

Lending club case study

- BY

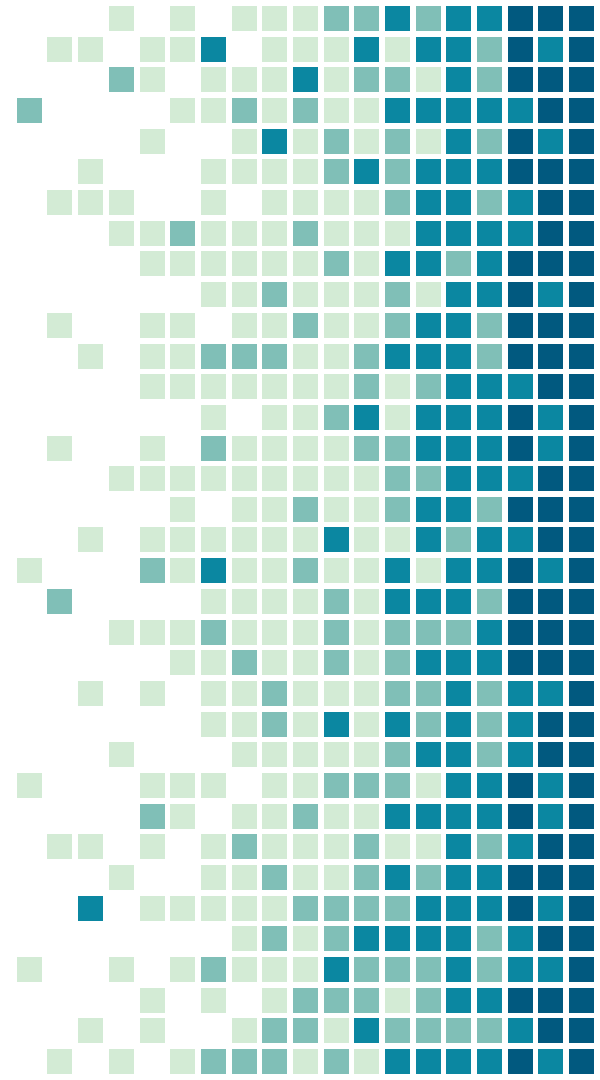
Nikhileshwar Reddy B
nikhil1234nikhi@gmail.com

BUSINESS UNDERSTANDING

- You work for a **consumer finance company** which specializes in lending various types of loans to urban customers.
- When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two **types of risks** are associated with the bank's decision:
 - ❑ If the applicant is **likely to repay the loan**, then not approving the loan results in a **loss of business** to the company.
 - ❑ If the applicant is **not likely to repay the loan**, i.e. he/she is likely to default, then approving the loan may lead to a **financial loss** for the company.

PROCESS FOLLOWED

- The first step is to import the libraries that are required for the project.
- Then followed by cleaning rows and columns.
- Remove all the columns that have same repetitive value or Nulls for the entire column.
- There are some columns that are not required for analysis remove them accordingly.
- After removing the columns then the next step is renaming the metadata in understandable format.
- Check if any columns have nulls and try to fix it.
- Some of the nulls are ignored for better and understandable data
- Then change the data type according to the required analysis.



“

UNIVARIATE ANALYSIS

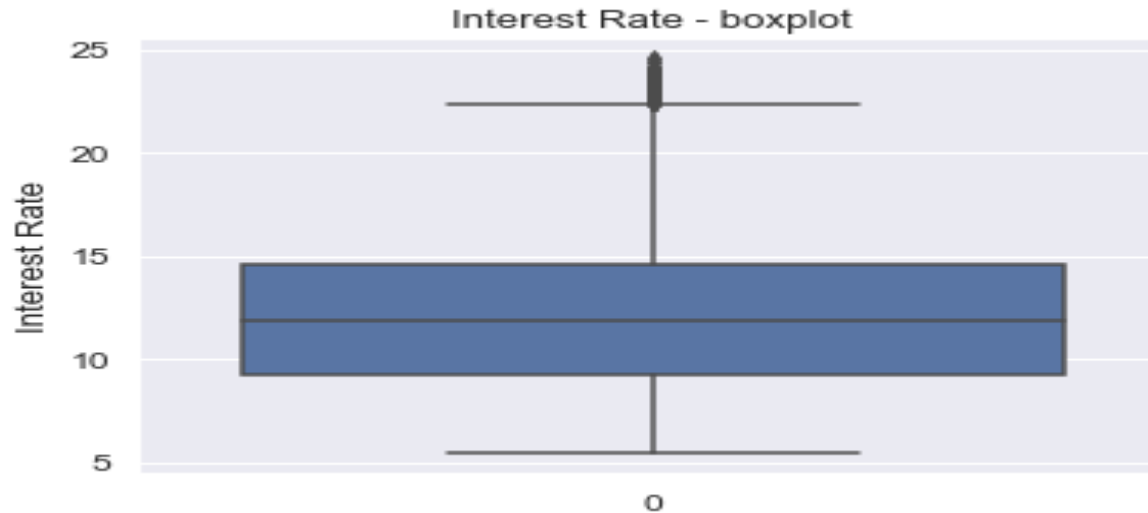
Observation :

- As we can see there are no outliers in this column
- most amount of loan applications lie in 5000-15000 range



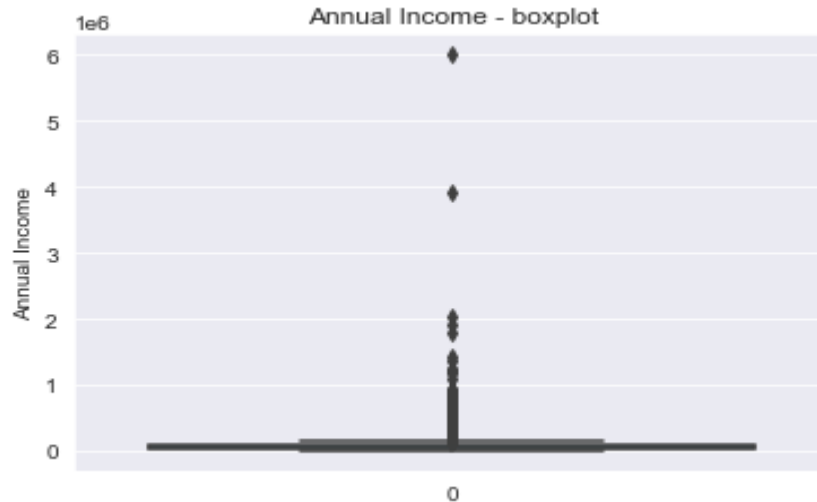
Observation :

- As we can see there are no outliers in this column.
- The average interest rate lies around 12% range.



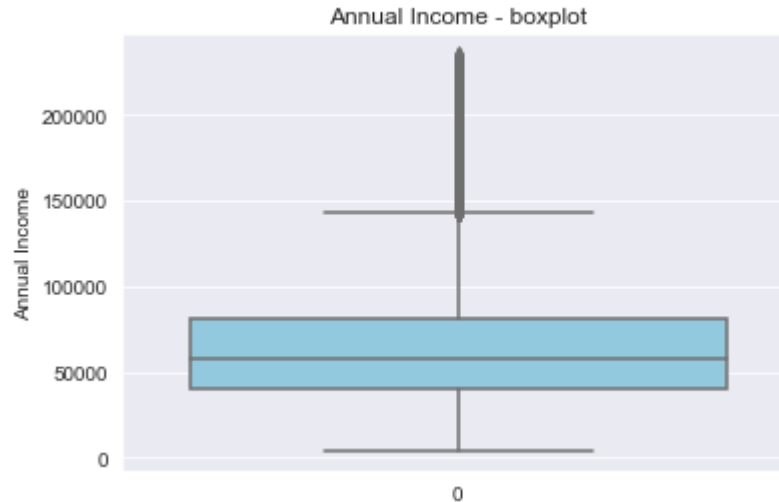
Observation :

- As we can see there are outliers in this column nearly above 5000000 mark.
- We have to remove the outliers by getting the right quantile percentage so that we don't lose data in this process.



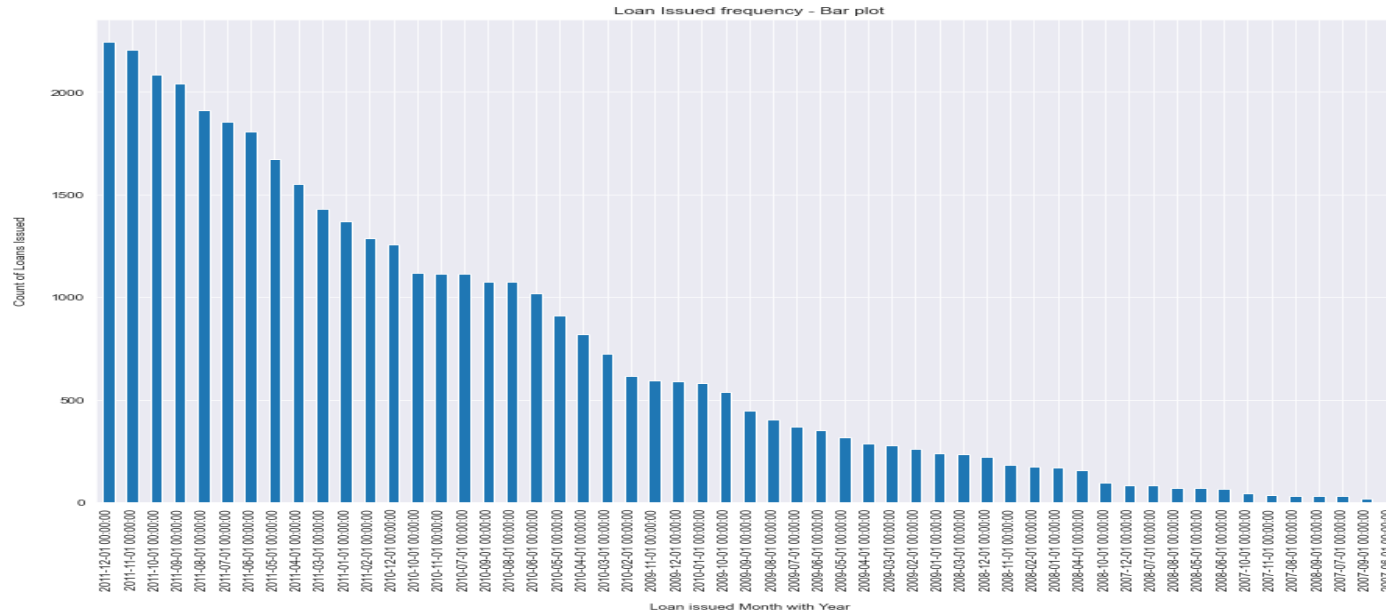
Observation :

- Now we removed the outliers by setting conditions and when we plot the graph again we don't get any outliers.
- As we can observe the average annual income range lies between 4000 - 8000.



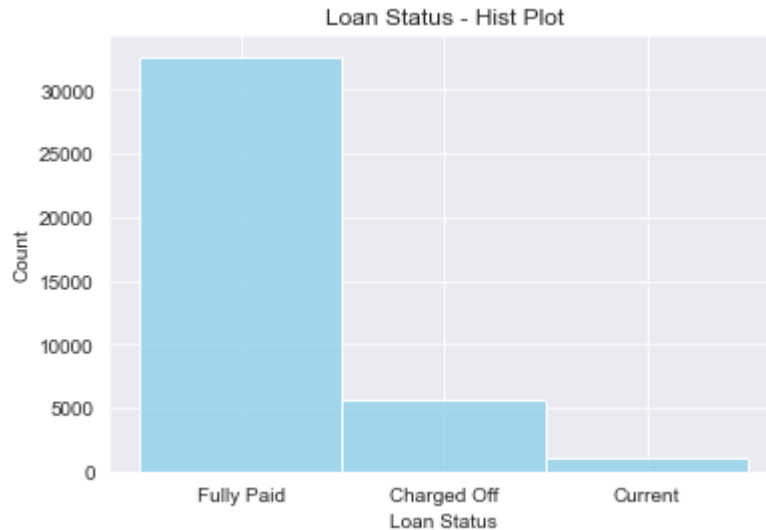
Observation :

- Here we plotted a graph between Loan issued month wise according to different years.
- There is gradual increase in the loan issued frequency over years.
- Peaking at December 2011 and least at June 2007



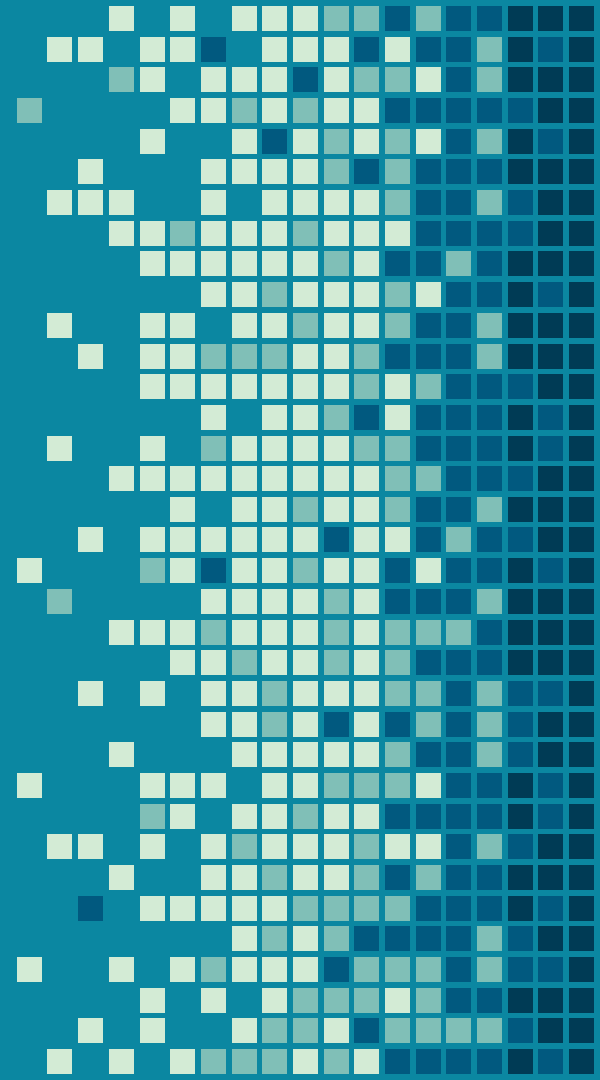
Observation :

- Here as we can observe that Fully paid applicants are more when compared with the other 2 categories.
- Detailed analysis on this can be done by using Segmented analysis.



“

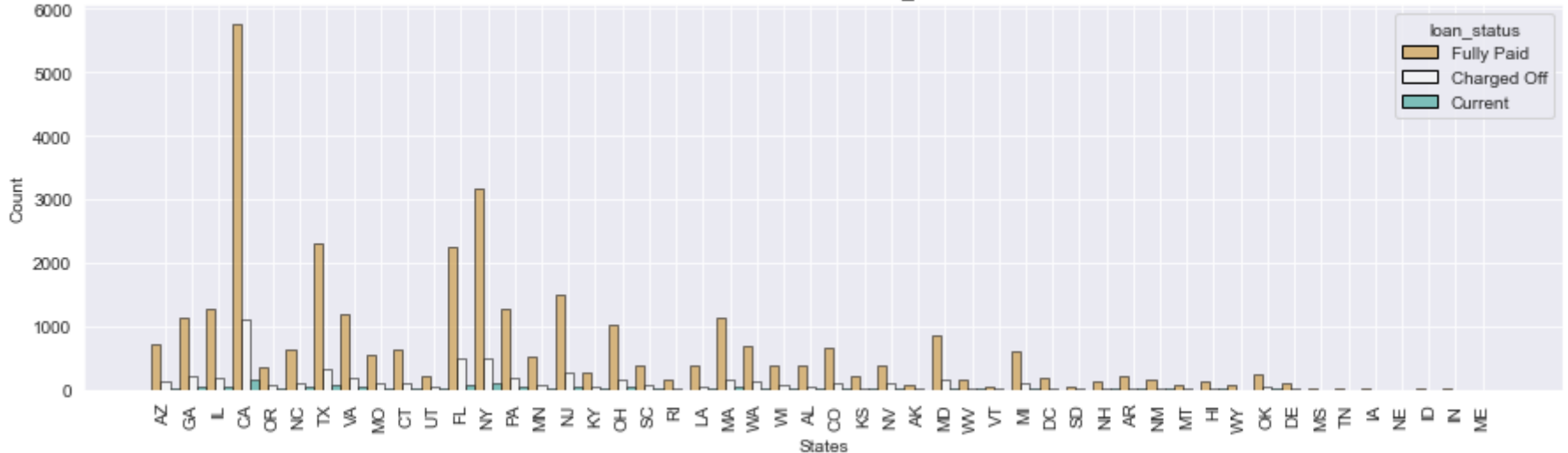
SEGMENTED UNIVARIATE ANALYSIS



Observation :

- Here we can see the Different states Loan status in different segments.
- People in California, Texas, Florida, New York and few other states comprises of most of the loans(Fully paid and charged off).

States vs Count of Loan_status



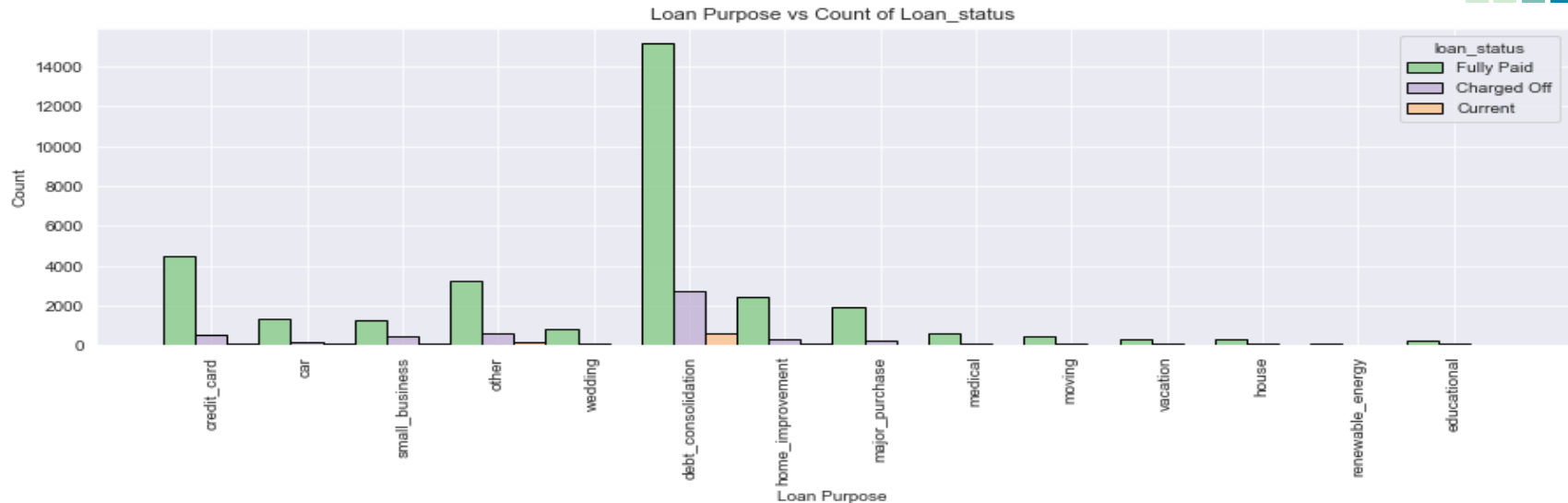
Observation :

- Here we can see the Terms of Loan status in different segments.
- There are no current loans in 36 months term applicants but there are more charged off applicants in 60 months when compared to 36 months.



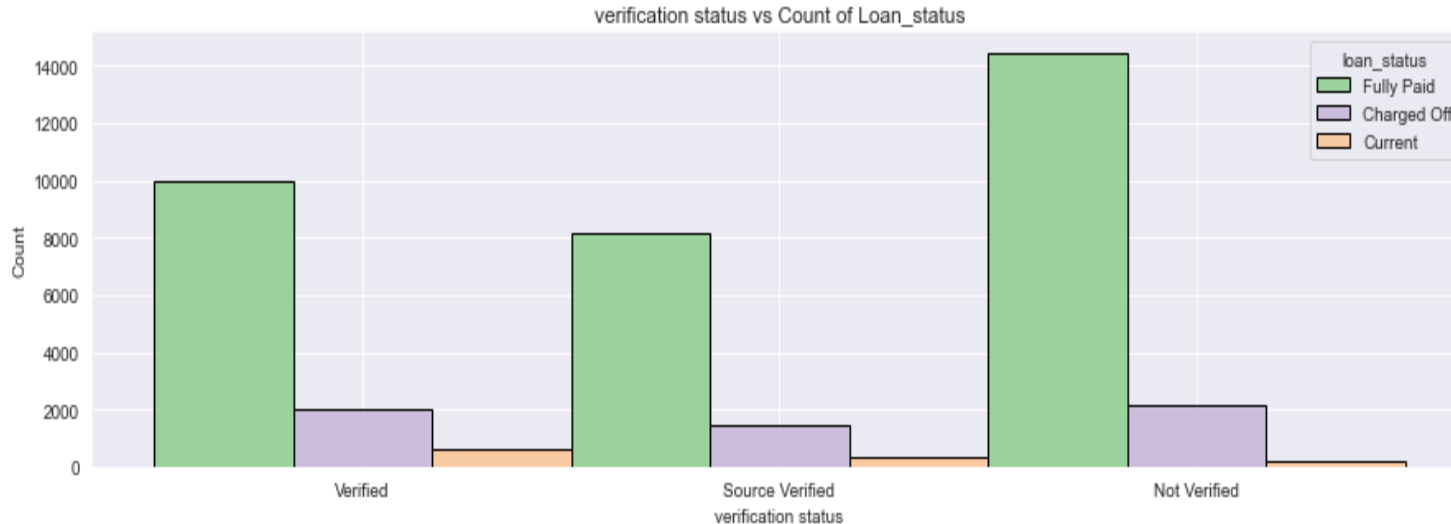
Observation :

- Here we can see the loan purpose of Loan status in different segments.
- Debt consolidation took most of the loans and have greater than 50% of Fully paid and charged off sector.



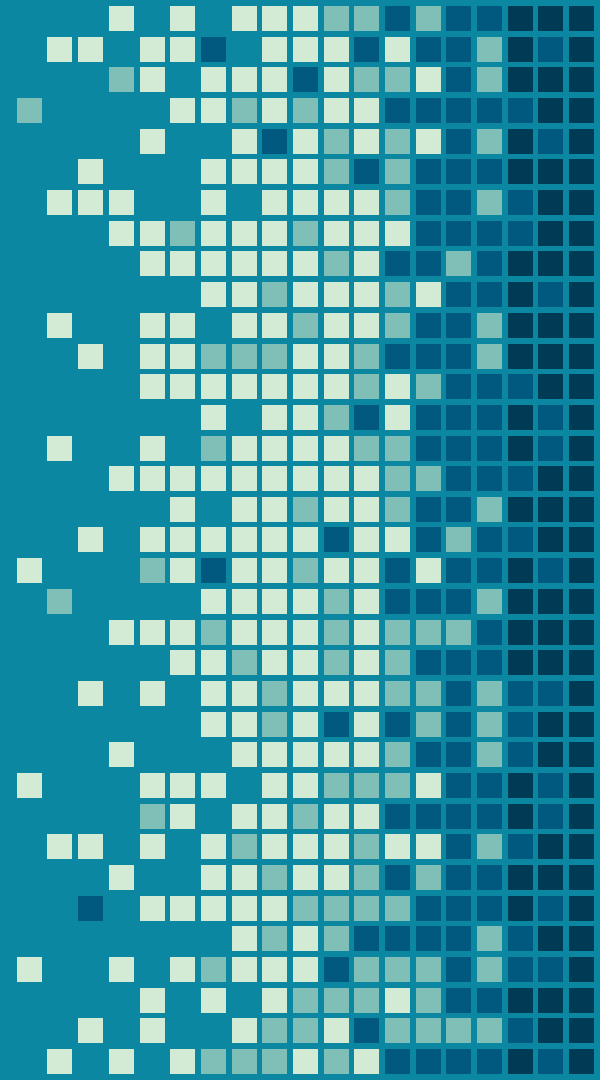
Observation :

- Here we can see the verification status of Loan status in different segments.
- Not verified people took most of the loans that are full paid but other 2 segments have a high charged off rate when compared.



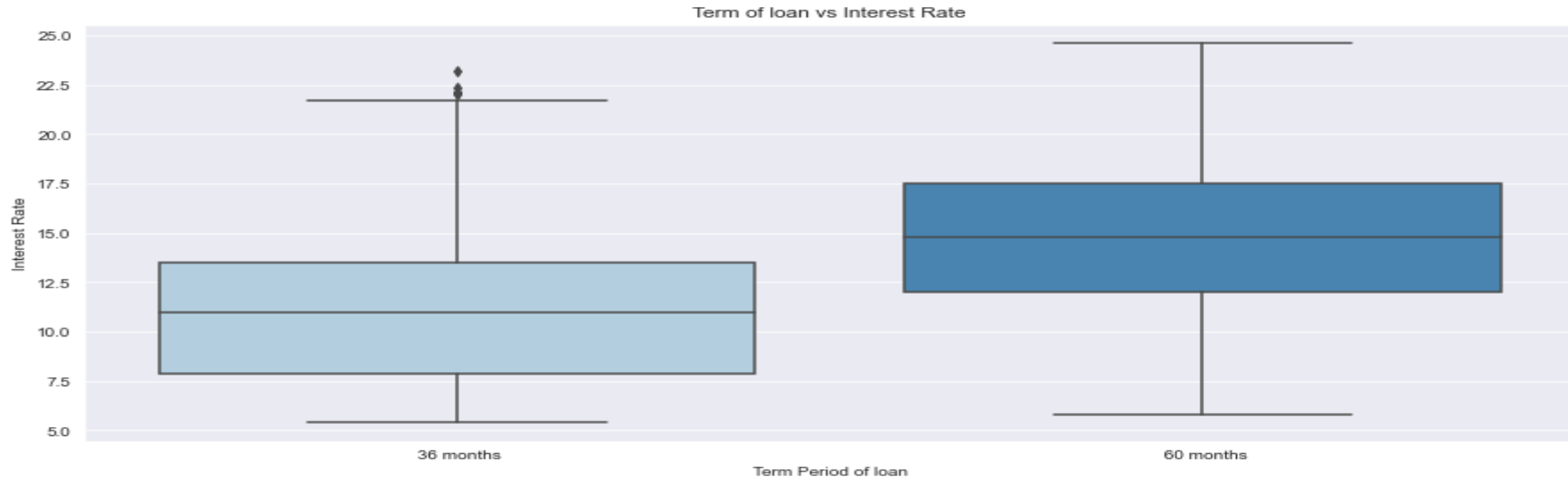
“

BI-VARIATE ANALYSIS



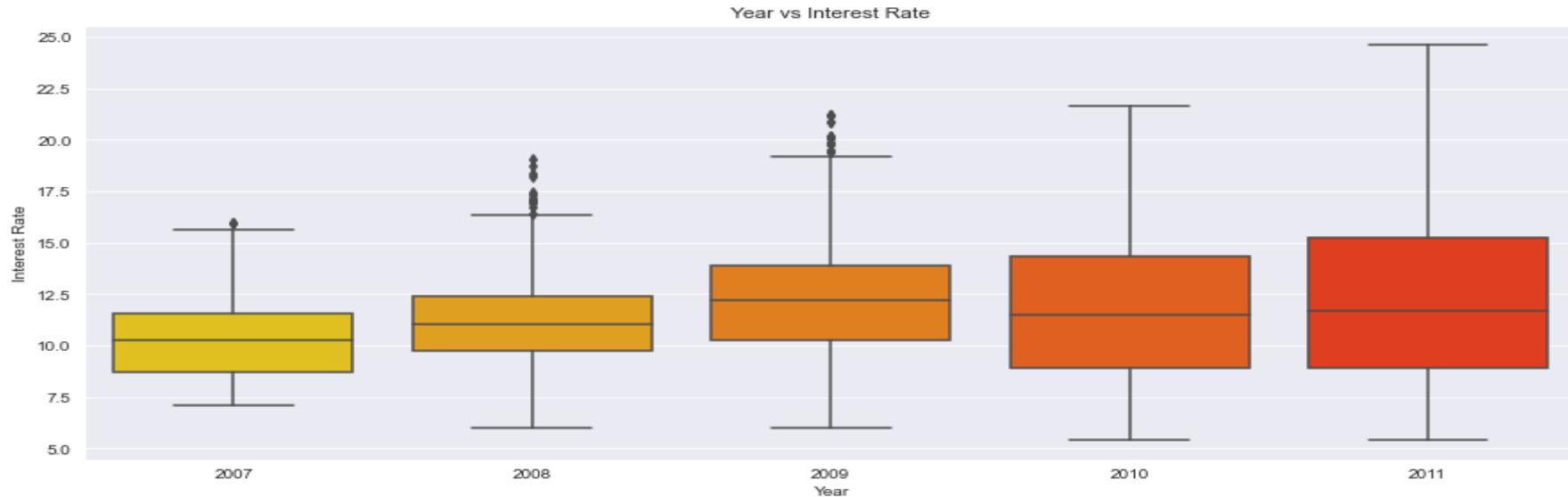
Observation :

- Here we can see Term of loan vs Interest rate.
- Average interest rate of 60 months is greater than 36 months.
- So there is a probability of more charged people here due to more interest rate.
- The interest rate varies heavily between 4%.



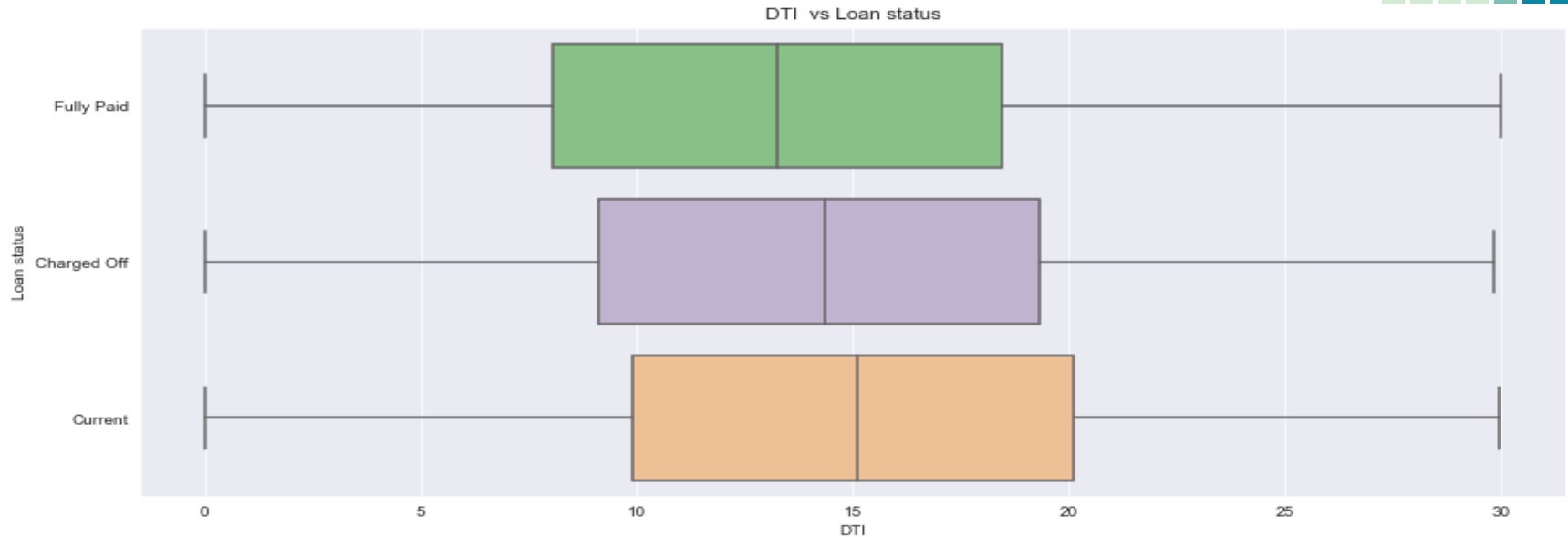
Observation :

- Here we can see Year of loan vs Interest rate.
- 2007 has the lowest avg interest rate and it gradually increases up and until 2009.
- Then there is gradual fall in the interest rate until 2011.



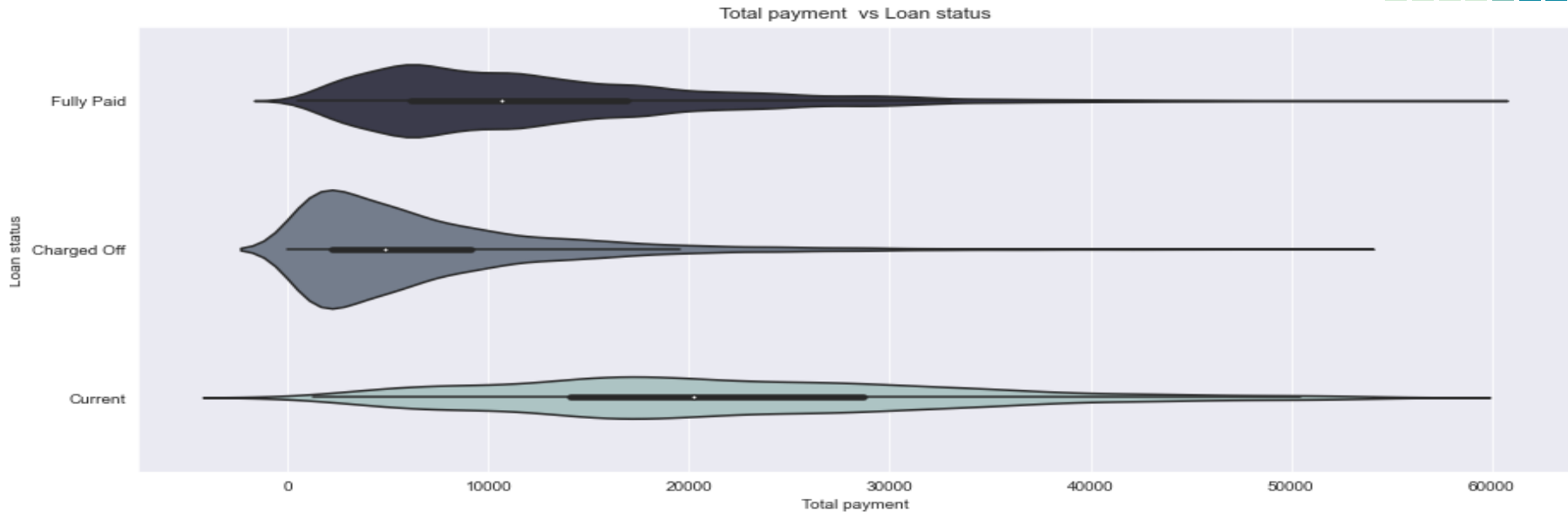
Observation :

- Here we can see DTI vs Loan status.
- charged off has a higher DTI when compared to Fullypaid.



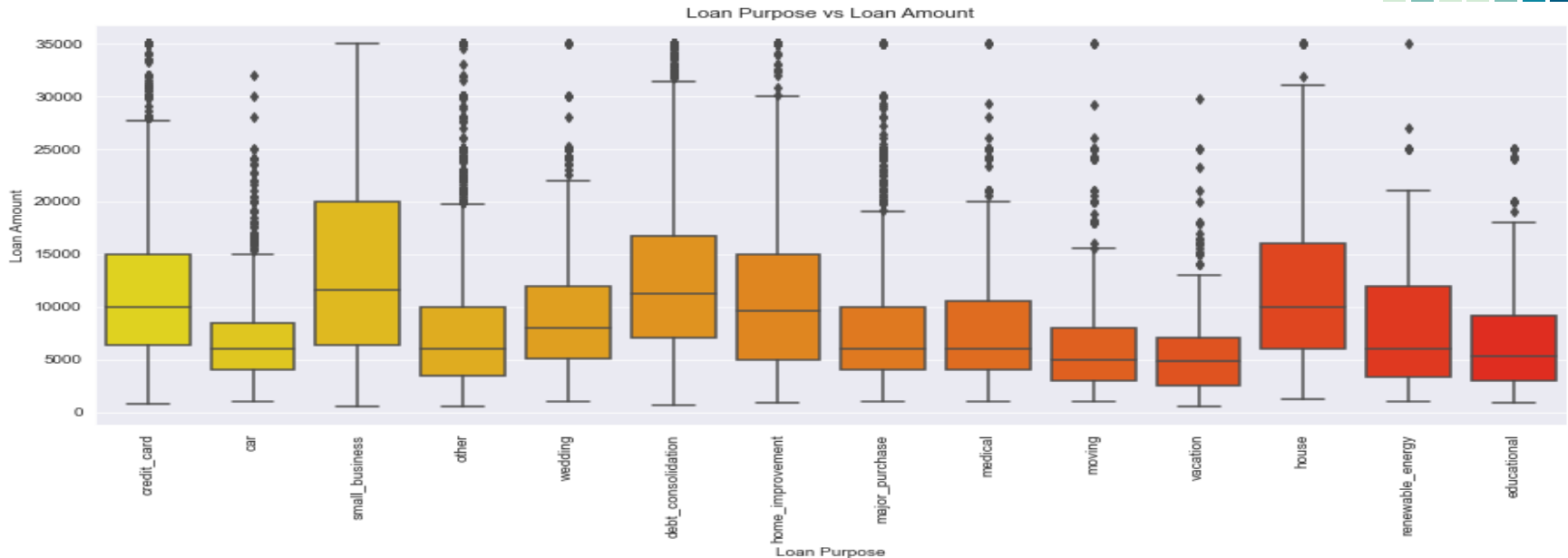
Observation :

- Here we can see Total payment vs Loan status.
- the average of total paid amount is far more less for Charged off people.



Observation :

- Here we can see Loan Purpose vs Loan Amount.
- small business is the most average loan amount.
- there is a lot of difference between q3 and 95th percentile in small business.



RECOMMENDATIONS

- The difference between average annual income of fully paid(66262) and charged off(59840) is nearly 6500.
- So set a minimum mark of average annual income of nearly around 70k so that there will be reduction in charged off cases.
- People between annual income of 50-150k are having a higher interest rate among charged off . Reduction in some interest rate may lead to reduction in charged off cases.
- Similarly Reduce the average loan amount the difference between both segments is nearly 2000 lesser the loan there will be a more chance for reduction in charged off cases.
- And also maximum number of charged off cases come from few states only.

THANKS!