

NSS Data Extraction and Regression Analysis.
(BREAKING BARRIERS: EXPLORING WOMENS WORK IN
INDIA’S VILLAGES AND CITIES)

NSSO 68th Round Employment & Unemployment

Muztaba Anjum (22227707120)

M.A. ECONOMICS, 2ND YEAR

Submitted To:
Professor Devesh Birwal



ACKNOWLEDGMENT

I would like to express my heartfelt gratitude towards Professor Devesh Birwal, Delhi School of Economics, for granting me this wonderful opportunity to work on this research project. I thank him for his indispensable guidance and feedback throughout the course of this Term paper

Overview

This study examines the contribution of women's in the labor force in India, unlocking several notable trends. The first one is, the educational attainment of women display a U-shaped curve, indicating difficulties faced by both those with low educations and those with the high educations. Secondly, the labor force participation rate (LFPR) is higher in rural regions than in urban settings. Thirdly, there is a noticeable decline in LFPR as women age. Forth one is with the exception of Buddhists, women from various religious backgrounds tend to have lower LFPR compared to their Hindu counterparts. Fifth one is the outcome of land ownership on female LFPR vary between rural and urban contexts; in rural areas, larger landholdings tally with increased participation, whereas the reverse is noticed in urban areas. Lastly, the analysis underscores the economic challenges faced by widows, as all other marital categories show an upward trend in LFPR. These insights illustrate the intricate relationships among education, geographic location, religion, land ownership, and marital status in influencing women's participation in the labor market in India.

Introduction

The participation of women in the labor force in India is critically low, ranking 120th out of 131 countries according to the International Labor Organization's 2013 report. Within the South Asian context, India is positioned just above Pakistan and Afghanistan, underscoring a significant obstacle to women's economic involvement. This low participation rate is vital as it influences family size, population management, and economic advancement.

Although nearly all women are engaged in some form of work, whether inside or outside their homes, much of this labor remains unacknowledged or unpaid. Conventional definitions of work often overlook these contributions, especially those related to self-consumption production and reproductive responsibilities. This limited perspective marginalizes the extensive, albeit informal, economic activities that women perform.

Enhancing female labor force participation is crucial for both fairness and efficiency. From an equity perspective, it improves women's economic status and fosters gender equality. From an efficiency standpoint, fully incorporating women into the workforce can enhance overall economic productivity and developmental potential.

A significant aspect of this issue in India is the contrast between rural and urban regions. Notably, female labor force participation tends to be higher in rural areas, where women often engage in agricultural work and family businesses. These contributions, while substantial, are frequently informally acknowledged and inadequately compensated. Conversely, urban areas, despite providing more job opportunities, impose greater barriers such as educational prerequisites, formal skill requirements, and restrictive social norms.

Grasping these dynamics is essential for developing effective policies. In rural regions, improving access to education and skills training, as well as recognizing informal work, can economically empower women. In urban environments, it is vital to address social norms, offer childcare support, and encourage flexible work arrangements. In urban settings, addressing social norms, providing childcare support, and promoting flexible work arrangements are crucial.

This study aims to explore these dynamics comprehensively, focusing on the rural-urban perspective to uncover the factors influencing female labor force participation in India. By addressing these challenges, we can enhance women's economic involvement, driving sustainable economic growth and development.

Literature review

In India, women disproportionately shoulder the responsibility of unpaid domestic tasks, including childcare and caring for the sick or elderly. Those who are part of the formal workforce face challenges such as lower wages, a lack of childcare options, and limited job security, particularly within the informal sector (Oxfam India, 2019). Olsen et al. (2006) argue that social and cultural obstacles in a largely patriarchal society like India influence women's employment decisions. A prevalent view is that rising incomes in rural areas have enabled more women to opt for staying at home, which is often seen as a preferable choice in such a societal context. Ghose (2016) highlights that gender inequality in employment remains a significant issue in India, with women's labor force participation at only 22%, compared to 54% for men, and a mere 15% of women engaged in the workforce in urban settings.

The inquiry into "Why is female labor force participation declining so sharply in India?" delves into the causes behind the recent steep decline and the long-term stagnation of female labor force participation in the country. It identifies factors such as improved education, increased household incomes, and changes in survey methodologies as contributors to this decline. A critical long-term challenge is the scarcity of employment opportunities for women, largely due to occupational segregation (Kapsos S., Silberman A., & Bourmpoula E., 2014). Kapsos et al. (2016) note that in India, female labor force participation exhibits a U-shaped relationship with educational attainment. Women with secondary education tend to have the lowest participation rates, whereas those with tertiary education show a significant rise in participation. Despite economic growth, declining fertility rates, and improvements in wages and education, research by Stephan Klasen and Janneke Pieters indicates that female labor force participation in urban India has remained unexpectedly low and stagnant at around 18 percent. Their cross-sectional studies suggest that both supply and demand factors have contributed to this stagnation. On the supply side, increasing household incomes and the education level of husbands, as well as a decrease in the selectivity of highly educated women entering the workforce, have contributed to the low participation rates. On the demand side, sectors that typically employ female workers have not expanded significantly, and changes in the employment sector structure alone would have likely led to a further decline in participation rates.

Methodology

The project's methodology focuses on analyzing the factors influencing female labor force participation rates, drawing on data from the NSSO 68th Employment & Unemployment round survey. The dependent variable, which is the female labor force participation rate, is divided into two primary categories: individuals who are part of the labor force and those who are not. For women's within the labor force, the involvement rate is calculated by combining those who are employed with those who are unemployed, whereas the remaining individuals are grouped as females outside the labor force.

DATA EXTRACTION

From NSSO 68th round of Employment & Unemployment data from Block - 3 , Block - 4 & Block - 5_1 has been utilized for the analysis.

Religion, Social_Group, Land_Owned has been taken from block-3 which provides Household level data.

Sex, Age, Marital_Status, General_Education, Sector has been taken from block-4 which provides Individual level data.

Usual_Principal_Activity_Status has been taken from block-5 which is again individual level data.

In block-4 & block-5 individual household key was already present , but for block-3 individual household key was generated (new variable) using command paste0 with the variables FSU_Serial_No, Hamlet_Group_Sub_Block_No, Stratum_No, Sample_Hhld_No

Then all three blocks were merged by using individual household key , then na.omit command was utilized to remove NAs and 'dplyr' package has been used to filter out categories which are of our use.

From NSSO data under Usual_Principal_Activity_Status code 11, 12, 21, 31, 41 & 51 refer to the 'employed', 81 to the 'unemployed' and the remaining viz. 91 to 97 refer to the 'not in labour force'.

Code 11, 12, 21, 31, 41, 51 & 81 was coded as 1 (in the labor force)

Code 91 to 97 was coded as 0 (out of labor force)

Since paper tries to find impact on Female labor force participation hence , **sex==2**
i.e., Only female was filtered out from the data (using dplyr package)

To analyze the determinants, various **independent variables** are considered:

1. General education, categorized with "Not literate" as the reference category.

Codes for different level of data are as follow

Not literate	01
Literate:	
primary.....	06
middle	07
secondary	08
higher secondary	10
diploma/certificate course.....	11
graduate.....	12
postgraduate and above.	13

2. Age, treated as a continuous variable.

(Age) ² was taken as an interaction term.

3. Religion, a categorical variable with "Hinduism" as the reference category.

Codes for different religions are as follows

Hinduism.....	1
Islam	2
Christianity.....	3
Sikhism.....	4
Jainism	5
Buddhism.....	6
Zoroastrianism.....	7
others	9

4. Social group, a categorical variable with "others" as the reference category.

Codes for different social groups are as follows

scheduled tribe	1
scheduled caste.....	2
other backward classes ...	3
others	9

5. Marital status, a categorical variable with "never married" as the reference category.

Codes for Marital status are as follows

never married	1
currently married	2

widowed.....	3
divorced/ separated.....	4

6. Land owned, treated as a continuous variable.

*(Following the recommendation, **attempts were made to categorize land ownership** into groups such as those with less than 5 hectares, between 5-10 hectares, and so forth. However, due to the highly right-skewed distribution of land ownership, with a maximum value of 38.413 hectares, a minimum of 0, and a median of 0.418 hectares, creating these categories did not yield statistically significant results. Consequently, categories for land ownership were not established).*

Separate regressions are conducted for the rural and urban sectors to provide a sector-wise analysis of the determinants affecting female labor force participation.

While analyzing rural sector , **sector==1** was filtered
& consequently for urban sector , **sector==2** was filtered

Additionally, a national-level regression, incorporating both rural and urban areas, is performed, with rural areas serving as the reference category.

As the dependent variable has a binary outcome (either in the labor force or not), logistic regression is employed. Additionally, probit regression is also conducted to compare and validate the results obtained from logistic regression.

The programming language used for data analysis is R Studio.

Description of models used

This project involves the application of three statistical models: C Log-Log regression, logistic regression and probit regression, given that the dependent variable is binary. Below is a detailed description of all three C Log-Log regression, Logistic regression and Probit regression.

C Log-Log Regression

C Log-Log (complementary log-log) regression is used to model binary outcome variables. It differs from logistic and probit regression in the link function it uses.

Assumptions:

- The dependent variable is binary.
- Observations are independent.
- Little multicollinearity among predictors.

The probability of the dependent variable $Y = 1$ is modelled as:

$$P(Y=1|X) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}$$

Logistic Regression

Logistic regression is a statistical method used for analyzing datasets in which the outcome or dependent variable is binary (i.e., it has two possible outcomes). It is used to model the probability that a given input point belongs to a particular category.

Assumptions:

- The dependent variable is binary.
- The log-odds of the outcome are a linear combination of the independent variables.
- Observations are independent of each other.
- There is no or little multicollinearity among the independent variables.

In logistic regression, the probability of the dependent variable \square taking on the value 1 is modeled as a function of one or more independent variables, \square .

The model can be expressed as:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}$$

where,

$P(Y=1|X)$ is the probability of the dependent variable, \square being 1 given the predictors, X .

\square_0 is the intercept term.

$\square_1, \square_2, \dots, \square_k$ are the coefficients of the predictors $\square_1, \square_2, \dots, \square_k$

The coefficients $\square_0, \square_1, \dots, \square_k$ are estimated using the method of maximum likelihood estimation (MLE), which finds the parameter values that maximize the likelihood of observing the given sample data.

Probit Regression

Probit regression, like logistic regression, is used to model binary outcome variables. The key difference lies in the link function used to model the relationship between the dependent variable and the independent variables.

Assumptions:

- The dependent variable is binary.
- The cumulative probability of the outcome is modeled as a standard normal CDF of a linear combination of the predictors.

- Observations are independent.
- There is no or little multicollinearity among the independent variables.

In probit regression, the probability of the dependent variable, y taking on the value 1 is modeled using the cumulative distribution function (CDF) of the standard normal distribution. The model can be expressed as:

$$P(y=1|x) = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)$$

Where, Φ denotes the CDF of the standard normal distribution.

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ are the coefficients of the model.

The coefficients in probit regression are also estimated using maximum likelihood estimation. The likelihood function is derived from the normal distribution, and the coefficients are chosen to maximize this likelihood function.

Comparison of Logistic and Probit Regression

1. Link Function

- **Logistic Regression:** Uses the **logit** link
 $\text{Log } P/(1-P)$
- **Probit Regression:** Uses the **standard normal CDF (Φ)**
 $\Phi^{-1}(P)$
- **Cloglog Regression:** Uses the **complementary log-log** link
 $\log(-\log(1-P))$

2. Symmetry

- **Logistic:** Symmetric S-shaped curve
- **Probit:** Symmetric, but slightly flatter in the middle and steeper at the tails
- **Cloglog:** **Asymmetric**; rises sharply and then levels off

3. Interpretation

- **Logistic:** Odds-based interpretation
- **Probit:** Interpreted in terms of Z-scores from the standard normal distribution
- **Cloglog:** Often used for modelling **rare events** or **time-to-event** data (hazard modelling)

4. Probability Range Behavior

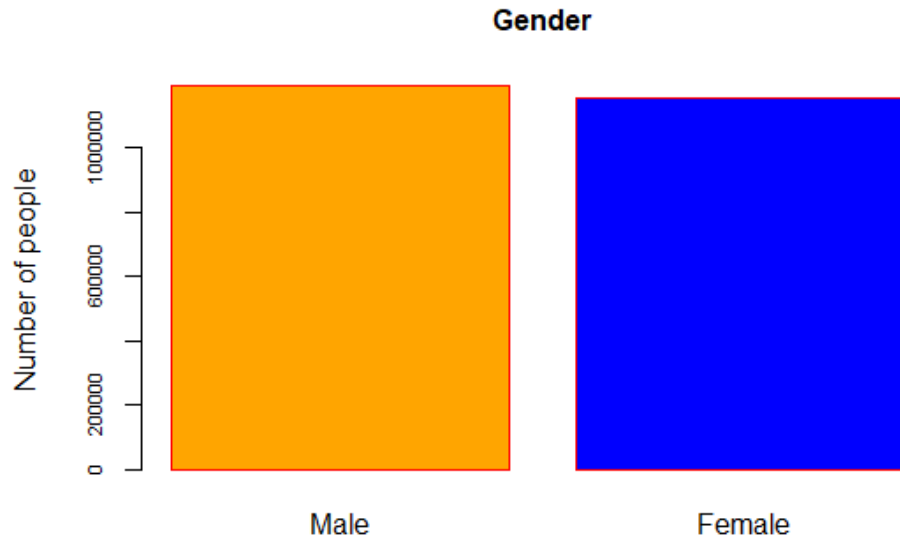
- **Logistic & Probit:** Probability approaches 0 and 1 symmetrically
- **Cloglog:** Probability approaches 1 more gradually, suitable when the event is skewed (e.g., mostly 0s or mostly 1s)

5. Use Cases

- **Logistic:** General-purpose binary classification
- **Probit:** When normality assumption of latent variable is justified
- **Cloglog:** When event probabilities are **skewed**, or in **survival analysis**

DEMOGRAPHIC STRUCTURE OF THE PARTICIPANTS

1. SEX

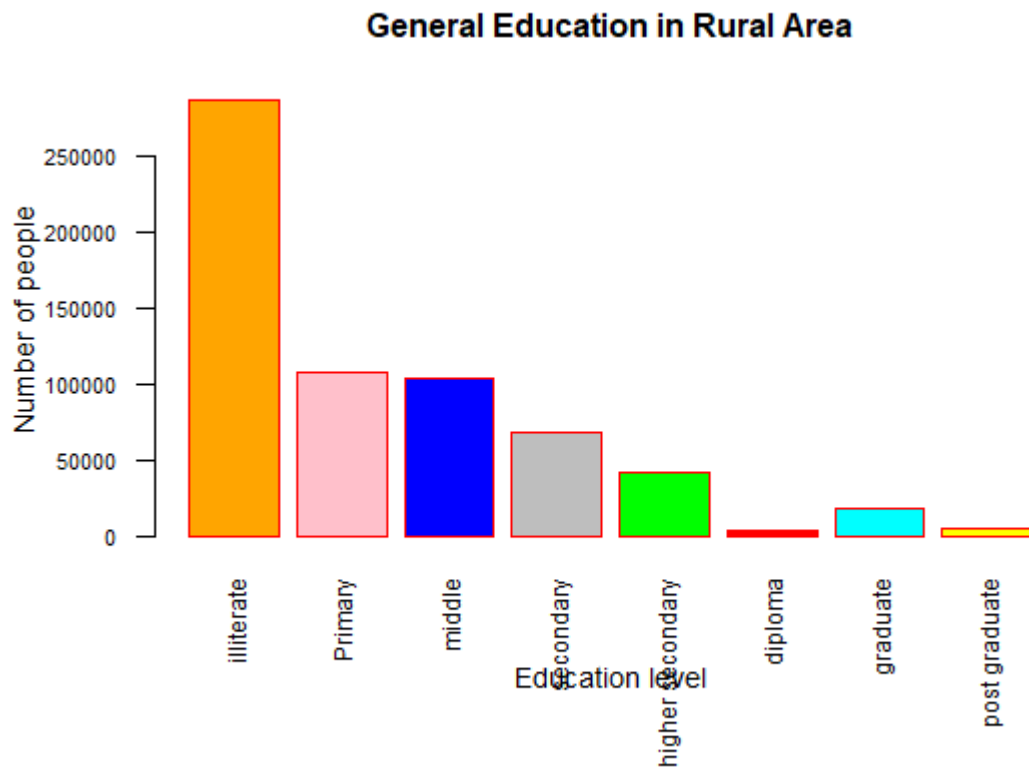


The study encompasses a near-balanced representation of both males and females, with a marginally larger proportion of males. However, given that the primary objective of the study pertains to the analysis of female labor force participation, exclusively female data was utilized for this purpose.

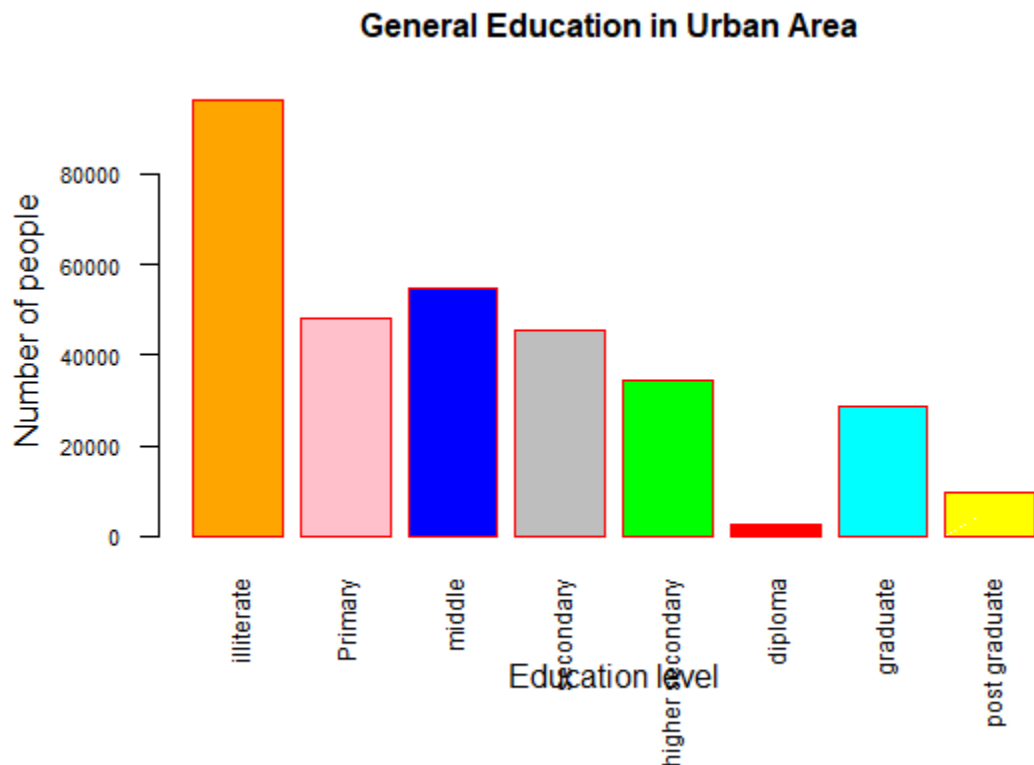
2. General Education

The study incorporates general education as an independent variable, encompassing various levels including illiterate, primary, middle, secondary, higher secondary, diploma, graduate, and postgraduate qualifications. The reference category within this framework is established as the illiterate

2.1 General Education in rural area



2.2 General education in urban area



In rural areas, female illiteracy rates surpass those in urban areas. Conversely, urban regions exhibit higher levels of female educational attainment across primary, middle, secondary, and higher levels of education.

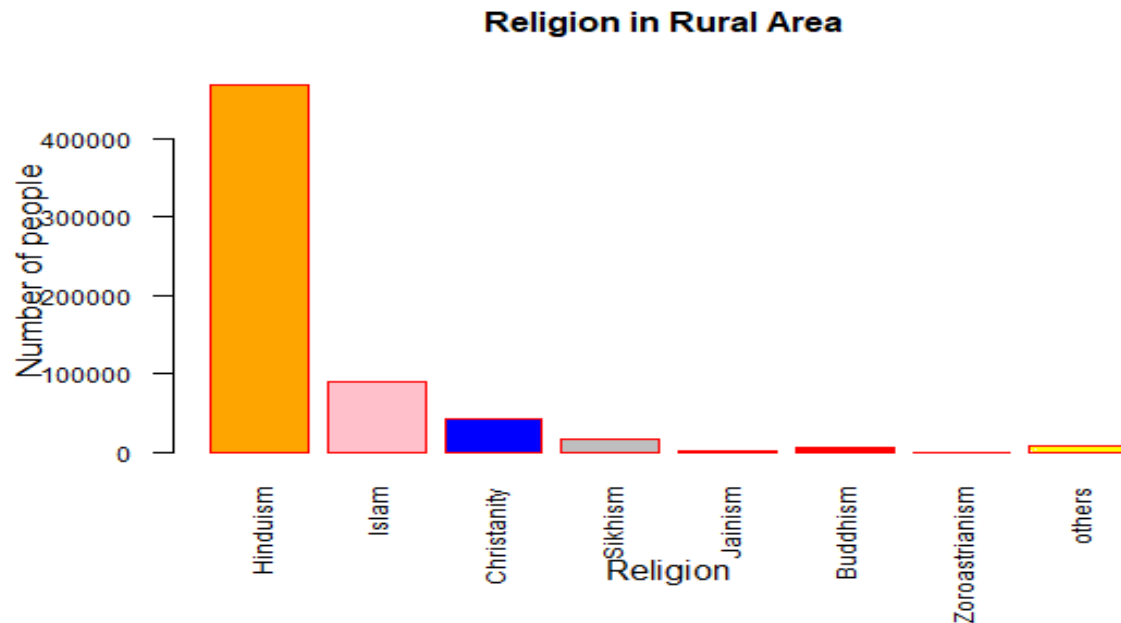
3. Age

The study incorporates age as an independent variable, which is a continuous variable.

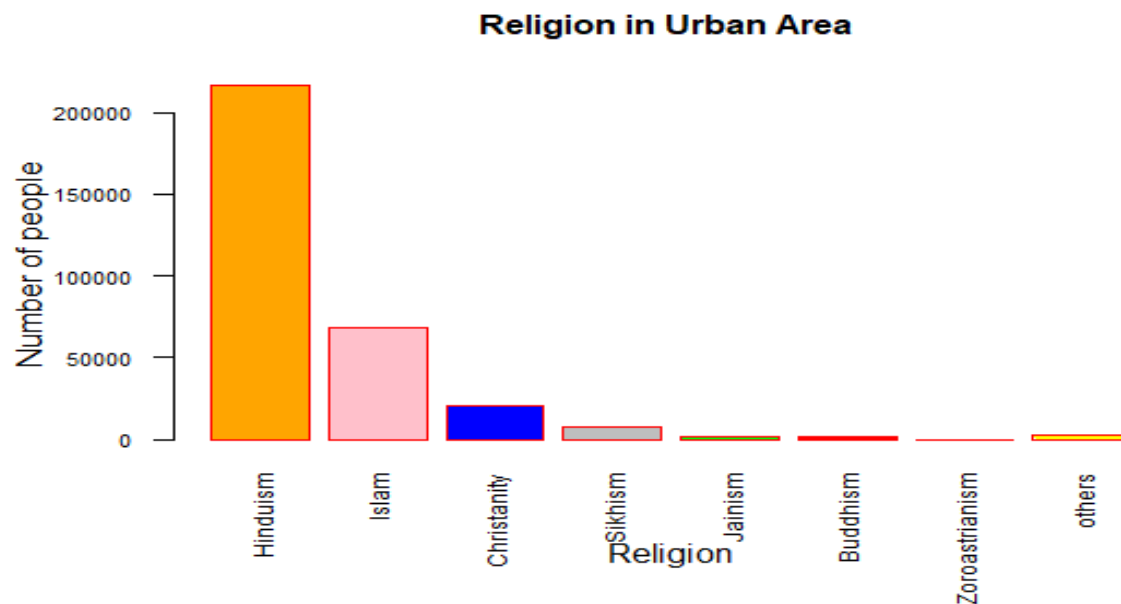
4. Religion

The study integrates religion as a categorical variable, delineating a spectrum of categories of hinduism, islam, christianity, sikhism, jainism, buddhism, zoroastrianism, and others. Within this analytical framework, the benchmark category is firmly established as Hinduism.

4.1 Religion in rural area



4.2 Religion in urban area

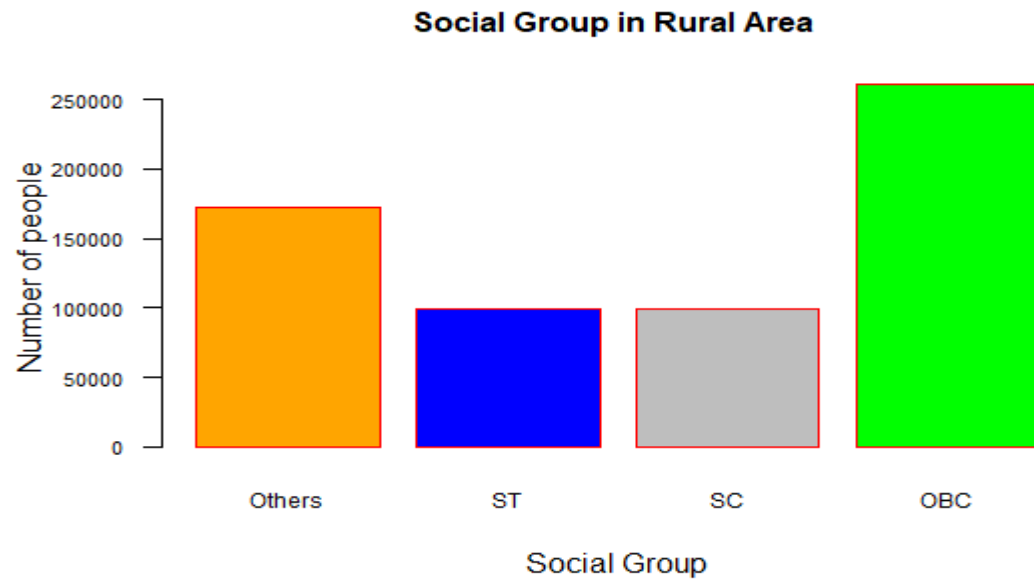


In rural areas, the Hindu population outweighs that in urban areas, whereas urban regions have a higher concentration of Islamic individuals compared to rural areas.

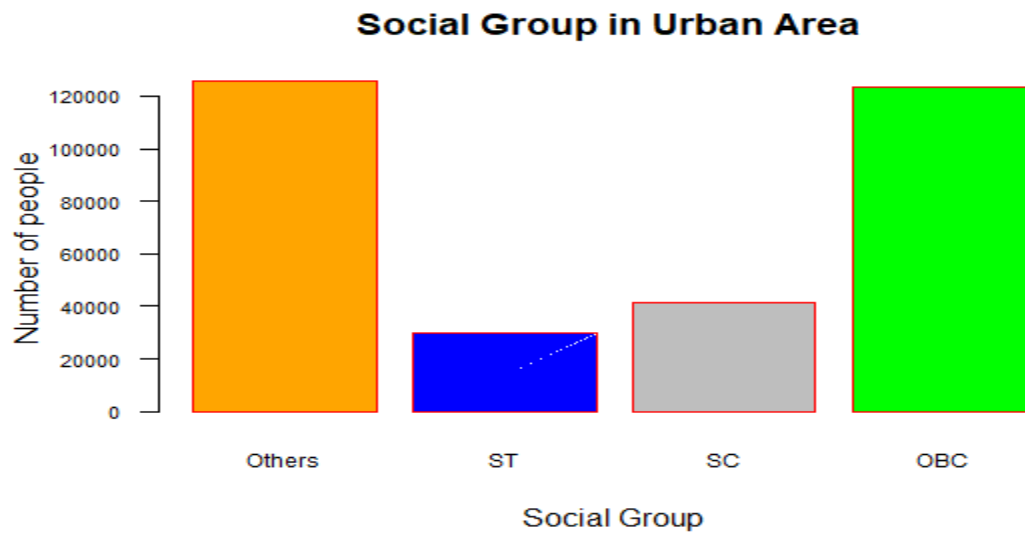
5. Social group

The study incorporates social group as a categorical variable, delineating various categories such as SC, ST, OBC, and the general population. Within this analytical framework, the reference category is firmly established as the general population.

5.1 Social group in rural area



5.2 Social group in urban area

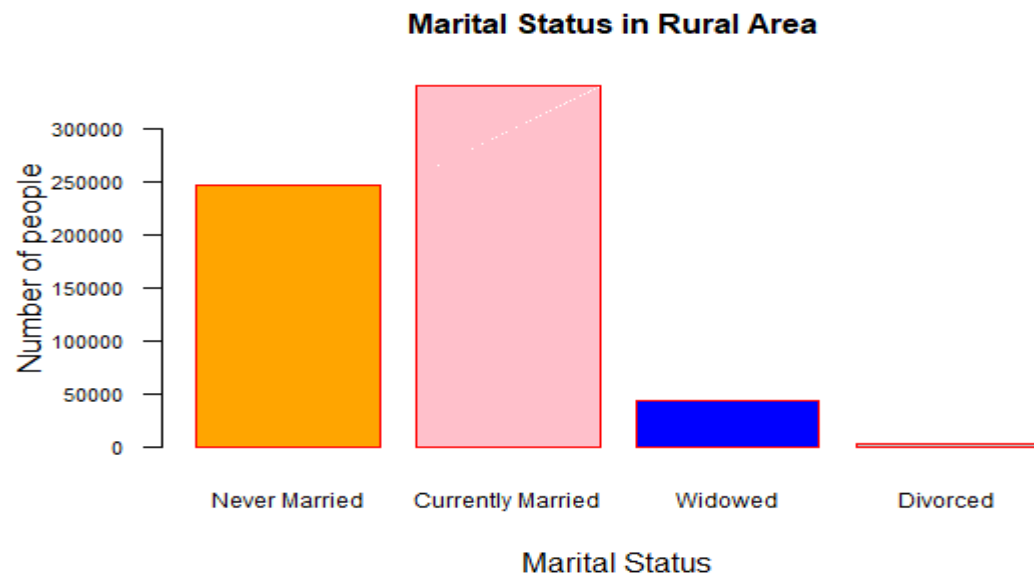


In rural locales, the proportion of individuals belonging to the Other Backward Classes (OBC) demographic exceeds that of urban areas, whereas urban centers exhibit a higher prevalence of the general category populace compared to rural regions.

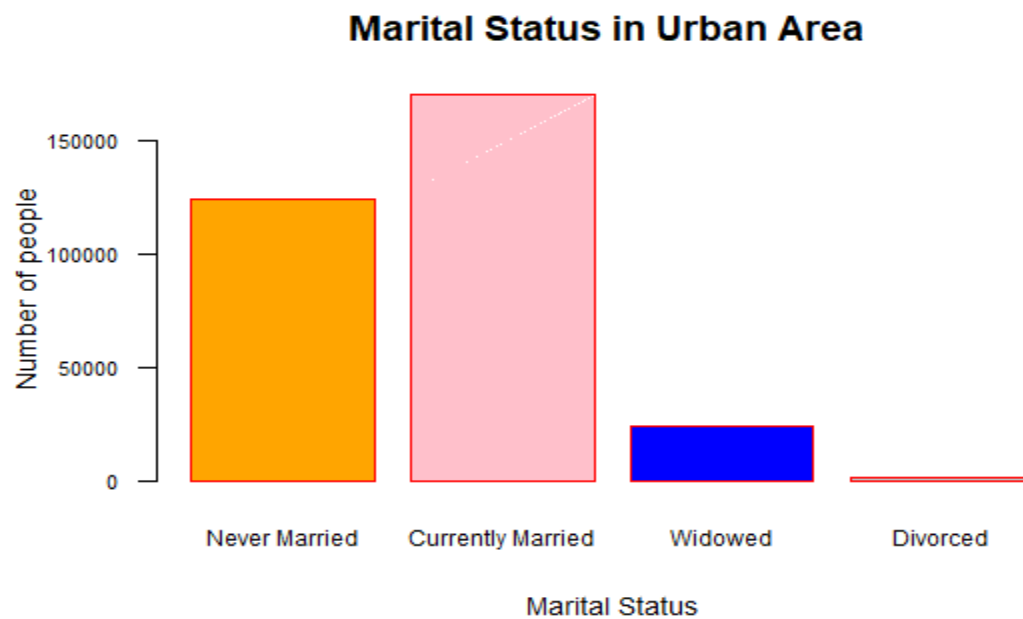
6. Marital Status

The research includes marital status as an independent variable, covering different levels such as never married, currently married, widowed, and divorced. Within this framework, the reference category is set as never married.

6.1 Marital status in rural area



6.2 Marital status in urban area



In rural areas, there is a higher prevalence of unmarried, currently married, and widowed females compared to urban areas.

Results

1. C Log-Log model rural sector

```
> summary(cloglog_reg_Rural)
```

Call:

```
glm(formula = Female_Labor_Force_Participation ~ General_Education +  
    Age + Age^2 + Religion + Social_Group + Land_Owned + Marital_Status,  
    family = binomial(link = "cloglog"), data = Labor_Force)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4319	-0.9394	-0.8646	1.4100	1.6656

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.9933253592	0.0072018334	-137.927	< 0.00000000000000002	***
General_Education06	-0.0314443904	0.0063761545	-4.932	0.000000815753856170	***
General_Education07	-0.0080238998	0.0064023794	-1.253	0.000000000000194	***
General_Education08	-0.0051133725	0.0075060460	-0.681	0.00000000039250	***
General_Education10	0.0081797135	0.0091624904	0.893	0.0000576788930	***
General_Education11	0.1898190775	0.0285291783	6.654	0.000000000028619089	***
General_Education12	0.1010458135	0.0129943943	7.776	0.000000000000007479	***
General_Education13	0.1999722224	0.0245965536	8.130	0.000000000000000429	***
Age	0.0028961511	0.0001883692	15.375	< 0.00000000000000002	***
Religion2	-0.2319672908	0.0070288171	-33.002	< 0.00000000000000002	***
Religion3	-0.0164973227	0.0098179040	-1.680	0.0929	.
Religion4	-0.1592917171	0.0147542051	-10.796	< 0.00000000000000002	***
Religion5	-0.1600267626	0.0712116137	-2.247	0.0246	*
Religion6	0.2191434525	0.0204254569	10.729	< 0.00000000000000002	***
Religion7	-0.6072128625	1.0034888897	-0.605	0.5451	
Religion9	-0.1369498580	0.0204694600	-6.690	0.000000000022248840	***
Social_Group1	0.1822750319	0.0078972128	23.081	< 0.00000000000000002	***
Social_Group2	0.0011485097	0.0072675179	0.158	0.8744	
Social_Group3	-0.0111295145	0.0055533756	-2.004	0.0451	*
Land_Owned	0.0000098384	0.0000008284	11.877	< 0.00000000000000002	***
Marital_Status2	0.0787456151	0.0067636130	11.643	< 0.00000000000000002	***
Marital_Status3	-0.0128776294	0.0131172196	-0.982	0.02861	*
Marital_Status4	0.2710588697	0.0351120333	7.720	0.0000000000000011649	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The regression results are:

- Females who have completed primary, middle, secondary education experiences a reduction in the log odds of participating in the labor force by 0.031, 0.008 and 0.005 respectively, compared to illiterate females. Conversely, females with diploma, graduate, and postgraduate levels see an increase in the log odds by 0.008, 0.189, and 0.101 respectively, compared to illiterate females, possibly due to increased participation at higher education levels.
- An increase in age by one year is associated with an increase in the log odds of labor force participation by 0.002, contrary to the hypothesis.
- Females adhering to Islam, Christianity, Sikhism, Jainism, Zoroastrianism, and other religions experience a reduction in the log odds of labor force participation by 0.231, 0.016, 0.159, 0.160, 0.607, and 0.136 respectively, compared to Hindus. Buddhists show an increase by 0.219, with all findings significant except for Christianity and Zoroastrianism.

- Females from Scheduled Castes (SC) and Other Backward Classes (OBC) have increased log odds by 0.182 and 0.001 respectively, while Scheduled Tribes (ST) show an decrease by 0.011, compared to the general category. All are significant except SC, possibly due to historical discrimination in SC and OBC communities versus economic necessity in ST.
- Widowed females experience a reduction in log odds by 0.012, while married and divorced females show increases by 0.078 and 0.271 respectively, compared to never-married females. All are significant except married females, potentially due to emotional and social factors for widows and financial needs for divorcees.
- A one-unit increase in land owned increases the log odds by 0.000009, suggesting slight encouragement of labor force participation due to agricultural demands.

2.C Log-Log Urban sector

```
> summary(cloglog_reg_Urban)
```

Call:

```
glm(formula = Female_Labor_Force_Participation ~ General_Education +  
  Age + Age^2 + Religion + Social_Group + Land_Owned + Marital_Status,  
  family = binomial(link = "cloglog"), data = Labor_Force)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4659	-0.8892	-0.8522	1.4532	1.6488

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.021091898	0.010247444	-99.644	< 0.0000000000000002	***
General_Education06	-0.066564384	0.010241828	-6.499	0.00000000000807	***
General_Education07	-0.031893642	0.009768635	-3.265	0.001095	**
General_Education08	-0.039114800	0.010464816	-3.738	0.000186	***
General_Education10	-0.014229488	0.011472022	-1.240	0.214841	**
General_Education11	0.272261496	0.031438793	8.660	< 0.0000000000000002	***
General_Education12	0.111709270	0.011995473	9.313	< 0.0000000000000002	***
General_Education13	0.251244623	0.017678219	14.212	< 0.0000000000000002	***
Age	0.002843464	0.000274466	10.360	< 0.0000000000000002	***
Religion2	-0.114604519	0.008382421	-13.672	< 0.0000000000000002	***
Religion3	0.134157523	0.015475382	8.669	< 0.0000000000000002	***
Religion4	-0.024424338	0.021066674	-1.159	< 0.0000000000000002	***
Religion5	0.030188747	0.040147577	0.752	0.452084	
Religion6	-0.059020912	0.040023698	-1.475	0.140307	
Religion7	0.835633464	0.269154934	3.105	0.001905	**
Religion9	-0.130983164	0.039034213	-3.356	0.000792	***
Social_Group1	0.058883480	0.014220259	4.141	0.0000346072268	***
Social_Group2	0.052110351	0.010588668	4.921	0.0000008595709	***
Social_Group3	0.011252454	0.007315852	1.538	0.124026	
Land_Owned	0.000001857	0.000002589	-0.717	0.473305	
Marital_Status2	-0.002368249	0.009639425	-0.246	0.0000001052	***
Marital_Status3	-0.063239196	0.018218652	-3.471	0.000518	***
Marital_Status4	0.092969889	0.041684864	2.230	0.025727	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The regression results are:

- Females who have completed primary, middle, secondary education experiences a reduction in the log odds of participating in the labor force by 0.066, 0.031, 0.039 and 0.014 respectively, compared to illiterate females. Conversely, females with diploma, graduate, and postgraduate levels see an increase in the log odds by 0.272, 0.111, and 0.251 respectively, compared to illiterate females, possibly due to increased participation at higher education levels.
- An increase in age by one year is associated with an increase in the log odds of labor force participation by 0.002, contrary to the hypothesis.
- Females adhering to Islam, Christianity, Sikhism, Jainism, and other religions experience a reduction in the log odds of labor force participation by 0.114, 0.024, 0.059, and 0.130 respectively, compared to Hindus. Buddhists show an increase by 0.134, with all findings significant except for Christianity and Zoroastrianism.
- Females from Scheduled Castes (SC), Other Backward Classes (OBC) and Scheduled Tribe (ST) have increased log odds by 0.058, 0.052 and 0.011 respectively. All are significant except SC, possibly due to historical discrimination in SC and OBC communities versus economic necessity in ST.
- Widowed and divorced females experience a reduction in log odds by 0.002 and 0.063 while married females show increases by 0.092 respectively, compared to never-married females. All are significant.
- A one-unit increase in land owned increases the log odds by 0.0000018, suggesting slight encouragement of labor force participation due to agricultural demands.

3. Logit model Rural sector

```
> summary(logistic_reg_Rural)
```

```
Call:
```

```
glm(formula = Labor_Force$Female_Labor_Force_Participation ~  
    Labor_Force$General_Education + Labor_Force$Age + I(Labor_Force$Age^2) +  
    Labor_Force$Religion + Labor_Force$Social_Group + Labor_Force$Marital_Status +  
    Labor_Force$Land_Owned, family = binomial(link = "logit"),  
    data = Labor_Force)
```

```
Deviance Residuals:
```

```
      Min       1Q   Median       3Q      Max  
-1.4229  -0.9423  -0.8606   1.4024   1.7743
```

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.877637788	0.009792437	-89.624	< 0.0000000000000002	***
Labor_Force\$General_Education06	-0.076283951	0.008088445	-9.431	< 0.0000000000000002	***
Labor_Force\$General_Education07	-0.055732883	0.008305461	-6.710	0.00000000000194	***
Labor_Force\$General_Education08	-0.056974794	0.009677537	-5.887	0.00000000039250	***
Labor_Force\$General_Education10	-0.047304679	0.011761157	-4.022	0.0000576788930	***
Labor_Force\$General_Education11	0.187241068	0.036758920	5.094	0.0000003510334	***
Labor_Force\$General_Education12	0.067637812	0.016667675	4.058	0.0000494898698	***
Labor_Force\$General_Education13	0.192150324	0.031851151	6.033	0.0000000016118	***
Labor_Force\$Age	0.015136535	0.000714435	21.187	< 0.0000000000000002	***
I(Labor_Force\$Age^2)	-0.000149527	0.000008742	-17.104	< 0.0000000000000002	***
Labor_Force\$Religion2	-0.284620283	0.008369639	-34.006	< 0.0000000000000002	***
Labor_Force\$Religion3	-0.016167132	0.012385635	-1.305	0.19179	
Labor_Force\$Religion4	-0.191185206	0.017703449	-10.799	< 0.0000000000000002	***
Labor_Force\$Religion5	-0.188696816	0.086208113	-2.189	0.02861	*
Labor_Force\$Religion6	0.287584695	0.027069372	10.624	< 0.0000000000000002	***
Labor_Force\$Religion7	-0.718871790	1.154750549	-0.623	0.53359	
Labor_Force\$Religion9	-0.172162055	0.025336309	-6.795	0.00000000000108	***
Labor_Force\$Social_Group1	0.212383396	0.009933544	21.380	< 0.0000000000000002	***
Labor_Force\$Social_Group2	-0.008226797	0.008927628	-0.921	0.35679	
Labor_Force\$Social_Group3	-0.018356569	0.006784735	-2.706	0.00682	**
Labor_Force\$Marital_Status2	0.001952097	0.009966173	0.196	0.84471	
Labor_Force\$Marital_Status3	-0.039088216	0.016264489	-2.403	0.01625	*
Labor_Force\$Marital_Status4	0.251914921	0.045891156	5.489	0.0000000403301	***
Labor_Force\$Land_Owned	0.000013182	0.000001049	12.565	< 0.0000000000000002	***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The regression results are:

- Females who have completed primary, middle, secondary, and higher secondary education experience a reduction in the log odds of participating in the labor force by 0.076, 0.055, 0.056, and 0.047 respectively, compared to females who are illiterate. Conversely, females who have attained diploma, graduate, and postgraduate levels of education see an increase in the log odds of labor force participation by 0.187, 0.067, and 0.192 respectively, in comparison to their illiterate counterparts. **This phenomenon may be attributed to the fact that only after attaining higher levels of education, females exhibit increased participation in the labor force.**

- An increase in age by one year for females is associated with an increase in the log odds of participating in the labor force by 0.015. However, this result contradicts our hypothesis, which posits that female labor force participation should decrease with age. Therefore, the model included (Age) ² as an additional predictor variable. The coefficient associated with (Age) ² is -0.001, indicating that an increase in age by one year for females is associated with a decrease in the log odds of participating in the labor force, aligning with our hypothesis.
- Females adhering to Islam, Christianity, Sikhism, Jainism, Zoroastrianism, and other religions experience a reduction in the log odds of participating in the labor force by 0.284, 0.016, 0.191, 0.188, 0.718, and 0.172, respectively, when compared to their Hindu counterparts. In contrast, females who are Buddhists exhibit an increase in the log odds of labor force participation by 0.287, relative to Hindu females. These findings are statistically significant for all categories except females who practice Christianity and Zoroastrianism.
- Females belonging to Scheduled Castes (SC) and Other Backward Classes (OBC) experience a reduction in the log odds of participating in the labor force by 0.008 and 0.018, respectively, compared to females from the general category. Conversely, females from Scheduled Tribes (ST) exhibit an increase in the log odds of labor force participation by 0.212, in comparison to their counterparts from the general category. These findings are statistically significant for all categories except females belonging to Scheduled Castes (SC). **This is probably because SC and OBC communities have historically faced systemic discrimination, limited access to quality education, and fewer economic opportunities, which can contribute to lower workforce participation, however socio-economic structures within ST communities might necessitate higher labor force participation rates to sustain their livelihoods.**
- Females who are widowed experience a reduction in the log odds of participating in the labor force by 0.039 compared to females who have never married. Conversely, females who are currently married and divorced exhibit an increase in the log odds of labor force participation by 0.001 and 0.251, respectively, relative to those who have never married. These findings are statistically significant for all categories except females who are currently married. **Widowed females may experience a reduction in labor force participation as loss of a spouse often leads to emotional distress. Furthermore, social stigma associated with widowhood can lead to marginalization and reduced opportunities in the labor market, further discouraging participation. Conversely, divorced females often exhibit increased labor force participation as there is a need for financial independence and stability after divorce. They may also have fewer caregiving responsibilities or may share custody of children, allowing them more**

time and motivation to engage in the workforce.

- A one-unit increase in the amount of land owned by females is associated with a very slight increase in the log odds of participating in the labor force, specifically by 0.00001. **Owning more land leads to higher labor force participation as it necessitates increased agricultural engagement. Larger landholdings require more labor for cultivation, maintenance, and harvesting, thus compelling landowners and their families to participate actively in the workforce.**

4Logit model urban sector

```
> summary(logistic_reg_Urban)
```

Call:

```
glm(formula = Labor_Force$Female_Labor_Force_Participation ~
    Labor_Force$General_Education + Labor_Force$Age + I(Labor_Force$Age^2) +
    Labor_Force$Religion + Labor_Force$Social_Group + Labor_Force$Marital_Status +
    Labor_Force$Land_Owned, family = binomial(link = "logit"),
    data = Labor_Force)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4331	-0.8923	-0.8420	1.4371	1.8616

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.965143429	0.014133875	-68.286	< 0.0000000000000002	***
Labor_Force\$General_Education06	-0.144137268	0.012700558	-11.349	< 0.0000000000000002	***
Labor_Force\$General_Education07	-0.121931615	0.012515644	-9.742	< 0.0000000000000002	***
Labor_Force\$General_Education08	-0.137999785	0.013429963	-10.276	< 0.0000000000000002	***
Labor_Force\$General_Education10	-0.122680789	0.014867953	-8.251	< 0.0000000000000002	***
Labor_Force\$General_Education11	0.230281839	0.040712590	5.656	0.0000000155	***
Labor_Force\$General_Education12	0.028108180	0.015768224	1.783	0.07465	.
Labor_Force\$General_Education13	0.200059837	0.023139614	8.646	< 0.0000000000000002	***
Labor_Force\$Age	0.022555387	0.001046976	21.543	< 0.0000000000000002	***
I(Labor_Force\$Age^2)	-0.000242766	0.000012634	-19.215	< 0.0000000000000002	***
Labor_Force\$Religion2	-0.150641293	0.010020979	-15.033	< 0.0000000000000002	***
Labor_Force\$Religion3	0.169154898	0.019180371	8.819	< 0.0000000000000002	***
Labor_Force\$Religion4	-0.030921659	0.025479013	-1.214	0.22490	
Labor_Force\$Religion5	0.044670252	0.049353916	0.905	0.36541	
Labor_Force\$Religion6	-0.070047924	0.048274155	-1.451	0.14677	
Labor_Force\$Religion7	1.136890291	0.421904456	2.695	0.00705	**
Labor_Force\$Religion9	-0.151839562	0.046355190	-3.276	0.00105	**
Labor_Force\$Social_Group1	0.050034713	0.017414249	2.873	0.00406	**
Labor_Force\$Social_Group2	0.039480014	0.012923157	3.055	0.00225	**
Labor_Force\$Social_Group3	0.003729896	0.008835050	0.422	0.67290	
Labor_Force\$Marital_Status2	-0.132771346	0.013491970	-9.841	< 0.0000000000000002	***
Labor_Force\$Marital_Status3	-0.117918462	0.022175353	-5.318	0.0000001052	***
Labor_Force\$Marital_Status4	-0.014508060	0.052273120	-0.278	0.78136	
Labor_Force\$Land_Owned	-0.000001468	0.000003142	-0.467	0.64019	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The regression results are:

- Females who have completed primary, middle, secondary, and higher secondary education experience a reduction in the log odds of participating in the labor force by 0.144, 0.121, 0.137, and 0.122 respectively, compared to females who are illiterate. Conversely, females who have attained diploma, graduate, and postgraduate levels of education see an increase in the log odds of labor force participation by 0.23, 0.028, and 0.2 respectively, in comparison to their illiterate counterparts. These findings are statistically significant for all categories except females who have attained graduate level of education. **This phenomenon may be attributed to the fact that only after attaining higher levels of education, females exhibit increased participation in the labor force.**
- An increase in age by one year for females is associated with an increase in the log odds of participating in the labor force by 0.022. However, this result contradicts our hypothesis, which posits that female labor force participation should decrease with age. Therefore, the model included (Age)² as an additional predictor variable. The coefficient associated with (Age)² is -0.0002, indicating that an increase in age by one year for females is associated with a decrease in the log odds of participating in the labor force, aligning with our hypothesis.
- Females adhering to Islam, Sikhism, Buddhism and other religions experience a reduction in the log odds of participating in the labor force by 0.150, 0.030, 0.07 and 0.151, respectively, when compared to their Hindu counterparts. In contrast, females who are adhering to Christianity, Jainism and Zoroastrianism exhibit an increase in the log odds of labor force participation by 0.169, 0.044 and 1.136 respectively relative to Hindu females. These findings are statistically significant for all categories except females who are practicing Sikhism, Jainism and Buddhism.
- Females belonging to Scheduled Tribes (ST), Scheduled Castes (SC) and Other Backward Classes (OBC) experience an increase in the log odds of participating in the labor force by 0.05, 0.039 and 0.003, respectively, compared to females from the general category. These findings are statistically significant for all categories except females who belong to Other Backward Classes (OBC). **In urban areas, SC & ST communities often exhibit higher labor force participation than the General category due to economic necessity and limited access to wealth and resources. These communities are more likely to engage in manual and low-skilled labor to meet their livelihood needs.**
- Females who are currently married, widowed and divorced experience a reduction in the log odds of participating in the labor force by 0.132, 0.117 and 0.014 respectively compared to females who have never married. These findings are statistically significant

for all categories except females who are divorced. **Married women often withdraw from the workforce because families with husbands in well-paying, white-collar jobs view it as a matter of pride and status to have their daughters-in-law stay at home, signifying that they are well provided for and do not need to work. For widowed women, social stigma and marginalization further reduce their participation in the labor force. The societal perception of widowhood can lead to isolation and decreased opportunities, making it more challenging for them to engage in employment. Conversely, divorced females often exhibit increased labor force participation as there is a need for financial independence and stability after divorce.**

- A one-unit increase in the amount of land owned by females is associated with a very slight decrease in the log odds of participating in the labor force, specifically by 0.000001, but it is not statistically significant. **Females who own large amounts of land typically possess significant wealth, reducing their financial necessity to work. With ample wealth, they can achieve their desired consumption levels without labor income, leading to a preference for more leisure time.**

Similar Results were found in probit model, **refer to appendix**

.

REFERENCES

- NSSO 68th Round Employment & Unemployment :
<https://microdata.gov.in/nada43/index.php/catalog/127>
- Oxfam India. (2019). *Mind the gap: The state of employment in India*.
<https://d1ns4ht6ytuzzo.cloudfront.net/oxfamdata/oxfamdatapublic/2019-03/Full%20Report%20-%20Low-Res%20Version%20%28Single%20Pages%29.pdf>
- <https://fid4sa-repository.ub.uni-heidelberg.de/3883/1/Job%20Opportunities%20Rural-Urban%20Gradation%20India.pdf>
- Ghose A. K. (2016). *India employment report 2016*. Oxford University Press.
- https://scholar.google.com/scholar_lookup?title=Why+is+female+labour+force+participation+declining+so+sharply+in+India%3F&author=S.+Kapsos&author=A.+Silberman&author=E.+Bourmpoula&publication_year=2014
- <https://www.ilo.org/resource/article/paradox-low-female-labour-force-participation#:~:text=Firstly%2C%20there%20is%20a%20U.education%20do%20participation%20rates%20increase.>
- *The World Bank Economic Review*, Volume 29, Issue 3, 2015, Pages 449–478,
<https://doi.org/10.1093/wber/lhv003>

APPENDIX

1. PROBIT MODEL RURAL SECTOR

```
> summary(probit_reg_Rural)
```

Call:

```
glm(formula = Labor_Force$Female_Labor_Force_Participation ~  
    Labor_Force$General_Education + Labor_Force$Age + I(Labor_Force$Age^2) +  
    Labor_Force$Religion + Labor_Force$Social_Group + Labor_Force$Marital_Status +  
    Labor_Force$Land_Owned, family = binomial(link = "probit"),  
    data = Labor_Force)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4194	-0.9424	-0.8607	1.4025	1.7798

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.5422893251	0.0059501747	-91.138	< 0.0000000000000002 ***
Labor_Force\$General_Education06	-0.0469714345	0.0049446264	-9.499	< 0.0000000000000002 ***
Labor_Force\$General_Education07	-0.0342085441	0.0050895433	-6.721	0.00000000000180 ***
Labor_Force\$General_Education08	-0.0350180424	0.0059319564	-5.903	0.00000000035633 ***
Labor_Force\$General_Education10	-0.0290006838	0.0072134839	-4.020	0.0000581133145 ***
Labor_Force\$General_Education11	0.1158912586	0.0228335025	5.075	0.0000003864947 ***
Labor_Force\$General_Education12	0.0417231711	0.0102932039	4.053	0.0000504639360 ***
Labor_Force\$General_Education13	0.1189850750	0.0197919831	6.012	0.0000000018350 ***
Labor_Force\$Age	0.0091758836	0.0004361657	21.038	< 0.0000000000000002 ***
I(Labor_Force\$Age^2)	-0.0000906807	0.0000053433	-16.971	< 0.0000000000000002 ***
Labor_Force\$Religion2	-0.1725537287	0.0050451899	-34.202	< 0.0000000000000002 ***
Labor_Force\$Religion3	-0.0094767589	0.0076544796	-1.238	0.21569
Labor_Force\$Religion4	-0.1164508504	0.0107329255	-10.850	< 0.0000000000000002 ***
Labor_Force\$Religion5	-0.1151647071	0.0524851137	-2.194	0.02822 *
Labor_Force\$Religion6	0.1791182654	0.0168909480	10.604	< 0.0000000000000002 ***
Labor_Force\$Religion7	-0.4365804613	0.6813490514	-0.641	0.52168
Labor_Force\$Religion9	-0.1051544853	0.0155671140	-6.755	0.00000000000143 ***
Labor_Force\$Social_Group1	0.1303592563	0.0061241824	21.286	< 0.0000000000000002 ***
Labor_Force\$Social_Group2	-0.0052451476	0.0054650445	-0.960	0.33717
Labor_Force\$Social_Group3	-0.0115036913	0.0041427119	-2.777	0.00549 **
Labor_Force\$Marital_Status2	0.0023183883	0.0060999695	0.380	0.70390
Labor_Force\$Marital_Status3	-0.0228703265	0.0099970090	-2.288	0.02215 *
Labor_Force\$Marital_Status4	0.1569152730	0.0285249623	5.501	0.0000000377684 ***
Labor_Force\$Land_Owned	0.0000081647	0.0000006485	12.590	< 0.0000000000000002 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

2. PROBIT MODEL URBAN SECTOR

```
> summary(probit_reg_Urban)
```

Call:

```
glm(formula = Labor_Force$Female_Labor_Force_Participation ~
    Labor_Force$General_Education + Labor_Force$Age + I(Labor_Force$Age^2) +
    Labor_Force$Religion + Labor_Force$Social_Group + Labor_Force$Marital_Status +
    Labor_Force$Land_Owned, family = binomial(link = "probit"),
    data = Labor_Force)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4331	-0.8923	-0.8421	1.4373	1.8704

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.5940470431	0.0085192872	-69.730	< 0.0000000000000002	***
Labor_Force\$General_Education06	-0.0880581137	0.0077038286	-11.430	< 0.0000000000000002	***
Labor_Force\$General_Education07	-0.0743403579	0.0076169280	-9.760	< 0.0000000000000002	***
Labor_Force\$General_Education08	-0.0839230994	0.0081723660	-10.269	< 0.0000000000000002	***
Labor_Force\$General_Education10	-0.0744414528	0.0090547961	-8.221	< 0.0000000000000002	***
Labor_Force\$General_Education11	0.1429700487	0.0252848372	5.654	0.0000000156	***
Labor_Force\$General_Education12	0.0177725354	0.0096724015	1.837	0.06614	.
Labor_Force\$General_Education13	0.1240742126	0.0143240353	8.662	< 0.0000000000000002	***
Labor_Force\$Age	0.0135487541	0.0006345026	21.353	< 0.0000000000000002	***
I(Labor_Force\$Age^2)	-0.0001458717	0.0000076495	-19.069	< 0.0000000000000002	***
Labor_Force\$Religion2	-0.0912858318	0.0060523569	-15.083	< 0.0000000000000002	***
Labor_Force\$Religion3	0.1038179687	0.0117847671	8.810	< 0.0000000000000002	***
Labor_Force\$Religion4	-0.0191079958	0.0155058217	-1.232	0.21783	
Labor_Force\$Religion5	0.0270267482	0.0302473602	0.894	0.37158	
Labor_Force\$Religion6	-0.0426796200	0.0293401921	-1.455	0.14577	
Labor_Force\$Religion7	0.7084606373	0.2607434541	2.717	0.00659	**
Labor_Force\$Religion9	-0.0916755568	0.0279424664	-3.281	0.00103	**
Labor_Force\$Social_Group1	0.0303290073	0.0106390383	2.851	0.00436	**
Labor_Force\$Social_Group2	0.0238632551	0.0078817806	3.028	0.00246	**
Labor_Force\$Social_Group3	0.0019379266	0.0053669578	0.361	0.71804	
Labor_Force\$Marital_Status2	-0.0787374222	0.0082238580	-9.574	< 0.0000000000000002	***
Labor_Force\$Marital_Status3	-0.0697750813	0.0135465915	-5.151	0.000002594	***
Labor_Force\$Marital_Status4	-0.0065363529	0.0321647881	-0.203	0.83897	
Labor_Force\$Land_Owned	-0.0000008819	0.0000019148	-0.461	0.64510	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

