

LENDING CLUB CASE STUDY

- Jimmy Ha

- Mrunmai Khedekar

Business Objective

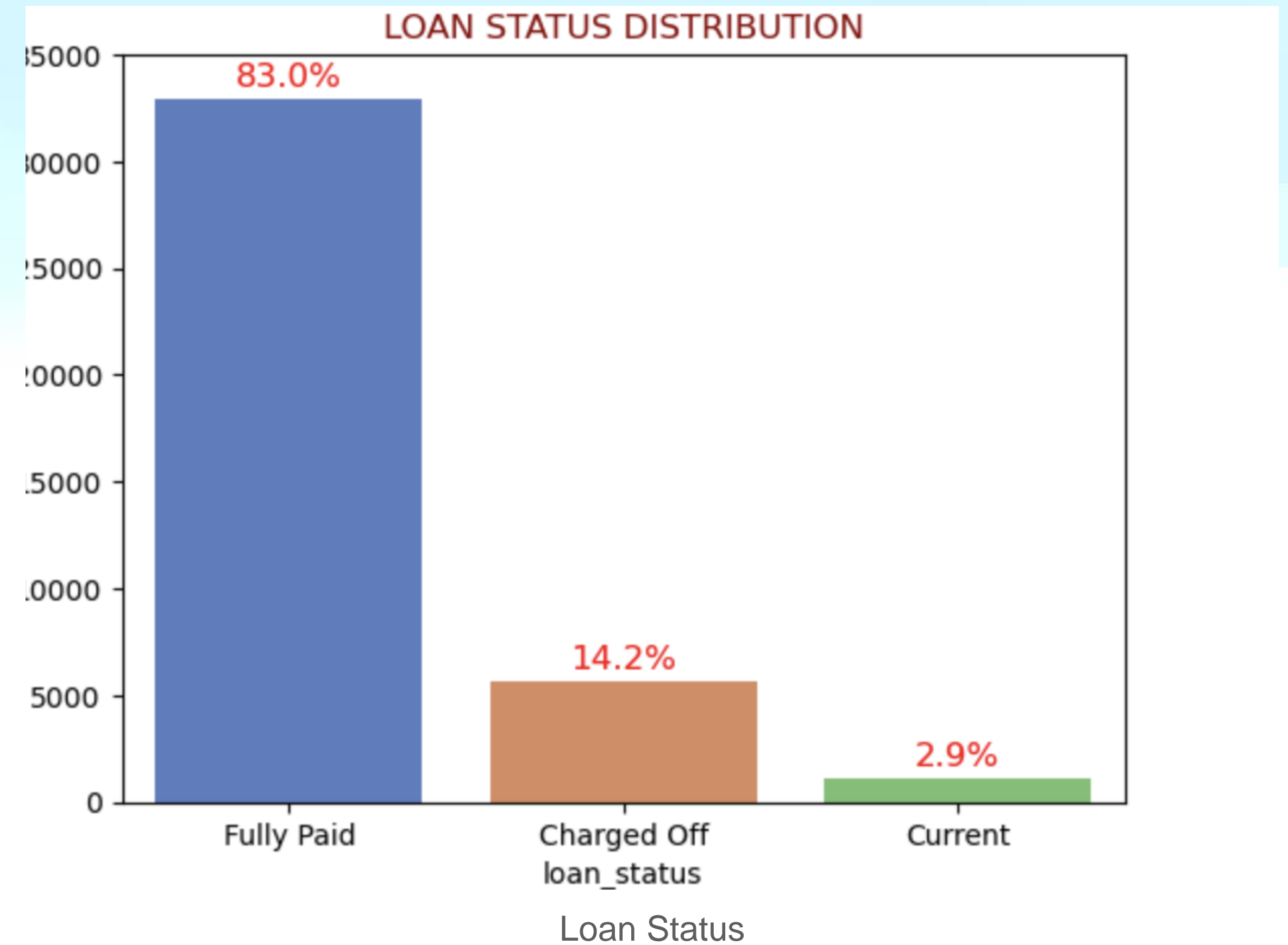
- The company wants to understand the driving factors (or driver variables) behind loan default
- The company can utilise this knowledge for its portfolio and risk assessment.
- The company wants to understand the strong indicators of default

Data Cleaning

- The original data has 39717 rows and 111 columns
 1. 54 Empty Columns
 2. 9 columns with only 1 unique value across rows
 3. 14 unnecessary fields
 4. 4 fields with value 0 in most rows and some outliers
 5. 2 fields with more than 90% null values
- Dropped all the above columns
- After cleaning up, the df_loan has 25 columns left.

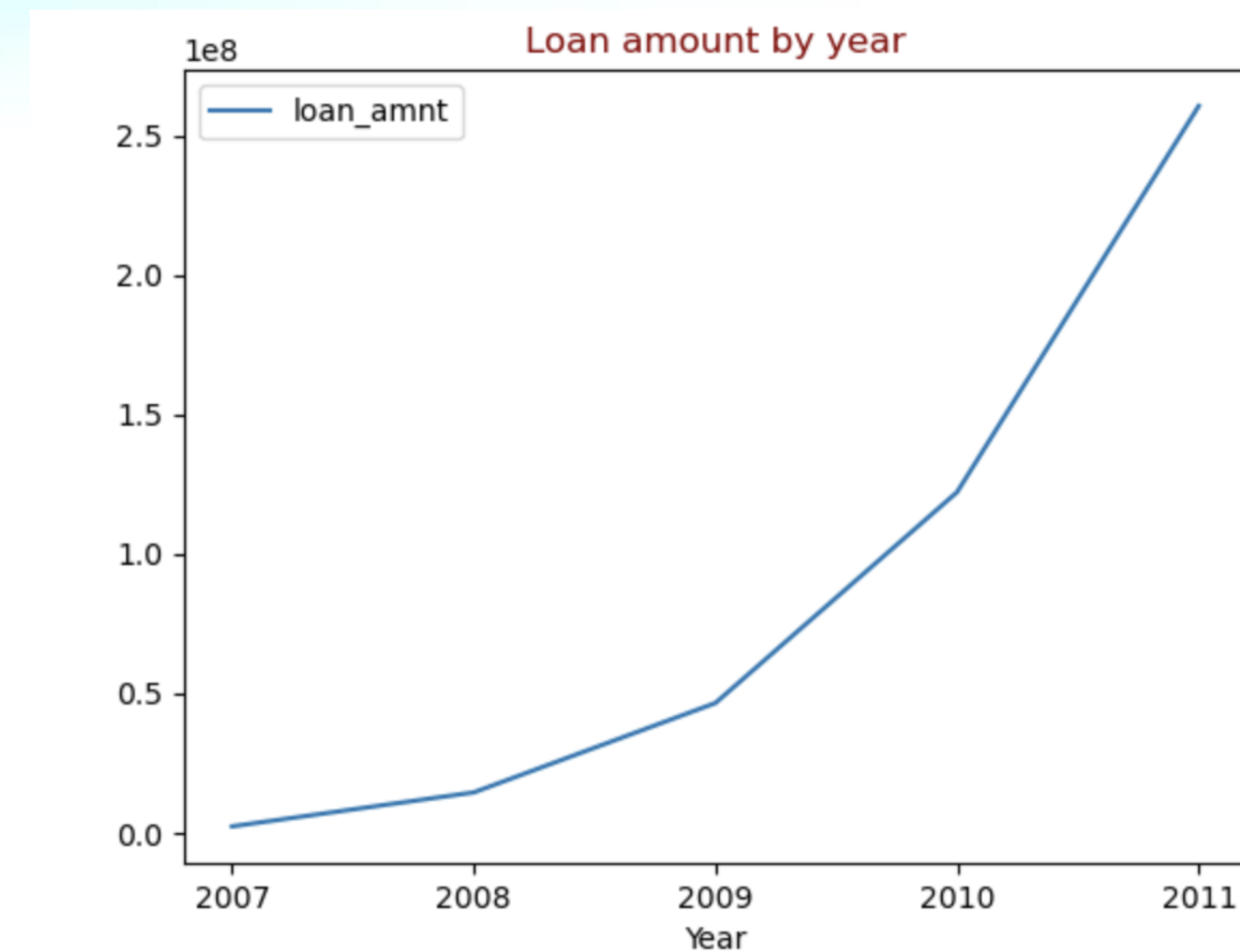
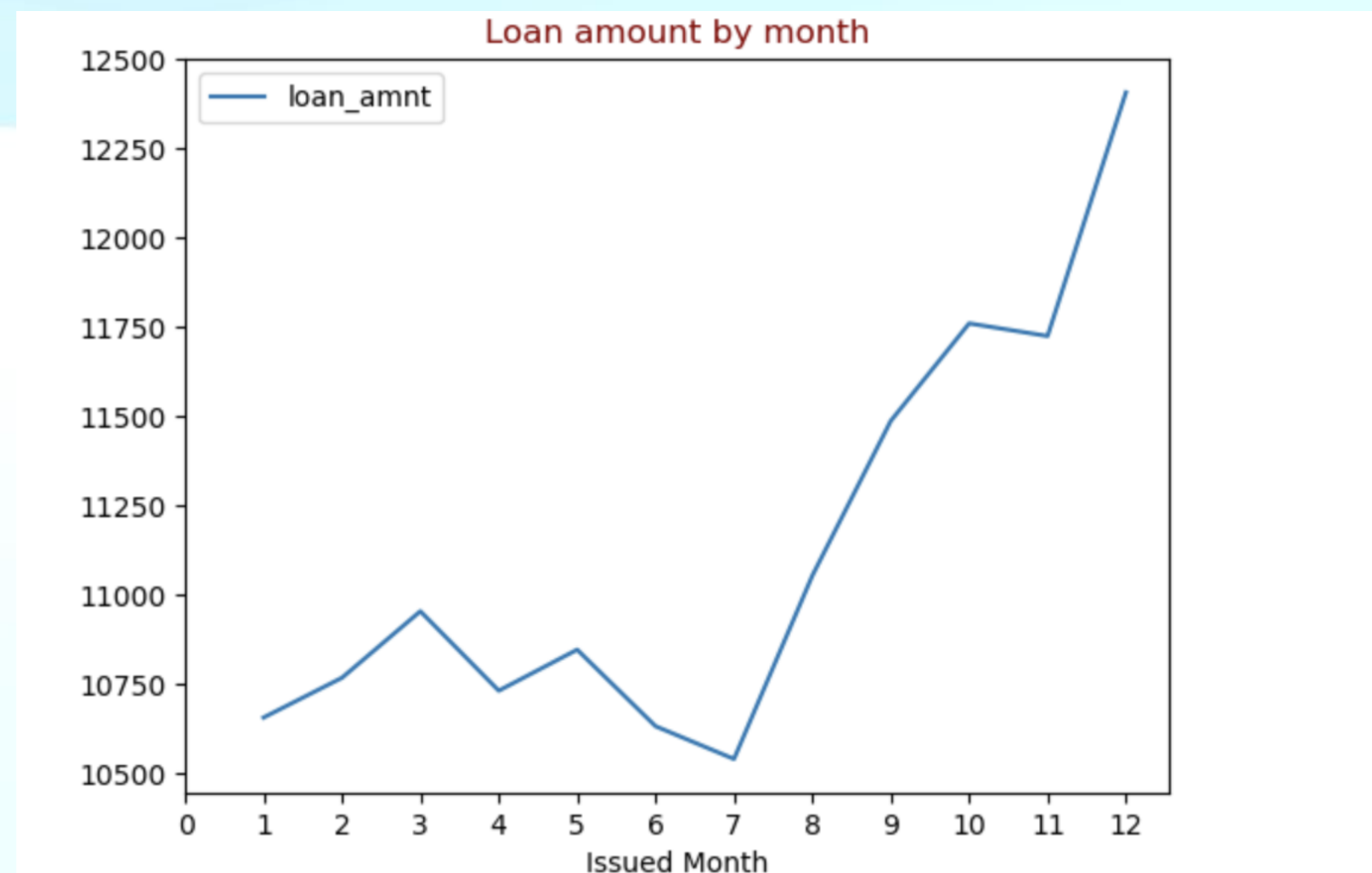
Loan Status

- 14.2 % of loans are defaulted



Loan vs Time by Year

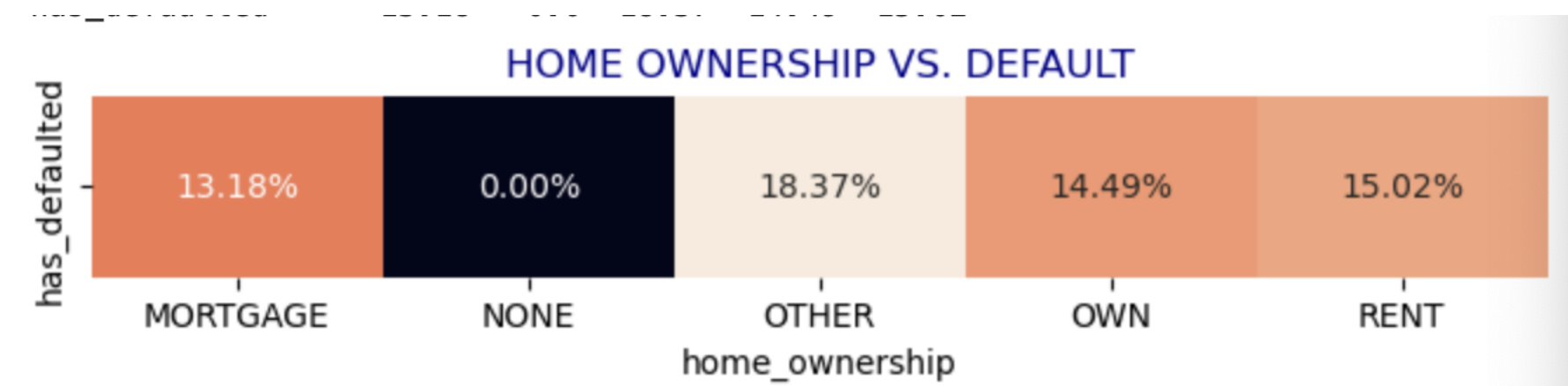
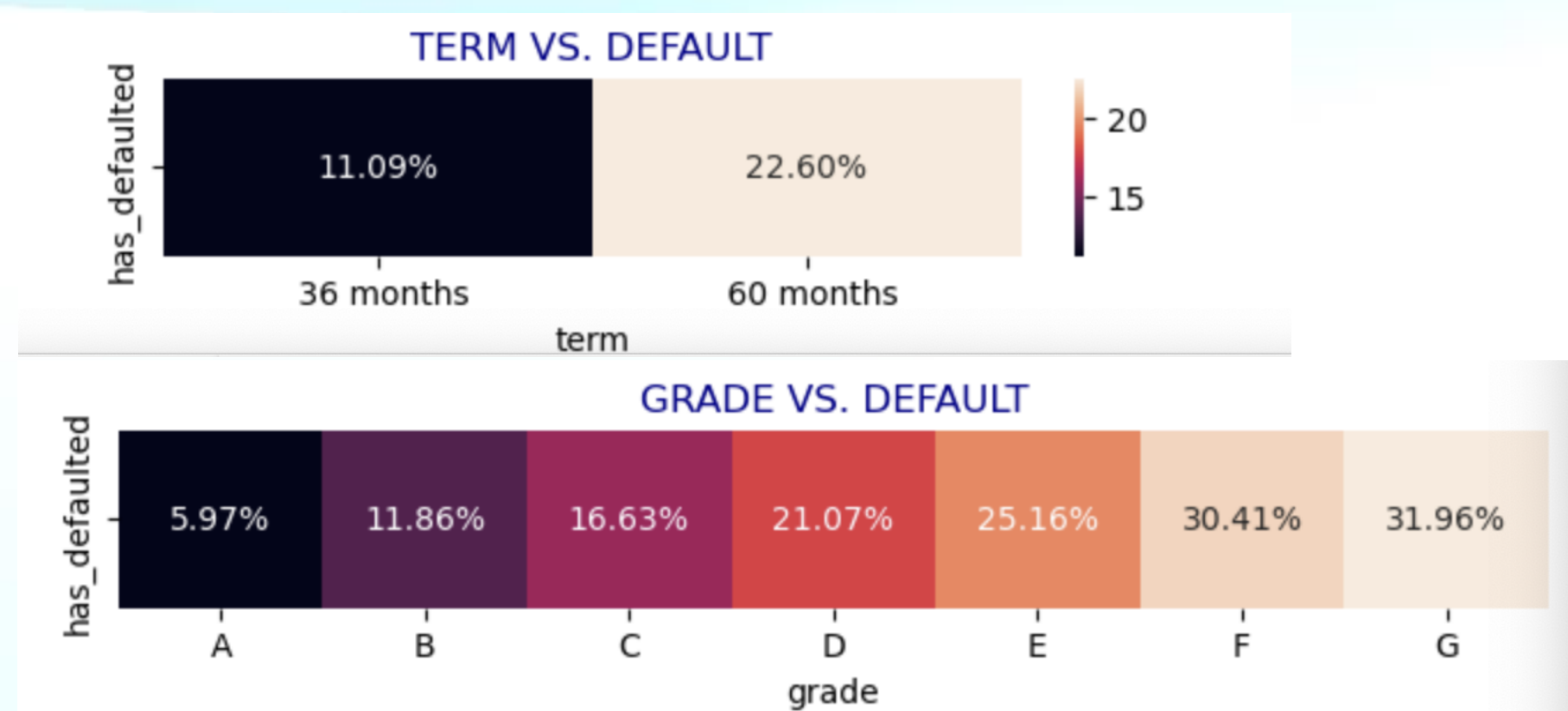
- Number of loan applications are increasing over time
- The number of applications seem to increase after 7th month



Term Distribution, Grade, Home Ownership

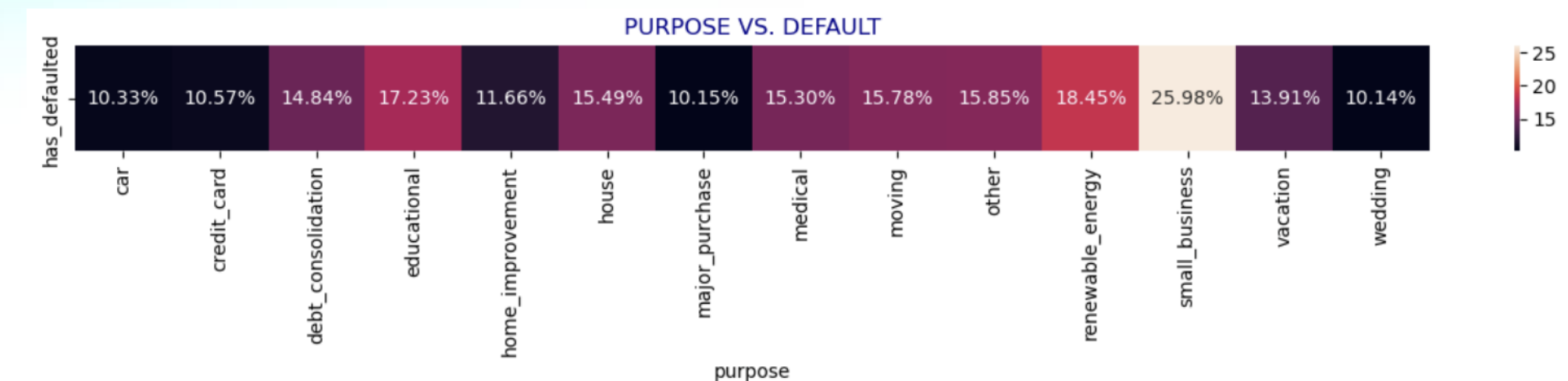
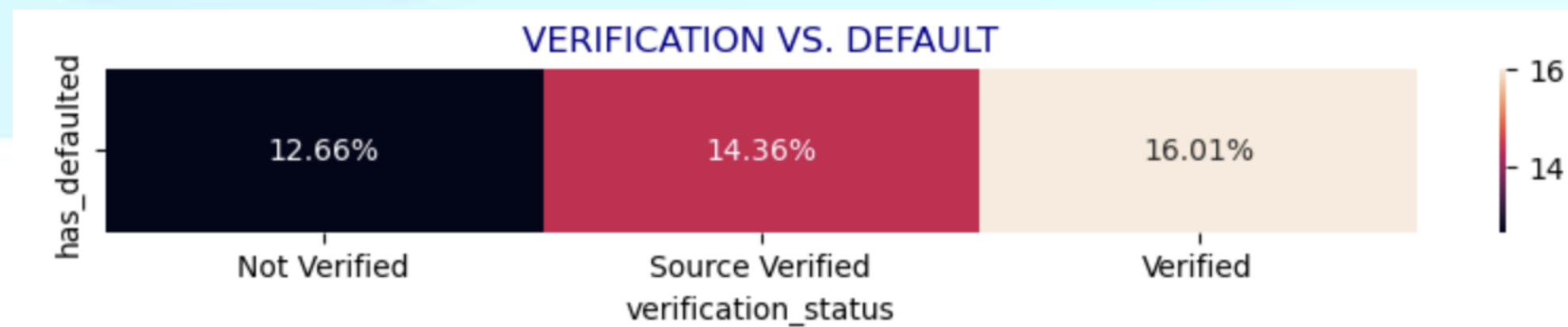
For Default loans

- The loan borrowers with 60 months(22.60 %) term have comparatively defaulted the loan then 36 months (11.09 %)
- The loan borrowers in Grade G(31.96 %), F (30.41 %), E(25.61%) have defaulted comparatively to other grades
- The other(18.37%) home ownership defaulter is greater than rental (15.02%) and mortgage(13.18%)



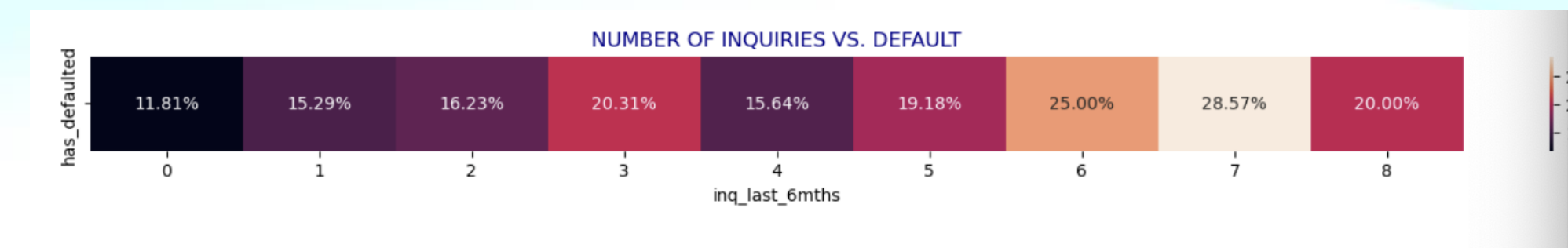
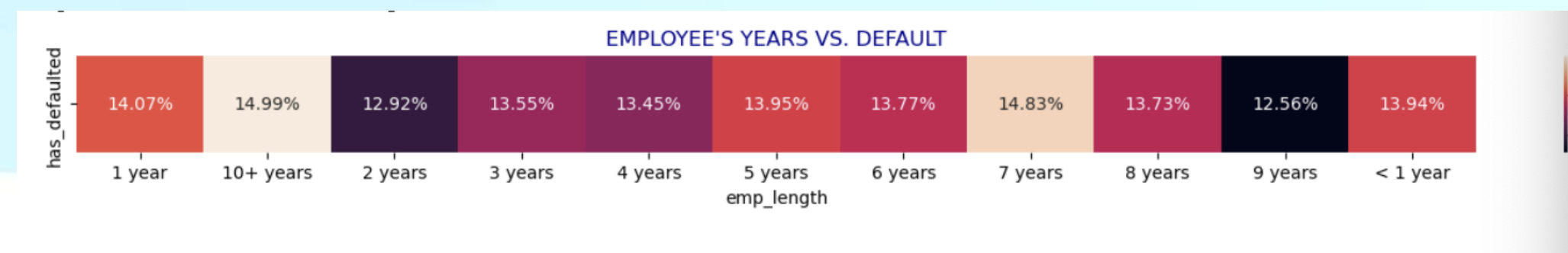
Loan Verification, Purpose for Default loans

- Loan defaulters is greater in verified(16.01%) source than not verified (12.66%)
- The loan borrower defaulted had loan purpose as small business (25.98%)



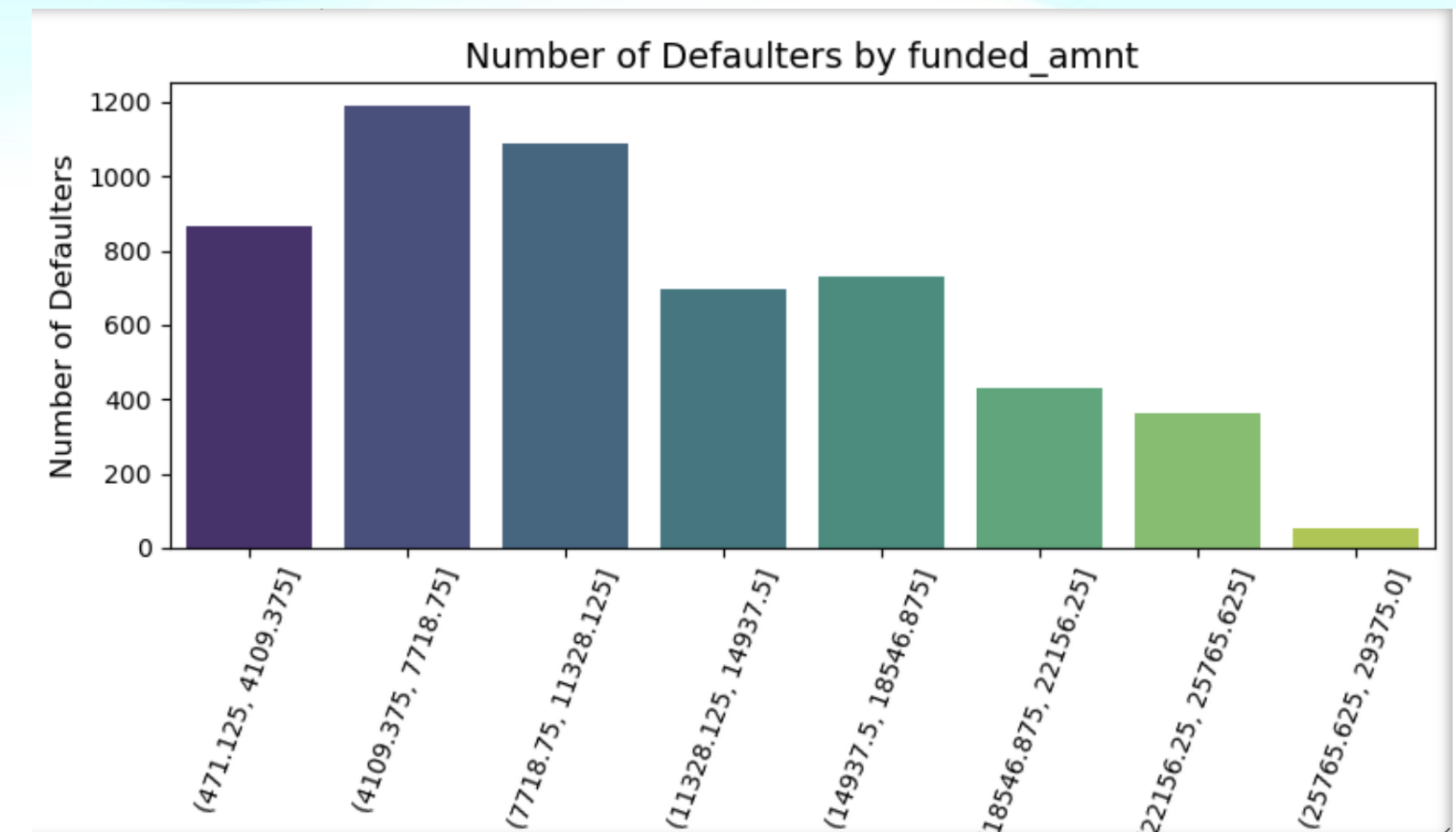
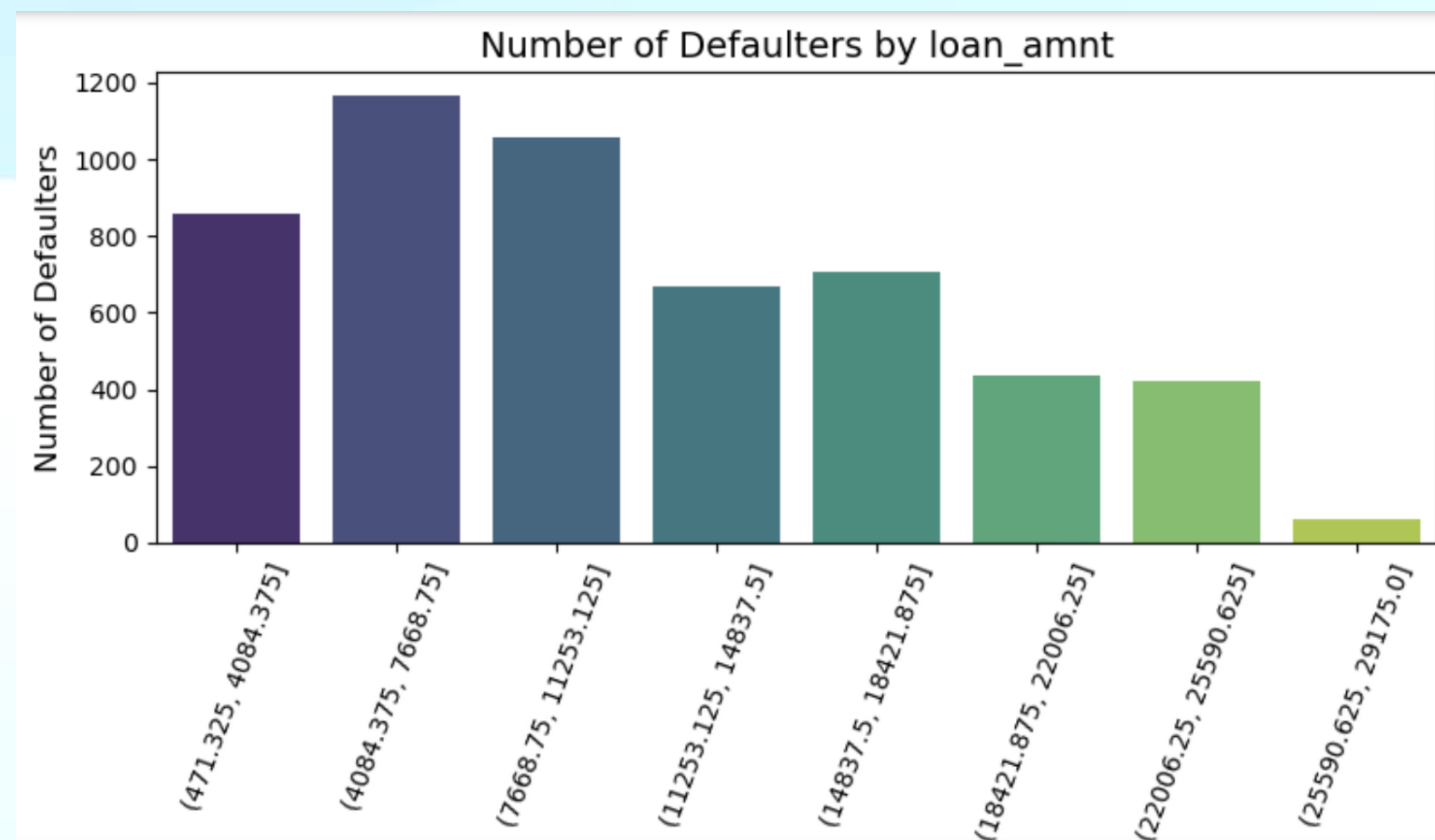
Employee length and number of enquires

- The loan defaulting is greater in employee length 10+ years (14.99%)
- The number of inquires 3 in last 6 months is greater 20.31%



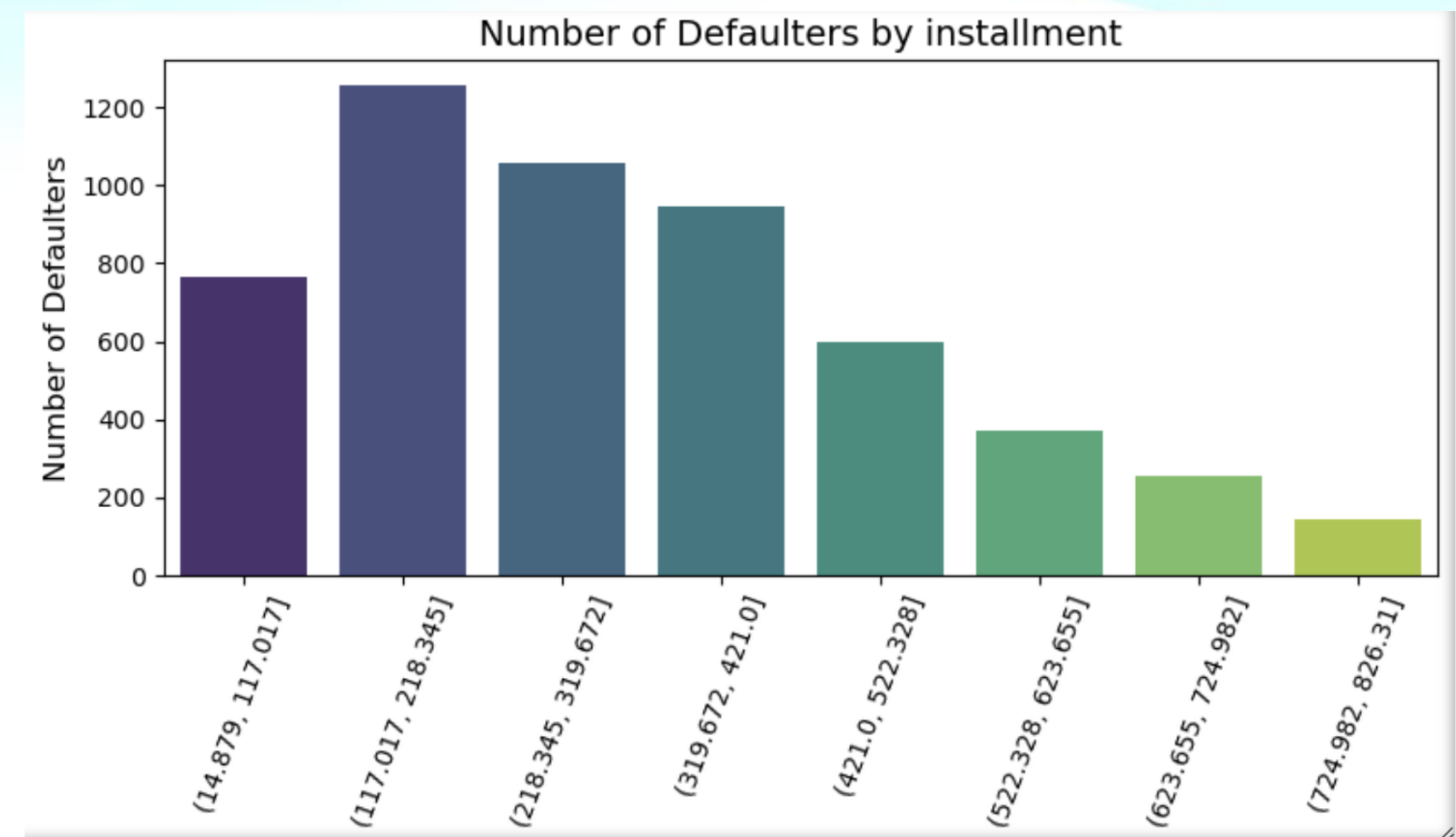
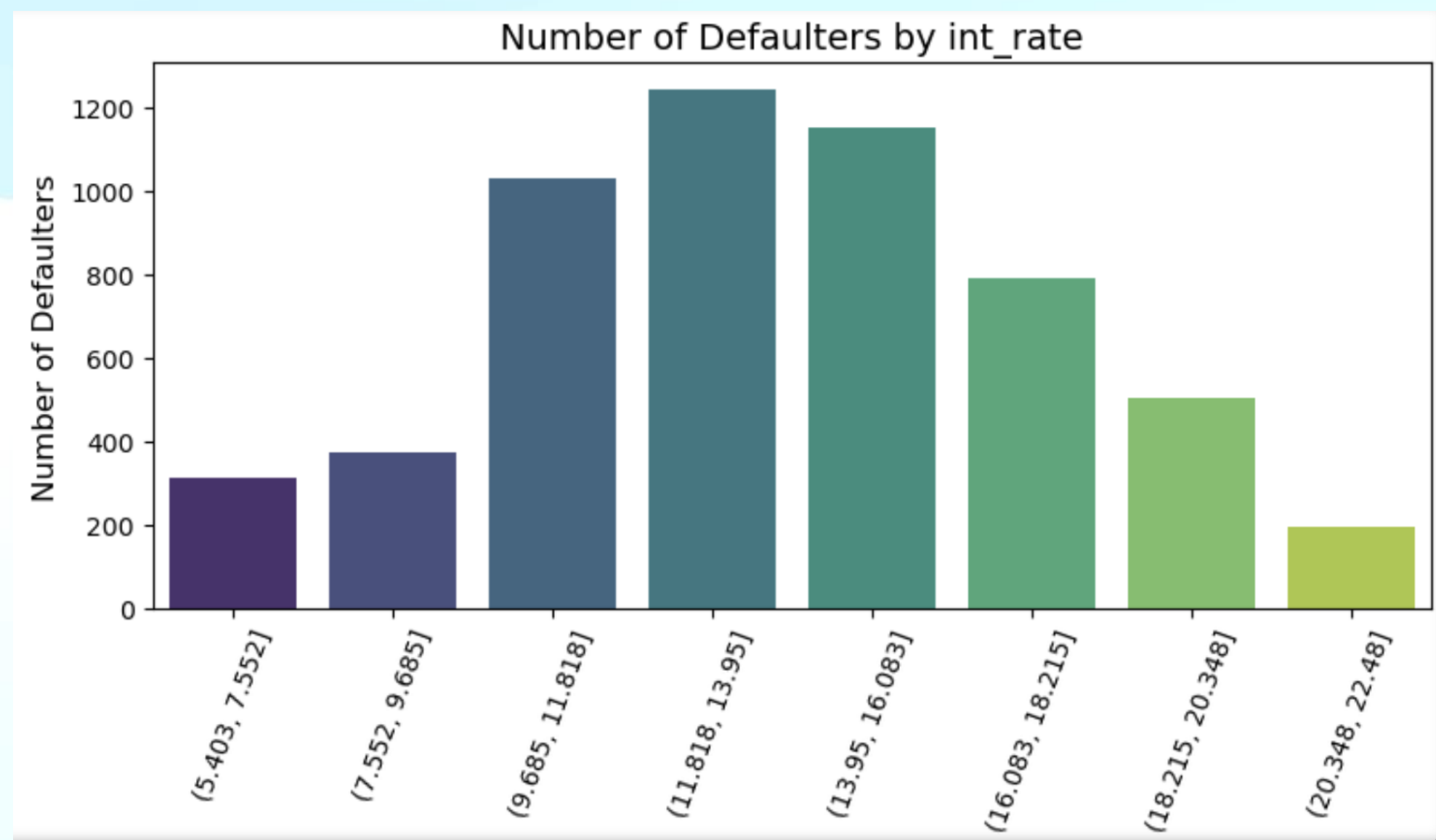
Loan Amount, Funded Amount for default loan

- The number of defaulters is comparatively high in loan amount range 4084.375 to 7668.75
- The number of defaulters is comparatively high in funded amount range 4109.375 to 7118



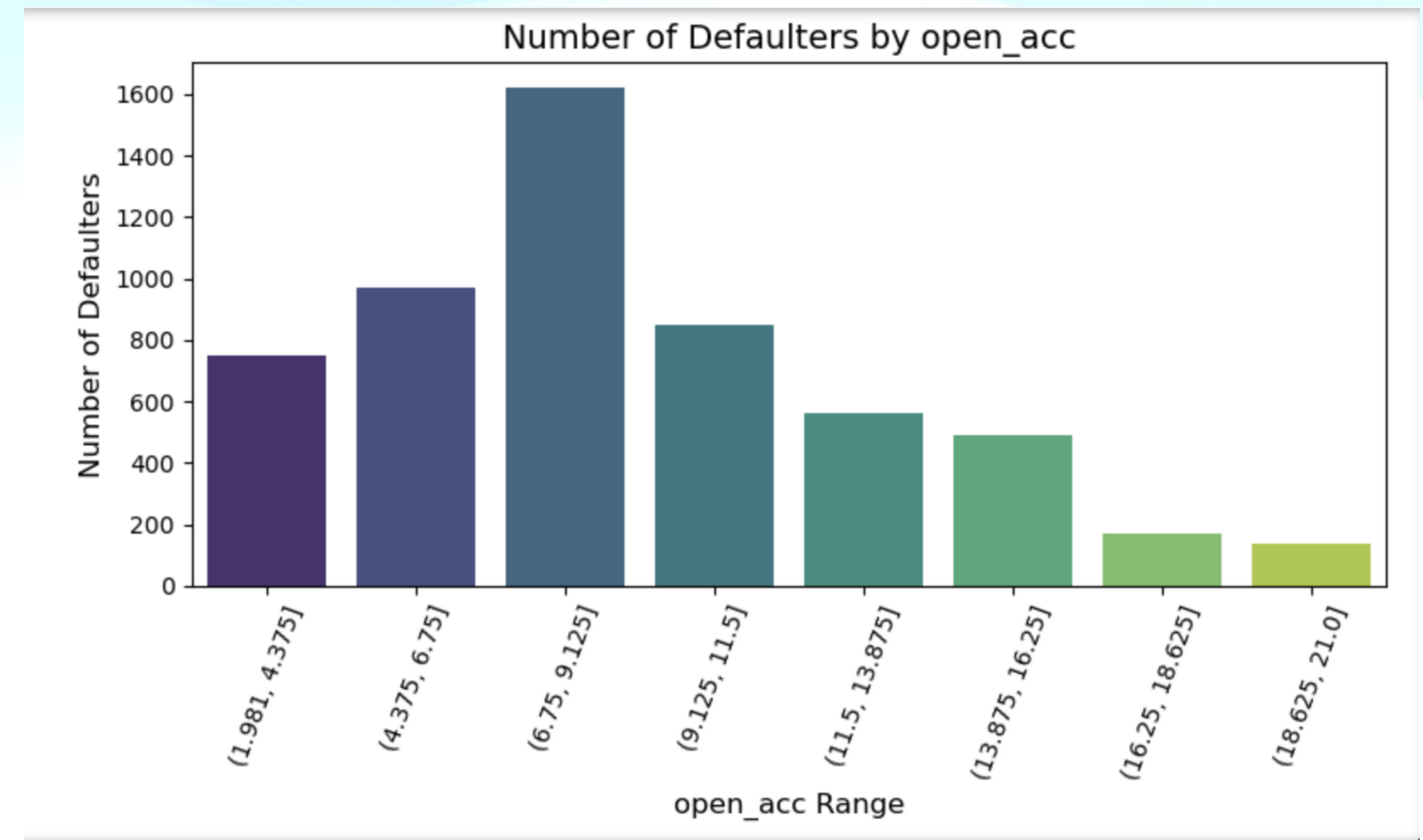
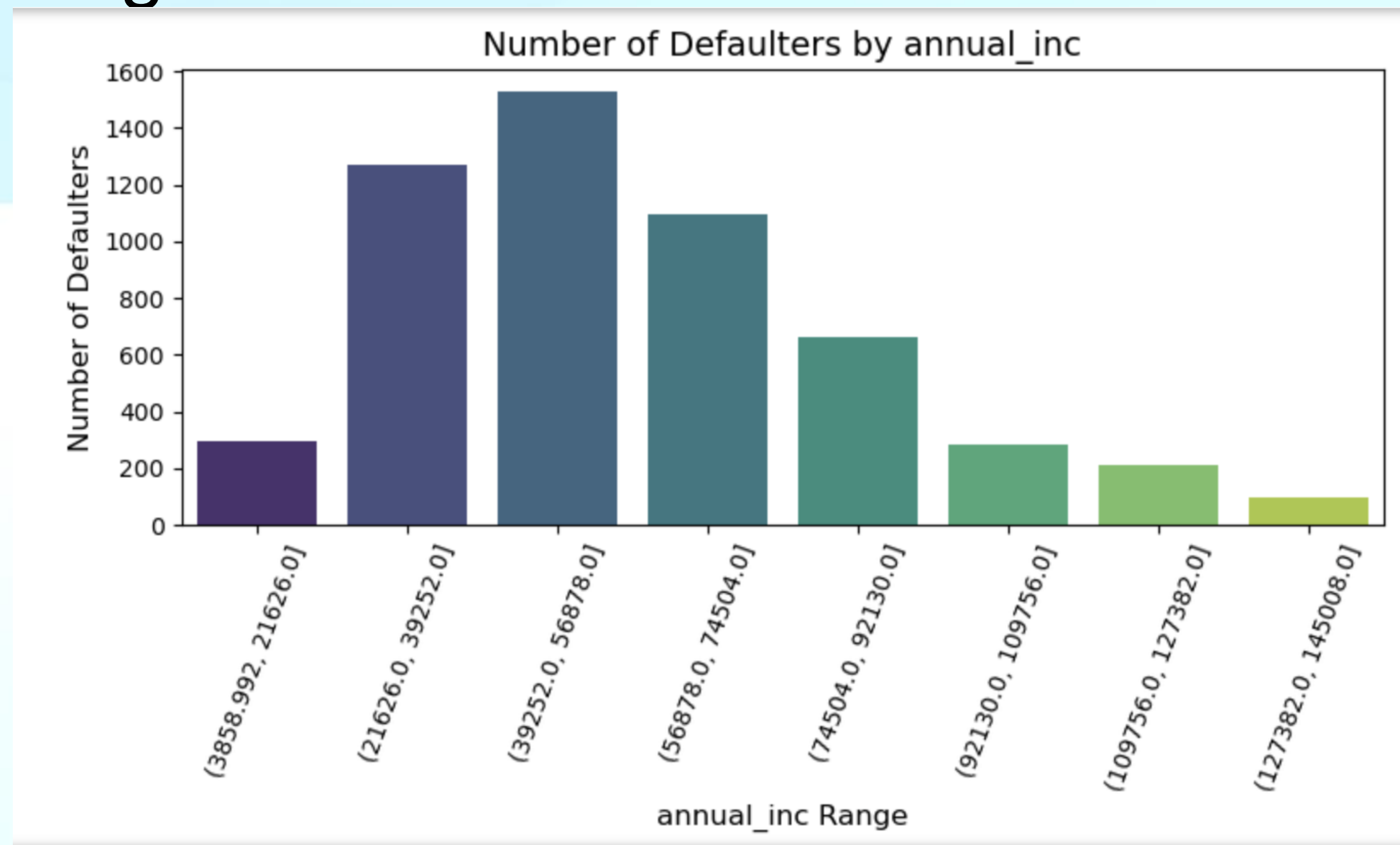
Interest rate, Installment for default loan

- Interest rate in range 11.8 to 13.95 has comparatively high defaulters
- For installment range 117.017 to 218.354 has high defaulters



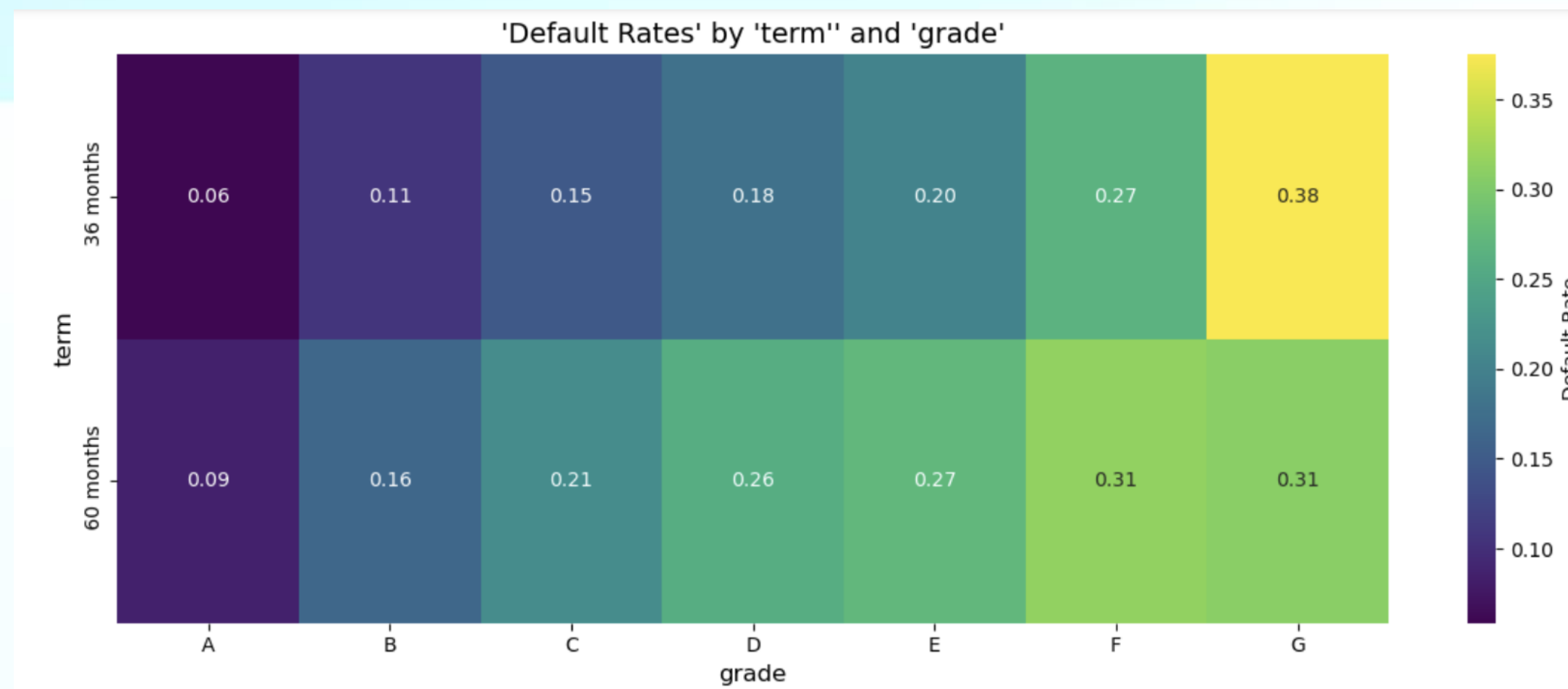
Annual Income, Open Accounts for loan defaulter

- The loan borrowers having annual income in range 39252.0 - 566878.0 have comparatively higher defaulters
- The loan borrowers having open account in range 6 to 9 have comparatively higher defaulters



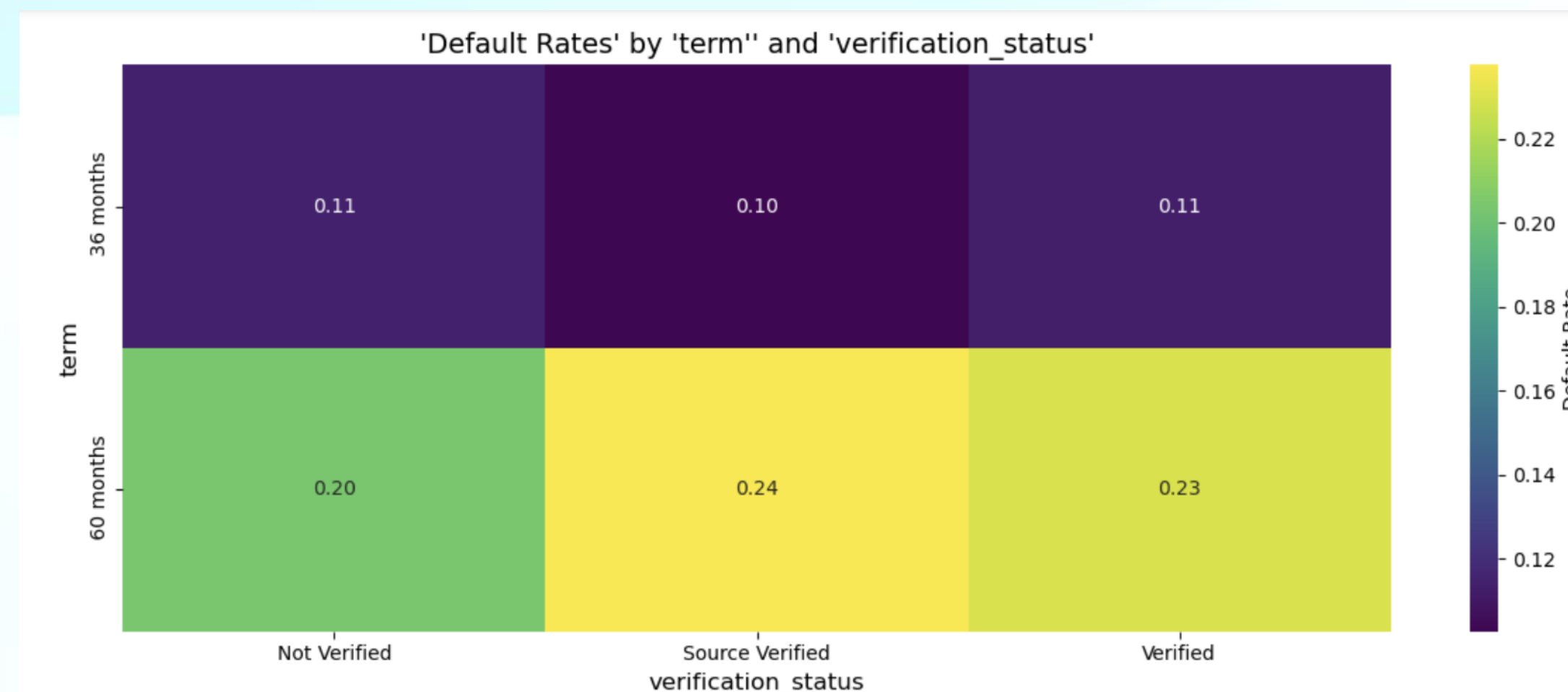
Default rate by term and grade

- For the term 36 months and grade G shows higher default rate -0.38
- For term 60 months and grade G shows higher default rate - 0.31



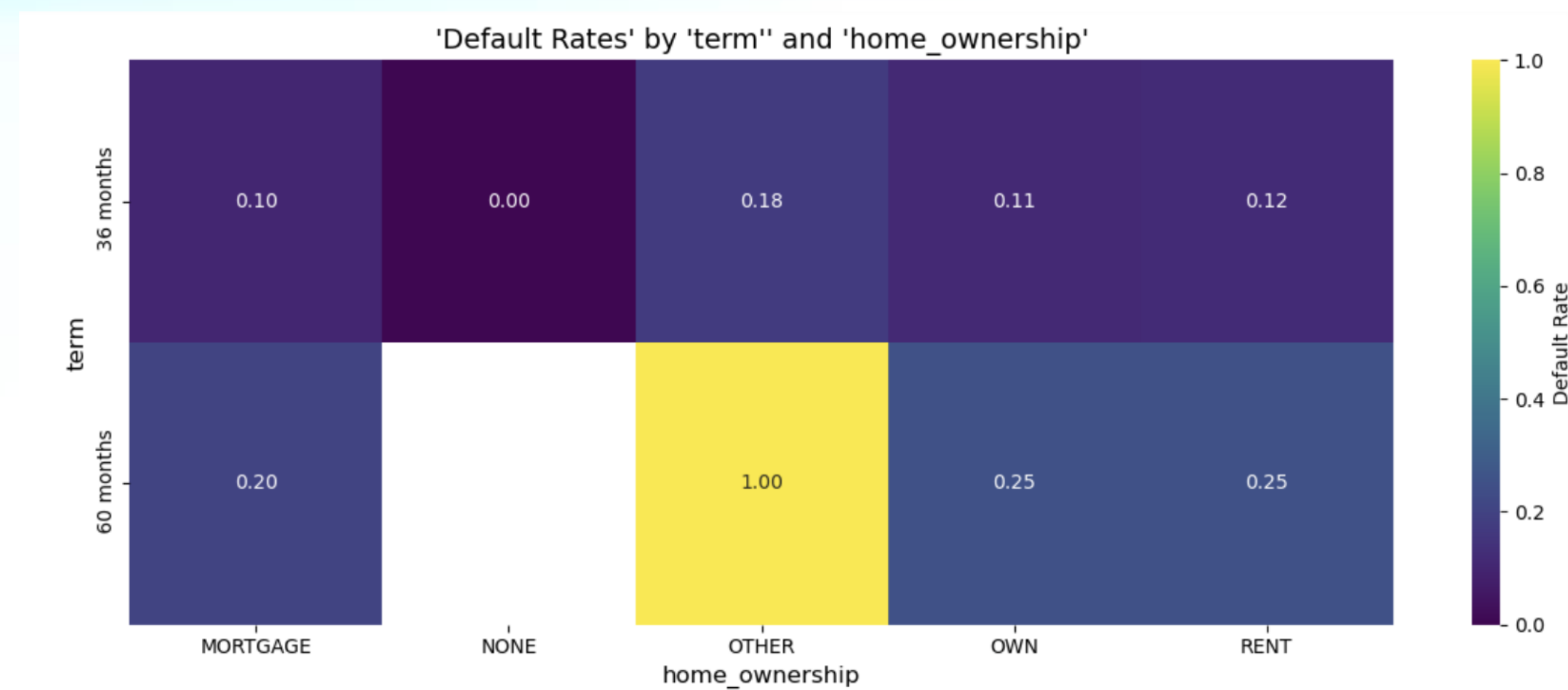
Default rate by term and verification status

- For term 60 months and source verified the default rate is higher 0.24 followed by verification source as verified which 0.23



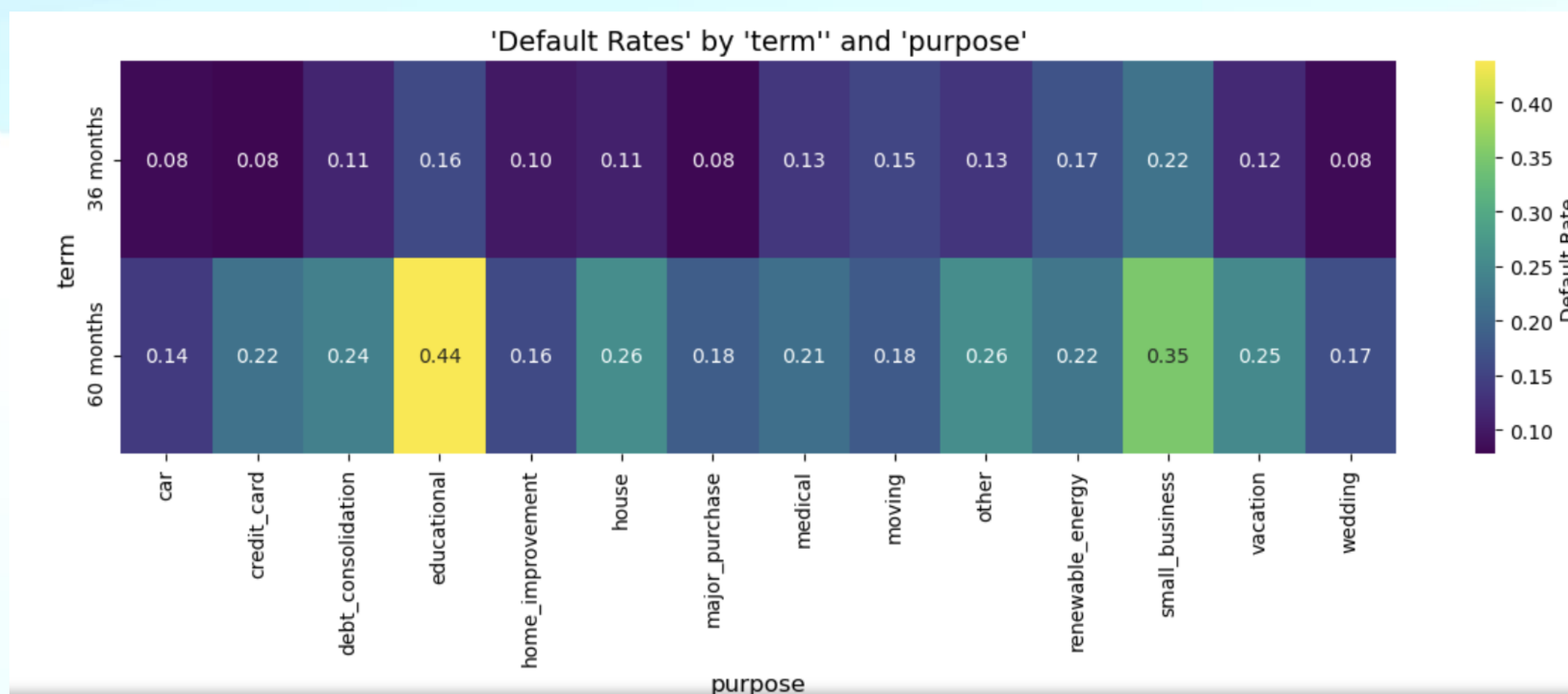
Default rate by term and home ownership

- The default rate for term and home ownership is comparatively higher for home ownership as other 1.00



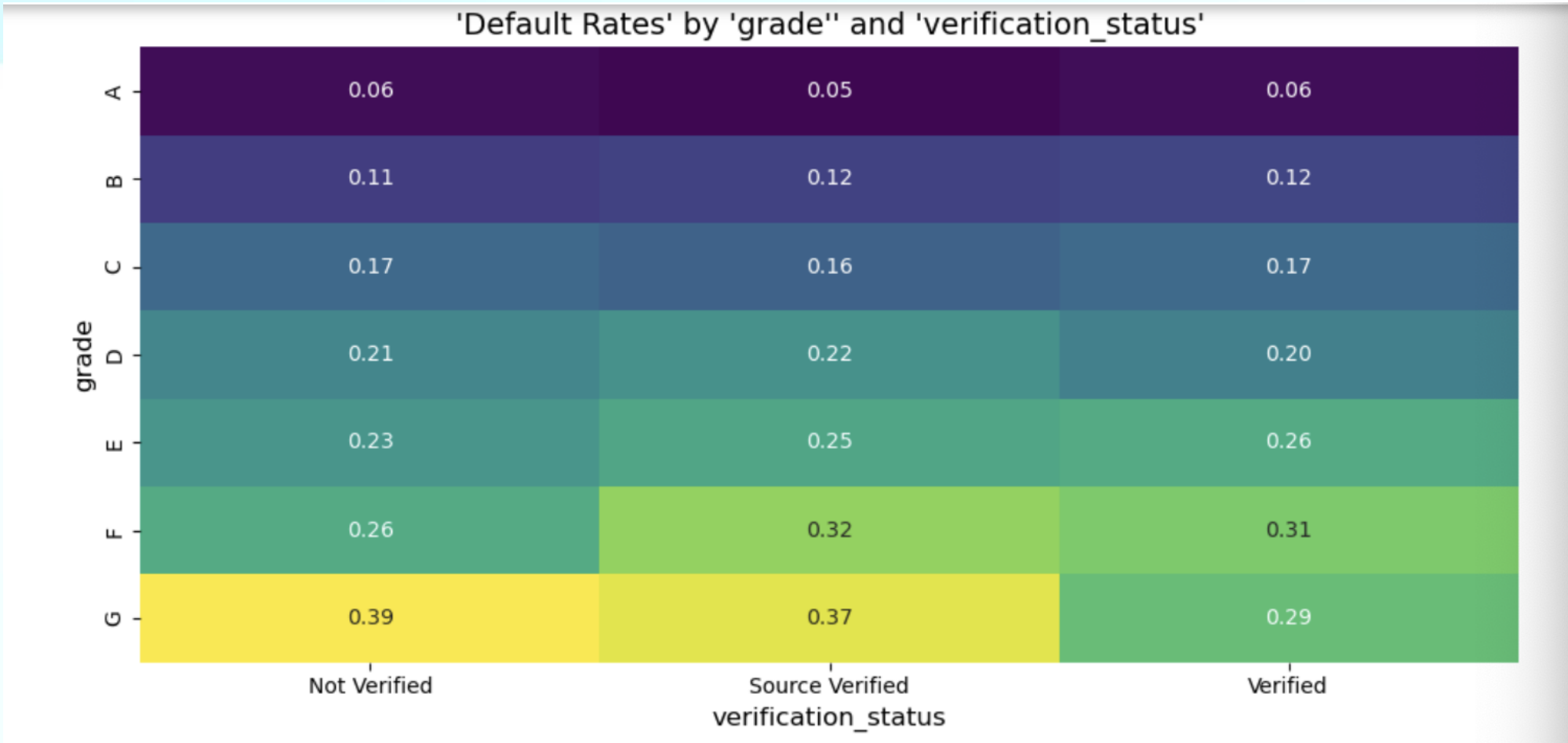
Default rate by term and purpose

- Defaulter rate is significantly higher for purpose 'educational' for term 60 months



Default rate by grade and verification status

- Default rate is high for Grade G for not verified status, followed by source verified and verified



Key Insights

- The loan borrowers from Grade G, F and E have comparatively defaulted the loan
- Interest rate with range form 11-19 % showed more defaulters
- The default rate is high for grade G for 60 months term
- The default rate is high for home ownership 'other' for 60 months term