Lending Club Case Study:
Assignment Submission

Prepared and submitted by

Shraddha Bhoir

Krishna Murthi B

Course : Executive PG Programme in Machine Learning & AI Feb 2022
Date : 06-Apr-2022

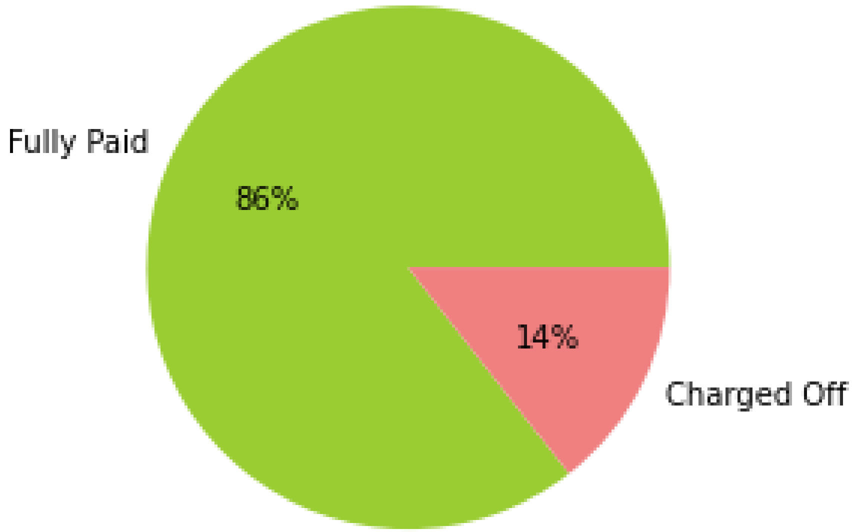
Problem Statement

- The company is a consumer finance company which specialises in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile.
- 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.

Approach

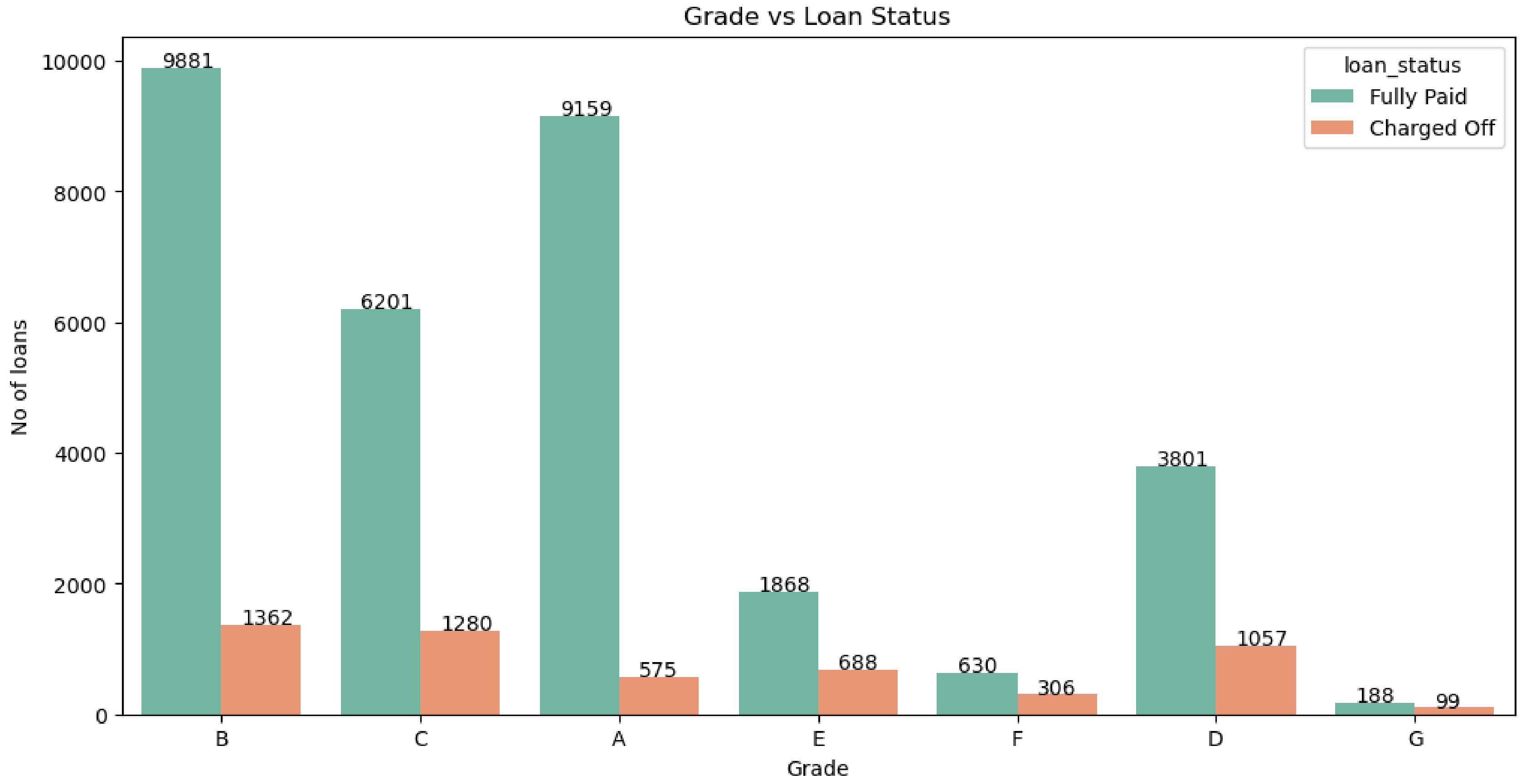
- Step 1:
  - Data Understanding,
  - Cleaning,
  - Manipulation,
  - Treating missing values,
  - Missing value imputation,
  - Handling outliers and
  - Derived metrics
- Step 2: Univariate Analysis
- Step 3: Segmented Univariate Analysis
- Step 4: Bivariate/Multivariate Analysis
- Step 5: Results

Analysis – Fully paid vs Charged Off



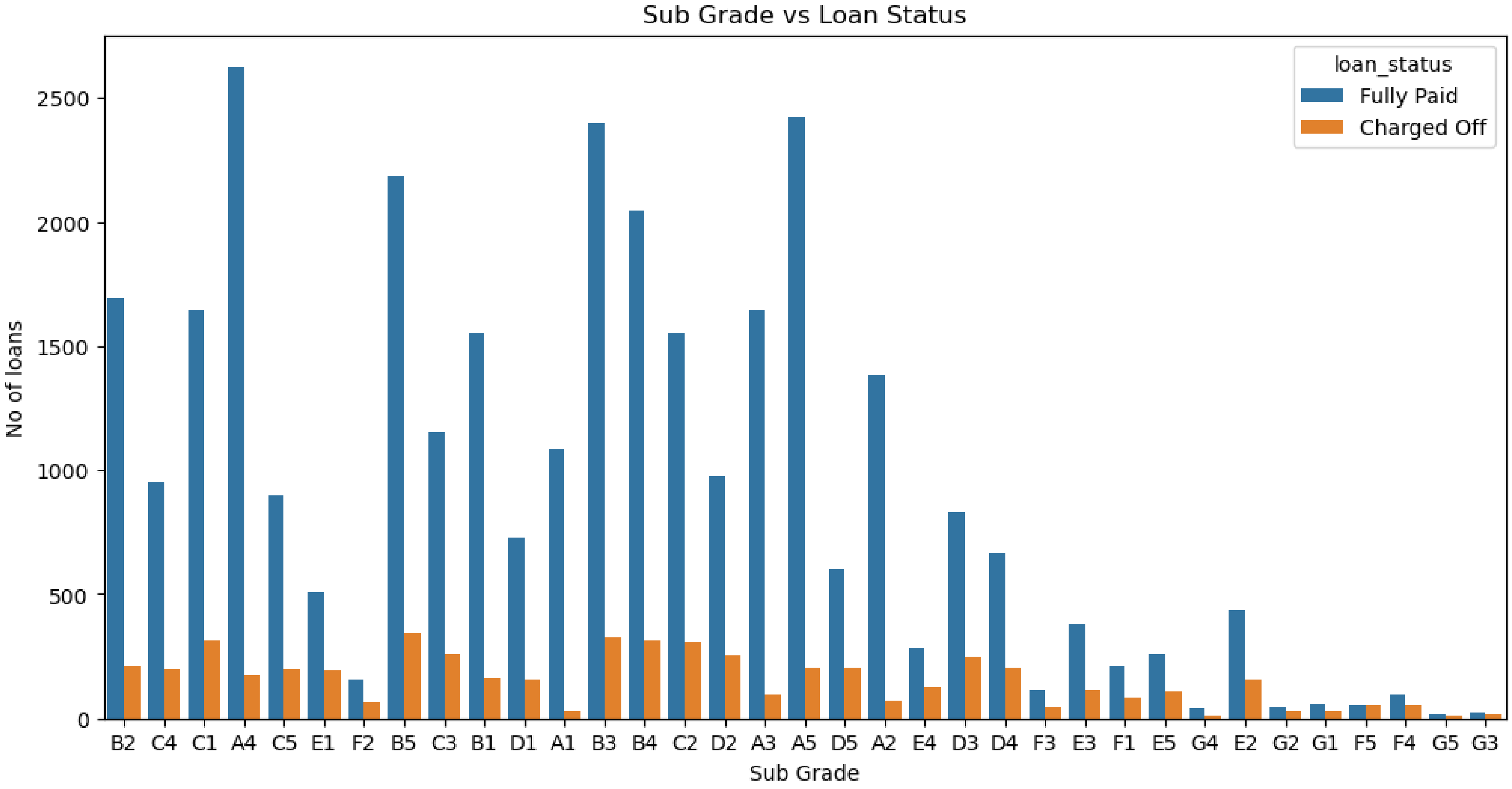
- 15 % of the total loans are charged off

Analysis – Grade



- Based on the counts, Grade B, C and D are top three in Charged Off

Analysis – Sub Grade

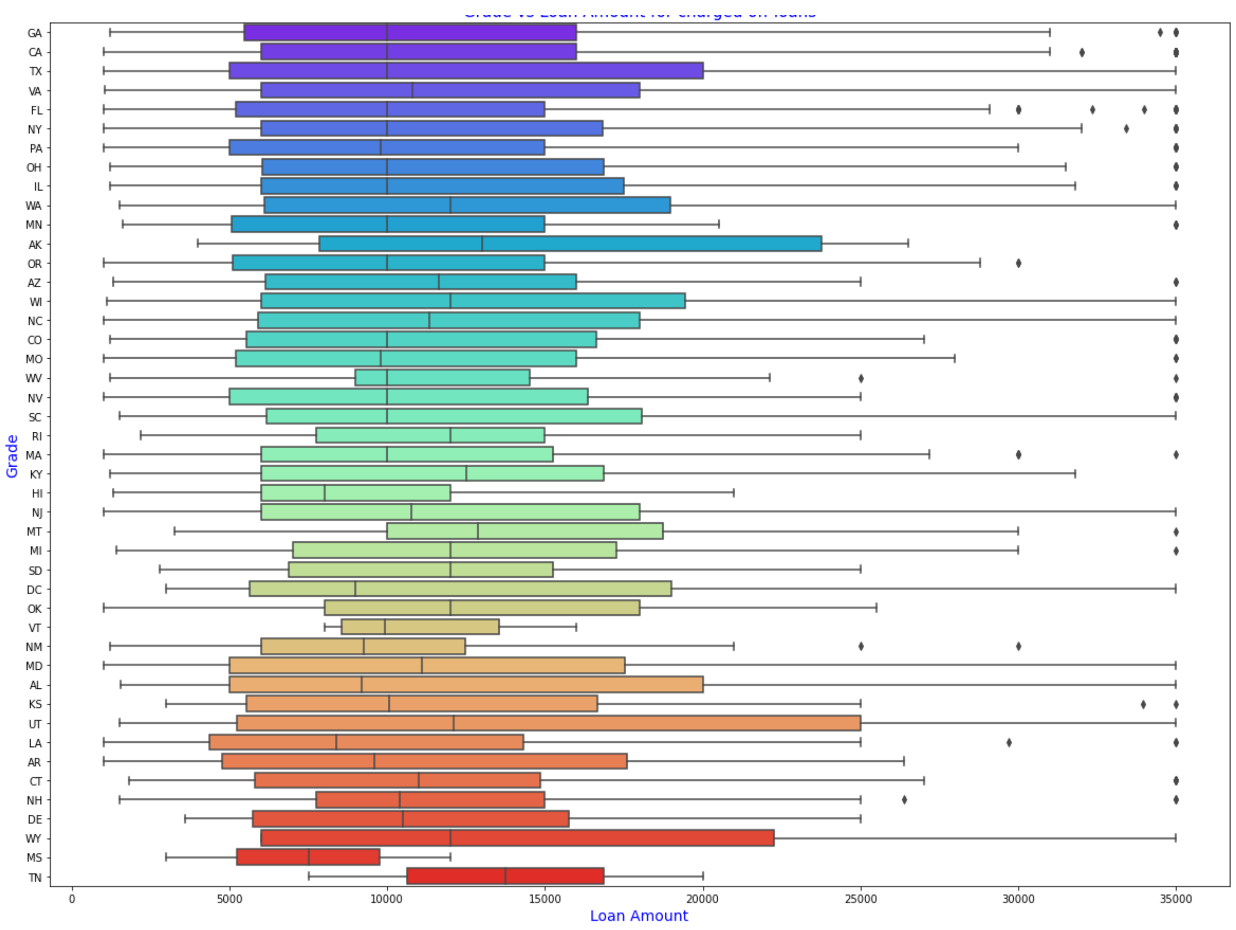


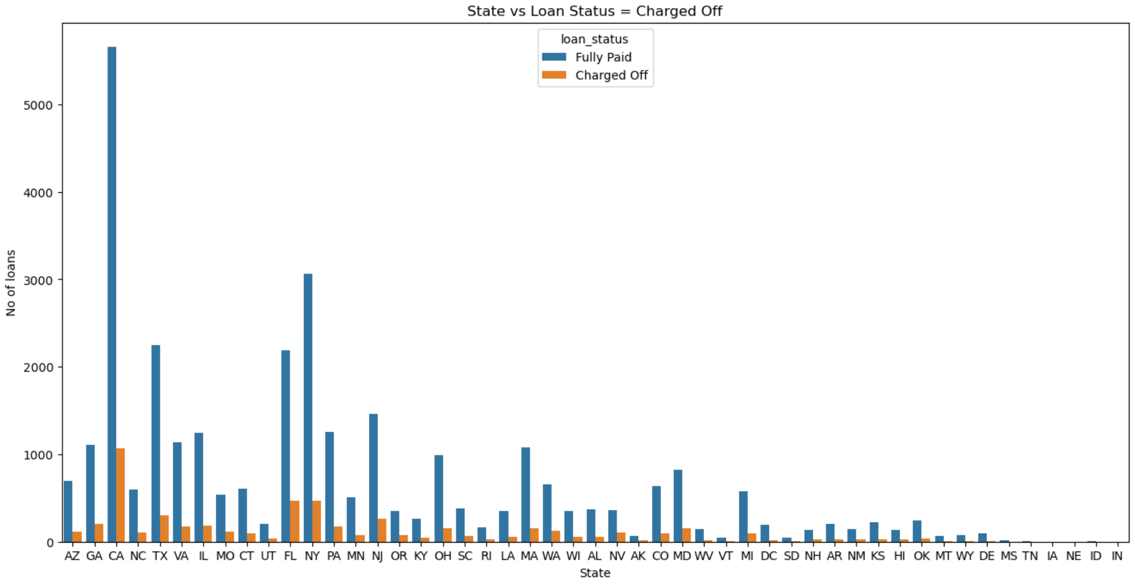
- Based on the counts, Grade B3,B4,B5, C1,C2, D3, D4 top sub grades in Charged Off

Analysis – Emp Length

- Employment length of 10 years got more loans and they are the maximum defaulters considering the count

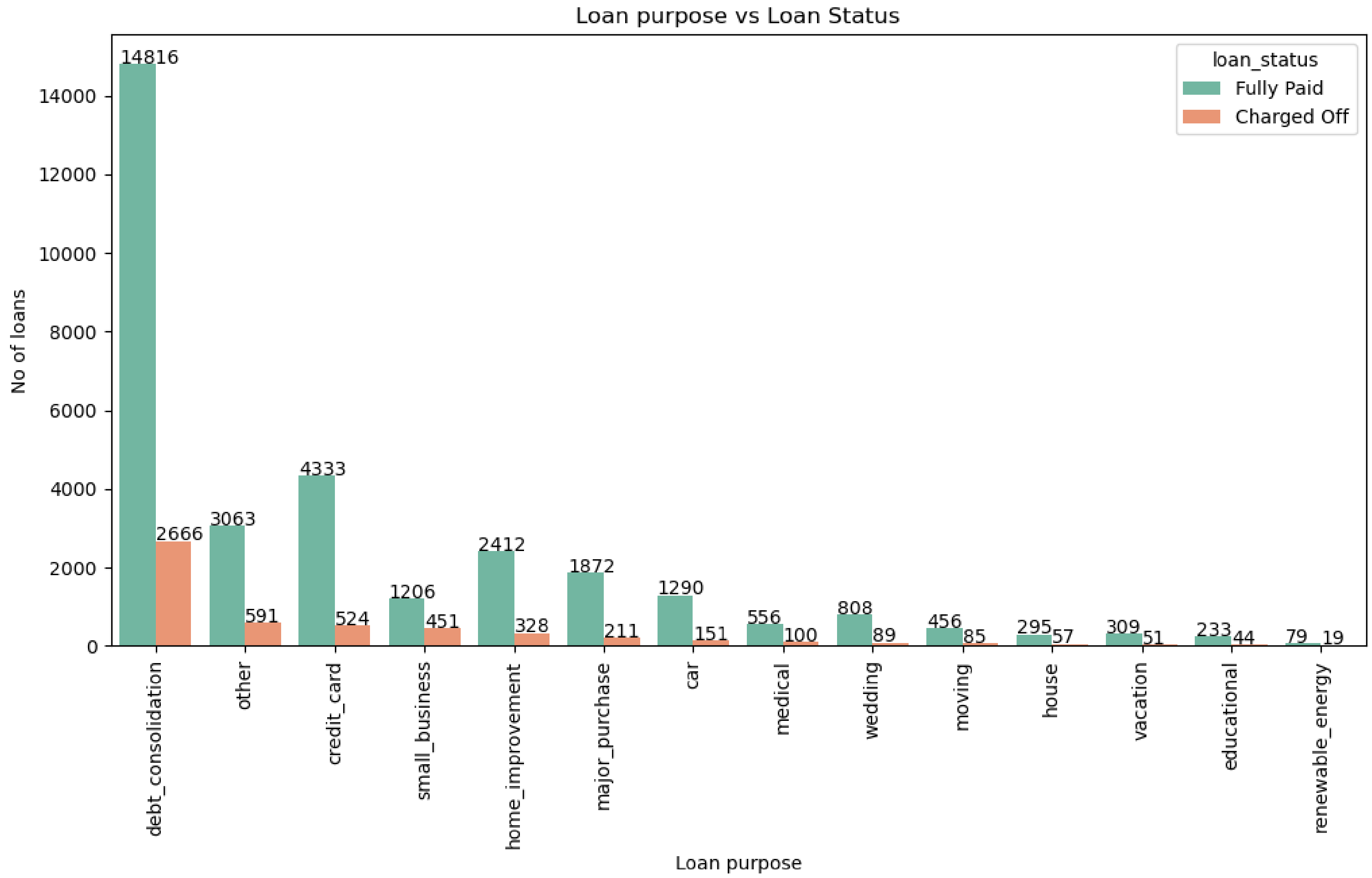
Analysis – State





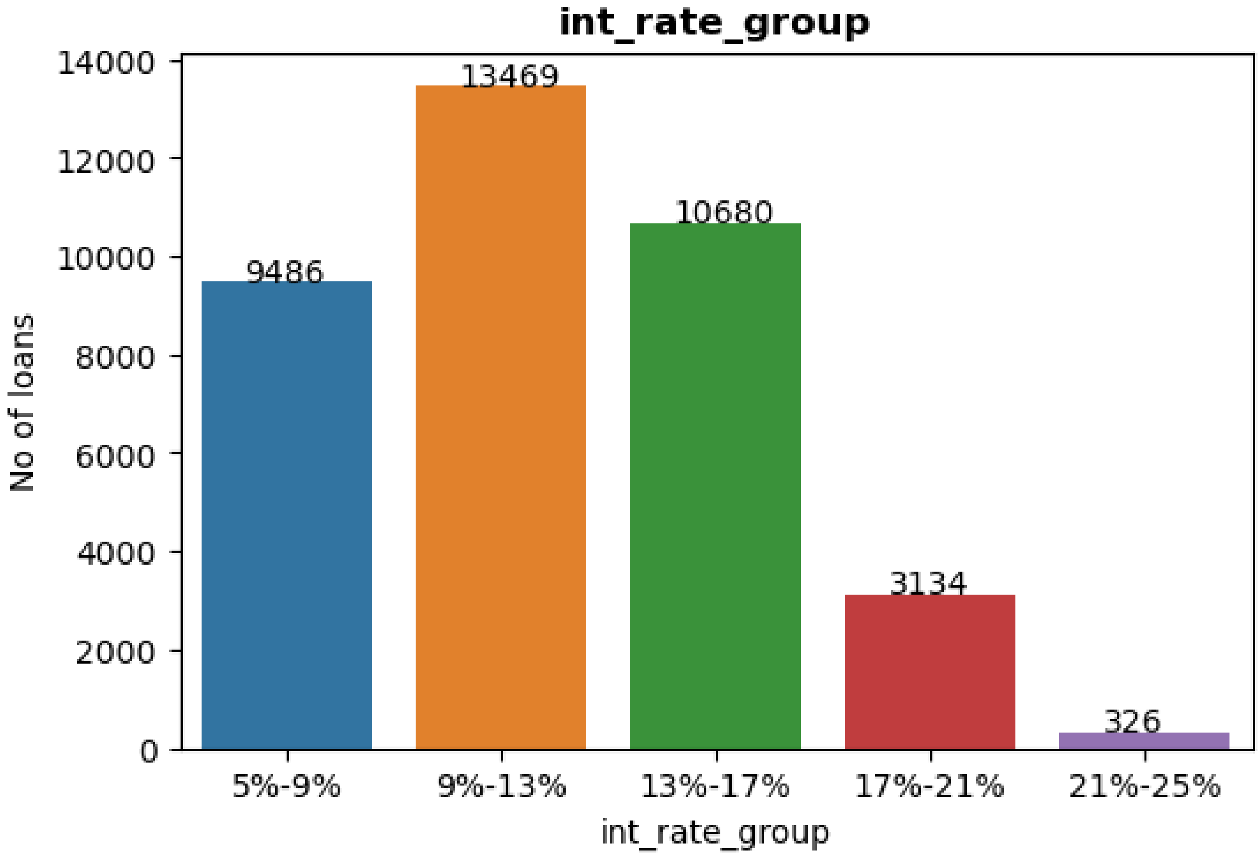
- Many states median is at 10k for charged off loan
- State AK, MT, KY, UT, TN are marginally above the 10k level and at 13k
- States AK, UT the Q3 is at 23k

Analysis – Purpose



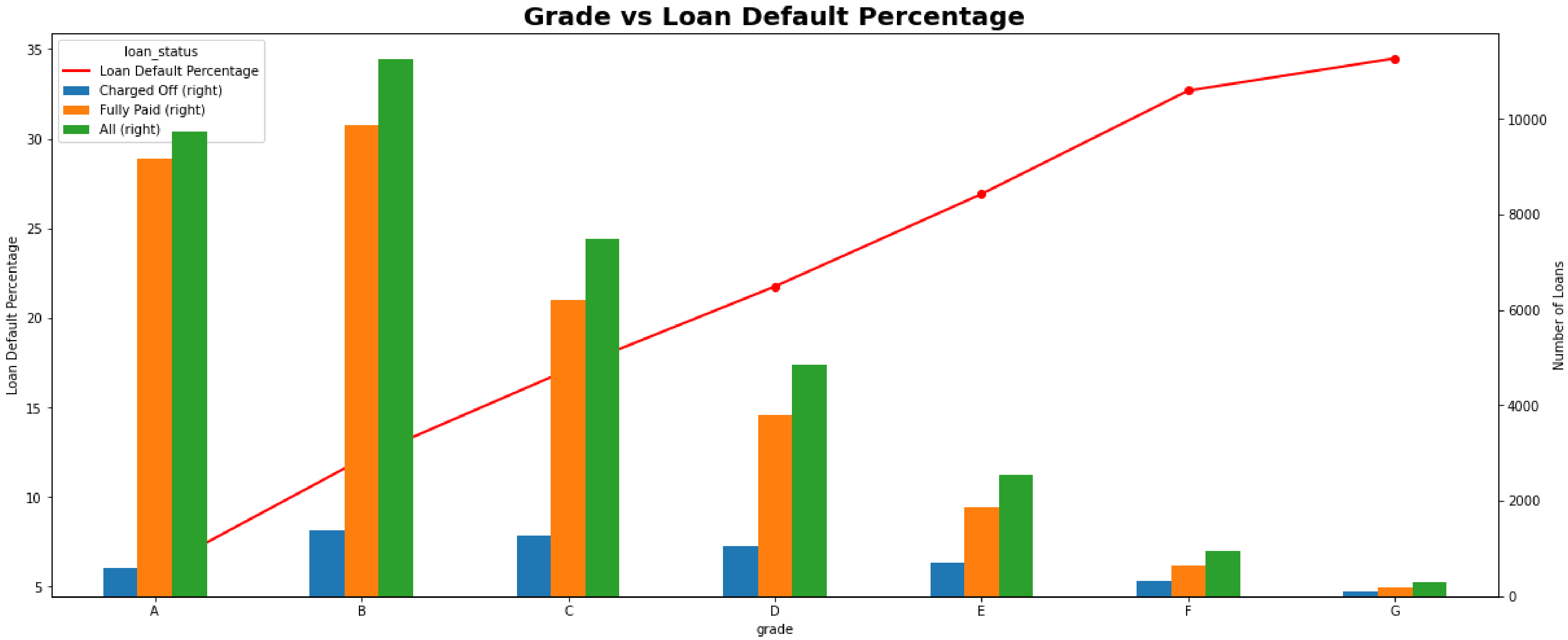
- Loans with purpose debt consolidation, other, credit crd and home improvement categories have failed to pay the loan compared with education / renewable energy

Analysis – Interest rate



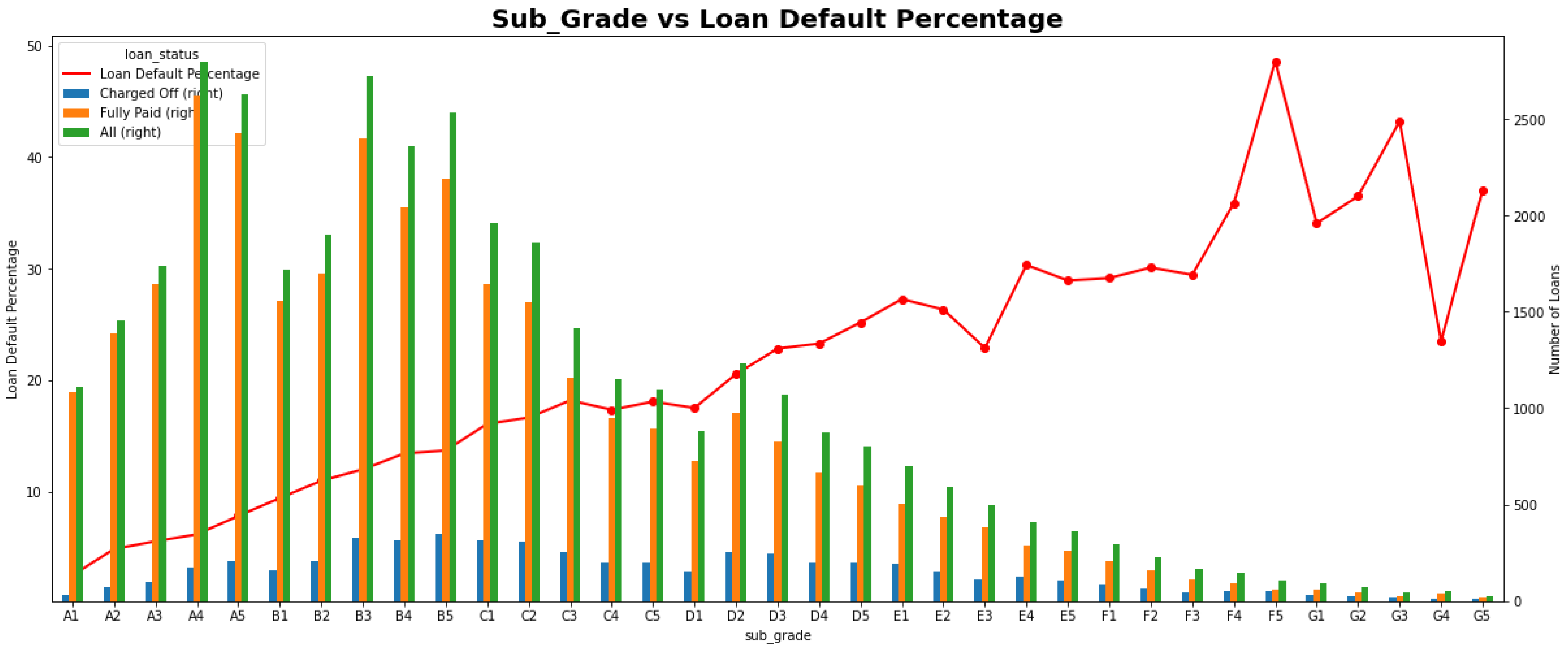
- Interest rate range 9 to 13 is the range where maximum loans have been issued

Analysis – Grade vs Loan default percent



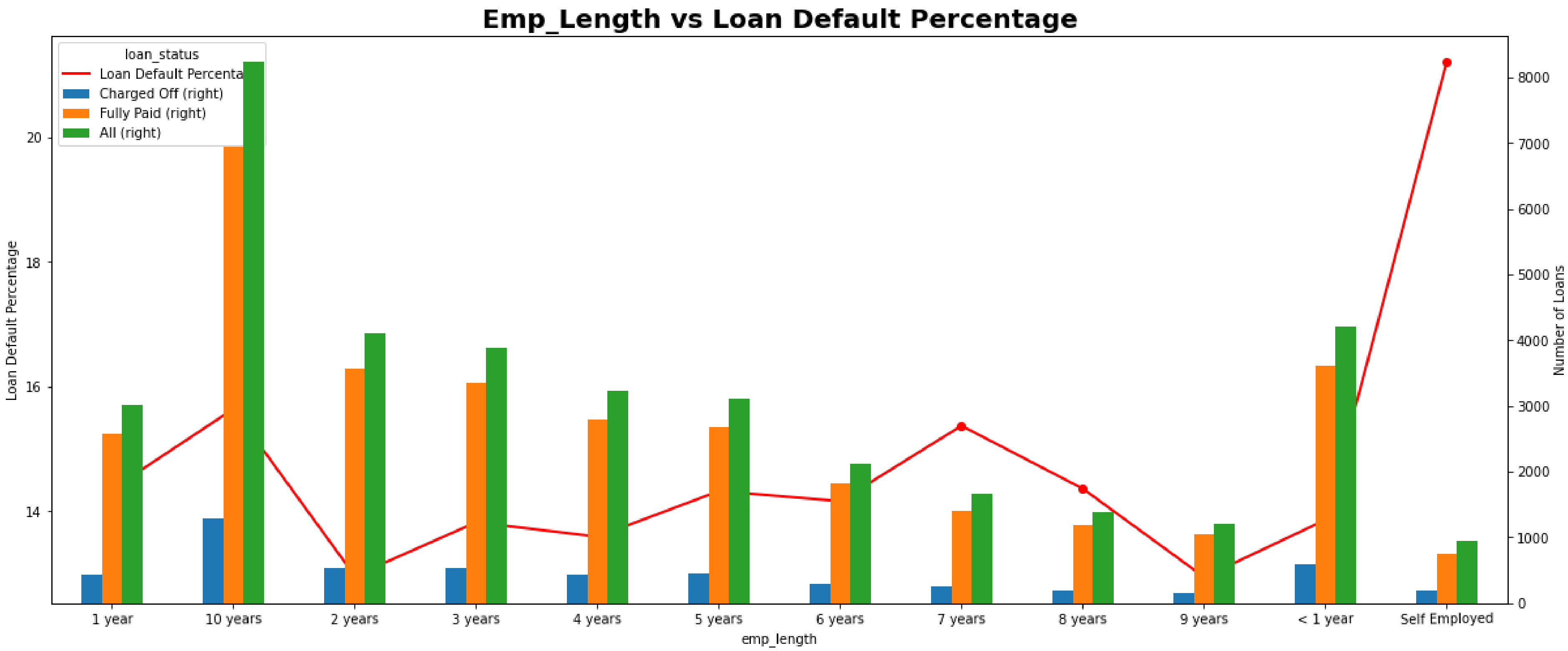
- Tendency to default the loan is increasing from Grade A to Grade G

Analysis – Sub Grade vs Loan default percent



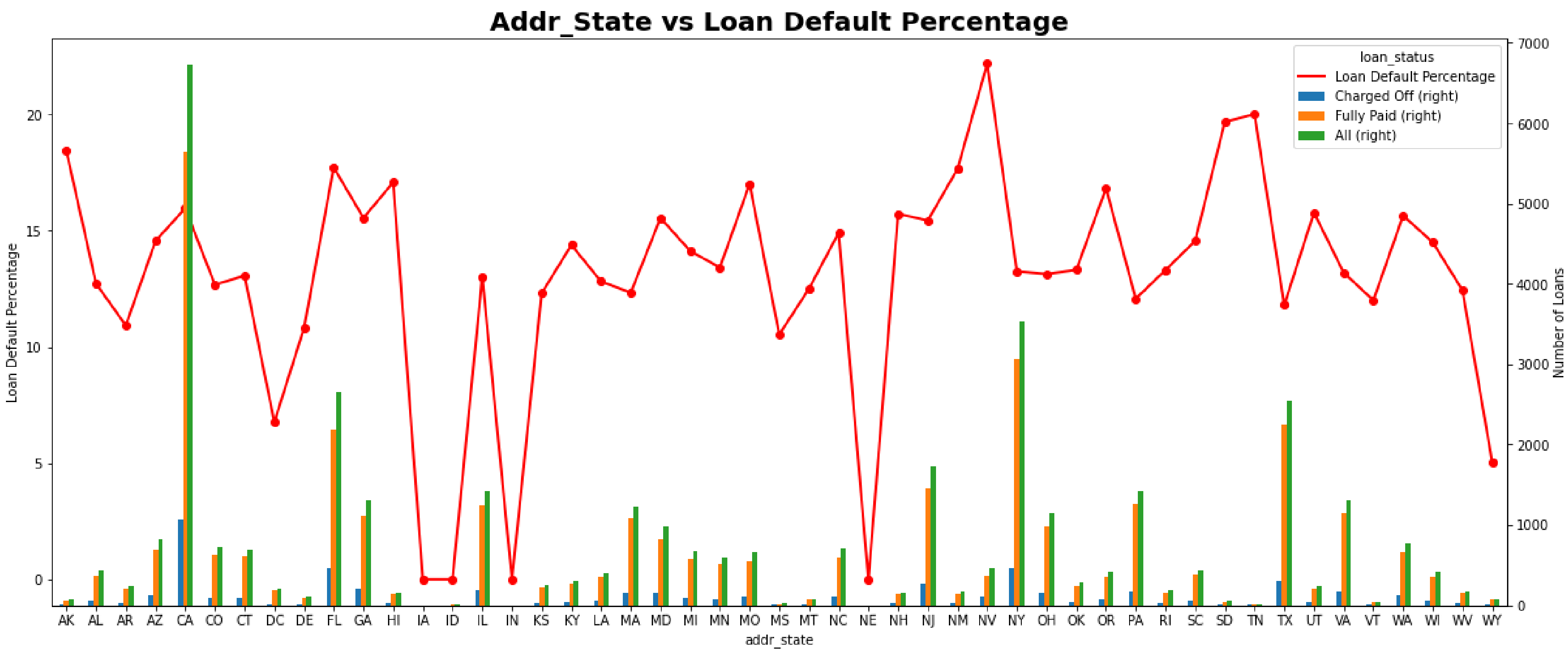
- Borrowers from sub grade F5, G3 and G5 have maximum tendency to default.

Analysis – Emp length vs Loan default percent



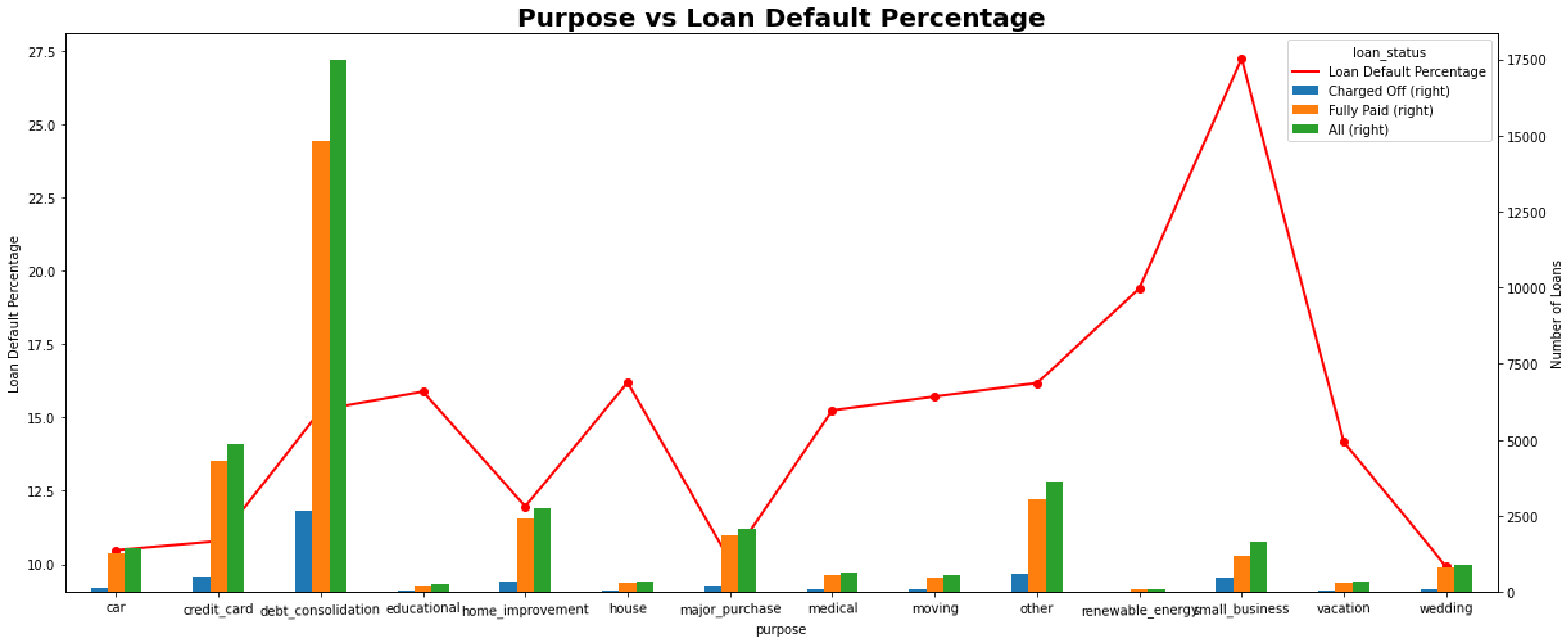
- Self employed, 10 years and 7 years categorys has maximum tendancy to default the loan.

Analysis – addr state vs Loan default percent



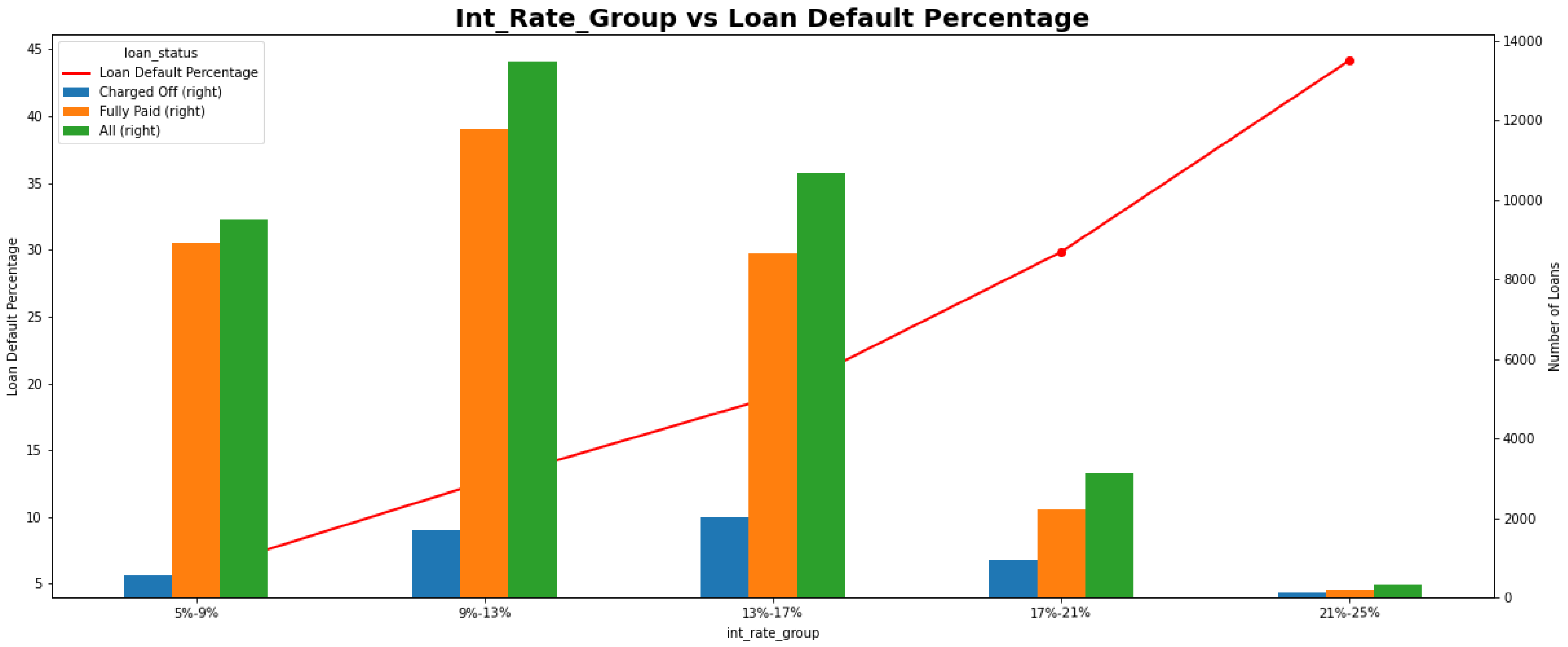
- Borrowers from states NV, TN, SD, AK, FL and HI have maximum tendency to deafult the loan.

Analysis – Purpose vs Loan default percent



- Small business category has the highest chances to default the loan

Analysis



- The tendency to default the loan is increasing with increase in the interest rate.

Conclusion

- Based on the extensive analysis, graphs and observations we conclude the following are the drive variables which may tend for a loan becoming ‘Charged Off’ (not in any particular order)
  - Grade
  - Sub Grade
  - Employee length
  - State
  - Month
  - Purpose
  - Loan Amount to Annual income ratio
  - Interest rate