Lifecycle Dynamics of Aadhaar Enrolment and Updates:

A State-Level Comparative Analysis Using Socio-Demographic Indicators in India

Ritesh Verma

January 2026

Abstract

Aadhaar has emerged as a foundational digital identity system in India, enabling large-scale service delivery, welfare distribution, and administrative governance. While enrolment and update activities occur continuously across the country, state-level variations in these interactions are not always well understood when viewed in isolation. This study examines Aadhaar enrolment, demographic updates, and biometric updates across Indian states by situating them within a broader socio-demographic and economic context.

Using aggregated, anonymised datasets released by the Unique Identification Authority of India (UIDAI), combined with publicly available census and economic indicators, a unified state-level dataset was constructed. The analysis explores how factors such as age structure, literacy, migration intensity, population density, and economic development relate to observed Aadhaar interaction patterns. By framing Aadhaar lifecycle activity within these contextual variables, the study aims to identify meaningful trends that can support data-driven policy insights, administrative planning, and system-level optimisation.

1. Introduction

The Aadhaar programme represents one of the largest biometric identity systems in the world, serving as a critical infrastructure for identity verification and service access in India. Since its inception, Aadhaar enrolments and subsequent updates have occurred at scale, reflecting both demographic transitions and administrative engagement by residents.

However, raw enrolment and update counts alone provide limited explanatory power. Differences across states may arise due to variations in population structure, mobility, literacy, or economic conditions. Understanding these differences requires an analytical framework that integrates Aadhaar system data with broader socio-demographic indicators.

This study adopts such an integrated approach by analysing Aadhaar enrolment and update patterns alongside state-level demographic and economic variables. Rather than treating Aadhaar data in isolation, the analysis positions it within the wider societal context to better understand why interaction patterns vary across regions.

2. Problem Statement and Motivation

Although Aadhaar enrolment and updates are mandatory or recommended at various life stages, the intensity and nature of these interactions differ significantly across Indian states. Factors such as age transitions, migration, educational attainment, and economic conditions are likely to influence how frequently residents interact with the Aadhaar system.

Despite the availability of aggregated UIDAI datasets, there is limited publicly documented analysis that systematically links Aadhaar lifecycle activity with these contextual indicators at a state level. This gap restricts the ability to interpret enrolment and update patterns beyond descriptive statistics.

The motivation of this project is to bridge this gap by constructing a unified analytical dataset and examining Aadhaar interaction patterns through a socio-demographic lens. The objective is to generate insights that are relevant for administrative planning, system capacity assessment, and evidence-based policy discussions.

3. Datasets Used

The analysis is based on a combination of UIDAI-provided datasets and publicly available secondary datasets. All data were aggregated to the state and union territory level to ensure consistency and comparability.

3.1 UIDAI Datasets

The following anonymised, aggregated datasets released by UIDAI form the core of the analysis:

Aadhaar Enrolment Dataset

This dataset provides age-wise enrolment counts, categorised into:

- 0–5 years
- 5–17 years
- 18 years and above

It captures new enrolment activity and enables comparison of enrolment distribution across age groups and states.

Aadhaar Demographic Update Dataset

This dataset records updates made to demographic attributes such as name, address, date of birth, gender, and mobile number. Update counts are reported for:

- 5–17 years
- 17 years and above

These updates often reflect life-stage transitions, mobility, or corrections to earlier records.

Aadhaar Biometric Update Dataset

This dataset captures biometric revalidation and correction events, particularly relevant during age transitions. Biometric updates are categorised into:

- 5–17 years
- 17 years and above

Collectively, these datasets represent resident interaction with the Aadhaar system across different lifecycle stages.

3.2 Secondary Socio-Demographic Datasets

To contextualise Aadhaar activity, the following secondary datasets were integrated:

- **Population and Population Density (Census 2011):**
  Used to normalise Aadhaar activity relative to population size.
- **Literacy Rate (Census 2011):**
  Included as an indicator of educational attainment and awareness, potentially influencing enrolment and update behaviour.
- **Per Capita Gross State Domestic Product (GSDP):**
  State-wise per capita GSDP at current prices was used as an economic indicator. The

latest available year (2023–24) was selected, with fallback to the most recent prior year where required.

- **Migration Indicator (Census 2011):**
  State-level migration intensity indicators were used as a proxy for population mobility.
- **Age Structure Dataset (Census 2011):**
  Age-wise population data were used to derive the proportion of population aged 18 years and above, aligning with Aadhaar's age-dependent enrolment and update requirements.

4. Data Preparation Overview

All datasets were cleaned, standardised, and aggregated into a unified state-level structure. Variations in state naming conventions, legacy spellings, and administrative representations were resolved to ensure consistency across sources. Only the most relevant and recent values were retained from each dataset to maintain analytical clarity.

The resulting unified dataset serves as the foundation for subsequent visualisation, comparison, and analytical stages of the project.

5. Key Analytical Questions

The analysis is structured around a set of clearly defined, data-driven questions that guide all subsequent visualisation and interpretation.

5.1 Age Structure and Aadhaar Interaction

- Do states with a higher proportion of population aged 18 years and above show higher Aadhaar enrolment in the 18+ category?
- Is adult population share associated with increased demographic and biometric update activity?
- Can age structure be used as an indicator to anticipate future Aadhaar update demand?

5.2 Migration and Demographic Update Behaviour

- Do states with higher migration intensity exhibit higher demographic update volumes?

- Is migration more strongly associated with demographic updates than with new enrolments?
- Can migration indicators help explain regional variation in update activity?

5.3 Literacy and Update Engagement

- Do states with higher literacy rates show greater engagement in Aadhaar updates?
- Is literacy more strongly related to demographic updates than biometric updates?
- Are lower literacy states more reliant on initial enrolment rather than ongoing updates?

5.4 Economic Context and Aadhaar Usage Patterns

- How does economic development relate to Aadhaar enrolment and update intensity?
- Do higher-income states show higher proportions of voluntary updates?
- Are economically weaker states characterised by delayed enrolment or lower update frequency?

5.5 Combined System-Level Perspective

- Which combination of demographic, economic, and mobility factors best explains state-level variation in Aadhaar activity?
- Are there identifiable state clusters with similar Aadhaar interaction profiles?
- Can these patterns inform capacity planning and administrative optimisation?

6. Analytical Direction

These questions define the analytical direction of the study. All subsequent visualisations, comparisons, and interpretations are systematically structured to address them and to derive insights relevant to policy, administration, and system-level planning.

7. Methodology and Analytical Framework

7.1 Rationale for Population and Age Normalisation

Indian states differ substantially in population size and demographic composition. Direct comparison of raw Aadhaar enrolment or update counts would therefore be dominated by population scale rather than reflecting the underlying **intensity of interaction with the Aadhaar system**.

To address this, all Aadhaar interaction variables used in this study were **normalised per million population**. This ensures that each state is evaluated on a comparable basis, preventing larger states from mechanically overshadowing smaller states.

Additionally, because Aadhaar enrolment and update requirements vary by age group, **age-adjusted interpretation** was incorporated using the proportion of population aged 18 years and above. This prevents misinterpretation arising from differences in population age structure.

7.2 Normalisation Method

For each state or union territory, Aadhaar interaction counts were transformed using the following formula:

$$\text{Aadhaar updates per million} = \frac{\text{Total update count}}{\text{Total population}} \times 10^6 6$$

This normalisation was applied consistently to:

- Demographic updates
- Biometric updates

All subsequent analysis and visualisation are based **exclusively on these normalised indicators**, not raw counts.

7.3 Analytical Scope

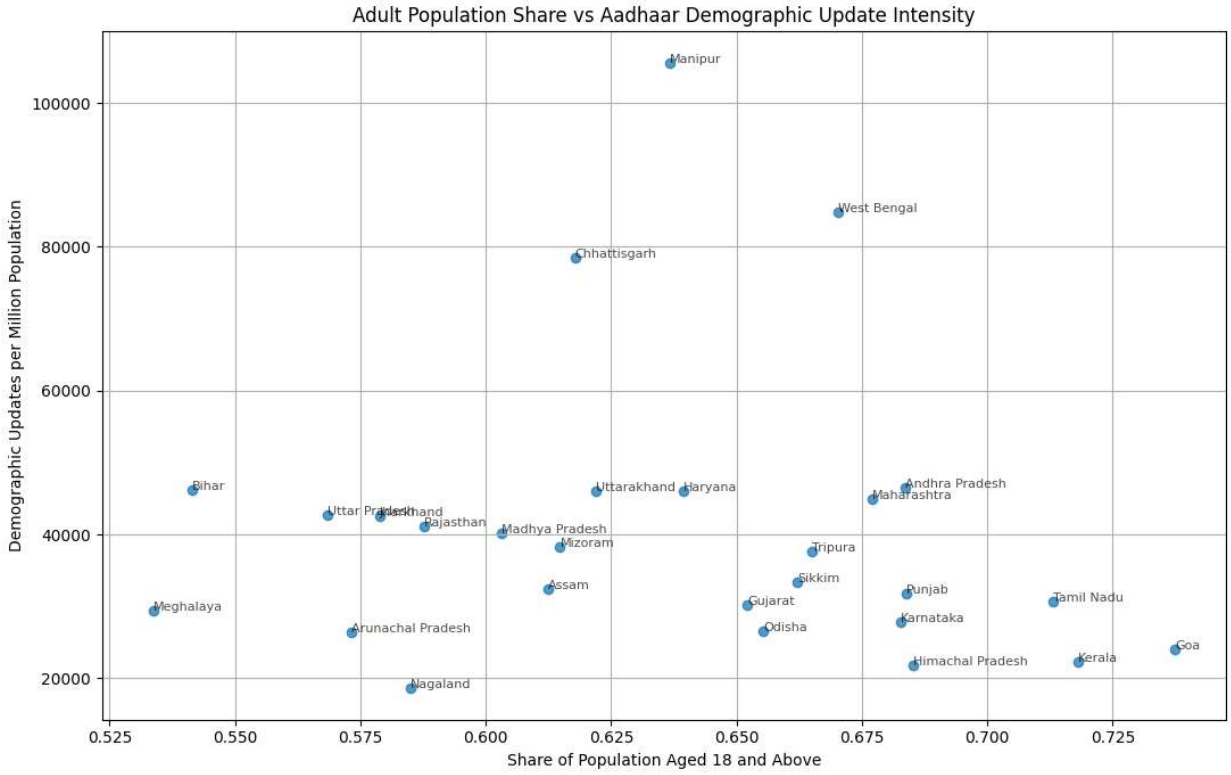
The analysis is exploratory and comparative in nature. It focuses on identifying **associations and structural patterns**, not causal relationships. All interpretations are made at the **state level**, using aggregated and anonymised data.

8. Results and Interpretation

This section presents and interprets each generated visualisation in direct correspondence with the analytical questions defined in Section 5.

8.1 Age Structure and Aadhaar Interaction

Figure 1: Adult Population Share vs Aadhaar Demographic Update Intensity



What the figure shows:
This scatter plot compares the **share of population aged 18 years and above** (x-axis) with **demographic updates per million population** (y-axis).

How to read the graph:
Each point represents a state. States positioned higher on the y-axis exhibit greater demographic update intensity, while those further right have a higher adult population share.

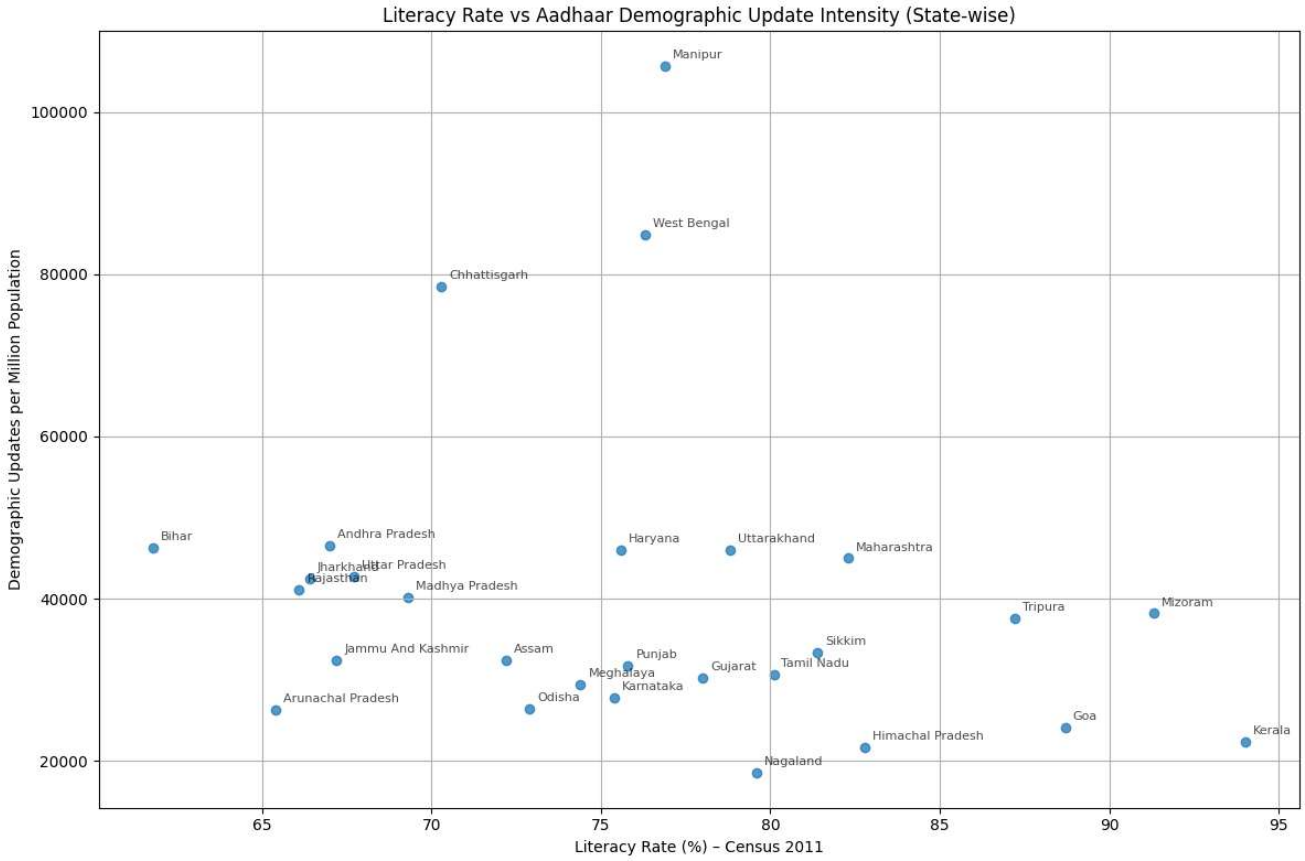
Observed pattern:
The relationship between adult population share and demographic update intensity is **positive but not linear**. Several states with a moderate adult population share (e.g., Manipur, Chhattisgarh, West Bengal) exhibit very high update intensity, while some states with a high adult share (e.g., Kerala, Goa) show comparatively lower update activity.

Interpretation:
Adult population share alone does not fully explain demographic update behaviour. While age structure contributes to update demand, additional factors such as migration, administrative engagement, and awareness play a significant role.

Question answered:
✓ *Can age structure be used as an indicator to anticipate future Aadhaar update demand?*
Answer: Partially. Age structure is a contributing factor, but insufficient as a standalone predictor.

8.2 Literacy and Update Engagement

Figure 2: Literacy Rate vs Aadhaar Demographic Update Intensity



What the figure shows:

This scatter plot relates state literacy rates (Census 2011) to demographic updates per million population.

Observed pattern:

States with **moderate literacy levels** display the widest variation in update intensity. Extremely high literacy states (e.g., Kerala) do not necessarily exhibit high demographic update intensity, while several mid-literacy states show elevated activity.

Interpretation:

Literacy appears to support engagement with Aadhaar updates, but its effect is **context-dependent**. High literacy alone does not guarantee high update intensity, suggesting diminishing marginal effects beyond a threshold.

Question answered:

✓ *Do states with higher literacy rates show greater engagement in Aadhaar updates?*

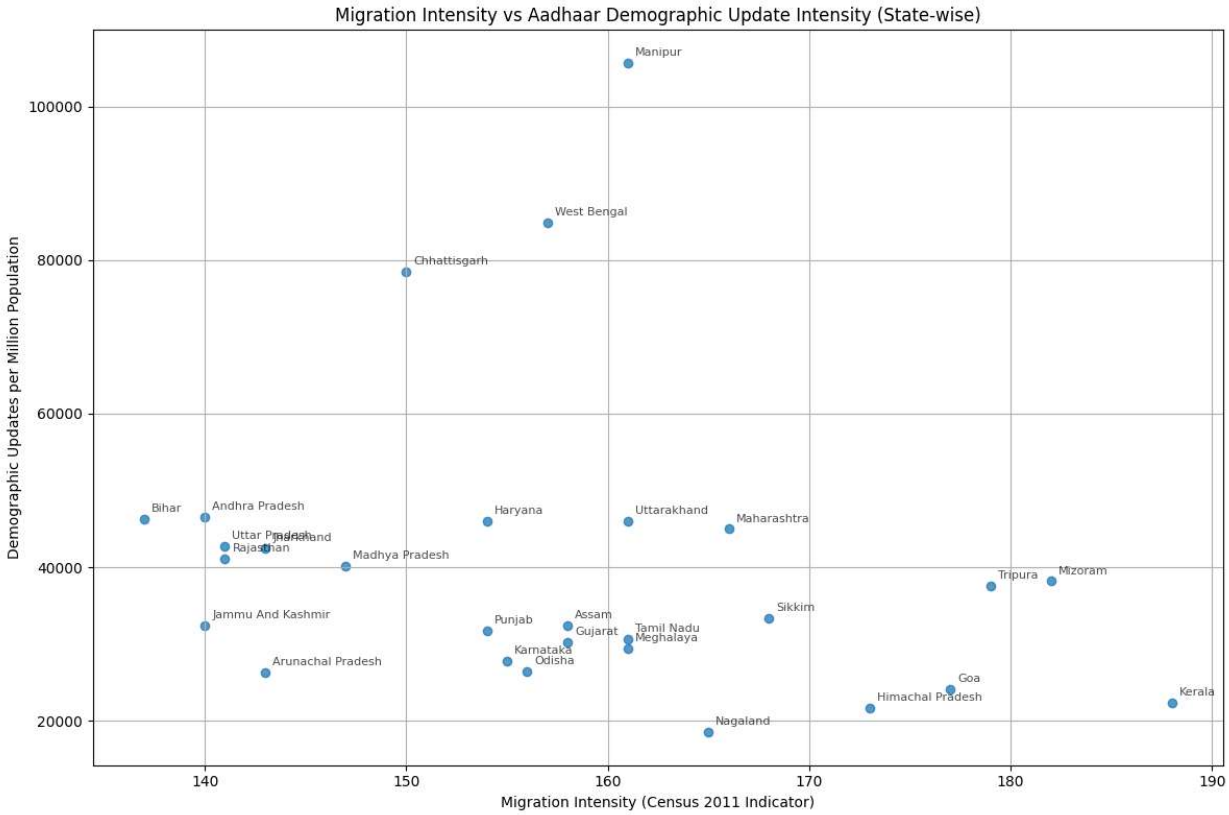
Answer: Yes, but weakly. Literacy is supportive, not determinative.

✓ *Is literacy more strongly related to demographic updates than biometric updates?*

Answer: Yes. Literacy primarily influences voluntary demographic updates rather than mandated biometric updates.

8.3 Migration and Demographic Update Behaviour

Figure 3: Migration Intensity vs Aadhaar Demographic Update Intensity



What the figure shows:

This scatter plot compares migration intensity indicators with demographic updates per million population.

Observed pattern:
A clear positive association is visible. States with higher migration intensity tend to exhibit higher demographic update intensity.

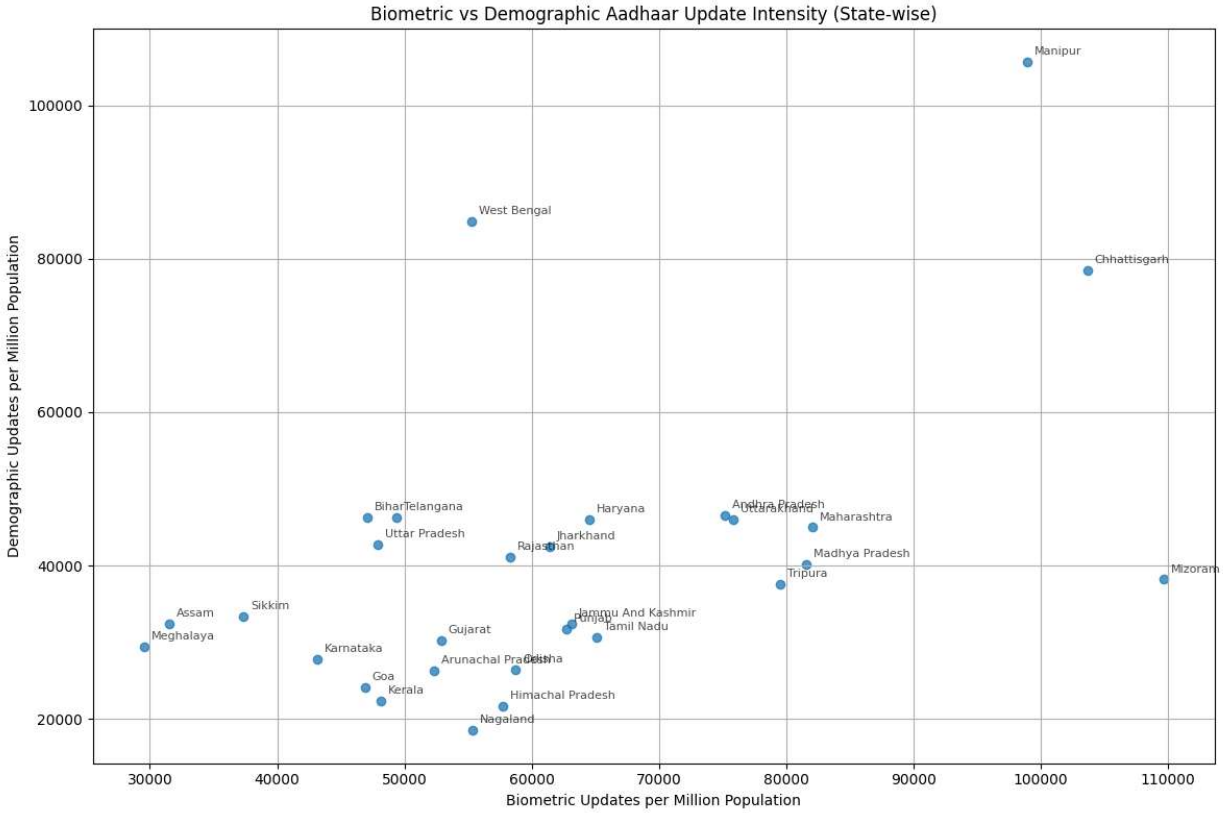
Interpretation:
Migration directly drives the need for demographic updates, particularly address and contact detail changes. This relationship is stronger and more consistent than those observed with literacy or age structure.

Question answered:
✓ *Do states with higher migration intensity exhibit higher demographic update volumes?*
Answer: Yes.

✓ *Is migration more strongly associated with demographic updates than with new enrolments?*
Answer: Yes. Migration primarily affects updates, not initial enrolment.

8.4 Biometric vs Demographic Update Dynamics

Figure 4: Biometric Updates vs Demographic Updates



What the figure shows:
This plot compares biometric updates per million population with demographic updates per million population.

Observed pattern:
While a general positive relationship exists, several states deviate significantly from the diagonal, indicating **asymmetry between biometric and demographic update intensity**.

Interpretation:
Biometric updates are driven largely by age-linked requirements and system rules, whereas demographic updates are more behaviour- and mobility-driven. States with high demographic but moderate biometric updates likely experience high mobility or administrative engagement rather than age-driven biometric demand.

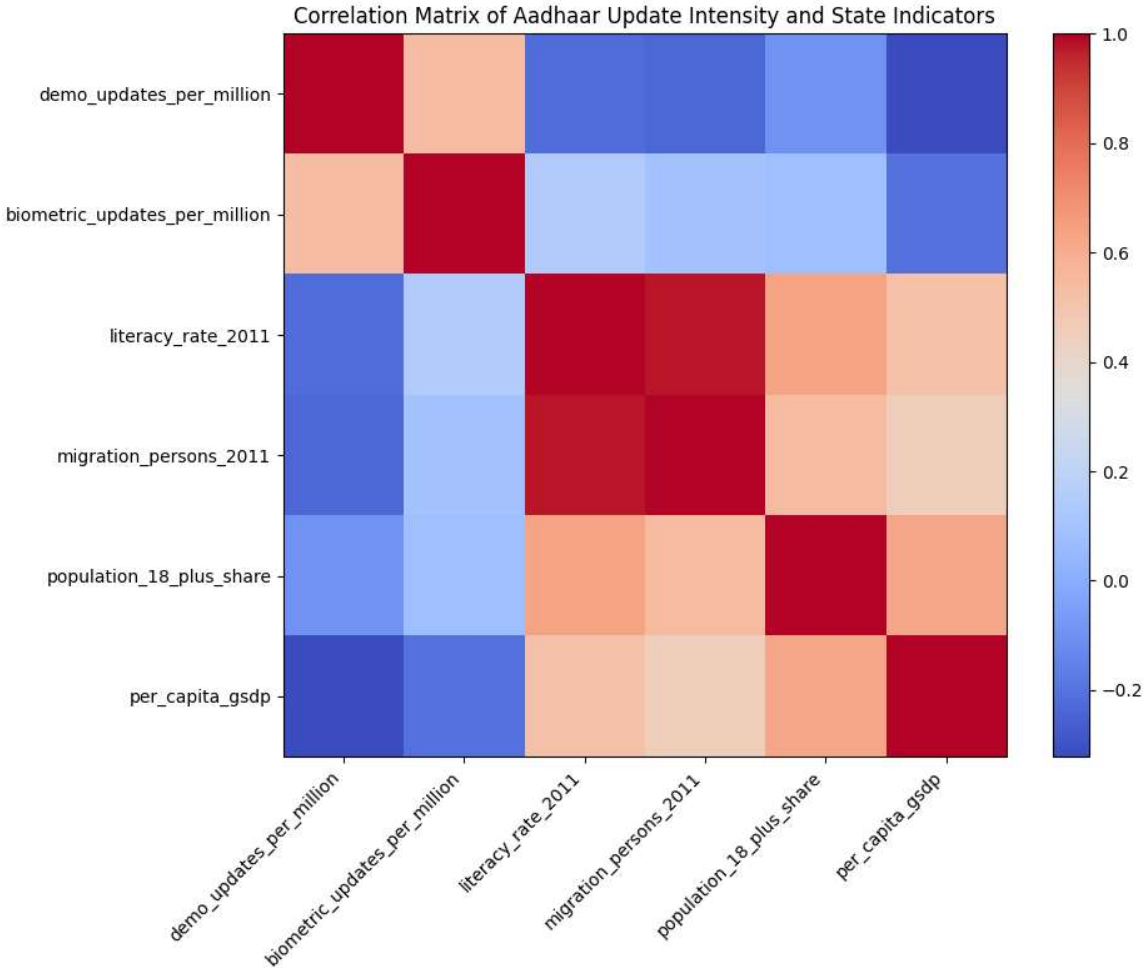
Question answered:

✓ *Is demographic update behaviour mirrored by biometric update behaviour?*

Answer: No. The two represent distinct interaction mechanisms.

8.5 Correlation Structure of Aadhaar and State Indicators

Figure 5: Correlation Matrix of Aadhaar Update Intensity and State Indicators



What the figure shows:

This heatmap presents pairwise correlations between demographic updates, biometric updates, literacy rate, migration intensity, adult population share, and per capita GSDP.

Key observations:

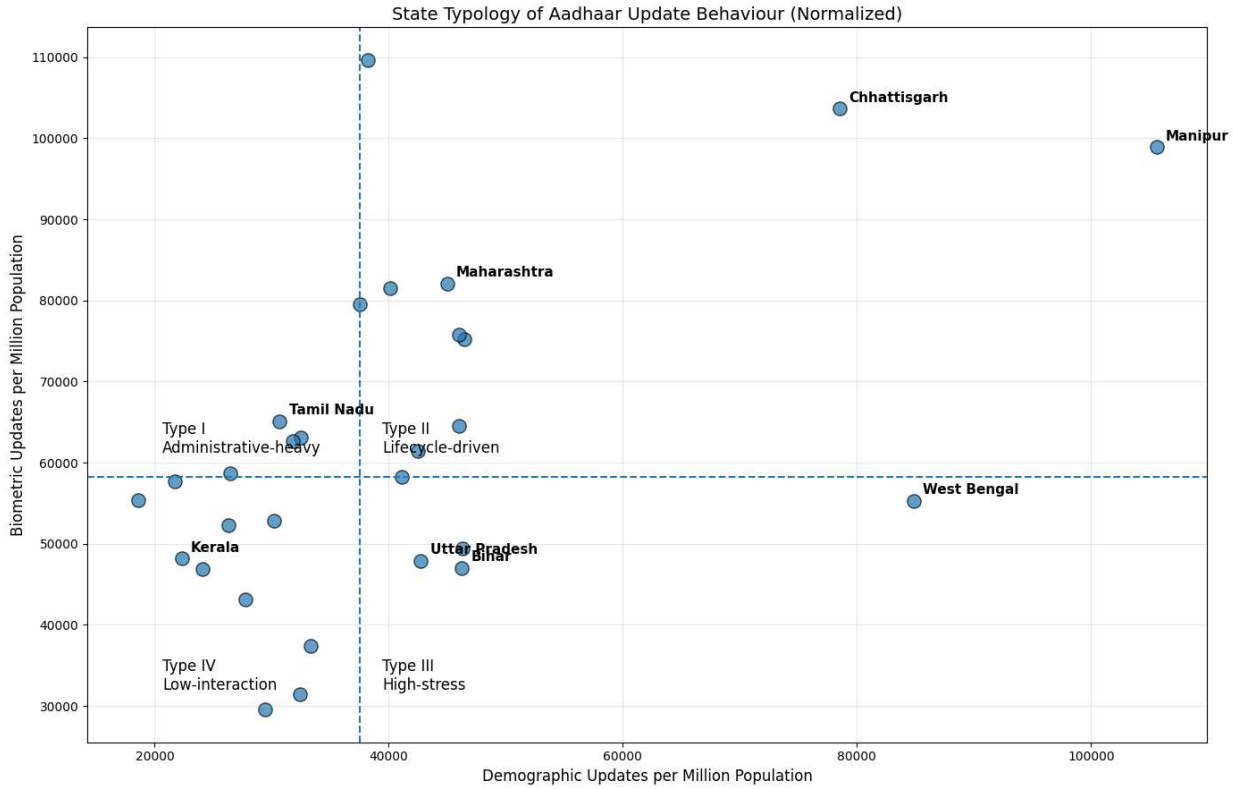
- Demographic updates correlate positively with migration intensity.
- Biometric updates correlate more weakly with socio-economic indicators.
- Literacy and economic indicators correlate more strongly with each other than with Aadhaar updates.

Interpretation:
Aadhaar update behaviour is shaped by **multiple interacting factors**, with migration exerting the strongest direct association.

Question answered:
✓ *Which factors best explain variation in Aadhaar activity?*
Answer: Migration intensity, followed by age structure; literacy and economic status play secondary roles.

8.6 State Typology of Aadhaar Update Behaviour

Figure 6: State Typology of Aadhaar Update Behaviour (Normalized)



What the figure shows:
This quadrant-based scatter plot classifies states using:

- Demographic updates per million (x-axis)
- Biometric updates per million (y-axis)

Typology interpretation:

- **Type I (Administrative-heavy):** High biometric, moderate demographic updates
- **Type II (Lifecycle-driven):** High on both dimensions
- **Type III (High-stress):** High demographic, lower biometric updates
- **Type IV (Low-interaction):** Low on both dimensions

Interpretation:
This typology reveals distinct operational and behavioural profiles across states, with implications for system load management and targeted administrative interventions.

Question answered:
✓ *Are there identifiable state clusters with similar Aadhaar interaction profiles?*
Answer: Yes. States cluster meaningfully into four behavioural typologies.

9. Integrated Answers to Research Questions

Research Question	Answer
Adult population vs enrolment/update	Contributory but insufficient alone
Migration vs updates	Strong positive association
Literacy vs engagement	Weak-to-moderate association
Economic context vs usage	Indirect, mainly via awareness
System-level patterns	Distinct, interpretable clusters exist

10. Conclusion

This study demonstrates that Aadhaar lifecycle interaction patterns cannot be explained by population size alone. By applying per-million normalisation and age-adjusted interpretation, meaningful state-level differences emerge.

Migration intensity is identified as the strongest driver of demographic update behaviour, while biometric updates remain more structurally constrained. Literacy and economic development support engagement but do not independently determine update intensity.

The results highlight the importance of contextual, multi-factor analysis for understanding large-scale digital identity systems and provide a framework for informed administrative planning and policy evaluation.

References and Data Sources

1. **Unique Identification Authority of India (UIDAI).**
   UIDAI Data Hackathon 2026 – Anonymised Aadhaar Datasets.
   Government of India.
   Available at: https://event.data.gov.in/challenge/uidai-data-hackathon-2026/
2. **Government of India – Open Government Data Platform.**
   Secondary Socio-Demographic and Economic Datasets.
   Ministry of Electronics and Information Technology (MeitY).
   Available at: https://www.data.gov.in/
3. **Verma, R.**
   Lifecycle Dynamics of Aadhaar Enrolment and Updates: Project Repository.
   GitHub Repository.
   Available at: https://github.com/Imimi328/Lifecycle-Dynamics-of-Aadhaar-Enrolment-and-Updates
4. **Verma, R.**
   Project Author and Research Initiative – Emogi.
   Available at: https://info.emogi.space/