

Lending Club Case Study

Kajal Kaspate

Aarti Rani Bhatia

Abstract

An Exploratory Data Analysis was carried out on the subset of loans data to understand factors driving borrowers to default.

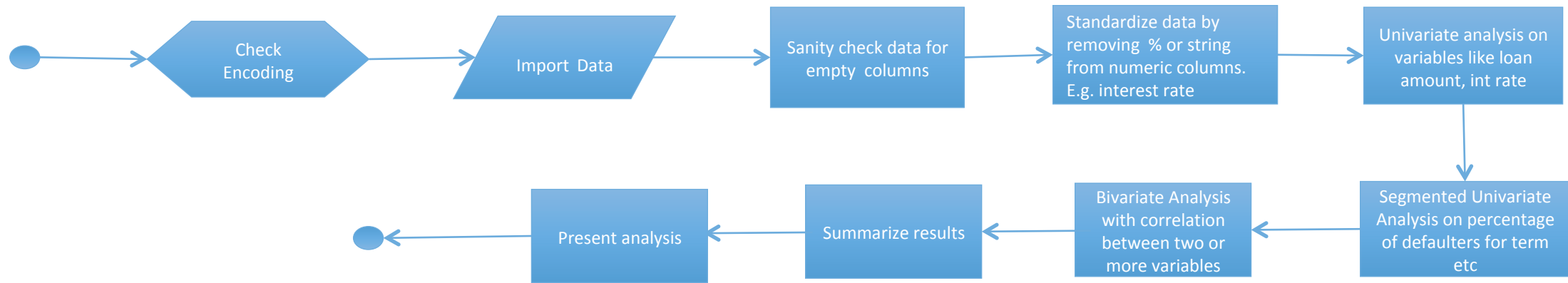
We have considered below loans –

- Loans that are ‘Fully Paid’ off
- Loans that are ‘Charged Off’

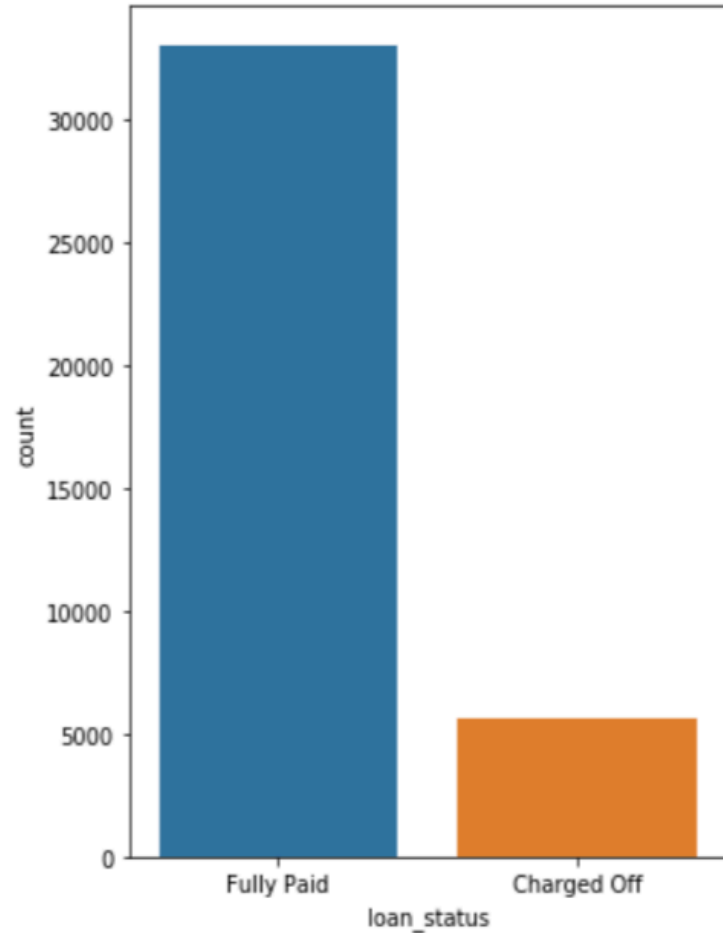
Data points used to find the driving factors are both customer and loan related.

- Customer demographic attributes like annual income, home ownership etc
- Loan related attributes like loan ammount, term etc.

Problem solving methodology

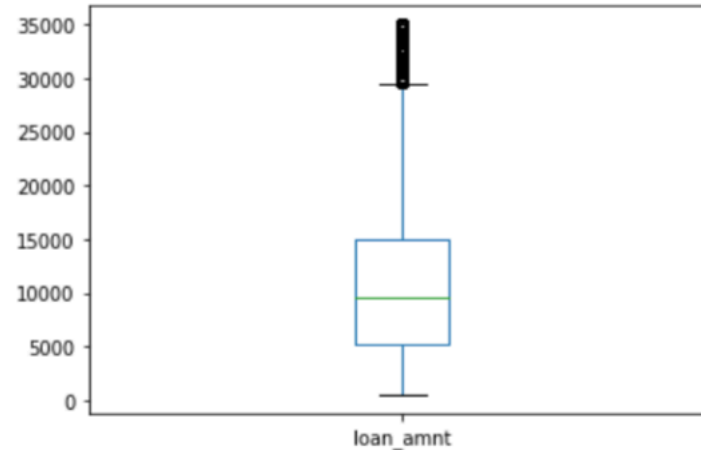
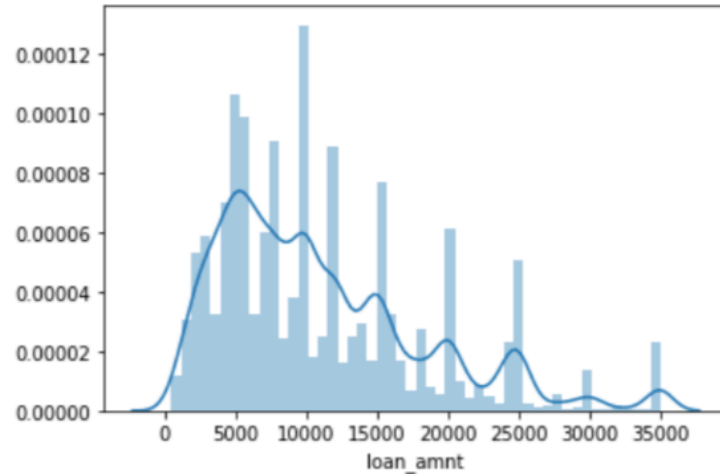


Univariate Analysis – Loan Status

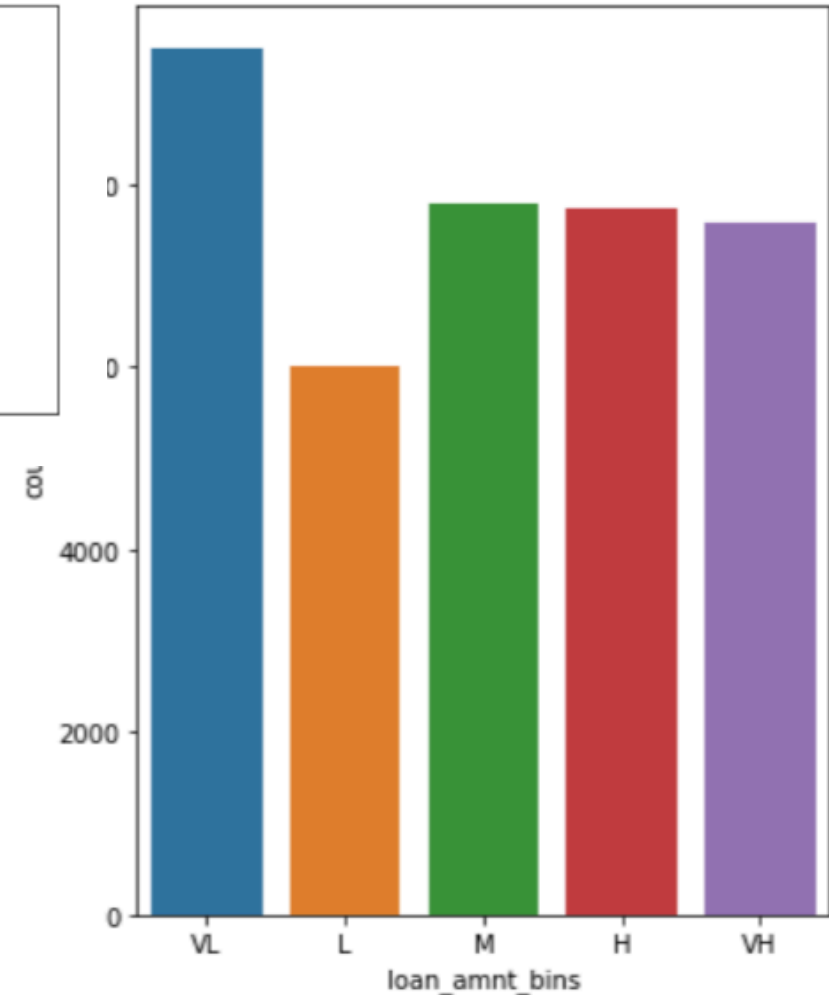


- Total loans considered in this data are 38563
- 32937 loans are fully paid while 5626 loans are charged off.
- 14.5% loans are charged off

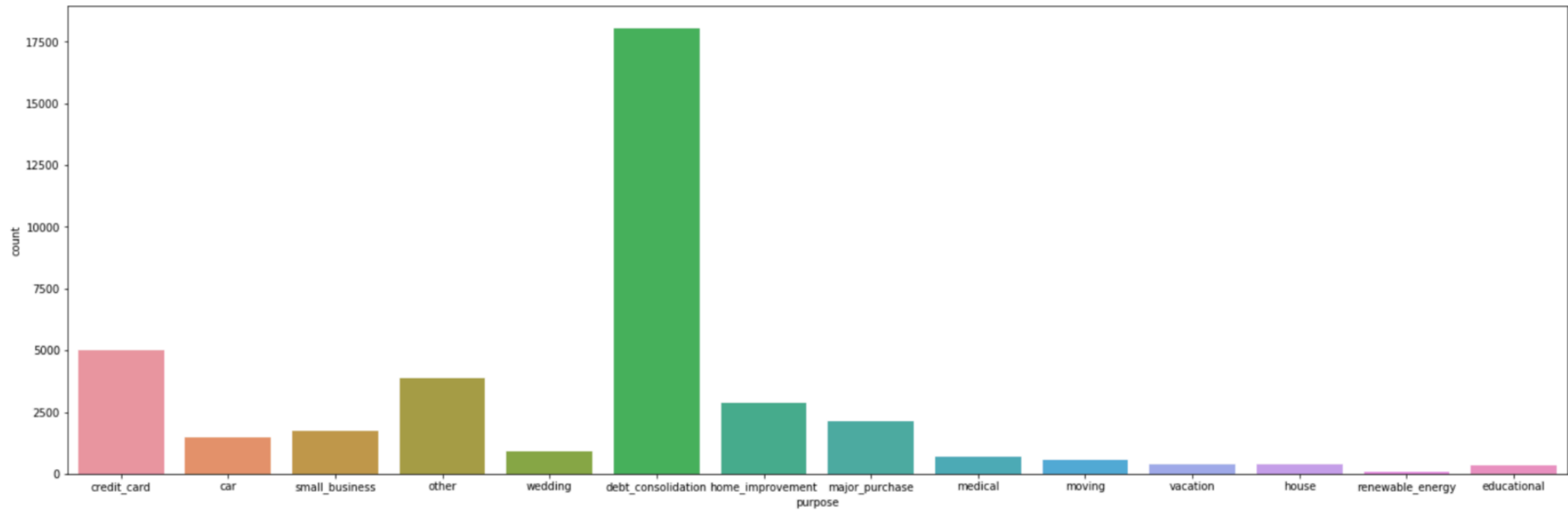
Univariate Analysis – Loan Amount



- 46% of loans are with the loan amount below 40th percentile
- Mean loan amount is 11047.2 and has std dev of 7348.60

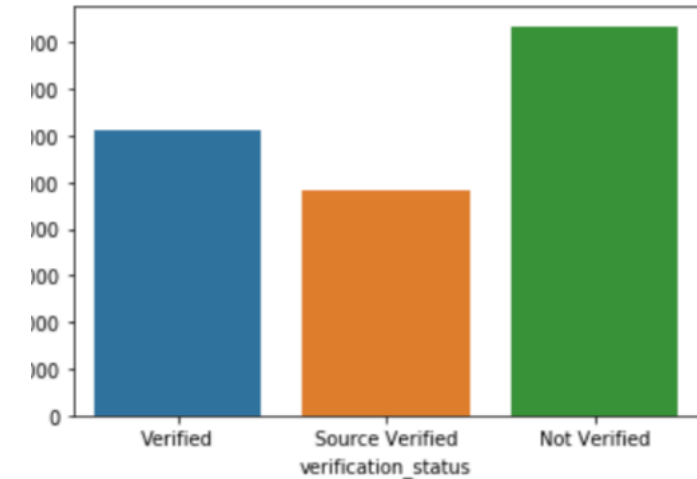
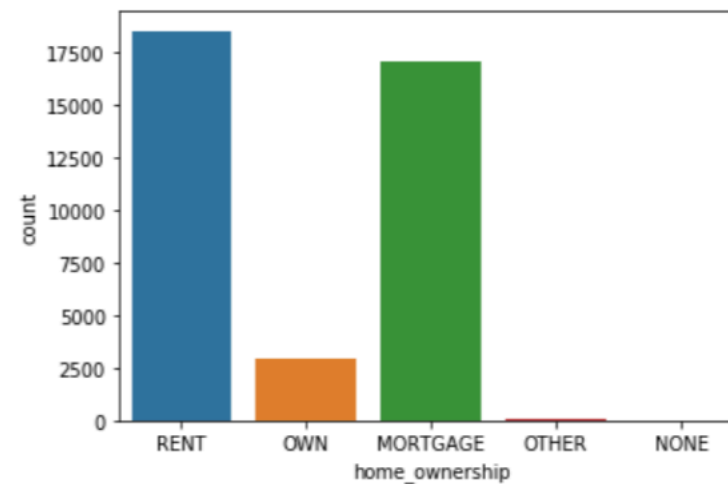
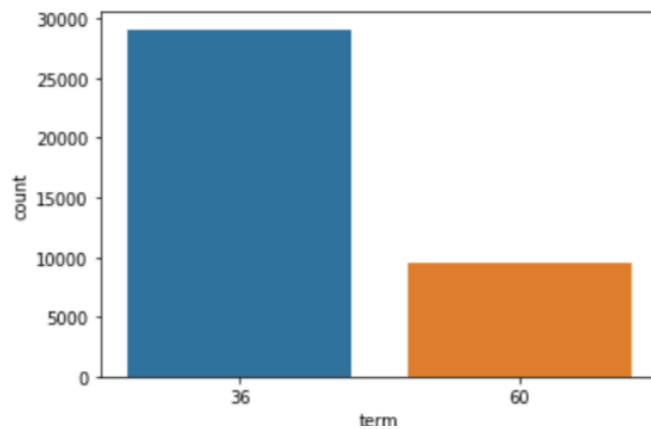


Univariate Analysis – Purpose



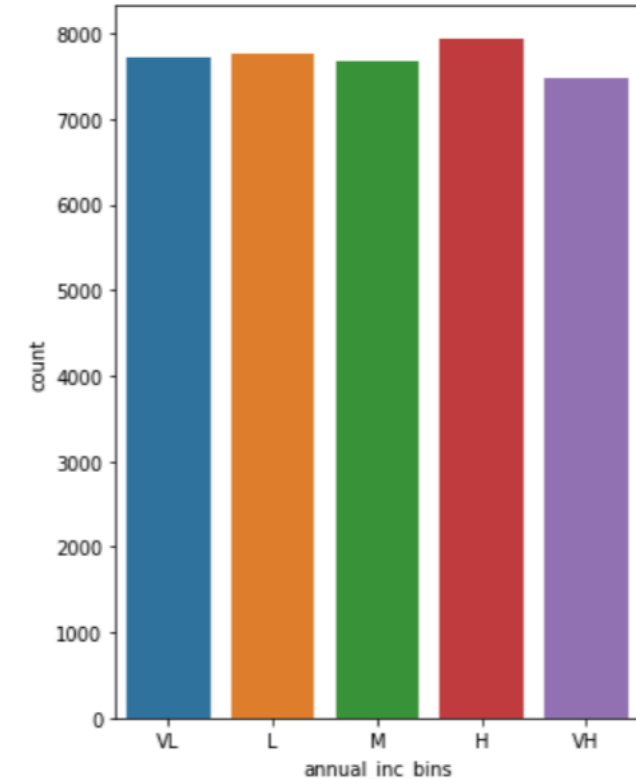
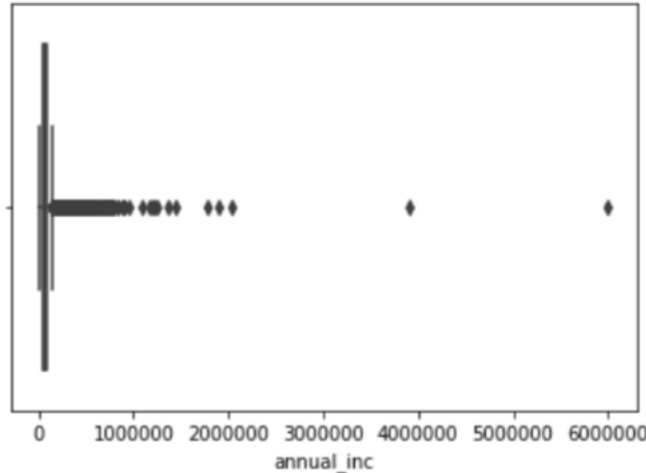
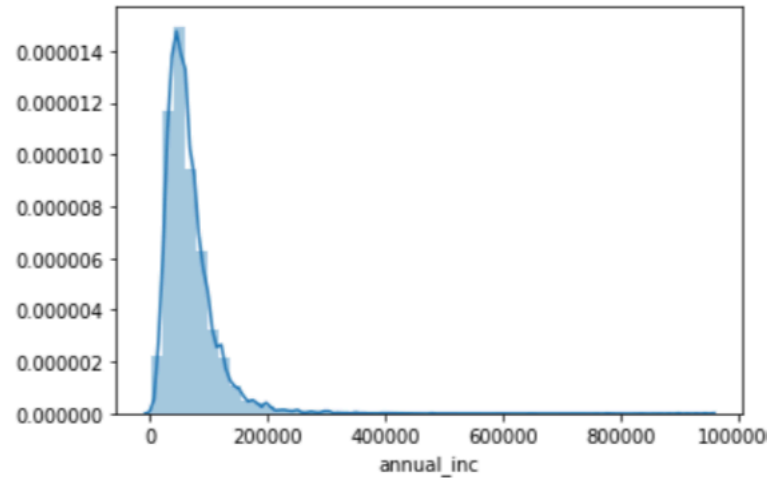
- Purpose of 46.9% loans is debt consolidation

Univariate Analysis – Term, Home Ownership, Verification Status



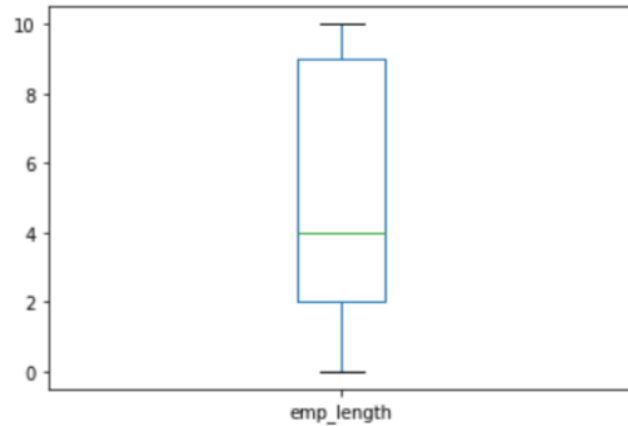
- Loans with term as 36 months are 75.42% of total loans
- Number of customers who Rent and have Mortgage are 92% of total customers
- 43.27% of customers' income was not verified while lending money

Univariate Analysis – Annual Income

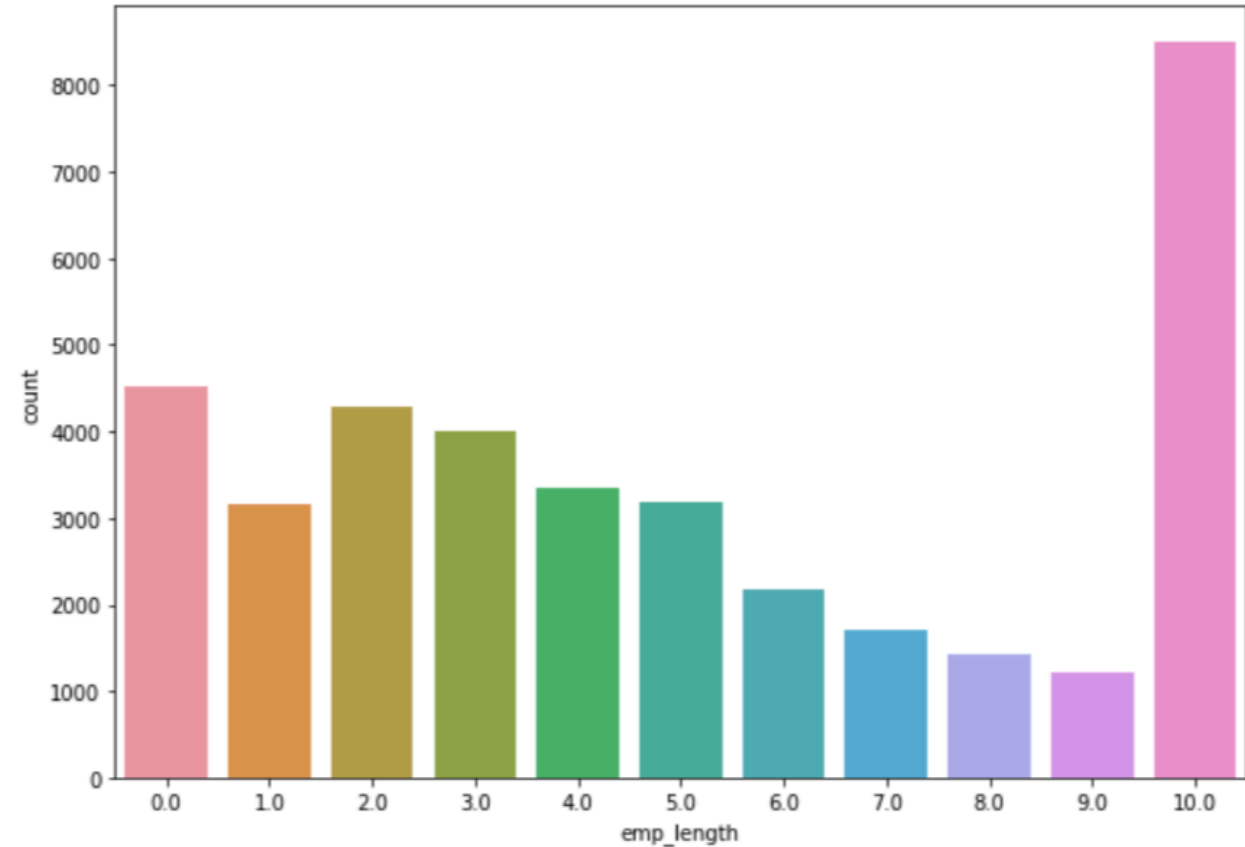


- Annual income for 60% customers is below 60th percentile
- After binning annual incomes in 5 buckets for percentiles – 20th, 40th, 60th, 80th and 100th, it seems customers are divided almost equally.
- Range of annual income is very high and it is 944000

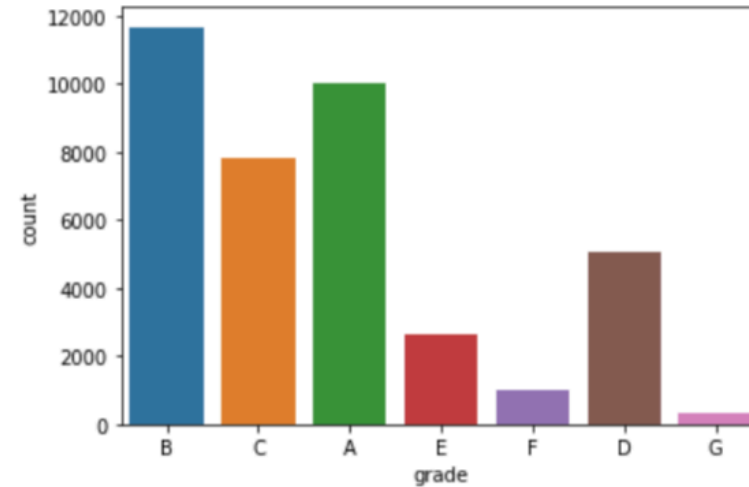
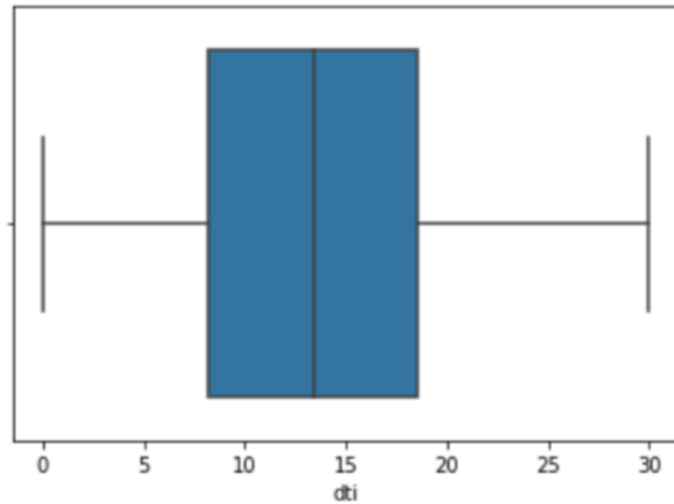
Univariate Analysis – Employment Length



- Employment length of customers have a median of 4 years.
- Customers with more than 10 years of employee length have around 8483 loans out of the total loans of 38536 (22.01%).

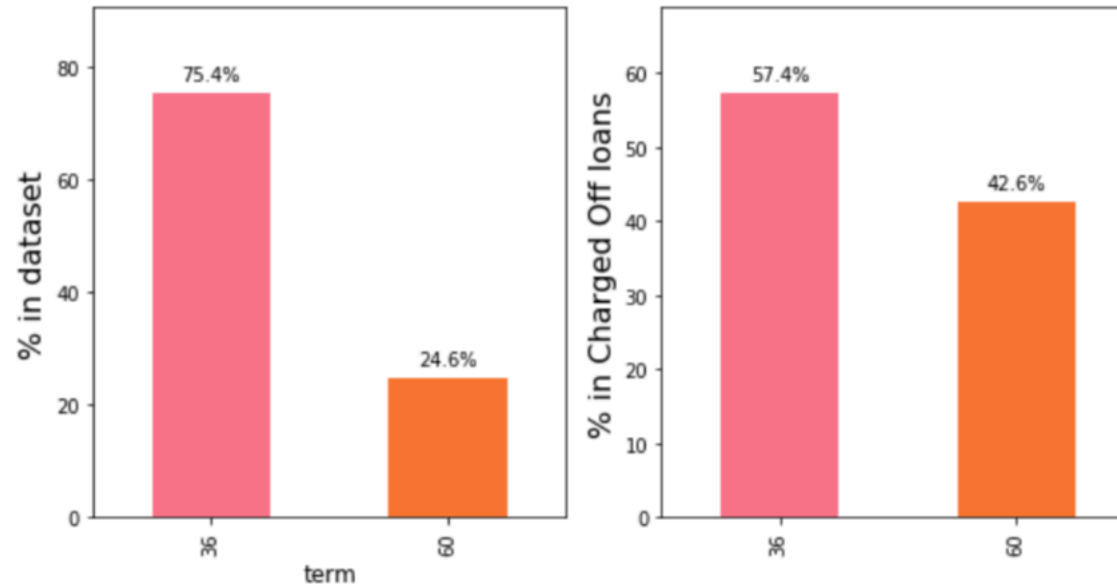


Univariate Analysis – DTI, Grade



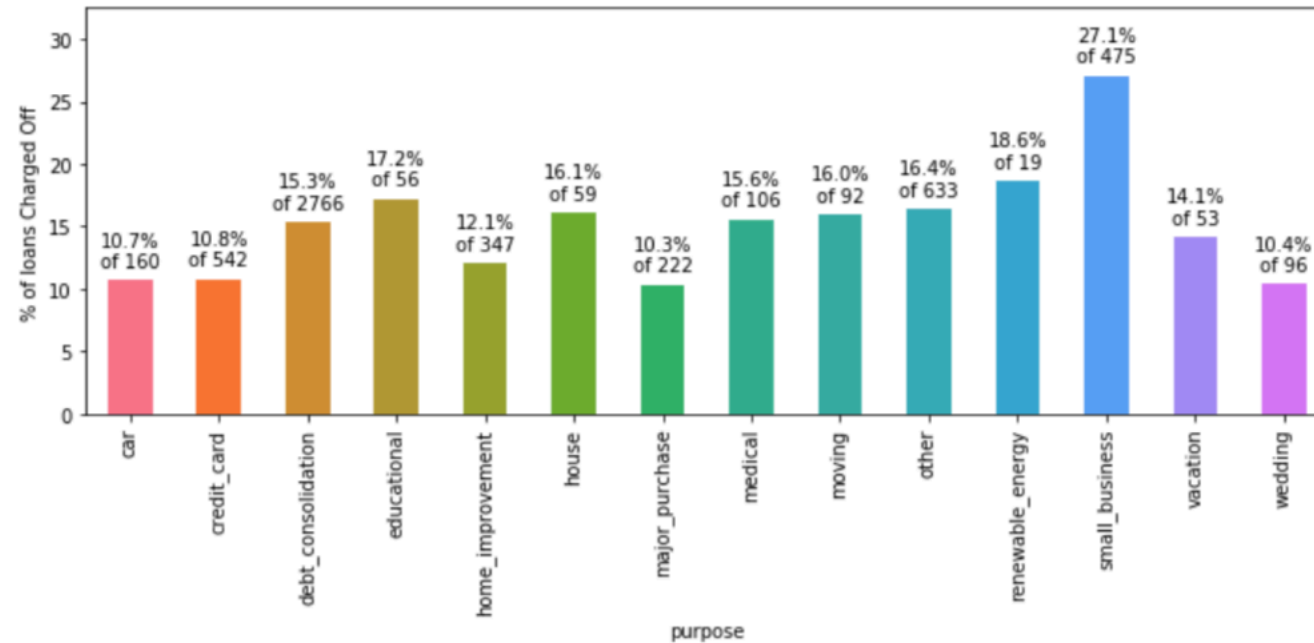
- Debt to Income mean is 13.28 and boxplot shows the median falls between 13 and 14.
- 75% of loans fall below debt to ratio of 20%
- Number of grade A, B and C loans are 76.76% of total loans

Segmented Univariate Analysis – Term



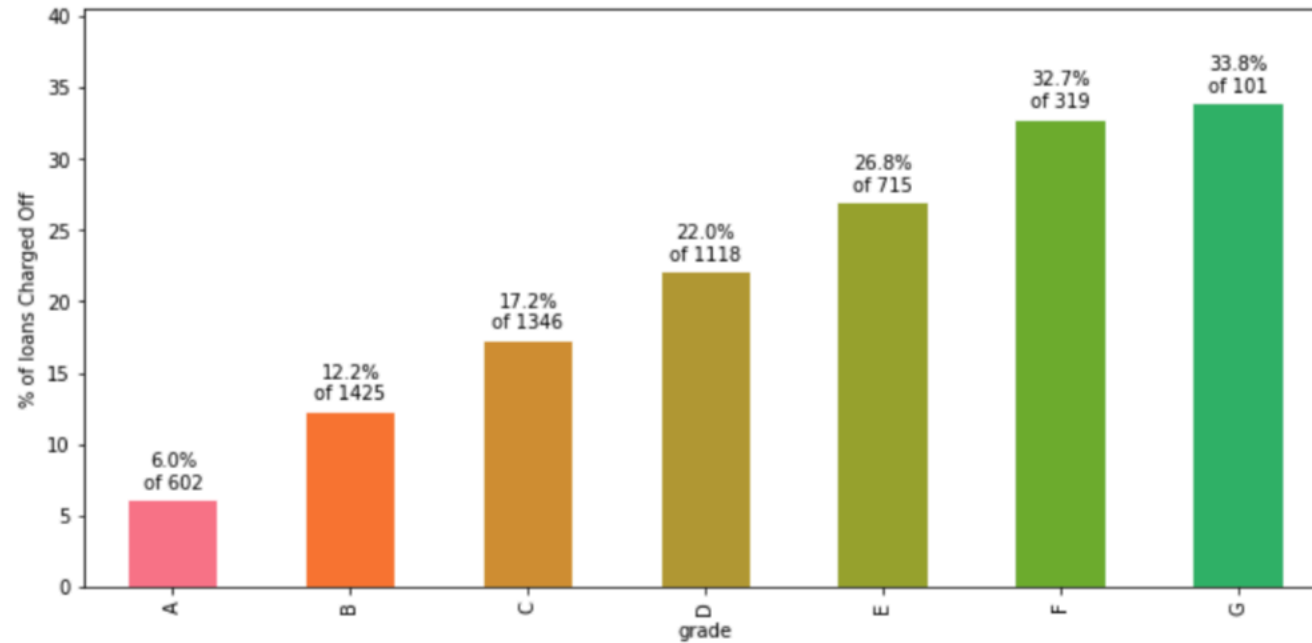
- As we see in above graph, 57.4% of charged off loans have term of 36 month and 42.6% of charged off loans have 60 month as term.
- % of loans 60 month term increases to 42.6% hence loans with 60 month term are more likely to default

Segmented Univariate Analysis – Purpose



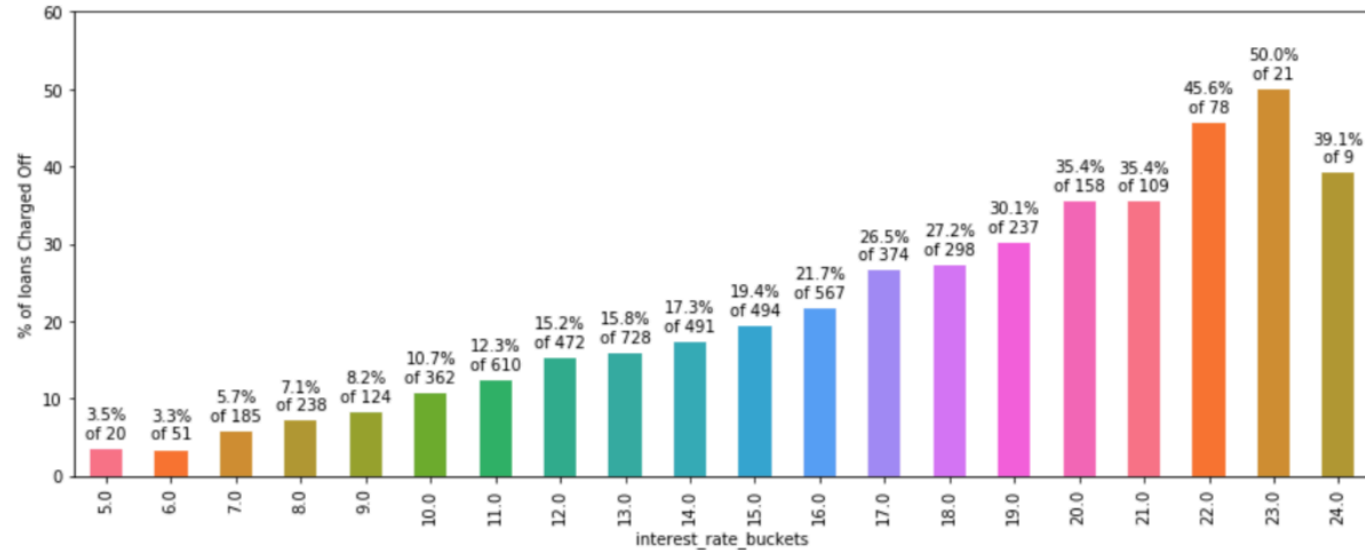
- 27.1% of Small Business Loans are defaulted making them more risky for lending

Segmented Univariate Analysis – Grade



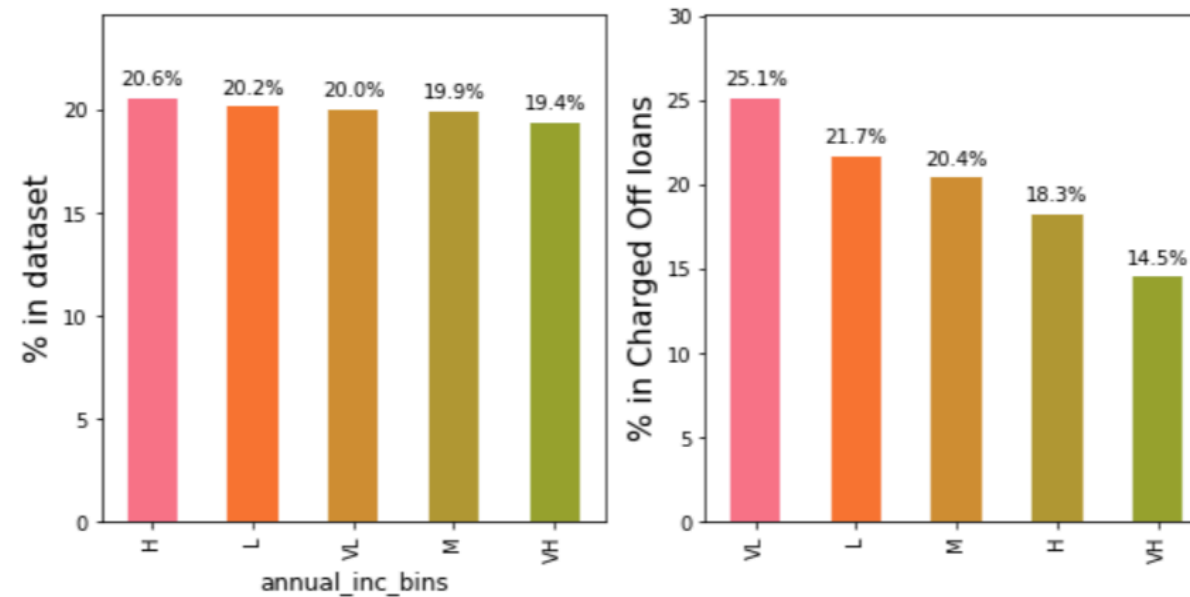
- 33.8% of Grade G loans are defaulted which is highest making them more risky
- 32.7% of Grade F loans are defaulted as well which is second highest

Segmented Univariate Analysis – Interest Rate



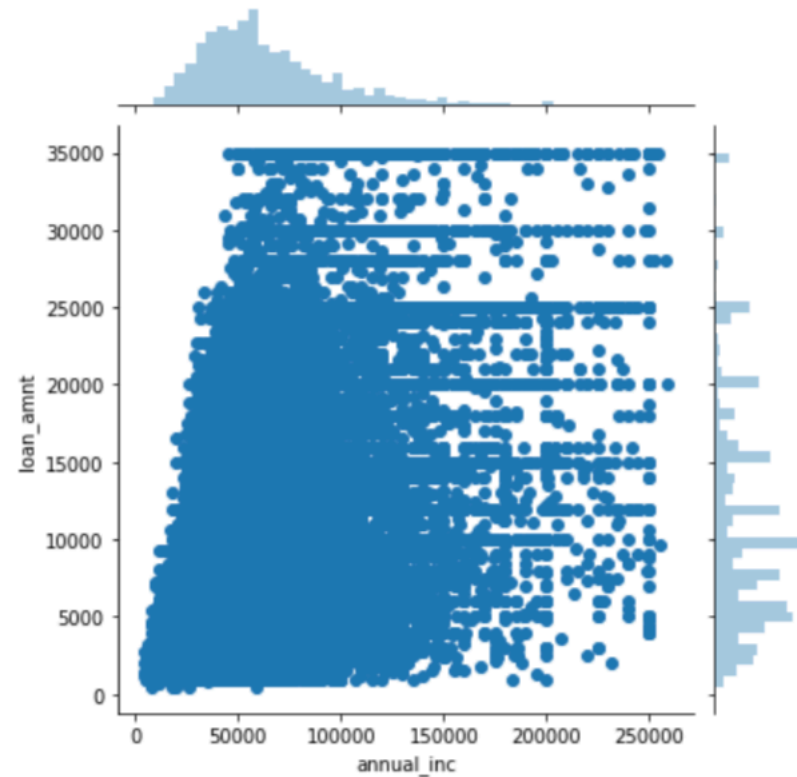
- 50% of loans with interest rate ~23% have defaulted.

Segmented Univariate Analysis – Annual Income



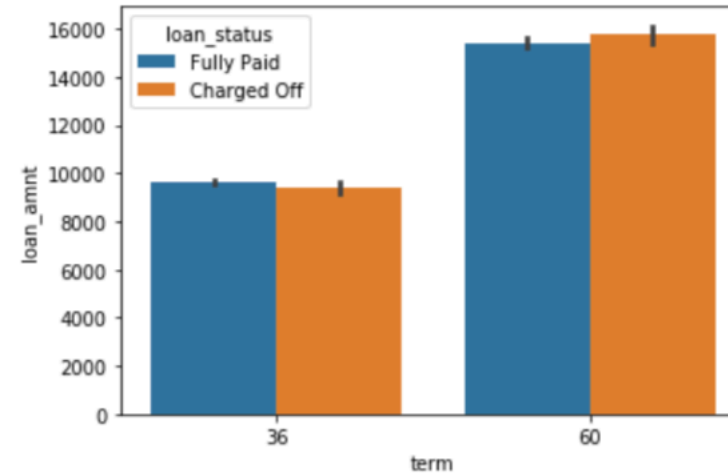
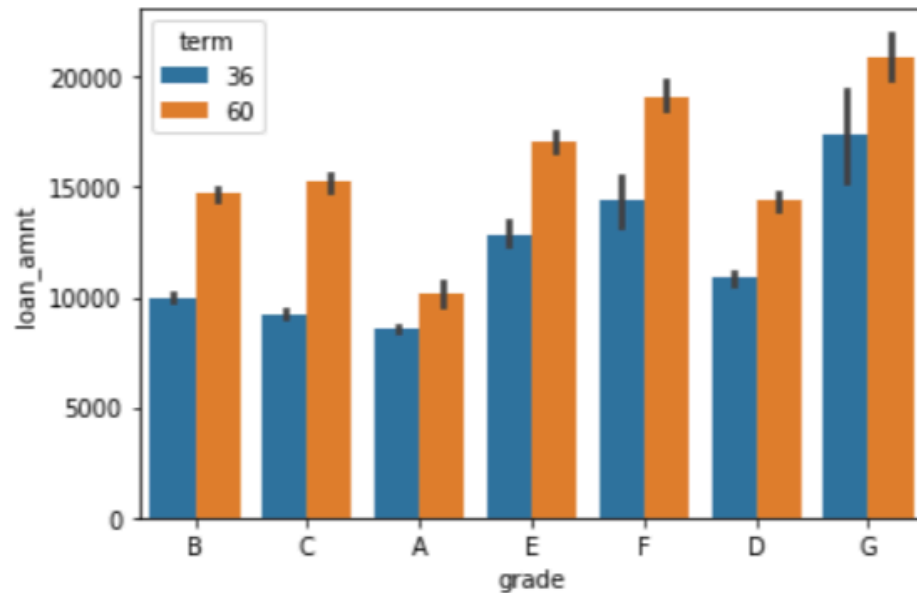
47.2% of VL, L and M annual income customers have charged off loans. Annual income below 60th percentile increases risk of defaulting

Bivariate Analysis – Annual Income vs Loan Amount



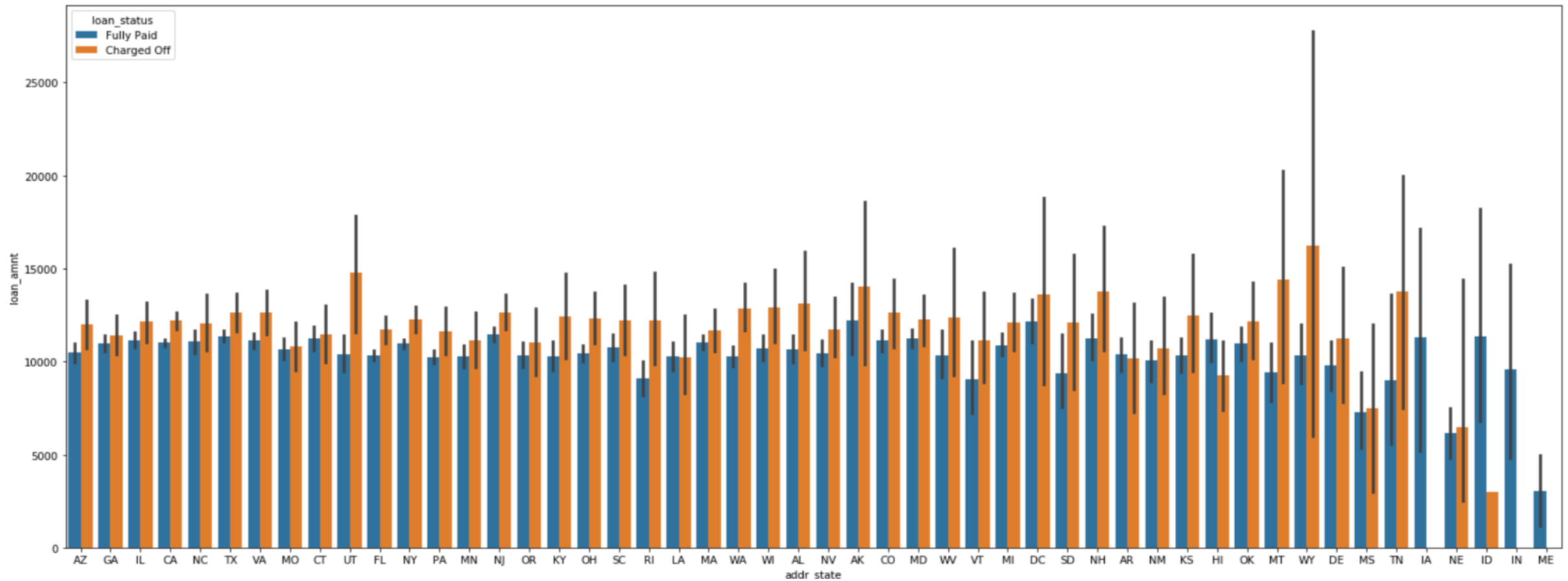
We can see some customers with low annual income have been offered high amount loans which is risky

Bivariate Analysis – Loan Amount, Term, Loan Status/Grade



Higher loan amounts are associated with higher grade and longer term. Longer the term, higher is the charged off rate.

Bivariate Analysis –Loan Status, Loan Amount and Address State



Wyoming has the highest mean loan amount that was charged off which needs further external investigation

Conclusions

- Loans with higher Loan Amount are associated with longer term and more default rate
- Loans taken for Small Businesses are more risky with higher default rate
- Higher the loan amount, lower is the grade with higher default rate. These should have higher interest rates.
- Wyoming has highest loan amount defaulted and it needs to be investigated.