Lending Club Case Study

Business Objective

To identify variables which are strong indicators of default and potentially use the insights in approval / rejection decision making.

About Data

Types of variables

- Customer (applicant) demographic
- Loan related information & characteristics
- Customer behaviour (if the loan is granted)

Customer Information:

- 1.emp_title
- 2.emp_length
- 3.home_ownership
- 4.annual_inc
- 5.verification_status
- 6.addr_state
- 7.zip_code
- 8.title
- 9.purpose

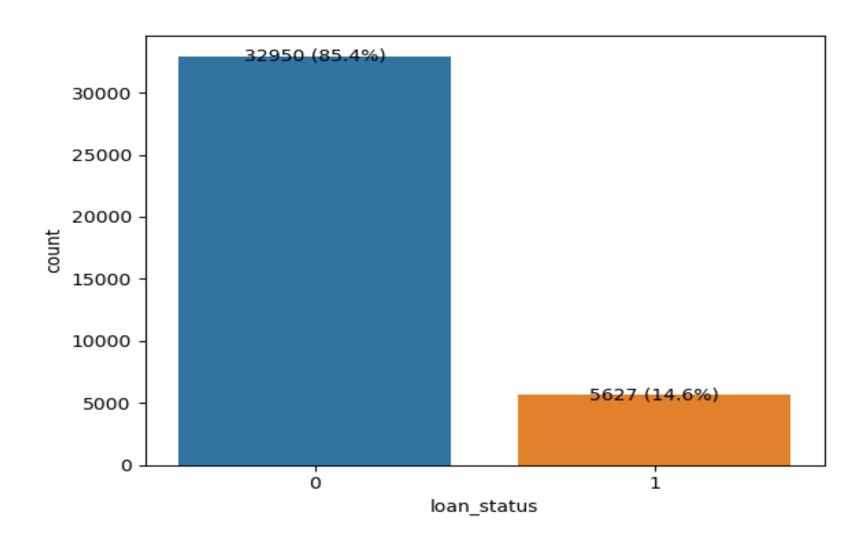
Loan Information:

- 1.funded_amnt
- 2.int_rate
- 3.loan_status
- 4.grade
- 5.sub_grade
- 6.dti
- 7.loan_issue_d
- 8.term
- 9.installment

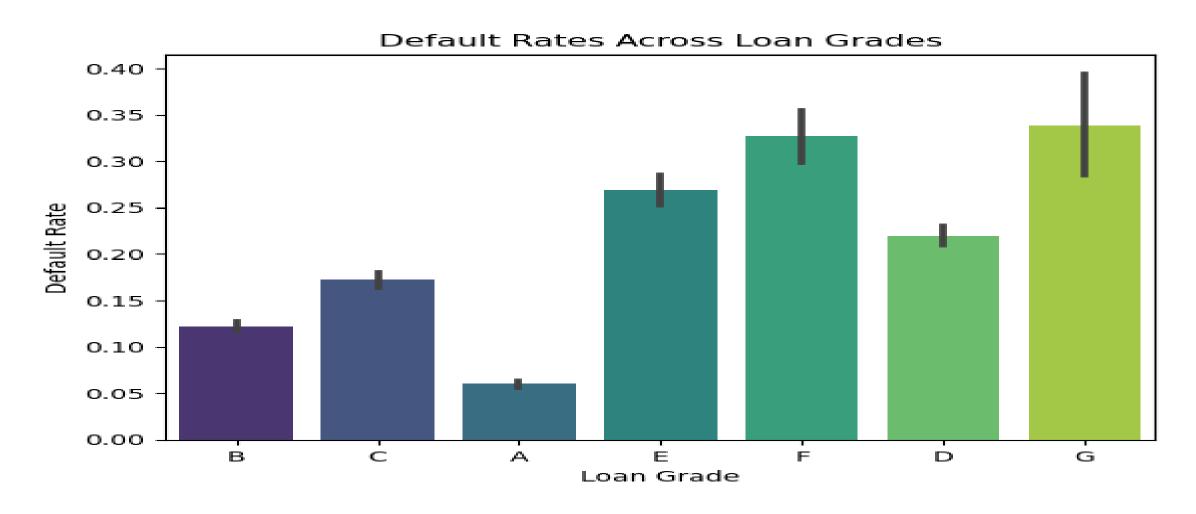
Customer Behavorial Metrics:

- 1.delinq_2yrs
- 2.earliest_cr_line
- 3.inq_last_6mths
- 4.open_acc
- 5.pub_rec
- 6.revol_bal
- 7.revol_util
- 8.total_acc
- 9.out_prncp
- 10.out_prncp_inv

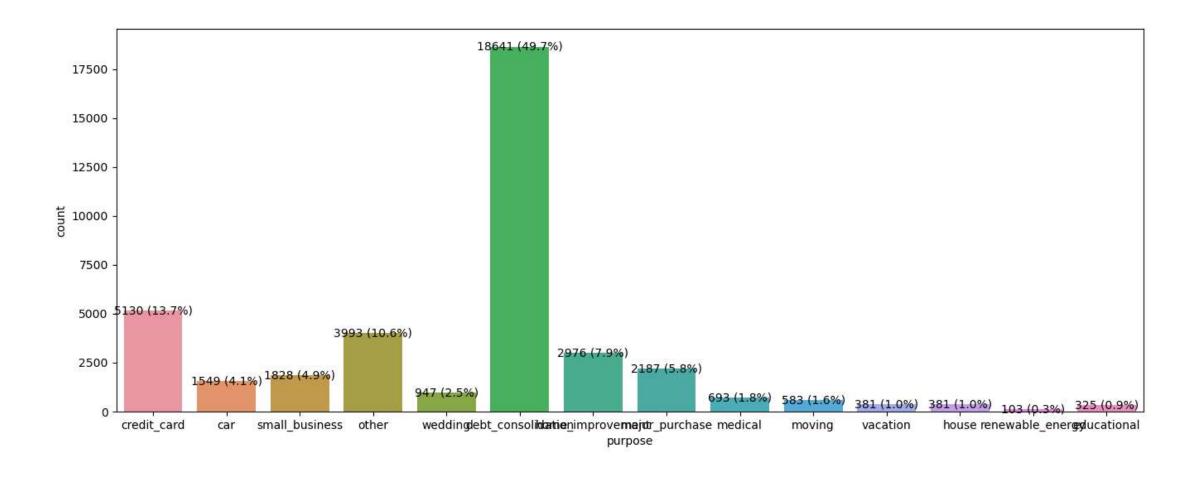
Overall Default Rate is 14%



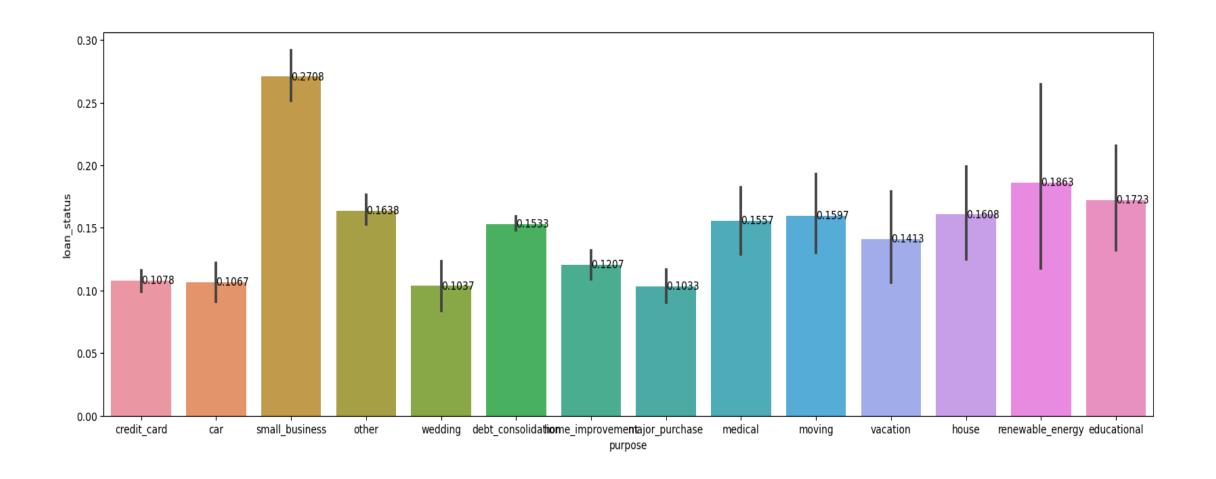
Most of the peoples are interest to take E,F,G grade loans



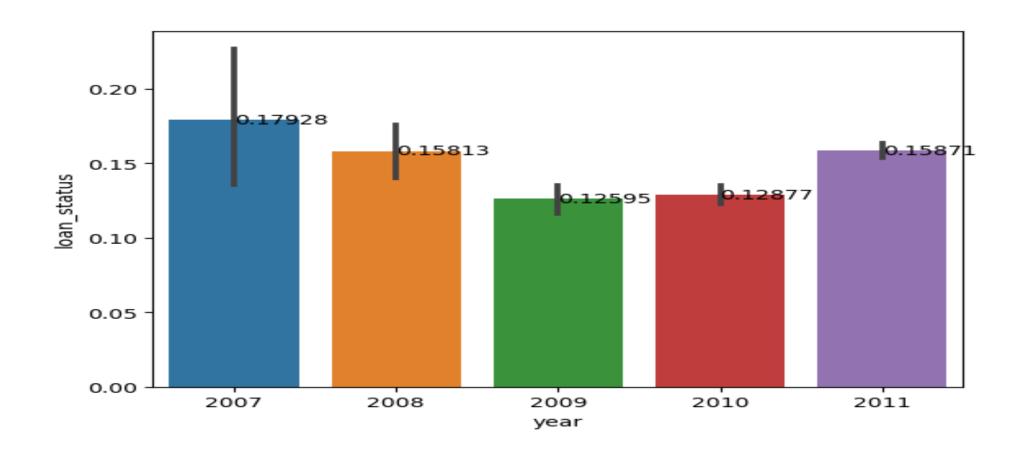
lets first look at the number of loans for each type (purpose) of the loan Most loans are debt consolidation (to repay otehr debts), then credit card, major purchase etc.



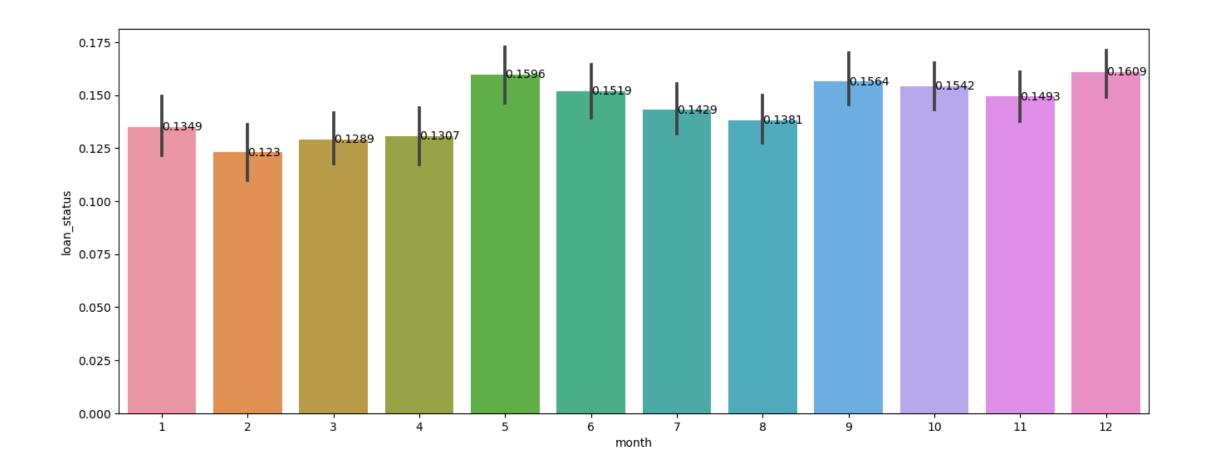
purpose: small business loans defualt the most, then renewable energy and education



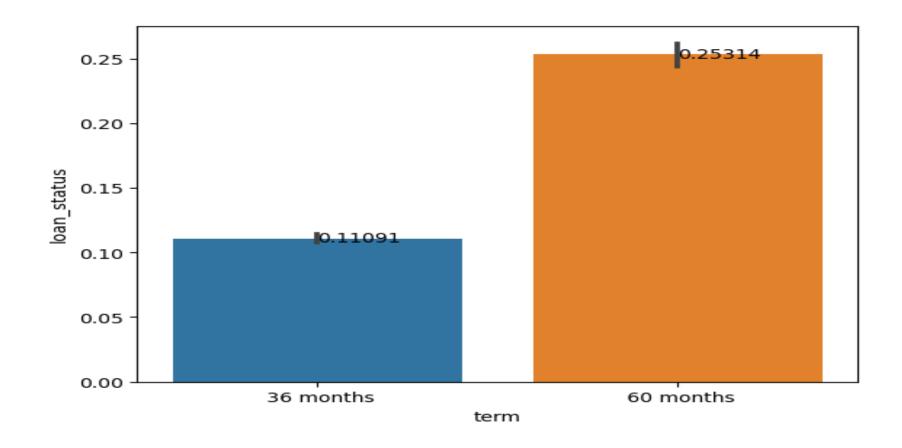
the default rate had suddenly increased in 2011, inspite of reducing from 2008 till 2010



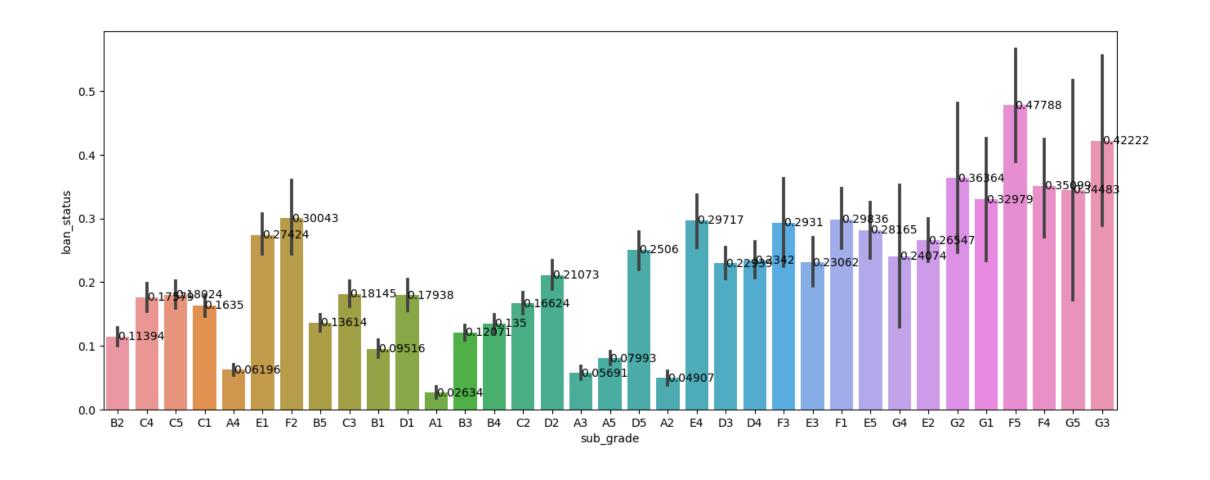
comparing default rates across months: not much variation across months



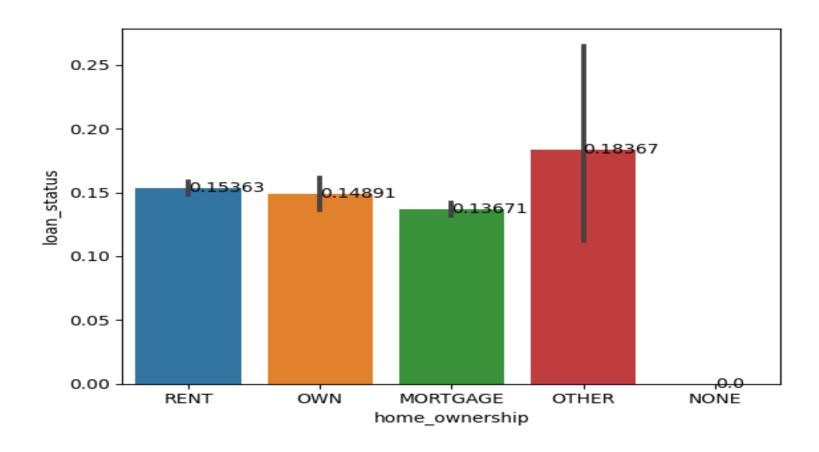
Most of the loan defaulted when they are in 60 months terms



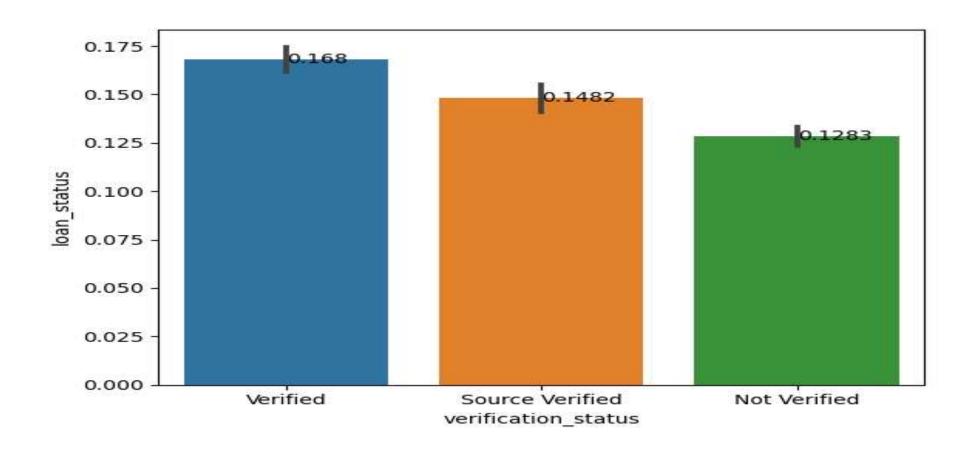
sub-grade: as expected - A1 is better than A2 better than A3 and so on



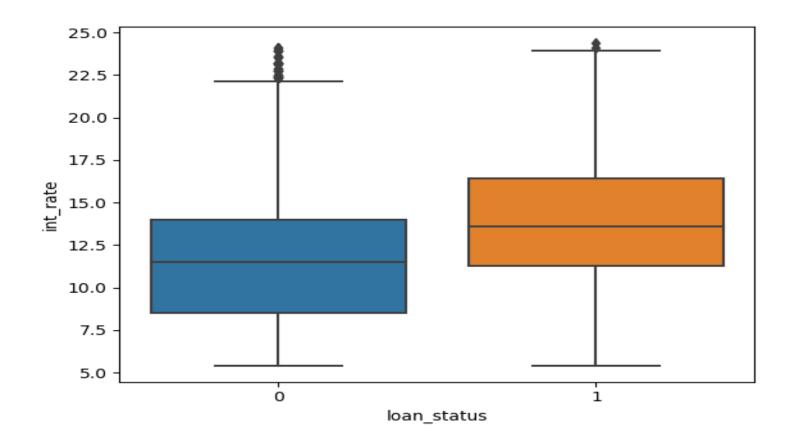
home ownership: not a great discriminator why because there is no distribution b/w home_ownwership



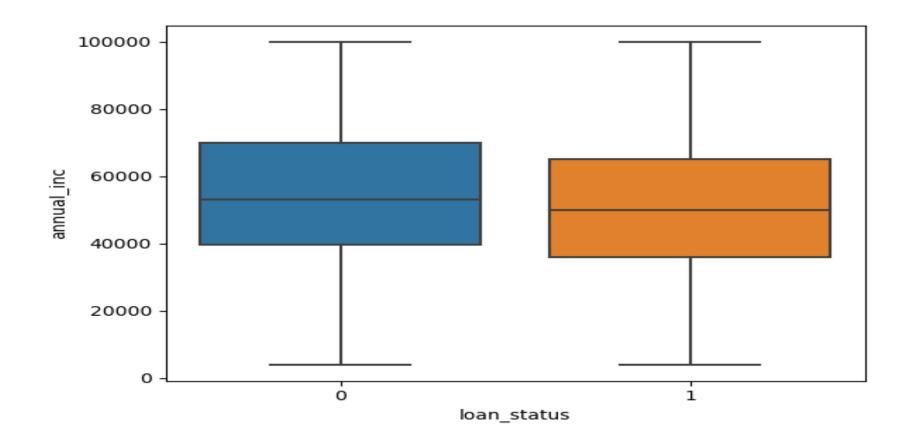
verification_status: surprisingly, verified loans default more than not verifiedb

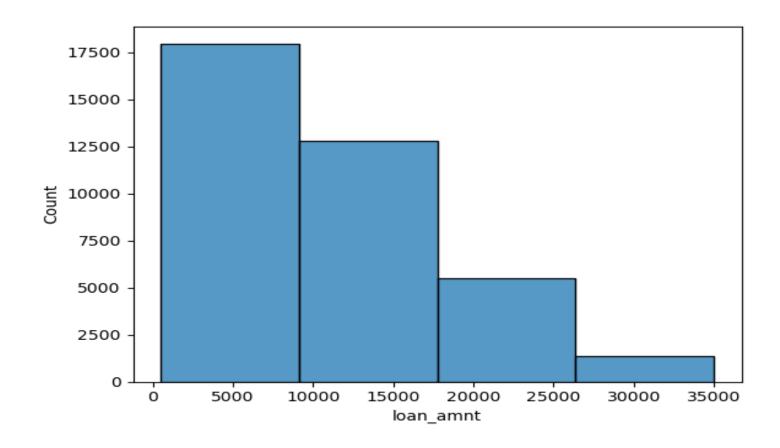


customers mostly defaulted which has higher interest rate for loans

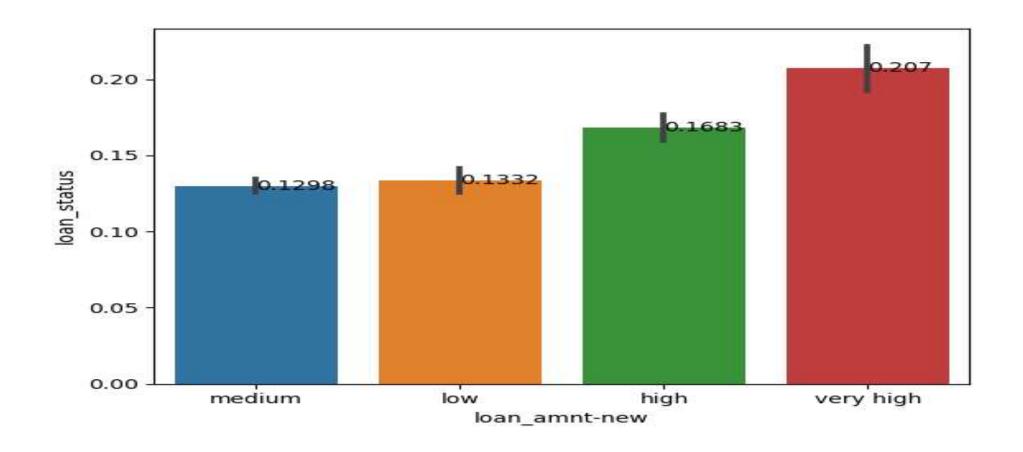


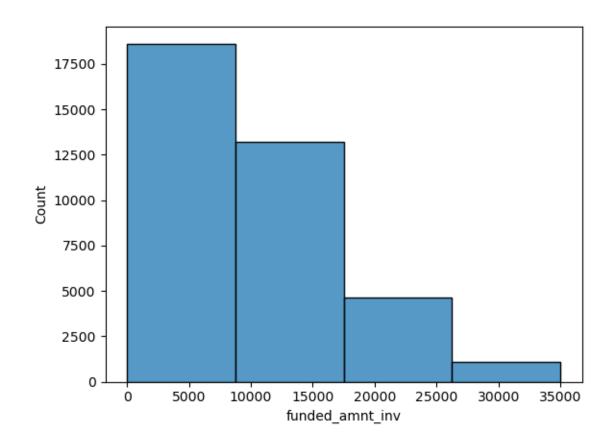
Defaulters salary is less than not defaulters salary



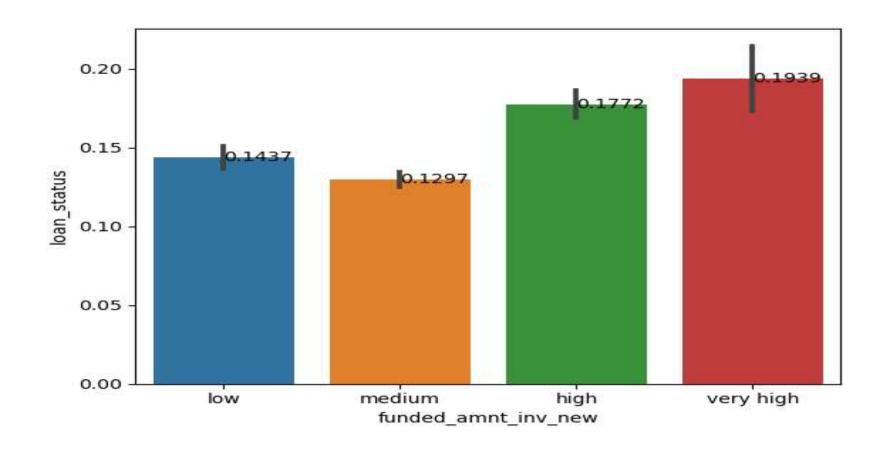


most of the customers defaulted when thay have higher amount of loan

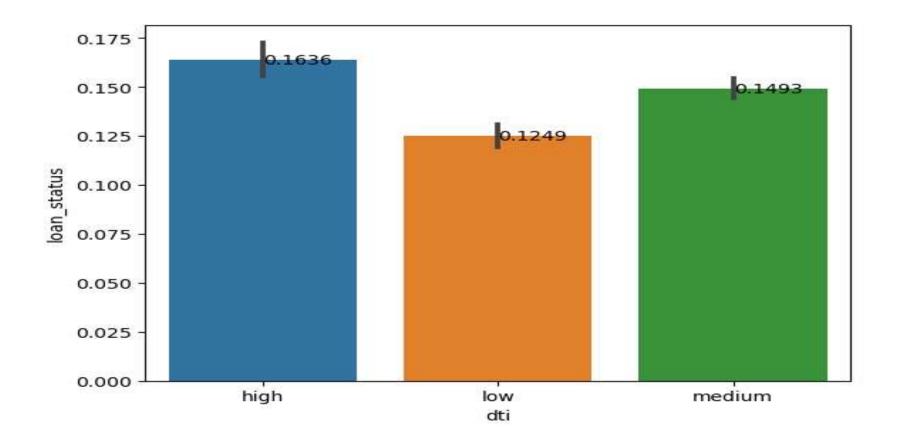




the customers defaulted when thay hav higher amount of funded_amt_inv



high dti translates into higher default rates, as expected



the higher the installment amount, the higher the default rate

