

## **Documentation for Vintage 2006 Bridged-Race Postcensal Population Estimates for Calculating Vital Rates**

The Vintage 2006 bridged-race postcensal population estimates file contains estimates of the resident population of the United States as of July 1, 2000 to July 1, 2006 by county, single-year of age (0, 1, 2,..., 85 years and over), bridged-race category (White, Black or African American, American Indian or Alaska Native, Asian or Pacific Islander), Hispanic origin (not Hispanic or Latino, Hispanic or Latino), and sex (1). One set of files has national estimates for persons 85 years and over by single-year of age. The estimates on these files are based on Census 2000 and result from bridging the Vintage 2006 postcensal estimates with 31 race groups (the 31 race groups used in Census 2000 in accordance with the 1997 Office of Management and Budget (OMB) standards for the collection of data on race and ethnicity) to the four race categories specified under the 1977 OMB standards (2, 3). The bridged-race postcensal estimates were produced by the Population Estimates Program of the U.S. Census Bureau in collaboration with the National Center for Health Statistics (NCHS). This file was released by NCHS on August 16, 2007.

### **Background**

Race bridging refers to making data collected using one set of race categories consistent with data collected using a different set of race categories, to permit estimation and comparison of race-specific statistics at a point in time or over time. More specifically, race bridging is a method used to make multiple-race and single-race data collection systems sufficiently comparable to permit estimation and analysis of race-specific statistics.

In 1997, OMB issued "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity," which supersedes the 1977 Statistical Policy Directive 15, "Race and Ethnic Standards for Federal Statistics and Administrative Reporting" (2,3). Both documents specify rules for the collection, tabulation, and presentation of race and ethnicity data within the Federal statistical system. The race categories specified in both standards represent a social-political construct and are not anthropologically or biologically based. The revised standards increased the minimum number of race categories to be used by Federal agencies from four (White, Black, American Indian or Alaska Native, and Asian or Pacific Islander) to five (White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander). In addition, the revised standards require Federal data collection programs to allow respondents to select more than one race category when responding to a query on their racial identity. This provision means that under the revised standards there are potentially 31 race groups (5 single-race and 26 multiple-race), depending on whether an individual selects one, two, three, four, or all five of the race categories.

During the transition to full implementation of the 1997 standards (see paragraph below), these two different standards for the collection of race and ethnicity data are both being used, creating incomparability across data systems. Further, within a given data system, the change in the race standards results in incomparability across time, thus making it difficult to perform trend analyses. The OMB recognized that race-bridging approaches would be needed to make race data collected under the 1997 standards comparable to race data collected under the 1977 standards. Therefore, the OMB issued “Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity” (4). The guidance document contains a detailed discussion of various bridging methods.

Vital rates are based on information obtained from vital records collected through the state-based Vital Statistics Cooperative Program (numerators) and population estimates derived from the U.S. Census (denominators). Because of differences in the timing of implementation of the 1997 standards, beginning with the 2000 data year, the numerators and denominators of vital rates have incompatible race data. The question on race on the 2000 census was based on the revised OMB standards and so allowed respondents to select more than one race category. As a result, population estimates for 2000 and beyond have five single-race categories and up to 26 multiple-race categories. Implementation of the 1997 standards within the Vital Statistics Cooperative System started in 2003, on an individual state basis, and is expected to proceed slowly as states implement revised birth and death certificates which incorporate the revised OMB standards. As of 2006, most states had not revised the race question on their birth or death certificates and were still collecting race data using the 1977 race categories. Therefore, at this time, the calculation of post-2000 race-specific birth and death rates requires population estimates with the 1977 race categories. When a sufficient number of states have adopted the revised birth and death certificates, rates will be presented using population estimates that comply with the 1997 standards.)

### **NCHS Regression Bridging Method**

Bridging methodology developed by NCHS bridges the multiple-race group population counts to four single-race categories (5,6). Information from the pooled 1997-2000 National Health Interview Survey was used to develop the bridging methodology. Regression models with person-level and county-level covariates were used to generate the probability of selecting each single-race category possible for a multiple-race group. The probabilities generated from the fitted regression models are referred to as the NHIS bridging proportions. The Census Bureau applied the NHIS bridging proportions generated by NCHS to the Census 2000 Modified Race Data Summary file (7). This application resulted in a bridged population count for each of the four single-race categories (White, Black or African American, American Indian or Alaska Native, and Asian or Pacific Islander) by county, single-year of age, Hispanic origin group, and sex, for April 1, 2000. Annually, the bridging proportions are applied to the latest vintage of postcensal estimates with 31 race groups (the 31 race groups used in Census 2000 in accordance with the 1997 OMB standards) to obtain the

bridged-race population estimates for the four single-race categories. The bridging methodology is described in detail in the report, “United States Census 2000 Population with Bridged Race Categories” (which is available for download from this site) (5).

### **Bridged-race postcensal population estimates**

Postcensal population estimates are estimates made for the years following a census, before the next census has been taken. Postcensal estimates are derived by updating the resident population enumerated in the decennial census using various measures of population change. The components of population change used in the derivation of the postcensal estimates include: births to U.S. resident women, deaths to U.S. residents, net international immigration, net movement of U.S. Armed Forces and civilian citizens of the U.S, and migration within the U.S. The Census Bureau annually produces a series of postcensal estimates that includes estimates for the current data year and revised estimates for earlier years. Estimates for earlier years in a given series are revised to reflect changes in the components of population change data sets (for example, a preliminary natality file is replaced with a final natality file). The last year in a series is used to name the series. For example, the Vintage 2002 postcensal series has estimates for July 1, 2000, July 1, 2001, and July 1, 2002. The Vintage 2003 series has estimates for July 1, 2000, July 1, 2001, July 1, 2002, and July 1, 2003. The July 1, 2000, July 1, 2001, and July 1, 2002 estimates from the Vintage 2002 and Vintage 2003 series differ.

The Census bureau annually produces a postcensal series of estimates of the resident population of the United States with 31 race groups (in accordance with the 1997 OMB standards). The Census 2000 Modified Race Data Summary File serves as the base data for these post-2000 postcensal series (7). These series initially had estimates for 31 race groups, in accordance with the 1997 race and ethnicity standards (2). Under a collaborative arrangement with NCHS, the Population Estimates Program of the U.S. Census Bureau applies the NHIS bridging proportions to the 31-race postcensal population estimates to produce bridged-race postcensal estimates (estimates for the four single-race categories: White, Black or African American, American Indian or Alaska Native, and Asian or Pacific Islander).

### **Variance of bridged-race population estimates**

Population estimates generally are assumed to be fixed and do not contribute to the variance of rates. However, this is not true for bridged-race population estimates. Methodology to compute variances for bridged-race population estimates has been developed (8).

### **Geographic codes in the bridged-race population files**

The Vintage 2006 files have population estimates for 3,141 counties and county equivalents. The State and county FIPS codes in the Vintage 2006 files are the same as those in the Vintage 2002 and Vintage 2005 files. The Vintage 2003 and Vintage 2004 files do not have population estimates for Broomfield County, Colorado; the Vintage 2002, Vintage 2005, and Vintage 2006 files do.

**Broomfield County, Colorado** (FIPS code=08014) was created effective November 15, 2001 from parts of four Colorado counties: Adams, Boulder, Jefferson, and Weld. Vintage 2003 and Vintage 2004 did not contain estimates for Broomfield. Note that data for Broomfield County does not appear on NCHS birth or mortality files until data year 2003.

There are no estimates for **Clifton Forge County, Virginia** (FIPS code=51560) on the bridged-race postcensal population files. On July 1, 2001, Clifton Forge city, Virginia, formerly an independent city, merged with Alleghany county (FIPS code=51005). Note that estimates for Clifton Forge city appear on NCHS birth and mortality files prior to data year 2003; beginning with the 2003 data year, no data for Clifton Forge city appear on the birth and death files.

### **Release of bridged-race population estimates**

In response to the need for bridged estimates by a wide range of users, NCHS is making the bridged-race population estimates available for download from this web site (see **Data Files and Documentation**). The report detailing the bridging methodology is available for download from this site (see "United States Census 2000 Population with Bridged Race Categories" under **Methodology**).

NCHS is currently using the bridged-race postcensal population estimates to calculate birth and death rates. Previously published reports that used 1990-based postcensal population estimates to calculate rates for 2001 have been re-issued in whole or in part; new reports use the bridged-race estimates.

Although efforts were made to use the best available data and methods to produce these estimates, the modeling process introduces error into the estimates. The potential for error will be greatest for the smallest population groups, particularly the smaller race groups and county level estimates.

NCHS would appreciate receiving feedback on the usefulness of the estimates as well as notification of any problems that have been identified. Please provide comments via e-mail to: [PopEst@cdc.gov](mailto:PopEst@cdc.gov).

### **Suggested citation**

National Center for Health Statistics. Postcensal estimates of the resident population of the United States for July 1, 2000-July 1, 2006, by year, county, age, bridged race,

Hispanic origin, and sex (Vintage 2006). Prepared under a collaborative arrangement with the U.S. Census Bureau. Available on the Internet from: [http://www.cdc.gov/nchs/nvss/bridged\\_race.htm](http://www.cdc.gov/nchs/nvss/bridged_race.htm). August 16, 2007.

## References

1. National Center for Health Statistics. Postcensal estimates of the resident population of the United States for July 1, 2000-July 1, 2006, by year, county, age, bridged race, Hispanic origin, and sex (Vintage 2006). Prepared under a collaborative arrangement with the U.S. Census Bureau. Available from: [http://www.cdc.gov/nchs/nvss/bridged\\_race.htm](http://www.cdc.gov/nchs/nvss/bridged_race.htm). August 16, 2007.
2. Office of Management and Budget. Revisions to the standards for the classification of Federal data on race and ethnicity. Federal Register 62FR58781-58790, October 30, 1997. Available from: <http://www.whitehouse.gov/omb/fedreg/1997standards.html>.
3. Office of Management and Budget. Race and ethnic standards for Federal statistics and administrative reporting. Statistical Policy Directive 15. May 12, 1977.
4. Office of Management and Budget. Provisional guidance on the implementation of the 1997 standards for the collection of Federal data on race and ethnicity. December 15, 2000. Available from: [http://www.Whitehouse.gov/omb/inforegr/r&e\\_guidance2000update.pdf](http://www.Whitehouse.gov/omb/inforegr/r&e_guidance2000update.pdf).
5. Ingram DD, Parker JD, Schenker N, Weed JA, Hamilton B, Arias E, Madans JH. United States Census 2000 population with bridged race categories. Vital Health Stat 2(135). Hyattsville, Maryland: National Center for Health Statistics. 2003. Available from: [http://www.cdc.gov/nchs/data/series/sr\\_02/sr02\\_135.pdf](http://www.cdc.gov/nchs/data/series/sr_02/sr02_135.pdf)
6. Parker JD, Schenker N, Ingram DD, Weed JA, Heck KE, Madans JH. Bridging between two standards for collecting information on race and ethnicity: an application to Census 2000 and vital rates. Public Health Reports 119(2): 192-205. 2004
7. U.S. Census Bureau. Census 2000 Modified Race Data [MR(31)-CO.txt], prepared by the U.S. Census Bureau, 2002. Available on the internet from: <http://www.census.gov/popest/archives/files/MR-CO.txt>.
8. Schenker, N. Assessing variability due to race bridging: application to Census counts and vital rates for the Year 2000, J American Statistical Association, 98, 818-28. 2003.

## File layout for the Vintage 2006 Bridged-Race Postcensal Files, 2000-2006

### Estimates for persons 0 to 85+ years of age (0,1, ..., 85 years and over)

There is one file for the full Vintage 2006 postcensal series with data for all seven years in the series (July 1, 2000 to July 1, 2006) and two files with only the Vintage 2006 estimates for July 1, 2006. The files contain bridged-race postcensal estimates of the July 1 resident population of the United States by year, county, age (0, 1, ..., 85 years and over), bridged-race category (White, Black or African American, American Indian or Alaska Native, Asian or Pacific Islander), Hispanic origin (not Hispanic or Latino, Hispanic or Latino), and sex. There is a record on the files for each combination of county, age, race, sex, and Hispanic origin. The population estimates on the files were derived by the U.S. Census Bureau using the Census 2000 Modified Race Data Summary File as the base file (9). The files were released by NCHS on August 16, 2007.

### Estimates for persons 85 to 100+ years of age (85, 86, ... 100 years and over)

The Vintage 2006 estimate series (April 1, 2000, July 1, 2000-July 1, 2006) is available as text and SAS files (all years in the series are contained on the same file). The files contain bridged-race postcensal estimates of the resident population of the United States by year, age (85, 86, ..., 100 years and over), bridged-race category (White, Black or African American, American Indian or Alaska Native, Asian or Pacific Islander), Hispanic origin (not Hispanic or Latino, Hispanic or Latino), and sex. There is a record on each file for each combination of month, year, age, race, Hispanic origin, and sex. The population estimates on the files were derived by the U.S. Census Bureau using the Census 2000 Modified Race Data Summary File as the base file (13). This file was released by NCHS on May 6, 2011.

#### Control totals for Vintage 2006 data files

File name	Number of records <sup>1</sup>	Month, Year	Population count
Pcen_v2006.txt	4,322,016		
		July 1, 2000	282,216,952
		July 1, 2001	285,226,284
		July 1, 2002	288,125,973
		July 1, 2003	290,796,023
		July 1, 2004	293,638,158
		July 1, 2005	296,507,061
		July 1, 2006	299,398,484
Pcen_v2006_y06.txt	4,322,016	July 1, 2006	299,398,484
Pcen_v2006_y06.sas7bdat	<u>4,322,016</u>	July 1, 2006	299,398,484
pcen_v2006_85to100.txt	2,048	April 1, 2000	2,239,674
pcen_v2006_85to100.sas7bdat		July 1, 2000	4,286,137
		July 1, 2001	4,417,646
		July 1, 2002	4,546,944
		July 1, 2003	4,715,738

		July 1, 2004	4,848,477
		July 1, 2005	5,077,302
		July 1, 2006	5,296,817

For the files with estimates for persons aged 0 to 85+ there is one record for each county, race, sex, Hispanic origin, and age combination. For the files with estimates for persons aged 85 to 100+ there is one record for each race, sex, Hispanic origin, and age combination.

File layout for text file with estimates for persons aged 0 to 85 years and over, estimate years 2000 to 2006:

**pcen\_v2006.txt:**

Location	Field size	Item and Code Outline	Format
1-4	4	<u>Series vintage</u> (2006)	Numeric
5-6	2	<u>FIPS State code</u>	Numeric
7-9	3	<u>FIPS county code</u>	Numeric
10-11	2	<u>Age</u> (0, 1, 2,..., 85 years and over)	Numeric
12	1	<u>Race-sex</u> 1=White male 2=White female 3=Black or African American male 4=Black or African American female 5=American Indian or Alaska Native male 6=American Indian or Alaska Native female 7=Asian or Pacific Islander male 8=Asian or Pacific Islander female	Numeric
13	1	<u>Hispanic origin</u> 1=not Hispanic or Latino 2=Hispanic or Latino	Numeric
14-21	8	<u>Population count for July 1, 2000</u>	Numeric
22-29	8	<u>Population count for July 1, 2001</u>	Numeric
30-37	8	<u>Population count for July 1, 2002</u>	Numeric
38-45	8	<u>Population count for July 1, 2003</u>	Numeric
46-53	8	<u>Population count for July 1, 2004</u>	Numeric
54-61	8	<u>Population count for July 1, 2005</u>	Numeric
62-69	8	<u>Population count for July 1, 2006</u>	Numeric



File layout for text file with estimates for persons aged 0 to 85 years and over, 2006  
 estimate year: **pcen\_v2006\_y06.txt:**

Location	Field size	Item and Code Outline	Format
1-4	4	<u>Series vintage</u> (2006)	Numeric
5-6	2	<u>FIPS State code</u>	Numeric
7-9	3	<u>FIPS county code</u>	Numeric
10-11	2	<u>Age</u> (0, 1, 2, ..., 85 years and over)	Numeric
12	1	<u>Race-sex</u> 1=White male 2=White female 3=Black or African American male 4=Black or African American female 5=American Indian or Alaska Native male 6=American Indian or Alaska Native female 7=Asian or Pacific Islander male 8=Asian or Pacific Islander female	Numeric
13	1	<u>Hispanic origin</u> 1=not Hispanic or Latino 2=Hispanic or Latino	Numeric
14-21	8	<u>Population count for July 1, 2006</u>	Numeric

File layout for SAS files with estimates for persons aged 0 to 85 years and over, 2006  
 estimate year: **pcen\_v2006\_y06.sas7bdat**

Variable	Item and code outline	Format
VINTAGE	Series Vintage (2006)	Numeric
ST_FIPS	State FIPS code	Numeric
CO_FIPS	County FIPS code	Numeric
AGE	<u>Age</u> (0, 1, 2, ..., 85 years and over)	Numeric
RACESEX	<u>Race</u> 1=White male 2=White female 3=Black male 4=Black female 5=American Indian or Alaska Native male 6=American Indian or Alaska Native female 7=Asian or Pacific Islander male 8=Asian or Pacific Islander female	Numeric
HISP	<u>Hispanic origin</u> 1=not Hispanic or Latino 2=Hispanic or Latino	Numeric
Pop	July 1, 2006 resident population	Numeric

File layout for text file with estimates for persons aged 85 to 100 years and over:  
pcen\_v2006\_85to100.txt

Location	Field size	Item and code outline	Format
1-4	4	<u>Series vintage</u> (2006)	Numeric
5-8	4	<u>Estimate year</u> 2000... or 2006	Numeric
9	1.	<u>Estimate month</u> 4=April 7=July	Numeric
10-12	3	<u>Age</u> (85, 86... 100 years and over)	Numeric
13	1	<u>Race-sex</u> 1=White male 2=White female 3=Black or African American male 4=Black or African American female 5=American Indian or Alaska Native male 6=American Indian or Alaska Native female 7=Asian or Pacific Islander male 8=Asian or Pacific Islander female	Numeric
14	1	<u>Hispanic origin</u> 1=not Hispanic or Latino 2=Hispanic or Latino	Numeric
15-22	8	<u>Population estimate</u>	Numeric

---

File layout for SAS file with estimates for persons aged 85 to 100 years and over:  
**pcen\_v2006\_85to100.sas7bdat**

---

Variable name	Item and code outline	Format
VINTAGE	Series Vintage (2006)	Numeric
YEAR	<u>Estimate year</u> (2000... or 2006)	Numeric
MONTH	<u>Estimate month</u> 4=April 7=July	Numeric
AGE	<u>Age</u> (85, 86... 100 years and over)	Numeric
RACESEX	<u>Race-sex</u> 1=White male 2=White female 3=Black male 4=Black female 5=American Indian or Alaska Native male 6=American Indian or Alaska Native female 7=Asian or Pacific Islander male 8=Asian or Pacific Islander female	Numeric
HISP	<u>Hispanic origin</u> 1=not Hispanic or Latino 2=Hispanic or Latino	Numeric
Pop	<u>Population estimate</u>	Numeric

---