Lending Club Case Study

By:

Ankit Bhatnagar

Bhavesh Bihani

Overview

- Lending Club is a P2P lending company in operation since 2007.
- They work as intermediary between a investor (loaner) and prospective loanee.
- Lending Club as intermediary offers ease of access to a individual to seek funds.
- while at the same time offering a loner tantalising investment opportunity.
- Lending Club on its part undertakes verification of a prospective loanee and categorises the loanee into grades.
- Since, the money is being loaned by an investor and not by Lending Club, the risk of default on part of loanee is borne by the investor.

Understanding the Client

Problem Statement

Aim

Identify patterns in data which indicate if a potential loanee is likely to default, identification of which will enable the investor to apply suitable risk avoidance/ mitigation Strategy.

Objective

Based on data made available, understand the driving variables available at the time on application and historic data to identify strong indicators of default.

Approach

Utilise EDA concepts learnt to undertake risk analytics in banking and financial sector and understand use of data to minimise risk to clients.

Problem Solving Methodology

| Data Cleaning	Data Understanding	Univariate Analysis	Segmented Univariate Analysis	Bivariate Analysis	Recommendation
Removing the null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.	Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses	Analysing each column, plotting the distributions of each column.	Analysing the continuous data columns with respect to the categorical column	Analysing the two variable behaviour like term and loan status with respect to loan amount.	Analysing all plots ,recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

Lending Trend

Tread 1
There has been a steep increase in amount of money lended by LC since 2007
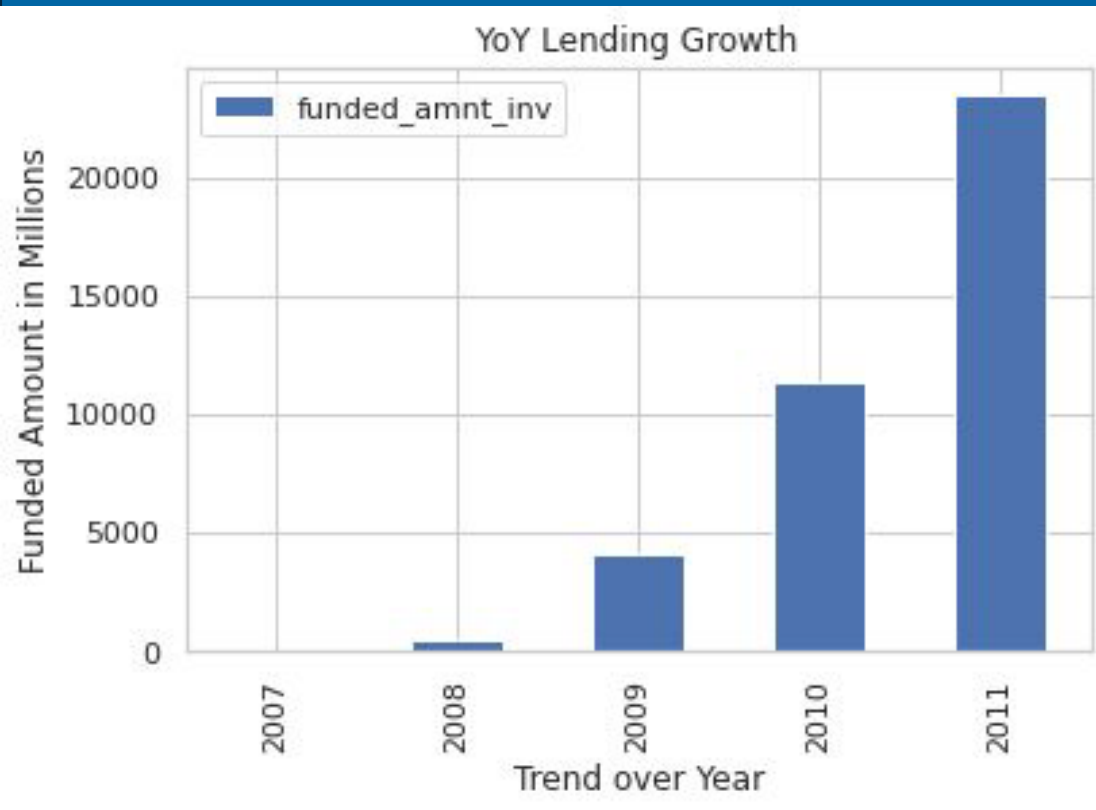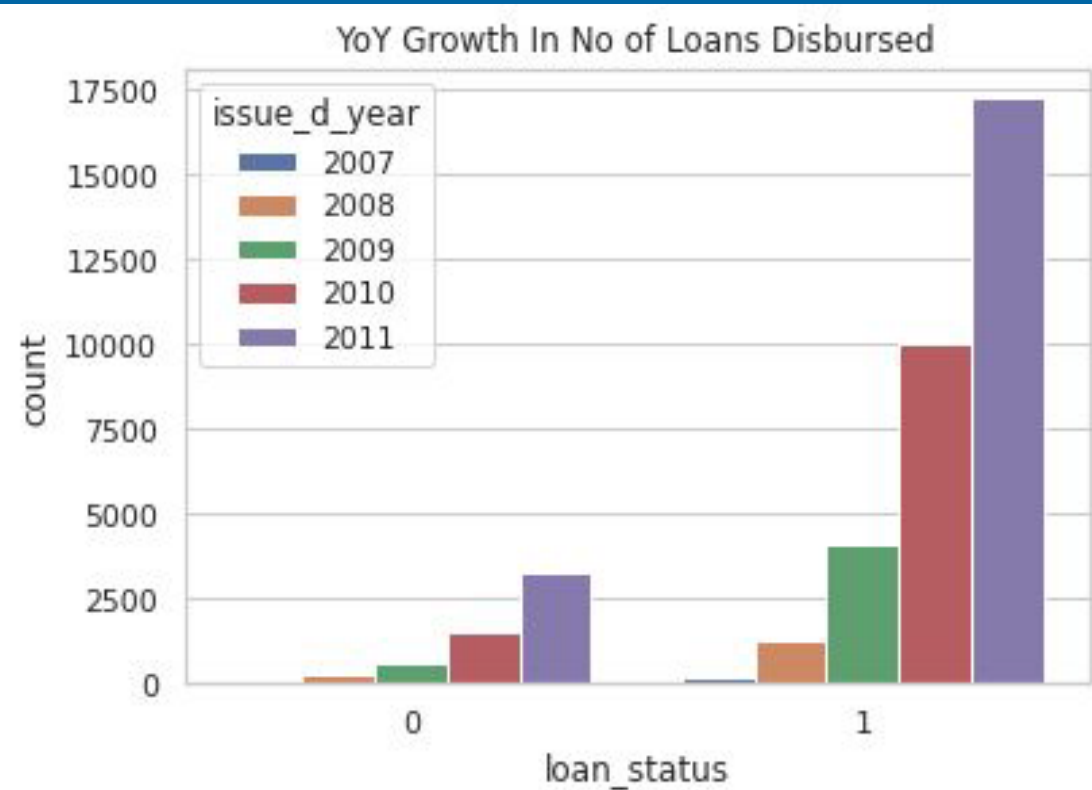
Tread 2
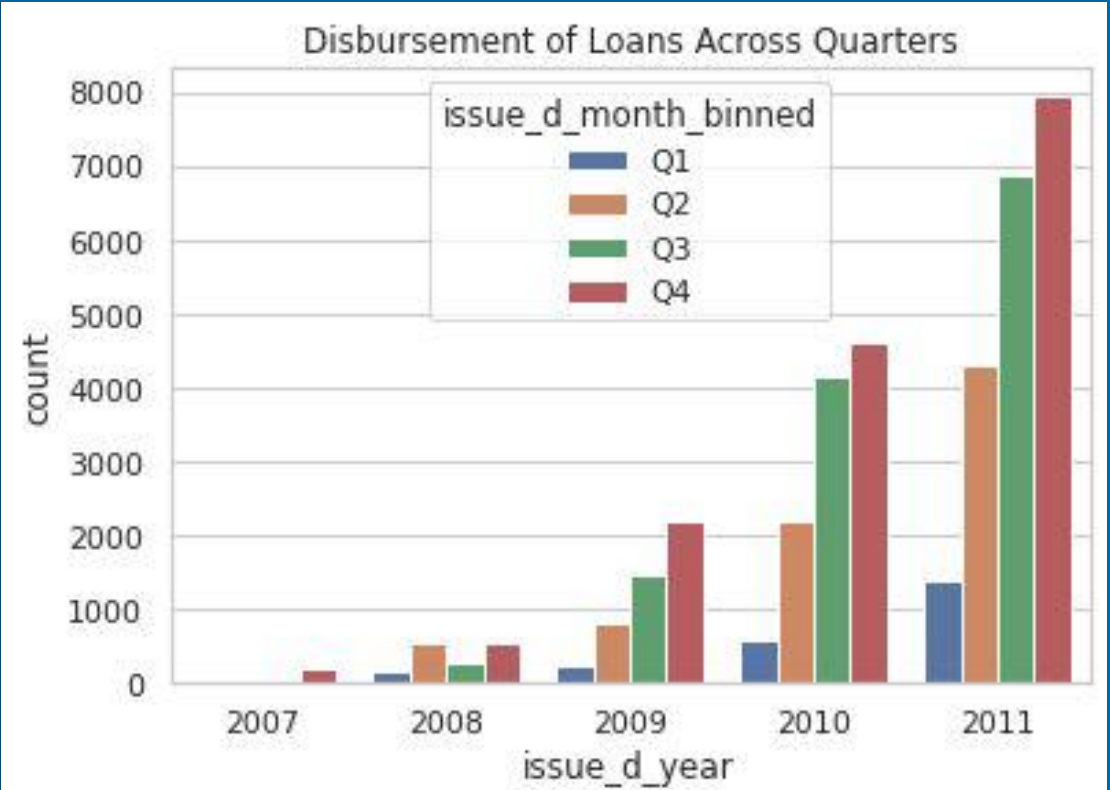There has also been a corresponding increase in the amount of default
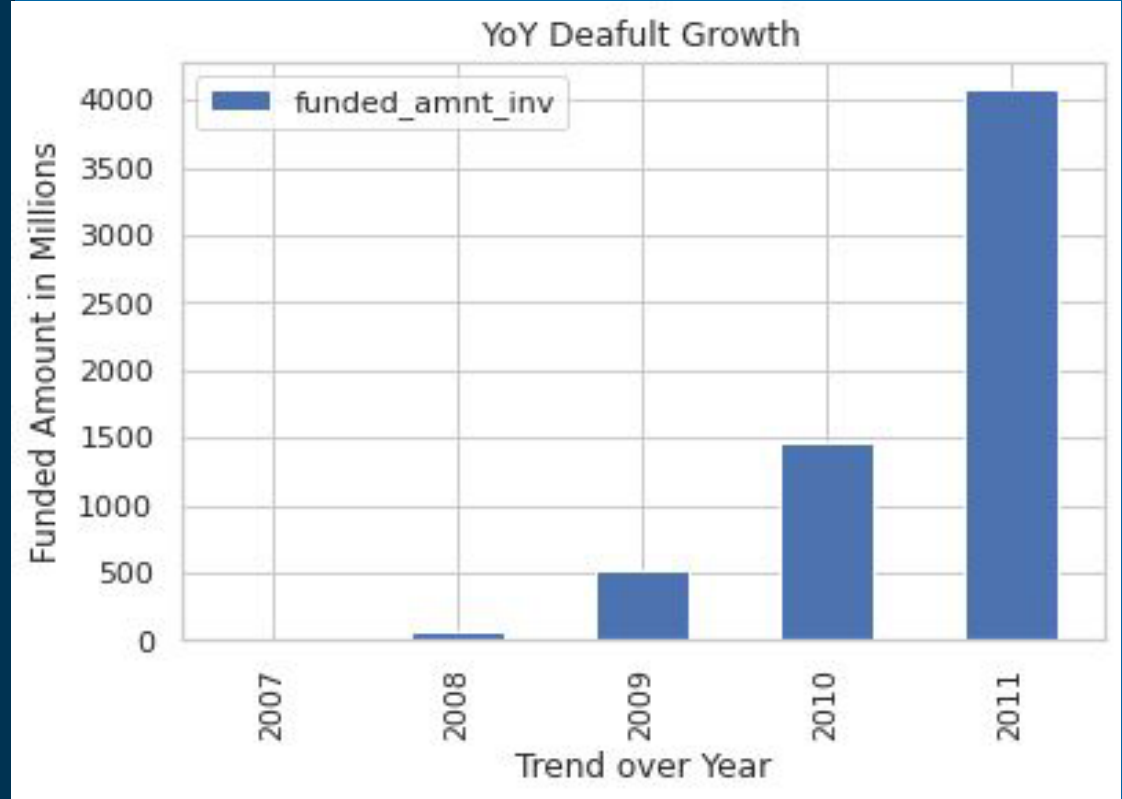
Lending Trend

Tread 3
There has been a steep increase in number of loan disbursals.

YoY Growth In No of Loans Disbursed

Tread 4
Disbursement of loans increases towards end of financial year.

Disbursement of Loans Across Quarters

Caveats to Analysis

- Only historical data and data available at the time of application considered.

- Running debt schemes not considered for purpose of analysis.

- Columns having null data dropped.

- Demographic markers utilized for data analysis.

- Bespoke binning of highly segmented variables undertaken for current case and will therefore affect analysis.

- Vertical (Column) and Horizontal (Row) data aggregated to get a better picture.

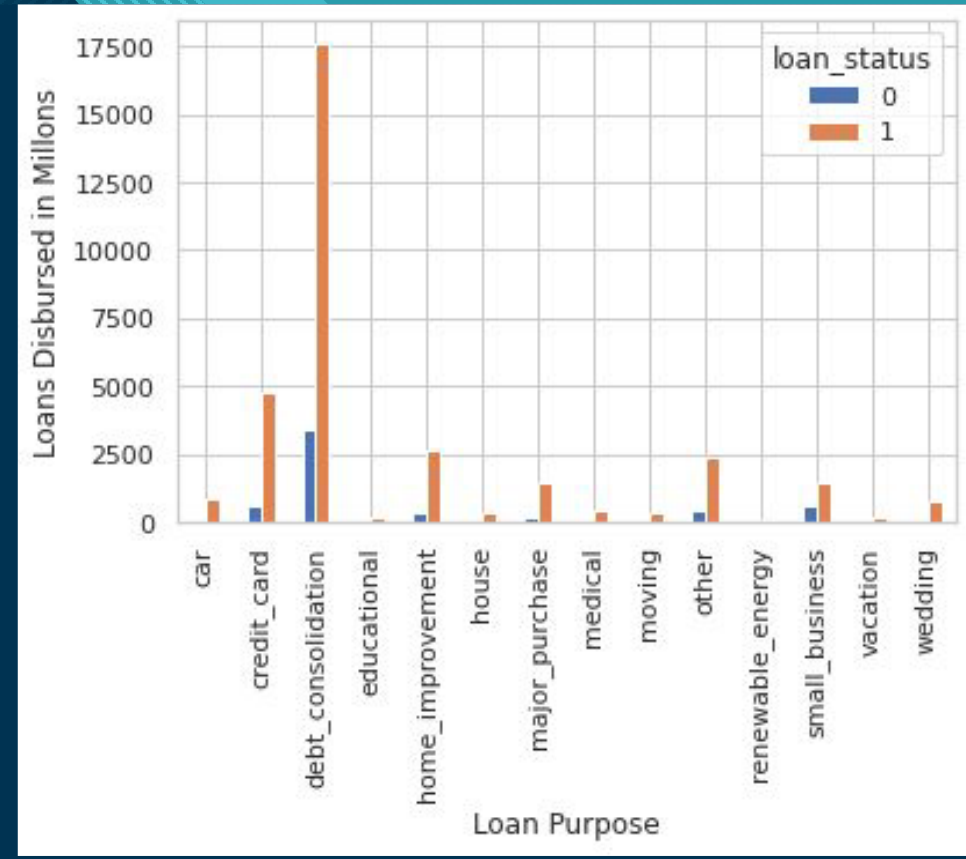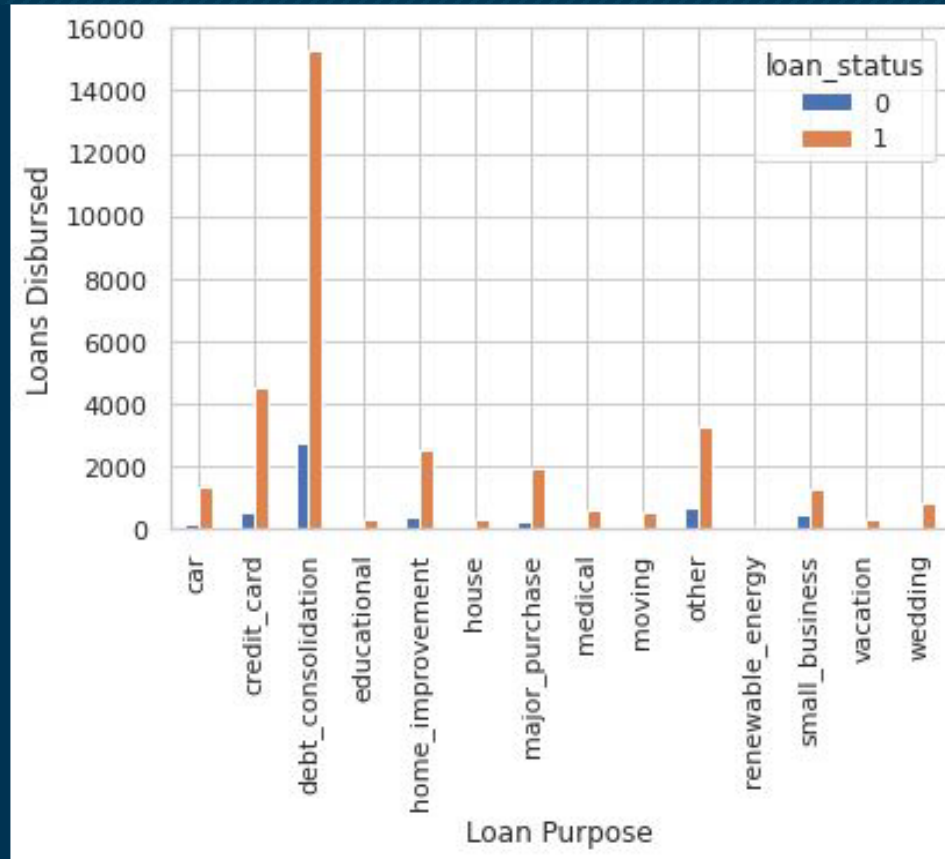- Outlier treatment not undertaken as already a lot of data has been culledand binned

Trends Analysis – By Purpose

Findings:

Most of the loans being sought are for debt consolidation.

Clients Implications:

No direct correlation observed between purpose of loan sought and possibility

Trend analysis – by Purpose

Findings

Notwithstanding the value, loans disbursed for small businesses , education and renewable energy have high default percentage amongst its category

Client Implications:

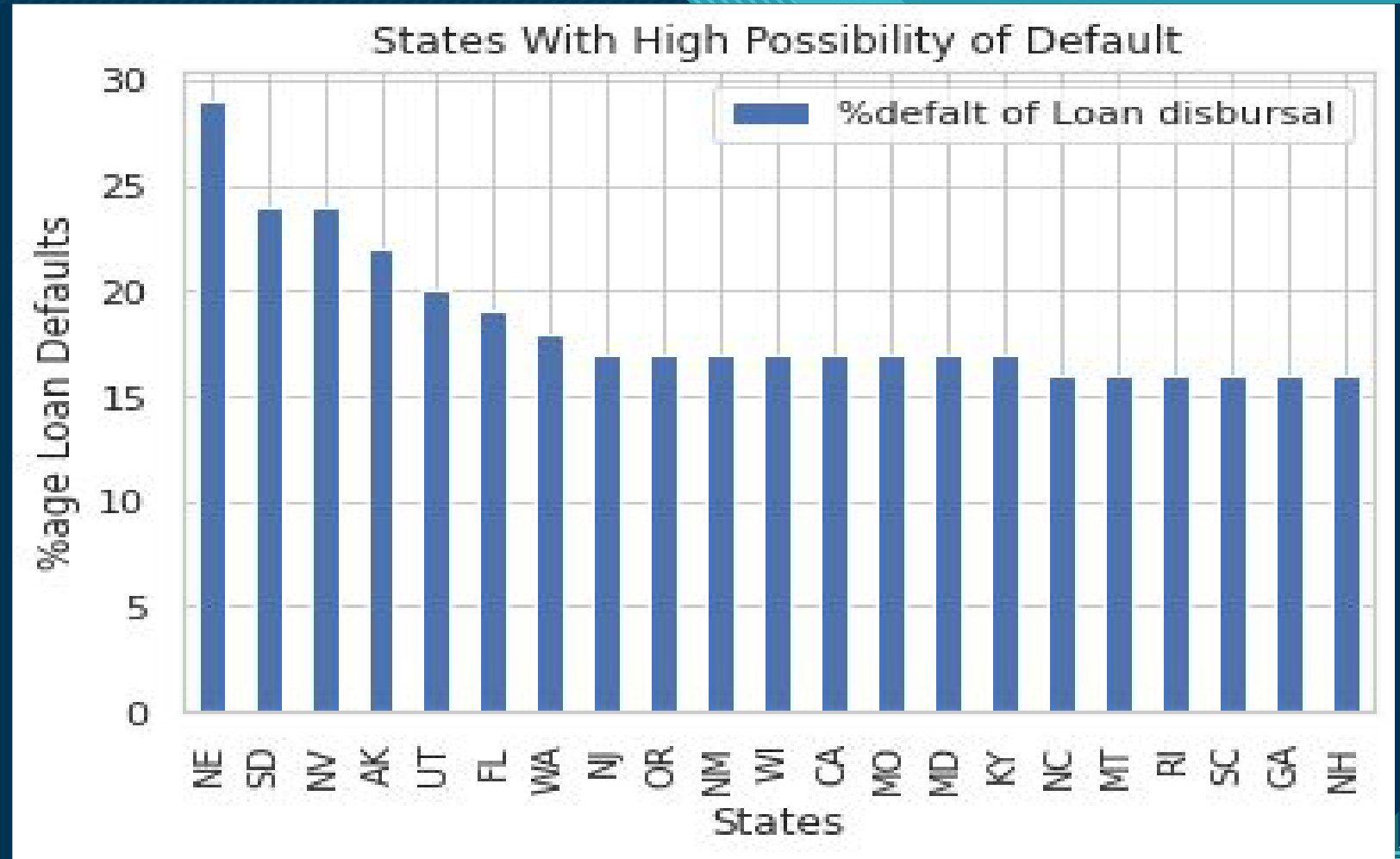Stricter vetting of loanees to be considered for indicated purposes

Loan Purposes With High Possibility of Default
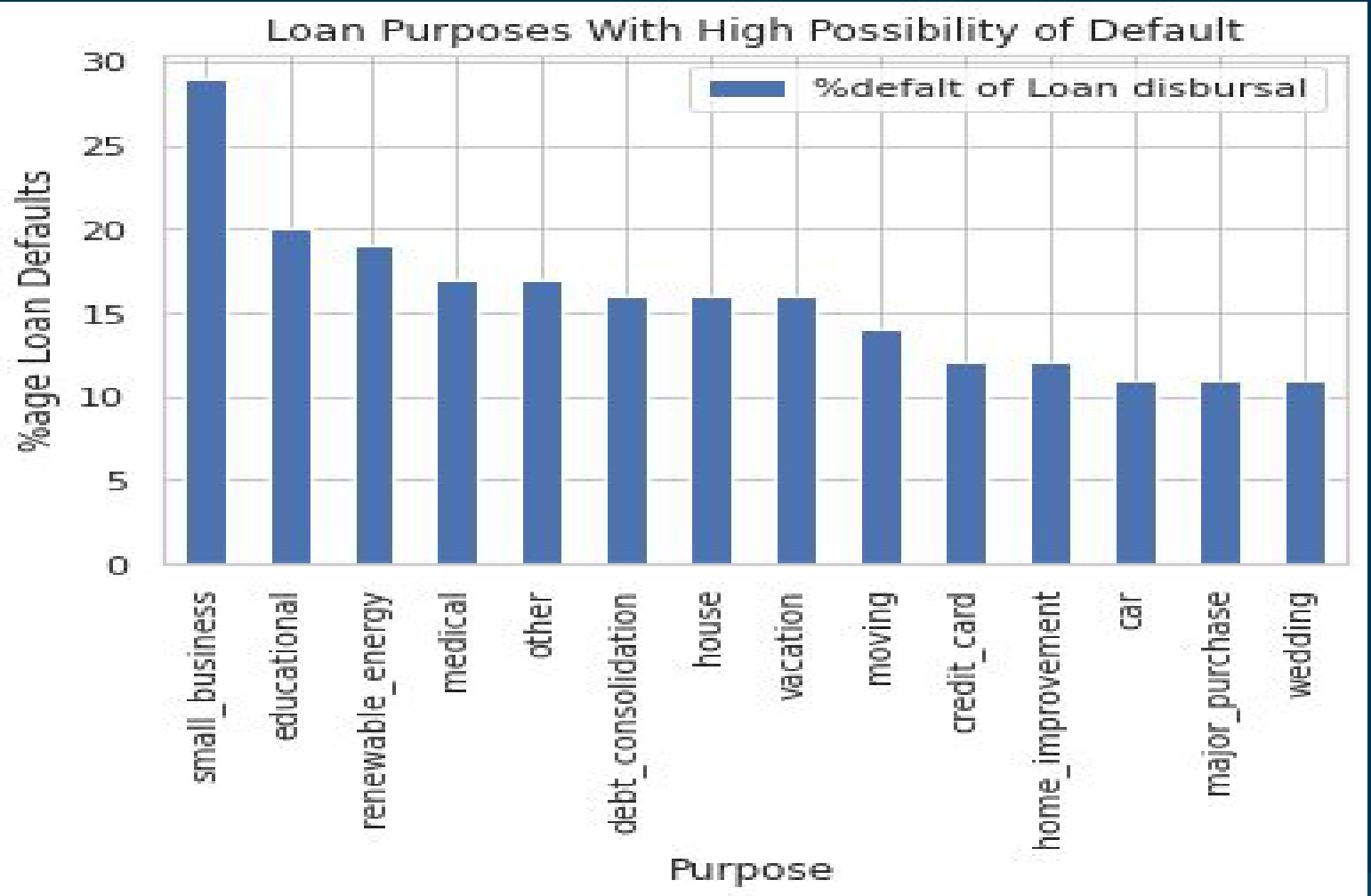
Trend analysis - by State

Findings

There are individuals belonging to a few states that exhibit a higher tendency to default

Client Implications:

Stricter vetting of loanees to be considered for Nebraska, South Dakota, Nevada, Alaska etc.

States With High Possibility of Default

Trend analysis – by Zip Code

Findings

There are individuals belonging to a ZIP Codes that exhibit a higher tendency to default

Client Implications:

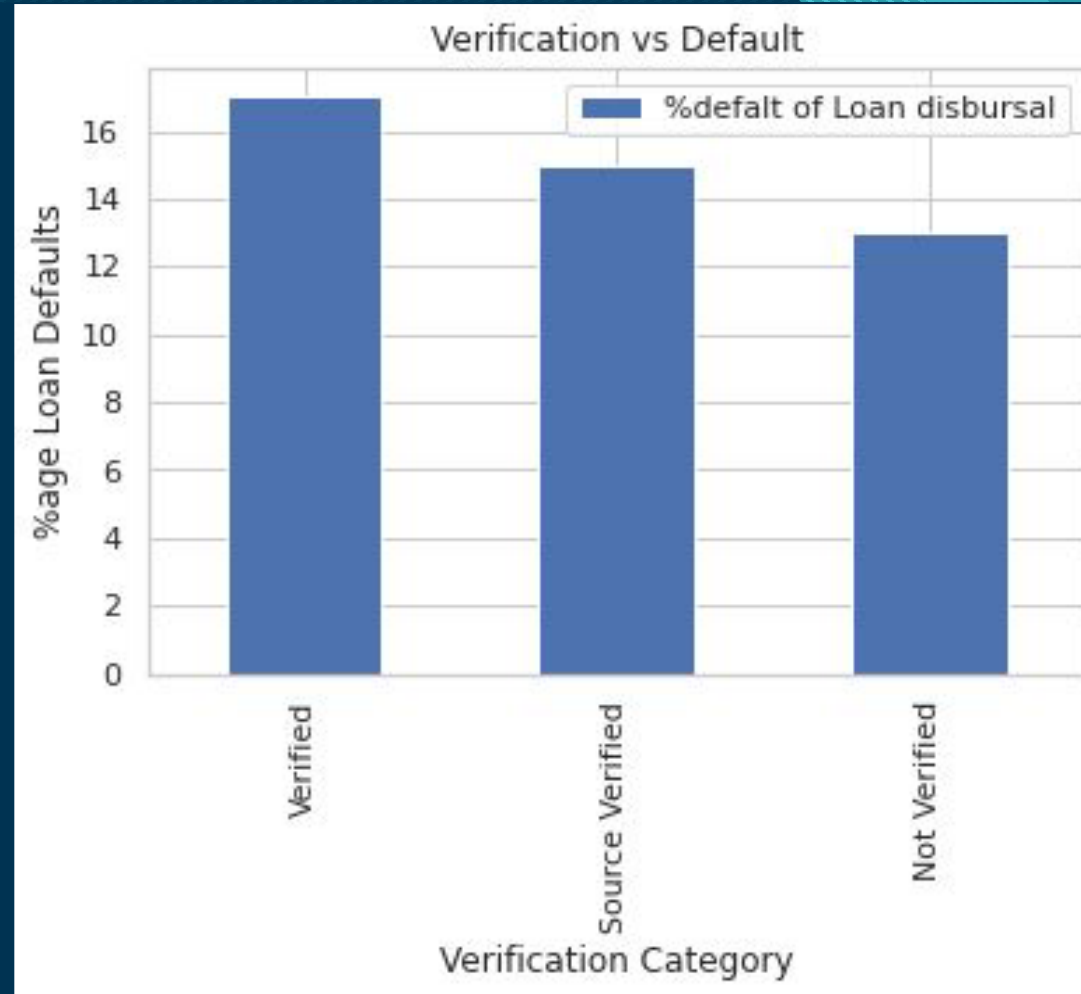Stricter vetting of loanees to be considered for indicated Zip Codes.

Zip Codes With High Possibility of Default
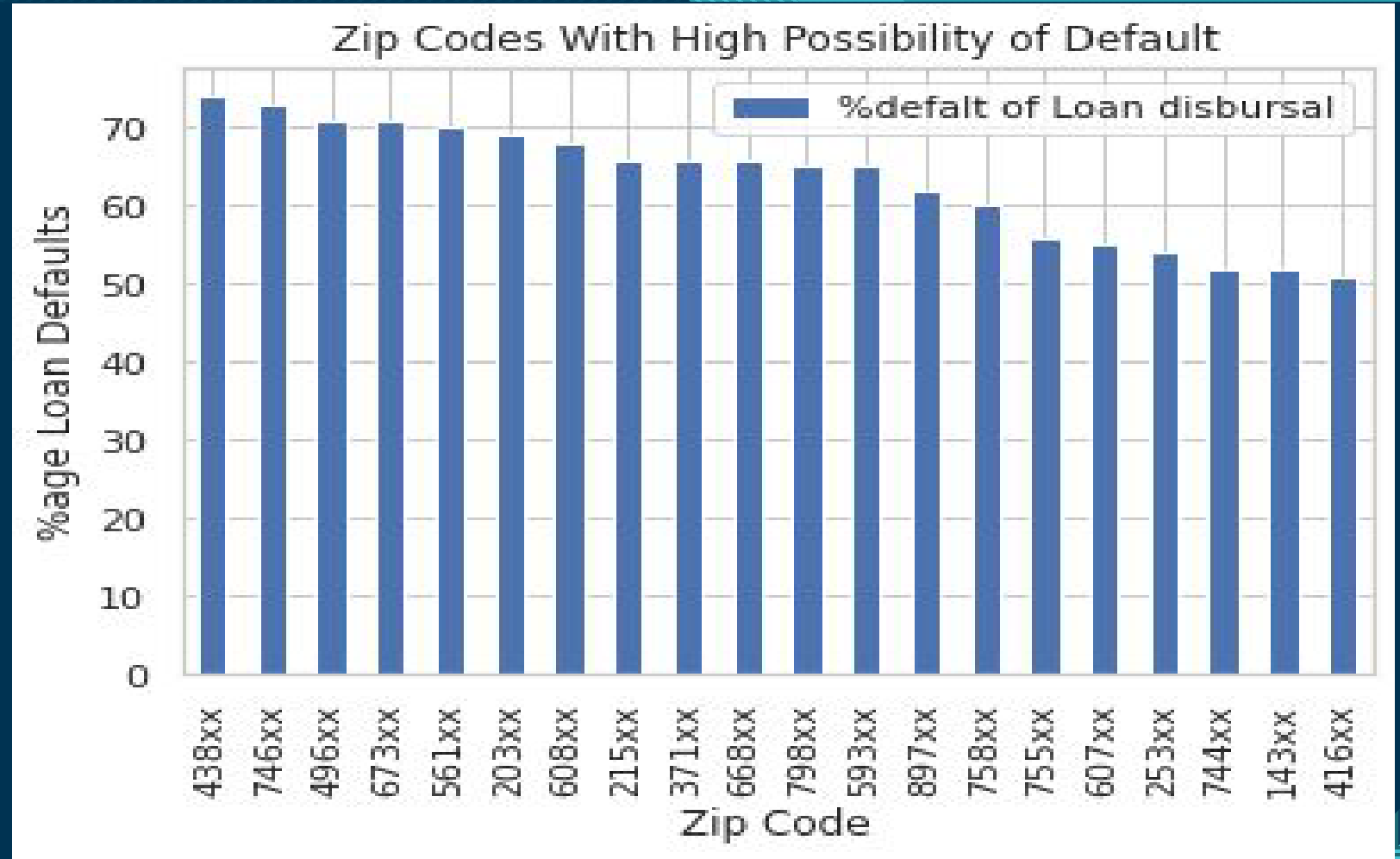
Trend analysis - by Verification

Findings

No significant variation observed in default against different verification categories

Client Implications:

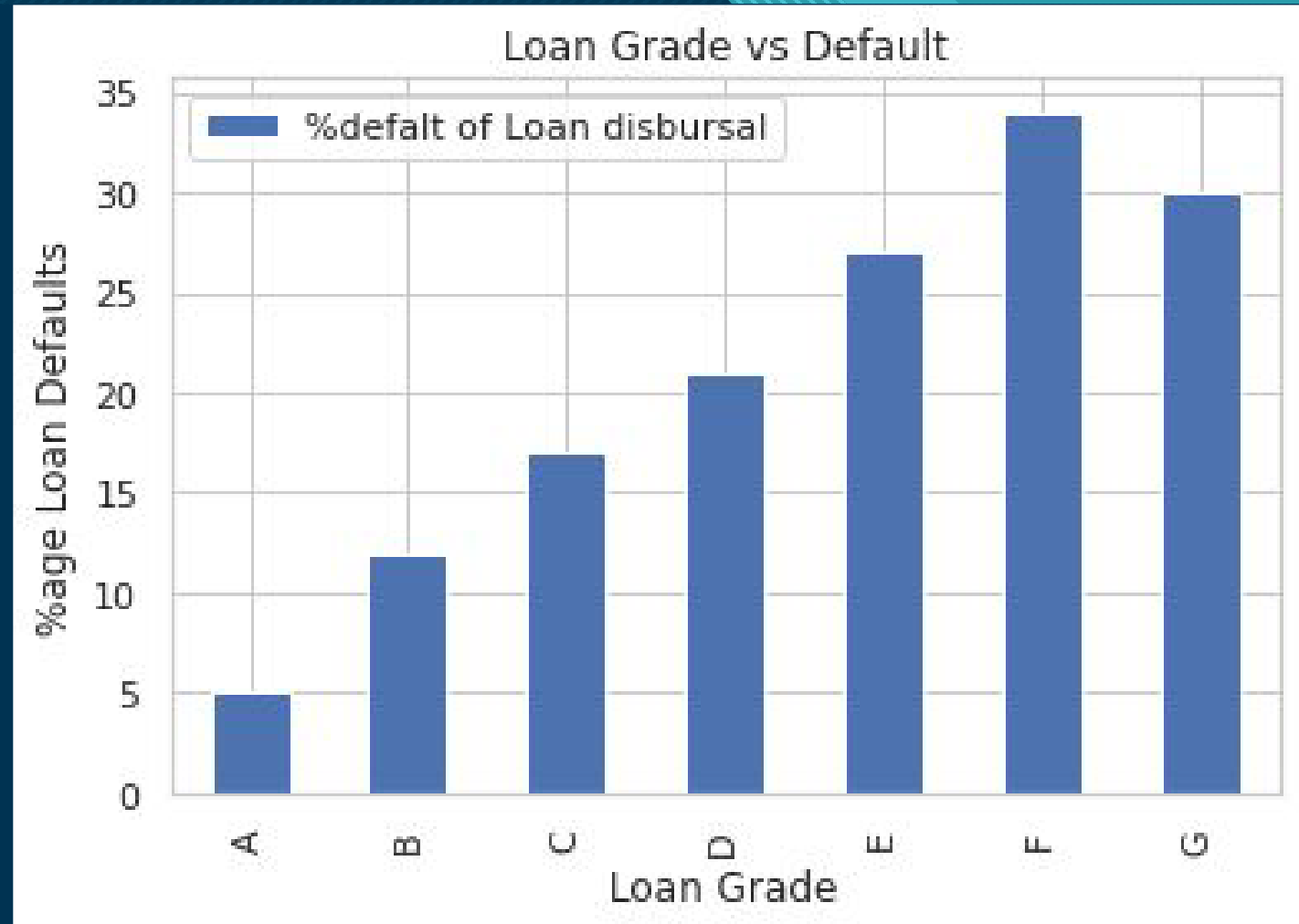Existing venting mechanism appears to be accepted.

Trend analysis - by Grade

Findings

There is a direct correlation between degreasing grade of loan and increasing incidents of defaults

Client Implications:

Existing loan grading mechanism appears to be working satisfactorily.

Loan Grade vs Default
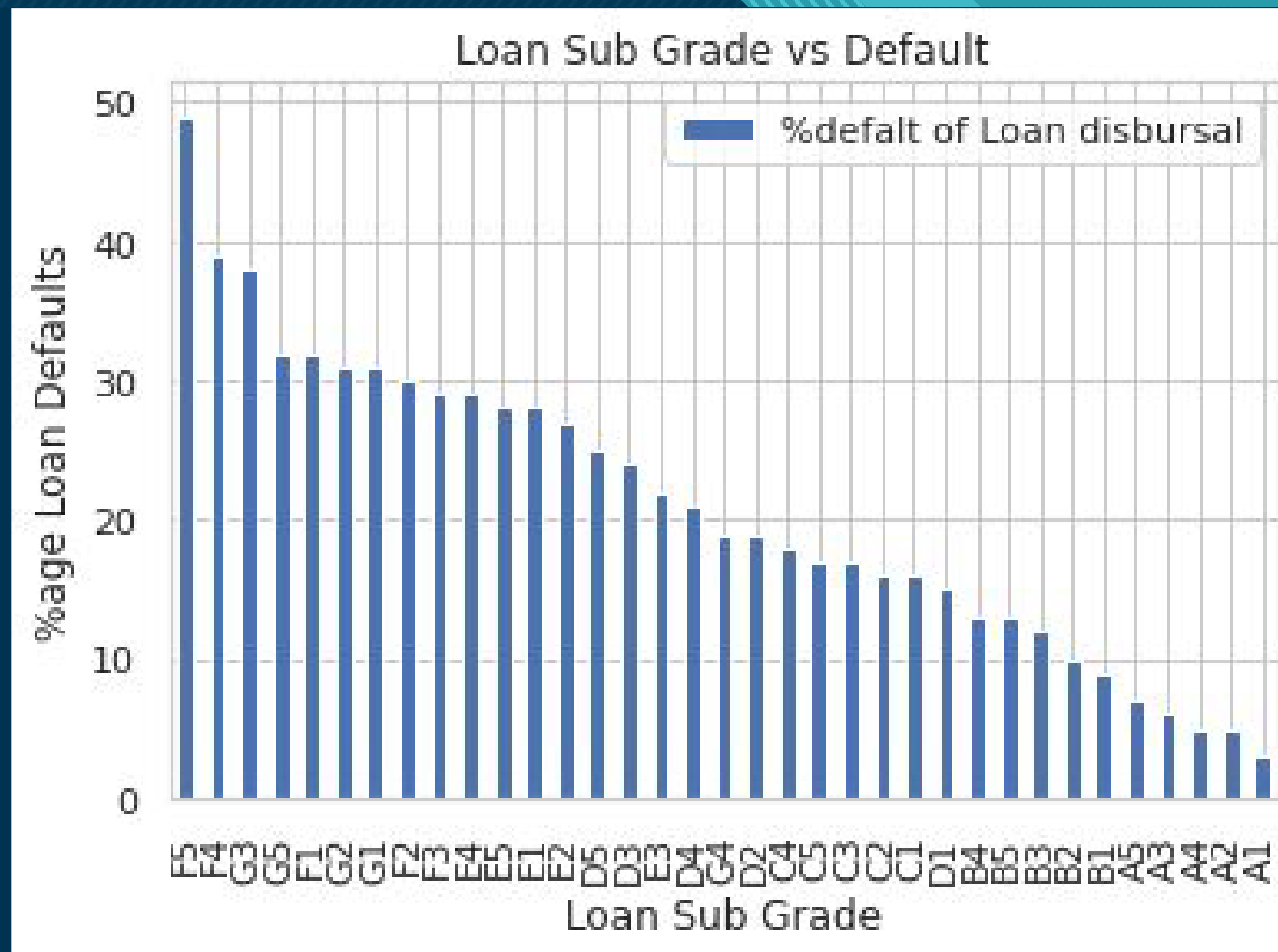
Trend analysis – by Sub Grade

Findings

There is a direct correlation between degreasing sub grade of loan and increasing incidents of defaults

Client Implications:

Existing loan grading mechanism appears to be working satisfactorily.

However, loans classified with sub grade F4, F5 and G3 required careful scrutiny
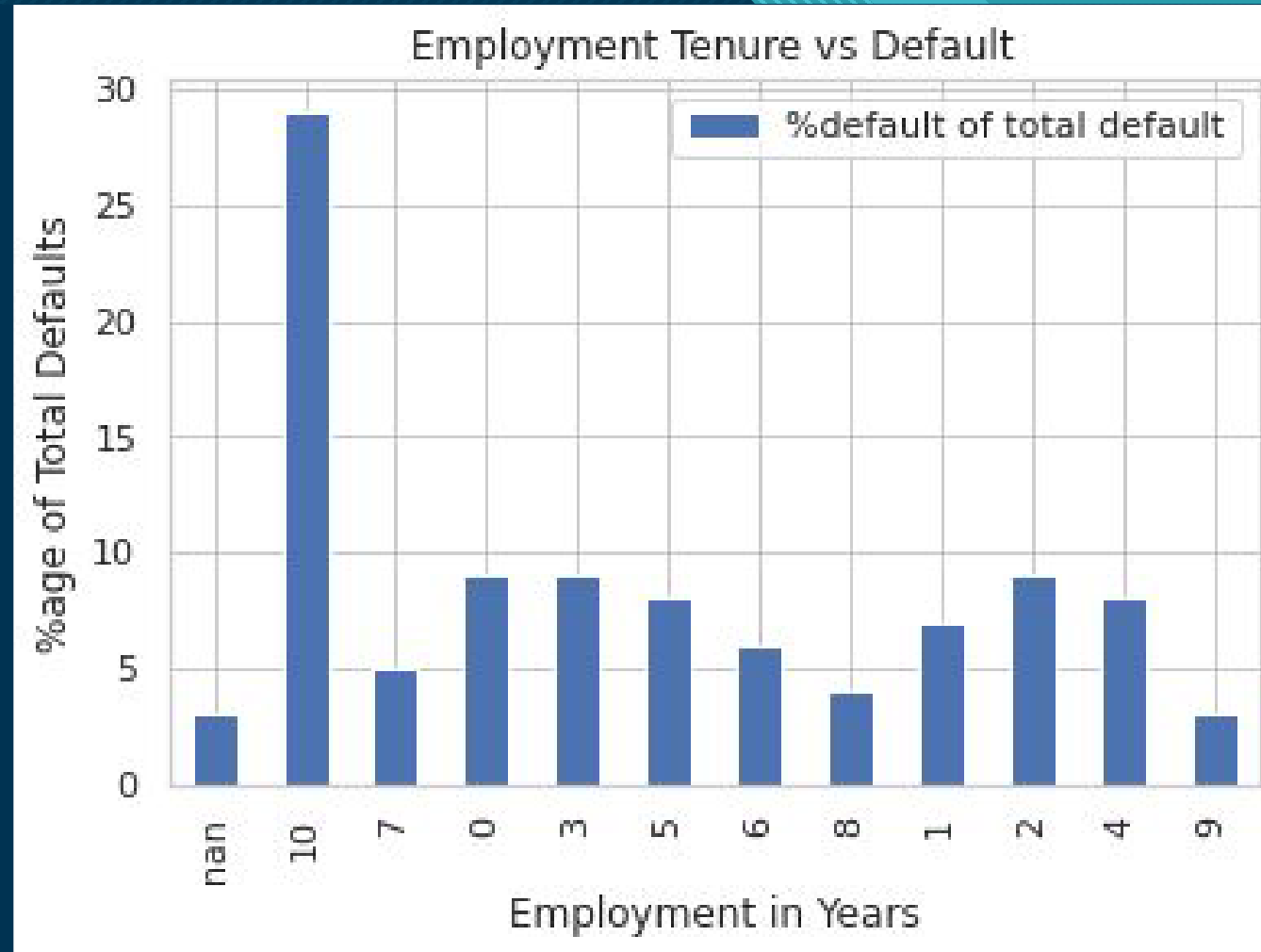
Trend analysis - by Employment Tenure

Findings

It is observed that people having employment tenures longer than 10 years are likely to default more

Client Implications:

Greater scrutiny recommended for loanees having employment more than 10 years.

Trend analysis – by Debt to Income Ratio DTI

Findings

It is observed that people having employment tenures longer than 10 years are likely to default more

Client Implications:

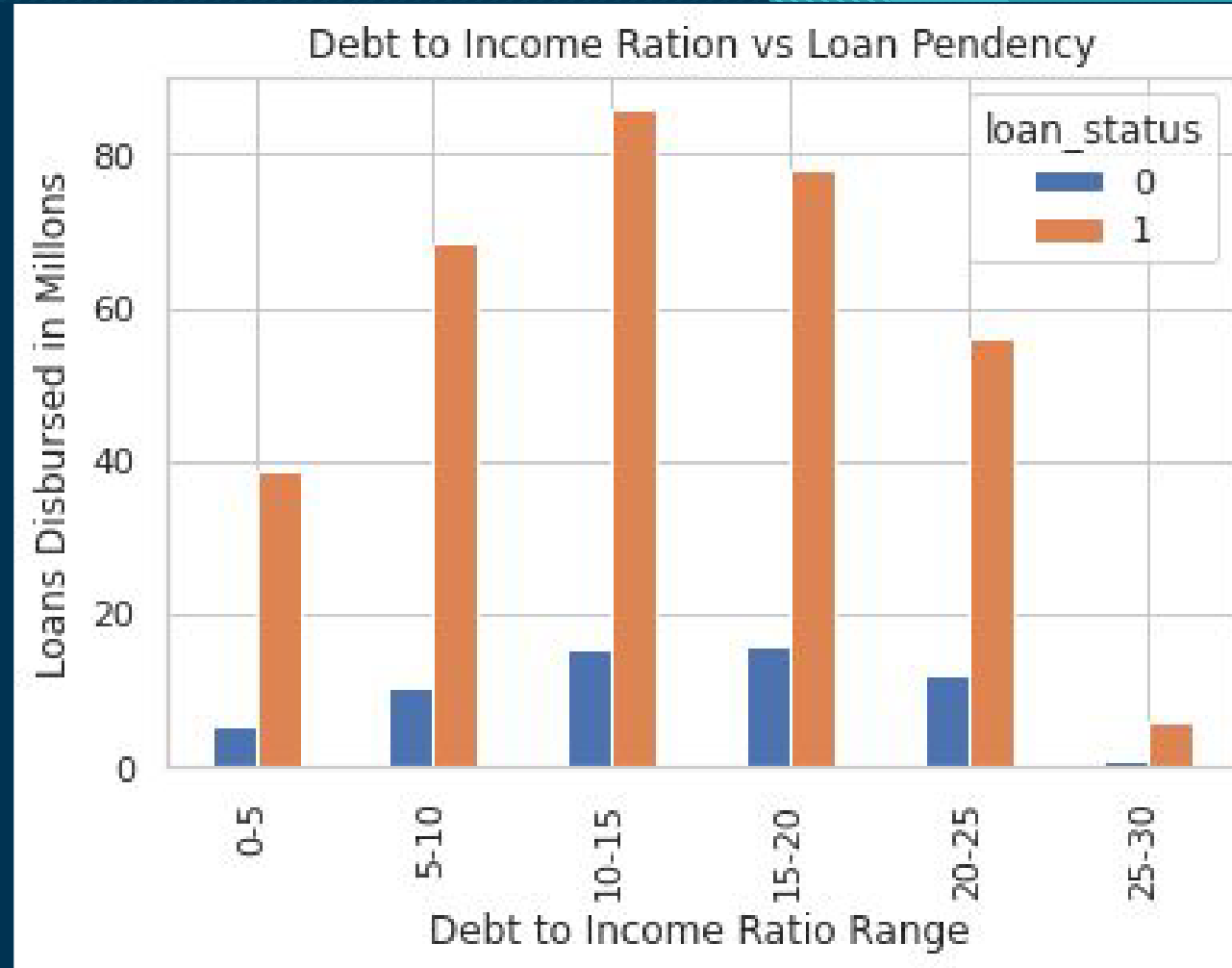Greater scrutiny recommended for loanees having employment more than 10 years.
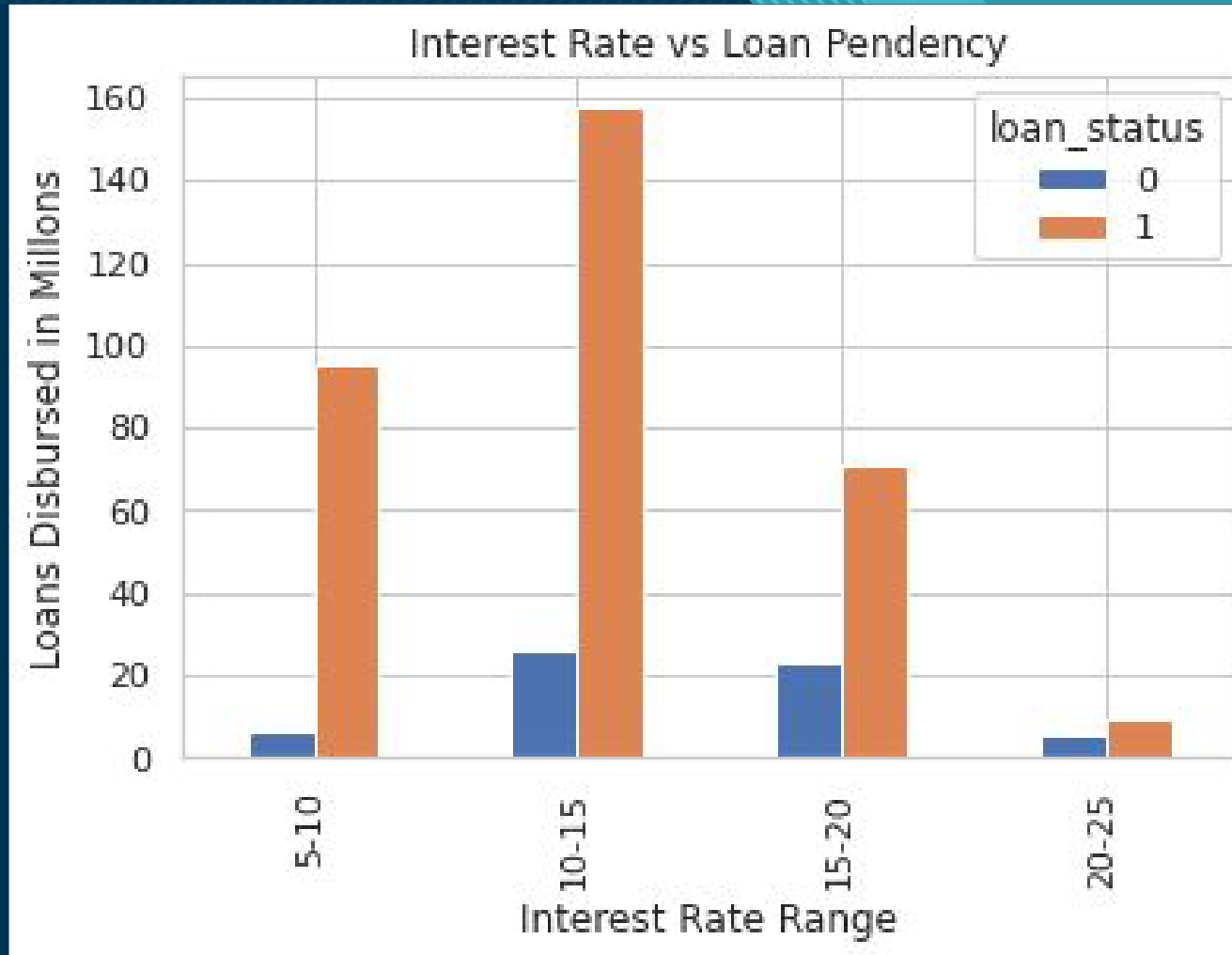

Debt to Income Ration vs Loan Pendency

Trend analysis - by Interest Rate

Findings

It is observed that chances of default in middle interest rate region is highest

Client Implications:

Lending Club may consider flexible interest rate in case loan appears to default

Trend analysis - by Employment Tenure

Findings

It is observed that chances of default in middle interest rate region is highest

Client Implications:

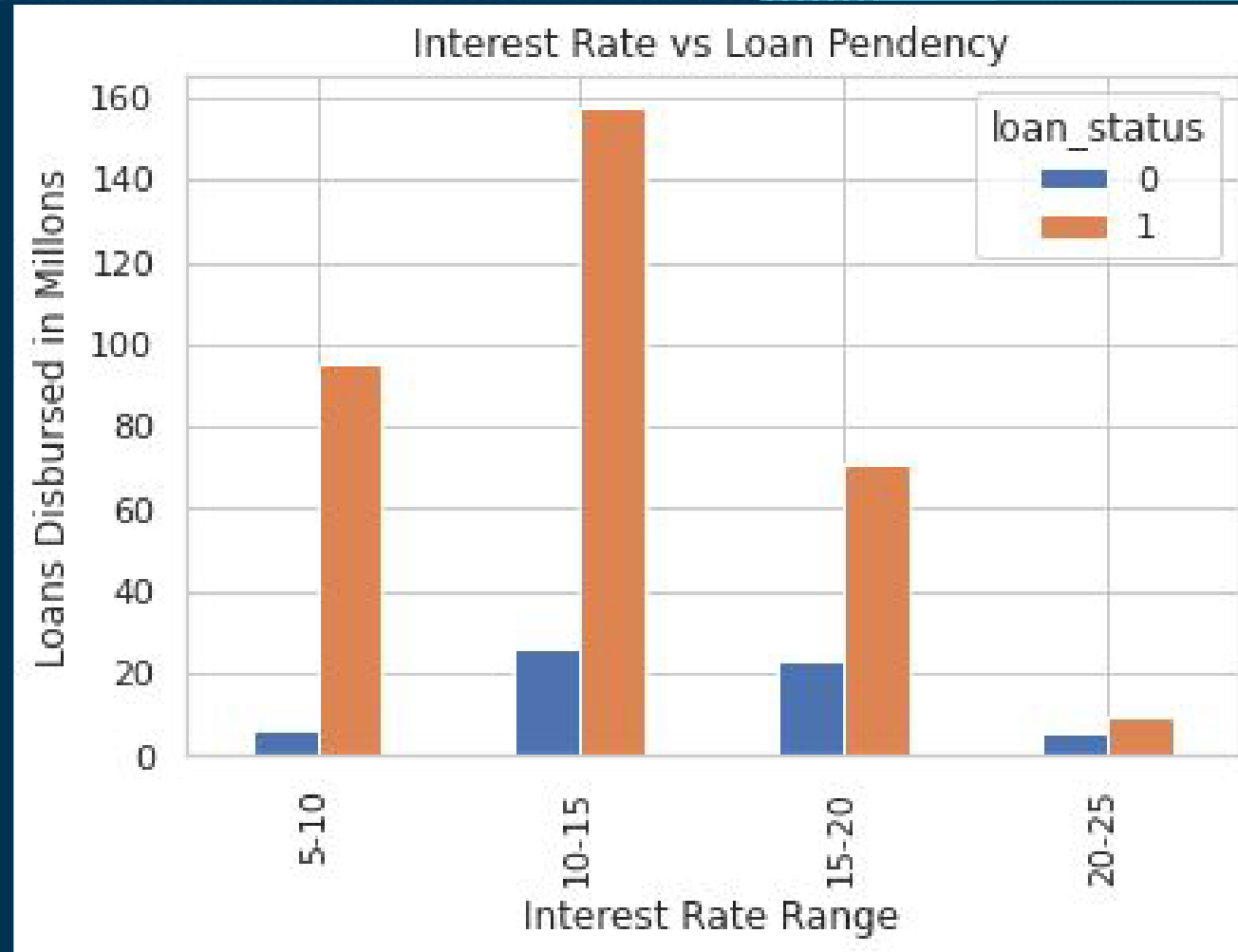Lending Club may consider flexible interest rate in case loan appears to default
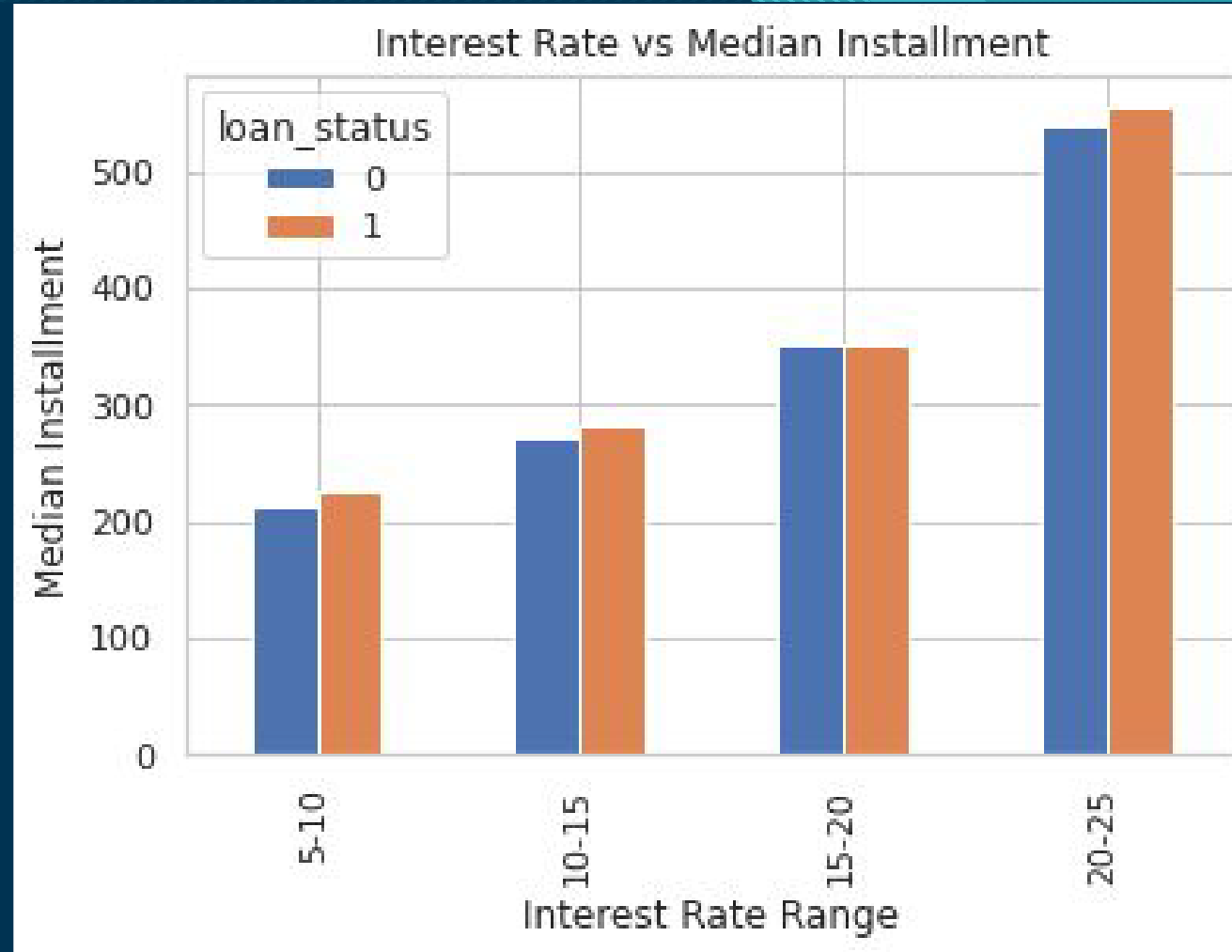
Trend analysis - by Employment Tenure

Findings

There is a direct correlation between interest rate and median installment for loan

Client Implications:

Investors may consider creating another interest bracket below 20% as there is a significant jump in default beyond the same


Interest Rate vs Median Installment

Multivariate Analysis

Findings

There is a positive correlation between some entities and a negative correlation between some. Chances of default are strongly correlated with term and interest rate

Client Implications:

Investors may modulate the variables plotted in the heat map to tailor loans to individual needs to reduce chances of default and ensure guaranteed returns.

Correlation Heatmap
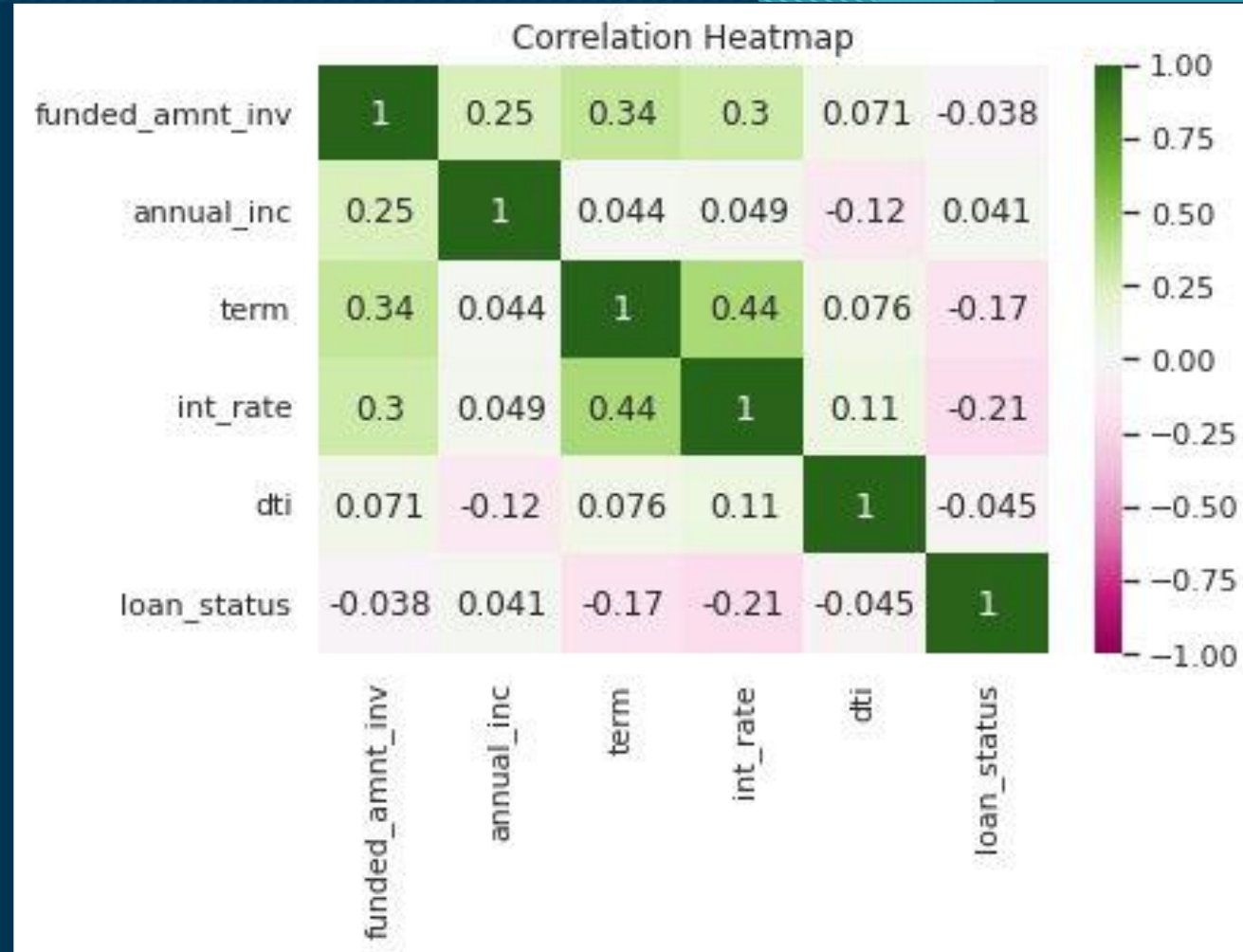
Recommendations

Customer Demographics
- Customer demographics are a good indicator of ability of customer to pay loans

Credit Verification
- Existing schemes to grade loans and customer verification seem to be satisfactory

Loan Metrics
- Lower interest rates coupled with short term and easy EMI decrease the chances of default

Loan Disbursal
- Investors may closely scrutinize high value-high interest rate loans

Thank You