UNITED NATIONS ECONOMIC COMMISSION FOR AFRICA

REPORT of the SEMINAR on the ORGANIZATION and CONDUCT of CENSUSES of POPULATION and HOUSING

Addis Ababa 17-28 June 1968

Vol. II

Part II

REPORT OF SEMINAR ON ORGANIZATION AND CONDUCT OF CENSUSES OF POPULATION AND HOUSING

Addis Ababa - 17 - 28 June 1968

V O L U M E I I

P a r t II*

SEMINAR PAPERS *

---

* Due to binding limitations of the Commission facilities, Volume II is made up in two parts: part I contains the papers listed as numbers I through VII, while part II of Volume II, bound separately, contains those papers which are designated as VIII-XVI (see Table of Contents, page iii).

REPORT OF SEMINAR ON ORGANISATION AND CONDUCT
OF CENSUSES OF POPULATION AND HOUSING
Addis Ababa, 17 - 28 June 1968

Volume II: Seminar papers*

CONTENTS

Paper Number		Pages	Document Number
VIII	Age Data in African Censuses and Surveys	225 - 251	E/CN.14/CPH/13
IX	Cartography for Census Purposes	253 - 291	E/CN.14/CPH/6 ST/STAT/25
X	Locality Listing and Delimination of Enumeration Areas	293 - 309	E/CN.14/CPH/9
XI	Preparation and Use of Census Control Lists	311 - 347	E/CN.14/CPH/5 ST/STAT/22
XII	Mapping for the 1966 Census of Swaziland	349 - 362	E/CN.14/CPH/19
XIII	Pilot Surveys and Pre-tests of Census Procedures	363 - 374	E/CN.14/CPH/12
XIV	Enumeration Problems and Procedures	375 - 388	E/CN.14/CPH/11
XV	Electronic Processing of Census Data	389 - 422	E/CN.14/CPH/7 E/CN.14/CPH/7/Add.1 ST/STAT/33
XVI	Methods of Evaluating the Reliability of Population and Housing Census Data	423 - 454	E/CN.14/CPH/4 ST/STAT/19

* Due to binding limitations of the Commission facilities, Volume II is made up in two parts: part I contains the papers listed as numbers I through VII, while part II of Volume II contains the papers VIII - XVI. (see table of contents, page iii)

E/CN.14/CPH/13

VIII. AGE DATA IN AFRICAN CENSUSES AND SURVEYS

(Prepared by the ECA Secretariat)

AGE DATA IN AFRICAN CENSUSES AND SURVEYS

CONTENTS

	Paragraphs
INTRODUCTION	1 - 2
I. The patterns of age mis-statement in African censuses and surveys	3 - 17
An African population pyramid	5 - 9
The directions of age mis-statement	10 - 14
Conclusions	15 - 17
II. The collection of age data	18 - 27
The form of the age question	18 - 22
Techniques of estimating age	23 - 27
III. The rectification of the reported age distributions	28 - 36
The Brass-carrier models	29 - 30
Selection of a model age distribution	31 - 32
Sex ratio models	33 - 34
Adjustments for imigration	35 - 36
Notes and references	
Appendix. Brass-Carrier Stable Population Models: Model Age Distribution per 1,000 of all ages	

INTRODUCTION

1. After the figures of total population, the most important information which any census or demographic survey should aim to collect is that on the distribution of the population by age and sex. This information is essential for almost every aspect of development planning - for studies of employment, manpower, and the economic structure of the population, for educational planning, for the analysis of fertility, mortality and population growth and for the construction of population projections. Unfortunately in Africa the accurate estimation of ages constitutes one of the most difficult problems which face the census administration, simply because of the widespread ignorance of age in the orthodox sense of the completed number of years.

2. In recognition of the peculiar importance and difficulty of the problem, the First Working Group on Censuses of Population and Housing which met in Addis Ababa in June 1965 recommended that "the United Nations should sponsor a study of mis-statements of age in African censuses and of possible methods of improving the data." This paper has been prepared in response to this recommendation. It is divided into three main parts: the first describes the patterns of age mis-statement evident in the data obtained from African censuses and surveys; the second discusses methods of collecting the age data in the field; the third suggests methods of rectifying the figures after they have been collected and tabulated.

I. THE PATTERNS OF AGE MIS-STATEMENT IN AFRICAN CENSUSES AND SURVEYS

3. This study is not the only analysis of patterns of age mis-statement to have been made in recent years. Since the recommendation by the Working Group, similar studies have been undertaken by other agencies, notably the Office of Population Research in Princeton and the Institut National de la Statistique et des Etudes Economiques (INSEE), working in conjunction with the Institut National d'Etudes Demographiques, in Paris. The results of the Princeton study have been incorporated in the new United Nations Manual on methods of estimating population[1], and comprise analyses not only of African age distributions but also those of Asian and Latin

American countries. The INSEE study[2] was confined to the former French territories in Africa south of the Sahara and covered the distributions by sex and five-year age groups obtained from censuses and demographic surveys of fifteen countries: viz. Guinea (1955); Ivory Coast (1957-59); Central African Republic (1959-60); Niger (1959-60); Senegal (1960-61); Mali (1960-61); Upper Volta (1960-61); Congo (Brazzaville) (1960-61); Gabon (1960-61); Dahomey (1961); Togo (1961); Chad (1963-64); Cameroon (1960-65); Mauritania (1964-66); and Madagascar (1957-61).

4. The data for the fifteen countries covered by the INSEE study have therefore been incorporated here, together with analyses of the age-sex data for a further fifteen African countries: United Arab Republic (1960); Morocco (1960); Libya (1964); Algeria (1966); Ghana (1960); Sierra Leone (1963); Liberia (1962); Congo (Kinshasa) (1955-57); Burundi (1965); Kenya (1962); Angola (1960); Republic of South Africa (African population only) (1960).
of Africa either no data have been collected (e.g. Ethiopia and Somalia); or they have been collected but not yet published (Nigeria 1963, Malawi 1966, Tunisia 1966 and Tanzania 1967); or they were collected in age groups too broad to be of use for this study (Rhodesia 1962, Zambia 1963, Uganda 1959 and Sudan 1955).

5. On the basis of the reported age-sex distributions for the thirty countries listed above, an attempt is made to present here a composite age-sex distribution which would reflect their principal features. It must be emphasised that this composite distribution does <u>not</u> purport to represent the age-sex structure of the population of the whole continent of Africa: in the first place many African countries were excluded from the study, for the reasons given above; and secondly in the construction of the composite distribution the figures for the individual countries have not been weighted in proportion to their total populations. The sole aim of the exercise is to illustrate the patterns of age mis-statement which usually afflict African census and survey data. The percentage age distribution of each sex and the sex ratios by age group of the composite population are shown

in Table 1, and the figures are illustrated by the population pyramid in the accompanying figure.

Table 1

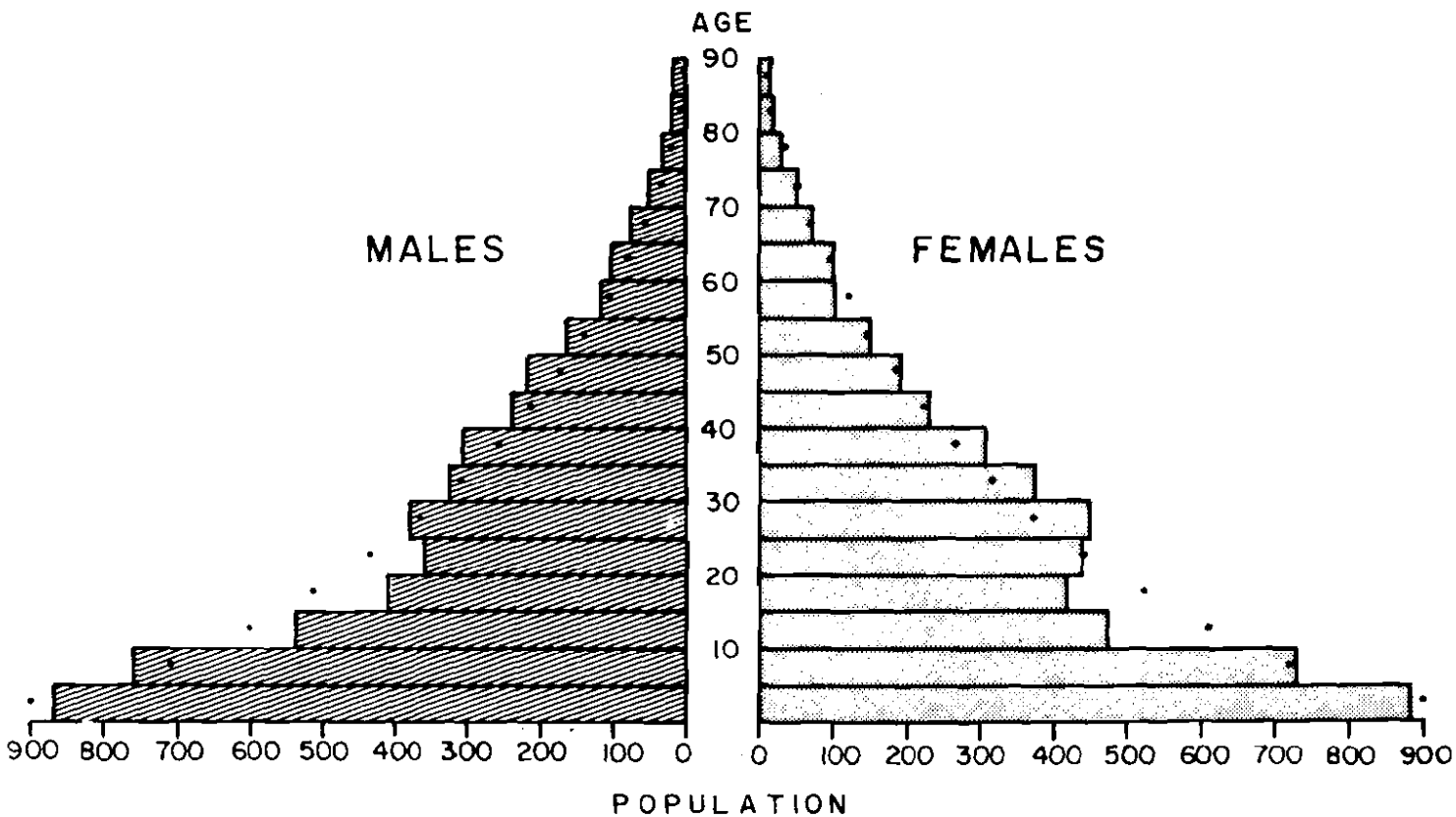
Percentage Distribution of Each Sex and Sex Ratios (Males per 100 Females) of a Composite African Age-Sex Distribution

Age group	Percentage age distribution		Sex ratios
	Males	Females	
0-4	17.5	17.5	99
5-9	15.3	14.6	103
10-14	10.8	9.4	113
15-19	8.2	8.3	97
20-24	7.2	8.7	82
25-29	7.7	8.9	85
30-34	6.5	7.4	87
35-39	6.1	6.1	99
40-44	4.8	4.6	101
45-49	4.3	3.8	113
50-54	3.3	3.0	108
55-59	2.3	2.0	114
60-64	2.0	2.0	103
65-69	1.5	1.4	106
70-74	1.1	1.0	102
75-79	0.6	0.6	103
80-84	0.4	0.4	105
85 +	0.4	0.3	106
Total	100.0	100.0	99

AN AFRICAN POPULATION PYRAMID

(BASED ON REPORTED AGE—SEX DISTRIBUTIONS OF THIRTY AFRICAN COUNTRIES)

Dots Represent Graduated Values



6. The distribution is characterised by a broad base with relatively large numbers of children in the 0-4 and 5-9 age groups. The difference between numbers in these two age groups is frequently rather less than might be expected as a result of the known patterns of child mortality and population growth, and for some countries (e.g. Botswana and Lesotho) the numbers shown as aged 5-9 were even greater than those shown as aged 0-4.

7. The relatively large numbers under 10 are followed by a pronounced trough in the 10-14 and 15-19 age groups. This trough was a marked feature of almost all the age distributions studied, and the only countries where it was not clearly apparent were the U.A.R., Algeria, Lesotho and Swaziland. It is generally more pronounced for females than for males, and the sex ratios for the 10-14 age group showed more males than females in all the populations examined except Swaziland and Congo (Brazzaville); this excess of males was generally of the order of between 10 and 20 per cent, but in several countries it was over 30 per cent, and in Niger it was as high as 56 per cent. In the 15-19 age group no consistent bias in the sex ratios was apparent, some countries showing excesses of males and others excesses of females.

8. The trough in the 10-14 and 15-19 age groups is followed by a bulge in the 20-24 and 25-29 age groups. This bulge tends to be more pronounced for females than for males; indeed in many countries virtually no such bulge for males was apparent, while in others it was confined to the 25-29 age group, following a distinct trough at 20-24.[3/] The greater prominence of the female bulge is reflected by the marked fall in the sex ratios, large excess numbers of females being reported in both the 20-24 and the 25-29 age groups. Of the thirty countries covered in this study, only three (U.A.R., Libya and Burundi) showed more males than females in the 20-24 age group, and only two (Libya and the Republic of South Africa) in the 25-29 age group.[4/]

9. Above the age of thirty the numbers of both sexes show a progressive if somewhat irregular decline with age. The sex ratios, however, which in view of the normal pattern of excess male mortality might have been expected to decline with age, show instead a marked increase, and the great majority

of the populations examined showed more males than females over the age of 40 years.[5]

The Directions of Age Mis-Statement

10. In order to account for these anomalies in the shape of most reported African age distributions, the composite distribution described above was graduated, using the stable population and sex ratio models described in the third section of this paper. By matching the graduated distributions against the reported ones, it has been possible to estimate the directions in which the ages of persons in the various age groups tend to be mis stated. Several different models were used but in every case the same conclusions were reached on the directions of age mis-statement, except in two unitaries mentioned below. The values of the graduated distributions are represented by the dots on the population pyramid.

11. It has sometimes been maintained[6] that the relatively large numbers reported as aged 0-9, followed by the sharp cut-back in the 10-14 age group may be, in part at least, a genuine feature of most African age distributions, and is attributable to drastic reductions in the levels of infant and child mortality during the decade preceding the census or survey. While it is undoubtedly true that such reductions in mortality will cause an increase in the proportion of children, we nevertheless believe that the principal cause is age mis-statement. In the first place, it appears to be a feature of almost all the reported age distributions, irrespective of the date of the census or survey; for example it is noticeable both in the data for Guinea obtained in 1955, and in those for Mauritania where the survey was conducted some ten years later. Secondly, in those countries where data from more than one census or survey are available, there seems to be no tendency for the bulge to move up into the 10-14 and 15-19 age groups (except perhaps in Madagascar where the standard of age reporting is officially higher than in most countries on the African continent). Thirdly, this type of distortion has been found, not only in African age data, but also in those of Asian countries with longer histories of census-taking; it was a prominent feature, for example, of the age distributions obtained from the Indian censuses of 1901 and 1911, when relatively little improvement in the levels of infant and child mortality had been effected.[7]

12. On the assumption, therefore, that the sharp drop in numbers between the 5-9 and 10-14 age groups is attributable chiefly to age mis-statement, the comparison with the stable population models indicated first that the 5-9 age group had been inflated by the inclusion both of children really aged 0-4[8] and of those aged 10-14. In contrast the 10-14 age group appears to have been depleted, not only by the under-statement of ages of some, causing them to be recorded in the 5-9 age group, but also in the case of females at any rate by the over-statement of ages of others, who are pushed up into the 15-19 group. In the case of males the position is somewhat obscure, the conclusions drawn being sensitive to the shape of the models used for the graduation. It has been maintained that the ages of some males in the 15-19 group tend to be under-stated, pushing them down into the 10-14 age group. The results of the present study, however, suggest that this movement may be less extensive than has sometimes been supposed, and that the high sex ratios in the 10-14 age group should be attributed primarily to the heavy two-way loss of females from this group.

13. Over the age of 15 the general over-statement of ages of females becomes increasingly pronounced. Between a quarter and a half of those really aged 15-19 are shown as aged 20-24, and a similar proportion of those really aged 20-24 are recorded as 25-29, thus creating the bulge in the latter age group noted above. Over the age of 30, however, the position as regards females is again somewhat obscure. It has been suggested that the ages of women in their late 30's tend to be under-stated,[9] but the evidence on this point is by no means conclusive. It is, however, quite clear that the general tendency towards over-statement which so distorts the data for women between the ages of 15 and 30 generally eases up, and as Coale and Demeny have pointed out,[10] the proportion of the total female population recorded as aged under 35 years tends to be one of the most reliable indices of the general shape of the age distribution. Somewhere about age 55, however, the general over-statement of female ages seems to be resumed, though the dimensions of this movement cannot be accurately estimated in view of our woeful ignorance of the patterns of African adult mortality and of past rates of growth.

14. For males the general over-statement of ages appears to begin in the 15-19 or 20-24 age group and shows no easing in the early 30's such as is

evident for females. The proportions of each age group recorded in higher groups becomes progressively larger, and after the age of 60 less than a quarter of the male population seem to have been recorded in their correct quinquennial age group.[11] The greater exaggeration of ages for males in comparison with females thus causes more pronounced heaping at older ages and the high sex ratios noted above.

Conclusions

15. The mis-statements of age which afflict the data obtained in African censuses and surveys are of a massive and deep-seated nature and are not confined to "digital preference" or rounding to the nearest numbers ending in 0, 5, 2 or 8.

16. Generally similar patterns of age mis-statement are evident throughout the continent. Indeed it may be observed that these patterns are not peculiar to Africa, and, as Coale and Demeny have shown, similar distortions are to be found in the data for Asian and Latin American countries.

17. Although the recorded age distributions for males may appear to be smoother than those for females, the inaccuracies may in fact be more serious, since the female age reporting suffers from alternating over-and under-statement, whereas that for males has a consistent and massive bias in one direction.

II. THE COLLECTION OF AGE DATA

1. The Form of the Age Question

Age or Date of Birth?

18. The censuses or surveys conducted in the majority of African countries have generally asked for the age of each individual enumerated either in completed years or in years and months. The only countries where the question on age has been replaced by one on the date of birth are Madagascar (1957-61), Rhodesia (1962), Zambia (1963) and Malawi (1966). In Rhodesia and Zambia the questions did not require the date of birth by individual years for the whole population; persons over 16 in Rhodesia and over 21 in Zambia were merely allocated to broad groups. The data obtained are therefore insufficiently detailed to permit a critical evaluation of the procedure. No results from

Malawi were available at the time of the preparation of this paper, and it is felt that conditions in Madagascar are in any case a typical of Africa as a whole.

19. When a question on date of birth has been substituted for one on age in Western countries (e.g. the U.S.A.) the information obtained is generally thought to have shown an appreciable improvement. It does not necessarily follow, however, that a similar improvement would be effected in Africa. When the majority of the population do not know their date of birth, its inclusion in the census or survey may simply impose on the enumerators a double task: they will first estimate the age, and then calculate the date of birth by subtracting the age from the date of the inquiry. On the other hand advocates of this approach argue that some respondents may in fact have a more accurate idea of their date of birth, which is fixed, than of their age, which changes from year to year, and that it might induce the enumerators to probe the question with the use of event calendars rather than simply to guess the person's age from his or her appearance.

20. Faced with this conundrum, some countries - notably Congo (Brazzaville) and Mauritius - have asked both for age and for date of birth, while others - notably Morocco (1960), Algeria (1966), Tunisia (1966) and Madagascar (1966) - have included both questions on an alternative basis: persons who knew their precise date of birth should record it, otherwise an estimate of the age should be entered. It is difficult to make an objective assessment of the success or failure of these procedures, and more data are needed before a final judgment can be passed. But it may be noted that the application of accuracy tests[12/] has not indicated that the age data for the Congo (Brazzaville) are any more accurate than those for the neighbouring countries of the Central African Republic, Gabon and Cameroun, where only age was asked; nor were the data for Morocco and Algeria noticeably better than those for Libya and the United Arab Republic. In these circumstances it seems open to doubt whether the burdening of the questionnaire with the additional column for date of birth is justified, at least in rural areas.

Single Years or Broad Groups?

21. In many African censuses the infromation on age has been obtained in the form of broad groups, the enumerators being required to allocate the

population to these groups rather than to estimate the ages (or dates of birth) in single years. This procedure was adopted, for example, in Nigeria in 1952-53, in Sudan in 1955, in Tanzania in 1957-58, in Uganda in 1959, in Rhodesia in 1962 and in Zambia in 1963. In its favour it may be argued that it simplifies the enumerators' work, and that the data in single years are so inaccurate that all that can be gleaned from them is a rough indication of the general shape of the age distribution which could be obtained equally well from broad groups. In some cases careful interpolations of these broad groups into the conventional quinquennial groups (or any other required break-down) have been made, and it may well be that these interpolated distributions are just as close to the truth as those based on the graduation of data collected in single years.

22. There are however, valid reasons for preferring the single year approach. The age distribution obtained from a census or survey is not simply of value per se, but also because it may be cross-tabulated against the other characteristics under investigation; and even though the age data may be subject to considerable errors, it is nevertheless possible to establish the general relationship between age and these characteristics. When the original information on age is restricted to broad groups the value of such cross-tabulation is greatly reduced. The United Nations therefore firmly endorses the recommendation made by the two Working Groups on Censuses of Population and Housing that age data should be collected in single years, and welcomes the fact that the resort to broad age groups is now apparently becoming unfashionable.

2. Techniques of Estimating Age

Event Calendars

23. The use of calendars of notable events to assist enumerators in pin-pointing the dates of birth of their respondents is well known and need not be elaborated here.[13/] There is however a mounting scepticism among persons concerned in this field of study as to the value of these event calendars. This scepticism rests on three major criteria. First, it has frequently been found that the population being enumerated is totally unfamiliar with many of the events listed in the calendars. This is particularly true of

political events which may have seemed of paramount importance to the educated persons drawing up the calendars, but which in fact had had relatively little impact on the lives of the majority of the people, particularly those in remote rural areas. The second shortcoming of event calendars arises from the way in which they are used. Thus enumerators have sometimes been instructed first to estimate a person's age by looking at him, and then to refine this estimate by questioning him as to whether he remembers, or associates his birth with, the appropriate events in the calendar. But all too often a man (or woman) who appears to an enumerator to be, say, 50 years old in 1968 might readily acquiesce with the suggestion that he or she was born in the year of the influenza epidemic (1918) when in fact the respondent has not the faintest idea of what the enumerator is talking about, or even knows him to be wrong. Thirdly, the effective use of event calendars is a lengthy and laborious business, and the enumerators simply do not have the time to continue probing until a really accurate estimate of age has been reached.[14/]

Conversion from a Locally-Recognised Method of Age Reckoning

24. The difficulties of obtaining accurate age statistics in Africa appear in some ways to be ironical, because although the population may be ignorant of their ages in the sense of the number of completed years, they are nevertheless acutely conscious of the general concept of age, which in Africa is associated with status in a way unknown in western society. This awareness of age is sometimes reflected in the existence of tribal "age grades" or "age sets" - groups of individuals who have undergone circumcision or some other form of rite de passage at the same time or within a specified period and which play an important role in the hierarchy of society. If the relationship between these age grades and age in years can be established, useful indications of the latter can thus be obtained.[15/]

25. This method has the considerable advantage that most people when asked what age grade they belong to, can answer immediately and unequivocally, so that no probing is required. It is however of limited applicability. Many tribes and peoples have no age grade systems, and among others the age grades may be so broad as to be virtually useless for this purpose.[16/]

Elsewhere there may be important local variations in the system and nomenclature of the age grades even within the same tribe, and considerable care and research for small geographical areas is necessary before the conversion tables showing the relationship between the age grades and age in years can be drawn up.

Relative Ages of Persons in the Same Household or Locality

26. The relationship between age and status discussed above is sometimes also reflected in a general awareness of the relative ages of persons in the same community: although people may not know their ages in years, they may nevertheless be conscious of whether they are older or younger than other members of their household or village. This fact may be exploited: having determined the ages of one or two persons with a fair degree of precision (e.g. they may be in possession of birth or baptismal certificates), those of their relatives and neighbours can be estimated in relation to them. This method is perhaps one which has received insufficient attention to date, and its systematic use may well give valuable results. In particular attention should be paid to the relative ages of persons of opposite sex, since, as has been shown in the first section of this paper, the data for males and for females tend to suffer from different types of bias. It is, however, a technique which again requires time and probing if it is to be utilised effectively.

The Importance of Pre-Census Age Education

27. It will be seen from the foregoing that one of the great difficulties which hampers the collection of accurate age statistics in Africa is that estimation procedures are laborious and time-consuming, so that enumerators cannot reasonably be expected to exploit them efficiently. The only solution is therefore that the estimation should be done before the enumeration takes place, and an "age education campaign" should constitute an important element of the census publicity. The details of such a campaign cannot be discussed here, but two important features deserve mention. First, much can be done through the schools: if the techniques of age estimation are taught in class, the school-children can then implement them for their families and neighbours. Secondly, the establishment of informal committees for small

geographical areas, comprising such persons as the chiefs, headmen, schoolmasters and priests, can be useful for the compilation of really meaningful event calendars or age grade conversion tables, or for the pinpointing of the ages of prominent persons in the locality in relation to which those of others can be established.

III. THE RECTIFICATION OF THE REPORTED AGE DISTRIBUTIONS

28. It has been recognised by demographers for several years that the conventional methods which have been adopted for the graduation of the age distributions of western populations, and which aimed at no more than the elimination of a barely discernible digital preference, are entirely useless for the correction of the massive errors normally apparent in African data. Other more drastic methods of smoothing[17] sometimes been used successfully, but in recent years informed opinion has increasingly favoured the fitting of stable or quasi-stable population models. To take but two examples, the age distribution for the African population of Kenya obtained from the 1962 Census was corrected by the fitting of a stable population model based on the United Nations model life tables,[18] and in the recent INSEE study a quasi-stable model based on the Coale-Demeny "North" model life tables was adopted for the rectification of the age distributions of the fifteen French-speaking African countries listed in paragraph 3.[19]

The Brass-Carrier Models

29. Although the United Nations and Coale-Demeny models have often given satisfactory results, they suffer from the limitation that the underlying mortality models are essentially of the "single-parameter" variety - i.e., if the level of mortality at any age is determined the mortality rates at all other ages are also fixed.[20] This feature gives rise to a certain inflexibility in the stable population age distributions derived from them. Greater flexibility may however be achieved by the use of Brass's model life table system,[21] which is based on two parameters and thus permits variations in the relationship between adult and child mortality. It also has the advantage that it is based primarily on African mortality data, whereas

in the case of the United Nations and the Coale-Demeny models no African data whatever had been incorporated in their construction.

30. A wide range of stable population models based on Brass's life table system have been constructed by Mr. N.H. Carrier of the London School of Economics. A selection of these models is reproduced in the Appendix to this paper.[22] They are arranged in groups showing the percentages under 15 (denoted by P) and over 45 (denoted by Q). They thus constitute ready-made graduated age distributions which may be taken "off the peg", and the main problem is how to select the most appropriate model. Experience shows that there is in fact no fool-proof, hard-and-fast rule for the selection of the best model; the problems vary greatly from population to population and a method which gives good results in one case may be useless in another; the element of subjective judgment can thus never be wholly eliminated.

Selection of a Model Age Distribution

31. There are basically two different approaches to the problem of selecting the best model, which we may term the "mathematical" and the "demographic" approaches. The mathematical approach is that which selects a given model on the criterion that it is a close mathematical fit to the recorded distribution - e.g. the sum of the squared deviations between the model and the recorded distribution are a minimum. With the demographic approach, on the other hand, the recorded distribution is ignored altogether, and the model is selected on the basis of its underlying parameters of fertility and mortality, which should accord most closely with those estimated for the population concerned on the basis of other information. There are serious shortcomings inherent in both approaches. Thus the mathematical approach has been criticised on the grounds that "fitting by statistical methods such as least squares is not necessarily satisfactory since the assumption is that the age errors are divided fairly equally between positives and negatives over the range; the distortion may indeed be mainly in one direction."[23] On the other hand the demographic approach sometimes cannot be used at all, because no other data are available which will enable independent estimates of fertility and mortality to be made.

32. In these circumstances no precise procedure can be laid down, but we would like to suggest the following general rules:

(i) The initial selection of the model should not be made simply on the basis of the P and Q values on the recorded distribution; the percentage age distribution of the latter should be cumulated so as to show percentage aged under 5, under 10, etc., up to age 50, and these compared with the corresponding values of the cumulated models. In this way a "short list" of possible models may be compiled. The sums of the squared differences between these models and the recorded distribution should then be calculated to find which one will give the best fit on the basis of the "mathematical" approach.

(ii) The parameters of fertility, mortality and growth underlying the selected model should at least appear to be plausible for the population concerned. For this purpose three such parameters are shown in the Appendix for each of the models - the rate of population growth (per cent per annum), the infant mortality rate and the expectation of life at birth.

(iii) The corrections which the application of the selected model will make to the recorded distribution should appear to be reasonable. For example, if the model contains a higher proportion of old people than the recorded population (thus indicating that ages had been under-rather than over-stated), the use of that model should be distrusted.

(iv) In view of the fact that the general bias in the reported ages for males tends to be more serious than in those for females, and that male age distributions also tend to be more distorted by migration, it is often advisable to select the model by comparison with the female distribution only, and to graduate the male distribution by applying model sex ratios (see below) to the corrected females. If a separate fitting for males is made, the resulting pattern of sex ratios must be examined for plausibility.

(v) Despite the apparent embarras de richesses of models, improved results can sometimes be obtained by averaging two or more models, or by making small adjustments to selected models by linear regression of the recorded distribution on the models, provided that rules (ii) and (iii) above are not violated.

Sex Ratio Models

33. For the purpose outlined in (iv) above, we have constructed a set of sex ratio models for use with African data. These models are shown in Table 2, for various assumed sex ratios at birth (males live births per 100 female). These models were based on a study of sex differentials in age-specific mortality rates obtained in twelve African countries: Cameroon,

Senegal, Mali, Upper Volta, Togo, Guinea, Dahomey, Chad, Central African Republic, Nigeria, Burundi and Congo (Kinshasa). Despite considerable irregularities for individual countries and age groups, the general pattern of these differentials appeared to be plausible and in general accordance with corresponding patterns derived from non-African populations.[24] Checks on the sex differentials in infant and child mortality were also obtained from figures of child-survival by sex available for nine countries: Mali, Niger, Guinea, Togo, Dahomey, Cameroon, Central African Republic, Congo (Brazzaville), and Gabon. The sex differentials in age-specific rates were then combined with Brass's "African Standard Life Table" to obtain the sex ratios of the life table population.

Table 2

African Sex Ratio Models: Males Per 100 Females in Each Age Group

Age group	Sex Ratio at Birth					
	100	102	103	104	105	106
0-4	97.6	99.5	100.5	101.5	102.4	103.4
5-9	96.8	98.8	99.7	100.7	101.7	102.6
10-14	96.7	98.6	99.6	100.6	101.5	102.5
15-19	96.8	98.7	99.7	100.6	101.6	102.6
20-24	96.8	98.7	99.7	100.7	101.7	102.6
25-29	96.8	98.7	99.7	100.6	101.6	102.6
30-34	96.6	98.5	99.5	100.5	101.4	102.4
35-39	96.2	98.1	99.1	100.1	101.0	102.0
40-44	95.2	97.3	98.3	99.2	100.2	101.2
45-49	94.1	96.0	96.9	97.8	98.8	99.7
50-54	91.9	93.7	94.7	95.6	96.5	97.4
55-59	88.8	90.6	91.5	92.4	93.3	94.2
60-64	85.3	87.0	87.9	88.7	89.6	90.4
65-69	81.7	83.4	84.2	85.0	85.8	86.6
70-74	78.0	79.5	80.3	81.1	81.9	82.7
75-79	73.3	74.8	75.5	76.3	77.0	77.7
80-84	66.6	67.9	68.6	69.3	69.9	70.6
85 +	52.9	54.0	54.5	55.0	55.5	56.1

34. There is a great shortage of reliable data on the sex ratio at birth in Africa,[25] but there is reason to suppose that it may be somewhat lower than in most European, Asian and Latin American countries. In the absence of any trustworthy indications to the contrary, it is suggested that for countries south of the Sahara the model based on the sex ratio at birth of 103 should be adopted, while for North African countries 105 or 106 may be more appropriate.

Adjustments for Migration

35. The age distribution and sex ratio models described above have of course been constructed on the assumption that there is no migration. Where such migration in fact exists, adjustments must be made for its effects on the age-sex distributions. In practice these adjustments cannot be made unless the appropriate information is obtained in the census or survey. In populations experiencing appreciable emigration, questions on members of the household outside the country by age and sex although inadequate as a means of recording the total number of emigrants, have nevertheless been found to go a long way towards reducing the effects of such emigration on the age distribution. For countries experiencing net immigration a simple question on birthplace will enable the immigrants to be excluded, but will not wholly eliminate the effects of migration on the age structure, since the presence of local-born children of immigrants will still cause the distribution to diverge from the stable pattern. The classification of the population by "father's country of birth", such as was made in Ghana in 1960, should permit a closer approximation to stable or quasi-stable conditions.

36. In Africa, a great part of the existing migratory movements consists of temporary labour migration of unattached males. The effects of this type of movement on the age distribution are relatively easy to eliminate. The problem arising from the migration of whole family units, however, is more intractable, and it must be admitted that no satisfactory solution has as yet been found. There is a great need for further research on African migration and its effects on age-sex structure; if such research results in the construction of useable migration models, a considerable advance will have been achieved.

NOTES AND REFERENCES

1. United Nations, Methods of Estimating Basic Demographic Measures from Incomplete Data (by Ansley J. Coale and Paul Demeny), New York 1967 (ST/SOA/Series A/42), pp. 17-22.

2. Afrique Noire, Madagascar, Comores: Demographié Comparée. 9-10 - Structures par age, actuelle et future, by F. Gendreau and R. Nadot, Paris 1967.

3. The INSEE study notes that the pattern for the countries formerly constituting French Equatorial Africa (Chad, Central African Republic, Congo and Gabon) is somewhat different, the main trough for males being situated in the 15-19 age group and the main bulge in the 35-39 or 40-44 age group.

4. The sex ratios in the 20's and 30's are of course particularly liable to be affected by migration, which may well have been instrumental in producing the excess of males in the 25-29 age group in Libya and the Republic of South Africa.

5. The only exceptions were Botswana, Lesotho and Swaziland, and Togo, Congo (Brazzaville), and Gabon. Migration again may have been a factor.

6. See, for example, the INSEE study, pp. 32, 56-63.

7. For further discussion of this point, see Etienne van de Walle, "Some Characterisitic Features of Census Age Distributions in Illiterate Populations", The American Journal of Sociology Vol. LXXI no. 5 (March 1956).

8. Apparent deficits of children in the 0-4 age group have frequently been attributed to under-enumeration, but there is a good case for believing that the major factor was really age mis-statement, which is itself attributable, in part at least, to the western method of reckoning age "at the last birthday". Thus according to the western definition a child aged 4 years 364 days is still only 4 and not 5 years old. But if the date of birth is not known precisely, there will be a natural tendency for the ages to be rounded up, thus leaving deficits at the youngest ages. It is often instructive to examine the distributions by single years of age, the patterns of mis-statement among children then being clearly apparent. Thus it occasionally happens that the "under 1" age group is inflated, since the attention of the enumerators may have been specially drawn to it, and there is a tendency for them to include all unweaned babies at this age. There is then a marked deficit at age 1, followed by inflated numbers at 2, 3, 4, 5 and 6.

9. See, for example, W. Brass, "Uses of Census or Survey Data for the Estimation of Vital Rates", Paper presented to the ECA Seminar on Vital Statistics, Addis Ababa, December 1964, paragraph 18. The figures shown in the INSEE study also suggest under-statement at these ages.

10. Op. cit., p.23.

11. Confirmation of a different sort of the extreme unreliability of the ages reported for old people is available from the Ghana 1960 data. Matching of a sub-sample of records obtained from the Post-Enumeration Survey with those from the main Census showed that less than 25% of males over 65 gave ages in the same quinquennial age group in the two enumerations, which were separated by only a few months.

12. It should be noted, however, that none of the standard accuracy tests is, in our opinion, suitable for application to African data. The United Nations test based on age and sex ratios (which was in fact used in the present instance) only measures the smoothness of the data and takes no account of general bias. Indeed in some cases the results of this test may be actively misleading. The same goes for the tests of digital preference (e.g. Myer's index or Whipple's index). An ad hoc test which was also applied in the present case consisted simply in summing the divergences of the recorded sex ratios from those of the sex ratio models. This test is also clearly unsatisfactory in that the results, may be biased by migration, but it may nevertheless be more meaningful than the other types of test.

13. For an account of the use of event calendars, see, R. Blanc, Manuel de Recherche Démographique en Pays Sous-Développé.

14. Probably the most detailed and systematic use which has ever been made of event calendars was that adopted in the second round of the 1961-63 multi-purpose sample survey of Morocco, an account of which is given by Christopher Scott and Georges Sabagh in a forthcoming paper - "The Historical Calendar as a Method of Estimating Age", shortly to be published in Population Studies. Even in this case, however, it was by no means certain that the use of the calendars effected a material improvement in the quality of the data.

15. For an account of the use of age grades among the Kikuyu and Bukusu peoples in Kenya, see J.G.C. Blacker, "Use of Sample Surveys to obtain data on age structure of the population where respondents in a regular census enumeration cannot give accurate data", World Population Conference 1965 Vol. III, pp. 126-130.

16. Even fairly broad age grades, however, may help to prevent the wholesale exaggeration of ages among old people. Thus in the 1966 Census of Swaziland use was made of the Swazi "Regiments" for the estimation of ages. The results showed remarkably little evidence of over-statement of ages among old people, and there is reason to suppose that the use of the regiments may have been instrumental in preventing this bias.

17. See, for example, N.H. Carrier and A.M. Farrag, "The Reduction of Errors in Census Populations for Statistically Underdeveloped Countries", Population Studies Vol. XII no. 3 (March 1959).

18. Kenya Population Census. 1962, Vol. III, pp. 27-30, 89. For the U.N. model life tables, see Methods for Population Projections by Sex and Age (ST/SOA/Series A/25) New York 1956.

19. Gendreau and Nadot, op. cit., pp. 58-66. For the Coale-Demeny model life tables, see Ansley J. Coale and Paul Demeny, Regional Model Life Tables and Stable Populations, Prineton 1966.

20. The Coale-Demeny tables provide a greater flexibility than the U.N. models in that they present four different "families" of model life tables and stable populations, with differing patterns of mortality.

21. See W. Brass, "Uses of Census or Survey Data....", paragraphs 30-37.

22. Thanks are due to Mr. Carrier for permission to reproduce these models. Responsibility for the selection of the particular models shown here, and for the comments on their use, rests with the authors of this paper.

23. W. Brass, "The Use of Existing Data", in The Population of Tropical Africa (ed. J.C. Caldwell and C. Okonjo) Longmans 1968, pp. 174-5.

24. See, for example, the discussion of sex differentials in mortality in the United Nations monograph, Age and Sex Patterns of Mortality: Model Life Tables for Under-Developed Countries (ST/SOA/Series A/22) New York 1955, pp. 16-20.

25. One of the major problems in obtaining information of the sex ratio at birth is that it is subject to large random errors, so that relatively large numbers are needed. Accurate information on sex relating to about a quarter of a million births are needed before any worth-while conclusions can be drawn.

APPENDIX

Brass-Carrier Stable Population Models: Model Age Distributions Per 10,000 of All Ages

Percentage Under 15 (P)	Percentage Over 45 (Q)	Age Distribution Per 10,000																Rate of Natural Increase (Per Cent Per Annum)	Infant Mortality Rate Per 1,000	Expectation of Life (ae)
		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75+			
35	18	1331	1125	1044	984	904	820	741	664	587	506	420	331	241	157	88	57	0.2	151	31
	19	1334	1128	1038	968	885	801	723	649	575	501	423	343	260	180	110	83	0.6	163	34
	20	1333	1133	1034	954	867	732	704	632	562	493	423	351	277	203	134	120	1.0	167	38
	21	1329	1139	1032	941	850	754	686	614	546	482	418	354	289	224	159	173	1.4	153	45
36	17	1376	1157	1067	1000	913	822	737	655	574	490	402	313	224	143	79	48	0.3	145	30
	18	1380	1160	1061	983	893	802	719	640	564	486	407	326	244	166	99	72	0.7	159	33
	19	1380	1164	1056	968	874	783	700	624	551	479	407	335	261	189	122	107	1.1	166	38
	20	1377	1170	1053	955	857	755	682	606	536	469	404	339	274	210	147	156	1.5	157	44
37	17	1425	1191	1083	998	900	803	715	632	551	472	390	308	227	152	89	62	0.8	152	33
	18	1427	1195	1078	982	881	784	697	616	540	466	392	319	245	175	111	93	1.2	162	37
	19	1425	1201	1074	968	863	766	678	599	526	457	390	324	260	196	135	138	1.6	158	43
38	16	1417	1220	1103	1012	908	805	711	624	540	457	374	291	211	138	78	52	0.8	158	32
	17	1472	1227	1100	997	888	785	693	608	529	452	377	303	230	160	100	79	1.3	154	37
	18	1472	1233	1096	982	870	766	674	592	515	444	376	310	245	182	123	121	1.7	148	43
	19	1466	1240	1094	969	853	748	658	574	500	433	370	312	255	201	148	182	2.1	131	52
39	15	1527	1249	1124	1027	916	806	707	615	528	442	357	273	194	123	68	42	0.8	158	30
	16	1525	1255	1119	1010	896	786	689	601	518	439	362	286	213	146	88	67	1.3	163	35
	17	1518	1265	1118	996	877	767	671	584	505	432	362	295	230	168	111	103	1.8	147	43
	18	1523	1266	1111	981	859	749	653	568	491	422	358	298	241	186	135	160	2.1	170	47
40	14	1570	1282	1148	1044	925	808	703	606	515	425	339	255	177	110	58	34	0.9	142	31
	15	1577	1284	1139	1024	903	788	685	593	507	425	346	270	197	131	77	55	1.3	167	33
	16	1572	1292	1136	1009	883	768	667	577	495	420	348	280	214	153	99	87	1.8	162	40
	17	1570	1298	1132	994	865	749	649	561	482	411	345	285	227	173	122	137	2.2	164	46
41	14	1628	1313	1159	1039	910	789	681	585	496	411	330	253	180	117	66	44	1.3	164	32
	15	1617	1325	1159	1023	890	769	663	570	485	407	334	264	199	139	87	71	1.9	146	40
	16	1616	1330	1154	1007	871	750	645	554	472	399	333	271	213	159	110	114	2.3	148	46
42	13	1677	1342	1181	1054	918	790	677	576	484	396	313	235	164	102	55	34	1.3	156	31
	14	1669	1353	1178	1036	897	770	660	562	475	394	319	249	183	124	75	57	1.9	149	38
	15	1671	1358	1172	1020	877	751	642	547	463	388	320	257	199	145	97	94	2.3	163	43
43	12	1724	1373	1203	1070	927	792	673	568	471	381	296	217	147	88	46	26	1.3	145	30
	13	1728	1378	1195	1049	904	771	656	555	464	381	304	232	167	109	63	44	1.8	164	34
	14	1724	1386	1190	1032	883	751	638	540	454	377	307	243	184	130	84	75	2.3	169	40
44	12	1776	1408	1216	1064	912	772	652	547	453	367	288	215	150	95	52	33	1.8	153	33
	13	1776	1414	1210	1045	890	752	635	533	445	365	293	228	168	116	71	58	2.3	166	38
	14	1768	1425	1207	1029	870	733	617	518	433	359	293	235	183	136	94	101	2.8	154	47
45	11	1830	1435	1235	1079	920	774	648	539	441	352	271	198	133	80	42	24	1.7	150	31
	12	1826	1444	1230	1059	897	753	631	526	434	353	279	212	153	101	59	43	2.3	156	37
	13	1822	1453	1225	1042	876	733	613	511	424	348	281	222	169	121	80	78	2.8	158	44
46	11	1883	1470	1247	1073	904	754	627	518	424	339	264	196	136	86	48	31	2.2	158	34
	12	1874	1481	1244	1054	882	734	610	505	415	337	268	208	154	107	67	58	2.8	152	42
47	10	1936	1497	1267	1088	912	755	623	510	412	325	247	178	119	72	37	22	2.1	150	31
	11	1924	1511	1265	1068	889	735	606	498	405	325	255	193	139	92	55	41	2.8	139	40
	12	1930	1514	1256	1049	868	715	589	484	395	321	257	202	153	111	75	80	3.2	173	46
48	10	1990	1532	1279	1080	896	736	602	491	395	312	240	177	122	77	43	29	2.6	158	35
	11	1982	1544	1275	1061	874	715	585	477	387	311	245	188	139	97	62	57	3.2	158	44
49	9	2050	1555	1295	1095	904	737	598	483	384	298	224	159	105	63	32	19	2.4	158	31
	10	2041	1568	1291	1073	880	716	582	471	378	299	232	174	124	83	49	39	3.1	161	39
	11	2038	1577	1285	1055	858	696	565	457	369	295	234	182	138	101	70	80	3.6	182	48
50	9	2095	1595	1310	1087	887	717	578	463	368	287	217	158	108	68	37	25	3.0	150	36
	10	2089	1607	1305	1067	864	697	561	451	360	285	222	170	125	87	56	54	3.6	155	46

E/CN.14/CPH/6
ST/STAT/25

IX. CARTOGRAPHY FOR CENSUS PURPOSES

(This document, issued by the Statistical Office of the United Nations in collaboration with the Cartography Section of the Resources and Transport Division of the Department of Economic and Social Affairs, was prepared by Marvin F. Gordon, Associate Professor, Department of Geography and Regional Science, George Washington University, United States of America, as Consultant to the Secretariat.)

CARTOGRAPHY FOR CENSUS PURPOSES

CONTENTS

		Paragraphs
I.	Need for Maps in census operations	1 - 4
II.	Functions of an office of cartography and geography	5 - 19
	A. Work assignments and objectives	5 - 13
	B. Staffing requirements and problems	14 - 19
III.	Census map requirements	20 - 41
	A. General considerations	20 - 31
	B. Characteristics of enumerator's maps	32 - 36
	C. Needs of census offices	37 - 41
IV.	Map acquisition	42 - 43
V.	Operational steps in a mapping program	44 - 53
	A. Inventory	44 - 48
	B. Base map preparation	49 - 51
	C. Field work program	52 - 53
VI.	Map and graphics publication	54 - 59

ANNEX I Glossary of terms

ANNEX II Statistical areas

ANNEX III Organization of a geography office

ANNEX IV Table of organization, cartography and geography office

ANNEX V Bibliography

I. NEED FOR MAPS IN CENSUS OPERATIONS

1. A census of population and housing must have a spatial framework for which data are to be gathered, tabulated and reported. Areas must therefore be delineated for this purpose; it is the task of the map maker to provide a tool of high quality which will facilitate this kind of work. Accordingly, a strong case can be made for the compilation and use of maps which are current and precise, since they can render certain tasks easier and ensure an order of census accuracy not normally possible without them. Maps are used to establish the location of many types of boundaries, to prevent omission or duplication of information, to aid enumerator canvassing, to facilitate coordination between various offices and between office and field staff, to settle administrative jurisdictional problems, to identify segments used in area sampling, to help determine the number and distribution of enumerators and supervisors to be employed and to foster comparability of data from census to census. Maps furnish some indications of the speed with which the enumeration can take place; they help provide a framework in which to measure internal migration and the rate of urbanization. Maps gathered during initial phases of the work also may be helpful in providing background information which can then be used in planning and operational activities. They can, for example, indicate the capabilities of the transportation network, show settlement types and location of housing or reveal land use patterns. The organization encharged with the responsibility of preparing maps and establishing area controls can also lay out statistical areas, develop maps for publication and perform a variety of post-census tasks.

2. It is for the foregoing reasons that statistical gathering agencies are becoming increasingly aware of the advantages accruing to the use of maps of high quality in all types of censuses. Accordingly, numerous reports and meetings have called attention to the need for maps and

accurately defined areas.[1] It need hardly be added that the time required to do cartographic work for any upcoming censuses in 1970 is growing short.. Some of the most troublesome mapping problems must be faced by many emerging nations, whose areal coverage is often poor, whose finances are limited and whose staff may be too small or lack the specific skills needed for the job. Nevertheless, the situation shows marked signs of improvement, since sizeable portions of Latin America, Africa and Asia have been, or are in the process of being photographed and mapped.

3. Whereas it is the objective of this paper to describe cartographic operations for population and housing censuses, it should be pointed out that similar tasks and objectives may often be involved in other kinds of enumerations, such as those dealing with agriculture or certain economic activities. Inter-censal sample surveys also require maps, usually of a high order of accuracy.

4. This paper will concern itself with the cartographic and geographic aspects of census taking. It will cover the following general subjects: the functions of a geography office; census map requirements; map acquisition; operational steps in a mapping program; and map and graphics publication. In Appendix A will be found a Glossary of some of the terms used in the text of this paper.

II. FUNCTIONS OF AN OFFICE OF CARTOGRAPHY AND GEOGRAPHY

A. Work assignments and objectives

5. A geography office given the job of supporting a national statistical office would have a considerable number of tasks assigned to it. In general terms, it would have two objectives: to delimeate boundaries and

---

1/ Secretaria Permanente del Tratado General de Integración Económica Centroamericana. Primer Seminario de Cartografía para Estadística, Panama, 1964 and Segundo Seminario de Cartografía para Estadística. Guatemala, 1966. Report of the Expert Group on Problems of Internal Migration and Urbanization (United Nations document SA/Dem./EGIM/L.22), 1967.

supply maps of those areas for which the census is to report data and to aid in making data more reliable. In more specific terms, its work can be described under three different categories, dealing with development of maps and determination of boundaries, the establishment of statistical areas and the publication of maps.

6. The first job is to prepare maps for use in censuses of varying types and frequency. These maps would be of many different kinds, depending upon the needs of the user. Those destined for regional offices, for example, would require less detail than those needed by local field units. Maps for rural areas would differ in a number of ways from those required for urban centers; those used for sampling would have to be more accurate than average. Preparation of these maps implicitly assumes that the geography office will prepare a complete listing of all places for which data are to be reported, by means of a geographic code. The listing would, of course, be developed in accordance with the decisions and specifications of the statistical authorities involved.

7. It is also necessary for the geography office to ascertain the location of the boundaries of all places for which the census is to report information. This is often a difficult chore, since administrative boundaries may be poorly defined. Honduras distritos, Panama corregimientos, Vietnam communes and U.S.A. magisterial district (Kentucky) boundaries can all be cited as past cases in point. Indeed, clarification of administrative lines may often be considered a side benefit which stems from census work. However, the statistical agency can normally do no more than point to boundary discrepancies; decisions to rectify the situation are normally not within the scope of its delegated authority.

8. A wide range of data may often be sought from censuses in order to formulate programs for specific areas. Such information might be used, for example, in economic development work or in national or urban planning. It would be useful if the places for which data are gathered, regardless of the type of inquiry, were kept the same. This would facilitate the establishment of a statistical series covering a wide range of topics for

the same area. A geography office obviously can and should assist in work of this type by attempting, if possible, to retain the boundaries of areas used in other inquiries. However, there are two problems associated with such a meritorious undertaking, which should be kept in mind. First, each subject matter division planning a census specifies to the geography office the kinds of areas it wishes delineated. Thus, it is only within this restrictive framework that the map maker can try to retain boundaries of enumeration areas which were used for other censuses. Cooperation between the various subject matter offices involved in such cases would, of course, be critical. Second, physical, sociceconomic and political changes in an area may often require that new boundaries be chosen for it. This is especially true in such places as urban centers which may be experiencing rapid growth.

9. Another important function of a geography office is map reproduction. It is imperative that copies be obtained cheaply, quickly and in sufficient quantities to meet the operational demands of the census agency.

10. The delineation of statistical areas, undertaken in accordance with the specifications laid down by the census authorities, may often become another task assigned to a geography office. These areas are usually defined for data gathering purposes and do not therefore correspond to governmental or administrative divisions, which are officially designated areal components of the state.[2/] Most developing countries, however, cannot afford tabulations of both statistical and administrative areas. Furthermore, definition and mapping of the latter is a job of sizeable magnitude and must remain the prime imperative of countries with poor coverage and limited funds. For this reason, the description of statistical areas has been placed at the end of this paper, in Appendix B; it is not here implied that establishment of these types of reporting areas

2/ A. Archer. The Use of Maps in Census Work. U.S. Bureau of the Census. ms., 1958.

is unimportant, but rather that, on a priority basis, there is much less urgency in doing so, in many emerging nations.

11. The preparation of published maps, graphs and charts which help reveal the spatial relationships of statistical data, remains another task of the geography office. If presented in meaningful fashion, such graphics enhance the value and usefulness of the census figures. One further requirement, in this regard, is to prepare maps which show the boundaries of areas and the location of places for which data are being reported. These are fairly simple to construct and inexpensive to publish.

12. Post-census operations involve the updating and qualitative improvement of maps and a continuing program of map acquisition and compiling. Contact is maintained with the government agencies of those places for which data must be reported, so that the geography office can keep abreast of all boundary changes. Field work may be also undertaken to establish statistical areas for the next census. Maps may be prepared for sample surveys and other special enumerations. It might be noted that these chores often are held to a bare minimum as a result of budgetary cuts. The impact of such decisions are felt, however, when an attempt is made to construct new maps within a comparatively short period of time before the enumeration.

13. The geography office must finally consider, as a mandatory task, the establishment of close working relationships with other mapping agencies in its government. It should also maintain active contact with private and international organizations which can help supply it with maps and with information which will be of value in taking a census.

B. Staffing requirements and problems

14. The table of organization of a cartography and geography branch of a statistical agency is shown in Appendixes C and D. The former briefly indicates the functions which might be performed by such an office. Since it is generally of a non-permanent type, and since skilled personnel may be hard to find, its staffing problems, particularly for many emerging

nations, will doubtless be a source of much concern. Necessity may require that all sorts of expedients be attempted to circumvent this difficulty. In the final analysis, however, there can be no substitute for the competent professional employee; the quality of the work will normally suffer as a result of staff inadequacies.[3/] Accordingly, training programs for census geographers and cartographers should be started as soon as is feasible. Appropriate faculty from institutions of higher education might be used profitably in order to teach or upgrade skills of employees.

15. There are generally four kinds of workers whose skills are required in census mapping operations. The first is the draftsman, who needs no special geography training, although such a background is naturally desirable. The second is the cartographer, who must have a good working knowledge of techniques of map compilation, editing, reproduction and publication. Such an individual must be well aware of current technological changes in his field. Since good cartographers are in short supply, emerging nations, in particular, may have to rely on persons who may simply have taken a cartography course or are skilled and able draftsmen. On the job training and refresher courses at institutions of higher learning can often effectively improve the quality of such a neophyte staff. The third type of employee with special skills is the geographer. He should, ideally, have a good background in cartography, quantitative geography, demography and economics. Since demand will probably outstrip supply, it may be necessary to resort to further formal instruction during, or in advance of the census mapping program, in order to satisfy the need for skilled personnel of this type. The last kind of employee, the clerk, can often be trained on the job in order to perform the necessary tasks. Such persons must be able to demonstrate that they can read maps easily.

16. It should be noted that many personnel can, and probably will, of necessity, be shifted from one type of operation to another. Hence,

---

3/ J. Zarur. Geography and Cartography for Census Purposes in Latin America. Inter American Statistical Institute, Washington, D.C., 1947, p. 6.

their training should, preferably, not be restricted to work of a limited scope. Cartographers, for example, may have to do the work of geographers and vice versa; draftsmen, as noted, may have to substitute as cartographers.

17. On the job training might include subjects such as drafting, aerial photography interpretation, boundary delimitation techniques and methods of map reproduction. Acquisition of an adequate, albeit limited, background of this sort is possible for most persons with the equivalent of a secondary school education. However, one difficulty which may arise in this regard is that the personnel qualified to give such instruction may be limited in number and busily engaged in many other activities because of the shortage of competent employees.

18. Map acquisition is another important operation in census mapping and it is assumed that training individuals for such jobs would probably be necessary in many emerging nations. On-the-job training would be time consuming, however.

19. Aside from secondary schools and colleges, other sources for recruiting qualified employees might include organizations having a map facility, surveying agencies, planning departments, construction and transportation firms and public utilities. It may also be possible, under certain special conditions, to make use of personnel in the cartographic branches of the armed forces.

III. CENSUS MAP REQUIREMENTS

A. General considerations

20. Maps may differ widely in terms of the subject covered, the size of area encompassed and the mode of presentation. They can show such varied types of information as topography, weather conditions, ethnic distributions or statistical patterns for large and for small areas. There are maps prepared for many different types of users; these people may be navigators, blind persons, petroleum prospectors or law enforcement authorities. A census organization is no different in this regard. It needs many kinds of maps, depending upon such variables as the type of unit being enumerated,

the kind of area in which the canvass is to take place and the needs of the different offices and employees of the census agency. There is not, then, a standard map which can be used by an enumerator for all censuses or often for the same census. For example, what is an adequate map in a rural area may not be so in a village; what is adequate in a small urban place may not be in a large city. Such considerations, however, do not negate the possibility that many maps could be used interchangeably, for various kinds of censuses, with little or no change.

21. Certain general statements can be made in regard to all census maps. First, distance and accuracy are less important than the portrayal of features which are in proper relationship to each other. Whether a dwelling unit lies fifty or one hundred meters from the road is less important than whether it is placed, on the map, on the correct side of the road. Second, all boundaries which are needed to demarcate separate reporting areas must be shown as clearly and as accurately as possible in relation to other features of the landscape. Third, maps should be simple and show little information which is not germane to the needs of the user, as this will only serve to confuse; most individuals involved in census operations are normally not well versed in map use. Fourth, features and names must be shown clearly, since most maps will be reproduced many times. If copies are blurred, the utility of the map may decrease sharply. Fifth, map distance should be shown by means of a graphic (horizontal) scale, since it then does not have to be altered in case the map is enlarged or reduced in size. Last, it is important that all names of agglomerations and important features be spelled correctly and reflect any recent change; place names must be those which are commonly used by local people.

22. It might also be worth observing that the map paper and the ink used should be of a sort which withstands rough and often protracted use. This is a very practical and important consideration, since poor quality material may impair the usuability of the map and delay operations until it is replaced. Lamination is a protective device which could be used on

those maps which might have to withstand a great deal of abuse. In this process, a thin plastic film is appended to the paper. However, it should not normally be necessary to protect maps to this degree in census work. Regardless of the materials used, however, it would probably be beneficial to include some mention of the need for careful handling of maps during the enumerator's training period.

23. As a general rule, features should be included on a census map if they assist the user in determining his position, if they help to identify the boundaries of a designated area, or if they show the location of the census units being enumerated - in this case, people and housing. Some of these features are itemized below.

24. (a) Physical features. Topography is shown for the reasons noted above. It will often have a direct bearing on the size chosen for an enumeration district. Boundaries at times can be drawn along crest lines if they are distinct barriers to transverse travel and are sparsely populated. It is often hard, however, to connect such lines with features in surrounding low lying areas. Rivers and streams generally are good boundaries, although care should generally be exercised to avoid intermittent streams, if possible. Water bodies and marshes also can serve as convenient lines of demarcation.

25. (b) Cultural (non-physical) features. Highways, roads, streets and trails should be shown. Those which pass through the center of a village are not always the best boundaries, since they normally will then bisect a socioeconomic unit. Quality of the roads is important, since it provides some indication of the ability to sustain vehicular traffic and can also help in estimating the travel time needed for canvassing. This information is also important if the time of enumeration coincides with particularly cold weather or a rainy season. In the event that road names are unavailable, they must be described ("road from town x to town y"). Other features which can be useful, if shown are transmission lines; telephone and telegraph lines; railroads; canals and large, permanent irrigation and drainage ditches.

26. Normally, the best boundary is one which can be seen easily on the ground. Nevertheless, since it is necessary to adhere to administrative area limits, many imaginary lines must be used. In such cases, however, a respondent will often know where he lives in relation to such a boundary. At times, fence lines or even property lines may be used to divide areas; off-set lines (short, imaginary lines connecting visible features) can also be drawn, if necessary, to complete a boundary.

27. If several names are commonly used for one place, they should all be shown on the map, if possible. Since it is important to obtain correct names on census maps, attention should also be called to sources of information in this regard. For example, the gazeteers and other publications of the U.S. Department of the Interior, Board of Geographic Names, and the U.K. Ordnance Survey, Permanent Committee of Geographic Names might be helpful. However, it should be understood that large scale maps, such as those used by enumerators, may require place names which often cannot be obtained from sources such as those mentioned.

28. Since boundaries of places in some countries may be subject to fluctuation, it may be desirable to establish a cut-off date, after which further changes will not be reflected in the census statistics. If this is not done, a great deal of time and effort can be expended in keeping the maps current. One difficulty in using a terminal date, is that it may confuse both enumerator and respondent, since the map will not then represent the true situation. It is therefore more desirable to attempt to pass a law or issue some type of edict which would prevent boundary changes from taking place after a certain date, except under rather unusual circumstances; several countries concerned with such problems (such as Brazil) have taken measures of this type to fix boundaries prior to the census.

29. In addition to the features noted above, census maps must show areas peripheral to the place which has been delineated. The purpose of this requirement is to help the map user to locate boundaries on the ground. If recognizable points outside such lines are shown on a map, the limits of an area can be easily ascertained.

30. Settlement patterns should also be shown. Detailed map information dealing with villages in rural areas can be shown as insets or enlarged and shown on separate sheets. If the location of buildings is not definitely known, it is safer to omit them from the map; users become more confused by erroneous information than by lack of features. However, in such a case, an estimate of the living quarters in the area should be obtained from field personnel.

31. Maps prepared for special purposes usually require special care. Sample surveys taken on an area basis fall into such a category. Since fewer maps are involved, however, this does not present any major problems. In the U.S.A., the enumeration area can often form the frame from which sample segments are obtained. Choice of such districts are up to the subject division, however. Special efforts are usually expended to upgrade the quality of the map once a selection has been made.

B. Characteristics of enumerator's maps

32. There are more enumerator's maps prepared than any other kind. It is important that they be of good quality so that a high order of accuracy in data gathering is maintained. Following are some of the features which must be shown on these maps:

(a) A north arrow, for use as a directional guide.

(b) A graphic scale (see Appendix A, Glossary of Terms). The enumerator is not trained to interpret other representations of scale, such as fractions or word statements (for example, one inch to one mile, or 1:63,360).

(c) An identifying geographic code number which will also indicate those major and minor civil divisions which are used in the census, and in which the enumeration district is located.

(d) A legend which explains all features on the map.

(e) Peripheral area, to help fix boundary locations.

(f) Names of physical and cultural features, printed clearly.

(g) (Roads with double lines). This makes clear that a boundary, running down the middle of the street, includes one side but not the other. If single line roads were used as boundaries, the enumerator might not remember his instructions and could canvass both sides. An alternative, to be followed only when necessary, is to include an explanatory statement on the map.

(h) Boundaries are shown by a colored line, preferably red or orange. Since this must be applied manually, it is time consuming. Aside from the original map in the central office, all other enumerator's maps can show this boundary by means of a black, wiggly line with tight meanders; this reproduces adequately for whatever use it is put by other personnel. While not mandatory, a word description of the boundary is desirable and should be included.

(i) All buildings, as well as other structures occupied as living quarters, should be shown if their location can be fairly accurately ascertained. Thus the map should ideally include detached dwellings, multi-dwelling buildings and non-residential buildings such as schools, industrial plants, commercial buildings or barns, and it may also include other structures or places occupied as living quarters, such as ruins or caves. If possible, symbols should be employed for buildings which contain at least one set of living quarters and special symbols should be used for multi-dwelling buildings and marginal housing units.

33. It is the responsibility of supervisors to number city blocks for canvassing and to field check new construction, areas of recent rapid growth and doubtful boundary lines. Arrangements should be made so that supervisor's training includes practical experience of this sort.

34. Several common mistakes in construction of enumeration maps are listed below. First, the map may be made of paper which deteriorates rapidly with use. Second, it may be of a scale which makes it too large to handle easily in the field. Third, the features and names shown may be too small, and they may therefore be hard to read or subject to misinterpretation. Such maps may often not reproduce well. Fourth, if the peripheral zone does not extend out far enough to cover an identifiable point beyond the area being enumerated, the location of the boundary may not be clear. Fifth, doubtful boundaries, particularly those following imaginary lines, must be field checked; if this is not done, the margin for error is increased. Sixth, single line roads, as noted, enhance the chances for inaccurate enumeration. Finally, if dwelling units and other buildings are incorrectly located on the map, they tend to confuse the interviewer. Under such conditions, the possibilities of omission or duplication of information are enhanced. This situation also applies to the misplacement of other physical and cultural features on the map. It has been pointed out that there is a certain tolerance of error on census maps as regards distance, but not in terms of the proper relationship of one feature with another. It must be stressed that it is the responsibility of the supervisor to check out all major map problems; the enumerator can make simple changes on his map but he must not be relied upon to interpret ambiguous features. For example, he can often determine whether a house is located within an urban center bounded by imaginary lines by asking the respondent; he can add or delete roads, bridges, houses, and other features as long as such change is minor. However, if a boundary has been changed and is not shown on the map, or if a new road complex has been built, or if a great deal of new construction has obliterated or masked the old settlement pattern, then he needs assistance from his supervisor.

35. The size of the enumeration area will vary as a result of a number of different factors. For example, the type of terrain and the ability to traverse it during the season of the year in which the census is taken, will have a bearing on size. The quality of the roads, the kind

of vehicle which might be used for transporting enumerators, and the length of the enumeration period must be taken into consideration. The level of education of the enumerator and the respondents, the length of the questionnaire and the density of the area will also affect size. Finally, rural enumeration areas almost invariably tend to be larger than their urban counterparts. Basically, however, the size of the enumeration area will reflect the desires of the subject matter division conducting the census. Old enumeration areas can be subdivided into acceptable sizes for a sampling survey by means of standard map chunking and segmenting methods.

36. In view of the foregoing paragraphs, it should be fairly easy to determine the extent of map training which enumerators should receive. They should be taught something about scale, orienting maps, canvassing techniques, boundary recognition, legend symbol identification and standard procedures for changing and correcting maps. The latter task might include the addition of new information on the maps as well as the deletion or alternation of incorrectly placed features. Apropos of this, it also should be stressed that all maps, without exception, should be returned to the Geography Office.

C. Needs of census offices

37. The central office requires a map of the entire country, showing major population agglomerations, terrain features and drainage and the transportation network. It must be clear enough so that jurisdictional boundaries of census regions can be outlined, in order to avoid overlap. Copies should normally be sent to regional offices. Simple maps must also be available which show boundaries of administrative areas and the relative location of all places for which data are to be reported and published.

38. Regional offices require full map coverage, including contiguous areas. Maps of intermediate and minor civil divisions, as well as those of supervisors, crew leaders and enumerators are also needed. Originals of all these maps are filed at the central headquarters.

39. District or local offices require coverage of pertinent administrative areas, plus copies of crew leader's and enumerator's maps. Copies of the latter can be used to replace maps in the field which are lost, mutilated or destroyed. Supervisor's maps should show most pertinent physical and cultural features of the area involved, although the detail would not necessarily be as great as that depicted on enumeration district maps.

40. The magnitude of the map reproduction job is indicated by the foregoing paragraphs. For example, a copy of an enumerator's map must be made available to the central, regional and district offices as well as to the enumerator himself. Furthermore, since peripheral areas have to be included, the base map from which a "cut-out" map comes, must in turn be reproduced several times. This is because the base map, which contains a certain number of enumeration districts, could not be cut up to furnish the exact same number of enumerator maps.

41 Scale denotes the relationship between distance on the map and actual distance on the ground. Generally speaking, small scale maps cover comparatively large areas and vice versa. Enumerator's maps are large scale; as a rule, they vary from 1:5,000 to 1:12,000 in urban areas to 1:10,000 to 1:25,000 in rural areas. If the scales are smaller than those indicated, names often are not legible; if larger, they are too big and unwieldy to handle in the field. (The representative fractions noted indicate the actual map and ground distances, regardless of the unit of measure. Thus, 1:5,000 could be defined as one inch to 5,000 inches or one centimeter to 50 m.). As a result of the differences in areal sizes of enumeration areas, the line used to demarcate the end of an urban place and the beginning of a rural region also represents a transition from one scale of map to another.

IV. MAP ACQUISITION

42. Maps can be obtained from a variety of sources. Government agencies, for example, may have maps, sketches, aerial photographs or special information which may be of value to census operations. Highway agencies

can have maps of the road network; departments of interior or conservation may have maps of special areas such as forests, river basins or reservations; an agriculture ministry may have soil and land use maps; a geological survey agency may have topographic sheets, or maps showing locations of mining camps; defense agencies may have coverage of training areas and other installations of an unclassified nature, as well as maps of ports, coastal areas and boundary regions. Maps showing construction housing starts, hydrography or postal routes can also often be available and helpful. Other sources of information may be planning departments, development banks and area development organizations, and health and education agencies. Local governments might furnish maps from cadastral records, city files, or from police, fire, highway, health and sanitation departments. Other sources might include air, rail and trucking firms, construction companies, utilities, engineering organizations, farm cooperatives, libraries, insurance companies and cartographic firms. International agencies and geographic institutes may also have usable maps. The American Geographical society, the International Geographical Union or the Pan American Institute might be cited in this regard. Exchanges with other countries is also possible. As for emerging nations, the general impression seems to be that adequate maps may be harder to obtain. However, this situation may apply only to internal sources; other nations and international agencies may often have good coverage for large areas. Here the skills of an individual well versed in map acquisition are needed. Large portions of Latin America and Africa, for example, have been photographed and mapped.

43. Maps can be obtained by contract or by cooperative effort with other agencies. If private organizations are involved, care must be exercised to ensure that copyright laws or other legal privileges are observed. Often, a promise is made that these maps will not be used for anything but census purposes.

V. OPERATIONAL STEPS IN A MAPPING PROGRAM

A. Inventory

44. The first step in determining the magnitude of the job of preparing maps for a census is to take a complete inventory. This entails the establishment of a list of all areas for which data are to be reported. It is not always a simple task, since many of these areas may not be known. For example, if the urban definition is couched in terms of agglomeration size, then those villages on the lower end of the population spectrum may or may not be included in the final report. In order to ensure coverage, all marginal cases have to be considered as eligible. Armed with this listing, the map collection is then examined in order to determine both quantitative and qualitative coverage. It should be understood that the inventory must be made in close collaboration with the subject division in charge of the census; only such an office can determine the framework in which the enumeration is to take place.

45. The inventory table should have a listing of all administrative or statistical areas for which data are required. This stub is arranged in hierarchical order from major to minor civil divisions. Columns must be included which will furnish pertinent information in regard to scale, dates of coverage, sources and boundaries. In the last column, a qualitative assessment is made of each map and recommendations are then formulated. For example, maps may be rated as satisfactory or they may be noted as requiring up-dating, field work, further compiling or correcting. Areas for which maps are missing are also listed. The entire table should be designed so as to conform to a geographic code which is then to be used in facilitating all data processing operations. The code is established by the geography office in conformity with the requirements of those subject departments concerned with categorizing data by specified type of area.

46. The inventory at this juncture can provide the geography office with some fairly clear indications of the time required, the cost of the operation, the size of the field staff, the equipment needed and the

amount of field work probably entailed. It also can furnish the priority order of work, since the first tasks should normally be concerned with those areas having sizeable populations but poor or non-existent map coverage.

47. Only general statements can be made in regard to the financial aspects of census mapping, since expenditures depend on a large number of variables. However, as a general rule, the cost of the maps needed will range from five to fifteen per cent of the amount set aside for the census operations up to and including the period of enumeration. One of the major cost items would be field work. It accounts, for example, for almost one half of the amount spent in producing expensive U.S.A. quandrangle (topographic) maps from an aerial photo base. Some countries have decreased these costs by relying heavily upon the work of local individuals in supplying necessary information. In addition to these considerations, it can be pointed out that the delineation of statistical areas will also raise the cost of the operation, as will the use of aerial photography.

48. It is difficult to develop an accurate time schedule for mapping operations. Obviously, there will probably be a close relationship between time expended and the magnitude of the job. Batchelet and Archer,4/ two U.S. geographers, maintain that the time needed must depend on the number of employees, the size of the country and the results of the map inventory. It is their contention that a three year period, on an average, is needed to do the work; however, as noted, this estimate is subject to change for many reasons. In any event, allowance of too short a time span for the work enhances the chances that the quality of the final product will deteriorate. As a rule of thumb, an inventory should probably be made two to three years in advance of the enumeration, and map collection should start at least two years before the terminal date. All maps

---

4/ C. Batchelet and A. Archer. Comments on Geography and Cartography for Census Purposes. Inter-American Statistical Institute, IV Session of the Committee on Improvement of National Statistics, 1956, p. 1.

should be available at least three months before the census is to be taken[5] and maps needed for establishing enumeration areas should be available at least six months before the enumeration. It is generally wise to allow more time than appears to be needed if much field work or outside contracting is involved.

B. Base map preparation

49. If the original census map or one subsequently acquired is suitable in terms of scale, mode of presentation of features and age, it is then used as a base map. It may also be used when there is insufficient time to prepare a better map; this is often the case if the size of the job, the skills of the staff and the allotted time are all at variance. Maps may also be serviceable as bases if the changes required are not numerous. New maps are usually constructed if the previous coverage is extremely poor and if new sources of information, such as aerial photographs, are made available. Updated information is obtained in many ways. By way of illustration, it can come from corrections on previous census maps; from surveys taken by various agencies, such as health or education; from new maps which cannot be used as bases; from census field offices; from local officials. In the last census for Uruguay, for example, local authorities were asked to enter the number of living quarters, by street and block, on maps for their locality. This information was then used in delineating enumeration areas. Compilation, tracing, editing and other techniques used to develop base maps are not described here, since they are no different from standard cartographic methods. The end product of the process for field work is a "cut-out" map which can be used for enumeration purposes.

5/ Principles and Recommendations for the 1970 Population Censuses (United Nations publication, Sales no.: 67.XVII.3), para. 53. Principles and Recommendations for the 1970 Housing Censuses (United Nations publication, Sales no.: 67.XVII.4), para. 60.

50. Aerial photographs, if current, show physical and cultural features in their correct perspective and may be of great help in base map preparation. However, their use as maps is limited because of the constraints imposed by time and costs. For example, whether the photos are used as such or whether they are converted into maps, all pertinent names, administrative boundaries and scales have to be indicated. This usually involves enlargement of the photo and costly field work by staff from the central office. Furthermore, if the photos are used directly as enumeration district maps, an extra effort must be expended in training the interviewer to read them. It must be assumed, in such cases, that the enumerator has an adequate education or is capable of absorbing such training and that the additional time and money for this training are available.

51. Some countries prepare census control lists some time before the census enumeration and use them to help determine the size of the enumeration areas. If time allows, it would be advantageous if the enumeration areas could be delineated prior to the listing and finally determined after information is received from local authorities or regional census offices. However, close working arrangements between census and geography personnel would be required; given the limitations on time involved in map preparation, it may not always be feasible to attempt such an undertaking, desirable as it might be. Apropos of this type of work, it might be mentioned that the outer boundaries of large urban centers (Urbanized Areas) in the U.S.A. in 1960 were defined after the census was taken.[6/] This method of boundary determination is unusual however, and requires a good deal of coordination between census offices.

C. Field work program

52. Field work is required under a variety of conditions: if acceptable maps are not available; to obtain locations of administrative boundaries

---

6/ R.C. Klove and A. Archer. "The Development of Urbanized Area Base Maps for the 1960 Census". Surveying and Mapping, Volume XXIII, March, 1963, p. 86.

and names; to establish certain kinds of statistical areas. There are side benefits to this work, since culture can be added to the maps during the course of the operation. Thus, for example, living quarters can be placed on the map and roads can be classified. It should be mentioned, as a general precaution, that a field team should not normally be sent into an area without at least some rudimentary sort of map from which to start its work.

53. It has been stated that some field work can be done by local authorities, such as police, teachers, government officials or soldiers. If the program is carefully conceived and administered, the requisite information can be obtained at comparatively low cost. This was done, for example, in Bolivia and Panama. Often, however, the quality of the end product is very uneven. Furthermore, if aerial photos are used, few local people can be relied upon to have the requisite skills to work with them. Nevertheless, given the limited budgets, particularly of some emerging nations, this type of field work may represent a feasible solution; information can be obtained cheaply and can be, comparatively speaking, fairly accurate. In the U.S.A., it is anticipated that some local officials will be called upon to help in map editing and the preparation of an address coding guide for certain places in 1970.[7]

7/ W.T. Fay. The Geography of the 1970 Census. U.S. Bureau of the Census. ms., 1966.

VI. MAP AND GRAPHICS PUBLICATION

54. One of the required tasks of a geography office in a census organization is to prepare area identification maps for publication. These show the boundaries of areas and the location of places for which data are to appear in print. A location code guide is set up so that a place identified by a grid number in a table can be found easily on an appropriate accompanying map. No other cultural or physical detail need be shown.

55. The publication of data in map and chart form is a fairly inexpensive yet advantageous way to show some of the results of the enumeration. Generally, the viewer can readily comprehend the patterns revealed by these statistical maps and other graphics. The subjects are also not limited solely to population numbers and distribution or settlement types. Data can be shown which deal with vital statistics, growth rates, urban-rural dichotomies, ethnic structure, economically active or inactive populations, migration, education, age and sex, housing conditions and land use. Although most methods should conform to cartographic convention, modifications permit a fairly wide latitude of choice in presenting census information. Charts, for example, can be superimposed on maps; the results can be gratifying if combinations are constructed with a discerning eye. Census information can also be used to make maps in which data are not simply described, but analyzed within a spatial framework. These maps, for example, may show various kinds of migration streams; centers of population gravity or potentials (shown by isopleths, which are lines passing through points having a constant value); market areas; or degree of settlement dispersion.

56. It should be kept in mind that a proper presentation of the geographical distribution of statistical data requires professional staff skills of a high order. Ability to present information in map form is certainly not a capability unique to geographers; accordingly, many hands are often involved in the formulation and final selection of maps and graphics. At times, as a result, some census atlases tend to be overly simple or, on occasion, much too complex.

57. The objective of a map showing statistical information is to reveal distribution patterns which are not as easily grasped in tabular presentation of data. If the map only shows what is already common knowledge and does that in overly simple terms, then it contributes little. By contrast, if it attempts to show too much or if the scale of the published map is not adequate to the task, the map becomes hard to read and loses much of its value.

58. Some maps can be misleading. For example, it would be incorrect to stress area size if population size were implied. Thus, if the area of Kiruna, Sweden, one of the largest cities in the world in terms of acreage, were shown on a map depicting urban centers in the country, the conclusion might be reached that its population exceeded that of Stockholm. Density maps, commonly shown in census atlases, can tend to project erroneous impressions also. The usual (choropleth) map of this type shows population distribution as evenly spread throughout a certain designated area; in actual fact, such an apportionment is more unusual than usual. Furthermore, since simple arithmetic density does not refine attributes of people or type and use of land area, any correlation between this measure and other characteristics of the landscape which may be inferred from the map is not necessarily of a causal nature Dot maps, which show location and density as well, tend to be more accurate in showing distribution.

59. Perhaps the only conclusion which can be reached in regard to map and chart publication is that extreme care must be taken to make sure that the presentations are meaningful. Maps can reveal spatial patterns with relative precision if they are judiciously constructed. It needs to be stressed that emergent nations, particularly, should seriously consider the mapping of census results. Costs are comparatively low and the benefits to the data user can be appreciable.

Glossary of Terms

1. Choropleth Map. A quantitative areal map in which a symbol representing a certain value is applied to a defined place. A wide variety of subjects can be used. For example, population or housing density per square kilometer, sex ratio or percentage urban might be represented on maps of intermediate or minor civil divisions.

2. Graphic Scale. A map scale which depicts distance by means of a line of designated size which is divided into standard lengths. It is also called a horizontal scale or a long line scale.

3. Isopleth. A line on a map passing through points having the same values. It can be used to show a variety of types of distribution, both physical and cultural.

4. Map Culture. Map information relating to man made features, such as roads, houses, bridges, transmission lines or boundaries of administrative areas.

5. Map Scale. Expresses the relationship of distance between features on a map with the actual distance on the earth's surface. It may be represented in graphic form (see Graphic Scale), as a word statement (such as "one inch to one mile",) or as a representative fraction (such as 1:100,000). The latter refers to equivalent areas on the map and on the ground, using the same type of measure. Thus, in the fraction 1:100,000, one inch (or centimeter) measured on the map is equal to 100,000 inches (or centimeters) on the surface of the earth.

6. Quadrangle Maps. Maps of varying scale which show selected physical and cultural features. Rigid standards are set by the developing agency so that a high order of accuracy in cartographic representation is obtained. Maps of this kind generally appear in a series.

Statistical Areas

1. Statistical areas are defined by criteria established by a national statistical office. They are delineated for data gathering purposes and differ from administrative areas. The latter are civil divisions of a country which are defined and controlled by an agency of the government. Most emerging nations have done little work in developing reporting areas of this kind.

2. Statistical areas are usually defined in terms of some sort of cultural or economic homogeneity. Boundaries are very often drawn by a national statistical office in conjunction with the recommendations of local officials or important members of the community or communities involved.

3. There are several different types of statistical areas. For example, Census Divisions, such as those found in Canada, define areas which have similar social characteristics and market functions. They vary in size, being comparable to counterpart intermediate and minor administrative divisions. Economic Areas and Sub-regions, such as those found in the U.S.A., also delineate areas having similar socioeconomic conditions; in this case, countries (intermediate size civil divisions) tend to be the smallest units used. Connurbations are large urban areas, defined in the U.K., whose boundaries are determined on the basis of urban land use. Tracts are small homogeneous portions of cities which are also defined for data gathering purposes.

4. The advantages of having statistical areas are manifold. For example, it should be possible to obtain a more meaningful series of data if they are gathered for homogeneous areas. Thus, urban planners could make good use of separate tabulations for a central business district or for the squatter settlements (bidonvilles, callampas, barriadas, etc.) which are characteristic of many of the larger cities in Africa, Asia, and Latin America. Furthermore, as noted, since boundaries do not readily change for statistical areas, comparable data can be provided from census

to census. This is important mainly in those countries in which boundary changes of administrative areas are common. Lastly, the boundaries of statistical areas are chosen so that they generally follow visible features on the ground. Since these lines are used in delimiting enumeration areas, they can then serve as an aid in making sure that the interviewer finds his boundaries easily.

5. It has been noted that many countries have done little work in defining statistical areas. As a result, urban places are often not clearly bounded, so that data suffer accordingly. This is one of the reasons why information on rural-urban migration tends to be less reliable than it otherwise might be. Work also needs to be done in defining areas in those countries where agglomerations are large but are not classed as urban.

Organization of a Geography Office

1. A geography and cartography office must be flexible in terms of its table of organization. This is particularly true if the office is a small one in which personnel are often switched from job to job. Certain functions must be performed, however, regardless of staff size or functional subdivision. For this reason, the various duties of the geography office, as well as staffing requirements, are categorized under several different types of activities. A simple table of organization is shown in chart form in Appendix D.

2. Cartography operations. This type of work would involve the compiling, editing and reproduction of maps. It would require personnel skilled in preparing and interpreting maps and aerial photographs. Employees would also have to be able to develop maps and graphics for publication, often in the form of a census atlas. Attached to this core of professionals would be a complement of draftsmen whose knowledge of mapping techniques, concepts and skills need not be of a very high order. The job of measuring areas in order to obtain density data could be handled by this group; the work would mainly be of a clerical nature.

3. Administrative areas delimitation. The collection and revision of maps showing officially defined administrative boundaries would be one of the major chores involved in this assignment. Accordingly, communications must be established with all organizations controlling places for which data are needed; if boundaries are altered, or are in question, field work may be required. People doing this work must also keep track of the boundaries of special areas, such as designated forest regions, territories, reservations or large plantations. One of the most important overall tasks however, would be to ensure that there are no omissions or duplications of areas shown on maps of intermediate and minor civil divisions. This operation should be headed by a geographer, but much of the work can be done by a clerical staff; professionals may become involved in questions dealing with the resolution of boundary problems.

4. Enumeration area delineation. The drawing up of boundaries from the base map prepared by the cartography group requires some professional judgement, so that it would generally be advisable to have trained geographers do the work. For example, the following decisions require some exercise of professional skill: the determination as to when old enumeration area lines should be adhered to and when they should not; the ability to know when to depend on map information and the degree of such dependence; the delineation of boundaries so that homogeneous areas may result. These determinations must, of course, accord to the general specifications laid down by the proper census authorities. The tasks involved are usually given to clerks, however, since professionals are normally in short supply at this time. About the best way to maintain control over the situation is to make sure that geographers serve as supervisors and review all work.

5. Statistical areas delineation. This work would involve the determination and testing of criteria for establishing boundaries of statistical areas. Staff members in such operations should be geographers with backgrounds in statistics. Assignments involve field as well as office work.

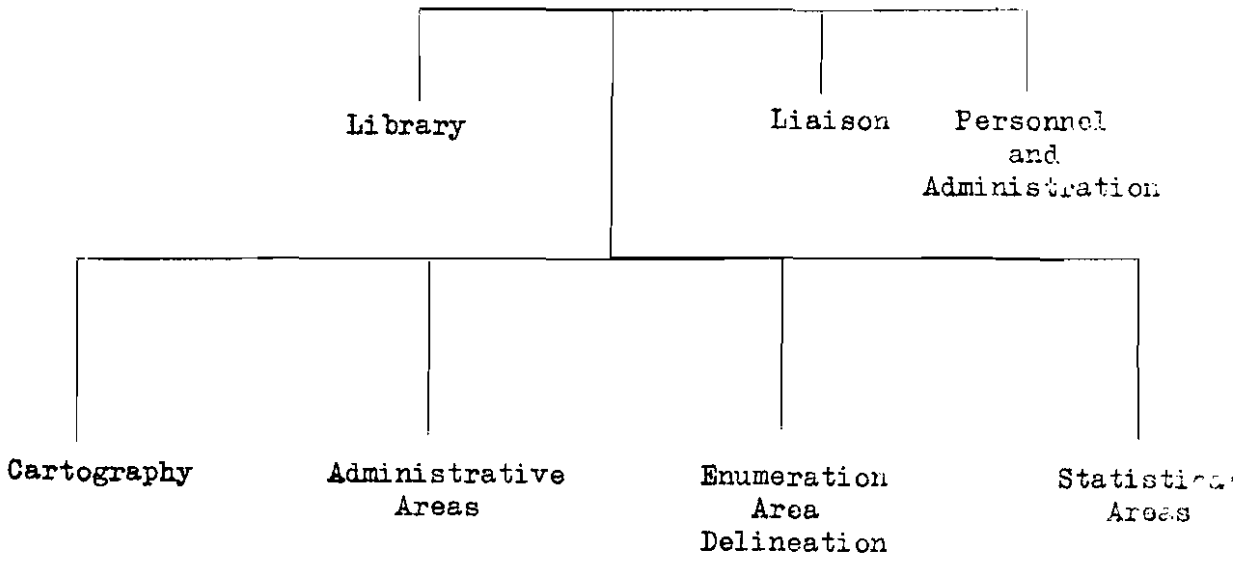
6. Headquarters functions. Some sort of centralized operations must be performed, not only for administration and personnel management, but for a variety of other duties as well. For example it is important to maintain liaison with other organizations and to operate a library in which maps can be filed and stored. It may also be necessary to assist in map training for enumerators and to engage in various types of special field work activities.

7. The equipment for the office is made up of those cartographic tools and machines which are needed to make and reproduce maps. Their complexity ranges from the simple and inexpensive to the complicated and costly. It is not within the scope of this paper to list and describe these machines. Nevertheless, it should be pointed out that, as a minimum, access to cheap and fast duplicating equipment is mandatory. Map filing cabinets are desirable, but low cost shelving can be made to substitute for them if funds are not available.

Cartography and Geography Office

Table of Organization

Headquarters



BIBLIOGRAPHY

Literature dealing with cartography for census purposes is limited. Included in the list below are several general readings which do not deal directly with census geography but which are still good references for anyone planning a map acquisition and development program.

Archer, A. "The Use of Maps in Census Work". ms., U.S. Census Bureau, Washington, D.C., 1958.

Batchelet, C. and Archer, A. "Comments on Geography and Cartography for Census Purposes". Inter-American Statistical Institute, IV Session of the Committee on Improvement of National Statistica, 1956.

Burgdorfer, F. (ed.) World Atlas of Population. Falk, Hamburg, 1954.

Carls, N. How to Read Aerial Photographs for Census Work. U.S. Census Bureau, Washington, D.C., 1947.

Clarke, J. Population Geography. Pergammon Press, London, 1965.

Cooper, S. "The CCD: A Major Revision of the MCD", Professional Geographer, Vol. XV, July, 1963, pp. 4-8.

Fay, William T. "The Geography of the 1970 Census - a Cooperative Effort". Paper presented at Annual Meeting of American Society of Planning Officials, Philadelphia, 1966.

Garnier, B.J. Practical Work in Geography. St. Martin's Press, London, 1963.

Hagood, Margaret J. "Statistical Methods for the Delimitation of Regions applied to Data on Agriculture and Population". Social Forces, Vol. 21, 1943, pp. 287-307.

Jenks, George F. "Generalization in Statistical Mapping", Annals of the Association of American Geographers, Vol. 53, 1963, pp. 15-26.

Klove, Robert C. "Developments in Thematic Mapping at the Census Bureau". Paper presented at Annual Meeting of American Congress of Surveying and Mapping, Washington, D.C., 1965.

Klove, Robert C. and Archer, A. "The Development of Urbanized Area Base Maps for the 1960 Census", Surveying and Mapping, Vol. XXIII, 1963, pp. 79-87.

Prothero, R.M. "Problems of Population Mapping in an Underdeveloped Territory", Nigerian Geographical Journal, Vol. 3, 1960.

Robinson, Arthur H. Elements of Cartography. Wiley and Sons, Inc., New York, 1960.

Secretaria Permanente del Tratado General de Integracion Economica Centroamericana. Primer Seminario de Cartografia para Estadistica. Panama, 1964.

_____. Segundo Seminario de Cartografia para Estadistica. Guatemala, 1966.

Schmid, C.F. and MacConnell, E.H. "Theory of Isopleth Mapping", Journal of the American Statistical Association, Vol. 50, 1955, pp. 220-239.

Ullman, Morris and Klove, Robert C. "The Geographic Area in Regional Economic Research", in Regional Income, National Bureau of Economic Research, Princeton University Press, Princeton, N.J., 1957.

United Nations. Principles and Recommendations for the 1970 Housing Censuses. (United Nations publication, Sales No.: 67.XVII.4).

_____. Principles and Recommendations for the 1970 Population Censuses. (United Nations publication, Sales No.: 64.XVII.3).

_____. Report of the Expert Group on Problems of Internal Migration and Urbanization. (United Nations document SA/Dem./EGIM/L.22), 1967.

_____. Handbook of Population Census Methods. (United Nations publication, Sales No.: 58.XVII.6, vol. 1, 2, 3).

United States, Bureau of the Census, Geography Division. "How to Establish Enumeration Districts", in Instruction Manual, part III, Washington, D.C., 1949, pp. 1-27.

_____. "The Ground You Cover". Training film strips and records numers 4 and 5 for use in the 1960 U.S. Census of Population and Housing. Washington, D.C.

Zarur, Jorge. Geography and Cartography for Census Purposes in Latin America. Inter American Statistical Institute, Washington, D.C., 1947.

Zelinsky, Wilbur. Bibliographic Guide to Population Geography. University of Chicago Press, Chicago, 1962.

X. LOCALITY LISTING AND DELIMITATION OF ENUMERATION AREAS

(Prepared by the ECA Secretariat)

LOCALITY LISTING AND DELIMITATION OF ENUMERATION AREAS

CONTENTS

	Paragraphs
Locality listing	1 - 7
The "village approach"	8 - 18
Nomadic populations	19 - 22
The contents of locality lists	23 - 26
The use of enumeration areas in a census	27 - 33
The size of an enumeration area	34 - 38
Delimitation of enumeration areas	39 - 43
Identification of enumeration areas	44 - 49

Locality listing

1. One of the first steps in the preparations for a population census is the compilation of a list of all inhabited places (localities) in the country. The framework for this list is provided by the civil divisions of the country. It is within this frame that the localities have to be listed, separately for each smallest civil division.

2. According to U.N. recommendations, a locality, for census purposes "should be defined as a distinct population cluster (also designated as inhabited place, populated centre, settlement, etc.) of which the inhabitants live in neighbouring buildings and which has a name or a locally recognized status." (Para. 232 in Principles and Recommendations for the 1970 Population Censuses). Thus every hamlet, camp, farm, etc.is a separate locality. Likewise a cattle-post where some members of the household, separated from their family, are herding cattle, as well as huts on cultivated lands far from the main village which are occupied during the sowing and harvesting seasons, should be considered localities if inhabited during the time of the census.

3. In some African countries the population is so heavily agglomerated and the administration so developed that each locality as a rule forms its own civil division and thus is a recognised administrative entity on a permanent location and with more or less well defined boundaries. Such situations are met particularly in many parts of North and West Africa. In practically every country at least some larger localities form their own civil divisions.

4. In contrast to this are the situations where the population is widely scattered and the smallest units which can be properly called civil divisions include a large number (tens and even hundreds) of villages, hamlets and isolated homesteads. Such cases are very common in Africa.

5. The settlements in many African countries are also very changeable. There is a constant process of their breaking up and joining together, isolated farms or satellite hamlets being built and these latter acquiring an independent status. There are seasonally inhabited places even among the settled agricultural population. Entire villages may be abandoned while the inhabitants move to another place. In such countries and regions, understandably, the villages have no geographically determined boundaries.

6. Such fluidity of localities makes their listing both difficult and very important. It also follows that any existing lists soon become obsolete and have to be brought up to date shortly before the census.

7. Ideally, all localities should be listed, marked on a map and their approximate populations estimated in order to prepare for the next stage of census ground-work, the delimitation of Enumeration Areas.

The "village approach"

8. The above-quoted para. 232 of Principles and Recommendations foresees the possibility that some countries will depart from the recommended definition of locality.

9. In many parts of Africa, indeed, situations exist where circumstances make it impracticable or where it would be unduly expensive and time-consuming to apply the recommended definition of locality because it would require extensive field investigation by a large number of properly trained staff. The benefit to be derived may also be small. Though locality statistics give interesting information on settlement patterns they may be of little practical importance in view of the rather volatile nature of the smallest localities.

10. There is another unit which in the African society has more importance than the locality. Its name may vary but it can be generally understood as a village. In order to define what we mean by a village (in itself a somewhat vague word) we might describe it as an inhabited rural place or area which has a headman who is directly responsible to the administrator of the smallest civil division. If there are other, minor headmen placed under a village headman, we can say that the village includes a corresponding number of sub-units.

11. Sometimes a village may be identical with the smallest civil division and sometimes identical with a locality. A more common case in African societies, however, is that the village is part of the smallest civil division and comprises several localities. Also its sub-units, if any, may be composed of several localities.

12. On the other hand, two or more villages or parts of them may have grown to form one single agglomeration, i.e., locality. A further complication arises when elements of two villages dovetail or intermingle in the terrain. The concept of a village as determined by headmanship is therefore different from the concept of locality if strictly applied.

13. To mention a few examples from tropical Africa where the smallest civil division usually contains a large number of villages and localities, such are the Meketil Woredas and, after their gradual abolition, the Woredas in Ethiopia, the Wards in Zambia, the Cantons in Niger, the Circonscriptions in Togo, the Local Councils in Southern Nigeria, the Native Authorities in Northern Nigeria, the Parishes in Uganda.

14. An imaginary example of this type of case is given in the attached sketch map. It is not given as a model of a map that should be prepared for a census but only as an illustration of the relationships between a civil division, villages, sub-units and localities. As a matter of fact, maps giving similar information are seldom available and it may not be possible to prepare such for the census. It should also be noted that village limits would be only roughly drawn so as to group together

all parts of the same village. It is characteristic of these cases that the geographical limits of the smallest civil divisions are determined and fixed but those of the villages are not.

15. While for some countries the strict application of the locality concept is suitable, there are others where the village approach is preferable. Three reasons for this are prevalent in Africa.

16. First, the village is a living entity in African life and development. Both administrative practice and the people themselves think and act in terms of the village. Population data by village are therefore needed.

17. Secondly, no census-type operation in rural Africa is conceivable without the knowledge of the village headman. It is desirable, natural and often quite necessary that the village headman accompany the enumerator in his rounds introducing him to the villagers and ensuring their co-operation. It is therefore advantageous or even necessary to build the census organization into the village system.

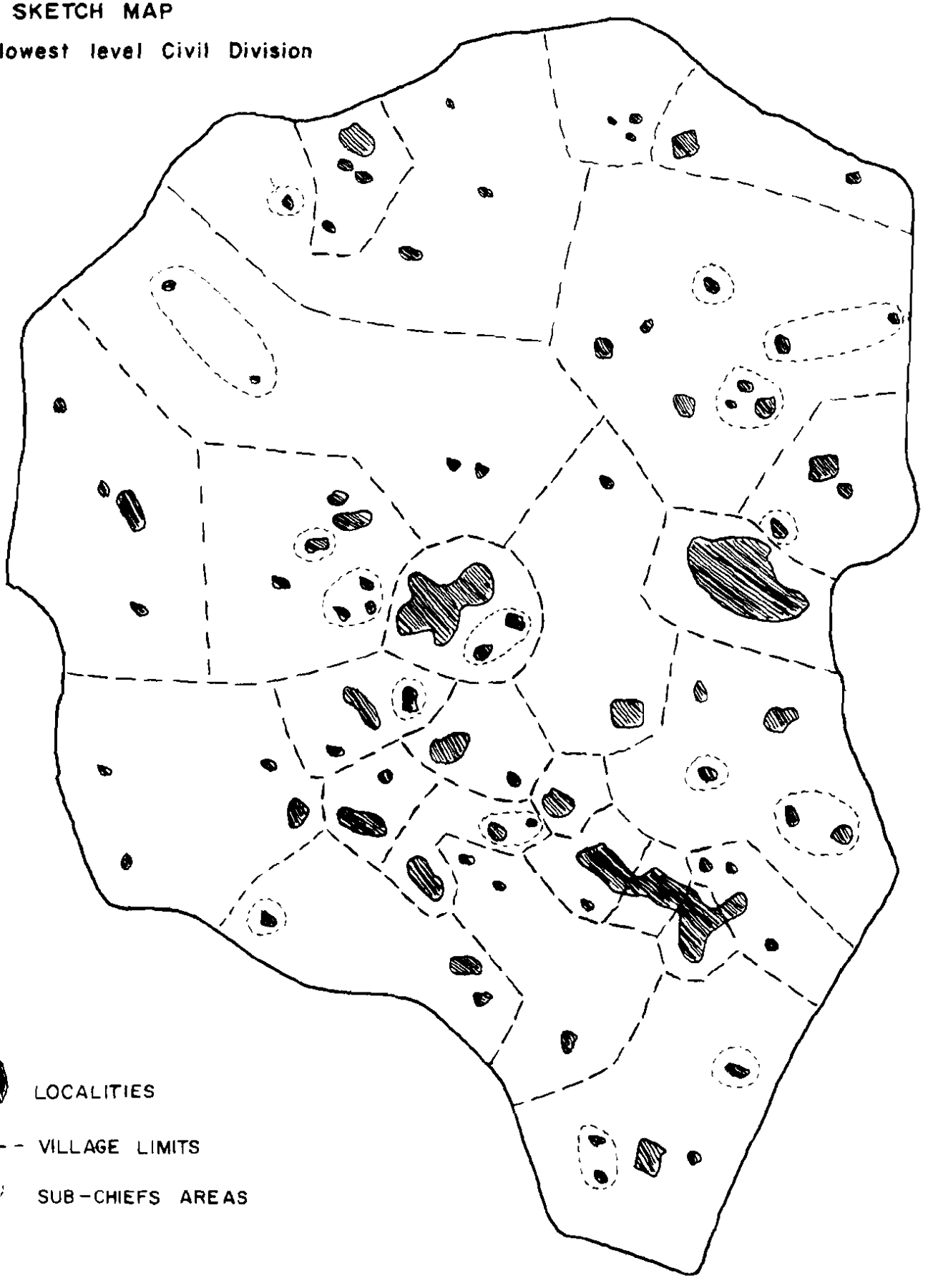
18. Thirdly, in areas for which accurate maps showing all inhabited places cannot be prepared or where the enumerators are not able to use them, the village approach offers better assurance against omission of scattered localities. Omissions and inaccuracies easily occur in the preparation of lists and the geographical locations may be seriously in doubt. The headman is the person best placed to help to prepare a sketch map or a description of the village or a list of its component parts if such are required. He is also the man to guide the enumerator to all part of his village.

Nomadic populations

19. Since the nomads do not live in permanent settlements and do not form localities, an alternative to locality listing has to be found in order to prepare for their enumeration.

SKETCH MAP

Of a lowest level Civil Division



20. In regions where water is very scarce a census of the nomads could be attempted at the height of the dry season by sending enumerators to all places where water is found even at that time. In such case a listing and mapping of all such water sources (wells, bore-holes, water holes, tanks etc.) is required. This would thus be only a variation of locality listing.

21. More often, however, a better prospect of success is offered by a hierarchical approach which uses the traditional structure of society. This is normally based on kinship. Pastoral nomads usually have very strong clan and family ties and the authority of the chiefs is firmly established. The different groups and sub-groups as well as the names of their chiefs should therefore be listed.

22. Among some primitive hunter and gatherer peoples (e.g. pygmies and bushmen) the institution of chiefs may not be clearly developed; also the administration may have little or no contact with these peoples. Such groups may be identified and listed by their clan name, by language or dialect, by the area where they live or even by some cultural characteristics. Though these populations can rarely be properly enumerated, a listing of groups will facilitate the estimation of their numbers.

The contents of locality lists

23. A locality list should carry the following information:
- the official or accepted "census name" of the locality;
- known alternative names and spellings;
- name of headman if any;
- approximate population (or number of houses, households or compounds);
- map co-ordinates if available.

24. If the "village approach" is used or if both the villages and the localities are listed, there will probably be units of different levels. In such cases, naturally, under each primary unit (e.g. village) its component sub-units (hamlets, farms etc.) should be listed. If separate enumeration of the sub-units is wanted, individual information on them (headman, population, etc.) is desirable. If this is not wanted, it will suffice to annotate the village list with the names of its component parts.

25. Should it not be possible to prepare maps showing the localities or villages, it would be useful to prepare descriptions of them or instructions how to reach them.

26. When nomads are to be enumerated using the administrative (hierarchical) approach, it will be necessary to record information on their seasonal movements and, if possible, supplement it with sketch maps.

The use of Enumeration Areas in a census.

27. The main purpose of the division of the country into Enumeration Areas is to split the existing local units into census units of manageable size and to assign one enumerator to each such unit (E.A.). This will facilitate the enumeration, the supervision and possible post-enumeration checking.

28. Another purpose is to provide a suitable area sampling unit, whether for a "built-in" census sample, for a post-enumeration check or for any subsequent survey which may use the census as a frame.

29. A necessary requirement for E.A.s is that they should never cut across the limits of civil divisions or such smaller units for which census data are required. If such units are too small to occupy one enumerator, two or more units may be assigned to the same enumerator who will, however, enumerate them separately under separate code numbers.

30. As already mentioned it is customary to assign one enumerator to each E.A. This is economical use of the census manpower and it minimizes the likelihood of omissions and double counts.

31. In some countries enumerators prefer to work in pairs. This may be a habit acquired in the course of sample surveys. It is obviously bad economy in a census. When enumerators work in an unfamiliar or insecure area this may be justified as an added protection against being lost or against robbers or wild animals and it may help to overcome unco-operative attitudes of respondents. However, the provision of an escort is in such cases a better solution.

32. Occasionally entire teams of enumerators have worked together in each area. In the Northern Rhodesia 1963 census the EA.s had usually 5,000 to 8,000 and sometimes over 10,000 inhabitants and often covered huge areas. They were canvassed by a team of 6 to 12 enumerators working their way from one end of the E.A. to the other. After the census it was concluded that this was not a practical or efficient arrangement.

33. In the Bechuanaland 1964 census mobile census teams moved from village to village. Smaller villages were enumerated by one team (4 to 6 enumerators) or part of a team while in larger villages two or more teams operated simultaneously. In such cases the village was divided either beforehand or ad hoc into E.A.s and in each of them the team leader assigned houses to his team members. This method was necessitated by the shortage of suitable field staff and by transportation difficulties and is not to be preferred when division into small, one-enumerator E.A.s is possible.

The size of an Enumeration Area.

34. At the present stage of development the size of an E.A. generally suitable in African conditions is about 500 people which equals about 100 households. Using a questionnaire of 10 to 20 questions and possibly also a succinct housing questionnaire, one enumerator can usually well enumerate such an area in one week. The rate of enumeration is not

necessarily faster in towns because the saving in walking time is offset by greater diversity in entries, e.g. on education and economic characteristics, which therefore take more time.

35. A smaller E.A size would speed up the enumeration and thus make it more nearly simultaneous. On the other hand it would increase the cost of the census by complicating organizational, training and other aspects. It might also lead to recruitment of less qualified staff and so have an adverse effect on the quality of the collected data. It appears that a target E.A size smaller than 500 persons is not generally recommendable.

36. In many countries it is hardly possible to find one suitable and available enumerator for every 500 persons. Such lack of suitable personnel may occur in some remote areas of almost any African country. This may be remedied by transfer of staff to such areas from elsewhere but this is not always practicable.

37. If as many as 1,000 or more persons have to be assigned to each enumerator, it may be advisable to keep the E.A. size smaller and to assign more than one E.A. to each enumerator. Otherwise subsequent sampling, if any is planned, could become less efficient.

38. While a certain target size for an E.A. is selected and maintained, it may exceptionally be necessary to form much smaller E.A.s in places where the settlements are widely scattered and the communications particularly poor. Moreover, considerable flexibility as to size, both downward and upward, should be allowed in order to satisfy an important requirement: easy identification of the E.A.'s in the terrain.

Delimitation of Enumeration Areas

39. After the localities or other local units have been listed, their approximate populations estimated and the target size of the E.A.s decided upon, the next step is to determine how many E.A.s each local unit shall have. This may be done tentatively on a rather high level but the finalization of the number of E.A.s is better left to those

who actually delimit them. These could be middle-level census staff: officials in charge of smallest or next-smallest civil divisions or their assistants or field supervisors but not the immediate supervisors of the enumerators.

40. Whenever a local unit has too large a population to form one single E.A., it shall be split into an appropriate number of E.A.s. While the E.A.s of one local unit should be approximately equal in size, it is still more important that they are easy to delimit and to identify. In towns which have regular streets, the limits should as far as possible follow these so that an E.A. would consist of one or more entire city blocks. In villages and more omplicated parts of towns the limits should follow roads, tracks, open spaces, rivers and other clear landmarks.

41. Mapping of E.A.s, when there are more than one in a local unit, is strongly recommended. For many towns and some villages it may be possible to obtain aerial photographs on which the E.A.s can be shown. When these are not available, sketch maps can be drawn showing relevant landmarks and the E.A. limits in relation to them but not necessarily individual houses. In very simple cases a description may make even a sketch map superfluous (e.g. E.A.1: North of the road. E.A.2: South of the road).

42. The procedures described above are best suited to densely populated areas and are not always practicable in large, sparsely inhabited civil divisions with scattered settlements. In such cases (similar to that shown in the attached sketch map) it may not be feasible to delimit the E.A.s either on maps or on the ground. It would then only be possible to carry out what could be called conceptual delimitation. This will simply mean indicating which sub-units of a village form each E.A. An example is provided by the delimitation of E.A.s in the 1967 census of Tanzania by reference to the TANO party's"ten-house chairmen."

43. When nomads are enumerated using the administrative approach, the E.A.s have no geographical base at all and are purely conceptual. Groups or sub-groups are assigned to individual enumerators on the basis of the information listed.

Indentification of Enumeration Areas

44. When the E.A.s have been delimited, it must be possible for the supervisors and enumerators to identify them. Identification must also be possible for higher supervisory staff during the census, for those who make a post-enumeration survey and for the staff of any subsequent census-based sample survey.

45. The role of mapping in this respect is of primary importance. However, the mapping may not be complete and even if it is, it does not solve the whole problem. It cannot be expected that all field staff can make use of maps. Depending on the type of settlement and on the quality of the supervisors, these latter may or may not be able to identify the E.A.s marked on maps. Consequently, they may or may not need to be shown them on the ground by their superior officers.

46. As to the enumerators, the strict rule should be followed that each supervisor has to show to every enumerator of his team the limits of his E.A. on the ground. This applies of course only when the E.A.s are geographically and not merely conceptually delimited. Where feasible, the supervisor and the enumerator should actually walk together around the E.A. When this is done, a map in the hand of the enumerator usually becomes superfluous and perhaps confusing.

47. If a pre-enumeration house-numbering is carried out, it will facilitate the identification of the E.A.s. The house numbers should be marked on the E.A. lists and, if necessary, on maps. Each enumerator will be given the house numbers belonging to his E.A.

48. When the houses (buildings or living quarters or both) are numbered during the enumeration and the numbers visibly attached or painted at the entrances, this will, as the work proceeds, help enumerators to avoid straying into neighbouring areas. It will also help to check the completeness of the census and assist subsequent indentification.

49. When the "village approach" is used and a local unit needs to be divided between two or more enumerators, these will have to work in very close contact with each other and with the village headman to ensure that no inhabited places are left out in the process of dividing the work because systematic canvassing of large areas is not always practicable.

E/CN.14/CPH/5
ST/STAT/22

XI. PREPARATION AND USE OF CENSUS CONTROL LISTS

(Prepared by the
Statistical Office of the United Nations)

PREPARATION AND USE OF CENSUS CONTROL LISTS

CONTENTS

	Paragraphs
INTRODUCTION	1 - 3
PART I. USES OF THE LISTS	4 - 29
A. To assist in achieving complete coverage in carrying out censuses of population and housing	5 - 7
B. To record and control the movement of questionnaires	8 - 9
C. As a control or frame for statistical inquiries other than censuses of population and housing	10 - 13
D. For planning the census operations	14 - 16
(1) Estimates of the time required for the enumeration, determination of the size of enumeration districts and the number of enumerators required	15
(2) Estimates of census equipment required	16
E. For the collection of statistical data on population, housing and other topics	17 - 26
(1) Housing data	20 - 22
(2) Population data	23 - 24
(3) Other data	25 - 26
F. For advance census results	27 - 29
PART II. COVERAGE, FORMAT, BY WHOM PREPARED AND WHEN PREPARED	30 - 37
A. Coverage of the lists	30
B. Format of the lists	31 - 32
C. By whom prepared	33 - 35
D. When prepared	36 - 37

		Paragraphs
PART III.	CONTENT OF THE LISTS	38 - 42
PART IV.	METHOD OF PREPARATION	43 - 59
	A. Identification of the units to be recorded	46 - 54
	B. Measures to facilitate re-identification.	55 - 57
	C. Recording of the units identified	58 - 59
PART V.	HANDLING OF COMPLETED LISTS	60
ANNEX	Country experience in the preparation and use of census control lists	
Table 1	Timing and use of census control lists	
Table 2	Topics included in census control lists	
Table 3	Topics arranged in rank order by frequency of inclusion in census control lists.	

INTRODUCTION

1. During censuses of population and housing, irrespective of whether they are carried out separately or as a combined operation, all places which could conceivably be occupied by people must be investigated. Lists of these places, prepared prior to, or during the enumeration, serve as a control to ensure that all living quarters in the territory concerned, and/or their occupants, are included in the censuses. Because inadequacies in the control lists thus compiled may adversely affect the entire census operation, the function, content and method of preparation of these lists should be carefully considered in planning a census. Where, as sometimes occurs, the list is to serve as a basis for other inquiries in addition to censuses of population and housing, the consequences of a poorly prepared list are multiplied and careful preparation is even more vitally important.

2. Ninety-one countries out of 128 for which census material was examined, compiled a control list as part of the census operation.* In practice, the control lists have been given various titles, Liste de contrôle. Hoja de Control de Empadronamienta. House list. Enumeration Record Book, Visitation Record etc. The term "Census control list" is used here because it describes the principal function involved but, in fact, the lists may serve several important purposes in the course of the planning and execution of population and housing censuses.

3. No predetermined role or standard format for the lists can be recommended which would be applicable under all circumstances; however, there are certain factors which it will be useful to consider in deciding how the control lists are to be prepared and the function which they are to perform during the census operation. These factors include (i) the possible uses of the list, (ii) coverage, format, by whom prepared, when prepared, (iii) content, (iv) method of preparation, and (v) the handling of completed lists. Each of these factors will be discussed below. A record of country experience in the preparation and use of census control lists is shown in tabulated form in an Annex.

---

* Some of the remaining 37 countries may also have utilized control lists but this could not be established from the material available.

I. USES OF THE LISTS

4. An examination of the census control lists prepared in 91 countries (see Table 1 of the Annex) indicates that the lists may be used inter alia, to assist in achieving complete coverage in carrying out censuses of population and housing, as a control or frame for other statistical inquiries, for certain aspects of census planning, for the collection of statistical data and as a means of providing advance census results. Which of these purposes the list is to serve in any particular census must be decided in the country concerned; however, its function as a factor to ensure complete coverage in the census, represents without question a most important aspect of the census operation. This and its other possible uses are discussed in the following paragraphs.

A. To assist in achieving complete coverage in carrying out censuses of population and housing

5. Census control lists are used by supervisors to assign and check the work of enumerators and by the enumerators themselves to check the completeness of their own enumeration.

6. In compiling a census control list the basic units recorded may be buildings, structures or places occupied as living quarters, living quarters and/or households or compounds (concessions).[1/] The list is compiled by seeking out whichever of the units is selected as the control and recording it in the listing with sufficient identifying information for it to be readily located before, during the course of, and after the census enumeration.

7. If, as usually occurs, the censuses of population and housing are being carried out as a combined operation, the identification and recording of living quarters serves a dual function; it ensures that all living quarters are included in the housing census and at the same time it tends to provide for a complete enumeration of the population living in them. If a population census is being carried out independently of the housing

---

1/ For information on the topics which have been included in preparing census control listings reference may be made to Table 2 of the Annex.

census, the identification and recording of structures or living quarters, or even compounds, would still be a prerequisite for the systematic enumeration of the population.

B. To record and control the movement of questionnaires

8. Little would be gained by ensuring that all living quarters and persons are located during the census operation if the control of questionnaires is not carefully carried out. In the case of self-enumeration, census control lists provide a convenient means of recording the fact that a blank questionnaire has been issued to a particular household or set of living quarters and that the completed questionnaire has subsequently been collected or received. Where the canvasser method is employed, the list may be used to record the completion of each questionnaire by the enumerator. The control may be stablished either by recording the number of questionnaires delivered to each set of living quarters or to each household or by assigning to each questionnaire a serial number which is recorded in the listing.

9. Irrespective of the method of enumeration employed, the need for call-backs will sometimes arise and the control list may be used to record the living quarters or households that need to be revisited because the occupants were absent or because of some other difficulty, the dates of subsequent visits and whether the enumeration was finally completed.

C. As a control or frame for statistical inquiries other than censuses of population and housing

10. Although the listing is usually prepared to serve primarily as a control for censuses of population and housing, it may also be used as a convenient means of controlling other censuses and it provides a useful frame for the selection of sample units, either as part of population and housing censuses or for subsequent sample surveys.2/

2/ The use of census control lists to relate the population and housing census to other inquiries in this way is also touched upon in the working paper entitled Co-ordination between housing censuses and population censuses and of these censuses with other statistical inquiries and compilations.

11. Since all structures must be investigated during the preparation of the list, information on the use of the structures or the activities being carried on within them can be easily recorded. Such information can furnish a basis for any number of inquiries; principal among these is the census of industry but a census of manufacturing, schools, hospitals, etc. represent other possibilities. For a future census of industry or manufacturing, information may be recorded on the number and type of establishments located in the various structures recorded and the subsequent inquiry is based on the units thus identified.

12. In some censuses, only certain housing or population topics are collected on the basis of a complete enumeration while others are investigated from a sample of living quarters or households. In this case, the form which is to be used for compiling the control list may be provided with a system (numbers, spacing, bold lines etc.) which automatically identifies those units in the listing which are to be included in the sample.

13. Obviously, the control lists will be equally useful for sample surveys carried out subsequent to the census. Housing surveys or household surveys represent the most usual type of inquiry for which the lists are used but a list of living quarters is sometimes employed for carrying out surveys of other kinds. Living quarters as a frame have the advantage over households in that they are relatively stable and it is usually easier to up-date a list of living quarters than a list of households. In either case, close attention should be paid to up-dating so that all living quarters constructed since the original listing was prepared are included and those that have disappeared (through demolition, destruction, conversion etc.) are removed from it or, in the case of households, so that those that have moved away or dissolved are deleted and newly formed households, or households that have moved into the area since the original listing was compiled, are added.

D. For planning the census operations

14. Information may be obtained from a census control list for planning certain aspects of the census operation such as estimates of the time required for the enumeration, determination of the size of enumeration districts, the number of enumerators required and estimates of the census equipment needed.

(1) Estimates of the time required for the enumeration, determination of the size of enumeration districts and the number of enumerators required

15. The size of enumeration districts and the number of enumerators required will be influenced by the time required by one enumerator to collect information on the census topics for each household and/or each set of living quarters and the dispersion or concentration of these units in the area concerned. The time required for the enumeration of a single household or set of living quarters will be determined in the course of pre-census tests. The compilation of census control lists may be used to provide useful information on the dispersion or concentration of households and living quarters and a general idea of the time which will be required for the total enumeration using a given number of enumerators and available transportation facilities.

(2) Estimates of census equipment required

16. In planning a census it will be necessary, not only to estimate the total quantities required of various forms, cards, pencils, folders etc. but also to estimate the number of these that need to be made available in various parts of the country. A miscalculation could cause delays during the enumeration which might have serious consequences. If prepared sufficiently in advance of the enumeration, estimates of population and living quarters obtained on the basis of census control lists may serve a useful purpose either for estimating national requirements and/or for indicating the quantitative distribution of census materials that should be made to field offices.

E. For the collection of statistical data on population, housing and other topics

17. In its most essential form the census control list is used to identify and record the basic units covered or to be covered by the population and/or housing census. The list may also be used, however, to collect statistical data for the units thus identified or to gather statistical data on other topics.

18. It should be noted that, where data are collected by means of the listing and the listing is compiled prior to the enumeration, the information collected needs to be verified and/or corrected at the time of the enumeration. This process of verification enables changes which may have occurred between the time of a pre-census listing and the date established for the census to be taken into account and it also serves to increase the validity of the data since it could be expected to reveal errors which may have occurred during the initial collection.

19. For the topics which have been included in census control listings reference may be made to Tables 2 and 3.

(1) Housing data

20. Since the basic units of enumeration utilized in a housing census are also the principal control units recorded in the lists, information such as the number of structures, number of living quarters and number of compounds, becomes available almost automatically from the listing. In addition to these, the housing topics for which information is most frequently collected by means of the lists are (1) occupancy (2) number of rooms (3) type of living quarters (4) use of the structure and (5) material of construction of building.

21. The topic concerning the use of structures may be employed to show the kind of buildings in which living quarters are located and it may also be used as a starting point for inquiries related to non-residential uses such as a census of industry or manufacturing (see paragraph 11).

22. The collection of data at the listing stage may serve to reduce the burden on the enumerator and the respondent and may be an especially useful procedure when a large number of topics are included in a combined population and housing census operation.

(2) Population data

23. The control topic most closely related to the population census is the household since the household is a unit of investigation as well as the general framework within which individuals are identified.

24. If the control list is a list of households, the number of households becomes available automatically. If, however, the control list is basically a list of structures, living quarters or compounds, the number of households in each of these must be added as additional information. The other population topics for which information is most frequently collected at the listing stage are: (1) the number of persons, (2) sex, (3) race and (4) nationality of each person.

(3) Other data

25. Statistical data, other than population and housing information, collected during the preparation of the listing refer most frequently to establishments and agriculture. As noted in paragraph 21, the topic on use of the structure is frequently used to indicate whether there is an establishment located within it. Where an establishment is identified, information may also be collected on the name of the owner or operator, the kind of goods or services produced, number of employees etc.

26. The questions on agriculture may identify agricultural holdings and are sometimes directed at collecting information on the name of the owner or operator, number of employees, livestocks, main crop etc.

F. For advance census results

27. From the description of the topics most frequently collected during the preparation of census control lists (paras. 17 to 26), it is clear that the lists would provide a convenient means of producing advance

census results. Advance results on the population side are usually restricted to basic data such as the distribution by geographic area of the total population by sex, on the housing side the information most frequently made available on a priority basis is the number of living quarters by type and number of occupants. In each case, this information could be readily obtained by means of census control lists either prepared during the enumeration or prepared prior to the enumeration and carefully brought up to date during the enumeration.

28. A difficulty which should be kept in mind, however, is that, since advance totals of this type have not been subjected to evaluation, they may contain serious errors. In the 1960 census of Brazil, for example, the first estimate of the size of the total population, based on lists, was 70,967,185. The revised estimate, after deletion of duplication of persons reported as members of the household where they usually lived and also of households where they spent the census night, was 600,000 persons less.

29. Computation of enumeration district totals by enumerators and of accumulated totals by supervisors can begin immediately following the listing so that it does not interfere with or delay the processing of information obtained in the census questionnaires.

II. COVERAGE, FORMAT, BY WHOM PREPARED AND WHEN PREPARED

A. Coverage of the lists

30. The coverage of a control list usually corresponds to that of the census to which it refers. Where the distance to be travelled in sparsely populated areas presents a problem and two visits to each household or to each set of living quarters are not feasible, the preparation of lists prior to the enumeration may be limited to urban or built-up areas. As a means of preventing omissions in sparsely populated area; a pre-census list is sometimes compiled on the basis of postal records or from information obtained from local doctors, missionaries, chiefs etc. Even a partial list may help to prevent omissions. This preliminary listing may be corrected and completed as the enumeration in these areas proceeds.

Although this is a bigger operation than up-dating a more complete list, it is essentially the same procedure.

B. Format of the lists

31. Because of its purpose and content, the control list is always prepared in columnar form. One list is usually prepared for each enumeration district and provision should be made in the heading or on the cover for appropriate identification.

32. For example, provision should be made to record major civil division, minor civil division, city, town, village or rural district, enumeration area and name of enumerator.

C. By whom prepared

33. If the census control listing is prepared in the weeks or days immediately preceding the census enumeration, then the census enumerators will be the logical persons to compile the control listing. In addition to providing a control for the census enumeration, the listing procedure serves to familiarize the enumerator with his district. If the same persons can be used for listing as well as enumeration there will also be a substantial saving on training. Persons who prepare the listings will have to be thoroughly trained in the definitions of the basic units to be identified for the census and, if statistical data are also being collected by means of the listing, they should be trained in enumeration techniques also. If the same persons are not used for the enumeration much of this training would have to be duplicated.

34. In cases where the list is prepared several months in advance of the enumeration this continuity may not be possible. However, where listers cannot be continuously employed up to the time of the enumeration there would be obvious advantages in attempting to rehire the same persons to serve as enumerators.

35. Where the control list is compiled at the same time as the enumeration, the census enumerator will automatically be responsible for both the questionnaire and the listing.

D. When prepared

36. Census control lists are most frequently prepared prior to the enumeration but very often they are prepared during the enumeration. Table 1 of the Annex shows that out of 89 censuses for which information is available, a control list was prepared for 48 prior to the enumeration, for 40 during the enumeration and for one following the enumeration. The length of time between the preparation of the listing and the enumeration may range from a few days to several months.

37. The timing of the control list will depend upon the purpose for which the listing is to be used and also upon the estimated time required to prepare it. If the listing is to be used for planning purposes, for example to estimate the number of enumerators and other personnel required or the quantities of census materials needed, then, sufficient time should be allowed for the ordering of materials, printing of forms, recruitment and training of personnel. It should be remembered, however, that although a long time period between listing and enumeration may be desirable from some points of view, there may be double training requirements if the same persons cannot be used for both the listing and the enumeration (see para. 33) and there is also the question of the data collected at the listing stage becoming invalid for statistical purposes unless care is taken to verify the information as of the census moment (see para. 18).

III. CONTENT OF THE LISTS

38. Clearly, the content of the lists will depend upon which of the purposes outlined in Part I above it is intended to serve. Where the purpose is purely regulatory, it must be decided which units are most likely to serve efficiently as controls for the enumeration. For example, if living quarters in a particular country are frequently grouped within compounds or concessions it must be decided whether compounds will serve as a useful unit of control. Buildings are normally recorded for purposes of control, and arrangements should be made so that living quarters which are not located in buildings are also included in the lists. If a housing census is being carried out

simultaneously with a census of population, living quarters should be individually recorded; it would also be useful to record them even in those cases where only a population census is being taken. The recording of households in addition to living quarters will be necessary since, by definition, more than one household may occupy a single set of living quarters and the recording of at least the number of households in each set of living quarters will be of assistance to the enumerator. The further identification of each household by recording the name of the head provides an even more secure means of ensuring that households are easily located at the time of, or following, the enumeration.

39. Provision should also be made in the listing to record questionnaires distributed, completed, or collected and to indicate living quarters or households that need to be revisited.

40. If it is proposed to use the listing to furnish advance census results a decision must be made concerning the topics for which information is required on a priority basis and these must be included.

41. Where topics which are to be used in connexion with a census of agriculture or a census of establishments are to be included, determination of the information to be collected should be made in consultation with the persons responsible for carrying out these inquiries.3/

42. Decisions concerning other housing or population topics to be included in the lists will depend upon the relative efficiency of collection and processing by means of the listing as opposed to the census questionnaire. In the case of India and Pakistan, all housing topics included in the 1960 censuses were collected during the preparation of the lists. Where only certain topics are included; it has to be borne

---

3/ For further information concerning the relationship and means of co-ordinating population and housing censuses with other statistical inquiries, reference may be made to the working paper entitled Co-ordination between housing censuses and population censuses and of these censuses with other statistical inquiries and compilations.

in mind that questionnaires and listings may need to be related during processing if cross-tabulation of topics collected in the listing with topics collected in the questionnaire is required. However, in the case of hand-processing, and provided the question or cross-tabulations is taken into account, it may be time-saving to have two processing documents.

IV. METHOD OF PREPARATION

43. The census control list will often be the first form to be completed during the census operation. In many cases, the information entered on this form will serve as the basis for the collection of all information in the census and therefore the utmost care must be exercised in completing it.

44. The personnel responsible for compiling the lists need to be thoroughly trained in the procedures to be followed and they should be provided with clear instructions concerning all aspects of their work. Certain aspects of the instructions will depend upon the particular circumstances in the country concerned and special instructions may have to be provided according to the way in which living quarters are arranged.

45. In general, however, there are three basic functions which have to be carried out during the listing procedure and these should be covered by instructions: (1) identification of the units to be included in the listing (2) the taking of measures to facilitate re-identification; and (3) recording of the units identified. These are discussed in the following paragraphs.

A. Identification of the units to be recorded

46. It is essential for those responsible for compiling the lists to be perfectly clear about the units which they are to identify. The definitions of basic concepts such as "compound," "building," "structure," "living quarters," "households" should be included in their instructions and must be understood beyond any shadow of doubt. Errors made in the initial identification of these units may seriously affect the results of the census.

47. In addition to a thorough understanding of the definitions, the lister must also have a knowledge of when and when not to include some of the units in the listing. For example, it should be clearly indicated in his instructions whether all buildings are to be included or only those containing at least one set of living quarters. Since all buildings will have to be investigated, a tighter control will be provided if they are all listed. On the other hand, if the listing is compiled immediately preceding or during the enumeration, it may be necessary to list only those containing living quarters.

48. In preparing instructions, it should be borne in mind that living quarters are not always located in buildings and some provision must be made for recording living quarters which are tents, boats, natural shelters, etc. If only buildings or other structures containing at least one set of living quarters are to be included, then the procedure to be followed for vacant units must be indicated. It may be noted that according to the recommended definitions, all living quarters except dwellings have to be occupied in order to fail within the scope of the census.

49. The instructions should indicate at what stage of construction buildings under construction are to be included. This will be affected by the time, if any, which is to elapse between the listing and the enumeration. If the time is substantial then probably buildings at whatever stage of construction should be included since they may have reached the required stage of completion by the time of enumeration. Buildings which are occupied as living quarters, even though they have not reached the prescribed stage of completion, should in any event be included.

50. Buildings in the course of, or scheduled for, demolition should not be included in the listing unless they are occupied or there is a possibility that they may be occupied at the time of the census. In these cases, they should be included in the listing but not in the census unless they are occupied at the time of the enumeration.

51. In some circumstances it may be useful to make a separate count, for national purposes, of the living quarters originally intended for habitation but used wholly for other purposes at the time of the census.

In these cases, specific instructions need to be issued with respect to the identification and recording of these units.

52. The relationship among the various units with which the lister, is to deal should be explained. For example, he should understand that a single set of living quarters may occupy more than one building, that it may be conterminous with a single building or that it may occupy only part of a building; that more than one household may occupy a single set of living quarters or that a household may occupy more than one set of living quarters. In the case of countries where the compound is to be recorded in addition to buildings, living quarters and households, these relationships are apt to be particularly confusing and need to be clearly explained. The differences between households and persons not living in households should be made clear.

53. The instructions should specify, as far as possible, places which are likely to be omitted during the preparation of lists, for example, living quarters which are in natural shelters such as caves, or which are located under bridges or in ruined structures; the living quarters of janitors or caretakers in office buildings, stores, garages or warehouses. Attention should be drawn to the fact that housing units may sometimes be located on the grounds or within the building housing an institution, camp, etc.

54. Since the order in which control units are recorded and enumerated will correspond to the route followed by the lister, he should be instructed concerning the route which he is to follow.

B. Measures to facilitate re-identification

55. The effectiveness of control lists depends upon the ease with which the physical units concerned may be re-identified, prior to, during or following the enumeration, from the description of the units recorded in the listing. The description usually includes the kind of unit (compound, building, etc.), an identifying number within a series for the enumeration area, an identifying number within a series limited to a

street, road or block, the name or number of the street or road, and/or any other information which distinguishes a particular unit from others in the same area. In the case of households, the name of the household head provides the identifying link between the unit and the listing.

56. Before the units are recorded, it will be necessary to consider whether sufficient means of identification exist or whether steps need to be taken to provide them. Where street names or numbers are lacking, it may be necessary for names or numbers to be designated. Where compounds, buildings or living quarters do not have permanent numbers or where there are gaps in the series, numbers will have to be assigned. Very often, the assigning and affixing of numbers is carried out jointly with the preparation of the listing. Consideration should be given to providing permanent identification to streets and buildings which can be used for successive censuses and for other purposes.

57. The numbering system to be used in the lists needs to be carefully considered so that it clearly identifies the units being included in the list and also facilitates, as far as possible, the carrying over of totals from one page of the listing to another and the provision of grand totals for the enumeration area concerned. Limitations on the system selected may be imposed by the system of permanent numbersif these already exist.

C. Recording of the units identified

58. If the preceding steps of identifying the units and ensuring that they can be easily re-identified have been properly carried out, recording them should present no great difficulty.

59. However, clear instructions need to be issued to the lister and/or the enumerator concerning the procedure that he is to follow in making additions to the list after all or part of it has been completed. For example, it may be found that although only one household has been recorded as occupying a particular set of living quarters, actually two households occupy these living quarters and the additional household has to be recorded. Should the additional household be inserted after the last household recorded and, if so, how should it be identified with the

living quarters to which it belongs which is further up the list? Upon checking the completed list, it may be discovered that some of the living quarters in a building have been omitted. They will need to be added to the list and the instructions should indicate how they are to be related to the building in which they are located.

V. HANDLING OF COMPLETED LISTS

Since the control lists, when compiled, will represent the only complete record of all the living quarters and households for which housing and population census data have been collected, they are valuable documents and should be treated as such. Their processing will depend upon the use for which they are intended but, since their primary purpose is one of control, at least one copy should be kept with the questionnaires to which they refer. Where the lists are used to compile advance census results, totals should be extracted from the lists by supervisors and entered on forms provided for this purpose thus freeing the lists themselves to be transmitted to the processing centre with the questionnaires.

Country experience in the preparation and use of census control lists

The following tables have been compiled in order to discern significant patterns in methodology with respect to the use, content and timing of census control lists. The information contained in the tables has been derived from census material made available by countries to the United Nations Statistical Office or from published census reports and refers to censuses carried out during the period 1955-1964.

Table 1 includes information on the timing and use of census control lists in 90 countries. The table shows whether the lists were prepared during the enumeration or before the enumeration and, if before, an effort was made to determine the length of time which elapsed between the completion of the list and the census date. In many cases, however, available information was insufficiently precise for the time interval to be established. The uses which appear in the table are those stated or implied in the material examined.

Table 2 shows, for 74, countries, the topics included in census control lists. In order to demonstrate as clearly as possible, the kind of data recorded or collected in the lists, topics for each country are grouped according to whether they served mainly as controls for the census, or whether they were included to collect population information, housing information or other data.

The frequency with which the topics were included in census control lists of the 74 countries is shown in Table 3. Topics are listed in rank order by frequency of collection within subject groups.

Page 2

Table 1

Countries by timing and use of census control listings in censuses of population and housing carried out during the period 1955-1964

Country and census date		When prepared		Use						
		Before enumeration (length of time if information is available)	During enumeration	To establish size of enumeration areas	To estimate census materials	To control enumeration	As a frame for sampling	To collect statistical data	For advance census results	Other
AFRICA										
Algeria	PH-1960	2 weeks	-	...	...	...	...	X	...	...
Angola	PH-1960	immediately before	-	...	...	X	...	X	...	X 1/
Cape Verde Islands	PH-1960	5 months	-	...	...	...	...	...	...	...
Chad 2/	PH-1962	X	-	...	...	X	...	...	X	...
Equatorial Guinea	PH-1960	-	X	...	...	X	...	X	X	X 3/
Ethiopia 4/	PH-1961	-	X	...	...	X	X	...	...	...
Ghana	PH 5/ -1960	-	X	...	...	X	...	X	X	...
Ifni	PH-1960	-	X	...	...	X	...	X	X	X 3/
Ivory Coast 6/	PH-1956/57	-	X	X	...	X	...	...	X	...
Liberia 7/	PH-1956	-	X	...	...	X	...	X	...	...
Libya	PH-1964	1 week	-	...	...	...	...	...	...	...
Mauritius	PH-1962	6 weeks	-	...	...	X	...	X	...	...

1/ To provide an inventory of buildings, housing units and establishments.

2/ Fort-Lamy only.

3/ Used to identify agricultural holdings.

4/ City of Addis Ababa only.

5/ Housing information referred to number of housing units and occupancy only.

6/ Housing of African population in cities of Abengourou, Abboville, Dimbokro and Man only.

7/ City of Monrovia only.

Country and census date		When prepared		Use						
		Before enumeration (length of time if information is available)	During enumeration	To establish size of enumeration areas	To estimate census materials	To control enumeration	As a frame for sampling	To collect statistical data	For advance census results	Other
AFRICA (CONT.)										
Mozambique	PH-1960	3 weeks	-	...	...	X	...	...	...	...
Rhodesia (S.), Zambia, Malawi	PH-1956	-	X	...	...	X	...	...	X	...
Sao Tomé & Principe	PH-1960	2 months	-	...	X	X	...	X	...	...
Seychelles	PH-1960	3 months	-	X	...	X	...	X	X	X 8/
South Africa	PH 9/ -1960	-	X	...	...	X	...	X	X	...
Spanish North Africa	PH-1960	-	X	...	...	X	...	X	X	X 3/
Spanish Sahara	PH-1960	-	X	...	...	X	...	X	X	X 3/
Sudan 10/	H-1960	X	-	...	...	X	...	X	...	...
Togo 11/	PH-1958/59	X	-	...	...	X	...	X	...	X 12/
Tunisia 13/	PH-1956	-	X	...	...	X	...	X	X	...
United Rep. of Tanzania 14/	PH-1958	-	X	X	...	X	...	X	...	...
Totals	23	11	12	3	1	20	1	16	11	7

3/ Used to identify agricultural holdings.

8/ Used as a check on final enumeration.

9/ Excluding housing of Bantu population.

10/ City of Wadi Halfa only.

11/ Communes of Lomé, Anécho, Tsévie, Palimé, Atakpamé, Sokodé and Bassari only.

12/ Used to identify "concessions" (compounds).

13/ Listing prepared for urban areas only.

14/ Former Zanzibar only.

Country and census date		When prepared		Use						
		Before enumeration (length of time if information is available)	During enumeration	To establish size of enumeration areas	To estimate census materials	To control enumeration	As a frame for sampling	To collect statistical data	For advance census results	
AMERICA, NORTH										
Antigua	PH-1960	3 weeks	-	...	...	X	...	X[15]	X	..
Bahamas	PH-1963	2-3 weeks	-	...	...	X	...	X[15]	X	..
Barbados	PH-1960	5 days	-	...	...	X	...	X[16]	X	X[1]
British Honduras	PH-1960	3 weeks	-	...	...	X	...	X[15]	X	..
British Virgin Islands	PH-1960	3 weeks	-	...	...	X	...	X[15]	X	..
Canada	PH-1961	-	X	...	...	X	X	X	X	..
Cayman Islands	PH-1960	3 weeks	-	...	...	X	...	X[15]	X	..
Dominica	PH-1960	5 days	-	...	...	X	...	X[16]	X	X[1]
Grenada	PH-1960	5 days	-	...	...	X	...	X[16]	X	X[1]
Jamaica	PH-1960	3 weeks	-	...	...	X	...	X[15]	...	..
Montserrat	PH-1960	3 weeks	-	...	...	X	...	X[15]	X	..
Nicaragua	PH-1963	-	X	...	...	X	...	X	X	X[1]
Panama[19]	PH-1960	2 years	-	X[20]	...	...	...	X	...	X[2]
Puerto Rico	PH-1960	-	X	...	...	X	X	X	...	..
St. Kitts-Nevis-Anguilla	PH-1960	3 weeks	-	...	...	X	...	X[15]	X	..

5/ Vacant living quarters recorded.

6/ Information referred to buildings and date of occupancy.

7/ Acres operated and livestock kept.

8/ To record agricultural holdings enumerated.

9/ Pre-census investigation for urban areas only.

10/ And time required to enumerate each block.

11/ Identification of agricultural holdings.

Country and census date		When prepared		Use						
		Before enumeration (length of time if information is available)	During enumeration	To establish size of enumeration areas	To estimate census materials	To control enumeration	As a frame for sampling	To collect statistical data	For advance census results	Other
AMERICA, NORTH (CONT.)										
St. Lucia	PH-1960	5 days	-	...	...	X	...	X16/	X	X17/
St. Vincent	PH-1960	5 days	-	...	...	X	...	X16/	X	X17/
Turks & Caicos Is.	PH-1960	3 weeks	-	...	...	X	...	X15/	X	...
United States	PH-1960	-	X	...	...	X	X	X	...	...
Totals	19	15	4	1	...	18	3	19	15	7
AMERICA, SOUTH										
Argentina	PH-1960	X	-	X	...	X	...	X	...	...
Chile	PH-1960	-	X	...	...	X	...	X	X	X22/
Ecuador	PH-1962	-	X	...	...	X	...	X	...	...
Guyana	PH-1960	5 days	-	...	...	X	...	X16/	X	X17/
Paraguay	PH-1962	-	X	...	...	X	...	...	...	...
Peru13/	PH-1961	X	-	...	...	X	...	X	...	...
Uruguay	PH-1963	-	X	...	...	X	...	...	...	...
Venezuela13/	PH-1961	X	X	...	...	X	...	X	...	...
Totals	8	4	5	1	...	8	...	6	2	2

13/ Listing prepared for urban areas only.

15/ Vacant living quarters recorded.

16/ Information referred to buildings and date of occupancy.

17/ Acres operated and livestock kept.

22/ Agricultural producers were identified.

Country and census date		When prepared		Use						
		Before enumeration (length of time if information is available)	During enumeration	To establish size of enumeration areas	To estimate census materials	To control enumeration	As a frame for sampling	To collect statistical data	For advance census results	Other
ASIA										
Bahrain	PH-1959	X	–	...	...	X	...	...	...	...
Brunei	PH[23]-1960	1 month	–	X	...	X	...	X	...	X[24]
Cyprus	PH-1960	2 months	–	X	...	X	...	...	...	...
Hong Kong	PH-1961	–	X	...	...	X	...	...	X	...
India	H-1960 P-1961	3 months	–	X	X	X	...	X	...	...
Indonesia	PH-1961	X	–	...	...	X	...	X	...	X[25]
Israel	PH-1961	...	...	...	...	X	...	X[15]	X	...
Japan	PH-1960	...	...	...	X	...	...	...	X	X[26]
Jordan	PH-1961	–	X	...	...	X	...	X	X	...
Kuwait	PH-1957	X	–	...	...	X	...	X	...	X[27]
Macau	PH-1960	1 month	–	...	X	X	...	X	...	X[27]
Malaysia, East										
Sabah	P-H[28]-1960	1 month	–	X	...	X	...	X	...	X[24]
Sarawak	P-H[29]-1960	1 month	–	X	...	X	...	X	...	X[24]
Pakistan	H-1960 P-1961	5 months	–	X	...	X	...	...	X	...

15/ Vacant living quarters recorded.

23/ Towns of Brunei, Tutong, Seria and Belait only.

24/ Used to collect information for the census of agriculture.

25/ Used to collect information for the census of agriculture and the census of establishment.

26/ Included information on farm household, fishery household, mining manufacturing etc.

27/ Used to collect information for the census of establishment.

28/ Towns of Jesseldon, Sandakan, Tawaua and Victoria (Labuan) only.

29/ Towns of Kuching, Simanggang, Sarikei, Sibu, Bintulu, Miri, Lutong and Limbang only.

Country and census date		When prepared		Use						
		Before enumeration (length of time if information is available	During enumeration	To establish size of enumeration areas	To estimate census materials	To control enumeration	As a frame for sampling	To collect statistical data	For advance census results	Other
ASIA (CONT.)										
Philippines	PH-1960	3 months	-	...	...	X	...	...	...	...
Portuguese Timor	PH-1960	-	X	...	X	X	...	X	...	X 27/
Syria 30/	H-1958	X	-	...	...	X	...	...	...	X 31/
Turkey 32/	PH-1960	-	X	...	...	...	...	X	...	...
Totals	18	12	4	6	4	16	...	11	5	9
EUROPE										
Austria	PH-1961	-	X	X	...	X	...	...	...	X 33/
Belgium	PH-1961	...	...	...	...	X	...	...	...	...
Bulgaria	PH-1956	5 months	-	X	...	...	...	X	...	...
Denmark	PH-1960	-	X	X	X	X	...	...	...	...
France	PH-1962	34/	34/	...	...	X	...	15/	...	...
Greece	PH-1961	X	X	X	...	X	...	X	X	...
Hungary	PH-1960	-	X	X	X	X	...	...	...	...
Iceland	PH-1960	X	-	X	...	...	...	...	X	...
Ireland	PH-1961	-	X	X	...	...	...	...	X	...

15/ Vacant living quarters recorded.

27/ Used to collect information for the census of establishment.

30/ Housing census for cities only.

31/ Used to collect information for the census of building.

32/ Housing of the population within the municipal boundaries of province and district centres.

33/ List used to distribute schedules for census of commerce and industry.

34/ Following enumeration.

Country and census date		When prepared		Use						
		Before enumeration (length of time if information is available)	During enumeration	To establish size of enumeration areas	To estimate census materials	To control enumeration	As a frame for sampling	To collect statistical data	For advance census results	Other
EUROPE (CONT.)										
Netherlands	PH-1960	-	X	...	X	...	...	...	X	...
Portugal	PH-1960	X	-	X	X	...	...	...	X	...
Spain	PH-1960	-	X	...	...	X	...	X	X	...
Sweden	PH-1960	X	-	X	X	...	...	...	...	...
Switzerland	PH-1960	-	X	X	...	X	...	15/	...	...
United Kingdom										
England and Wales	PH-1961	-	X	...	...	X	...	X	X	...
Northern Ireland	PH-1961	-	X	...	...	X	...	X	X	...
Scotland	PH-1961	-	X	...	...	X	...	X	X	...
Yugoslavia	PH-1961	-	X	...	...	X	...	X	X	...
Totals	18	5	12	10	7	12	...	7	10	1
OCEANIA										
Australia	PH-1961	-	X	...	...	X	...	...	X	...
Fiji	PH-1956	-	X	...	...	X	...	...	X	...
New Zealand	PH-1961	-	X	...	...	X	...	15/	...	...
Totals	3	-	3	...	...	3	...	...	2	...
USSR										
USSR 35/	H-1960	X	...	...	...	X	...	...	...	...

15/ Vacant living quarters recorded.

35/ List compiled for privately owned housing only.

Table 2

Countries according to the topics included in census control lists prepared in connexion with censuses of population and housing carried out during the period 1955-1964

Country and census date		Control topics							Population topics					Housing topics						Other topics		
		Location	Compounds	All structures	Structures with living quarters	Living quarters	Household	Questionnaires	Number of persons	Sex	Race	Nationality	Other	Occupancy	Type of living quarters	Construction material	Number of rooms	Use of structure	Other	Establishments	Agriculture	Other
AFRICA																						
Algeria	PH-1960	X	-	X	-	X	X	-	X	-	X	X	X	-	-	-	-	X	-	X	-	-
Angola	PH-1960	X	-	X	-	X	X	X	X	-	-	-	-	-	-	X	-	X	X	X	-	-
Chad 1/	PH-1962	X	X	-	-	X	X	-	X	-	-	-	X	-	-	-	-	-	-	-	-	-
Equatorial Guinea	PH-1960	X	-	-	-	X	X	X	X	X	-	-	-	X	X	-	-	-	-	-	-	-
Ethiopia 2/	PH-1961	X	-	-	-	-	X	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ghana	PH 2a/ -1960	X	X	-	X	-	-	X	X	-	-	-	-	X	-	-	-	-	-	-	-	-
Ifni	PH-1960	X	-	-	-	X	X	X	X	X	-	-	-	X	X	-	-	-	-	-	-	-
Ivory Coast 3/	PH-1956/57	X	X	-	-	X	X	-	X	-	-	-	-	-	-	-	-	-	X	-	-	-
Liberia 4/	PH-1956	X	-	-	X	-	X	-	X	X	-	-	X	-	-	-	X	-	-	-	-	-
Libya	PH-1964	X	-	-	X	X	X	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-
Mauritius	PH-1962	X	-	X	-	-	-	-	X	X	X	-	-	-	-	X	X	X	X	X	-	-
Rhodesia (S.), Zambia, Malawi	PH-1956	X	-	-	-	X	X	X	X	X	X	-	X	-	X	-	-	-	-	-	-	-
Sao Tomé & Principe	PH-1960	X	-	X	-	X	X	X	-	-	-	-	X	X	-	X	X	-	X	X	-	-

1/ Fort-Lamy only.
2/ City of Addis Ababa only
2a/ Housing information referred to number of housing units and occupancy only.
3/ Housing of African population in cities of Abengourou, Agboville, Dimbokro and Man only.
4/ City of Monrovia only.

Country and census date		Control topics							Population topics					Housing topics						Other topics		
		Location	Compounds	All structures	Structures with living quarters	Living quarters	Households	Questionnaires	No. of persons	Sex	Races	Nationality	Other	Occupancy	Type of living quarters	Construction material	No. of rooms	Use of structure	Other	Establishments	Agriculture	Other
AFRICA (CONT.)																						
Seychelles	PH-1960	X	–	–	X	–	X	–	X	X	–	–	–	–	–	–	–	–	–	–	–	–
South Africa	P-H 5/ -1960	X	–	–	X	–	X	X	X	–	X	–	–	–	–	–	–	–	–	–	–	–
Spanish North Africa	PH-1960	X	–	–	–	X	X	X	X	X	–	–	–	X	X	–	–	–	–	–	–	–
Spanish Sahara	PH-1960	X	–	–	–	X	X	X	X	X	–	–	–	X	X	–	–	–	–	–	–	–
Sudan 6/	H-1960	X	–	–	X	–	X	–	X	–	–	–	–	–	–	X	X	–	X	–	–	–
Tunisia 7/	PH-1956	X	–	–	X	X	X	–	X	–	X	–	–	–	–	–	–	–	X	–	–	–
United Rep. of Tanzania 8/	PH-1958	X	–	–	X	–	X	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Totals	20	20	3	4	8	12	18	10	17	8	5	1	5	6	5	4	4	3	6	4	–	–
AMERICA, NORTH																						
Antigua	PH-1960	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
Bahamas	PH-1963	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
British Honduras	PH-1960	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
British Virgin Islands	PH-1960	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
Canada	PH-1961	X	–	–	–	X	X	–	X	–	–	–	–	X	X	–	–	–	X	–	–	–
Cayman Islands	PH-1960	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
Jamaica	PH-1960	X	–	–	–	–	X	X	X	–	–	–	–	X	–	–	–	–	–	–	–	–

5/ Excluding housing of Bantu population
6/ City of Wadi Halfa only.
7/ Listing prepared for urban areas only.
8/ Former Zanzibar only

Country and census date		Control topics							Population topics					Housing topics						Other topics		
		Location	Coordinates	All structures	Structures and living quarters	Living quarters	Households	Coordinates of dwellings	No. of persons	Sex	Race	Nationality	Other	Occupancy	Type of living quarters	Construction material	No. of rooms	Use of structure	Other	Establishments	Agriculture	Other
AMERICA, NORTH (CONT.)																						
Montserrat	PH-1960	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
Puerto Rico	PH-1960	X	–	–	X	X	X	–	X	–	–	–	–	X	–	–	–	–	–	–	–	–
St. Kitts-Nevis-Anguilla	PH-1960	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
Turks & Caicos Is.	PH-1960	X	–	–	–	–	X	–	X	X	–	–	–	X	–	–	–	–	–	–	–	–
United States	PH-1960	X	–	–	X	X	X	–	X	–	–	–	–	X	–	–	–	–	–	–	–	–
Totals	12	12	–	–	2	3	12	1	12	8	–	–	–	12	1	–	–	–	1	–	–	–
AMERICA, SOUTH																						
Argentina	PH-1960	X	–	–	–	X	X	–	X	X	–	X	–	–	–	–	X	–	–	–	–	–
Chile	PH-1960	X	–	X	–	–	–	X	X	X	–	–	–	–	–	–	–	X	–	X	X	–
Ecuador	PH-1962	X	–	–	X	X	X	–	X	X	–	–	–	X	–	–	–	–	X	–	–	–
Paraguay	PH-1962	X	–	X	–	–	X	–	X	X	–	–	–	–	–	–	–	X	–	X	–	–
Uruguay	PH-1963	X	–	X	–	X	X	–	X	X	–	–	–	–	–	–	–	X	–	X	–	–
Venezuela 1/	PH-1961	X	–	–	–	X	X	X	X	X	–	–	–	X	–	–	–	–	X	–	–	–
Totals	6	6	–	3	1	4	5	2	6	6	–	1	–	2	–	–	1	3	2	3	1	–
ASIA																						
Bahrain	PH-1959	X	–	–	X	–	X	–	X	X	–	–	–	–	–	–	–	–	–	–	–	–

1/ Listing prepared for urban areas only.

Country and census date		Control topics							Population topics					Housing topics						Other topics		
		Location	Compounds	All structures	Structures with living quarters	Living quarters	Households	Questionnaires	No. of persons	Sex	Race	Nationality	Other	Occupancy	Type of living quarters	Construction material	Number of rooms	Use of structure	Other	Establishments	Agriculture	Other
ASIA (CONT.)																						
Brunei	P-H 9/ -1960	X	–	–	X	–	X	–	X	–	–	–	–	–	–	–	–	–	–	–	X	–
Cyprus	PH-1960	X	–	–	–	X	X	–	–	–	–	–	–	X	–	–	–	–	X	–	–	–
Hong Kong	PH-1961	X	–	–	–	X	X	X	X	X	–	–	–	–	X	–	–	–	–	–	–	–
India	H-1960 P-1961	X	–	X	–	–	X	–	X	X	–	–	–	–	–	X	X	X	X	X	–	–
Indonesia	PH-1961	X	–	X	–	–	X	–	X	–	–	X	X	–	–	X	X	–	X	X	X	–
Israel	PH-1961	X	–	X	–	X	X	X	X	–	–	–	–	X	–	–	–	–	–	–	–	–
Japan	PH-1960	X	–	–	–	–	X	X	X	X	–	–	X	–	–	–	–	–	–	–	–	–
Jordan	PH-1961	X	–	X	–	–	X	–	X	X	–	–	–	–	–	–	–	X	X	X	–	–
Kuwait	PH-1957	X	–	X	–	–	X	–	X	X	–	X	–	–	X	–	X	X	X	X	–	–
Macau	PH-1960	X	–	X	–	X	X	X	X	–	–	X	–	–	–	X	–	X	–	X	–	–
Malaysia, East																						
Sabah	P-H 10/ -1960	X	–	–	X	–	X	–	X	–	–	–	–	–	–	–	–	–	–	–	X	–
Sarawak	P-H 11/ -1960	X	–	–	X	–	X	–	X	–	–	–	–	–	–	–	–	–	–	–	X	–
Pakistan	H-1960 P-1961	X	–	–	X	–	X	–	X	X	–	–	X	X	–	X	X	–	X	–	–	–

9/ Towns of Brunei, Tutong, Seria and Selait only.

10/ Towns of Jessalton, Sandakan, Tawaua and Victoria (labuan) only.

11/ Towns of Kuching, Simanggang, Sarikei, Sibu, Bintulu, Miri, Lutong and Limbang only.

Country and census date		Control topics							Population topics					Housing topics						Other topics		
		Location	Compounds	All structures	Structures with living quarters	Living quarters	Households	Questionnaires	No. of persons	Sex	Race	Nationality	Other	Occupancy	Type of living quarters	Construction material	Number of rooms	Use of structure	Other	Establishments	Agriculture	Other
ASIA (CONT.)																						
Philippines	PH-1960	X	-	-	X	-	-	-	-	-	-	-	X	-	-	-	-	-	-	-	-	-
Portuguese Timor	PH-1960	X	-	X	-	X	X	X	X	-	-	-	-	-	-	-	X	X	-	X	-	-
Syria 12/	H-1958	X	-	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-	X	-	-	-
Totals	17	17	-	7	7	5	15	5	14	7	-	3	4	3	2	4	5	5	7	6	4	-
EUROPE																						
Austria	PH-1961	X	-	-	-	X	-	X	X	-	-	-	-	-	-	-	-	-	X	-	-	-
Belgium	PH-1961	X	-	-	X	X	X	X	-	-	-	-	-	X	X	-	-	-	-	-	-	-
Bulgaria	PH-1956	X	X	-	X	X	X	-	X	-	-	-	X	X	-	-	-	-	X	-	-	-
Denmark	PH-1960	X	-	-	-	X	X	X	-	-	-	-	-	X	-	-	-	-	X	-	-	-
France	PH-1962	X	-	-	X	X	-	X	X	-	-	-	-	X	-	-	-	-	-	-	-	-
Greece	PH-1961	X	-	-	X	X	X	X	X	-	-	-	-	-	-	-	-	-	-	-	X	-
Hungary	PH-1960	X	-	-	X	X	-	X	X	X	-	-	-	X	X	-	X	-	X	-	-	-
Iceland	PH-1960	X	-	X	-	-	-	X	X	-	-	-	-	-	-	-	-	-	-	-	-	-
Ireland	PH-1961	X	-	-	-	-	-	-	X	X	-	-	-	-	-	-	-	-	-	-	-	-
Netherlands	PH-1960	X	-	-	X	X	-	X	-	-	-	-	-	-	X	-	-	-	-	-	-	-
Portugal	PH-1960	X	-	X	-	X	X	X	-	-	-	-	-	X	-	-	X	X	X	X	-	-
Spain	PH-1960	X	-	-	-	X	X	X	X	X	-	-	-	X	X	-	-	-	-	-	-	-
Switzerland	PH-1960	X	-	-	X	-	X	X	-	-	-	-	X	-	-	-	-	-	X	-	-	-

12/ Housing census for cities only.

Country and census date		Control topics							Population topics					Housing topics						Other topics		
		Location	Compounds	All structures	Structures with living quarters	Living quarters	Households	Questionnaires	Number of persons	Sex	Race	Nationality	Other	Occupancy	Type of living quarters	Construction material	Number of rooms	Use of structure	Other	Establishments	Agriculture	Other
EUROPE (CONT.)																						
United Kingdom																						
England & Wales	PH-1961	X	-	X	-	X	X	X	X	X	-	-	-	-	-	-	X	-	X	X	-	-
Northern Ireland	PH-1961	X	-	X	-	X	X	X	X	X	-	-	-	X	-	-	X	-	-	-	-	-
Scotland	PH-1961	X	-	X	-	X	-	X	X	X	-	-	X	-	X	-	X	-	X	-	-	-
Yugoslavia	PH-1961	X	-	-	-	X	-	-	-	-	-	-	-	-	X	-	-	-	X	-	-	-
Totals	17	17	1	5	7	14	9	14	11	6	-	-	3	8	6	-	5	1	9	2	1	-
OCEANIA																						
Australia	PH-1961	X	-	-	-	-	X	X	X	-	X	-	-	X	-	-	-	-	-	-	-	-
Totals	1	1	-	-	-	-	1	1	1	-	1	-	-	1	-	-	-	-	-	-	-	-
USSR																						
USSR 13/	H-1960	X	-	-	X	-	-	-	X	-	-	-	X	-	-	X	X	-	X	-	-	-
Totals	1	1	-	-	1	-	-	-	1	-	-	-	1	-	-	1	1	-	1	-	-	-

13/ List compiled for privately owned housing only.

Table 3

Topics included in the census control lists of 74 countries in which population and housing censuses were carried out during the period 1955-64 (topics arranged in rank order by frequency of collection)

Control topics	
Location	74
Compounds	4
All structures	19
Structures in which living quarters are located	26
Living quarters	38
Households	6[illegible]
Questionnaires	33
Population topics	
Number of persons	62
Sex	35
Name of occupants	6
Race	6
Nationality	5
Age	4
Type of household	3
Socio-economic characteristics of household	2
Language	1
Literacy	1
Marital status	1
Relationship to head of household	1
Live births during last year	1
Deaths during last year	1
Infant deaths during last year	1
Number of family members of owner	1
Social group of house property owner	1
Household occupying whole or portion of dwelling	1

Housing topics

Occupancy status		32
Number of rooms		16
Type of living quarters		14
Use of structure		12
Construction materials: of outer walls	9	
of roof	6	
of floor	3	9
Ownership		7
Purpose for which building is intended		6
Number of floors in the building		6
Tenure		5
Water supply system		4
Electricity		4
Dwellings under construction		3
Type of building		2
Toilet facilities		2
Gas		2
Heating		2
Sewerage system		2
Owner of compound		2
Bathing facilities		2
Number of stables		1
Wall around house		1
Period of construction		1
Number of apartments in structure		1
Floor space		1
Rental value		1
Household occupying which floor of building		1
Name of person entitled to occupy the dwelling		1
Open verandah		1

Establishment topics	
Use of structure	12
Name of establishment, owner or operator	5
Number of persons employed	3
Branch of activity	3
Number of stores in structure	2
Kind of fuel or power used	1
Whether equipped with machinery	1
Number of manufacturers in building	1
Agricultural topics	
Livestock	4
Acres of land operated	4
Main crops	3
Owner of agricultural or livestock holding	1
Number of agricultural workers	1

E/CN.14/CPH/19

XII. MAPPING FOR THE 1966 CENSUS OF SWAZILAND

(Prepared by HUW M. JONES
Census Commissioner, 1966 Census of Swazilands)

MAPPING FOR THE 1966 CENSUS OF SWAZILAND

CONTENTS

	Paragraphs
INTRODUCTION	1 - 3
The use of maps	4 - 5
Existing map coverage	6
Plotting and mapping	7 - 9
Timing	10
Staff	11 - 12
Cost	13
Delineation of enumeration areas	14 - 17
Analytical mapping	18

Introduction

1. Swaziland, a small country some 6,700 square miles (17,300 square kilometres) in extent in south-eastern Africa, has a comparatively long history of census taking, the first being conducted in 1898 and followed by a further six censuses before 1966. Unfortunately, however, the data which they provided were suspect primarily because they were administrative censuses for which the 'census by assembly' technique was used. This method was emphatically rejected in favour of the household canvasser method when initial plans for the 1966 census were made, but its adoption specifically implied the division of the country into enumeration areas, which, with one possible exception in 1904, did not appear to have been attempted previously. Thus, whilst there were no precedents upon which to base preliminary proposals, a fresh approach to the problem was possible.

2. Unlike Kenya with its sub-locations, Uganda with its parishes or the Sudanese omodias, Swaziland has no small administrative units which could have been readily adapted for use as enumeration areas. Nor are there easily identifiable localities on which to base such units. The Swazi do not live in nucleated settlements but in homesteads (imiti, sg. umuti) which are residential units of territorially distinct collections of huts, varying in number from one to several hundred with the status of the homestead head (umnumzane), scattered across the country side. Although the individual was the ultimate enumeration unit, the homestead was chosen as the operational unit of identification. The homestead is the basic unit of the Swazi traditional system of administration with economic, political and social significance and may comprise one or more households belonging to the family of the umnumzane together with other dependants not directly related to him. Even the names which the Swazi give to areas of land do not encompass readily definable boundaries and frequently one area is known by a variety of names. Nor do tax lists permit spatial analysis, except by administrative districts, and they are notoriously

inaccurate. A further factor which made the division of the country into enumeration areas difficult was the complex tenurial pattern. There are three types of tenure in Swaziland which present a fragmented spatial pattern and comprise:

a) urbanized areas, which cover legally defined urban areas and their peri-urban precincts,

b) land held by individuals and companies in freehold or on long leases, and

c) rural areas, comprising land held in trust for the Swazi Nation by the Ngwenyama (King) and on which a traditional semi-communal type of tenure obtains.

It was important at all stages of the census to recognise these types of tenure and observe their boundaries as far as administrative convenience allowed.

3. The ultimate factors which determine the size and shape of an enumeration area, however, are the size and distribution of the population. But the uncertain basis of enumeration of the previous census in 1956 and the unreliable nature of the data which were collected gave little idea of how many people would probably be counted and even less as to how they were distributed throughout the country. As an essential planning preliminary, therefore, it was necessary to discover the size of the problem before enumeration areas could be delineated.

The use of Maps

4. Whilst cartographic techniques are essential for the successful planning and execution of any census, for the 1966 Swaziland census they were vital in providing solutions to the two twin problems outlined above. It was decided that the best way to obtain a quick and reasonably accurate idea of the size and distribution of the population in the urban and rural areas where the bulk of the population lived was to plot and map the position of every homestead in these two tenurial types. In parallel with this, another exercise using simple questionnaires sent

out to each owner or occupier was designed to discover how many people might be expected to be enumerated on freehold and leasehold properties. From the resultant data the delineation of enumeration areas would be possible.

5. Thus maps of varying types and scales would be required not only for plotting homesteads and subsequently mapping them but also for delineating enumeration areas, for planning and for administrative control. The importance of maps in analyzing the spatial patterns presented by the data was also realized and the cartographic programme was therefore planned as a complete operation from the stage of providing preliminary data through the use of maps in the enumeration phase to the analysis and illustration of the final results.

Existing Map Coverage

6. A report on cartography was, in fact, the first document to issue from the census office and its main purpose was to list all the available maps which might be of use. It covered three main classes, the first of which comprised township diagrams, both the official plans of the Surveyor General as well as maps of urban areas and their precincts compiled from aerial photographs and showing a variety of detail at scales of about 1:2,500. Secondly there were large-scale territorial maps of which the 1:50,000 topographic map of the country in 31 sheets prepared by the Directorate of Overseas Surveys was the most important. Although the relief and hypsometric detail was of the highest accuracy, these maps carried misleading locality information, no cadastral detail and out-dated road alignments. The Department (now Ministry) of Agriculture had, however, annotated a set of these maps with cadastral detail, recent road re-alignments, the line of secondary roads and tracks, the boundaries of a variety of agricultural service areas as well as many locality and feature names. They were ideal for census purposes and had been photographically reproduced so that further reproduction at suitable scales was possible. The third class then available comprised a variety of territorial maps, mostly at a scale of

1:250,000, illustrating the distribution of medical, education and administrative services, land tenure, physiographic features and communications. In addition, the territory was covered by sets of aerial photographs, the most recent flight having been made in 1961.

Plotting and Mapping

7. It was then for decision as to which types of map and scale were the most suitable for each particular purpose. Foroverall planning and administrative control the annotated topo-cadastral maps were used at a scale of 1:76,000. In urban and peri-urban areas the township maps derived from aerial photographs were used and these proved suitable with the exception of two densely populated areas of uncontrolled settlement; here aerial photography was planned but the warren-like nature of the settlement pattern would probably have defied adequate definition, even had this solution not been prohibitively expensive. In these two areas the boundaries of enumeration areas were pointed out to enumerators and a closer degree of supervision introduced.

8. After early suggestions to enlarge the topo-cadastral maps to field scale had been rejected for a variety of reasons, principally time and labour, it was decided to use aerial photographs as photographic maps for plotting homesteads in the field and for this purpose a section of maps covering rural areas were enlarged to 1:10,000. It may be considered that sufficiently accurate results could have been obtained by plotting directly from photographs but there were two main reasons why this method would not have been completely reliable. Firstly, there were good reasons for presuming that there is a continual and significant movement of homesteads from place to place and the interval of five years between photography and enumeration would have produced an incorrect settlement pattern. Secondly it has been found difficult to identify all homesteads from photographs because the huts blend well with the landscape. Particularly in areas characterized by granitic outcrops in the form of exfoliated boulders and in the Lowveld where the bush is dense, the beehive shape of the traditional Swazi hut defies ready identification, even using

the most modern photogrammetric equipment. And whatever the interpretive difficulties, it is never possible to determine with absolute certainty whether a homestead is occupied or abandoned because in many cases homesteads are left as they stood, the roofs collapsing straight down and the well trodden courtyards resisting the encroachment of grass for many years. For these reasons some ground control would have been necessary and it was decided to use photographic maps and annotate them with up to date settlement information. For mapping homesteads and delineating enumeration areas outside the urban areas a twice linear enlargement of the 1:50,000 annotated topo-cadastral maps was made and photographic positives of the enlargements obtained so that copies could be run off quickly and easily.

9. Apart from the photographic maps, which because of their size (27½" square), comulative weight and the need to keep them as clean as possible, had to be kept at district headquarters and issued as required, each supervisor was issued with a hard-board, covered in strong polythene to keep materials clean, and pens with felt or bamboo nibs which were found superior to any other type. Instructions for homestead plotting indicated the symbols to be used and the type of information to be collected for planning purposes. Existing homesteads still in position as on the map were ringed, abandoned homesteads were crossed out and new homesteads indicated by a square. The maps were also annotated with the sites of stores, schools, churches and any other information which might eventually be of assistance to enumerators. The instructions advised that in hilly country where the enumerator could see long distances and was accompanied by a guide there was no need to visit each homestead personally but in areas where visibility was restricted, in the thickly bushed Lowveld, for example, a call had to be made at each one. In transferring this information to the 1:25,000 maps, conventional dots were used for homesteads and other symbols for churches, schools and stores etc.

Timing

10. Census moment had been set as midnight on the 24th May, 1966, and the enumeration planned in two stages; the preliminary enumeration was to start on 10th May and final enumeration was to be done on 25th May. In order to allow the maximum possible time for ordering stores, recruiting enumerators and training, tasks which depended upon the provision of satisfactorily accurate preliminary data, homestead plotting was timed to begin in mid-December, 1965, and mapping the first available photographic maps early in January. The 1st March was set as the date by which enumeration areas had to be delineated but this was put back first by eighteen days, allowing a full seventy days, and then again deferred to 15th April. The reasons for the setbacks were various and cumulative in their effect; they included:

a) The short period allowed for planning and preparation which allowed little latitude in a tight schedule,

b) Serious delays in sanctioning advance expenditure which delayed recruitment of staff and the ordering of materials until a period which very inconveniently for the logistics of the operation coincided with a National Festival,

c) Inclement weather which included a severe cyclone,

d) An initial mis-appreciation of the speed at which the homesteads could be plotted, and

e) A lack of a sense of urgency about preliminary census operations by district administration staff responsible for field operations.

In the event, by mounting a crash programme of mapping and working long hours in the field the final dateline was met with only one serious effect. It was by this time too late to distribute maps to schools, so that the school children who were being trained as enumerators could not be taught to use the enumeration area maps.

This reduced their usefulness during the enumeration period and for post enumeration analysis because it was impossible to insist that the instruction to number homesteads on the maps in the order in which they were enumerated should be followed. Many enumerators were able to use their maps without training, however, but it did mean that supervisors were doubly careful in pointing out the boundaries of enumeration areas; the maps were valuable in that checks on coverage could be made on the spot by supervisors well versed in map reading, using enumerators' maps. With the checks which were introduced, the six boundary transgrassions which did occur were quickly spotted and corrected.

Staff

11. The plotting of homesteads onto the photographic maps was done by twenty census supervisors, semi-permanent staff recruited for a variety of field jobs of which this was the most important. One of the main objects of the first training course for census staff was to provide instruction in the use of photographic maps and the simple techniques of homestead plotting. In addition to lectures, a minimum of four hours was set aside for practical work and once the formal course was completed supervisors were split into two groups to plot homesteads in Manzini and Mbabane Urban Areas, a shake-down exercise under field conditions which provided additional plotting practice under close control. Instructions for plotting homesteads were also issued. The country was divided for preliminary census work into twenty areas, one supervisor being allocated to each, with District Commissioners responsible for those areas which fell within their districts. Transport specifically for census work was provided for each district.

12. It had been hoped to recruit a draughtsman for the cartographic programme but this proved impossible and it was therefore necessary for this work to be undertaken by an existing government drawing office. The Director of Geological Survey and Mines co-operated and the senior draughtsman in his departmental drawing office was appointed to supervise census cartographic work. From January until early March a trainee

draughtsman worked alone on mapping and was then assisted by a trained draughtsman whom it had been possible to recruit. Because of the serious delays in the field and failure to maintain an even supply of photographic maps to the cartographic office, a serious backlog piled up and to meet the final deadline a crash mapping programme was organized. In addition to six trained draughtsmen and tracers from government drawing offices, nine students were also employed for short periods during April to help to complete the mapping programme and produce enumeration area maps.

Cost

13. As in all census operations there is a hidden cost element of expenditure and the following mapping costs represent only those directly attributable to the scheme. It is not possible, for example, to cost the time of census supervisors on plotting because they were at the same time undertaking publicity and other preliminary work.

Additional Cartographic Staff	R840.00
Photographic Positives of EA maps	R370.00
Enlargement of Aerial Photographs	R2100.00
Freight Charges	R16.00
Drawing materials (say)	R100.00
Total	R3426.00

This figure represents some 5 per cent of the total expenditure directly credited to the census scheme.

Delineation of Enumeration Areas

14. The delineation of enumeration areas in rural tenure areas was done by the central census office and in other tenure areas by district administration. Amongst the factors used to determine the boundaries of enumeration areas were:

a) the need to analyze data by tenure,

b) the need to analyze data by geographical region

c) the need for clarity - boundaries had to be features clearly recognizable by enumerators; there was insufficient time to provide written descriptions and boundaries had to be identifiable from cartographic records; arbitrary lines drawn on maps which could not be translated into reality on the ground had to be avoided,

d) some enumerators had to be housed in camps and for economy the fewer the camps which would serve enumeration areas the better,

e) the need for boundaries to make sense in the future as well as for the 1966 census, and

f) the need to relate boundaries to those of rural development areas in which crash agricultural programmes were proceeding.

15. Amongst the determinants which governed the size of enumeration areas were:

a) the density of homesteads,

b) the type of terrain,

c) transport facilities,

d) shape, and

e) future use.

16. Once delineated each enumeration area was given a code number and name although the latter was of use only at local level. Each census supervisor's area had been given a code letter and within this area enumeration areas were numbered in sequence. This was done by district administration staff and when the schedules were being edited and batched before processing a new five-digit code number was assigned to each. This was to join certain enumeration areas on freehold farms which had very small populations, to renumber others to conform with the area of administrative districts, certain adjustments having been made for census purposes, to sort out sequences with numbers missing and generally correct the mistakes made in numbering in the field prior to tabulation.

17. In all 885 enumeration areas were delineated, 550 in rural areas, 278 on freehold and leasehold farms and 57 in urban areas. The average size was 7.5 square miles and the average population enumerated in each was 423 persons. The size of enumeration areas varied with tenurial type and with geographical region. In rural areas the average size was 6.4 square miles and the average number of persons enumerated was 477.

Analytical Mapping

18. Mapping the characteristics of the population to determine the spatial pattern was regarded as an essential analytical operation and also of importance in illustrating census data for administrative and planning purposes. The following maps were produced using a very accurate and detailed territorial topo-cadastral map at a scale of 1:250,000 as a base. They were compiled and draughted by the Census Commissioner and drawn for reproduction by draughtsmen in the Geological Survey drawing office.

a) a dot distribution map using one 0.8 mm dot for twenty persons and proportionate circles for heavy rural and all urban concentrations,
b) a choropleth density map based on enumeration areas,
c) An isarithmic density map using enumeration area densities at a scale of 1:500,000,
d) an isarithmic density map using the dot distribution map as a base and plotting values for every four square miles, at a scale of 1:250,000, and
e) a map of the distribution of the areas of origin of temporary absentees from the country.

HUW M. JONES
Census Commissioner

XIII. PILOT SURVEYS AND PRE-TESTS OF CENSUS PROCEDURES

(Prepared by the ECA Secretariat)

PILOT SURVEYS AND PRE-TESTS OF CENSUS PROCEDURES

CONTENTS

	Paragraphs
General remarks	1 - 3
Survey methods	4 - 8
Pre-testing of enumeration procedures	9
Collection of data on characteristics of the population	10 - 11
Pre-testing the questionnaires	12 - 13
Training of enumerators	14
Public reaction	15 - 17
Conclusion	18 - 20

General remarks

1. A population census is a complex undertaking, which marshals quite considerable resources in a short space of time. Inevitably, the planning of such an operation is liable to defects which must be corrected as far as possible before the actual census takes place since it is then usually impossible to do so. It was therefore not long before the need was felt to conduct trial censuses or pilot surveys on a reduced scale to make sure that they worked smoothly in practice to test the methods of data collection and, generally speaking, acquire the means of improving the results of the final operation.

2. The use of pilot surveys has often been recommended by the United Nations. In particular, the recommendations of the Statistical Office of the United Nations for the 1970 Population Census, stresses the usefulness of pilot surveys which are "essential for countries without a long history of census taking...". The industrialized countries carry out pilot surveys regularly before taking a population census. This is true now of Great Britain which is preparing the 1971 census. The Fifth Conference of African Statisticians also recognized the value of pilot surveys in relation to population censuses.

3. Pilot survey targets extend in principle to the testing of all the aspects of a census, and may be grouped under a few heads:

(a) testing the enumeration procedures:

(i) the organization of the enumeration,

(ii) the methods of obtaining data in characteristics of the population,

(iii) training of enumerators,

(iv) public reaction to the census;

(b) testing of methods of processing the census data;

(c) testing of methods of evaluating results.

Survey methods

4. It may be supposed that the ideal method is to conduct pilot surveys in a probability sample representative of the population where the census is to take place In this way, it might be possible to obtain a population estimate, which is not always available, before the actual census. In the case of a sample census, the pilot survey would facilitate the evaluation of variability. In actual fact, the use of probability sampling is not practicable in the majority of cases. The possibilities in this field are limited by the resources of the African Statistical Offices. It is either impossible to draw a sufficiently large sample, or else the geographical distribution of the sample entails too heavy expenses.

5. Consequently, the method of purposive sampling is the one in most frequent use. The geographical areas are chosen for the survey in such a way as to be "representative" of the different conditions in which the census is to be carried out. As a matter of fact, population density, transport facilities, urban or rural characteristics, etc., vary from one region to another, affecting various parameters such as the duration of the census, census costs, and the like. A typical example of this process is to be found in the report on the 1966 Swaziland population census. The pilot survey was carried out in three selected zones, the first in a rural area, the second in an urban area, and the third in a forest plantation which was considered typical of individual holdings. In this particular instance, it seems as if an attempt was made to reproduce in the pilot survey the actual census conditions by including zones each one of which was typical of areas where the environment for census purposes was presumed to be different.

6. The 1953 pilot population census in the Sudan provides an example of a combination between purposive sampling and probability sample methods. Eleven pilot zones were selected from 9 regions, covering various types of population. Within the pilot zones a sample was taken, the final number of persons covered being 39,202, or 3.6 per cent of the population in the selected zones. In the population census of Uganda (1969) it is proposed to select from each of the 180 districts in the country two parishes, in such a way as to cover something like 90,000 people for a pilot survey.

7. In the African countries where different regions have very different conditions as regards accessibility, population density, distribution, etc., there is every reason to make a careful distribution of the zones selected for pilot surveys, so as to obtain a complete picture of the actual conditions for carrying out the census.

8. The number of persons to be included in a pilot survey should be at least of the order of 20,000 to 30,000. Nor is it necessary to concentrate all the details for pre-testing in a single survey. Several pilot surveys may be required. For instance, the population census of Ghana in 1960 included 2 pilot surveys, the "First Field Test" and the "Trial Census". Finally, the pilot survey must be carried out fairly well ahead, in practice at least one year before the census, so that there may be sufficient time to draw the necessary conclusions. In Great Britain, the 1966 sample census was conducted on the basis of a pilot survey carried out in 1964.

Pre-testing of enumeration procedures

9. Pre-testing the smooth running of the census organizational machinery is one of the most important objectives of a pilot survey. This includes a number of heads, the main ones being as follows:

(a) Pre-testing the division into enumeration zones (EAs) care must be taken to see that zones do not overlap, are not too large for the enumerators to cover, etc.

In Ghana, when preparations for the 1960 census were being made, a number of errors could thus be detected from the maps and locality lists given to the enumerators. It is of fundamental importance to have an exhaustive list of the various localities.

(b) Pre-testing the organization of work: the success of a census depends on the way in which the various aspects of work have been distributed, the efficiency with which the supervisors carry out their work, the degree of accuracy with which the work of each officer has been defined etc. An effort must be made to ascertain whether liaison between the various organs is proceeding satisfactorily, for instance, verifying whether the documents have been distributed in time, fully collected without loss of time, whether the supervisors have actually gone into the field and were able to supervise the work of the census officers, etc.

(c) Pre-testing the duration of the census and costs: the census plan rests on assumptions concerning the number of persons a census officer can enumerate in a given period of time. It is of the utmost importance that the accuracy of these assumptions should be checked, since the duration and costs of the census depend upon them. In Africa, it is generally considered that a census officer can count 500 people a week. However, it is possible in some cases to strike lower rates, for instance in the Sudan in 1953 the number of persons enumerated per day per enumerator varied from 33 to 156. In some African countries where density often falls below one inhabitant per $km.^2$, particularly in the forest zones, much lower rates of enumeration may be obtained.

(d) Detecting and identifying statistical units: During the census, the enumerators should be in a position to identify without difficulty or ambiguity such units as a village, a household, a dwelling, etc. The pilot survey affords a valuable opportunity for determining whether the definitions worked out in the office can be applied in the field.

(e) Census period: a pilot survey should in principle be conducted during the same season as that selected for the census. The census management would then be in a position to assess the wisdom of its choice bearing in mind points observed during the pilot survey, for example, on population movements, transport facilities, agricultural work etc.

Collection of data on characteristics of the population

10. Experience has shown that some population characteristics are difficult to observe, whether intrinsically or for reasons peculiar to the terrain. In Africa, illiteracy and the late introduction of the registration of births and deaths, have resulted in the fact that few people have any clear idea of their age, and this leads to serious difficulties in determining that particular basic characteristic.

11. A number of countries have included in their pilot survey programmes, the testing of methods for determining ages: in Ghana, for instance, the survey revealed that only 14.3 per cent of ages were determined with reference to a historical calendar. The pilot survey can facilitate the comparison of the advantages and disadvantages of the different procedures in estimating age and as such is of considerable importance. Data on fertility and mobility are generally not well known in Africa. Taboos and other religious beliefs are often an obstacle in securing a complete statement of deaths. During the pilot survey some idea of the importance of such problems may be gained.

Pre-testing the questionnaires

12. Generally speaking in Africa, questionnaires are filled out by the census officer. Even so, these questionnaires should be drawn up with a great deal of care. It must be steadily borne in mind that they should be clear beyond shadow of doubt to the enumerator, present no physical difficulties by way of awkward format or type-face and should be such as to make the processing of information as easy as possible.

13. Pilot surveys have been of great assistance in the preparation of questionnaires. Almost invariably after such surveys, improvements have been made in drawing them up and these have sometimes been substantial. Processing of information from questionnaires gives valuable information on public reaction to various questions, and throws some light on their varying degrees of difficulty.

Questionnaires are tried out as a rule in part before the pilot survey, as such. The first draft questionnaires are tried out on restricted samples, and this facilitates the avoidance of serious errors. In Ghana in 1960, the "First Field Test" was used to sound out the draft census questionnaire as well as other forms.

Training of enumerators

14. During the pilot survey, the enumerators and other officers have to work in conditions that are practically the same as those of an actual census. This exercise thus makes it possible to check the quality of the training given them and determine the problems they are up against, and from that point to improve their training. There is no doubt that the pilot survey is useful in training the staff who will direct the operations of the actual census. Such staff is given the opportunity in practice to gain first-hand knowledge of problems which are hardly likely to occur in the office. This was particularly noticeable in Ghana where one of the objectives of the pilot survey was to enable the management personnel to realize the type of problems that might arise in the course of the actual census.

Public reaction

15. Public reaction to the census is a decisive factor, and that is why it has been recommended that very wide publicity should be given to the census, and no opportunity lost of educating the people. Their fears and suspicions are sometimes justified by the past experience some of them have had in administrative censuses carried out during the pre-independence period.

16. The pilot survey might serve as a means of determining the real impact of publicity, comparing the effectiveness of different publicity props: radio, newspapers, and the like. In Ghana, it was possible in this way to collect data on what the public knew about the census, and this according to the type of information media. It revealed the fact that in the capital and the surrounding areas, the percentage of people who were well-informed was abnormally low.

17. The pilot survey can also reveal the kind of obstacle that may make a census unpopular. For example, it will be observed that some groups in the population systematically oppose a census, as it is the case in Congo (Brazzaville) in connextion with the religious sect known as the "Matswanistes".

Conclusion

18 During the preparation and implementation of the pilot survey, it must constantly be borne in mind that it is fairly difficult to make any rigid separation in the field between the testing of the various census aspects. The training of enumerators, the structure of the questionnaires, the method used for observing demographic variables, and the like all constitute a living entity, the analysis of which is rather delicate and arbitrary. Consequently the judgement of the demographer or statistician will often be qualitative, often even synthetic. However, it is still useful to separate the testing of the various aspects of the census, and some provision must be made for this by way of a procedure for recording field observations.

19. The pilot survey will also facilitate checking of the system of processing the results. The rapid and efficient functioning of this last activity determines whether the census figures are quickly published, the value of which is greatly reduced if the delay in publishing them is too long. There is every reason why the standards of work set for the coder should be tested to see whether they are realistic. Codes and various types of instruction, envisaged for the processing of the data should be tested and, if necessary, corrected. Experience reveals that the establishment of good codes is greatly facilitated by testing them in a population sample. The testing of computer programmes should not be neglected. Experience has shown that the development of such programmes takes a fairly long time and if the appropriate steps are not taken, their execution may cause considerable delay in the publication of the results.

20. Most of the African countries have conducted censuses, and seem to have carried out one or more pilot surveys in advance. It is eminently desirable that this practice should be made general, and thus contribute to an improvement in the quality of data. It would therefore be most useful if the national statistical services would as far as possible, publish the methods used and the results obtained from these pilot surveys as well as other relevant data, so that their experience when made known might lend a handle to progress in demography in Africa and the other regions.

XIV. ENUMERATION PROBLEMS AND PROCEDURES

(Prepared by the ECA Secretariat)

ENUMERATION PROBLEMS AND PROCEDURES

CONTENTS

		Paragraphs
INTRODUCTION		1 - 2
I.	The physical environment, dispersion of settlement and mobility of population	3 - 12
II.	The respondents	13 - 17
III.	The enumerators	18 - 21
IV.	Implications for the general census plan	22 - 25

INTRODUCTION

1. The accuracy of a population census in respect of (a) the number of persons enumerated and (b) the characteristics of these persons, depends almost entirely on the degree to which obstacles are foreseen and methods devised to minimize their effects. Similar considerations apply in a housing census but the emphasis in this paper is on population censuses: in practice, of course, complete coverage of population implies complete coverage of all the places in which people may be living even if the living quarters are not themselves recorded.

2. The essential preliminaries in Africa, as elsewhere, are a census law which gives adequate powers to the census organization (and preferably guarantees the confidentiality of the census results in relation to individuals), and a budgetary allocation which makes possible the operations of the census organization. Needless to say, the law cannot automatically secure complete support from the public and, in fact, census planning must build up this public cooperation and try to counterbalance genuine inability to produce answers to essential questions, at the same time as it attempts to provide for the difficulties arising from diversity of physical and social and cultural conditions. The census preparations must therefore consist in a judicious mixture of psychological, substantive and geographical planning which may have, according to local circumstances, to overcome any or all of the problems referred to below.

(I) The physical environment, dispersion of settlement and mobility of population

3. Urban areas. It may be thought that a satisfactory enumeration is relatively easy to ensure because of the greater concentration of settlement together with the possibility of dividing towns into regular blocks and utilizing (or organizing a system of) street names and house numbers. It is true that in parts of many towns in Africa, such conditions prevail and enumerators can be instructed to proceed systematically in a specified direction around the blocks allocated to them, attaching labels or other signs to indicate that each structure (or, if preferred, each set of living quarters) has been visited. Higher levels of education, closer contact with administrative services, expectation of further improvements in

facilities, accessibility to publicity media, availability of higher-level enumerators together with limited travel and close supervision, certainly favour a relatively complete and accurate enumeration. Yet the difficulties are numerous: many adults are absent during working hours and enumerators may have to work mainly at night, especially to find single-person households; homeless persons, beggars, residential schools, hotels and so on pose special problems, as do complicated households such as those composed of several adults sharing the expenses of house-keeping; mobility of population within the town and also as between the town and the rural areas may confuse both respondent and enumerator unless the enumeration can be conducted in a very short period of time; occupation, ethnic group, religion, educational levels vary far more than is the case in rural areas; there may be concentrated settlements of particular religious, ethnic or language groups whose attitude to the census is unfavourable. Often greater care and more time are required in the enumeration of an urban household than are required in a rural enumeration.

4. When households in shanty-town areas are being enumerated, the difficulties are even greater: in addition to the standard problems of variability in characteristics of urban households, there is the difficulty in systematic canvassing of Enumeration Areas where living quarters consist of very informal structures or parts thereof and where there is no pattern of regular blocks.

5. The solutions lie in pre-listing of households and living quarters[1/] and, where possible, in pre-numbering (of structures or of living quarters); in close supervision; in insistence on call-backs to find acceptable respondents; and even in the imposition of a curfew which would permit the enumeration to be carried out at night.

6. Rural areas. The difficulties increase according to whether the population of specific areas is (a) sedentary and in concentrated settlements (villages, hamlets, localities), (b) sedentary but not organized into villages or comparable units, (c) semi-nomadic or (d) nomadic.

1/ See Preparation and Use of Census Control Lists (E/CN.14/CPH/5; ST/STAT/22).

7. The active support of the leaders of the traditional society can do much to overcome the obstacles. The preparation of locality lists and the delimitation of Enumeration Areas depends largely on their help, if it is to be really effective; and at the stage of the enumeration proper the local leaders can show enumerators the boundaries of their areas on the ground. (See Locality Listing and Delimitation of Enumeration Areas, E/CN.14/CPH/9, para. 14-18). Their functions can be extended even further with great advantage where the possibility of mass publicity is limited by lack of facilities: in such cases, the cooperation of the local leaders is essential in passing down to each household head, and so to each individus within his sphere of authority, his support of the census and his reasons for asking his people to answer specific questions and to facilitiate in whatever way possible the work of the enumerators. This has been done in many African inquiries, and useful examples, of equal applicability in complete enumerations, are to be found in the accounts of the sample surveys in French-speaking West and Equatorial Africa, where the line of authority from village chief to chef de concession to chef de menage was an intrinsic part of the organization of the inquiry as well as of the method of recording the data on the questionnaires. It is undoubtedly one of the most practical measures which can be taken to allay fears and resentments which may seriously hinder the operation.

8. Enumerating a rural sedentary population in concentrated settlements is perhaps simpler than an urban enumeration, provided that the terrain is fairly open, as it is in much of coastal North and West Africa, and that adequate geographic preparations and publicity have been undertaken. If, however, as in forest areas in Central Africa and in the mountains of Ethiopia, physical difficulties result in a discontinuous pattern of settlement associated with poor land communications, the risk of omissions is great because, where it is impossible to canvass an Enumeration Area systematically, whole villages, hamlets or other agglomerations may be entirely missed. Similarly, where the pattern of settlement consists of isolated (or at least separated) homesteads or plots without a unifying village organization, as in parts of Eastern Africa, and in Southern Sudan,

it is very easy to omit individual units. Various attempts have been made in Africa to pre-list and even to map isolated households (Tunisia and Swaziland, 1966) and this can no doubt be effective where the village leader can be counted upon to ensure completeness or near-completeness in the list Otherwise, it may be necessary to leave the initiative to the enumerator to inquire where he can about households which are out of sight and to which there is no visible way of access.

9. The problems of finding people at home are sometimes serious in the rural areas, and this is particularly so where the terrain is such as to discourage, or actually to prohibit in the time available, return visits. If it is possible for local or other leaders to arrange that certain households stay at home from their fields on a specified day, all to the good; otherwise enumerators must work in the very early morning and in the evenings to complete the enumeration in the agreed period of time. If this is combined with extensive travel between households, compounds (concessions), or hamlets, the effect of fatigue on the enumerator may be harmful to the quality of the work he can do. This will be even more marked if it is found necessary to go out to the fields to interview households who remain away from their homes for several nights at a time. For reasons related to what has just been said, adequate supervision is hard to arrange in rural areas, except in closely-settled areas, and the supervisor may not be able to do more than a final check on the work of each enumerator. He can, of course, even within these limitations, insist on re-enumeration of households where the quality of the work is clearly unsatisfactory.

10. By implication, semi-nomads are sedentary part of the year, moving as necessity requires in search of food and water for their stock. They should be enumerated in whatever place seems most convenient on the basis of advance knowledge of their habits. Even then, a problem may arise from the absence of individual members of a household, or employees of the household, who may be absent watering or pasturing stock; these absentees should be enumerated either at the cattle-posts or other watering or pasture-grounds, if this is both possible and convenient, otherwise on the

basis of information given by their households.

11. The enumeration of nomads is the most difficult problem in African censuses. Administrative pre-listing or special attempts to obtain lists through the leadership system of the tribes or equivalent groupings (compare E/CN.14/CPH/9, paras. 19-22) offers a first approach, similar to that used in the Sudan in 1955-56. Listing of watering places is another method, but it should be noted that in the experiment in Niger this did not prove as successful as expected because the list of watering places was incomplete and also because movements of the tribes even over small distance created great difficulties. Enumerating at places where the tribes congregate has obvious advantages, but does not seem sufficient in itself. The experience of the 1946 enumeration of the Beersheba Bedouins is a good example of a psychological approach to the traditional leaders and then to the people, the enumeration being conducted on the basis of tribal groups and sub-groups rather than on the basis of geographically-determined EAs.[1/]

12. Other types of spatial movement which complicate African censuses are international and internal migration, and the movements of migrant and seasonal labourers. The two former are dealt with in the African Recommendations for the 1970 Population Censuses (E/CN.14/CAS.6/1) paras. 44-52. Brief mention is made here of census investigations of migrant labourers because this is a real issue in such countries as Botswana, Lesotho and Malawi, all of which have inquired about household members absent on work contracts in neighbouring countries. The information so obtained may be disorted by reports relating to the same person from more than one household (for example, a mother and a wife may both report the same man), and because the information given may be less accurate than that provided by those present at the interview. Valuable though the data are for particular purposes, the tabulations should be made separately for such absentees, Seasonal labourers constitute a similar problem in a number of countries, and techniques have been applied in French-speaking West Africa which would be equally applicable to the investigation of migrant labourers: that is, the enumeration of separate categories of Residents present,

1/ Sami W, Dajani: "The Enumeration of the Beersheba Bedouins in May 1946" in Population Studies, December, 1947.

Residents absent and Visitors in each household, with additional questions on duration of absence (or visit), place of present sojourn (or of usual residence), reason for absence (or visit). Such questions are, however, more suitable to sample surveys (in which they were in practice used) than to the general censuses which are the subject of the present paper. Nevertheless, the questions could in principle be employed in a sample of population only, additional to the questions asked in the general census.

(II) The respondents

13. The qualities and attitudes of the population to be enumerated in Africa vary enormously, as does the physical environment. There are hindrance to accurate replies on the part of respondents and sometimes motives for deliberate evasion of enumeration or for falsification of replies during enumeration, arising from a great variety of sources. Examples are too many to list exhaustively, but the following gives an impression of the problems: superstitious fears (reluctance to mention all children born or to refer to deaths); fear of taxation or of military or civil service, mainly a relict of administrative counts; failure to comprehend the question, especially where language is a difficulty and where interpretation may not be sufficiently precise to clarify the meaning, or where equivalents may not exist in the vernacular (questions on marital status may be unanswerable in any precise sense where marriage is a process rather than a suddenly-acquired state and where widows are inherited on the death of the husband by a male relative); forgetfulness (omission of very young children); resentment of the status of the interviewer (sex, caste, ethnic group, religion, language); fear of revealing adherence to an unpopular minority (religious, ethnic, national, religious); personal vanity (up-grading of educational or occupational status). The question of names of individuals is an additional well-recognized hazard in enumerations, because a person may have different names which are used in different contexts and the names may be changed over time, or may differ in spelling. Tribal names also constitute a hazard, because the respondent may give a sub-tribe or clan name which does not fit into the enumerator's list

14. Some of these factors are already of minor significance such as the superstition attaching to reports of dead persons, but others need to be combatted by census "education", which, by emphasizing the potential value of the census in the extension of health services, schooling, and so on, and by calling on national pride and prestige, may inculcate a sense that everyone should cooperate to the full. Others again, such as the fear of violation of religious traditions which require the seclusion of Muslim women, must be overcome by provisions which meet census needs without offending the beliefs of adherents; for example, by the use of female enumerators, as in the post-enumeration check on the numbers of young children in the Sudan in 1956.

15. There are naturally some questions, such as the reporting of age, where the amount that can be done is limited until levels of education are raised; this does not preclude the use of historical calendars of events and attempts to persuade the population, before the census, to do their best to get some evidence of their age, from written records if these are anywhere available or from comparison with other people in their local community whose age IS known, and whose relative status would at least determine whether A is older or younger than B, or approximately the same age. (See Age Data in African Censuses and Surveys, E/CN.14/CTH/13).

16. For the rest, the solutions must lie mainly in (a) pre-testing questions to ensure that there is nothing offensive or misleading in the content or form of the question and (b) in training enumerators in techniques of eliciting replies which, if not precisely correct, seem at least to be reasonable approximations.

17. Typically, the schools in Africa have played a large part in census education and the continuation of this channelling of information and ideas from children to parents offers prospects of further development of understanding of what a census aims at achieving and how the population can cooperate.

(.) The enumerators

18. Complementary to the information which is to be given by respondents is the function of the enumerators in eliciting and recording it. The

standard of work attained by the enumerator constitutes the essential success or failure of the census: an accurate enumeration or one which is unreliable and perhaps positively misleading in the conclusions which can be drawn from the results. While it is impossible here to go into details of recruitment and training of enumerators, some reference must be made to the sources of error which may derive from the enumerators themselves. Lack of sufficient educational background is a common trouble, but not necessarily the essential one. Lack of training and lack of adequate supervision are more serious in some instances, while in others training and the preparation of instructions to field workers may have been too ambitious and may have confused rather than clarified the issues at stake.

19. Whatever the reasons, enumerators frequently fail to carry out instructions on such matters as filling in every column even if the entry is "nil" or "not applicable" or "not stated"; using ditto marks and failing to alter them if a correction has had to be made somewhere down the column; assuming that all members of a household have the same nationality or ethnic origin; failing to check consistency between entries (for example, a very young woman with an unbelievably large number of children), entering occupation or industry in a way which cannot be subsequently classified ("labourer" or "fish" or something such). More serious again is the failure in some cases to take proper measures to see that they have canvassed the whole of their EA and failure to make sufficient effort to ensure complete coverage o membership of the households they enumerate.

20. It is a fairly elementary consideration that enumerators, like respondents, must have a cooperative attitude to the census, and one way of achieving this is to convince them of the value and prestige of the job they are doing; the other is to arrive at a system of payments which they regard as acceptable. The payment need not be high but it should offer some compensation for the hard work involved and it should not be capable of manipulation in a way which would encourage over-enumeration (payment by numbers enumerated) or under-enumeration (payment by number of days worked irrespective of the numbers enumerated). A lump sum, with perhaps a bonus

for difficult or dangerous conditions, seems the most reasonable method.

21. Regional, and sometimes local, ambitions have been known to produce motives for inflation of numbers, which, however it is achieved in practice, is likely to be based on inflation of numbers in individual EAs. This tendency, related to political rivalries and the search for greater representation of particular parties or areas in parliament, is something fairly new in Africa and should be avoided at all costs. There has been an unfortunate move to hold censuses close before parliamentary elections, and the connection between the two is prejudicial to the objectivity of the census. Unless the provisions of the law are inescapable there should be a reasonable time-interval between the two operations. Where for particular reasons, this separation in time is impossible, rigid controls need to be exercised of the work of each enumerator in the field followed by field reconciliation of discrepancies after the census has been concluded. If inflation of numbers is thought to have occurred at other levels than the EA, then the totals in each enumerator's census returns (or totals of groups of enumerators' returns) should be compared with the totals for the corresponding area as eventually established for publication or for electoral use.

(IV) Implications for the general census plan

22. It is considerations such as those outlined above that have influenced the African Recommendations for the 1970 Population Censuses to give preference to an enumeration of the de facto rather than the de jure population. The main advantage is that the present-in-area population is so much more easily defined and that enumerators do not have to hold in mind complicated instructions about temporary visitors, temporary absentees, and persons who are not to be included at all because they are not "residents" of the national area.

23. Where resources are really scarce, it is not always possible to complete the enumeration in as short a time as is desirable. The ideal plan would be to do it in one day, in this way minimizing the effects of population movements in the census period. In practice, some African

countries have been forced to permit the enumeration to continue for two or more weeks or even longer, because of the shortage of enumerators, and this may give rise to double counts or to omissions of persons who have changed residence in the period. Such an extension of time must be tolerated on occasion, but every other possible means should be taken to simplify the enumeration: simple and clear-cut definitions of household, and other census concepts, and simple and clear cut instructions for recording household members and their characteristics; a questionnaire design which facilitates to the maximum the task of making the entries in an unambiguous fashion; elimination through pre-testing of questions which may not be fully understood or which may arouse antagonism.

24. Despite all of this, the advantages of a simple de facto enumeration may have to be jettisoned where the effects of mobility cannot be eliminated, for example, where the general enumeration extends beyond, say, one month, or where it is a question of enumerating nomads.

25. Finally, the importance of the choice of the census date cannot be over-emphasized, since this too can be used as a means of reducing the effects of mobility, provided a date can be fixed at which movement is minimal throughout the various sectors of the population. The obvious examples relate to choosing a date when nomads and semi-nomads are in known and accessible places. Unfortunately, here again, African conditions may defeat such optimism, and the date which is convenient for one sector of the population may not be so for others.

E/CN.14/CPH/7
E/CN.14/CPH/7 Add.1
ST/STAT/33

XV. ELECTRONIC PROCESSING OF CENSUS DATA

ELECTRONIC PROCESSING OF CENSUS DATA

PART I

Machine Processing of Census Data

(Prepared by the Secretariat of the Economic Commission for Asia and the Far East for the Seminar on the Organization and Conduct of Population and Housing Censuses held at Bangkok, Thailand, 24 November - 1 December 1967, for the countries of the ECAFE region, and issued at ECAFE as United Nations document E/CN.11/ASTAT/SPHC/L.8, 2 November 1967.)

ELECTRONIC PROCESSING OF CENSUS DATA

PART I

Machine Processing of Census Data

CONTENTS

		Paragraphs
I.	Introduction	1 - 4
II.	Advantages of computer processing	5 - 13
III.	Need for early planning	14 - 16
IV.	Influence of equipment on planning	17 - 21
V.	System design and programming	22 - 26
VI.	Conclusion	27 - 29

I. INTRODUCTION

1. With few exceptions the results of the 1960 round of censuses of population in the ECAFE region were tabulated with conventional punched-card machines. Judging by the rate at which the statistical work in the region has become computer-oriented in recent years, it is evident that all but a small number of countries are planning to use computers for processing the 1970 round of censuses. For those few statistical offices which will not have access to a computer in their own countries, it may be possible to consider the use of computer facilities of a nearby country or of the proposed ECAFE regional computer centre. For organization and planning of the forth-coming censuses it is, therefore, assumed in this paper that electronic data processing techniques will be used almost exclusively throughout the region.

2. Considerable documentation giving guidance in the use of punched-card methods in processing census data was prepared for the 1960 population censuses. In particular, many of the principles enunciated in the informative publication "Handbook on Data Processing Methods" published jointly by the United Nations and the United Nations Food and Agriculture Organization in 1959 are directly applicable to electronic data processing. Some materials designed to aid in computer processing of censuses have also been prepared by the countries.

3. Among the series of activities planned for the 1970 round of population and housing censuses it has been proposed to hold seminars dealing specifically with the technical considerations in the application of computers to census data processing. For the ECAFE region, this is being considered for the first half of 1969. The documentation from this and similar meetings, and the services of advisers from international and other organizations with previous experience in computer processing of population census data will be available for assistance to ECAFE countries in preparing for the 1970 round. The discussion in the present paper will be confined, therefore, to general principles and aspects of planning for computerized processing of the population and housing censuses.

4. For many countries of the region the 1970 censuses will represent an initial exposure to the use of computer techniques. It may, therefore, be in order to give a brief description of the main characteristics of computer processing, particularly as applied to census work.

II. ADVANTAGES OF COMPUTER PROCESSING

5. The use of punched-cards for recording statistical data and sorting, counting and tabulating machines for compiling and summarizing the information recroded in the cards was developed originally in response to the needs for tools to handle the massive task of processing the census. The advent of the electronic computer has provided the potential for far greater speed, versatility and accuracy than the previous methods of processing census work. At the present time, electronic equipment is used extensively for mass processing of statistical programmes in virtually all modern governmental statistical agencies throughout the world.

6. In general, the principal attributes of electronic data processing which offer impressive superiority over conventional punched-card methods include:

(i) Tremendously improved speed in computations and transfers of data.

(ii) Memory storage which provides the capacity for "remembering" and retrieving at high access speeds the data stored in the memory.

(iii) The capacity for automatic execution of sequential instructions.

(iv) The capability of making logical decisions automatically during the course of processing.

(v) Considerably improved accuracy and self-checking features.

The combination of these characteristics of electronic computing equipment thus renders it possible to solve complex data-handling

problems in one continuous process at electronic speeds, as opposed to the segmented handling of data as performed in the conventional punched-card system. The capacity to process vast quantities of data and the flexibility of the new equipment represent a qualitative difference from the earlier devices that make practicable new techniques in statistical operations.

7. For example, the computer is commonly used to excellent advantage in census work for editing and correction of the data prior to or in combination with tabulation. The computer can be directed to automatically assign, or impute, values when information on personal or housing characteristics is omitted in the questionnaire or when certain information reported (or encoded) is inconsistent or impossible. Most computers are capable of applying to the data all the rules that a group of clerks would use in detecting and reconciling inconsistencies and inferring missing information. A common situation encountered, for example, is when a given characteristic is inconsistent with other characteristics reported for that person. The general procedure under these circumstances may be to draw a value of the characteristic from a distribution of that characteristic for the appropriate sub-group in the population. Thus, a person who is reported as a male relative of the household head, but for whom marital status is not reported, may be assigned by the computer a marital status from a distribution of marital status for male household head relatives in the same age group.

8. Similar procedures may be used for the detection of missing or impossible data. In many ways, the computer performs these editing tasks strikingly better and faster than when they are done by manual inspection. It will apply the editing rules individually to each person in the enumeration. It will do this with strict consistency and uniformity, and can be directed to keep a record of how often each such rule was applied, thus providing a measure of the quality of the enumerated data.

9. The use of a computer in census processing also provides the potential for development of more detailed or more informative statistics. The flexibility of the computer makes possible far greater utilization of data for derived statistics and by-products that become available for the first time. Key ratios and relationships that formerly could not be obtained without costly reprocessing of data or vast amounts of manual computation can be programmed as an incidental by-product of computer tabulations. Data can more easily be rearranged by more variables of classification and more detailed cross-classifications which, by other methods, would be prohibitive in time and cost. It is also technically possible to combine and relate census data with data from other sources concerning the same area of inquiry. Thus, it may be possible to match and combine census data with those from other sources which, for one reason or another, cannot be collected in the census. For example, vital registration data might be used for inter-censal updating of some of the population characteristics compiled in the current census.

10. Another advantage of processing census data on computers is that it often permits preparation of the final tables in suitable form directly from the machine-prepared tables for printing by offset printing processes. With earlier mechanical tabulating equipment, the results usually cannot be prepared in a form in which they can be sent to a printer. It is generally necessary to do a substantial amount of hand posting from the machine sheets and then to type the posted tables in final publication form. Preparation of the final tables by computer for direct printing eliminates the need for manuscript, type-setting, and proofreading. On this basis, the final results of the census could be released to the public many months earlier than with conventional printing methods.

11. Where physical space and storage facilities are at a premium, as is commonly the case throughout the region, computers require considerably less floor space than that required for conventional equipment. In countries using tape computer systems, the storage of census records on magnetic tape requires only a fraction of the space normally needed for card storage.

12. There are, of course, some disadvantages and difficulties involved in the transition to computers. The use of this kind of equipment brings new problems which require that the processing of census data be organized in a more systematic manner than for conventional punched-card equipment. One requirement in applying computers is that very detailed and explicit instructio must be provided to the computer before it will proceed with any operation. For example, in the machine-edit procedures, the machine must be instructed to provide for every relevant combination of entries it may encounter. Failure to consider all contingencies may result either in bringing the computer to a halt every few minutes or in rejecting large numbers of records because of some flaw in the data for which adequate provision was not made in the computer programme. For this reason, the programming and supervisory staff must be highly competent and well trained. Personnel with these attributes are difficult to find and much in demand. To attract qualified personnel it is often necessary to offer inducements in the form of premium salaries and special allowances. Preparation of physical facilities for the computer are more demanding than with conventional punched-card equipment, requiring air-conditioning, dust control, special flooring, etc.

13. Above all, it must be recognized that the availability of sophisticated data processing equipment will not, in itself, insure efficient and accurate census results. Successful computer application demands a much higher degree of precision and discipline in planning the census processing operations than was recuired with previous techniques.

III. NEED FOR EARLY PLANNING

14. The methods and equipment used in processing are directly related to the objectives and results of the census, and have a broad impact in virtually all operations. Each step in the census is closely linked with the previous operation and, in turn, influences subsequent operations. Thus, it is of paramount importance that the data processing function be represented in the planning and organization of the census from the very beginning as an integral member of the over-all census planning group.

15. The data processing representative in the planning group should, for example, actively participate in the design of the questionnaire to insure that the final design meets the requirement for efficient processing. He should assist in the development of control procedures, editing and coding specifications, the tabulation requirements, schedules, and other factors so that they may be considered with realism by the census planning authorities.

16. The introduction of the computer to census processing, by its nature, increases the time necessary for preparation for statistical processing. A great premium is therefore placed on careful advance planning. In the past, those responsible for processing the data, have, in many cases, been forced to do their detailed planning by improvisation and on the spur of the moment. This type of planning when using electronic computers is not only more difficult but perilous from the point of view of the end results. Computer programming is extremely intricate and requires planning of a high order and painstaking attention to detail. The cost, effort and time involved in changing computer procedures makes it undesirable to change those once established. Thus, if the desired time schedules are to be achieved, it requires that complete and final tabulation plans and specifications must be available to technical programmers at a much earlier date than was required by previous processing systems. The possibilities for flexibility and improvisation of tabulation plans under conditions of tight schedules are extremely limited. This, of course, suggests that over-all planning of the census should be accomplished well in advance. Many countries, even those with long histories of census taking, devote five or more years to active census planning. The United States, with 180 years of census experience, has been actively planning the 1970 census since before 1964.

IV. INFLUENCE OF EQUIPMENT ON PLANNING

17. The capacity, speed, and configuration of the computer to be used for processing the census is naturally the factor having the greatest bearing on the processing plan. Equipment may range from low speed card-operated computers with very limited core storage to installations with large storage capacities and a wide range of peripheral equipment. These considerations will

influence and, in many cases, dictate the entire approach to the census plan, including design of the enumeration form, field work, coding, input preparation editing and tabulation. The impact of equipment considerations on the whole plan for the census may best be illustrated by the far-reaching effects which the use of optical scanning techniques in the enumeration would have on all census operations.

18. For planning purposes, therefore, it is essential to know, at least in general terms, the equipment on which the census will be processed. The specific distinctions that are needed for planning are as between equipment systems operating with (i) card only; (ii) card and magnetic tape; and (iii) card, tape and optical reading devices.

(i) Card only

19. Use of computers with only card input and output capabilities will speed up the statistical compilation work as compared with the use of conventional punched-card machines; but it will actually represent a minimum change from the earlier methods. Computers provide greater computational speed and data-edit capabilities than conventional punched-card machines, but the magnitude of the key punch work in preparing the input cards remains the same as previou However, the higher speed card reading devices (as much as 800 cards per minute versus approximately 150 per minute with conventional tabulators) permit each pass on the computer to be made more quickly while the memory storage permits tabulation of more variables on each pass. The memory storage facility of these machines also provides the capability for effectively editing the recorded data for invalid codes, inconsistency and reasonableness. However, sorting, collating and other operations on the cards must continue to be done on punched-card equipment in the conventional manner.

(ii) Card and magnetic tape

20. The most common type of equipment on which the forthcoming round of censuses will be processed within the region is likely to be a medium scale system with card and magnetic tape devices, printer, and possibly magnetic disc or drum storage peripherals. After the cards have been punched, converte to tape, edited and corrected where necessary, it is possible on such installa tions to sort and tabulate the results with considerable speed. In these

cases, the cost and delay of manually punching the data before it can be transferred to magnetic tape for high speed processing in the computer must still be considered in the processing plan. The impact of such computers will be in the improved speed of computation and final tabulation of the data. The tabulation work, however, may not be begun until a "clean" tape is achieved and, therefore, special emphasis in census planning should be placed on expediting and controlling the quality of the card punching, card-to-tape, and editing operations.

(iii) Card, tape and optical reading devices

21. Optical scanning systems are automatic methods of preparing punched cards or magnetic tape records from the enumeration form and are of relatively recent development. Some countries in the ECAFE region are considering the use of these devices in the 1970 censuses. In these cases, the planning for allocation of resources will be affected by elimination of the need for large scale punching operations. On the other hand, additional resources and effort will be required to assure the quality of field work necessary to obtain a high percentage of machine-readable enumeration forms. Optical scanning may be by mark-sensing or character recognition equipment. In the latter category, although developments have been encouraging, the techniques has not yet been advanced to the point at which it may be generally considered for population censuses. Character-reading equipment which will reliably read the hand-written or non-standardized characters encountered in this work is not yet suitable for general purpose application, but equipment for reading position marking has been and can be successfully applied. Further experience with these techniques, in areas where the field work is more amenable to control, is desirable.

V. SYSTEM DESIGN AND PROGRAMMING

22. Considerable system design and computer programming work will be required for the machine-edit and tabulation phases of the census. There is a general tendency in all countries to underestimate the magnitude of the effort required to programme and install computer systems. Because of the interaction of computer processing with virtually all other elements of the census,

it is essential that the computer programmes be prepared and fully tested beforehand. It is not sufficient for the programmes merely to be ready by the time the main processing is scheduled to begin.

23. The more difficult part of the programming lies in the machine editing procedures, that is, in having the computer check for inconsistencies such as improbable code combinations. Translating the editing rules into steps sufficiently simple and precise for the computer to follow is usually more difficult than anticipated. It is of particular importance to check out these programmes well in advance using realistic data. Machine editing requires that non-responses and inconsistencies be detected and corrected by having the computer impute the corrected data in so far as possible. Early trial runs of the editing programmes will disclose other possible phases of the census operation which are the likely sources of editing problems. As a result, the need for corrective action in such areas as the enumeration form, enumerator instructions, coding, key-punching or supervision may be brought into prominence.

24. If thorough and early testing of the complete data processing phase is not conducted, it is likely that during the actual operations the computer will reject excessively large numbers of records with one defect or another. This, in turn, creates the problem of reinstating the records that have been rejected and corrected as a result of the editing operation. The effect is likely to be unexpectedly time consuming, requiring more computer time than originally planned for.

25. Pre-testing of the entire range of computer operation, that is, the edit and tabulation programmes, will also provide much better advance estimates of the computer hours needed. Estimates based entirely on theoretical factors have often proved to be greatly in error. This consideration is important where an installation is heavily loaded and precise scheduling is necessary.

26. A great deal of time is required for programming for a computer and the lack of experienced staff is a common limiting factor in new computer installations. It requires trained personnel, and their training may have to be as much as one year for adequate proficiency. Thus, it is especially

important that sufficient time be allowed in the census plan to permit the training of programmer staff early enough for completion and testing of the edit and tabulation programmes prior to the census date so that full-scale processing can begin immediately after the enumeration.

VI. CONCLUSION

27. The increased power and versatility of the present generation of machines over their predecessors will give those statisticians processing the 1970 census substantial advantages as compared with 1960. This equipment will introduce a marked improvement in the census procedures of the countries of the region, as did the introduction of punched-card equipment years ago. For efficient use of the equipment, the design of the processing needs to be integrated with the other aspects of the over-all census design. Integration of the capabilities of computer processing into the census operations as a whole entails a balance in supervision, quality control, editing, and the capacity to rectify detected errors. Imbalances may lead to serious problems.

28. It is urged, therefore, that the computer programming be completed well in advance of actual processing and that very thorough testing be performed on data obtained through pilot-tests. Though the outcome of this possibly may be an unchanged computer system design, it may well lead to strengthening of supervision and quality controls in the key stages of the census.

29. With the general application of computers to the 1970 round of population censuses and with careful planning of all stages of processing, administrators and planners of the censuses in the ECAFE region may look forward with reasonable confidence to reducing the interval between collection of data and publication to at least half the time required for the 1960 censuses.

ELECTRONIC PROCESSING OF CENSUS DATA

Part II

Some Principles of Computer Processing of Census Data

(Prepared by

the Statistical Office of the United Nations)

ELECTRONIC PROCESSING OF CENSUS DATA

PART II

Some Principles of Computer Processing of Census Data

CONTENTS

		Paragraphs
A.	INTRODUCTION	1
B.	OPTIMUM USE OF CAPACITY	2 - 8
C.	THE DETECTION AND CORRECTION OF ERRORS	9 - 21
D.	PRODUCTION BY COMPUTER OF FINAL TABLES FOR PHOTO OFFSET	22 - 24
E.	SUMMARY	25 - 26

ANNEX. Example of tabulation on a computer

A. Introduction

1. Almost all censuses in the 1970 round - and afterwards - will be processed by computers. In some cases, the input-output operations will be highly automated through the use of mark readers or optical character readers for input operations, and electronic photo-composers for out-operations. Some principles of the processing technique will be discussed here with a special view to explanations and guidelines needed for those countries that are about to enter this field for the first time.

B. Optimum use of capacity

2. Common to all computers is a central processing unit, hereafter referred to as a "CPU". The CPU contains the central memory, that is the device where the programme to be executed is stored together with the data to be manipulated (the input data, the intermediate data) and the final data resulting from the manipulations (the output data). Attached to the CPU is a configuration of input-output devices, or so-called peripherals, such as units for reading or punching cards or paper tapes, magnetic tape drives, direct access storage devices (magnetic disk, drum units, etc.), optical character readers, printers, etc.

3. Handling of data inside the central memory is much faster than reading, punching and writing of data on the peripherals. It is good programming technique to arrange, as far as possible, for the central memory to be kept busy and not idling, waiting for a record to be read in or written out. Reading and punching of cards or paper tapes, as well as reading marks or characters, is much slower than reading and writing magnetic records (on tapes, disks, etc). The general practice is, therefore, to write input-data, that have to be used more than once, on magnetic tapes or other magnetic storing media. There are even cases where a special, smaller, computer is used for such comparatively slow operations as "card to tape", "tape to card" and "tape to print", leaving the bigger computer to communicate with the "outer" world with

magnetic tapes only.

4. In tabulation of census data, the standard method is to use part of the central memory for the simultaneous creation of a set of different tables (see annex). In certain cases, areas can be reserved in the central memory for all required tables, and if so, only a single input operation for each record is required (for the tabulation) and no sorting at all of the input records is needed. If there is not space enough in the central memory for the tabulation to be made in one run, the following methods can be used, separately or in combinations:

(1) Reruns, without sorting, of the input-records, with modified programmes, until all basic tables are created. A table is said to be basic if it is not a sum of other tables.

(2) Sorting the input records before and/or between the different tabulations. In the extreme case, sorting can be used so that only a single area in the central memory is required for the tabulation. This single area is then used for the accumulation for one table-cell at a time. The extreme case has been mentioned here only to illustrate the thesis that the larger the memory, the less sorting -- the smaller the memory, the more sorting.

(3) Attaching a so-called Direct Access Storage Device, a DASD, to the central memory. On a DASD, records can be accessed directly, as in the central memory, rather than serially, as is the case with magnetic tapes. Magnetic disks, drums and even magnetic strips or cards, such as the IBM Data Cell Drive or the NCR Card Random Access Memory (CRAM) are examples of DASD.

5. Before further discussion of the three methods, the proportions of a typical census case will be given here. In the United Nations "Principles and Recommendations for the 1970 Population Censuses" there are 22 tables ranked as "first priority". The total number of locations, or cells, in these tables is around 5,000. In this count, no totals are included. For instance, the requirements for Table 7 is taken as 40 and not 63, as is required when the marginal distributions are included (for explanation of "marginal distribution" see the annex). Many of the tables will appear on the lowest level in the hierarchy of geographical subdivisions. It can therefore, easily be seen that the required number of table-locations, taking into account the geographical distribution, might run into the millions or even tens of millions. If such cases have to be dealt with, without sorting and rerunning, the only way is to use a kind of DASD. A census, where the material is collected and prepared in an order that has no relation whatsoever to the tabulation order, might be a case for a DASD. Generally speaking, a DASD offers an attractive solution where the input material as a whole has to be taken as a unit for processing and where no tabulation can be finalized before the very last input record has been prepared. In many cases, however, the way in which the census-data are collected results in a useful presorting. This is normally the case in a de facto population census. The data for an enumeration area can be processed, independently of the data for other areas, as soon as they have been prepared. This is called batch-processing. Assuming the above-mentioned recommended tables, only some 5,000 locations are needed for a complete tabulation in one run. If an additional presorting according to, for instance, sex and age is made, the required capacity comes down to the order of a few hundreds. (The capacities quoted here are for the tables only, and are in addition to the requirements for the actual programme and a possible supervisory programme.)

6. The three methods can, as already mentioned, be combined. It should also be noted that a DASD can play a very important role, even if it is not large enough to store all required tables simultaneously.

7. Storing in a DASD is slower than storing in the central memory. It might, therefore, be economical in some cases to programme so that the most frequent cases are tabulated in the central memory and the less frequent in the DASD. The computer can be programmed to find out where the tabulation should take place in order to minimize the processing time. Input data for which no destination is reserved in the central memory or in a DASD (because no DASD is attached or its capacity is surpassed) can be "dumped" on a magnetic tape for future processing. In some cases, it might even pay its way to dump excessive data on punched cards. Inherent in the dumping technique is the possibility of reducing the volume of input data from one run to another.

8. The tabulation of an input record consists of finding, for each table, whether any data from the record should be added to any location in the table. If data have to be added to the table, the problem is to find the serial numbers of the locations within the table that have to be "updated". When the numbers have been found, the additions take place. A serial number calculated for a table must, of course, fall within the limits of the table. One way of ensuring this would be to reserve for each table, a residual group where the records with erroneous codes could be added. This would, however, create chaotic discrepancies between tables and would not solve the problem in the case where it is not clear whether a record belongs to a table or not. The method normally chosen is to ensure "processability" through a special checking programme. <u>This imposes a kind of precision in processing that goes far beyond what is statistically required.</u>

C. The detection and correction of errors

9. Precision is achieved through volume control and editing. The standard method is to let the enumerator summarize - on a checklist - a few basic data from the questionnaires, such as number of persons, households, living quarters, industrial enterprises, agricultural holdings, etc. By summarizing the checklists throughout the hierarchy of geographical subdivisions, two things are achieved. Firstly, a few, provisional, statistical data are quickly made available. Secondly, a framework is created for checking completeness and uniqueness. This is the volume control and it is of undamental importance throughout all the handling of material and data.

10. Editing consists of two phases: the detection of errors and the correction of errors. Some editing is, and must be, performed manually. However, experience has shown that manual editing is not efficient enough. Furthermore, the records must be error-free as they appear in the central memory. Errors might be introduced in punching, in mark sensing or in optical character reading. The computer can be programmed for the first phase only or for the two phases combined. In the first case, the computer will print a list of the erroneous records and the character of each error. This procedure might be combined with rejection of the erroneous input records. The corrections will, in this case, be made manually. In the second case the appearance of an error will initiate an automatic correction procedure, which is either deterministic or stochastic. A deterministic procedure is one where a given condition defines a unique value. For instance, the rule always to set sex as female whenever a number of children born alive is reported, is a deterministic procedure. A stochastic procedure is one where a given condition defines a frequency distribution of possible values for an erroneous (or missing) code. The actual value is determined by a random process. For instance, a rule to determine a missing age by random drawing of a number with a given distribution, is a stochastic process.

11. The correction procedures must, of course, be constructed so that no new errors or inconsistencies are introduced. A simple way to formulate the stochastic rules is to assume that missing or erroneous data are distributed as the correct data. But this is one of the more questionable methods to use. Errors sometimes appear in a very systematic way. A coder might constantly misunderstand certain answers (perhaps because of language differences) or he might have memorized certain codes incorrectly. People of one religion might be more unwilling to state their religion than people of other religions. It is a very tricky problem to formulate the rules for automatic corrections. But once it has been done, the gains are considerable. Verification of coding and punching can be skipped or reduced to sample operations.

12. Automatic corrections must be under constant statistical control. The editing programme must report the number of errors by type by coder by punch operator for each batch, and from time to time the prerequisites for the stochastic rules must be checked.

13. Not all errors can be corrected automatically. For instance, errors resulting from misplacing of questionnaires are of this nature.

14. Editing is not always limited to the individual record per se, its codes and the interrelations between its codes. Editing might cover the interrelations between records of the same group, such as records for all members of a household. Editing is sometimes combined with the creation of summary records as well as with transfer of data from master records to detail records or transfer between matching returns in different censuses.

15. Editing is a prerequisite for the functioning of the tabulation-programme. This does not mean that editing only before the tabulation starts is a sufficient procedure to ensure acceptable tables, nor does it mean, for practical reasons, that

all editing must be done before the tabulation is started. The following two examples will clarify the two points.

16. Example 1. In a census of population, a question about the kind of diploma of vocational training was asked. The answer "Driver's licence" was given the code for "Other diploma". In the individual record per se the code for "Other diploma" was acceptable. In the tabulation, the high frequency for "Other diploma" raised doubt and resulted in the detection of the error.

17. Example 1 shows that a study of the proportions within a table can reveal errors that cannot be, or have not been, detected in the individual record per se. The computer can be programmed to analyze the tables and to signal cases to be investigated.

18. Example 2. In a census of population the tabulation plan might have a table, for the country as a whole, with complete cross-classification of industry by occupation. The capacity might very well be sufficient for testing the industrial and occupational codes independently of each other on the batch level but insufficient for a practical and economical testing of the combination of them on the same level. By excluding the cross-classification from the original testing and tabulation, the batch-processing might be undertaken with a fair degree of accuracy. When the batch-processing is over, the cross-classifications can be sorted and tested against a file with permissible combinations. After corrections, the countrywide table can be produced.

19. Example 2 raises the question of whether the errors detected in the second round of editing will affect the tables made directly after the first round of editing. If the end of the tabulation has been reached, the rigorous requirements set by the computer are no longer at hand and the problem is reduced to one of statistical significance; if any possible decision based upon the tables can be reasonably suspected to be influenced by an additional round of corrections, the corrections have to be made; otherwise they need not be made.

20. Corrections can be done systematically by running the whole tabulation programme with the erroneous input records as negative components and their possible replacements as positive components. The tables, consisting only of the balances between the two components are then used for updating of the table file. It is thus not necessary to rerun all input records in order to make the corrections.

21. The basic tables, that is, the tables created directly from the input records for the individual cases, are in their turn added together to new tables in one or more hierarchies. The editing of tables is not limited to the basic round; it has to be made after each step in the aggregation. The editing of the tables is not only a hunt for errors, but it is, as a matter of fact, an advanced analysis and interpretation of the statistical material.

D. Production by computer of final tables for photo-offset

22. It is well known that the computer can be used for printing tables with headings, stubs, page numbers, etc., in a form that can be used for reproduction by the offset method. There are, for certain computers, character sets available for simultaneous printing with different alphabets, such as Latin-Greek, Latin-Arabic, Latin-Hebrew etc. This means that bi-alphabetic tables can be printed in one run.

23. Perhaps less known is a method called electronic photo-composition. In one such system, now in use, a computer is programmed to direct a beam of light through a selected image of a typographical character, thus projecting the character on a photographic film. In another system, now being developed, the computer is programmed to direct the beam in a TV-tube to compose a page on the screen. The screen is photographed and the film is, as in the other system, used for offset reproduction. A big variety of type-fonts are available and the results can hardly be distinguished from ordinary printing. The advantages with the automatic production of the tables are: higher speed, lower cost (at least in developed countries) and higher accuracy than with conventional methods It has previously been said

that editing is a prerequisite for the tabulation programme. Editing is also a prerequisite for the automatic production of tables.

24. The edited results of the tabulation programme are normally not completely reproduced. The reasons are many. For convenience in processing, some tables are created for the smallest geographical sub-divisions but are meaningful only higher up in the hierarchy. Some errors can only be found but not corrected, resulting in scrapped tables. Some tables might contain all possible entries, including those for which no amounts are reported and those entries are not to appear in the printed tables. The selection and modifications of the tables to be printed are to a certain extent governed by the results of the tabulation and can therefore not be fully planned in advance.

E. Summary

25. At the center of the processing of census data stands the tabulation. In tabulation, the computer is used as an extremely fast accounting machine with a large number of registers. Verification of coding and possible punching can to a considerable extent be taken over by the computer. The computer can read not only punched cards or papertapes but also marks and characters (typed or handwritten). The computer can perform the first phase of the editing: the detection and reporting of errors. To a certain extent, the computer can also perform the second phase of the editing: the corrections. Every second step in the processing is an editing step, where the results of the previous step are analyzed and the way for the next following step is cleared. Interwoven is a volume control to ensure that no material is mistakenly processed more or fewer times than required. The computer can be programmed to prepare the source for offset reproduction of the tables.

26. The degree of automation a country should apply must be judged from case to case. Generally speaking, the computer must have capacity enough to keep pace with the coding and other preparations of the

input records. The clerical operations and the coding system must be under continuous control by the computer through a kind of "early warning system". The returns might show unexpected characteristics that call for changes in the coding system. All of this means that the computer and the programmes must be in operating shape when the census is taken. This, according to experience, means that the preparations for the processing must start between 12 to 18 months before the census day. If a pilot census is taken it is preferable to let it go through the processing phase. This means that the reference day for the start of the preparations will be the pilot census day, or approximately so. Ample time for the development of the processing system is of greatest importance. The processing aspect must be given attention from the very start of the census planning and throughout its development.

Example of Tabulation on a Computer

In the central memory of the computer, 400 words (locations, areas, cells) numbered 4200 thru 4599 are (in this example) available for tabulation. Each of these words can be used as a counter. In this example, the available part of the central memory will be used for the creation of:

Table 1: Population by single year of age and sex

Table 2: Population by marital status, age and sex

As can be seen below, there are 102 age-codes, 2 sex-codes and 5 codes for marital status. Thus, 204 words are needed for Table 1, and consequently the words 4200 thru 4403 have been reserved for Table 1. For Table 2, the ages form 15 age-groups. The required number of words is therefore 150 and the words 4404 thru 4553 have been set aside for Table 2. Observe that words have been reserved for the basic parts of the tables only and not for any total or so-called marginal distribution. This concept is illustrated in the following version of Table 2. The shaded areas represent the marginal distributions.

Age group	Male						Female						Total					
	Marital status						Marital status						Marital status					
	1	2	3	4	9	TOTAL	1	2	3	4	9	TOTAL	1	2	3	4	9	TOTAL
less than 15																		
15-19																		
20-24																		
....																		
75 and over																		
Not stated																		
TOTAL																		

The words 4554 thru 4599 are not used. For each input record, a "one" has to be added to the proper word in the area reserved for Table 1 and to the proper word in the area reserved for Table 2. The number of the proper word, that is the address, is a simple function of the codes for sex, age and marital status, but not of the external version of the codes but of an internal version. The external version is the one used in coding and punching the questionnaires. The internal version is calculated by the computer. The internal version may be the same as the external and the internal version may vary from table to table, as illustrated below. In this case the two functions are:

Address for Table 1 = (sex - 1). 102 + (age - 1) + 4200

Address for Table 2 = (sex - 1). 75 + (age - 1) .5 + (marital status - 1) + 4404

The different versions of the codes are as follows:

External code for age	Internal code for age, Table 1	Internal code for age, Table 2
00 for age below 1 year	001 for age below 1 year	01 for age below 15
01-99 for ages 01-99	002-100 for ages 01-99	02 for ages 15-19
XX for age 100 and over	101 for age 100 and over	03 for ages 20-24
YY for age not stated	102 for age not stated	.....
		14 for age 75 and over
		15 for age not stated

External code for sex = Internal code for sex

1 for male

2 for female

External code

External code for marital status	Internal code for marital status, Table 2
1 for single	1 for single
2 for married	2 for married
3 for widowed	3 for widowed
4 for divorced	4 for divorced
9 for not stated	5 for not stated

Sample computation of addresses:

	External codes			Internal codes and addresses						
				Table 1			Table 2			
	Age	Sex	Mar. status	Age	Sex	Address	Age	Sex	Mar. status	Address
Record 1	21	1	2	22	1	4221	3	1	2	4415
Record 2	YY	2	9	102	2	4403	15	2	5	4554
Record 3	75	2	4	76	2	4377	14	2	4	4547

On the following page, the available part of the centrab memory is illustrated. The words where a "one" has to be added in the tabulation of the above three records contain the symbol "plus 1".

	0	1	2	3	4	5	6	7	8	9	
420	START OF TABLE 1										
421											
422		PLUS 1									
423											
424											
425											
426											
427											
428											
429											
430											
431											
432											
433											
434											
435											
436											
437								PLUS 1			
438											
439											
440				PLUS 1	START OF TABLE 2						
441						PLUS 1					
442											
443											
444											
445											
446											
447											
448											
449											
450											
451											
452											
453											
454								PLUS 1			
455				PLUS 1	START OF UNUSED AREA						
456											
457											
458											
459											

E/CN.14/CPH/4
ST/STAT/19

XVI. METHODS OF EVALUATING THE RELIABILITY OF POPULATION AND HOUSING CENSUS DATA

(Prepared by the
Statistical Office of the United Nations)

METHODS OF EVALUATING THE RELIABILITY OF POPULATION AND HOUSING CENSUS DATA

CONTENTS

	Paragraphs
I - INTRODUCTION	1 - 3
II - ERRORS IN CENSUS ENUMERATION	4 - 22
A. Errors of coverage	4 - 20
1. Omission or under-enumeration	6 - 15
a. Omission of living quarters	7 - 18
b. Omission of households	9 - 10
c. Omission of person within households	11 - 14
d. Other likely omissions	15
2. Duplication or over-enumeration	16 - 18
3. Erroneous inclusions	19 - 20
B. Errors of content	21 - 22
III - GENERAL CONTROL OF CENSUS PROCEDURES	23 - 24
IV - METHODS OF EVALUATION	25 - 78
A. Objectives of an evaluation programme	26 - 28
B. Direct methods of evaluation	29 - 63
1. Field Checks	30 - 53
a. *Ad hoc* post-enumeration sample field checks	31 - 49
b. Labour-force surveys	50 - 53
2. Records from other sources	54 - 63
a. Population registers	58 - 59
b. Birth registers	60 - 63
3. Internal checks of duplication	64 - 65
C. Indirect methods of evaluation	66 - 79
1. Total population	69 - 74
2. Selected attribute distributions	75 - 79
V - CONCLUSION	80 - 82

ANNEX. National Reports on the Methodology of Ad-hoc Post Enumeration Sample Field Checks.

I. INTRODUCTION

1. It is an accepted fact that, even though the programme for a population census or a housing census makes every provision to guarantee the completeness of the enumeration and the accuracy of the statistical results, there will always be some errors. It is also recognized that, although the census results cannot be 100 per cent accurate, they are still of value if the errors are kept within reasonable bounds.

2. While the existence of errors is accepted, therefore, there is also a growing realization of the need to evaluate the results of all statistical inquiries in order to determine the magnitude of the error. The importance of census results and their influence on economic and social development programmes makes it imperative to ascertain the degree of reliability of the information they contain, so that it can be effectively used. The results of a census cannot be regarded as reliable if they are not properly checked, nor can they be used with confidence for purposes of study and analysis unless they have been previously subjected to some kind of verification. Furthermore, an understanding of the sources and causes of error in the census data is of great value for the improvement of future censuses.

3. In any census, errors can occur at the time of enumeration and during the processing of the raw data. The possibility of error at either stage can be greatly reduced by the application of sound principles of census taking. Some obvious enumeration errors can be corrected during processing as can most processing errors. Many enumeration errors, however, cannot be detected at this point. It is generally assumed, therefore, that the bulk of the errors in the results originates during the enumeration. Accordingly this paper is devoted primarily to the causes of enumeration errors and their evaluation.

II. ERRORS IN CENSUS ENUMERATION

A. Errors of coverage

4. Coverage, or quantitative errors are those which affect the total population figure or the total figure of living quarters and also the distribution of the population and/or living quarters among the geographic divisions of the country. In regard to population, they may also affect the data on such characteristics as sex, age, occupation and national and/or ethnic group if the coverage errors occur primarily within one sex or within particular age, occupation or ethnic groups. Even further, the errors in the enumeration of a particular group of the population may distort the results for many other characteristics, if the group is markedly different from the remainder of the population in regard to these characteristics.

5. Errors of coverage are attributable to (1) omission or under-enumeration (2) duplication or over-enumeration and (3) erroneous inclusions.

(1) Omission or under-enumeration

6. Omission, or errors of under-enumeration, occurs where persons and/or living quarters which should have been included in the enumeration have been omitted because of deficiencies in some of the census operations or because of other factors (political, social, economic, climatic, etc.) that may give rise to unforeseen problems. Omissions may occur in respect of living quarters (and, hence, all of the households therein), households (and, hence, all of their members), individuals in households and particular groups of the population or of living quarters as a whole.

a. Omission of living quarters

7. Omission of living quarters results from failure to cover the entire territory of a country or to identify all of the living quarters in that part of the territory which is covered.

8. Failure to cover the entire territory of a country may take place where the cartographic material is incomplete or inaccurate, so that certain areas are overlooked, where the census staff is inexperienced, as may happen where there has been a long inter-censal interval or where a first census is being taken, and where the access to some areas is hampered by natural obstacles such as high mountains, extensive forests, deserts, swamps or large bodies of water. Such omissions may also occur in the case of areas that can be easily overlooked and of living quarters in apparently uninhabited structures.

b. Omission of households

9. The omission of households, whether one-person or multi-person is directly and primarily due to the omission of the living quarters referred to in the preceding paragraph. Other households likely to be omitted are those having more than one place of residence, those living in a boarding house or hotel, one-person households of which the single occupant is usually working away from home during the day, those which have changed residence during the enumeration period, those without fixed places of abode (such as those accustomed to sleep on the streets or in other places not constituting living quarters) and those in transit at the time of the census.

10. There is, in addition, the possibility that the respondent or enumerator, not fully understanding the definition of a household, will fail to make the proper distinction between two or more households occupying a single set of living quarters. Although omissions of the latter kind do not affect the total population figures, they do have an important bearing on the total number of households revealed by the census.

c. Omission of persons within households

11. These omissions result from ignorance, forgetfulness or an inadvertent or deliberate oversight on the par of the respondent, and from the improper application by the enumerators of the prescribed definitions.

12. A paramount factor in omissions of this kind is how the place of enumeration is defined for the purposes of the census, i.e., whether persons are to be enumerated at the place where they usually reside or at the place where they happen to be at the time of the census. The duration of the enumeration is also of fundamental importnce, for the longer the period of enumeration, the greater the possibilities of error, specially if persons are to be enumerated where they are found at the time of the census.

13. Among the individuals who are likely to be omitted are those who have transferred from one household to another during the enumeration period, regardless of whether or not this involves a change in residence, and also those in transit. Furthermore, when persons are enumerated at their usual place of residence, there may be a tendency for the respondent to fail to mention temporarily absent members of the household, such as those in hospitals and prisons, and those away on short visits, who the respondent does not realize should be included as members of the household. In some instances, there is evidence of deliberate failure to mention certain individuals, such as young men of military age, whose existence it is desired to conceal from the authorities. Some countries have found large omissions of women who are kept in seclusion within the household and of domestic servants who sleep in a structure separate from the main dwelling of the household being enumerated.

14. Special mention should be made of the likely omission of infants and young children of either sex, the under-enumeration of whom is a common occurrence in developed as well as developing countries. In some of the latter, this may result from a superstitious fear that mentioning a child will bring it bad luck.

d. Other likely omissions

15. Apart from persons in those groups already mentioned, there are other groups of persons who are particularly apt to be omitted in a census. These are nomads and other groups of persons living in remote areas, of whom the census officials may have little exact knowledge. Both the persons and their living quarters are likely to be overlooked. It cannot be over-emphasized that these groups, as well as those mentioned earlier, must always be given special attention in the original enumeration.

(2) Duplication or over-enumeration

16. Errors of duplication, or over enumeration, result from the inadvertant or deliberate inclusion of persons and/or living quarters more than once, as a result of which the totals appear to be greater than they really are.

17. The most common causes of over-enumeration are overlapping of enumeration districts because of inadequate cartogrphic-work, misinterpretation or ignorance of the census definitions concerning who is to be enumerated, and an excessively long enumeration period which makes it difficult for the respondents to remember, and the enumerators to elicit, the actual state of affairs on the date designated as the census day. In some cases, in addition, countries have found that enumerators who are paid according to the number of persons or of living quarters they enumerate will tend to inflate the figures by adding fictional persons or living quarters.

18. The duplication of persons is most apt to occur in the case of those who move immediately after being enumerated, those travelling at the time of the census, patients in hospitals, or similar institutions, students living at their schools, persons working away from home and persons temporarily staying in a boarding house or hotel. The first group of persons may be enumerated twice because they fail to inform the enumerator at their new residence that they were previously enumerated at their old residence. Persons in the remaining groups may be enumerated both at their usual residence and at the place where they are found at the time of the census.

(3) Erroneous inclusions

19. In addition to the errors of omission or duplication mentioned above, there may be certain erroneous inclusions. Although duplications are also erroneous inclusions, the cases dealt with here concern enumerations which are defective in form or which, according to the census definitions, should not be made at all.

20. A distinction can be drawn between those cases which affect the national population total and those which do not. Cases which affect the total population figure are the incorrect inclusion of children born between the census date and the time of enumeration and of persons who died immediately prior to the census date. Cases which do not affect the total population figure, but will affect the geographic distribution of the population, are the inclusion of persons in the wrong enumeration district and their corresponding exclusion from the appropriate enumeration district.

B. Errors of content

21. Errors of content include mistakes in reporting and/or recording information concerning the characteristics of living quarters, households and individuals. These are errors which affect the quality of the results, as contrasted with quantitative errors, which affect their magnitude. They include inadvertent or deliberate misstatement of fact, as well as complete absence of response about some characteristics.

22. The causes of such errors include ignorance of the facts, misunderstanding of the questions, deliberate misstatement and carelessness. Ignorance may be due to forgetfulness on the part of the respondent or to the fact that information is furnished by children, neighbours and other improper respondents who are not aware of the facts. Misunderstanding may result from the inclusion of difficult questions, the use of a poorly designed questionnaire that is hard to understand and manage, or the employment of improperly trained or poorly educated enumerators. Deliberate misstatement may be made in an effort to conceal certain facts, such as age, or to enhance the status of the respondent, as by attributing to him an occupation which is considered of a higher status than his actual occupation.

Carelessness on the part of the enumerator may result in mistakes in entries for which the respondent supplied the correct information. This would include, for example, transposing of digits for the entry on age or checking the wrong entry in a series of pre-coded replies. Erroneous entries may also result from failure of the enumerator to pay sufficient attention to the replies received so that he does not record adequate details on such topics as occupation or educational attainment.

III. GENERAL CONTROL OF CENSUS PROCEDURES

23. The conscientious application of certain procedures which are a technical and administrative concomitment of census data-collection and processing can eliminate a certain amount of error and inconsistency which impair the completeness and quality of the data collected and tabulated. These procedures may be briefly summarized as (a) pre-censal testing of procedures and material; (b) control and correction in the field by enumerators and supervisors to spot errors and omissions and to correct them by re-enumeration; (c) strict control of the receipt and handling of questionnaires in the office, including a careful review to ascertain that all parts of the country have been covered by the enumeration; (d) careful editing of questionnaires for the correction of obvious inconsistencies and for the completion of incomplete questionnaires by the addition of missing data about which logical assumptions can be made, such as the allocation of sex on the basis of other entries on the questionnaire, where sex has not been recorded; (e) quality control of coding and of transfer of numerical data to data-processing equipment; and (f) the utilization of checks of internal consistency and of correction techniques built into the machine-processing procedures.

24. These procedures cannot, however, contribute, except indirectly, to an evaluation of the reliability of the census returns. The principal methods of evaluating reliability are discussed in part IV.

IV. METHODS OF EVALUATION

25. Several methods have been used for evaluating both the quantitative and the qualitative accuracy of census results. Since no single method can produce a perfect measure of accuracy, countries should take advantage of as many methods as their resources make feasible.

A. Objectives of an evaluation programme

26. The basic objectives of a complete evaluation programme are: (1) to determine the level of accuracy of the census results in regard to their coverage and content, (2) to indicate to the users of the data the limitations of the results and, if possible, to correct some of the errors; and (3) to determine the causes of errors and the characteristics of the living quarters, households and persons involved.

27. The first two objectives provide the information concerning the reliability of the published data which are needed by the users of the statistics. This not only makes the data easier to use but may also increase public confidence in the results. The third objective is conducive to improvements in the methodology of future censuses and other statistical inquiries, and thus to the solution of some of the difficulties inherent in the compilation of social and economic statistics. It should be noted, however, that it is usually only the total population figure and the total living quarters figure which can be corrected. To attempt to correct the tabulated results for particular characteristics is an impossible task because most census tabulations involve the cross-classification of two or more characteristics. Therefore, although the estimated corrected total population figure should be shown in the census publications, the detailed tabulations will of necessity be based on only the enumerated population and/or living quarters.

28. Just as a census as a whole is not complete until the information collected is made available to potential users in a form suited to their needs, so the evaluation programme is not complete until its results are set forth clearly for the users of the census data and its technical and methodological aspects (including the difficulties met with) are explained both for the users of the data and for the technicians who will prepare the next census.

B. Direct methods of evaluation

29. The direct methods of evaluation are those which involve the checking of census returns against independently obtained records to measure the degree of consistency between the two sets of records. The direct methods which have been used are: (1) field checks, including ad hoc post-enumeration sample field checks and labour force surveys, (2) comparison with other statistics such as those from registers of population, housing, births, deaths, military personnel, etc., from school enrolment lists, from electoral rolls and from tax rolls, and (3) internal checks of duplication. The extent to which any or all of these methods can be used in an evaluation programme will depend on the resources, statistical development and other characteristics of the country, the purpose of the evaluation and the possibilities offered by each method.

(1) Field checks

30. Field check procedures are those which involve the comparison of records from the census with records independently obtained by field methods.

a. Ad hoc post-enumeration sample field checks

31. Theoretically, any independent random sample survey of living quarters and/or households should yield records adequate for matching against the records from the census provided that the sample frame is such that the unit of enumeration is the same in each investigation and that living quarters, households and individuals which failed to be enumerated in one or the other inquiry can be identified.

32. The most efficient method is replication of the census under intensive conditions of precision. This replication is known as a post-enumeration sample field check which is defined as the independent re-enumeration of a representative sample of the population in such a way as to provide a reliable measure of (a) the number of living quarters, households and persons omitted from, or erroneously included in, the original census

count and (b) the magnitude and nature of response errors[1]. Some countries have, at the same time, collected additional data which could not be included in the original enumeration but this procedure is not an integral part of the check.

33. It must be remembered, of course, that the difference between the census results and those of the post-enumeration check do not represent an absolute error in the former because the latter will also be subject to some response error, no matter how carefully it is planned and administered. Nevertheless, the post-enumerations check is likely to be one of the best methods of evaluating the accuracy of the completeness and content of the census results.[2]

(i) Criteria to be met

34. To accomplish its purpose, a true post-enumeration field check must meet three specific criteria, namely, (a) be independent of the original enumeration (b) be representative of the whole country and all population groups, and (c) invove one-to-one matching of records to produce an identical group from each investigation.

35. To achieve independence in the post-enumeration field check, the ultimate sampling units for re-enumeration should be clearly defined and operationally convenient geographic areas. They should be chosen not from a list of living quarters, households or names enumerated at the census being evaluated, but from a comprehensive frame consisting of geographic area units, preferably the enumeration districts used in the original census, the probable population-size and/or number of living quarters of which is already known and within which every set of living quarters and/or every household will be re-canvassed.

1/ It has sometimes been necessary, because of indications that the census enumeration has been very seriously deficient in particular parts of a country, to carry out a complete re-enumeration in these areas. Since the results of a re-enumeration are substituted as a whole for those of the original one, they do not serve as a post-enumeration check and should, ideally, be evaluated in the same way as are the results for the remainder of the country.

2/ For detailed descriptions of recent ad hoc post enumeration sample field checks in the Republic of Korea, Canada, Chile, Costa Rica, France, France, Guatemala, Honduras, Panama, United States and Uruguay, see the references listed in the Annex.

36. Representativeness in respect of the whole geographic area and all population groups and/or living quarters is also an important criterion. Recommended procedure is for the post-enumeration field check to be carried out in a well-designed random cluster sample, small enough to minimize cost but large enough to give at least an evaluation of the enumeration in the country as a whole and in its major civil divisions, since determination of quantitative measures of under-and over-enumeration is one of the important objectives of the procedure. Moreover, the random cluster sample will provide a scientific basis for future investigations and, for this reason alone, it would be desirable to employ random sampling.

37. The one-to-one matching of census schedules to post-enumeration field check schedules is the essence of the re-interview method of census evaluation. Name-to name matching is not an easy task in any country and in cultures where names lack uniqueness it may be extremely difficult. Even identification of the living quarters and the household may be difficult in those parts of the world where street names and house numbers are practically non-existent Nevertheless, since the efficacy of the post-enumeration field check as a remedial measure in census taking rests upon the identification of errors and their correction, comparisons must be made on a one-to-one basis to the degree possible. Naturally, comparison can be made simply between gross numbers enumerated in the census and those revealed in the post-enumeration check but the possibility of compensating errors might invalidate some conclusions based on evidence derived in this way.

38. Because of the importance of matching, some means of facilitating identification of persons, households and living quarters enumerated in both inquiries must be found. The most promising approach appears to be the mechanization of the process, using high speed electronic computers to effect the match.

(ii) Other requirements

39. Since the post-enumeration field check is an additional enumeration, it should be provided for in the census regulations so that its scope and its confidentiality are clearly defined, as are the legal obligations of the public and the enumerators. In order to avoid possible adverse reaction ~~to a re-enumeration of~~ only a part of the population, the census publicity should explain the purpose of the field check, the manner in which the sample will be selected, the approximate timing of the re-enumeration and the importance of public co-operation.

40. Among the technical aspects which must be kept in mind are: (a) the necessity of conducting the check as close as possible to the date of the original enumeration so that there will be a minimum of differences resulting from births, deaths, marriages, population movements, construction or demolition of living quarters and structural changes in living quarters which occurred in the interval, respondents will still have in mind the information as of the date of the census, and the population will be prepared to collaborate with the new inquiry because it has not forgoten the publicity and other activities connected with the census, (b) the advisability of selecting the best of the supervisors and enumerators from the regular census to serve on the check and ensuring that they receive a most adequate and vigorous training, (c) the importance of using as respondents responsible persons, preferably the head of the household or the individual about whom the questions are asked, even if this means repeat visits; (d) the necessity of using a questionnaire designed to elicit the most exact responses and to simplify the subsequent matching procedure; (e) the need for exact determination of enumeration areas to ensure the representativeness of the sample chosen for re-enumeration and the possibility of matching questionnaires, and (f) the importance of maintaining the questionnaires from the original enumeration in the proper enumeration area order to facilitate the matching process.

41. To meet all the above requirements is not an easy task. If they are not met, however, the time and money invested in the check may be completely wasted because the goal of accurate evaluation of census results will not be achieved.

(iii) Alternate frame for selection of sample

42. One of the drawbacks of a traditional ad hoc post-enumeration sample field check is that certain proportions of the persons, households and living quarters omitted from the original enumeration are also likely to be missed in the re-enumeration because of their particular characteristics. For example, it has been estimated that the post-enumeration check of the 1950 census of the United States found not much more than half of the original under-enumeration of the population. One of the reasons for this appeared to be that both investigations missed persons without a close attachment to a multi-person household or to a particular set of living quarters.[3/]

43. One way of avoiding this situation is to use an independent set of records as the frame for the sample to be re-enumerated, rather than the census enumeration districts. This procedure was used in the post-enumeration sample field checks of the 1963 population census of Uruguay[4/], of the 1960 census of the United States[5/]; and the 1961 census of Canada[6/].

3/ United States, Department of Commerce, Bureau of the Census, Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960, Series ER. 60, No.1, p. 5.

4/ Uruguay, Ministerio de Hacienda, Direccion General de Estadistica y Censos, un Ensayo de Evaluacion del IV Censo General de Poblacion, 1963.

5/ United States, Department of Commerce, Bureau of the Census, op. cit., No.2.

6/ Canada, Dominion Bureau of Statistics, paper on Final Report on Project III of the Quality Analysis of the 1961 Census.

44. In Uruguay the sample was composed of three sub-samples, derived from a national population register compiled in 1957 and the register of births between 1957 and 1963. The Canadian sample was drawn from the 1956 census records. The United States drew probability samples of persons from four sets of records, that is, (1) persons enumerated in the 1950 census, (2) aliens who were registered with the Immigration and Naturalization Service, (3) children born during the intercensal period, whose and births were registered and (4) persons missed by the 1950 census but detected by the 1950 post-enumeration field check. "It was hoped that this approach would be particularly useful for measuring the coverage of some population groups having a high risk of underenumeration. For example, the sample would identify from the 1950 census records a group of persons who were 8-14 years old in 1950 - a group for which the risk of underenumeration was relatively low-- but who in 1960 would be 19-24, a group for which the risk of underenumeration would be relatively high."[7]

45. In each case the first step after the selection of the sample was to trace each person in order either to locate him or to determine if he had died or emigrated. For each person located, sufficient information had then to be obtained in a re-enumeration to make it possible to find his census return or to make sure that no census return existed for him. The tracing procedure, the re-enumeration and the treatment of persons who died between the census and the post-enumeration check, varied among the three countries. They are described in the publications cited in the Annex.

(iv) Recent post-enumeration field checks in the African region

46. The following table shows, for 12 countries in the African region, the purpose of post-enumeration field checks held in connexion with recent censuses of population and housing, and the availability of results of the checks, according to information received at the Statistical Office of the United Nations.

[7] United States, Department of Commerce, Bureau of the Census, op. cit., No.1, p.5.

Purpose of recent post-enumeration sample field checks, availability of results, and interval between population census enumeration and holding of field checks, in twelve countries in Africa

["x" indicates that the specified purpose was intended. "-" that it was not intended, "..." that information is not available. Letters in Parenthesis indicate availability of results as follows: (a) results available, (b) results not available]

Country	Census date	Subject and purpose of check and availability of results						Interval between completion of population census enumeration and holding of field check
		POPULATION			HOUSING			
		To evaluate completeness	To evaluate content errors	To collect additional data	To evaluate completeness	To evaluate content errors	To collect additional data	
Algeria	4 IV 1966	x(b)	x(b)	x	x(b)	-	-	Immediately
Gabon	8 X 1960-V 1961	x(b)	-	-	-	-	-	...
Ghana	20 III 1960	x(a)	x(a)	x	x(a)	-	x	2-1/2 months
Kenya	15 VIII 1962	x(a)1/	x(a)1/	x1/	-	-	-	2 weeks
Liberia	2 IV 1962	x(a)	-	-	-	-	-	4 months
Malawi	9 VIII 1966	x(b)	...	...	-	-	-	...
Morocco	18 VI 1960	x(b)2/	-	-	-	-	-	...
Sierra Leone	1 IV 1963	x(a)	-	-	-	-	-	about 3 months
Togo	XI 1958 XII 1960	x(b)	-	x	-	-	-	1 year
Tunisia	1 II 1956	x(b)	x(b)	-	-	-	-	2 months
Uganda	18 VIII 1959	x(a)1/	-	x1/	-	-	-	2 weeks
United Republic of Tanzania Tanganyika	19 VIII 1957 1/	x(a)	-	x	-	-	-	2 weeks

1/ Only for African population in rural areas

2/ Only for rural population.

47. It may be noted, in the first place, that post-enumeration checks have been much more commonly used in connexion with enumerations of population than with enumerations of housing. Of the 10 countries for which only a check of population is shown, 5 collected housing information as well as information on population at the time of the census. It is also evident that considerably more attempts have been made to check completeness of coverage than to check content accuracy. Also at least half of the countries combined the post-enumeration field check with a survey designed to collect the additional information on a sample basis to supplement the data gathered in the original enumeration.

48. In regard to the availability of the results, it may be noted that 6 of the countries have completed the evaluation as planned. For the remaining countries, the results are not yet available.

49. Another point of interest about the checks conducted in connexion with population censuses in the region is the time interval between the completion of the original enumeration and the date of the post-enumeration field check. In one country, the check is reported to have taken place immediately after the original enumeration; in 3 countries the interval was reported to be two weeks; for five others, it was respectively, two months, two and a half months, three months, four months and one year. Information is not available for 3 countries.

b. Labour force surveys

50. A labour-force survey held shortly after a general census, or even prior to it, is sometimes used in place of an *ad hoc* post-enumeration field check especially for evaluating quality of response. In Japan, for example, data on type of activity, industry and occupation from the monthly labour-force survey were used to evaluate the accuracy of responses on these topics in the 1950, 1955 and 1960 population censuses.[8/]

8/ Japan, Office of the Prime Minister, Bureau of Statistics, Comparison of the Results of the 1960 Population Census and the Labour Force Survey, Research Paper No. 3, January 1964 (in Japanese)

Israel, the Federal Republic of Germany and the United States[9] have also used information from labour-force surveys for evaluating some of the results of their recent censuses.

51. The information provided by a labour-force survey is generally considered to be of higher quality than that collected in a general census for a number of reasons, including the fact that enumerators for labour-force surveys are usually more intensively trained and more experienced than census enumerators, that more detailed data are collected in the survey and that the enumerators can be more closely supervised. It will be noted that these are the same factors which are believed to make a post-enumeration field check correspondingly more accurate than the general census enumeration.

52. On the assumption that the replies received in a labour-force survey are nearer to the truth than those obtained in a census, the former will provide a good basis of comparison with the census information relating to the characteristics investigated in both inquiries. Furthermore, it is possible to effectuate one-to-one matching of the records from both operations.

53. There are, however, limitations which reduce the effectiveness of the labour-force survey as a tool for evaluation of the census results. In the first place, children below working age are not usually included in a labour-force survey; it cannot, therefore, be used for evaluating coverage except in the labour force ages. In addition, because the labour-force survey is primarily planned and conducted for purposes other than evaluation of the census results, it may be especially difficult completely to match the records from the two investigations. Also, the lack of a close correspondence in time/between the two enumerations leads to the possibility of considerable changes in the composition and characteristics of the population in the interval between the two inquiries. Furthermore

9/ United States, Department of Commerce, Bureau of the Census, Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Accuracy of Data on Population Characteristics as Measured by CPS-Census Match, Series ER 60, No. 5, Washington, D.C., 1964

there may be a lack of uniformity or definitions. Finally, it must be borne in mind that, as with the results of a post-enumeration field-check, the labour-force survey results also cannot be regarded as completely accurate.

(2) Records from other sources

54. Another direct method of evaluating the coverage and content of census results involves comparison of the census data for individuals with those obtained from other sources.

55. The extent to which such comparisons can be made depends on the availability and the completeness of the various independent sources of information. Among these sources, those which have recently been used the most are the continuous population registers, housing registers, registers of births, registers of deaths, military registers, school enrolment lists, electoral rolls, alien registration lists, and tax rolls.

56. With the possible exception of the population registers and some housing registers, all these registers and lists are limited in scope and pertain only to certain sectors of the population. There is thus no point in comparing the census results with these limited sources of data where the purpose is to evaluate the coverage of the census for all sectors of the population throughout the country. Another drawback to the effective use of these registers and lists is that the information they contain on each person is apt to be wanting in accuracy and completeness and is thus not fully satisfactory for evaluating the results of a national census; they may, however, prove useful in making partial comparisons. Some of the registers and lists may, in fact, be particularly useful in evaluating coverage errors because they provide a frame of persons who are considered most likely to be missed in the regular census enumeration (see para. 44)

57. As an illustration of the use of these sources of information for evaluation of census results, the employment of information from population registers and from birth registers is briefly described in paragraphs 58-63.

a. Population registers

58. Because population registers constitute the most comprehensive and reliable of the independent sources of data for evaluating census results, they have been used for this purpose in a number of countries. It must, however, be borne in mind that where discrepancies are encountered, the only way of determining which of the two sources is correct is to trace and interview the persons involved.

59. The use of these registers for the evaluation of coverage is more common than their use for evaluation of content errors, but where, as is not very common, the register information on each person is entirely up-to-date, it can also be used for the latter purpose. Incomplete registers may be used in the areas where they are known to be fairly accurate.

b. Birth registers

60. Although registers of births are in the nature of partial registers and thus, as already mentioned, are not a satisfactory means for evaluating the total coverage of the census, they are used in many countries for evaluating the census coverage in respect of children and conversely the census returns are used for evaluating the coverage of the birth registers.

61. Registers of infants born during the 3-month period preceding the census were used as the universe for testing completion of enumeration of infants at the 1950 census of the United States[10/]. In similar procedures in connexion with the 1964 population census of Guatemala[11/] and the 1961 census of Honduras,[12/] samples were drawn from registers

10/ Shapiro, Sam and Schachter, Joseph, Birth Registration Completeness: United States, 1950, Public Health Reports (U.S. Public Health Service) Vol. 67, No.6, p. 523, June 1952.

11/ Guatemala, Direccion General de Estadistica, Estudio Post-enumerativo Censal, 1965, p. 23.

12/ Honduras, Direccion General de Estadistica y Censos, Estudio de Enumeracion Post-censal de Poblacion y Vivienda de Honduras, 1961, p.4.

of births during one month preceding the census. The two latter countries have reported many difficulties in the attempt at one-to-one matching because of differences between the census information and that of the birth register in respect of name and address.

62. Birth registers have likewise been used for verifying data on individual characteristics, such as age and sex. One way in which this is done is to compare the age of an individual obtained in a census, with the date of birth shown on his birth certificate. Except however in the case of young infants, who are likely to have been enumerated in the area where they were born, matching census records to civil registers of birth is an arduous and costly process, inasmuch as it depends on allocating census returns back to place of birth and of subsequently identifying individuals.

63. In areas where there is a high degree of mobility among the population, this method may, therefore, be very difficult to apply. It may, however, prove useful for application to a small representative sample of the census returns. This was done in connexion with the 1951 census of England and Wales[13/].

(3) Internal checks of duplication

64. In addition to field checks and comparisons with records from other sources another direct method has been used in some countries for analysing the extent of duplication in the census enumeration. The method consists of punching for a sample of the population (e.g. all persons born on a particular date, or all persons whose surnames began with a particular letter), a special set of cards containing only identifying information (name, surname, sex, date of birth marital status and so on). These cards can then be examined for duplications.

13/ United Kingdoms, England and Wales, General Register Office, Census 1951, England and Wales: General Report, London, Her Majesty's Stationary Office, 1958, pp. 36-40.

65. This method is especially appropriate for census counts based on the concept of resident populations, where there is a likelihood of double enumeration of such individuals as persons with more than one residence, students away at school, hospital patients, travellers or persons in military service. It suffers, however, from the possibility that some persons may be overlooked in the punching and that some duplications may be hidden by errors in punching or by discrepancies in the identifying information about the same person collected from two different respondents.

C. Indirect methods of evaluation[14/]

66. The indirect method of detecting unreliability in the enumeration consists of the critical analysis of the internal consistency of census results and the way in which these results conform to expected values obtained from other sources. The methods employed are essentially those of demographic analysis, utilizing census data and related statistics, especially those of natural and physical movement of population, that is, statistics of birth, death, immigration and emigration for the country as a whole for its major civil divisions and, if appropriate, for major national and/or ethnic groups.

67 The essence of these methods is the determination of the results which would be expected on the basis of other information and the comparison of these with the actual census results. It is predicated upon the principle that, in the absence of exceptional circumstances (war, significant migration, or natural catastrophes) changes in the size of the population, its geographic distribution and its other characteristics occur at a fairly regular pace. Unusual deviations from this pace must be carefully examined, and, if they cannot be satisfactorily explained, it must be assumed that

14/ Most of the indirect methods of assessment of accuracy have been described in detail in Manuals on Methods of Estimating Population: Manual I, Methods of Estimating Total Population for Current Dates; Manual II, Methods of Appraisal of Quality of Basic Data for Population Estimates; Manual III, Methods for Population Projections by Sex and Age (United Nations publications, Sales Nos.: 52.XIII.5, 56.XII, 56XII.3). See also "Notes on Availability of National Population Census Data and Methods of Evaluating their Reliability", Demographic Year book 1962 (United Nations publication, Sales No.: 63.XIII.1), Chapter I, pp. 1-11.

there are deficiencies in the current or older census data or in the other sources of information.

68. These methods cannot, in themselves, produce an assessment of reliability; all they can do is produce a clue to possible unreliability which must be followed up by additional analyses designed to prove or disprove the possibility.

(1) Total population

69. The simplest numerical relationship which can be investigated is that of the latest cansus result - that is, the total count for the entire country or geographic sub-divisions - with that from a previous census or a series of previous censuses, for the calculation of annual inter-censal rates of growth or decrease. Abnormally high or low rates are an indication of the need for further investigation.

70. As a next step, the total may be compared with the regular population estimate for the census date if the estimate has been constructed according to accepted methods and the components are accurate.

71. A more refined approach is that of the "balancing equation", which is based on the fact that the population of an area during a given time interval can grow only by the number of births and immigrants and decrease only by the number of deaths and emigrants. Therefore, if these elements are known with some degree of reliability, their use in a balancing equation should produce a reliable estimate of the population, which, in turn, should be in close agreement with the results of a census taken at the same time.

72. The equation to be used takes the following form:

$$P_1 = P_0 + B + I - D - E$$

where P_1 is the result of the most recent census

P_0 is the result of the previous census

B is the number of briths during the intercensal period

I is the number of immigrants during the intercensal period

D is the number of deaths during the intercensal period

E is the number of emigrants during the intercensal period.

73. Problems arising from a lack of or deficiency in birth statistics and unreliable statistics on deaths of infants and very young children may be overcome by using a slightly different form of the balancing equation which eliminates persons under 5 years of age at the time of the first census. This simplified form is based on the fact that the population above a certain age enumerated at a recent census must already have been alive at the time when the previous census was taken. These persons would have aged by the number of years elapsed between the censuses, been augmented by persons of this age group immigrating, and been diminished by deaths and emigration which occurred in the interval. Thus, by utilizing only the population 5 years of age and over at the first census together with deaths and migration in the inter-censal period one may test the reliability of the recent enumeration of population 5 + x years of age and over, x being the number of inter-censal years.

74. A further simplification of the equation can be made if it can be determined that migratory movement in the inter-censal period has been negligible. In this case, the migration elements drop out and the equation becomes simply:

$$P_1 = P_o - D$$

where P_1 is the population 5 + x years of age and over at the most recent census

P_o is the population 5 years of age and over at the previous census

D is the number of deaths of persons 5 years of age and over during the intercensal period.

(2) Selected attribute distributions

75. Other analytical studies of certain selected attribute distributions of the population, such as the distribution of the population by size of household, by sex, by age and sex, and by age, sex and marital status, and of the availability of specific housing facilities can also indicate possible errors in coverage.

76. In regard to the size of household, for example, results which indicate a larger average size of household in urban areas than in rural areas may indicate an error of enumeration because the average size of household is usually larger in rural areas than in urban areas.

77. The sex distribution of the population should show a more or less equal number of males and females. The ratio of males per 100 females should not contain inexplicable deviations among major sub-divisions of the country. The sex/age distribution should produce a configuration in which the transitions from one group to the next are orderly, with no unexplained abrupt fluctuations from one age group to the next. Sex ratios should show fairly regular decreases with advancing age because male mortality rates exceed female rates. Balancing equations (see para.71) can test the adequacy of the enumeration of each age/sex cohort above the youngest.

78. Examination of the age/sex/marital status distribution of the population can reveal such abnormalities as a lower average age among married men than among married women or a preponderance of married men over married women.

79. Housing census results which show a greater frequency of piped water or of electricity, for example, in rural areas than in urban areas would obviously be immediately suspect and call for further investigation.

V. CONCLUSION

80. It is clear that the evaluation of the reliability of census results is a prime requisite for their proper use and also for the improvement of future censuses. The scope of an evaluation programme will be determined by the resources, the statistical development and other characteristics of the country the purpose of the evaluation and the possibilities offered by each method.

81. Among the different methods which have been used., the ad hoc post-enumeration sample field check is generally accepted as the most satisfactory. Countries which lack the resources to carry out a field check, to evaluate coverage as well as the quality of the major census results, can benefit from a simple check of coverage or from a coverage and content check of special segments of the population in which errors are more likely to occur. Even partial results are of great value in indicating different types of errors, their possible causes and the characteristics of the living quarters, households and persons involved.

82. In addition to the post-enumeration field check, countries should use any other method of evaluation which is feasible.

ANNEX

National Reports on the Methodology of Ad-hoc
Post Enumeration Sample Field Check

A. AFRICA

Ghana, Census Office, 1960 Population Census of Ghana, Vol. V, Administrative Report Post Enumeration Survey, Part. II, and Census Evaluation Part. III. pp 313-401, Accra, 1964.

Kenya, Statistics Division, Ministry of Economic Planning and Development, Population Census, 1962, African Population Vol. III, October 1966.

Liberia, Office of National Planning, Bureau of Statistics, 1962 Population Census of Liberia, Monrovia, 1965.

Sierra Leone, Central Statistics Office, 1963 Population Census of Sierra Leone, Vol. 2, Social Characteristics, Freetown, 1965.

Tanganyika, African Census Report, 1957.

Uganda, Statistics Branch, Ministry of Economic Development, Uganda Census 1959, African Population, November, 1961.

B. AMERICA, NORTH

Canada, Dominion Bureau of Statistics, paper on Final Report on Project III of the Quality Analysis of the 1961 Census.

Costa Rica, Ministerio de Economia y Hacienda, Direccion General de Estedistica y Censos, Evaluacion Censos Población y Vivienda y Agropecuario, 1963.

Guatemala, Direccion General de Estedistica, Estudio Post-Enumerative Censal, 1965.

Honduras, Direccion General de Estadistica y Censos, Estudio de Enumeracion Post-Censal de Población y Vivienda de Honduras, 1961.

Panama, Contraloria General de la Republica, Direccion General de Estadistica y Censos, Informe General sobre el levantamiento de los Censos Macionales 1960-1962, Parte III, Cap. IV, péginas 87-103.

United States, Department of Commerce, Bureau of the Census, Evaluation and Research Programme of the US Censuses of Population and Housing, 1960, Series ER 60, Nos. 1-4, Washington D.C., 1963 and 1964.

C. AMERICA, SOUTH

Chile, Kappas Barrientos, Heccor and Marks, Fli S., "Evaluation of the Accuracy of the 1960 Censuses of Chile" Estadistica, Journal of the Inter-American Statistical Institute. Vol. XIX, No. 72, September 1961, pp 492-505.

Uruguay, Ministerio de Hacienda, Dirección General de Estadistica y Censos Un Ensayo de Evaluación del IV Censo General de Población, 1963.

Venezuela, Ministerio de Fomento, Dirección General de Estadística y Censos Nacionales, Encuesta por muestreo para medir el grado de omision censal en el Área Metropolitana de Caracas, referente al IX Censo de Población de 1961 Muestreo en Venezuela, Caracas 1962, páginas 49-78.

D. ASIA

Hong Kong, Report on the 1961 Census, Vol. II Chapter XVIII.

Republic of Korea, Bureau of Statistics, Economic Planning Board, An Evaluation Study for the Accuracy of the 1960 Population and Housing Census of Korea, Seoul, 1966.

E. EUROPE

France, Institut National de la Statistique et des Etudes Economiques, Recensement Demographique de 1962, Etude de l'exaustivite du Denombrement, Paris, 1963.