

## Glossary of Technical Terms and Abbreviations

Included in this glossary are technical terms associated with the collection, processing, and tabulation of 1970 census data; terms used in the technical description of census summary files on magnetic tape; and other terminology which may be used in meeting requests for census data or special services. Also included are frequently used abbreviations, many for geographic areas are defined in the Census Users' Dictionary.

**ACG (Address Coding Guide).** A computer listing of streets by block sides. Each record identifies a single block side by a range of addresses, block numbers, tract numbers, and other geographic codes. This geographic tool, created for the city mail delivery areas of 145 SMSA's, does not contain confidential information and is available to the general public on a cost basis.

**Address Register.** A listing of all addresses in mail-out/mail-back areas for housing units (occupied or vacant) and other living quarters. It was used in the preparation of questionnaire mailing labels, for drawing samples of dwelling units, etc. This file is confidential and cannot be made available to any other governmental agency or the public.

**Address Serial Number (ASN).** The unique identifier for addresses in both the census basic records and in the address register. Since ASN's are common to the two files, records in the files may be linked by matching ASN's.

**ADMATCH (Address Matching System).** A computer program package designed for use in assigning census geographic codes to local records with street addresses by using a DIME or similar geographic base file.

**Advance reports.** Advance reports present selected final 1970 census figures in advance of their publication in final reports. Most appear 1 or 2 months prior to the complete census report.

**Allocation.** Assignments of acceptable entries in place of unacceptable or missing entries on a basic record. Such assignments are made where an entry on a given item is lacking or where the

information reported for a person on that item is inconsistent with other information for the person.

**Alphanumeric.** Terminology applied to a character set containing a combination of numbers and letters, and sometimes other characters.

**ASCII (American Standard Code for Information Interchange).** A 7-bit code developed by a committee established in the United States of America Standards Institute as the standard code for representing information to be interchanged with another party.

**ASN.** (See *Address Serial Number.*)

**Base maps.** Master maps which show all the political and statistical boundaries to be observed during the census enumeration.

**Basic record.** Set of data for an individual, household, or housing unit carried on census basic record tapes. (See *Logical record.*)

**Basic Record Tape (BRT).** Computer tapes containing the records of edited census information about each housing unit and each person. Neither names of household members nor addresses are included in these files. These data records are cross-referenced to the Address Register records by address serial number. The BRT and Address Register files are confidential and access to them is restricted to Census personnel for use in developing statistical summaries.

**BCD (Binary Coded Decimal).** A system of character coding in which decimal digits, alphabetic letters, and special characters are represented in terms of binary digits, usually in sets of six bits per character.

**Binary.** A numbering system based on 2 digits, 0 and 1. It permits representation of information on magnetic tape and in computers. A binary condition exists if there is or is not an electrical pulse or a magnetic force at a particular location in a computer memory or on a tape.

**Bit (Binary Digit).** The smallest element of binary machine language. A bit is represented by a magnetized spot on a recording surface or a magnetized element of a storage device. A bit may represent 0 or 1.

**Block.** 1. A physical record on magnetic tape.  
2. A city block. A piece of land generally bounded by four intersecting streets.

**Block length.** The total number of characters contained in one block (physical record).

**Breaker sheets.** A special form prepared and inserted as the first page of each enumeration book to identify the ED during processing, to record preliminary counts of population and housing, and to separate one ED from another.

**BRT** (See *Basic Record Tape*)

**CBD** (Central Business District)

**CCC**(Central County Code)

**CCD** (Census County Division)

**Cell.** Tabulations are often referred to as being composed of tally cells. One cell contains a number representing the count (tabulation) of some kind of unit, such as a person, house, or business, possessing some kind of characteristic; e.g., a certain age or marital status. Refer to *Data Access Description CT-1* (revised).

**Census.** An enumeration or count of an entire population. (See *Complete-count*.)

**Census Use Study.** A unit of the Bureau of the Census which did extensive small area data research in New Haven, Connecticut after the 1967 special census there. The study explored the uses of census statistics at the local level and the integration of census information with data from other sources.

**Central Users' Service.** An office within the Bureau of the Census responsible for handling requests for most 1970 census data products.

**Centroid.** The approximate center of a given geographic area.

**Channel.** One of the parallel tracks on computer tape used for the storage of data.

**Character.** Collections of bits representing symbols which are intelligible when displayed on a printer or other computer output device.

May mean the symbol itself. Characters may represent alphabetic, numeric, and other symbols such as + or - signs.

**Character set.** A systemized group of characters, such as the English alphabet.

**CINCH** (Components of Inventory Change). See the *Census Users' Dictionary*.

**COBOL** (Common Business Oriented Language). A computer programming language.

**Coding.** An operation in the processing of census data in which handwritten entries on the questionnaires (occupation, place of work, etc.) are assigned numeric codes by filling appropriate circles which can be then processed through FOSDIC. A coder assigns a code for each entry according to a detailed set of instructions and fills appropriate circles on the census questionnaire.

**Compilations.** Data presentations which are constructed from existing tabulations. If a user's needs may be met by reorganizing existing tabulations or calculating measures from existing tabulations, the work is termed a compilation.

**Complete-count.** Refers to questions asked of all the people and housing units and the resulting data, as opposed to sample questions and sample data.

**Component summaries.** Census data tallies for portions of States and SMSA's. For example, four types of SMSA component summaries follow each SMSA summary in file B of the Second Count summary tape. They include: (1) Each county inside SMSA; (2) each central city or city of 50,000 or more population inside SMSA; (3) urban balance of SMSA; and (4) rural balance of SMSA.

**Computer graphics.** Techniques for graphically displaying data using a computer and associated output devices such as a printer, pen plotter, or cathode ray tube. Through the use of computer graphics, large amounts of data can be easily visualized or displayed.

**Confidentiality.** Answers to census questions are held in strictest confidence by law. All Bureau of the Census employees are under oath to comply with the law guaranteeing confidentiality.

**Conversion.** The process of changing information from one physical form of represen-

tation to another, such as from the bit arrangements required by one type of computer to that required by another. At the Census Bureau, the term usually has reference to operations performed on tapes produced by the Bureau computers so that they will be compatible for use on other computers.

**Coordinates.** A group of numbers used to identify the location of a point on an x-y plane.

**Core.** Internal computer storage device, in which binary data is represented by the direction of magnetization. Its size is defined in terms of the number of bits usually expressed as so many "K" (thousands).

**Count, First--Sixth.** The term used by the Census Bureau to identify a set of specified data tabulations for certain kinds of geographical units. Numbering refers to the order in which they are released.

**CPI (Characters Per Inch).** A measure of character density on computer tapes. (See *Density and Recording density*.)

**CPS (Current Population Survey).**

**CRT (Cathode Ray Tube).** A television tube used as a computer output for graphics.

**Cross-tabulations.** Tabulations of data structured by other data characteristics. For example, years of age by sex and race.

**CUD (Census Users' Dictionary).** An appendix to the Census Users' Guide.

**DAD (Data Access Descriptions).** Papers issued occasionally by the Bureau of the Census covering specific topics of data access and use.

**Data.** A general term used to denote any or all facts or quantities represented by numbers, letters, or symbols. It also denotes basic elements of information that can be processed or produced by a computer.

**Data item.** A cell of data appearing in a tabulation; sometimes referred to as a tally cell. One of the numbers appearing in a table.

**Data processing.** The execution of one or several computer operations (sorting, calculating, summarizing, etc.) on data.

**DAUL (Data Access and Use Laboratory).** A research and development division of the Bureau

of the Census primarily concerned with making census data more accessible and useful to users in government, private industry, and the academic community.

**DAULLIST.** Programs developed by the Bureau of the Census to display data from the 1970 census summary tapes.

**Deck.** A set of punched cards.

**Density.** The number of characters or groups of bits recorded on an inch of tape. Frequently used densities are 556, 800, or 1600 cpi (characters per inch). A particular computer system is capable of reading only certain specified densities of magnetic tape. (See *CPI* and *Recording density*.)

**Diary.** A set of records produced during computer operations showing significant processing events, used particularly to read frequency of edits, allocations, and inconsistencies.

**Diary review.** The professional and clerical review of editing diaries to determine the acceptability of the data and/or the corrective action to be taken.

**Digitizer.** A machine used for assigning x-y geographic coordinates to points on a map.

**DIME (Dual Independent Map Encoding).** The name given to a technique of creating a geographic base file. Also, a term used synonymously with "geographic base file."

**DO (District Offices).** Temporary offices established for the conduct of the 1970 census enumeration.

**DOS (Disk Operating System).** A monitor system for some computers.

**Dump.** A printout of the contents of a computer tape, disk, or core storage.

**EBCDIC (Extended Binary Coded Decimal Interchange Code).** A character set using 8-bit codes, designed for use with IBM 360 computers.

**ED (Enumeration District).**

**EDP (Electronic Data Processing).**

**Editing.** The operations in the processing of census data which involve checking the data for completeness, consistency, and meaning. For example, household heads must be at least 14 years old, wives must be married, medical

doctors must have college degrees, etc. Editing is usually done by the computer. (See *Allocation*.)

**Enumeration.** The act of taking a census. A census taker is called an enumerator.

**ESR (Economic Subregion).**

**Field.** A reserved area in a record which serves a similar function in all records of that group of records; a specified area of a record used for a particular category of data. For example, the codes indicating which geographic area a tabulation represents may be located in a field at the beginning of that tabulation record.

**File.** A collection of many logical records with common identifying features contained in any media.

**File maintenance.** The periodic updating of a file.

**FIPSPUB (Federal Information Processing Standards Publication).** National Bureau of Standards publication.

**Flag.** A signal used in data processing. Usually a simple symbol expressing a complex condition.

**Format.** The pre-determined arrangement of the characters, fields, lines, etc., of a single record or file.

**FORTRAN (Formula Translation).** A computer programming language.

**FOSDIC (Film Optical Sensing Device for Input to Computer).** A general term for that part of a census document reader capable of reading the contents of microfilm by optical methods and converting the information codes to magnetic tape.

**GBF (Geographic Base File).** A file which describes the geographic attributes of an area, normally in terms of street segments with associate descriptors and codes.

**General tabulations.** Tabulations developed as part of the regular census program. They include the contents of 1970 Census Summary Tapes, printed reports, and microform products.

**Geospace plotter.** A commercially available computer mapping device used by the Census Use Study experimentally in the New Haven SMSA.

**GICS (Geographic Identification Code Scheme).** A listing of geographic codes used with the 1960 census.

**GQ (Group Quarters).**

**GRIDS.** A computer mapping software system designed to display data on the printer of a small-scale computer.

**Hard copy.** A printed copy of machine output in an eye-readable form; for example, printed reports, listings, and documents.

**Hardware.** Computers and their peripheral equipment.

**Header record (or label).** Both a logical and physical record that contains descriptive information for a group of logical records which are to follow. Header records for census summary and basic record tapes generally contain codes identifying the geographic areas represented by the subsequent file. Volume label is one type of header label.

**Hexadecimal.** A computer numbering system in which numbers are expressed to the base 16.

**HSP (High Speed Printer).**

**HU (Housing Unit).**

**ICR (Individual Census Report).** A special census questionnaire used to obtain population information for persons in group quarters.

**Input data.** Data which is to be processed by a computer.

**I/O (Input/output).**

**Interblock spacing.** The separation of physical records on magnetic tape by the use of blank spaces.

**List.** To print relevant items of data.

**Listing books.** Books that contain addresses of the residents in an enumeration district. Listing books were prepared to improve the coverage of the census enumeration.

**Logical record.** In reference to Census Bureau summary tapes, a set of data items for a particular geographic area accompanied by a set of geographic identifiers. A logical record may be a part of a physical record or may consist of one or more physical records. A logical record may also refer to a set of data for any unit of analysis such as an individual, household, or housing unit as carried on census basic record tapes.

**Logical record size.** The total number of characters contained in one logical record.

**Long form.** A term used in reference to the 1970 census sample questionnaire (the questions range in number from 66 to 74) which was distributed to 1 out of every 5 households or 20 percent of the population. There were two versions of the long form: the "15 percent long form" sent to 15 out of every 100 households and the "5 percent long form" sent to 5 out of every 100 households.

**Magnetic tape.** A tough plastic film that has one of its sides coated with iron oxide that can be magnetized. It is used as both an input and output media with computers.

**Mail-out/mail-back.** A collection technique used in the 1970 census by which persons, usually those living in large metropolitan areas, were asked to mail back their completed questionnaires rather than waiting for an enumerator to pick them up.

**Matching study.** A procedure involving matching individual census records to a list of names and addresses and preparing statistical summaries of census characteristics from the records. The confidentiality of replies to census questions is preserved in the preparation of matching studies.

**Matrix.** A statistical table, usually cross-classifying two or more variables.

**MCD (Minor Civil Division).**

**MEDList (Master Enumeration District List).** A hierarchical code list which relates each State, county, MCD, and place segment by name to all relevant census geographic codes.

**Microfiche.** A sheet of microfilm containing microimages in a grid pattern.

**Microfilm.** A ribbon-like film that has reduced photo images of printed or tabulated material.

**MMS (Metropolitan Map Series).**

**MRC (Major Retail Center).**

**MRF (Master Reference File).** A Census Bureau work tape which is the single source of geographic codes for all areas recognized in regular census tabulations. A modified version (MEDList) is available to the public.

**MSP (Married, Spouse Present).**

**Node.** As used in creating geographic base files, a dot, uniquely numbered, placed on a map showing the point at which map features either intersect, end, or curve sharply.

**Nonmail area (Conventional enumeration area).** Areas where census enumerators visit each housing unit.

**Numeric code.** A code which only contains numbers.

**Octal.** A computer numbering system which expresses numbers to the base 8.

**Outer line.** The potential urbanized area boundary.

**Output data.** Data which has been processed by a computer and transferred from memory to another device.

**Padding.** Entering dummy data or blanks to fill out a recording block on tape.

**Parity.** An even or odd bit configuration used to check the validity of information on a storage media.

**Pass.** A full cycle of performance of all the computer processes involving all of a particular file; e.g., to "pass" the sample basic record file to produce a certain set of tabulations. (See *Run.*)

**Pen plotter.** A computer mapping output device which automatically moves a pen over a piece of paper, under program control, to produce graphical output.

**Peripheral devices.** Input and output equipment which may be placed under the control of a computer's central processing unit.

**Physical record.** A set or block of characters on magnetic tape which is read into a computer as a unit; may be referred to as a block or a recording block. Physical records are separated by interblock spacing on the tape. Each physical record may contain one or more logical records or some fraction of a logical record requiring more than one physical record.

**Preliminary Reports.** Unofficial population and housing counts compiled in the census district offices.

**Printout.** A paper copy of data records reported from the computer system through a printing device.

**Program.** The complete set of instructions which determine the sequence and type of computer actions.

**Programing language.** A language used by programmers to prepare computer instructions.

**Public Use Sample.** A sample of census basic records for individual persons and households, made available for statistical use outside the Census Bureau. Identifying information, such as name and detailed geographic location, are omitted to preserve the confidentiality of the census records.

**Q.C. (Quality Control).** Procedures for checking the reliability of produced results.

**Recode.** A code representing information derived from two or more existing codes.

**Record (See *Basic record, Logical record, and Physical record.*)**

**Record layout.** A description of the arrangement of data in a record.

**Recording block.** (See *Physical record.*)

**Recording density.** The number of machine characters recorded on an inch of magnetic tape measured in characters per inch. (See *CPI and Density.*)

**Releasability analysis.** A procedure for determining whether or not a general or special census product may be released to the public. Decisions are based on concern for confidentiality, adequacy of documentation, and statistical validity.

**Run.** The act of operating on a file in the computer with a program.

**SADNotes (Small-Area Data Notes).** A newsletter issued monthly by the Bureau of the Census.

**Sample data.** The data collected from a random selection of the reporting units in some specified universe.

**Sample survey.** A collection of data from a part or sample of a population used to make estimates for the entire population.

**SC-4411.** A special machine which "reads" a summary tape, operates on specific formatting instructions, and arranges the tapes' contents for photographing. It is used in producing microfilm copy of the First and Third Count summary tapes.

**SCA (Standard Consolidated Area).**

**Schedule.** The document or questionnaire on which answers to census or survey questions are recorded. The 1970 census used three kinds of schedules: the "short form," which includes only the complete-count questions; and two versions of the "long form" which include both complete-count and sample questions.

**SCRIS (Southern California Regional Information Study.)** A local and Bureau of the Census statistical research project.

**SEA (State Economic Area).**

**Self-enumeration.** A procedure whereby a questionnaire is filled out by the respondent rather than having the respondent personally interviewed by an enumerator.

**Sentinel.** A group of characters on a tape used to denote the termination of a file.

**Short form.** A term used in reference to the 1970 census questionnaire which was distributed to 4 out of every 5 households or 80 percent of the population. The short form schedule consisted of 23 basic questions.

**SIC (Standard Industrial Classification).** Numeric code scheme for industries.

**SLA (Standard Location Area).**

**SLEUTH (Symbolic Language for the Univac 1107 Thin Film Computer).** Programing language used to process the 1960 and 1970 census.

**SMSA (Standard Metropolitan Statistical Area).**

**Software.** Computer programs and procedures.

**Special areas.** Territorial units not recognized in the standard census codes, such as school districts, traffic zones, and planning areas.

**Special project.** Term used to designate jobs or special services undertaken for persons or institutions outside the Census Bureau on a reimbursable or cost basis. These jobs may involve the copying of summary tape files or the preparation of photocopies of tabulations, etc. (See *Special tabulations.*)

**Special tabulations.** Retabulations of the census basic record tapes (BRT's) at users' expense to obtain data summaries for geographic areas not recognized in general tabulations and/or

including subject breakdowns or cross-classifications not appearing in general tabulations.

**Stratification.** Grouping data into class intervals.

**Stripping.** An operation of selecting and copying specified data for only a few geographic areas and/or tabulations from a data source; for example, copying selected data from a complete summary tape to a deck of cards.

**STUM (Summary Tape User Memoranda).** A series of memoranda issued frequently by the Bureau of the Census as information becomes available relevant to summary tape use.

**Sub-block.** In reference to Census Bureau magnetic tapes, a 120-character part of a physical record consisting in most cases of complete 8- or 16-character data fields.

**Summary tape.** Magnetic tape recording of strings of numbers in the form of magnetized dots or bits which represent a series of accumulations or tallies of responses from a set of census questionnaires.

**Summary Tape Processing Center.** An organization which intends to obtain census summary tapes and offer services to 1970 census data users.

**Suppression.** The exclusion from public release of certain data items which might violate the confidentiality of census information.

**TA (Tracted Area).**

**Table generator.** A computer program which aggregates unit records into summary table reports through generalized table specifications supplied by the requester.

**Tabulations.** Summaries of basic records.

**Tally.** To count or aggregate raw data.

**Tape drive.** A device which moves magnetic tape past computer sensing and recording mechanisms.

**Tape mark.** A special character used in the gap between files. This term often refers to the gap itself.

**Technical Documentation (TD).** Descriptions of the contents of particular tape files.

**Terminal.** Any mechanism which can send and/or receive data through a communications or systems network.

**Time sharing.** Multiple use of available computer time, often via terminals.

**Track.** A path parallel to the edge of the magnetic tape used for the storage of data. (See *Channel*.)

**Trailer record (or label).** A physical record which follows a file and its tape mark and contains data related to the file.

**UA (Urbanized Area).**

**UAC (Universal Area Code).**

**Volume label.** Both the first logical and physical record on a header-labeled tape, consisting of 80 characters containing the four characters VOL 1, a 6-character tape reel number, and 70 spaces.

**Weighting.** A weight is the value assigned to each unit in a sample which expresses the number of cases in the universe which it represents. Special ratio estimation procedures are used to determine the weights.

**Word.** A set of characters or bits that occupies one storage location on tape or in the computer and is treated by the computer circuits as a unit. Word lengths may vary from one kind of computer to another.

**XS-3 (Excess-three Code).** A BCD code in which numbers are represented by the binary equivalent of that number plus three. XS-3 codes are used in Univac programs at the Bureau. Their alphabetic and symbolic parts are not equivalent to IBM BCD codes.

**Abbreviations for Selected United States Government  
Agencies, Offices, and Other Units**

- BDSA - Business and Defense Services Administration  
 BIC - Bureau of International Commerce  
 BLS - Bureau of Labor Statistics  
 BOB - Bureau of the Budget  
 BPR - Bureau of Public Roads  
 CEN - Bureau of the Census
- APSD - Administrative and Publication Services Division  
 DAUL - Data Access and Use Laboratory  
 DSD - Demographic Surveys Division  
 FLD - Field Division  
 GEO - Geography Division  
 GRD - General Reports Division (formerly Statistical Information Division)  
 HSG - Housing Division  
 PIO - Public Information Office  
 POP - Population Division  
 PROC - Processing Division  
 SMD - Statistical Methods Division  
 SRD - Statistical Research Division  
 SYS - Systems Division
- CSC - Civil Service Commission  
 DOD - Department of Defense
- DOT - Department of Transportation  
 EDA - Economic Development Administration  
 ESSA - Environmental Science Services Administration  
 FHA - Federal Housing Administration  
 GAO - General Accounting Office  
 GPO - Government Printing Office  
 GSA - General Services Administration  
 HEW - Department of Health, Education, and Welfare
- HUD - Department of Housing and Urban Development  
 IRS - Internal Revenue Service  
 NBS - National Bureau of Standards  
 NCES - National Center for Educational Statistics  
 NCHS - National Center for Health Statistics  
 NIH - National Institutes of Health  
 NIMH - National Institute of Mental Health  
 NSF - National Science Foundation  
 OBE - Office of Business Economics  
 OCD - Office of Civil Defense  
 OEO - Office of Economic Opportunity  
 PHS - Public Health Services  
 PO - Post Office  
 SBA - Small Business Administration  
 SSA - Social Security Administration  
 VA - Veterans Administration