

README

Census 2000 Summary File 2 Supplement Delivered via FTP

Contents

File Description

About the FTP Application

File Naming Conventions

Technical Documentation

Data Records and Segmentation

File Record Layout

FTP File Transfer

Spreadsheet and Data Base Aids

File Description

The Summary File 2 (SF 2) Supplement contains matrix PCT37, which looks identical to SF 2 matrix PCT5 (see File Record Layout below). The only difference between matrix PCT37 and matrix PCT5 is that the data in PCT37 are tallied according to the race, Hispanic or Latino origin, or American Indian or Alaska Native tribe of the *person*. Matrix PCT5 was tallied according to the race, Hispanic or Latino origin, or American Indian or Alaska Native tribe of the *householder*.

About the FTP Application

- The application is intended for experienced users of census data and spreadsheet/database software.
- FTP users should have a fast file transfer capability.
- Users of the FTP application need to download the file and then import it into the spreadsheet/database software of their choice for data analysis and table presentation.

File Naming Conventions

- The naming convention for geographic header files is *ssgeo.sf2*.
 - *ss* is USPS state abbreviation
 - The codes are in technical documentation on page 7-1, located at <http://www.census.gov/prod/cen2000/doc/sf2.pdf>
 - *geo.sf2* is a constant across SF2 geographic header names.
- Naming convention for SF2 data files is *ssiiyy.sf2*.
 - *ss* is USPS state abbreviation
 - *iii* is the characteristic iteration (total population, race groups, American Indian and Alaska Native tribes, and Hispanic/Latino groups)
 - Characteristic iteration codes are in the Appendix H of the technical documentation, which is available at <http://www.census.gov/prod/cen2000/docs/sf2.pdf> - page=278.
 - *yy* is the number of the file
 - Valid code is 01. See below for distribution of tables across files.
 - *.sf2* is a constant across SF2 data file names

Technical Documentation

- The complete technical documentation for SF2 is available at <http://www.census.gov/prod/cen2000/doc/sf2.pdf>.

Data Records and Segmentation

- If the total population (iteration 001) for a geographic entity is less than 100, there will be no geographic header record and no data records for that area.
 - *Example:* The population of One Road township, Roberts County, South Dakota is 54. There will be no records, either geographic header or data, for this area.
- If the population for a race, tribe, or Hispanic/Latino group (iterations 002-463) within a geographic entity is 100 or greater, geographic header and data records will be present for the group.
 - *Example:* The Black or African American alone population for Bremer County, Iowa is 110. There will be a geographic header record and data records for the Black/African American iteration (iteration 004) for this area.
- If the population for a race, tribe, or Hispanic/Latino group (iterations 002-463) within a geographic entity is less than 100, geographic header and data records will not be present for that group.
 - *Example:* The Hispanic/Latino population for South Harding CCD, Harding County, New Mexico is 19. There will be no data records for the Hispanic/Latino iteration (iteration 400) for this area.

- Table distribution across data files is as follows:

File name	Number of data items	Matrix number
<i>ssiii01_sf2</i>	49	PCT37

- Five fields are carried over from the geographic header file into each data file.
 - These fields are file identification (FILEID), state abbreviation (STUSAB), characteristic iteration (CHARITER), characteristic iteration file sequence number (CIFSNUM) and logical record number (LOGRECNO).
 - These five fields appear in the geographic header record are in a fixed field format.
 - These five fields appear in the data file in a comma delimited format.
 - These five fields are used to “match” records in the data file for a particular characteristic iteration to the geographic information in the geoheader.
 - A file set structure schematic appears below:

Figure 2–1a. **File Set Structure Schematic — File Set for Total Population Characteristic Iteration**

Geographic Header File File name=algeo.sf2	File 01 File name=al00401.sf2
<p><i>Record 1</i> FILEID (uSF2) STUSAB (AL) CHARITER (000) CIFSN (00) LOGRECNO (0000001) Remainder of geographic header record for geographic area x</p> <p><i>Record 2</i> FILEID (uSF2) STUSAB (AL) CHARITER (000) CIFSN (00) LOGRECNO (0000002) Remainder of geographic header record for geographic area y</p>	<p>FILEID (uSF2) STUSAB (AL) CHARITER (001) CIFSN (01) LOGRECNO (0000001) Tables PCT37 (49 cells)</p> <p>FILEID (uSF2) STUSAB (AL) CHARITER (001) CIFSN (01) LOGRECNO (0000002) Tables PCT37 (49 cells)</p>

Figure 2–1b. **File Set Structure Schematic — File Set for Black or African American Alone Population Characteristic Iteration**

Geographic Header File File name=algeo.sf2	File 01 File name=al00401.sf2
<p><i>Record 1</i> FILEID (uSF2) STUSAB (AL) CHARITER (000) CIFSN (00) LOGRECNO (0000001) Remainder of geographic header record for geographic area x</p> <p><i>Record 2</i> FILEID (uSF2) STUSAB (AL) CHARITER (000) CIFSN (00) LOGRECNO (0000002) Remainder of geographic header record for geographic area y</p>	<p>FILEID (uSF2) STUSAB (AL) CHARITER (004) CIFSN (01) LOGRECNO (0000001) Tables PCT37 (49 cells)</p> <p>FILEID (uSF2) STUSAB (AL) CHARITER (0041) CIFSN (01) LOGRECNO (0000002) Tables PCT37 (49 cells)</p>

File Record Layout

- For a layout of the table in each file, see the following:

PCT37. SEX BY AGE [49]

Universe: Population in households

Total:

Male:

Under 5 years
5 to 9 years
10 to 14 years
15 to 17 years
18 and 19 years
20 years
21 years
22 to 24 years
25 to 29 years
30 to 34 years
35 to 39 years
40 to 44 years
45 to 49 years
50 to 54 years
55 to 59 years
60 and 61 years
62 to 64 years
65 and 66 years
67 to 69 years
70 to 74 years
75 to 79 years
80 to 84 years
85 years and over

Female:

(Repeat AGE)

FTP File Transfer

- Summary File 2 (SF2) Supplement FTP directory is at ftp2.census.gov/census_2000/datasets/Summary_File_2_Supplement.
 - Each state directory provides all files available for the identified state.
 - The directory for each state has a geographic header file and four files for each iteration (total population, race, tribe, or Hispanic/Latino group) having a population of 100 or more.
 - There is a potential (but not likely) for 251 files for each state—one geographic header file and one data for each of the 250 iterations. This occurs only when each of the 250 groups has a population of 100 or more.
 - The chart on page 3 of this document lists the table number available in the file. The file record layout is shown on page 4 of this document.
 - The files are in a flat ASCII format.
 - No software is provided.
- Downloading multiple files
 - UNIX environment—“mget” subcommand allows transfer of multiple files using the wildcard character.
 - Example: ftp> prompt off
ftp>mget ne* (for this example, Nebraska is selected).
 - Windows Environment—many FTP products have been developed which have the capability to download multiple files with a single command.
 - We used the ws_ftp product in testing the download.
 - A demonstration copy is available at <http://www.ipswitch.com/> .
 - An Internet search using the term “download multiple files” yielded other similar products.
 - For step-by-step instructions with screen shots, please see <http://www.census.gov/support/2000/SF2/> ..

Spreadsheet and Data Base Aids

- For step-by-step instructions with screen shots for moving the data and structure to a spreadsheet, please see www.census.gov/support/SF2ASCII.html .
- Structure files in Access97 and other formats are available at <http://www.census.gov/support/2000/SF2/> .
- We are unable to provide one-on-one support for applications of the data to specific spreadsheets or data base software.