

Appendix C.—ACCURACY OF THE DATA

SOURCES OF ERROR	App-52
EDITING OF UNACCEPTABLE DATA	App-52
ALLOCATION TABLES	App-53
SAMPLE DESIGN	App-54
RATIO ESTIMATION	App-55
SAMPLING VARIABILITY	App-56

SOURCES OF ERROR

Human and mechanical errors occur in any mass statistical operation such as a decennial census. Errors during the data collection phase can include failure to obtain required information from respondents, obtaining incorrect or inconsistent information, and recording information in the wrong place or incorrectly. Errors can also occur during the field review of the enumerator's work, the clerical handling of the questionnaires, and the various stages of the electronic processing of the material. Careful efforts are made in every census to keep the errors in each step at an acceptably low level. Quality control and check measures are utilized throughout the census operation.

In the manual editing and coding operation, a sample of the work of each coder was verified, and a coder who showed consistently high error rates was retrained. A coder who still did not produce work of acceptable quality after retraining was removed from the coding operation. In addition, provision was made for correction of any work units for which the error rate exceeded an acceptable level. Information on error rates will be given in later publications.

As was done for the 1950 and 1960 censuses, evaluative material on many aspects of the 1970 census will be

published as soon as the appropriate data are accumulated and analyzed. A major concern in the evaluation work is to ascertain, insofar as possible, the degree of completeness of the count of both population and housing units.

EDITING OF UNACCEPTABLE DATA

The objective of the processing operation is to produce a set of statistics that describes the population as accurately and clearly as possible. To meet this objective, certain unacceptable entries were edited.

In the field, questionnaires were reviewed for omissions and certain inconsistencies by a census clerk or enumerator, and a followup was made for missing information. The major review occurred in the central processing office, where the editing and coding operation provided an opportunity to correct obvious errors in the respondents' entries for those items which required manual processing. In coding relationship to household head, for example, the clerk made use of written entries, which the computer is not able to read, in determining the correct entry where the machine readable item was blank or contained conflicting information. For a few of the items, the respondents' entries were reviewed for reasonableness or consistency on the basis of other information on the questionnaire. As a rule, however, editing was performed by hand only when it could not be done effectively by machine.

Because of limitations of computer capacity and other resources, a number of complicated editing steps were not introduced when the effect upon the final data was considered to be small. Thus, there may be a small

number of cases having unlikely combinations of characteristics.

As one of the first steps in mechanical editing, the configuration of marks on the questionnaire was scanned electronically to determine whether it contained information for a person or merely spurious marks. If the questionnaire contained entries for at least two of the basic characteristics (relationship, sex, race, age, marital status), or for at least two relevant sample characteristics, the inference was made that the marks represented a person. Names were not used as a criterion of the presence of a person because the electronic scanning was unable to distinguish between a name and any other entry in the name space.

If any characteristics for a person were missing, they were, in most cases, supplied by allocation. Allocations, or assignments of acceptable codes in place of unacceptable entries, were needed most often where an entry for a given item was lacking or where the information reported for a person on that item was inconsistent with other information for the person. As in earlier censuses, the general procedure for changing unacceptable entries was to assign an entry that was consistent with entries for other persons with similar characteristics. Thus, a person who was reported as a 20-year-old son of the household head, but for whom marital status was not reported, was assigned the same marital status as that of the last son processed in the same age group. The assignment of acceptable codes in place of blanks or unacceptable entries, it is believed, enhances the usefulness of the data. The allocation technique may be illustrated by the procedure used in the assignment of wage or salary income. The allocation of this item was carried out in the following steps:

1. The computer stored reported wage or salary income entries, by sex, age, relationship to household head, race, major occupation group, class of worker, and number of weeks worked in 1969, for persons 14 years old and over who worked in 1969.

2. Each stored wage or salary entry was retained in the computer only until a succeeding person having the same characteristics and having wage or salary income reported was processed through the computer during the mechanical edit operation. Then the reported wage or salary income entry of the succeeding person was stored in place of the one previously stored.

3. When the wage or salary income of a person 14 years old or over who worked in 1969 was not reported or the entry was unacceptable, the wage or salary income assigned to this person was that stored for the last person who otherwise had the same characteristics.

This process insured that the distribution of wage or salary income assigned by the computer for persons of a given set of characteristics would correspond closely to the wage or salary income distribution of persons who had reported that item in the current census.

The editing process also includes another type of correction; namely, the assignment of a full set of characteristics for a person. When there was indication that a housing unit was occupied but the questionnaire contained no information for any person, a previously processed household was selected as a substitute and the full set of characteristics for each substitute person was duplicated.

These duplications fall into two classes: (1) "persons substituted due to noninterview," e.g., a housing unit indicated as occupied but the occupants were not listed on the questionnaire and (2) "persons substituted due to mechanical failure," e.g., where the questionnaire page on which persons are listed was not properly microfilmed.

Specific tolerances were established for the number of computer allocations and substitutions that would be permitted. If the number of corrections was beyond tolerance, the questionnaires in which the errors occurred were clerically reviewed. If it was found that the errors resulted from damaged questionnaires, from improper microfilming, from faulty reading by FOSDIC of undamaged questionnaires, or from other types of machine failure, the questionnaires were reprocessed.

ALLOCATION TABLES

The extent of the various editing and allocation procedures and their effect on each of the subjects is shown in tables B-1 to B-5 (which follow table 39) and tables C-1 to C-4 (which follow table 137). Information on the number of substitutions for each State, with separate counts of those for enumeration reasons and those for processing reasons, is presented in the United States Summary report, PC(1)-B1. Tables B-1 to B-5 relate to 100-percent tabulations and tables C-1 to C-4 relate to sample tabulations.

There is a difference in the method of counting allocations between tables B-1 and B-5 on the one hand, and tables B-3 and B-4 on the other hand. In tables B-1 and B-5, a person with one or more allocations whose record is duplicated for substitution purposes

is counted twice (i.e., both "originally" and as a "substitute"). In tables B-3 and B-4, such a person is counted only once.

The sum of the percentages of persons having assignments in each population characteristic is greater than the number of persons with one or more allocations because some persons had allocations for more than one characteristic. Not tallied, and therefore not included in these tables, are the allocations for missing information on quarter of year of birth; these allocations were made on a random basis.

The number of persons and housing units enumerated in the sample is shown in table C-1. The extent of allocations or substitutions for non-response or inconsistency is shown for selected characteristics in tables C-2 and C-4; and distributions as they appeared before allocations or substitutions are shown in table C-3.

In table C-1, the "unweighted sample count" of persons and housing units represents a count of one for each sample person and each sample housing unit readable by the computer; thus, it is the unweighted universe before ratio estimation.

In table C-2 "persons with two or more sample characteristics reported" are persons with acceptable entries in two or more relevant sample characteristics (for example, entries in weeks worked in 1969 and wage or salary income for persons 14 years old and over). The column "persons with sample information" in table C-4 refers to the same information for specific places and counties. The characteristics listed in tables C-2 and C-4 cover most of the items that are published in this report. For each subject in tables C-2 and C-4 the universe applicable to the charac-

teristic is described, along with the percent of persons in the relevant universe for whom nonresponses were allocated. In table C-2, the number of persons in each universe is also shown.

The figures in table C-3, representing distributions for each subject as they appeared before substitution and allocation for nonresponse, may be compared with the corresponding statistics in the appropriate detailed table, category by category, to measure the net effect of substitution and allocation.

The allocation rates shown in these tables are generally, but not always, comparable with the rates shown in similar tables in the 1960 reports. Certain types of response allocation are included as allocations in 1970 but were not so included in 1960. For example, allocations of race are counted for household members when allocating from the race of the household head, and allocations of marital status and sex are counted for persons identified as wives of household heads or as heads of households with wife present. Moreover, the characteristics of sample persons in substituted households are counted as allocations in 1970 but were not so counted in 1960 (except when the person in the substituted household had a nonresponse on a given characteristic).

In tables C-2 and C-4 the allocated characteristics resulting from substituted households include only those substitutions made when a housing unit enumerated as occupied contained no information for any person. Other substitutions made in the processing of the 100-percent data are mainly accounted for in the sample processing by adjusting the weights of the sample persons.

The nonresponse rates shown in tables C-2 to C-4 do not necessarily

reflect omissions on the questionnaire. Processing difficulties, particularly as related to the income items, resulted in the loss of some answers and the subsequent treatment of those entries as blanks.

SAMPLE DESIGN

For persons living in housing units at the time of the 1970 census, the housing unit, including all its occupants, was the sampling unit; for persons in group quarters identified in advance of the census, the sampling unit was the person. In nonmail areas, the enumerator canvassed his assigned area and listed all housing units in an address register sequentially in the order in which he first visited the units whether or not he completed the interview. Every fifth line of the address register was designated as a sample line, and the housing units listed on these lines were included in the sample. Each enumerator was given a random line on which he was to start listing and the order of canvassing was indicated in advance, although the instructions allowed some latitude in the order of visiting addresses. In mail areas, the list of housing units was prepared prior to Census Day either by employing commercial mailing lists corrected through the cooperation of the post office or by listing the units in a process similar to that used in nonmail areas. As in other areas, every fifth housing unit of these lists was designated to be in the sample. In group quarters, all persons were listed and every fifth person was selected for the sample.

This 20-percent sample was subdivided into a 15-percent and a 5-percent sample by designating every fourth 20-percent sample unit as a member of the 5-percent sample. The remaining sample units became the

15-percent sample. Two types of sample questionnaires were used, one for the 5-percent and one for the 15-percent sample units. Some questions were included on both the 5-percent and 15-percent sample forms and therefore appear for a sample of 20 percent of the units in the census. Other items appeared on either the 15-percent or the 5-percent questionnaires. The sample rates for various subjects collected in the census are given in Appendix B.

Although the sampling procedure did not automatically insure an exact 20-percent sample of persons or housing units in each locality, the sample design was unbiased if carried through according to instructions; generally for larger areas the deviation from 20 percent was found to be quite small. Biases may have arisen, however, when the enumerator failed to follow his listing and sampling instructions exactly. Quality control procedures were used throughout the census process, however, and where there was clear evidence that the sampling procedures were not properly followed, some enumerators' assignments were returned to the field for resampling. The percent of the total population and housing units enumerated in the sample is shown in table C-1, chapter C. The computation of these proportions excluded several classes of the population for which no attempt at sampling was made. These were the relatively small numbers of persons and housing units (in most States, less than one percent) added to the enumeration from the post-census post office check, the various supplemental forms, and the special check of vacant units. However, the ratio estimation procedure described below adjusts the sample data to reflect these classes of population and housing units.

APPENDIX C—Continued

<p>RATIO ESTIMATION</p> <p>The statistics based on 1970 census sample data are estimates made through the use of ratio estimation procedures, which were applied separately for population and for housing data for each of the 5-, 15-, and 20-percent samples. The first step in carrying through the ratio estimates</p>	3	3-person household	<p>operation was performed for each of the 19 groups in the first stage, then for the two groups in the second stage and finally for the 24 groups in the third stage. As a rule, the weighted sample counts within each of the 24 groups in the third stage should agree with the complete counts for the weighting areas. Close, although not exact consistency can be expected for</p>
	.	.	
	6	6-or-more-person household	
		<i>Male head without own children under 18</i>	
	7-12	1-person to 6-or-more-person households	

The estimates realize some of the gains in sampling efficiency that would have resulted had the population been stratified into the groups before sampling. The net effect is a reduction in both the sampling error and possible bias of most statistics below what would be obtained by weighting the results of the sample by a uniform factor (e.g., by weighting the 20-percent sample results by a uniform factor of 5). The reduction in sampling error will be trivial for some items and substantial for others. A byproduct of this estimation procedure is that estimates for this sample are, in general, consistent with the complete count for the population groups used in the estimation procedure. A more complete discussion of the technical aspects of these ratio estimates will be presented in a separate report.

SAMPLING VARIABILITY

The estimates from the 20-, 15-, and 5-percent sample tabulations are subject to sampling variability. The standard errors of these estimates can be approximated by using the data in tables D through F. The chances are about 2 out of 3 that the difference (due to sampling variability) between the sample estimate and the figure that would have been obtained from a complete count of the population is less than the standard error. The chances are about 19 out of 20 that the difference is less than twice the standard error and about 99 out of 100 that it is less than 2½ times the standard error. The amount by which the estimated standard error must be multiplied to obtain other odds deemed more appropriate can be found in most statistical textbooks. The sampling errors may be obtained by using the factors shown in table F

in conjunction with table D for absolute numbers and in conjunction with table E for percentages. These tables reflect the effect of simple response variance, but not of bias arising in the collection, processing, and estimation steps nor of the correlated errors enumerators introduce; estimates of the magnitude of some of these factors in the total error are being evaluated and will be published at a later date.

Table D shows approximate standard errors of estimated numbers for most statistics based on the 20-percent sample. In determining the figures for this table, some aspects of the sample design, the estimation process, and the population of the area over which the data have been compiled are ignored. Table E shows standard errors of most percentages based on the 20-percent sample. Linear interpolation in tables D and E will provide approximate results that are satisfactory for most purposes.

Table F provides a factor by which the standard errors shown in tables D and E should be multiplied to adjust for the combined effect of the sample design, the estimation procedure, and the sample size (i.e., whether 20-percent, 15-percent, or 5-percent). Some items collected from the 20-percent sample are tabulated from the 15-percent or 5-percent samples in specific tables. Table B shows the tabulation rate for data published in tables 40-137 (chapter C). Table C shows the tabulation rate for data published in tables 138-216 (chapter D).

To estimate the standard error for a given characteristic based on the 15- or 5-percent sample, or for a more precise estimate for the 20-percent sample, locate in table F the factor applying to the characteristic and sample size used to tabulate the data and multiply this

factor by the standard error found in table D or E. If the estimate is not identified in table F, use the factor shown for "all other." Where data are shown as cross-classifications of two characteristics, locate and use the larger factor. Similarly, if an item, although collected on one sample basis, has been tabulated for a smaller sample, use the factor appropriate for the smaller sample.

The standard errors estimated from these tables are not directly applicable to differences between two sample estimates. In order to estimate the standard error of a difference, the tables are to be used somewhat differently in the three following situations:

1. For a difference between the sample figure and one based on a complete count (e.g., arising from comparisons between 1970 sample statistics and complete-count statistics for 1960 or 1950), the standard error is identical with the standard error of the 1970 estimate alone.
2. For a difference between two sample figures (that is, one from 1970 and the other from 1960, or both from the same census year), the standard error is approximately the square root of the sum of the squares of the standard errors of each estimate considered separately. This formula will represent the actual standard error quite accurately for the difference between estimates of the same characteristics in two different areas, or for the difference between separate and uncorrelated characteristics in the same area. If, however, there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error. The approximate standard

error for the 1970 sample figure is derived directly from tables D through F. The standard error of a 25-percent 1960 sample figure may be obtained from the relevant 1960 census report or an approximate value may be obtained by multiplying the appropriate value in table D or E by 0.9.

3. For a difference between two sample estimates, one of which represents a subclass of the other, the tables can be used directly with the difference considered as the sample estimate.

The sampling variability of the medians presented in certain tables (median age, median years of school completed, and median income) depends on the size of the base and on the distribution on which the median is based. An approximate method for measuring the reliability of an estimated median is to determine an interval about the estimated median such that there is a stated degree of confidence the true median lies within the interval. As the first step in estimating the upper and lower limits of the interval (that is, the confidence limits) about the median, compute

one-half the number on which the median is based (designated $\frac{N}{2}$). From table D, following the method outlined in other parts of this section, compute the standard error of an estimated number equal to $\frac{N}{2}$. Subtract this standard error from $\frac{N}{2}$. Cumulate the frequencies (in the table on which the median is based) until the total first exceeds the difference between $\frac{N}{2}$ and its standard error and by linear interpolation obtain a value corresponding to this number. In a corresponding manner, add the standard error to $\frac{N}{2}$, cumulate the frequencies in the table, and obtain a value in the table on which the median is based corresponding to the sum of $\frac{N}{2}$ and its standard error. The chances are about 2 out of 3 that the median would lie between these two values. The range for 19 chances out of 20 and for 99 in 100 can be computed in a similar manner by multiplying the standard error by the appropriate factors before subtracting from and adding to one-half the number reporting the characteristic. Interpolation to obtain the

values corresponding to these numbers gives the confidence limits for the median.

The sampling variability of a mean, such as the number of children ever born per 1,000 women or mean income, presented in certain tables, depends on the variability of the distribution on which the mean is based, the size of the sample, the sample design (for example, the use of households as the sampling unit), and the use of ratio estimates.

An approximation to the variability of the mean may be obtained as follows: compute the standard deviation of the distribution on which the mean is based; divide this figure by the square root of one-fifth of the total units in the distribution; multiply this quotient by the factor from table F appropriate to the statistic and the actual sample rate on which the mean is based. If the distribution is not published in the detailed tables, calculate the standard deviation from a comparable distribution for a larger area or for a similar population group; divide by the square root of one-fifth of the units on which the mean of interest is based; multiply the quotient by the factor from table F.

TABLE B. Tabulation Rates for Subjects in Tables 40 to 137 (Chapter C)
(Cross-classifications of two or more items are tabulated on the smaller rate)

Subject	Tabulation rate (percent)	Subject	Tabulation rate (percent)	Subject	Tabulation rate (percent)
Sex	20	Mother tongue	15	Activity 5 years ago	20
Race	20	Spanish surname	15	Place of work	15
Age	20	Year moved into present house	15	Means of transportation to work	15
Household relationship	20	Residence in 1965	15	Occupation	20
Family composition	20	School enrollment	¹ 15	Industry	20
Families and subfamilies	20	Years of school completed	20	Class of worker	20
Type of group quarters	20	Vocational training	5	Income	20
Marital status	20	Veteran status	15	Poverty status	20
Marital history	5	Disability	5	Tenure of housing unit	20
Children ever born	20	Employment status	20	Farm residence	20
State of birth	20	Labor force participation	20	Value of housing unit	20
Country of origin	15	Weeks worked in 1969	20	Gross rent	20
Spanish origin or descent	5			Plumbing facilities	20
Nativity and parentage	15				

¹The data under the heading "Males 16 to 21 years old not attending school" and the data on school enrollment for persons not in the labor force are based on question 22, which has a 20-percent sample tabulation rate.

TABLE C. Tabulation Rates for Subjects in Tables 138 to 216 (Chapter D)

Table number	Tabulation rate (percent)	Table number	Tabulation rate (percent)	Table number	Tabulation rate (percent)
138-142	15	159-160	5	179-181	¹ 20
143, 144	5	161	¹ 20	182	15
145, 146	15	162	(²)	183-189	¹ 20
147, 148	¹ 20	163-168	¹ 20	190, 191	15
149, 150	5	169	5	192-214	¹ 20
151	15	170-177	¹ 20	215	15
152	5	178	5	216	¹ 20
153-158	¹ 20				

¹Data for the Spanish heritage population are tabulated from the 15-percent sample.

²In this table, data on marital history are based on the 5-percent sample; data on State of birth and ethnic group are based on the 15-percent sample; and all other data are based on the 15-percent sample when shown for persons of Spanish heritage and on the 20-percent sample when shown for other subgroups and the total.

TABLE D. Approximate Standard Error of Estimated Number Based on 20-Percent Sample

(Range of 2 chances out of 3; for factors to be applied see table F and text)

Estimated number ¹	Number of persons, families, or households in area ²								
	1,000	10,000	25,000	100,000	250,000	1,000,000	3,000,000	5,000,000	20,000,000
50.....	15	15	15	15	15	15	15	15	15
100.....	20	20	20	20	20	20	20	20	20
250.....	30	30	30	30	30	30	30	30	30
500.....	30	45	45	45	45	45	45	45	45
1,000.....	...	60	60	65	65	65	65	65	65
2,500.....	...	90	95	100	100	100	100	100	100
5,000.....	...	100	130	140	140	140	140	140	140
10,000.....	150	190	200	200	200	200	200
15,000.....	150	230	240	240	240	240	240
25,000.....	270	300	310	310	320	320
50,000.....	320	400	440	440	440	450
75,000.....	270	450	520	540	540	540
100,000.....	490	600	620	630	630

¹For estimated numbers larger than 100,000, the relative errors are somewhat smaller than for 100,000.

²This relates to the smallest complete geographic area to which the estimate under consideration pertains. Thus, the area may be the State, city, county, standard metropolitan statistical area, urbanized area, or the urban or rural portion of the State or county. The rural farm or rural non-farm persons in the State or county, the Negro families, etc., do not represent complete areas.

TABLE E. Approximate Standard Error of Estimated Percentage Based on 20-Percent Sample

(Range of 2 chances out of 3; for factors to be applied see table F and text)

Estimated percentage	Base of percentage						
	500	1,000	2,500	10,000	25,000	100,000	250,000
2 or 98.....	1.3	0.9	0.6	0.3	0.2	0.1	0.1
5 or 95.....	2.0	1.4	0.9	0.4	0.3	0.1	0.1
10 or 90.....	2.7	1.9	1.2	0.6	0.4	0.2	0.1
25 or 75.....	3.9	2.7	1.7	0.9	0.5	0.3	0.2
50.....	4.5	3.2	2.0	1.0	0.6	0.3	0.2

TABLE F. Factor to Be Applied to Standard Error
 (For cross-classifications of two or more subjects, locate the factor for each subject at the appropriate tabulation rate and use the largest)

Subject	Factor if tabulation rate is—			Subject	Factor if tabulation rate is—		
	20 percent	15 percent	5 percent		20 percent	15 percent	5 percent
Race	1.7	2.1	3.8	Labor force status or participation	0.7	0.9	1.5
Age	0.8	1.0	1.8	Unemployed	1.1	1.3	2.3
Household relationship	0.5	0.6	1.1	Hours worked	0.7	0.9	...
Families and subfamilies ¹	0.6	0.7	...	Weeks worked in 1969	0.7	0.9	...
Unrelated individuals	1.3	1.5	...	Year last worked	0.7	0.9	...
Type of group quarters	0.6	0.7	...	Activity 5 years ago	0.9	1.6
Marital status	0.6	0.7	1.3	Place of work	1.3	...
Marital history	2.0	Means of transportation to work	1.3	...
Own children under 5 years old ...	0.8	1.0	1.8	Occupation	1.0	1.2	2.2
State of birth	1.3	1.6	2.8	Industry	1.1	1.3	...
Country of origin	1.6	...	Class of worker	1.1	1.3	...
Country of birth	1.6	2.8	Income in 1969			
Spanish origin or descent ²	2.8	Persons	1.0	1.2	2.2
Spanish heritage ²	1.6	...	Families ¹	1.1	1.3	2.3
Nativity and parentage	1.7	3.1	Poverty status in 1969			
Mother tongue	1.8	...	Persons	1.8	2.1	3.9
Year moved into present house	1.9	...	Families ¹	1.1	1.3	2.3
Residence in 1965	2.0	3.7	Housing characteristics			
Rural farm-nonfarm residence ...	1.9	2.2	4.1	Tenure ¹	0.2	0.3	...
School enrollment	0.8	0.9	...	Other ¹	1.0	1.1	...
Years of school completed	1.0	1.2	2.2	All other	1.0	1.2	2.2
Vocational training	1.6				
Veteran status	0.9	...				
Disability	2.5				

¹When determining the standard error of a number relating to families or households, use the number of families or households in the area for selecting the appropriate column in table D.

²Tabulations of characteristics for persons of Spanish heritage (Puerto Rican birth or parentage, Spanish language or Spanish surname) are based on the 15-percent sample, and the appropriate factor is found in the 15-percent column for the subject or for Spanish heritage, whichever is the larger. Tabulations of characteristics for persons of Spanish origin or descent are based on the 5-percent sample and the appropriate factor is found in the 5-percent column for the subject or for Spanish origin or descent, whichever is the larger.

ILLUSTRATION. Assume table 150 shows that a total of 100,000 males in an area of 2,000,000 have 4 years of high school and some vocational training. This represents a cross-classification of years of school completed and vocational training, and the tabulation rate is 5 percent, according to table C. The largest of the factors for those subjects at the 5-percent tabulation rate is found in table F to be 2.2. Interpolation in table D shows that the approximate standard error of an estimate of 100,000 in an area of 2,000,000 when based on a 20-percent sample is about 610. The product of 2.2 times 610 is 1342 which means the chances are about 2 out of 3 the results of a complete census will not differ by more than 1342 from the estimated 100,000 when based on the 5-percent sample. It also follows there are only 5 chances in 100 that a complete census would differ by as much as 2684; that is, by about 2 times the number estimated from tables D and F.

Appendix D.—PUBLICATION AND COMPUTER SUMMARY TAPE PROGRAM

The results of the 1970 Census of Population and Housing are being issued in the form of printed reports, microfiche copies of the printed reports, computer summary tapes, computer printouts, and microfilm. Listed below are short descriptions of the final report series and computer tapes, as currently planned. More detailed information on this program can be obtained by writing to the Publications Distribution Section, Bureau of the Census, Washington, D.C. 20233.

Population Census Reports

Volume I. CHARACTERISTICS OF THE POPULATION

This volume consists of 58 "parts"—number 1 for the United States, numbers 2 through 52 for the 50 States and the District of Columbia in alphabetical order, and numbers 53 through 58 for Puerto Rico, Guam, Virgin Islands, American Samoa, Canal Zone, and Trust Territory of the Pacific Islands, respectively. Each part, which is a separate clothbound book, contains four chapters designated as A, B, C, and D. Each chapter (for each of the 58 areas) is issued as an individual paperbound report in four series designated as PC(1)-A, B, C, and D, respectively. The 58 PC(1)-A reports have been specially assembled and issued in a clothbound book, designated as Part A.

■ Series PC(1)-A.

NUMBER OF INHABITANTS

Final official population counts are presented for States, counties by urban and rural residence, standard metropolitan statistical areas (SMSA's), urbanized areas, county subdivisions, all incorporated places, and unincorporated places of 1,000 inhabitants or more.

■ Series PC(1)-B.

GENERAL POPULATION CHARACTERISTICS

Statistics on age, sex, race, marital status, and relationship to head of household are presented for States, counties by urban and rural residence, SMSA's, urbanized areas, county subdivisions, and places of 1,000 inhabitants or more.

■ Series PC(1)-C.

GENERAL SOCIAL AND ECONOMIC CHARACTERISTICS

Statistics are presented on nativity and parentage, State or country of birth, Spanish origin, mother tongue, residence 5 years ago, year moved into present house, school enrollment (public or private), years of school completed, vocational training, number of children ever born, family composition, disability, veteran status, employment status, place of work, means of transportation to work, occupation group, industry group, class of worker, and income (by type) in 1969 of families and individuals. Each subject is shown for some or all of the following areas: States, counties (by urban, rural-nonfarm, and rural-farm residence), SMSA's, urbanized areas, and places of 2,500 inhabitants or more.

■ Series PC(1)-D.

DETAILED CHARACTERISTICS

These reports cover most of the subjects shown in Series PC(1)-C, above, presenting the data in considerable detail and cross-classified by age, race, and other characteristics. Each subject is shown for some or all of the following areas: States (by urban, rural-nonfarm, and rural-farm residence), SMSA's, and large cities.

Volume II.

SUBJECT REPORTS

Each report in this volume, also designated as Series PC(2), concentrates on a particular subject. Detailed information and cross-relationships are generally provided on a national and regional level; in some reports, data for States or SMSA's are also shown. Among the characteristics covered are national origin and race, fertility, families, marital status, migration, education, unemployment, occupation, industry, and income.

Housing Census Reports

Volume I.

HOUSING CHARACTERISTICS FOR STATES, CITIES, AND COUNTIES

This volume consists of 58 "parts"—number 1 for the United States, numbers 2 through 52 for the 50 States and the District of Columbia in alphabetical order, and numbers 53 through 58 for Puerto Rico, Guam, Virgin Islands, American Samoa, Canal Zone, and Trust Territory of the Pacific Islands, respectively. Each part, which is a separate clothbound book, contains two chapters designated as A and B. Each chapter (for each of the 58 areas) is issued as an individual paperbound report in two series designated as HC(1)-A and B, respectively.

■ Series HC(1)-A.

GENERAL HOUSING CHARACTERISTICS

Statistics on tenure, kitchen facilities, plumbing facilities, number of rooms, persons per room, units in structure, mobile home, telephone, value, contract rent, and vacancy status are presented for some or all of the following areas: States (by urban and rural residence), SMSA's, urbanized areas, places of 1,000 inhabitants or more, and counties.

■ Series HC(1)-B.

DETAILED HOUSING CHARACTERISTICS

Statistics are presented on a more detailed basis for the subjects included in the Series

HC(1)-A reports, as well as on such additional subjects as year moved into unit, year structure built, basement, heating equipment, fuels, air conditioning, water and sewage, appliances, gross rent, and ownership of second home. Each subject is shown for some or all of the following areas: States (by urban, rural-nonfarm, and rural-farm residence), SMSA's, urbanized areas, places of 2,500 inhabitants or more, and counties (by rural and rural-farm residence).

Volume II.

METROPOLITAN HOUSING CHARACTERISTICS

These reports, also designated as Series HC(2), cover most of the 1970 census housing subjects in considerable detail and cross-classification. There is one report for each SMSA, presenting data for the SMSA and its central cities and places of 50,000 inhabitants or more, as well as a national summary report.

Volume III.

BLOCK STATISTICS

One report, under the designation Series HC(3), is issued for each urbanized area showing data for individual blocks on selected housing and population subjects. The series also includes reports for the communities outside urbanized areas which have contracted with the Census Bureau to provide block statistics from the 1970 census.

Volume IV.

COMPONENTS OF INVENTORY CHANGE

This volume will contain data on the disposition of the 1960 inventory and the source of the 1970 inventory, such as new construction, conversions, mergers, demolitions, and other additions and losses. Cross-tabulations of 1970 and 1960 characteristics for units that have not changed and characteristics of the present and previous residence of recent movers will also be provided. Statistics will be shown for 15 selected SMSA's and for the United States and regions.

Volume V.

RESIDENTIAL FINANCE

This volume will present data regarding the financing of privately owned nonfarm residential properties. Statistics will be shown on amount of outstanding mortgage debt, manner of acquisition of property, homeowner expenses, and other owner, property, and mortgage characteristics for the United States and regions.

**Volume VI.
ESTIMATES OF "SUBSTANDARD"
HOUSING**

This volume will present counts of "substandard" housing units for counties and cities, based on the number of units lacking plumbing facilities combined with estimates of units with all plumbing facilities but in "dilapidated" condition.

**Volume VII.
SUBJECT REPORTS**

Each report in this volume will concentrate on a particular subject. Detailed information and cross-classifications will generally be provided on a national and regional level; in some reports, data for States or SMSA's may also be shown. Among the subjects to be covered are housing characteristics by household composition, housing of minority groups and senior citizens, and households in mobile homes.

Joint Population-Housing Reports

**Series PHC(1).
CENSUS TRACT REPORTS**

This series contains one report for each SMSA, showing data for most of the population and housing subjects included in the 1970 census.

**Series PHC(2).
GENERAL DEMOGRAPHIC TRENDS FOR
METROPOLITAN AREAS, 1960 TO 1970**

This series consists of one report for each State and the District of Columbia, as well as a national summary report, presenting statistics for the State and for SMSA's and their central cities and constituent counties. Comparative 1960 and 1970 data are shown on population counts by age and race and on such housing subjects as tenure, plumbing facilities, value, and contract rent.

**Series PHC(3).
EMPLOYMENT PROFILES OF SELECTED
LOW-INCOME AREAS**

This series consists of 76 reports, each presenting statistics on the social and economic characteristics of the residents of a particular low-income area. The data relate to low-income neighborhoods in 51 cities and seven rural poverty areas. Each report provides statistics on employment and unemployment, education, vocational training, availability for work, job history, and income, as well as on value or rent and number of rooms in the housing unit.

Additional Reports

**Series PHC(E).
EVALUATION REPORTS**

This open series will present the results of the extensive evaluation program conducted as an integral part of the 1970 census program, and relating to such matters as completeness of enumeration and quality of the data on characteristics.

**Series PHC(R).
PROCEDURAL REPORTS**

This open series presents information on various administrative and methodological aspects of the 1970 census, and will include a comprehensive procedural history of the 1970 census.

Computer Summary Tapes

The major portion of the results of the 1970 census are produced in a set of six tabulation counts. To help meet the needs of census users, these counts are designed to provide data with much greater subject and geographic detail than it is feasible or desirable to publish in printed reports. The data so tabulated are generally available—subject to suppression of certain detail where necessary to protect confidentiality—on magnetic computer tape, printouts, and microfilm, at the cost of preparing the copy.

First Count—source of the PC(1)-A reports; contains about 400 cells of data on the subjects covered in the PC(1)-B and HC(1)-A reports and tabulated for each of the approximately 250,000 enumeration districts in the United States.

Second Count—source of the PC(1)-B, HC(1)-A, and part of the PHC(1) reports; contains about 3,500 cells of data covering the subjects in these reports and tabulated for the approximately 35,000 tracts and 35,000 county subdivisions in the United States.

Third Count—source of the HC(3) reports; contains about 250 cells of data on the subjects covered in the PC(1)-B and HC(1)-A reports and tabulated for approximately 1,500,000 blocks in the United States.

Fourth Count—source of the PC(1)-C, HC(1)-B, and part of the PHC(1) reports; contains about 13,000 cells of data covering the subjects in these reports and tabulated for the approximately 35,000 tracts and 35,000 county subdivisions in the United States; also contains about 30,000 cells of data for each county.

Fifth Count—contains approximately 800 cells of population and housing data for 5-digit ZIP code areas in SMSA's and 3-digit ZIP code areas outside SMSA's; the ZIP code data are available only on tape.

Sixth Count—source of the PC(1)-D and HC(2) reports; contains about 260,000 cells of data covering the subjects in these reports and tabulated for States, SMSA's, and large cities.

The tapes are generally organized on a State basis. To use the First Count and Third Count tapes, it is necessary to purchase the appropriate enumeration district and block maps.

The term "cells" used herein to indicate the scope of subject content of the several counts refers to each figure or statistic in the tabulation for a specific geographic area. For example, in the Third Count, there are six cells for a cross-classification of race by sex: three categories of race (white, Negro, other race) by two categories of sex (male, female).

In addition to the above-mentioned summary tapes, the Census Bureau makes available for purchase certain sample tape files containing population and housing characteristics as shown on individual census records. These files contain no names or addresses, and the geographic identification is sufficiently broad to protect confidentiality. There are six files, each containing a 1-percent national sample of persons and housing units. Three of the files are drawn from the population covered by the census 15-percent sample and three from the population in the census 5-percent sample. Each of these three files provides a different type of geographic information: One identifies individual large SMSA's and, for the rest of the country, groups of counties; the second identifies individual States and, where they are sufficiently large, provides urban-rural and metropolitan-nonmetropolitan detail; and the third identifies State groups and size of place, with each individual record showing selected characteristics of the person's neighborhood.

TABLE FINDING GUIDE—Subjects by Type of Area and Table Number

This guide lists all subjects covered in this report, but does not indicate all cross-classifications (e.g., by age, race, or sex) or the historical data shown in some tables. An asterisk (*) indicates that the table presents data for Negroes exclusively; the symbol (†) means the table presents data for persons of Spanish heritage exclusively. Data on allocation rates appear in chapter B, tables B-1 to B-5, and chapter C, tables C-1 to C-4.

Subject	The State		Metropol- itan- nonmetro- politan residence	Standard metro- politan statistical areas Places of 50,000 or more (or central cities) Urbanized areas	Places with fewer than 50,000	Counties
	Total	Urban Rural nonfarm Rural farm Size of place				
NUMBER OF INHABITANTS	1,2,4,5,14,15	1-3,5	4	6-8,10-14	6-8,10	9,10 ¹
SUMMARY CHARACTERISTICS	16	16	—	16,40,41	16,32,40-42	16,33 ¹ ,43,44
GENERAL AND SOCIAL CHARACTERISTICS						
Age:						
Single years of age	19	—	—	—	—	—
Age groups	20,21,48,59,138	20,48,59,138	70	24,96†,138	28,31,112†	35,38,129†,134,136
Race	17,18,139	17,18	17	23	27,31	34,38
Nativity and parentage	45,49,60,138-142	49,60,138	71	81,138,140-142	102	119
Place of birth	45,50,61,140	50,61	72	82,91*,97†,140	102,108*,113†,117	119,125*,130†
Country of birth or country of origin	45,49,60,141,144	49,60	71	81,141,144	102	119
Mother tongue	49,60,142	49,60	71	81,142	102	119
Citizenship	143,144	—	—	143,144	—	—
Year of immigration	144	—	—	144	—	—
Residence in 1965	45,50,61,145	50,61	72	82,91*,97†	102,108*,113†,117	119,125*,130†
By selected classes of migrants	50,61,145,178	50,61	72	82	102	119
Year moved into present house	50,61	50,61	72	82	—	—
School enrollment	45,51,62,146,166,215	51,62,146,166,215	73	83,91*,97†,146,166,215	103,108*,113†,117	120,125*,130†
By percent enrolled	51,62,146	51,62,146	73	83,91*,97†,146	103,108*,113†	120,125*,130†
Years of school completed	46,51,62	51,62	73	83,91*,97†	103,108*,113†,117	120,125*,130†,134,136
Persons 16 to 24 years old	147	147	—	147	—	—
Family heads	158	158	—	158	—	—
By age	52,63,148	52,63,148	74	84,148	103	120
By occupation	179	—	—	—	—	—
By income	197,202	197	—	197,202	—	—
By poverty status	211,216	211,216	—	211,216	—	—
Percent by level of school completed	51,62	51,62	73	83	—	—
Vocational training	51,62,149,150	51,62	73	83	—	—
Work disability	52,63,169	52,63	74	84,169	—	—
Veteran status	50,61,151	50,61	72	82	102	119
Marital status	22,152,155,165	22,152,155,165	—	26,152,155,165	30	37
Marital history	52,63,152,159,160	52,63,152	74	84,152	—	—
Fertility:						
Children ever born	45,52,63,161,162	52,63,161	74	84,91*,97†,161	103,108*,113†	120,125*,130†,134,136
Own children under 5 years old	52,63,163	52,63,163	74	84,163	—	—
Households and household relationship	22,39,48,59,153	22,48,59,153	70	25,26,39,96†,153	29-32,39,112†	33 ¹ ,36-39,129†,134,136
Group quarters	22,48,52,59,63,154	22,48,52,59,63,154	70,74	25,26,84,96†,154	29-32,103,112†	33 ¹ ,36-38,120,129†,134,136
Inmates of institutions	22,48,52,59,63,154,155	22,48,52,59,63,154,155	70,74	25,26,84,96†,154,155	29-31,103,112†	36-38,120,129†,134,136
Families:						
By presence and number of own children under 18 years old	22,48,52,59,63,155,156	22,48,52,59,63,155,156	70,74	25,84,91*,96†,97†,155,156	29,103,108*,112†,113†	36,120,125*,129†,130†
By type and composition	22,48,52,59,63,155-158	22,48,52,59,63,155-158	70,74	25,84,91*,96†,97†,155-158	29,103,108*,112†	36,120,129†
By characteristics of head and wife	158	158	—	158	—	—
By income	47,57,68,198-205	57,68,198-201,205	79	89,94*,100†,198-205	107,111*,116†,118	124,128*,133†,135,137
By poverty status	58,69,207-214	58,69,207-214	80	90,95*,101†,207-214	107,111*,116†,118	124,128*,133†,135,137
Subfamilies	52,63,155	52,63,155	74	84,155	103,108*,113†	120,125*,130†
Unrelated individuals	22,48,59	22,48,59	70	25,26,96†	29-31,112†	36-38,129†,134,136
By age	153	153	—	153	—	—
By marital status	155	155	—	155	—	—
By income	198,200,205	198,200,205	—	198,200,205	—	—
By poverty status	207,213	207,213	—	207,213	—	—

Subject	The State		Metropol- itan- nonmetro- politan residence	Standard metro- politan statistical areas Places of 50,000 or more (or central cities) Urbanized areas	Places with fewer than 50,000	Counties
	Total	Urban Rural nonfarm Rural farm Size of place				