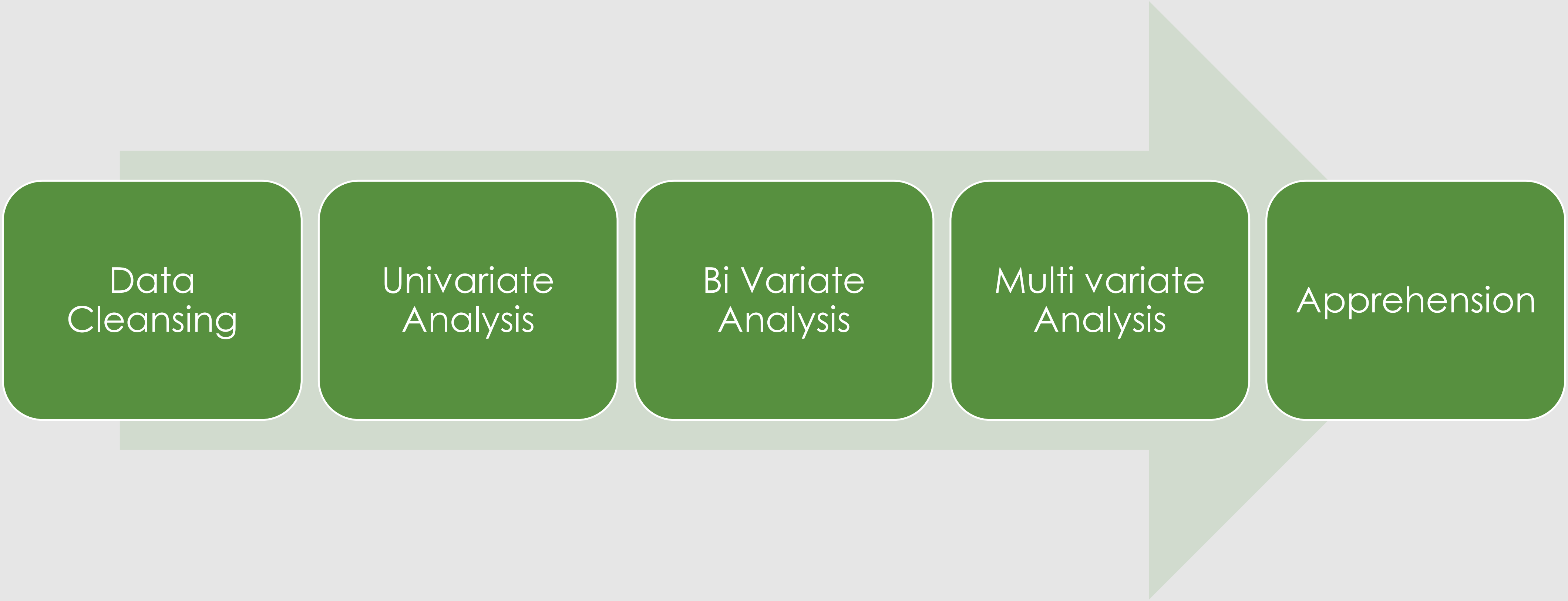
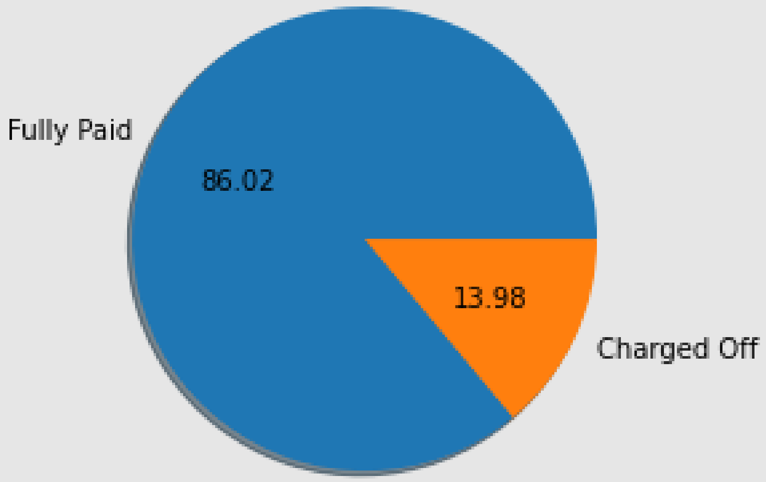
EDA CASE STUDY
LENDING CLUB

Process followed



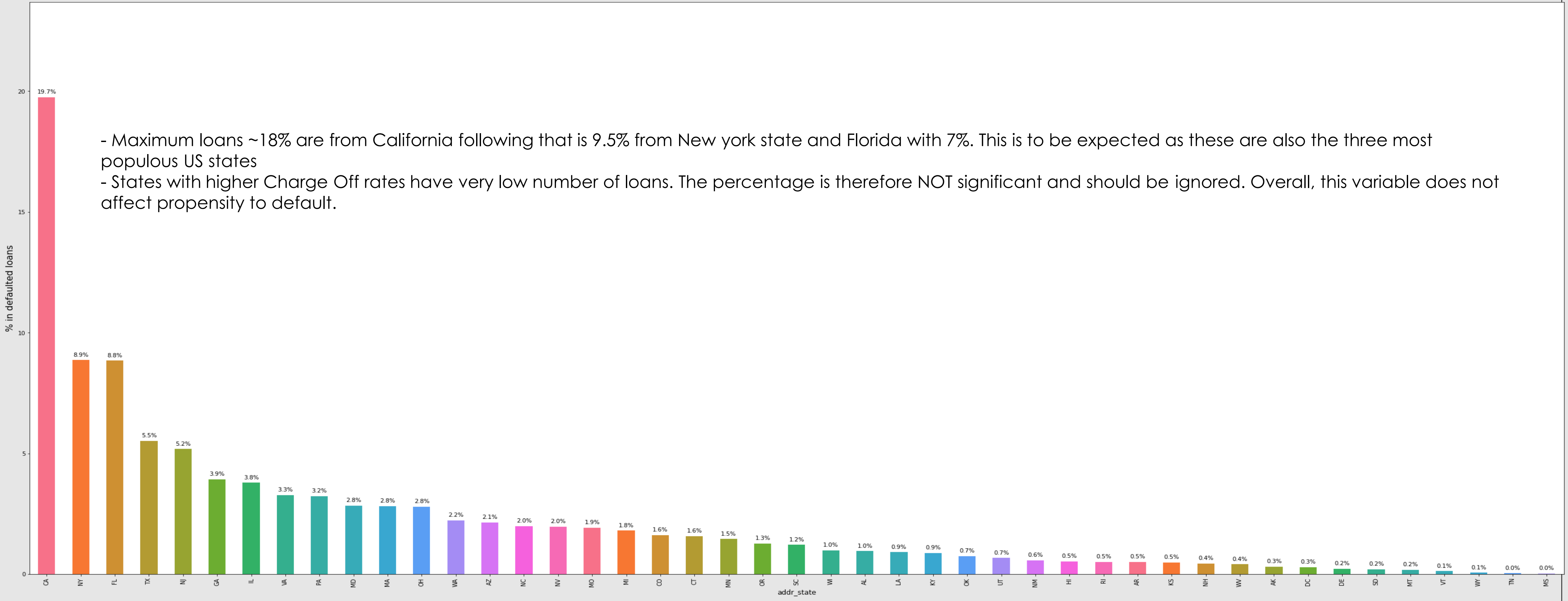
Overview



Approximately 14% of loans in the dataset are defaulted.

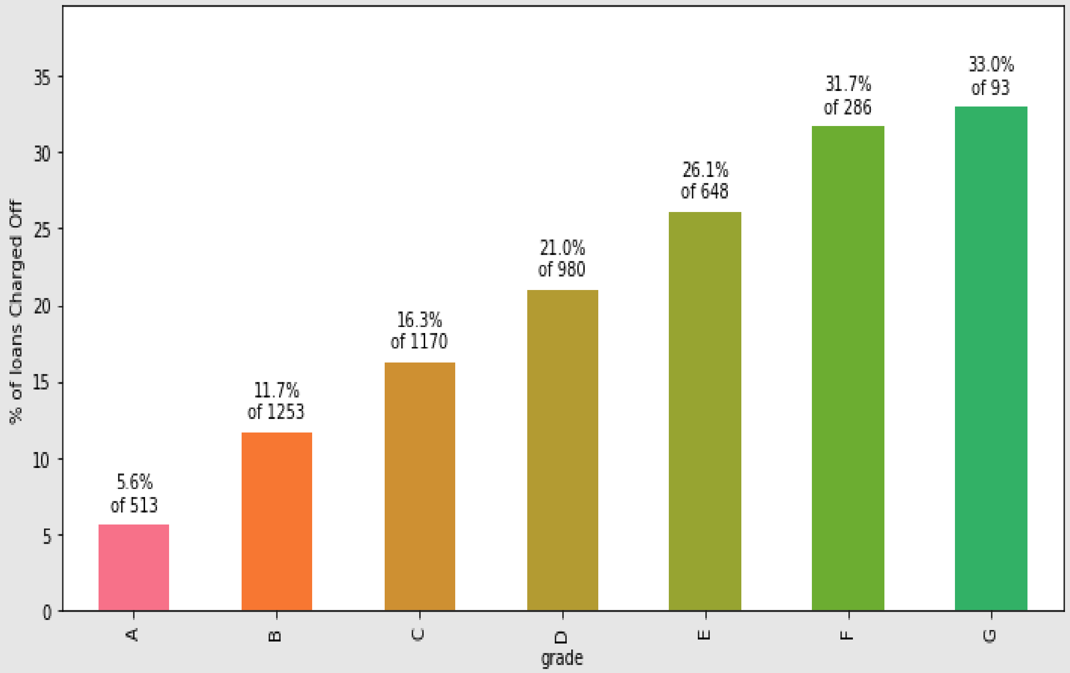
Any variable that increases percentage of default to higher than 20% should be considered a business risk. As this will increase the risk reasonably more than now.

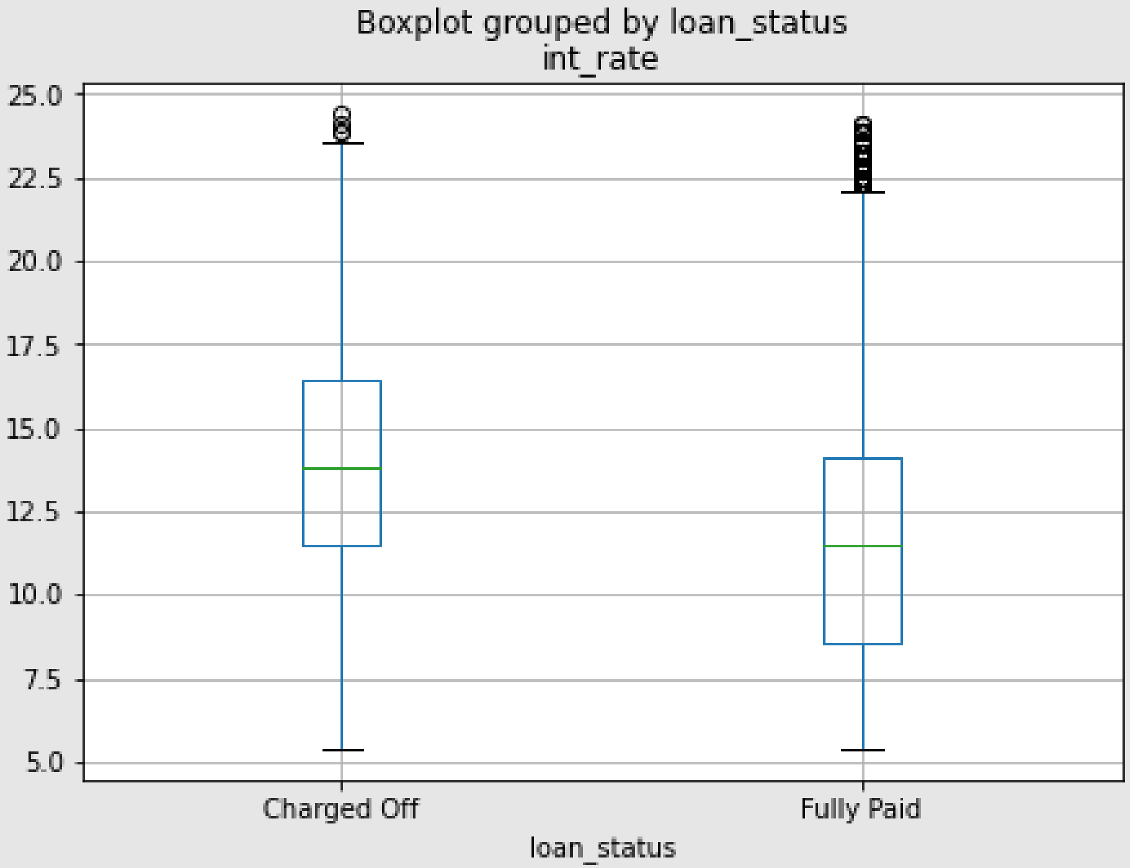
Analysis – I - States



- Maximum loans ~18% are from California following that is 9.5% from New york state and Florida with 7%. This is to be expected as these are also the three most populous US states
- States with higher Charge Off rates have very low number of loans. The percentage is therefore NOT significant and should be ignored. Overall, this variable does not affect propensity to default.

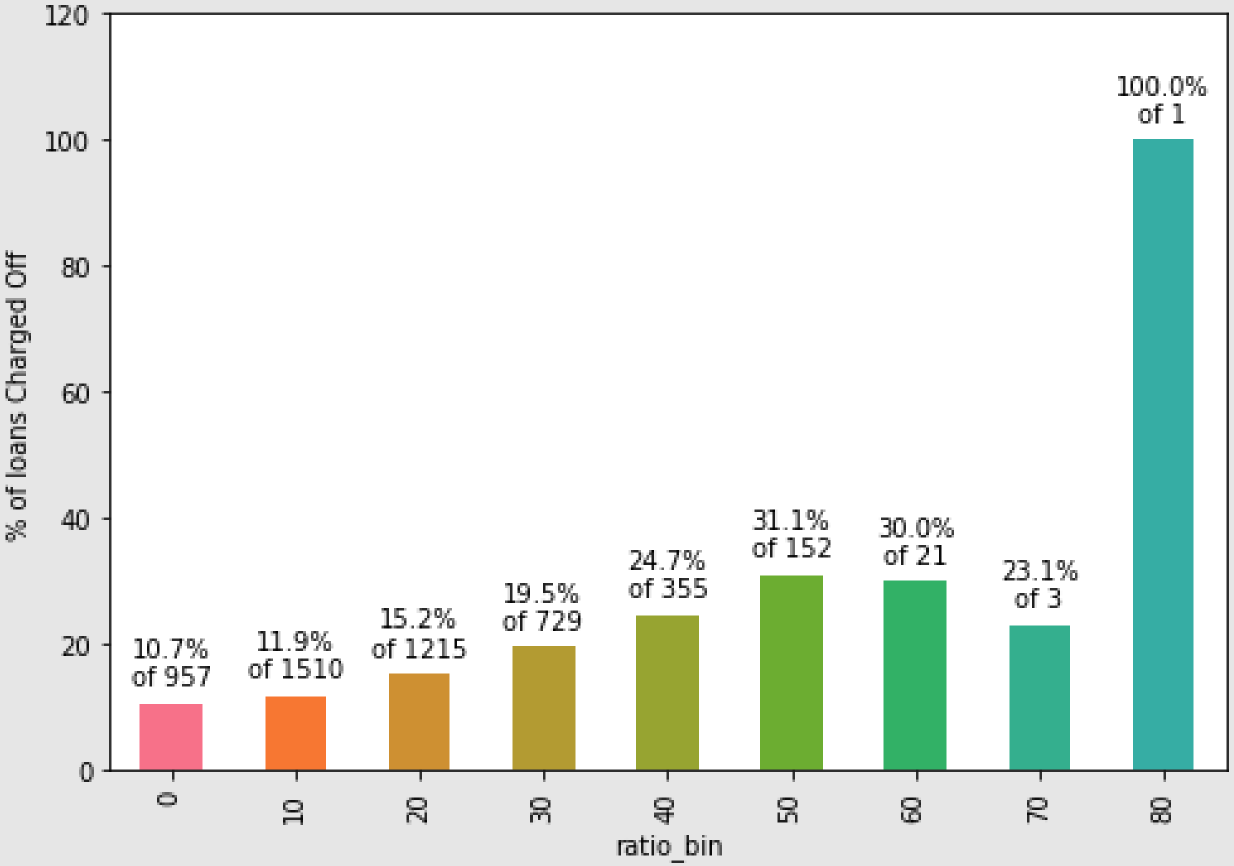
Analysis – II - grade





Nearly 30% of all loans in Grades F and G see a default, which are already considered as risky.

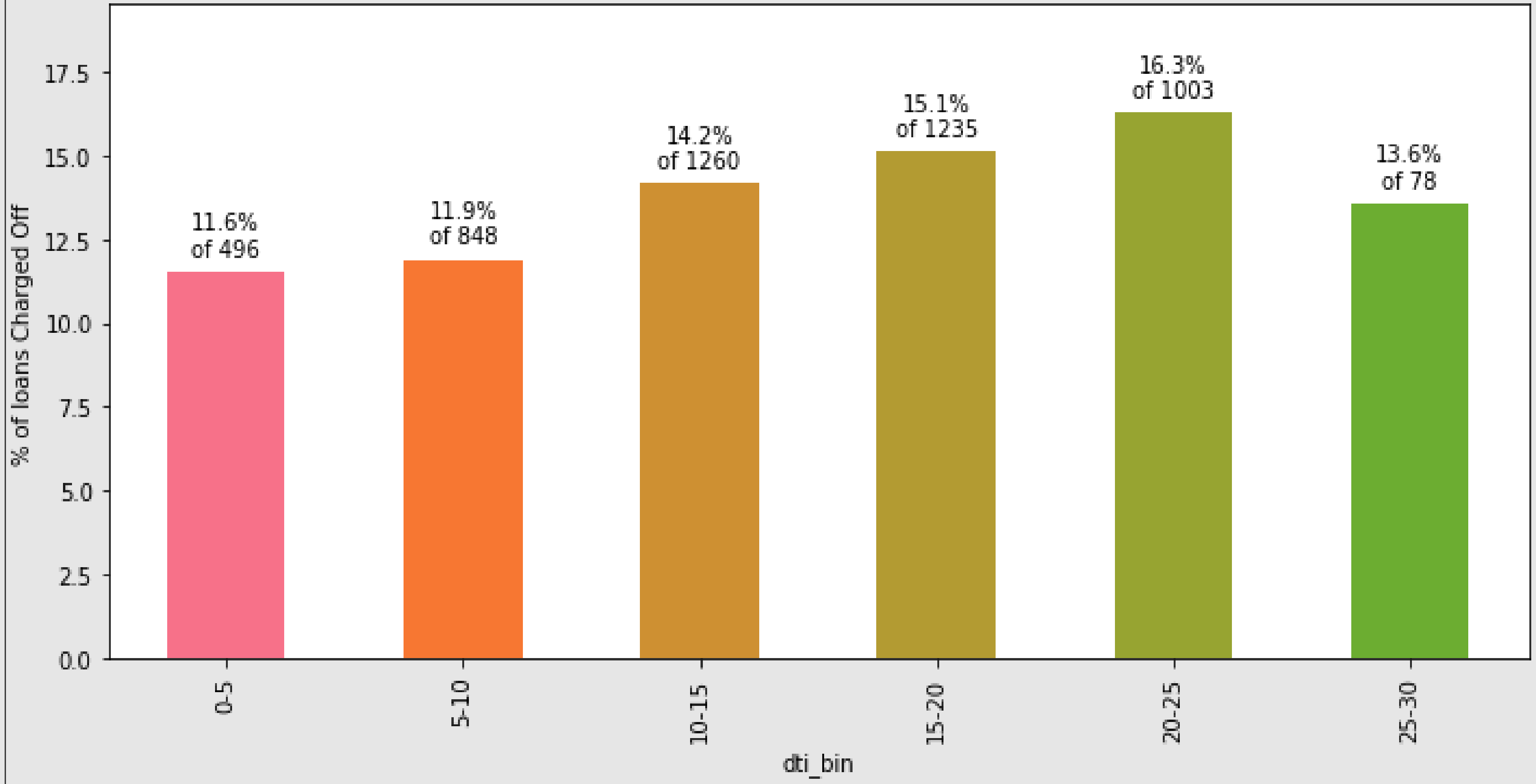
Analysis – III – Annual Income



(Excluding the outliers 60 -80) as number of loans are less.

Loan amounts of 30% of annual income or higher see a high rate of default.

Analysis – IV - dti



- When the dti or dept payment to income ratio is higher than 20, higher percentage of loans are Charged Off
- Higher the dti higher the chances of loan being Charged Off

Summary

From the analysis :

- It is observed that ‘dti’, ‘annual income’, seems to be factors that are co-relating to the charged-off Loans.
- In the future loans, the weightage of these factors can be increased while sanctioning loan and/or deciding loan amount.