



# Lending Club Case Study

*Group Members:*

- Pavan Kumar Kamisetti
- Saumya Priyam

# Introduction

Lending money to applicants who might not repay it is a major cause of financial loss, known as credit loss. Credit loss means the money a lender loses when borrowers don't pay back their loans or run away without repaying.

Our main goal here is to find these risky loan applicants effectively. This will help us reduce such loans and lower the overall credit loss. We'll use a method called Exploratory Data Analysis (EDA) to identify these applicants, and that's the main focus of this case study.

## Objective

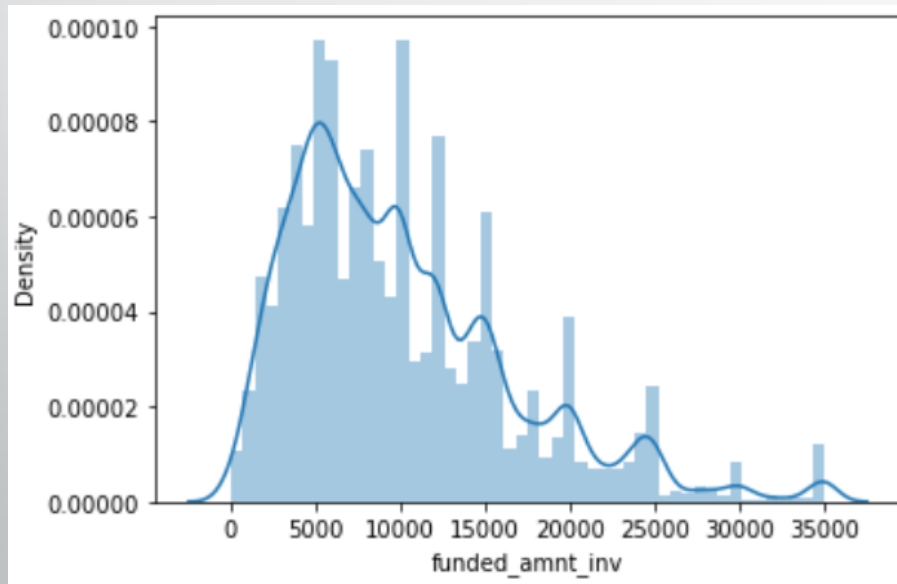
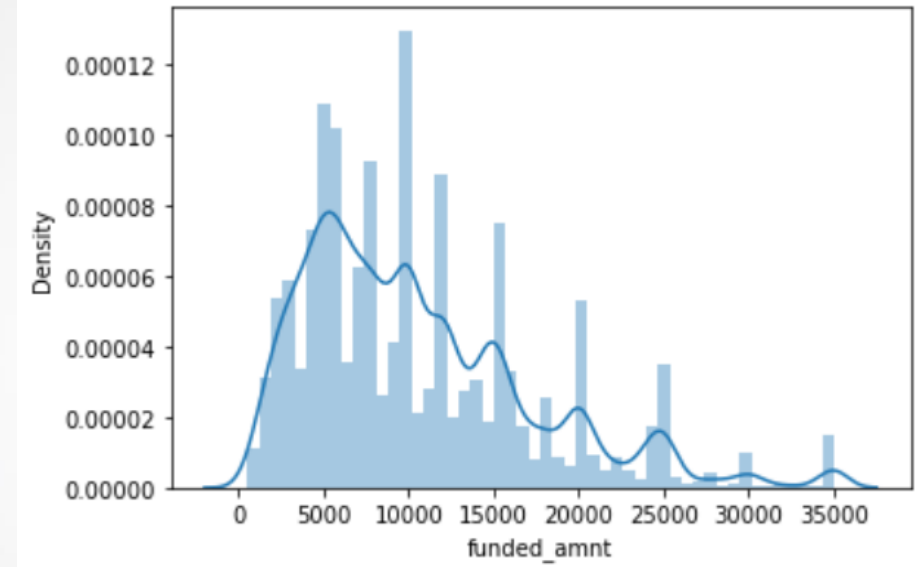
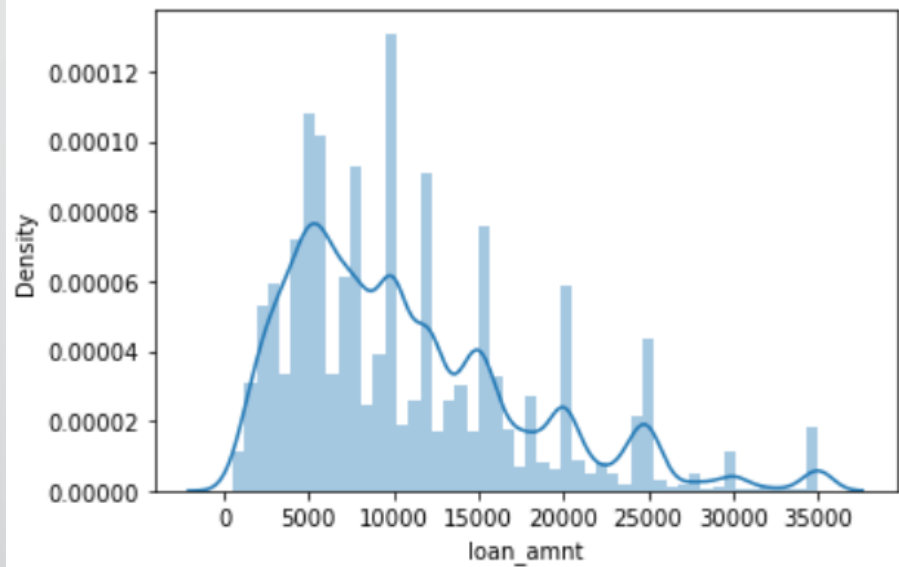
This analysis is all about finding out what factors make loans go bad. We call these factors 'driver variables' because they strongly indicate the risk of a loan not being repaid. The company can use this knowledge to improve how it manages its loans and assess the risk involved.

# Methodology

The following steps outline the process:

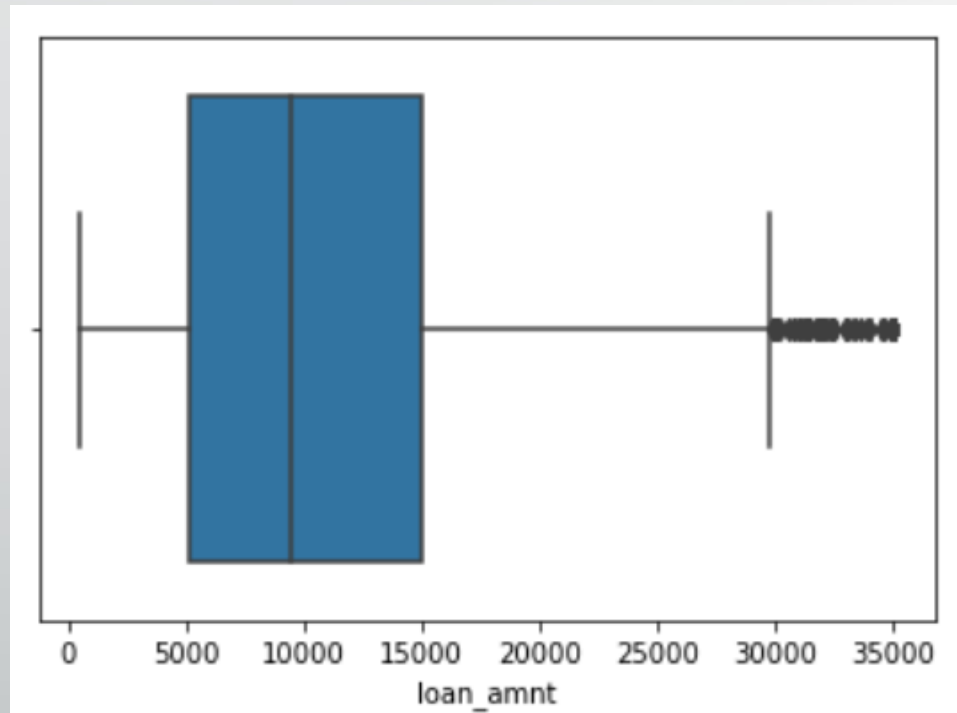
- 1. Data Cleaning:** In this initial step, the data will be cleaned and prepared for analysis, ensuring that it is accurate and free from inconsistencies.
- 2. Univariate Analysis:** This step involves examining individual variables one at a time to gain insights into their distributions and characteristics.
- 3. Segmented Univariate Analysis:** Here, the focus shifts to studying variable segments, helping identify patterns within specific groups or categories.
- 4. Bivariate/Multivariate Analysis:** This step explores relationships between variables, including how multiple factors may influence loan default.
- 5. Recommendations:** The final step presents the findings and conclusions drawn from the analysis, providing actionable insights for the company's portfolio and risk assessment strategies.

# ANALYSIS

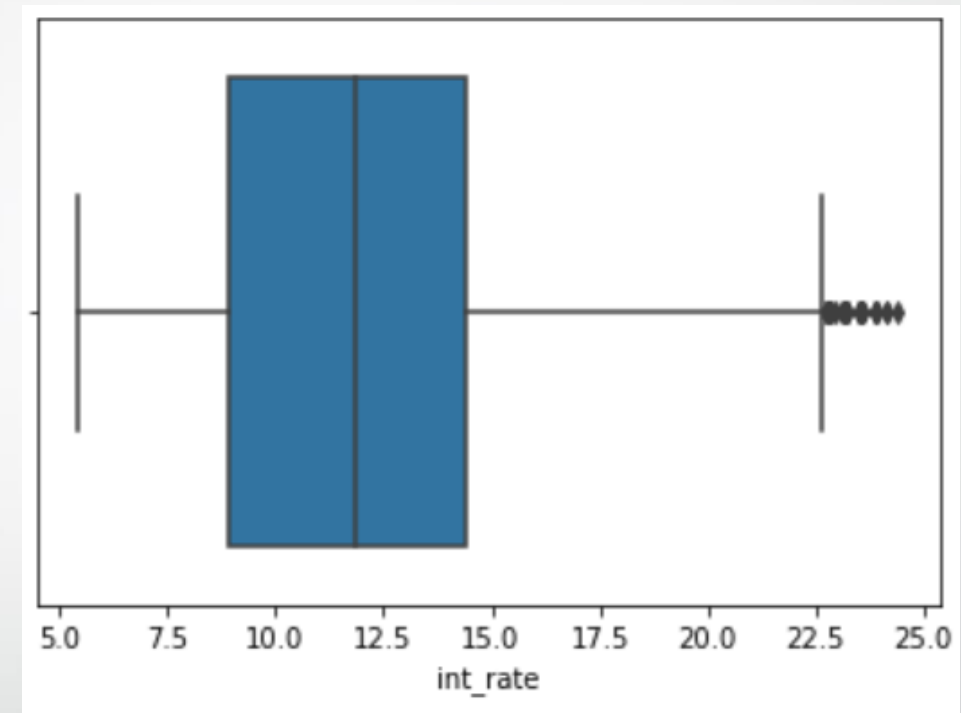


The plots look almost the same, we can opt for one for further analysis. We would use 'loan amount' for further analysis.

# Analysing Trends

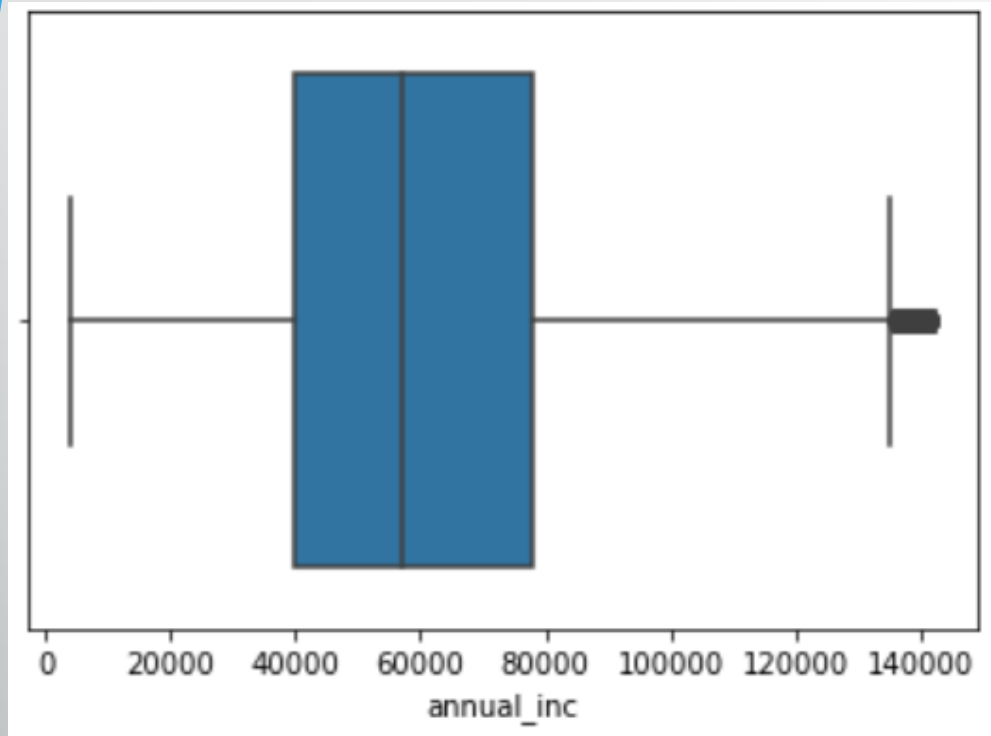


It can be deduced that most of the loan are in the range 5000 to 15000.

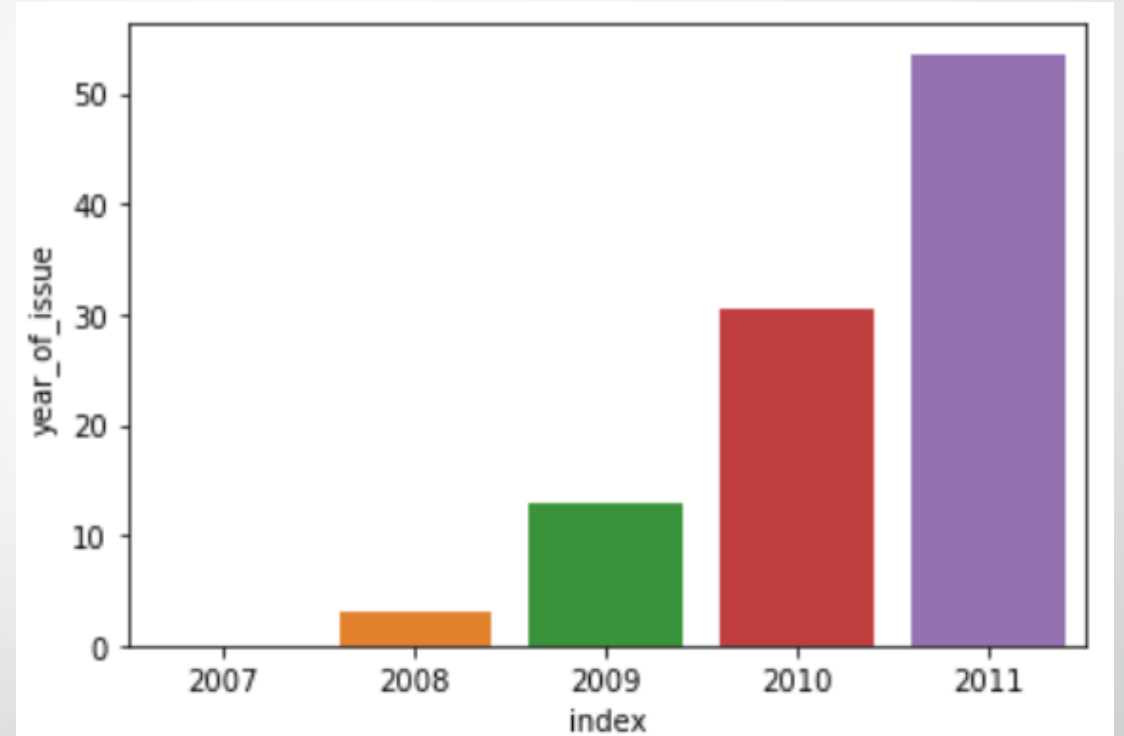


It can be deduced that most of the interest rate on the loan is between 8.5% and 14.5%

# Analysing Trends

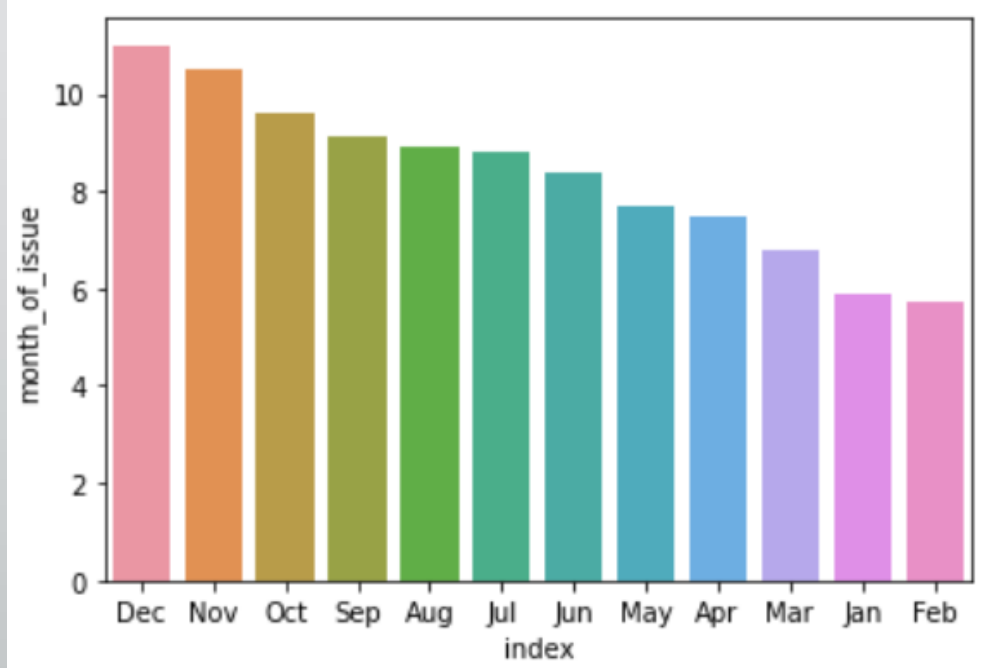


It can be seen that most of the income ranges from 40000 to approximately 80000

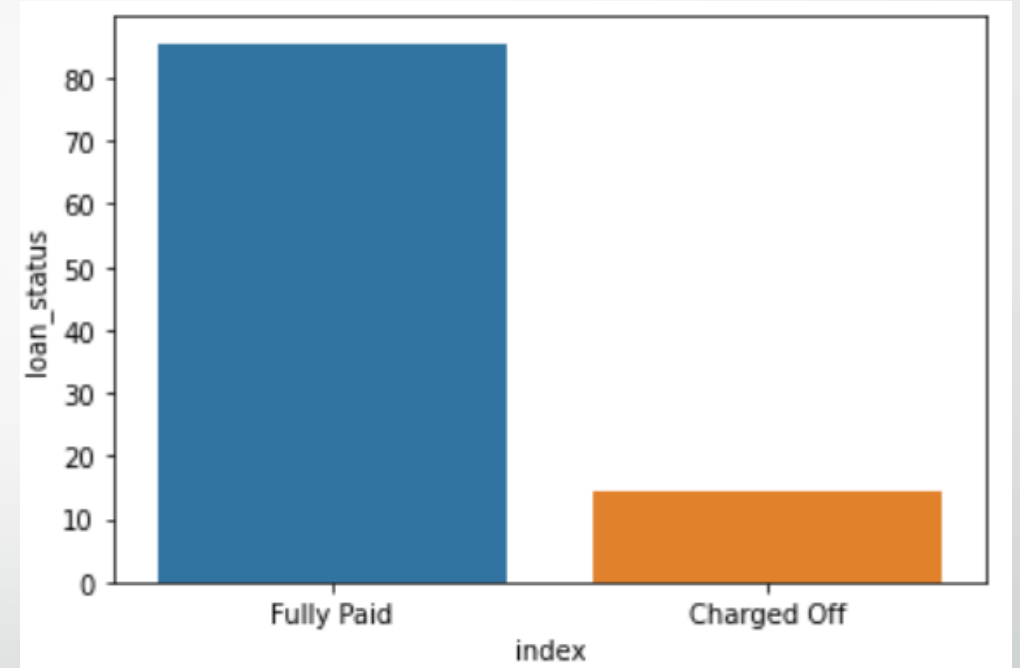


It can be seen that as the loan amount disbursal has increased over the years

# Analysing Trends

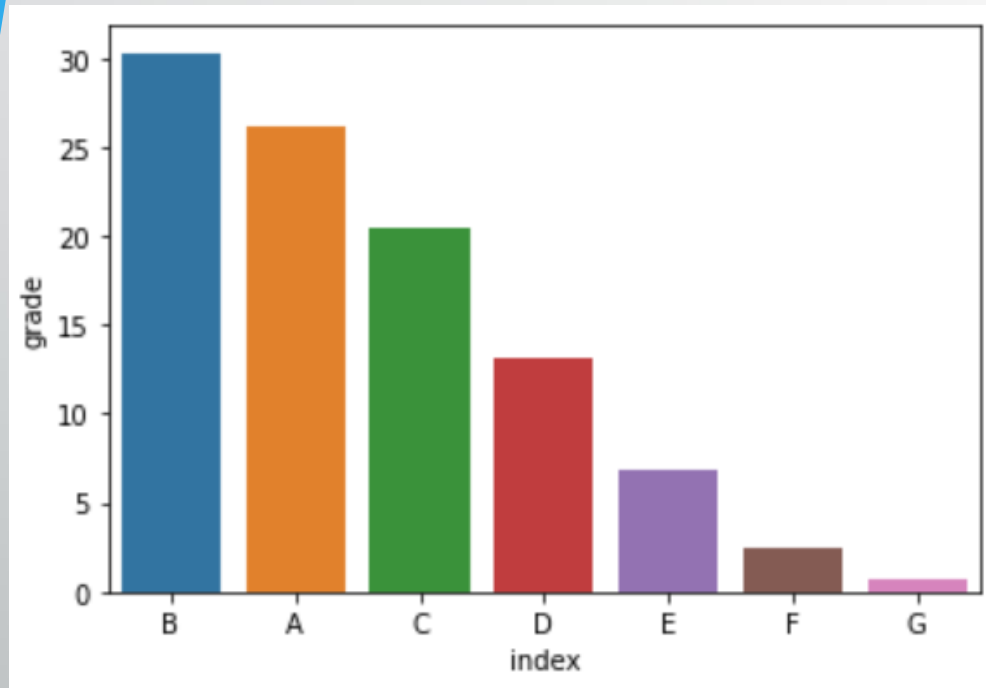


It can be seen that the loan disbursement increases as the year end approaches. Maybe to achieve loan targets, or due to festive mood.

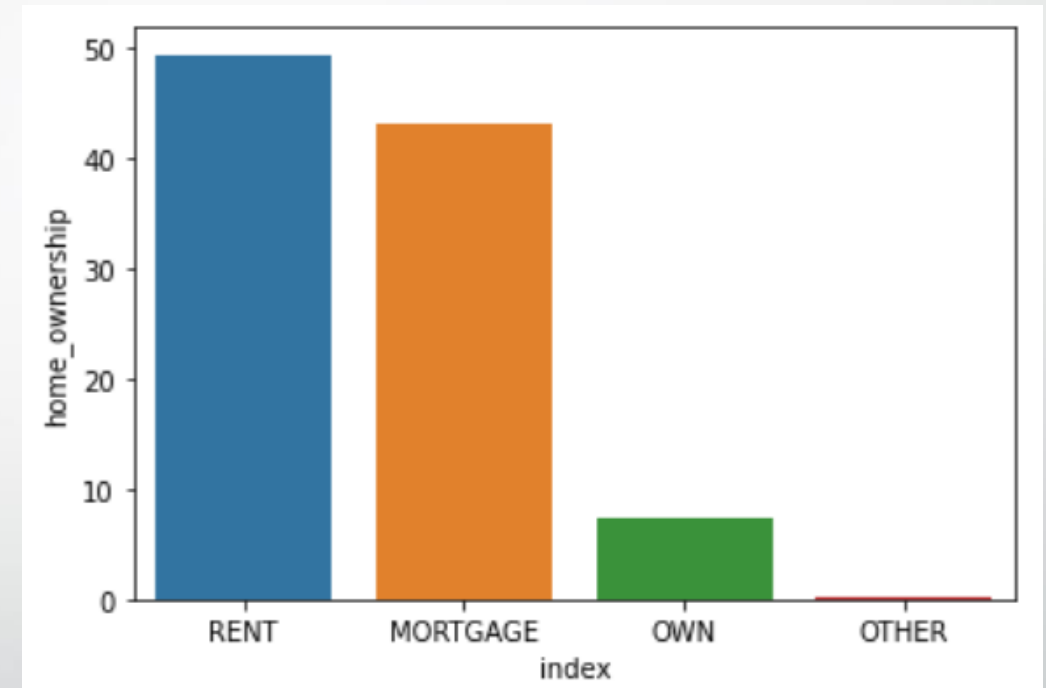


85.5% have fully paid the loan, and 14.5% have been charged off.

# Analysing Trends



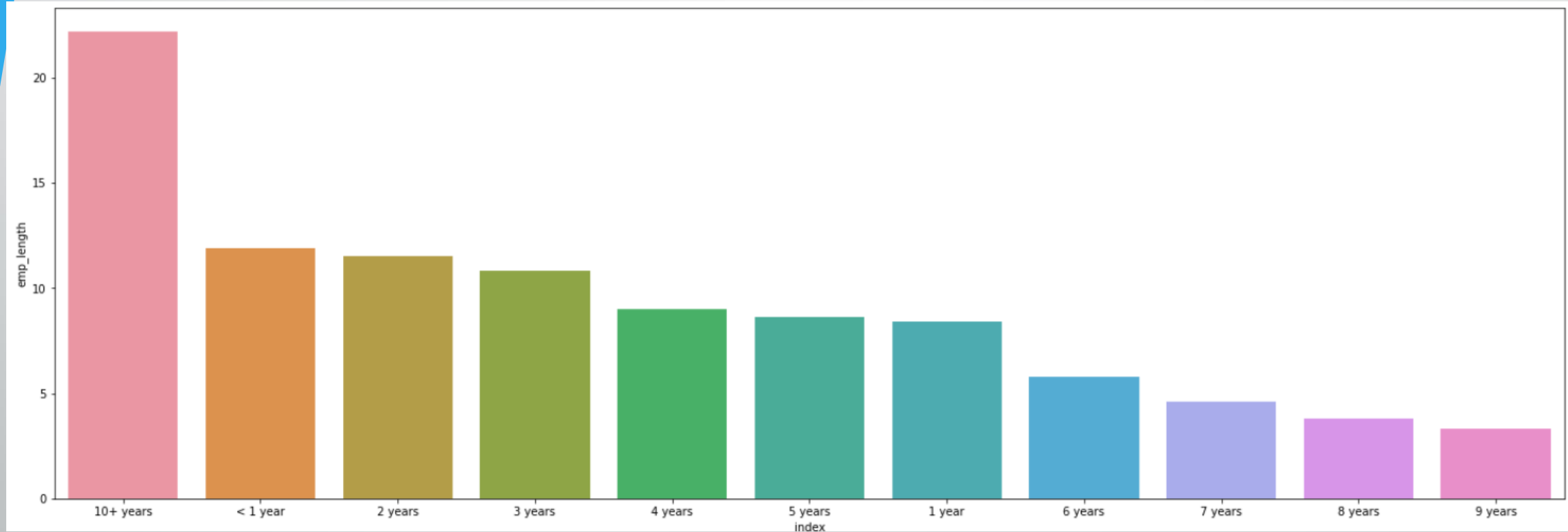
Most of the loan is by group B.



Borrowers who have rented home or mortgage have taken more loans compared to others.

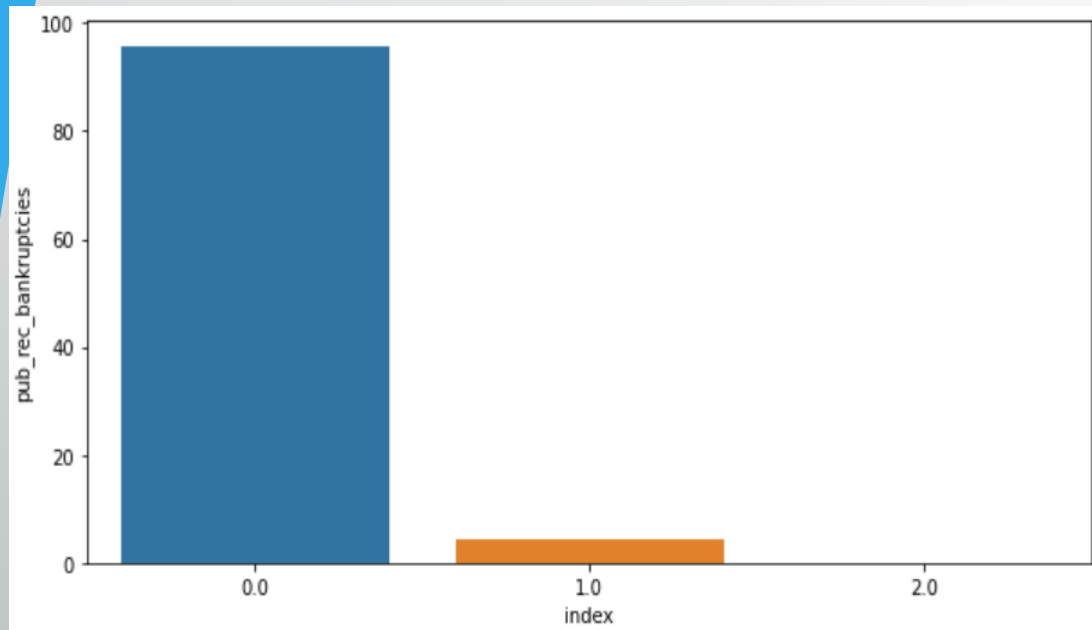


# Analysing Trends

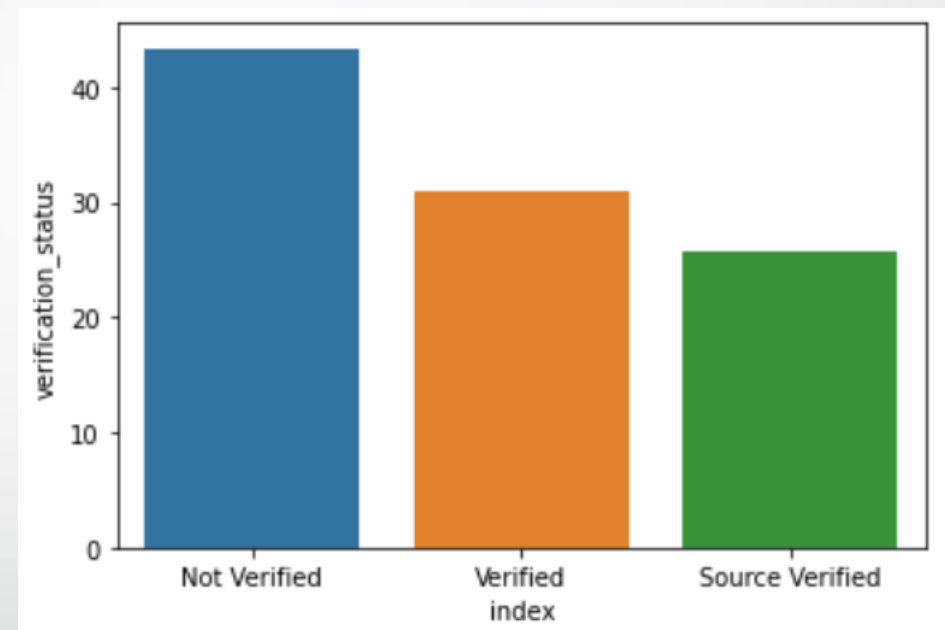


Most of the loan was by the ones who have been in employment for more than 10 years.

# Analysing Trends

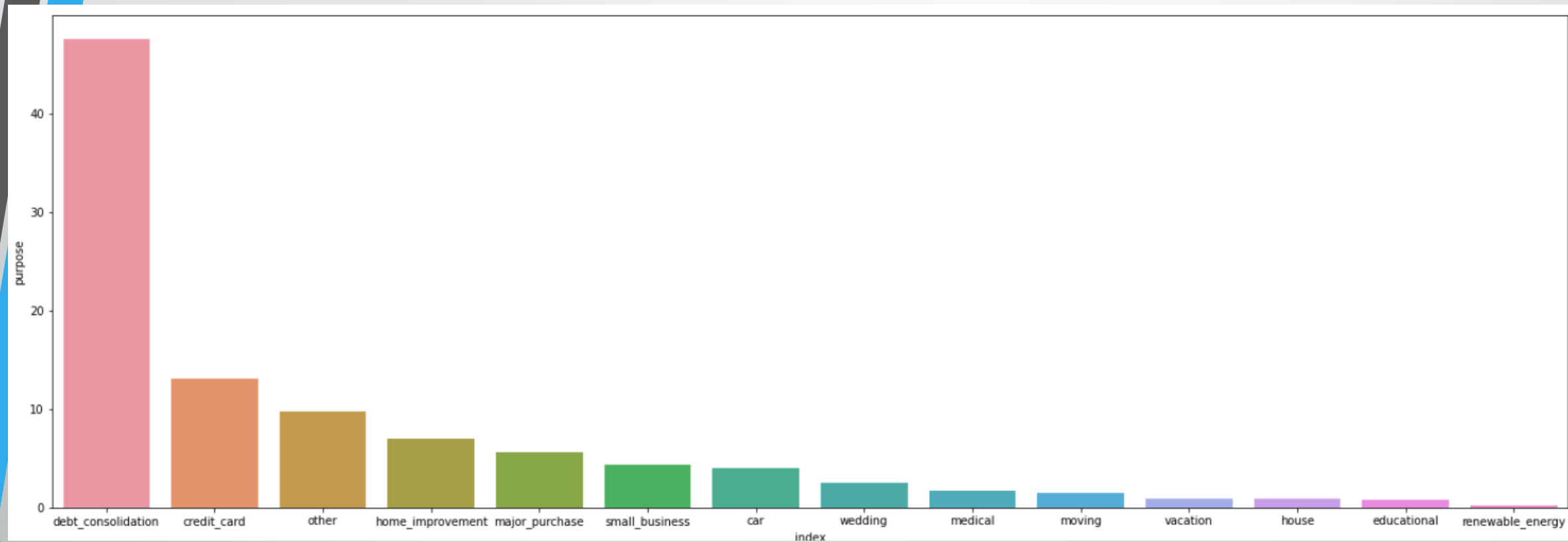


The number of public record bankruptcies which is 0, is very high.



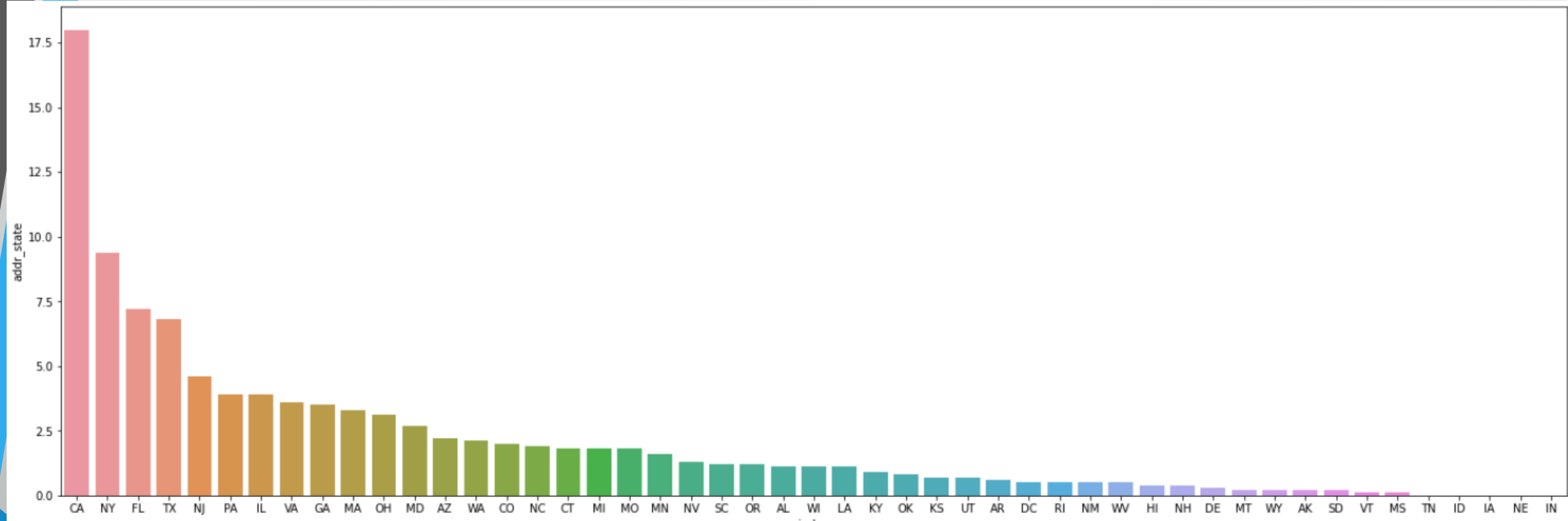
Even though the number of Not Verified is highest, the difference is not much.

# Analysing Trends



Majority of the bad loan is used for debt consolidation

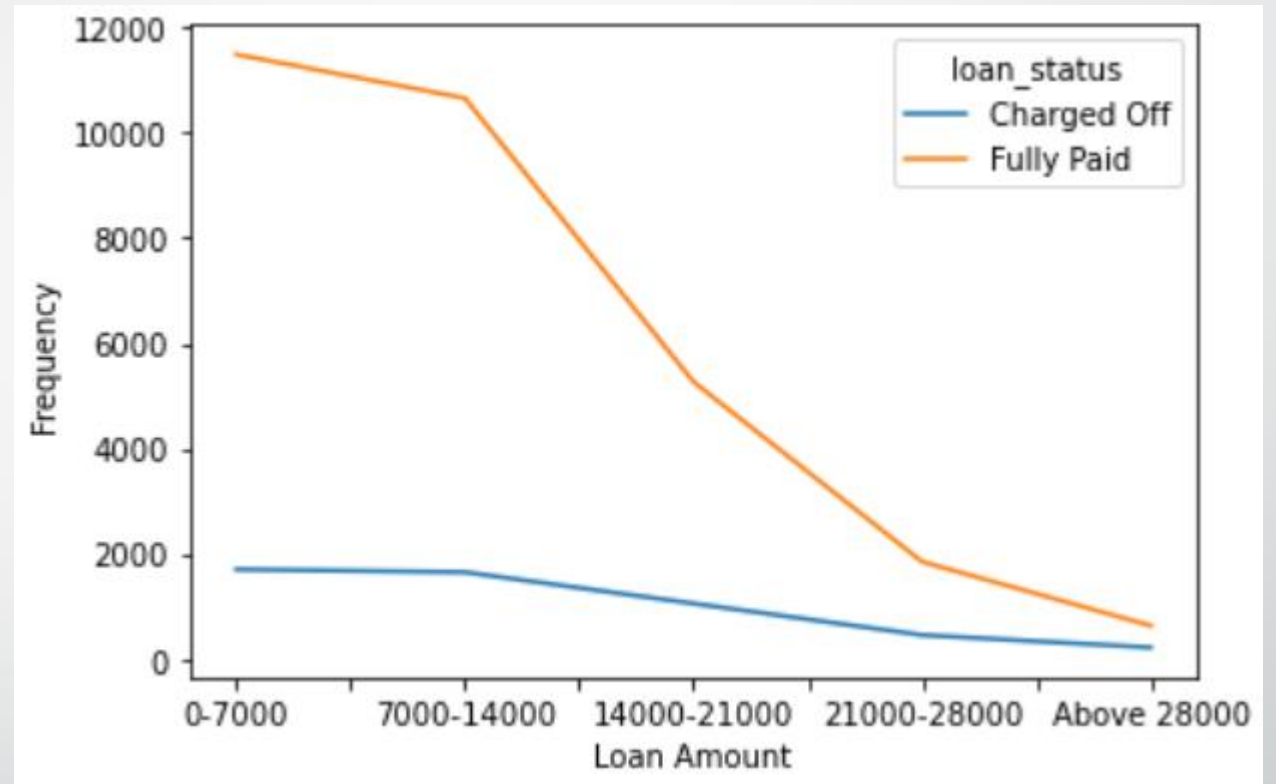
# Analysing Trends



Major chunk of the bad loan is from the states CA, NY, FL, TX and NJ.

# Loan Amount v/s Loan Status

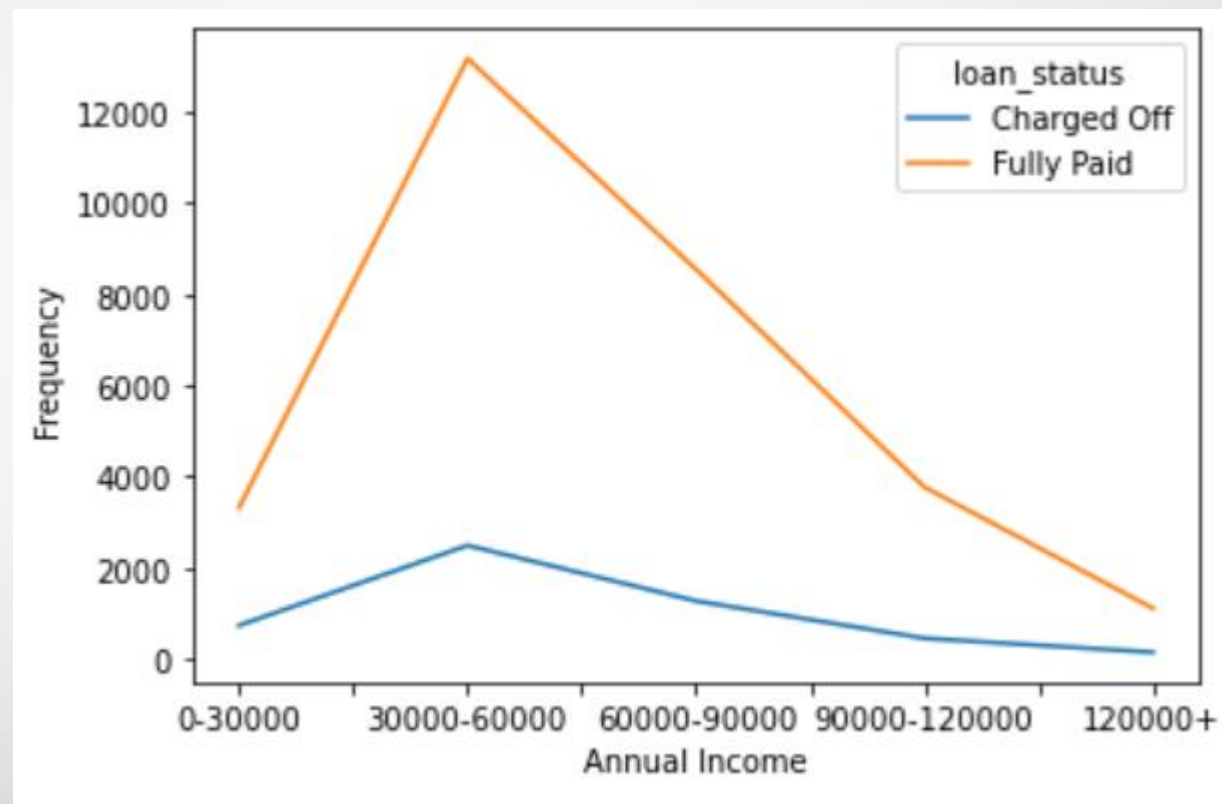
grp_loan_amnt	loan_status	
0-7000	Fully Paid	11495
	Charged Off	1701
7000-14000	Fully Paid	10664
	Charged Off	1648
14000-21000	Fully Paid	5268
	Charged Off	1052
21000-28000	Fully Paid	1846
	Charged Off	452
Above 28000	Fully Paid	630
	Charged Off	216



from the above chart, it can be observed that most loans which were charged off are under 21000.

# Annual Income v/s Loan Status

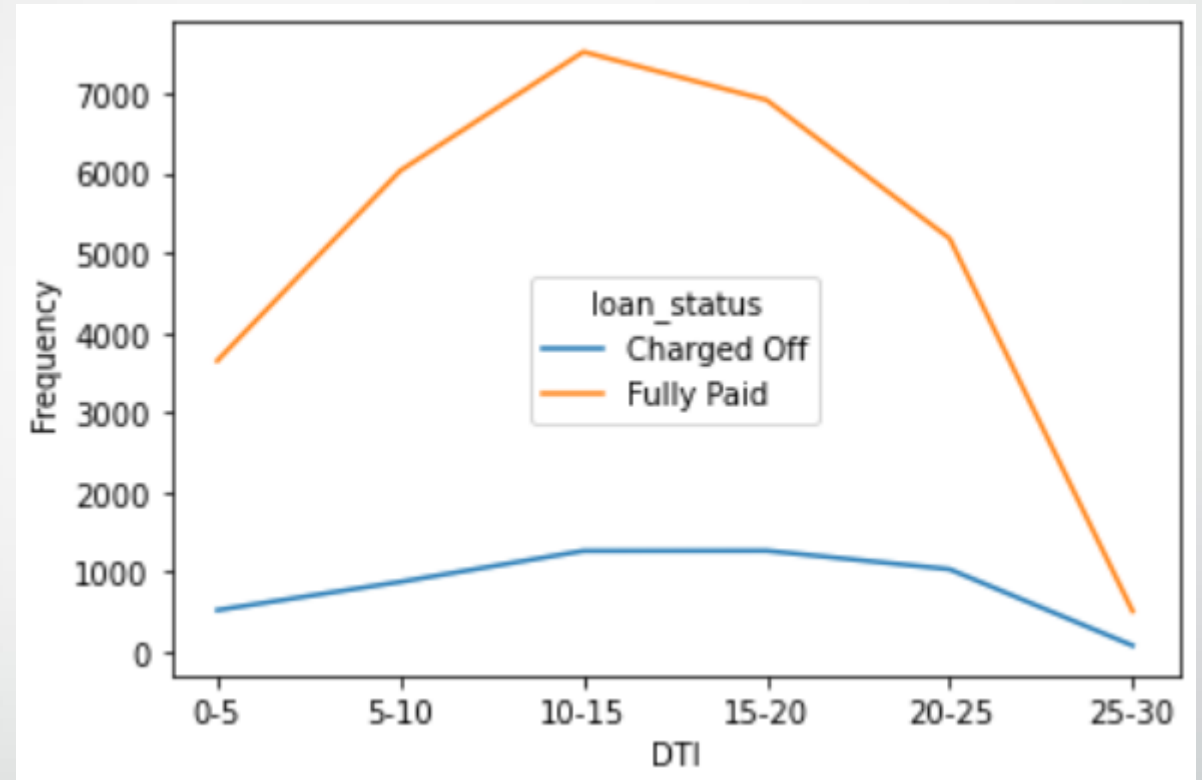
grp_annual_inc	loan_status	
0-30000	Fully Paid	3316
	Charged Off	729
30000-60000	Fully Paid	13170
	Charged Off	2485
60000-90000	Fully Paid	8546
	Charged Off	1267
90000-120000	Fully Paid	3765
	Charged Off	447
120000+	Fully Paid	1106
	Charged Off	141



Most borrowers have annual income in the range 30000-60000

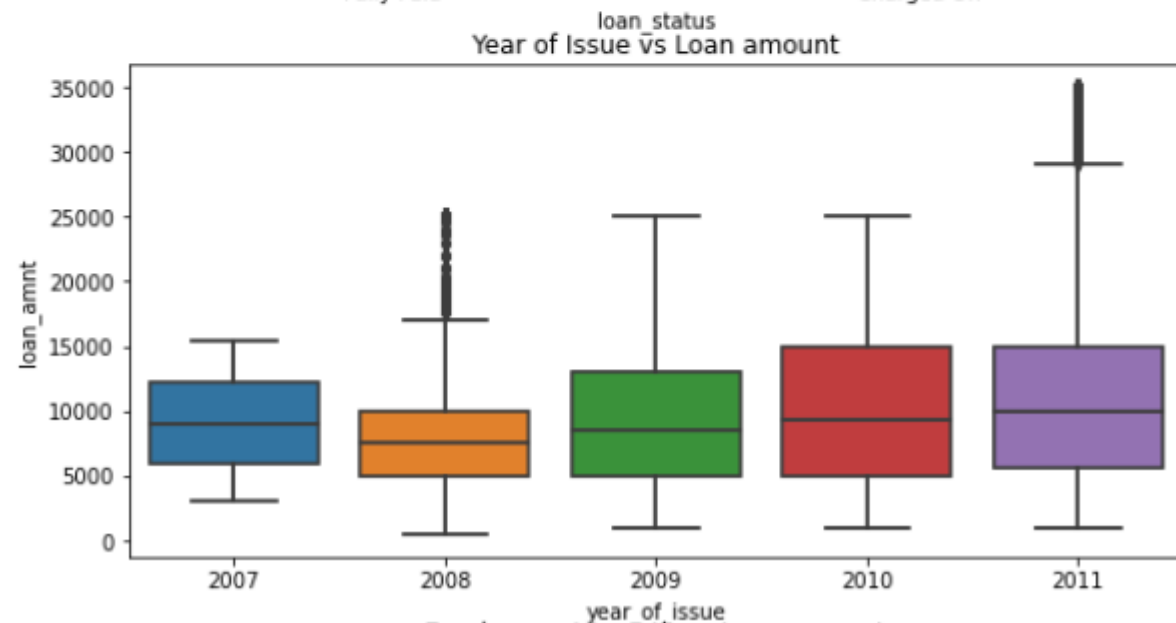
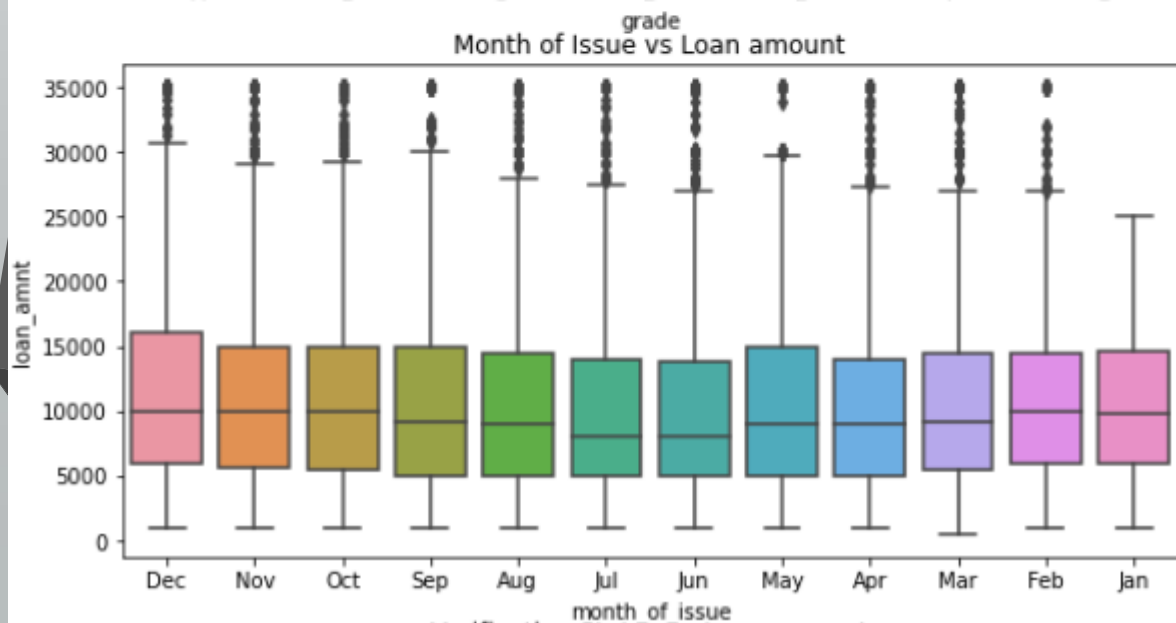
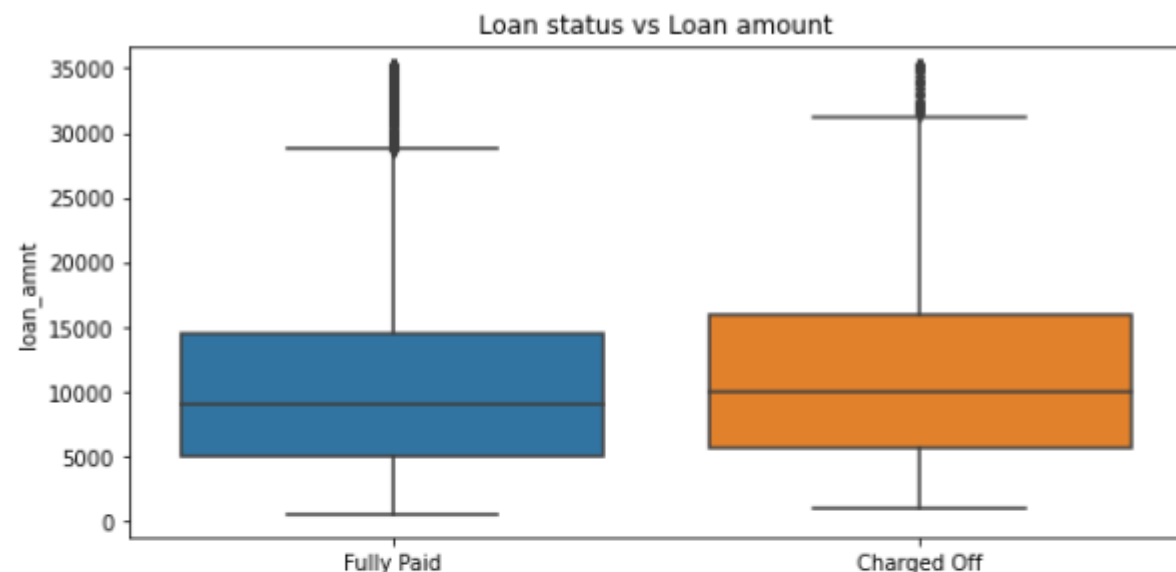
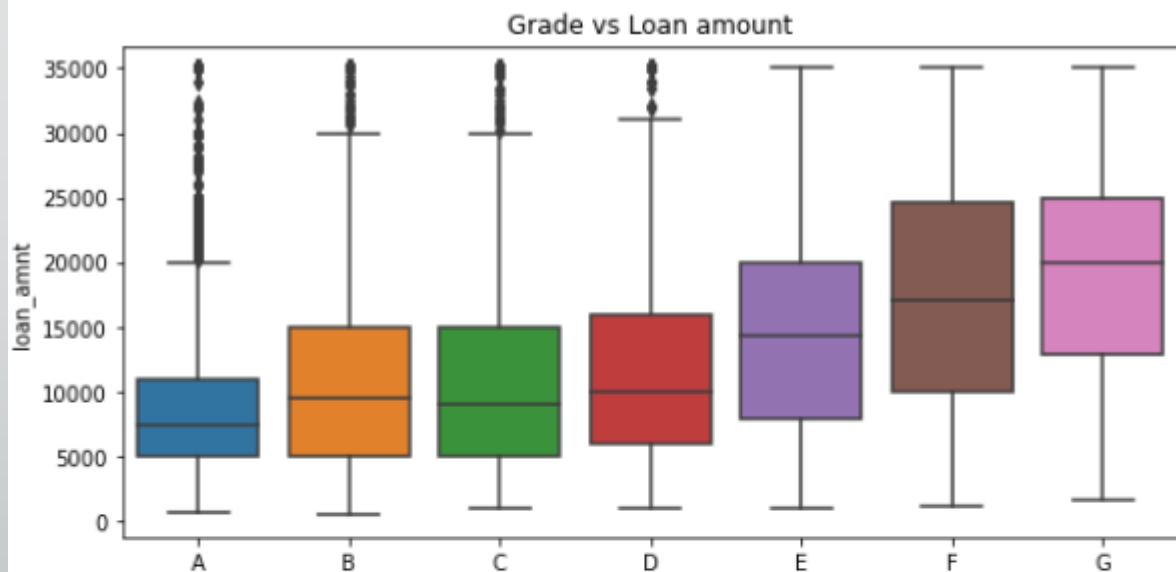
## DTI v/s Loan Status

grp_dti	loan_status	
0-5	Fully Paid	3646
	Charged Off	523
5-10	Fully Paid	6027
	Charged Off	877
10-15	Fully Paid	7515
	Charged Off	1265
15-20	Fully Paid	6912
	Charged Off	1267
20-25	Fully Paid	5178
	Charged Off	1035
25-30	Fully Paid	509
	Charged Off	79



The most charged off loans have a DTI between 10 and 25

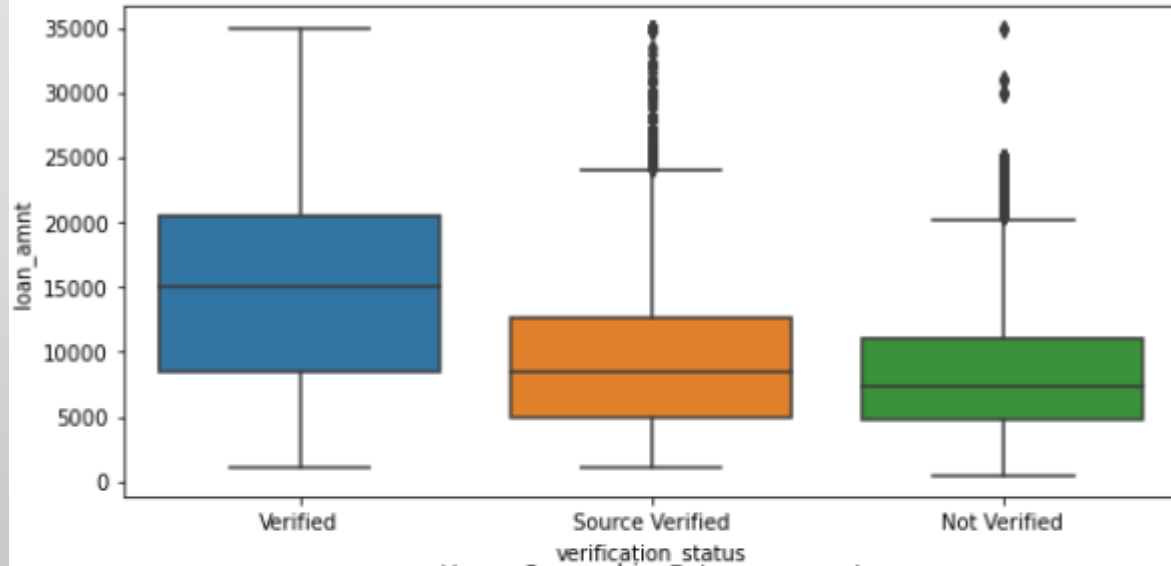
# Analysing Loan Amount with every Variable



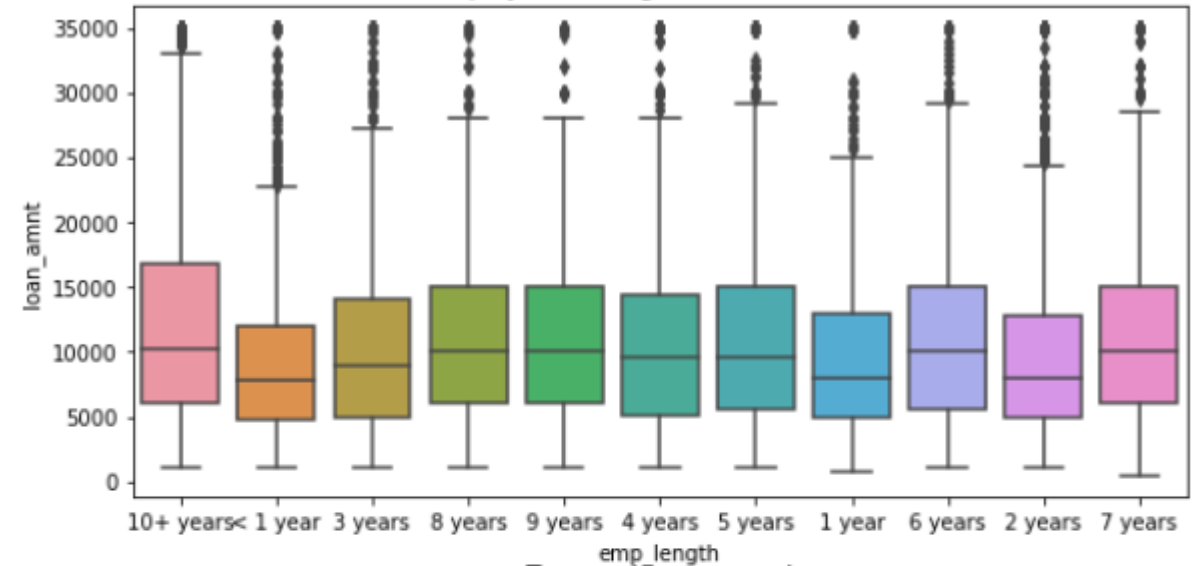


# Analysing Loan Amount with every Variable

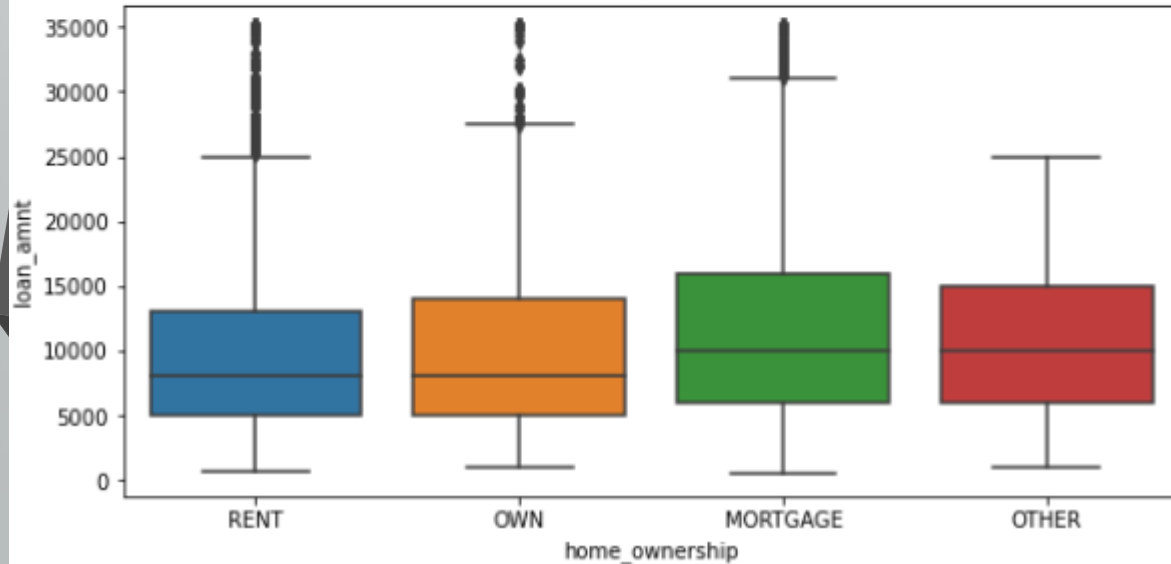
Verification Status vs Loan amount



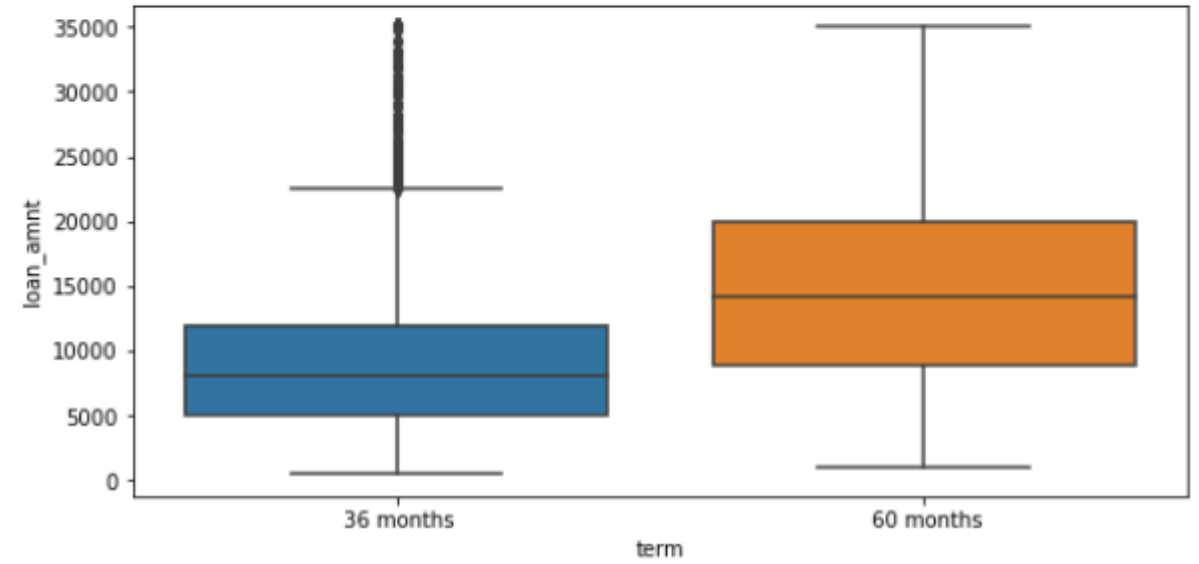
Employment Length vs Loan amount



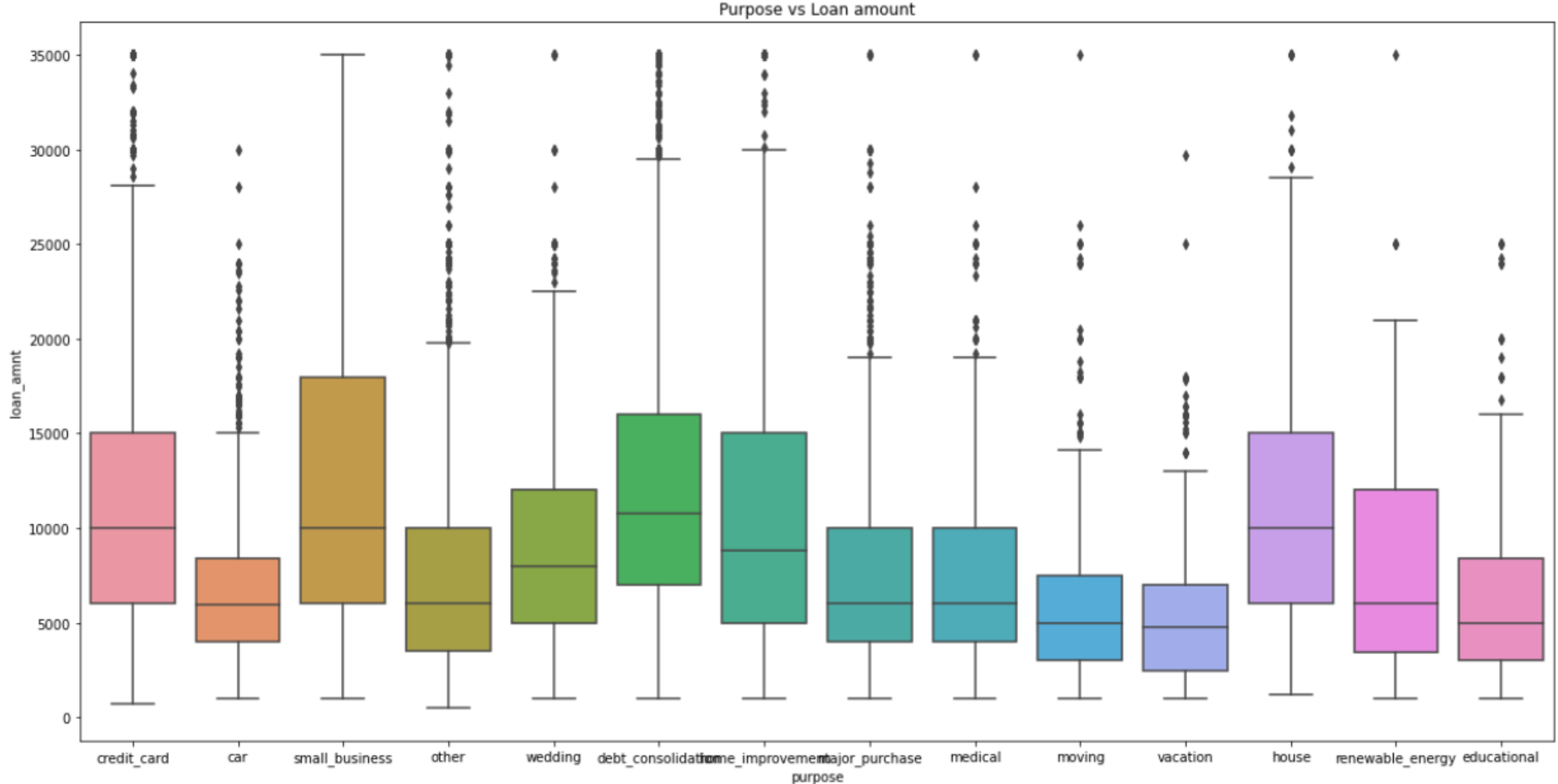
Home Ownership vs Loan amount



Term vs Loan amount



# Analysing Loan Amount with every Variable

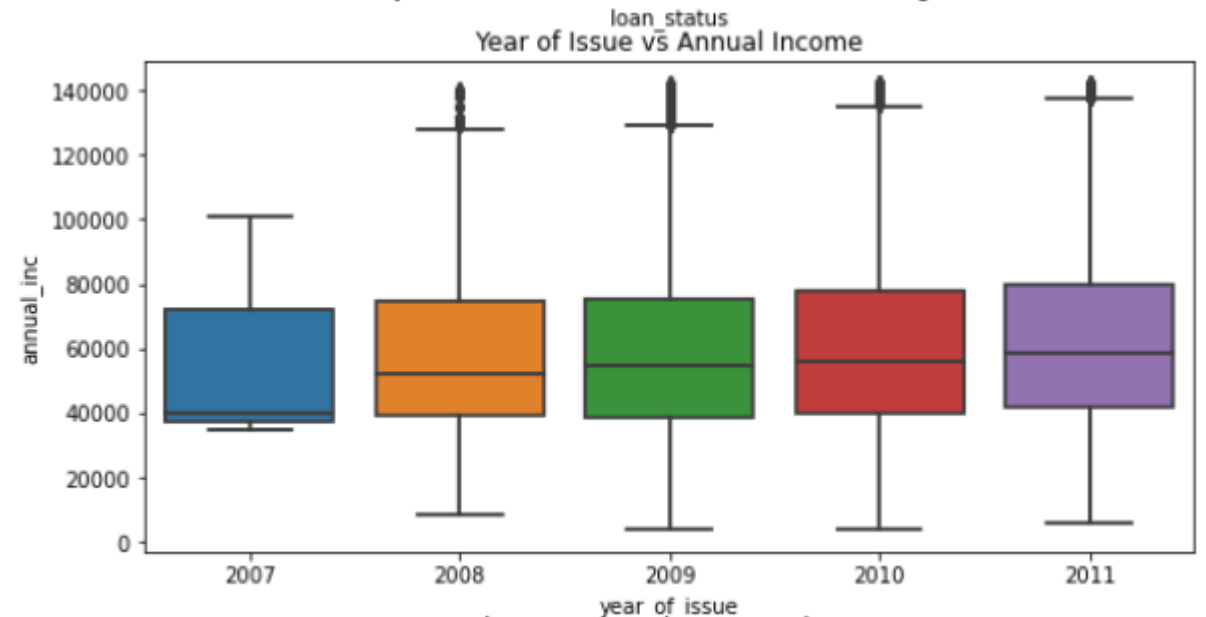
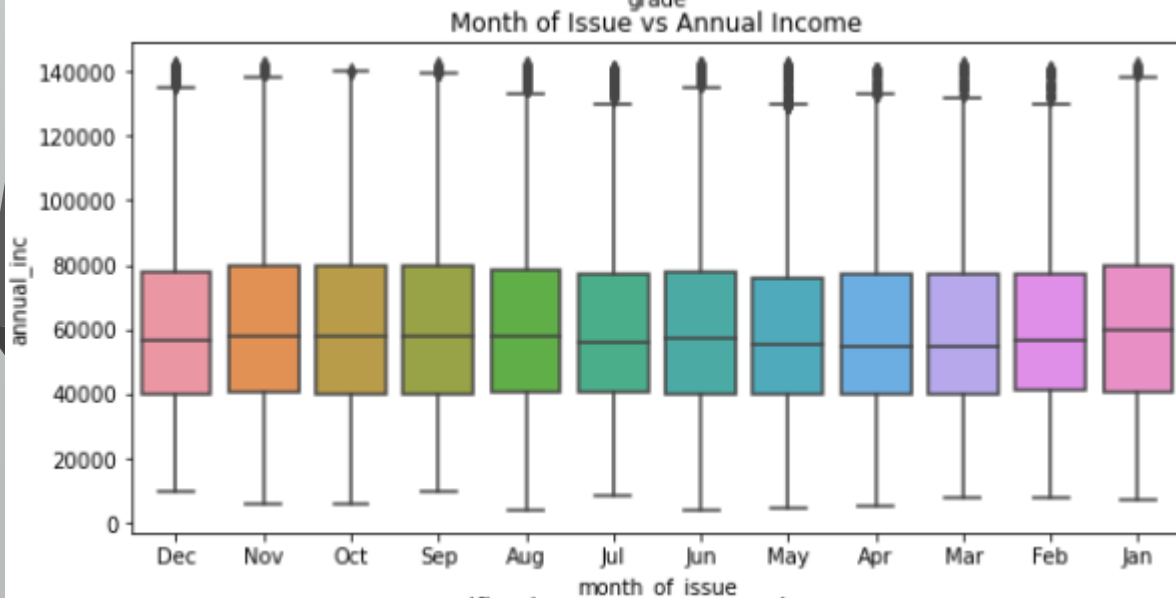
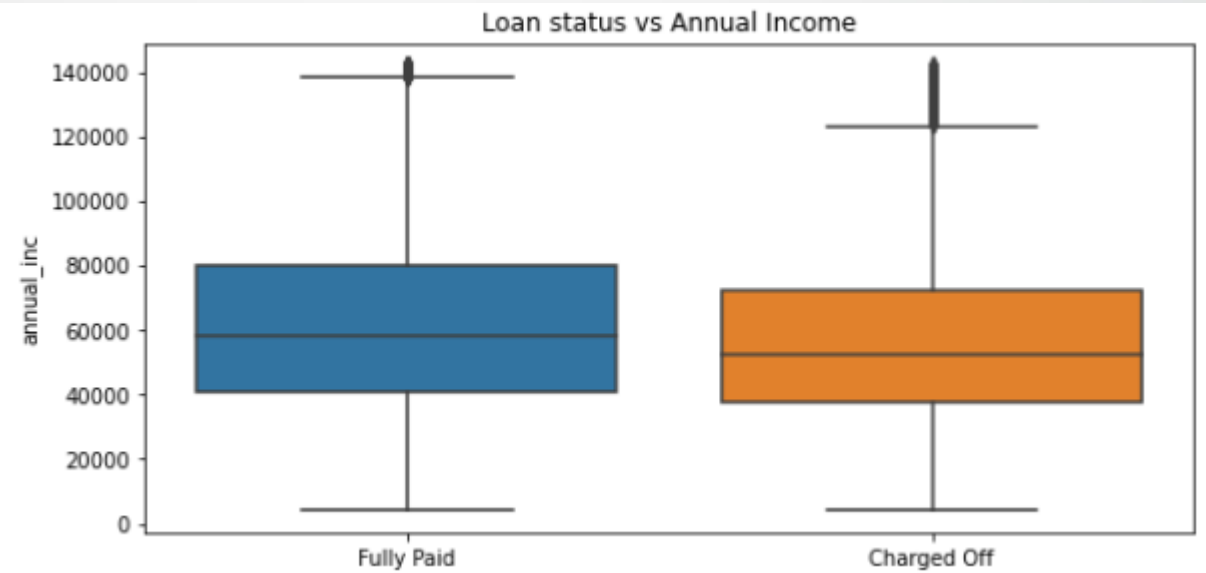
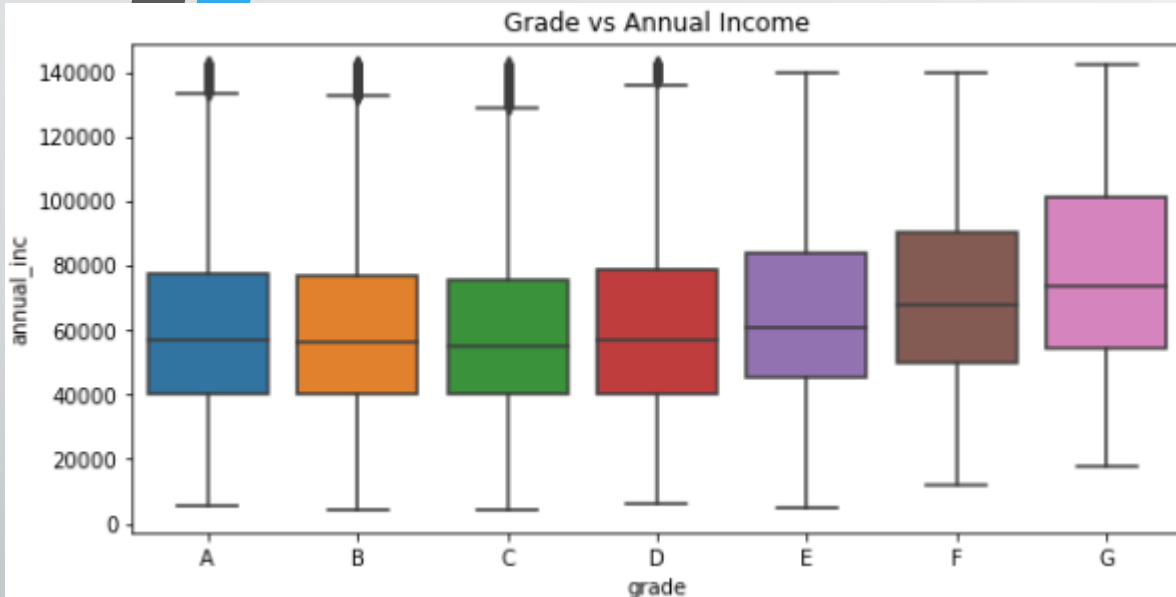


# Analysing Loan Amount with every Variable

## OBSERVATIONS:

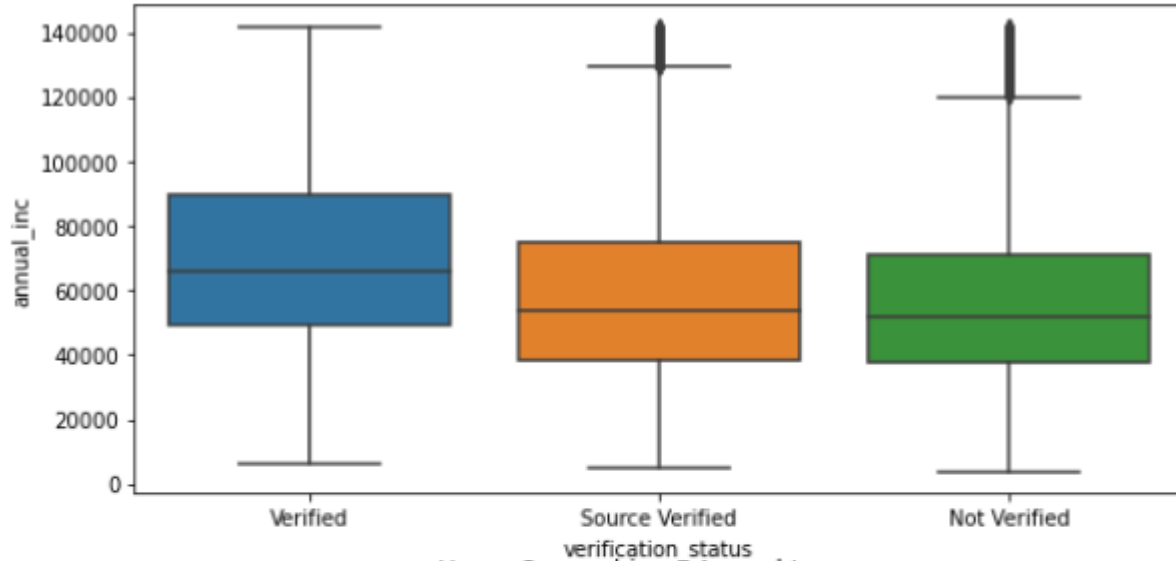
- The Grades G, F and E have higher loan amount.
- The loan amount has almost remained same through the years 2007 to 2011.
- The *Verified* borrowers have more loan amount than *Source Verified* and *Not Verified*.
- The loan amount almost remains same throughout a year (over the months).
- The Employment length of 10+ years have most loan amount. However, the average is same for almost all the employment length.
- Home Ownership does not have much effect on the loan amount, and it almost remains same throughout.
- Borrowers who are taking loan for 60 months are taking more loan amount, than with 36 months.
- Highest loan amount is needed by the small business. However, there are others purposes like house, credit card and debt consolidation which are high as well.

# Analysing Annual Income with every Variable

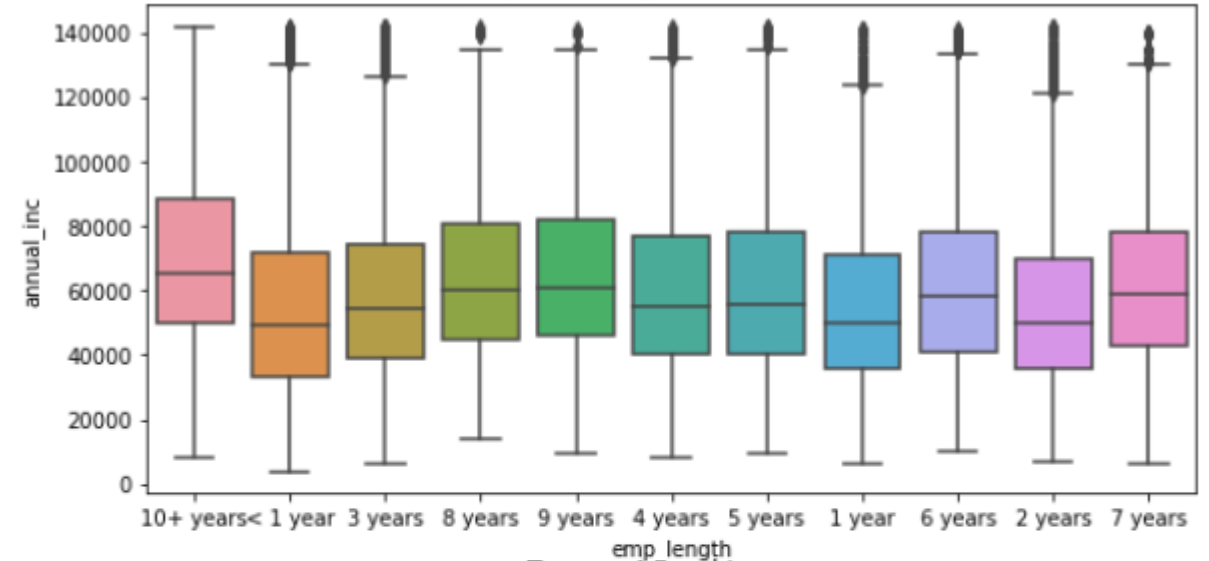


# Analysing Annual Income with every Variable

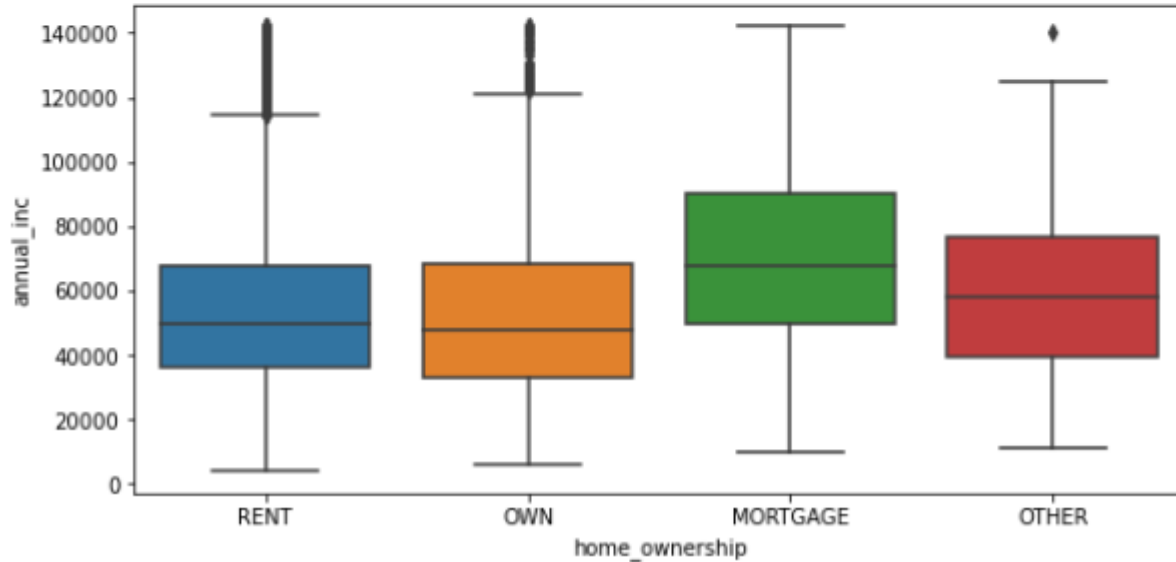
Verification Status v/s Annual Income



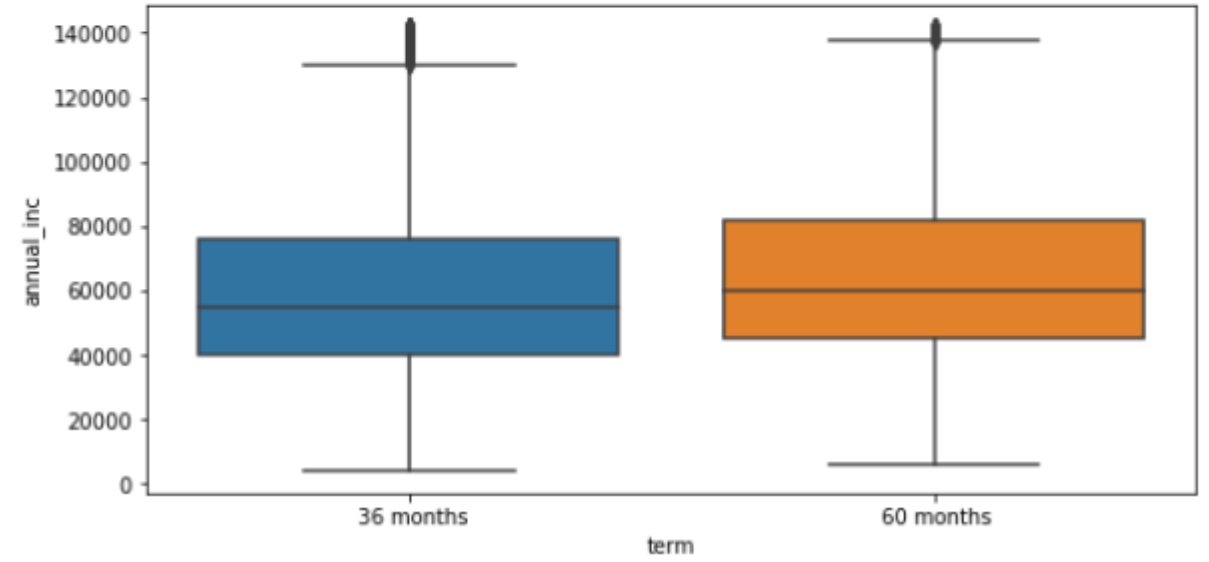
Employment Length vs Annual Income



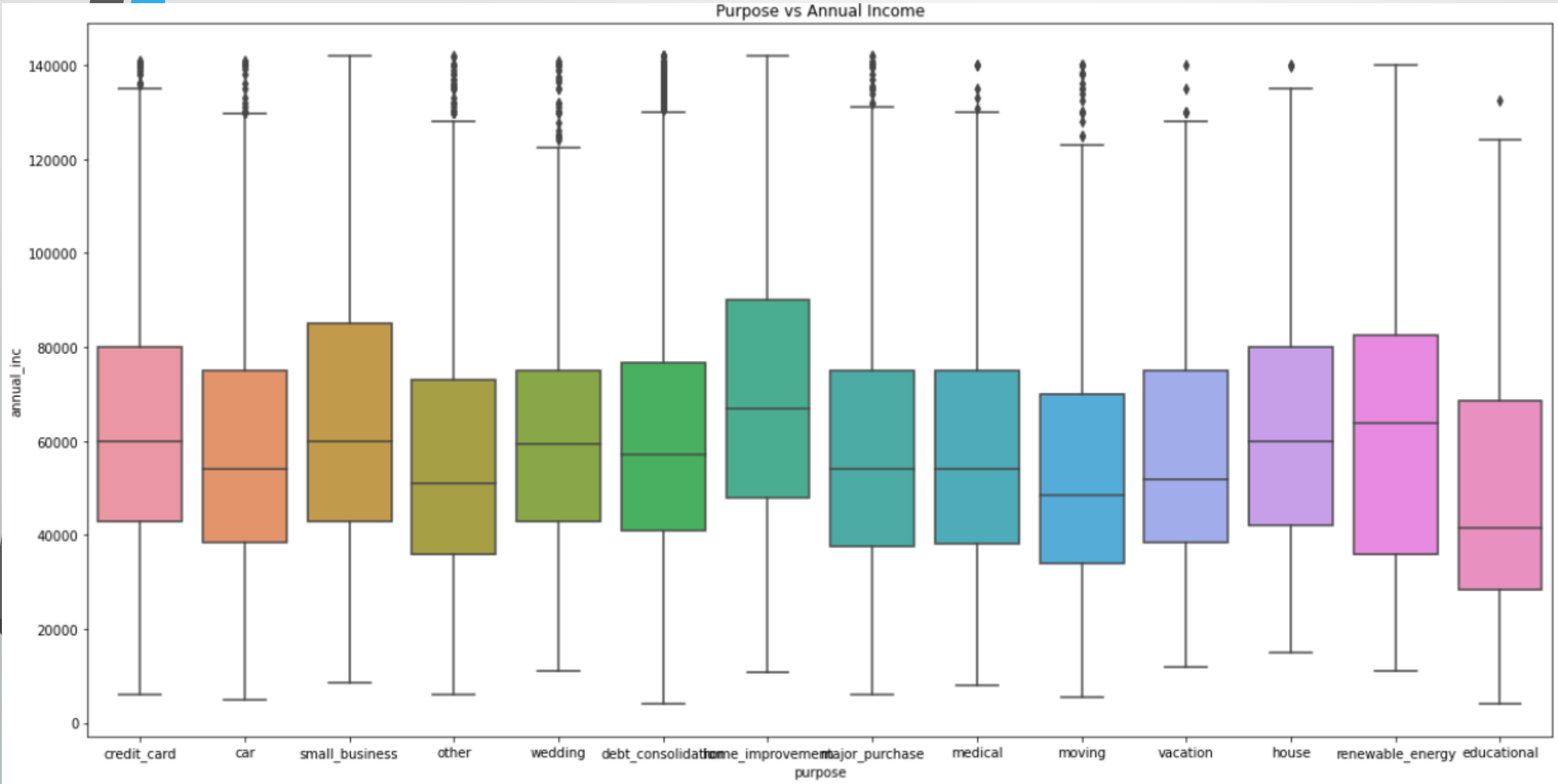
Home Ownership v/s Annual Income



Term vs Annual Income



# Analysing Annual Income with every Variable

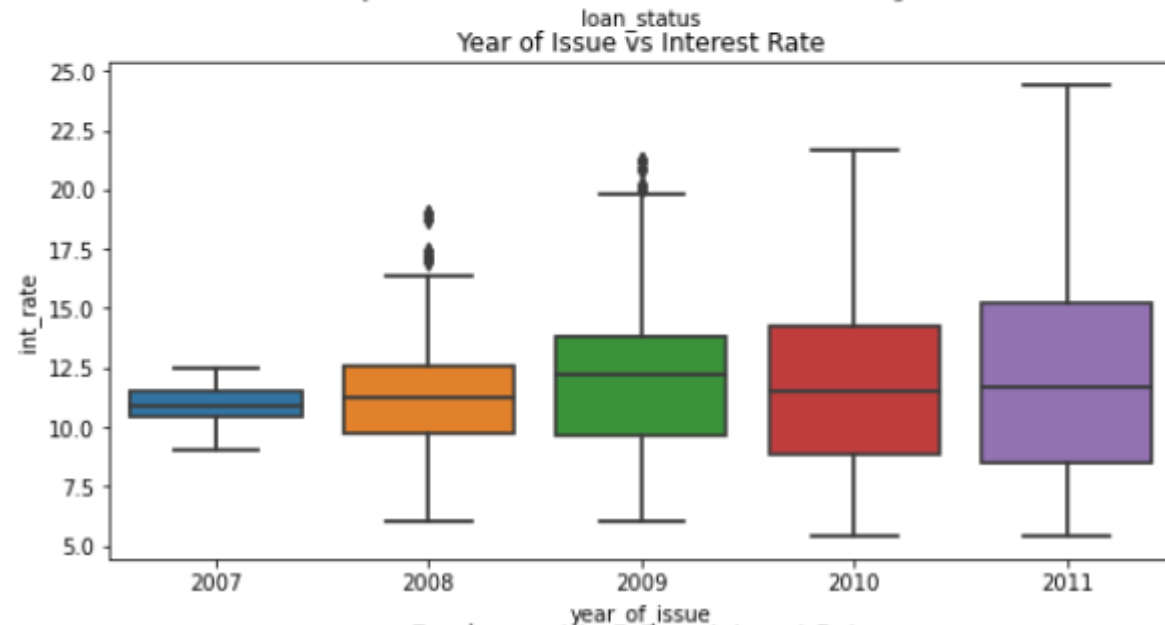
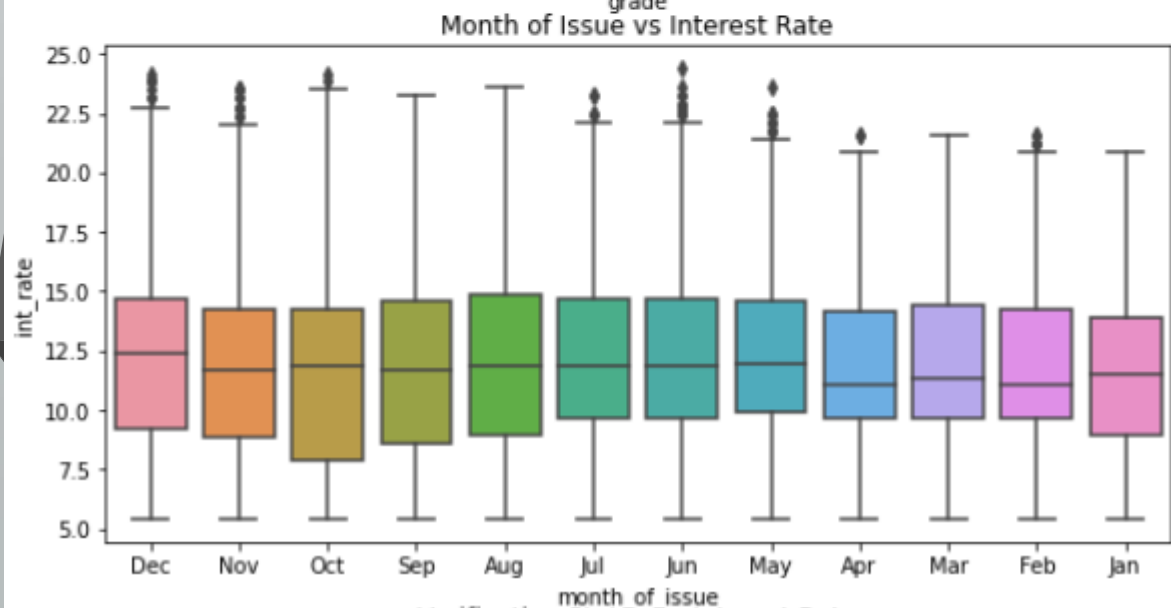
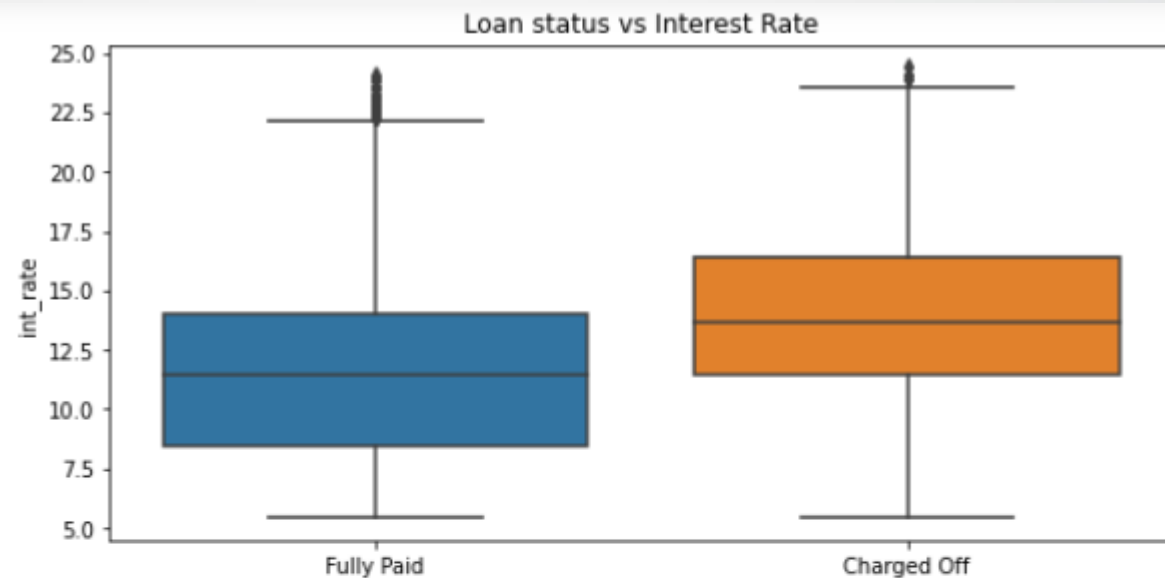
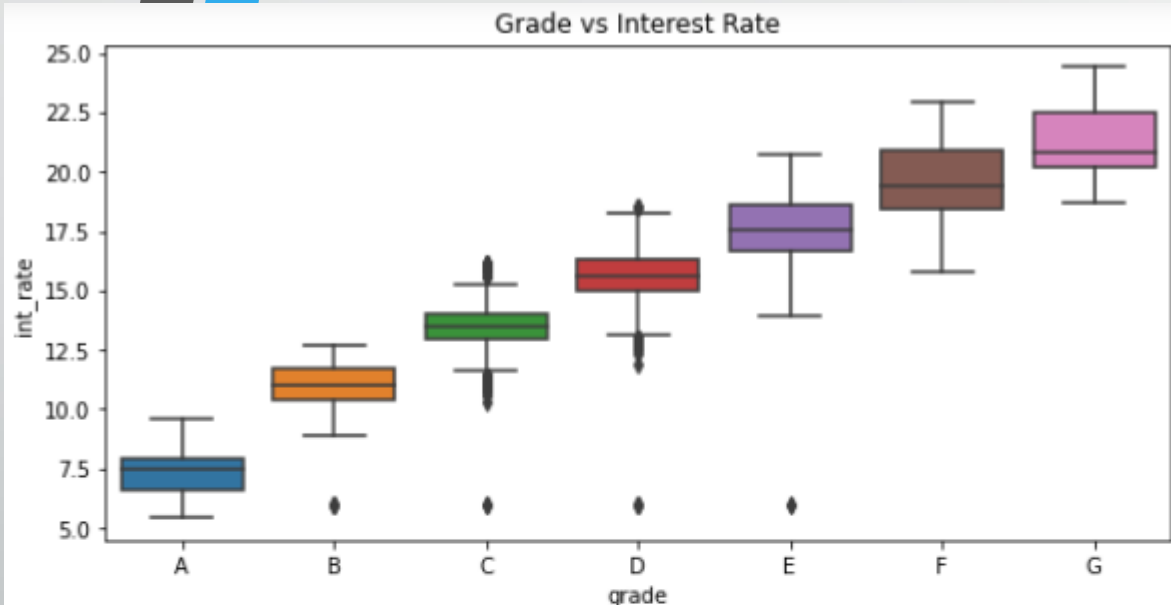


# Analysing Annual Income with every Variable

## **OBSERVATIONS:**

- Grade G has the highest annual income.
- Fully paid loan borrowers have slightly higher income than charged off borrowers.
- Month and year have almost no effect on annual income.
- Verified borrowers have higher annual income compared to others.
- Annual income increases as employment duration increases.
- Borrowers with Home Ownership Mortgaged. have most annual income.
- Borrowers with 60 months loan period have slightly more annual income.
- Borrowers with purpose Home Improvement has higher income, Followed by Small Business.

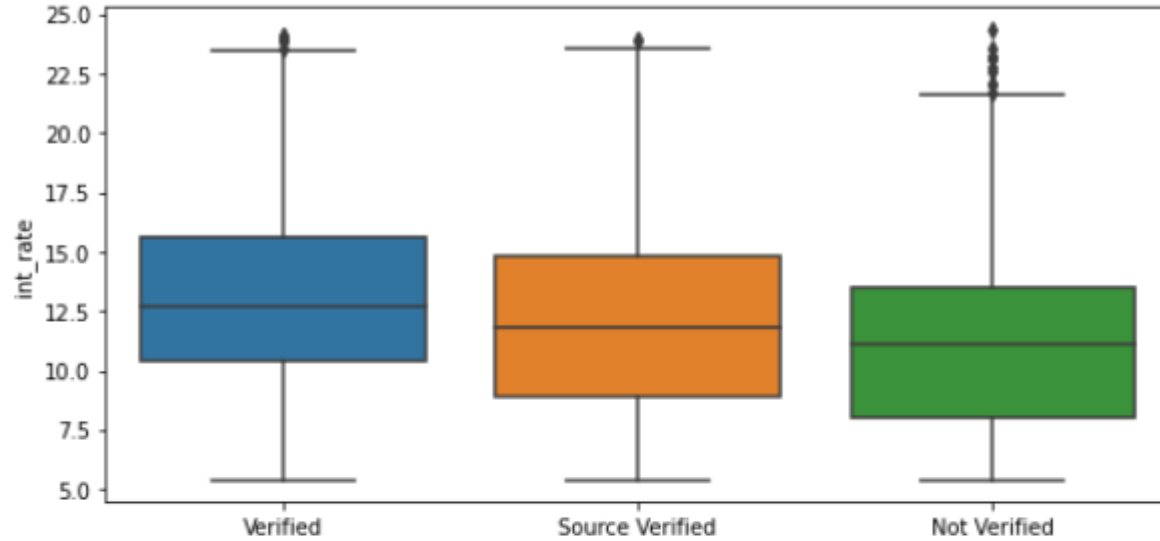
# Analysing Interest Rate with every Variable



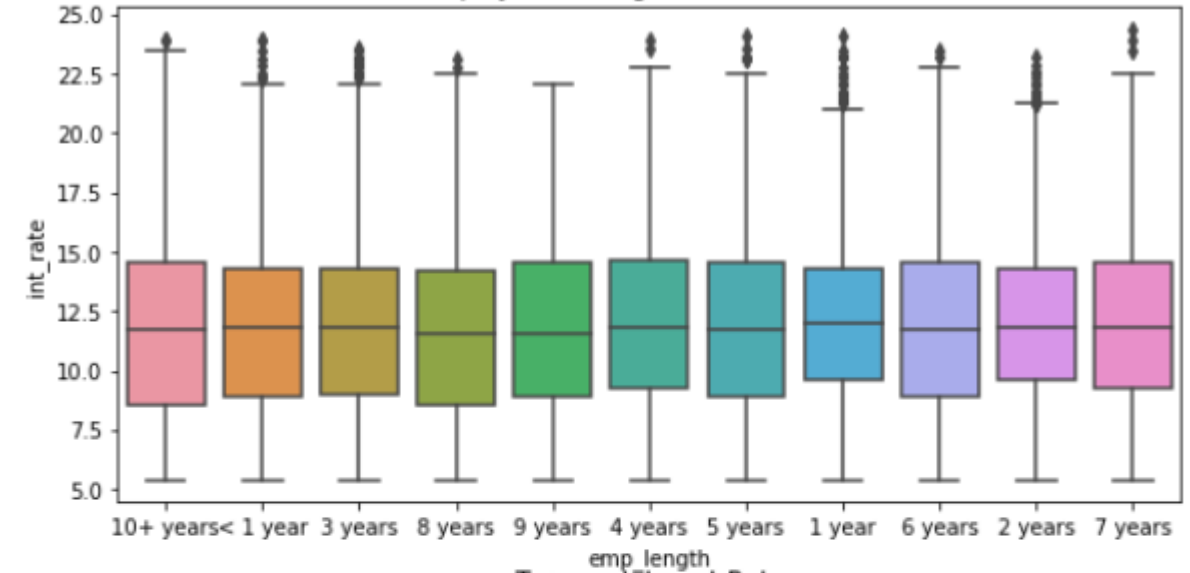


# Analysing Interest Rate with every Variable

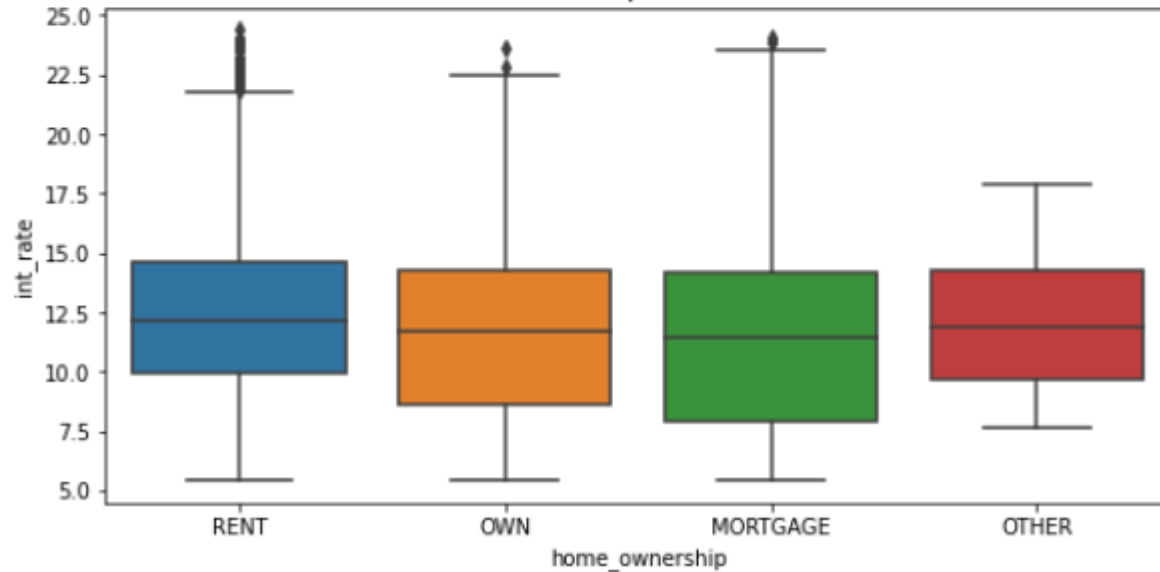
Verification Status vs Interest Rate



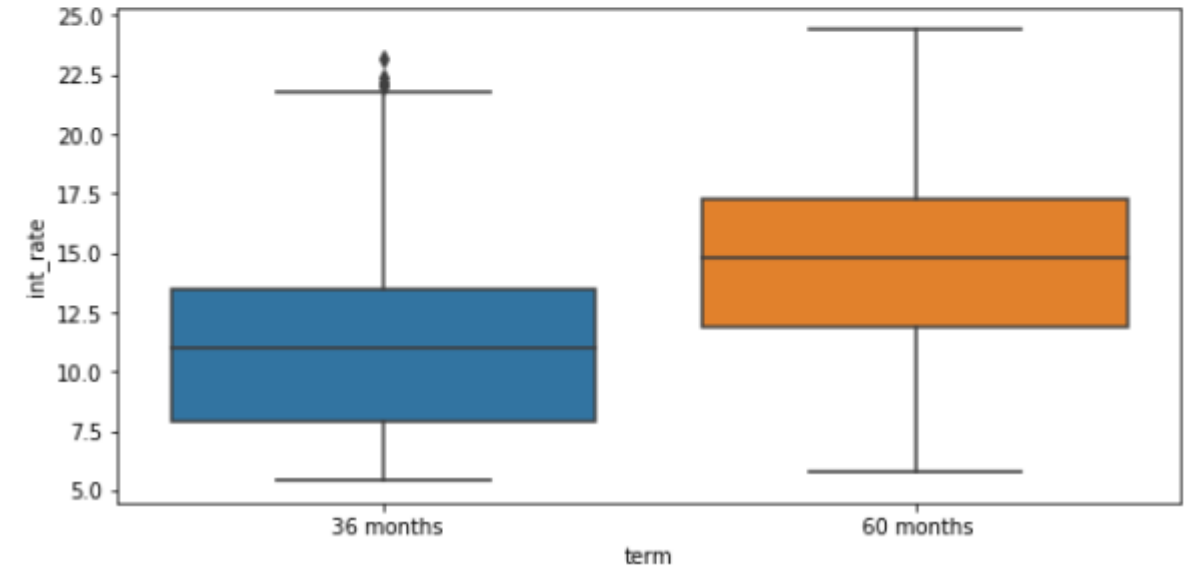
Employment Length vs Interest Rate



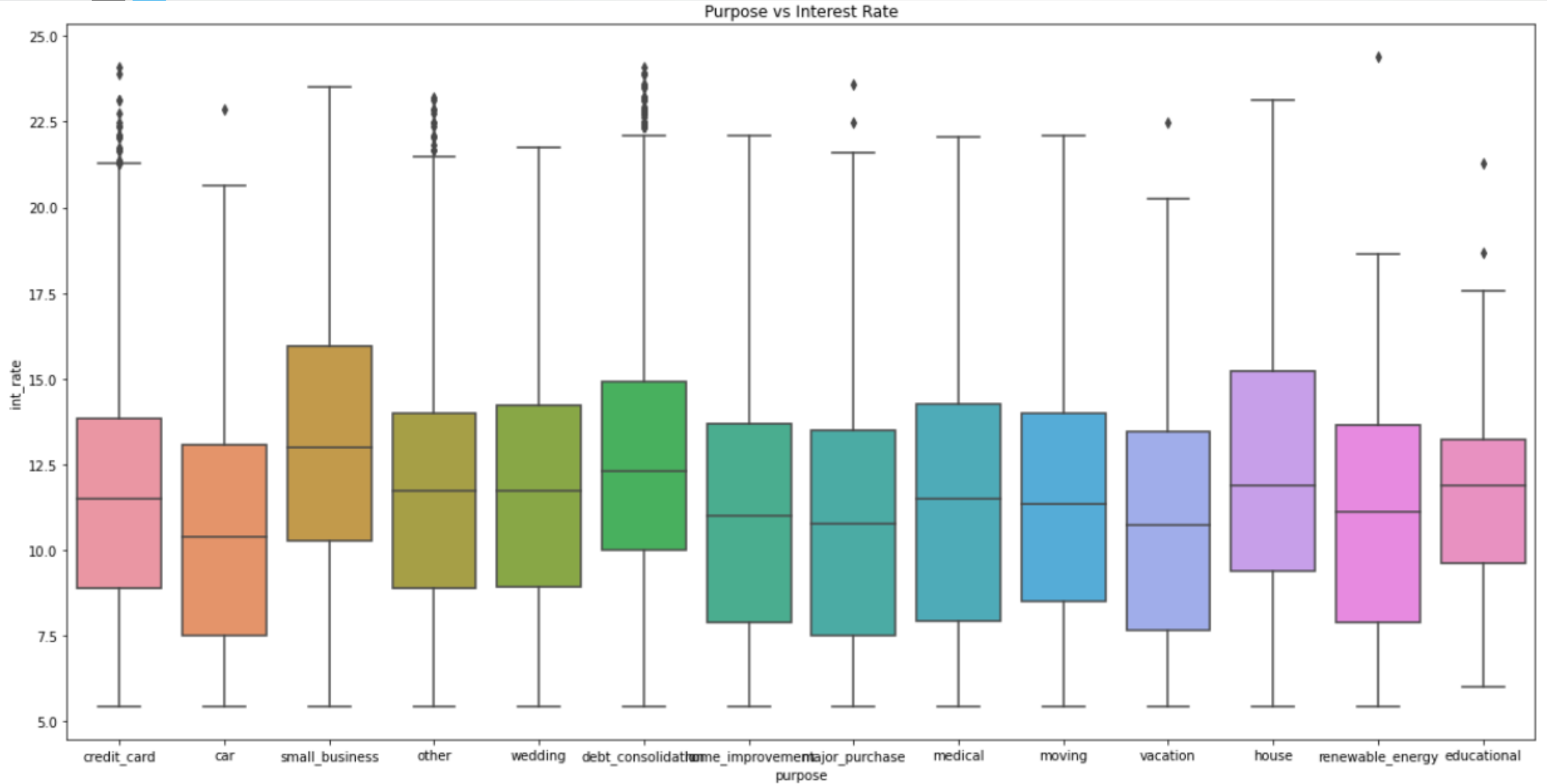
Home Ownership vs Interest Rate



Term vs Interest Rate



# Analysing Interest Rate with every Variable



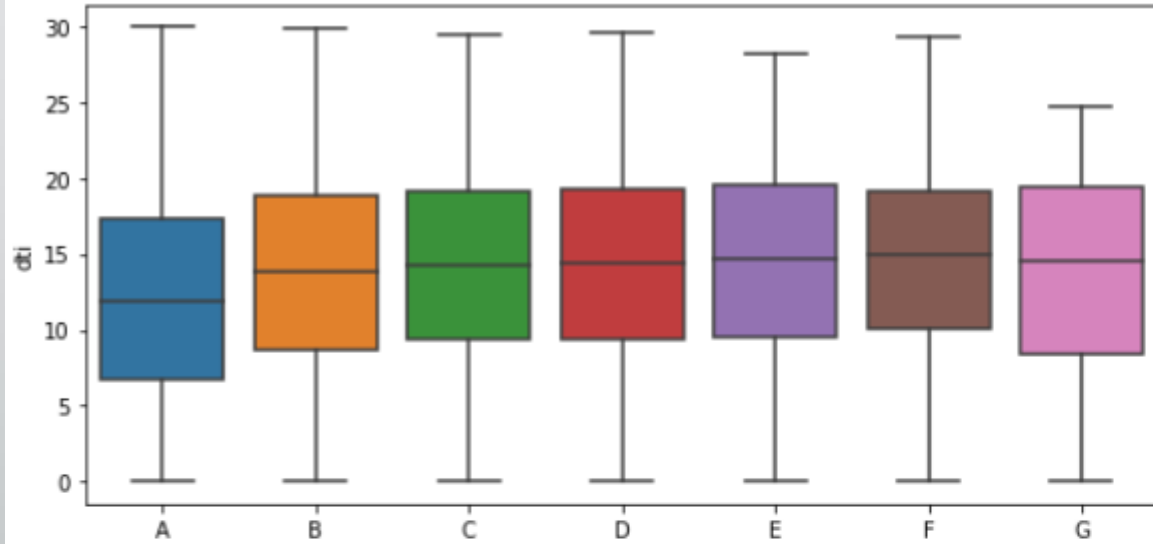
# Analysing Interest Rate with every Variable

## OBSERVATIONS:

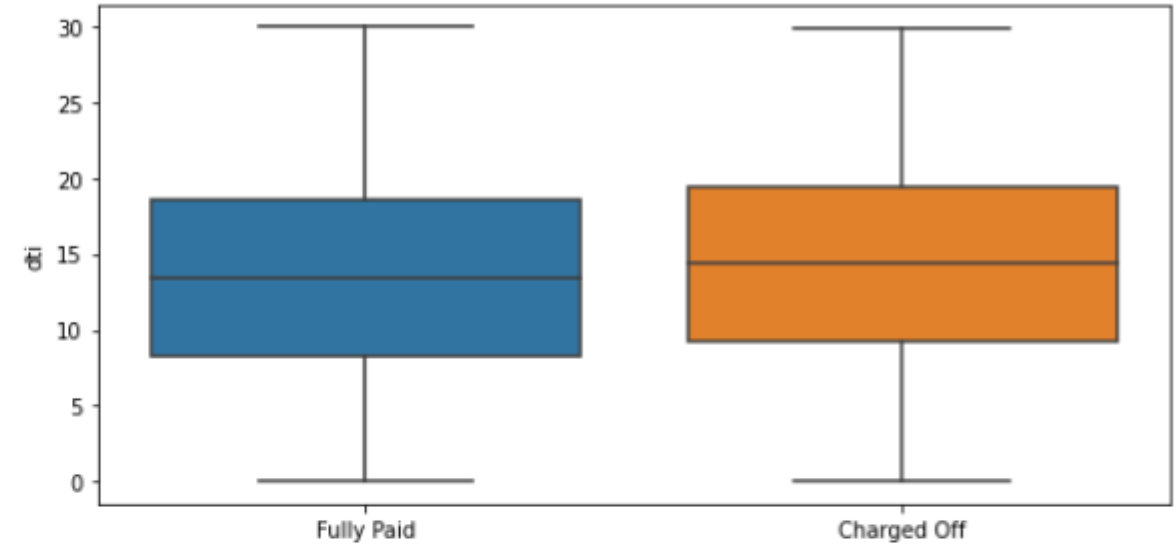
- Grade G pays the highest interest rate and the rate then reduces as the grade reduces.
- Fully paid loan borrowers pay less interest than charged off borrowers.
- Month and employment length have almost no effect on annual income. (Median is almost the same)
- Verified borrowers have higher interest rate compared to others.
- As the year goes on, the interest range is becoming more diverse.
- Borrowers who are in rent have a higher interest rate. The borrower with own and mortgage home are getting loan with slightly less percentage.
- Borrowers with 60 months loan period have significantly higher interest rate.
- Small business, Debt consolidation and House Loan are getting loan at high interest rate than others.

# Analysing DTI with every Variable

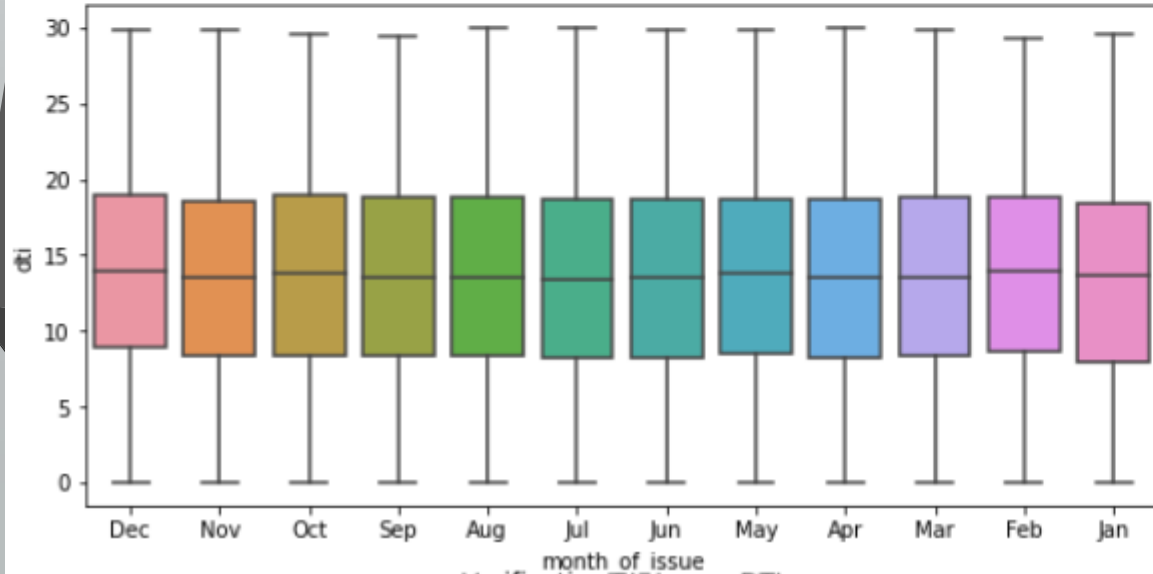
Grade vs DTI



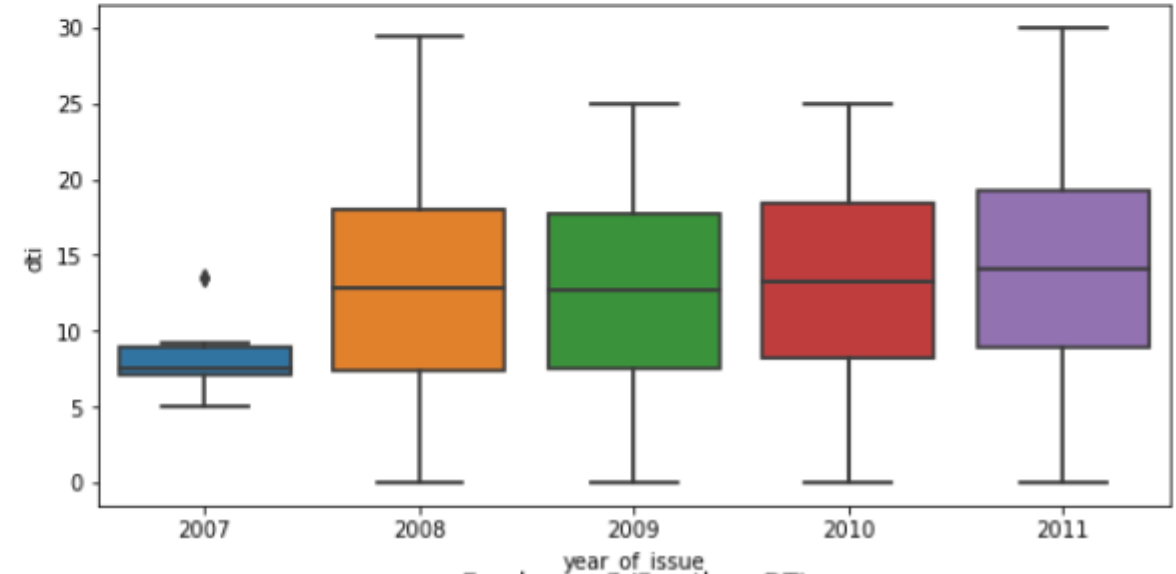
Loan status vs DTI



Month of Issue vs DTI

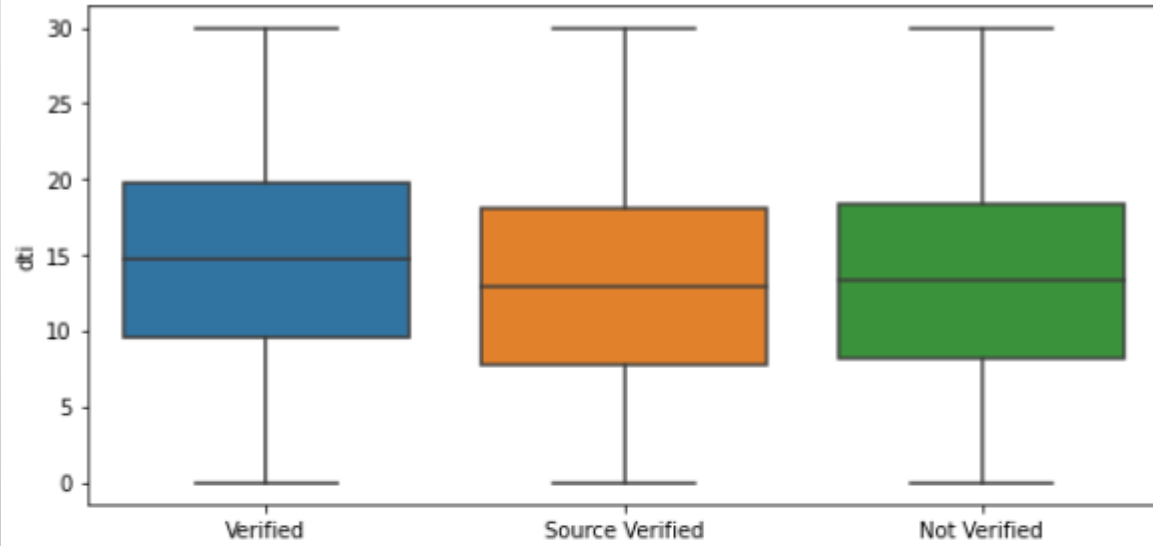


Year of Issue vs DTI

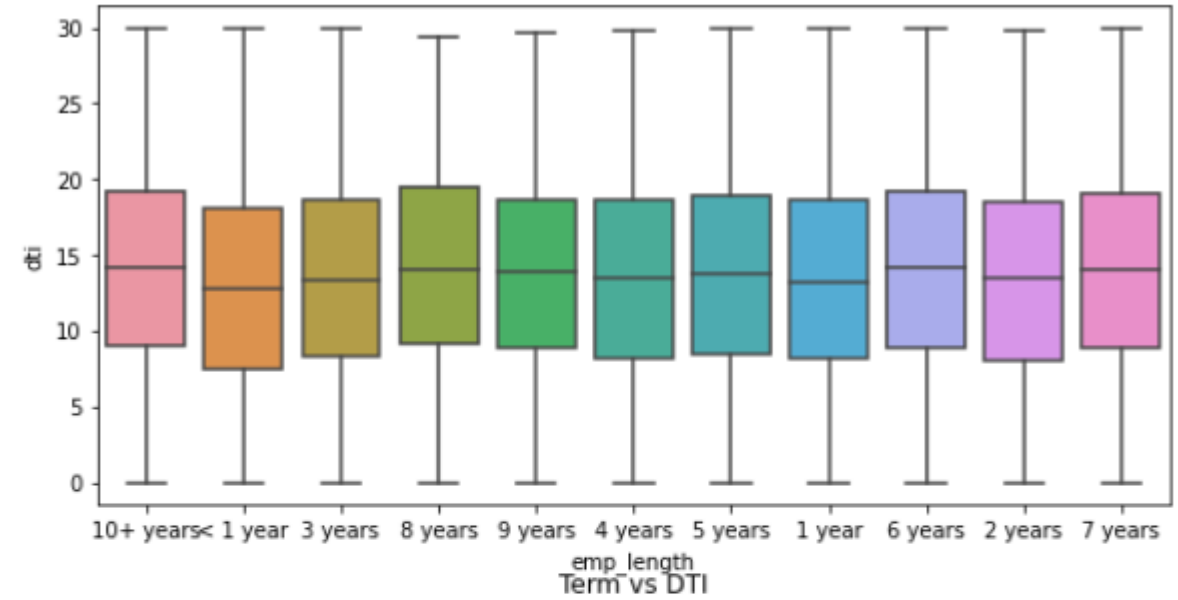


# Analysing DTI with every Variable

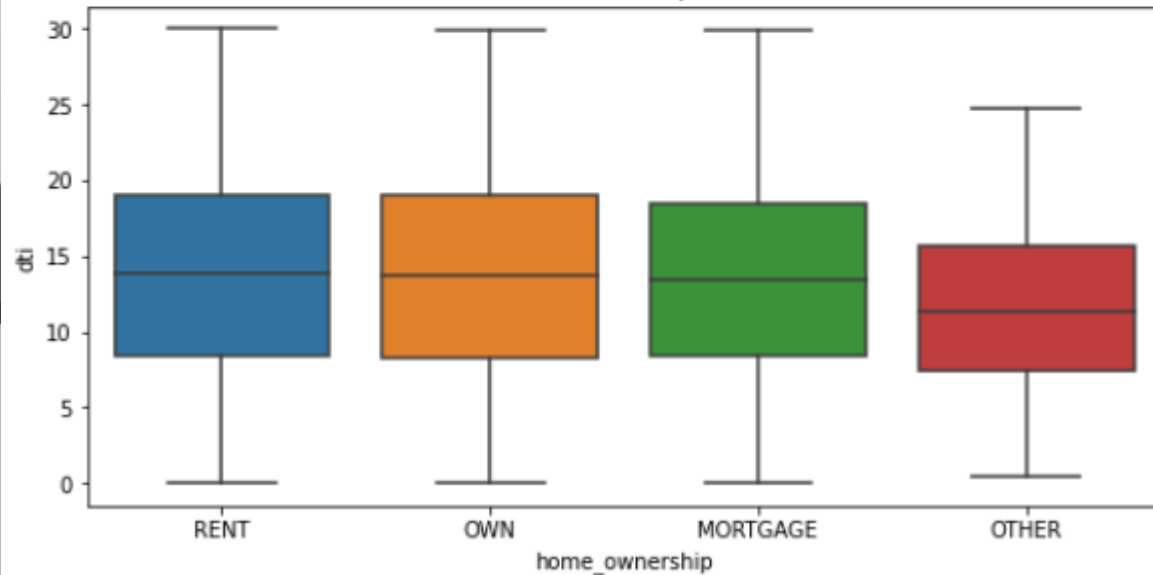
Verification Status vs DTI



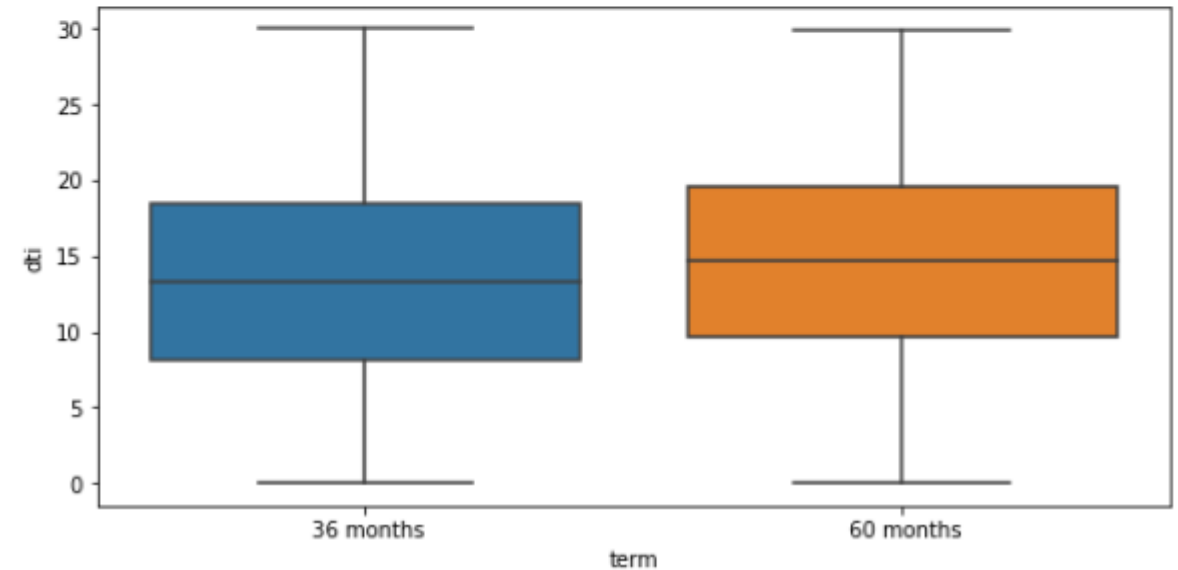
Employment Length vs DTI



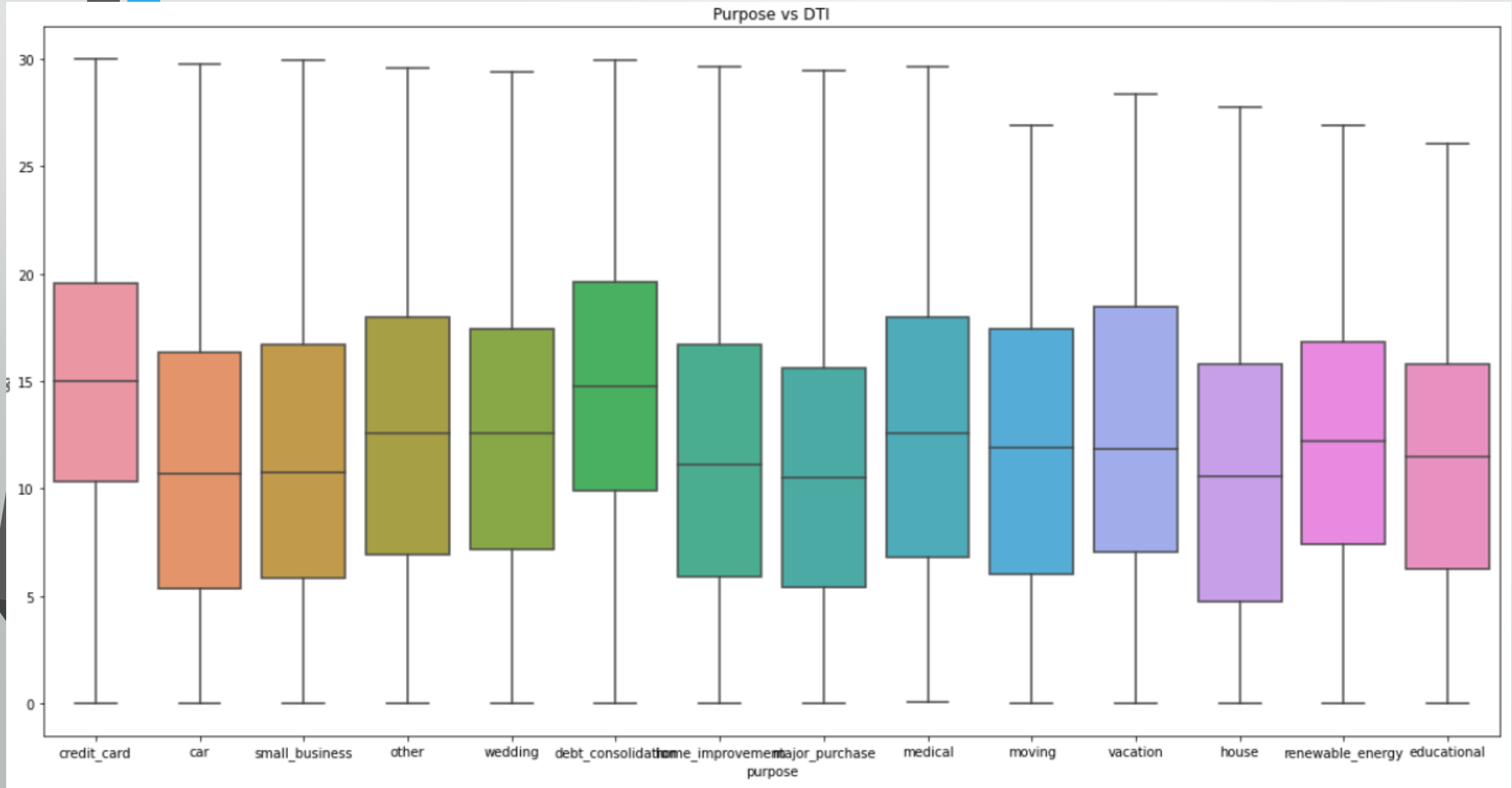
Home Ownership vs DTI



Term vs DTI



# Analysing DTI with every Variable

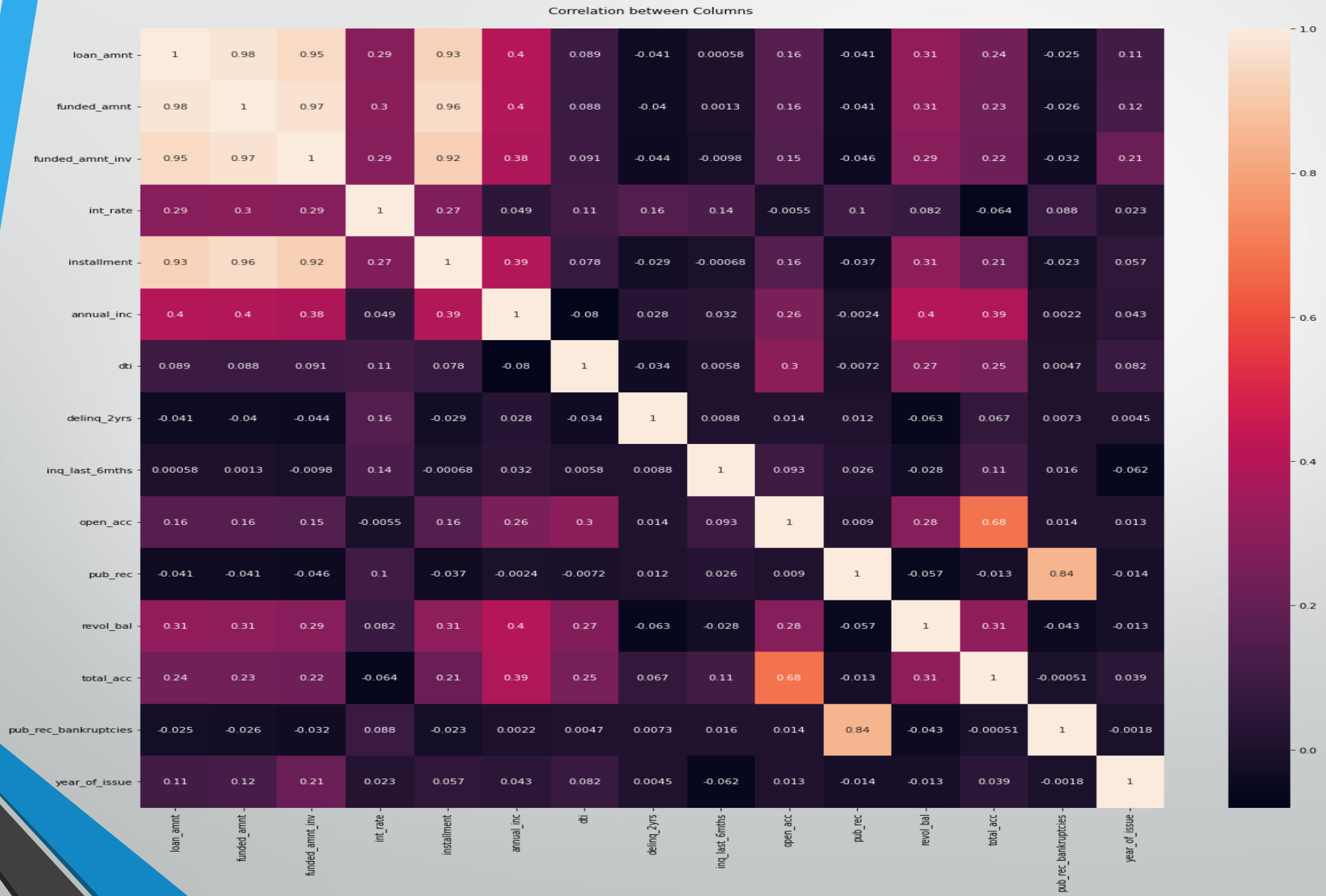


# Analysing DTI with every Variable

## OBSERVATIONS:

- The borrower with 60 months tenure has higher DTI.
- The verified borrowers have higher DTI.
- The charger off borrowers have higher DTI.
- DTI Ratio increases as the grade decreases.
- Employee length does not affect DTI.
- Borrower with rent and own house ownership have higher DTI.
- Month does not have any effect on the DTI.
- DTI increases gradually with the years.
- Debt consolidation, credit card and vacation have higher DTI.

# Correlation





# Correlation

## **OBSERVATIONS:**

- Interest Rate and total account balance is negatively correlated.
- installment, funded\_amnt, funded\_amnt\_inv and loan amount are highly correlated.
- DTI and Annual Income are correlated negatively.
- So, if Annual\_inc is low, the debt is high.

# Recommendations

**From above analysis we can say that, there is a high probability of defaulting the loan when:**

1. The borrower belongs to a lower grade, like F or G; and especially the subgrade F5, G3, G5.
2. The loan amount is in the range 0 to 21000.
3. The annual income of borrower is between 30000 and 60000.
4. The borrower has a rented home and the loan is for the purpose of debt consolidation.
5. The loan period is of 60 months and the loan status is 'verified'.
6. The interest rate is in the range 15-20%.
7. The purpose of the loan is small business.