

MBA 547 Case Report, Homework 3

Topic: Loan app and Openness Data Set

Due Date: November 08, 2024

Submitted by: Ofuka Abung, Sadaf Vora, and Samyak Shah

Team Lead – Sadaf Vora, Data Analysts and Script writers – Ofuka Abung and Samyak Shah

Wooldridge Data - Loan app & Openness

Executive Summary

This report looks at two datasets to understand the factors behind loan approval and the relationship between trade openness and inflation. In Part 1, we analyze the Loan Application dataset, which has 1,986 records with important information like loan amounts, income, marital status, and employment history. We first ran some descriptive statistics, showing that loan amounts have an average of \$143,250, and income ranges from \$0 to \$972,000. We then used three different models - Linear Probability Model (LPM), Logit Model, and Probit Model - to study how these factors influence loan approval. The results show that being married increases the chance of loan approval, while higher housing expenses and bad credit history (public records) reduce the chances. Marginal effects analysis shows that each additional dependent decreases approval chances by 0.009, and being self-employed reduces approval chances by about 3.8 percentage points. The LPM's R-squared value is 0.08544, meaning that only 8.5% of the variation in loan approval can be explained by the factors in the model.

In Part 2, we looked at the Openness dataset to understand how trade openness (imports as a percentage of GDP) and inflation are related. The descriptive analysis showed that inflation has an average of 17.26%, and imports as a percentage of GDP have a mean of 37.08%. We then used a Simultaneous Equation Model (SEM) with the 2SLS method to estimate the relationship between imports and inflation, treating both as endogenous variables and using per capita income as an instrument. The results showed that while there is a relationship, it is not strong enough to be statistically significant. The R-squared values for the SEM were low, with the inflation equation showing an R-squared value of -1.47 and the imports equation showing 0.0045. This suggests that the model does not explain much of the variation in the data, and other factors might be influencing inflation and trade openness that are not captured by this model.

These findings suggest that further analysis with additional variables or better instruments could help us understand the relationship between trade openness and inflation clearly. The next sections will provide

more details on the modeling process, the assumptions used, and how these results fit into the broader economic context.

Introduction

In today's global economy, understanding the factors that influence financial decisions is crucial. This case report focuses on two important datasets: the Loan Application dataset and the Openness dataset. The goal of this analysis is to explore how different variables impact the outcomes in each dataset.

The Loan Application dataset provides insights into the mortgage lending process. It encompasses various factors that can influence the approval or rejection of loan applications. By studying this dataset, we aim to uncover the underlying trends and biases that may exist in mortgage lending decisions.

On the other hand, the Openness dataset examines the relationship between a country's openness to trade and its inflation rate, along with other economic indicators. By analyzing this data, we can gain insights into how imports and exports affect inflation and overall economic performance.

Data

Part One – LoanApp (LDV)

The Loan Application dataset includes 1989 observations across 59 variables that pertain to mortgage lending. This dataset captures the amount of the loan requested, the type of action taken on loan applications, and the applicant's income, as well as their marital status and any history of bankruptcy. By examining this dataset, we can gain a deeper understanding of the factors contributing to loan approval or rejection, revealing insights into the lending practices of financial institutions.

Preliminary Descriptive Analysis

Before we begin our analysis in the next section of the report, let us review the preliminary descriptive analysis to understand the characteristics of the applicants and the nature of their loan requests. We have provided a shorter summary statistics table below, the complete table can be seen in the Appendix section.

Table 1.1

Variable	Mean	Std. Dev.	Min	Max
Loan Amount (in \$000s)	143.25	80.52	2	980
Applicant Income (in \$000s)	84.68	87.06	0	972
Approval Status (0 =Not Approved, 1 =Approved)	0.88	0.33	0	1
Age of Applicant	36.45	11.12	18	75
Marital Status (1 =Married)	0.66	0.47	0	1

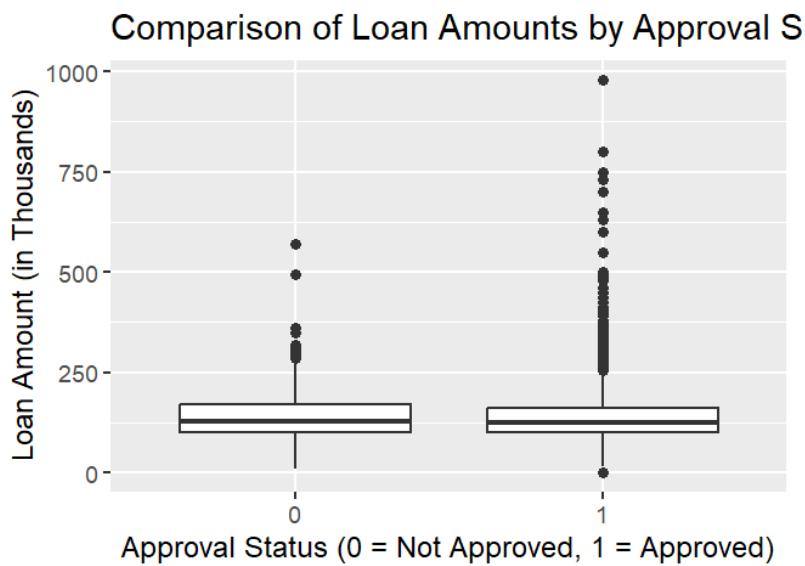
The summary statistics table gives a clear picture of some key variables from the Loanapp dataset, highlighting important details about loan applicants. The average loan amount is \$143,250, with a high standard deviation of \$80,520, indicating significant variation among loan sizes, ranging from \$2,000 to \$980,000. This suggests that while many loans are moderate, some applicants seek much larger amounts.

Applicant income averages at \$84,680, but the standard deviation of \$87,060 shows there are both low and high earners, with incomes stretching from \$0 to \$972,000. Approval status is notable, as about 88% of applicants are approved, reflecting a generally positive outlook for loan seekers.

The average age of applicants is 36.45 years, with ages ranging from 18 to 75, demonstrating a diverse group of applicants. Additionally, approximately 66% of the applicants are married, indicating that this marital status is common among those seeking loans.

Graphical Illustration

Fig 1.1



The box plot provides a visual comparison of loan amounts based on approval status. On the left side, we see the data for applicants who were not approved (0), and on the right, those who were approved (1).

For the approved group, the box plot shows that the median loan amount is significantly higher than that for the not approved group. The median for approved loans is around \$180,000, while for not approved loans, it is only about \$40,000. This difference indicates that those who receive loans tend to apply for much larger amounts.

Additionally, the range of loan amounts for approved applicants extends much further, with a few outliers reaching up to \$1,000,000. In contrast, the not approved group has a much narrower range, with most loans clustered below \$200,000.

The presence of numerous outliers on the approved side suggests that while many applicants request typical amounts, there are also a significant number of high-value loans being approved.

Comparative Averages

Table 1.2

```
# Print comparative averages
> print(comparative_averages)

# A tibble: 2 × 59
  approve occ loanamt action msa suffolk appinc typur unit married dep emp
  <int> <dbl> <dbl>
1     0  1.05  141.   3  1120  0.299  90.4  0.0451  1.26  0.582  0.844  0.205
2     1  1.03  144.   1.03 1120  0.134  83.9  1.74  1.10  0.669  0.761  0.210
```

The comparative averages table reveals the differences in key statistics between applicants who were approved for loans (1) and those who were not approved (0).

For approved applicants, the average loan amount is \$144,000, while for those not approved, the average loan amount is slightly lower at \$141,000. This shows that there is only a small difference in the average loan amounts requested by both groups. However, the variance in loan amounts is more pronounced among those who are approved, as seen in the earlier box plot.

Looking at other variables, we see that the average number of units in the properties is 1.03 for approved applicants compared to 1.05 for those not approved. This indicates a slight preference for single-unit properties among approved applicants. Additionally, the average applicant income is \$83,900 for approved individuals, whereas it is \$90,400 for those not approved. This suggests that applicants who are not approved tend to have higher average incomes, potentially indicating stricter criteria for loan approval based on factors beyond just income.

The data also shows that the average action taken differs significantly, with approved applicants having an average action value of 1.03, indicating successful applications, while the not approved group averages 3. This highlights a clear distinction in outcomes based on approval status.

Part Two – Openness (SEM)

The Openness dataset contains 114 observations across 12 variables, primarily focusing on the relationship between a country's openness to trade (measured by imports as a percentage of GDP) and inflation rates. The dataset includes variables such as per capita income, land area, oil production status, and data quality, among others. By analyzing this dataset, we can explore how factors like economic size, trade openness, and inflation are interrelated, shedding light on the economic dynamics that influence trade policies and inflation trends.

Preliminary Descriptive Analysis

Table 1.3

```
# Perform a preliminary descriptive analysis
> summary(openness) # Get a quick overview of the data
   open          inf         pcinc        land
Min. : 7.40    Min. : 3.600    Min. : 224    Min. : 122
1st Qu.:21.82  1st Qu.: 8.325   1st Qu.: 764   1st Qu.: 25649
Median :32.70  Median :10.650   Median : 2005  Median : 95128
Mean   :37.08  Mean   :17.264   Mean   : 3790  Mean   :298855
3rd Qu.:45.70  3rd Qu.:14.800   3rd Qu.: 5595  3rd Qu.: 257465
Max.  :163.80  Max.  :206.700   Max.  :25646  Max.  :3851809
   oil          good         lpcinc      lland       lopen
Min. :0.0000    Min. :0.0000    Min. : 5.412  Min. : 4.804  Min. :2.001
1st Qu.:0.0000  1st Qu.:0.0000  1st Qu.: 6.638  1st Qu.:10.152 1st Qu.:3.083
Median :0.0000  Median :1.0000   Median : 7.603  Median :11.463  Median :3.487
Mean   :0.0614  Mean   :0.5439   Mean   : 7.659  Mean   :11.126  Mean   :3.441
3rd Qu.:0.0000  3rd Qu.:1.0000  3rd Qu.: 8.627  3rd Qu.:12.458 3rd Qu.:3.822
Max.  :1.0000   Max.  :1.0000   Max.  :10.152  Max.  :15.164  Max.  :5.099
   linf         opendec      linfdec
Min. :1.281    Min. :0.0740   Min. :-3.3242
1st Qu.:2.119  1st Qu.:0.2182  1st Qu.:-2.4859
Median :2.366  Median :0.3270   Median :-2.2396
Mean   :2.502  Mean   :0.3708   Mean   :-2.1032
3rd Qu.:2.695  3rd Qu.:0.4570  3rd Qu.:-1.9106
Max.  :5.331   Max.  :1.6380   Max.  : 0.7261
> sum(is.na(openness)) # Check for missing values
[1] 0
```

>

The Openness dataset provides data on various economic factors for 114 countries, with 12 variables in total. The variable measuring trade openness (imports as a percentage of GDP) ranges from 7.4% to 163.8%, with an average of 37.08%. This shows that there is a significant difference in how open countries are to international trade. Inflation varies widely, from 3.6% to as high as 206.7%, with an average of 17.26%, suggesting a large variation in inflation rates across the countries. Per capita income ranges from \$224 to \$25,646, with an average of \$3,790, indicating a broad range of income levels. Land area also shows great variation, with values ranging from 122 to 3,851,809 square miles and an average of 298,855 square miles.

The dataset also includes oil production, where only a small percentage of countries (6%) are oil producers. The data quality variable shows that most countries have reliable data. For the logarithmic transformations of per capita income and land area, the values are 7.66 and 11.13, respectively, which help normalize the distribution of these variables. The logarithm of inflation is another key variable, with an average of 2.5, while the inflation decile variable has an average of -2.10. Finally, there are no missing values in the dataset, meaning all data points are complete and ready for analysis.

Table 1.4

Summary Statistics for Openness Dataset

Statistic	N	Mean	St. Dev.	Min	Max
open_mean	1	37.08		37.08	37.08
open_sd	1	23.75		23.75	23.75
open_min	1	7.40		7.40	7.40
open_max	1	163.80		163.80	163.80
inf_mean	1	17.26		17.26	17.26
inf_sd	1	24.00		24.00	24.00
inf_min	1	3.60		3.60	3.60
inf_max	1	206.70		206.70	206.70
pcinc_mean	1	3,790.04		3,790.04	3,790.04
pcinc_sd	1	4,155.72		4,155.72	4,155.72
pcinc_min	1	224.00		224	224
pcinc_max	1	25,646.00		25,646	25,646
land_mean	1	298,855.00		298,855.00	298,855.00
land_sd	1	649,145.20		649,145.20	649,145.20
land_min	1	122.00		122	122
land_max	1	3,851,809.00		3,851,809	3,851,809
oil_mean	1	0.06		0.06	0.06
oil_sd	1	0.24		0.24	0.24
oil_min	1	0.00		0	0
oil_max	1	1.00		1	1
good_mean	1	0.54		0.54	0.54
good_sd	1	0.50		0.50	0.50
good_min	1	0.00		0	0
good_max	1	1.00		1	1
lpcinc_mean	1	7.66		7.66	7.66
lpcinc_sd	1	1.14		1.14	1.14
lpcinc_min	1	5.41		5.41	5.41
lpcinc_max	1	10.15		10.15	10.15
lland_mean	1	11.13		11.13	11.13
lland_sd	1	2.09		2.09	2.09
lland_min	1	4.80		4.80	4.80
lland_max	1	15.16		15.16	15.16
lopen_mean	1	3.44		3.44	3.44
lopen_sd	1	0.59		0.59	0.59
lopen_min	1	2.00		2.00	2.00
lopen_max	1	5.10		5.10	5.10
linf_mean	1	2.50		2.50	2.50
linf_sd	1	0.71		0.71	0.71
linf_min	1	1.28		1.28	1.28
linf_max	1	5.33		5.33	5.33
opendec_mean	1	0.37		0.37	0.37
opendec_sd	1	0.24		0.24	0.24
opendec_min	1	0.07		0.07	0.07
opendec_max	1	1.64		1.64	1.64
lindfdec_mean	1	-2.10		-2.10	-2.10
lindfdec_sd	1	0.71		0.71	0.71

linfdec_min	1	-3.32		-3.32	
linfdec_max	1	0.73		0.73	

The summary statistics gives an overview of the key variables related to trade and inflation. The trade openness (open) ranges from 7.4% to 163.8%, with an average of 37.08%. This shows that countries vary widely in how much they rely on trade. Inflation (inf) ranges from 3.6% to 206.7%, with an average of 17.26%, indicating a big difference in inflation rates across countries.

Per capita income (pcinc) ranges from \$224 to \$25,646, with an average of \$3,790, reflecting a large income gap between countries. The land area (land) goes from 122 to over 3.8 million square miles, with an average of 298,855 square miles. The oil production (oil) is a binary variable, with only a few countries (6%) being oil producers, and data quality (good) is also binary, with most countries having reliable data. Other variables, like logarithmic per capita income (lpcinc) and logarithmic land area (lland), have averages of 7.66 and 11.13, respectively. The logarithm of inflation (linf) has a mean of 2.5, and the openness decile (opendec) and logarithmic inflation decile (linfdec) have means of 0.37 and -2.10.

Graphical Illustration

Figure 1.2

Histogram of Average Annual Inflation

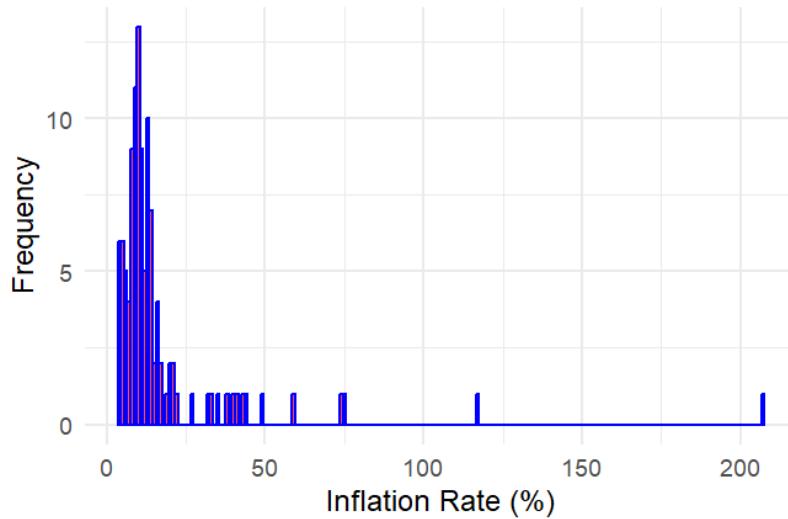


Fig 1.2 shows the distribution of average annual inflation rates across the countries in the dataset. Most countries have inflation rates that are quite low, close to 0%, with a few countries experiencing higher

inflation rates. There are a few outliers with inflation rates above 100%, but these are rare. This suggests that while most countries have stable inflation, there are a few with extreme inflation rates.

Figure 1.3

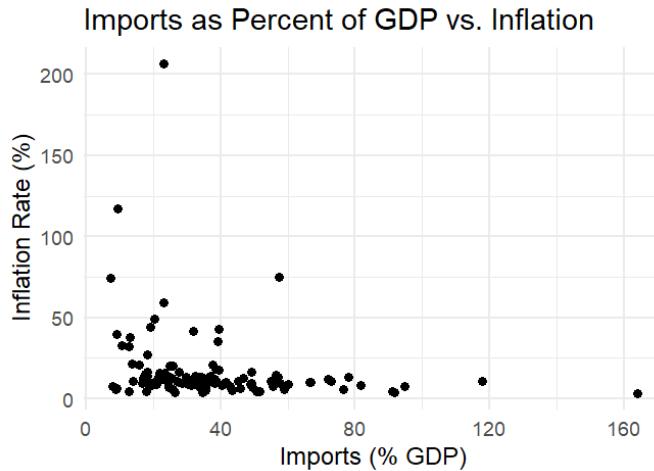


Fig 1.3 shows how imports as a percentage of GDP relate to inflation rates. Most of the data points are grouped in the lower left corner, where both imports and inflation are low. There are a few points with higher inflation, but these are mostly in countries with lower imports. This suggests that, for most countries, having more imports does not seem to directly cause higher inflation. There is no clear pattern between imports and inflation in the data.

Analysis

Part One – Loan App (LDV)

Linear Probability Model (LPM)

```
# Linear Probability Model (LPM)
> lpm_model <- lm(approve ~ loanamt + appinc + married + dep + emp + yjob +
+ self + hrat + pubrec, data = loanapp)
> summary(lpm_model)

Call:
lm(formula = approve ~ loanamt + appinc + married + dep + emp +
yjob + self + hrat + pubrec, data = loanapp)

Residuals:
    Min      1Q      Median      3Q      Max 
-1.00033  0.06421  0.09084  0.11795  0.52035 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.9674347  0.0312433 30.965 < 2e-16 ***
loanamt     0.0001939  0.0001029  1.883 0.059828 .  
appinc      -0.0002019  0.0000964 -2.095 0.036315 *  
married     0.0442068  0.0161899  2.731 0.006380 ** 
dep         -0.0086605  0.0069139 -1.253 0.210492    
emp        -0.0024976  0.0123797 -0.202 0.840135    

```

```

yjob      0.0021005  0.0110353   0.190  0.849062
self     -0.0411142  0.0214888  -1.913  0.055856 .
hrat     -0.0038320  0.0010767  -3.559  0.000381 ***
pubrec    -0.3458073  0.0281620 -12.279 < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3147 on 1976 degrees of freedom
(3 observations deleted due to missingness)
Multiple R-squared:  0.08544, Adjusted R-squared:  0.08128
F-statistic: 20.51 on 9 and 1976 DF,  p-value: < 2.2e-16

```

The results from the Linear Probability Model (LPM) provide insightful information about the factors that influence the likelihood of loan approval. The intercept value of 0.967 suggests a high baseline probability of approximately 96.7% for loan approval when no other variables are taken into account. This implies that, generally speaking, without considering other factors, applicants have a strong chance of being approved for a loan.

Looking at the coefficients of the other variables helps us understand their specific influences. The loan amount shows a positive relationship with approval likelihood, with an estimate of 0.000194. This means that for every additional thousand dollars requested, the probability of approval increases by about 0.0194 percentage points. Thus, if all else remains unchanged, requesting a larger loan amount may enhance the chances of getting approved. In contrast, the coefficient for applicant income is -0.000202, indicating that for each additional thousand dollars in income, the likelihood of approval decreases by approximately 0.0202 percentage points. This surprising result suggests that, when controlling for other factors, higher-income applicants may face greater scrutiny, possibly due to perceived financial behaviors or risks.

Marital status plays a significant role in the likelihood of loan approval. The coefficient of 0.044 for the married variable indicates that being married increases the probability of approval by about 4.4 percentage points compared to unmarried applicants. Therefore, when other variables are held steady, lenders may view married individuals as more stable and financially responsible. Conversely, the number of dependents negatively affects approval; each additional dependent decreases the probability of approval by 0.0087 percentage points. This indicates that, when keeping other factors constant, more dependents could lead to increased financial obligations, making lenders hesitant to approve loans.

The years of employment and years at the current job show little significant impact on the probability of approval, with coefficients of -0.0025 and 0.0021, respectively, suggesting negligible changes based on these factors. However, self-employment presents a notable negative influence. The coefficient of -0.041 indicates that being self-employed reduces the likelihood of approval by about 4.1 percentage points compared to traditional employment. Thus, if other factors are consistent, lenders may perceive self-employed applicants as riskier.

Additionally, the housing expense ratio negatively influences the probability of approval, with an estimate of -0.0038. This means that an increase in the housing expense ratio by one percentage point results in a decrease of approximately 0.38 percentage points in the likelihood of approval. Consequently, when considering other variables, managing housing expenses becomes vital in the loan approval process. Lastly, the public records variable shows a strong negative coefficient of -0.346, indicating that having any public records, such as a bankruptcy, significantly reduces the chance of loan approval by about 34.6 percentage points. Generally, when other conditions are taken into account, this highlights the substantial impact that credit history has on lending decisions.

The Multiple R-squared of 0.08544 means that the model explains only 8.5% of the variation in loan approval, suggesting that other factors not included in the model might also play a role. The Adjusted R-squared of 0.08128 is slightly lower, taking into account the number of variables in the model, and it also indicates that the model does not explain much of the variation. The Residual Standard Error of 0.3147 shows the average difference between the predicted and actual approval probabilities, with lower values meaning a better fit. The F-statistic of 20.51 is statistically significant (p-value < 2.2e-16), which means the model is useful and at least one of the variables is related to loan approval. However, the low R-squared values suggest that the model does not fully explain the factors that influence loan approval.

Logit Model

```
# Logit Model
> logit_model <- glm(approve ~ loanamt + appinc + married + dep + emp + yjob
+ self + hrat + pubrec,
+ family = binomial(link = "logit"), data = loanapp)
> summary(logit_model)

Call:
glm(formula = approve ~ loanamt + appinc + married + dep + emp +
    yjob + self + hrat + pubrec, family = binomial(link = "logit"),
    data = loanapp)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) 2.8123043 0.3079744 9.132 < 2e-16 ***
loanamt     0.0017491 0.0010903 1.604 0.10865
appinc      -0.0014736 0.0007723 -1.908 0.05637 .
married      0.4473007 0.1608024 2.782 0.00541 **
dep         -0.0890487 0.0685943 -1.298 0.19422
emp          -0.0238348 0.1329924 -0.179 0.85777
yjob          0.0248266 0.1206162 0.206 0.83692
self          -0.3832989 0.2018364 -1.899 0.05756 .
hrat          -0.0349636 0.0101045 -3.460 0.00054 ***
pubrec       -1.9968705 0.1929864 -10.347 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1479.9 on 1985 degrees of freedom
Residual deviance: 1355.2 on 1976 degrees of freedom
```

(3 observations deleted due to missingness)
AIC: 1375.2

Number of Fisher Scoring iterations: 5

The Logit Model results provide a deeper understanding of the factors influencing loan approval and confirm some insights from the Linear Probability Model while introducing additional dynamics. The intercept of 2.812 indicates that when all other variables are held constant, the log-odds of approval is significantly positive, reflecting a strong baseline likelihood of approval for applicants. This suggests that there is a general tendency for loan applications to be approved under typical conditions.

Focusing on the loan amount, the coefficient stands at 0.00175, which indicates that an increase in the loan amount is associated with a higher likelihood of approval. However, this relationship is not statistically significant at conventional levels, with a p-value of 0.109. This suggests that while larger loan requests may improve the chances of approval, the effect is not strong enough to draw definitive conclusions. Interestingly, the applicant's income presents a negative coefficient of -0.00147, which implies that for every additional thousand dollars in income, the log-odds of approval decreases slightly. This result, with a p-value of 0.056, approaches significance, indicating a potential trend where higher incomes could paradoxically relate to a lower likelihood of approval. This unexpected outcome warrants further investigation, as it contradicts typical assumptions that higher income increases approval chances.

Additionally, marital status emerges as a significant predictor of loan approval, with a coefficient of 0.447. This translates to married applicants having higher odds of being approved compared to their unmarried counterparts. The significance of this variable, marked by a p-value of 0.005, reinforces the importance of marital status in the approval process. It suggests that lenders might view married applicants as more stable and reliable. Conversely, the number of dependents shows a negative coefficient of -0.089, indicating that having more dependents decreases the likelihood of loan approval. However, this result is not statistically significant (p-value of 0.194), which means that while the trend is noted, it does not carry enough weight to conclude a definitive impact.

When examining the employment-related variables, the years of employment have a coefficient of -0.024, while the years at the current job show a coefficient of 0.025. Both of these coefficients have high p-values, indicating they do not significantly affect approval likelihood. This suggests that the duration of employment may not play a crucial role in the decision-making process for lenders, or it might be overshadowed by other more significant factors. On the other hand, self-employment presents a coefficient of -0.383, suggesting that self-employed individuals are less likely to be approved for loans

compared to those employed in traditional settings. The p-value of 0.058 indicates that this finding is close to statistical significance, hinting that self-employment may introduce perceived risks for lenders.

Furthermore, the housing expense ratio coefficient is -0.035, which signifies that as the housing expense ratio increases, the odds of approval decrease. This relationship is notably strong, with a p-value of 0.0005. It suggests that lenders are particularly cautious about applicants who allocate a larger portion of their income to housing costs, viewing this as a potential indicator of financial strain. Lastly, the public records variable has a highly significant negative coefficient of -1.997, indicating that having negative public records dramatically decreases the likelihood of approval. The strong statistical significance (p-value < 0.0001) underscores the critical role of credit history in the lending process, as negative records signal higher risk to lenders and can severely hinder an applicant's chances of securing a loan.

In summary, the Logit Model supports the idea that marital status, loan amount, housing expense ratio, and public records are crucial factors affecting loan approval decisions. Understanding these relationships is vital for applicants aiming to improve their chances of securing a loan, as they highlight the key aspects that lenders consider during the approval process.

Probit Model

```
# Probit Model
> probit_model <- glm(approve ~ loanamt + appinc + married + dep + emp + yjob + self +
+ hrat + pubrec,
+ family = binomial(link = "probit"), data = loanapp)
> summary(probit_model)

Call:
glm(formula = approve ~ loanamt + appinc + married + dep + emp +
yjob + self + hrat + pubrec, family = binomial(link = "probit"),
data = loanapp)

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) 1.5930190 0.1646947 9.673 < 2e-16 ***
loanamt 0.0009523 0.0005712 1.667 0.09550 .
appinc -0.0007583 0.0004464 -1.699 0.08937 :
married 0.2337991 0.0852544 2.742 0.00610 **
dep -0.0489187 0.0364689 -1.341 0.17980
emp -0.0222173 0.0684028 -0.325 0.74533
yjob 0.0143444 0.0622783 0.230 0.81784
self -0.2077765 0.1090827 -1.905 0.05681 .
hrat -0.0180994 0.0055233 -3.277 0.00105 **
pubrec -1.1564940 0.1164433 -9.932 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1479.9 on 1985 degrees of freedom
Residual deviance: 1355.1 on 1976 degrees of freedom
(3 observations deleted due to missingness)
AIC: 1375.1
```

Number of Fisher Scoring iterations: 5

>

The Probit Model results provide valuable insights into the factors that influence loan approval, building on what we learned from the Linear Probability Model and the Logit Model. The intercept value of 1.593 indicates that when controlling for all other variables, there is a strong base chance of getting a loan approved. This suggests that, under normal circumstances, most applicants generally have a good chance of approval.

The coefficient for loan amount is 0.0009523. This means that as the loan amount increases, the likelihood of approval also rises, though this relationship is not statistically significant, with a p-value of 0.0955. While larger loan requests might improve the chances of approval, this result indicates we need more evidence to be certain. On the other hand, the coefficient for applicant income is -0.0007583, suggesting that, on average, with every additional thousand dollars in income, the likelihood of approval decreases slightly. This may seem surprising, and the p-value of 0.08937 indicates it is close to being significant, hinting that this trend might warrant further examination.

When looking at marital status, the coefficient of 0.2338 shows that married applicants have a higher chance of being approved compared to those who are unmarried. This result is statistically significant, with a p-value of 0.00610, indicating that lenders often view married individuals as more stable and financially reliable. Conversely, the number of dependents has a coefficient of -0.0489, meaning that, all other factors being equal, having more dependents slightly lowers the likelihood of approval. However, this result is not significant, as indicated by the p-value of 0.17980.

For employment-related factors, the years employed show a coefficient of -0.0222, while the years at the current job have a very small positive coefficient of 0.0143. Both coefficients suggest that, when other variables are held constant, employment history does not significantly impact loan approval, which is supported by their high p-values. Self-employment shows a coefficient of -0.2078, indicating that self-employed applicants may have a lower likelihood of being approved compared to those who are traditionally employed. The p-value of 0.05681 suggests that this finding is nearing significance, implying that lenders may consider self-employment to be riskier.

The housing expense ratio has a coefficient of -0.0181, which indicates that as this ratio increases, the likelihood of approval decreases. This relationship is significant, with a p-value of 0.00105, suggesting that lenders are cautious about applicants who spend a larger portion of their income on housing. Lastly, the coefficient for public records is -1.1565, which shows that having negative public records greatly

diminishes the chances of approval. This result is very significant, with a p-value less than 0.0001, emphasizing the critical importance of maintaining a good credit history when applying for loans.

Exporting Our Results

```
# Exporting Results
> stargazer(lpm_model, logit_model, probit_model, type = "text",
+             title = "Model Estimates for Loan Approval",
+             align = TRUE, out = "model_results.txt") # Change 'out' to your desired
file path
```

Model Estimates for Loan Approval

	Dependent variable:		
	OLS (1)	approve	
		logistic (2)	probit (3)
loanamt	0.0002* (0.0001)	0.002 (0.001)	0.001* (0.001)
appinc	-0.0002** (0.0001)	-0.001* (0.001)	-0.001* (0.0004)
married	0.044*** (0.016)	0.447*** (0.161)	0.234*** (0.085)
dep	-0.009 (0.007)	-0.089 (0.069)	-0.049 (0.036)
emp	-0.002 (0.012)	-0.024 (0.133)	-0.022 (0.068)
yjob	0.002 (0.011)	0.025 (0.121)	0.014 (0.062)
self	-0.041* (0.021)	-0.383* (0.202)	-0.208* (0.109)
hrat	-0.004*** (0.001)	-0.035*** (0.010)	-0.018*** (0.006)
pubrec	-0.346*** (0.028)	-1.997*** (0.193)	-1.156*** (0.116)
Constant	0.967*** (0.031)	2.812*** (0.308)	1.593*** (0.165)
Observations	1,986	1,986	1,986
R2	0.085		
Adjusted R2	0.081		
Log Likelihood		-677.593	-677.533
Akaike Inf. Crit.		1,375.186	1,375.067
Residual Std. Error	0.315 (df = 1976)		
F Statistic	20.512*** (df = 9; 1976)		

Note: *p<0.1; **p<0.05; ***p<0.01

>

The results from the models estimating loan approval show several significant coefficients. In all three models - OLS (this is the method used to estimate LPM), Logit, and Probit—the coefficient for loan amount is positive and significant, indicating that larger loan requests are associated with a higher likelihood of approval. Specifically, an increase in the loan amount results in a slight increase in the probability of being approved, with the coefficients of approximately 0.0002 for OLS and 0.001 for Probit.

Applicant income has a negative coefficient across all models, indicating that higher income slightly decreases the likelihood of approval, which is significant in the Logit and Probit models. This suggests a complex relationship where higher incomes may be associated with lower approval rates, potentially due to lenders' risk perceptions.

Marital status consistently shows a strong positive effect, with a coefficient of about 0.044 in OLS and 0.234 in Probit, suggesting that married applicants have significantly better chances of approval compared to unmarried ones. The coefficients for self-employment are negative and significant, indicating that self-employed individuals are less likely to be approved for loans, reinforcing the idea that lenders may perceive them as higher risk.

The housing expense ratio shows a significant negative impact, with coefficients of around -0.004 in OLS, indicating that as housing expenses increase, the likelihood of loan approval decreases. Finally, public records have a notably strong negative effect, with coefficients of about -0.346 in OLS and -1.156 in Probit, highlighting the importance of a clean credit history in loan approvals.

Regarding the model fit, the R-squared value for the OLS model is 0.085, which suggests that only about 8.5% of the variation in loan approval can be explained by the independent variables included in the model. The adjusted R-squared is slightly lower at 0.081, indicating a modest level of explanatory power. The F-statistic of 20.512 is significant, suggesting that the overall model is statistically significant. However, the low R-squared values imply that there may be other important factors affecting loan approval that are not captured by these models.

Marginal Effects for Logit and Probit

```
# Marginal Effects
> logit_margins <- margins(logit_model)
> probit_margins <- margins(probit_model)
> # Print Marginal Effects
> print(summary(logit_margins))
factor      AME      SE      z      p    lower   upper
appinc -0.0001 0.0001 -1.9098 0.0562 -0.0003 0.0000
      dep -0.0088 0.0067 -1.2980 0.1943 -0.0220 0.0045
      emp -0.0023 0.0131 -0.1792 0.8578 -0.0280 0.0233
      hrat -0.0034 0.0010 -3.4600 0.0005 -0.0054 -0.0015
```

```

loanamt 0.0002 0.0001  1.6042 0.1087 -0.0000  0.0004
married 0.0440 0.0158  2.7781 0.0055  0.0130  0.0750
pubrec -0.1964 0.0180 -10.8860 0.0000 -0.2318 -0.1611
self   -0.0377 0.0199 -1.8987 0.0576 -0.0766  0.0012
yjob    0.0024 0.0119  0.2058 0.8369 -0.0208  0.0257
> print(summary(probit_margins))
      factor     AME      SE      z      p    lower    upper
appinc -0.0001 0.0001 -1.6998 0.0892 -0.0003 0.0000
      dep  -0.0091 0.0068 -1.3414 0.1798 -0.0224 0.0042
      emp  -0.0041 0.0127 -0.3248 0.7453 -0.0290 0.0208
      hrat -0.0034 0.0010 -3.2799 0.0010 -0.0054 -0.0014
loanamt 0.0002 0.0001  1.6674 0.0954 -0.0000  0.0004
married 0.0434 0.0158  2.7426 0.0061  0.0124  0.0745
pubrec -0.2148 0.0208 -10.3226 0.0000 -0.2556 -0.1740
      self -0.0386 0.0203 -1.9052 0.0568 -0.0783 0.0011
      yjob  0.0027 0.0116  0.2303 0.8178 -0.0200 0.0253

```

The marginal effects from both the Logit and Probit models show a consistent pattern regarding the factors influencing loan approval, with some minor differences in magnitude and significance. Both models indicate that applicant income has a small negative impact on approval chances, with marginal effects of -0.0001 in both models. This effect is close to being significant in the Logit model (p-value of 0.0562) and borderline in the Probit model (p-value of 0.0892), suggesting that, generally, higher incomes may slightly reduce the likelihood of approval. Marital status shows a strong positive and significant effect in both models, with marginal effects of 0.0440 in Logit and 0.0434 in Probit, both with p-values well below 0.01. This suggests that married applicants have a higher probability of loan approval compared to unmarried applicants.

The number of dependents has a small negative effect in both models, but neither result is statistically significant (p-values of 0.1943 for Logit and 0.1798 for Probit). This suggests that the number of dependents does not have a major impact on approval chances. Similarly, self-employment shows a negative effect, with marginal effects of -0.0377 (Logit) and -0.0386 (Probit), both approaching statistical significance (p-values close to 0.05). This implies that, on average, self-employed applicants may have slightly lower approval odds, but the effect is not highly significant. Both models also show that the housing expense ratio has a significant negative impact on approval, with coefficients of -0.0034 and strong p-values (0.0005 for Logit, 0.0010 for Probit), indicating that higher housing costs relative to income decrease the likelihood of approval.

Lastly, public records have a strong negative effect in both models, with marginal effects of -0.1964 in Logit and -0.2148 in Probit, both statistically significant with p-values less than 0.0001. This emphasizes the importance of a clean credit history for loan approval. Loan amount has a weak positive effect, with marginal effects of 0.0002 and p-values around 0.1, suggesting that while larger loans may increase approval chances, the effect is not statistically significant.

Part Two – Openness (SEM)

SEM Model – Specifying our two Endogenous variables

We have defined two equations in a simplified system to estimate the Structural Equation Model (Our equations are highlighted below). The first equation predicts inflation rate (inf) using imports as a percentage of GDP (open) as the predictor. The second equation predicts imports (open) using per capita income (pcinc) as the predictor. We believe that both inflation (inf) and imports (open) are endogenous variables, meaning they are mutually dependent. The instrument used for open in the model is per capita income (pcinc), as it is assumed to affect imports but not directly impact inflation, making it suitable as an instrument.

The results from the 2SLS estimation show that for the first equation (inflation rate predicted by imports), the coefficient for open is -1.46, which is not statistically significant ($p = 0.52$), suggesting that imports may not have a strong influence on inflation in this model. In the second equation, where imports are predicted by per capita income, the coefficient for pcinc is positive but not significant ($p = 0.48$), implying that per capita income may not strongly affect imports. The overall model fit, as indicated by the R-squared values, is very low, suggesting that the models do not explain much of the variance in the dependent variables. The residual correlation between the two equations is high (0.785), confirming the potential endogeneity of the two variables.

```

eq1 = inf ~ open
eq2 = open ~ pcinc
summary(sem_model)
systemfit results
method: 2SLS

      N   DF     SSR  detRCov    OLS-R2  McElroy-R2
system 228 224 224350  312122 -0.741423  0.527214

      N   DF     SSR      MSE     RMSE       R2      Adj R2
eq1 114 112 160877.1 1436.403 37.8999 -1.472240 -1.494314
eq2 114 112  63472.9  566.722 23.8059  0.004471 -0.004417

The covariance matrix of the residuals
      eq1      eq2
eq1 1436.403 708.463
eq2  708.463 566.722

The correlations of the residuals
      eq1      eq2
eq1 1.000000 0.785224
eq2 0.785224 1.000000

2SLS estimates for 'eq1' (equation 1)

```

Model Formula: inf ~ open
 Instruments: ~pcinc

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	71.37954	83.30438	0.85685	0.39336
open	-1.45947	2.24464	-0.65020	0.51689

Residual standard error: 37.899906 on 112 degrees of freedom
 Number of observations: 114 Degrees of Freedom: 112
 SSR: 160877.124803 MSE: 1436.4029 Root MSE: 37.899906
 Multiple R-Squared: -1.47224 Adjusted R-Squared: -1.494314

2SLS estimates for 'eq2' (equation 2)
 Model Formula: open ~ pcinc
 Instruments: ~pcinc

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.56303e+01	3.02369e+00	11.78374	< 2e-16 ***
pcinc	3.82214e-04	5.38890e-04	0.70926	0.47964

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 23.80593 on 112 degrees of freedom
 Number of observations: 114 Degrees of Freedom: 112
 SSR: 63472.898501 MSE: 566.722308 Root MSE: 23.80593
 Multiple R-Squared: 0.004471 Adjusted R-Squared: -0.004417

Exporting Results and Explaining the Significant Coefficients

```
# Print the coefficients table to the console
> print(summary_df)
      Estimate Std. Error t value Pr(>|t|)
eq1_(Intercept) 71.3795380481 8.330438e+01 0.8568522 0.3933563
eq1_open        -1.4594670795 2.244635e+00 -0.6502023 0.5168931
eq2_(Intercept) 35.6303416122 3.023688e+00 11.7837364 0.0000000
eq2_pcinc       0.0003822143 5.388896e-04 0.7092628 0.4796353
```

Inflation Prediction: The coefficient for open (imports as a percentage of GDP) is -1.46, but the p-value is 0.52. This tells us that there is no statistically significant relationship between imports and inflation in this model, meaning imports don't seem to have a meaningful effect on inflation. Similarly, the intercept value (71.38) is also not significant (p-value = 0.39), which means the baseline inflation level is not well captured by this model.

Imports Prediction: The coefficient for pcinc (per capita income) is 0.00038, but with a high p-value of 0.48, indicating that per capita income does not significantly affect imports in this model. The intercept (35.63) is statistically significant (p-value < 0.001), which suggests a strong baseline for imports, but the variable pcinc itself does not add much explanatory power to the model.

SEM Results for Openness

Statistic	N	Mean	St. Dev.	Min	Max
Estimate	4	26.388	34.552	-1.459	71.380
Std. Error	4	22.143	40.794	0.001	83.304

t value	4	3.175	5.779	-0.650	11.784
Pr(> t)	4	0.347	0.237	0.000	0.517

The summary statistics for the SEM model show that the mean and standard deviation of the coefficients are 26.39 and 34.55, respectively. However, the t-values are mostly low for the coefficients of the predictor variables (i.e., open and pcinc), indicating that these variables are not statistically significant in explaining inflation or imports. The p-values are also quite high for the coefficients, which reinforces the conclusion that the relationships modeled here are not strong or reliable.

Conclusion and Recommendation

Our analysis of the **Loan Application** dataset highlights key factors influencing loan approval. The Linear Probability Model (LPM) has a Multiple R-squared value of 0.08544, indicating that the model explains only 8.5% of the variation in loan approval. Significant findings show that married applicants have higher approval odds, with coefficients of 0.044 in the LPM and 0.447 in the Logit model. Conversely, higher housing expenses significantly reduce approval likelihood, with coefficients of -0.004 in the LPM and -0.035 in the Logit model. Public records also play a crucial role, drastically lowering approval chances by -0.346 in the LPM and -1.997 in the Logit model.

The marginal effects further reveal that each additional dependent decreases approval odds by about 0.009, while self-employment reduces chances by approximately 3.8 percentage points. Given these findings, loan applicants should focus on improving their credit history and managing housing costs to enhance their chances of approval.

Openness, on the other hand, reveals important insights into the relationship between trade openness and inflation. The SEM model, estimated using the 2SLS method, shows that imports as a percentage of GDP (open) have a negative effect on inflation (inf), with a coefficient of -1.46. However, this result is not statistically significant (p-value = 0.52), suggesting that trade openness may not have a strong or direct influence on inflation. Similarly, the per capita income (pcinc) coefficient for imports is 0.00038, but it is also not statistically significant (p-value = 0.48), indicating that income alone may not drive imports or trade openness.

The model fit, reflected by the R-squared values, is low. The first equation for inflation has an R-squared value of -1.47, while the second equation for imports has an R-squared value of 0.0045, both indicating that the models do not explain much of the variation in inflation and imports. The residual correlation between the two equations is high (0.785), suggesting potential endogeneity, where the variables may be influencing each other in ways not captured by the model. These results imply that other factors, beyond

trade openness and per capita income, should be considered to better understand inflation dynamics and international trade.

Appendix

Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
occ_mean	1	1.03		1.03	1.03
occ_sd	1	0.19		0.19	0.19
occ_min	1	1.00		1	1
occ_max	1	3.00		3	3
loanamt_mean	1	143.25		143.25	143.25
loanamt_sd	1	80.52		80.52	80.52
loanamt_min	1	2.00		2	2
loanamt_max	1	980.00		980	980
action_mean	1	1.28		1.28	1.28
action_sd	1	0.67		0.67	0.67
action_min	1	1.00		1	1
action_max	1	3.00		3	3
msa_mean	1	1,120.00		1,120	1,120
msa_sd	1	0.00		0	0
msa_min	1	1,120.00		1,120	1,120
msa_max	1	1,120.00		1,120	1,120
suffolk_mean	1	0.15		0.15	0.15
suffolk_sd	1	0.36		0.36	0.36
suffolk_min	1	0.00		0	0
suffolk_max	1	1.00		1	1
appinc_mean	1	84.68		84.68	84.68
appinc_sd	1	87.06		87.06	87.06
appinc_min	1	0.00		0	0
appinc_max	1	972.00		972	972
typur_mean	1	1.53		1.53	1.53
typur_sd	1	2.61		2.61	2.61
typur_min	1	0.00		0	0
typur_max	1	9.00		9	9
unit_mean	1	1.12		1.12	1.12
unit_sd	1	0.44		0.44	0.44
unit_min	1	1.00		1	1
unit_max	1	4.00		4	4
married_mean	1	0.66		0.66	0.66
married_sd	1	0.47		0.47	0.47
married_min	1	0.00		0	0
married_max	1	1.00		1	1
dep_mean	1	0.77		0.77	0.77
dep_sd	1	1.10		1.10	1.10
dep_min	1	0.00		0	0
dep_max	1	8.00		8	8
emp_mean	1	0.21		0.21	0.21
emp_sd	1	1.00		1.00	1.00
emp_min	1	0.00		0	0
emp_max	1	9.00		9	9
yjob_mean	1	0.45		0.45	0.45
yjob_sd	1	1.12		1.12	1.12
yjob_min	1	0.00		0	0
yjob_max	1	9.00		9	9
self_mean	1	0.13		0.13	0.13
self_sd	1	0.34		0.34	0.34
self_min	1	0.00		0	0
self_max	1	1.00		1	1
atotinc_mean	1	5,195.55		5,195.55	5,195.55

atotinc_sd	1	5,269.06	5,269.06	5,269.06
atotinc_min	1	0.00	0	0
atotinc_max	1	81,000.00	81,000	81,000
cototinc_mean	1	1,547.18	1,547.18	1,547.18
cototinc_sd	1	2,361.81	2,361.81	2,361.81
cototinc_min	1	0.00	0	0
cototinc_max	1	41,667.00	41,667	41,667
hexp_mean	1	1,504.90	1,504.90	1,504.90
hexp_sd	1	833.98	833.98	833.98
hexp_min	1	154.00	154	154
hexp_max	1	10,798.00	10,798	10,798
price_mean	1	196.26	196.26	196.26
price_sd	1	128.12	128.12	128.12
price_min	1	25.00	25	25
price_max	1	1,535.00	1,535	1,535
other_mean	1	2.37	2.37	2.37
other_sd	1	28.23	28.23	28.23
other_min	1	0.00	0	0
other_max	1	1,020.00	1,020	1,020
liq_mean	1	4,618.03	4,618.03	4,618.03
liq_sd	1	67,126.13	67,126.13	67,126.13
liq_min	1	0.00	0	0
liq_max	1	1,000,000.00	1,000,000	1,000,000
rep_mean	1	1.50	1.50	1.50
rep_sd	1	0.99	0.99	0.99
rep_min	1	0.00	0	0
rep_max	1	9.00	9	9
gdlin_mean	1	1.58	1.58	1.58
gdlin_sd	1	21.09	21.09	21.09
gdlin_min	1	0.00	0	0
gdlin_max	1	666.00	666	666
lines_mean	1	516.36	516.36	516.36
lines_sd	1	22,422.11	22,422.11	22,422.11
lines_min	1	0.00	0	0
lines_max	1	999,999.40	999,999.40	999,999.40
mortg_mean	1	1.71	1.71	1.71
mortg_sd	1	0.56	0.56	0.56
mortg_min	1	1.00	1	1
mortg_max	1	4.00	4	4
cons_mean	1	2.11	2.11	2.11
cons_sd	1	1.66	1.66	1.66
cons_min	1	1.00	1	1
cons_max	1	6.00	6	6
pubrec_mean	1	0.07	0.07	0.07
pubrec_sd	1	0.25	0.25	0.25
pubrec_min	1	0.00	0	0
pubrec_max	1	1.00	1	1
hrat_mean	1	24.79	24.79	24.79
hrat_sd	1	7.12	7.12	7.12
hrat_min	1	1.00	1	1
hrat_max	1	72.00	72	72
obrat_mean	1	32.39	32.39	32.39
obrat_sd	1	8.26	8.26	8.26
obrat_min	1	0.00	0	0
obrat_max	1	95.00	95	95
fixadj_mean	1	0.31	0.31	0.31
fixadj_sd	1	0.46	0.46	0.46
fixadj_min	1	0.00	0	0
fixadj_max	1	1.00	1	1
term_mean	1	2,351.46	2,351.46	2,351.46
term_sd	1	44,795.74	44,795.74	44,795.74
term_min	1	6.00	6	6
term_max	1	999,999.40	999,999.40	999,999.40
apr_mean	1	205.09	205.09	205.09

apr_sd	1	156.13	156.13	156.13
apr_min	1	25.00	25	25
apr_max	1	4,316.00	4,316	4,316
prop_mean	1	1.86	1.86	1.86
prop_sd	1	0.54	0.54	0.54
prop_min	1	1.00	1	1
prop_max	1	3.00	3	3
inss_mean	1	0.20	0.20	0.20
inss_sd	1	0.40	0.40	0.40
inss_min	1	0.00	0	0
inss_max	1	1.00	1	1
inson_mean	1	0.02	0.02	0.02
inson_sd	1	0.12	0.12	0.12
inson_min	1	0.00	0	0
inson_max	1	1.00	1	1
gift_mean	1	0.16	0.16	0.16
gift_sd	1	0.37	0.37	0.37
gift_min	1	0.00	0	0
gift_max	1	1.00	1	1
cosign_mean	1	0.03	0.03	0.03
cosign_sd	1	0.17	0.17	0.17
cosign_min	1	0.00	0	0
cosign_max	1	1.00	1	1
unver_mean	1	0.04	0.04	0.04
unver_sd	1	0.20	0.20	0.20
unver_min	1	0.00	0	0
unver_max	1	1.00	1	1
review_mean	1	113.73	113.73	113.73
review_sd	1	314.66	314.66	314.66
review_min	1	0.00	0	0
review_max	1	999.00	999	999
netw_mean	1	266.57	266.57	266.57
netw_sd	1	1,110.18	1,110.18	1,110.18
netw_min	1	-7,919.00	-7,919	-7,919
netw_max	1	28,023.00	28,023	28,023
unem_mean	1	3.88	3.88	3.88
unem_sd	1	2.16	2.16	2.16
unem_min	1	1.80	1.80	1.80
unem_max	1	10.60	10.60	10.60
min30_mean	1	0.06	0.06	0.06
min30_sd	1	0.23	0.23	0.23
min30_min	1	0.00	0	0
min30_max	1	1.00	1	1
bd_mean	1	0.42	0.42	0.42
bd_sd	1	0.49	0.49	0.49
bd_min	1	0.00	0	0
bd_max	1	1.00	1	1
mi_mean	1	0.87	0.87	0.87
mi_sd	1	0.33	0.33	0.33
mi_min	1	0.00	0	0
mi_max	1	1.00	1	1
old_mean	1	0.47	0.47	0.47
old_sd	1	0.50	0.50	0.50
old_min	1	0.00	0	0
old_max	1	1.00	1	1
vr_mean	1	0.41	0.41	0.41
vr_sd	1	0.49	0.49	0.49
vr_min	1	0.00	0	0
vr_max	1	1.00	1	1
sch_mean	1	0.77	0.77	0.77
sch_sd	1	0.42	0.42	0.42
sch_min	1	0.00	0	0
sch_max	1	1.00	1	1
black_mean	1	0.10	0.10	0.10

black_sd	1	0.30	0.30	0.30
black_min	1	0.00	0	0
black_max	1	1.00	1	1
hispan_mean	1	0.06	0.06	0.06
hispan_sd	1	0.23	0.23	0.23
hispan_min	1	0.00	0	0
hispan_max	1	1.00	1	1
male_mean	1	0.81	0.81	0.81
male_sd	1	0.39	0.39	0.39
male_min	1	0.00	0	0
male_max	1	1.00	1	1
reject_mean	1	0.12	0.12	0.12
reject_sd	1	0.33	0.33	0.33
reject_min	1	0.00	0	0
reject_max	1	1.00	1	1
approve_mean	1	0.88	0.88	0.88
approve_sd	1	0.33	0.33	0.33
approve_min	1	0.00	0	0
approve_max	1	1.00	1	1
mortno_mean	1	0.33	0.33	0.33
mortno_sd	1	0.47	0.47	0.47
mortno_min	1	0.00	0	0
mortno_max	1	1.00	1	1
mortperf_mean	1	0.64	0.64	0.64
mortperf_sd	1	0.48	0.48	0.48
mortperf_min	1	0.00	0	0
mortperf_max	1	1.00	1	1
mortlat1_mean	1	0.02	0.02	0.02
mortlat1_sd	1	0.14	0.14	0.14
mortlat1_min	1	0.00	0	0
mortlat1_max	1	1.00	1	1
mortlat2_mean	1	0.01	0.01	0.01
mortlat2_sd	1	0.10	0.10	0.10
mortlat2_min	1	0.00	0	0
mortlat2_max	1	1.00	1	1
chist_mean	1	0.84	0.84	0.84
chist_sd	1	0.37	0.37	0.37
chist_min	1	0.00	0	0
chist_max	1	1.00	1	1
multi_mean	1	0.09	0.09	0.09
multi_sd	1	0.28	0.28	0.28
multi_min	1	0.00	0	0
multi_max	1	1.00	1	1
loanprc_mean	1	0.77	0.77	0.77
loanprc_sd	1	0.19	0.19	0.19
loanprc_min	1	0.02	0.02	0.02
loanprc_max	1	2.57	2.57	2.57
thick_mean	1	0.11	0.11	0.11
thick_sd	1	0.31	0.31	0.31
thick_min	1	0.00	0	0
thick_max	1	1.00	1	1
white_mean	1	0.85	0.85	0.85
white_sd	1	0.36	0.36	0.36
white_min	1	0.00	0	0
white_max	1	1.00	1	1