

Summary and introduction

This dataset records monthly biometric enrollment counts at the district level, split into two age groups (5–17 and 17+), across India. It covers 9 months, 34 states, 729 districts and 6,371 district-month records. Totals are large and similar across age groups (about 30.87M for age 5–17 and 31.25M for age 17+), so overall volume is balanced and the two series move closely together.

Key takeaway — overall shape

- The national totals for the two age groups are nearly equal: 30.9M (age 5–17) vs 31.2M (age 17+), showing broad parity in coverage.
- Coverage: 34 states, 729 districts, 9 months, 6,371 rows — broad geographic and temporal scope supports both national and local insights.

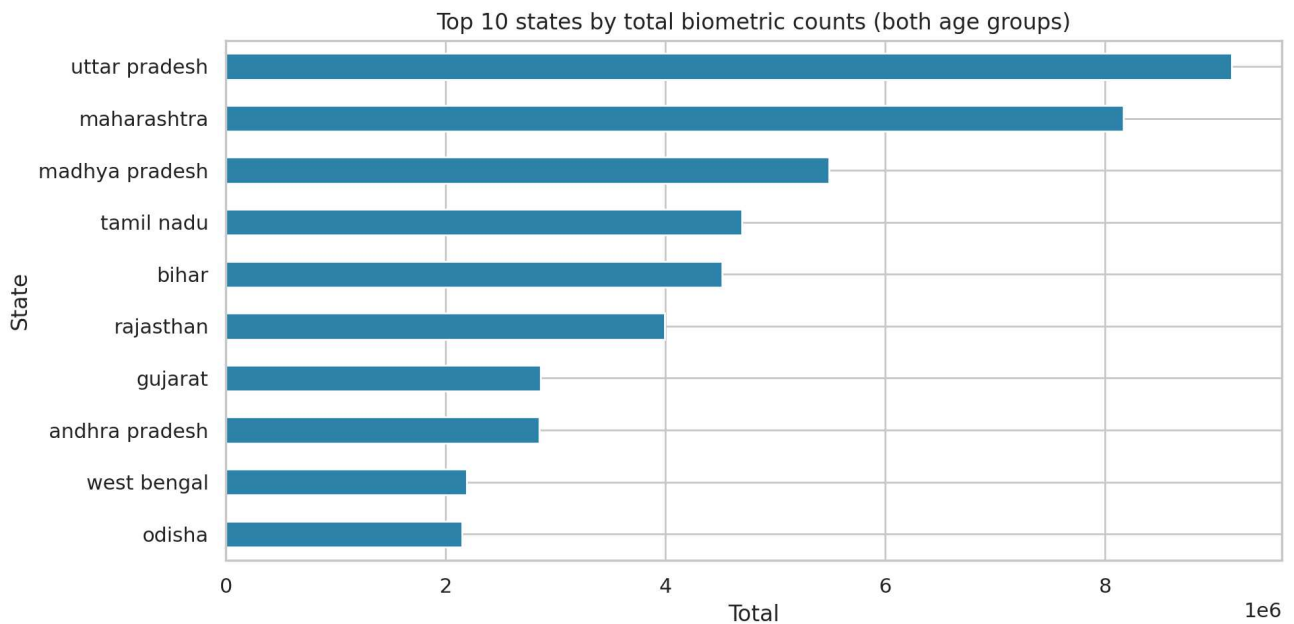
Insight 1 — A small number of states account for a large share

- Finding: The top states concentrate most activity; Uttar Pradesh, Maharashtra and Madhya Pradesh are the biggest contributors.

- Supporting data (top states table):

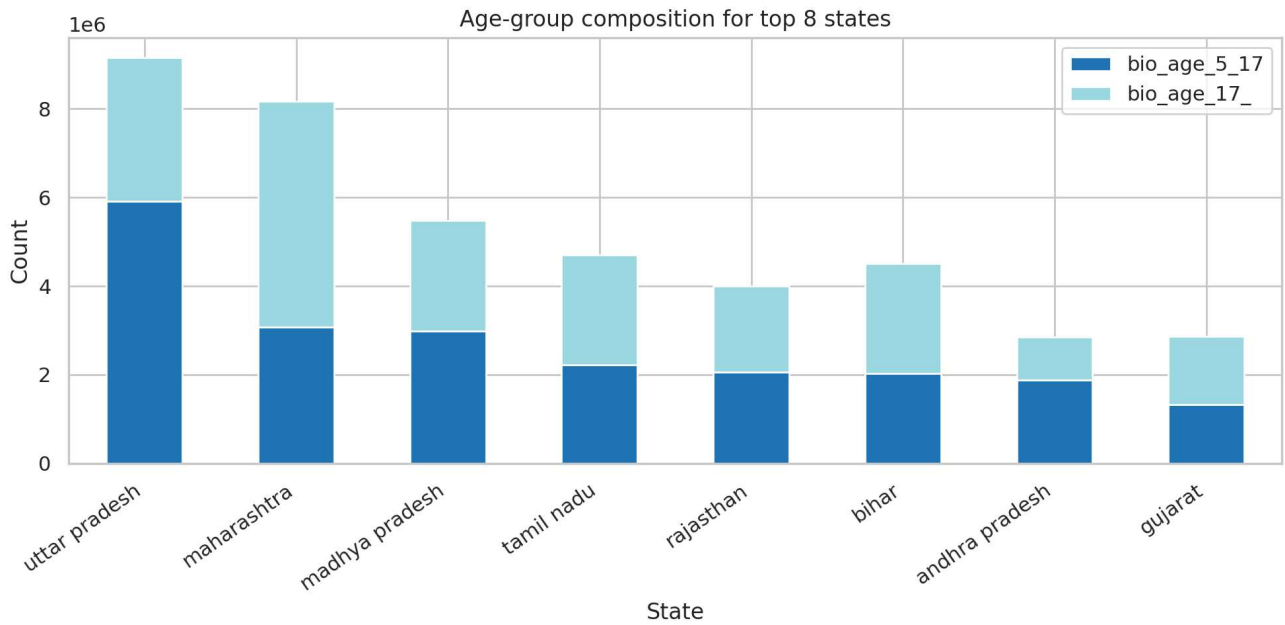
bio_age_5_17 bio_age_17_ total
---: ---: ---: ---:
uttar pradesh 5918481 3232833 9151314
maharashtra 3069596 5098958 8168554
madhya pradesh 2986220 2498978 5485198
tamil nadu 2227251 2470860 4698111
rajasthan 2066630 1927925 3994555
bihar 2035703 2477847 4513550
andhra pradesh 1876951 972795 2849746
gujarat 1324712 1537270 2861982
odisha 1074233 1073407 2147640
west bengal 926636 1261455 2188091

- Visual: Top states bar chart (shows relative dominance of top 10)

[Download](#)

Insight 2 — Age-group mix varies by state

- Finding: While national totals are balanced, some states lean more to one age group. The stacked bar view shows relative composition across the top states.
- Visual: Stacked bar chart for top states (age-group mix)

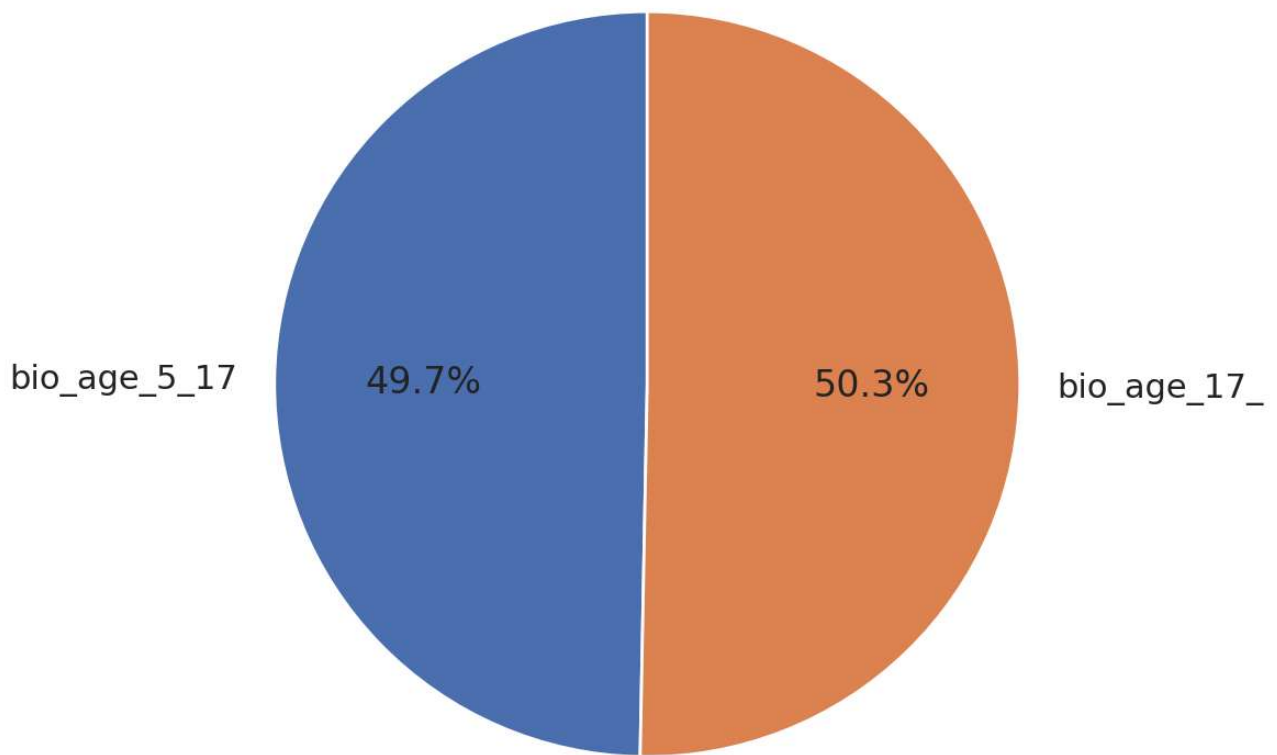
[Download](#)

Insight 3 — National age split is roughly 50/50

- Finding: Overall share between the two age groups is close to equal, confirming the parity noted above.
- Visual: National share pie chart

[Download](#)

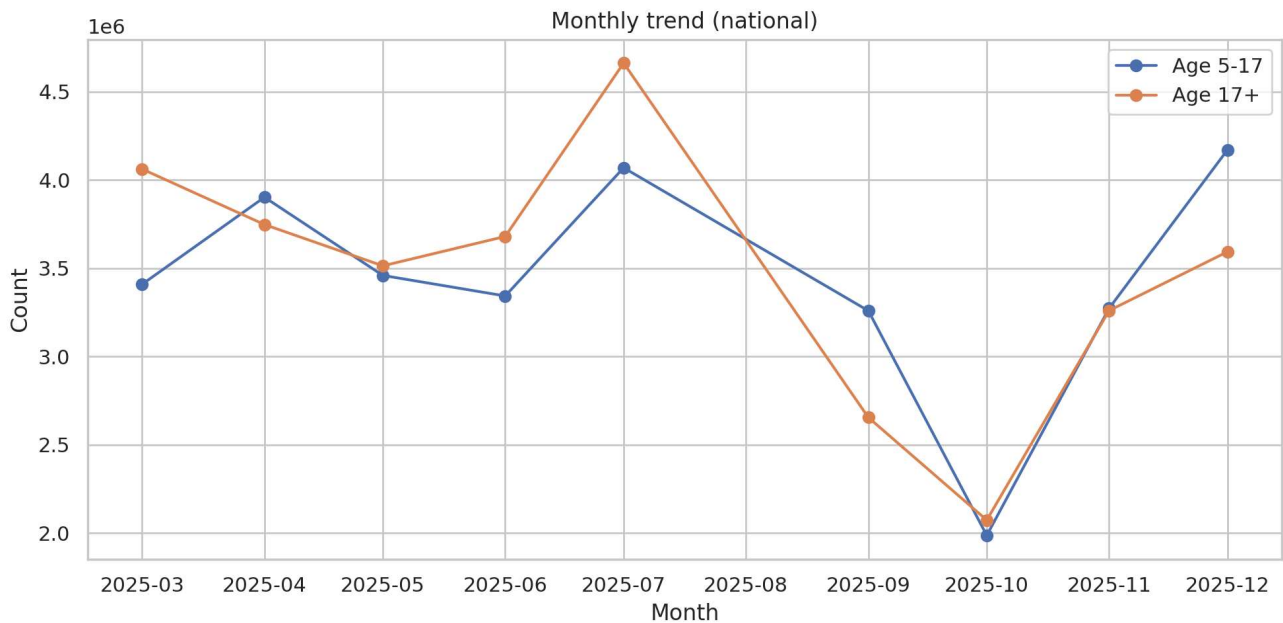
National share by age group



Insight 4 — Trends over time move together

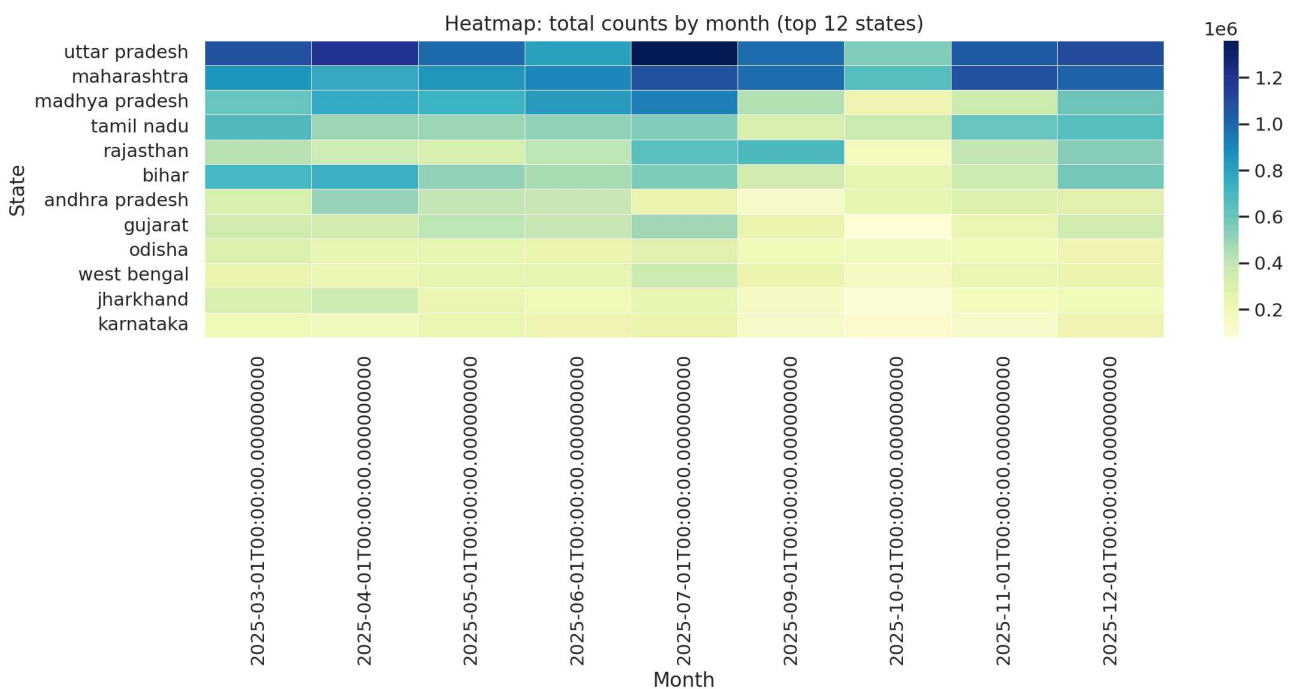
- Finding: Monthly national trends for both age groups follow similar patterns, suggesting common operational drivers (campaigns, seasonality).
- Visual: Monthly national trend line chart

[Download](#)



Insight 5 — State-month “hot spots” and bursts

- Finding: Some states show clear month-to-month bursts or sustained high activity. The heatmap highlights which states and months drive peaks.
- Visual: Heatmap of top 12 states by month




Insight 6 — District-level concentration: a few districts dominate

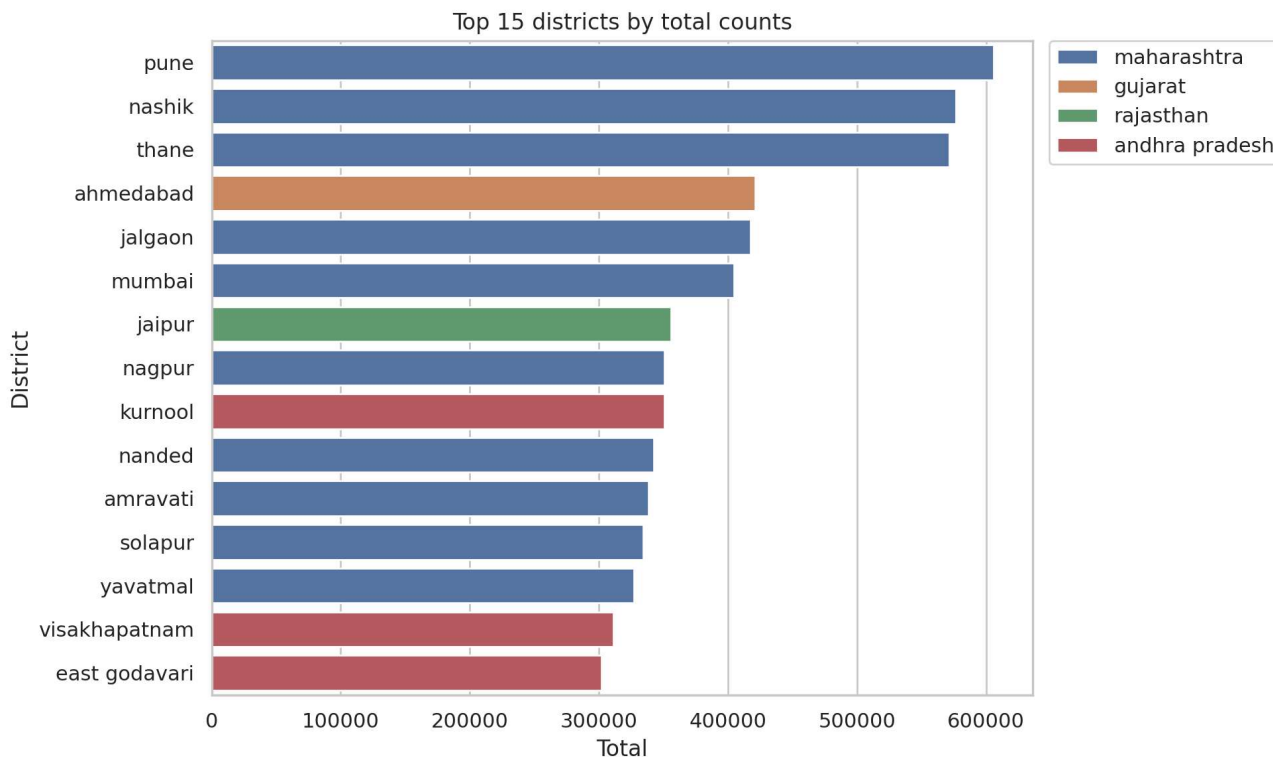
- Finding: Top districts account for very high totals; Pune, Nashik, Thane and Ahmedabad appear among the largest individual district totals.

- Supporting data (top districts table):

state_norm	district_resolved	bio_age_5_17	bio_age_17_	total
---	---	---	---	---
maharashtra	pune	283480	322282	605762
maharashtra	nashik	213100	363506	576606
maharashtra	thane	202896	368377	571273
gujarat	ahmedabad	166849	254323	421172
maharashtra	jalgaon	148853	268531	417384
maharashtra	mumbai	119615	284744	404359
rajasthan	jaipur	158293	197591	355884
maharashtra	nagpur	112851	238072	350923
andhra pradesh	kurnool	242992	107641	350633
maharashtra	nanded	124730	217810	342540

- Visual: Top 15 districts bar chart

 Download

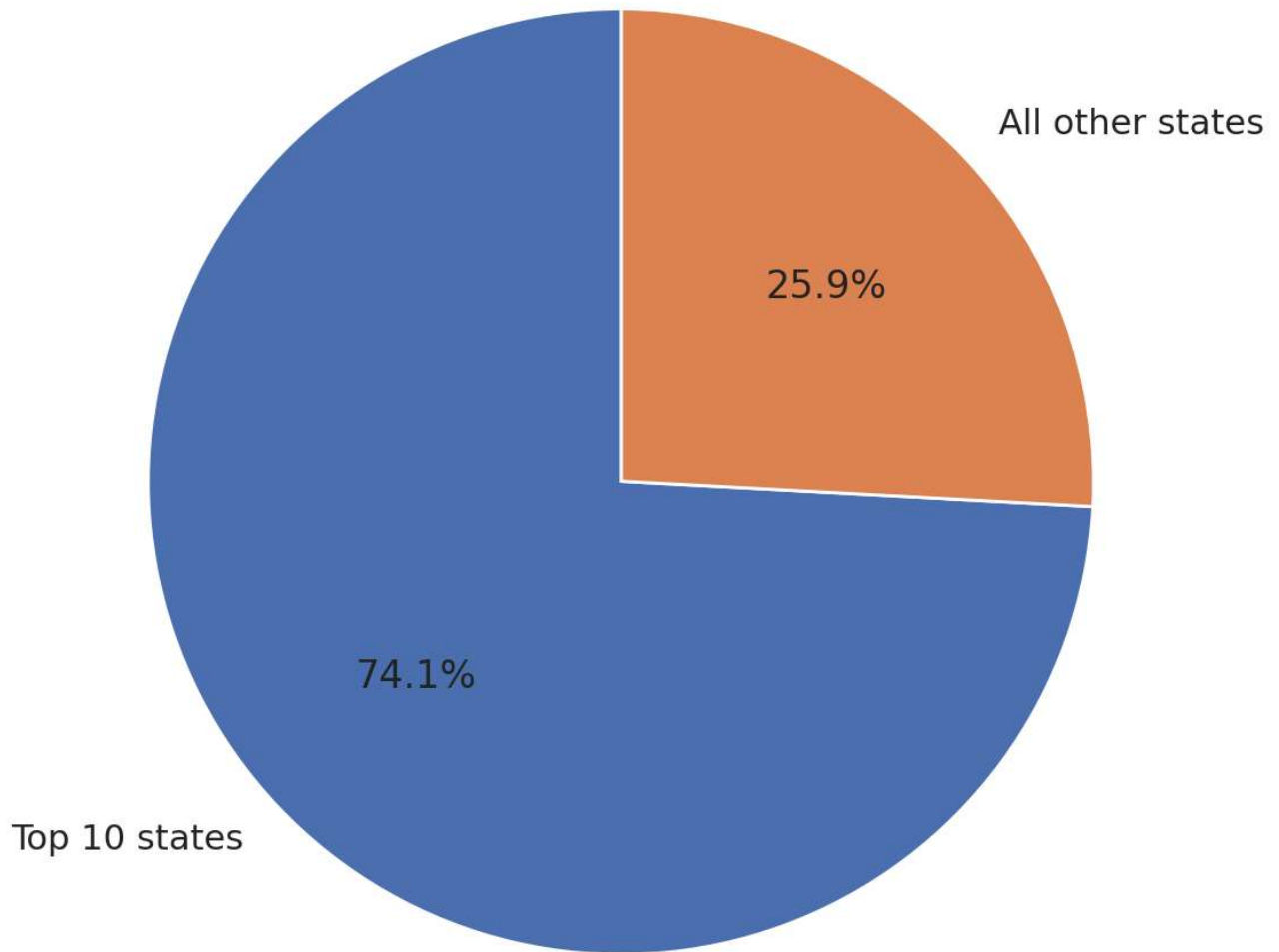


Insight 7 — Concentration metric: top 10 states vs rest

- Finding: The top 10 states capture a large portion of national totals; the concentration pie highlights that imbalance.
- Visual: Pie — top 10 states vs rest

[Download](#)

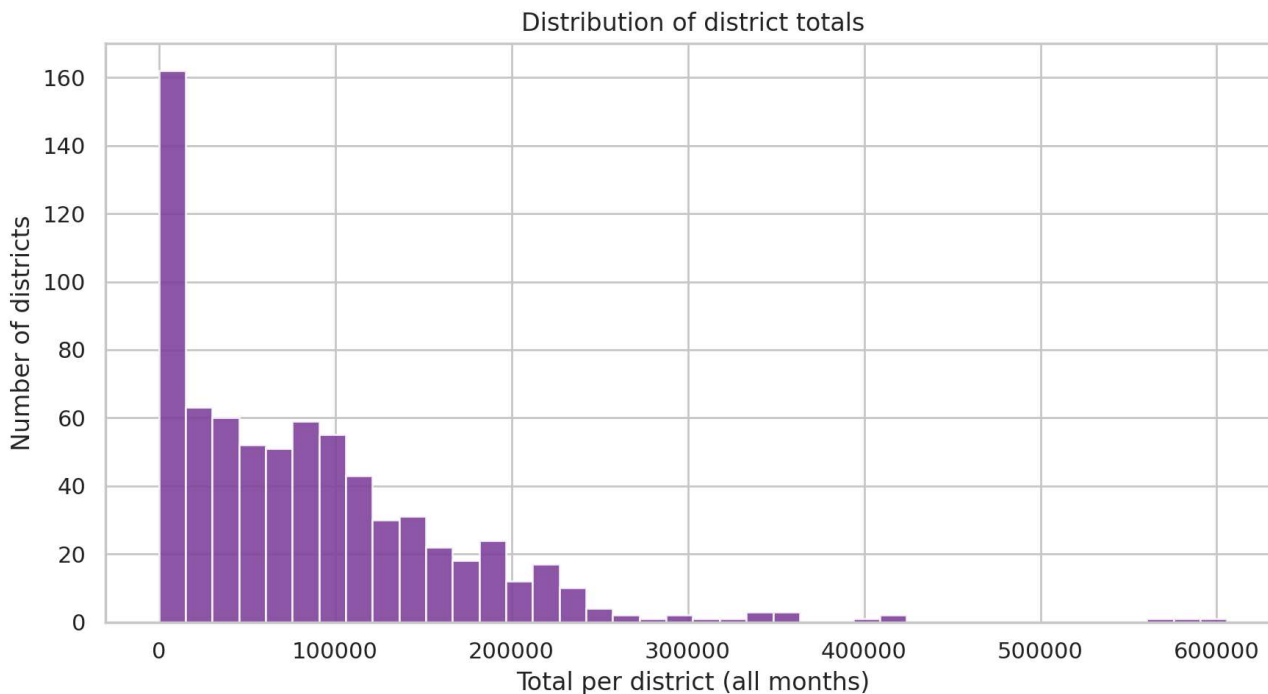
Concentration: Top 10 states vs rest



Insight 8 — Distribution is long-tailed across districts

- Finding: Many districts have moderate totals while a small number have very large totals — typical long-tail. This matters for resource allocation and monitoring.
- Visual: Histogram of district totals

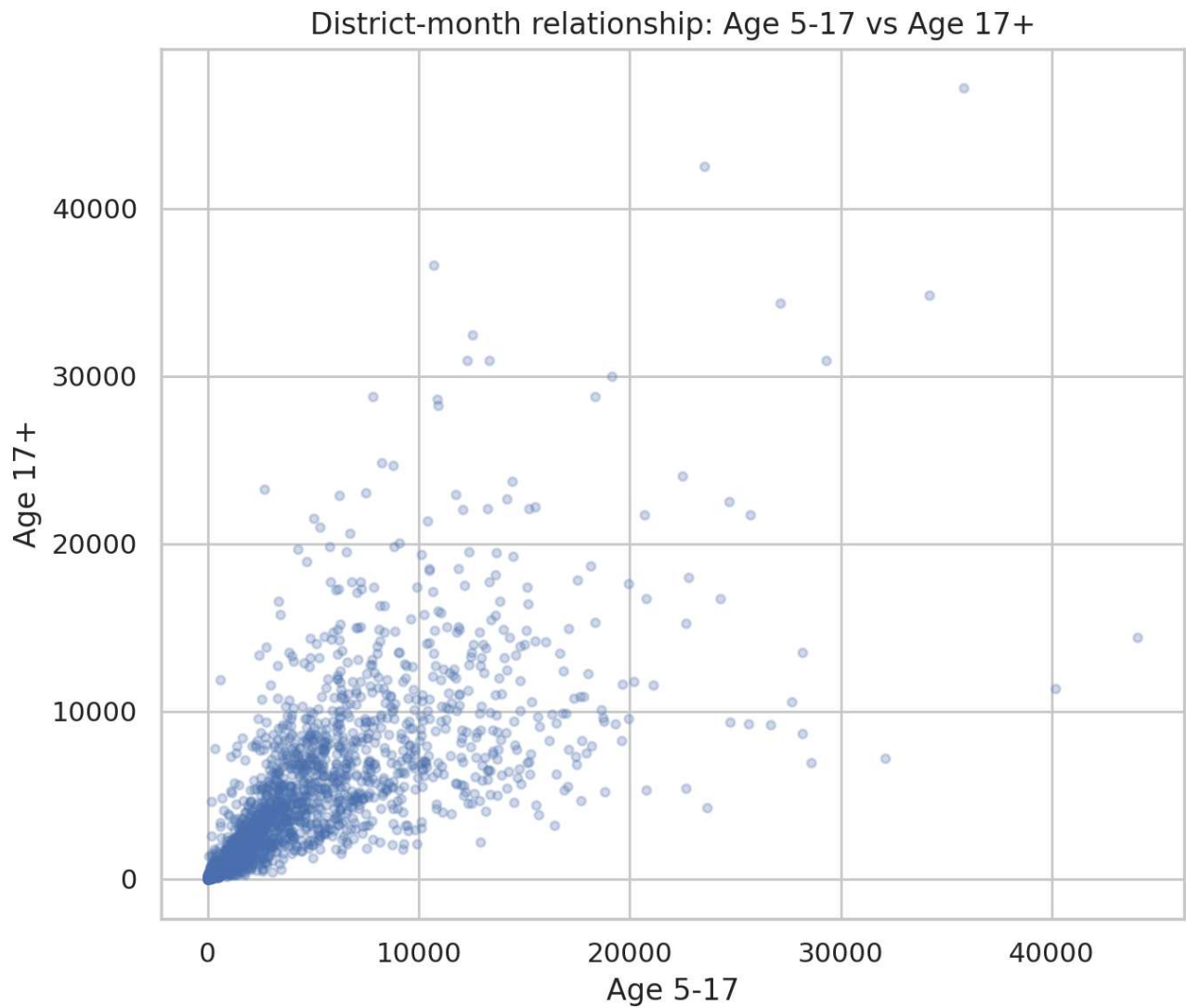
[Download](#)



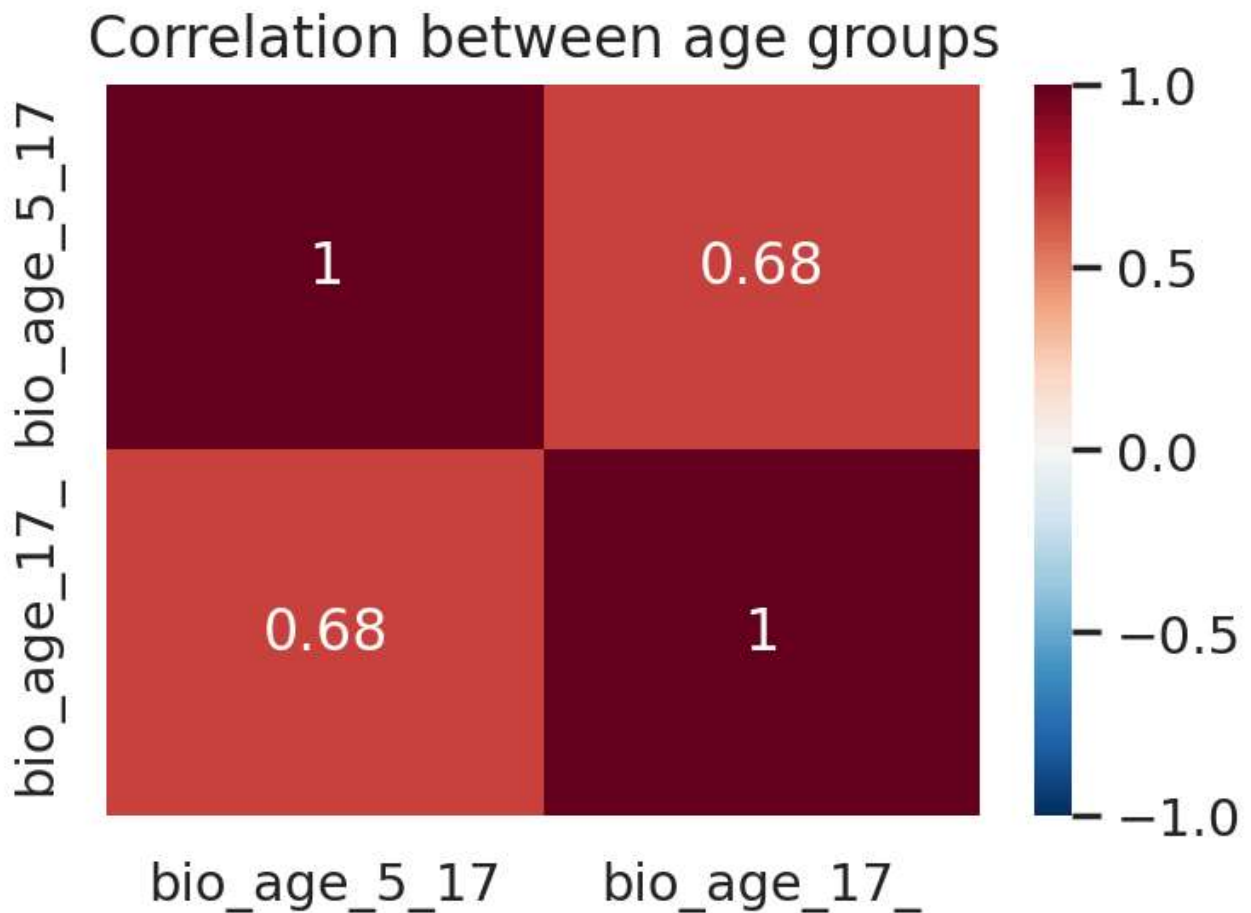
Insight 9 — Age groups are highly correlated

- Finding: District-month records show a strong positive relationship between the two age-group counts, indicating the same on-ground activities affect both groups.
- Visual: Scatter (sample) and correlation heatmap

[Download](#)



[Download](#)



What went wrong earlier (brief)

- The initial code attempted to convert non-existent columns named value_1/value_2. The actual numeric columns are bio_age_5_17 and bio_age_17_. After switching to the correct names, numeric conversion and visualization proceeded successfully.

Final summary (concise)

- Nationwide biometric counts are large and nearly balanced between 5–17 and 17+ age groups.
- A small set of states and districts drive most of the volume (Uttar Pradesh, Maharashtra, Madhya Pradesh, Pune/Nashik/Thane, etc.).
- Monthly patterns show synchronized movement across age groups, and the district-level distribution is long-tailed.

- Recommendation for readers: focus monitoring and resource allocation on the identified top states/districts while keeping an eye on month-to-month bursts shown in the heatmap.

Conclusive note: the data supports both national-level summaries (balanced age coverage) and targeted operational actions (top states/districts concentrate the workload).