Lending Club Case Study

Submitted by-

Rajdip Paul

Rajesh K

Background

Lending Club stands as the most extensive peer-to-peer platform linking borrowers and lenders. Prospective borrowers submit their applications via an online portal, where an internal rating is designated to them. Lenders then make two pivotal decisions: 1) whether to extend a loan, and 2) the specifics of the loan arrangement, including elements like the interest rate, monthly payments, and loan duration.

Business Objective

To find out drivers of default and potentially apply these findings in the process of making decisions regarding approval or rejection.

Understanding Data

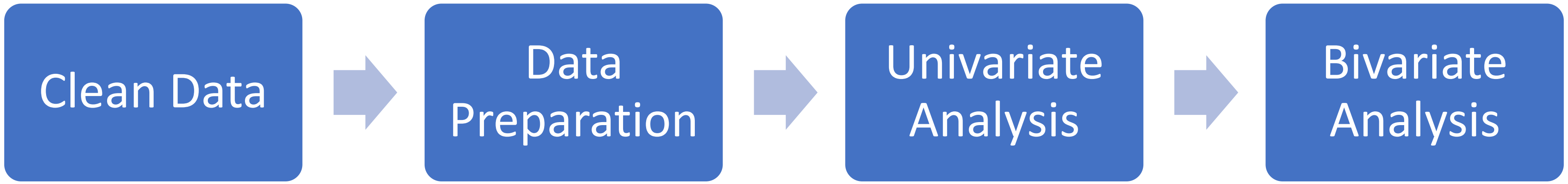
Types of Data:

- Consumer Profiles
    - Employment tenure
    - Employment Title
    - Annual Income
    - Zip code
    - Description

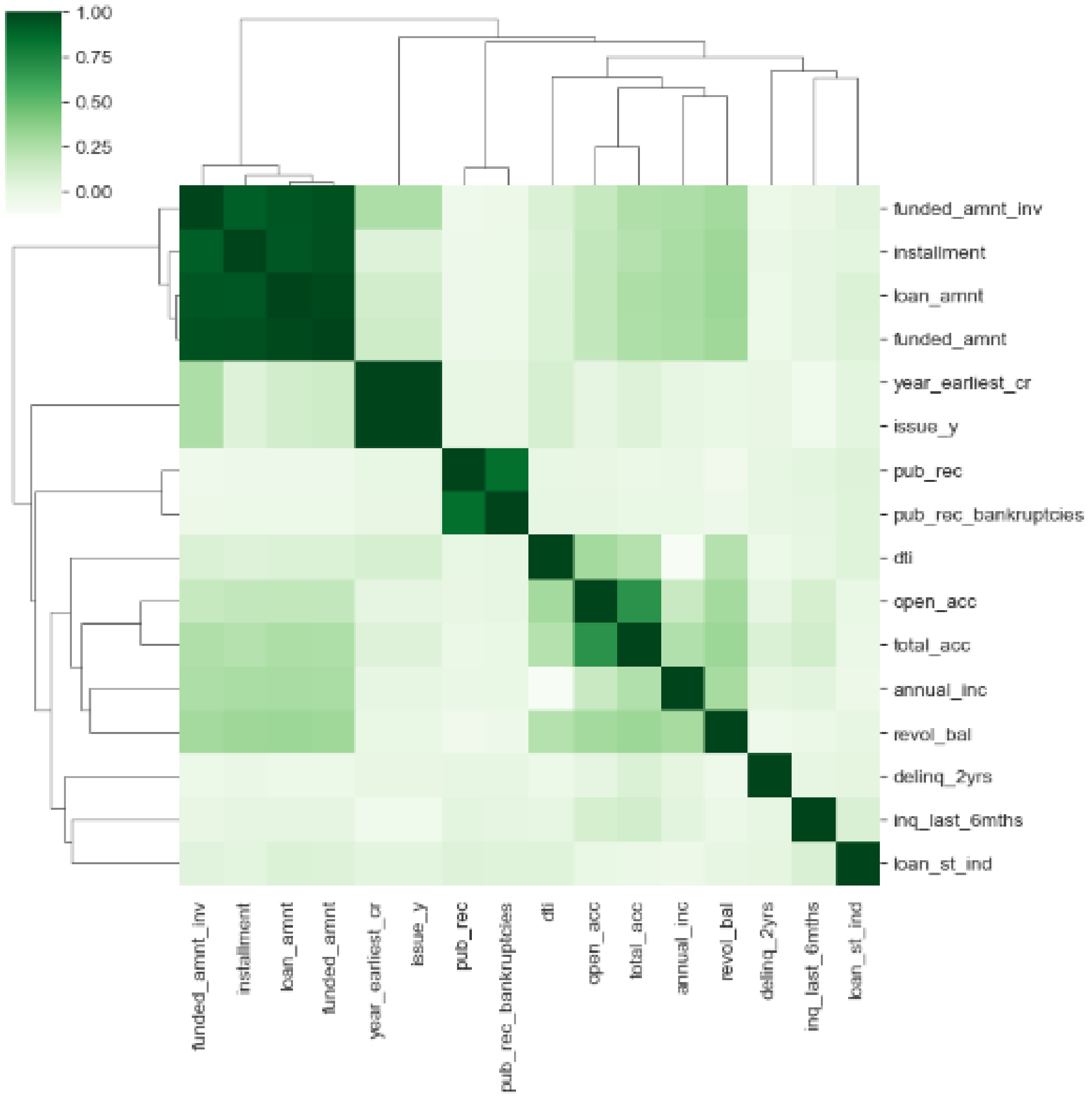
- Information about Loans
    - Loan Amount
    - Funded Amount
    - Interest Rate
    - Loan Status
    - Loan Grade

- Customer actions
    - Delinquency Year -2
    - Earliest Credit Line
    - Revolving Balance
    - Recoveries
    - Application Type
    - Loan Purpose

Analysis approach

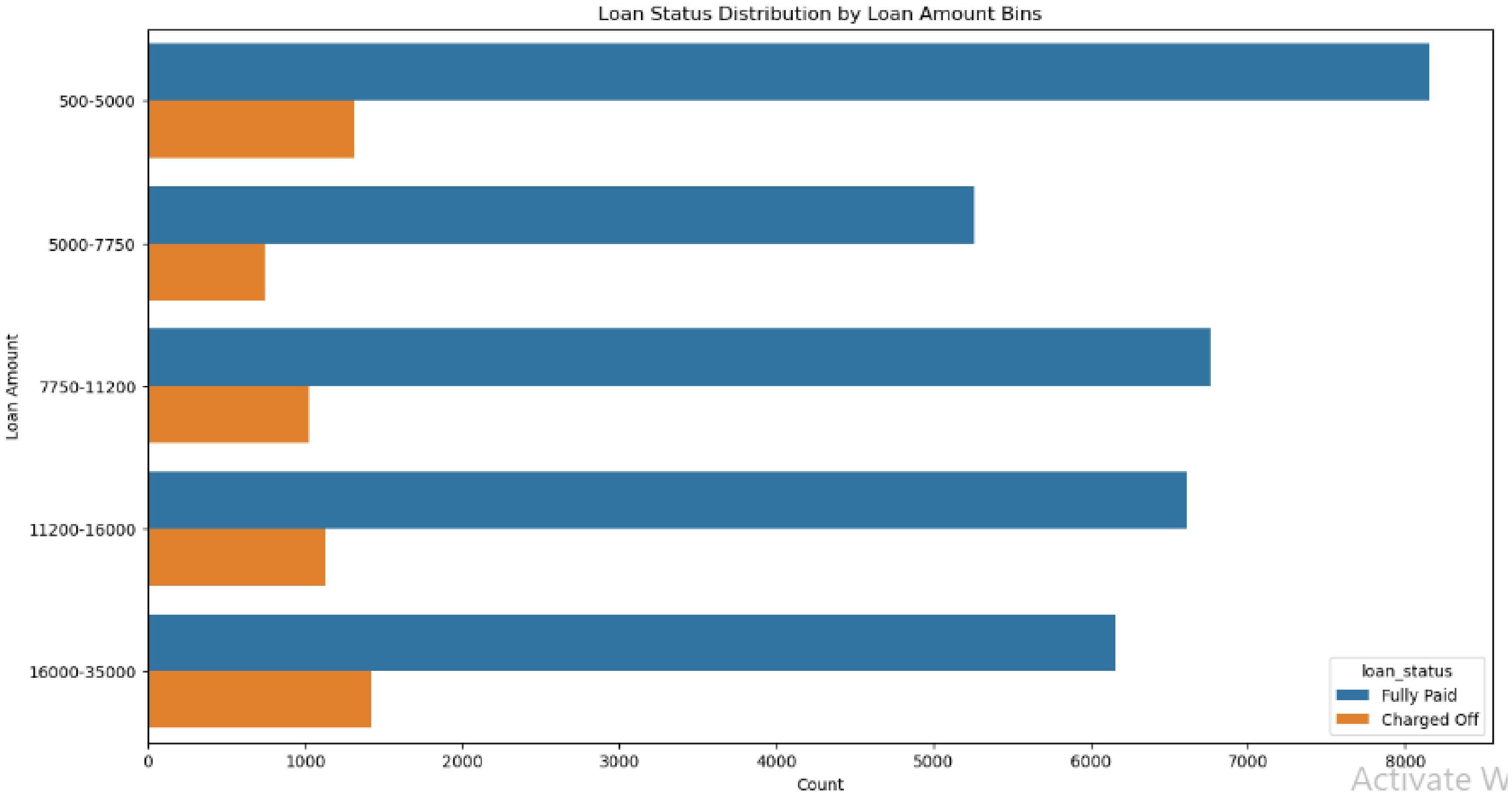


Let us check the correlation between the various numeric fields to see the relationships

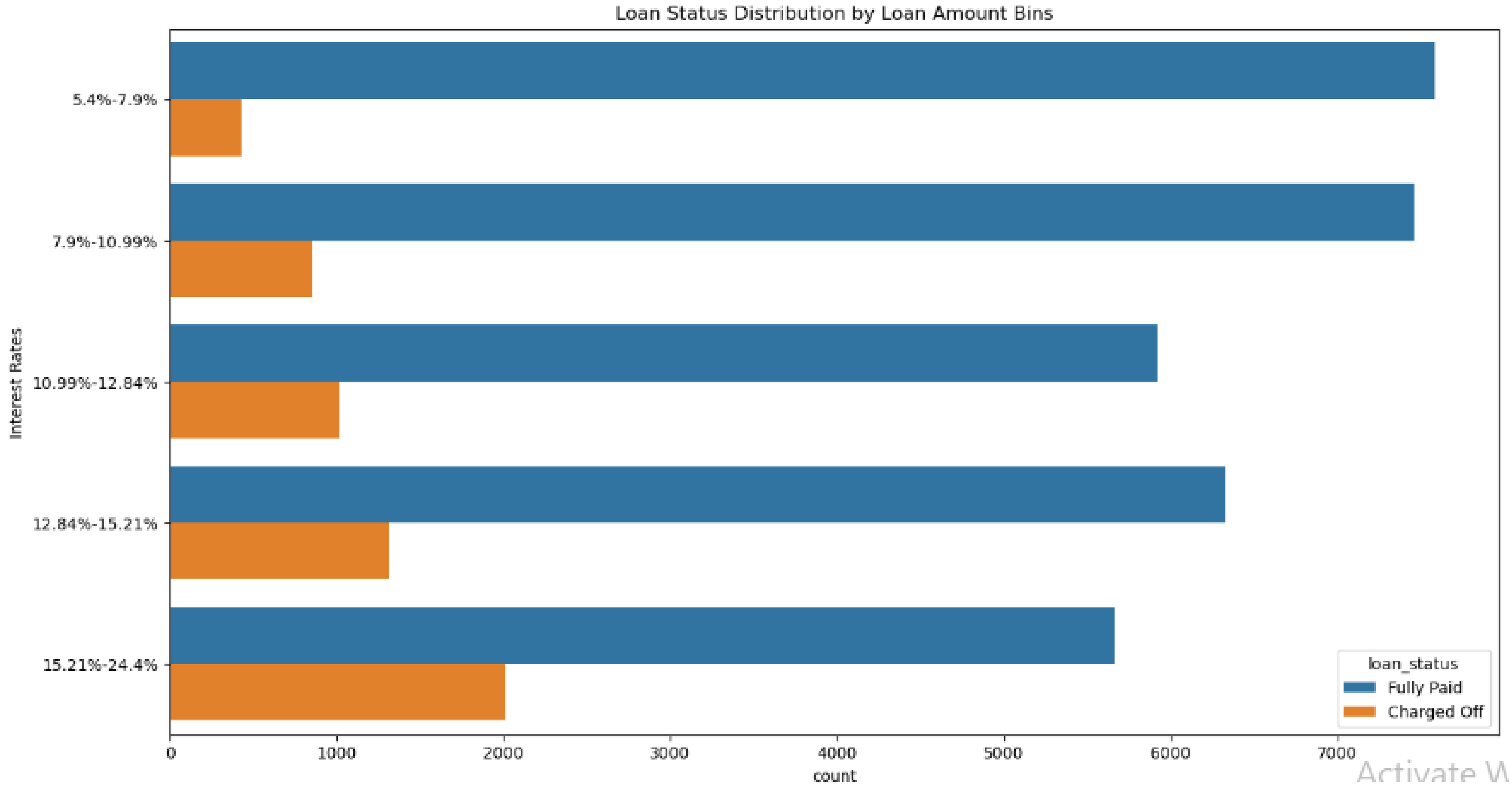


From the cluster map we found out that funded_amnt_inv, installment, loan_amnt, funded_amnt have big correlation.

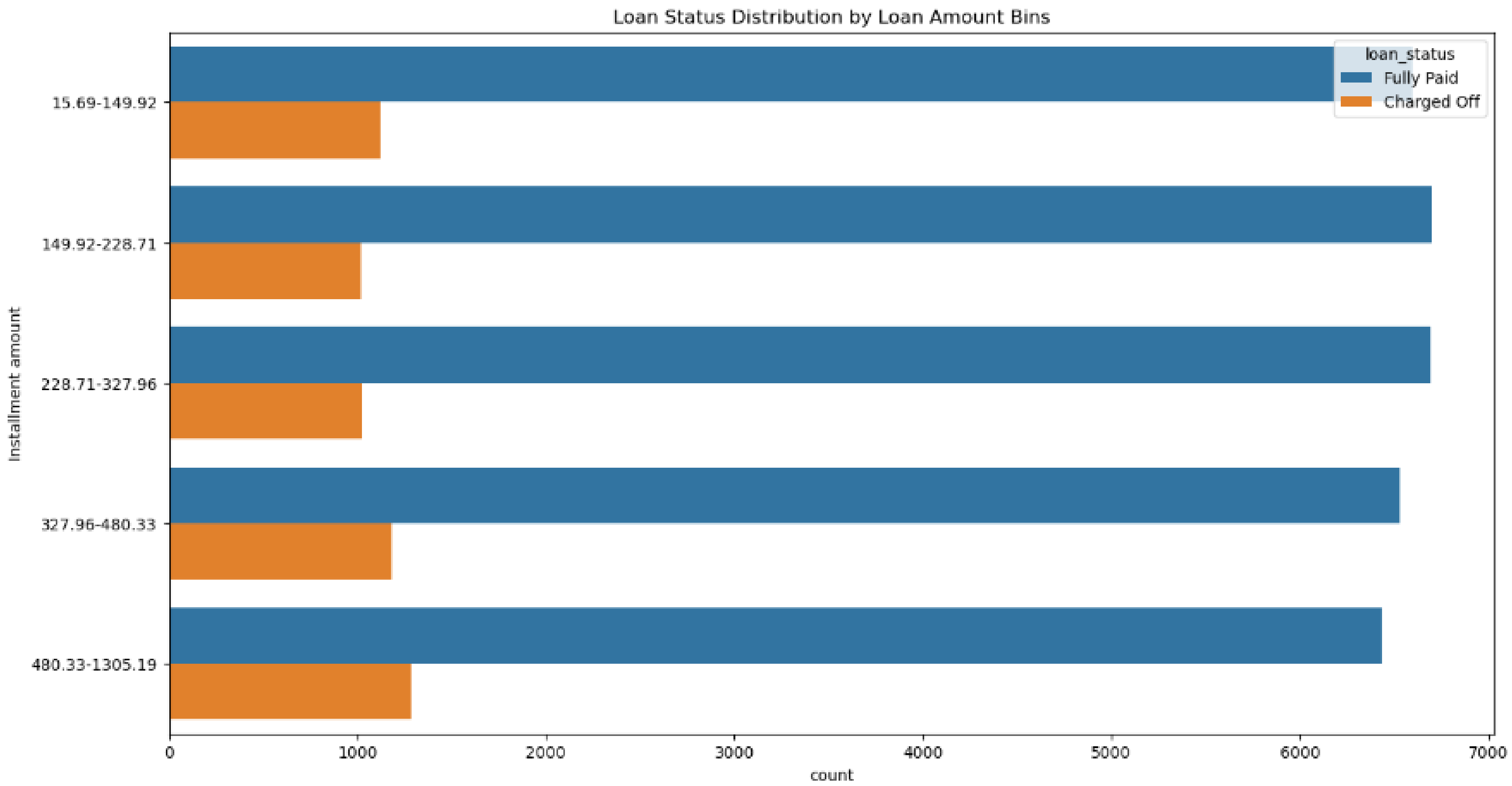
Comparing loan_status column with other numerical conitnuous values



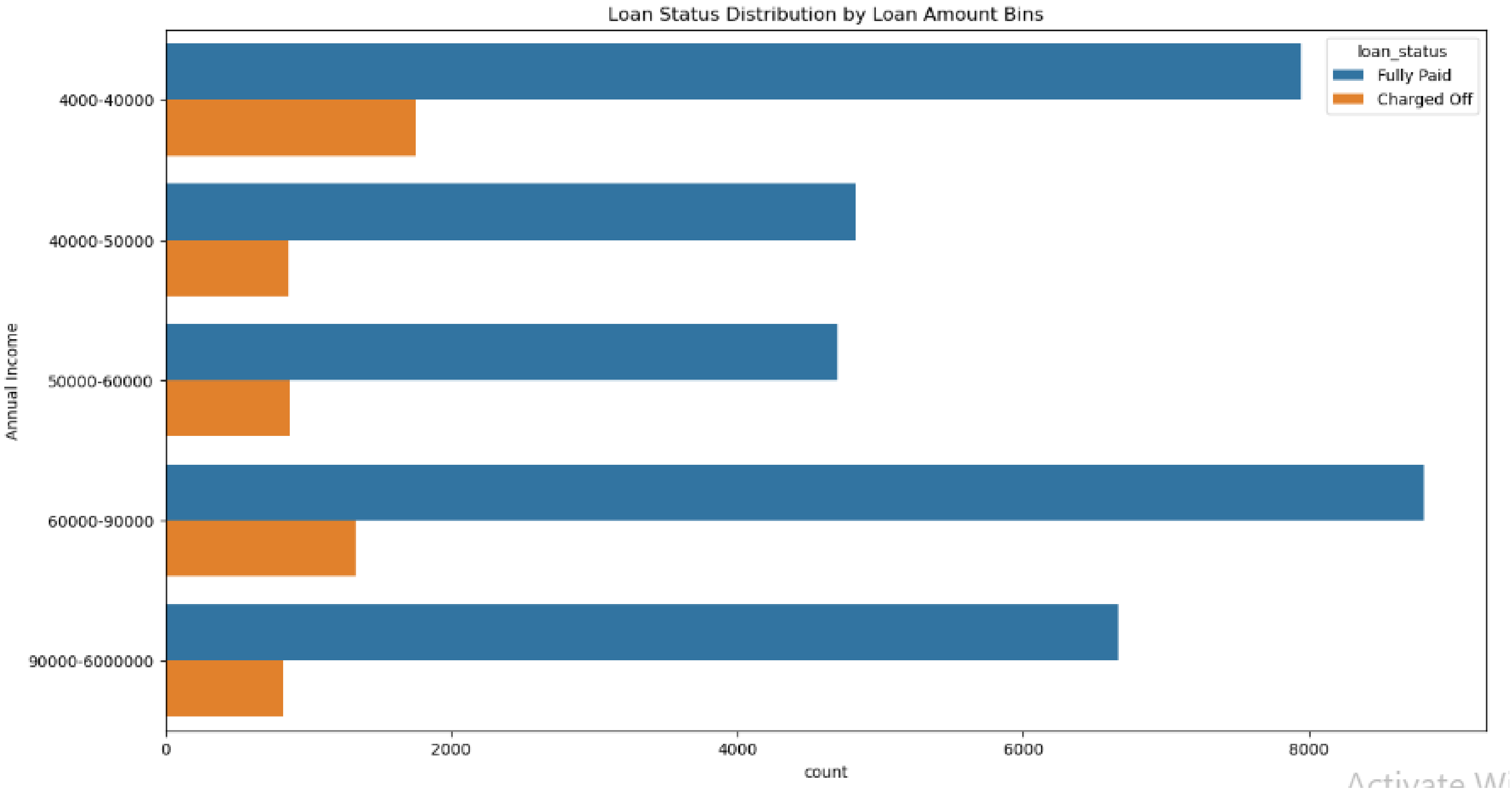
We can see higher loan amounts lead to more charge offs



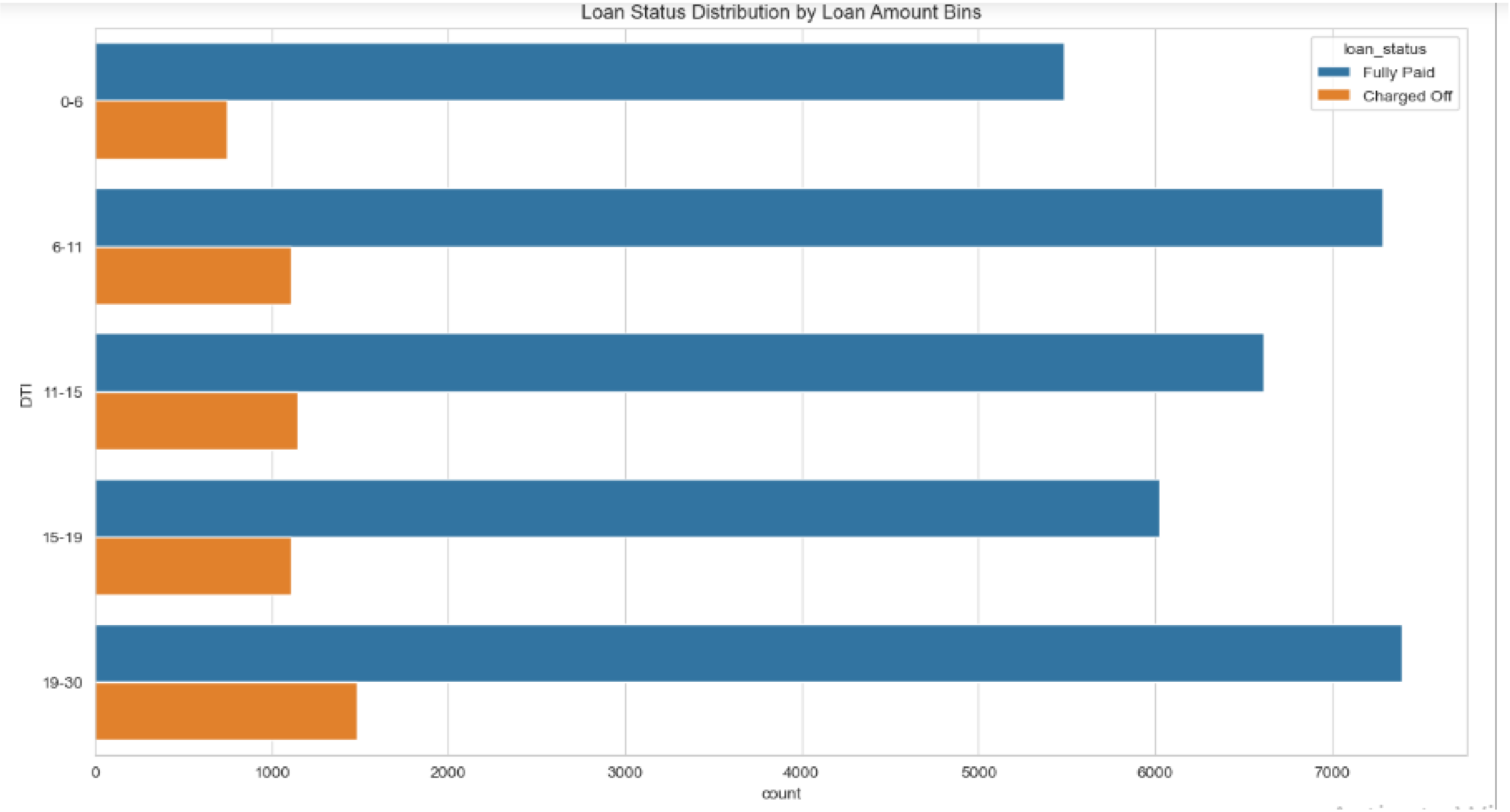
We can see that higher interest rates also contribute to charge offs shooting up.



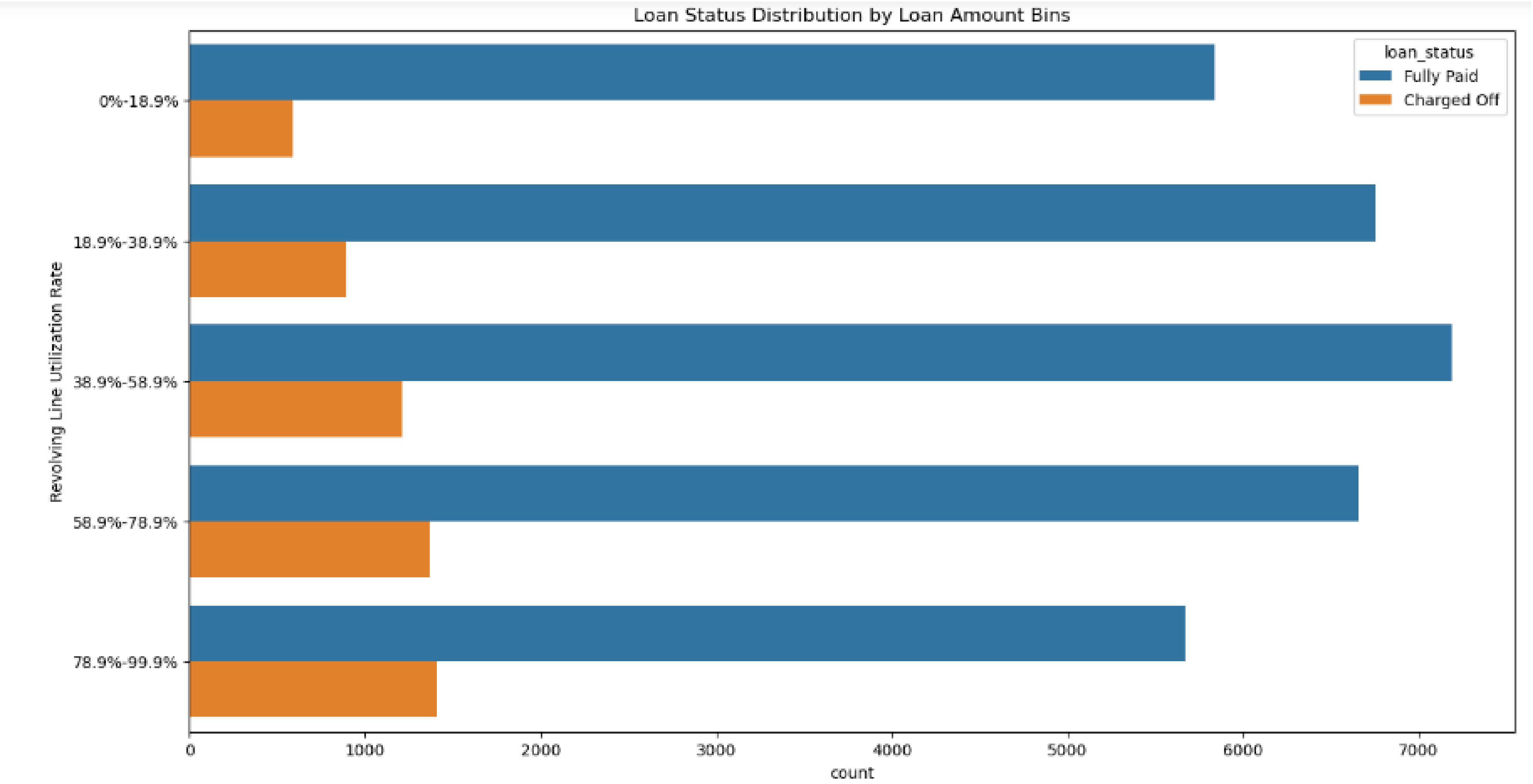
From above plot we can see that higher interest amounts leads to higher defaults.



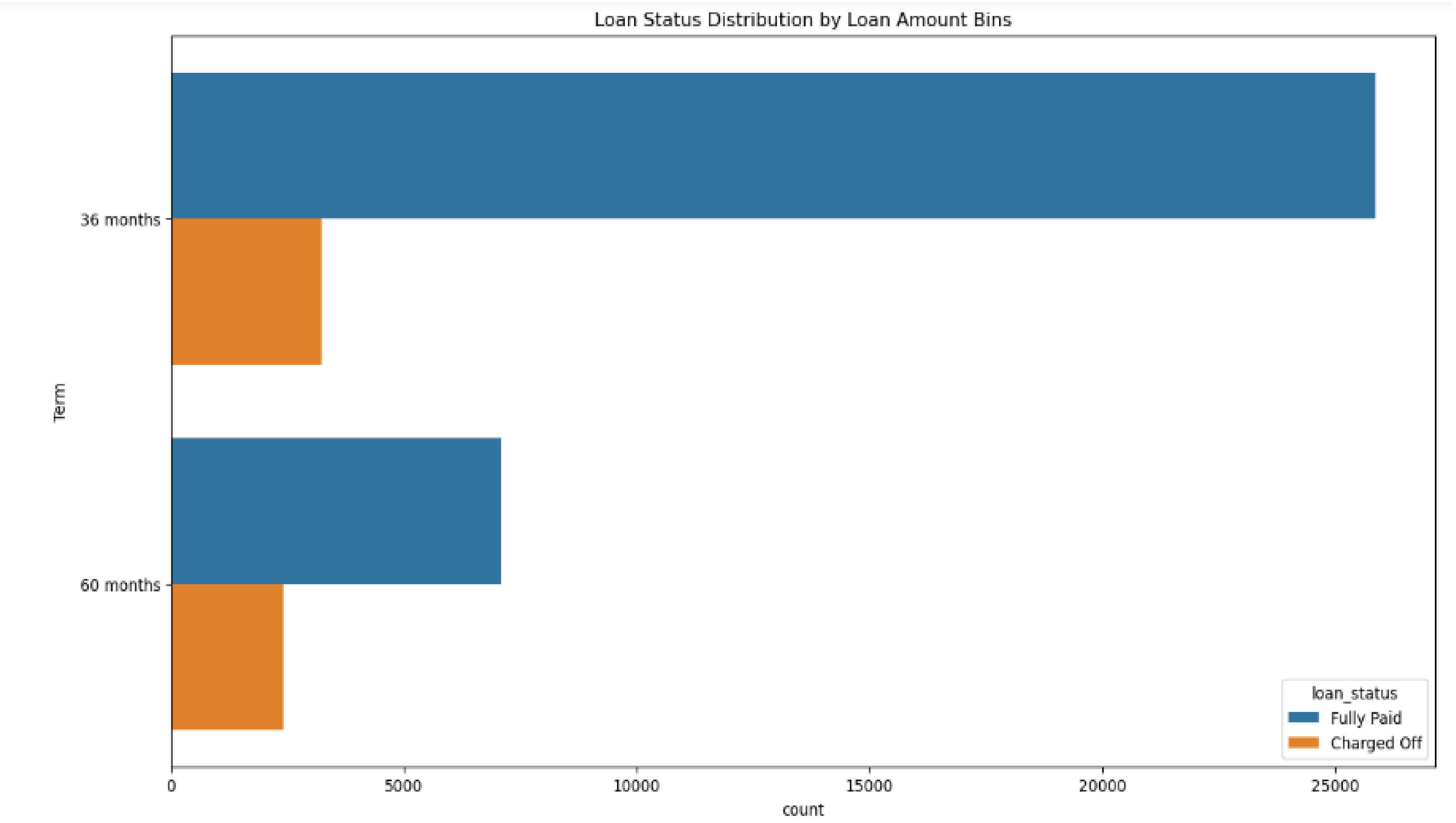
We can see that lower income has higher default chances, higher income has least default chances



Higher DTI leads more defaults

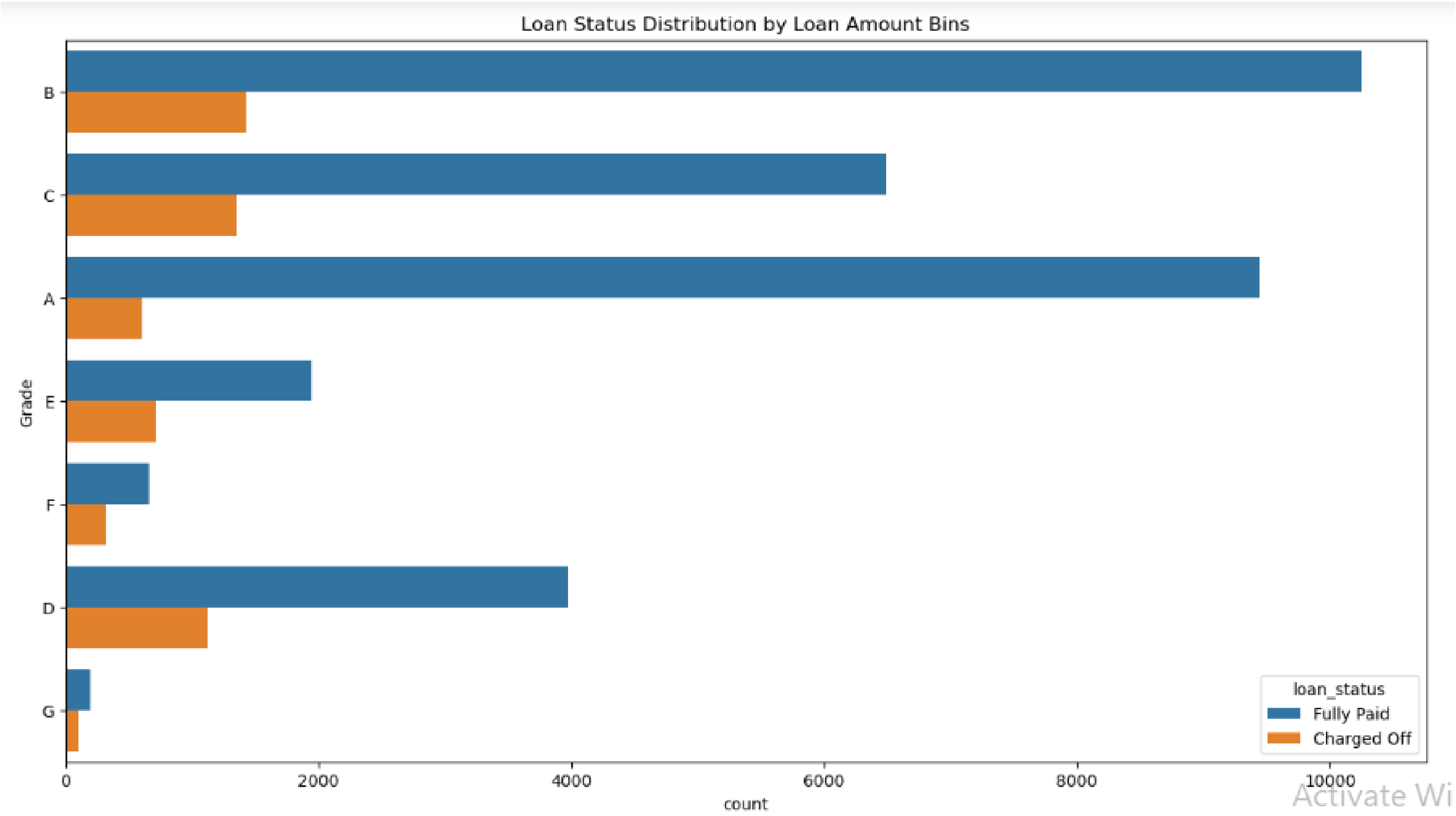


Higher Revolving line utilization rate indicates higher default chances.



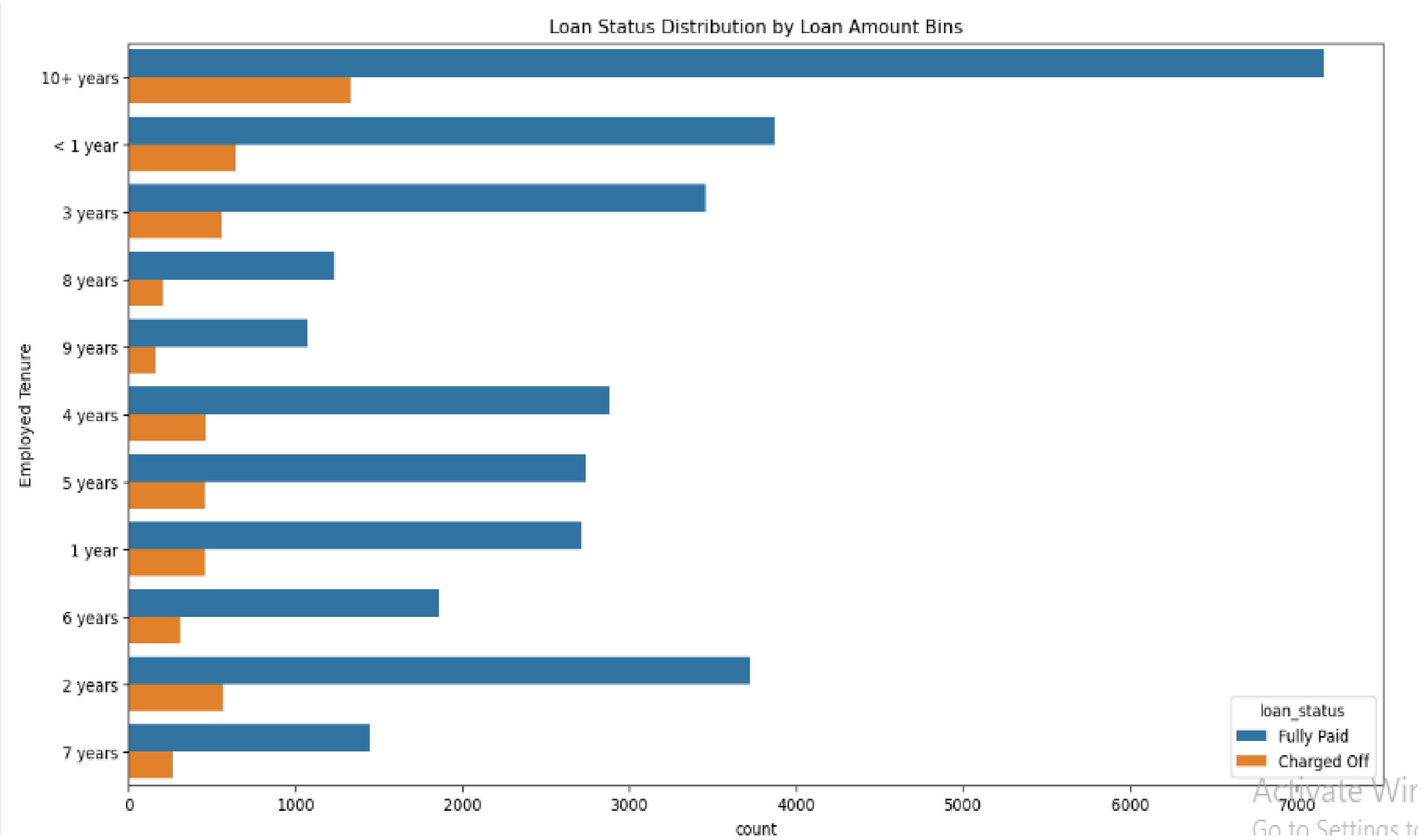

```
For 36 months term:
Fully Paid     0.89
Charged Off    0.11
Name: loan_status, dtype: float64
For 60 months term:
Fully Paid     0.75
Charged Off    0.25
Name: loan_status, dtype: float64
```

We can see that for 36 months the charge of rate is 11%, while for 60 months it is 25%. Hence there is higher chance of efaulting loans for higher loan repayment time.



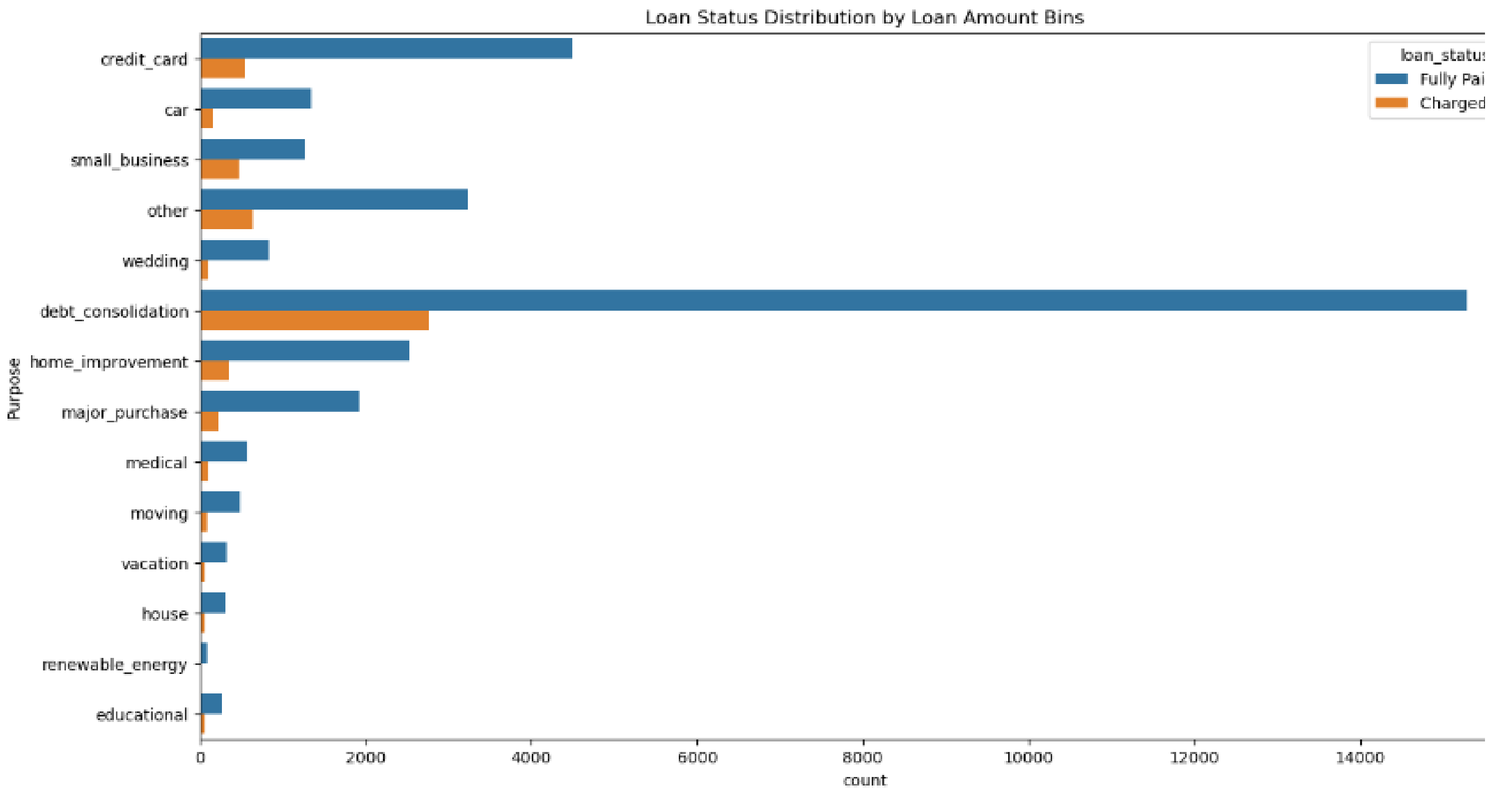

```
Grade: A
Fully Paid     0.94
Charged Off    0.06
Name: loan_status, dtype: float64

Grade: B
Fully Paid     0.88
Charged Off    0.12
Name: loan_status, dtype: float64

Grade: C
Fully Paid     0.83
Charged Off    0.17
Name: loan_status, dtype: float64

Grade: D
Fully Paid     0.78
Charged Off    0.22
Name: loan_status, dtype: float64

Grade: E
Fully Paid     0.73
Charged Off    0.27
Name: loan_status, dtype: float64
Grade: F
Fully Paid     0.67
Charged Off    0.33
Name: loan_status, dtype: float64

Grade: G
Fully Paid     0.66
Charged Off    0.34
Name: loan_status, dtype: float64
```

We can see that for lower grades the charge off percent shoots up




```
Employee length: < 1 year
Fully Paid     0.858252
Charged Off    0.141748
Name: loan_status, dtype: float64

Employee length: 1 year
Fully Paid     0.856106
Charged Off    0.143894
Name: loan_status, dtype: float64

Employee length: 2 years
Fully Paid     0.867863
Charged Off    0.132137
Name: loan_status, dtype: float64

Employee length: 3 years
Fully Paid     0.861665
Charged Off    0.138335
Name: loan_status, dtype: float64

Employee length: 4 years
Fully Paid     0.861759
Charged Off    0.138241
Name: loan_status, dtype: float64

Employee length: 5 years
Fully Paid     0.856606
Charged Off    0.143394
Name: loan_status, dtype: float64

Employee length: 6 years
Fully Paid     0.858395
Charged Off    0.141605
Name: loan_status, dtype: float64

Employee length: 7 years
Fully Paid     0.846289
Charged Off    0.153711
Name: loan_status, dtype: float64

Employee length: 8 years
Fully Paid     0.858537
Charged Off    0.141463
Name: loan_status, dtype: float64

Employee length: 9 years
Fully Paid     0.871126
Charged Off    0.128874
Name: loan_status, dtype: float64

Employee length: 10+ years
Fully Paid     0.84319
Charged Off    0.15681
Name: loan_status, dtype: float64
```

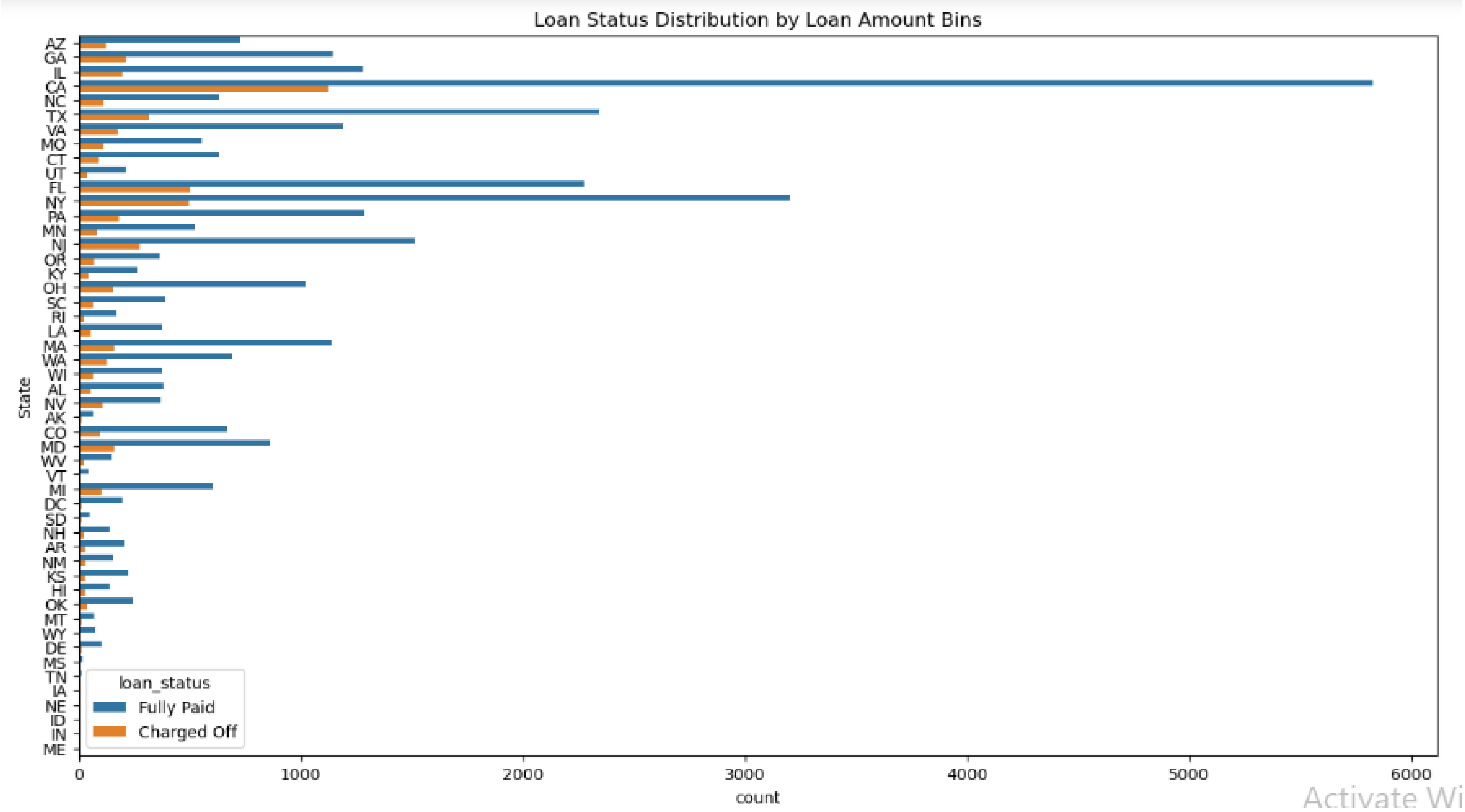
Though there is not much evident data, we can say that people with less than 1 year employment are likely to default more



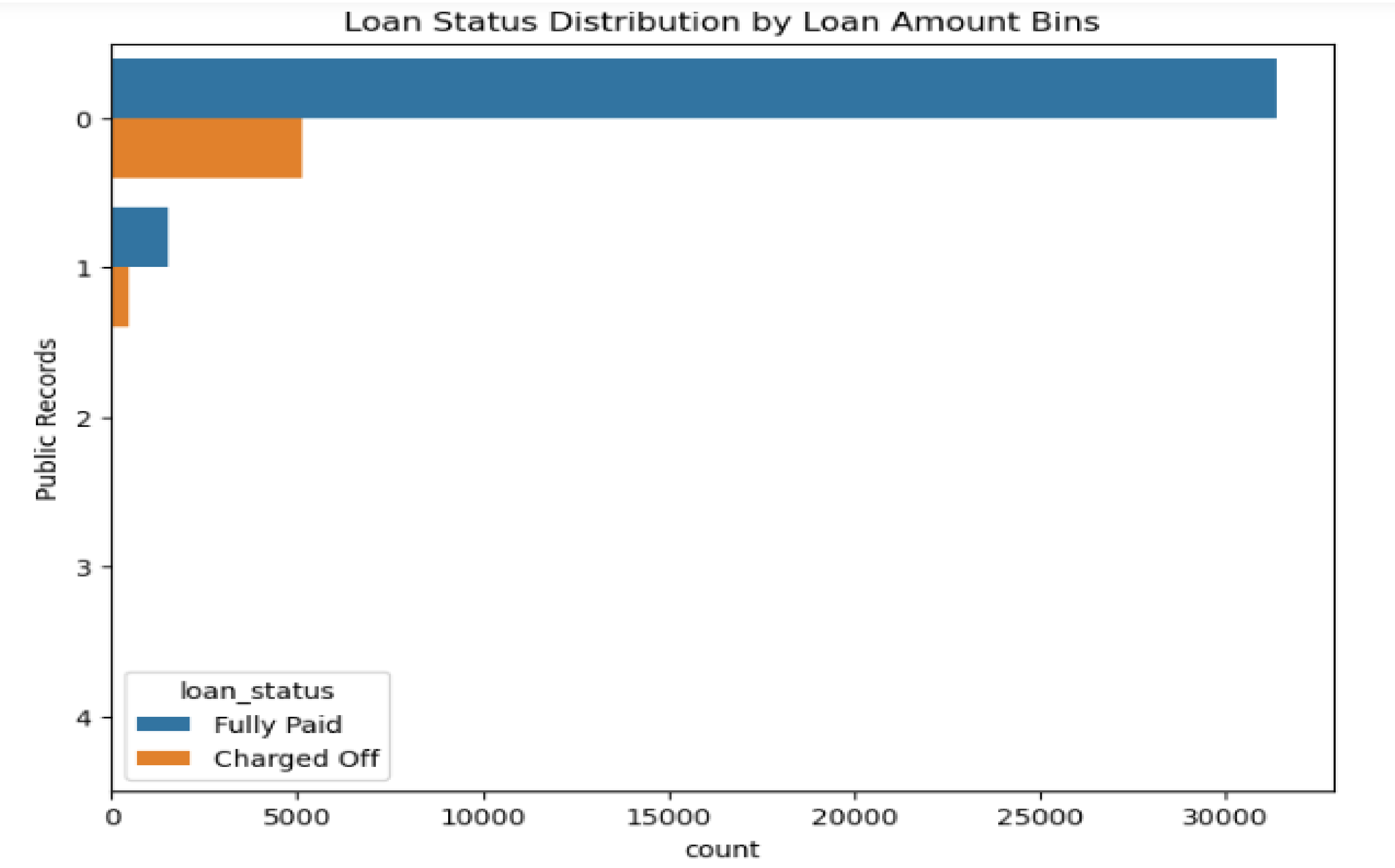

```
Purpose: debt_consolidation
Fully Paid     0.85
Charged Off    0.15
Name: loan_status, dtype: float64

Purpose: credit_card
Fully Paid     0.89
Charged Off    0.11
Name: loan_status, dtype: float64

Purpose: other
Fully Paid     0.84
Charged Off    0.16
Name: loan_status, dtype: float64

Purpose: home_improvement
Fully Paid     0.88
Charged Off    0.12
Name: loan_status, dtype: float64

Purpose: major_purchase
Fully Paid     0.9
Charged Off    0.1
Name: loan_status, dtype: float64

Purpose: small_business
Fully Paid     0.73
Charged Off    0.27
Name: loan_status, dtype: float64
```

```
Purpose: car
Fully Paid     0.89
Charged Off    0.11
Name: loan_status, dtype: float64

Purpose: wedding
Fully Paid     0.9
Charged Off    0.1
Name: loan_status, dtype: float64

Purpose: medical
Fully Paid     0.84
Charged Off    0.16
Name: loan_status, dtype: float64

Purpose: moving
Fully Paid     0.84
Charged Off    0.16
Name: loan_status, dtype: float64

Purpose: vacation
Fully Paid     0.86
Charged Off    0.14
Name: loan_status, dtype: float64
Purpose: house
Fully Paid     0.84
Charged Off    0.16
Name: loan_status, dtype: float64

Purpose: educational
Fully Paid     0.83
Charged Off    0.17
Name: loan_status, dtype: float64

Purpose: renewable_energy
Fully Paid     0.81
Charged Off    0.19
Name: loan_status, dtype: float64
```

Apparently loans granted for small business,renewable energy, educational are riskier.

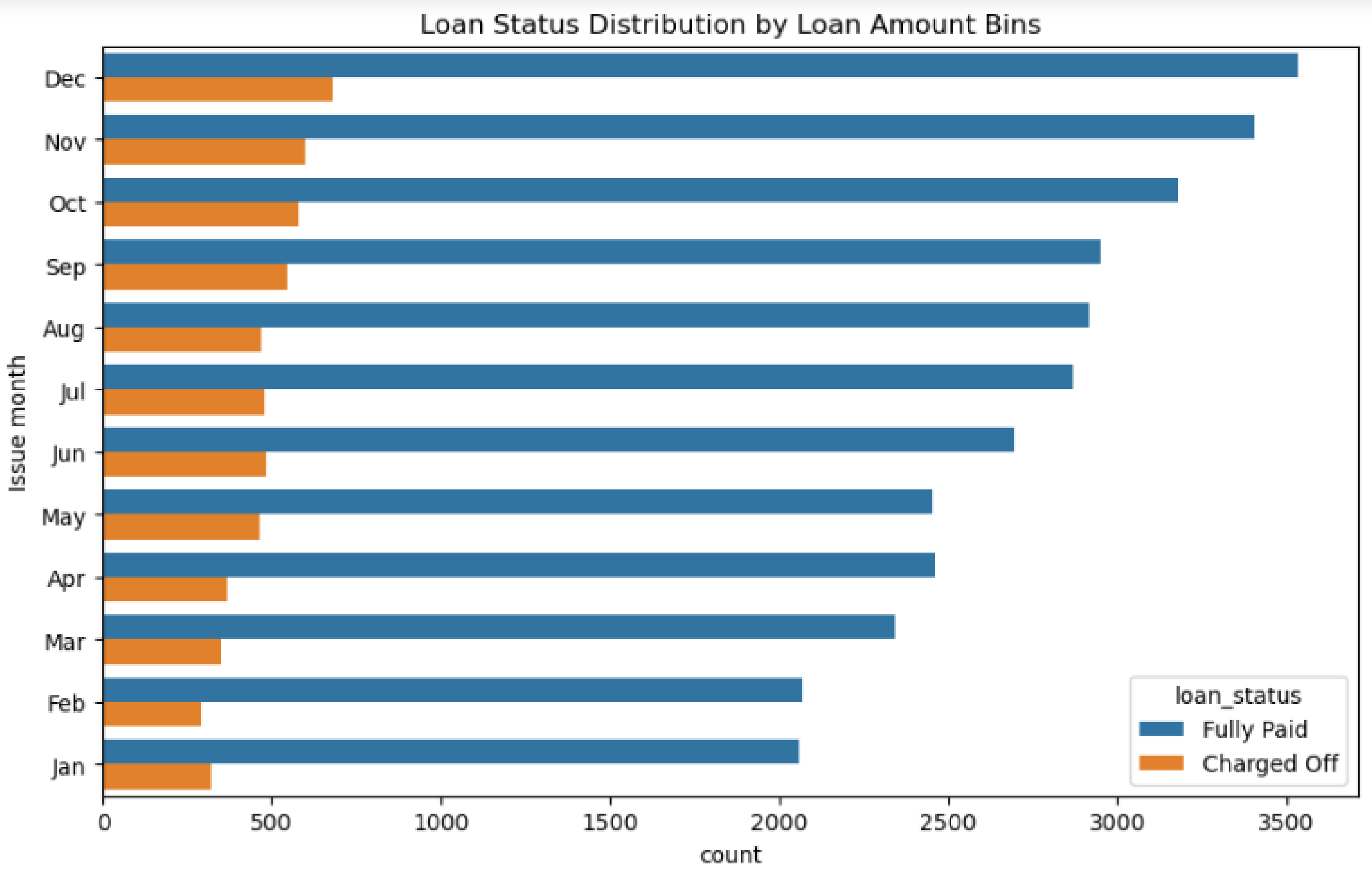


We can find out certain states are riskier.



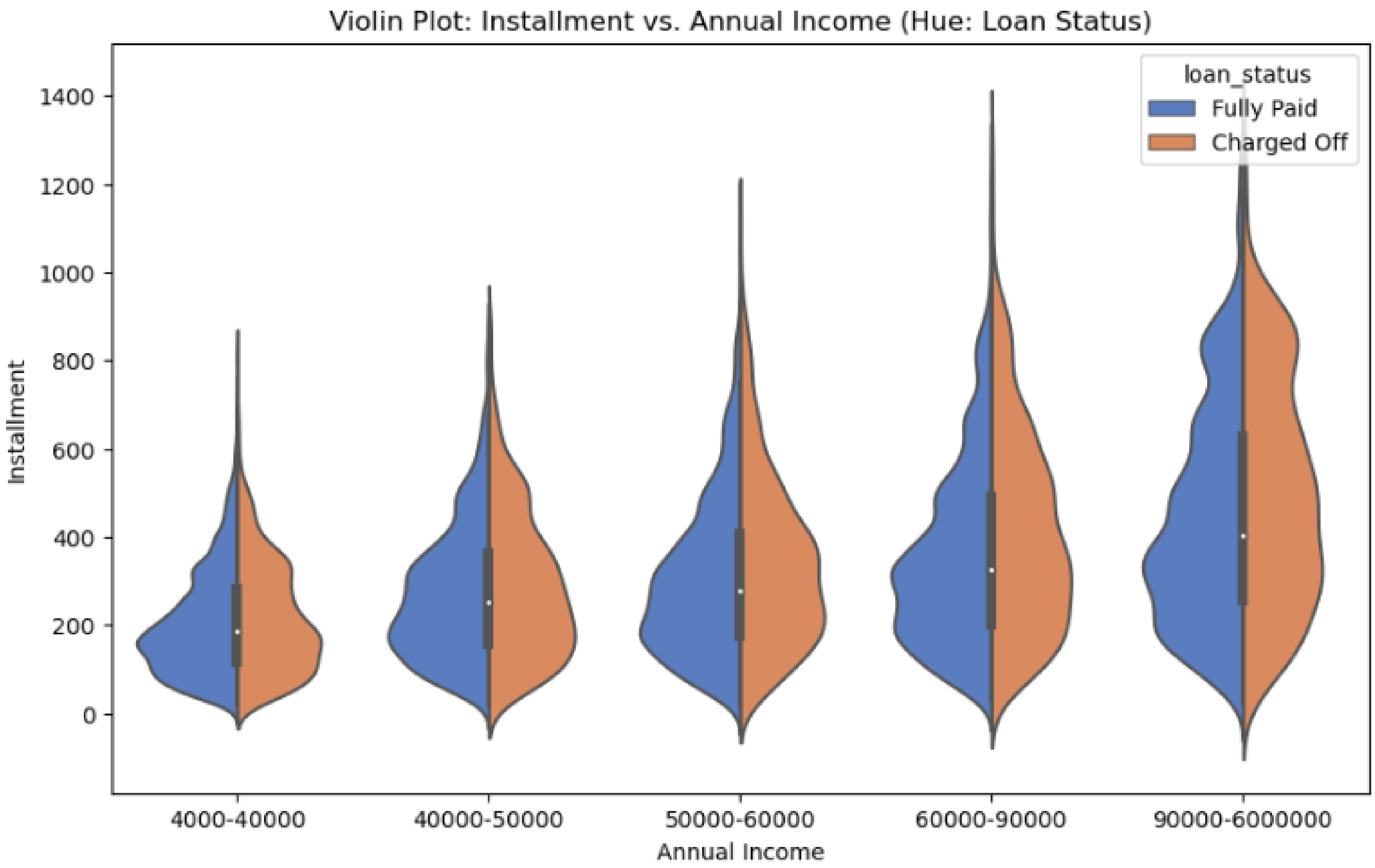

```
Public Records: 0
Fully Paid     0.86
Charged Off    0.14
Name: loan_status, dtype: float64

Public Records: 1
Fully Paid     0.77
Charged Off    0.23
Name: loan_status, dtype: float64

Public Records: 2
Fully Paid     0.79
Charged Off    0.21
Name: loan_status, dtype: float64

Public Records: 3
Fully Paid    1.0
Name: loan_status, dtype: float64

Public Records: 4
Fully Paid    1.0
Name: loan_status, dtype: float64
```

We can see that the applicants with non zero dereogatory record have higher chance of charge off.

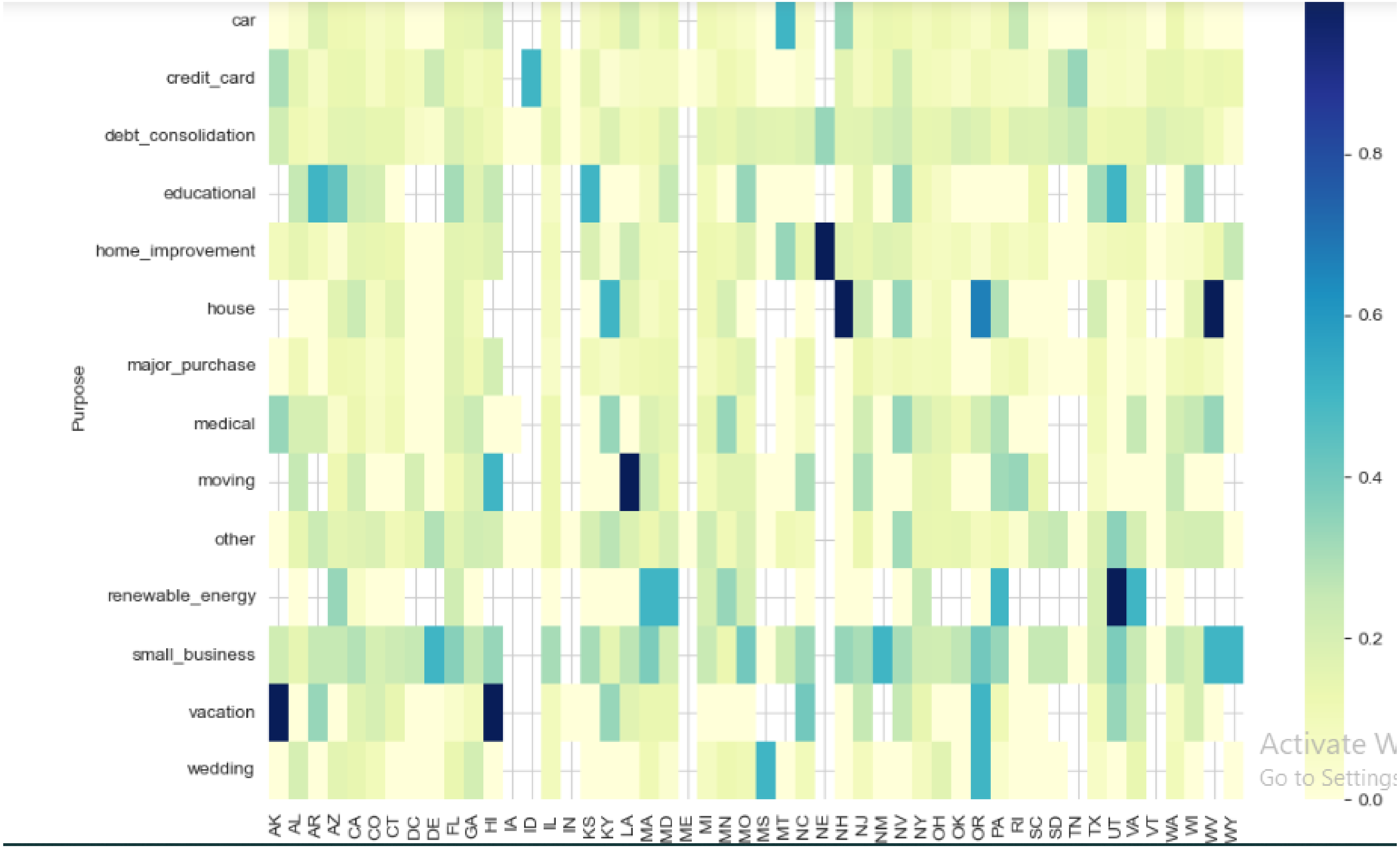



```
Issue month: Jan
Fully Paid      0.865069
Charged Off     0.134931
Name: loan_status, dtype: float64

Issue month: Feb
Fully Paid      0.877014
Charged Off     0.122986
Name: loan_status, dtype: float64

Issue month: Mar
Fully Paid      0.871052
Charged Off     0.128948
Name: loan_status, dtype: float64

Issue month: Apr
Fully Paid      0.869304
Charged Off     0.130696
Name: loan_status, dtype: float64

Issue month: May
Fully Paid      0.840356
Charged Off     0.159644
Name: loan_status, dtype: float64

Issue month: Jun
Fully Paid      0.848113
Charged Off     0.151887
Name: loan_status, dtype: float64

Issue month: Jul
Fully Paid      0.857058
Charged Off     0.142942
Name: loan_status, dtype: float64

Issue month: Aug
Fully Paid      0.861865
Charged Off     0.138135
Name: loan_status, dtype: float64

Issue month: Sep
Fully Paid      0.843625
Charged Off     0.156375
Name: loan_status, dtype: float64

Issue month: Oct
Fully Paid      0.845786
Charged Off     0.154214
Name: loan_status, dtype: float64

Issue month: Nov
Fully Paid      0.850724
Charged Off     0.149276
Name: loan_status, dtype: float64

Issue month: Dec
Fully Paid      0.839146
Charged Off     0.160854
Name: loan_status, dtype: float64
```

We can see Dec has the highes number of loan applications and also highest charge off percentage. May,Sep,Jun are other months where max number of defaults are seen. Maybe holiday season is the reason people apply for loans and then fail to pay back.



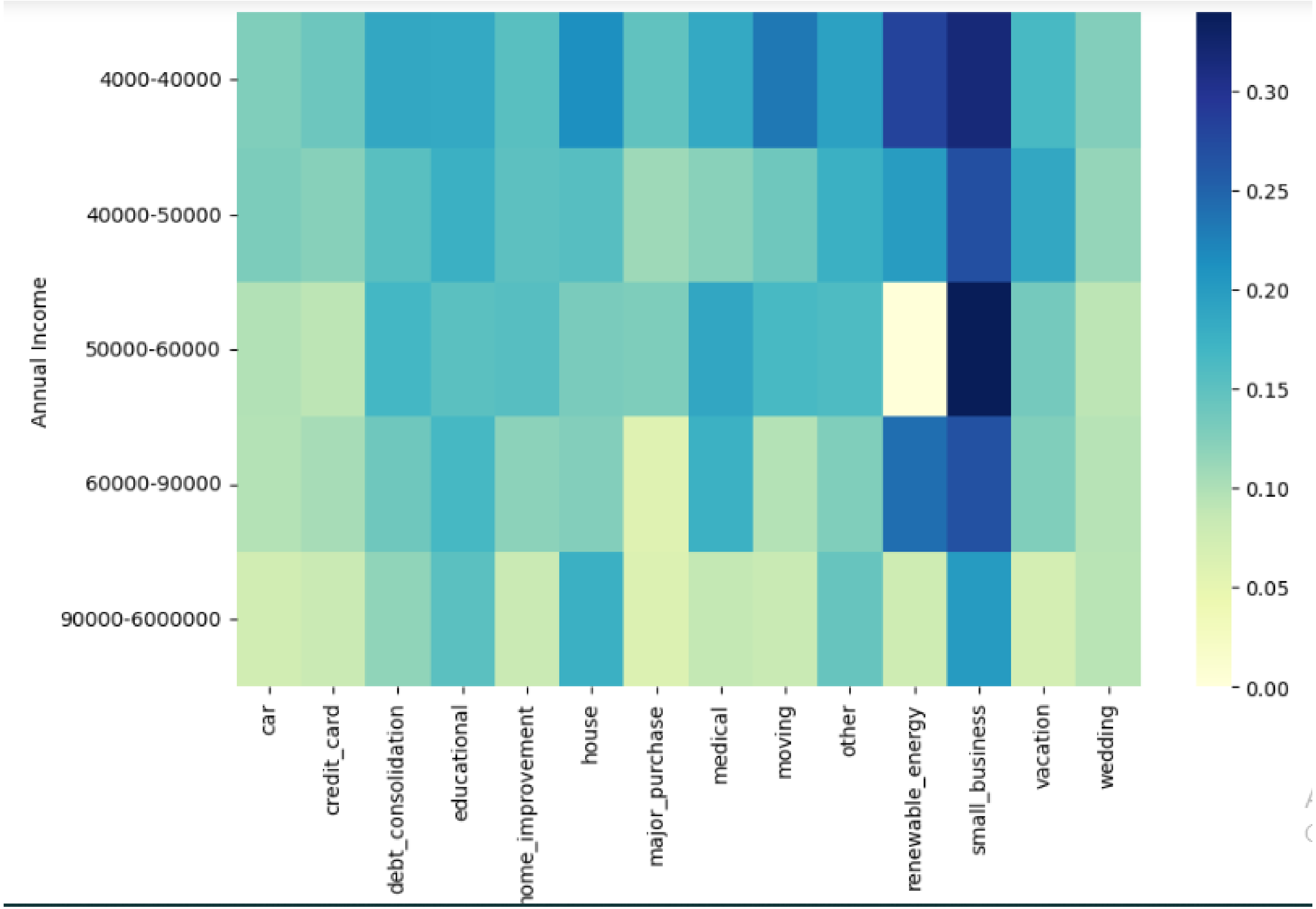
Higher installments lead to higher default rates



car
credit_card
debt_consolidation
educational
home_improvement
house
major_purchase
medical
moving
other
renewable_energy
small_business
vacation
wedding
Purpose
AK AL AR AZ CA CO CT DC DE FL GA HI IA ID IL IN KS KY LA MA MD ME MI MN MO MS MT NC NE NH NJ NM NV NY OH OK OR PA RI SC SD TN TX UT VA VT WA WI WV WY
0.8
0.6
0.4
0.2
0.0
Activate W
Go to Settings

Below are the most risky loan lending for states with purposes are listed below:

- Vacation loans in AK, HI
- Education loans in AR, KS, UT, AZ
- Small business loans in DE, NM, WV, WY, MO
- Wedding loans in MS, OR
- Moving loans in LA, HI
- Renewable energy loans in UT, VA, PA, MD, MA
- House loans in NH, OR, WV
- Home improvement in NE



Small business loans for lowest and middle level annual income
Renewable energy for lowest and high level annual income are risker loans

Findings

- Higher installment amount (above 327)
- Lower annual income (below 37K)
- Higher debt to income ratio (above 15%)
- Applicant’s address state (NV, SD, AK, FL, etc.)
- Loan issue month (Dec, May, Sep, Jun)
- Higher interest rate (above 13%)
- Higher revolving line utilization rate (above 58%)
- Repayment term (5 years)

Findings

- Loan grade & sub-grade (D to G)
- Missing employment record
- Loan purpose (small business, renewable energy, educational)
- Derogatory public records (1 or 2)
- High loan amount & interest rate for lower income group
- High installment and longer repayment term
- Residential state and loan purpose
- Income group and loan purpose