LendingClub Case Study ANALYSIS

ANIRUDHA SAHU
ANKIT SHRIVASTAVA

Brief description of the project

LendingClub is a financial services company headquartered in San Francisco, California.The goal of this project is to conduct an exploratory data analysis (EDA) and visualize trends, patterns, and insights in the LendingClub's loan dataset. This will help the client better understand loan distributions and identify potential defaulters in the future based on the currently available data so that it can take necessary reformations in its operations to minimize credit loss.

Exploratory Data Analysis (EDA) Steps

1. Data Import & Setup: Loaded loan.csv, used meta-excel dictionary, adjusted display settings.

2. Data Cleaning: Dropped empty, high-missing, single-value, ID, text, and behavioral columns.

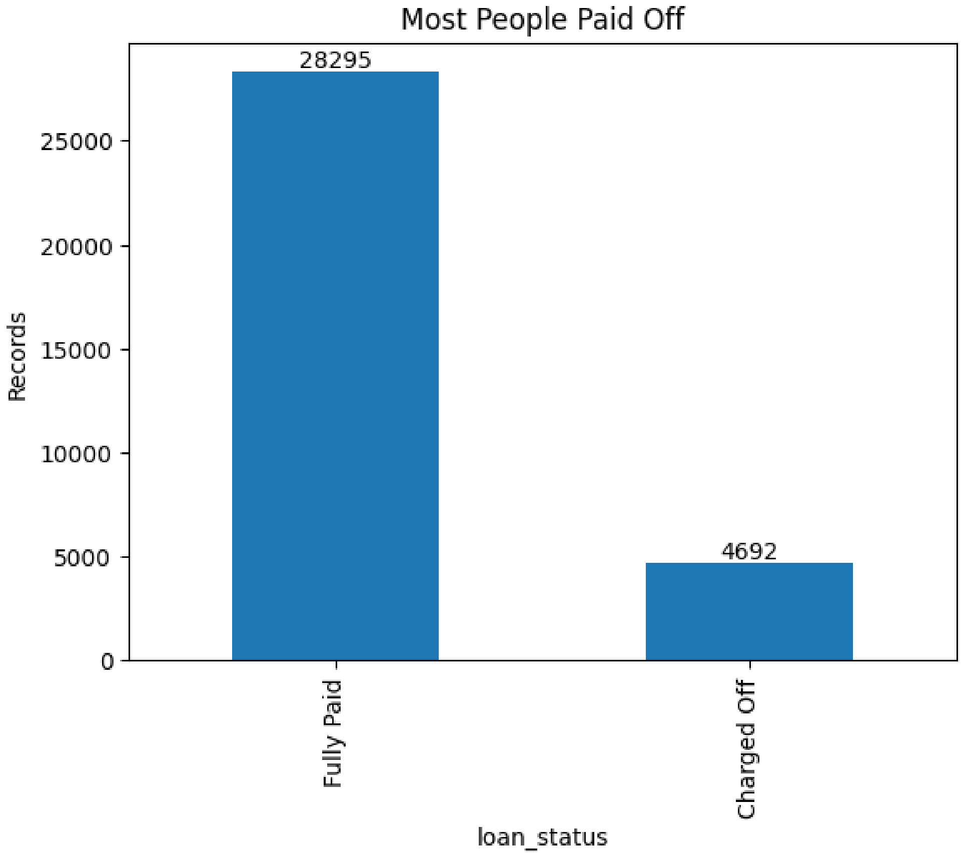
3. Preprocessing: Removed in-progress loans, converted data types, handled missing values and outliers.

4. Charged-Off Proportion: Calculated as Charged-Off / (Charged-Off + Fully Paid) and used for analysis.

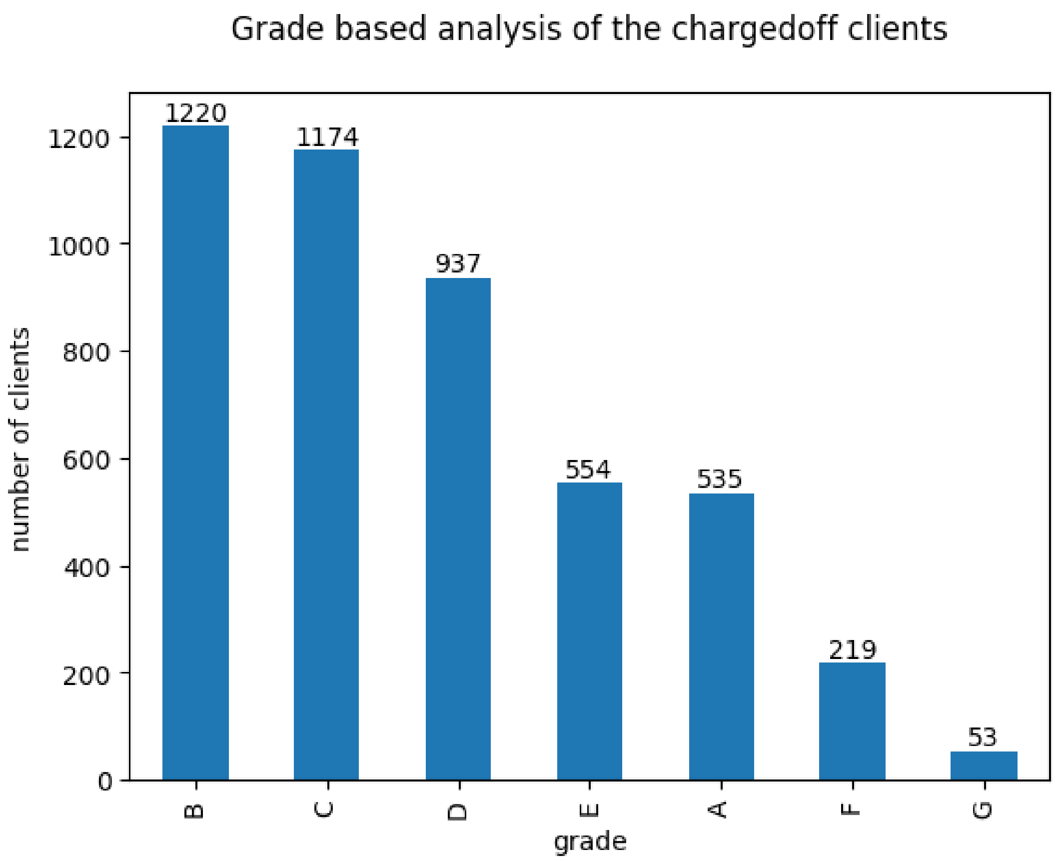
5. Bivariate Analysis Prep: Bucketed annual_inc, loan_amnt, int_rate, and dti, saved cleaned data (df_cleaned_bivariate).

TABLE : CATEGORIZATION OF VARIABLES

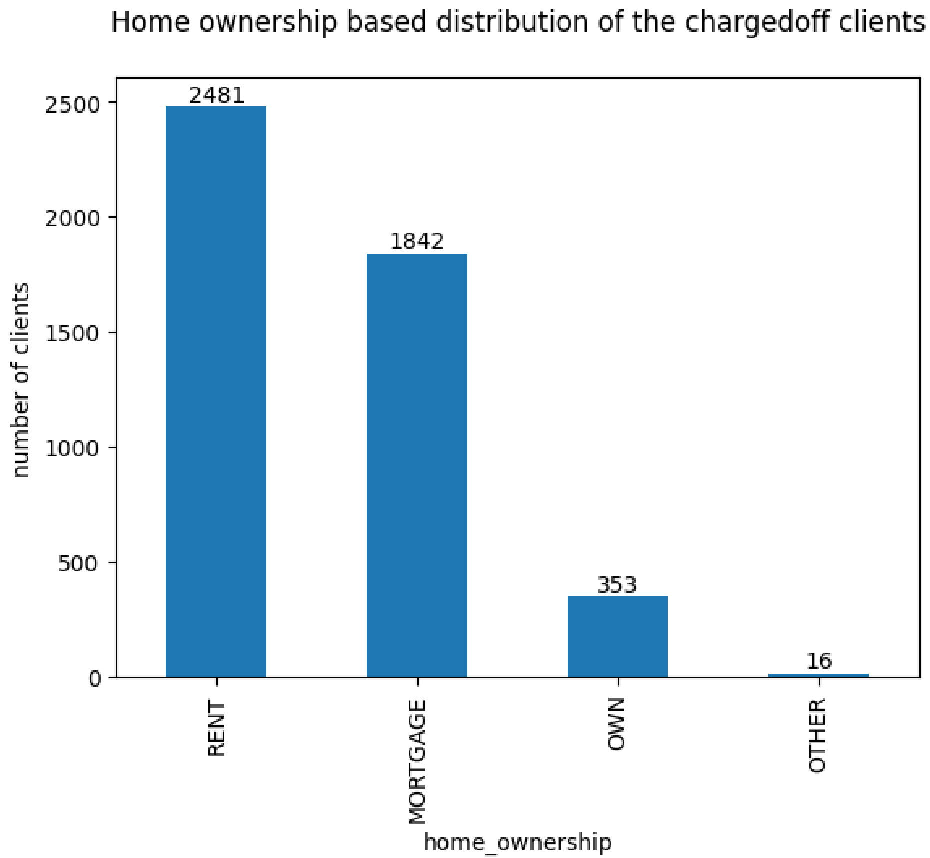
TABLE	Numerical Variable	Ordered Categorical Variable	Unordered Categorical Variable
Loan Info	loan_amt, fund_amt, fund_amt_inv, int_rt, inst	term, grade, sub_gr	id, mem_id
Borrower Info	ann_inc	emp_len	emp_tit, home_own, verify_stat, issue_d, loan_stat, desc, purpose, titl, zip_cd, addr_stat
Credit & Financial Info	dti, del_2yrs, inq_6mths, mths_delinq, mths_record, open_acc, pub_rec, rev_bal, rev_util, tot_acc		earl_cr_line
Payment & Receivables Info	out_prn, out_prn_inv, tot_pymnt, tot_pymnt_inv, tot_rec_prn, tot_rec_int, tot_rec_late, recov, coll_rec_fee, last_pymnt_amt, coll_12mths, chgoff_12mths, pub_rec_bnk, tot_rev_lim		last_pymnt_d, last_cr_pull_d, verif_stat_joint



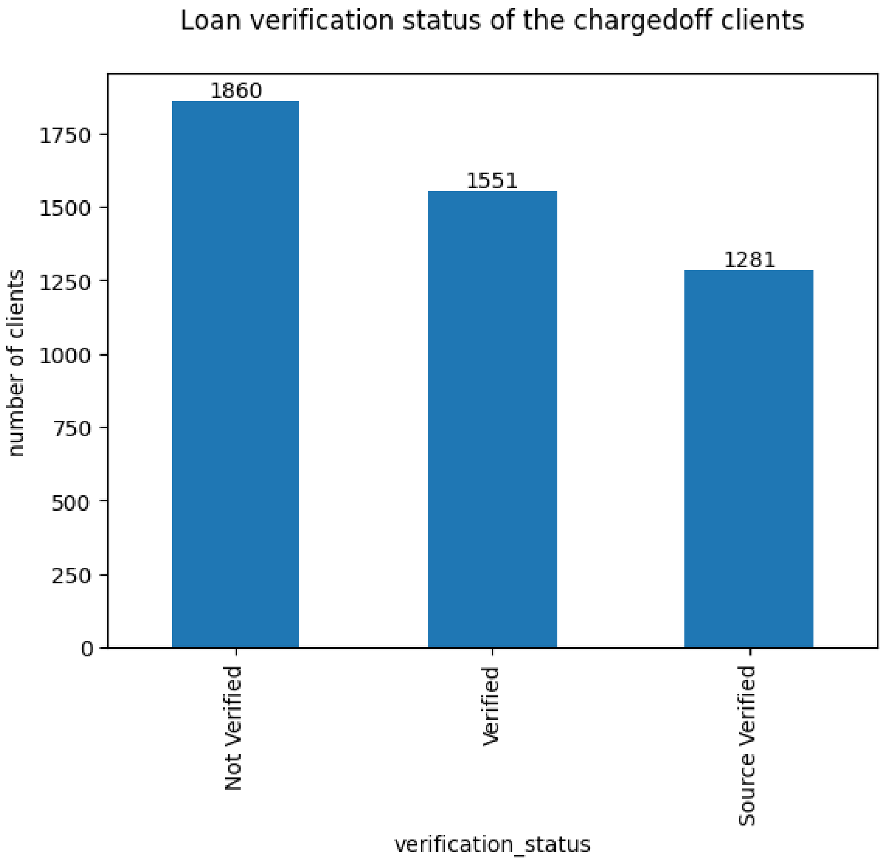
INFERENCE : It looks like most people who took loans has paid off their loans.



The chart indicates that Grades B and C have the highest charge-offs (1220 and 1174), implying higher default risk. In contrast, Grades F and G show the lowest charge-offs (219 and 53), suggesting lower risk. This insight highlights the importance of careful credit evaluation for clients in Grades B and C to mitigate potential losses.



Clients with RENT (2481) and MORTGAGE (1842) statuses have the highest charge-offs, indicating they pose a higher default risk. Clients with OWN (353) and OTHER (16) statuses have lower charge-offs, suggesting they are less risky. This insight emphasizes the importance of careful credit evaluation based on housing status to mitigate potential losses



The chart shows that loans categorized as Not Verified have the highest charge-offs (1860), followed by Verified (1551) and Source Verified (1281). This suggests that verifying loan information reduces the risk of charge-offs, highlighting the importance of thorough verification processes in lending practices to minimize defaults.

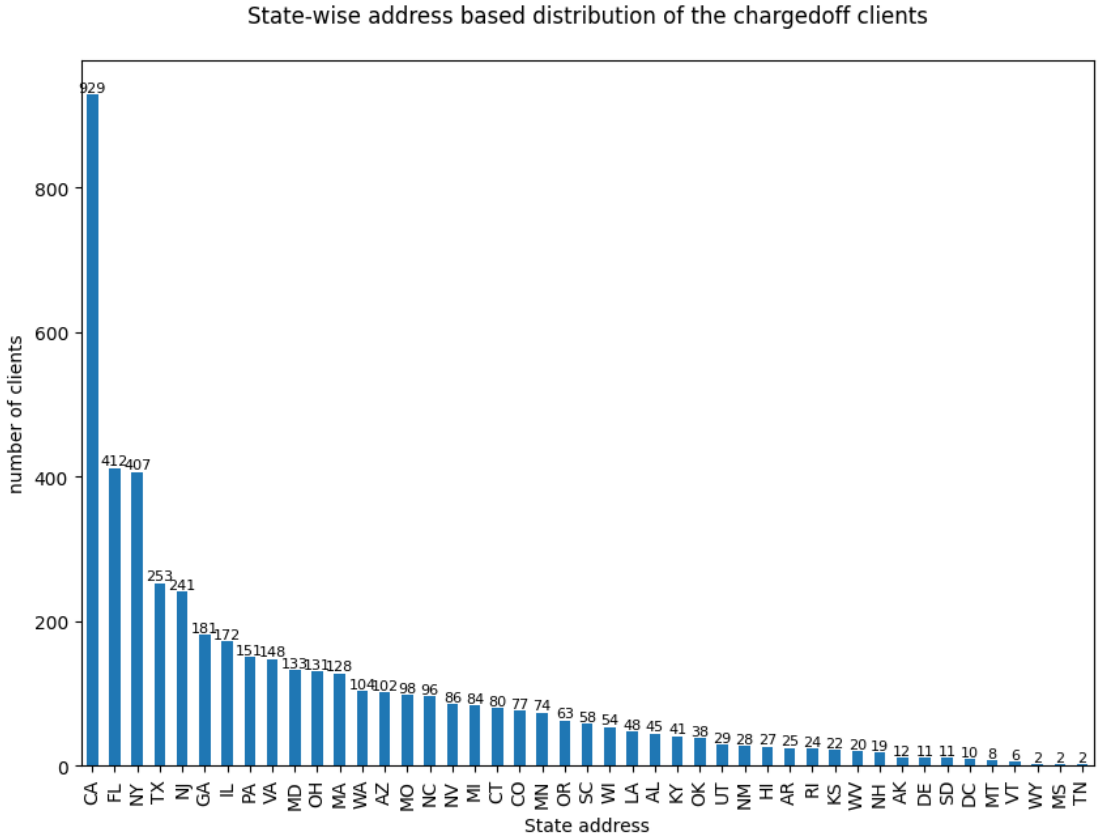
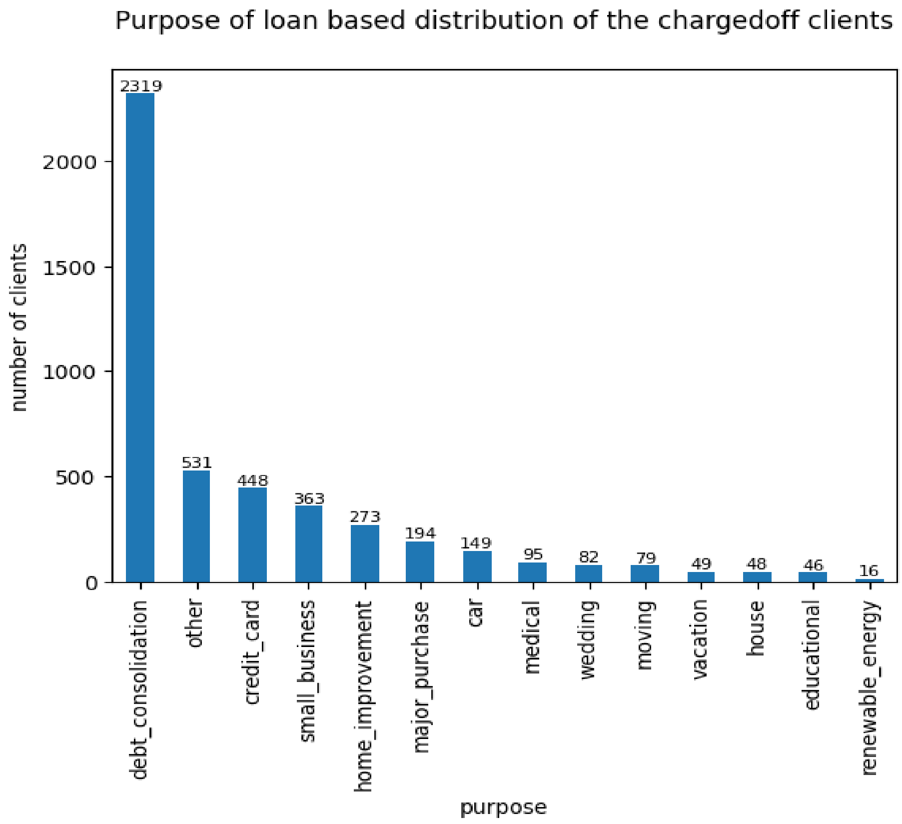
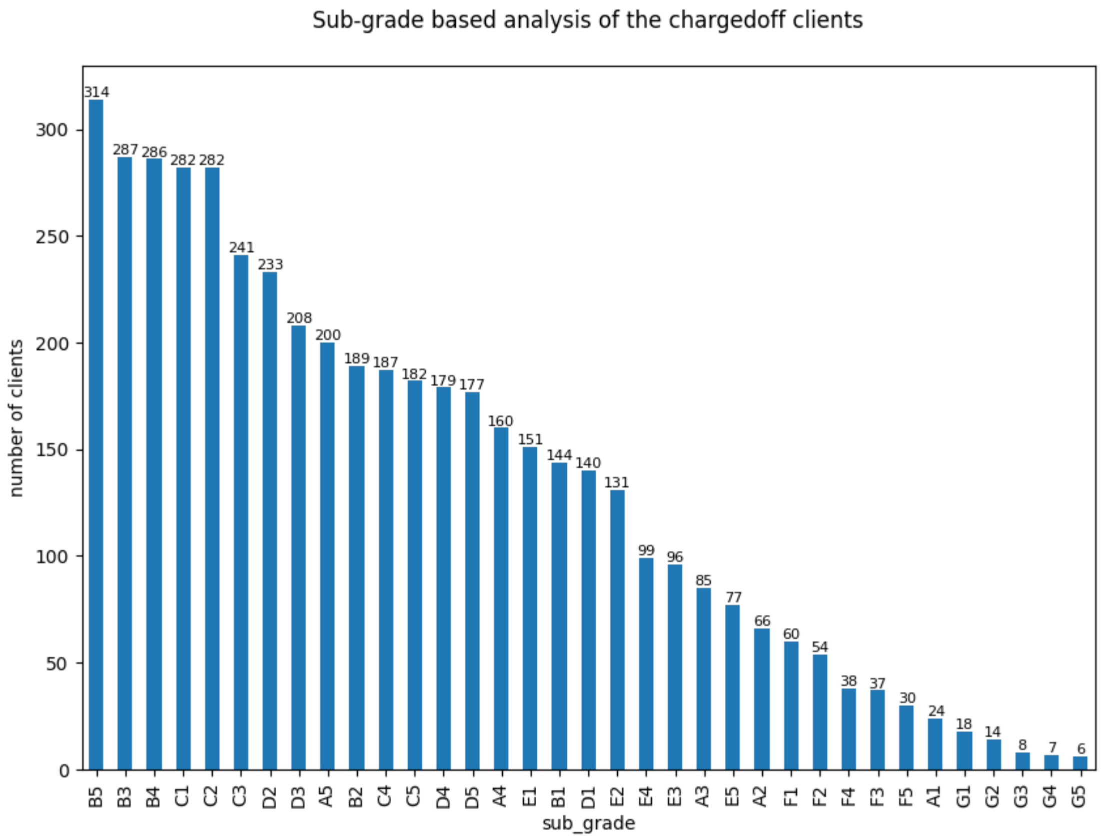


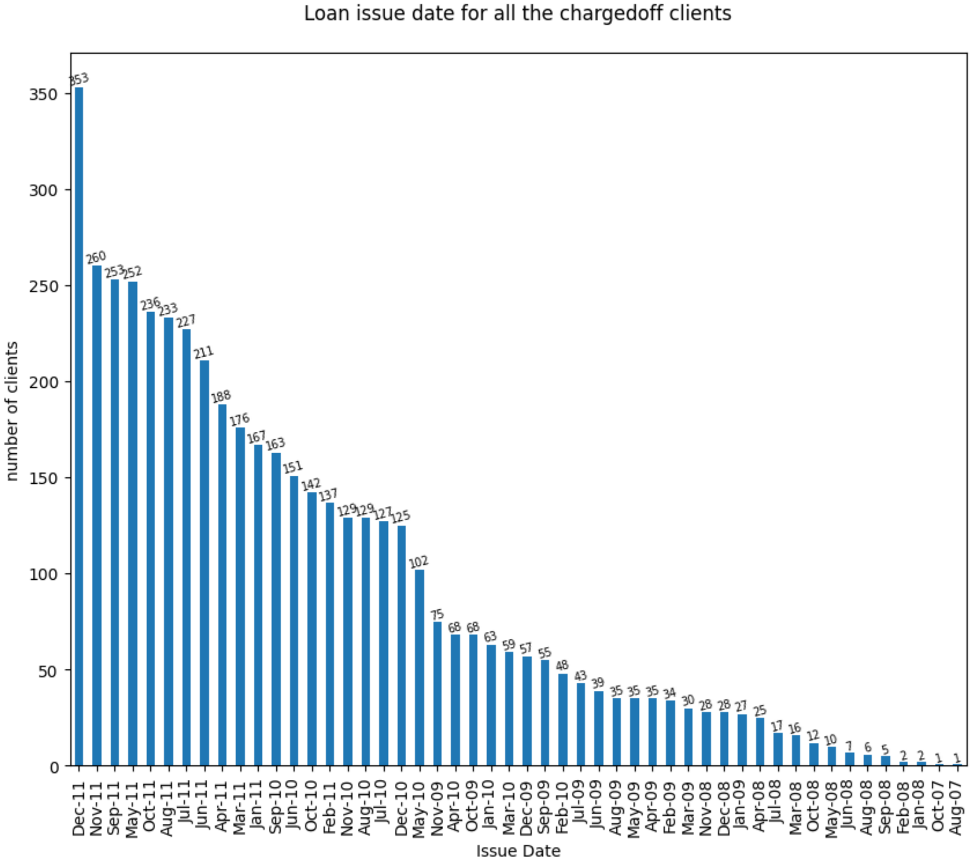
Chart shows California (CA) with the highest number of charged-off loans (929), followed by Florida (FL) with 412 and New York (NY) with 407. These states also have the highest loan approvals and applications, indicating a strong correlation between high loan activity and default rates. This suggests that while these states have a high volume of lending, they also face significant default risks, emphasizing the need for rigorous credit evaluation and risk management practices in high-activity regions to mitigate potential losses.



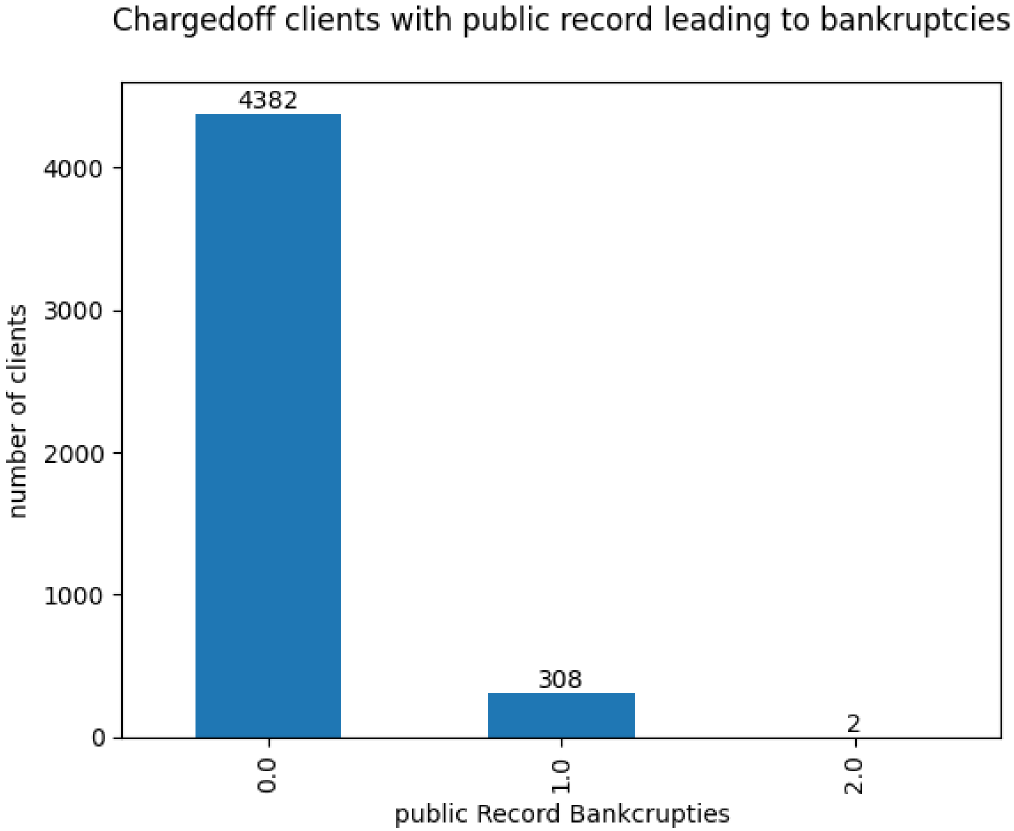
The chart highlights that debt consolidation has the highest number of charged-off clients (2319), followed by other purposes (531) and credit card loans (448). This indicates a significant risk associated with debt consolidation loans, suggesting that borrowers in this category may be struggling with their financial obligations. Financial institution should implement stricter credit evaluations and offer targeted support measures for these high-risk loans to mitigate potential losses and assist borrowers in managing their debt more effectively.



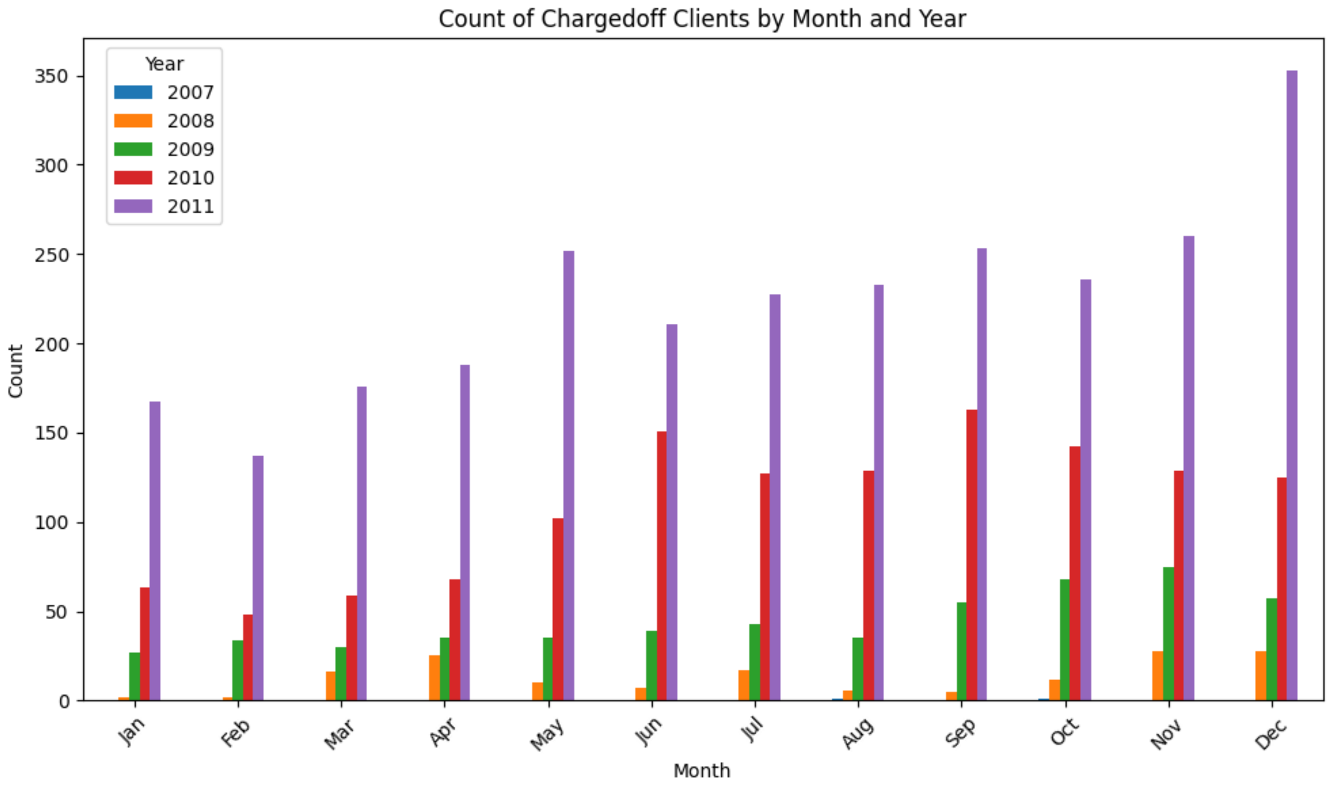
The chart reveals that sub-grade B5 has the highest number of charged-off clients (314), followed closely by B3 (287) and B4 (286). This indicates that sub-grades in the B category are at a higher risk of default. Sub-grades G1 to G5 have the lowest charge-offs, suggesting they pose minimal risk. Financial institutions should focus on improving credit assessments and risk management strategies for sub-grades within the B category to mitigate potential defaults and enhance financial stability



The chart indicates that loans issued in December have the highest charge-offs (353), suggesting increased default risk during this month. January, March, July, and November also show higher charge-offs, while February, May, September, and August have lower risks. Overall, year-end and early months are riskier periods for loan defaults, highlighting the need for lenders to reassess strategies and enhance risk management during these times to reduce potential losses.



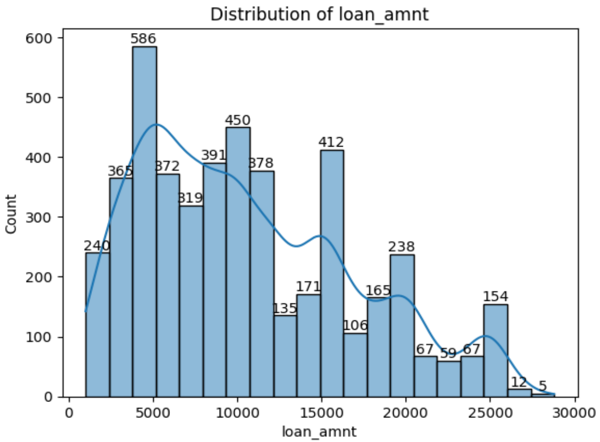
The chart indicates that clients with no public record bankruptcies (4382) have the highest charge-offs, highlighting significant default risk in this group. In contrast, clients with one public record bankruptcy (308) and none with two public bankruptcies suggest that a single bankruptcy might deter further defaults. Financial institutions should focus on proactive risk management for clients without bankruptcies to mitigate potential losses.



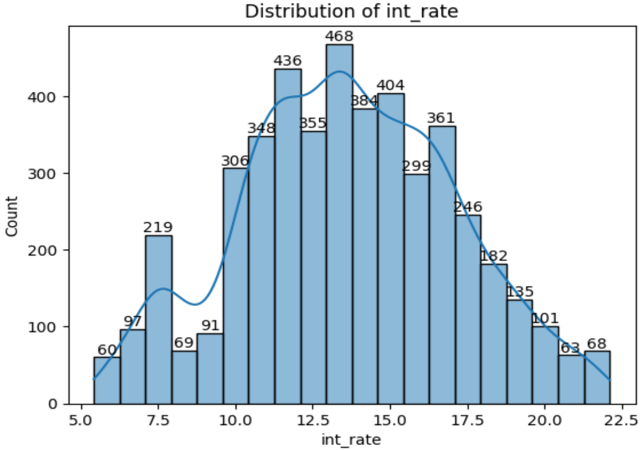
Inferences:

1.There is an yearly increase in the number of applicants those have been chargedoff.
This might imply that either more people have started using the bank recently or the bank is not able to optimize the protocol for screening trust worthy loan applicants.

1.



2.



Inferences:

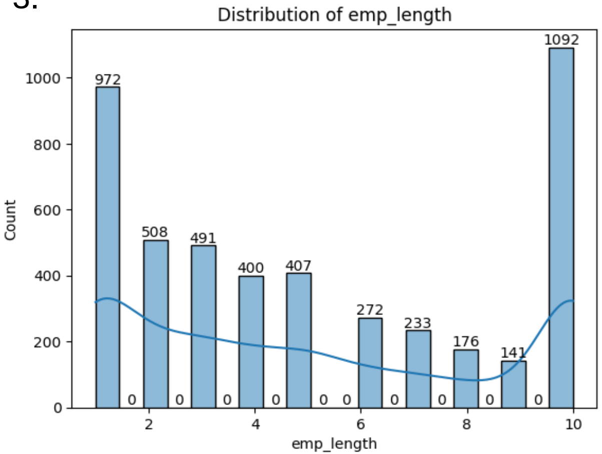
1.Most clients who are being charged have taken a loan amount in between 3K to 5K, followed by 10K, followed by in between 15K and 16K

2.The interest rate for most loans availed by the chargedoff clients are in between 10% to 17.5%

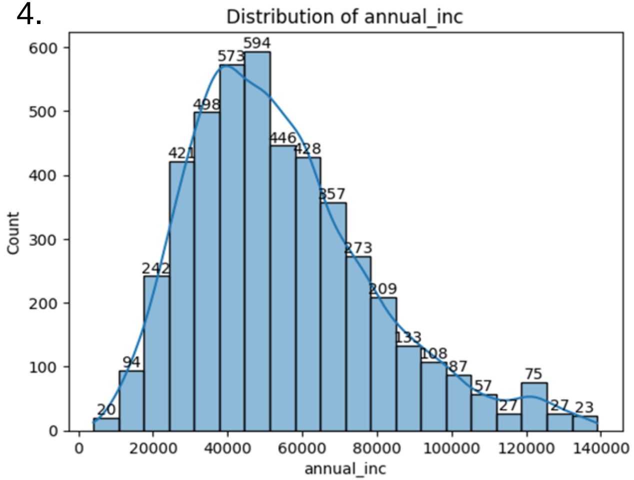
3.Most clients those who had applied for the loan and have chargedoff are either freshers or have experience above 10 years.

4.Most clients those earn in the range of 20K and 60K have the most chances to default.

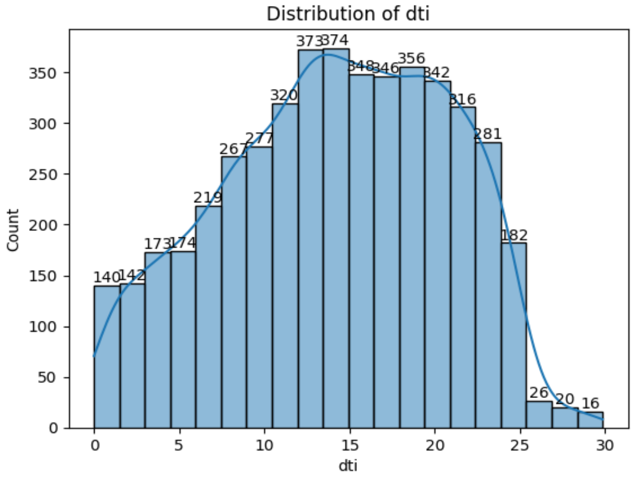
3.



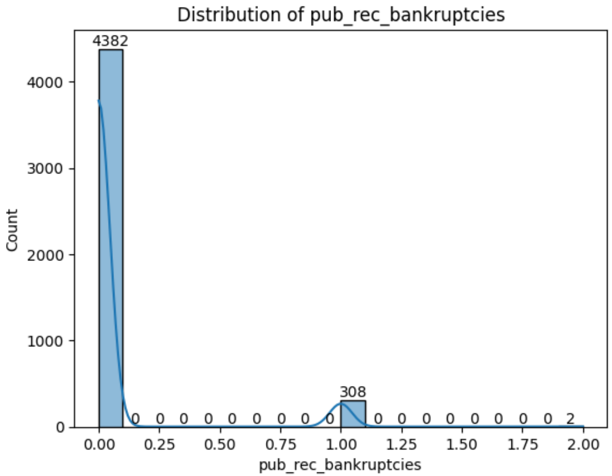
4.



1.



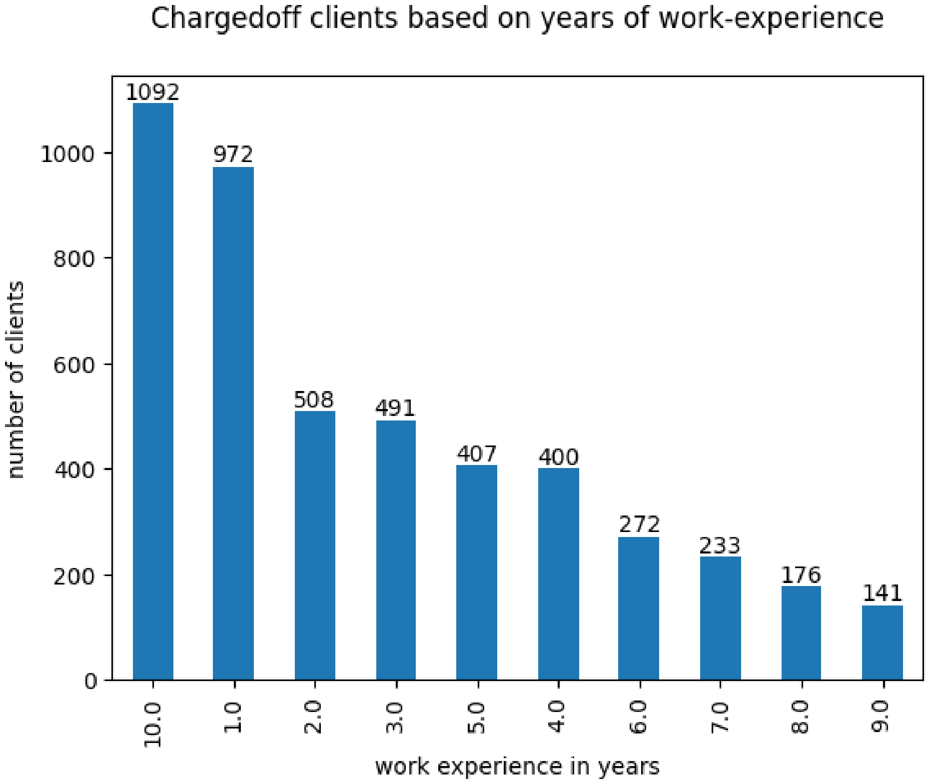
2.



Inferences:

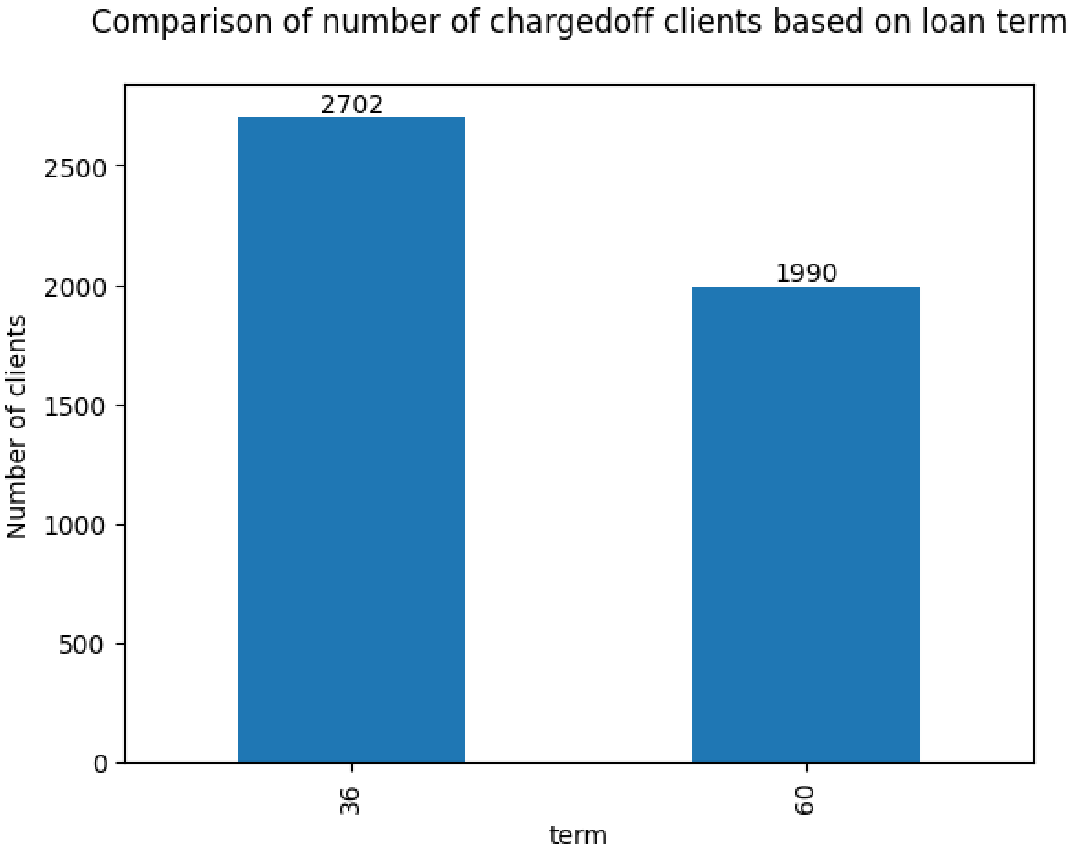
1.The debt:income ratio of most of the chargedoff clients lie in between 11 and 19.

2.Most chargedoff clients do not have a public record of bankcrupties.

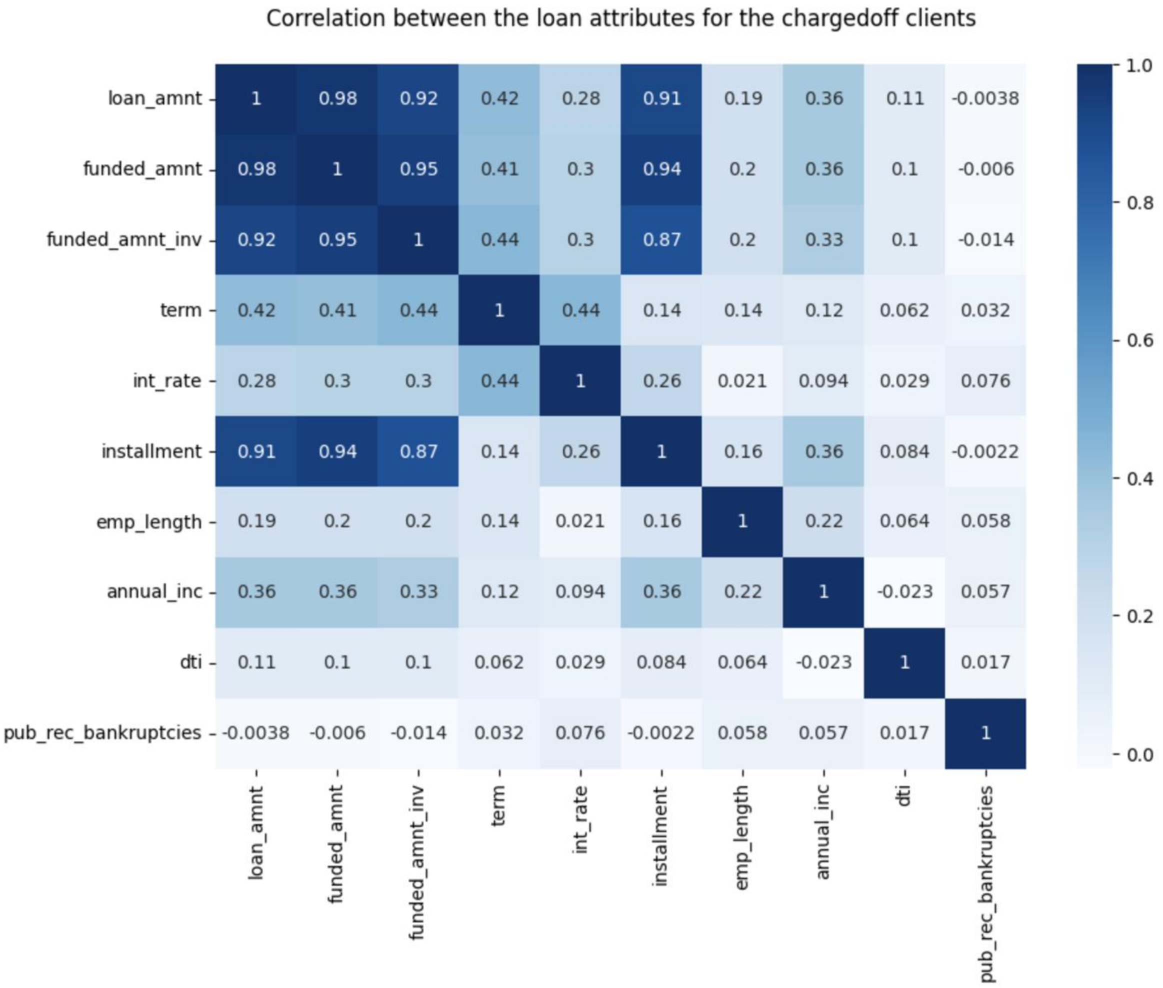


Note: 10.0 refers to >10 years and 1.0 refers to <1 year of work experience

The chart reveals that clients with over 10 years of work experience have the highest charge-offs (1092), indicating significant default risk. Clients with less than 1 year also show high charge-offs (972), suggesting early-career clients are vulnerable. Financial institutions should implement targeted risk management strategies for both experienced and early-career clients to mitigate potential defaults



The chart shows that 36-month loan terms have higher charge-offs (2702) than 60-month terms (1990). This suggests that shorter loan durations may be associated with higher default risk.



Inferences:
1.Positive correlations:

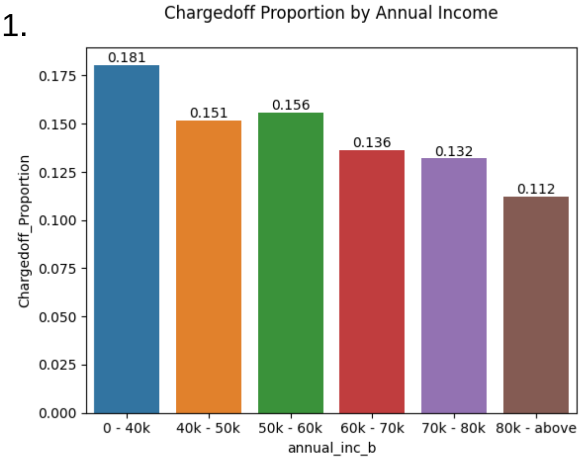
•loan_amnt, funded_amt, funded_amnt_inv and installment are highly correlated with each other

•All the correlations above are positive and therefore are directly proportional with each other

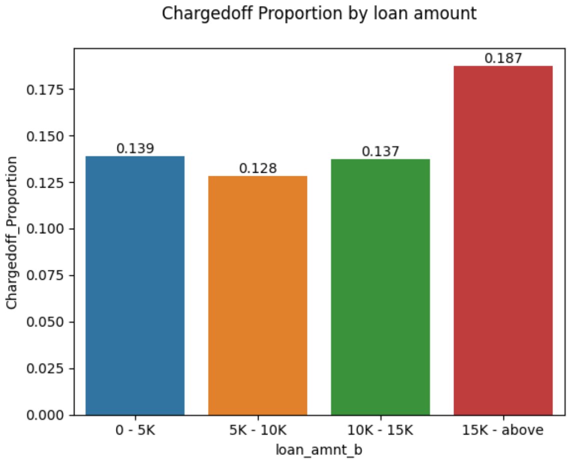
2.Negative correlations:

•loan_amnt, funded_amt, funded_amnt_inv and installment are negatively correlated with pub_rec_bankrupticies.

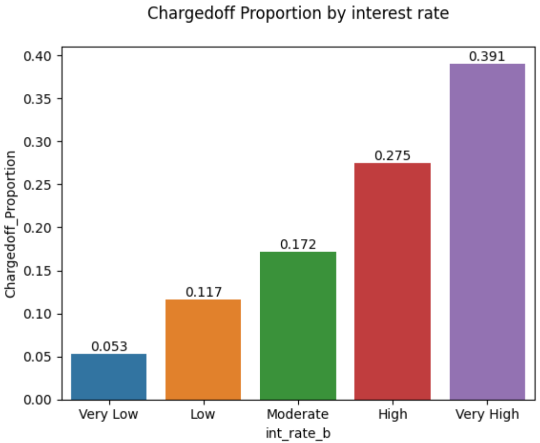
1.



2.

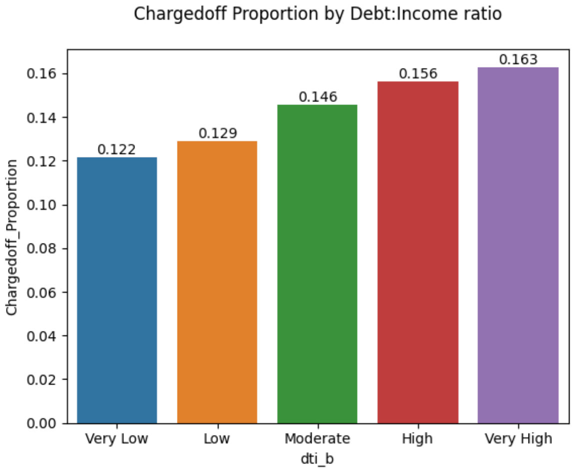


3.



Very Low: 0% to 9% | Low: 9% to 11% | Moderate: 11% to 13% | High: 13% to 15% | Very High: Above 15%

4.



Very Low: 0 to 8 | Low: 8 to 12 | Moderate: 12 to 16 | High: 16 to 20 | Very High: Above 20

Inferences:

1.1.Annual income between 0-40K have the most number of defaulters.

1.2.Clients with lower annual income have more chances of defaulting.

2.Clients those availed loan at a higher interest rate have more chances to default

3.Clients those have availed a higher loan amount have more chances to default

4.Clients those have higher debt to income ratio have more chances to default

Thank you