

Types of demographic data: census (records of all individual data at one time), survey (specially collecting data of [usually a part of] population) and registration systems (records of events within particular period).

The purposes of data collection

- Most demographic data are by-product of other purposes like administrative exercises, including vital registration systems, censuses, migration controls, public health campaigns, population control programs and special surveys.
- Vital registration systems: primarily for legal purposes like the production of birth and death certificates, but also for demographic analysis like mortality and fertility data.
- Censuses: mainly for aid-planning, provision of basic service needs, but also used for population distribution and structure.
- Surveys: designed with more strongly research-oriented objectives.

Censuses

- Enumerating all individuals in the nation.
- "***de facto***" (現在人口) and "***de jure***" (常住人口) **population**: the former is simply the population actually living at the moment of census. the latter is somewhat complicated.
- Japan conducts national census on October 1st every 5 years, when the enumerator (temporarily employed by the government) counts "***de facto***" population at past, but "***de jure***" now (At each place, count the individuals who are to live there more than 3 months).
 - Especially in developing countries, some enumerators fill the sheet without any confirmation, so that age misreporting is frequently seen.
- UN recommends that censuses should be held at least at ten-yearly intervals, preferably in years ending in 0.
- Each country take somewhat diversified method. Several kinds of "population" are enumerated: Whether military, diplomat, visitors are enumerated or not? Distinguishing migrants from visitors/travelers depends on how long they stay.
- **Chief characteristics**: Complete recording of entire population within a geographically well-defined area. No sampling. Each person must be enumerated separately. It must have legal basis. It must relate to a single point in time, not a period.
- Source of errors: poor response quality, missing answers, refusal, misclassification.
- **To collect reliable census data, properly trained enumerator is essential.**
- (Identification of individual is essential for longitudinal survey.)
- The census statistics data in Japan are available from e-Stat (See, <http://www.e-stat.go.jp/SG1/estat/eStatTopPortalE.do>).
- IPUMS (<https://international.ipums.org/international/>) provides census data from around the world. The ipumsr package (<https://cran.r-project.org/web/packages/ipumsr/>) makes easy to use IPUMS data.

Vital registration

- Source of vital statistics (number of birth, death, ...).
- In developing countries, the system is not working well. Delay, missing and duplication are usual.
- **In Japan, major vital events including birth, marriage, divorce, and death are reported and registered to the municipal office**, which are summarized by health center, then **reported to the Ministry of Health, Labor and Welfare**. The information is mostly correct, but the accuracy of cause of death may be not enough because the doctors usually do not conduct anatomical investigation after the death of the patient (and the number of pathologists and forensic doctors are relatively smaller than that of other countries): The proportion of the determination of cause of death based on anatomical investigation is less than 5% in most prefectures.
- The vital statistics data in Japan are also available from e-Stat.

Sample surveys

- Taking a sample of the population, generating complex data by conducting object-oriented questions.
- Sampling error is main possible disadvantage.
- Prospective survey costs much.
- Retrospective survey may be more biased. To avoid bias, demographers use some sophisticated techniques.