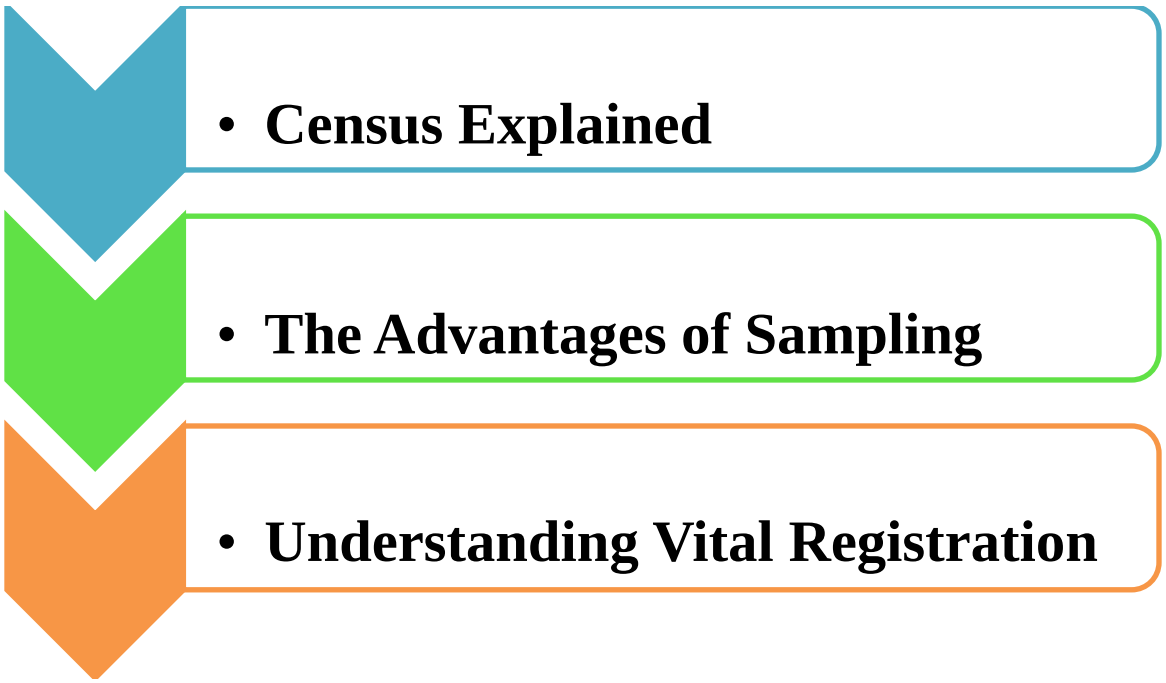
LESSON 2

Population Data Sources

Aynalem Adugna , July 2014

EthioDemographyAndHealth.Org

Learning Objectives:



Introduction

There are three sources of population data:

- Censuses
- Vital Registration Systems
- Sample Surveys

I - POPULATION CENSUSES

Censuses taking started nearly 6000 years ago [1]. Babylonians are said to be the pioneers in the field, followed later on by Persians and other civilizations, including the Greeks, Romans and Chinese dynasties. The practice also underlies a fundamental belief in Christianity regarding Jesus Christ's place of birth. "It was the five-yearly census ordered by Caesar Augustus which required every man in the Roman Empire to return to his place of origin, thus ensuring that

Joseph and Mary travelled to Bethlehem for the birth of Jesus" [2]. The 1841 census of England and Wales is widely regarded as the first truly modern census.

The following paragraphs are based on United Nations recommendations on census taking [3]:

Census: Definition

Population

" A population census is the total process of collecting, compiling, evaluating, analysing and publishing or otherwise disseminating demographic, economic and social data pertaining, at a specified time, to all persons in a country or in a well delimited part of a country" [3]

Housing

"A housing census is the total process of collecting, compiling, evaluating, analysing and publishing or otherwise disseminating statistical data pertaining, at a specified time, to all living quarters1 and occupants thereof in a country or in a well-delimited part of a country" [3].

A census must have the following essential features:

Individual enumeration – each individual and living quarter has to be enumerated separately
Universality – A census must cover every individual or housing unit present within the defined census area.
Simultaneity – Each person and housing unit must be canvassed within a defined point in time.
Defined periodicity - There should be a defined time gap between censuses. The most commonly used interval is 10 years.

Why take population censuses?

- To provide facts to policy makers and planners
- To allow policy refinements in management/evaluation of programs
- For gerrymandering or redistricting – delimitation of election boundaries to insure adequate representation
- For scientific research
- Industry – to determine consumer demand and availability of labor

Other Censuses

- Agriculture
- Livestock
- Industry and commerce

Two Types of enumeration:

- Canvasser (enumerator) method – information on each individual or housing unit is entered by a census official.
- Household method: responsibility for entering information given to an individual in the housing unit.

Census Questionnaire: List of Topics

1. **Geographical and internal migration characteristics**
(a) Place of usual residence
(b) Place where present at time of census, **Locality**
(c) Place of birth
(d) Duration of residence
(e) Place of previous residence
(f) Place of residence at a specified date in the past
2. Household and family characteristics
(a) Relationship to head or other reference member of
(b) Household and family composition household
(c) Household and family status
3. Demographic and social characteristics
(a) Sex
(b) Age
(c) Marital status
(d) Citizenship
(e) Religion
(f) Language
(g) National and/or ethnic group
4. Fertility and mortality
(a) Children ever born
(b) Children living
(c) Date of birth of last child born alive
(d) Deaths in the past 12 months
(e) Maternal or paternal orphanhood

Principles and recommendations for population and housing censuses
Topics collected directly Derived topics
(f) Age, date or duration of first marriage
(g) Age of mother at birth of first child born alive
5. Educational characteristics
(a) Literacy
(b) School attendance
(c) Educational attainment
(d) Field of education and educational qualifications
6. Economic characteristics
(a) Activity status
(b) Time worked
(c) Occupation
(d) Industry
(e) Status in employment
(f) Income
(g) Institutional sector of employment
(h) Place of work
7. International migration characteristics
(a) Country of birth
(b) Citizenship
(c) Year or period of arrival
8. Disability characteristics
(a) Disability
(b) Impairment and handicap
(c) Causes of disability

Source: [3]

II. VITAL REGISTRATION SYSTEMS

"Although local or parish registers were kept by some churches in Europe from the 14th century onwards, civil or state registration systems did not develop until the 19th and 20th centuries…….Unlike censuses that describe the state of the population at a fixed point in time, vital statistics are collected on a continuous basis" [4]

Life events registered under a complete registration system include [5]:

- Live Births
- Deaths
- Foetal deaths
- Marriages
- Annulments/ Legal separations
- Adoptions

Important principles of a Vital Registration system:

All of the paragraphs below come from a single source [3] which we have found to be a good guide for understanding civil registration systems.

Universal coverage

A vital statistics system should include all vital events occurring in every geographic area and in every group comprising the national population.

Continuity

"The principle of continuity in the collection and compilation of vital statistics should be observed in order that the data may reflect short-term fluctuations, including seasonal movements, as well as longer-term movements."

Confidentiality

"Confidentiality of personal information in registration records and any associated statistical reports should be safeguarded insofar as consistent with the intended uses of these records for specific administrative and statistical purposes."

Regular dissemination

"The compilation of vital statistics should have as its ultimate minimum goal (*a*) the provision of total monthly or quarterly summary counts of live births, deaths, foetal deaths, marriages and

divorces on a time schedule prompt enough to provide information for health intervention and population estimation programs, administrative uses or other needs, and (*b*) the production of detailed annual tabulations of each type of vital event cross classified by its demographic and socioeconomic characteristics."

Definition of Vital Events

"LIVE BIRTH is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which after such separation, breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered liveborn (all live-born infants should be registered and counted as such, irrespective of gestational age or whether alive or dead at the time of registration, and if they die at any time following birth they should also be registered and counted as deaths)."

"DEATH is the permanent disappearance of all evidence of life at any time after live birth has taken place (postnatal cessation of vital functions without capability of resuscitation) (This definition excludes foetal deaths, which are defined separately below)."

"FOETAL DEATH [DEADBORN FOETUS] is death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy; the death is indicated by the fact that after such separation the foetus does not breathe or show any other evidence of life, such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles25 (note that this definition broadly includes all terminations of pregnancy other than live births, as defined above)"

"MARRIAGE is the act, ceremony or process by which the legal relationship of husband and wife is constituted. The legality of the union may be established by civil, religious or other means as recognized by the laws of each country."

"DIVORCE is a final legal dissolution of a marriage, that is, that separation of husband and wife which confers on the parties the right to remarriage under civil, religious and/or other provisions, according to the laws of each country."

"ANNULMENT is the invalidation or voiding of a marriage by a competent authority, according to the laws of each country, which confers on the parties the status of never having been married to each other."

"SEPARATION, JUDICIAL is the disunion of married persons, according to the laws of each country, without conferring on the parties the right to remarry."

"ADOPTION is the legal and voluntary taking and treating of the child of other parents as one's own, in so far as provided by the laws of each country."

"LEGITIMATION is the formal investing of a person with the status and rights of a person born in wedlock, according to the laws of each country."

"RECOGNITION is the legal acknowledgment, either voluntarily or compulsorily, of the paternity of a child born out of wedlock."

Summary Definition of Civil Registration Systems

"Civil registration is defined as the continuous, permanent, compulsory and universal recording of the occurrence and characteristics of vital events pertaining to the population as provided through decree or regulation in accordance with the legal requirements of a country. Civil registration is carried out primarily for the purpose of establishing the legal documents provided by the law. These records are also a main source of vital statistics. Complete coverage, accuracy and timeliness of civil registration are essential for quality vital statistics." [6]

Locality

According to this UN report a locality is defined as "a distinct population cluster (also designated as inhabited place, population centre, settlement etc.), in which the inhabitants live in neighboring sets of living quarters and which has a name or a locally recognized status" These entities should not be confused with the smallest administrative l divisions of a country (a Kebele, in the case of Ethiopia). There might be an overlap between the two in some instances. But, "in others, however, even the smallest civil division may contain two or more localities". [5]

The recommended classification of localities by size-class is as follows:

- All localities
- 500,000 or more inhabitants
- 100,000 - 499,000 inhabitants
- 50,000 - 99,999 inhabitants
- 20,000 - 49,999 inhabitants
- 10,000 - 19,999 inhabitants
- 5,000 - 9,999 inhabitants
- 2,000 - 4,999 inhabitants
- 1,000 - 1,999 inhabitants
- 500 - 999 inhabitants
- 200 - 499 inhabitants
- Less than 200 inhabitants
- Population not in localities

Vital registration in Ethiopia: a Brief History

Tentative and largely symbolic efforts to establish a registration system in Ethiopia include the following: [6]

- The 1900 proclamation by Emperor Menelik to institute a registration system. This was well-intentioned but failed to materialize.
- The 1960 Civil Code: This too was not implemented for lack of institutional support and enforcement.
- The city of Addis Ababa launched birth registration in 1942. It then started registering marriages and deaths in 1953 and 1970 respectively. However, this took place at the whim of individual residents with legal and other needs for a certificate.
- A 1980 proclamation stipulated that the Central Statistical Authority (CSA) will begin to undertake registration of vital events. All of the preliminary efforts and attempts to lay the groundwork for a national registration system proved costly and the plan was shelved in 1999.
- The 1983/88 proclamations by the Derg to register births, deaths, marriages and population numbers did not come to fruition.
- The 1995 FDRE constitution considered the naming of a child and record of his/her birth a fundamental right, "but the law on civil registration has not come into effect." [6].

Ongoing activities:

- Continuing efforts by CSA to develop the ground rules for a national program is being assisted by the UN and other organizations and donor agencies.
- CSAs continued efforts at model building, and testing.
  - Continued canvassing by tens of thousands of health care workers, who are supplying vital health information to a national database.
  - The 2006 National Conference on registration of vital events entrusted the Ethiopian Human Rights Commission with the task of setting up a national task force.

III SAMPLE SURVEYS

The following are obtained from a UN "Studies in Methods" series on population sampling [26]. "Household surveys provide a cheaper alternative to censuses for timely data and a more relevant and convenient alternative to administrative record systems".

Sample surveys are used for the collection of detailed data on:

- Socio-demographic characteristics
- Conditions under which people live
- Their well-being
- Activities in which they engage
- Demographic characteristics and cultural factors which influence behavior,
- Social and economic change

Sample surveys provide the structure within which other variables such as education, health, labor force, disability, nutritional status, migration, fertility, mortality and seemingly unquantifiable issues such as criminal victimization and culturally sensitive topics including age at first sexual activity, frequency of intercourse, number of sexual partners, sexual orientation, etc. are studied. Survey data often complements those obtained from vital registration records, and from national censuses.

Planning and execution of surveys requires all of the following steps below and many more that have to precede the training of interviewers, such as the selection and specification of the subject matter, development of survey design, design and printing of questionnaires, pre-testing, and preparation of instructional and training materials for field use:

- Training interviewers
- Data collection
- Field administration
- Data processing
    - Systems planning
    - Computer programming
    - Clerical coding
    - Key-to-disk operations
- Data review and publication

The Sample design should insure the following [3]:

- The sample must be conducted in stages to identify accurately the locations where interviews are to be conducted and to select the study households efficiently.
- It has to be stratified to insure coverage of targeted geographic sub-areas and population sub-groups
- It has to make use of clusters of study households in order to keep costs down to a manageable level.
- The size of the study sample should be just right in order to optimally balance the competing needs of cost-cutting on the one hand and accuracy of results on the other.

History of Sample Surveys in Ethiopia

Census taking is a costly and time-consuming undertaking. The continuous and complete registration of viral events – births, deaths, marriages, etc. is even costlier and beyond the financial and technical reach of a developing country like Ethiopia. The solution has been use of sampling. What is sampling? It is a statistical technique defined as "the analysis of a group by determining the characteristics of a significant percentage of its members chosen at random." [7]

The organization charged with the collection, analysis, and dissemination of sample and census data in Ethiopia is the Central Statistical Authority (CSA) which was established in 1960 (formerly known as the Central Statistical Office - CSO). In the agency's own words, its functions include "…. running a National Integrated Household and Enterprise Survey Program (NIHESP), undertaking ad-hoc surveys, conducting census, and compilation of secondary data from administrative records" [8].

The Agency has conducted several socio-economic and demographic surveys as well as agricultural surveys, surveys featuring the price of goods and services, household income, consumption and expenditure, welfare monitoring, large and medium scale manufacturing industries, production of electricity, small scale manufacturing industries, cottage industries, construction, mining and quarrying, transport and communications, the informal sector, distributive trade and services, manpower, demography, family and fertility, health and nutrition, and child labor. Current surveys conducted by the CSA cover 2,072 rural Enumeration areas (EA) and 790 urban EAs [8].

The first national demographic surveys were conducted between 1964 and 67 (first round) and 1968 – 1969 (second round). Others ample surveys hitherto undertaken by the agency include the:

- Addis Ababa manpower and housing survey 1976
- Population and housing characteristics of 17 major towns 1978
- The rural labour force survey 1981/2
- The rural labour force survey 1987/8
- National Rural Nutrition Survey of 1992
- 1998 Health and Nutrition Survey
- National Labour force survey 1999

- Disability Survey of Selected Weredas
- Demographic and Health Survey 2000
- Child labor survey 2001
- Biannual employment 2003
- Biannual employment 2004
- Demographic and Health Survey 2005

Data Quality

Just how reliable are the data put out by the CSA? Given the socioeconomic level of the country, the educational background of its population, and the likely impacts these would have on data collection, a quick answer would be, not very reliable. But, once again, this would have to be viewed within the context of the social, educational, and economic environments within which the agency operates. None are favorable for the production of high quality data. How is quality measured anyway? The answer lies in the following crucial ingredients which form the corner-stones of a high-quality survey data. [9].

Relevance: This refers to the degree to which the data produced meets all of the real objectives of the survey.

Accuracy: This measures the extent to which the data "…..correctly describes the phenomena it was designed to measure" [9]. The usual spoilers include errors in statistical estimation, bias (or systematic errors) and variance (often described as random error), and other errors including interviewer error, respondent error, and non-response.

Timeliness: This pertains to the time-gap between the reference point (or the end of the survey) to which the information relates, and the date at which the data is made available.

Accessibility This refers to the ease or difficulty with which data can be accessed and the form or medium of accessibility. To some users, the cost of acquiring the data becomes an important aspect of accessibility.

Interpretability Listed underneath the interpretability heading are issues such as the underlying concepts, types and number of variables, classifications used, and methodologies employed to collect the data.

Coherence: " The *coherence* of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time" [9].

A cursory evaluation of the quality of some of the survey data for Ethiopia can be performed using a basic consistency test (see examples below).

Example 1.
Sex Ratio: Rural of the Population by Region

Region	Sex Ratio = Males per100 Females
Tigray	**97.2**
Afar	**125.0**
Amara	**99.9**
Oromia	**99.6**
Somali	**116.0**
Benishangul /Gumuz	**101.0**
SNNPR	**98.9**
Gambela	**104.0**

Source: Based on [8]

The rural population is selected for review of sex ratios to avoid confusion with urban populations' uneven balance of sexes arising from regional differences in rural-urban migratory streams and counter streams. The first review (above) of the data put out by CSA earns the agency a failing grade as there are no realistic explanations on the ground for the regional differences in the ratios shown. This has to be an artifact of data errors, and by no means a true reflection of gender balance within regions, or of differences between regions.

Example 2

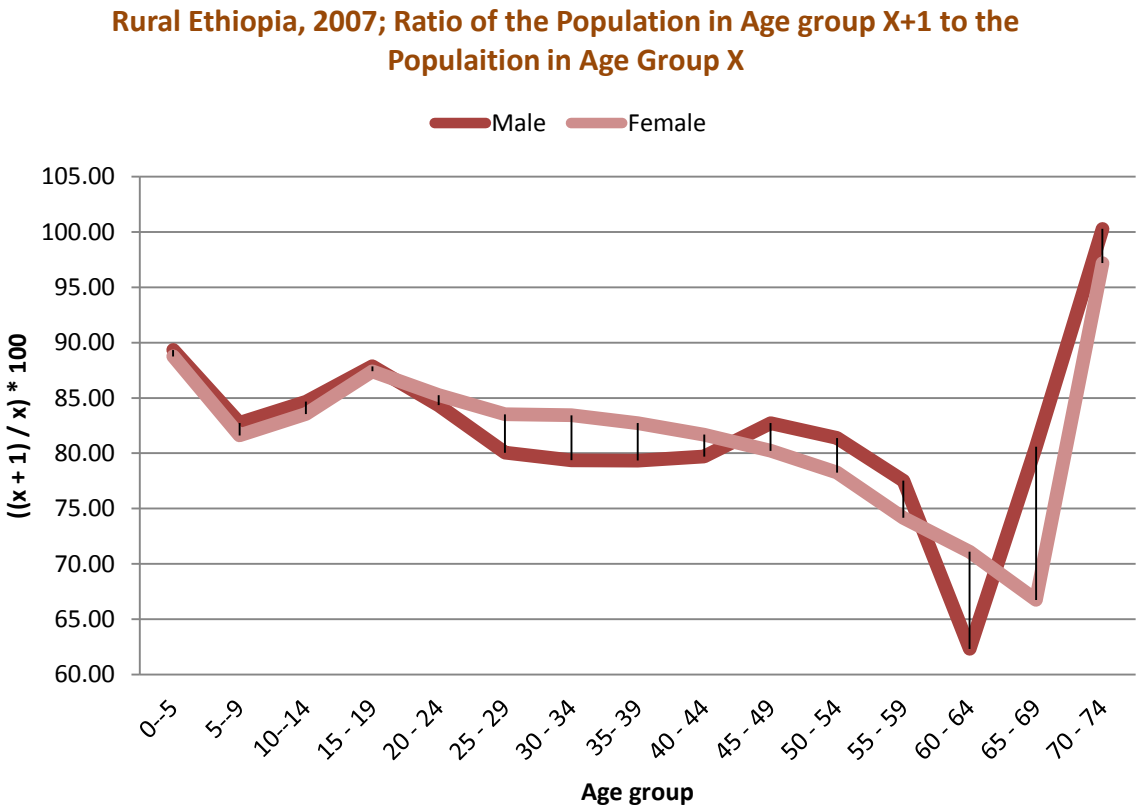
Ratio of populations in five-year age groups, to the preceding group

The table below shows the number of persons in an x+1 age group for every 100 persons in the x age group. For example there are 89 males in the 5 – 9 age group for every 100 males in the 0 – 4 age group (see the graph below). A similar table and graph are added at the bottom (United States 2000) for comparison. The table can also be regarded as a crude indicator survivorship probabilities at higher ages but migration can complicate the picture. It cannot be use as such at lower ages especially when fertility levels have changed in the immediate years prior to the census date. For example, the ratio of boys in rural Ethiopia in the 5 – 9 age group surviving to those in the 0 - 4 age group is 0.89 suggesting an 89/100 survivability rate. The ratio in the United States is 107/100 because there are more boys in the 5 – 9 age group than in the 0 – 4 age group due to higher fertility in the 5 – 9 years prior to the year 2000 census than the last five years before the census. The graph and table for rural Ethiopia, show that the female "mortality advantage" commonly observed among virtually all human populations is absent with the exception of the four age groups between age 25 and 45 and in the 60 - 64 age group. Note, also, the very prominent female "mortality advantage" above age 65 displayed in the US graph which is shown for rural Ethiopia as a sizeable male advantage.

Rural Ethiopia, 2007: Ratio of Age category (x+1) to x, 2007

Age	Category	Males	Females	Ratio ((x+1)/x)*100	
Group	(x)			Males	Females
0 - 5	1	5477074	5415921	89.33403	88.75098
5 - 9	2	4892891	4806683	82.74292	81.60611
10 - 14	3	4048521	3922547	84.66386	83.55071
15 - 19	4	3427634	3277316	87.85991	87.40829
20 - 24	5	3011516	2864646	84.35027	85.25207
25 - 29	6	2540222	2442170	80.04466	83.51532
30 - 34	7	2033312	2039586	79.36805	83.43453
35- 39	8	1613800	1701719	79.34205	82.72635
40 - 44	9	1280422	1407770	79.69951	81.68948
45 - 49	10	1020490	1150000	82.72095	80.22913
50 - 54	11	844159	922635	81.38597	78.25955
55 - 59	12	687027	722050	77.51675	74.16509
60 - 64	13	532561	535509	62.30817	71.0974
65 - 69	14	331829	380733	80.58247	66.74415
70 - 74	15	267396	254117	100.2805	97.19853

Source: Based on [28]



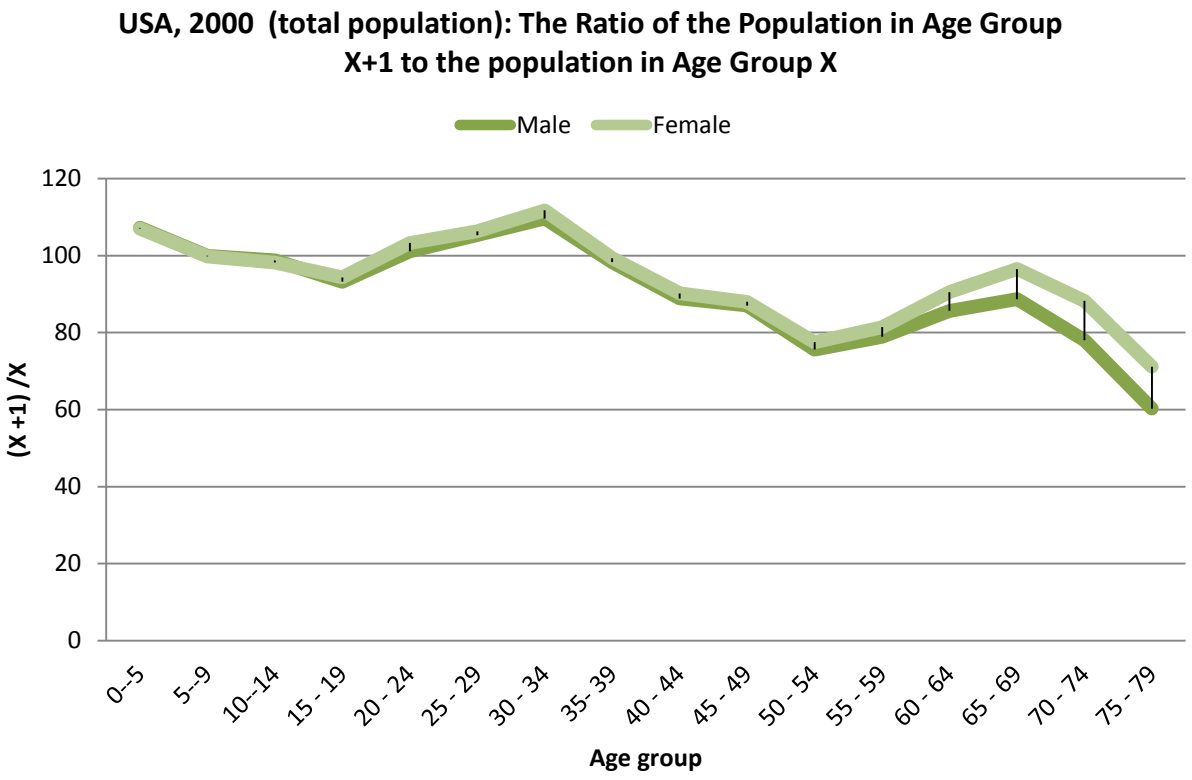
The ratio of successive age groups ((x + 1) / x) would look like the following (see graph and table below) when based on high quality data.

Click here to return to the main page → www.EthioDemographyAndHealth.Org

USA, year 2000: Ratio of Populations in Age category (X+1) to Populations in X

United States , 2000				((x+1)/x)*100	((x+1)/x)*100
Age group	Males	Females	Category	Males	Females
0-5	9,810,733	9,365,065	**1**	107.2629	107.0599
5- 9	10,523,277	10,026,228	**2**	99.97073	99.81695
10 -14	10,520,197	10,007,875	**3**	98.77195	98.21152
15-19	10,391,004	9,828,886	**4**	93.2327	94.37679
20- 24	9,687,814	9,276,187	**5**	101.1452	103.303
25-29	9,798,760	9,582,576	**6**	105.3375	106.3244
30- 34	10,321,769	10,188,619	**7**	109.6585	111.7715
35- 39	11,318,696	11,387,968	**8**	98.32495	99.33959
40- 44	11,129,102	11,312,761	**9**	88.86167	90.18928
45- 49	9,889,506	10,202,898	**10**	87.03897	87.99288
50- 54	8,607,724	8,977,824	**11**	75.61498	77.53001
55- 59	6,508,729	6,960,508	**12**	78.91905	81.44262
60- 64	5,136,627	5,668,820	**13**	85.66637	90.55117
65- 69	4,400,362	5,133,183	**14**	88.69525	96.51963
70- 74	3,902,912	4,954,529	**15**	78.00473	88.22952
75- 79	3,044,456	4,371,357	**16**	60.27011	71.15571
80- 84	1,834,897	3,110,470	**17**	66.87013	96.85318

Source: Based on [10]



References:

1. http://en.wikipedia.org/wiki/Census
2. http://www.statistics.gov.uk/census2001/cb_8.asp
3 Principles and Recommendations for, Population and Housing Censuses Revision 1, Department of Economic and Social Affairs, Statistics Division, Statistical Papers Series M No. 67/Rev.1 United Nations, New York, 1997
4. http://unstats.un.org/unsd/pubs/gesgrid.asp?ID=127
5. http://adsri.anu.edu.au/pubs/BAPS/BAPSChap2b.pdf
6. Tilaye Geressu Birru, Gender *Statistics and the Status of Civil Registration and Vital Statistics in Ethiopia*. Global Forum on Gender Statistics. ESA/STAT/AC.140/8.2 December 2007.
7. http://www.pcmag.com/encyclopedia_term/0,2542,t=sampling&i=50790,00.asp
8. http://www.csa.gov.et/
9. Gordon Brackstone, *Managing Data Quality in a Statistical Agency*, Statistics Canada, Survey Methodology, Catalogue No. 12-001-XPB, Vol. 25 No. 2, December 1999.
10. http://www.censusscope.org/us/chart_age.html