Lending Club Case Study
SUBMISSION

Team Members:
Sourabh Mishra (Group Facilitator)

Mahesh Chintha

Business Understanding:

▪ Lending club is an online market place to enable peer to peer lending of money between the borrowers and lenders. Borrowers are individuals who come to the LC to avail loans at a lower interest rate. Investors who lend on the LC get solid returns based on the interest rate they choose to lend.

▪ 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. In this case, the customers labelled as 'charged-off' are the 'defaulters'

Business Objective:

▪ 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 utilise this knowledge for its portfolio and risk assessment. Data categorization: Loan accepted: If the company approves the loan, there are 3 possible scenarios described below: Fully paid: Applicant has fully paid the loan (the principal and the interest rate) Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'. Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan

Data categorization:

Loan accepted: If the company approves the loan, there are 3 possible scenarios described below: Fully paid: Applicant has fully paid the loan (the principal and the interest rate) Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan

<Problem solving methodology>

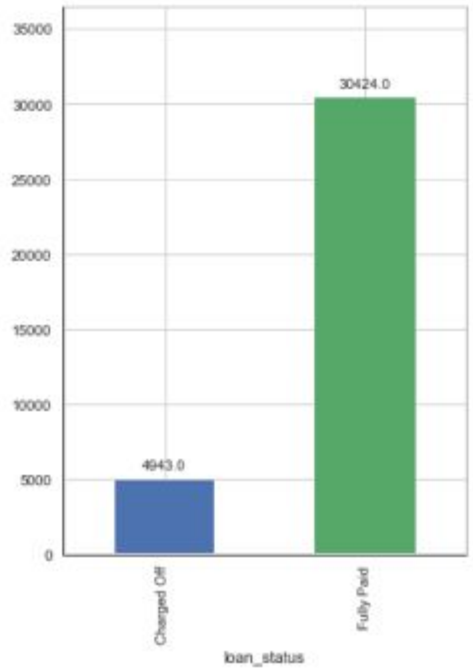
- Data Cleaning:
    - Drop columns with null values all random values or single category value
    - Convert values to proper int, float, date representations
- Univariate Analysis:
    - Check distributions and frequencies of various numerical and categorical variables.
    - Create derived variables.
- Segmented Univariate Analysis:
    - Analyzed variables against segments of other variables.
    - Create derived variables
- Bi-Variate Analysis:
    - Do correlation analysis check how two variables affect each other or a third variable
    - Analyze joint distributions.
- Summarize Solution:
    - Produce insights and observations.

Analysis - Overall Loan Status

Total Loans

Approximately 14% of loans are defaulted

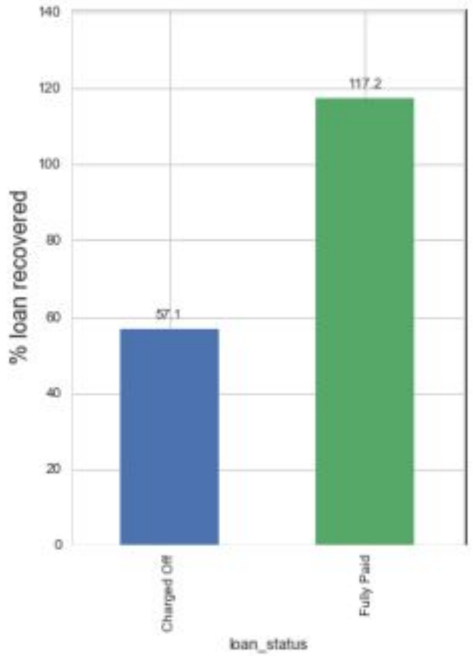
Any variable that increases percentage of default to higher than 16.5% should be considered a business risk.



Total Money Earned

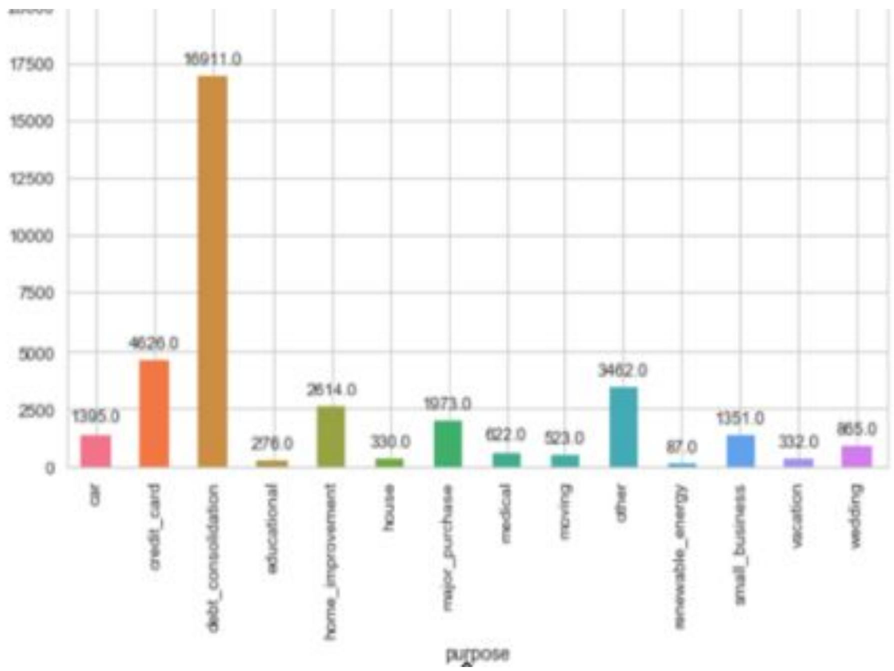
Lending Club only recovers 57% of the loan amount when loans are **defaulted**.

On fully paid up loans, the company makes 17% profit.

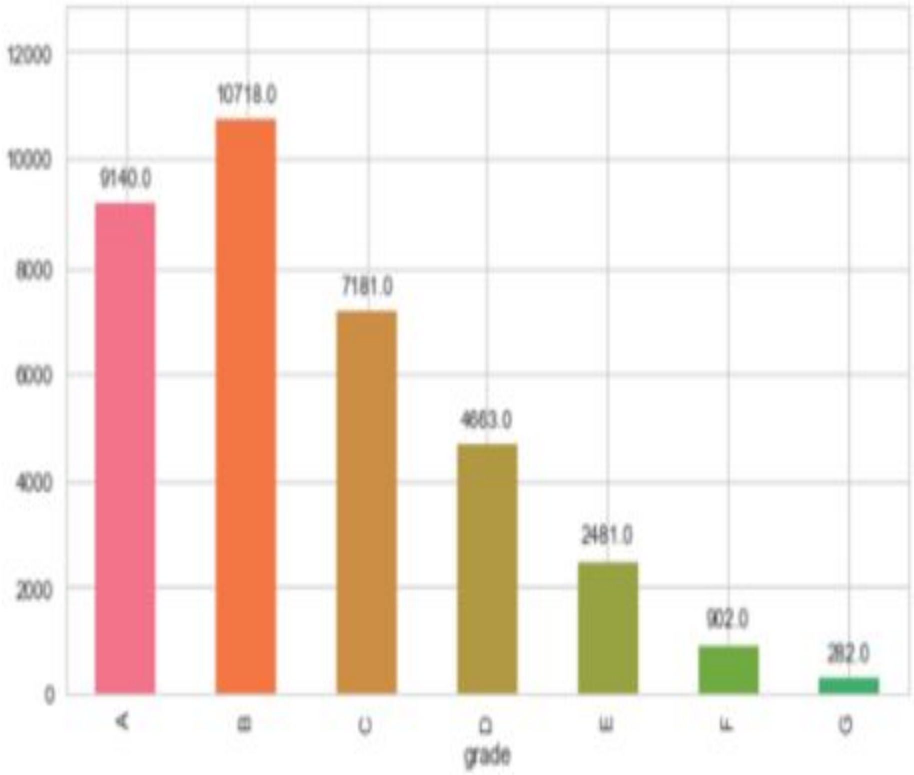


Analysis - Understanding Loans

- max no of loans are for debt consolidation followed by credit card.

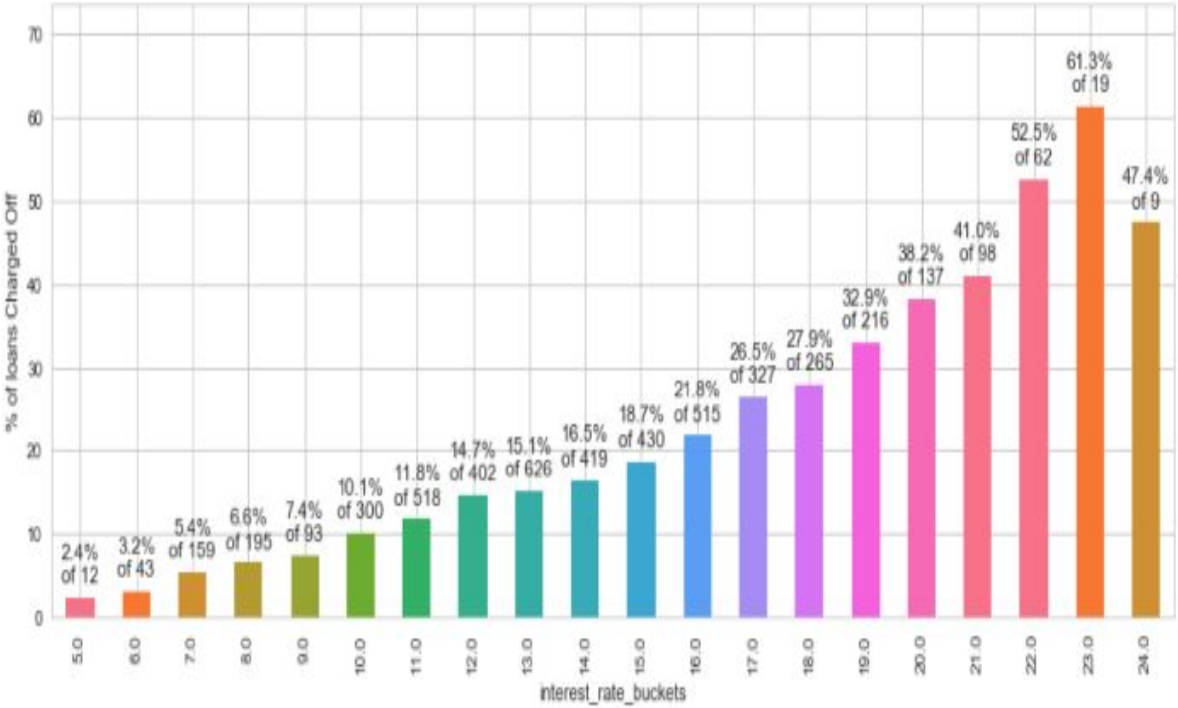


- most loans are high quality, with A grade or B.



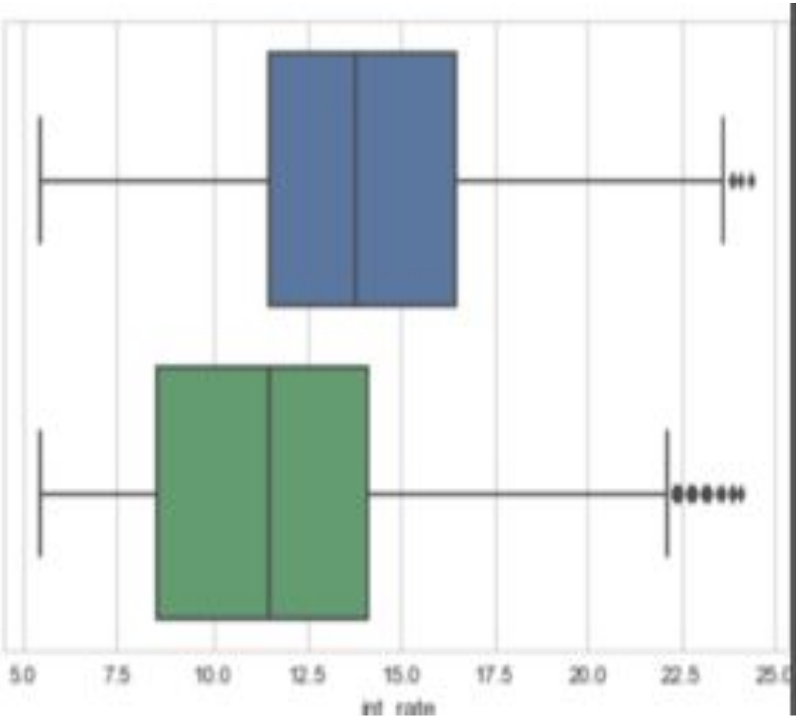
Analysis - Default by interset rate

- percentage of default increases monotically with higher interest dates.



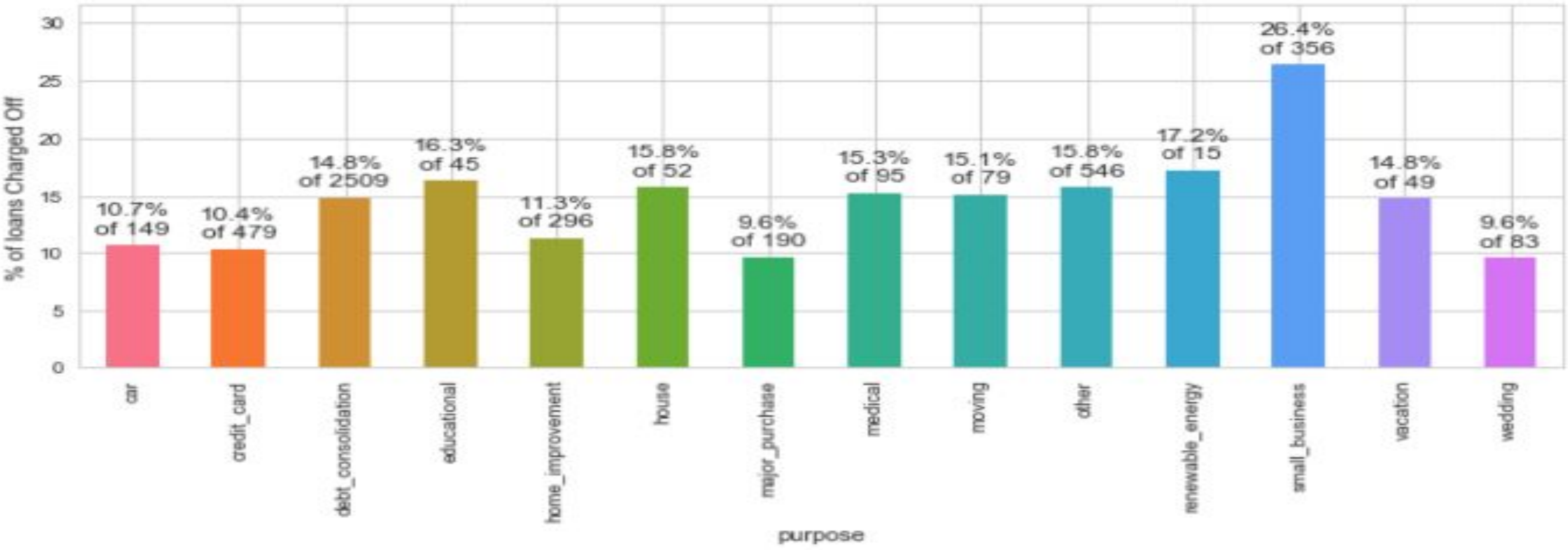
Charged off

Fully paid



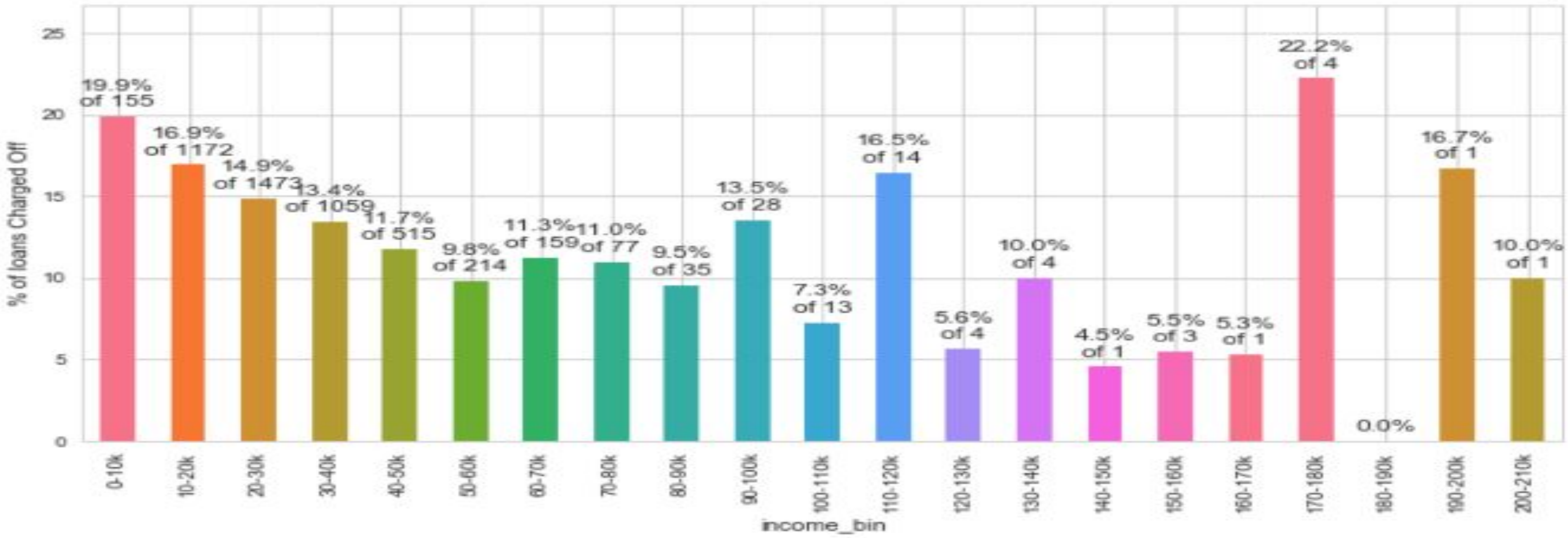
Analysis - Default by loan purpose

More than a quarter of loans taken for the purpose of running a small business see defaults.

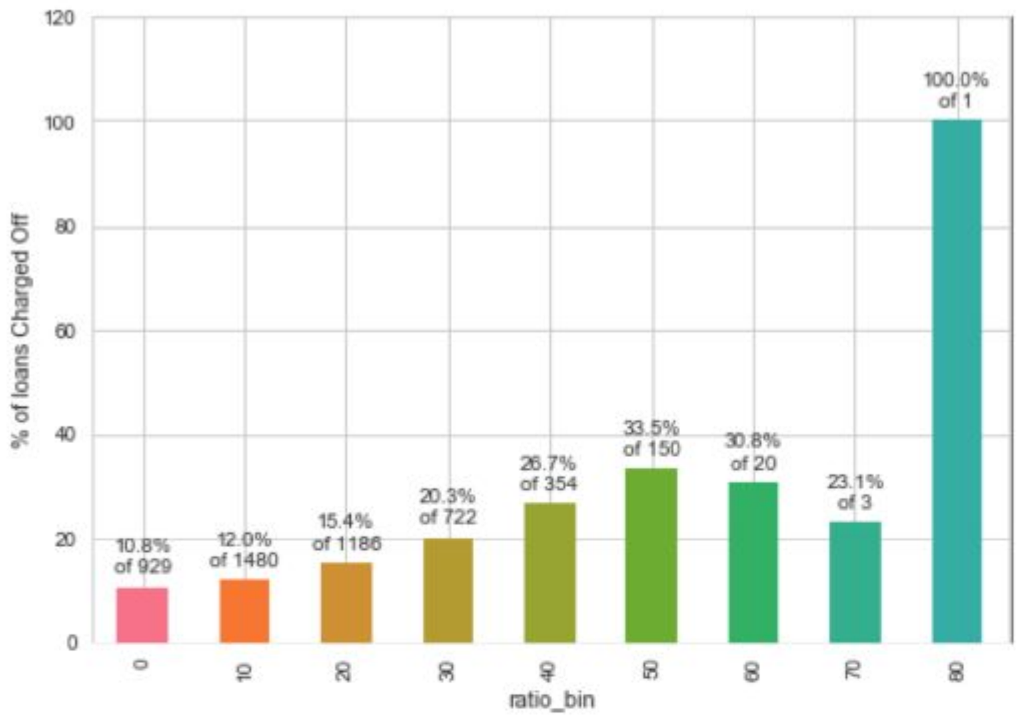


Analysis - Default by Borrower's income

Borrowers having annual income less than 20000 default on their loans at much higher rates.
Loan default decreases with higher annual income.
As we will see on next slide – the ratio of amount to income is more important.



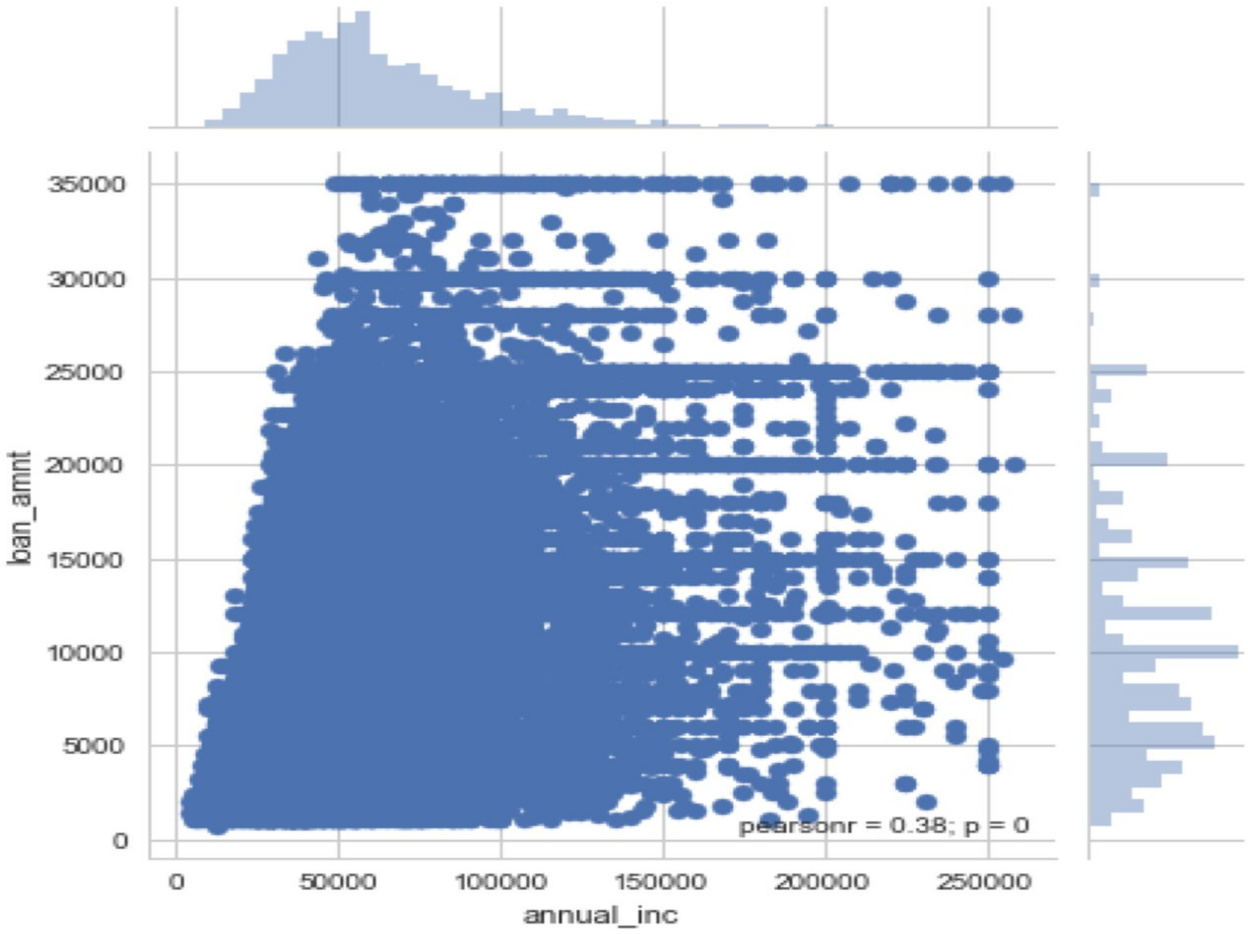
Analysis - Default by ratio of amount to income



X-axis is the % of loan amount versus annual income

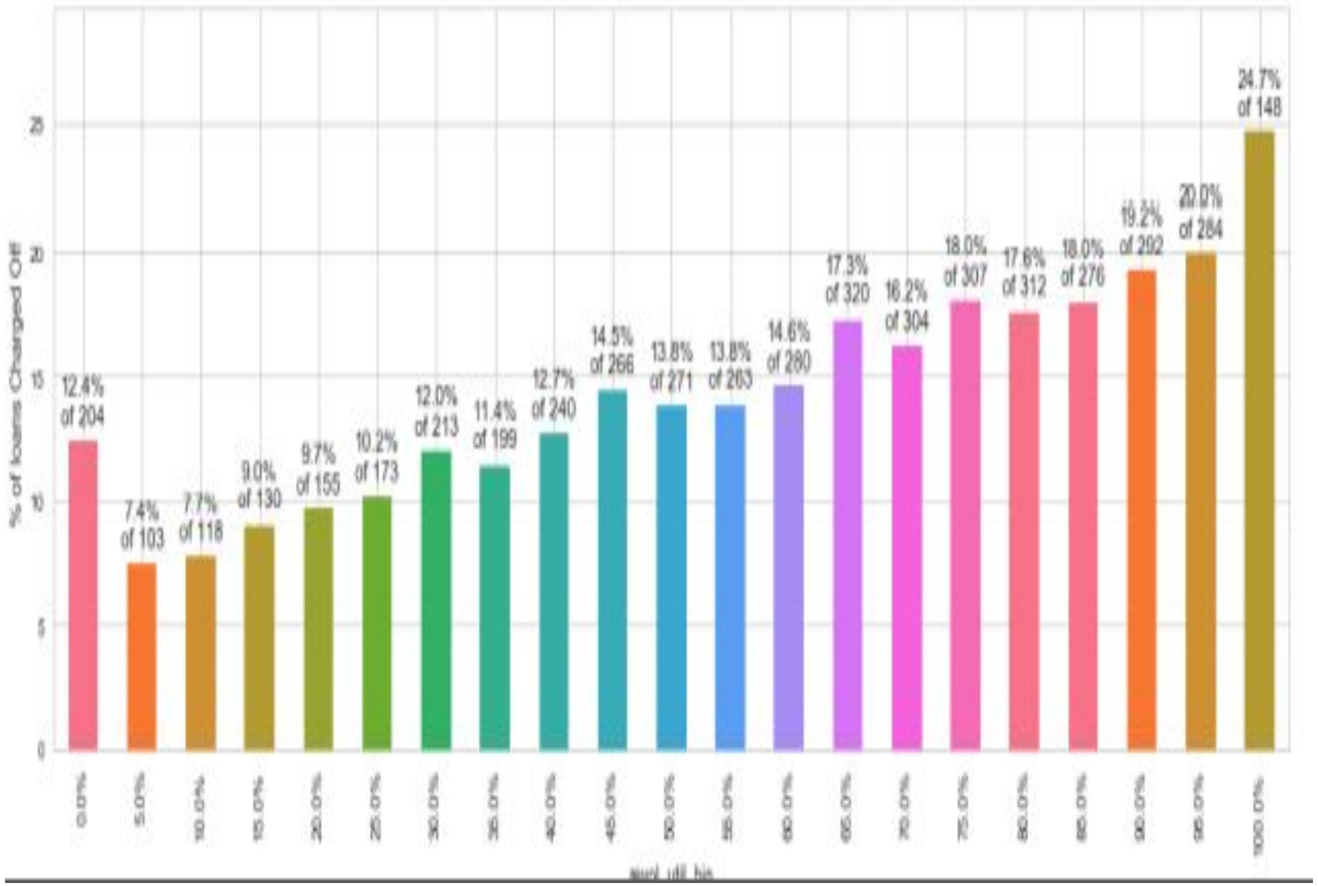
As long as loan amount is less than 20% of annual income, defaults are low.
Loan amounts of 30% of annual income or higher see a high rate of default.

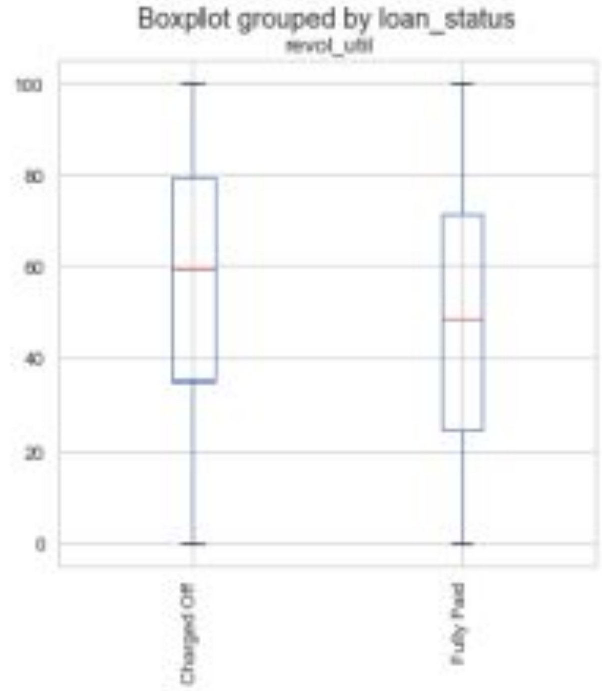
Analysis - Default by ratio continued.



We see here that Lending Club has
extended high-value loans to people with low
income.
There are many cases of people with income
50000 or less getting loans of 25000 or
more.
This practice should be curtailed.

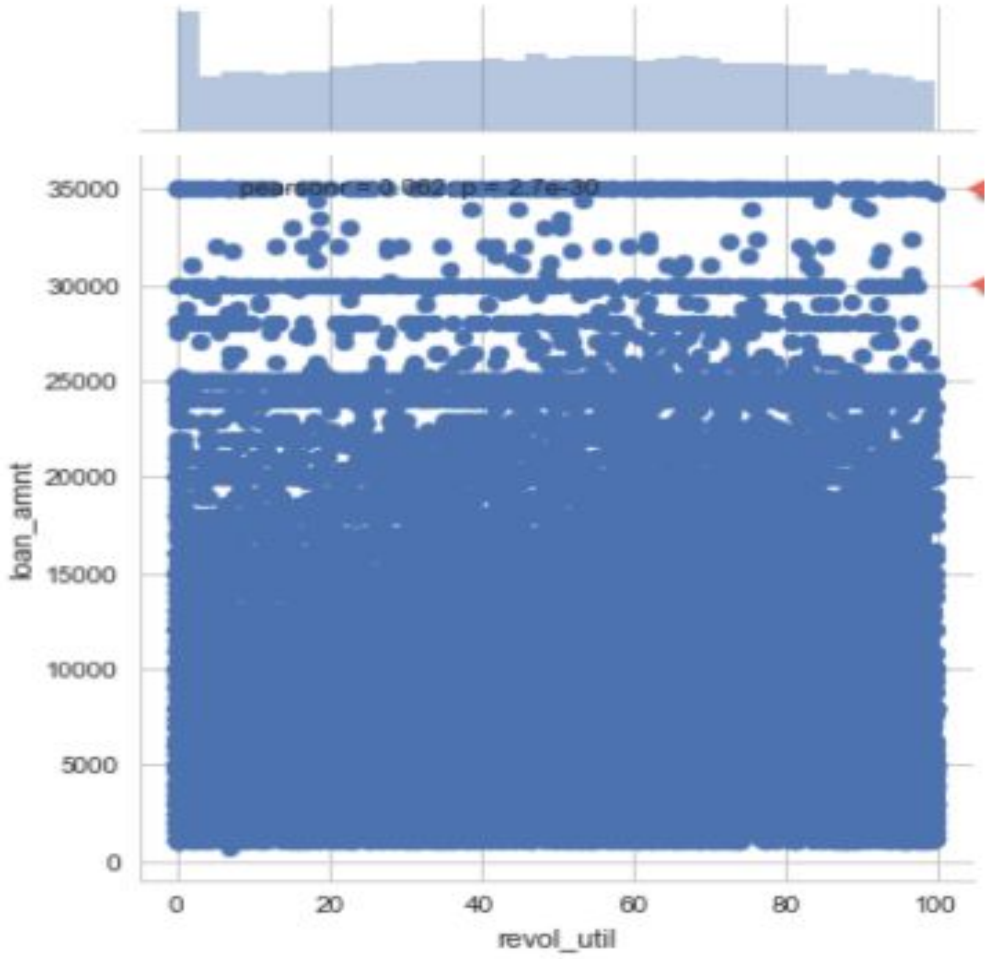
Analysis - Defaults by Revolving Line Util Rate





People with high utilization of Revolving Line of Credit at the time of taking loan default more.
Loans with utilization > 75% are risky.

Analysis - Defaults by Revolving Line Util Rate continued



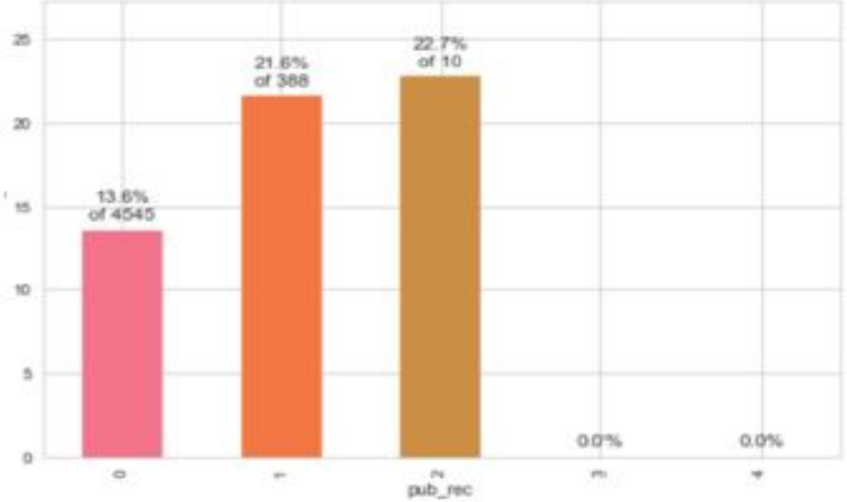
There have been some high value loans extended to borrowers with revolving line utilization rate of higher than 75%.

This practice should be stopped. Density of low value loans is also high. They should be approved less often.

Analysis - Defaults by prior bad record

94% have no Public derogatory records.

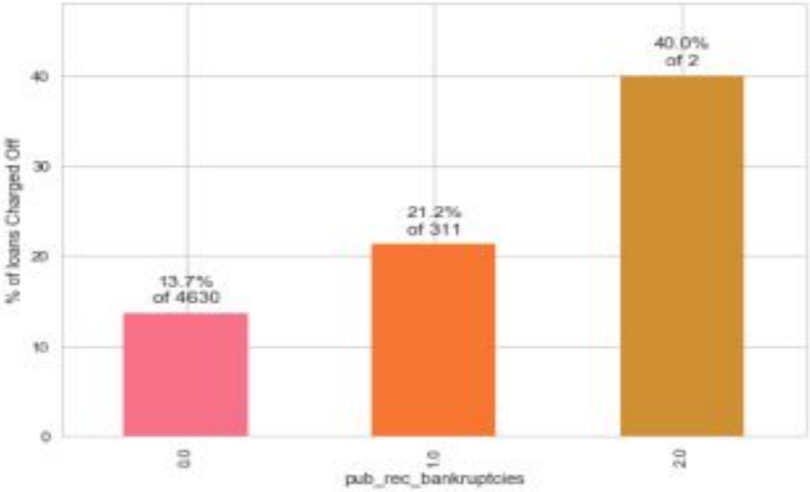
● Having even 1 derogatory record increases the chances of Charge Off significantly.



● 96% have no bankruptcy record.

● Having even 1 bankruptcy record increases the chances of Charge Off significantly.

● Public Derogatory Record and Public Bankruptcy records have 83% correlation. We can use any one.



Recommandations:

1. Stop – approving loans where amount/income is higher than 30%

2. Reduce – number of approvals where purpose is small business

3. Stop – approving high-value loans when revolving line utilization rate greater than 75%

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

5. Start – charging higher interest rates for loans with dti greater than 20