

README File for Census 2010 National Summary File of Redistricting Data 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
Other Sources of the Data
Table 1 - Estimated File Sizes

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 below.

The FTP Directory

The FTP directory is at http://www2.census.gov/census_2010/01-Redistricting_File--PL_94-171. When the Census 2010 National Summary File of Redistricting Data are added to its directory, it will contain a single zip package with the three 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. A SAS program and an ACCESS database structure have been made available at http://www2.census.gov/census_2010/01-Redistricting_File--PL_94-171 for converting the data files. 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. Both the geographic header and the data files contain geographic linking fields. 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.

Segmented Data

The data in the National Summary File of Redistricting Data and other Census 2010 summary files are segmented. This is done so that individual files will not have more than 255 fields, facilitating exporting into spreadsheet or database software. In short, **to get a complete set of the National data summary files, users must download all three files in the package.**

The files are as follows:

- Geographic Header file
- File01 (Tables P1 and P2)
- File02 (Tables P3, P4 and H1)

It is easiest to think of the file set as a logical file. This logical file consists of three physical files: the geographic header file, file01, and file02. **This structure is the same as the Census 2000 Decennial data summary files.**

The explanation below for linking the three data files requires specific location information for the geographic header. These are located in Chapter 2 of the Technical Documentation <http://www.census.gov/prod/cen2010/doc/nsfrd.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. Additional identifying fields are also carried over from the geographic header file to the table files. The fields are: file identification (FILEID), state/U.S. abbreviation (STUSAB), and characteristic iteration (CHARITER).

The geographic header record layout is identical across all electronic data products from Census

2010. Since the National Summary File of Redistricting Data is quite simple, some of the fields, including some geographic header fields that appear in all three files (the geographic header, segments 1 and 2) are constant. For example, the characteristic iteration (CHARITER) field will be used in Summary File 2 but in the National File Summary of Redistricting Data, it is always coded as 000.

File Record Layout

For a layout of the individual tables for each file, see <http://www.census.gov/prod/cen2010/doc/nsfrd.pdf>, Select Chapter 6, *Data Dictionary*.

Other Sources of the Data

The Census Bureau releases most Census 2010 data on a state-by-state basis. Tables are available in American FactFinder (factfinder2.census.gov) upon release of the designated state/National data. Within American FactFinder, individual tables can be downloaded in database-compatible or presentation-ready formats. Users have the option to save the file in a comma delimited (.csv), Microsoft Excel (.xls), or PDF format.

Table 1 - Estimated File Sizes

Postal	Name	Zip File (K)	Geoheader (MB)	Segment1 (MB)	Segment2 (MB)
US	National	600	1	1	1