Lending Club Case Study

Submitted By : Durgesh Chaubey

(Batch: Oct_2021)

Index

- Overview on thought process

- Observation during Data Cleaning

- Bivariate Analysis

- Univariate Analysis

- Conclusion

Overview on thought process

▶ Import all Required libraries such as pandas, numpy, matplotlib.pyplot, seaborn, datetime, warning, etc.

▶ Data cleaning to be done with respect to null, unique and missing value.

▶ Identify and delete column which is not required for analysis.

▶ Identify and convert datatype, category into required one.

▶ Perform reverse EDA, i.e. first multivariate, then Bivariate and then Univariate Analysis.

▶ Conclude based on curve, plots and comparison. Provide driving variables of loan defaulters to company.

Observation during Data Cleaning

▶ Available data before cleaning is **39717** rows and **111** columns

▶ After removing null, unique, missing values and non useful columns data reduced to before cleaning is **38577** rows and **18** columns

Bivariate Analysis

loan Status vs fund amount invested
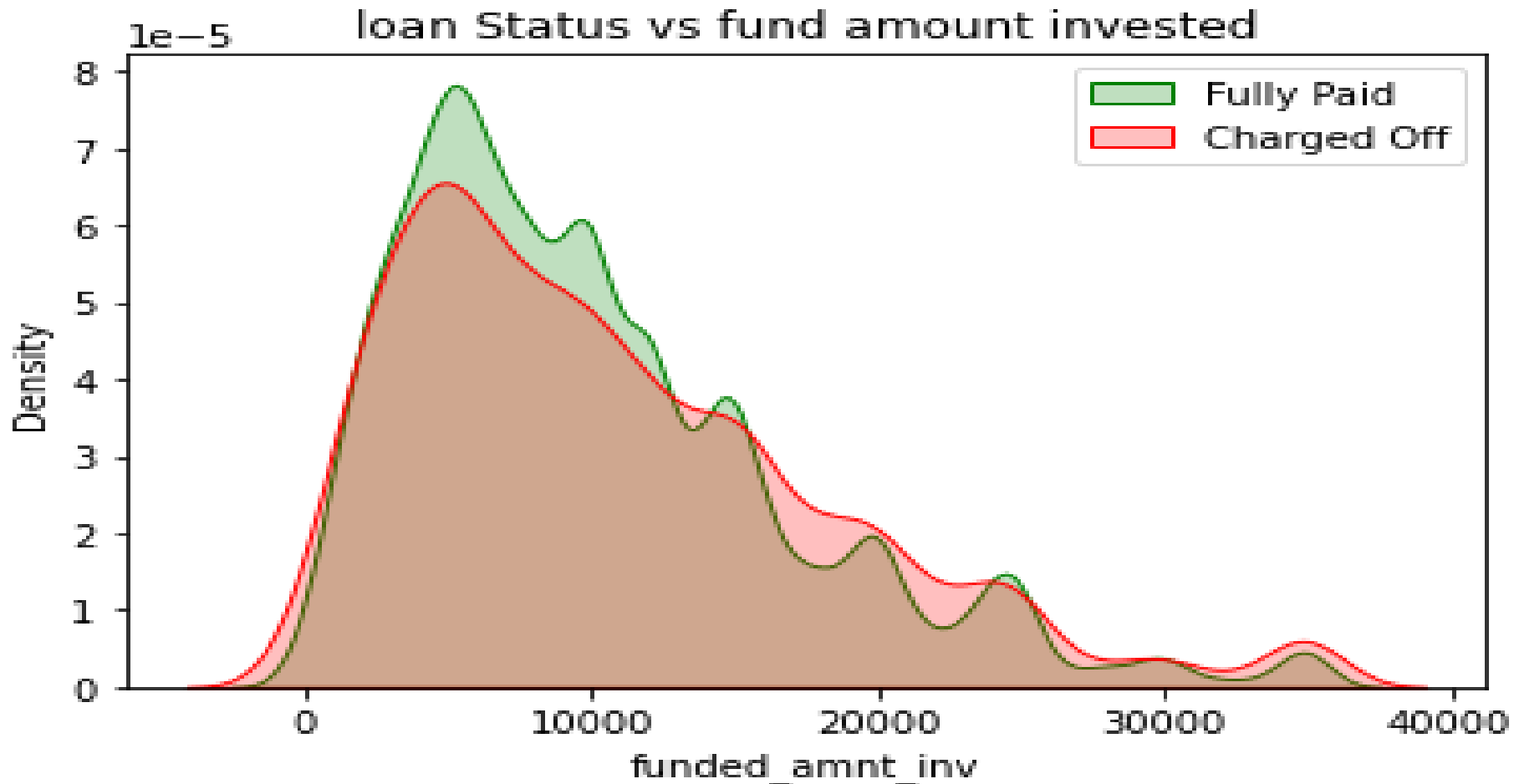
▶ With increase in funding amount invested, there is a slight increase in density of people becoming defaulter.

# Comparison of Interest Rate and Grade vs loan status

Defaulters proportion vs Loan Grade
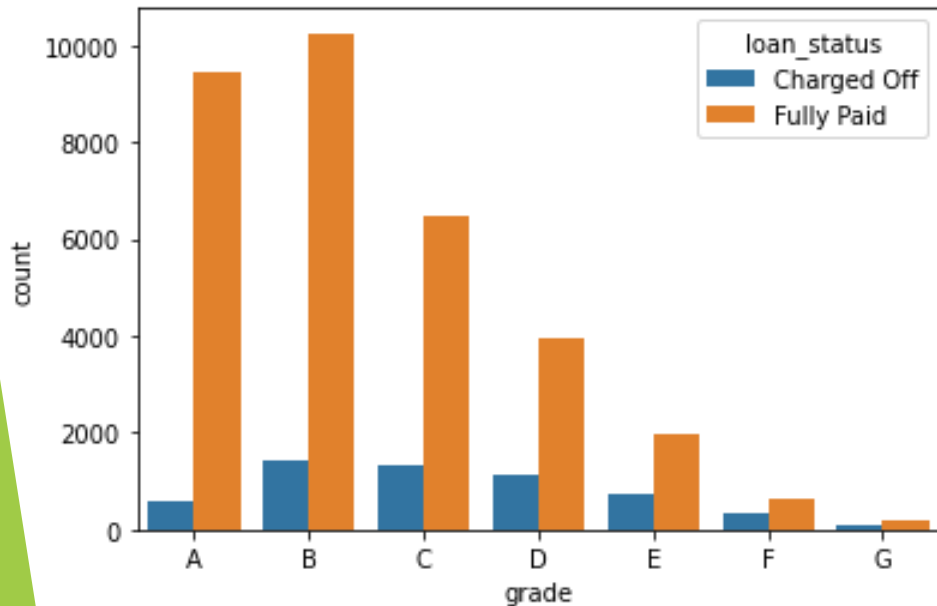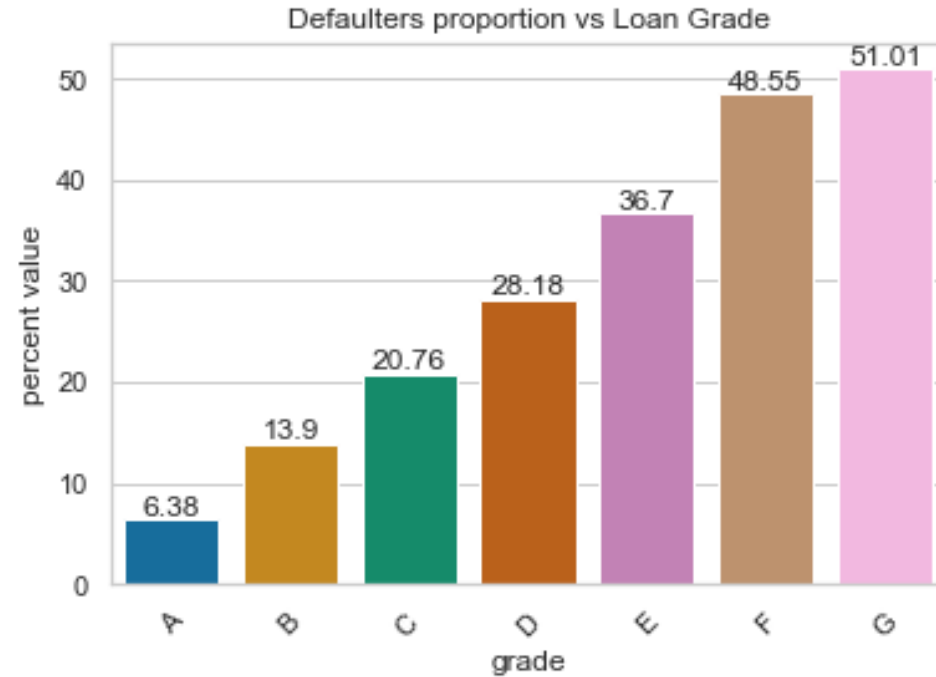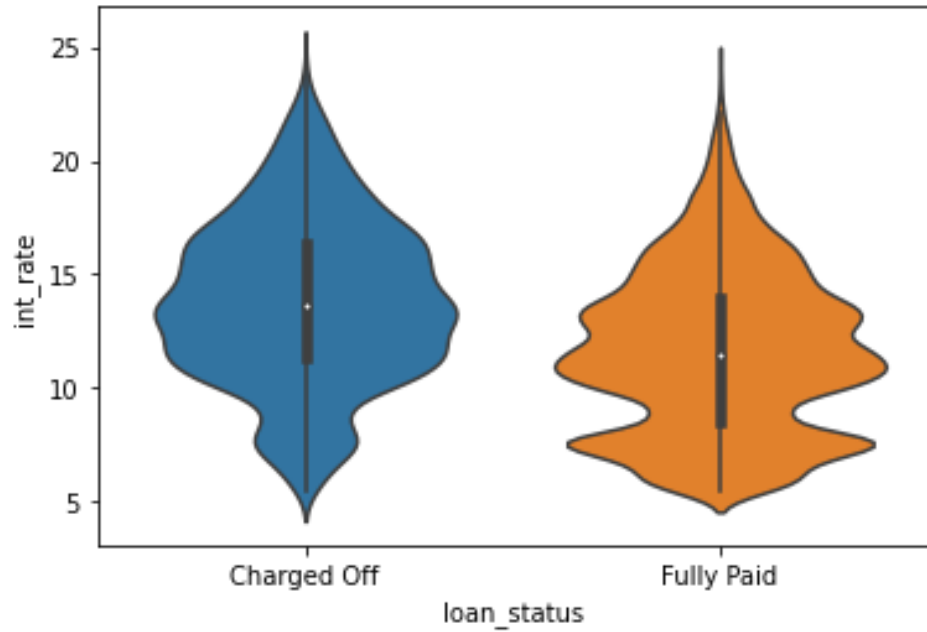

Conclusion from above plots are:

1. Lower the interest rate lower the number of defaulters.

2. The number of defaulters started increasing from interest rate 12 to 13 percent .

3. Around 50 percent of people defaulted against grade F and G.

Employee Work Experience vs Loan Status
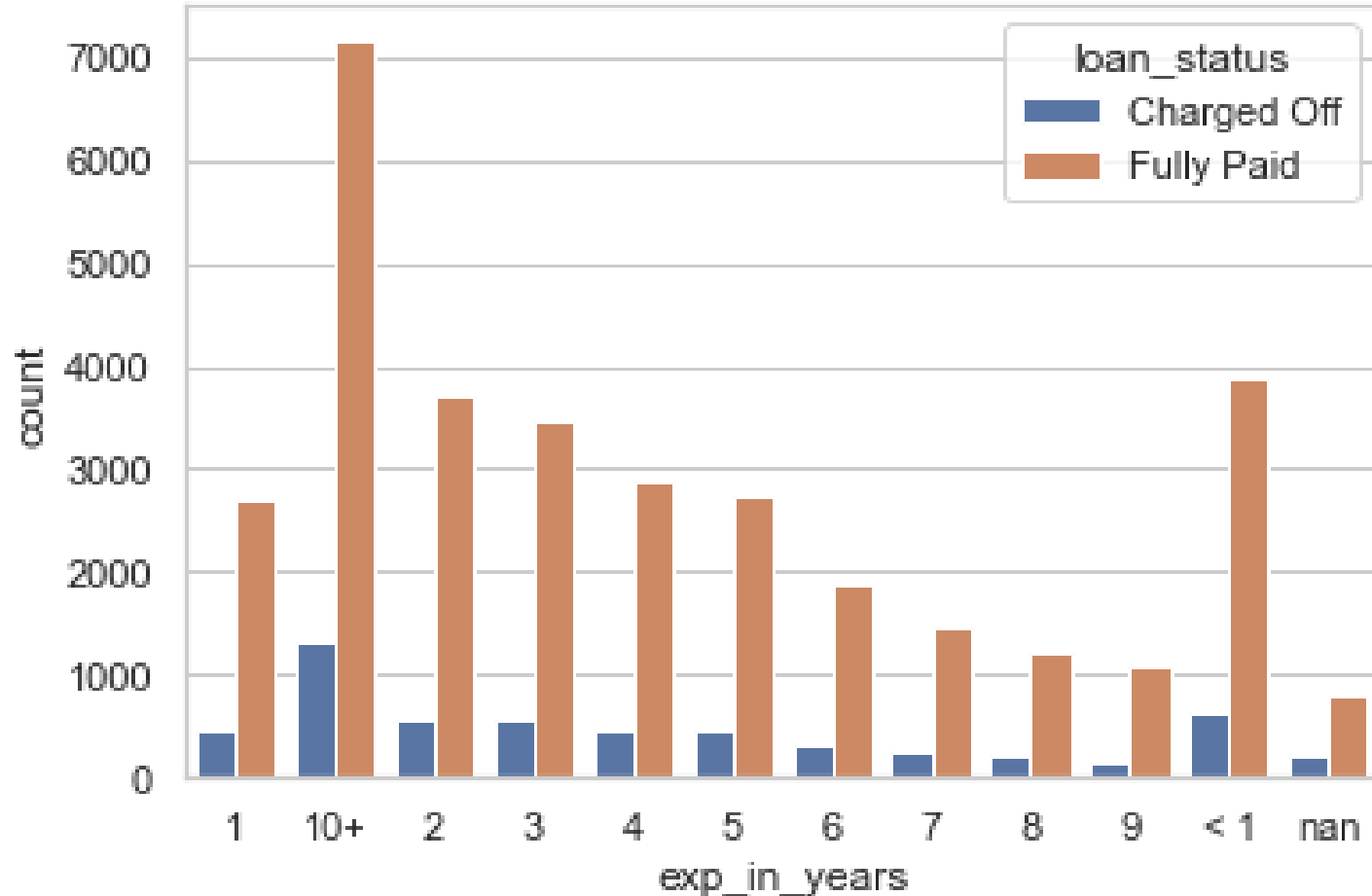
There is not a single line judgement regarding relation of work experience vs loan defaulter. We have a bracket of one to ten years where people are less defaulter. People with less then one year of experience and more then ten years of experience are more prone to loan default.

Home Ownership vs Loan Status
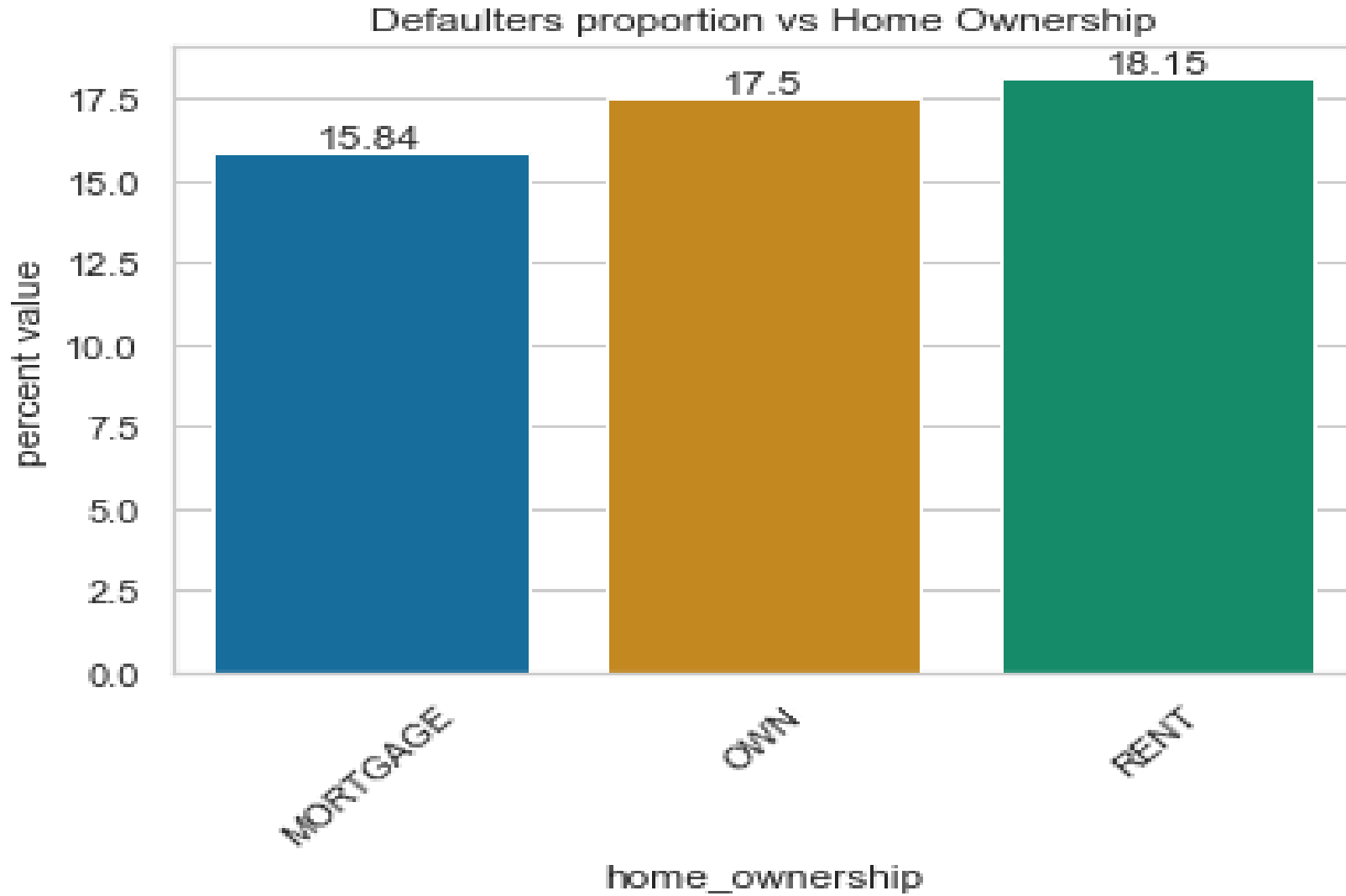
▶ As per plot rented people are more prone to loan default than owner and mortgage are least prone to loan default.
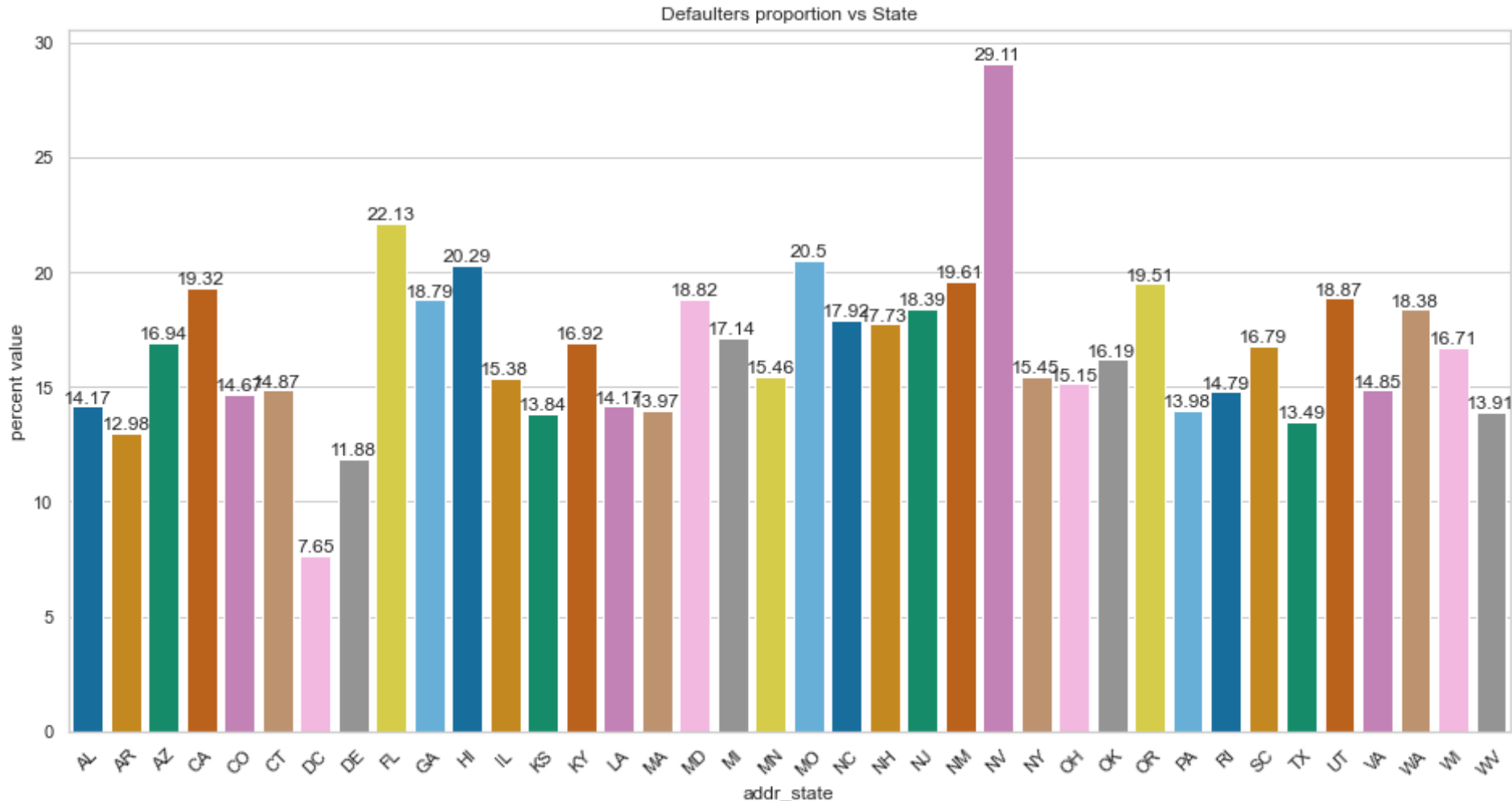
Loan purpose vs Loan Status

- Loan for Small business have max. no's of default (37.09%)

- Renewable energy (22.89%) & Education (21.05%).

- Loan for car, credit card, major purchase and wedding show are least prone to default

Defaulters proportion vs Purpose of loan

State vs Loan Status

Defaulters proportion vs State
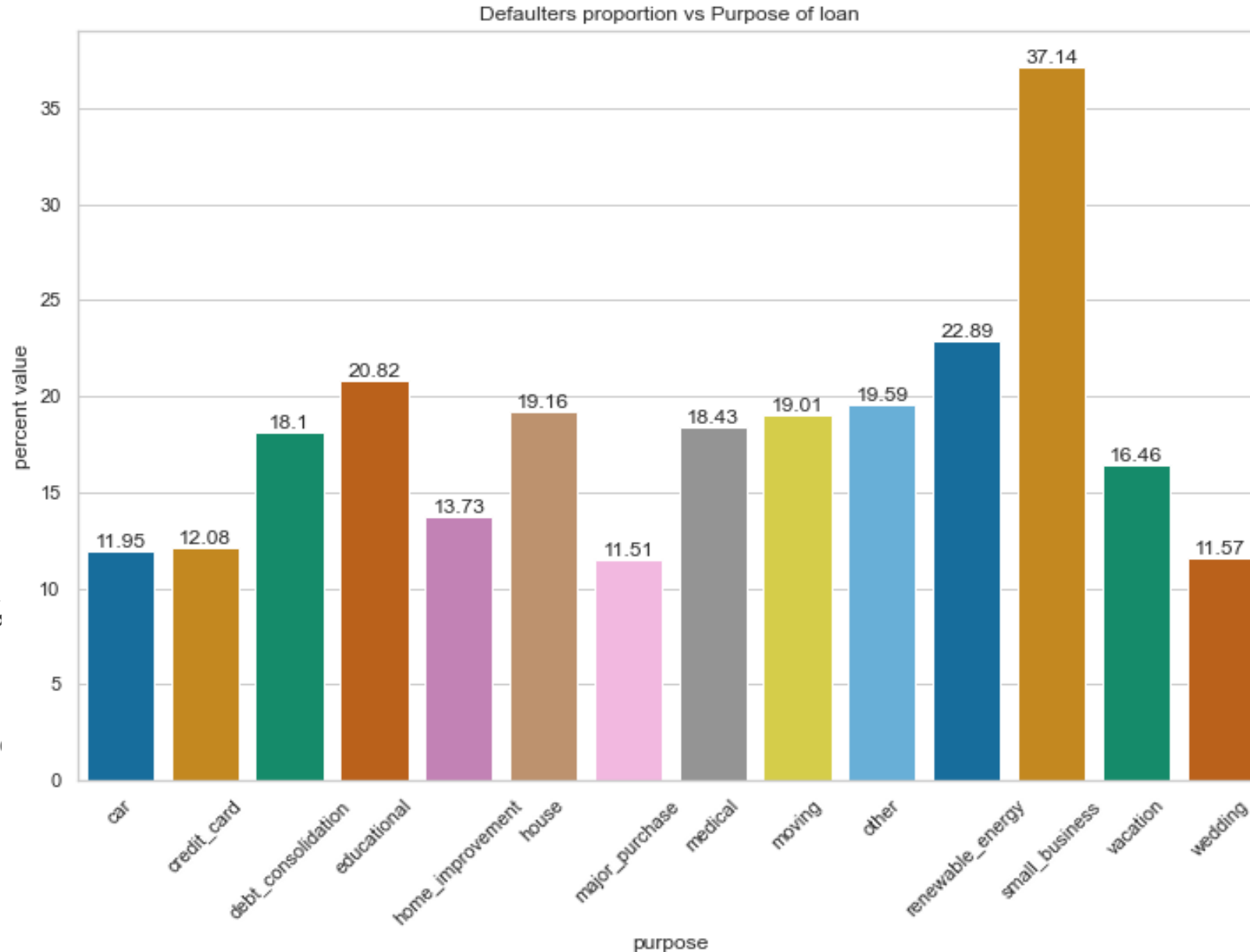
▶ Plot Says State Nevada(NV) have max default (29.19%)

▶ States as Florida (22.09%), Hawaii(20.44%) & Missouri (20.32%) have major default.

▶ Washington DC (7.73%) & State of Delaware (10.89%) have min default

Univariate Analysis

Loan amount distribution
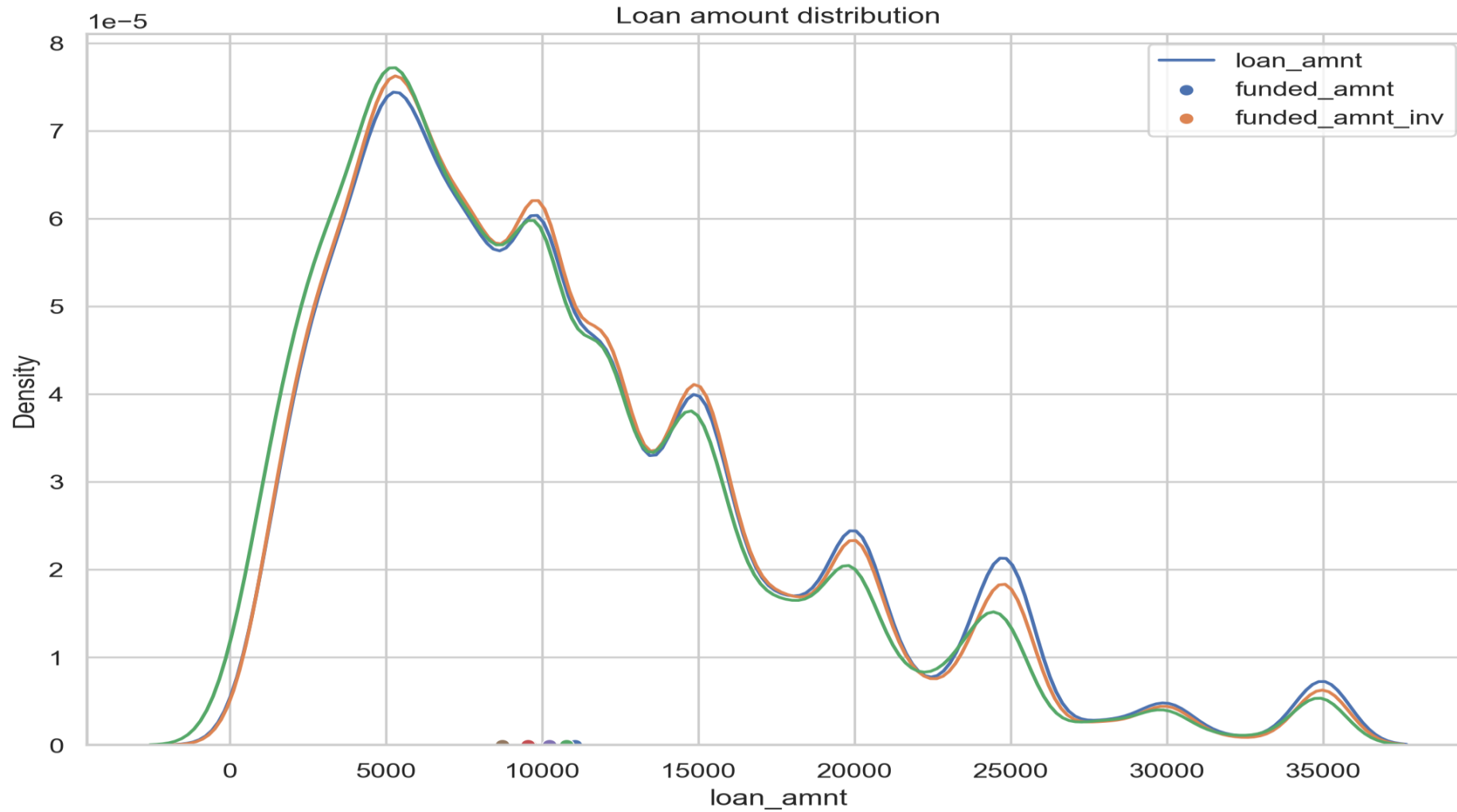
Amount of loan given to maximum no of people is 5000 USD.

Conclusions

- Avoid approving loan amount more than 10,000 USD.

- Keep interest rates below 12% .

- Immediately stop F and G grade loans approval.

- Loan to fresher to company is not good idea and also avoid person with more than 10 years in same company

- Rented applicant is have slightly higher probability to default, prefer one who is owner or mortgage.

- Avoid loan application for small business, renewable and education. Prefer car loan, credit card payment and wedding shows etc.

- Avoid loan application from Nevada, Florida , Hawaii & Missouri states. Prefer Washington DC and State of Delaware instead.

- Loan applicant with delinquency is prone to default loan, avoid approval.

- Make sure to note down date along with month for record for better analysis.

thank you