# Statistical Analysis of Census Data

Summer Research Internship Program - 2024,
Dyal Singh College (Univesity of Delhi), Lodhi Road, New Delhi - 110003
Is being submitted to: Internal Quality Assurance Cell (IQAC), DSC

By
*Gaurav Godara, Lehar Bansal, Shubham Tiwari, Yash Kumar Singh*
*Dyal Singh College (University of Delhi).*

*Stuti Gupta*
*Kalindi College (University of Delhi), Delhi.*

**Mentors**:

Harleen Kaur Daber, Department of Geography,

Dinesh Kumar, Department of Mathematics,

Naveen Gaur, Department of Physics,

Nikhil Jain, Department of Political Science,

Dyal Singh College (University of Delhi) Delhi-110003, India

# Acknowledgement

This summer project was undertaken in the Dyal Singh College (University of Delhi), Lodhi Road, New Delhi-110003 during the period June-August 2024. We thank the IQAC cell of the college for the support. We also thank Sachin N. (Department of English, Dyal Singh College) for the discussions.

Gaurav Godara          Lehar Bansal          Shubham Tiwari          Stuti Gupta          Yash Kumar Singh
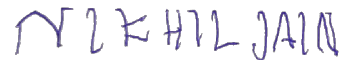
Harleen Kaur Daber          Dinesh Kumar          Naveen Gaur          Nikhil Jain

**Disclaimer**: We do not claim the originality of the work.
Authors and mentors are listed in alphabetical order.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

*"Census as an institution goes back to the remote past, but it is no longer a mere counting of the heads; it involves extraction of information which plays a vital role in determining many of our administrative policies. The facts elicited during this study yield valuable scientific data of sociological importance. In many matters, it provides a useful guide for the effectiveness or otherwise of our economic policies. The theory of population is in itself an interesting part of economics. The census helps us to test and adapt that theory to facts."*

- Sardar Vallabhbhai Patel, 1st Home Minister of India on the 1951 Census [Office of the Registrar General & Census Commissioner, 2023])

To gauge the progress of a nation and the impact of its policies on the ground and people it is essential to have an in-depth analysis of the actual data. The closest we can get to the representation of the nation is through its census data. It is a rich source of the actual representation of its people spanning various variables and policies. Over the decades, the Indian census has evolved, incorporating more comprehensive data collection methods and expanding the range of information gathered.

If we specifically talk about India, the census here is conducted every ten years, with the first census being conducted in 1872 under Viceroy Lord Mayo and the last one was conducted in 2011.

Due to the global pandemic starting in 2019, the Registrar General and Census Commissioner of India were not able to conduct census in 2021. Hence for this study, we have worked with the data from the most recent census conducted in 2011. In the present report, we have tried to analyze literacy and its dependence on some factors using 2011 census data.

## 1.1 Literacy

*"The waves of the ocean, before reaching the child, break against four walls, which encompass the water of his education and crystallization - father, mother, brothers and sisters, and a few extra people, are his forming-world and mold. But, all this deducted, we must remember in education that its power, like that of the spirit of the age, which must not be measured by individuals but by the concentrated mass or majority, must be judged, not by the present, but by the future."*

- Richter, A General Report, Census of India, 1891 [Office of the Registrar General
& Census Commissioner, 1891]

To transform society into a meaningful and productive resource, a country needs to work on educating society. The minimum required is that it can read and write in one language. The effects of education may not be immediately visible but they create long-term benefits for society, enriching all aspects of life.

Before embarking on the education journey, literacy is the essential first step. In common terms, literacy is the ability to read and write at least a simple message in any language; conversely, illiteracy refers to the lack or absence of this ability. The literate society is the one in which most people are literate. But this does not mean that a person with just the ability to read and write a simple message can meet all the challenges of this complex world. Therefore, we in this report also try to analyze the formal education pattern as given by census data.

The definition of literacy is given on Page - 122 [Hassan, 2020] and is:
*Information on the extent of literacy forms an integral part of census enumeration in countries where census counts are taken regularly. Differences in the definition of literacy and enumeration procedures render data on literacy across different countries incomparable. However, now most countries are switching to the definition proposed by the United Nations Population Commission- 'the ability of people to read and write a simple message in any language with some understanding'. This has made it easier to compare literacy rates, internationally. The Indian Census has been using the UN definition. However, the problem persists as the different countries use different tabulation techniques for literacy statistics. Some countries compute the rate of literacy taking into account the total population. India, too, followed this technique till the 1981 census. It is, however, argued that since children, particularly in the early age groups, do not possess the potential to acquire literacy in the true sense of the term to qualify as functionally literate, they should be excluded from the population while computing literacy levels. In some countries, therefore, the population below five years of age is ignored while computing the proportion of literate. In India, from 1991 onwards, the population in the age group 0–6 years is excluded while calculating the literacy rate.*

### 1.1.1 Benefits of Literacy

According to the Education For All (EFA) Global Monitoring Report, 2006 published by UN-ESCO, Literacy brings wide-ranging benefits, but measuring these benefits precisely is difficult. Research often conflates literacy with general schooling or adult literacy programs and few studies rigorously assess literacy's direct impact on cognitive skills and long-term effects. Additionally, benefits such as political awareness and empowerment are challenging to quantify. Considering all these challenges, we will look into some of the benefits offered by literacy:

- **Human Benefits:** Literacy enhances self-esteem, confidence, and personal empowerment, leading to greater individual and collective action. It is associated with increased civic engagement and political involvement.

- **Cultural Benefits:** The cultural benefits of literacy are harder to define but include fostering critical thinking and challenging stereotyped traditional attitudes. Literacy programs can lead to new norms and attitudes, particularly among women. Additionally, literacy in minority languages enhances cultural diversity by helping people engage with and preserve their cultural heritage.

- **Social Benefits:** Improving literacy yields significant social benefits. Women who participate in literacy programs have better knowledge of health, and family planning and are more likely to adopt preventive health measures. Educated parents, including those from adult programs, are more supportive of their children's education. Literacy programs often target women, enabling them to access male-dominated areas and gain a stronger voice in their households. Some programs also address gender issues at the community level.

- **Economic Return:** Education's economic benefits are well-documented, particularly regarding individual income and overall economic growth. The data from many countries indicate that improving literacy levels have a positive impact on economic growth. For example, literacy differences explain a significant portion of growth variations among OECD (Organization for Economic Cooperation and Development) nations and positively affect the Gross Domestic Product (GDP) in African and Islamic developing countries. Although primary education is often seen as more cost-effective, investing in adult literacy programs can be just as beneficial. Some studies find that the costs of adult literacy programs are much lower compared to primary education, yet they deliver strong economic returns.

### 1.1.2  Literacy in India

In India, ever since the first census in 1872, it has served as the main source of literacy data. Over time, however, the definitions, understanding, and classification of literacy and education have evolved significantly. The rules for filling the columns required for getting information in Census 1891 [Office of the Registrar General & Census Commissioner, 1891]:

*Col.12. Enter in this column against each person, whether grown-up, child, or infant, either 'Learning', 'Literate', or 'Illiterate'. Enter all those as learning who are under instruction, either at home or at school or college. Enter as Literate who can both read and write, but are not under instruction as above. Enter as Illiterate those who are not under instruction, and who do not know how to both read and write, or who can read but not write, or can sign their name but not read.*

Therefore, until the Census of 1891, there were three categories to classify the population. However, due to some inherent problems with this classification, the census of India abandoned this method and switched to a two-fold classification: literates and Illiterates. Although this sharp, binary distinction made classification easier, as a person would now be either literate or illiterate, with no intermediate categories, still since complexity is a fundamental aspect of the real world; things in the real world are inherently complex and hence this complexity extends to the simple binary classifications of literacy also, which often fail to capture the full range of abilities and skills involved as literacy spans various levels of proficiency, requiring a more nuanced understanding and measurement to fully grasp individuals' capabilities and the level of literacy.

**Literacy From Past to Present**

Even though India is a land known for its knowledge and education where the world's first residential university, home to nine million books that attracted around ten thousand students from all over the world, was situated, the country's population was marked with an abysmally low level of literacy at the time census operation began in the country. *The reasons for this generally include factors like the caste-based social system, primarily agricultural-based economy, low rate of population mobility, predominantly rural society, unfavorable social values towards education and literacy, prejudice against female education and mobility, extreme poverty, and above all rapid population growth* - (Page-129 [Hassan, 2020]).

Before the 1991 census of India, the method used to calculate literacy rates was *Crude Literacy Rate* which included the total population in the denominator but from the census 1991, the base population was changed from the total to "population aged 7 years and above." and this method is called Effective Literacy Rates. Also, from the 1961 to 1981 censuses, all children below the age of 5 years were considered out of the literate calculation while after the census in 1981, this age was changed to seven years instead of five.

There has been a significant improvement in literacy rates in India over the last century.

However, comparing these rates across such a long period can be problematic due to changes in the definition of literacy and varying methodologies. To make meaningful comparisons, it is crucial to establish a consistent base by using a single method, whether the crude literacy rate or the effective literacy rate. This approach ensures that the data is comparable and accurately reflects the progress made in literacy over time. However, before 1981, age and sex-wise data on literacy and educational attainment were tabulated in five-year age groups. Specifically, literates aged 5 and 6 years could not be distinguished from those in the 5-9 years age group due to the data aggregation methods used. This limitation makes it infeasible to compute effective literacy rates according to the 1991 Census definition for periods before 1981.
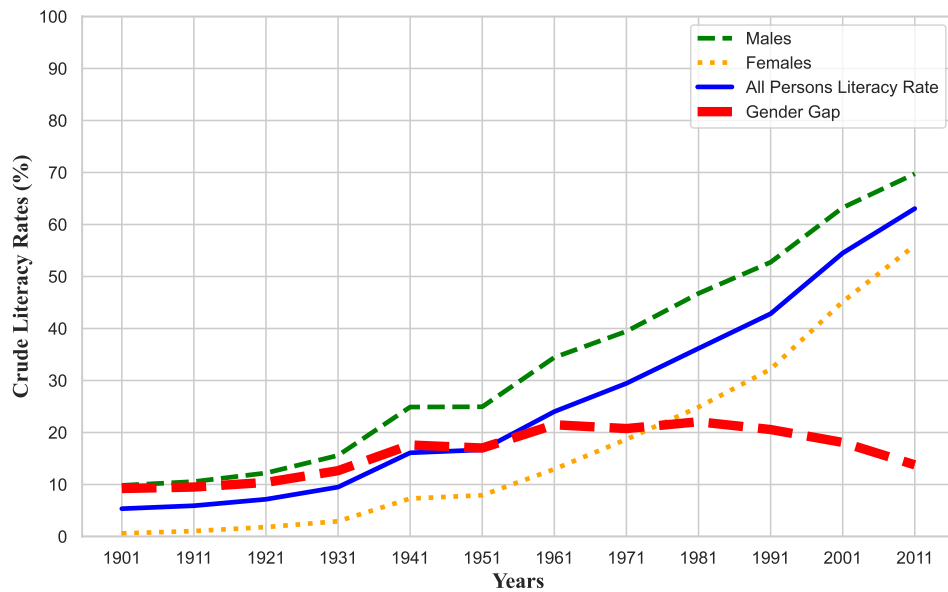


Figure 1.1: *Crude Literacy rate as a function the years when Census of India was conducted*

Figure 1.1[1] provides a visual representation of the data to illustrate the trends in crude literacy rates over the past century.

1. Figures from 1901 to 1941 are for undivided India.

---

[1]Sources of data for the figure

1. Provisional Population Table, Census of India,1891, Series-1 [Padmanabha, 1971]

2. Census of India,1991 [Office of the Registrar General & Census Commissioner, 1991]

3. Census of India,2001 [Office of the Registrar General & Census Commissioner, 2001]

4. Census of India,2011 [Office of the Registrar General & Census Commissioner, 2011d]

2. Figures for 1981 exclude Assam and those for 1991 exclude Jammu and Kashmir as no census could be conducted in Assam in 1981 and in Jammu and Kashmir in 1991.

3. For the years 1991,2001 and 2011, there is no information on literate children aged 5 and 6 years, hence in computing literacy rates, the crude literacy rates are an underestimate by a small amount.

The use of crude literacy rates allows for consistent historical comparisons despite methodological limitations and changes in definitions over time. This graph 1.1 offers a clear depiction of progress, allowing for easy comparison across different periods. It highlights the overall upward trend in literacy and notable milestones, making it evident how literacy rates have evolved in India.

The figure effectively illustrates the progressive trends in crude literacy rates in India from 1901 to 2011, highlighting significant improvements across genders over time. Initially, women's literacy was notably poor, with a literacy rate of just 0.60% in 1901, starkly contrasting with the male literacy rate of 9.83%.

The gender gap in literacy rates kept on increasing till 1961 and thereafter started to show a marginal and consistent decline, indicating significant strides toward gender parity in education. This figure encapsulates India's educational journey, highlighting both achievements and ongoing challenges in bridging the gender literacy gap.

Having seen the condition of literacy over the decades, we decided to conduct a detailed analysis of the census data to identify patterns, distribution, and disparities within literacy in India at the zonal council level in the census year 2011. This approach provides a novel perspective, differing from the usual state or national-level analyses found in the existing literature.

The analysis at the zonal council level allows for a more granular examination of literacy trends and disparities. By focusing on these administrative divisions, we aim to uncover regional variations that may be masked in broader analyses. This can help to identify specific areas that require targeted interventions and policies. Our methodology involves collecting and interpreting data from the official census 2011 data and comparing literacy rates among different zones and also within the zone among male, female, rural, urban, and different religious communities.

To illustrate our findings, we have presented results from the 2011 census data:

1. **Total Literacy Rate over the Zonal Councils**: This overview of literacy rates across different zonal councils provides a comparative analysis of educational attainment in each zone. By highlighting these variations, we can better understand the regional disparities in literacy.

2. **Male and Female Literacy Rates in Rural/Urban Areas at the Zonal Council**

**Level**: Here we break down literacy rates by gender and area of residence within each zonal council. It helps in identifying urban-rural divides and gender disparities, offering insights into the specific needs and challenges faced by different demographic groups within each zone.

3. **Zone-wise Literacy Rates of All Religious Communities**: They provide a detailed view of literacy rates among different religious communities within each zonal council. This analysis aims to uncover any disparities in educational attainment among various religious groups, contributing to a deeper understanding of the sociocultural factors influencing literacy.

In this report, we will present our findings through graphs, and tables that elucidate the literacy landscape across the different zonal councils. We will delve into the factors influencing literacy rates, such as socioeconomic conditions, urban -rural divides, and gender disparities. By providing a zonal perspective, we hope to contribute a fresh and insightful analysis that can inform policymakers and educational planners in their efforts to enhance literacy across India. Also, note that further in our report we have used the term 'literacy' to denote *effective literacy rate*.

This report is organized as follows: In Chapter 2 the conceptual framework used in the report. Chapter 3 is devoted to the methodology used for analyzing the census data. Finally, we will conclude in Chapter 4 with our results and analysis.

# Chapter 2

# Conceptual Framework

In this chapter, we will look into the conceptual framework that undermines this report. This includes terminologies and the description of the zones according to which the states of India are divided.

## 2.1  Theoretical Framework

In this section, we have given the definitions and the nomenclature used in our report.

### 2.1.1  Literates

According to the 2011 Census, Literates are defined as per the following definition ( [Office of the Registrar General & Census Commissioner, 2011f]) :

1. Aged 7 years and above

2. Can read AND write in any language

3. Need not have:

    - Any formal education
    - Minimum educational standard
    - Probing of this aspect is particularly important for women and girls

4. A blind person who can read Braille is literate.

### 2.1.2  Illiterates

According to the 2011 Census, Illiterates are defined as per the following definition: ( [Office of the Registrar General & Census Commissioner, 2011f])

1. Anyone less than 7 years of age even if he/she knows how to read and write

2. One who can read but not write in any language

3. One who can only sign his/her name

### 2.1.3 Literacy Rate

Throughout the years the Literacy Rate was calculated in India through the following measures ( [Das, 2012])

- Literacy Rate was calculated through the following formula before 1981. This is also known as the *Crude Literacy Rate*:

$$\left(\frac{\text{Number of Literate persons aged 5+ years}}{\text{Total Population}}\right) \times 100$$

- After 1981, the Literacy Rate was calculated using the formulae below. This is also known as the *Effective Literacy Rate*:

$$\left(\frac{\text{Total Literates}}{\text{Total Population Age 7 and Above}}\right) \times 100$$

### 2.1.4 Urban Areas

According to the 2011 Census, Urban Area is defined as follows : ( [Office of the Registrar General & Census Commissioner, 2011g])

1. All places with a municipality, corporation, cantonment board, notified town area committee, etc.

2. All other places which satisfied the following criteria:

   - A Minimum Population of 5000;

   - At least 75 percent of the male working population engaged in non-agricultural pursuits; and

   - A density of population of at least 400 per sq. km.

### 2.1.5 Rural Areas

According to the 2011 Census, Rural Area is defined as per the following definition: ( [Office of the Registrar General & Census Commissioner, 2011g])

   All areas that are not categorized as Urban area are areas which are not categorized as Urban area are considered Rural Areas.

### 2.1.6 Other Religions

According to the 2011 Census, the category of "Others Religion" includes the following religions and persuasions: ( [Office of the Registrar General & Census Commissioner, 2011h])

Addi Bassi, Adi, Aka, Animist, Apo Rangang, Baiga, Bhil, Bhoi, Bhumia, Birsa, Bodo/Boro, Bori, Chang Naga, Dongi, Doni Polo/Sidonyi Polo, Dupub, Fralung, Garo, Gond/Gondi, Hajong, Halba, Heraka, Hill Miri, Ho, Idu/Idu Mishmi, Kaman/Miju Mishmi/Kaman Mishmi/Miju, Karbi/Mikir, Katkari, Kharwar, Khasi, Kisan, Korku, Koyatur, Krupa, Marangboro, Munda, Nani Intiya, Nature Religion, Niam Shnong, Niamtre, Nocte, Non Christians, Nyarino, Oraon, Pagan, Paharia, Pardhi, Santal, Saranath, Sari Dharma, Sarna, Sarnam, Songsarek, Swarna, Tadvi, Tana Bhagat, Tribal Religion, Yumasam, Budhadeo, Intaya, Rangfra, Bamanya, Hidmaraj, Subba, Rangkho Thak, Tikao Ragong, Paniyar, Mannan, Baigani Dharam, ADI KURUM, Adim Dhamm, A.C., Bahai/Bahais, Jews/Judaism, Nirankari, Parsi/Zoroastrian, Sadri, Sanamahi, Traditional Religion, Dera Sarsa, ADI DHARM, Bidin, and Atheist.

## 2.2 Zonal Councils

In this report, we have presented results for various Zonal Councils. In this section, we have presented the definition of Zonal Councils.

### 2.2.1 Background

The idea of the creation of Zonal Councils was mooted by the first Prime Minister of India, Pandit Jawahar Lal Nehru in 1956 when during a debate on the report of the States Reorganisation Commission. He suggested that the States proposed to be reorganized may be grouped into four or five zones having an Advisory Council *to develop the habit of cooperative working* among these States.

### 2.2.2 Composition of Zonal Councils

In the light of the vision of Pandit Nehru, five Zonal Councils were set up vide Part III of the States Re-organisation Act, 1956. The present composition of each of these Zonal Councils is given below. By composition, we means the name of the states that comes under respective zones.

1. The Northen Zonal Council: Haryana, Himachal Pradesh, Jammu & Kashmir, Punjab, Rajasthan, National Capital Territory of Delhi and Union Territory of Chandigarh;

2. The Central Zonal Council: Chhattisgarh, Uttarakhand, Uttar Pradesh and Madhya Pradesh;

3. <u>The Eastern Zonal Council</u>: Bihar, Jharkhand, Orissa, and West Bengal;

4. <u>The Western Zonal Council</u>: Goa, Gujarat, Maharashtra and the Union Territories of Daman & Diu and Dadra & Nagar Haveli;

5. <u>The Southern Zonal Council</u>: Andhra Pradesh, Karnataka, Kerala, Tamil Nadu and the Union Territory of Puducherry.

6. <u>The North Eastern States</u>: The states Assam, Arunachal Pradesh, Manipur, Tripura, Mizoram, Meghalaya, and Nagaland are not included in the Zonal Councils and their special problems are looked after by the North Eastern Council, set up under the North Eastern Council Act, 1972. The State of Sikkim is also included in the North Eastern Council vide North Eastern Council (Amendment) Act, 2002 notified on 23rd December 2002. [Commission, 1956]

# Chapter 3

# Methodology

This chapter will discuss the methodology used to analyze the raw data files of the 2011 Census of India. A detailed flowchart on how we extracted the information required for our work from the raw census data files is presented. After extracting the data we have discussed how the data is analyzed and the tools used for the analysis/presentation of the data.

## 3.1 Data Sources

The data for this research is sourced from the official Census India website, which provides comprehensive census reports for the years 1991, 2001, and 2011. These reports include detailed information on the literacy level and the total population for each state and union territory. The raw data files used in our work can be downloaded from the following link:

`https://censusindia.gov.in/census.website/data/census-tables`

This chapter is organized as follows: In section 3.2 we have explained how raw data from the Census of India website was processed to provide us with the data files we used for our analysis. Section 3.3 is devoted to how we have analyzed the data files, the techniques, and the software/libraries used. Finally, the section 3.4 is regarding the presentation of our data.

## 3.2 Data Collection and Processing

The following steps were taken to collect and process the raw data taken from the Census of India website (link given above) :

1. **Data Extraction:** Census data from 1991, 2001, and 2011 were obtained from the Census India website. The relevant data on literacy rates and population figures was extracted and organized into a structured format for analysis.

2. **Data Grouping:** Due to India's vast diversity, with multiple States and Union Territories, it was out of our scope (due to the time constraints) to study each state and union territory individually. Therefore, regions were grouped according to the Zones defined by the State Reorganization Act of 1956, facilitating a more manageable analysis.

3. **Data Filtering and Calculation:** The 2011 census data was processed using the `Pandas` library in `Python`. The following steps were undertaken:

   - Filtering the dataset to include only individuals aged 7 and above, thereby excluding younger age groups to align with the study's objectives.

   - Grouping the data by State and Region (Urban and Rural) and calculating the total and literate population sums.

   - Further calculations were performed to analyze literacy rates by gender and religion, providing a more nuanced understanding of the data.

   - Pivoting the grouped data to place Urban and Rural data side by side, enabling a clear comparison between these regions.

   - Adding columns to calculate the total population and literate population for both Urban and Rural areas.

   - The data was then further categorized into zones: Northen Zone, Southern Zone, North Eastern Zone, Western Zone, Eastern Zone, and Central Zone, based on regional divisions.

   - The processed data was saved in both Excel and CSV formats on GitHub, ensuring accessibility and version control for future analysis.

4. **Excel Calculations:** After processing the data in Python, Excel formulas were applied to calculate literacy rate percentages for Urban and Rural areas, further breaking down the data by gender and religion.

5. **Data Storage:** The processed data was stored and managed using GitHub, ensuring proper version control and accessibility for future analysis.

## 3.3   Analytical Tools and Techniques

The analysis was conducted using Python, leveraging several powerful libraries:

- **Jupyter Notebook**: `Jupyter Notebook` is an open-source web application that enables the use of data analytics and visualization using `Python`.

- **Jupyter Lab:** Provided an interactive development environment for data analysis.

18

- **Pandas:** Used for data manipulation and analysis, offering data structures and functions necessary for working with census data.

- **Numpy:** Utilized for numerical computations and efficient handling of large data arrays.

- **Seaborn:** Employed for creating statistical graphics and visualizing data distributions and relationships.

- **Matplotlib:** Used to create static and interactive, visualizations.

- **Plotly:** Enabled the creation of interactive plots and visualizations.

- **Mitosheet:** Facilitated spreadsheet-like interaction with data within the `Jupyter Notebook` environment.

## 3.4 Data Visualization

To effectively communicate the findings, various visualizations were created:

- **Bar Graphs:** Bar graphs are created for each zonal council to display their literacy rates for the year 2011. This provided a clear visual representation of the changes in literacy rates over time.

- **Line Graphs:**: A line graph was superimposed on the bar graphs to compare each state's literacy rate with the national average literacy rate. This allowed for an easy comparison between regional and national trends.

## 3.5 Documentation and Reproducibility

All codes, data, and visualizations have been documented and uploaded to `GitHub`. To see the coding and processing done to make graphs from the census files, one can visit the link- `https://github.com/navgaur/SP-Census-2024`. This ensures transparency and allows others to reproduce or build upon the analysis for further studies.

# Chapter 4

# Results and Analysis

In this chapter, we will explore the data through various lenses, highlighting trends, patterns, and significant variations. Subsequently, a detailed analysis of the results is presented including comparisons across groups and with the National Average Literacy Rate. All the results presented in this chapter are based on the Census of India conducted in the year 2011.
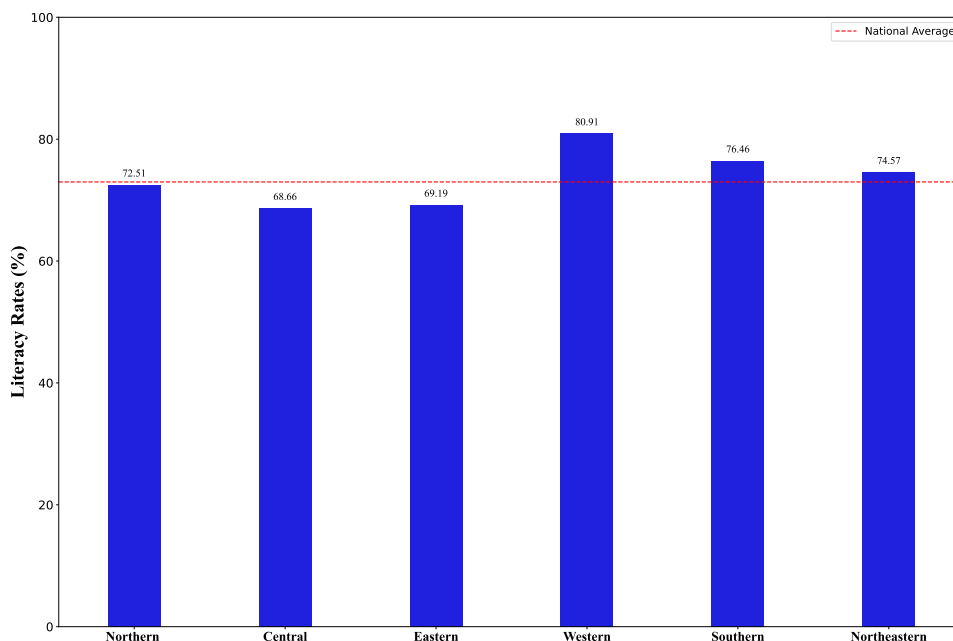
## 4.1   Zone wise Literacy Rates



Figure 4.1: *Total Literacy rate of different zones.*

The bar chart Fig. 4.1 depicts the Literacy Rate (%) across the Zones of India according

to the 2011 Census. The dotted horizontal line in the plot represents the National average Literacy Rate. The Western Zone has the highest literacy rate whereas the Central Zone has the lowest. Interestingly the Southern zone, which consists of Kerala a state with the highest literacy rate (Table A.6) has the second highest literacy rate.

The high literacy rate of the Western Zone can be attributed to States like Goa and Maharashtra which have the 4th and 12th (Table A.5) highest literacy rates and also Union Territories like Daman and Diu which have the 6th (Table A.5) highest literacy rate. On the other hand, the low literacy rate of the Central Zone can be attributed to states like Chhattisgarh, Madhya Pradesh, and Uttar Pradesh which have some of the lowest literacy rates in India ranked 27th, 28th, and 29th (Table A.2) respectively.

We can observe that Zones like the Northen, the Eastern, and the Central are all below the National Literacy Rate and zones like the North Eastern Council, Southern, and Western are above the National Literacy Rate.
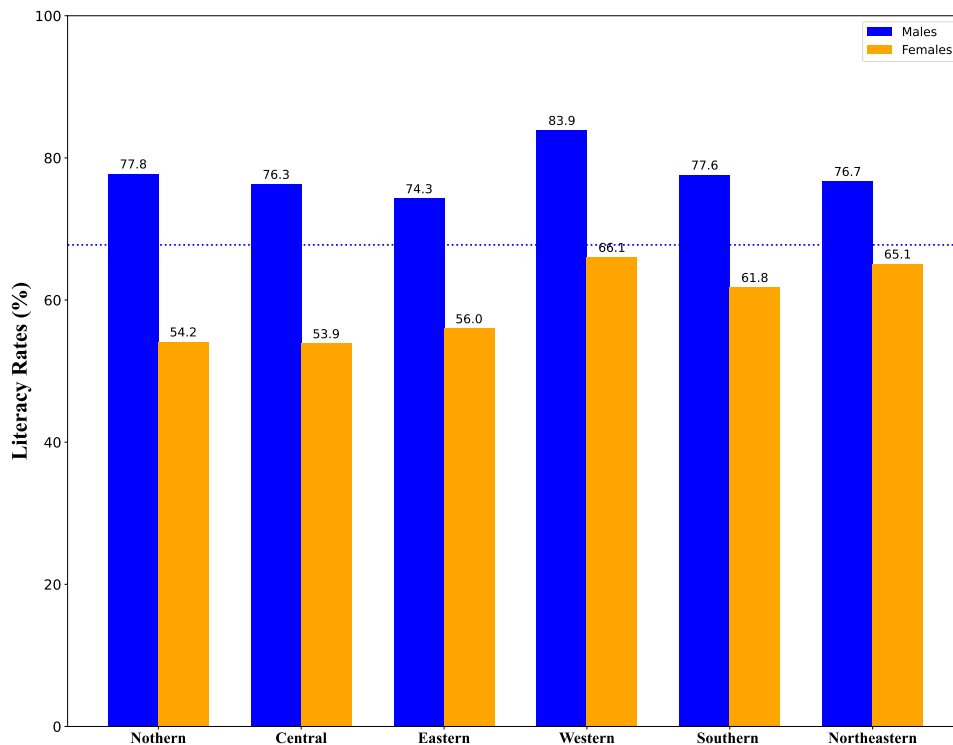
## 4.2 Rural/Urban disparities in Literacy Rates



Figure 4.2: *Rural Literacy rate. The horizontal blue dotted line represents the average national rural literacy rate which is 67.7%.*
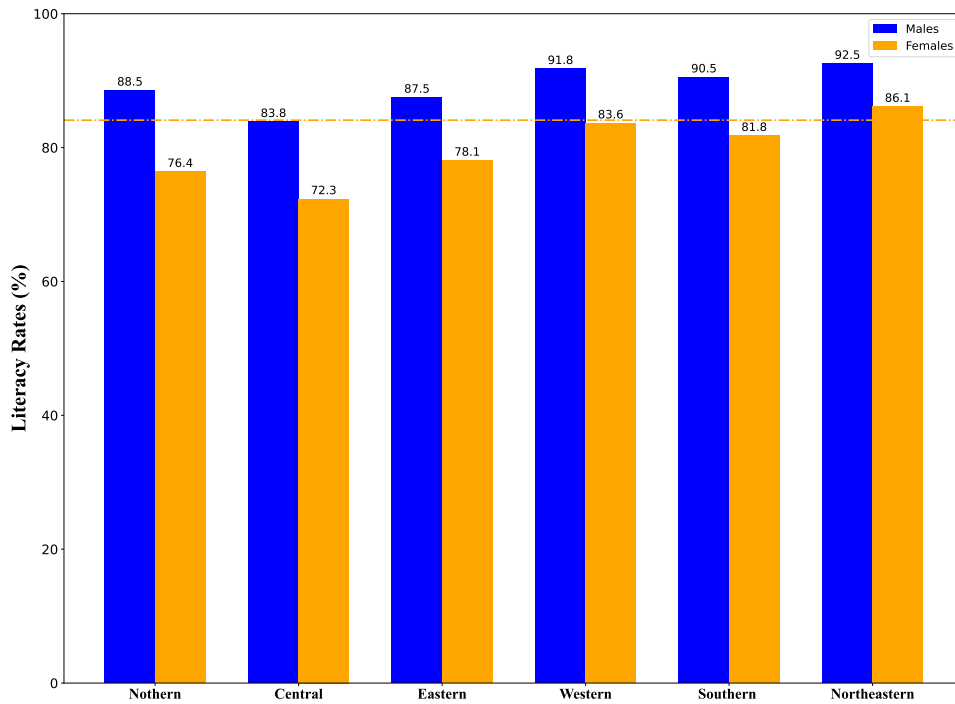
Figure 4.3: *Urban Literacy rate. The horizontal dotted line represents the average national urban literacy rate which is 84.1 %.*

In continuation of our analysis of literacy rates across different zonal councils, we now delve deeper into the demographic disparities in literacy rates. The accompanying graphs, 4.2 and 4.3, delineate the literacy rates of males and females in rural and urban areas, respectively. These graphs provide a detailed perspective on how literacy varies within these demographics.

As can be seen from Fig 4.2 the rural literacy rate for Males in all the zones is above the National Rural Literacy Rate but the same cannot be said for their counterparts as the Female Rural Literacy Rate is below the National Rural Literacy Rate showcasing the Gender disparities across India.

In some zones, this disparity is at the low end of the spectrum for example in the North Eastern Zone or the Southern Zone where the Gender Gap in the rural literacy rate is only 11.6% and 15.8% respectively. On the other hand in Zones like the Northen Zone and the Central Zone, this gap is at the high end of the spectrum and is as large as, 23.6% and 22.4%, respectively. The Western Zone has the highest male and female literacy in India. But gender disparity is glaring here also. The male literacy at 83.9% and female literacy at 66.1%, a gender gap is about 17.8% is observed in the zone which is also quite remarkable.

As evident from Fig 4.3 Central Zone still has the lowest urban literacy rate for both males and females followed by the Eastern and Northen Zone have 2nd lowest male and female literacy rates, respectively. Unlike the Rural Literacy Rate we see here, there is a variation, all the Male

Urban Literacy Rates are above the National Urban Literacy Rate except for the Central Zone. On the other hand for females despite all the other zones having urban literacy rates below the National Urban Literacy Rate, the North Eastern Zone literacy rate for females is above the same.

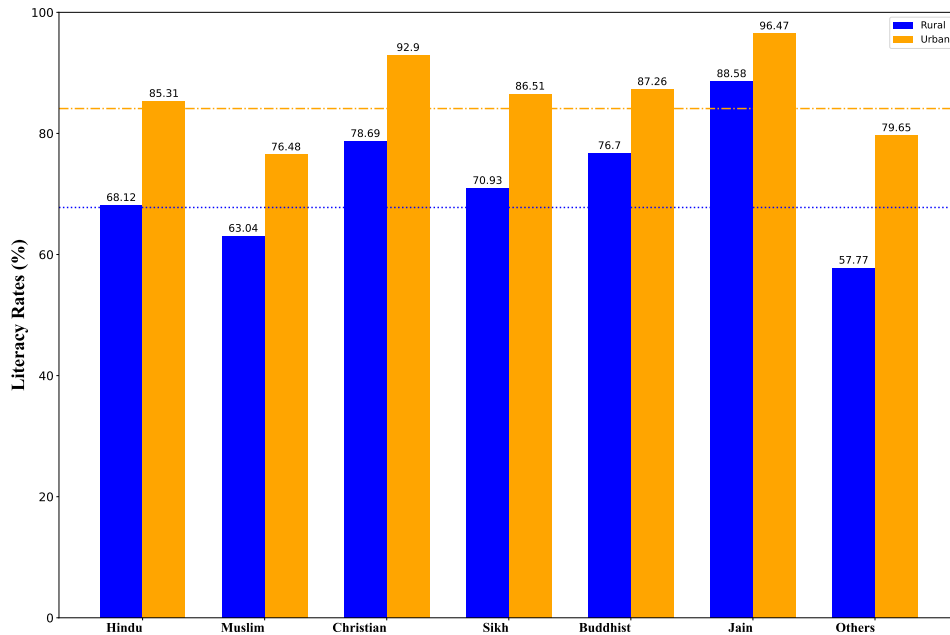## 4.3  Religion wise Literacy Rates



Figure 4.4: *Literacy rates by Religion in National Level. The horizontal lines represent the average national rural (blue dotted line) and urban (orange dashed-dotted line) literacy rates which are 67.7 % and 84.1 % respectively.*

Following the analysis of male and female literacy rates across rural and urban areas in various zonal councils, the examination of literacy rates among different religious communities further elucidates the complex landscape of educational attainment in India.

Figure 4.4 reveals significant disparities among religious groups nationally, with literacy rates consistently higher in urban areas for all communities. Notably, minority religious communities such as Jains and Christians exhibit higher literacy rates compared to the Hindu majority and the Muslim community, the second-largest religious group in India.

**Northen Zone :** Following the national level interpretation, we now examine the literacy rates within each zonal council to find the disparities within each zone itself. Starting with the Northen Zonal Council we can see in figure 4.5 that the data almost shows the same trend
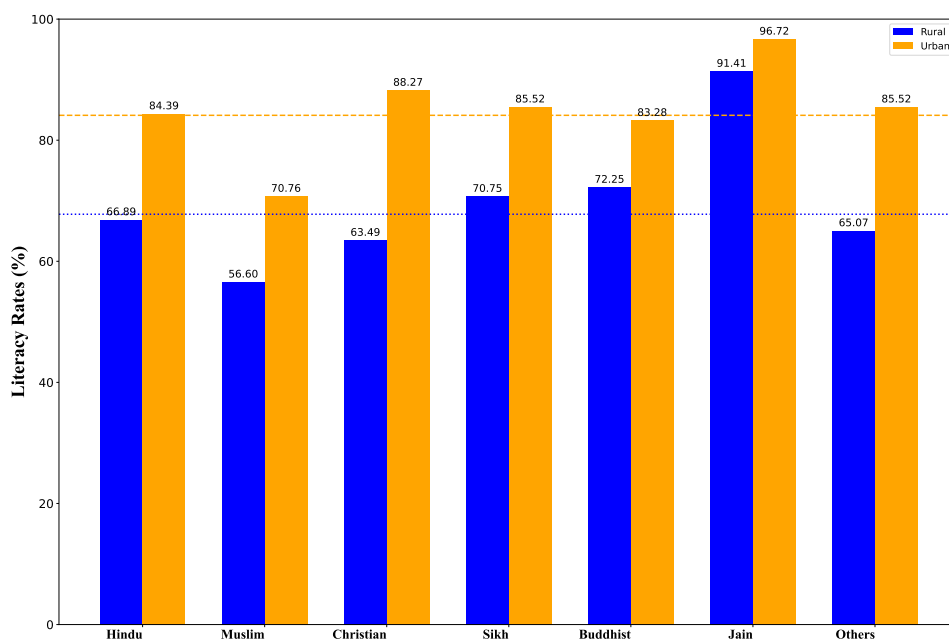
Figure 4.5: *Literacy rates by Religion in Northen Zone. The horizontal lines are as described in Fig. 4.4.*

as at the national level. Still, it is noteworthy that Christian's rural literacy rate is below the national rural average literacy rate. Also, Buddhist's urban literacy rates is less than the national urban literacy rate unlike the rural literacy rate, which surpasses the national average rural literacy rate. The Jain community exhibits the highest literacy rates, both in rural and urban areas. The Muslim community has the lowest literacy rates, both in rural and urban areas. Christians and Sikhs demonstrate relatively high literacy rates. These findings emphasize the diverse educational outcomes across religious communities in the Northen Zonal Council, revealing how regional and cultural factors influence literacy rates in this part of India.

**Central Zone :** As evident from Fig 4.3 this zone has a literacy rate lower than the national average. The religious profile of this zone is given in the figure 4.6. Both the Hindu and Muslim communities fall below the national average literacy rates in both rural and urban regions. Interestingly, in this region, the gender gap in the Muslim community is relatively low. The Jain community shows very high literacy levels with very low gender gap. These things require a deeper understanding of the underlying issues.

**North Eastern Zonal Council :** This zone has a relatively better educational landscape compared to other zones. The Hindu community exhibits significantly improved literacy rates, particularly in rural areas, with a rate of 76%, outperforming their counterparts in other regions.
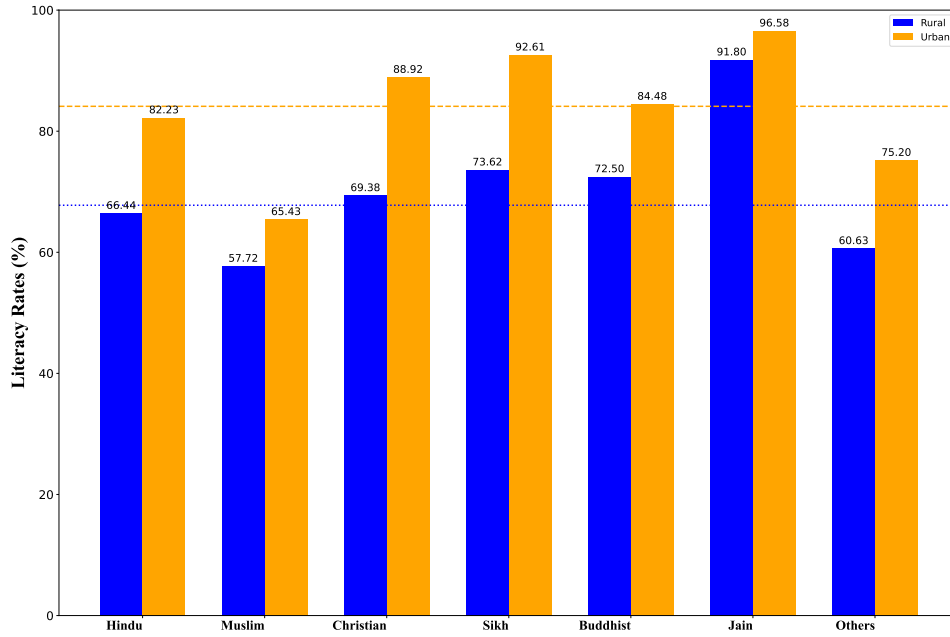
24

Figure 4.6: *Literacy rates by Religion in Central Zone. The horizontal lines are as described in Fig. 4.4.*

However, the Muslim community remains below the national average in both rural (61 %) and urban (78%) areas, reflecting ongoing educational challenges. Interestingly, the Buddhist community's rural literacy rate in the North Eastern Zone is 65 %, which falls below the national rural average literacy rate. The Sikh community shows a minimal gap between rural (90 %) and urban (93 %) literacy rates, indicating a more uniform and favorable educational scenario across different settings. The Christian community, which has a substantial presence in the North East, demonstrates commendably high literacy rates, with 73 % in rural areas and 93 % in urban areas, benefiting from the region's strong emphasis on education. These observations underscore the diverse educational outcomes in the North Eastern Zonal Council, with notable improvements among Hindus and Christians, consistent high performance among Jains, and persistent challenges for Muslims and Buddhists. This regional analysis highlights the importance of tailored educational policies to address specific community needs and promote inclusive literacy advancements.

**Eastern Zone :** In the Eastern Zone, we can see that for the urban areas, the Literacy Rate is highest for the people in the Jain religion. In contrast, it is lowest for the people categorized as "Others" religion followed by the people in the Muslim religion. In rural areas, just like the urban areas, the literacy rate is highest for the Jain Religion and lowest for the people in the "Others" Religion is followed by the people in the Muslim religion. Interestingly though the
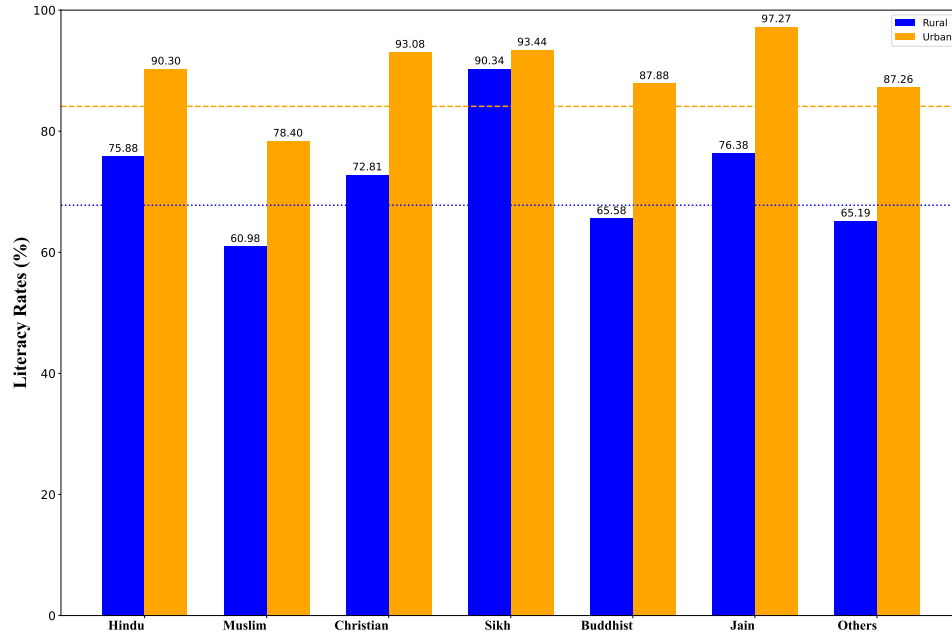
Figure 4.7: *Literacy rates by Religion in North Eastern Zone. The horizontal lines are as described in Fig. 4.4.*

difference between the Rural and Urban Literacy Rates are lowest in the Muslim religion and highest in the Christian Religion standing at 11 % and 22% respectively. When comparing these rates to the National Urban Literacy Rate, all religions in the Eastern Zone, except for Muslims and "Others" are above the national average. In contrast, when comparing the National Rural Literacy Rate to individual religions' literacy rates in the Eastern Zone, only Sikhs, Buddhists, and Jains surpass the national average, while Hindus, Muslims, Christians, and those in the "Others" category fall below it.

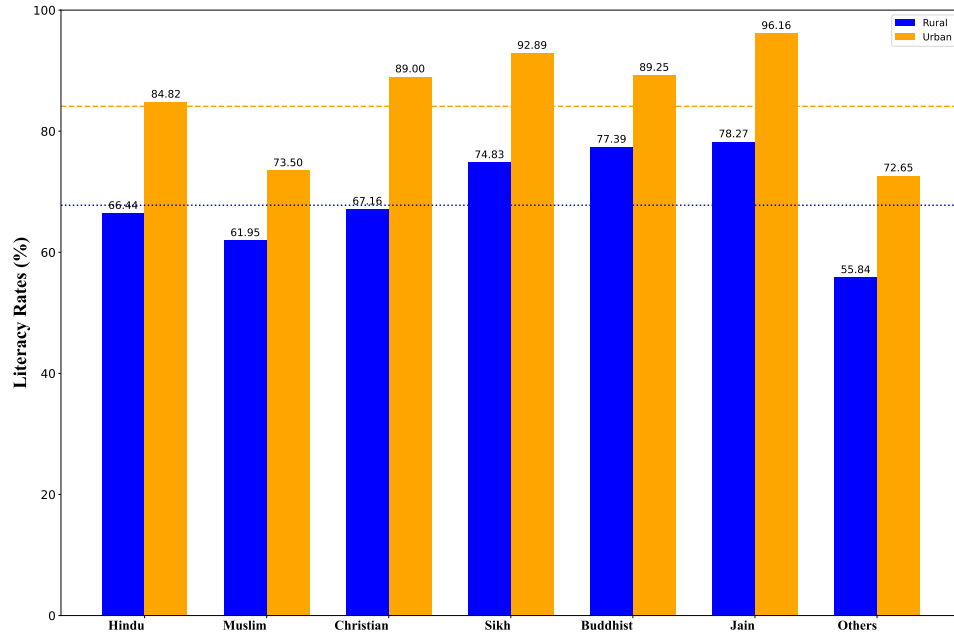Figure 4.8: *Literacy rates by Religion in Eastern Zone. The horizontal lines are as described in Fig. 4.4.*
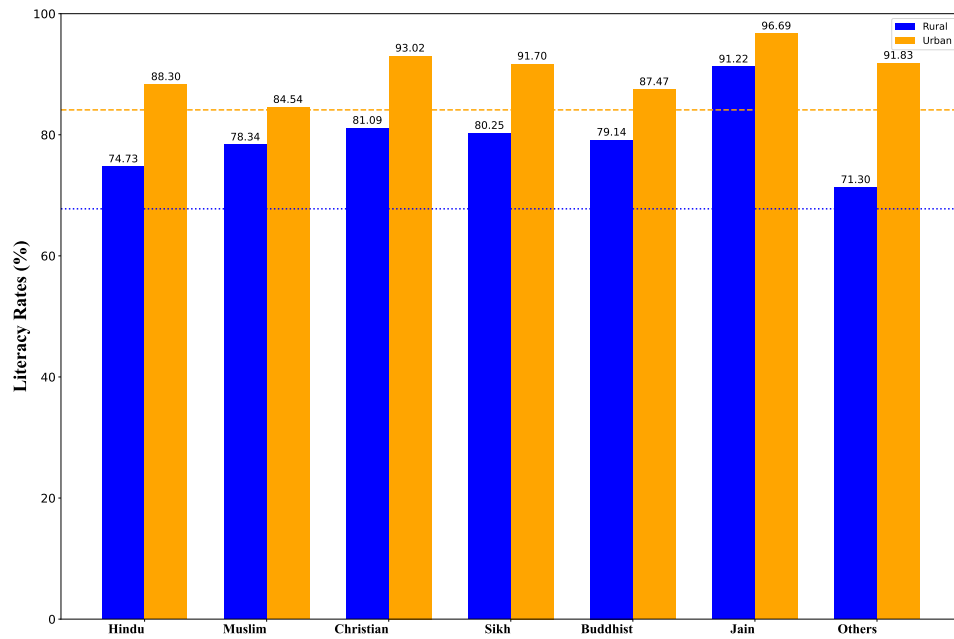


Figure 4.9: *Literacy rates by Religion in Western Zone. The horizontal lines are as described in Fig. 4.4.*

**Western Zone :** The highest literacy rate in both urban and rural areas is observed among people of the Jain religion. Conversely, the lowest rural literacy rate is found among individuals categorized under the "Others" religion, followed closely by people of the Hindu religion, which holds the second-lowest rural literacy rate. In urban areas, the lowest literacy rate is seen among people of the Muslim religion, with the Hindu religion again coming in second-lowest. When examining the gap between urban and rural literacy rates, it is most pronounced among those in the "Others" religion, at 20 %, and least pronounced among people of the Jain religion, at 5%. When comparing the rural literacy rates of different religions with the National Rural Literacy Rate, we find that all religions have literacy rates above this national benchmark. However, the scenario is different for urban literacy rates. When compared to the National Urban Literacy Rate, it is observed that all religions, except for the Muslim religion, have literacy rates above the national average.
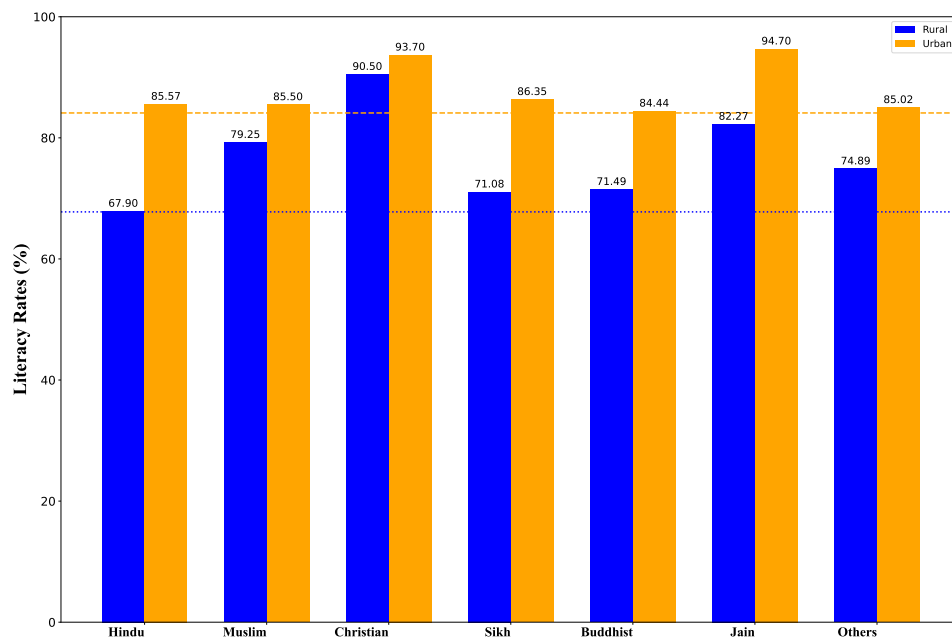


Figure 4.10: *Literacy rates by Religion in Southern Zone. The horizontal lines are as described in Fig. 4.4.*

**Southern Zone :** In this zone also the highest urban literacy rate is found among people following the Jain religion, while the highest rural literacy rate is observed among people following the Christian religion, surpassing even the National Urban Literacy Rate. On the other hand, the lowest urban literacy rate is seen among people of the Buddhist religion, and

the lowest rural literacy rate is among people of the Hindu religion. The literacy gap between the urban and rural areas is lowest for the people following the Christian religion and highest for the people in the Hindu religion standing at 3% and 17% respectively. All religions, except for Buddhism, have urban literacy rates above the National Urban Literacy Rate. In contrast, all religions, have their rural literacy rates above the National Rural Literacy Rate line.

While comparing all the zones we can see that on average people following the Jain religion tend to have the highest urban literacy rate which aligns with what we saw in figure 4.4 but for the rural literacy rate except for the North Eastern and the Southern Zone the rural literacy rate for the people following the Jain religion is still the highest.

The Literacy gap between the rural and urban areas for certain zones like the North Eastern Zone, the Eastern Zone, and the Northen Zone is quite high but comparatively zones like the Central Zone, the Southern Zone, and the Western Zone have a low literacy gap. This gap can be attributed to geographical areas, infrastructure across the zones, education institutions, education levels, etc.

## 4.4 Conclusion and prospects

Overall, these findings highlight the educational disparities not only between the regions, with higher literacy rates in urban areas compared to lower rates in rural regions but also between genders, with male literacy significantly outpacing female literacy. Additionally, when examining literacy rates across religious communities, a notable gap exists, with the Jain community exhibiting the highest literacy rates and the Muslim community the lowest emphasizing the multifaceted nature of literacy challenges in India.

Addressing the disparities in literacy rates across different regions, genders, and religions requires significant effort and concrete policy measures. While our study does highlight certain sections of society with low literacy rates it also highlights those sections that are doing increasingly well. By conducting a more in-depth examination of the successful segments, we can gain valuable insights and draw inspiration to address the challenges faced by less advantaged groups. For the development of the nation, it is very important for us that all the sections of society are doing increasingly well and the first step towards that is improvement in the literacy rate and further in the education level of the people of our nation coming from all regions, background, religion, and gender.

In Appendix A we have given the urban and rural literacy rates of the states. The tables are clubbed according to the zones. The All India Ranking in the tables refers to the rank of the state according to the total literacy rate. As can be seen from the results even within a particular zone there can be a major change in the literacy pattern and hence one can study the

correlations between the states and try to understand the reasons for the same. One can also try to study why some of the zones/states show different patterns of religion-specific literacy rates. We leave these topics for future studies.

**Report documentation:** This report has been prepared using $\LaTeX\,2_\varepsilon$ documentation system using Overleaf[1] platform.

---

[1]https://www.overleaf.com

# Appendix A

# Literacy rates in states and Union Territories

The literacy rate across states is included highlighting their Rural, Urban, and Total Literacy Rates with their All India Ranking is divided according to their respective zones. These tables are presented to facilitate an easier and more thorough understanding of the work and its outcomes.

| S.No. | States/Union Territories | Rural Literacy Rate | Urban Literacy Rate | Total Literacy Rate | All India Ranking |
|-------|--------------------------|---------------------|---------------------|---------------------|-------------------|
| 1 | Haryana | 71 | 83 | 75 | 22 |
| 2 | Himachal Pradesh | 81 | 91 | 82 | 11 |
| 3 | Jammu & Kashmir | 63 | 77 | 67 | 30 |
| 4 | Punjab | 71 | 83 | 75 | 21 |
| 5 | Rajasthan | 61 | 79 | 66 | 33 |
| 6 | NCT of Delhi | 81 | 86 | 86 | 8 |
| 7 | Chandigarh | 80 | 86 | 86 | 9 |

Table A.1: Literacy Rate of the states and UTs in the Northen Zone

| S.No. | States/Union Territories | Rural Literacy Rate | Urban Literacy Rate | Total Literacy Rate | All India Ranking |
|-------|--------------------------|---------------------|---------------------|---------------------|-------------------|
| 1 | Chhattisgarh | 65 | 84 | 70 | 27 |
| 2 | Uttarakhand | 76 | 84 | 78 | 16 |
| 3 | Uttar Pradesh | 65 | 75 | 67 | 29 |
| 4 | Madhya Pradesh | 63 | 82 | 69 | 28 |

Table A.2: Literacy Rate of the states and UTs in the Central Zone

| S.No. | States/Union Territories | Rural Literacy Rate | Urban Literacy Rate | Total Literacy Rate | All India Ranking |
|-------|--------------------------|---------------------|---------------------|---------------------|-------------------|
| 1 | Assam | 69 | 88 | 72 | 26 |
| 2 | Arunachal Pradesh | 59 | 82 | 65 | 34 |
| 3 | Manipur | 73 | 85 | 76 | 18 |
| 4 | Tripura | 84 | 93 | 87 | 5 |
| 5 | Mizoram | 84 | 97 | 91 | 3 |
| 6 | Meghalaya | 69 | 90 | 74 | 24 |
| 7 | Nagaland | 75 | 89 | 79 | 15 |
| 8 | Sikkim | 78 | 88 | 81 | 13 |

Table A.3: Literacy Rate of the states and UTs in the North Eastern Zone

| S.No. | States/Union Territories | Rural Literacy Rate | Urban Literacy Rate | Total Literacy Rate | All India Ranking |
|-------|--------------------------|---------------------|---------------------|---------------------|-------------------|
| 1 | Bihar | 60 | 77 | 62 | 35 |
| 2 | Jharkhand | 61 | 82 | 66 | 32 |
| 3 | Orissa | 70 | 86 | 73 | 25 |
| 4 | West Bengal | 72 | 85 | 76 | 19 |

Table A.4: Literacy Rate of the states and UTs in the Eastern Zone

| S.No. | States/Union Territories | Rural Literacy Rate | Urban Literacy Rate | Total Literacy Rate | All India Ranking |
|-------|--------------------------|---------------------|---------------------|---------------------|-------------------|
| 1 | Goa | 87 | 90 | 89 | 4 |
| 2 | Gujarat | 72 | 86 | 78 | 17 |
| 3 | Maharashtra | 77 | 89 | 82 | 12 |
| 4 | Daman & Diu | 81 | 89 | 87 | 6 |
| 5 | Dadra & Nagar Haveli | 64 | 90 | 76 | 20 |

Table A.5: Literacy Rate of the states and UTs in the Western Zone

| S.No. | States/Union Territories | Rural Literacy Rate | Urban Literacy Rate | Total Literacy Rate | All India Ranking |
|-------|--------------------------|---------------------|---------------------|---------------------|-------------------|
| 1 | Andhra Pradesh | 60 | 80 | 67 | 31 |
| 2 | Karnataka | 69 | 86 | 75 | 23 |
| 3 | Kerala | 93 | 95 | 94 | 1 |
| 4 | Tamil Nadu | 74 | 87 | 80 | 14 |
| 5 | Puducherry | 80 | 88 | 86 | 10 |

Table A.6: Literacy Rate of the states and UTs in the Southern Zone

# Bibliography

[Commission, 1956] Commission, S. R. (1956). Zonal councils. Available at: `https://www.mha.gov.in/sites/default/files/ZCS-CitiCharter-130710_1.pdf`.

[Das, 2012] Das, M. K. P. D. N. (2012). *Population Of India 2011.* B.R.Publishing Corporation.

[Global Education Monitoring Report Team, 2005] Global Education Monitoring Report Team, U. (2005). Education for all: literacy for life; efa global monitoring report, 2006, summary. Available at: `https://doi.org/10.54676/BQLU6071`.

[Hassan, 2020] Hassan, M. I. (2020). *Population Geography:A Systematic Exposition.* Routledge.

[Office of the Registrar General & Census Commissioner, 1891] Office of the Registrar General & Census Commissioner, I. (1891). General report of census of india. Available at: `https://censusindia.gov.in/nada/index.php/catalog/28412/download/31594/31444_1891_REP.pdf`.

[Office of the Registrar General & Census Commissioner, 1991] Office of the Registrar General & Census Commissioner, I. (1991). 1991:c-2 age sex and educational level - all areas. Available at: `https://censusindia.gov.in/nada/index.php/catalog/35544/download/39202/1991-C02T-0100.xlsx`.

[Office of the Registrar General & Census Commissioner, 2001] Office of the Registrar General & Census Commissioner, I. (2001). 2001: C-08: Educational level by age and sex for population age 7 and above (all). Available at: `https://censusindia.gov.in/nada/index.php/catalog/44754/download/48427/PC01_C08_00.xls`.

[Office of the Registrar General & Census Commissioner, 2011a] Office of the Registrar General & Census Commissioner, I. (2011a). 2011: B-01: Main workers, marginal workers, non-workers and those marginal workers, non-workers seeking/available for work classified by age and sex (total). Available at: `https://censusindia.gov.in/nada/index.php/catalog/12597/download/15710/DDW-0000B-01-Census.xls`.

[Office of the Registrar General & Census Commissioner, 2011b] Office of the Registrar General & Census Commissioner, I. (2011b). 2011: C-09: Educational level by religious community and sex for population age 7 and above, 2011. Available at: `https://censusindia.gov.in/nada/index.php/catalog/2493/download/5570/DDW-0000C-09.xlsx`.

[Office of the Registrar General & Census Commissioner, 2011c] Office of the Registrar General & Census Commissioner, I. (2011c). 2011:c-01: Population by religious community. Available at: `https://censusindia.gov.in/nada/index.php/catalog/11361/download/14474/DDW00C-01%20MDDS.XLS`.

[Office of the Registrar General & Census Commissioner, 2011d] Office of the Registrar General & Census Commissioner, I. (2011d). 2011:c-08: Educational level by age and sex for population age 7 and above (total). Available at: `https://censusindia.gov.in/nada/index.php/catalog/31608/download/34789/26583_1981_PPT.pdf`.

[Office of the Registrar General & Census Commissioner, 2011e] Office of the Registrar General & Census Commissioner, I. (2011e). A-02: Decadal variation in population 1901-2011. Available at: `https://censusindia.gov.in/nada/index.php/catalog/43333/download/47001/00%20A%202-India.xls`.

[Office of the Registrar General & Census Commissioner, 2011f] Office of the Registrar General & Census Commissioner, I. (2011f). Census of india 2011-questions 12, 13 and 14 related to literacy & education. Available at: `https://censusindia.gov.in/nada/index.php/catalog/42588/download/46225/Census%20of%20India%202011-Questions%2012%2C%2013%20And%2014%20Related%20To%20Literacy%20%20Education.ppt`.

[Office of the Registrar General & Census Commissioner, 2011g] Office of the Registrar General & Census Commissioner, I. (2011g). Census of india 2011-rural urban distribution of population. Available at: `https://censusindia.gov.in/nada/index.php/catalog/42617`.

[Office of the Registrar General & Census Commissioner, 2011h] Office of the Registrar General & Census Commissioner, I. (2011h). Details of religious community shown under 'other religions and persuasions' in main table c01 - 2011. Available at: `https://censusindia.gov.in/nada/index.php/catalog/11398/download/14511/DDW00C-01%20Appendix%20MDDS.xlsx`.

[Office of the Registrar General & Census Commissioner, 2023] Office of the Registrar General & Census Commissioner, I. (2023). *A Treatise on Indian Censuses since 1981*. Office of the Registrar General and Census Commissioner, India. Available at: `https://censusindia.gov.in/ebook/Com_Eng/mobile/index.html`.

[Padmanabha, 1971] Padmanabha, P. (1971). Provisional population totals, series-1 , india - census 1981. Available at: `https://censusindia.gov.in/nada/index.php/catalog/44790/download/48463/DDW-0000C-08.xlsx`.