Lending Club Case Study

Pallavi Dasgupta

Problem statement & Objective

- **Problem Statement -** The company specializes in lending various types of loans to urban customers. When a loan application is received, the company must decide whether to approve the loan based on the applicant's profile. The decision carries two types of risks:
    - **Loss of Business**: If the applicant is likely to repay the loan and the company does not approve it, the company loses potential business.
    - **Financial Loss**: If the applicant is likely to default and the loan is approved, it results in a financial loss for the company.

- **Objective**: The goal is to identify patterns that indicate if a person is likely to default on a loan. This will help the company take data-driven actions, such as denying the loan, reducing the loan amount, or lending at a higher interest rate to risky applicants. Ultimately, the company aims to reduce financial losses from defaults and improve its risk assessment strategies.

Business Understanding

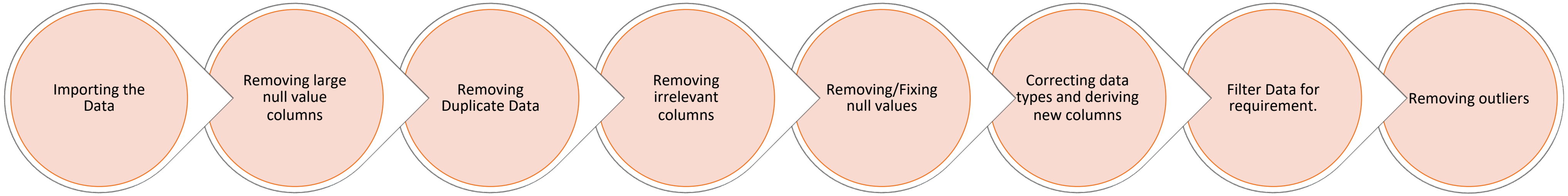
You work for a consumer finance company specializing in lending various types of loans to urban customers. When a loan application is received, the company must decide whether to approve the loan based on the applicant's profile. Two main risks are associated with this decision:

- Loss of Business: If the applicant is likely to repay the loan and the company does not approve it, it loses potential business.
- Financial Loss: If the applicant is likely to default and the loan is approved, it results in a financial loss for the company.

Dataset: **Loan Decision Outcomes**:

- **Loan Accepted**:
  - **Fully Paid**: Applicant has repaid the loan fully.
  - **Current**: Applicant is still repaying the loan (not yet completed).
  - **Charged-off**: Applicant has defaulted on the loan.
- **Loan Rejected**: No transactional history available.

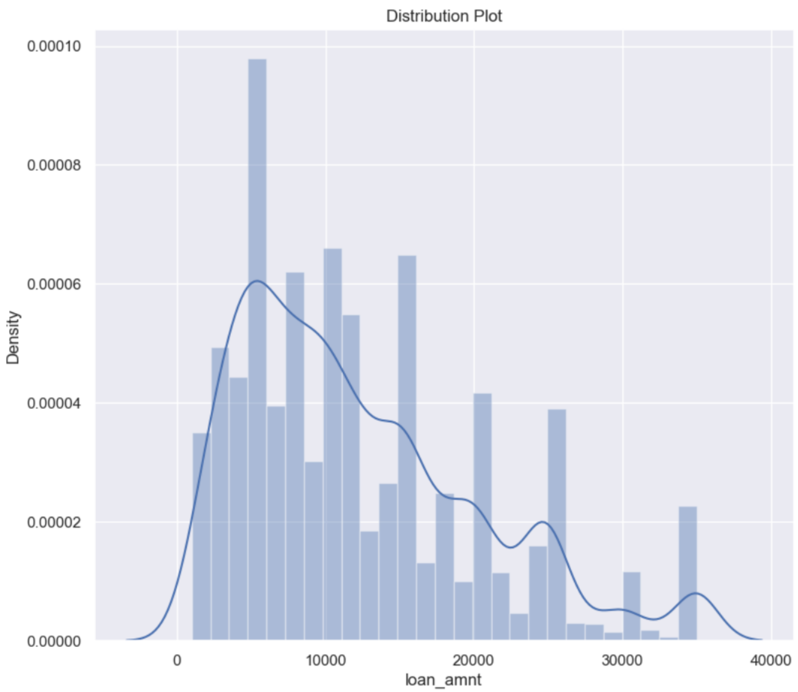
Data Clean-up and preparation process:



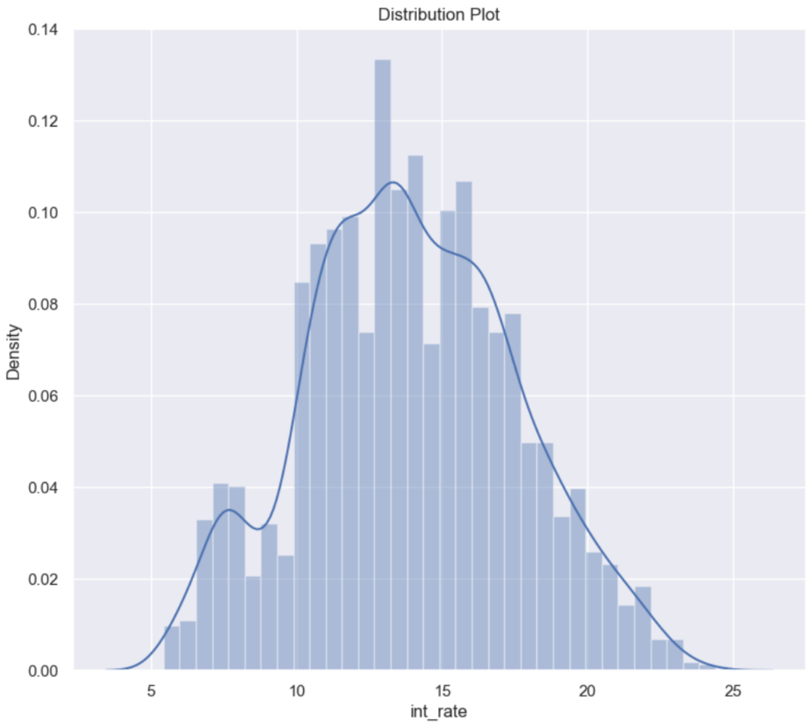
Approach

- **Solution Approach -** Using exploratory data analysis (EDA), the aim is to understand how consumer attributes (e.g., income, home ownership) and loan attributes (e.g., interest rate, loan amount) influence the tendency of loan default. By identifying the significant driver variables behind loan defaults, the company can optimize its loan approval process to minimize financial risks.
- **Desired Outcomes**:
  - Identify risky applicants who are likely to default.
  - Provide actionable insights to reduce financial loss and improve portfolio quality.
  - Optimize lending strategies to balance risk and business opportunities.

Data Insights-I

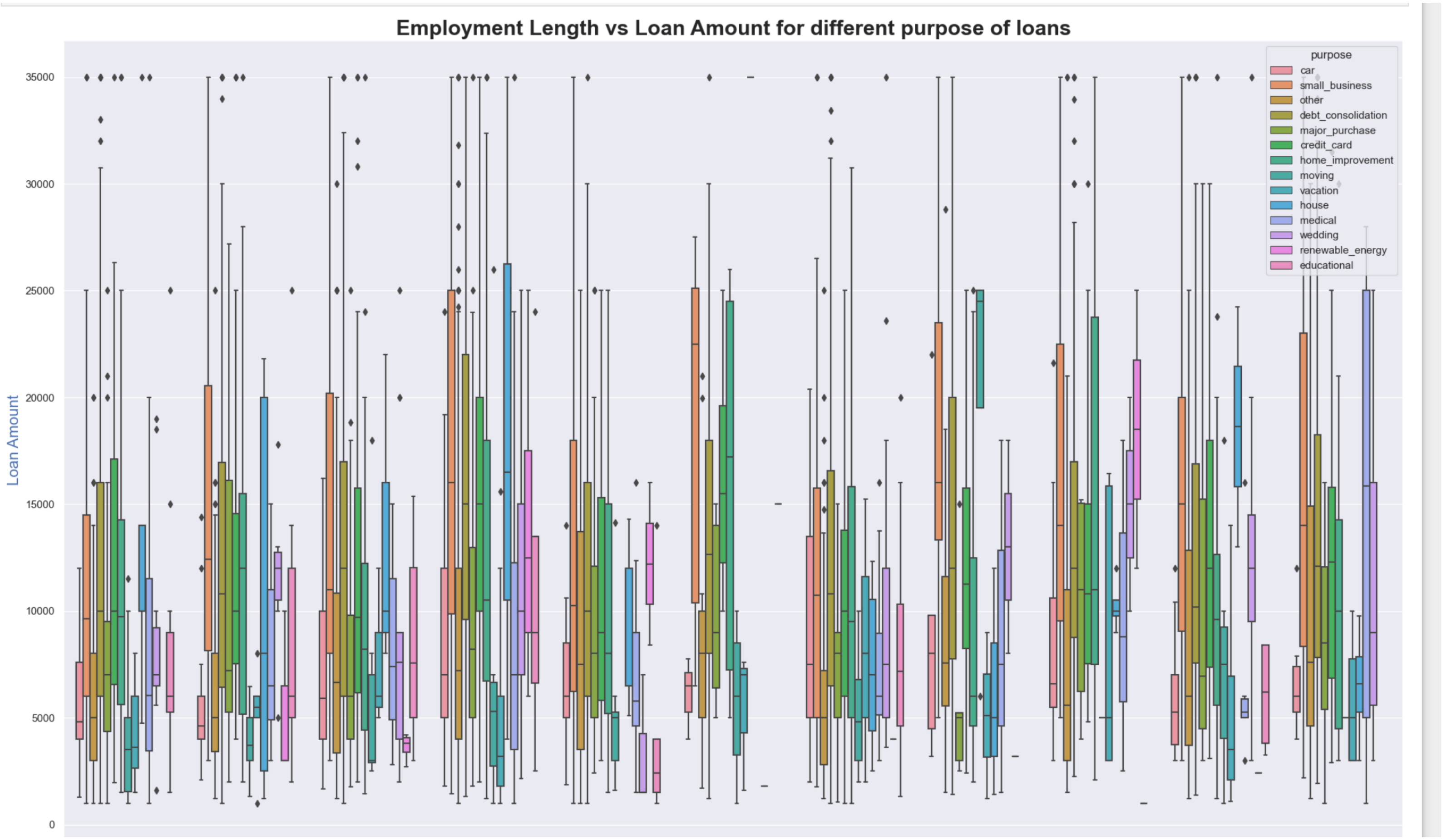


Most of the defaulted loan amounts are distributed between 6000 to 16000 USD.



Most of the defaulted loan interest rates are distributed between 11% to 17%.

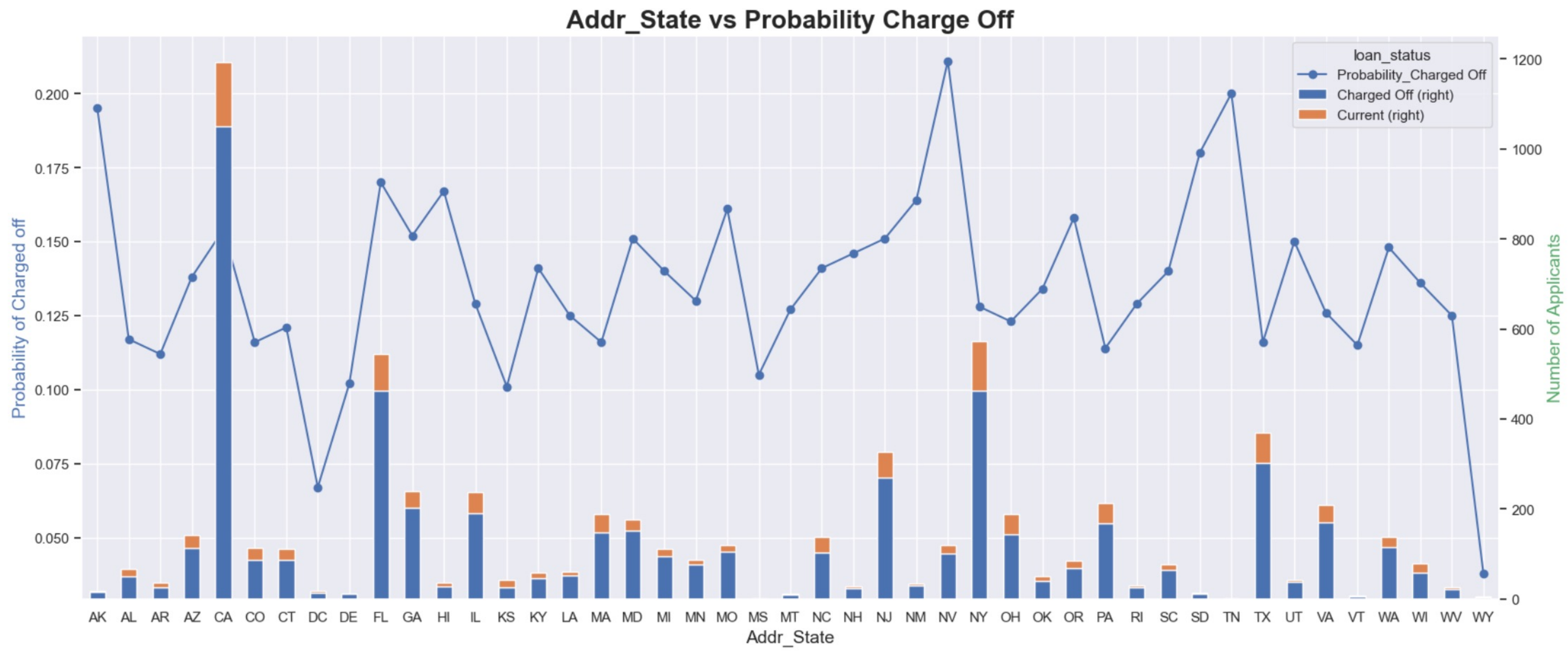
Data Insights-II



Insights:

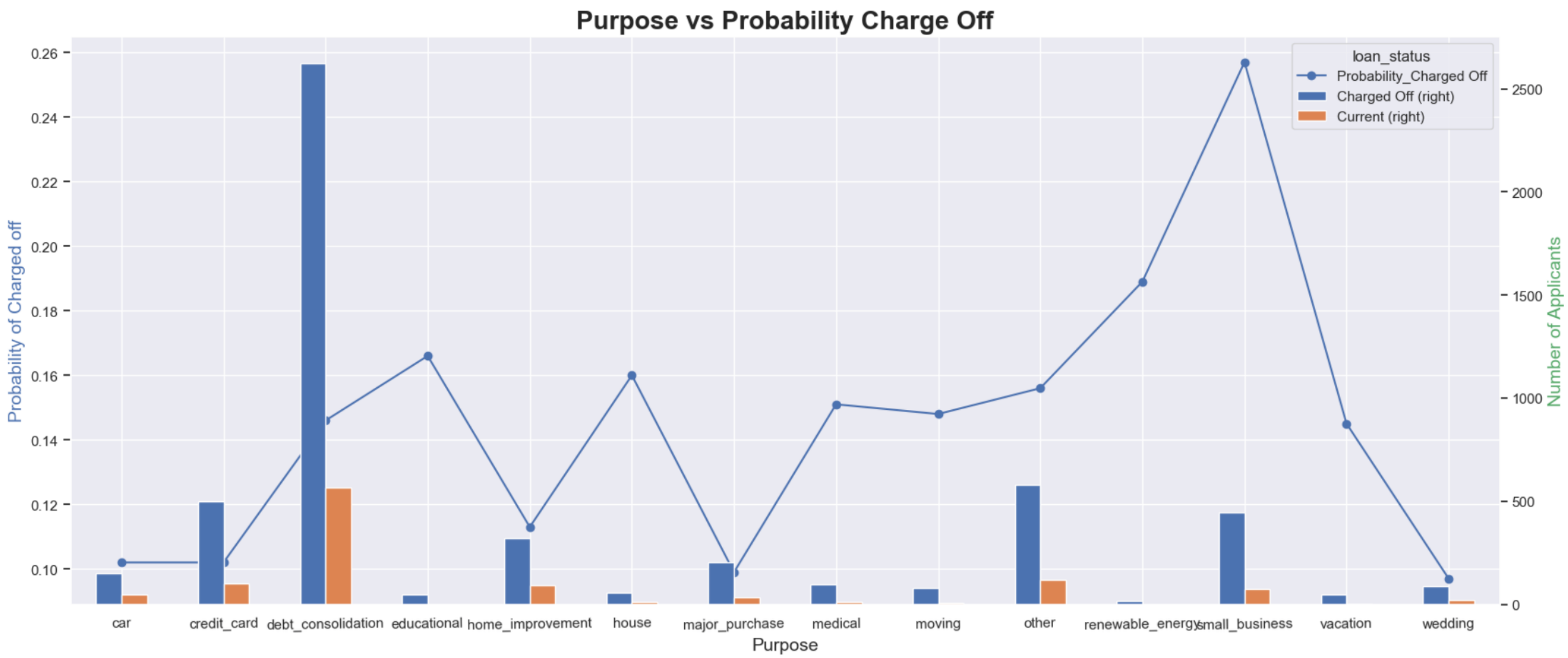
Customers with more than 10+ years of employment tenure and loan amounts more than 150K USD specifically applied for small businesses and houses are likely to be defaulters more than the rest of the lot.

Data Insights-III



For large metropolitan cities we see large number of loans, with higher number of defaulted loans from California, New York, Texas, Florida.

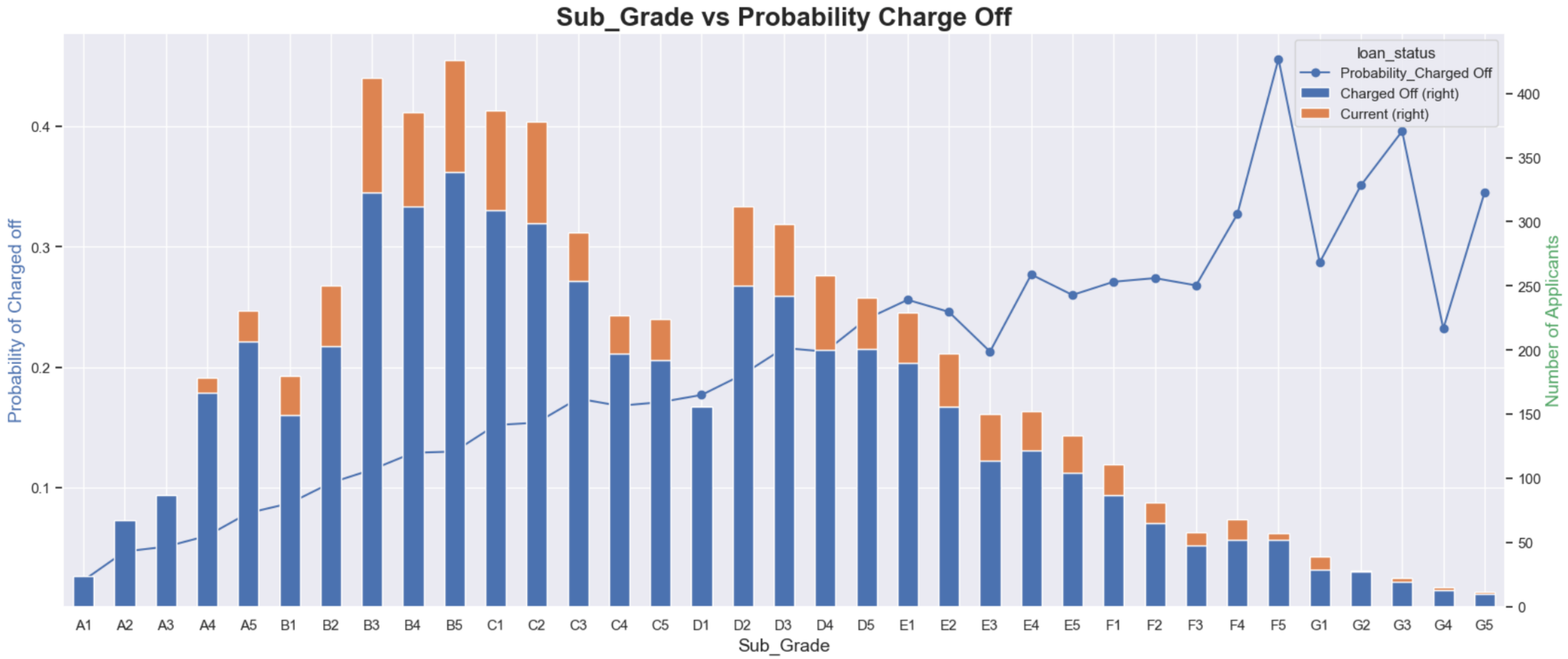
Data Insights-IV



Observations:

Applicants who have loan for 'small business' has the highest probabilty of charge off of 26%. So bank should take extra caution like take some asset or guarentee while approving the loan for purpose of 'small business'

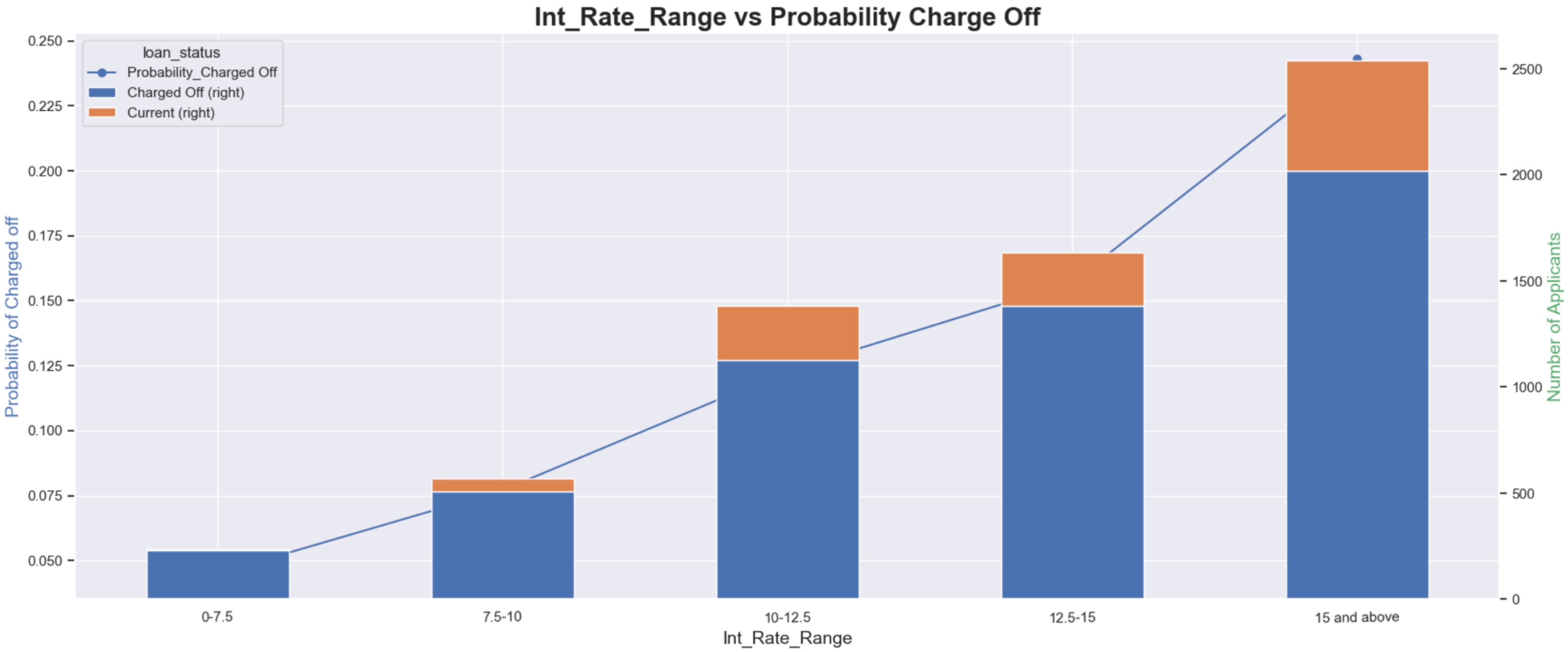
Data Insights-V



Insights:

As we move from Grade A to G, probability that cutomers will charged off is increasing.

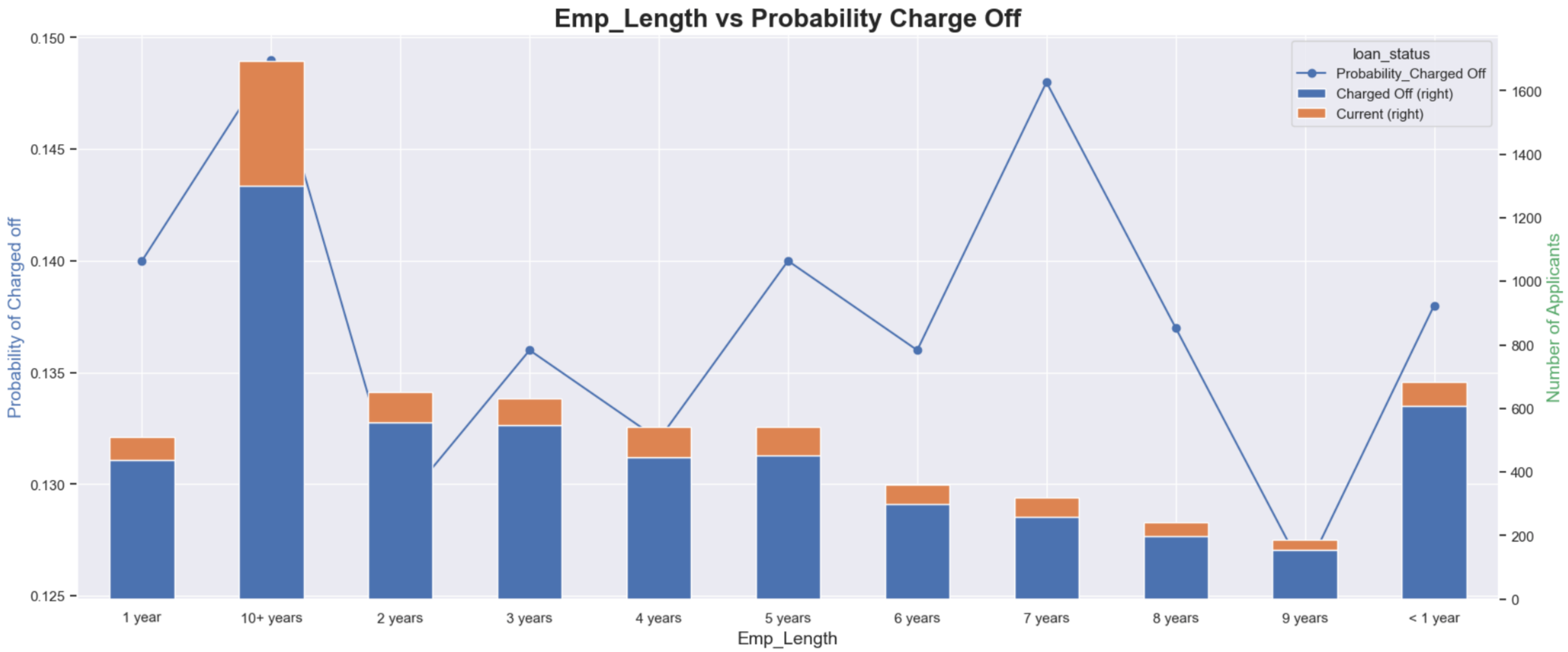
Data Insights-VI



Insights:

As the interest rate is increasing the probability that person will default is increasing with highest at 23% at 15% & above bracket.

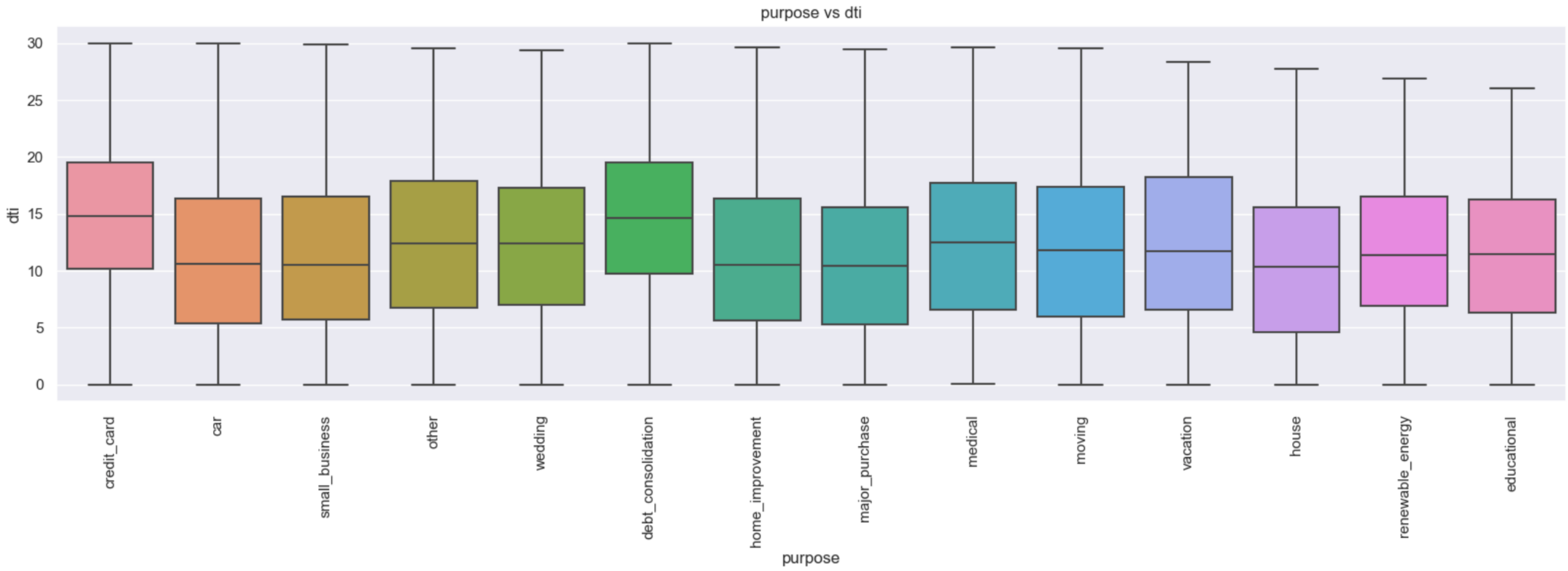
Data Insights-VII



Insights:

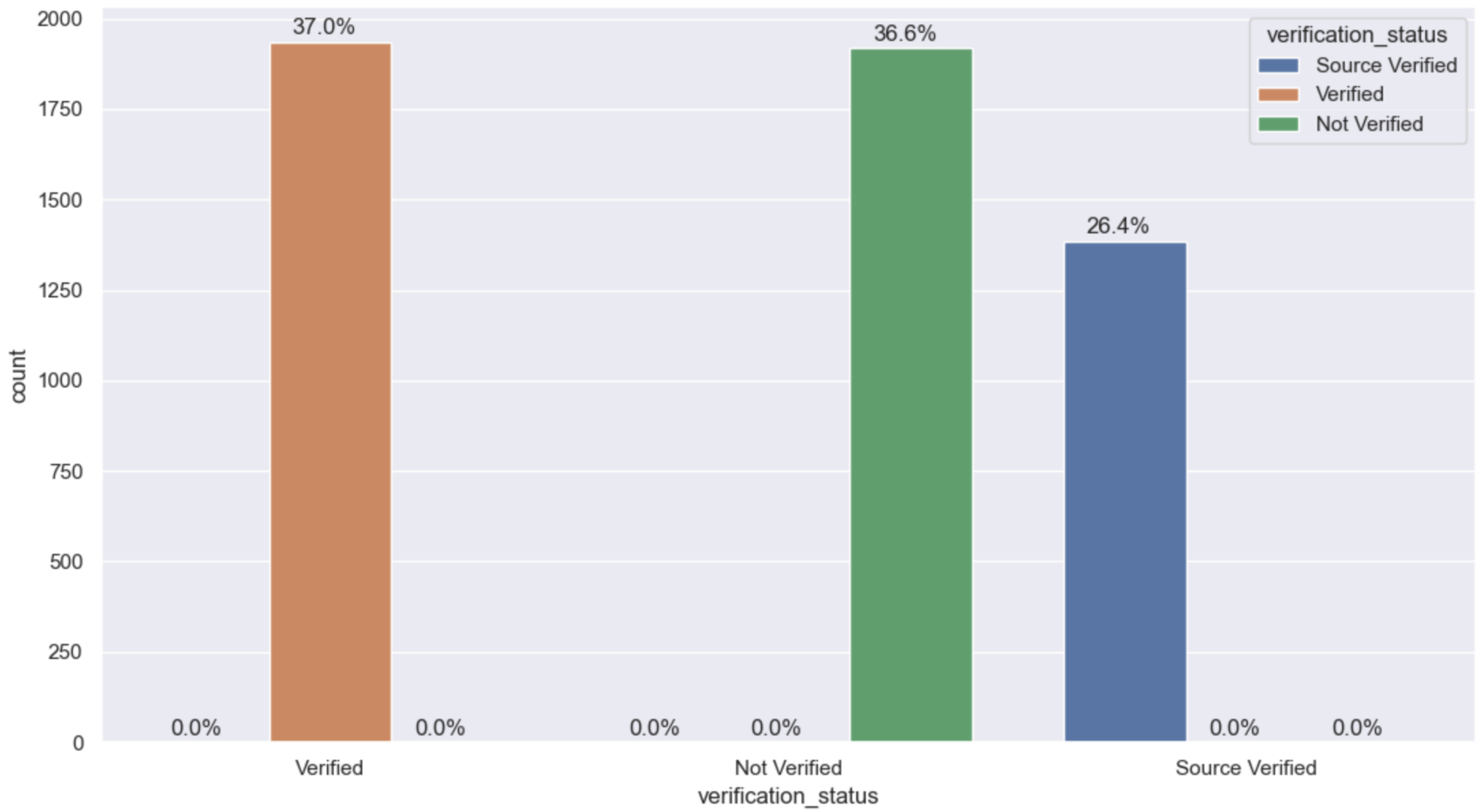
Applicants who have 1 & less than 1 year of experience are more likely to default, along with the highest amount defaulters at 7 and 10+ years of employment length

Data Insights-VIII



Observations:
People who took loan for *credit card* and *debt consolidation* purpose has more DTI than other purposes.

Data Insights-IX



Insights:
Majority of loans were given without verification of applicants income.

Summary & Recommendations

Driving factors which can be used to predict the chance of defaulting and avoiding Credit Loss:

- Customers with more than 10+ years of employment tenure and loan amounts more than 150K USD specifically applied for small businesses and houses are likely to be defaulters more than the rest of the lot.
- Applicants who have loan for 'small business' has the highest probability of charge off of 26%. So bank should take extra caution like take some asset while approving the loan for purpose of 'small business’
- Most of the defaulted loan interest rates are distributed between 11% to 17%.
- Applicants who have 1 & less than 1 year of experience are more likely to default, along with the highest amount defaulters at 7 and 10+ years of employment length

Other considerations for 'defaults' :

1. Borrowers not from large urban cities like California, new york, texas, florida etc.
2. Borrowers having annual income in the range 50000-100000.
3. Borrowers having Public Recorded Bankruptcy.
4. Borrowers with least grades like E,F,G which indicates high risk.
5. Borrowers with very high Debt to Income value.