

## Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression Approach

**\*Dr. Savitri Chandoliya**

### Abstract

This research aims to explore the determinants that influence life expectancy in Indian states, including gender, per-capita income, age, and area. A literature review was conducted to validate the selected indicators for Indian states. Secondary data from reliable sources were collected, and regression analysis was used to identify the key determinants affecting life expectancy in various Indian states. The findings indicate that literacy rate, access to doctors in rural areas, income, and geography are the main factors that influence life expectancy. Females have a higher life expectancy than males living in urban areas, and literacy rate and net income play a significant positive role in affecting life expectancy in both urban and rural areas across all states. The results also suggest that people living in mid and south India have a higher life expectancy than those in the northern states. The study highlights the need for policy measures to enhance life expectancy and quality of life across Indian states. Future research can examine the impact of specific diseases on life expectancy in greater detail.

**Keywords:** Life expectancy, Quality of life, Policy, Indian states, Regression modelling

### Introduction

Life expectancy is a statistical measure that represents the average number of years a person is expected to live based on their year of birth, their current age, and other demographic factors such as gender and country of residence. Life expectancy is often used as an indicator of the health and well-being of a population, and it is influenced by a wide range of factors, including access to healthcare, lifestyle behaviors, social and economic factors, environmental factors, and genetic factors. Improvements in healthcare, sanitation, and living conditions have contributed to increasing life expectancy in many countries over time. According to various empirical studies, factors such as caste, gender, income, and diseases that lead to mortality, such as heart diseases and cancer, have been found to impact life expectancy. This study aims to examine the statistical health economic factors, including the availability of hospitals, physicians, and doctors, among others, in Indian states to provide implications for health policy.

---

**Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression  
Approach**

*Dr. Savitri Chandoliya*

**Objective of study**

1. The study aims to investigate life expectancy in Indian states.
2. The study seeks to determine how various characteristics such as gender, per capita income, age, and geography impact life expectancy.

**Review of Literature**

The articles in this part that examine the research provide reliable multiple regression-derived measures of life expectancy for both rich and developing nations. With articles concentrating on Indian states, the impact of life expectancy and its connection to economic development is also depicted. A few articles describe the different factors, such as geographical patterns and a person's well-being, that influence life expectancy in diverse circumstances.

The Gini coefficients used in [1] were generated by running life expectancy regressions that were derived for each of the 50 US states and separately for men and women. The authors acquire several key conclusions using this cutting-edge method of calculating life expectancy, including an increase in female life expectancy across states that rise with pollution, poverty rate, and the number of white people. The findings applied to both men and women. The impact of life expectancy characteristics on economic growth is explored in [2]. Their research aims to calculate the impact of life expectancy on a wide range of economic factors, including GDP and population. The study's findings show that while GDP increased along with life expectancy, population growth is not matched by a corresponding rise in GDP or vice versa. Income per capita continues to decline as life expectancy rises. The authors also point out that there is no proof that an increase in life expectancy would result in higher income per capita. The authors have recommended a certain level of caution in relation to the findings since illnesses that are prevalent now could not have the same effects as those that were prevalent 40–50 years ago. Investigate the link between life expectancy and quality of life, as measured by dietary, economic, and environmental factors, in [3]. The findings were obtained using a multivariate analysis. Additional linear regression models were created to relate factors like the literacy rate, per capita health spending, and access to drinking water to life expectancy. The correlation obtained from the linear regression was 0.88. [4] examined the impact of gender and education for various racial and ethnic groups in the USA in 1970, 1980, and 1990. The results showed that there are significant racial disparities in life expectancy at lower levels of schooling. In [5], an empirical research of industrialized nations, especially the OECD nations, was undertaken using OLS regression. Their research shows that using pharmaceutical products lengthens life expectancy, particularly in middle and old age. The research also shows that lifestyle variables including intake of alcohol, cigarettes, fruits, vegetables, and butter have an impact on life expectancy rates. The significance of "Healthy life expectancy" and its indicators, which are impacted by changes in the population's mental and physical health, are discussed in [6]. Furthermore, it considers the relationship between these metrics and the success of political programs. According to [7], a community health center, smoking and poverty have statistically less significant effects on life expectancy than other characteristics

---

**Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression Approach**

*Dr. Savitri Chandoliya*

including weight, race, and physical activity.

### Research Methodology

To evaluate the validity of indicators, a thorough literature research from different nations, both developed and developing, was conducted first. Relevant variables that can be centered around Indian states were chosen. Secondary data from trusted sources and archives (for example, India Stat) were gathered for this research. Regression modeling using linear multiple regression and stepwise regression using Minitab software has been widely utilized to identify the important variables of life expectancy in different Indian states.

### Research Question

In this study, we examine the relationship between "Life Expectancy" across Indian states from 2010 to 2014 with a particular emphasis on how factors like literacy and the accessibility of public health support, such as the total number of medical professionals available in rural or urban areas of Indian states, affect "Life Expectancy".

### Data Sources

India Stat, one of the most complete databases with regard to socioeconomic statistics data, provided the data for this paper's analysis, which included information particular to India from 56 clustered locations. These statistics offer the academic and business research community with an accurate and thorough collection of socio-economic secondary statistical data about India and its states. They are also state-specific, region-specific, and sector-specific. For the convenience of the reader, the factors taken into consideration in this research have been specified in Tables 1 and 2.

Table 1. Definition of dependent variable used in study

Dependent Variable	Definition	Data source	Year
Life Expectancy	Denotes the actual expectation of life at an age $x$ viz.  The mean number of years in future, a person will live if the current trends of mortality continued	India Stat& Office of Registrar General and Census Commissioner, India	2010-14

---

## Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression Approach

Dr. Savitri Chandoliya

Table 2. Definition of key independent Variables used in study

Independent Variable	Definition	Data source	Year
Literacy Rate	The total percentage of the population of a state at a particular year of time when the person is aged seven years or above who can read and write with understanding. The denominator denotes the overall population aged seven years or more. (Due to data constraints, we have taken five or more years*)	India Stat& Office of Registrar General and Census Commissioner, India	2010-14
Per capita Income	Per capita or average income measures the mean income earned per person in a given region (here we have taken state) in a specified year. The calculation is done by dividing the area's total income by its total population.	Same as Above	2014
Primary Health Care Doctors	Primary Health Centre (PHCs), also referred to as public health centres, is a state's rural health care facilities. The doctors available in these basic units are known as PHC doctors	Same as Above	2014
Sex/Gender	Being female or male	Same as Above	2010-14
Age - Group	Number of people classified in the same age	Same as Above	2010-14
Area	Urban or Rural based on population density	Same as Above	2010-14
State	Local authorities within the territory of India based on linguism, culture etc. (20 States are considered for study**)	Same as Above	2010-14

### Assumptions

1. The underlying assumption was that there exists a linear and random relationship between the quantitative independent variables and Life Expectancy.
2. Although the intention was to gather secondary data for all 29 Indian states, only data from 20 states was collected due to certain states such as the north-eastern states, Telangana, and Goa not providing statistics for the required variables. It is assumed that these twenty states are representative of the Indian population in general.
3. Certain data, such as per capita income, literacy rate, and availability of physicians in primary health centers, were collected during a specific time period from 2010 to 2014 due to inconsistent data availability. However, it is presumed that there were minimal changes in

---

### Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression Approach

Dr. Savitri Chandoliya

social and economic conditions during that time, making the data set appropriate for these variables.

### Initial Model

The initial model aimed to establish a relationship between the dependent variable "Life Expectancy" and independent variables such as location (represented by 20 states), age group (19 age groups), and per capita income. However, during the regression analysis, the variable "per capita income" showed a high Variation Inflation factor of 29, which suggested that there may be an issue of multicollinearity between categorical variables and per capita income. This was due to the strong correlation between per capita income, age group, and the states. Therefore, per capita income was eliminated, and the regression was rerun using only categorical variables. The regression findings are presented in Appendix A, along with residual plots that show normally distributed residuals symmetrically centered around zero.

The model's high R square value in the model summary suggests that the model is highly accurate, despite the potential endogeneity between Life Expectancy and Age-Group. Although there may be a link between the two, the model's validity is demonstrated by the high R square value.

#### Model Summary

S	R-sq	R-sq (adj)	R-sq (pred)
0.694307	99.91%	99.90%	99.89%

The following model included regressing life expectancy on age and gender. This produced outcomes identical to the previous one once again, and the findings are shown in Appendix B. In contrast to the previous model, which included age group and "states" as independent variables, this model also included gender. When these two variables were regressed separately, it was found that while states did not have any statistical significance ( $p > 0.05$ ) or any R square value when regressed independently to Life Expectancy, gender did show some percentage contribution because it was statistically significant ( $p < 0.05$ ) and had a medium R square value.

### Multiple regression analysis to explore life expectancy at different stages of life, specifically at birth (0-4 years), mid-age (45-50 years), and old age (85+ years)

The findings revealed a strong relationship between life expectancy and a number of factors, including place, gender, urban/rural regions, literacy rate, and per capita income. For instance, inhabitants in various states, including as J&K, Kerala, Punjab, and Rajasthan, showed a tendency to live longer than those in Chhattisgarh, especially mid-aged females from metropolitan regions. States like Haryana, AP, HP, Punjab, Rajasthan, J&K, and Uttarakhand have favorable aging circumstances, and women tended to live longer there. According to the results of the regression study, states like Jharkhand, J&K, HP, and Bihar had a higher life expectancy at birth than other states like Tamil Nadu, Maharashtra, and Gujarat. The models' R-square values ranged from modest to high, making them

---

### Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression Approach

*Dr. Savitri Chandoliya*

useful for further research from diverse angles.

### Findings

Based on empirical research, the availability of medical professionals in rural areas, income levels, and the preferred states of residence for Indians (including their geography) are crucial factors that contribute to improving life expectancy in Indian states. Women in urban areas of most Indian states have a higher life expectancy than men. Moreover, research indicates that both literacy rate and net income have a positive impact on life expectancy at birth in both urban and rural areas across all Indian states. Additionally, the state's geography plays a significant role in determining life expectancy. People residing in the southern and central regions of India are likely to have a longer life expectancy than those living in the northern regions. The study's findings provide insights for policymakers in designing healthcare and social policies that can enhance the quality of life and longevity of individuals in Indian states. Future studies can examine other factors, such as caste, gender, income, and diseases that lead to mortality, in more detail to determine their influence on life expectancy in Indian states.

### Conclusion

The research highlights the value of life expectancy as a key indicator of public health and the need for policymakers to concentrate on fostering conditions that prolong life. The study's data shows that a number of variables, including states, gender, wealth, literacy rates, and access to physicians in remote regions, have an impact on life expectancy. Policymakers at the state and federal levels may create plans to enhance public health based on the results. The report makes recommendations for implementing best practices in urban regions in states with greater life expectancies, fostering health literacy and education, and developing a plan to improve access to physicians in rural areas. The research also emphasizes how important geographical location and state choices are in shaping life expectancy. Future studies may dive more into other elements that might affect life expectancy, such as smoking and certain illnesses.

**\*Lecturer**  
**Department of Economic**  
**Government College**  
**Tonk (Raj.)**

### Reference

1. Israelsen, A.J., & DwilghtIsraelsen, L. (2006). The distribution of life expectancy within US States. Economic research Institute Study papers, 1-13.
2. Acemoglu, D., & Johnson, S. (2007). Disease and development: the effect of life expectancy on economic growth. Journal of Political Economy, 115(6), 1-61.

---

### **Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression Approach**

*Dr. Savitri Chandoliya*

3. Crimmins, E.M., & Saito, Y. (2001). Trends in healthy life expectancy in the United States 1970-1990: gender racial and educational differences. *Social Science & Medicine*, 52(11), 1629-41.
4. Gulis, G. (2000). Life expectancy as an indicator of environmental health. *European Journal of Epidemiology*, 16(2), 161-5.
5. Shaw, J.W., Horace, W.C., & Vogel, R.J. (2005). The Determinants of Life Expectancy: An analysis of the OECD Health Data. *Southern Economic Journal*, 71(4), 768-783.
6. Robine, J.M., & Ritchie, K. (1991). Health life expectancy: evaluation of global indicator of change in population health. *The Lancet*, 302(6774), 457-460.
7. Park, J.Y. (2011). Factors that influence – life expectancy: multivariate regression analysis of US counties. *Statistical Method in Policy Analysis*, 1-26.
8. Ghosh, M. (2006). Economic growth and human development in Indian States. *Economic and Political Weekly*, 41(30), 3321-3327.
9. Kabir, M. (2008). Determinants of life expectancy in developing countries. *The journal of Developing Areas*, 41(2), 185-204.
10. Chen, M., & Ching, M. (2000). A statistical analysis of life expectancy across countries using multiple regressions, 1-41.

---

**Investigating the Main Influencers of Life Expectancy in India's States: A Multiple Regression Approach**

*Dr. Savitri Chandoliya*