

LEVELS AND MODES OF DISSEMINATING CENSUS DATA IN PUBLIC DOMAIN

J. K. Banthia

1. While it is relatively easy to organize and conduct a census, it is a bit difficult to process the data and still more difficult to disseminate data quickly, cheaply and widely to the entire range of data users. This is partly due to the fact that the census organisation gets a bit tired while undertaking the first two activities of conducting and processing the census data. Further, many census organisations may not have the requisite in-house capabilities to utilize the wide choices of the technology now available for dissemination of data. Several organisations are additionally hampered by lack of expertise to do any meaningful analysis of the census data and at times engage only in compilation and publication of the bare minimum results. This background paper deals with the specific issue of the 'Levels and modes of disseminating Census data in public domain', with a view to share the experience of several countries in this sphere.

2. While most countries customarily release only the final population totals along with several socio economic and demographic characteristics, a few countries have an advance well defined programme of releasing the provisional population totals and certain specified population characteristics. This initial release is followed by the normal publication of the final population totals etc. The key question is whether it is good policy to release provisional data, particularly, if it does not suffer from inaccuracies and is neither a financial or administrative burden on the organisation. This approach is particularly helpful where the gap between the conduct of census and release of the bare population totals is six months or more.

3. Conceptually, although it is necessary that the census tabulation plan includes all the characteristics of the population canvassed in the census schedule, it is likely that the data on some of the population characteristics collected are not exploited fully. This is particularly important when the population characteristics are tabulated only as a single variable and not cross tabulated with other population characteristics except of sex, place of residence (rural or urban) and in terms of geo-political administrative units. Usually all population characteristics are tabulated by sex (including the combined total population), by residential status in terms of rural/urban classification and well defined geographical units such as country, states/province, district/county and town. However, it is not known whether even the bare population totals are published for all the existing lowest administrative levels such as villages in rural areas or wards in towns, for example. Various countries have different

administrative units and therefore it would be important to know whether for all the administrative levels the population details are published (or not published). The importance of small area statistics and the strength and ability of the census to provide this important data set is one of the most important argument put forward for conducting census while seeking financial allocations from the Government. It is important to critically analyse as to what extent this unique feature of census for producing small area statistics is being complied with.

4. While it is customary for most of the countries to produce the basic population totals on a 100% count at the country, state/province, district, cities/towns and villages, several other tables are neither produced on a 100% counts nor at all levels. The choice of whether the tabulation plan should be on 100% basis or on a sampling basis largely depends upon two factors. First, the cost involved in processing and tabulating 100% of the data collected. Second, the critical time factor since processing, compilation and validation of the census data on a 100% basis is a time consuming process and leads to considerable delay in release of the census data.

5. There has been a debate that if the census schedule includes several questions which are not subsequently processed and tabulated on a 100% basis, then, why in the first instance the population characteristics be collected of all individuals. This obviously is a big drain on the designing of the questionnaire, training of the enumerating agency and finally an avoidable burden on the participating population and the enumerator. It is, therefore, important to critically examine whether there is a mis-match in collection of the population characteristics on a 100% basis but tabulating it on a sampling basis. Some countries have already taken note of this and introduced the short form and the long form as part of the census exercise. Whereas the short form has certain basic questions to be canvassed universally among the population, the long form contains very detailed questions on a wide variety of subjects but is canvassed only to a select pre-determined population on a sample basis. It may be important to identify the population characteristics which are collected on a universal basis but tabulated on a sample basis.

6. While it is true that geo-political or administrative unit is the basic criterion to decide the levels at which the census tables are generated, usually there is a demand to make available data for certain other configuration of the geo political units not commonly utilized. These specially customized tables may or may not be generated by census organisation and therefore leave behind a data gap. For example, in India the normal level of publication of the census tables is determined by the administrative boundaries in terms of country, the

state and thereafter districts/cities/towns and villages. The political boundaries of the elected members to the national Parliament and the other Legislatures are not considered for census tabulation purposes in India. Similarly at the local level, the census tabulation plan does not ordinarily meet the requirements of the constituents of the rural local governments (such as Zila Parishad, Panchayat Samities and Village Panchayats). Even for urban local governments such Municipal Corporations of cities/towns, this requirement is not fulfilled. Similarly, the police districts, the telephone districts etc., have different jurisdiction from the administrative boundaries and though theoretically it is possible to tabulate data for any particular combination of the administrative unit, until now it has not been done. In such cases, however, the combination of various small administrative areas is required for generating other desired levels of tabulation and advance work needs to be done in arranging the census data base at the time of its processing. Are similar problems faced in your country of several levels of geo-political and administrative boundaries overlapping each other? How is this problem dealt with while planning and producing census tables and results?

7. Gender issues in planning the development of a society have become very prominent in recent times and usually the census publishes data for each sex separately. However, certain countries may have developed specialized features for tabulating gender based data on population characteristics. This is a relatively new field and it would be important to exchange the work done in this area. For the 2001 Census tabulation plan, India has proposed several new gender based tables for the first time (Annexure B). These special tables help in focusing the various efforts undertaken to improve the status of women and also highlight the areas where females have a clear disadvantage as compared to the males. A related issue is preparations of tables for female headed household. This obviously means that the census tables are generated for the household also in addition to tables based on individual characteristics of the population. Usually the household tables are generated more for a socio-economic investigation rather than limiting the inquiry into the nature of the composition and size class of the household. It would be important to know the number of tables generated based on individual characteristics and those based on the household characteristics.

8. The traditional approach for dissemination of the basic census data has been by way of printing hard copies. It is supplemented by various population charts and special analytical volumes from time to time. However, since the 1980s huge changes have occurred in the computer hardware and software industry, and it has become increasingly possible to disseminate data on the electronic media through floppies and subsequently

through CDs etc. Now several organisations have started disseminating census data through internet and dedicated web sites. Another effective mode of data dissemination has been through the GIS. While all these modes of data dissemination are pretty attractive yet they could be extremely expensive. A number of countries have therefore introduced a dual policy of releasing census data partly free of cost and partly at a certain cost determined by their organisation. It is the second part which varies considerably from country to country depending upon the philosophy as to whether the public at large needs to pay heavily again for the census data since conducting of the census has already been financed through the taxpayers money. However, in certain countries the approach is now to become increasingly commercial and most of the data is made available only at cost even when it has not been particularly customized. Customized data in any case is an extremely expensive proposition and generally demanded only by the commercial organisations. It may be important to share the various policies formulated by respective countries on dissemination of data free of cost, at some cost and on a customized basis. Certain countries have also a policy of making available the raw processed but un-tabulated data for research purposes. This is largely a pre-determined sample and ensures that the confidentiality of the individual is not revealed. However, the policy for making the raw data available including methodology of concealing the identity particulars are not well known particularly among the developing countries. It would be good idea to share the knowledge on the methodology adopted for concealing the individual identity particulars and the availability of the raw data for purposes other than research.

9. To recapitulate, It is also important to know the time frame by which the basic population totals and population characteristics are made available in the public domain. Is there a pre determined time table for this and who authors it - the census organization or the Government? Does the census organisation produce independent analytical reports on the census data or merely confines itself to the publication and release of the basic census tables? What are the inhouse capabilities created on a permanent basis to assist the organisation in analysis of the data? Ultimately, is it left to the several government departments to analyse the data themselves or the census organization provides them a helping hand. Are some of the analytical reports contracted out to various consultants? Is there a working relationship between census organisation and the universities to produce quality analytical reports on the census data? These are some of the important and critical questions to understand so that the power and potential of the census data is utilized fully and effectively. A suitable responses on these issues will help in deciding the future of the census heavily since governments are finding it increasingly difficult to fund such an

expensive exercise and are looking for best value for the money spent in census undertakings.

- Note:-
1. Census Organisation here includes both an independent Census Organisation and the National Statistical office.
 2. Census includes both the population and the housing census, although for the sake of simplicity only population characteristics have been discussed.

Encl:- Annexure A and B

A FEW QUESTIONS ON CENSUS

Country:

Represented by:

Date:

01. When was the last population census conducted in your country?

02. Who releases the first Census results?
(Tick mark ✓)

- (a) President/Prime Minister (...)
(b) Minister of Statistics (...)
(c) Any other Minister
(Specify) _____
(d) Head of the Census Division (...)
(e) Any other person
(Specify) _____

03. Does your country release provisional census results?

Yes/No

If yes, how many weeks after the date of conduct of census?

04. What is the time lag between the date of conducting census and the

(Tick in appropriate place)

(a) Release of first result?

- (i) Up to 3 months (...)
(ii) 3 to 6 months (...)
(iii) 6 to 12 months (...)
(iv) 12 to 24 months (...)
(v) 24 to 36 months (...)
(vi) Above 36 months (...)

Note: All responses are with reference to the most recent Census conducted in your country

04 (Contd). What is the time lag between the date of conducting census and the

(Tick in appropriate place)

(b) Release of last result?

- (i) Up to 3 months (...)
(ii) 3 to 6 months (...)
(iii) 6 to 12 months (...)
(iv) 12 to 24 months (...)
(v) 24 to 36 months (...)
(vi) Above 36 months (...)

05.* How many population characteristics (Questions) were canvassed in the recently concluded Census (including both long and short form)?

06.** For how many population characteristics only a univariate table is produced?

07.*** Was there any characteristic not finally tabulated and published? (Tick mark ✓)

0 (...), 1 (...), 2 (...),

3 (...), More than 3 (...)

08. Is there a published Census Tabulation Plan indicating the various tables and levels of presentation?

Yes

No

09. Which are the geo-political (or administrative) levels at which the Census tables are published? (Give equivalent levels as applicable in your country) (Tick 'Yes' or 'No')

Level 1 : Country	Yes/No
Level 2 : State	Yes/No
Level 3 : District	Yes/No
Level 4 : Sub-district	Yes/No
Level 5 : City	Yes/No
Level 6 : City Ward	Yes/No
Level 7 : Village	Yes/No

10. How many Census tables will be produced in the recently concluded census?

11. How many Census tables for small area statistics (the lowest administrative level) will be published?

Note:

* - Count characteristics as follows:

Name – 1, Sex – 2, Literacy Status – 3, Level of education – 4 etc.

** - Census tables are usually generated for each characteristic by sex, geographical area and rural/urban status.

*** - For example in 1991 Census although, we in India, canvassed a question on ex-defence personnel, finally no table was published.

12.	Are all the Census Tables produced at all these levels on 100% sample basis?	Yes/No
13.	If answer to Question 12 is 'No' then how many tables are produced on a sample basis ?	_____
14.	Does your country canvass only one form (Long Form) of Questionnaire/Schedule?	Yes/No
15.	If the answer to Question 14 is 'No', then does your country canvass both 'Long Form' and 'Short Form'?	Yes/No
16.	Does your country canvass Long Form on 100 % or sample basis?	(1) 100 % basis (2) Sample basis (<i>Specify</i>)
17.	Does your country prepare any exclusive table on women? (<i>If yes, please provide list</i>)	Yes/No
18.	Does your organisation only publishes the Tables or also brings out analytical reports? (<i>Please Tick</i>)	(1) Publishes only Tables (2) Also publishes Analytical Reports
19.	Does your organisation have in-house capacity for preparing analytical reports or is the work allotted to outside agency?	(1) Have in-house capacity (2) Entrusted to outside agency
20.	Which of the following modes of data dissemination is (are) used by your organisation? (<i>Please tick as many as applicable</i>)	1. Hard Copy in printed books 2. Charts 3. Data Sheets 4. On the Internet 5. CD- Rom 6. Customised tables as requested by users 7. Any other (<i>Specify</i>)
21.	Does your organisation have their own website for disseminating census data on the Internet?	Yes/No

-
- | | | |
|-----|---|---|
| 22. | If answer to above question is 'Yes', then indicate restrictions in providing data free of cost at the site. (<i>Please tick</i>) | 1. No data is provided free
2. Only country level data are provided free
3. Complete census data sets are provided free
4. Only tables generated on users requested are provided at a cost, rest are free
5. Other (<i>Specify</i>) |
| 23. | Does your organisation provide record level data on a sample basis after deleting sensitive fields to data users? | Yes/No |
| 24. | Does your organisation disseminate census data by way of GIS based thematic maps? | Yes/No |
-

Thank you for your time and effort

Office of the Registrar General, India
Government of India
2A, Mansingh Road, New Delhi 110011, India
rgoffice@censusindia.net

List of new gender based tables proposed for the 2001 Census

- ❖ Households by number of female literates among female members of the household age 7 years and above (Also separately for Scheduled Castes and Scheduled Tribes - historically marginalised groups)
- ❖ Households with at least one member age 15 years and above and with or without educational level matriculation and above by household size
 - Number of households with at least one female matriculate
 - Number of households with at least one female graduate.
- ❖ Households by marital status, sex and age of the head of household
- ❖ Households having female children in the age-group 5-14 years by number of female children attending school
- ❖ Households by size and number of female members seeking/ available for work
- ❖ Households by number of disabled females and household size

HH-9 APPENDIX B

HOUSEHOLDS BY NUMBER OF FEMALE LITERATES AMONG FEMALE MEMBERS OF THE HOUSEHOLD AGE 7 YEARS AND ABOVE

Number of female literates in the household	Total households	Number of female members in household age 7 years and above				
		None	1	2	3	4+
1	2	3	4	5	6	7

India/State/District
Total/Rural/Urban

Total
None
1
2
3
4+

Note : This table excludes institutional households.

HH-9 SC APPENDIX B

HOUSEHOLDS BELONGING TO SCHEDULED CASTES BY NUMBER OF FEMALE LITERATES AMONG FEMALE MEMBERS OF THE HOUSEHOLD AGE 7 YEARS AND ABOVE

Number of female literates in the household	Total households	Number of female members in household age 7 years and above				
		None	1	2	3	4+
1	2	3	4	5	6	7

India/State/District
Total/Rural/Urban

Total
None
1
2
3
4+

Note : 1. This table excludes institutional households.
2. Data for districts will be available only in electronic format.
3. A household has been presumed to be a Scheduled Caste household when the head of the household belongs to a Scheduled Caste.

HH-9 ST APPENDIX B

HOUSEHOLDS BELONGING TO SCHEDULED TRIBES BY NUMBER OF FEMALE LITERATES AMONG FEMALE MEMBERS OF THE HOUSEHOLD AGE 7 YEARS AND ABOVE

Number of female literates in the household	Total households	Number of female members in household age 7 years and above				
		None	1	2	3	4+
1	2	3	4	5	6	7

India/State/District
Total/Rural/Urban

Total
None
1
2
3
4+

Note : 1. This table excludes institutional households.
2. Data for districts will be available only in electronic format.
3. A household has been presumed to be a Scheduled Tribe household when the head of the household belongs to a Scheduled Tribe.

HH-13 HOUSEHOLDS WITH AT LEAST ONE MEMBER AGE 15 YEARS AND ABOVE AND WITH OR WITHOUT EDUCATIONAL LEVEL MATRICULATION AND ABOVE BY HOUSEHOLD SIZE

Number of matriculates and above in the household	Total households	Household size with at least one member age 15 years and above				
		1	2	3-6	7-10	11+
1	2	3	4	5	6	7

India/ State/ District
Total/ Rural/ Urban

1. Households with no matriculate and above
2. Households with at least one matriculate and above
3. Households with at least one female matriculate and above
4. Households with at least one graduate and above
5. Households with at least one female graduate and above

Notes : This table excludes institutional households.

HH-6 HOUSEHOLDS BY MARITAL STATUS, SEX AND AGE OF THE HEAD OF HOUSEHOLD

Age of the head of household (in years)	Total households		Marital status of the head of household									
			Never married		Currently married		Widowed		Divorced/ Separated		Unspecified	
	Male head	Female head	Male head	Female head	Male head	Female head	Male head	Female head	Male head	Female head	Male head	Female head
1	2	3	4	5	6	7	8	9	10			11

India/State/District
Total/Rural/Urban

All ages
Less than 20
20-29
30-39
40-49
50-59
60-69
70-79
80+

Age not stated

-
- Notes:**
1. This table excludes institutional households.
 2. Data for districts will be available only in electronic format.

HH-14 APPENDIX

HOUSEHOLDS HAVING FEMALE CHILDREN IN THE AGE-GROUP 5-14 YEARS BY NUMBER OF FEMALE CHILDREN ATTENDING SCHOOL

Number of female children in age-group 5-14	Total number of households	Number of female children attending school				
		None	1	2	3	4+
1	2	3	4	5	6	7

India/ State/ District
Total/ Rural/ Urban

Total
None
1
2
3
4+

Note : This table excludes institutional households.

HH-16 APPENDIX

HOUSEHOLDS BY SIZE AND NUMBER OF FEMALE MEMBERS SEEKING/ AVAILABLE FOR WORK

Female members in the household seeking/ available for work	Total households	Number of female members in the household						
		1	2	3	4	5	6	7+
1	2	3	4	5	6	7	8	9

India/State/District
Total/Rural/Urban

Total
None
1
2
3
4+

Note : This table excludes institutional households.

HH-17 APPENDIX B*HOUSEHOLDS BY NUMBER OF DISABLED FEMALES AND HOUSEHOLD SIZE*

Number of disabled females in the household	Total Households	Household size			
		1	2	3	4+
1	2	3	4	5	6

India/State
Total/Rural/Urban

Total
None
One
More than one

Note : This table excludes institutional households.
