

LENDING CLUB CASE STUDY

SUBMISSION

Name: Amit Goel

Name: Abdulazeez Basheer

Cohort: C9, June 2019

Business Objective and Goal of EDA

Business Objective

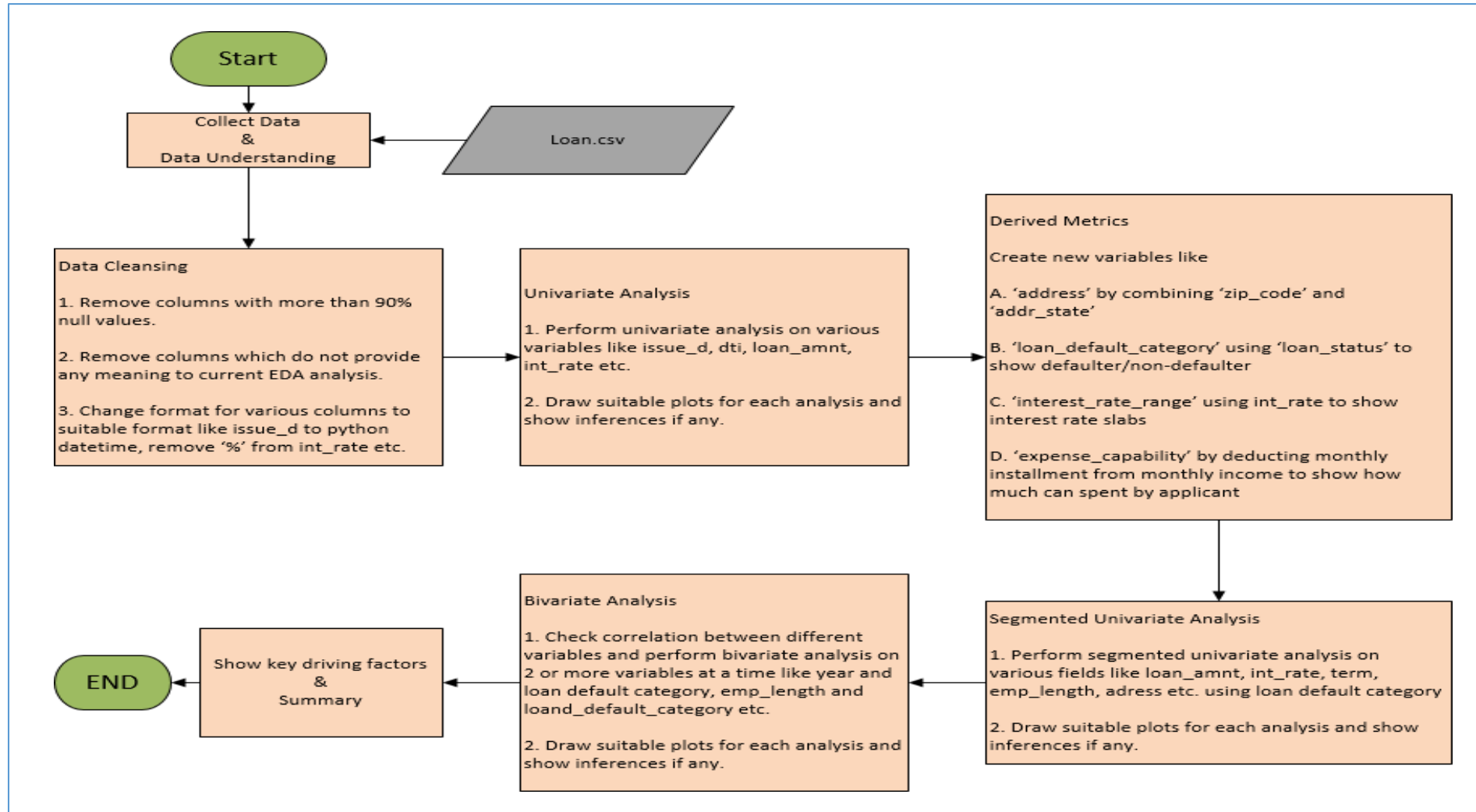
Company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

Goal of Exploratory Data Analysis (EDA)

Company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface. Like most other lending companies, lending loans to 'risky' applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders.

Company should be able to identify risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.

Problem Solving Methodology

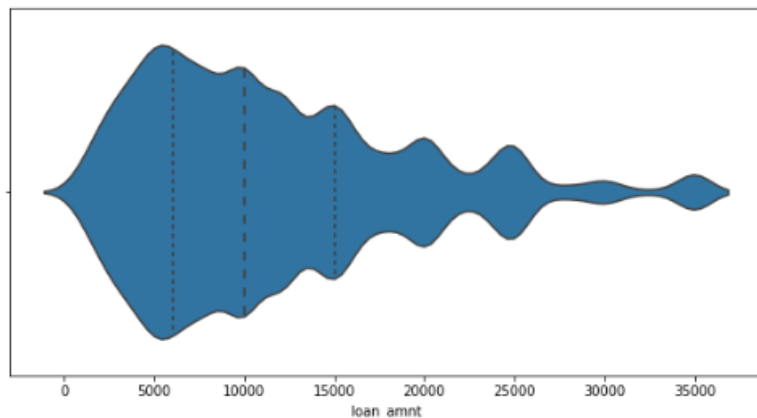


Understand Data Set and Data Cleansing

- ☐ Calculating column's missing percentage
- ☐ Removing columns with high NA percentage
- ☐ Removing Columns with only one unique values which are not having real predictive power
- ☐ Removing columns which are not having real predictive power. Variables like id, member_id, url etc.
- ☐ Removing the post loan attributes, these are basically generated post loan. Variables like collection_recovery_fee, recoveries, total_pymnt, total_pymnt_inv, total_rec_prncp etc.

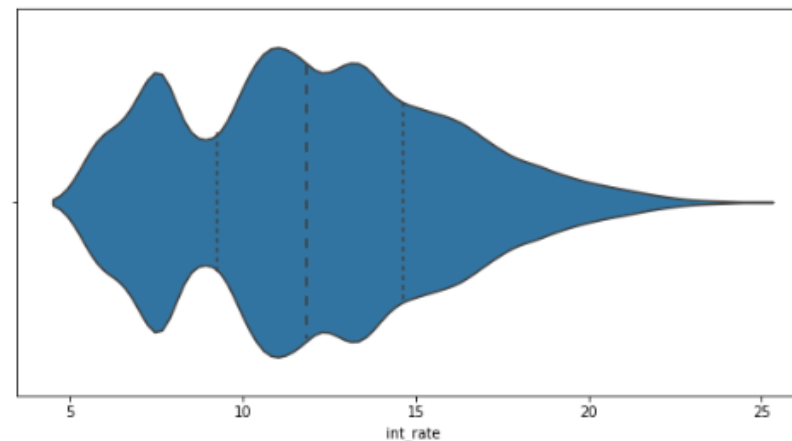
Univariate Analysis

Spread and central Tendency of loan_amnt



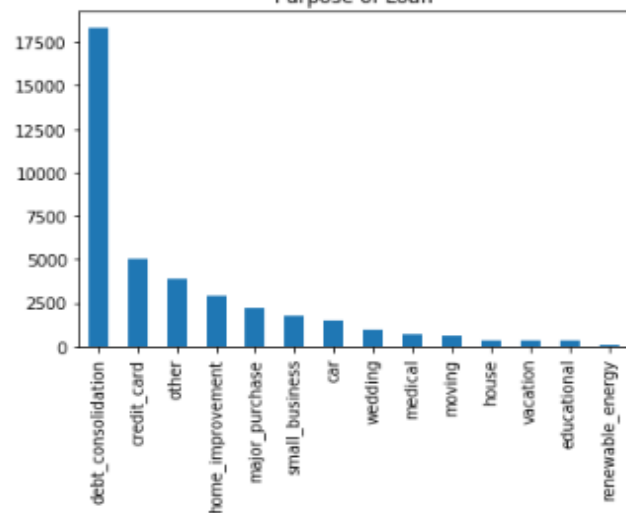
Insights: Most of the loan amount are 6000 to 15000 and we don't have more Outliers

Spread and central Tendency of int_rate



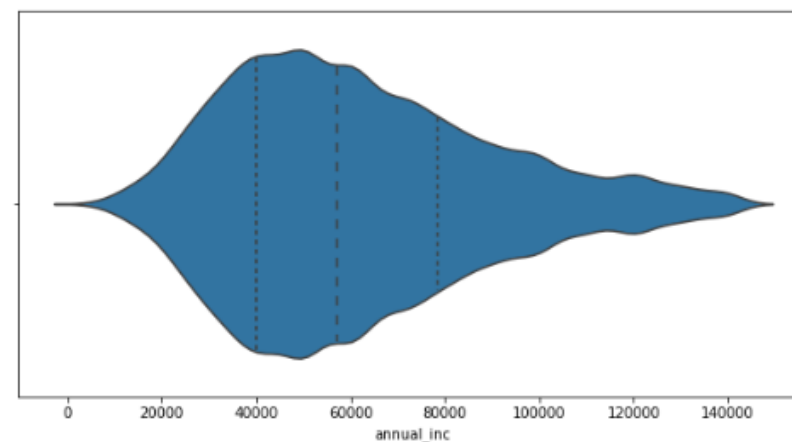
Insights: Most of the loan interest are distributed between 9 to 15 and we don't have more Outliers

Purpose of Loan



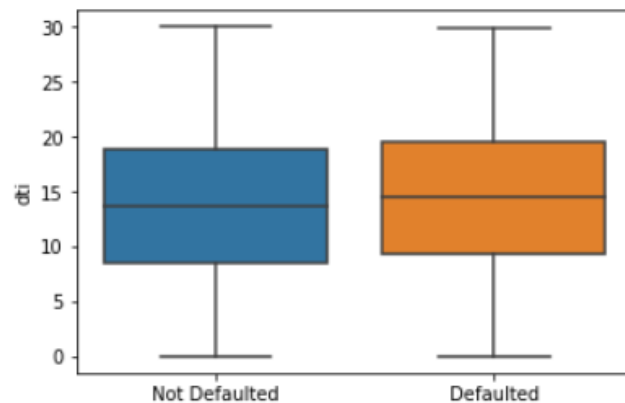
Most of the loan seeking reason is to pay the other debts(debt_consolidation)

Spread and central Tendency of annual_inc

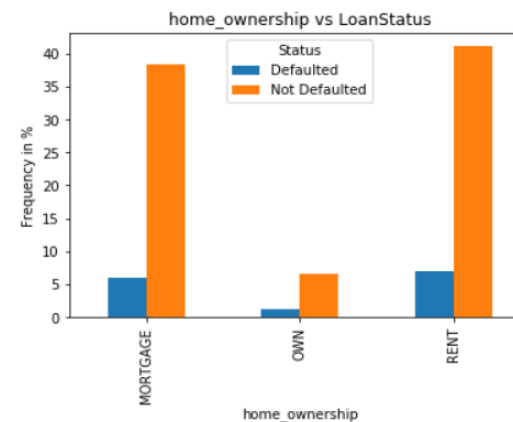


Insights: Most of the applicants earn between 40000 to 75000 annually

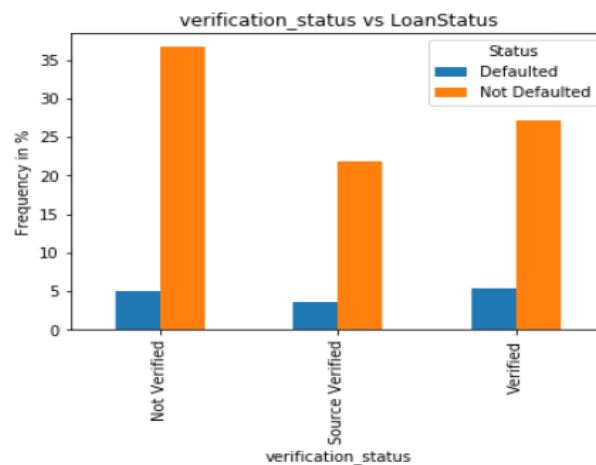
Segmented Analysis



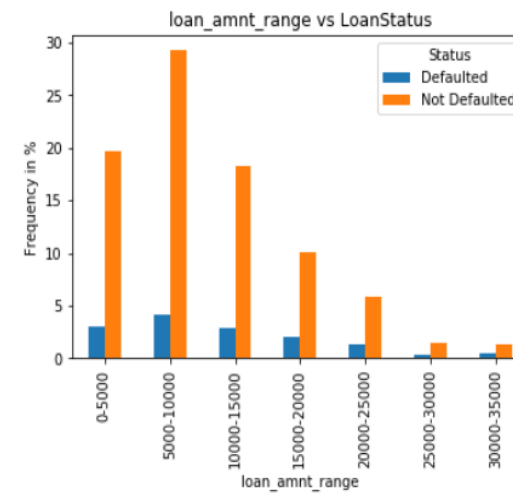
Insights: Loan Defaults happened in Range of 10 to 20 and Defaulters debt to income ratio mean is high



Insights: Most defaulters are in Rented Home

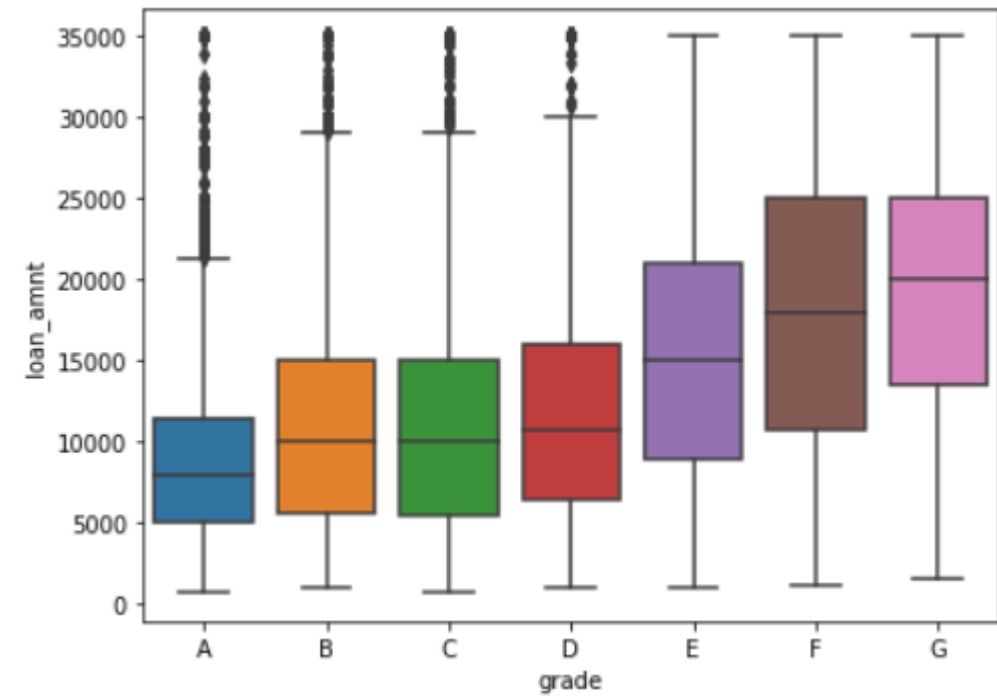
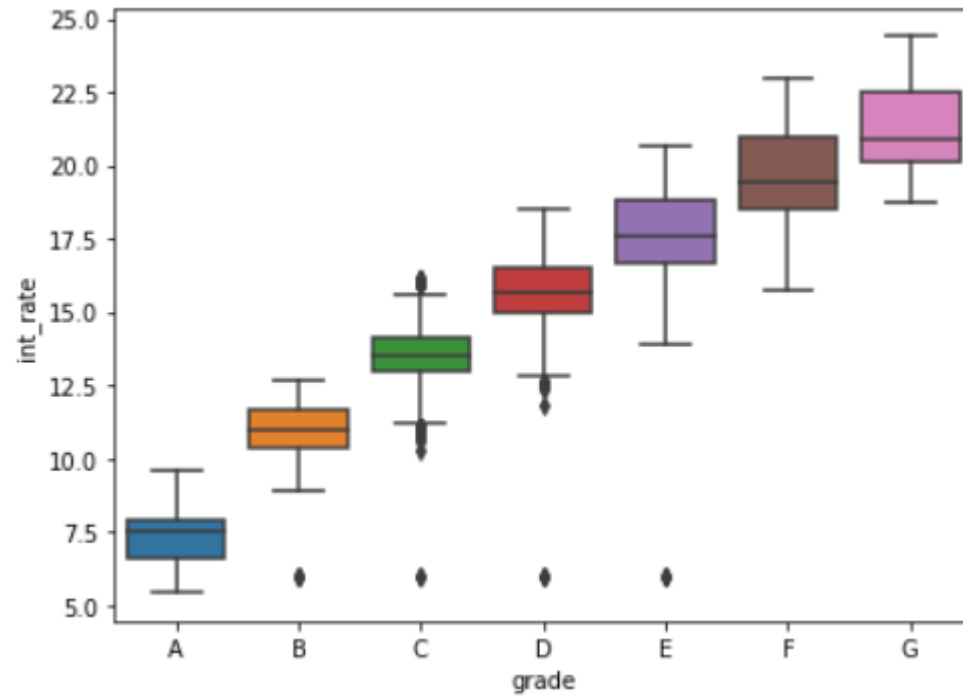


Insights: No major Significant identified



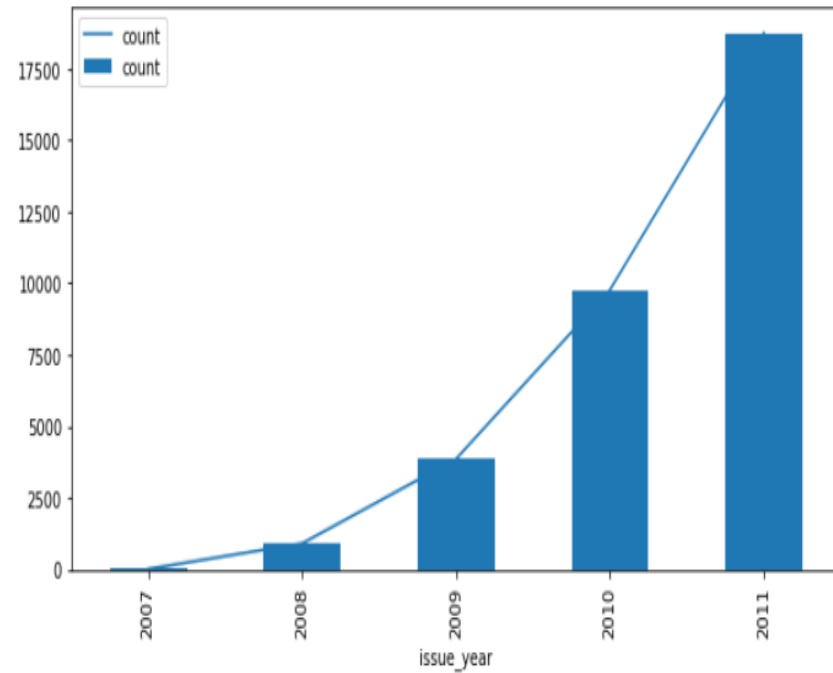
Insights: Most of the defaults happens in Loan amount between 5000 to 100000

Segmented Analysis

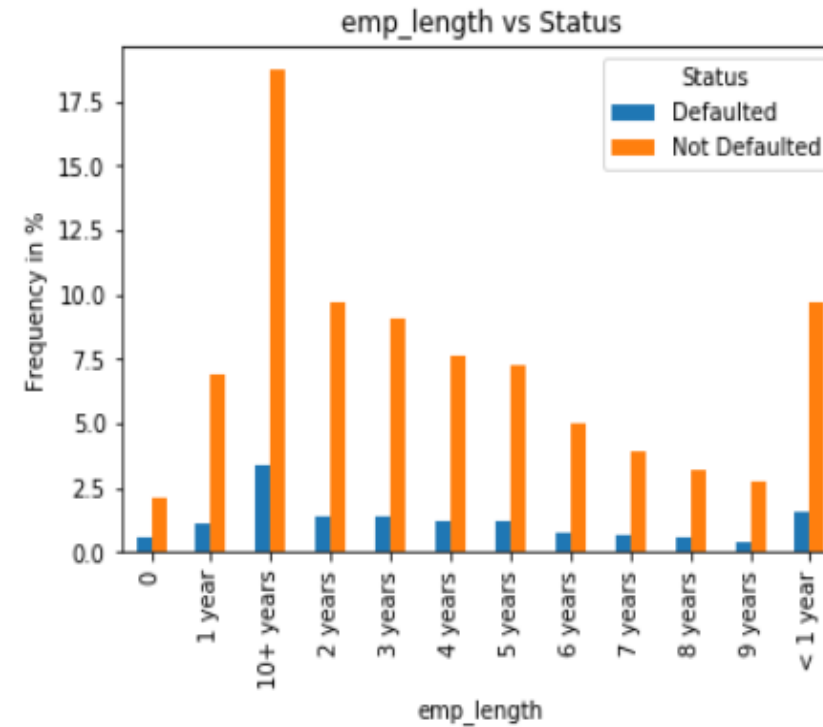


Insights: As the grade Increase the Intrest Rate also Increases.

Segmented Analysis

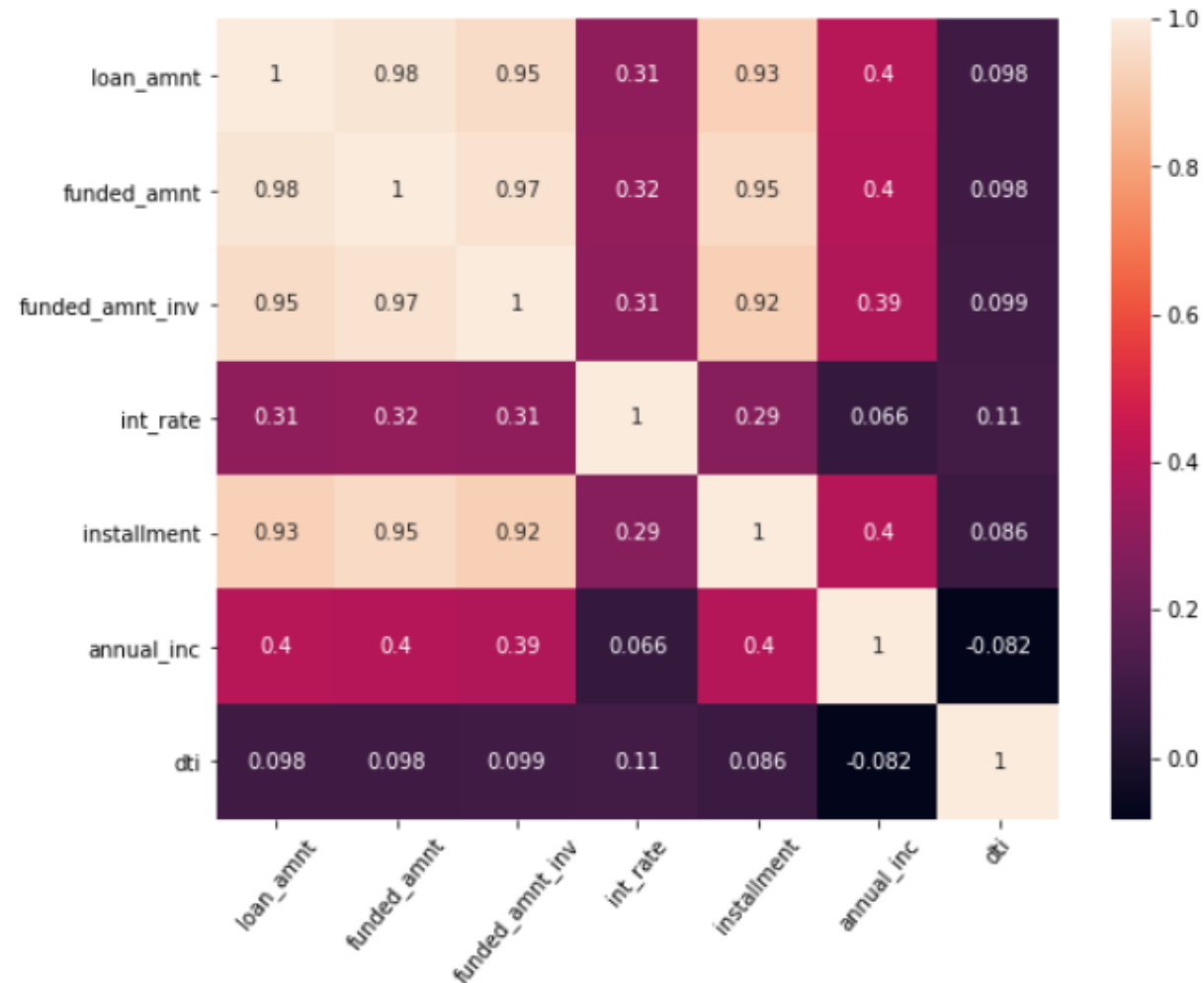


Insights: Number of applicants doubles every year



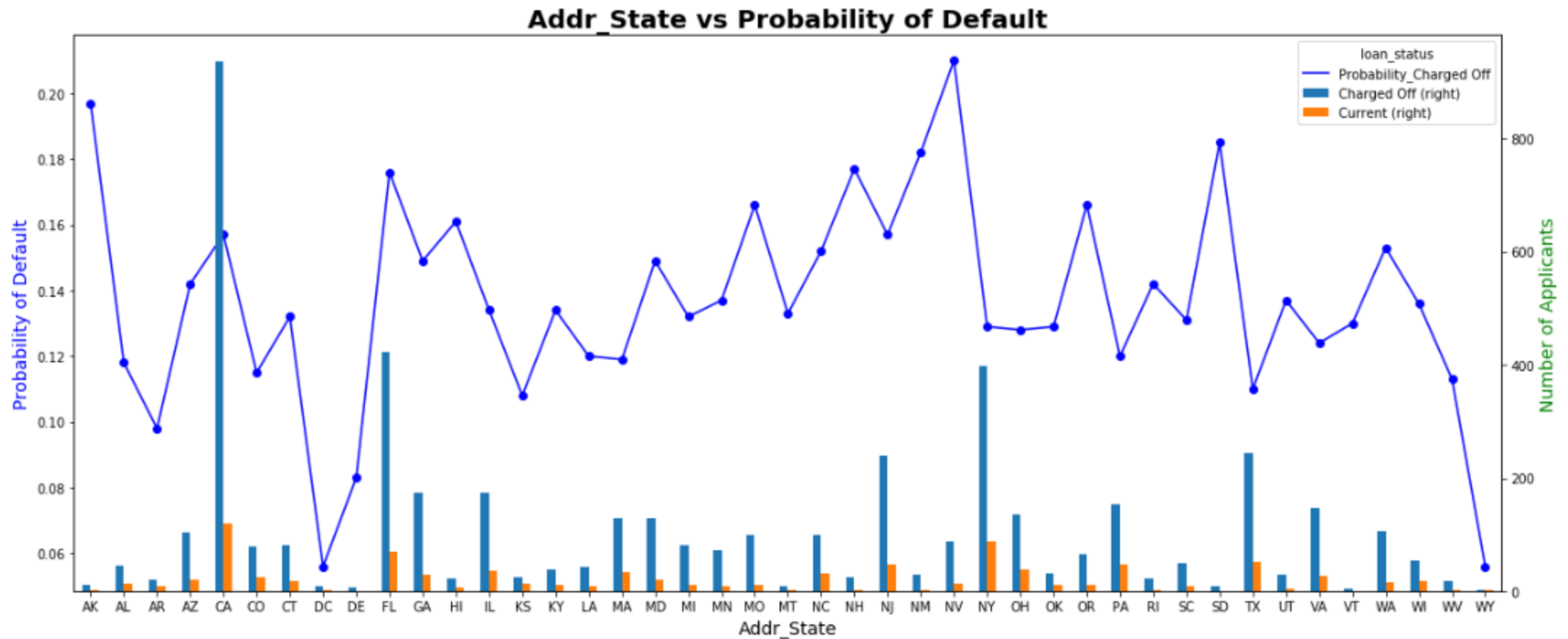
Insights: Lower Experience employees tend to default more on loan payment than higher experience with respect to probability

Bivariate / Multivariate Analysis



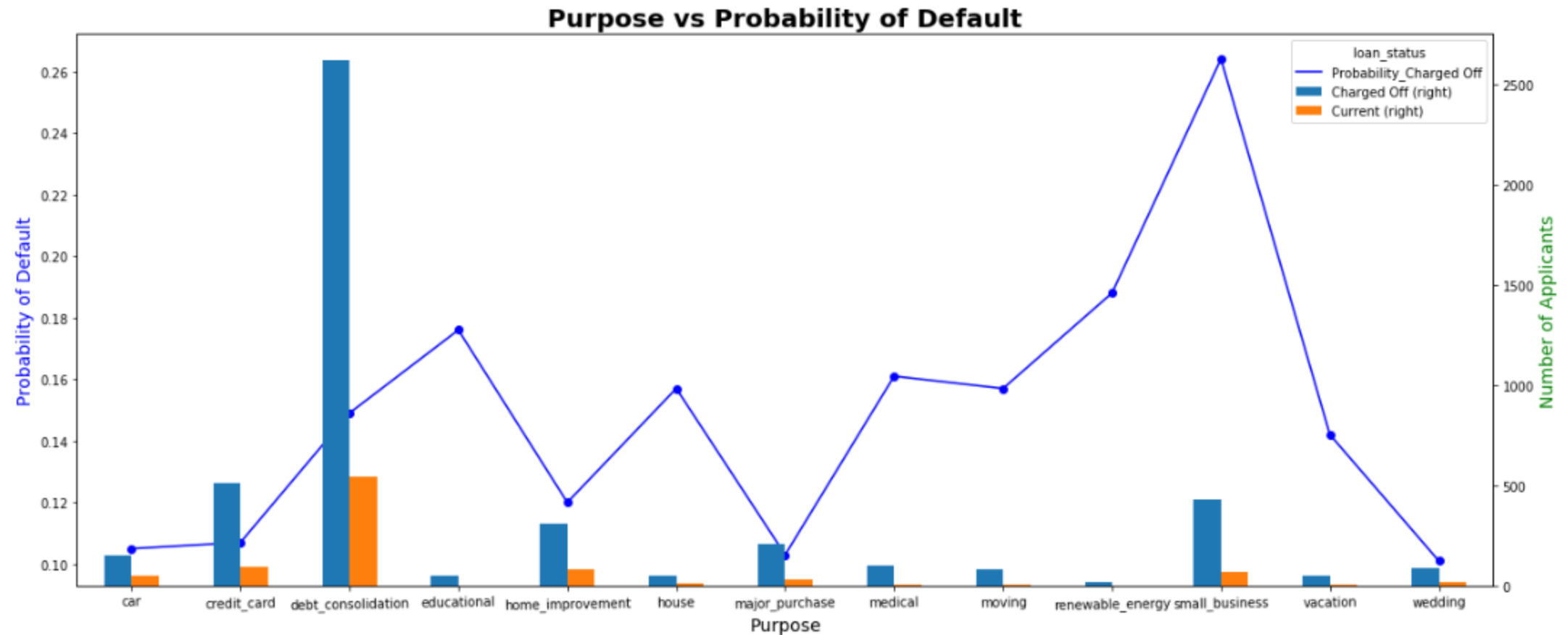
Insights: It is clear from the Heatmap that how 'loan_amnt', 'funded_amnt' & 'funded_amnt_inv' are closely interrelated.

Bivariate / Multivariate Analysis



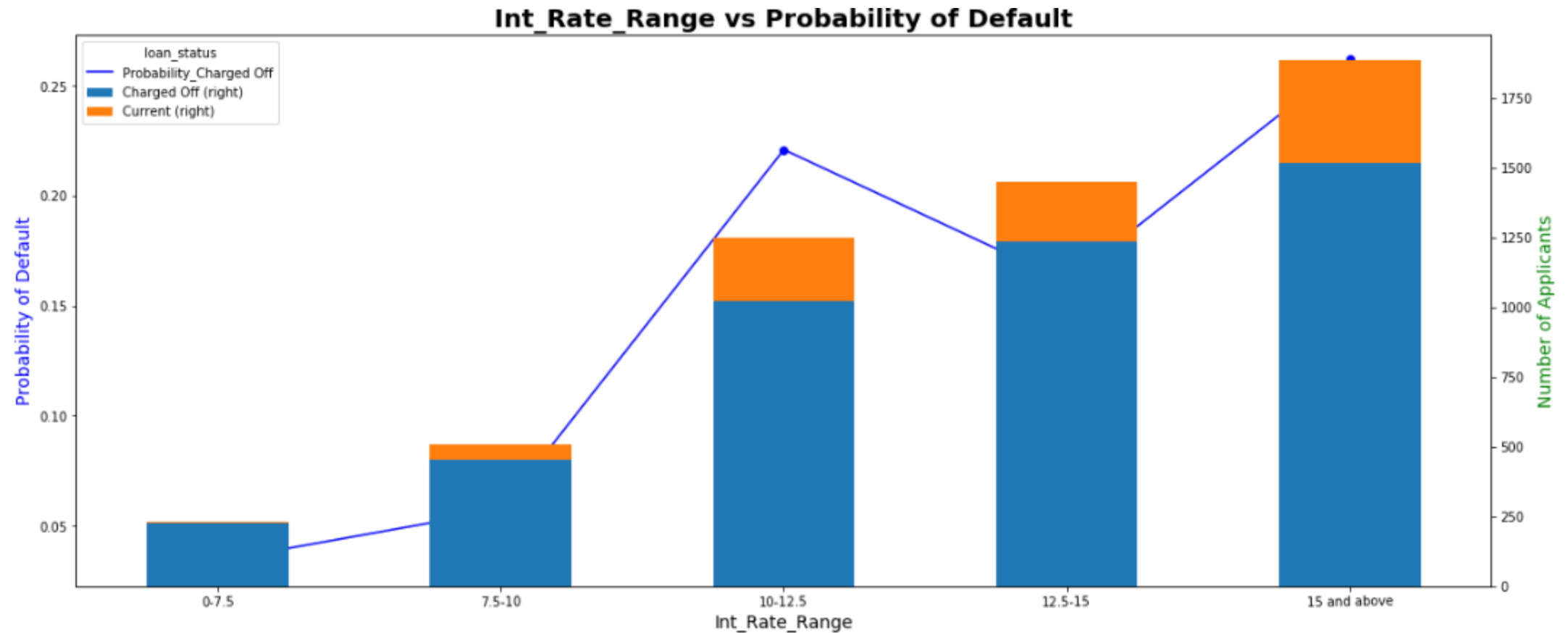
Insights: There are multiple States/Provinces with high probability of charge, highest being 'NV'.

Bivariate / Multivariate Analysis



Insights: Applicants who has taken the Loan for 'Small Business' has the highest probability of charge off of 26%. So bank should take extra caution like take some asset or guarentee while approving the loan for purpose like 'Small Business'

Bivariate / Multivariate Analysis



Insights: As the interest rate is increasing the probability that person will default is increasing with highest of 11% at 15% & above bracket.

Key Driving Factors

- ✓ Purpose of the loan (***purpose***)
- ✓ Interest Rate for loan (***int_rate / interest_rate_range***)
- ✓ Term for loan (***term***)
- ✓ Location(***address_state***)
- ✓ Debt to income ratio (***dti***)

Summary

Purpose of Loan:

Applicants who has taken the Loan for 'Small Business' has the highest probabiltiy of charge off of 26%. So bank should take extra caution like take some asset or guarentee while approving the loan for purpose like 'Small Business'

Location:

There are multiple States/Provinces with high probability of charge, highest being 'NV'

Interest Rate:

As the interest rate is increasing the probability that person will default is increasing with highest of 11% at 15% & above bracket

dti:

If Debt to income ratio is more then applicants have higher changes of defaulting. For example, in interest rate range of 20-25% and term 60 months, if dti is more than 14% then applicant will have higher chances of defaulting.

Term for loan (term):

Loans with term of 60 months, have higher chances of defaulting.



THANK YOU