LENDING CLUB CASE STUDY

PAVAN ANDAGI

SUDEEP PAUL

OVERVIEW

- Lending Club is a P2P lending company in operation since 2007
- They work as intermediary between a investor (loaner) and prospective loanee
- Lending Club as intermediary offers ease of access to a individual to seek funds while at the same time offering a loner tantalising investment opportunity
- Lending Club on its part undertakes verification of a prospective loanee and categorises the loanee into grades
- Since, the money is being loaned by an investor and not by Lending Club, the risk of default on part of loanee is borne by the investor

AIM

- Identify patterns in data which indicate if a potential loanee is likely to default, identification of which will enable the investor to apply suitable risk avoidance/ mitigation strategy

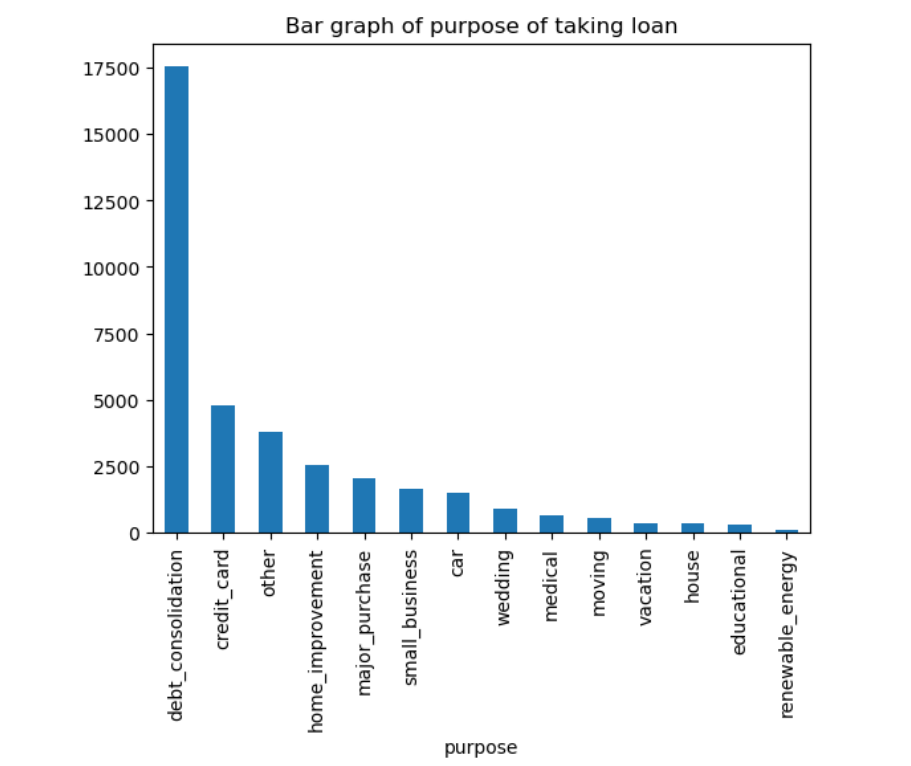
OBJECTIVE

- Based on data made available, 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 utilize this knowledge for its portfolio and risk assessment.

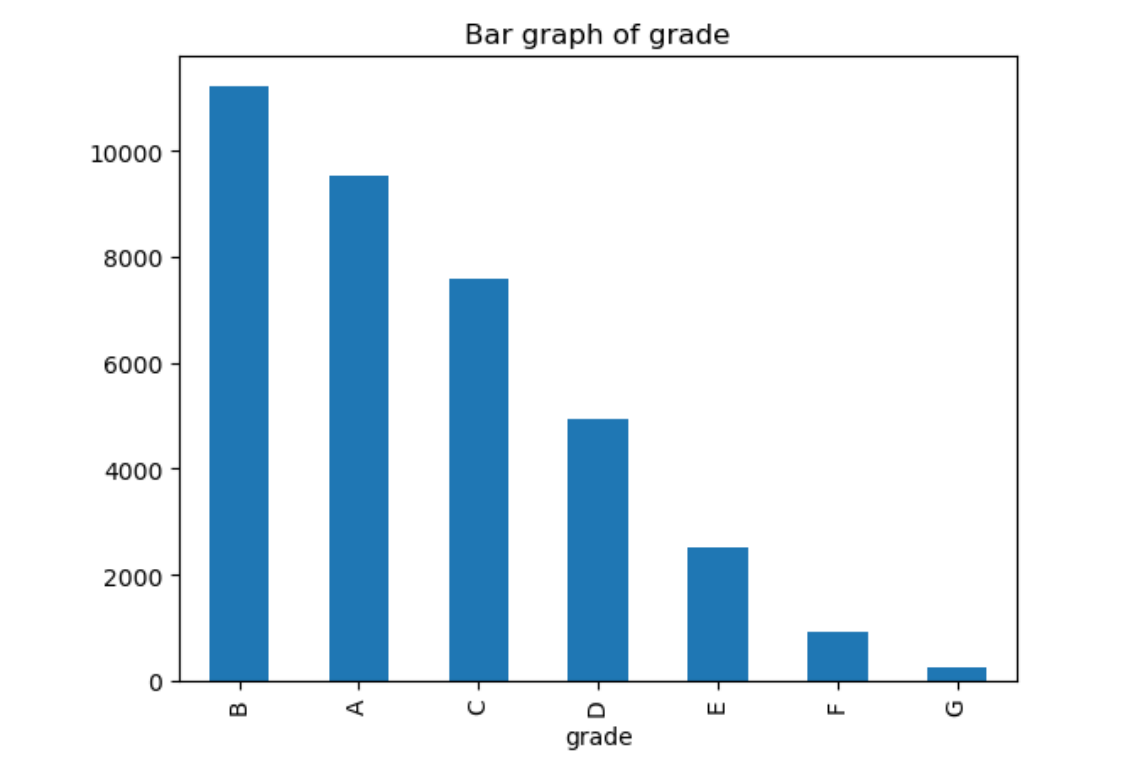
APPROACH

- Utilize EDA concepts learnt to undertake risk analytics in banking and financial sector and understand use of data to minimize risk to clients

BAR GRAPH OF PURPOSE OF TAKING LOAN

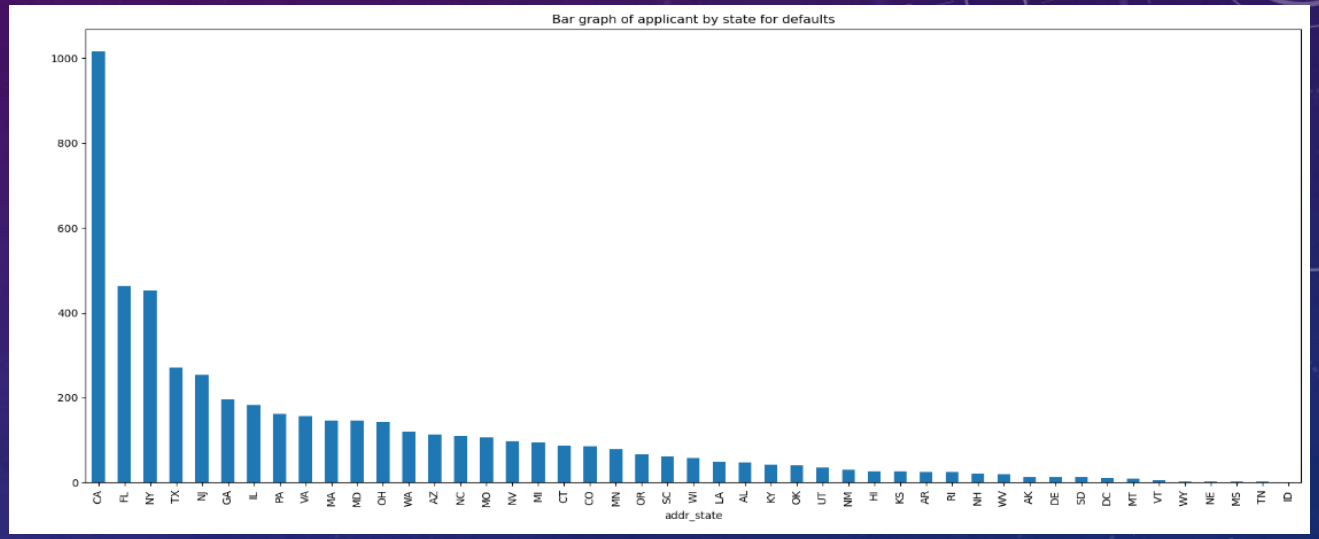


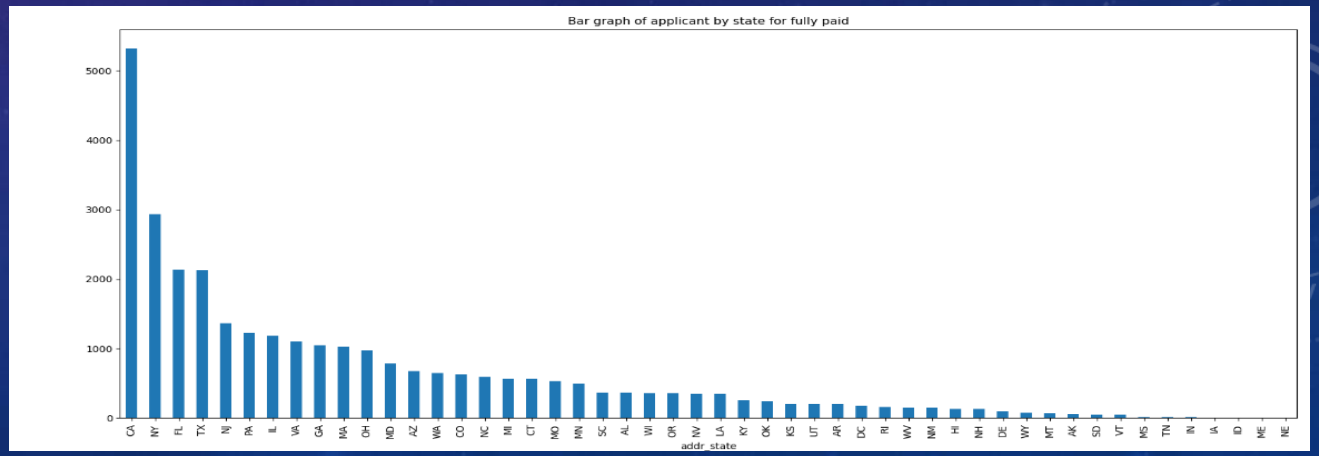
BAR GRAPH OF GRADE



BAR GRAPH OF APPLICANT BY STATE FOR FULLY PAID

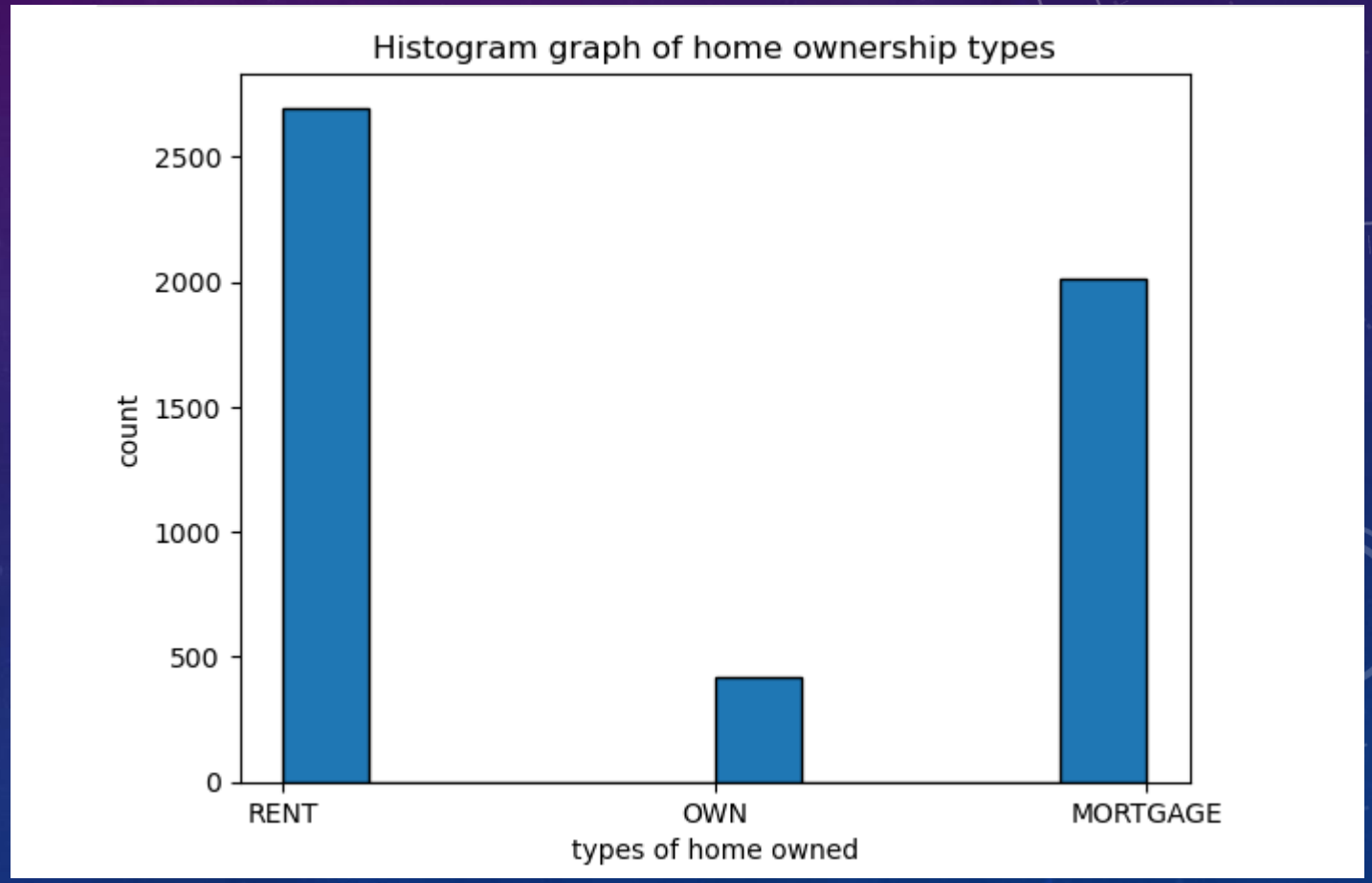
- we can see that states like CA NY TX are preforming better compared to states like FL.





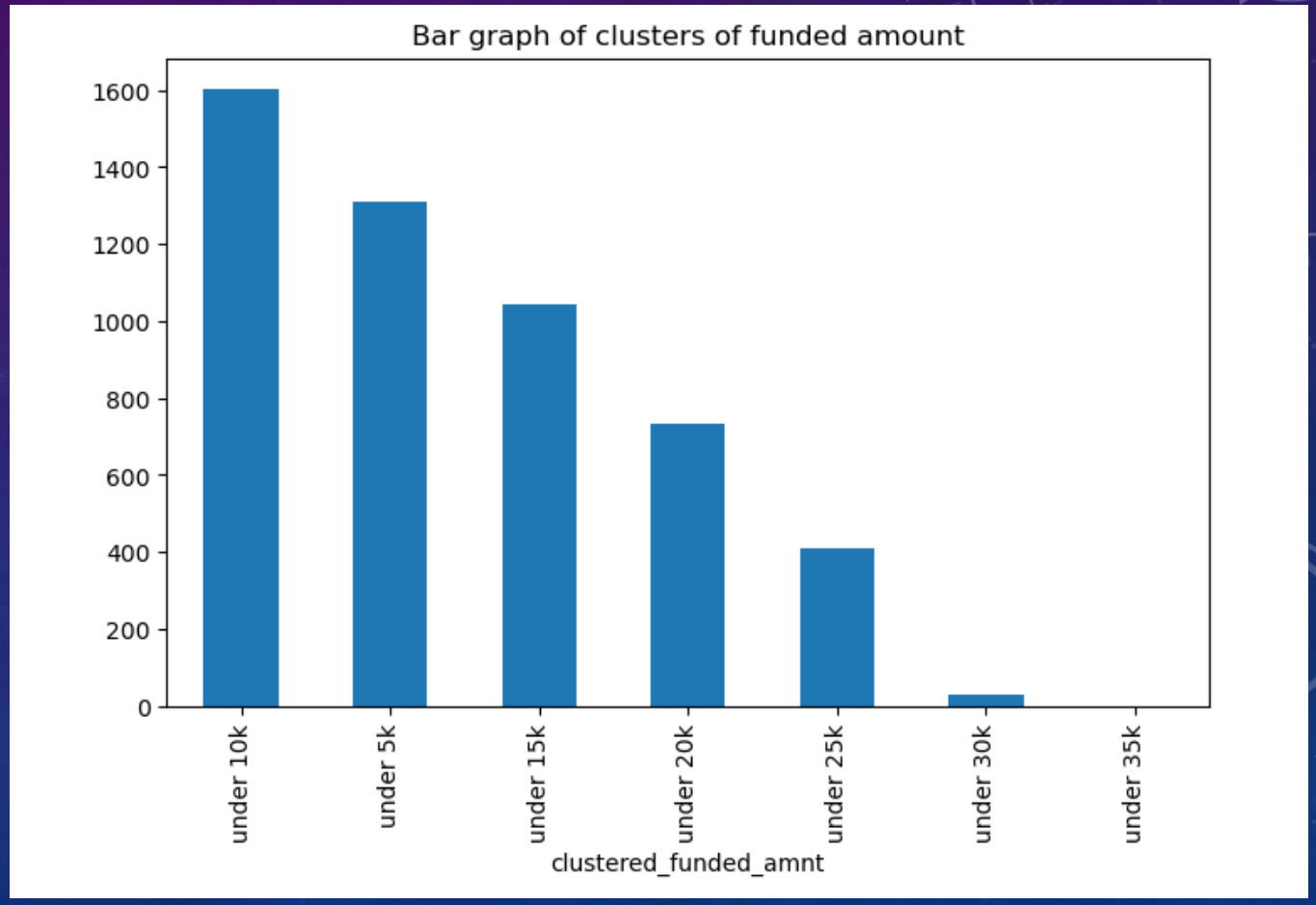
HISTOGRAM GRAPH OF HOME OWNERSHIP TYPES

- The graph says the number of people who has own type are tend to have less defaulted than those who have rented or Mortgage



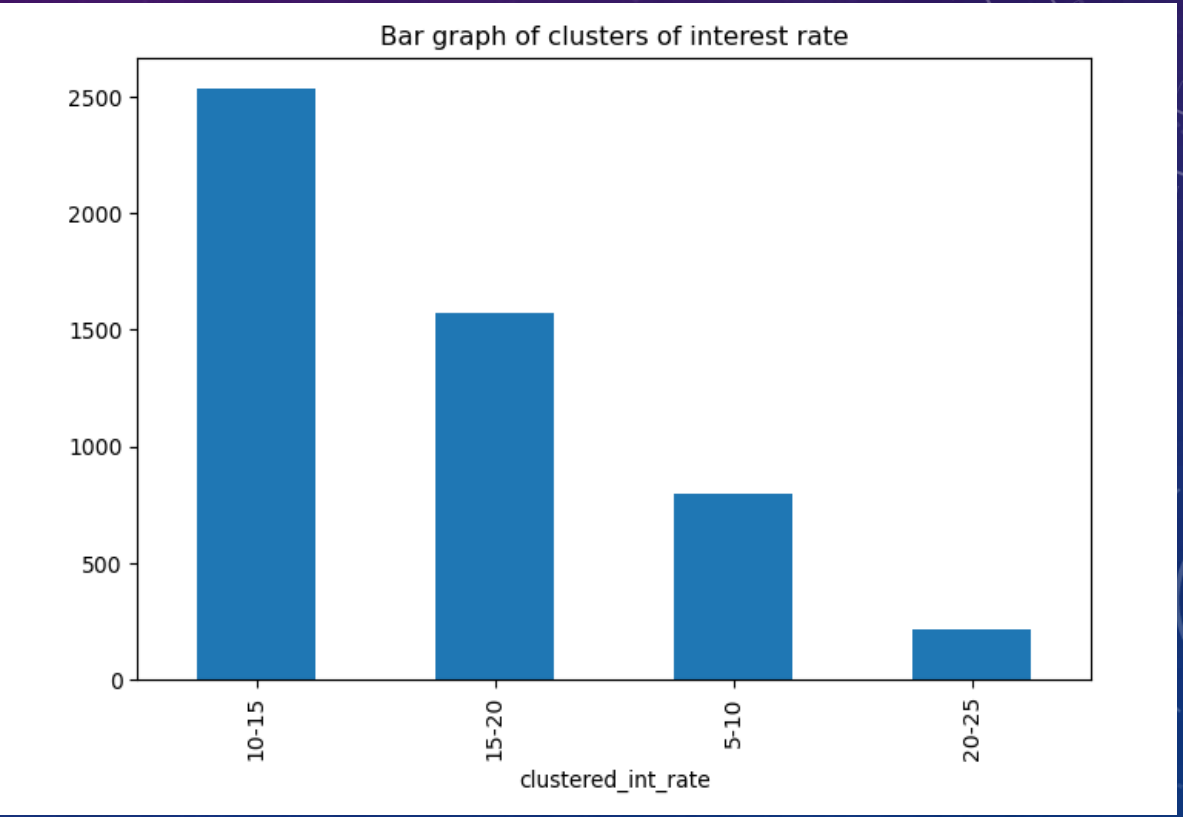
BAR GRAPH OF CLUSTERS OF FUNDED AMOUNT

- most of the loans given ranges from 5-10k, followed by below 5k, then 10-15k, then 15-20k



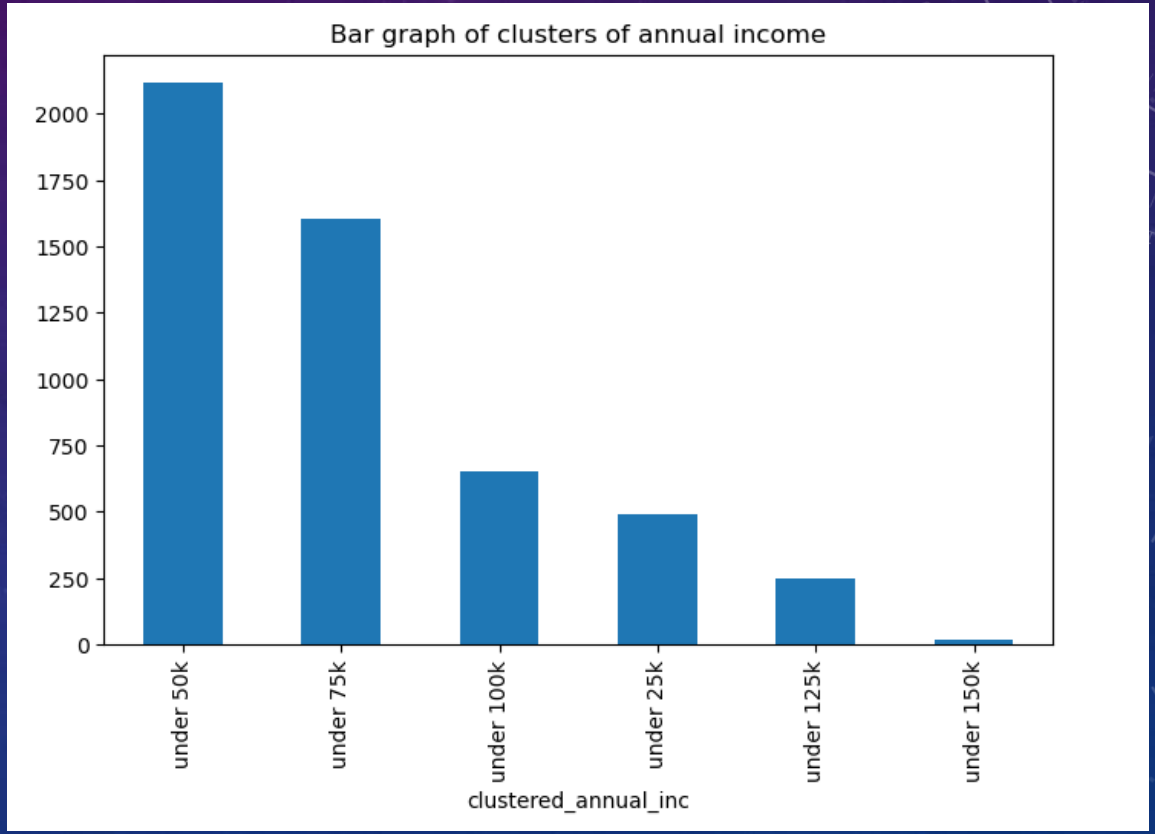
BAR GRAPH OF CLUSTERS OF INTEREST RATE

- The highest number of defaulters are given the loan with the interest rate of 10-15 then 15-20 followed by 5-10. this gives clear indication that most of the people are not able to pay the loan because of the higher interest.



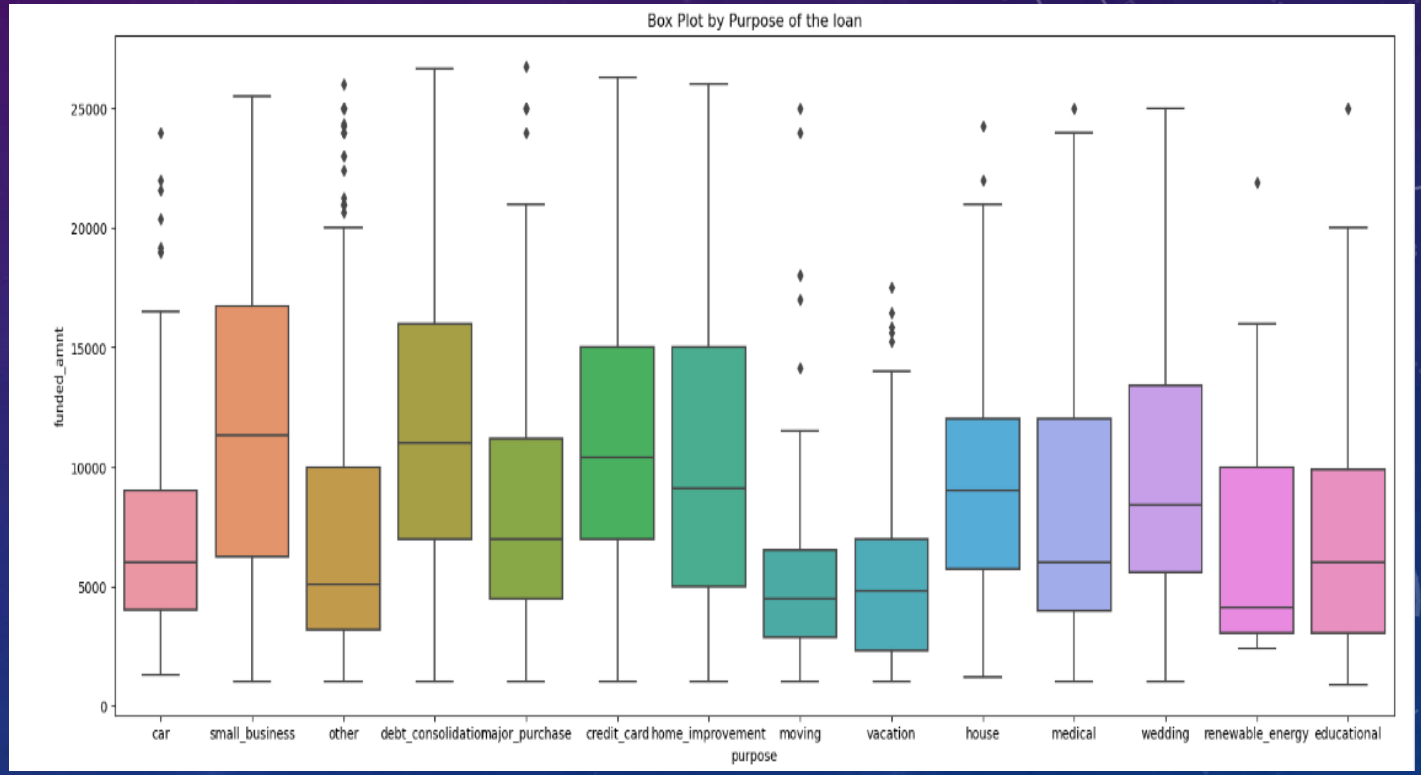
BAR GRAPH OF CLUSTERS OF ANNUAL INCOME

- borrowers whose income is less are tend to be defaulters than those whose income is higher. This can be analyzed clearly with the fund to income and installment to income ratio. This we can do it in the segmented univariant analysis in the next session



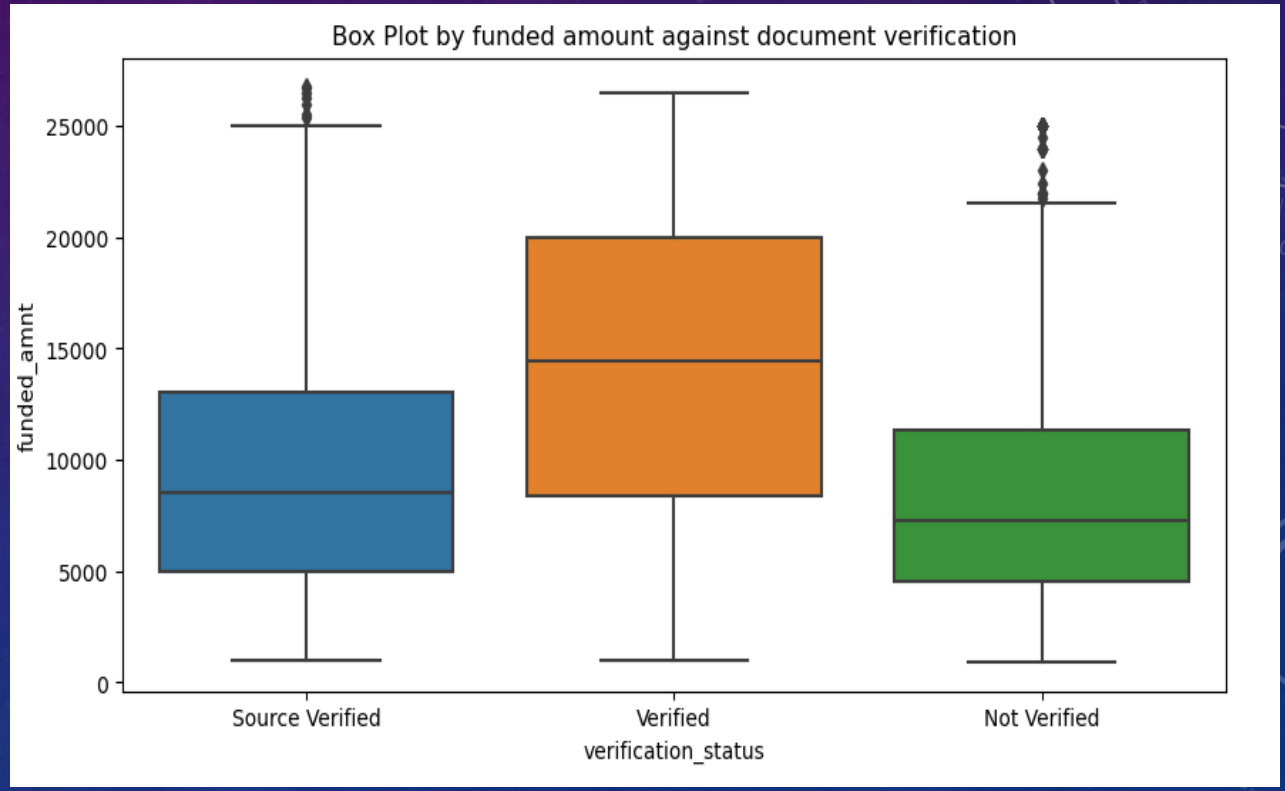
BOX PLOT BY PURPOSE OF THE LOAN

- Credit card, Small business, debt consolidation home improvement and wedding are the reasons generally higher loan are expected. As expected The loan mostly taken for general expense are charged off



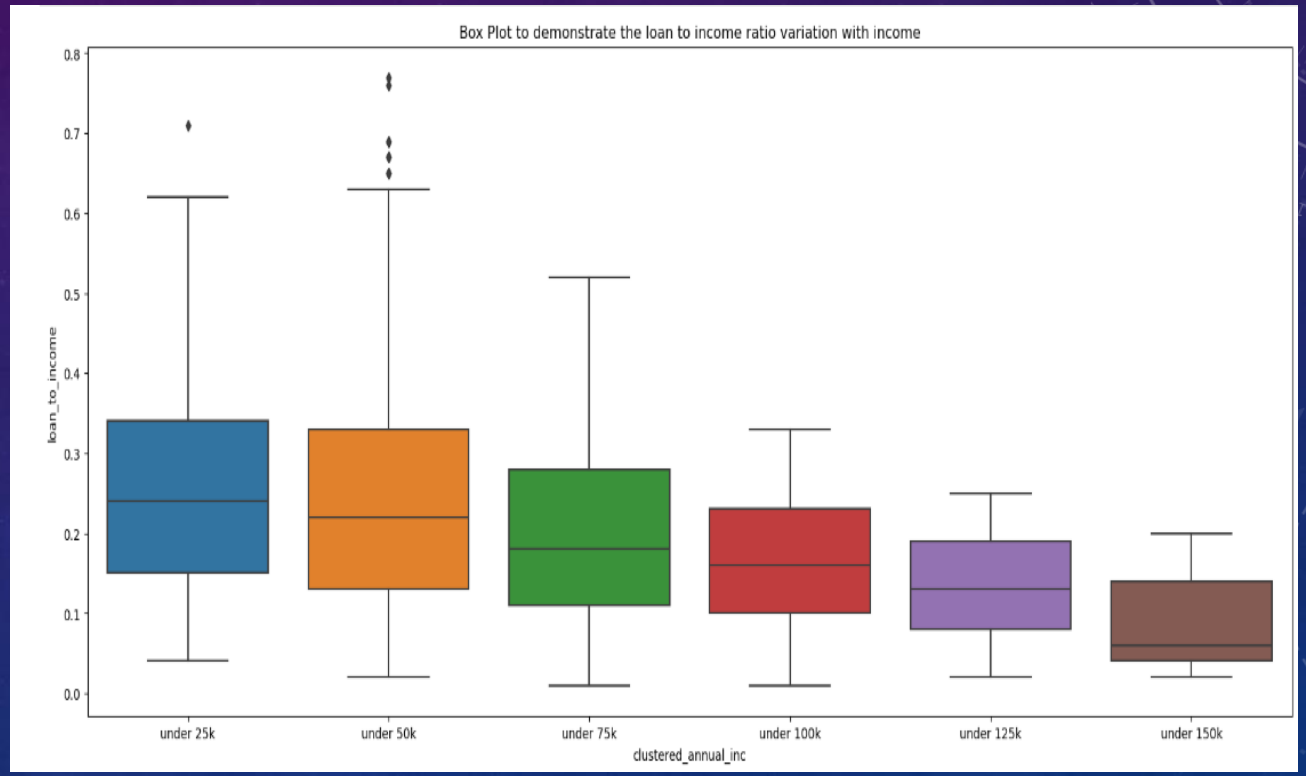
BOX PLOT BY FUNDED AMOUNT AGAINST DOCUMENT VERIFICATION

- As expected the verification is done completely if the funded amount is more. Even though it is verified the defaulters are still seen if the amount is big irrespective of if the docs are verified or not.



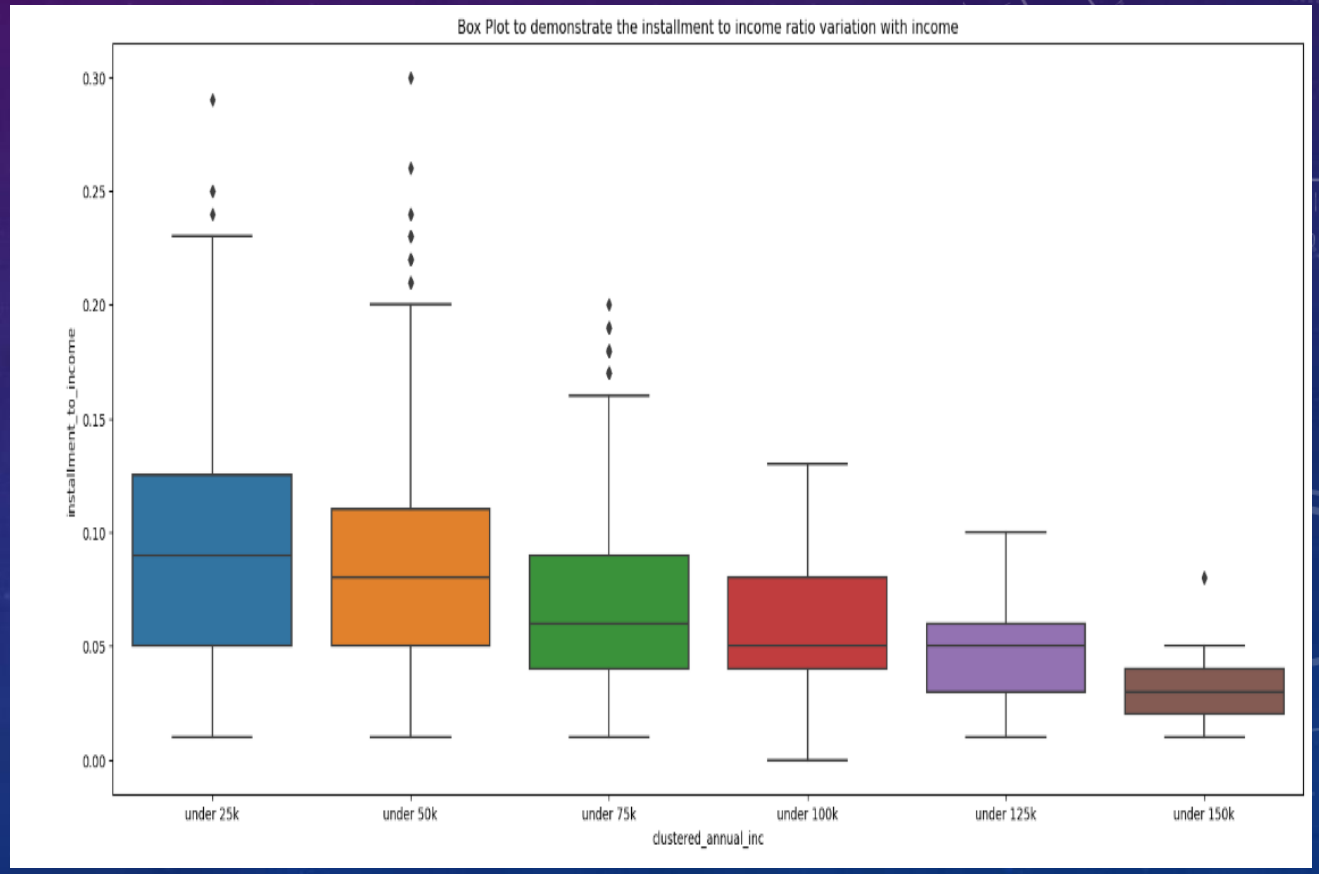
BOX PLOT TO DEMONSTRATE THE LOAN TO INCOME RATIO VARIATION WITH INCOME

- when it comes to borrowers whose income is less are given bad loans than those who has higher income



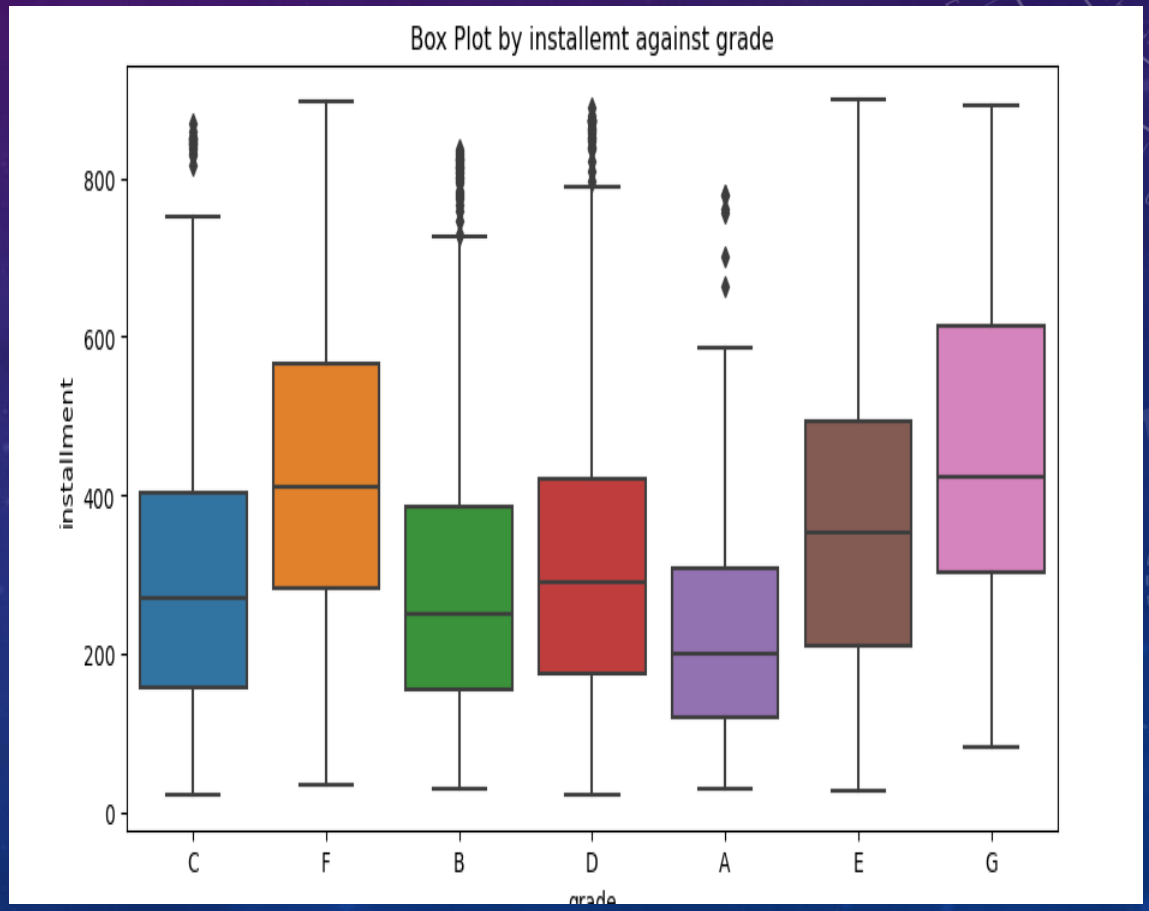
BOX PLOT TO DEMONSTRATE THE INSTALLMENT TO INCOME RATIO VARIATION WITH INCOME

- Again the observation is same, when it comes to borrowers whose income is less are given bad loans than those who has higher income



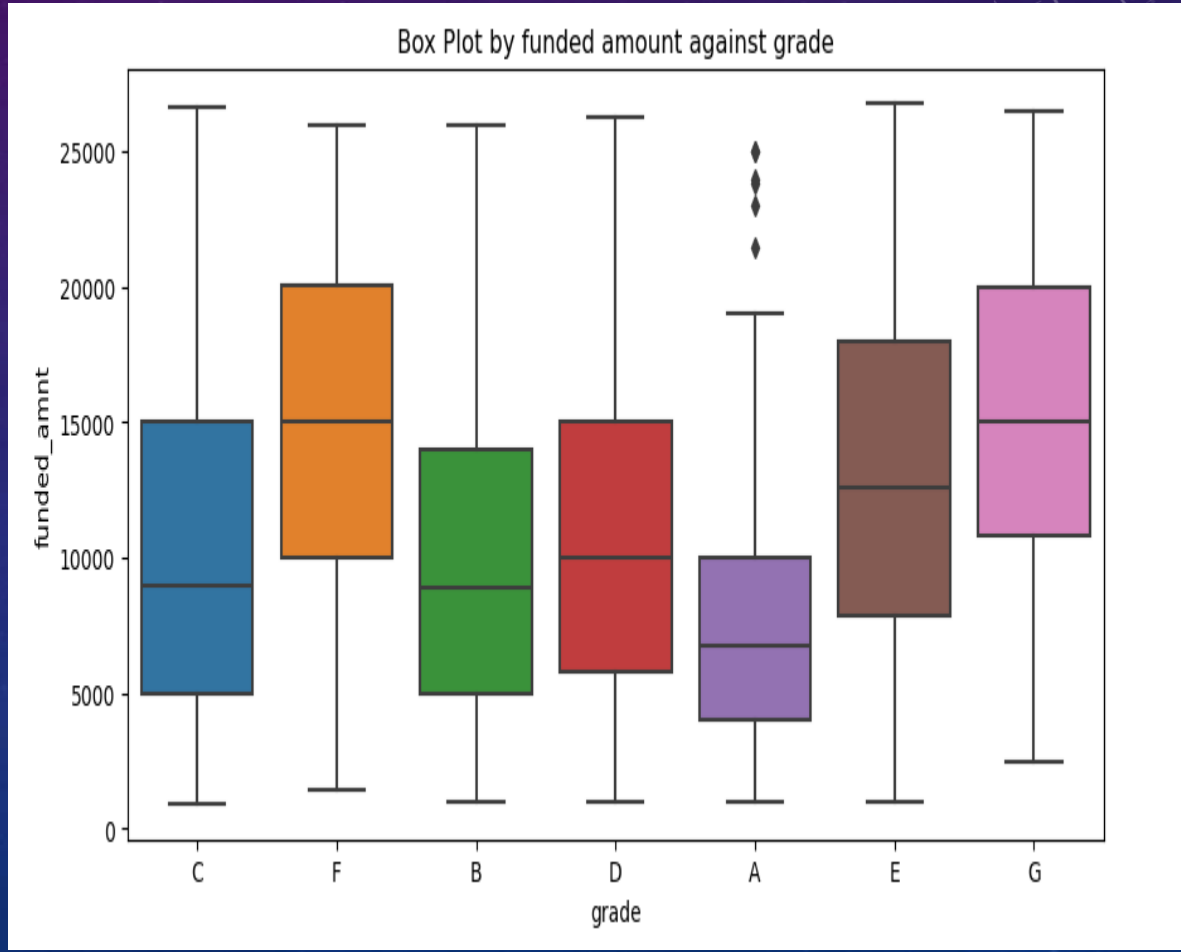
BOX PLOT BY INSTALLMENT AGAINST GRADE

- As expected the performance of the applicant is inversely related to installment to be paid per month. More the installment lesser is the performance



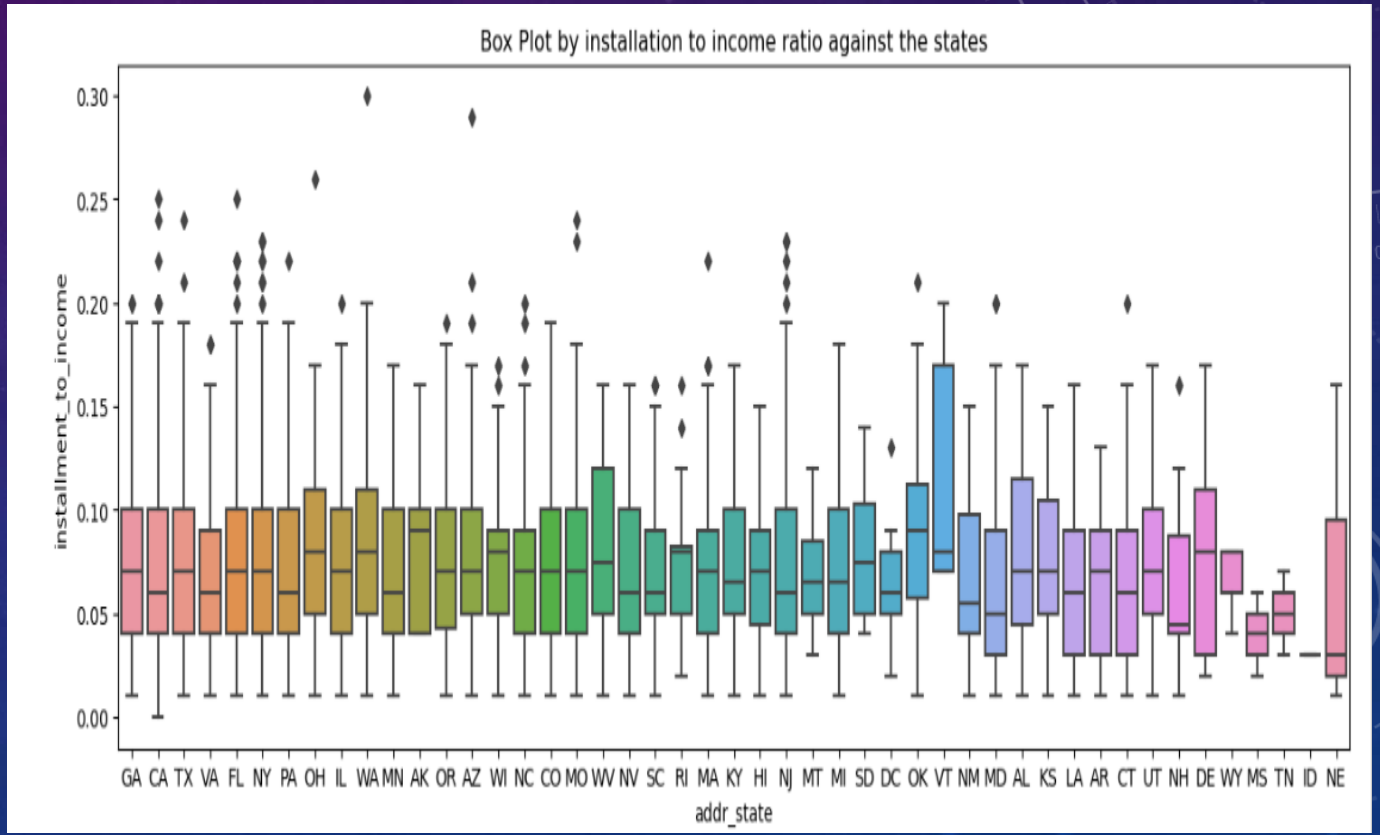
BOX PLOT BY FUNDED AMOUNT AGAINST GRADE

- More the funded amount high chance that the performacne of the loan bearer reduces



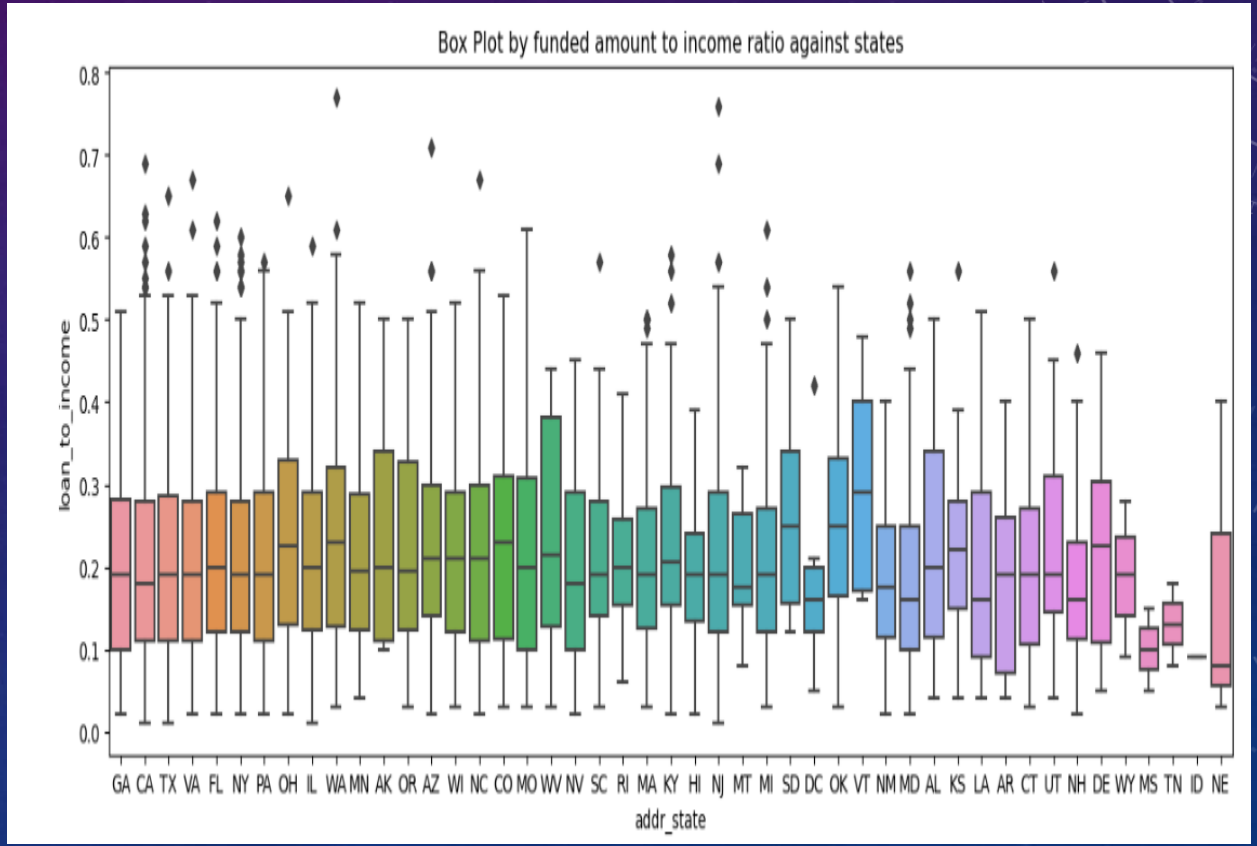
BOX PLOT BY INSTALLATION TO INCOME RATIO AGAINST THE STATES

- States like VT DE WV AL OK are given bad loans compared to the states like DC MS TN



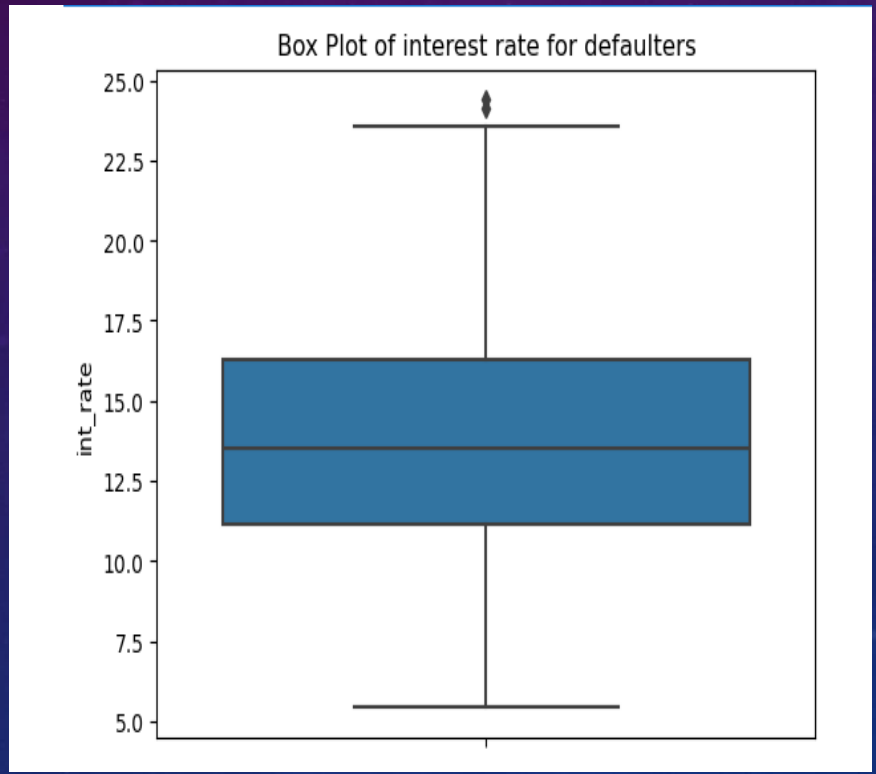
BOX PLOT BY FUNDED AMOUNT TO INCOME RATIO AGAINST STATES

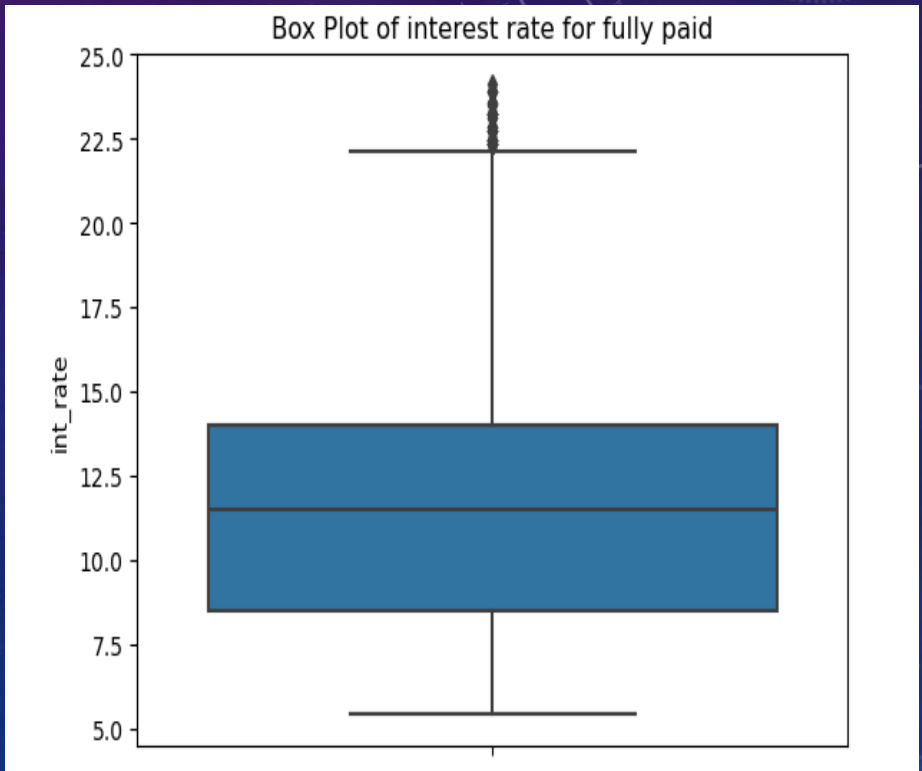
- States like VT WV AL OK are given bad loans compared to the states like DC MS TN



BOX PLOT OF INTEREST RATE FOR DEFAULTERS AND FULLY PAID

- we can see the clear difference between the fully paid are those who got lesser interest rates compared to defaulters





ANALYSIS BETWEEN FUNDED AMOUNT AND ANNUAL INCOME

- the count of higher amount funded for those whose annual income is around 50k to 70k are tend be more than whose annual income more than 80k adding to that, initially the graph goes as a funnel but at the end it becomes flat. If we can try to reduce the slope of the graph in the beginning by doing more strict verification and reduce the funded amount those who has less income we can reduce the huge amount of the defaults