CHAPTER 8

Migration and Urbanization (Ethiopia focus plus global lens)

Aynalem Adugna, October 2025

Suggested citation: Aynalem Adugna, Chapter 8. Migration and Urbanization (Ethiopia focus plus global lens, www.EthioDemographyAndHealth.org, October 2025.

CONTENT

1. **Population Censuses**

- Ethiopia's census history, mandates, uses, challenges (hard-to-reach, displacement, pastoralist mobility).

- International standards (UN Principles & Recommendations), de-facto vs de-jure, digital census trends.

2. **Civil Registration and Vital Statistics (CRVS)**

- Births, deaths, causes of death, marriages/divorces; legal basis and institutional roles in Ethiopia.

- Global CRVS completeness benchmarks; ICD coding; linking CRVS to health systems and national ID.

3. **Household Sample Surveys**

- DHS, MICS, HICES/WMS, LSMS-ISA, labor force and intercensal surveys in Ethiopia.

- Sampling frames, weights, non-sampling errors; harmonization with international survey programs.

4. **Administrative Data Systems**

- Health (HMIS), education (EMIS), social protection, national ID, tax, voter rolls.

- Data quality (coverage, duplication), interoperability, and how countries leverage admin data between censuses.

5. **Population Registers & Sample Registration Systems**

- Feasibility and phased pathways in Ethiopia; pilots for continuous demographic measurement.

- Nordic-style registers and India's SRS as global reference models.

6. **Migration & Displacement Data**

- Internal migration, pastoralist mobility, seasonal labor; refugee/IDP data (camp/settlement vs non-camp).

- International sources: UN DESA, IOM, UNHCR; measuring emigration and remittances.

7. **Geospatial, Remote Sensing & Gridded Population Datasets**

- Using satellite imagery, building footprints, night-lights, and gridded surfaces (e.g., WorldPop, GHS).

- Ethiopia use cases: small-area estimation, service catchments, drought/conflict monitoring.

8. **Big Data & Alternative Digital Traces**

- Mobile CDRs, social media, merchant data—promise and pitfalls.

- Governance, bias, representativeness, and ethical constraints; examples from LMIC settings.

9. **International Compendia & Modelled Estimates**

- UN World Population Prospects, World Bank, UNICEF, WHO, IHME.

- How global estimates reconcile national data and how Ethiopia's submissions flow into these series.

10. **Data Quality Assessment & Triangulation**

- Completeness, coverage error, age heaping, displacement effects, post-enumeration surveys.

- Triangulating census/CRVS/surveys/admin data; plausibility checks and reconciliation.

11. **Indirect Demographic Methods**

- Brass techniques, child survivorship, orphanhood, maternal mortality sisterhood, model life tables.

- When and how to use them in Ethiopia's data landscape; uncertainty communication.

12. **Small-Area Estimation & Spatial Demography**

- SAE and Bayesian approaches to woredas/kebeles; integrating survey + covariates + geospatial layers.

- Equity-focused mapping for underserved and mobile populations.

13. **Data Integration & Interoperability**

- Standards (DDI, SDMX), master frames, entity resolution, linkage (CRVS-ID-HMIS).

- Architectures for a national data ecosystem aligned with global interoperability practices.

14. **Ethics, Privacy & Data Governance**

- Consent, confidentiality, re-identification risks in sparse settings; data-sharing agreements.

- International norms (GDPR-inspired principles) adapted to Ethiopian law and practice.

15. **Dissemination, Open Data & Visualization**

- National portals, microdata access, documentation; accessibility & multilingual dissemination.

- Global exemplars; reproducible pipelines and codebooks for researchers and planners.

16. **Institutional Arrangements & Capacity Building**

- Roles of the national statistical office and sector ministries; financing and staffing.

- Partnerships with universities, donors, and international agencies; training and sustainability.

17. Use Cases for Policy & Planning

- Health, education, urban planning, disaster risk, elections, SDGs—Ethiopian scenarios with global parallels.

- From data to decisions: evidence uptake, M&E, and feedback loops.

1) Population Censuses

1.1 What a census is (and why it matters). A population and housing census is a complete count of all persons and dwellings in a country on a specified reference date. It underpins decisions on electoral apportionment, fiscal transfers, and the placement of schools, clinics, and roads. Internationally, the United Nations' Principles and Recommendations for Population and Housing Censuses (Rev.4) set the standard for scope, definitions, topics, operations, quality assurance, and dissemination.

1.2 Ethiopia's census history (overview). Ethiopia has completed three modern Population and Housing Censuses (PHC): 1984 (first modern PHC), 1994 (second PHC), and 2007 (third PHC). The 2007 PHC used staggered fieldwork (May and November) due to operational/security constraints in the Somali Region. Official totals are listed below.

Table 1. Ethiopia—Official population counts by census year

| Census year	Total population	Male	Female	Notes
1984	39,868,572			First modern PHC; coverage challenges documented
1994	53,477,265			Second PHC; improvements but continued undercount concerns
2007	73,918,505	37,296,657	36,621,848	Third PHC; May & Nov enumeration; delays in Somali Region

Sidebar: De facto vs. De jure residence (Ethiopia examples)

Definitions. De facto counts people where they are physically found on census night. De jure counts them at their usual place of residence, regardless of where they happen to be on the reference date.

Why it matters in Ethiopia. In settings with seasonal migration, pastoral mobility, university students, domestic workers, and displaced populations, the choice affects small-area totals, service planning, and fiscal allocations. Ethiopia's 2007 census used staggered fieldwork in some regions; clarity about the applied residence concept and reference date helps users interpret results.

• **Pastoralist households (Afar, Somali).** A de facto approach during a dry-season census might miss households that have temporarily moved to wet-season grazing areas; a de jure approach anchored to 'usual residence' with careful listing rules can mitigate omissions.

• **University students (Addis Ababa, Mekelle, Hawassa).** Students physically present in dorms on census night would be counted there under de facto; under de jure, they are counted at their usual household if that better reflects long-term service needs.

• **Seasonal agricultural workers (Oromia, SNNP).** Harvest-time de facto counts can inflate populations in high-demand farm zones; de jure rules help maintain accurate home-woreda totals for budgeting and service delivery.

• **Internally displaced persons (IDPs) and refugees.** Emergency movements can shift large populations; consistent, well-communicated residence rules, plus coordinated frames with humanitarian registries (IOM/UNHCR), are essential to reduce double-counting or omission.

Figure 1. Ethiopia—Total population by census year

Ethiopia: Total Population by Census Year
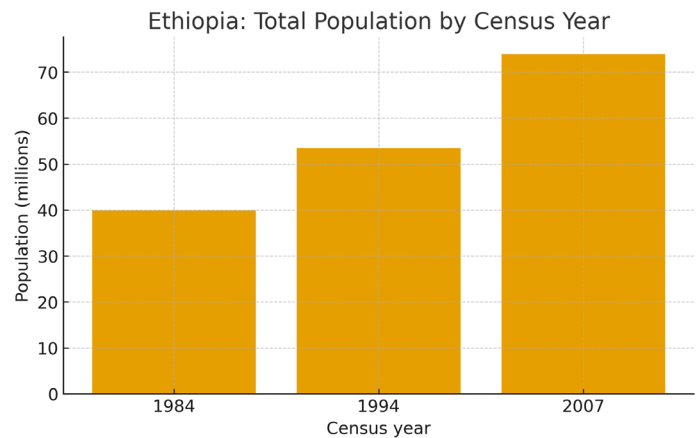
1.3 International standards and Ethiopia-specific operational challenges. Core international standards include universal enumeration on or around a fixed reference date, transparency in choosing de facto vs. de jure residence, rigorous mapping and pilot testing, and a post-enumeration survey (PES) to measure coverage error. In Ethiopia, common challenges have included access limitations in insecure areas, high mobility and displacement, pastoralist enumeration, rapid peri-urban growth, and data-quality artifacts (age heaping, sex-ratio anomalies).

1.4 Future directions for Ethiopia's next census. Priorities include a digital census with GPS-enabled enumeration areas and real-time field monitoring; tailored strategies for pastoralist and displaced populations; a robust, published post-enumeration survey (PES); stronger integration with CRVS and administrative data for coverage evaluation; and transparent dissemination of microdata and documentation.

Selected references

- United Nations (2025). Principles and Recommendations for Population and Housing Censuses, Rev.4.

- Ethiopia Population and Housing Census Commission (2007). Summary and Statistical Report of the 2007 Population and Housing Census.

- UNECA/AfDB national digital census toolkits and guidance (various).

2) Civil Registration and Vital Statistics (CRVS)

2.1 What CRVS is (and why it matters). Civil Registration and Vital Statistics (CRVS) is the continuous, permanent, compulsory, and universal registration of births, deaths (with cause), marriages, and divorces—followed by the production of vital statistics. CRVS provides legal identity and routine, high-value statistics for planning and accountability.

2.2 Ethiopia's legal and institutional framework. Ethiopia established its modern CRVS via the 2012 Registration of Vital Events & National Identity Card Proclamation and a 2017 amendment. A UNECA-led comprehensive assessment in 2018–2019 helped set a modernization roadmap. Core institutions: VERA and regional bureaus (registration and certificates), Ethiopian Statistics Service (vital statistics), Federal Ministry of Health and facilities (event notification, medical certification of cause of death), and courts/police for legal and medico-legal functions.

Table 1. Legal and institutional milestones

Year	Instrument/Activity	Reference	Purpose/Notes
2012	Registration of Vital Events & National Identity Card Proclamation	Proc. No. 760/2012	Establishes compulsory CRVS; links to national ID
2017	Amendment Proclamation	Proc. No. 1049/2017	Addresses service delivery gaps; clarifies roles/procedures
2018–2019	Comprehensive national CRVS assessment	UNECA-led	Guides modernization roadmap

Table 2. CRVS core events and minimum data items (examples)

Vital event	Minimum statistical items (examples)
Birth	Child name, sex, date/place of birth; parents' names/IDs/residence; informant; registrar; registration date; certificate number
Death	Name, sex, date/place of death, age/date of birth, usual residence; medical certification with ICD; registration date; certificate number
Marriage	Spouses' names, ages, IDs; date/place of marriage; witnesses; officiant; registration date; certificate number
Divorce	Parties' names/IDs; date/place of divorce; court/authority; registration date; certificate number

Table 3. Institutional responsibilities (summary)

Institution	Primary CRVS role (summary)
VERA / Regional bureaus	Civil registration (birth, death, marriage, divorce); certificate issuance; registries maintenance
Ethiopian Statistics Service (ESS)	Compilation and dissemination of vital statistics; standards and coordination
Federal Ministry of Health & facilities	Event notification; medical certification of cause of death; HMIS linkage
Courts / Police	Legal decrees (divorce); medico-legal death certification in specific cases
National ID Program (NIDP)	Interoperability with CRVS for identity verification and deduplication

2.3 International standards (what "good" looks like). UN guidance emphasizes universal, continuous registration, unique identification, ICD-compliant medical certification of cause of death, and routine dissemination of vital statistics. Quality should be tracked via completeness, timeliness, accuracy, and internal consistency across sources (CRVS, surveys, censuses, administrative data).

2.4 Performance and data quality in practice. Countries track birth and death registration completeness, certificate issuance times, and the share of medically certified deaths with ICD codes. For Ethiopia, users should consult the latest ESS/VERA bulletins, UNICEF/World Bank indicators for birth registration, and WHO notes on cause-of-death data.

2.5 Future directions for Ethiopia's CRVS. End-to-end digital registration with offline capability; integration with the National ID program and health information systems; nationwide medical certification of cause of death with ICD training and audit; regular subnational vital statistics publication; and public dashboards with metadata and data dictionaries.

Selected references:

- Proclamation on the Registration of Vital Events & National Identity Card (Proc. No. 760/2012).

- Amendment Proclamation No. 1049/2017 (Vital Events Registration & National Identity).

- UNECA Global CRVS Group: Comprehensive assessment of Ethiopia's CRVS (2018–2019).

- UN ESCAP: Principles and Recommendations for a Vital Statistics System.

- WHO guidance on medical certification and ICD coding for cause of death.

- UNICEF & World Bank indicators on birth registration completeness (country pages).

- WHO country data portal—cause-of-death and mortality statistics (Ethiopia).

3) Household Sample Surveys

3.1 What they are (and why they matter). Household sample surveys are statistically designed studies that collect standardized information from a probability sample of households and individuals. They generate high-value, disaggregated indicators between censuses—on fertility, mortality, health and nutrition, poverty and consumption, labor, and agriculture—and provide the backbone for SDG and sector monitoring when CRVS or administrative systems are still maturing.

3.2 Major survey programs used in Ethiopia (with a global frame). Ethiopia participates in several international survey programs while running national rounds tailored to policy needs. The table below summarizes the most common platforms and what they measure.

*Table 1. Major household survey programs in Ethiopia—scope and topics*

Program	Implementers (Ethiopia)	Typical frequency	Primary units	Core topics
Demographic and Health Survey (DHS)	Ethiopian Statistics Service (ESS) & MoH (with ICF support)	≈5 years; mini rounds between	Households, women/men, children	Fertility, mortality, MCH, FP, nutrition, WASH, HIV, malaria
LSMS-ISA / Ethiopia Socioeconomic Survey (ESS)	ESS & World Bank	Panel waves every 2–3 years	Households & individuals	Income, consumption, agriculture, labor, shocks, assets (panel)
HICES / WMS	ESS	≈3–5 years	Households	Income, consumption/expenditure, poverty, welfare indicators
Labor Force & Employment Surveys	ESS (with MoLSA)	≈1–2 years / periodic	Individuals & households	Employment/unemployment, informality, earnings, time use (select rounds)
Malaria Indicator Survey (MIS)	MoH / EPHI / partners	≈3–5 years (as needed)	Households & eligible individuals	Malaria prevalence, nets, treatment, vector control
PMA (Performance Monitoring for Action)	Addis Ababa Univ. / MoH / partners	≈1–2 years / rapid	Women of reproductive age, facilities (modules)	Family planning access/use, service delivery

3.3 Sampling design essentials (how precision is achieved). Most rounds use a two-stage, stratified cluster design and publish weights. Key elements are summarized below.

Table 2. Sampling design essentials

Design element	Best-practice note
Frame	Up-to-date list of enumeration areas (EAs) from the latest census or master sample; verify coverage for new settlements and displaced populations.
Stratification	Urban/rural and regional strata to ensure representativeness and domain precision.
Clustering	Two-stage design is common: select EAs (PPS) then households systematically within EAs.
Sample size	Driven by key indicators (e.g., modern contraceptive prevalence), desired precision, design effect, and budget.
Weights	Account for selection probabilities, nonresponse adjustment, and post-stratification to population controls.
Nonresponse	Minimize via revisits, callbacks, and field supervision; document rates and patterns.
Questionnaires	Standard modules + Ethiopia-specific modules (e.g., pastoralism, shocks, IDPs) with careful translation/back-translation.
Ethics	Informed consent, privacy, interviewer safety; sensitive modules handled with protocols.

Figure 2. Illustrative sample allocation by domain

Illustrative allocation by domain
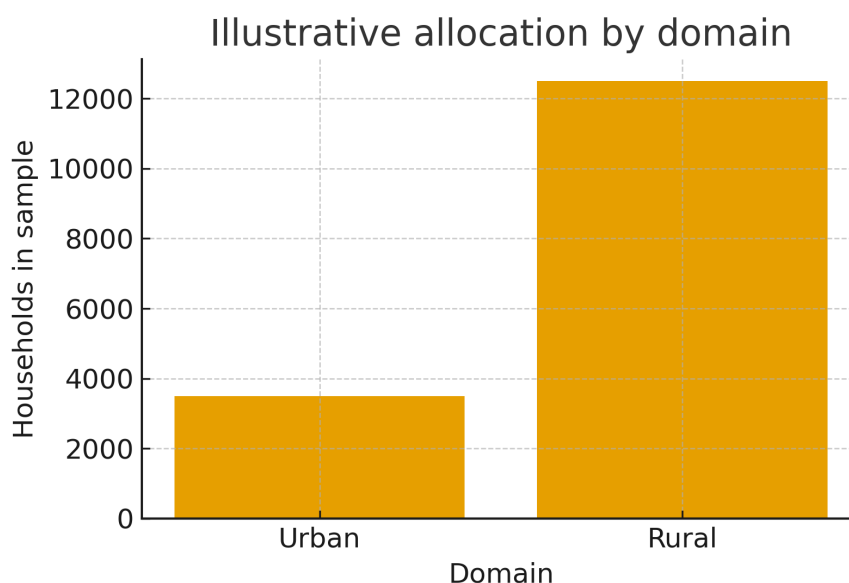
3.4 Data quality and comparability. Quality hinges on frame currency, interviewer training, field supervision, and digital data capture with validation checks. For comparability over time, maintain consistent indicator definitions and document any questionnaire or universe changes. For equity analyses, ensure sample sizes allow precision at regional and urban/rural levels; consider oversampling special populations where feasible.

Table 3. Common data-quality risks and mitigations

Risk	Mitigation or practice
Coverage error	Use updated frames; supplement with satellite listing/mapping; special operations for pastoral & displaced populations.
Measurement error	CAPI with range/logic checks; pilot tests; interviewer training & certification.
Recall bias	Shorten recall windows; use event calendars and visual aids.
Non-sampling error documentation	Field paradata dashboards; publish response rates, substitutions, and protocol deviations.
Comparability over time	Track questionnaire changes; harmonize variable names/labels; maintain codebooks & version control.
Linkages	Enable geo-referencing (with displacement) for merging with climate, market, or service-access layers (proper anonymization).

3.5 Indicators typically derived (examples).

Table 4. Indicators commonly derived from Ethiopia's survey platforms

Domain	Illustrative indicators
Demography	Total fertility rate, age-specific fertility, child mortality (U5MR/IMR), adolescent birth rate
Maternal & child health	ANC visits, skilled birth attendance, immunization coverage, nutrition (stunting, wasting)
Water & sanitation	Improved water and sanitation access, safely managed services (where measured)
Socioeconomic	Poverty rates (consumption-based), inequality, food security, education levels
Labor	Employment/unemployment rates, labor force participation, informal employment share
Agriculture (LSMS-ISA)	Crop/livestock production, input use, land tenure, shocks, productivity

Sidebar: Panel vs. cross-sectional surveys (Ethiopia examples)

• Panel (e.g., LSMS-ISA/ESS): Follows the same households across waves to study dynamics—poverty transitions, shocks, and productivity.

• Repeated cross-sections (e.g., DHS/MIS): New, representative samples each round to track national and regional trends with consistent indicators.

Selected references (for users of the site):

- Ethiopian Statistics Service (ESS): DHS/EMDHS, HICES/WMS, labor statistics documentation.

- World Bank & ESS: Ethiopia Socioeconomic Survey (LSMS-ISA) technical reports and questionnaires.

- MoH/EPHI: Ethiopia Malaria Indicator Survey (EMIS) instruments and reports.

- UN & ICF international survey program manuals (DHS, MICS, LSMS).

4) Administrative Data Systems

4.1 What they are (and why they matter). Administrative data are records generated as part of delivering public services—health visits, school enrollment, vital events, taxes, social protection, and more. Unlike surveys, they are continuous and potentially exhaustive, enabling high-frequency indicators for planning, budgeting, and accountability when quality and coverage are strong.

4.2 Ethiopia's administrative data landscape (selected systems). The table below summarizes commonly used systems, custodians, reporting flows, and core data elements. Coverage can vary across regions and over time; integration and standards are critical to maximize value.

Table 1. Selected administrative systems in Ethiopia—custodians, flows, and content

System	Primary custodian	Flow	Reporting frequency	Core data
HMIS (DHIS2)	Federal Ministry of Health (FMoH)	Facilities -> Woreda -> Zone -> Region -> Federal	Monthly (most), weekly for select programs	Service volumes, morbidity, staffing, logistics
EMIS	Ministry of Education	Schools -> Woreda/Zone -> Region -> Federal	Annual census + term modules	Enrollment, completion, staffing, infrastructure
Civil Registration (VERA) & National ID	VERA / NIDP	Local registry -> Woreda -> Region -> Federal	Continuous	Births, deaths, marriages, divorces; ID issuance
Tax/VAT	MoF / Revenues	Enterprises/individuals -> Federal databases	Monthly/quarterly	Tax accounts, VAT returns, payroll
Social Protection (PSNP)	MoLSA & partners	Kebele/woreda registries -> Federal	Continuous + periodic recert.	Beneficiary registries, payments, compliance
Agriculture (MoA)	Ministry of Agriculture	Kebele/woreda -> Region -> Federal	Seasonal/annual	Production, inputs, livestock, extension
Voter Roll	NEBE	Kebele/woreda -> Region -> Federal	Periodic (electoral cycles)	Eligible voters, polling places

4.3 Data quality: dimensions and metrics. Key dimensions and how to track them are summarized below.

Table 2. Data quality dimensions and metrics (admin data)

Dimension	Illustrative metric(s)
Coverage/completeness	Registered events ÷ expected events; facility catchment coverage; % duplicate IDs
Timeliness	Median days from event to entry; % reports on time
Accuracy/validity	Outlier checks; logical consistency (age/sex/program rules); verification rates
Consistency	Trends over time; cross-source concordance with surveys/census
Integrity	Audit trails; role-based access; error-correction protocols
Disaggregation	Availability by sex, age, region, urban/rural, disability (where collected)

4.4 Interoperability: identifiers, exchange, and metadata. Adopting shared identifiers, exchange standards, and metadata ensures data are linkable, discoverable, and reusable.

Table 3. Interoperability layers and standards (examples)

Layer	Standards/artifacts (examples)
Identifiers	National ID, facility ID, school ID, health worker ID; privacy-preserving linkage keys
Data exchange	APIs; SDMX for indicators; HL7 FHIR for clinical resources; CSV/Parquet for bulk
Metadata	DDI (surveys/admin tables), DCAT; code lists and data dictionaries
Geospatial	Admin boundaries, EA codes, facility/school coordinates; versioned gazetteers

4.5 Priority linkages and policy use-cases.

Table 4. High-value linkages and what they enable

Linkage	Illustrative policy use
CRVS + HMIS	Birth/death notification flows; reconcile facility deaths with civil registration; improve maternal and neonatal statistics
CRVS + National ID	Unique identity anchoring vital events; de-duplication and service eligibility checks
EMIS + Population (census/grids)	Compute school-age denominators; plan classroom & teacher allocation
HMIS + EMIS	School health programs; vaccination catch-up via school rosters
Social Protection + Population/ID	Targeting, payments, grievance redress; shock-responsive registries
Tax/VAT + Business Registry	Formalization, MSME targeting, revenue analytics

4.6 Governance and roles. Clear accountability reduces duplication and protects privacy.

Table 5. Governance (RACI) — who leads what

Area	Lead roles (R=Responsible, A=Accountable, C=Consulted)
Data standards	ESS (A), Sector custodians (R), Digital agency (C), Legal/Privacy office (C)
Interoperability platform	Digital agency (A/R), Sector custodians (C), ESS (C)
Quality assurance & audit	ESS (A/R), Sector custodians (R), Supreme audit/Inspectorates (C)
Data sharing & access	Legal/Privacy office (A), Custodians (R), ESS (C)
Open data & documentation	ESS (A/R), Custodians (R), Communications (C)

4.7 Risks and mitigations.

Table 6. Selected risks, impacts, and mitigations

Risk	Impact	Mitigation
Security/access constraints	Data gaps from non-reporting areas	Alternative reporting, remote collection, modeling
ID quality/duplicates	Double counting; poor linkage	ID verification, dedup algorithms, periodic audits
Data protection gaps	Privacy risk; re-identification	DPIAs, minimization, access controls, legal agreements
Legacy systems	Inconsistent formats; siloed data	ETL pipelines, standards adoption, phased migration
Staff turnover	Loss of know-how	SOPs, documentation, training, mentoring

Sidebar: From siloed admin files to a national data ecosystem

Start with shared IDs and a master geo frame, publish data dictionaries and APIs, stand up a small interoperability platform, and run regular data-quality audits. Prioritize a few high-value linkages (e.g., CRVS↔HMIS, EMIS↔Population) before scaling to others. Ensure privacy by design and transparent governance.

Selected references (for users of the site):

- Ethiopian Statistics Service (ESS) and sector ministries: HMIS/DHIS2, EMIS, and administrative data manuals.

- UN Statistics Division: Handbook on the Use of Administrative Data for Statistics; SDMX & DDI guidance.

- WHO/UNICEF: Health administrative data quality toolkits; HMIS quality audit guidelines.

- World Bank & UNECA: Interoperability, CRVS–ID linkages, and administrative data modernization resources.

5) Population Registers & Sample Registration Systems

5.1 Why these systems (and how they fit). Population registers and sample registration systems (SRS) complement censuses and CRVS. Registers provide a continuously updated person-level infrastructure for administration and statistics, while SRS offers timely vital rates from a representative sample until CRVS/registers are complete.

5.2 Global registers: models to learn from. A variety of models exist; key features are summarized below.

Table 1. Population register models—global references and features

Model	Key features
Nordic model (e.g., Sweden, Norway, Denmark)	Central population register linked via unique personal ID; legal obligations to update events.
Hybrid model (register + CRVS)	CRVS remains legal source; register serves as master person index, synchronized via APIs.
Residency-based municipal registers	Local registers aggregate to a national backbone; strong address frames are key.

5.3 Ethiopia: Sample Registration System (SRS) — design elements. Key design choices for a nationally representative SRS in Ethiopia are summarized below.

Table 2. SRS design—sample, operations, and outputs

Element	Specification (illustrative)
Goal	Continuous measurement of births/deaths (with causes) in a representative sample; bridge gaps while CRVS matures.
Sample	Multi-stage stratified selection (region × urban/rural; pastoralist strata); ≈200–500 EAs illustrative.
Operations	Resident enumerators or periodic visits; verbal autopsy for out-of-facility deaths; physician/algorithm-coded causes.
Data flows	Event books/mobile apps → woreda → region → national; monthly/quarterly syncs; QA and audits.
Outputs	Annual birth/death rates; U5MR; cause-of-death distributions; subnational estimates for larger domains.

5.4 Phased pathway for Ethiopia (register + SRS).

Table 3. Phased roadmap—from pilots to scale

Phase	Core actions (illustrative)
Phase 1 (2025–2026)	Legal basis; data model & metadata; link CRVS with National ID; pilot SRS in 2–3 diverse regions.
Phase 2 (2026–2027)	Roll out unique ID to priority sectors; expand SRS; publish annual vital rates.
Phase 3 (2027–2029)	Scale register coverage; enable routine record linkage; strengthen cause-of-death certification; retire SRS as systems mature.

5.5 Benefits, risks, and mitigations.

Table 4. Benefits, risks, and mitigations—registers vs SRS

Approach	Benefits	Risks	Mitigations
Population register	Up-to-date person-level backbone; supports service delivery & statistics; small-area estimates.	Complex startup; high governance/privacy bar; requires reliable ID/address systems.	Phased rollout; strong legal safeguards; DPIAs; independent oversight; audits.
SRS	Timely vital rates without full CRVS; lower cost than census; strengthens cause-of-death data.	Sample-only; sustained supervision; generalization limits.	Embed QA; integrate with CRVS; invest in verbal autopsy quality.

5.6 Minimum data model for a population register.

Table 5. Person-level data layers and fields (illustrative)

Layer	Fields (illustrative)
Identifiers	Unique personal ID (linked to National ID), alternative keys; checksum rules.
Core person	Name(s), sex, date/place of birth, citizenship, usual residence, parents' identifiers.
Lifecycle events	Birth, death (date/place/cause link), marriage, divorce, migration in/out, ID issuance/expiry.
Linkage fields	Address codes, geo identifiers, facility/school IDs; timestamps; source system; status flags.
Security	Access roles, consent flags (where applicable), audit trails; retention/archiving rules.

Sidebar: When to use SRS vs. when to push for a full register

Use SRS to quickly obtain national and regional birth/death rates and causes when CRVS completeness is low; invest in the register when unique ID services, address systems, and CRVS interoperability are maturing. A pragmatic path is to launch SRS pilots while building the legal/technical foundations of the register.

Selected references (for users of the site):

- Nordic population register documentation (e.g., Statistics Sweden, Statistics Norway).

- India Sample Registration System (SRS) design and vital rates bulletins.

- UN Principles and Recommendations for Population and Housing Censuses (for frames & linkages).

- UN Guidelines on Statistical Business Registers and population registers linkages.

- WHO verbal autopsy standards; ICD coding guidance for cause of death.

6) Migration & Displacement Data

6.1 Why these data matter. Migration and displacement reshape Ethiopia's population distribution, labor markets, service demand, and vulnerability profiles. Robust, ethically collected data are essential for planning education and health services, designing safety nets, coordinating humanitarian response, and engaging the diaspora.

6.2 Core sources—national and international. The table below summarizes key sources, custodians, frequency, and what each provides.

Table 1. Data sources for migration and displacement (Ethiopia + global)

Source	Custodian (Ethiopia/Global)	Frequency	What it provides
Population & Housing Census	ESS	≈10 years	Lifetime migration, recent migration, place of previous residence; household roster; locality codes
Demographic & Health Survey (DHS/EMDHS)	ESS/MoH (with ICF)	≈5 years (+ minis)	Migration module (selected rounds), household roster demographics
Labor Force Survey (LFS)	ESS (with MoLSA)	periodic	Place of residence history, labor mobility, commuting (where measured)
HICES / WMS	ESS	≈3–5 years	Household moves, shocks; can support welfare analysis by migrant status
Administrative: Passports, work permits, visas	Immigration/NIDP/MFA	continuous	Applications/issuances—partial proxy for emigration
Administrative: National ID / CRVS	NIDP / VERA	continuous	Place of birth/residence, events; potential for movement tracking (with safeguards)
IOM Displacement Tracking Matrix (DTM)	IOM	quarterly/biannual	IDP/returnee stocks & flows; site profiles; reasons/needs
UNHCR Refugee Statistics	UNHCR	monthly/annual	Refugee/asylum-seeker stocks & flows by country of origin; camp/non-camp
UN DESA International Migration	UN DESA	biennial/quadrennial	Modeled migrant stocks by origin/destination; global comparability
World Bank Remittance Data	World Bank/KNOMAD	annual	Remittance inflows/outflows; costs

Figure 1. Illustrative displacement trends—IDPs and refugees hosted

Illustrative: Ethiopia—displacement stocks over time
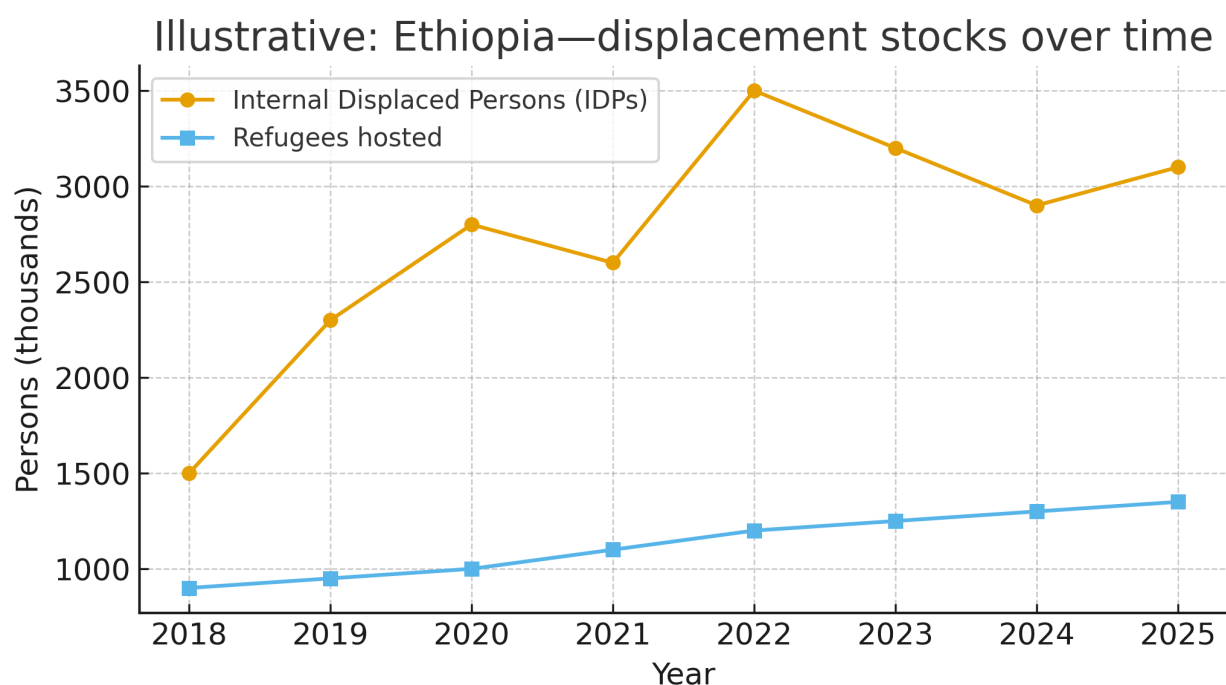
6.3 Indicators and definitions (speak the same language).

Table 2. Key indicators and standard definitions (short)

Indicator/term	Definition (short)
Lifetime migrant	Person living in a region/place different from their place of birth.
Recent migrant	Moved residence within a specified period (e.g., last 5 years or 12 months).
Net migration	In-migrants minus out-migrants for a region over a period.
Origin–destination flows	Counts or rates of moves between regions/kebeles; often from census/surveys.
Refugee/Asylum-seeker	Cross-border displacement categories under international protection (1951 Convention).
Internally Displaced Person (IDP)	Forced to flee home but remain within national borders.
Remittances	Transfers from emigrants to residents, measured via balance-of-payments and surveys.

6.4 Internal migration: origin–destination analysis. Origin–destination matrices capture flows between regions/woredas and help planners anticipate pressure on services. The heatmap below shows an illustrative matrix for communication.

6.5 Data quality, protection, and ethics.

Table 3. Common challenges and practical mitigations

Challenge	Mitigation
Coverage (hard-to-reach, mobile)	Use mixed frames (satellite-informed), outreach via local structures; allow proxy reporting with verification.
Timing & recall bias	Short recall windows; event calendars; triangulate survey/census/admin records.
Duplication & definition mismatch	Harmonize definitions; deduplicate with IDs/linkage keys; be explicit on usual residence.
Insecurity & nonresponse	Adaptive fieldwork, remote data capture, partner data (IOM/UNHCR); model-based small-area estimates.
Ethics & protection	Do-no-harm protocols, sensitive data handling, aggregation thresholds, consent and risk assessment.

6.6 Methods toolbox for Ethiopia.

Table 4. Analytical methods and how they help

Method	Use in Ethiopia context (examples)
Small-area estimation (SAE)	Combine survey/census with covariates (nightlights, roads, conflicts) for woreda-level indicators.
Origin–destination modeling	Gravity/spatial interaction models; Bayesian flows with priors from geography and networks.
Capture–recapture	Estimate hidden or duplicated populations by overlapping lists (with strong safeguards).

Record linkage	Privacy-preserving linkage to track movements between admin systems (ID, HMIS, EMIS).
Nowcasting with geospatial data	Use remote sensing and mobility proxies to update displacement between rounds.

6.7 Policy use-cases (turning data into decisions).

Table 5. What decision-makers can do with these data

Use-case	What data enable
Urban planning & services	Plan schools, clinics, water points for migrant-dense neighborhoods.
Labor & skills	Match training and job creation to migrant/returnee profiles.
Health & protection	Vaccination catch-up, MHPSS for displaced persons, GBV services.
Disaster risk & resilience	Target drought/conflict hotspots; pre-position supplies; support returns.
Remittances & finance	Leverage diaspora flows; reduce costs; financial inclusion initiatives.

Sidebar: Definitions at a glance (Ethiopia examples)

• Migrant (internal): Person whose usual residence has changed from one administrative area to another.
• Migrant (international): Person who changes country of usual residence (short-/long-term definitions vary).
• Refugee / Asylum-seeker: Person outside their country with a well-founded fear of persecution (under international law) / seeking recognition.
• Internally Displaced Person (IDP): Person forced to flee home but who remains within Ethiopia's borders.
• Returnee: Formerly displaced person who returns to their area of origin or resettles elsewhere.

Selected references (for users of the site):

- Ethiopian Statistics Service (ESS): Census, DHS/EMDHS, HICES/WMS, LFS documentation.

- IOM Displacement Tracking Matrix (DTM) Ethiopia: site assessments and emergency tracking.

- UNHCR Refugee Statistics (Ethiopia country page).

- UN DESA International Migration Stock and related methodological notes.

- World Bank/KNOMAD remittance datasets and migration briefs.

7) Geospatial, Remote Sensing & Gridded Population Datasets

7.1 Why geospatial data matter for population analysis. Remote sensing and gridded population products allow planners to move beyond administrative averages and see where people actually live—supporting service catchments, emergency response, and equity mapping between census rounds.

7.2 Major datasets used in Ethiopia (with global sources).

Table 1. Key geospatial and gridded population datasets

Dataset / Layer	Native resolution (approx.)	Notes (Ethiopia-relevant usage)
WorldPop (top-down)	≈100 m	Disaggregate census counts using covariates (night lights, buildings, land cover) via ML; annual estimates; constrained/unconstrained variants.
WorldPop (bottom-up)	≈100 m (varies)	Household micro-surveys + building footprints + spatial covariates; model up to national totals; useful where census is outdated.
GHS-POP (EC-JRC)	≈100 m	Population grids from census + GHSL built-up; epoch years; harmonized across countries.
HRSL (Meta/World Bank)	≈30 m (settlement mask), pop grid ≈30–100 m	Buildings detected from high-res imagery; population distributed to buildings.
LandScan	≈1 km	Ambient population (24-hour average presence) using multiple covariates.
WSF & GHS-SMOD (settlement layers)	10–100 m	Built-up extent/settlement classification from radar/optical sources.
VIIRS/DMSP Night Lights	≈500 m–1 km	Human activity proxy useful as covariate; saturation & blooming issues in bright cores.
Microsoft/Meta Building Footprints	~meter-level polygons	Open building footprints for Ethiopia (coverage varies); useful to constrain population to structures.

7.3 Methods toolbox—how numbers get into the grid.

Table 2. Methods used to create and apply population grids

Method	What it does / Why it helps
Areal weighting	Spread admin totals uniformly within polygons; baseline for comparison.
Dasymetric mapping	Constrain or weight by land cover/settlement masks/buildings to avoid uninhabited areas.
ML-based top-down disaggregation	Use covariates (night lights, roads, slope, buildings) and models to apportion counts into grid cells.
Bottom-up modeling	Predict densities from sampled micro-areas, then scale to national totals; powerful where census data are sparse.
Small-area estimation (Bayesian/SAE)	Fuse surveys + geospatial covariates to predict indicators at kebele/woreda scales.
Accessibility/catchment modeling	Network travel times to clinics/schools; allocate demand and plan service coverage.

7.4 Ethiopia use-cases and decision support.

Table 3. Common Ethiopia use-cases with examples

Use-case	Example for Ethiopia
Health service planning	Compute realistic catchments for facilities; estimate population within 30/60/120-minute travel bands.
Education & WASH	Map school-age children and water access gaps; prioritize fast-growing settlements.
Pastoralist & mobile groups	Use settlement layers + seasonal water points/NDVI to anticipate shifts and plan outreach.
Crisis & displacement	Combine DTM/UNHCR sites with gridded baselines to monitor hosting pressure.
Urban growth & SDG 11	Track built-up expansion (WSF/GHSL) and densification; guide land-use decisions.
Equity mapping	Overlay poverty/health indicators with high-resolution population to target underserved kebeles.

7.5 Quality, ethics, and communication.

Table 4. Common issues and good practice

Issue	Good practice
Uncertainty	Provide confidence intervals or sensitivity bands where available; note model years vs. reference years.
Bias & omission	Buildings-only masks can miss informal settlements or nomadic camps; validate with local knowledge.
Versioning	Document dataset version/date; results can change when providers update models or inputs.
Privacy	Avoid re-identification; aggregate appropriately; follow data-sharing agreements.
Triangulation	Cross-check with census/admin totals, survey totals, and local registers before decisions.

Sidebar: Choosing the right grid for Ethiopia

• For national planning and SDG tracking, 100 m products (WorldPop, GHS-POP) offer a good balance of resolution and stability.

• For urban analysis or settlements, augment with building footprints and settlement layers (WSF, GHSL).

• For mobile/pastoral contexts, pair grids with seasonal covariates (NDVI, water points) and field validation.

Selected references (for users of the site):

- WorldPop project documentation (top-down and bottom-up methods).

- EC-JRC Global Human Settlement Layer (GHS-POP, GHSL, GHS-SMOD).

- Facebook/Meta HRSL and Microsoft Building Footprints (Ethiopia coverage).

- LandScan ambient population methods documentation.

- WHO/UN guidelines on geospatial data integration for health services planning.

8) Big Data & Alternative Digital Traces

8.1 What counts as "big data" for population work. Digital exhaust from telecoms, apps, sensors, and satellites can complement traditional sources. Used responsibly, these signals provide timely, small-area insights—especially valuable between censuses and while CRVS matures.

8.2 Sources and signals.

Table 1. Common big-data sources and what the signals reveal

Source	What the signal can show
Mobile call detail records (CDRs)	Tower-to-tower movements, population presence, disaster evacuation patterns, commuting proxies
App-based mobility (aggregated)	Relative changes in visits to points of interest; home/work time; travel radius
Social media (public posts/ads)	Geo-tagged activity; audience size estimates; sentiment topics (limited representativeness)
Search and web activity (aggregated)	Interest in symptoms, services, prices; early signals of shocks (requires careful validation)
Payments/merchant data	Spending patterns by location and category; resilience of local markets
Satellite proxies (night lights, SAR)	Economic activity, infrastructure damage, flood extent; complements ground data

Figure 1. Illustrative scores: coverage vs. bias by data source

Illustrative scores: coverage vs. bias by data source
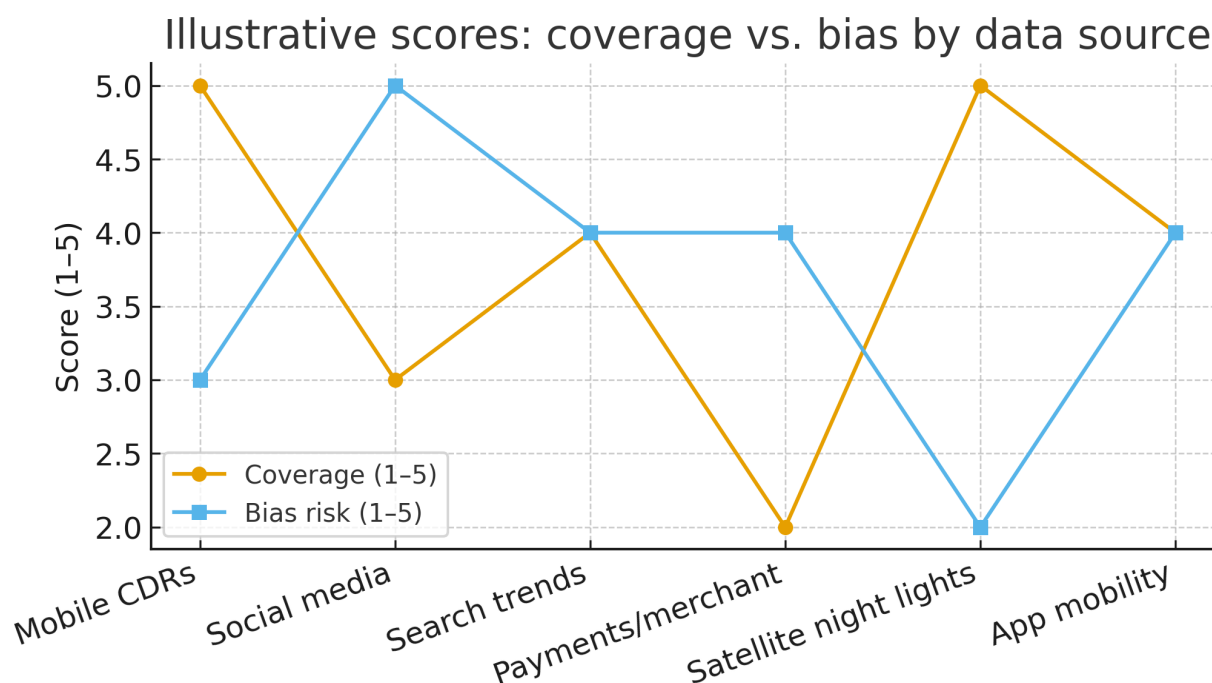
8.3 Strengths and limitations.

Table 2. Strengths and limitations by source

Source	Strengths	Limitations/risks
Mobile CDRs	Near-real-time, broad spatial coverage	Operator access & privacy; uneven SIM ownership; rural tower spacing
App mobility	Timely, granular movement patterns	Smartphone bias; opt-in requirements; urban skew
Social media	Rapid thematic insights	Non-representative users; bot/noise; platform policy changes
Search trends	Early indicators for demand	Ambiguous intent; platform sampling changes; urban bias
Payments/merchant	Objective spending data	Financial inclusion gaps; proprietary access
Night lights / SAR	Independent physical proxy	Cloud/saturation issues; indirect measure of people

8.4 Methods toolbox (how to analyze safely and usefully).

Table 3. Methods for analysis, bias correction, and privacy

Method	Use in practice
Anomaly detection	Identify deviations from seasonal baselines in mobility or spending indices
Origin–destination matrices	Infer flows and effective catchments using CDR or app traces
Downscaling & fusion	Combine big data with census/admin for small-area estimates
Domain adaptation	Correct for selection bias via weighting/transport models
Differential privacy	Add noise or apply thresholds to protect individuals
Nowcasting	Short-term estimates of demand or displacement before surveys arrive

Figure 2. Illustrative mobility index around a disruption (weekly)

Illustrative mobility index around a disruption (weekly)
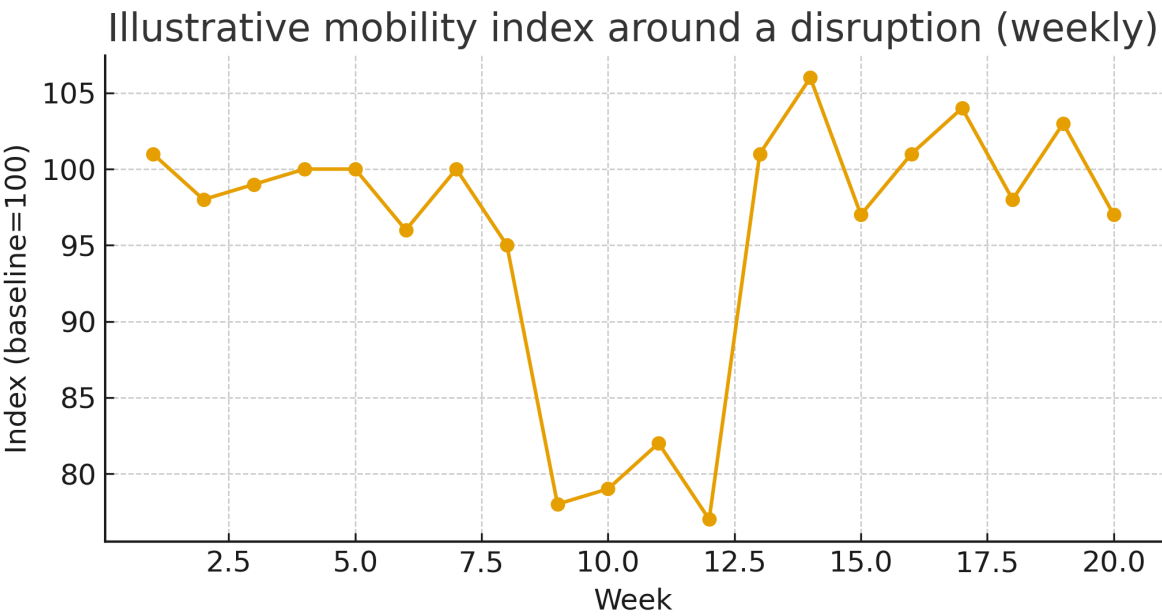
8.5 Governance, ethics, and partnerships.

Table 4. Governance and ethics checklist

Area	Good practice checklist
Legal basis & contracts	Data sharing agreements with privacy-by-design clauses; defined retention and permitted uses
Data protection impact assessment (DPIA)	Assess re-identification and harm risks; mitigation plans
Aggregation & thresholds	Release only aggregated/thresholded indicators; suppress small cells
Equity & representativeness	Measure and correct biases (device, gender, rural/urban); publish limitations
Local value & transparency	Co-develop methods with ESS/ministries; document algorithms and versioning
Sustainability	Plan for ongoing access fees, staffing, and infrastructure

8.6 Ethiopia-focused use-cases.

Table 5. What decision‑makers can do with alternative digital traces

Use-case	What to do with the data
Shock monitoring (conflict/drought)	Track mobility drops and returns; prioritize humanitarian corridors
Health services	Plan mobile clinics using day/night population patterns; monitor vaccination catch-up
Urban transport & planning	Estimate peak flows and congestion; test scenarios for BRT or minibus routes
Food security & markets	Monitor local spending proxies and night-lights to flag market stress
Tourism & migration	Use anonymized roaming statistics to gauge cross-border flows
Program targeting	Fuse big data with survey/admin to refine beneficiary maps (with safeguards)

8.7 Validation & evaluation (trust but verify).

Table 6. Practical framework to validate big-data indicators

Component	Practical approach
Face validity	Patterns align with known events (holidays, harvest, road closures)
Ground truthing	Compare with survey/census/admin; field checks where feasible
Back-testing	Would the method have detected known shocks earlier?
Error decomposition	Quantify selection vs. measurement error; report uncertainty
Reproducibility	Version data and code; publish methods summaries

Sidebar: Minimum privacy safeguards for population analytics

• Use aggregated indicators with thresholds (e.g., k-anonymity ≥ 10) and noise where needed.
• Avoid exact coordinates for sensitive groups; use buffers or hex bins.
• Publish limitations and uncertainty; invite independent review.
• Ensure data access is auditable and time-bound.

Selected references (for users of the site):

- UN Global Working Group on Big Data for Official Statistics—guidance and case studies.

- GSMA & ITU resources on mobile big data for development and disaster response.

- WorldPop/Flowminder methods on mobility and population estimation from telecom data.

- OECD & World Bank briefs on alternative data, privacy, and policy use.

9) International Compendia & Modelled Estimates

9.1 What these sources are (and why they matter). International compendia reconcile diverse national inputs into comparable, time-consistent series. They are essential for cross-country work, SDG reporting, and long-term planning—while national series remain the authoritative reference for local decisions.

9.2 Major compendia (who publishes what).

Table 1. Key compendia, custodians, and content

Compendium	Custodian	What it provides
UN World Population Prospects (WPP)	UN DESA/Population Division	Population levels & projections (by age/sex), fertility, mortality, migration—standard global series.
World Bank World Development Indicators (WDI)	World Bank	Curated indicators from national & international sources; population, socio-economic, health, poverty.
UNICEF global databases	UNICEF	Child & maternal indicators; under-5 and adolescent metrics; MICS integration.
WHO GHO/GHE	WHO	Health indicators; mortality causes; service coverage; modeled time series.
IHME Global Burden of Disease (GBD)	IHME	Modeled estimates for mortality, morbidity (DALYs, YLLs, YLDs), risk factors.
UN DESA Migration & Urbanization	UN DESA	International migrant stocks/flows; World Urbanization Prospects.
UNESCO UIS	UNESCO	Education indicators, enrolments, completion; SDG4.
FAO, ILOSTAT, OECD, UNSD SDG	FAO/ILO/OECD/UNSD	Sectoral indicators—food & ag, labor, economics; SDG official repository.

9.3 How international series are created (high-level).

Table 2. Harmonization and modeling steps

Step	What happens
Harmonization	Convert national definitions to common standards; adjust breaks; apply age-heaping/under-reporting corrections where needed.
Modeling	Bayesian/ensemble models fill gaps; smooth volatility; ensure coherence across age/sex/time; incorporate covariates.
Consistency rules	Enforce demographic balance (births–deaths+net migration); align totals with census rounds; cross-walk survey/admin sources.
Uncertainty	Produce intervals/versions; document input sources and methodological changes.

9.4 Reading footnotes and vintages (practical tips).

Table 3. What to check before citing numbers

What to check	Why it matters
Reference year / vintage	Check the edition (e.g., WPP 2024) and the data 'as of' date; figures may change next release.
Definition alignment	See how indicators are defined vs. national series; look for changes in questionnaire or universe.
Gap-filling model	Identify whether values are observed or imputed; consult uncertainty notes.
Country comments	Review country notes for Ethiopia—special treatments, conflict-affected areas, or rebasings.
Comparability with national	If disseminating nationally, explain any differences and cite both sources.

9.5 Triangulating Ethiopia's indicators.

Table 4. Example triangulation matrix for Ethiopia

Indicator	Sources to triangulate
Population level (mid-year)	UN WPP (global consistency); ESS census/projections; World Bank WDI
Under-5 mortality rate (U5MR)	UN IGME (UNICEF/WHO/WB/UN DESA) vs DHS/MICS Ethiopian rounds vs IHME GBD
Maternal mortality ratio (MMR)	UN MMEIG vs national facility data/CRVS vs IHME GBD
Life expectancy at birth	UN WPP/WHO vs DHS indirect estimates; CRVS (when available)
International migrants	UN DESA migration vs immigration admin data; IOM/UNHCR (context-specific)

9.6 Why sources disagree—and what to do.

Table 5. Common reasons for differences across sources

Cause	Explanation
Different cut-off dates ('data as of')	One series includes a newer survey/census; another does not yet.
Definition & coverage	Different age definitions, universes (e.g., usual vs de facto), or geographic coverage.
Model assumptions	Priors and covariates differ (e.g., child mortality spline choices; excess mortality treatments).
Revision policy	Some series back-cast when new baselines arrive; others only move forward.
Rounding and presentation	Totals rounded differently; base year differences (calendar vs mid-year).

9.7 Use-cases for Ethiopian planning and policy.

Table 6. How to apply modeled series in practice

Use-case	How to apply
Medium-term planning	Use WPP for population denominators; scenario analysis for school/clinic placement.
Health burden & financing	Use GBD/WHO for disease burden and coverage; cross-check with DHS and HMIS.
Child & gender programs	Use UNICEF/IGME for U5MR, stunting, immunization; align with MICS/DHS microdata.
Migration & urbanization	Use UN DESA migration and WUP for urban shares; validate with census and admin.
International benchmarking	Use WDI/SDG database for cross-country comparisons; cite national exceptions.

Sidebar: National vs. international numbers—how to cite on your website

• For national planning documents, lead with Ethiopian Statistics Service (ESS) figures; add international numbers as context.
• Always include the vintage (e.g., "UN WPP 2024") and a note if figures differ from ESS due to modeling/harmonization.
• When small differences matter (e.g., denominators for coverage rates), choose one source consistently and explain why.

Selected references (for users of the site):

- UN DESA—World Population Prospects (WPP) methodological notes.

- World Bank—World Development Indicators documentation.

- UNICEF, WHO, World Bank—IGME child mortality methods.

- IHME—Global Burden of Disease methods summary.

- UN DESA—International Migration and World Urbanization Prospects metadata.

- UNSD—SDG Global Database methodology notes.

10) Data Quality Assessment & Triangulation

10.1 Why quality assessment matters. For Ethiopia's users—planners, researchers, and partners—credible numbers require transparent diagnostics and triangulation across sources. This section offers practical tools and examples to evaluate completeness, content error, and coherence, and to reconcile differences.

10.2 Age heaping example (Whipple index). Using a synthetic age distribution with heaping at digits 0/5, a Whipple-like index on ages 23–62 equals approximately 125.9. Values near 100 indicate no preference; higher values indicate increasing heaping.

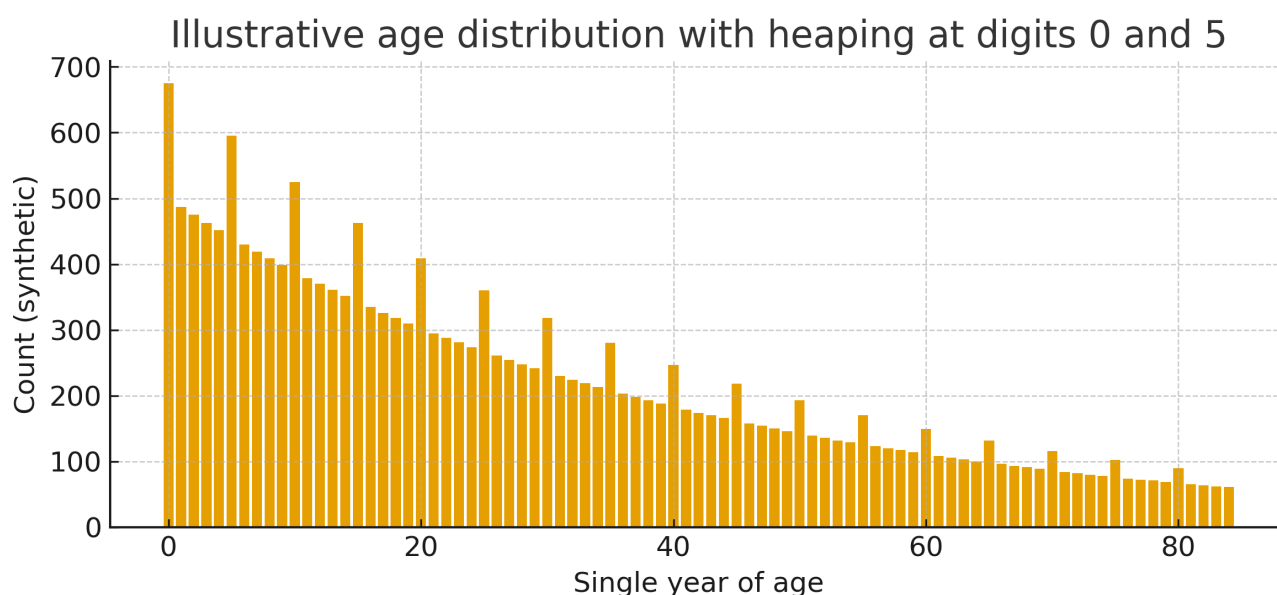*Figure 1. Illustrative age distribution with heaping at digits 0 and 5*

Table 1. Quality dimensions and operational indicators

Dimension	Operational indicators
Completeness/coverage	Reporting rate, % under/over-count, PES net coverage error
Content accuracy	Age heaping indices (Whipple/Myers), heaped dates, digit preference
Timeliness	Lag between reference and publication; lag in admin reporting
Consistency	Sex ratio by age; cohort consistency over time; trend breaks
Coherence	Cross-source alignment (census–survey–admin); demographic balancing
Accessibility & metadata	Availability of microdata, codebooks, methods, PES reports

10.3 Sex ratio and cohort checks. Sex ratios by age can reveal selective undercount, sex-differential migration, or data errors. Cohort traces across rounds should evolve plausibly after accounting for mortality and migration.

Figure 2. Illustrative sex ratio by age group (males per 100 females)

Illustrative sex ratio (males per 100 females) by age group
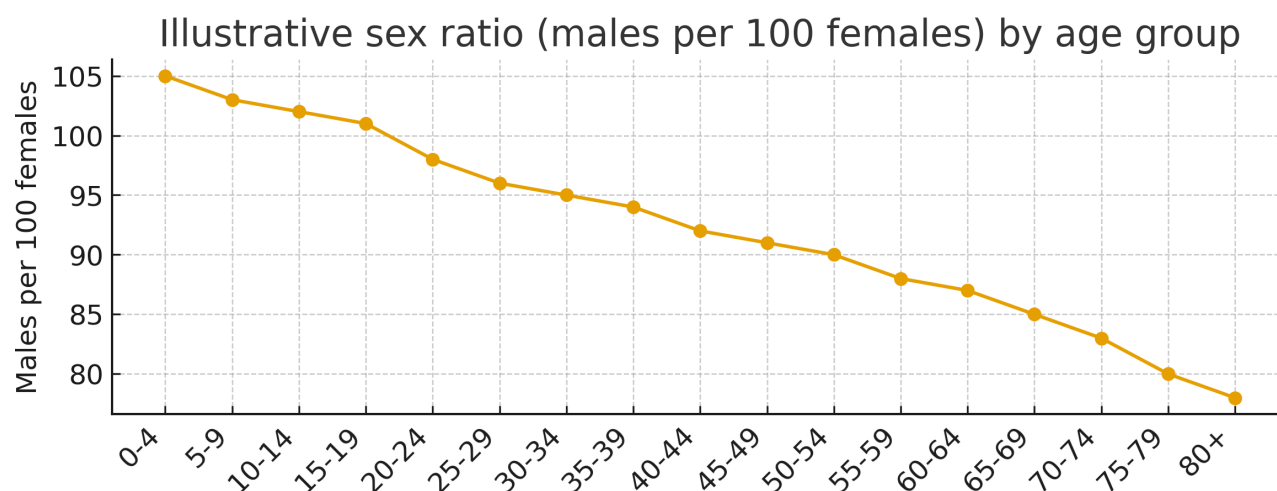
10.4 Triangulation across sources. Compare census, surveys, CRVS/admin, and modelled series—aligning reference dates and definitions—and then examine residual differences. A simple scatter along the 45° line provides a quick visual check for regional consistency.

Figure 3. Illustrative triangulation: census vs survey (regional totals)

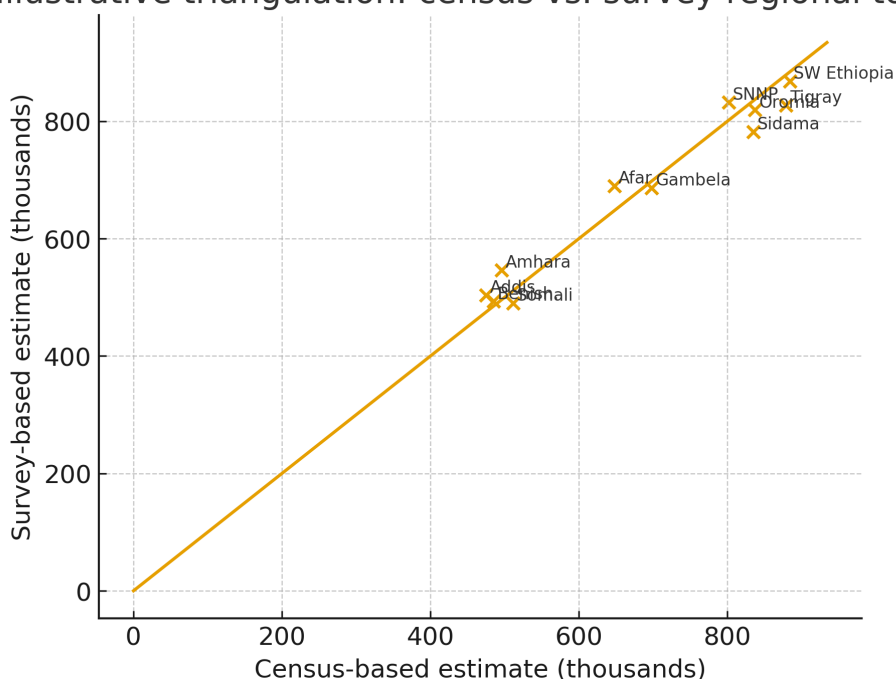Illustrative triangulation: census vs. survey regional totals

Table 2. Triangulation playbook—what to compare and what to check

Question	Sources to compare	What to check
Population totals (region)	Census (baseline), Survey (post-stratified), Admin (ID/CRVS where available)	Check reference date, universe (de facto/de jure), migration timing; reconcile to baseline with documented adjustments
U5MR / MMR	DHS/MICS (direct/indirect), HMIS/CRVS (cause of death), UN IGME/MMEIG, IHME	Beware model vs. observed; sample size/recall; facility-only bias; present uncertainty bands
Fertility (TFR)	DHS/MICS vs. WPP/GBD modeled series	Check age-displacement; birth histories quality; recent shock effects
Net migration	Census lifetime/recent migration, UN DESA migration, IOM/UNHCR & admin	Align definitions; account for displacement; cross-border vs internal flows

10.5 Post-Enumeration Survey (PES)—measuring coverage error.

Table 3. PES essentials—design, matching, estimation, transparency

Component	Key points
Design	Independent EA samples; matched to census EAs; adequate size for national + domain estimates
Matching	Strict rules to link persons/households; investigate mismatches (omissions/duplications)
Estimation	Dual-system estimation for net/gross coverage error; publish standard errors
Diagnostics	Coverage profiles by sex/age/region/urban-rural; reasons for misses
Transparency	Release PES report, microdata (as feasible), and reconciliation steps

Table 4. Common anomalies and diagnostics—what to look for

Anomaly	Diagnostic & response
Age heaping at digits 0/5	Compute Whipple/Myers; smooth or flag; emphasize correct age reporting in future rounds
Sex ratio anomalies	Check male deficits in young adult ages (migration or undercount); examine consistency across sources
Heaped dates (admin)	Look for spikes on 1st/15th of month; fix with validations; audit facilities
Regional breaks	Compare cohort traces across censuses/surveys; inspect cartography or boundary changes
Outliers in indicators	Use robust stats; verify with microdata re-tabulation; check skip logic effects

10.6 Reconciling differences—practical workflow.

Table 5. Reconciliation steps and documentation

Step	Action
1. Frame the discrepancy	Specify indicator, geography, period, tolerance
2. Align definitions	Residence concept, reference date, age universes, denominators
3. Inspect quality signals	PES, heaping, nonresponse, design effects, digit preference
4. Weight evidence	Give priority to higher-quality/closer-to-event sources
5. Document and publish	Explain adjustments; provide code/notebooks and version

Sidebar: Quick diagnostics to run on any new dataset

• Plot sex ratio by age; check cohort consistency vs prior rounds.

• Compute age-heaping indices; inspect date spikes; review missingness patterns.

• Cross-check key totals with another source (survey/admin/UN series) and note reference dates.

• Document findings in a short QA note attached to the dataset.

Selected references (for users of the site):

- UN Principles and Recommendations for Population and Housing Censuses—quality evaluation annexes.

- UNSD Handbook on the Use of Administrative Data for Official Statistics—quality frameworks.

- WHO/UNICEF and DHS/MICS methodological guides—indicator quality and comparability.

- Literature on Whipple/Myers indices and dual-system estimation for PES.

11) Indirect Demographic Methods

11.1 Why indirect methods? When civil registration is incomplete or censuses/surveys have gaps, indirect demographic methods provide principled estimates of fertility, mortality, and population structures. They rely on parsimonious inputs and modeling assumptions, and must be paired with transparent diagnostics and triangulation.

11.2 Toolbox overview—what each method needs and produces.

Table 1. Indirect methods—outputs, inputs, and assumptions

Method	Outputs	Inputs	Key assumptions
Brass P/F ratio	Total fertility level (TFR) adjustment	Parity by age (P), ASFRs (F) from survey/census	Fertility pattern stable; reporting consistent; no major tempo shifts
Own-Children Method (OCM)	ASFRs and recent TFR from household rosters	Children linked to mothers; age distributions	Co-residence sufficient; mortality/fostering addressed
Child survivorship (Brass/Trussell)	Under-5 mortality (q(5))	Children ever born and surviving by mother's age	Stationarity or model relationships; reporting quality acceptable
Orphanhood	Adult mortality (e.g., 45q15)	Proportions maternal/paternal orphans by age	Stable mortality; correct parental survival reporting
Sisterhood (direct/indirect)	Maternal mortality ratio (MMR)	Reports of sisters' survival and pregnancy-related deaths	Adequate sample; recall/truncation adjusted
General Growth Balance (GGB)	Coverage of adult deaths & census	Two censuses + adult deaths by age	Closed population or corrected; roughly stable coverage
Synthetic Extinct Generations (SEG)	Adult mortality & completeness	Age distribution of deaths and population	Stationarity at older ages; minimal migration at older ages
Relational logit (Brass; Gompertz)	Full life table from q(x) points	Observed survivorship points; standard life table	Shape approximates chosen standard
Model life tables (UN, Coale–Demeny)	Mortality schedule & life expectancy	Few indicators (e0, q5, 45q15)	Population resembles selected family

*Figure 1. Brass P/F ratios by age group (illustrative)*

Illustrative Brass P/F ratios by age group
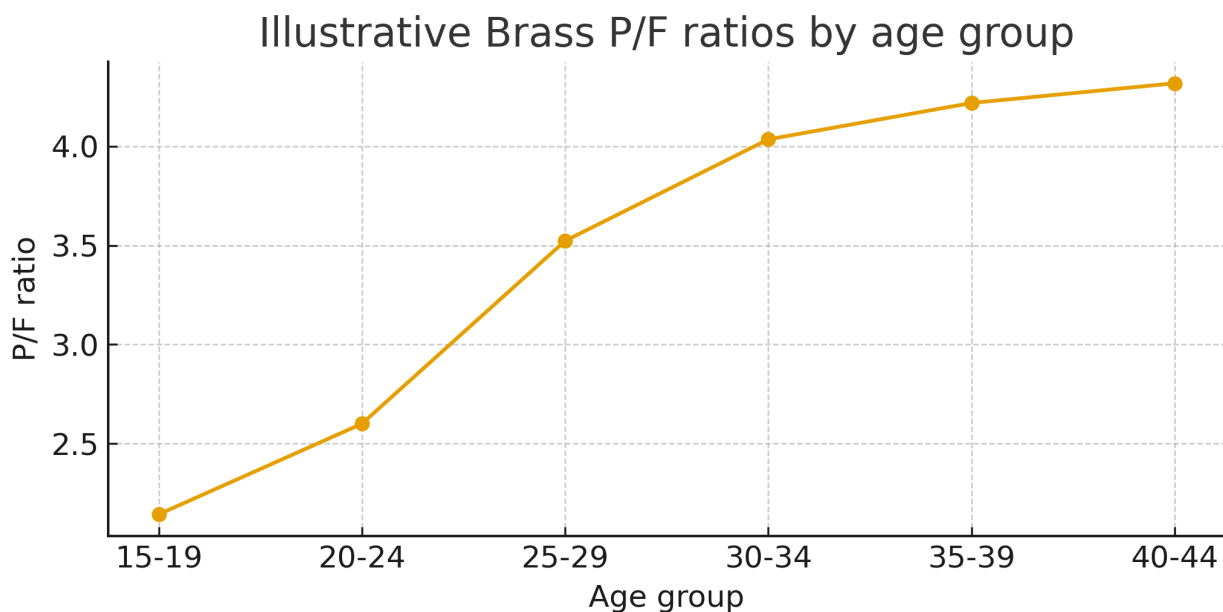
11.3 Ethiopia—when to use which method.

Table 2. Ethiopia context → recommended approach

Situation in Ethiopia context	Recommended approach
Weak birth/death registration; recent DHS available	Use Brass P/F or OCM to benchmark TFR; cross-check with DHS direct estimates
No recent census; strong household roster	OCM with adjustments for child mortality and fostering; validate with school enrolment
Conflict-affected or mobile populations	Orphanhood and child survivorship can bias; prefer survey modules + triangulation
Sparse adult death data	Combine GGB with SEG (GGB–SEG) for completeness; test sensitivity to migration
No adult mortality data; limited survivorship	Relational life table to infer full schedule; compare with UN model families

*Figure 2. Model life-table style survivorship l(x) (illustrative)*

Illustrative survivorship l(x) — model life-table styles
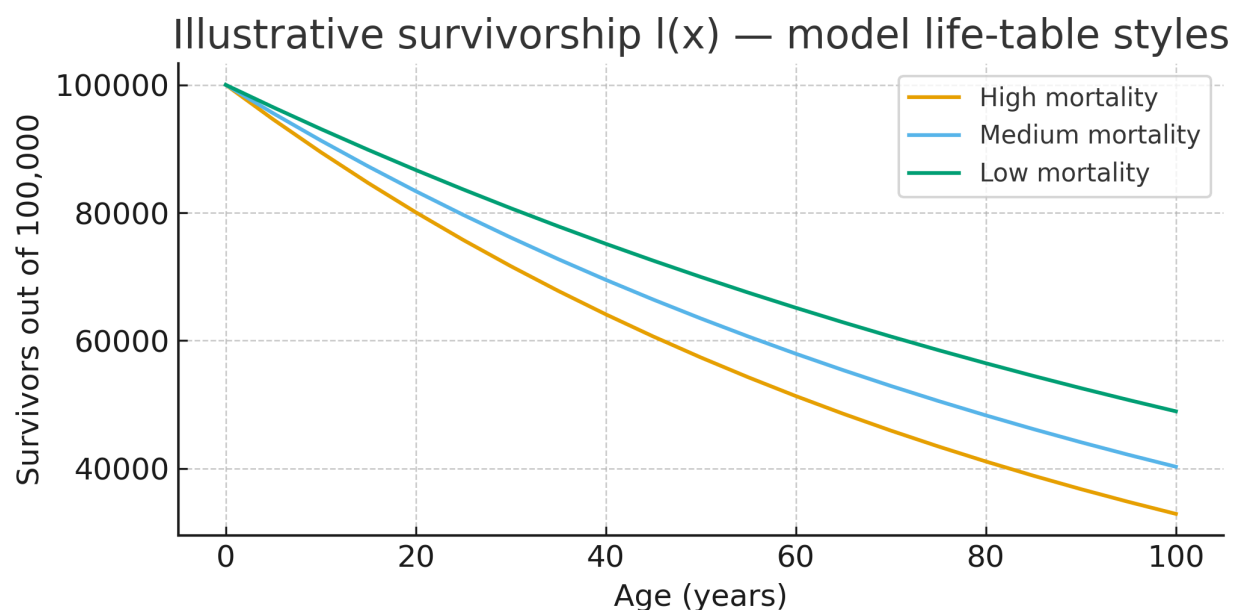
11.4 Quick workflow for fertility and mortality.

Table 3. Step-by-step workflows

Target	Steps
Fertility (TFR)	Compute ASFRs → P/F adjustment → compare with DHS/UN WPP → document uncertainty
Under-5 mortality	Brass child survivorship (q(5)) → compare with DHS direct + IGME → reconcile
Adult mortality	Use orphanhood or GGB–SEG (if data exist) → compare with model life tables/GBD
Life table	Fit relational logit to observed points → generate l(x), m(x), e(x) → sanity checks

Figure 3. GGB–SEG diagnostic (illustrative)
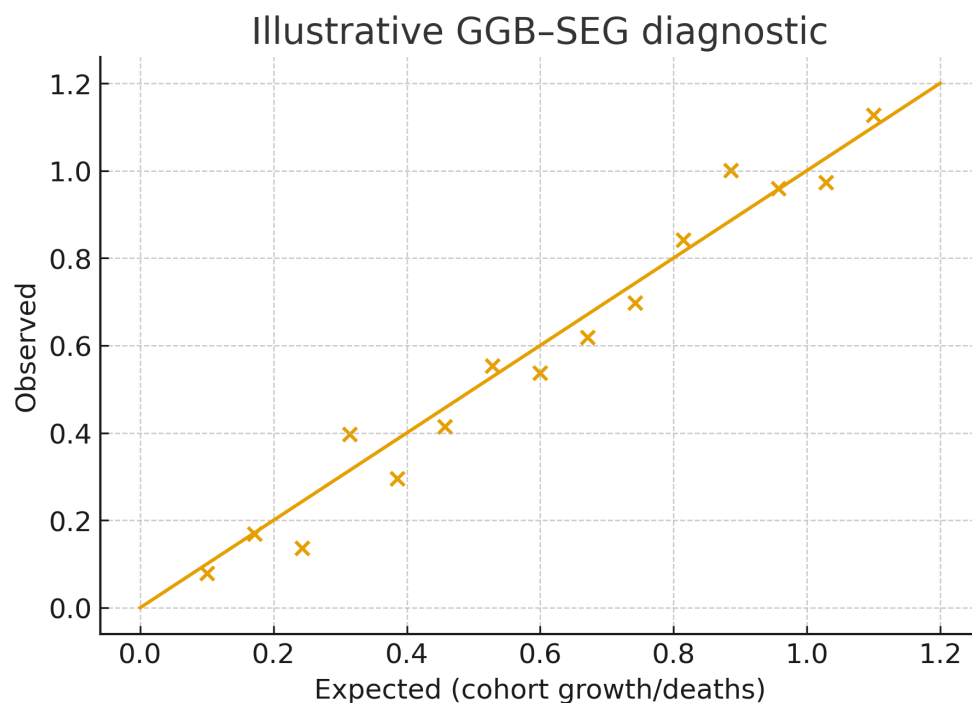
11.5 Common pitfalls and mitigations.

Table 4. Pitfalls to watch for and how to mitigate them

Pitfall	Mitigation
Age misreporting / heaping	Use smoothing; group ages; apply P/F on central reproductive ages
Fostering / non-co-residence	OCM adjustments; sensitivity analysis; check orphanhood biases
Migration (adult)	Limit GGB–SEG to ages less affected; adjust/model migration where possible
Recall bias (sisterhood)	Use shorter reference; larger samples; interviewer training
Model mismatch	Try multiple life-table families; report ranges

Sidebar: Communicating uncertainty with indirect methods

• Report plausible ranges using sensitivity to model choices (e.g., different life-table families or Brass variants).

• Show how estimates change when excluding suspect age groups (e.g., heaped ages) or adjusting for migration.

• Clearly label all outputs as indirect estimates and cite inputs and software used.

Selected references (for users of the site):

- Brass, W. Methods for Estimating Fertility and Mortality from Limited Data.

- UN Manual X & Manual II—Indirect Techniques for Demographic Estimation; Model Life Tables.

- Coale & Demeny (Regional Model Life Tables).

- Preston, Heuveline & Guillot—Demography: Measuring and Modeling Population Processes.

- Hill, Timæus & Zlotnik—Adult mortality estimation (GGB/SEG).

12) Data Governance, Privacy & Ethical Use

12.1 Why governance matters. Trusted population statistics depend on lawful collection, strong stewardship, privacy-by-design, and accountable use. This section offers an Ethiopia-first framework aligned with international good practice.

12.2 Governance principles and how to operationalize them.

Table 1. Core principles → how to implement

Principle	Operationalization
Lawfulness & purpose limitation	Collect and process data under clear legal mandates aligned to public interest; limit to stated purposes.
Data minimization & proportionality	Only collect fields essential for statistics/service delivery; avoid sensitive data unless necessary.
Quality & accountability	Document sources, methods, QA results; appoint accountable owners and stewards.
Security & confidentiality	Apply access controls, encryption, audit logs; protect microdata and PII rigorously.
Fairness & inclusion	Assess representativeness and equity impacts; avoid discriminatory uses.
Transparency	Publish metadata, data dictionaries, revision notes, and user-friendly explanations.

12.3 Roles and accountability (illustrative mapping).

Table 2. Ethiopia institutions and primary governance roles

Institution	Primary governance role
Ethiopian Statistics Service (ESS)	Lead on statistical standards, quality frameworks, microdata curation, and dissemination.
Sector custodians (MoH/EMIS/VERA etc.)	Own administrative systems; ensure lawful collection, data quality, and secure sharing.
National ID Program (NIDP)	Provide unique identifiers and deduplication services; manage ID–data link governance.
Data protection/legal authority	Issue guidance, review DPIAs, and enforce privacy/security obligations.
Digital agency/ICT	Run interoperability platforms (APIs), identity services, security operations.
Independent oversight (audit/ethics)	Review high-risk projects; handle complaints; publish annual reports.

12.4 Data lifecycle: artifacts you should always have.

Table 3. Lifecycle stages and minimum documentation

Stage	Required documents/artifacts
Plan	Business case, legal basis memo, DPIA screening, stakeholder map
Design	Data dictionary, code lists, consent/notices, security plan, QA plan
Collect	SOPs, training, consent scripts, incident response playbook
Process	ETL specs, pseudonymization keys, QA checks, version control
Analyze	Method notes, bias assessment, reproducible code notebooks
Disseminate	Release checklist, disclosure control, licensing, metadata
Archive/Retire	Retention schedule, deletion certificates, archival metadata

12.5 Privacy techniques and access controls.

Table 4. Practical privacy & disclosure‑control toolbox

Area	Techniques / examples
Access control	Role-based access, least privilege, multi-factor authentication, audit logs
De-identification	Pseudonymization, tokenization, hashing with salts, separation of keys
Statistical disclosure control	Top/bottom coding, cell suppression, noise infusion, differential privacy (where applicable)
Geoprivacy	Blur coordinates, aggregate to hex/buffer zones, mask rare categories
Data sharing tiers	Public (aggregates), Restricted (approved users/microdata), Confidential (on-site/virtual enclave)
Incident management	Breach detection, reporting timelines, user notification templates

12.6 Data sharing agreements (DSAs) and DPIAs.

Table 5. Essential clauses for DSAs

Clause	What to include
Purpose & legal basis	Clearly defined processing purpose(s) and statute/regulation references
Data specification	Variables, sensitivity classification, time coverage, updates
Security & confidentiality	Controls, location of processing, encryption at rest/in transit
Access & retention	Who can access, for how long; retention schedule and deletion protocol
Permitted/Prohibited uses	Research vs. operational use; bans on re-identification or enrichment without approval
Accountability	Contacts, audit rights, penalties for breach, publication & attribution rules

Table 6. DPIA checklist—what to answer before processing

Section	Key questions
Scope & context	Data subjects, categories, volumes, flows, cross-border transfers
Necessity & proportionality	Is each element necessary? Consider less intrusive alternatives
Risk analysis	Identify harms (discrimination, security, reputational); likelihood×impact
Mitigations	Controls, consent/notices, safeguards, governance measures
Residual risk & sign-off	Decision to proceed; conditions; monitoring plan

Sidebar: Three-tier data access model for Ethiopia

• Public: Aggregated indicators and anonymized micro-samples with strong disclosure control.

• Restricted: Researcher access to detailed microdata under DSA, ethics approval, and secure enclave conditions.

• Confidential: Identifiable data only for authorized officials on secured platforms with strict logging and purpose limitation.

Selected references (for users of the site):

- UN Fundamental Principles of Official Statistics; UN Handbook on Privacy-Preserving Microdata Dissemination.
- OECD & World Bank data governance guidance; SDMX/DDI standards for metadata and exchange.
- WHO/UNICEF guidance on health data governance and disclosure control.
- International best practices on DPIAs, DSAs, and secure research environments.

Appendix — Glossary & References

A. Glossary of Terms

Table A1. Glossary

Term	Definition
Administrative data	Records generated as part of delivering public services (e.g., HMIS, EMIS, tax).
Age heaping	Over-reporting of ages ending in certain digits (often 0 or 5).
Areal weighting	Distributing counts uniformly within polygons to derive gridded estimates.
Big data / digital traces	Large, passively generated datasets (e.g., mobile CDRs, app mobility, social media).
CRVS	Civil Registration and Vital Statistics—continuous registration of vital events.
De facto / De jure	Where a person slept on census night vs. their usual residence.
Differential privacy	Adding noise or other techniques to limit disclosure while preserving statistics.
DHS / EMDHS	Demographic and Health Survey / Ethiopia Mini DHS—household surveys on health and population.
GHS-POP / WorldPop	Global gridded population products derived from census + covariates.
GGB / SEG	General Growth Balance / Synthetic Extinct Generations—methods to assess death/census completeness.
Household frame	A list or master sample of areas/households used to select survey samples.
IDP / Refugee	Internally displaced person / person displaced across borders under international protection.
Indicator	A measurable statistic derived from raw data (e.g., U5MR, TFR).
Interoperability	Systems' ability to exchange/use information through standards (APIs, code lists).
LSMS-ISA / ESS	Living Standards Measurement Study—Ethiopia Socioeconomic Survey; panel survey platform.
Model life table	Standard mortality schedule used to infer full life table from limited inputs.
P/F ratio	Brass parity/fertility ratio used to adjust fertility levels.

PES	Post-Enumeration Survey—independent survey to measure census coverage error.
Register (population)	Continuously updated person-level system linked via unique ID; used for administration & stats.
Sankey / Flow matrix	Visuals to depict origin–destination flows of migrants or resources.
SRS	Sample Registration System—continuous measurement of births/deaths in a sampled population.
Triangulation	Comparing multiple sources/methods to increase confidence in estimates.
Usual residence	Place where a person normally lives for the majority of time.
Whipple's index	Measure of age heaping focusing on digits 0 and 5 among ages 23–62.

B. Acronyms

Table A2. Acronyms used in Chapter 1

Acronym	Expanded form
CRVS	Civil Registration and Vital Statistics
DHS / EMDHS	Demographic and Health Survey / Ethiopia Mini DHS
DTM	Displacement Tracking Matrix (IOM)
EMIS	Education Management Information System
ESS	Ethiopian Statistics Service (also Ethiopia Socioeconomic Survey context)
GBD	Global Burden of Disease (IHME)
GHO / GHE	Global Health Observatory / Global Health Estimates (WHO)
GHSL / GHS-POP / GHS-SMOD	Global Human Settlement Layer / Population / Settlement Model
HMIS	Health Management Information System
HRSL	High Resolution Settlement Layer
IGME	Inter-agency Group for Child Mortality Estimation (UNICEF/WHO/WB/UN DESA)
LSMS-ISA	Living Standards Measurement Study—Integrated Surveys on Agriculture
MMEIG	Maternal Mortality Estimation Inter-agency Group
NIDP	National ID Program
PES	Post-Enumeration Survey
PSNP	Productive Safety Net Program

SDMX / DDI	Statistical Data and Metadata eXchange / Data Documentation Initiative
SEG / GGB	Synthetic Extinct Generations / General Growth Balance
UN WPP	United Nations World Population Prospects
UNHCR	UN Refugee Agency
UNSD SDG	UN Statistics Division—SDG Global Database
WHO	World Health Organization
WDI	World Development Indicators

C. References & Further Reading

C1. Methods & Demographic Estimation

- Brass, W. (various). Methods for Estimating Fertility and Mortality from Limited Data.

- United Nations, Manual X / Manual II. Indirect Techniques for Demographic Estimation; Model Life Tables.

- Preston, Heuveline & Guillot. Demography: Measuring and Modeling Population Processes.

C2. International Series & Compendia

- UN DESA. World Population Prospects (latest vintage) — methodology notes.

- World Bank. World Development Indicators — documentation.

- UNICEF/WHO/WB/UN DESA (IGME). Child mortality methods.

- IHME. Global Burden of Disease — methods summary.

- UN DESA. International Migration and World Urbanization Prospects — metadata.

C3. Systems, Governance & Interoperability

- UNSD. Handbook on the Use of Administrative Data for Official Statistics.

- WHO/UNICEF. Health data governance and disclosure control guidance.

- UNECA/World Bank. CRVS–ID linkages and interoperability resources.