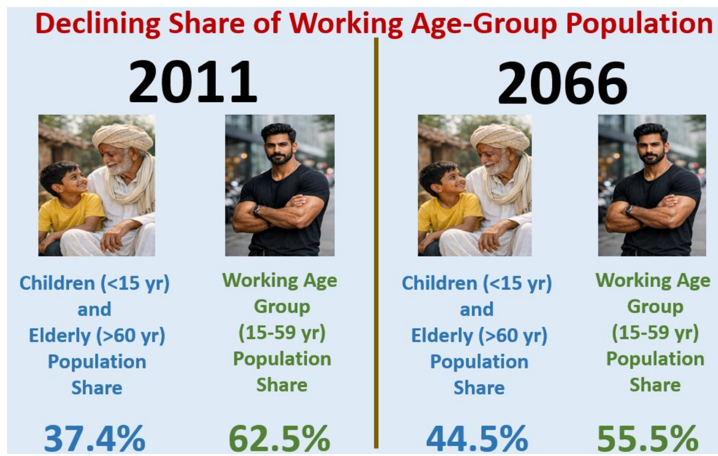
India's Working Age-group Population is Declining Rapidly

Projecting India's Population Using Markov Chain Model Till 2066

By Soumyodeep Mukhopadhyay[1] and Dripto Mukhopadhyay[2]

1. Why Predicting Population Accurately is Vital

India's population surpassed that of China in 2023, according to UN population projection. According to UN mid-year population of India in 2026 is ~142 crore. Since population census has not conducted in India since last 15 years, the actual population of the country is not known. However, demographic dividend of India has always been a topic of discussion during past last couple of decades or so. It refers to the potential for immense economic growth driven by its large number of young populations in working age group. According to 2011 Census, as per India's actual data, roughly 62% of the population was recorded in the working-age bracket (15–59 years). The median population age in 2011 was close to 28 years.



Recently, Andhra Pradesh Government has proposed monetary benefits for larger families. This includes financial incentives for third and fourth child (e.g., ₹30,000 for a third child and ₹40,000 for a fourth) and IVF treatments. The reason cited in various sources suggests that the current trend of population growth could lead to significant drop of population in working age group. This would lead to severely straining economic productivity by 2047. In view of this, we would like to share our recent work, which suggests a possible decline in population at the country level in late 2060s/2070 onward. Since Census of India's population projection is available till 2031, our prediction focused on the period 2036

[1] Soumyodeep Mukhopadhyay is a doctoral student in Applied Mathematics, Stellantis & École des Mines Saint-Étienne, France.

[2] Dripto Mukhopadhyay is the Founder Director of Ascension Centre for Research and Analytics (ACRA)

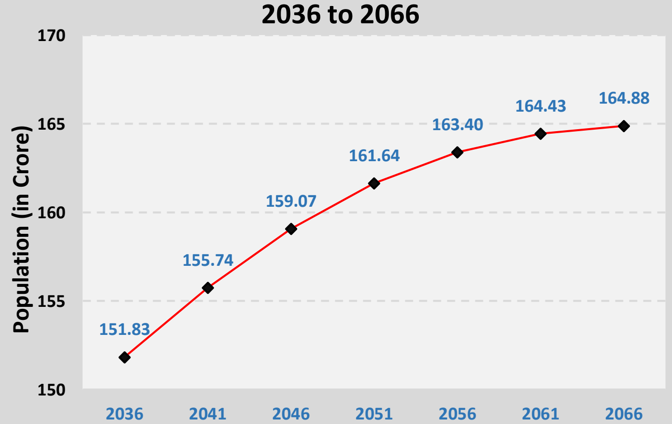
onward. The projections suggest that India's population in working age-group would be reduced to around 55.5% compared to 62.6% in 2011.

Debates continue regarding impacts of India's huge population pressure, both positive and negative. A segment of opinions reveal that this huge population boosts the economy, but the rest consider this population pressure as reasons for unemployment, crumbled infrastructure including basic amenities and the similar ones. While the debate on whether India's huge population is a boon or a burden for the country's sustainable development path may continue, it is important to visualise the future population trajectory in a scientific manner. Any planning, especially socio-economic, requires accurate predictions to produce desired results. We, at ACRA, projected India's population till 2066 using **Markov Chain model.** This is the most recommended model for predicting population globally. This article presents a few crucial trends regarding India's population, as obtained from our study, pertinent to India's economic development.

1. Key Takeaways from Our Estimations

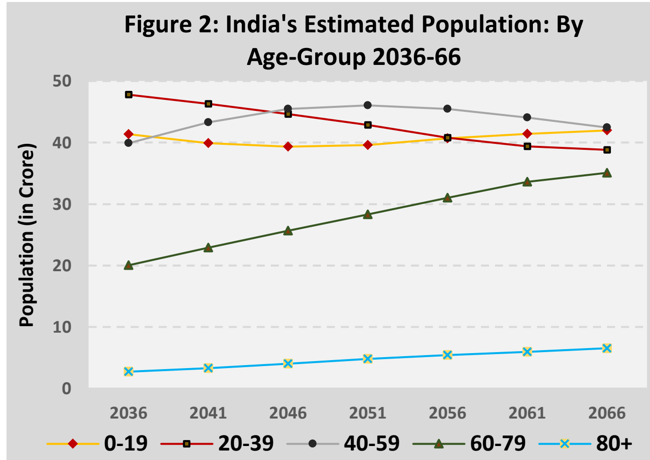
- India's population would stop increasing beyond 2066. As seen from the graph below **(Figure 1)**, the growth rate is expected to reach near zero beyond 2056. The estimated population in 2061 and 2066 is more or less similar. This declining growth trajectory reveals that the population would start declining within a few years beyond 2066.

Figure 1: India's Estimated Total Population 2036 to 2066

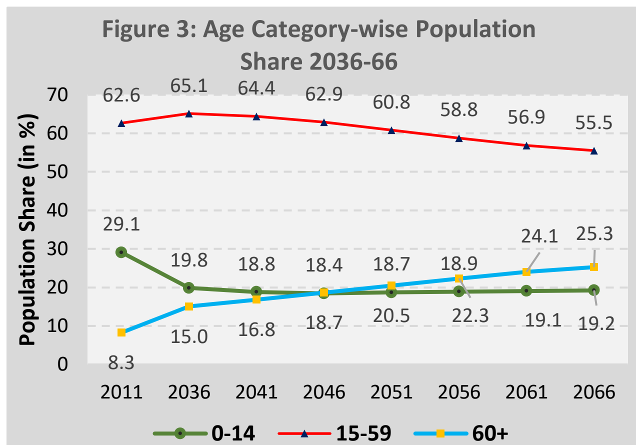


- India's population stops growing beyond 2066. As seen from the graph below **(Figure 2)**, the age category-wise population growth is expected to decline significantly for working age-category. Though our original estimation is executed at a 5-year interval for male and female, for the benefit of the readers that graph is presented in a broader age-category of 20 years. It clearly depicts that the population in younger age-categories have shown significant decline over time.

- Implications of the above are shown in the graph below (**Figure 3**). For an easier comparison, 2011 population data is also shown in this visual. It clearly shows that the population share of the working age-category (15 to 59 years) increased from 62.6% in 2011 to 65.1% in 2036. Since 2036 it started declining and by 2066 it would reduce to 55.5%. the share of children and elderly population (non-working) would reach close to 45%.



- This implies that the proportion of the dependent population would increase significantly. With the current unemployment trend, this would be a critical challenge for the country. At present, a significant share of the population live in below Population projection, especially by age-group and gender, is not simple task since it needs to recon large number of parameters poverty line and survive on government support on multiple fronts. The predictions in this article calls for a long-term demographic and economic development policy for the country to feed such huge dependent population.



- Since any government support for free food and other schemes poses significant fiscal challenge and debt burden on the governments, delays in proactive and meaningful policy formulation may lead to an unrecoverable poverty trap by 2070.

2. Methodology – Markov Chain Prediction Model Explained

We used Python programming to develop Markov Chain model. Population projection, especially by age-group and gender, is not simple task since it requires to consider large number of parameters impacting population. Markov chains have been used in ecology as a discrete-time model to quantify dynamics of an age-structured, population distribution. The model involves constructing a Leslie matrix composed of state-transition probabilities depending only on the current state and treating age groups as states of the Markov chain. However, a key difference is that the Leslie matrix does not satisfy the unit row-sum property (right stochastic matrix), violating the law of total probability. It is more accurate to interpret the matrix entries as survival coefficients rather than as true probabilities. These coefficients are a means of representing the following set of linear equations:

$$v_i(t+1) = \rho_{i-1} \; v_{i-1}(t); \quad \text{for } i = 1, \ldots, N \; , \tag{1}$$

$$v_0(t+1) = \sum_{i=0}^{N} f_i \; v_i(t) \; , \tag{2}$$

where v represents the population vector at time t[1], i is the ith component of v denotes the age-group from a set of N +1 possible age groups. The coefficients ρ_i, f_i are the survival coefficient and the fertility rate of the i^{th} age group, respectively. Equations (1) and (2) can be written more compactly as,

$$v(t+1) = \begin{bmatrix} f_0 & f_1 & \ldots & f_a \\ p_0 & 0 & \ldots & 0 \\ 0 & p_1 & 0 & 0 \\ . & . & . & . \\ 0 & \ldots & p_{N-1} & 0 \end{bmatrix} v(t) \tag{3}$$

The matrix in eq. (3) is the Leslie matrix, L. Given the initial population distribution at time t = 0, v(0) = v0, we can compute the population distribution after T time steps as,

$$v(T) = L^T v_0 \tag{4}$$

This is a direct consequence of eq. (3). We exploit this model to project population dynamics of the Indian population over a period of 30 years from the latest government projections Government census data: Population Projections for India and States 2011 - 2036. Leslie matrix for the Indian population prediction has been investigated in the past, most notably in [1]. We adopt a similar framework, differing only in the computation of newborns, i.e. age group i = 0 (0-4 years). Traditionally, the Leslie matrix computation is applied only to the female population which is why you have well-defined fertility rates for each age group. Handling a mixed-sex population is more complicated, in this aspect. Research studies [1] used fertility rates from 2011 census data to calculate the Leslie matrix as usual. In contrast to the traditional approach, we use the projected crude birth rates and sex ratio for 2036 to enable predictions of the mixed population vector. The modified set of linear equations are as following:

$$v_0^f(t+1) = CBR \cdot \left(\frac{SR}{SR+1000}\right) \cdot \sum_{i=0}^{N} v_i^f(t) + v_i^m(t) \ , \tag{5}$$

$$v_0^m(t+1) = CBR \cdot \left(\frac{1000}{SR+1000}\right) \cdot \sum_{i=0}^{N} v_i^f(t) + v_i^m(t) \ , \tag{6}$$

$$v_i^f(t+1) = \rho_{i-1}^f v_{i-1}^f(t); \quad \text{for } i = 1, \ldots, N \ , \tag{7}$$

$$v_i^m(t+1) = \rho_{i-1}^m v_{i-1}^m(t); \quad \text{for } i = 1, \ldots, N \ , \tag{8}$$

where v^f and v^m denote the female and male population vectors, CBR is the crude birth rate and SR is the sex ratio. Writing equations (5)-(8) as a matrix equation,

$$\begin{bmatrix} v^f(t+1) \\ v^m(t+1) \end{bmatrix} = \begin{bmatrix} L_f & \mathbf{0} \\ \mathbf{0} & L_m \end{bmatrix} \begin{bmatrix} v^f(t) \\ v^m(t) \end{bmatrix} \ , \tag{9}$$

$$L_f\,[ij] \coloneqq \begin{cases} 1 & \text{if } \ i = 0 \ \ \text{(first row)} \\ \rho_{i-1}^f & \text{if } \ j = i-1 \ \ \text{(sub-diagonal)} \\ 0 & \text{otherwise} \end{cases} \ , \tag{10}$$

$$L_m\,[ij] \coloneqq \begin{cases} 1 & \text{if } \ i = 0 \ \ \text{(first row)} \\ \rho_{i-1}^m & \text{if } \ j = i-1 \ \ \text{(sub-diagonal)} \\ 0 & \text{otherwise} \end{cases} \ , \tag{11}$$

$$i, j = 1, \ldots, N \ ,$$

where the block-diagonal matrix in eq. (9) is of size, 2N × 2N, expressed in terms of two matrices L_f and L_m of size N × N each, as defined in equations (10) and (11).

To estimate the survival coefficients, we used the change in projected data from 2031-35, Assuming constancy of survival coefficients, crude birth rate (CBR = 13 × step-size) and sex ratio (SR = 909) over the projection period, we calculate six time-steps of this model with step-size = 5 years from 2036 to 2066.

$$\hat{\rho}_{i-1}^f \coloneqq \frac{v_i^f(t+1)}{v_{i-1}^f(t)} \tag{12}$$

$$\hat{\rho}_{i-1}^m \coloneqq \frac{v_i^m(t+1)}{v_{i-1}^m(t)} \tag{13}$$

Suggested Readings

1. Vikas Kumar, Vanshika Jain, Garima Tomar, N.S. Chauhan, Vinod Chandra, and Shankar Lal. Application of leslie matrix model in human population dynamics and public health. South Eastern European Journal of Public Health, pages 444–459, 12 2024.

2. P. H. Leslie. On the use of matrices in certain population mathematics. Biometrika, 33(3):183–212, 1945.

3. J. R. Norris. Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.