

# Appendix C.

## Data Collection and Processing Procedures

---

### CONTENTS

|  | Page |
|--|------|
| Enumeration and Residence Rules .....                                    | C-1  |
| United States .....  | C-1  |
| Puerto Rico .....  | C-3  |
| Major Components of the Census 2000 Plan.....                            | C-4  |
| Master Address File .....  | C-4  |
| Public Outreach and Marketing.....                                       | C-5  |
| Questionnaire Mailout/Mailback.....                                      | C-6  |
| Collecting Data on Populations Living in Nontraditional Households ..... | C-7  |
| Collecting Long Form Data to Meet Federal Requirements .....             | C-7  |
| Retrieving and Processing the Data From Returned Forms .....             | C-7  |
| Matching and Unduplication.....  | C-7  |
| Geographic Database Development – TIGER® .....                           | C-8  |
| Field Offices and Staffing .....   | C-8  |
| Data Collection: Basic Enumeration Strategy .....                        | C-9  |
| Special Populations .....  | C-11 |
| Telecommunications Support and Automated Data Processing .....           | C-13 |
| Quality Assurance.....   | C-13 |
| The Census 2000 Dress Rehearsal in 1998 .....                            | C-13 |
| Data Dissemination Through the Internet.....                             | C-14 |
| Evaluation and Preparation for 2010 .....                                | C-14 |
| Glossary .....   | C-14 |

### ENUMERATION AND RESIDENCE RULES

In accordance with census practice dating back to the first U.S. census in 1790, each person was to be enumerated as an inhabitant of his or her “usual residence” in Census 2000. Usual residence is the place where the person lives and sleeps most of the time. This place is not necessarily the same as the person’s legal residence or voting residence. In the vast majority of cases, however, the use of these different bases of classification would produce substantially the same statistics, although there might be appreciable differences for a few areas.

The implementation of this practice has resulted in the establishment of rules for certain categories of people whose usual place of residence is not immediately apparent. Furthermore, this practice means that people were not always counted as residents of the place where they happened to be staying on Census Day (April 1, 2000).

#### United States

**Enumeration rules.** Each person whose usual residence was in the United States was to be included in the census, without regard to the person’s legal status or citizenship. As in previous censuses, people specifically excluded from the census were citizens of foreign countries temporarily traveling or visiting in the United States who had not established a residence.

Americans temporarily overseas were to be enumerated at their usual residence in the United States. With some exceptions, Americans with a usual residence outside the United States were not enumerated in Census 2000. U.S. military personnel and federal civilian employees stationed outside the United States and their dependents living with them, are included in the population counts for the 50 states for purposes of Congressional apportionment but are excluded from all other tabulations for states and their subdivisions. The counts of overseas U.S. military personnel,

---

federal civilian employees, and their dependents were obtained from administrative records maintained by the employing federal departments and agencies. Other Americans living overseas who were not affiliated with the U.S. government were not included in the census.

**Residence rules.** Each person included in the census was to be counted at his or her usual residence the place where he or she lives and sleeps most of the time. If a person had no usual residence, the person was to be counted where he or she was staying on Census Day.

People temporarily away from their usual residence on Census Day, such as on a vacation or business trip, were to be counted at their usual residence.

**Armed forces personnel in the United States.** Members of the U.S. Armed Forces were counted at their usual residence (the place where they lived and slept most of the time), whether it was on or off the military installation. Family members of armed forces personnel were counted at their usual residence (for example, with the armed forces person or at another location).

Personnel assigned to each Navy and Coast Guard vessel with a U.S. homeport were given the opportunity to report an onshore residence where they usually stayed when they were off the ship. Those who reported an onshore residence were counted there; those who did not were counted at their vessel's homeport.

**Personnel on U.S. flag merchant vessels.** Crews of U.S. flag merchant vessels docked in a U.S. port, sailing from one U.S. port to another U.S. port, or sailing from a U.S. port to a Puerto Rico port were counted at their usual onshore residence if they reported one. Those who did not were counted as residents of the ship and were assigned as follows:

- The U.S. port, if the vessel was docked there on Census Day.
- The port of departure, if the ship was sailing from one U.S. port to another U.S. port, or from a U.S. port to a Puerto Rico port.

Crews of U.S. merchant ships docked in a foreign port (including the U.S. Virgin Islands, American Samoa, the Commonwealth of the Northern Mariana Islands, and Guam), sailing from one foreign port to another foreign port, sailing from a U.S. port to a foreign port, or sailing from a foreign port to a U.S. port were not included in the census.

**People away at school.** College students were counted as residents of the area in which they were living while attending college, as they have been since the 1950 census. Children in boarding schools below the college level were counted at their parental home.

**People in institutions.** People under formally authorized, supervised care or custody, such as in federal or state prisons; local jails; federal detention centers; juvenile institutions; nursing or convalescent homes for the aged or dependent; or homes, schools, hospitals, or wards for the physically handicapped, mentally retarded, or mentally ill; or in drug/alcohol recovery facilities were counted at these places.

**People in general hospitals.** People in general hospitals or wards (including Veterans Affairs hospitals) on Census Day were counted at their usual residence. Newborn babies were counted at the residence where they would be living.

**People in shelters.** People staying on Census Day at emergency or transitional shelters with sleeping facilities for people without housing, such as for abused women or runaway or neglected youth, were counted at the shelter.

**People with multiple residences.** People who lived at more than one residence during the week, month, or year were counted at the place where they lived most of the time. For example, commuter workers living away part of the week while working were counted at the residence where they stayed most of the week. Likewise, people who lived in one state but spent the winter in another state with a warmer climate ("snowbirds") were to be counted at the residence where they lived most of the year.

---

**People away from their usual residence on Census Day.** Temporary, migrant, or seasonal workers who did not report a usual U.S. residence elsewhere were counted as residents of the place where they were on Census Day.

In some areas, natural disasters (hurricanes, tornadoes, flooding, and so forth) displaced households from their usual place of residence. If these people reported a destroyed or damaged residence as their usual residence, they were counted at that location.

People away from their usual residence were counted by means of interviews with other members of their families, resident managers, or neighbors.

## **Puerto Rico**

**Enumeration rules.** Each person whose usual residence was in Puerto Rico was to be included in the census, without regard to the person's legal status or citizenship. As in previous censuses, people specifically excluded from the census were citizens of foreign countries temporarily traveling or visiting in Puerto Rico who had not established a residence.

Americans usually living in Puerto Rico but temporarily overseas were to be enumerated at their usual residence in Puerto Rico. Americans with a usual residence outside Puerto Rico were not counted as part of the Puerto Rico resident population.

**Residence rules.** Each person included in the census was to be counted at his or her usual residence the place where he or she lives and sleeps most of the time. If a person had no usual residence, the person was to be counted where he or she was staying on Census Day.

People temporarily away from their usual residence were to be counted at their usual residence. People who moved around Census Day were counted at the place they considered to be their usual residence.

**Armed forces personnel in Puerto Rico.** Members of the U.S. Armed Forces were counted at their usual residence (the place where they lived and slept most of the time), whether it was on or off the military installation. Family members of armed forces personnel were counted at their usual residence (for example, with the armed forces person or at another location).

Personnel assigned to each Navy and Coast Guard vessel with a Puerto Rico homeport were given the opportunity to report an onshore residence where they usually stayed when they were off the ship. Those who reported an onshore residence were counted there; those who did not were counted at their vessel's homeport.

**Personnel on U.S. flag merchant vessels.** Crews of U.S. flag merchant vessels docked in a Puerto Rico port, sailing from one Puerto Rico port to another Puerto Rico port, or sailing from a Puerto Rico port to a U.S. port were counted at their usual onshore residence if they reported one. Those who did not were counted as residents of the ship and were attributed as follows:

- The Puerto Rico port if the vessel was docked there on Census Day.
- The port of departure if the ship was sailing from one Puerto Rico port to another Puerto Rico port or from a Puerto Rico port to a U.S. port.

Crews of U.S. merchant ships docked in a foreign port (including the U.S. Virgin Islands, American Samoa, the Commonwealth of the Northern Mariana Islands, and Guam), sailing from a Puerto Rico port to a foreign port, or sailing from a foreign port to a Puerto Rico port were not included in the census.

**People away at school.** College students were counted as residents of the area in which they were living while attending college, as they have been since the 1950 census. Children in boarding schools below the college level were counted at their parental home.

**People in institutions.** People under formally authorized, supervised care or custody, such as in federal or state prisons; local jails; federal detention centers; juvenile institutions; nursing or convalescent homes for the aged or dependent; or homes, schools, hospitals, or wards for the physically handicapped, mentally retarded, or mentally ill; or in drug/alcohol recovery facilities were counted at these places.

---

**People in general hospitals.** People in general hospitals or wards (including Veterans Affairs hospitals) on Census Day were counted at their usual residence. Newborn babies were counted at the residence where they would be living.

**People in shelters.** People staying on Census Day at emergency or transitional shelters with sleeping facilities for people without housing, such as for abused women or runaway or neglected youth, were counted at the shelter.

**People with multiple residences.** People who lived at more than one residence during the week, month, or year were counted at the place where they lived most of the time. For example, commuter workers living away part of the week while working were counted at the residence where they stayed most of the week.

**People away from their usual residence on Census Day.** Temporary, migrant, or seasonal workers who did not report a usual Puerto Rico residence elsewhere were counted as residents of the place where they were on Census Day.

In some areas, natural disasters (hurricanes, tornadoes, flooding, and so forth) displaced households from their usual place of residence. If these people reported a destroyed or damaged residence as their usual residence, they were counted at that location.

People away from their usual residence were counted by means of interviews with other members of their families, resident managers, or neighbors.

## **MAJOR COMPONENTS OF THE CENSUS 2000 PLAN**

The Census Bureau prepared the Census 2000 plan to ensure the most accurate decennial census legally possible. This plan included data collection from 100 percent of households and housing units. In addition, the plan included an extensive statistical operation to measure and correct overall and differential coverage of U.S. residents in Census 2000. This operation consisted of a scientific sample of approximately 300,000 housing units and used regional groupings to generate corrected counts. To ensure that Census 2000 will be both more accurate and more cost-effective than the 1990 Census, the Census Bureau reviewed its procedures with input from a wide array of experts. In addition, the Census Bureau and Department of Commerce officials held more than 100 briefings for the members of Congress and their staff on the plan for Census 2000. The result has been an innovative departure from past practices that substantially increased overall accuracy and addressed the differential undercount of children, renters, and minorities. At the same time, the new methods of enumeration saved money and delivered results more quickly. The major components of the plan for Census 2000 included:

### **1. The Master Address File**

To conduct Census 2000, the Census Bureau needed to identify and locate an estimated 118 million housing units in the Nation. The Census Bureau accomplished this goal by developing and maintaining the Master Address File (MAF). This vital operation took place with the assistance of the U.S. Postal Service (USPS); other federal agencies; tribal, state and local governments; community organizations; and by an intensive canvass of selected areas. The resulting file was more comprehensive than ever before.

In 1990, the Census Bureau relied on address lists purchased from vendors. As these lists were originally generated for marketing purposes, they proved to be less accurate in low-income areas. As a result, during the 1990 census, housing units were missed often enough to contribute notably to the undercount problem. Plans for Census 2000 were designed to address weaknesses found in the 1990 address list. The Census 2000 MAF started with the USPS address list, a list that does not discriminate against certain areas because of their marketing potential. Partnerships with state and local officials, community organizations, and tribal governments also played an important role in making sure the MAF is accurate; the local officials who knew the areas best helped develop the MAF. Finally, the Bureau made intensive efforts to create address lists in rural areas well in advance of the census.

---

**City-style addresses.** The USPS uses the term “city-style” for an address such as “123 Main Street,” even though such an address may occur in small towns and increasingly along country roads. In areas where the USPS delivers mail primarily to city-style addresses, the Census Bureau created the MAF by combining addresses from the 1990 Census Address Control File with those addresses in the USPS Delivery Sequence File (DSF). The DSF is a national file of individual delivery point addresses. As part of a cooperative agreement, the USPS provided the Census Bureau with updated DSFs on a regular basis. The Bureau then located these addresses in its computer mapping system called TIGER® (Topologically Integrated Geographic Encoding and Referencing). If an address could not be located, the location was researched and resolved through an office operation or through assistance from local partners. As a result of this research, the Bureau identified new features and corrected and added address ranges to the TIGER® database.

**Noncity-style addresses.** In late 1998 and early 1999, the Census Bureau launched a comprehensive effort to canvass areas where most residences did not have city-style addresses. Over 30,000 canvassers visited approximately 22 million residences without a street address to enter their locations in the TIGER® system. The combination of innovative use of computer data and technology along with these visits allowed the Bureau to construct the most accurate address list ever, giving field enumerators more time to meet other challenges presented by the 2000 count.

**Remote areas.** In a few extremely remote and sparsely settled areas, census enumerators created the address list at the time of the initial census data collection while canvassing their assignment area and picking up or completing unaddressed questionnaires that the USPS previously had delivered to each household.

**Nontraditional living quarters.** A separate operation built an inventory of all facilities that were not traditional living quarters; for example, prisons and hospitals. The Bureau interviewed an official at each location using a Facility Questionnaire. The responses to the questionnaire identified each group quarters and any housing units associated with the location. The Bureau classified each group quarters and its associated housing units at the location according to whether they would be enumerated as part of special place enumeration or through regular enumeration. The Bureau added these group quarters and housing units to the MAF and linked them to the TIGER® database.

**Local government partnerships.** The Bureau relied on local knowledge to build the MAF. State, local, and tribal governments; regional and metropolitan planning agencies; and related nongovernmental organizations were encouraged to submit locally developed and maintained city-style address lists to the Census Bureau to enhance the MAF. The Bureau matched the local lists both to the MAF and TIGER® database and verified the status of each newly identified address through ongoing matches to updated address information from the USPS, other independent sources, and its own field operations. The Local Update of Census Addresses (LUCA) program was a partnership that allowed local and tribal governments to designate a liaison to review the portion of the MAF that covered their jurisdiction to help ensure its completeness. After processing the LUCA input, the Census Bureau provided feedback on the status of the adds, deletes, and corrections of addresses to the liaisons. The updated address list then was used to deliver census questionnaires.

## **2. Public Outreach and Marketing**

In 1990, the mail response rate dropped in spite of the Census Bureau’s support of a public service announcement (PSA) effort that aired donated advertisements. Part of this drop was caused by the Bureau’s inability to ensure that PSAs were broadcast at optimum times and in appropriate markets. An evaluation of the 1990 PSA campaign noted that the ads were seldom placed at optimal times because decisions about when to air PSAs rested with local radio and television stations. Sixty percent of the U.S. population received 91 percent of the census advertising impact; 40 percent received only 9 percent. Based on its studies of prior outreach campaigns, the Bureau concluded that the professional control of a paid media campaign would produce the best results. Census 2000 launched a vigorous public outreach campaign to educate everyone about the importance of being counted. Among the improvements in public outreach and marketing were:

---

**Partnerships/targeted community outreach.** The Census Bureau built partnerships with local and tribal governments, businesses, and community groups to get the word out, to endorse the census, and to encourage constituents to respond. Beginning in 1996 and expanding in 1998, the Bureau hired government and community specialists to build relationships with local community and service-based organizations, focusing on groups representing traditionally undercounted populations. The Bureau deployed an extensive outreach program to reach schools, public sector employees, American Indians, and religious organizations. Businesses, nonprofit groups, and labor organizations also were asked to endorse participation and to publicize the census through employee newsletters, inserts with paychecks, and through communications with members and local chapters.

**Direct mail.** The census questionnaire and related materials delivered to individual addresses carried the same themes and messages as the overall campaign.

**Public relations.** The Census Bureau used public meetings and the news media to inform the public about the value of the census and to encourage response. Communications specialists were assigned to each field office to perform media outreach, to respond to media inquiries, and to coordinate the dissemination of the Census 2000 message. In many communities, the Census Bureau established local broadcaster/news director committees to emphasize Census 2000 to television viewers and radio listeners through broadcast segments and editorials in newspapers.

**Paid advertising.** The Census Bureau planned a targeted campaign to reach everyone through ads in newspapers, magazines, billboards, posters, radio, and television. A private advertising firm designed and implemented the Census 2000 advertising campaign. The Census Bureau conducted a first-ever paid advertising campaign, including a national media campaign aimed at increasing mail response. The campaign included advertising directed at raising mail response rates among historically undercounted populations, with special messages targeted to hard-to-enumerate populations. Advertising also focused on encouraging cooperation during the nonresponse follow-up procedures.

**Media public relations.** The Census Bureau assigned media specialists to the regional census centers to cultivate local press contacts and respond to local media inquiries.

**Promotion and special events.** A variety of special events, including parades, athletic events and public services television documentaries were cosponsored by state, local, and tribal governments and by community organizations and businesses to motivate people to respond.

**More ways to respond.** In 2000, in addition to mailing the census questionnaires, the Census Bureau made the forms available in stores and malls, in civic or community centers, in schools, and in other locations frequented by the public. A well-publicized, toll-free telephone number was available for those who wished to respond to the census by telephone. People also had the option to respond to the short form via the Internet.

**Multiple languages.** In 2000, as in all prior decennial censuses, questionnaires were in English (the Census Bureau has made Spanish-language questionnaires available in the past). However, for the first time in a decennial census, households had the option to request and receive questionnaires in five other languages (Spanish, Chinese, Korean, Tagalog, and Vietnamese). In addition, questionnaire assistance booklets were available in 49 languages.

### **3. Questionnaire Mailout/Mailback**

In Census 2000, the questionnaire mailout/mailback system was the primary means of census-taking, as it has been since 1970. The short form was delivered to approximately 83 percent of all housing units. The short form asked only the basic population and housing questions, while the long form included additional questions on the characteristics of each person and of the housing unit. The long form was delivered to a sample of approximately 17 percent of all housing units.

USPS letter carriers delivered questionnaires to the vast majority of housing units that had city-style addresses. In areas without such addresses, enumerators hand delivered addressed census questionnaires to each housing unit. In very remote or sparsely populated areas, enumerators visited each housing unit and picked up or completed unaddressed questionnaires that the USPS previously delivered to each unit.

---

#### **4. Collecting Data on Populations Living in Nontraditional Households**

During a decennial census, the Census Bureau not only counts people living in houses and apartments, but also must count people who live in group quarters and other nontraditional housing units, as well as people with no usual residence. These units include nursing homes, group homes, college dormitories, migrant and seasonal farm worker camps, military barracks or installations, American Indian reservations, and remote areas in Alaska.

Some of the methods that were used for these special populations are listed below:

- The Census Bureau designed an operation for Census 2000 called Service-Based Enumeration (SBE) to improve the count of individuals who might not be included through standard enumeration methods. The SBE operation was conducted in selected service locations, such as shelters and soup kitchens, and at targeted outdoor locations.
- Another special operation counted highly transient individuals living at recreational vehicle campgrounds and parks, commercial or public campgrounds, marinas, and even workers' quarters at fairs and carnivals.
- The Census Bureau worked with tribal officials to select the appropriate data collection methodologies for American Indian reservations.
- Remote areas of Alaska, often accessible only by small airplanes, snowmobiles, four wheel-drive vehicles, or dogsleds, were enumerated beginning in mid-February. This special timing permitted travel to these areas while conditions are most favorable.
- The Census Bureau worked with the Department of Defense and the U.S. Coast Guard to count individuals living on military installations, and with the U.S. Maritime Administration to identify maritime vessels for enumeration.

#### **5. Collecting Long Form Data to Meet Federal Requirements**

The census is the only data gathering effort that collects the same information from enough people to get comparable data for every geographic area in the United States. The Census Bureau has used the long form on a sample basis since 1940 to collect more data, while reducing overall respondent burden. The Census 2000 long form asked questions addressing the same 7 subjects that appeared on the short form, plus an additional 27 subjects which were either specifically required by law to be included in the census or were required in order to implement other federal programs.

#### **6. Retrieving and Processing the Data From the Returned Forms**

The Census Bureau contracted with the private sector to secure the best available data capture technology. This technology allowed the Census Bureau to control, manage, and process Census 2000 data more efficiently.

The Census 2000 Data Capture System has been a complex network of operational controls and processing routines. The Census Bureau recorded a full electronic image of many of the questionnaires, sorted mail-return questionnaires automatically, used optical mark recognition for all check-box items, and used optical character recognition to capture write-in character based data items. The system allowed the Census Bureau to reduce the logistical burdens associated with handling large volumes of paper questionnaires. Once forms were checked in, prepared, and scanned, all subsequent operations were accomplished using the electronic image and data capture.

#### **7. Matching and Unduplication**

One of the main goals of Census 2000 was to make it simpler for people to be counted by having census forms available in public locations and providing multiple language translations. Responses also were accepted over the telephone and, for the short form only, on the Internet. These options made it easier for everyone to be counted, but increased the possibility of multiple responses for a given person and household. Advances in computer technology in the areas of computer storage, retrieval, and matching, along with image capture and recognition, gave the Census Bureau the flexibility to provide multiple response options without incurring undue risk to

---

the accuracy of the resulting census data. Unduplication of multiple responses in past censuses required massive clerical operations. Modern technology allowed the Census Bureau to spot and eliminate multiple responses from the same household.

## **8. Geographic Database Development—TIGER®**

The Census Bureau's TIGER® (Topologically Integrated Geographic Encoding and Referencing) system provided the geographic structure for the control of the data collection, tabulation, and dissemination operations for Census 2000. The TIGER® system links each living quarter to a spatial location, each location to a specific geographic area, and each geographic area to the correct name or number and attributes. The database constantly changes; for example, when new streets are built and the names and address ranges of existing streets change. To ensure that the TIGER® database is complete and correct, the Census Bureau works with other federal agencies; state, local and tribal governments; and other public and private groups to update both its inventory of geographic features and its depiction of the boundaries, names, and attributes of the various geographic entities for which the Census Bureau tabulates data.

The Census Bureau obtains updates to the features in the TIGER® system, including associated address ranges, from its various address list improvement activities, from partnership efforts like the Local Update of Census Addresses (LUCA) program, from digital files provided by some local and tribal governments, and from local and tribal governments in response to a preview of the census map of their jurisdictions.

As a part of updating the TIGER® system, the Census Bureau conducted boundary surveys in 1998 and 1999 to determine the boundaries that were in effect on January 1, 2000, which were the official Census 2000 boundaries for functioning governments. The Census Bureau also relied on other programs to update the TIGER® boundaries data, including a program that allowed local or tribal officials to review proposed Census 2000 boundaries a program that allowed local and tribal participants the opportunity to delineate Census 2000 participant statistical areas (block groups, census county divisions, census designated places, and census tracts) and additional programs that offered participants the opportunity to identify other areas for which the Census Bureau would tabulate data (for example, traffic analysis zones).

## **9. Field Offices and Staffing**

The Census Bureau opened a national network of temporary offices from which employees collected and processed the data for Census 2000. Establishing the office network required, for most offices, the leasing of office space, purchasing furniture and equipment, purchasing and installing computer hardware and software, and establishing voice and data line connections. The plan for the office structure included:

- **12 Regional Census Centers (RCCs).** Through a network of Census Field Offices, the RCCs managed all census field data collections operations, address listings, and address list enhancement for city-style address areas; coordinated the LUCA program; produced maps; updated TIGER®; worked with local participants in the Public Law 94–171 Redistricting Data Program; and recruited temporary staff.
- **402 Census Field Offices (CFOs).** Opened in September 1998, these offices helped with address listing; conducted local recruiting; and performed clerical review of completed field address listing work.
- **520 Local Census Offices (LCOs).** These offices produced enumerator maps and assignments; conducted local recruiting; conducted outreach and promotion; conducted group quarters and service-based enumeration activities; conducted update/leave and list/enumerate operations; conducted nonresponse follow-up, coverage improvement follow-up, and address verifications; and performed the block canvass operations.
- **3 New Data Capture Centers (DCCs).** These centers checked in mail returns, prepared questionnaires, and conducted data capture.
- **1 National Processing Center (NPC).** In addition to performing the functions of a Data Processing Center, it processed address listing data and performed coding of questionnaire data.

---

To conduct a successful Census 2000, the Census Bureau recruited and tested hundreds of thousands of applicants for a wide range of positions, such as local census office managers, enumerators, partnership specialists, media specialists, and clerks. This required an extraordinary recruiting effort throughout the country. Every job applicant was required to pass a written test and was screened for criminal history. Applicants selected for employment had to take an oath of office and sign an affidavit agreeing not to disclose census information.

Many factors converged to present the Census Bureau with unprecedented challenges in hiring, retraining, and training the necessary employees for Census 2000. To address this challenge, the Census Bureau implemented several new approaches:

- Innovative methods of setting pay and incentives.
- Expanding the potential labor force by working with other federal agencies and state agencies to reduce barriers presented by various income transfer programs, and encouraging recipients of these programs to work for the Census Bureau. Consistent with these efforts, the Census Bureau hired more welfare-to-work employees than any other federal agency.
- Earlier and expanded training for enumerators.

#### **10. Data Collection: Basic Enumeration Strategy**

To ensure that the Census Bureau obtained a completed questionnaire from every household, or as close to that as possible, the Census Bureau developed a ten-part, integrated enumeration strategy.

- The first part of this strategy ensured that a questionnaire was delivered to every housing unit, by one of three data collection methods:
  - **Mailout/mailback.** U.S. Postal Service delivered questionnaires to every “city style” housing unit with a street name and house number.
  - **Update/leave.** Census enumerators delivered questionnaires to housing units without street names and house numbers to be mailed back, mainly in rural areas, and corrected and updated the address list and maps for any additions or errors.
  - **List/enumerate.** In remote and sparsely populated areas, enumerators visited every housing unit and completed the enumeration as delivered.
- The second part of this strategy provided people with assistance, as needed, to complete and return their questionnaires.
  - **Telephone questionnaire assistance (TQA).** The Census Bureau operated a toll-free TQA system, in English, Spanish, and several other languages, providing automated touch-tone answers to common questions, personal operator answers to those requesting it, and special service for the hearing impaired to assist them in completing a short form. Callers also could request a questionnaire.
  - **Internet.** Respondents were able to access an Internet Web site to both receive assistance and, for short forms, submit their responses.
  - **Questionnaire assistance centers.** The Census Bureau opened Walk-In Questionnaire Assistance Centers in convenient locations to assist respondents with filling out questionnaires in person. Bilingual staff was available in these centers.
  - **Questionnaire assistance guides.** Questionnaire Assistance Guides were available in 49 languages.
- The third part of this strategy provided a means for people who believed they had not received a questionnaire or were not included on one. Part of this operation was targeted to members of historically undercounted groups. The major element of this operation was the distribution of “Be Counted Questionnaires.” The Census Bureau distributed these questionnaires at public locations, such as Walk-In Questionnaire Assistance Centers and some public and private facilities, staffed with bilingual competencies when appropriate. These forms were available in English, Spanish, Korean, Chinese, Vietnamese, and Tagalog.

- 
- The fourth part of this strategy was designed to enumerate people who did not live in traditional housing units, including group quarters situations, such as nursing homes and college dormitories; people living in migrant farm worker camps, on boats, on military installations; and federal employees living overseas. This part of the strategy was expanded further because the Census 2000 Dress Rehearsal results indicated that, compared to 1990, many more people did not live in traditional housing units.
    - **Group quarters enumeration.** This operation identified the location of all group living quarters and made advance visits to each group quarter. Census staff listed all residents in April 2000 and distributed questionnaire packets.
    - **Transient night operation.** Transient Night enumerated people living a mobile lifestyle by visiting and interviewing people at racetracks, commercial or public campgrounds and those for recreational vehicles, fairs and carnivals, and marinas.
    - **Remote Alaska enumeration.** This operation sent out enumerators to deliver and complete questionnaires for people living in outlying or remote settlements in Alaska.
    - **Domestic military/maritime enumeration.** The Census Bureau, in cooperation with the Department of Defense and U.S. Coast Guard, identified living quarters and housing units on military installations and ships assigned to a U.S. home port and used appropriate enumeration methods.
    - **Overseas enumeration.** The Census Bureau, in cooperation with the Department of Defense and other departments, counted federal employees assigned overseas (including members of the armed forces) and their dependents, for apportionment purposes.
  - The fifth part of this strategy targeted people with no usual residence or address. This operation was conducted at selective service locations, such as shelters and soup kitchens and non-sheltered outdoor locations.
  - The sixth part of this strategy deployed special data collection methods to improve cooperation and enumeration in certain hard-to-enumerate areas.
    - Regional Census Centers used the planning database and their knowledge of local conditions to identify appropriate areas for targeted methods. A team of enumerators then went to targeted areas, such as areas with high concentrations of multiunit buildings, safety concerns or low enumerator production rates, and conducted team enumerations.
    - Mail response rates and maps were available to local and tribal officials so they could work with Census Bureau staff to identify low-response areas and implement additional outreach and publicity efforts and targeted enumeration efforts.
    - In partnership with local and tribal governments and community-based organizations, local census offices established Walk-In Questionnaire Assistance Centers in locations, such as community centers and large apartment buildings, to provide assistance in English, Spanish, and other and foreign languages.
    - The Be Counted Program made unaddressed questionnaires available in the Walk-In Assistance Centers and other locations.
    - Letters were mailed to managers of large multiunit structures and gated communities informing them of upcoming census operations.
    - In preidentified census blocks, census enumerators canvassed the blocks, updated the address list, and delivered and completed census questionnaires for all housing units.
    - In preidentified blocks originally classified as “Mailout/Mailback” areas, enumerators delivered the questionnaire and updated the address list (Urban Update/Leave).
  - The seventh part of this strategy, coverage-edit and telephone follow-up, reviewed completed questionnaires for potential missing, incomplete, or inconsistent data.

- 
- **Coverage edit.** The Census Bureau checked completed questionnaires for discrepancies between the number of persons reported and the number of persons for whom information was provided, forms returned where population count was blank, and forms for certain households that contained complex living arrangements.
  - **Follow-up.** Telephone clerks contacted and reinterviewed the households with discrepancies identified after mail returns were data captured; field staff resolved discrepancies found on enumerator returned questionnaires.
  - **Content edit.** Computer operations identified missing or incomplete responses to population or housing units and used statistical imputation to complete the information.
  - The eighth part of this strategy, nonresponse follow-up (NRFU), was the effort to secure a response in Census 2000 from every housing unit and resident. One hundred percent of nonresponding households were followed up.
    - In the initial period, the Census Bureau used reminder publicity urging people to return their questionnaires.
    - Following the period of mail response, nonresponding households were identified and listed.
    - Enumerators visited all nonresponding addresses to obtain a completed questionnaire for each household.
    - In mailout/mailback areas, enumerators also followed up 100 percent of housing units identified as nonexistent or vacant by the U.S. Postal Service.
    - In update/leave areas, enumerators followed up 100 percent of housing units where the Census Bureau was unable to deliver questionnaires.
    - The Census Bureau conducted quality assurance checks of NRFU to ensure the completeness and accuracy of the operations.
  - The ninth part of strategy involved additional operations to improve the coverage of Census 2000.
    - In mailout/mailback areas, enumerators revisited addresses for which questionnaires were returned in NRFU reporting the housing unit as vacant or delete and which were not initially identified by the U.S. Postal Service as undeliverable as addressed.
    - In update/leave areas, enumerators revisited addresses for which a questionnaire was returned as vacant or nonexistent in NRFU, but the questionnaire was not returned as undeliverable during the update/leave operation.
    - In both mailout/mailback and update/leave areas, mail returns checked in but not data captured were rechecked and, if necessary, revisited.
  - The tenth part of this strategy was unduplication, which involved reviewing and selecting person information when more than one questionnaire data set was reported for a single address. Dress Rehearsal results showed that the multiple ways in which people could respond to the census increased the possibility of more than one response being submitted for a given person or household. Automated matching technologies allowed the Census Bureau to resolve situations where more than one form was received for an address.

## 11. Special Populations

### American Indian and Alaska Native Areas and Hawaiian Home Lands

The Census Bureau based its strategy for enumerating the populations in the American Indian and Alaska Native Areas (AIANAs) and Hawaiian home lands on building partnerships for:

- **Address list development.** The Census Bureau used U.S. Postal Service's Delivery Sequence Files in AIANAs and Hawaiian home lands where there were city-style addresses. In other areas, the Census enumerators used the "update/leave" method where a form is left with the respondent for return by mail. In more remote areas, the census enumerator actually delivered the

---

form and conducted the census interview all in one visit. Tribal governments had an opportunity to participate in the LUCA program. The Census Bureau worked with tribal officials to select the appropriate data collection methodology for each area.

- **Geographic programs.** There were many programs available to review and define geographic areas (see [Appendix A](#) for more details).
- **Marketing.** Census Bureau staff and tribal liaisons compiled lists of available media for paid advertising and promotion. The Census Bureau also enlisted the help of tribal liaisons and locally established “Complete Count Committees” to assist with promotional activities.
- **Field operations.** The Census Bureau worked with tribal governments to assist in all levels of field operations, including training local staff in cultural awareness, assisting in recruiting efforts, and identifying locations for census questionnaire assistance centers.
- **Data dissemination.** While most data were processed in the same way as data for rest of the nation, the Census Bureau worked with tribal governments to meet their data needs.

### **Puerto Rico**

The Census 2000 operations in Puerto Rico were comparable to activities in the 50 states and the District of Columbia. The Census Bureau worked in partnership with the government of Puerto Rico to ensure that Census 2000 data met the federal legal requirements.

- **Build partnerships at every stage of the process.** The Census Bureau entered a Memorandum of Agreement with the governor of Puerto Rico which outlined mutual roles and responsibilities. In consultation with the government of Puerto Rico, census questionnaire content was developed to meet the legislative and programmatic needs of Puerto Rico. A separate advertisement and promotion campaign was conducted in Puerto Rico to build awareness of the census and boost participation. Address list development allowed Puerto Rico to participate in the LUCA program.
- **Census questionnaires.** Census questionnaires were readily available in Spanish and also in English, if requested. In Puerto Rico, only update/leave method was used to distribute questionnaires. However, questionnaires also were placed in Walk-In Questionnaire Assistance Centers and other locations identified through consultation with local partners.
- **Use of technology.** The Census Bureau made use of the same technological advances that were used in the United States. Many operations performed clerically in 1990 were automated. Data users have access to Census 2000 data products through the Internet using the American FactFinder (AFF) system. The AFF offers a separate user interface utilizing the Spanish language for Census 2000 Puerto Rico data.
- **Special techniques to improve coverage.** The update/leave methodology for census data collection was used for the first time in Puerto Rico. Census enumerators updated the Master Address File for Puerto Rico while delivering questionnaires. Respondents had the opportunity to complete the census questionnaires and return them by mail.

### **Island Areas**

The Census Bureau conducted the Census 2000 operations in American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the U.S. Virgin Islands (collectively referred to as the “Island Areas”) in partnerships with the government of each area. These partnerships ensured that Census 2000 data met federal legal requirements, as well as the specific needs of each area. The Census 2000 operations in the Island Areas were built around the following:

- **Data collection.** Data collection in the Island Areas used the list/enumerate method. This decision was based on recommendations from Island Area representatives and an analysis of the various data collection methodologies. Unlike stateside list/enumerate procedures, the Census Bureau delivered Advance Census Reports before the list/enumerate operation and asked respondents to complete the form and hold it for enumerator to pick up.

- 
- **Build partnerships at every stage of the process.** The Census Bureau developed and signed a Memorandum of Agreement with the governor of each Island Area that outlined mutual roles and responsibilities. In consultation with the governments of the Island Areas, census questionnaire content was developed to meet the legislative and programmatic needs of each Island Area. A separate advertisement and promotion campaign was developed for each Island Area to build awareness of the census and boost participation.
  - **Census questionnaires.** Census questionnaires and other forms were readily available to respondents in convenient locations identified through consultation with local partners.
  - **Use of technology.** The Census Bureau made greater use of the telephone to provide assistance to respondents with questions about Census 2000. Data users have access to Census 2000 data and products through the Internet using the American FactFinder system.

## **12. Telecommunications Support and Automated Data Processing**

Using dedicated links and other secure lines, the Census 2000 telecommunications network linked all census offices including: Census Headquarters in Suitland, Maryland, the 520 Local Census Offices, the 12 Regional Census Offices, the 12 Regional Census Centers, the Puerto Rico Area Office, the Maryland Computer Center in Bowie, the National Processing Center in Jeffersonville, Indiana, and the three contracted Data Capture Centers (Phoenix, AZ, Pomona, CA, and Essex, MD). The Census Bureau also established communication links with planned commercial telephone centers to assist with the Telephone Questionnaire Assistance program and the coverage edit follow-up program.

The use of electronic imaging reduced the logistical and staffing requirements of handling large volumes of paper questionnaires. Some components of data capture were performed by private-sector partners. The Census Bureau used commercially available advanced hardware and software rather than limiting itself to creating in-house solutions.

The most significant features of the Data Capture System included (1) work divided among four centers, (2) full electronic imaging and processing of questionnaires, (3) automated sorting of mailed responses, (4) optical mark recognition for check-box data, (5) optical character recognition for write-in data with automated processes to resolve difficult cases, and (6) quality assurance checks.

## **13. Quality Assurance**

To detect, correct, and minimize performance errors in critical census operations, the Census Bureau developed individual quality assurance plans for all activities that could contribute to errors in outcome, such as misprinted census forms, inaccurate maps or address lists, faulty intelligent character recognition, inadequate training of enumerators, and miskeyed entries.

## **14. The Census 2000 Dress Rehearsal in 1998**

A good dress rehearsal is crucial to a successful census, and the key to any dress rehearsal is making it as much like the actual event as possible. The Census Bureau conducted Census 2000 Dress Rehearsal in three sites: Sacramento, California; Columbia, South Carolina along with 11 surrounding counties in north central South Carolina; and the Menominee American Indian Reservation in northeastern Wisconsin.

Since the summer of 1996, the Census Bureau worked closely with local officials and community-based organizations in each of the three sites to plan and build the various infrastructures needed to ensure a successful dress rehearsal. These joint activities included refining the geographic database, building and refining the address list, and working with community and tribal organizations to plan effective outreach and promotion efforts. Also, the Census Bureau recruited staff in all three sites to complete address list development and verification.

The dress rehearsal allowed for a thorough demonstration of the most critical procedures for Census 2000. These procedures included address list development; marketing and promotion; and data collection, processing, and tabulation. The dress rehearsal plan also demonstrated the use of statistical sampling in four major census operations: nonresponse follow-up, housing units designated as undeliverable as addressed by the U.S. Postal Service, integrated coverage measurement (ICM), and the long form survey.

---

## **15. Data Dissemination Through the Internet**

The census provides a wealth of data that researchers, businesses, and government agencies are eager to use. Taking advantage of modern computer and Internet capabilities, the Census Bureau planned to make data from Census 2000 more readily available than any previous decennial census data. The Census 2000 data are tabulated using the Data Products Production (DPP) system and disseminated using the American FactFinder (AFF) system on the Internet, in addition to CD-ROMs and DVDs. The AFF provides an interactive electronic system to allow data users to access data products, documents, and online help, as well as to build custom data products.

The Census Bureau solicited the advice and recommendations of data users throughout the planning, design, and testing stages of the AFF system (initially known as the Data Access and Dissemination System (DADS)). The system is accessible to the widest possible array of users through the Internet and all available intermediaries, including the nearly 1,800 data centers and affiliates, the 1,400 Federal Depository libraries and other libraries, universities, and private organizations. It also allows users to create customized products, such as tables, charts, graphs, and maps for census geographic areas of their choice, and access metadata that provide documentation and explanatory information for data subjects and geographic areas.

## **16. Evaluation and Preparation for 2010**

After the completion of Census 2000, the Census Bureau plans to conduct a variety of post census evaluation studies, as it has after all the previous censuses. These studies will help data users, both within and outside the Census Bureau, to assess the data and plan for the 2010 Census. The evaluation studies generally rely on demographic analysis, statistical methods, and ethnographic analyses.

## **GLOSSARY**

### **100-Percent Data**

Information based on a limited number of basic population and housing questions collected from both the short form and the long form for every inhabitant and housing unit in the United States.

### **100-Percent Edited Detail File (HEDF)**

Files composed of individual records of information on people and housing units for the 100-percent census data items from the census questionnaires. Estimation is included in these files. These files are used for tabulation purposes and are not released to the public.

### **Accuracy and Coverage Evaluation (A.C.E.)**

The Accuracy and Coverage Evaluation (A.C.E.) is a survey designed to measure the undercount/overcount of the census. The A.C.E. was designed to assess the size and characteristics of the population missed or double-counted in Census 2000, similar to the originally planned Integrated Coverage Measurement (ICM) Survey.

### **Advance Notice Letter/Reminder Card (ANL/RC)**

These are part of the questionnaire mailing strategy. In every area except list/enumerate, the Census Bureau sends an advance notice letter to every mailout address to alert households that the census form will be sent to them soon. Reminder Card is a postcard that is sent to addresses on the decennial Master Address File (see definition below) to remind respondents to return their census questionnaires or to thank them if they already have. All addresses in mailout/mailback areas receive a postcard. The Census Bureau also mails these postcards to postal patrons in update/leave areas.

### **American FactFinder (AFF)**

An electronic system for access and dissemination of Census Bureau data. The system is available through the Internet and offers prepackaged data products and the ability to build custom products. The system serves as the vehicle for accessing and disseminating data from Census 2000 (as well as economic censuses and the American Community Survey). The system was formerly known as the Data Access and Dissemination System (DADS).

---

## **Apportionment**

Apportionment is the process of dividing up the 435 memberships, or seats, in the House of Representatives among the 50 states. The Census Bureau has a dual responsibility in this connection. It conducts the census at 10-year intervals. At the conclusion of each census, the Census Bureau uses the results for calculating the number of House memberships each state is entitled to have. The latter process is the initial use of the basic results of each census.

## **Be Counted Enumeration and Be Counted Form**

The Be Counted enumeration procedure targets areas that are traditionally undercounted. Unaddressed census questionnaires (Be Counted forms) are placed at selected sites where people who believe they were not counted can pick them up, complete them, and mail them to the Census Bureau. The sites are in targeted areas that local governments and community groups, in conjunction with the Census Bureau, identify as traditionally undercounted.

## **Census 2000 Publicity Office (C2PO)**

An office at the Census Bureau which developed, implemented, and coordinated an integrated marketing program for Census 2000, including paid advertising, direct mail, public relations, partnerships, and local outreach.

## **Census Address List Improvement Act of 1994**

See [Program for Address List Supplementation \(PALS\)](#) below.

## **Census Edited File (CEF)**

This file contains the 100-percent edited characteristics/records for all households and people in the census. The edits include consistency edits and imputation for items or persons where the data are insufficient. See descriptions for 100-percent data and census unedited file.

## **Census Information Center (CIC)**

The Census Information Center Program (CIC) is the community-based component of the Census Bureau's data dissemination network. While census data are readily available on CD-ROM, the Census Bureau's Web site on the Internet, in its 12 Regional Offices, 1,400 Federal Depository Libraries, and 1,800 state and local government agencies participating in the State Data Center Program, the CICs provide access to local communities that might not have access through these traditional channels. CIC's goal is to provide efficient access to Census Bureau data and data products to organizations representing populations that have been traditionally undercounted in censuses and surveys.

## **Census Unedited File (CUF)**

A file created by merging the control file for the decennial master address file with the decennial response file of unedited data after the primary selection algorithm has been applied. This file contains the final housing unit and person counts. It is used to generate apportionment data as well as related "raw" or unedited census data.

## **Computer-Assisted Personal Interview (CAPI)**

A method of data collection consisting of the interviewer asking questions displayed on a laptop computer screen and entering the answers directly into the computer.

## **Computer-Assisted Telephone Interviewing (CATI)**

A method of data collection using telephone interviews in which the questions to be asked are displayed on a computer screen and responses are entered directly into the computer.

---

**Confidentiality**

The guarantee made by law (Title 13, United States Code) to individuals who provide census information regarding nondisclosure of that information to others.

**Confidentiality Edit**

The name for the Census 2000 disclosure avoidance procedure.

**Coverage Edit/Coverage Edit Follow-Up (CEFU)**

An edit performed on the mailback census response universe. Census staff make telephone calls to resolve forms that are incomplete or have other coverage discrepancies, such as a difference between the number of people reported in that household and the number of people for whom census information was provided on the form. This edit includes the large household follow-up.

**Coverage Improvement Adjustment**

This phrase was included in the table outlines and the technical documentation before the review, analysis, and recommendation on whether to adjust Census 2000 data for coverage improvement was completed. As the data are not adjusted, a zero (0) will appear. This phrase does not refer to any other outreach or collection operations which were introduced to improve coverage in Census 2000.

**Coverage Improvement Follow-Up (CIFU)**

A procedure for the traditional census in which housing units with conflicting status information are followed up.

**Data Access and Dissemination System (DADS)**

The system is now known as the American FactFinder (AFF).

**Data Capture Center (DCC)**

A decentralized facility that checks in questionnaires returned by mail, creates images of all questionnaire pages, and converts data to computer readable format. The DCCs also perform other computer processing activities, including automated questionnaire edits, work flow management, and data storage. There is one permanent DCC, the National Processing Center in Jeffersonville, Indiana. For Census 2000, the Census Bureau set up three temporary DCCs. The temporary facilities were provided and operated by a private contractor through the Data Capture Services contract.

**Data Capture System 2000 (DCS 2000)**

The DCS 2000 is a data capture system that is used to capture information from census forms. For Census 2000, this system processed more than 150 million incoming forms, digitally captured and processed billions of bits of information on the forms, converted automatically the image of the form to text-based data, and edited/repared data that the system was unable to decipher automatically.

**Decennial Census**

The census of population and housing, taken by the Census Bureau in years ending in 0 (zero). Article I of the Constitution requires that a census be taken every 10 years for the purpose of reapportioning the U.S. House of Representatives.

**Decennial Master Address File (DMAF)**

The decennial version of the Master Address File has features for controlling and tracking the long- and short-term operations and programs of the Census 2000. The DMAF contains the processing status information to support document mailouts; data capture progress control, tracking, and reporting; and field enumeration processes (notably follow-ups). The DMAF is limited to addresses that the Census Bureau has successfully linked to the TIGER® database. See [Master Address File](#).

---

### **Decennial Response File (DRF)**

Contains every response to the census from all sources. The primary selection algorithm is applied to this file to unduplicate people between multiple returns for a housing unit and to determine the housing unit record and the people to include at the housing unit. The DRF is then combined with the Decennial Master Address File to create the census unedited file (CUF).

### **Delivery Sequence File (DSF)**

A computerized file containing all delivery point addresses serviced by the U.S. Postal Service (USPS). The USPS updates the DSF continuously as its letter carriers identify addresses for new delivery points or changes in the status of existing addresses.

Demographic Analysis (DA)

A method the Census Bureau uses to measure coverage at the national level. It differs from survey coverage estimates, such as Post-Enumeration Survey, Integrated Coverage Measurement, or Accuracy and Coverage Evaluation, in that it does not rely on case-by-case matching of census records. To produce an estimate of the total population, DA relies on administrative records to provide estimates of births, deaths, immigration, and emigration. DA provides estimates on the national level only.

### **Derived Measures**

Census data products include various derived measures, such as medians, means, and percentages, as well as certain rates and ratios. Derived measures that round to less than 0.1 are normally indicated as 0.

### **Disclosure Avoidance (DA)**

Statistical methods used in the tabulation of data prior to releasing data products to ensure the confidentiality of responses.

### **Dual-System Estimation (DSE)**

The estimation methodology used for the Accuracy and Coverage Evaluation (A.C.E.). This operation uses a geographic sample of block clusters to find people missed by the census or A.C.E. and any errors from the census. The information is then processed using computer matching, clerical matching, and field follow-up to resolve discrepancies.

### **Family**

A group of two or more people who reside together and who are related by birth, marriage, or adoption.

### **Geocoding**

A code assigned to identify a geographic entity; to assign an address (such as housing unit, business, industry, farm) to the full set of geographic code(s) applicable to the location of that address on the surface of Earth.

### **Group Quarters**

A facility where people live that is not a typical household-type living arrangement. The Census Bureau classifies all individuals not living in households as living in group quarters. There are two types of group quarters institutional (for example, correctional facilities, nursing homes, and mental hospitals) and noninstitutional (for example, college dormitories, military bases and ships, hotels, motels, rooming houses, group homes, missions, shelters, and flophouses).

### **Heterogeneity**

Heterogeneity occurs when blocks of housing units assigned to sampling strata or groupings are not similar in terms of the likelihood of being included or missed by the census. Heterogeneity creates difficulty for the small area estimation process because the correction factor gets applied to all people with the specified characteristic in that sampling poststratum, even though some of them do not actually have the coverage characteristics.

---

## **Homogeneity**

The assumption of homogeneity expects that all people in a particular sampling stratum or grouping will be very much alike in terms of their likelihood of being included or missed by the census. The grouping of people in a particular stratum is called poststratum, such as all White, non-Hispanic male renters ages 18-22 in a rural area. A lack of homogeneity in a particular sample block is not an error, but it does create difficulty for the small area estimation process. This happens because the correction factor gets applied to all people with the specified characteristic in that poststratum, even though some of them do not exhibit the same coverage characteristics.

## **Household**

Household refers to all of the people who occupy a housing unit.

## **Housing Unit**

A housing unit is a house, an apartment, a mobile home or trailer, a group of rooms, or a single room occupied as a separate living quarters, or if vacant, intended for occupancy as a separate living quarters. Separate living quarters are those in which the occupants live separately from any other individuals in the building and which have direct access from outside the building or through a common hall. For vacant units, the criteria of separateness and direct access are applied to the intended occupants whenever possible.

## **Imputation**

When information is missing or inconsistent, the Census Bureau uses a method called imputation to assign values. Imputation relies on the statistical principle of “homogeneity,” or the tendency of households within a small geographic area to be similar in most characteristics. For example, the value of “rented” is likely to be imputed for a housing unit not reporting on owner/renter status in a neighborhood with multiunits or apartments where other respondents reported “rented” on the census questionnaire. In past censuses, when the occupancy status or the number of residents was not known for a housing unit, this information was imputed.

**Internet Questionnaire Assistance (IQA)** An operation which allows respondents to use the Census Bureau’s Internet site to (1) ask questions and receive answers about the census form, job opportunities, or the purpose of the census and (2) provide responses to the short form.

**Interpolation** Interpolation frequently is used in calculating medians or quartiles based on interval data and in approximating standard errors from tables. Linear interpolation is used to estimate values of a function between two known values. Pareto interpolation is an alternative to linear interpolation. In Pareto interpolation, the median is derived by interpolating between the logarithms of the upper and lower income limits of the median category. It is used by the Census Bureau in calculating median income within intervals wider than \$2,500.

## **List/Enumerate**

A method of data collection in which temporary field staff, called enumerators, list each residential address, spot the location of each on a census map, and interview the residents of the household during a single visit. This completes the census address list for these areas and provides the information needed to update the TIGER® database and Master Address File (see definitions below).

## **Local Update of Census Addresses (LUCA)**

A Census 2000 program, established in response to requirements of P. L. 103-430. It provided an opportunity for state, local, and tribal governments to review and update individual address information in the Master Address File and associated geographic information in the TIGER® database before using the addresses for questionnaire delivery. This improved the completeness and accuracy of both computer files and the census.

---

## **Long Form**

The decennial census questionnaire, sent to approximately one in six households, contains all questions on the short form, as well as additional detailed questions relating to the social, economic, and housing characteristics of each individual and household. Information derived from the long form is referred to as sample data and is tabulated for geographic entities as small as the block group level.

## **Mailout/Mailback (MO/MB)**

A method of data collection in which the U.S. Postal Service delivers addressed questionnaires to residents who are asked to complete and mail back the questionnaire to the appropriate Census Bureau office. This method is used for more than 80 percent of all households (usually with city-style addresses).

## **Master Address File (MAF)**

A computer file based on a combination of the addresses in the 1990 census address file and current versions, supplemented by address information provided by state, local, and tribal governments. The MAF is continually updated to provide a basis for creating the Census 2000 address list, the address list for the American Community Survey, and the address list for the Census Bureau's other demographic surveys.

## **Metadata**

Information about the content, quality, condition, and other characteristics of data.

## **Microdata**

Nonaggregated data about the units sampled. For surveys of individuals, microdata contain records for each individual interviewed; for surveys of organizations, the microdata contain records for each organization.

## **Nongovernment Organization**

The partnerships developed during Census 2000 planning include national and local organizations and community groups that are not governmental entities.

## **Nonresponse Follow-up**

A census follow-up operation in which temporary field staff, known as enumerators, visit addresses from which no response was received.

## **Nonsampling Error**

Errors that occur during the measuring or data collection process. Nonsampling errors can be the most serious types of errors because they yield biased results when most of the errors distort the results in the same direction. Unfortunately, the full extent of nonsampling error is unknown. Decennial censuses traditionally have experienced nonsampling errors, most notably undercount, resulting from people being missed in the enumeration processes.

## **Optical Character Recognition (OCR)**

Technology that uses an optical scanner and computer software to "read" human handwriting.

## **Optical Mark Recognition (OMR)**

Technology that uses an optical scanner and computer software to scan a page, recognize the presence of marks in predesignated areas, and assign a value to the mark depending on its specific location and intensity on a page.

---

## **Poststratum**

Information about the current occupants of each housing unit in the Accuracy and Coverage Evaluation (A.C.E.) survey found during the A.C.E. interview is used to form groupings called “poststrata.” This information, including the age of respondent, current owner/renter status, etc., is used to form homogeneous groupings and improve the estimation process. By contrast, the initial A.C.E. strata are formed using aggregate information about each block as of the 1990 census.

## **Primary Selection Algorithm (PSA)**

Computer program applied to the decennial response file (DRF) to eliminate duplicate responses and to determine the housing unit record and the people to include at the housing unit. After this procedure, the DRF is merged with the Decennial Master Address File to create the census unedited file.

## **Program for Address List Supplementation (PALS)**

A program providing all governmental units and regional and metropolitan agencies the opportunity to submit lists of individual addresses for their community to the Census Bureau for use in building the MAF. Ongoing submissions and feedback between the Census Bureau and local governments on this program, enabled by the Census Address List Improvement Act of 1994 (P.L. 103-430) help ensure the completeness and accuracy of the Master Address File and the TIGER® database.

## **Public Law (P.L.) 94-171**

Public Law (P.L.) 94-171, enacted in 1975, directs the Census Bureau to make special preparations to provide redistricting data needed by the 50 states. Within a year following Census Day, the Census Bureau must send the data agreed upon to redraw districts for the state legislature to each state’s governor and majority and minority legislative leaders.

To meet this legal requirement, the Census Bureau set up a voluntary program that enables participating states to receive data for voting districts (e.g., election precincts, wards, state house, and senate districts) in addition to standard census geographic areas, such as counties, cities, census tracts, and blocks.

## **Public Law (P.L.) 103-430**

Public Law (P.L.) 103-430, enacted in 1994, amends Title 13, United States Code, to allow designated local and tribal officials access to the address information in the Master Address File to verify its accuracy and completeness. This law also requires the U.S. Postal Service to provide its address information to the Census Bureau to improve the Master Address File.

## **Public Law (P.L.) 105-119**

Public Law (P.L.) 105-119, enacted in 1997, directs the Census Bureau to make publicly available a second version of Census 2000 data that does not include the corrections for overcounts and undercounts measured in the Accuracy and Coverage Evaluation (A.C.E.). The format, timing, geographic levels, and price of the P.L. 94-171 and these data are identical.

## **Public Use Microdata Area (PUMA)**

An area that defines the extent of territory for which the Census Bureau tabulates public use microdata sample (PUMS) data.

## **Public Use Microdata Sample (PUMS)**

Hierarchical files containing small samples (5% and 1%) of individual records from the census long form showing characteristics of the housing units and people included on those forms.

---

## **Quality Assurance (QA)**

Quality assurance represents a broad philosophy and specific procedures that are designed to build quality into the system, constantly improve the system, and integrate responsibility for quality with production.

## **Questionnaire Mailing Strategy**

For Census 2000, an advance notice letter, a questionnaire, and a reminder/thank you postcard were sent to every mailout address.

## **Reapportionment**

The redistribution of seats in the U.S. House of Representatives among several states on the basis of the most recent decennial census as required by Article 1, Section 2 of the Constitution. See apportionment and redistricting.

## **Redistricting**

The process of revising the geographic boundaries of areas from which people elect representatives to the U.S. Congress, a state legislature, a county or city council, a school board, and the like to meet the legal requirement that such areas be as equal in population as possible following a census. See apportionment and reapportionment.

## **Sample Census Edited File (SCEF)**

A file containing 100-percent and sample characteristics for housing units and people in the long form sample. Processing for the SCEF includes merging the results of industry and occupation coding and place of work and migration coding, coding several other items, and weighting the long forms.

## **Sample Edited Detail File (SEDF)**

A file containing 100-percent and sample characteristics for housing units and people in the long form sample. The file is used for tabulation purposes only and is not released to the public.

## **Sampling Error**

Errors that occur because only a part of the population is being contacted directly. With any sample, differences are likely to exist between the characteristics of the sampled population and the larger group from which the sample was chosen. However, sampling error, unlike nonsampling error, is readily measured.

## **Sampling Stratum**

A sampling stratum, as used in the A.C.E., is a grouping or classification that has a similar set of characteristics, based on the 1990 census. For example, one might define a stratum as all blocks in large central cities with a 1990 census population that was 30 percent or more Black renters.

## **Scanner**

Equipment used to capture images from documents for the purpose of entering the information into an electronic format. For Census 2000, scanners replaced some keying operations.

## **Seasonal/Recreational/Occasional Use**

A housing unit held for occupancy only during limited portions of the year, such as a beach cottage, ski cabin, or time-share condominium.

---

## **Separate Living Quarters**

Those living quarters in which the occupants live separately from any other individual in the building and which have direct access from outside the building or through a common hall. For vacant units, the criteria of separateness and direct access are applied to the intended occupants whenever possible.

## **Service-Based Enumeration (SBE)**

An operation designed to enumerate people at facilities where they might receive services, such as shelters, soup kitchens, healthcare facilities, and other selected locations. This operation targets the types of services that primarily serve people who have no usual residence.

## **Service Locations**

Locations where clients are enumerated during the service-based enumeration operation, such as emergency or transitional shelters, soup kitchens, regularly scheduled mobile food vans, and targeted nonsheltered outdoor locations.

## **Short Form**

The decennial census questionnaire, sent to approximately 5 of 6 households, that contains population questions related to household relationship, age, sex, relationship, race, Hispanic origin, and tenure (i.e., whether home is owned or rented). The questions contained on the short form also are asked, along with additional questions, on the long form.

## **Simplified Enumerator Questionnaire (SEQ)**

A questionnaire that enumerators use for transient, or T-night, enumeration and when conducting the nonresponse follow-up after the decennial census.

## **Soup Kitchens**

Includes soup kitchens, food lines, and programs distributing prepared breakfasts, lunches, or dinners. These programs may be organized as food service lines, bag or box lunches, or tables where people are seated, then served by program personnel. These programs may or may not have a place for clients to sit and eat the meal. These are service locations.

## **Special Place**

An institution that includes facilities where people live or stay other than the usual house, apartment, or mobile home. Examples are colleges and universities, nursing homes, hospitals, and prisons. Often the facilities that house people are group quarters, but they may include standard houses or apartments as well.

## **Special Place Facility Questionnaire (SPFQ)**

A questionnaire used to interview an official at a special place for the purpose of collecting/updating address information for the special place and any associated group quarters and housing units, determining the type of special place/group quarters, and collecting additional administrative information about each group quarters at the special place.

## **State Data Center (SDC)**

A state agency or university facility identified by the governor of each state and state equivalent to participate in the Census Bureau's cooperative network for the dissemination of census data. SDCs also provide demographic data to local agencies participating in the Census Bureau's statistical areas programs and assist the Census Bureau in the delineation and identification of statistical areas.

---

### **Summary File (SF)**

A series of census summary tabulations of 100-percent and sample population and housing data available for public use on CD-ROM and the Internet. In 1990, these files were available on computer tapes and, as a result, were known as summary tape files (STF).

### **Summary Table**

A collection of one or more data elements that are classified into some logical structure either as dimensions or data points.

### **Tabulation Block**

A physical block that does not have any legal or statistical boundaries passing through it; or each portion of a physical block after the Census Bureau recognizes any legal or statistical boundaries that pass through it.

### **Targeted Nonsheltered Outdoor Location (TNSOL)**

A geographically identifiable outdoor location open to the elements where there is evidence that people might be living without paying and who also do not usually receive services at soup kitchens, shelters, and mobile food vans. These sites must have a specific location description that allows a census enumeration team to physically locate the site and excludes pay-for-use campgrounds, drop-in centers, post offices, hospital emergency rooms, and commercial sites (including all-night theaters and all-night diners).

### **Telephone Questionnaire Assistance (TQA)**

A toll-free service that was provided by a commercial phone center to answer questions about Census 2000 and the Census 2000 questionnaire and to take interviews from people who prefer to be interviewed over the telephone.

### **Thematic Map**

A map that reveals the geographic patterns in statistical data.

### **Title 13 (United States Code)**

The law under which the Census Bureau operates and that guarantees the confidentiality of census information and establishes penalties for disclosing this information.

### **Topologically Integrated Geographic Encoding and Referencing (TIGER®)**

A computer database that contains a digital representation of all census-required map features (streets, roads, rivers, railroads, lakes, and so forth), the related attributes for each (street names, address ranges, etc.), and the geographic identification codes for all entities used by the Census Bureau to tabulate data for the United States, Puerto Rico, and the Island Areas. The TIGER® database records the interrelationships among these features, attributes, and geographic codes and provides a resource for the production of maps, entity headers for data tabulations, and automated assignment of addresses to a geographic location in a process known as “geocoding.”

### **Transient Night (T-Night)/T-Night Enumeration (TNE)**

A method of enumeration in which Census Bureau staff enumerate people at transient locations, such as campgrounds at race tracks, recreational vehicle campgrounds or parks, commercial or public campgrounds, fairs and carnivals, and marinas. Enumerators conduct a personal interview using Simplified Enumerator Questionnaire. No vacant units are generated by this operation.

### **Type of Enumeration Area (TEA)**

A classification identifying how the Census Bureau takes the decennial census of a geographic area. Examples of TEAs include (1) the area inside the “blue line” - this is the mailout/mailback and urban update/leave operations area, (2) address listing areas, (3) list/enumerate areas, and (4) remote areas of Alaska.

---

**Urban Update/Leave (UU/L)**

Update/leave procedures are used in targeted urban areas where mail delivery may be a problem, such as an apartment building where the mail carrier may leave the forms in a common area. Enumerators deliver census questionnaires for residents to complete and mail back, update the address register, and update the census maps.

**Usual Home Elsewhere (UHE)**

A housing unit that is temporarily occupied by a person(s) who has a usual home elsewhere.

**Usual Residence**

The living quarters where a person spends more nights during a year than any other place.

**Voting District (VTD)**

Any of a variety of areas, such as election districts, precincts, legislative districts, or wards, established by states and local governments for voting purposes.

**Whole Household Usual Home Elsewhere (WHUHE)**

See Usual Home Elsewhere.