Lending Club Case Study: Financial Credit Risk

Case Study By:

M. Bharat

Puneet Bansal

Abstract

Credit risk analysis and credit risk management is important to financial institutions which provide loans to businesses and individuals.
Credit risk can occur for various reasons such as bank mortgages (or home loans), motor vehicle purchase finances, credit card purchases, installment purchases, and so on. Credit loans and finances have risk of being defaulted.
To understand risk levels of credit users, credit providers normally collect vast amount of information on borrowers. Some predictive analytic techniques can be used to analyze or to determine risk levels involved on credits, finances, and loans, i.e., default risk levels

Synopsis

You work for a consumer finance company which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to decide for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

a) If the applicant is likely to repay the loan, then not approving the loan results is a loss of business to the company

b) If the applicant is not likely to repay the loan, i.e., he/she is likely to default, then approving the loan may lead to a financial loss for the company

When a person applies for a loan, there are two types of decisions that could be taken by the company:

Loan accepted: If the company approves the loan, there are 3 possible scenarios described below:

A. **Fully paid**: Applicant has fully paid the loan (the principal and the interest rate)
B. **Current**: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
C. **Charged-off**: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan

Loan rejected: The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset).

Business Objectives

This company 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.

Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicant's using EDA is the aim of this case study.

In other words, the company 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 utilise this knowledge for its portfolio and risk assessment.

Findings:

- For our analysis we only consider Loan_status as Fully paid or Charged off. Neglecting Loan_status as Current status.
- Clean the data to fill null values and outlier imputation.
- Drop unnecessary columns to keep strong driving factors (or driver variables).
- After cleaning and imputation process we found below listed variables which sounds strong variables in this analysis.

```
Data columns (total 22 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   id                      36837 non-null  int64
 1   member_id               36837 non-null  int64
 2   loan_amnt               36837 non-null  int64
 3   funded_amnt             36837 non-null  int64
 4   funded_amnt_inv         36837 non-null  float64
 5   term                    36837 non-null  object
 6   int_rate                36837 non-null  object
 7   installment             36837 non-null  float64
 8   grade                   36837 non-null  object
 9   sub_grade               36837 non-null  object
 10  emp_length              36837 non-null  object
 11  home_ownership          36837 non-null  object
 12  annual_inc              36837 non-null  float64
 13  verification_status     36837 non-null  object
 14  issue_d                 36837 non-null  object
 15  loan_status             36837 non-null  object
 16  purpose                 36837 non-null  object
 17  title                   36837 non-null  object
 18  addr_state              36837 non-null  object
 19  dti                     36837 non-null  float64
 20  total_acc               36837 non-null  int64
 21  pub_rec_bankruptcies    36837 non-null  float64
dtypes: float64(5), int64(5), object(12)
```

Hypothesis:

Study will confirm or reject the following hypotheses :

a) There is a relationship between the Lending Club loan decision and borrowers' employment length.

b) Grade of Borrower's working company has impact on interest rate and Loan decision.

c) Loan decision is impacted by Lending interest rate and home_ownership of borrowers

d) There is a mutual relationship among loan_amount , funded_amount, funded_amnt_inv and Monthly installment.

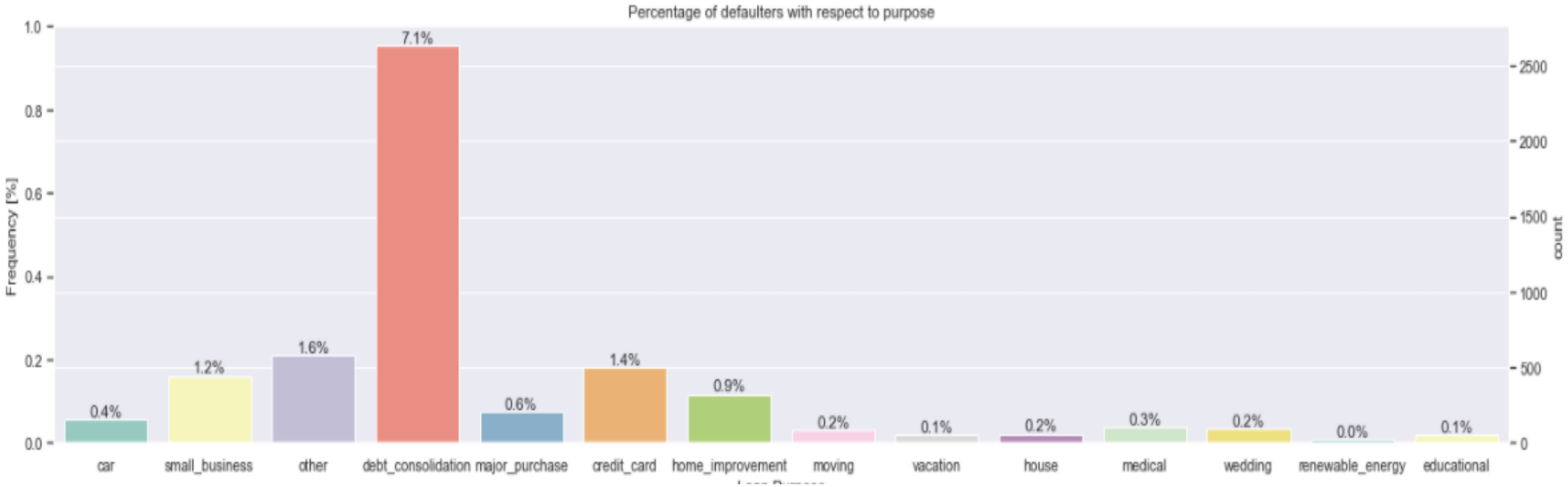
e) There is a relationship of employment length and Grades with Lending amount Interest rate

f) Loan status (higher default rate) is dependent on Loan term and monthly installment amount

g) Loan status is also dependent on Borrower's home_ownership and employment tenure.

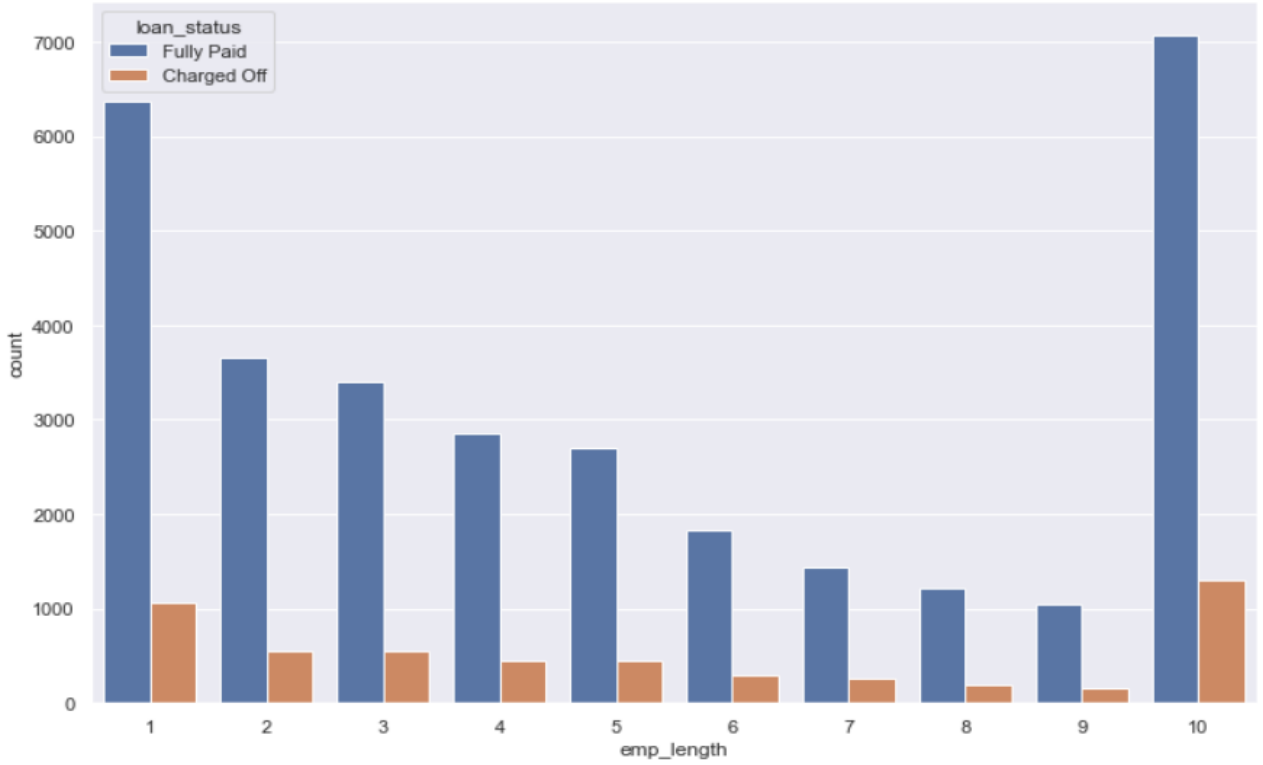
Results:

a) Analysis of Defaulted accounts with respect to Purpose.



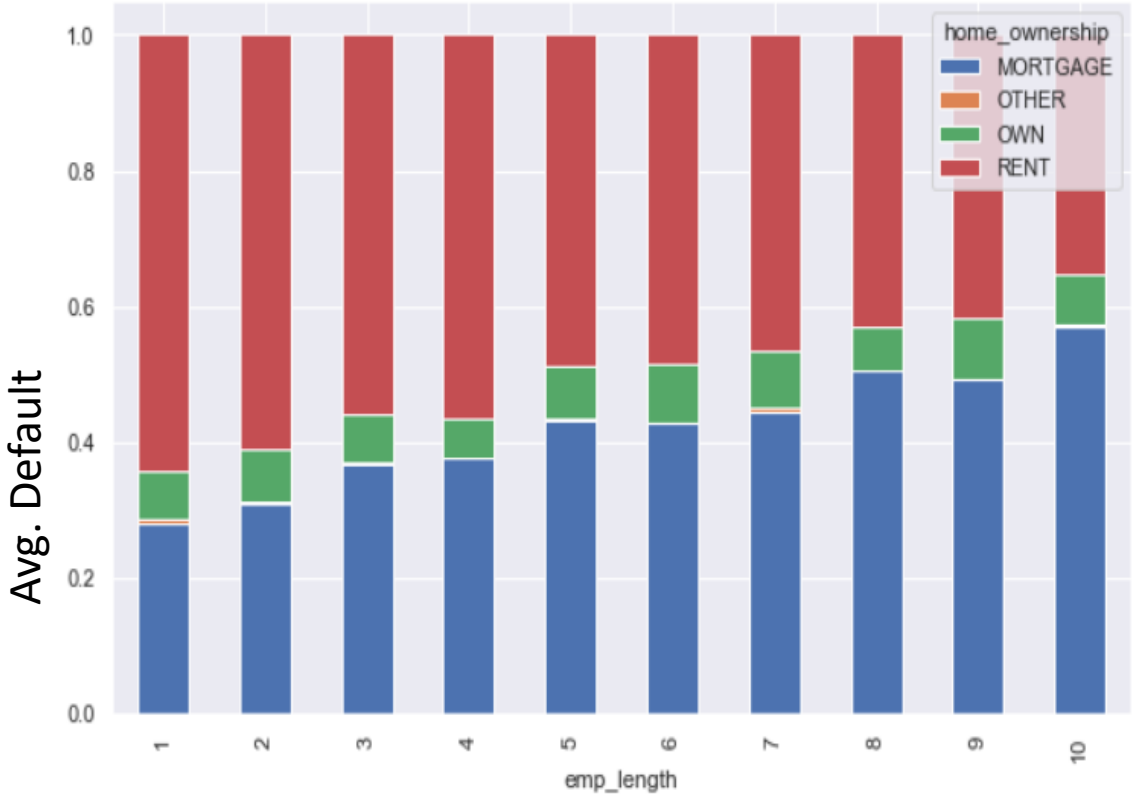
Conclusion：As we observe that most defaulters are having debt_consolidation purpose.

b) Analyzing impact of Employee Length (Employee exp.) on Loan status. → Univariate analysis



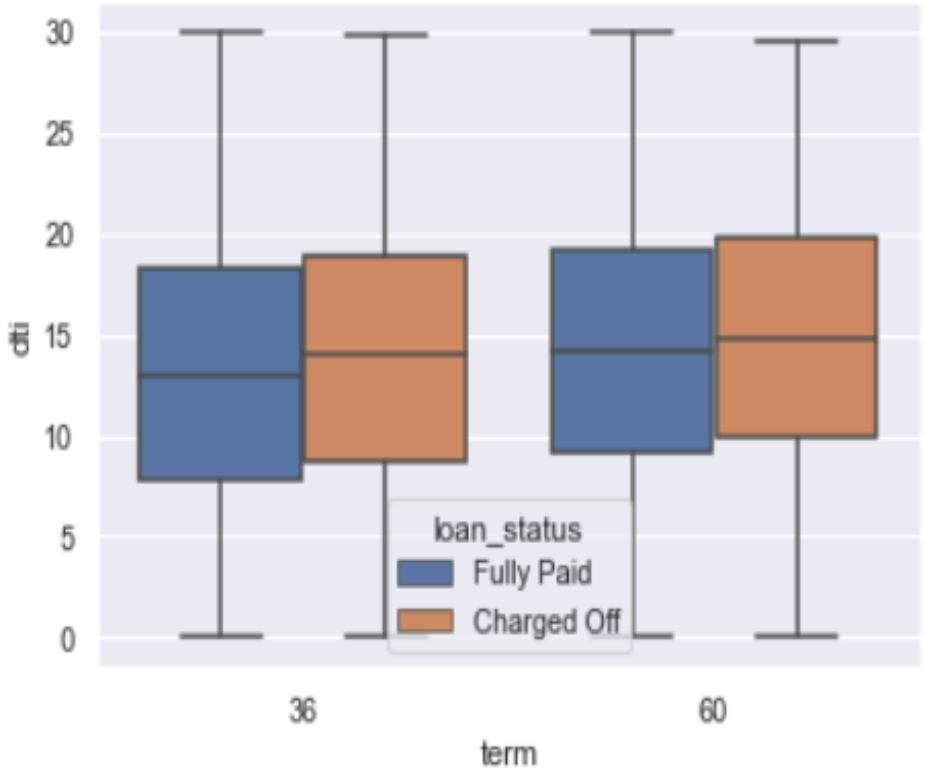
Conclusion : From above Bar graph, we can analyze that, bad Loan increases with employment Length 1 and 10/10+ years of experience.

c) Analysis with Employee_Length, Home_Ownership and Loan_Status as Defaulters.



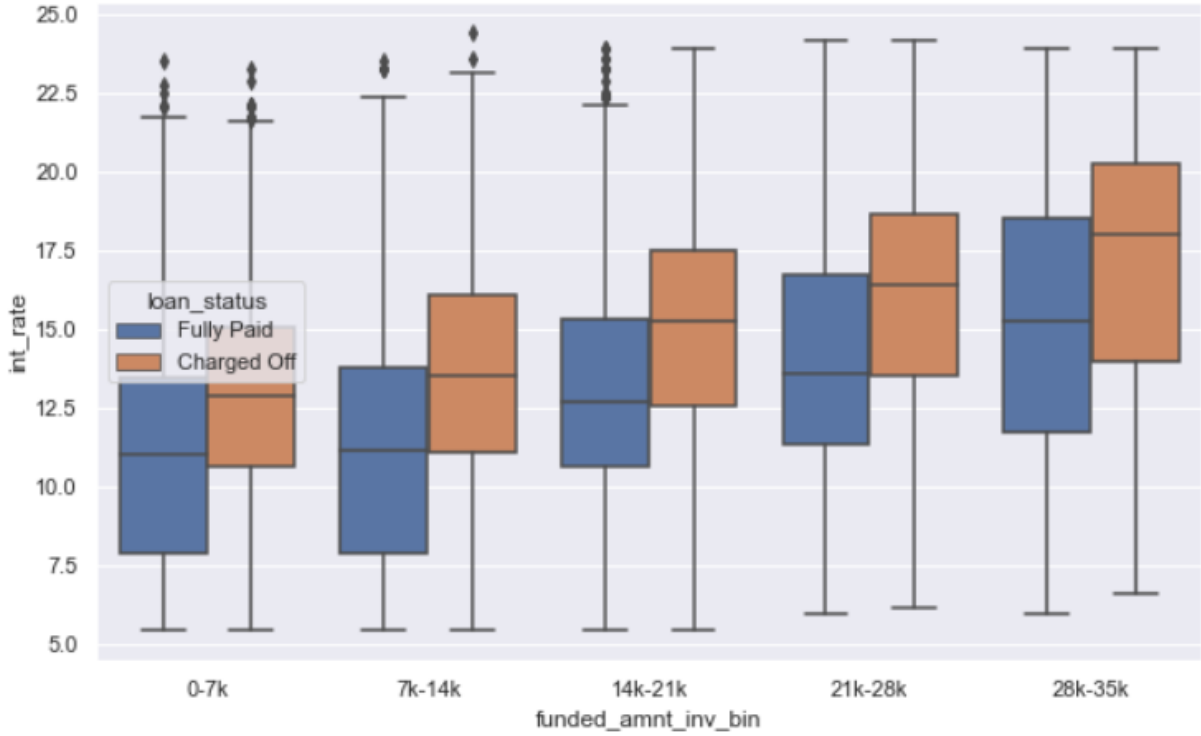
Conclusion : From the above bar plot, we can observe that for **employment length = 1** and borrowers living on "**Rent**" the **defaulter rate is highest**. If borrower own his house, then percentage of default is very less.

d) Analysis of Loan term (Tenure) with dti (Debt to income ratio) and Loan_Status.



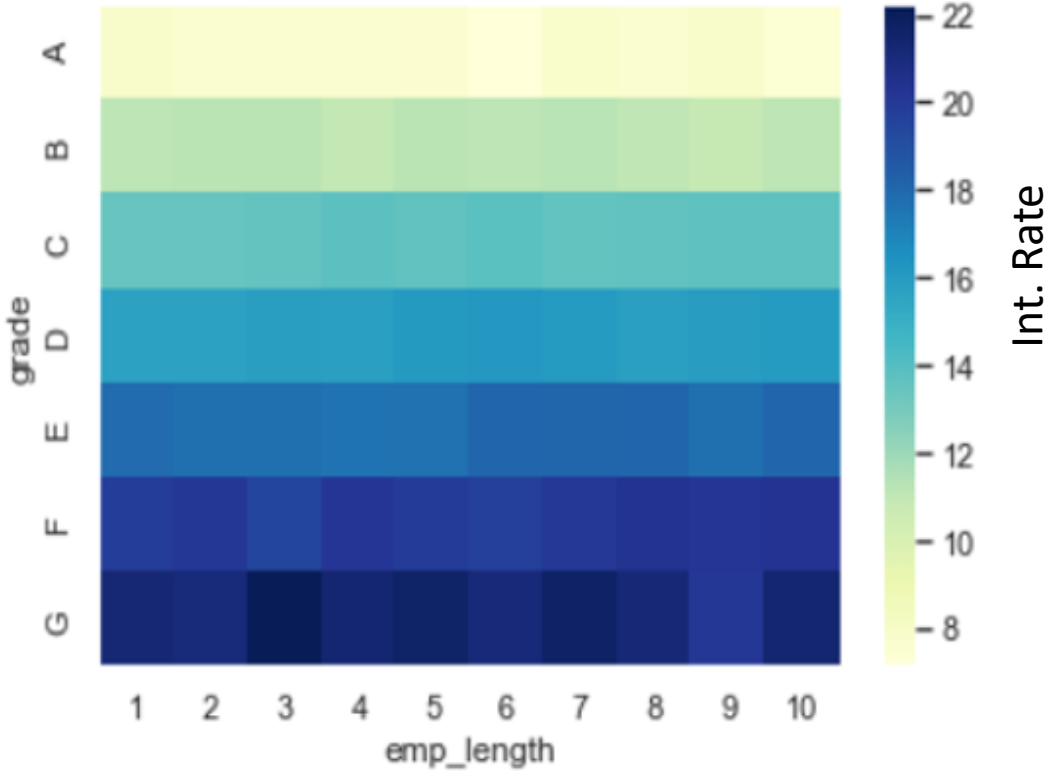
Conclusion：As we observe the trend in the above plot as the term increases with respect to dti there is more chances of becoming default.

e) Analysis on columns Int_Rate and funded_amnt_inv w.r.t Loan_Status



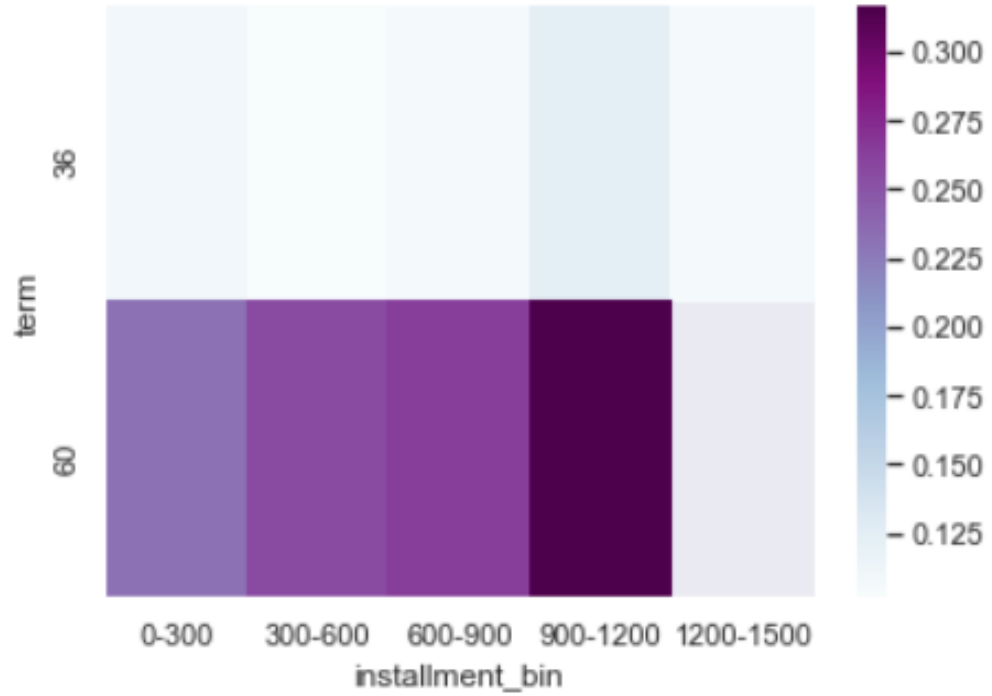
Conclusion：As we observe the trend in the above plot as the Interest rate and Funded amount increases , more chances of becoming default.

f) Analyzing an impact of Employee_Length (Employee exp.) and Grade on Loan Interest rate



Conclusion： From the below heatmap we have analyzed among variables Grade, Interest_Rate and Employee_length, that employee length **3** with grade **'G'** has the highest interest rate. Similarly, we have high rate of interest in 'G' grade for employee length 5,7 and 10.

g) Analysis of Loan term (tenure) with Monthly Installment and Loan status.



Conclusion： As the installment amount increases with number of terms increase(around 60) , there are more chances of becoming defaulter.

Conclusion:

- Factor “**Purpose**” has highest defaulters for “**debt_consolidation**” reason, loan approval for “**debt_consolidation**” should reduce.
- States with state_code CA, FL and NY has more cases of defaults ,lending Club should control number of loan approval to borrowers.
- There are a greater number of loans for employee having 1 and 10/10+ yrs exp. Lending club should avoid providing loans to borrowers having employee length around 1 yr and living on Rent. Loan can be approved for higher employee exp like 10/ 10+ yrs as 10+ yrs has high Fully_Paid rate.
- We should reduce providing loans to borrowers seeking more Loan Term period with high DTI (Debt to income ratio).
- We need to avoid approving loans where the funded amount provided by investor and interest rate is more.
- Reduce approving loans to borrowers with employee length 3, 5 ,7 and 10 having grade 'G’ with high interest rate which might make a loss.
- Avoid approving loans to borrowers taking longer term loan (around 60) and having high monthly installment amount.