

# Lending Club Case Study

---

GROUP –

1. ADITYA SONI

2. AKANKSHA PODUTWAR

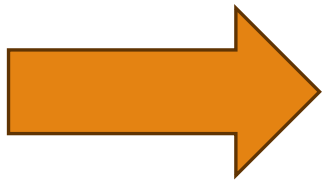
# Objective

---

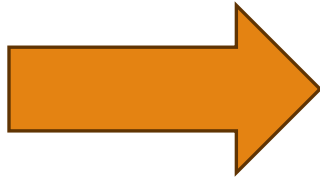
- The objective of this case study is to conduct Exploratory Data Analysis (EDA) on a real-world problem, with the aim of deriving meaningful insights and enhancing the understanding of the underlying data patterns.
- Developing foundational knowledge of risk analytics in the banking and financial services sectors.
- Demonstrating how data is leveraged to reduce financial losses in the lending process.
- Enhancing our comprehension of data visualization techniques and the appropriate charts to use for analyzing real-world data.

# Problem Solving Methodology

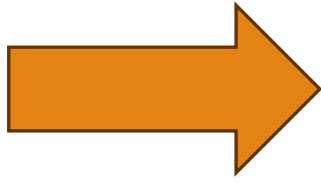
---



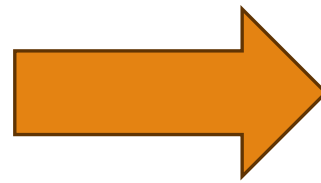
Data Sourcing



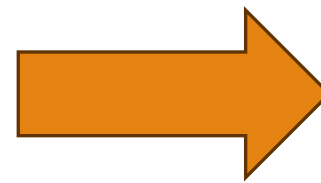
Data Cleaning



Univariate Analysis



Bivariate Analysis



Derived Metrics

**1.Data Sourcing:** Identifying and collecting relevant data.

**2.Data Cleaning:** Processing and refining the data to ensure accuracy and consistency.

**3.Univariate Analysis:** Examining individual variables to understand their distributions and characteristics.

**4.Bivariate Analysis:** Analyzing the relationships between pairs of variables to identify correlations and patterns.

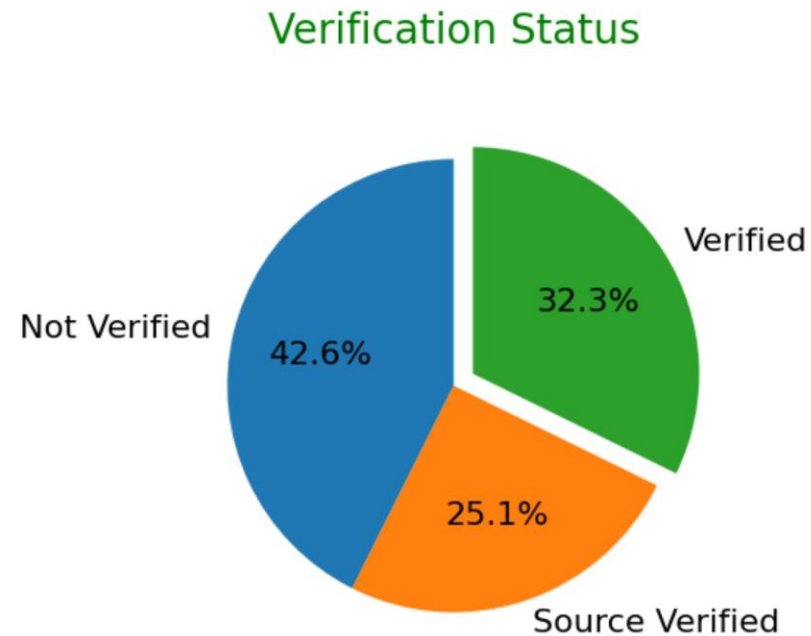
**5.Derived Metrics:** Creating additional metrics from the data to enhance insights and inform further analysis.

---

We have conducted Univariate Analysis on the following variables –

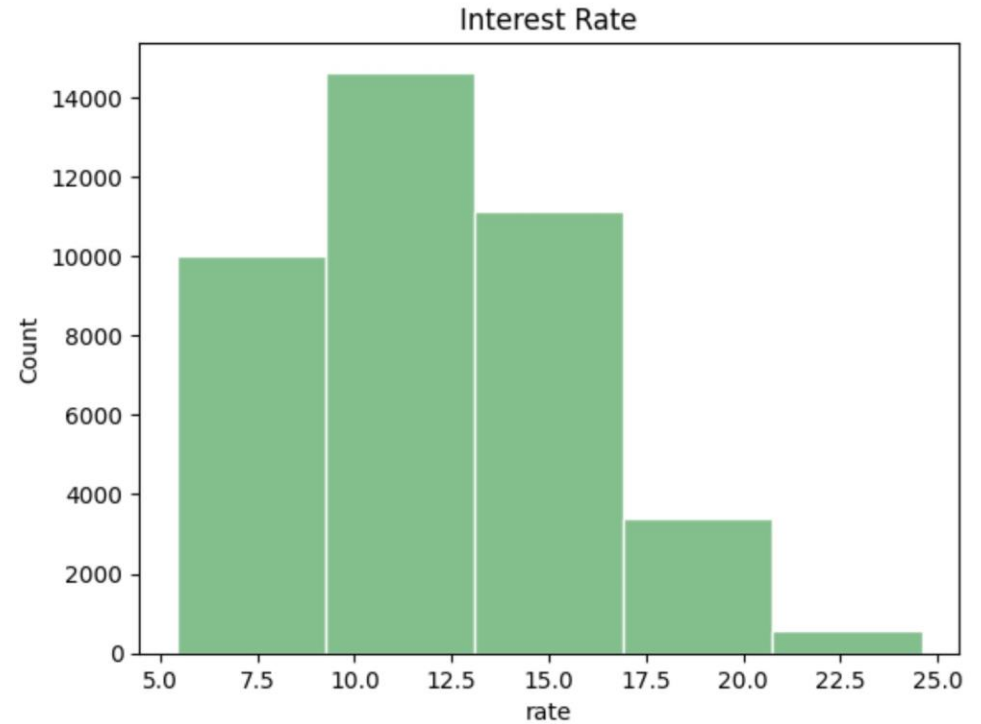
1. Loan Status
2. Interest Rate
3. Loan Issued
4. Purpose of Loan
5. Grade

# Verification Status



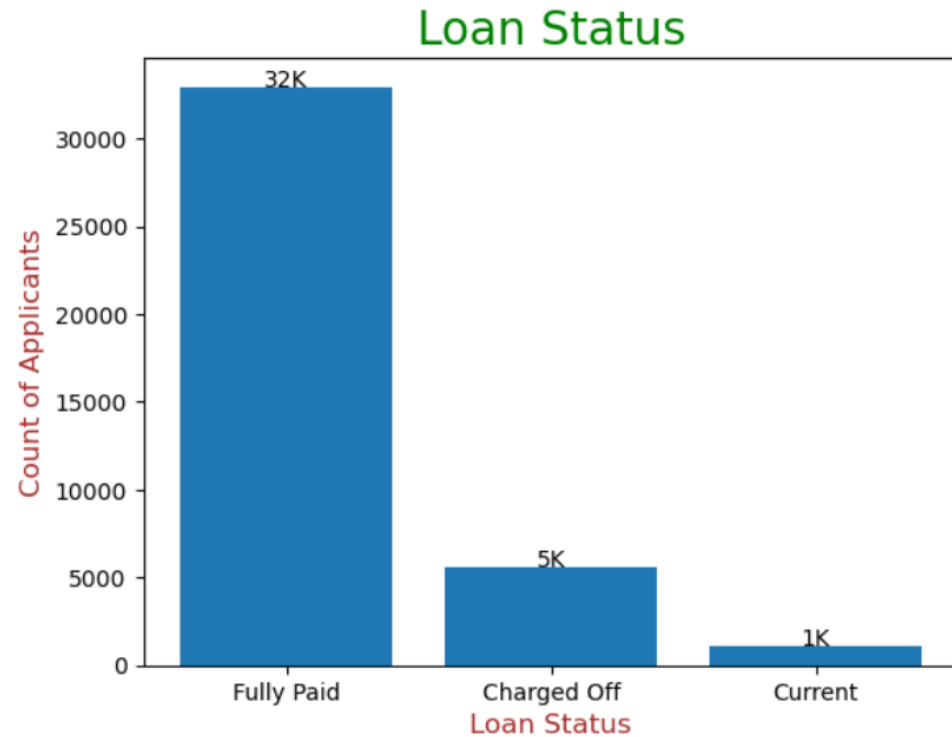
- 32.3% of loan applicants are verified, 42.6% are not verified and 25.1% are source verified.

# Interest Rate



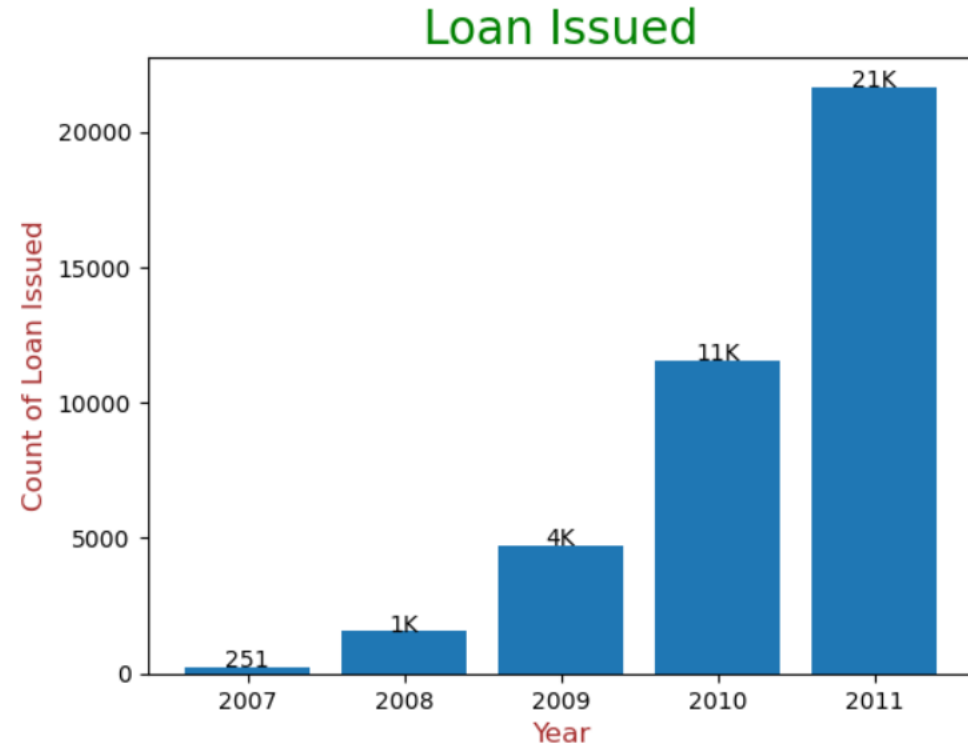
- Loans with interest rate 10% to 12% are borrowed with higher frequency.

# Loan Status



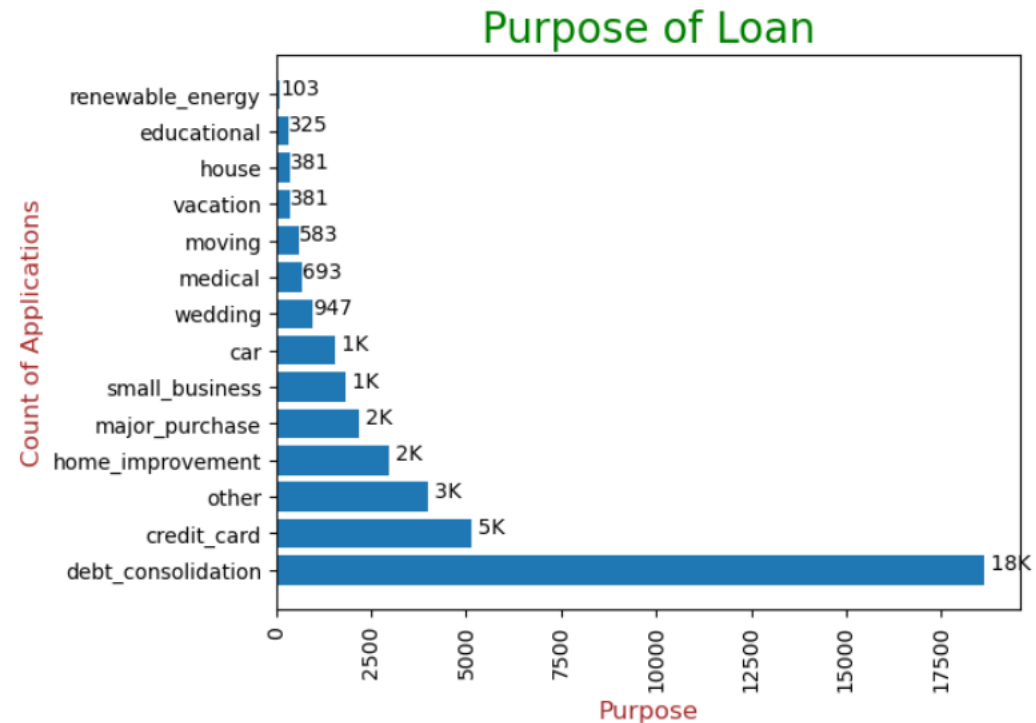
- The Fully Paid borrowers are 83%, charged off are 14% and Current are 3%.
- Majority of borrowers (83%) successfully repaid their loans in full.

# Loan Issued



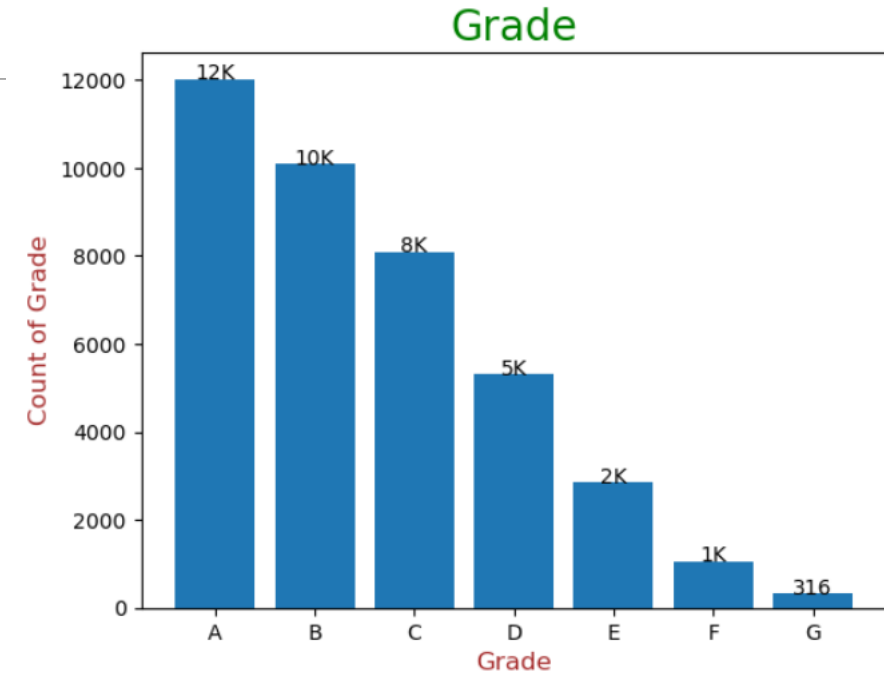
- Loan Issued in Year 2007 is 0.6%, 2008 is 4%, 2009 is 12%, 2010 is 29% and in 2011 is 55%.
- Trend reflects a growing demand for loans.
- Most substantial growth occurred between 2010 and 2011, with the percentage of loans issued increasing from 12% to 55%.

# Purpose of Loan



- Primary reason for loan issuance is debt consolidation, nearly 47% of all loans.
- Followed by credit card (13%), other (10%), home improvement (7%).

# Grade



- Loans classified under grades A, B, and C have a higher frequency compared to loans in other grades.

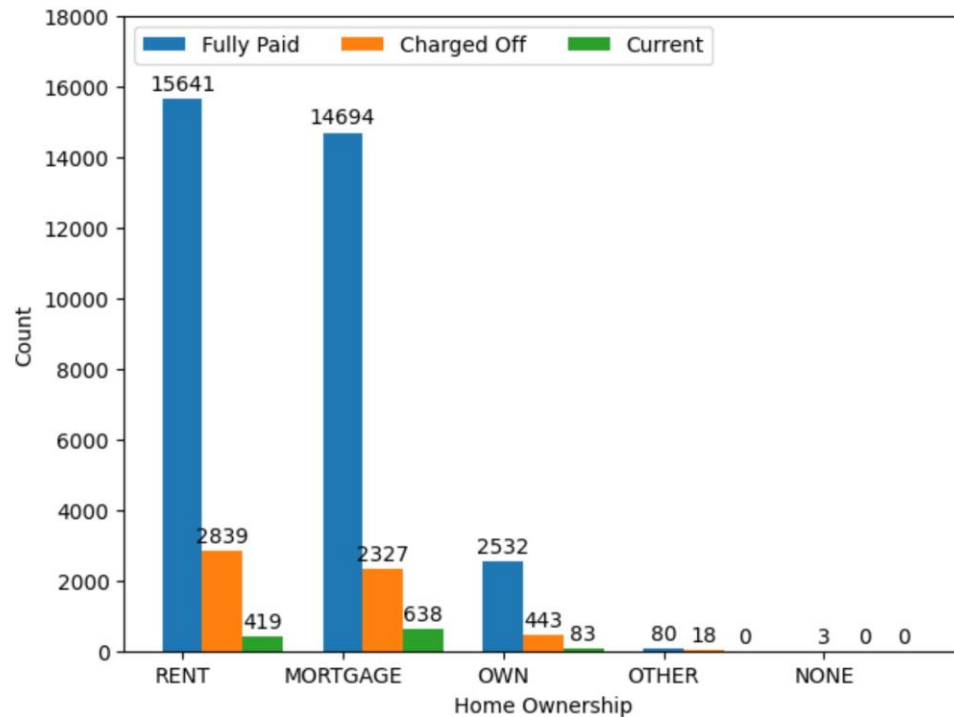
---

We have conducted bivariate analysis on the following variables –

1. Home Ownership and Loan Status
2. Loan Amount and Loan Status
3. Correlation between Loan Amount, Installment, Annual Income, Funded Amount, Funded Amount Invested
4. Loan Status and Grade
5. Subgrade and Grade
6. Address State and Loan Status

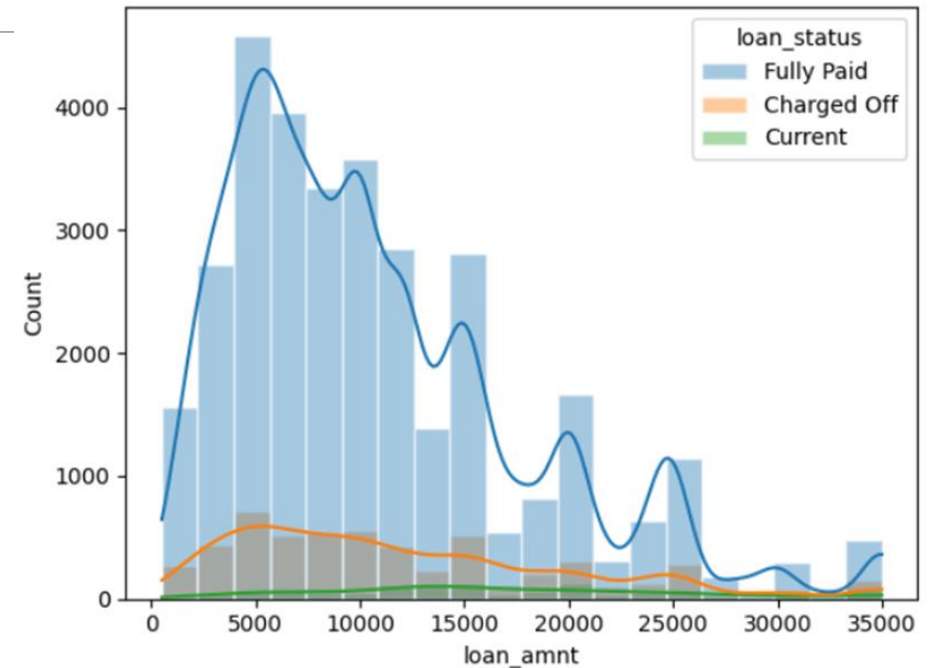


# Home Ownership and Loan Status



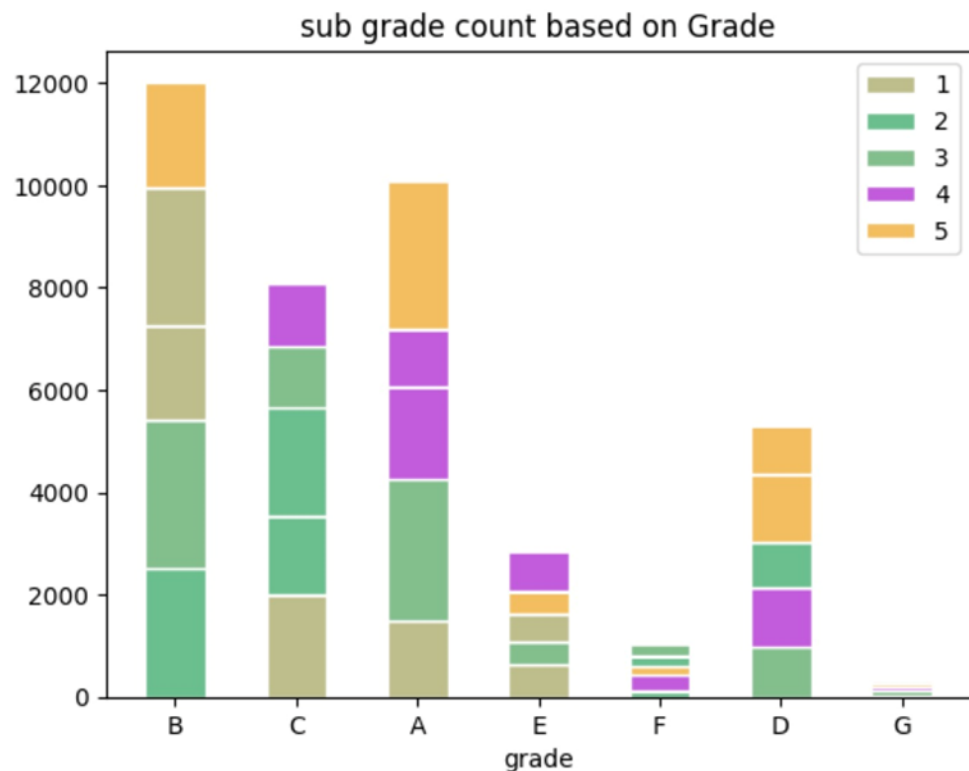
- The majority of clients do not own property and are either renting or have a mortgage, which may increase their likelihood of defaulting on their loans.

# Loan Amount and Loan Status



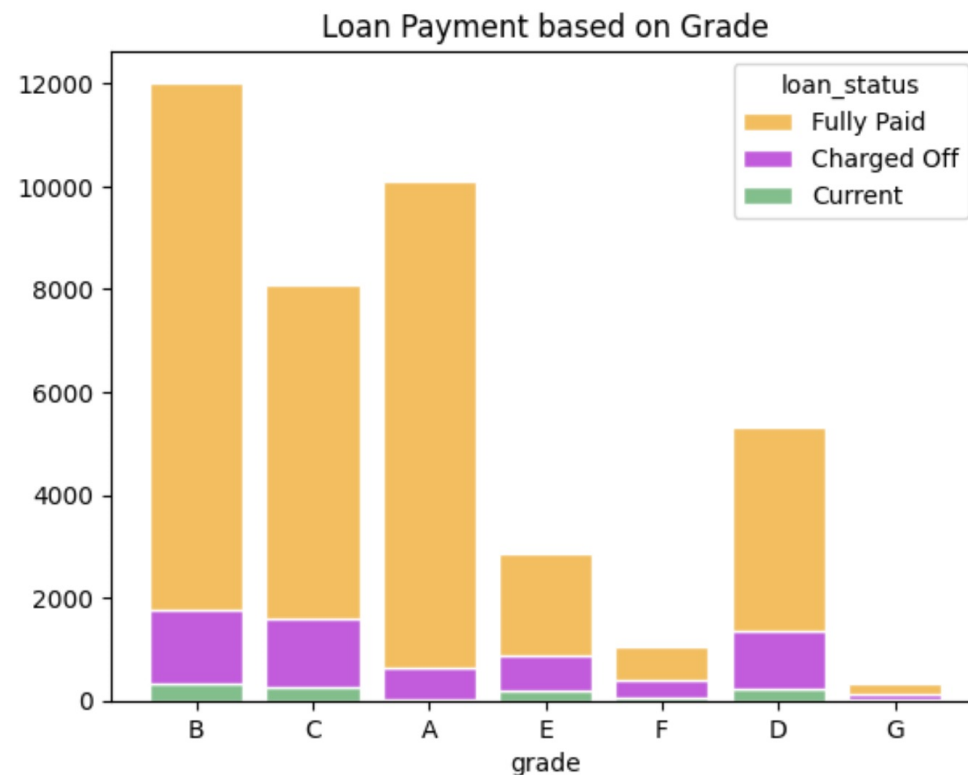
- Majority of clients have borrowed loan in range of 5000-15000 and fully paid.

# Subgrade and Grade



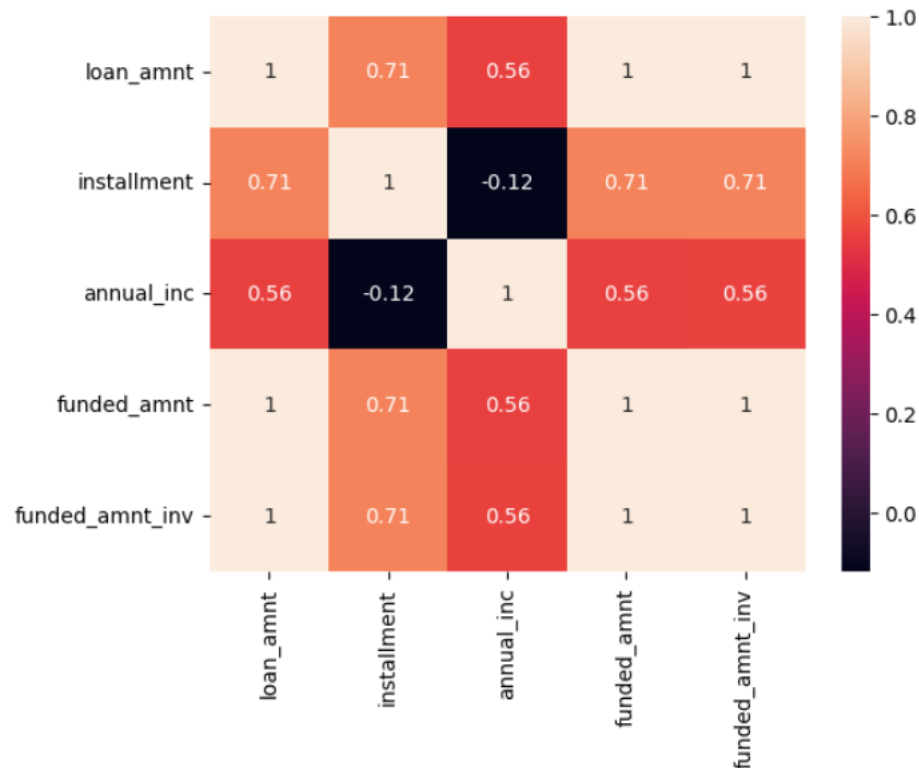
- Grades are further subdivided into subgrades such as A1, A2, A3, A4, and A5.
- In the legend, the numbers 1, 2, 3, 4, and 5 correspond to these respective subgrades.

# Loan Status and Grade



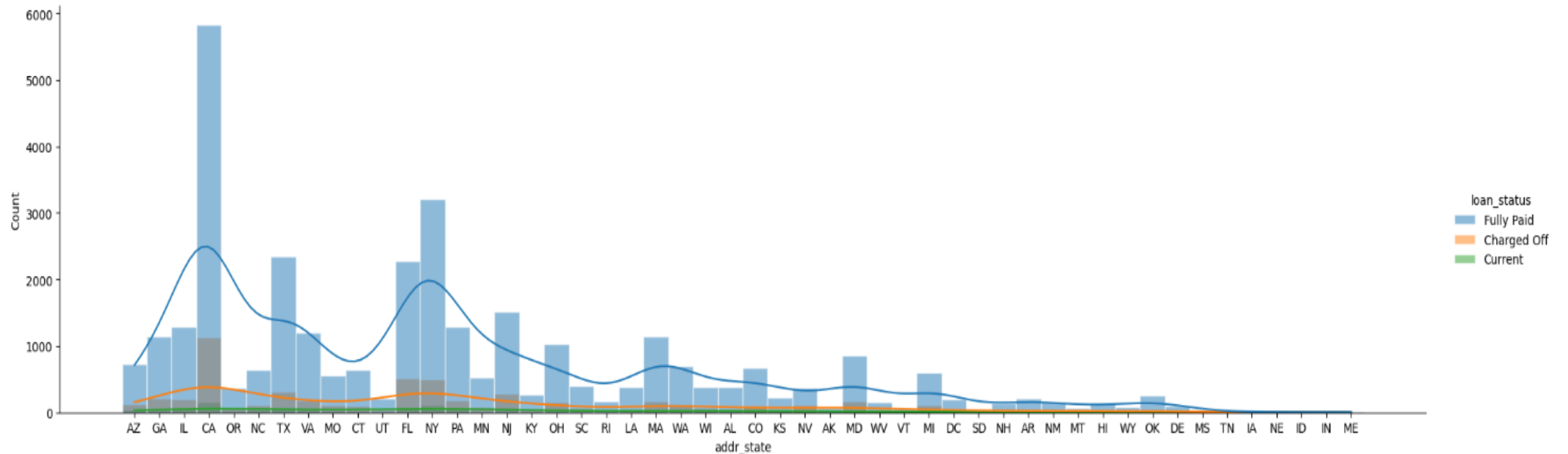
- Loans categorized under grades A, B, and C demonstrate a higher chances of repayment compared to loans in other grade classifications.

# Correlation between Loan Amount, Installment, Annual Income, Funded Amount, Funded Amount Invested



- Correlation between the numerical variables in dataset.
- Higher correlation is between loan amount and funded amount and funded amount investment.
- Lower correlation is between annual income and installment.

# State and Loan Status



- In large States such as California, New York, Texas, and Florida, there is a high volume of loans, including a notable number of defaults. However, the overall likelihood of defaulting in these regions remains relatively low.

# Conclusion

---

1. Lending Club should control their number of loan issues to borrowers who are from CA, FL and NY to make profits.
2. Borrowers with mortgage home ownership are taking higher loans and defaulting the approved loans.
3. Grades are good metric for detecting defaulters. Lending club should examine more information from borrowers before issuing loans to Low grade (G to A).