*i*Journals: International Journal of Software & Hardware Research in Engineering (IJSHRE)
ISSN-2347-4890
Volume 10 Issue 4 April 2022

Projecting Teenage Pregnancy and Motherhood and Fertility Rates: An Analytical and Machine Learning Study of Maharashtra, India

Authors: Anya Jajodia[(1)], Reetu Jain[(2)]

[(1)]Class of 2023, Jamnabai Narsee International School (JNIS), Narsee Monjee Bhavan, Narsee Monjee Marg, N.S Road No 7, JVPD Scheme, Vile Parle (W), Mumbai 400 049, Maharashtra, India.
Email id: anya.jajodia@jnis.ac.in

[(2)]Mentor, On My Own Technology Pvt. Ltd., 2A, Pace House, 7 Swastik Society, Gulmohar Road, Vile Parle West, Mumbai - 400049, Maharashtra, India.
Email-id: reetu.jain@onmyowntechnology.com

Abstract

This study determines the impact of factors on teenage pregnancy, motherhood and fertility rates. The study is divided into two halves: the first part determines the fertility trends of teenage pregnancy and motherhood in Maharashtra, while the second half focuses on the correlations of residency, age, marital status, education, religion, and caste and tribe. The datasets were extracted from 5 phases of National Family Health Survey (NFHS) from the years 1992–2021 and integrated. The findings suggest that higher rates of pregnancy are related to rural areas, lower income levels, lesser education, and being a part of a scheduled tribe. A machine learning modeling was employed and the performance of the built model was calculated using five metrics: correlation coefficient, mean absolute error, root mean squared error, relative absolute error, and root relative squared error. The values of the performance metrics were 85.37%, 2.19, 2.72, 49.442% and 51.4241%, suggesting that the regression model is accurate.

Keywords: teenage, pregnancy, motherhood, fertility, analysis, Maharashtra

1. Introduction

Teenage pregnancy or adolescent pregnancy is when a female aged between 13–19 years, who has not completed her core education, has very little or no marketable skills, is financially dependent on their parents or guardians and/or continues to live at home is impregnated (Kumar et. al., 2007). When a teenager decides to keep their baby and starts childbearing then it is called teenage motherhood (Wilson and Huntington, 2006). According to the World Health Organization (WHO), about 16 million girls aging between 15 and 19 years and about two million girls younger than 15 years give birth every year (Idele et. al., 2014).

According to the National Child Development Study (NCDS) report, even after controlling for the influence of family background, teenage mothers were more likely to have negative outcomes than other women. In fact, the majority of the apparent difference in risk was due to the woman's age when she had her first child, rather than her childhood experiences (Luong, 2008).

Teenage pregnancy in many parts of the world is a multifaceted problem with many contributing factors such as poverty, gender inequality, gender-based violence, substance use, poor access to contraception and issues with

pregnancy termination, inconsistent and incorrect use of contraception and a limited number of healthcare providers (Jonas et. al., 2016). Infants born from teenage mothers are more likely to be born prematurely with lower birth weights and with higher neonatal mortality. While mothers have higher rates of postpartum depression and are less likely to initiate breastfeeding (Chen et. al., 2007; Kumar et. al., 2007). Teenage mothers experience high levels of stress which leads to increased mental health problems. It is found that the teenage mothers have higher rates of depression and suicidal tendencies (LePlatte et. al., 2012).

Analyzing the correlation of different socioeconomic factors and its impact on teenage pregnancy and motherhood and deriving a rational conclusion is a case of Exploratory Data Analysis (EDA). It is an important initial step for any knowledge discovery process, in which data scientists interactively explore unfamiliar datasets by issuing a sequence of analysis operations (Milo and Somech, 2020).

The main purpose of EDA is to help look at data before making any assumptions. It can help identify obvious errors, as well as better understand patterns within the data, detect outliers or anomalous events, and find interesting relations among the variables (Velleman and David, 2012). EDA makes it simple to comprehend the structure of a dataset, making data modeling easier. The primary goal of EDA is to clean the data implying that it should be devoid of redundancies. It aids in identifying incorrect data points so that they may be readily removed and the data cleaned (Karamanci, 2009). Furthermore, it aids the analyst in comprehending the relationship between the variables, providing them with a broader view of the data and allowing them to expand on it by leveraging the relationship between the variables. It also aids in the evaluation of the dataset's statistical measurements (Mukhiya and Ahmed, 2020).

Outliers or abnormal occurrences in a dataset can have an impact on the accuracy of machine learning models. The dataset might also contain some missing or duplicate values. EDA may be used to eliminate or resolve all of the dataset's undesirable qualities (Karr et al., 2006). Hence EDA and machine learning (ML) are very important and effective from the perspective of deriving a rational and reasonable conclusion.

1.1. Motivation and Novelties

11% of the world's teenage pregnancies or 16 million teenage pregnancies happen in India (Fulpagare et al., 2019). Keeping this into consideration, we have performed a data analysis in order to find the correlations between various factors with respect to teenage pregnancy and motherhood. To make these correlations and conclusions, we employed the data from all five phases of NFHS extracted specifically for the state of Maharashtra, India from the year 1992 to 2021. NFHS are different phases of surveys conducted throughout India to bring out reliable data on emerging health and family welfare issues (Rajan et al., 2004).

All five NFHS surveys were carried out under the auspices of the Government of India's Ministry of Health and Family Welfare (MoHFW). The International Institute for Population Sciences (IIPS) in Mumbai was designated as the survey's nodal agency by the MoHFW. The Government of India provided funding for NFHS-5. The USAID-supported Demographic and Health Surveys Program, ICF, USA, provided technical assistance and additional funding for NFHS-5. The ICMR and the National AIDS Research Institute (NARI), Pune, assisted with some of the Clinical, Anthropometric, and Biochemical (CAB) tests. The 2019–21 National Family Health Survey (NFHS-5), provides information on population, health, and nutrition for India and each state. 29% of women and 33% of men were in the 15–24 age group, while 32 and 30% of women and men respectively, were in the 25–34 age group. 52% of both women and men were in rural areas. This study will provide trends in teenage pregnancy and motherhood of rural Maharashtra. It will correlate with the variables such as residency, age, marital status, education, TFR, religion, and caste and tribe

Furthermore, in the next part of the study, we made use of machine learning modeling using linear regression. Machine learning is the process of using mathematical models of data to assist a computer in learning without direct instruction. Algorithms are used to identify patterns within data, which are then used to create a data model that can make predictions. Machine learning's adaptability makes it a great choice in scenarios where the data is

constantly changing — in this case, fertility and teenage pregnancy and motherhood rates, making it a great fit for this paper. Under machine learning, a linear regression model has been used.

In this paper, we built correlations and looked for pattern analysis for the age at marriage, total fertility rate, and teenage motherhood, which will help in setting benchmarks for the Sustainable Development Goals at the sub-national level. The 2030 Agenda for Sustainable Development, which was adopted by all UN Member States in 2015, provides a shared framework for stability and prosperity for people and the planet now and in the future. The 17 Sustainable Development Goals (SDGs) are at the heart of it, and they are an urgent call by all nations developed and developing in a global partnership. They recognize that eradicating poverty and other deprivations must be accompanied by actions to improve health care and education, reduce inequality, and stimulate economic growth all while combating climate change and working to protect our oceans and forests. As the next part of the study involves machine learning modeling, we employed a linear regression machine learning algorithm. It is a very powerful technique and can be used to understand the factors that influence profitability used to forecast sales in the coming months by analyzing the sales data for previous months. It can also be used to gain various insights about customer behavior. Mirroring these uses, we employed this algorithm in our study between variables teenage marital status and teenage pregnancy/motherhood.

2. Methods and Materials

This section of the paper briefly discusses the dataset collected for the study. Also gives a brief description of the research questions and the methodology adopted for the study.

2.1 Data Source

For the study, data were gathered from the National Family Health Survey (NFHS) India. Based on the objective of the paper, the required data were specifically extracted for Maharashtra state. NFHS is a large-scale survey which is conducted through different states and households throughout the country providing state and national information on fertility, infant and child mortality, the practice of family planning, maternal and child health, reproductive health, nutrition, anemia, and utilization and quality of health and family planning services (Borkotoky et. al., 2014).

2.2 Study Setting

The present study analyzes the change in the fertility trends among the teenagers in the state of Maharashtra, India over the past few decades. The study identify the causation factors and its relation to the teenage pregnancy and childbearing age. It also identifies the extent to which the causal factors are interconnected with regards to teenage pregnancy, childbearing, and motherhood. Above that, the paper studies the impact of teenage childbearing on the health of the mother and the babies. Finally, the paper

In order to achieve the objectives of the paper, the study is divided into two phases. The first phase is about gathering information of the fertility trends from the NFHS dataset. The data from the five different phases of NFHS viz. NFHS-1, NFHS-2, NFHS-3, NFHS-4 and NFHS-5 for the state of Maharashtra were extracted and integrated to analyze the trends. The factors responsible for teenage pregnancy are identified and establish the causal relationship from the analysis. The second phase analyzes the pattern and then interprets it contextually.

2.3. Variable Description:

The variables included are:

A. **For the fertility trends:** Age of women (15, 16, 17, 18, and 19); NFHS phases (NFHS-1, NFHS-2, NFHS-3, NFHS-4, NFHS-5); Total fertility Rate (TFR) expressed per woman.

B. For teenage pregnancy and motherhood: Age group of women (15, 16, 17, 18, and 19); Residence (Urban or Rural); Marital Status (Never Married or Currently Married); Schooling (no schooling, >5 years complete, 5–9 years, 10–11 years, 12 years or more); Religion (Hindu, Muslim, Buddhist/Neo-Buddhist, and other); Caste and Tribe (Scheduled Caste, Scheduled Tribes, Other Backward Class); Count of Women; Percentage of women pregnant with first child; Percentage of women who have begun childbearing.

2.4 Methodology:

This paper has made use of empirical analysis, pattern analysis, correlations, and trend analysis. Empirical analysis is any research where conclusions of the study are strictly drawn from concretely empirical evidence, and therefore, "verifiable" evidence. There are two ways of gathering empirical evidence, namely qualitative and quantitative (Diefenbach, 2009). Only quantitative research methods have been used in this study. Under quantitative research methods secondary survey research, correlational research, and causal comparative have been adapted for analyzing the dataset. One of the biggest advantages of using quantitative research is that it can focus on facts or a series of information (Sukamolson, 2007). Another benefit of employing secondary research under quantitative methods is availability of similar research in the domain and deriving more useful and consequential information from the analysis (Johnston, 2017).

2.4.1. Exploratory Data Analysis (EDA)

This paper mainly used EDA to analyze and investigate data sets and summarize their main characteristics, often employing data visualization methods. It helps to determine the best way to manipulate data to answer the desired questions, aiding in discovering patterns, spotting anomalies, testing hypotheses, or checking assumptions. Its main purpose is to visualize the data to help arrive at any conclusions (Myatt, 2007).

To achieve the desired result and answer the questions posed, this paper makes use of Multivariate graphical EDA for the first part regarding the fertility trends, and Multivariate and Univariate graphical EDA for the second part concerning teenage pregnancy and motherhood. Multivariate graphical EDA uses graphs and diagrams representing the relationships between more than 2 variables usually displayed with the help of grouped bar charts. Univariate graphical EDA uses graphs and diagrams to show the relationship between any two variables.

3. Results

The section of the paper briefly describes the result obtained after analyzing the data sets by EDA approaches.

3.1 NFHS Teenage Fertility Trends

In this section of the paper that data collected are analyzed to evaluate the teenage fertility trend for the state of Maharashtra. Figure 3 represents the trends in TFR according to residence. It is a multivariate graph which shows the correlation between the Urban, Rural, and Total rates of fertility throughout the 5 phases. The total column can be interpreted as the average fertility rates throughout each phase.

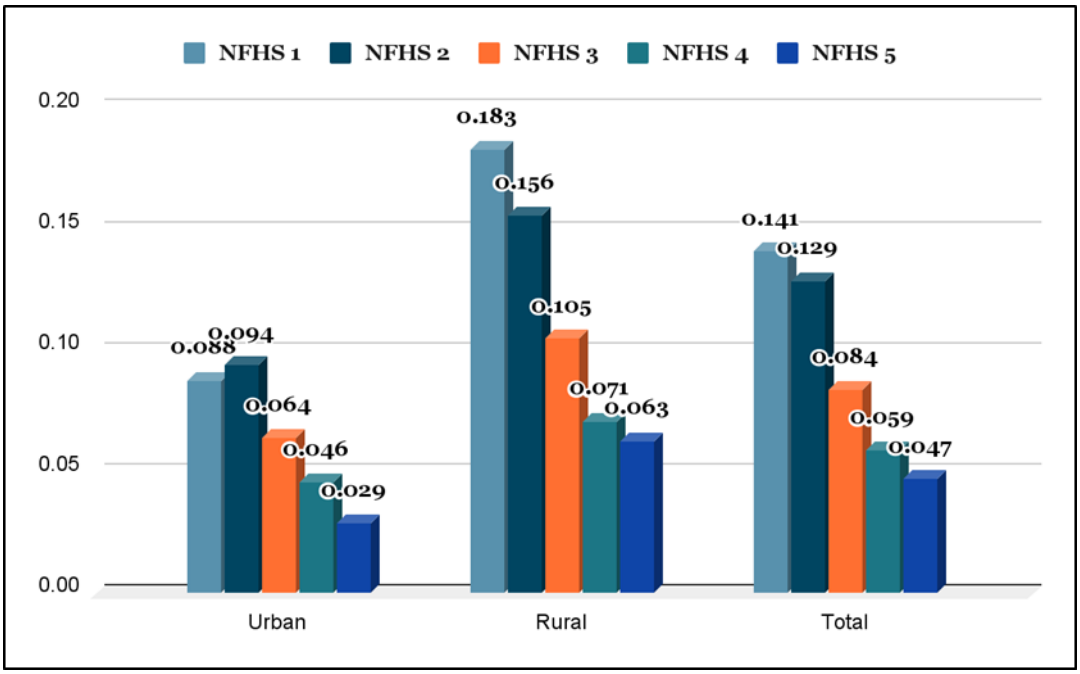

Figure 1: Trends in Teenage TFR per Woman by Residence for Maharashtra.

From the trends of Figure 1, it is clear that the teenage fertility and pregnancy rates in rural areas are higher in each phase when compared to the Urban rates. In the last 23 years i.e., from NFHS-2 to NFHS-5, the graph shows a clear decline for the urban rates of fertility compared to the earlier time period between 1992 to 1999 i.e., NFHS-1 to NFHS-2, where the fertility rates increased from 0.088 to 0.094 (8.8% to 9.4%). The overall fertility slowdown from 0.088 to 0.029 (8.8% to 2.9%) throughout the 5 phases is mainly due to plateauing of fertility in urban areas (Figure 1).

As opposed to the Urban rates, in Rural areas the fertility rate at no given point increases, it shows a steady decrease in the rates right from NFHS-1 to NFHS-5 showing a clear reduce of the rate at which fertility rates were decreasing over the periods of 29 years during which this study took place.

TFR (Figure 1) in Maharashtra is 0.047 children per woman (NFHS-5) which is below replacement level fertility. The TFR for teenagers has reduced from NFHS-1 to NFHS-5 with the greatest rate of decline being from NFHS-2 to NFHS-3 i.e., from the years 1998 to 2006. After this, due to the plateauing of fertility in urban areas, the average rate of decrease steadily decreased from 0.045 from NFHS-2 to NFHS-3 to 0.025 from NFHS-3 to NFHS-4 and finally to 0.012 in the final phase of the study.

Figure 2 shows the bar plot for the teenage TFR per woman by residence for India and figure 3 shows the teenage TFR per 1000 women for the different Indian states.

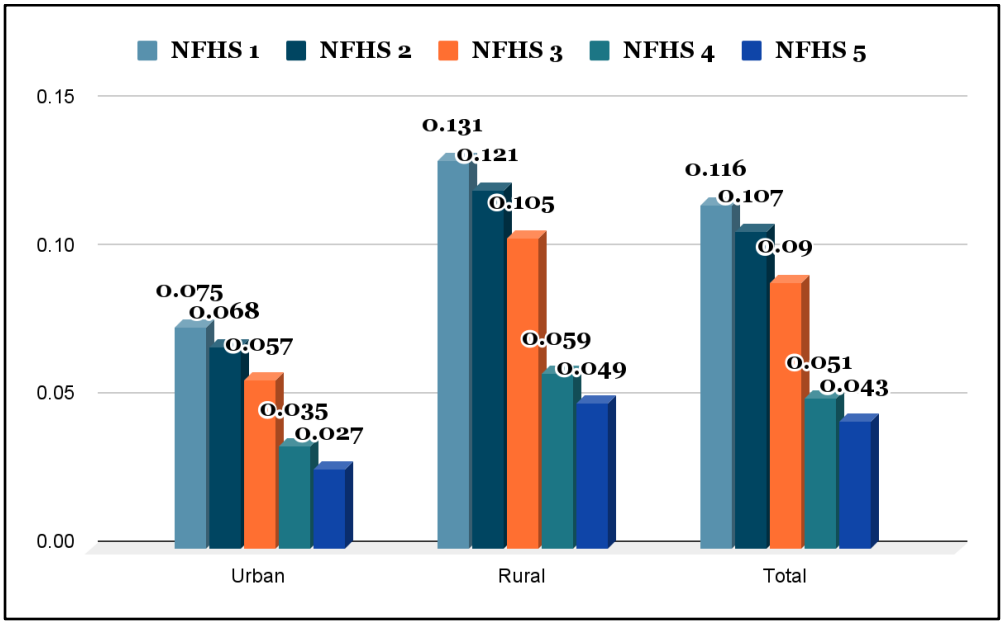

Figure 2: Trends in Teenage TFR per Woman by Residence for India.

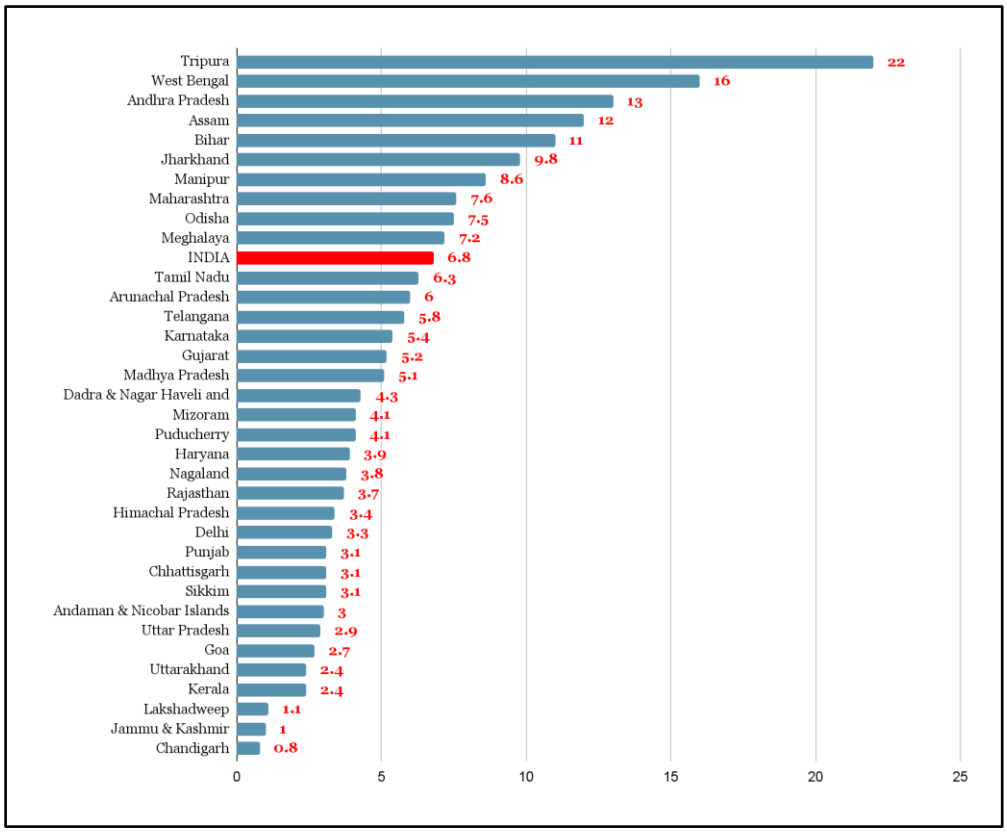

Figure 3: Teenage TFR per 1000 Women for different states of India.

There is an evident decline seen in the TFR of teenager in each phase of the survey. As seen in figure 3, the average TFR for teenagers is 6.8, however rates in Maharashtra among a few other states exceeds that. In India, the TFR of teenagers is decreasing over the years which can be verified by the data available from the papers by (Rao et. al., 1976); 10.6%, (Bhattacharya and Chaudhary, 1986): 14.7%, (Chhabra, 1991): 8.2% and (Mahavarkar, 2008): 10%. Despite the decrease in the TFR, it is very high compared to countries like the UK, Germany, France and Netherlands therefore making teenage childbearing a major problem in India (Mahavarkar,

2008). Teenage childbearing is vastly dependent on cultural and socio-economic factors. In India, despite legislation against child marriage and teenage pregnancy in place, such as the 1978 amendment to the Sharada Act, it continues to be an issue posing a challenge to healthcare policy makers till this day. Teenage pregnancy is a common occurrence amongst the rural population in different states of India.

In Urban areas of Maharashtra, 1.3% teenagers between the age of 15 and 19 are pregnant with their first child and 3.9% of them have started childbearing. In contrast, the numbers in rural areas are much higher. There, 3.1% of the teenagers are pregnant with their first child while 10.6% of them have started childbearing. One of the reason for such high rates of teenage pregnancy in rural areas of Maharashtra could be due to various cultural practices, early menarche, and a fairly high fertility rate in this age group. This can also related to the poor literacy rates and the social role of women in our society. With increasing education and socioeconomic status of rural female population the incidence seems to be reducing, but it still remains a problem leading to the in-depth analysis of teenage pregnancy.

3.2. Teenage Pregnancy and Motherhood

This subsection of the paper briefly identifies the different factors and its impact on teenage pregnancy and motherhood.

3.2.1. Age and Pregnancy/Motherhood relationship

Table 1 represents the count and percentage of women per age category for teenage pregnancy and motherhood. It helps gain a more holistic image of comparisons between the ages and how the values are close together for the ages from 15–19. The pregnancy and motherhood rates shoot up for the age group of 19 and are the highest through the age groups analyzed.

Table 1: Age vs Count of Women for Teenage Pregnancy and Motherhood.

Age	Number of Women	% of Count of Women per Age
15	964	20.50
16	926	19.69
17	911	19.37
18	889	18.90
19	1013	21.54

Figure 4 (a) and figure 4 (b) present the line graphs representing the correlation of age and teenage pregnancy and age and teenage motherhood respectively.

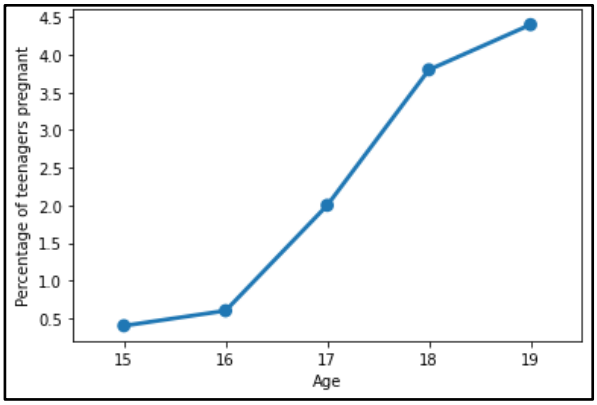

Figure 4(a): Age vs Teenage Pregnancy

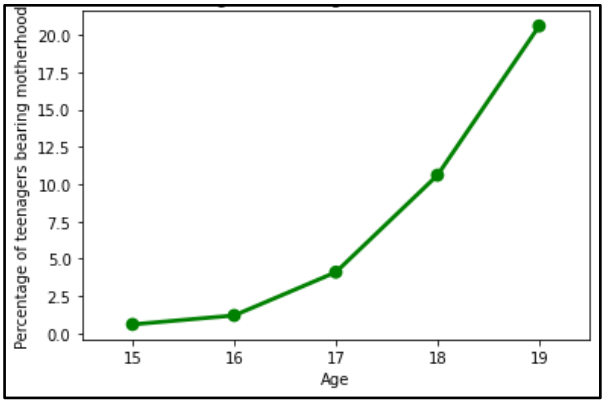

Figure 4(b): Age vs Teenage Motherhood

The percentage of pregnant teenagers only increase with age as seen in Figure 4(a). The rate at which teenage pregnancies increase is low between the ages of 15 and 16 compared to the trends between ages 16–18. It reduces once again from 18 to 19. This trendline reflects a certain psychological phase of adolescence and is called "rebellion in late adolescence" (Pickhardt, 2009). This behavior of teenagers as they are exploring the world around them causes them to be rebellious therefore leading to higher number of pregnancies but lower rates of childbearing at this age.

As represented in figure 4(b), the relationship between age and teenage motherhood is clearly exponential, increasing at a faster rate as the ages progress from 15 to 19. This alludes to the maturity as one grows up, and thus their increasing ability to financially support another human being. The gradient of the straight line joining the two points on the graph increases as the age increases, therefore signifying the increasing rate of increase for the childbearing abilities of teenagers as they grow.

Among young women aged 15–19 in Maharashtra, 2.24% of teenagers are pregnant which is down from 2.40% in NFHS-4, 7.42% have already begun childbearing down from 9.44% in NFHS-4. The proportion of women who have started childbearing rises sharply from 5 percent at age 17 years to 13 percent among women aged 18 years and to 21 percent among women aged 19 years.

The TFR in teenagers is declining due to a number of reasons, some of them listed below:

I. Increasing rates of education for girls, proactive government investments in adolescent girls, and strong public awareness against child marriage and teenage fertility.
II. Decline in child marriage rates from 47% to 27% between 2005–2006 and 2015–2016 (Paul, 2020).
III. Better education and inclusion on the topic of sexuality is now a part of India's school curriculum (Malik, 2022).
IV. Easy access and increased usage of contraception (Mehta et. al., 2020). In the state of Maharashtra around 66% increase in the use of contraception is observed between the NFHS phase 4 to NFHS phase 5.

3.2.2. Education and Teenage Pregnancy/Motherhood relationship

The literacy rate of the female population of Maharashtra has increased from 43.5% in 1981 to 75.48% in 2011. **Table 2** shows the literacy rate for Maharashtra from 1981 to 2011.

Table 2: Literacy rate (expressed in percentage) of Maharashtra

Census	Female	Male	Total
1981	43.5	70.06	57.24
1991	52.32	76.56	64.87
2001	67.03	85.97	76.88
2011	75.48	89.82	82.91

*Data: (Census of India 2011)

The rise in the female literacy rate can be credited till some extent to the 1993 scheme in which the Maharashtra government provides incentives such as Attendance Allowance to females of economically disadvantaged communities aimed at encouraging female education. The rise in female literacy rate plays a crucial role in declining the TFR in the state of Maharashtra.

Although the TFR of teenagers is declining, it remains an issue in the state. This is to be statistically scrutinized in order to get a better in depth of the scenario. The count per schooling level from the age of 15–19 is categorized as follows:

- No schooling
- < 5 years complete
- 5–9 years complete
- 10–11 years complete
- > 12 years complete

Figure 5 represents the above values as percentages divided into 2 parts, pregnancies and motherhood, for each schooling category.

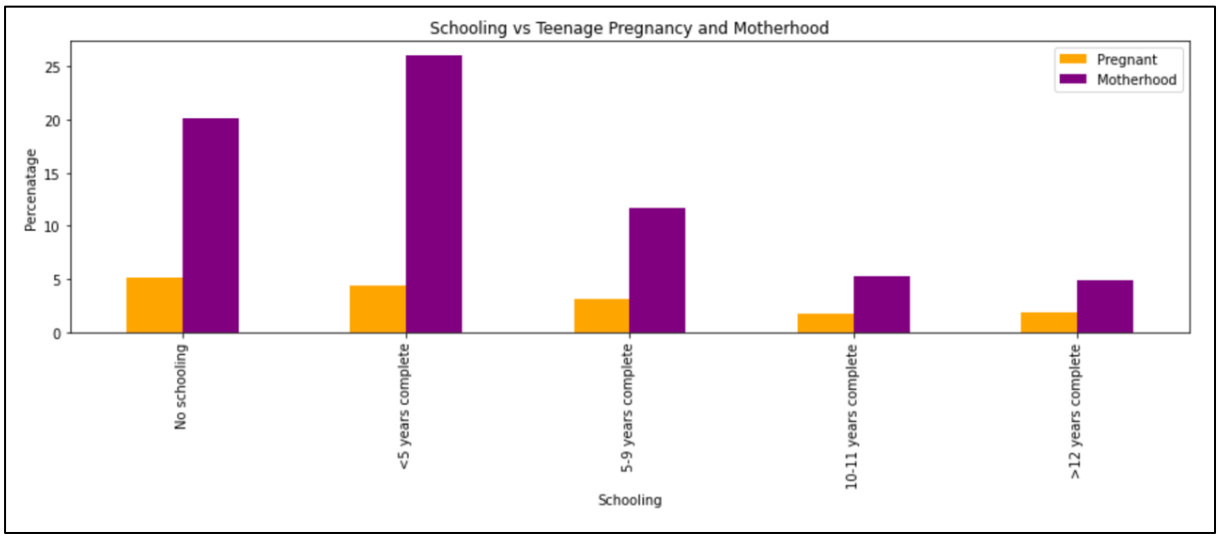

Figure 5: Schooling vs Teenage Pregnancy and Motherhood.

The percentage of women who were pregnant with their first child when they had no schooling (5.1%) was higher in comparison to the pregnancies of other 5 divisions (4.4%, 3.1%, 1.7% and 1.9% respectively). This implies the connection between illiteracy and pregnancy rates — illiteracy indirectly leads to higher rates of teenage pregnancies. This directly relates to their residences as only 77.09% of the rural population is educated as opposed to 89.84% in rural areas. Therefore, from Figure 5, it can be inferred that women with a far lesser educational background are more prone to adolescent pregnancy.

The percentage of teenagers who had started childbearing and had no schooling (15%) and the percentage of childbearing teenagers with lesser than 5 years of education (21.6%) both are almost double than the childbearing percentages of teenagers with 5–9 years of schooling (8.5%), 10–11 years complete (3.7%), and more than 12 years of schooling completed (2.9%). The downward sloping trendline after the '<5 years complete' category proves that the more educated the woman is, the lesser the chance of them bearing a child as a teenager.

3.2.3. Economic perspective of Child Marriage and Teenage Pregnancy/Motherhood relationship

According to a report by the UN, India could add $7.7 billion every year to its economic productivity if its young girls are able to study and work till their 20s instead of becoming mothers at an adolescent age (Modi, 2017). Analyzing the global data revealed that the girls from the poorest 20 percent of families are twice as likely to marry before the age of 18 when compared to girls from the richest 20 percent of families. This stems from the traditional perception that girls are financial burdens rather than potential wage earners in rural areas. Families living in poverty with several children use child marriage as a way to reduce their economic burden. To them, one

less daughter means one less person to feed, clothe, and educate. Families often use child marriage as a strategy to evade food insecurity. Girls are even used as a substitute for money to offset debts and settle conflicts. In cultures where the bride's family is expected to pay a dowry, early marriage equates to a lower bride price. In cultures where the groom's family pays the dowry in exchange for the bride, younger girls fetch a higher price. The families that cannot afford to raise their daughters may perceive child marriage as the next best alternative and a source of income (Wahhaj, 2015). However, due to the COVID-19 lockdown, the women and child development (WCD) department has revealed that child marriages in the state have increased by 78.3%. It can be speculated that the increase in the number of child marriages is due to strict lockdown and the impact it has had on the rural economy (Chakraborty, 2020).

In the state of Maharashtra, menarche at young age increases the risk of child marriage (Raj et. al., 2015). The frequency of child marriage in the state of Maharashtra is lower than the national average and mostly occurs among girls between 16 and 18 years. Youth Study findings underscore the prevalence of early marriage among young women in the state. Of those aged 20–24, almost one in 10 young women were married before age 15 and over one-third before age 18. Young women in rural areas were twice as likely as young women in urban areas to be married before age 18; 46% of young women in rural areas compared to 22% of young women living in urban areas were married before age 18. In contrast, just 2% of men aged 20–24 were married before age 18 (Ram, 2008). Figure 6 shows the bar plot for the percentage of young women aged 20–24 who were married before selected ages.

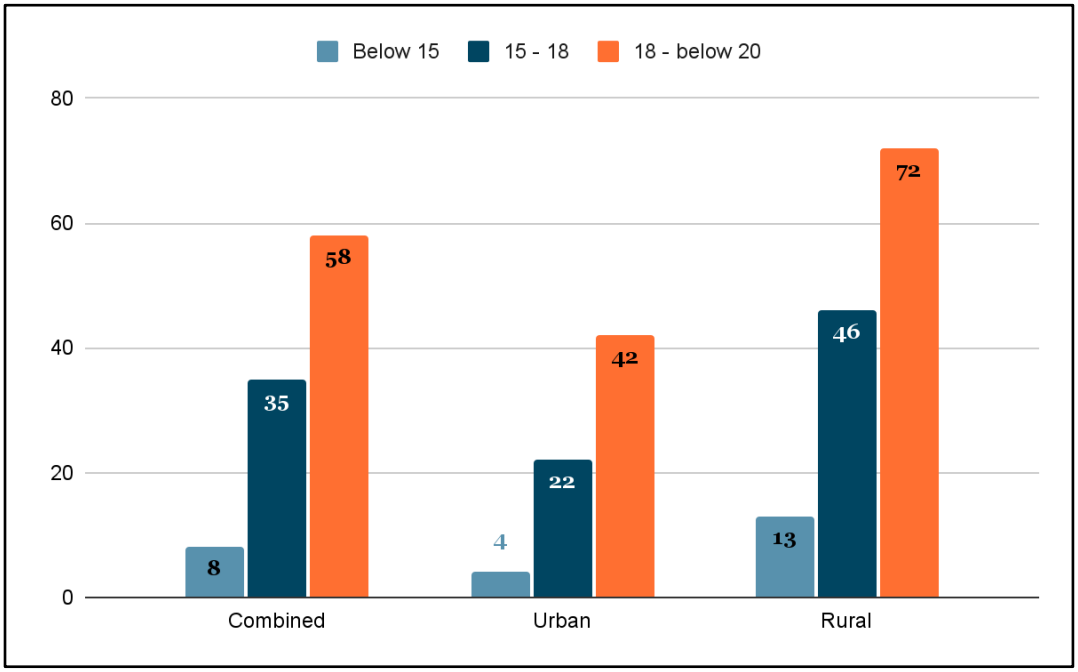

Figure 6: percentage of young women who were married before selected ages

The striking difference in the magnitude of early marriage by years of schooling completed by young women is of note. About four in five young women with no formal schooling were married before age 18. The proportion of young women married before age 18 declined to 33% among those who had some secondary education and 4% among those who had completed at least 12 years of schooling. This finding is also summarized in the section 3.2.2.

The above figure shows a decline in the trends for child marriage in Maharashtra. This is due to the steps taken by the government to stop child marriage, which has also resulted in a decline of the adolescent fertility rates (AFR).

3.2.4. Religion and Teenage Pregnancy and Motherhood relationship

Figure 7 shows the relation between the percentage of pregnant teenagers in each religion in Maharashtra. In Maharashtra, out of the total teenage pregnant girls it is estimated that 40.54% are Muslims, 29.73 % are Hindus and Buddhist or Neo-Buddhist. However, in India, the number of teen pregnancies was highest among Muslims, where 9% of teenage women were mothers, and it was the lowest among Jains (Vinathi, 2021). The most probable reason for high teenage pregnancy among Indian Muslims is that Muslim girls can marry after their puberty starts which is at the age of 15 years, while other religious communities cannot.

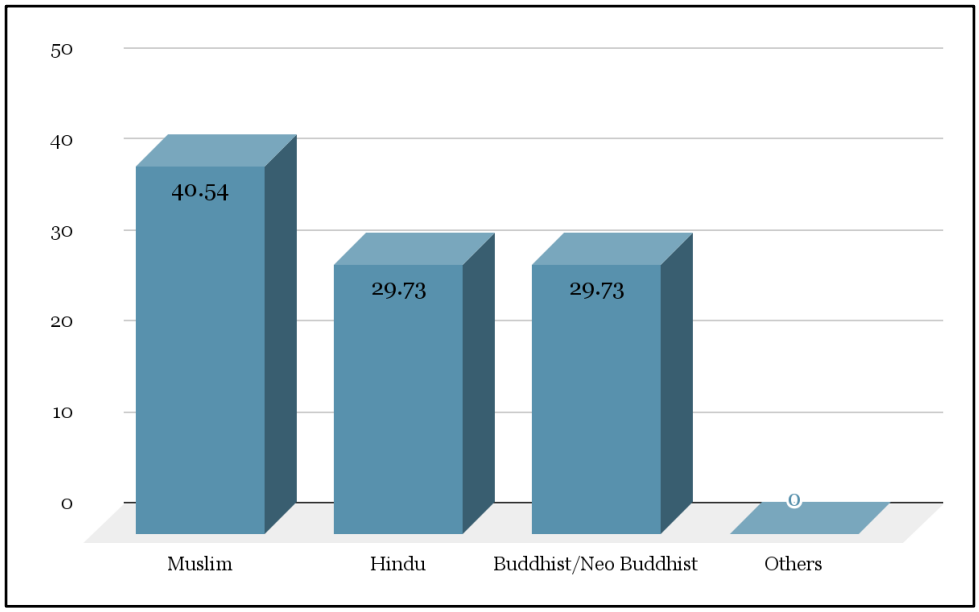

Figure 7: Religion vs Teenage Pregnancy.

3.2.5. Caste/Tribe vs Teenage Pregnancy and Motherhood:

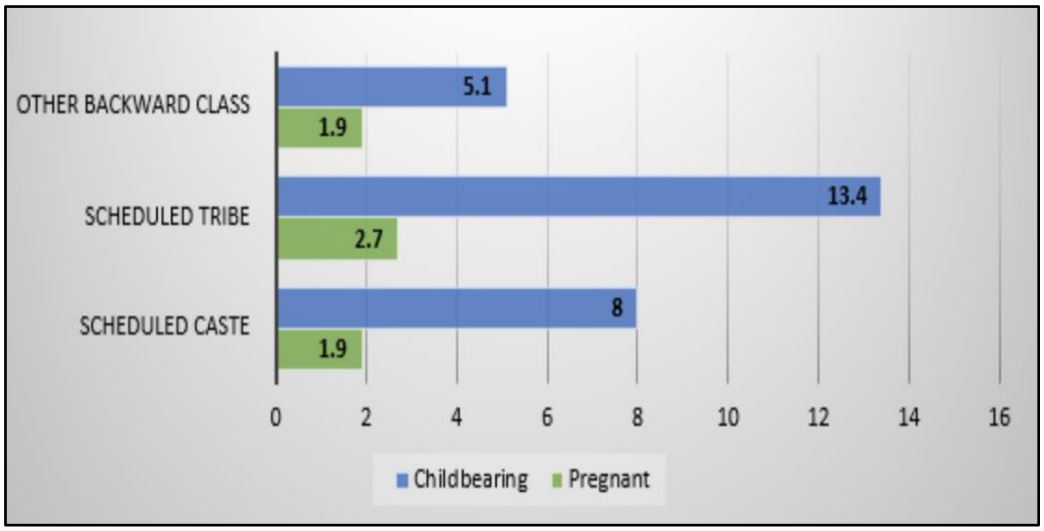

Figure 8: Class/Tribe vs. Pregnancy and Childbearing.

As seen in figure 8, the rates for childbearing and teenage pregnancy in scheduled caste is 8% and 1.9%, scheduled tribe 13.4% and 2.7% and 5.1% and 1.9% in other backward classes respectively.

4. Machine Learning Modeling

Machine learning, and more precisely predictive modeling, is primarily concerned with minimizing a model's error or producing the most accurate forecasts attainable. Linear regression algorithm is studied as part of machine learning as a model for comprehending the relationship between numerical input and output variables and is a combination of statistics and a machine learning method.

Linear Regression is a supervised machine learning algorithm using regression models. Based on independent variables, regression models a target prediction value mostly used to determine the relationship between variables. Different regression models differ in terms of the type of relationship they evaluate between dependent and independent variables and the number of independent variables they employ. The linear regression model establishes a linear relationship between a dependent variable and one or more independent variables, thus the name. Because linear regression demonstrates a linear relationship, it identifies how the dependent variable's value changes in response to the independent variable's value.

In this study, linear regression algorithms are used to determine the relationship between variables teenage marital status and teenage pregnancy/motherhood. The results are presented as follows:

4.1 Correlation Analysis:

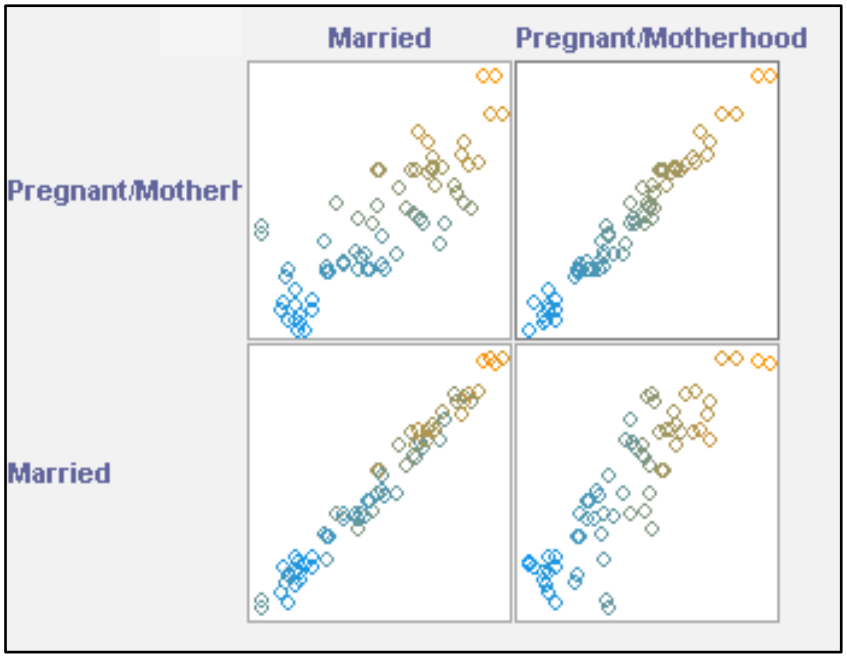

Figure 9: Correlation Matrix.

Figure 9 presents the plot for correlation built between teenage marital status and teenage pregnancy/motherhood. From Figure 9 it can be inferred that both these variables are positively correlated. Positive correlation is when an increase in the dependent variable on the Y-axis leads to an increase of the independent variable on X-axis. In this case, pregnant/motherhood is the dependent variable and marriage is the independent variable.

4.2: Model Building:

With a positive correlation established between the pregnant/motherhood and married variables represented by Figure 9, we present the linear regression modeling results in Table 3.

Table 3: Machine Learning Modeling Results.

Modeling properties	Values
Correlation coefficient	0.8537
Mean absolute error	2.1918
Root mean squared error	2.7215
Relative absolute error	49.442%
Root relative squared error	51.4241%

1. **Correlation coefficient:** The linear correlation coefficient is a numerical value generated from data given that indicates the strength of the linear relationship between two variables: *x* and *y*. The direction of the linear relationship between *x* and *y* is indicated by the sign of the linear correlation coefficient.
2. **Mean absolute error:** This is a metric for evaluating regression models. The mean absolute error of a model in relation to a test set is equal to the mean of the absolute values of the individual prediction errors on all instances in the test set.
3. **Root mean squared error:** Root Mean Squared Error is the square root of Mean Squared error. It quantifies the residuals' standard deviation.
4. **Relative absolute error:** A ratio expressing the difference between a mean error (residual) and the errors produced by a trivial model. A good forecasting model generates a ratio near to zero; a bad forecasting model generates a ratio greater than one.
5. **Root relative squared error:** The root relative squared error is the error that would have occurred in the absence of a simple predictor. More precisely, this straightforward predictor is simply the average of the observed values.

As a result of the preceding findings, we can conclude that the correlation coefficient of 0.85 indicates a strong association between these two variables. It indicates that teenage pregnancy/motherhood and marriage are substantially connected and follow a linear trajectory. Additionally, the lower mean absolute error of 2.19 and root mean squared error of 2.72 values indicate that the developed regression model is more accurate. Thus, root mean squared error indicates the accuracy with which a linear regression model fits the NFHS data.

With relative absolute error of 49%, approximately 0.5 in terms of ratio, and root relative squared error value of 51% (0.5 in terms of ratio), shows the classifier errors. The root relative squared error quantified a model's performance in comparison to the mean of the true values. As a result, the developed model outperforms the simple model, as the root relative squared error is less than one. Hence, the lower value of root relative squared error indicated that an accurate model was created.

5. Discussion

Every year, an estimated 21 million girls between the ages of 15 and 19 become pregnant in underdeveloped countries, with around 12 million of them giving birth. In poor nations, at least 777,000 adolescent girls under the age of 15 give birth every year (Ursachi, 2020). Over the last 20 years, the estimated global adolescent-specific fertility rate has decreased by 11.6%. However, there are significant regional variances in these rates. For example, East Asia has a 7.1 adolescent fertility rate whereas Central Africa has an adolescent fertility rate of 129.5 (Monteiro et. al., 2021).

In this study, we saw that for all the five phases, the TRF was decreasing at an increasingly slow rate and the fertility trends were always higher in rural areas when compared to urban areas of Maharashtra. These trends were a result of the higher illiteracy rates in rural areas and lack of awareness about implications of teenage pregnancies. Another reason for this is the traditionalist mindset in villages where women get married as teenagers most of the time.

According to a UN report, India bears economic losses of $7.7 billion a year due to teenage pregnancies. An earlier estimate by the health ministry suggested economic losses of teenage pregnancies at 12% of the gross domestic product (GDP). In this study, the pattern of exponentially increasing rates of childbearing and motherhood as ages increase from 15 to 19 is seen. The increasing rates of pregnancy seen between the ages of 16 and 18 also signify the late rebellious stage instilling in teenagers.

According to a WHO report, adolescent pregnancies are a huge problem. In a UNICEF report published in 2013, it was said that teen pregnancies occur mostly in marginalized communities due to poverty, illiteracy and reducing employment opportunities. Many times, girls choose to get married and get pregnant at a young age because of their limited access to educational and employment prospects. In numerous societies, marriage and childbearing is the best option available to many young mother's where motherhood is encouraged by society. In the less developed countries, at least 39% of girls marry before the age of 18 years and 12% before the age of 15 (Godha et. al., 2013). According to the Observer Research Foundation, these pregnancies not only make teenagers very vulnerable both physically and mentally, but it also places them and their babies at risk as their bodies are not prepared for it yet. This leads to higher rates of miscarriages, abortions, and other adverse outcomes (Suri, 2020).

In this performed research, it was seen that as the level of schooling increases, the rates of pregnancy and motherhood decreases, and that pregnancy and motherhood rates are double in rural areas when compared to urban areas which can be inferred as lack of education. The study also indicates that none of the childbearing or pregnant teenagers are not married, also alluding to breaking the law of the legal age of marriage.

According to an research conducted by the Observer Research Foundation, in Maharashtra, women aged 20 to 49 had a median age of first marriage of 19.6 years (Suri, 2020). In NFHS-4, nearly a quarter of women in their 20s and 30s (22%) married before attaining the legal minimum age of 18, compared to 26% in NFHS-4. In Maharashtra, 40% of women between the ages of 20 and 24 have never married, compared to 86% of men in the same age range, demonstrating that women marry at a younger age than men.

Although the trend of early marriages is declining, at least 27% of women marry before reaching the legal marriage age of 18. The trend is higher in West Bengal (44%), Bihar (42%), Jharkhand (39%), and Andhra Pradesh (36%), and lower in Lakshadweep (5%), Jammu and Kashmir (39%), and Kerala (36%). (9%) (Suri, 2020). This demonstrates the link between high rates of adolescent marriage and low levels of education. Most of these places are less developed when compared to other cities in India such as Delhi and Mumbai, they can be considered as the more rural areas of the country. Taking that into consideration, it also shows a link between lower levels of education, rural areas and higher rates of early marriage and childbearing.

From the results procured from this study, it is clear that the majority of the teenagers who wound up pregnant were Hindus. In each individual religion as well, 70% or more of the women surveyed who either were pregnant, or were childbearing. The fact that most of the teenagers pregnant were Hindus can be a result of the illegal custom of child marriage in rural areas where fathers sell off their daughters to other men at a young age (8–15). Religion, caste/tribe, and education have the greatest differences in fertility. At current fertility rates, women with no education will have 0.6 more children than women with 12 or more years of education. Muslim women will have 0.4 more children than Hindu women (a TFR of 2.1 vs. 1.7) and 0.7 more children than Buddhist/ Neo-Buddhist women (a TFR of 1.4) (Patra and Singh, 2013).

Early adolescent pregnancies have major health consequences for adolescent mothers and their babies, according to (Neal et al., 2012) Prenatal and perinatal issues are the leading cause of death among girls aged 15–19 worldwide, with low and middle-income nations accounting for 99 percent of global maternal deaths among women aged 15–49. According to a WHO paper published in 2016, adolescent mothers aged 10–19 years have a higher risk of eclampsia, puerperal endometritis, and systemic infections than women aged 20–24 years. Furthermore, approximately 3.9 million risky abortions among girls aged 15–19 years are performed yearly, contributing to maternal deaths, morbidity, and long-term health problems. According to an ORF report,

approximately 54.1% of adolescent girls aged 15–19 are anemic, with a relatively increased occurrence in rural areas than in urban areas. This is due to the fact that only 28.1 percent of teenage expecting mothers took iron and folic acid tablets, which are critical in preventing anemia during pregnancy. Similarly, at least 42 percent of teen girls in India have a BMI of less than 18.5. As a result, pregnancy in such a vulnerable state reinforces the cycle and causes neonatal and infant deaths. A recent study discovered that children born to teenage mothers have greater stunting and underweight rates. Low education, poor nutritional status, and high anemia rates all contribute to the vicious cycle. Low food intake with a diet deficient in fruits and milk, as well as wide variation in meals from required food groups across states, contributes to women's poor nutritional status (Berthoud and Robson, 2001).

According to Raj and Boehmer (2013), social implications for unmarried pregnant adolescents also include stigma, rejection, or violence from partners, parents, and peers. Girls who become pregnant prior to actually reaching the age of 18 are more likely to experience violence in a marriage or partnership. According to the WHO paper 'Making health services adolescent friendly: Developing national quality standards for adolescent friendly health services,' adolescent pregnancy and childbearing frequently leads to girls dropping out of school, and while efforts are being made in some places to allow them to return to school after childbirth, this may jeopardize girls' future education and employment opportunities (Berthoud and Robson, 2001).

From the trends analyzed in the schooling patterns, it can be inferred that if female education is given a boost, the rates of adolescence pregnancy and motherhood can considerably get reduced. Providing education for females also has other benefits apart from just reducing teenage pregnancy rates as it ensures a higher chance of them having a stable and independent future. Free education or education at a lower cost provided by the government can aid this development considering the scale of this issue. A sex education program being included in existing curriculums will encourage females to be more open about the subject. Knowing more about the topic will educate them on different methods of contraception and will aid in reducing the pregnancy rates in wealthier regions/cities. If everyone is well educated about the heinous consequences of teenage pregnancy and motherhood, they will be more likely to practice safe intercourse and be more careful.

The evidence consistently shows that, after controlling for other factors, the age at first birth has no direct impact on female earnings. However, it appears that there are some indirect effects. Card (1977) discovered that when only background variables were controlled for, teen child bearers earned less than later child bearers or those childless at all follow-ups. These intervening effects have also been identified by other researchers (Hofferth et al., 1987).

The provision of a sex education program, as mentioned in a paper written by Patra and Singh (2013), can be beneficial in increasing awareness about numerous contraception methods and consequences of teenage pregnancy and motherhood (Patra and Singh, 2013). In this regard, example of Sweden's success can be cited (as supported by a paper explaining trends in teenage childbearing in Sweden written by Santow G and Bracher M) in reducing teenage pregnancy rates and birth rates which is credited to both improved sex education and improved provision of contraceptives to teenagers (Berthoud and Robson, 2001).

From this and the results of this study combined, it can be inferred that early birth appears to increase family size, which decreases the proportion of years worked and hours worked in the previous year, lowering earnings at the age of 27. Early birth brings down schooling, which minimizes the proportion of years worked, hours worked, and earnings at the age of 27. Adding all the effects up, early age at first birth is associated with reduced earnings, but this is because it is associated with reduced schooling and increased family size.

With the machine learning results, the study indicates that there was a strong relationship between teenage pregnancy/motherhood and marriage and followed a linear trajectory. The lower mean absolute error and root mean squared error values indicated that the developed regression model is more accurate. Also, the root relative squared error quantified a model's performance in comparison to the mean of the true values. As a result, the

developed model outperformed the trivial model, as the root relative squared error was found to be less than one. Hence, the lower value of root relative squared error indicated that an accurate model was built.

6. Conclusion

From this study, it can be concluded that although the average fertility rates have decreased in the past 29 years, it is still a huge problem. From the results and discussion, it can be inferred that avoiding teenage pregnancy and motherhood allows women to complete their education and secure a stable income and job later on therefore reaping them long term benefits. This will also reduce infant and maternal mortality rates considering the number of teenagers aged 15–19 die each year due to premature childbearing. Having a child at a young age also has severe social consequences especially if the child bearer is still reliant on their parents or guardians financially. The study's findings indicate that, in order to protect adolescent women's reproductive rights and monitor their health complications (due to early pregnancy), state-level programs and strategies must be implemented, with a focus on teenage women from economically disadvantaged backgrounds.

References

Ahmed, Usman, and Suresh Kumar Mukhiya. 2020. *Hands-On Exploratory Data Analysis with Python*. Packt Publishing Ltd.

Berthoud, Richard, and Karen Robson. 2001. *The Outcomes of Teenage Motherhood in Europe*. Innocenti Working Paper 86. Florence: UNICEF Innocenti Research Centre.

Bhattacharya, Anima, and Nandita Chowdhury. 1986. 'Teenage Primigravidae'. *Journal of Obstetrics & Gynecology* 36: 660–62.

Borkotoky, Kakoli, and Sayeed Unisa. 2014. 'Indicators to Examine Quality of Large Scale Survey Data: An Example through District Level Household and Facility Survey'. Edited by Thomas Eisele. *PLoS ONE* 9 (3): 1–11. https://doi.org/10.1371/journal.pone.0090113.

Census of India. 2011. 'Single Year Age Returns by Residence, Sex and Literacy Status (India & States/UTs)'. Government of India. https://censusindia.gov.in/census.website/data/census-tables.

Chakraborty, Rupsa. 2020. 'Child Marriages in Maharashtra Surge by 78.3% amid Lockdown as Families Reel under Poverty'. *Hindustan Times*, 7 October 2020.

Chen, Xi-Kuan, Shi Wu Wen, Qiuying Yang, and Mark C. Walker. 2007. 'Adequacy of Prenatal Care and Neonatal Mortality in Infants Born to Mothers with and without Antenatal High-Risk Conditions'. *The Australian and New Zealand Journal of Obstetrics and Gynaecology* 47 (2): 122–27. https://doi.org/10.1111/j.1479-828X.2007.00697.x.

Diefenbach, Thomas. 2008. 'Are Case Studies More than Sophisticated Storytelling?: Methodological Problems of Qualitative Empirical Research Mainly Based on Semi-Structured Interviews'. *Quality & Quantity International Journal of Methodology* 43 (6): 875–94.

Fulpagare, Prakash H., Abhishek Saraswat, Konsam Dinachandra, Nikita Surani, Rabi N. Parhi, Sourav Bhattacharjee, Somya S, et al. 2019. 'Antenatal Care Service Utilization Among Adolescent Pregnant Women-Evidence From Swabhimaan Programme in India'. *Frontiers in Public Health* 7: 1–13. https://doi.org/10.3389/fpubh.2019.00369.

Godha, Deepali, David R. Hotchkiss, and Anastasia J. Gage. 2013. 'Association Between Child Marriage and Reproductive Health Outcomes and Service Utilization: A Multi-Country Study From South Asia'. *Journal of Adolescent Health* 52 (5): 552–58. https://doi.org/10.1016/j.jadohealth.2013.01.021.

Gurijala, Vinathi. 2021. 'Teen Pregnancies Drop In India, Urban Rural Divide Still High'. *BehanBox* (blog). 5 February 2021. https://behanbox.com/2021/02/05/teen-pregnancies-drop-in-india-urban-rural-divide-still-high/.

Hofferth, Sandra L. 1987. 'Social and Economic Consequences of Teenage Childbearing'. In *Risking the Future: Adolescent Sexuality, Pregnancy, and Childbearing*, edited by Sandra L. Hofferth and Cheryl D. Hayes, II: Working Papers and Statistical Appendices:123–44. Washington, D.C.: National Academies Press.

Idele, Priscilla, Amaya Gillespie, Tyler Porth, Chiho Suzuki, Mary Mahy, Susan Kasedde, and Chewe Luo. 2014. 'Epidemiology of HIV and AIDS among Adolescents: Current Status, Inequities, and Data Gaps'. *Journal of Acquired Immune Deficiency Syndromes (1999)* 66 (Suppl 2): 144–53. https://doi.org/10.1097/QAI.0000000000000176.

Johnston, Melissa. 2017. 'Secondary Data Analysis: A Method of Which the Time Has Come'. *Qualitative and Quantitative Methods in Libraries* 3 (September 2014 Issue): 619–26.

Jonas, Kim, Rik Crutzen, Bart van den Borne, Ronel Sewpaul, and Priscilla Reddy. 2016. 'Teenage Pregnancy Rates and Associations with Other Health Risk Behaviours: A Three-Wave Cross-Sectional Study among South African School-Going Adolescents'. *Reproductive Health* 13 (50): 1–14. https://doi.org/10.1186/s12978-016-0170-8.

K. S., James, and Irudaya Rajan Sebastian Nil. 2004. 'Second National Family Health Survey: Emerging Issues'. *Weekly*, February, 647–51.

Karamancı, Kaan. 2009. 'Exploratory Data Analysis for Preemptive Quality Control'. Massachusetts Institute of Technology.

Karr, Alan F., Ashish P. Sanil, and David L. Banks. 2006. 'Data Quality: A Statistical Perspective'. *Statistical Methodology* 3 (2): 137–73. https://doi.org/10.1016/j.stamet.2005.08.005.

Kumar, Ashok, Tej Singh, Sriparna Basu, Sulekha Pandey, and V. Bhargava. 2007. 'Outcome of Teenage Pregnancy'. *The Indian Journal of Pediatrics* 74 (10): 927–31. https://doi.org/10.1007/s12098-007-0171-2.

Leplatte, Dayna, Katherine Lisa Rosenblum, Emily Stanton, Nicole Miller, and Maria Muzik. 2012. 'Mental Health in Primary Care for Adolescent Parents'. *Mental Health in Family Medicine* 9 (1): 39–45.

Mahavarkar, S. H., C. K. Madhu, and V. D. Mule. 2008. 'A Comparative Study of Teenage Pregnancy'. *Journal of Obstetrics and Gynaecology: The Journal of the Institute of Obstetrics and Gynaecology* 28 (6): 604–7. https://doi.org/10.1080/01443610802281831.

Mehta, Nikisha R., Shrinivas S. Darak, and Ritu S. Parchure. 2020. 'The Use of Emergency Contraceptive Pills in India: A Meta-Analysis'. *Indian Journal of Public Health* 64 (2): 178–85. https://doi.org/10.4103/ijph.IJPH_494_19.

Milo, Tova, and Amit Somech. 2020. 'Automating Exploratory Data Analysis via Machine Learning: An Overview'. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, 2617–22. Portland OR USA: ACM. https://doi.org/10.1145/3318464.3383126.

Modi, Armene. 2017. 'Giving Girls Wings to Fly Tools to Empower Adolescent Girls in Rural Communities in India'. *ECHIDNA GLOBAL SCHOLARS PROGRAM*, November.

Monteiro, Denise Leite Maia, Ida Peréa Monteiro, Márcia Sacramento Cunha Machado, Zenilda Vieira Bruno, Filomena Aste da Silveira, Marta Francis Benevides Rehme, Albertina Duarte Takiuti, and Nádia Cristina Pinheiro Rodrigues. 2021. 'Trends in Teenage Pregnancy in Brazil in the Last 20 Years (2000–2019)'. *Revista Da Associacao Medica Brasileira (1992)* 67 (5): 759–65. https://doi.org/10.1590/1806-9282.20210265.

Myatt, Glenn J. 2007. *Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining*. Hoboken, N.J: Wiley-Interscience.

Neal, Sarah, Zoë Matthews, Melanie Frost, Helga Fogstad, Alma V. Camacho, and Laura Laski. 2012. 'Childbearing in Adolescents Aged 12–15 Years in Low Resource Countries: A Neglected Issue. New Estimates from Demographic and Household Surveys in 42 Countries: Childbearing in Adolescents Aged 12–15 Years'. *Acta Obstetricia et Gynecologica Scandinavica* 91 (9): 1114–18. https://doi.org/10.1111/j.1600-0412.2012.01467.x.

Patra, Shraboni, and Rakesh Kumar Singh. 2013. 'Levels, Trends, Determinants and Consequences of Adolescent Pregnancy in India'. *IUSSP International Population Conference*, August, 1–17.

Paul, Pintu. 2020. 'Child Marriage Among Girls in India: Prevalence, Trends and Socio-Economic Correlates'. *Indian Journal of Human Development* 14 (2).

Pickhardt, Carl E. 2009. 'Rebel with a Cause: Rebellion in Adolescence'. 6 December 2009.

Raj, Anita, and Ulrike Boehmer. 2013. 'Girl Child Marriage and Its Association with National Rates of HIV, Maternal Health, and Infant Mortality across 97 Countries'. *Violence Against Women* 19 (4): 536–51. https://doi.org/10.1177/1077801213487747.

Raj, Anita, Mohan Ghule, Saritha Nair, Niranjan Saggurti, Donta Balaiah, and Jay G. Silverman. 2015. 'Age at Menarche, Education, and Child Marriage among Young Wives in Rural Maharashtra, India'. *International Journal of Gynaecology and Obstetrics: The Official Organ of the International Federation of Gynaecology and Obstetrics* 131 (1): 103–4. https://doi.org/10.1016/j.ijgo.2015.04.044.

Sukamolson, Suphat. 2007. 'Fundamentals of Quantitative Research'. *Maejo University*, January.

Suri, K. B. 1994. 'The Problem of Teenage Pregnancy: An Educational Imperative'. *Journal of Multicultural Social Work* 3 (3): 35–48. https://doi.org/10.1300/J285v03n03_03.

Suri, Shoba. 2020. 'There's a Need to End Teenage Pregnancies in India, It's Harming the National Economy: Teenage Pregnancy Is Almost Double in Rural Areas, 9.2%, as Compared to Urban, 5%, in India'. *Observer Research Foundation*, 14 January 2020, sec. Commentaries. https://www.orfonline.org/research/theres-a-need-to-end-teenage-pregnancies-in-india-its-harming-the-national-economy-60307/.

Ursachi, Nina. 2020. 'Particularities of the Evolution of Pregnancy and Birth in Primiparous Women with Extreme Ages'. In *MedEspera: The 8th International Medical Congress for Students and Young Authors: Abstract Book*. Chisnau, Moldova: Nicolae Testemitann State University of Medicine and Pharmacy. https://medespera.asr.md/wp-content/uploads/ABSTRACT-BOOK.pdf.

Velleman, P. F., and D. C. Hoaglin. 1983. *Applications, Basics, and Computing of Exploratory Data Analysis.* 1981: Boston, Duxbury Press.

Wahhaj, Zaki. 2014. *A Theory of Child Marriage*. School of Economics Discussion Papers 1520. Kent: University of Kent.

Wilson, Helen, and Annette Hunnigton. 2005. 'Deviant (M)Others: The Construction of Teenage Motherhood in Contemporary Discourse'. *Journal of Social Policy* 35 (1): 59–76.

Zer, Shiran, Tamar Wainstock, Asnat Walfisch, and Eyal Sheiner. 2019. 'Perinatal Outcomes and Long-Term Health in Offspring of Teenage Mothers'. *Journal of Pediatric and Adolescent Gynecology* 32 (6): 622–27. https://doi.org/10.1016/j.jpag.2019.08.003.