Model #101: Credit Card Default Model
Model Development Guide

Andrew Watson

1. Introduction
This study uses consumer credit card data to predict which customers will eventually default. Due to the low proportion of individuals who end up defaulting, specific attention needs to be paid to the true positive rate. Four different models are fit to the data to make predictions: Random Forest, Gradient Boosting Machine, Logistic Regression, and a Neural Network. All of the models demonstrate similar classification accuracies. The Random Forest model better identifies individuals who will default, but this comes at the expense of an increase false positive rate.

2. The Data

2.a Data Dictionary
The Default of Credit Card Clients Data Set comes from the UCI Machine Learning Repository. Attributes were collected for 6 months in 2005 from 30,000 credit card clients in Taiwan. The data were used to measure the likelihood of default.

Table 1: Data Dictionary

Variable	**Definition**
LIMIT_BAL	Credit limit (NT dollars)
SEX	1 = male, 2 = female
EDUCATION	1 = graduate school; 2 = university; 3 = high school; 4 = others
MARRIAGE	1 = married, 2 = single, 3 = others
AGE	Age in years
PAY_1	Repayment delay in months for September, 2005
PAY_2	Repayment delay in months for August, 2005
PAY_3	Repayment delay in months for July, 2005
PAY_4	Repayment delay in months for June, 2005
PAY_5	Repayment delay in months for May, 2005
PAY_6	Repayment delay in months for April, 2005
BILL_AMT1	Bill statement (NT dollars) for September, 2005
BILL_AMT2	Bill statement (NT dollars) for August, 2005
BILL_AMT3	Bill statement (NT dollars) for July, 2005
BILL_AMT4	Bill statement (NT dollars) for June, 2005
BILL_AMT5	Bill statement (NT dollars) for May, 2005
BILL_AMT6	Bill statement (NT dollars) for April, 2005
PAY_AMT1	Amount of previous statement paid (NT dollars) for September, 2005
PAY_AMT2	Amount of previous statement paid (NT dollars) for August, 2005
PAY_AMT3	Amount of previous statement paid (NT dollars) for July, 2005
PAY_AMT4	Amount of previous statement paid (NT dollars) for June, 2005
PAY_AMT5	Amount of previous statement paid (NT dollars) for May, 2005
PAY_AMT6	Amount of previous statement paid (NT dollars) for April, 2005
DEFAULT	Yes = 1, No =0

2.b Data Quality Check
While the dataset contained no missing values and the labels generally matched the data dictionary, a few minor alterations applied. First, PAY_1 was incorrectly labeled as PAY_0 and had to be corrected. Next, EDUCATION contained 345 counts that were either 0, 5, or 6. These values were reclassified as 4 for others. Similarly, MARRIAGE contained 54 counts with a value of 0. These values were reclassified as 3 for others. Table 1A displays the original variables after cleaning. To compare to the original raw data, please see Appendix Table 1A.

Table 2: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
LIMIT_BAL	30,000	167,484.30	129,747.70	10,000	50,000	140,000	240,000	1,000,000
SEX	30,000	1.60	0.49	1	1	2	2	2
EDUCATION	30,000	1.84	0.74	1	1	2	2	4
MARRIAGE	30,000	1.56	0.52	1	1	2	2	3
AGE	30,000	35.49	9.22	21	28	34	41	79
PAY_1	30,000	-0.02	1.12	-2	-1	0	0	8
PAY_2	30,000	-0.13	1.20	-2	-1	0	0	8
PAY_3	30,000	-0.17	1.20	-2	-1	0	0	8
PAY_4	30,000	-0.22	1.17	-2	-1	0	0	8
PAY_5	30,000	-0.27	1.13	-2	-1	0	0	8
PAY_6	30,000	-0.29	1.15	-2	-1	0	0	8
BILL_AMT1	30,000	51,223.33	73,635.86	-165,580	3,558.8	22,381.5	67,091	964,511
BILL_AMT2	30,000	49,179.08	71,173.77	-69,777	2,984.8	21,200	64,006.2	983,931
BILL_AMT3	30,000	47,013.15	69,349.39	-157,264	2,666.2	20,088.5	60,164.8	1,664,089
BILL_AMT4	30,000	43,262.95	64,332.86	-170,000	2,326.8	19,052	54,506	891,586
BILL_AMT5	30,000	40,311.40	60,797.16	-81,334	1,763	18,104.5	50,190.5	927,171
BILL_AMT6	30,000	38,871.76	59,554.11	-339,603	1,256	17,071	49,198.2	961,664
PAY_AMT1	30,000	5,663.58	16,563.28	0	1,000	2,100	5,006	873,552
PAY_AMT2	30,000	5,921.16	23,040.87	0	833	2,009	5,000	1,684,259
PAY_AMT3	30,000	5,225.68	17,606.96	0	390	1,800	4,505	896,040
PAY_AMT4	30,000	4,826.08	15,666.16	0	296	1,500	4,013.2	621,000
PAY_AMT5	30,000	4,799.39	15,278.31	0	252.5	1,500	4,031.5	426,529
PAY_AMT6	30,000	5,215.50	17,777.47	0	117.8	1,500	4,000	528,666
DEFAULT	30,000	0.22	0.42	0	0	0	0	1

2.c Data Split
The data were split into a Train group for model fitting, a Test group for model evaluation, and Validation group for performance monitoring.

Table 3: Train, Test, Validation Split

Group	Counts	% of Dataset
Train	15,180	50.60
Test	7323	24.41
Validation	7497	24.99

3. Feature Engineering

Table 4 lists features that were engineered from the default variables. This new metrics are used in place of LIMIT_BAL, PAY, BILL_AMT, and PAY_AMT for fitting and evaluating models. Additionally, the monthly values for the Payment Ratio and Utilization features were not used with the models, as the Maximum, Minimum, and Average values for each were engineered and applied in their place.

Table 4: Engineered Feature Dictionary

Feature	Definition
Average Bill Amount (Avg_Bill_Amt)	Mean bill amount taken over 6 months.
Average Payment Amount (Avg_Pmt_Amt)	Mean payment amount taken over 6 months.
Payment Ratio (Pmt_Ratio)	Payment amount divided by the bill amount for each month. If the bill amount was zero, the corresponding ratio was defined as 1.
Average Payment Ratio (Avg_Pmt_Ratio)	Mean payment ratio taken over 6 months.
Utilization (Util)	Bill amount divided by the credit limit for each month.
Average Utilization (Avg_Util)	Mean utilization taken over 6 months.
Balance Growth Over 6 Months (Bal_Growth_6mo)	Bill amount from April was subtracted from the bill amount from September.
Utilization Growth Over 6 Months (Util_Growth_6mo)	Utilization from April subtracted from the utilization from September.
Max Bill Amount (Max_Bill_Amt)	Maximum bill amount over 6 months.
Max Payment Amount (Max_Pmt_Amt)	Maximum payment amount over 6 months.
Max Delinquency (Max_DLQ)	All the Pay_X variables with values of -1 and -2 were set to zero. From there, the max delinquency was taken over 6 months.
Max Utilization (Max_Util)	Maximum utilization over 6 months.
Max Payment Ratio (Max_Pmt_Ratio)	Maximum payment ratio over 6 months.
Min Bill Amount (Min_Bill_Amt)	Minimum bill amount over 6 months.
Min Payment Amount (Min_Pmt_Amt)	Minimum payment amount over 6 months.
Min Delinquency (Min_DLQ)	All the Pay_X variables with values of -1 and -2 were set to zero. From there, the min delinquency was taken over 6 months.
Min Utilization (Min_Util)	Minimum utilization over 6 months.
Min Payment Ratio (Min_Pmt_Ratio)	Minimum payment amount over 6 months.

Table 5 shows the summary statistics for the new features engineered in this section.

Table 5: Summary Statistics for Engineer Features

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Avg_Bill_Amt	30,000	44,976.95	63,260.72	-56,043	4,781.3	21,051.8	57,104.4	877,314
Avg_Pmt_Amt	30,000	5,275.23	10,137.95	0.00	1,113.29	2,397.17	5,583.92	627,344.30
Pmt_Ratio1	30,000	0.50	26.10	-498	0.04	0.1	1	4,444
Pmt_Ratio2	30,000	0.71	38.81	-385	0.04	0.1	1	5,001
Pmt_Ratio3	30,000	-2.30	475.12	-82,150	0.04	0.1	1	4,444
Pmt_Ratio4	30,000	-0.003	37.21	-4,307	0.04	0.1	1	130
Pmt_Ratio5	30,000	0.47	5.24	-185	0.04	0.1	1	691
Avg_Pmt_Ratio	30,000	-0.13	96.60	-16,429.80	0.05	0.17	0.81	2,667.20
Util1	30,000	0.42	0.41	-0.62	0.02	0.31	0.83	6.46
Util2	30,000	0.41	0.40	-1.40	0.02	0.30	0.81	6.38
Util3	30,000	0.39	0.40	-1.03	0.02	0.27	0.76	10.69
Util4	30,000	0.36	0.37	-1	0.01	0.2	0.7	5
Util5	30,000	0.33	0.35	-1	0.01	0.2	0.6	5
Util6	30,000	0.32	0.35	-2	0.01	0.2	0.6	4
Avg_Util	30,000	0.37	0.35	-0.23	0.03	0.28	0.69	5.36
Bal_Growth_6mo	30,000	12,351.57	43,922.42	-428,791	-2,963	923	19,793.8	708,323
Util_Growth_6mo	30,000	0.11	0.30	-1.83	-0.03	0.01	0.18	5.31
Max_Bill_Amt	30,000	60,572.44	78,404.81	-6,029	10,060	31,208.5	79,599	1,664,089
Max_Pmt_Amt	30,000	15,848.23	37,933.56	0	2,198	5,000	12,100	1,684,259

Max_DLQ	30,000	0.68	1.07	0	0	0	2	8
Max_Util	30,000	0.49	0.43	-0.10	0.07	0.43	0.92	10.69
Max_Pmt_Ratio	30,000	1.06	39.05	-1	0.1	0.4	1	5,001
Min_Bill_Amt	30,000	31,722.21	54,719.58	-339,603	0	8,664.5	39,180.5	551,702
Min_Pmt_Amt	30,000	1,243.65	2,472.83	0	0	17	1,500	63,758
Min_DLQ	30,000	0.08	0.37	0	0	0	0	4
Min_Util	30,000	0.26	0.32	-2	0	0.1	0.5	4
Min_Pmt_Ratio	30,000	-3.17	475.91	-82,150	0	0.04	0.1	2

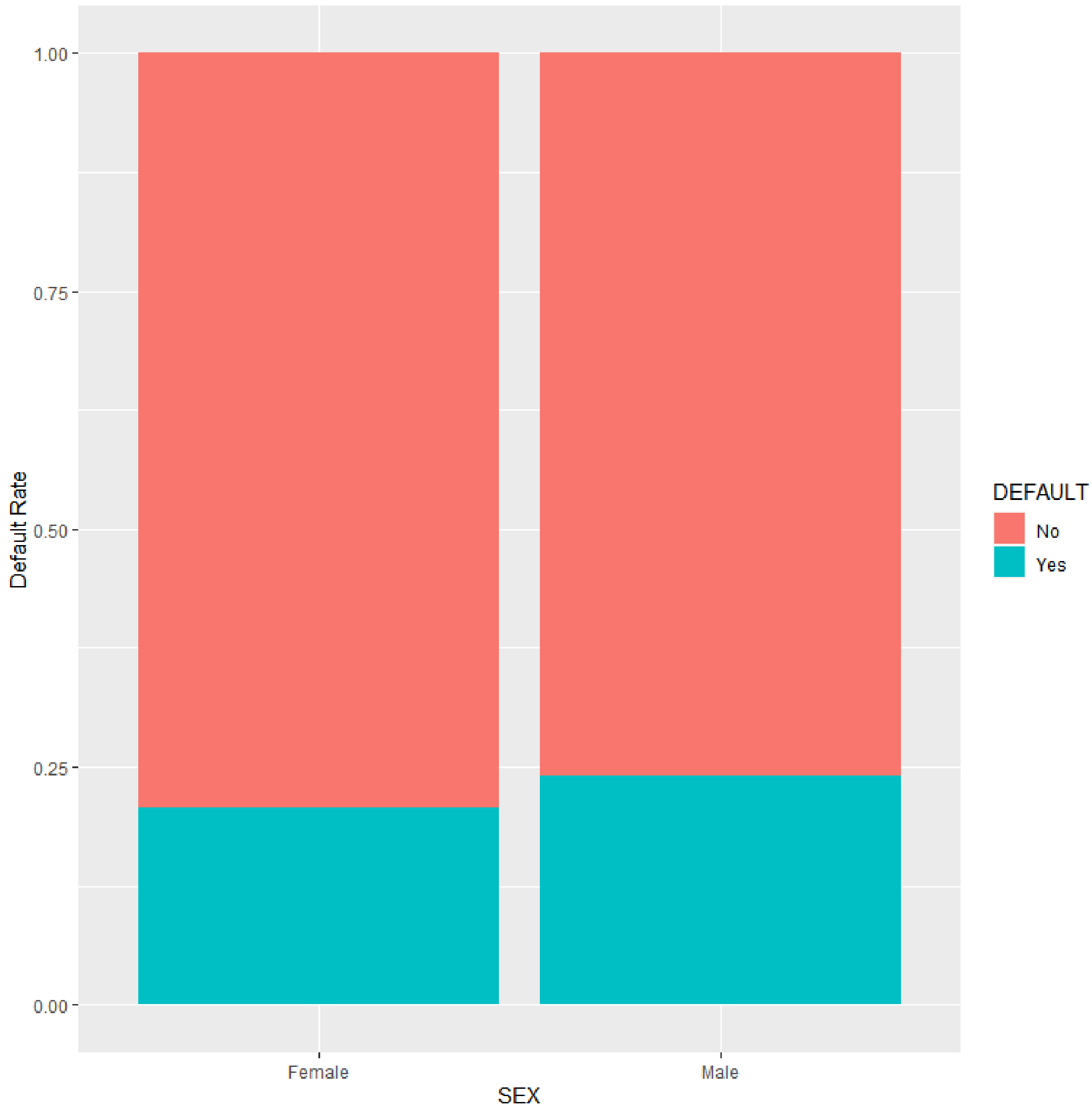
4. EDA

4a. Traditional EDA

Weight of Evidence binning was used to discretize continuous variables for the EDA. For purposes of comparison, the overall default rate is 22.12%.

SEX

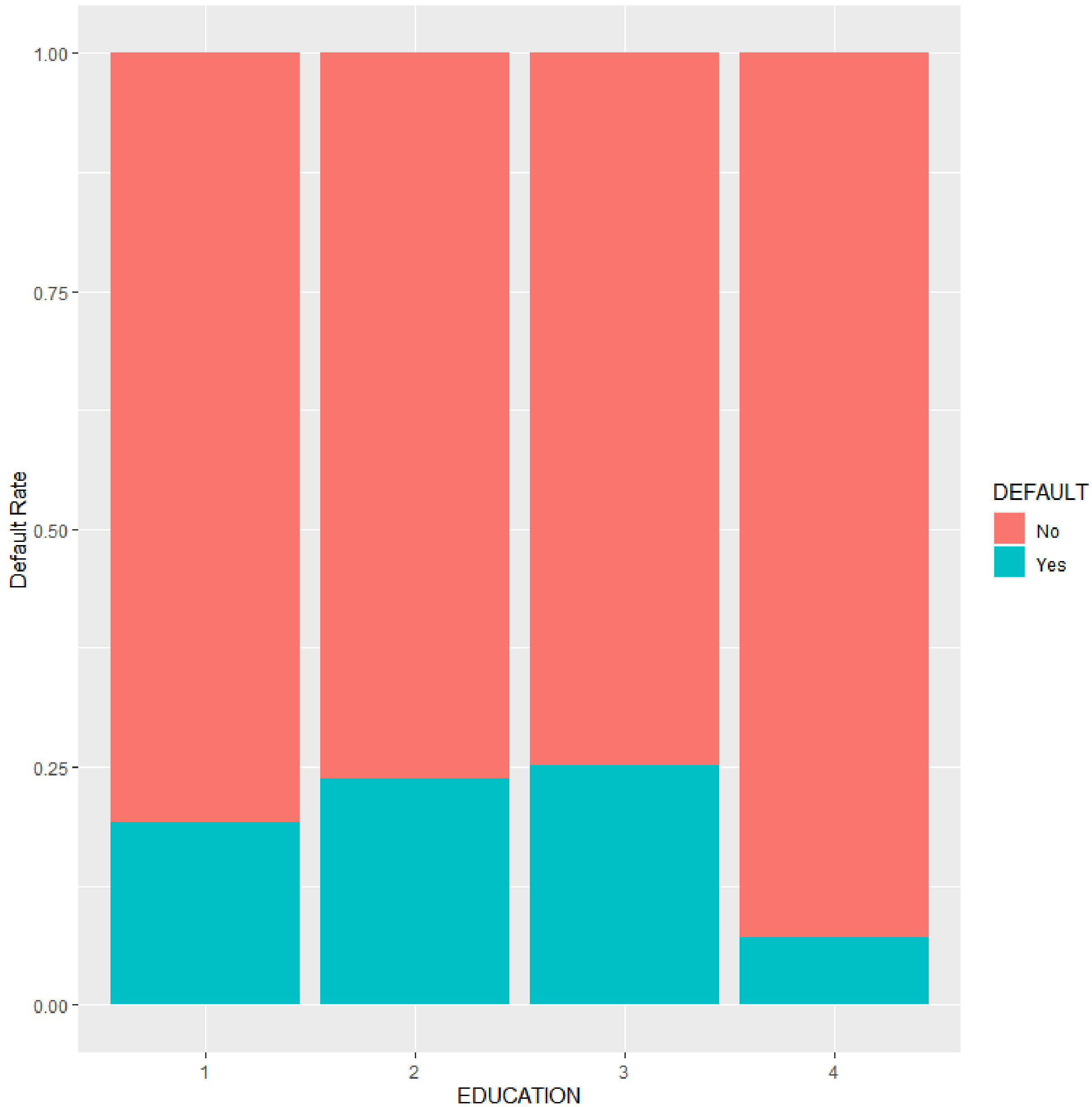
Figure 1: Comparison of Default Rates



Notes:
Though there was not a big discrepancy, Males were slightly more likely than Females to default.

EDUCATION

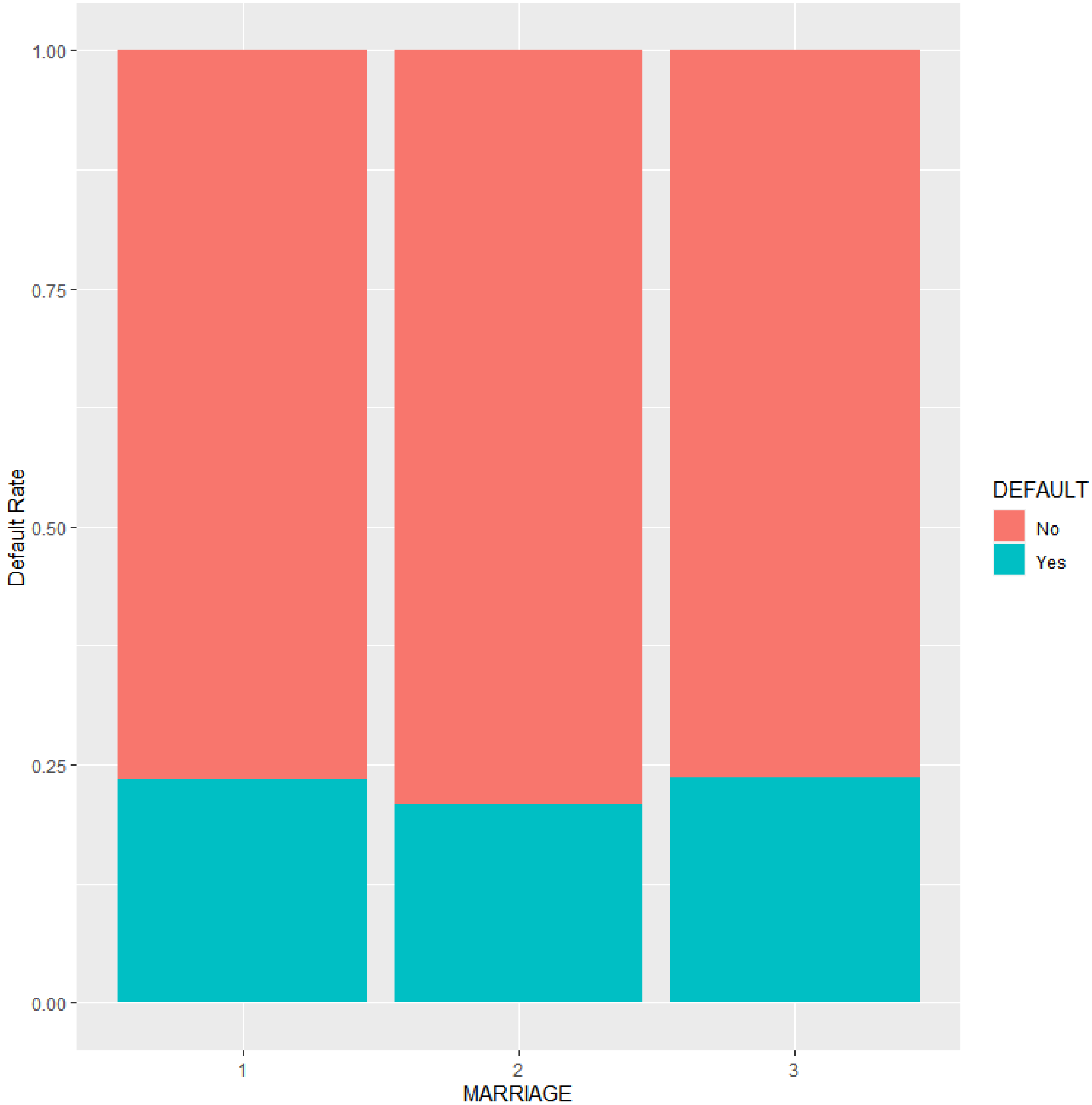
Figure 2: Comparison of Default Rates



Notes:
Perhaps due to a small sample size, the EDUCATION = 4 (other) group was less likely to default.

MARRIAGE

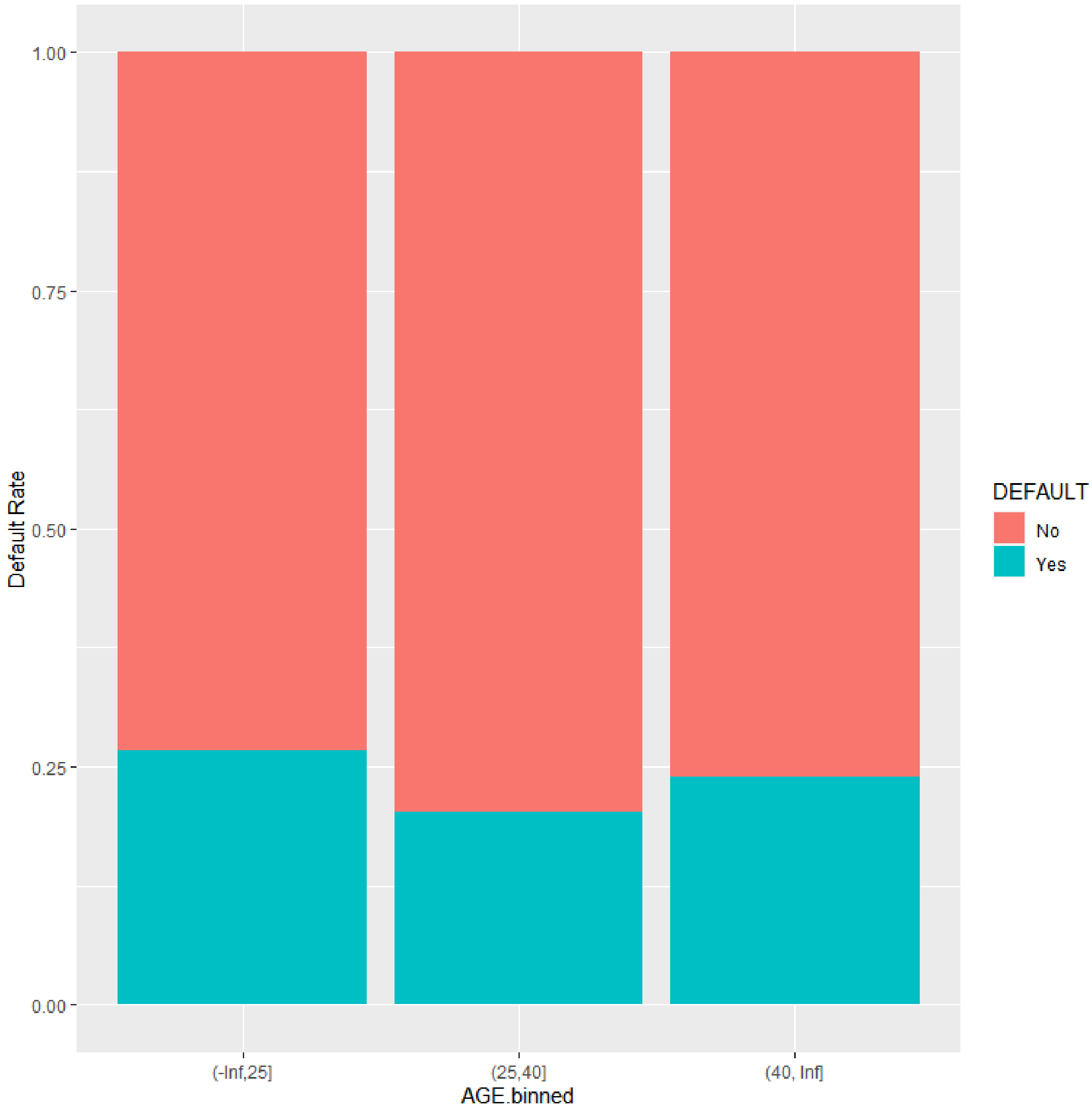
Figure 3: Comparison of Default Rates



Notes:
Married individuals were slightly more likely to default than single individuals.

AGE

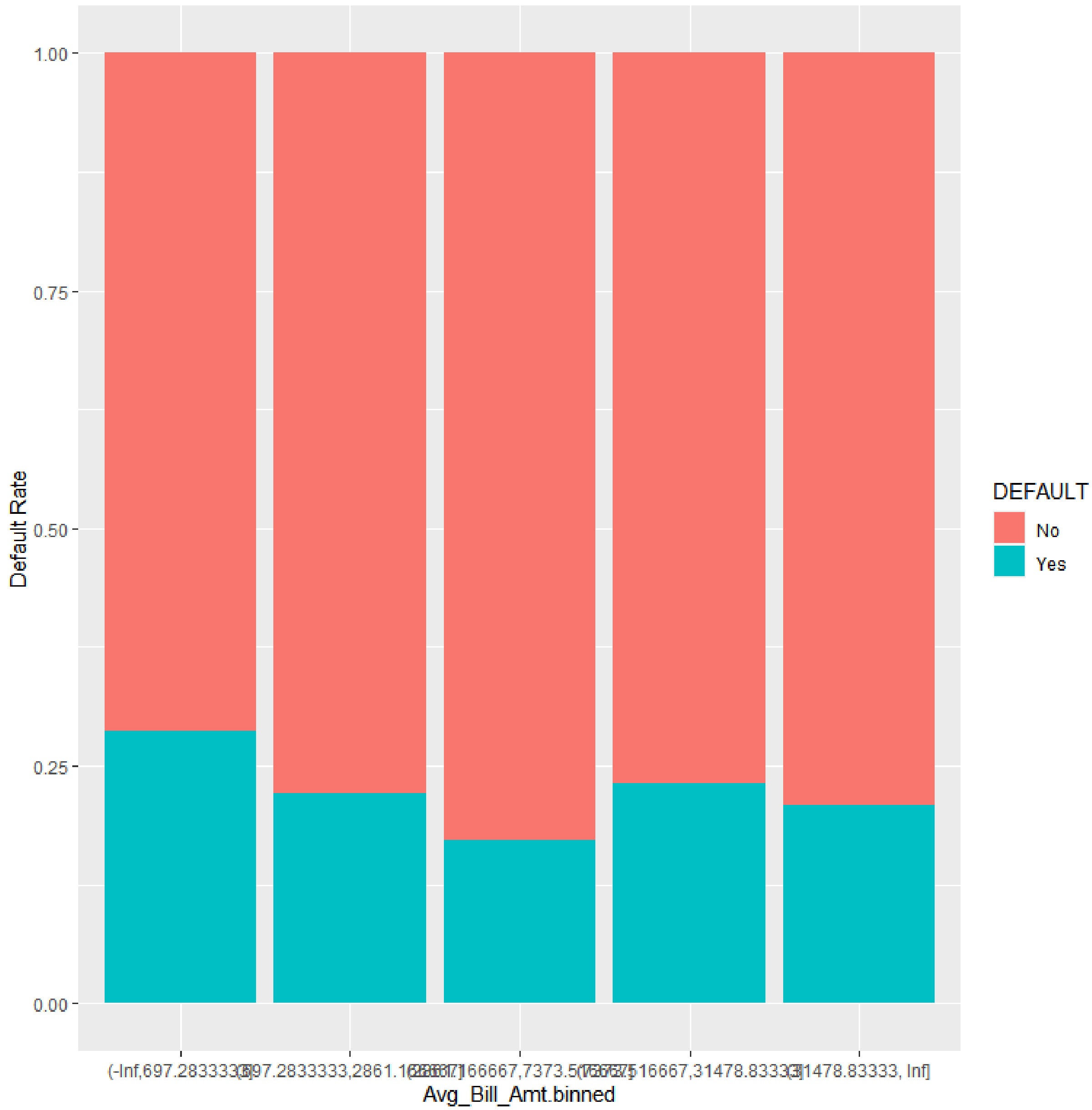
Figure 4: Comparison of Default Rates



Notes:
People between the ages of 25-40 were less likely to default than both younger and older groups.

Avg_Bill_Amt

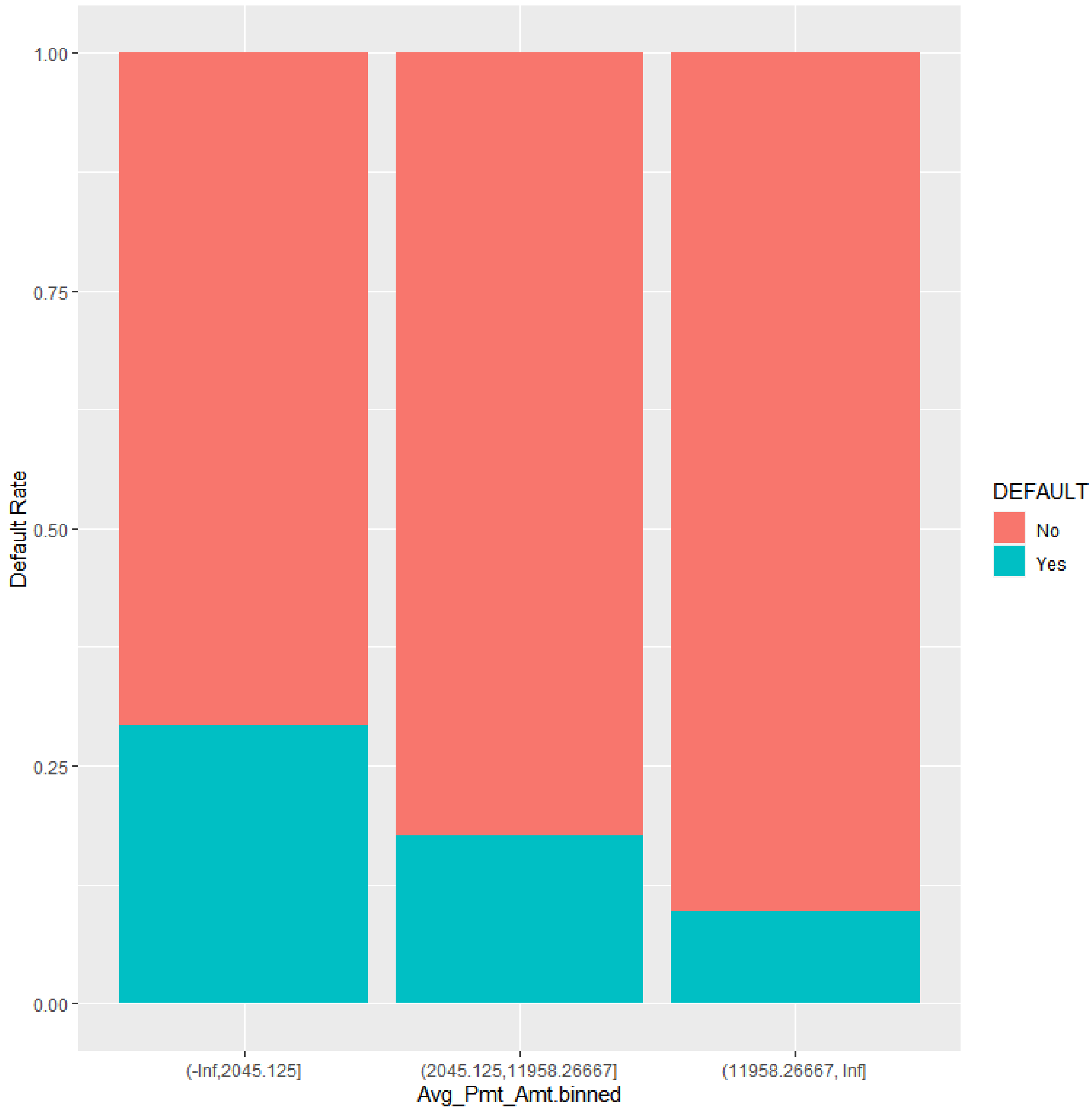
Figure 5: Comparison of Default Rates



Notes:
The default rate did not vary greatly by Avg_Bill_Amt.

Avg_Pmt_Amt

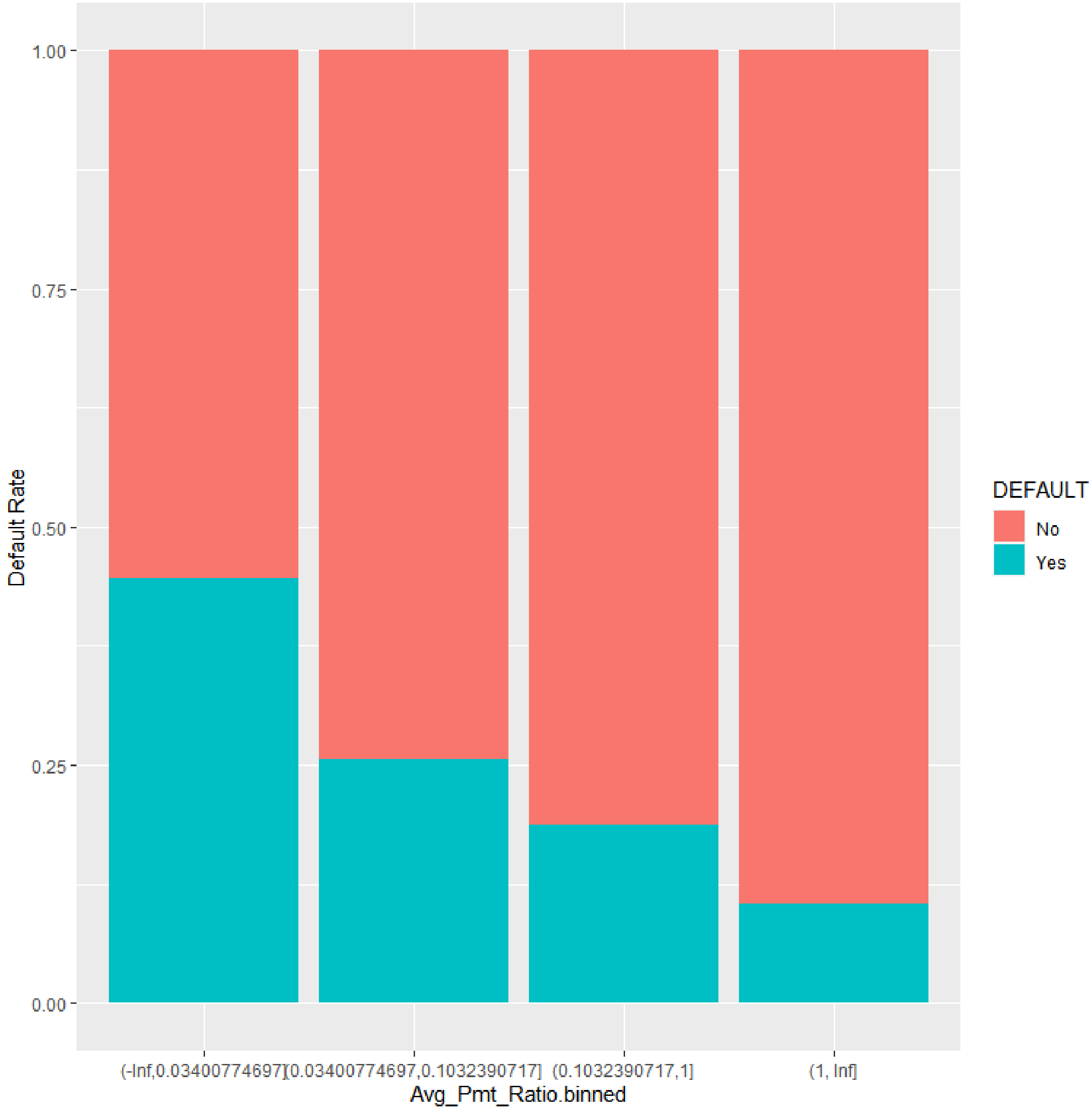
Figure 6: Comparison of Default Rates



Notes:
Individuals who made larger average payments were less likely to default.

Avg_Pmt_Ratio

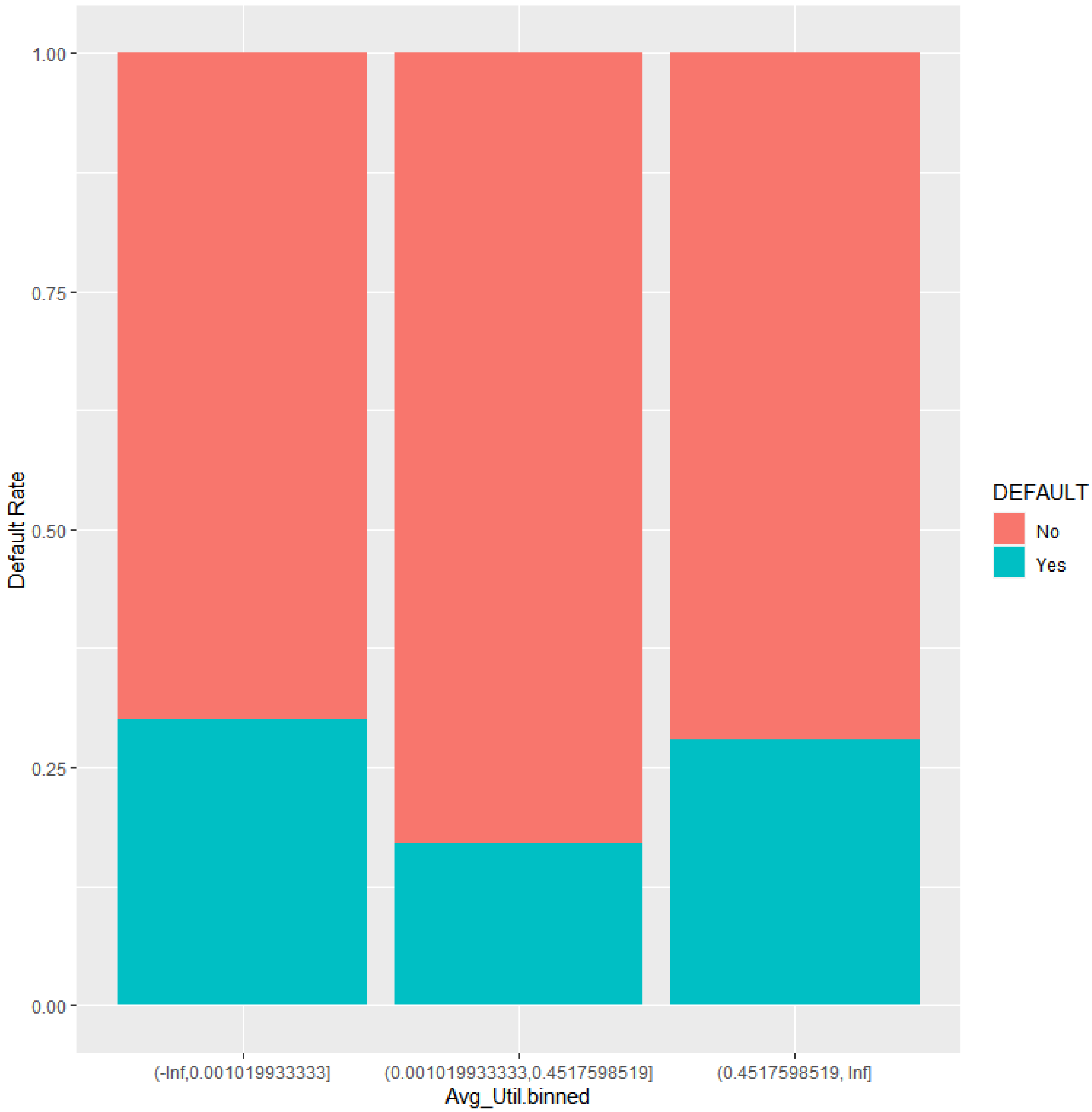
Figure 7: Comparison of Default Rates



Notes:
Individuals with a higher Avg_Pmt_Ratio were less likely to default.

Avg_Util

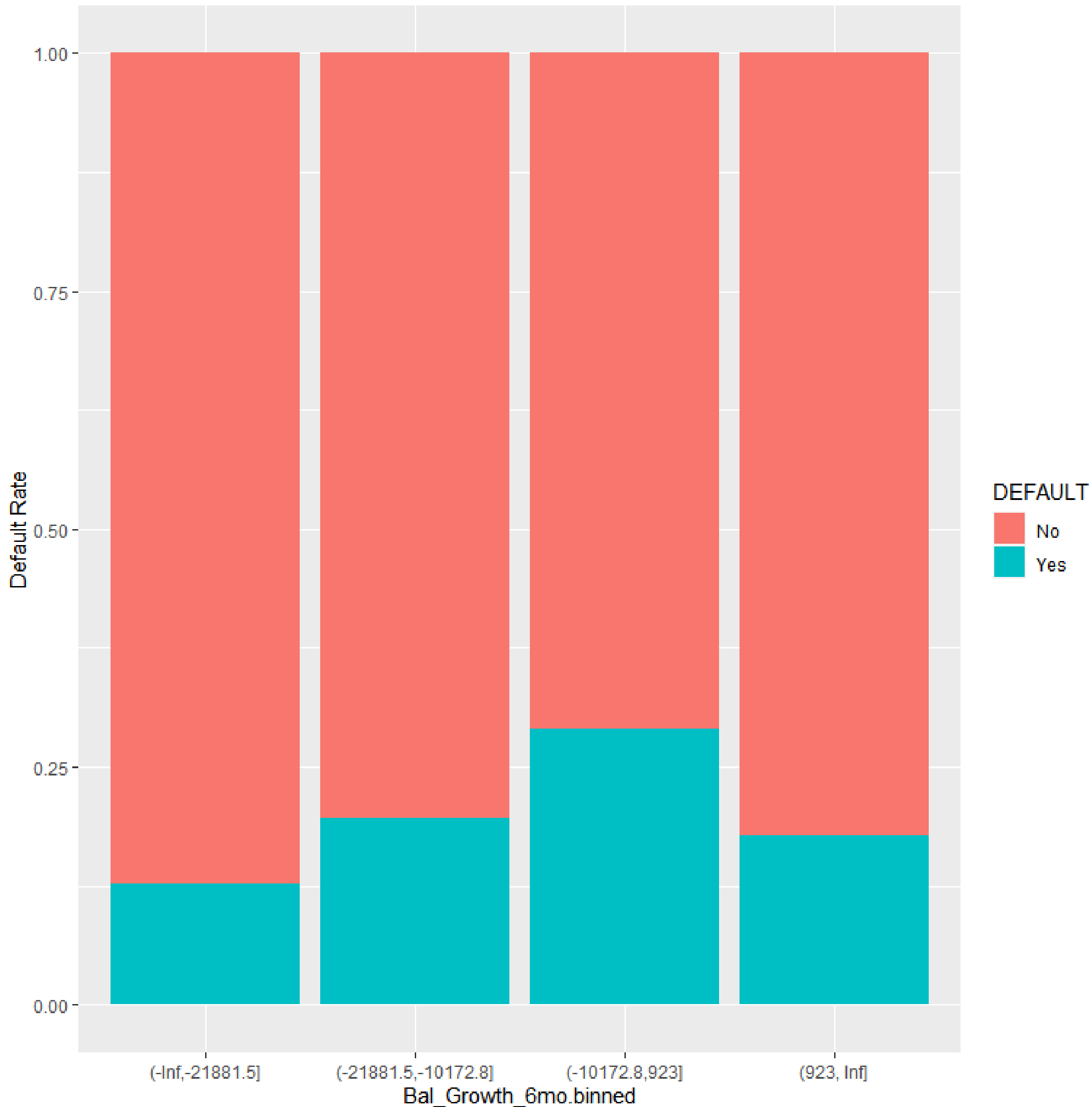
Figure 8: Comparison of Default Rates



Notes:
Individuals with an Average Utilization between 0.001 and 0.452 were less likely to default than those with higher or lower Average Utilization scores. It would appear that a small but consistent utilization of credit was correlated with lower default rates.

Bal_Growth_6mo

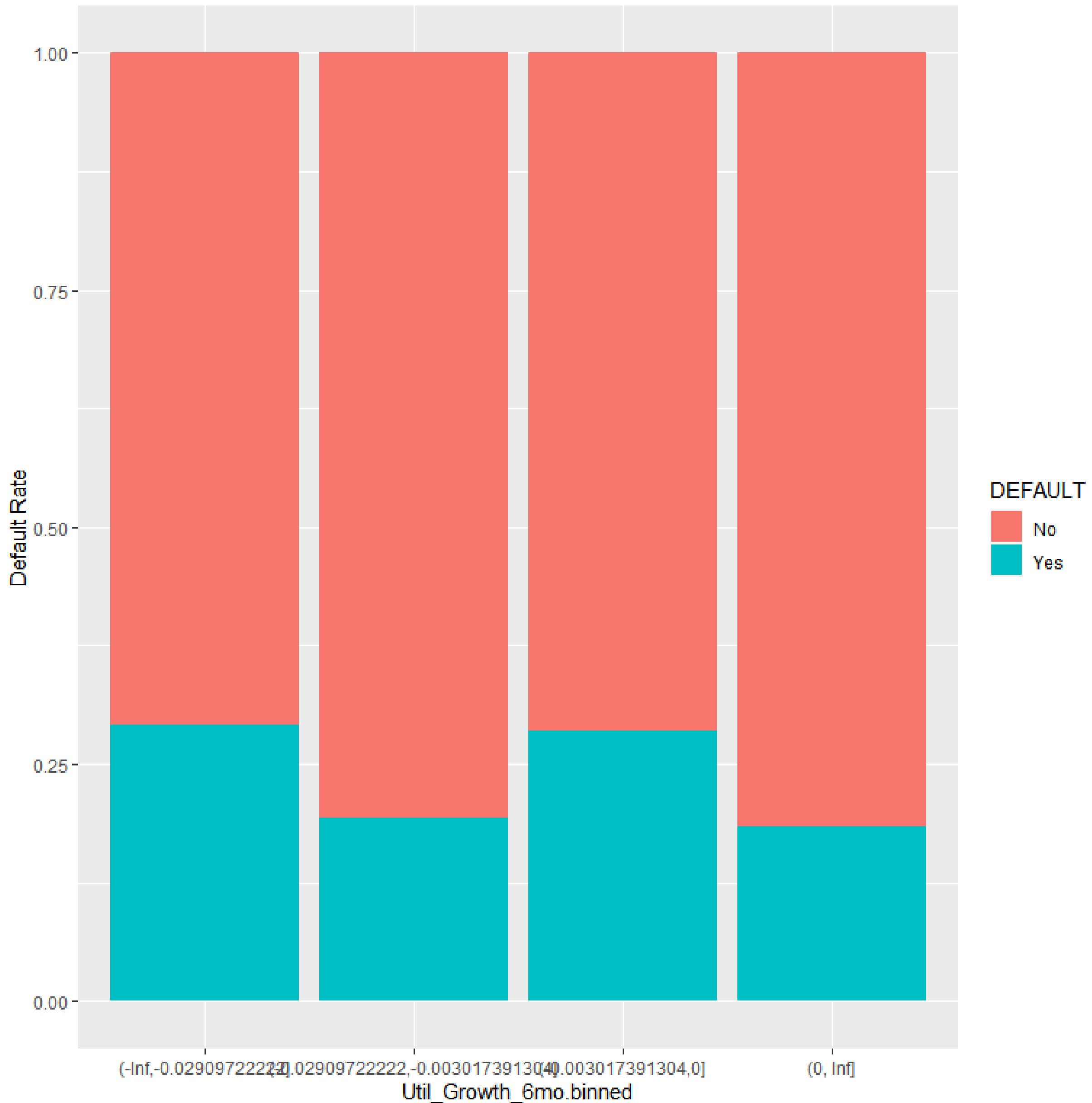
Figure 9: Comparison of Default Rates



Notes:
Individuals with lower 6 Month Balance Growth values were less likely to default.

Util_Growth_6mo

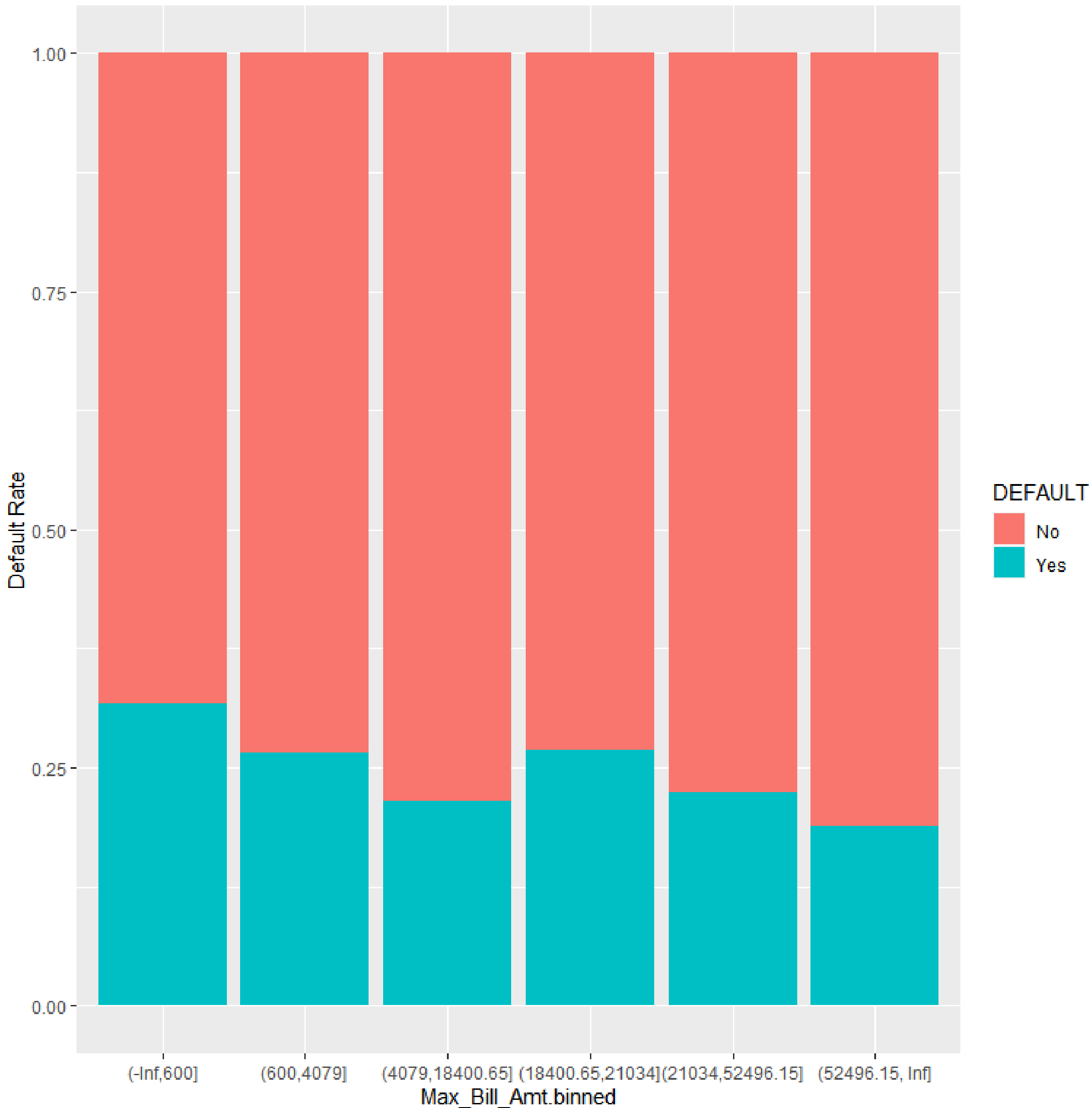
Figure 10: Comparison of Default Rates



Notes:
There did not appear to be much correlation between Utilization Growth and default rates.

Max_Bill_Amt

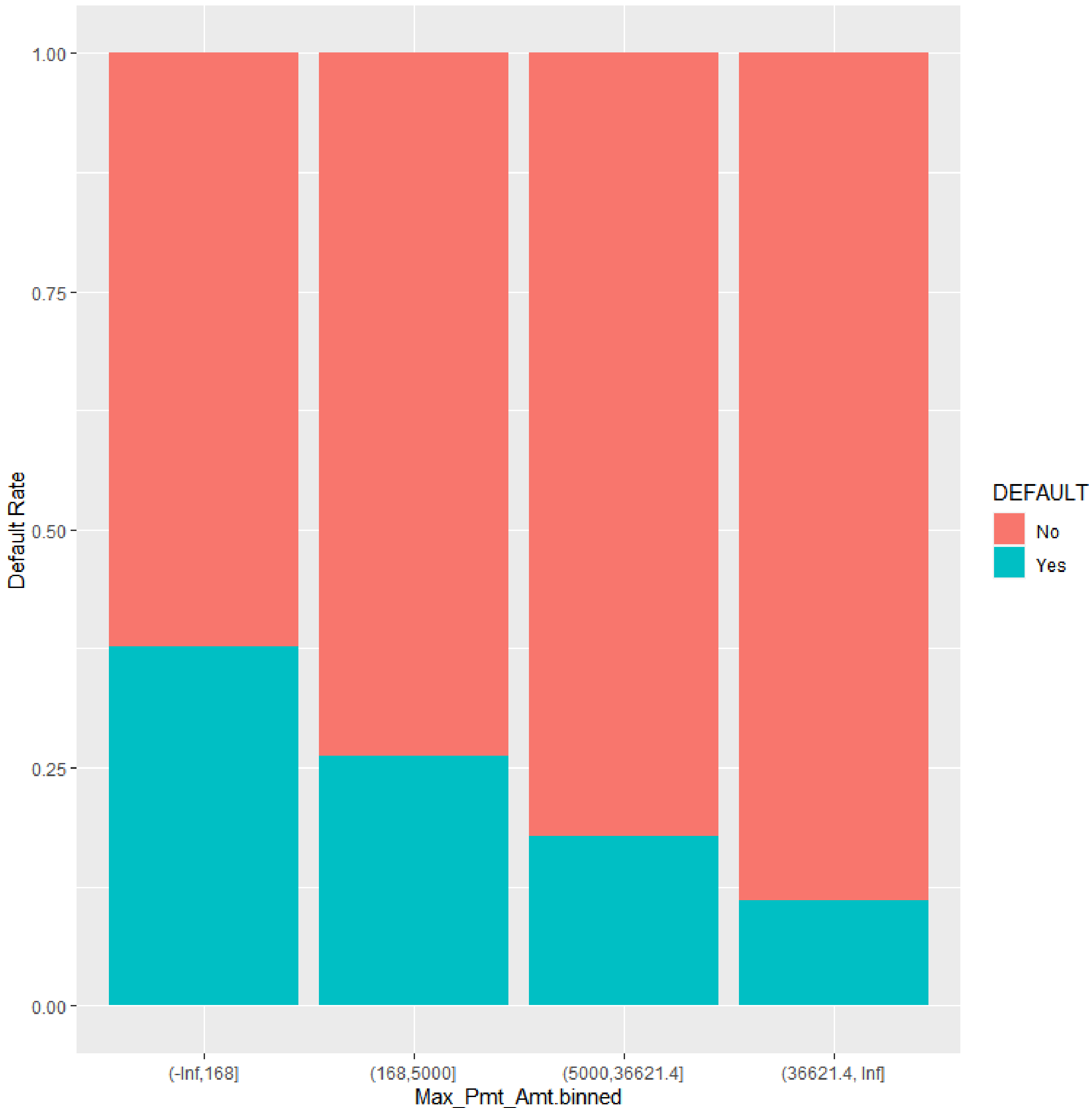
Figure 11: Comparison of Default Rates



Notes:
There did not appear to be much correlation between the Max Bill Amount and default rates.

Max_Pmt_Amt

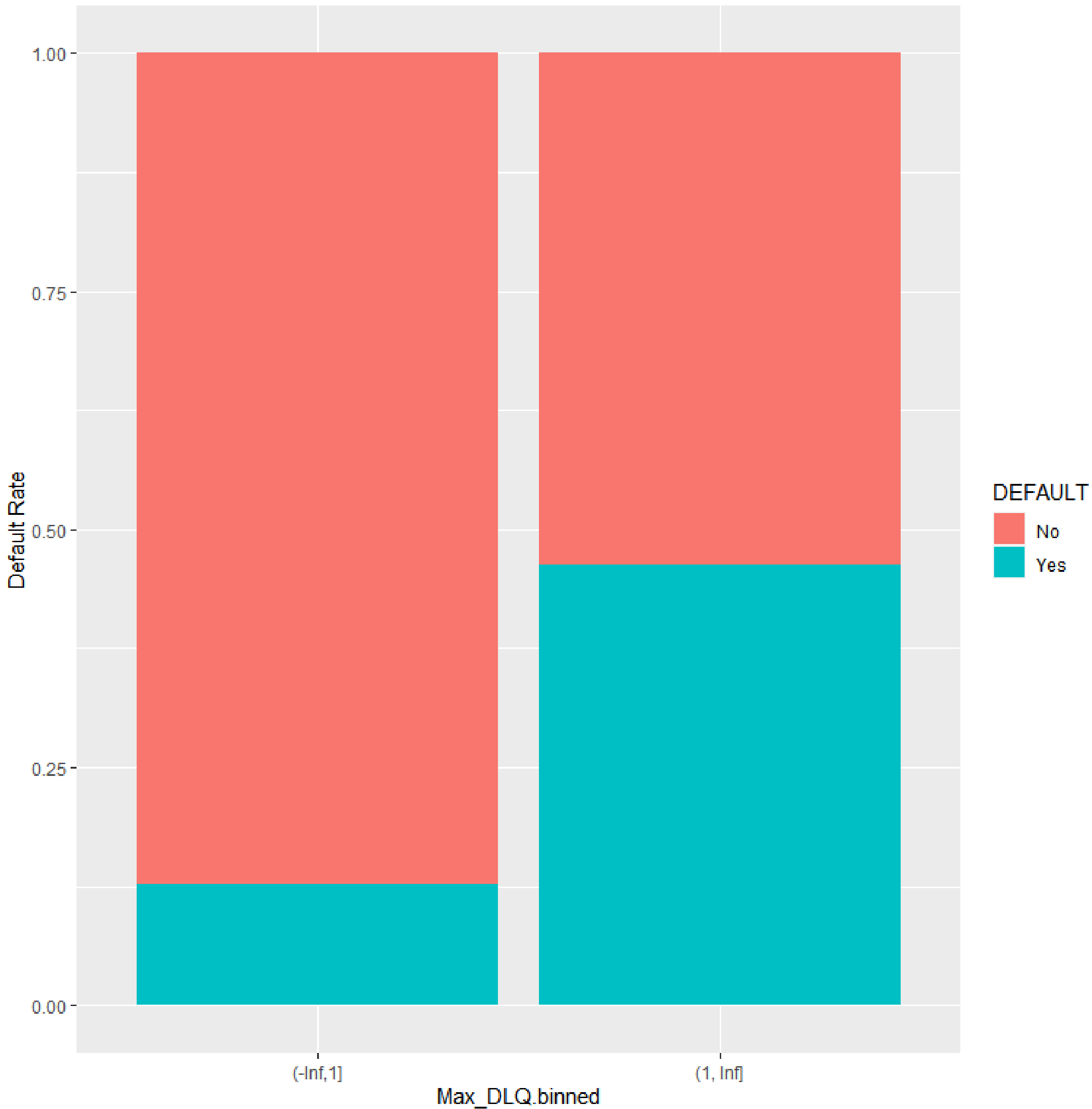
Figure 12: Comparison of Default Rates



Notes:
Individuals with higher Max Payment Amounts were less likely to default.

Max_DLQ

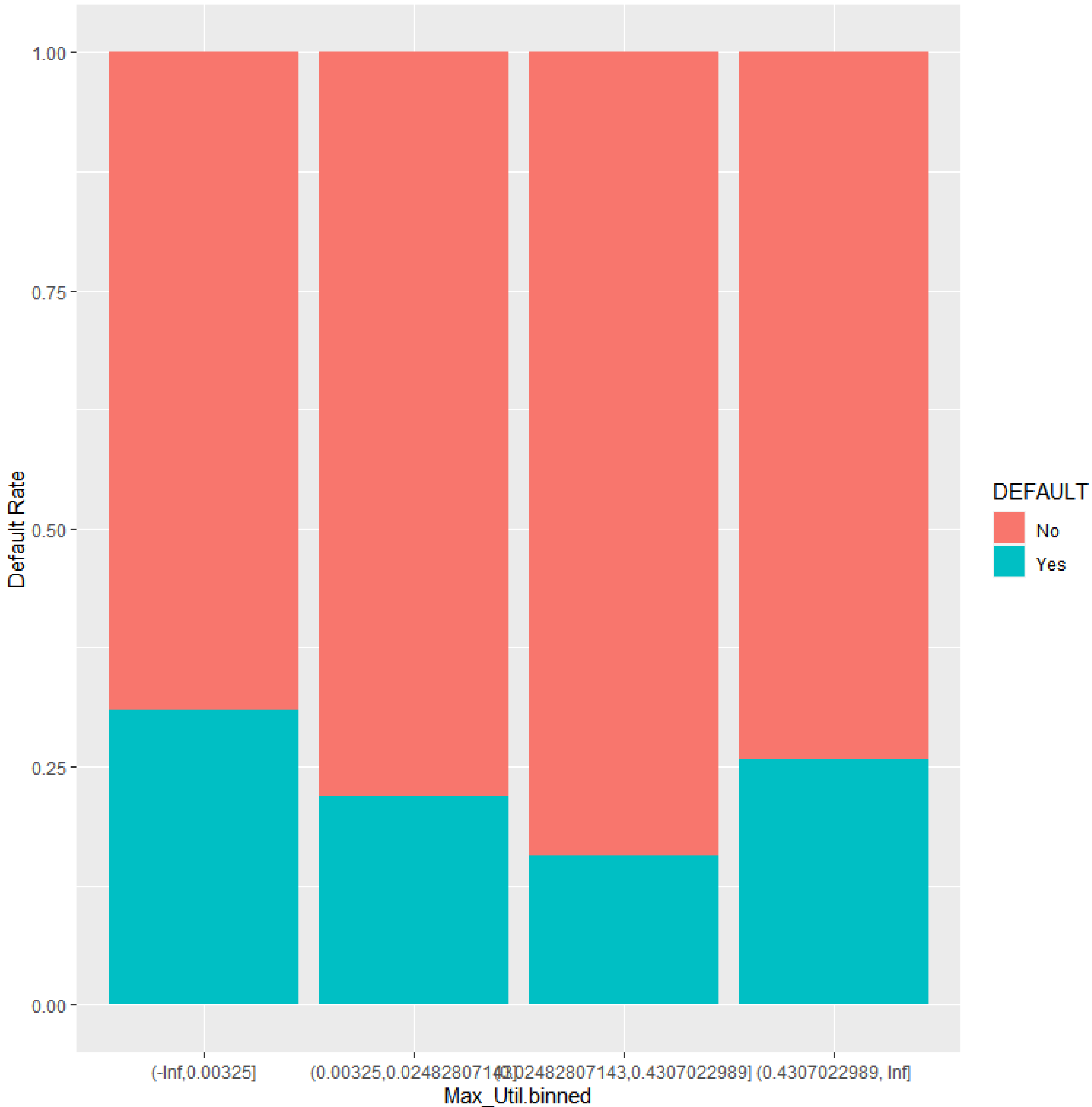
Figure 13: Comparison of Default Rates



Notes:
Individuals with a Max Delinquency greater than 1 month were significantly more likely to default.

Max_Util

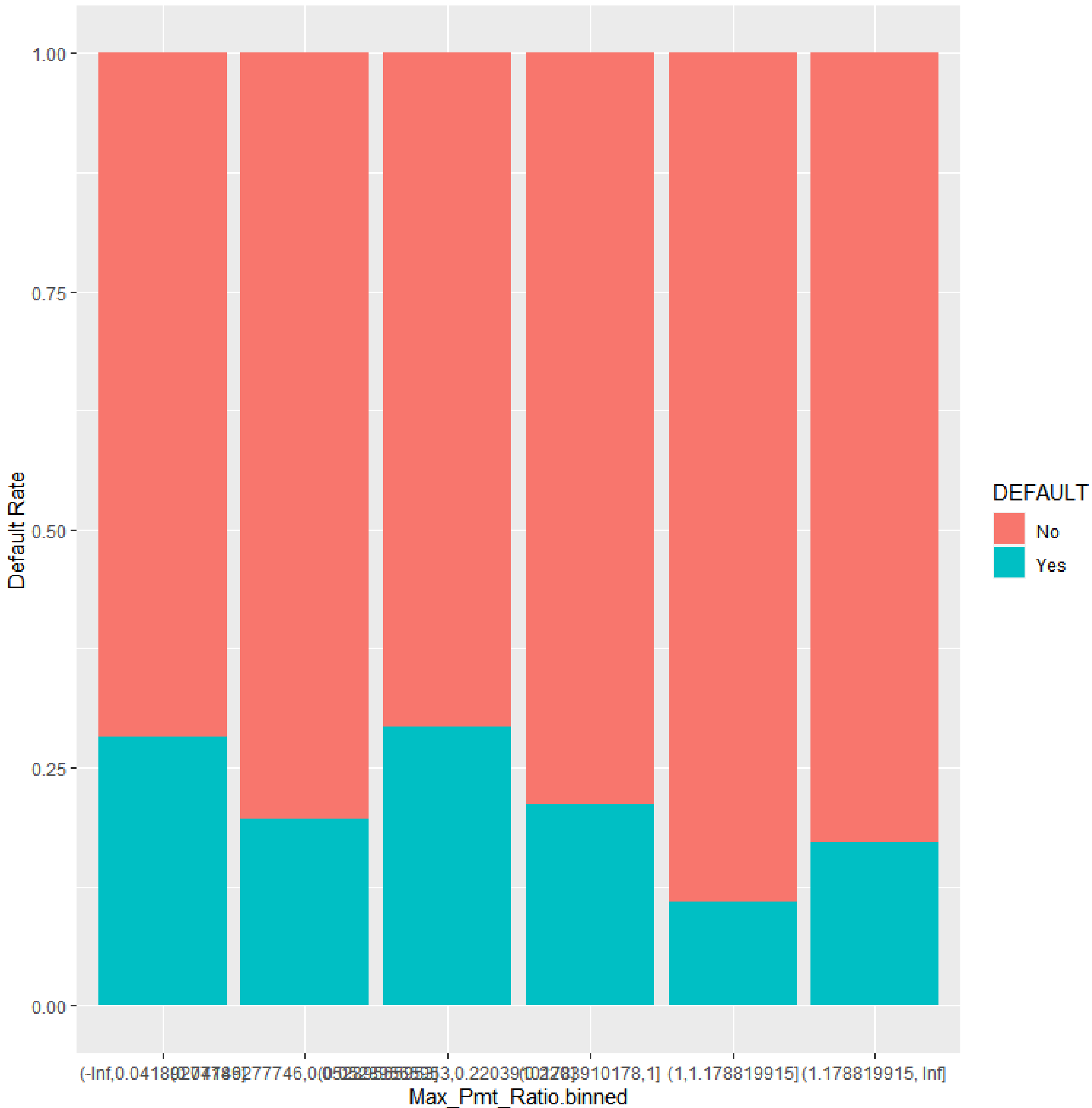
Figure 14: Comparison of Default Rates



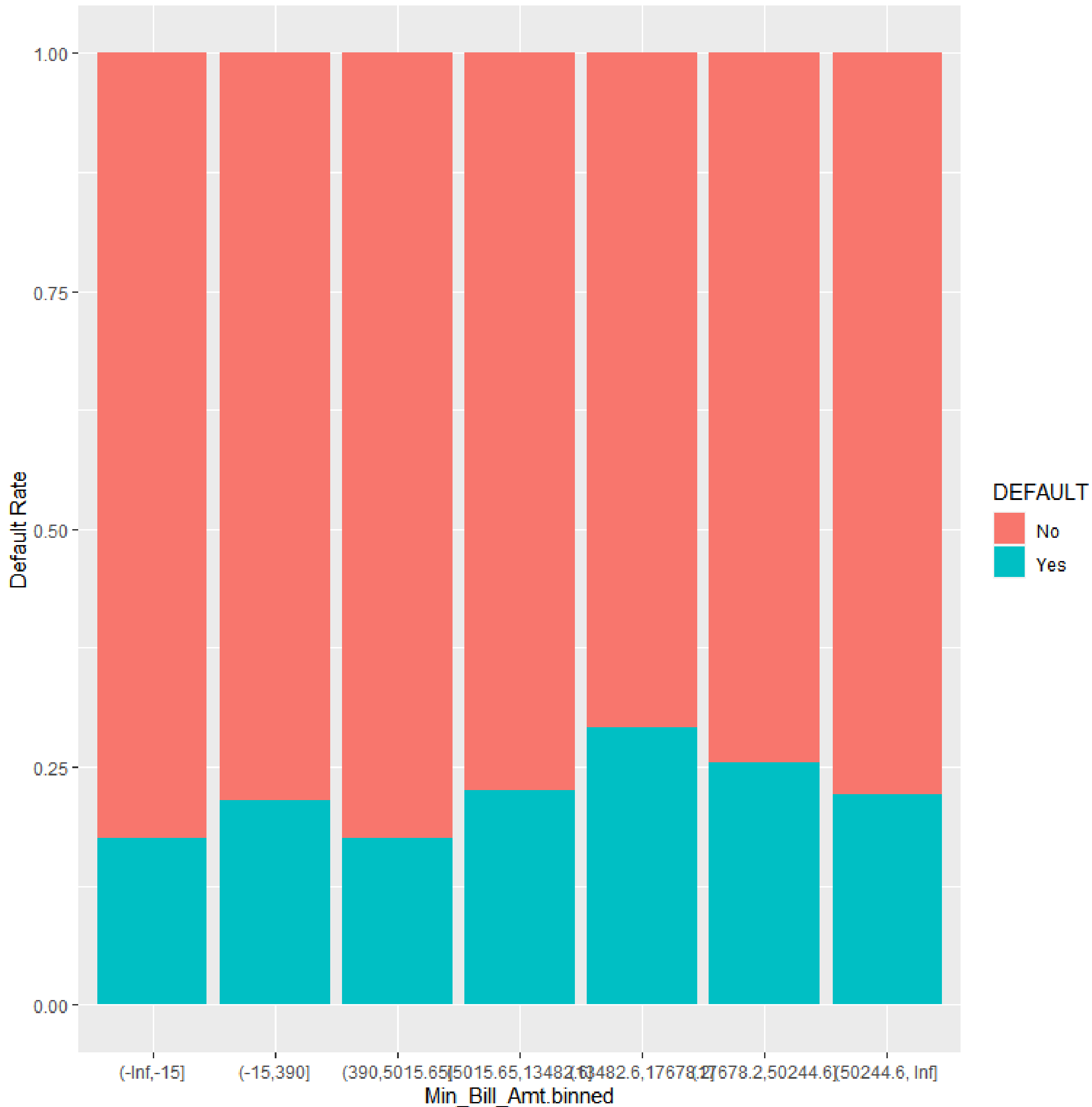
Notes:
Similar to Avg_Util, individuals in middle groups of Max Utilization were less likely to default. It would appear that a small but consistent utilization of credit was correlated with lower default rates.

Max_Pmt_Ratio

Figure 15: Comparison of Default Rates

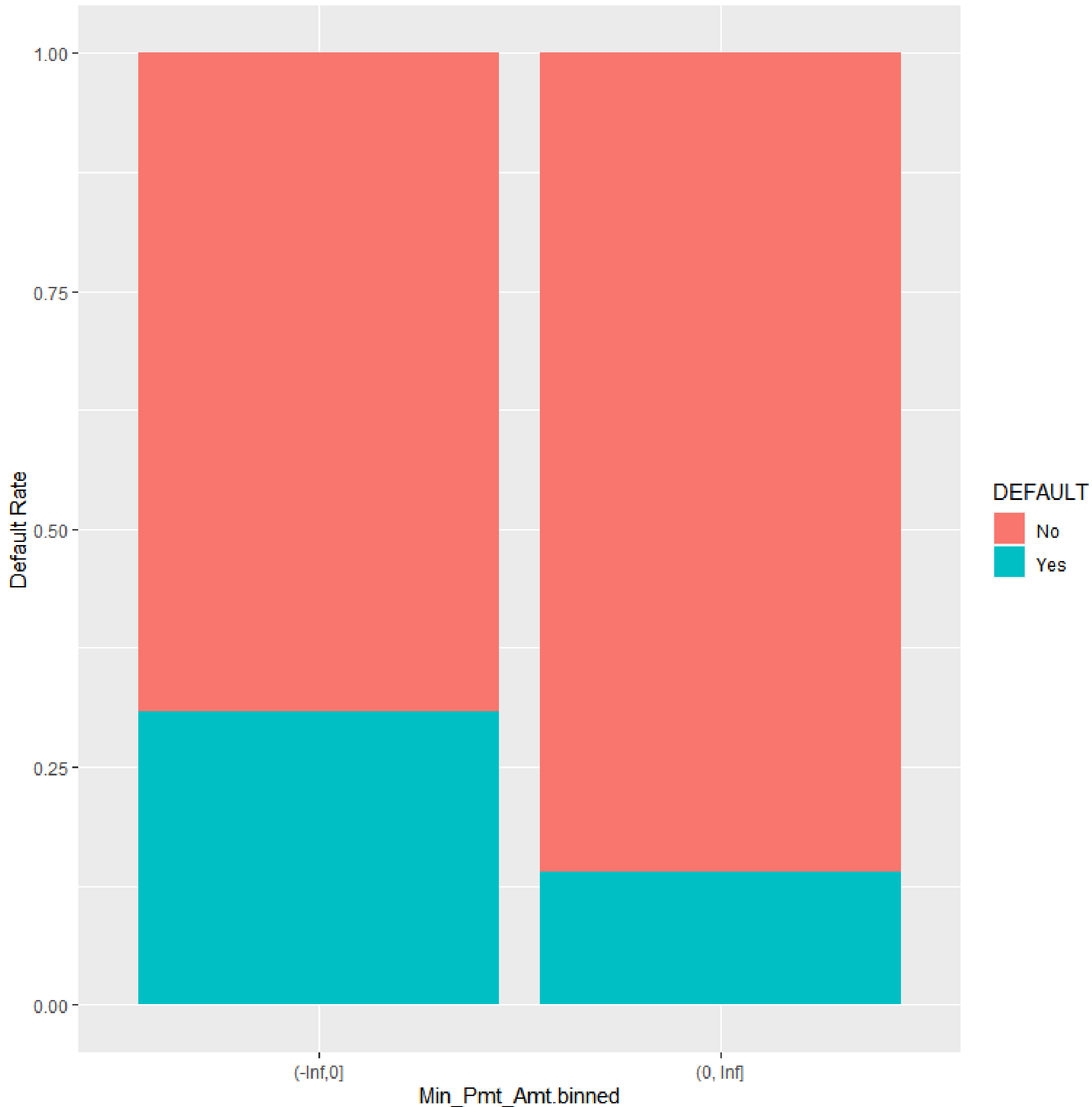


Notes:
There did not appear to be a strong correlation between the Max Payment Ration and the default rate.

Min_Bill_Amt
Figure 16: Comparison of Default Rates

Notes: There did not appear to be a strong correlation between the Min Bill Amount and the default rate.

Min_Pmt_Amt

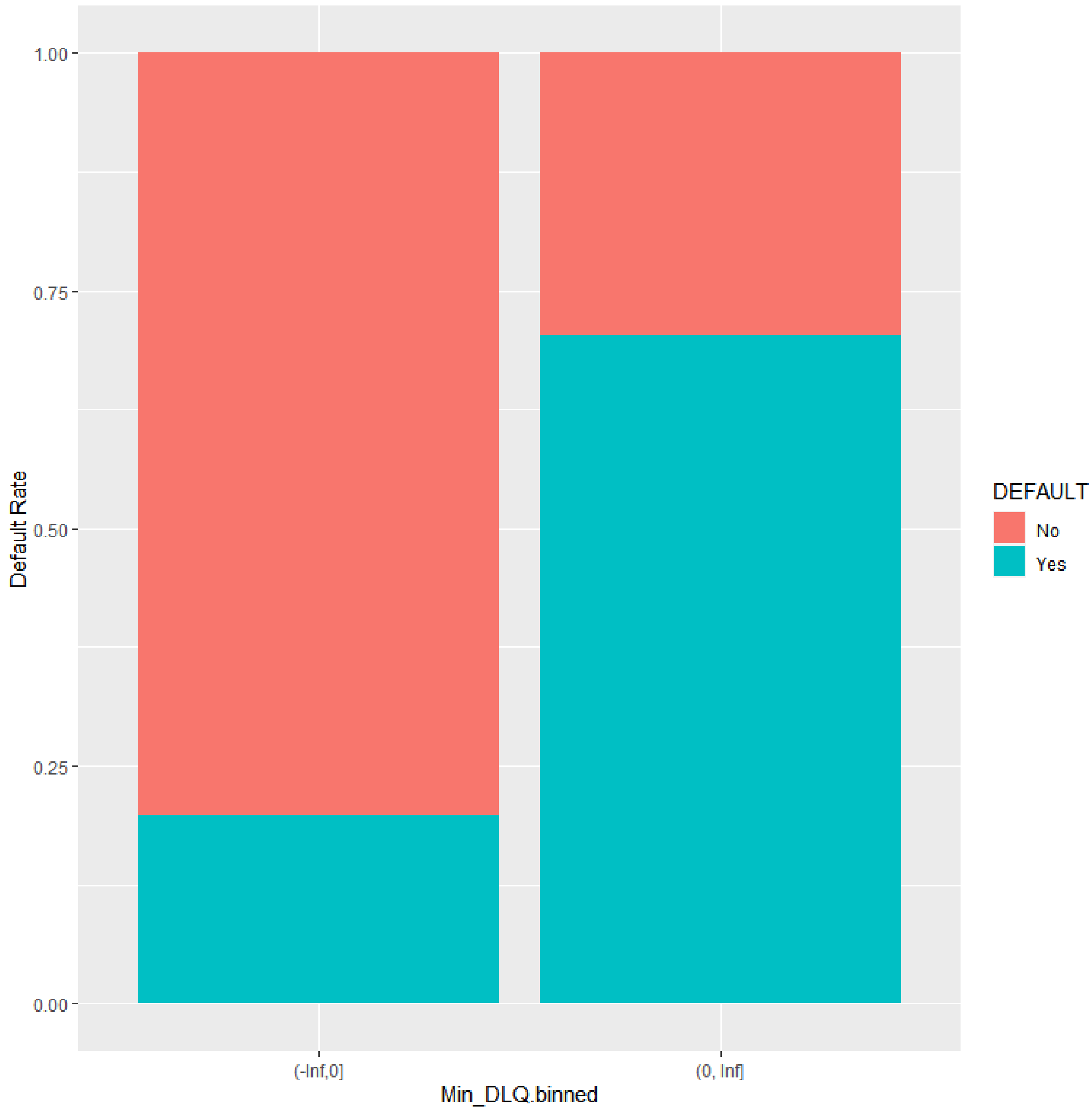
Figure 17: Comparison of Default Rates



Notes:
Individuals with a Min Payment of 0 were more likely to default than those actually made a payment each month.

Min_DLQ

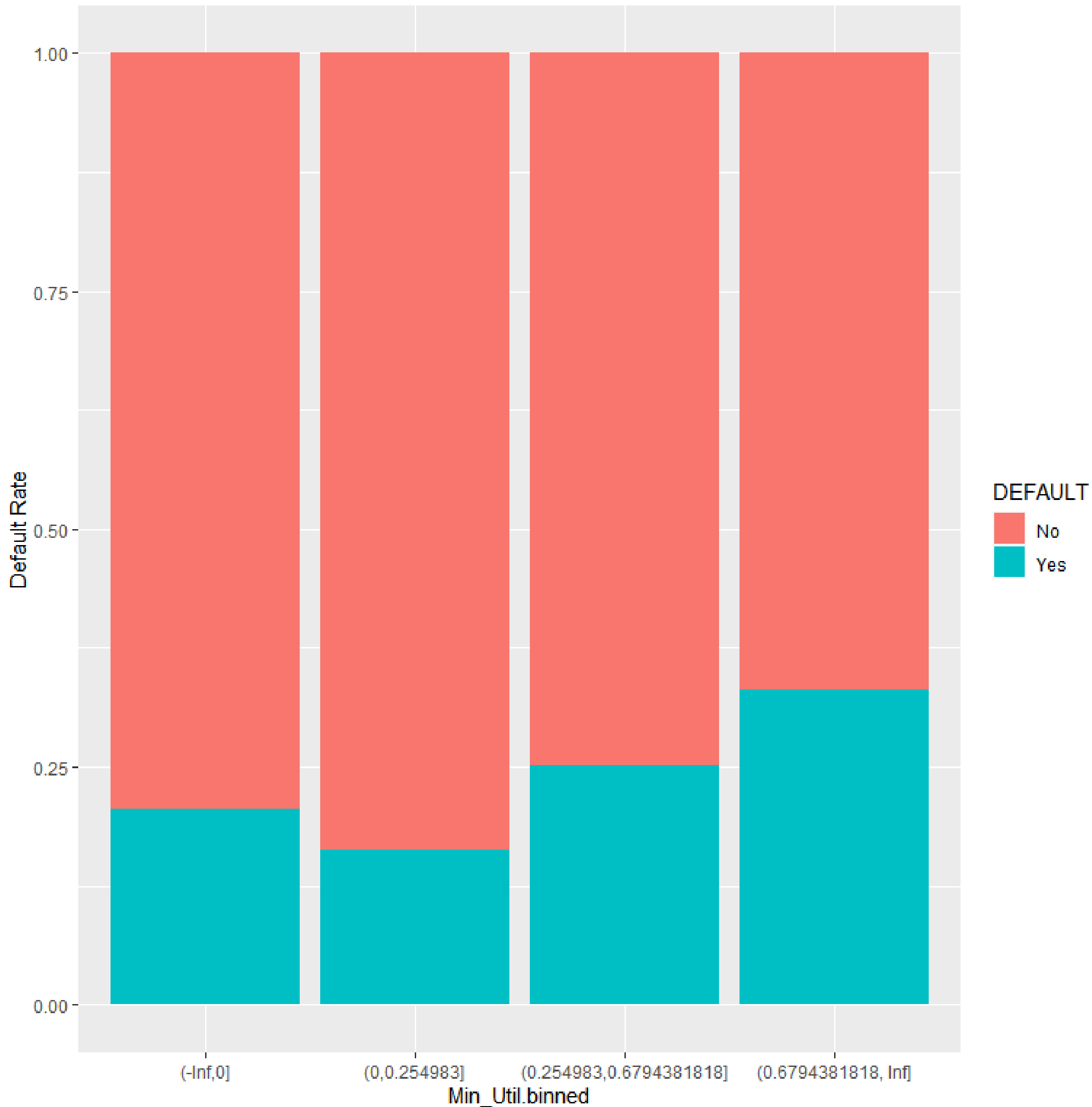
Figure 18: Comparison of Default Rates



Notes:
Individuals who were never delinquent were significantly less likely to default.

Min_Util

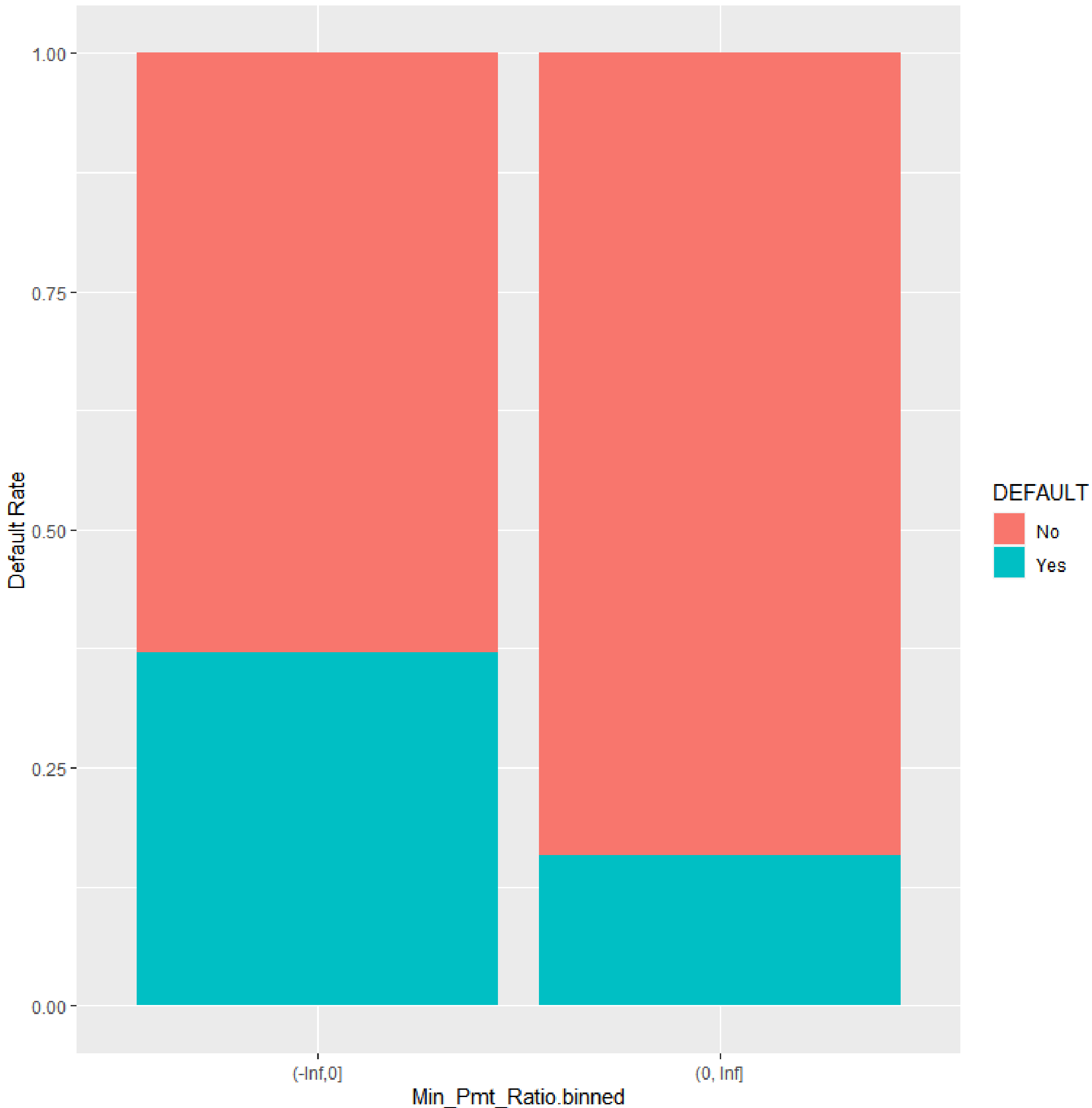
Figure 19: Comparison of Default Rates



Notes:
There did not appear to be a strong correlation between the Min Utilization and default rates. Perhaps individuals with a higher Min Utilization were slightly more likely to default.

Min_Pmt_Ratio

Figure 20: Comparison of Default Rates

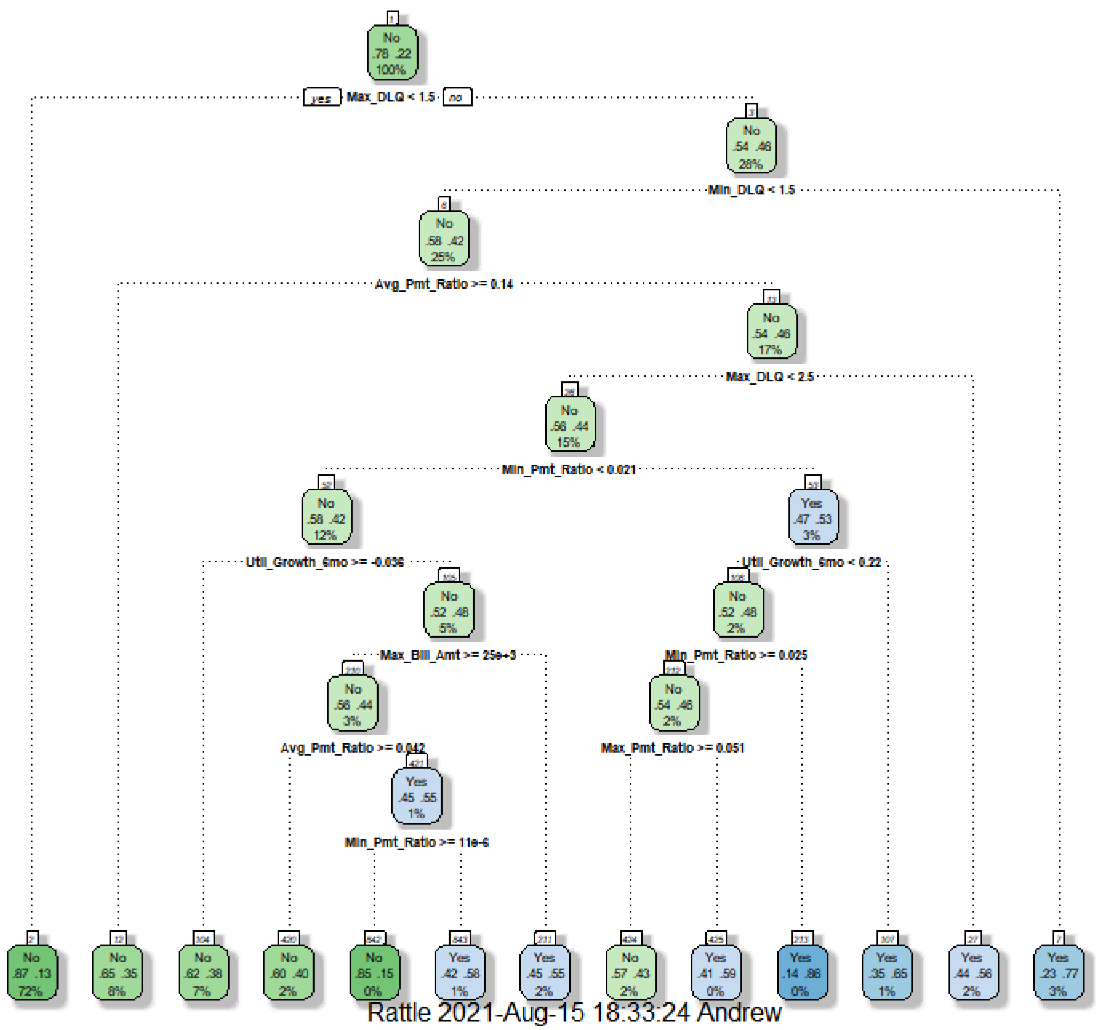


Notes:
Individuals with a Min Payment Ratio of 0 were more likely to default. This observation was consistent with Min Payment Amount, where individuals who did not make a payment were more likely to default.

4b. Decision Tree EDA

A decision tree was fit to the data in order to determine which features could be most insightful for making predictions. Unfortunately, a complete decision tree would be too complicated and visually taxing for this report, so a complexity parameter of 0.002 implemented. Figure 21 below shows the dendrogram. The Maximum and Minimum Delinquencies appear to have played a critical role in predicting whether an individual will default. Also of importance were Payment Ratios and the Utilization Growth Over 6 Months.

Figure 21



4b. OneR EDA

OneR was used to fit a simple model to the data. Table 6 shows accuracies for a model based on each feature. Min_DLQ was chosen to find build the model. Ranked just behind were Max_DLQ and Bal_Growth_6mo. In both the decision tree and OneR models, delinquency was found to be a key predictor in determining whether an individual will default.

Table 6

Attribute	Accuracy
1 * Min_DLQ	79.7%
2 Max_DLQ	78.26%
3 Bal_Growth_6mo	77.89%
4 SEX	77.88%
4 EDUCATION	77.88%
4 MARRIAGE	77.88%
4 AGE	77.88%
4 Avg_Bill_Amt	77.88%
4 Avg_Pmt_Amt	77.88%
4 Avg_Pmt_Ratio	77.88%
4 Avg_Util	77.88%
4 Util_Growth_6mo	77.88%
4 Max_Bill_Amt	77.88%
4 Max_Pmt_Amt	77.88%
4 Max_Util	77.88%
4 Max_Pmt_Ratio	77.88%
4 Min_Bill_Amt	77.88%
4 Min_Pmt_Amt	77.88%
4 Min_Util	77.88%
4 Min_Pmt_Ratio	77.88%

Table 7 denotes the classification rules employed by the OneR model. If the Minimum Delinquency is 0 months, then the model predicts that the person will not default. However, if the person with a Minimum Delinquency of 1 or greater, the model predicts that they will default.

Table 7

Rules:
If Min_DLQ = 0 then DEFAULT = No
If Min_DLQ = 1 then DEFAULT = Yes
If Min_DLQ = 2 then DEFAULT = Yes
If Min_DLQ = 3 then DEFAULT = Yes
If Min_DLQ = 4 then DEFAULT = Yes

Overall, the accuracy was 79.7%, just slightly higher than if everyone was predicted not to default. The model demonstrated a true positive or sensitivity rate 14.2%, meaning it failed to identify the vast majority of individuals who will default. On the other end, it only demonstrated a false positive rate of 1.7%. This analysis forbode that the identification of true positives would be challenge with future models.

5. Predictive Modeling: Methods and Results

5.a Random Forest

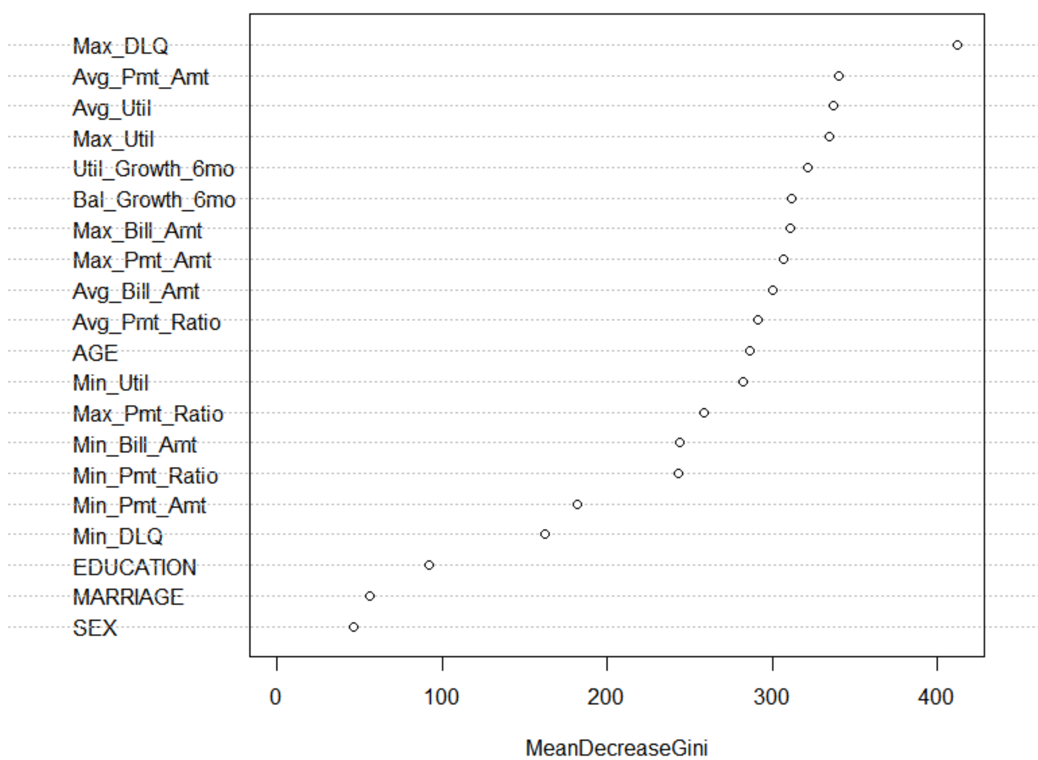
A random forest with 500 trees was fit to the Train Dataset and evaluated on the Test Dataset. Tables # below shows the True Positive, False Positive, and Accuracy metrics. There was no issue with overfitting, as the random forest model demonstrated similar accuracies between the train and test datasets. However, as predicted by the EDA, the model struggled to accurately individuals who would actually default. It only picked up on 27.9% of the true positives when given the test data.

Table 8

Metric	**Train Data**	**Test Data**
True positive rate (%)	27.8	27.9
False positive rate (%)	05.9	06.3
Accuracy (%)	79.1	79.7

Figure shows the mean decrease in Gini coefficient for each variable. Similar to the models used in the EDA, Max_DLQ had the largest impact on the random forest.

Figure 21



5.b Gradient Boosting

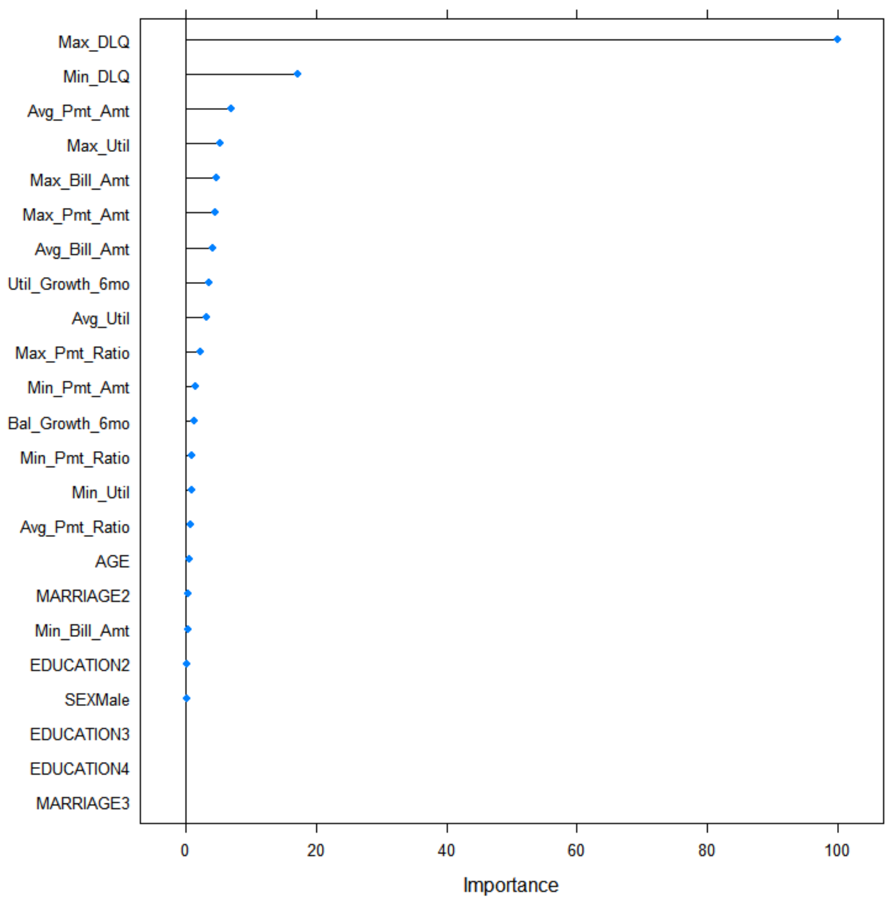
Next, the GBM package was used for classification. As noted in Table 9, the model struggled identify true positive positives, only identifying 19.1% of the cases where the individual would default.

Table 9

Metric	Train Data	Test Data
True positive rate (%)	20.0	19.1
False positive rate (%)	02.6	03.0
Accuracy (%)	79.9	80.4

Figure 22 shows the importance of each variable considered the GBM. Again, Max_DLQ provided the most insight, followed my Min_DLQ and Avg_Pmt_Amt.

Figure 22



5.C Logistic Regression with Variable Selection

Backward stepwise regression was used to build a logistic regression model. Table 10 shows the variables that were retained and their respective coefficients and p-values.

Table 10: Logistic Regression Coefficients and p-Values

	Dependent variable:
	DEFAULT.Yes
SEX.Male	0.13^{***}
	(0.04)
EDUCATION.4	-0.88^{***}
	(0.24)
MARRIAGE.2	-0.19^{***}
	(0.04)
MARRIAGE.3	-0.29
	(0.19)
Avg_Bill_Amt	0.0000^{***}
	(0.0000)
Avg_Pmt_Amt	-0.0001^{***}
	(0.0000)
Bal_Growth_6mo	0.0000^{*}
	(0.0000)
Max_Bill_Amt	-0.0000^{**}
	(0.0000)
Max_Pmt_Amt	0.0000^{***}
	(0.0000)
Max_DLQ	0.60^{***}
	(0.02)
Max_Util	0.23^{***}
	(0.06)
Min_Pmt_Amt	-0.0001^{***}
	(0.0000)
Min_DLQ	0.62^{***}
	(0.06)
Constant	-1.70^{***}
	(0.05)

Table 11 contains the accuracy metrics for the regression model.

Table 11

Metric	**Train Data**	**Test Data**
True positive rate (%)	19.3	19.0
False positive rate (%)	02.7	03.1
Accuracy (%)	79.7	80.4

5.D Neural Network

Finally, a neural network was fit to the train dataset. The model contained 2 dense layers with 2,000 nodes at each layer. Dropout was also utilization to prevent overfitting. The model summary is presented below with Figure 23.

Figure 23

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 2000)              54000
_________________________________________________________________
dropout (Dropout)            (None, 2000)              0
_________________________________________________________________
dense_1 (Dense)              (None, 2000)              4002000
_________________________________________________________________
dropout_1 (Dropout)          (None, 2000)              0
_________________________________________________________________
output_layer (Dense)         (None, 2)                 4002
=================================================================
Total params: 4,060,002
Trainable params: 4,060,002
Non-trainable params: 0
```

Similar to the other machine learning models, the neural network obtained a test accuracy of 80.2%. The model continued to struggle with a low true positive rate at 18.6%.

Table 12

Metric	**Train Data**	**Test Data**
True positive rate (%)	17.2	18.6
False positive rate (%)	02.2	3.1
Accuracy (%)	79.5	80.2

To measure the impact of each feature, the permutation importance was run. Table 13 shows the importance of each variable and Figure 24 represents these values on a bar plot. Max_DLQ, Min_DLQ, and Min_Pmt_Amt had the greatest impact on the model.

Figure 24

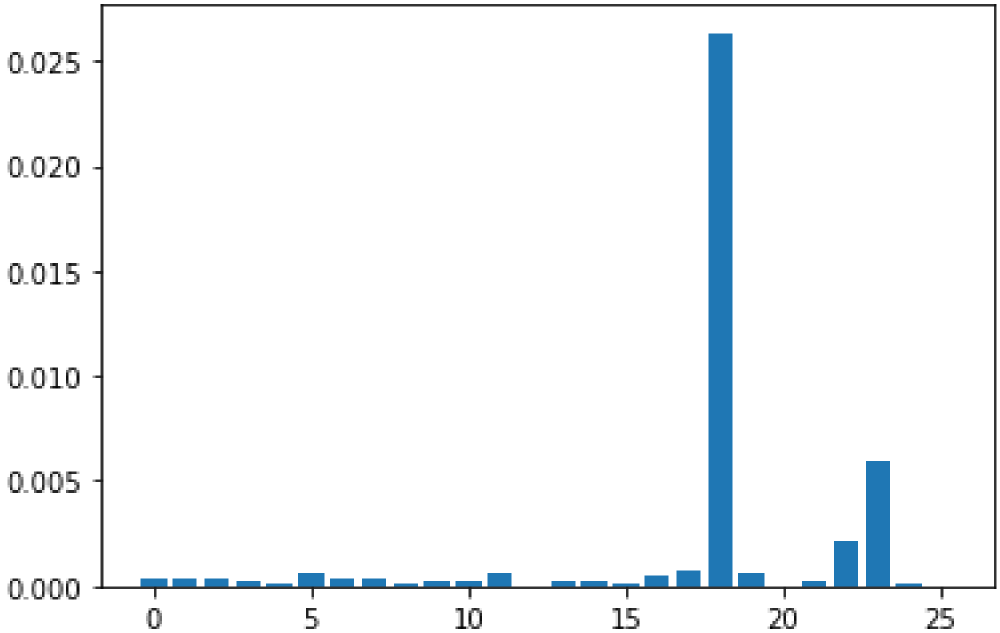


Table 13

```
Feature: 0, Score: 0.00034 SEX.Female
Feature: 1, Score: 0.00029 SEX.Male
Feature: 2, Score: 0.00030 EDUCATION.1
Feature: 3, Score: 0.00025 EDUCATION.2
Feature: 4, Score: 0.00014 EDUCATION.3
Feature: 5, Score: 0.00059 EDUCATION.4
Feature: 6, Score: 0.00040 MARRIAGE.1
Feature: 7, Score: 0.00032 MARRIAGE.2
Feature: 8, Score: 0.00016 MARRIAGE.3
Feature: 9, Score: 0.00022 AGE
Feature: 10, Score: 0.00022 Avg_Bill_Amt
Feature: 11, Score: 0.00065 Avg_Pmt_Amt
Feature: 12, Score: 0.00003 Avg_Pmt_Ratio
Feature: 13, Score: 0.00025 Avg_Util
Feature: 14, Score: 0.00024 Bal_Growth_6mo
Feature: 15, Score: 0.00005 Util_Growth_6mo
Feature: 16, Score: 0.00042 Max_Bill_Amt
Feature: 17, Score: 0.00068 Max_Pmt_Amt
Feature: 18, Score: 0.02635 Max_DLQ
Feature: 19, Score: 0.00065 Max_Util
Feature: 20, Score: 0.00002 Max_Pmt_Ratio
Feature: 21, Score: 0.00022 Min_Bill_Amt
Feature: 22, Score: 0.00208 Min_Pmt_Amt
Feature: 23, Score: 0.00592 Min_DLQ
Feature: 24, Score: 0.00012 Min_Util
Feature: 25, Score: 0.00002 Min_Pmt_Ratio
```

6. Comparison of Results

Table 14 lists the test performance metrics for each model employed in this study.
Table 14

Model	True positive rate (%)	False positive rate (%)	Accuracy (%)
Random Forest	27.9	06.3	79.7
GBM	19.1	03.0	80.4
Logistic Regression	19.0	03.1	80.4
Neural Network	18.6	03.1	80.2

Despite their diverse methodologies, all of the models presented similar results. Each one demonstrated a classification accuracy around 80%. In particular, the GBM, Logistic Regression, and Neural Network models had strikingly similar classification metrics with true positive rates around 19% and false positive rates around 3%. This could possibly be due to the fact that they all focus on similar variables. Max_DLQ or Min_DLQ followed by the Avg_Pmt_Amt or Min_Pmt_Amt were the most informative features for each model.

The Random Forest model was the most successful at identifying individuals who would default, with a true positive rate of 27.9%. However, in the sensitivity/specificity tradeoff, this same model had double the false positive rate of the other models at 6.3%.

7. Conclusion

While the models were not strong predictors of who would default, they did identify some key variables to help make that determination. Delinquency is a major key into default rates, as people who fell behind were more likely to eventually fall into default.

All the models employed in this study demonstrated roughly an 80% classification accuracy for predicting which individuals would default. Due to their low prevalence in the dataset, the models consistently performed at a low true positive rate. This measurement could be improved lowering the threshold probabilities for the GBM and Logistic Regression models. However, this would come at the cost of an increased false positive rate.

8. Appendix

Table 1A: Unclean Variables Summary

Statistic	N	Mean	St. Dev.	Min	P(25)	Median	P(75)	Max
LIMIT_BAL	30,000	167,484.30	129,747.70	10,000	50,000	140,000	240,000	1,000,000
SEX	30,000	1.60	0.49	1	1	2	2	2
EDUCATION	30,000	1.85	0.79	0	1	2	2	6
MARRIAGE	30,000	1.55	0.52	0	1	2	2	3
AGE	30,000	35.49	9.22	21	28	34	41	79
PAY_0	30,000	-0.02	1.12	-2	-1	0	0	8
PAY_2	30,000	-0.13	1.20	-2	-1	0	0	8
PAY_3	30,000	-0.17	1.20	-2	-1	0	0	8
PAY_4	30,000	-0.22	1.17	-2	-1	0	0	8

PAY_5	30,000	-0.27	1.13	-2	-1	0	0	8
PAY_6	30,000	-0.29	1.15	-2	-1	0	0	8
BILL_AMT1	30,000	51,223.33	73,635.86	-165,580	3,558.8	22,381.5	67,091	964,511
BILL_AMT2	30,000	49,179.08	71,173.77	-69,777	2,984.8	21,200	64,006.2	983,931
BILL_AMT3	30,000	47,013.15	69,349.39	-157,264	2,666.2	20,088.5	60,164.8	1,664,089
BILL_AMT4	30,000	43,262.95	64,332.86	-170,000	2,326.8	19,052	54,506	891,586
BILL_AMT5	30,000	40,311.40	60,797.16	-81,334	1,763	18,104.5	50,190.5	927,171
BILL_AMT6	30,000	38,871.76	59,554.11	-339,603	1,256	17,071	49,198.2	961,664
PAY_AMT1	30,000	5,663.58	16,563.28	0	1,000	2,100	5,006	873,552
PAY_AMT2	30,000	5,921.16	23,040.87	0	833	2,009	5,000	1,684,259
PAY_AMT3	30,000	5,225.68	17,606.96	0	390	1,800	4,505	896,040
PAY_AMT4	30,000	4,826.08	15,666.16	0	296	1,500	4,013.2	621,000
PAY_AMT5	30,000	4,799.39	15,278.31	0	252.5	1,500	4,031.5	426,529
PAY_AMT6	30,000	5,215.50	17,777.47	0	117.8	1,500	4,000	528,666
DEFAULT	30,000	0.22	0.42	0	0	0	0	1