Lending Club Case Study

Prepared By
Rohinee Kanta Mishra
Arun Kumar

Agenda

- Problem Statement

- Analysis Approach

- Data Preparation

- ED Analysis

- Conclusion

Problem Statement

Lending Club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.

The 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.

As a data scientist working for Lending Club analyze the dataset containing information about past loan applicants using EDA to understand how consumer attributes and loan attributes influence the tendency of default.

Analysis Approach

| Clean data	Univariate Analysis	Segmented Univariate Analysis	Conclusion
->Drop columns with null values, all random values or single category value ->Convert values to proper int, float, date representations	->Check distributions and frequencies of various numerical and categorical variables ->Create derived variables	->Analyze variables against segments of other variables,	->Publish insights and observations

Analysis – Understanding Loans

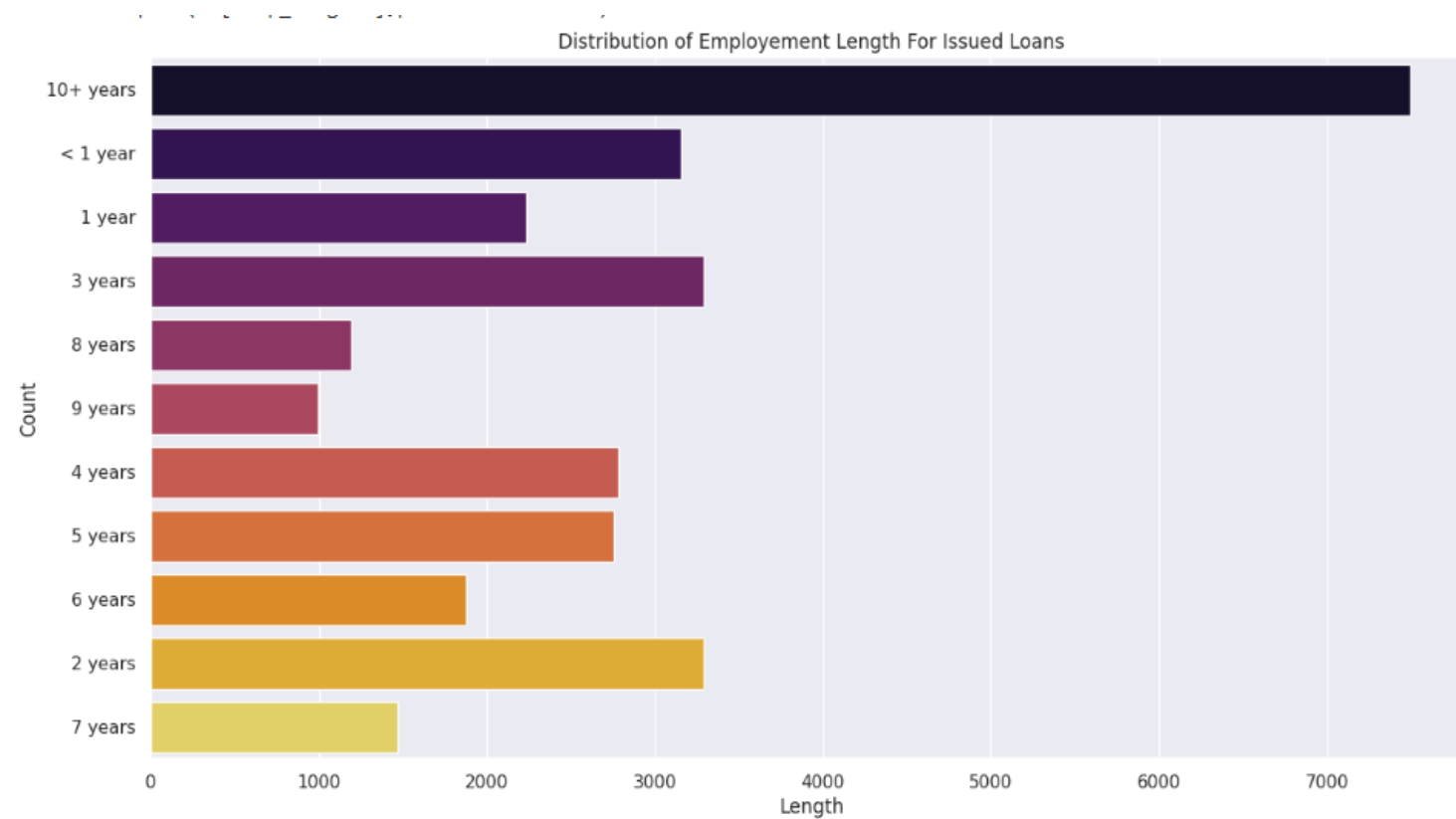# Most of the borrower are able to pay their Loan on time i.e. 85.4% are Fully paid and 14.6 % are defaulter
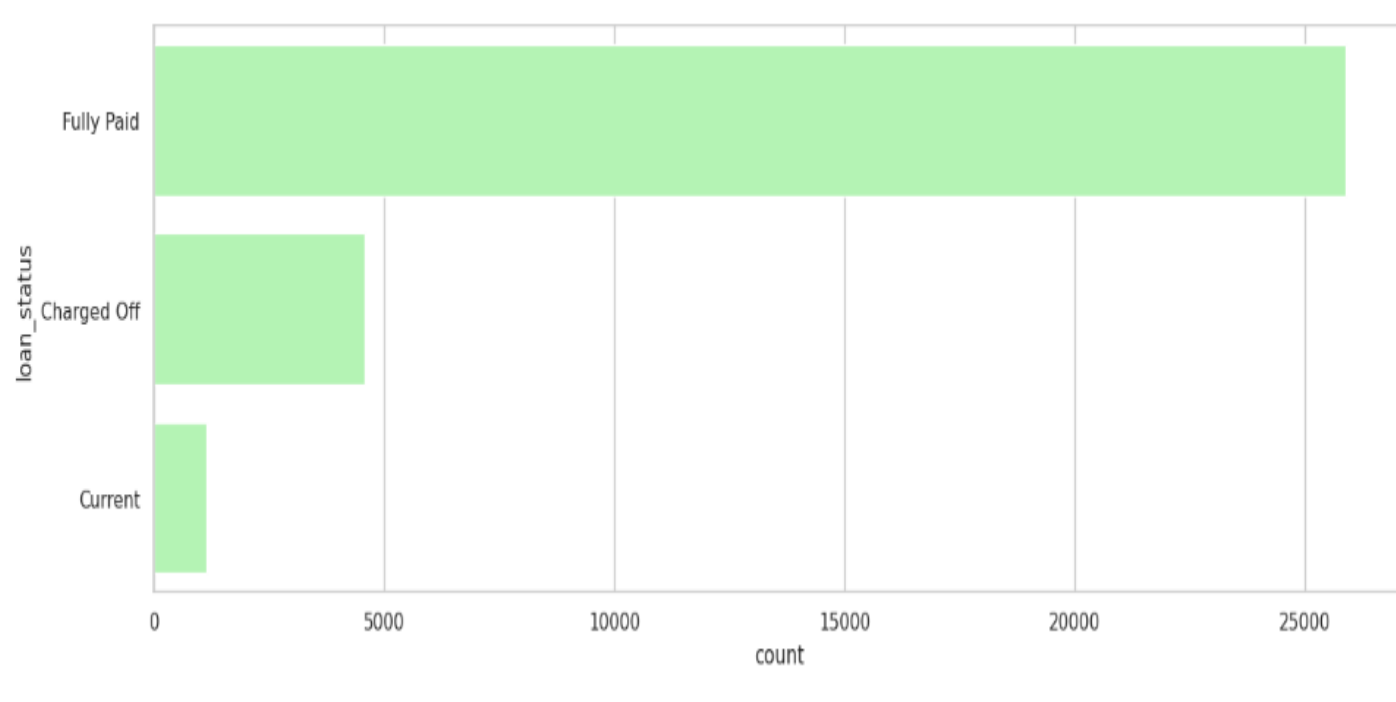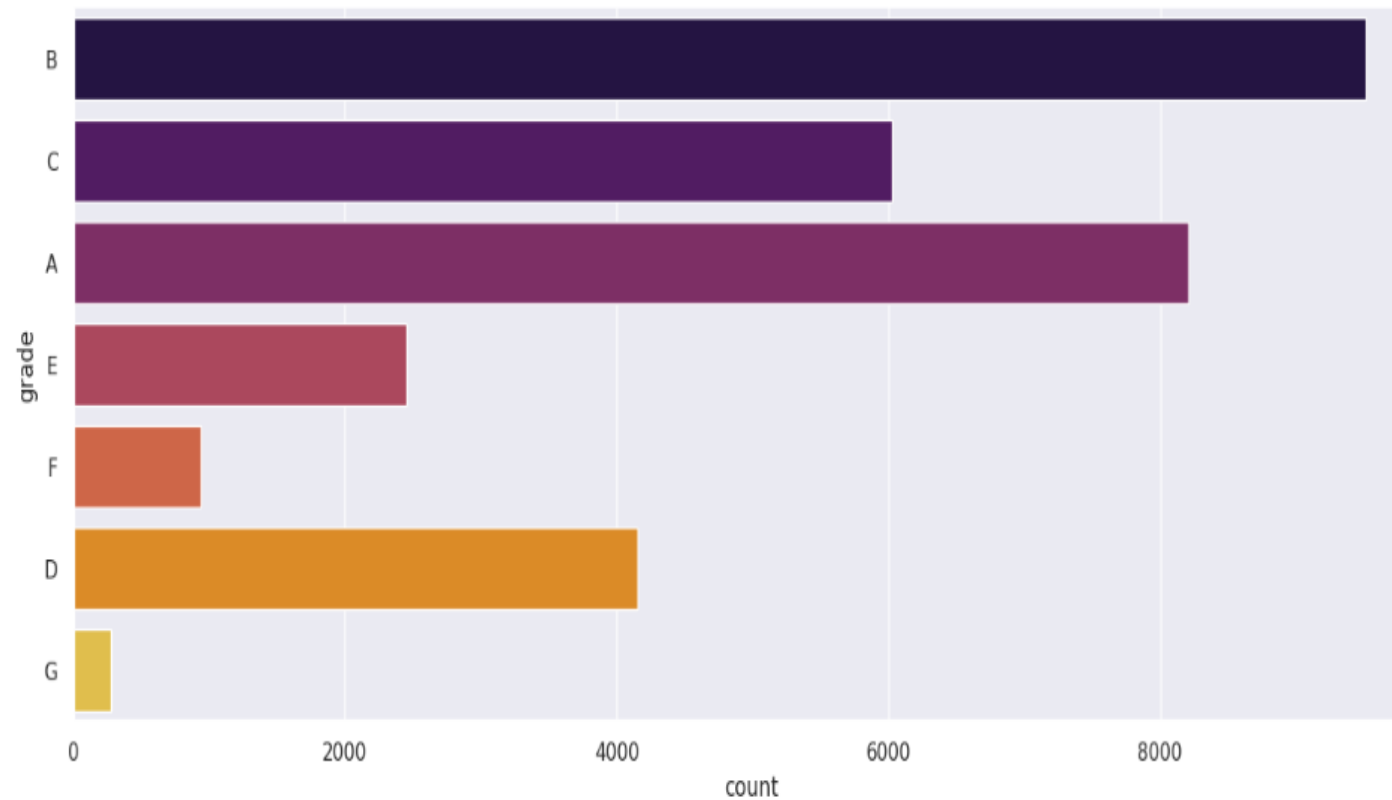
Analysis - Understanding Loans
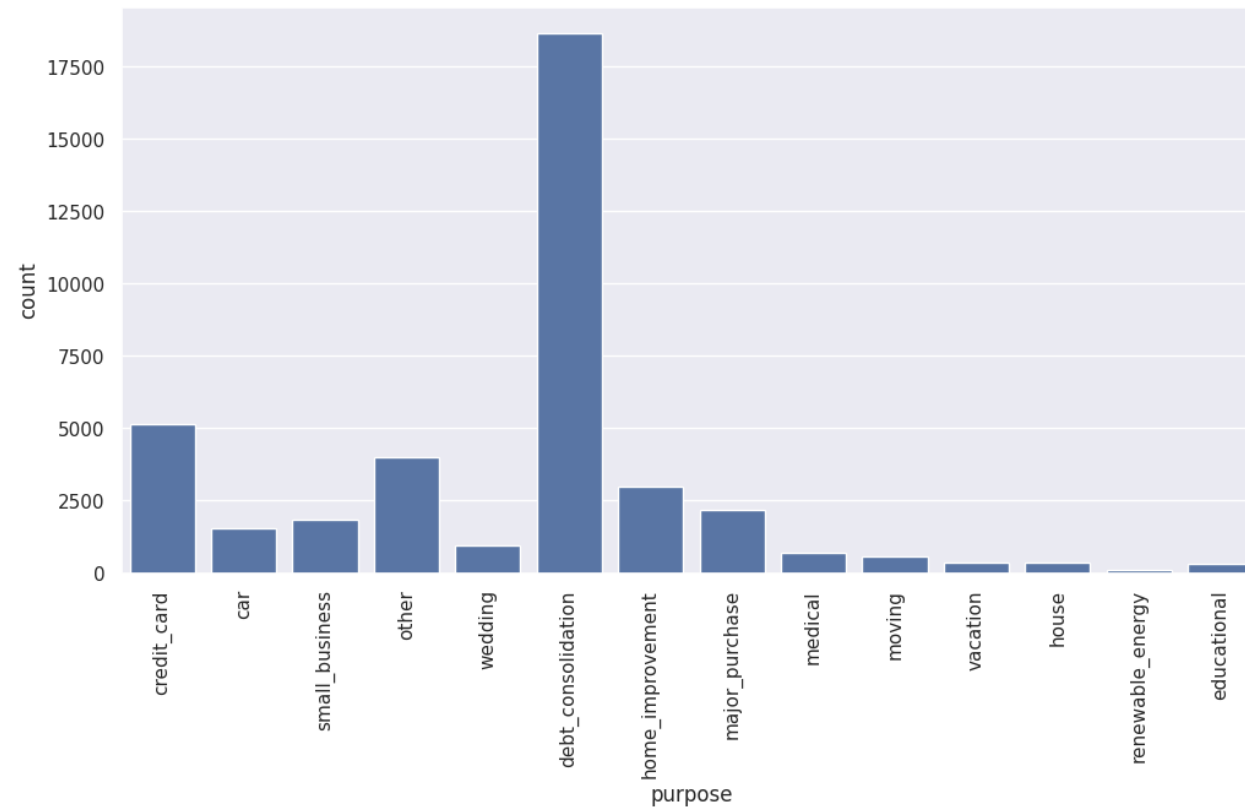
Most loans were taken by customers for debt_consolidation and credit card

Analysis - Understanding Loans

Distribution of Employement Length For Issued Loans

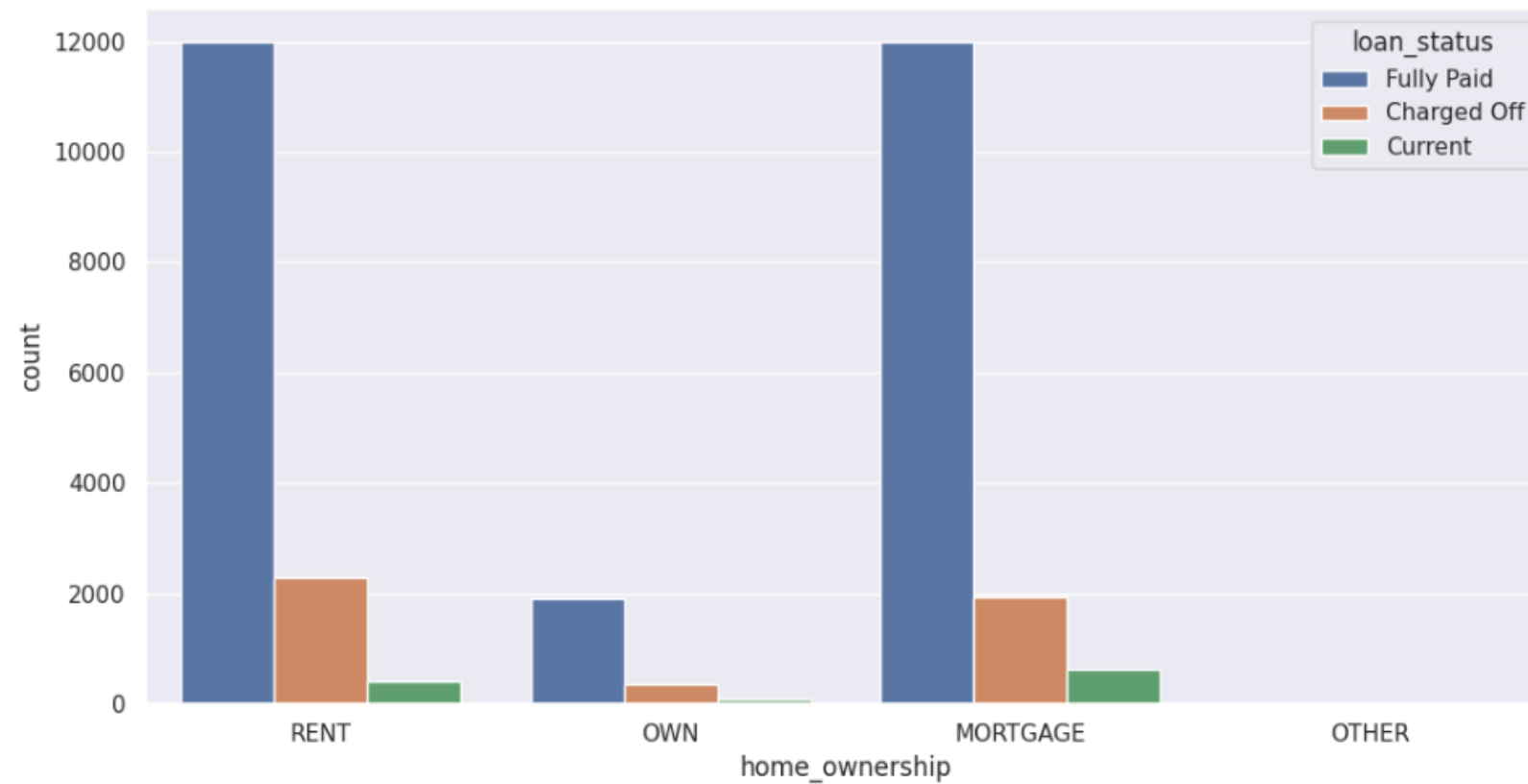More Loans are taken by employee with more than 10 years of experience

Analysis - Understanding Loans

Grade A,B,C seems to be applied for most loans compared to other grades
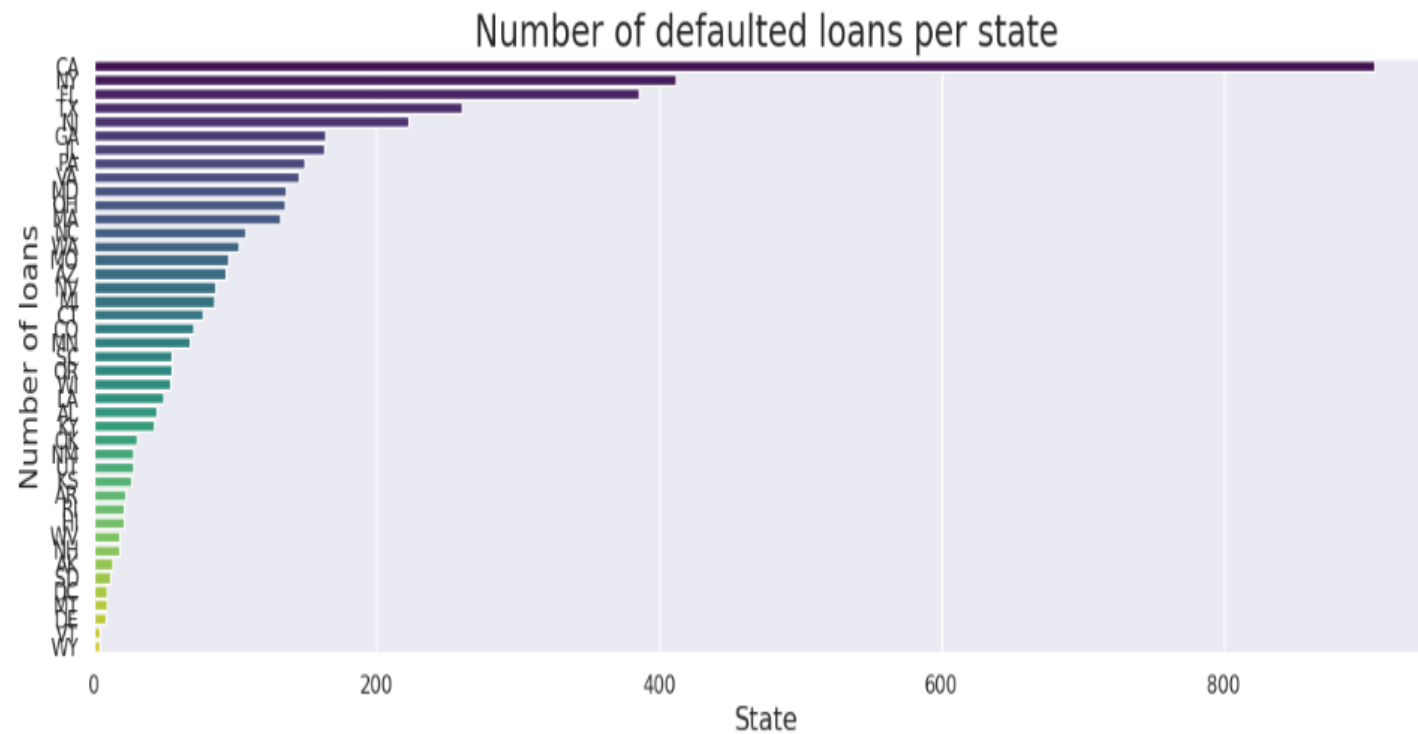
With Grade F and G being the lowest borrower
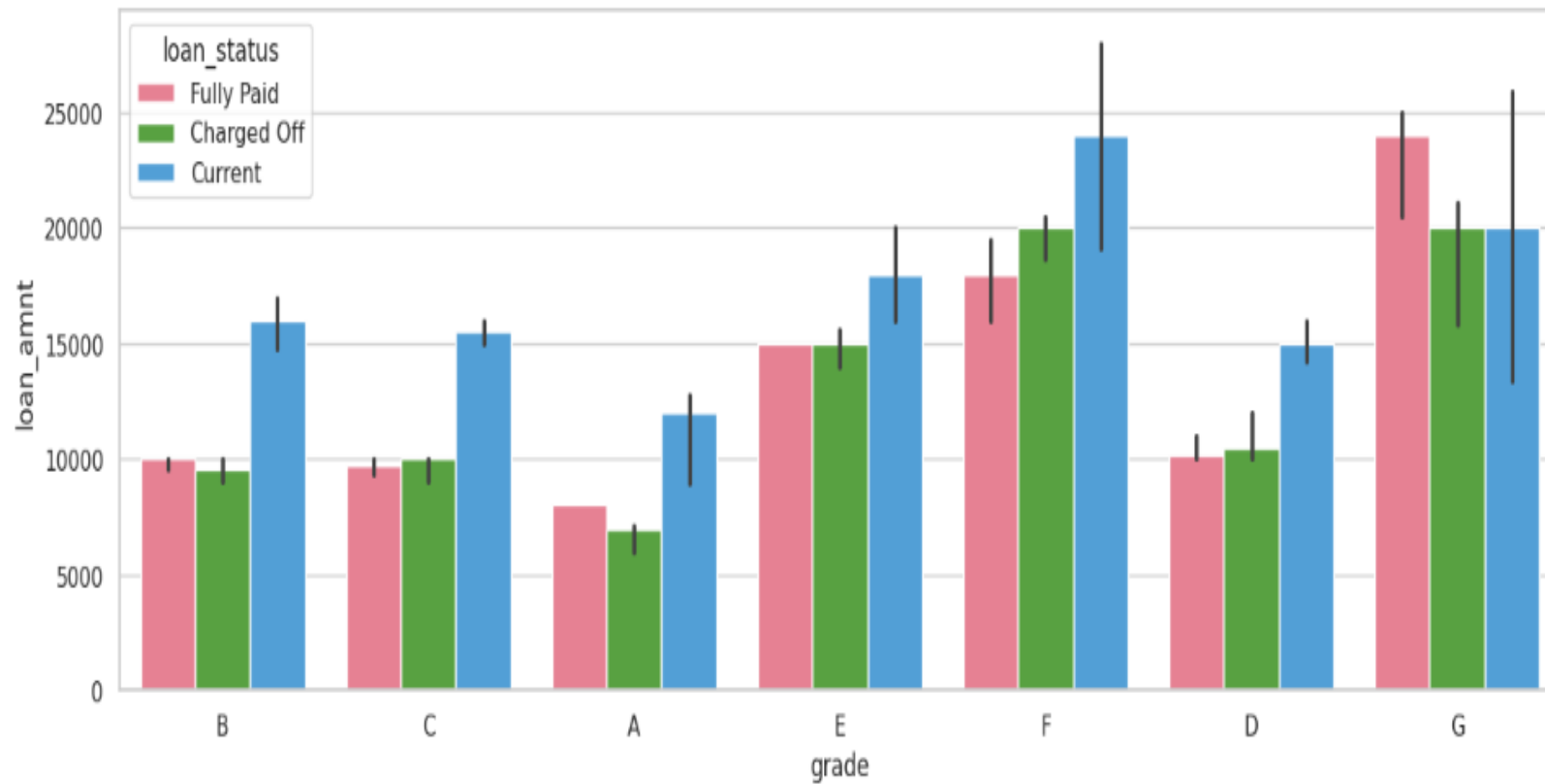
Analysis - Understanding Loans

People who has Mortgaged or rented has the highest amount of Loan

Analysis - Understanding Loans
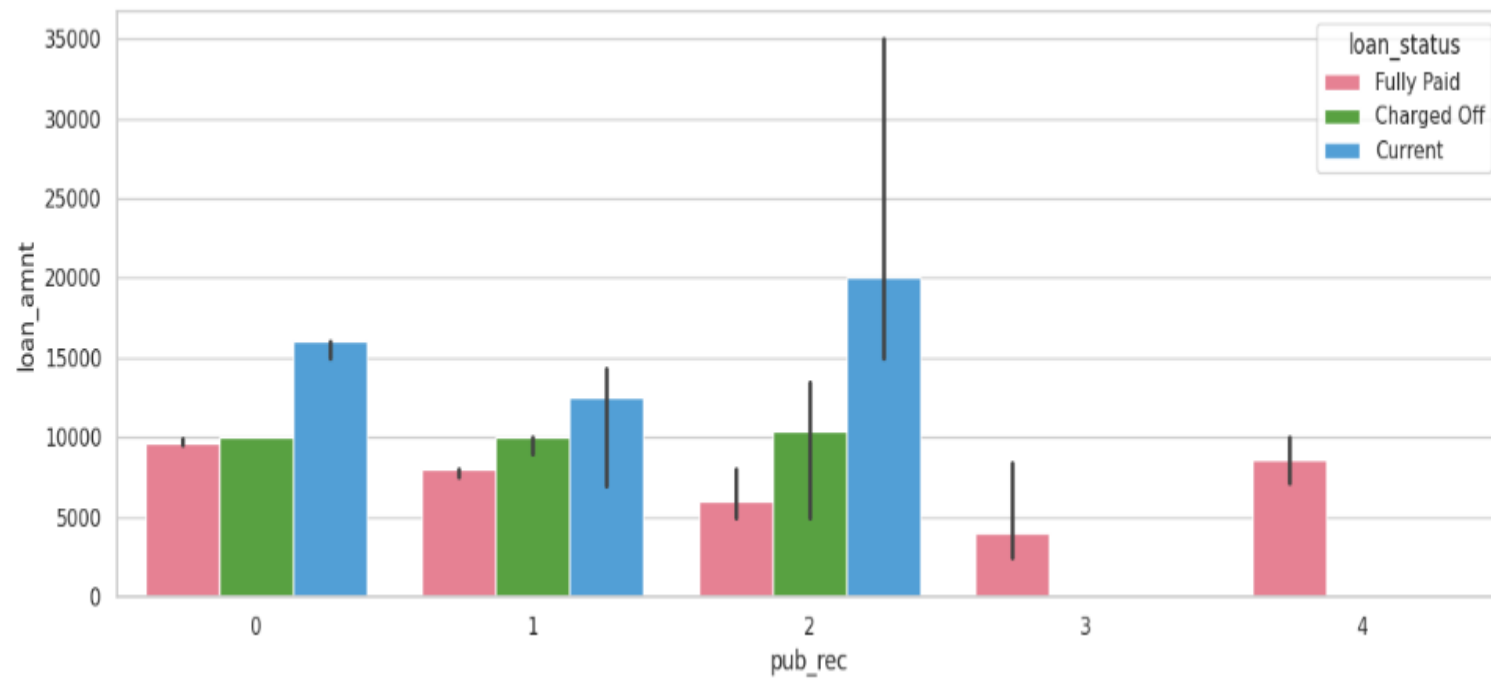

Number of defaulted loans per state

Shows top cities who have taken the highest
Loans I.e. CA,NY and FL
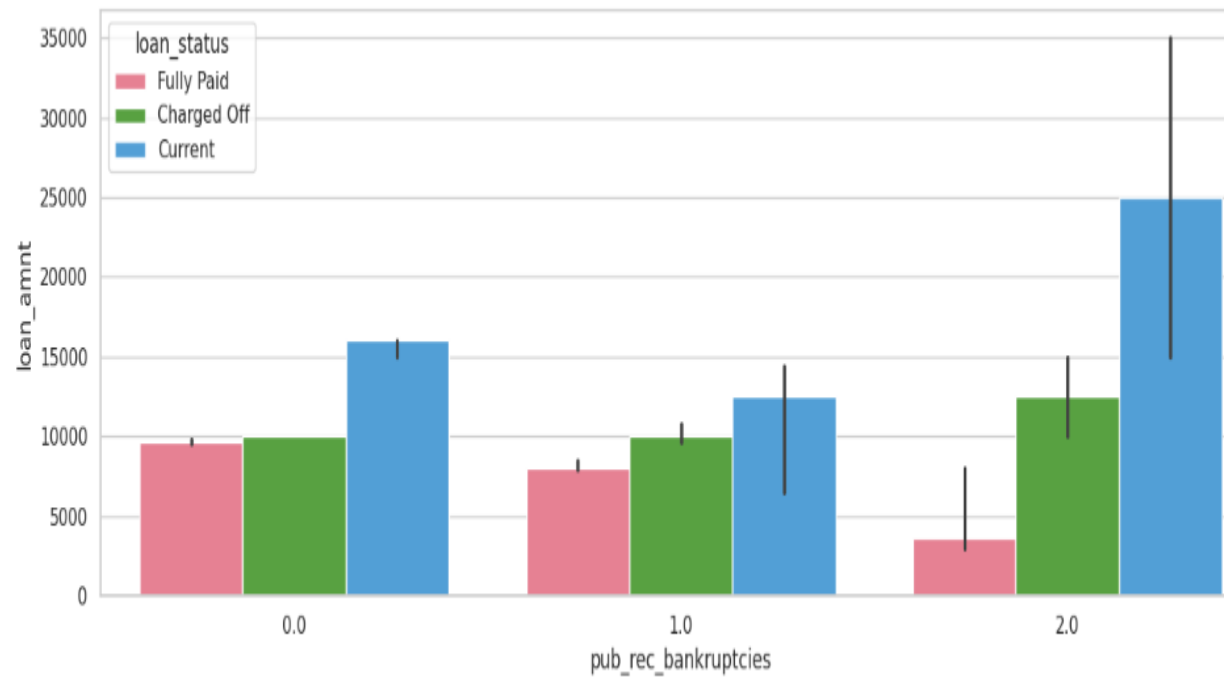
Analysis – Default by Grades

It shows that even if grades 'F' & 'G' were lowest loan borrowers, they were still the highest defaulters.
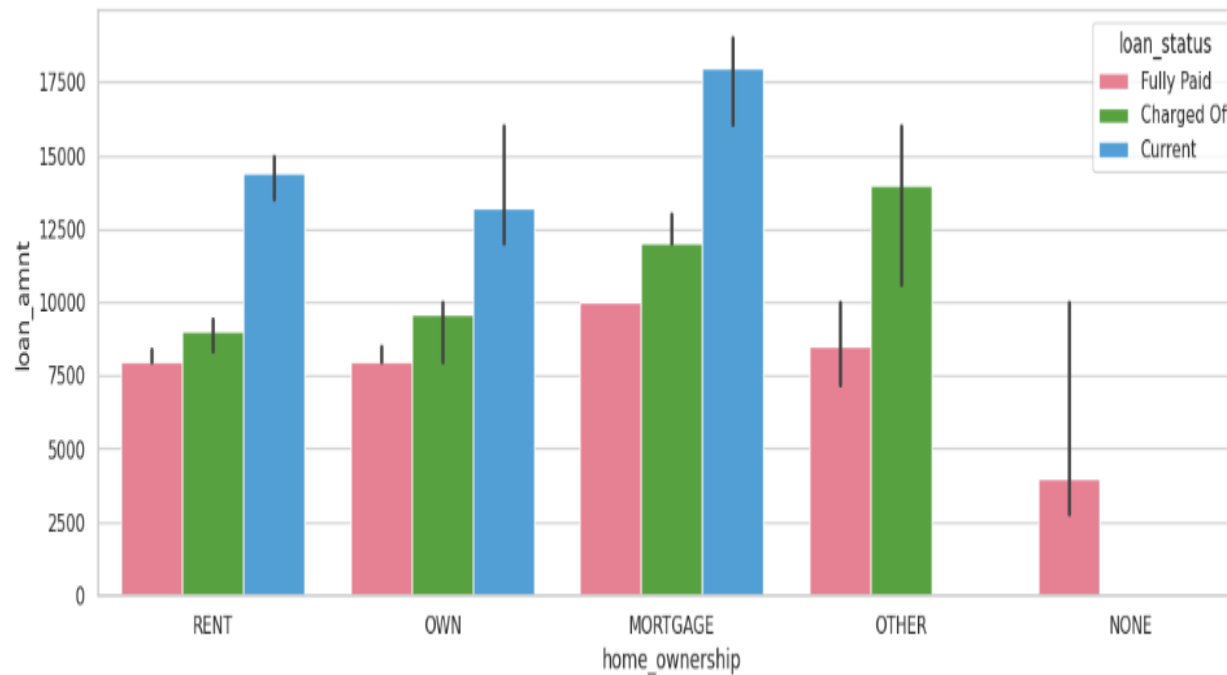
Analysis – Default by Public Records

It clearly shows that if a person has a number of derogatory public records, then there is high possibility of him being defaulted.
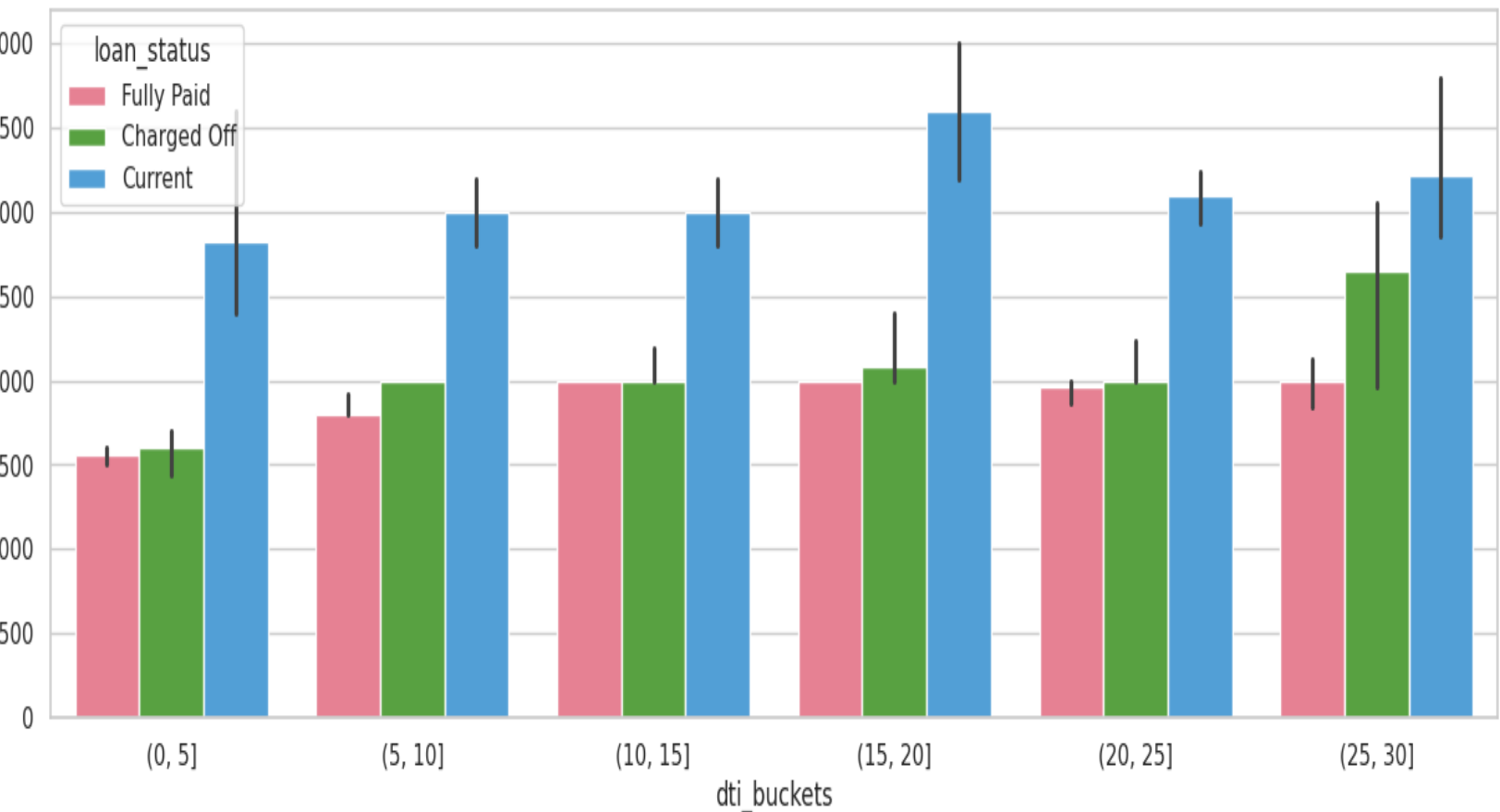
Analysis – Default by Bankruptcies

Customers who have higher records of bankruptcies are bound to be charged-off more.
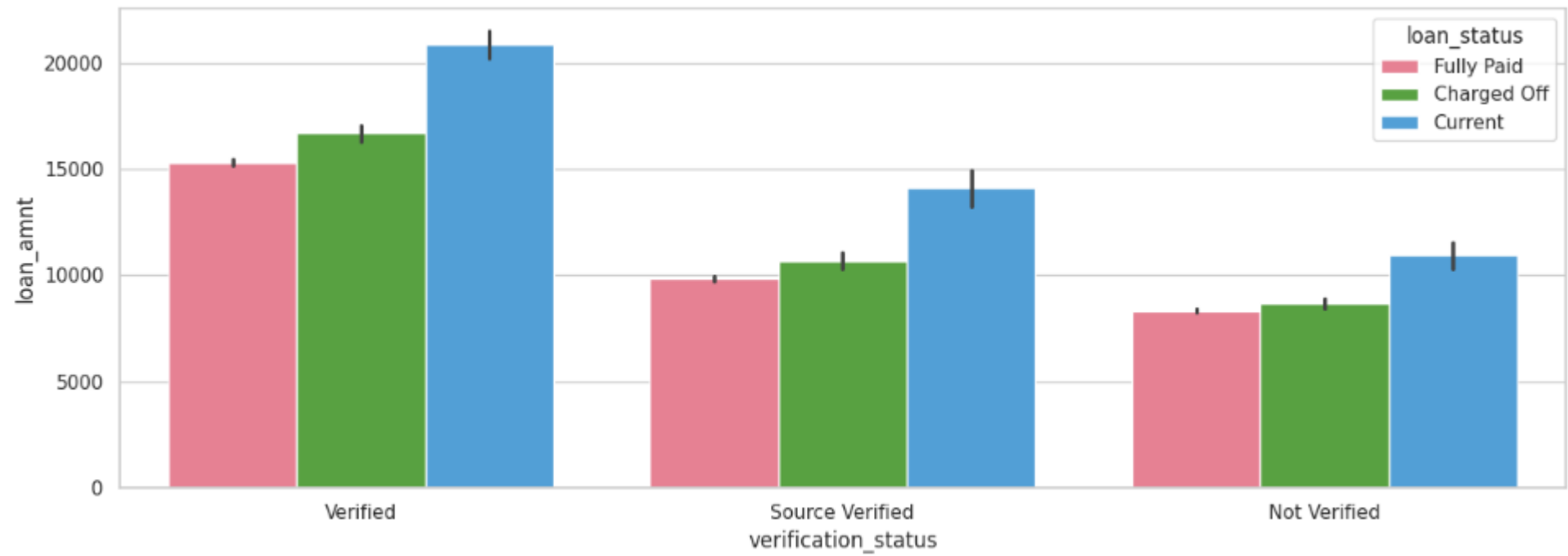
Analysis – Default by home ownership

Customers who don't seem to have a fixed home ownership are likely to default more. 'Other' is having high Defaulter
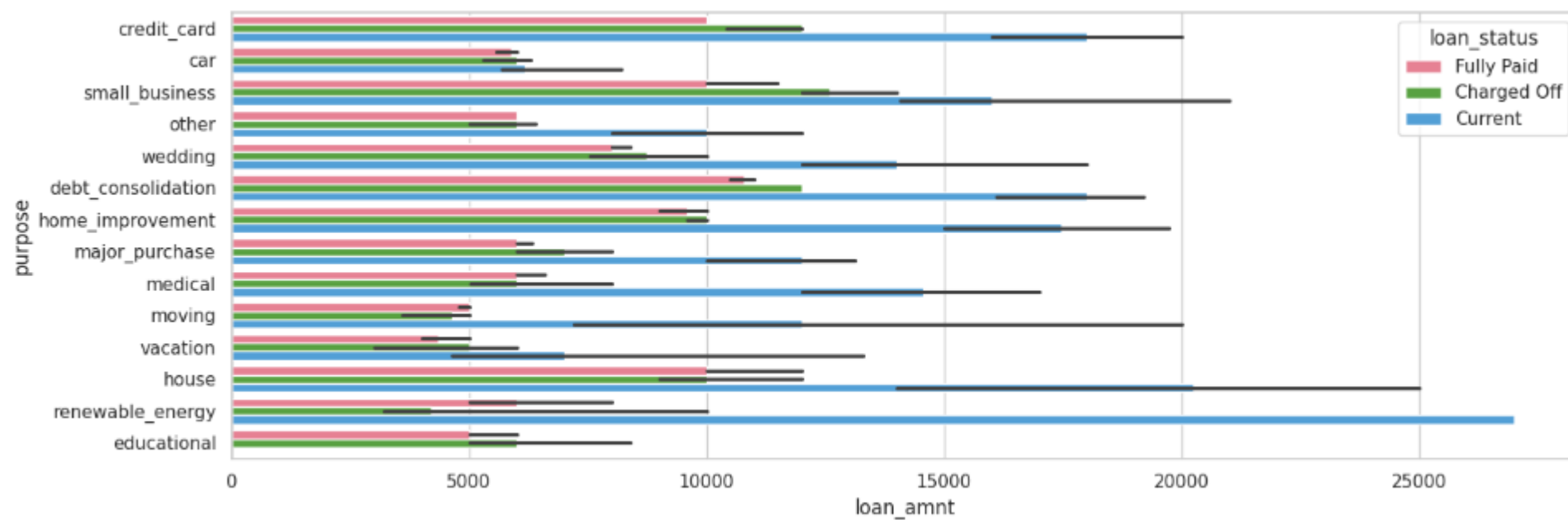
Analysis – Default by DTI

As the dti ratio rises above 20, the loans become risky

Analysis – Defaulter by Verification status

Analysis – Defaulter by Purpose

#Most defaulters were people who had the purpose of small business, renewable energy and debt consolidation for their loans

Takeaways

- Reduce number of approvals where purpose is small business and debt -.consolidation

- Start charging higher interest rates for loans with DTI greater than 20.

- Stop approving loans to people with prior bad record. Or at least stop approving high-value loans.

- Customers who have higher records of bankruptcies are bound to be charged-off more.

- Customers who don't seem to have a fixed home ownership are likely to default more.

- People being a defaulter is more when they have a less employment experience.

- The top cities who have taken the highest number of loans are also the ones who have higher number of defaulters.

Thank you

Rohinee Kanta Mishra

&

Arun Kumar