

Census 2000 Special EEO Tabulation Technical Documentation for SAS Datasets

Appendix 4.0: Using Appendices 4.1 to 4.7C

The Census 2000 Special EEO Tabulation is composed of three types of SAS datasets: (1) Residence datasets (Dataset 1 to Dataset 7); (2) Work Site/Residence (Work Flow) datasets (Dataset 8 to Dataset 19); and (3) Work Site Only datasets (Dataset 20 to Dataset 24). Each type has two versions: One for the United States, excluding Hawaii, and one for Hawaii only (except for dataset 20). See Appendix 4.3 for the contents of all these SAS datasets.

The Residence and Work Site Only datasets have the same basic structure, which differs from the Work Site/Residence (Work Flow) datasets. To process the Residence and Work Site Only datasets correctly, the user must understand the purpose of the Summary Level variable (**sumlevel**) and the Combined Geographic Area Code(s) variable (**geocode**). The values of **sumlevel** are three-digit codes that identify the types of geographic areas presented in a dataset (see Appendices 4.6A and 4.6B for a list of summary levels in each dataset). These codes, which are assigned by the Census Bureau, allow the user to subset the dataset into various geographic summary level groups (010 = United States, 040 = States, 050 = Counties, etc).

The value of the **geocode** variable, on the other hand, is a single geographic area code (FIPS State, FIPS Metropolitan Area, County Set Area, etc.) or a combined geographic area code (FIPS State and FIPS County, FIPS State and FIPS Place or FIPS State and Census Place) that identifies a unique geographic area within a particular summary level. For example, a state-level observation in Table1.sas7bdat will have a value of 040 for **sumlevel** and a FIPS State code of 01 to 56, excluding 15 (Hawaii) for **geocode**. The FIPS State code scheme does not use the values of 03, 07, 14, 43, and 52.

The **sumlevel** and **geocode** variables are not on any of the Work Site/Residence (Work Flow) datasets, which are organized by Place of Work Area/Residence Area flows rather than by summary levels. Every observation in each of the Work Site/Residence (Work Flow) datasets has two geographic code variables: one for the place of work component of the flow and the other for the residence component of the flow.

Appendix 4.1 to Appendix 4.7C provide information on all the SAS datasets in the Census 2000 Special EEO Tabulations.