Glossary of Terms

100 Percent Data. A term used in 2000 to describe the data that were asked of "100 percent" of the population in Census 2000. That is, questions that were collected for all people on both the census short-form and long-form questionnaires. In 2000, this included sex, relationship, age/date of birth, Hispanic origin, race, and tenure.

Accessibility. One of four key dimensions of survey quality, accessibility refers to the ability of the data users to readily obtain and use survey products.

Acceptability Index. The average number of basic ACS items reported per person, including sex, age (counted double), relationship, marital status, Hispanic origin, and race. A questionnaire for an occupied unit must have an acceptability index of 2.5 or greater to be considered an interview.

Accuracy. One of four key dimensions of survey quality. Accuracy refers to the difference between the survey estimate and the true (unknown) value. Attributes are measured in terms of sources of error (for example, coverage, sampling, nonresponse, measurement, and processing).

Address Control File. The residential address list used in the 1990 census to label questionnaires, control the mail response check-in operation, and determine the nonresponse follow-up workload.

Address Corrections from Rural Directories. A post-Census 2000 Master Address File (MAF) improvement operation where Census Bureau staff reviewed commercial directories for 300 rural counties in 10 Midwestern states to obtain new city-style addresses for MAF records that did not contain a city-style address. Conducted in 2002, over 15,000 city-style addresses were associated with MAF records that previously lacked a city-style address.

Address Listing. A Census 2000 field operation to develop the address list in areas with predominantly non-city-style mailing addresses. A lister captured the address and/or a physical/location description for each living quarters within a specified assignment area. The lister marked the location of each residential structure on a block map by placing a spot on the map indicating its location and assigning a map spot number. The lister also updated and corrected features on the map if necessary. This activity was called "prelist" in the 1990 census.

Administrative Entities. Geographic areas, usually with legally defined boundaries but often without elected officials, created to administer elections and other governmental functions. Administrative areas include school districts, voting districts, ZIP codes, and nonfunctioning Minor Civil Divisions (MCDs) such as election precincts, election districts, and assessment districts.

Allocation. Imputation method required when values for missing or inconsistent items cannot be derived from the existing response record. In these cases, the imputation must be based on other techniques such as using answers from other people in the household, other responding housing units, or people believed to have similar characteristics. Such donors are reflected in a table referred to as an allocation matrix.

American Community Survey (ACS) Alert. This periodic electronic newsletter informs data users and other interested parties about news, events, data releases, congressional actions, and other developments associated with the ACS.

American Community Survey Demonstration Program. The full set of testing, research, and development program activities that started in 1994 and continued until the ACS was fully implemented in 2005.

American Community Survey Full Implementation. The period beginning in January 2005 during which the ACS interviewing of its housing unit sample was conducted in every county and Puerto Rico *municipio* as well as all American Indian and Alaska Native Areas and Hawaiian Homelands. The full implementation initial sample size is approximately 3 million addresses each year, and includes group quarters (GQ) facilities which were added beginning in January 2006.

American Community Survey Test Sites. The ACS demonstration program expanded from an initial four test counties in 1996 to 36 test counties in 1999. When the term ACS test site is used, it refers to data from these 36 counties.

American FactFinder (AFF). An electronic system for access and dissemination of Census Bureau data on the Internet. The system offers prepackaged data products and user-selected data tables and maps from Census 2000, the 1990 Census of Population and Housing, the 1997 and 2002 Economic Censuses, the Population Estimates Program, annual economic surveys, and the ACS.

American Indian Area, Alaska Native Area, Hawaiian Homeland (AIANAHH). A Census Bureau term referring to the following types of areas: federal and state American Indian reservations, American Indian off-reservation trust land areas (individual or tribal), Oklahoma tribal statistical areas (in 1990 tribal jurisdictional statistical area), tribal designated statistical areas, state designated American Indian statistical areas, Alaska Native Regional Corporations, Alaska Native village statistical areas, and Hawaiian Homelands.

Assignment. Imputation method in which values for a missing or inconsistent item can be derived from other responses from the sample housing unit or person. For example, a first name can be used to determine and assign the sex of a person.

Automated Address Unduplication. An ongoing MAF improvement activity completed twice a year (coinciding with the delivery sequence file (DSF) refresh of the MAF) where, through automated means, pairs of city-style addresses are identified as identical based on house number, street name, five-digit ZIP code, and within structure identifier (if one exists). These addresses are linked for future operations to control duplication.

Automated Clerical Review. The ACS program run on raw mail return data to determine whether or not a case goes to failed-edit follow-up. The name reflects the fact that it was originally done clerically. The operator checks for missing content and for large households (more than five members) and for coverage inconsistencies.

Automated Editing. Editing that is accomplished using software, as opposed to being done clerically.

Automated Listing and Mapping Instrument (ALMI). Software used primarily by Census Bureau field representatives for the purpose of locating an address or conducting an address listing operation. The ALMI combines data from the MAF and the Topologically Integrated Geographic Encoding and Referencing (TIGER®) System database to provide users with electronic maps and associated addresses. ALMI functionality allows users to edit, add, delete, and verify addresses, streets, and other map features, view a list of addresses associated with a selected level of geography, and view and denote the location of housing units on the electronic map.

Automated Review Tool (ART). A Web-based computer application designed to help subject matter analysts quickly review and approve ACS estimates.

Automated Review Tool II (ART II). The next generation of the ART. It is aimed at providing analysts with reports at a more detailed level than the previous version.

Base Tables. Tables that provide the most detailed estimates on all topics and geographic areas from the ACS. Base tables also include totals and subtotals. These tables form the data source for the "Derived Products." Base tables are also known as detailed tables.

Base Weight. The base weight for an address is equal to the inverse of the probability with which the address was selected for the sample as determined by the sample design. Since these weights are based only on the initial probability of selection, they are known as *a priori* to the data collection phase. This is the weight for a housing unit before any adjustments are made. The base weight is also known as the unbiased weight.

Be Counted Enumeration and Be Counted Questionnaire. The Be Counted program provided a means for people who believed they were not counted to be included in Census 2000. The Census Bureau placed Be Counted questionnaires at selected sites that were easily accessible to and frequented by large numbers of people. The questionnaires also were distributed by the Questionnaire Assistance Centers and in response to requests received through Telephone Questionnaire Assistance.

Blaise. An authoring application that produces an instrument used to collect data using computer-assisted telephone interviewing (CATI) or computer-assisted personal interviewing (CAPI).

Block. A subdivision of a census tract (or, prior to 2000, a block numbering area), a block is the smallest geographic entity for which the Census Bureau tabulates decennial census data. Many blocks correspond to individual city blocks bounded by streets, but blocks—especially in rural areas—may include many square miles and may have some boundaries that are not streets. The Census Bureau established blocks covering the entire nation for the first time in 1990. Previous censuses back to 1940 had blocks established only for part of the nation. Over 8 million blocks were identified for Census 2000.

Block Canvassing. A Census 2000 field operation to ensure the currency and completeness of the MAF within the mailout/mailback area. Listers traveled in their assignment areas to collect and verify information to ensure that their address listing pages (derived from the MAF) contained a mailing address for every living quarters. They especially looked for hidden housing units (such as attics, basements, or garages converted into housing units) and houses that appeared to be one unit but actually contained multiple housing units. They also updated and corrected their Census Bureau maps.

Block Group. A subdivision of a census tract (or, prior to 2000, a block numbering area), a block group is a cluster of blocks having the same first digit of their four-digit identifying number within a census tract.

Boundary and Annexation Survey (BAS). An annual survey of all counties and statistically equivalent entities, all or selected incorporated places and minor civil divisions, all or selected federally recognized American Indian reservations and off-reservation trust lands, and Alaska Native Regional Corporations, to determine the location of legal limits and related information as of January 1 of the survey year.

Case Management. A tool used by field representatives that allows them to manage their interview assignments on their laptops.

Census 2000 Supplementary Survey (C2SS). The C2SS was an operational test conducted as part of the research program in Census 2000, and used the ACS questionnaire and methods to collect demographic, social, economic, and housing data from a national sample. This evaluation study gave the Census Bureau essential information about the operational feasibility of converting from the census long-form sample to the ACS.

Census County Division (CCD). A subdivision of a county that is a relatively permanent statistical area established cooperatively by the Census Bureau and state and local government authorities. Used for presenting decennial census statistics in those states that do not have well-defined and stable minor civil divisions that serve as local governments.

Census Designated Place (CDP). A statistical entity that serves as a statistical counterpart of an incorporated place for the purpose of presenting census data for a concentration of population, housing, and commercial structures that is identifiable by name, but is not within an incorporated place. CDPs usually are delineated cooperatively with state, Puerto Rico, Island Area, local, and tribal government officials, based on the Census Bureau guidelines. For Census 2000, CDPs did not have to meet a population threshold to quality for the tabulation of census data.

Census Geography. A collective term referring to the types of geographic areas used by the Census Bureau in its data collection and tabulation operations, including their structure, designations, and relationships to one another.

Census Information Center (CIC). The CIC program is a cooperative activity between the Census Bureau and the national nonprofit organizations representing interests of racial and ethnic communities. The program objective is to make census information and data available to the participating organizations for analysis, policy planning, and for further dissemination through a network of regional and local affiliates.

Census Sample Data. Population and housing information collected only on the census long form for a sample of households.

Census Tract. A small, relatively permanent statistical subdivision of a county delineated by a local committee of census data users for the purpose of presenting data. Census tract boundaries normally follow visible features, but may follow governmental unit boundaries and other nonvisible features; they always nest within counties. Designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions at the time of establishment, census tracts average about 4,000 inhabitants.

City-Style Address. An address that consists of a house number and street or road name; for example, 201 Main Street. The address may or may not be used for the delivery of mail, and may include apartment numbers/designations or similar identifiers.

Coding. The process of associating numeric codes with write-in strings. For example, the write-in associated with Place of Birth is turned into a three-digit code.

Coefficient of Variation (CV). The ratio of the standard error (square root of the variance) to the value being estimated, usually expressed in terms of a percentage (also known as the relative standard deviation). The lower the CV, the higher the relative reliability of the estimate.

Cold Deck Values. The values used to initialize matrices used for hot-deck allocation.

Collapsing. Reducing the amount of detail shown in a base table to comply with data release rules.

Community Address Updating System (CAUS). A post-Census 2000 MAF improvement program that provides a systematic methodology for enhancement and update of address and feature information. Designed to provide a rural counterpart to the update of the city-style addresses received from the U.S. Postal Service's Delivery Sequence File, CAUS identifies and conducts listing operations in selected geographic areas suspected of experiencing growth that is either not available from or appears to be incomplete in the U.S. Postal Service's Delivery Sequence File. Address and feature updates collected for CAUS are added to the MAF and the TIGER® System.

Comparison Profile. Comparison profiles are available from the ACS for 1-year estimates beginning in 2007. These tables are available for the United States, the 50 states, the District of Columbia, and geographic areas with a population of more than 65,000.

Complete Interview. The ACS interview is classified as complete when all applicable questions have been answered on the mail form, or during a CATI or CAPI interview. The interview may include responses of "Don't Know" and "Refused" to specific questions.

Computer-Assisted Personal Interviewing (CAPI). A method of data collection in which the interviewer asks questions displayed on a laptop computer screen and enters the answers directly into a 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 a computer.

Confidence Interval. The sample estimate and its standard error permit the construction of a confidence interval that represents the degree of uncertainty about the estimate. Each ACS estimate is accompanied by the upper and lower bounds of the 90 percent confidence interval, or the 90 percent margin of error, from which a confidence interval can be constructed. A 90 percent confidence interval can be interpreted roughly as providing 90 percent certainty that the interval defined by the upper and lower bounds contains the true value of the characteristic.

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

Congressional Tool Kit. A collection of documents developed for members of Congress that explain how and why the ACS is conducted, its benefits, and how to obtain additional information. The Tool Kit originally was distributed as hard copies in 3-ring binders and is now available as a series of online portable document format (PDF) files.

Consumer Price Index (CPI). The CPI program of the Bureau of Labor Statistics produces monthly data on changes in the prices paid by urban consumers for a representative basket of goods and services.

Control File. A file which represents the current status of any case in sample in the ACS.

Controlled. During the ACS weighting process, the intercensal population and housing estimates are used as survey controls. Weights are adjusted so that ACS estimates conform to these controls.

Count Question Resolution (CQR). A process followed in Census 2000 whereby state, local, and tribal government officials could ask the Census Bureau to verify the accuracy of the legal boundaries used for Census 2000, the allocation of living quarters and their residents in relation to those boundaries, and the count of people recorded by the Census Bureau for specific living quarters.

Cross Tabulation. The joint distribution of two or more data characteristics, where each of the categories of one characteristic is repeated for each of the categories of the other characteristic(s). A cross-tabulation in a base table is denoted where "BY" is used as the conjunction between characteristics; for example, "AGE BY SEX" or "AGE BY SEX BY RACE."

Current Population Survey (CPS). Monthly sample survey of the U.S. population that provides employment and unemployment estimates as well as current data about other social and economic characteristics of the population. Collected for the Bureau of Labor Statistics by the Census Bureau.

Current Residence. The concept used in the ACS to determine who should be considered a resident of a sample address. Everyone who is currently living or staying at a sample address is considered a resident of that address, except people staying there for two months or less. People who have established residence at the sample address and are away for only a short period of time are also considered to be current residents.

Custom Tabulations. The Census Bureau offers a wide variety of general purpose data products from the ACS. These products are designed to meet the needs of the majority of data users and contain predefined sets of data for standard census geograpic areas, including both political and statistical geography. These products are available on the American FactFinder and the ACS Web site.

For users with data needs not met through the general purpose products, the Census Bureau offers "custom" tabulations on a cost-reimbursable basis, with the ACS Custom Tabulation program. Custom tabulations are created by tabulating data from ACS microdata files. They vary in size, complexity, and cost depending on the needs of the sponsoring client.

Data Capture File. The repository for all data captured from mail return forms and by CATI and CAPI Blaise instruments.

Data Collection Mode. One of three ACS methods (mail, telephone, personal visit) of data collection.

Data Profiles. Data products containing estimates of key demographic, social, economic, and housing characteristics. Data swapping is done by editing the source data or exchanging records for a sample of cases. A sample of households is selected and matched on a set of selected key variables with households in neighboring geographic areas that have similar characteristics. Because the swap often occurs within a neighboring area, there is usually no effect on the marginal totals for the area or for totals that include data from multiple areas.

De Facto Residence Rules. De facto means "in fact." A de facto residence rule would define survey residents as all people living or staying at the sample address at the time of the interview without considering other factors such as the amount of time they will be staying there. Such a rule would exclude people away from a regular residence even if they were away for only that one day. The ACS is using a de facto residence rule when determining the residents of GQ facilities eligible to be sampled and interviewed for the survey.

Delivery Sequence File (DSF). A U.S. Postal Service (USPS) computer file containing all mailing addresses serviced by the USPS. The USPS continuously updates the DSF as its letter carriers identify addresses for new delivery points and changes in the status of existing addresses. The Census Bureau uses the DSF as a source for maintaining and updating its MAF.

Demographic Area Address Listing (DAAL). A post-Census 2000 program associated with coverage improvement operations, address list development, and automated listing for the CAUS and demographic household surveys. The program uses automated listing methods to update the inventory of living quarters, and also updates the street network in selected blocks.

Derived Products. Derived products are informational products based largely on estimates from the base tables.

Detailed Tables. See **Base Tables**.

Disclosure Avoidance (DA). Statistical methods used in the tabulation of data prior to releasing data products to ensure the confidentiality of responses. See **Confidentiality**.

Disclosure Review Board (DRB). A board comprised of Census Bureau staff who review and must approve all data products based on disclosure avoidance rules before they can be released to the public.

Edit. To subject data to program logic to check for missing data and inconsistencies.

Edit Management and Messaging Application (EMMA). An Internet application used by ACS subject-matter analysts to show the status of edit review and to relay analyst's relevant comments.

Estimates. Numerical values obtained from a statistical sample and assigned to a population parameter. Data produced from the ACS interviews are collected from samples of housing units. These data are used to produce estimates of the actual figures that would have been obtained by interviewing the entire population using the same methodology.

Evaluation Studies. Research and evaluation conducted by Census Bureau staff and external experts to assess a broad set of topics including the feasibility and the quality of the data products produced by the ACS.

Failed Edit Follow-Up (FEFU). Data collection activity of mail response records designed to collect missing information. Mail returns failing the automated clerical review edit are contacted by telephone.

Federal Agency Information Program (FAIP). A long-term program of information and technical partnership with federal agencies. The FAIP is designed to establish a relationship with each agency that will identify the unique opportunities and challenges it faces in using ACS data. The program targets assistance based on the needs and resources of each federal agency in order to help the agency make a smooth transition to using ACS data.

Federal Government Unit (FGU). Any of a variety of civil divisions; places and is used for sampling.

Federal Register Notice. Published by the Office of the Federal Register, National Archives and Records Administration (NARA), the *Federal Register* is the official daily publication for rules, proposed rules, and notices of federal agencies and organizations, as well as executive orders and

other presidential documents. Information describing proposed data collection must be posted on the *Federal Register* for public review and comment for a 30-day period and must take place before the Office of Management and Budget (OMB) can provide final clearance for the data collection.

Federal-State Cooperative Program for Population Estimates (FSCPE). FSCPEs are state-level organizations, designated by their respective governors, to work cooperatively with the Census Bureau's Population Estimates Program in the production of subnational population estimates and in making data broadly available to the public.

Field Representative (FR). A Census Bureau employee who interviews people to obtain information for a census or survey.

File Transfer Protocol (FTP). A process that allows a user to download large files and datasets from American FactFinder.

Final Outcome Code. A code assigned to a CATI or CAPI case at the conclusion of the data collection which characterizes the status of the case, such as "completed occupied interview" or "respondent refusal noninterview."

First Stage Sample. ACS first stage sampling maintains five 20 percent partitions of the MAF by determining which addresses were in the first stage sample 4 years prior and excluding them. This ensures that no address is in sample more than once in any 5-year period. The first phase sample is the universe from which the second phase sample is selected.

Five-Year Estimates. Estimates based on 5 years of ACS data. These estimates are meant to reflect the characteristics of a geographic area over the entire 60-month period and will be published for all geographic areas down to the census block group level.

Functioning Governmental Unit (FGU). A general purpose government that has the legal capacity to elect or appoint officials, raise revenues, provide surveys, and enter into contracts.

General Coding. The process whereby write-in answers to Hispanic origin, race, ancestry, and language are categorized into codes. This is accomplished using an automated system approach, relying on a set of growing dictionaries of write-ins against which responses are computer matched. Responses that are not found in the dictionaries are sent to subject matter experts who code them. These new responses are added to the computer dictionaries for subsequent use.

Geocoding. The assignment of an address, structure, key geographic location, or business name to a location that is identified by one or more geographic codes. For living quarters, geocoding usually requires identification of a specific census block.

Geographic Summary Level. A geographic summary level specifies the content and the hierarchical relationships of the geographic elements that are required to tabulate and summarize data. For example, the county summary level specifies the state-county hierarchy. Thus, both the state code and the county code are required to uniquely identify a county in the United States or Puerto Rico.

Government Printing Office (GPO). A federal agency responsible for producing, procuring, and disseminating printed and electronic publications of the Congress as well as the executive departments and establishments of the federal government.

Governmental Unit Measure of Size (GUMOS). The smallest measure of size associated with a given block. It is used in the sample selection operation to determine the initial sampling rate at the block level.

Group Quarters (GQ) Facilities. A GQ facility is a place where people live or stay that is normally owned or managed by an entity or organization providing housing and/or services for the residents. These services may include custodial or medical care, as well as other types of assistance. Residency is commonly restricted to those receiving these services. People living in GQ facilities are usually not related to each other. The ACS collects data from people living in both housing units and GQ facilities.

Group Quarters Facilities Questionnaire (GQFQ). A Blaise-based automated survey instrument that field representatives (FRs) use to collect new or updated information about a GQ facility. Questions in this survey include facility name, mailing and physical address, telephone number, GQ contact name and telephone number, special place name, and the GQ facility's maximum occupancy and current number of people staying in the GQ facility.

Group Quarters Geocoding Correction Operation. A post-Census 2000 MAF improvement operation implemented to correct errors (mostly census block geocodes) associated with college dormitories in MAF and TIGER®. Conducted by Census Bureau staff, source materials for over 20,000 dormitories were reviewed and used to identify and correct MAF/TIGER® errors.

Group Quarters Listing Sheet. This form is preprinted with information such as GQ name and control number for sample GQ facilities. It is used by FRs when the GQ administrator is unable to provide a list of names or occupied bed locations for person-level sample selection.

Group Quarters Measure of Size (GQMOS). The expected population of a given GQ facility divided by 10. It is used in the sample selection operation to determine the universe of sample units to be sampled. A sample unit is a cluster or group of 10 people.

Hot Deck Imputation. An approach for filling in missing answers with information from like households or persons, with donors determined by geographic location or specific characteristics reported. Hot deck imputation continually updates matrices with data from donors with acceptable data and then provides values from such matrices to recipients who need data.

Household. A household includes all the people who occupy a housing unit that meet all the residence rules of a survey or census.

Housing Unit (HU). A house, apartment, mobile home or trailer, a group of rooms, or a single room occupied as separate living quarters, or if vacant, intended for occupancy as separate living quarters. Separate living quarters are those in which the occupants live separately from any other individuals in the building and 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 imputation methods to assign or allocate 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.

Interactive Voice Recognition (IVR). An automated telephone application which allows the caller to hear prerecorded responses to frequently asked questions. The caller may proceed through the application by entering numbers from the telephone key pad or by speaking responses to select which messages he/she wants to hear. The caller may also elect to speak to an interviewer instead of listening to the recorded responses.

Intercensal Estimates. Official Census Bureau estimates of the population of the United States, states, metropolitan areas, cities and towns, and counties; also official Census Bureau estimates of housing units (HUs).

Interim Codes. These are codes assigned to a sample GQ assignment in the GQFQ system by a field representative when scheduling a personal visit to a sample ACS GQ facility, when additional research is needed to locate the GQ facility, or when a return visit to the GQ facility is needed to obtain additional survey information.

Interpolation. Interpolation is frequently 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.

Interview Monitoring. A process in which CATI supervisors, for quality control purposes, listen to interviewers while they are conducting interviews with respondents to assure that the interviewer is following all interviewing procedures correctly. The interviewer is not told when the supervisor is listening, but is given feedback on his/her performance after the monitoring.

ACS Design and Methodology

Item Nonresponse. The failure to obtain valid responses or responses consistent with other answers for individual data items.

Iterations. Subgroups of the original tabulation universe, especially by race, Hispanic origin, ancestry, and tribal groups. For example, many ACS base tables are iterated by 9 race and Hispanic origin groups.

Joint Economic Edit. An edit which looks at the combination of multiple variables related to a person's employment and income, thereby maximizing the information used for filling any missing related variables.

Key-From-Image (KFI). An operation in which keyers use a software program to capture questionnaire responses by typing responses directly into the scanned image of a questionnaire displayed on their work station screen.

Key-From-Paper (KFP). An operation in which keyers use a software program to capture questionnaire responses from a hard-copy of the questionnaire.

Legal Entity. A geographic entity whose origin, boundary, name, and description result from charters, laws, treaties, or other administrative or governmental action, such as the United States, states, the District of Columbia, Puerto Rico, the Island Areas, counties, cities, boroughs, towns, villages, townships, American Indian reservations, Alaska Native villages, congressional districts, and school districts. The legal entities and their boundaries that the Census Bureau recognizes are those in existence on January 1 of each calendar year.

List/Enumerate. A method of decennial census data collection in some of the more remote, sparsely populated areas of the United States and the Island Areas, where many of the households do not have mail delivery to city-style addresses. Enumerators list the residential addresses within their assignment areas on blank address register pages, map spot the location of the residential structures on Census Bureau maps, and conduct an interview for each household.

Local Update of Census Addresses (LUCA). A Census 2000 program, established in response to requirements of Public Law 103-430, that provided an opportunity for local and tribal governments to review and update individual address information or block-by-block address counts from the MAF and associated geographic information in the TIGER® database. The goal was to improve the completeness and accuracy of both computer files. Individuals working with the addresses had to sign a confidentiality agreement before a government could participate. Also called the Address List Review Program.

Long Form. The decennial census long-form questionnaire was used to survey a sample of the U.S. population. It contained the questions on the census short form and additional detailed questions relating to the social, economic, and housing characteristics of each individual and household.

Lower Bound. Represents the low end of the 90 percent confidence interval of an estimate from a sample survey. A 90 percent confidence interval can be interpreted roughly as providing 90 percent certainty that the true number falls between the upper and lower bounds.

Mailing Address. The address used by a living quarters, special place, business establishment, and the like for mail delivery by the USPS. It can be a house number and street or road name, which may be followed by an apartment, unit, or trailer lot designation; a building or apartment complex name and apartment designation; a trailer park name and lot number; a special place/GQ facility name; a post office box or drawer; a rural route or highway contract route, which may include a box number; or general delivery. A mailing address includes a post office name, state abbreviation, and ZIP Code. A mailing address may serve more than one living quarters, establishment, and so on.

Mailout-Mailback. A method of data collection in which the USPS delivers addressed questionnaires to housing units. Residents are asked to complete and mail the questionnaire to a specified data capture center.

Main Phase Sample. The annual ACS sample is chosen in two phases. During the first phase, referred to as the main phase, approximately 98 percent of the total ACS sample is chosen. The main phase sample addresses are allocated to the 12 months of the sample year. The second phase, referred to as supplemental sample selection, is implemented to represent new construction.

Master Address File (MAF). The Census Bureau's official inventory on known living quarters (housing units and GQ facilities) and selected nonresidential units (public, private, and commercial) in the United States. The file contains mailing and location address information, geocodes, and other attribute information about each living quarters. The Census Bureau continues to update the MAF using the USPS DSF and various automated, computer-assisted, clerical, and field operations.

Master Address File Geocoding Office Resolution (MAFGOR). An operation in which census staff try to find the location of addresses from the USPS that did not match to the records in the TIGER® database. Staff use atlases, maps, city directories, and the like to locate these addresses and add their streets and address ranges to the TIGER® database.

Master Address File/TIGER® Reconciliation. A post-Census 2000 MAF improvement activity where census staff reviewed and corrected map spot inconsistencies in over 1,800 counties. Over 75,000 MAF records in nonmailout/mailback blocks were corrected. The most common types of MAF corrections were the assignment of map spots to MAF records such that they are consistent with the TIGER® database, and the identification and linkage of duplicate MAF records.

Margin of Error (MOE). Some ACS products provide an MOE instead of confidence intervals. An MOE is the difference between an estimate and its upper or lower confidence bounds. Confidence bounds can be created by adding the MOE to the estimate (for the upper bound) and subtracting the MOE from the estimate (for the lower bound). All published ACS MOEs are based on a 90 percent confidence level.

Measure of Size (MOS). A generic term used to refer to the estimated size of a specific administrative or statistical area. It is used in the sample selection operation to determine the initial sampling rate at the block level.

Measurement Error. Also referred to as "response error," measurement error occurs when the response received differs from the "true" value as a result of the respondent, the interviewer, the questionnaire, the mode of collection, the respondent's record-keeping system(s) or other similar factors.

Median. This measurement represents the middle value (if n is odd) or the average of the two middle values (if n is even) in an ordered list of data values. The median divides the total frequency distribution into two equal parts: one-half of the cases fall below the median and one-half of the cases exceed the median. Medians in the ACS are estimated using interpolation methods.

Metadata. Information about the content, quality, condition, and other characteristics of data. Metadata related to tables presented in American FactFinder can be found by clicking on column headings or by clicking "Help" and then "Census Data Information."

Minor Civil Division (MCD). A primary governmental and/or administrative subdivision of a county, such as a township, precinct, or magisterial district. MCDs exist in 28 states and the District of Columbia. In 20 states, all or many MCDs are general-purpose governmental units: Connecticut, Illinois, Indiana, Kansas, Maine, Massachusetts, Michigan, Minnesota, Missouri, Nebraska, New Hampshire, New Jersey, New York, North Dakota, Ohio, Pennsylvania, Rhode Island, South Dakota, Vermont, and Wisconsin. Most of these MCDs are legally designated as towns or townships.

Multiyear Estimates. Three- and five-year estimates based on multiple years of ACS data. Three-year estimates will be published for geographic areas with a population of 20,000 or more. Five-year estimates will be published for all geographic areas down to the census block group level.

Municipio. Primary legal divisions of Puerto Rico. These are treated as county equivalents.

Narrative Profile. A data product that includes easy-to-read descriptions for a particular geography.

National Processing Center (NPC). The permanent Census Bureau processing facility in Jeffersonville, Indiana. Until 1998, it was called the Data Preparation Division.

Non-City-Style Address. A mailing address that does not use a house number and street or road name. This includes rural routes and highway contract routes, which may include a box number; post office boxes and drawers; and general delivery.

Noninterview/Nonresponse. A sample address which was eligible for an interview, but from which no survey data was obtained.

Nonresponse Error. Error caused by survey failure to get a response to one or possibly all of the questions. Nonresponse error is measured in the ACS by survey response rates and item nonresponse rates.

Nonresponse Follow-Up. An operation whose objective is to obtain complete survey information from housing units for which the Census Bureau did not receive a completed questionnaire by mail. In the ACS, telephone and personal visit methods are used for nonresponse follow-up.

Nonsampling Error. Total survey error can be classified into two categories—sampling error and nonsampling error. Errors that occur during data collection (for example, nonresponse error, response error, and interviewer error) or data capture fall under the category of nonsampling error.

Office of Management and Budget (OMB). OMB assists the President in the development and execution of policies and programs. OMB has a hand in the development and resolution of all budget, policy, legislative, regulatory, procurement, e-government, and management issues on behalf of the President. OMB is composed of divisions organized either by agency and program area or by functional responsibilities. However, the work of OMB often requires a broad exposure to issues and programs outside of the direct area of assigned responsibility. In accordance with the Paperwork Reduction Act of 1995, the Census Bureau submits survey subjects, questions, and information related to sampling, data collection methods, and tabulation of survey data to OMB for approval and clearance.

Operational Response Rates. Response rates for data collection operations conducted in the ACS—Mail, CATI, CAPI, and FEFU operations.

Optical Mark Recognition (OMR). Technology that uses a digital image of a completed questionnaire and computer software to read and interpret the marking of a response category and to convert that mark into an electronic response to the survey question.

Overcoverage. Extent to which a frame includes units from the target population more than once, giving the unit multiple chances of selection, as well as the extent to which the frame includes units that are not members of the target population.

Period Estimates. An estimate based on information collected over a period of time. For ACS the period is either 1 year, 3 years, or 5 years.

Point-in-Time Estimates. An estimate based on one point in time. The decennial census long-form estimates for Census 2000 were based on information collected as of April 1, 2000.

Population Controls. Intercensal estimates used in weighting ACS sample counts to ensure that ACS estimates of total population and occupied housing units agree with official Census Bureau estimates.

Primary Sampling Unit (PSU). The PSU for the housing unit sample selection is the address. For the GQ sample selection it is groups of ten expected interviews. For the small GQ sample selection operation it is the GQ facility. All residents of small GQ facilities in sample are included in the person sample.

Processing Error. Error introduced in the postdata collection process of taking the responses from the questionnaire or instrument and turning those responses into published data. Thus, processing error occurs during data capture, coding, editing, imputation, and tabulation.

Public Use Microdata Area (PUMA). An area that defines the extent of territory for which the Census Bureau releases Public Use Microdata Sample (PUMS) records.

Public Use Microdata Sample (PUMS) Files. Computerized files that contain a sample of individual records, with identifying information removed, showing the population and housing characteristics of the units and people included on those forms.

Public Use Microdata Sample (PUMS) Management and Messaging Application (PMMA). This system is the PUMS version of EMMA, and is used by analysts to communicate with the data processing team about their review of the PUMS files.

Puerto Rico Community Survey (PRCS). The counterpart to the ACS that is conducted in Puerto Rico.

Quality Assurance (QA). The systematic approach to building accuracy and completeness into a process.

Quality Control (QC). Various statistical methods that validate that products or operations meet specified standards.

Quality Index. A measure of the quality of a particular return which is used when there are multiple returns for a particular sample unit.

Quality Measures. Statistics that provide information about the quality of the ACS data. The ACS releases four different quality measures with the annual data release: 1) initial sample size and final interviews; 2) coverage rates; 3) response rates, and; 4) item allocation rates for all collected variables.

Raking. An iterative procedure whereby a series of ratio adjustments are performed and then repeated. Each ratio adjustment corresponds to a dimension of the raking matrix. The goal of the procedure is to achieve a high degree of consistency between the weighted marginal totals and the control totals used in the ratio adjustment. The raking ratio estimator is also known as iterative proportional fitting.

Ranking Table. Ranking tables are tables and related graphics that show the rank order of a key statistic or derived measure across various geographic areas, currently states, counties, and places.

Recodes. Variables on data files that are the result of combining values from more than one variable.

Reference Period. Time interval to which survey responses refer. For example, many ACS questions refer to the day of the interview; others refer to "the past 12 months" or "last week."

Regional Office (RO). One of 12 permanent Census Bureau offices established for the management of all census and survey operations in specified areas.

Relevance. One of four key dimensions of survey quality. Relevance is a qualitative assessment of the value contributed by the data. Value is characterized by the degree to which the data serve to address the purposes for which they are produced and sought by users (including mandate of the agency, legislated requirements, and so on.)

Remote Alaska. Rural areas in Alaska which are difficult to access. In these areas, all ACS sample cases are interviewed using the personal visit mode. Field representatives attempt to conduct interviews for all cases in specific areas of remote Alaska during a single visit. All sample cases in remote Alaska are interviewed in either January through April or September through December.

Residence Rules. The series of rules that define who (if anyone) is considered to be a resident of a sample address for purposes of the survey or census.

Respondent. The person supplying survey or census information about his or her living quarters and its occupants.

Respondent Errors. The respondent's failure to provide the correct answer to a survey question for any reason, such as poor comprehension of the question meaning, low motivation to answer the question, inability to retrieve the necessary information, or an unwillingness to answer the question truthfully.

Response Categories. The response options for a particular survey question shown on the paper questionnaire, read to the respondent in a CATI interview or read or presented on a flash-card to the respondent in a CAPI interview.

Response Errors. Also referred to as measurement error, response error is any error that occurs during the data collection stage of a survey resulting in a deviation from the true value for a given survey question or questions. Errors made by respondents, interviewer errors such as misreading a question or guiding the response to a particular category, and poorly designed data collection instruments or questionnaires all contribute to response error.

Rolling Sample. A rolling sample design jointly selects *k* nonoverlapping probability samples, each of which constitutes $1/F$ of the entire population. One sample is interviewed each time period until all of the sample has been interviewed after *k* periods.

Sample Month. The first month of a sample's 3-month interview period.

Sampling Entity. Geographic and statistical entities eligible to be used in determining the sampling strata assignment.

Sampling Error. Errors that occur because only part of the population is directly contacted. 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.

Sampling Frame. Any list or device that, for purposes of sampling, delimits, identifies, and allows access to the sampling units, which contain elements of the sampled population. The frame may be a listing of persons, housing units, businesses, records, land segments, and so on. One sampling frame or a combination of frames may be used to cover the entire sampled population.

Sampling Rate. Proportion of the addresses in a geographical area, or residents of a GQ facility, who are selected for interview in a particular time period.

Sampling Variability. Variation that occurs by chance because a sample is surveyed rather than the entire population.

Second Stage Sample. The set of addresses selected from the first phase sample using a systematic sampling procedure. This procedure employs seven distinct sampling rates.

Selected Population Profiles (SPPs). An ACS data product that provides certain characteristics for a specific race or ethnic group (for example, Alaska Natives) or other population subgroup (for example, people aged 60 years and over). SPPS are produced directly from the sample microdata (that is, not a derived product).

Short Form. The decennial census short-form questionnaire includes questions on sex, age/date of birth, relationship, Hispanic origin, race, and tenure.

Single-Year Estimates. Estimates based on the set of ACS interviews conducted from January through December of a given calendar year. These estimates will be published for geographic areas with a population of 65,000 or more.

Size Thresholds. Population sizes of geographical areas that determine when data products will first be released for that area; for example, areas with 65,000 or greater populations will get single-year profiles in 2006 and every year thereafter; areas with 20,000 or greater populations will receive 3-year data products in 2008 and every year thereafter. There are no population size thresholds applied to the 5-year data products other than those imposed by the DRB.

Small Area Income and Poverty Estimates (SAIPE). Census Bureau program that prepares mathematical model-based estimates of selected characteristics of the United States, states, and school districts.

Special Census. A federal census conducted at the request and expense of a local governmental agency to obtain a population count between decennial censuses.

Special Place (SP). A special place is an entity that owns and/or manages one or more GQ facilities. A special place can be in the same building or location as the GQ facility or it can be at a different location than the GQ facility it manages or oversees.

Special Sworn Status (SSS) or Special Sworn Status (SSS) Individual. Individuals with SSS are defined as non-Census Bureau personnel who require access to census information or confidential data. An SSS individual is bound by Census Bureau confidentiality requirements, as authorized by Title 13, United States Code.

Standard Error. The standard error is a measure of the deviation of a sample estimate from the average of all possible samples.

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.

Statistical Areas. Defined and intended to provide nationally consistent definitions for collecting, tabulating, and publishing federal statistics for a set of geographic areas.

Statistical Significance. The determination of whether the difference between two estimates is not likely to be from random chance (sampling error) alone. This determination is based on both the estimates themselves and their standard errors. For ACS data, two estimates are "significantly different at the 90 percent level" if their difference is large enough to infer that there was a less than 10 percent chance that the difference came entirely from random variation.

Strata. See **Stratum.**

Stratum. A grouping or classification that has a similar set of characteristics.

Subsampling. Refers to the sampling of a sample. The cases that are not completed by mail or through a telephone interview become eligible for CAPI interviewing. This winnowing of the sample is referred to as subsampling.

Subject Tables. Data products organized by subject area that present an overview of the information that analysts most often receive requests for from data users.

Successive Differences Replication (SDR). A variance estimation methodology to be used for surveys with a systematic sample. The initial sampling weights are multiplied by sets of 80 predetermined factors, and then reprocessed through the weighting system to produce 80 new sets to replicate weights. The 80 replicate weights and the final production weights are used to estimate the variance of ACS estimates.

Sufficient Partial Interview. A sufficient partial interview means that the Census Bureau accepts an interview as final even if the respondent did not provide a valid response for all applicable items.

Summary File 3 (SF 3). This file presents base tables on population and housing characteristics from Census 2000 sample topics, such as income and education. It also includes population estimates for ancestry groups and selected characteristics for a limited number of race and Hispanic or Latino categories.

Summary File 4 (SF 4). This file presents data similar to the information included on Summary File 3. The data from Census 2000 are shown down to the census tract level for 336 race, Hispanic or Latino, American Indian and Alaska Native, and ancestry categories.

Supplemental Sample. The sample that is selected from new addresses (primarily new construction) and allocated to the last 9 months of the sample year. This is done in January of the sample year.

Survey. A data collection for a sample of a population. Surveys are normally less expensive to conduct than censuses, hence, they may be taken more frequently and can provide an information update between censuses.

Survey of Income and Program Participation (SIPP). A longitudinal survey conducted by the Census Bureau that collects data periodically from the same respondents over the course of several years. The SIPP produces data on income, taxes, assets, liabilities, and participation in government transfer programs.

Survey Quality. The four key elements of survey quality include relevance, accuracy, timeliness, and accessibility.

Survey Response Rates. A measure of total response across all three modes of data collection, calculated as the ratio of the estimate of the interviewed units to the estimate of all units that should have been interviewed. The ACS weights the survey response rate to reflect the sample design, including the subsampling for the CAPI.

Swapping. See **Data Swapping.**

Systematic Errors. Errors or inaccuracies occurring in data consistently in one direction, which can distort survey results. By definition, any systematic error in a survey will occur in all implementations of that same survey design.

Tabulation Month. The month associated with a sample case which is used in producing estimates. Also known as the Interview Month, it reflects the response month, which may or may not be the same as the sample month.

Tabulation Universe. The specific category of people, households, or housing units on which estimates are based; for example, people aged 25 and over or occupied housing units.

Targeting. In the context of the ACS language program, this refers to the identification of geographic areas warranting specific language tools.

Telephone Questionnaire Assistance (TQA). A process which allows respondents to call a toll-free telephone number to receive help when completing the survey questionnaire. This process also allows respondents to complete the survey over the telephone with an interviewer.

Thematic Maps. Data products that show the geographic patterns in statistical data. Thematic maps are a complement to the ranking tables, and are a tool to visually display on a map the geographic variability of a key summary or derived measure.

Three-Year Estimates. Estimates based on 3 years of ACS data. These estimates are meant to reflect the characteristics of a geographic area over the entire 36-month period. These estimates will be published for geographic areas with a population of 20,000 or more.

Timeliness. One of four key dimensions of survey quality. Timeliness refers to both the length of time between data collection and the first availability of a product and to the frequency of the data collection.

Title 13 (U.S. Code). The law under which the Census Bureau operates and that guarantees the confidentiality of census information and establishes penalties for disclosing this information.

Topcoding. A disclosure avoidance practice whereby extremely low or high values are masked by replacing them with a value that represents everything above or below a certain value.

Topologically Integrated Geographic Encoding and Referencing (TIGER®) System or Database. A digital (computer-readable) geographic database that automates the mapping and related geographic activities required to support the Census Bureau's census and survey programs.

Tract. See **Census Tract.**

Undeliverable-As-Addressed (UAA). A USPS notification that a mailing piece could not be delivered to the designated address.

Undercoverage. The extent to which the sampling frame does not include members of the target population thus preventing those members from having any chance of selection into the sample.

Unit Nonresponse. The failure to obtain the minimum required data from a unit in the sample.

Unmailable. A sample address that is inadequate for delivery by the USPS.

Update/Leave (U/L). A method of data collection used in Census 2000 and other censuses, whereby enumerators canvassed assignment areas and delivered a census questionnaire to each housing unit. At the same time, enumerators updated the address listing pages and Census Bureau maps. The household was asked to complete and return the questionnaire by mail. This method was used primarily in areas where many homes do not receive mail at a city-style address; that is, the majority of United States households not included in mailout/mailback areas. U/L was used for all of Puerto Rico in Census 2000.

Upper Bound. Represents the high end of the 90 percent confidence interval of an estimate from a sample survey. A 90 percent confidence interval can be interpreted roughly as providing 90 percent certainty that the true number falls between the upper and lower bounds.

Urbanizacion. An area, sector, or residential development, such as a neighborhood, within a geographic area in Puerto Rico.

Urbanized Area (UA). A densely settled territory that contains 50,000 or more people. The Census Bureau delineates UAs to provide a better separation of urban and rural territory, population, and housing in the vicinity of large places.

Usual Residence. The concept used to define residence in the decennial census. The place where a person lives and sleeps most of the time.

Voluntary Methods Test. A special test conducted at the request of Congress in 2002 to measure the impact on the ACS of changing the data collection authority from mandatory to voluntary.

WebCATI. A control system which is used to track and assign cases to individual telephone interviewers. WebCATI evaluates the characteristics of each case (for example, the date and time of the previous call) and the skills needed for each case (for example, the need for the case to be interviewed in Spanish), and delivers the case to the next available interviewer who possesses the matching skill.

Web Data Server (WDS). A research tool for reporters, SDCs, CICs, ROs, and internal Census Bureau analysts. WDS features a user-friendly interface that allows users to quickly access, visualize, and manipulate ACS base tables.

Weighting. A series of survey adjustments. Survey data are traditionally weighted to adjust for the sample design, the effects of nonresponse, and to correct for survey undercoverage error.