LENDING CLUB ASSIGNMENT

SUBMISSION

Name:

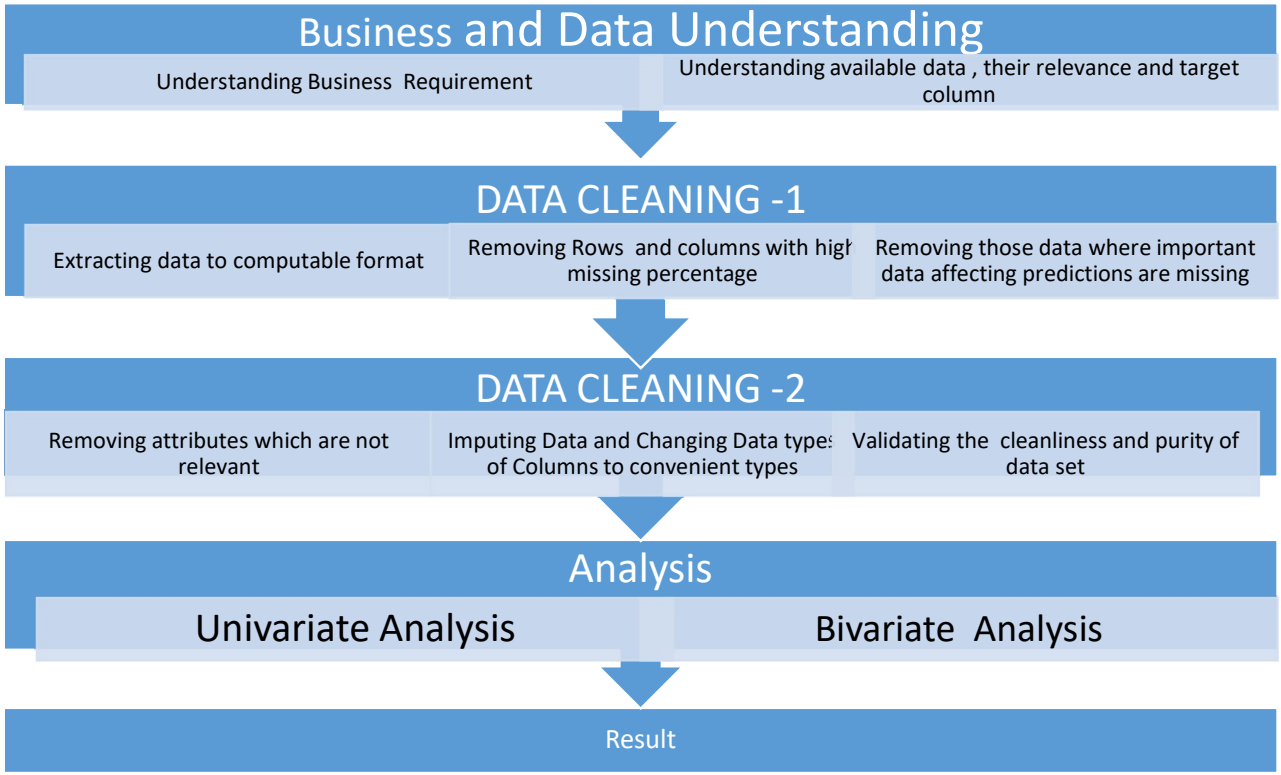
1. Murali Divya Teja Gummadidala
2. Alex Cherian Paul

Abstract

Lending Club 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. Lending loans to 'risky' applicants is the largest source of financial loss (called credit loss).

The given data set contains the information about past loan applicants and whether they 'defaulted' or not So our **objective** 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.

<Problem solving methodology>



Analysis:

For neat and good analysis we need to understand the data .The data can be divided into three categories

1. Customer Information (Information related to his financial status , employment , assets etc.)
2. Loan Characteristics (Type of Loan , Reason for loan ,Interest Rate, Loan Status)
3. Customer-Loan Behaviour (next payment details, payment mode etc.)

The Customer – Loan Behaviour Data can be discarded along with other missing attributes as it wont be available at the time of disposing the loan or they wont contribute to our scope of study. The target variable for our analysis is Loan Status .Those attributes with loan status as current are also discarded as we don't know the future of these loans so we are discarding entries with loan payment status as current. The target variable have to 2 categorical data and it is changed to numerical (0 | 1) .

We are analysing the target variables with various variables like Term, Interest Rate, Grades, Subgrades, Reason for loan etc.

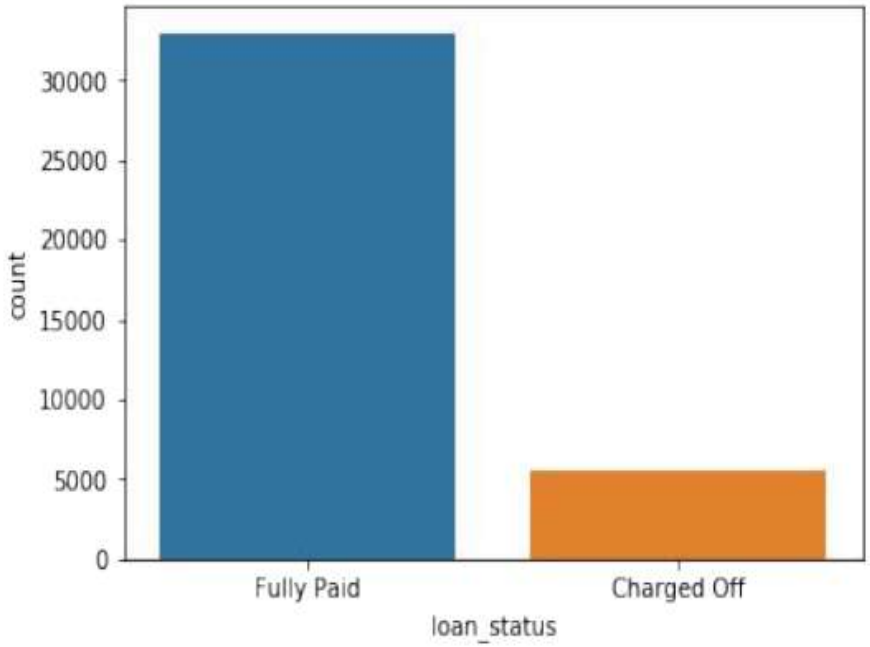
Data Clean Up Steps

- Delete columns: Delete unnecessary columns.
- Missing values: Treat missing values with appropriate approach.
- Duplicate data: Remove identical rows, remove rows where some columns are identical.
- Binning-columns: Bin few columns which are very much useful for Analysis.
- Filter rows: Filter by segment, filter by date period to get only the rows relevant to the analysis.

Loan Status Analysis

Observations :

1. Most of the loans are Fully Paid.
2. bout 14% of loan are having status as defaulters.



Analysis for Grade and Default Rate

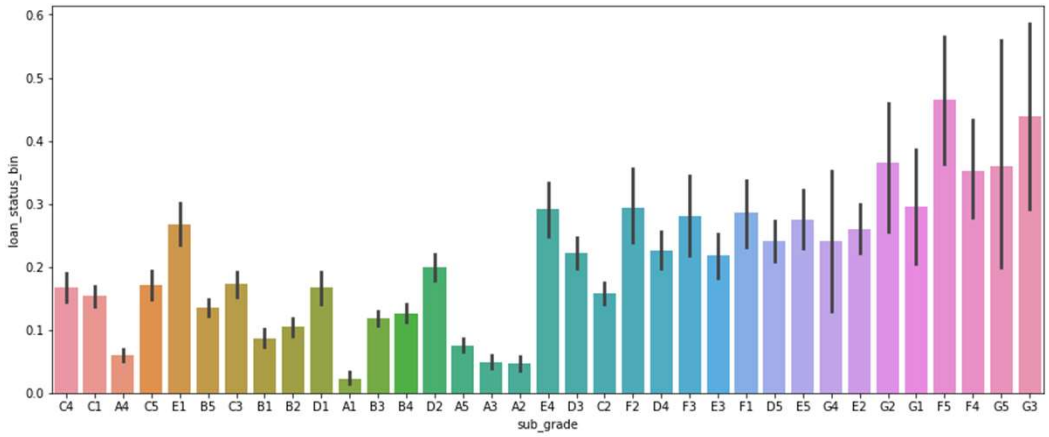
Major Observations

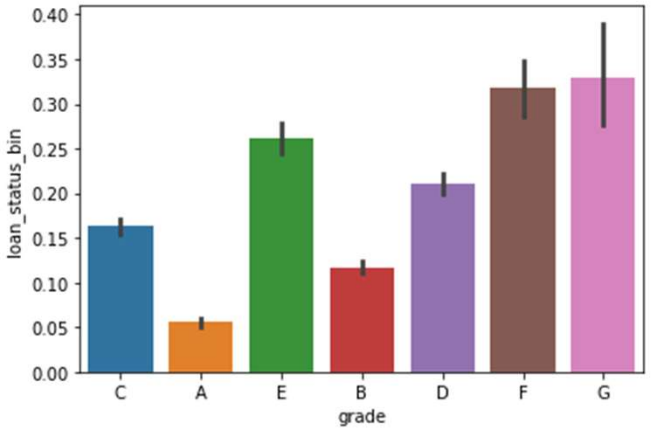
1. As the grade changes from **A to G, the default rate is increased significantly.**

   It can be observed that less than 5% with A grade defaults where as more than 35% people in G grade defaults their loan.

2. We can recommend that as the grade of loan changes there more likely defaulted.

Note: Loan_status_bin is the Default Rate.

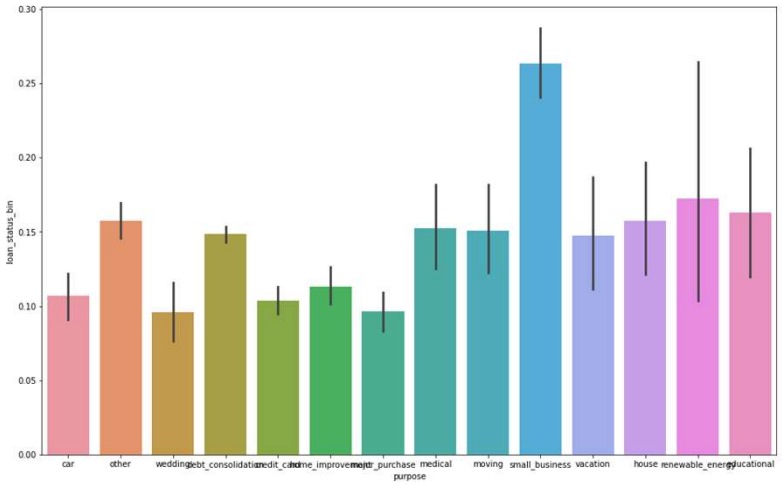


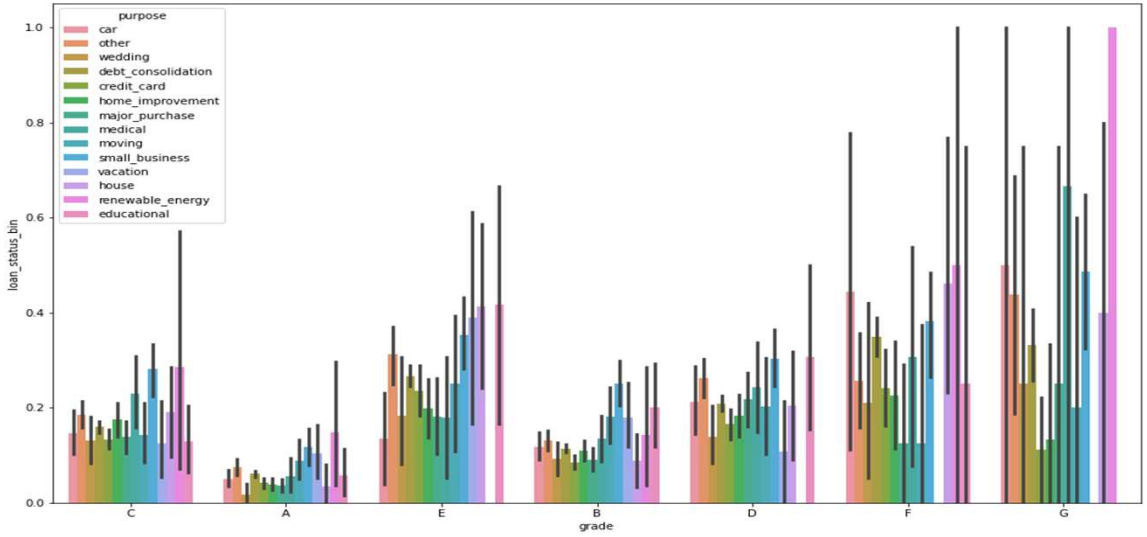


Analysis for Purpose and Default Ratio

Major Observations

1. Bivariate Analysis on Grade and purpose .We can observe that loans for “small_business” (Mean Default Rate>33%)seems more risky and its risk increases from A->G, while loans for “credit_card”, “home_improvement”,” major_purchase”,” vacation” are less risky (Mean Default Rate<=17%).
2. We can recommend that loan for small business is more likely to be defaulted more than any thing.

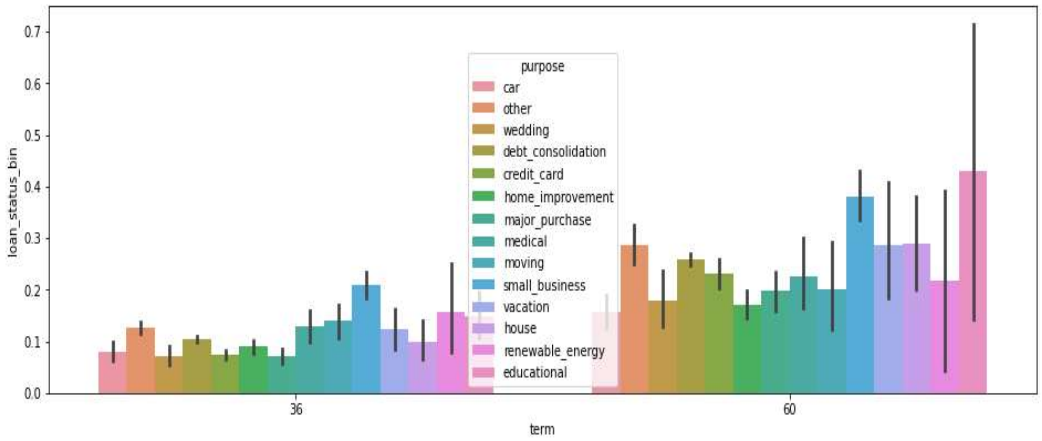


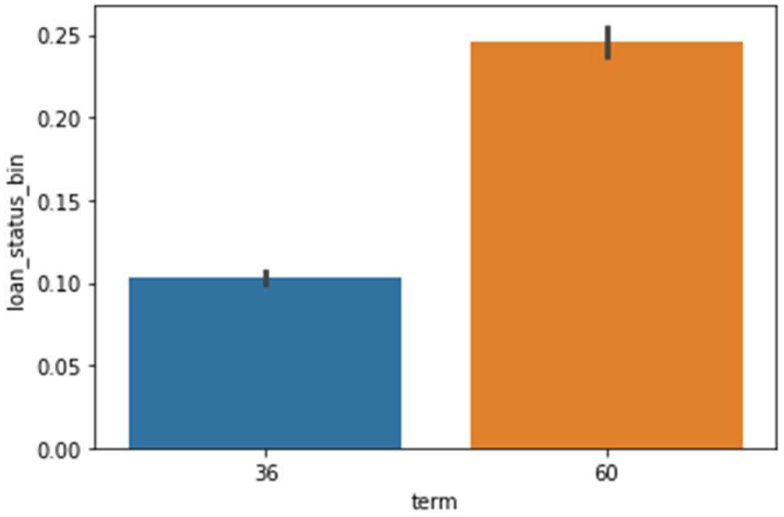


Analysis for Term and Default ratio

Major Observations:

1. On Uni-variate analysis between loan period and default rate. Loans with longer duration tends to have larger Default rate.
2. We can recommend that loan with long period may likely defaulted.

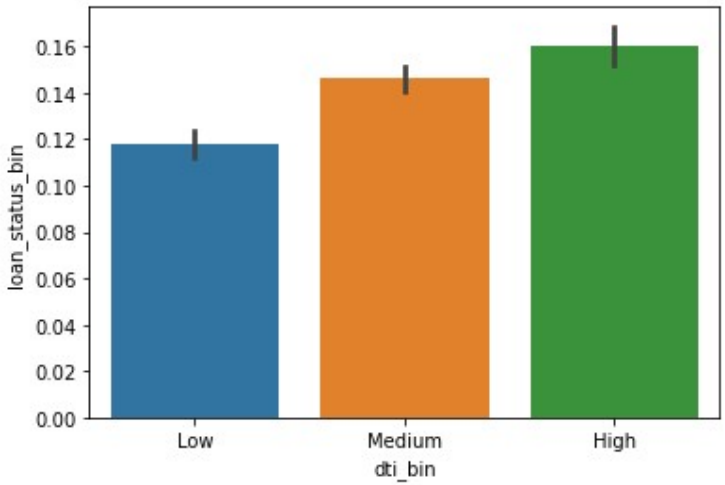


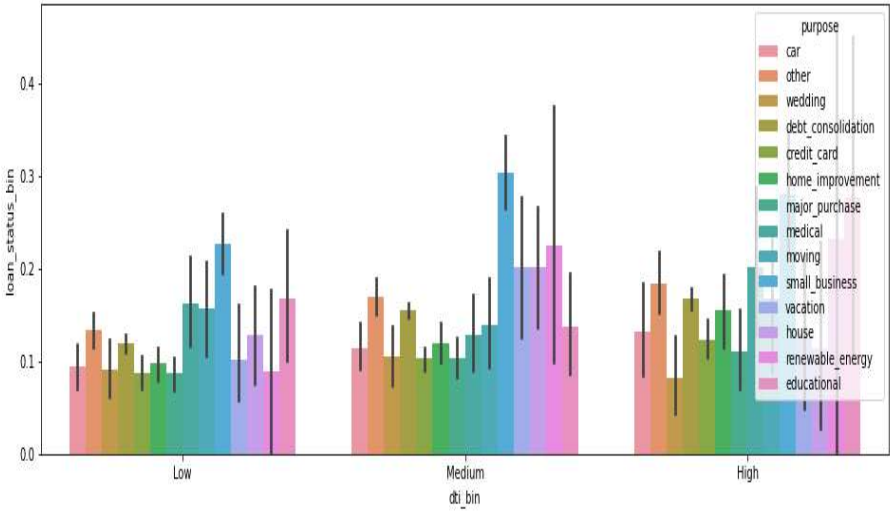


Analysis For DTI and Default ration

Major Observations:

- From the Graph we can say that as Debt to income ratio is increased the default ratio is also more.
- We can recommend that the who has highest DTI ratio is more likely defaulted.

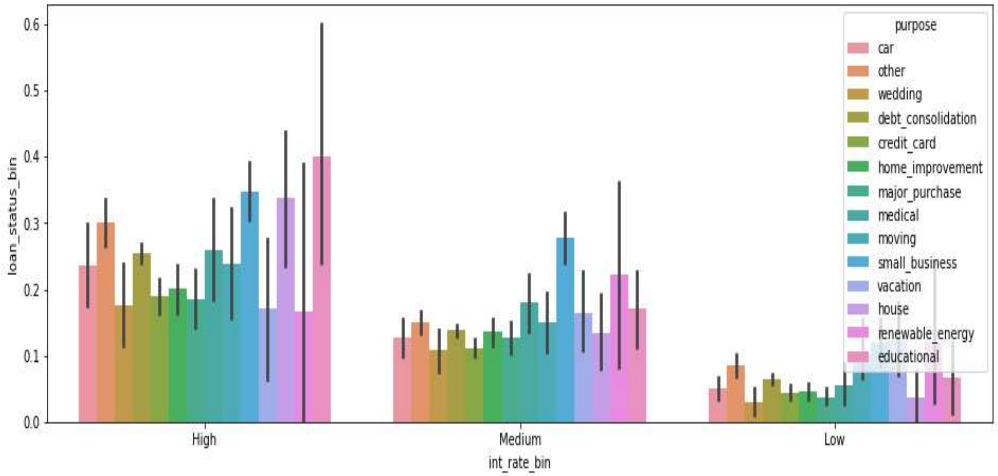


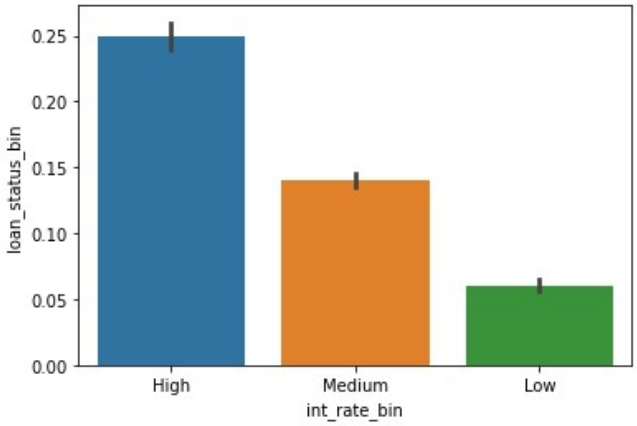


Analysis for Interest Rate vs Default Ratio

Major Observations:

- As the rate of interest getting decreased the default ratio also getting decreased we can confidently say from graph.
- We can recommend that charging more interest rate may more likely defaulted.

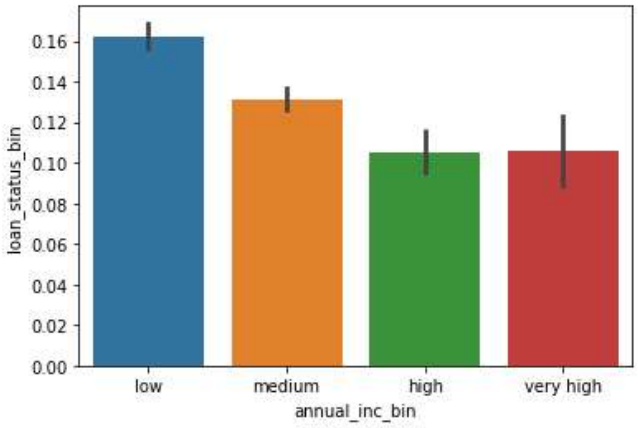


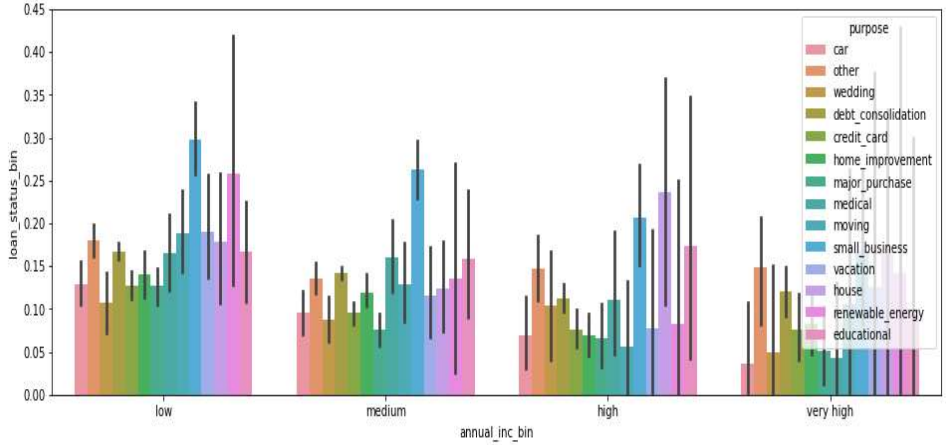


Analysis for Annual Income Vs Default Ratio

Major Observations:

From below graphs we can say that the lower income will have highest default rate than any other.

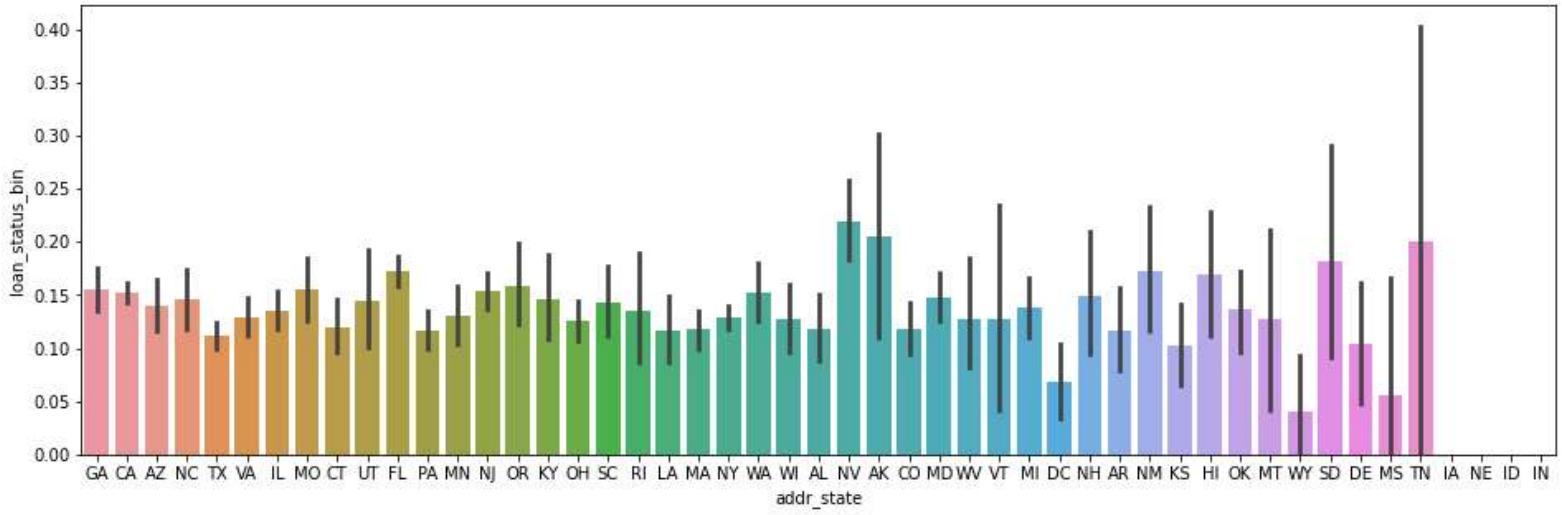




Analysis for state address and default ratio

Observations:

From below graph We can recommend that the state NV is having default ratio.



Conclusion:

Results:

1. High grade loans have high tendency to default. Grading system is working as expected.

2. Loans having higher interest rate have more defaulters. Check the background of applicant thoroughly if interest rate is high.

3. Check the term loan borrowing as the term is more like to be defaulted

4. When the purpose is debt consolidation check applicant thoroughly as it has high tendency to default.

5. Check Annual Income of applicant thoroughly as it has high tendency to default.

6. Deriving factors for loan are:

- Annual Income
- Grade of Loan
- Interest rate
- purpose
- Term