DASC 1223H
Dr. Schubert

Fully Paid Loan Analysis

By Team 2: Ethan Ericson, Harvy Maza, Andrew Mendez, and Ethan Styles.

Goals

What indicates a good borrower?

How can we improve applicants who are unlikely to succeed?

Hypotheses

We hypothesized that the following variables would indicate someone's ability to fully pay off their loan.

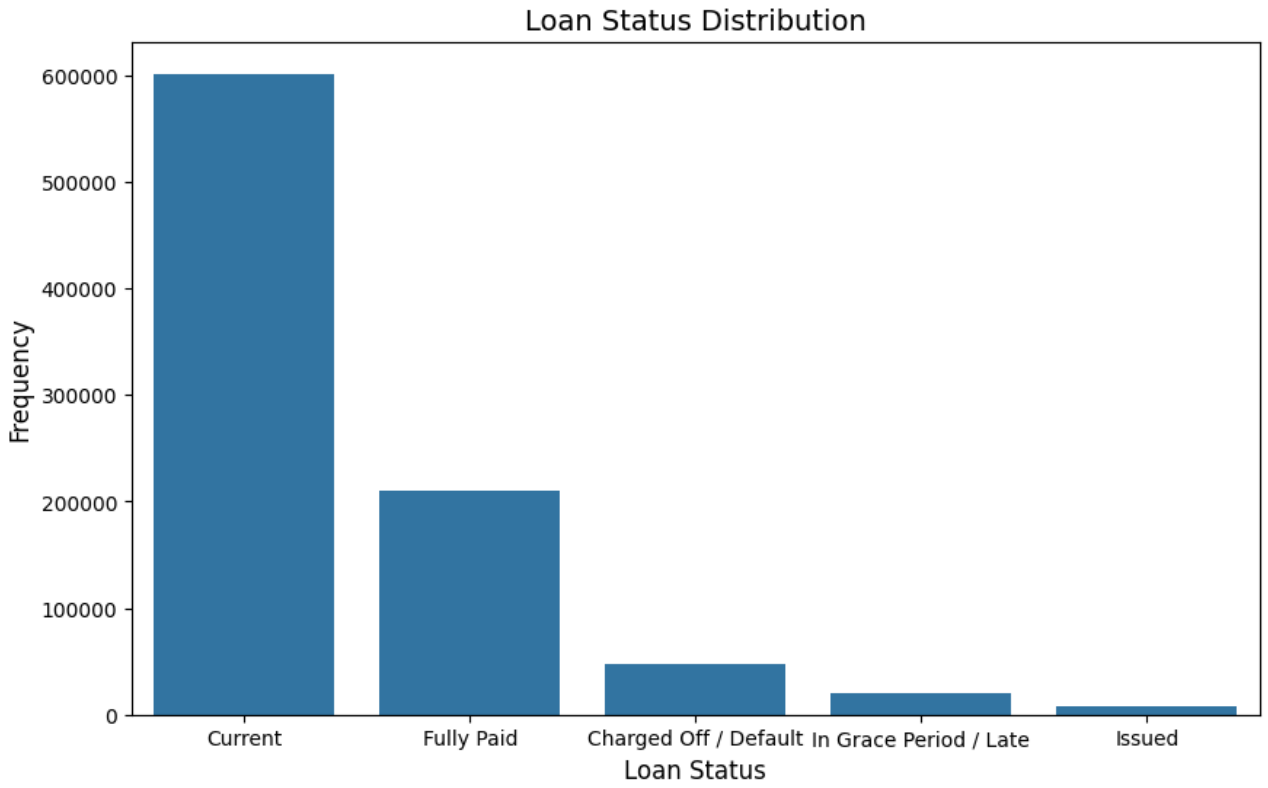
Harvy	Andrew	Ethan S.	Ethan E.
Loan-related Factors	Credit History	Personal Demographics	Financial Indicators
Loan Amount Interest Rate	Number of Delinquencies/Derogatory Public Records	Homeownership Status Verification Status State	Annual Income Debt-Income Ratio Employment Length

Data Set Overview

Loan Status Overview

Loan Status: The majority of loans are current loans, then fully paid.

Statuses: Fully Paid are borrowers that fully pay off their loans, Charged Off / Default are borrowers that fail to do so.



Loan Related Factors

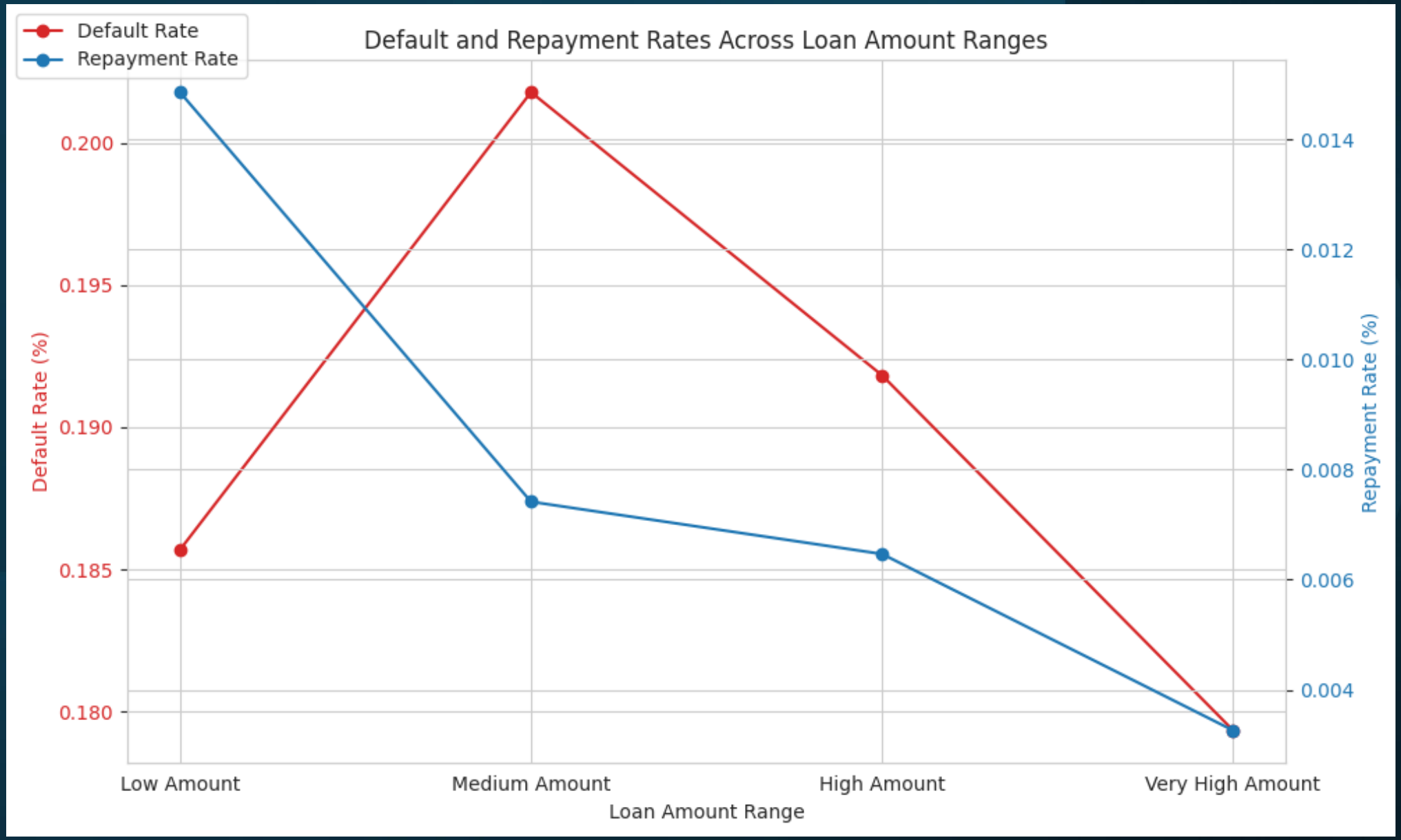
Loan Amount Impact on Payment Status: Exploring the Relationship

- Importance:

It can help determine if borrowers with larger initial loan amounts are likely to default compared to those with smaller loan amounts

Making Connections:

- Loan Amount Payment status (repayment ratio), and default rate.
- Interest Rates and Repayment ratio
- Interest and Default Rate.


Default and Repayment Rates Across Loan Amount Ranges
Default Rate
Repayment Rate
Default Rate (%)
Repayment Rate (%)
0.200
0.195
0.190
0.185
0.180
0.014
0.012
0.010
0.008
0.006
0.004
Low Amount
Medium Amount
High Amount
Very High Amount
Loan Amount Range

Repayment Likelihood and Loan Amounts: Implications for Lending Strategies

This information helps in the identification of reliable borrowers by:

1. Recognizing the inverse relationship between default and repayment rates

2. Observing the trend which indicates that risk escalates with loan amount

3. The impact on repayment likelihood diminishes as loan amounts increase.

The decreasing repayment rate across higher loan amount ranges underscores the potential challenges borrowers may face in managing and repaying larger loans

Regression model

OLS Regression Results

Dep. Variable:	repayment_status	R-squared:	0.000
Model:	OLS	Adj. R-squared:	0.000
Method:	Least Squares	F-statistic:	7.368
Date:	Wed, 24 Apr 2024	Prob (F-statistic):	6.21e-05
Time:	00:05:19	Log-Likelihood:	2.4155e+05
No. Observations:	208942	AIC:	-4.831e+05
Df Residuals:	208938	BIC:	-4.830e+05
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.9962	0.001	1633.434	0.000	0.995	0.997
x1	-8.04e-05	3.9e-05	-2.061	0.039	-0.000	-3.94e-06
x2	-8.217e-08	2.25e-08	-3.649	0.000	-1.26e-07	-3.8e-08
x3	3.527e-09	3.22e-09	1.096	0.273	-2.78e-09	9.83e-09

Omnibus:	367231.440	Durbin-Watson:	1.987
Prob(Omnibus):	0.000	Jarque-Bera (JB):	246848413.797
Skew:	-12.975	Prob(JB):	0.00
Kurtosis:	169.375	Cond. No.	3.36e+05

While the overall model performance indicates minimal variance explanation, individual coefficients offer nuanced understandings.

Regression Model Statistics

The findings from the regression model provide valuable insights into the characteristics of clients and their loan repayment behavior, informing tailored lending strategies.

- Default Rate: 5%
- Repayment Rate: 95%
- Average Loan Amount: $10,000
- Average Interest Rate: 8%

Understanding client characteristics and their implications for loan repayment involves considering a comprehensive range of factors beyond the variables examined in this analysis.

Highlights from regression Model

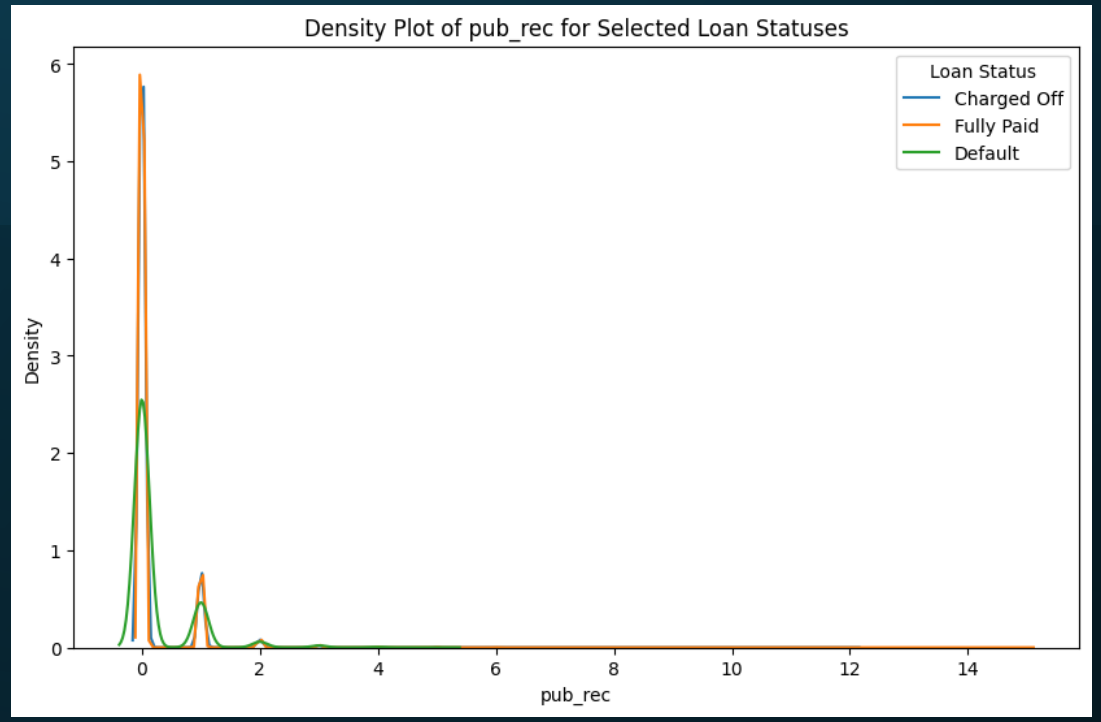
Our Regression statistic model provide us multiple key findings about our Fully Paid borrowers' behavior:

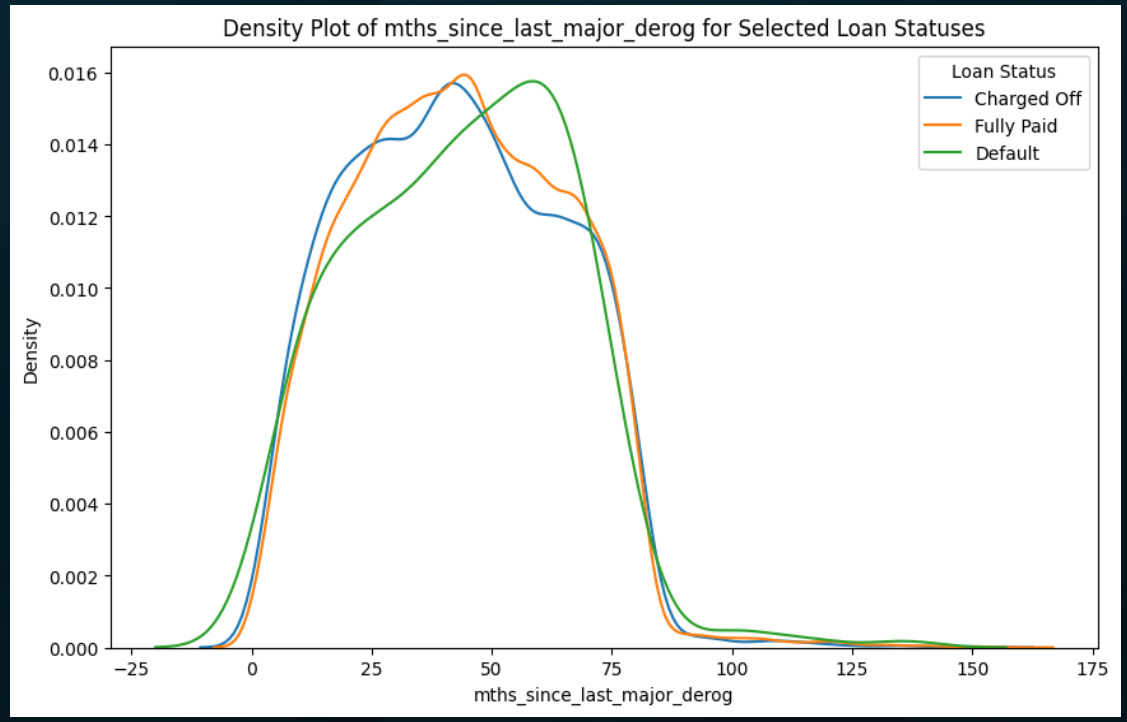
- Moderate Loan Amounts
- High Repayment ratio
- Low default rate
- Reasonable Interest Rates

These characteristics collectively demonstrate that good borrowers exhibit responsible financial behavior and manage debt effectively, making them reliable candidates for lending institutions.

Credit History

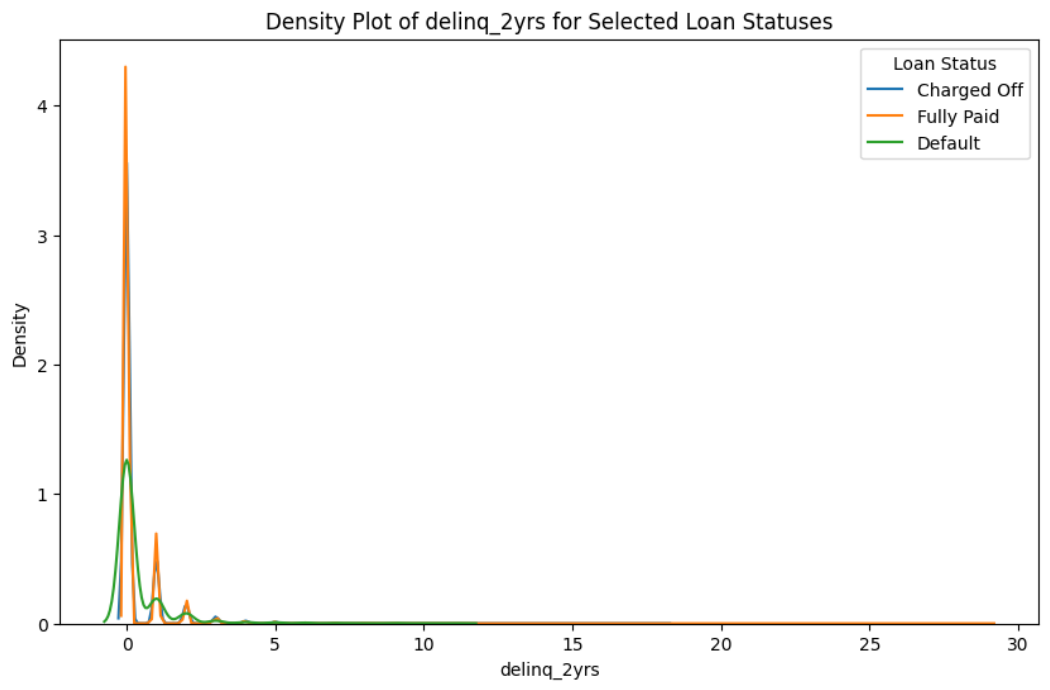
Derogatory for Fully Paid, Charged Off, and Defaulted Loans

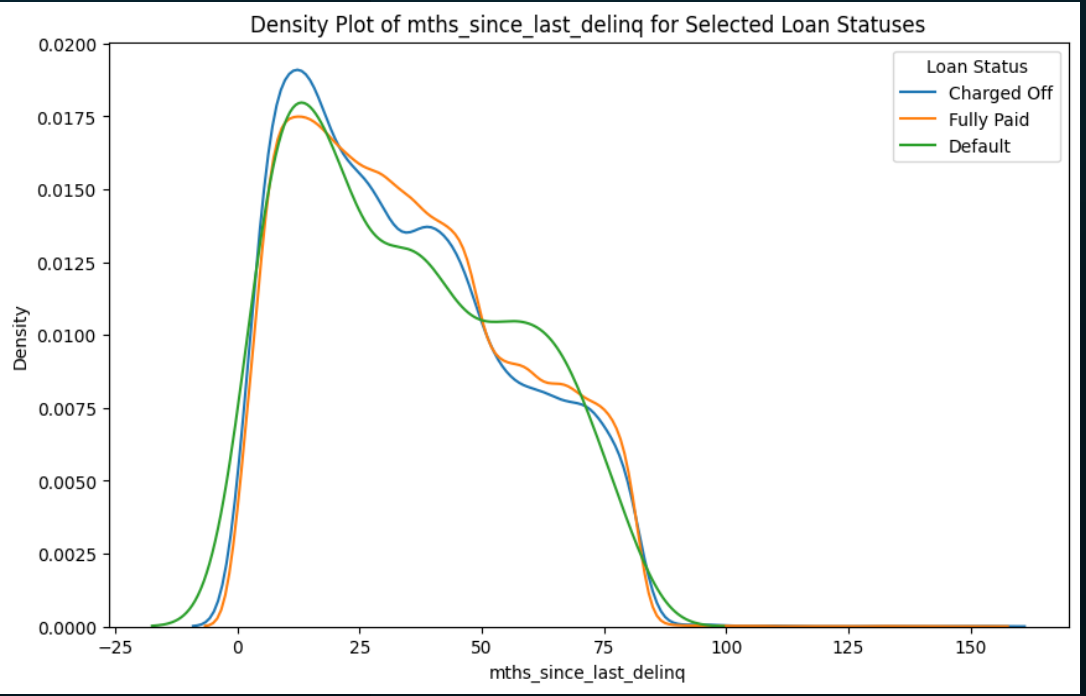




Pub_Rec: Total amount of Derogatory on record
mnths_since_last_major_derog: Months since the most recent 90-day or worse rating

Delinquincy for Fully Paid, Charged Off, and Defaulted Loans

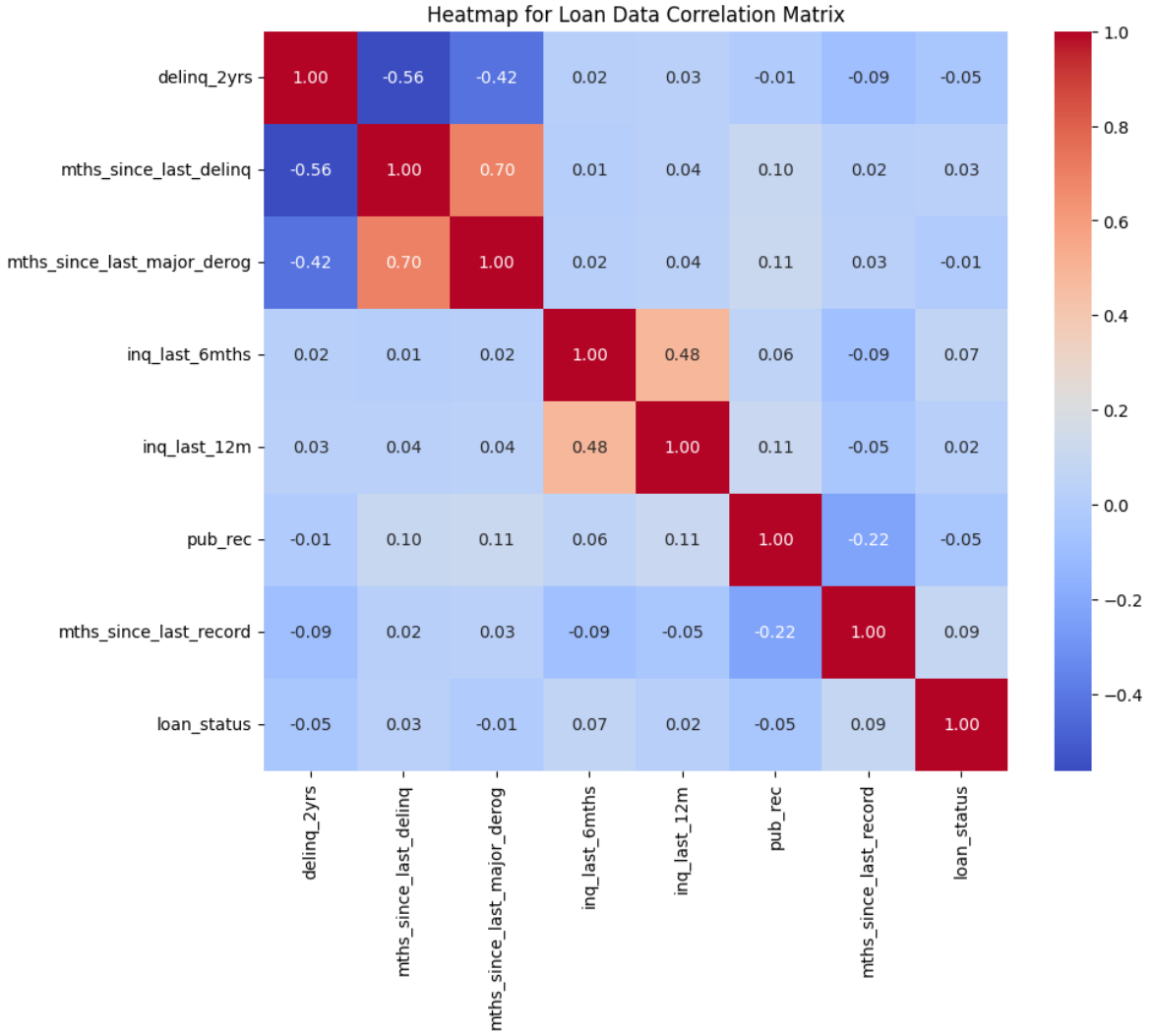




Delinq_2yrs: The number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years

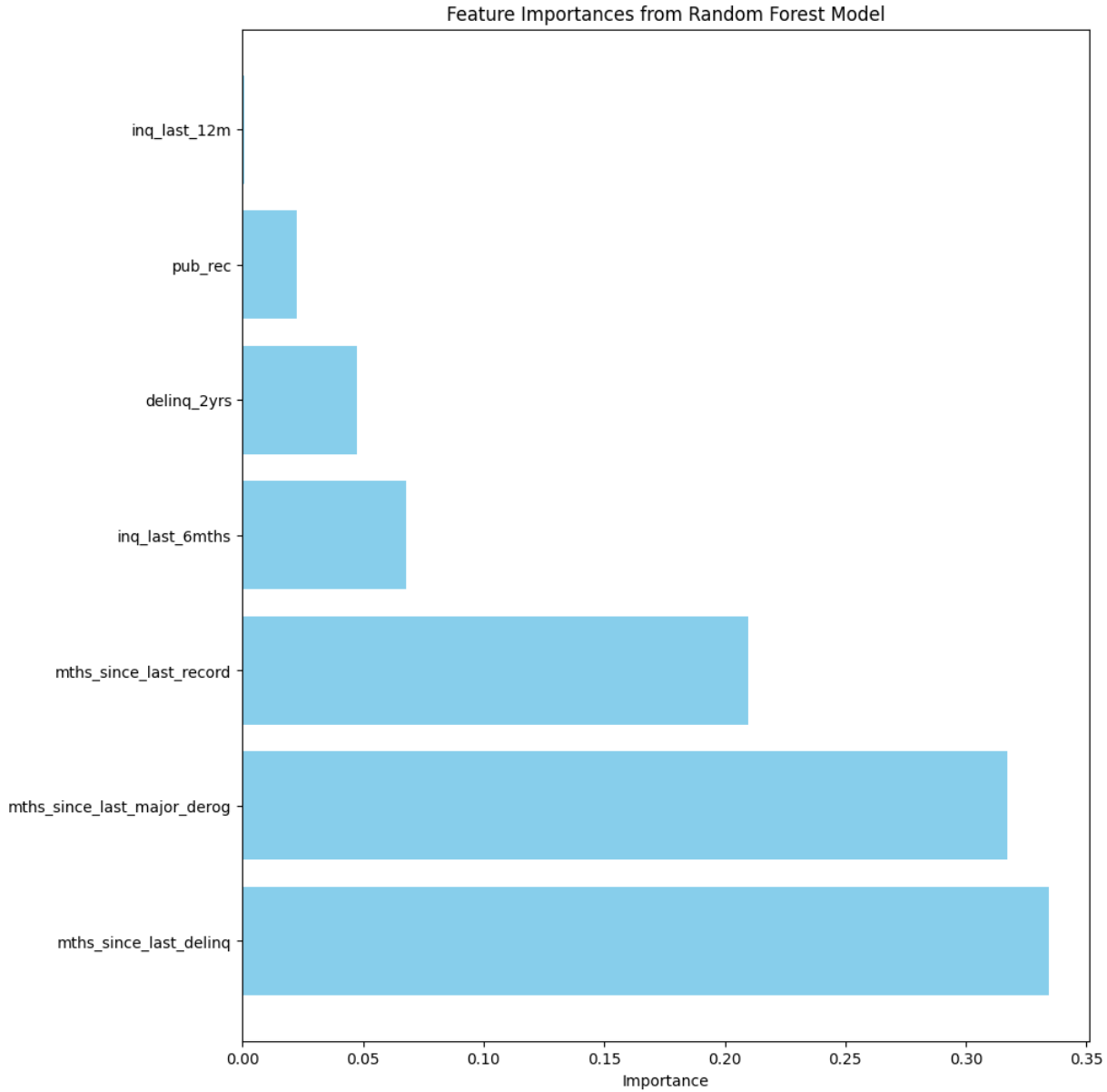
mths_since_last_delinq: The number of months since the borrower's last delinquency.

Loan Correlations to Fully Paid Loans



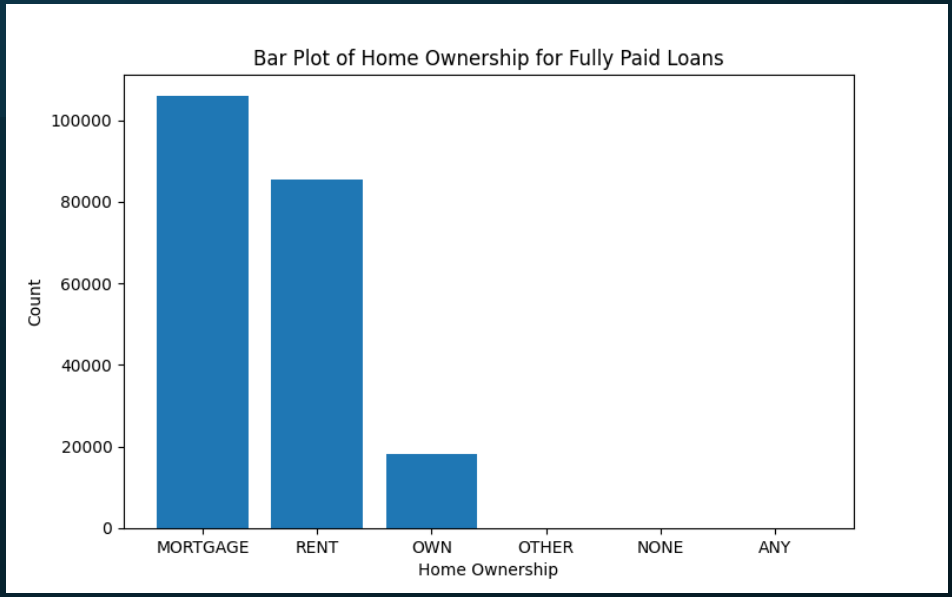
Loan Random Forest Model for Fully Paid, Charged Off, and Default

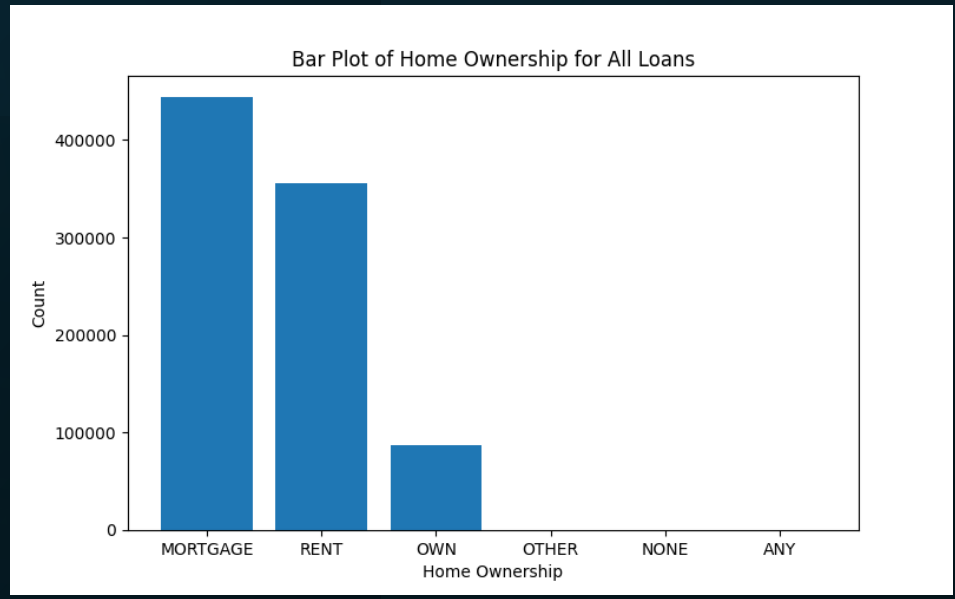
Conclusion: The history of a borrower isn't enough by itself to predict if a loan will fully be paid off, but the biggest indicators are last delinquency and derogatory



Personal Demographics

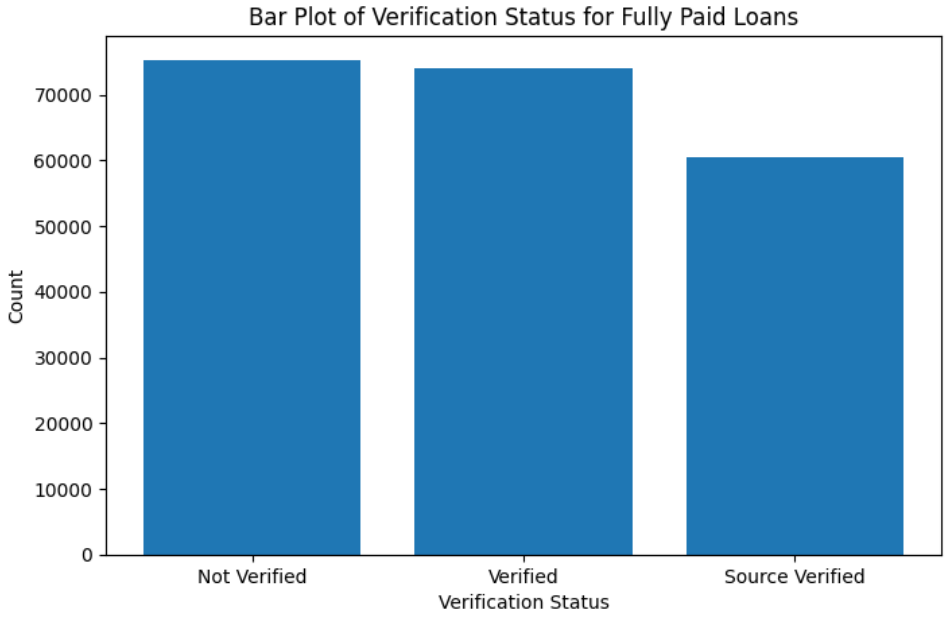
Homeownership

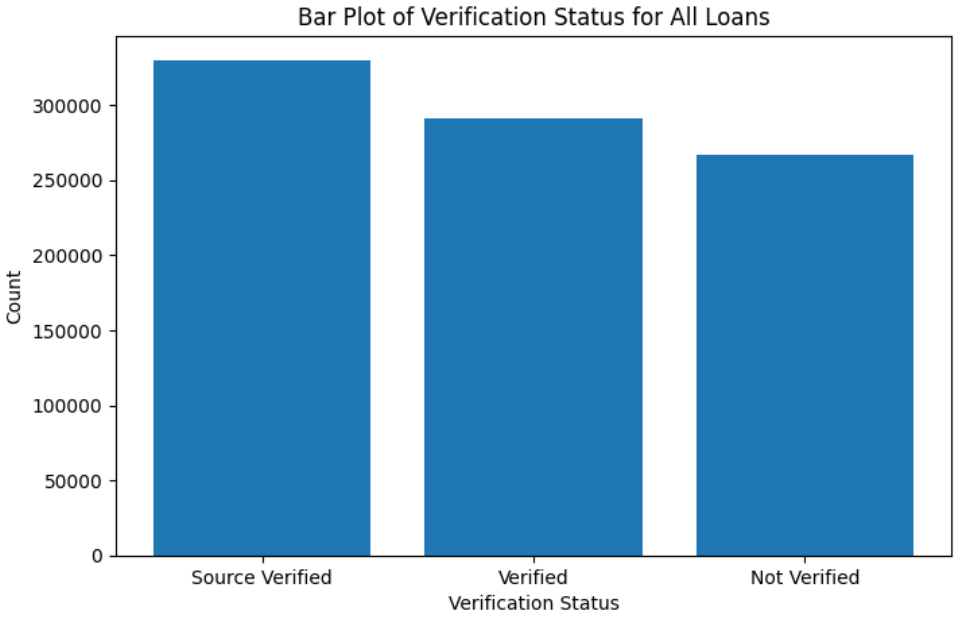




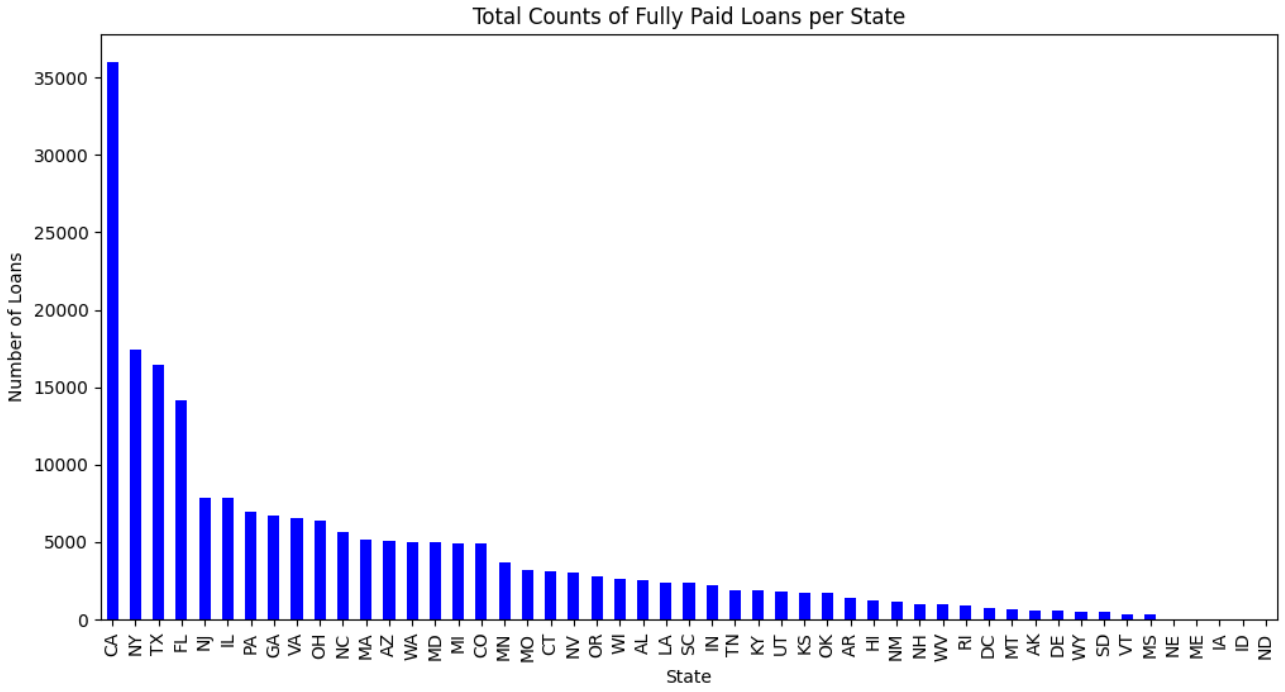
Practically the same distribution

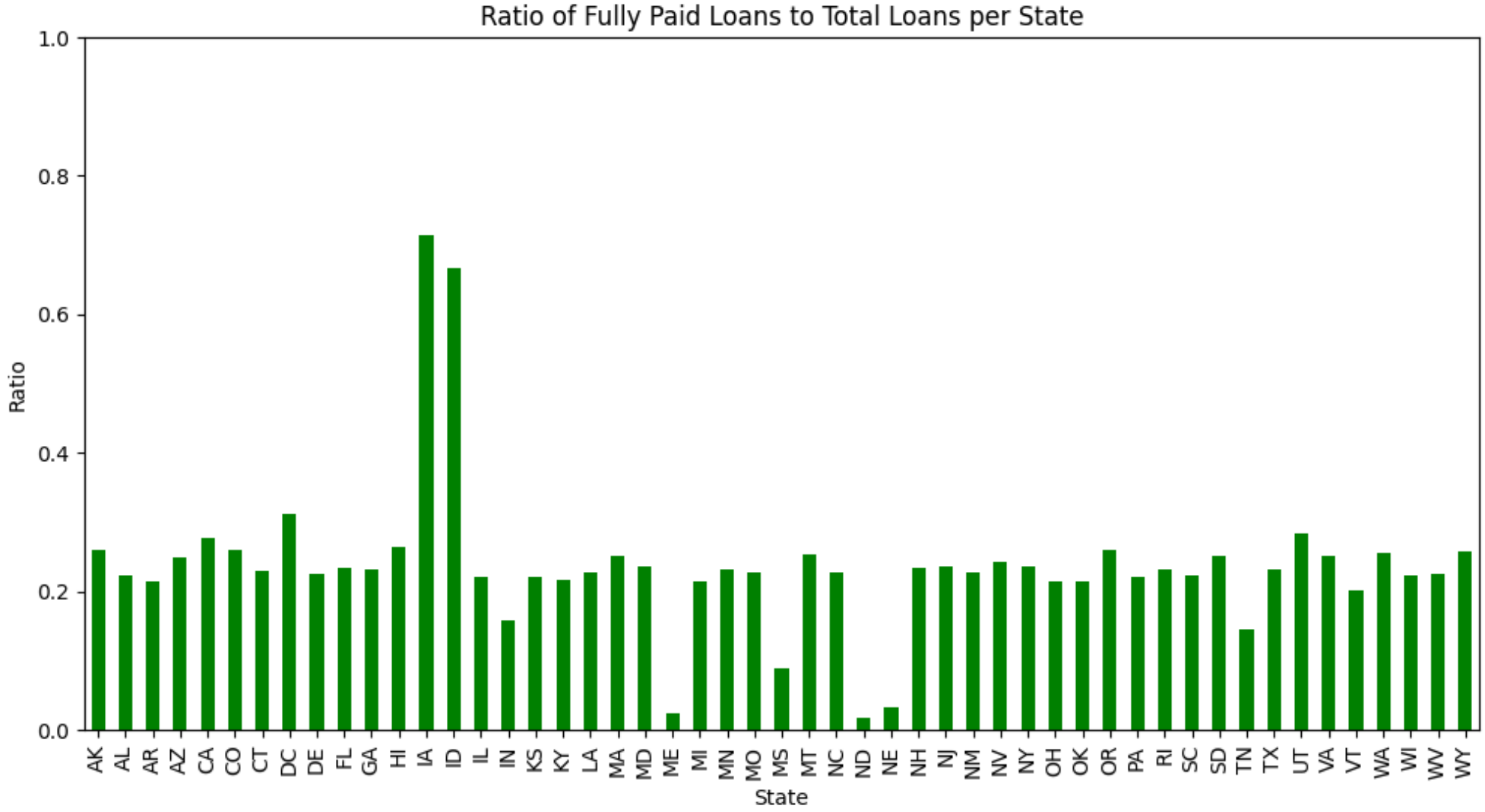
Verification Status





State


Total Counts of Fully Paid Loans per State
Number of Loans
0
5000
10000
15000
20000
25000
30000
35000
CA NY TX FL NJ IL PA GA VA OH NC MA AZ WA MD MI CO MN MO CT NV OR WI AL LA SC IN TN KY UT KS OK AR HI NM NH WV RI DC MT AK DE WY SD VT MS NE ME IA ID ND
State


Ratio of Fully Paid Loans to Total Loans per State
Ratio
1.0
0.8
0.6
0.4
0.2
0.0
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 ND NE NH NJ NM NV NY OH OK OR PA RI SC SD TN TX UT VA VT WA WI WV WY
State

Financial Indicators

Procedure

Goal: Determine if the following areas impact a borrower's ability to fully pay off their loans.

Annual Income

Debt to Income

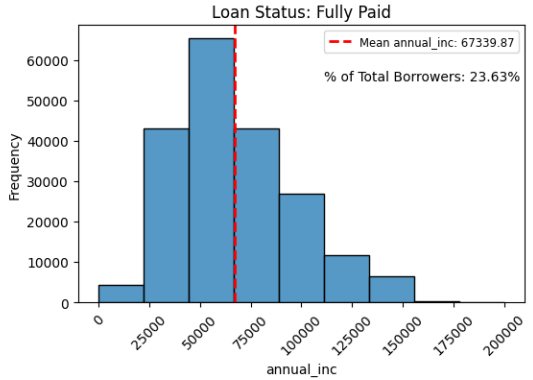
Employment Length

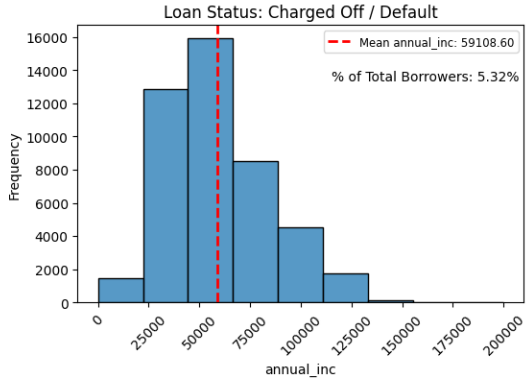
Annual Income

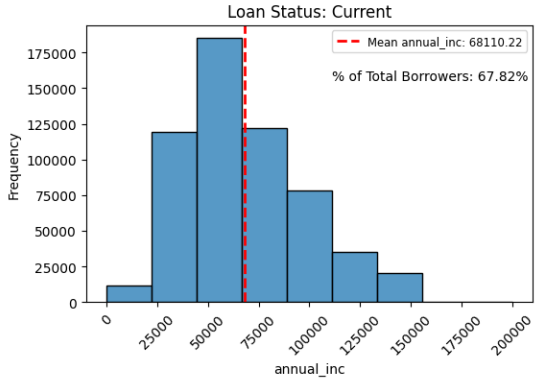
Income Disparity: There is around a $8,000 difference between the mean annual incomes of borrowers who fully pay and those who are charged off/default. This indicates annual income does have an impact.

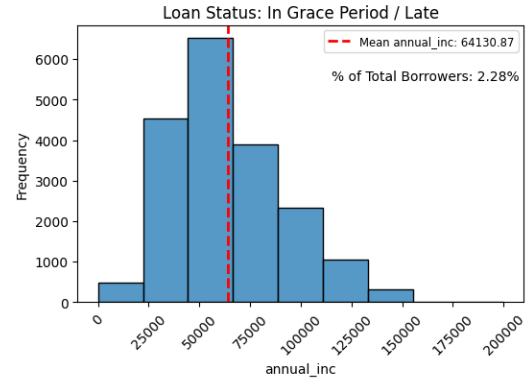
% of Total Borrowers: Indicates the relative weights of loan status groups

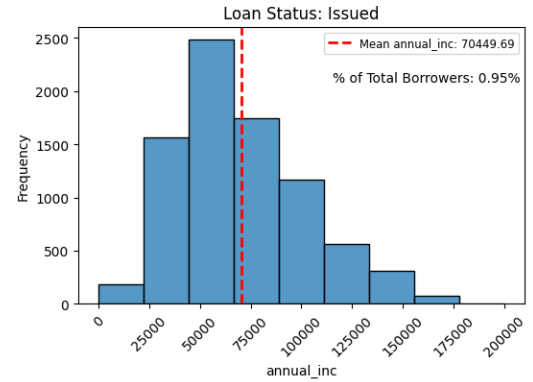
Mean Distribution: Indicates how spread out the mean annual incomes are between different loan statuses.

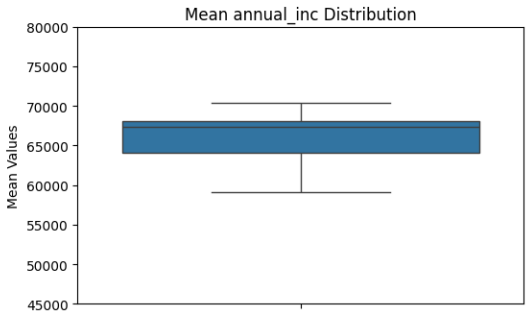








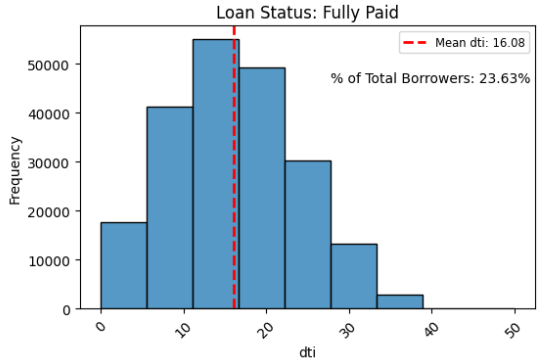


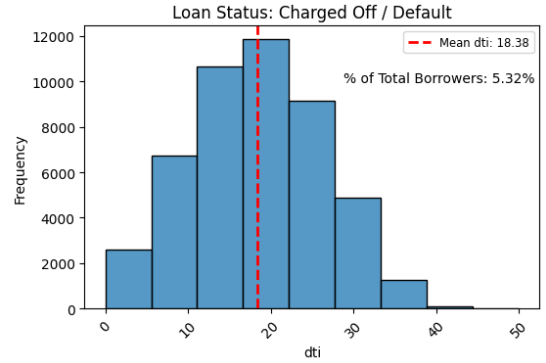


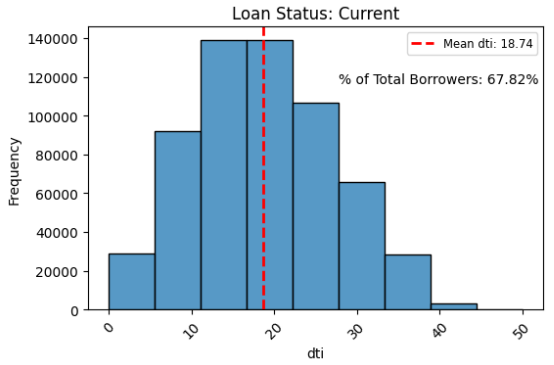
Debt to Income

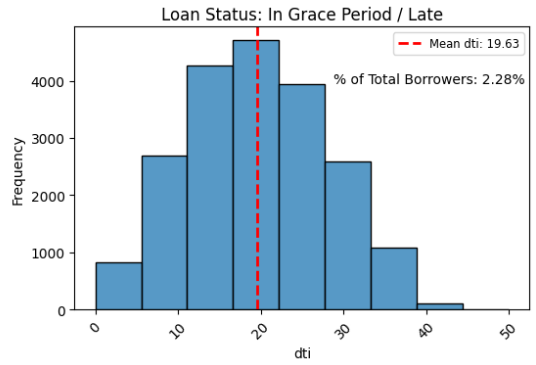
DTI Disparity: There is about a 2% difference in mean DTI between borrowers with fully paid off loans and borrowers who had been charged off / defaulted, indicating that DTI has an impact.

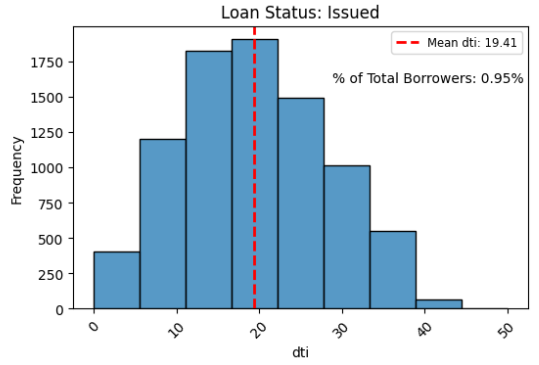
Mean Distribution: Small scale of DTI units makes the mean DTI distribution look clustered, but in reality, there exist significant differences, especially between the fully paid status and charged off / defaulted status.

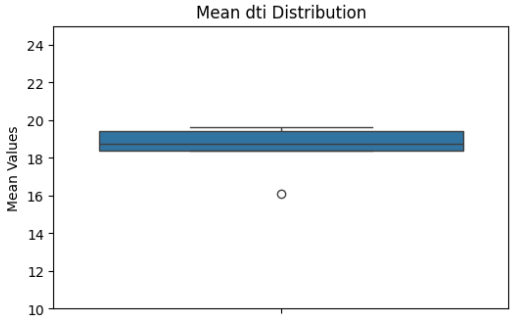






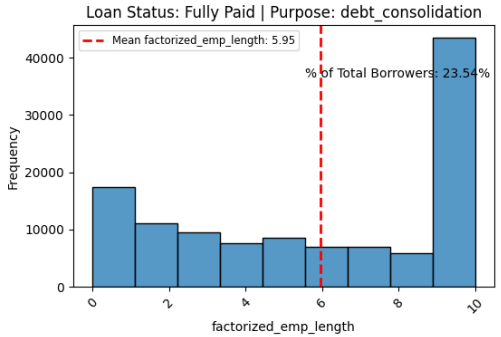


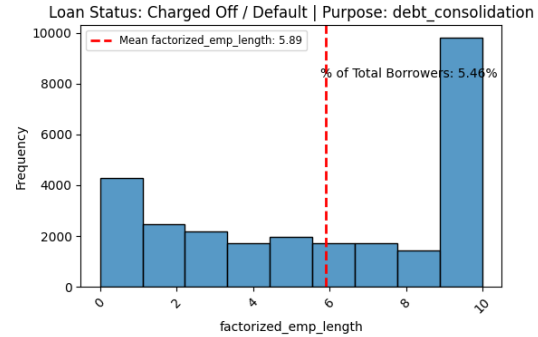


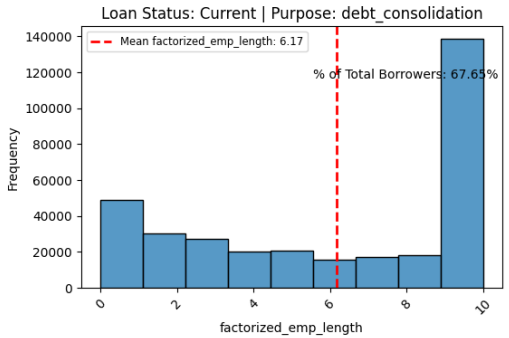


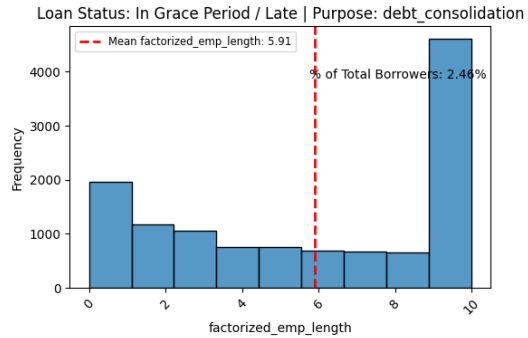
Employment Length

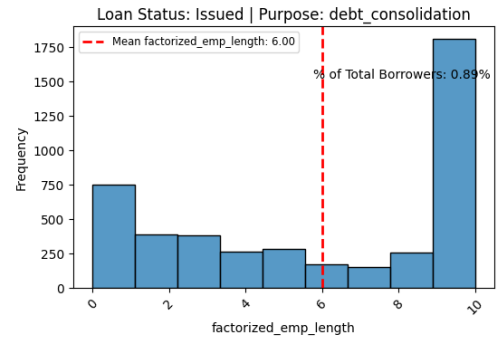
Lack of Disparity: Looking at employment length, which is measured in years, we see barely any difference in the means between status, indicating employment length is not a strong predictor of a borrower being able to fully pay off their loan.

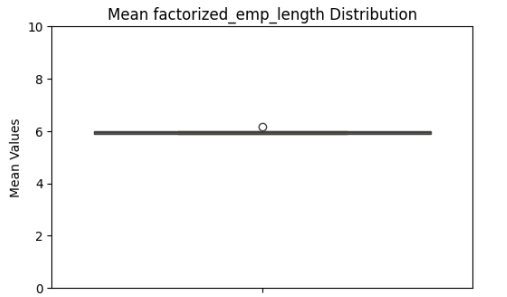






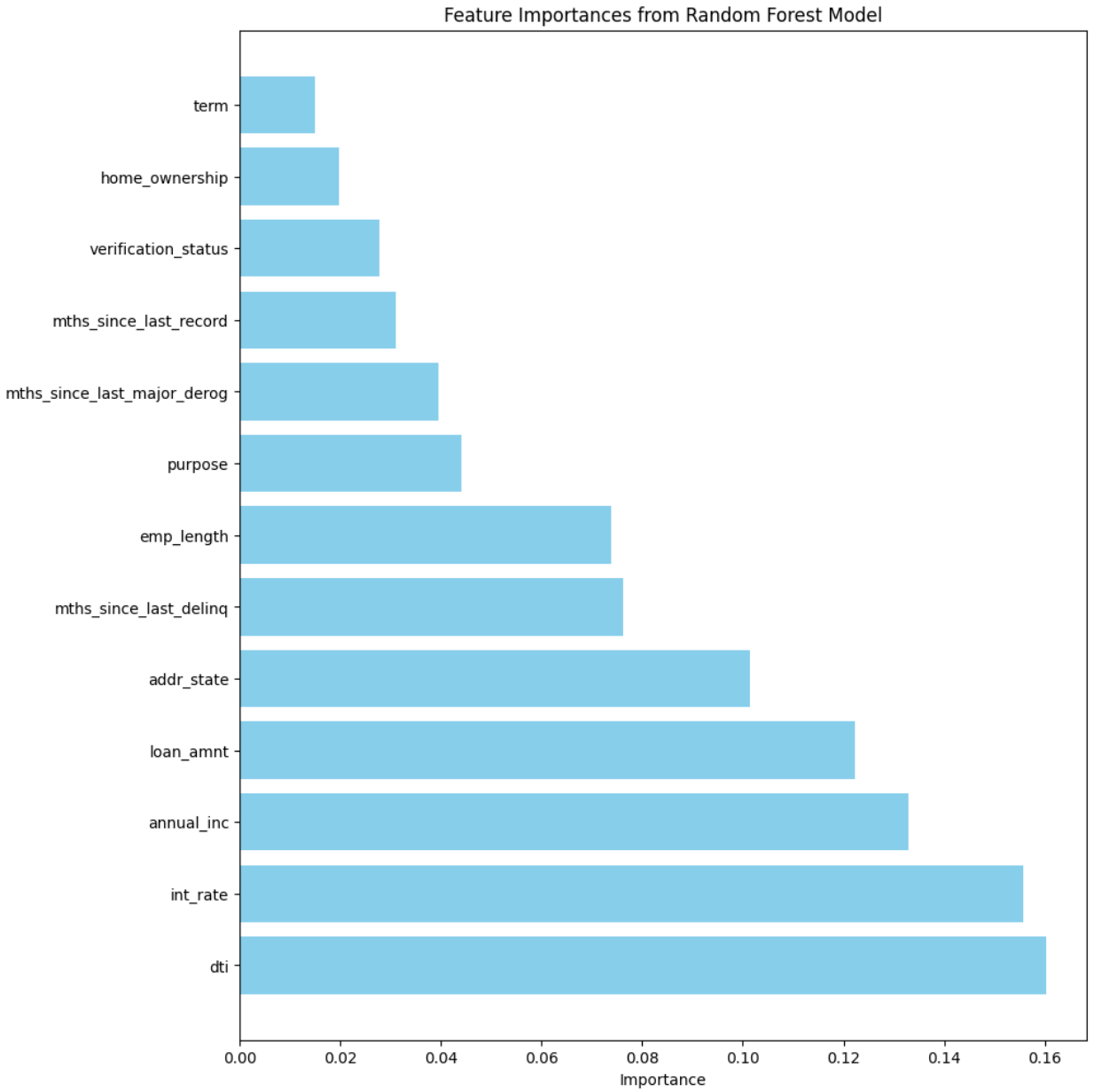






Machine Learning

Attributes from all our EDAs using the Random Forest Model



“Good” Borrower Characteristics

- These characteristics are the relative thresholds that indicate a borrower will fully pay off their loan.
  - Loan Indicators
    - Payment History: High Repayment Rate (95%)
    - Loan Interest: Is fine with a reasonable Interest rate (29 - 40%)
    - Loan Amount: Ask for a moderate Loan Amount (avg. $10,000)
  - Credit Indicators
    - Months since the last Derogatory: 48+ months
    - Months since the last Delinquency: 24+ months
  - Demographic Indicators
    - State: Varies by Population
    - Less Verified the better
  - Financial Indicators
    - Annual Income: $60,000 or more
    - Debt to Income: Within 1% of 16% DTI – This can change based on annual income

Recommendations & Conclusions

- Marketing Department
  - Use our defined “Good: Borrower Characteristics to market loans to groups more likely to pay them off.
- Revolving Lines of Credit Department
  - Use our characteristics as benchmark for current borrowers.

Questions?