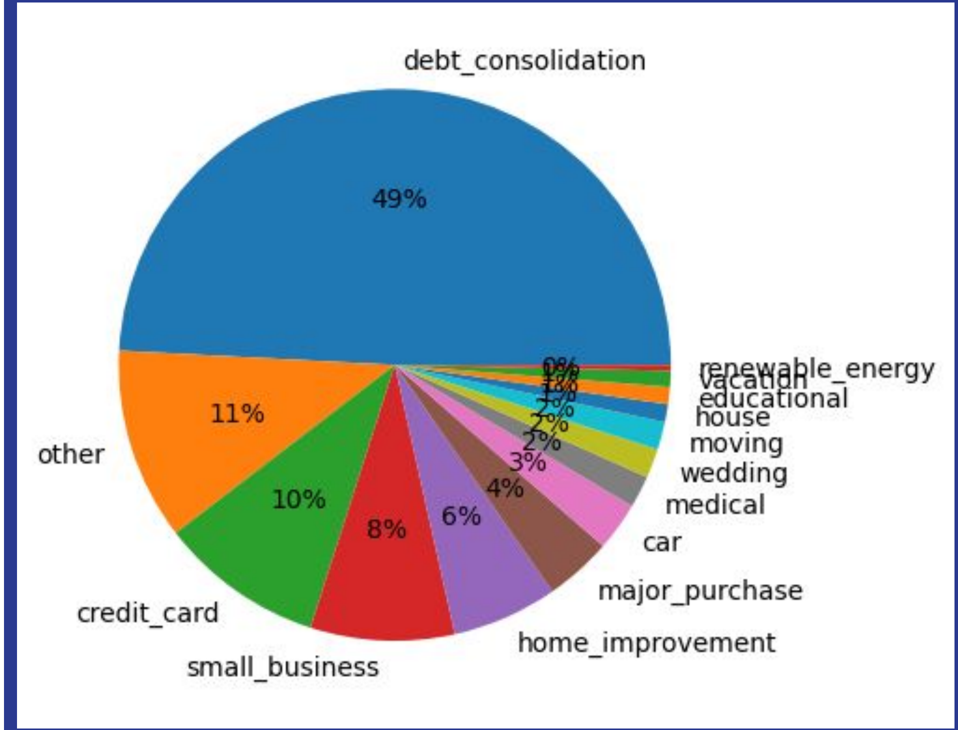


Purpose Insights

- half of the loan applications were for **debt consolidation**, meaning loans were taken to clear other loans

- amongst the other half, credit card debt, small businesses and 'other' purposes were significant

- minimum loan amounts were asked for moving and vacations

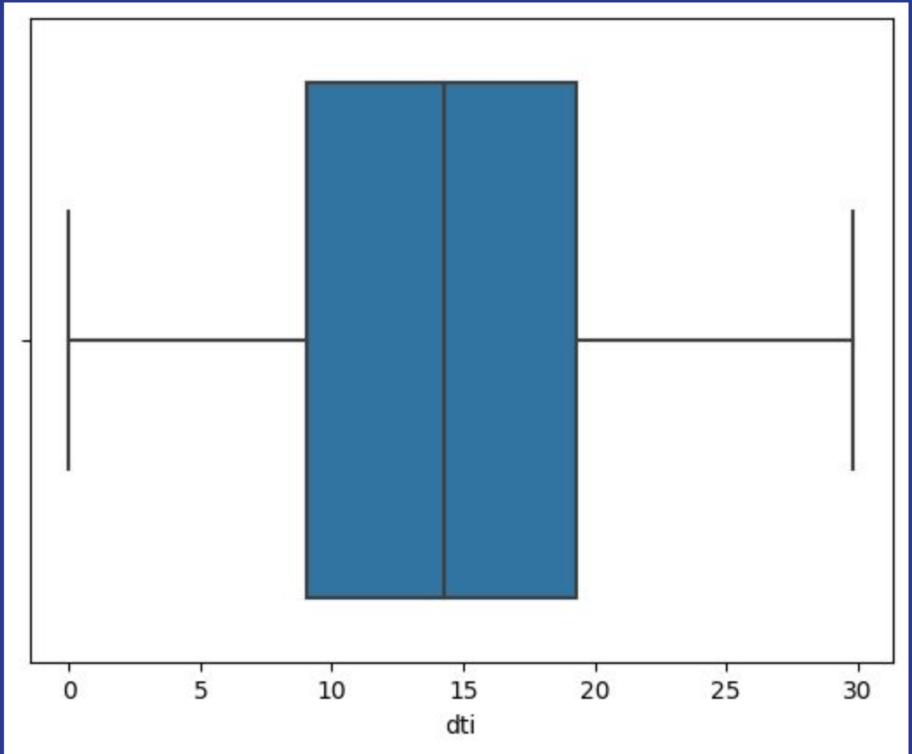


DTI Insights

- amongst the charged off loans, DTI (debt to income ratio) was distributed mostly between 7.5 to 25%

- according to [Wikipedia], a DTI of 43% is typically the highest ratio a borrower can have and still get qualified for a mortgage, but lenders generally seek ratios of no more than 36%.

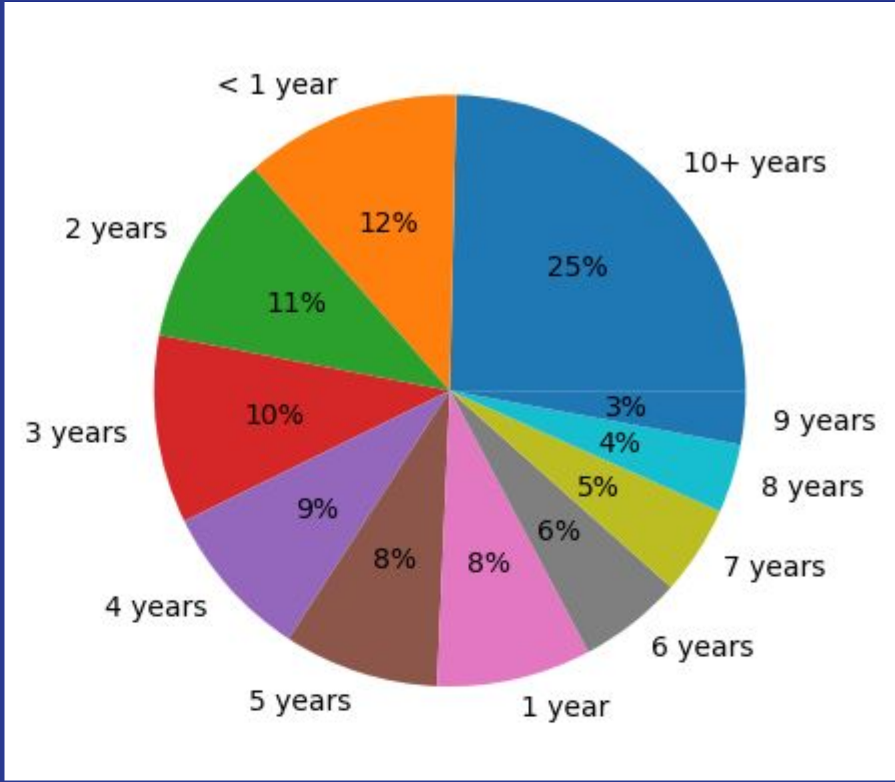
- this shows that, while applying for the loans their DTI were up to the standard.



Employment Insights

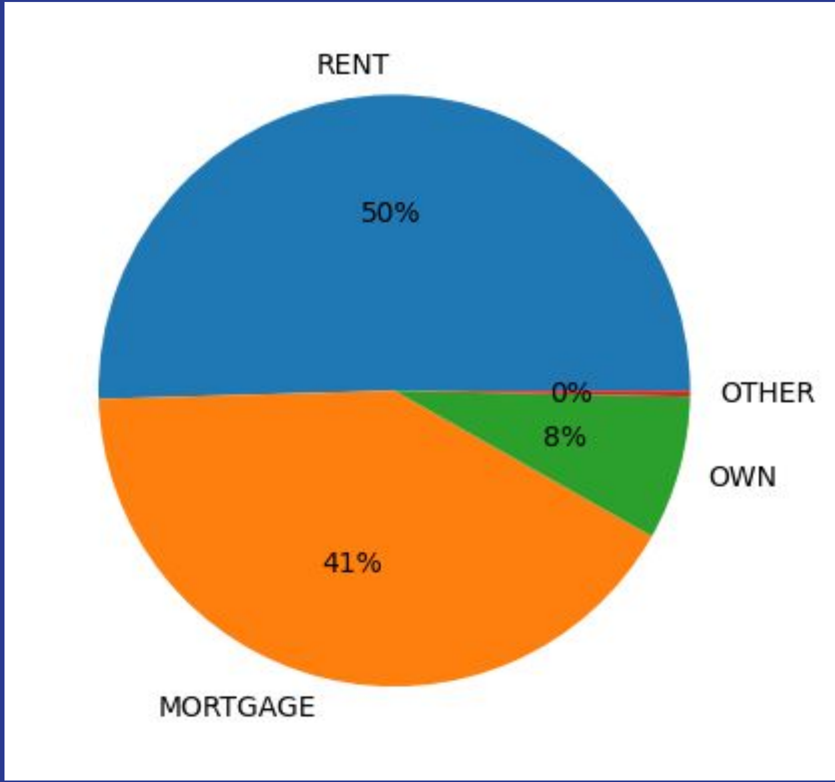
- amongst the loans, about a quarter of loans had applicants with employment of more than 10 years

- more than half of the applications had employment of more than 5 years



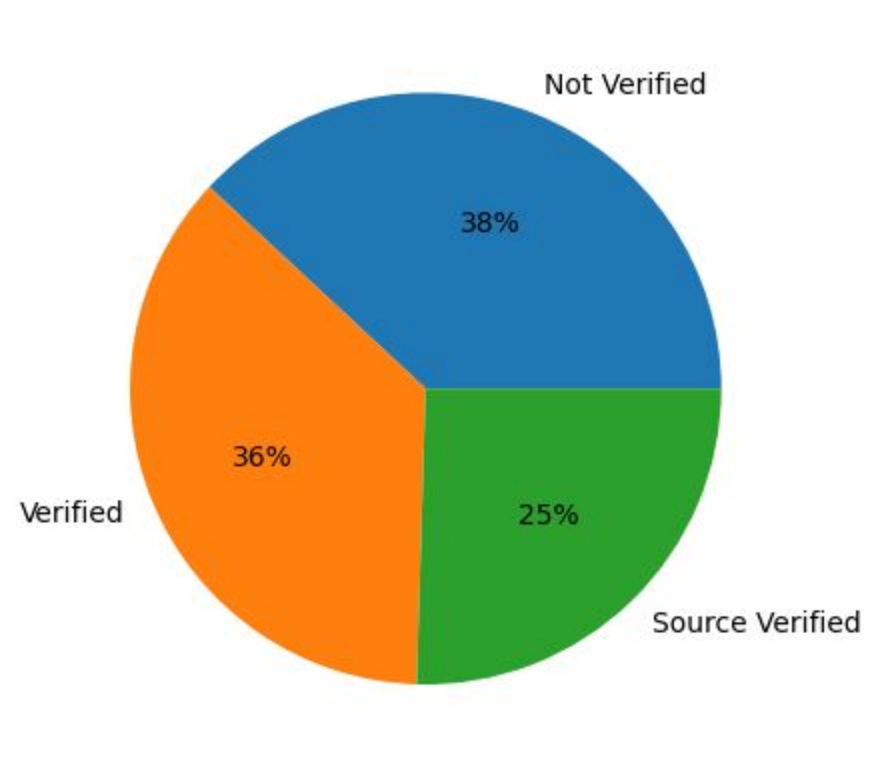
Home Ownership Insights

- more than 90% of loans don't own a home, they are either living on rent or have an running mortgage



Income Verification Status Insights

- about 38% of charged off loans don't have their income verified



Income Verification Status Insights

- verified applicants received close to the amount requested

- surprisingly, unverified income source applicants got loans at a lower interest rate

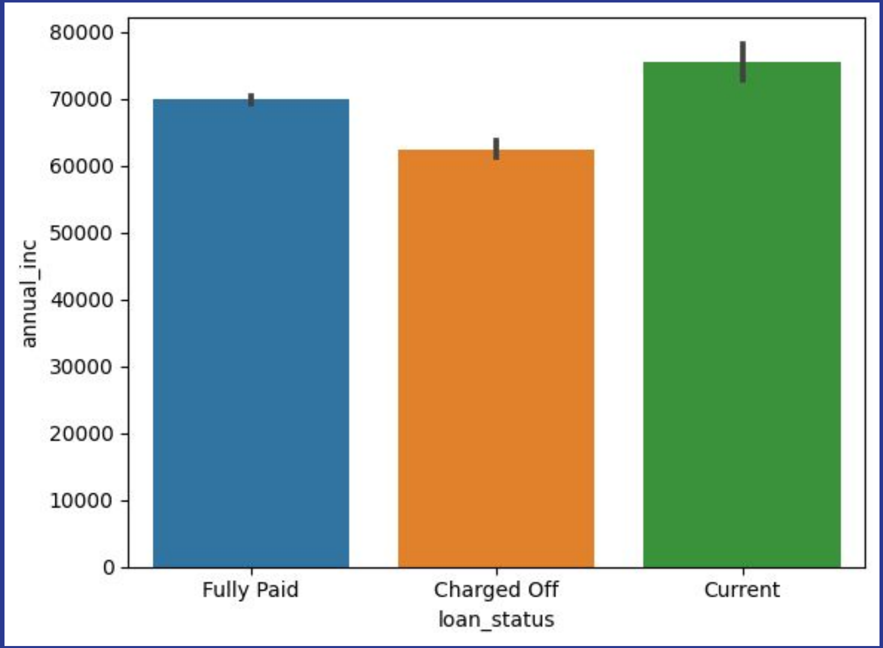
	loan_amnt	funded_amnt_inv
verification_status		
Not Verified	8667.238562	7308.538055
Source Verified	10686.506276	10040.727229
Verified	16685.372989	15154.251650

```
verification_status
Not Verified       12.689253
Source Verified    14.133312
Verified           14.783042
```

Comparing good loans with bad

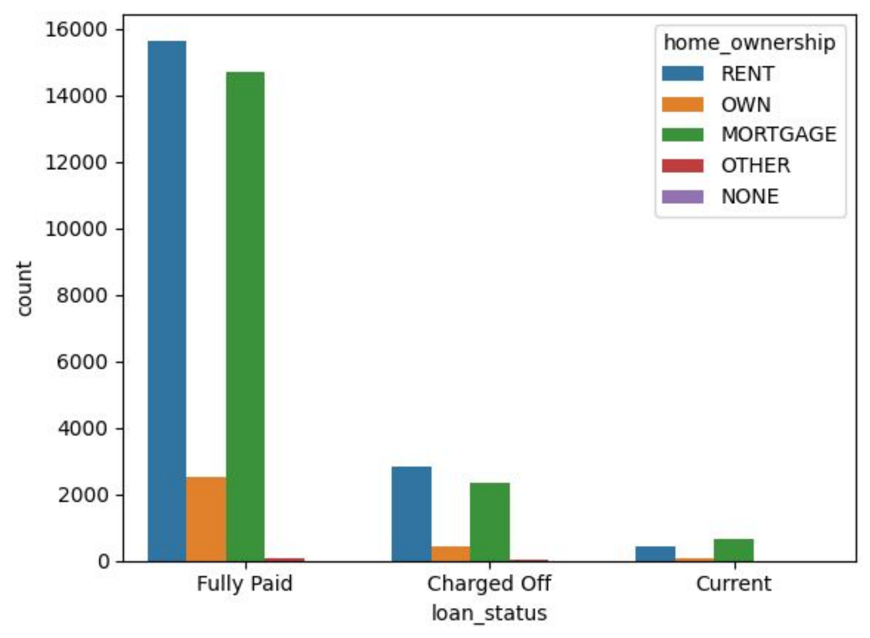
Annual Income Insights

- income under 70000 have defaulted more



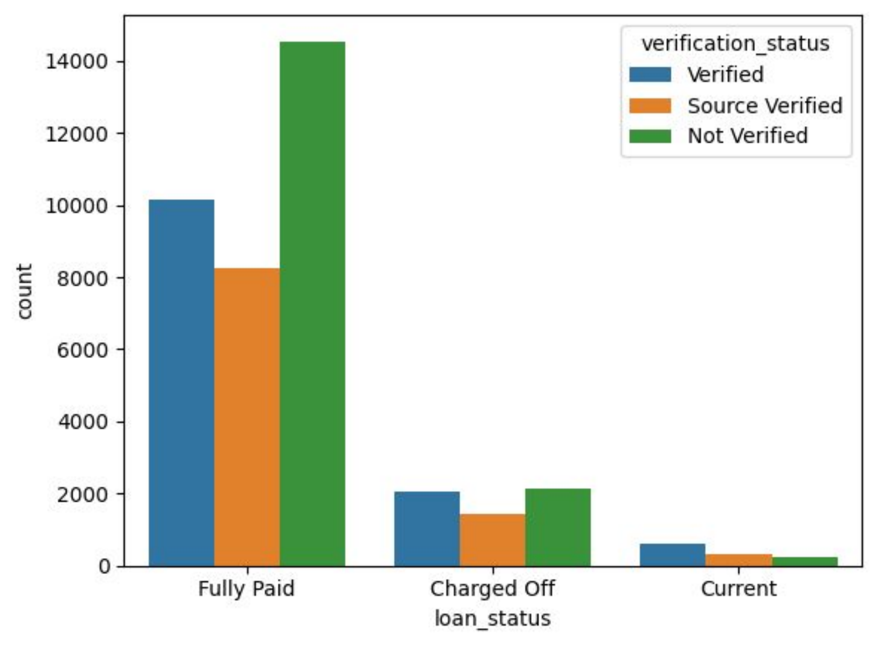
Home Ownership Insights

- people who are already having a mortgage loan or are paying rents have defaulted more



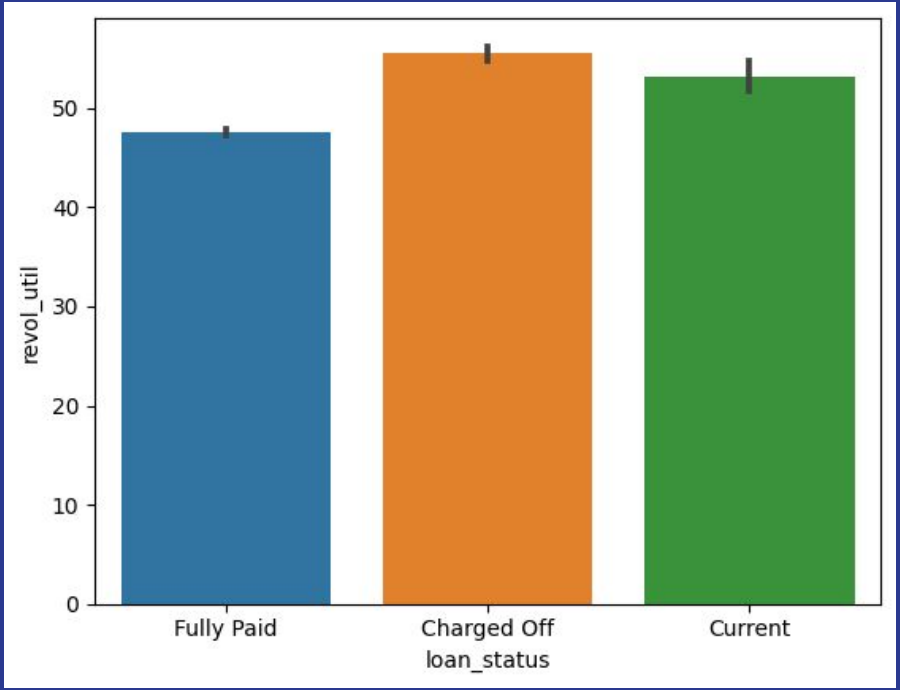
Verification Status Insights

- verified and unverified loans have same chances of defaulting



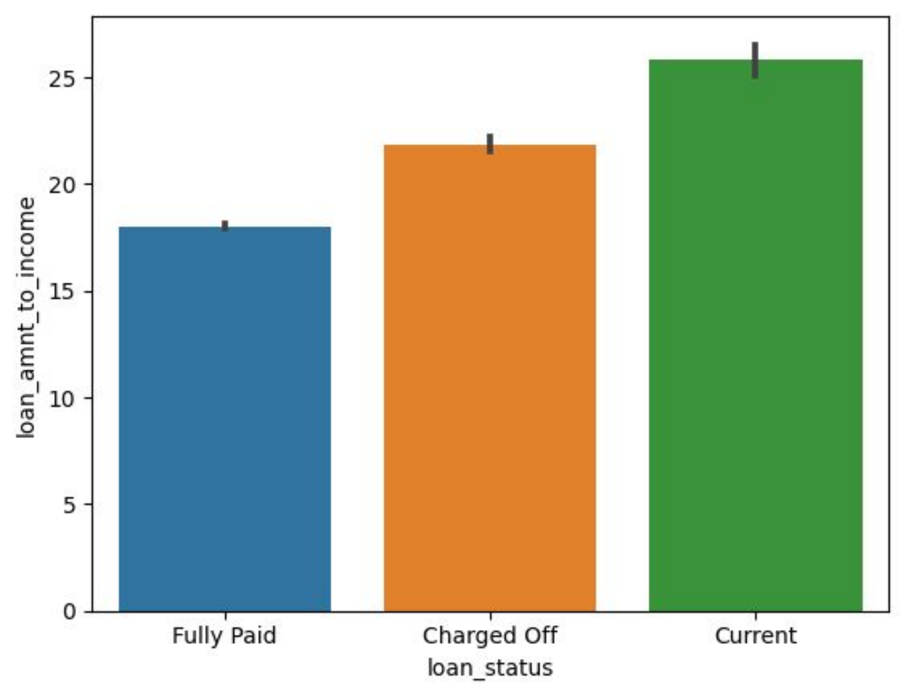
Revolving Line Utilization Rate Insights

Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit.
- loans with revol_util more than 50% have defaulted more



Loan Amount to Annual Income* Ratio Insights

- When ratio is higher, loans have defaulted more



*this is a derived metric

Recommendations

Key drivers of risky loan applications

- Lower interest rates
- Purpose of the loans
- Home ownership
- Annual income
- Revolving utilization of credit
- Debt to Income ratio
- Income source verification

Identifying risky loan applications

- care needs to be taken while providing loans to **unverified** income sources at **lower interest rates**
- debt consolidation and credit card loans seem to be **high risk** as they are serving to clear other loans which only increases the DTI
- verify income sources when home ownership in currently under **mortgage** as this loan would only add to the debt
- when annual income is less than 70000, loans have defaulted more - check **loan to income** ratio to decide
- revolving utilization of credit is a good measure to check likelihood of defaulting

Thank you