UIDAI Data Hackathon

Aadhaar Demographic Insights

Simple insights, problems, and solutions from Aadhaar age data

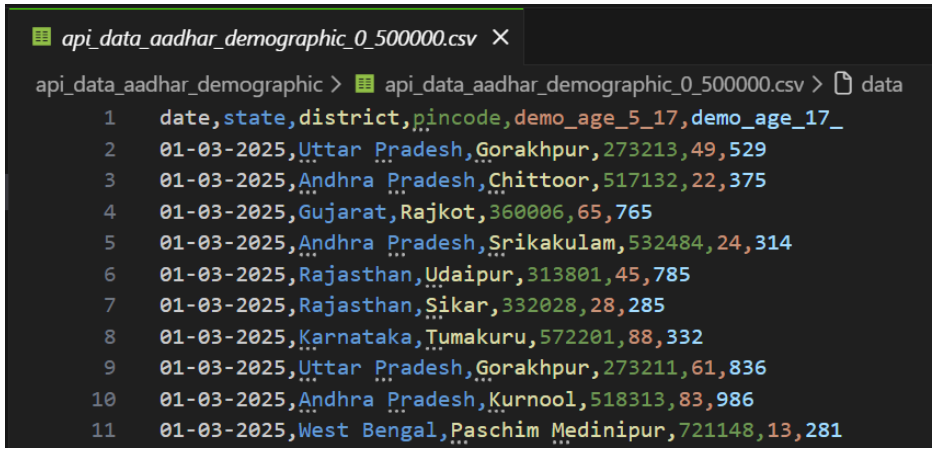
Problem Statement

Government planners need clear, data-driven insight into population structure to design education, workforce, and welfare policies. Raw Aadhaar demographic data is fragmented, noisy, and not decision-ready.

Objective: Transform raw UIDAI demographic datasets into clean, ranked, and visual intelligence showing:

1. Child vs Adult population structure
2. State-wise demographic imbalance
3. National demographic composition

Dataset Overview



UIDAI Aadhaar Demographic Datasets (Public Data):

1. api_data_aadhar_demographic_0_500000.csv
2. api_data_aadhar_demographic_500000_1000000.csv
3. api_data_aadhar_demographic_1000000_1500000.csv
4. api_data_aadhar_demographic_1500000_2000000.csv
5. api_data_aadhar_demographic_2000000_2071700.csv

Data from UIDAI Aadhaar demographic updates:

- Fields: Date, State, District, Pincode
- Age groups: Children (5–17) and Adults (17+)
- Goal: Understand patterns and improve Aadhaar systems

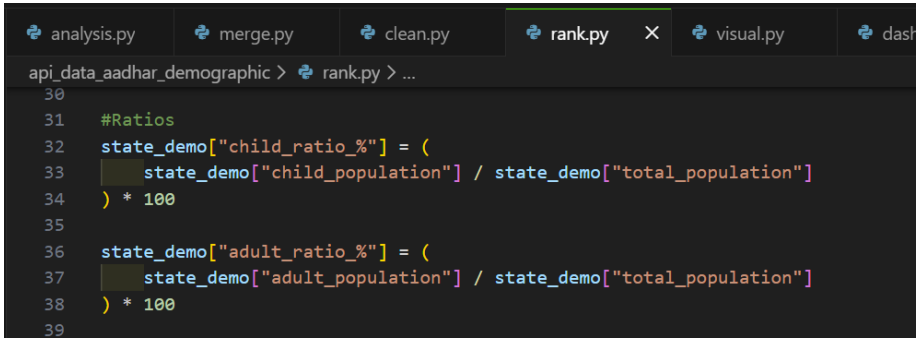
Methodology

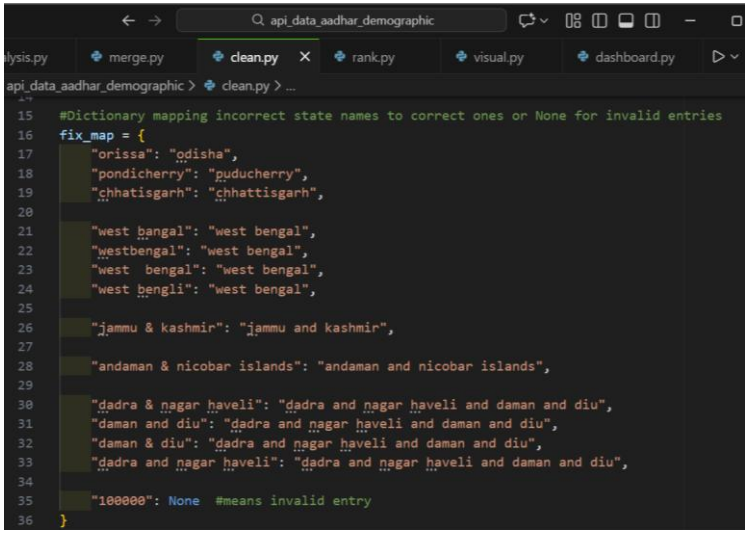
Raw CSVs → Merging → Cleaning → Intelligence → Visualization

Pipeline: merge.py merges all UIDAI

CSV files; clean.py standardizes state names and removes invalid entries; rank.py computes child/adult ratios and state rankings; visual.py generates analytical charts.

This ensures reproducibility and accuracy.





Logic:

- Child Ratio (%) = demo_age_5_17 / (demo_age_5_17 + demo_age_17_) × 100
- Adult Ratio (%) = demo_age_17_ / (demo_age_5_17 + demo_age_17_) × 100
- States are ranked using these ratios to identify demographic dominance.

Analysis:

1. Univariate: National child vs adult share

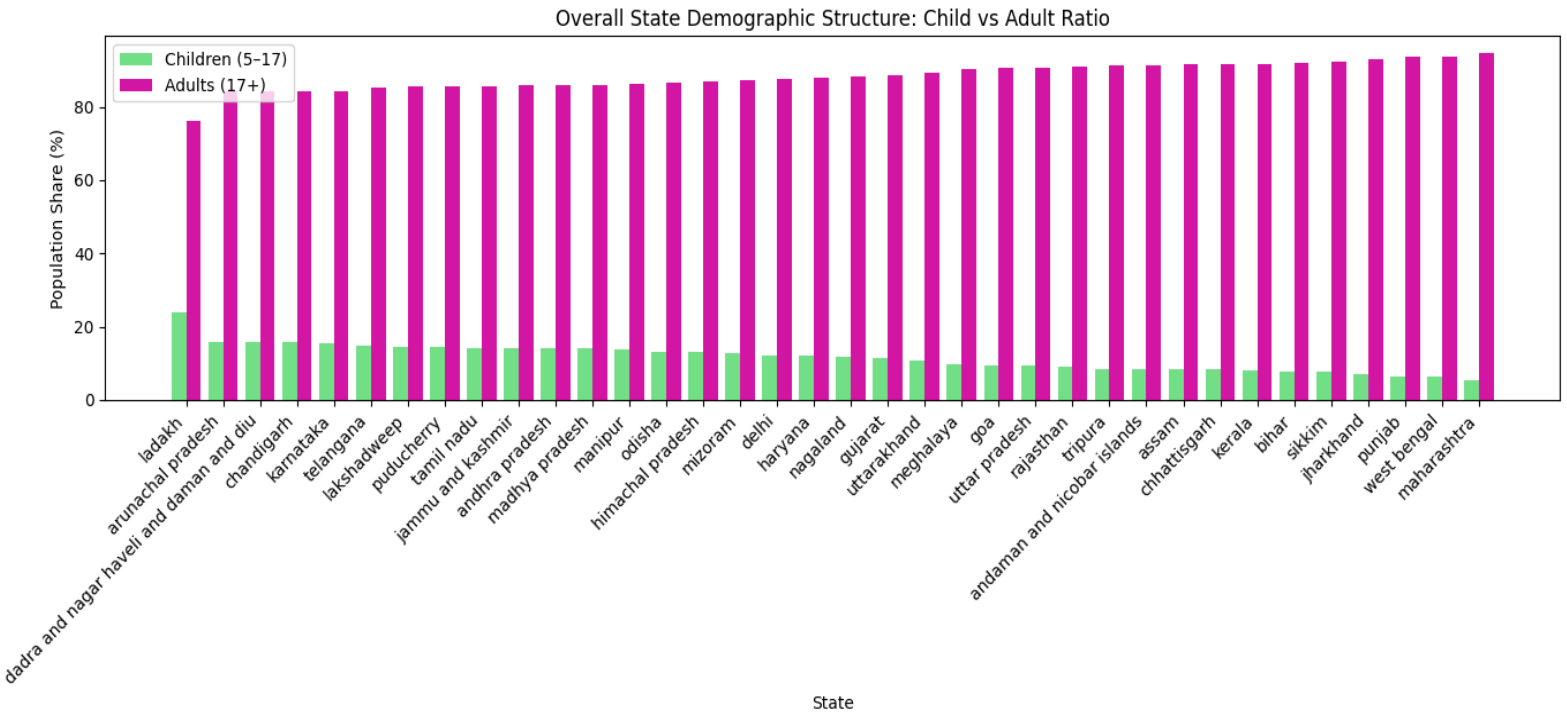
2. Bivariate: State vs population ratio

3. Comparative: Top 10 Child-Dominant and Top 10 Adult-Dominant states

Aadhaar is Adult-Dominant

- About 90% Aadhaar holders are adults
  - Only ~10% are children
  - India is mostly already enrolled
- Meaning: Focus shifts from enrollment to maintenance
  - Problem: Aadhaar data becomes outdated over time
  - Solution: Periodic reminders for address update

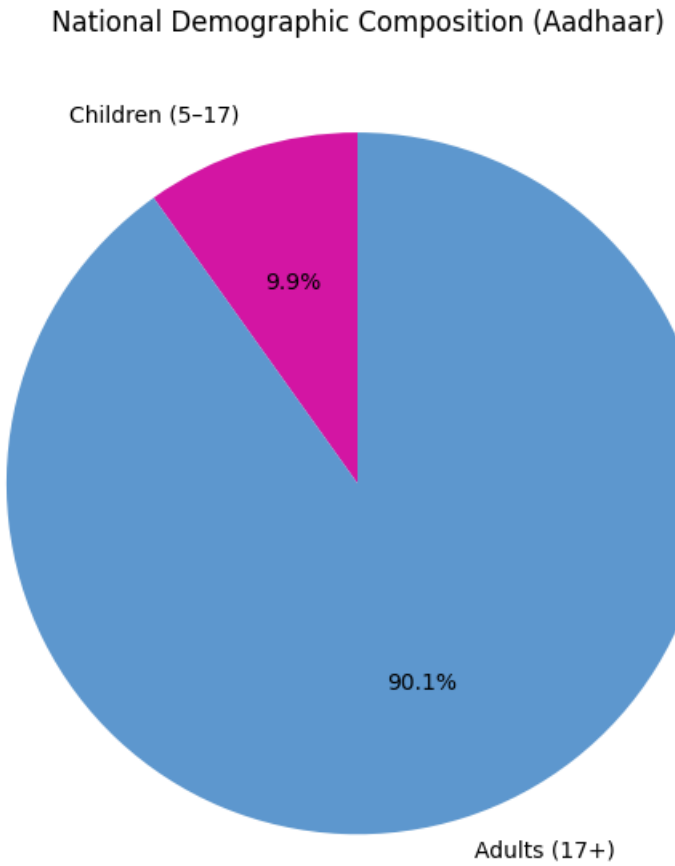
National Aadhaar Composition



States Are Not Demographically Equal

- Large variation between states
  - One policy cannot fit all
- Meaning: Different Aadhaar needs by region
  - Problem: Uniform policy reduces efficiency
  - Solution: State-specific Aadhaar strategies

Child vs Adult by State

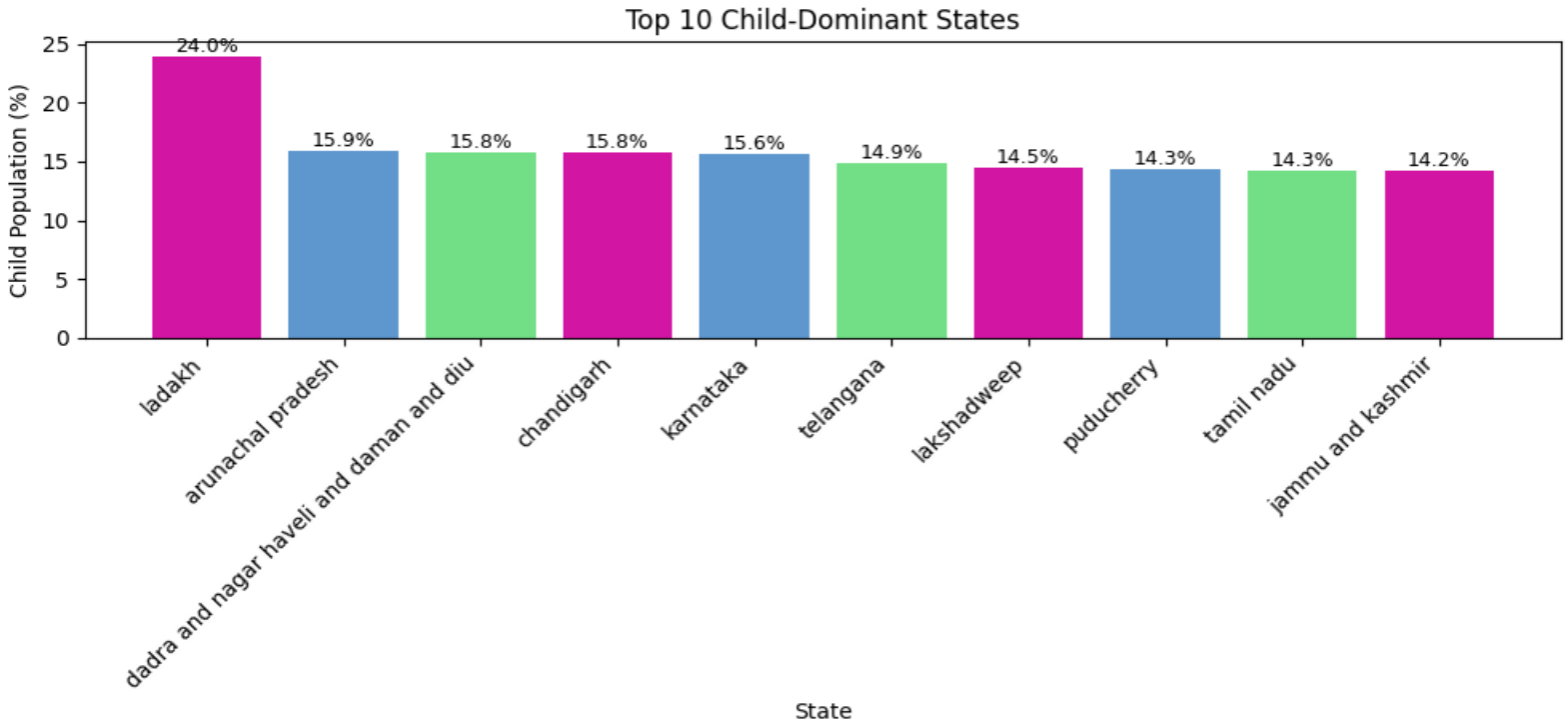


Child-Heavy Regions

- States/UTs like Ladakh, Arunachal Pradesh
  - High percentage of children (5–17)
- Meaning: Many future Aadhaar updates coming
  - Problem: Sudden overload at age 15–18
  - Solution: Predictive planning & early reminders

```python
#Child-Dominant Ranking
child_rank = state_demo.sort_values(
    by="child_ratio_%",
    ascending=False
).reset_index(drop=True)

child_rank["Child_Rank"] = range(1, len(child_rank) + 1)

print("\n Top 10 Child-Dominant States:")
print(
    child_rank[["Child_Rank", "state_", "child_ratio_%"]]
    .head(10)
    .to_string(index=False)
)

#Adult-Dominant Ranking
adult_rank = state_demo.sort_values(
    by="adult_ratio_%",
    ascending=False
).reset_index(drop=True)

adult_rank["Adult_Rank"] = range(1, len(adult_rank) + 1)

print("\n Top 10 Adult-Dominant States:")
print(
    adult_rank[["Adult_Rank", "state_", "adult_ratio_%"]]
    .head(10)
    .to_string(index=False)
)
```

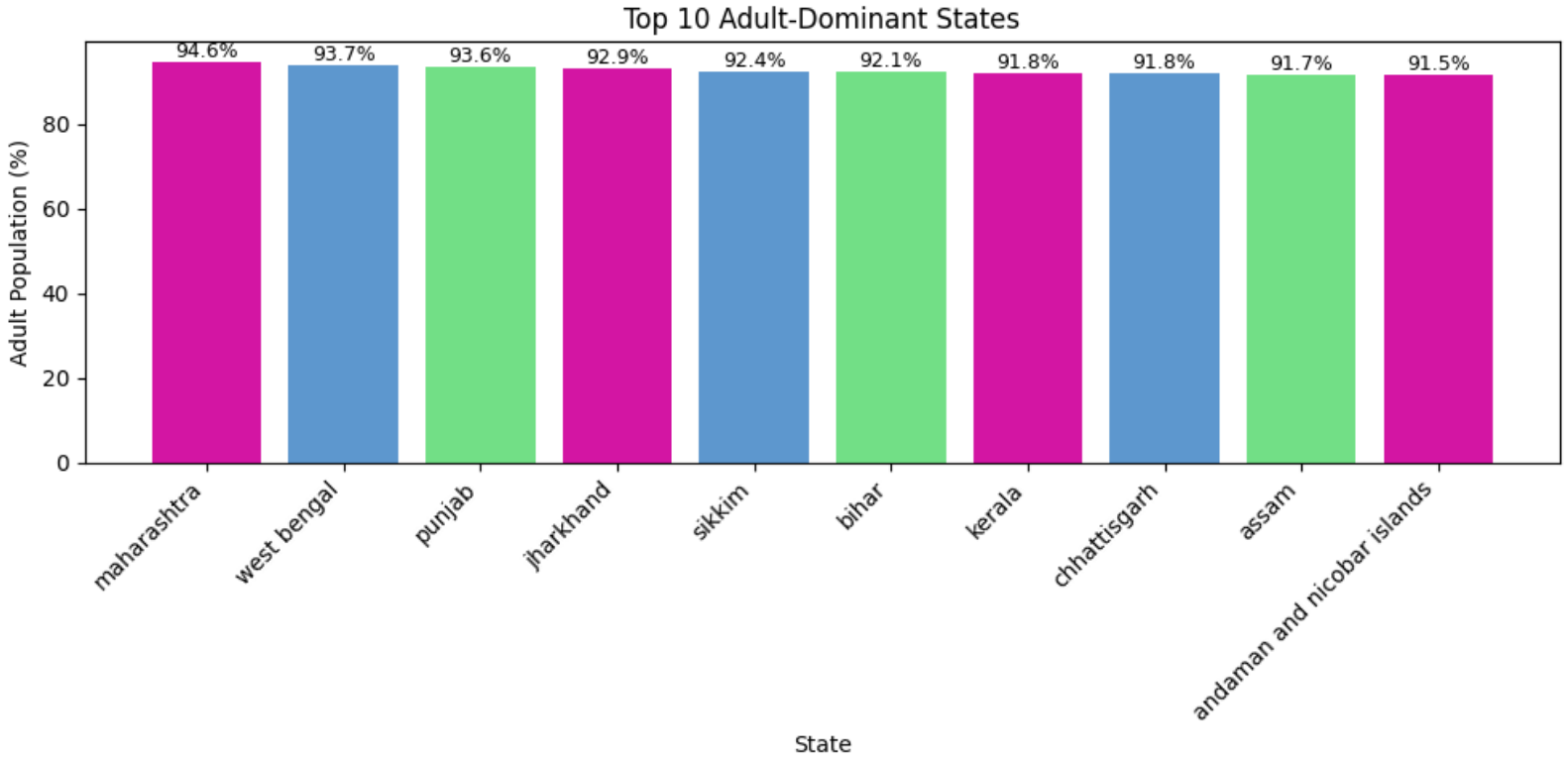
Child-Dominant States



Adult-Heavy States

- States like Maharashtra, Punjab, Kerala
  - Low child population, high adult population
- Meaning: High migration and aging population
  - Problem: Address mismatch & biometric failures
  - Solution: Targeted update drives in these states

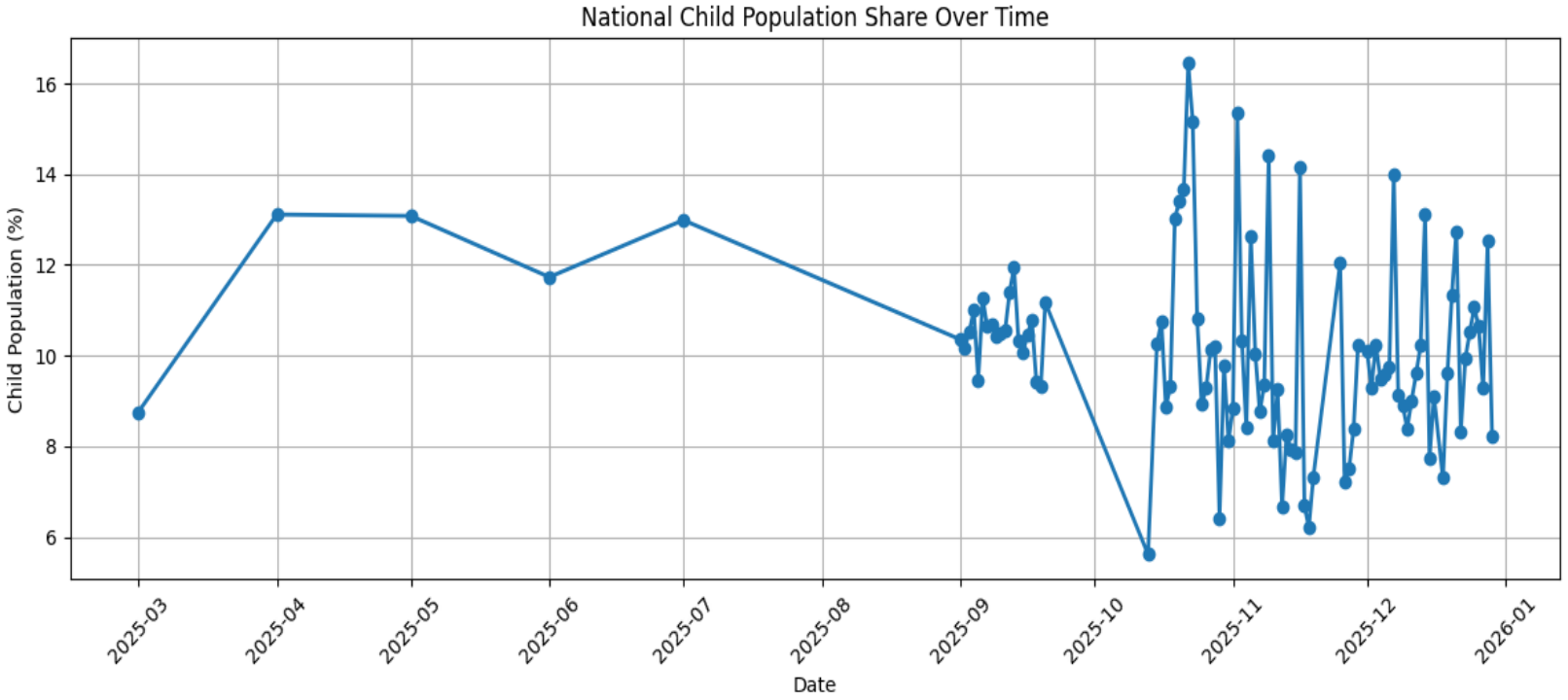
Adult-Dominant States



Sudden Spikes in Child Aadhaar Data

- Child percentage rises and falls sharply
  - Not natural population change
- Meaning: Enrollment drives cause spikes
  - Problem: UIDAI reacts late to workload
  - Solution: Real-time monitoring dashboard

Child Population Over Time



Insight 6: Future Risk Zones

- These regions will need many updates soon
  - Ignoring now causes future failures
- Meaning: High future update demand
  - Problem: Service disruption risk
  - Solution: School-linked Aadhaar update programs

Dashboard

Aadhaar Demographic Intelligence Dashboard

National Overview

Children (5–17)

4,863,424

Adults (18+)

44,431,745

State Demographic Summary

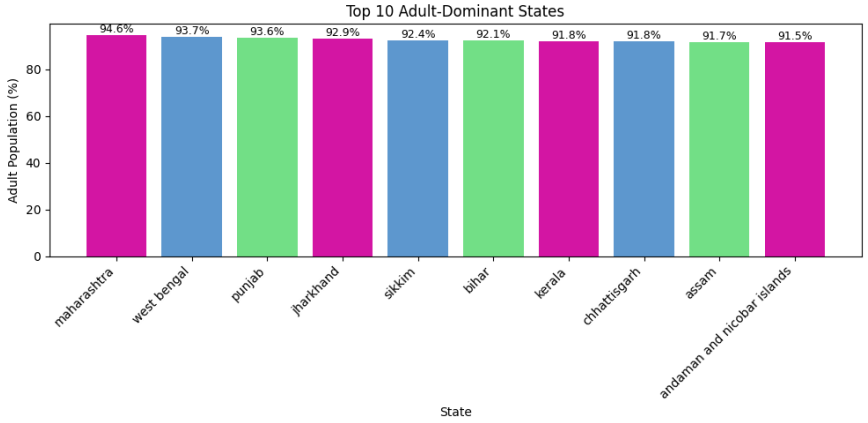
	state_	total_population	demo_age_5_17	demo_age_17_	child_population	a
0	andaman and nicobar islands	7246	617	6629	617	
1	andhra pradesh	2295582	321148	1974434	321148	
2	arunachal pradesh	36443	5783	30660	5783	
3	assam	1012578	84480	928098	84480	
4	bihar	4814350	380023	4434327	380023	
5	chandigarh	83361	13133	70228	13133	
6	chhattisgarh	2005438	165207	1840231	165207	
7	dadra and nagar haveli and dar	12204	1923	10281	1923	
8	delhi	1438934	175535	1263399	175535	
9	goa	35120	3275	31845	3275	

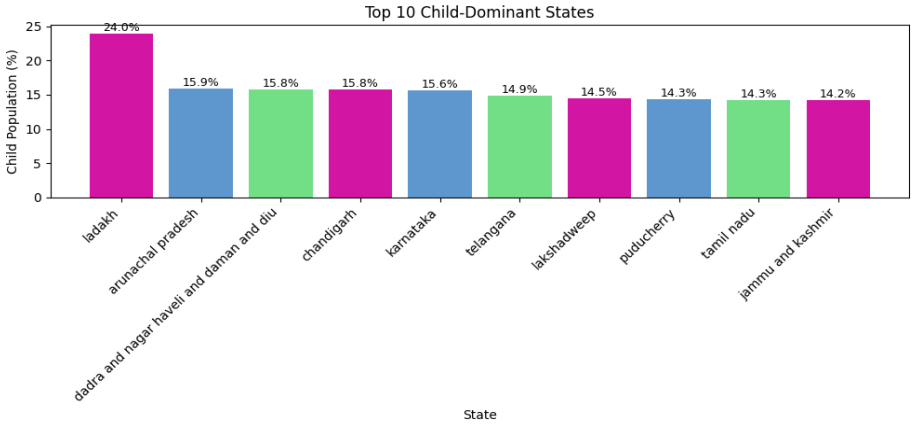
Top 10 Child-Dominant States

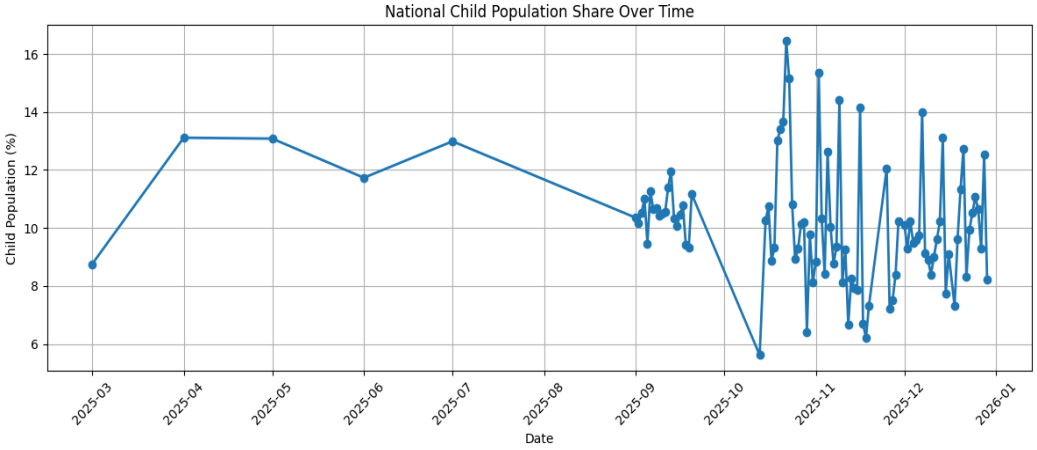
	state_	total_population	demo_age_5_17	demo_age_17_	child_population	adult_pop
0	ladakh	5735	1375	4360	1375	
1	arunachal pradesh	36443	5783	30660	5783	
2	dadra and nagar have	12204	1923	10281	1923	
3	chandigarh	83361	13133	70228	13133	
4	karnataka	1695285	264981	1430304	264981	1
5	telangana	1629908	242259	1387649	242259	1
6	lakshadweep	1176	170	1006	170	
7	puducherry	32763	4696	28067	4696	
8	tamil nadu	2212228	315638	1896590	315638	1
9	jammu and kashmir	407202	57873	349329	57873	

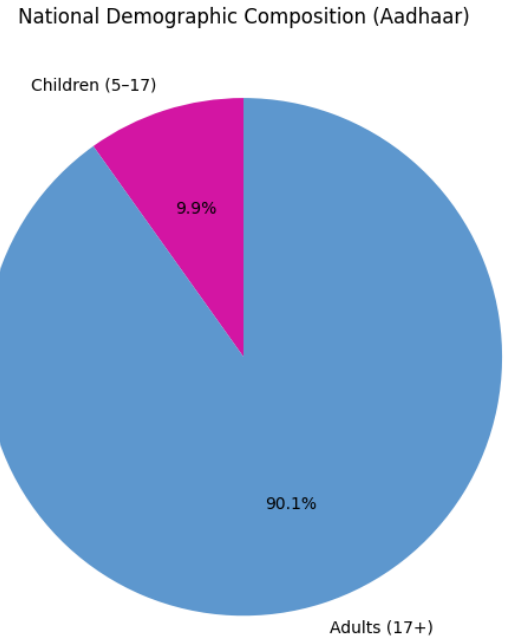
Top 10 Adult-Dominant States

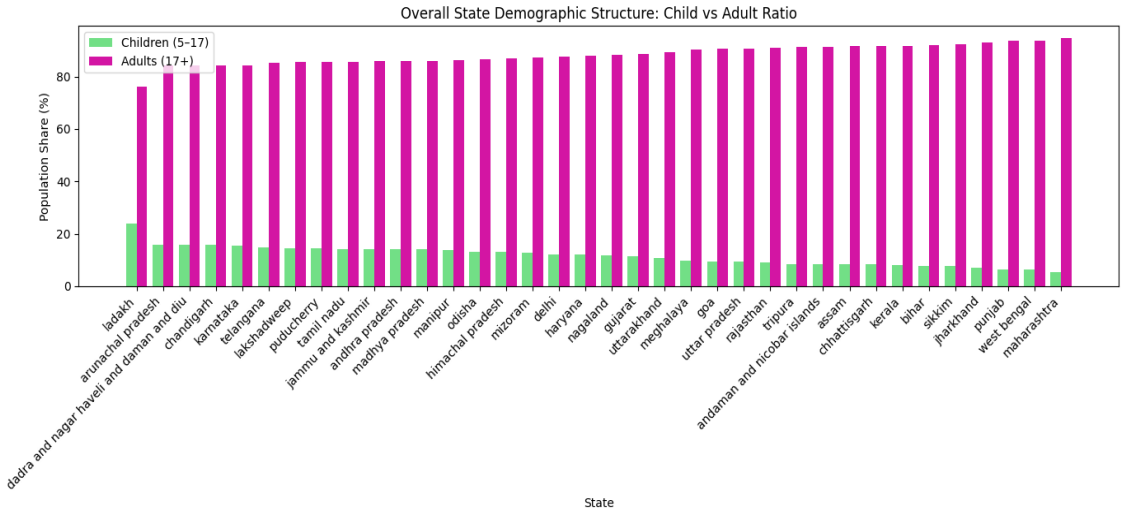
	state_	total_population	demo_age_5_17	demo_age_17_	child_population	adult_population
0	maharashtra	5054602	273322	4781280	273322	4781280
1	west bengal	3872737	242561	3630176	242561	3630176
2	punjab	881895	56866	825029	56866	825029
3	jharkhand	1401189	99376	1301813	99376	1301813
4	sikkim	20340	1555	18785	1555	18785
5	bihar	4814350	380023	4434327	380023	4434327
6	kerala	744952	61064	683888	61064	683888
7	chhattisgarh	2005438	165207	1840231	165207	1840231
8	assam	1012578	84480	928098	84480	928098
9	andaman and	7246	617	6629	617	6629


Top 10 Adult-Dominant States
Adult Population (%)
94.6%
93.7%
93.6%
92.9%
92.4%
92.1%
91.8%
91.8%
91.7%
91.5%
maharashtra
west bengal
punjab
jharkhand
sikkim
bihar
kerala
chhattisgarh
assam
andaman and nicobar islands
State


Top 10 Child-Dominant States
Child Population (%)
24.0%
15.9%
15.8%
15.8%
15.6%
14.9%
14.5%
14.3%
14.3%
14.2%
ladakh
arunachal pradesh
dadra and nagar haveli and daman and diu
chandigarh
karnataka
telangana
lakshadweep
puducherry
tamil nadu
jammu and kashmir
State


National Child Population Share Over Time
Child Population (%)
2025-03
2025-04
2025-05
2025-06
2025-07
2025-08
2025-09
2025-10
2025-11
2025-12
2026-01
Date


National Demographic Composition (Aadhaar)
Children (5–17)
9.9%
90.1%
Adults (17+)


Overall State Demographic Structure: Child vs Adult Ratio
Children (5–17)
Adults (17+)
Population Share (%)
ladakh
arunachal pradesh
dadra and nagar haveli and daman and diu
chandigarh
karnataka
telangana
lakshadweep
puducherry
tamil nadu
jammu and kashmir
andhra pradesh
madhya pradesh
manipur
odisha
himachal pradesh
mizoram
delhi
haryana
nagaland
gujarat
uttarakhand
meghalaya
goa
uttar pradesh
rajasthan
tripura
andaman and nicobar islands
assam
chhattisgarh
kerala
bihar
sikkim
jharkhand
punjab
west bengal
maharashtra
State

Key Insights

1. **Aadhaar is Adult-Dominant:** ~90% adults, ~10% children. Focus shifts from enrollment to maintenance. **Problem:** Data becomes outdated. **Solution:** Periodic update reminders.

2. **Adult-Heavy States:** Maharashtra, Punjab, Kerala show aging and migration effects. **Problem:** Address mismatch & biometric failures. **Solution:** Targeted update drives.

3. **Child-Heavy Regions:** Ladakh, Arunachal Pradesh have high child share. **Problem:** Overload at age 15–18. **Solution:** Predictive planning & early reminders.

4. **Spikes in Child Data:** Caused by enrollment drives. **Problem:** Reactive workload management. **Solution:** Real-time monitoring dashboard.

5. **States Are Not Equal:** One policy does not fit all. **Solution:** State-specific Aadhaar strategies.

Overall Conclusion

- Aadhaar enrollment is mostly complete
  - Main challenge is data freshness
  - Age data helps predict future workload
  - Proactive planning improves UIDAI efficiency

Project Repository:
https://github.com/PAnand04/api_data_aadhar_demographic

UIDAI Data Hackathon 2026

By:
Pallak Anand
Ridhima Tripathi
Harsimar Singh
Aniket Singhal