

LENDING CLUB CASE STUDY SUBMISSION

Group:

1. Ram Krishn Mishra
2. Sai Raju Kota

Business Objectives

- The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. and With the help of **loan, consumer attributes** and **loan attributes** influence the tendency of default.
- The lending club company can clearly understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default. Lending company can utilize this knowledge for its portfolio and risk assessment.

Problem solving methodology

The data analysis consists four main parts:

- Data understanding
- Data cleaning (cleaning missing values, removing duplicate columns and so)
- Data Analysis
- Recommendations

Data analysis approach

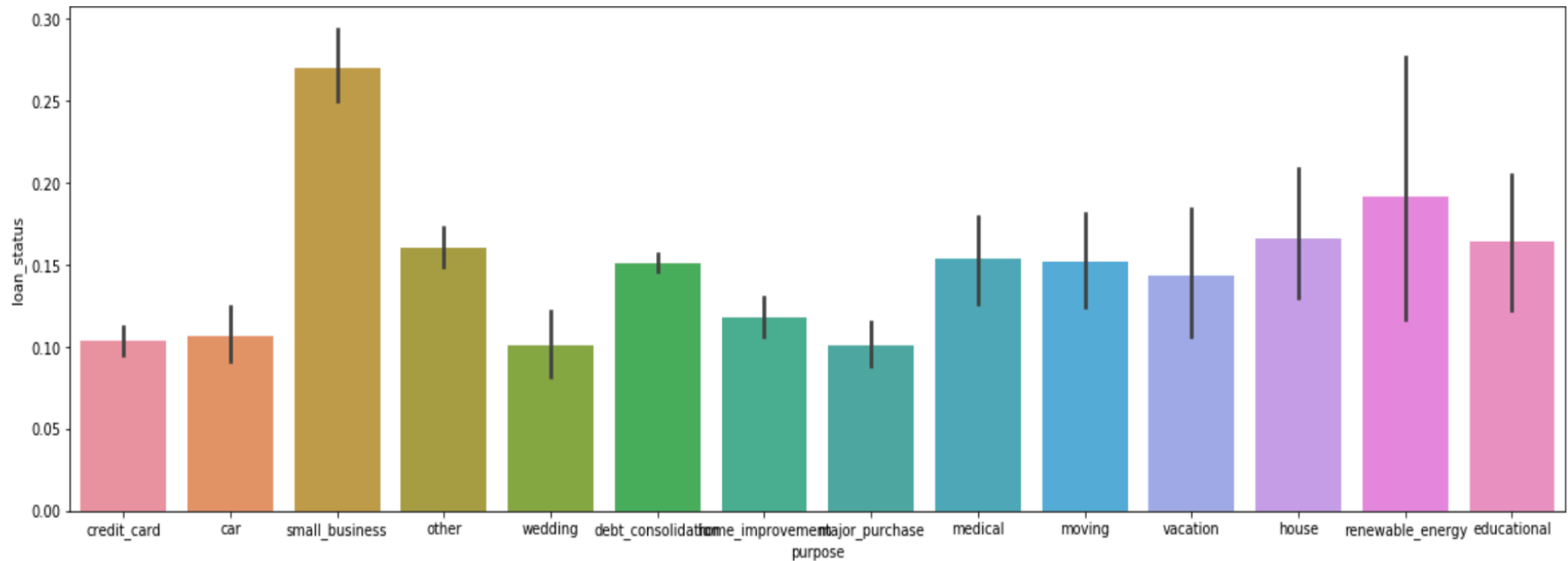
- ✓ Load the CSV file in loan data frame
- ✓ Explore the Dictionary file to understand the columns in the loan csv and its significance
- ✓ Remove all the columns which has more than 90 % missing values.
- ✓ Fix the data discrepancy such as data type, blanks and duplicates.
- ✓ Create bins/categories for loan characteristics'.
- ✓ Create derived columns from Issue_date to check year and month wise.
- ✓ Concluded column required for analysis.
- ✓ Plot the Univariate analysis and observe.
- ✓ Plot out Multivariate analysis and observe.
- ✓ Do Recommendation's

Behavior variables by type

- Consumer – (Annual income, Employee Length, Employee Title)
- Loan information & characteristics - (Loan Amount, Loan Status, Loan Grade/ Sub Grade, Amount investment, Interest rate)
- Customer behavior – (delinquency year-2, earliest, purpose, recoveries, earliest cred line, application type, revolving balance)

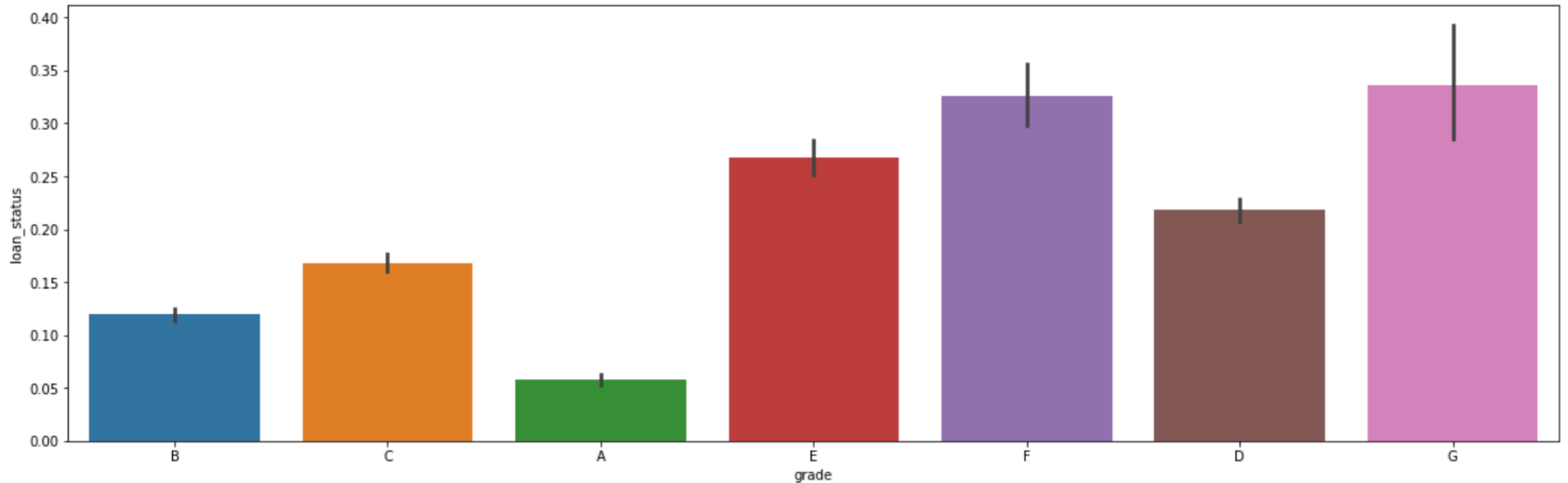
Univariate Analysis (Purpose vs Loan Status)

- the overall loan default rate is 14 % and non- default rate is 86%



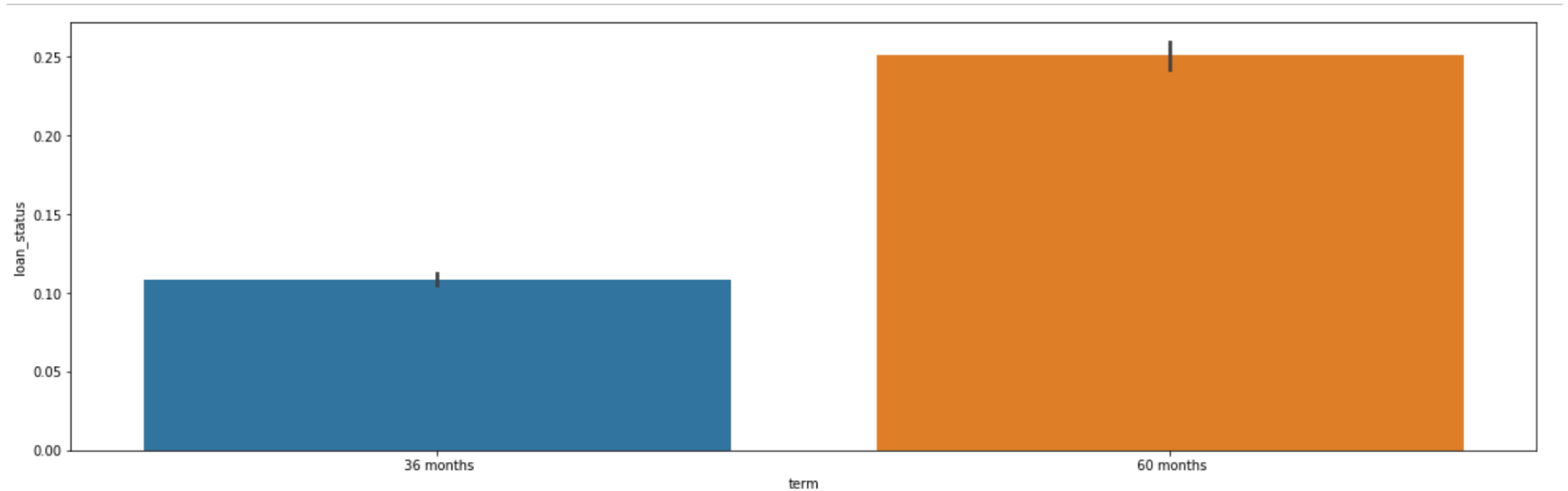
Observation: small business loans default the most, then renewable energy, education and houses

Grade and loan status



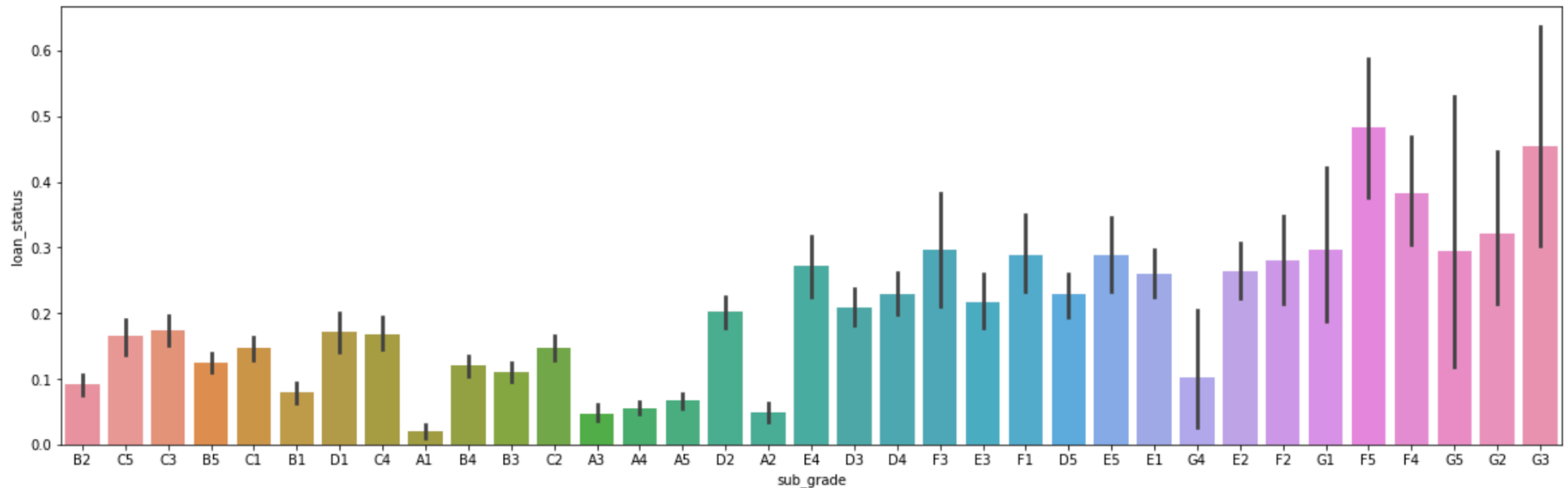
Observation: from A to G, the default rate increases which is expected because the grade is decided by Lending Club based on the risk undertaking of the loan.

Term and loan status



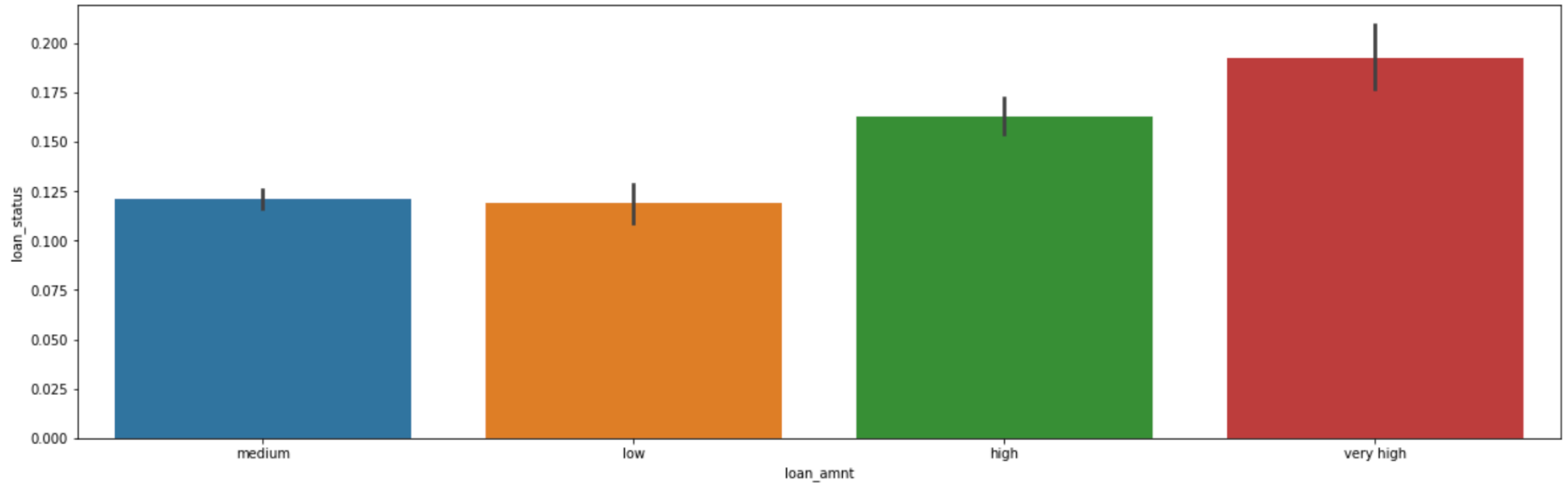
Observation : 60 months loans default more than 36 months loans. So consumers preferring 60 months term.

sub-grade and loan status



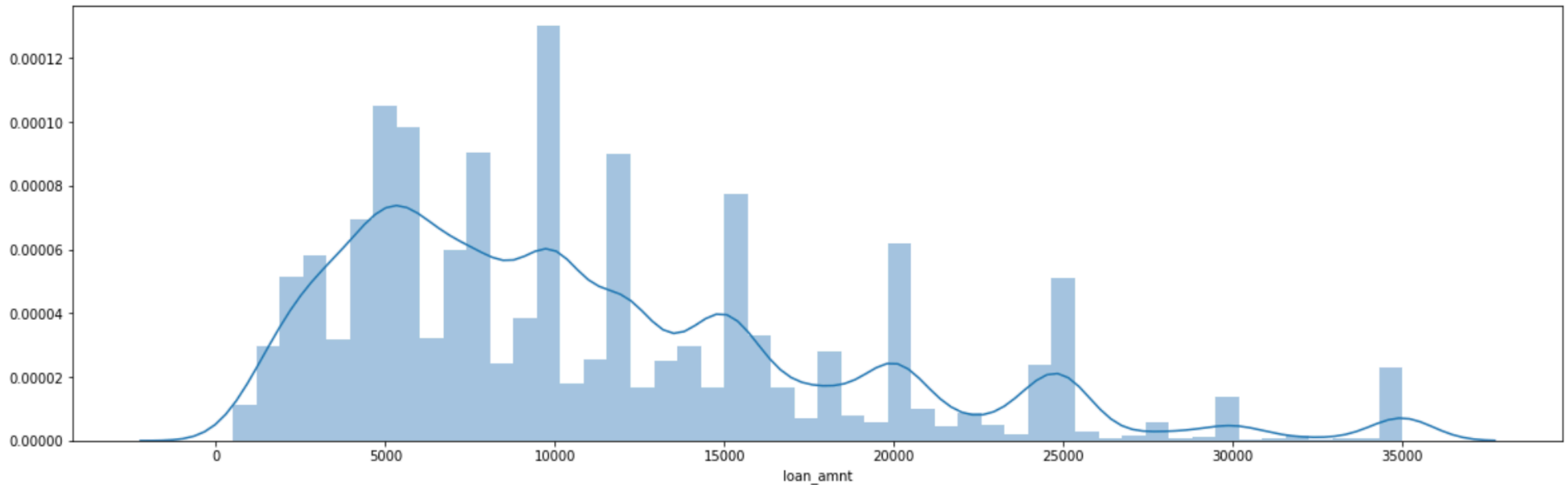
Observation: A1 is better than A2 better than A3 and so on in ascending manner .It got varied in F4 & F5 which highest and more than G5 & G3 and Median is F5 sub-grade.

funded amount invested and loan status



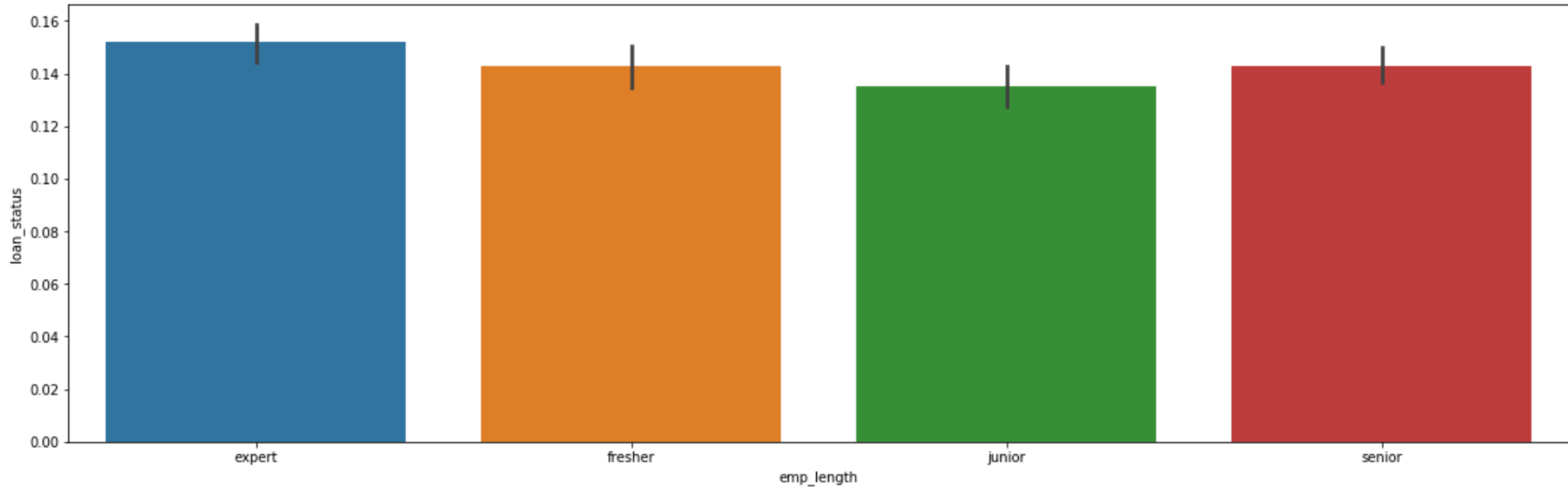
Observation : higher the loan amount, higher the default rate

Loan amount analysis



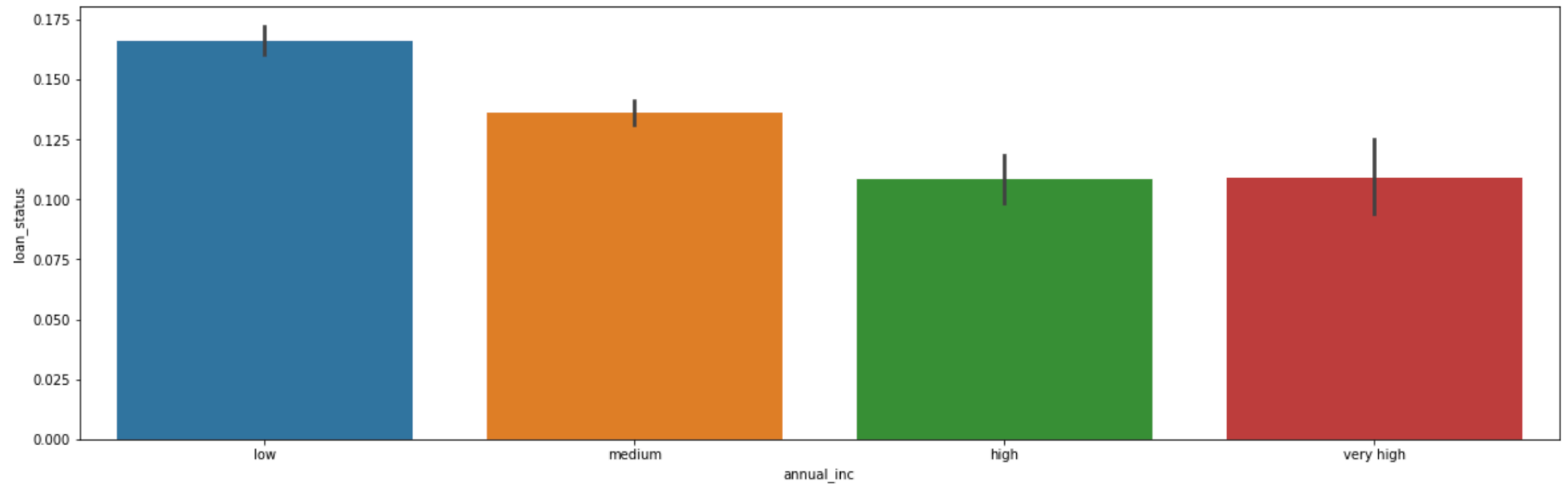
Observation: the median loan amount is around 10,000

Employee length and default rate



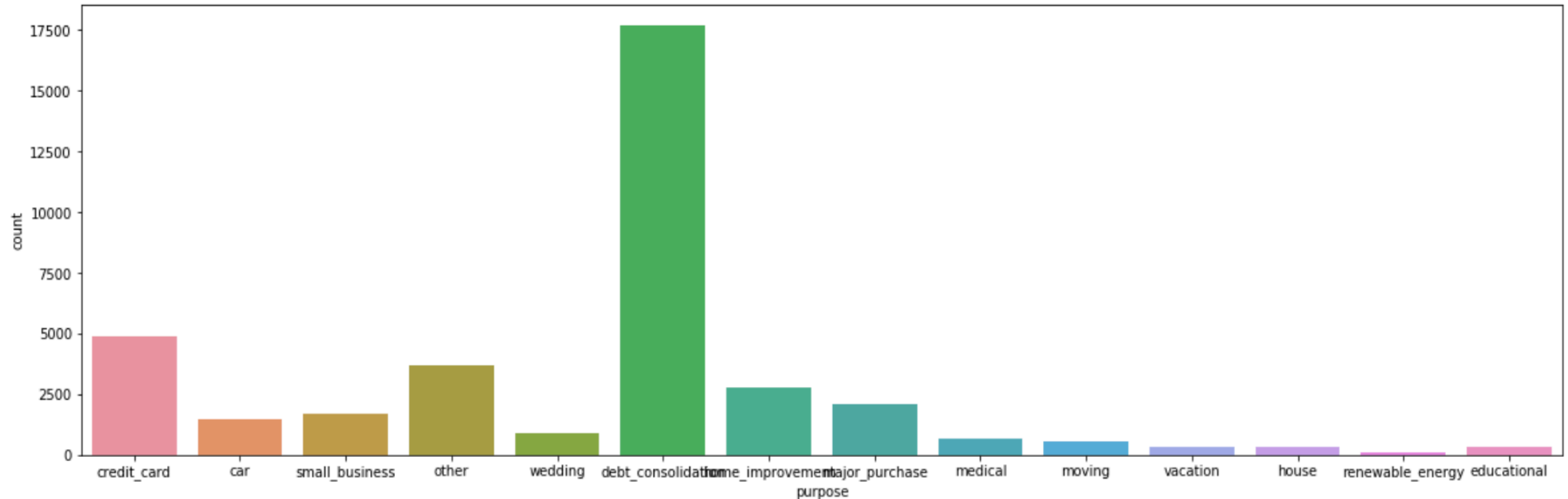
Observation: not much of a predictor of default rate

Annual income and default rate



Observation: lower the annual income, higher the default rate

Purpose and No of type of loans



Observation: the top 4 types of loans based on purpose: consolidation, credit card, home improvement and major purchase.

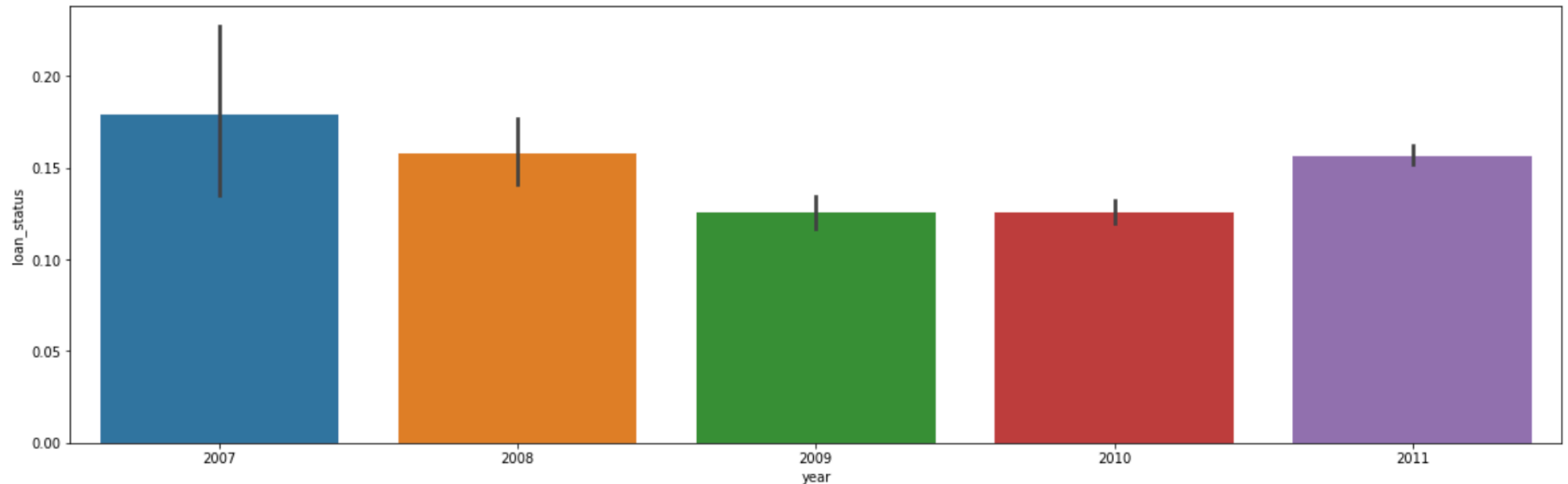
No of debt_consolidation - 17675

No of credit_card - 4899

No of home_improvement – 2785

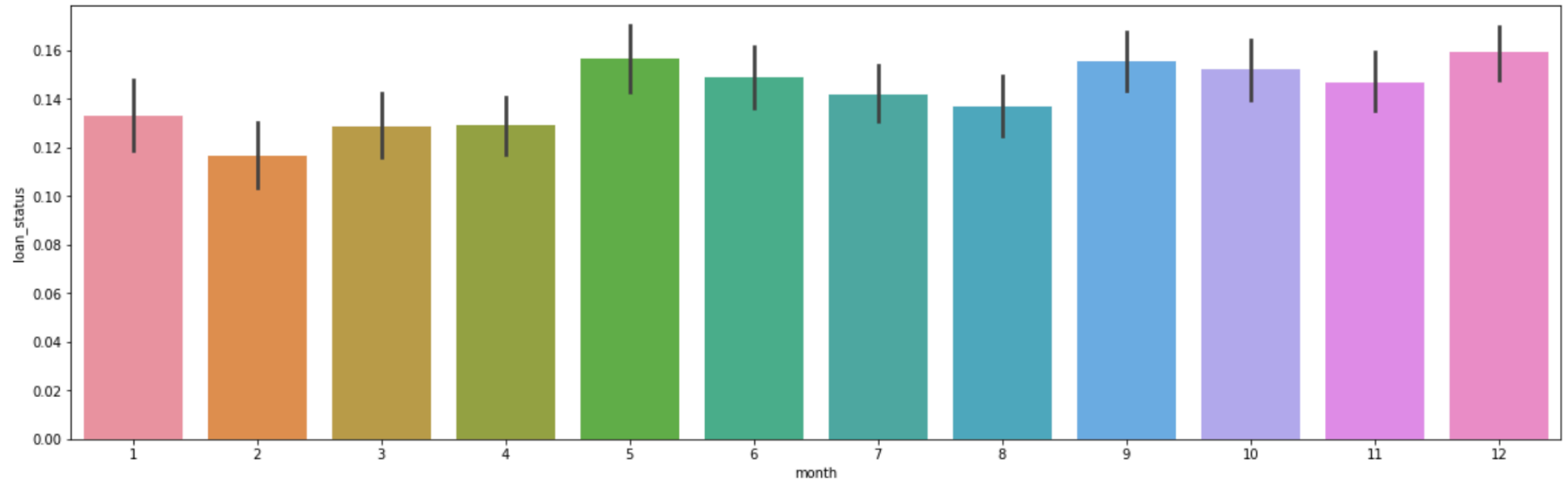
No of major_purchase - 2080

default rates and years



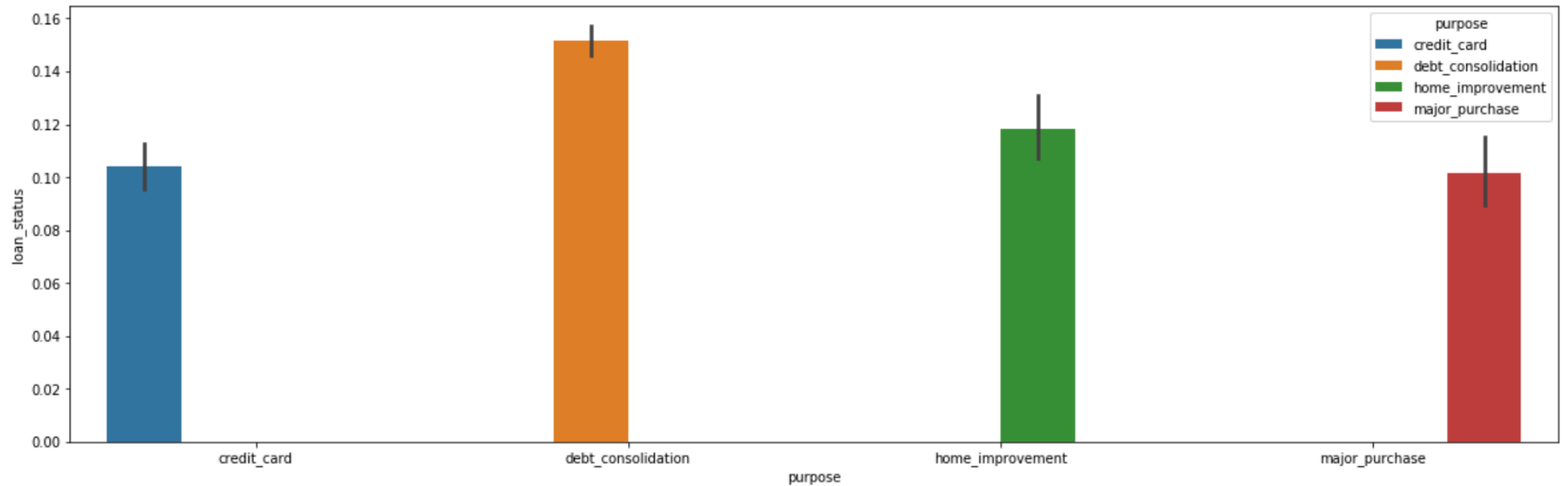
Observation: the default rate had suddenly increased in 2011, in spite of reducing from 2008 till 2010. and we can see that the number of loans has increased steadily across years.

default rates and years

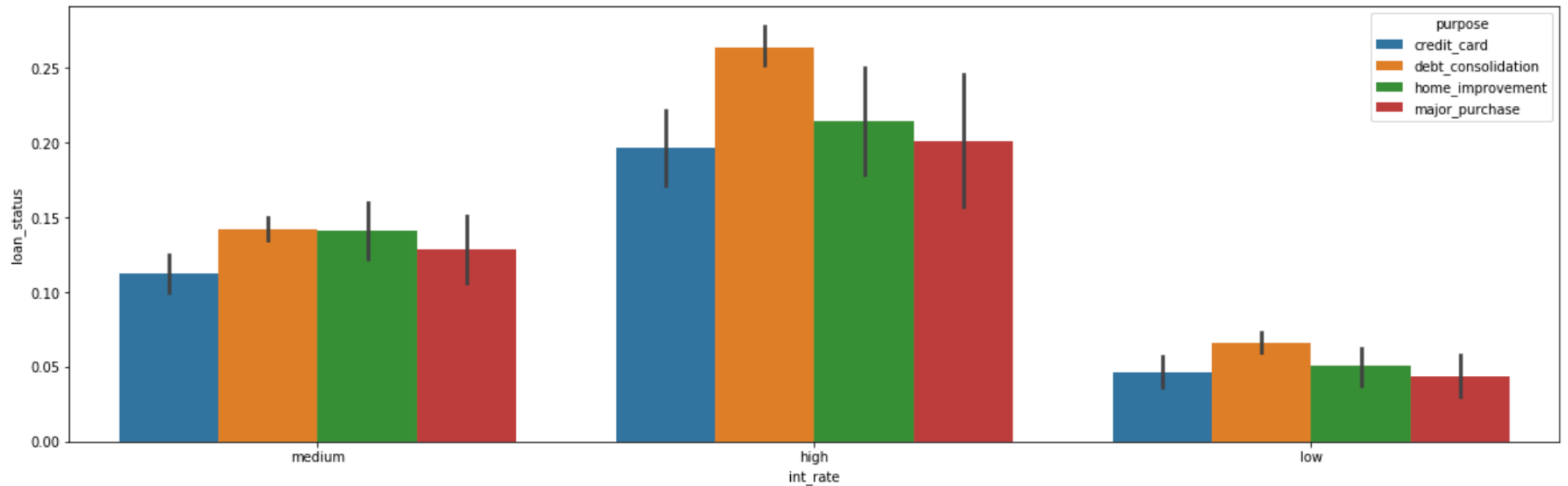


Observation: Most loans are granted in December, and in general in the latter half of the year.

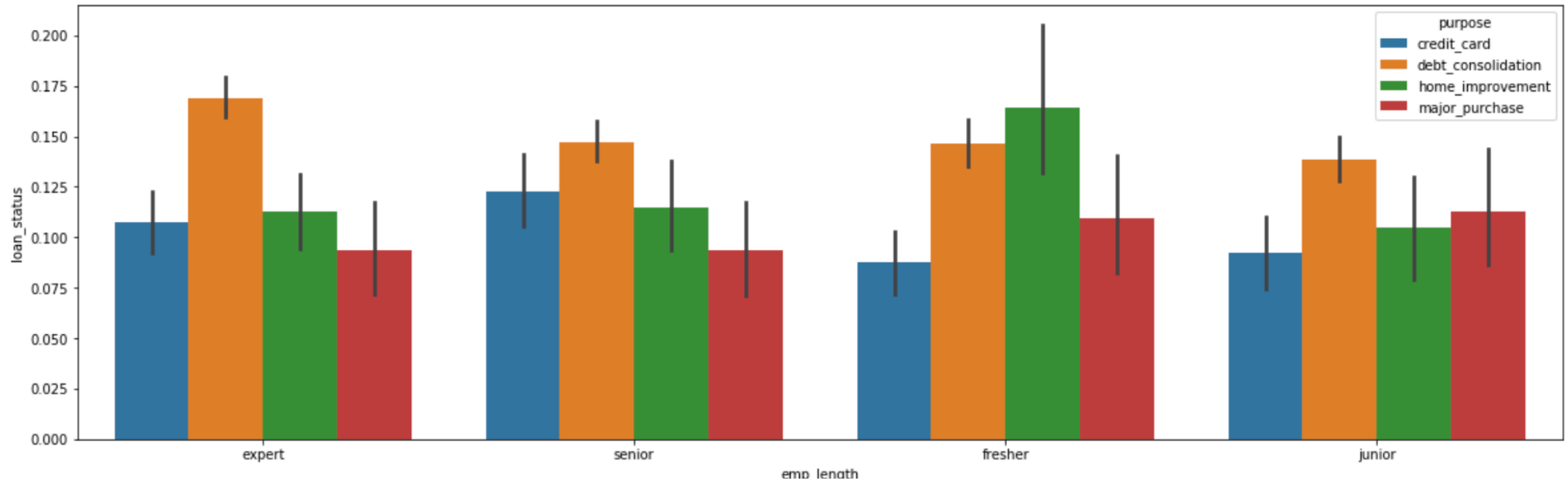
Multi variate analysis: status and purpose



interest rate and loan status



Employee length(service) and loan status



Conclusions and Recommendations

- For better results to reject the loan application which has high risk of defaulting the loan for achieving the same following attributes are recommended to use while processing an loan application
 - *Attributes:*
 - **Purpose:** Small business loans default the most, then renewable energy, education and houses and tell that purpose contributes to the charged off.
 - **Sub-grade/grade:** The analysis tells that increasing rates increases the default rate
 - **Emp_length** (service of the applicant): we have observed that consumers who has more 10 years experienced loans are charged off most.
 - **Varification status** : As analysis of verified category resulted more defaults than other categories.
 - **Term:** Analysis tells that consumers preferring 60 months term than 36 months