Lending Club Case Study

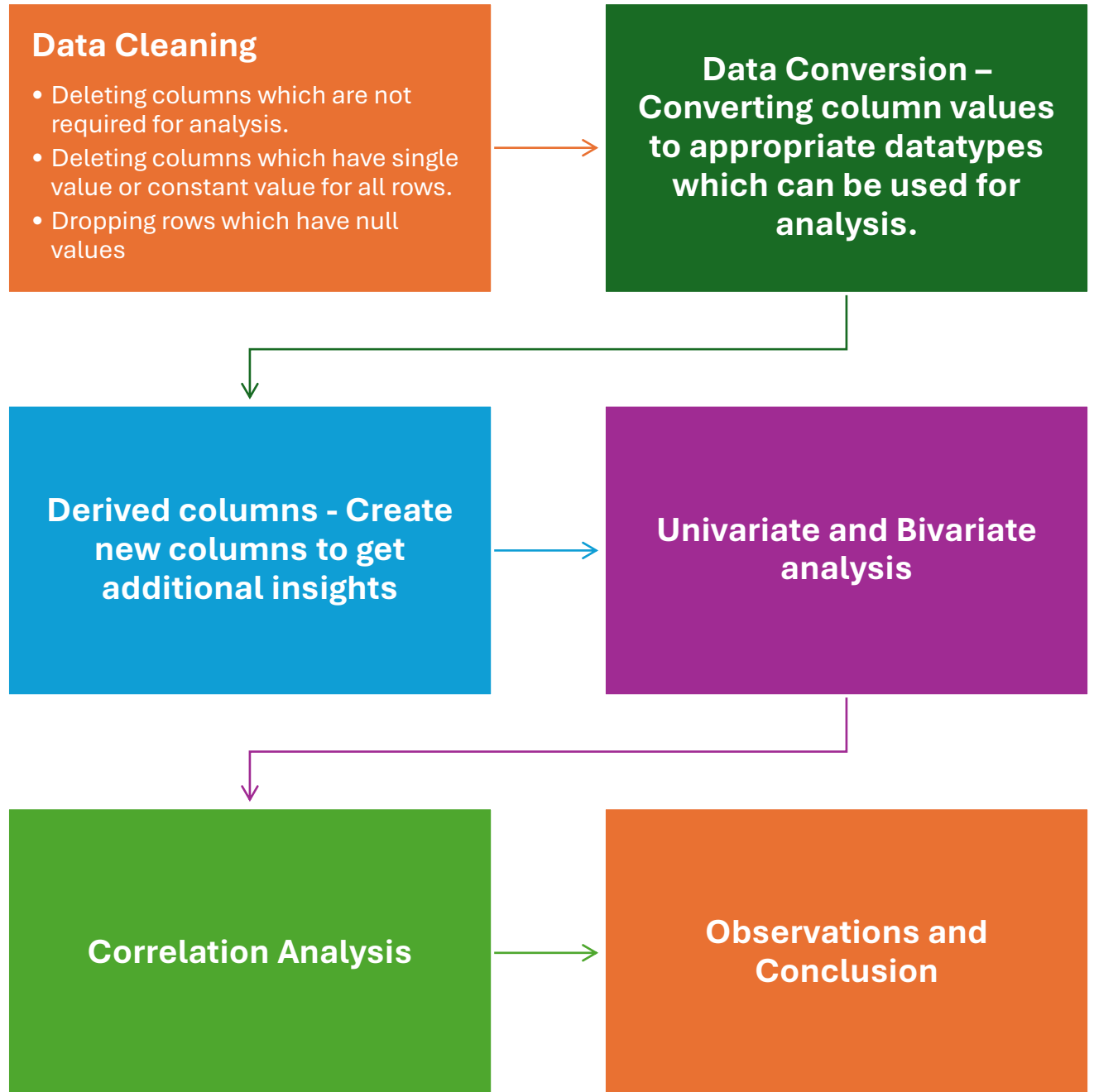By Ravi Theja and Bhushan Pande

Problem Statement

To identify patterns from data provided which indicate if a person who applied for loan is likely to default. Based on the patterns company can take decisions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

In other words, the company wants to understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

Steps for EDA

Data Cleaning

- Deleting columns which are not required for analysis.
- Deleting columns which have single value or constant value for all rows.
- Dropping rows which have null values

Data Conversion – Converting column values to appropriate datatypes which can be used for analysis.

Derived columns - Create new columns to get additional insights

Univariate and Bivariate analysis

Correlation Analysis

Observations and Conclusion

Driver variables for loan default

- Interest Rate
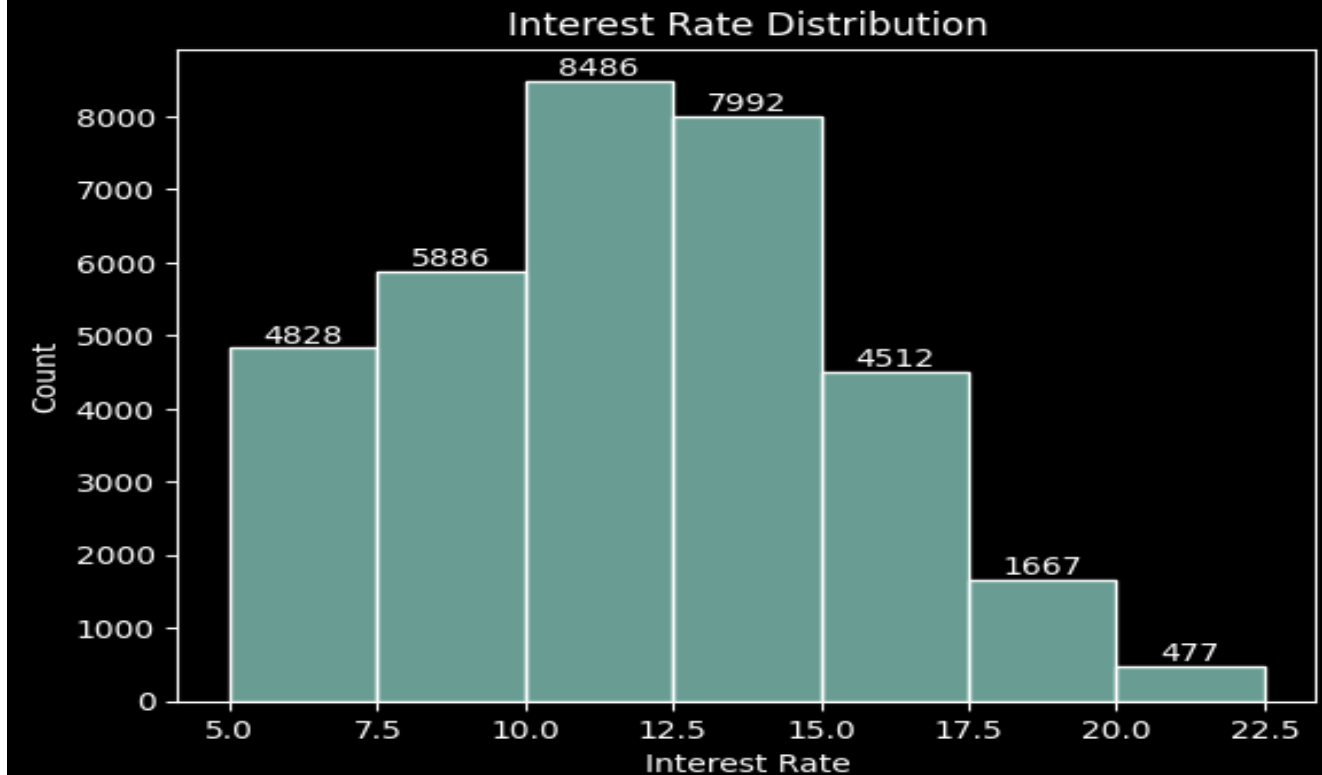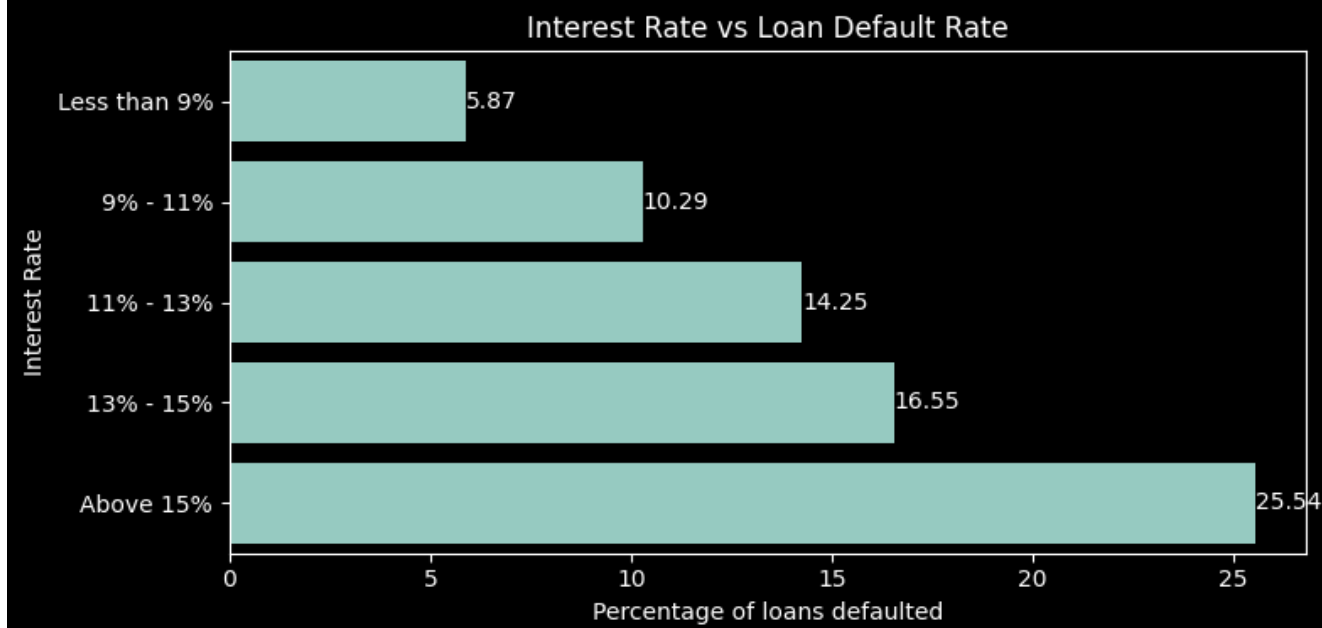- Annual Income
- Loan Amount
- Grade of the Loan
- Purpose of the Loan
- Public record for bankruptcies
- Employment Length

Interest Rate

- Average loan interest rate is **11.82%**.

- Most approved loans have an interest rate between **10-12.5%**.

- Loans with interest rates less than 9% have relatively less chances of default.

- Interest rates are starting from minimum 5.42 %.

- Loans with interest rates above 15% have relatively higher chances of default.

- The rate of loan defaults increases as the interest rate rises.

Interest Rate vs Loan Default Rate


Interest Rate Distribution

Annual Income

- The median annual income is 55.5k.
- Most of the loan applicants have an annual income of 40k-60k.
- <mark>The rate of loan defaults decrease with increase in borrower's annual salary.</mark>
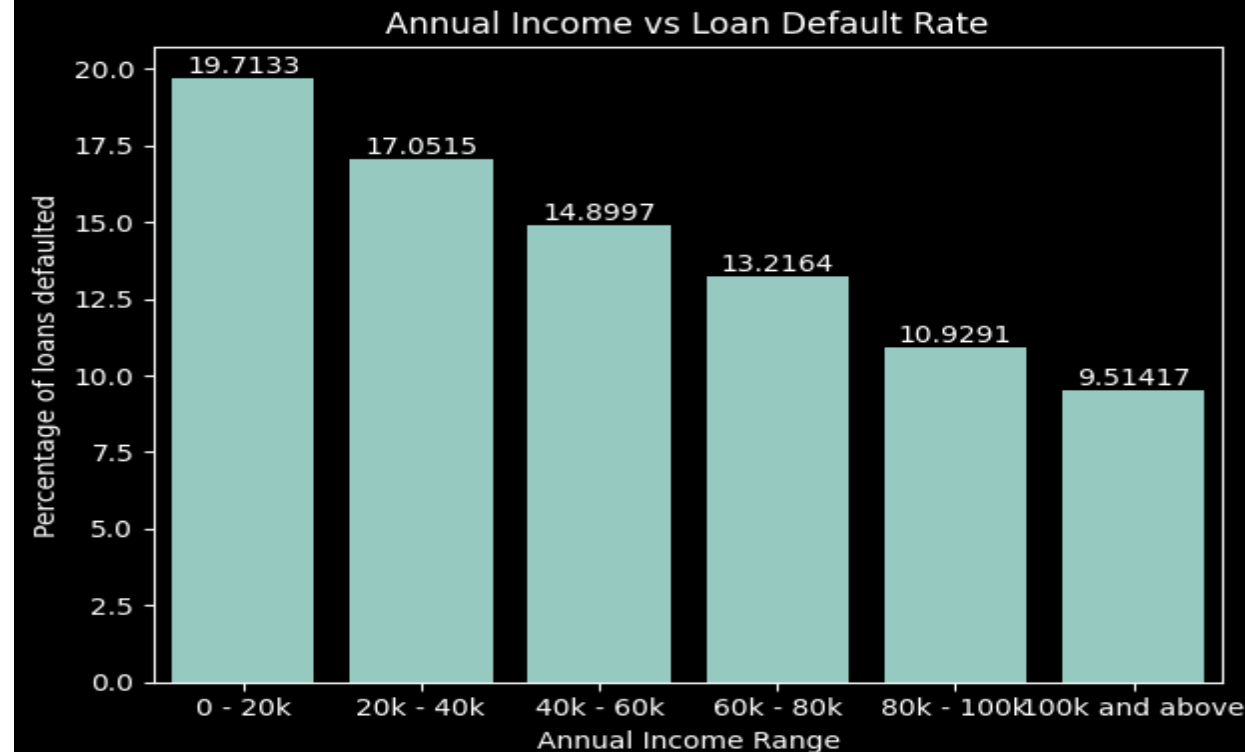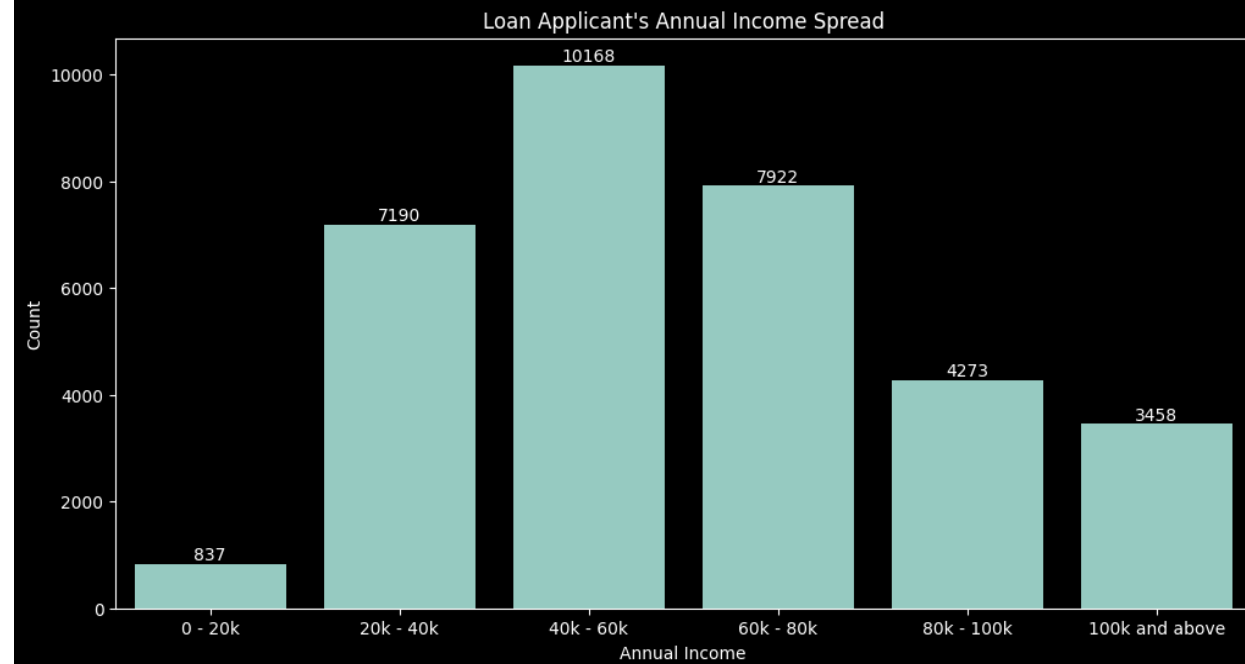
Loan Amount

- Most loans fall within the **5k to 10k** range, but these loans do not have the highest default rates, indicating that borrowers may be more capable of repaying smaller loans.

- The highest rate of default is observed in the **25k and above** category, with a default rate of **20.05%.**

- The default rate generally increases with the loan amount.
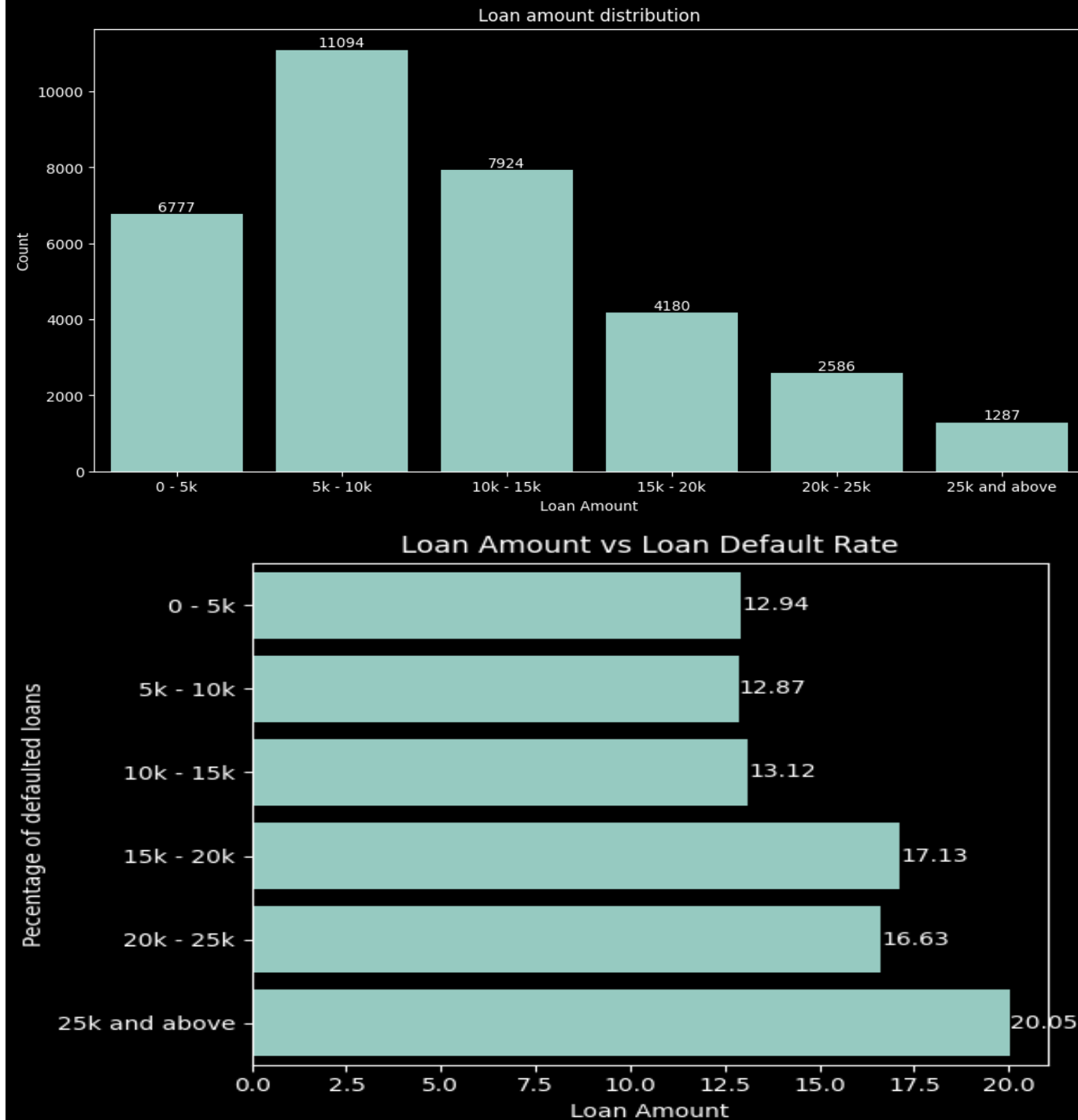
Purpose of loan

- **Debt Consolidation** is the most popular purpose (**47.3% of total loans).**

- Small business loans and renewable energy loans are the riskiest loans with 26.19% and 18% default rate, respectively.
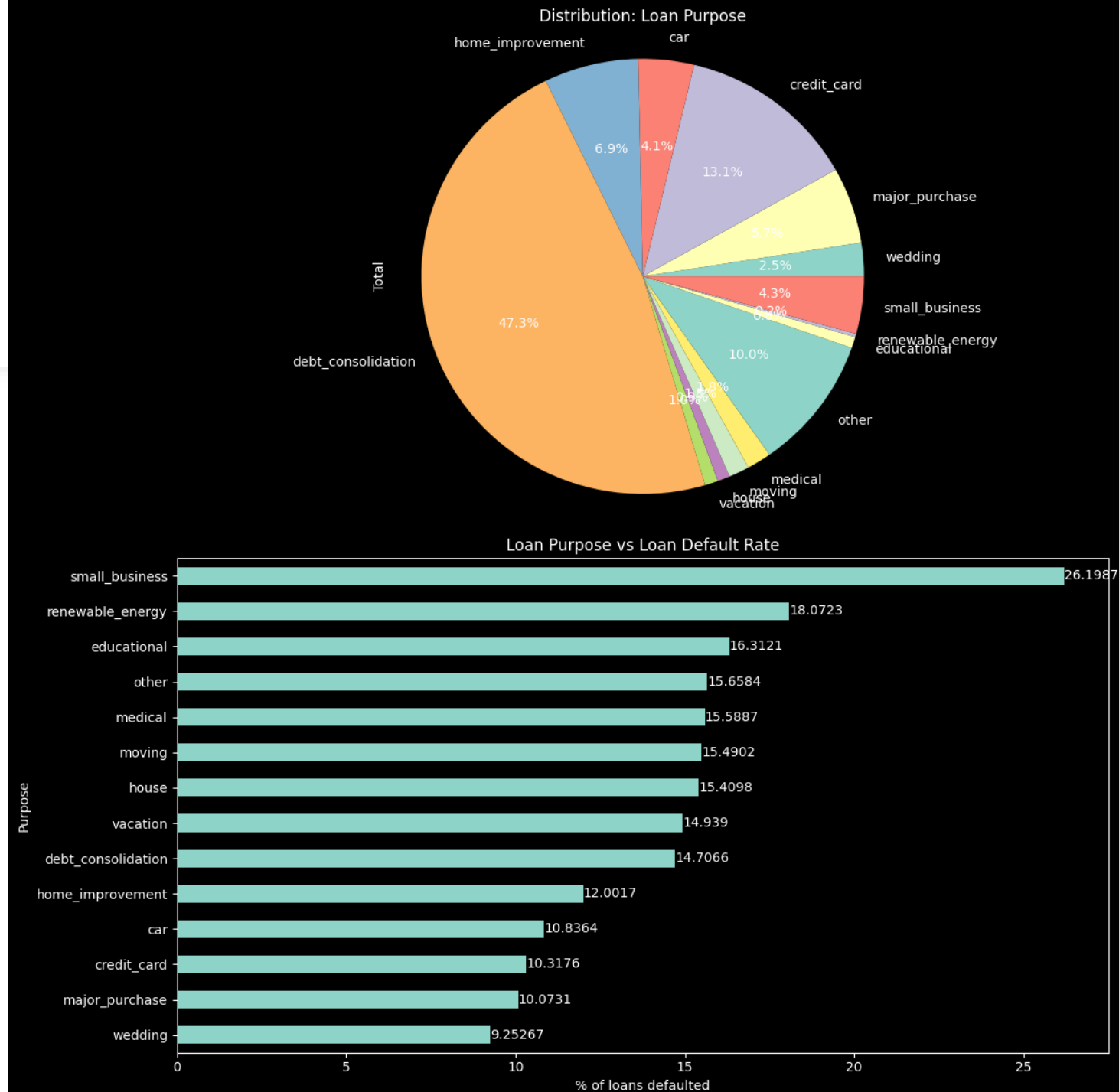
Grade of the Loan

- Lower grades have significantly higher default rates compared to higher grades.
- This correlates with the interest rate analysis where the default rate increases with increase in interest rate.
- There is an inverse correlation between loan grade and rate of default. The lower the grade, higher the chances of default.
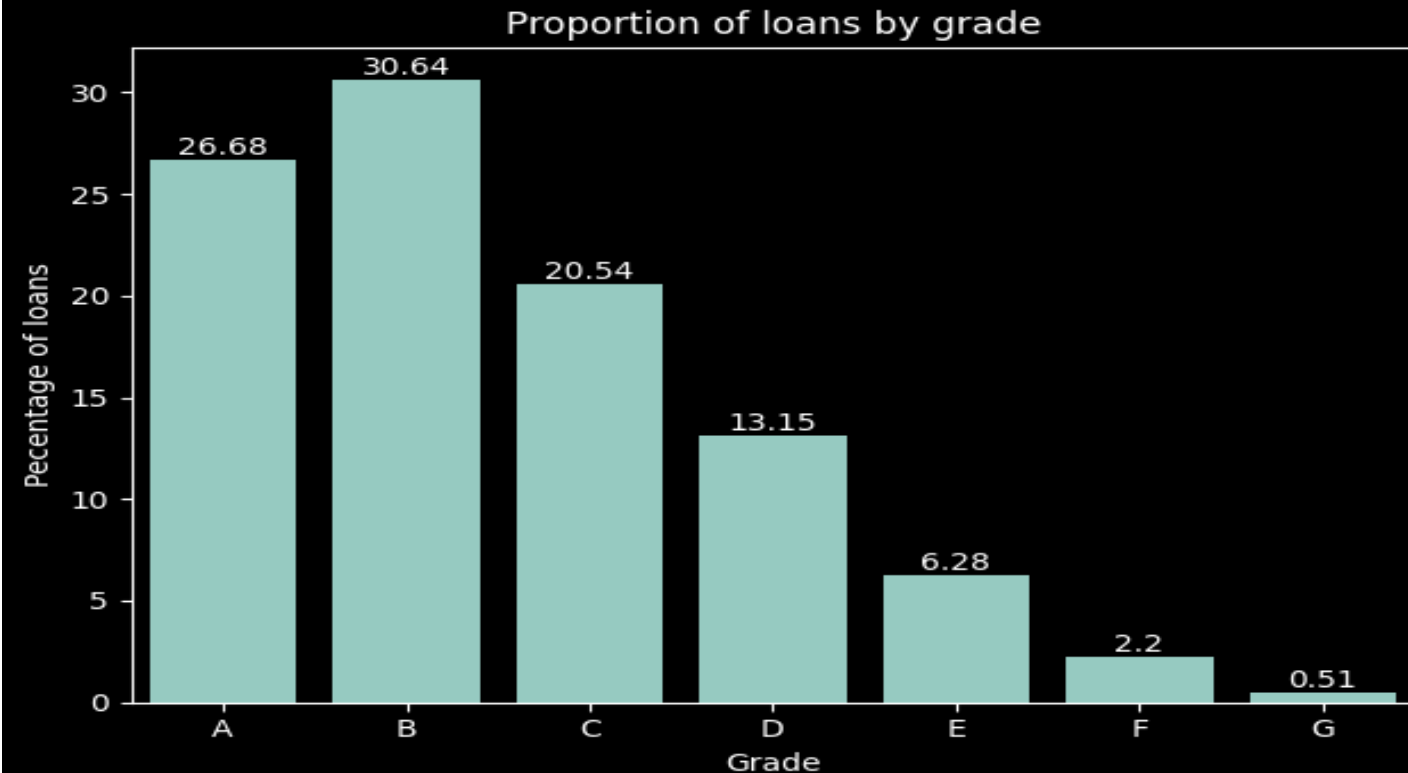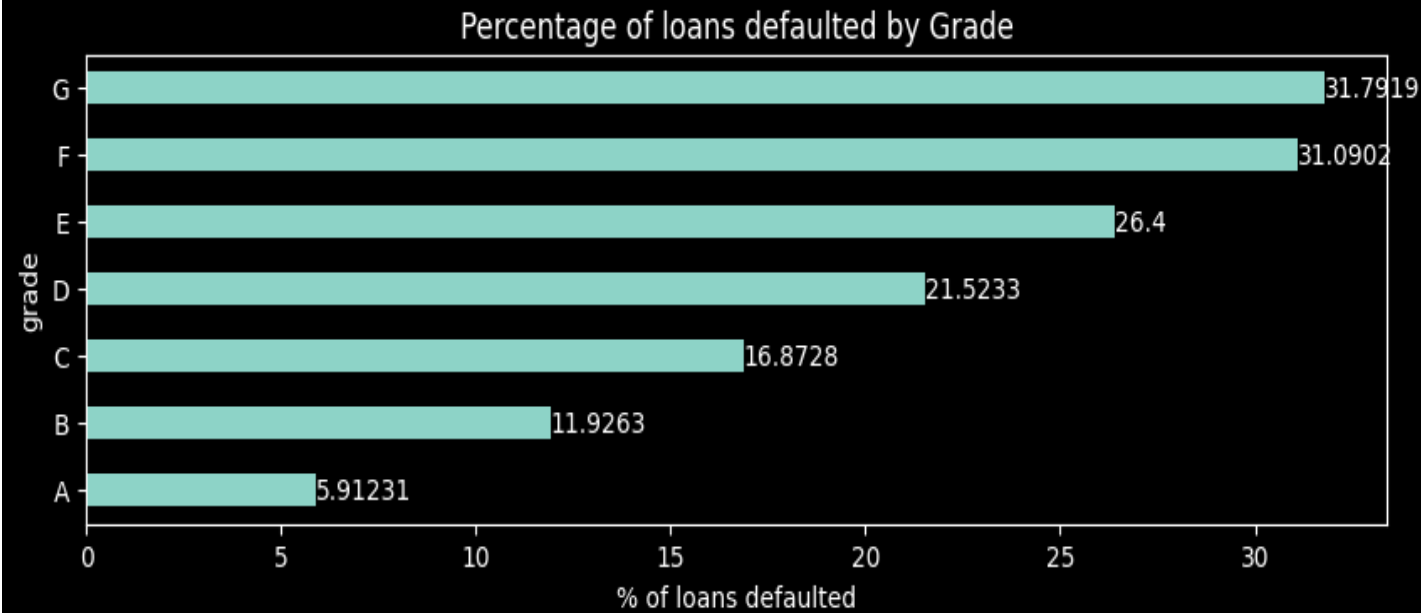
Public record of bankruptcies

- 4.3% of borrowers had 1 publicly recorded bankruptcy.
- The percentage of defaults for borrowers with 1 past bankruptcy is 6.5%
- Borrowers with at least one past bankruptcy are more likely to default than borrowers who don't have a past bankruptcy.
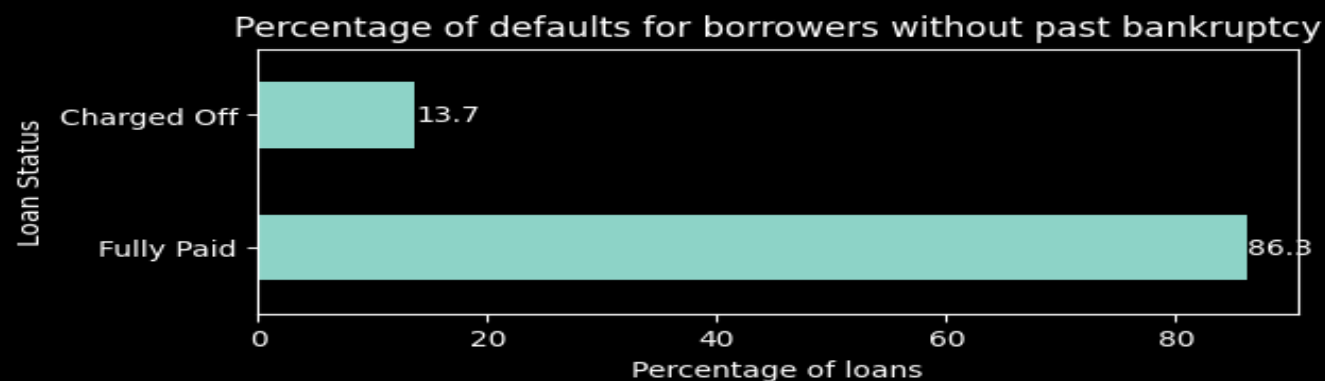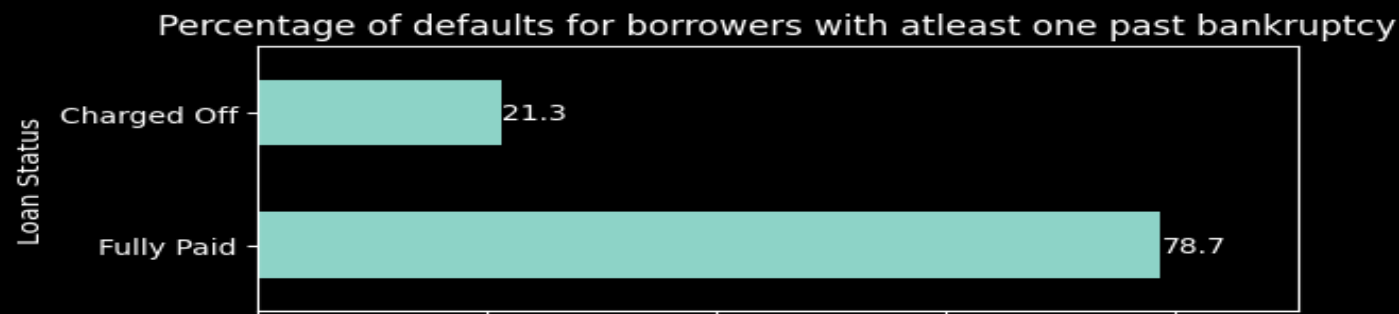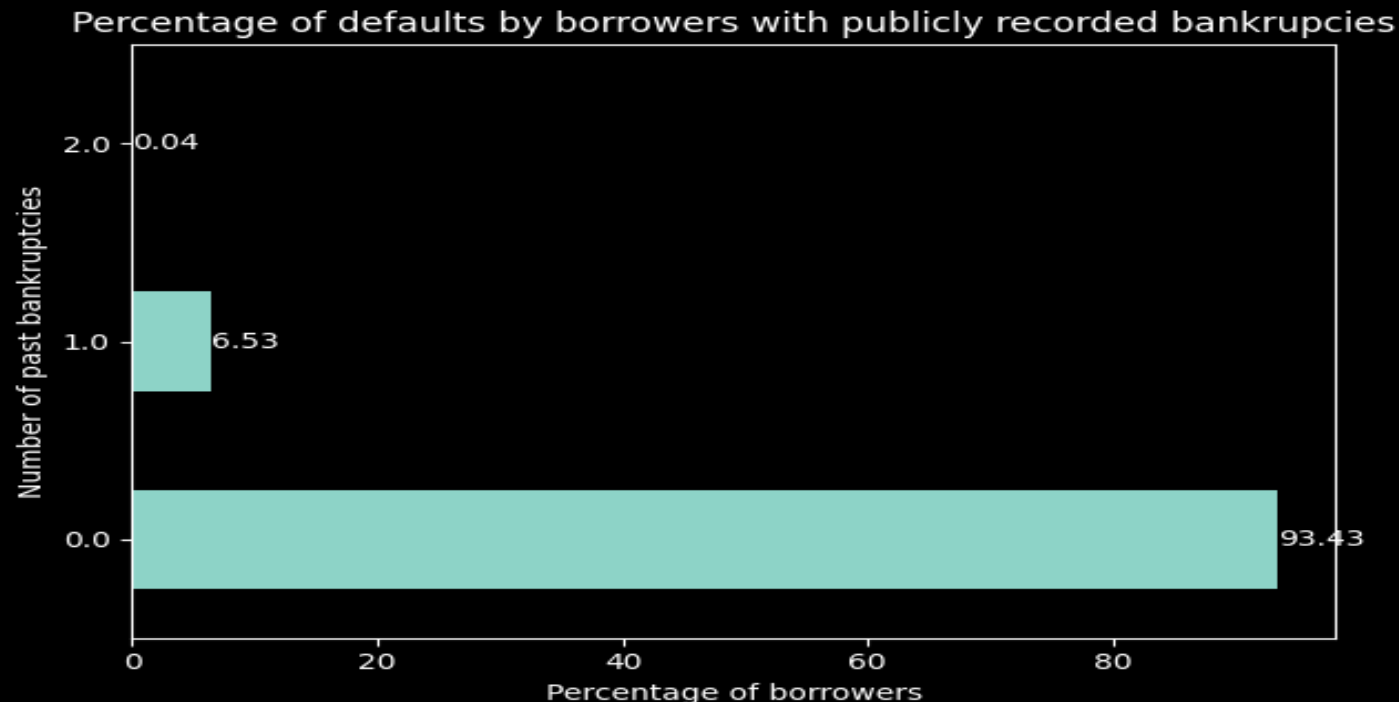
Employment Length

• Contrary to popular belief, longer employment doesn't necessarily lead to lower default rates as the default rate increases for borrowers who 10+ years of employment length.

• Further drill-down analysis can be done to identify micro trends resulting in default based on the employment length.
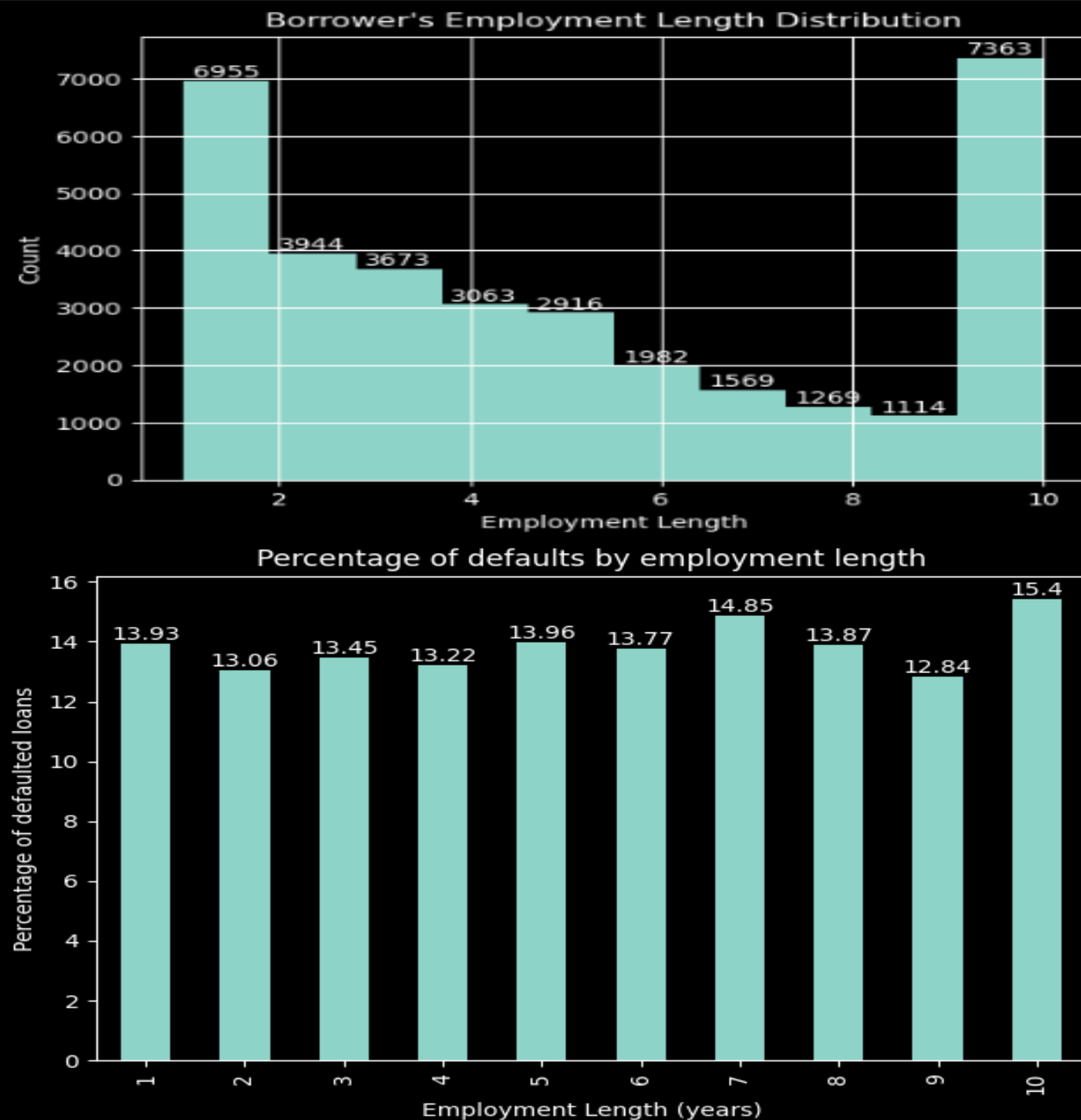
Below correlation matrix can be used to identify additional insights

Correlation Matrix
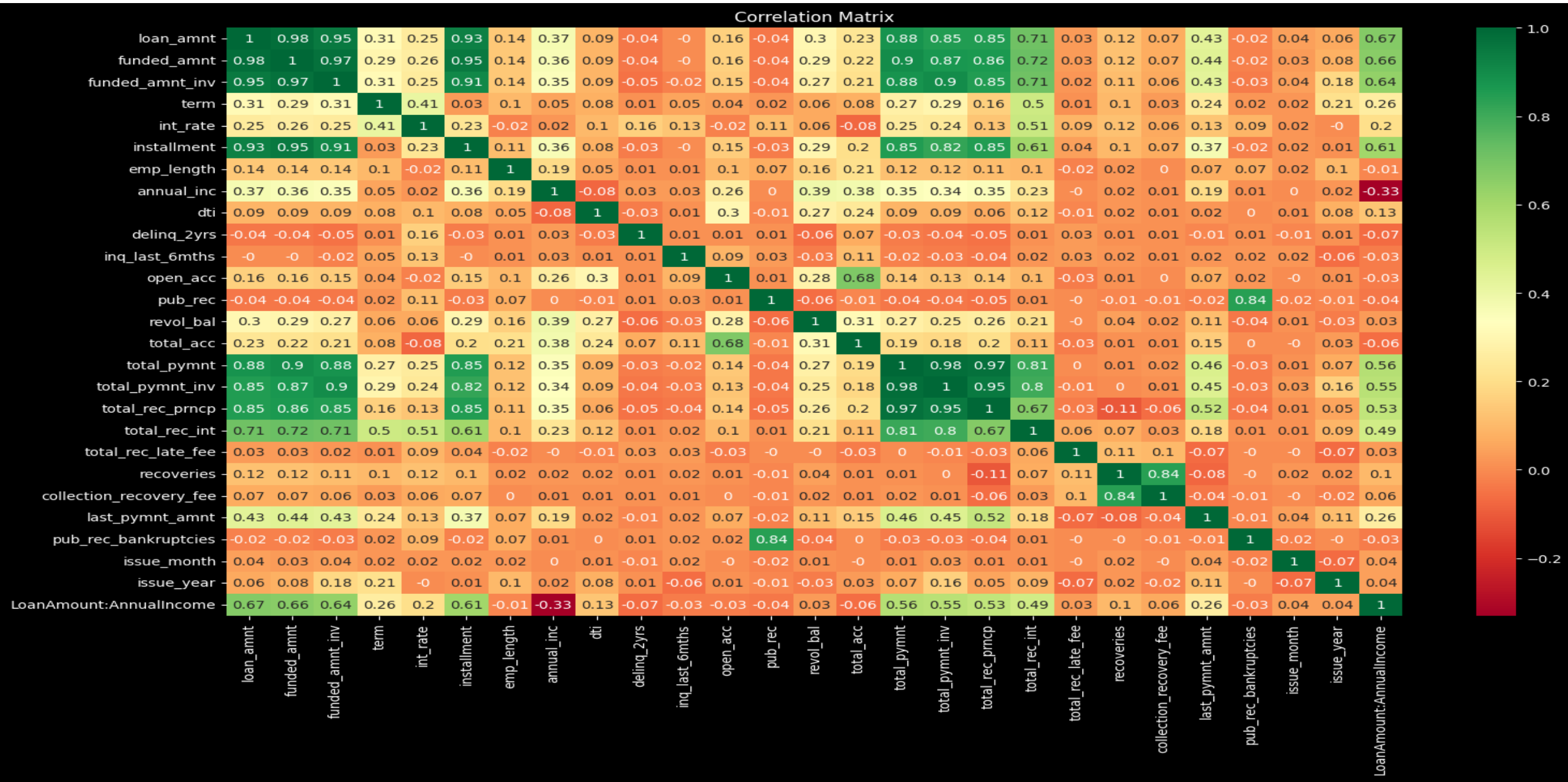
Patterns resulting in loan defaults

- The rate of loan defaults increases as the interest rate rises.
- The rate of loan defaults decrease with increase in borrower's annual salary.
- The default rate generally increases with the loan amount.
- There is an inverse correlation between loan grade and rate of default. The lower the grade, higher the chances of default.
- Borrowers with at least one past bankruptcy are more likely to default than borrowers who don't have any past bankruptcy.
- Small Business and renewable energy loans tend to have a higher default rate.

Recommendations

- Keep the loan interest percentage between 10-12.5, to attract more customers.

- Target customers with annual income between 40k-60k with attractive offers as they are more likely to apply for a loan.

- Secure loans with better collateral quality if the applicant has past bankruptcy records or loan amount is greater than 25k or if the purpose for loan is "small business", "renewable energy" or "debt consolidation".

Thank you