

Lending Club Case Study

Kedar Lachke

Task Performed

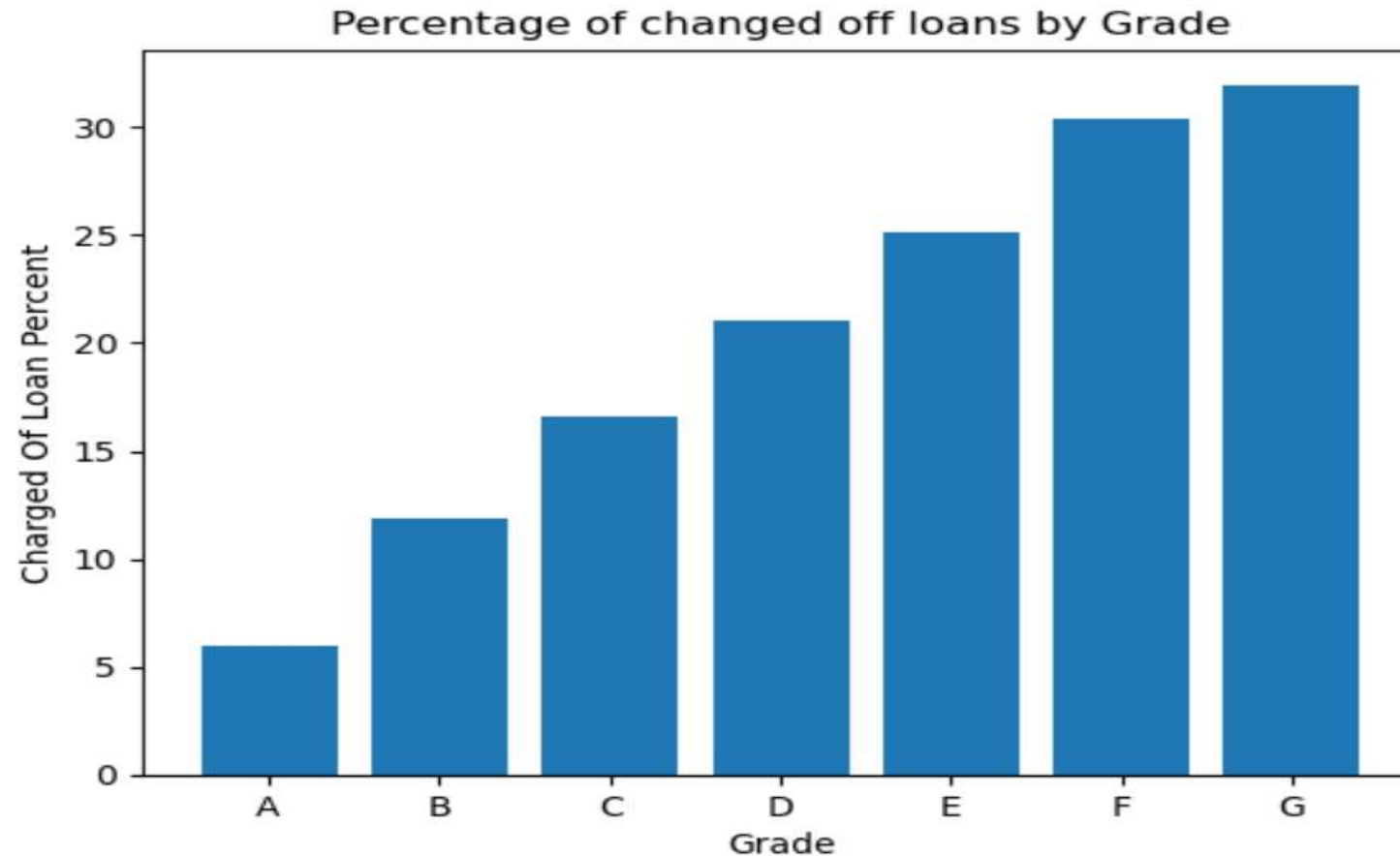
- Data Cleaning
- Univariate Analysis
- Segmented Analysis
- Bivariate Analysis

Data Cleaning

- Deleted all the columns containing 'No Value or Null'
- Derived new columns from existing columns (emp_length convert emp_tenure_num)
- Manipulated data of column int_rate. Data was in string format. So converted to numerical data
- Removed rows having more than 5 blank values

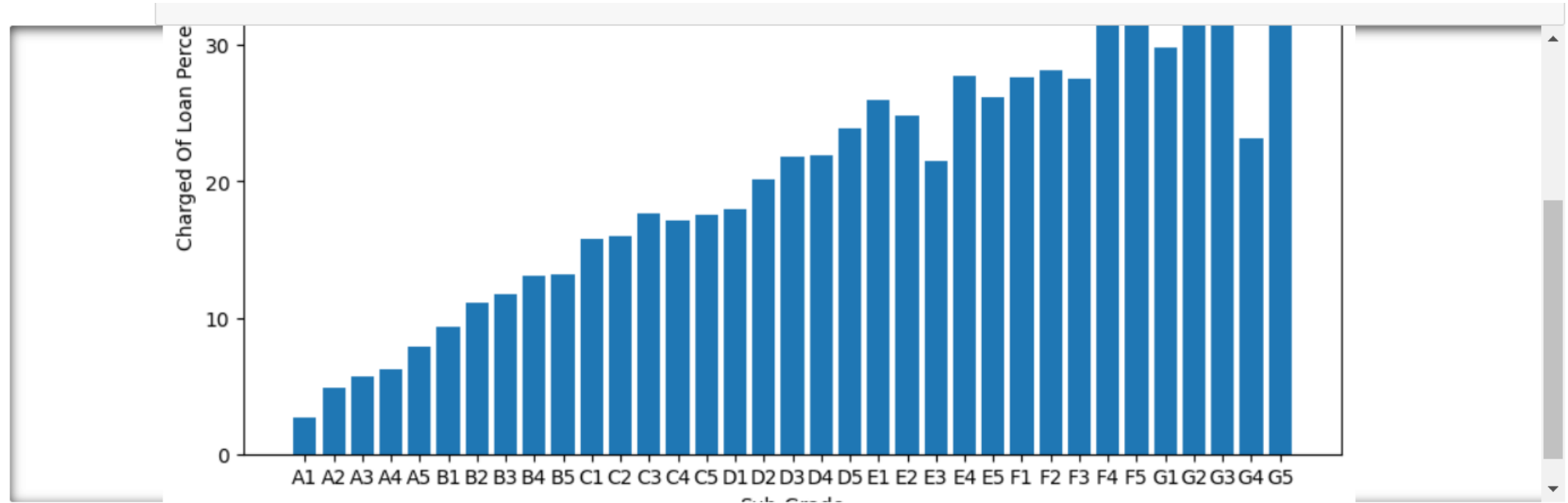
Univariate Analysis

1. Charge off loan with Grade first and the Sub-Grade on the bar chart for comparison



After analysing the above graph, we observe that the as the grading of the applicant decreases, charged Off loan increases. while approving the loan Grade of the applicant is one of the most important parameter to consider.

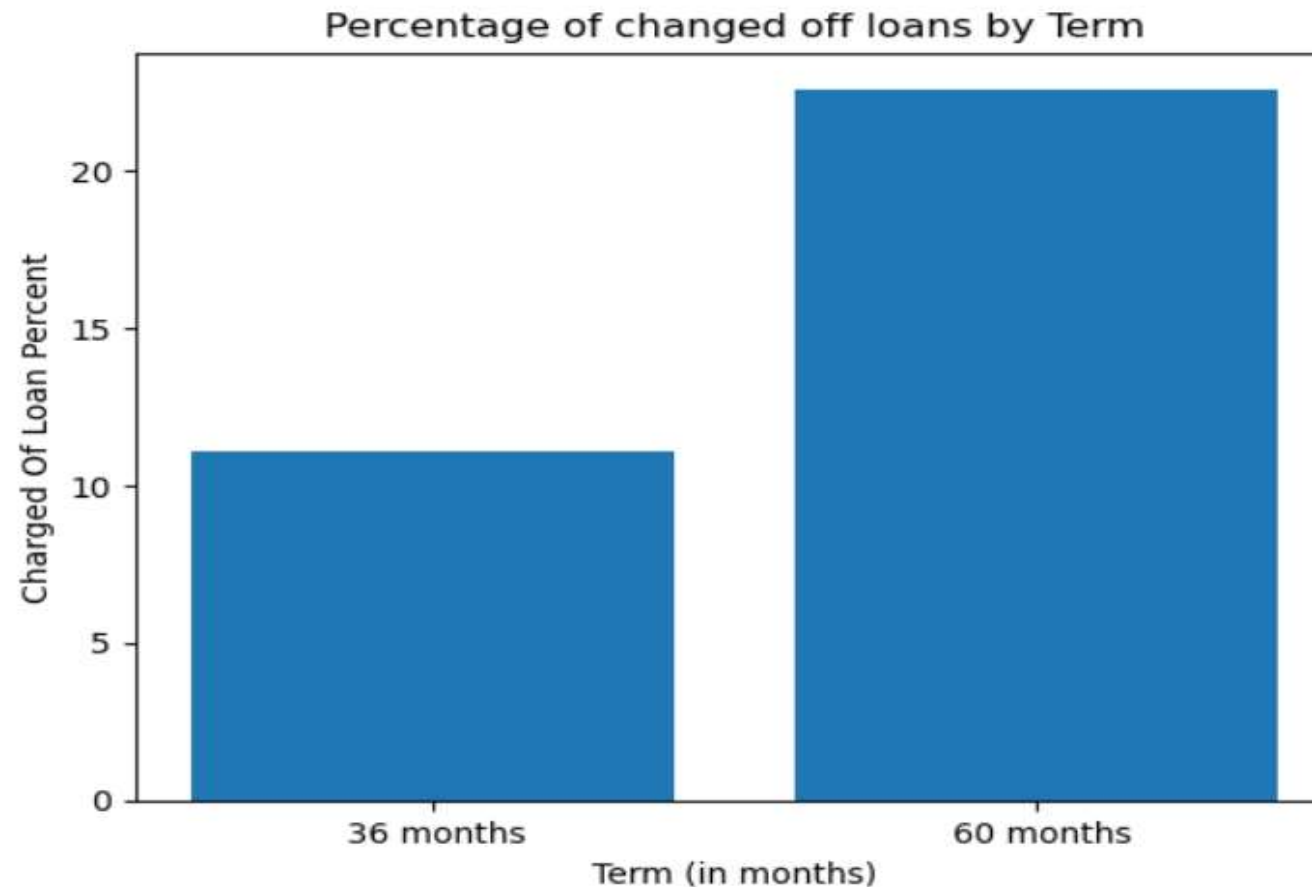
Sub-Grade



After analysing the above graph, we observed that as the grade decreases the rate of charged off loan increases. There is minor anomaly in case of G Sub-

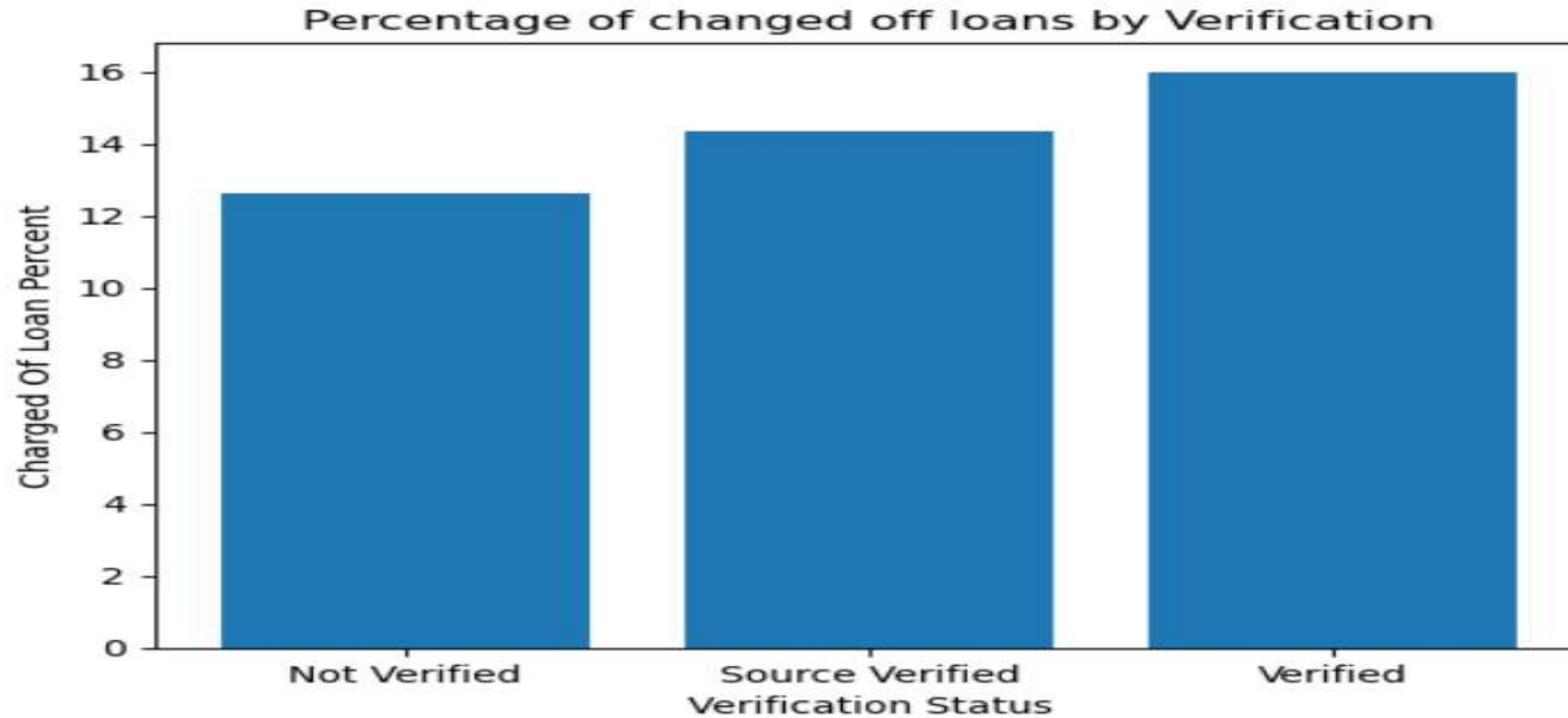
- After analysing the above graph, we observed that as the grade decreases the rate of charged off loans increases. There is a minor anomaly in the case of the G Sub-Grade group, but the trend is the same
- The anomaly is due to less number of applicant in G Sub-Grade group. The details are as below
- ***Total Applicants in G Sub-Grade group***
- G1 104
- G2 78
- G3 48
- G4 56
- G5 30
- ***Total Applicants in G Sub-Grade group***
- G1 31
- G2 28
- G3 19
- G4 13
- G5 10
- ***So we can conclude that Grade and charged off loans are negatively co-related and Grade and Sub-Grade parameters should be consider at the time of approving the loans***

2. Charge off Loans By Loan-term



After analysing the above graph, we observed that as the term increases the rate of charged off loan increases. But only consider Term parameter is not appropriate. Loan Amount and rate of interest should also be consider. But yes Term is and import parameter which should be consider at the time of approving the loan

3. Charge off Loans by Verification Status



After analyzing we observed that verification of the applicant does not make significant impact for charged off the loan. The verification parameter should be used in combination with other parameter like e, Term, Loan amount

Let us analyze the Verification parameter a little more using pivot table.

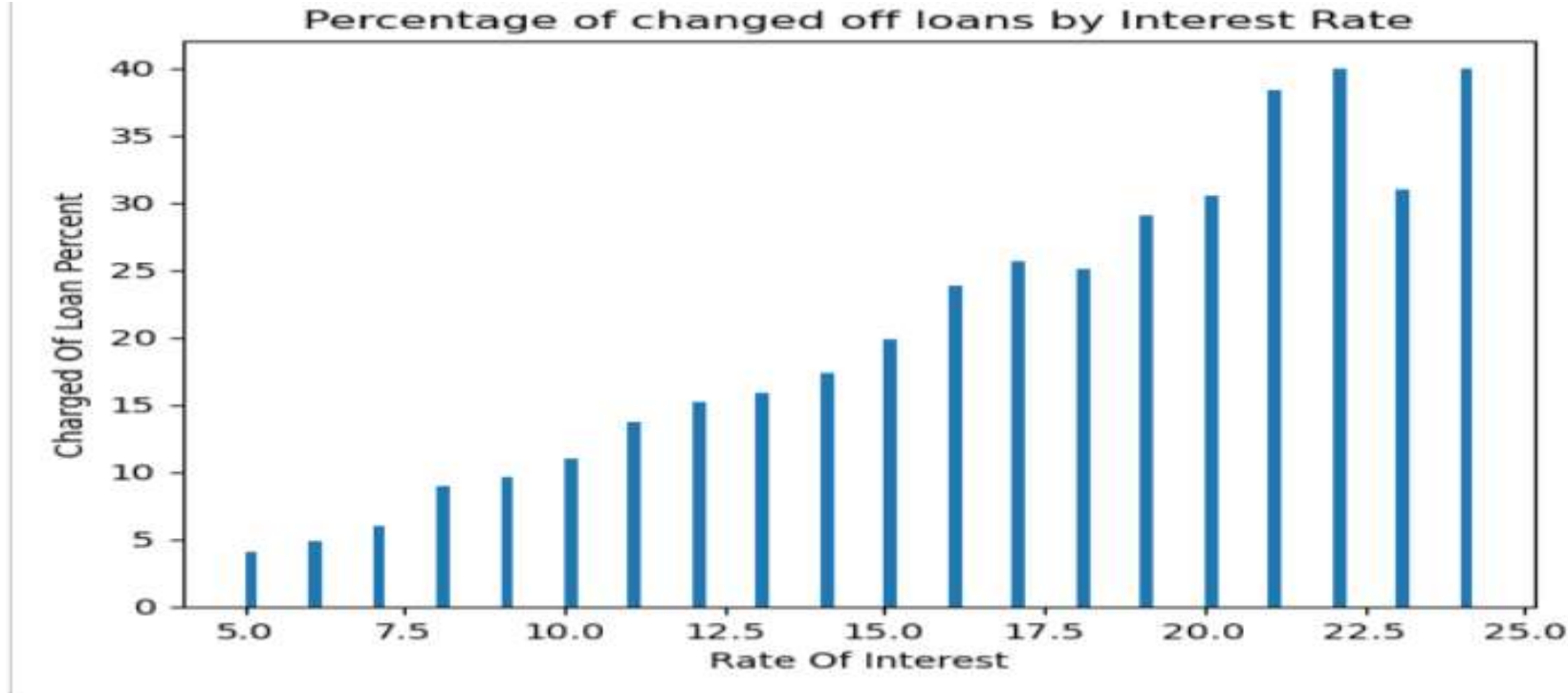
grade	loan_status	verification_status	A	B	C	D	E	F	G	A%	B%	C%	D%	E%	F%	G%
0	Charged Off	Not Verified	316	617	586	408	150	45	17	52.491694	43.328652	43.601190	36.493739	21.008403	14.106583	16.831683
1	Charged Off	Source Verified	140	354	316	311	194	93	26	23.255814	24.859551	23.511905	27.817531	27.170868	29.153605	25.742574
2	Charged Off	Verified	146	453	442	399	370	181	58	24.252492	31.811798	32.886905	35.688730	51.820728	56.739812	57.425743

Instead of calculating the percentage on total verification categories(from cleaned data frame i.e newloan_df) we calculate on total charged off dataframe data frame(i.e chargedoff_df)

After analyzing the data above(in the scenario above), we observe that the verification parameter is important when consider only charged off the Loan frame.

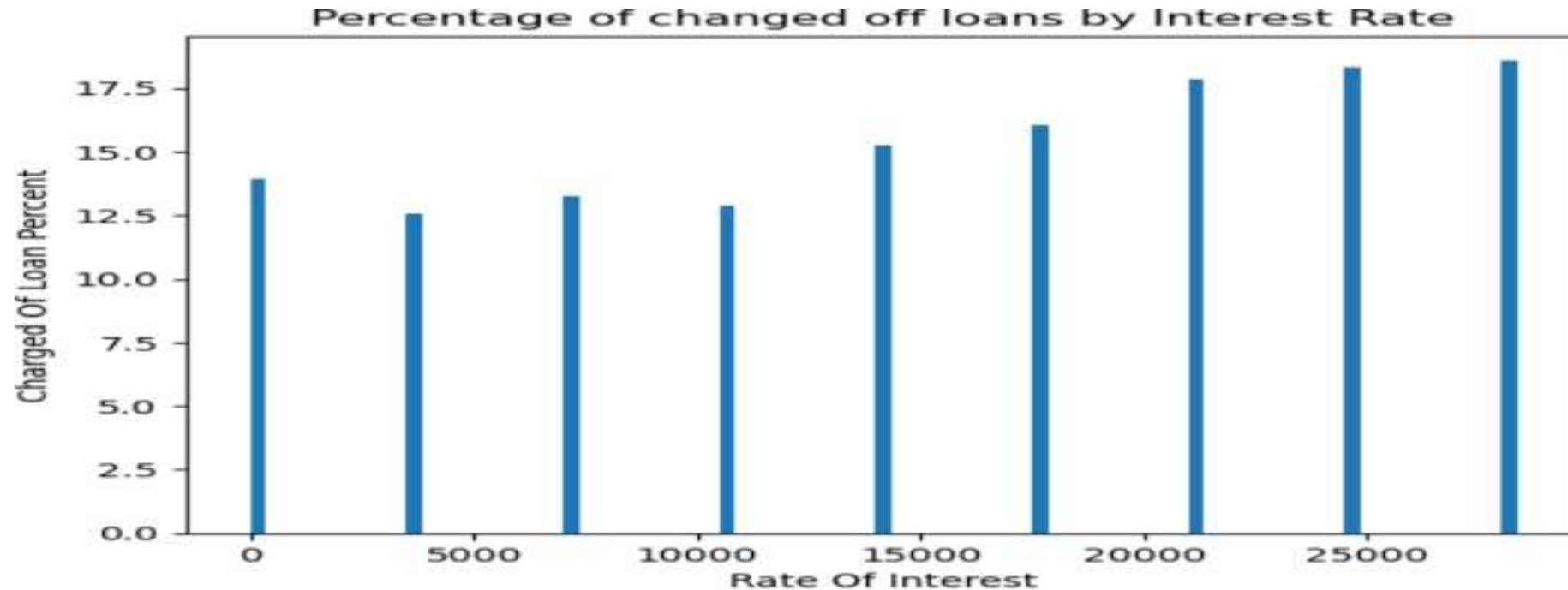
There is an anomaly in E, G, F grade as the number of applicant is less

4. Charge off Loans by Interest Rate



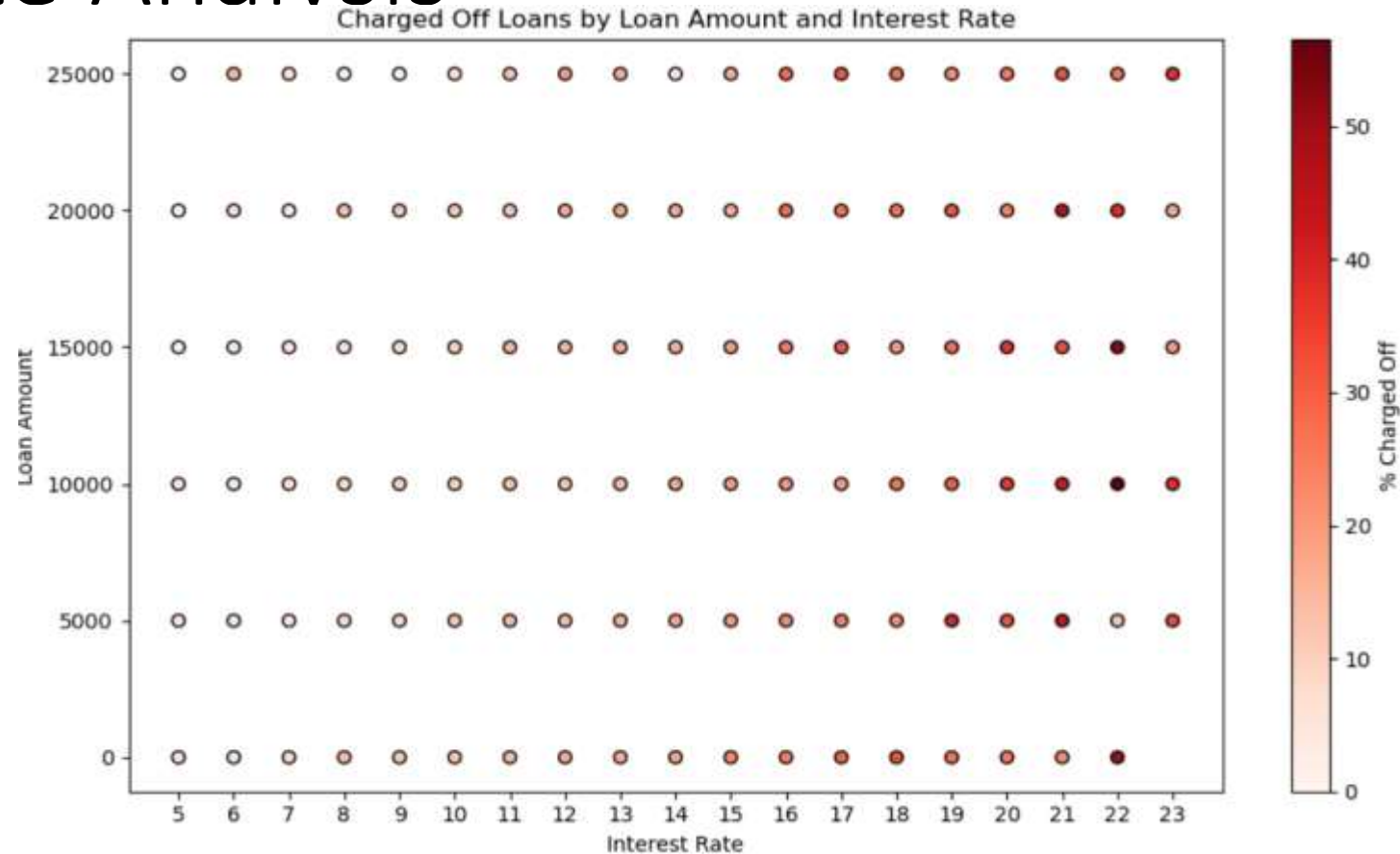
After analyzing the graph, it is observed that with an increase in the rate of interest, the charged off loans increase. But interest should be considered with other parameters like loan amount, and term. After observing the graph we can say that rate of interest and charged off are positively co-related

5. Charge off Loans by Loan Amount



After observing the above graph, we observe that the loan amount does not make much significant impact on charged of loan. So with loan amount other parameter should be consider like interest rate, term, verification

Bivariate Analysis



In the above graph, it is observed that, with increase in interest rate and loan mount, the charged of loan also also increases. i.e with high rate of interest rate for large amount the changes of applicant defaulting is more

Conclusion

- So the 5 parameter that will help us avoiding giving loan to probable defaulter are
- Loan Amount
- Rate Of interest
- Term
- Grade and Sub-Grade of Applicant
- Verification status of Applicant