

# Assessing Infant Mortality Rate: Problems stemming from Household Living Conditions, Women's Education and Health\*

Trends in India from the 1998 Demographic Health Survey

Jacob Yoke Hong Si

27 April 2022

## Abstract

What areas can be improved in order to promote the well-being of women in India and hence, reduce the infant mortality rate? Utilizing the data from the 1998-1999 India National Family Health Survey provided by the Demographic and Health Survey (DHS) program, we look to depict the demographics of Indian women and infants in different states of India. We have found that the root causes of poor infant mortality rates stem from having poor living conditions that affect the likelihood of women to attain education and understand the importance of antenatal care and birth delivery assistance. We also explore other factors such as potentially inheritable traits (unhealthy body weight and anaemia disease) as well as an infant's diet. These factors are crucial in the development of an infant and the reduction of the infant mortality rate.

**Keywords:** demographics, health, living conditions, infant mortality rate, education, antenatal care, body weight, anaemia, diet, birth delivery assistance.

## 1 Introduction

The Demographic and Health Surveys (DHS) is a nationwide household survey that provides data for a myriad of evaluation indicators with respect to health, nutrition, population and living standards. A few examples of topics include education, mortality rate and health conditions. In addition to the topics mentioned, each survey also collects comprehensive socio-demographic data including the respondents' age, marital status, residential area, etc.

The DHS program was established in 1984 and has provided technical assistance to over 300 demographic and health surveys in more than 90 countries. Its main objective is to enhance and develop data collection and utilization by host countries for monitoring and evaluating demographic indicators in order to devise high quality policy development decisions.

In these regards, the DHS program has been majorly successful with the data collected often serving as crucial assistance to the government's policy making to help improve the well-being of the population. Furthermore, the program data is also accessible to health care providers, researchers and post-secondary institutions that have informed research with respect to the public health of the citizens. The program has thus been an invaluable source for quantitative and statistical methods to further understand the demographics of the population holistically. For these aforementioned reasons, the DHS program serves an important and fundamental role within a country's statistical system (Wikipedia 2022).

The National Family Health Survey (NFHS-2: DHS program in India) on the state findings in India assesses key demographics of women and infants including their education, healthcare and living conditions. These

---

\*Code and data are available at: [https://github.com/jacobyokehongsi/Assessing\\_Infant\\_Mortality\\_Rate\\_in\\_1998\\_India](https://github.com/jacobyokehongsi/Assessing_Infant_Mortality_Rate_in_1998_India)

factors are of utmost importance in the growth of a human being. Women that are highly educated tend to pay more attention to their own health leading to a happier and healthier lifestyle. Furthermore, these characteristics of a women would be passed down generations, affecting the health conditions of their infants.

The NFHS on women and infants' health aims to provide the International Institute for Population Sciences, which is the organization responsible for coordinating the health survey in India, information that includes demographics and trends. With the available data, it would assist the organization to help the Indian government design social policies that will best inform the Indian female and infant population on improving their living standards. Thus, the DHS is one of the key aspects of the foundation in informing Indian women to have a healthy and functional support system as well as ushering new generations of Indians into a prosperous world.

In this paper, we will explore the background and characteristics of Indian women and infants using the data from the 1998-1999 India NFHS-2. Specifically, I will address the following research questions:

- How do various living conditions such as electricity, water and sanitation facilities affect the ability of women to attend school?
- How does education affect women's healthcare decisions and their attention to antenatal care?
- What are the factors that affect infant mortality rate?

The rest of the paper explores the following: Section 2 explores the origin of the data, methods used to obtain the data and its strengths and weaknesses. Section 3 explores the results and the relationships between the variables. Lastly, Section 4 discusses the results as well as potential solutions and policies that could be implemented to improve infant mortality rate.

## 2 Data

### 2.1 Data Source and Methodology

The data was obtained from India's second National Family Health Survey (NFHS-2) on women's and infants' health. The survey was conducted in 1998-1999 by the International Institute for Population Sciences (IIPS) using the NFHS questionnaire where households would be interviewed. The questionnaire was devised during workshops held in Mumbai where experts in population and health fields would partake in creating the questions. The target population was a sample of more than 90,000 ever-married women age 15-49 as well as infants born in the three years preceding the survey. The list of villages are used as a sampling frame and first stratified through several variables. Next, the participants were selected using systematic sampling with equal probability from a household list in each village area. The overall individual response rate was 96% with a total of 89,199 eligible women that were part of the 91,196 interviewed households.

The following is a brief outline of the questionnaire content on the 1998 India NFHS:

- Background characteristics
- Reproductive behavior and intentions
- Quality of care
- Knowledge and use of contraception
- Sources of family planning
- Antenatal, delivery, and postpartum care
- Breastfeeding and health
- Reproductive health
- Status of Women
- Knowledge of AIDS

Table 1: Several Key Features (in percent)

State	Females Age 6-14 Attending School	Households with Electricity	Infant Mortality Rate
Delhi	90.8	97.7	46.8
Haryana	85.5	89.1	56.8
Himachal Pradesh	97.3	97.2	34.4
Jammu & Kashmir	77.5	90.1	65.0
Punjab	90.0	95.5	57.1
Rajasthan	63.2	64.4	80.4
Madhya Pradesh	70.8	68.1	86.1
Uttar Pradesh	69.4	36.6	86.7
Bihar	54.1	18.2	72.9
Orissa	75.1	33.8	81.0

## 2.2 Survey Frame and Sampling Method

As mentioned in the previous section, participants of this survey were chosen via stratified sampling and systematic sampling. The survey frame that was created is as follows. The list of villages were first stratified by geographic location, then continuous regions and lastly further stratified using selected variables such as village size, female literacy and so on. Each state in India was stratified individually using these variables with a limit of 6 stratas for small states, 12 stratas for medium size states and 15 stratas for large states. From the list of stratified villages, they were then sampled systematically with probability proportional to the 1991 Census population of the village. The villages that had fewer than five households were not included in the sampling frame. Lastly, the individuals to be interviewed were selected with equal probability from the household list in each village area using systematic sampling.

## 2.3 Key features

The raw data of the state findings included 70 variables which were obtained from the questionnaire responses. Key features part of the raw data included education standards, accessibility to electricity, drinking water and toilets, infant mortality rate, mothers receiving antenatal care, anaemia statuses (mothers and infants), weight statuses (mothers and infants) and the infants' diet.

We use R (R Core Team 2020) and R packages tidyverse (Wickham et al. 2019) and pdftools (Ooms 2022) to extract, preprocess and analyze the data. When preprocessing the data, We also use R package pointblank (Iannone and Vargas 2022) to put together tests for class and content to ensure that the dataset passes. The key features analyzed in this paper are shown in table 1 and 2, which are made using R package knitr (Xie 2022).

## 2.4 Strengths and Weaknesses

### 2.4.1 Strengths

The primary strengths of this survey was the myriad of variables describing participants and households, as well as the large sample size where more than 99 percent of India's female population was surveyed. Additionally, all Indian states except for one was surveyed, covering the majority of India. The field data collected by interviewing teams were inputted into microcomputers to produce field-check tables, ensuring that errors were prevented when eliciting information and filling out questionnaires. The large amount of variables allowed for thorough data exploration of the relationships between them while the ample amount of data allowed concrete conclusions to be drawn. Furthermore, the sampling methodologies was explained in high detail as well as a comprehensive fact sheet put together to display key data from the survey.

Table 2: Several Key Features (in percent) Continued.

State	Mothers $\geq 1$ antenatal check-up	Underweight infant	Women BMI $< 18.5\text{kg/m}$
Delhi	83.5	34.7	12.0
Haryana	58.1	34.6	25.9
Himachal Pradesh	86.8	43.6	29.7
Jammu & Kashmir	83.2	34.5	26.4
Punjab	74.0	28.7	16.9
Rajasthan	47.5	50.6	36.1
Madhya Pradesh	61.0	55.1	38.2
Uttar Pradesh	34.6	51.7	35.8
Bihar	36.3	54.4	39.3
Orissa	79.5	54.4	48.0

### 2.4.2 Weaknesses

One of the drawbacks of the survey was that survey findings were obtained in all states but one, that was Tripura, due to the delay in fieldwork based on a local problem. Additionally, data was unable to be obtained from areas with union territories. Nonsampling errors were also present where surveyors were unable to locate and interview the correct household and the misunderstanding of the questionnaire by the interviewer or respondent. In addition, the collection of anaemia data was hindered since some female participants refused to partake in the anaemia testing and/or have their infants tested as well. These aforementioned weaknesses result in missing data thus, the inference we will make could be less diverse.

## 3 Results

The figures are made using R (R Core Team 2020), R package tidyverse (Wickham et al. 2019) and Lucidchart (*Lucidchart* 2022). The variables explored provide numeric values that vary between the states of India.

Figure 1 represents a directed acyclic graph that will outline the interrelationships between variables affecting Infant Mortality Rate.

### 3.1 Household conditions that affect females attending school

Figure 2 depicts the distribution of females age 6-14 attending school (in percent) across the different states of India. Here, we observe that majority of females attending school, are from states including Kerala, Himachal Pradesh and Goa whereas majority of females not attending school are from Bihar, Rajasthan and Uttar Pradesh. In the states where the majority of females are not attending school, we can also infer that these states tend to have a larger population, especially Uttar Pradesh.

Figure 3 depicts the relationship between households with electricity and females age 6-14 attending school (in percent). Here, we observe that as the percent of households with electricity increases, the percent of females age 6-14 attending school increases. This implies that households that can afford electricity tend to be more well-off and are able to send their children to school.

Figure 4 depicts the relationship between households with drinking water piped or from hand pump and females age 6-14 attending school (in percent). Here, we observe that as the percent of households drinking water piped or from hand pump increases, the percent of females age 6-14 attending school decreases slightly which is odd. Therefore, we further analyze this strange trend by investigating the relationship between population and households with drinking water. Looking at Figure 5, as the percent of households with

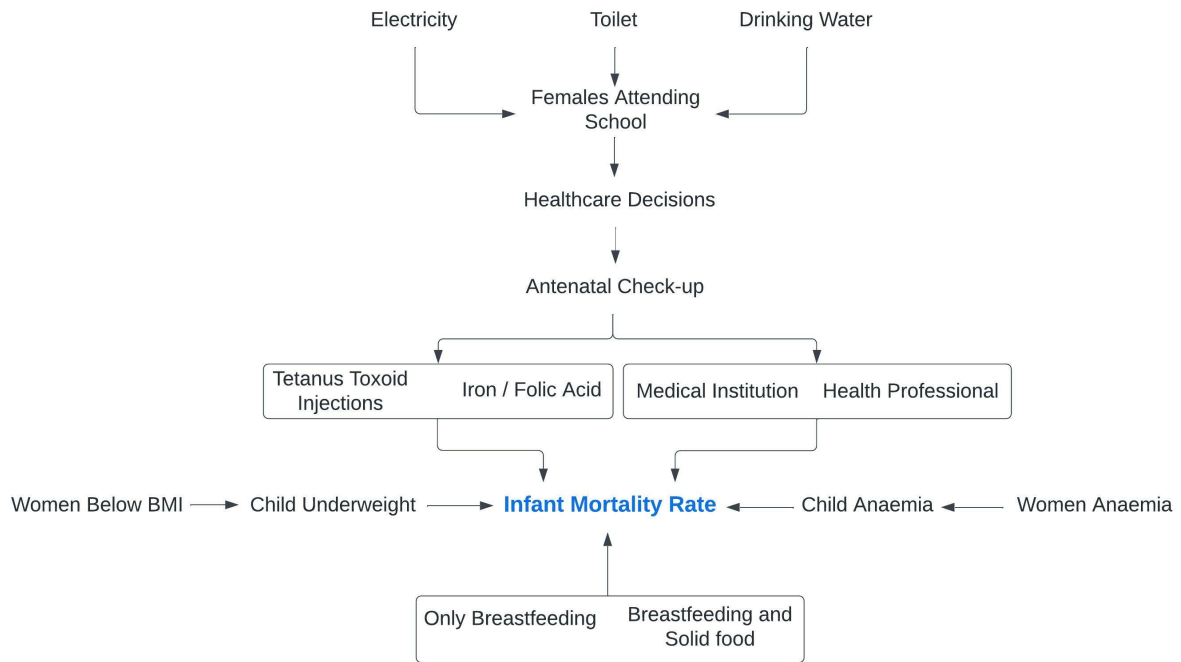


Figure 1: Interrelationships between variables affecting Infant Mortality Rate

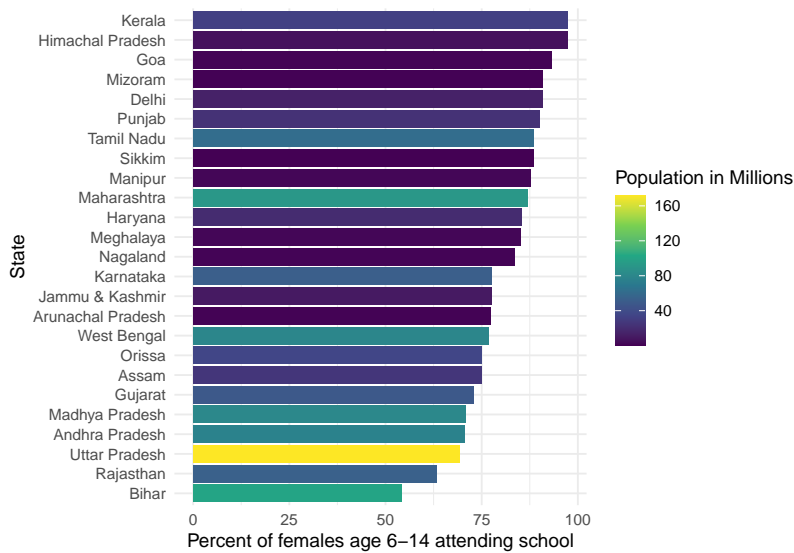


Figure 2: Distribution of the Percent of females age 6-14 attending school in various Indian States

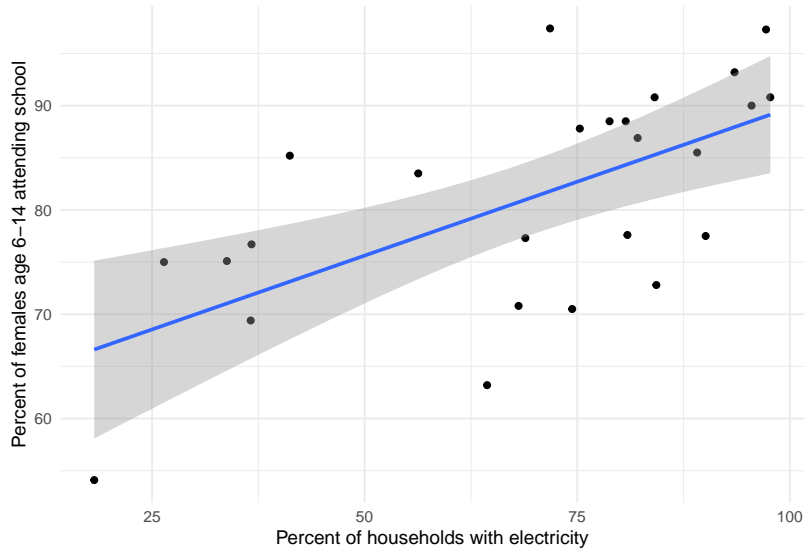


Figure 3: Relationship between households with electricity and females attending school (in percent)

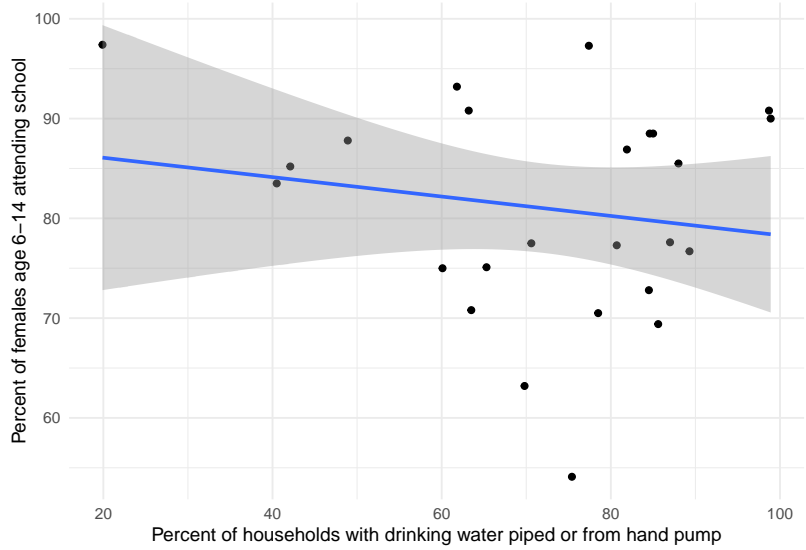


Figure 4: Relationship between households with drinking water piped or from hand pump and females attending school (in percent)

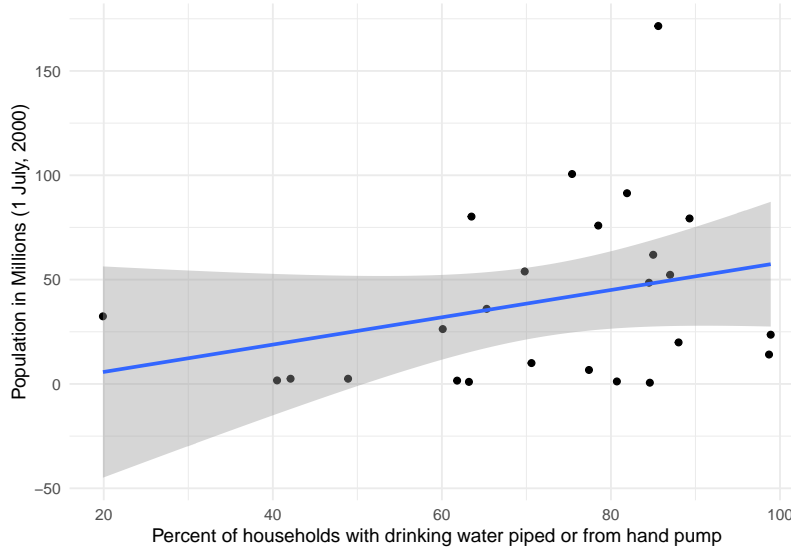


Figure 5: Relationship between households with drinking water piped or from hand pump and population in millions (in percent)

drinking water piped or from hand pump increases, the population increases. Thus, this aligns with the fact that in Figure 2, states with higher populations are also states where the majority of females are not attending school.

Lastly, figure 6 depicts the relationship between households with no toilet/latrine facility and females age 6-14 attending school (in percent). Here, we observe that as the percent of households with no toilet/latrine facility increases, the percent of females age 6-14 attending school decreases. This implies that households that have sanitation areas contribute to a better living environment leading to educated female individuals in the household.

### 3.2 Females attending school and Women involved in decisions regarding personal healthcare

Figure 7 depicts the relationship between females age 6-14 attending school and women involved in decisions about their own healthcare (in percent). Here, we observe that as the percent of females age 6-14 attending school increases, the percent of women involved in decisions about their own healthcare increases. This suggests that female individuals who are educated are more likely to take care of their own health and live a healthier lifestyle.

### 3.3 Women involved in decisions regarding personal healthcare and mothers receiving at least one antenatal check-up

Figure 8 depicts the relationship between women involved in their own healthcare decisions and mothers receiving at least one antenatal check-up for births in the three years preceding the survey (in percent). Here, we observe that as the percent of women involved in own healthcare decisions increases, the percent of mothers receiving at least one antenatal check-up for births increases. This suggests that female individuals who are informed of their personal health care are also wary of having antenatal check-up conducted.

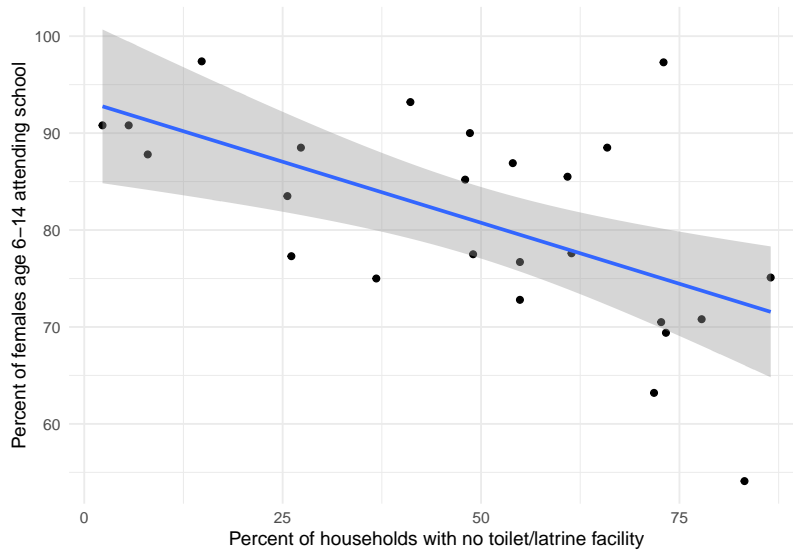


Figure 6: Relationship between households with no toilet/latrine facility and females attending school (in percent)

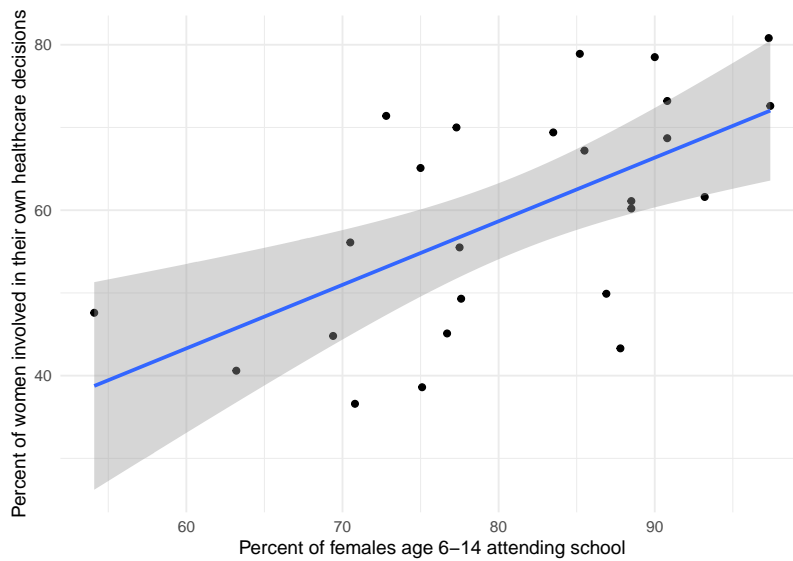


Figure 7: Relationship between females attending school and women involved in their own healthcare decisions (in percent)



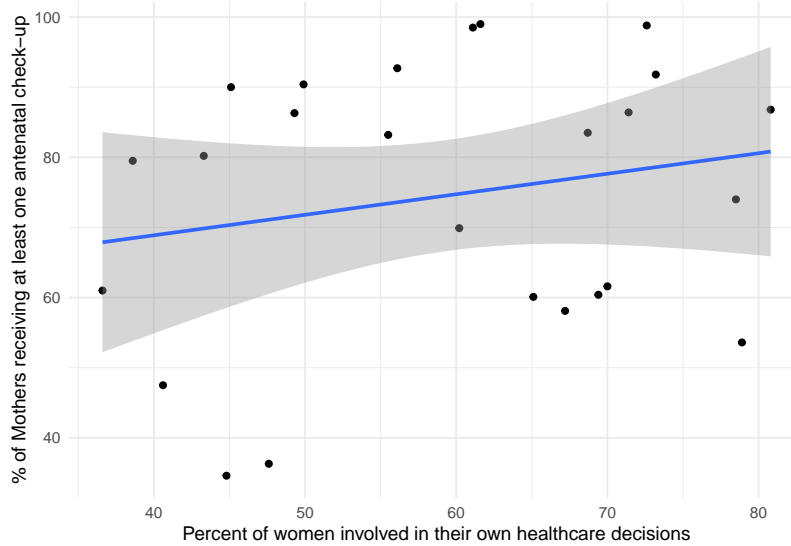


Figure 8: Relationship between women involved in their own healthcare decisions and mothers receiving at least one antenatal check-up (in percent)

### 3.4 Mothers that receive antenatal check-up, antenatal care and birth delivery assistance

Here, we explore how mothers that have antenatal check-up affect the antenatal care (supplements and injections) and birth delivery assistance they receive for births three years preceding the survey.

#### 3.4.1 Antenatal Care (Supplements and Injections)

Figures 9 and 10 depict the relationships between mothers receiving at least one antenatal check-up with mothers receiving iron and folic acid tablets or syrup as well as two or more tetanus toxoid injections (in percent) respectively. We observe that as the percent of mothers receiving at least one antenatal check-up increases, the percent of mothers receiving iron and folic acid tablets or syrup as well as two or more tetanus toxoid injections increases. This suggests that as mothers attend antenatal check-ups, doctors are likely to prescribe the mothers with iron and folic acid supplements as well as tetanus toxoid injections which boosts the health of the mother and her newborn infant.

#### 3.4.2 Birth Delivery Assistance

Figures 11 and 12 depict the relationships between mothers receiving at least one antenatal check-up with deliveries in medical institutions as well as deliveries assisted by a health professional (in percent) respectively. We observe that as the percent of mothers receiving at least one antenatal check-up increases, the percent of deliveries in medical institutions as well as deliveries assisted by a health professional increases. This suggests that when mothers go for their antenatal check-up, they likely have to approach a medical institution and have their deliveries conducted within the institution and/or by a health professional.

### 3.5 Factors that affect Infant Mortality Rate

In this subsection, we will explore the factors that will affect the infant mortality rate per 1,000 live births for the five years preceding the survey.

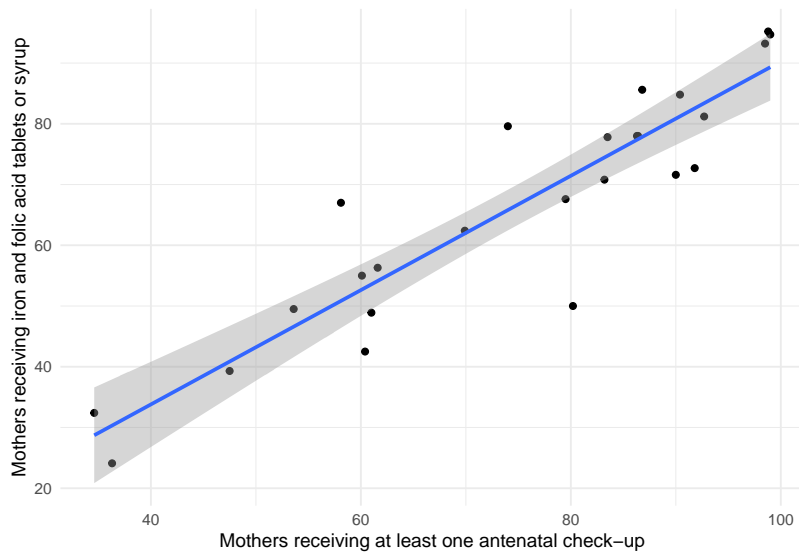


Figure 9: Relationship between mothers receiving at least one antenatal check-up and mothers receiving iron folic acid tablets or syrup (in percent)

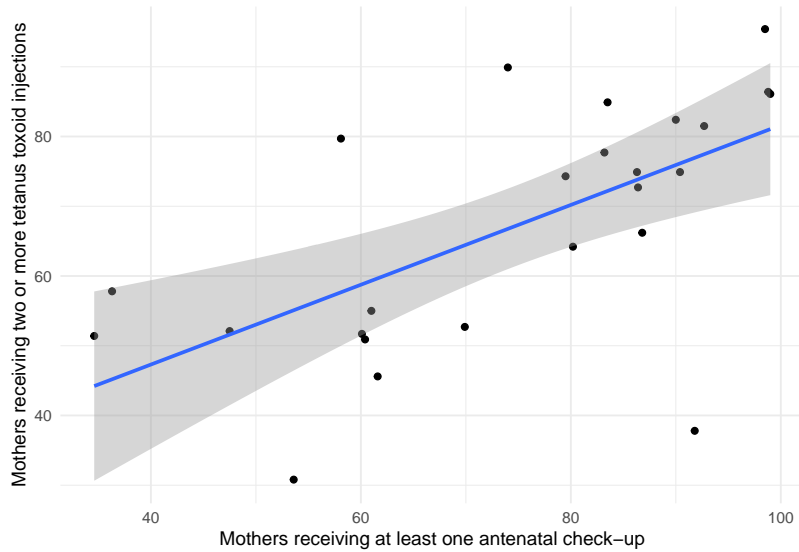


Figure 10: Relationship between mothers receiving at least one antenatal check-up and mothers receiving two or more tetanus toxoid injections (in percent)

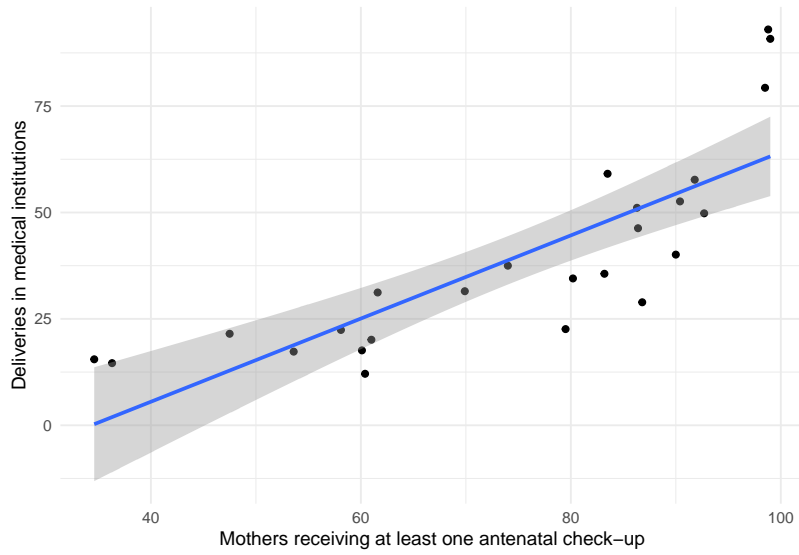


Figure 11: Relationship between mothers receiving at least one antenatal check-up and deliveries in medical institutions (in percent)

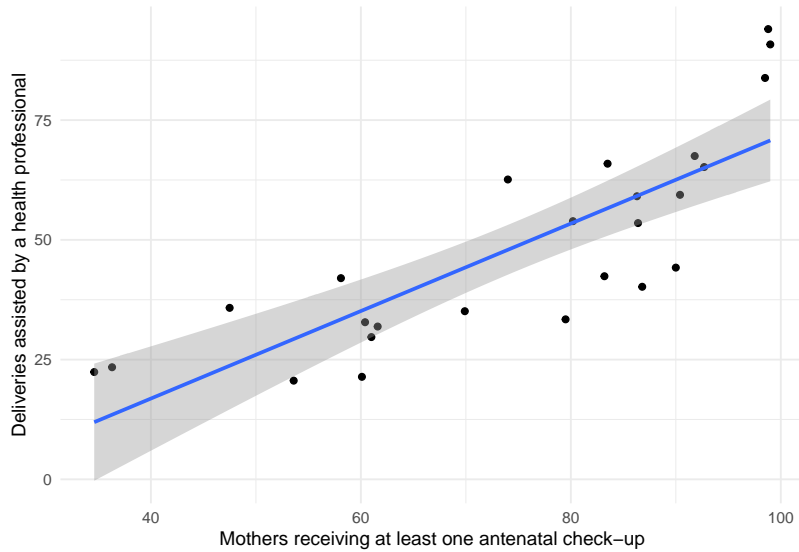


Figure 12: Relationship between mothers receiving at least one antenatal check-up and deliveries assisted by a health professional (in percent)

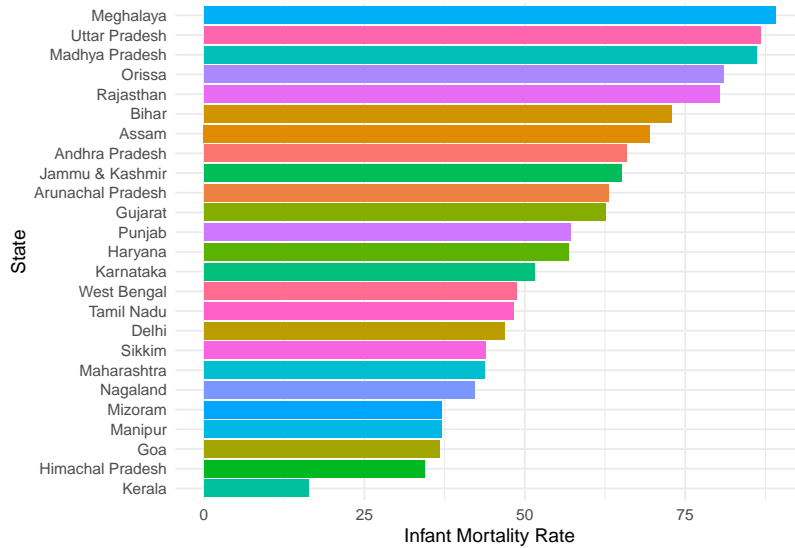


Figure 13: Distribution of Infant Mortality Rate in various Indian States

Figure 13 depicts the distribution of infant mortality rate across the different states of India. Here, we observe that states with the highest infant mortality rates include Meghalaya, Uttar Pradesh and Madhya Pradesh whereas states with the lowest infant mortality rates include Kerala, Himachal Pradesh and Goa.

### 3.5.1 Antenatal care and Delivery Assistance

Here, we explore how antenatal care such as supplements and injections as well as delivery assistance in the three years preceding the survey affect the infant mortality rate for births.

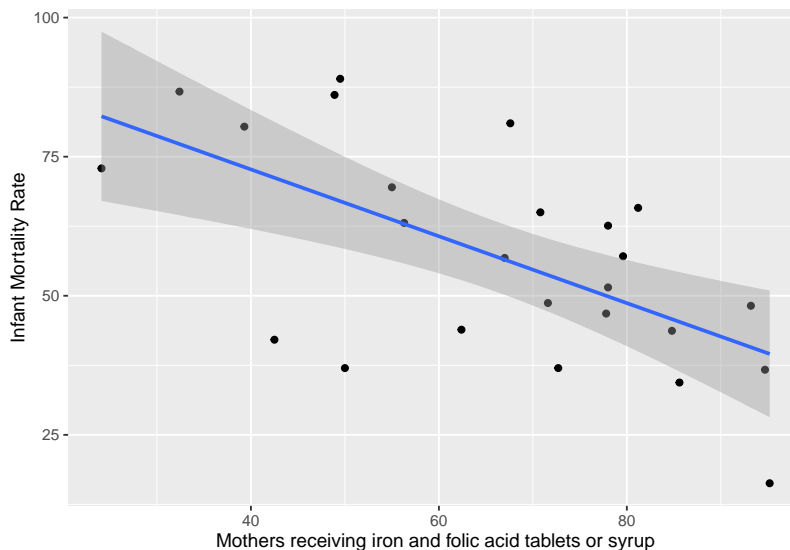


Figure 14: Relationship between mothers receiving iron and folic acid tablets or syrup and infant mortality rate (in percent)

Figures 14 and 15 depicts the relationship between mothers receiving iron and folic acid tablets or syrup, and two or more tetanus toxoid injections (in percent) with infant mortality rate respectively. Here, we observe

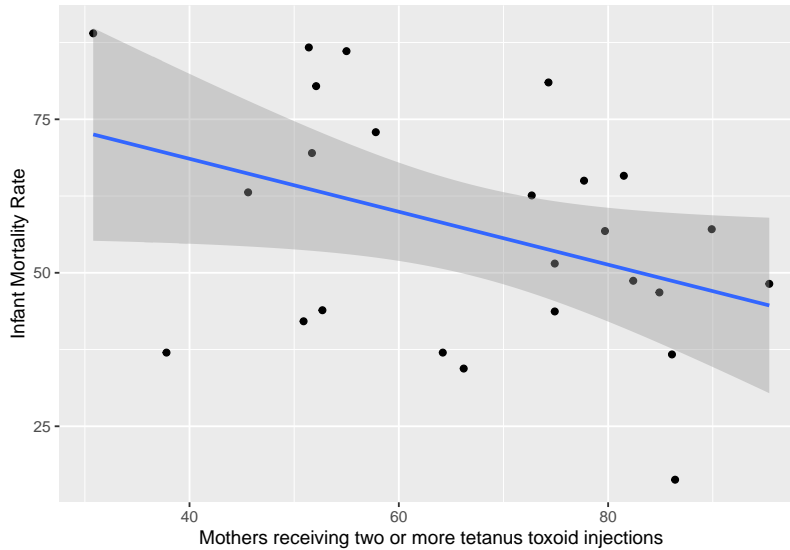


Figure 15: Relationship between mothers receiving two or more tetanus toxoid injections and infant mortality rate (in percent)

that as the percent of mothers receiving iron and folic acid tablets or syrup and two or more tetanus toxoid injections increases, the infant mortality rate decreases. This suggests that mothers who pay more attention to their own health and follow the doctor’s antenatal prescriptions will lead to a reduction in deaths of infants after they are born.

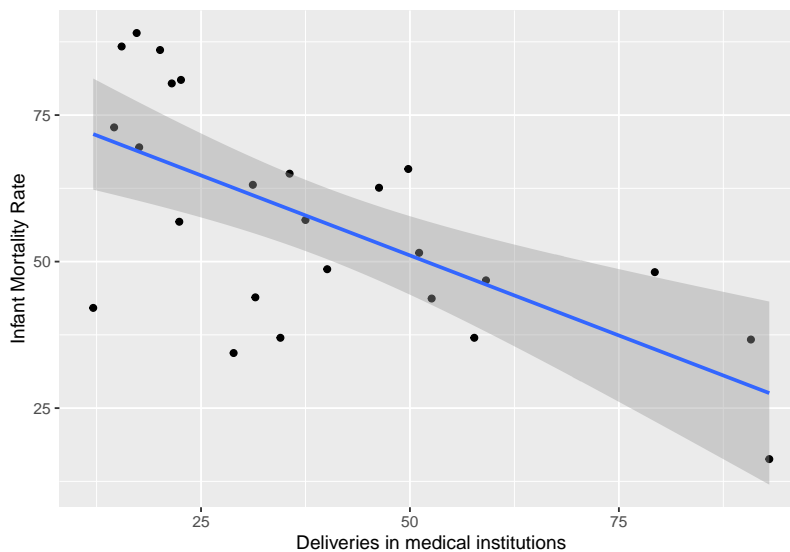


Figure 16: Relationship between deliveries in medical institutions and infant mortality rate (in percent)

Figures 16 and 17 depicts the relationship between deliveries in medical institutions and deliveries assisted by a health professional (in percent) with the infant mortality rate respectively. Here, we observe that as the percent of deliveries in medical institutions and deliveries assisted by a health professional increases, the infant mortality rate decreases. This implies that mothers who receive professional medical services during birth deliveries will also have a smooth child birth.

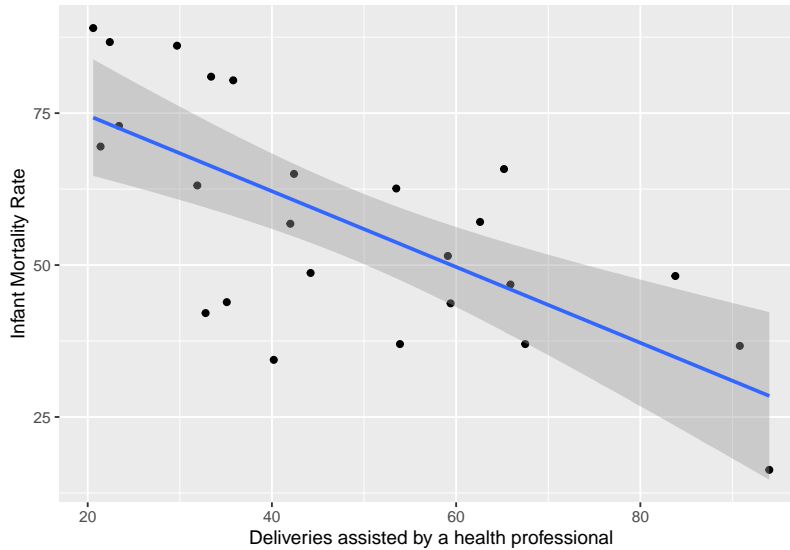


Figure 17: Relationship between deliveries assisted by a health professional and infant mortality rate(in percent)

### 3.5.2 Infants that receive exclusively breastmilk and both breastmilk and solid food

Figures 18 and 19 portrays the relationship between infant age 0-3 months exclusively breastfed, as well as infant age 6-9 months receiving breast milk and solid/mushy food (in percent) with the infant mortality rate respectively. Here, we observe that the percent of infant age 0-3 months that are exclusively breastfed does not have a significant impact on infant mortality rate. On the other hand, as the percent of infant age 6-9 months receiving breast milk and solid/mushy food increases, the infant mortality rate decreases. This indicates that regardless of breast milk or formula milk fed to infants at age 0-3 months, infant mortality rate is not affected. However, as infants grow to the age of 6-9 months, they should be provided solid/mushy food on top of breast milk to support their growth and reduce the risk of mortality.

### 3.5.3 Relationship between a women’s body mass index (BMI), infants’ weight and mortality rate

Figure 20 depicts the relationship between women with BMI below 18.5 kg/m and infants under age three years that are underweight (in percent). Women that have BMI below 18.5kg/m are considered to be underweight (C. Clinic 2022). Here, we observe that as the percent of women with BMI below 18.5 kg/m increases, the percent of infants under age three years that are underweight increases. This implies that mothers that are skinny and malnourished would also lead to offsprings that are underweight.

Figure 21 depicts the relationship between infants under age three years that are underweight (in percent) and infant mortality rate. Here, we observe that as the percent of infants under age three that are underweight increases, the infant mortality rate increases. From this we can infer that infants who are not underweight tend to be in good health which reduces infant mortality rate.

### 3.5.4 Relationship between women with anaemia, infants with anaemia and infant mortality rate

Figure 22 depicts the relationship between women age 15-49 with any anaemia and infants age 6-35 months with any anaemia (in percent). Here, we observe that as the percent of women age 15-49 with any anaemia increases, the percent of infants age 6-35 months with any anaemia increases. This is coherent with the fact

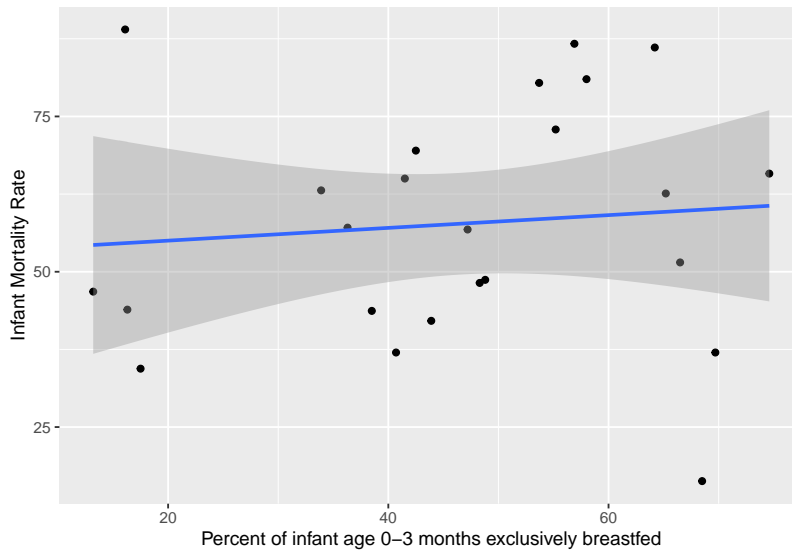


Figure 18: Relationship between infant age 0-3 months exclusively breastfed and infant mortality rate (in percent)

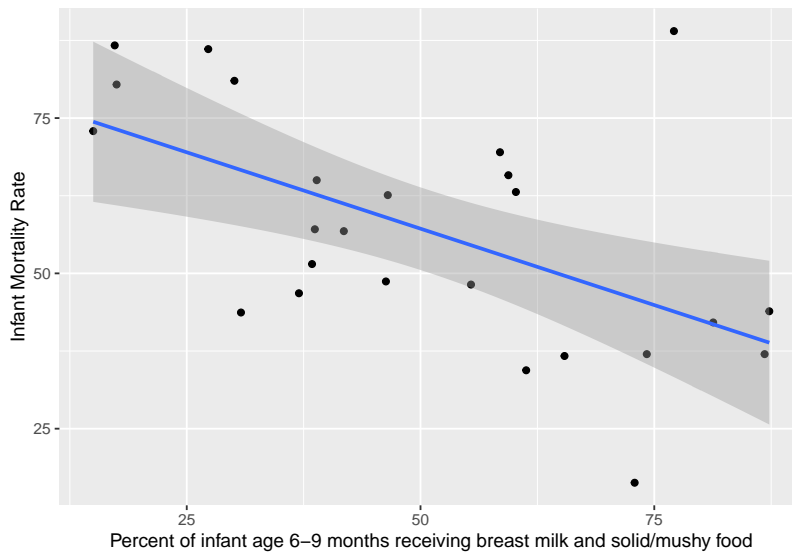


Figure 19: Relationship between infant age 6-9 months receiving breast milk and solid/mushy food and infant mortality rate (in percent)

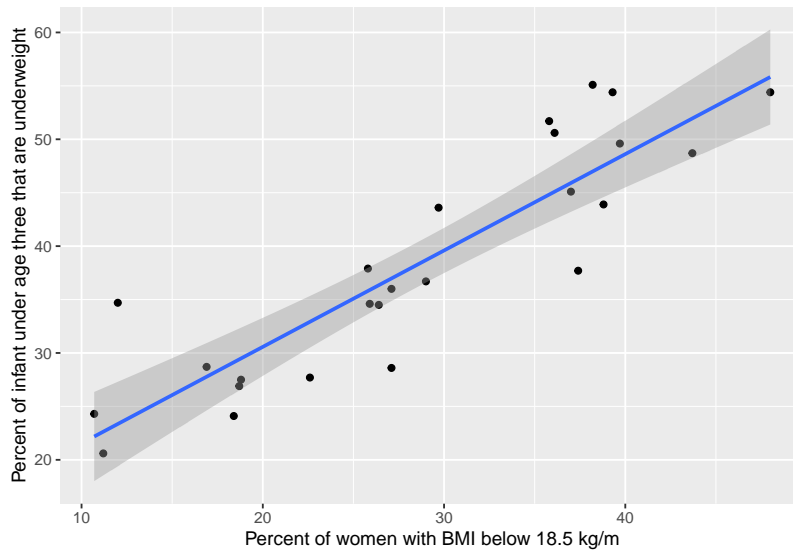


Figure 20: Relationship between women with BMI below 18.5 kg/m and infant under age three that are underweight (in percent)

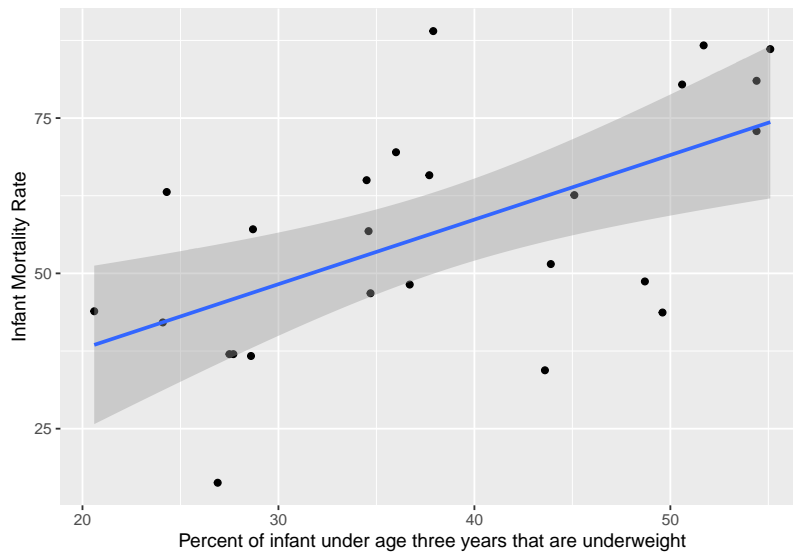


Figure 21: Relationship between infant under age three years that are underweight and infant mortality rate (in percent)



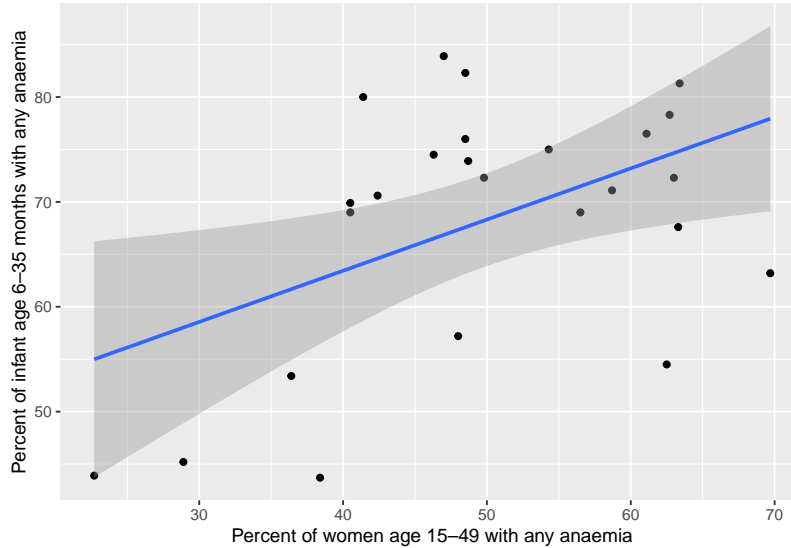


Figure 22: Relationship between women age 15–49 with any anaemia and infant age 6–35 months with any anaemia (in percent)

that since anaemia such as hemolytic anaemia can be inherited, mothers could pass on the anaemia to their offsprings.

Figure 23 depicts the infants age 6–35 months with any anaemia (in percent) and infant mortality rate. Here, we observe that as the percent of infants age 6–35 months with any anaemia increases, the infant mortality rate increases. From this we can infer that infants who have the anaemia disease would have symptoms that lead to critical health conditions and death.

## 4 Discussion

The ultimate goal of the paper is to assist policy makers in implementing appropriate and efficient socio-economic policies to help promote the steady growth of infants and reduce infant mortality rate in India. We will first discuss the root causes which are the household living environment, women’s education and their healthcare decisions which lead them to attend antenatal check-ups. Then, we examine the numerous factors that are affected by conducting antenatal check-ups which include antenatal supplements and injections, as well as birth delivery assistance. These are some of the medical factors that directly impact the infant mortality rate. Additionally, we also look at the importance of infants’ meals for their growth. Lastly, we examine potentially inheritable traits such as anaemia and unhealthy body weight that could be passed on from mother to infant and how they impact the infant mortality rate. The full data can be found in Appendix B.

### 4.1 Household Living Environment, Women’s Education and their Healthcare Decisions

The household living environment of individuals including access to electricity, clean water and sanitation facilities is crucial in providing a comfortable and productive living space. As observed previously, households with better living conditions will also lead to women who are able to attend school and be well educated. In our figures, we could infer that households in India with poorer living conditions also have fewer women attending school. Thus, emphasis should be put in developing the living conditions within households in India, especially states with poorer living conditions such as Bihar. In Bihar, only 18.2% of the households

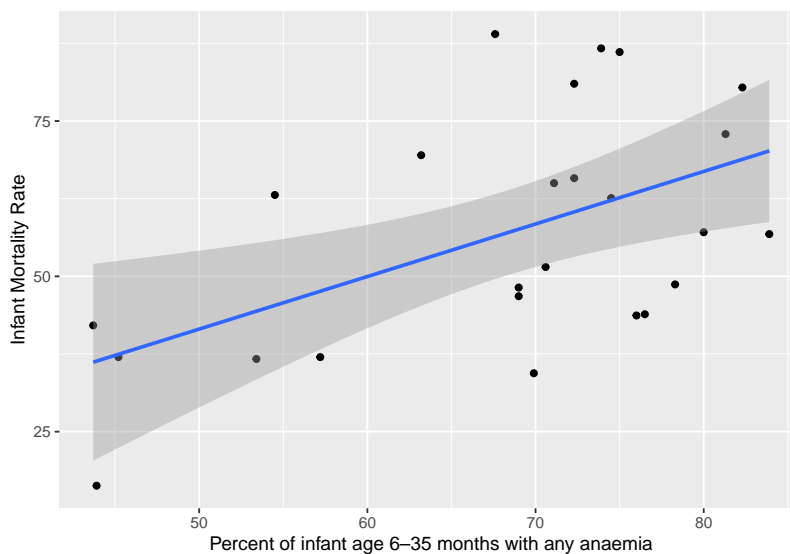


Figure 23: Relationship between infant age 6–35 months with any anaemia and infant mortality rate (in percent)

have access to electricity and 54.1% of females attending school. With such poor basic necessities and living standards, there is no doubt that the women are unable to attend school. By developing the living conditions, it will in turn improve the well-being of the population and encourage women to attend school. Another proposal could also be to provide direct subsidies for needy students in hopes to attain a well educated population.

Having an educated female population, we observe that women are more involved in making decisions regarding their own health care. By paying a higher attention to their personal health, women would devote more time and effort in ensuring a healthy lifestyle. In more developed states such as Himachal Pradesh, 97.3% of females age 6-14 are attending school. As these females become mothers, their personal knowledge on healthcare would be carried forward to having antenatal care which is shown as 80.8% of women is involved in decisions about their own health care in Himachal Pradesh.

## 4.2 Antenatal check-ups, antenatal care and delivery assistance to promote mothers and infants health

Attending antenatal check-ups is key for protecting the health of women and their unborn infants. During antenatal check-ups, it is likely that doctors prescribe supplements and procedures needed in order for a smooth child birth. This includes iron and folic acid supplements that supports the development of the placenta and fetus as well as preventing neural tube defects respectively (M. Clinic 2022b). In addition, tetanus toxoid injections help prevent neonatal tetanus (Disease Control and Prevention 2022). Furthermore, partaking in check-ups will also increase the likelihood of having child deliveries in medical institutions and/or by a health professional since the check-ups are conducted in a similar fashion. Child deliveries in this manner are safe and sanitary, ensuring that complications are minimized when the baby is delivered. In our data, we observe that in states where mothers often receive at least one antenatal check-up such as Goa (99.0% of mothers receiving at least one antenatal check-up), they also pay more attention to antenatal care where mothers receive two or more tetanus toxoid injections (86.1%), iron and folic acid tablets or syrups (94.7%) as well as having their infant delivered in medical institutions (90.8%) and/or by a health professional (90.8%).

### 4.3 Meals of infants and their effects on infant mortality rate

The diet of an infant is one of the most important fundamentals in a child's health. An infant has to be well-fed in order to receive the nutrients and vitamins required for a healthy body. To recap, we observe that for infants age 0-3 months, being exclusively breastfed does not have a significant impact on the infant mortality rate. On the other hand, for infants age 6-9 months, it is important that they receive solid/mushy food to supplement the breast milk they are consuming. States such as Rajasthan with low percent of infants age 6-9 months receiving breast milk and solid/mushy food (17.5%) also have a high infant mortality rate (80.4%). As an infant grows to more than 6 months old, the nutrients supplied by the breast milk are simply insufficient to develop a healthy immune system. Thus, it is vital that infants are supplied with ample amount of nutritious solid foods on top of breast/formula milk to promote a healthy growth.

### 4.4 Unhealthy body weight and anaemia of mothers affecting the infant mortality rate

Lastly, we will explore how potentially inheritable traits that mothers possess, such as unhealthy body weight and anaemia disease, are inherited by their infants and would be detrimental to their health. In the state Orissa, the percent of women who are underweight is roughly 48.0% which corresponds to 54.4% of infants being underweight. Hence, women that are underweight would increase the likelihood of their offsprings being underweight and malnourished.

On the other hand, infants could also inherit anaemia diseases from their mother. Anaemia is a condition in which the body does not have sufficient healthy red blood cells to carry adequate oxygen around the body. Symptoms of anaemia include tiredness, shortness of breath, irregular heartbeats and more (M. Clinic 2022a). As infants are born with this disease, they are fatigued and sick leading to their poor health and death. In the northeast state of Assam, the percent of women age 15-49 with any anaemia is approximately 69.7% which also corresponds to the high anaemia rate of 63.2% in infants age 6-35 months. The percentage of infants with anaemia is relatively close to the percentage of women with anaemia. From this, we can infer that infants are most likely to have inherited this disease from their mother.

Therefore, without a strong functional health system, the child would be weak and frail resulting in a higher rate of infant deaths. To solve these issues, it is important that mothers attend antenatal check-ups where doctors could discover these health conditions and prescribe them with the appropriate treatments.

### 4.5 Weaknesses and next steps

There are a few weaknesses in the paper that I will evaluate. Firstly, the data was not collected by myself and hence, we only know as much about the methodology behind it as we've been provided. Despite this being a potentially major issue, the IIPS did more than their due diligence by providing supple information in explaining every aspects of the survey including the survey methodology and sampling techniques as well as documenting detailed variable names. However, there could be potential biases that were unaccounted for by the IIPS that included the sampling or surveying process, and ethical considerations that impacted the data collection in which we have no information about them.

Another possible weakness in our paper is the uncertainty in which the factors are individually correlated to the infant mortality rate. For instance, infants that do not receive solid/mushy food is not the sole reason for the deaths of infants. Instead, it is the amalgamation of all the different factors analyzed that leads to their deaths. As a next step for this issue, we could conduct statistical tests such as t-tests to determine if individual variables are statistically significant when regressed on the infant mortality rate. Apart from the aforementioned variables, there are factors such as the household income, number of schools in a state etc. that could affect a female's ability of attending school. Hence, by collecting more data on more variables, we would be able to make even more concrete conclusions with regards to the factors that contribute to infant mortality.

To summarize, there are numerous factors that could affect infant mortality rate which include the unhealthy weight of an infant, whether an infant has anaemia, diet of an infant and the prenatal health of a mother. The health of a mother is affected by her education background and the living conditions she grew up in. Hence, it is crucial for the government to develop policies that solves the fundamental problems by first improving the living standards of households, starting with states such as Bihar with the poorest living conditions. The government could also build more schools as well as subsidize education expenses for poor families which encourages them to send their children to school. Having good living conditions and educated women would improve the health of mothers and her offsprings, resulting in a healthier future generation of India.

# Appendix

## A Data Sheet (obtained from (Gebru et al. 2018))

1. *For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.*
  - The dataset was created to enable analysis of policy makers and program administrators in India responsible for improving health and family welfare programs in the states.
2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
  - The International Institute for Population Sciences (IIPS) was designated as the agency for creating the dataset on behalf of the Ministry of Health and Family Welfare, Government of India, New Delhi.
3. *Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.*
  - It was funded by the Government of India.
4. *Any other comments?*
  - This second National Family Health Survey (NFHS-2) is created to further bolster the existing database from the first NFHS and further update the required health programs that should be implemented in the country.

### Composition

1. *What do the instances that comprise the dataset represent (for example, documents, photos, people, countries)? Are there multiple types of instances (for example, movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.*
  - The instances that comprise the dataset represent the states of India. This includes North, Central, East, Northeast, West and South states of India.
2. *How many instances are there in total (of each type, if appropriate)?*
  - North: 6
  - Central: 2
  - East: 3
  - Northeast: 7
  - West: 3
  - South: 4
  - Total:  $25 + 1$  (India) = 26
3. *Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (for example, geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (for example, to cover a more diverse range of instances, because instances were withheld or unavailable).*
  - The dataset contain all possible instances (states of India).
4. *What data does each instance consist of? “Raw” data (for example, unprocessed text or images) or features? In either case, please provide a description.*
  - Each instance consists of 1 categorical variable representing the state and 18 continuous variables representing variables such as school attendance rate, percent of households with electricity, infant mortality rate etc.

5. *Is there a label or target associated with each instance? If so, please provide a description.*
  - There is no target associated with each instance.
6. *Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (for example, because it was unavailable). This does not include intentionally removed information, but might include, for example, redacted text.*
  - Under the instance Goa and variable percent of infant age 0-3 months exclusively breastfed, information is missing because the data collected is based on fewer than 25 unweighted cases.
7. *Are relationships between individual instances made explicit (for example, users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.*
  - There are no relationships between individual instances.
8. *Are there recommended data splits (for example, training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.*
  - There are no recommended data splits
9. *Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.*
  - There are no errors, sources of noise, or redundancies in the dataset.
10. *Is the dataset self-contained, or does it link to or otherwise rely on external resources (for example, websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (that is, including the external resources as they existed at the time the dataset was created); c) are there any restrictions (for example, licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.*
  - The dataset is self-contained.
11. *Does the dataset contain data that might be considered confidential (for example, data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.*
  - The dataset does not contain data that might be considered confidential.
12. *Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.*
  - The dataset does not contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety.
13. *Does the dataset identify any sub-populations (for example, by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset.*
  - The dataset identifies women and infants below age three.
14. *Is it possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset? If so, please describe how.*
  - It is not possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset.
15. *Does the dataset contain data that might be considered sensitive in any way (for example, data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? If so, please provide a description.*

- The dataset does not contain data that might be considered sensitive in any way.

16. *Any other comments?*

- N/A

## Collection process

1. *How was the data associated with each instance acquired? Was the data directly observable (for example, raw text, movie ratings), reported by subjects (for example, survey responses), or indirectly inferred/derived from other data (for example, part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.*
  - The data associated with each instance is acquired through interviews conducted in 25 Indian states.
2. *What mechanisms or procedures were used to collect the data (for example, hardware apparatuses or sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?*
  - Questionnaires are first filled out using manual human curation then entered into microcomputers to produce field-check tables which ensures that the data is free from errors.
3. *If the dataset is a sample from a larger set, what was the sampling strategy (for example, deterministic, probabilistic with specific sampling probabilities)?*
  - The dataset is not a sample from a larger set.
4. *Who was involved in the data collection process (for example, students, crowdworkers, contractors) and how were they compensated (for example, how much were crowdworkers paid)?*
  - The fieldwork in each state was carried out by a number of interviewing teams, each team consisting of one field supervisor, one female field editor, four female interviewers, and one health investigator. Compensation data is not provided.
5. *Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (for example, recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.*
  - The data collection was carried out in two phases, starting in November 1998 and March 1999. The timeframe matches the creation timeframe of the data associated with the instances.
6. *Were any ethical review processes conducted (for example, by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.*
  - Ethical review processes were not conducted
7. *Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?*
  - The data is obtained from the Demographic and Health Survey program website. [https://dhsprogram.com/publications/publication-frind2-dhs-final-reports.cfm?cssearch=467922\\_1](https://dhsprogram.com/publications/publication-frind2-dhs-final-reports.cfm?cssearch=467922_1)
8. *Were the individuals in question notified about the data collection? If so, please describe (or show with screenshots or other information) how notice was provided, and provide a link or other access point to, or otherwise reproduce, the exact language of the notification itself.*
  - The individuals in question were notified as they voluntarily provide data to the interviewers. The exact language of the notification is not available.

9. *Did the individuals in question consent to the collection and use of their data? If so, please describe (or show with screenshots or other information) how consent was requested and provided, and provide a link or other access point to, or otherwise reproduce, the exact language to which the individuals consented.*
  - The individuals in question consent to the collection and use of their data. For example, the health investigator read a detailed informed consent statement to the respondent, informing her about anaemia, describing the procedure to be followed for the test, and emphasizing the voluntary nature of the test. The exact language is as follows. “May I ask you now to give your consent to have the test(s) done. If you decide not to have the test(s), it is your right, and we will respect your decision. Now please tell me whether you agree to have the test(s)(and allow me to test your child).”
10. *If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? If so, please provide a description, as well as a link or other access point to the mechanism (if appropriate).*
  - Yes. In the questionnaire, they are allowed to revoke their consent.
11. *Has an analysis of the potential impact of the dataset and its use on data subjects (for example, a data protection impact analysis) been conducted? If so, please provide a description of this analysis, including the outcomes, as well as a link or other access point to any supporting documentation.*
  - Analysis of the potential impact of the dataset and its use on data subjects has not been conducted.
12. *Any other comments?*
  - N/A

### Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.*
  - Preprocessing/cleaning/labeling of the data was done. The original format of the file is in pdf and the data was extracted using R package pdftools and R in order for data analysis to be conducted.
2. *Was the “raw” data saved in addition to the preprocessed/cleaned/labeled data (for example, to support unanticipated future uses)? If so, please provide a link or other access point to the “raw” data.*
  - The raw data is saved in inputs/data/raw\_data1.csv, inputs/data/raw\_data2.csv and inputs/data/raw\_data3.csv
3. *Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.*
  - R software is available at <https://www.r-project.org/>
4. *Any other comments?*
  - N/A

### Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
  - The dataset has not been used for other tasks yet.
2. *Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.*



- [https://github.com/jacobyokehongsi/Assessing\\_Infant\\_Mortality\\_Rate\\_in\\_1998\\_India](https://github.com/jacobyokehongsi/Assessing_Infant_Mortality_Rate_in_1998_India)
3. *What (other) tasks could the dataset be used for?*
    - Tasks could include analyzing the demographics of women and infants in India during 1998-1999.
  4. *Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (for example, stereotyping, quality of service issues) or other risks or harms (for example, legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?*
    - The preprocessing is curated just for this specific dataset and cannot be applied to other datasets.
  5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*
    - N/A
  6. *Any other comments?*
    - N/A

## Distribution

1. *Will the dataset be distributed to third parties outside of the entity (for example, company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.*
  - The dataset will not be distributed to third parties outside of the entity on behalf of which the dataset was created.
2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*
  - The dataset will be distributed on Github.
3. *When will the dataset be distributed?*
  - The dataset will be distributed during April 2022.
4. *Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/ or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.*
  - The dataset will be distributed under the MIT license.
5. *Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.*
  - No third parties have imposed IP-based or other restrictions on the data associated with the instances.
6. *Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.*
  - No export controls or other regulatory restrictions apply to the dataset or to individual instances.
7. *Any other comments?*
  - N/A

## Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
  - Jacob Yoke Hong Si
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*
  - Through email at [jacobyh.si@mail.utoronto.ca](mailto:jacobyh.si@mail.utoronto.ca).
3. *Is there an erratum? If so, please provide a link or other access point.*
  - There is no erratum.
4. *Will the dataset be updated (for example, to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (for example, mailing list, GitHub)?*
  - There is no plan to update the dataset. Any updates will be communicated to dataset consumers via Github.
5. *If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (for example, were the individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.*
  - There are no applicable limits on the retention of the data associated with the instances.
6. *Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.*
  - Older versions will not be hosted. Its obsolescence is communicated via commit history on Github.
7. *If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.*
  - The repository on Github can be cloned to extend/augment/build on/contribute to the dataset. Contributions will not be validated/verified since I am not responsible for any other dataset extensions.
8. *Any other comments?*
  - N/A

## B Full data tables

### NFHS-2 FACT SHEET - STATES

State	Population, 1 July, 2000 (in millions) <sup>1</sup>	Percent of females illiterate (age 6+)	Percent of females age 6–14 attending school	Percent of ever-married women age 15–49 who are:		Percent of households				Percent of women involved in decisions about own health care	Percent of women age 20–24 married by exact age 18
				Urban	Not regularly exposed to any media	With electricity	With drinking water piped or from hand pump	With no toilet/latrine facility	Using adequately iodized salt <sup>2</sup>		
<b>India<sup>#</sup></b>	1002.1	48.6	73.7	26.2	40.3	60.1	77.9	64.0	49.3	51.6	50.0
<b>North</b>											
Delhi	14.1	21.7	90.8	92.1	7.3	97.7	98.7	5.6	89.2	68.7	19.8
Haryana	19.9	42.7	85.5	28.8	33.1	89.1	88.0	60.9	71.0	67.2	41.5
Himachal Pradesh	6.7	31.3	97.3	9.1	16.3	97.2	77.4	73.0	90.5	80.8	10.7
Jammu & Kashmir	10.0	55.3	77.5	21.5	25.6	90.1	70.6	49.0	52.9	55.5	22.1
Punjab	23.6	35.1	90.0	30.8	18.0	95.5	98.9	48.6	75.3	78.5	11.6
Rajasthan	53.9	62.9	63.2	24.2	63.1	64.4	69.8	71.8	46.3	40.6	68.3
<b>Central</b>											
Madhya Pradesh	80.2	55.5	70.8	25.3	45.2	68.1	63.5	77.8	56.7	36.6	64.7
Uttar Pradesh	171.5	57.3	69.4	20.0	54.7	36.6	85.6	73.3	48.8	44.8	62.4
<b>East</b>											
Bihar	100.6	65.2	54.1	10.2	72.7	18.2	75.4	83.2	46.9	47.6	71.0
Orissa	36.0	48.7	75.1	11.0	55.7	33.8	65.3	86.5	35.0	38.6	37.6
West Bengal	79.3	42.6	76.7	23.8	38.6	36.7	89.3	54.9	61.7	45.1	45.9
<b>Northeast</b>											
Arunachal Pradesh	1.2	43.0	77.3	15.9	36.7	68.9	80.7	26.1	84.1	70.0	27.6
Assam	26.3	40.9	75.0	8.5	47.4	26.4	60.1	36.8	79.6	65.1	40.7
Manipur	2.5	41.3	87.8	33.7	16.2	75.3	48.9	8.0	87.9	43.3	9.9
Meghalaya	2.5	33.2	85.2	20.0	37.3	41.2	42.1	48.0	63.0	78.9	25.5
Mizoram	1.0	10.6	90.8	52.9	16.9	84.1	63.2	2.3	91.3	73.2	11.6
Nagaland	1.7	31.7	83.5	20.3	35.7	56.3	40.5	25.6	67.2	69.4	22.9
Sikkim	0.6	35.6	88.5	14.2	21.5	80.7	84.6	27.3	79.1	60.2	22.3
<b>West</b>											
Goa	1.6	25.2	93.2	41.6	11.6	93.5	61.8	41.1	41.9	61.6	10.1
Gujarat	48.5	46.4	72.8	42.5	33.8	84.3	84.5	54.9	56.1	71.4	40.7
Maharashtra	91.4	38.6	86.9	41.3	29.6	82.1	81.9	54.0	60.1	49.9	47.7
<b>South</b>											
Andhra Pradesh	75.9	54.0	70.5	24.9	23.7	74.4	78.5	72.7	27.4	56.1	64.3
Karnataka	52.3	44.5	77.6	34.8	21.4	80.9	87.0	61.4	43.4	49.3	46.3
Kerala	32.4	14.9	97.4	23.1	11.5	71.8	19.9	14.8	39.3	72.6	17.0
Tamil Nadu	61.9	41.7	88.5	34.6	20.3	78.8	85.0	65.9	21.2	61.1	24.9

<sup>1</sup>Registrar General of India

<sup>2</sup>Cooking salt that has an iodine content of at least 15 parts per million (ppm)

<sup>#</sup>Excludes Tripura

Figure 24: Data Tables

**NFHS-2 FACT SHEET - STATES (Contd.)**

State	For births in the three years preceding the survey <sup>10</sup> , percent of:							Percent of children		
	Infant mortality rate <sup>9</sup>	Under-five mortality rate <sup>9</sup>	Mothers receiving at least one antenatal check-up	Mothers receiving two or more tetanus toxoid injections	Mothers receiving iron and folic acid tablets or syrup	Deliveries in medical institutions	Deliveries assisted by a health professional	Age 12-23 months who have received all vaccinations <sup>11</sup>	Age 12-35 months who have received at least one dose of Vitamin A	Age 0-3 months exclusively breastfed
<b>India</b>	67.6	94.9	65.4	66.8	57.6	33.6	42.3	42.0	29.7	55.2
<b>North</b>										
Delhi	46.8	55.4	83.5	84.9	77.8	59.1	65.9	69.8	32.7	13.2
Haryana	56.8	76.8	58.1	79.7	67.0	22.4	42.0	62.7	45.2	47.2
Himachal Pradesh	34.4	42.4	86.8	66.2	85.6	28.9	40.2	83.4	71.1	17.5
Jammu & Kashmir	65.0	80.1	83.2	77.7	70.8	35.6	42.4	56.7	36.0	41.5
Punjab	57.1	72.1	74.0	89.9	79.6	37.5	62.6	72.1	56.5	36.3
Rajasthan	80.4	114.9	47.5	52.1	39.3	21.5	35.8	17.3	17.6	53.7
<b>Central</b>										
Madhya Pradesh	86.1	137.6	61.0	55.0	48.9	20.1	29.7	22.4	24.4	64.2
Uttar Pradesh	86.7	122.5	34.6	51.4	32.4	15.5	22.4	21.2	13.9	56.9
<b>East</b>										
Bihar	72.9	105.1	36.3	57.8	24.1	14.6	23.4	11.0	10.2	55.2
Orissa	81.0	104.4	79.5	74.3	67.6	22.6	33.4	43.7	42.0	58.0
West Bengal	48.7	67.6	90.0	82.4	71.6	40.1	44.2	43.8	43.4	48.8
<b>Northeast</b>										
Arunachal Pradesh	63.1	98.1	61.6	45.6	56.3	31.2	31.9	20.5	20.9	(33.9)
Assam	69.5	89.5	60.1	51.7	55.0	17.6	21.4	17.0	15.4	42.5
Manipur	37.0	56.1	80.2	64.2	50.0	34.5	53.9	42.3	38.4	69.7
Meghalaya	89.0	122.0	53.6	30.8	49.5	17.3	20.6	14.3	24.7	16.1
Mizoram	37.0	54.7	91.8	37.8	72.7	57.7	67.5	59.6	70.6	40.7
Nagaland	42.1	63.8	60.4	50.9	42.5	12.1	32.8	14.1	6.8	43.9
Sikkim	43.9	71.0	69.9	52.7	62.4	31.5	35.1	47.4	45.8	16.3
<b>West</b>										
Goa	36.7	46.8	99.0	86.1	94.7	90.8	90.8	82.6	78.0	*
Gujarat	62.6	85.1	86.4	72.7	78.0	46.3	53.5	53.0	51.9	65.2
Maharashtra	43.7	58.1	90.4	74.9	84.8	52.6	59.4	78.4	64.7	38.5
<b>South</b>										
Andhra Pradesh	65.8	85.5	92.7	81.5	81.2	49.8	65.2	58.7	24.8	74.6
Karnataka	51.5	69.8	86.3	74.9	78.0	51.1	59.1	60.0	48.4	66.5
Kerala	16.3	18.8	98.8	86.4	95.2	93.0	94.0	79.7	43.6	68.5
Tamil Nadu	48.2	63.3	98.5	95.4	93.2	79.3	83.8	88.8	16.2	48.3

( ) Based on 25-49 unweighted cases  
\*Percentage not shown; based on fewer than 25 unweighted cases  
<sup>9</sup>Per 1,000 live births for the five years preceding the survey (1994-98)  
<sup>10</sup>Includes only the two most recent births  
<sup>11</sup>BCG, measles, and three doses each of DPT and polio vaccines

Figure 25: Data Tables

NFHS-2 FACT SHEET - STATES (Contd.)

State	Percent of children			Percent of children under age three years <sup>12</sup>			Percent of women with BMI <sup>14</sup> below 18.5 kg/m <sup>2</sup>	Percent of women age 15-49 with any anaemia	Percent of women reporting a reproductive health problem <sup>15</sup>	Percent of women age 15-49 who have heard of AIDS
	Age 6-9 months receiving breast milk and solid/mushy food	Age 1-35 months with diarrhoea who received ORS <sup>13</sup>	Age 6-35 months with any anaemia	Underweight	Stunted	Wasted				
<b>India</b>	33.5	26.8	74.3	47.0	45.5	15.5	35.8	51.8	39.2	40.3
<b>North</b>										
Delhi	37.0	39.1	69.0	34.7	36.8	12.5	12.0	40.5	36.5	79.2
Haryana	41.8	25.7	83.9	34.6	50.0	5.3	25.9	47.0	38.2	44.3
Himachal Pradesh	61.3	45.6	69.9	43.6	41.3	16.9	29.7	40.5	33.7	60.9
Jammu & Kashmir	38.9	47.5	71.1	34.5	38.8	11.8	26.4	58.7	60.5	31.9
Punjab	38.7	42.3	80.0	28.7	39.2	7.1	16.9	41.4	28.3	54.6
Rajasthan	17.5	20.3	82.3	50.6	52.0	11.7	36.1	48.5	43.2	20.8
<b>Central</b>										
Madhya Pradesh	27.3	29.8	75.0	55.1	51.0	19.8	38.2	54.3	44.9	22.7
Uttar Pradesh	17.3	15.8	73.9	51.7	55.5	11.1	35.8	48.7	38.1	20.2
<b>East</b>										
Bihar	15.0	15.4	81.3	54.4	53.7	21.0	39.3	63.4	44.2	11.7
Orissa	30.1	35.1	72.3	54.4	44.0	24.3	48.0	63.0	27.5	39.0
West Bengal	46.3	40.5	78.3	48.7	41.5	13.6	43.7	62.7	45.3	26.4
<b>Northeast</b>										
Arunachal Pradesh	(60.2)	40.2	54.5	24.3	26.5	7.9	10.7	62.5	42.1	60.4
Assam	58.5	37.1	63.2	36.0	50.2	13.3	27.1	69.7	50.6	33.7
Manipur	86.8	50.7	45.2	27.5	31.3	8.2	18.8	28.9	56.0	92.9
Meghalaya	77.1	22.4	67.6	37.9	44.9	13.3	25.8	63.3	66.9	44.2
Mizoram	(74.2)	44.7	57.2	27.7	34.6	10.2	22.6	48.0	52.5	93.2
Nagaland	81.3	29.7	43.7	24.1	33.0	10.4	18.4	38.4	45.6	72.4
Sikkim	87.3	27.0	76.5	20.6	31.7	4.8	11.2	61.1	48.6	53.6
<b>West</b>										
Goa	(65.4)	55.6	53.4	28.6	18.1	13.1	27.1	36.4	40.2	76.3
Gujarat	46.5	28.9	74.5	45.1	43.6	16.2	37.0	46.3	28.6	29.8
Maharashtra	30.8	33.2	76.0	49.6	39.9	21.2	39.7	48.5	40.0	61.1
<b>South</b>										
Andhra Pradesh	59.4	39.6	72.3	37.7	38.6	9.1	37.4	49.8	48.5	55.3
Karnataka	38.4	34.3	70.6	43.9	36.6	20.0	38.8	42.4	18.8	58.1
Kerala	72.9	47.9	43.9	26.9	21.9	11.1	18.7	22.7	42.4	86.9
Tamil Nadu	55.4	27.9	69.0	36.7	29.4	19.9	29.0	56.5	27.8	87.3

( ) Based on 25-49 unweighted cases  
<sup>12</sup> Oral rehydration salts  
<sup>13</sup> Underweight assessed by weight-for-age, stunting assessed by height-for-age, wasting assessed by weight-for-height; undernourished children are those more than two standard deviations below the median of the International Reference Population, recommended by the World Health Organization.  
<sup>14</sup> Body mass index  
<sup>15</sup> Currently married women with abnormal vaginal discharge, symptoms of a urinary tract infection, painful intercourse, or bleeding after intercourse

Figure 26: Data Tables

## References

- Clinic, Cleveland. 2022. *BMI for Women: How It Works and What It Reveals about Your Health*. <https://health.clevelandclinic.org/bmi-for-women/>.
- Clinic, Mayo. 2022a. *Anemia*. <https://www.mayoclinic.org/diseases-conditions/anemia/symptoms-causes/syc-20351360>.
- . 2022b. *Prenatal Vitamins: Why They Matter, How to Choose*. <https://www.mayoclinic.org/healthy-lifestyle/pregnancy-week-by-week/in-depth/prenatal-vitamins/art-20046945#:~:text=Ideally%2C%20you'll%20begin%20taking,of%20healthy%20red%20blood%20cells>.
- Disease Control, Centers for, and Prevention. 2022. *Tdap Vaccine Safety for Mother and Infant*. <https://www.cdc.gov/pertussis/pregnant/hcp/vaccine-safety.html#:~:text=Pregnant%20women%20have%20been%20getting,1960s%20to%20prevent%20neonatal%20tetanus>.
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna M. Wallach, Hal Daumé III, and Kate Crawford. 2018. “Datasheets for Datasets.” *CoRR* abs/1803.09010. <http://arxiv.org/abs/1803.09010>.
- Iannone, Richard, and Mauricio Vargas. 2022. *Pointblank: Data Validation and Organization of Metadata for Local and Remote Tables*. <https://CRAN.R-project.org/package=pointblank>.
- Lucidchart. 2022. <https://www.lucidchart.com/pages/>.
- Ooms, Jeroen. 2022. *Pdftools: Text Extraction, Rendering and Converting of PDF Documents*. <https://CRAN.R-project.org/package=pdfutils>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wikipedia. 2022. *Demographic and Health Surveys*. [https://en.wikipedia.org/wiki/Demographic\\_and\\_Health\\_Surveys](https://en.wikipedia.org/wiki/Demographic_and_Health_Surveys).
- Xie, Yihui. 2022. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.