LENDING CLUB CASE STUDY

Presenter:
Shweta Rao
Pragadeesh S

Case Study Description
Finding the Defaulter:

Identification of Loan Applicant traits that tend to default

Step 1:

Finding the Driving factors for the defaults

Step 2:

Providing the finding to Organisation to regularise and take actions agains each of them attribute causing the defaulters

Step 3:
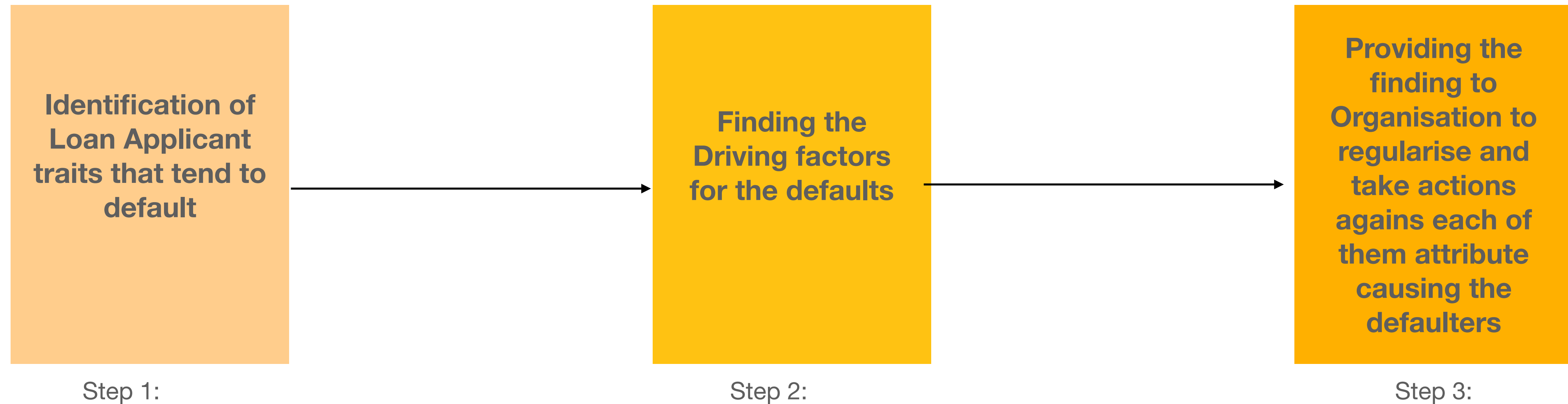
Steps followed to get to the Conclusion

Data Cleaning Steps

- **Delete Columns** : Delete unnecessary columns i.e. having all values as Null/NaN/0 , irrelevant columns like url,'desc' etc

- **Delete Rows:** Delete data from row with 5% and more null value.

- **Remove outliers:** Divide data into quantile which will give high and low values that would be inappropriate /disproportional.

- **Filter rows:** Filter by segment, filter by date period to get only the rows relevant to the analysis.
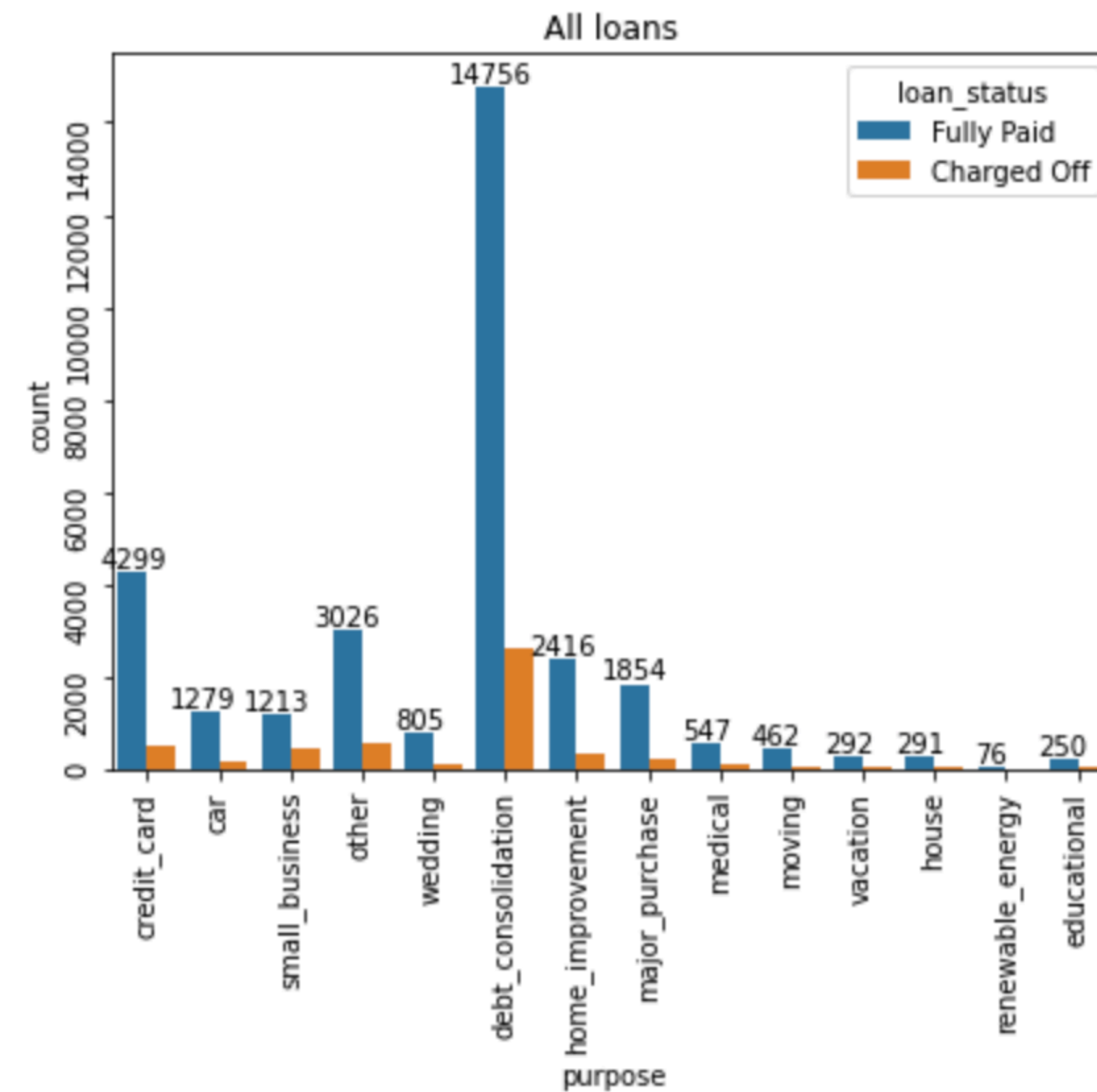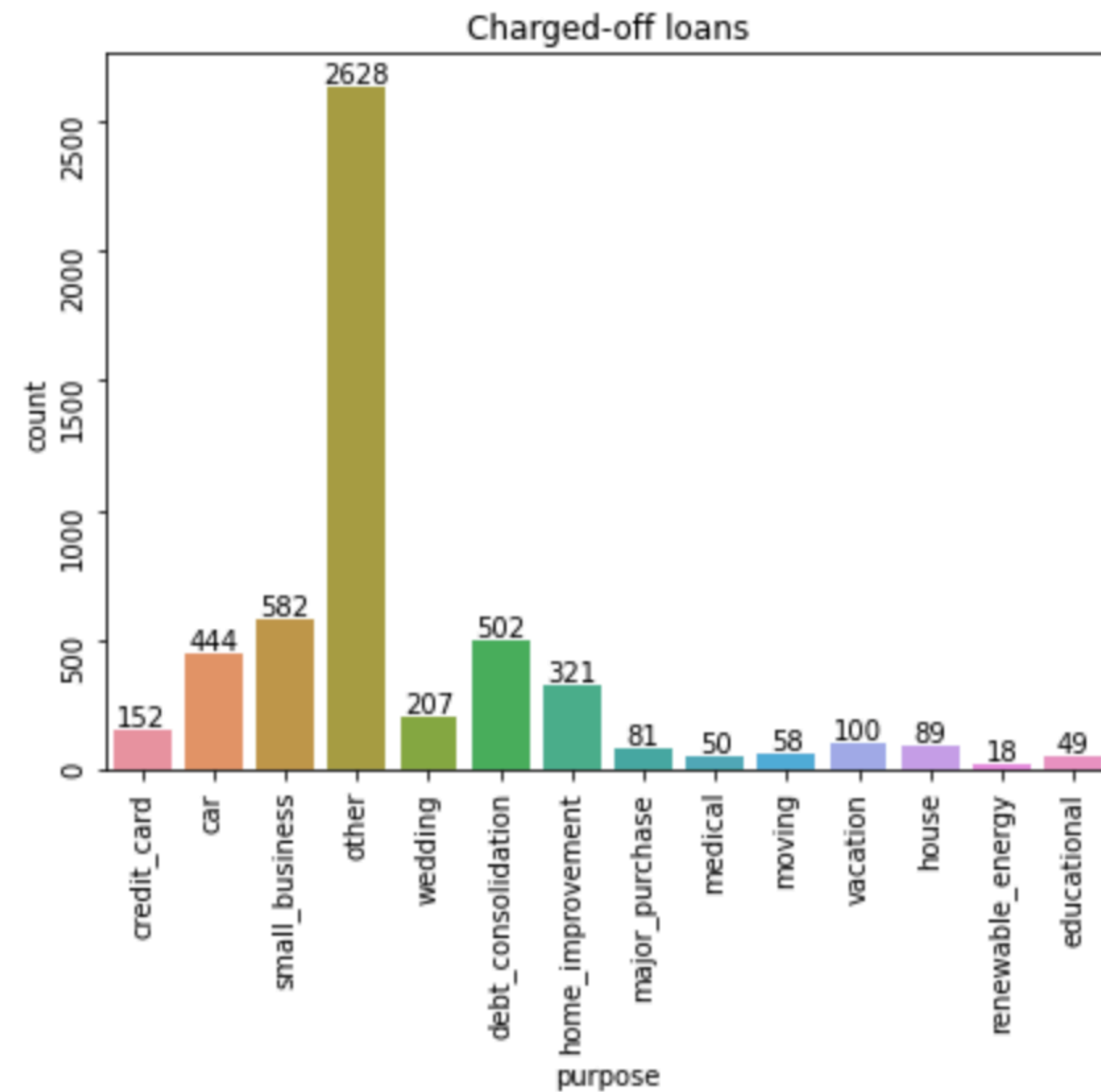
Loan Status

```
1  # checking Loan_status count
2  ax=sns.countplot(x='loan_status', data=df)
3  for container in ax.containers:
4      ax.bar_label(container)
```
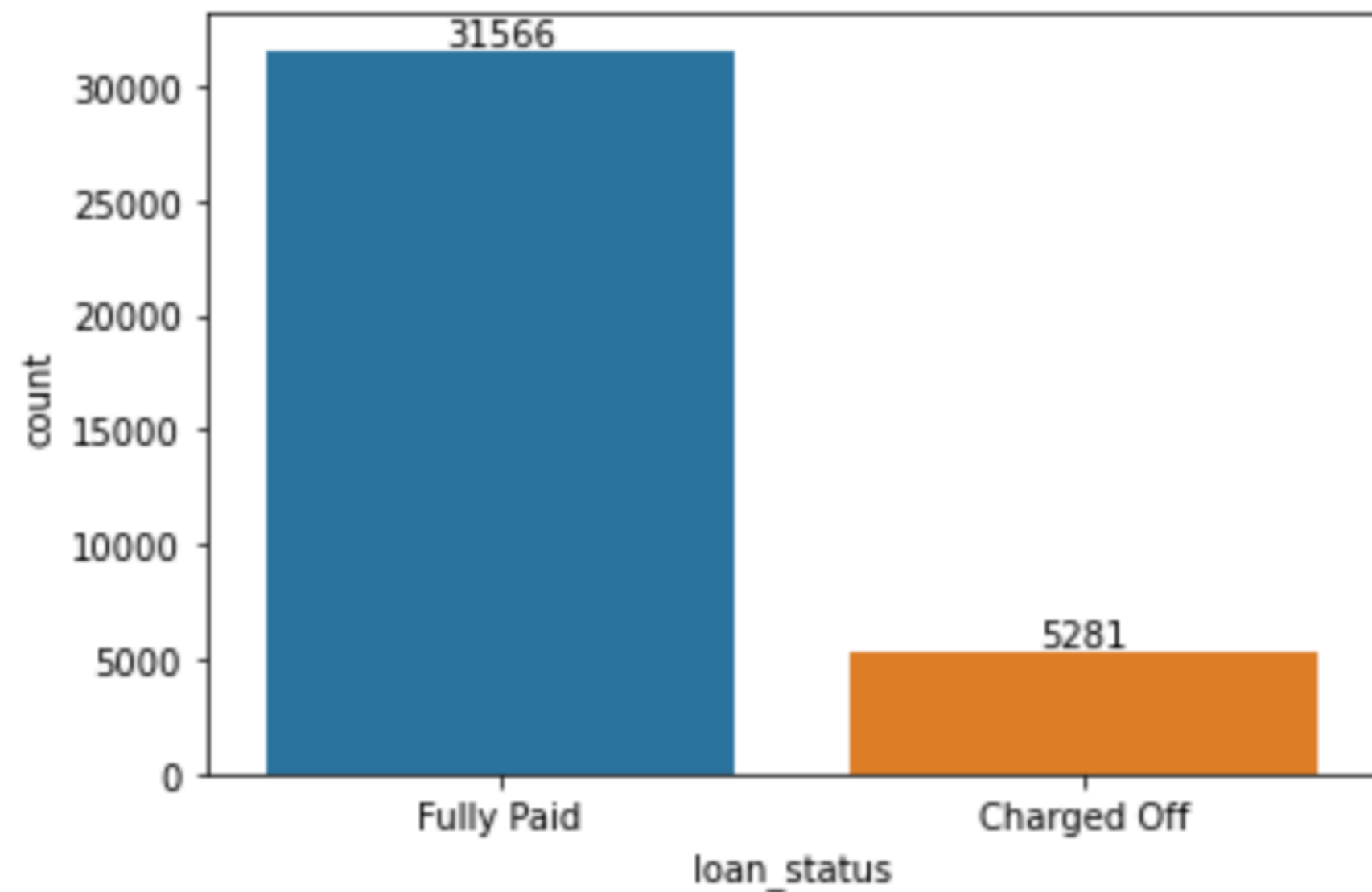

- **Observations:**

  - Most of the loans are Fully Paid

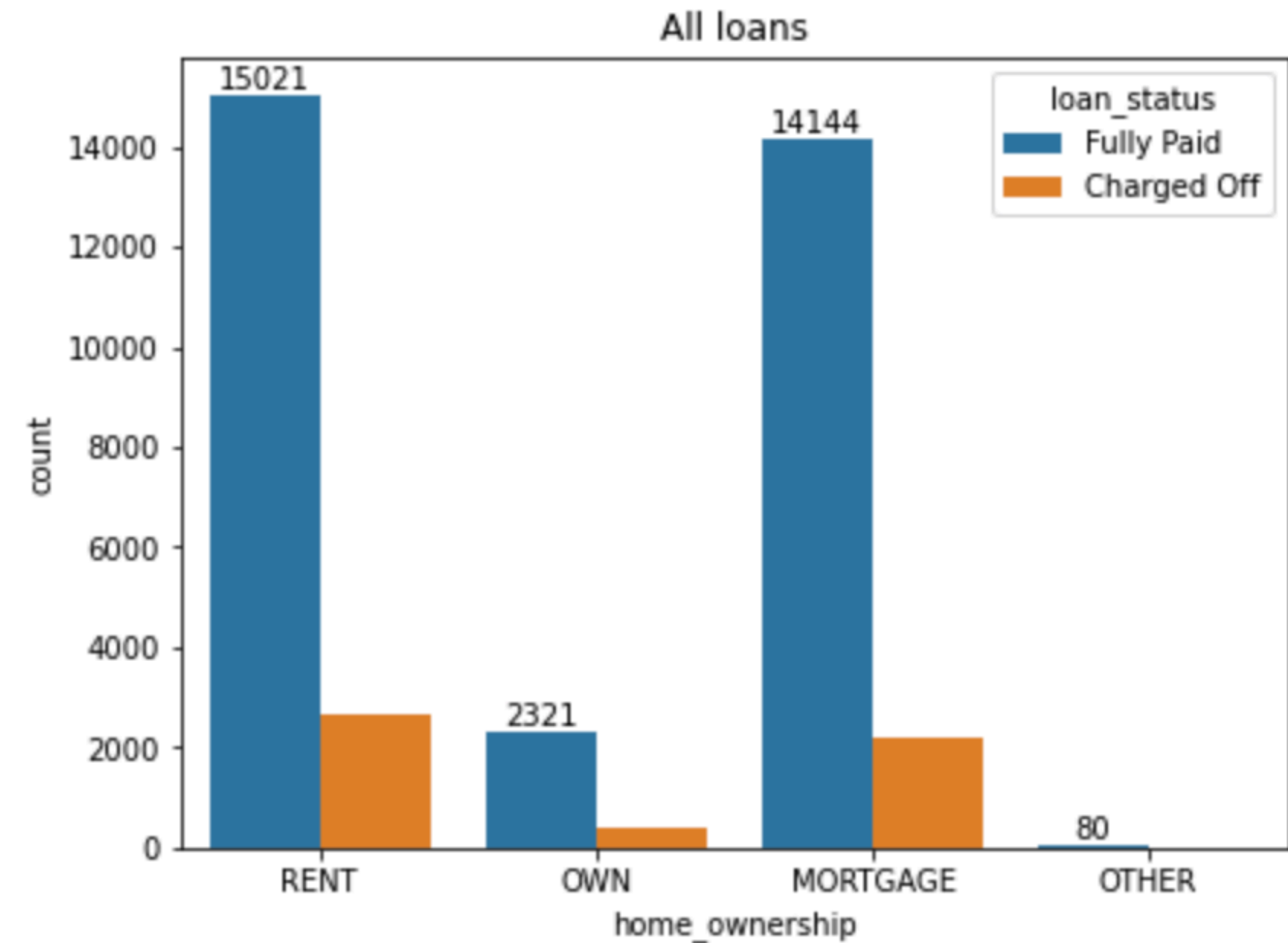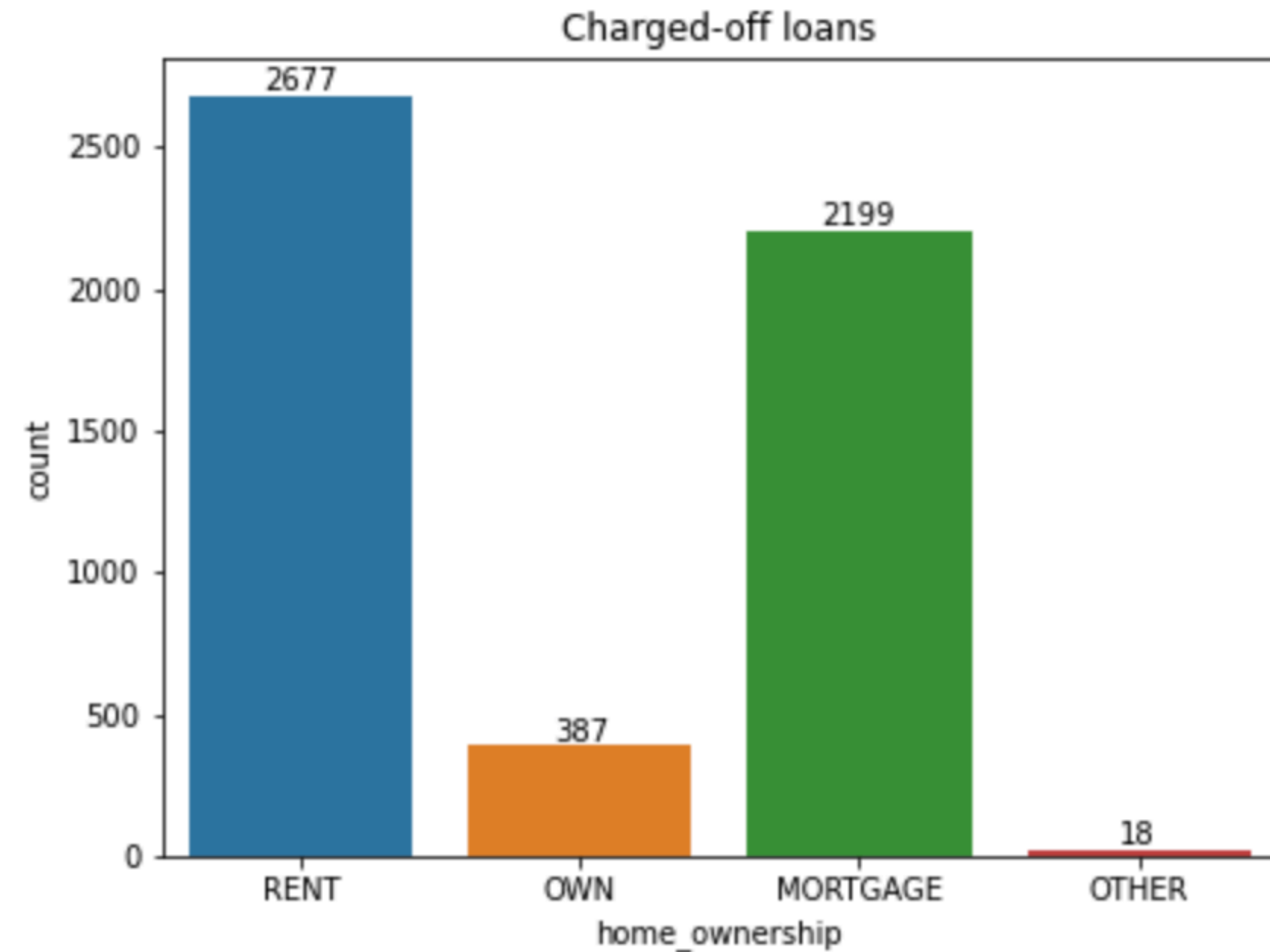 - About 14% of the loans are defaulters

Loan Purpose Status

- **Observations:**

  - Most of the Loans that is Charged-off is with the purpose of 'others'

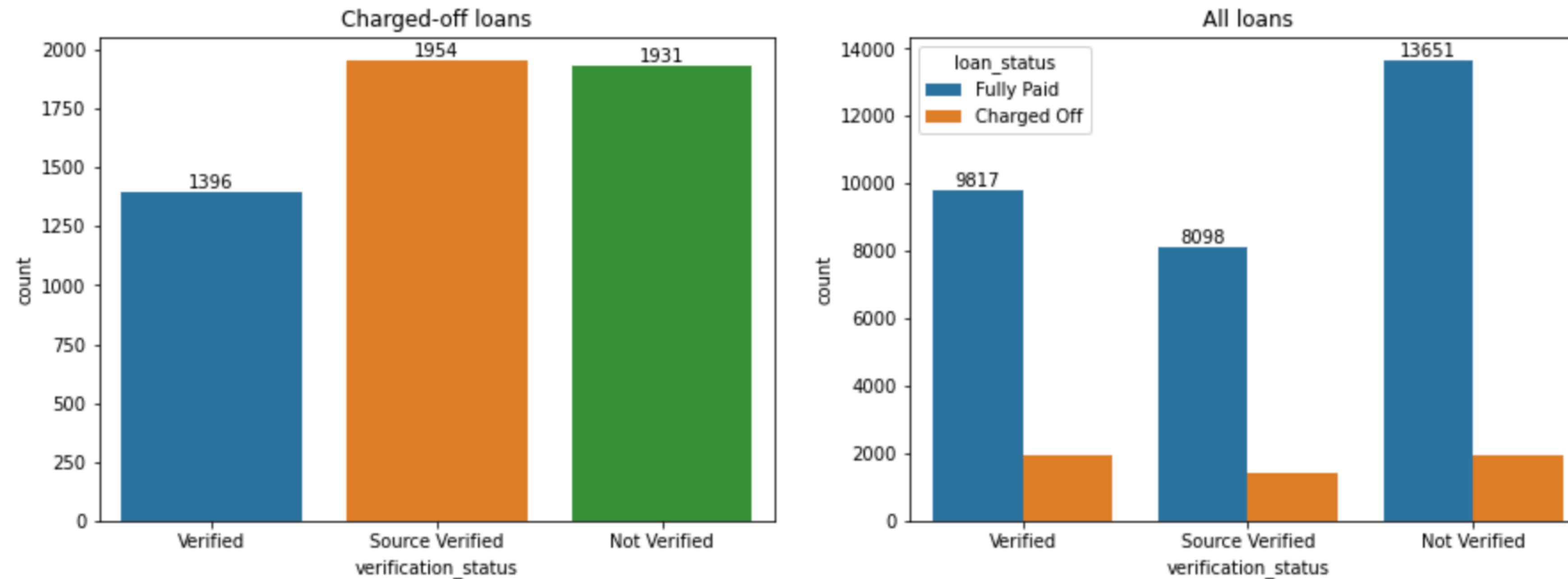 - Maximum loan defaulters fall under the purpose of 'debt_consolidation'

Loan Home Ownership Status

- **Observations:**

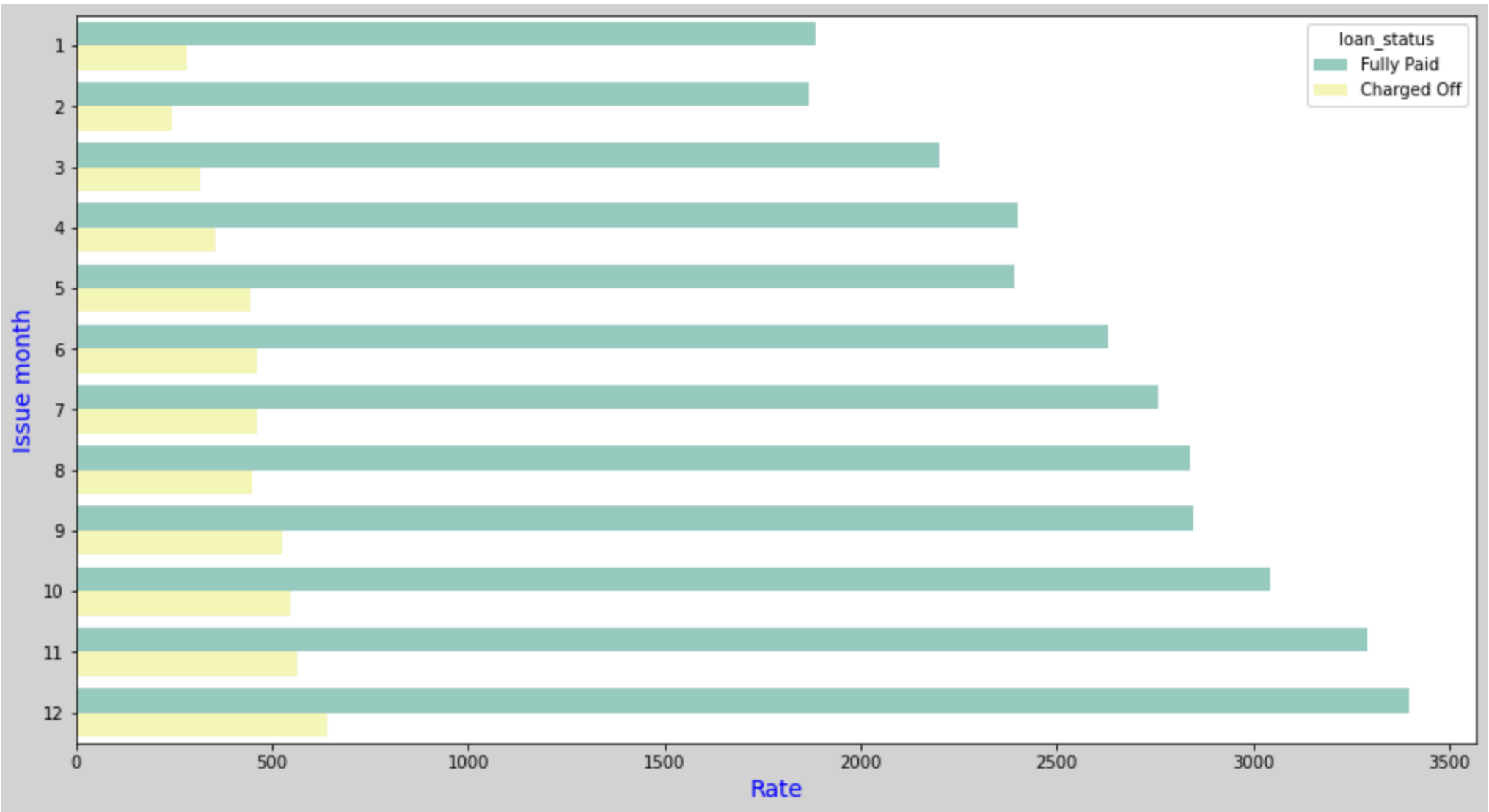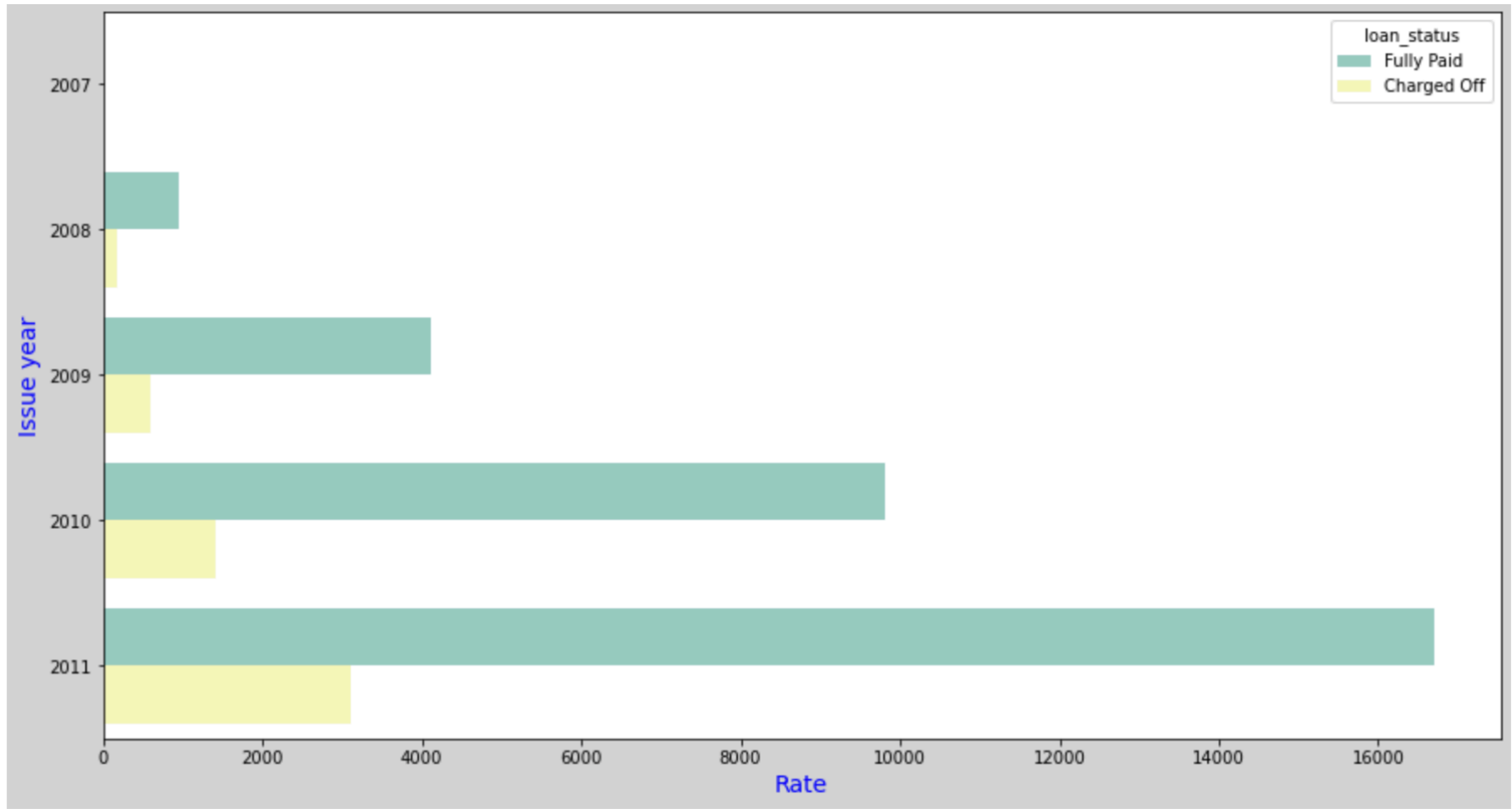 - Most of the Loans that is Charged-off is with the 'home_ownership' is due to **'RENT'** and **'MORTGAGE**
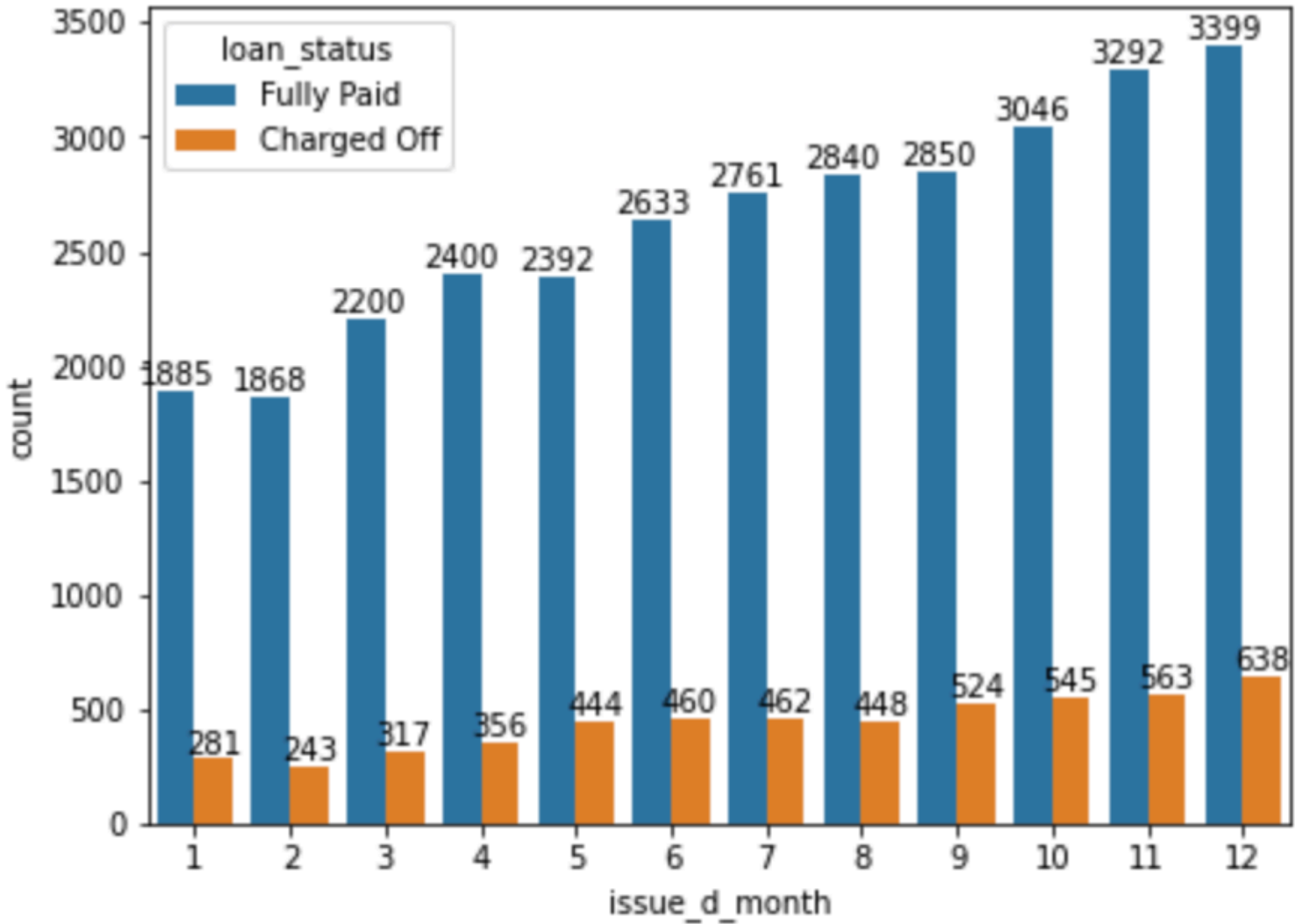
Loan VerifiedStatus
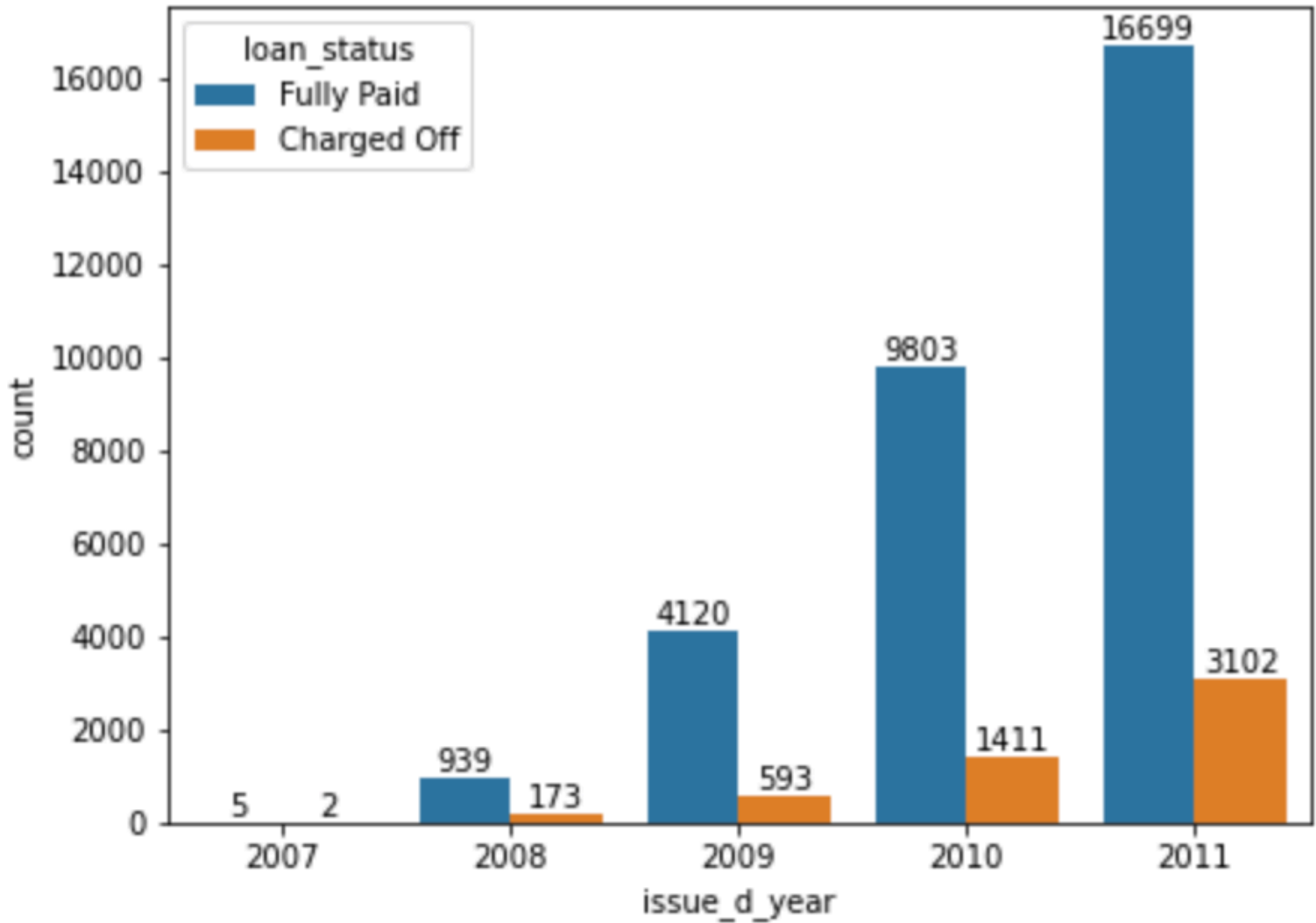

- ## Observations:

  - Most of the Loans that is Charged-off is with the 'verification' is due to **'Source Verified"**
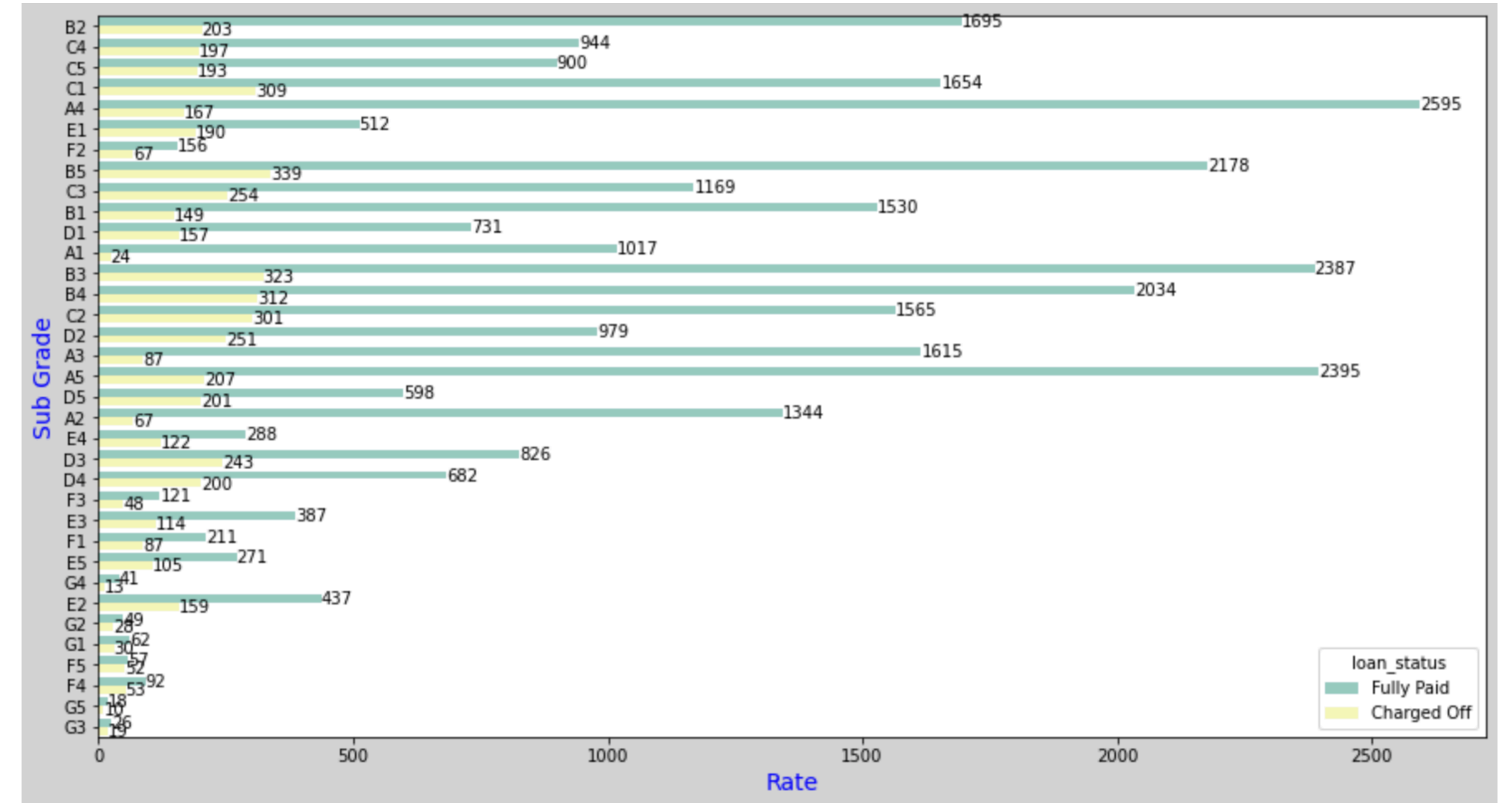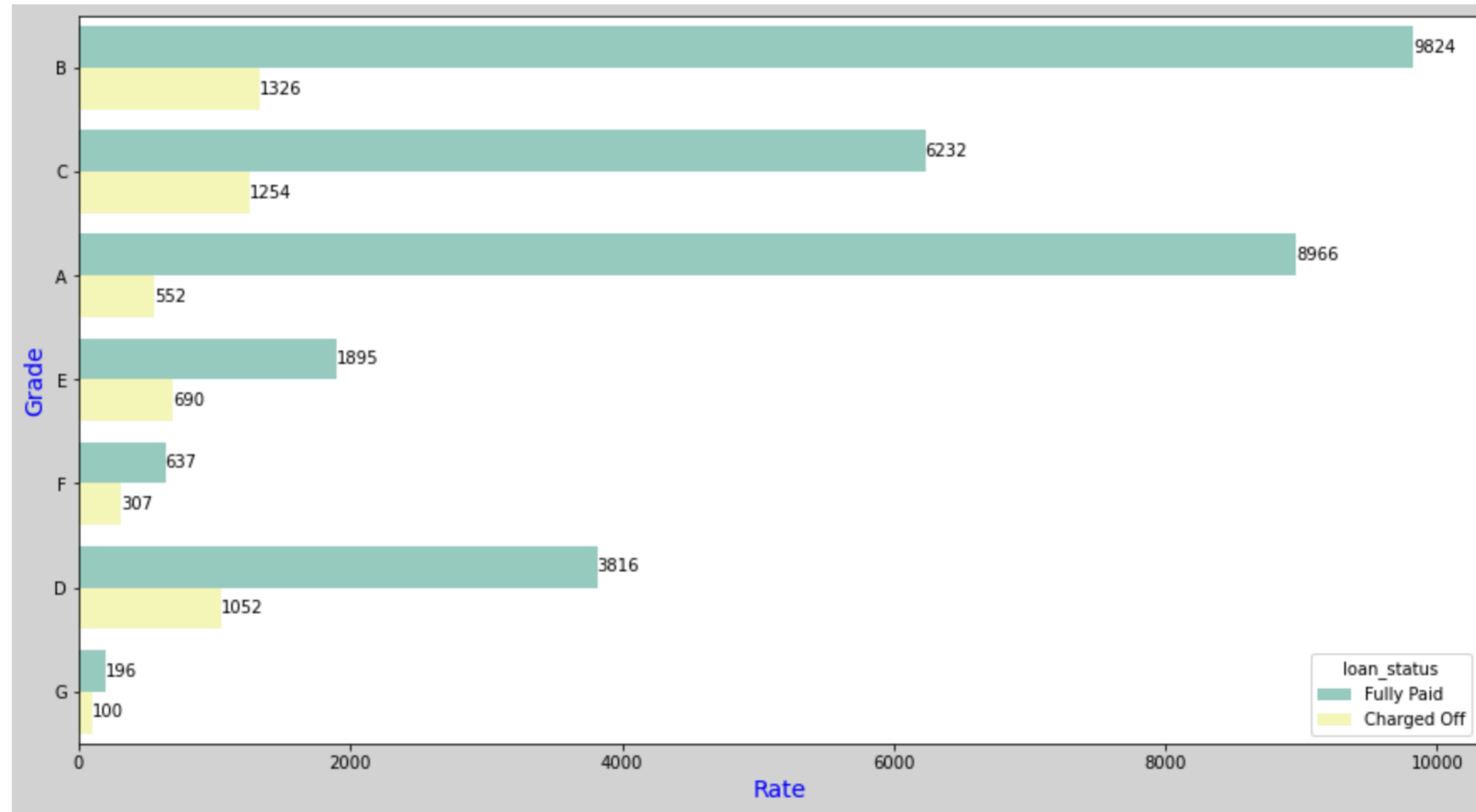
Loan Issue year and Issue month

- **Observations:**

  - Most of the Loans that is Charged-off is with the loan 'issue_year' is "2011" and "issue_month" is the year ending month Nov and Dec

Loan grade and Sub_grade

- **Observations:**
  - Grades B & C has almost 25% of defaulters for each grades.
  - Grades F & G, the number of loans are less but more than 50% of them are defaulted
  - From grade analysis, we know that A & B have many loans and B & C have high defaulter rate.
  - Sub_groups B3,B4,B5,C1,C2 have high defaulter rate.
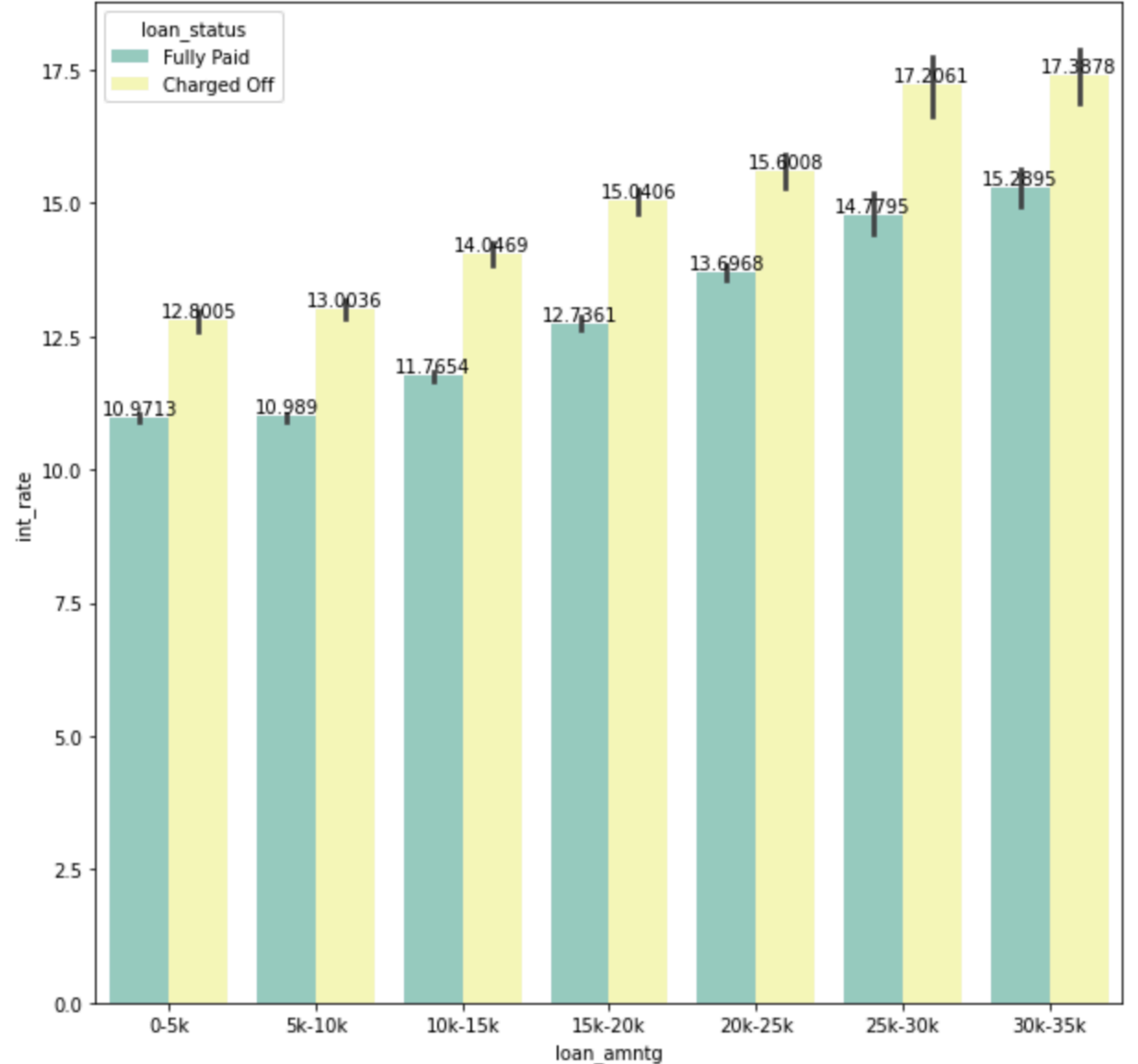
Loan analysis for loan_amnt and purpose

- **Observations:**
  - Category loan for purpose 'small_business' , 'house' and 'debt_consolidation' type against loan_amt are likely to default.
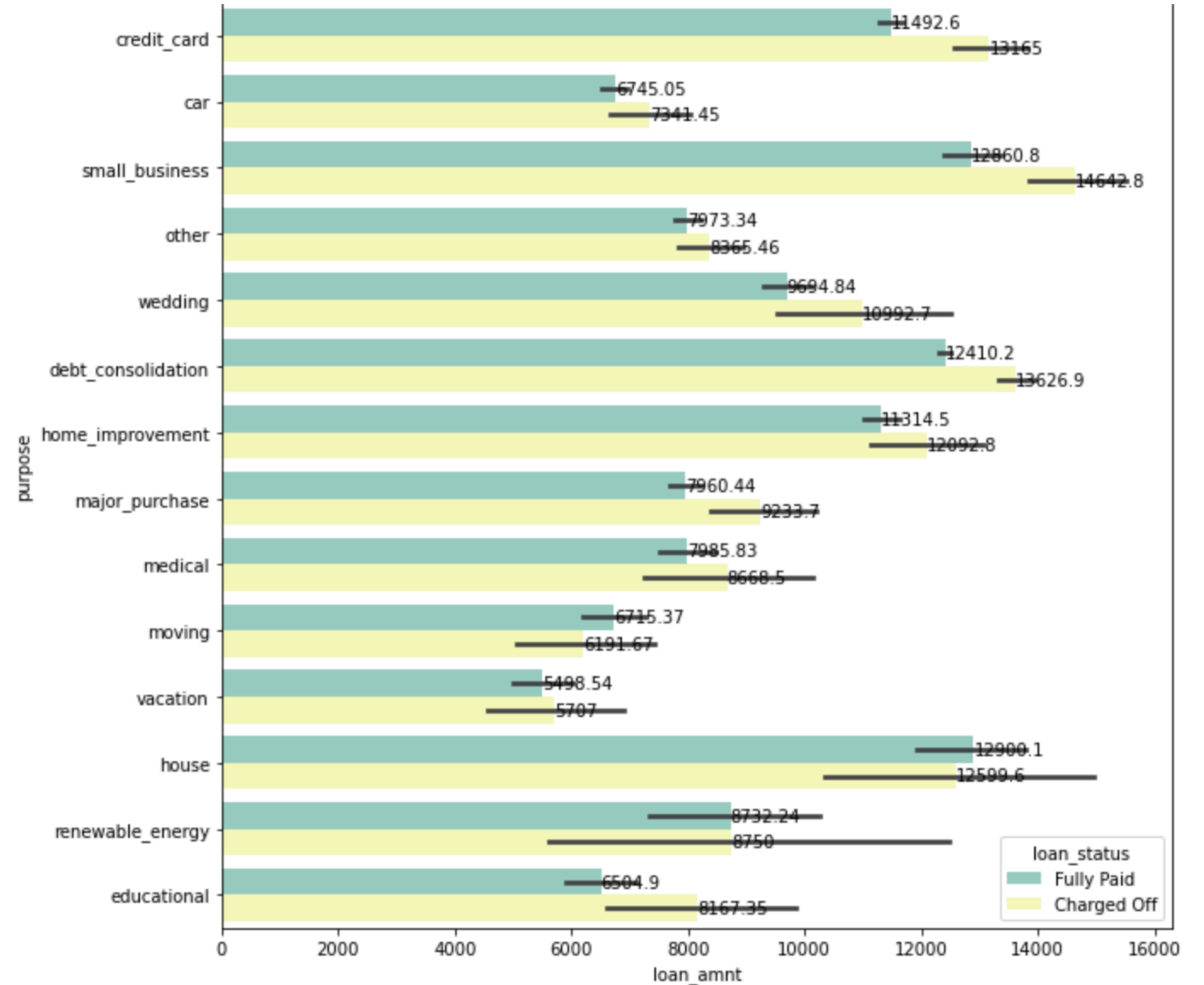
Loan analysis for loan_amnt and int_rate

- **Observations:**
  - Loan amount with the range asc'25k-30k' and '30k-35k' and with 'int_rate' of over and above 16.5

Loan analysis for annual_increment and loan_amnt

- **Observations:**
  - Loan applicants with annual income less than 50k getting loans above 25k which is 50% of their annual income. These is a high chance for these type of loan getting defaulted.
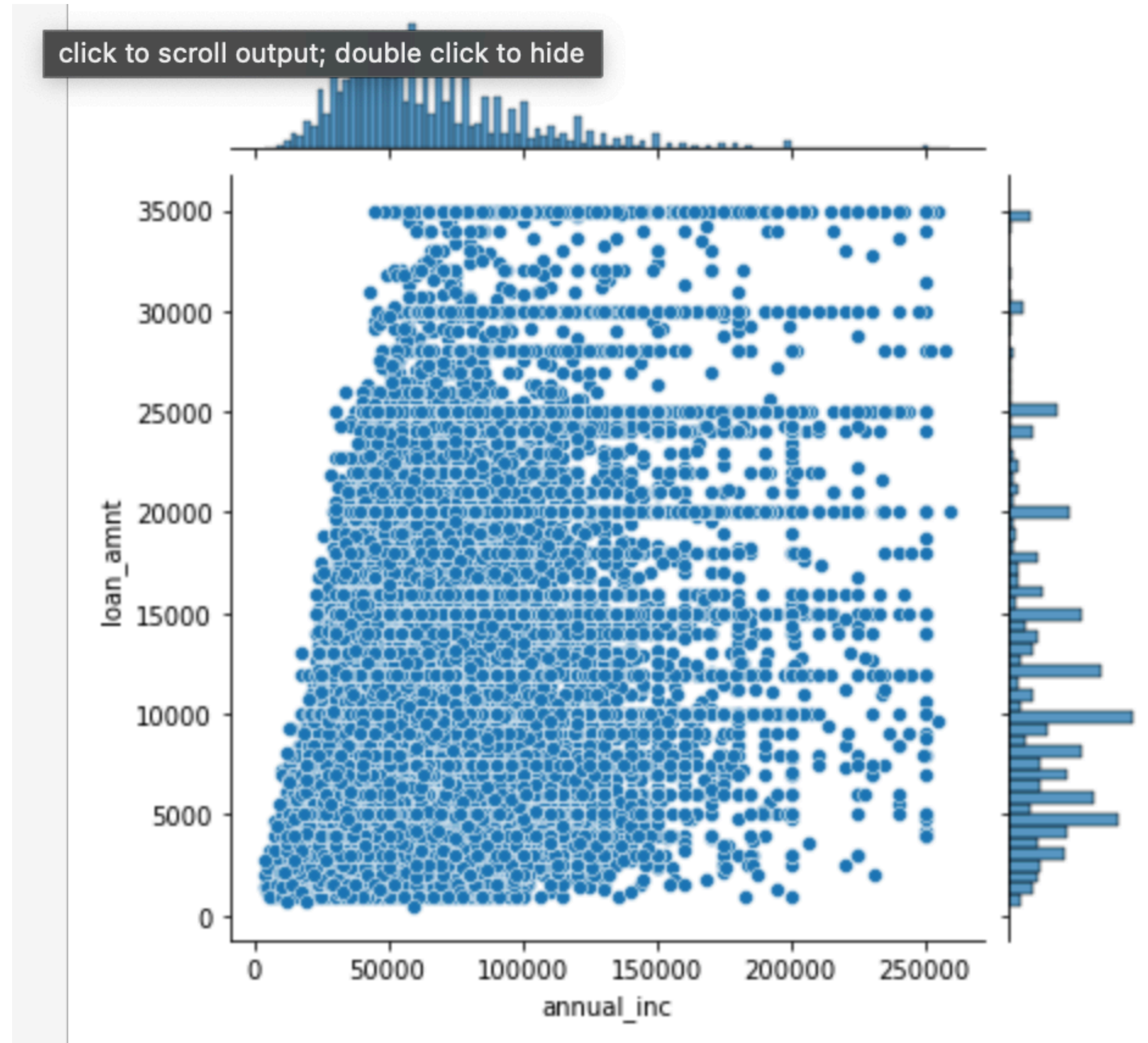
Conclusion

- For Grading and Sub-Grading: B&C have higher Defaulter rate, 3/4/5 B and 1/2 C sub-Grade are likely to default.
- Maximum Defaulter are with purpose of 'debt-consolidation' , analysing the application and applicant is important.
- Loans having high interest rate is risky are they tend to default.Check the background of the applicant .
- Applicant with annual income of 50k gets the loan of half of the amount which is a risky approval.These applicant tends to default.Try to have a higher margin of the loan coverage with respect to the annual income.