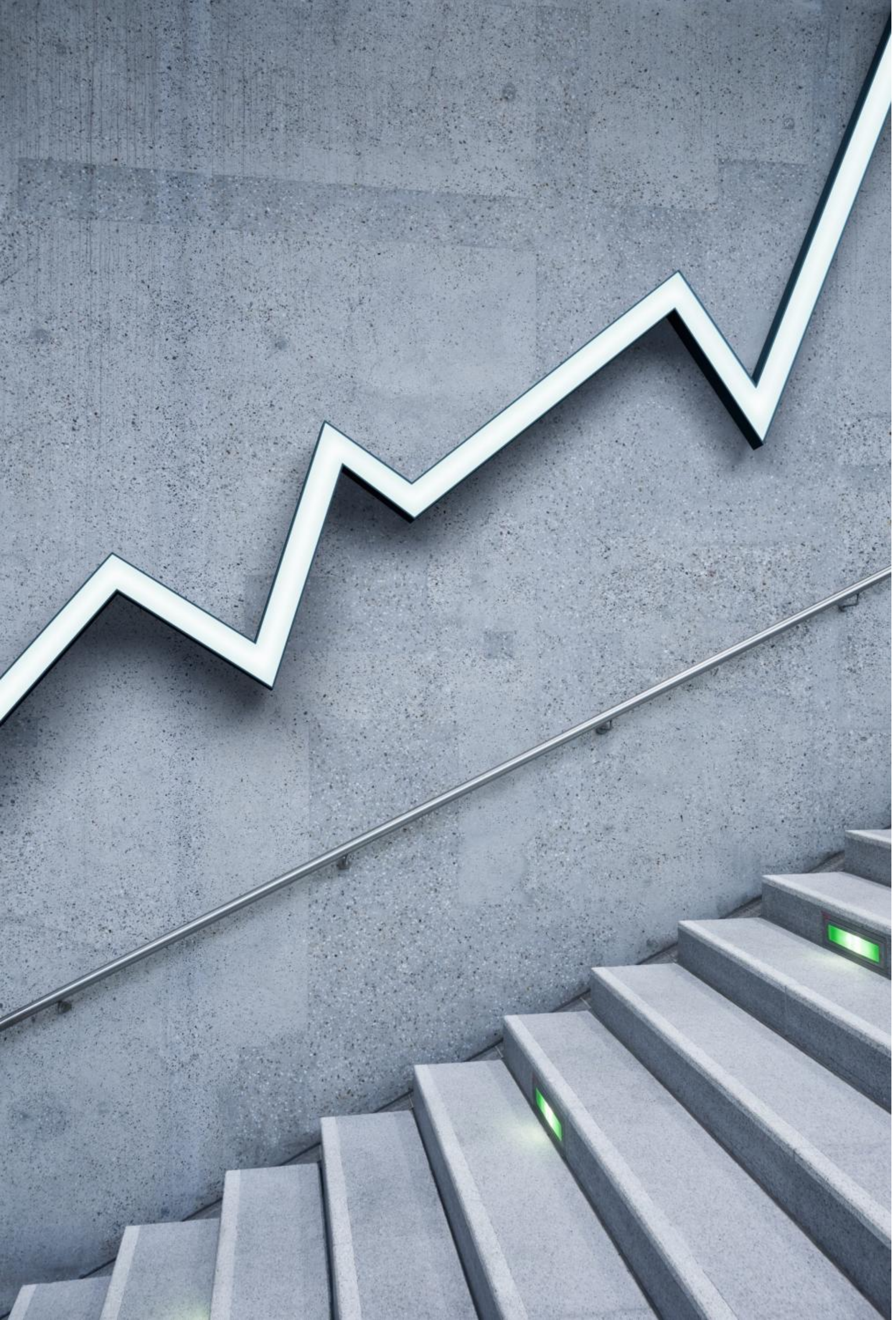


Lending Club Case Study

TEAM

- VIJAY BHASKAR

- VIDHYADHAR PANDIT

Table of Contents

Problem statement

Objective of case study is to understand the driving factors (or driver variables) behind loan default i.e. the variables which are strong indicators of default

Assumptions

- The analysis is done only for loan status that has values Fully Paid and Charged Off. The data that has loan status value "current" is removed from the dataset

Approach

- Step 1: Data Understanding
  - Total number of Rows - 39717
  - Total number of Columns – 111
  - Conclusion: There are many columns with null values for all the rows. Hence we need data cleaning

Approach

- ❖ Step 2: Data Cleaning – Strategy
  - ❖ Columns having high percentage of missing values, directly discard those column from the analysis
  - ❖ Columns having missing values under acceptable range we prefer to impute them(replace with some values)
  - ❖ For numerical columns we prefer to use mean or median (Median is preferred)
  - ❖ Median is not affected with the outliers
  - ❖ For categorical column we refer to use mode
  - ❖ If we have very less missing values in any column in that case you can drop rows as well
  - ❖ For a target variable - if there are missing values, drop the row
  - ❖ Removing columns which has all values as null
  - ❖ Drop columns where same value is present in all or most of the rows

Approach

- Step 3: Data Analysis
  - : For performing the analysis, the percentage of charged off loans in the overall dataset is taken as the baseline and data in various segments is compared against this baseline to find out the strong indicators of default
  - Identify Categorical and Numeric variables
  - Univariate Analysis
    - Categorical columns
    - Numeric Columns
  - Bivariate Analysis
    - Categorical columns
    - Numeric Columns

Insights

RECOMMENDATION AND CONCLUSION

Relationship between loan amount and charged off loans

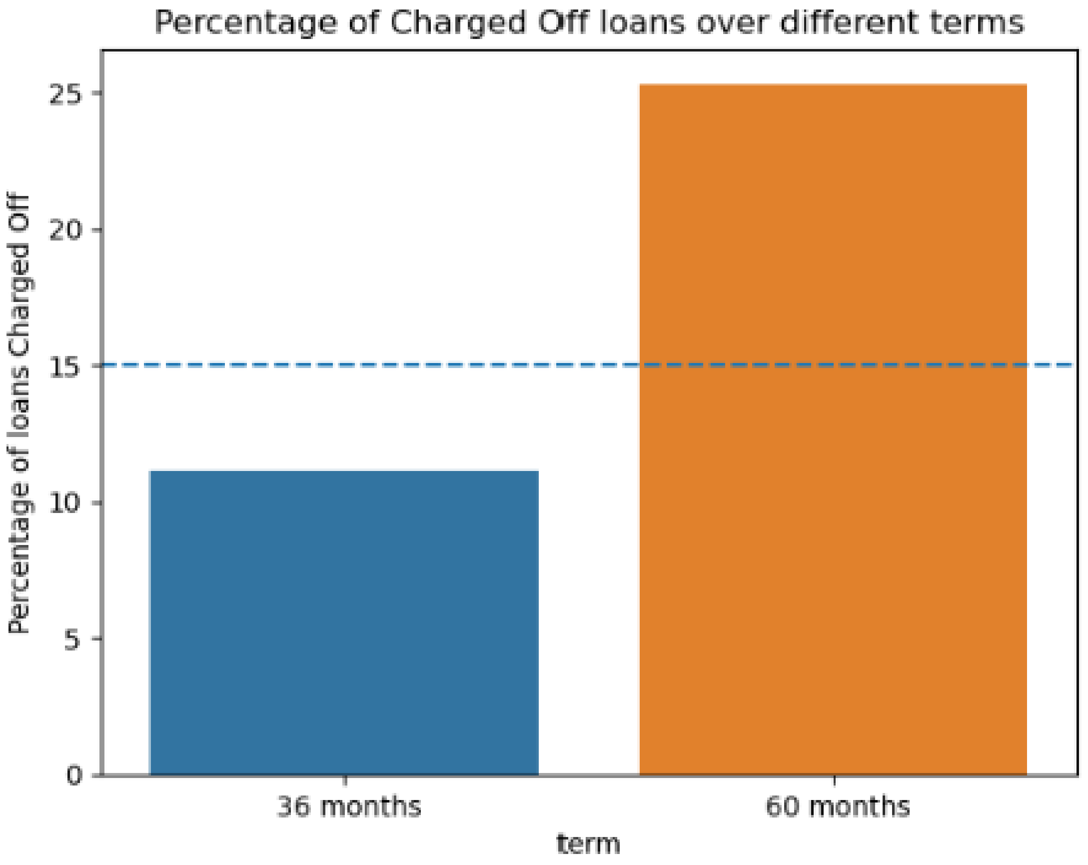
- The analysis is done for the data beyond 90th percentile of loan amount.
- In this dataset, it is observed that the percentage of charged off loans increased by 5 percentage points. While the percentage of charged off loans in the overall dataset is 15%, the same in higher loan amount dataset is 20%
- **Observation:**
  - Higher the loan amount, higher are the chances for the loan getting charged off

Relationship between interest rate and charged off loans

- ❖When the analysis is done for the dataset beyond 90th percentile of the interest rate, it is observed that the percentage of the charged off loans in the total loans in this dataset is 30%. This is double when compared to the percentage of charged off loans in the overall dataset(15%) indicating that at higher interest rate there is high chance of loan getting charged off.
- ❖Similarly, when the analysis is done for the data below the 10th percentile of interest rate, it is observed that the percentage of the charged off loans in the total loans in this dataset is 4%. This is very less when compared to the percentage of charged off loans in the overall dataset(15%) indicating that, at lower interest rate there is high chance of loan getting fully paid.
- ❖Observation:
  - ❖Higher the interest rate, the higher are the chances for the loan getting chargedoff

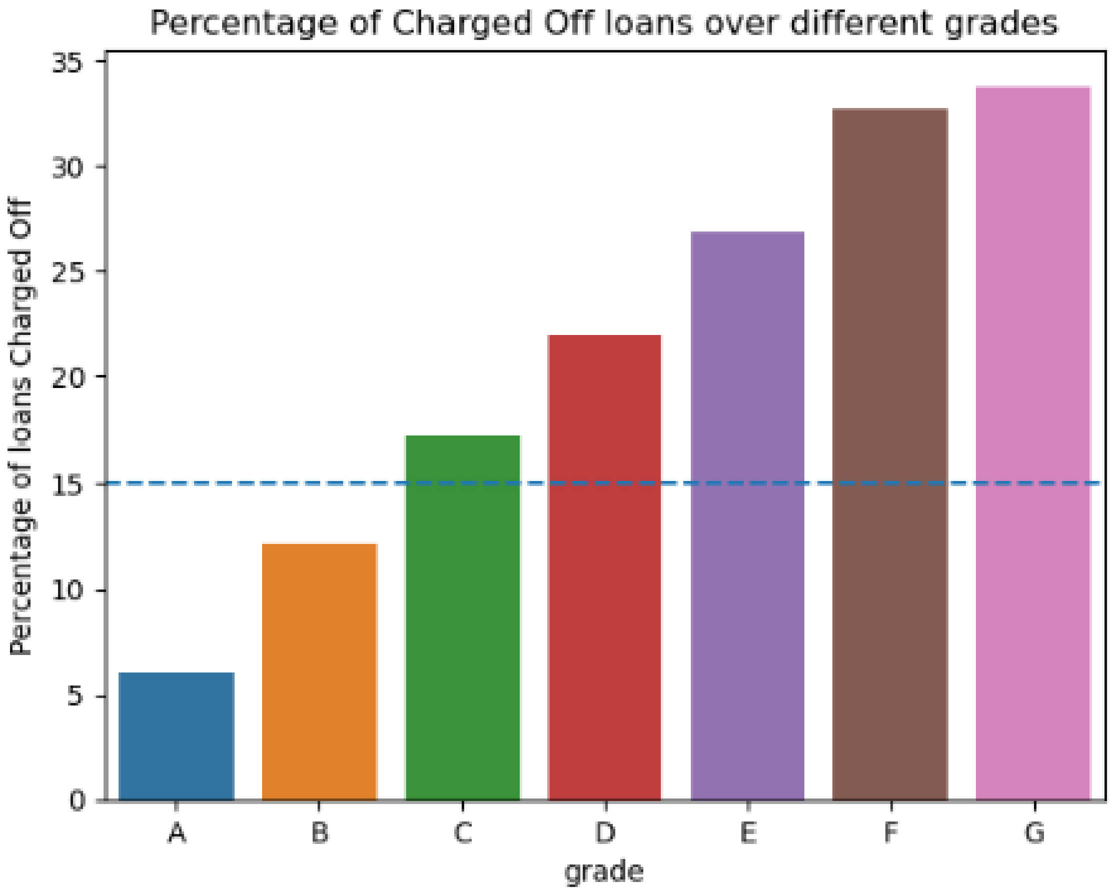
Loan term and Charged Off loans

- ❖When the loan term is 60 months, 25% of the loans in this dataset got charged off. This is higher than the overall average of 15%
- ❖When the loan term is 36 months, only 11% of the loans in this dataset got charged off. This is lower than the overall average of 15%
- ❖Observation:
  - ❖Lending for longer term has high chances for loan getting charged off



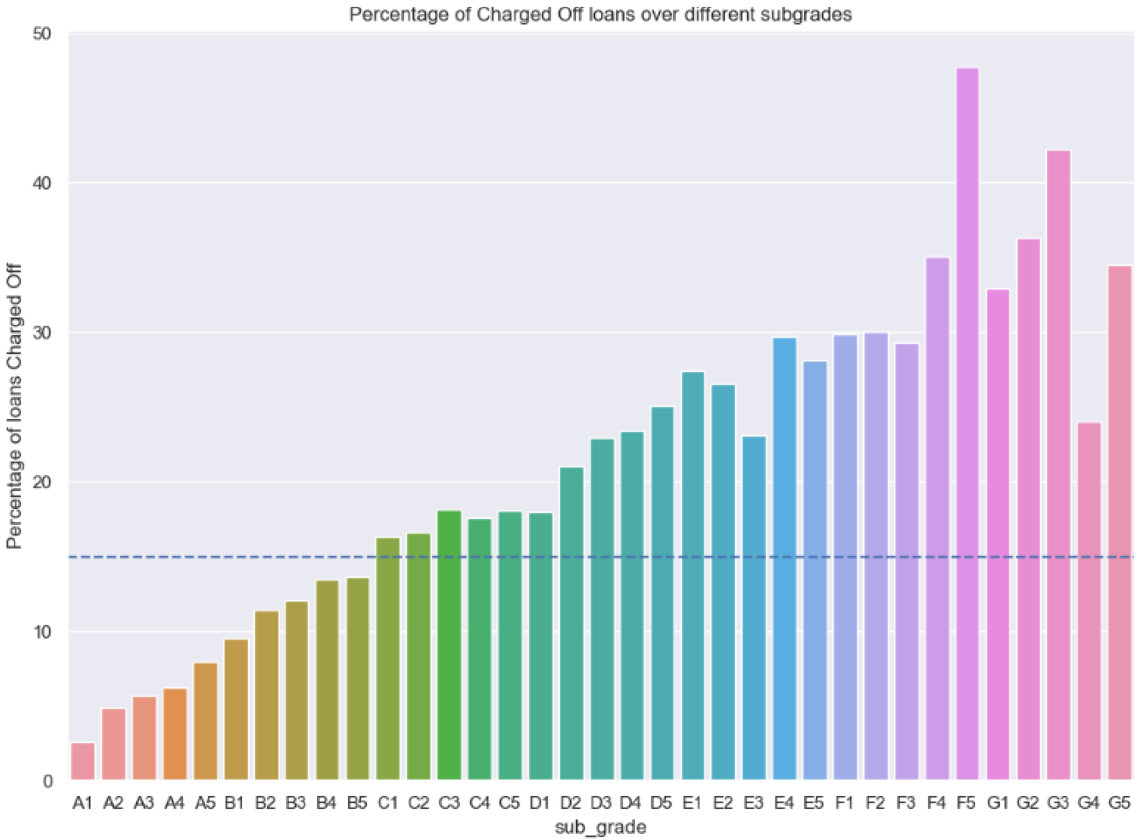
Loan Grade and Charged Off loans

- ❖As the grade of loan moves from A to G, the percentage of charged off loans keeps increasing. While only 6% of the grade A loans ended up in charged off, in grade G over 33% of the loans are charged off
- ❖**Observation:**
  - ❖The higher the grade of the loan(E,F,G), the higher are the chances of loan getting charged off. It is advised to give loans with grades (A,B,C) to minimize the credit loss.



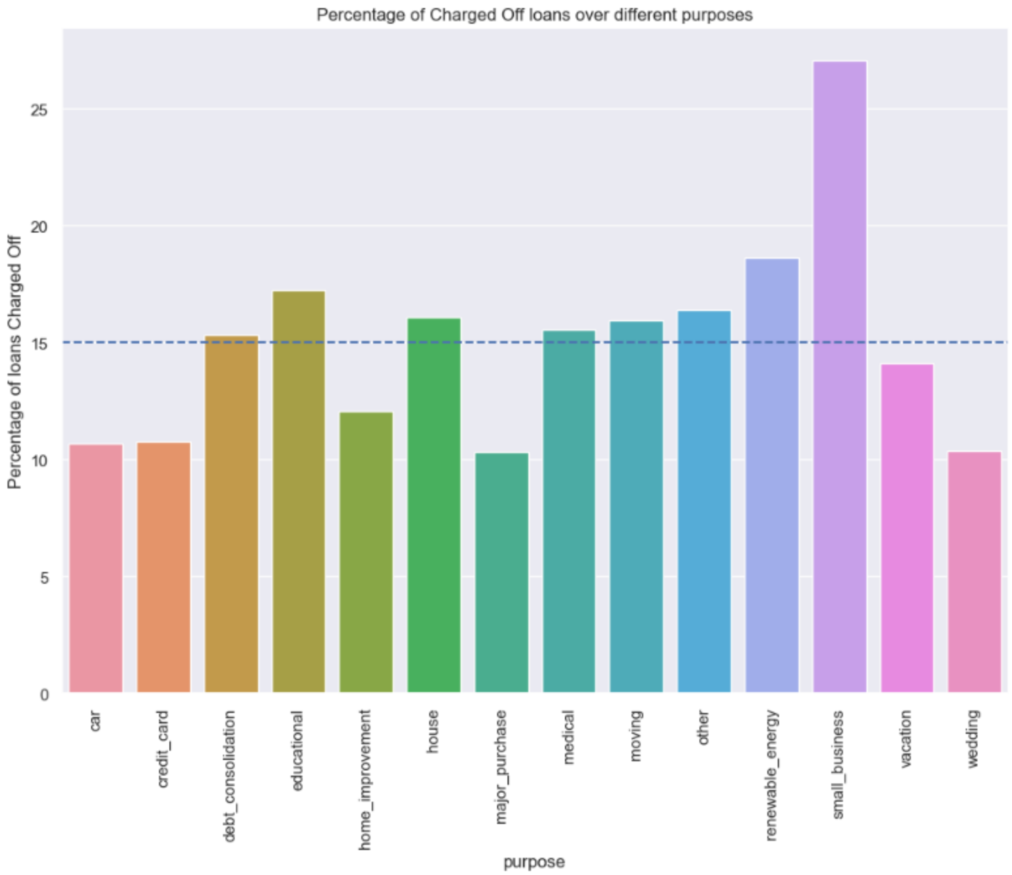
Loan sub grade and Charged Off loans

- In the F5 subgrade, almost 48% of the loans resulted in charged off loans. In these subgrades F5,G3,G2,F4,G5,G1 the charged off loans percentage is very high(more than 33%. At least One in every three loans got charged off)
- **Observation:**
  - The loans in subgrades of E,F,G have higher chances of default



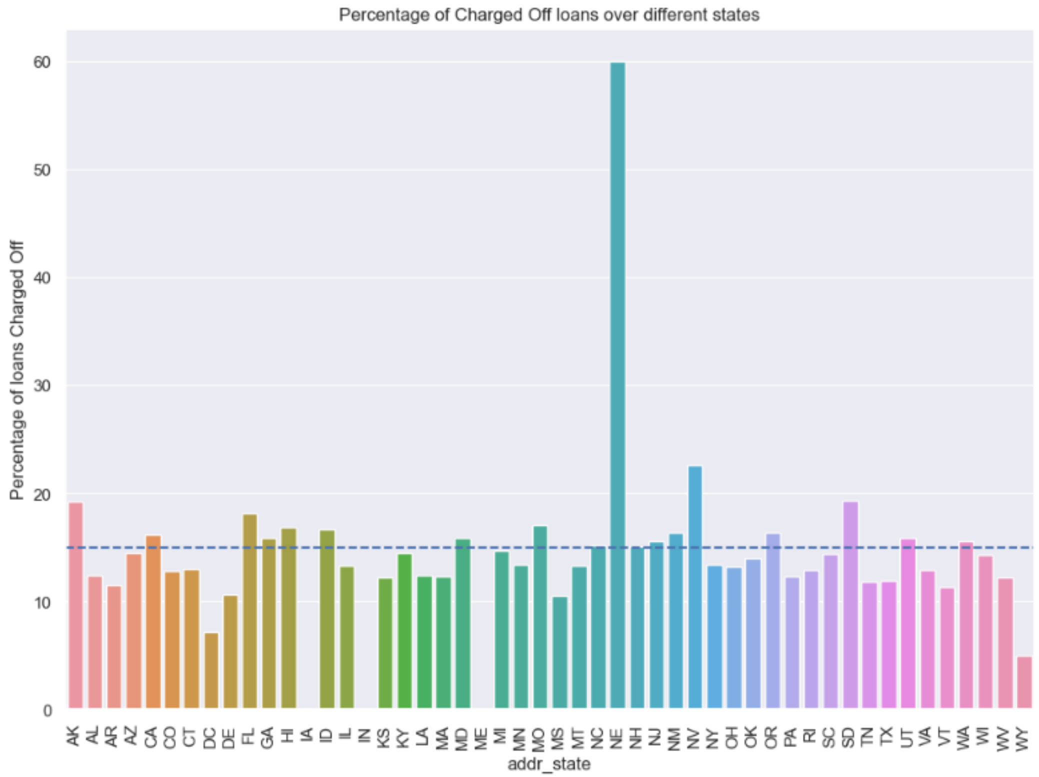
Purpose of the loan and Charged Off loans

- Lending for the purpose of small business turned out to be risky. More than 27% of the total loans resulted in charged off, this is way more than the average of 16%
- **Observation:**
  - Lending to small businesses is risky when compared to loans for other purposes



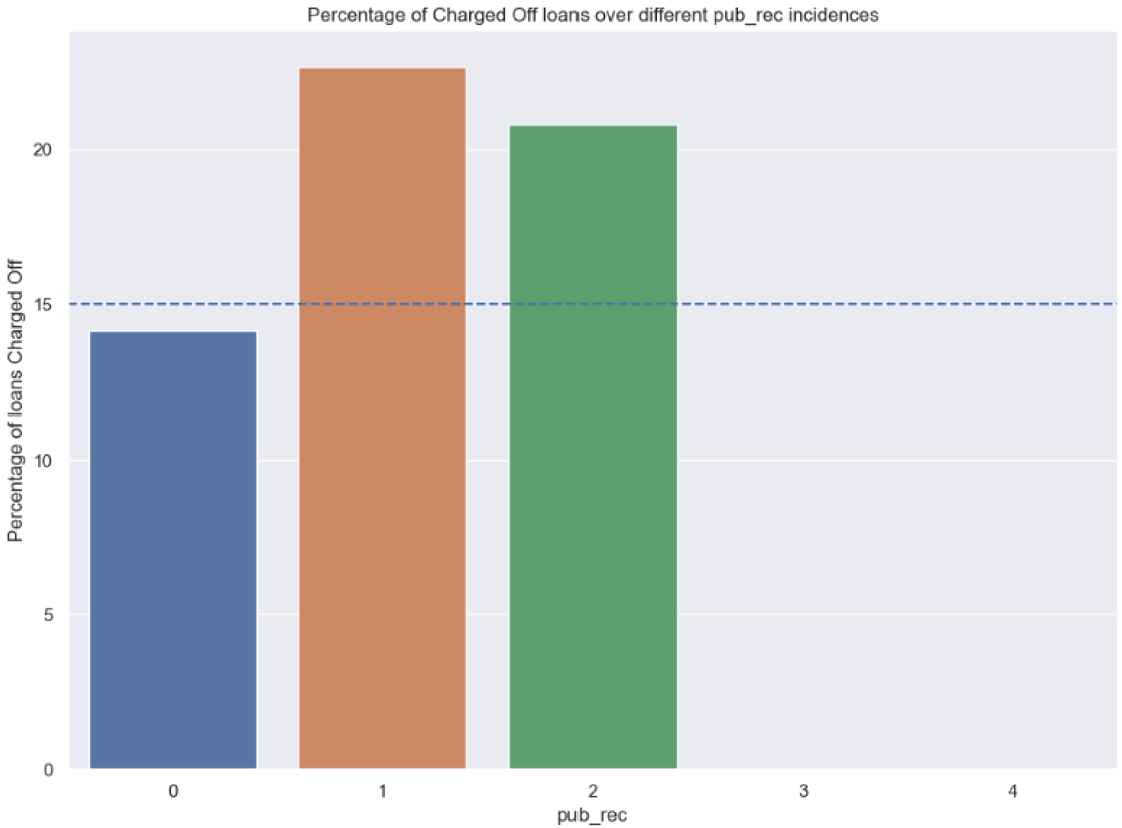
Charged Off loans across states

- In the state of DC, the percentage of charged off loans is only 7% when compared to national average of 15%
- In the state of NV, it is 22.5%,this is higher than the national average of 15%.
- In the state of NE, it is 60% way higher than the national average.
- **Observation:**
  - For the loans in the states of NE and NV,the chances of loan getting charged off is higher.



Charged Off loans when there are derogatory public records

- When there are derogatory public records against a loan applicant, there is a higher chance of loan getting defaulted



Charged Off loans when there are publicly recorded bankruptcies

- When there are publicly recorded bankruptcies against a loan applicant, there is a higher chance of loan getting defaulted