

Aadhaar Data Hackathon — Analysis & Deliverables

Identifying Regional and Age■Group Patterns in Aadhaar Updates to Improve Service Delivery Efficiency

Author: Dev Trivedi

Date: 05 January 2026

Executive summary

States such as Uttar Pradesh and Maharashtra consistently generate the highest Aadhaar update volumes across both demographic and biometric datasets, indicating persistent service demand that warrants additional update centers or temporary camps. Within Gujarat, Ahmedabad and Surat dominate district-level update requests, pointing to urban concentration and frequent address or mobile changes. Short-term forecasts (12-week horizon) show continued elevated demand in these regions; operational actions — targeted mobile camps, temporary staffing increases during forecasted peak weeks, and biometric quality revalidation programs — will mitigate service delays and improve efficiency.

Problem Statement

Identifying Regional and Age■Group Patterns in Aadhaar Updates to Improve Service Delivery Efficiency

Why Aadhaar update demand matters

Aadhaar update volumes indicate where citizens are seeking changes (address, mobile, DOB, gender) and where services are in demand. High update demand can create service bottlenecks and delays that affect citizen experience.

How regional and age■based patterns affect service load

Regions with concentrated updates require more operational capacity. Age transitions and biometric revalidations create predictable demand spikes.

Why data■driven insights help UIDAI optimize resources

Quantitative insights enable targeted allocation of staff and mobile units, informed timing for outreach campaigns, and prioritization of infrastructure scaling in high■load districts.

Datasets Used

Primary dataset: Aadhaar Demographic Update Dataset
Focuses on updates (not enrolments). Includes region (state/district/PIN) and time data; supports service demand analysis.

Secondary dataset (bonus): Aadhaar Biometric Update Dataset
Captures biometric revalidation trends; useful for age-group service demand analysis.

Optional context: Aadhaar Enrolment Dataset (used for context only)

Methodology

Raw CSVs explored from `api_data_aadhar_demographic/` and `api_data_aadhar_biometric/`. Analytics pipeline implemented in `analysis/analytics.py`.

Script steps:

- Reads CSVs in chunks
- Normalizes column names
- Aggregates totals by state and district
- Generates charts and CSV summaries
- Resamples data weekly
- Applies Holt–Winters forecasting (12-week horizon)
- Saves outputs in `analysis/outputs/`

State-level analysis

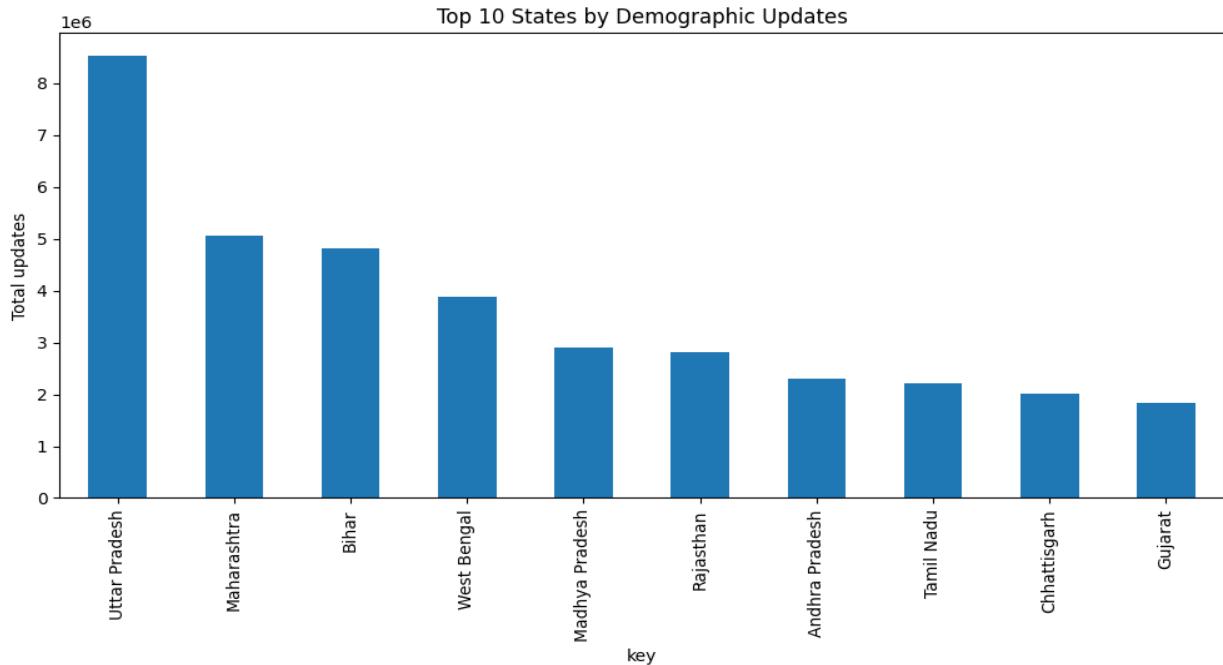


Figure: Top states by demographic updates (see analysis/outputs/state_demographic_top10.png)

Explanation: Top states by demographic updates; regions with highest update volumes (e.g., Uttar Pradesh, Maharashtra) indicate priority need for resource allocation and service scaling.

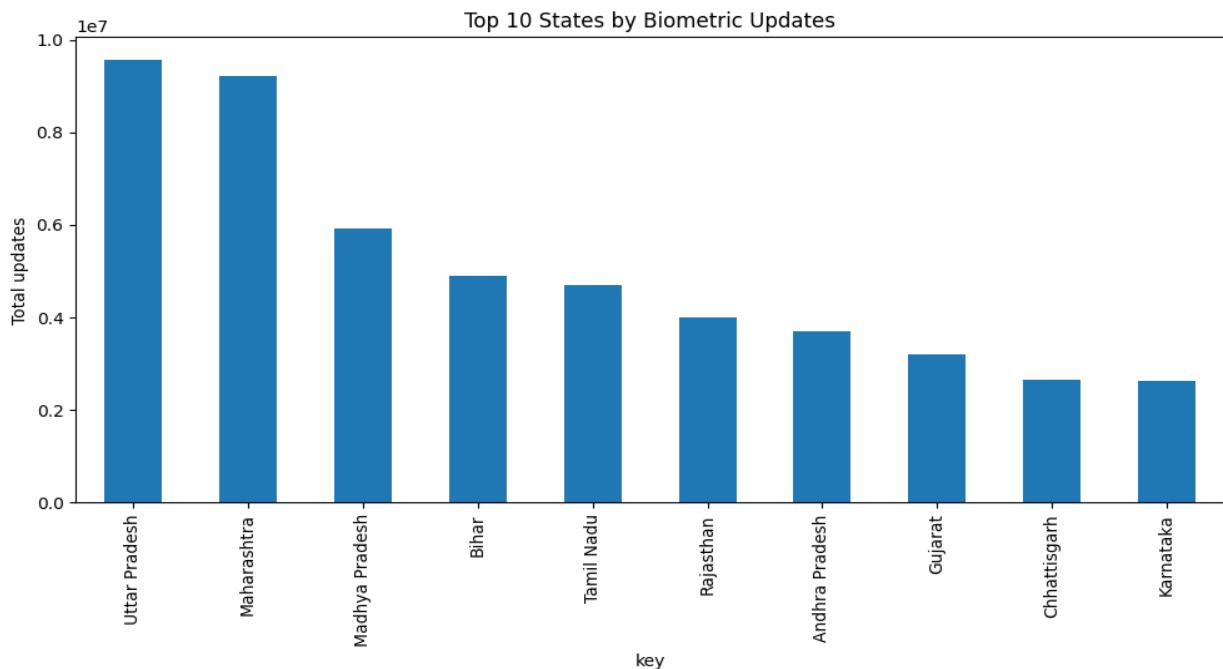


Figure: Top states by biometric updates (see analysis/outputs/state_biometric_top10.png)

Explanation: Biometric revalidation pressure; age transition and revalidation needs may be concentrated in the top states shown.

Relevant CSVs: `analysis/outputs/state_summary_demographic.csv`,
`analysis/outputs/state_summary_biometric.csv`

District-level analysis — Gujarat focus

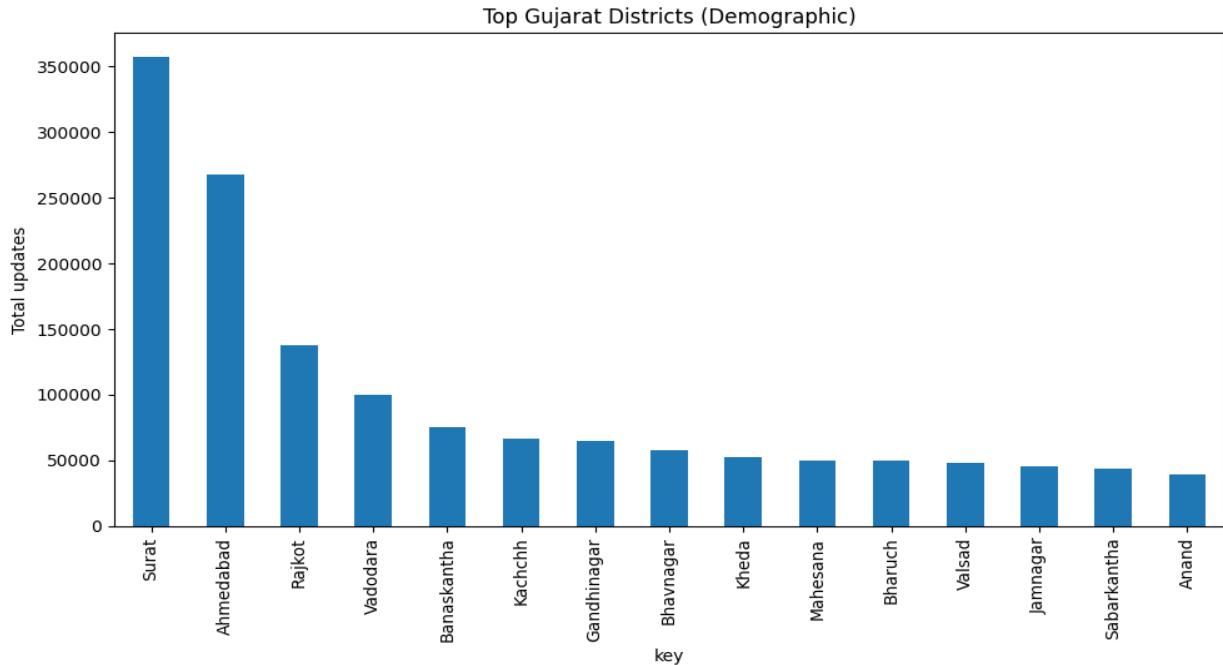


Figure: Gujarat — top districts (demographic) (see analysis/outputs/gujarat_demographic_top15.png)
Explanation: Urban concentration and migration-driven updates (Ahmedabad, Surat).

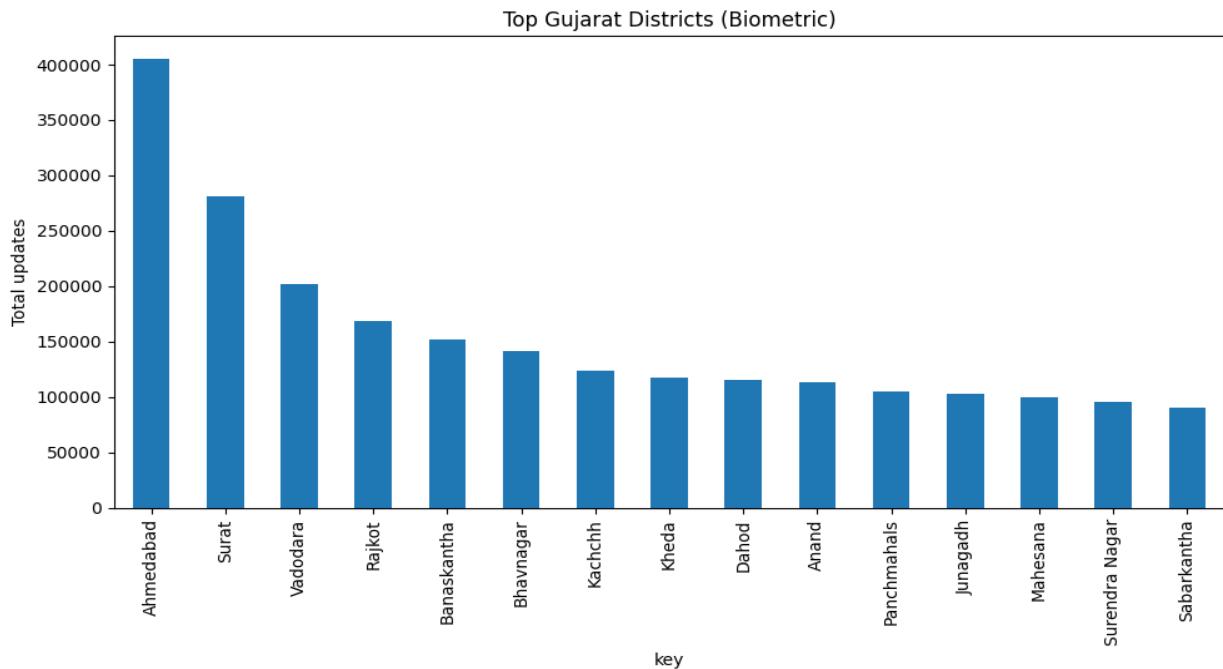


Figure: Gujarat — top districts (biometric) (see analysis/outputs/gujarat_biometric_top15.png)
Relevant CSVs: `analysis/outputs/gujarat_demographic_by_district.csv`,
`analysis/outputs/gujarat_biometric_by_district.csv`

Forecasting & Future Demand

Forecasting approach:

- Weekly resampling
- Holt-Winters seasonal model
- 12-week horizon

Example forecast CSV (preview): `analysis/outputs/forecasts/state_Uttar_Pradesh_forecast.csv`

```
,historical,forecast
2025-03-02,2199725.0,
2025-03-09,0.0,
2025-03-16,0.0,
2025-03-23,0.0,
2025-03-30,0.0,
2025-04-06,180864.0,
2025-04-13,0.0,
2025-04-20,0.0,
2025-04-27,0.0,
2025-05-04,183385.0,
2025-05-11,0.0,
2025-05-18,0.0,
2025-05-25,0.0,
2025-06-01,144904.0,
2025-06-08,0.0,
2025-06-15,0.0,
2025-06-22,0.0,
2025-06-29,0.0,
2025-07-06,329067.0,
2025-07-13,0.0,
2025-07-20,0.0,
2025-07-27,0.0,
2025-08-03,0.0,
2025-08-10,0.0,
2025-08-17,0.0,
2025-08-24,0.0,
2025-08-31,0.0,
2025-09-07,436425.0,
2025-09-14,533627.0,
2025-09-21,430173.0,
2025-09-28,0.0,
2025-10-05,0.0,
2025-10-12,0.0,
2025-10-19,287885.0,
2025-10-26,230977.0,
2025-11-02,336847.0,
2025-11-09,393281.0,
2025-11-16,491008.0,
2025-11-23,183222.0,
2025-11-30,332229.0,
2025-12-07,457106.0,
2025-12-14,479188.0,
2025-12-21,378373.0,
2025-12-28,466646.0,
2026-01-04,67396.0,
2026-01-11,,189829.51111111112
2026-01-18,,189829.51111111112
2026-01-25,,189829.51111111112
2026-02-01,,189829.51111111112
2026-02-08,,189829.51111111112
2026-02-15,,189829.51111111112
2026-02-22,,189829.51111111112
2026-03-01,,189829.51111111112
2026-03-08,,189829.51111111112
2026-03-15,,189829.51111111112
2026-03-22,,189829.51111111112
2026-03-29,,189829.51111111112
```

Explanation: Continued elevated demand; forecasts support proactive planning but are trend-based.

Key Findings

- Uttar Pradesh & Maharashtra show sustained highest update volumes
- Ahmedabad & Surat dominate Gujarat updates
- High biometric update counts indicate age-related revalidation needs

Recommendations

- Deploy temporary Aadhaar update camps in high-load districts
- Increase staffing during forecasted peak weeks
- Run biometric quality & revalidation awareness programs
- Prioritize urban districts for infrastructure scaling

Limitations

- Update-type breakdown (address/mobile/DOB/gender) unavailable due to dataset header limitations
- Forecasts are short-term and trend-based
- External covariates not included

Reproducibility

```
python3 -m venv .venv
source .venv/bin/activate
pip install -r analysis/requirements.txt
python analysis/analytics.py
python analysis/make_presentation.py
```

Deliverables

- analytics_report.md
- Generated charts (PNG)
- Forecast CSVs & PNGs
- Aggregated CSV summaries
- PPT (Aadhaar_analytics_presentation.pptx)

Conclusion

Data-driven planning and short-term forecasting enable targeted operational responses that improve service delivery and citizen experience.