Analysis of MSME Loan Program Data for deAsra

Dr. Yugank Goyal
FLAME University

24.8.2021

Acknowledgement

We would like to thank Gromor Finance for providing this research with the requisite data, which served as a crucial base for the analysis. We would also like to express gratitude to the participating Financial Institutions (FIs): Gromor, Happy Loans, Micrograam, Janaseva, Saraswat for being a part of this program. We would also like to thank Dr. Yugank Goyal for his contribution to this study.

Why MSME Loan Program?

- deAsra conducted a survey in April 2020 to understand challenges faced by small business owners in the midst of a pandemic
- Responses were collected from 1000 small business owners
- Around 50% of the respondents stated lack of finance as the biggest hurdle
- In order to address this issue, deAsra tied up with multiple Financial Institutions (NBFCs, Cooperatives, Banks etc.) to provide access to capital to these entrepreneurs.
- deAsra encouraged users to apply for the MSME Emergency Loans during the first lockdown period of Covid-19 (starting from April 2020).

Step by Step Approach of MSME Loan Program

Interested Entrepreneurs enters his/her details basis which deAsra checks the eligibility across the 5 FIs

deAsra shows the names of the FI - that one is eligible for

Users can then apply to 1 or all the eligible FI

deAsra collects the application details

Details communicated to FIs

FIs interact with users and collect required documentary evidence

Based on the evidence and other selection criteria (may differ across FI), application is approved/rejected

If approved, loan is sanctioned as per payment disbursement schedule

Repayment starts - could be weekly/monthly depending on the FI

Participating Financial Institutions (FIs): Gromor, Happy Loans, Micrograam, Janaseva, Saraswat

Impact Study in partnership with Dr. Yugank Goyal

Ph.D. (Law and Economics): University of Hamburg, Erasmus University Rotterdam, University of Bologna.
Yugank is Associate Professor in Public Policy at FLAME University, Pune. He sits on the board of Indian School of Public Policy. Earlier he has been a founding faculty member of O.P. Jindal Global University. He is also a visiting faculty at IIMA and IIMK.

For this research, Dr Goyal was assisted by Mr. Mayank Patel, Ms. Nikhilla B., Ms. Ketki Balyan and Mr. Swapnil Doke.

Background

- This analysis is done on the datasheet containing some details of the MSME Loan Program of deAsra, and a specific datasheet on the Gromor loan applicants.
- The overall program stretches from June 2020 to October 2020.
- Timeline for the FI involvement is as follows. Gromor: June-August 2020, Happy: August 2020 onwards, and Micrograam, Janaseva, Saraswat: September 2020
- Each FI had its own eligibility requirements based on the following factors: *pin code, business type, loan amount, business age, age, business address proof, bank statement, UPI ID, co-borrower*.
- *Gender* and *business sector* were the only other variables collected by deAsra.
- The Gromor data sheet gives information on loans taken in months June-Oct in 2020.
- This is more useful datasheet which has personal details of loanees, their enterprise details of address, sector and size in terms of people employed, loan details and socio-economic detail of house ownership and education level. It gives loan details as days past due and total Overdue amounts as on 4th Aug 2021.

Methodology

For the MSME Loan Program General Dataset:

- The data required extensive and significant cleaning using email address and phone numbers.
- After the removal of duplicated entries, the final unique observations was 3162 (the original dataset had 4099 entries)
- Imported and examined the data in R and Excel to check summary, relationship between various variables and eligibility and performed regressions to assess which factors influenced which FIs, and which FIs mattered for which types of entrepreneurs.

For the Gromor dataset:

- Cleaned datasheet for only valid RFIDs in both sheets to get 131 valid observations.
- Imported in STATA14 and checked summary, correlation and regression for numeric variables.
- Performed cross tabulation of non-numeric variables to find relation between variables and loan defaulting tendencies.

Considerations

For the MSME Loan Program General Dataset:

- There were 33 exceptional applicants, who, despite being ineligible are tabulated under eligible since deAsra persuaded two Fis (Happy (29) and Saraswat(4)) to consider them. They are highlighted as ineligible, under the eligible categories.
- There were 78 applicants who did not go ahead with the program even though they were overall eligible (these are considered as eligible for overall analysis).
- There are 10 applications who were eligible for 2 or more FIs but we do not know the basis of eligibility for a specific FI.
- The evaluation is done for five FIs (LendingKart is not included since it had only 3 data points of relevance)
- Age categorization here is taken to be 17-30, 30-40, 40-50 and 50+

For the Gromor dataset:

- A loan default for more than 90 days makes it a Non Performing Asset (NPA) for the bank, hence analysis looks at 90+ days of default as bad.
- Loan amount for 131 observations have been categorized on basis of principle amount as <50,000 and >50,000 as per categorization of Mudra loan.
- Age has been categorized as 20-30,30-40,40-50 and 50-60 years.
- Size of enterprise (including the owner as employee) have been categorized as <5, 5-20 & >20.

MSME Loan Program

Overview of MSME Loan Programme

- Total number of participants: 3162
- Almost half of the entrepreneurs were eligible in at least one or more FIs (Eligible: 1549, Ineligible: 1613)
- Total number of borrowed loans: 181
- Total amount of loan disbursed: ₹ 109.26 lakhs

MSME Loan Programme

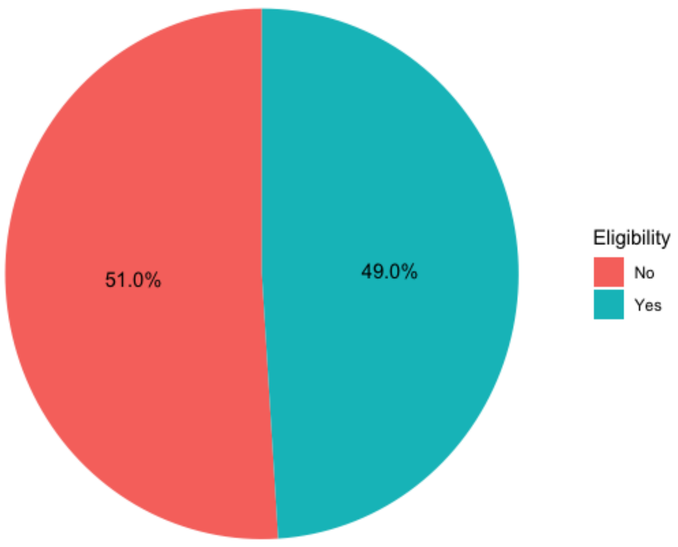
- For the purpose of this analysis the data (MSME Loan Applications) has been divided into the following sections:

- Demographics of applicants
    - Age
    - Gender
    - Region
- Business Profile
    - Business sector
    - Business age
    - Business type
    - Business address proof consent
- Other aspects
    - Loan amount required
    - Other variables (Coborrower, NEFT and bank statement)
- Strength of applications (eligibility and FI)
- FI wise individual criteria
- Regression Results

Overall Eligibility Count

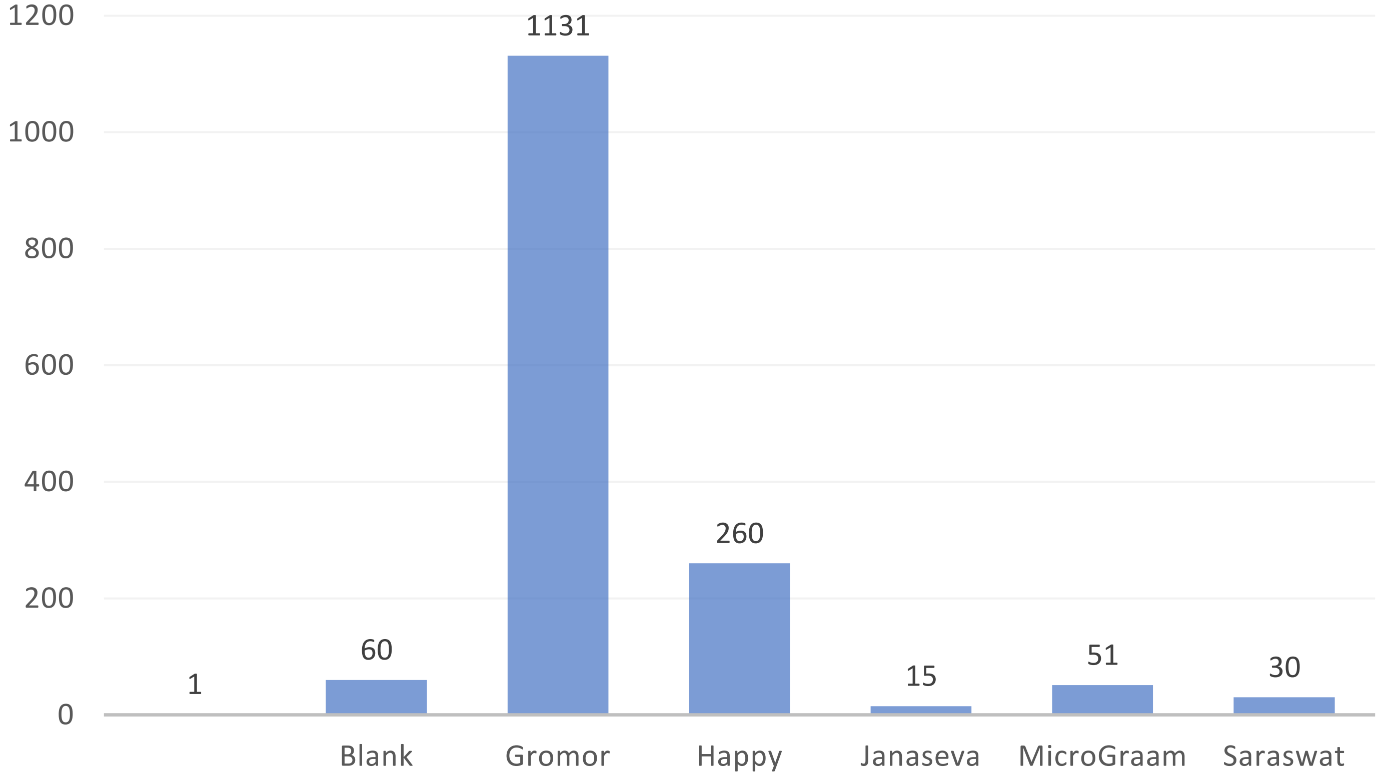
Ineligible: 1613
Eligible: 1549

Almost half of the entrepreneurs were eligible in at least one or more FIs.



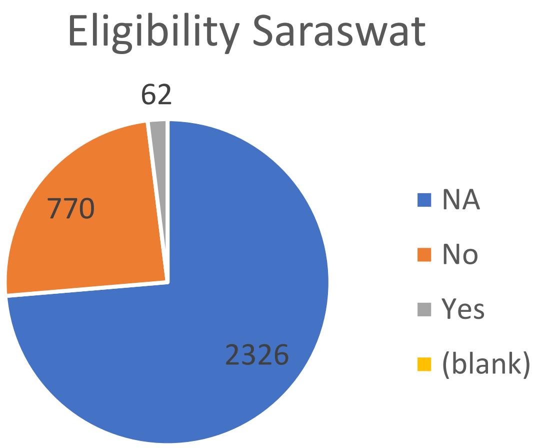
Eligibility criteria and Financial Institution

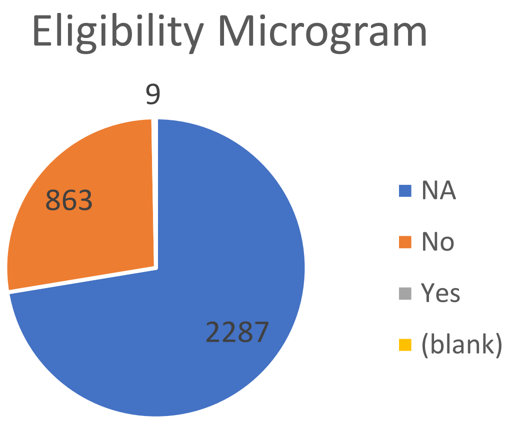
Majority have become eligible for Gromor, least for Janaseva

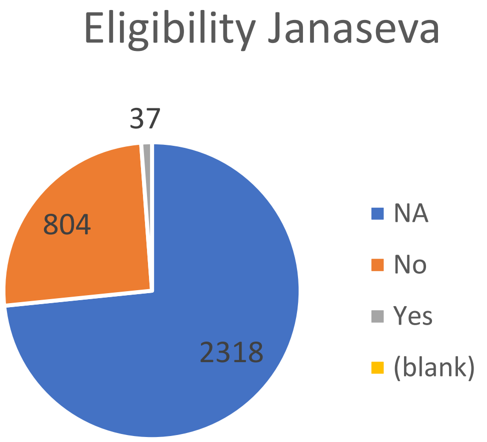


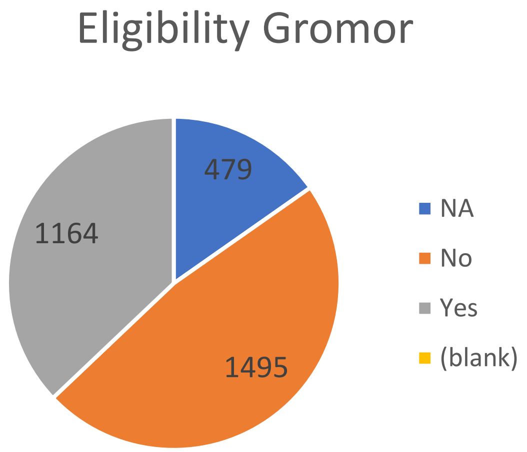
Eligibility distribution within each FI shows

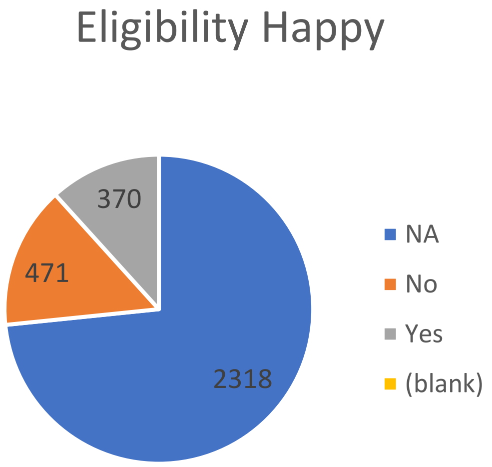
NA shows the MSME entrepreurs applied when the given FI was not onboarded. Hence this does not indicate a comparative strictness in qualification.





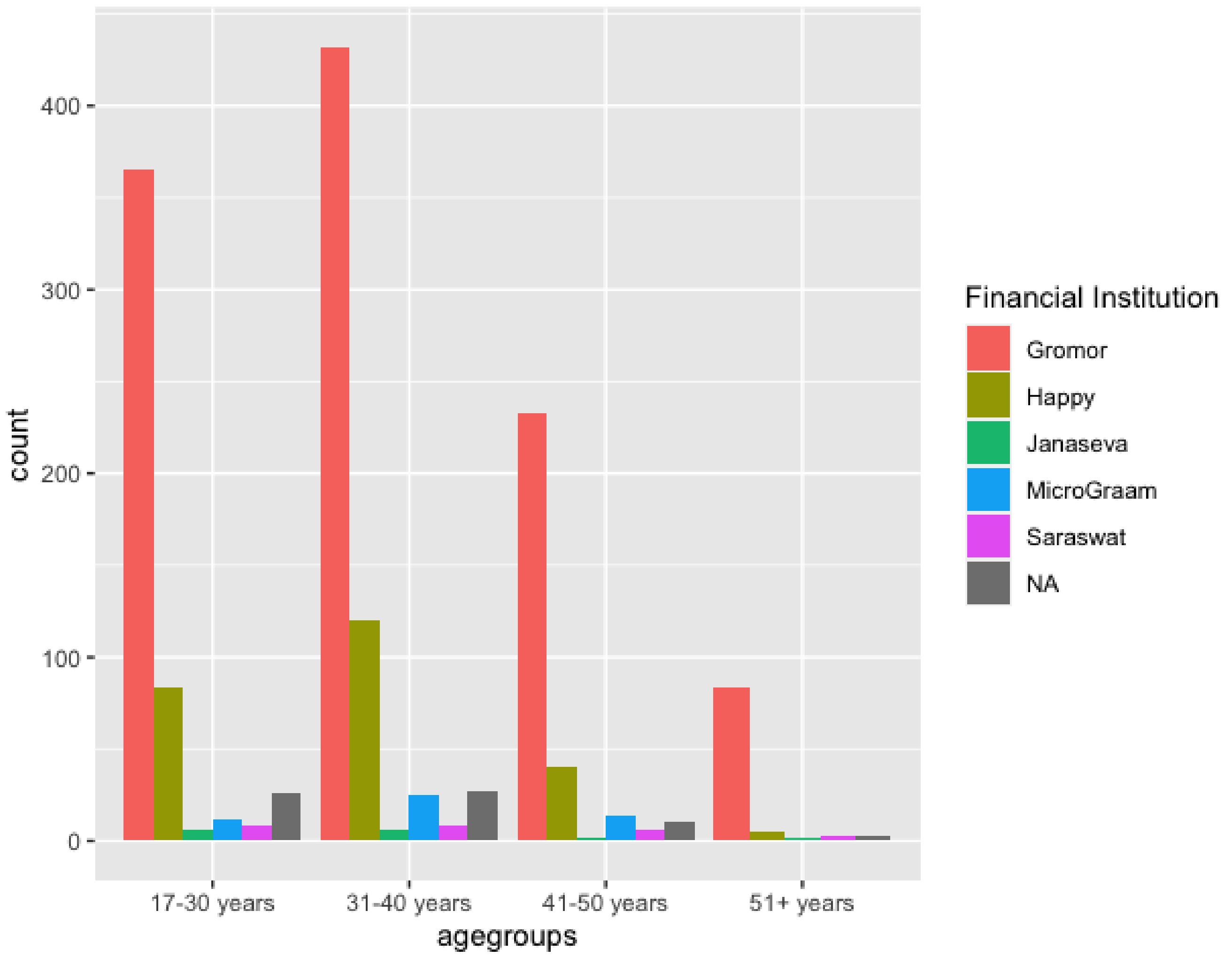






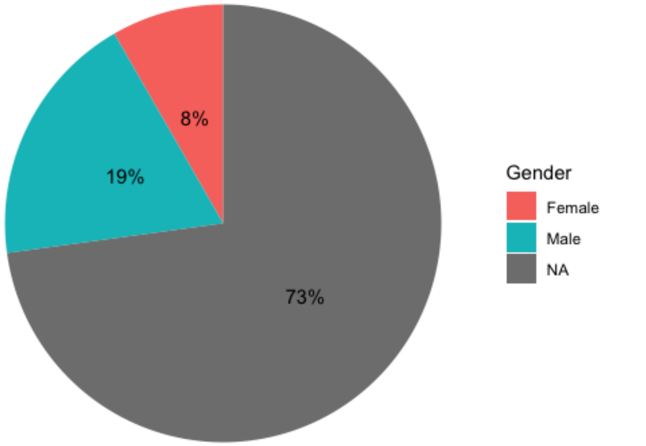
Age

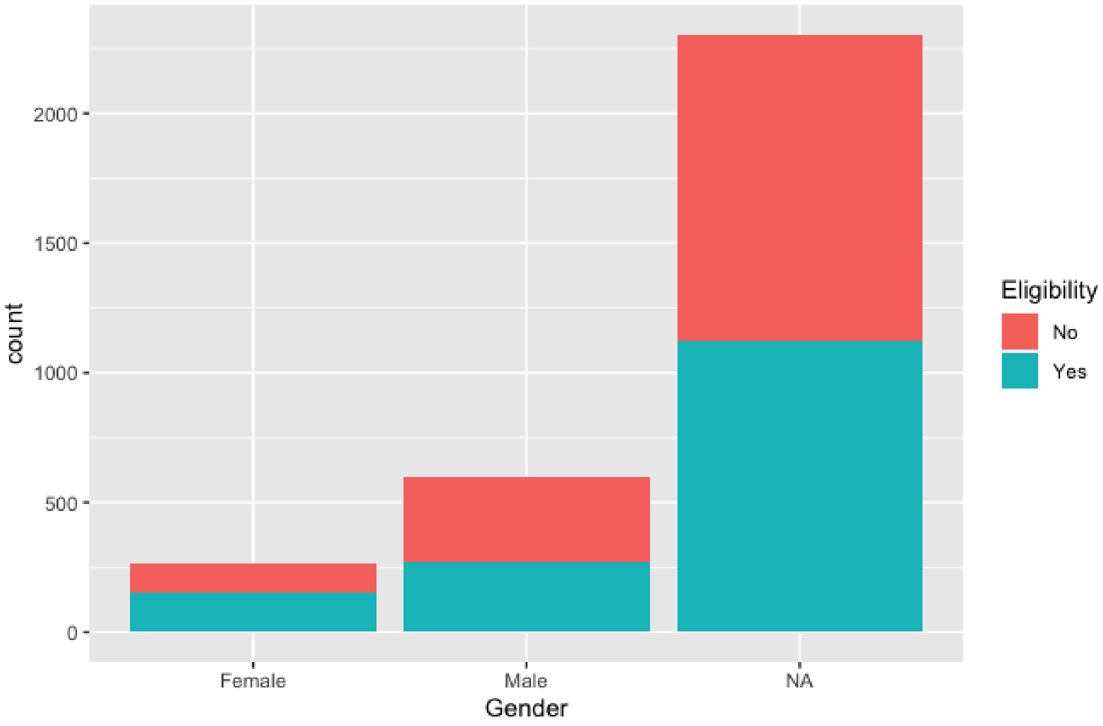
N=1531

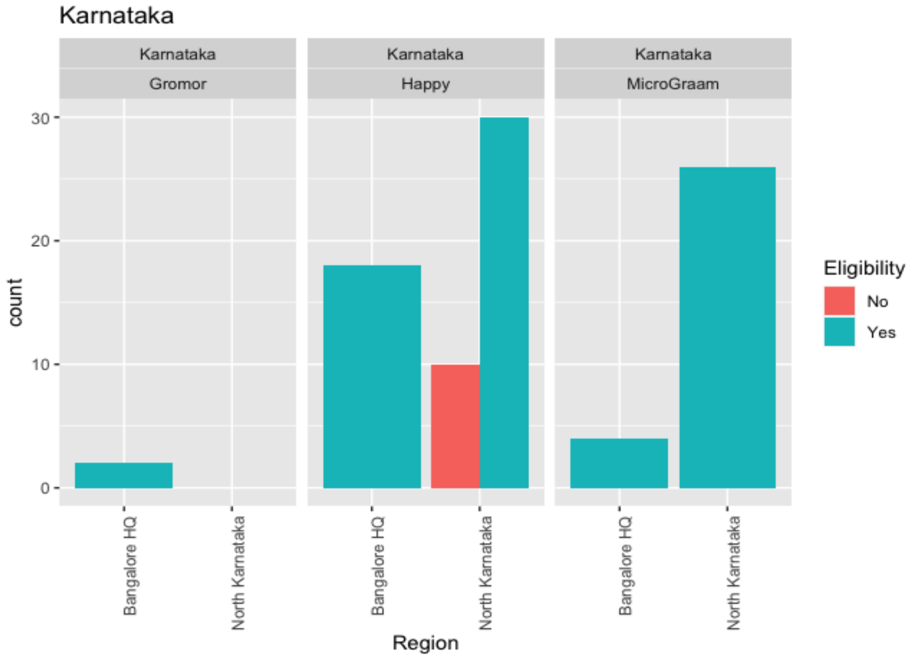


Gender

Almost 70% didn't fill this entry

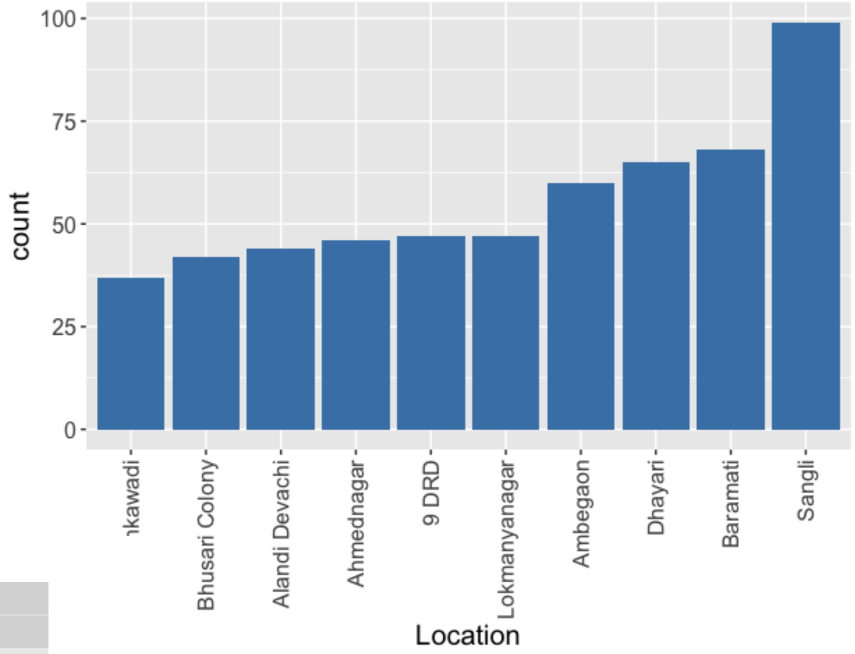


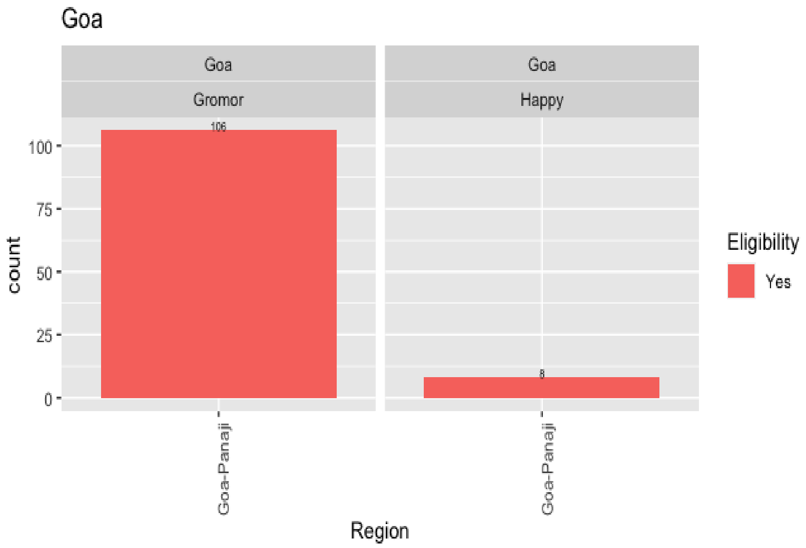


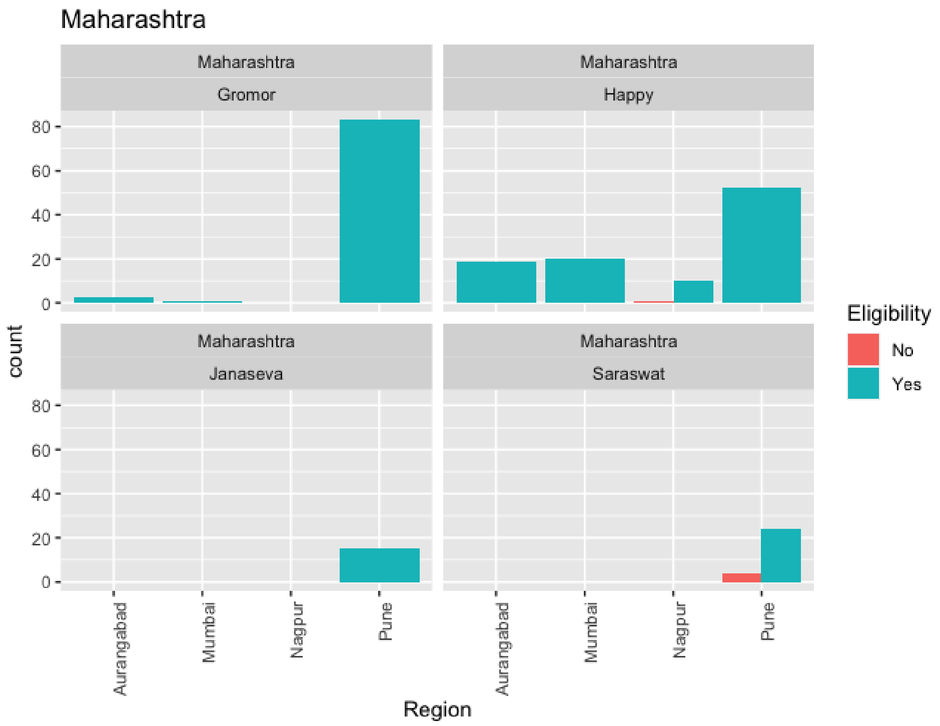


N (Maharashtra): 232
N (Goa): 114
N (Karnataka): 90

Top 10 pin-codes with most applications (all from Pune district)







Regional Share

Business Profiles

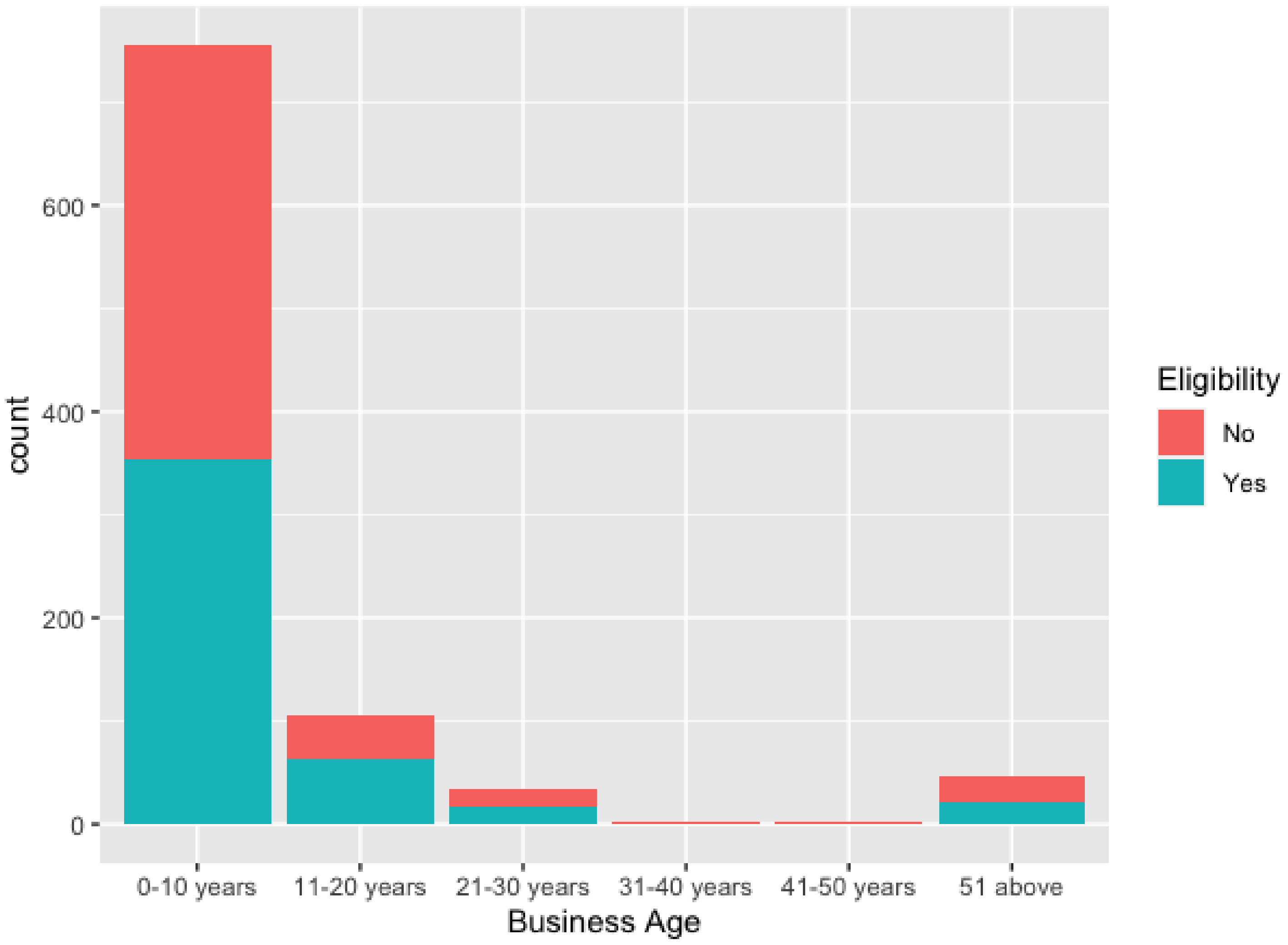
N = 410

Loans are distributed over sectors.

Gromor is the only FI catering to almost all business types.

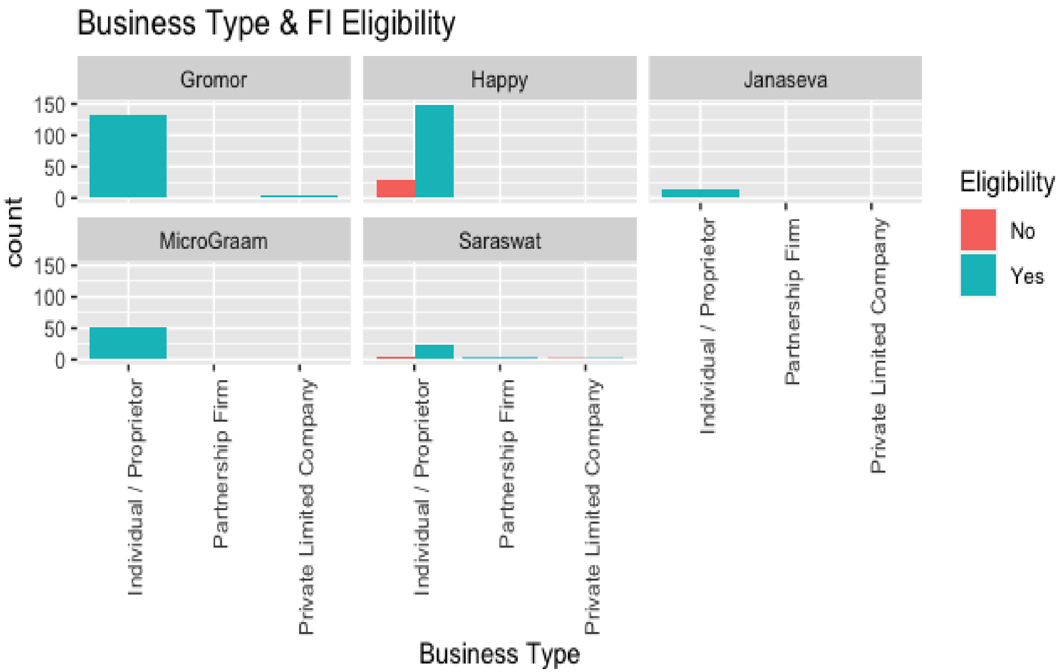
Business Age

N=944

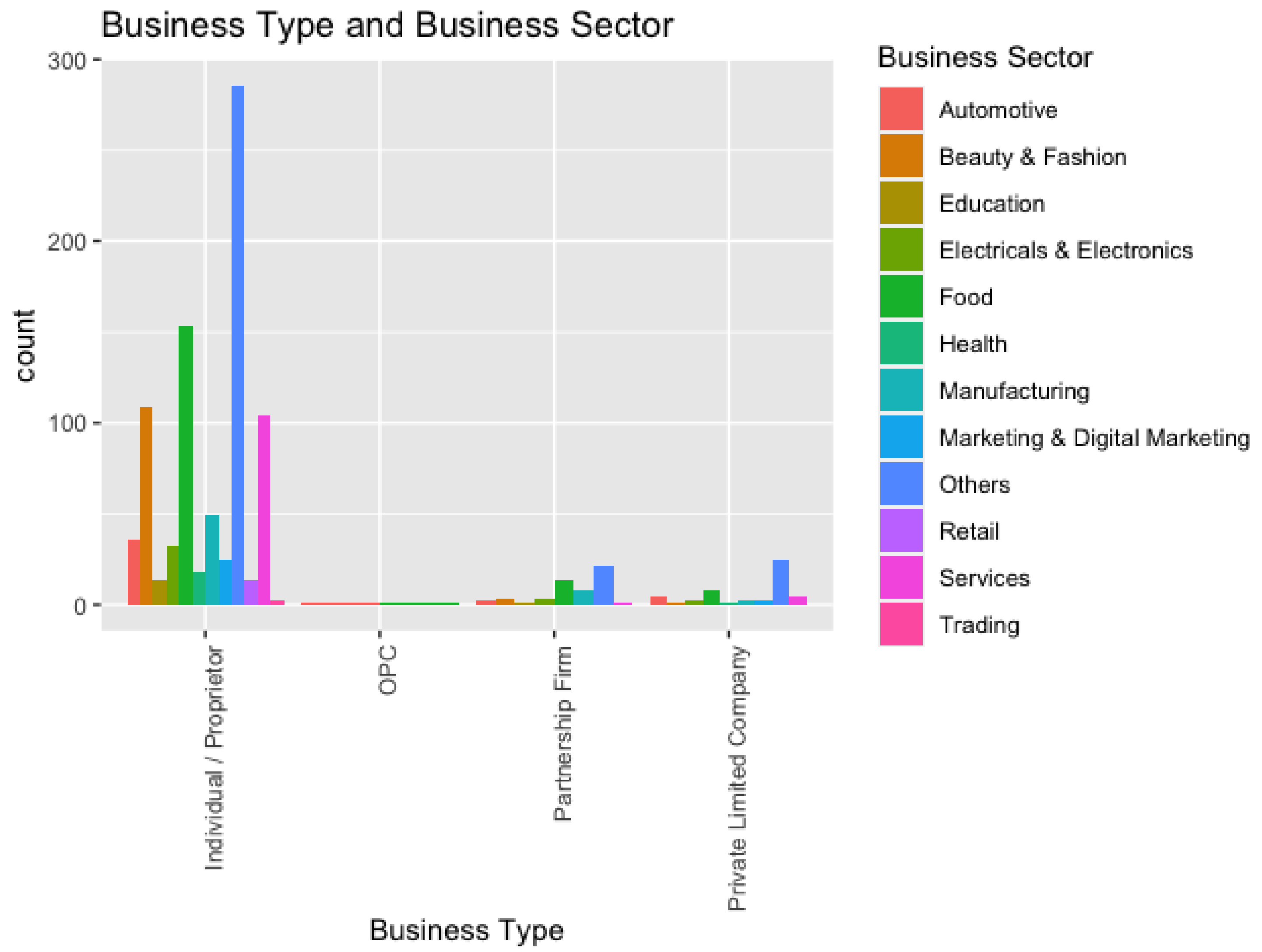


Business Type

N=410

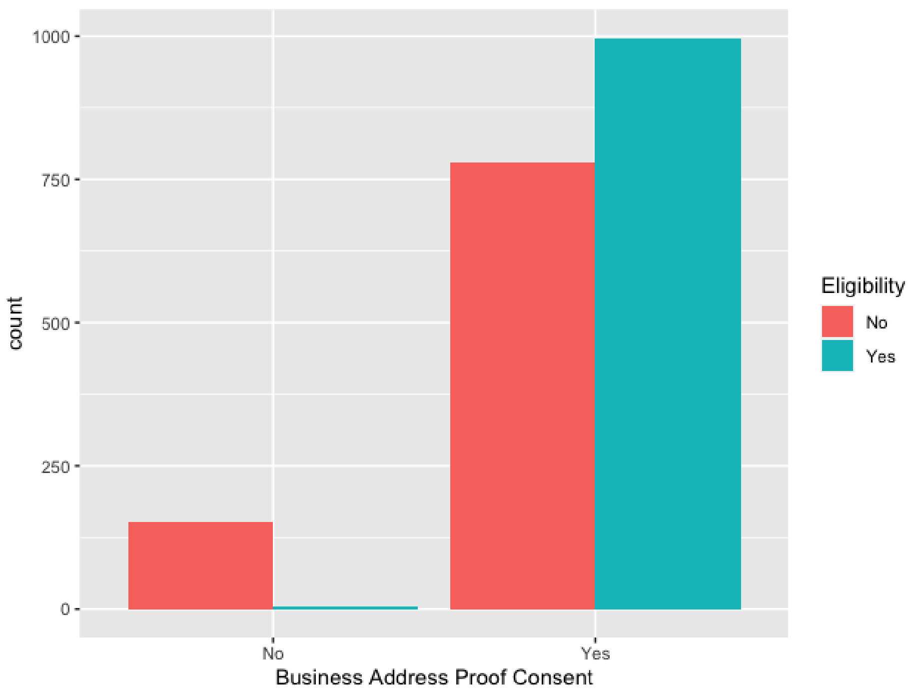


Business Type and Sector (N=944)

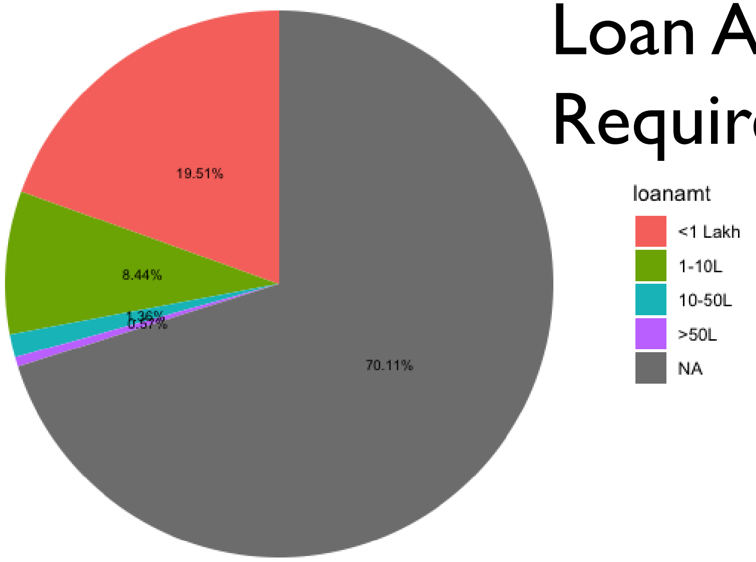


Business Address Proof

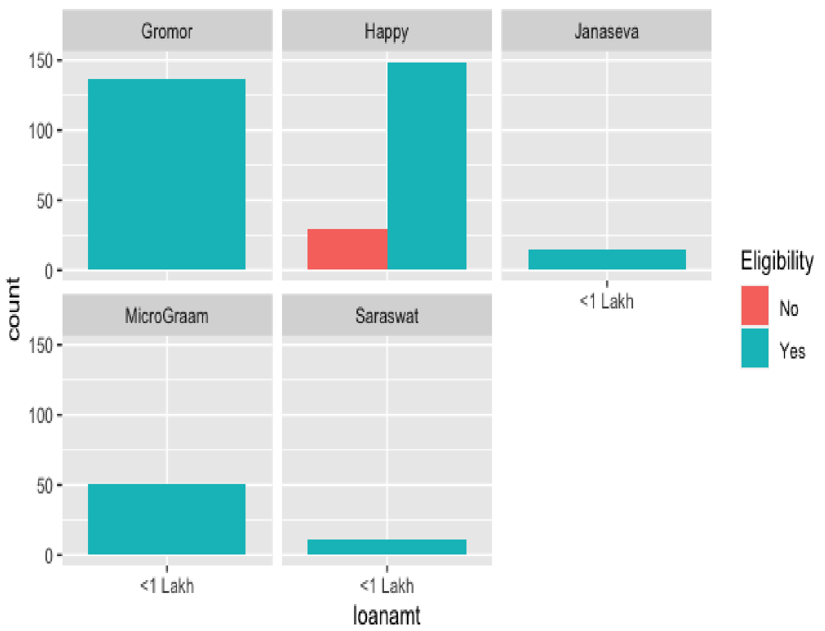
N=1932



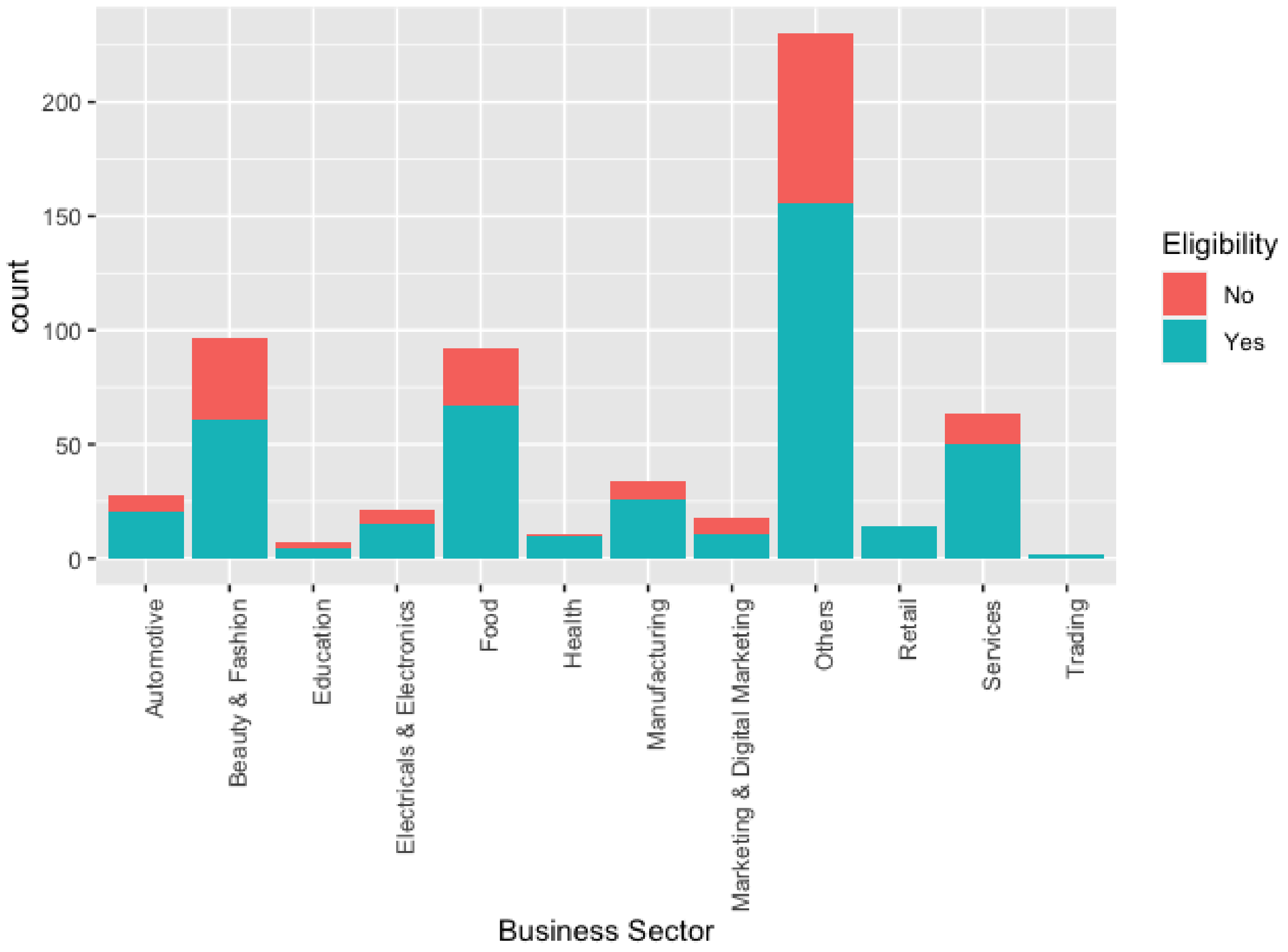
Loan Amount Required



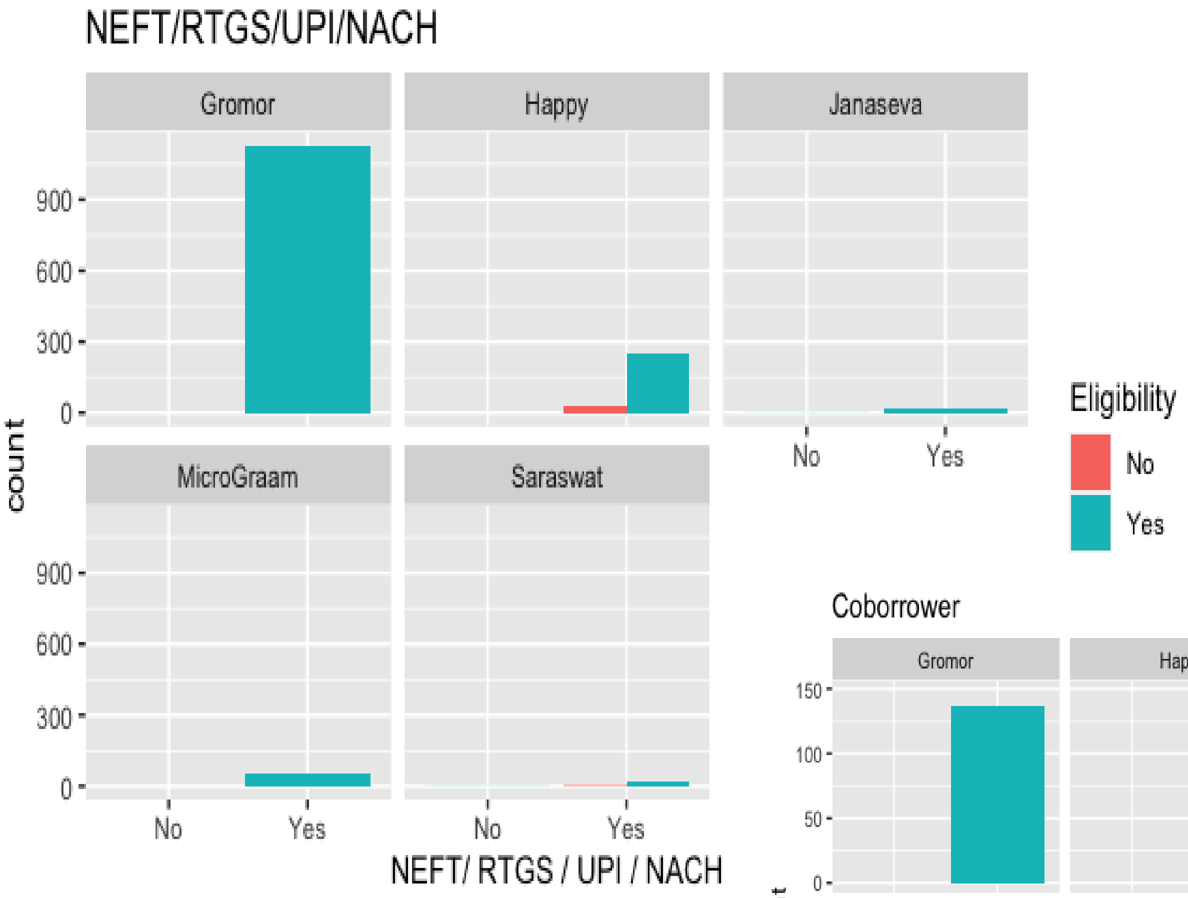
Loan Amount < 1 Lac Eligibility (N=392)



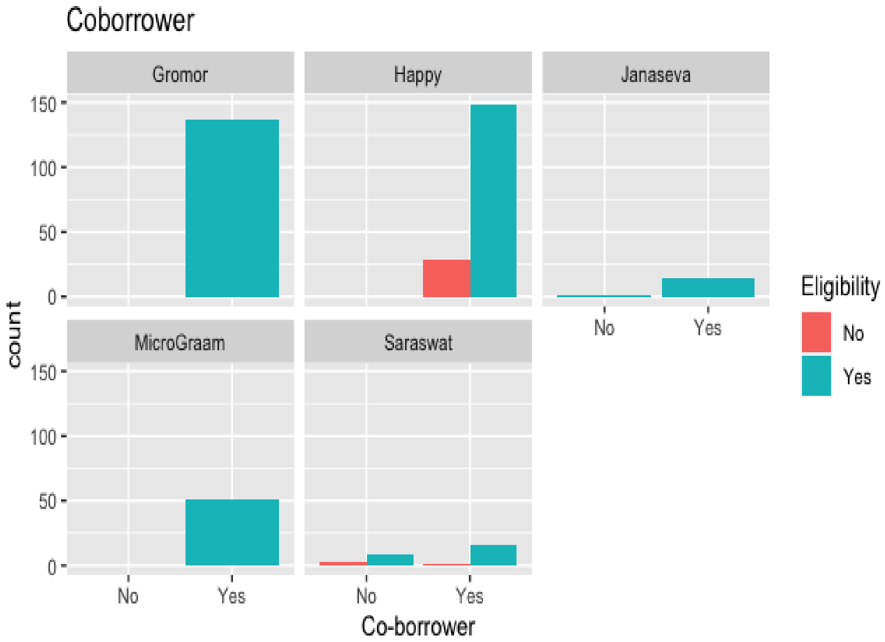
Loan Amount (< 1 lac) Eligibility and Business Sectors

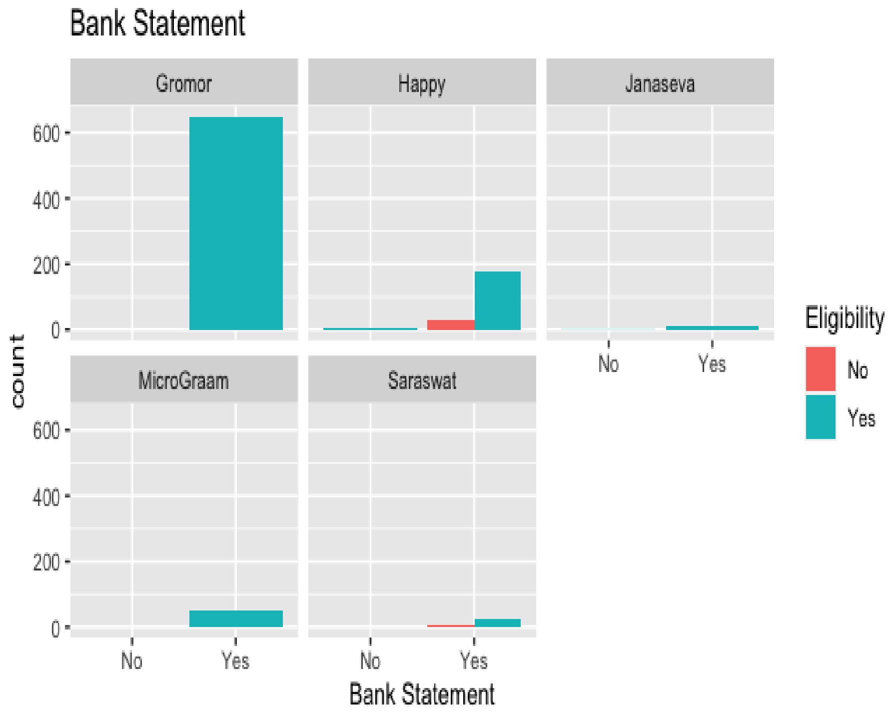


N = 1504



N = 410





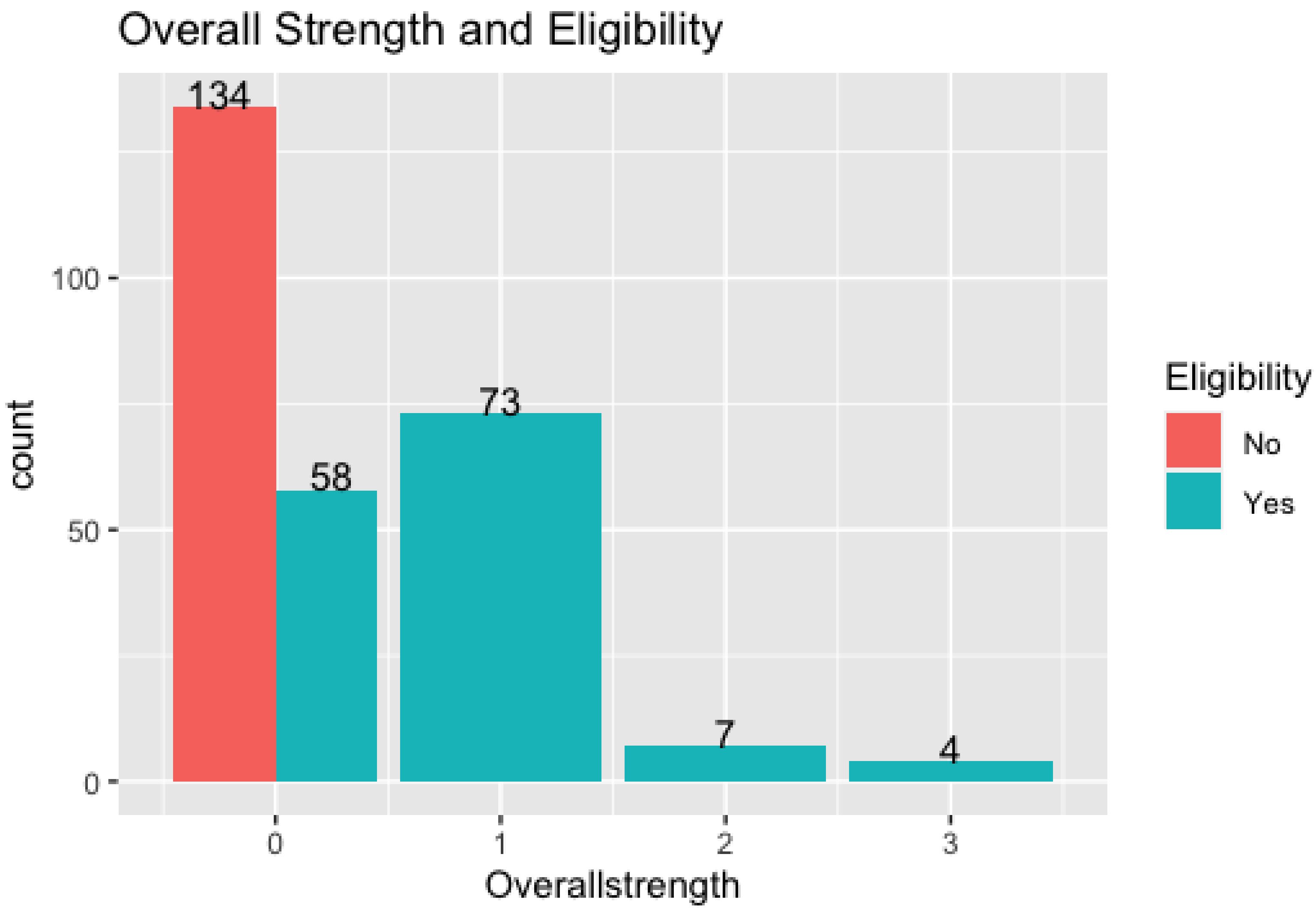
N =957

Other Variables

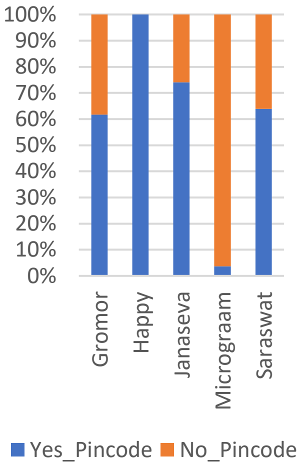
Strength of Applicants

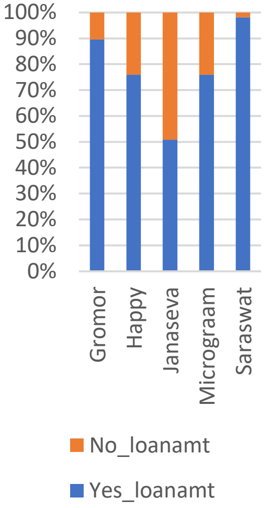
This index measures the eligibility strength of the applicants, by estimating total number of FIs that they were found eligible. The numbers on x-axis mentions mean-number of FIs they were eligible for.

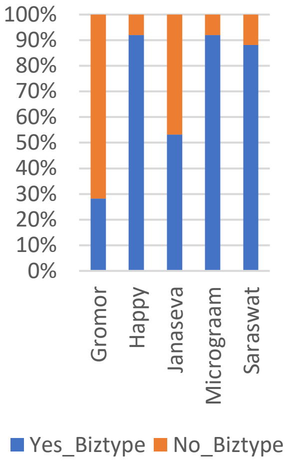
Only the application filed in September 2020 (when all the four FIs were on-board) were taken into consideration.

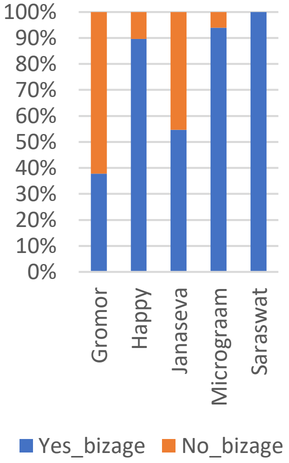


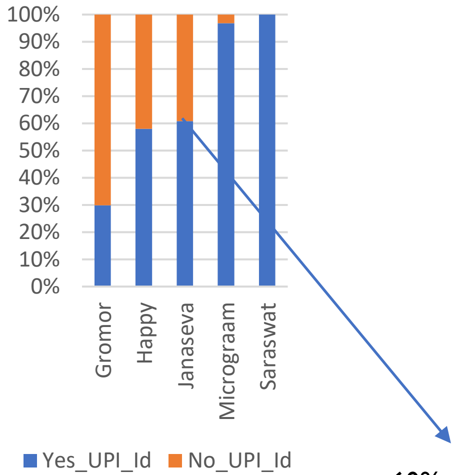
Analysis of eligibility criteria against each FI (how difficult are these criteria for entrepreneurs)



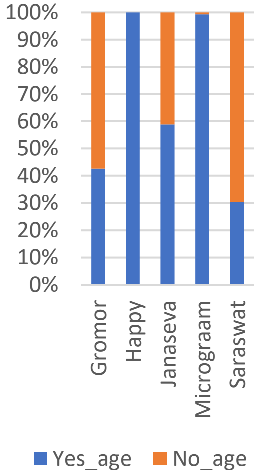


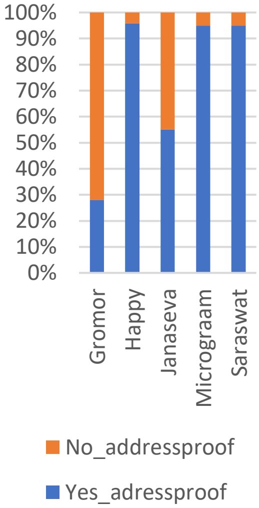


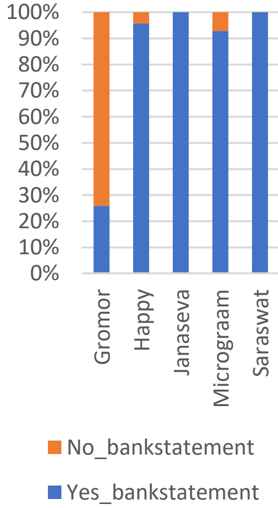


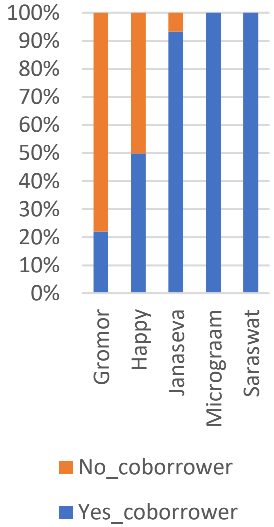


60% of the total applicants were eligible for Janaseva's UPI Id category



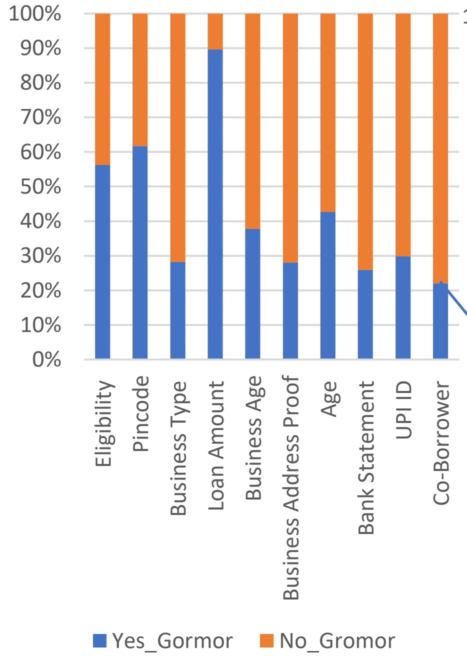


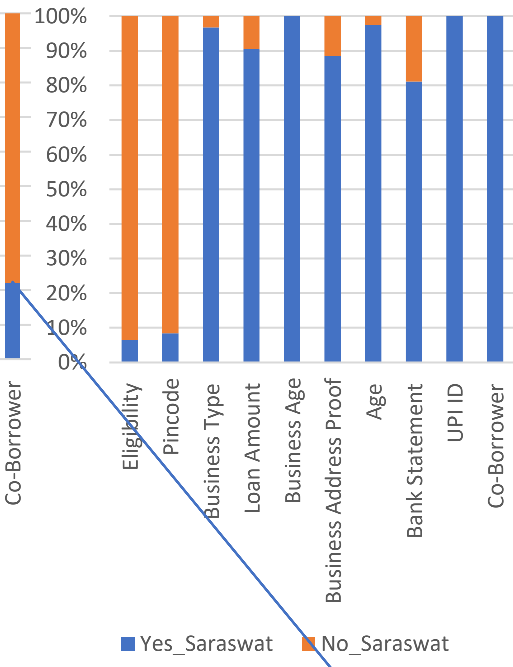


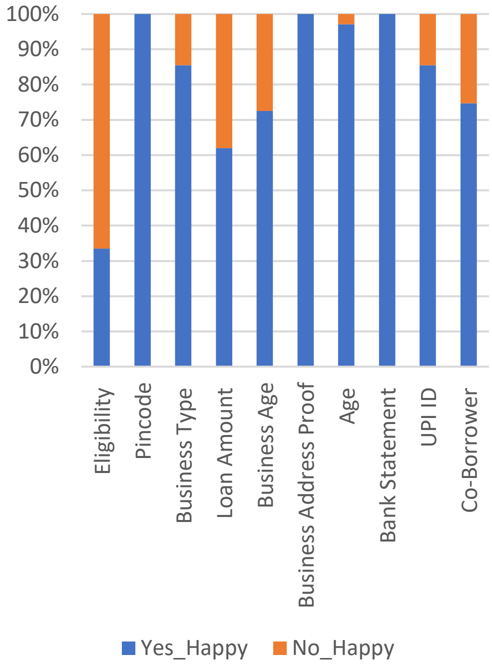


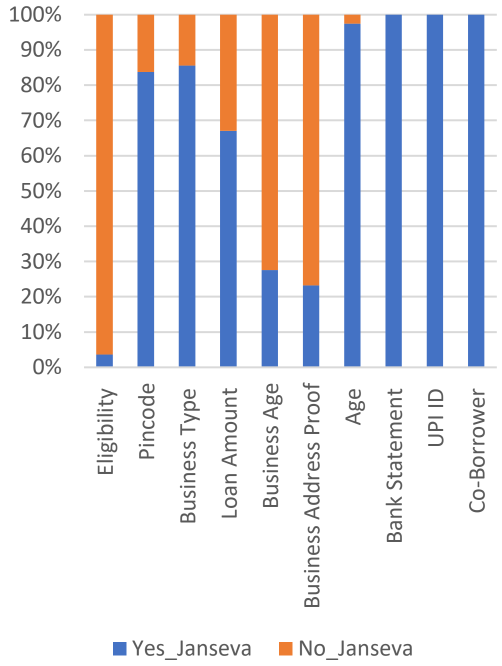
There were 96 applicants under NA for age criteria across all 5 FIs

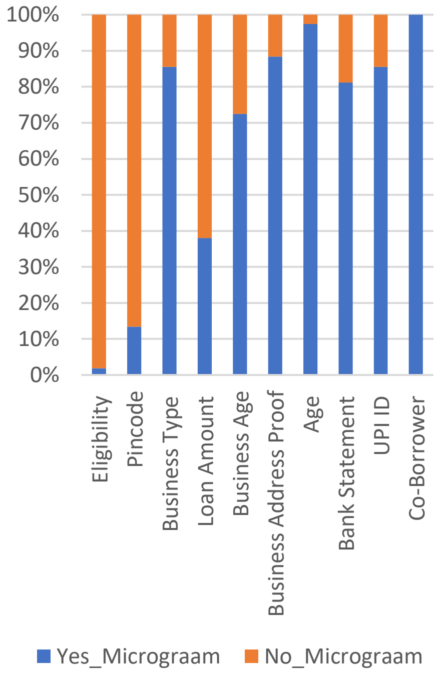
FI-Analysis against various factors that determine eligibility (which FIs were generally more difficult in their criteria)











Almost 20% of the total applicants were eligible for Gromor's co-borrower category

Eligibility for Gromor and details shared by applicant.

Eligibility	Yes	
Financial Institution	Gromor	
Gromor_pincode	(All)	
Gromor_Loan_Amount	(All)	
Gromor_Age	(All)	
Gromor_Business_Age	(All)	
Gromor_Co_Borrower	(All)	
Gromor_Bank_Statement	(All)	
Gromor_Business_Address_Proof	(All)	
Gromor_UPI_Id	(All)	
Gromor_Business_Type	(All)	
Count of EligibilityGromor	% Diff	
1131	0	

- With eligibility as 'Yes' and FI as Gromor, there are 1131 applicants.
- We check the relevance of each 'detail' by changing it to 'NO' in filters and look at the difference in total eligible numbers for Gromor.
- Variables are designated as red orange and yellow based on decreasing criticality to get eligible at gromor.
- Pincode, Loan Amount, Applicants Age are very critical for getting eligibility at Gromor: Red
- Business age is relatively less critical: Orange
- Co Borrower, bank statement, business address proof, UPI ID and business type are least critical in being eligible at Gromor: Yellow.

Eligibility	Yes
Financial Institution	Gromor
Gromor_pincode	No
Gromor_Loan_Amount	(All)
Gromor_Age	(All)
Gromor_Business_Age	(All)
Gromor_Co_Borrower	(All)
Gromor_Bank_Statement	(All)
Gromor_Business_Address_Proof	(All)
Gromor_UPI_Id	(All)
Gromor_Business_Type	(All)
Count of EligibilityGromor	% Diff
3	99.73475

Eligibility	Yes
Financial Institution	Gromor
Gromor_pincode	(All)
Gromor_Loan_Amount	No
Gromor_Age	(All)
Gromor_Business_Age	(All)
Gromor_Co_Borrower	(All)
Gromor_Bank_Statement	(All)
Gromor_Business_Address_Proof	(All)
Gromor_UPI_Id	(All)
Gromor_Business_Type	(All)
Count of EligibilityGromor	% Diff
	100

Eligibility	Yes
Financial Institution	Gromor
Gromor_pincode	(All)
Gromor_Loan_Amount	(All)
Gromor_Age	No
Gromor_Business_Age	(All)
Gromor_Co_Borrower	(All)
Gromor_Bank_Statement	(All)
Gromor_Business_Address_Proof	(All)
Gromor_UPI_Id	(All)
Gromor_Business_Type	(All)
Count of EligibilityGromor	% Diff
695	38.54996

Eligibility	Yes
Financial Institution	Gromor
Gromor_pincode	(All)
Gromor_Loan_Amount	(All)
Gromor_Age	(All)
Gromor_Business_Age	No
Gromor_Co_Borrower	(All)
Gromor_Bank_Statement	(All)
Gromor_Business_Address_Proof	(All)
Gromor_UPI_Id	(All)
Gromor_Business_Type	(All)
Count of EligibilityGromor	% Diff
737	34.83643

Eligibility	Yes
Financial Institution	Gromor
Gromor_pincode	(All)
Gromor_Loan_Amount	(All)
Gromor_Age	(All)
Gromor_Business_Age	(All)
Gromor_Co_Borrower	No
Gromor_Bank_Statement	(All)
Gromor_Business_Address_Proof	(All)
Gromor_UPI_Id	(All)
Gromor_Business_Type	(All)
Count of EligibilityGromor	% Diff
994	12.11317

Eligibility	Yes
Financial Institution	Gromor
Gromor_pincode	(All)
Gromor_Loan_Amount	(All)
Gromor_Age	(All)
Gromor_Business_Age	(All)
Gromor_Co_Borrower	(All)
Gromor_Bank_Statement	No
Gromor_Business_Address_Proof	(All)
Gromor_UPI_Id	(All)
Gromor_Business_Type	(All)
Count of EligibilityGromor	% Diff
994	12.11317

Note: Remaining 3 variables also show only 12% difference when set to 'No'.

Criticality of Variables for Other FIs Eligibility

- Janaseva and Saraswat need all variables to fit their requirements.
- Microgram shows some leniency with pin-code data only.
- After Gromor, only Happy shows some leniency with few variables highlighted in orange.

Eligibility	Yes
Financial Institution	Janaseva
Janaseva_pincode	(All)
Janaseva_Loan_Amount	(All)
Janaseva_Age	(All)
Janaseva_Business_Age	(All)
Janaseva_Co_Borrower	(All)
Janaseva_Bank_Statement	(All)
Janaseva_Business_Address_Proof	(All)
Janaseva_UPI_Id	(All)
Janaseva_Business_Type	(All)
Count of EligibilityJanaseva	% Diff
14	0

Eligibility	Yes
Financial Institution	MicroGraam
Micrograam_pincode	(All)
MicroGraam_Loan_Amount	(All)
MicroGraam_Age	(All)
MicroGraam_Business_Age	(All)
MicroGraam_Co_Borrower	(All)
MicroGraam_Bank_Statement	(All)
MicroGraam_Business_Address_Proof	(All)
MicroGraam_UPI_Id	(All)
MicroGraam_Business_Type	(All)
Count of EligibilityMicroGraam	% Diff
51	0

Eligibility	Yes
Financial Institution	Saraswat
Saraswat_pincode	(All)
Saraswat_Loan_Amount	(All)
Saraswat_Age	(All)
Saraswat_Business_Age	(All)
Saraswat_Co_Borrower	(All)
Saraswat_Bank_Statement	(All)
Saraswat_Business_Address_Proof	(All)
Saraswat_UPI_Id	(All)
Saraswat_Business_Type	(All)
Count of EligibilitySaraswat	% Diff
30	0

Eligibility	Yes
Financial Institution	Happy
Happy_pincode	(All)
Happy_Loan_Amount	(All)
Happy_Age	(All)
Happy_Business_Age	(All)
Happy_Co_Borrower	(All)
Happy_Bank_Statement	(All)
Happy_Business_Address_Proof	(All)
Happy_UPI_Id	(All)
Happy_Business_Type	(All)
Count of EligibilityHappy	% Diff
260	0

Some Inferences

- The overall entrepreneur sample is good, with almost 50% being eligible.
- Female applicants and those in the age of 30-40 yrs stand higher chance of eligibility for loan.
- Individual proprietary enterprises must be encouraged more to apply for loans.
- Loan eligibility is high for small amounts of loans. This necessitates training.

Methodology of regression

- Logistic regression was used for analysis of the entire dataset, with eligibility variables has been considered as the dependent variable.
- We ran two sets of iterations: one set which included individual eligibility criteria for each FI and another without them.
- The regression attached in the following slide offers the best model out of all iterations that we ran
- The other 7 models (total number of models= 8) were rejected on the basis of confusion matrix and chi square values of the model.

$$Probability\ (Eligibility) = \frac{e^{\beta_0 + \beta_1\ `Business\ Type` + \beta_2\ Co{-}borrower + \beta_3\ `Bank\ Statement` + \beta_4\ Gromorpincode + \beta_5\ Mircrograampincode + \beta_6\ Saraswatpincode + \beta_7\ Jasevapincode + \beta_8\ HappyLoanAmount + \beta_9\ Jaseva\_Loan\ Amount + \beta_{10} HappyBusiness\_Age + U_i}}{1 + e^{\beta_0 + \beta_1\ `Business\ Type` + \beta_2\ Co{-}borrower + \beta_3\ `Bank\ Statement` + \beta_4\ Gromorpincode + \beta_5\ Mircrograampincode + \beta_6\ Saraswatpincode + \beta_7\ Jasevapincode + \beta_8\ HappyLoanAmount + \beta_9\ JasevaLoan\ Amount + \beta_{10} HappyBusinessAge + U_i}}$$

Regressing Eligibility against Other factors

The significant variables of this model are: business type, business address proof consent*, co-borrower***, bank statement***, Gromor_pincode***, MicroGram_pincode*, Saraswat_pincode***, Janseva_pincode*, Happy_loanamount***, Janseva_loanamount***, Happy_businessage***

```
Call:
glm(formula = Eligibility ~ `Business Type` + `Business Address Proof Consent` +
    `Co-borrower` + `Bank Statement` + Gromor_pincode + Micrograam_pincode +
    Saraswat_pincode + Jaseva_pincode + Happy_Loan_Amount + Jaseva_Loan_Amount +
    Happy_Business_Age, family = "binomial", data = train)

Deviance Residuals:
    Min        1Q    Median        3Q       Max
-3.3302   -0.1454   -0.0006    0.3927    3.4598

Coefficients:
                                       Estimate Std. Error z value Pr(>|z|)
(Intercept)                            -15.5852     1.5486 -10.064  < 2e-16 ***
`Business Type`OPC                      -5.4770   911.6053  -0.006  0.99521
`Business Type`Partnership Firm         -4.7221     0.9343  -5.054 4.32e-07 ***
`Business Type`Private Limited Company  -2.6784     0.8772  -3.053  0.00226 **
`Business Address Proof Consent`1        2.2195     0.8720   2.545  0.01092 *
`Co-borrower`1                           2.8808     0.4742   6.074 1.24e-09 ***
`Bank Statement`1                        3.2844     0.6586   4.987 6.12e-07 ***
Gromor_pincode1                          1.9767     0.3641   5.429 5.68e-08 ***
Micrograam_pincode1                      1.0261     0.4711   2.178  0.02941 *
Saraswat_pincode1                        2.9408     0.7290   4.034 5.48e-05 ***
Jaseva_pincode1                          1.1590     0.5839   1.985  0.04717 *
Happy_Loan_Amount1                       1.6618     0.4036   4.117 3.84e-05 ***
Jaseva_Loan_Amount1                      5.0032     0.5619   8.905  < 2e-16 ***
Happy_Business_Age1                      2.7441     0.5556   4.939 7.87e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1032.4  on 744  degrees of freedom
Residual deviance:  330.0  on 731  degrees of freedom
  (1780 observations deleted due to missingness)
AIC: 358

Number of Fisher Scoring iterations: 14
```

Some Inferences

The log odds of OPC being eligible is 5.47 times lower, of a partnership firm being eligible is 4.72 times lower, and of a private limited firm being eligible is 2.68 times lower than individual / proprietorship firm.

Individual/Proprietorship Firm is more likely to get loans

The log odds of eligibility of an applicant who has business address proof is 2.22 times higher, of an applicant who has co-borrower is 2.88 times higher and of an applicant who has bank statement is 3.28 times higher than those who do not have either of these.

Applicant with business address proof, co-borrower and bank statement are more likely to get the loan

The log odds of eligibility of an applicant who is eligible for Gromor, the pincode is 1.98 times higher, the one who is eligible for MicroGraam, the pincode is 1.03 times higher, the one who is eligible for Saraswat, the pincode is 2.94 times higher and the one who is eligible for Janaseva, the pincode is 1.16 times higher than those who do not.

Pincode is important for Gromor, MicrGraam, Saraswat and Janaseva, and most for Saraswat

The log odds of eligibility of an applicant who is eligible for Happy loan amount is 1.66 times higher, for Janaseva is 5 times higher than those who do not.

Loan Amount matters most for Janaseva, and somewhat to Happy

The log odds of eligibility of an applicant who is eligible for Happy business age is 2.74 times higher than those who do not.

Business Age is important to Happy.

Gromor Data

Data Analysis

Question:

1. Are defaulting of loans related to Age of loanee and original principal amount of loan?
2. How are factors like Age, Gender, Purpose of loan, City, Enterprise size and original principal amount distributed across the category of days past due for these loans?

Note:

Days past due and Total overdue amount are considered as outcome variable for Loans. The lesser these two are, better the loans will be. Age, Number of people employed and principal amount of loan are taken as independent variables.

Summary of variables					
Variable	**Observations**	**Mean**	**Std. Dev.**	**Min**	**Max**
Age	131	36.67	7.93	24	60
No. of people employed	127*	2.46	3.56	0	28
Principle amount of Loan	131	64786.26	25864.14	10000	150000
Current Days Past Due	131	65.82	61.87	0	219
Total Over due amount	131	13030.43	13790.63	0	58469

Note*: *Observations for number of people employed had 4 entities with missing data.*

Correlation table					
	Age	No. of people employed	Principle amount of Loan	Current Days Past Due	Total Over due amount
Age	1				
No. of people employed	-0.0384	1			
	0.6679				
Principle amount of Loan	-0.0195	0.103	1		
	0.8251	0.2492			
Current Days Past Due	-0.1787*	0.1265	-0.022	1	
	0.0411	0.1564	0.8027		
Total Over due amount	-0.1915*	0.1706	0.3199*	0.8764*	1
	0.0284	0.0552	0.0002	0.000	

Age shows inverse weak relation with current days past due and total overdue amount. Hence although weakly related, a higher age might show less defaulting possibilities for loans.

Total overdue amount has positive relation with principal amount and current days past due. Hence higher principal amounts and higher defaults tend to show higher total overdue amount. Default in days past due is strongly related whereas principal amount relatively isn't.

Source	SS	df	MS	No. of obs	=	131
				F(3, 127)	=	322.48
Model	2.19E+10	3	7.28E+09	Prob > F	=	0
Residual	2.87E+09	127	22590302	R-squared	=	0.884
				Adj R-squared	=	0.8812
Total	2.47E+10	130	1.9E+08	Root MSE	=	4752.9
Total Over due amount	Coef.	Std. Err.	t	P>t	[95% Con	Interval]
Current Days Past Due	195.9021	6.850222	28.6	0.0000	182.347	209.4575
Age	-48.3851	53.41084	-0.91	0.367	-154.08	57.30528
Principle amount	0.180604	0.016126	11.2	0.0000	0.14869	0.212514
Constant	-9789.51	2396.27	-4.09	0.0000	-14531	-5047.72

Regression table for Total overdue amount

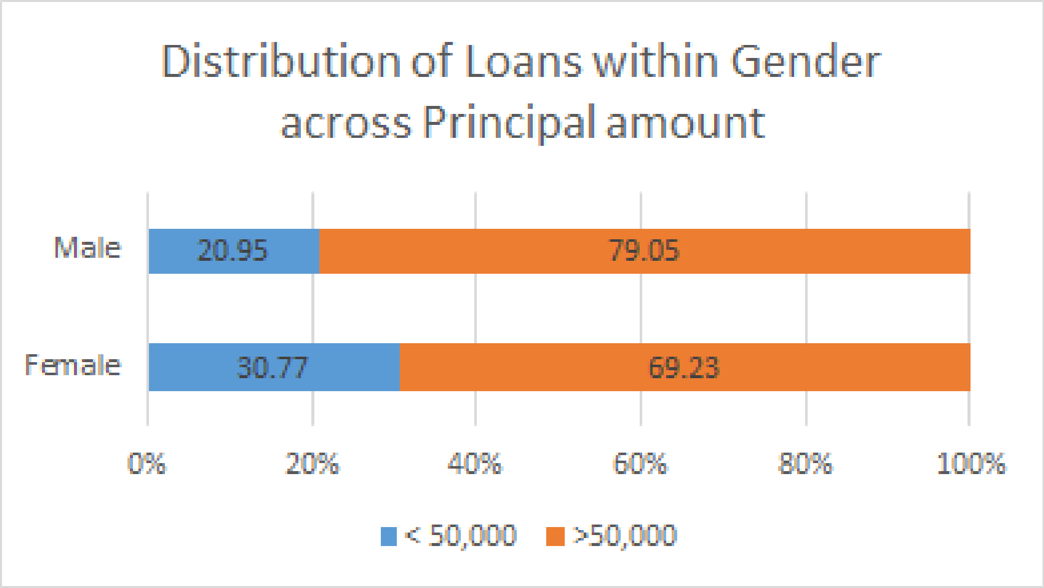
- P of F test is less than 0.1, this shows our model is statistically significant.
- R squared value is also closer to 1, hence shows our findings are valid.
- P value of t test is not significant for Age.
- P value is significant for days past due. 1 unit change in days past due has high value change for total overdue amount.
- Principal amount is significant as per t test, but the relation factor is quite low. **Thus the higher loans need not necessarily result in loan defaults**.

Source	SS	df	MS	No. of obs	=	131
				F(3, 127)	=	283.23
Model	432923.6	3	144307.9	Prob > F	=	0
Residual	64707.96	127	509.5115	R-squared	=	0.87
				Adj R-square	=	0.8669
Total	497631.6	130	3827.935	Root MSE	=	22.572
Current Days Past Due	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Age	0.025923	0.254464	0.1	0.919	-0.4776151	0.529462
Principle amount of Loan	-0.00081	8.09E-05	-9.97	0.0000	-0.0009662	-0.00065
Total Over due amount	0.004419	0.000155	28.6	0.0000	0.0041127	0.004724
Constant	59.52242	10.89186	5.46	0.0000	37.96939	81.07545

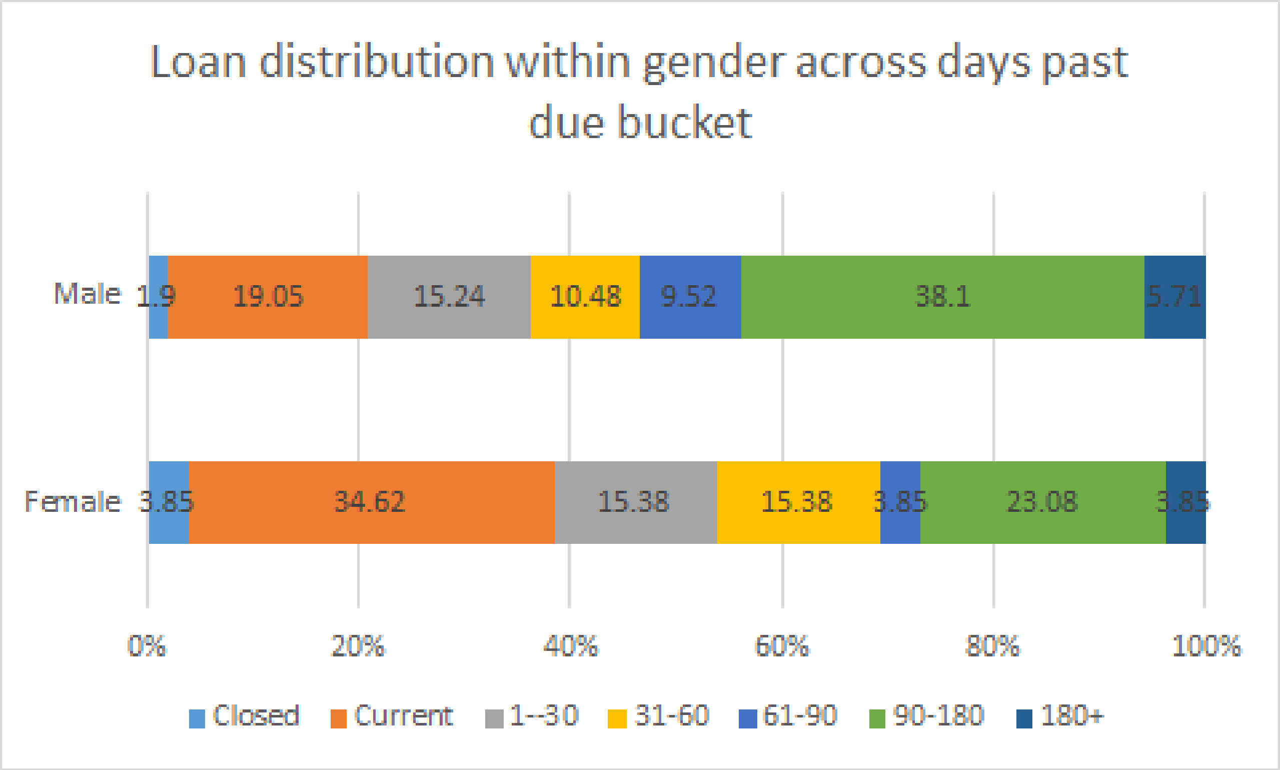
Regression table for Current days past due

- P of F test is less than 0.1, this shows our model is statistically significant.
- R squared value is also closer to 1, hence shows our findings are valid.
- P value of t test is not significant for Age.
- P value is significant for principal amount of loan, showing a higher value of loan amount may have a bleak chance of lower days past due.
- Total overdue amount is also significant as per t test, but the relation factor is quite low. Thus **amount that is due in a loan affects days past due in a bleak manner.**

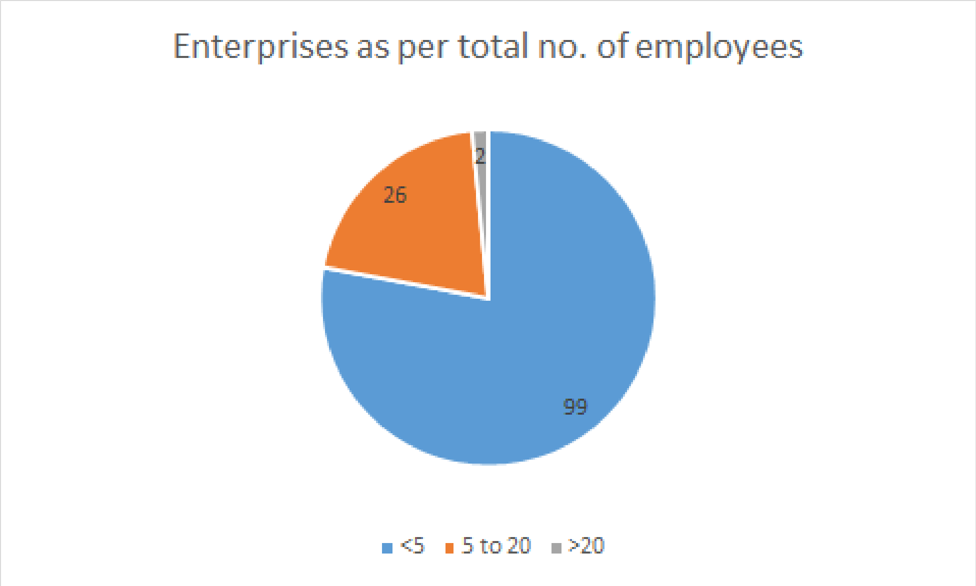
Distribution Charts



Female loanees have taken more number of smaller loans.

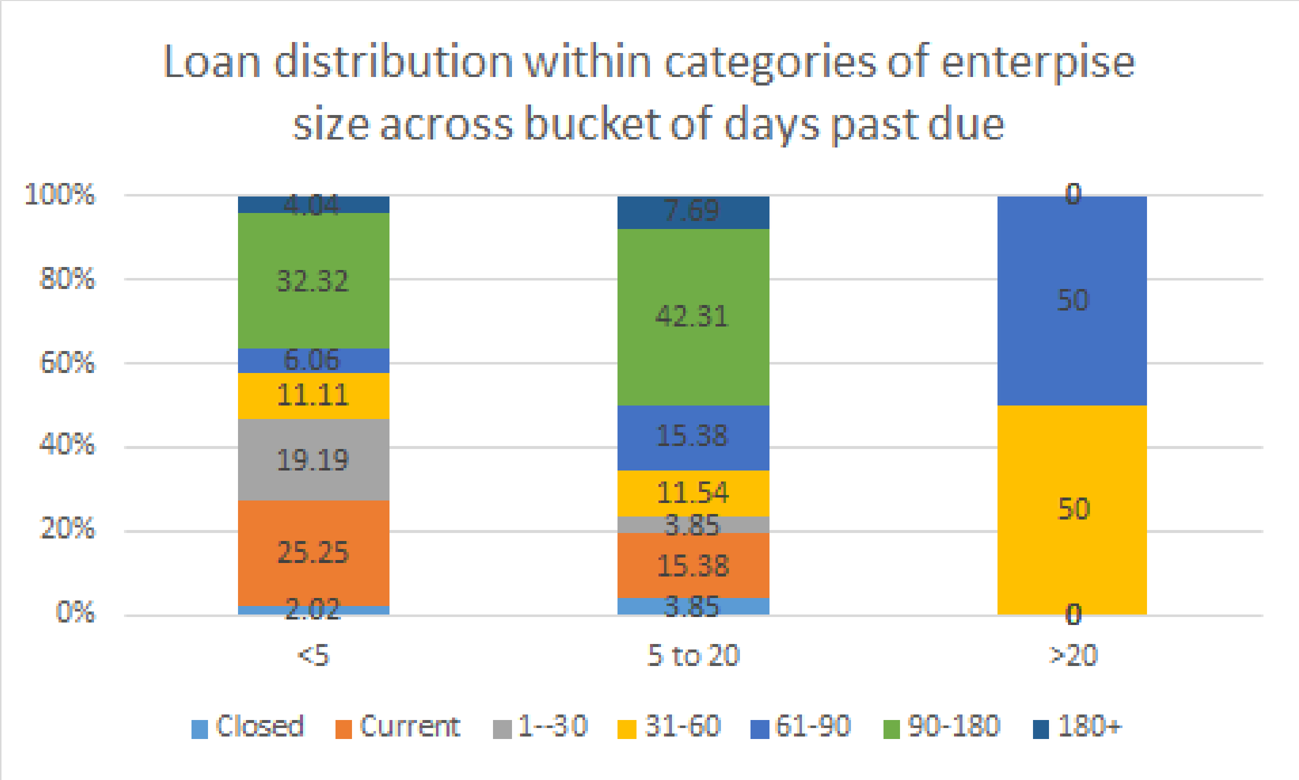


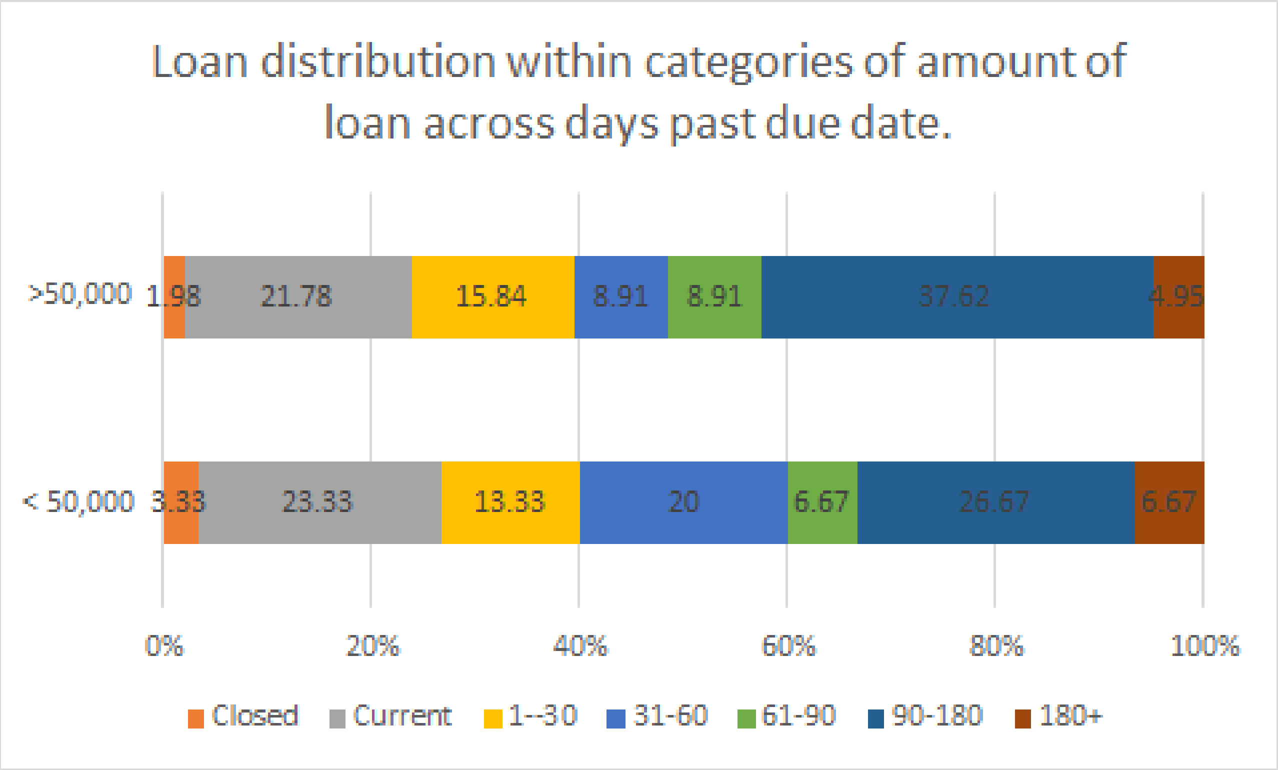
Females tend to default loans beyond 90 days relatively lesser (~27%) than male counterparts (~44%)



More than 75% of loans are taken by enterprises with less than 5 employees. Around 24% by employees between 5 to 20. A miniscule percent by those with more than 20 employees.

Enterprises with more than 20 people have not defaulted beyond 90 days. Those with 5 to 20 people have a majority ~50% of defaulters beyond 90 days. Those with less than 5 people have relatively lesser ~36% defaulters.





Higher value loans(>50,000) have higher share of defaults beyond 90 days.
But correlation table has shown this relation is not significant.

Distribution within age groups shows that those in ages 20-40 have higher number of loans defaulted beyond 90 days than those in age group 40-60.

Age group and days past dues								
age_group	Closed	Current	1--30	31-60	61-90	90-180	180+	Total
20-30	4	20	20	4	4	44	4	100
30-40	3.39	20.34	11.86	3.39	10.17	44.07	6.78	100
40-50	0	26.32	13.16	26.32	7.89	21.05	5.26	100
50-60	0	22.22	33.33	22.22	11.11	11.11	0	100
Total	2.29	22.14	15.27	11.45	8.4	35.11	5.34	100

Cities across DPD bucket							
City	Closed	Current	1--30	31-60	61-90	90-180	180+
Ahmed Nagar	5.26	31.58	26.32	10.53	0	21.05	5.26
Bengaluru	0	0	0	0	0	100	0
Kolhapur	0	0	0	0	50	50	0
Mumbai	0	0	0	0	0	100	0
Nashik	0	0	50	0	50	0	0
Navi Mumbai	0	0	0	0	0	100	0
Pune	1.67	23.33	18.33	13.33	6.67	33.33	3.33
Sangli	2.44	21.95	7.32	9.76	12.2	39.02	7.32
Thane	0	0	0	25	0	50	25
Total	2.29	22.14	15.27	11.45	8.4	35.11	5.34

Ahmad Nagar (19), Pune (60) and Sangli (41) have majority of loans among the 131 observations, and among these Pune and Sangli haver a higher number of defaults beyond 90 days. Sangli has a higher percentage ~46% in spite of less observations than Pune.

Purpose of loan and DPD Bucket									
Purpose of loan		Closed	Current	1--30	31-60	61-90	90-180	180+	Total
Expansion Capital	Freq	0	1	2	1	1	0	0	5
	percent	0	20	40	20	20	0	0	100
Working capital	Freq	3	28	18	14	10	46	7	126
	percent	2.38	22.22	14.29	11.11	7.94	36.51	5.56	100
	Total	3	29	20	15	11	46	7	131
		2.29	22.14	15.27	11.45	8.4	35.11	5.34	100

Loan taken for expansion of capital is lesser in frequency but has shown zero defaults beyond 90 days, whereas those for working capital have more than 40% loans defaulted.

Some inferences

- Distribution tables show that lesser age and higher principal amount have higher percentage of loan defaults but our correlation and regression tables have shown that these relations are not so significant.
- Defaulting behaviour is also regionally affected as Sangli shows higher rates of default
- Based on gender, women are seen to have taken smaller loans and defaulted less than men.

Additional Data Required

For assessing Impact of Loans

- Turnover/sales of the borrowing company
- Net profit
- Value of inventory
- Any existing loan, amount and name of lending bank
- Does the applicant/borrower know about the MUDRA loan scheme? Has he availed a loan under that?
- Number of employees
- Amount of fixed assets

For assessing S-ROI (in addition to the data required to assess Impact of Loans:

- Cost incurred by deAsra in implementing the Program

Thank You!