

README File for 2010 Census Summary File 1 Delivered via FTP

Note: *Users processing these FTP files in a Windows environment should read carefully the File Information section of this document.*

Contents

About the FTP Application

The FTP Directory

File Information

Segmented Data

File Record Layout

About the FTP Application

This FTP (File Transfer Protocol) application is intended for experienced users of census data, zip files, and spreadsheet/database software. It provides quick access for data users such as State Data Centers and news media needing to begin their analysis immediately upon data release. Due to the size of the files, the FTP user should have a fast file transfer capability. See Table 1. Estimated File Sizes on pages 4-5.

The FTP Directory

The FTP directory is at http://www2.census.gov/census_2010/04-Summary_File_1. When the 2010 Census Summary File 1 are added to each state's directory, it will contain a single zip package with the geoheader file, 47 data segments and a packing list which provides information about the file's creation and size. See below for more information on Segmented Data.

Users of the FTP application need to unzip the package after downloading, then import the data into the spreadsheet/database software of their choice for data analysis and table presentation. *We are unable to provide one-on-one support for applications of the data to specific spreadsheets or data base software.*

File Information

Once the package is unzipped, the files are in flat ASCII format. The geographic header file contains fixed fields while the data files are in comma-delimited format. These are text files however the file extension is not '.txt'. The user will need to rename the files with a .txt extension for import into some software packages, e.g. Microsoft Access. These files have been constructed in a LINUX environment. They use an ASCII linefeed, chr(10), to indicate a new record.

For successful use with many programs running in a Windows environment, these files need to be modified to use the ASCII carriage return/linefeed sequence, chr(13) + chr(10) as a record terminator. This is an easy step in the UnZIP process using any UnZIP software which offers the conversion option. Winzip for Windows, version 14.5 has been tested.

The resulting file will meet the ANSI MS-DOS/Windows standard used by Access 2003 and Access 2007 and other MS Windows-based programs. If the data are being processed in a LINUX environment, they can be unzipped using any standard Linux ZIP/UnZIP package.

Note to Users of Microsoft Access:

Due to the FieldSize property limitations within Microsoft Access, modifications to field types are required when importing the Geographic Header Record file and File01 through File47:

- Fields classified as numeric (N) should be imported as long integers unless the field has been described as having decimals. These fields should be imported as double.
- AREALAND and AREAWATR should be imported as text.
- Nine (9) data segments between File01 and File47 contain more than 255 fields. See highlighted segments in Table 2 below. Also see Access database structure and instructions (0HowToUseMSAccessWithSummaryFile1.pdf) for converting the data at http://www2.census.gov/census_2010/04-Summary_File_1.
- Failure to make these changes may result in missing data upon import.

Segmented Data

The data in the 2010 Census Summary File 1 and other 2010 Census summary files are segmented. This is done to manage the volume of data and to facilitate exporting into spreadsheet or database software. The data and the corresponding geographic information for an individual state are known as the file set. Because of the large size of the tables, the file set will be broken into 49 files: a geographic header record file and 47 data segment files and a packing list. To get a complete set of the 2010 Census Summary File 1, users must download the geographic header file and all the data file segments in the package.

Table 2 below provides the file/table details. When using Table 2 you must take into consideration that the first 5 fields of the data file contains identification information (FILEID, STUSAB, CHARITER, CIFSNO, and LOGRECNO) which makes some segments have more than 255 fields. These segments are highlighted in Table 2. Some manipulation of the files will be required for software with a field limit of 255.

Table 2. File/Table Segmentation

| Data file segment number | Number of data cells | Starting matrix number | Ending matrix number |
|--------------------------|----------------------|------------------------|----------------------|
| 1 | 1 | P1 | P1 |
| 2 | 6 | P2 | P2 |
| 3 | 194 | P3 | P9 |
| 4 | 239 | P10 | P14 |
| 5 | 245 | P15 | P30 |
| 6 | 254 | P31 | P49 |
| 7 | 251 | P50 | P12E |
| 8 | 253 | P12F | P17A |
| 9 | 249 | P17B | P28I |
| 10 | 252 | P29A | P29I |
| 11 | 254 | P31A | P34E |
| 12 | 251 | P34F | P38E |
| 13 | 240 | P38F | P39H |
| 14 | 20 | P39I | P39I |
| 15 | 251 | PCT1 | PCT8 |
| 16 | 59 | PCT9 | PCT11 |
| 17 | 209 | PCT12 | PCT12 |
| 18 | 188 | PCT13 | PCT20 |
| 19 | 216 | PCT21 | PCT22 |
| 20 | 209 | PCT12A | PCT12A |
| 21 | 209 | PCT12B | PCT12B |
| 22 | 209 | PCT12C | PCT12C |
| 23 | 209 | PCT12D | PCT12D |
| 24 | 209 | PCT12E | PCT12E |
| 25 | 209 | PCT12F | PCT12F |
| 26 | 209 | PCT12G | PCT12G |
| 27 | 209 | PCT12H | PCT12H |
| 28 | 209 | PCT12I | PCT12I |
| 29 | 209 | PCT12J | PCT12J |
| 30 | 209 | PCT12K | PCT12K |
| 31 | 209 | PCT12L | PCT12L |
| 32 | 209 | PCT12M | PCT12M |
| 33 | 209 | PCT12N | PCT12N |
| 34 | 209 | PCT12O | PCT12O |
| 35 | 245 | PCT13A | PCT13E |
| 36 | 245 | PCT13F | PCT19B |
| 37 | 237 | PCT19C | PCT20E |
| 38 | 254 | PCT20F | PCT22F |
| 39 | 63 | PCT22G | PCT22I |
| 40 | 234 | PCO1 | PCO6 |
| 41 | 156 | PCO7 | PCO10 |
| 42 | 1 | H1 | H1 |
| 43 | 6 | H2 | H2 |
| 44 | 249 | H3 | H11F |
| 45 | 255 | H11G | H17C |
| 46 | 126 | H17D | H17I |
| 47 | 74 | HCT1 | HCT4 |

The explanation below for linking the two data files requires specific location information about the geographic header. These are located in Chapter 2-How to use this Product, of the Technical

Documentation <http://www.census.gov/prod/cen2010/doc/sf1.pdf>.

A unique logical record number (LOGRECNO) in the geographic header is assigned to all files for a specific geographic entity; all records for that entity can be linked together across files. The geographic header **record layout** is identical across all electronic data products from the 2010 Census. However, the content is product specific. Some header fields that appear in both file types (geographic header and fileXX) are not used. For example, the characteristic iteration (CHARITER) field will be used in the 2010 Census Summary File 2 data product but in the 2010 Census Summary File 1, it is always coded as 000.

File Record Layout

For a layout of the data table, see <http://www.census.gov/prod/cen2010/doc/sf1.pdf>, select Chapter 6, *Data Dictionary*.

Table 1. Estimated File Sizes

| State | STUSAB | Zipped Estimate Gb | Unzipped Estimate Gb |
|-------|--------|--------------------------|----------------------------|
| 01 | AL | 0.19 | 2.8 |
| 02 | AK | 0.03 | .65 |
| 04 | AZ | 0.20 | 2.7 |
| 05 | AR | 0.25 | 2.2 |
| 06 | CA | 0.95 | 8.8 |
| 08 | CO | 0.17 | 2.3 |
| 09 | CT | 0.10 | .92 |
| 10 | DE | 0.03 | .33 |
| 11 | DC | 0.01 | .11 |
| 12 | FL | 0.53 | 5.7 |
| 13 | GA | 0.38 | 3.4 |
| 15 | HI | 0.04 | .42 |
| 16 | ID | 0.18 | 1.5 |
| 17 | IL | 0.61 | 5.6 |
| 18 | IN | 0.35 | 3.1 |
| 19 | IA | 0.29 | 2.7 |
| 20 | KS | 0.15 | 2.7 |
| 21 | KY | 0.22 | 2.0 |
| 22 | LA | 0.27 | 2.4 |
| 23 | ME | 0.09 | .89 |
| 24 | MD | 0.20 | 1.9 |
| 25 | MA | 0.21 | 1.9 |

| | | | |
|--------------|----|--------------|--------------|
| 26 | MI | 0.44 | 4.1 |
| 27 | MN | 0.35 | 3.3 |
| 28 | MS | 0.13 | 2.0 |
| 29 | MO | 0.45 | 4.0 |
| 30 | MT | 0.16 | 1.4 |
| 31 | NE | 0.10 | 2.1 |
| 32 | NV | 0.11 | 1.0 |
| 33 | NH | 0.06 | .61 |
| 34 | NJ | 0.23 | 2.2 |
| 35 | NM | 0.21 | 1.9 |
| 36 | NY | 0.50 | 4.7 |
| 37 | NC | 0.32 | 3.6 |
| 38 | ND | 0.18 | 1.6 |
| 39 | OH | 0.50 | 4.5 |
| 40 | OK | 0.20 | 3.1 |
| 41 | OR | 0.25 | 2.2 |
| 42 | PA | 0.43 | 5.2 |
| 44 | RI | 0.03 | .35 |
| 45 | SC | 0.24 | 2.1 |
| 46 | SD | 0.12 | 1.2 |
| 47 | TN | 0.32 | 2.8 |
| 48 | TX | 1.16 | 10.1 |
| 49 | UT | 0.15 | 1.3 |
| 50 | VT | 0.04 | .44 |
| 51 | VA | 0.37 | 3.3 |
| 53 | WA | 0.26 | 2.4 |
| 54 | WV | 0.17 | 1.5 |
| 55 | WI | 0.34 | 3.1 |
| 56 | WY | 0.10 | .91 |
| 72 | PR | 0.12 | 1.2 |
| | | | |
| TOTAL | | 13.49 | 133.4 |