Consumer Finance Company Loan Data Analysis

EDA Case Study

upGrad – IIITB ML and AI Program - Feb 24 batch

Siva Vaddadi & Yukta Bajaj

Analysis Approach

- Data Understanding: Load and inspect the loan dataset.
- Data Cleaning: Handle missing values and drop irrelevant columns.
  - Drop columns having more than 35% null values
  - Drop columns having 100% duplicate values
  - Fill null values with median or mode
  - Do not consider rows with outlier values such as annual income > 0.2 million
  - Remove text from numeric columns such as percent sign
- Univariate Analysis: Explore the distribution of each attribute after cleaning up such as loan amounts and interest rates, etc.
- Bivariate Analysis: Investigate the relationship between columns such as loan amount and loan status.
- Insights & Recommendations: Draw insights and provide recommendations for mitigating default risk.

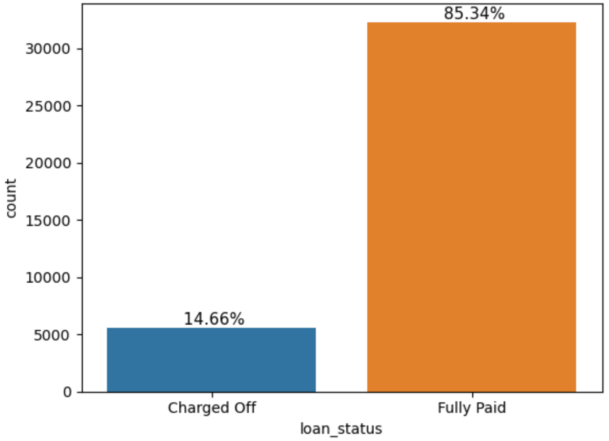
Explanation of terms

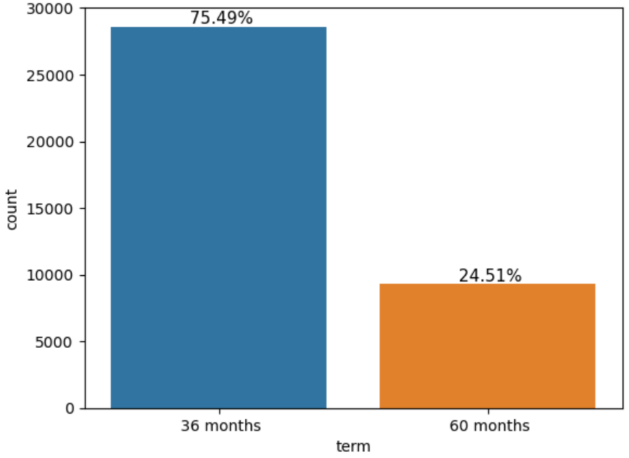
- The expressions "most occurrences" or "mainly falling within" suggest the Interquartile Range (IQR) or the range spanning from Q1 to Q3 on the box plot for that particular column.
- “Average” means Median value of the column/attribute
- Grade “A” is lowest, and “F” is highest. “A” is lower than “B” which is lower than “C” and so on
- “~” (tilde symbol) represents around that number. For example, ~25 means around 25
- Most of the numbers and percentages are rounded off to nearest integer for easy reading

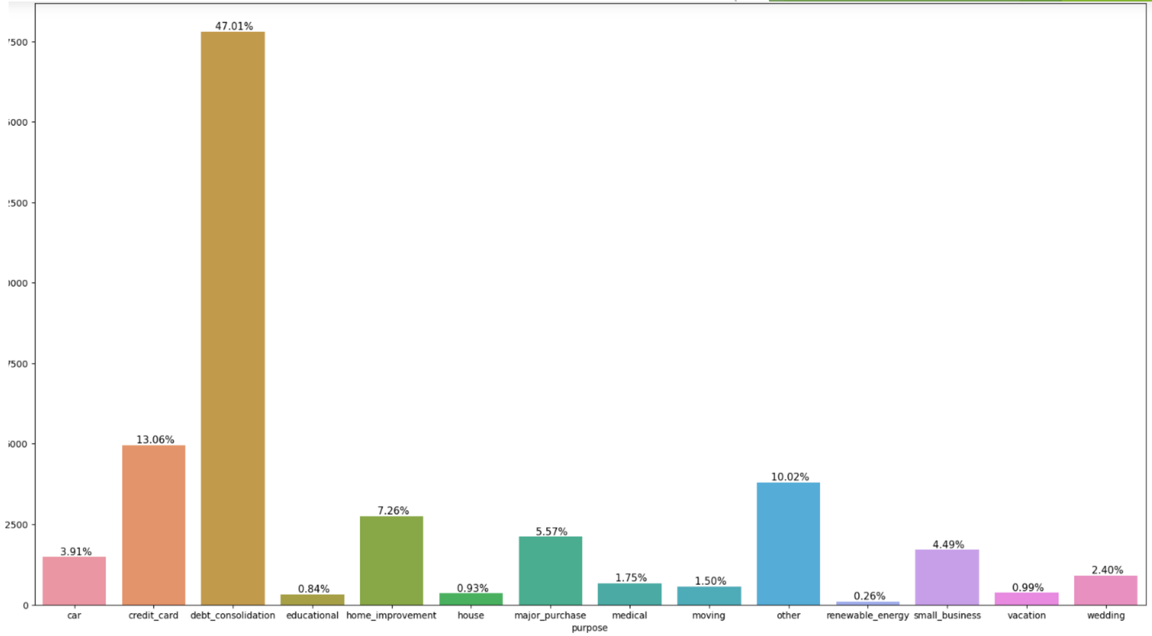
Observations about Loan data

- Charged off loans stand at 14.66% and this needs to be brought down.
- The Consumer Finance Company has been giving loans mostly in the range of $5 to $15k with a spike at every 5k. Average loan amount is $9600
- Interest rates mostly range between ~9% to 14%. Average interest rate is around 12%
- Most of the Installment amounts range between $165 to $420. Average installment is ~$275
- Loans are typically extended to borrower's whose annual income ranges from $40,000 to $80,000, with an average annual income of $58,000.
- The number of loans granted with a 36-month tenure at ~75% is nearly three times the number granted with a 60-month tenure (~25%).
- Most number of loans are taken for the purpose of debt consolidation (47%), followed by credit card debt (13%). Small business, debt consolidation, house, credit card purposes got big loan amounts, at higher interest rates and installment amounts. Home improvement, small biz, renewable energy and wedding borrowers have higher annual incomes.

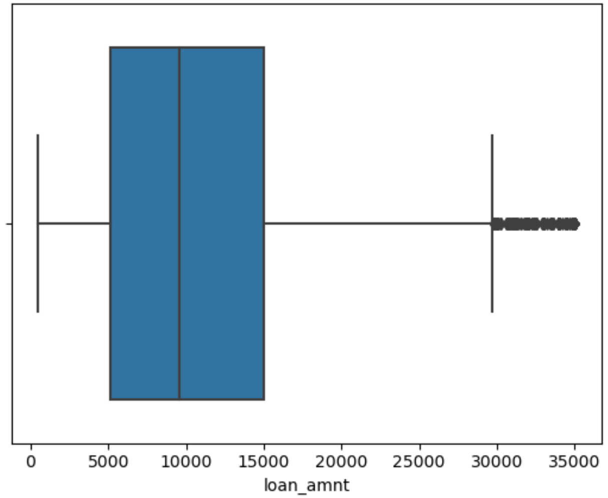
Observations about Loan data (Charts)



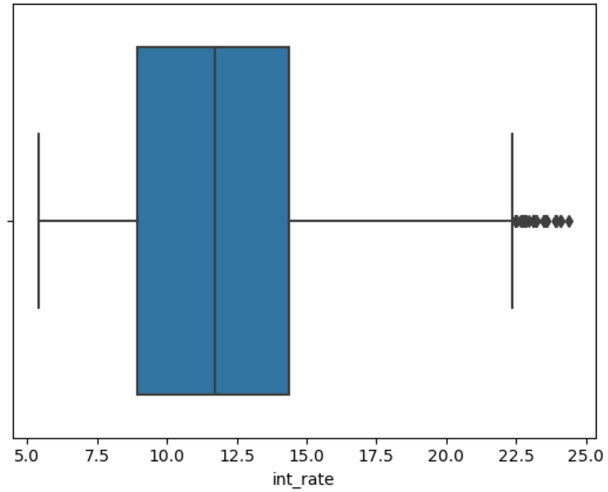




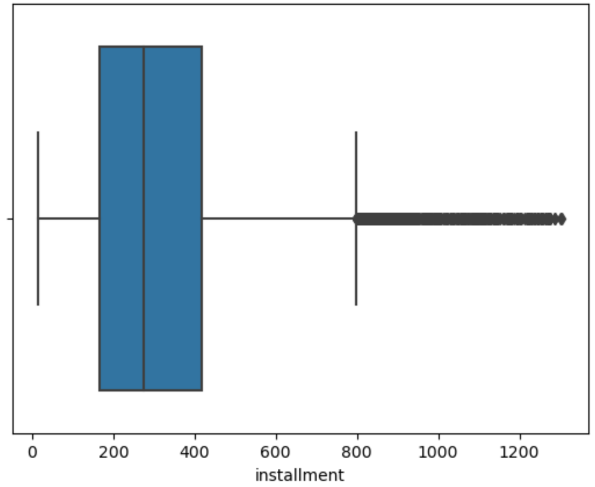

```
Median: 9600.00
q1: 5150.00
q3: 15000.00
```



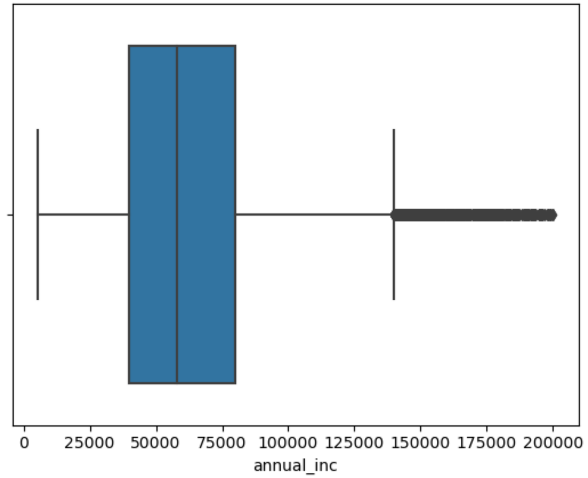

```
Median: 11.71
q1: 8.94
q3: 14.35
```




```
Median: 275.28
q1: 164.86
q3: 418.50
```



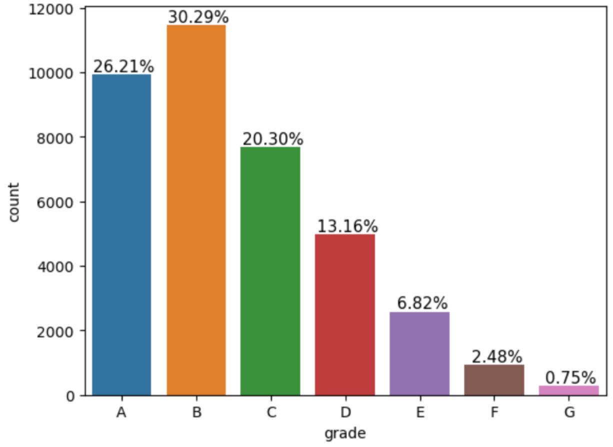

```
Median: 57906.00
q1: 40000.00
q3: 80000.00
```

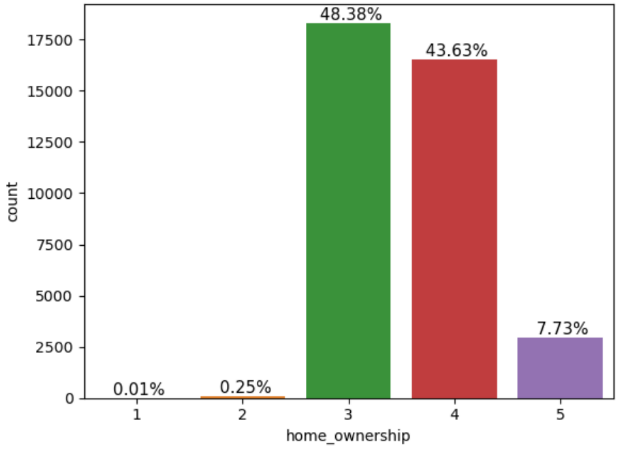


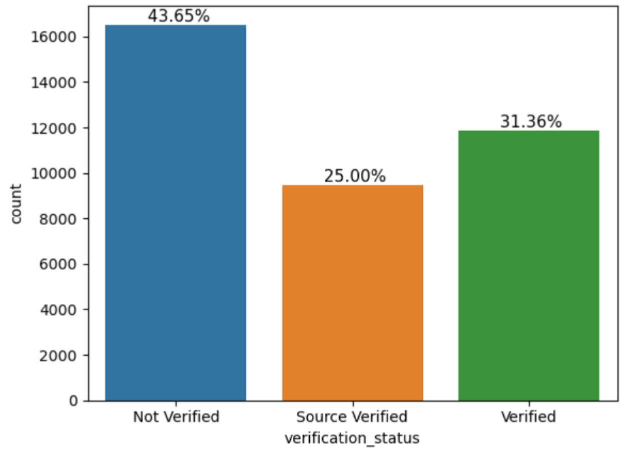
Observations about Consumer data

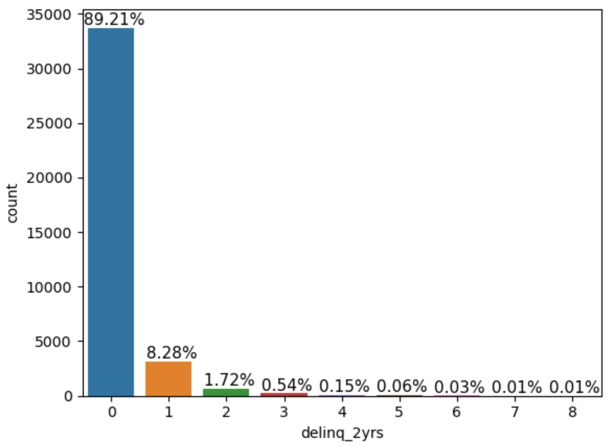
- Borrower Debt To Income ratio mostly ranges between ~8 to 18. Average is 13.47
- Most borrowers are rated as B grade (30%), followed by A (26%), C (20%), D (13%), E (7%), F (3%), and G (1%).
- The majority of loans are allocated to borrowers who either pay rent (48%) or a mortgage (44%), with homeowners comprising the remaining 8%.
- Most loans are extended to *unverified borrowers (43%)*, followed by verified borrowers (32%), and source verified borrowers (25%).
- 95% of borrowers have 0 public records, while 5% have 1. Additionally, 96% have 0 public bankruptcy records, with 4% having 1.
- 89% of borrowers do not have 30+ days past due incidences in last 2 years. 8% have had one such incident, 2% have had two, and 1% have had three.
- Over the last 6 months, 49% of borrowers did not make any inquiries, while 28% made one, 15% made two, and 8% made three inquiries.

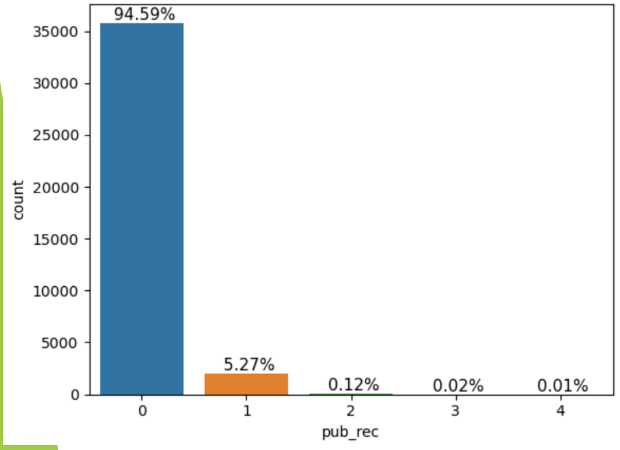
Observations about Consumer data (Charts)

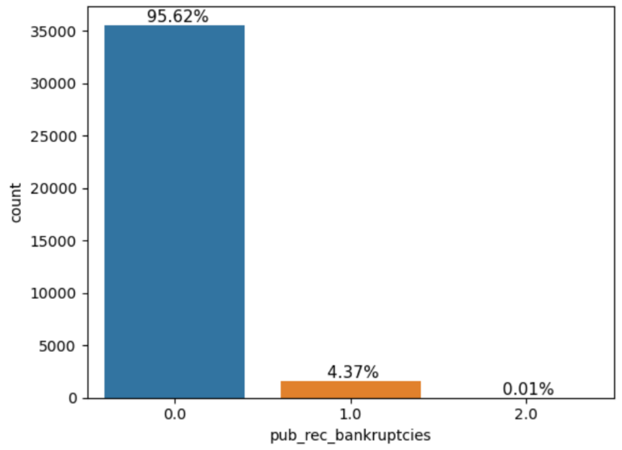


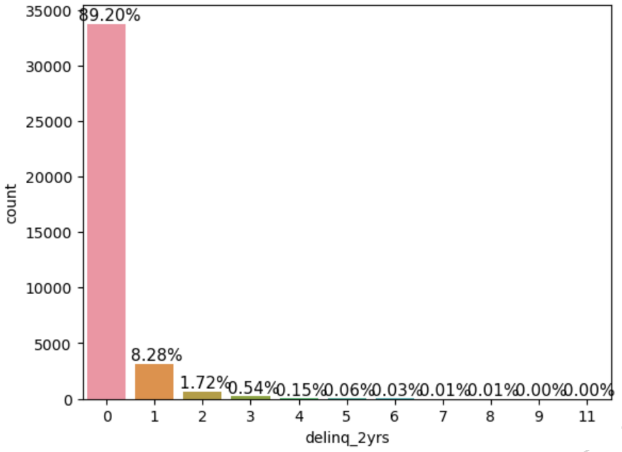


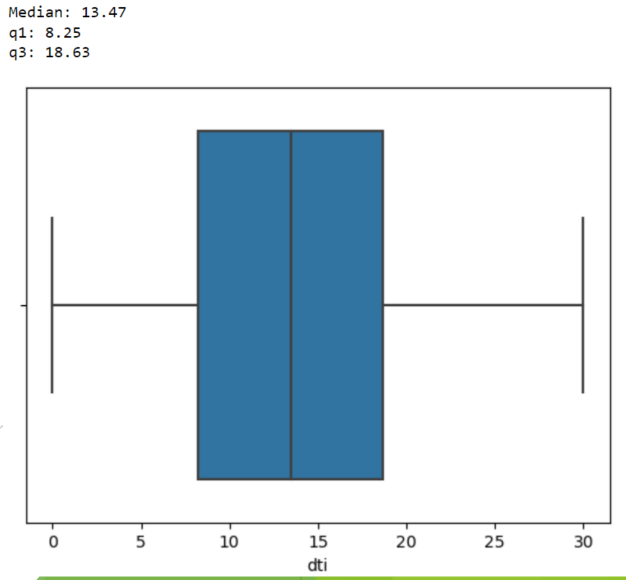






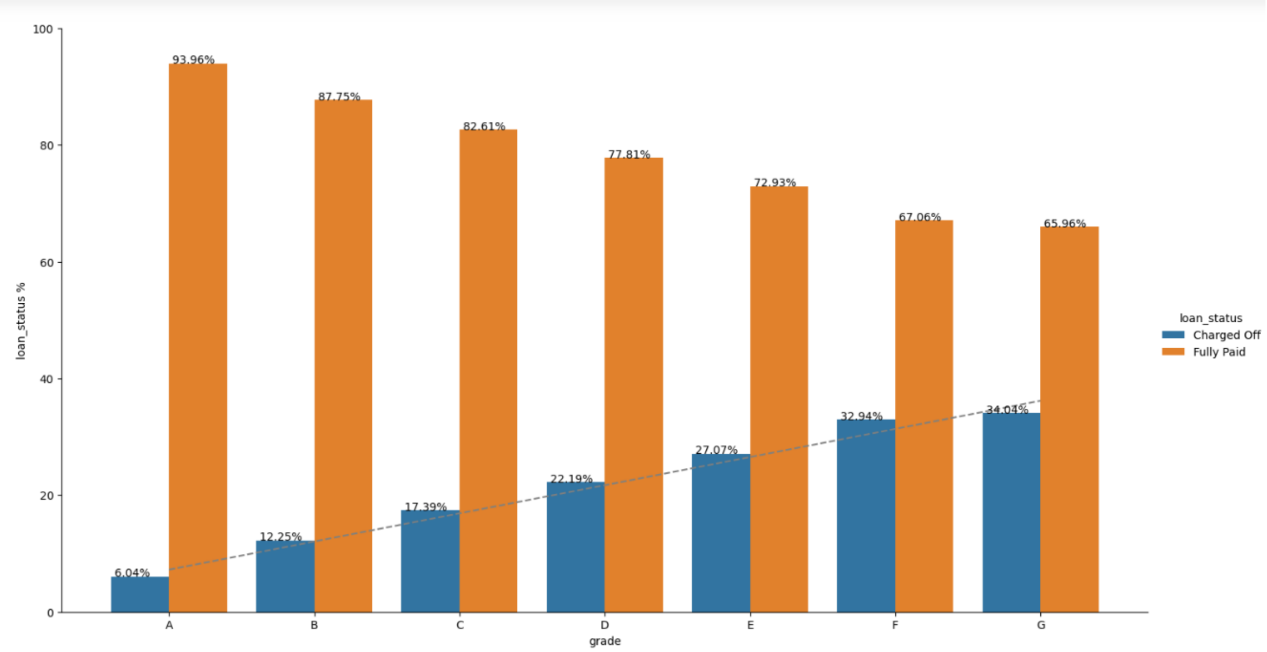


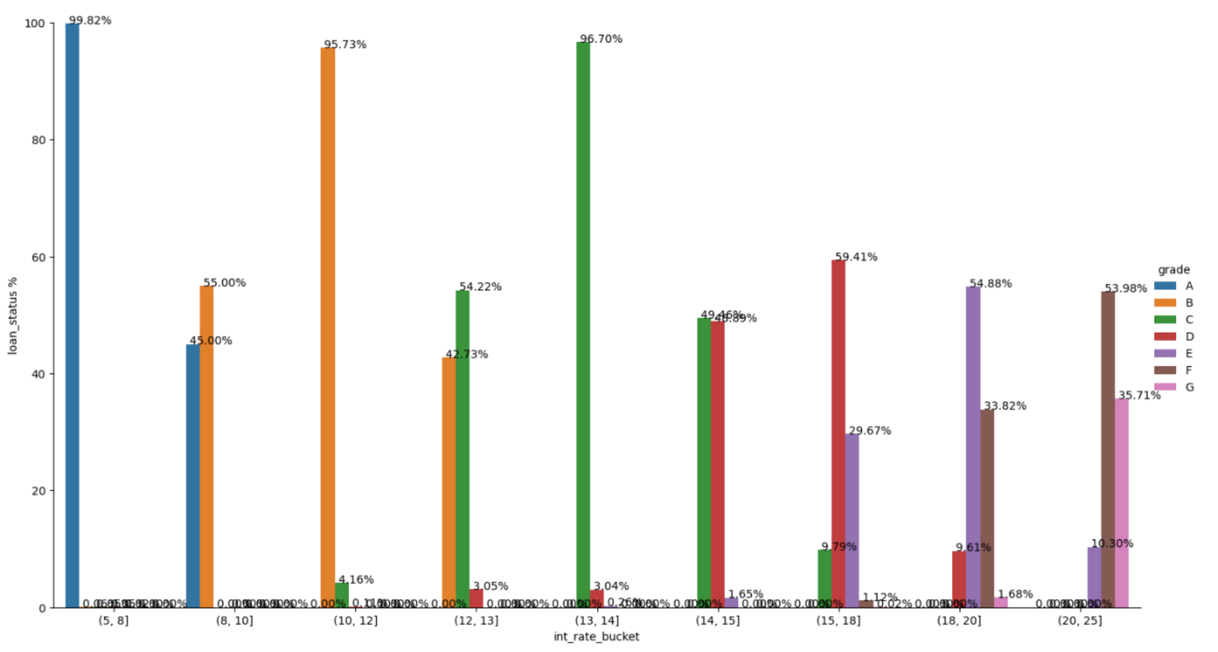




Conclusions

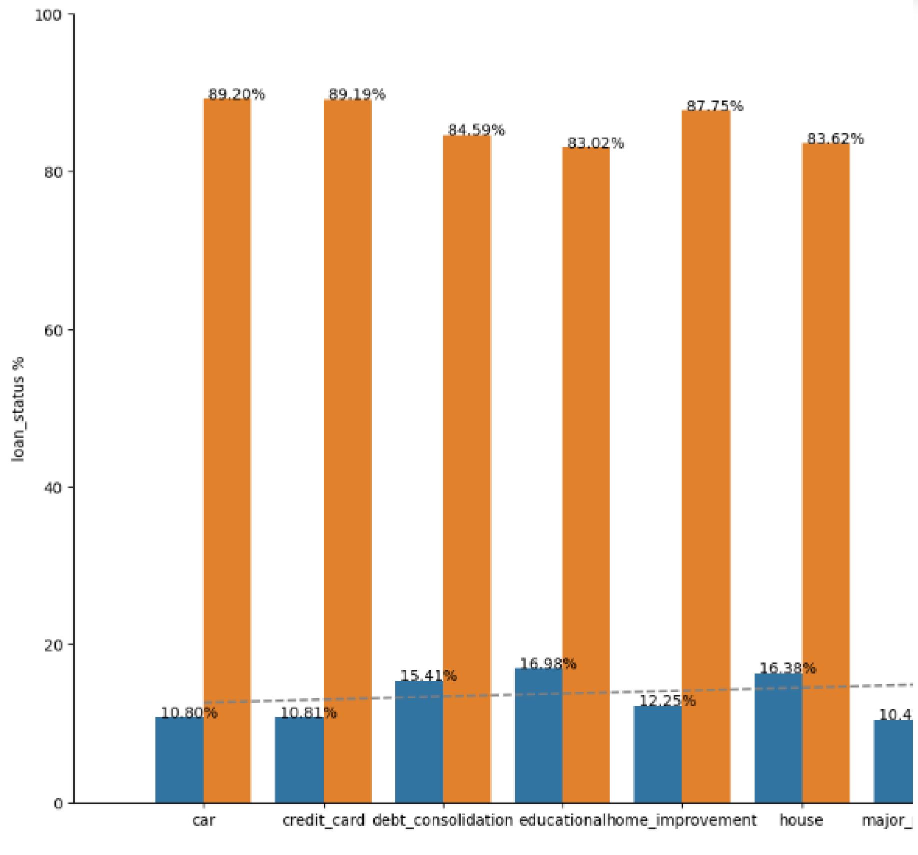
- C, D, E, F, and G grade borrowers exhibit a higher default rate compared to A and B grade borrowers. Despite having annual incomes comparable to those of A and B grade borrowers, these higher-grade borrowers were extended larger loan amounts at interest rates above 12%. In contrast, A and B grade borrowers received loans with interest rates not exceeding 12% for most part. Default rate steadily increases with interest rate.

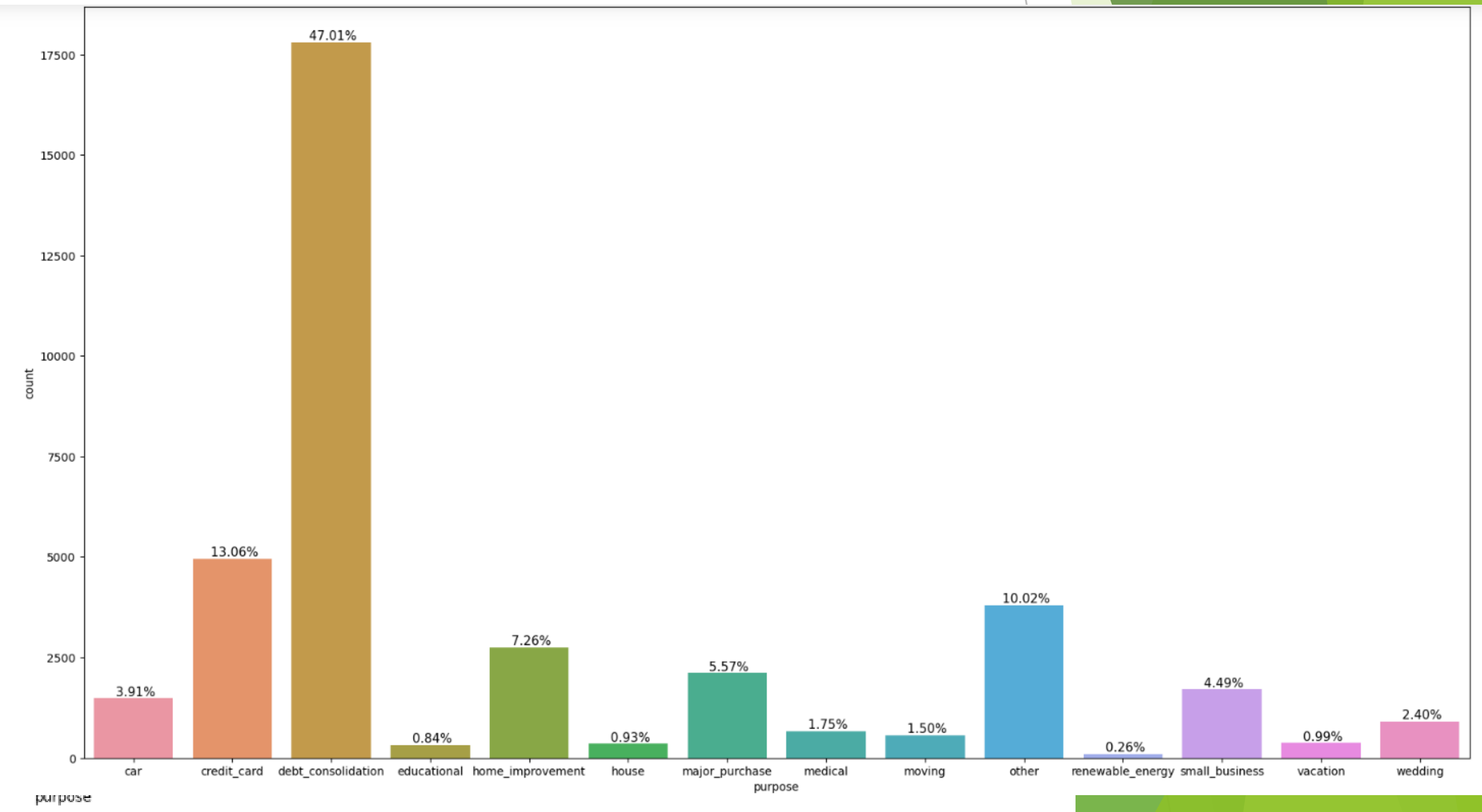




Conclusions

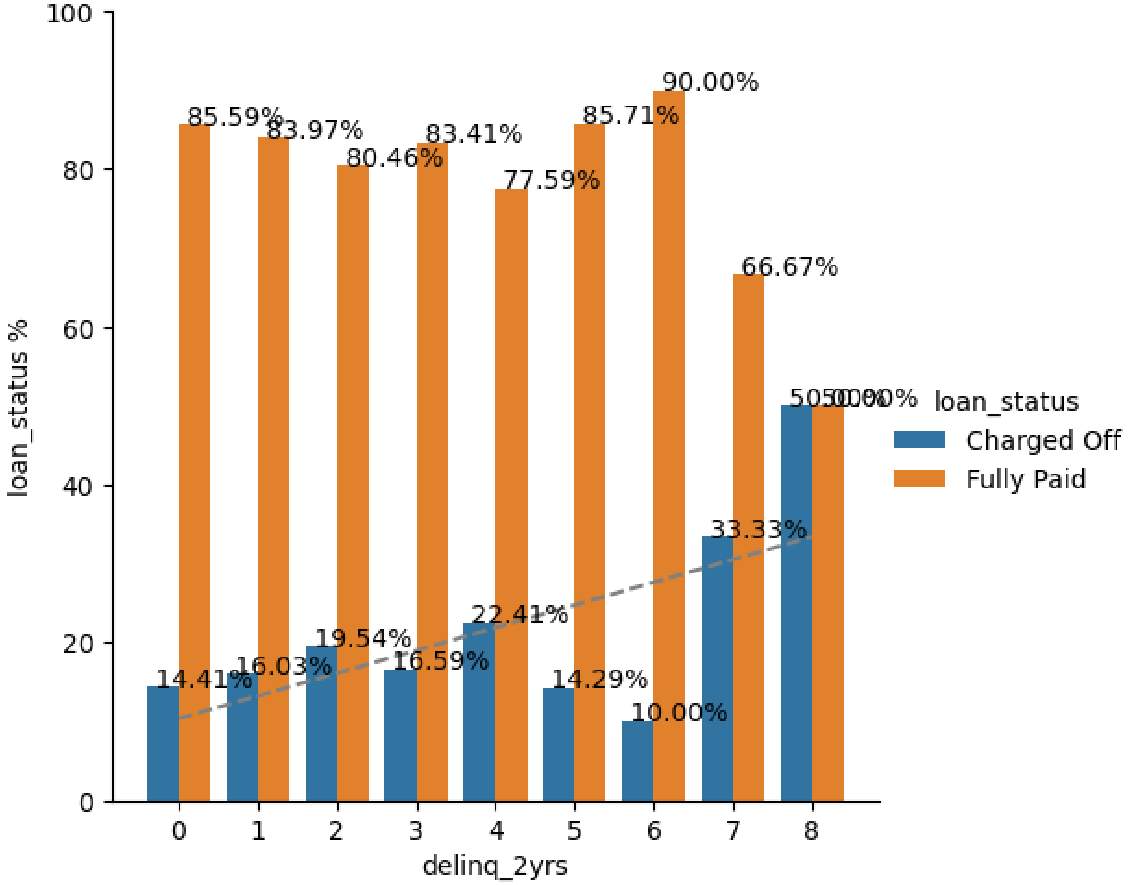
- Debt consolidation loans amount to 47% of all loans given and have a significant default rate at 15.41%.

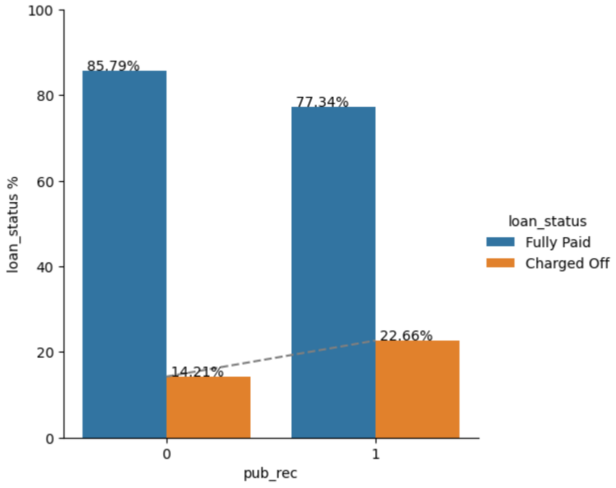


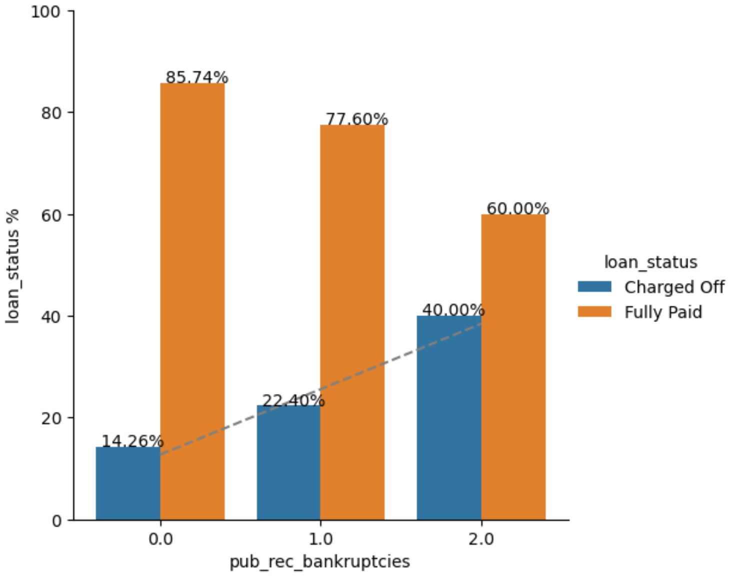


Conclusions

- The default rate tends to rise with borrowers having 30+ days past due incidences (delinquencies), derogatory public records or public bankruptcy records.

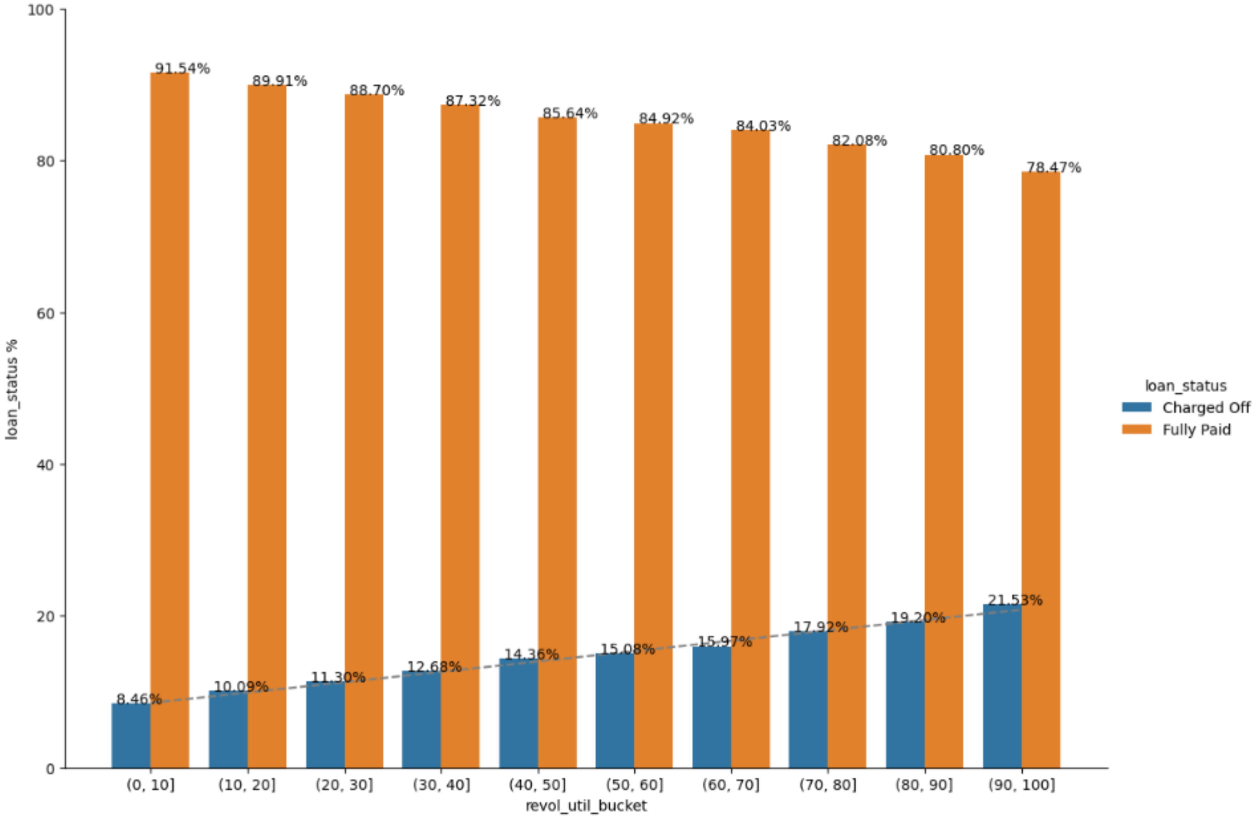


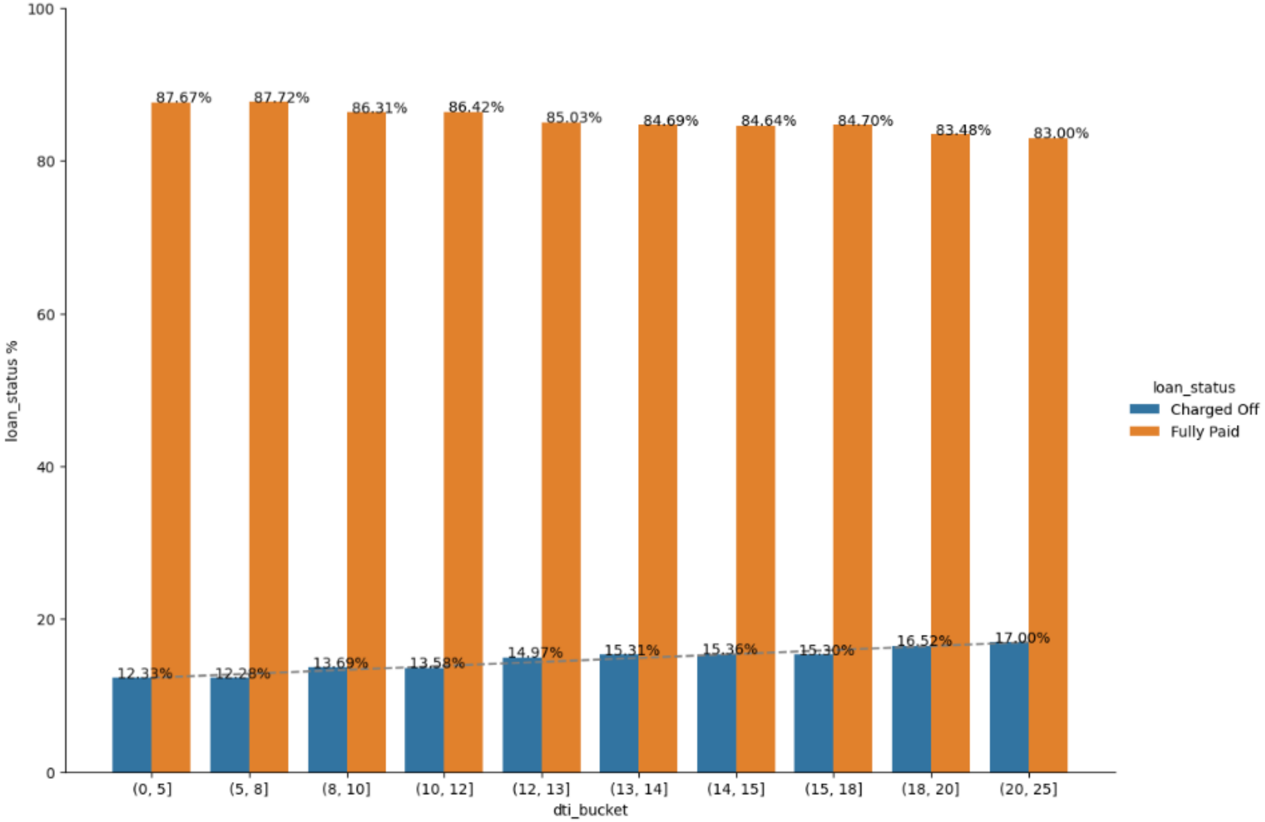




Conclusions

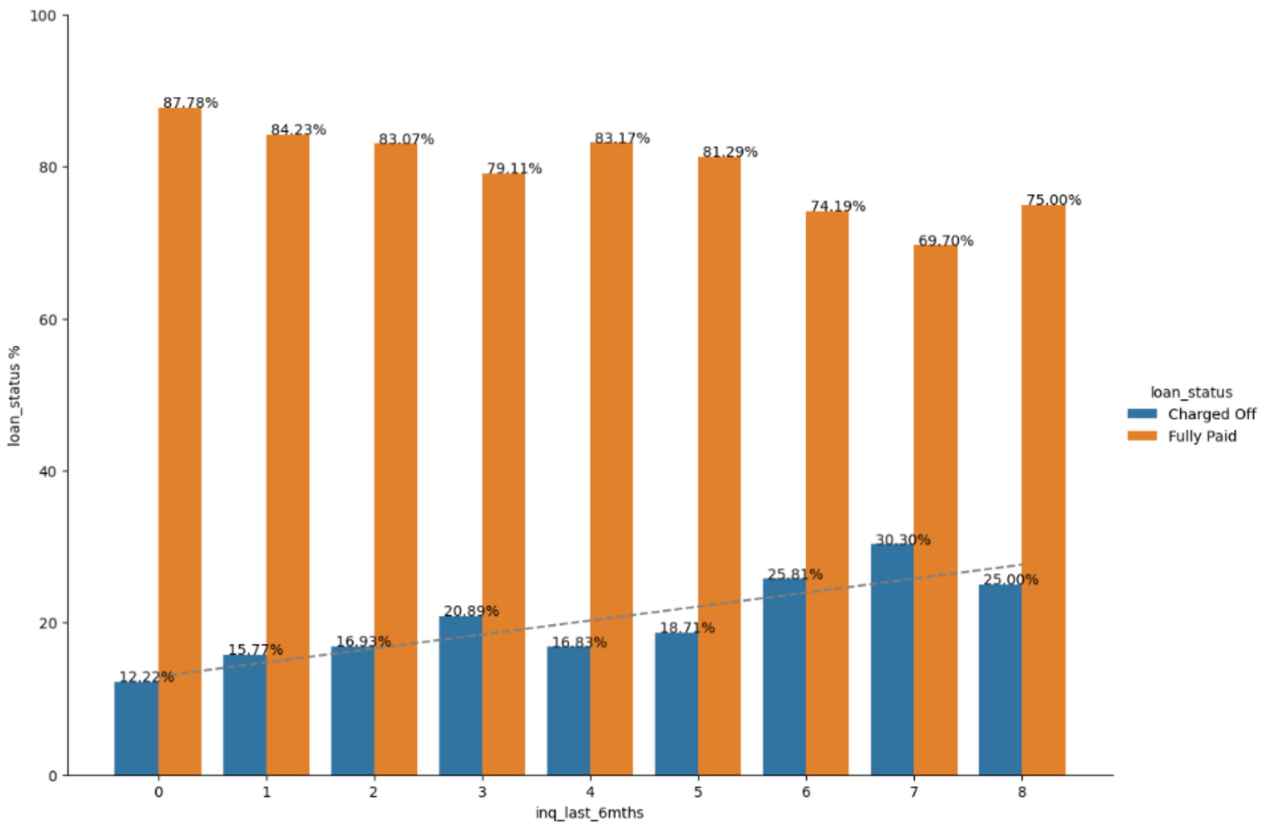
- Default rate increases with revolving utilization rate. Above 30% utilization rate the chances of default increase significantly.
- Default rate increases with DTI, anything above 8 has significantly higher default rate.

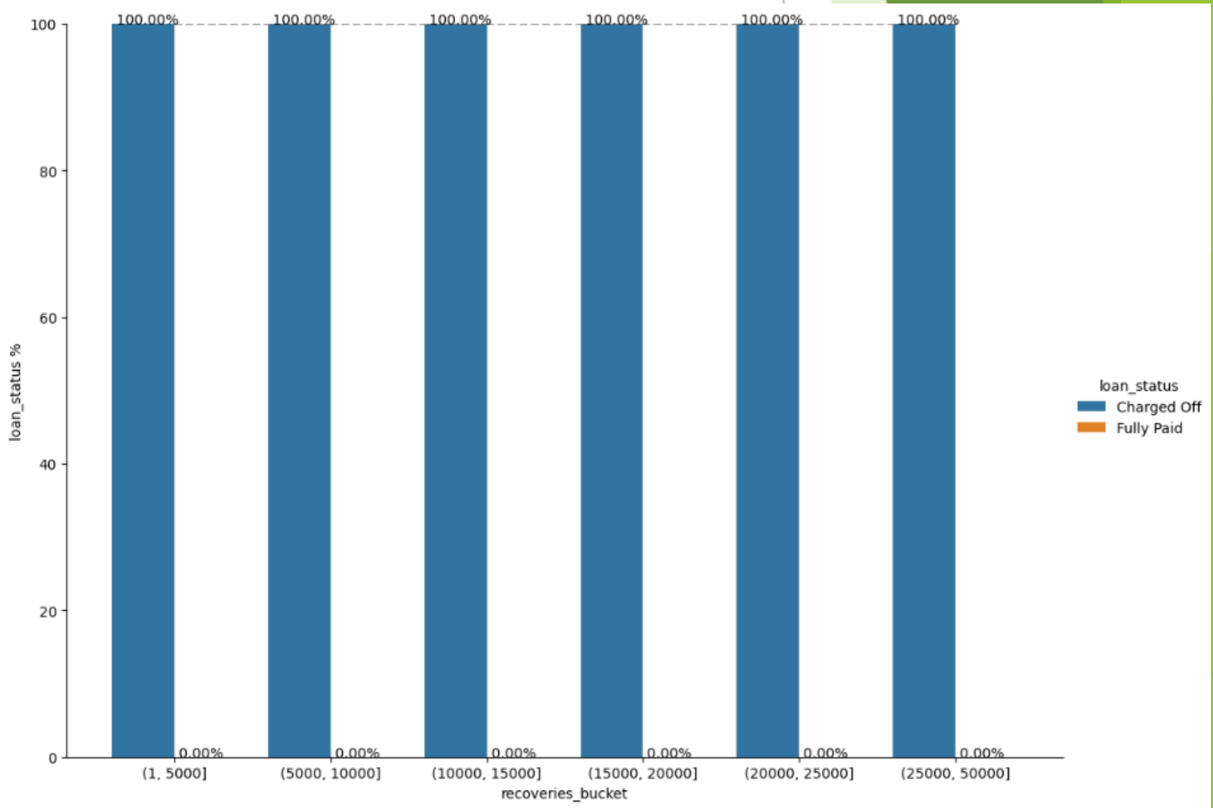




Conclusions

- Default rate increases with number of inquiries in last 6 months in general.
- Recovery fees paid by borrower is a lead indicator of loan default.





Recommendations

- The primary objective of this analysis is
  - Not to attack the main revenue generator for the company
  - Keep the recommendations few so company can implement them

Recommendations

- To mitigate default risk, aim to keep the interest rate below 12%.
- The company should contemplate diversifying loan purposes by decreasing the overall percentage of loans earmarked for debt consolidation. Instead, it should increase loan allocations toward categories such as car purchases, credit card payments, home improvements, weddings, and major purchases. These categories were found to have lower default rates, thereby are good options to mitigate default risk.
- Borrowers falling into the following categories are at a heightened risk of default:
  - Individuals with a history of 30+ days past due incidences (delinquencies), derogatory public records, or public bankruptcy records.
  - Those with a loan revolving utilization rate exceeding 30%.
  - Borrowers with a debt-to-income ratio (DTI) surpassing 8.
  - Individuals who have made more than one loan inquiry in the last 6 months.
  - Borrowers with a record of paying recovery fees.

For future loans extended to borrowers meeting the aforementioned conditions, it is advisable to mitigate loan default risk by considering strategies such as offering smaller loan amounts and/or higher interest rates. These measures can help offset the increased default risk associated with these borrower profiles.

Thank you