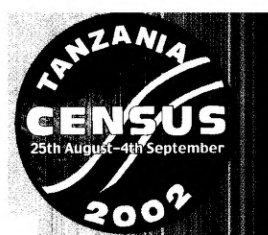


HA
4697
. A4
2002
v. 1



2002 Population and Housing Census

Volume

Methodology Report

**Central Census
National Bureau**

**President
Planning and**

**Dennis Salaam
September, 2003**

CHAPTER 6

CENSUS DOCUMENTS

6.1 Introduction

The population census is the main source of information on the number and important demographic, social and economic characteristics of the population in any country. The uses of census data are innumerable. They are used in formulating policies for national development such as manpower development, formulation of housing policies and programmes, planning for social welfare programmes, and formulation of policies for investment (United Nations, 1992 etc.).

In Tanzania, a population census is the major source of population data; hence there is always a need to be careful in determining the scope and content of the census. In the 2002 Population and Housing Census, as with all previous censuses, great care was taken to ensure that the census collected the highest priority information for use by planners, researchers, etc. It was in particular recognized that improving the timeliness of production of the results was a key performance indicator.

6.2 Census Content

One of the early decisions which was made, concerned the topics to be included in the 2002 population and housing census. The major factor, which was used to determine the choice of topics, was users' needs for accurate and comprehensive statistical information. A list of topics was drawn up in the Central Census Office by taking into consideration the resources available and the contents of the previous two censuses, which provided the basis for understanding the long-term trends, UN and SADC recommendations on the census basic topics.

A key meeting to set out the content of the census was held in April 2001¹. Participants at the meeting included data users from government departments, University of Dar es Salaam, research institutions, individuals and international organizations particularly the United Nations Agencies. The main purpose of this meeting was to agree on the topics to be included in the census questionnaire and to consider the tabulations to be adopted for the responses. The participants were free to indicate which topics they wished to be retained, dropped or added. A first draft list of standard tabulations was also circulated for discussion.

6.3 Questionnaire Design: Questions and Responses

The 2001 Pilot Census questionnaire provided the basis for preparing the final questionnaire which was used to collect and record information obtained from the respondents. In order to collect accurate information the questionnaire was well designed so that the questions did solicit the right answers, and did not offend the people providing the information or those recording and processing the information. It was thus crucial that they were well tested before the main census was taken. The census questionnaire is shown in Appendix 6.1.

¹ 2002 Population and Housing Census: A report on the data users meeting held at the Prime Minister's Office Conference Room on 20 April 2001 Central Census Office, Dar es Salaam, April 2001.

When the final list of topics was adopted, one of the early decisions to be made regarding the questions was whether to be simply a topic heading (for example Marital status) or a complete, scripted question (for example What is the person s marital status?) The latter, which is usually referred to as verbatim question was adopted and, subsequently, the questions were developed.

There were two major benefits from the use of verbatim questions. First, verbatim questions achieve more uniformity and thus reduce the bias that results from permitting enumerators to interpret and phrase questions according to their individual understanding and experiences. Secondly, they allow the enumerators to proceed smoothly with the interview and relieve them of the burden of formulating the questions themselves.

The initial question design was based upon the enumerators writing pre-assigned codes on an A3 page questionnaire. However, following a decision to test the use of the Optical Mark Recognition (OMR) technology in the Pilot Census the questionnaire was extensively revised to meet the standards required by the OMR technology.

6.3.1 Physical Structure of the Questionnaire for the Pilot Census

After the content and forms of questions were determined the immediate task on questionnaire design was to decide on the layout of the forms. Two types of OMR questionnaires were designed to be administered in the field². The short questionnaire covered one side of an A4 sheet (providing enough space for the responses for 10 people). The long questionnaire (with 37 questions) required a folded A3 sheet (in effect a booklet of 4 pages of A4 size) to accommodate the responses for 6 people. A relatively smooth flow of questions was arranged so that people dropped out when they reached the end of the questions applicable to them (for example, responses for children under 5 years were up to question 12; and those for men were up to question 20 - other than the head who was expected to answer household level questions on mortality and housing conditions). However, it was found necessary to fit the questions on housing conditions early in the questionnaire and this caused some problems in the field and during data processing. Some small problems were to be expected given the speed required to revise and print the questionnaires in time for the Pilot Census.

Another key issue was the choice of colour for the questionnaires. It was essential that all printing near a response area be done in a colour that was not picked up by the scanner. The advice from the supplier was to use questionnaires which were light red in colour and this was done.

6.3.2 Lessons Learnt from the 2001 Pilot Census

The 2001 Pilot Census was a major success and clearly showed that data capture using OMR technology was feasible. The short questionnaire worked very well. However, a number of issues emerged from the 2001 Pilot Census relating to the design of the questionnaires. These included:

- i) There was no provision for supervisor s identification mark after editing the questionnaires;
- ii) There was no provision for a mark to show that the household responses continued in more than one questionnaire;

² An English language version of the long questionnaire was also created to provide assistance if the interviewee was unable to respond to questions in Kiswahili. Even where this was used it was necessary to record the responses on a Kiswahili questionnaire.

- iii) The English version of the questionnaires was not available;
- iv) The way questions appeared on the questionnaires was not as was originally designed. Some were mere captions and font size was rather small;
- v) The format for questions number 1 - 8 on both the short and long questionnaires was not the same ;
- vi) The questions were not serially arranged (e.g. Household questions appeared on page 1 of the questionnaire) and thus caused considerable confusion to enumerators; and
- vii) The red colour was too bright and had a blurring effect. Light green and purple colours were recommended.

As a result of these findings the questionnaire was restructured and it was found that reducing the intensity of the colour produced good results. During this restructuring much use was made of e-mail to transfer successive drafts of the questionnaire between the printer, based in the UK and the Central Census Office in Dar es Salaam. Details of the 2001 pilot census questionnaire are to be found in Chapter 7.

6.4 Manuals

Population census is an enormous exercise and requires instruction manuals at various stages so as to ensure uniformity of the exercise. During the 2002 Population and Housing Census exercise there were Training Manuals, Instructions to Regional Census Coordinators (RCCs) and District Census Executive Officers (DCEOs), Instructions to Supervisors and Instructions to Enumerators.

The training manual was a document which was to be followed by all trainers at national, regional and lower levels. This document was prepared with great care; it includes relevant examples and solutions. It was prepared before the Pilot Census and it was tested during Training of Trainers to ensure that it was satisfactory for the main training.

The instructions to RCCs and DCEOs were documents which were prepared to guide them in their day-to-day activities before, during and after enumeration. It included instructions on how to complete control forms which were expected to be filled in by them.

The Instructions to Supervisors document, included all quality control checks in the field during enumeration as well as what role the supervisors had to play before, during and after the enumeration. It also explained procedures to be followed by the supervisor during enumeration and how to complete control forms which were expected to be completed by them.

The Instructions to Enumerators document was a reference book not only to the enumerators but also to supervisors and the rest in the field; it covered the background and purpose of the census, administrative requirements for enumerators, techniques for interviewing and detailed technical instructions on how to handle the questionnaires and other related forms.

6.4.1 Control Forms

All Census project documents are important; therefore their movements to and from the field had to be monitored. To do this, the use of control forms is highly recommended. These forms have a variety of functions, but they were used primarily in connection with record keeping, recording of movement of materials from one operation to another and progress reporting. They were designed

in such a way that all the documents which were to be used during enumeration were known as well as who was responsible. It was important to have quality control forms during enumeration to check if the supervisors followed quality control instructions. Also, use of control forms reminded officials about the type and quantity of materials which were required in the field and what had to be collected from the field. The list of the 2002 Population and Housing Census control forms is in appendix 6.2

6.5 Problems Faced and Recommendations for Future Censuses

6.5.1 Problems

There was delay in printing of census questionnaires because of the need to conform to a new Procurement Act which involved elaborate and lengthy procedures. This resulted in expensive delivery as the questionnaires had to be airlifted at extra cost.

Some of the control forms were not well understood by the trainers and later enumerators. As a result some were not filled in at all and those filled in, some were not filled in properly.

The late decision on procurement of OMR technology disrupted the workplan on production of other census documents, such as Instruction Manuals.

6.5.2 Recommendations

- i) In order to avoid delays, if any technology is tested and approved during the pilot exercise, arrangements for it to be taken as a solution for the main census should be agreed on early in the process; and
- ii) Adequate time should be allocated for training in the use of control forms.

CHAPTER 7

PILOT CENSUSES

7.1 Introduction

Preparations for any census take necessarily a long time and involve quite many distinct activities. These activities include conducting a pilot census, an exercise which aims at replicating at a smaller scale the actual conditions of undertaking the actual census. The testing of various aspects including logistics of the census plan prior to the enumeration is a crucial component of delivering high quality results.

This chapter describes the pilot censuses which were designed to test all aspects of the Population and Housing Census. Two such pilot censuses were carried out; the first was held in 1998 as a test for the census planned for 1999 and the second in 2001 as a test for the census taken in 2002. Both are described below, with more weight being given to the 2001 Pilot Census since that had much greater relevance to the procedures used in 2002.

7.2 The 1998 Pilot Census

7.2.1 Objectives

The main objectives of the pilot census were to:

- i) Test the suitability and adequacy of the census instruments such as enumeration area maps, questionnaires, control forms, instruction manuals, enumeration procedures, quality control procedures, editing specifications and editing;
- ii) Evaluate the quality of data on important items such as age, fertility, mortality, migration, economic activity, education and housing conditions;
- iii) Provide information on the average time required for enumerating a single household, which was an important input in estimating the staff and financial requirements for the main census;
- iv) Test the adequacy of field and central organization;
- v) Assess the data processing methods and quality control systems;
- vi) Serve as a practical training in census operations;
- vii) Assess the population's awareness of the census and their reaction to census questions;
- viii) Assess the census logistics such as storage facilities, transport and communication problems in the field; and
- ix) Assess the possibility of including agriculture and gender questions in the main census.

7.2.2 Coverage

The first pilot census was held in August 1998. Enumeration areas from five districts were covered. The districts were Bagamoyo, Kinondoni, Makete, Kahama and Chake Chake.

A total of 150 EAs were sampled. The purposive sampling method was employed to draw 30 EAs from each district. This type of sampling method was used to make sure that as many different categories of the population were represented in the sample.

7.2.3 Dispatch of Materials to and from Regions/districts

Finance and Supplies Unit of the Central Census Office dispatched materials from the Headquarters to the Regional/District Offices. Due to the small quantities of materials needed for the pilot census in each district, no special arrangement was needed as most of these materials could be and were sent to the regions using office vehicles which were taking trainers to regions/districts. The same method was used to bring back the materials to Dar es Salaam after enumeration. Dispatch of materials in the pilot census did not provide the actual picture (in terms of cost, logistics and tonnage) for the proper census when tons of materials had to be ferried to each region and back.

7.2.4 Recruitment and Training

Recruitment of enumerators and supervisors was done jointly by Regional Census Officers and Regional Statistical Officers with instructions and guidelines from the Central Census Office. Supervisors/enumerators were mainly recruited from among primary school teachers and a few from secondary school teachers.

There were two levels of training; level one was training of trainers, which was held in Morogoro; and level two training was held in the selected districts. Trainers at level one training came from the National Bureau of Statistics, Ministry of Agriculture, Ministry of Youth and Labour, Ministry of Education, President's Office, Planning and Privatisation and Building Research Unit.

Training involved both theory and practice. Supervisors and enumerators were taught how to identify the enumeration areas and how to handle effectively the interviews. They were also taught how to fill in different control forms and how to control the quality of information.

7.2.5 The Questionnaire

Only the detailed questionnaire was used in the pilot census. It had a total of 48 questions divided into 10 sections, namely:

- i) Identification;
- ii) Questions for all persons;
- iii) Questions on all persons aged five years and above;
- iv) Questions on all persons aged ten years and above;
- v) Questions on all women aged twelve to forty nine years;
- vi) Deaths;
- vii) Housing conditions;
- viii) Gender;
- ix) Agriculture; and
- x) Total Population.

7.2.6 Enumeration

The Pilot Census was carried out on the last Sunday of September 1998, exactly one year before the intended census date. Enumeration was based on *de facto* approach with some characteristics of *de jure* for specific questions. Persons were enumerated on the basis of where they spent the pilot census night. Enumerators spent about three days familiarizing themselves with their areas. This period was used to introduce themselves to local leaders, make appointments and to identify their enumeration areas. They also updated the list of the heads of households. Enumeration took a

maximum of seven days in most cases. However, in a few cases enumeration took more time than the intended period of seven days.

7.2.7 Data processing

Questionnaires and other pilot census questionnaires were received at the Central Census Office mid October 1998. Data capture was done manually as planned for the census itself.

7.2.8 Information, Education and Communication

The Central Census Office used several means to sensitize the public in general. Means used were:

- i) Meetings with Regional and District leaders involved in the pilot census;
- ii) Press conferences: A total of three conferences were held between the Commissioner of Population Census and the media personnel;
- iii) Advertisements through radio stations, newspapers and television stations. On the eve of the pilot census, The Minister of State, President's Office, Planning Commission delivered a speech that was aired by Radio Tanzania, Dar es Salaam and was televised by Independent Television Limited. Similarly, question and answer sessions were held with Radio Tumaini and Sauti ya Tanzania, Zanzibar. These were conducted by the Commissioner and Commissar of Population Census for Tanzania Mainland and Tanzania Zanzibar respectively;
- iv) Posters and leaflets were another kind of mobilization used. About 5,000 leaflets and 700 posters were printed and distributed in the piloted areas; and
- v) To complement the use of electronic media in Dar es Salaam the public address system was also used.

7.2.9 Analysis

A team of analysts produced a report titled The 1998 Pilot Census for the 1999 Population and Housing Census. Topics covered in this report were:

- i) Introduction;
- ii) Evaluation and Analysis of the Age Structure;
- iii) Disabled and Marital Status;
- iv) Migration;
- v) Literacy and Education;
- vi) Economic Characteristics;
- vii) Gender;
- viii) Fertility Levels and Patterns;
- ix) Mortality Levels and Patterns;
- x) Housing Characteristics;
- xi) Agriculture;
- xii) Data Processing; and
- xiii) Information, Education and Communication Activities.

The conclusion reached in the Pilot Census Report³ was that the planned census operations appeared to be feasible, although improvements would be desirable in some areas.

³ The 1998 Pilot Census for the 1999 Population and Housing Census Central Census Office, Dar es Salaam, March 2001

7.3 The 2001 Pilot Census

7.3.1 Introduction

The 1998 pilot census was meant to test the adequacy of preparations for the 1999 Population and Housing Census, which was postponed to 2002. With a postponement period of three years, there was a need for carrying out another pilot census in 2001. Most of the procedures and techniques used in 2001 pilot census were more or less similar to those used in 1998 with one major exception. The 2001 pilot census was used to assess the possibility of using the new data capture technology known as OMR (Optical Mark Recognition) in the main census in 2002.

7.3.2 Objectives

Objectives of the 2001 pilot census were similar to those of the 1998 pilot census with the major exception of evaluating the new data capture technology.

The Government had earlier decided that the Optical Mark Recognition (OMR) be tested as a method of data capture instead of the traditional manual data entry approach. The 2001 pilot census was used to test the new technology before the final decision could be made.

The DRS company of the UK was given the responsibility of designing and printing questionnaires. The company was also given the responsibility of providing scanners and technical support.

7.3.3 Coverage

Unlike the 1998 pilot census, the 2001 pilot census was conducted in all districts whereby three enumeration areas were selected randomly in each district. A total of 369 EAs were covered of which 123 were rural and the remaining 246 were from urban areas. The high proportion in urban areas reflected an expectation that they were more difficult to enumerate and thus required greater investigation.

7.3.4 Questionnaires

Two questionnaire designs were employed, just as intended for the main census. The key feature was that the designs had been created to utilise the OMR technology. The short questionnaire comprising one A4 sheet, printed on one side only had sufficient space to record details of 10 persons. The long questionnaire with 37 questions was an A3 sheet printed on both sides and folded to give 4 A4 pages.

The detailed questionnaire had seven sections with a total of 37 questions. Topics covered were like those in the 1998 pilot census minus the agriculture and gender questions. The topics were:

- i) Identification;
- ii) Questions for all persons;
- iii) Questions for all persons aged five years and above;
- iv) Questions for all women aged 12 years and above (fertility);
- v) Questions for number of deaths in the households;
- vi) Questions for housing conditions;
- vii) Summary for total persons in the household.

The short questionnaire had a total of eight questions only. The topics covered were:

- i) Identification;

- ii) Questions for all persons;
 - a) Serial number;
 - b) Name of the household member;
 - c) Age;
 - d) Relationship to head of household;
 - e) Sex;
 - f) Marital status;
 - g) Disability;
 - h) Nationality; and
 - i) Summary of total persons in the household.

7.3.5 Recruitment and Training

As planned for the main census, and following established practice, enumerators were primary school teachers who lived in, or were very near to the selected enumeration areas. The role of the supervisor was taken by staff members from the Central Census Office assisted by staff members from a range of other Ministries (some of whom had administered the training program for the Pilot). The Central Census Office issued guidelines for recruiting enumerators. The guidelines instructed Regional Census Coordinators, Regional Statistical Officers and District Census Executive Officers to recruit capable personnel to do the job. Like during the previous pilot census, Grade A Primary School teachers were given preference.

Training was conducted at two levels. Level one was Training of Trainers, which was held in Morogoro. Level two training involved Regional Census Coordinators, Regional Statistical Officers, District Census Executive Officers, supervisors and enumerators and was held in the respective zones.

Training involved both theory and practice. The training had a special session on how to shade and handle the OMR questionnaires.

7.3.6 Enumeration

This pilot census was conducted in August 2001, precisely one year before the 2002 census. This timing coincided with plans for the actual census and thus tested issues such as whether people were at their usual residence, as expected, at this time of the year.

7.3.7 Data Processing

To test the OMR data capture method, technical assistance, including borrowing of some personnel from the National Examinations Council of Tanzania who had experience in using the OMR technology in marking examinations was sought. They also provided the required hardware. The scanning operation was undertaken at their premises at Mikocheni while data cleaning and other analytical work was done at the CCO.

The overall results showed that good quality data had been achieved during enumeration with only a low proportion of questionnaires requiring remedial action prior to data capture. While some enumerators performed better than others there was no general evidence of significant difficulty in carrying out the shading requirements for responses as required by the OMR technology.

7.3.8 Analysis

In terms of data analysis the results of the Pilot Census showed that all the core indicators gave results which were comparable with those obtained from the results of earlier censuses and relevant surveys (adjusted for known changes in policies and behaviour). Some problems known to exist in African census data (notably age heaping and digital preference) were also observed — with levels similar to those experienced in previous censuses.

Recommendations recorded in the report of the pilot census mainly relate to improvements in the administration of the census.

7.3.9 Other Recommendations Based on the 2001 Pilot Census were

- i) The training period of 10 days for enumerators administering the long questionnaire was not enough;
- ii) Enumeration period of 10 days was not enough in some places;
- iii) Enumerators and supervisors should be paid bus fare during enumeration period. They should also be paid for the days spent on familiarization. Lunch allowance was recommended for this;
- iv) EA maps and their descriptions should be improved;
- v) One rural Supervisory Area (SA) should have a maximum of three EAs while an urban SA should have a maximum of five EAs;
- vi) The allowances to enumerators and supervisors should be increased;
- vii) Publicity should be intensified in all areas; and
- viii) The color and layout of the questionnaire should be changed.

7.3.10 Conclusion

The organization and conduct of the pilot censuses enabled the Central Census Office Staff (the technical staff in particular) to identify, and resolve deficiencies in readiness for the 2002 Population and Housing Census.

These pilot censuses provided important information on how various aspects of the Census Project should be conducted and processed. It is considered that experience gained from the pilot censuses contributed greatly to the successful conduct of the 2002 Population and Housing Census.

A key point was that the 2001 Pilot Census showed that it was feasible to use OMR technology for data capture in the main census. As a result of this experience major efforts were directed towards developing techniques and procedures using the OMR technology.

7.4 Problems Faced and Recommendations for Future Censuses

7.4.1 Problems

- i) Remuneration for enumerators and supervisors during the pilot census were different from those paid during the main census. This was a major contradiction and defeated the whole purpose of testing the census mechanism. This also had a budget implication;
- ii) Some of the recommendations from the pilot census were not adopted in the main census. For example, it was recommended after the pilot census that trainers should arrive at the training centers one day before the training starts, but this was not done during the main census;

- iii) Adopting new technology after the pilot census, for example use of aerial photos in identifying enumeration areas was adopted after the pilot census had been conducted and as such they were not included in the training manuals to enumerators and supervisors;
- iv) A pilot census is not a good measure for testing logistics in full because the pilot census involves only a few EAs compared to the main census;
- v) The pilot census did not test the financial transactions. Mode of payment to enumerators and supervisors during the pilot census was quite different from the main census; and
- vi) Having no PES for the pilot census meant that some key instructions were omitted from the main census.

7.4.2 Recommendations

- i) Pilot census should really be a measure of the main census. Therefore recommendations from the pilot census should be taken on board;
- ii) Careful planning and budgeting is needed for logistics during the main census. Observations from the pilot census on this approach should only be taken as a guide for the main census;
- iii) No new technology or procedure should be used during the main census if it was not tested in the pilot census; and
- iv) In future censuses, a PES should be carried out for the pilot census.

CHAPTER 8

SAMPLE DESIGN

8.1 Introduction

During the planning of the 2002 Population and Housing Census it was decided to collect information on education, economic activity, disability, migration, fertility, mortality, housing conditions and ownership of selected assets through a sample survey within the framework of the census. This followed the strategy employed in previous post independence censuses of Tanzania.

Generally, in the context of a developing country like Tanzania¹, there are several arguments in favour of a sample approach for collection of detailed statistical data of persons, households and dwellings as follows:

- i) Needs for information related to statistics at several levels, e.g., national, regional, district and smaller geographic levels. At the lower geographic levels there is only limited capacity for analysing and using the data;
- ii) The expected presence of measurement problems which should be kept under control, as far as possible, favour a limited scope for data being collected and processed;
- iii) Costs of collecting and processing the data are reduced by a sampling strategy; and
- iv) The timeliness of the results is improved through handling less data.

The Central Census Committee decided to use two questionnaires in the 2002 Population and Housing Census. These were:

- i) A short questionnaire with 8 questions for all households; and
- iii) A long questionnaire with a further 29 questions for a sample of households.

As for the 1988 census, results from the 2002 Population and Housing Census are presented at district level, broken into rural and urban domains. In the 1978 and earlier censuses, a smaller sample size was used and the lowest level of geographic breakdown for which reliable data was possible was for regions.

Since most of the subsequent parts of this chapter draw heavily on the report on sample design² it is suggested that the readers interested in further technical details may refer to the document.

It is important to note that the Enumeration Areas (EAs) which covered institutions like hospitals, guest houses, schools, fishing and mining camps etc. were excluded from the sample design for 2002 Population and Housing Census. The same practice was followed in earlier censuses.

¹ It should be noted that many of the more developed countries (including the USA, UK and Canada) also employ sampling as part of their census taking strategy.

² TANSTAT 27th May, 1988, Sampling in Tanzania, A mission report, by R. Petterson, 9th April — 5th May, 1988.

8.2 Sample Design for the Rural Areas

The sample of households for the long questionnaire was selected in clusters formed by all households within an EA (Ultimate Area Unit). The target size of rural EAs in the 2002 Population and Housing Census was about 800 people. This size was still considered on the high side for Ultimate Area Units at least for socio-economic variables. However, the advantage of using an existing area as an Ultimate Area Unit is so great that it was felt that this would offset the loss in precision. Since rural parts of the district as domain of study varied in size, various sample sizes were used. These are summarised in the following table.

Table 8.1: Selection Criteria for Sample Enumeration Areas

Total EAs in the domain of the study	Number of EAs selected in each domain of the study
< 30	All EAs
30 — 199	30
200 — 399	40
400+	50

The assumption made was that each EA had equal probability of being selected. The method adopted for drawing a sample was Systematic Simple Random Sampling (SSRS).

8.3 Sample Design for Urban Areas

The size of the urban EA in the 2002 Population and Housing Census was about 400 people, as it was for the 1978 and 1988 Population and Housing Censuses. In the previous censuses investigations had concluded that there were benefits from a design with a larger sample per district in the urban areas than in the rural areas.

It was concluded that the sample allocation used in 1988 was still basically appropriate. For the cities of Mwanza and Dar es Salaam and the municipality of Zanzibar, it was decided to take 70 EAs in the sample. For urban areas in other districts a sample of 50 EAs per district was taken. Where the number of EAs in the urban domain was less than 50, all EAs were covered by the long questionnaire.

As with the rural domain each EA had equal probability of being selected and SSRS was used in the selection.

8.4 Selected and Utilized Enumeration Areas

Overall, about 99 percent of all the selected EAs were utilized during the post enumeration survey. These were the areas in which the long questionnaire was used during the main census. Urban areas were slightly better on utilization than rural areas. Table 8.2 below gives the breakdown of utilization of selected enumeration areas by rural/urban localities.

Table 8.2: Selected and Utilized Enumeration Areas

Area	Mainland			Zanzibar			Tanzania		
	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total
Selected Eas	4,474	4,988	9,462	270	260	530	4,744	5,248	9,992
Utilized Eas	4,422	4,955	9,377	270	260	530	4,692	5,215	9,907
% Covered	98.84	99.34	99.10	100.00	100.00	100.00	98.90	99.37	99.15

Appendix 8.1 shows the distribution of total EAs and number of selected EAs in which the long questionnaire was used in respective regions; Tanzania Mainland, Tanzania Zanzibar and Tanzania as a whole.

8.5 Problems Faced and Recommendations for Future Censuses

8.5.1 Problems

Overall the sample design used for the selection of the sample and **estimation of different characteristics**, was appropriate. Most of the estimated characteristics that were obtained by using the design were of reasonable quality. However, some problems were encountered as follows:

- i) At a later stage in the sample selection process it was necessary to revise the sampling frame to accommodate the split of the previous Arusha Region into the two regions of Arusha and Manyara; The Cartographic Unit undertook the allocation of EAs to the new regions once the boundaries were declared as the official boundaries;
- ii) In some cases EAs were incorrectly included in the urban or rural domains. This led to some urban EAs being selected as rural or vice versa. Since this was only discovered after enumeration the problem had to be tolerated and appropriate adjustments made to the weights. The extent of errors was not significant and mostly affected the rural sample for Kinondoni Municipality in Dar es Salaam Region.
- iii) Each DCEO was furnished with the list of EAs selected to be covered with the long questionnaire. As was the case with the 1988 Census, not all the districts utilized the sample listing properly. In some cases EAs not in the sample used the long questionnaire, while in others the selected EAs had been enumerated with the short questionnaire. These errors had limited adverse effect on the quality of the results; since in most cases the difference was only plus or minus 1.
- iv) There were incidences where names of villages were not clear.

The impact of the problems described above was more for the rural areas than for the urban areas.

8.5.2. Recommendations

To avoid the shortcomings which were observed in the 2002 Population and Housing Census it is recommended that:

- i) For future censuses efforts should be made to find ways of accommodating changes to administrative boundaries without having to reselect the sample;
- ii) To avoid the errors which were evident in allocation of EAs to rural and urban domains the Cartographic Unit should, as part of its GIS work, establish a data base in which the nature of an EA is an attribute of that entity, rather than just the EA number;
- iii) Special category population i.e. fishing camps, mining camps, nomadic populations, etc, should be included in the sample design; and
- iv) Use of one reasonably designed questionnaire for the entire population will reduce problems, which are related to long and short questionnaires.

CHAPTER 9

IEC/ADVOCACY FOR THE 2002 POPULATION AND HOUSING CENSUS

9.1 Introduction

Information, Education and Communication (IEC) activities were amongst the major strategic components of the 2002 Population and Housing Census. Since 1998, when the planned census failed to take off due to financial constraints, IEC was a regular feature of the planning process leading to the 2002 census. In the period before the launch of the publicity campaign on June 2, 2002, IEC activities focused more on informing the government and development partners and hence the use of more interactive communication channels. A regular monthly publication, coupled with regular meetings by the various interest/stakeholder groups sufficed to sustain their continued interest in the census project.

First amongst IEC strategies was to identify the various stakeholders that would be addressed. These were found to be:

- i) Individuals in urban and rural areas;
- ii) Census implementers all over the country;
- iii) The various supervising committees at the national, regional, district, ward and village levels;
- iv) Media practitioners and other mass communicators; and
- v) The donor community, which provided financial support for the census to augment government funding.

The IEC activities, messages and choice of channels of communication were tailored to suit the specific needs of these target groups. The publicity aspect was to make the public and stakeholders fully aware of the importance of the census and the data to be generated from the exercise. The people were to be educated on the content of the questionnaires, and the role of the various authorities during the enumeration exercise. The advocacy aspect was to maintain a continuous support from the government, donors and social partners, so that the requisite resources could be assured. The ultimate goal of all this was to facilitate the collection of complete and accurate information and data on socio-economic characteristics of the population during the 2002 census.

9.2 Objectives

- i) To sensitize and mobilize Tanzanians so that they would support, co-operate and participate in the 2002 Population and Housing Census;
- ii) To secure continuous support and commitment by the government at all levels, development partners and the private sector; and
- iii) To promote acceptance and extensive use of the census results by various users.

9.3 Outputs, Strategies and Activities

The IEC was designed to achieve specific outputs through linking of various strategies and activities.

Output 1:

Strengthened capacity to undertake effective publicity in support of the 2002 census

- i) The Central Census Office secured the attachment of professional mass communications officers from the Tanzania Information Services and the Zanzibar Department of Information who were later on assisted by a UNFPA sponsored consultant from Kenya;
- ii) Census mobilization committees were formed in all the regions and districts;
- iii) Members of Parliament and the House of Representatives attended special workshops whereby they received advocacy techniques, materials and orientation for the census mass mobilization;
- iv) A directive was issued assigning leaders at all administrative levels the duty of ensuring that all people were fully mobilized for the exercise; regional and district leaders were to be held accountable for any flaws; and
- v) Advocacy booklet containing all-important information and messages was produced and distributed to all mobilizers as a guiding document.

Output 2:

Enhanced partnership with the media

- i) A National Media Advisory Committee was formed to advise the IEC Secretariat on the campaign messages and publicity strategies. The committee comprised representatives from the public and private media;
- ii) Media owners and managers were briefed at a luncheon, and reporters had an opportunity of discussing the census at workshops held in Dar es Salaam and Zanzibar, where they had a chance to discuss their role in the coverage of the census. The aim was to sensitize them and win their support as important partners in publicity and mobilization;
- iii) The media were invited and facilitated to participate in all census functions and events;
- iv) Census up-dates in the form of press releases and press briefings on the progress of the exercise, were done throughout the census period; and
- v) During the enumeration period, press briefings were held almost every day.

Output 3:

Continuous commitment and interest by the government and development agencies

- i) Regular briefing meetings on the progress of the census were held by the Central Census Office with the Central Census Committee, Census Advisory Committee, Collaborators Forum, the Census Technical Committee, the Regional Census Coordinators and District/City/Municipal/Town Council Census Executive Officers;
- ii) A monthly newsletter on the progress of the census exercise was published and distributed to all relevant authorities; and
- iii) Top ranking government officials were invited to officiate at census functions.

Output 4:

Increased capacity for sensitization by other support groups

- i) Regional and District mobilization committees were facilitated;
- ii) Poetry and theatre groups were sensitized through workshops and given the guiding document with relevant messages to help them with their creations and compositions;
- iii) A cartoonist was sensitized and contracted to draw educative cartoons that were distributed to mainstream newspapers; and
- iv) Pop music, choir and traditional *ngoma* groups were commissioned to produce songs on the census and copies were distributed to broadcasting stations.

Output 5:

Enhanced public awareness and popularity of the exercise

- i) The public sensitization campaign was launched in a bang style on 2nd June, 2002 during the start of the Uhuru Torch Race at Sumbawanga, Rukwa Region. The event was officiated by the Vice President and witnessed by several Regional Commissioners and government officials from the neighboring districts of Zambia and Malawi. The national radio broadcaster covered it live;
- ii) Over 2 million leaflets containing key messages were distributed throughout the country;
- iii) The Uhuru Torch Race was used to disseminate census information through a special publication called Uhuru Torch Message ;
- iv) The private sector was requested to display the census logo on their products, and several of them did;
- v) Radio and television phone-in programmes were arranged, whereby the Commissioner/ Commissar of Population Census answered listeners questions and explained issues on the census;
- vi) Slogans were developed and broadcast on key radio and television stations;
- vii) A series of radio programmes were broadcast over the national radio and a few collaborating private stations;
- viii) Special television programmes were aired on major TV stations;
- ix) Over one million stickers were distributed all over the country;
- x) Information articles were sponsored in 8 national newspapers;
- xi) A demo video was produced for showing in buses and travel vessels;
- xii) Sponsored commemorative postage stamps were issued with the census logo;
- xiii) *Khangas*, *vitenges*, T- shirts, caps, sun visors, waving flags with the census logo and important messages were also distributed widely; and
- xiv) The President of the United Republic of Tanzania gave a speech on the eve of the census date, as part of his monthly speeches in which he emphasized the importance of the census to the nation in order to rally public support for the census. The speech of H.E. the President of the United Republic of Tanzania when launching the census on the census night is attached as Appendix 9.1.

Output 6:

Increased dissemination and use of the 2002 census data

- i) The 2002 Population and Housing Census General Report was launched at a press conference and distributed free of charge to all key institutions, development and social partners;
- ii) Production of fact documents on the themes relevant to the interests of various stakeholders and planning requirements of the government and the private sector; e.g. district profiles, etc.;
- iii) Re-packaging in appropriate formats and publishing for distribution in friendly forms of analyzed data by using simple standard tables on hard copy, complex tables on CD ROMS, and floppy discs etc.;
- iv) Holding of end users meeting with national government, regional and district administrators; academics, schools, private sector, political parties, NGOs and international organizations;
- v) Development of a marketing strategy for census data that would take into account the relevance to various users, accuracy, cost etc;
- vi) Posting the data on the internet. It is on the national website www.tanzania.go.tz/census; and
- vii) Introduction of super cross programme for use at regional and district levels,

For details on the logframe matrix of IEC activities see Appendix 9.2.

9.4 Problems Faced and Recommendations for Future Censuses

- i) IEC is very important for a country to carry out a successful census. For 2002 Population and Housing Census staff members from the Information Department, Prime Minister's Office and the Zanzibar Department of Information were seconded to the CCO to carry out Population and Housing Census IEC activities. However, not all the seconded staff were competent enough. It is therefore recommended that one of the staff members at the recommended permanent CCO should be an expert on IEC. This person must have background knowledge in statistics;
- ii) Choice of media was also problematic because of the multiplicity of newspapers, radio and television stations. However, not many of these did reach all corners of the country. Since most media outlets are expensive, in order to get the maximum coverage at a reasonable cost, a proper trade off must be made between coverage and cost. However, lower costs should not be achieved at the expense of coverage;
- iii) The 2002 Population and Housing Census faced problems in getting the proper interpretations of the census exercise in particular because of the environment of multiparty politics. In some cases census was viewed as an activity of the ruling party. This posed a problem of possible boycott by members of other parties. It is therefore recommended that census activities should be politically neutral in order to get everybody fully involved. Thorough IEC work is very important in this area;
- iv) The 2002 Population and Housing Census faced problems of people refusing to be enumerated on the basis of religion. They were arguing that their religion forbids them to be counted. In such a case it is important to invoke the Presidential Order issued under the Statistics Act, 2002 which requires everybody to be enumerated. It is the role of the IEC expert to explain this order to the public well in advance in order to achieve maximum cooperation. For a sample of the 2002 Presidential Order see Appendix 9.3.

- v) The 2002 Population and Housing Census faced a problem of people refusing to be counted unless they were given meat, and local brew. In such a case it is recommended that the government must be prepared to comply with their reasonable demands, in order to win their cooperation; and
- vi) The 2002 Population and Housing Census benefited substantially from total commitment to the census by national, regional and district government authorities. The census organization chart included regional and district census committees, which were serviced by RCCs and DCEOs respectively. Moreover, during the meeting which was chaired by the President of United Republic of Tanzania (URT) and attended by regional and district authorities just one week before the census, the president directed them to give top priority to the census exercise. This directive was a big boost to IEC work. It is recommend that future censuses should involve fully national, regional and district authorities in the census operations.

CHAPTER 10

ENUMERATION PROCEDURES

10.1 Introduction

Enumeration in a Population Census in simple terms, means counting the people of a particular area with recognized geographical boundaries. To be more scientific, enumeration is a procedural activity, whereby the information about the people, needed for planning, research, etc. are collected from primary sources.

During the planning stage, enumeration activities were given special attention in order to make thorough preparations before enumeration. The importance of the administrative machinery during enumeration cannot be over-emphasized. The complexity and delicate nature of the work is evidenced by the enormous manpower and materials involved.

10.2 Unit of Enumeration

The basic unit of enumeration in 2002 census was the person. However, for purposes of social and economic analysis it was desirable to associate each person with a human group of which he/she is part (household) and with specific living quarters.

10.3 Enumeration Approach

Population may be enumerated in two different ways. One is **de Jure** census (a person is enumerated at the place of usual residence) and the other is **de facto** census (a person is enumerated where he/she spent the census night) approach.

The **de facto** approach was used for the 2002 census due to the following reasons;

- i) For continuity and comparison reasons; since all post-independence censuses had been conducted by using the **de facto** approach.
- ii) Comparatively, this approach costs less than the **de jure** approach.
- iii) Technically, there were fears that if strict care was not taken, the **de jure** approach would have resulted into having more omissions compared with the **de facto**.

The reference night for the 2002 census was the night of 24/25 August 2002. This date relates to the last Sunday of August 2002. This means that a household was enumerated as it was on that reference date.

10.4 Special Enumeration Arrangements

The geographical work had resulted in a sub-division of the country in enumeration areas; called EAs. An EA in urban areas in Tanzania Mainland has an estimated population of about 400 persons and in rural areas about 800 persons. In Zanzibar (Unguja and Pemba) both rural and urban EAs have an estimated population of about 400 persons.

In addition to this, there were special population categories, which by definition did not belong to any demarcated EA, or for other reasons required special arrangements to be enumerated. These special population categories include ship s crew, homeless, passengers on trains, buses, aero planes etc, unstable population in institutions and migratory populations. Special Enumeration procedures were required to enumerate these populations, since they did not fall under normal EAs. These categories of population were enumerated using the short questionnaire and male enumerators were used. District Census Executive Officers (DCEOs) local knowledge was required to make these special arrangements.

10.5 Enumeration of Private Households

For the purpose of the 2002 population and housing census a private household was a group of persons who lived together and shared living expenses. Usually these were husband, wife and children. Other relatives, boarders, visitors and servants were included as members of the household, if they were present in the household on the census night.

If one person lived and ate by himself/herself, then he/she was a one-person household even if he/she stayed in the same house with other people (these cases were more prevalent in the urban areas). Household members staying in more than one house were enumerated as one household if they ate together.

During enumeration, persons were enumerated where they spent the census night. Two types of questionnaires were used. A long questionnaire was used to enumerate private households in sampled EAs, while the short questionnaire was used to cover other private households in non-sampled EAs.

10.5.1 Enumeration of Special Populations

10.5.1.1. Ship s Crew

The DCEOs in harbour/port towns contacted in advance harbour/port authorities, and explained about census requirements regarding visiting ships. Arrangements were made for enumerators to start the enumeration in the early morning of the 25th of August, 2002. Crews on all ships docked in the harbour port on the **census night** were enumerated. Ships docked on 25th August 2002 were not included.

10.5.1.2. Collective or Institutional Households

A collective household comprised groups of persons in camps, boarding schools, hospitals, hotels, prisons, etc. Again, only the persons who spent the census night in such places were enumerated.

10.5.1.3 The Homeless

Some people in urban areas are outdoor sleepers with no fixed dwellings. They are usually found in and around markets, at railway stations, or under trees and the roadside. DCEOs in advance noted where such homeless groups could be found. Male enumerators, working in pairs and accompanied by plain clothes Police Officers (who were required to stay at a distance during enumeration so as to get maximum co-operation from respondents) covered these people. Enumerators were sent out very **early on the morning on 25th of August 2002.**

10.5.1.4 Passengers on Trains, Buses, Aeroplanes, etc

Some people embarked on trains, buses, etc, on 24th of August, 2002, and traveled on the **census night**. DCEOs in advance, contacted authorities concerned to find out transport timetables, and agreed on necessary arrangements for enumerating passengers. Enumerators were placed at the main stations from the morning on 25th of August, for at least 24 hours. They enumerated people who embarked or disembarked during this period. In order to avoid double counting or omission, individuals enumerated in such places were given a **Census Card**, showing that he/she has been enumerated (*Nimehesabiwa*). The short questionnaire was used to enumerate this category.

10.5.1.5 Migratory Population

This category comprises people who have no permanent living residence. They usually move from one living place to another. The reason for this could be that they are looking for new fishing areas, grazing grounds, etc. Before enumeration, the DCEOs contacted their local leaders so as to make special arrangements on how to locate and enumerate such a population. All persons belonging to this category were enumerated where they spent the census night.

10.5.1.6 Scattered Population

DCEOs in advance got information about scattered populations in remote areas in their districts. All persons belonging to this group were enumerated where they spent the census night. Again, male enumerators were used to enumerate this group.

10.6 Eligibility Criteria

The following were the eligibility criteria during the enumeration:

- i) List of all usual members of the household. For the purpose of the population census, a household was a group of persons who usually live and eat together (share their living expenses). Usually this will be the husband, wife and children. Other relatives, visitors and servants were included as members of the household, if they were present in the household on the census night;
- ii) Enumerators were asked to enumerate all persons in households in their respective areas. These included Household members who were out in the fields, fishing, attending funerals, or away on a night shift, etc. were also enumerated even though they were not present when the enumerator visited the household;
- iii) Visitors and servants living in the same premises were enumerated; and
- iv) Persons who spent the census night in hotels, camps, prisons, hospitals (sick people) etc. were enumerated in these places and were excluded from their normal households.

10.6.1 Other persons who were included in the enumeration were

- i) Those who slept in a household but were absent when the enumerator arrived, e.g. those who left their households in the early morning to tend their gardens, farms, etc;
- ii) All visitors who spent the census night in a household;
- iii) Those who did not spend the census night in the household because they were on night shifts, attending funerals, etc (nurses, watchmen, etc);
- iv) Those who spent the census night in Tanzania but crossed the border into neighbouring countries, the next morning for various reasons such as business, work, etc;
- v) Those who were not citizens, but spent the census night in Tanzania;

- vi) Those who died after census night;
- vii) Those who had no permanent residence but usually do shift from one place to another provided they were in the country during the census night; and
- viii) Tanzanian diplomats who happened to have spent the census night in the country.

10.6.2 Persons who were not included in the Enumeration were

- i) Foreign diplomats and their families, except Tanzanians who worked or lived with them as relatives, visitors, servants etc.;
- ii) Foreign military attaches;
- iii) Children who were born after the census night;
- iv) Those who died before census night; and
- v) Those who were residents in neighboring countries or other countries but crossed into the country in the morning of census date or later.

10.7 Enumeration Staff and Their Functions

It is essential to bear in mind that every Supervisor and Enumerator enumerated himself/herself in the area where he/she conducted the enumeration.

During enumeration, the fieldwork group comprised enumerators, supervisors, DCEOs, RCCs and RSOs. Enumeration started on the night of 24/25 August 2002. However, before the census night, enumerators and supervisors spent three days updating lists of heads of households (where available), EA boundaries, Supervision Area maps, etc.

This job was done with the help of local leaders like Village/Area Executive Officers, Vitongoji chairpersons, etc. The role of enumerators was to enumerate all household members in their respective EAs.

The role of supervisors was to ensure that the enumeration was properly done by the enumerators as well as, deal with problems faced by the enumerators in the field.

As part of quality control, supervisors had the task of checking and correcting the questionnaires which were filled in by enumerators. During and after enumeration, the major part of the workload was borne by the **DCEOs**. The primary function of **RCCs/RSOs** was to co-ordinate all census activities within their respective regions. RCCs/RSOs, however, had other specific engagements as follows:

- i) Participated in the training at level II as trainees;
- ii) Ensured that arrangements for training at level III were properly done by the DECOs;
- iii) Ensured that the geographical materials prepared in their offices - particularly EA and SA maps - were correct, complete and ready, well in advance for the start of the training at all levels;
- iv) Co-ordinated the activities for counting special categories in the region to ensure that they were properly enumerated; and
- v) Kept the Regional Census Committees well informed about the census.

On the other hand, the DCEOs performed the following functions during enumeration;

- i) They were trained (together with their assistants), by the Central Census Office Staff on the duties of supervisors and enumerators, on their own duties, and on how to train supervisors and enumerators (training level II at every regional headquarters);
- ii) Recruited supervisors and enumerators and arranged training venues;
- iii) Received and checked materials required for the training and enumeration from the Central Census Office;
- iv) Prepared geographical materials for supervisors and ensured that the supervisors distributed the materials to enumerators;
- v) Distributed materials to supervisors and ensured that the supervisors distributed them to enumerators.
- vi) Made special arrangements for enumerating special population groups. (homeless, nomads, passengers on trains, buses etc.);
- vii) Checked the work of supervisors and enumerators;
- viii) Received materials from supervisors, when enumeration was completed;
- ix) Packed the questionnaires;
- x) Handed over the materials to the Regional Census Office;
- xi) The DCEOs with his/her assistant also worked on materials after the enumeration, check-ups, and compilation of population figures; and
- xii) DCEOs kept District Officials well informed about the census. They organized meetings of the District Census Committees.

10.8 Recruitment and Training for Enumeration

Proper planning with statistical figures, requires intensive and extensive training and well coordinated data collection from the village to national level. Population census is a single, extensive, and probably the most expensive statistical operation undertaken by any country. In order to accomplish this task to the required standard it was necessary to ensure that proper recruitment and training of personnel involved in the Census Project were given the attention they deserved.

Enumeration of the total population of the country required services of approximately 70,287 persons (about 55,285 Enumerators and 15,002 Supervisors). Experience from the previous censuses showed that, primary and secondary school teachers were used as junior and senior enumerators.

It was decided by the Central Census Committee that teachers should again be engaged as enumerators and supervisors during the enumeration phase for the following reasons:

- i) Teachers are normally accepted by the community in which they live; and
- ii) Teachers are employees of the government, thus the public expects them to be responsible enough while conducting census enumeration.

10.8.1 Training Organization

The purpose of the training was to equip all the census staff with requisite knowledge and skills to carry out the fieldwork during and after enumeration. In order to standardize the training and reduce the risk of every trainer having his/her own interpretation as how to carry out the enumeration, various instructions were used during the training.

The technical staff from the Central Census Office worked on the organization and conduct of the training. The training involved both theory sessions in the classrooms as well as practical field exercises, including mock interviews.

The training programme was based on the Enumerator s Instructions Manual. Other materials included Enumeration and Supervision Area maps, and Quality Control Forms.

There were three training levels namely; national, regional and divisional levels.

10.8.1.1 National Level

The Training of Trainers (ToT) was conducted by the Central Census Office Staff and involved about fifty-three people (two trainers per region). The training was held for 10 days. The trainees at this level served as trainers at the regional level.

10.8.1.2 Regional Level

Trainees at this level comprised Regional Census Coordinators, Regional Statistical Officers, District Census Executive Officers, District Education Officers, District Supplies and Logistics Officers, District Adult Education Officers, Educational Coordinators and teachers from secondary schools. In total, about 1370 trainees were trained at this level. The training was held for 12 days. Subsequently, the trainees served as trainers at division level.

10.8.1.3 Division Level

Enumerators and supervisors were trained at this level. It was decided to organize the training at the divisional level in order to avoid problems in terms of logistics and reduce costs.

Trainees at this level comprised Education Coordinators, Head teachers and Grade A primary school teachers (to work as supervisors and senior enumerators), and primary school teachers (to work as junior enumerators).

Supervisors and senior enumerators were trained for 10 days while junior enumerators were trained for 6 days.

10.9 Guidelines for Interviewing

Clear instructions and well defined procedures guided the enumerator in contacting households and conducting interviews. The following points were among those that were covered in the instruction manual:

- i) How to ask the questions and how to record answers;
- ii) How to plan the routing and sequence of visits to ensure that the enumeration will be efficient in terms of distance and that all units will be covered;
- iii) Whom to interview;
- iv) What to do when no one is at home;
- v) Under what conditions callbacks should be arranged;
- vi) How many call backs should be made to one address and when they should be made;
- vii) How to handle refusals; and
- viii) How to deal with vacant units.

10.10 Equipment for the Enumerator

The basic enumeration instrument used in the field was the census questionnaire. An enumerator, was supposed to be thoroughly familiar with the **census questionnaire** as well as the enumerators Instructions Manual. In addition, the following materials were also issued.

- i) 1 copy of the EA map;
- ii) 1 copy of the description of the EA;
- iii) 2 blank copies of the household listing forms (EA2) for up-dating the list (where available);
- iv) Oath of secrecy;
- v) Special bag;
- vi) HP pencils;
- vii) Special erasers;
- viii) 1 clip, 2 blue biro;
- ix) 1 Note book;
- x) 1 file cover;
- xi) 2 pieces of chalk; and
- xii) 1 Notebook; the notebook was an official census document. Enumerators were required to write their names on the front page together with the names and codes of their EAs, i.e. region, district, ward/shehia and village, street/kitongoji. The following four pages of the notebooks were used for the timetable. The rest of the notebook was used as follows:
 - a) In case he/she had to return to the household to enumerate absent members, he/she wrote down the name and address of the household and time which he/she was expected to go back; and
 - b) Wrote down all other relevant comments on the enumeration, e.g. questions for which he/she needed additional help from his/her supervisor.

10.11 Enumerator in his/her Enumeration Area

Each enumerator was assigned a single enumeration area. A supervisor who was in-charge of about five enumerators was responsible for checking and coordinating the work of all enumerators within his/her areas. Each person was guided by EA/SA map, boundary description and a list of heads of households if available. The local leader played a very crucial role at this stage.

An enumerator arrived in his/her EA two or three days before the enumeration started. These days were used to:

- i) Contact local officials — Ward/Shehia Executive Officer, Village Executive Officer, etc and brief them on the census and the necessary assistance needed from them;
- ii) Check the accuracy of the household list with local officials together with the boundaries of that EA ; and
- iii) Introduce themselves to their households and prepare a tentative timetable such as the one below:

Table 10.1: Enumerators Timetable

Area/Village Executive Officer	Address	Planned time for enumeration	Real time for enumeration
		(eg. 30/08/2003) 07:10 to 08:30am	

10.11.1 Role of the Enumerator

An enumerator was required to make every effort to obtain complete and accurate answers and to record them, correctly. The quality of his/her work determined the quality of the census. It was important that all enumerators working for the census follow carefully the procedures laid down by the Central Census Office. Supervisors were supposed to be in contact with enumerators all the time. More details are provided in the junior and senior enumerators Instruction Manuals.

10.11.2 Role of the Supervisor

In order to ensure the quality of the enumerator s work, the supervisor did the following during enumeration:

- i) Scrutinized in detail all the completed questionnaires, to ensure that each interview had been completed properly;
- ii) Sport-checked some of the interviews; and
- iii) Met with the enumerators each day to discuss their performance. More details are provided in the Supervisor s Instruction Manual.

10.11.3 Checking Completed Questionnaires

After an interview had been completed, the enumerator reviewed the questionnaire he/she had just completed. This meant going over again the entire interview by reading carefully through all relevant questions and answers.

While checking, he/she might correct spelling mistakes or clarify on the answers given. This check was to be done before leaving the respondent. It was most important to check if information for identification of the enumeration area had been provided.

10.12. Problems Faced and Recommendation for Future Censuses

10.12.1 Problems Experienced

- i) **Enumeration Period:** Duration of the enumeration was too long, this caused memory lapse among respondents.
- ii) **Economic Activity Questions:** These questions were not clear to some enumerators;
- iii) **Disability Questions:** Definition of disability was not well understood by some of the enumerators.
- iv) **Confidentiality:** In some cases, confidentiality (i.e. local leaders not to be around during interviews) was not followed as instructed in the manuals.
- v) **Calendars of Events:** In some districts, calendars of local events were not available.

10.12.2 Recruitment and training for Enumerators and Supervisors

Main problems faced in this area include:

- i) Some enumerators were not physically fit;
- ii) Lack of enough qualified enumerators and supervisors in some areas; and
- iii) The method used in selecting senior enumerators and supervisors caused frustration among teachers with higher ranks who were not selected as supervisors.

10.12.3 Training

- i) There was over-crowding and poor facilities ie. poor accommodation, no water, electricity, etc at levels II and III of training;

- ii) In some areas, time allocated for practical exercises was not enough;
- iii) Inadequate Instruction Manuals particularly at levels II and III;
- iv) Discontinuation of some trainees due to religious beliefs and other social problems.
- v) Insufficient number of trainers in some centers;
- vi) Reserve senior enumerators were too few in some areas;
- vii) Almost all RCCs/RSOs and DCEOs did not participate fully in the training due to administrative work; and
- viii) It was very difficult to identify different special groups in the final output because of lack of unique EA codes.

10.12.4 Enumeration of Special Population Groups (i.e homeless, nomadic etc.)

- i) There was no budget provision for police escort;
- ii) Special forms for guest houses; -some forms were not filled in properly;
- iii) Lack of knowledge by guest house owners or attendants in filling in the forms
- iv) Editing; - Supervisors did not do edit as required;
- v) Some enumerators did not follow the enumeration procedures; and
- vi) In the final output it has been difficult to identify the different groups of populations because of lack of identification codes.

10.12.5 Recommendations

- i) To avoid memory lapse, enumeration period should be reduced from the current 10 days;
- ii) For economic activity questions there is a need of revisiting them and if possible this information should be collected in the labour force survey otherwise the questions should be reduced to two questions that is current and usual economic activity;
- iii) On disability, these questions should be revisited in close collaboration with NGOs dealing with disability i.e. CCBRT, etc,. Basic questions should be included and detailed information should be collected separately (carrying out a survey of its own);
- iv) Concerning confidentiality, IEC should pay adequate attention to this issue;
- v) About the calendar of events: the permanent Central Census Office should make a follow up on updating the calendar of events in the respective regions and districts;
- vi) Recruitment and training of enumerators and supervisors; instructions or procedures should be adhered to;
- vii) Maximum number of trainees per class should be 45 persons;
- viii) Training at level III should be conducted at the District Headquarters when a comfortable and cost effective training venue is not available at division level;
- ix) Time allocated for practical exercises should be observed;
- x) Instruction manuals and other materials should be delivered at the training venue well in advance;
- xi) There should be at least two trainers per class;
- xii) All reserve enumerators should be trained as senior enumerators;
- xiii) RCCs/RSOs and DCEOs should attend training and administrative work assigned to them should be planned in advance;
- xiv) There should be budget allocation for police escort;

- xv) Enumerators should educate Guest House owners/attendants on the use of forms during the familiarization time;
- xvi) Delineation of EAs should consider geographical distribution of population (i.e. scattered population);
- xvii) Delineation of EAs should be done close to the census date;
- xviii) Structure counting method of delineating EAs should be discouraged;
- xix) Population estimates should range between 300-500 in both urban and rural EAs;
- xx) The role of the supervisors should be revisited during preparations for the next census;
- xxi) Trainers from CCO, RCCs/RSOs and DCEOs should make sure that enumerators strictly follow the enumeration procedures;
- xxii) Definitions of head of household as well as household itself should be revisited; and
- xxiii) Assign a code for each special population group (i.e. nomadic population, boarding schools etc.).

CHAPTER 11

DATA PROCESSING

11.1 Introduction

Due to space constraint at the National Bureau of Statistics and the Central Census Office, data processing of the 2002 Population and Housing Census was carried out at the National Council for Technical Education (NACTE) compound. The Government facilitated the acquisition of the site and renovation of two warehouses. It was important to construct two structures namely: the data processing centre and the container shed, in order for the site to be ready to receive and process the questionnaires. Although the generator house was provided and plans to install a generator were made, due to very late delivery the intended generator was never installed.

11.2 Recruitment and Training

11.2.1 Recruitment

Two types of personnel were charged with the task of processing the 2002 Population and Housing Census: permanent and temporary staff. Permanent staff were recruited from the National Bureau of Statistics, Tanzania Mainland and from the Office of the Chief Government Statistician, Tanzania Zanzibar. Some of the permanent staff were seconded from other Government institutions, namely, the President's Office, Planning and Privatisation; Eastern Africa Statistical Training Centre; and the National Examinations Council of Tanzania. Temporary staff members were recruited from the open market by advertising in the various local media.

The selection criterion was based on **oral and practical interviews on use of personal computers** and thereafter the best ones were recruited. This strategy was necessary due to the fact that the time allocated for data capture and processing of results was limited; the time allocated for training did not permit basic computer course rather it was allocated for the computer specific data capture application training; and the census output was required to be ready at the shortest possible time. The results of the recruitment were to have a team that would work around the clock in three shifts as such individual discipline was critical for the whole exercise. The following is the summary of the recruitment.

Table 11.1: Number of Data Processing Staff

SNO:	SECTION	SHIFT			TOTAL
		A	B	C	
1.	Supervisors	3	3	3	9
2.	Network Administrator	1	1	1	3
2.	Database Administrator	1	1	1	3
3.	Scanner Technician	1	1	1	3
4.	Scanner Support Team	3	4	3	10
5.	Data Entry Clerks	3	3	3	9
6.	Control Clerks	2	2	2	6
7.	Computer Data Editors	40	40	40	120
8.	Scanner Operators	33	33	33	99
9.	Casual Labourers	10	10	11	31
10.	Cleaners	3	2	2	7
	Total	100	100	100	300

11.2.2 Training

From past experience, and in particular from the processing of the 1988 Population and Housing Census whereby use of the Census and Surveys Processing package developed and maintained by the International Programs Center of the US Bureau of Census was made, the same package was adopted again. However, there was a need to update the programmers on new features and enhancements of the same. One permanent staff attended a four week workshop at the center that was conducted just three months before the enumeration period. The staff member had also an opportunity to attend a two week work/study session at the centre where support from the Bureaus staff was obtained to start setting up the data entry programs, just in case manual data entry was adopted. This time was also used to develop the computer editing programs and tabulation programs which were based on the United Nations standards.

Another capacity building training programme for three permanent staff on the use and maintenance of scanners was carried out at Milton Keynes, England for a period of two weeks. It was found necessary to familiarize the staff with the new technology so that they could be of help to others during the actual exercise. This training also enabled CCO staff members to develop ownership of the system rather than depending totally on the supplier of the technology.

A one week workshop for ten (10) selected supervisors was carried out in advance of the general training for the rest of the temporary staff. This was done in order to increase the number of supervisors due to the fact that permanent staff were few and sections which needed supervision were many. This team was later given supervisory responsibilities over various sections and the permanent staff became overall supervisors.

A two week crash program was carried out for temporary staff at the Data Processing Center. This training was mainly based on the familiarization of the questionnaires, the process involved in the data capture and flow of the work at the site. During the last three days before the start of scanning, a dress rehearsal was conducted for all shifts by practical testing of all subsystems with questionnaires flowing from one subsystem to the other and repeating as many times as necessary in order to build confidence of the staff.

11.3 Receipt of Questionnaires from the Field and Storage

The receipt of questionnaires began with those from Ilala and Kinondoni Districts on September 16th 2002. During the receipt period, the Data processing Center was receiving questionnaires from one, two, or three regions in a day. The peak day was on September 28, 2002 when the centre received questionnaires from five regions. The last region to submit its questionnaires was Tabora, and that was on October 8, 2002, seven days before scanning was scheduled to begin.

Most regions followed CCO instructions for packing questionnaires for each Enumeration Area (EA) by starting with unused questionnaires, followed by destroyed questionnaires and finally used ones with the types of questionnaires kept apart by separators. The EA boxes were then labelled on the outside with the respective identification names and codes. A group of six EA boxes were packed in a bigger box with codes of the content EAs on each.

Activities that were being carried out during the receipt of questionnaires were:

- i) Physical verification of EAs from the regions against the inventory that was developed by the cartographic unit;

- ii) Removing the unused and destroyed questionnaires from the EA boxes;
- iii) Resizing, re-taping and re-labelling the EA boxes;
- iv) Arranging the EA boxes serially by region, district, ward and Enumeration Area on the racks that were installed in the two renovated warehouses; and
- v) Sorting out cartographic materials and control forms as well as all unused stationery that were delivered together with the questionnaires from the regions.

11.4 Scanning of Questionnaires and Archiving

11.4.1 Introduction

In general, questionnaire handling in the field was good as care was taken to ensure that they were kept clean and tidy. Shading was also so good that even the scanning technology suppliers remarked that they had not seen forms of the same nature anywhere. There was only one exception out of over 55,200 Enumeration Areas (EA) where an EA was entered manually after failing to be scanned. The questionnaires in this EA from Pwani Region seem to have come into contact with water.

11.4.2 Equipment Used

Data Capture was done using eleven (11) CD 800 series scanners under Optical Mark Reader Technology (OMR) attached to COMPAQ PCs running under Windows XP professional operating system with flat screens. An APC Uninterruptible Power Supply unit was attached to each set of equipment. Although 12 scanners were purchased for the purpose, one was damaged in transit from UK to Dar es Salaam. These sets of equipment were connected in a Local Area Network to a dual processor server running under windows 2000 operating system.

11.4.3 Software Used

Development of the software used to scan the census forms was done in collaboration between the permanent staff of the Data Processing Unit in the Central Census Office and Data & Research Services (DRS) of the United Kingdom. DRS were the technology suppliers who developed the scanning edit specifications as well as several verification procedures which were tested at DRS, UK and Mikocheni Data Processing site in Dar es Salaam. This helped build confidence that the technology was manageable after having an opportunity to see the conditions of questionnaires that were being scanned at DRS. The scanning software was developed in Delphi programming language.

A separate Batch Registration System was developed in Microsoft Access to register and produce batch headers, in specific Enumeration Areas.

11.4.4 Stages Involved in Scanning

The following were a few but necessary activities that had to be carried out just before the scanning of the questionnaires:

- i) Counting the number of used forms in an EA, splitting the long questionnaires and writing the count and the type of form on top of the EA box. This was necessary to facilitate checking of the number of forms during the scanning process;

- ii) Registering the EA received from the field and producing two batch headers for use by the scanning section to label the accepted and rejected forms as well as preparing envelopes that were used to store rejected forms; and
- iii) Registering the batches ready for scanning.

Scanning of census forms began on October 15, 2002 with three shifts in operation for an average of eight (8) hours each.

During the scanning session, questionnaires which had been accepted by the scanner were returned to their original boxes and sent back to the warehouse. Few rejected ones were sent to the data validation section.

In the validation section, the rejected forms were recalled from the system using barcodes and handheld scanners and the necessary corrections were done. The rejected forms which were kept in envelopes were later sent back to warehouses and put in their respective EA boxes.

11.4.6 Completion

The Scanning exercise was completed on November 9, 2002 after 26 days that is a scanning average of a region a day. Since questionnaires from all regions were on site before starting, management at the Data Processing Center decided to follow the ascending order of regional codes to scan forms and easily monitor progress. Rescanning of rejected forms was conducted in parallel with the main scanning exercise and was completed five days later.

11.5 Comparison with Manual Counts

During the scanning process comparison of scanning count that had undergone basic editing of the sex variable was done with that from the field which was derived manually after enumeration. The outcome of the comparison was that there was a very small difference between the respective counts. For more details see Appendix 11.1.

11.6 Data Cleaning and Preparation of the General Report

After successful completion of scanning, data processing activities continued with the first activity which was to enable the production of the General Report that had the numbers of population counts by sex for use by planners.

The cleaning of scanned data images began on November 15, 2002. During that exercise, the following set of activities were carried out:

- i) Export of data files by region. This exercise was necessary to convert questionnaire scanned images into ASCII format files as per format that was developed ready for use by the application package of Census and Surveys Processing (CSPPro);
- ii) Resolving mismatch of identification blocks by subjecting the data through a basic program written using CONCOR package and identifying the offending records by barcode;
- iii) Creating regional backups of the exported and aligned data;
- iv) Merging of the regional files to create one master file for the whole country;
- v) Development of a basic program to edit the data and distribute the same by sex in order to prepare the General Report;

- vi) Development and production of the basic tables using the CENTS package that segregated the data by region, district and ward/shehia; and
- vii) Saving the above output in text formats and later editing the same in Excel, making them ready for use.

11.7 Computer Editing

The main computer editing began after the preparation of the General Report that enabled the announcement, by the President of United Republic of Tanzania, of the basic counts of population on the 2003 new year s eve.

There were two types of edit specifications that were developed. One type was based on United Nations Guidelines that points out how each missing value should be handled by looking through other variables in the questionnaire. At this stage subject specialists/users developed the other set of edit specifications. In this set all missing values were supposed to be assigned the code of 9 meaning not stated . It was later agreed that all types of edit specifications be utilized and later the subject matter specialists/users will decide which set of results will be useful to them. Hence it was necessary to have this part of processing split into two phases whereby the first would deal with the structure of the collected data and the next phase would carter for content editing.

11.7.1 Structural editing

The main objectives of the first phase of structural editing were to work on the following areas:

- i) General ordering of hierarchy, looking for duplicates and gaps;
- ii) Relationship between total counts on the housing record and sum of population records total, and by sex. Methods of correcting the problems, including determination of the number of entries to trigger a person;
- iii) Problem of extra housing records for continuation forms;
- iv) Problem of housing records for collective households;
- v) Problem of non-blanks for short forms, collective households and whether these were to be deleted or to be given a second look; and
- vi) Problem of merged households that created the problem of extra heads of households.

In order to accomplish the above listed objectives, several small programs were written using the CONCOR editing package. The main purpose of the said programs were:

- i) To identify and eliminate blank questionnaires. There were two types of such questionnaires: One type was of forms, which did not have any identification block filled in. The second type was of forms that had the identification block filled in but no housing or person records. There was an opinion that the main cause of such forms being found in the enumeration area was that of enumerators mixing blank and filled in forms as well as filling in many forms with the identification block in advance before reaching the house to be enumerated and forgetting to remove them from the lot when they had not been utilized;
- ii) To identify and eliminate the ghost records. These were determined by ensuring that each person had at least the first basic variables of Relationship, Sex and Age filled in. These were found to be crucial in accepting one as a person in the database. As a result after this edit the population count reported in the General Report decreased by 0.36 percent;

iii) To certify that each private household had a head of household. This person had to have a Relationship code filled in as 1. While this certification was going on, five sets of cases were identified:

- a) **A case when there was no head.** This case was treated by assigning the first person listed in the house as the head and other members of that household retaining their reported relationships;
- b) **A case with only one head.** This case was accepted as a valid private household;
- c) **A case with only one head but under age.** This case was resolved by assigning the young head as child and locating the oldest person in the house who was related to the young as head, that is with relationship code other than 7 that was assigned to non-relative to replace the young head. Since some of the young heads were as old as zero years it was decided that the cut —off age for a child to be head of household should be 5 years.
- d) **A case with more than one head.** This case was caused by two main reasons namely miss-shades on the identification block up to the last identification field of form number; and the miss-shade on the relationship code. A split was done where necessary. Duplicate households were determined by first searching for extra head of households. Where it was found, another check was done to determine whether it was a private or collective household. All collective households were not supposed to have heads.

Once it was determined that they were private, yet another check was done to find whether there was more than one barcode and if not, that house was not split. If extra heads and extra barcodes were found, an investigation was done to a person under serial number 7 for long forms and serial number eleven for short forms. The aim was to determine whether it was a continuation form with age of the head looked at and the respective relationship if it is that of head. For cases where there was no continuation form, the house was split and the first part of the house retained the same household number with the rest being assigned an alpha on third position ready for proper allocation of numeric household number.

Used household numbers were saved in an array for each EA and those unused were given to the ones with alphas once a match on an EA was found; and

- e) **A case of collective households having heads.** Collective households include fishing camps, hospitals, schools, hotels, guest houses, migratory populations and vagrants. In actual fact, most of these people are not related in any way hence it was not necessary to assign heads of households in such houses. It was finally agreed to have the relationship code of 7, reserved for non-related people assigned to all those identified to have the relationship code of 1 that was reserved for head of household in private households;
- iv) To certify that if there is a spouse in the house, that is a person assigned a relationship code of 2 is not under age, it was decided to have a cut-off age of 10 years for spouses and those below the cut-off age were changed to children of the head of household by assigning them a relationship code of 3;
- v) To ensure that in a household where there is no head, there is no spouse as well. If a spouse is found in that house, the relationship code for that person is changed to 6 implying that the person is changed to another relative;

- vi) To ensure that in case a house had more than one spouse, all spouses, however many, were checked individually for their age. All were not to be under 10 years;
- vii) To ensure that distinction between private and collective households was clear. This implied that we had to ensure that all Enumeration Areas and household numbers for these groups of households were known; and that only checks relevant to private households were carried out on such households. Generally, collective households were all those with the following identifications:
 - a) **Enumeration Areas from number 501 to 699** were mainly used to identify people who were resident in the Kagera Region refugee camps; number **788** was used to identify vagrants and those who were travellers on the census night; from number **801 to 899** were used to identify migratory population and fishing camps; and from number **901 to 999** were used to identify schools, hospitals, etc.; and
 - b) **Household Numbers from 901 to 999**. It was earlier expected that the collective household numbers would be in the range of **951 to 999** but the reality from the field was that the range had changed from the expected to 901 to 999. There was an opinion that, either the enumerators confused these numbers with the enumeration area range of 901 to 999; or the allocated range of 951 to 999 was not enough since with this range it was possible to assign 49 households only. It should also be borne in mind that some Enumeration Areas were big creating a possibility of having more than 49 collective households.

It should be noted that, in some areas where the above enumeration areas were demarcated, a few private households were among households enumerated in such enumeration areas. Such households were allocated household numbers starting from 001 as for other private households.

11.7.2 Content Editing

As explained earlier, content editing was done following the United Nations Guidelines of determining missing values by investigating the rest of the data files in the respective record or records in a household. It was necessary to determine needed records, and to write programs to add them on to the CONCOR editing program. Mainly, first it was necessary to find out whether for each field valid values as assigned in the questionnaire were reported, that is range checks. If they were missing, determination of the values was done by examining other variables. Secondly, consistency among reported values was examined and if necessary the recommended guidelines were applied to resolve the inconsistencies.

11.8 Tabulation of Results

The development of the tabulation plan started in early 2001. This tabulation plan was discussed by the stakeholders and their suggestions were accommodated wherever possible. Since the demands for data are many and varied it was agreed that there should be two types of tables. Group A tables were to contain basic data and hence were termed as priority tables. Group B tables were to contain data of lower priority. Group A tables were ready by December, 2002. Group B tables were at an advanced stage of preparation then.

11.9 Creation of Database

Steps had already been taken by the National Bureau of Statistics to establish a national database as way back as early 2001. These steps included:

- i) Dialogue with other institutions in the country which are involved in data collection with the view to harmonizing data collection procedures and finally sharing electronically the data already collected;
- ii) Establishment of Tanzania Social and Economic Database (TSED); and
- iii) Negotiations with JICA in order to get assistance in establishing Integrated Database Management System for all the data, including census data, that is available at the NBS.

11.10 Technical Assistance

In preparation for scanning the questionnaires, the data processing unit obtained technical assistance from Mr. Mwale of the Examinations Council of Zambia and from DRS of the UK. In carrying out the processing of the 2002 Population and Housing Census data, Technical Assistance was obtained from the International Programs Centre (IPC) of the US Bureau of Census. The assistance began when one programmer attended a four week Workshop on Processing of Censuses and Surveys at the Centre and later a work-study session for two weeks. The knowledge gained assisted the programmer in developing most of the editing and tabulation programs. After data capture, a one-week mission by an IPC member of staff was conducted. The outcome of that visit was a development of a fixed programme of work to accomplish the data processing activities in the shortest possible timeframe. A relay plan as outlined below, was put in place and implemented. A three week mission on structure edits to sort out structural issues in the data; development of the remaining table sets; and lastly, a final mission to finalise the tabulation programs as well as editing programs. During the last mission an investigation was done on the correctness of the data by deriving various demographic indices. It should be noted that together with the above mentioned technical assistance, nationals carried out the entire data processing exercise.

11.11 Problems Faced and Recommendations for Future Censuses

11.11.1 Problems

- i) Problems that were faced while receiving the questionnaires from the field were:
 - a) A number of big boxes did not have labels that showed the contents;
 - b) Labels of other big boxes did not match with the content labels;
 - c) Some EA numbers did not match with inventory that existed;
 - d) Mixing of EAs from various wards in one big box thus creating delay in sorting them out;
 - e) Missing expected EAs and getting others in duplicates compared to the inventory;
 - f) Some of the enumerators did not follow the packing instructions by interchanging the layers for used and unused forms;
 - g) Duplicate EAs due to poor labelling; and
 - h) Missing and poorly filled in control forms that were hardly useable.
- ii) The following were major technical problems which the data processing unit faced during scanning:
 - a) There was no standby generator for use in case of electricity failure. However, temporary staff were trained in managing all subsystems hence in a case of power failure, they were assigned to do other activities;

- b) System failures i.e. the system became very slow or sometimes hanged up completely when more than 50 computers were operating simultaneously. In such cases the staff were reduced and assigned other duties;
 - c) Fear of virus attacks due to late delivery of the server anti virus software. In order to prevent this from happening, all diskette drives were disabled;
 - d) Mixing of used and unused questionnaires and that of long and short questionnaires in an EA box. This problem was experienced in a few EAs in each region. Since scanners were programmed to accept one type of questionnaire at a time, the data processing site Management Team decided to transfer all filled in short questionnaires into long ones whenever such problems were encountered; and
 - e) File corruption due to power fluctuations that could not be prevented due to non-installation of UPSs before scanning operations began. Missed batches due to this effect were rescanned.
- iii) Problems which the Data Processing Unit faced during the production of the General Report:
- The basic problem faced during the production of the General Report was how to resolve mis-shaded identification codes causing some data to appear as if they were from another region, district, ward or enumeration area while that wasn't the case. Since problems of this nature were confined to each region at a time it was possible to detect the cause of the problem and take corrective action. An example of such a case was as follows; say region **code 16** which was meant for **Kigoma Region** was shaded as **code 06** which was allocated to **Pwani Region**. A re-export exercise was then necessary to effect the necessary correction.
- iv) Problems Faced During Data Processing:
- Some problems which were faced during this stage of data processing delayed the completion of the edit programs as scheduled. Such problems were as follows:
- a) Treatment of duplicate records. During training for enumerators and supervisors it was emphasized that, if thorough editing was to be done in the field then special attention should be paid to the identification block. The reality showed that this was not the case and the role of supervisors was not played well. Most of the duplicate cases in the identification block would have been reduced if supervisors had done their job well;
 - b) It was not made clear to all levels of supervision that collective households had a changed range of codes from the previously allocated one. It was after checking the actual data during the structure edits that this fact became known; and
 - c) The Data processing Teams view was that supervision during enumeration was not done well. Considering the various levels of supervision that existed, a number of problems faced during data processing like mis-shades, un-shaded responses etc. could have been corrected in the field.

11.11.2 Recommendations

A data processing exercise is a major undertaking that needs commitment and strong support from all those involved. A series of lessons that were learnt are narrated below. We expect the managers of the next census to make use of the same to assist them in getting the results out much earlier.

i) **Preparation of the Data Processing Site**

Good plans to have the site ready for the task ahead were there. However, implementation was not on time. Such necessary items as a generator, computer tables and office furniture, curtains, UPS, anti virus software, fire extinguishers etc. were not put in place on time. Some items like the generator was not delivered at all while others were received well after the data capture exercise was complete. The series of problems that were being experienced would not have occurred if the implementation phase was executed as planned.

ii) **Recruitment and Training**

- a) It is recommended that the recruitment process be as transparent as possible to avoid complaints;
- b) It is recommended that both oral and practical interviews be used;
- c) Experience showed that those workers recruited from other working places were more disciplined than those who were direct from school. Therefore it is recommended that one of the qualifications for data processing recruitment should include past working experience; and
- d) It is recommended that an optimum time and positioning of theoretical and practical training should not be missed out. Too early or too late training before the actual work should be avoided. This must be followed by a dress rehearsal for all processes as a lot of mishaps can be corrected before live runs.

iii) **Supervision in the field, questionnaire handling and packing**

- a) It is recommended that the various levels of supervision should be revisited during the next census;
- b) It is recommended that used questionnaires should be kept separate from those unused and those destroyed should be completely cancelled across; and
- c) It is recommended that used and unused questionnaires should be counted by enumerators and the count written on top of the Enumeration Area box. The boxes should also be identified as long questionnaire (LQ) or short questionnaire (SQ) before handing the boxes over to the supervisors.

iv) **Employing Past and Other Countries Experiences**

It will save a great deal of time and money if enough research on similar projects from other countries is done well in advance and results put to use. Experiences gained and ways used to solve them helps the project supervisors not to re-invent the wheel and wherever possible to leapfrog. During the 2002 Population and Housing Census a lot of experience was tapped and employed from other countries by means of correspondence and personal contacts with people who worked on similar projects in Ghana, Kenya, Zambia and South Africa. The gained knowledge shows that the following points need to be taken into consideration:

- a) Enough questionnaires and of the right quality should be printed once a decision to use scanning technology is reached;
- b) A Standby generator must be in place before data processing begins, because constant power supply is not guaranteed and thus this can cause considerable delay to the whole process;

- c) It is advisable to have a single supplier for all systems. This package should also include technical support and in-house capacity building. Depending fully on outside technical support creates another danger of knowledge gap and causes lack of continuity as well as creates confusion and much delay in some cases;
- d) A smallest unit of batch once decided for processing should not be split;
- e) Barcodes that are printed on the forms should be captured and carried on board throughout the data processing exercise;
- f) The earmarked Data Processing Team should work hand in hand with the technology suppliers as ownership and responsibility of the system can only be achieved in this manner;
- g) Data Processing Unit should be established well in advance and work hand in hand with the rest of the Census Project Team starting from the stage of questionnaire design;
- h) The established management team of the Census Project should not be changed until the project is complete. This will ensure continuity in carrying out various activities of the project; and
- i) The Data Processing Team must be ahead in testing every piece of functionality to identify anticipated problems that might arise in order to be able to take necessary actions in time.

v) **Data Archiving**

- a) The whole exercise will be meaningless if no data set is archived for retrieval during disaster. In this regard, it is recommended that all data sets be backed up and stored in fireproof data safes at separate locations.
- b) Good progress in the data processing exercise can only be achieved with a spirit of teamwork between various parties, especially between programmers and subject matter specialists. These two groups need to work together right from the beginning and maintain this spirit to the finishing line. A considerable amount of time can be saved if this arrangement is established early.

Total Number of EAs, Number of Selected EAs and Number of EAs Utilized Using Long Questionnaire.

Region	District	Total Number of EAs			Number of Selected EAs			Number of Utilized EAs		
		Total EAs	Rural	Urban	Rural	Urban	Total	Rural	Urban	Total
01 Dodoma	1. Kondoa	530	478	52	50	50	100	50	50	100
	2. Mpwapwa	335	286	49	40	50	90	39	49	88
	3. Kongwa	276	240	36	40	36	76	40	36	76
	4. Dodoma (R)	534	512	22	50	22	72	50	22	72
	5. Dodoma (U)	538	210	328	40	50	90	41	49	90
	TOTAL	2213	1726	487	220	208	428	220	206	426
02 Arusha	1. Monduli	338	278	60	40	50	90	40	48	88
	2. Arumeru	730	544	186	50	50	100	50	50	100
	3. Arusha	629	19	610	19	50	69	15	50	65
	4. Karatu	268	248	20	40	20	60	40	20	60
	5. Ngorongoro	190	174	16	30	16	46	30	16	46
	TOTAL	2155	1263	892	179	186	365	175	184	359
03 Kilimanjaro	1. Rombo	359	327	32	40	32	72	39	32	71
	2. Mwanga	201	134	67	30	50	80	29	50	79
	3. Same	433	261	172	40	50	90	41	48	89
	4. Moshi (R)	521	501	20	50	20	70	50	20	70
	5. Hai	477	406	71	50	50	100	52	48	100
	6. Moshi (U)	318	0	318	0	50	50	0	50	50
	TOTAL	2309	1629	680	210	252	462	211	248	459
04 Tanga	1. Lushoto	537	501	36	50	35	85	49	35	84
	2. Korogwe	387	273	114	40	50	90	40	48	88
	3. Muheza	377	302	75	40	50	90	40	50	90
	4. Tanga	481	78	403	30	50	80	29	50	79
	5. Pangani	69	53	16	30	16	46	30	16	46
	6. Handeni	287	245	42	40	42	82	42	42	84
	7. Kilindi	30	30	0	30	0	30	29	0	29
	TOTAL	2168	1482	686	260	243	503	259	241	500
05 Morogoro	1. Kilosa	816	530	286	50	50	100	50	49	99
	2. Morogoro	390	370	20	40	20	60	40	20	60
	3. Kilombero	503	268	235	40	50	90	40	50	90
	4. Ulanga	289	217	72	40	50	90	43	48	91
	5. Morogoro (U)	559	31	528	10	50	60	12	47	59
	6. Mvomero	405	337	68	40	50	90	40	50	90
	TOTAL	2962	1753	1209	220	270	490	225	264	489
06 Pwani	1. Bagamoyo	306	209	97	40	50	90	40	48	88
	2. Kibaha	244	96	148	30	50	80	29	51	80
	3. Kisarawe	156	121	35	30	35	65	30	35	65
	4. Mkuranga	283	233	50	40	49	89	40	49	89
	5. Rufiji	350	227	123	40	50	90	41	49	90
	6. Mafia	60	43	17	31	16	47	30	16	46
	TOTAL	1399	929	470	211	250	461	210	248	458
07 D Salaam	1. Kinondoni	2918	57	2861	30	70	100	15	85	100
	2. Ilala	1700	56	1644	30	70	100	30	73	103
	3. Temeke	2140	64	2076	30	70	100	31	69	100
	TOTAL	6758	177	6581	90	210	300	76	227	303
08 Lindi	1. Kilwa	292	251	41	40	42	82	41	41	82
	2. Lindi (R)	397	328	69	40	50	90	40	50	90

Region	District	Total Number of EAs			Number of Selected EAs			Number of Utilized EAs		
		Total EAs	Rural	Urban	Rural	Urban	Total	Rural	Urban	Total
	3. Nachingwea	236	180	56	30	50	80	30	44	74
	4. Liwale	121	81	40	30	40	70	30	40	70
	5. Ruangwa	209	158	51	30	48	78	29	46	75
	6. Lindi (U)	89	17	72	17	50	67	16	51	67
	TOTAL	1344	1015	329	187	280	467	186	272	458
09 Mtwara	1. Mtwara (R)	308	272	36	40	37	77	40	36	76
	2. Newala	384	311	73	40	50	90	39	50	89
	3. Masasi	785	542	243	50	50	100	49	50	99
	4. Tandahimba	375	265	110	40	50	90	40	50	90
	5. Mtwara (U)	220	17	203	17	50	67	17	50	67
	TOTAL	2072	1407	665	187	237	424	185	236	421
10 Ruvuma	1. Tunduru	298	243	55	40	50	90	40	50	90
	2. Songea (R)	186	164	22	30	20	50	30	22	52
	3. Mbinga	491	436	55	50	50	100	50	50	100
	4. Songea (U)	263	33	230	30	50	80	31	49	80
	5. Namtumbo	235	207	28	40	29	69	43	27	70
	TOTAL	1473	1083	390	190	199	389	194	198	392
11 Iringa	1. Iringa (R)	370	353	17	40	17	57	37	17	54
	2. Mufindi	389	315	74	40	50	90	39	50	89
	3. Makete	185	169	16	30	16	46	31	16	47
	4. Njombe	671	474	197	50	50	100	48	48	96
	5. Ludewa	188	157	31	30	31	61	30	31	61
	6. Iringa (U)	257	10	247	10	50	60	10	50	60
	7. Kilolo	304	262	42	40	42	82	40	42	82
	TOTAL	2364	1740	624	240	256	496	235	254	489
12 Mbeya	1. Chunya	305	271	34	40	23	63	39	24	63
	2. Mbeya (R)	382	289	93	40	50	90	40	47	87
	3. Kyela	246	193	53	30	50	80	30	50	80
	4. Rungwe	419	374	45	40	45	85	39	45	84
	5. Ileje	151	145	6	30	6	36	30	6	36
	6. Mbozi	649	556	93	50	50	100	50	49	99
	7. Mbarali	307	230	77	40	50	90	39	47	86
	8. Mbeya (U)	588	46	542	30	50	80	31	50	81
	TOTAL	3047	2104	943	300	324	624	298	318	616
13 Singida	1. Iramba	473	411	62	50	50	100	49	50	99
	2. Singida (R)	573	532	41	50	41	91	50	41	91
	3. Manyoni	246	180	66	30	50	80	30	50	80
	4. Singida (U)	208	76	132	30	50	80	30	50	80
	TOTAL	1500	1199	301	160	191	351	159	191	350
14 Tabora	1. Nzega	579	521	58	50	50	100	49	49	98
	2. Igunga	358	311	47	40	48	88	40	47	87
	3. Uyui	267	256	11	40	11	51	40	11	51
	4. Urambo	443	385	58	50	45	95	40	50	90
	5. Sikonge	182	163	19	30	19	49	30	19	49
	6. Tabora (U)	373	94	279	30	50	80	30	54	84
	TOTAL	2202	1730	472	240	223	463	229	230	459
15 Rukwa	1. Mpanda	651	505	146	50	50	100	54	46	100
	2. Sumbawanga (R)	553	504	49	50	49	99	44	49	93
	3. Nkansi	292	218	74	40	50	90	36	50	86
	4. Sumbawanga (U)	233	77	156	30	50	80	29	50	79
	TOTAL	1729	1304	425	170	199	369	163	195	358

Region	District	Total Number of EAs			Number of Selected EAs			Number of Utilized EAs		
		Total EAs	Rural	Urban	Rural	Urban	Total	Rural	Urban	Total
16 Kigoma	1.Kibondo	348	313	35	40	35	75	40	35	75
	2.Kasulu	572	500	72	50	50	100	50	50	100
	3.Kigoma (R)	498	444	54	50	50	100	50	52	102
	4.Kigoma (U)	344	16	328	16	50	66	16	51	67
	TOTAL	1762	1273	489	156	185	341	156	188	344
17 Shinyanga	1.Bariadi	822	746	76	50	50	100	48	48	96
	2.Maswa	390	328	62	40	47	87	38	49	87
	3.Shinyanga (R)	345	342	3	40	3	43	40	3	43
	4.Kahama	809	659	150	50	50	100	48	51	99
	5.Bukombe	576	498	78	50	50	100	50	50	100
	6.Meatu	324	309	15	40	15	55	40	15	55
	7.Shinyanga (U)	256	77	179	30	50	80	30	50	80
	8.Kishapu	322	287	35	40	35	75	40	35	75
	TOTAL	3844	3246	598	340	300	640	334	301	635
18 Kagera	1.Karagwe	460	455	5	50	5	55	50	5	55
	2.Bukoba (R)	547	535	12	50	11	61	50	11	61
	3.Muleba	419	402	17	50	12	62	49	12	61
	4. Biharamulo	546	469	77	50	50	100	50	50	100
	5. Ngara	415	408	7	40	7	47	39	7	46
	6. Bukoba (U)	127	20	107	20	50	70	20	50	70
	TOTAL	2514	2289	225	260	135	395	258	135	393
19 Mwanza	1. Ukerewe	332	270	62	40	50	90	40	50	90
	2. Magu	497	413	84	50	50	100	49	50	99
	3. Nyamagana	69	0	69	0	70	70	0	69	69
	4. Kwimba	367	333	34	40	33	73	40	33	73
	5. Sengerema	545	466	79	50	50	100	50	50	100
	6. Geita	845	700	145	50	50	100	55	45	100
	7. Misungwi	318	275	43	40	43	83	40	43	83
	8. Ilmela	426	96	330	30	50	80	23	42	65
	TOTAL	3399	2553	846	300	396	696	297	382	679
20 Mara	1. Tarime	724	535	189	50	50	100	50	50	100
	2. Serengeti	247	208	39	40	46	86	40	39	79
	3. Musoma (R)	408	385	23	40	23	63	39	23	62
	4. Bunda	395	257	138	40	50	90	40	50	90
	5. Musoma (U)	243	4	239	4	50	54	4	50	54
	TOTAL	2017	1389	628	174	219	393	173	212	385
21 Manyara	1. Babati	468	380	88	40	50	90	40	50	90
	2. Hanang	303	254	49	40	49	89	40	49	89
	3. Mbulu	336	299	37	40	38	78	41	37	78
	4. Simanjiro	232	129	103	30	50	80	28	51	79
	5. Kiteto	190	152	38	30	38	68	30	38	68
	TOTAL	1529	1214	315	180	225	405	179	225	404
TOTAL MAINLAND		50760	32505	18255	4988	4474	9462	4955	4422	9377
51 North Unguja	1. North A	197	194	3	30	3	33	30	3	33
	2. North B	112	109	3	30	3	33	30	3	33
	TOTAL	309	303	6	60	6	66	60	6	66
52 South Unguja	1. Central	137	135	2	30	2	32	30	2	32
	2. South	67	58	9	30	9	39	30	9	39
	TOTAL	204	193	11	60	11	71	60	11	71
53 Urban West	1. West	391	155	236	30	50	80	30	50	80
	2. Town	442	0	442	0	70	70	0	70	70
	TOTAL	833	155	678	30	120	150	30	120	150

Region	District	Total Number of EAs			Number of Selected EAs			Number of Utilized EAs		
		Total EAs	Rural	Urban	Rural	Urban	Total	Rural	Urban	Total
54 North Pemba	1. Wete	208	159	49	30	49	79	30	49	79
	2. Micheweni	158	148	10	30	10	40	30	10	40
	TOTAL	366	307	59	60	59	119	60	59	119
55 South Pemba	1. Chake Chake	171	130	41	30	41	71	30	41	71
	2. Mkoani	179	156	23	30	23	53	30	23	53
	TOTAL	350	286	64	60	64	124	60	64	124
TOTAL ZANZIBAR		2062	1244	818	270	260	530	270	260	530
TOTAL TANZANIA		52822	33749	19073	4744	5248	9992	4692	5215	9907