

# National Aadhaar Enrollment Analysis Dashboard

Data-Driven Insights into Aadhaar Enrollment Trends & Operational Pressure



Name:

Harsh Soni



Hackathon Name:

UIDAI Data Hackathon



Tools Used:

- Power BI
  - Power Query
  - DAX
- (Measure Functions used for KPIs, Aggregations & Insights)



Data Source:

event.data.gov.in



# Problem–Solution Analysis: Aadhaar Enrollment Challenges & Insights

## 1: No Centralized View of Aadhaar Enrollment Trends

Aadhaar enrollment data is large and spread across multiple files and time periods. Decision-makers lack a single, consolidated view to understand overall enrollment trends.

### Solution

I built a centralized Power BI dashboard by appending and cleaning multiple datasets using Power Query, providing a unified view of national Aadhaar enrollment.

## 2: Difficulty in Identifying Peak Enrollment Periods

Enrollment centers face sudden crowding due to the absence of insights into peak days and months, making staff and resource planning inefficient.

### Solution

The dashboard analyzes day-wise and month-wise enrollment patterns, clearly highlighting peak enrollment periods to support better operational planning.

## 3: Lack of Visibility into Age-Group Driven Demand

There is limited clarity on which age groups contribute the most to Aadhaar enrollment, making it difficult to design targeted policies.

### Solution

I performed age-group segmentation (0–5, 5–17, 18+) and visualized enrollment dominance, revealing that early-age (0–5 years) registrations drive the majority of enrollments.

## 4: Inefficient Resource Allocation Across States

Without state-wise enrollment insights, resources are often distributed uniformly, leading to under-servicing in high-pressure regions.

### Solution

The project includes a state-wise enrollment pressure analysis, enabling identification of high-demand states for focused resource allocation.

## 5: Inability to Distinguish Normal vs High-Pressure Periods

All enrollment periods are treated equally, even though some periods experience unusually high operational load.

### Solution

I classified enrollment periods into low, moderate, high, and extreme pressure categories, allowing administrators to proactively prepare for high-pressure phases.

## 6: Limited Understanding of Institutional vs Public-Driven Enrollment

It is unclear whether enrollments are driven by public walk-ins or institutional processes such as schools and welfare programs.

### Solution

By analyzing weekday vs weekend enrollment patterns, the dashboard shows strong weekday dominance, indicating institution-driven enrollment behavior.

## 7: Lack of Insight into Policy-Driven Enrollment Impact

Government schemes often influence enrollment spikes, but their impact is not easily measurable.

### Solution

The dashboard includes policy impact segmentation, helping assess how welfare and administrative programs affect enrollment volumes.

 “This project transforms raw Aadhaar enrollment data into a decision-support dashboard that helps identify demand patterns, operational pressure, and demographic drivers for better government planning.”





# Project Objective

## Objective:

The objective of this project is to analyze Aadhaar enrollment data and transform it into meaningful insights using data visualization and analytics techniques. The project supports data-driven decision-making by identifying enrollment patterns, demographic dominance, and operational pressure across different regions and time periods.

## Specific Objectives

- ✓ Consolidate and prepare Aadhaar enrollment data using Power Query
- ✓ Highlight state-wise high-pressure regions
- ✓ Analyze enrollment trends across days, months, and years
- ✓ Classify enrollment activity into pressure and momentum categories
- ✓ Identify age-group-wise enrollment dominance
- ✓ Present insights through an interactive Power BI dashboard



**One-Line Objective:** To build a data-driven Power BI dashboard that analyzes Aadhaar enrollment patterns across time, age, and geography to support informed government planning and resource allocation.





# Data Description

## Dataset Summary:

The dataset used in this project consists of three Aadhaar enrollment CSV files sourced from the Government of India's open data platform (event.data.gov.in). Each file captures Aadhaar enrollment activity across different time periods. The datasets were combined to enable continuous, nation-wide analysis of enrollment behavior across India.

## Data Source Section



**Source:**

event.data.gov.in



**Data Type:**

Government  
Open Data



**Domain:**

Aadhaar  
Enrollment  
Records



**Consolidated  
Master Dataset**

## Data Integration Process

The three datasets were appended using Power Query to create a single consolidated master dataset. This integration ensures consistency in analysis and allows enrollment trends to be examined seamlessly across time without fragmentation.

## Data Content Overview



**Temporal Attributes**

Enrollment Date, Day Name,  
Month, Quarter, Day Type  
(Weekday / Weekend)



**Geographical Attributes**

State, District, Location,  
Pincode



**Demographic Attributes**

Age Group, Population Structure,  
Working Age Presence,  
Dominant Age Group



**Enrollment Metrics**

Enrollment Count, Total  
Enrollment, Age Enrollment  
Value, Enrollment Category



**Analytical & Derived Fields**

Demand Pressure, Enrollment  
Intensity, Enrollment Momentum,  
Policy Impact Zone, Child  
Dependency Area, Age  
Priority Rank



## Data Preparation

The appended dataset was cleaned and transformed using Power Query, including column standardization, data type conversion, handling missing and duplicate values, and creating calculated fields required for advanced analysis.



## Purpose of the Data:

The prepared dataset supports detailed analysis of Aadhaar enrollment trends across time, geography, and population segments, enabling data-driven insights into operational pressure, demographic dominance, and policy-driven enrollment behavior.





# Data Preparation



**ETL Overview:** The data preparation process was carried out using Power Query in Power BI to ensure Aadhaar enrollment data was clean, consistent, and analysis-ready. Since the project involved multiple datasets, a structured ETL (Extract, Transform, Load) approach was followed.



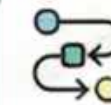
## Data Ingestion & Integration

- Imported three Aadhaar enrollment datasets (CSV format) into Power BI
- Appended all datasets using Power Query to create a single consolidated master dataset
- Ensured continuity of records across time periods and eliminated data fragmentation



## Data Cleaning

- Standardized column names across datasets
- Assigned correct data types for dates, numeric values, and categorical fields
- Handled missing or invalid values in critical fields (enrollment date, state, age)
- Identified and removed duplicate records created during dataset integration



## Data Transformation

- Extracted analytical date attributes such as day name, month, quarter, and weekday/weekend
- Categorized age values into defined groups (0–5, 5–17, 18+ years)
- Cleaned and standardized geographic attributes for accurate state and district-level analysis



## Feature Engineering

- Created derived analytical fields including Enrollment Intensity, Demand Pressure, Enrollment Momentum, Policy Impact Zone, and Dominant Age Group
- Developed additional classification fields to identify population structure and working-age presence



## Data Validation

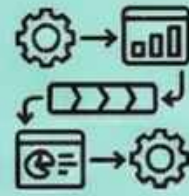
- Reviewed final dataset for accuracy, completeness, and consistency
- Performed summary checks to confirm correct record counts after appending and cleaning
- Validated dataset before using it for dashboard visualization



## Outcome:

The data preparation process resulted in a clean, consolidated, and enriched dataset, enabling detailed analysis of Aadhaar enrollment patterns across time, geography, and demographic segments, and forming a reliable foundation for all dashboard insights.





# Tools & Technologies



CSV Data



Power Query (ETL)



Power BI (Visualization)



$f_x$  DAX (Insights)

**Overview:** The following tools and technologies were used to design, develop, and analyze the Aadhaar enrollment dashboard. Each tool played a specific role in the data analytics workflow, from data preparation to insight generation.



## Power BI



Used as the primary platform for data **visualization** and dashboard development. It enabled the creation of **interactive** and **dynamic** dashboards that present **enrollment** trends, demographic patterns, and **operational insights** in an intuitive and user-friendly manner.



## Power Query

Served as the ETL layer of the project, handling data ingestion, cleaning, and transformation. Multiple datasets were appended, standardized, cleaned, and prepared for analysis, including data type conversion, duplicate removal, and derived field creation.



## DAX (Data Analysis Expressions)

Used to create calculated measures and analytical logic within Power BI, enabling computation of key metrics such as total enrollment, enrollment intensity, demand pressure, momentum classification, and age-group dominance for advanced insights.



## CSV Data Format

The datasets were provided in CSV format, ensuring easy integration into Power BI and compatibility during data ingestion and transformation.



## Government Open Data Platform

Data was sourced from [event.data.gov.in](http://event.data.gov.in), a Government of India open data portal, ensuring reliability, transparency, and real-world relevance of the analysis.







# Data Transformation & Feature Engineering

## ◆ STEP 1: Age Enrollment Value

```
Age Enrollment Value =  
Fact_Aadhaar_Enrollment_All[age_0_5] +  
Fact_Aadhaar_Enrollment_All[age_5_17] +  
Fact_Aadhaar_Enrollment_All[age_18_greater]
```

Combined all age-wise values to calculate total age enrollment per record.

## ◆ STEP 2: Age Group (Derived)

```
Age Group =  
SWITCH (  
    TRUE(),  
    Fact_Aadhaar_Enrollment_All[age_0_5] > 0, "0-5",  
    Fact_Aadhaar_Enrollment_All[age_5_17] > 0, "5-17",  
    "18+"  
)
```

Derived age group based on dominant age-wise enrollment.

## ◆ STEP 3: Day Name

```
Day Name =  
FORMAT ( Fact_Aadhaar_Enrollment_All[date], "dddd" )
```

Extracted day name from date for daily trend analysis.

## ◆ STEP 4: Day Type

```
Day Type =  
IF (  
    WEEKDAY ( Fact_Aadhaar_Enrollment_All[date], 2 ) >= 6,  
    "Weekend",  
    "Weekday"  
)
```

Classified enrollments into weekday and weekend.

## ◆ STEP 5: Month Name

```
Month Name =  
FORMAT ( Fact_Aadhaar_Enrollment_All[date], "MMMM" )
```

Extracted month name for seasonal trend analysis.





# Data Transformation & Feature Engineering

## STEP 6: Total Enrollment

```
Total Enrollment =  
SUM ( Fact_Aadhaar_Enrollment_All[Age Enrollment Value] )
```

🔊 Calculated total Aadhaar enrollment volume.

## STEP 7: Enrollment Intensity

```
Enrollment Intensity =  
SWITCH (  
    TRUE(),  
    [Total Enrollment] < 50000, "Low",  
    [Total Enrollment] < 150000, "Moderate",  
    [Total Enrollment] < 300000, "High",  
    "Very High"  
)
```

🔊 Classified enrollment activity based on volume.

## STEP 8: Demand Pressure

```
Demand Pressure =  
SWITCH (  
    TRUE(),  
    [Total Enrollment] < 50000, "Low Pressure",  
    [Total Enrollment] < 150000, "Moderate Pressure",  
    [Total Enrollment] < 300000, "High Pressure",  
    "Extreme Pressure"  
)
```

🔊 Identified operational pressure levels.

## STEP 9: Enrollment Momentum

```
Enrollment Momentum =  
SWITCH (  
    TRUE(),  
    [Total Enrollment] < 50000, "Low Momentum",  
    [Total Enrollment] < 150000, "Stable Momentum",  
    [Total Enrollment] < 300000, "Strong Momentum",  
    "Year-End Surge"  
)
```

🔊 Captured enrollment growth momentum.

## STEP 10: Population Structure

```
Population Structure =  
IF (  
    Fact_Aadhaar_Enrollment_All[Age Group] = "0-5",  
    "Child Heavy",  
    IF (  
        Fact_Aadhaar_Enrollment_All[Age Group] = "18+",  
        "Working Heavy",  
        "Balanced"  
    )  
)
```

🔊 Analyzed demographic dominance in enrollment.

## ✅ FINAL HONEST SUMMARY

The dataset already contained date, location, and age-segmented values. After loading the data using Power BI Get Data, I created analytical columns and measures using DAX to derive enrollment trends, pressure levels, and demographic insights.





# Dashboard Overview

**Overview:** The Power BI dashboard provides a comprehensive and interactive view of Aadhaar enrollment data, designed to support both strategic and operational decision-making. The dashboard is organized into focused analytical sections, each addressing a specific aspect of enrollment behavior.



## Executive Summary Dashboard

High-level KPIs such as total enrollments, dominant age group, and enrollment momentum, providing a quick snapshot of national performance.



## Aadhaar Enrollment Activity Snapshot

Day-wise and month-wise enrollment analysis to identify peak periods and recurring trends impacting operational planning.



## State-wise Enrollment Pressure Analysis

State and district-level visualization highlighting high, moderate, and low enrollment pressure regions for resource and infrastructure planning.



## National Enrollment Trend Analysis

Nationwide enrollment trends focusing on seasonality, weekday vs weekend behavior, and enrollment momentum classification.



## Population & Age-wise Enrollment Analysis

Demographic-focused analysis identifying age-group dominance, population structure, and child vs adult enrollment behavior to support policy and welfare planning.

**Purpose of the Dashboard:** Transform Aadhaar enrollment data into actionable insights, helping stakeholders understand demand patterns, identify pressure zones, and support data-driven planning and policy evaluation.



# National Aadhaar Enrollment – Executive Summary

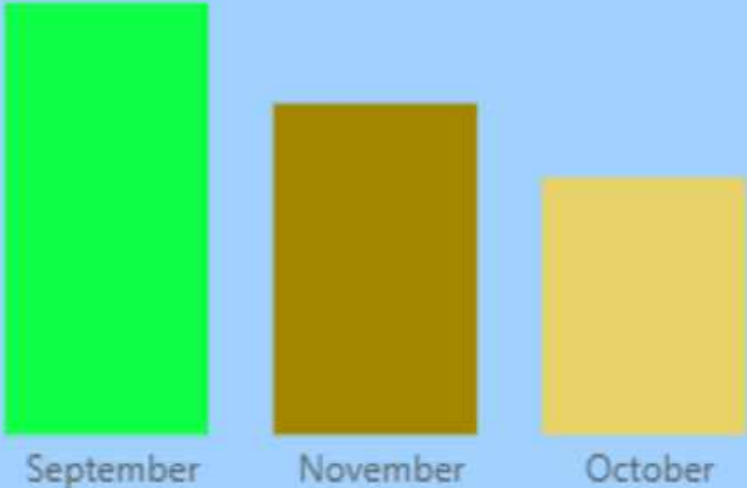


Source : [event.data.gov.in](https://event.data.gov.in)

## DATA SNAPSHOT

<b>Total Enrollment Records</b> <small>Based On Enrollment Records</small>	1.012058M
<b>Top Enrolling Age Group</b> <small>Based On Enrollment Records</small>	0-5
<b>Peak Enrollment Month</b> <small>Based On Enrollment Records</small>	September
<b>Peak Enrollment Day</b> <small>Based On Enrollment Records</small>	Friday

### Enrollment Records by Month



#### 1 Child-Centric Enrollment Dominance

Aadhaar enrollment in India is overwhelmingly driven by the 0-5 age group, indicating deep institutional integration with birth registration and child welfare systems rather than voluntary adult adoption.

#### 2 Operational Stability with Predictable Pressure

The enrollment system operates primarily under low to moderate demand pressure, reflecting strong capacity planning, while high-pressure spikes are rare, policy-driven, and manageable.

#### 3 Institution-Driven Temporal Behavior

Strong weekday dominance and program-aligned monthly spikes confirm that Aadhaar enrollment is led by institutions and administrative cycles, not random citizen walk-ins.

#### 4 Concentrated yet Scalable Geography

Enrollment activity is geographically concentrated in high-population states and specific districts, enabling targeted expansion and efficient resource deployment rather than uniform scaling.



#### Aadhaar Enrollment Activity Snapshot



#### State-wise Aadhaar Pressure Snapshot



#### National Aadhaar Enrollment Analysis



#### Population & Age-wise Enrollment Analysis

### Insights

India's Aadhaar enrollment activity is primarily driven by early-age (0-5 years) registrations, accounting for the majority of records, highlighting strong adoption at birth and early childhood stages.

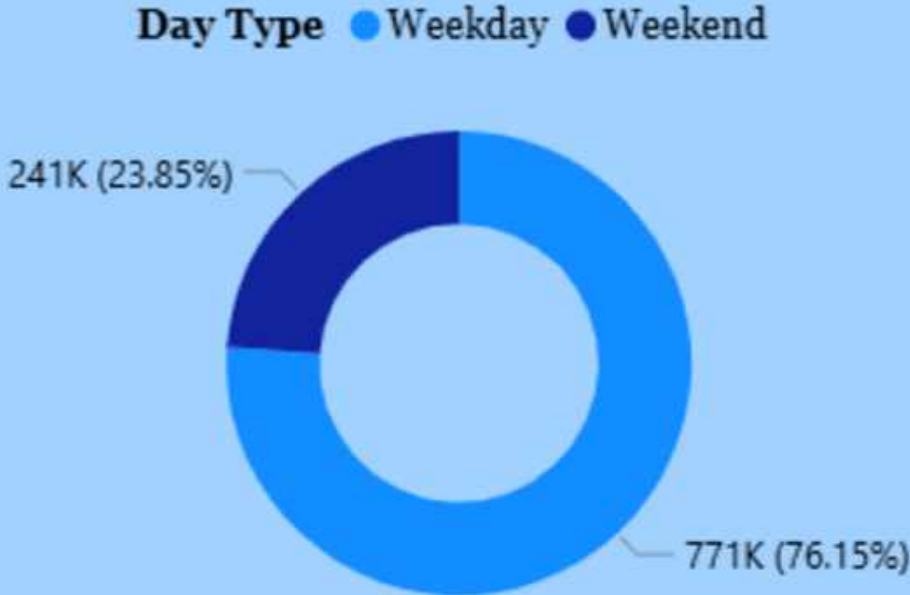


# Aadhaar Enrollment Activity Snapshot

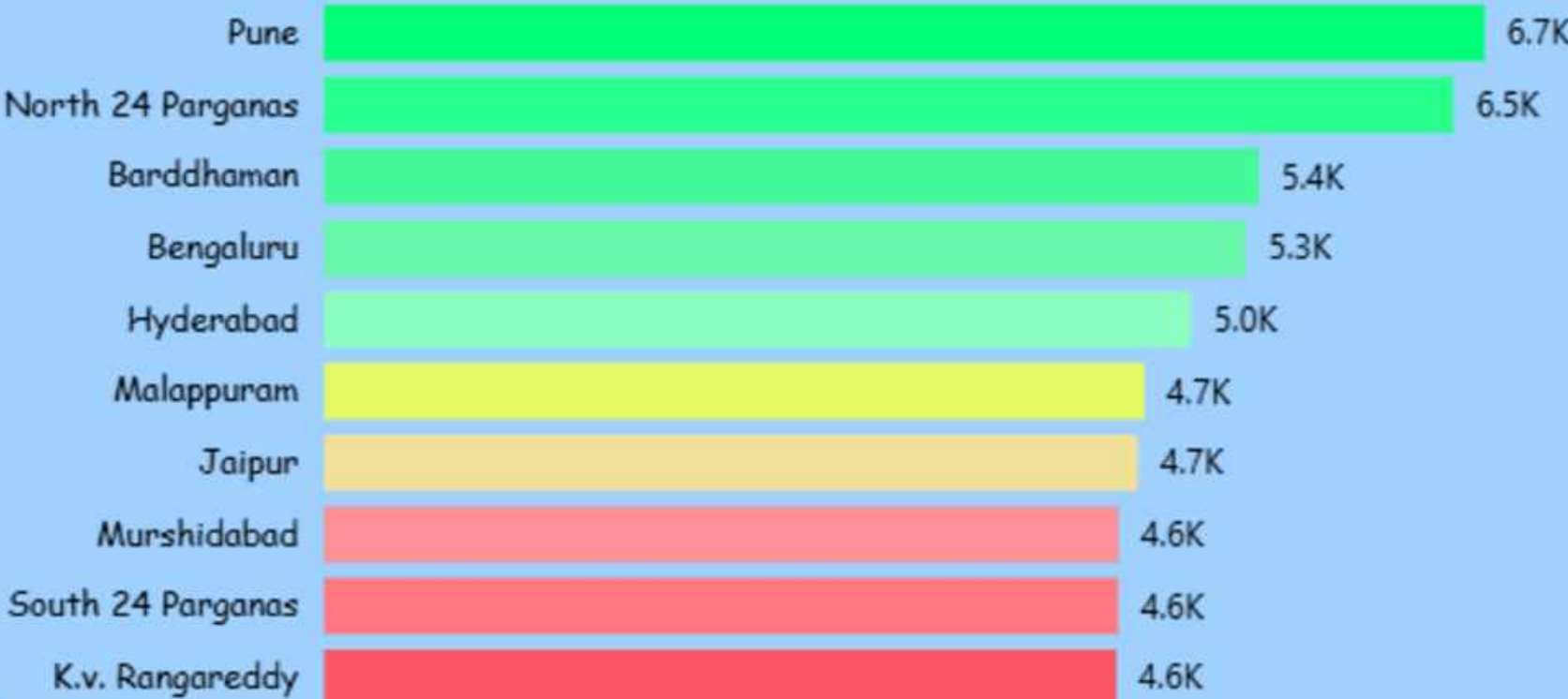
Daily Distribution of Enrollment Records



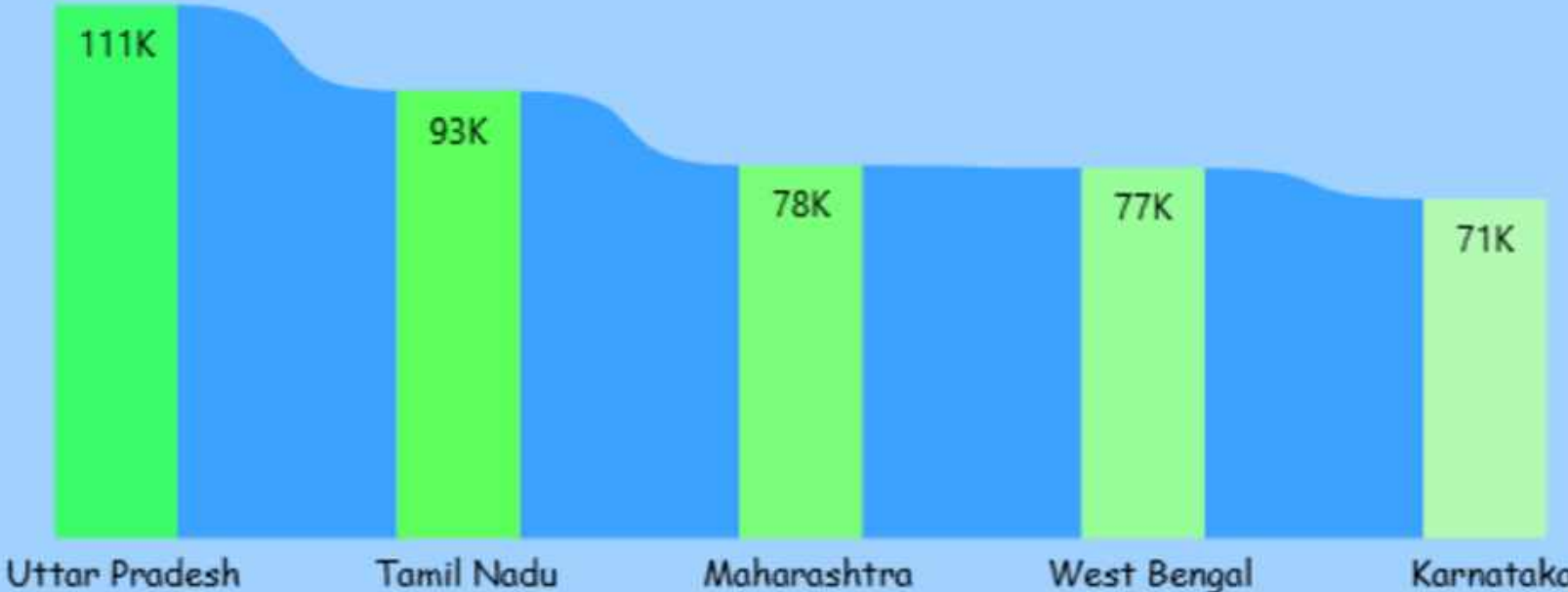
Enrollment Records: Weekday vs Weekend



Top Districts by Enrollment Records



Enrollment Records by Enrollment Intensity



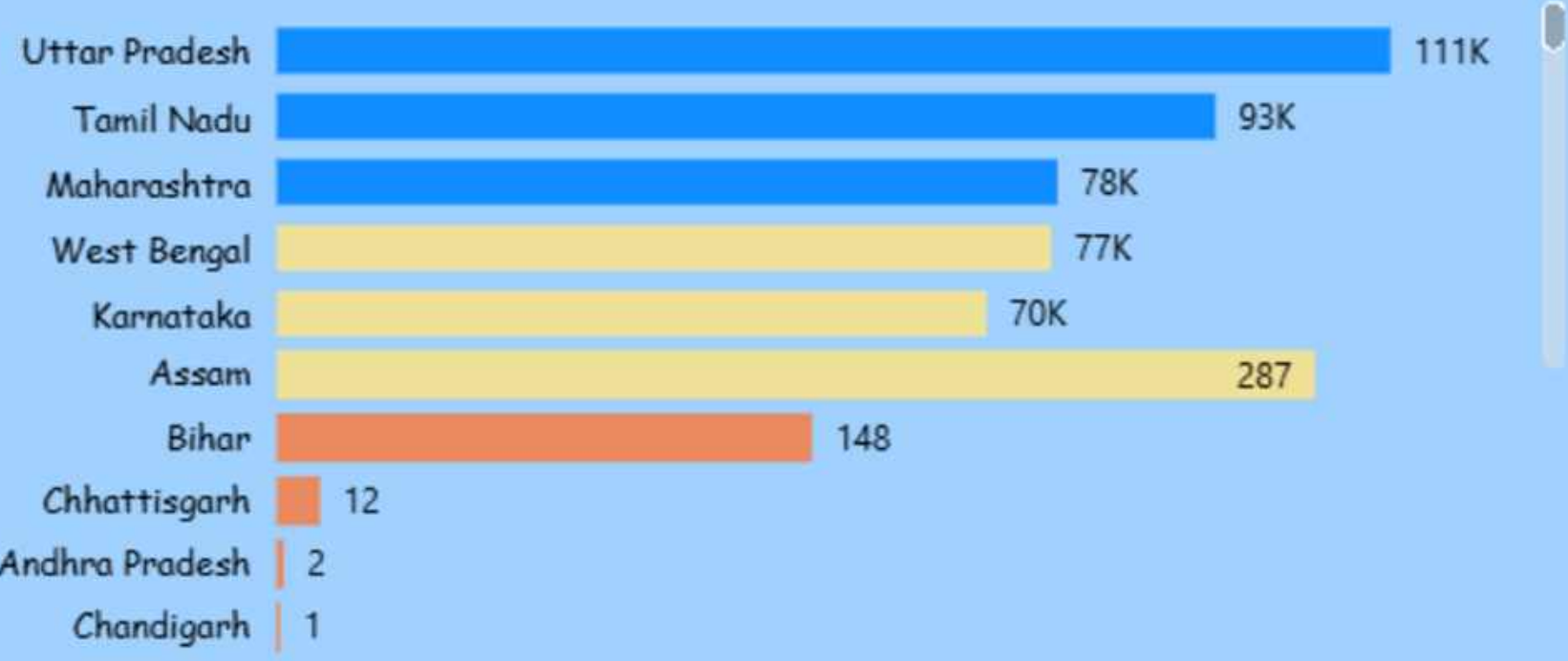
## Insights

Weekday enrollments dominate over weekends, indicating institution-driven enrollment activity.

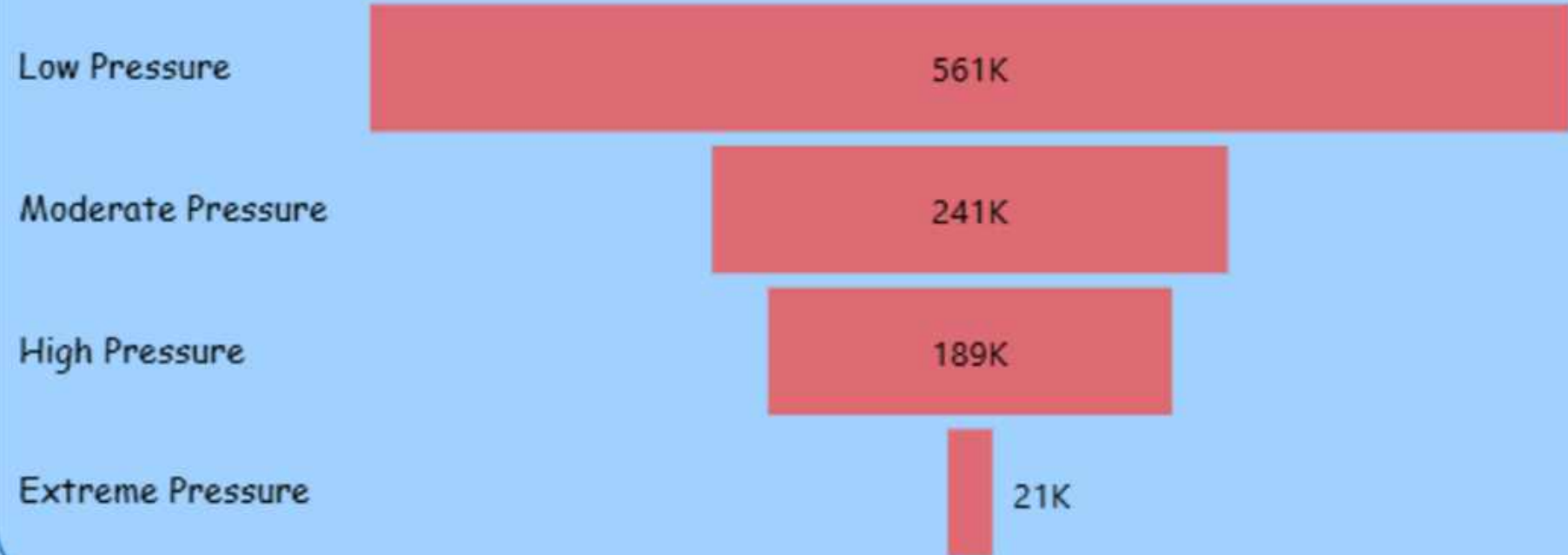


State-wise Aadhaar Enrollment Pressure Snapshot

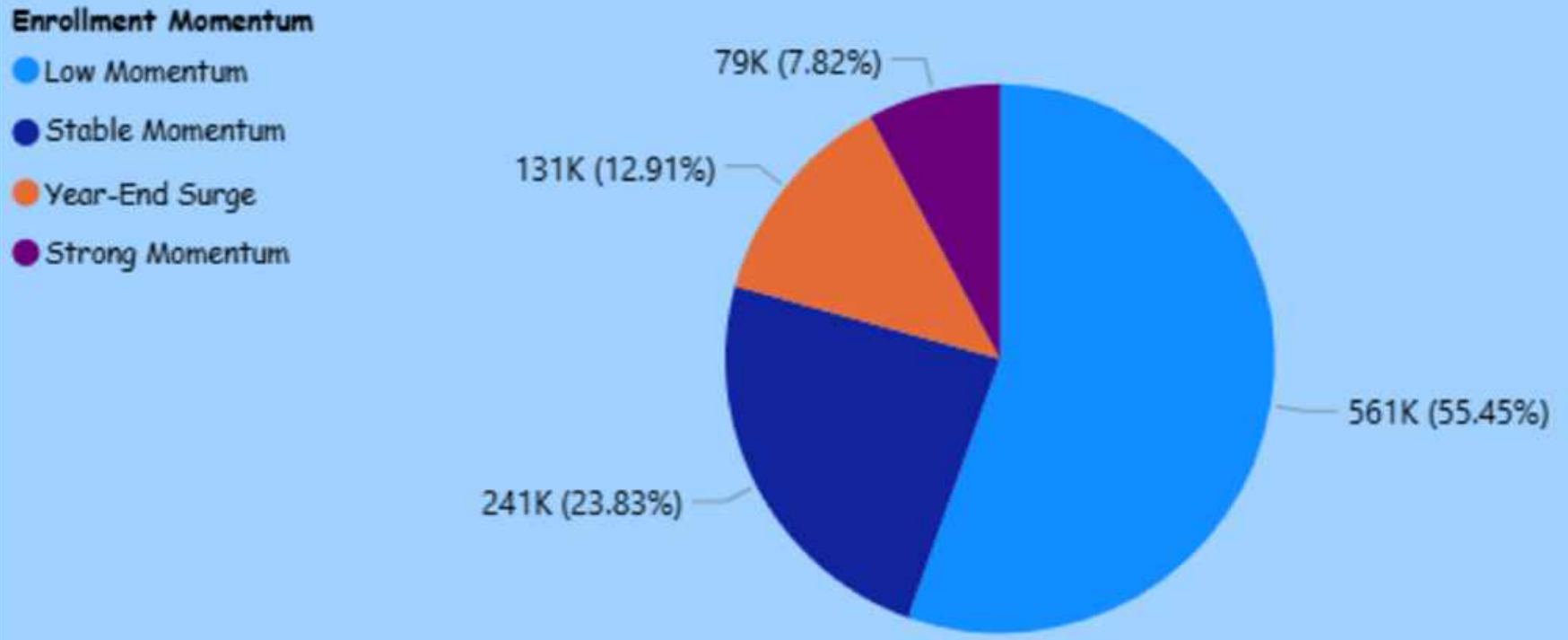
State-wise Aadhaar Enrollment Volume



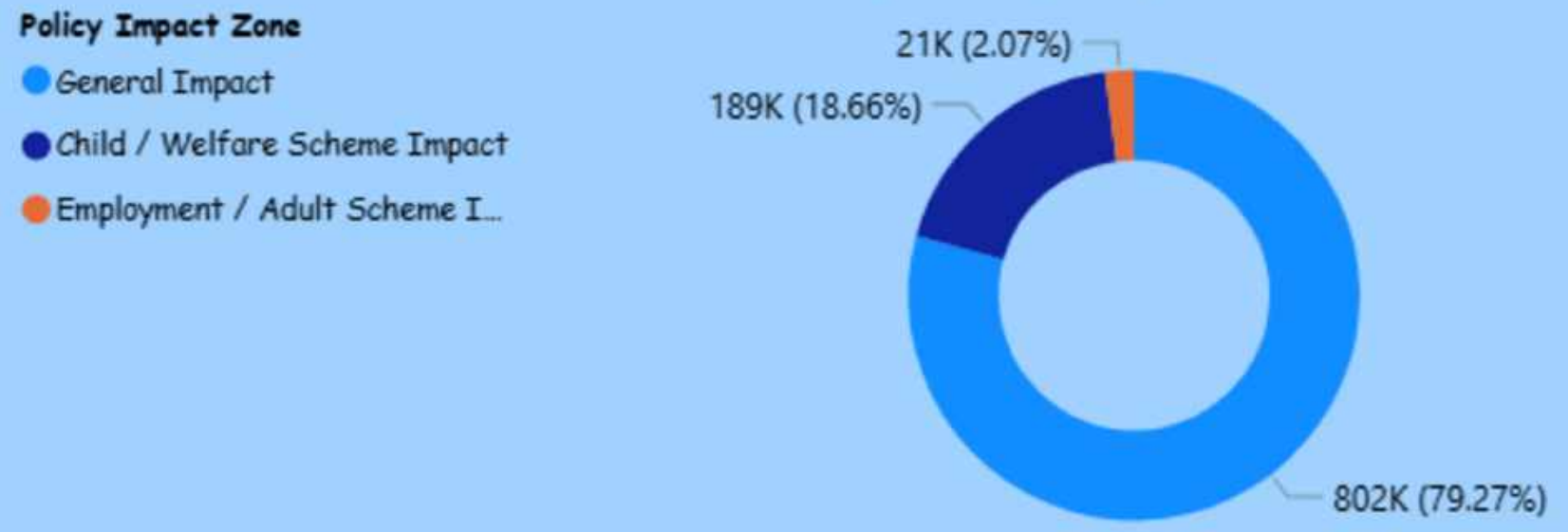
Enrollment Distribution by Demand Pressure Level



Enrollment Momentum Classification



Enrollment Records by Policy Impact Zone



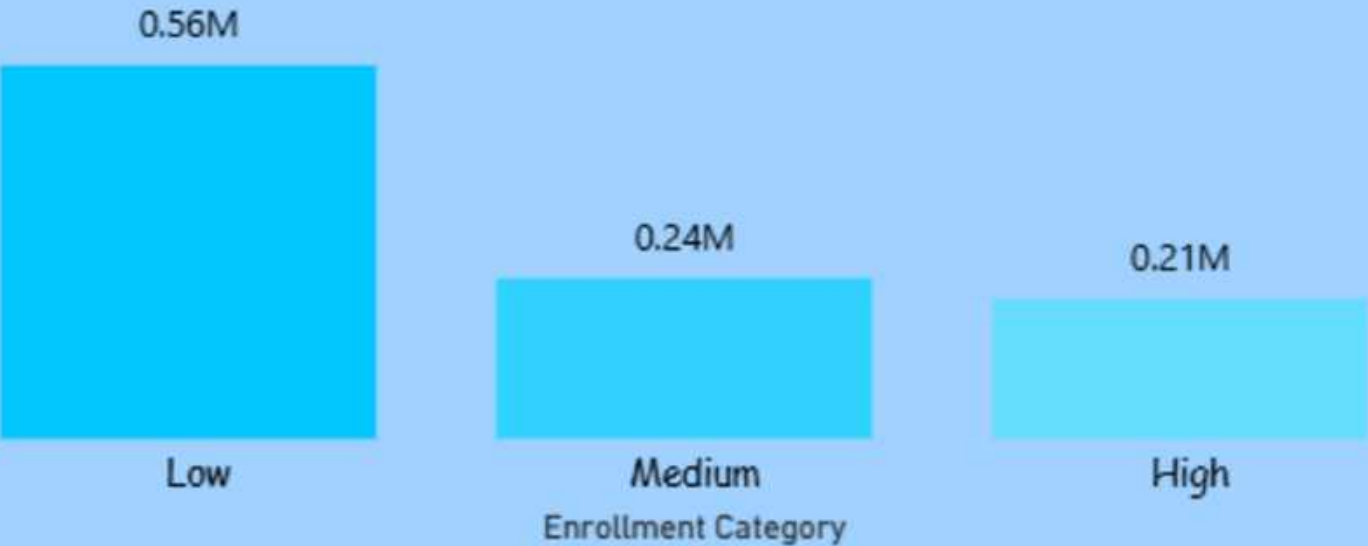
Insights

Most enrollments occur under low operational pressure, with fewer but critical high-pressure spikes.

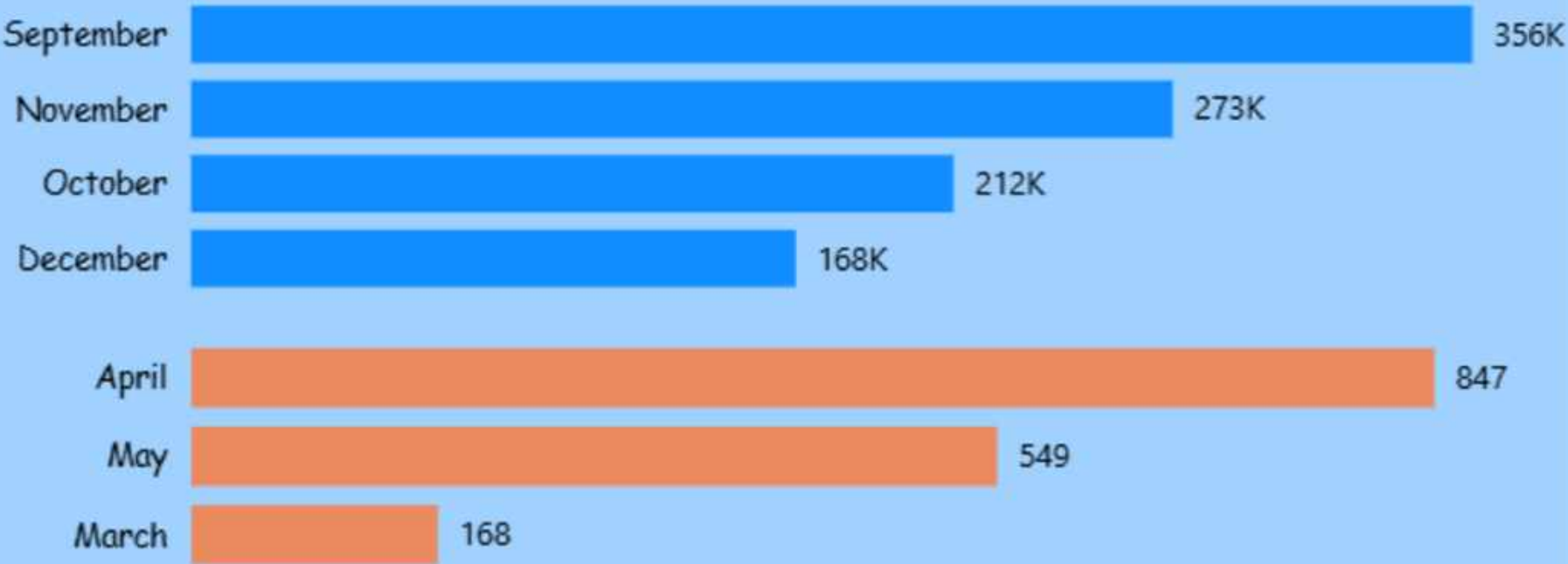


# National Aadhaar Enrollment Analysis

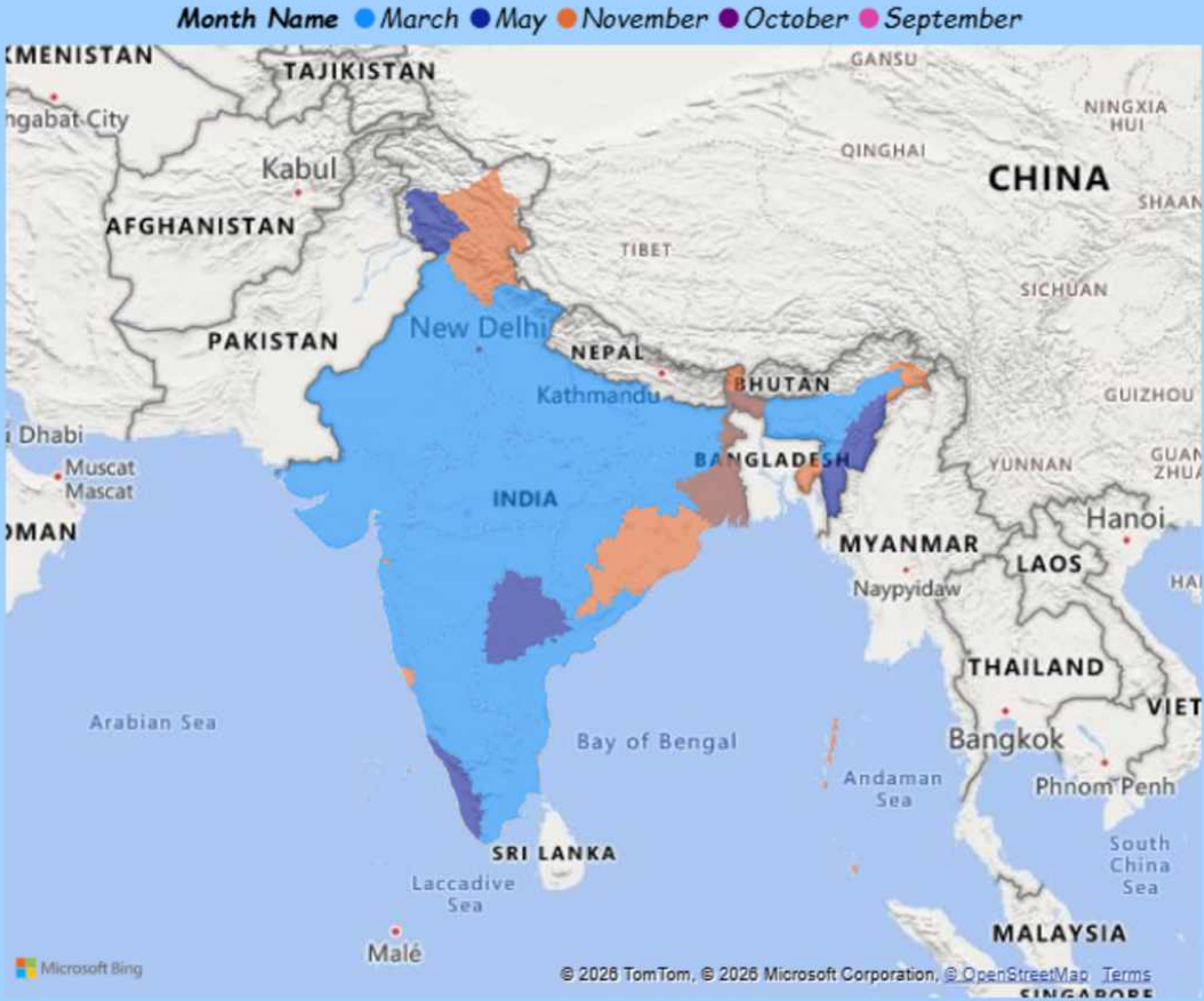
Enrollment Records by Enrollment Category



Monthly Distribution of Enrollment Records



State-wise Aadhaar Enrollment Distribution



## Insights

Low enrollment category accounts for the highest number of records, indicating frequent low-volume enrolments, while high enrollment periods are fewer but operationally significant.



# Population & Age-wise Enrollment Analysis

## Days

- ☐ Monday
- ☐ Tuesday
- ☐ Wednesday
- ☐ Thursday
- ☐ Friday
- ☐ Saturday
- ☐ Sunday

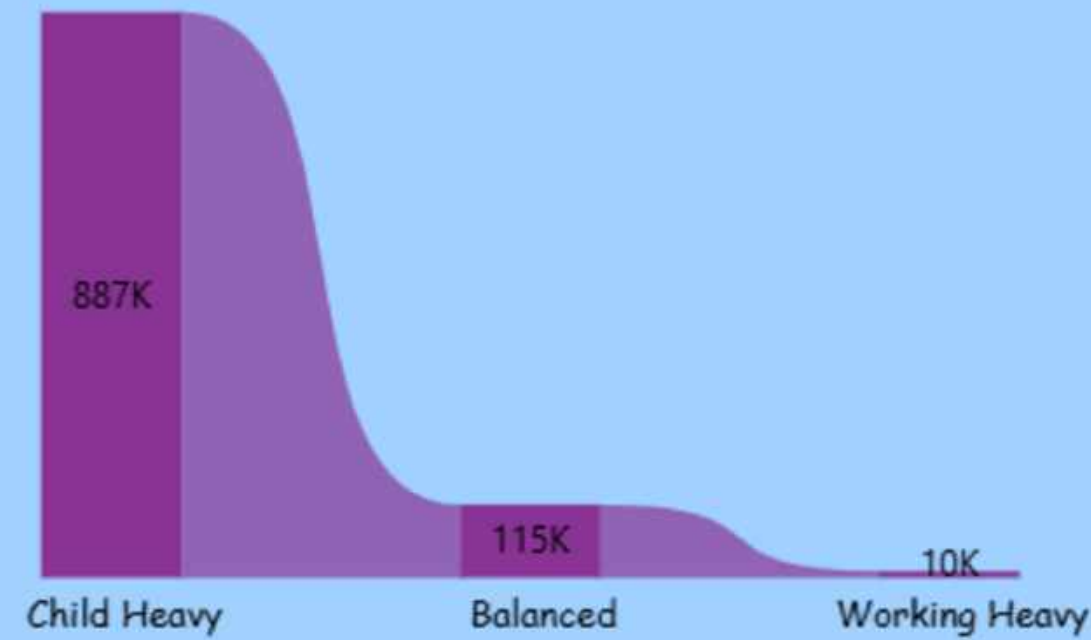
## Intensity

- ☐ High Activity
- ☐ Low Activity
- ☐ Moderate Activity
- ☐ Very High Activity

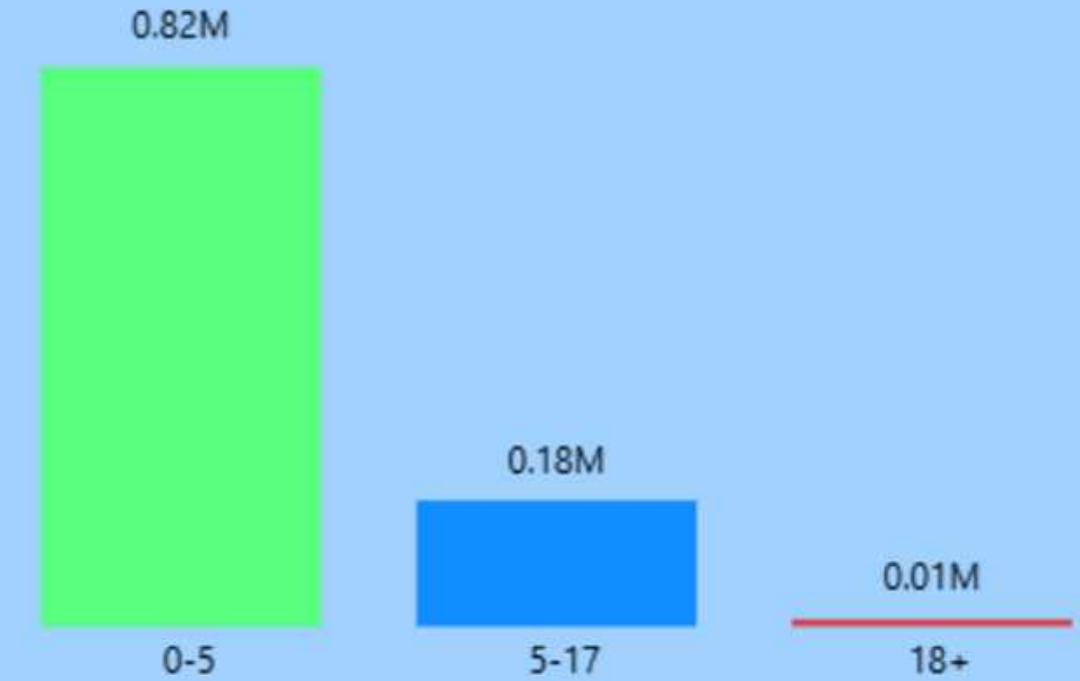
## Momentum

- ☐ Low Momentum
- ☐ Stable Momentum
- ☐ Strong Momentum
- ☐ Year-End Surge

## Enrollment Records by Population Structure



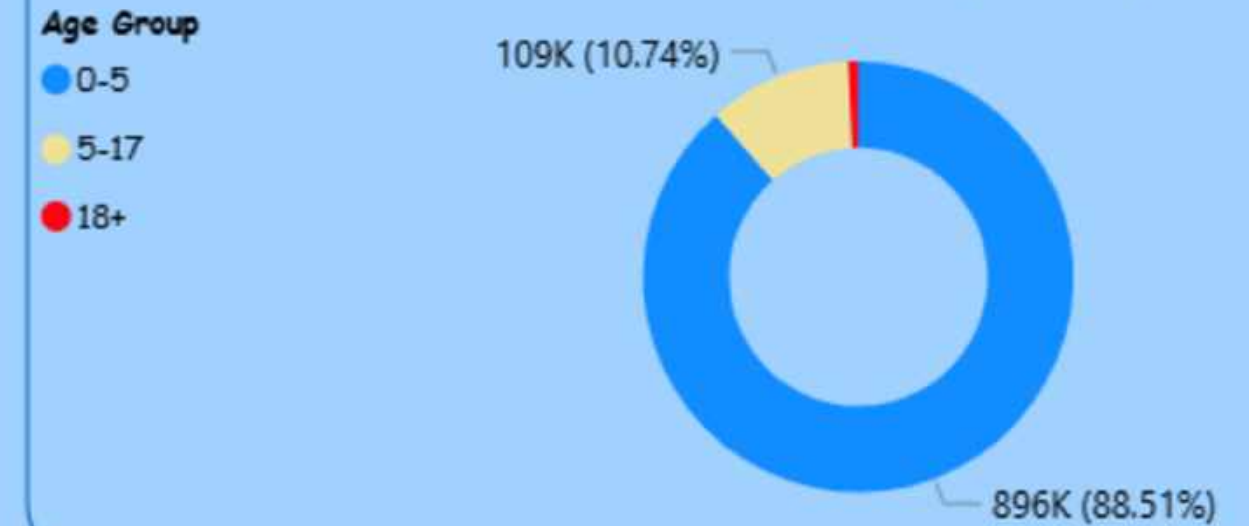
## Distribution of Dominant Age Groups



## Quarter-wise Enrollment by Working Age Presence



## Enrollment Records by Age Group



## Insights

Child-heavy populations and the 0-5 age group drive the majority of Aadhaar enrollments, indicating strong early-age adoption.





# Impact & Recommendations

## Project Impact:

This project transforms raw Aadhaar enrollment data into actionable insights that support operational planning, resource optimization, and policy evaluation.

By analyzing enrollment patterns across time, geography, and demographics, the dashboard enables stakeholders to shift from reactive decision-making to a data-driven approach.

## ✓ Key Impact Highlights

✓ Improved visibility into when and where enrollment demand is highest

✓ Better understanding of age-group-driven enrollment behavior

✓ Identification of low- and high-pressure operational periods

✓ Enhanced ability to interpret policy-driven enrollment spikes

## 📌 Recommendations

### 1 Optimize Enrollment Center Operations

- Increase staffing and resources during peak enrollment months and weekdays
- Reduce operational load during consistently low-activity periods
- Use pressure classification to plan temporary enrollment camps

### 2 Target Child-Centric Enrollment Programs

Since the 0–5 age group dominates enrollments, strengthen integration with:

- Birth registration systems
- Anganwadi and child welfare centers
- Ensure Aadhaar enrollment is completed early in the child lifecycle

### 3 Region-Specific Resource Allocation

- Deploy additional enrollment kits and manpower in high-pressure states and districts
- Avoid uniform resource distribution across low-demand regions
- Use district-level insights for micro-level planning

### 4 Plan for Policy-Driven Enrollment Spikes

Anticipate enrollment surges during:

- Welfare scheme launches
- Administrative deadlines
- Prepare contingency plans for short-term high-pressure periods

### 5 Adopt Data-Driven Monitoring

- Regularly update enrollment dashboards to track trends
- Use momentum and intensity indicators to monitor operational health
- Encourage data-backed decision-making at regional and national levels

## 🏆 Overall Recommendation:

Government agencies and administrators can leverage this dashboard as a decision-support tool to improve enrollment efficiency, enhance service delivery, and support evidence-based policy planning.





## Future Scope & Predictive Ideas

### Overview:

While the current project focuses on historical analysis of Aadhaar enrollment data, the same **analytical framework** can be **extended to support predictive analytics** and **proactive** decision-making for improved planning and policy execution.



### Future Enhancements

#### 1 Enrollment Demand Forecasting

- Use historical enrollment trends across months, weekdays, and age groups
- Apply time-series forecasting models to estimate future demand
- Enable proactive staffing and resource planning at enrollment centers

#### 2 Early Warning System for High-Pressure Periods

- Monitor enrollment momentum and intensity indicators
- Introduce alert mechanisms to flag sudden demand spikes
- Support early preparation for policy-driven or deadline-based surges

#### 3 Region-Specific Capacity Planning

- Combine historical data with state and district-level trends
- Identify regions with consistently rising enrollment momentum
- Prioritize infrastructure expansion and manpower allocation

#### 4 Policy Impact Simulation

- Analyze historical policy-driven enrollment spikes
- Simulate expected enrollment load before new welfare scheme rollouts
- Support evidence-based policy planning and risk assessment

#### 5 Integration with Population & Census Data

- Link Aadhaar enrollment data with population or census datasets
- Improve forecast accuracy
- Identify under-enrolled regions or age groups for inclusive strategies

🌟 **Overall Future Vision:** The dashboard can evolve from a descriptive analytics tool into a predictive decision-support system, enabling early planning, efficient resource utilization, and evidence-based policy formulation.





## Get in Touch

### Overview:


Thank you for reviewing this project. This Aadhaar Enrollment Analysis Dashboard was developed as part of a hackathon to demonstrate end-to-end data analysis, data preparation, and insight generation using industry-standard tools. If you would like to explore the dashboard in detail or connect for feedback, collaboration, or discussion, please feel free to reach out.


### About the Creator


**Name:** Harsh Soni


**Role:** Aspiring Data Analyst | B.Tech (Computer Science) Student


### Technical Skills

 **Power BI:** Dashboard design & data visualization

 **Power Query:** Data loading, cleaning, and transformation

 **DAX:** Calculated columns, measures, and analytical logic

 **SQL:** Data querying and analysis

 **Python:** Data cleaning, preprocessing, and dataset handling

### Live Power BI Dashboard:



[https://app.powerbi.com/view?  
r=eyJrIjoieYmlwZWU3Y2ltNDc5OC00Y2U2LTg3MjMtZmU1Yjc4NGl0ZjU0liwidCI6IjRhNzhmOWQwLWFiZGUtNDBjNC1hMDg4LTBiOTg5NTk5M2M0YSJ9](https://app.powerbi.com/view?r=eyJrIjoieYmlwZWU3Y2ltNDc5OC00Y2U2LTg3MjMtZmU1Yjc4NGl0ZjU0liwidCI6IjRhNzhmOWQwLWFiZGUtNDBjNC1hMDg4LTBiOTg5NTk5M2M0YSJ9)

(Publicly accessible • Best viewed on desktop)

### Contact Information



**Email:**  
[harshnewa@gmail.com](mailto:harshnewa@gmail.com)



**LinkedIn:**  
<https://www.linkedin.com/in/harsh-soni-data-analyst>

I welcome feedback, suggestions, and opportunities to further improve this project and apply **data analytics** to **real-world problem solving**.  
Thank you for your time and consideration.





# Thank You

Thank you for taking the time to review this project.  
This Aadhaar Enrollment Analysis Dashboard was created as part of a hackathon to demonstrate practical application of data analytics, data preparation, and insight generation using Power BI and related tools.



## **Key Takeaway:**

The project showcases how data-driven insights can support informed decision-making, operational planning, and policy evaluation.



For feedback, collaboration, or discussion, feel free to connect.