PUMSPrep Quick Reference

Version 4.0.1

Revision History
1/8/2010 Edited by AECOM Consult, Inc.
4/15/2010 Edited by RSG, Inc.

Syntax:

PUMSPrep [-flag] [control_file]

Purpose:

1. Split one or more Census PUMS data files into household and population files;
2. Add Census PUMS data from an additional state or PUMA to existing household and population files;
3. Select PUMAs to include in the output file using a state-PUMA list;
4. Select household records to include in the output file using household type codes;
5. Define household and population data fields to include in the output file; and
6. Use a program script to extract and/or manipulate data fields in the Census PUMS file for output to the household and population files.

Required Keys

PUMS_DATA_FILE (1)	[project_directory]*filename*
NEW_PUMS_HOUSEHOLD_FILE	[project_directory]*filename*
NEW_PUMS_POPULATION_FILE	[project_directory]*filename*
HOUSEHOLD_DATA_FIELD_*	*field_name, type, offset, length* [*, decimal*] (3)
POPULATION_DATA_FIELD_*	*field_name, type, offset, length* [*, decimal*] (4)

Optional Keys

TITLE	Text
REPORT_FILE	*Filename*
REPORT_FLAG	FALSE {true/false/yes/no/1/0}
MAX_WARNING_MESSAGES	100,000
MAX_WARNING_EXIT_FLAG	TRUE {true/false/yes/no/1/0}
PROJECT_DIRECTORY	*Pathname*
DEFAULT_FILE_FORMAT	VERSION3 {(2)}
PUMS_DATA_FILE_* (1)	[project_directory]*filename*
PUMS_HOUSEHOLD_FILE	[project_directory]*filename*
PUMS_POPULATION_FILE	[project_directory]*filename*
STATE_PUMA_LIST	All {SSPPP, …} (5)
SELECT_HOUSEHOLD_TYPES	0 {0, 1, 2, 3} (6)
CONVERSION_SCRIPT (7)	[project_directory]*filename.shp*

PUMS_HOUSEHOLD_FORMAT	VERSION3 {(2)}
PUMS_POPULATION_FORMAT	VERSION3 {(2)}
NEW_PUMS_HOUSEHOLD_FORMAT	VERSION3 {(2)}
NEW_PUMS_POPULATION_FORMAT	VERSION3 {(2)}

Reports

PUMSPREP_REPORT_#	CONVERSION_SCRIPT
	CONVERSION_STACK

Notes

1	At least one PUMS data file is required. The PUMS_DATA_FILE and/or PUMS_DATA_FILE_* key can be used.
2	{VERSION3, BINARY, FIXED_COLUMN, COMMA_DELIMITED, SPACE_DELIMITED, TAB_DELIMITED, CSV_DELIMITED, DBASE, LANL, SQLITE3}
3	This key defines new fields to add to the household file. The values assigned to these fields are set equal to a field in a PUMS household record or are initialized to zero or blank and set using a conversion script. The key must include at least four comma separated values. The first is the field name. This is followed by the field type code for an integer (I, INTEGER), floating point (R, REAL, D, DOUBLE), or string (S, STRING, C, CHARACTER). The third field is the column number in the PUMS household record were the associated data is found. If this column number is zero, the field is created and initialized to zero or blank. The fourth field is the number of characters in the PUMS data field (or the output field length if the column number is zero). If the field is a floating point number, a fifth field can be provided to identify the number of decimal places in the data. The default value is zero.
4	The same field definition codes as outlined above are used to define population fields using data from the PUMS person records.
5	The state FIPS code followed by a PUMA code. For example, the code for PUMA in Alexandria (PUMA=200), Virginia (FIPS=51) is 51200.
6	Zero selects all household types. 1=family households, 2=non-family households, 3=group quarter households.
7	The conversion script key is a file name that includes a TRANSIMS User Program script. Any household or population field defined for the output file can be referenced using the file label New_HH or New_Pop. All fields are zeroed when a new household record is read from the PUMS file. The program is executed for the household record and each person record within the household. The household fields are not written to the output file until all of the person records for the household are read and processed. This enables the script to set household attributes based on calculations done with the person records (e.g., number of workers, age of the householder, etc.)