

1980 Census of Population and Housing County Boundary File. Technical Documentation

Note: This technical documentation was issued at same time as the corresponding data file and does not necessarily reflect changes to the data file that may have been made later. We offer it as a service to our customers who have purchased the file, but we do not in general offer technical support for past data files. Note that although the documentation may refer to data on tape, the U.S. Census Bureau no longer releases data on magnetic tape; customers purchasing files previously released on tape will receive the file in ASCII on CD-ROM. For information about this and other technical documentation, please contact the Customer Services Center at 301-763-INFO (4636).

Additionally, Census Bureau organization, telephone numbers, and staff cited in this technical documentation are not current. For updated information, see our Telephone Contacts site at <http://www.census.gov/contacts/www/contacts.html> or Staff Search at <http://www.census.gov/cgi-bin/main/email.cgi>.

D1-D80-CTYB-14-TECH

U.S. Census Bureau
Marketing Services
Office

USCENSUSBUREAU
Helping You Make Informed Decisions • 1902-2002

U.S. Department of
Commerce
Economic and Statistics
Administration
U.S. Census Bureau



Census of Population and Housing, 1980:
County Boundary File

Technical Documentation

The documentation consists of the following materials:

Attachment 1	Abstract
Attachment 2	The 1980 Census Digital County Boundary File: A Description
Attachment 3	Example of File Record Sequence

NOTE

Questions about the accompanying documentation should be directed to Data User Services Division, Data Access and Use Staff, Bureau of the Census, Washington, D.C. 20233. Phone: (301) 763-2074.

Questions about the tape should be directed to Data User Services Division, Customer Services (Tapes), Bureau of the Census, Washington, D.C. 20233. Phone: (301) 763-4100.

Questions about the subject matter should be directed to Geography Division, Bureau of the Census, Washington, D.C. 20233. Phone: (301) 763-7856.

ATTACHMENT 1

Abstract

Census of Population and Housing, 1980:
County Boundary File [machine-readable data
file] / prepared by the Bureau of the Census.
--Washington: The Bureau, [producer and
distributor], 1982.

TYPE OF FILE:

Geographic reference (geographic spatial data base).

SUBJECT-MATTER DESCRIPTION:

A series of records containing FIPS State and county codes; alphabetic State and county names; 1980 census population count; 1980 area measurement figures for land, water, and land/water total; and a series of coordinate records for each geographic area boundary.

GEOGRAPHIC COVERAGE:

State and counties.

FILE SIZE:

474,130 logical records; 132 character logical record length.

REFERENCE MATERIALS:

"Census of Population and Housing, 1980: County Boundary File Technical Documentation" (this document). The documentation includes this abstract and a detailed description of the file.

FILE AVAILABILITY:

The file is available from Data User Services Division, Customer Services (Tapes), Bureau of the Census, Washington, D.C. 20233. The cost is \$140 per reel. This file is on 2 reels at 1600 bpi, and on 1 reel at 6250 bpi. When ordering, please refer to file number Cu GEO 80 002.

**Customer Services
(Microfiche, Tapes and Maps)**

Name		RETURN TO: Data User Services Division Customer Services Bureau of the Census Washington, D.C. 20233 Phone: 301/783-4100			Date	
Organization					CHECK ONE: <input type="checkbox"/> Enclosed is check or money order, payable to "Commerce-Census"	
Address					<input type="checkbox"/> Charge to Census Bureau Deposit Account No. _____	
City, State, ZIP						
Telephone						
Name of Data File, Documentation* Map (Series and Areas)		Order No.	No of Reels	No. of Map Sheets	No. of Document copies	Cost
*One copy of the Technical Documentation is sent at no extra charge With tape orders; additional copies are \$5.00 each.					TOTAL \$	
Characteristics of Tape (Check One): <input type="checkbox"/> 9 track, 1600 bpi, EBCDIC <input type="checkbox"/> 9 track, 6250 bpi, EBCDIC <input type="checkbox"/> 9 track, 1600 bpi, ASCII <input type="checkbox"/> 9 track, 6250 bpi, ASCII				Labeling (Check One): <input type="checkbox"/> Standard <input type="checkbox"/> Unlabeled		
Other characteristics are available by special arrangement.						

The 1980 Census
Digital County Boundary File:
A Description

by
Frederick R. Broome
and
Roy F. Borgstede

September 12, 1981

INTRODUCTION

Digital cartographic data bases have evolved over the past 2 decades from inflexible, highly structured sequential files containing limited spatial information into the present flexible, more freely formed files containing spatial coordinates, as well as selected geographic relationships and information. The first file of this modern structure was the the Geographic Base File/Dual Independent Map Encoding^{1/} (GBF/DIME) files produced by the United States Bureau of the Census as a by-product of the address coding for the 1970 Decennial Census. A similar file structure was used to produce the DIME county boundary file (DIMECO). Now the Bureau has developed a file structure which is even more flexible than that used for the GBF/DIME while simultaneously carrying more geographic relationships and information than either the DIMECO or the GBF/DIME files. It is one which can be reduced easily to a simpler structure level if desired. The purpose of this paper is to describe the new geographically encoded string structure as it has been applied to the 1980 digital county boundary file.

BACKGROUND

The 1980 digital county boundary file was developed as a by-product of another Census project, namely the Bureau of the Census project to remeasure the land and water area of the states and counties of the United States.^{2/} Since each county boundary had to be traced with a measuring instrument, it was decided to perform the tracing using an electronic digitizing device. The output of this device was a series of xy-coordinates, that is, coordinate strings which described a portion of a boundary. The computer then used the digitized strings to compute the areas the boundary strings enclosed.

The tracings for measurements were done on the largest scale maps available. Usually these were 1:24000 United States Geological Survey (USGS) 7.5 minute quadrangle sheets. Some 1:63360, 1:100000, or 1:250000 sheets were used where larger scale sheet coverage was not available. The boundary information on the USGS quadrangles was updated to reflect the 1980 county boundaries as shown on the 1980 Boundary and Annexation Survey maps.^{3/} The results of this large scale measurement of county boundaries is the most accurate digital county boundary file ever created of the entire U.S.

PROJECT DESCRIPTION

The area measurement project had two purposes: first was the remeasurement of the areas of the states and counties, or their equivalents, for the U.S. and second was the production of a digital cartographic data base containing the area measurement information along with the boundary coordinates.

The problem was to develop a file structure which could take advantage of this new source of digital data. It was decided that the new file structure should have the following characteristics:

- Maintainability
- Maximize the geographic relationship information
- Minimize reliance on external documentation
- Simplicity of use
- Ease of conversion to other file structures

The most important characteristic was determined to be maintainability. The inability to update digital cartographic files has been a major reason that caused many otherwise good files to fall into disuse. At the very least the cumbersome maintainability severely limited file use. Therefore, the file structure decided upon by the Bureau of the Census strongly reflects the importance of easy updates.

The geographically encoded string structure (GESS) described herein not only achieves the critical goal of maintainability, but ensures the other four desired characteristics, as well.

FILE DESCRIPTION

The 1980 county boundary digital geographic data base is logically founded upon the relationship of the boundary strings to the areas they enclose and conversely the areas are related to the strings bounding them. The physical file structure reflects these relationships, but is constrained in two ways. First, for computer processing it is a fixed length record format. Second, for partitioning the data, the strings are maintained in their original USGS map parts, as will be discussed later. It is important to note that the physical file structure described here is designed for ease of file transfer, and like most other file structures, may not be the desired structure when installed on a particular computer. The physical structure is the common, easily readable format for file transfer between computer systems and is a base structure from which others may be derived. Therefore, understanding the GESS structure is fundamental to taking advantage of all the capabilities of the file.

The GESS file physical characteristics are,

- sequentially organized file of varying record types
- fixed length physical records (130 characters)
- variable length logical records of 130 character units
- alphanumeric encoding

The record type is given in the first character of each record; there are currently eight record types:

- 0 - File description records.
- 1 - Total file parameters record.
- 2 - State parameters records.
- 3 - County parameters records.
- 4 - County additional attributes records.
- 5 - County boundary string pointer index records.
- 6 - Boundary string parameters records.
- 7 - String information records.
- 8 - Boundary coordinates records.

The different records are presented on page 8; they appear in the file in the following way:

```

Char.    11111111112222222223                               1
123456789012345678901234567890.....3
                                                0

0      <FILE DESCRIPTION>text....
1      <TOTAL FILE>text.....
2 ss   <STATE>data.....
3 sccc<COUNTY>data.....
4 sccc<ATTRIBUTES>data....
5 sccc<ppPOINTER>data....
6nnnnnn<BN DY STRING>data....
7nnnnnn<STRING INFO>data....
8nnnnnnmmmmcoordinates.....

```

where "ss" is the Federal Information Processing Standard (FIPS) state code, "ccc" is the FIPS county code, "pp" is the number of that pointer record, "nnnnn" is the string identification number ranging from 100000 to 999999, and "mmm" is the string record number which goes from 000 to 999.

The file, with the exception of the FILE DESCRIPTION records, can be sequentially sorted by keying first on character positions 2-7, second on character 1, and third on characters 8-10 and then sorting in ascending sequence.

DETAILED RECORD TYPE DESCRIPTION

Record Type 0 - File Description Record

These are text records which describe the contents of the entire file and each individual record type. These records may become jumbled if sorted as part of the whole file.

Record Type 1 - Total File Parameters Record

This record is a historical record of the total file. It records the date the file was created and the date of the last update. Included in this record is the highest string identification (ID) assigned in the file.

Record Type 2 - State Parameters Record

This record registers selected attributes of the state such as the FIPS state code, the alphabetic state name, and the maximum/minimum latitude/longitude of the state, that is, the state envelope.

Record Type 3 - County Parameters Record

This record is similar to record type 2, above. It records the FIPS state and county code, the alphabetic county name, and the county envelope. This record type also is used for county equivalents.

Record Type 4 - County Attributes Record

This record type is used to record the attributes which can be ascribed to the entire county such as the 1980 population or the total land area measurement. It carries the FIPS state and county code to indicate which county is being referenced. This record also is used to give point information pertaining to the entire area, for example, a population centroid. If the data are compiled from some special date, say 1970, it is so indicated. Thus one county may have several type 4 records.

Record Type 5 - County Boundary String Pointer Record

The type 5 record contains the FIPS state and county codes, an index record sequence number for this county and a series of string IDs for the strings that make up the county boundary. The record, or records as the case may be, gives the string IDs number of each coordinate string bounding the county as a positive number if the string is to be read from beginning to end and as a negative number if the string is to be read from end to beginning. The sequence of these IDs is the order in which the strings link to form the county, though not necessarily always clockwise or counter clockwise. If a county is in multiple parts, for example, islands, then an ID of 000000 separates the parts. If multiple pointer records are used, the number of coordinate records is listed. There can be up to 100 pointer records for an individual county.

Record Type 6 - Boundary String Record

This record presents the string as an encoded description of the boundary between two geographic areas. It contains the left and right FIPS state and county codes. Left and right are determined by imagining yourself to be standing at the beginning point of the string and looking along the string toward the end point. The record also contains the maximum and minimum coordinates enclosing the string, that is, the envelope. This record identifies the string by a 6-digit string identification number in character positions 2 through 7. In addition, the record gives the number of points in the string.

Record Type 7 - String Information Record

This record is used to provide other information about the string such as the name of the map from which the string was digitized and the date it was verified for accuracy.

Record Type 8 - String Coordinates Record

This record contains the coordinates of the string in decimal degrees of longitude/latitude to the fourth decimal place (equivalent to approximately 10 meters). Each type 8 record can contain up to seven points. Each string is assigned a unique string ID number between 100000 and 999999. Since there may be more than seven points to a string, the three digits following the string identification number are string subrecord numbers. A maximum of 1,000 coordinate substring records for a string is possible. Their numbers would be between 000 and 999. The total number of coordinate pairs (points) in a boundary string is stored in the strings associated record type 6. Both records will have the same string ID number. No string can have more than 7,000 points.

FILE MAINTAINABILITY

The basic structure of the file is a string of coordinates for each map. Should the boundary change, or a new and more accurate map become available for that area, or some similar occurrence indicate the need to change a string in the file, the following record types would be affected: The type 5 record pointers for the two geographic areas bounded by the string would have to be assigned the new string ID; record types 6, 7, and 8 would have to be created for the new string (the old string records 6, 7, and 8 could remain in the file for historical purposes); the type 1 record would have to be updated to show the highest string ID used and the date of last change.

Because the strings are maintained in a national frame of reference in terms of USGS quadrangles and are not kept as whole strings for a common county boundary, updates can be confined to the USGS quad(s) affected. Thus, it is not necessary to extract those sections of a larger string which must be changed and link the change into the larger string. The USGS map edges serve as built-in link points in the GESS structure.

A possible procedural sequence for updating boundaries is as follows:

1. The Boundary and Annexation Survey indicates a change common to two or more counties.
2. The change is located on USGS quadrangle maps.
3. The digital cartographic file is scanned for type 7 records matching the USGS coverage.
4. The string identification of the type 7 record is used to recall the type 6 records identifying the actual counties involved (this functions as a quality assurance check).
5. The type 5 records for the affected counties are retrieved.
6. The new string is digitized and assigned a new string ID. Record type 1 is updated with the new string number and date of change.

7. The type 5 records have their pointers assigned to the new string.
8. New type 6, 7, and 8 records are inserted into the file for the new boundary string.
9. The update is complete.

SIMPLICITY OF USE

The GESS structure is designed to permit both random and sequential file accesses. If one has data to be mapped, GESS provides,

- area figures for use in density computations so that separate area files will not have to be referenced.
- population count for per capita computations.
- pointers sequenced in order around polygons.
- information of which areas are adjacent to each other.
- predetermined envelopes for rapid spatial searches such as nearest neighbor or windowing for special displays.
- centroids of population and geographic areas.

In short, much of the work of geographic computation has already been performed. The user then is free to keep or strip away as much of the information as required for a project or as limited by the hardware on which the file will be maintained. The user may concentrate on file applications.

CONVERSION TO OTHER FILE STRUCTURES

Frequently, it is desirable to keep an old applications program intact in order to avoid rewriting it to accept a new cartographic data base. The GESS structure is easy to convert to the two most popular file structures currently in use, polygons and segments. Further, it is easy to reduce the number of points in the file by "thinning" to a smaller scale resolution.

For example: one may wish to produce a DIMECO file with approximately the same number of points as were in the 1970 file and in the same structure. The 1970 file has about 100,000 points and represents a map at a scale of 1:5,000,000. The procedure might go as follows:

- Read the strings from the GESS file one at a time.
- Apply a scaling factor to each point to transform them to a scale of 1:5,000,000.
- Compute the distance between each pair of points along the string.
- If the difference is less than 1.0, ignore that segment.
- If the difference is greater than 1.0, store that segment. It will become part of the new "DIMECO" file.
- Repeat the procedure until all strings have been processed.

With only little modification, the above procedure will result in a DIMECO type county boundary file of approximately 100,000 points.

SUMMARY

The geographically encoded string structure (GESS) developed for the digital cartographic county boundary file of the U.S. Bureau of the Census is flexible, maintainable, and easy to use. It incorporates numerous pre-calculated attributes and data items which reduce the computational load for the typical digital cartographic file user. It can be used to keep historical track of changes and can be expanded to contain the strings of additional geographic areas. It is the most up-to-date file of the total U.S. county boundaries in existence. Finally, the architectural principles of this structure can be applied to a host of other cartographic files.

Footnotes

- 1/ Dual Independent Map Encoding (DIME) is a coding scheme based on the principles of graph theory. For more information, see U.S. Bureau of the Census Technical Paper No. 48.
- 2/ Historically, the area measurement of the United States has been the responsibility of the U. S. Bureau of the Census. The last complete remeasurement was conducted for the 1940 census. The measurements for the 1960 census areas accepted the 1940 state area figures as constants and then adjusted the other figures to conform with them. The 1980 remeasurement is the first complete measurement since 1940.
- 3/ The Boundary and Annexation Survey is a periodic survey of places and counties that identifies any changes (annexations or detachments) which have occurred in their political boundaries since the previous survey. An official representative of each place or county receives a map showing the boundaries as of the previous survey date. The official then records any changes, certifies the map, and returns it to the Census Bureau.

