

(This file based on a document provided by the U.S. Census Bureau in October 2008, and edited by Lars Vilhuber for the VirtualRDC on 2008-11-17)

Zero obs data sets

The “zero obs” files are SAS® datasets that contain no records. Their structure, however, corresponds to the internal historical decennial census microdata files available in the Census Bureau’s national network of Research Data Centers (RDC). Researchers can use the zero obs files to test programs before entering an RDC, although these programs would yield no output because the zero obs files contain no data. The structure of the VirtualRDC replicates the exact directory and filename structure of the Census RDC, allowing for program sequences to be developed and easily ported to the restricted environment of the RDCs.

Also, researchers can produce PROC CONTENTS of the zero obs files to obtain a complete list of all variables in the internal historical decennial census microdata files available to researchers. Doing so will provide a list of all variables within each dataset, including the variable name, type, length, label and the position of the variable within the records. For more detailed documentation on the internal versions of the historical decennial census microdata files, please refer to the documentation on the ICSPR website at <http://www.icpsr.umich.edu/>.

Availability

Historical decennial census microdata files are available for the 1970, 1980, 1990 and 2000 censuses, but zero obs files are currently available only for the 1970, 1980 and 1990 data. The datasets available for each decennial census year are the 100 percent (or “short form”) microdata files and the sample (or “long form”) microdata files. Each of these datasets is divided into separate files for each record type: geographic, household and person records.

File names

The naming convention for the (zero obs and actual) files is

`t13[state][census year][dataset][record type].sas7bdat`

Where

state is the FIPS/Postal abbreviation for the state (and 'us' is used for national files;

census year is given in four digits

- 1970
- 1980
- 1990

dataset is represented by a single character

- h - 100 percent files
- s - sample files

record type is represented by a single character

- g - geographic record
- h - household record
- p - person record

For example, the file for the person records for New York from the 1980 sample file is `t13ny1980sp.sas7bdat`

Directory structure

```
/decennial/  
  [census year]cen/  
    [census year][RECORD TYPE]/  
      microdata/
```

where [census year] and [record type] are defined as above (record type is in CAPITAL LETTERS). For example, the above NY 1980 sample file for person records can be found at

```
/decennial/1980cen/1980S/microdata/t13ny1980sp.sas7bdat
```

Sample code

The following code is a sample setup for a program that wishes to analyze the sample files for Florida, New York, Oregon, and Tennessee from the 1980 person files:

```
libname cen1980s '/decennial/1980cen/1980S/microdata';  
data cen1980;  
set cen1980s.t13fl1980sp  
    cen1980s.t13ny1980sp  
    cen1980s.t13or1980sp  
    cen1980s.t13tn1980sp  
;  
run;
```

The code will run without error on the VirtualRDC, yielding a well-defined (but empty) file 'WORK.CEN1980' which could be used in further 'analysis'.