LENDING CLUB CASE STUDY

Boddu Subham Rao

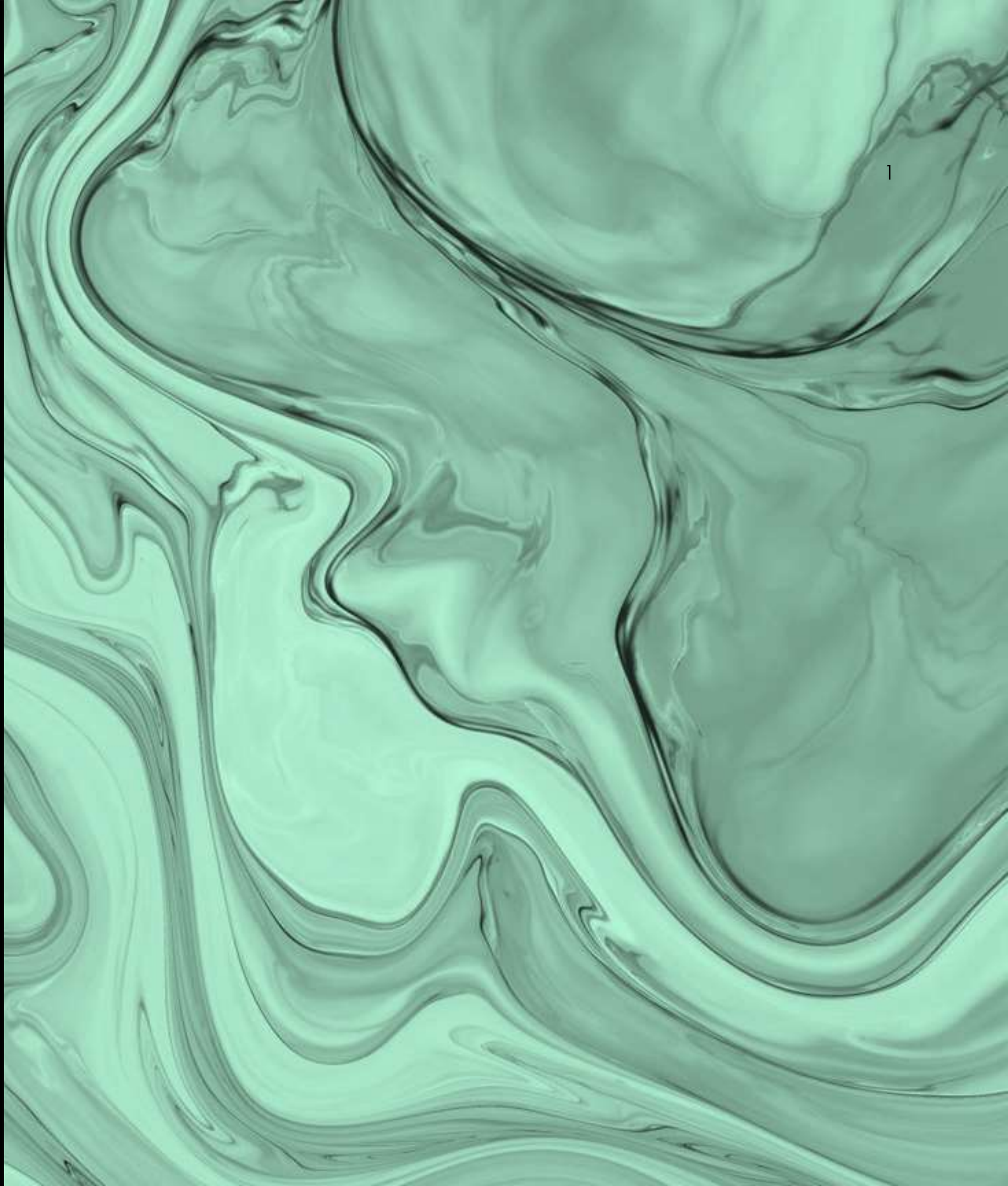


Table of Contents:

Objective

- The goal of this analysis is to **identify the key factors** that drive loan defaults, focusing on the **variables** that act as strong indicators of default risk. The main objectives are:
- **Analyze Lending Club's applicant data** from 2007 to 2011 to uncover patterns and trends associated with loan defaults.
- Minimize financial risk by **identifying** and avoiding loans to urban customers who exhibit
- characteristics linked to higher default rates.
- **Identify patterns and trends** that predict which applicants are more likely to default, enabling more informed and strategic lending decisions.
- **Prevent financial losses** by rejecting loan applications from high-risk applicants who are

Our Plan of Action

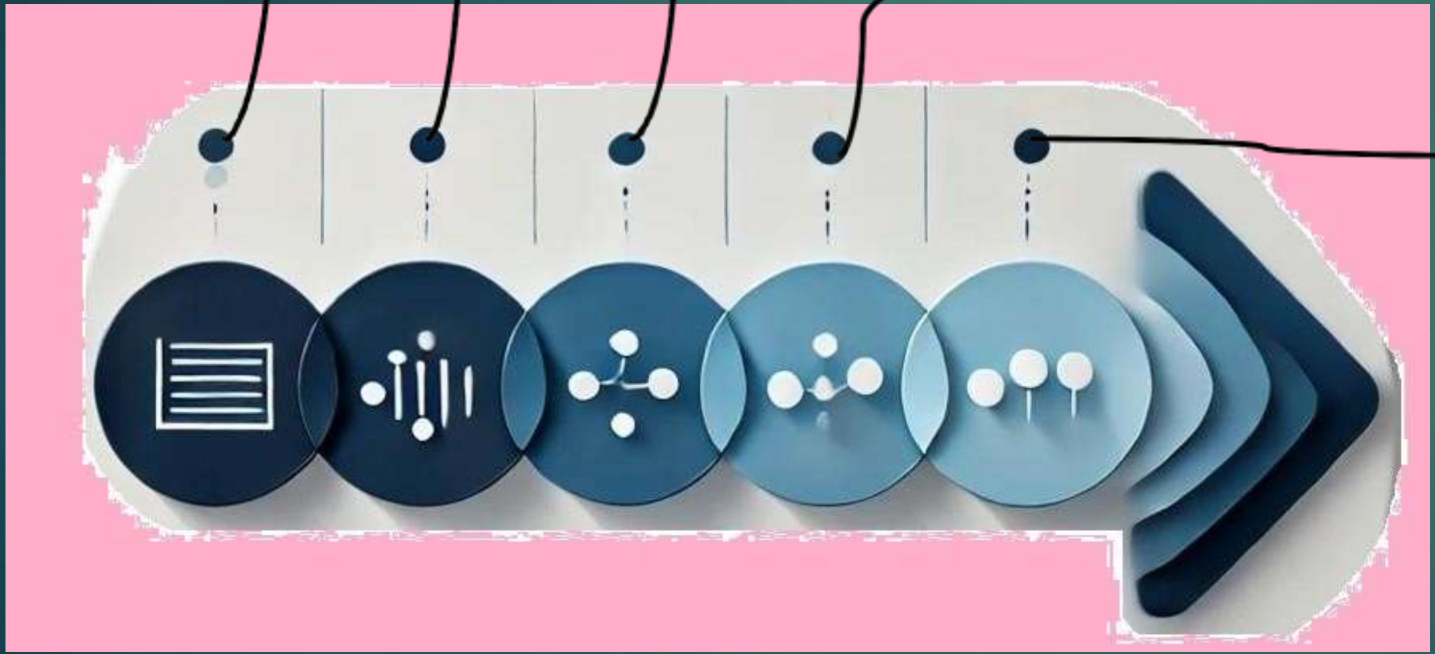
Data Exploration

Data Cleaning

Data Transformation

Analysis
- Univariate
- Bivariate
- Multivariate

Business Recommendation



Data Exploration

1. Understanding the Dataset

- Reviewed the dataset structure, including the number of rows, columns, and data types.
- Identified key variables related to loans (e.g., funded amount, interest rate, annual income, term, purpose, DTI).
- Mapped relationships between features and target variables (e.g., loan repayment status).

2. Data Cleaning

- Addressed missing values:
    - Used imputation for numerical variables.
    - Removed unnecessary or irrelevant variables.
- Handled duplicate records to avoid skewed analysis.
- Corrected inconsistencies in categorical data (e.g., standardized labels).

3. Data Transformation

- Created new features to gain deeper insights, such as:
    - Income-to-loan ratio.
    - Loan-to-interest rate categorization.
- Converted categorical data into numerical formats for analysis.

4. Exploratory Data Analysis (EDA)

- Performed statistical summaries to identify trends:
    - Mean, median, and standard deviation of key variables.
- Used visualizations to analyze distributions and correlations:
    - Histograms for continuous variables (e.g., annual income).
    - Heatmaps for identifying feature correlations.
    - Box plots for understanding variability across loan grades.

5. Outlier Detection

- Identified and treated outliers in critical variables like funded amount, annual income, and DTI.
- Applied appropriate techniques:
    - Capping extreme values for better model stability.

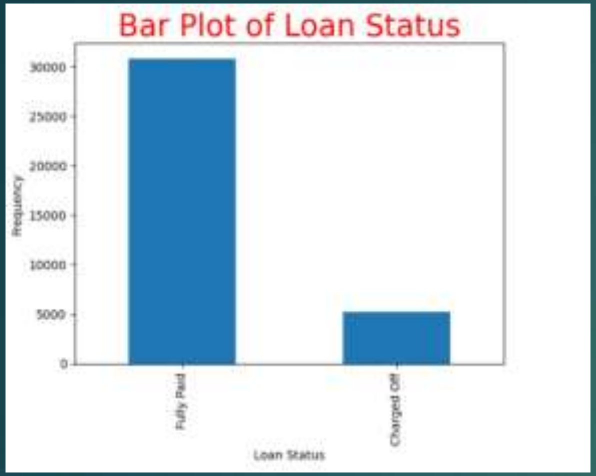
6. Segmentation and Grouping

- Segmented data based on key factors:
    - Loan grades, repayment status, income brackets.
- Analyzed repayment trends within each segment for actionable insights.

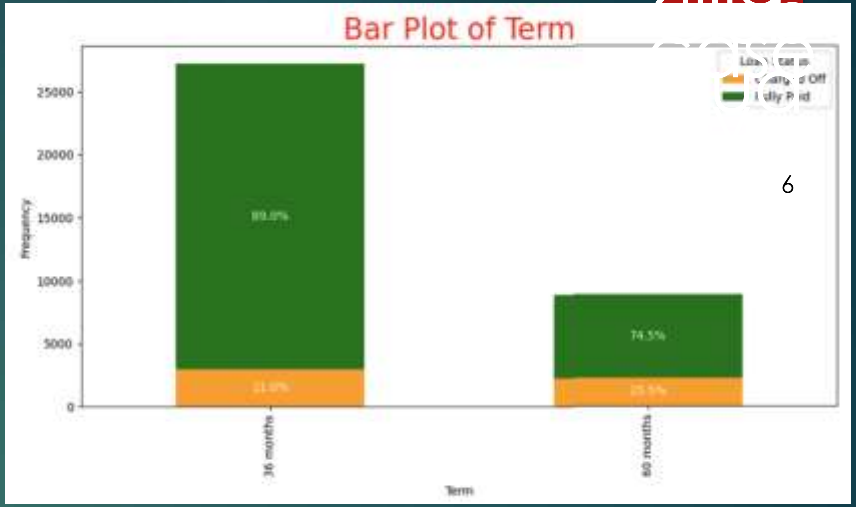
7. Initial Observations and Hypotheses

- Derived key trends and correlations:
    - Higher income borrowers had better repayment rates.
    - Smaller loan amounts were less likely to be charged off.
- Formulated hypotheses for further validation during model development.

- **Univariate Analysis**



From graph we can derive that **85.47% of loans** were **fully repaid** & **14.53% - Charged-off**

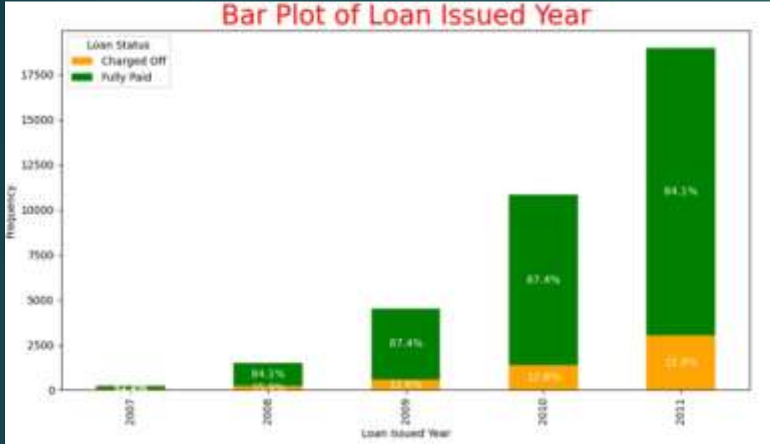


From above graph 89% of applicants with term-36months fully repay the loan, whereas only 74.5% of applicants with term-60 months fully repay the loan. Lend Money to applicants at short terms, here:36 months to avoid financial loss

Approximately 15% of the applicants charge off irrespective of employee length. It is safe to assume that Employee length does not impact the loan repayment.

The applicants who charged OFF is less in the months of January-April. Applicants taking loan in November or December default the highest. The reason for this could be Holiday(Black Friday, Christmas) season

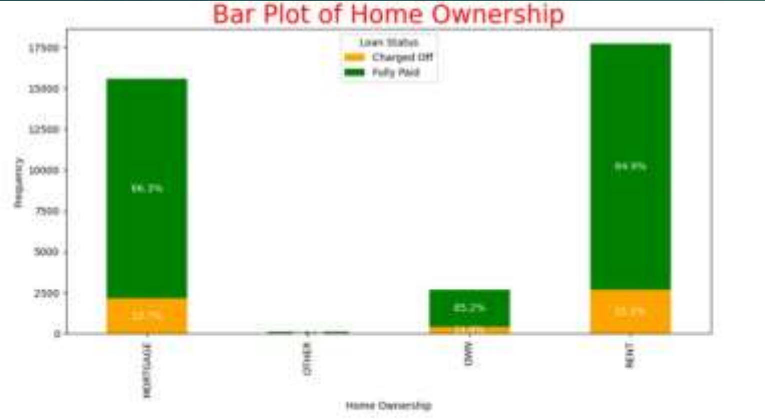
- **Univariate Analysis**



Approximately every year 15% of the applicants are charged-off.

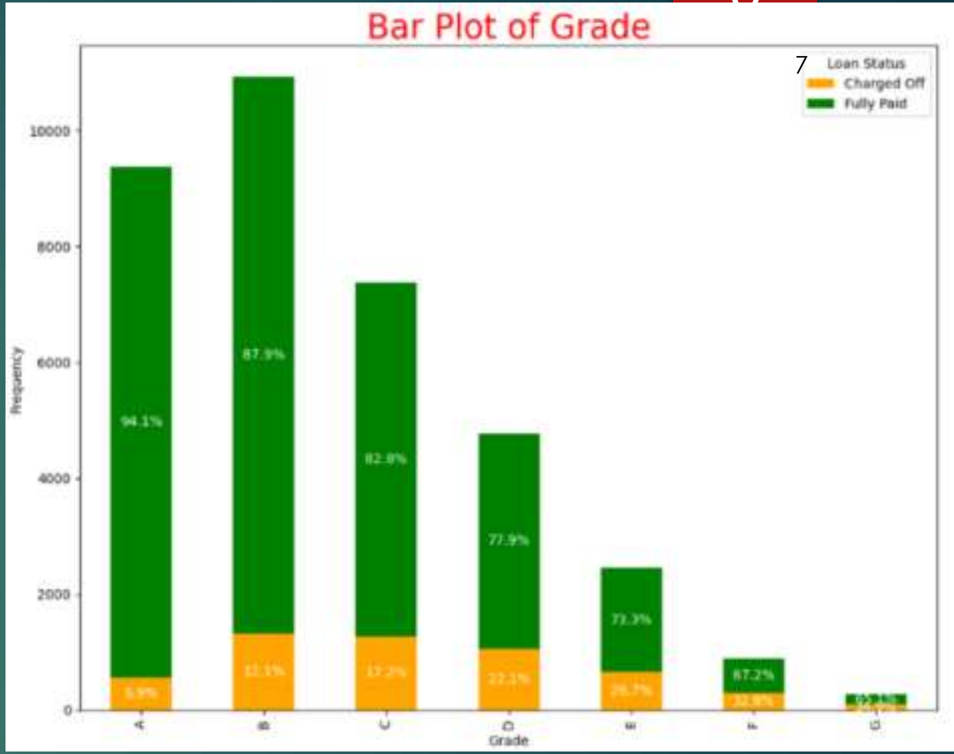
- It is safe to assume that Employee length does not impact the loan repayment.

From above graph we can derive that 85.47% - full repaid the loan and 14.53% - Charged-off



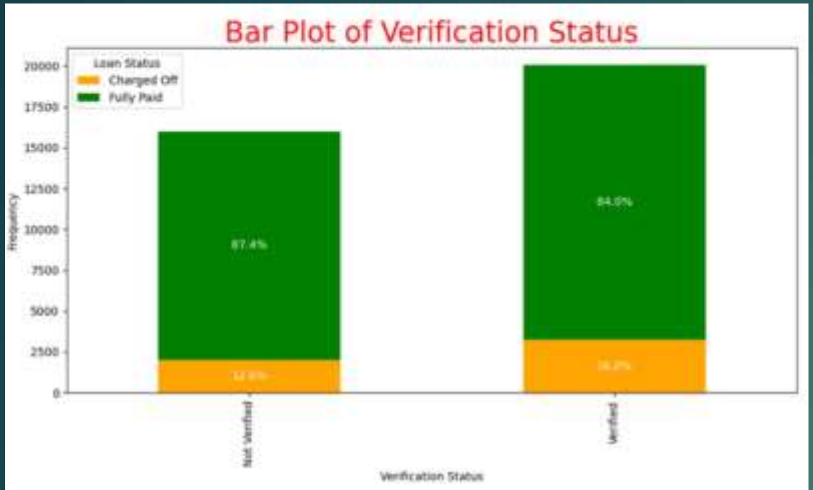
Approximately in all categories of Home Ownership 14% of the applicants are charged-off.

- It is safe to assume that Home Ownership does not impact the loan repayment.

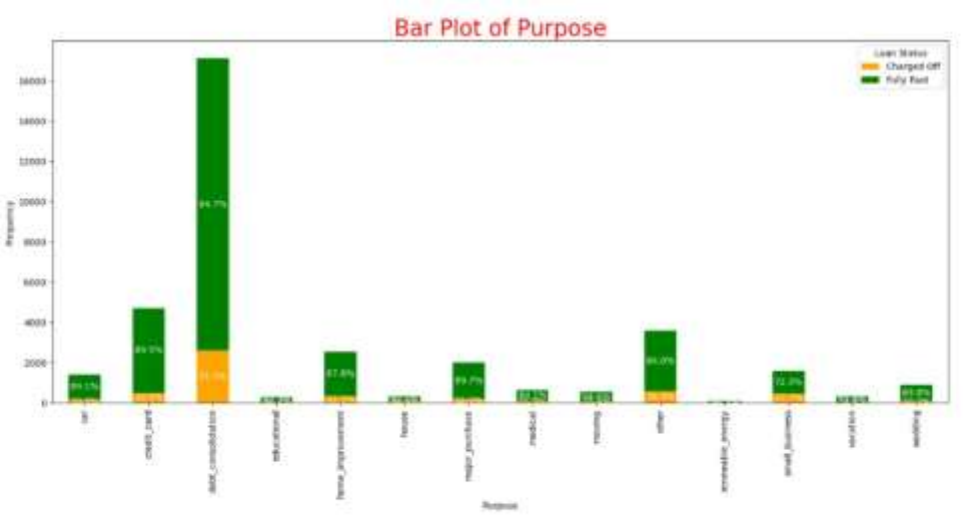


- Number of applicants who charge-off is significantly less when LC Grade is A,B.
- Number of applicants who charge-off is significantly higher when LC Grade is E,F and G.
- Lend Money to people who are in LC Grade A and B to avoid business loss, Lend Money to applicants who are in LC Grade C and D with shorter term – 36 months to reduce defaulters as per our earlier analysis.
- Reject applicants whose LC Grade are E,F and G

- **Univariate Analysis**



Non-verified applicants are repaying the loan better than Verified applicants.
But It can't be considered as a driving factor as the difference is not very significant

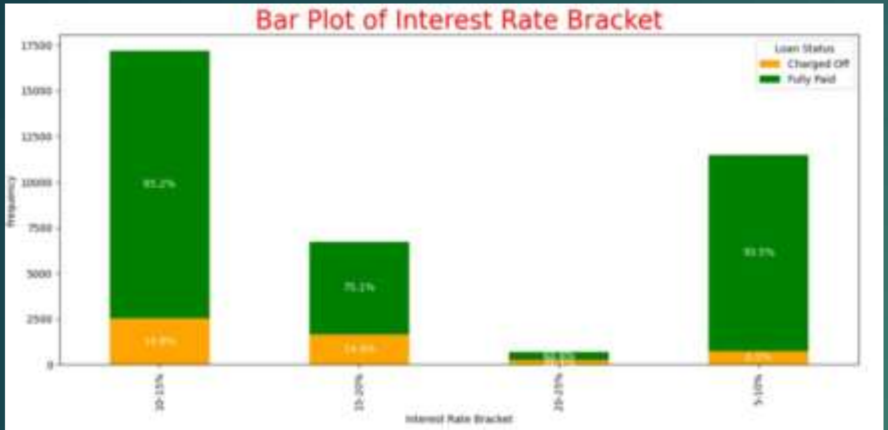


Applicants taking loan for small Business and renewable energy default the most.
For applicants in small Business and renewable energy provide loans with the shorter term to avoid financial losses

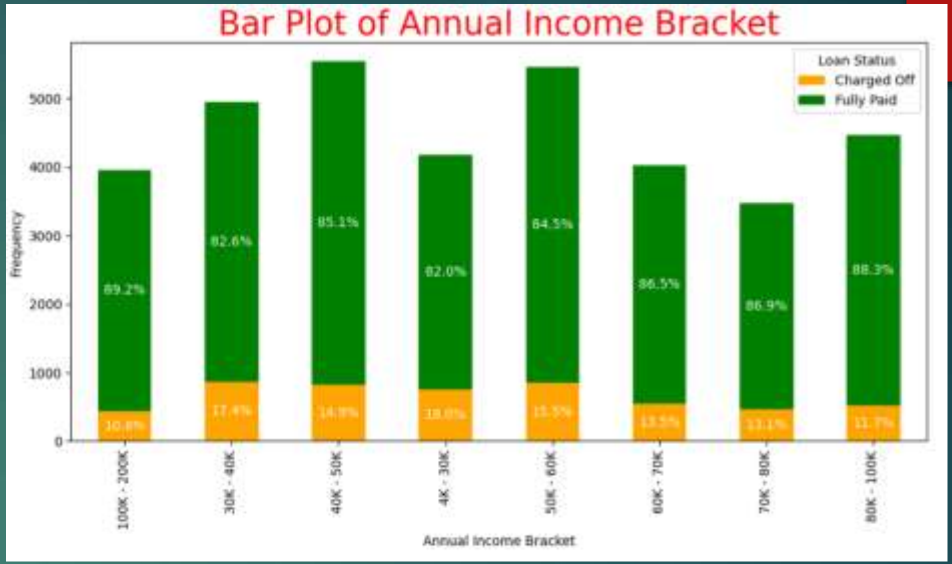
Approximately in all categories of Home Ownership 14% of the applicants are charged-off It is safe to assume that Home Ownership does not impact the loan repayment.

The funded amount is directly proportional to applicants being charged off.

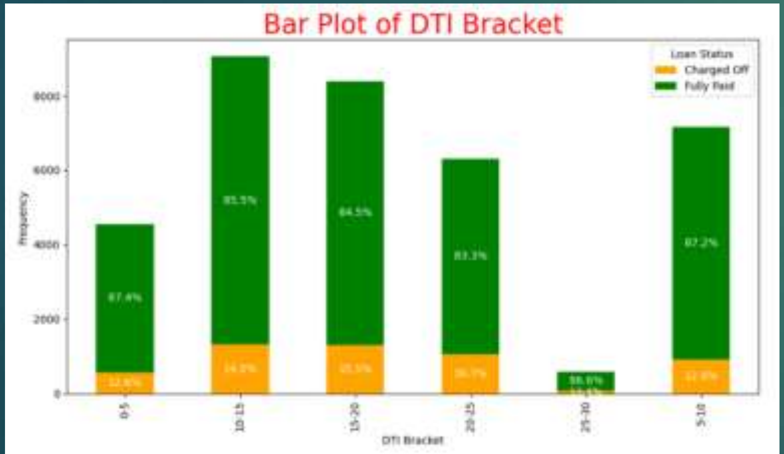
- **Univariate Analysis**



The Interest Rate is directly proportional to applicants being charged OFF

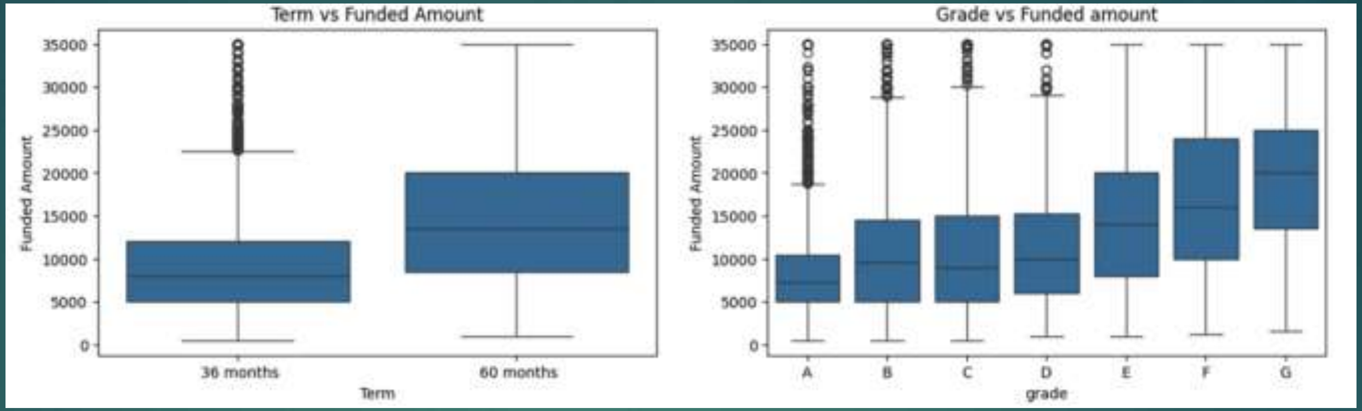


The Applicant's annual income is inversely proportional to applicants being charged OFF

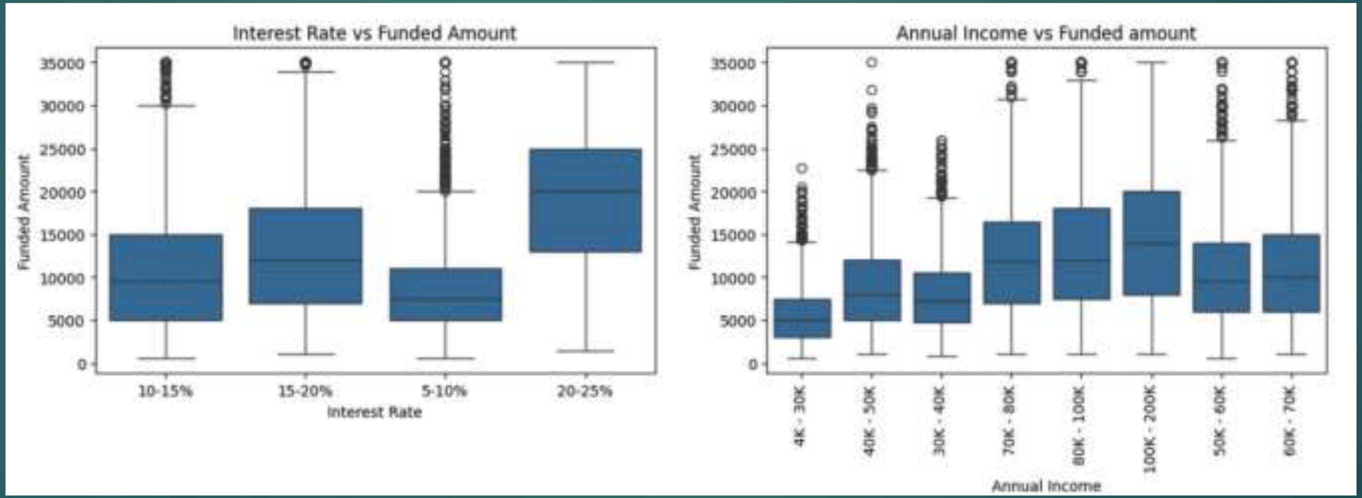


The Debt-to-Income Ratio is directly proportional to applicants being charged OFF

- Bivariate Analysis

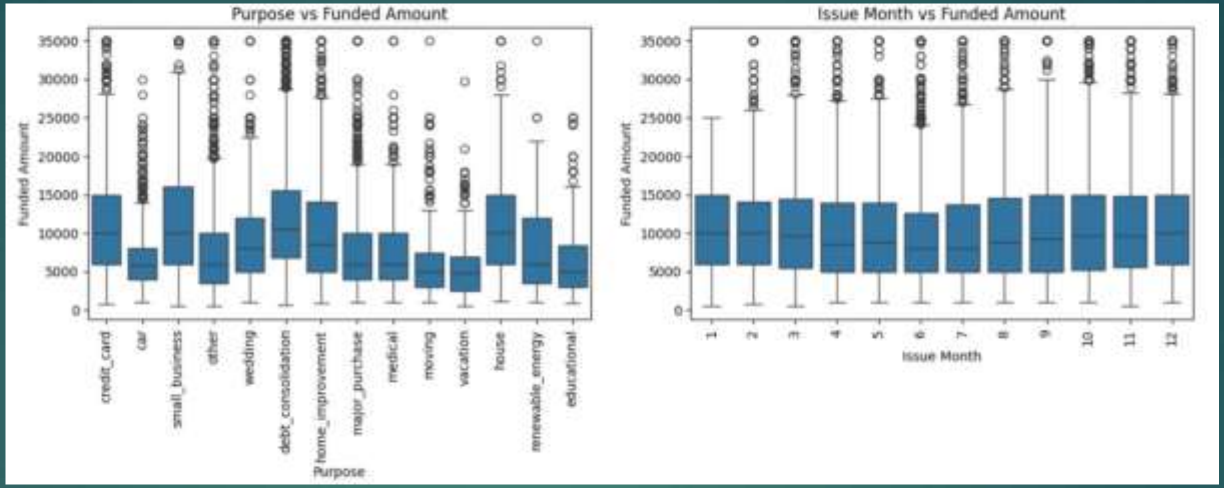


Funded Amount is directly proportional to term of the loan. Funded Amount is directly proportional to Grades

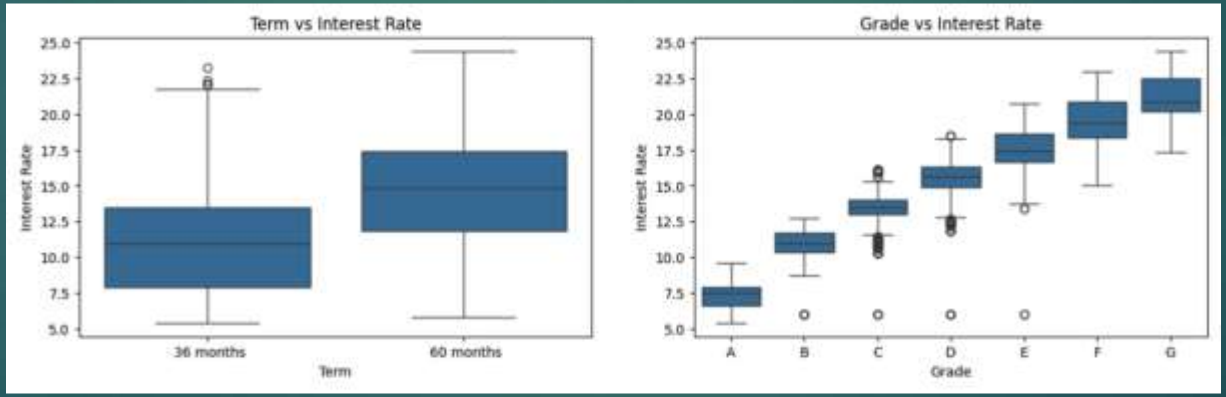


Funded Amount is directly proportional to interest rate. Funded Amount is directly proportion to Annual income

- **Bivariate Analysis**

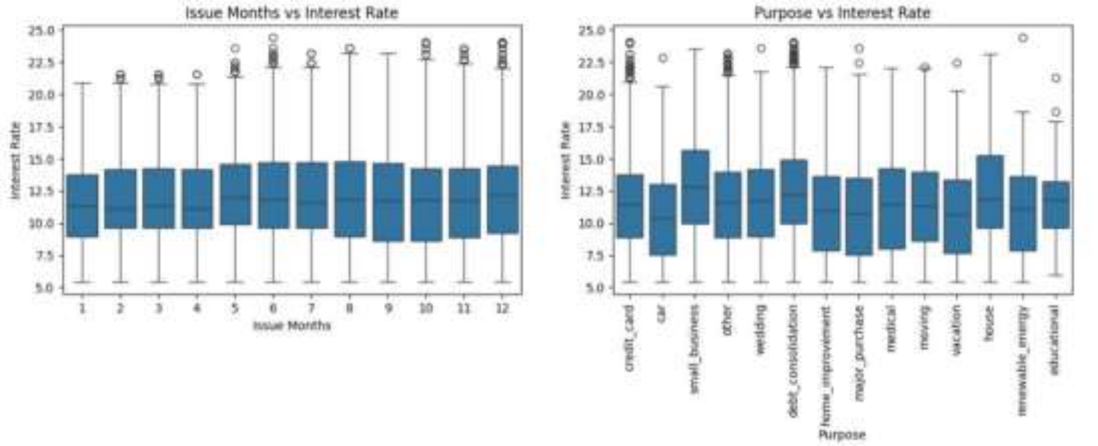


Funded Amount is higher for purposes like - Small business, debt-consolidation, credit card and home improvement.
Funded Amount is does not vary across months

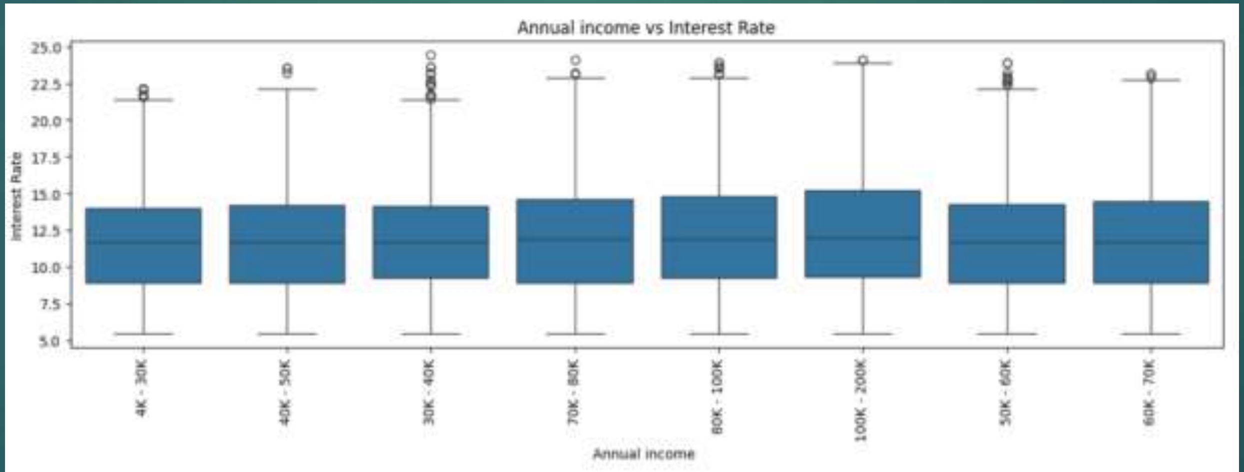


Interest Amount is directly proportional to Term and Grade

- **Bivariate Analysis**

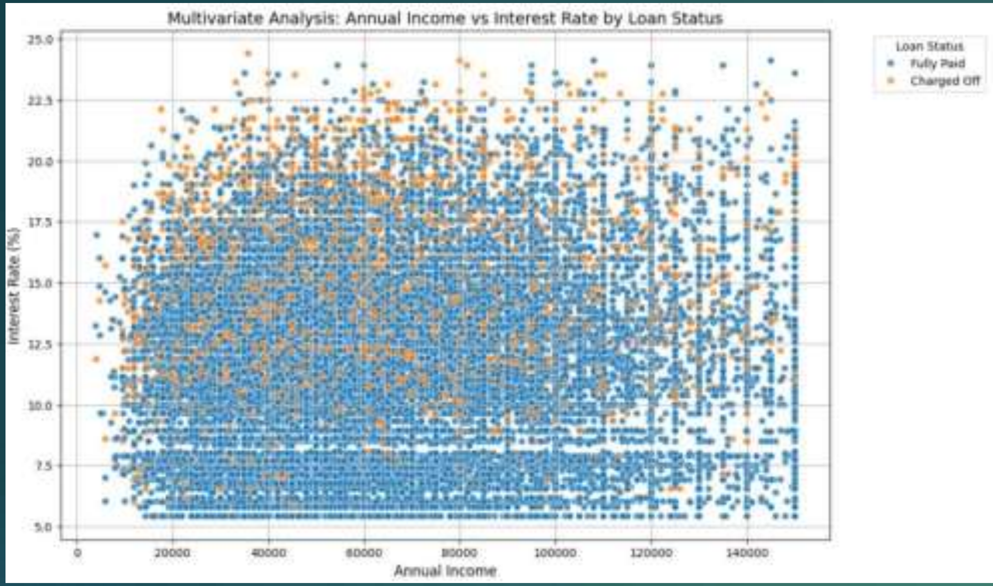


Interest rate is higher for small business, debt consolidation and house. Interest rate doesn't vary across months



Interest rate doesn't vary across annual income

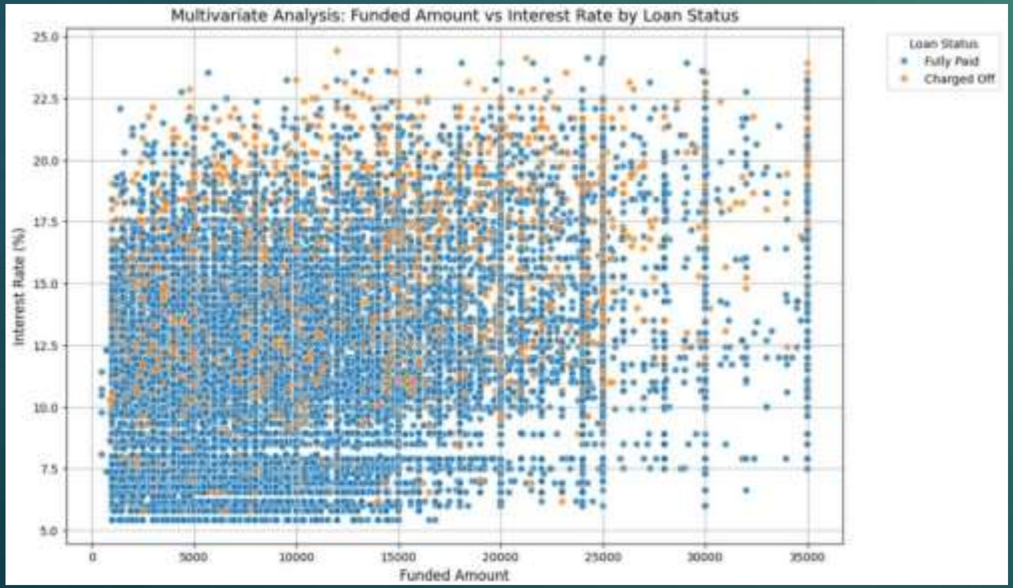
- # Multivariate Analysis



- The applicants tend to Charged OFF as the Interest Rate increases

- The applicants tend to pay off Loans are higher for the Interest Rate is less than or equal 10% across their Annual Income
The Interest Rate 10% to 15% Charged OFF gradually decreases as Annual Income increases

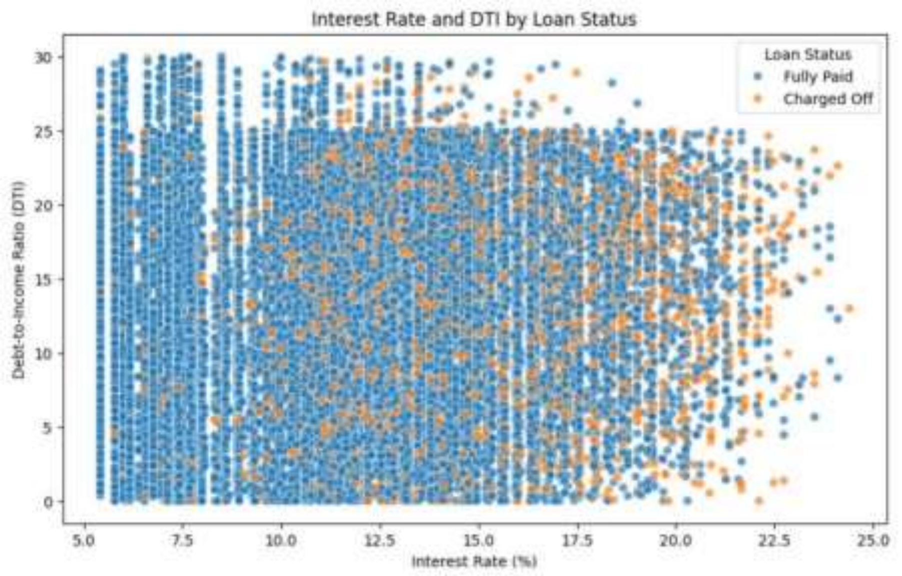
- Above 17.5% Interest Rate, the Applicants across all the Annual Income Bracket tend to Charged OFF



- Applicants tend to pay off loans if the amount and the interest rate are less for an example: the funded amount is as high as 35000 and the interest rate is 10%.

- The Charged OFF increases when the interest rate is above 15%

Multivariate Analysis



- The applicants pay off the loan when the funded amount is less than 15000

- The applicants Charged OFF when the funded amount is more than 15000 and the annual income is less than 80000

- The applicants pay off the loan when their annual income is more than 80000 and the funded amount is as high as 35000

- When interest rate is low, applicant across all DTIs are paying off the loans.

- The applicants Charged OFF when the interest rate increases across all DTIs

Business Recommendations

Avoid Business Loss

1. Funded Amount

- **Loans > $15,000**:
  - **Interest rate < 10%**: High loan amounts paired with low-interest rates reduce default risks.
  - **Annual income > $80,000**: Borrowers with higher incomes are more likely to repay large loans.

2. Interest Rate

- **Loans with rates > 15%**:
  - **Funded amount < $15,000**: Small loan amounts mitigate risks associated with high interest.
  - **Annual income > $80,000**: Ensures affordability despite higher interest rates.

3. Annual Income

- **Low-income borrowers**:
  - Avoid high-risk purposes like **home improvement, renewable energy, small business, or wedding**, as these categories may lead to financial strain.
  - Ensure **term = 36 months** for faster repayment cycles and reduced risk exposure.

4. Debt-to-Income Ratio (DTI)

- For **DTI > 10**:
  - Ensure **interest rate < 10%** to make repayments manageable.
  - Opt for **term = 36 months** to limit long-term risks.
  - Focus on borrowers with **Grade A or B**, as they represent lower risk.

5. Purpose

- For loans aimed at **small business, renewable energy, debt consolidation, or home improvement**:
  - Use a **term = 36 months** to reduce default likelihood.
  - Ensure **DTI < 10**, indicating better financial health.

6. Grade

- For borrowers in **Grades E, F, G**:
  - Restrict lending to loans with **interest rate < 10%** to minimize risk.
  - Ensure **term = 36 months** for quicker repayment.

7. Term

- For **60-month loans**:
  - Limit to **interest rate < 10%** to maintain affordability.
  - Funded amounts should be **< $15,000** to avoid excessive exposure.

Business Recommendations

Avoid Financial Loss

1. Funded Amount

- **Smaller loans (< $15,000)**:
  - Smaller loan amounts have a higher likelihood of being fully repaid, reducing the risk of charge-offs.

2. Interest Rate

- **Lower interest rates (< 10%)**:
  - Loans with lower interest rates are more affordable for borrowers, increasing repayment rates and decreasing defaults.

3. Annual Income

- **Income > $80,000**:
  - High-income borrowers tend to have better financial stability, leading to consistent loan repayments.

4. Debt-to-Income Ratio (DTI)

- **DTI < 10**:
  - Borrowers with lower DTI ratios have less financial burden, making them more reliable in repaying loans.

5. Purpose

- **Avoid high-risk purposes** such as:
  - **Small business, renewable energy, debt consolidation, and home improvement**.
  - These categories have a higher likelihood of defaults compared to others.

6. Grade

- **Priority Grades**:
  - **Grade A and B borrowers** should be prioritized for lending due to their lower default risks.
  - **Secondary focus** can be given to Grades C and D, but with caution.

7. Term

- **Shorter term loans (36 months)**:
  - Shorter loan terms reduce exposure to long-term financial risks and increase the likelihood of complete repayment.

Thank You