

Lending Club Case Study

GROUP MEMBERS -

1. SUBHABRATA GHOSH
2. MOHAMMED TARIQUE JAMAL

Background – Lending Club Case Study

Background

Lending club is the largest peer-to-peer marketplace connecting borrowers with lenders. Borrowers apply through an online platform where they are assigned an internal score. Lenders decide 1) whether to lend and 2) the terms of loan such as interest rate, monthly instalment, tenure etc.

Some popular products are credit card loans, debt consolidation loans, house loans, car loans etc.

Business Objective

To identify variables which are strong indicators of default and potentially use the insights in approval/rejection decision making.

The problem Statement

Company

Lending Club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.

Context

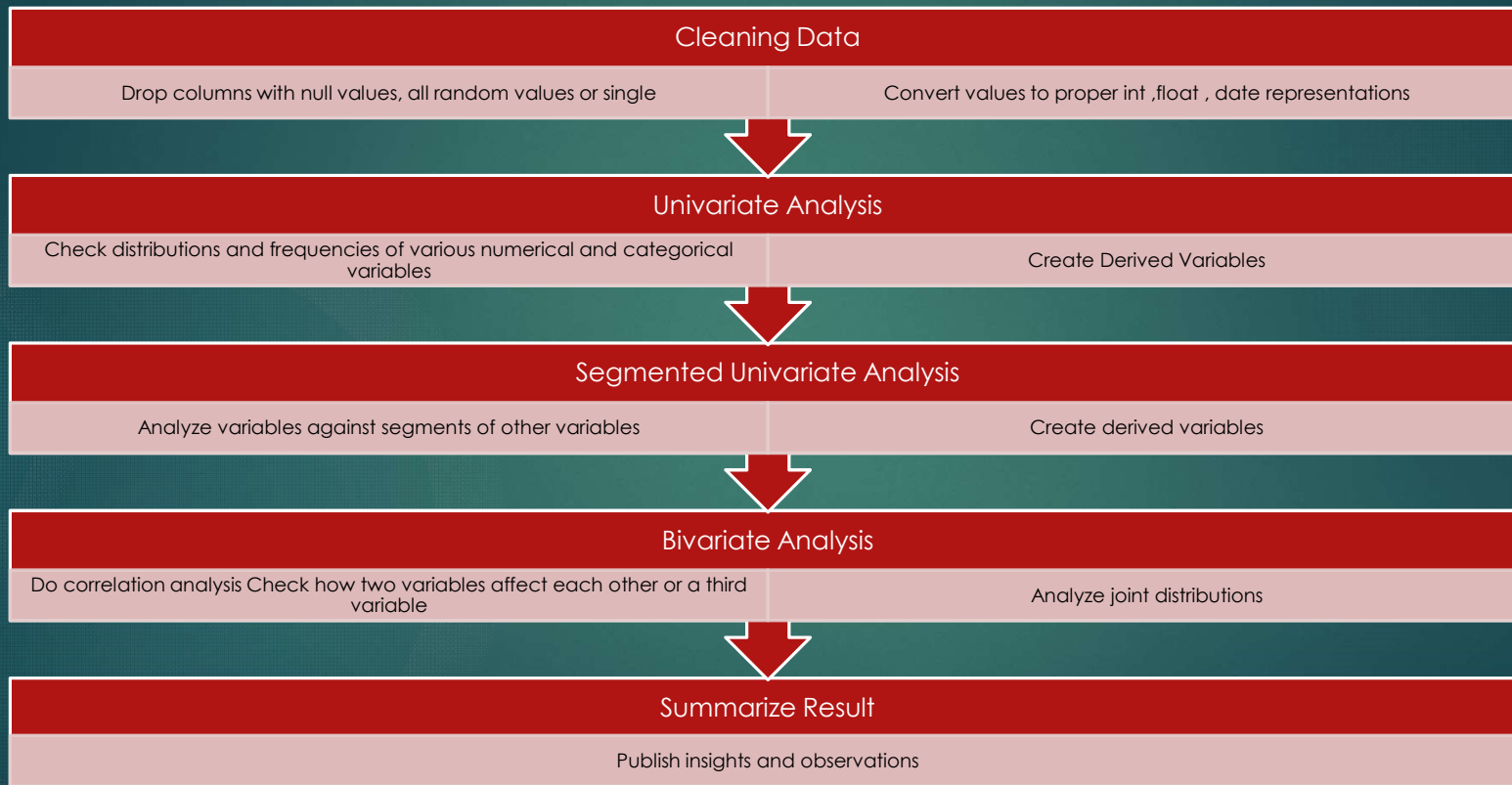
Lending Club wants to understand the driving factors behind loan default, i.e. the driver variables which are strong indicators of default.

The company can utilize this knowledge for its portfolio and risk assessment.

Problem Statement

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



Data Understanding

Types of variables

- Customer (applicant) demographic
- Loan related information & characteristics
- Customer behavior (if the loan is granted)

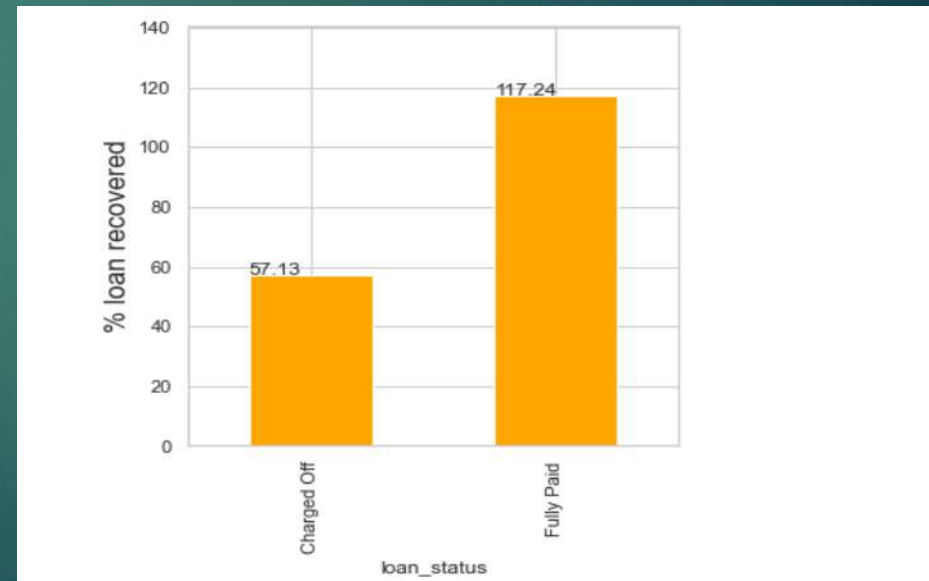
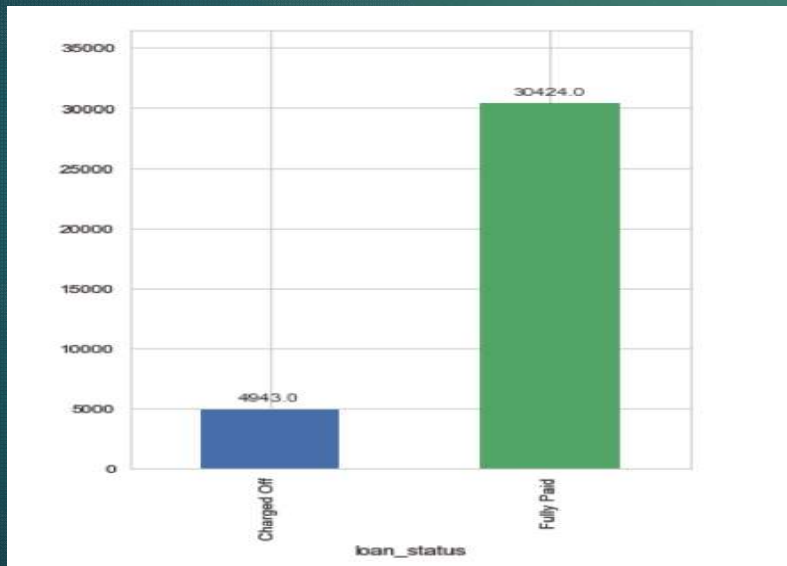
| Customer's Demographics |
|-------------------------|
| Employement Length |
| Employement title |
| Annual Income |
| Zip Code |
| Description |

| Loan Information & Characteristics |
|------------------------------------|
| Loan Amount |
| Funded Amount |
| Funded Amount Investment |
| Interest Rate |
| Loan Status |
| Laon Grade |

| Customer Behaviour variables |
|------------------------------|
| Delinquency year -2 |
| earliest credit line |
| Revolving balance |
| Recoveries |
| Application type |
| Loan purpose |

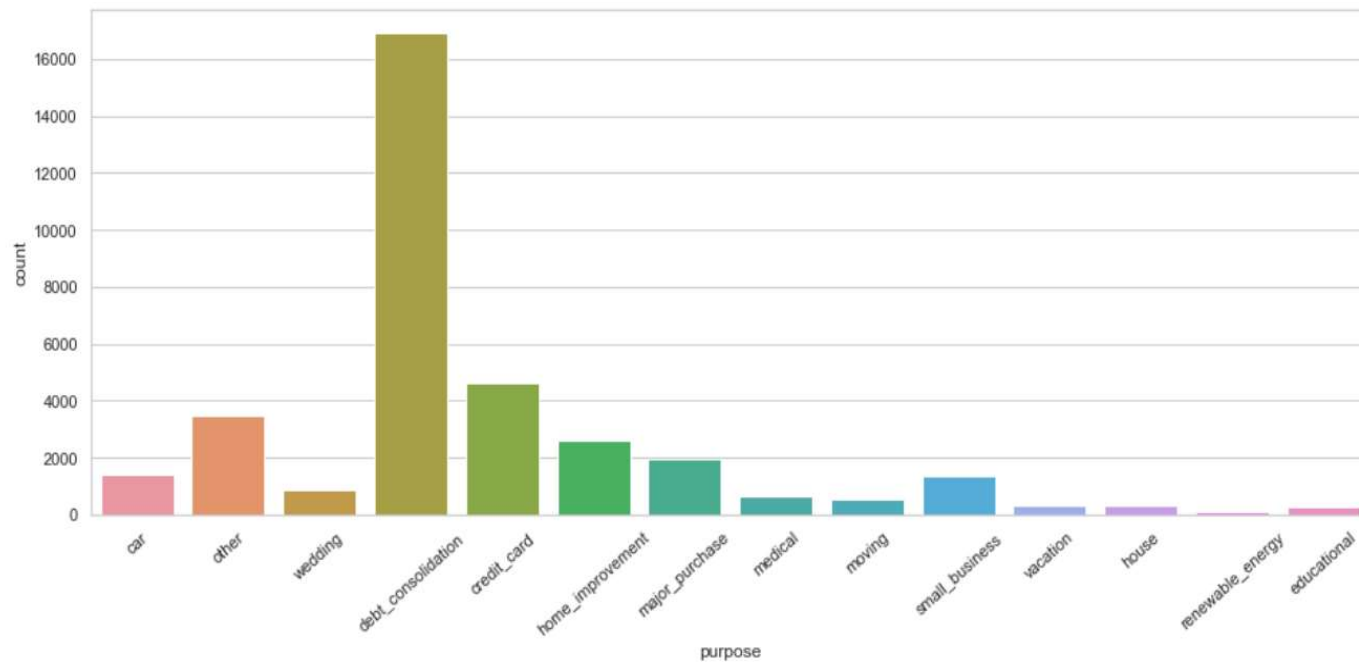
Analysis – Overall Loan Status

- ▶ Approximately 14% of the loans are defaulted
- ▶ Lending club recovers only 57% of the loan amount when it is defaulted.



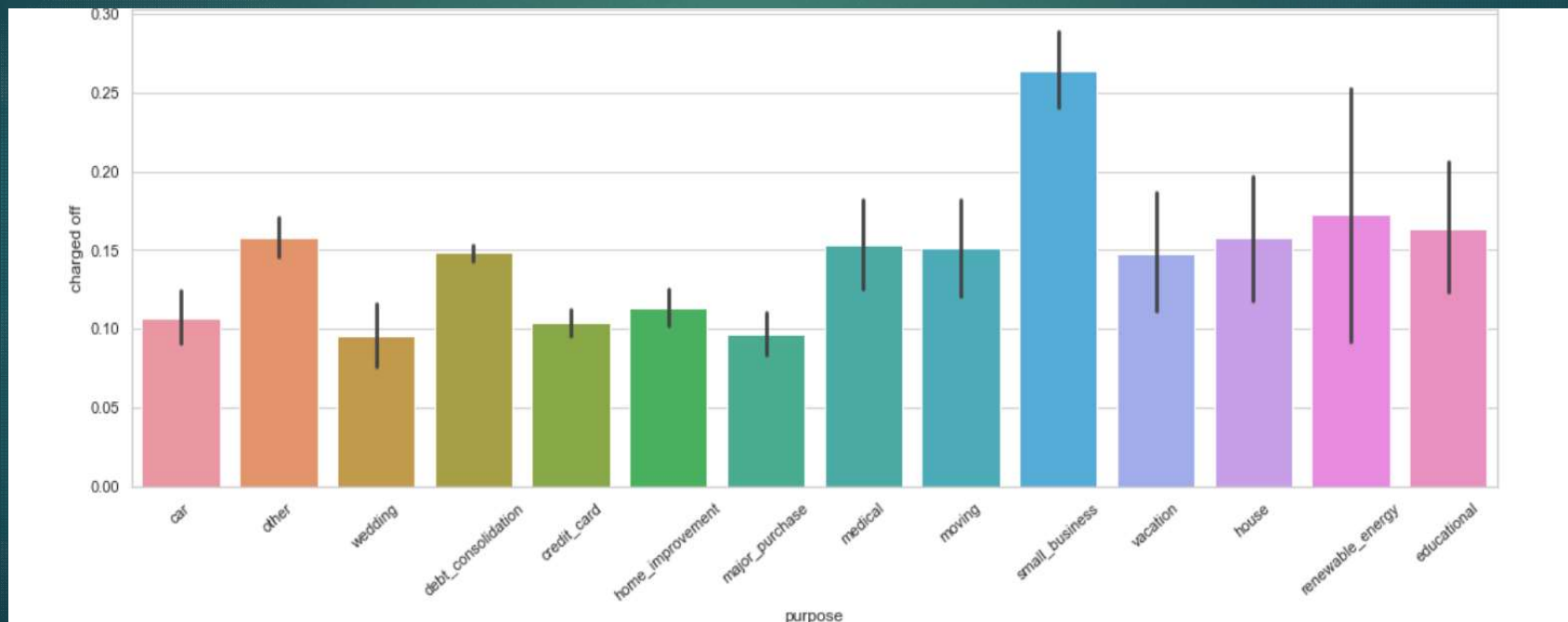
Analysis – Loan Understanding - Purpose

- ▶ Maximum loans are taken for debt consolidation followed by credit card, Home improvement, Major purchase.



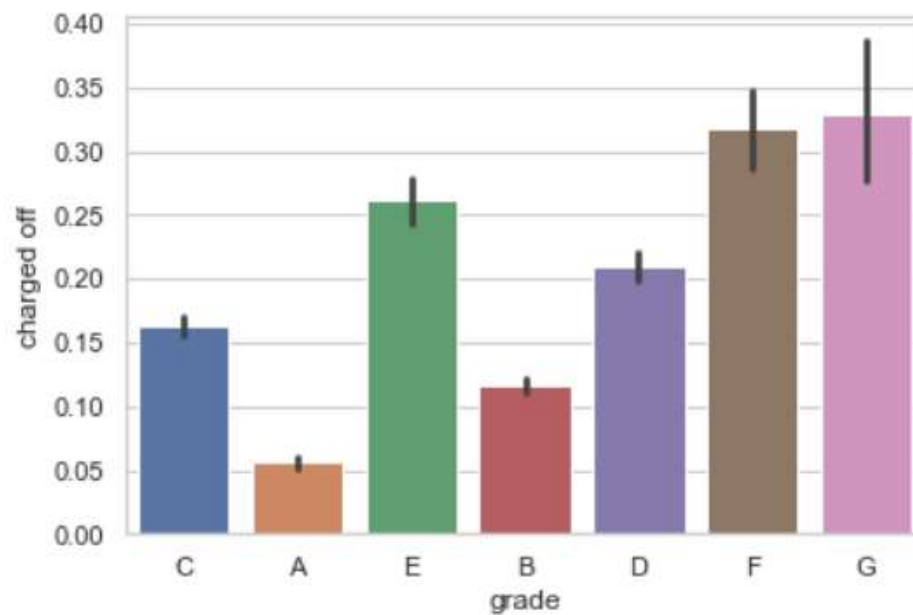
Defaults by Loan Purpose

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



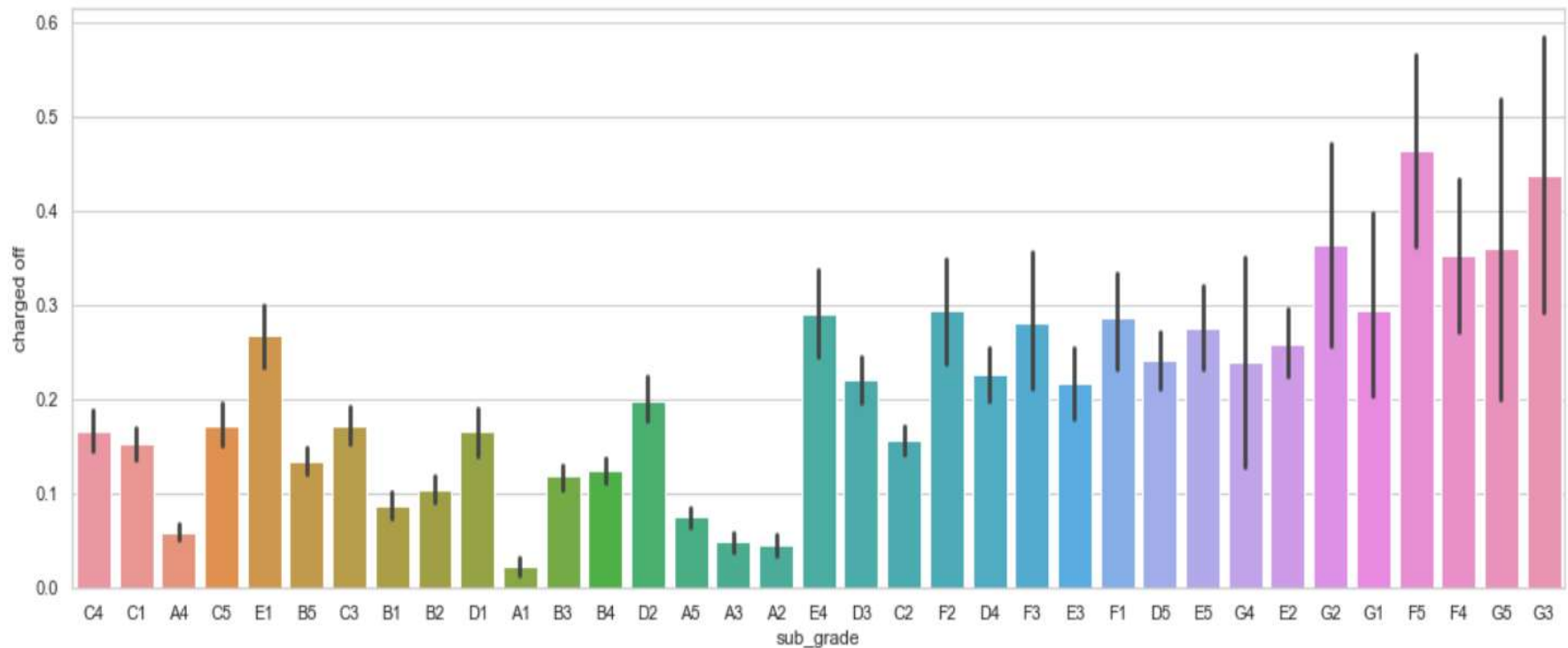
Defaults By Grade & Sub Grade

- The higher default rate happens for the lower grades. So, the grading system is working.



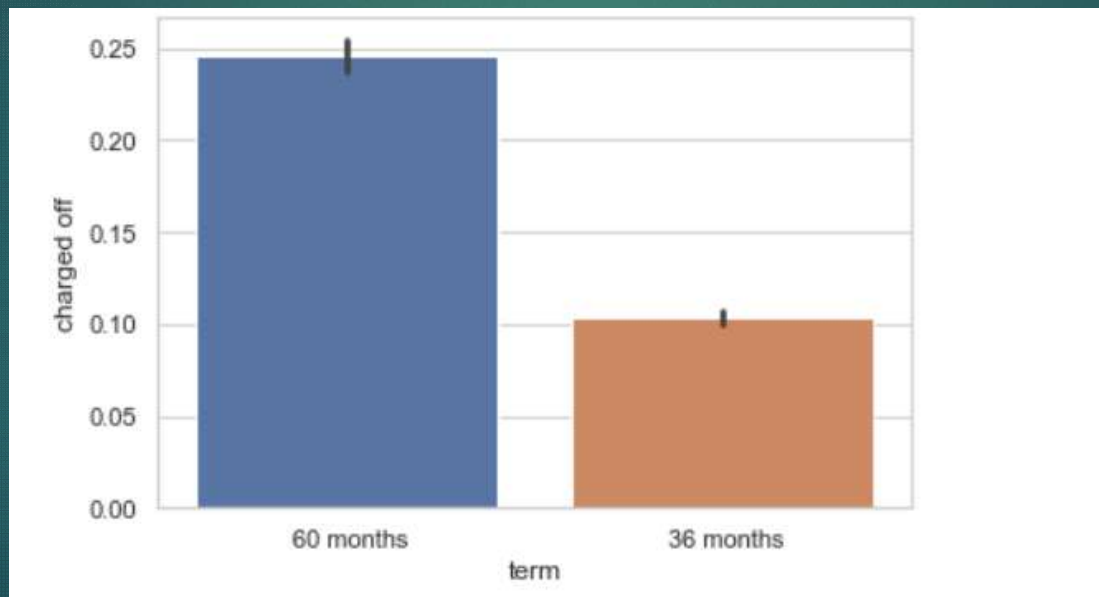
Defaults By Grade & Sub Grade

- ▶ The higher default rate for the lower sub grade. The sub grading system is also working fine.



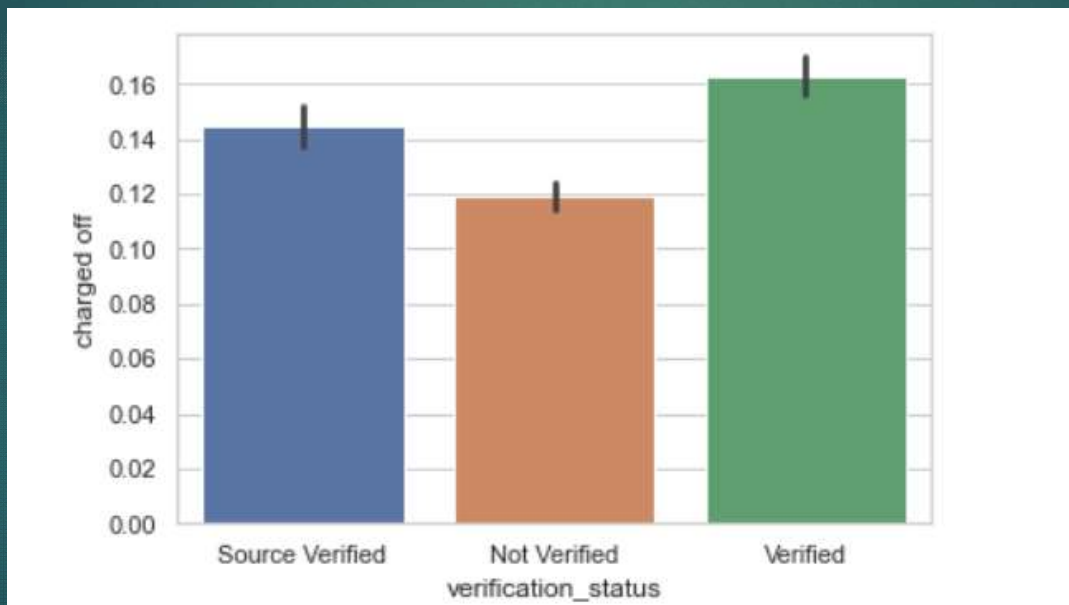
Defaults By Term

- ▶ The long term loans are more likely to get defaulted. The chances are more than double for the 60 months loans. So, Lending club should focus on giving short term loans.



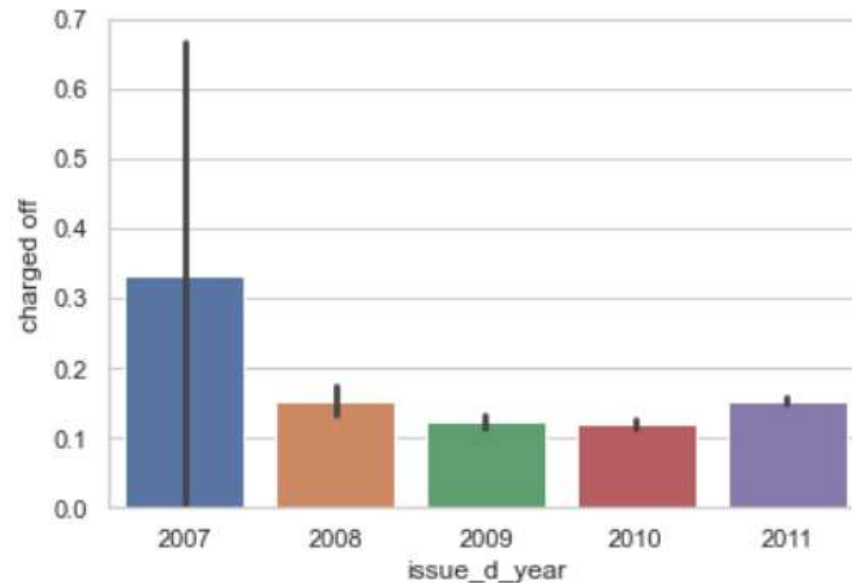
Defaults By Verification

- ▶ Most loan defaults happen for the verified loans. The verification system needs to be improved.



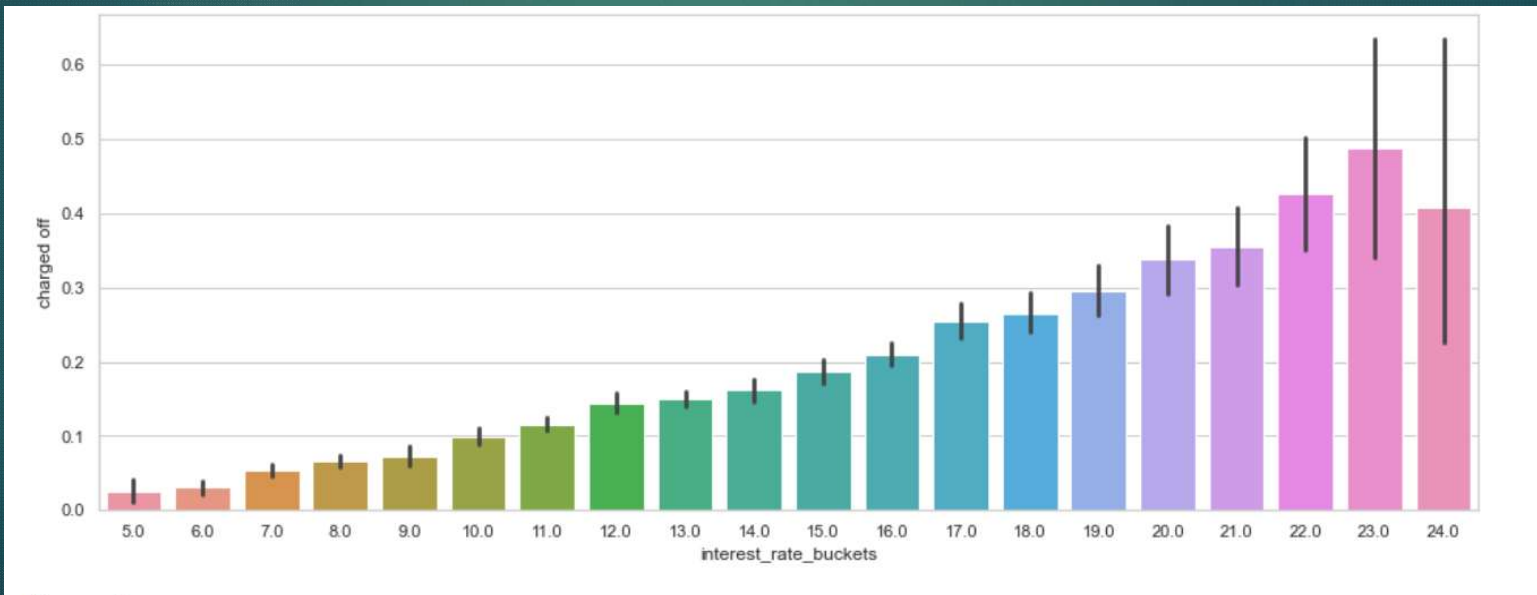
Defaults By Loan Year

- ▶ Most bad loans get approved in 2007.
- ▶ The default loans were decreasing from 2008 to 2010. However there is an increase in the year 2011.



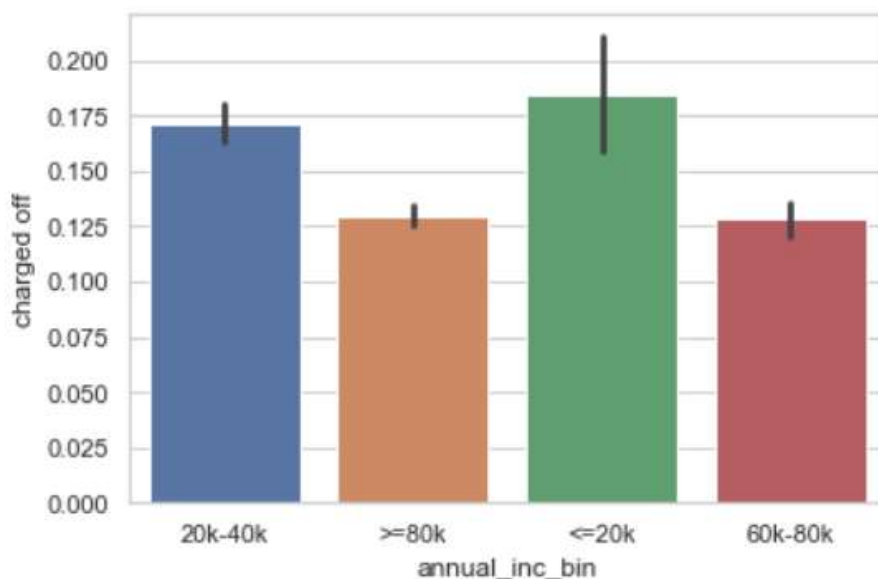
Defaults By Interest Rate

- ▶ Percentage of Defaults increases monotonically with higher interest rates. At rates of 19% and above, more than 33% of loans are Charged Off.



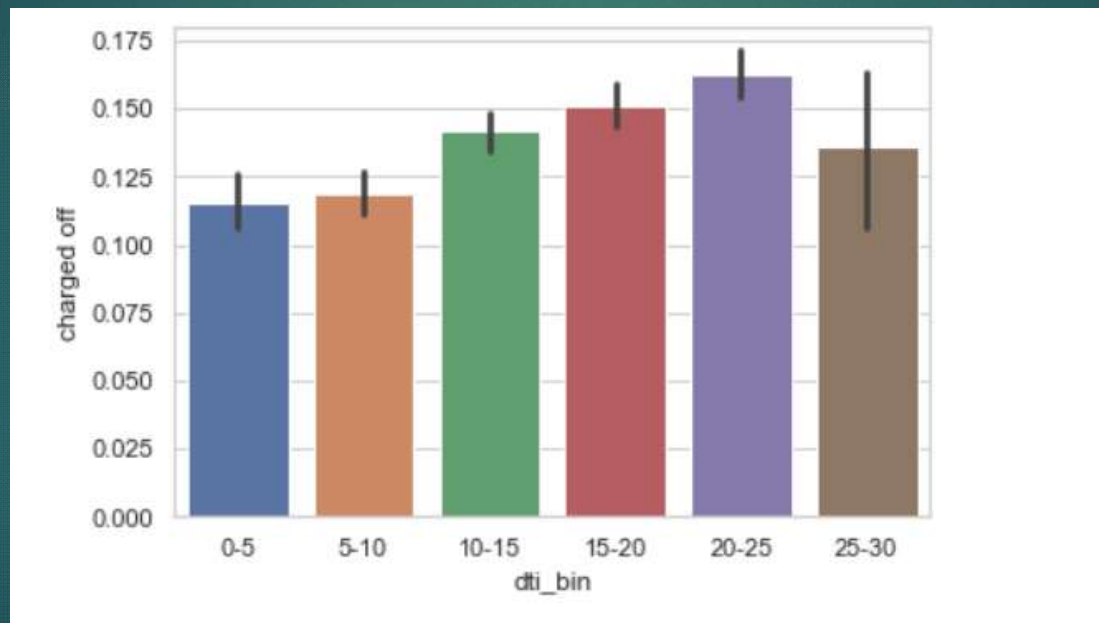
Defaults By Borrower's Income

- ▶ Borrowers having annual income less than 20000 default on their loans at much higher rates.
- ▶ Loan default decreases with higher annual income.



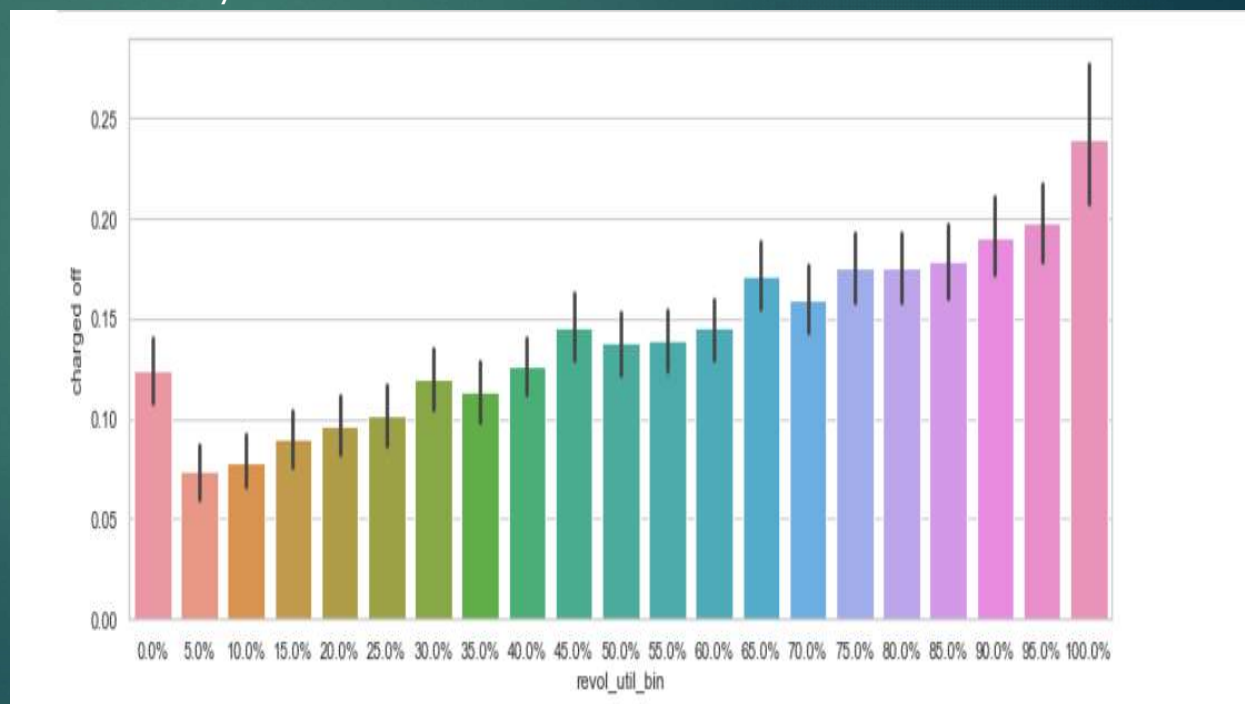
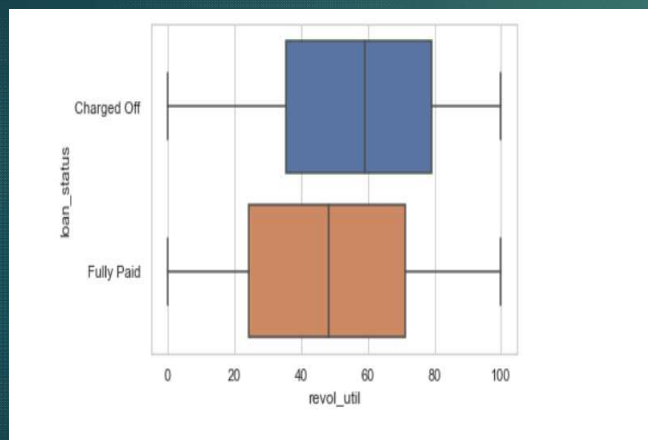
Defaults By Debt to Income Ratio

- ▶ Percentage of default rises with dti ratio.
- ▶ As the dti ratio rises above 20, the loans become risky.



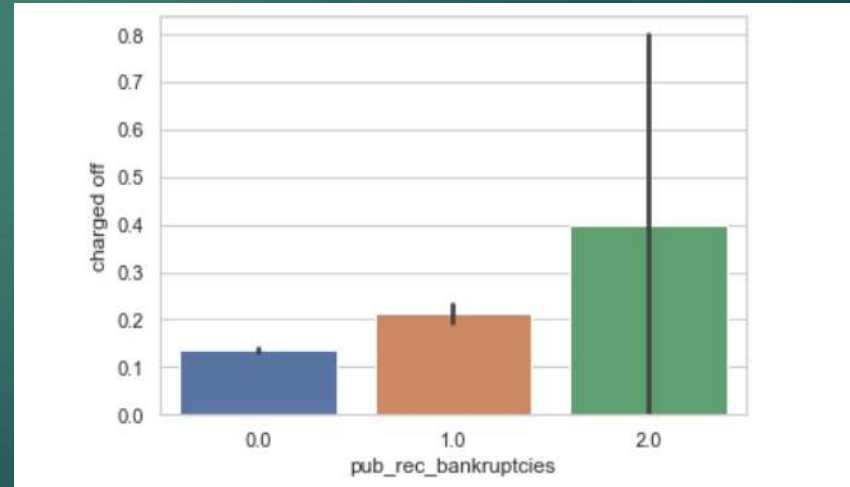
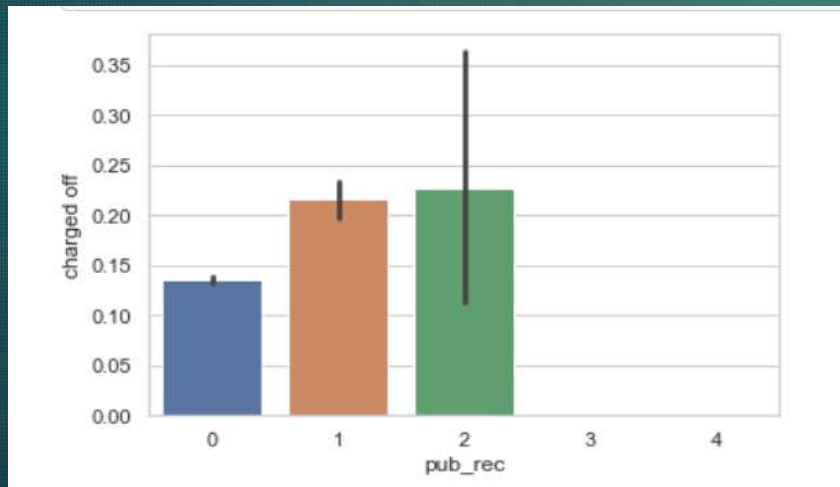
Defaults By Revolving Line Util Rate

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



Defaults by prior bad record

- ▶ 94.77% 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.



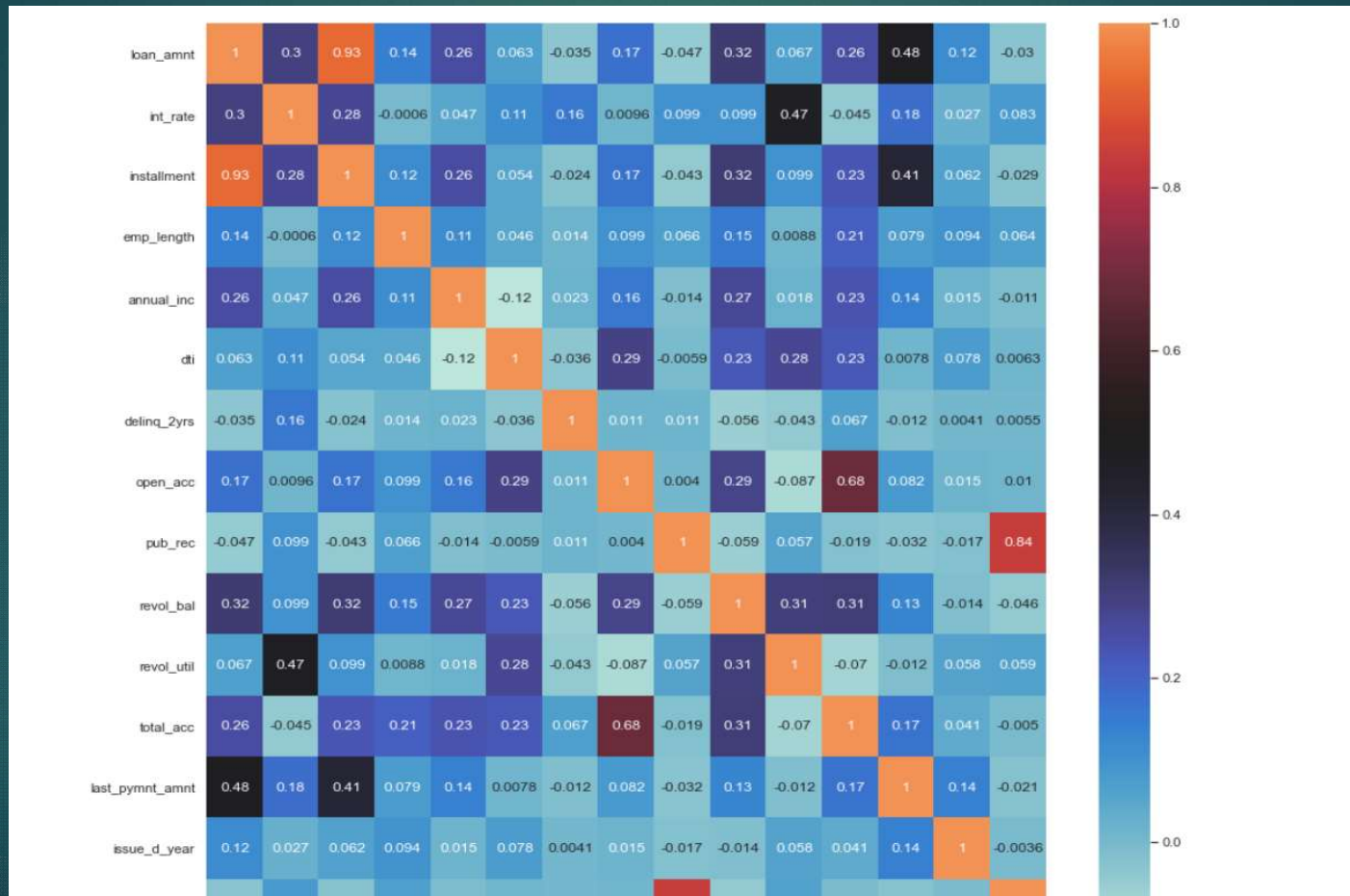
Recommendation(s)

- ▶ Stop – approving loans where amount/income is higher than 30%
- ▶ Reduce – number of approvals where purpose is small business
- ▶ Stop – approving high-value loans when revolving line utilization rate greater than 75%
- ▶ Stop – approving loans to people with prior bad record. Or at least stop approving high-value loans
- ▶ Start – charging higher interest rates for loans with dti greater than 20

Appendix - Correlations

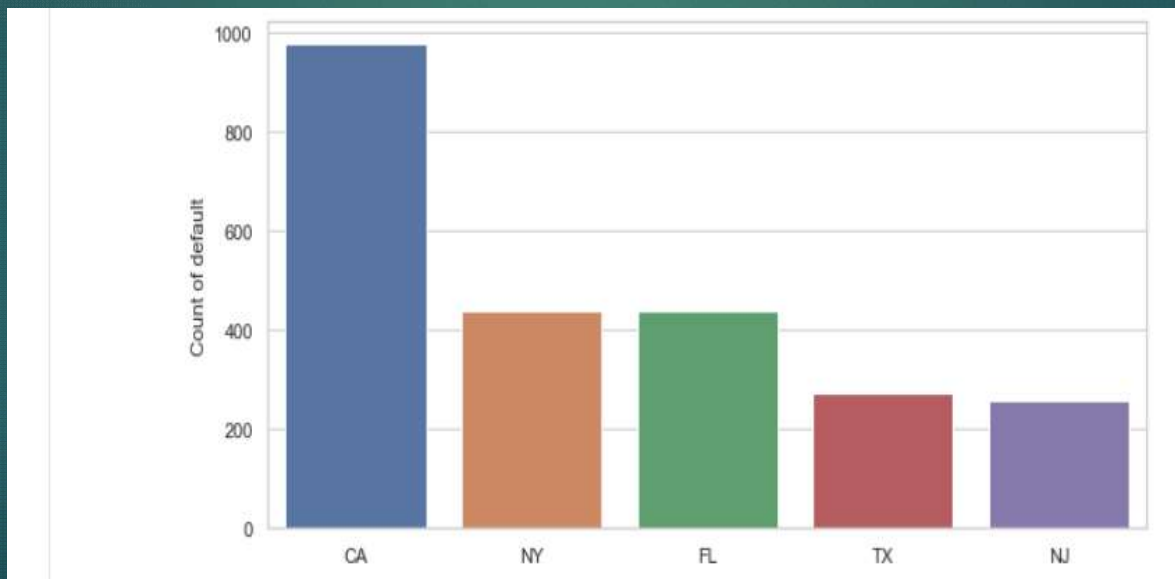
- ▶ Loan Amount is highly correlated (positive) with the installments , interest rates
- ▶ Debt-to-Income Ratio is high correlated (positive) to loan issue year and vice-versa.
- ▶ Interest rate is negatively correlated to total account
- ▶ Public record is high correlated (positive) to public record bankruptcies.

Appendix - Correlations



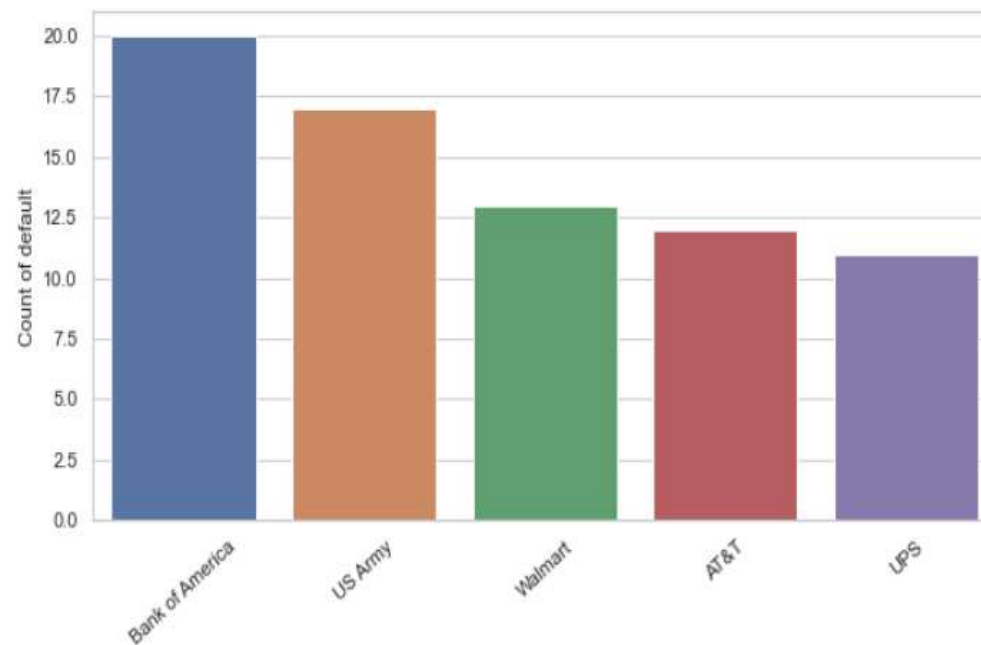
Appendix – Defaults By State

- ▶ Maximum defaulters are in populous states, California, New York, Florida and Texas.



Appendix – Defaults By Employer

- ▶ Maximum defaulters are from large organizations – Bank of America, US Army , Walmart , AT&T etc.



Appendix – Loan Amount

- ▶ The loan amount increases the chances of default are also increases.

