

LENDING CLUB CASE STUDY - EDA

SUBMISSION

Team Members:

- Aniket Gupta
- Darshjyot Singh Negi

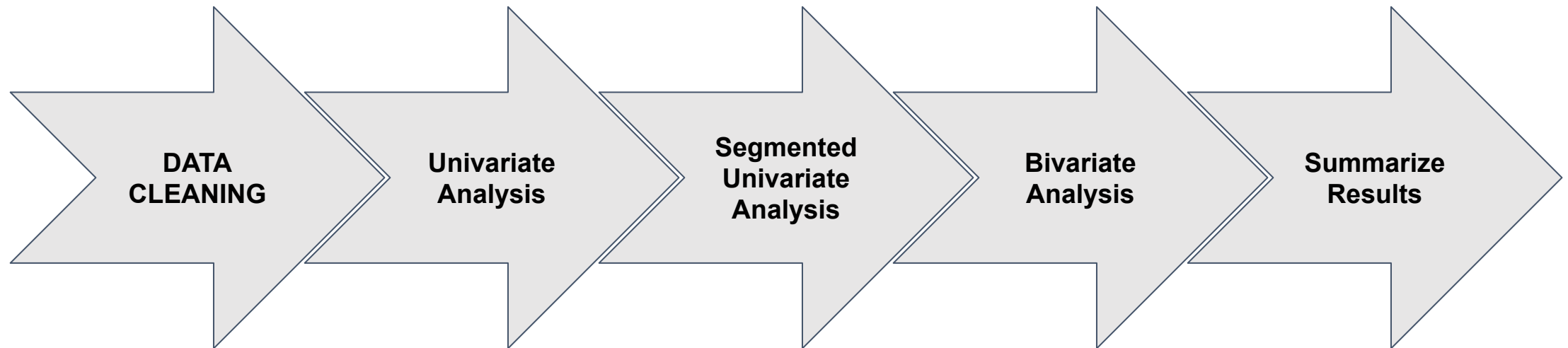
Problem Statement

When a person applies for a loan, there are two types of decisions that could be taken by the company:

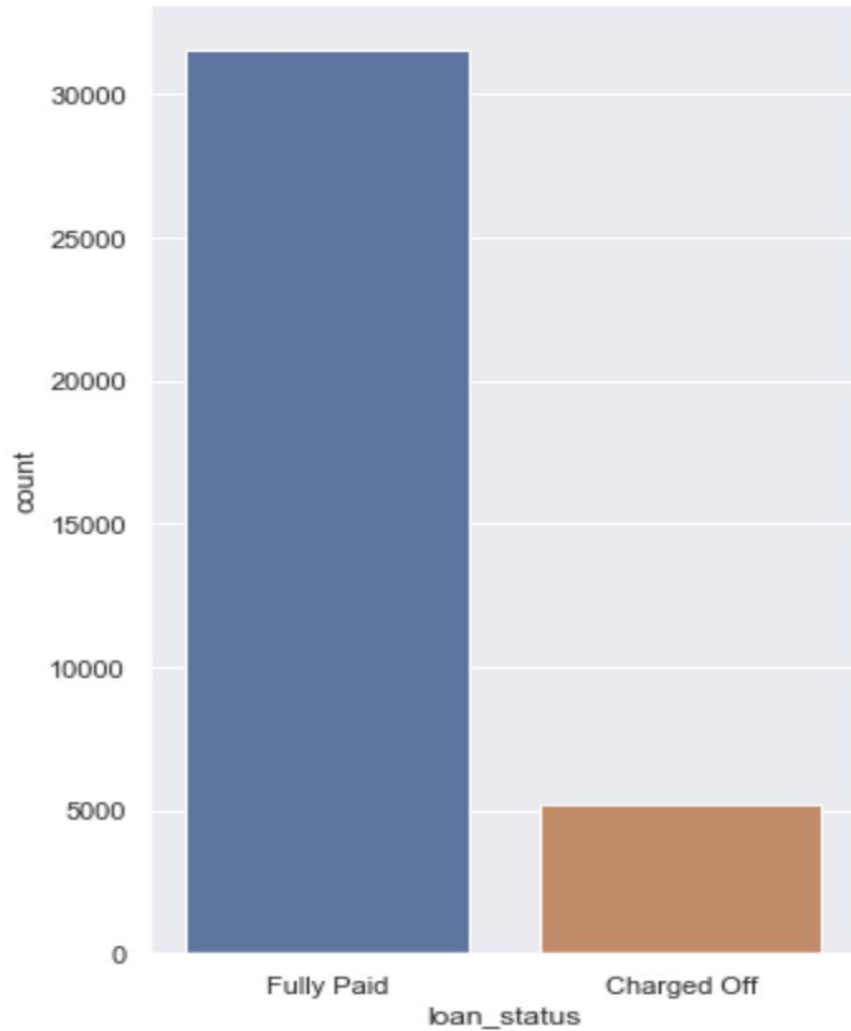
1. **Loan accepted:** If the company approves the loan, there are 3 possible scenarios described below:
 - **Fully paid:** Applicant has fully paid the loan (the principal and the interest rate)
 - **Current:** Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
 - **Charged-off:** Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan
2. **Loan rejected:** The company had rejected the loan.

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 utilise this knowledge for its portfolio and risk assessment.

Problem solving methodology

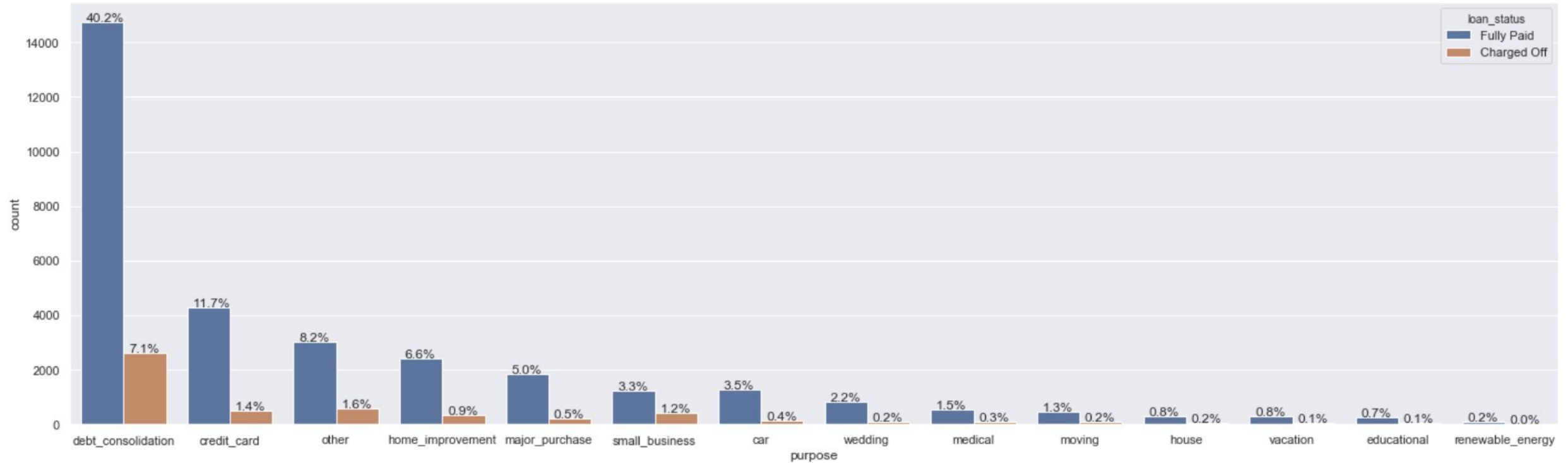


Analysis - Loan Status



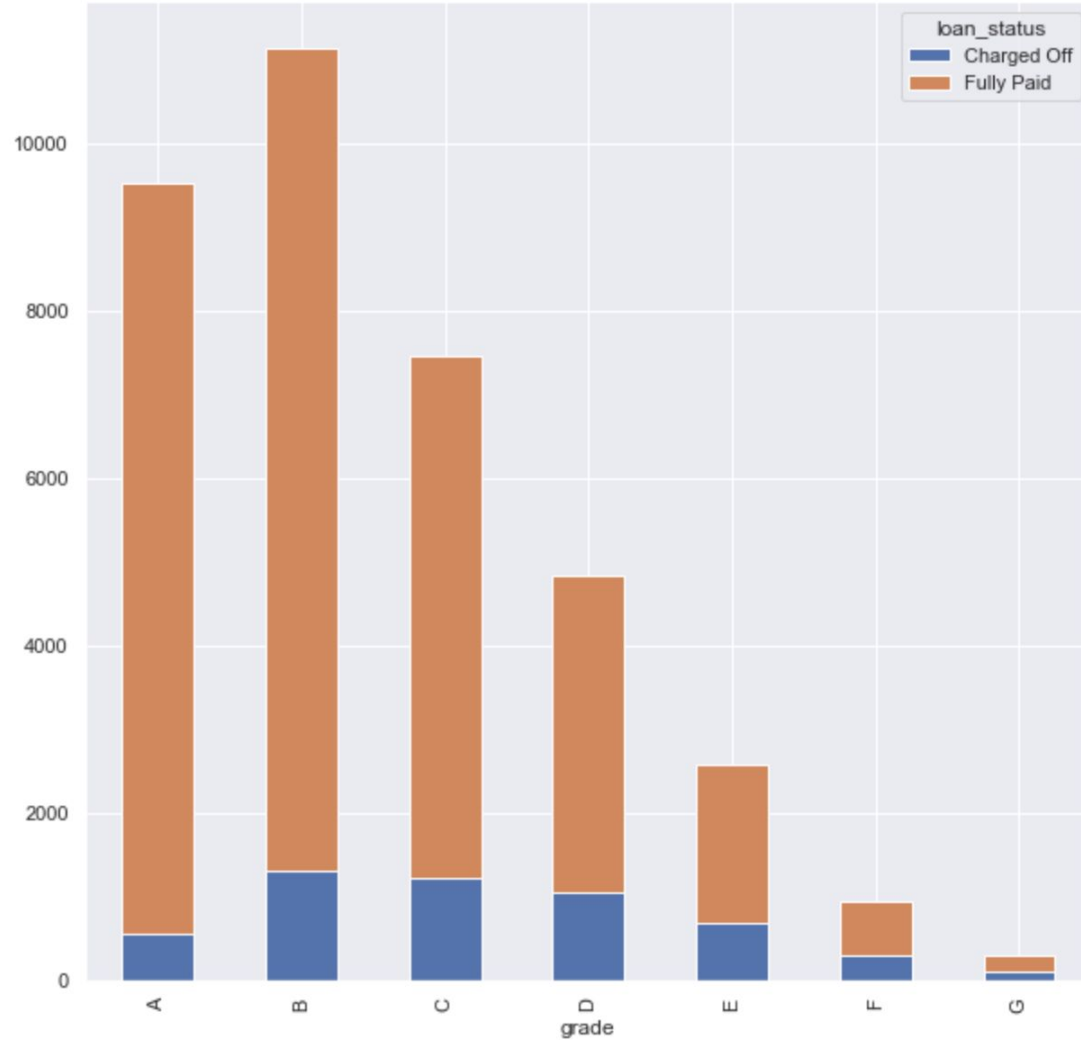
~14% out of total loans are defaulted.

Analysis - Understanding Data



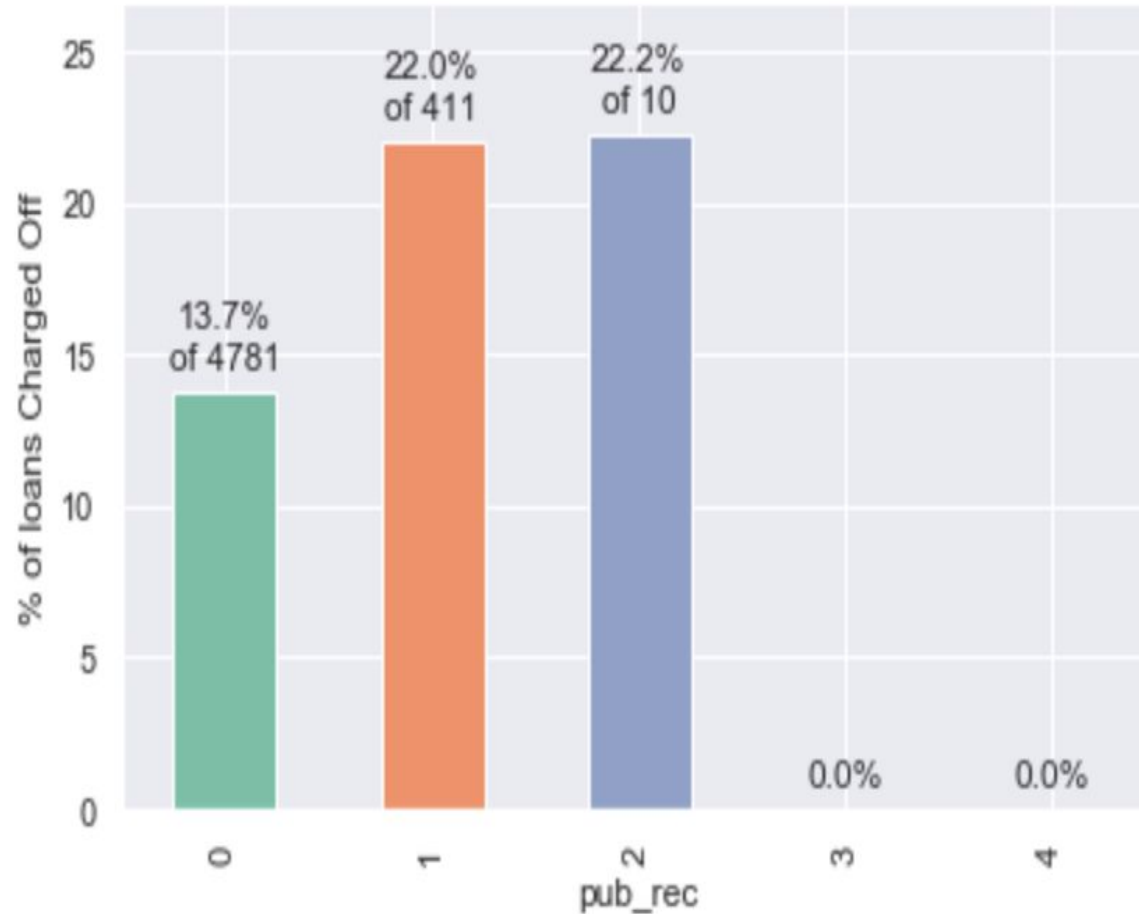
Maximum number of loans are applied for debt consolidation and credit card.

Analysis - Understanding Data



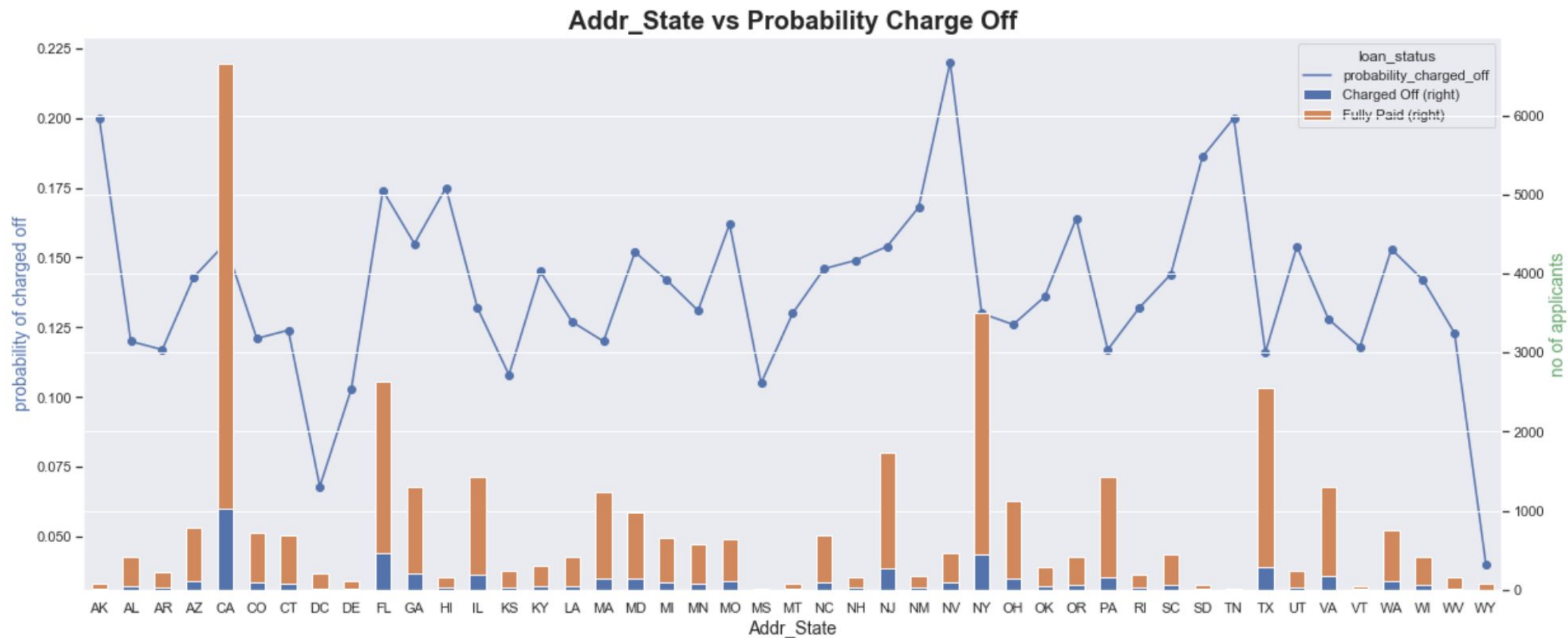
Lower grades(E,F,G) have higher percentage of defaults

Analysis



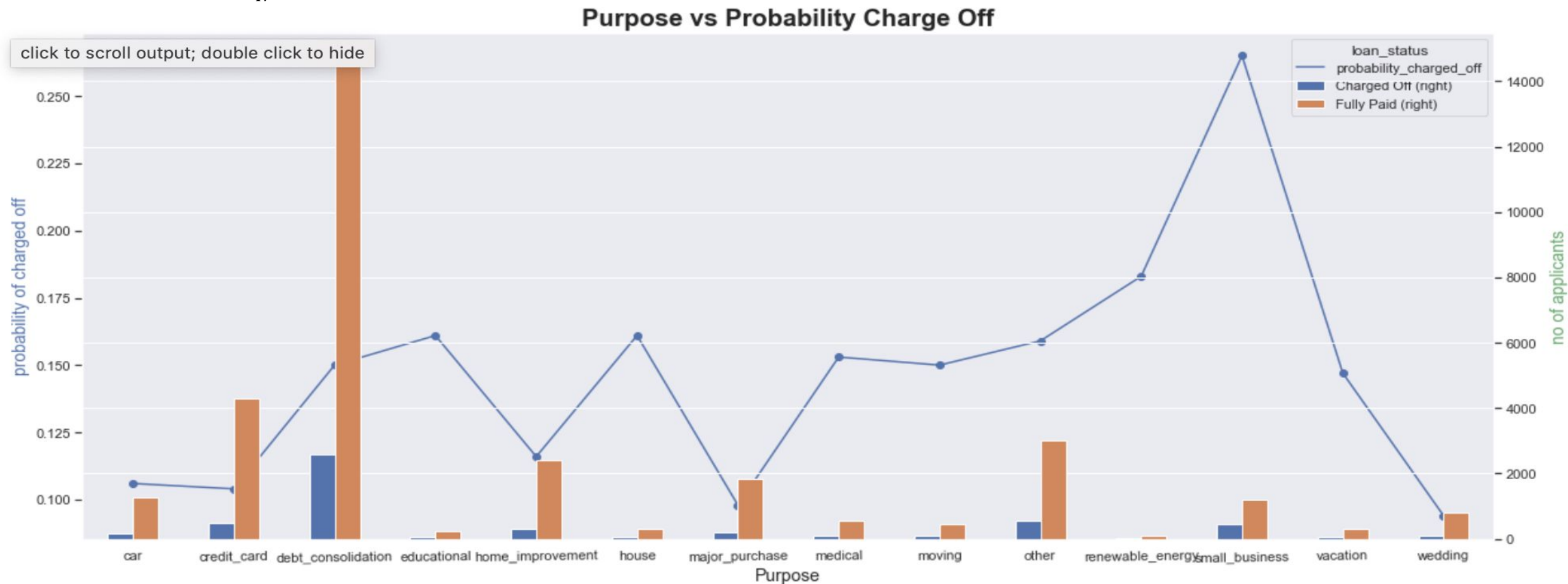
- 94% have no Public derogatory records. 5% have 1 derogatory record.
- Having even 1 derogatory record increases the chances of Charge Off significantly.

Analysis



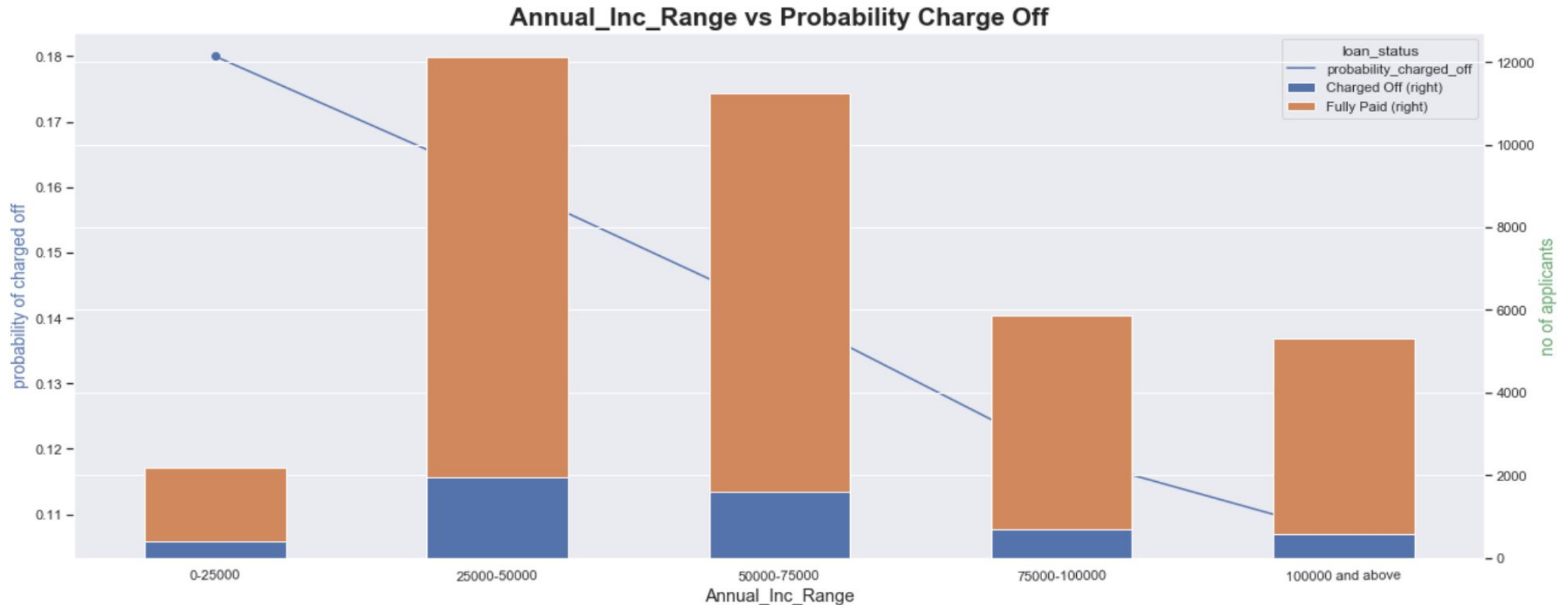
Observation: There are multiple states with high probability of charged off, highest is 'NV' at 22.5%

Analysis



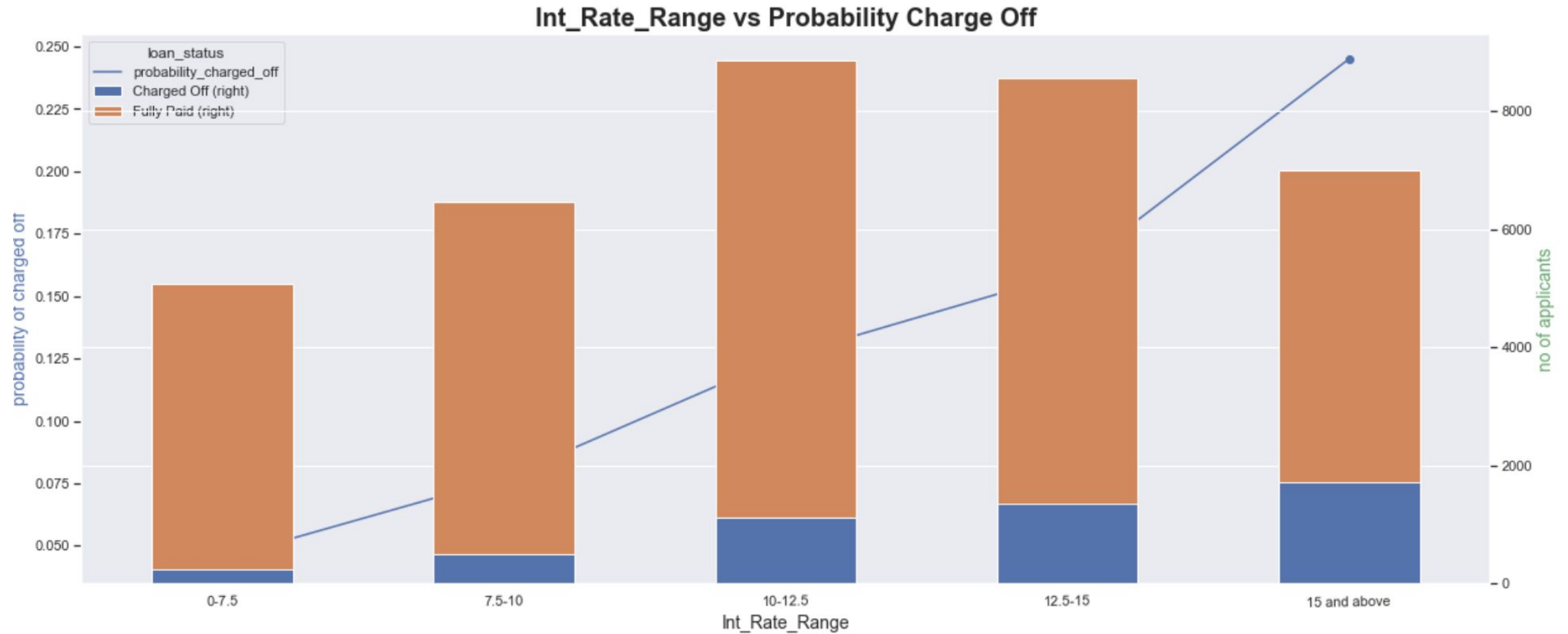
Applicants who has taken the Loan for 'small business' has the highest probability of charge off of 26.5%.

Analysis



As the annual income decreases the probability that applicants will default is increasing with highest of 18% at (0 to 25000) salary bracket.

Analysis



As the int_rate is increasing the probability that person will default is increasing with highest of 24.5% at 15% & above range.

Conclusions

Driving factors (or driver variables) behind loan default are:

- **Annual Income**
- **Purpose of loan**
- **Grade**
- **Number of derogatory public records**