UIDAI DATA HACKATHON 2026

Project Title:
Aadhaar Enrolment Analysis: National Overview

Team ID: UIDAI_8704

Prepared By:
Pruthviraj V Mariguddi

Submission Type:
Government Data Hackathon Project

Dataset:
Aadhaar Enrolment Dataset (Government of India)

INDEX

SECTION 1: PROBLEM STATEMENT & APPROACH

1.1 Problem Statement

The Aadhaar program is a foundational digital identity system designed to ensure inclusive access to government services across India. Despite widespread adoption, variations in enrolment coverage exist across states, districts, and demographic groups—particularly among children.

The challenge addressed in this project is to analyze Aadhaar enrolment data to identify geographic and demographic disparities, assess child versus adult enrolment trends, and highlight regions that require targeted enrolment interventions. A data-driven, visual approach is required to support effective policy planning and decision-making.

1.2 Approach

To address the problem, a business intelligence–driven analytical approach was adopted. The project uses interactive dashboards to analyze Aadhaar enrolment patterns across multiple dimensions such as geography and age groups.

The approach includes:

- Aggregating enrolment data at national, state, district, and PIN levels
- Creating age-wise enrolment metrics
- Visualizing patterns using maps, charts, and KPIs
- Deriving actionable insights for policy and operational planning

SECTION 2: DATASETS USED

2.1 Dataset Overview

The dataset used in this project is the **Aadhaar Enrolment Dataset** provided by the Government of India. It contains aggregated enrolment data across geographic and demographic levels.

This dataset is suitable for large-scale analysis as it is structured, clean, and policy-oriented.

2.2 Columns Used in the Analysis

Column Name	Description	Purpose
date	Enrolment snapshot date	Contextual reference
state	State / UT name	Geographic analysis
district	District name	Intra-state analysis
pincode	Postal code	Micro-level coverage
age_0_5	Enrolments for age 0–5	Child coverage
age_5_17	Enrolments for age 5–17	School-age coverage
age_18_greater	Enrolments for age 18+	Adult population

2.3 Why This Dataset Was Chosen

This dataset was chosen because:

- It provides **age-wise demographic segmentation**
- It enables **multi-level geographic analysis**
- It supports **policy-relevant insights**
- It is clean and ready for BI visualization

The dataset aligns well with the objective of analyzing Aadhaar coverage and identifying enrolment gaps.

SECTION 3: METHODOLOGY

3.1 Data Cleaning

The dataset underwent a quality assessment to ensure accuracy and consistency.

Key observations:

- No missing values
- No duplicate records
- Valid numeric values for enrolment counts
- Consistent geographic naming

As a result, no major data cleaning operations were required.

3.2 Data Pre-Processing

Pre-processing steps included:

- Validating data types (numeric and text)
- Setting correct geographic data categories in Power BI
- Merging multiple CSV files into a single dataset
- Creating relationships for filtering and slicing

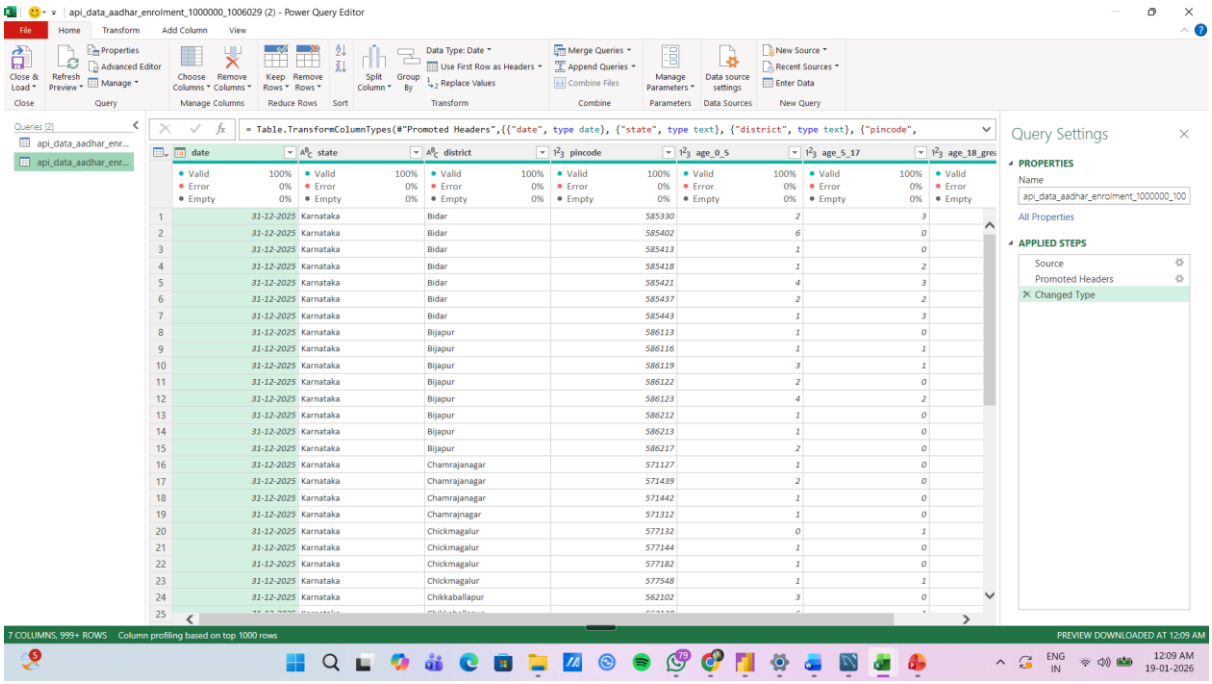
3.3 Data Transformation

Data transformation was carried out using **DAX measures** in Power BI to derive meaningful metrics.

Key transformations included:

- Total enrolments calculation
- Child and adult enrolment aggregation
- Percentage and ratio calculations
- Policy priority classification

These transformations enabled dynamic, filter-based analysis.


api_data_aadhar_enrolment_1000000_1006029 (2) - Power Query Editor
File
Home
Transform
Add Column
View
Close & Load
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Date
Use First Row as Headers
Replace Values
Merge Queries
Append Queries
Combine Files
Manage Parameters
Data source settings
New Source
Recent Sources
Enter Data
Queries [2]
api_data_aadhar_enr...
api_data_aadhar_enr...
= Table.TransformColumnTypes(#"Promoted Headers",{{"date", type date}, {"state", type text}, {"district", type text}, {"pincode",
date
state
district
pincode
age_0_5
age_5_17
Valid 100%
Error 0%
Empty 0%
1 31-12-2025 Karnataka Bidar 585330 2 3
2 31-12-2025 Karnataka Bidar 585402 6 0
3 31-12-2025 Karnataka Bidar 585413 1 0
4 31-12-2025 Karnataka Bidar 585418 1 2
5 31-12-2025 Karnataka Bidar 585421 4 3
6 31-12-2025 Karnataka Bidar 585437 2 2
7 31-12-2025 Karnataka Bidar 585443 1 3
8 31-12-2025 Karnataka Bijapur 586113 1 0
9 31-12-2025 Karnataka Bijapur 586116 1 1
10 31-12-2025 Karnataka Bijapur 586119 3 1
11 31-12-2025 Karnataka Bijapur 586122 2 0
12 31-12-2025 Karnataka Bijapur 586123 4 2
13 31-12-2025 Karnataka Bijapur 586212 1 0
14 31-12-2025 Karnataka Bijapur 586213 1 0
15 31-12-2025 Karnataka Bijapur 586217 2 0
16 31-12-2025 Karnataka Chamrajanagar 571127 1 0
17 31-12-2025 Karnataka Chamrajanagar 571439 2 0
18 31-12-2025 Karnataka Chamrajanagar 571442 1 0
19 31-12-2025 Karnataka Chamrajnagar 571312 1 0
20 31-12-2025 Karnataka Chickmagalur 577132 0 1
21 31-12-2025 Karnataka Chickmagalur 577144 1 0
22 31-12-2025 Karnataka Chickmagalur 577182 1 0
23 31-12-2025 Karnataka Chickmagalur 577548 1 1
24 31-12-2025 Karnataka Chikkaballapur 562102 3 0
Query Settings
PROPERTIES
Name
api_data_aadhar_enrolment_1000000_100
All Properties
APPLIED STEPS
Source
Promoted Headers
Changed Type
7 COLUMNS, 999+ ROWS
Column profiling based on top 1000 rows
PREVIEW DOWNLOADED AT 12:09 AM
ENG IN
12:09 AM 19-01-2026

SECTION 4: ANALYSIS, VISUALIZATION & KEY FINDINGS

4.1 Key Metrics (KPIs)

The dashboard includes the following KPIs:

- Total Aadhaar Enrolments
- Total Child Enrolments
- Total Adult Enrolments
- Child Enrolment Percentage
- Number of States Covered
- Number of Districts Covered

These KPIs provide a high-level summary of Aadhaar coverage.

4.2 Visualizations Used

State-wise Enrolment Map

- **Visual:** Filled Map
- **Purpose:** Identify geographic disparities
- **Insight:** Some states show significantly higher enrolment levels than others

Age Group Distribution

- **Visual:** Donut Chart
- **Purpose:** Analyze demographic composition
- **Insight:** Adult enrolments dominate, while child enrolments show gaps

State vs Age Group Comparison

- **Visual:** Stacked Column Chart
- **Purpose:** Compare age-wise enrolments by state
- **Insight:** Enrolment patterns vary significantly across states

Adult vs Child Enrolment Comparison

- **Visual:** Scatter Plot
- **Purpose:** Analyze demographic balance
- **Insight:** States with high adult enrolment do not always show strong child enrolment

Top 10 States & Districts

- **Visual:** Bar Charts
- **Purpose:** Rank regions by enrolment volume
- **Insight:** Highlights best-performing and under-performing regions

4.3 Key Findings

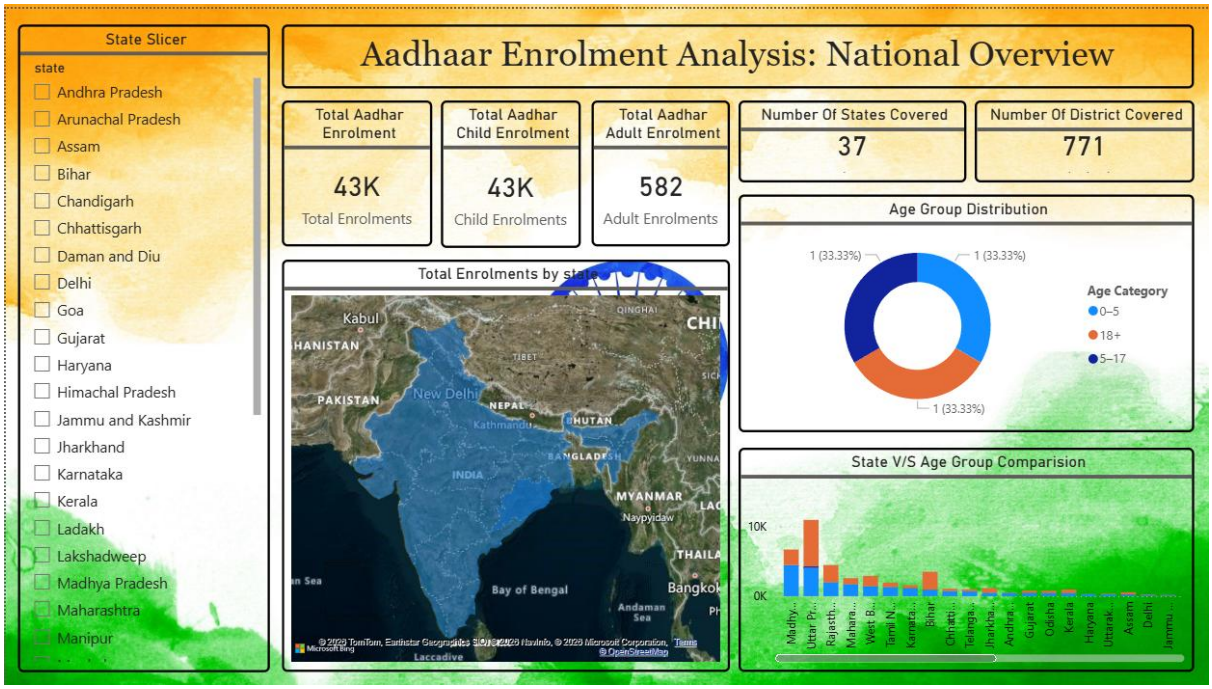
- Aadhaar enrolment coverage is strong at the national level
- Adult enrolments are consistently higher than child enrolments
- Significant variation exists across states and districts
- Child enrolment gaps indicate the need for targeted policy action
- Micro-level analysis (district and PIN) reveals localized disparities

4.4 Policy Implications

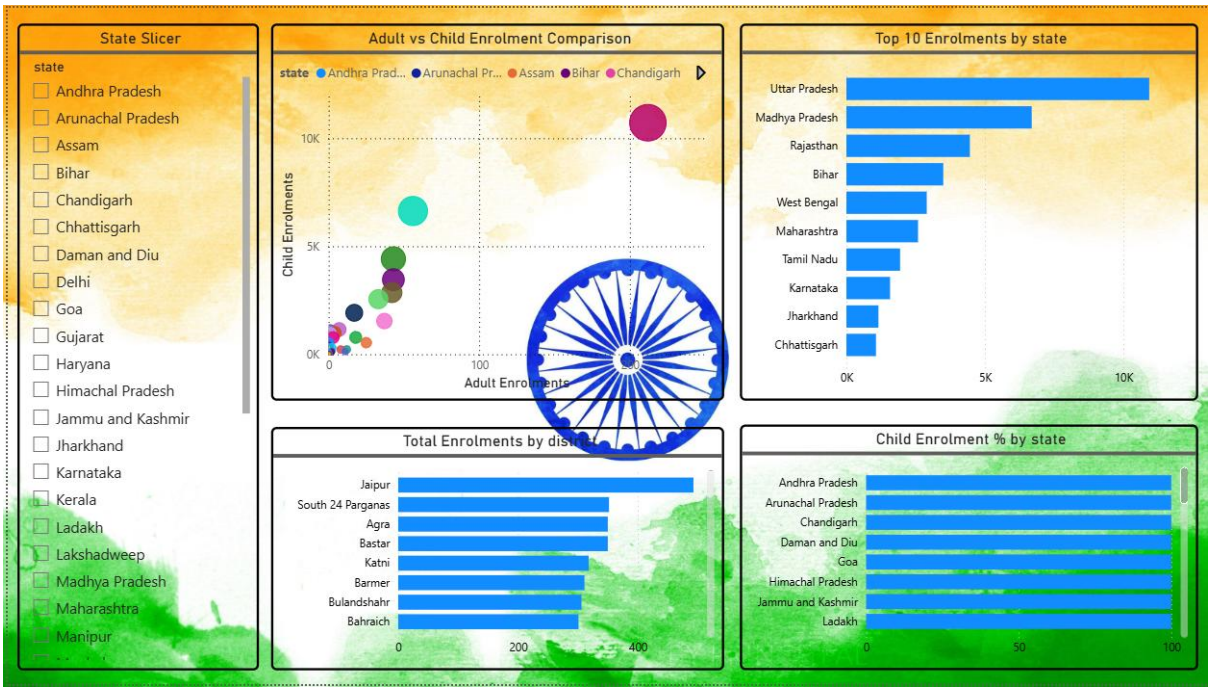
The analysis supports:

- Targeted child Aadhaar enrolment drives
- State-specific and district-level interventions
- Data-driven decision-making for government agencies

Dashboard Pg 1 :



Dashboard Pg 2 :



SECTION 5: DAX MEASURES & CALCULATIONS USED

To enable dynamic analysis and interactive visualizations, several **DAX (Data Analysis Expressions)** measures were created in Power BI. These measures aggregate raw enrolment data into meaningful metrics that update automatically based on slicers and filters.

5.1 Core Enrolment Measures

Total Aadhaar Enrolments

This measure calculates the total number of Aadhaar enrolments across all age groups.

```
Total Enrolments =
SUM('api_data_aadhar_enrolment_1000000_1006029'[age_0_5]) +
SUM('api_data_aadhar_enrolment_1000000_1006029'[age_5_17]) +
SUM('api_data_aadhar_enrolment_1000000_1006029'[age_18_greater])
```

Child Aadhaar Enrolments (0–17 Years)

This measure aggregates enrolments for children aged 0 to 17 years.

```
Child Enrolments =
SUM('api_data_aadhar_enrolment_1000000_1006029'[age_0_5]) +
SUM('api_data_aadhar_enrolment_1000000_1006029'[age_5_17])
```

Adult Aadhaar Enrolments (18+ Years)

This measure calculates total adult enrolments.

```
Adult Enrolments =
SUM('api_data_aadhar_enrolment_1000000_1006029'[age_18_greater])
```

5.2 Percentage-Based Measures

Child Enrolment Percentage

This measure shows the proportion of child enrolments relative to total enrolments.

Child Enrolment % =

DIVIDE([Child Enrolments], [Total Enrolments], 0) * 100

Adult Enrolment Percentage

This measure represents adult enrolments as a percentage of total enrolments.

Adult Enrolment % =

DIVIDE([Adult Enrolments], [Total Enrolments], 0) * 100

5.3 Geographic Coverage Measures

Number of States Covered

Counts distinct states present in the dataset.

Total States =

DISTINCTCOUNT('api_data_aadhar_enrolment_1000000_1006029'[state])

Number of Districts Covered

Counts distinct districts represented in the data.

Total Districts =

DISTINCTCOUNT('api_data_aadhar_enrolment_1000000_1006029'[district])

Number of PIN Codes Covered

Identifies the extent of micro-level geographic coverage.

Total PIN Codes =

DISTINCTCOUNT('api_data_aadhar_enrolment_1000000_1006029'[pincode])

5.4 Advanced & Policy-Oriented Measures

Child Coverage Gap Percentage

This measure highlights the shortfall in child Aadhaar enrolment.

```
Child Coverage Gap % =
100 - [Child Enrolment %]
```

Policy Priority Classification

States are classified based on child enrolment percentage to support targeted interventions.

```
Policy Priority =
IF(
    [Child Enrolment %] < 30, "High Priority",
    IF(
        [Child Enrolment %] < 60, "Medium Priority",
        "Low Priority"
    )
)
```

SECTION 6: KEY PERFORMANCE INDICATORS (KPIs) USED

KPIs provide a high-level snapshot of Aadhaar enrolment coverage and are displayed prominently at the top of the dashboard.

6.1 List of KPIs

KPI Name	Description	Business Value
Total Aadhaar Enrolments	Total enrolments across all age groups	Measures overall Aadhaar reach
Child Aadhaar Enrolments	Enrolments for ages 0–17	Tracks child inclusion
Adult Aadhaar Enrolments	Enrolments for ages 18+	Indicates workforce coverage
Child Enrolment %	Percentage of child enrolments	Identifies demographic gaps
Adult Enrolment %	Percentage of adult enrolments	Measures maturity of adoption
Number of States Covered	Total states & UTs	Shows geographic reach
Number of Districts Covered	Total districts	Indicates administrative penetration
Number of PIN Codes Covered	Micro-level coverage	Reflects last-mile reach
Child Coverage Gap %	Gap in child enrolments	Supports policy prioritization
High Priority States	States with low child enrolment	Enables targeted action

6.2 Importance of KPIs

These KPIs allow stakeholders to:

- Quickly assess Aadhaar penetration
- Compare demographic participation
- Identify regional disparities
- Monitor progress of enrolment initiatives
- Support evidence-based policymaking

CONCLUSION

The analysis reveals that while overall Aadhaar enrolment is strong across the country, noticeable disparities exist in child enrolment coverage. Adult enrolments significantly outnumber child enrolments in most regions, indicating the need for focused initiatives to ensure early inclusion. Geographic analysis further demonstrates uneven distribution of enrolments across states and districts, emphasizing the importance of region-specific strategies rather than a one-size-fits-all approach.

Through the use of KPIs, maps, and comparative visualizations, the dashboard enables policymakers and administrators to quickly identify high-priority regions and demographic gaps. The integration of DAX-based metrics ensures that insights remain dynamic and adaptable to changing filters and perspectives. Overall, this project demonstrates the effectiveness of business intelligence tools in transforming large-scale government datasets into actionable insights that support informed decision-making and policy formulation.

Future enhancements to this project may include time-series analysis using multi-year data, integration of population benchmarks for penetration analysis, and predictive modeling to forecast enrolment trends. Such extensions would further strengthen the analytical depth and policy relevance of the dashboard.

Project Github link:

https://github.com/Pruthviraj80888/aadhaar-enrolment-analysis-powerbi/tree/main

REFERENCES

1. Government of India – Unique Identification Authority of India (UIDAI). *Aadhaar Enrolment Statistics and Open Data Resources.*
2. UIDAI Official Website. *https://uidai.gov.in*
3. Open Government Data (OGD) Platform India. *https://data.gov.in*
4. Microsoft Power BI Documentation. *DAX Reference and Data Modeling Guidelines.*
5. Ministry of Electronics and Information Technology (MeitY), Government of India. *Digital India Initiative Reports.*
6. Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling.*
7. Few, S. (2012). *Show Me the Numbers: Designing Tables and Graphs to Enlighten.*