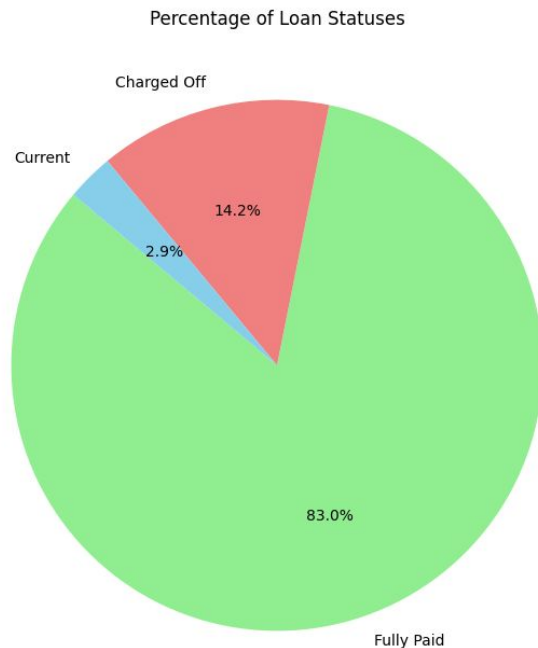# Lending Club Case Study

Supriyo Roy & Michael Thomas
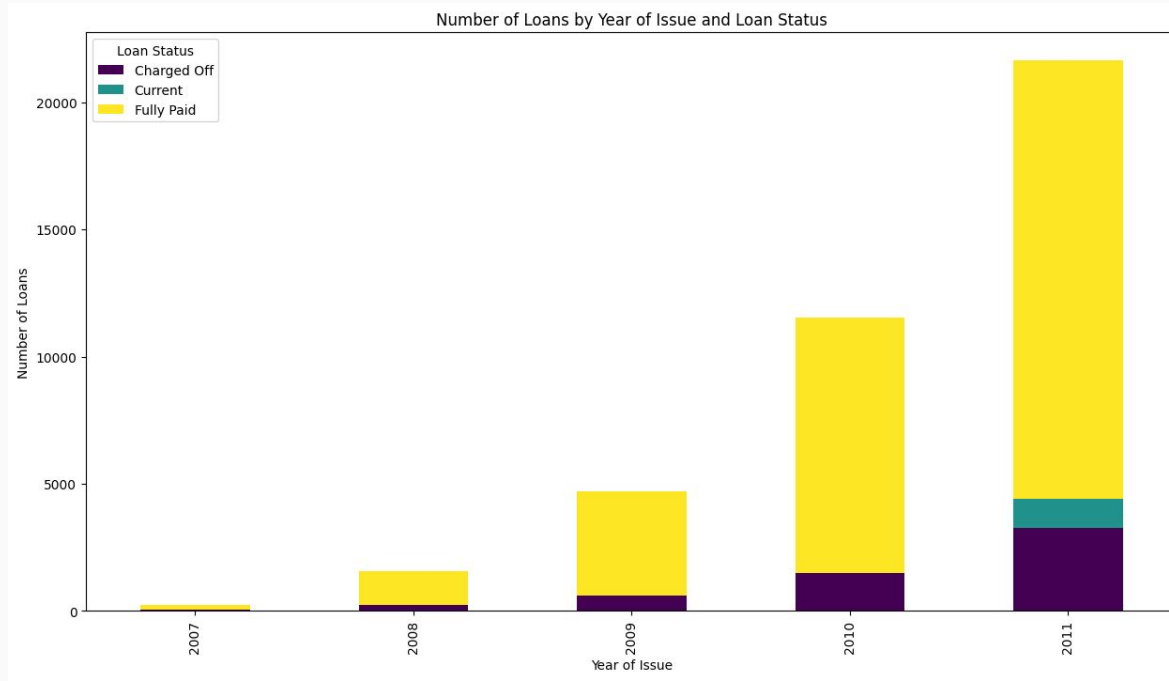
Intro
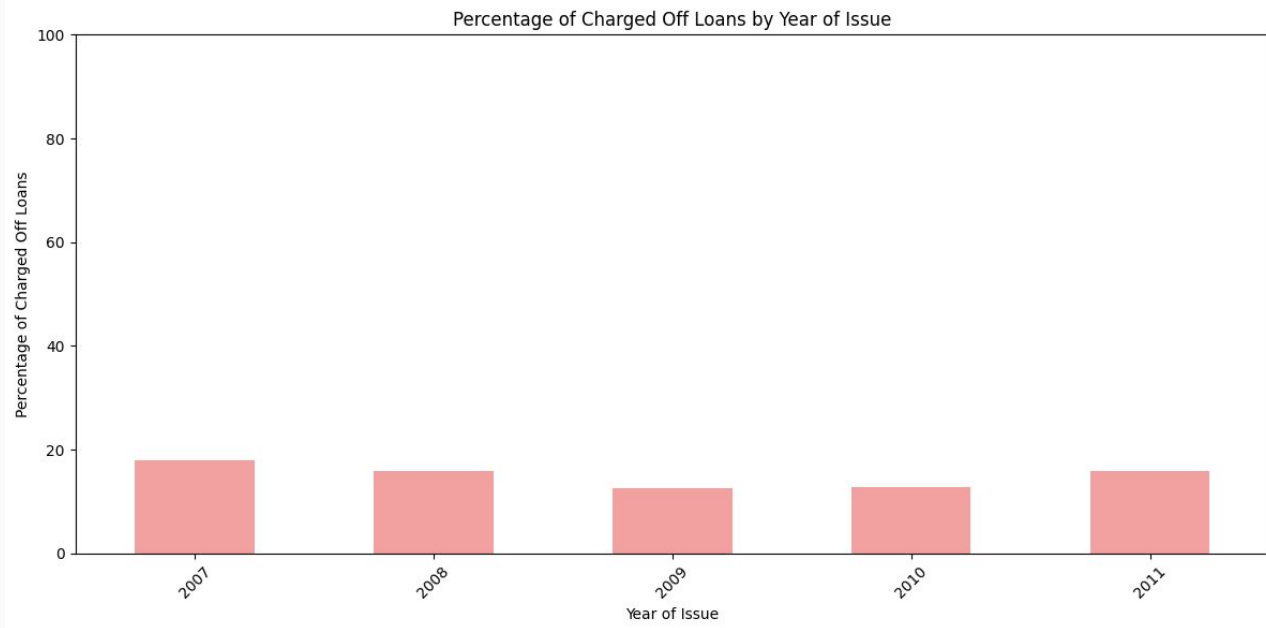
- Approximately 14% of loans are charged off
- The objective of this presentation is to assess how charge-offs can be reduced

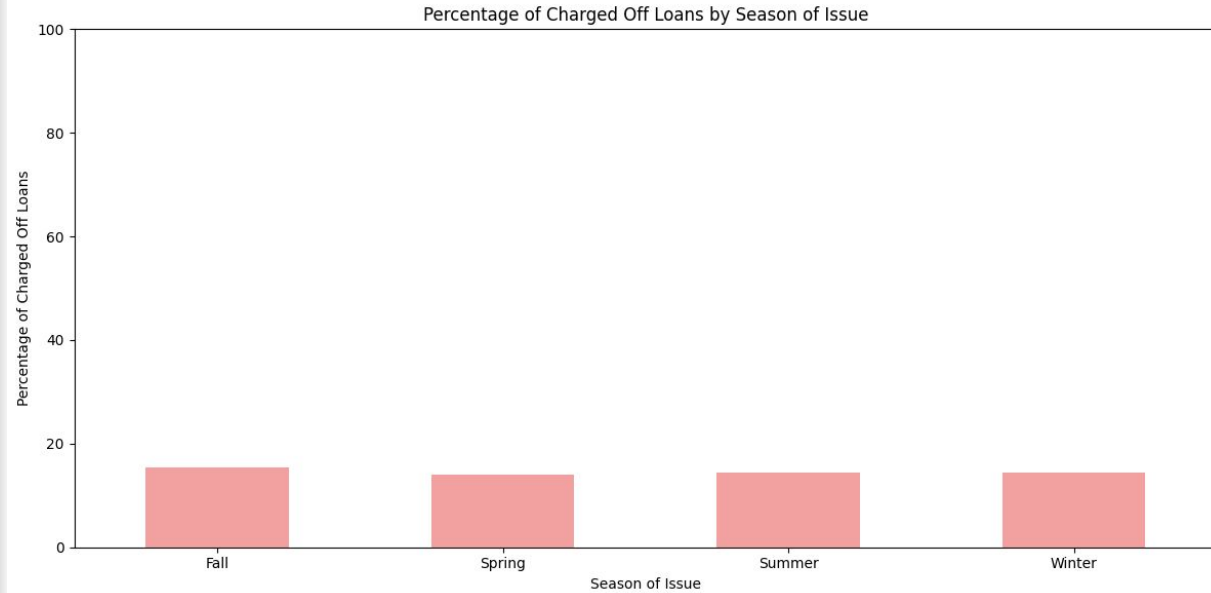

Percentage of Loan Statuses

Loans Have Significantly Grown between 2009-2017

Number of Loans by Year of Issue and Loan Status

Percentage of Charged Loans Remained Stable Over 2007-2011

Percentage of Charged Off Loans by Year of Issue

Charge Offs Do Not Show Signs of Seasonality

Percentage of Charged Off Loans by Season of Issue

Average Recoveries on Loans is Less than 5%

Histogram of Recovery Percentage for Charged Off Loans
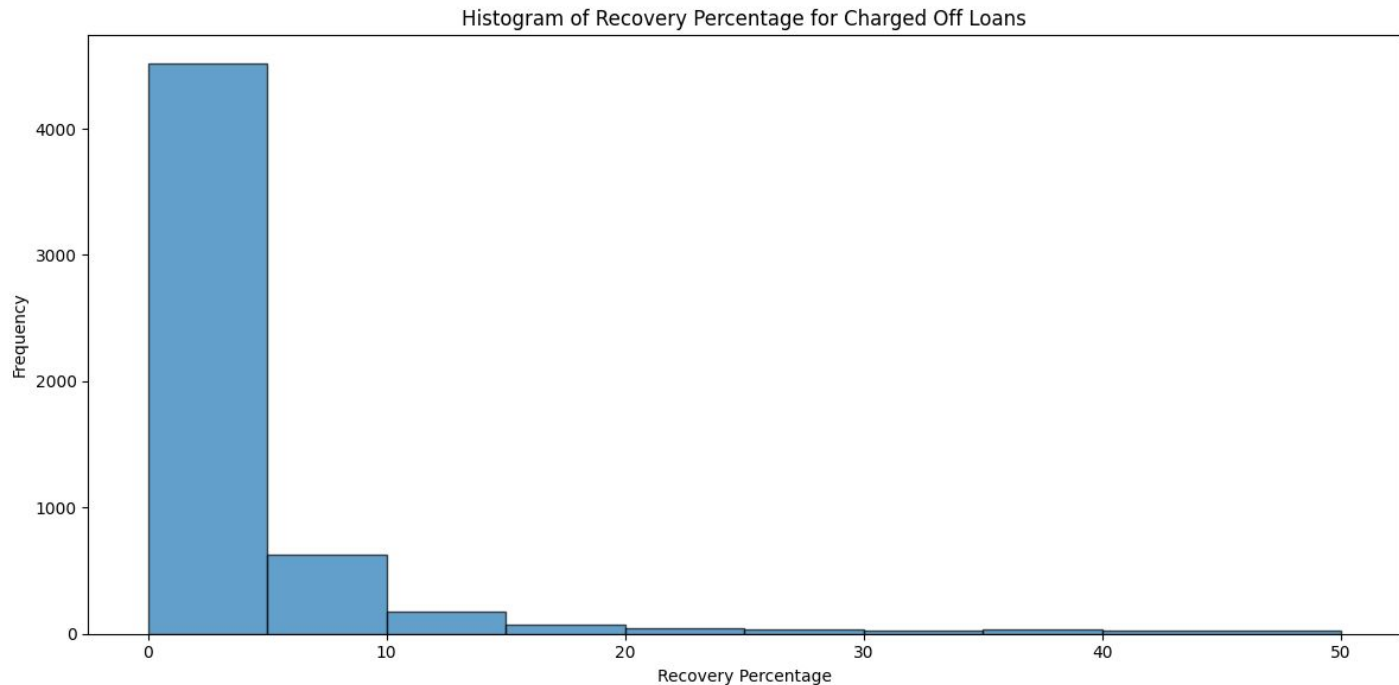
Scope

Data Type	Credit Data Existing When Loan was Made	Credit Data Arising After Loan was Made
Column Names	'annual_income_category' 'term', 'grade', 'sub_grade', 'emp_length', 'home_ownership', 'verification_status', 'issue_d', 'dti', - ratio of monthly paments to debt obligations is relevant 'addr_state', - e.g. are some areas may have more concentrated credit risk 'inq_last_6mths',	- 'loan_status', - 'delinq_2yrs', - 'mths_since_last_delinq', - 'open_acc', - 'total_acc', - 'pub_rec', - to check which column it goes in - 'earliest_cr_line', - 'int_rate', - 'total_rec_late_fee',
Decisions to be Made with Data	- Deny/ Accept Loan - Select an Interest Rate	- Reduce the Amount of the Loan - Change Interest Rate

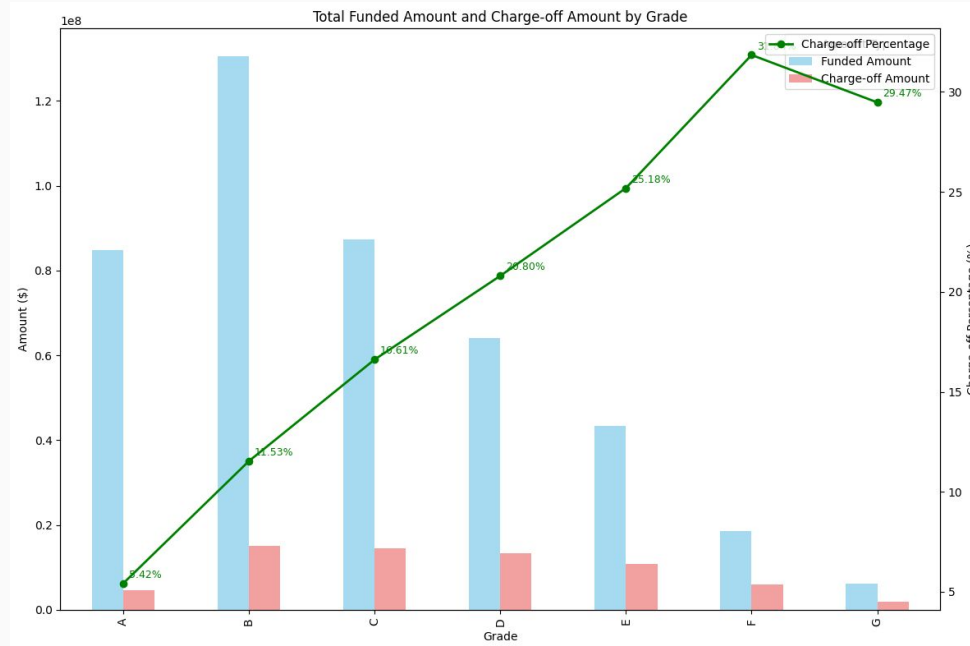# Credit Data Existing When Loan was Made

Relevant to
- Deny/ Accept Loan
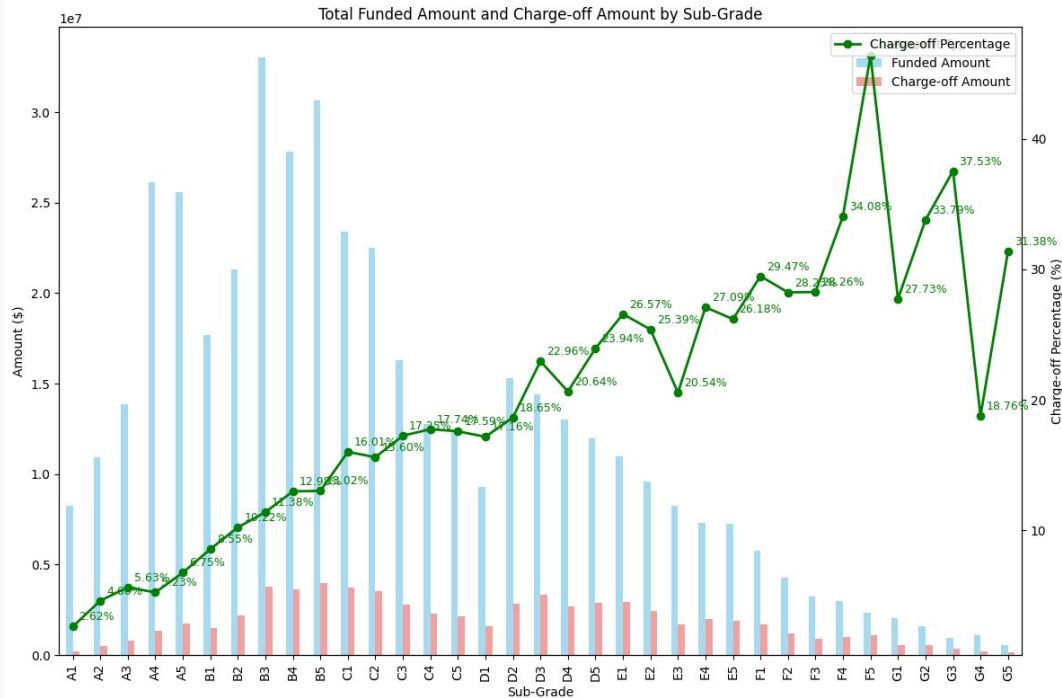- Choose an Interest Rate

Strong Relationship Between Grades and Charge-Offs

Total Funded Amount and Charge-off Amount by Grade

Strong Relationship Between Sub-Grades and Charge-Offs

Total Funded Amount and Charge-off Amount by Sub-Grade

Strong Relationship Between DTI and Charge-Offs

Total Funded Amount and Charge-off Amount by DTI
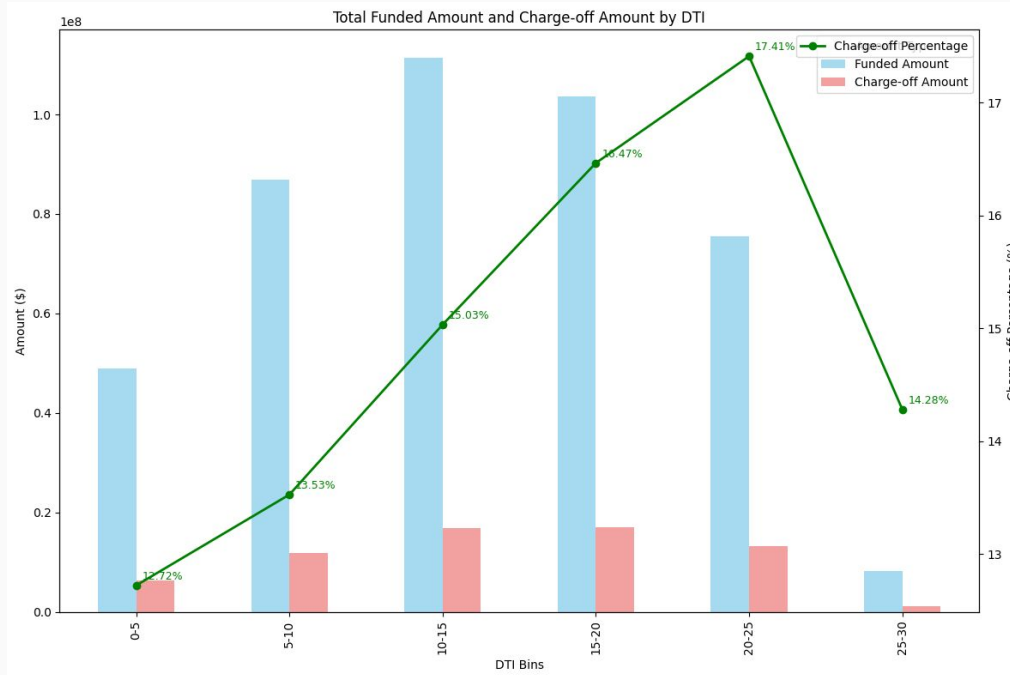
Strong Relationship Between Number of Credit Enquiries and Charge-Offs

Total Funded Amount and Charge-off Amount by Number of Credit Inquiries in Last 6 Months
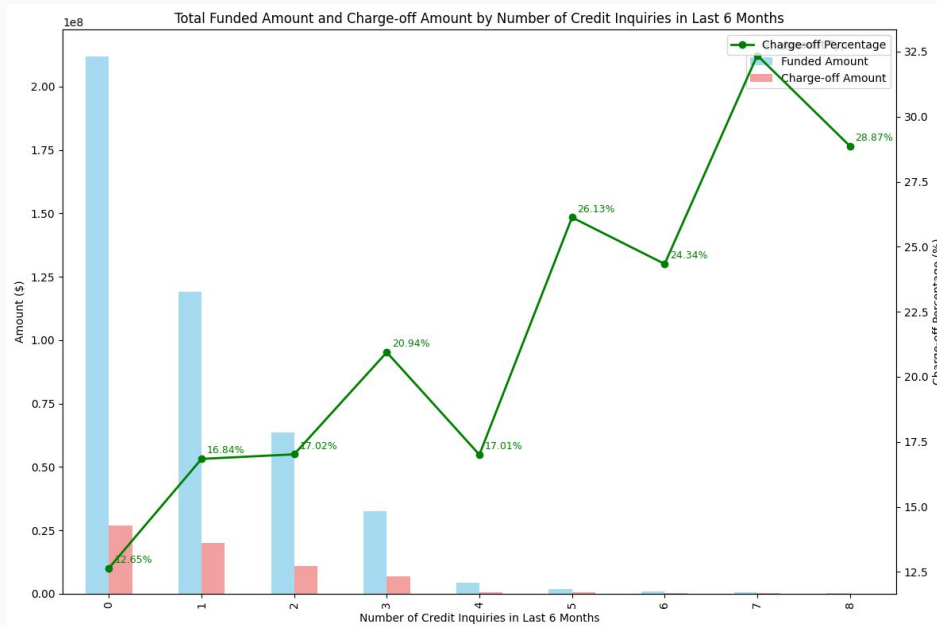
Lower Income Earners Have Higher Charge-Offs

Total Funded Amount and Charge-off Amount by Annual Income Category
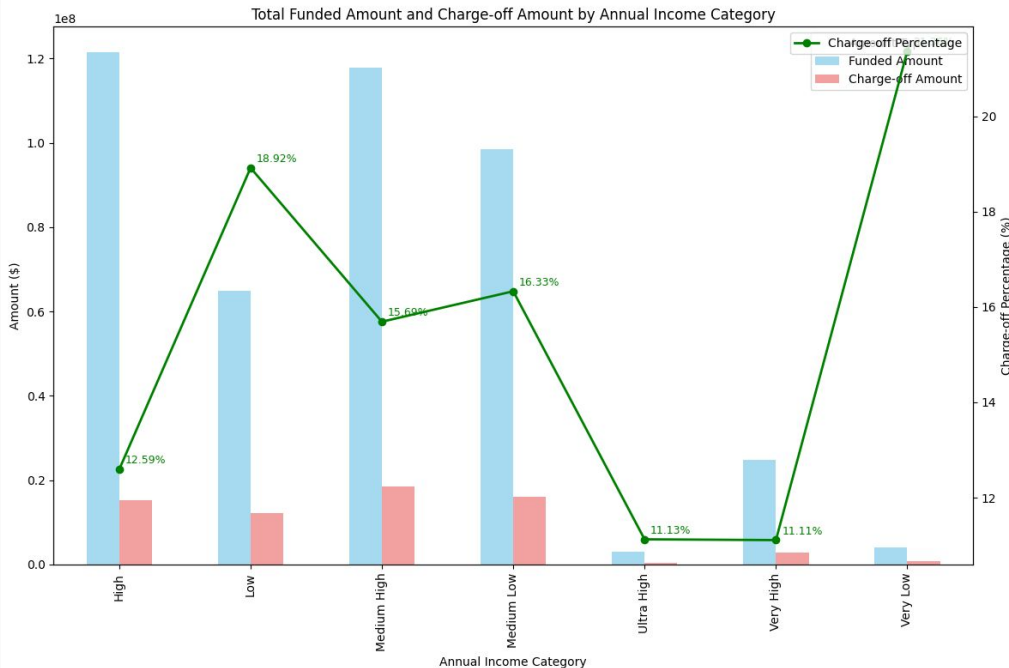
Longer Term Loans Have Higher Charge-Offs

Total Funded Amount and Charge-off Amount by Loan Term
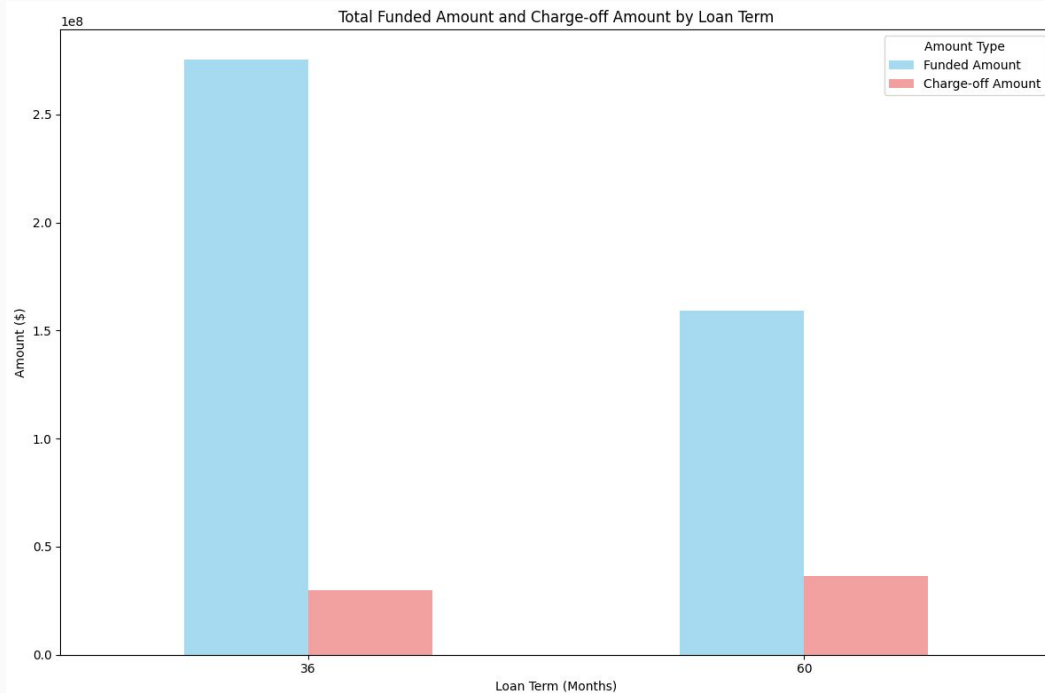
Highest Charge-Offs for Employee with No Work, Followed by Employees Who Worked for 10 Years

Total Funded Amount and Charge-off Amount by Employment Length
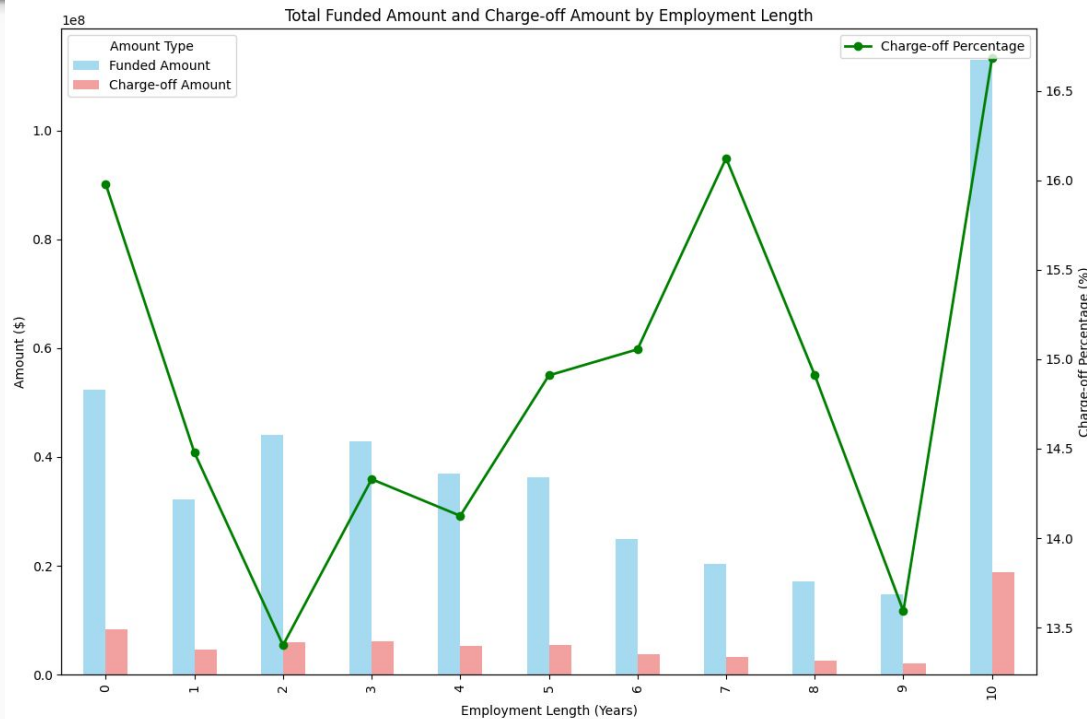
Charge-Offs Highest for Those with "Other" Home Ownership Status

Total Funded Amount and Charge-off Amount by Home Ownership
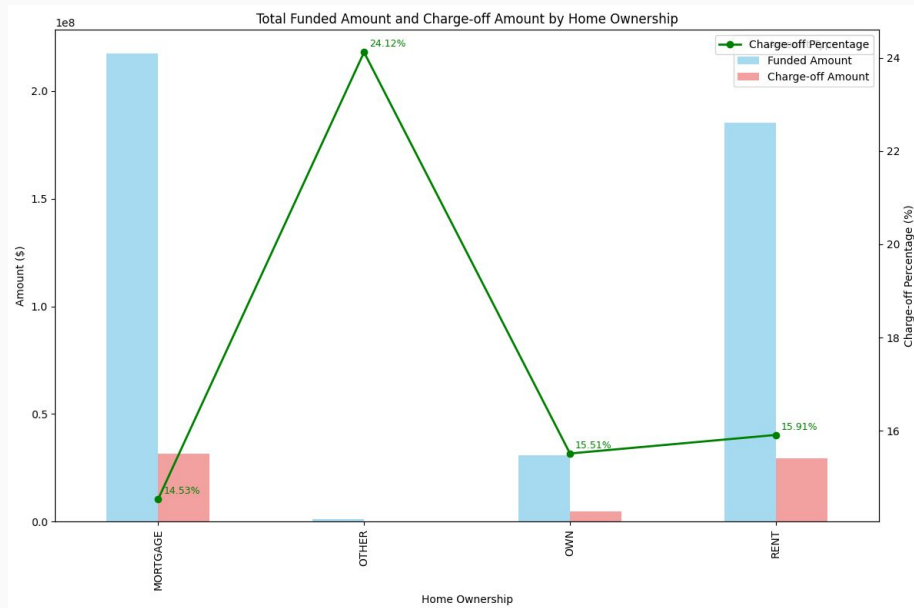
Charge-Offs Highest for Those Who are Verified

Total Funded Amount and Charge-off Amount by Verification Status
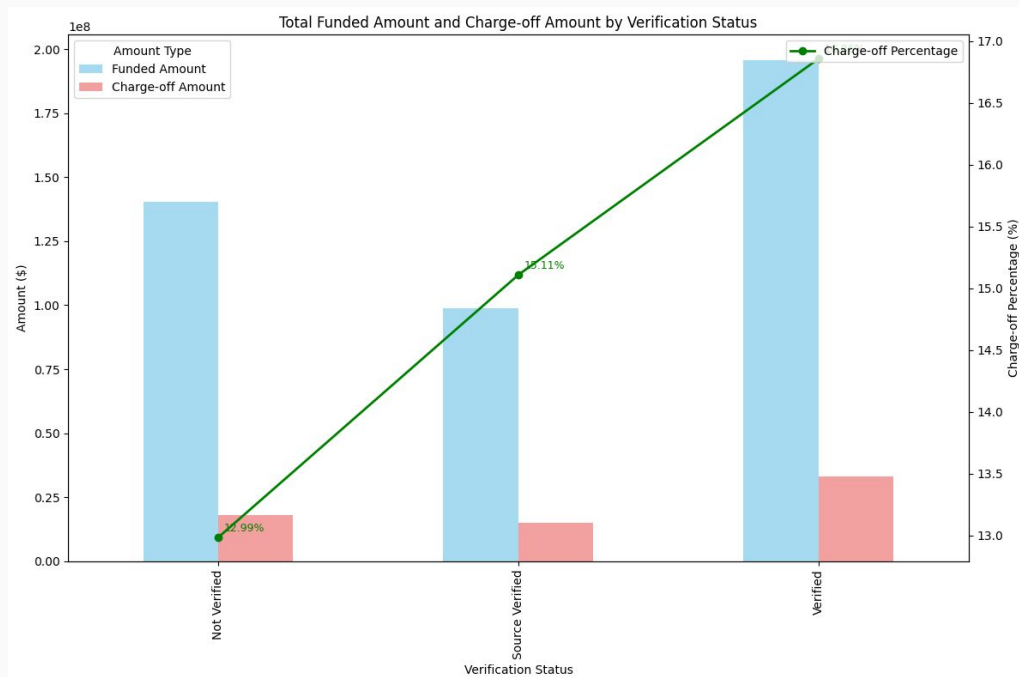
Charge-Offs Highest for Nebraska

Total Funded Amount and Charge-off Amount by Address State
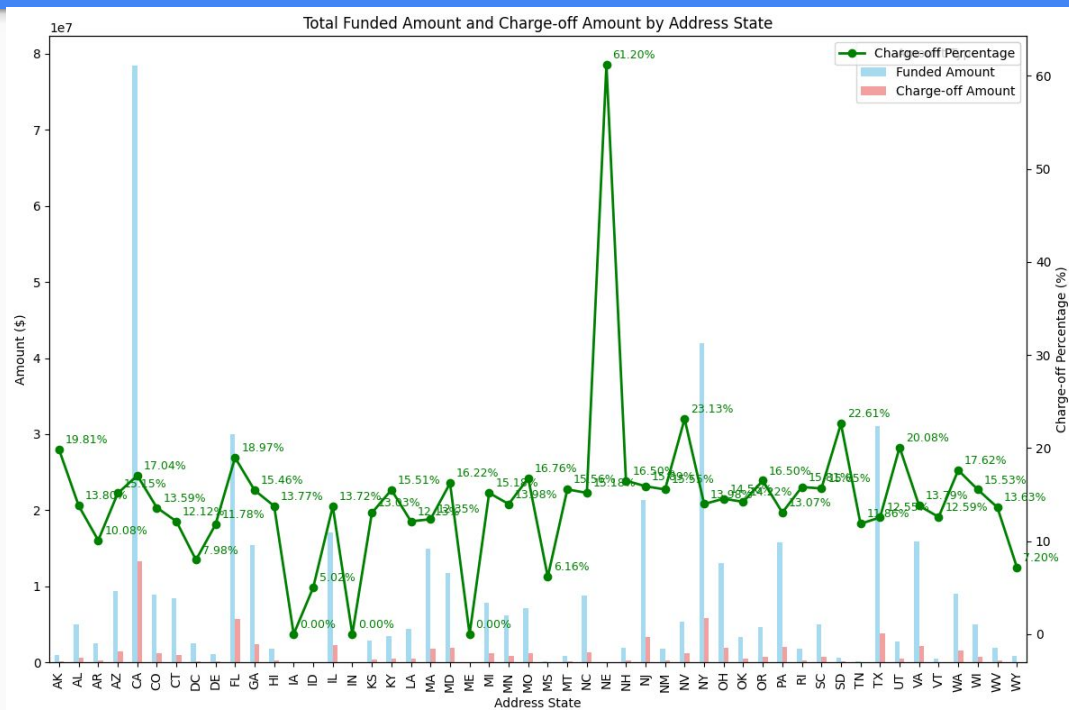
Credit Data Arising After When Loan was Made

Relevant to
- Reducing the Amount of the Loan
- Changing Interest Rate

Heatmap Showing Cross-Correlation Between Credit Factors
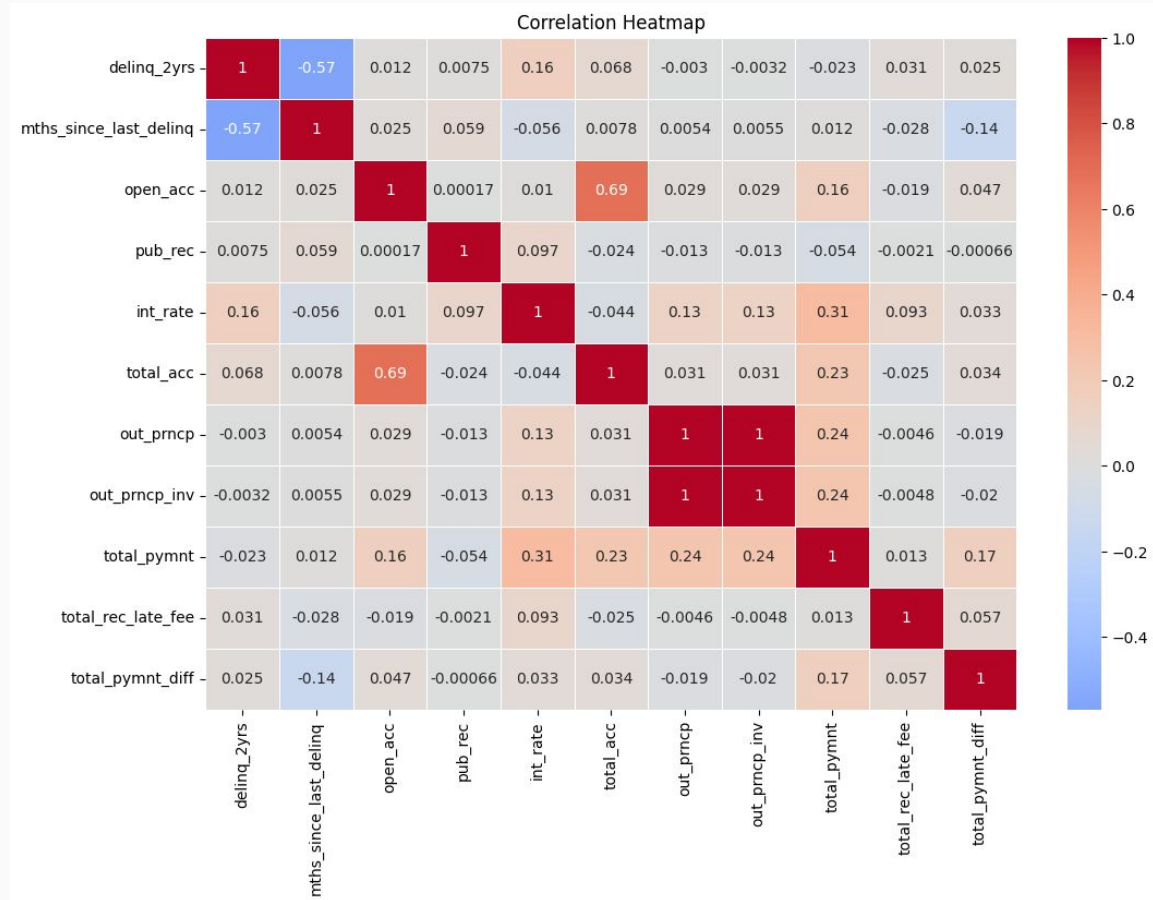
Positive Relationship Between Charge-offs and the Number of 30+ Days Past-Due Incidences of Delinquency

Total Funded Amount and Charge-off Amount by Delinquency in Last 2 Years
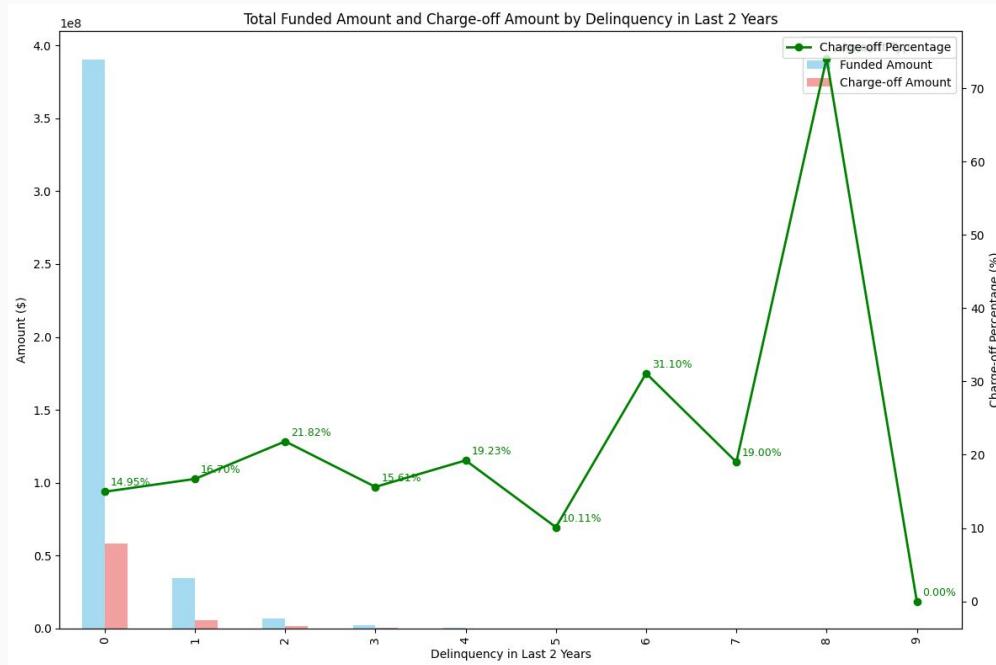
Interest rates increase as charge offs increase

Total Funded Amount and Charge-off Amount by Interest Rate
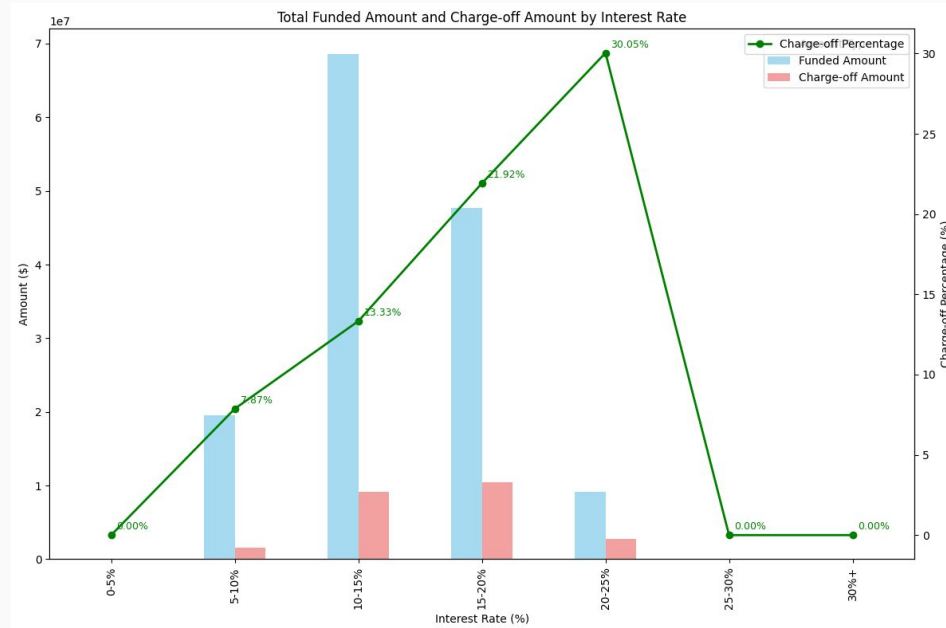
Positive relationship between late fees received to date and charge-offs
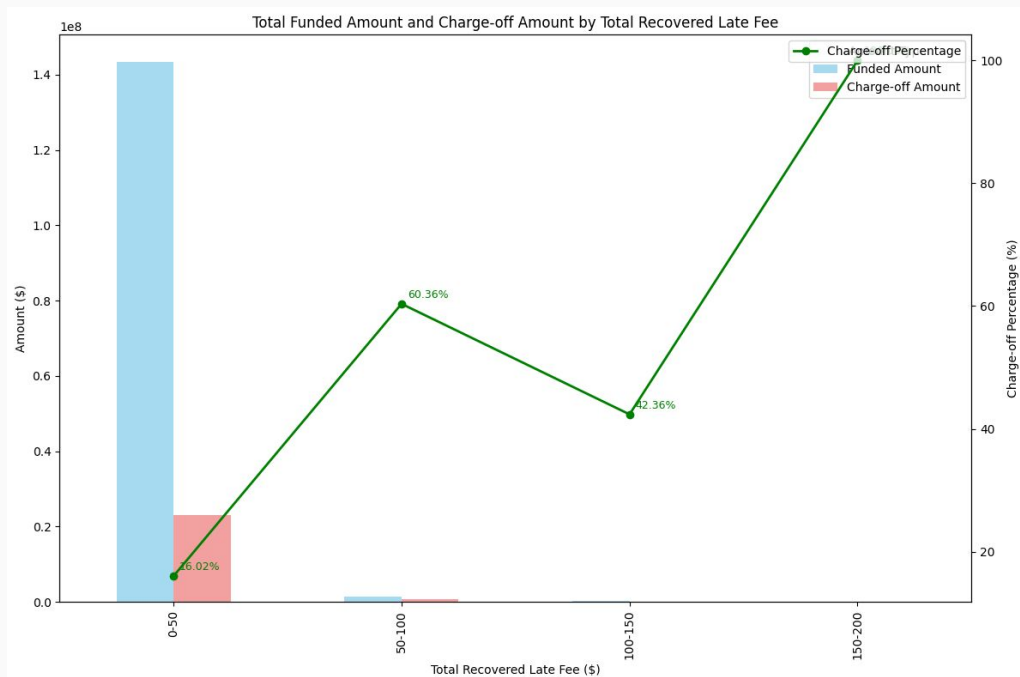
There are higher charge-offs for older credit lines.

Total Funded Amount and Charge-off Amount by Years Since Earliest Credit Line
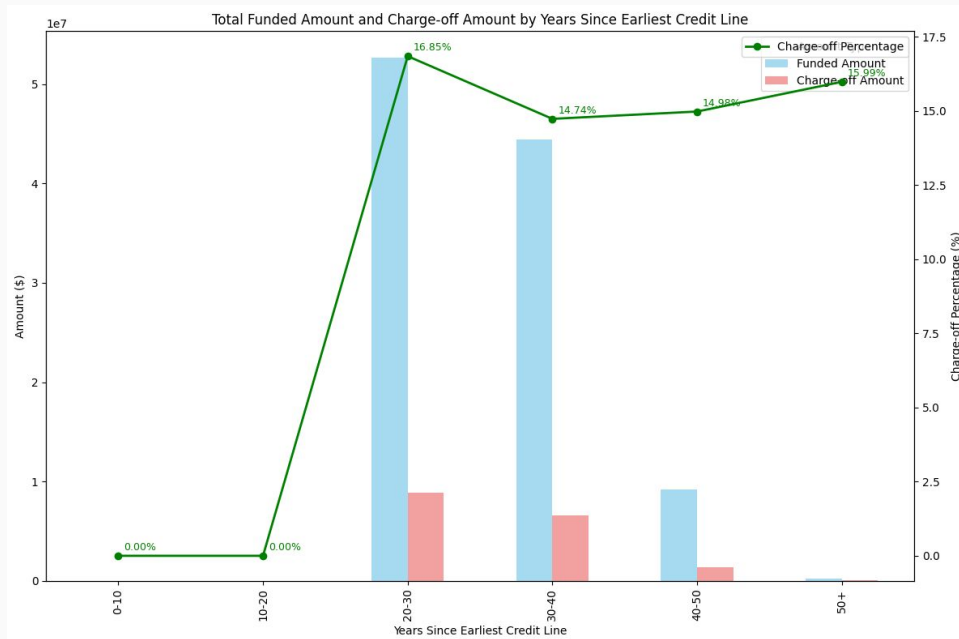
There is a slightly higher number of charge offs when a company has derogatory public records

Total Funded Amount and Charge-off Amount by Number of Public Records
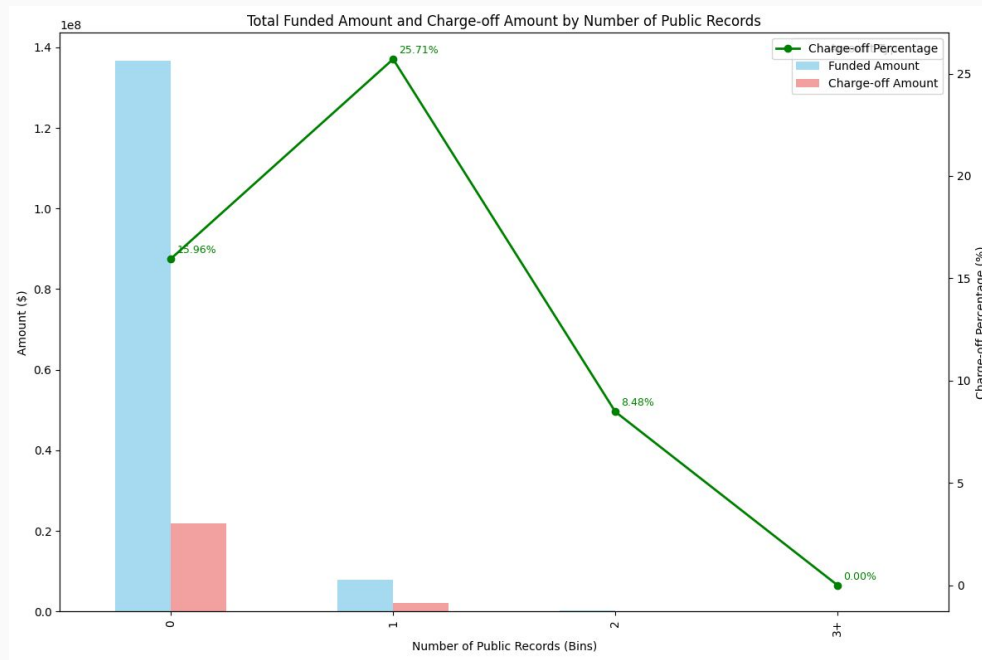
No clear relationship between the months since the last delinquency and charge-offs

Total Funded Amount and Charge-off Amount by Months Since Last Delinquency

No clear relationship between the number of open credit lines in the borrower's credit file and charge-offs

Total Funded Amount and Charge-off Amount by Number of Open Accounts
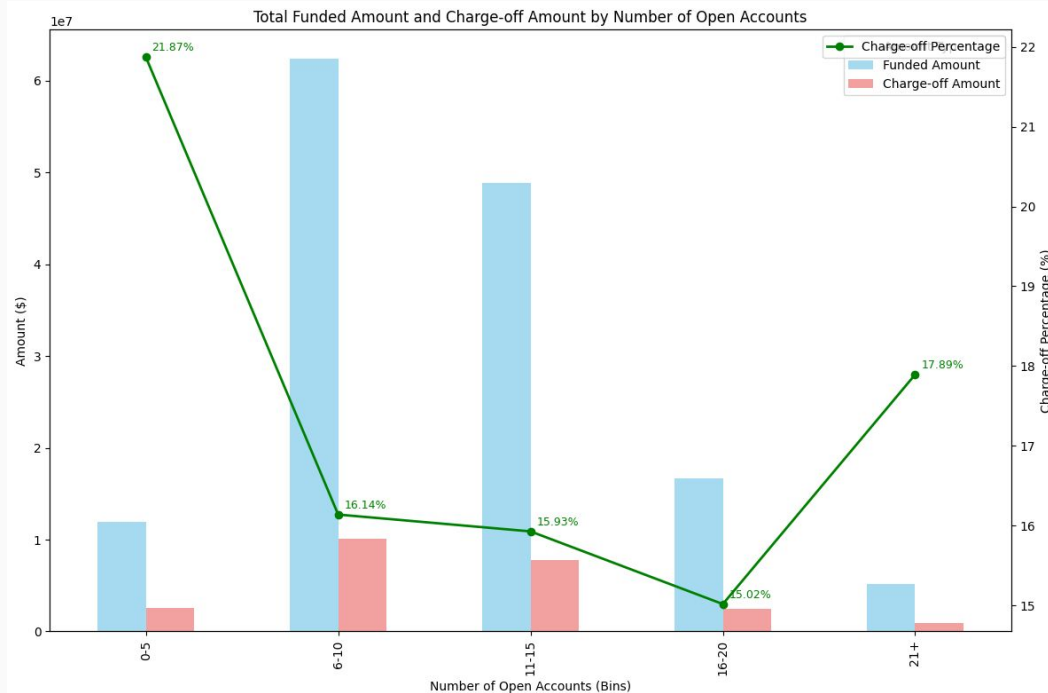
Conclusion

The Primary Factors for Better Decision-Making on Lending:

- **Grade** & **Sub-Grades:** Correlated with the risk of credit loss except for the final risk bucket which should be analysed
- **Inquiries in Last 6 Months (inq_last_6mths):** A surprising metric which shows that multiple credit inquiries may signal credit-seeking behavior and potential risk.
- **Home Ownership:** Homeowners may present lower risk.
- **Debt-to-Income Ratio (dti):** Higher DTI may indicate potential repayment issues.

Factors for Better Decision-Making on Loan Monitoring:

- **Delinquencies in Last 2 Years (delinq_2yrs):** An increase in the number of delinquencies indicate higher risk.
- **Earliest Credit Line (earliest_cr_line):** Regular credit reviews should be made on older accounts to ensure their credit limits are still correct
- **Interest Rate (int_rate):** Higher interest rates may correlate with higher risk loans due to the presence of penalty interest clauses
- **Total Recovered Late Fees (total_rec_late_fee):** Monitor late fees as an indicator of repayment behavior.

Appendix - Chi-Squared & P-Value Tests

Chi-square test p-value for grade: 5.0592742040116925e-127
Chi-square test p-value for sub_grade: 2.6788755805703413e-117
Chi-square test p-value for home_ownership: 9.935558557849398e-06
Chi-square test p-value for verification_status: 9.310717832316028e-25
Chi-square test p-value for issue_d: 4.657475601949828e-95
Chi-square test p-value for loan_status: 0.0
Chi-square test p-value for purpose: 6.061014196118496e-31
Chi-square test p-value for addr_state: 0.018510522404149146
Chi-square test p-value for earliest_cr_line: 0.21006850586093553
Chi-square test p-value for revol_util: 0.05832291177179975
Chi-square test p-value for next_pymnt_d: 1.0
Chi-square test p-value for application_type: 1.0
Chi-square test p-value for season: 4.877540322537689e-16
Chi-square test p-value for funded_amount_categories: 9.897358821118917e-46
Chi-square test p-value for annual_income_category: 9.612025809495913e-18
Chi-square test p-value for dti_bins: 3.4125197567086015e-06
Chi-square test p-value for delinq_bin: 0.04345261880181839
Chi-square test p-value for open_acc_bin: 9.347621253584592e-06
Chi-square test p-value for total_acc_bin: 0.00021603043980848716
Chi-square test p-value for pub_rec_bin: 6.068774355502321e-08
Chi-square test p-value for earliest_cr_bin: 0.0983373968265565
Chi-square test p-value for int_rate_bin: 1.4856173671186012e-154
Chi-square test p-value for total_rec_late_fee_bin: 9.926365436724708e-24

Appendix - Chi-Squared & P-Value Tests

Correlation between loan_status and term: 0.14146641942425675, p-value: 1.2233154484638262e-63
Correlation between loan_status and int_rate: 0.16523468338739422, p-value: 1.8770218223142662e-86
Correlation between loan_status and installment: 0.027614187261204865, p-value: 0.0010700921006587508
Correlation between loan_status and annual_inc: -0.05134635944778404, p-value: 1.1591468089352152e-09
Correlation between loan_status and dti: 0.03880633228272387, p-value: 4.261687994334477e-06
Correlation between loan_status and inq_last_6mths: 0.06211555758984623, p-value: 1.7786155378632959e-13

Thank you for your attention