UIDAI Data Hackathon 2026

Data-Driven Innovation on Aadhaar

Project Title:

Unlocking Societal Trends in Aadhaar Enrolment and Updates

1. Introduction

The objective of this project is to analyze anonymized Aadhaar enrolment and update datasets provided by UIDAI to identify meaningful patterns, trends, and anomalies. The insights derived from this analysis aim to support data-driven decision-making and system improvements for Aadhaar enrolment processes.

2. Datasets Used

The following anonymized datasets were used in this analysis:

• Aadhaar Enrolment Dataset
• Aadhaar Demographic Update Dataset
• Aadhaar Biometric Update Dataset

Each dataset contains records aggregated at the state, district, and pincode levels along with age-group-wise distributions.

3. Data Cleaning and Preprocessing

Prior to analysis, extensive data cleaning and preprocessing were performed to ensure data quality and consistency. The following steps were undertaken:

• Removal of duplicate records
• Standardization of state and district names
• Handling of missing and inconsistent values
• Verification of age-group distributions
• Date formatting and feature extraction (year and month)

4. National Level Insights
   4.1 Monthly Enrolment Trends (2025)

Analysis of the monthly Aadhaar enrolment data for the year 2025 reveals noticeable peaks and dips across different months. These variations suggest the presence of seasonal patterns, potentially influenced by factors such as school admission periods, awareness campaigns, and administrative drives.
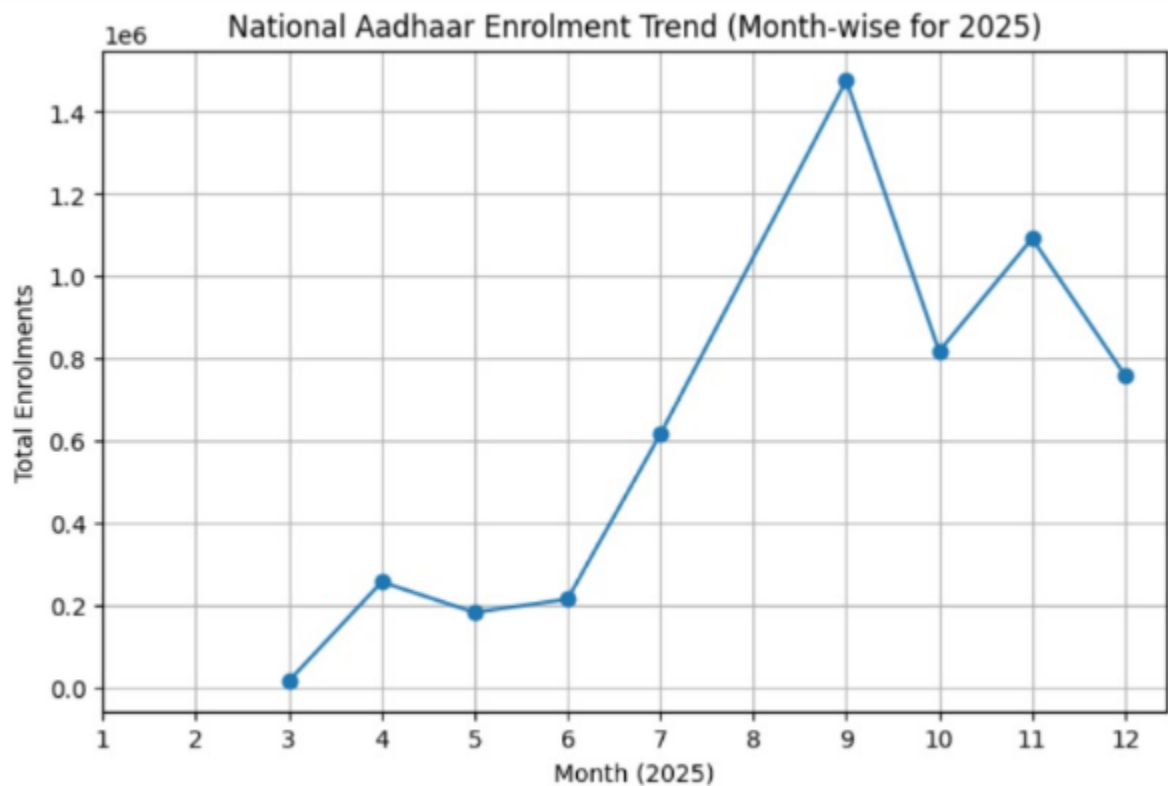

Figure 1: Monthly Aadhaar Enrolment Trend (2025)

4.2 Age-Group Distribution

The age-group-wise analysis indicates that the adult population (18 years and above) accounts for the majority of Aadhaar enrolments. In contrast, enrolments in the 0–5 age group remain comparatively low, highlighting a gap in early-age registration.

This trend suggests the need for focused initiatives to promote newborn and early childhood Aadhaar enrolment.
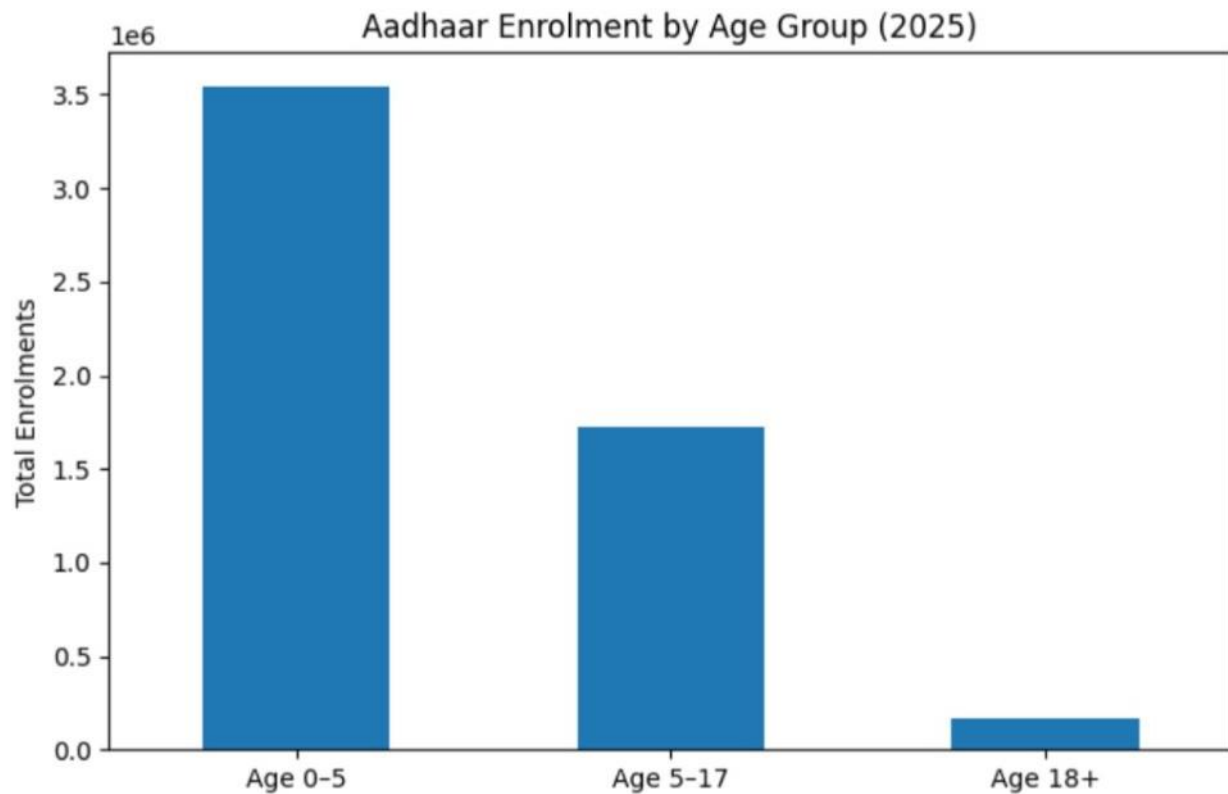
Figure 2: Age-Group-wise Aadhaar Enrolment Distribution

5. State Level Insights
5.1 Top and Bottom Performing States

State-wise analysis reveals significant disparities in Aadhaar enrolment volumes. States such as Karnataka and Maharashtra consistently record high enrolment numbers, reflecting effective administrative processes and higher population density.Conversely, smaller Union Territories and states such as Lakshadweep exhibit relatively low enrolment figures, which may be attributed to geographical constraints, smaller population sizes, or limited enrolment infrastructure.
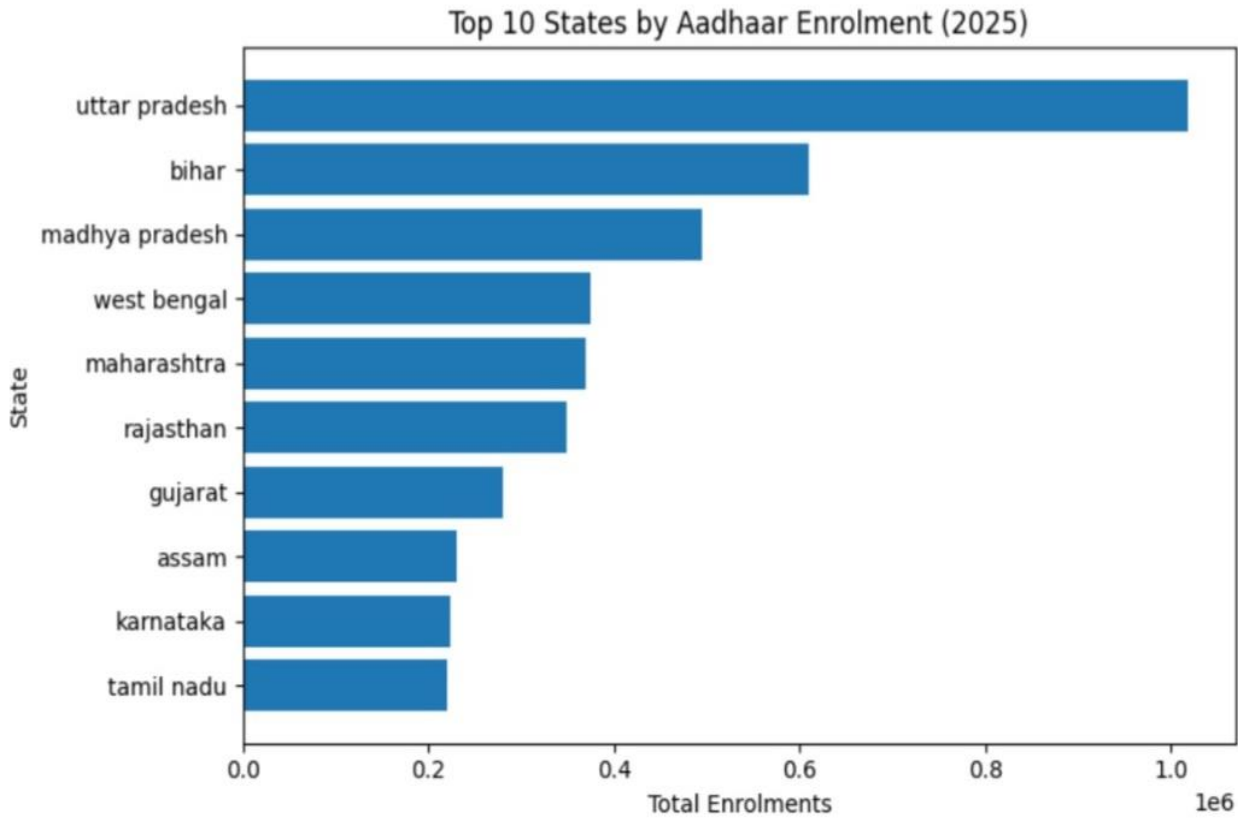
Figure 3: State-wise Aadhaar Enrolment Comparison

5.2 Age-Group Distribution Across States

Age-group analysis at the state level shows varying enrolment patterns across regions. Certain states demonstrate strong adult enrolment but lag in younger age groups, particularly in the 0–5 and 5–17 categories.

This insight enables the identification of under-enrolled age groups within specific states, facilitating targeted outreach efforts.
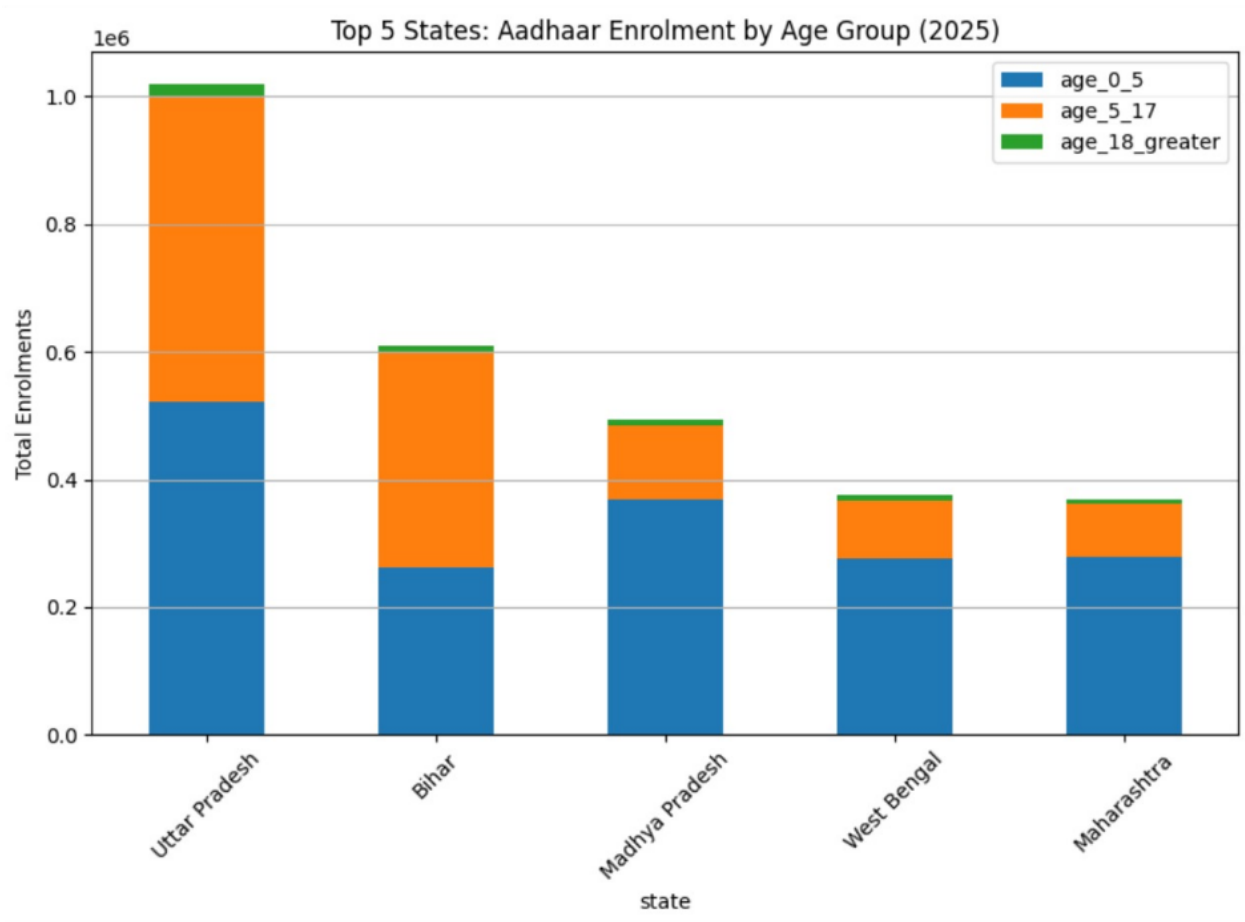
Figure 4: Age-Group-wise Enrolment by State

6. District Level Insights

6.1 District-wise Enrolment Distribution

District-level analysis highlights significant variation in enrolment volumes. Urban and densely populated districts emerge as high-enrolment hubs, while several rural or remote districts show extremely low enrolment counts.
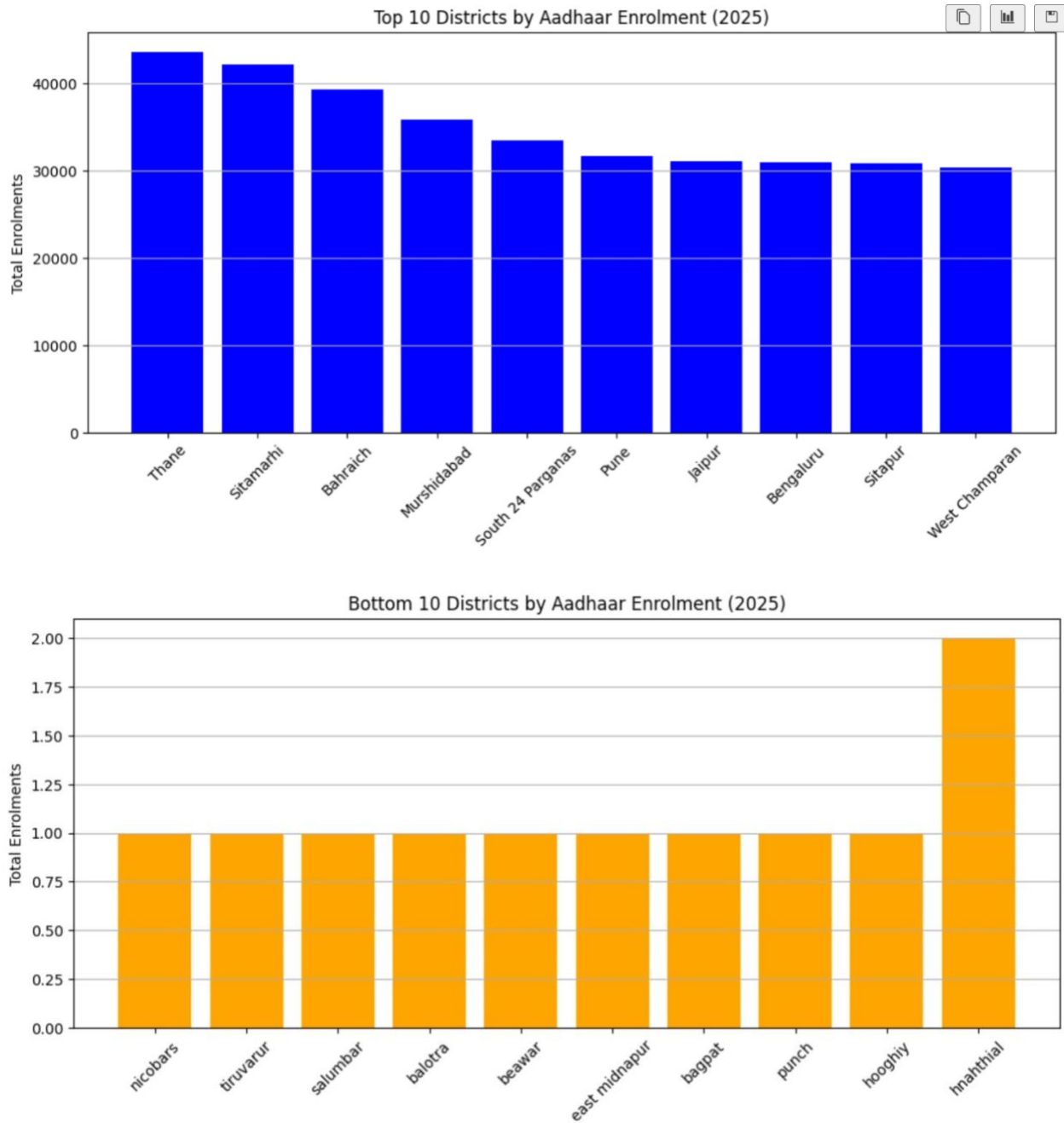
Figure 5: District-wise Aadhaar Enrolment Distribution

6.2 Anomalies and Irregularities

	biometric	enrolment	ratio
district			
serchhip	7312	102	71.686275
mamit	10241	171	59.888889
wardha	99008	1953	50.695341
gadchiroli	111394	2374	46.922494
mansa	50238	1124	44.695730
bhandara	83475	1903	43.864950
imphal east	48528	1132	42.869258
ratnagiri	109311	2603	41.994237
thoubal	44526	1067	41.730084
north and middle andaman	5382	132	40.772727

The biometric-to-enrolment ratio analysis reveals unusually high values in select districts. Such anomalies may result from delayed enrolment data reporting, migration-related updates, or repeated biometric corrections. These findings emphasize the importance of continuous data validation and district-level audits to maintain system accuracy.

Several anomalies were identified at the district level, including:

• Districts with zero or near-zero enrolment values
 • Unusually high enrolment figures in select districts
 • Duplicate or inconsistent district naming

These irregularities may indicate data reporting issues or operational gaps and warrant further verification.

7. Key Findings

• Aadhaar enrolment exhibits seasonal monthly trends
 • Adult population dominates total enrolment

- Significant enrolment disparities exist across states and districts
- Certain age groups and regions remain under-represented
- Data inconsistencies impact reporting accuracy

8. Conclusion and Recommendations

This study demonstrates the potential of data-driven analysis in uncovering enrolment patterns, demographic gaps, and operational anomalies within the Aadhaar ecosystem. The insights derived can assist UIDAI and policymakers in optimizing enrolment strategies, enhancing outreach for under-served age groups, and improving data quality through targeted interventions and monitoring mechanisms.