# Lending Club Case Study

Aditya Kintali

Nikhil Bansal

# AGENDA

- Problem Statement 01
- Key Objective 02
- Data Sourcing & Understanding 03
- Data Cleaning 04
- Data Analysis 05
- Conclusions 06

# PROBLEM STATEMENT

Underwriting risk analysis for a consumer finance company

The company has the below 2 types of risks associated with any given loan application:

1. If the applicant is likely to repay the loan, then not approving the loan will result in a loss of business.

2. If the applicant is not likely to repay the loan, then approving the loan may lead to a financial loss.

# KEY OBJECTIVE

Help the business in better decision making

## FIND DRIVING FACTORS BEHIND A LOAN DEFAULT

These factors will be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

# Data Sourcing & Understanding

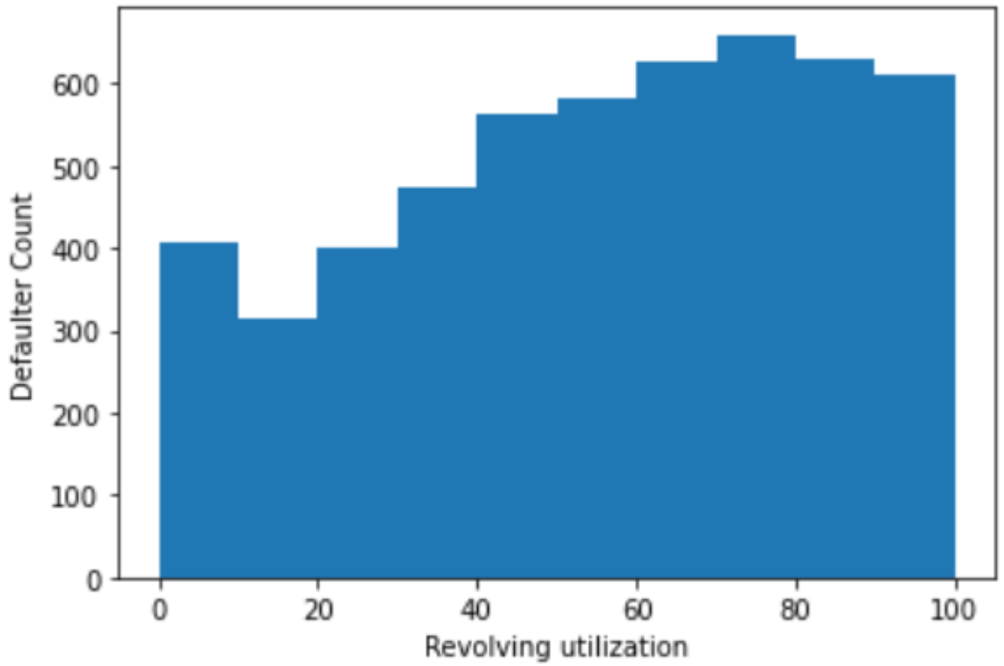
- The given dataset is all the loans issued between the time period 2007 to 2011
- The initial dataset has 39717 rows and 111 columns with mixed data types.
- The understanding started with analyzing the columns in the given data set:
    1. 54 columns have complete null values
    2. Further 4 columns have mostly null values
    3. Further 7 columns was found to have just one value
    4. Identified 4 columns that do not add any value to the analysis

# Data Cleaning

- The 69 columns (54+4+7+4) mentioned in the previous slide were dropped.
- The null values in “emp_length” column were replaced with “0” and removed the extra character “+” at the end.
- Removed extra “%” character from “int_rate” and “revol_util” columns.
- “Unknown” values in “title” and “pub_rec_bankruptcies” columns were handled.
- Converted relevant columns to numeric data type.
- Dropped the rows having just null values
- It finally leaves us with (39598 rows and 42 columns) for analysis.

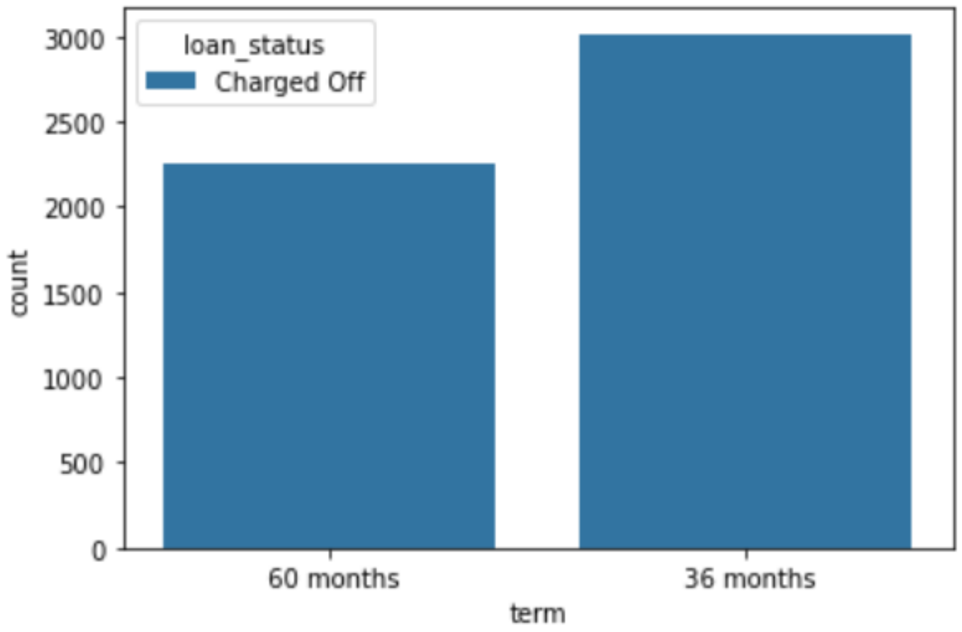
# Conclusion – 1

- Customers who are revolving their credit more are more likely to default.



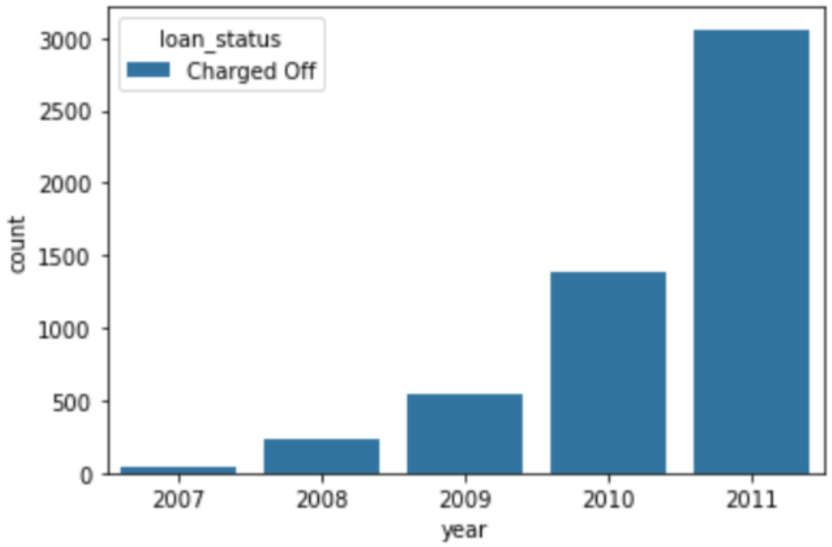
# Conclusion – 2

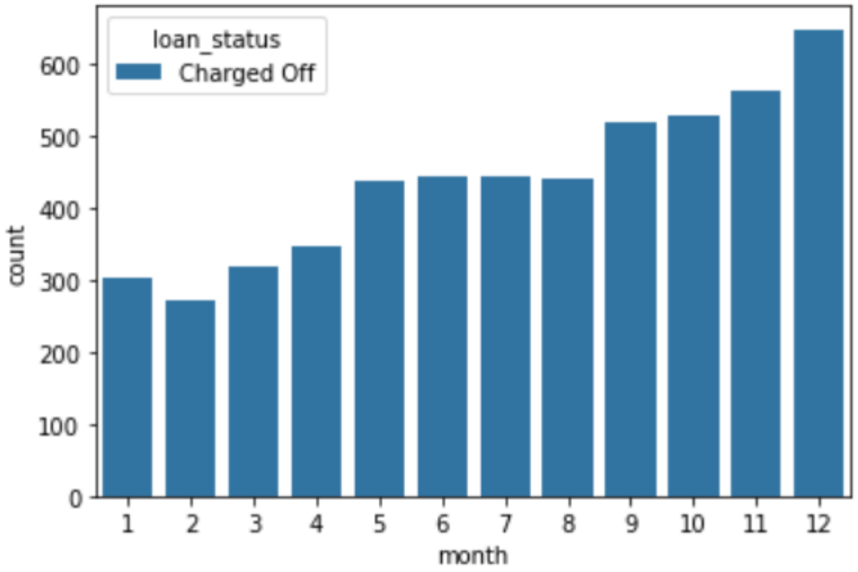
- The shorter the tenure of the loan, more is the likelihood of default.



# Conclusion – 3

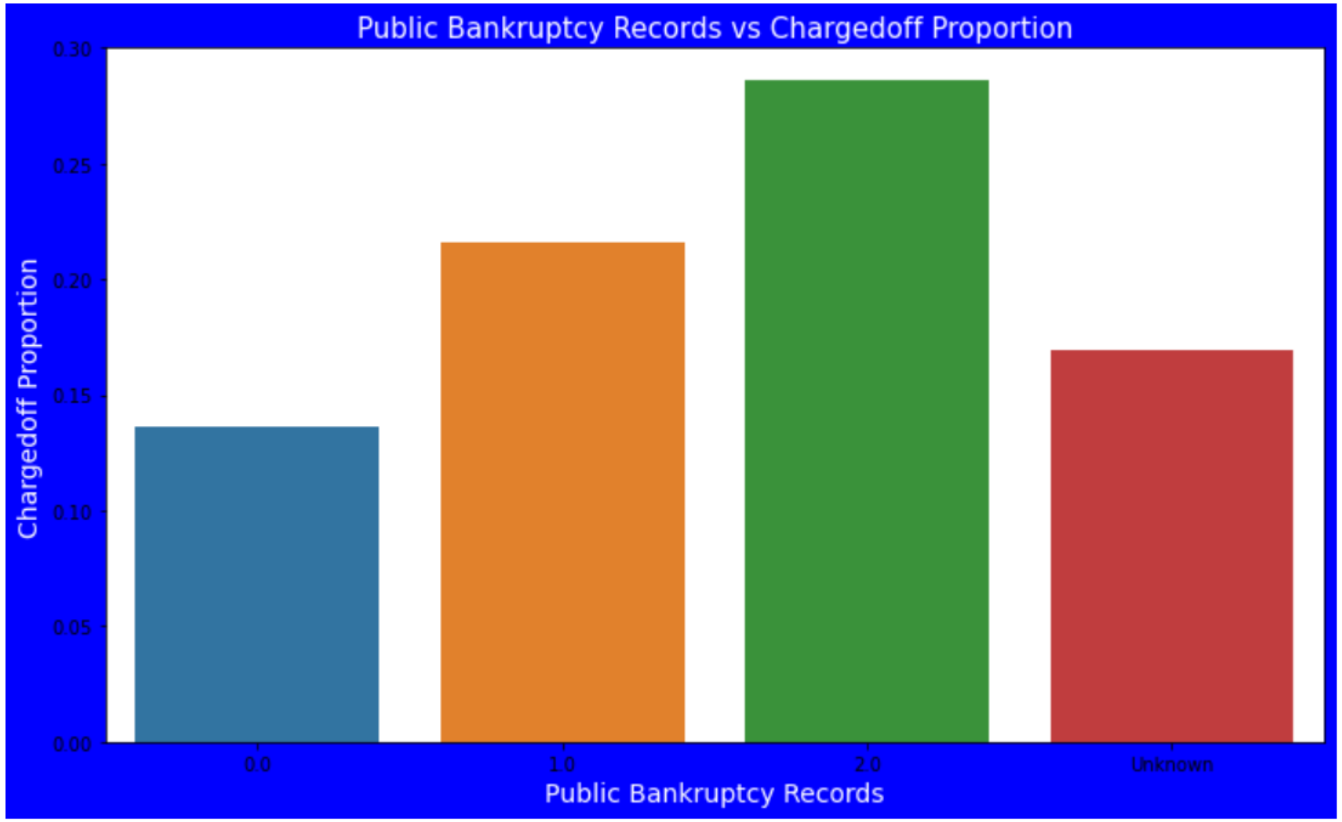
- Number of customers defaulting loan is almost increasing by double every year, which states the need to bring in better loan approval mechanism.
- Maximum numbers of loans are approved towards the financial year end.



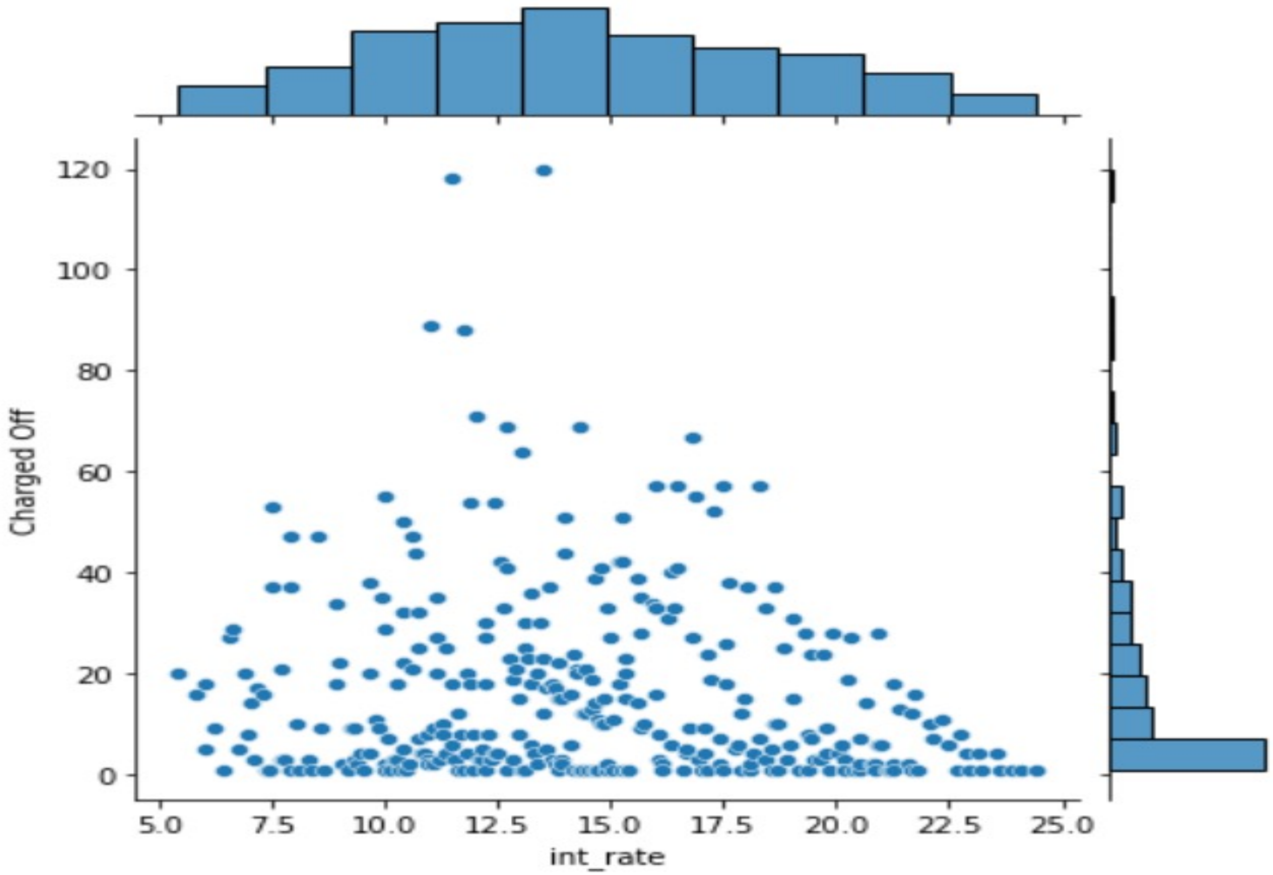
# Conclusion – 4

- Customers with previous bankruptcy records are more likely to default



# Conclusion – 5

- The defaulters are more for interest rate of around 15%



# THANK YOU