README File for 2010 Census American Indian and Alaska Native (AIAN) Summary File 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.

The FTP Directory

The FTP directory is http://www2.census.gov/census_2010/07-AIAN_Summary_File/National/. The 2010 Census AIAN Summary File is described as a single zip package (estimated 885M) with the geographic header file, 11 data segment files for each of the applicable 1570 iterations totaling as many as 17,270 files, 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 File11:

➢ 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.

➢ LOGRECNO, AREALAND and AREAWATR should be imported as text.

➢ Also see Access database structure and instructions for converting the data at http://www2.census.gov/census_2010/07-AIAN_Summary_File/National/.

➢ Failure to make these changes may result in missing data upon import.

Segmented Data

The data in the 2010 Census AIAN Summary File 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 or national product are known as the file set. Because of the large size of the tables, the file set of the 2010 Census AIAN Summary File will be broken into as many as 17,270 files: a geographic header record file, 11 data segment files for each of the applicable 1570 iterations, and a packing list. To get a complete set of the 2010 Census AIAN Summary File, 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, CIFSN, and LOGRECNO).

Table 2. File/Table Segmentation

Data file segment number	Number of data cells	Starting matrix number	Ending matrix number
1	1	PCT1	PCT1
2	6	PCT2	PCT2
3	212	PCT3	PCT4
4	223	PCT5	PCT18
5	233	PCT19	PCT31
6	124	PCT32	PCT38
7	219	PCT39	PCT47
8	235	PC01	PC07
9	89	PC08	PC10
10	6	HCT1	HCT1
11	172	HCT2	HCT14

The explanation below for linking the two data files require 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/aiansf.pdf.

 A unique logical record number (LOGRECNO) in the geographic header is assigned to all records 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 is used in the 2010 Census AIAN Summary File but not in Summary File 1, it is always coded as 000.

File Record Layout

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