



GRAMENER LOAN CASE STUDY USING EDA

SUBMITTED BY -

GOVIND SONI





P R O B L E M D E S C R I P T I O N

PROBLEM:-

THERE IS A COMPANY WHO USED TO LEND MONEY TO PEOPLE. THE PROBLEM IS THAT THEY WILL LEND MONEY BUT PEOPLE DON'T PAY THEIR MONEY BACK IN SOME CASES. OUR TASK IS TO FIND A SOLUTION TO MINIMIZE THE RISK OF LOSING MONEY.



● DATA INSPECTION PART

- THERE ARE TOTAL 39717 ROWS AND 111 COLUMNS IN THIS DATASET.
- THE DATA TYPES OF COLUMNS ARE AS FOLLOWS:-
 1. FLOAT64(74 COLUMNS)
 2. INT64(13 COLUMNS)
 3. OBJECT(24 COLUMNS)
- THESE COLUMNS START FROM ID TO TOTAL_IL_HIGH_CREDIT_LIMIT





DATA MANIPULATION

- THE COLUMNS WITH MORE THAN 50% NULL VALUES ARE REMOVED.(57 COLUMNS)
- AFTER REMOVING THEM, ONLY 54 COLUMNS LEFT.
- AGAIN CHECKED FOR MISSING VALUES AND IT IS FOUND THAT A COLUMN NAMED DESC HAVE 33% MISSING VALUE.
- SO THIS COLUMN IS ALSO REMOVED AND WE LEFT WITH 53 COLUMNS FOR ANALYSIS.
- THESE COLUMNS ARE ID TO TAX_LIENS.
- INT_RATE AND EMP_LENGTH COLUMN ARE STANDARDISE AND BY REMOVING % AND <>+ SIGNS RESPECTIVELY.
- MONTH AND YEAR IS EXTRACTED FROM ISSUE DATE COLUMN.





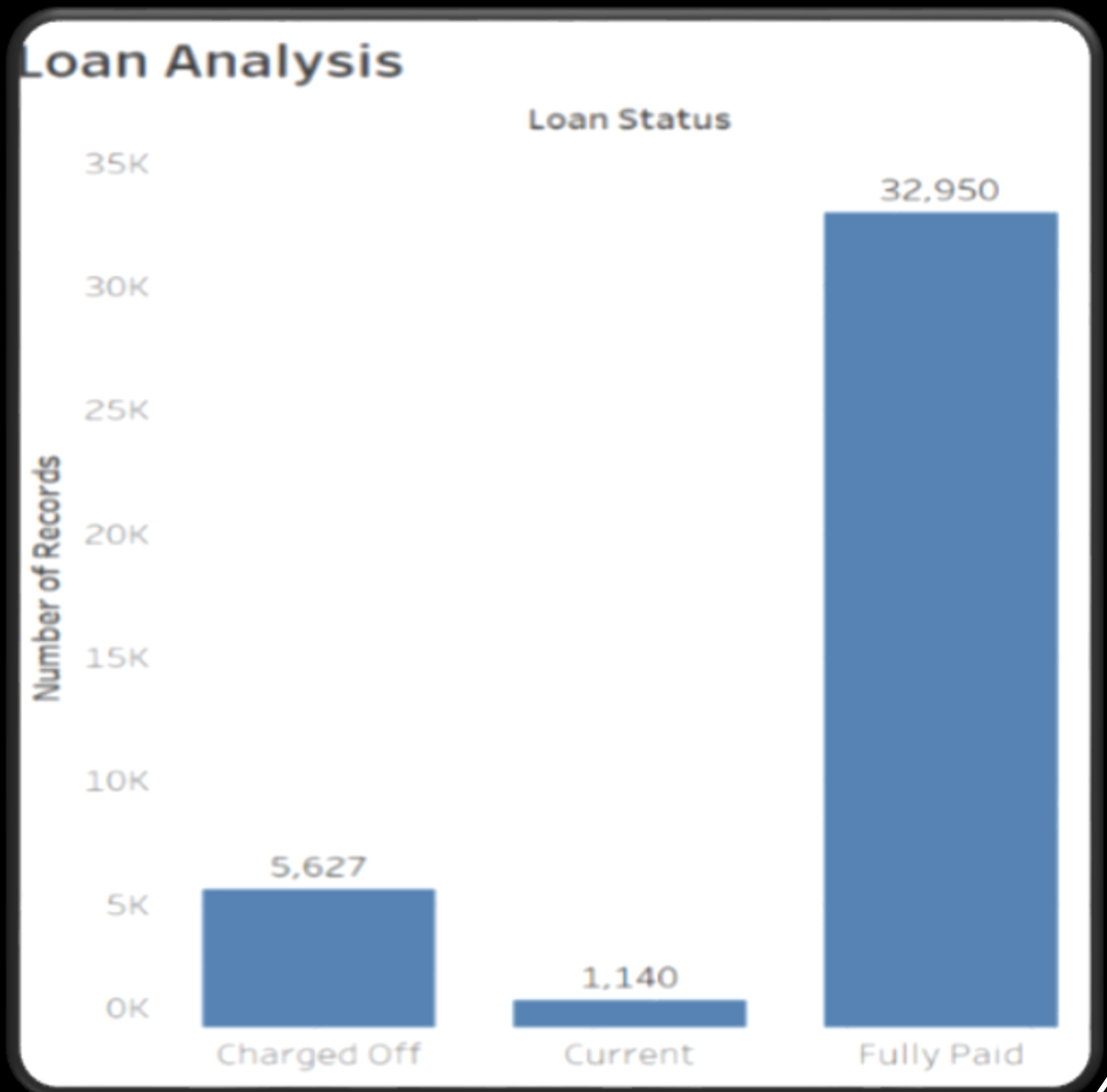
DATA ANALYSIS

- THIS DATA IS ANALYSED BY EDA PROCESS.
- IN THIS PROCESS, DATA IS ANALYSED IN THREE TYPES:-
 1. UNIVARIATE.
 2. SEGMENTED UNIVARIATE.
 3. BIVARIATE.
- IN THIS ANALYSIS, WE SPECIALLY PLOT LOAN_STATUS COLUMN AGAINST VARIOUS COLUMNS AND GET INSIGHTS.
- IN THIS COLUMN, WE SPECIALLY FOCUS ON CHARGED-OFF ENTRIES TO FULFILL OUR ANALYSIS' MOTO.
- HERE FEW PLOTS ARE ATTACHED.



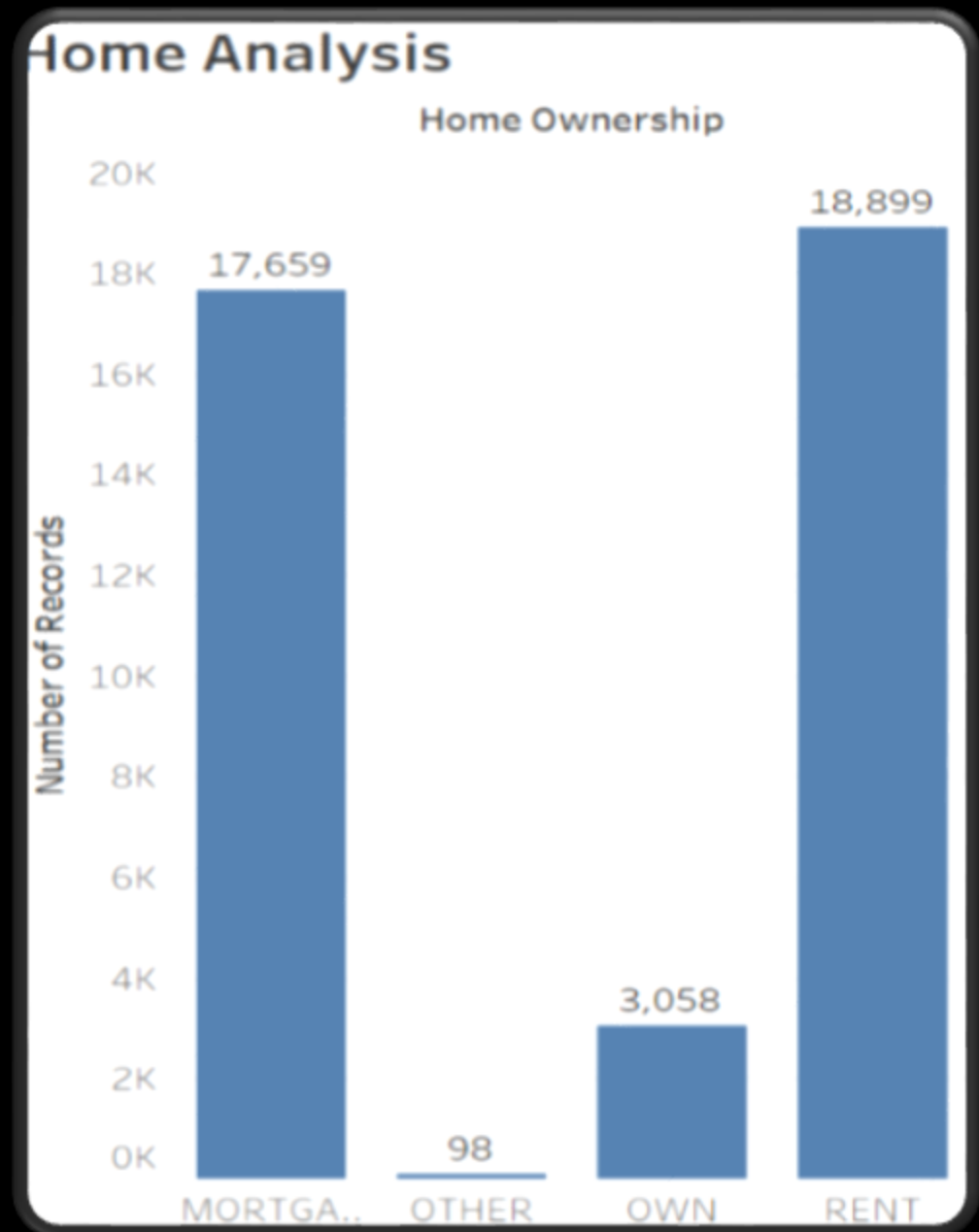
● LOAN ANALYSIS

- There are total 39717 loans were granted out of which:-
 1. 32950 are fully paid.
 2. 1140 are current.
 3. 5627 are charged off.



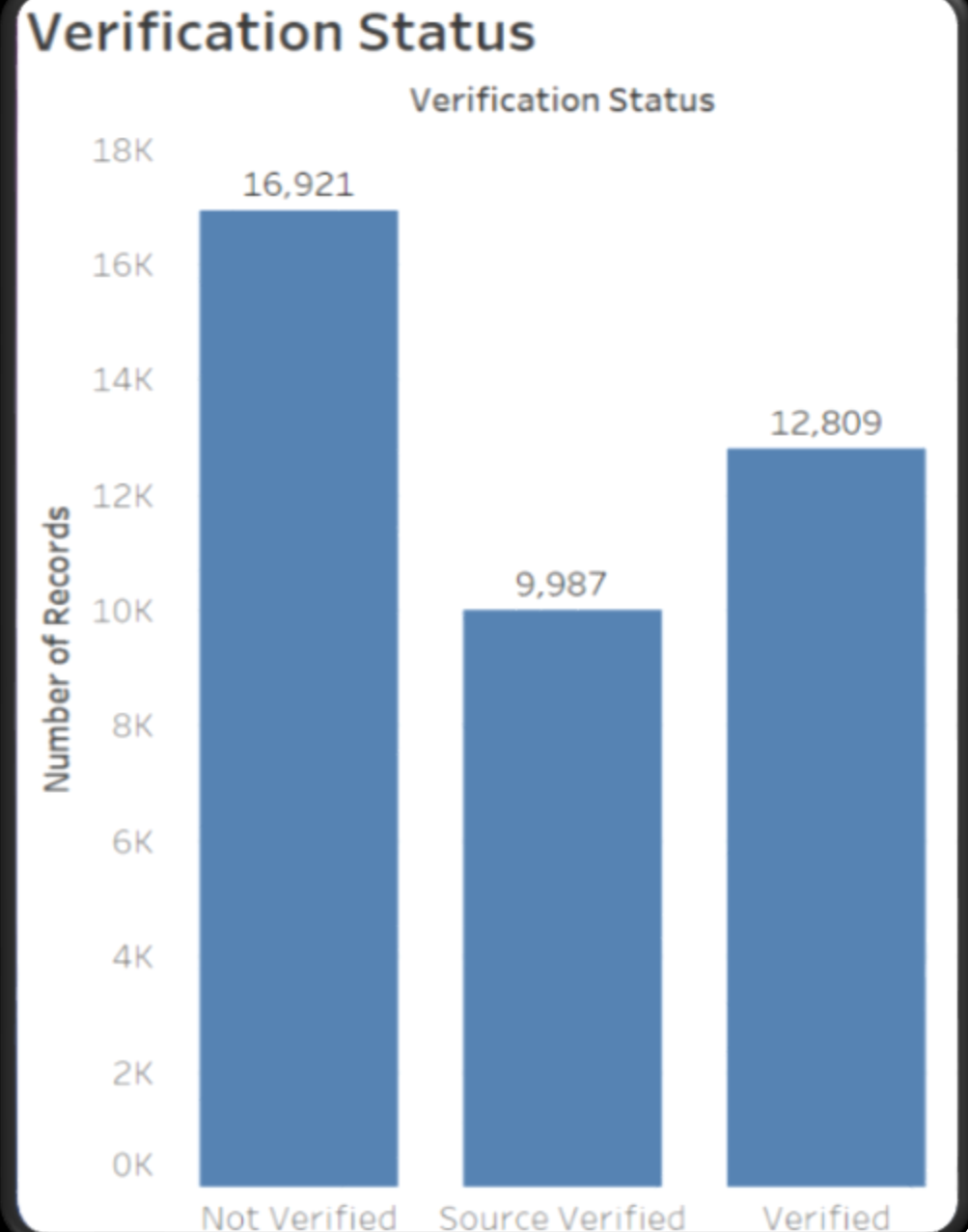
● CHARGED OFF- HOME OWNESHIP

- It is cleared from graph that most of the charged off people have rented and mortgage houses. Their values are as follows:-
 1. 18899 have rented houses.
 2. 17659 have mortgage houses.
 3. 3058 have their own houses.
 4. 98 have other.



● CHARGED OFF- VERIFICATION STATUS

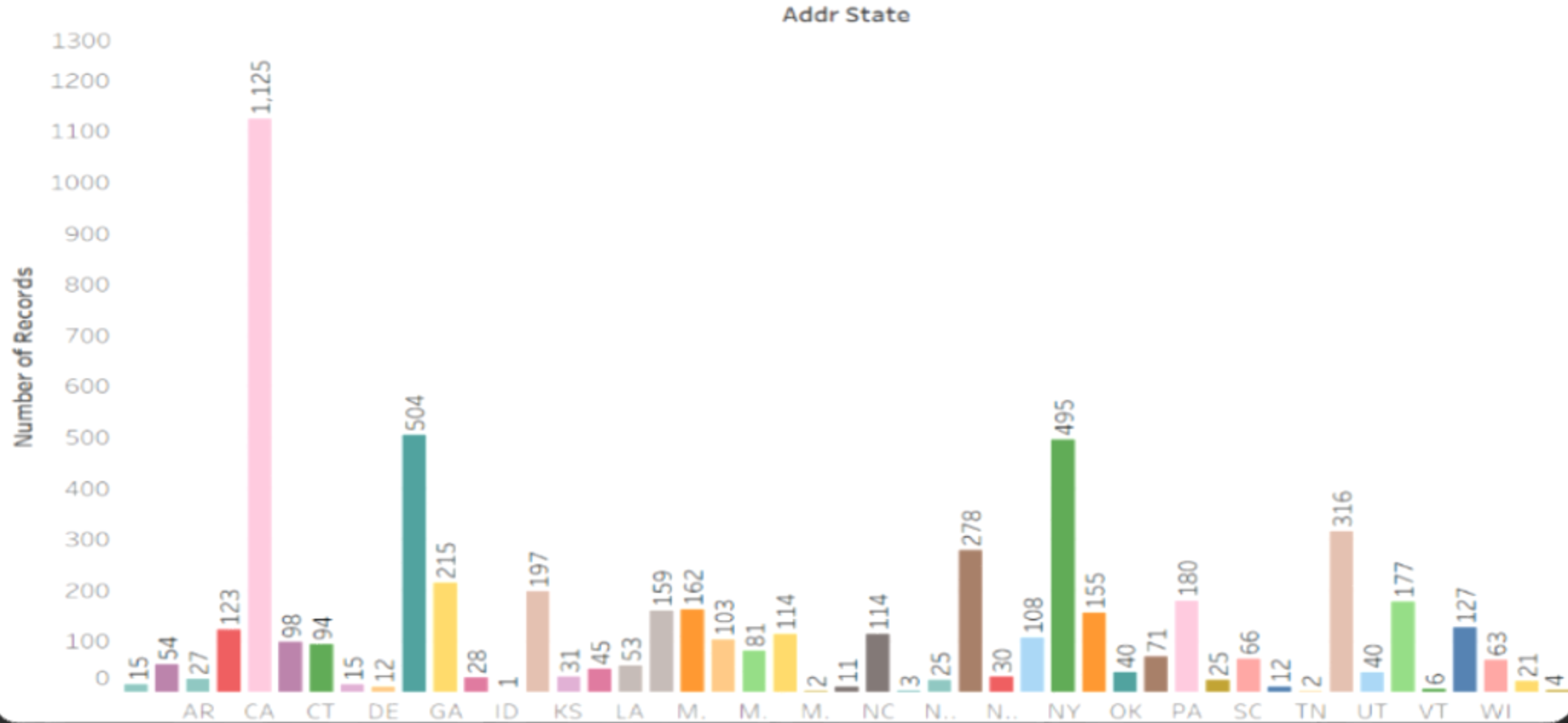
- This shows their verification status:-
 1. 16921 were not verified.
 2. 12809 were verified.
 3. 9987 source verified.
- 2124 members were granted for loan without verifying their income source.



CHARGED OFF - STATE

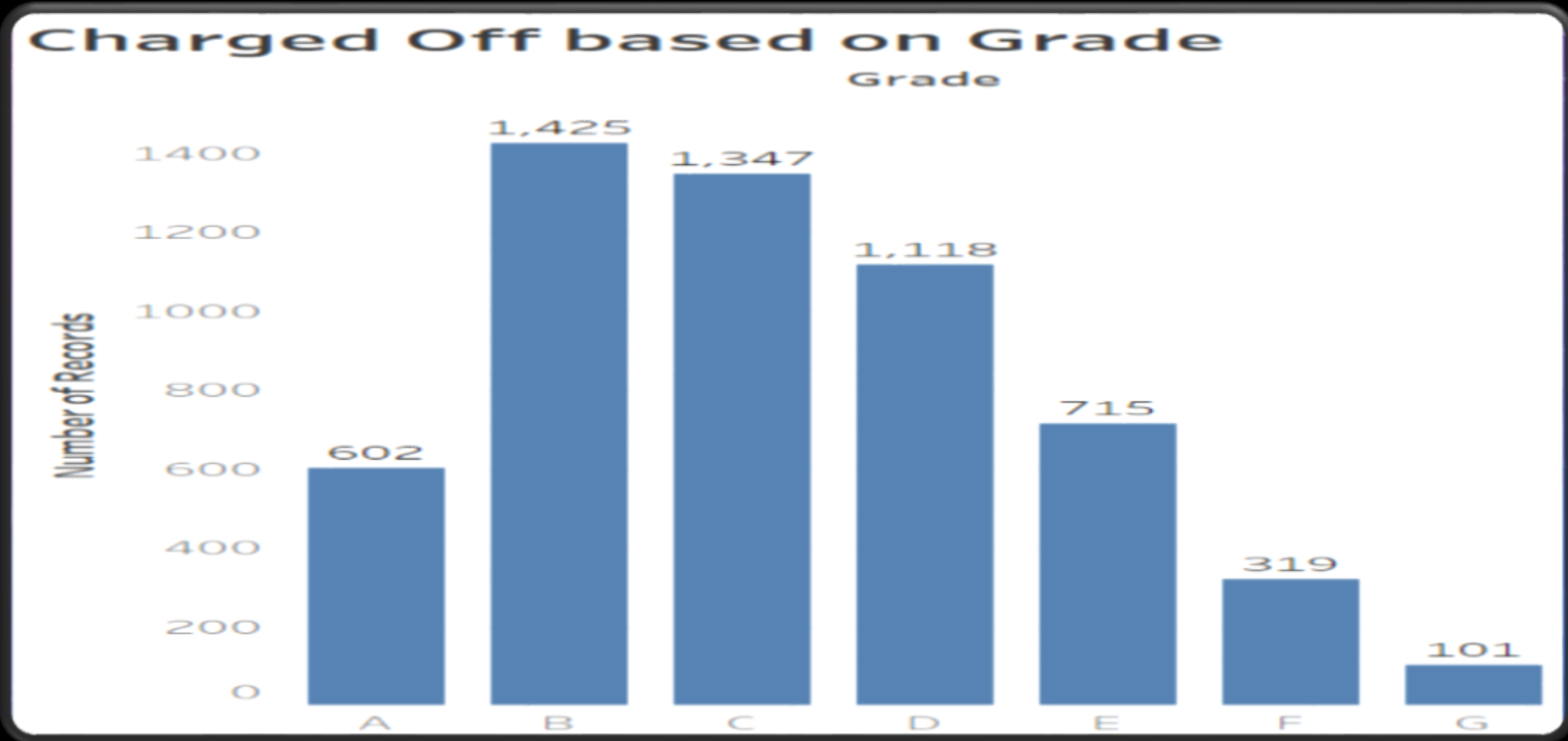
California State has the most cases of charged off

States Ratio with Charged Off Data



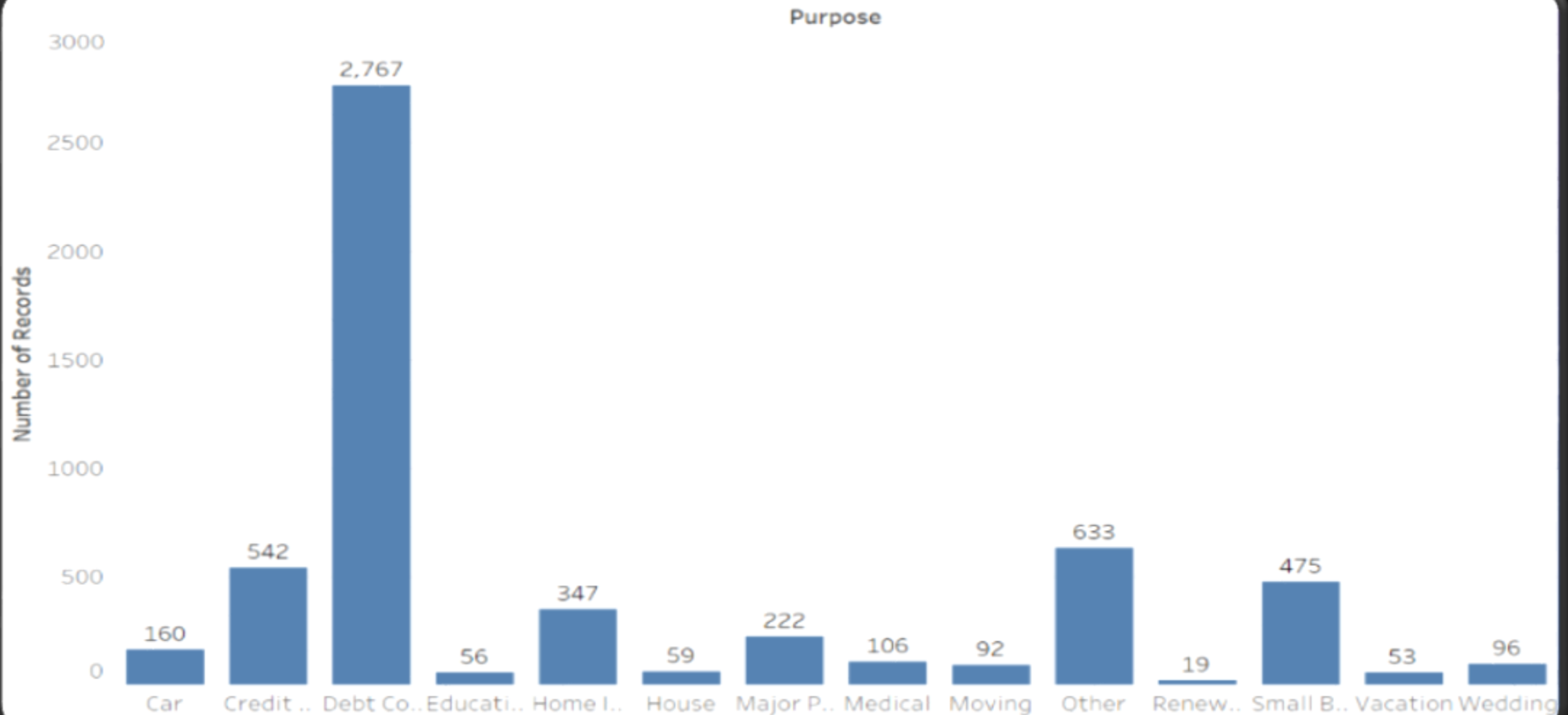
CHARGED OFF - GRADE

Most of the charged off occurred in B,C and D grade



CHARGED OFF - PURPOSE

Debt consolidation purpose have most charged off loan 2767



CONCLUSION

- IT IS SEEN THAT MOST OF THE DEFAULT CUSTOMER BELONGS TO:-
 1. RENTED AND MORTGAGE HOUSES.
 2. CALIFORNIA STATE.
 3. B, C AND D GRADE OF LOAN.
 4. DEBT CONSOLIDATION PURPOSE.
- IT IS CONCLUDED THAT IT WILL BE BENEFICIAL IF THESE PARAMETERS WERE FOCUSED.

