

# LENDING CLUB CASE STUDY

---

GROUP MEMBERS:

SIVARAMAKRISHNA BANDARU

SRAVANI KANUMURI

# Abstract

- Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.
- Borrowers can easily access lower interest rate loans through a fast online interface.
- The objective of analysis is to use the information about past loan applicants and find whether they 'defaulted' or not.



# Problem solving methodology:

**Data Cleaning :** Removing the null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.

**Data Understanding:** Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses.

**Univariate Analysis:** Analyzing each column, plotting the distributions of each column.

## Data Analysis

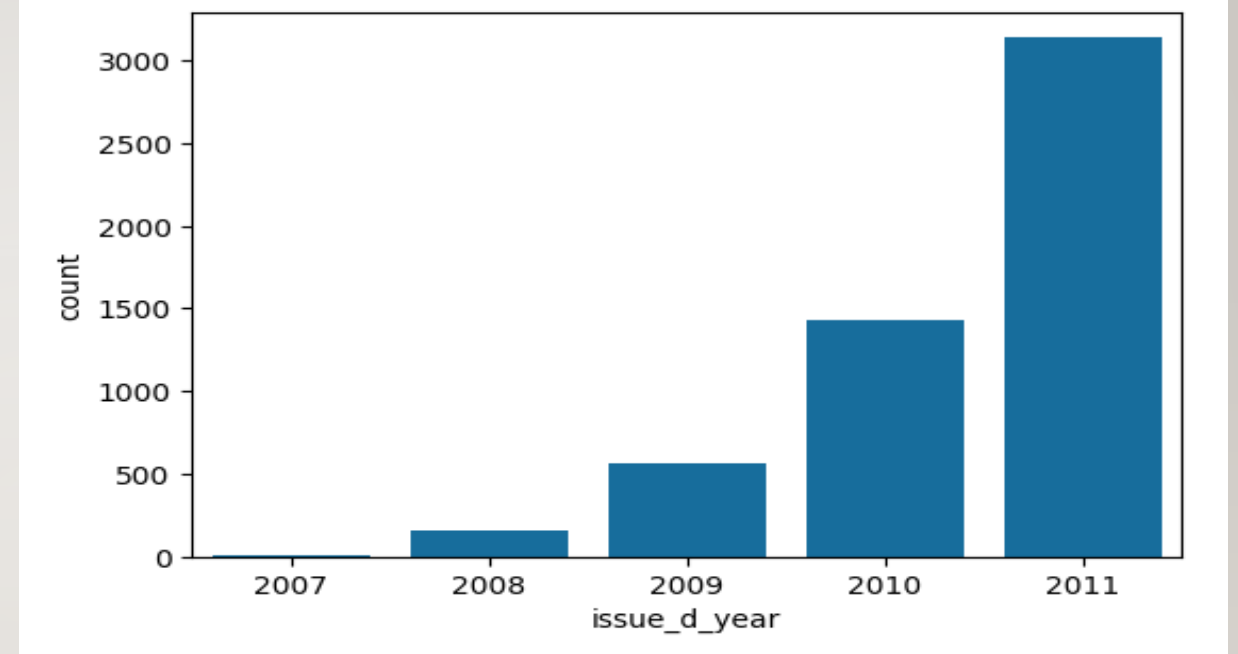
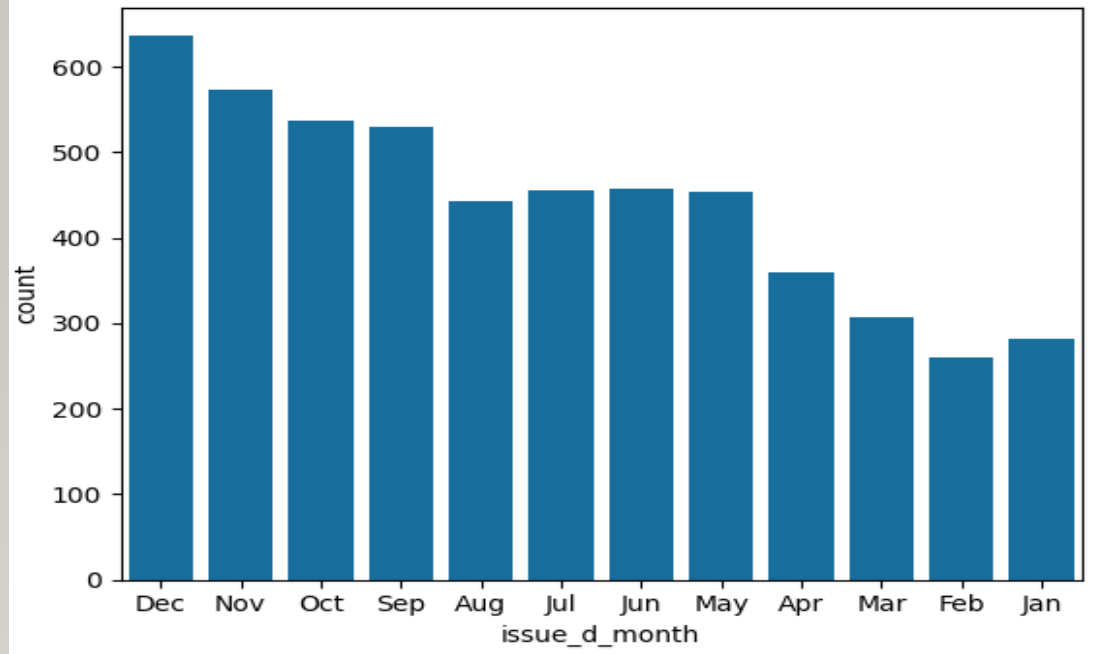
**Segmented Univariate:** Analysis Analyzing the continuous data columns with respect to the categorical column.

**Bivariate Analysis:** Analyzing the two-variable behavior like term and loan status with respect to loan amount.

**Recommendations:** Analyzing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

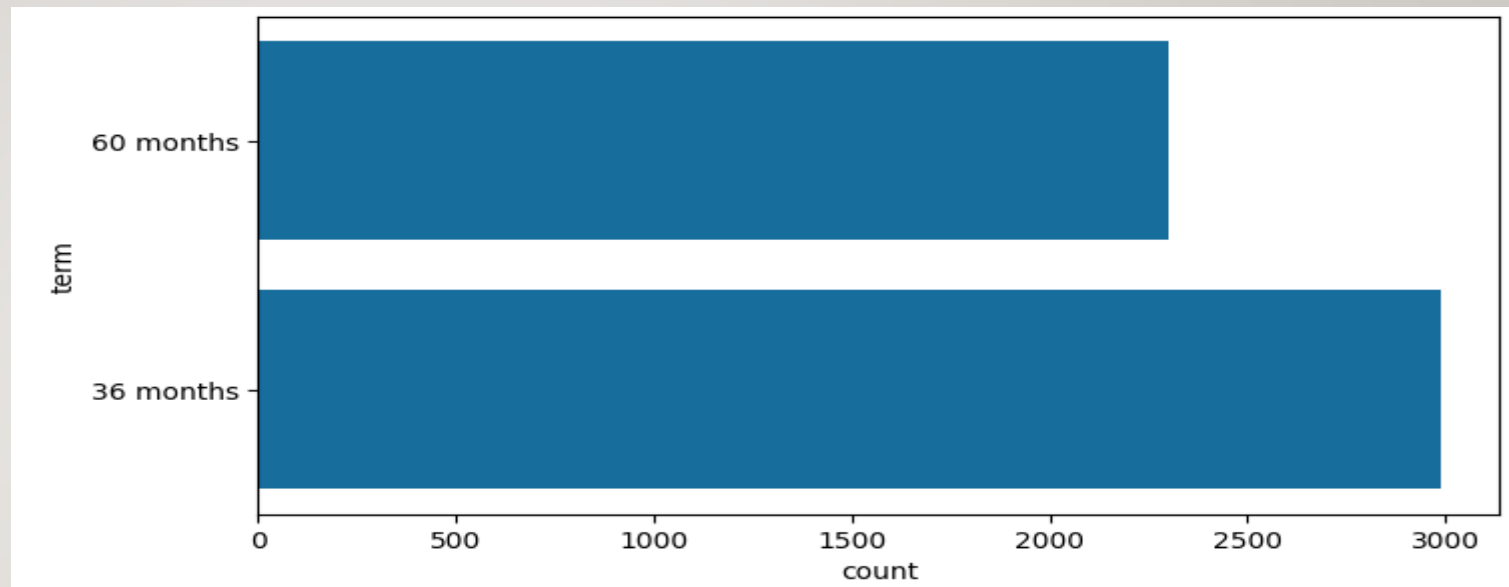
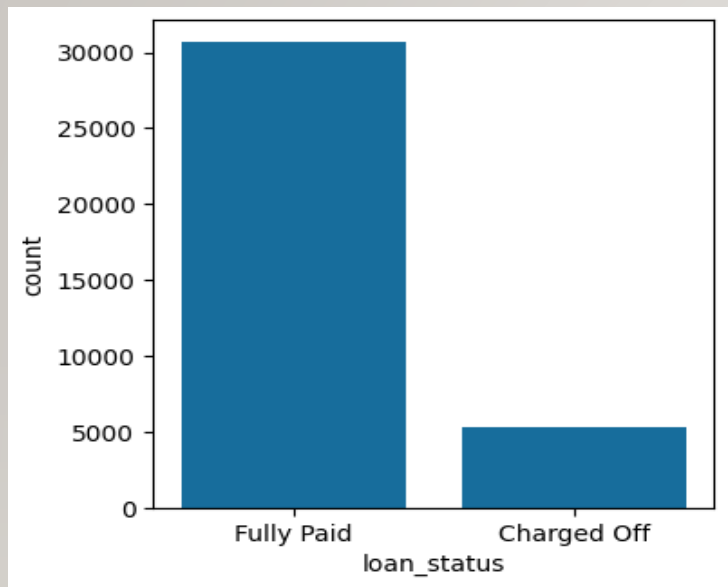


# Analysis



- Lending club has really expanded year by year, the number of loan issued are doubled every year.
- Also, the issued month of loans is also increasing from January to December. In the final quarter of year there are more loans issued this could be because of vacation and Christmas.

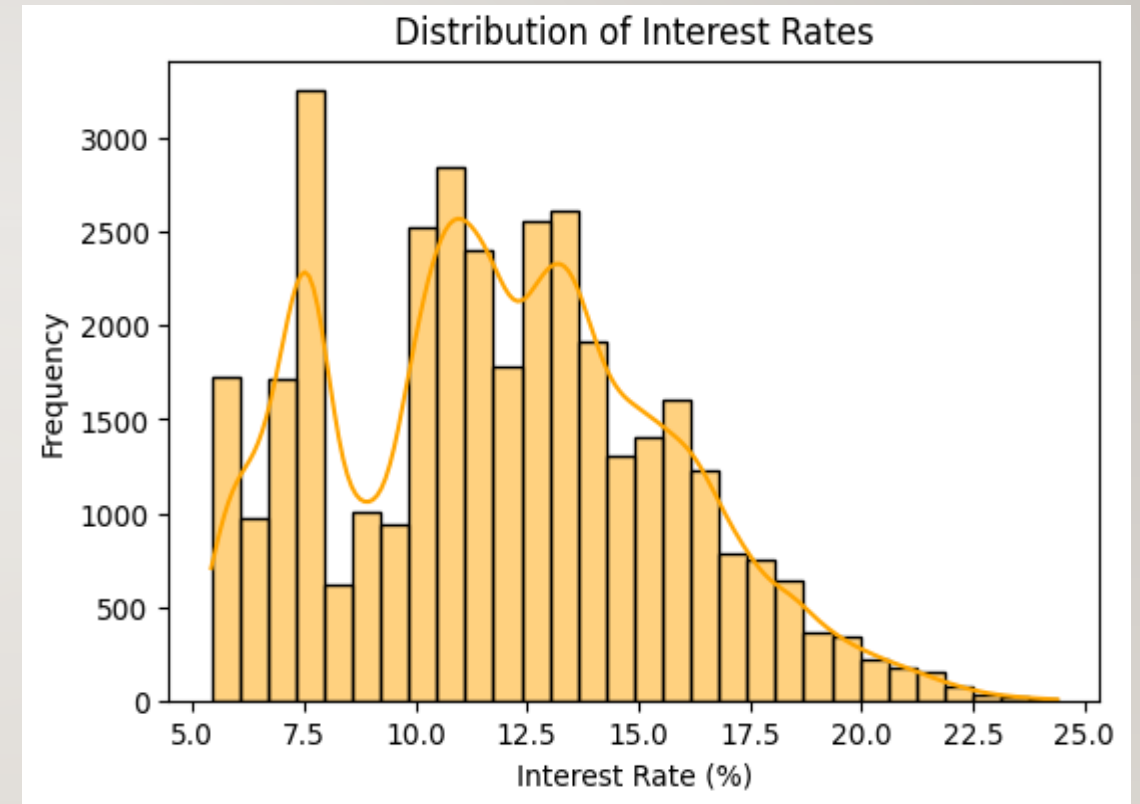
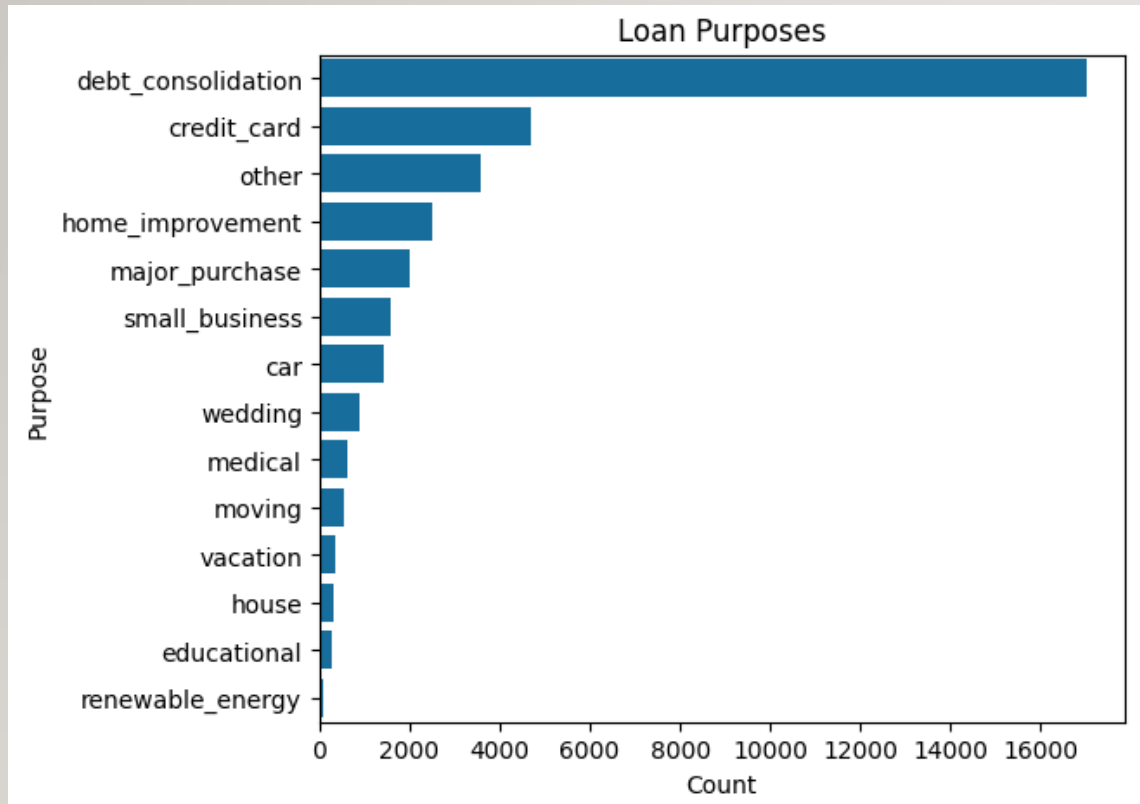
# Analysis



- There are only two loan terms 36 and 60 months. Around 75% borrowers took loans with 36 months term.
- The charged off borrowers are around 15% and fully paid is around 85% in the given data set.

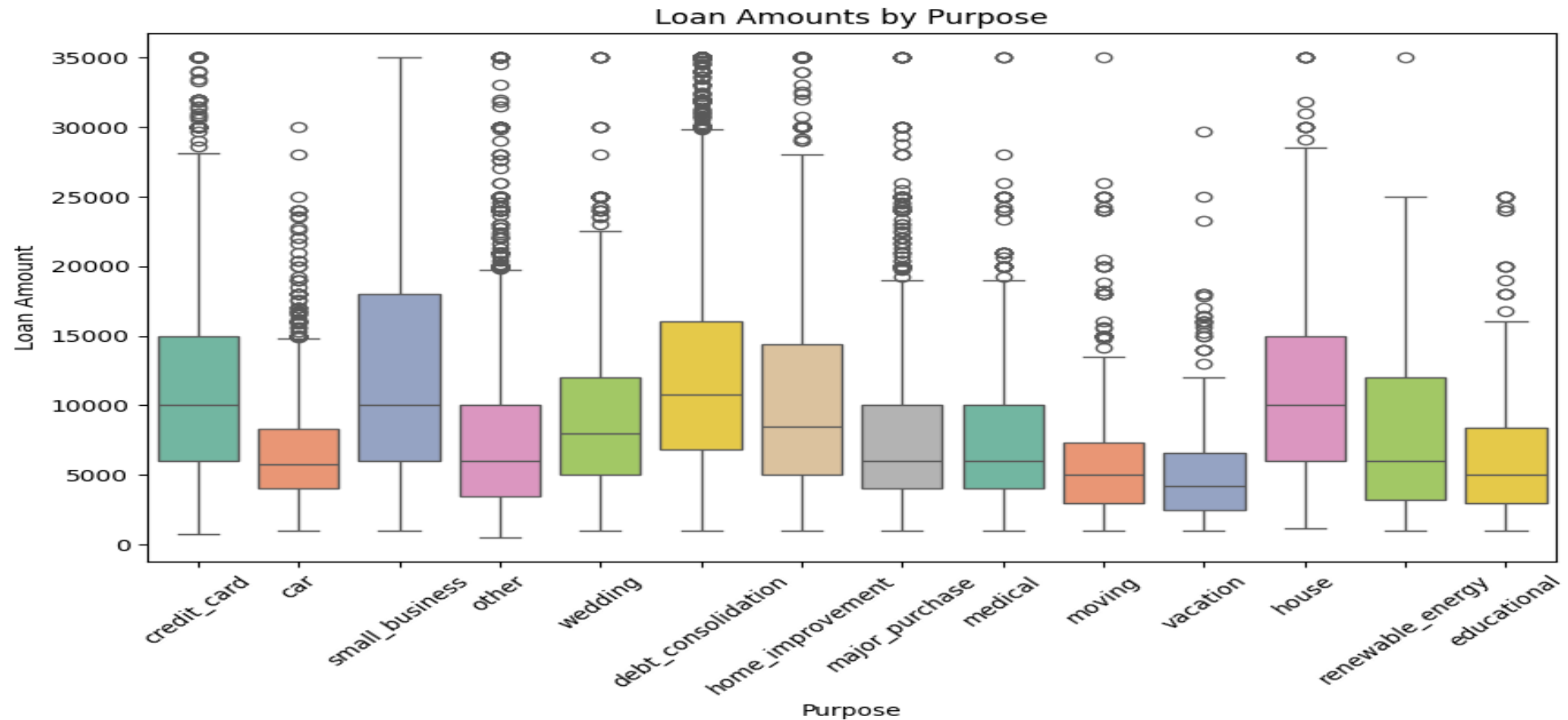


## Loan Purposes and the Distribution of Rate of Interest diagrams.

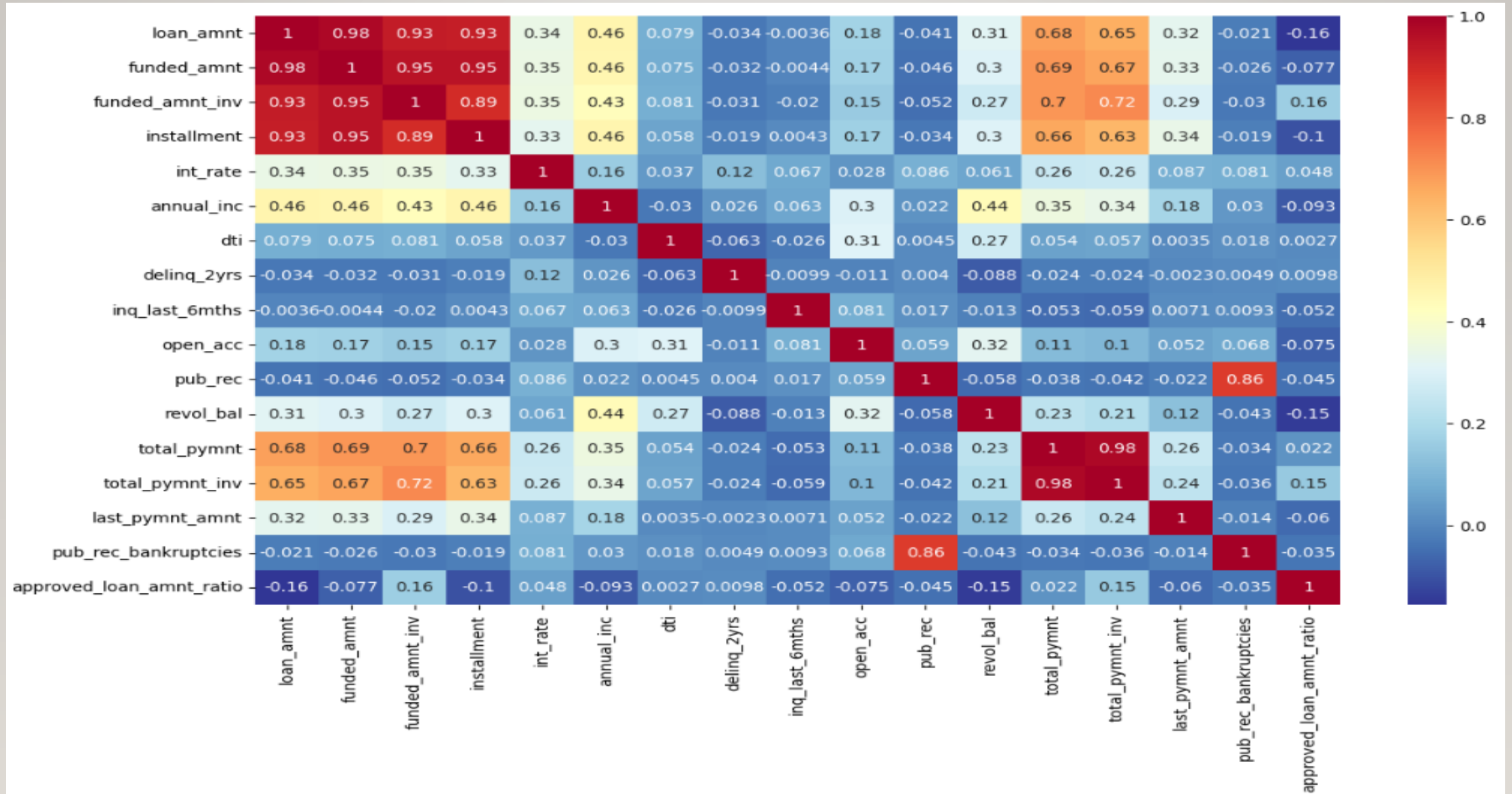


## Bivariate Analysis:

Relationships: Loan Amount vs. Purpose, Interest Rate vs. Loan Status, Income vs. Loan Amount



# Correlation Matrix





## Key Takeaways from this Correlation Matrix:

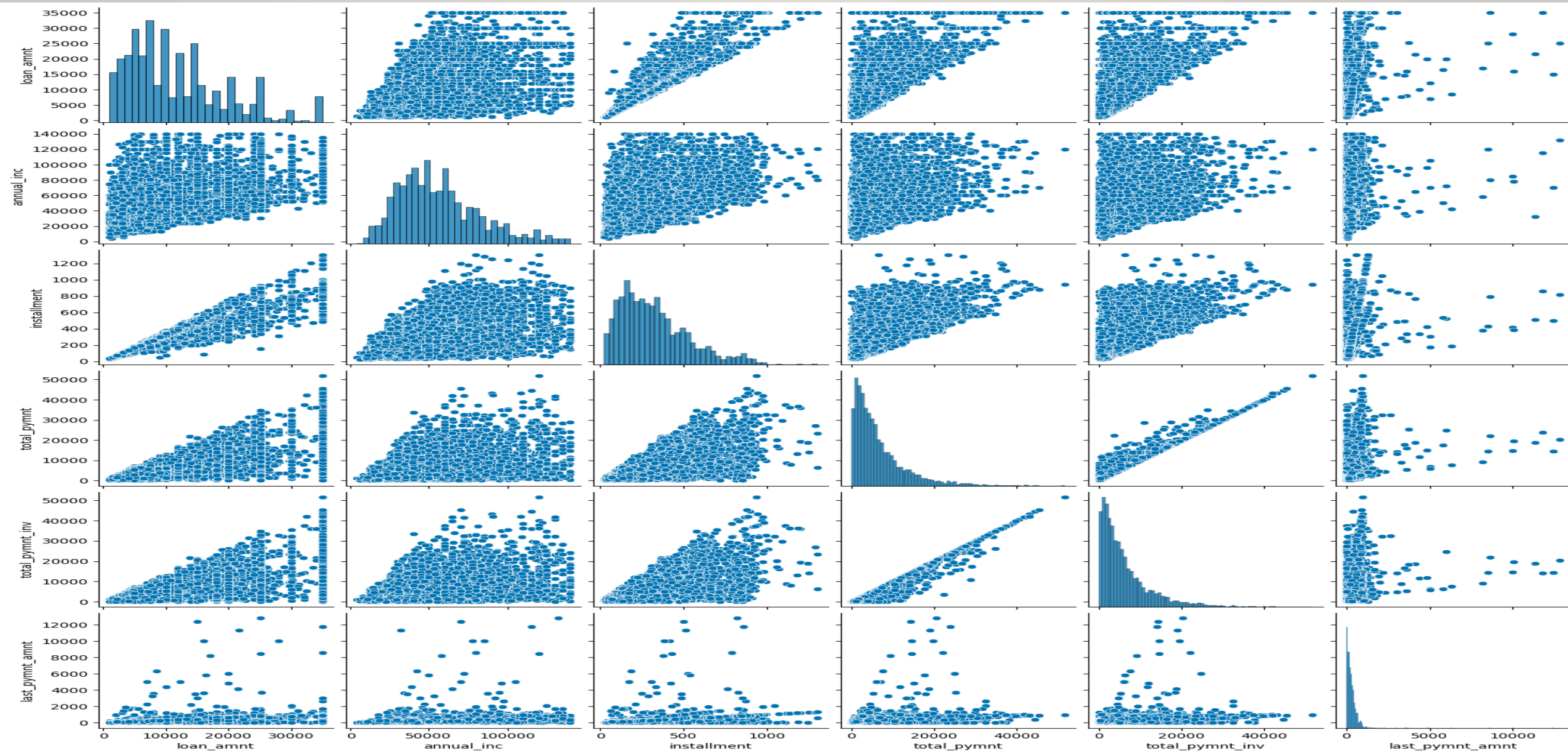
- **Strong Positive Correlations:** You can see strong positive correlations (red squares with values close to 1) between variables like:
  - loan\_amnt, funded\_amnt, funded\_amnt\_inv and installment (which makes sense, as the amount of the loan is closely related to the amount funded).
  - total\_pymnt and total\_pymnt\_inv (total payment received is correlated with the total payment invested).
- **Some Negative Correlations:** There are some negative correlations (blue squares), though they are generally weaker than the positive ones. For example, int\_rate (interest rate) has a slight negative correlation with last\_pymnt\_amnt (last payment amount).

This correlation matrix helps to understand the relationships between different factors in a loan dataset. It can be useful for tasks like:

- **Feature Selection:** Identifying highly correlated variables that might be redundant in a machine learning model.
- **Data Analysis:** Understanding the underlying relationships in your data.
- **Risk Assessment:** Identifying factors that might be associated with loan defaults (though this matrix alone doesn't provide that information directly).



# Multivariate analysis





# Summary and Business Insights

## Average Loan Amount by Purpose: purpose

- small\_business: 12763.262389
- debt\_consolidation: 12218.438768
- House: 11853.293413
- credit\_card: 11280.726168
- home\_improvement: 10472.712808
- Wedding: 9556.550343
- renewable\_energy: 8123.947368
- Medical: 7794.075829
- major\_purchase: 7738.958952
- Other: 7670.578801
- Car: 6652.603075
- Educational: 6565.034364
- Moving: 6037.662338
- Vacation: 5294.273743



# Conclusions

- Lending club should reduce the high interest loans for 60 months tenure, they are prone to loan default.
- Grades are good metric for detecting defaulters. Lending club should examine more information from borrowers before issuing loans to Low grade (G to A).
- Lending Club should control their number of loan issues to borrowers who are from CA, FL and NY to make profits.
- Small business loans are defaulted more. Lending club should stop/reduce issuing the loans to them.
- Borrowers with mortgage home ownership are taking higher loans and defaulting the approved loans. Lending club should stop giving loans to this category when loan amount requested is more than 12000.
- People with a greater number of public derogatory records are having more chance of filing a bankruptcy. Lending club should make sure there are no public derogatory records for borrower.

