

UTILIZATION OF STATISTICAL INFORMATION: THE CASE OF POPULATION STATISTICS*

By M. Khalid Siddiqui**

Introduction

Statistical information is essential for effective development planning, policy formulation, action programmes and monitoring of the progress made. It is fundamental to a democratic society for informed decision making by its institutions as well as individuals.

Historically, statistics were gathered by the state for administrative and planning purposes. The church, often affiliated with the state, also maintained records of baptism, marriages and burials which had the potential of providing population statistics. In fact, population data are by far the oldest among the statistics collected by the state as they provided the number, ethnic and other composition, and geographic distribution of the people. In the region covered by the Economic Commission for Asia and the Pacific (ESCAP), evidence suggests that some sort of record-keeping concerning people existed quite some time ago, for example in 668 AD in the United Sila Kingdom of Korea. In Asia, the colonial rulers introduced civil registration soon after their arrival, such as in India in the 1860s. In the ESCAP region most of the government census and statistics organizations were established in the twentieth century, while in some cases they had already been constituted towards the end of the nineteenth century.

Until the formal collection of statistical information for research began in the modern times, the collection, dissemination and use of data largely remained with the government departments. However, with the spread of democratic norms, statistics began to be perceived as public good. Slowly, the notions of 'government statistics office' and 'government statistician' began to wane and were replaced by the 'national statistical offices (NSOs)'. Now the national statistical offices in a democratic society strive to assist the process of decision-making, research and discussion not only within the government but also in the community. It is well

* Paper presented at „Asia-Pacific POPIN-IFLA (International Federation of Library Association) Joint Workshop on Demographic Information“ at Bangkok on 26th August, 1999.

** The views expressed in this paper are those of the author, and do not necessarily reflect those of the United Nations.

recognized that, the privacy of individuals notwithstanding, all potential users of data – including the general public, the civil society and the private sector – should have easy and equal access to basic official statistics.

Soon after statistics began to be perceived as public good, the demand for quantitative information rose. With limited resources at hand and multiplying demand for statistical information, NSOs introduced procedures for prioritizing their data collection through such institutions as data user committees for determining the topics to be collected in population and housing censuses and surveys. In some countries this process has become an extensive dialogue between the NSO and the community. Effective interaction between the data producers and data users is now considered an important element and a major aim of modern national statistical services. Now different interest groups are involved at a very early stage of data collection planning for which various user groups and advisory bodies are used. Similarly, users' views are taken into consideration at the dissemination stage. However, since the timing of the release and access to statistical information may have implications for the markets, election results and other processes in the society, many NSOs follow transparent procedures for data dissemination by announcing schedules of data releases in advance so as to give equal access to information to all concerned.

Focusing on the area of population statistics, this paper emphasizes the need for paying greater attention to the utilization of data. It argues that a great amount of data produced by the national statistical agencies remains grossly underutilized while the demand for new data has been multiplying at a fast rate in this era of information revolution. Therefore, concerted efforts are required by all concerned to promote the effective exploitation and fuller utilization of data. That will also help rationalize the production of additional data.

I Population Statistics

Strictly speaking, demography is the study of the size, distribution, structure and growth (or decline) of populations. Additional characteristics of the population, such as ethnic, social, and economic are also of interest, while the focus on reproductive health and family planning adds a host of new dimensions to population studies. Therefore, the term 'population statistics' has replaced 'demographic statistics' to reflect the wider sense of the field of demography.

Population statistics are not only vital for development planning, but are also utilized for decision making in the government, in industrial and other businesses, and for stimulating research and debate on important issues of concern to the society. Basically, there are three sources of population statistics: administrative

records, population and housing censuses, and household surveys. In most developing countries of the ESCAP region, administrative systems of data are scarce or inadequate for statistical purposes and many of the countries lack suitable programmes of household surveys to fully meet their data needs. Consequently, population census becomes the center of attention of many data users, perhaps more than the census statisticians realize. Under these circumstances, the users of data often experience frustration when the population census statistics do not meet their expectations. While the population census cannot meet all the wide ranging demand for data in these countries, such as for short-term monitoring, its dominant role must be duly recognized.

A population census is a process consisting of many phases involving extensive resources. To meet the broad requirements of potential data users, census organizations consult data users at the questionnaire designing stage as well as receive feedback on the needs of major groups of data users in terms of topics of interest, desired accuracy, timeliness, disaggregation and mode of dissemination. Being one of the most expensive statistical operations, the high cost of a population census would not be justified unless its data were fully utilized. In that regard, timely dissemination of data assumes utmost importance.

Population and housing censuses provide a consistent series of statistics on the size, distribution, and composition of the population of a country and for its geographic, administrative units and social groups, permitting monitoring of long-term changes of those variables. The population and housing censuses together with other sources of data, such as vital registration and household surveys, constitute an integrated system of statistics in modern societies. While for many conventional indicators the numerators come from other sources, the population census often provides the denominators, such as for per capita food consumption. One important strength of the population census lies in the fact that it can provide information for smaller administrative units and communities.

A. Changing shape of census reports

Until the decade of 1970, generally the census and national statistical offices released numerous and voluminous census reports which required huge shelf space in the libraries. However, trend towards reducing the size and volume of the published census reports gradually caught up, as it became possible to disseminate detailed data on electronic media such as magnetic tapes and on-line access. Now it is very common to release a census reports with some analytical commentary, accompanied by several thematic reports such as on fertility, mortality, population projections, women, the aged, the disabled and the family. Since the users who require detailed data on a given topic are relatively small in number and often

need special attention, their needs are met through special tabulations and supply of data on demand.

The analytical nature of the thematic reports varies from country to country. Some are prepared as monographs with detailed analytical chapters, while others focus on the release of more detailed tabulations in separate reports on topics such as demographic, social and employment characteristics. The inclusion in the main census report of some interpretation and analysis of summary tables has sharply increased. Apart from presenting summary tables and comparative data from previous censuses, some main census reports also include detailed analyses on selected topics of common interest such as demographic trends, internal migration and projections of major demographic and socio-economic aggregates.

II. Understanding the Nature of Population Data

Population Statistics and indicators are derived from various sources, either from a single source or a combination of several. In developed countries, the adequacy of the required data sources ensures the production of good quality information. However, in developing countries, the quality is not always assured due to gaps in the data collection systems, especially the civil registration and vital statistics systems. The United Nations publication entitled *Population and Vital Statistics Report* under the Statistical Papers Series (ST/ESA/STAT/SER.A) provides vital statistics for each country as well as information on the year of latest census and the extent of completeness of the registered data. The meta information included in this publication also reveal that even for simple measure as mid-year population, the estimates are not strictly comparable across countries. In many cases, they have to be estimated by the United Nations, utilizing information at hand and demographic techniques.

Despite the inadequacies of vital statistics systems, for some purposes, demographers are able to produce reasonable estimates for selected population parameters. However, often the data users fail to make a distinction between statistical estimates based on scientific models and the data produced from the sound statistical data sources. Attention to meta information is crucial for effective utilization of data. Again, the simple indicator of total population (by age or any other characteristic) may be any of the following: based on the census and adequate civil registration, an estimate that refers to mid year or beginning of the year, a figure based on population projections, or an estimate made by an outside agency. The situation becomes more complex with rates and indicators such as life expectancy. For the latter, the data requirements are relatively extensive. As a

rule of thumb, when the civil registration is inadequate, the basis of computation of a complex statistics should be reviewed and the estimate be corroborated.

III. Utilization of Poulations Data

The recognition of population policies and programmes as an integral parts of social, economic and cultural development has important implications for statistical data collection and analysis. While the demand for additional and more sophisticated data will continue to pose challenges for national statisticians and demographers, part of the demand for information can be met through further analysis and utilization of multiple sources of data.

The large quantity of census data produced by the statistical agencies remains under-utilized. In many developing countries this is mainly due to the lack of sufficient analytical capabilities, within the statistical office or even in planning departments and other potential research institutions. However, as will be discussed later, the timely availability and usable format of dissemination of census data also has important bearing on data utilization. Often the availability of census report becomes the guide for further research and that role of the census report perhaps has not been realized in many instances. Otherwise the commentary and interpretations included in the report, and the preliminary analysis of the topics of contemporary interest to the country, such as migration or urbanization, could be highlighted to invite researchers for further in-depth investigations. The introduction of analysis also adds value to the statistical products and satisfies some of the unmet needs of data users who may not have analytical capabilities, such as the non-governmental organizations(NGOs).

Table 1 provides some examples of potential uses of census data, by providing illustrative tabulations usually released by NSOs. Unless pointed out in this manner, the link between the tabulation and the uses may not be apparent to many consulting the census reports.

A. Factors Influencing Data Utilization: Experience in the ESCAP region

To be effective, census reports require adequate resources, attention and publicity to encourage their use. The inclusion of some commentary and interpretation has been found useful and interesting. The inclusion of information on data quality and limitations of census results would also facilitate the utilization and interpretation of the census results. Dissemination of information on definitions, standards and classifications used, and their comparison with past censuses, would also promote effective utilization of census data. In the absence of such information, there is danger of misinterpreting data. For example, the absence of

information on changes occurring in the boundaries of urban and rural areas over time would make it difficult to study population movement based on data from succeeding censuses. Many countries publish information on boundary changes, and they should be made available together with other census publications.

Tabulations produced should be usable and conveniently understood by users. Multidimensional cross-tabulations which are often included in census reports may at times be difficult to interpret. The utility of those tabulations may be enhanced if summary sub-tables at appropriate levels are also included. Otherwise data users not comfortable with complex tabulations may be discouraged to use the report. Also, researchers in some of the developing countries feel that many census statistics are not available in a form that make them useful, or statistics beyond published material are not accessible. While it may not be possible to release all possible tabulations from a census, adequate provisions should be made for the release of unpublished tabulations. For instance, the need for providing census tabulations containing data on the characteristics of data on small communities has been voiced. The availability of basic census tabulations for small communities have been found useful for research on population distribution and to analyses needed to aid planning in many fields, including housing, education, urban development, transportation, public utilities etc.

IV. Improving Data Utilization

A large amount of data remains underutilized for a variety of reasons: users may not be fully aware of the range of data that can be made available by the NSOs; producers may not publicize their products and services well enough; data may not be available when users need them most; statistical information may not conform to the form required by the users; poor interaction between the users and producers of statistics; access to data may not be easy; poor understanding of data limitations and strengths among the users; poor skills of research and analysis among the potential users. Considerable progress can be made by addressing the above mentioned issues and by establishing networks and partnerships.

A. Role of the producers of data

Effective communication between the producers of statistics with existing and potential users consists of involving users at all stages of the census, such as planning, questionnaire design and tabulation plan. Moreover, the cooperation of researchers and academics in the analysis of census data, conducting national and international seminars to disseminate census results, educating users and potential users on the strengths and limitations of data, and providing access to

databases containing data from current and past censuses are important. In some countries, the user inputs are sought in the production of draft chapters for the census report.

There is a need for the statisticians to play a more effective role in fostering population census data analysis. In the first place they should pay greater attention to the publicity of products and services. Information brochures, catalogues, press briefings, advance announcements of data releases, and user workshops are among the several approaches adopted by advanced statistical offices.

However, NSOs should also take the lead by demonstrating the use of statistical information for a better understanding of issues of public policy and various phenomena of interest to the society at large. The thematic publications now produced from the census data provide that opportunity. Moreover, by releasing data on a timely basis, the NSO would enhance the value and utility of their outputs.

B. Role of the users of data

The data users also play an important role through active participation in the consultative process and by providing feedback. They should take greater interest in understanding the strengths and limitations of the statistical system. The expression of the felt needs to the budgetary authorities would also ensure resources for priority areas. The greater use of meta information and demand for it would encourage NSOs to release such information.

The questionnaires used in censuses, surveys, and registration forms provide clues to the sort of data that could be available from the NSOs. By familiarizing themselves with those and other key documents, users will be better prepared to request data and tabulations.

C. Role of the libraries

Since many data users, including the general public, students and researchers, often turn to libraries for their data needs, libraries can play an important role in promoting effective data utilization. Census data catalogues, release calendars and press releases constitute important sources indicating the potential availability of data. Some NSOs also issue annual reports and forward work programme which may also provide updated information on recent and forthcoming products as well as contact addresses for obtaining information on request. In addition to census reports, efforts should also be made to obtain other relevant publications, such as on data evaluation and boundary changes. Now some advanced NSOs provide some meta-information on key outputs on their website.

While hard copy publications remain a dominant mode of data dissemination, NSOs are increasingly utilizing electronic media, such as the Internet services, floppy disk, CD-ROM, and magnetic tape. Of these, CD-ROM may turn out to be a preferred mode among the library users as it provides easy access to massive data. There is growing realization among the Internet users that, except for browsing small sets of data, it is quite burdensome and time consuming for downloading large amounts of data. Since in developing countries the access per capita to the computer facilities is exceedingly low, the availability of access to data on CD-ROM in the libraries would no doubt boost data utilization. In turn, the NSOs would be able to increase the number of CD-ROM produced.

Reference material to help data users may include dictionaries and encyclopedias for a better understanding of the technical terms. A glossary of selected terms is included in this paper as an illustration. A selected list of reference material follows:

R. Presat (1985). *Dictionary of Demography*

J. Ross (1982). International Encyclopedia of Population

Population Reference Bureau (1991). Population Handbook, 3rd edition.

United Nations (1983). Manual X: Indirect Techniques for Demographic Estimation (ST/ESA/SER.A/81)

United Nations (1998). *Principles and Recommendations for Population and Housing Censuses*, revision 1 (ST/ESA/STAT/SER.M/67/Rev.1)

Table 1

Illustrative Tabulations by Potential Use of Census Data

Illustrative tabulation	Potential use
<i>Total population by major and minor civil divisions, by sex</i>	<ul style="list-style-type: none"> - Studies of internal migration and other factors affecting the trends of population distribution. - Studies of the internal distribution of population in relation to economic and social facilities. - Base data for projections of the future regional and urban-rural distribution of the population.
<i>Decadal variation in population by major and minor subdivision, by sex</i>	<ul style="list-style-type: none"> - For studying population growth and trends.
<i>Population by locality size and sex</i>	<ul style="list-style-type: none"> - Studies of national pattern of concentration and dispersion of population. - Studies of urbanization.
<i>Population by geographic division and age, by sex</i>	<ul style="list-style-type: none"> - Construction of life tables. - Current population estimates and projections. - Estimation of the size and distribution of population groups, such as school age, voting age etc. - Analysis of population change.
<i>Population in households by relationship to head of household, marital status and sex</i>	<ul style="list-style-type: none"> - Studies on changing patterns of household compositions and of characteristics of heads of households.
<i>Population by type of activity, age and sex</i>	<ul style="list-style-type: none"> - Studies of the growth and distribution of manpower. - Studies of determinants of the size and composition of manpower. - Manpower projections. - Economic activity rates. - Studies of unemployment and underemployment.
<i>Population by duration of residence in locality and major civil subdivision, age and sex</i>	<ul style="list-style-type: none"> - Studies of internal migration. - Future population estimates by localities.

Glossary

BALANCING EQUATION. A basic demographic formula used to estimate total population change between two points in time. The balancing equation includes all components of population change: births, deaths, immigration, and emigration.

BIRTH RATE (or crude birth rate). The number of births per 1,000 population in a given year.

CHILD-WOMAN RATIO. The number of children under 5 years old per 1,000 women aged 15-49 years in a population.

CLOSED POPULATION. A population with no migratory flow in or out.

COHORT. A group of people sharing a common temporal demographic experience who are observed through time. For example, the birth cohort of 1900 would be the people born in that year.

DEATH RATE (or crude death rate). The number of deaths per 1,000 population in a given year.

DEPENDENCY RATIO. The ratio of the economically dependent part of the population to the productive part; arbitrarily defined as the ratio of the elderly (those 65 years and over) plus the young (those under 15 years of age) to the population in the "working ages" (those 15-64 years of age).

GENERAL FERTILITY RATE. The number of live births per 1,000 women aged 15-49 years in a given year.

GROSS REPRODUCTION RATE (GRR). The average number of daughters that would be born alive to a woman (or group of women) during her lifetime if she passed through her childbearing years conforming to the age-specific fertility rates of a given year.

INFANT MORTALITY RATE. The number of deaths to infants under one year of age in a given year per 1,000 live births in that year.

LIFE EXPECTANCY. The average number of additional years a person would live if current mortality trends were to continue.

LIFE TABLE. A tabular display of life expectancy and the probability of dying at each age for a given population, according to the age-specific death rates prevailing at that time.

MARITAL FERTILITY RATE. Number of legitimate live births per 1,000 married women aged 15-49 years in a given year.

MATERNAL MORTALITY RATE. The number of deaths to women due to pregnancy and childbirth complications per 100,000 live births in a given year.

NATURAL INCREASE. The surplus (or deficit) of births over deaths in a population in a given time period.

NET REPRODUCTION RATE (NRR). The average number of daughters that would be born to a woman (or group of women) if she passed through her lifetime from birth conforming to the age-specific fertility and mortality rates of a given year. This rate is

similar to the gross reproduction rate, but takes into account that some women will die before completing their childbearing years. An NRR of 1.00 means each generation of mothers is having exactly enough daughters to replace itself in the population.

PERINATAL MORTALITY RATE. The number of fetal deaths after 28 weeks of pregnancy (late fetal deaths) plus the number of deaths to infants under 7 days of age per 1,000 live births.

POPULATION PROJECTION. Computation of future changes in population numbers, given certain assumptions about future trends in the rates of fertility, mortality, and migration. Demographers often issue low, medium, and high projections of the same population, based on different assumptions of how these rates will change in the future.

POPULATION PYRAMID. A special type of bar chart that shows the distribution of a population by age and sex.

RATE OF NATURAL INCREASE. The rate at which a population is increasing (or decreasing) in a given year due to a surplus (or deficit) of births over deaths, expressed as a percentage of the base population.

REPLACEMENT LEVEL FERTILITY. The level of fertility at which a cohort of women on the average are having only enough daughters to "replace" themselves in the population. By definition, replacement level is equal to a net reproduction rate of 1.00. The total fertility rate is also used to indicate replacement level fertility; in the more developed countries, a TFR of 2.1 is considered to be replacement level.

SEX RATIO. The number of males per 100 females in a population.

STABLE POPULATION. A population with an unchanging rate of growth and an unchanging age composition, because of age-specific birth and death rates having remained constant over a sufficiently long period of time.

TOTAL FERTILITY RATE (TFR). The average number of children that would be born alive to a woman (or group of women) during her lifetime if she were to pass through her childbearing years conforming to the age-specific fertility rates of a given year.

ZERO POPULATION GROWTH. A population in equilibrium, with a growth rate of zero, achieved when births plus immigration equal deaths plus emigration.

Source: The Population Reference Bureau, *Population Handbook*, 3rd edition, 1991

References

1. Censuses of Population and Housing in Asia and the Pacific: Towards the 1990 Round (Bangkok, ESCAP, 1988)
2. Lee-Jay Cho and Robert L. Hearn, ed., Censuses of Asia and the Pacific (Honolulu, East-West Population Institute, East-West Center, 1984)

3. Report of the ESCAP/SPC Pacific Working Group on the 1990 World Population and Housing Census Programme, STAT/PWG/WPHCP/7 (Bangkok, ESCAP, 1987)
4. Report on Population Data and Utilization Seminar, 15-18 October 1985 (Apia, Samoa, Department of Statistics, 1985)
5. Report of the Regional Working Group on the 1990 World Population and Housing Census Programme, STAT/WPHCP/14 (Bangkok, ESCAP, 1986)
6. Status of Civil Registration and Vital Statistics in Asia and the Pacific (Bangkok, ESCAP, 1987).
7. Siddiqui, M. Khalid, "Analysis and utilization of population census data in the ESCAP region", paper presented at the Fifteenth Population Census Conference, Seoul, Republic of Korea, 17-20 August 1993.

M. Khalid Siddiqui
Statistics Division
ESCAP
The UN Building
Rajadamnern Ave.
Bangkok 10200
Thailand