

# Lending Club Case Study

## Group Members

- ▶ Prachi Jain
- ▶ Sachin Kumar



# Problem Statement:



- ▶ Lending Club company is 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.
- ▶ 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'.
- ▶ If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.
- ▶ Identification of such applicants using EDA is the aim of this case study.

# Analysis Approach:



Data Cleaning

Univariate  
Analysis

Segmented  
Univariate  
Analysis

Bivariate  
Analysis

Summary Of  
Result

# Applicant Default trend based on :

- ▶ **Ownership of Home**
- ▶ **Income verification by LC**
- ▶ **By states**
- ▶ **Purpose of Loan**
- ▶ **On LC Grade & Sub Grade basis**
- ▶ **Tenure of loan**
- ▶ **Years of employment**
- ▶ **Derogatory Public records**
- ▶ **No. of Open credit line Accounts**
- ▶ **No of Inquiries in last 6 months**
- ▶ **Earliest credit line**
- ▶ **Debt to Income Ratio**
- ▶ **Amount of loan**
- ▶ **Loan Interest Rate**

# No. of Defaulter under different parameters

## ►Ownership of Home -

Ownership	No. of Defaulter
RENT	2839
MORTGAGE	2327

## ►Income verification of LC-

Not Verified	2142
Verified	2051
Source Verified	1434

# No. of Defaulters under different parameters

- **By States** - The Top4 states are California, New York, Florida and Texas.

CA	1125
FL	504
NY	495
TX	316

- **Purpose of loan-**

debt_consolidation	2767
other	633
credit_card	542
small_business	475

# No. of Defaulters under different parameters

## ► On LC Grade & Sub Grade Basis -

grade	sub_grade	
B	B5	356
	B3	341
C	C1	336

## ► Tenure of Loan

Tenure in months	
36	3227
60	2400

# No. of Defaulters under different parameters

## ► Years of Employment-

10+	1331
<1	867

## ► Derogatory public records-

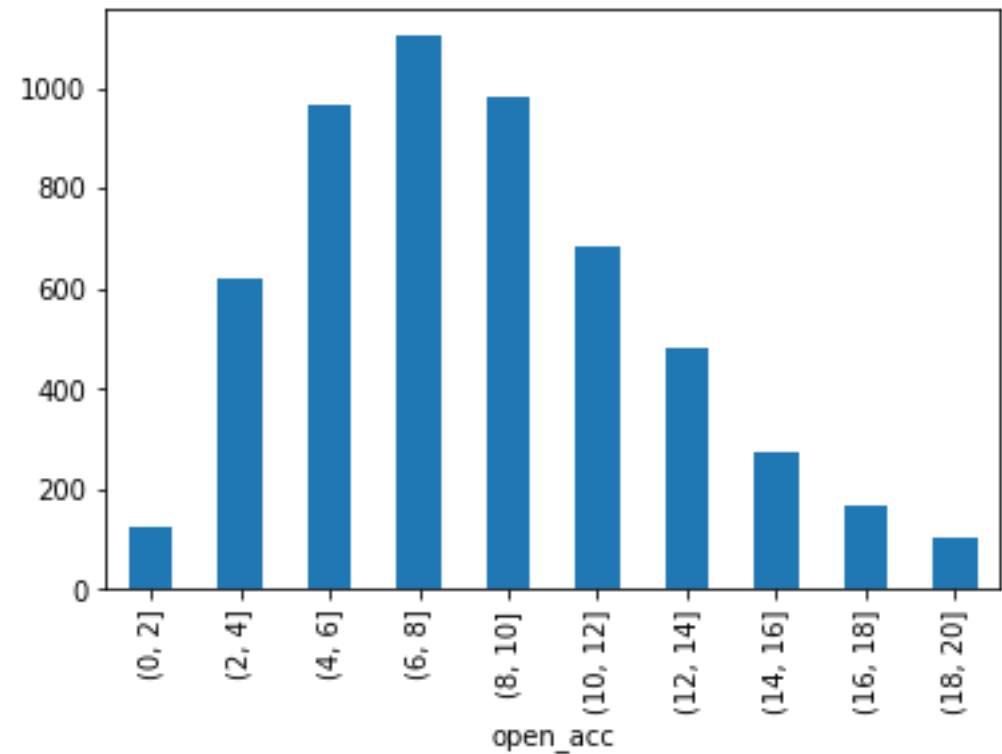
0	5160
1	457



# No. of Defaulters under different parameters

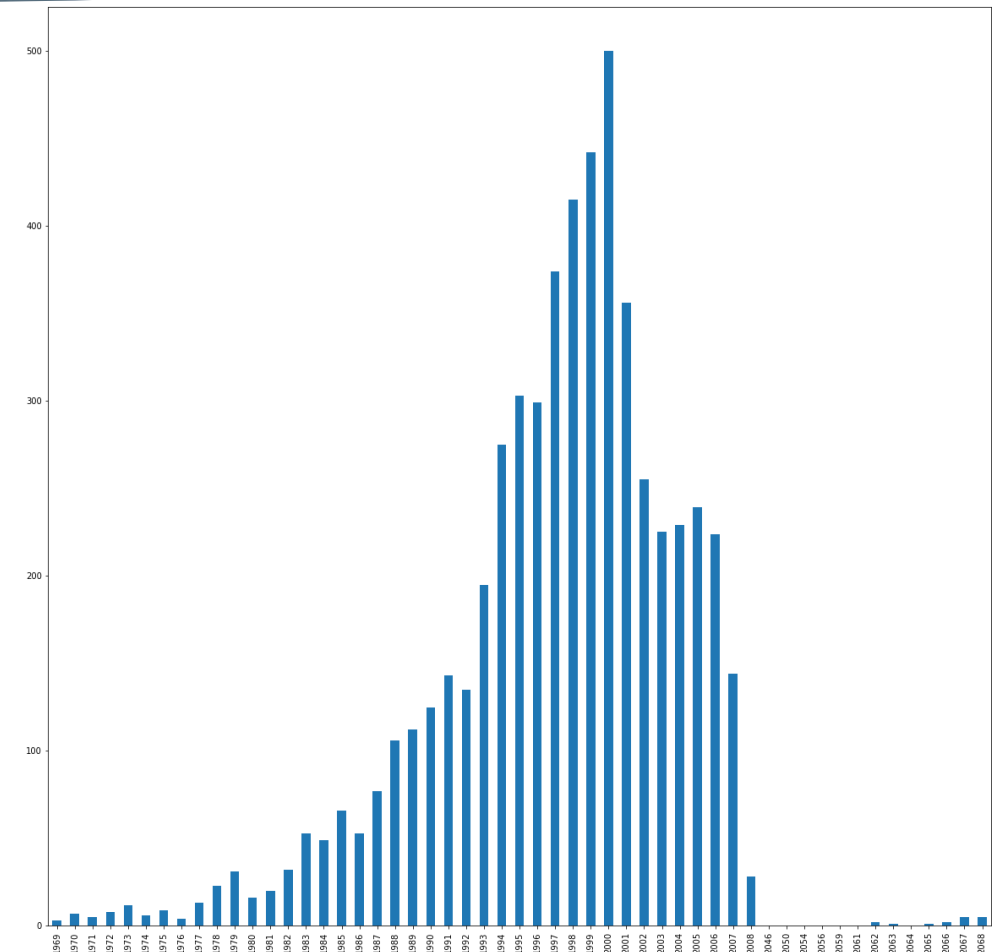
- ▶ No. of Open credit line Accounts
  - 6-8 Highest defaulters
- ▶ No of Inquiries in last 6 months

0	2280
1	1677



# No. of Defaulters under different parameters

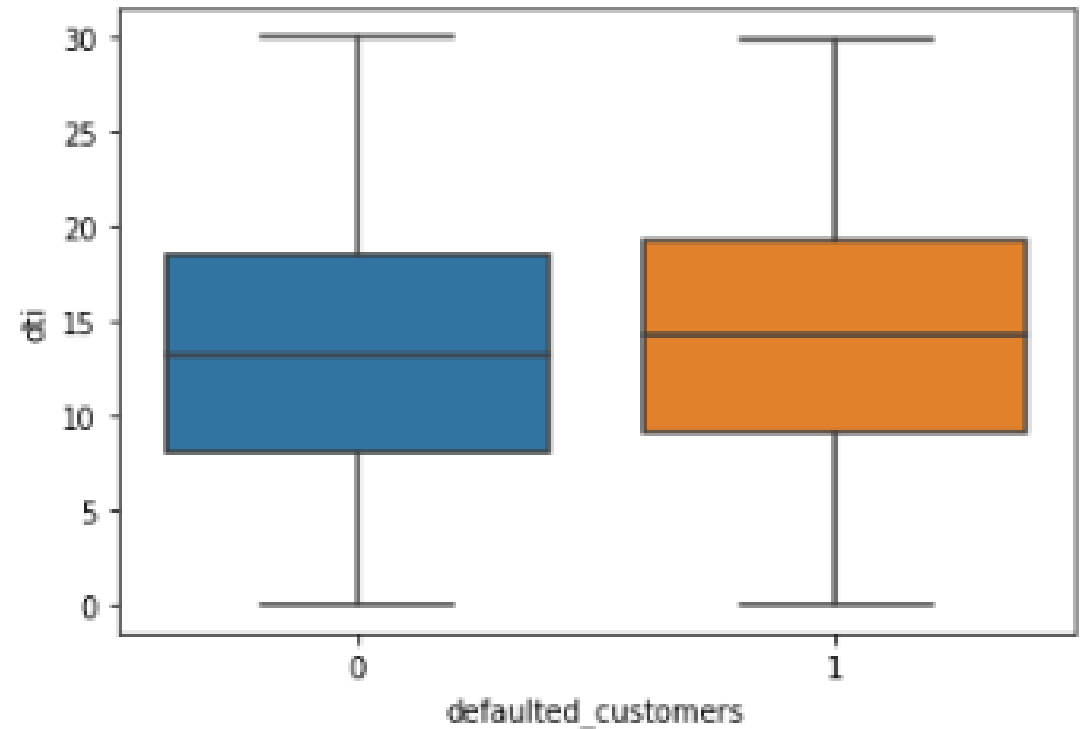
► Earliest credit line- The borrower who have started their credit line in the year 2000, i.e. 11 years from 2011 are risky borrowers



# No. of Defaulters under different parameters

- ▶ The applicants whose debt to income ratio is more are risky borrowers.

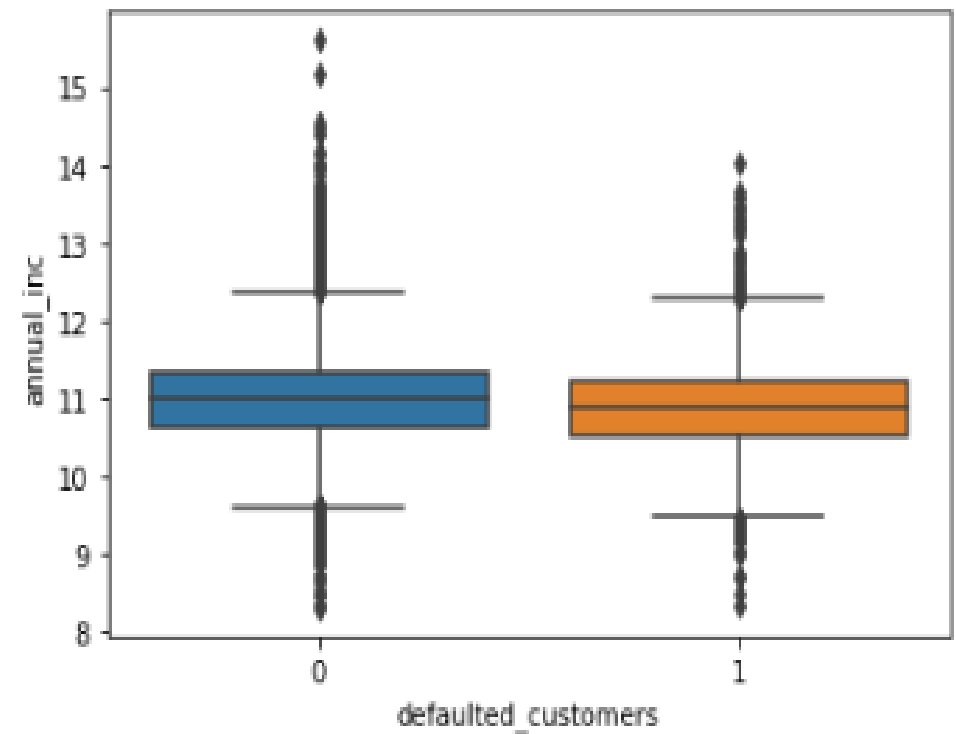
<b>Count</b>	5627
<b>mean</b>	14
<b>std</b>	6.58
<b>min</b>	0
<b>max</b>	29.85



# No. of Defaulter under different parameters

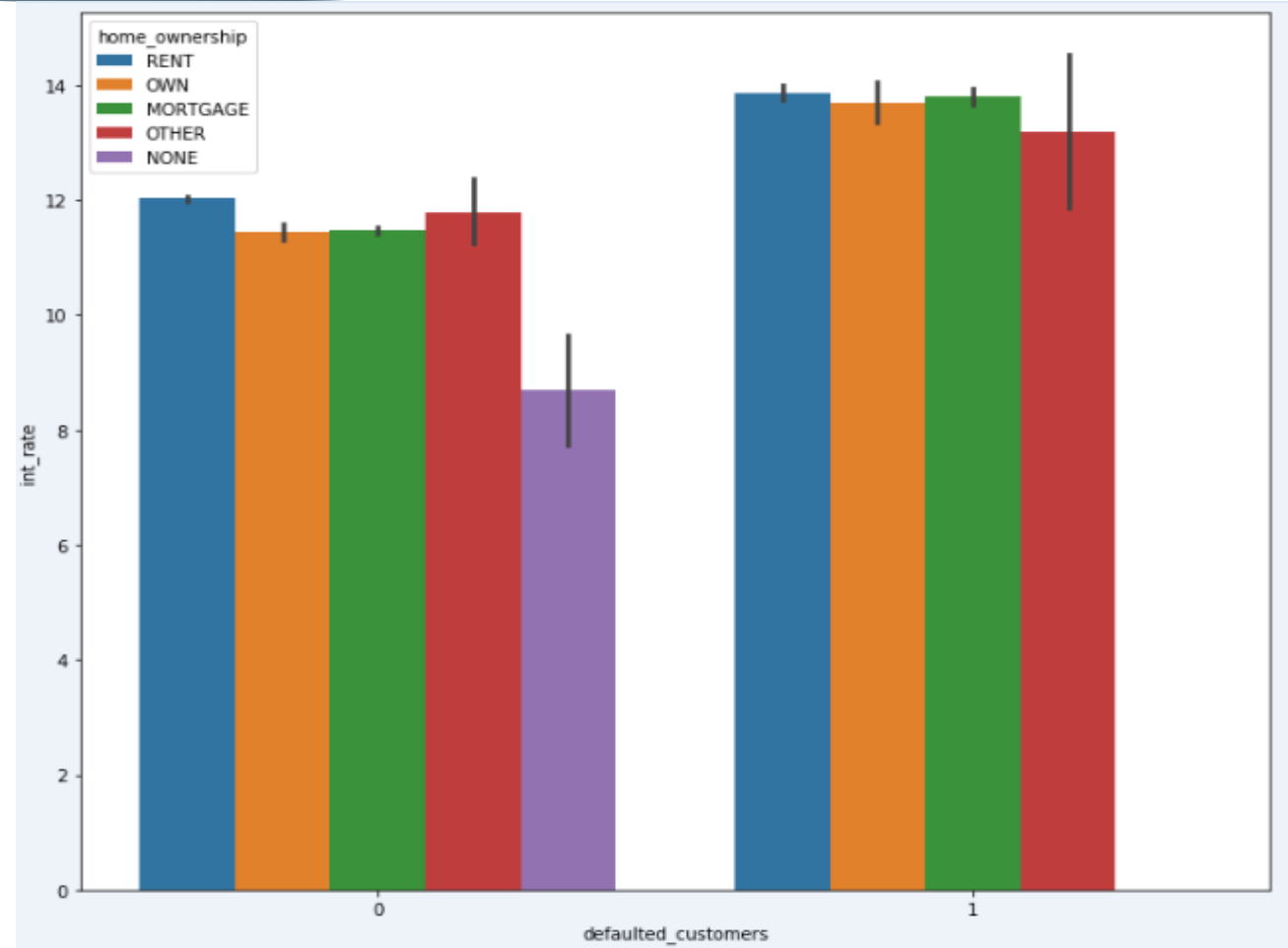
- ▶ The applicants whose annual income is less are risky borrowers.

<b>Count</b>	5627
<b>mean</b>	6247.30
<b>std</b>	47776.01
<b>min</b>	4080
<b>max</b>	1250000



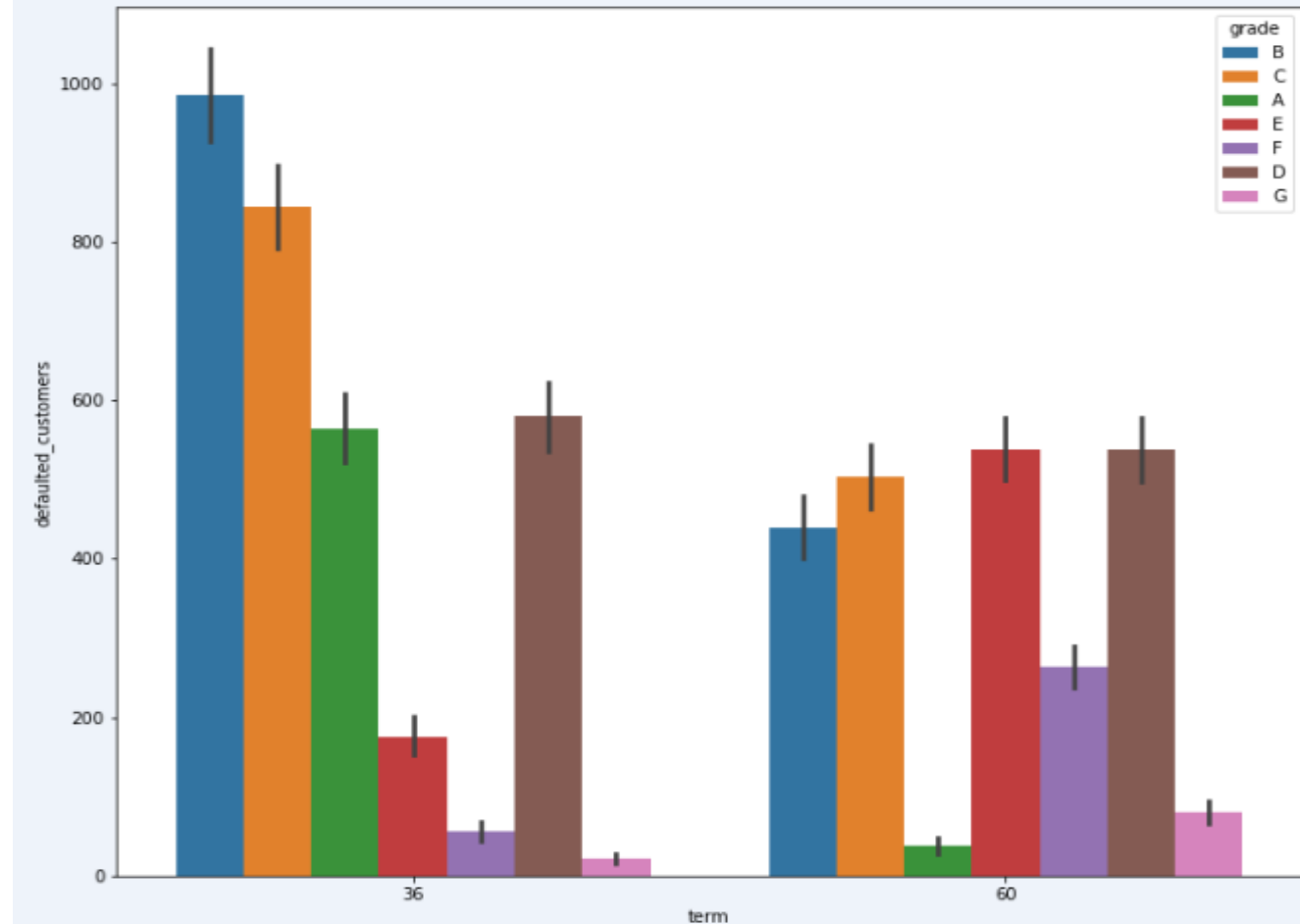
# Defaulters vs int\_rate vs home\_ownership

- ▶ Most of the defaulted customer are on Rent and the interest Rate is high.



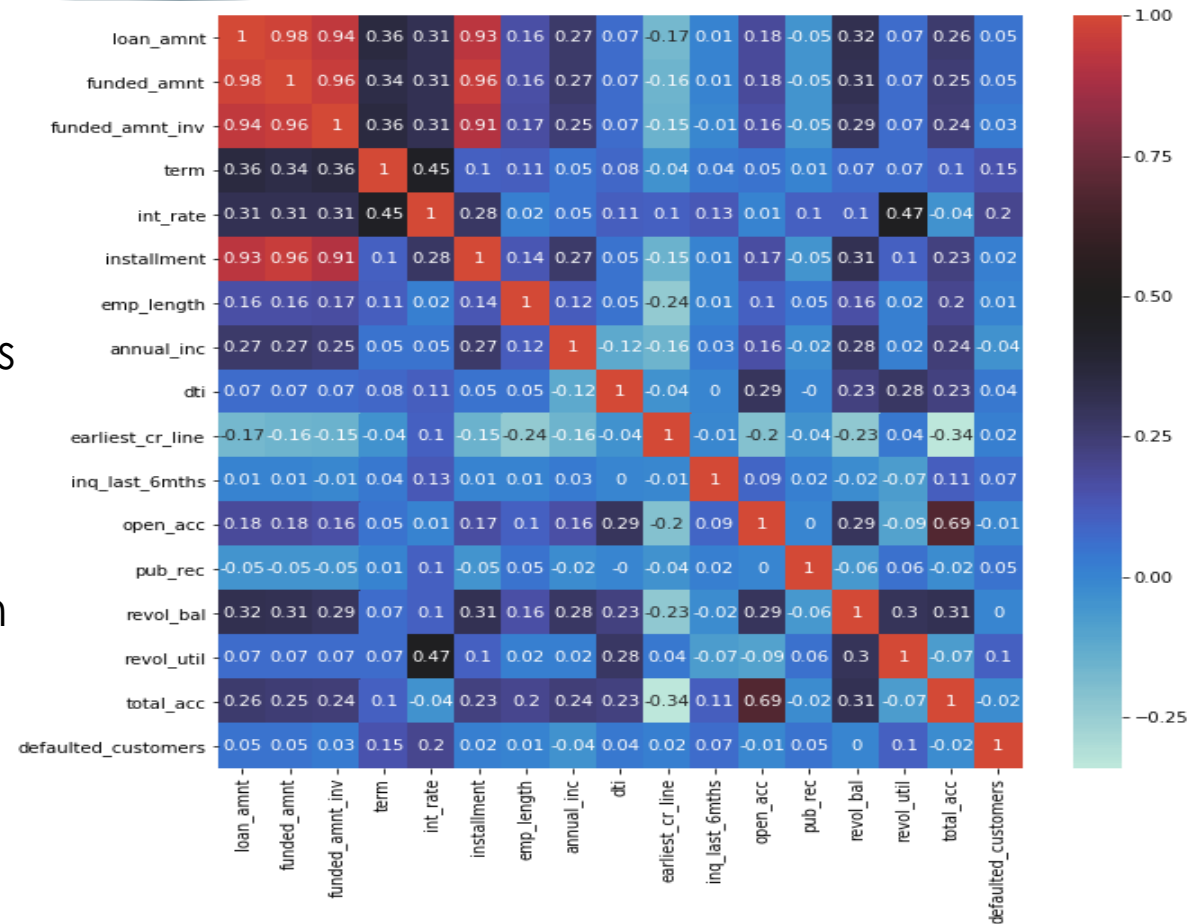
# Defaulters vs Term vs Grade

- ▶ Defaulter customers are very high when term is 36 for B & C Graded customers.
- ▶ For Term 60, grade E & D has most no. of defaulters.



# Heat map:

- ▶ Negative Correlation : When any one value among i.e. annual\_income & open\_acc & total\_acc increase the defaulted\_customers decrease.
- ▶ Since there is no correlation with revol\_bal, it has no impact.
- ▶ Rest attributes, i.e. loan\_amnt, funded\_amnt, funded\_amnt\_inv, term, int\_rate, installment, emp\_length, dti, earliest\_cr\_line, inq\_last\_6mths, pub\_rec, revol\_util has positive impact on defaulted\_customers.



# Summary of Results

• Home ownership as 'Rent' or 'Mortgage' are risky borrowers.



```
graph TD; A[• Home ownership as 'Rent' or 'Mortgage' are risky borrowers.] --> B[• Applicants living in California, New York or Florida are risky.]; B --> C[• Applicants having more the 10 years or less than 1 years of employment length are risky.]; C --> D[• Applicants whose income does not get verified are risky.]; D --> E[• Applicants applying loan for debt consolidation are risky.];
```

• Applicants living in California, New York or Florida are risky.

• Applicants having more the 10 years or less than 1 years of employment length are risky.

• Applicants whose income does not get verified are risky.

• Applicants applying loan for debt consolidation are risky.



# Summary of Results

• Applicants with no derogatory public records are riskier borrowers.

• The Applicants whose income is not verified by lending club are risky.

• Applicants with total Account between 10-15 are risky.

• Applicants with no inquiries in last 6 months are riskier borrowers than others.

• The applicants whose debt to income ratio is more are risky.

Thank You