



Decoding Digital Identity: A Data-Driven Analysis of Aadhaar Enrolment & Update Trends

Theme: Predictive Governance: Utilizing Enrolment Patterns to Anticipate and Manage Systemic Update Pressure

Event: UIDAI Data Hackathon 2026

A Submission for the UIDAI Data Hackathon 2026 Focus Area: Aadhaar Lifecycle Insights & Operational Resource Planning

Prepared by: Team SEMICOLON; M.Sc. Data Science | MGM University

Team ID: UIDAI_672 **Date:** 20 January 2026

Team Details:

- **Team Name:** Semicolon;
- **Team Lead:** Tejas Salunkhe | tejassalunkhe501@gmail.com
Role -Creating Dashboard
- **Member 2:** Mrudula Kamdi | kamdimrudula2003@gmail.com
Role – Cleaning Dataset

Tools & Technologies used:

- Microsoft Power BI
- Power Query
- DAX (Data Analysis Expressions)

1. Executive Summary:

India's Aadhaar ecosystem has entered a post-saturation phase, where the primary challenge is no longer mass enrolment but sustained maintenance, updates, and data accuracy. As Aadhaar becomes deeply embedded in welfare delivery and authentication systems, understanding who enters the system, who updates it, and where data corrections are concentrated becomes critical for proactive governance.

This project introduces a Lifecycle-based analytical framework to study Aadhaar data across three interconnected stages: Enrolment (Entry), Biometric Updates (Maintenance), and Demographic Updates (Accuracy). Using aggregated UIDAI public data and Power BI, we analyze age-wise participation, regional disparities, and update intensity patterns.

The dashboards reveal a strong adult-driven trend across enrolment and updates, indicating migration, workforce mobility, and address changes as key drivers of Aadhaar service demand. Child enrolment and youth updates, while present, show uneven geographic coverage, highlighting awareness and access gaps.

By transforming raw statistics into interactive dashboards and intensity indicators, this analysis supports a shift from reactive service response to anticipatory, data-driven decision-making. The resulting insights enable targeted policy interventions such as mobile Aadhaar units, early-age enrolment campaigns, and region-specific infrastructure planning.

2. Problem Statement & Objectives:

Problem Statement

- ✓ While Aadhaar enrolment coverage in India has largely matured, maintenance and data accuracy challenges are increasing. Biometric and demographic updates are unevenly distributed across age groups and regions, creating localized pressure on Aadhaar infrastructure and service delivery mechanisms.
- ✓ The absence of a lifecycle-oriented analytical view limits the ability of policymakers to anticipate future demand and optimize resources.

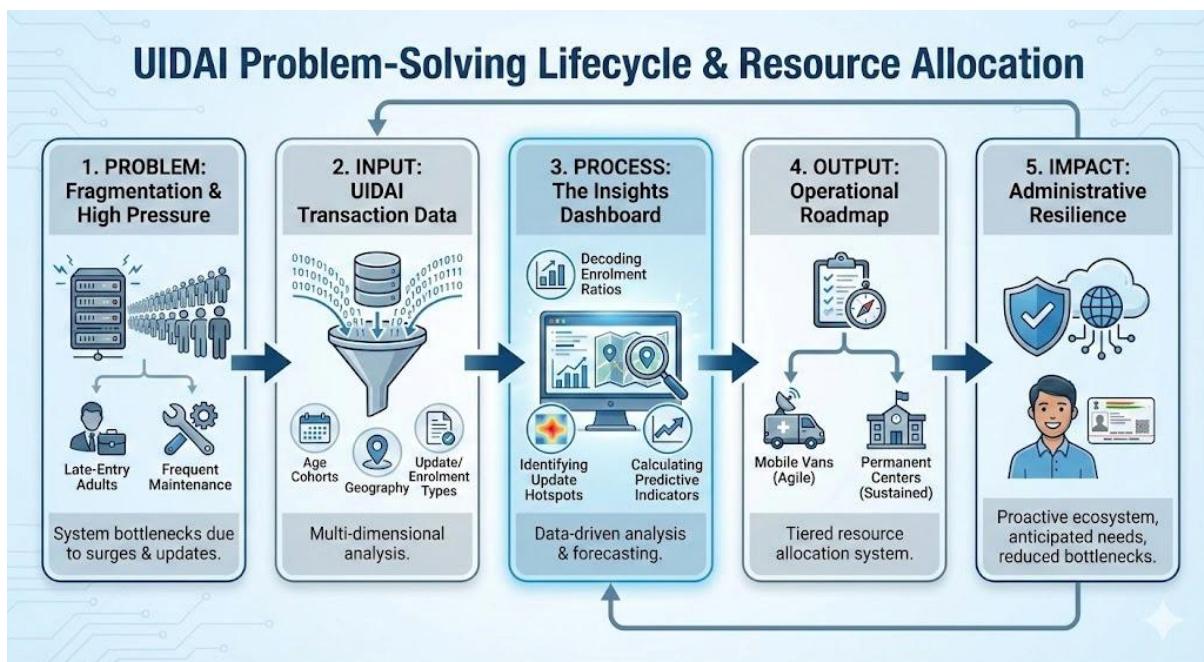
Objectives

- Identify age-wise and regional enrolment patterns.
 - Quantify biometric and demographic update pressure.
 - Detect dominant age groups driving Aadhaar maintenance.
 - Develop analytical indicators to support proactive governance decisions.
-

3. Datasets Used

1. Data Source: UIDAI Public Aadhaar Statistics (Aggregated, anonymized)
2. Datasets Utilized

- Aadhaar Enrolment Data
 - Biometric Update Data
 - Demographic Update Data
3. Key Dimensions/Key Data Components
- Geographical identifiers: State, District
 - Temporal fields: Date / Month
 - Operational metrics: Enrolment count, update count, and related Aadhaar service metrics
 - Age Group (Child, Youth, Adult)
4. Newly Added Columns for better data Understanding:
- Biometric Updates
 - Total Enrolments
 - Demographic Updates
5. Lifecycle Segmentation
- Entry Stage → Enrolment data
 - Maintenance Stage → Biometric updates
 - Accuracy Stage → Demographic updates



Column Mapping

4. Methodology

1. Data Processing Pipeline (Power BI)
2. Data ingestion using Power Query
3. Standardization of state and district names
4. Handling missing and inconsistent values
5. Age-group binning (Child, Youth, Adult)
6. Measure creation using DAX

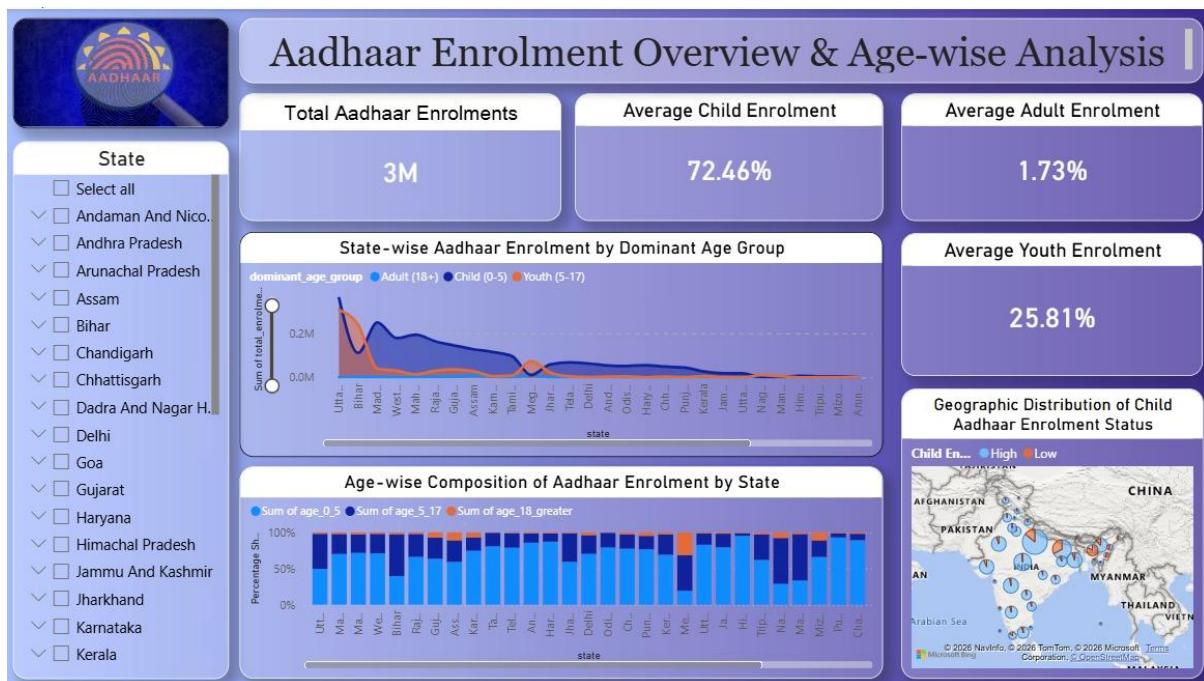
7. Interactive dashboard design and validation
8. Analytical Logic
9. Ratio-based indicators to normalize population differences
10. Dominant age group identification through numeric comparison
11. Filter-independent KPIs for national-level insights

All transformations and calculations were implemented entirely within Power BI, without the use of external programming languages.

5. Dashboard 1

1. Enrolment Insights (Entry Stage)

- i. Purpose: To analyze who is entering the Aadhaar ecosystem and identify enrolment gaps.
- ii. Key Insights.
 - Adult enrolment dominates across most states
 - Child enrolment shows regional disparities
 - Certain districts require targeted awareness efforts
- iii. Visuals Used
 - ✓ KPI Cards (Total & Average Enrolments)
 - ✓ State-wise Bar Charts
 - ✓ Geographic Distribution Map
 - ✓ 100% Stacked Age-wise Composition Chart

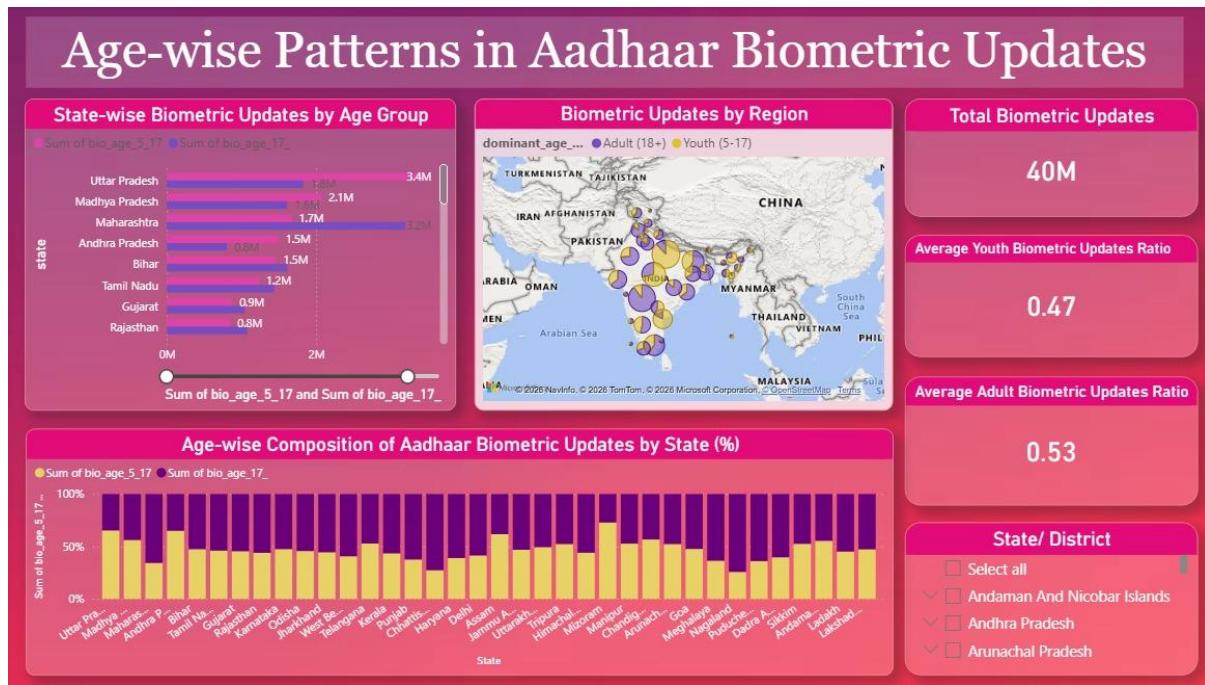


Enrolment DashBoard

6. Dashboard 2

Biometric Update Pressure (Maintenance Stage)

- i. Purpose: To assess service load and update demand on Aadhaar infrastructure.
- ii. Key Insights:
 - Biometric updates are predominantly adult-driven
 - Migration-heavy regions show higher update frequency
 - Update pressure varies significantly across states
- iii. Visuals Used:
 - ✓ KPI Cards
 - ✓ State-wise Biometric Update Bar Charts
 - ✓ Regional Bubble Map
 - ✓ Age-wise Composition Chart



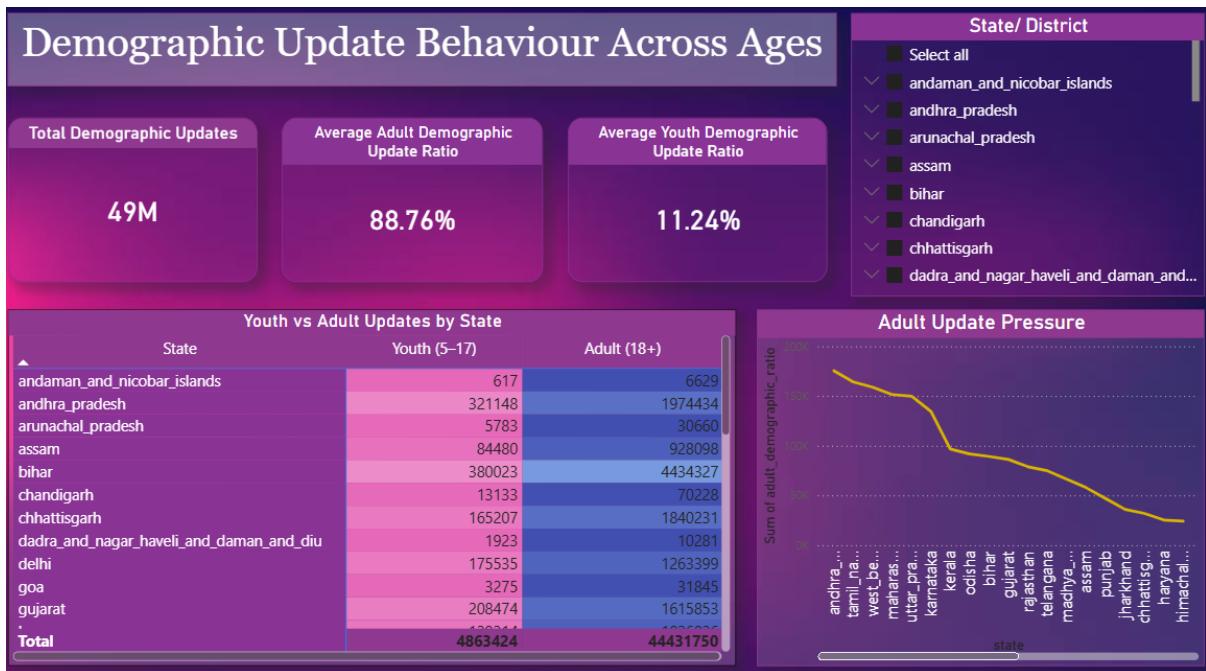
Biometric Dashboard

7. Dashboard 3

Demographic Update Behaviour (Accuracy Stage)

- i. Purpose: To evaluate data correction demand and awareness gaps.
- ii. Key Insights:
 - Demographic updates are heavily adult-dominated
 - Certain regions consistently show high correction intensity
 - Youth demographic updates remain comparatively low
- iii. Visuals Used

- ✓ Heatmap (Demographic Update Intensity by State & Age)
- ✓ Matrix Comparison (Youth vs Adult Updates)
- ✓ Trend-based Indicators



Demographic Dashboard

8. Technical Implementation

1. Platform Used: Microsoft Power BI
2. Core Components
 - 1) Power Query for data cleaning and transformation
 - 2) DAX Measures for:
 - Ratio calculations
 - Dominant age group logic
 - Filter-independent
 - KPIs Conditional formatting for heatmaps
 - Interactive slicers for regional analysis
 - 3) No external tools or scripting languages were used.

9. Creativity & Innovation

Lifecycle Dashboard Framework

This project introduces a Lifecycle-based Aadhaar Analytics Model, aligning data insights with governance stages:

- Entry (Coverage)
- Maintenance (Service Load)
- Accuracy (Data Quality)
- Innovation Highlights
- Lifecycle segmentation of UIDAI data
- Update intensity visualization
- Age-dominance indicators for decision support

This approach transforms static statistics into actionable governance intelligence.

10. Impact & Applicability

Policy Applications

- ✓ Deployment of mobile Aadhaar service units
- ✓ Targeted child and youth enrolment drives
- ✓ Region-specific infrastructure planning
- ✓ Improved data accuracy and authentication reliability

Impact

- ✓ Reduced operational bottlenecks
 - ✓ Proactive service planning
 - ✓ Enhanced inclusivity and data integrity
-

11. Conclusion & References

Conclusion:

The Aadhaar Lifecycle Analytics Dashboard demonstrates how aggregated public data can be leveraged to support evidence-based identity governance. By shifting focus from enrolment counts to lifecycle behavior, policymakers can better anticipate demand and optimize service delivery.

Future Scope:

- Integration with Census or migration datasets
- Time-series forecasting of update demand
- District-level vulnerability indices

References:

1. UIDAI Public Data Portal
2. Microsoft Power BI Documentation

