Lending Club
Case Study

In this project we are to analyse the risks for a consumer finance company while approving loans for applicants.

- The biggest risk for any lending company is to lend money to people who may not pay it back. This is called credit loss, and it happens when the borrower defaults on the loan. The customers who are marked as Charged Off are the ones who have defaulted on their loans.

- Analysing the data we need to find patterns and should be able predict which applicants in future are likely to pay back the loan, and which are more likely to default.
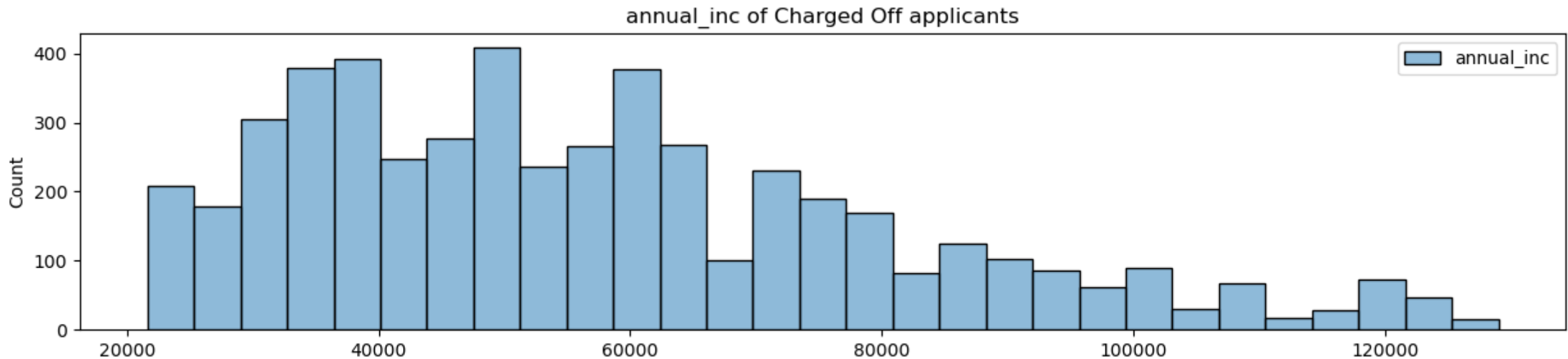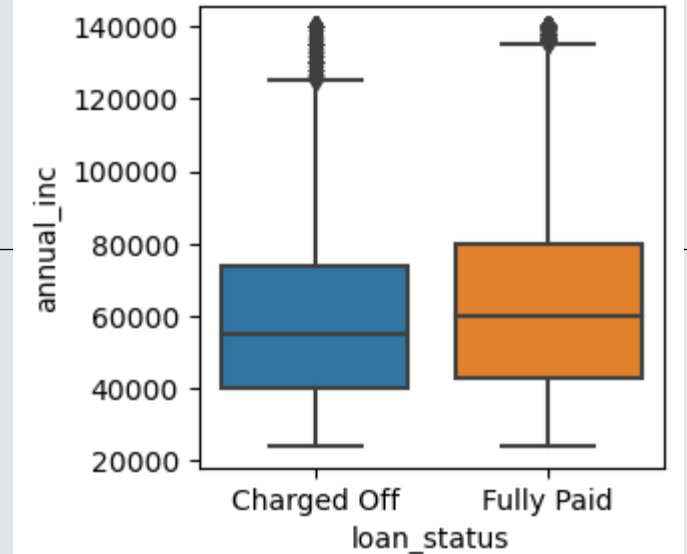
Numerical Data

Interest Rate (int_rate) is a strong indicator, rate between 11 and 16 indicates more chances of defaulting
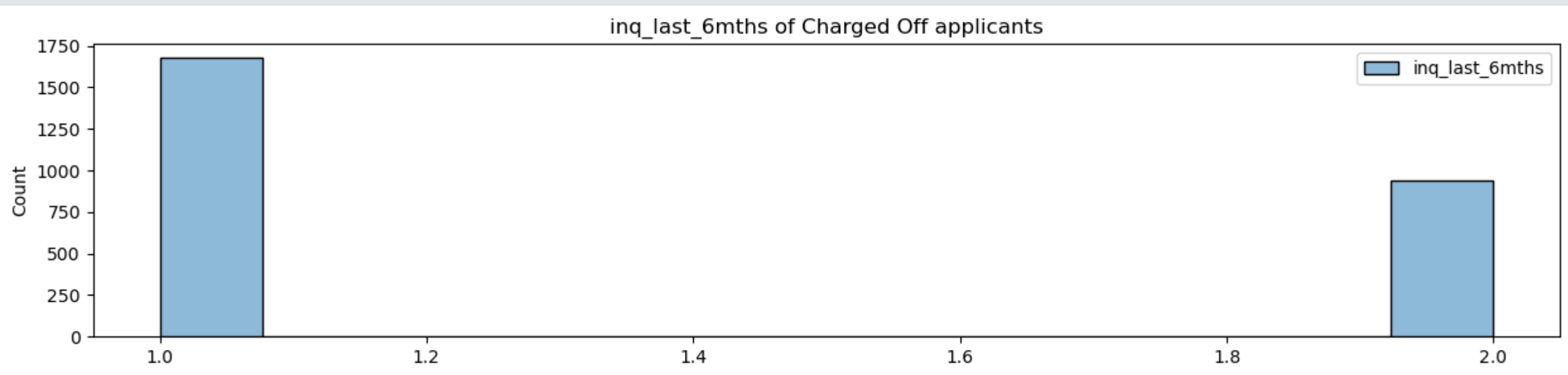

int_rate of Charged Off applicants
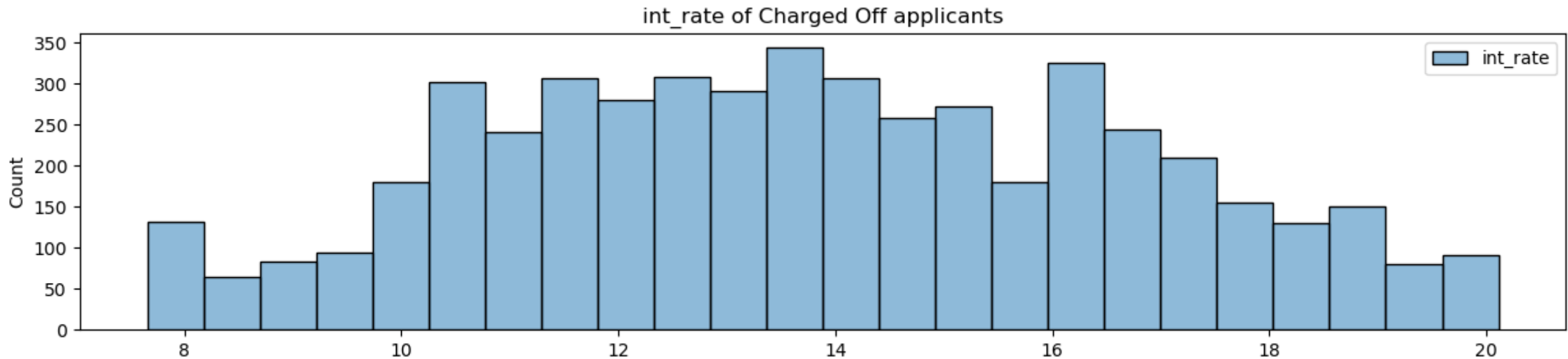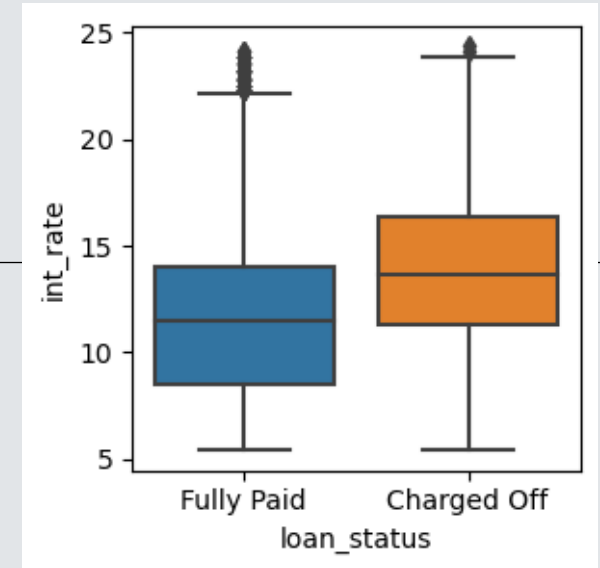
Annual Income (annual_inc) is a strong indicator, lower amount indicates more chances of defaulting

annual_inc of Charged Off applicants

Debt–To–Income Ratio (dti) is a strong indicator, higher ratio indicates more chances of defaulting

dti of Charged Off applicants
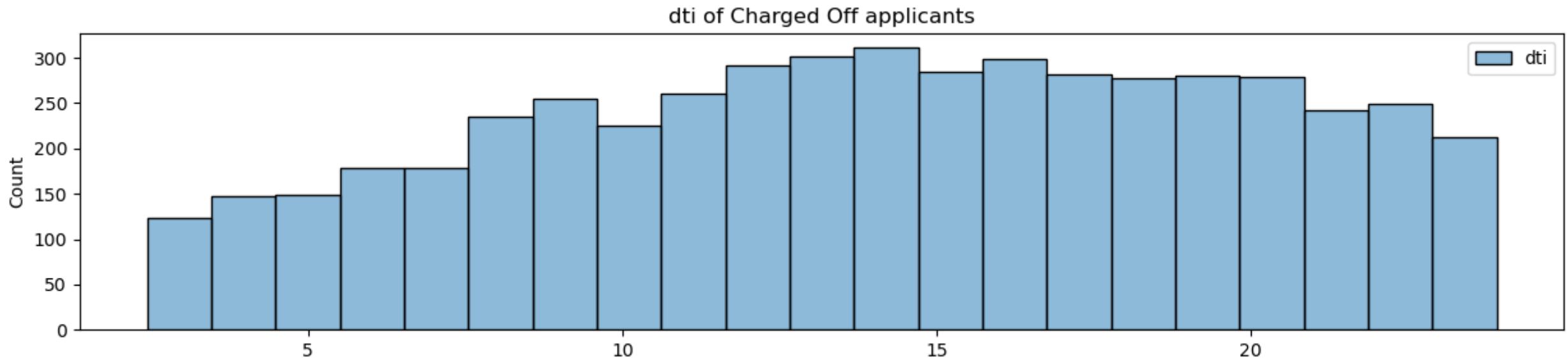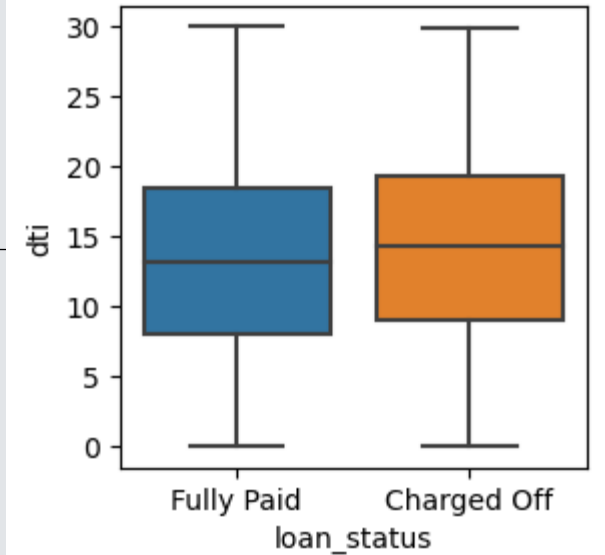
Inquiries in past 6 months (inq_last_6mths) is a strong indicator, more inquiries indicates more chances of defaulting
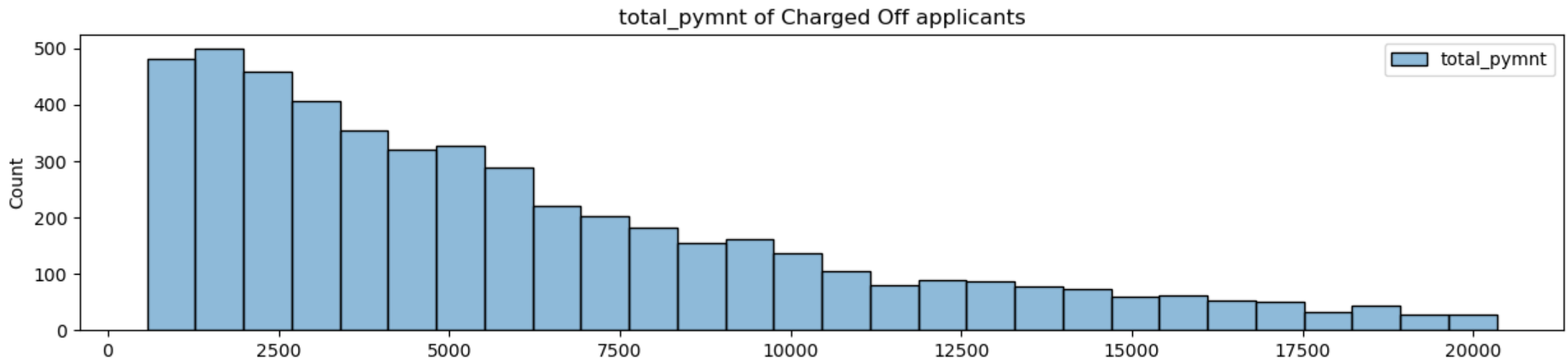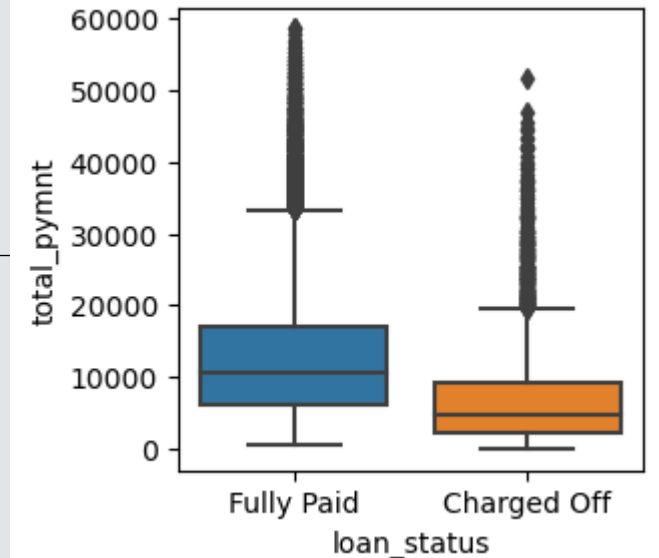

inq_last_6mths of Charged Off applicants
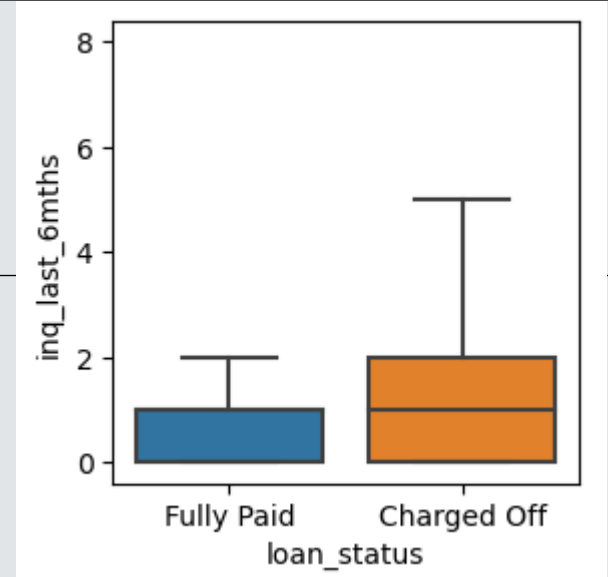
Revolving line utilization rate (revol_util) is a strong indicator, higher rate indicates more chances of defaulting
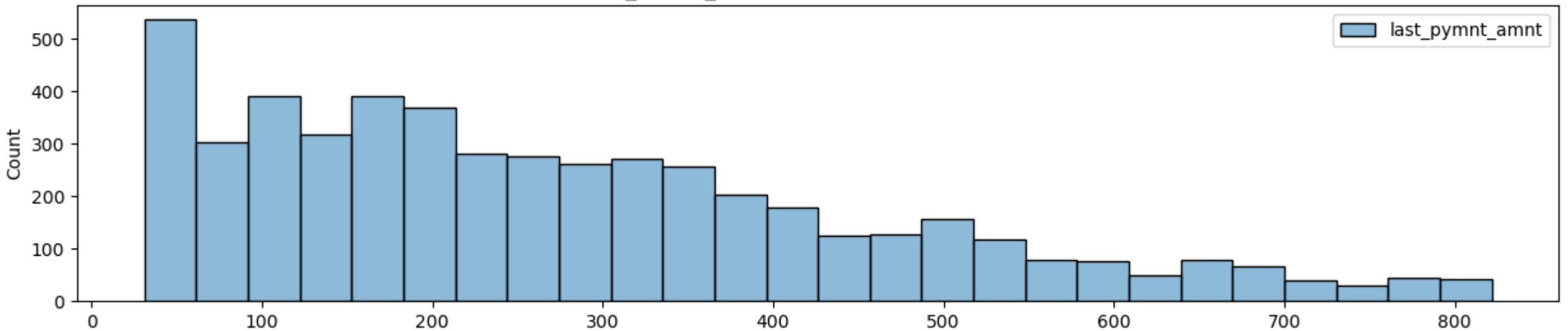

revol_util of Charged Off applicants
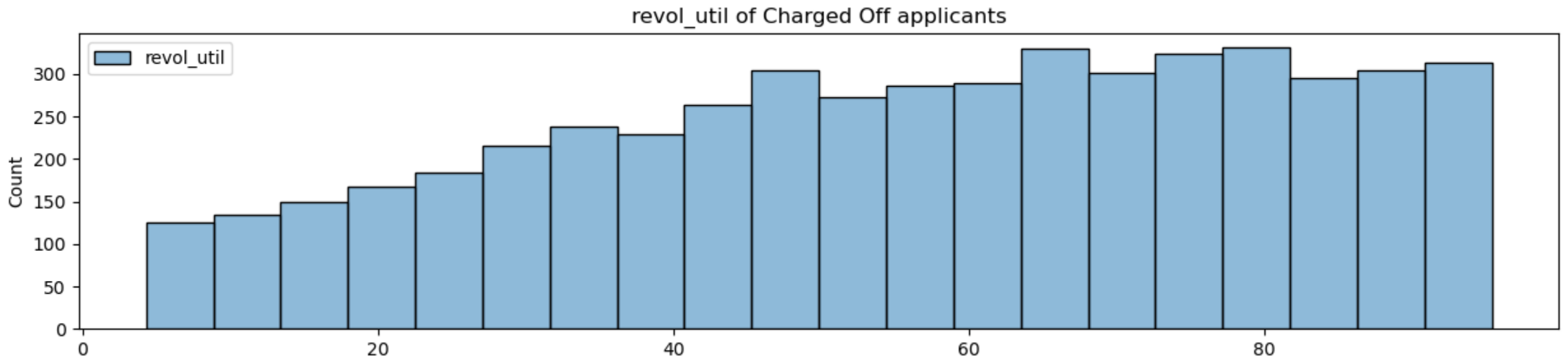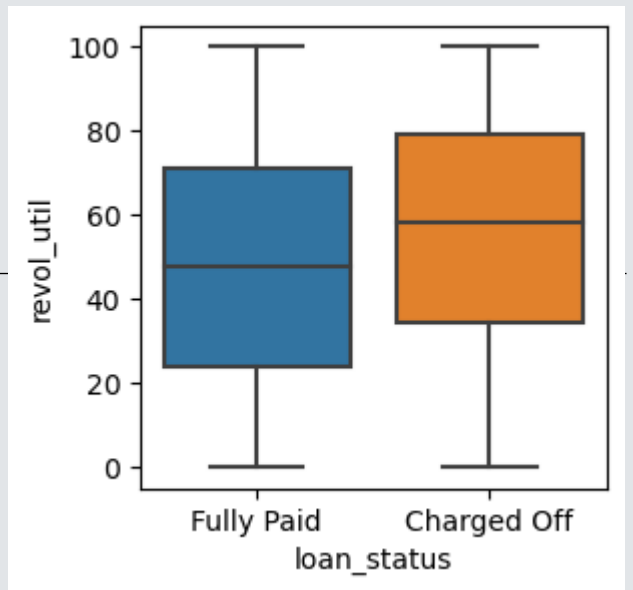
Payments received to date (total_pymnt) is a strong indicator, lower amount indicates more chances of defaulting

total_pymnt of Charged Off applicants

Last payment amount (last_pymnt_amnt) is a strong indicator, lower amount indicates more chances of defaulting
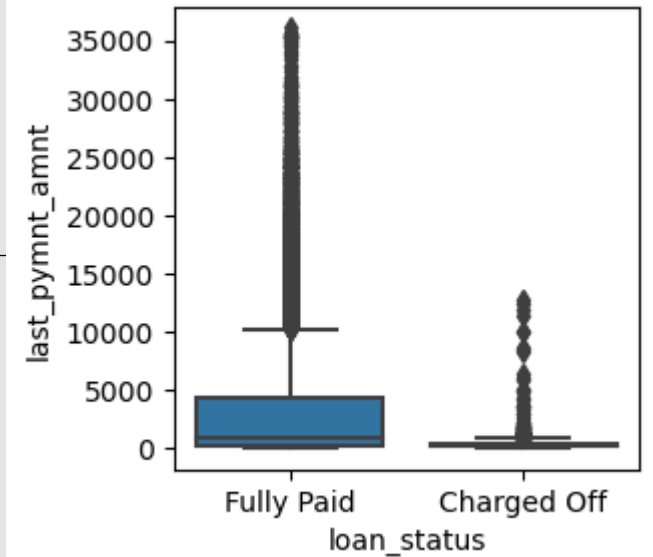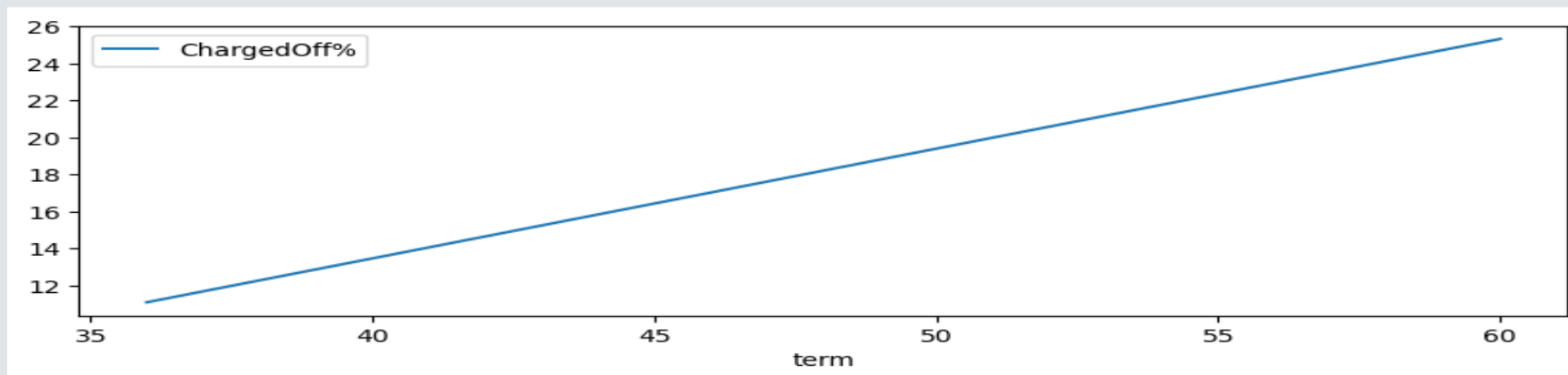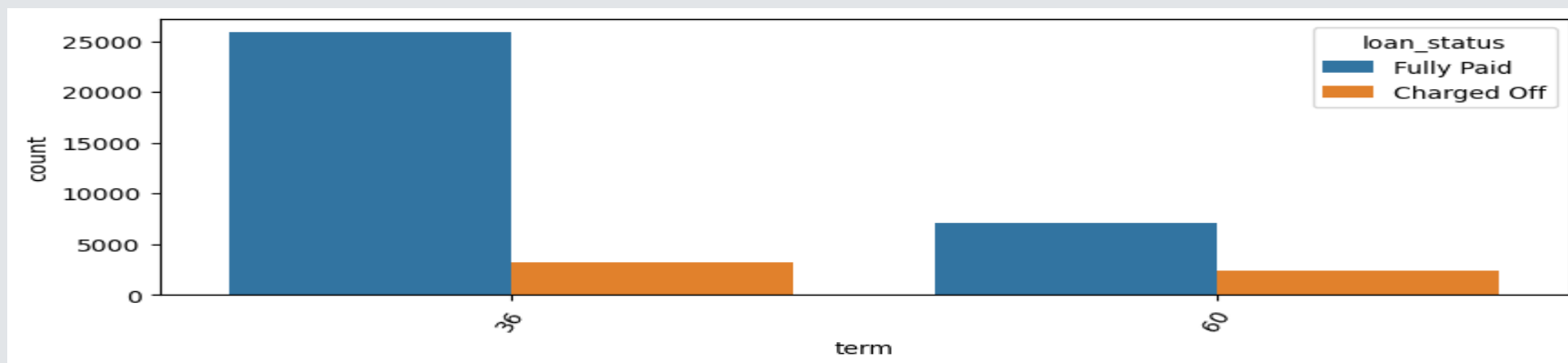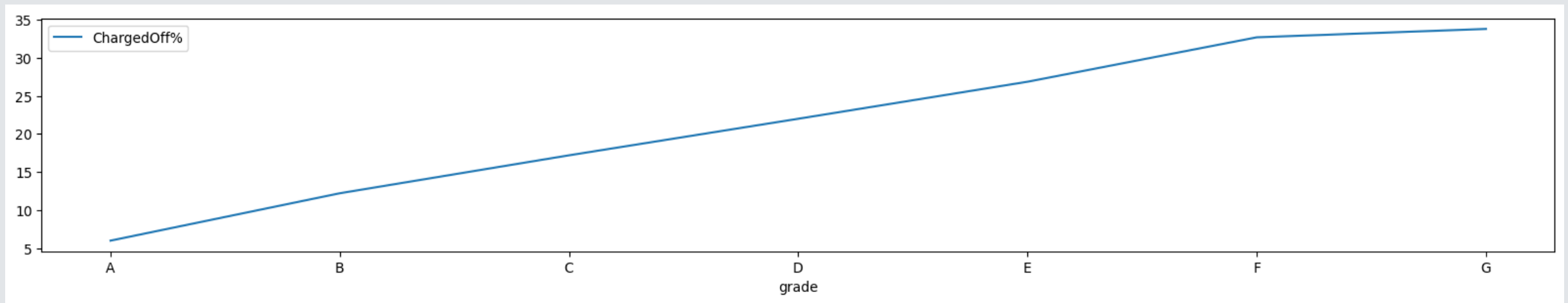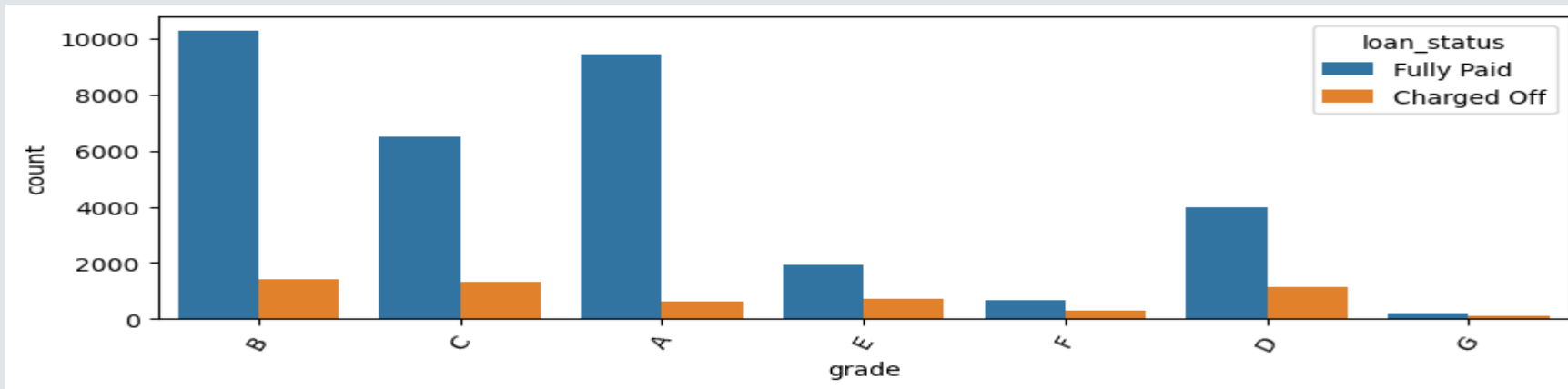

last_pymnt_amnt of Charged Off applicants

Categorical Data

# Term is a strong indicator, lower term indicates more chances of defaulting

# LC assigned loan grade (grade) is a strong indicator, defaulting percentage increase going from A to G

LC assigned loan subgrade (sub_grade) is a strong indicator, defaulting percentage increase going from 1 to 5 for higher grades like A,B,C,D but for lower grades like E and F sub_grade has low impact
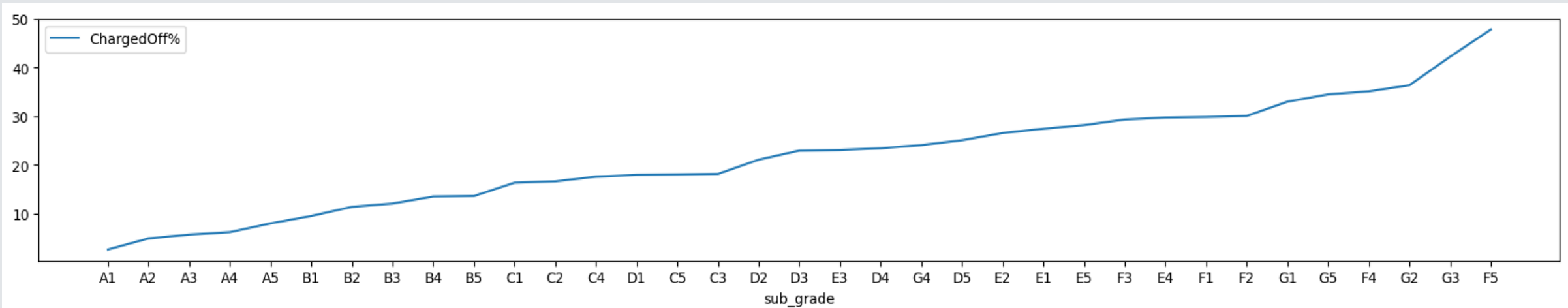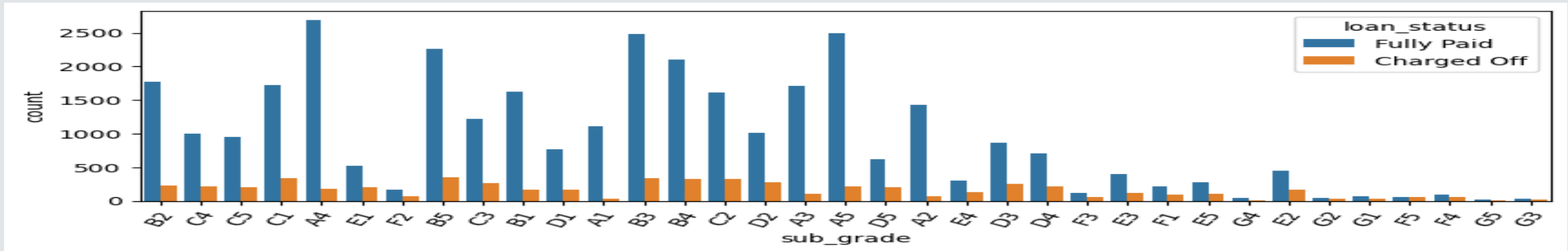
Home Ownership status (home_ownership) is a strong indicator, defaulting chance is low for Mortgage, higher for Own and Rent, and highest for Other
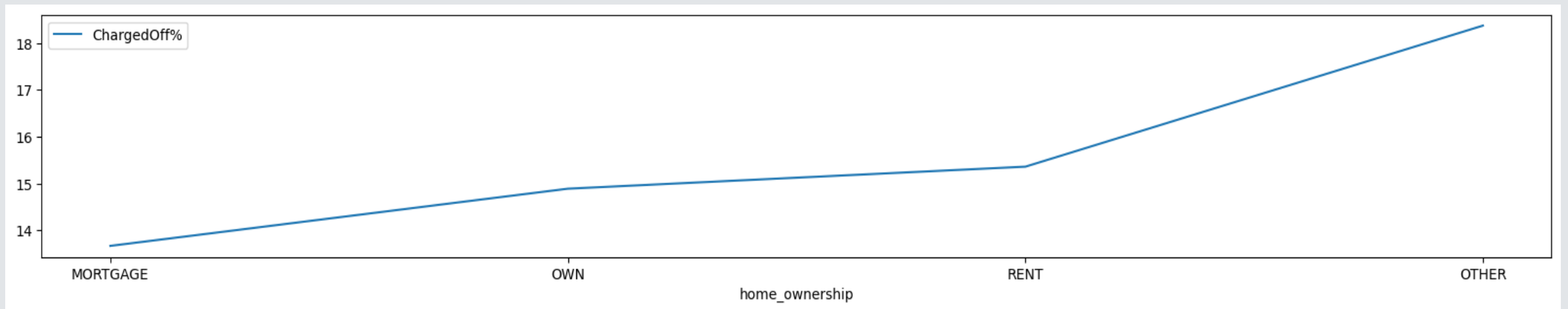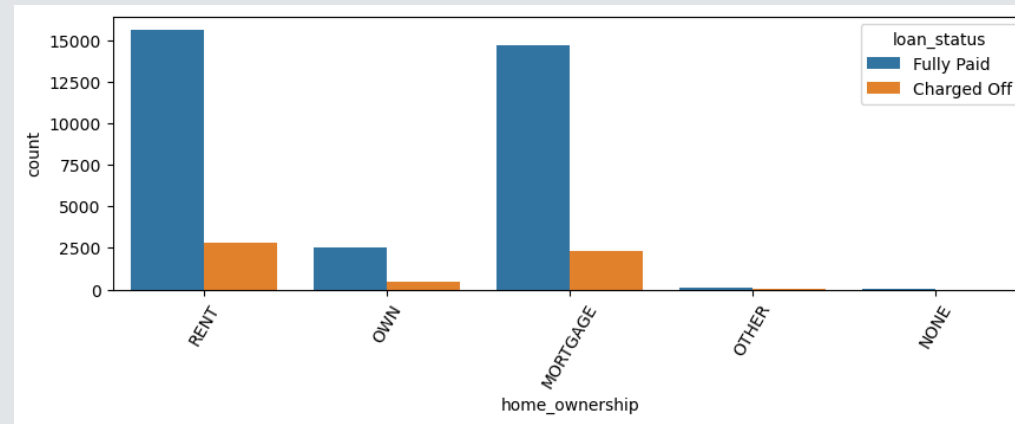
Verification Status (verification_status) is a strong indicator, defaulting chance is low for Not Verified, higher for Source Verified, and highest for Verified
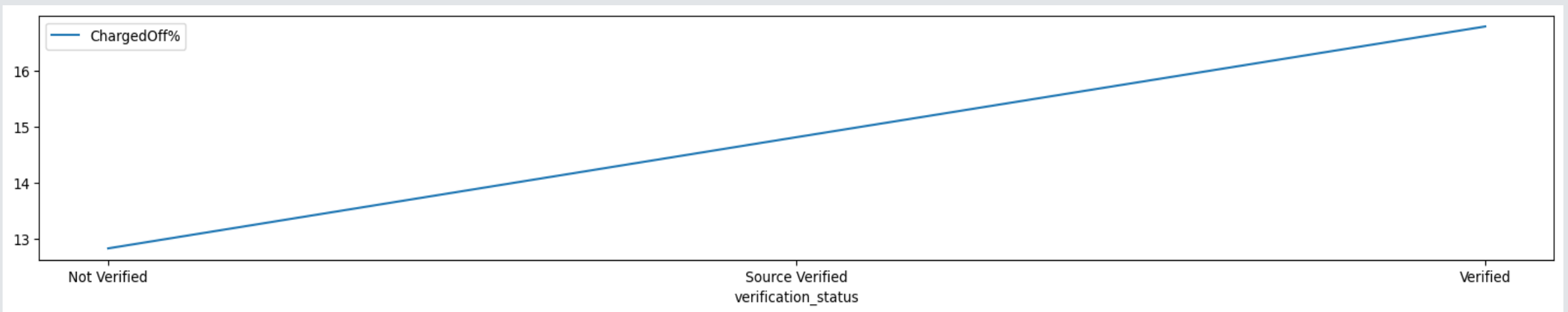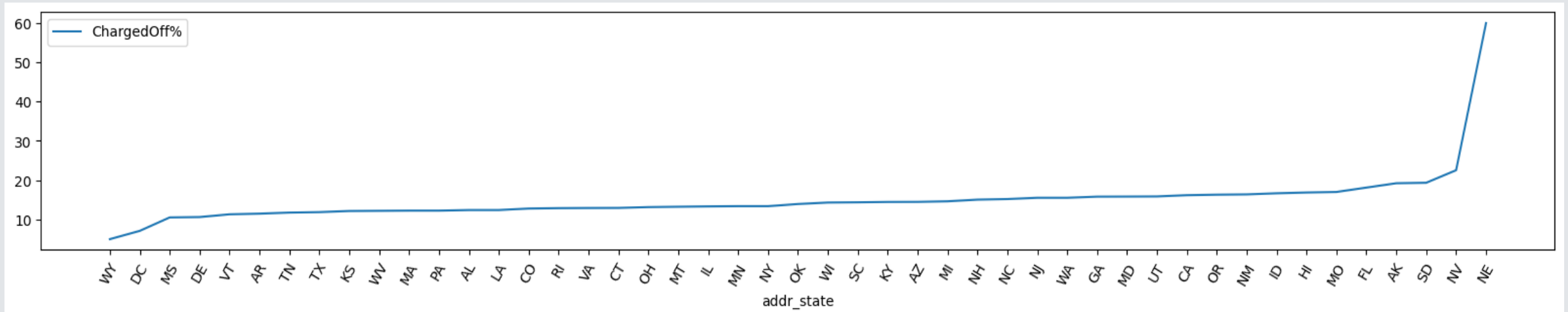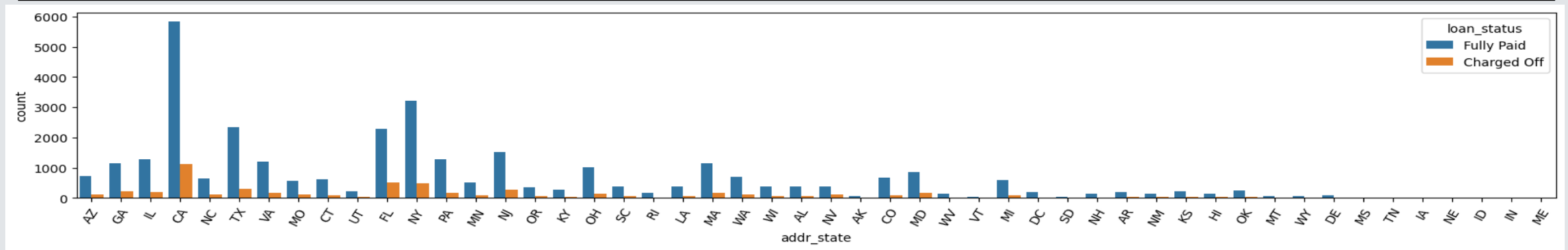
Purpose is a strong indicator.
- major_purchase, wedding, car, credit_card, home_improvement has low chance of defaulting
- vacation, debt_consolidation, medical, moving, house, other, educational has higher chance of defaulting
- renewable_energy, small_business has the highest chance of defaulting
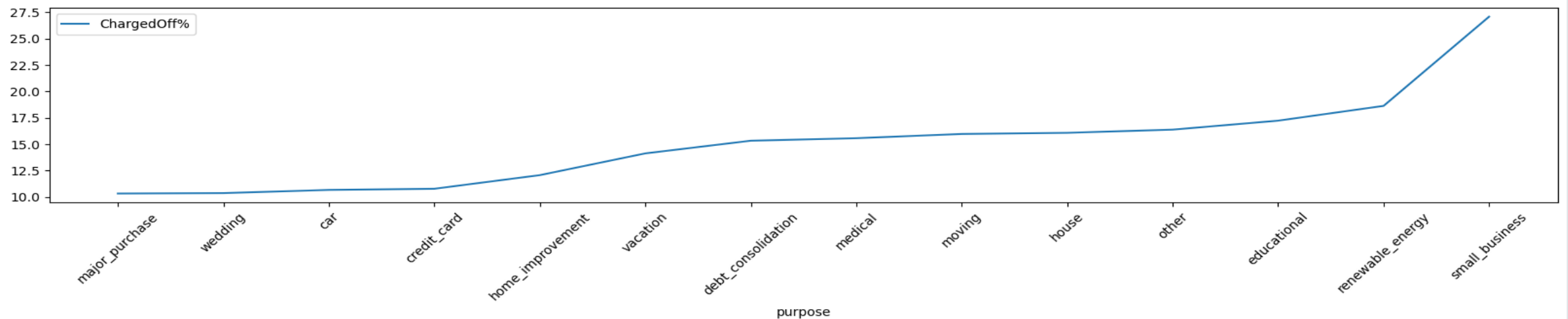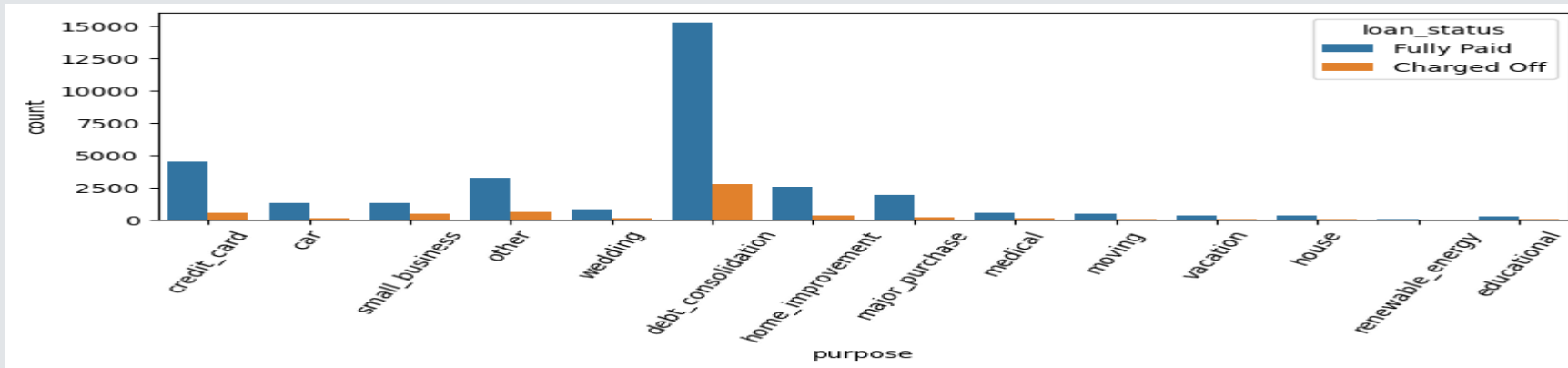
The state borrower lives in (addr_state) is a weak indicator, all states has 10-20% chance of defaulting accept, WY has very low chance of defaulting, while NE has very high chance of defaulting (though we have only 5 records from NE)

Thank You