Lending Club Case Study

A Lending company is about to take some business decision so we need to do some analysis on the basis of old data.

Data Understanding

- Take a look of data file manually.
- Get the shape of data frame.
- Get the overall description of data frame.
- Size of data frames.
- Get number of unique columns and rows.

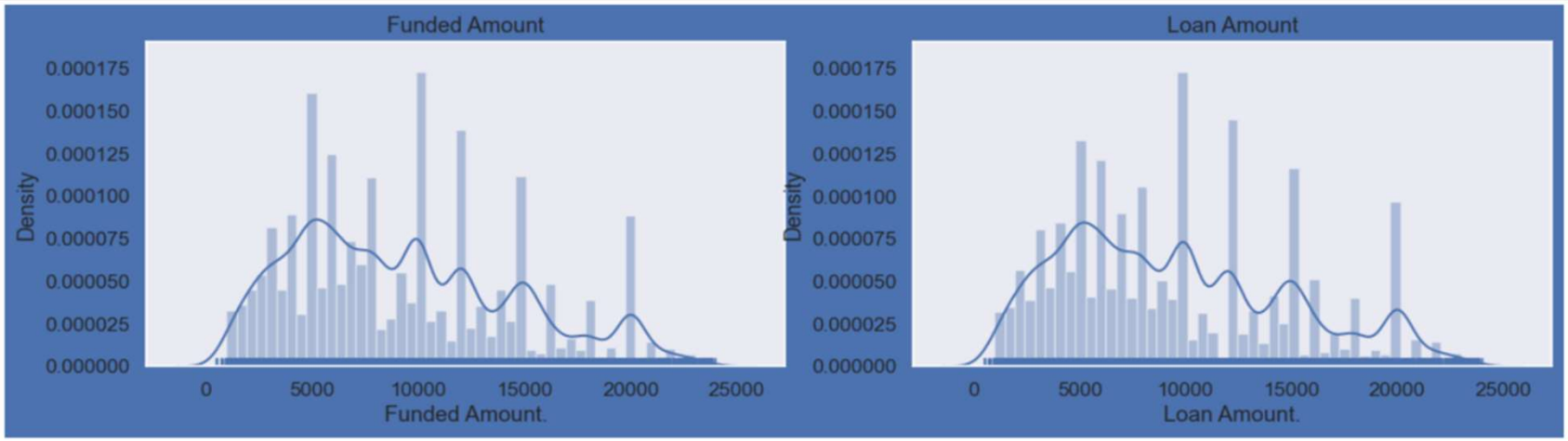
Data Cleaning

- Find the all duplicate values and remove it from actual dataframe and assign to it.
- Find all null and missing values and replace them with some usefull data e.g. int or float if we have missing data in numeric columns.
- Find outliner and try to remove it if not needed.
- Try to get numeric values for the important columns,we are gonna need in analysis.
- Try to merge two or more columns if needed.

Univariant Analysis

- I the field of Lending or Loans we have some fixed import columns like loan_amount , total_payment ,annual_income,loan_status.
- loan_amount , total_payment ,annual_income, should be in numeric format for sure.
- Did univariant analysis for loan_amount , total_payment ,annual_income,loan_status.
- Get relation between any two or three attribute and plot for each.

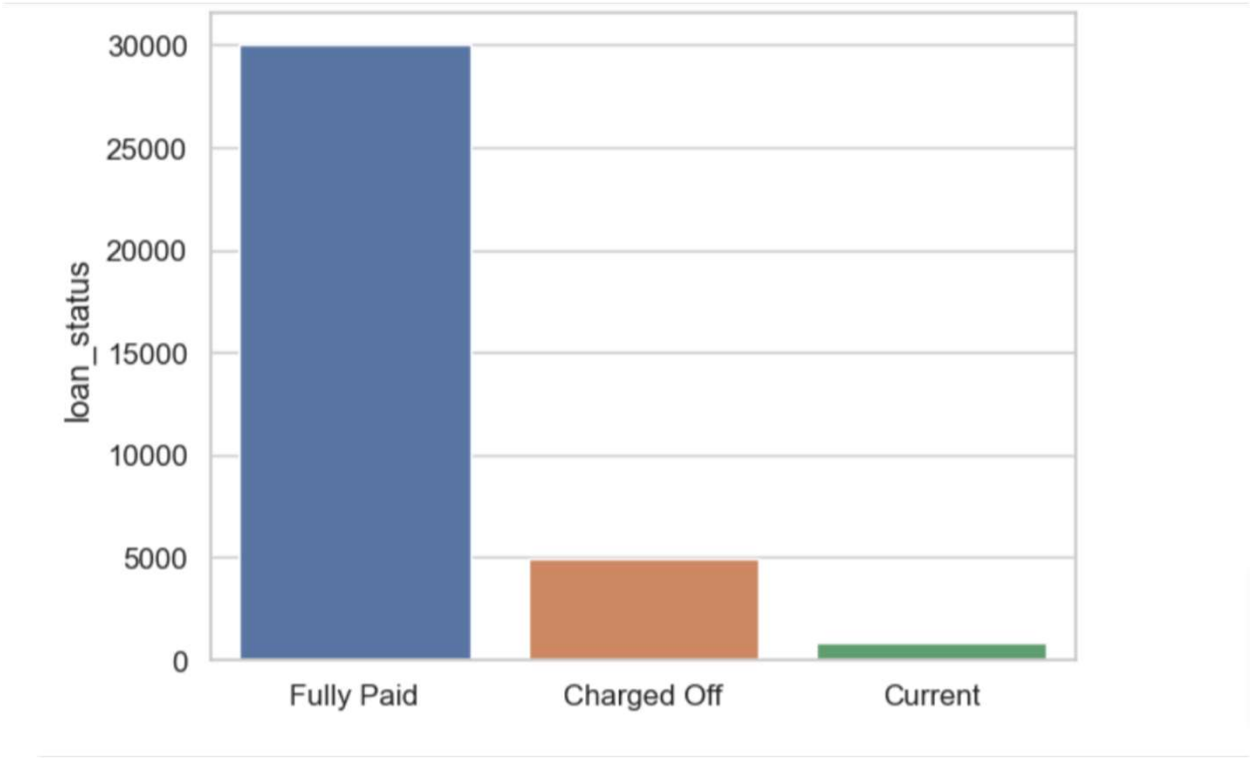
Loan Amount Vs Funded Amount



Conclusion

- Most of the people do payback there loans depending on the annual income and house owning condition.
- The people who have taken loan for depth consolidation are the higher in paying it back at full.
- The people who have taken loan for depth consolidation are also the higher in charged off.
- People with higher annual income are more likely to pay it back are less in defaulter.
- Less people lie between the full paying status and defaulter status.

Loan status vs number of loan application



Loan Status Vs Loan Purpose