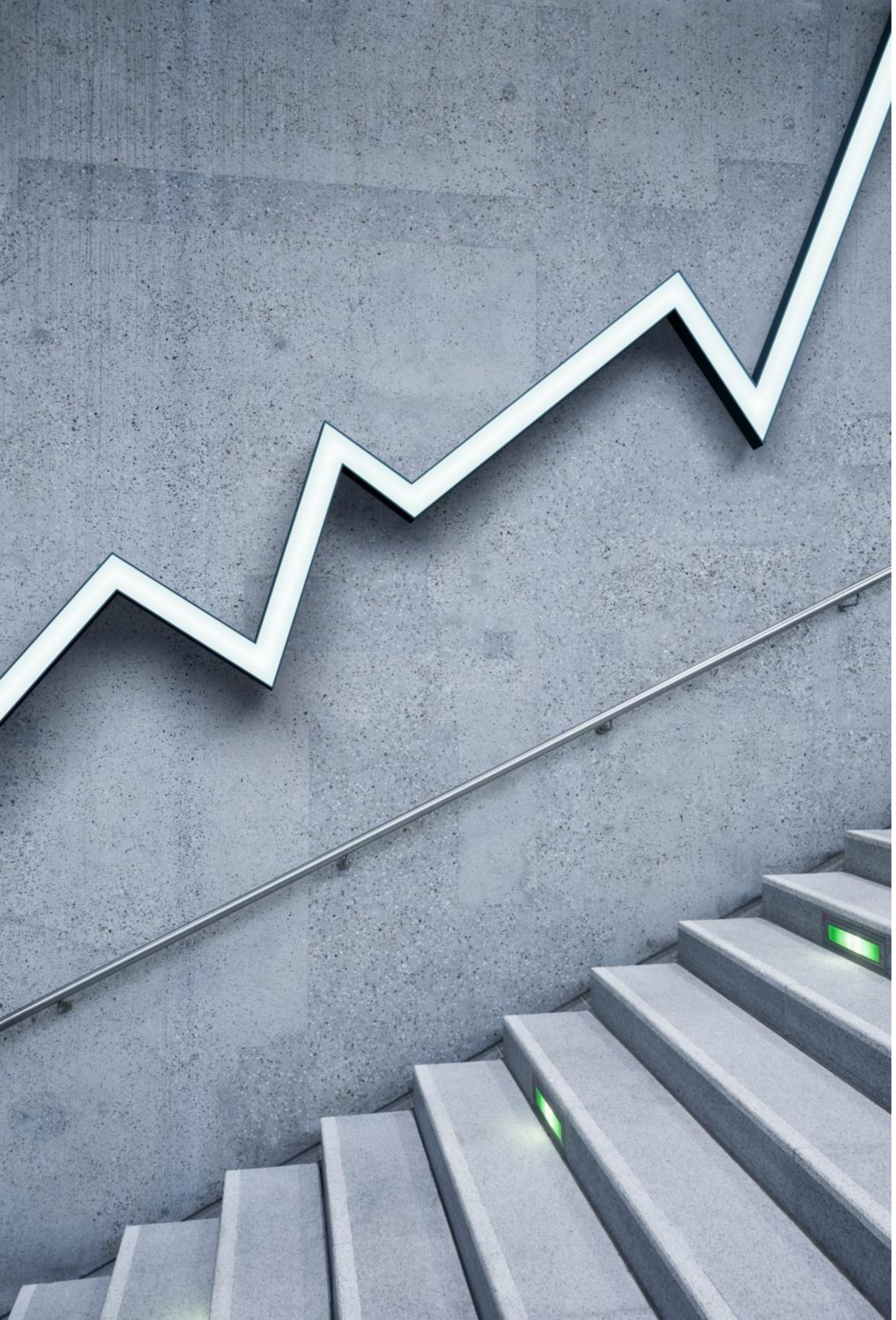


Lending Club Case Study

TEAM

- VIJAY BHASKAR
- VIDHYADHAR PANDIT

TOC

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

<need to add>

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
  - Conclusion
    - <need to add>

Approach

- Step 3: Data Analysis
  - Identify Categorical and Numeric variables
  - Univariate Analysis
    - Categorical columns
    - Numeric Columns
  - Bivariate Analysis
    - Categorical columns
    - Numeric Columns
  - Conclusion
    - <need to add>

Insights

RECOMMENDATION AND CONCLUSION

Insights

- Insight1: At higher loan amounts the percent of charged off loans over total loans increased by 5 percentage points i.e. one in every five loans charged off
- Insight2: At higher interest rates, charged off loans as a percentage of total loans doubled when compared to overall range of data. Three loans in every ten loans charged off, meaning there is correlation between higher interest rate and loan getting charged off
- Insight3: When the interest rate is very low, the percent of charged off loans is way to less when compared to the overall range. This means when the interest rate is very low, very few loans gets charged off
- Insight4: When the loan term is 60 months, 25% of the total loans got charged off
- Insight5: 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
- Insight6: 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)
- Insight7: 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%
- Insight8: The percentage of charged off loans is around 15% in most cases except DC-7% and NV-22.5%
- Insight9: When there is one pubrec, the percent of charged off loans is high when compared to average. The other sample sizes are very small, so not considering them.
- Insight10: When there is one pub_rec, bankruptcy, the percent of charged off loans is high when compared to average. The other sample sizes are very small, so not considering them

Insight 1

- At higher loan amounts the percent of charged off loans over total loans increased by 5 percentage points i.e. one in every five loans charged off

Insight 2

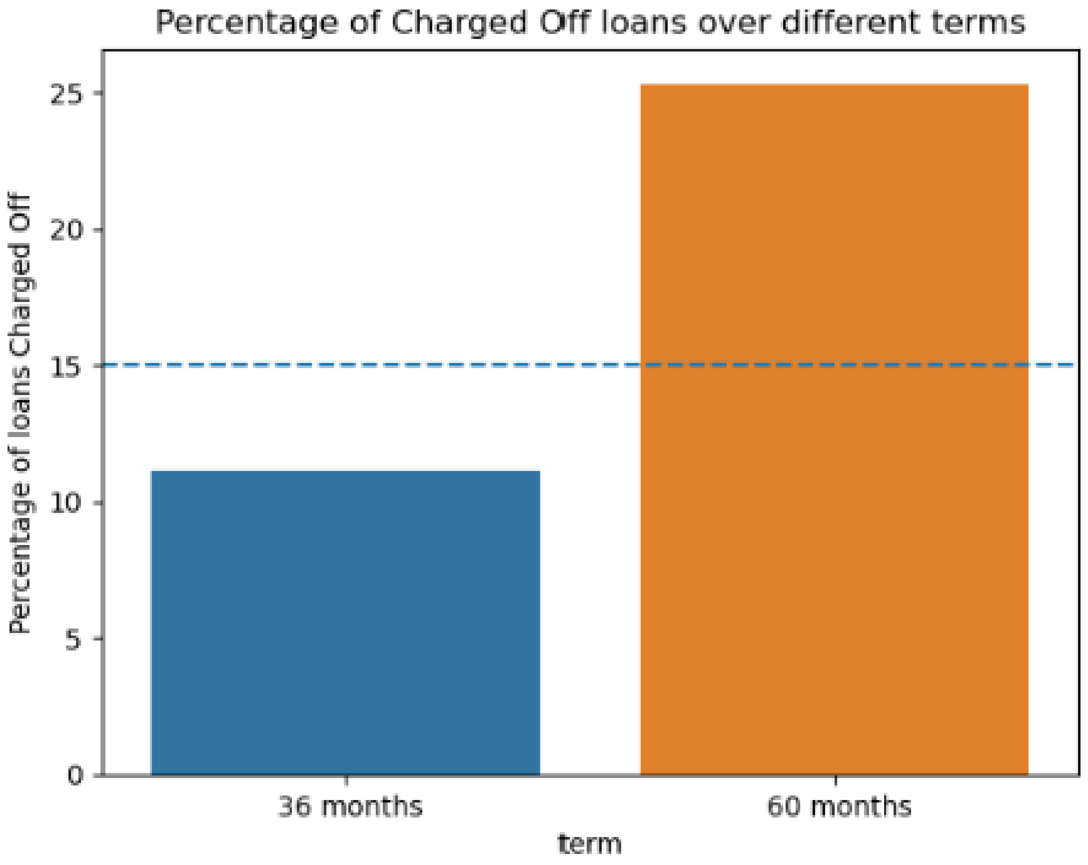
- At higher interest rates, charged off loans as a percentage of total loans doubled when compared to overall range of data. Three loans in every ten loans charged off, meaning there is correlation between higher interest rate and loan getting charged off

Insight 3

- When the interest rate is very low, the percent of charged off loans is way to less when compared to the overall range. This means when the interest rate is very low, very few loans gets charged off

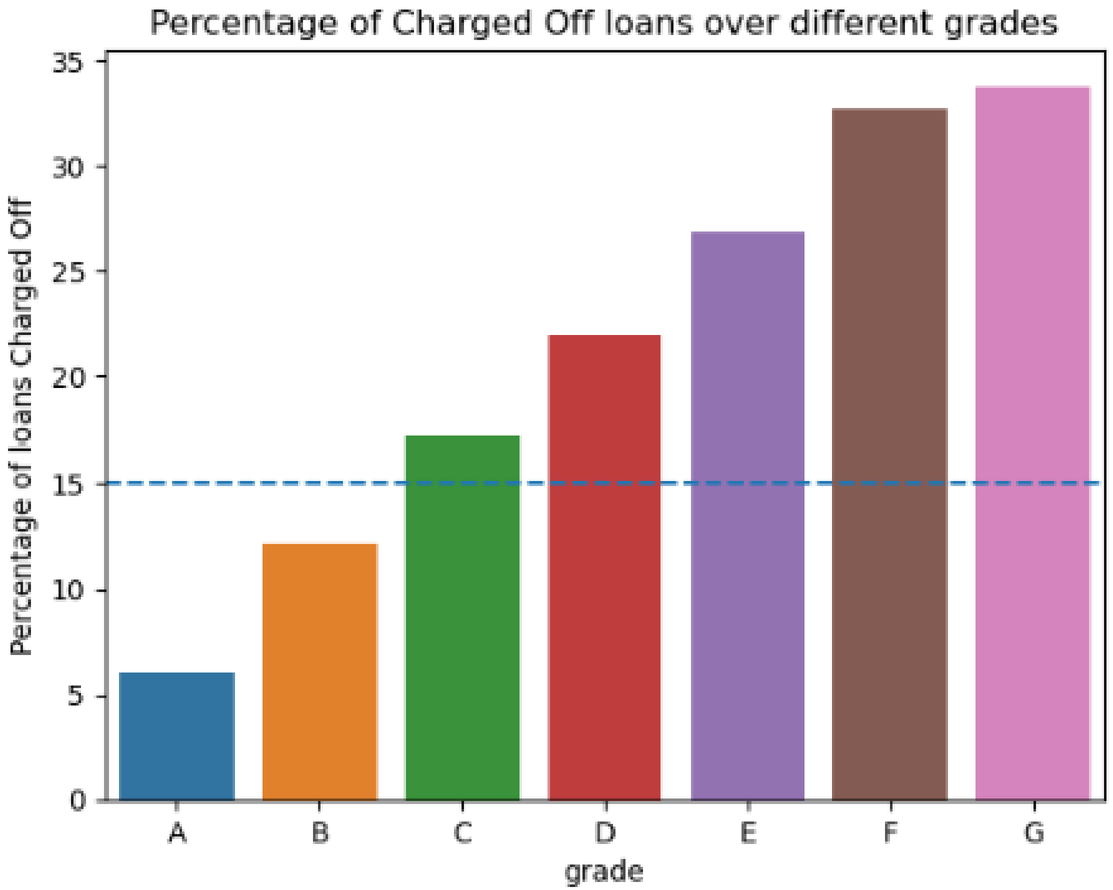
Insight 4

- When the loan term is 60 months, 25% of the total loans got charged off



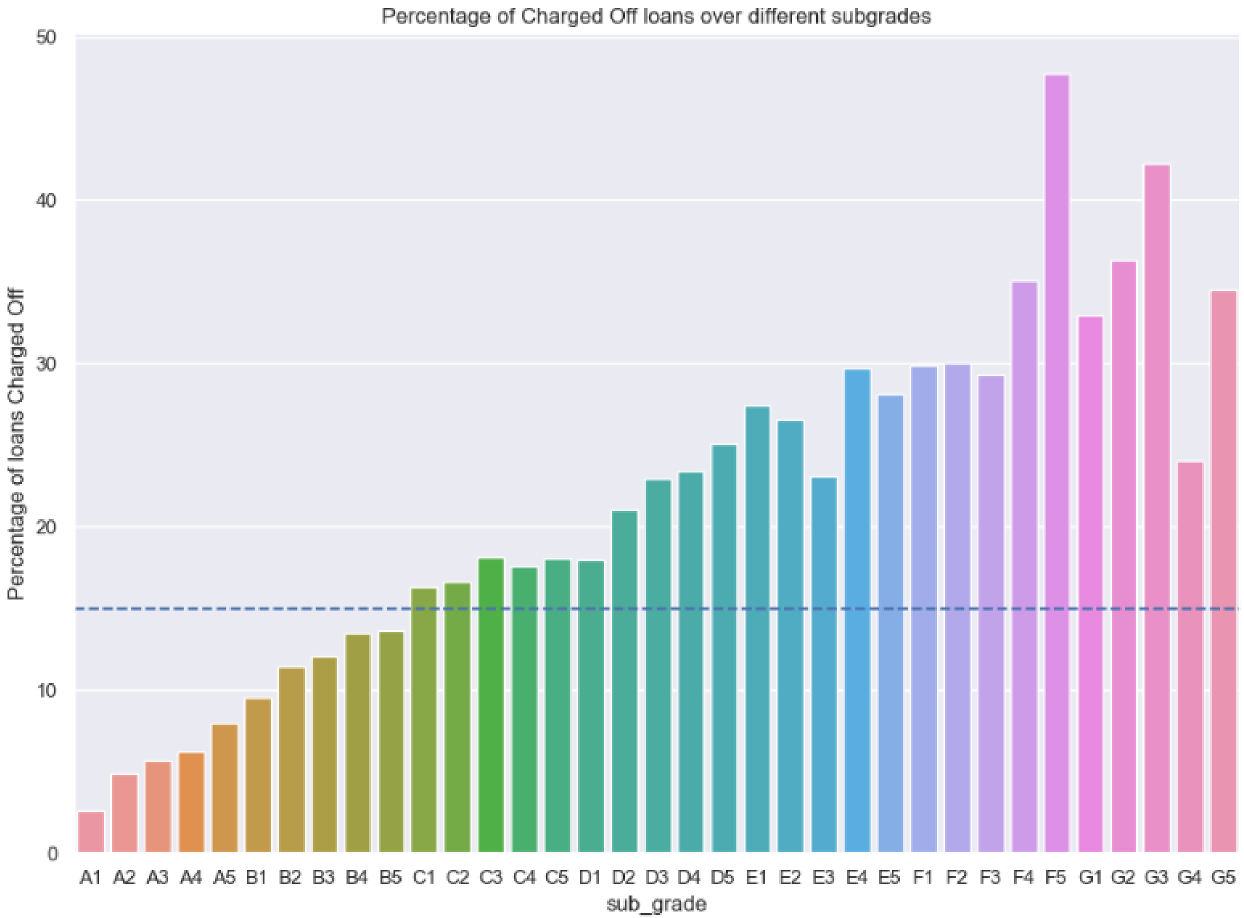
Insight 5

- 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



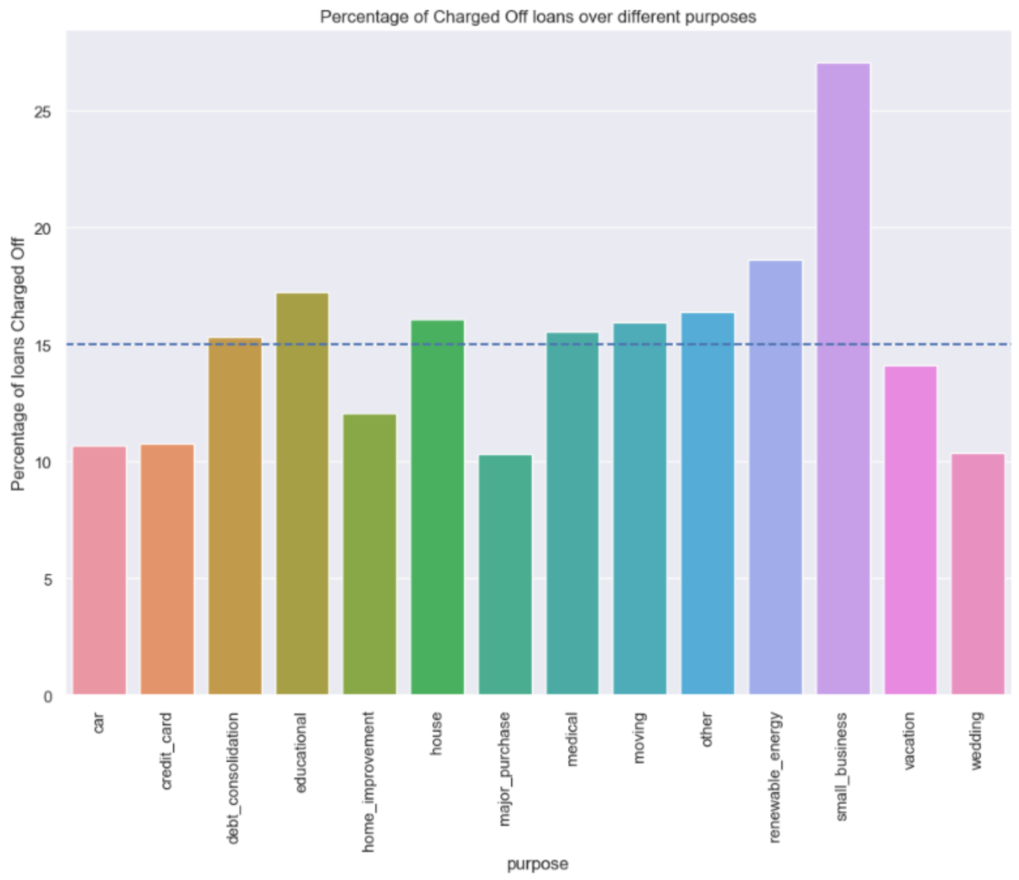
Insight 6

❖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)



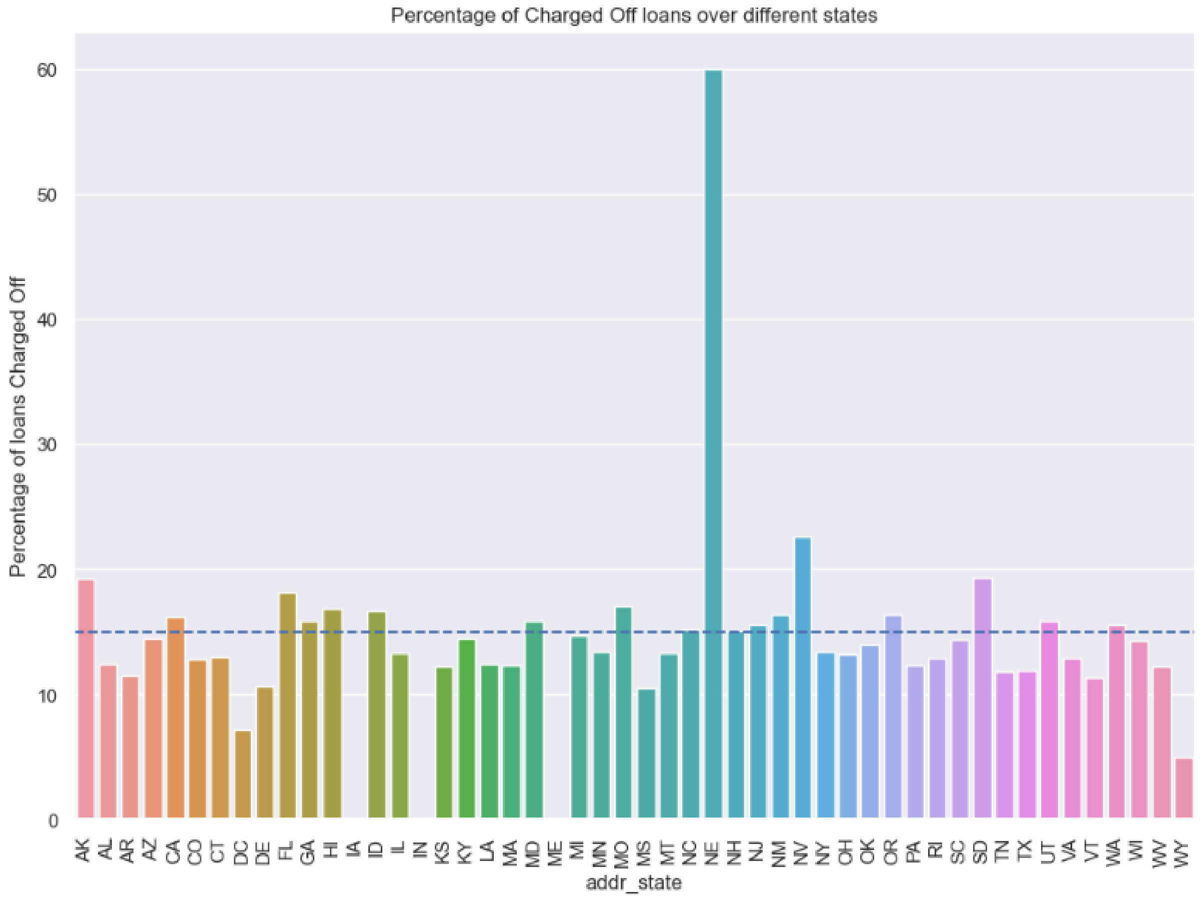
Insight 7

❖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%



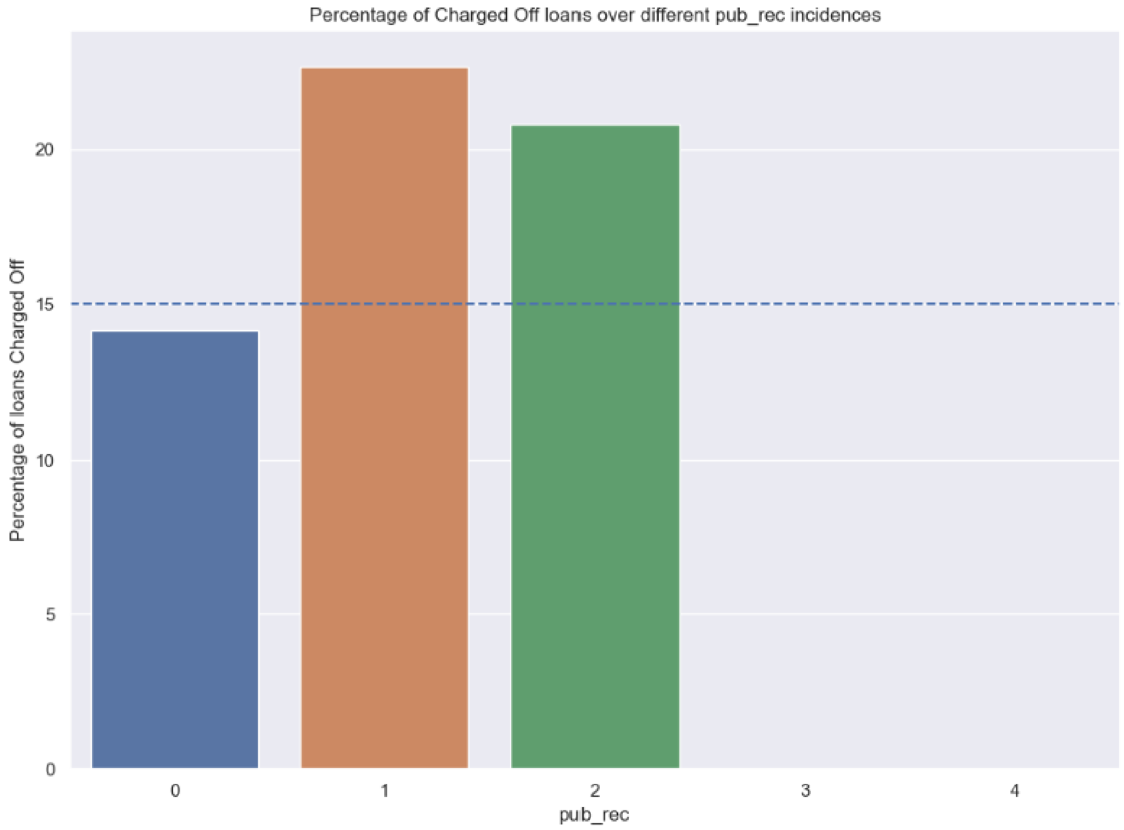
Insight 8

- The percentage of charged off loans is around 15% in most cases except DC-7% and NV-22.5%



Insight 9

- When there is one pubrec, the percent of charged off loans is high when compared to average



Insight 10

❖ When there is one pub_rec, bankruptcy, the percent of charged off loans is high when compared to average