NOTES AND ERRATA

2000

2000 Census of Population and Housing

SF/01-ER

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

United States Census 2000

CONTENTS

Program Overview

The Count Question Resolution (CQR) program handles external challenges to particular official Census 2000 counts of housing units and group quarters population received from state, local, or tribal officials of governmental entities or their designated representatives. For more information about the program, see the Count Question Resolution web page.

The corrected CQR counts will be reflected on a flow basis in the base for population (intercensal) estimates that will be released beginning in December 2002. An inventory of the corrections will be available on American FactFinder, but the base files for the census will remain unrevised, so that none of the standard Census 2000 data products will reflect the corrections.

Corrected Census 2000 Total Population, Group Quarters Population, Total Housing Unit, and Vacant Housing Unit Counts for the United States and Puerto Rico

Note: Corrected counts are a result of the Count Question Resolution (CQR) Program.

Federal Register/Vol. 66, No. 130/Friday, July 6, 2001 (Summary): *The CQR program is not a mechanism or process to challenge the March 6, 2001, decision of the Secretary of Commerce to release unadjusted numbers from Census 2000 for redistricting purposes; nor is it a mechanism or process to challenge or revise the numbers sent to the President on December 28, 2000, to be used to apportion the U.S. House of Representatives.*

This table presents census counts only when there is a CQR change that affects the state or Puerto Rico level. States with no CQR changes (and no tables) are Maine, New Hampshire, and Rhode Island. Use GU link to get to governmental units pages, or T/B link to get to tract/block listings.

Last updated 9/10/2004 — Number of governmental units affected by CQR = 1175

United States State and Puerto Rico			State FIPS	2000 Census Counts							
				Tabulation (Original)				Corrected (Revised)			
				Total Population	Group Quarters Population	Total Housing Units	Vacant Housing Units	Total Population	Group Quarters Population	Total Housing Units	Vacant Housing Units
United States				**281 421 906**	**7 778 633**	**115 904 641**	**10 424 540**	**281 424 603**	**7 780 475**	**115 904 472**	**10 424 715**
STATE											
Alabama	GU	T/B	01	4 447 100	114 720	1 963 711	226 631	4 447 351	114 720	1 963 834	226 662
Alaska	GU	T/B	02	626 932	19 349	260 978	39 378	626 931	19 349	260 963	39 364
Arizona	GU	T/B	04								
Arkansas	GU	T/B	05								
California	GU	T/B	06	33 871 648	819 754	12 214 549	711 679	33 871 653	819 757	12 214 550	711 679
Colorado	GU	T/B	08	4 301 261	102 955	1 808 037	149 799	4 302 015	102 955	1 808 358	149 819
Connecticut	GU	T/B	09	3 405 565	107 939	1 385 975	84 305	3 405 602	107 939	1 385 997	84 312
Delaware	GU	T/B	10								
District of Columbia	GU	T/B	11								
Florida	GU	T/B	12	15 982 378	388 945	7 302 947	965 018	15 982 824	388 945	7 303 108	965 033
Georgia	GU	T/B	13	8 186 453	233 822	3 281 737	275 368	8 186 816	233 822	3 281 866	275 381
Hawaii	GU	T/B	15								
Idaho	GU	T/B	16	1 293 953	31 496	527 824	58 179	1 293 956	31 496	527 825	58 179
Illinois	GU	T/B	17	12 419 293	321 781	4 885 615	293 836	12 419 647	321 773	4 885 744	293 845
Indiana	GU	T/B	18	6 080 485	178 154	2 532 319	196 013	6 080 517	178 154	2 532 327	196 013
Iowa	GU	T/B	19	2 926 324	104 169	1 232 511	83 235	2 926 382	104 169	1 232 530	83 237
Kansas	GU	T/B	20	2 688 418	81 950	1 131 200	93 309	2 688 824	81 950	1 131 395	93 332
Kentucky	GU	T/B	21	4 041 769	114 804	1 750 927	160 280	4 042 285	114 804	1 751 118	160 285
Louisiana	GU	T/B	22	4 468 976	135 965	1 847 181	191 128	4 468 958	135 965	1 847 174	191 129
Maine	GU	T/B	23								
Maryland	GU	T/B	24	5 296 486	134 056	2 145 283	164 424	5 296 507	134 056	2 145 290	164 424
Massachusetts	GU	T/B	25	6 349 097	221 216	2 621 989	178 409	6 349 105	221 216	2 621 993	178 410
Michigan	GU	T/B	26	9 938 444	249 889	4 234 279	448 618	9 938 480	249 889	4 234 252	448 625
Minnesota	GU	T/B	27	4 919 479	135 883	2 065 946	170 819	4 919 492	135 882	2 065 952	170 819
Mississippi	GU	T/B	28	2 844 658	95 414	1 161 953	115 519	2 844 656	95 414	1 161 952	115 519
Missouri	GU	T/B	29	5 595 211	162 058	2 442 017	247 423	5 596 683	163 534	2 442 003	247 425
Montana	GU	T/B	30								
Nebraska	GU	T/B	31	1 711 263	50 818	722 668	56 484	1 711 265	50 818	722 669	56 484
Nevada	GU	T/B	32								
New Hampshire	GU	T/B	33								
New Jersey	GU	T/B	34	8 414 350	194 821	3 310 275	245 630	8 414 347	194 821	3 310 274	245 630
New Mexico	GU	T/B	35								
New York	GU	T/B	36	18 976 457	580 461	7 679 307	622 447	18 976 821	580 833	7 679 307	622 447
North Carolina	GU	T/B	37	8 049 313	253 881	3 523 944	391 931	8 046 485	253 881	3 522 330	391 959
North Dakota	GU	T/B	38	642 200	23 631	289 677	32 525	642 200	23 631	289 678	32 526
Ohio	GU	T/B	39	11 353 140	299 121	4 783 051	337 278	11 353 145	299 121	4 783 066	337 279
Oklahoma	GU	T/B	40	3 450 654	112 375	1 514 400	172 107	3 450 652	112 375	1 514 399	172 107
Oregon	GU	T/B	41	3 421 399	77 491	1 452 709	118 986	3 421 436	77 491	1 452 724	118 986
Pennsylvania	GU	T/B	42	12 281 054	433 301	5 249 750	472 747	12 281 054	433 301	5 249 751	472 748
Rhode Island	GU	T/B	44								
South Carolina	GU	T/B	45	4 012 012	135 037	1 753 670	219 816	4 011 848	135 037	1 753 601	219 817
South Dakota	GU	T/B	46								
Tennessee	GU	T/B	47	5 689 283	147 946	2 439 443	206 538	5 689 267	147 946	2 439 435	206 537
Texas	GU	T/B	48	20 851 820	561 109	8 157 575	764 221	20 851 790	561 109	8 157 557	764 221
Utah	GU	T/B	49	2 233 169	40 480	768 594	67 313	2 233 198	40 480	768 603	67 313
Vermont	GU	T/B	50								
Virginia	GU	T/B	51	7 078 515	231 398	2 904 192	205 019	7 079 030	231 398	2 904 432	205 040
Washington	GU	T/B	53	5 894 121	136 382	2 451 075	179 677	5 894 141	136 382	2 451 081	179 677
West Virginia	GU	T/B	54	1 808 344	43 147	844 623	108 142	1 808 350	43 147	844 626	108 142
Wisconsin	GU	T/B	55	5 363 675	155 958	2 321 144	236 600	5 363 715	155 958	2 321 157	236 601
Wyoming	GU	T/B	56								
Puerto Rico	GU	T/B	72	3 808 610	46 774	1 418 476	157 151	3 808 603	46 774	1 418 474	157 151

American Indian/Alaska Native Areas T/B

- represents zero

Census 2000 Redistricting Data (Public Law 94-171) Summary File

INDEX TO PL 94-171 GEOGRAPHY NOTES

Note	Geographic area
1	Alaska
2	California
3	Connecticut
4	Florida
5	Georgia
6	Nebraska
7	Tennessee
8	Wisconsin

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 1

Alaska: 02

Nelson Lagoon Alaska Native village statistical area (ANVSA) (AIANHH 7025) erroneously contains block 2010, census tract 1 (000100) in Aleutians East census area (01598), Aleutians East Borough (013). This block should have not been coded to any ANVSA (9999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 02-003

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 2

California: 06

Los Angeles city (FIPS code 44000) erroneously contains block 1011, census tract 4002.03 (400203) in East San Gabriel Valley CCD (FIPS code 90810), Los Angeles County (FIPS code 037), CA (FIPS code 06). This block should have been coded to the place Balance of East San Gabriel Valley CCD (FIPS code 99999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP side.

Internal Errata ID 06-001

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 3

Connecticut: 09

The place record, Balance of Milford town (FIPS code 99999) erroneously contains block 2999, census tract 1502 (150200) in Milford town (FIPS code 47535), New Haven County (FIPS code 009), CT (FIPS code 09). This block should have been coded to place Milford city (balance) (FIPS code 47515). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 09-001

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 4

Florida: 12

Yeehaw Junction CDP (FIPS code 78975) in St. Cloud CCD (FIPS code 93029), Osceola County (FIPS code 097), FL (FIPS code 12) should be named Buenaventura Lakes with FIPS code 09415. In 1990, this area was named Buena Ventura Lakes (FIPS code 09415). The area that should have been Yeehaw Junction CDP was erroneously not defined and does not appear in any Census 2000 products.

Internal Errata ID 12-001

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 5

Georgia: 13

The place record Balance of Athens CCD (FIPS code 99999) erroneously contains blocks 2021 and 2023, census tract 1305 (130500) in Athens CCD (FIPS code 90138), Clarke County (FIPS code 059). Both blocks should have been coded to Bogart town (FIPS code 09068).

The place record Balance of Winterville CCD (FIPS code 99999) erroneously contains blocks 1008 and 1009, census tract 1406 (140600) in Winterville CCD (93402), Clarke County (FIPS code 059). Both blocks should have been coded to the place Athens-Clarke County (balance) (FIPS code 03440). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 13-001

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 6

Nebraska: 31

In the PL 94-171 and Summary File (SF) data products, Cisco CDP (FIPS code 09112) in Lisco precinct (FIPS code 91790), Garden County (FIPS code 069), NE (FIPS code 31) should be named Lisco with FIPS code of 28315.

Internal Errata ID 31-002

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 7

Tennessee: 47

The place record Balance of Metropolitan Government CCD (FIPS code 99999) erroneously contains blocks 1001 and 1008, census tract 171 (017100) in Metropolitan Government CCD (FIPS code 92200), Davidson County (FIPS code 037), TN (FIPS code 47). Both blocks should have been coded to place Nashville-Davidson (balance) (FIPS code 52006). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 47-001

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Geography Note 8

Wisconsin: 55

The county subdivision of Scott town (FIPS code 72200), in place Balance of Scott town (FIPS code 99999) erroneously contains blocks 2048, 2063, and 2064, census tract 203 (020300), Brown County (FIPS code 009), WI (FIPS code 55). These blocks should have been coded to county subdivision and place Pulaski village (FIPS code 65675).

The county subdivision of Pittsfield town (FIPS code 63075), in place Balance of Pittsfield town (FIPS code 99999) erroneously contains block 2049, census tract 203 (020300), Brown County (FIPS code 009). This block should have been coded to county subdivision and place Pulaski village (FIPS code 65675). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 55-001

May 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Technical Documentation Note 1

Chapter 2, How to Use This File

Page 2-2 was replaced because the second sentence under the heading "Geographic Hierarchy Primer" inadvertently references Figure 2-1. The sentence was corrected to read "Figure 2-2 at the end of this chapter provides an example of the various geographic hierarchies used, building from the block."

October 2001

Census 2000 Redistricting Data (Public Law 94-171) Summary File Technical Documentation Note 2

Appendix A, Geographic Definitions

Page A–7 was replaced because the first paragraph in the Area Measurement section stated that to convert square kilometers to square miles, divide by 2.58999. The correct number to divide by is 2.589988.

February 2002

Census 2000 Redistricting Data (Public Law 94-171) Summary File Technical Documentation Note 3

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

Summary File 1
Data Note 1

Summary File 1 state files contain erroneous data for selected geographic components[1] of Congressional Districts (summary level 500[2]). Geographic components are portions of the congressional district within specific types of geography, such as "In metropolitan statistical area (MSA)/consolidated metropolitan statistical area (CMSA)" or "In metropolitan statistical area (MSA)/ consolidated metropolitan statistical area (CMSA)—in MSA/CMSA central city." We plan to include the corrected data for the geographic components of Congressional Districts in the Final National Summary File 1, which is scheduled for public release in June 2002.

To summarize, Congressional District data are *correct* in all SF1 state files for:

- The Congressional District as a whole (summary level 500, geographic component code 00).
- All other Congressional District summary levels having a geographic component code of 00 (summary level 5nn, geographic component code 00).

Congressional District data are *in error* for:

- Congressional district records having a geographic component code *other than* 00 (summary level 500, geographic component codes 52-59, 64-71, 84, 89-95).

This note is applicable to the following data products:

- All Summary File 1 (SF1) state files available at the Census Bureaus FTP site.
- SF1 CD-ROMs (ASCII files only).
- Tables available on American FactFinder between June and September 2001. (Geographic components data for Congressional Districts were removed from American FactFinder on September 11, 2001.)

[1]Geographic components and their codes are listed in the *Census 2000 Summary File 1 Technical Documentation,* in Chapter 7 (Data Dictionary, Footnote Section, page 7-15).

[2]Summary level information is available in the *Census 2000 Summary File 1 Technical Documentation,* Chapter 4 (Summary Level Sequence Chart, page 4-1). The listing of the Congressional District summary levels in SF1 for states appears on page 4-2.

September 2001

Summary File 1
Data Note 2

In the Summary File 1 (SF 1) state files, the state geographic component records[1] contain errors in two geographic header fields. These fields are land area[2] and water area.

These errors appear in the geographic component records for the state (summary level[3] 040). Geographic components are portions of the state within specific types of geography, such as "In metropolitan statistical area (MSA)/consolidated metropolitan statistical area (CMSA)" or "In metropolitan statistical area (MSA)/ consolidated metropolitan statistical area (CMSA)—in MSA/CMSA central city."

The corrected data are included in the Advance National Summary File 1, which is scheduled for public release in November 2001.

To summarize, land area and water area are *correct* for:

- The state as a whole (summary level 040, geographic component code 00).

Land area and water area are *in error* for:

- State records having a geographic component code *other than* 00 (summary level 040, geographic components 52-59, 64-79, 84, 89-95).

This note applies to the following data products:

- All SF 1 state files available at the Census Bureau's FTP site.
- SF 1 state file CD-ROMs and DVDs.
- American FactFinder SF 1 detailed tables (geographic identifier for state geographic components).

---

[1]Geographic components and their codes are listed in the *Census 2000 Summary File 1 Technical Documentation* in Chapter 7 (Data Dictionary, Footnote Section, page 7-15).

[2]Land area (AREALAND) and water area (AREAWATR) appear in the geographic header portion of the data. The location is shown in the *Census 2000 Summary File 1 Technical Documentation* in Chapter 7 (Data Dictionary, Identification Section, pages 7-13 and 7-14).

[3]Complete summary level information is in the *Census 2000 Summary File 1 Technical Documentation* in Chapter 4 (Summary Level Sequence Chart, page 4-1).

October 2001

Summary File 1
Data Note 3

Data for two central city areas in the Summary File 1 (SF 1) state file are in error. These errors are in summary levels 375 and 391. Summary level[1] 375 is the record for the central city portion of a New England County Metropolitan Area (NECMA) within a state. Summary level 391 is a record for the central city portion of a Metropolitan Statistical Area/Consolidated Metropolitan Statistical Area (MSA/CMSA) within a state.

Equivalent records containing the correct data will be part of the Summary File 1 Advance National file. In the Advance National file, the equivalent records will have different summary levels. The correct data for summary level 375 will be in a summary level 372 record; the correct data for summary level 391 will be in a summary level 382 record.

Specifically, in summary level 375 data are *correct* for:

- All states except Massachusetts.
- All records for Massachusetts except the one record described below.

Data are *in error* in summary level 375 for:

- Yarmouth town, Massachusetts within the Barnstable-Yarmouth, MA NECMA. All data cells contain 0.

Data are *correct* in summary level 391 for:

- All records for all states except Massachusetts and New Jersey.
- All records for Massachusetts and New Jersey except the two listed below.

Data are *in error* in summary level 391 for:

- Yarmouth town, Massachusetts within the Barnstable-Yarmouth, MA MSA. All data cells contain 0.
- Dover township, New Jersey within the New York-Northern New Jersey-Long Island NY-NJ-CT-PA CMSA. All data cells contain 0.

This note applies to the following data products:

- All SF 1 state files available at the Census Bureau's FTP site.
- SF1 State file CD-ROMs and DVDs.
- American FactFinder SF 1 detailed tables.

[1]Complete summary level information is available in *Census 2000 Summary File 1 Technical Documentation* in Chapter 4 (Summary Level Sequence Chart, page 4-1). The sequence for summary levels 375 and 391 appears on page 4-2.

October 2001

Summary File 1
Data Note 4

Final National Summary File 1

Some medians in the Final National Summary File 1 may differ slightly from the medians for the same item that were released in the Advance National Summary File 1 or in the series of state files.

- Discrepancies are extremely rare;
- Discrepancies are due solely to the use of updated versions of the tabulation software with different rounding capabilities.

For further information about rounding methods, see the specific discussion of "Rounding" under DERIVED MEASURES in Appendix B, Definitions of Subject Characteristics in the Summary File 1 Technical Documentation.

October 2002

Summary File 1

INDEX TO SUMMARY FILE 1 GEOGRAPHY NOTES

Note	Geographic area
1	Alaska
2	California
3	Connecticut
4	Florida
5	Georgia
6	Nebraska
7	Tennessee
8	Wisconsin

Summary File 1
Geography Note 1

Alaska: 02

Nelson Lagoon Alaska Native village statistical area (ANVSA) (AIANHH 7025) erroneously contains block 2010, census tract 1 (000100) in Aleutians East census area (01598), Aleutians East Borough (013). This block should have not been coded to any ANVSA (9999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 02-003

May 2001

Summary File 1
Geography Note 2

California: 06

Los Angeles city (FIPS code 44000) erroneously contains block 1011, census tract 4002.03 (400203) in East San Gabriel Valley CCD (FIPS code 90810), Los Angeles County (FIPS code 037), CA (FIPS code 06). This block should have been coded to the place Balance of East San Gabriel Valley CCD (FIPS code 99999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP side.

Internal Errata ID 06-001

May 2001

Summary File 1
Geography Note 3

Connecticut: 09

The place record, Balance of Milford town (FIPS code 99999) erroneously contains block 2999, census tract 1502 (150200) in Milford town (FIPS code 47535), New Haven County (FIPS code 009), CT (FIPS code 09). This block should have been coded to place Milford city (balance) (FIPS code 47515). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 09-001

May 2001

Summary File 1
Geography Note 4

Florida: 12

Yeehaw Junction CDP (FIPS code 78975) in St. Cloud CCD (FIPS code 93029), Osceola County (FIPS code 097), FL (FIPS code 12) should be named Buenaventura Lakes with FIPS code 09415. In 1990, this area was named Buena Ventura Lakes (FIPS code 09415). The area that should have been Yeehaw Junction CDP was erroneously not defined and does not appear in any Census 2000 products.

Internal Errata ID 12-001

May 2001

Summary File 1
Geography Note 5

Georgia: 13

The place record Balance of Athens CCD (FIPS code 99999) erroneously contains blocks 2021 and 2023, census tract 1305 (130500) in Athens CCD (FIPS code 90138), Clarke County (FIPS code 059). Both blocks should have been coded to Bogart town (FIPS code 09068).

The place record Balance of Winterville CCD (FIPS code 99999) erroneously contains blocks 1008 and 1009, census tract 1406 (140600) in Winterville CCD (93402), Clarke County (FIPS code 059). Both blocks should have been coded to the place Athens-Clarke County (balance) (FIPS code 03440). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 13-001

May 2001

Summary File 1
Geography Note 6

Nebraska: 31

In the PL 94-171 and Summary File (SF) data products, Cisco CDP (FIPS code 09112) in Lisco precinct (FIPS code 91790), Garden County (FIPS code 069), NE (FIPS code 31) should be named Lisco with FIPS code of 28315.

Internal Errata ID 31-002

May 2001

Summary File 1
Geography Note 7

Tennessee: 47

The place record Balance of Metropolitan Government CCD (FIPS code 99999) erroneously contains blocks 1001 and 1008, census tract 171 (017100) in Metropolitan Government CCD (FIPS code 92200), Davidson County (FIPS code 037), TN (FIPS code 47). Both blocks should have been coded to place Nashville-Davidson (balance) (FIPS code 52006). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 47-001

May 2001

Summary File 1
Geography Note 8

Wisconsin: 55

The county subdivision of Scott town (FIPS code 72200), in place Balance of Scott town (FIPS code 99999) erroneously contains blocks 2048, 2063, and 2064, census tract 203 (020300), Brown County (FIPS code 009), WI (FIPS code 55). These blocks should have been coded to county subdivision and place Pulaski village (FIPS code 65675).

The county subdivision of Pittsfield town (FIPS code 63075), in place Balance of Pittsfield town (FIPS code 99999) erroneously contains block 2049, census tract 203 (020300), Brown County (FIPS code 009). This block should have been coded to county subdivision and place Pulaski village (FIPS code 65675). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 55-001

May 2001

Summary File 1
Technical Documentation Note 1

Chapter 5. List of Tables (Matrices)

The total number of data cells for matrices PCT16, PCT17, and PCT17A through PCT17I was incorrectly stated in Chapter 5, List of Tables (Matrices). The correct total number of data cells is as follows:

Table (matrix)	Total number of data cells
PCT16	52
PCT17	75
PCT17A—PCT17I	75

Chapter 6. Summary Table Outlines

"Emergency and transitional shelters (701–702)" was inadvertently included in matrices PCT16, PCT17, and PCT17A through PCT17I of Chapter 6, Summary Table Outlines. This line is now deleted.

June 2001

Summary File 1
Technical Documentation Note 2

The telephone number for Customer Services, U.S. Census Bureau has changed. The new number is 301-763-INFO (4636). Pages 1–3, 2–4, A–21, E–1, E–4, E–7, and F–1 were replaced to reflect this change.

Chapter 6, Table (Matrix) Outlines

Table (matrix) cell counts and codes were corrected on the following pages:

- Page 6–68

  PCT16 — cell count was changed to [52]

- Page 6–69

  "Other noninstitutional group quarters" — codes were changed to (604, 701-706, 904-905, 909, 911)

- Page 6–70

  PCT17 — cell count was changed to [75]

  "Other noninstitutional group quarters" — codes were changed to (604, 701-706, 801-810, 900-906, 908-909, 911)

- Page 6–84 through Page 6–88

  PCT17A through PCT17I — cell count was changed to [75]

  "Other noninstitutional group quarters" — codes were changed to (604, 701-706, 801-810, 900-906, 908-909, 911)

Chapter 7, Data Dictionary Table (Matrix) Section

- Page 7–48 was replaced because the continuation line, "Related child—Con.," inadvertently included the data dictionary reference name, segment, and MAX size.
- Page 7–87 was replaced because the continuation line, "In households—Con.," inadvertently included the data dictionary reference name, segment, and MAX size.
- The data in the following matrices include 1 or 2 expressed decimals as shown below:

P13.	1 expressed decimal	Page 7–41
P13A. – P13I.	1 expressed decimal	Pages 7–65 and 7–66
P17.	2 expressed decimals	Page 7–42
P17A. – P17I.	2 expressed decimals	Pages 7–68 and 7–69
P33.	2 expressed decimals	Page 7–49
P33A. – P3I.	2 expressed decimals	Pages 7–94 and 7–95
H12.	2 expressed decimals	Page 7–236
H12A. – H12I.	2 expressed decimals	Pages 7–242 and 7–243

- Page 7–236 was replaced because two lines in table (matrix) H14 did not show the data dictionary reference name, segment, and MAX size.

July 2001

Summary File 1
Technical Documentation Note 3

This user update is described on our Web site (www.census.gov) as:

Technical Note on Same-Sex Unmarried Partner Data From the 1990 and 2000 Censuses

The release of data in the SF 1 files from the 2000 census has brought with it a number of analyses documenting change that has occurred since the last census was conducted in 1990. While many of the variables and processes between the two censuses are comparable, some are not, and direct comparison of some estimates may lead to misleading conclusions. This note discusses one such topic, that of "unmarried partners," and advises that for some analyses — those involving unmarried same-sex partners — direct comparison of the 1990 and 2000 estimates is not substantively valid.

The household relationship item in both the 1990 and the 2000 censuses offered many ways of identifying how other people in the household were related to the householder (the person in whose name the house is owned or rented). Categories included spouse, child or other relative of the householder, housemate/roommate, roomer/boarder, and unmarried partner. In all circumstances, the respondent was asked to choose the category that best represented how other members of the household were related to the householder.

In both censuses, the "spouse" and "unmarried partner" response categories were defined and asked the same way. However, there were important differences in data processing that mean that some of the data are not comparable, limiting the usefulness of comparisons of the number of same-sex unmarried partners between these two censuses.

In both censuses, if a person was identified as the "spouse" of the householder and was the same sex as the householder, the "spouse" response was flagged for further review and allocation, that is, assignment of a value other than that originally reported, based on other data on the form. In 1990, the edit and allocation procedures did not allow same-sex "spouse" combinations to occur, thus resulting in the allocation of one of these two items in order to achieve editing consistency among the responses.

Processing steps were changed for Census 2000 for households that contained same-sex "spouses." If the person with the "spouse" category was the same sex as the householder and if neither person had their sex previously allocated, a relationship response of "spouse" was allocated as an "unmarried partner" response. Since marital status was no longer on the short form, its given value could not be considered (or modified) in this allocation procedure as it had been in 1990.

Data allocation is a standard statistical practice that is followed by most data collection agencies. Data on the relationship item (as other items) were subject to allocation in the census, as they are in virtually all Census Bureau surveys. In 1990, the marital status item was available on the 100 percent (short) form and aided in both the evaluation of the consistency of responses between the householder and the "spouse," and in the subsequent allocation procedure. The 1990 procedure allocated responses via a statistical model that distributed allocated responses from answers given by respondents in a proximate geographic area. This procedure used key demographic data from the census form, including marital status, as stratifying factors to provide a reasonable distribution of allocated responses. This procedure, while ensuring that no same-sex spouse response could be subsequently allocated, produced a set of allocated responses that could have included an "unmarried partner" response as well as any other response that was consistent with the age/sex/marital status profile of the respondent. This would include being allocated as a sibling or a relative, for example, or if the age differences were far enough apart (15 or more years), even a parent or child of the householder.

Three principal factors affected our decision to take this approach for Census 2000.

1. Same-sex spouse responses were flagged as invalid to comply with the 1996 Federal Defense of Marriage Act (H.R. 3396) passed by the 104th Congress. This act instructs all federal agencies only to recognize opposite-sex marriages for the purposes of enacting any agency programs. In order for Census Bureau data to be consistent with this act and the data requirements of other federal agencies, same-sex spouse responses were invalidated. The legislation defines marriage and spouse as follows:

   "In determining the meaning of any Act of Congress, or of any ruling, regulation or interpretation of the various administrative bureaus and agencies of the United States, the word 'marriage' means only a legal union between one man and one woman as husband and wife, and the word 'spouse' refers only to a person of the opposite sex who is a husband or wife."

   In order for the Census Bureau to be consistent with this act and the data requirements of other federal agencies, same-sex "spouse" responses were invalidated.

2. The second issue was statistical in nature. The principal basis of any good statistical allocation routine rests on the selection of the stratifying or input factors to provide a good statistical model. Without marital status data on the 100 percent form in Census 2000, the allocation routine would be relatively weak. Since many partners are roughly the same age, a statistical routine without marital status as one of its factors would have likely resulted in an overestimate of adult siblings or relatives, as the majority of people living in households are relatives, and this is the population from which we would draw our allocated responses. Additionally, if the same-sex partners were more than 15 years difference in age, the statistical routine would have likely allocated the invalidated "spouse" response as either a "child" or "parent" of the householder, as these types of relatives predominate in households in this age range of differences. This was an unacceptable outcome, as it would actually destroy the intent of the original "spouse" response, which clearly indicated a nonparental type of relationship. It should be noted that the "spouse" response on the form is assumed to be deliberate — not accidental — as it was the first response category on the question and was not placed between other possible response categories that may have been meant to be marked, such as housemates or roomers.

3. The third factor took into consideration that couples in long term same-sex relationships may consider themselves as "married partners" and thus respond as such on the census form. In addition, at the time of writing the editing program for Census 2000, there were several challenges in the courts concerning the legality of same-sex marriages. Clearly, we could not ignore the fact that same-sex spouse responses were going to be recorded during Census 2000. In light of these social and legal aspects — and the lack of a key variable in the statistical allocation routine (marital status) — the assignment of same-sex "married" couples to the same-sex "unmarried partner" category was the procedure chosen for the editing process. We were adverse to a randomized allocation of these responses after people had clearly marked a close relationship preference on the census form.

   As a result of these changes in the processing routine, estimates of same-sex unmarried partners are not comparable between the 1990 and 2000 census. We believe 2000 census estimates of this category are better estimates than those produced in 1990. It should also be noted that estimates of opposite-sex unmarried partners, however, were not affected by these editing procedures and changes and are comparable between the two censuses.

For further information on this topic, please contact the Fertility and Family Statistics Branch on 301-457-2416.

July 2001

Summary File 1
Technical Documentation Note 4

Chapter 4, Summary Level Sequence Chart

The following summary levels were corrected on the following pages:

- Page 4–3, Advance National Summary File 1

  060 was changed to—060 State-County-County Subdivision

  070 was changed to—070 State-County-County Subdivision–Place/Remainder

- Page 4–5, Final National Summary File 1

  060 was changed to—060 State-County-County Subdivision

  070 was changed to—070 State-County-County Subdivision–Place/Remainder

August 2001

Summary File 1
Technical Documentation Note 5

Alaskan Athabascan

The following corrections were made to the spelling of Alaskan Athabascan:

Chapter 6, Summary Table Outlines

- Page 6-60, Matrix PCT1 Alaska Athabaskan was changed to Alaskan Athabascan
- Page 6-61, Matrix PCT2
  Alaska Athabaskan was changed to Alaskan Athabascan
- Page 6-62, Matrix PCT3
  Alaska Athabaskan was changed to Alaskan Athabascan

Chapter 7, Data Dictionary

- Page 7-106, Matrix PCT1
  Alaska Athabaskan was changed to Alaskan Athabascan
- Page 7-107, Matrix PCT2
  Alaska Athabaskan was changed to Alaskan Athabascan
- Page 7-108, Matrix PCT3
  Alaska Athabaskan was changed to Alaskan Athabascan

Appendix B, Definitions of Subject Characteristics

- Page B-13
  Alaskan Athabaskan was changed to Alaskan Athabascan

Appendix G, Code Lists

- Page G-21
  Oregon Athabaskan was changed to Oregon Athabascan

August 2001

Summary File 1
Technical Documentation Note 6 – Updated January 2003

In October 2001, the technical documentation note below was issued. However, the number of data items for file 33 was incorrectly stated. The correct number of data items for file 33 is 228. Page 2–4 in Chapter 2, How to Use This File was replaced to reflect the change in Figure 2–2, File/Table Segmentation.

Appendix A, Census 2000 Geographic Terms and Concepts, Minor Civil Divisions

The following paragraphs were added to the description of Minor Civil Divisions on page A-14:

In eight MCD states (Illinois, Indiana, Kansas, Missouri, Nebraska, North Dakota, Ohio, and South Dakota) the MCD townships serve as general-purpose local governments but do not have the ability to perform all the governmental functions as incorporated places. This category also includes the counties in American Samoa. Missouri is exceptional in that it has a minority of townships that serve as general-purpose governments (the majority of townships in Missouri fall into the category described below).

In the remaining eight MCD states (Arkansas, Iowa, Louisiana, Maryland, Mississippi, North Carolina, Virginia, and West Virginia), the counties containing precincts in Illinois and Nebraska, the townships in Williamson County, Illinois, and the majority of townships in Missouri, the MCDs are geographic subdivisions of the counties, and are not governmental units. The MCDs in Puerto Rico and the Island Areas (except American Samoa) also fall into this classification.

Chapter 2, How to Use This File

The number of data items in Figure 2-2, File/Table Segmentation was incorrectly stated. The correct number of data items for files 04, 15, 33, 34, 35, and 36 follows. Page 2-4 was replaced to reflect these changes.

File name	Number of data items
04	149
15	196
33	225
34	225
35	225
36	75

Chapter 6, Summary Table Outlines

American Indian and Alaska Native tribe codes were corrected for matrices PCT1, PCT2, and PCT3. Pages 6-59 through 6-62 were replaced.

October 2001

Summary File 1
Technical Documentation Note 7

Table P26F

The universe for table P26F was corrected to add the word "race." It was corrected from "Universe: Households with a householder who is Some other alone" to "Universe: Households with a householder who is Some other race alone" in both Chapter 6, Summary Table Outlines (page 6-31) and Chapter 7, Data Dictionary (page 7-72).

October 2001

Summary File 1
Technical Documentation Note 8

Appendix A, Census Geographic Terms and Concepts

Page A–8 was replaced because the first paragraph in the Area Measurement section stated that to convert square kilometers to square miles, divide by 2.58999. The correct number to divide by is 2.589988.

February 2002

Summary File 1
Technical Documentation Note 9

Appendix B, Definitions of Subject Characteristics

Page B–14 was replaced because the last sentence in the section "Two or more races" was deleted as follows: "Additionally, in some data products, data showing characteristics of the population by race for people reporting the four most commonly reported race combinations will be shown without a population threshold."

June 2002

Summary File 1
Technical Documentation Note 10

Chapter 4, Summary Level Sequence Chart

Page 4–5 was replaced for the Final National File because summary level 276 was incorrectly aligned with summary level 275.

August 2002

Summary File 1
Technical Documentation Note 11

Chapter 1, Abstract

The National Files (Advance and Final) section under “Geographic Content” was corrected to indicate that the files provide summaries for all county subdivisions and places, not just those of 10,000 or more population.

August 2002

Summary File 1
Technical Documentation Note 12

Chapter 4. Summary Level Sequence Chart

The summary level sequence chart (Chapter 4) in the Summary File 1 technical documentation was corrected for Congressional Districts (summary level 500). The geographic components for Congressional Districts are now listed correctly as "00" for the state summary files and "00, 52-59, 64-71, 84, and 89-95" for the final national summary file.

September 2002

Summary File 1
Technical Documentation Note 13

Chapter 7. Data Dictionary

The Internet addresses listed in footnote 7, page 7–18 for Summary File 1 were corrected as follows:

General information about FIPS can be found on the Internet at:
http://geonames.usgs.gov/

The actual codes assigned to the many FIPS 55 entities can be found at:
http://geonames.usgs.gov/fips55.html.

April 2003

Summary File 1
Technical Documentation Note 14

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

Summary File 2
Data Note 1

Summary File 2, Table PCT5 provides data on the distribution by sex and age of people who live in households. When this table is shown for a particular race, Hispanic or Latino origin, or American Indian or Alaska Native tribe, the data are tallied according to the race, Hispanic or Latino origin, or American Indian or Alaska Native tribe of the householder. For example, when the table is presented for Asian alone, the data represent all people in households with an Asian alone householder, even if not all people in the household are Asian alone.

The presentation of data in SF 2, Table PCT5 is in contrast to Summary File 1, Tables PCT13(A-I), which show data on the distribution by sex and age. These data represent the race, Hispanic or Latino origin, or American Indian or Alaska Native tribe of each person in the household. For example, in SF 1, Table PCT13D, the data represent all people who live in households who are Asian alone, whether or not the householder is Asian alone.

May 2002

Summary File 2

INDEX TO SUMMARY FILE 2 GEOGRAPHY NOTES

Note	Geographic area
1	Alaska
2	California
3	Connecticut
4	Florida
5	Georgia
6	Nebraska
7	Tennessee
8	Wisconsin

Summary File 2
Geography Note 1

Alaska: 02

Nelson Lagoon Alaska Native village statistical area (ANVSA) (AIANHH 7025) erroneously contains block 2010, census tract 1 (000100) in Aleutians East census area (01598), Aleutians East Borough (013). This block should have not been coded to any ANVSA (9999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 02-003

May 2001

Summary File 2
Geography Note 2

California: 06

Los Angeles city (FIPS code 44000) erroneously contains block 1011, census tract 4002.03 (400203) in East San Gabriel Valley CCD (FIPS code 90810), Los Angeles County (FIPS code 037), CA (FIPS code 06). This block should have been coded to the place Balance of East San Gabriel Valley CCD (FIPS code 99999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP side.

Internal Errata ID 06-001

May 2001

Summary File 2
Geography Note 3

Connecticut: 09

The place record, Balance of Milford town (FIPS code 99999) erroneously contains block 2999, census tract 1502 (150200) in Milford town (FIPS code 47535), New Haven County (FIPS code 009), CT (FIPS code 09). This block should have been coded to place Milford city (balance) (FIPS code 47515). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 09-001

May 2001

Summary File 2
Geography Note 4

Florida: 12

Yeehaw Junction CDP (FIPS code 78975) in St. Cloud CCD (FIPS code 93029), Osceola County (FIPS code 097), FL (FIPS code 12) should be named Buenaventura Lakes with FIPS code 09415. In 1990, this area was named Buena Ventura Lakes (FIPS code 09415). The area that should have been Yeehaw Junction CDP was erroneously not defined and does not appear in any Census 2000 products.

Internal Errata ID 12-001

May 2001

Summary File 2
Geography Note 5

Georgia: 13

The place record Balance of Athens CCD (FIPS code 99999) erroneously contains blocks 2021 and 2023, census tract 1305 (130500) in Athens CCD (FIPS code 90138), Clarke County (FIPS code 059). Both blocks should have been coded to Bogart town (FIPS code 09068).

The place record Balance of Winterville CCD (FIPS code 99999) erroneously contains blocks 1008 and 1009, census tract 1406 (140600) in Winterville CCD (93402), Clarke County (FIPS code 059). Both blocks should have been coded to the place Athens-Clarke County (balance) (FIPS code 03440). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 13-001

May 2001

Summary File 2
Geography Note 6

Nebraska: 31

In the PL 94-171 and Summary File (SF) data products, Cisco CDP (FIPS code 09112) in Lisco precinct (FIPS code 91790), Garden County (FIPS code 069), NE (FIPS code 31) should be named Lisco with FIPS code of 28315.

Internal Errata ID 31-002

May 2001

Summary File 2
Geography Note 7

Tennessee: 47

The place record Balance of Metropolitan Government CCD (FIPS code 99999) erroneously contains blocks 1001 and 1008, census tract 171 (017100) in Metropolitan Government CCD (FIPS code 92200), Davidson County (FIPS code 037), TN (FIPS code 47). Both blocks should have been coded to place Nashville-Davidson (balance) (FIPS code 52006). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 47-001

May 2001

Summary File 2
Geography Note 8

Wisconsin: 55

The county subdivision of Scott town (FIPS code 72200), in place Balance of Scott town (FIPS code 99999) erroneously contains blocks 2048, 2063, and 2064, census tract 203 (020300), Brown County (FIPS code 009), WI (FIPS code 55). These blocks should have been coded to county subdivision and place Pulaski village (FIPS code 65675).

The county subdivision of Pittsfield town (FIPS code 63075), in place Balance of Pittsfield town (FIPS code 99999) erroneously contains block 2049, census tract 203 (020300), Brown County (FIPS code 009). This block should have been coded to county subdivision and place Pulaski village (FIPS code 65675). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 55-001

May 2001

Summary File 2
Technical Documentation Note 1

This user update is described on our Web site (www.census.gov) as:

Technical Note on Same-Sex Unmarried Partner Data From the 1990 and 2000 Censuses

The release of data in the SF 1 files from the 2000 census has brought with it a number of analyses documenting change that has occurred since the last census was conducted in 1990. While many of the variables and processes between the two censuses are comparable, some are not, and direct comparison of some estimates may lead to misleading conclusions. This note discusses one such topic, that of "unmarried partners," and advises that for some analyses — those involving unmarried same-sex partners — direct comparison of the 1990 and 2000 estimates is not substantively valid.

The household relationship item in both the 1990 and the 2000 censuses offered many ways of identifying how other people in the household were related to the householder (the person in whose name the house is owned or rented). Categories included spouse, child or other relative of the householder, housemate/roommate, roomer/boarder, and unmarried partner. In all circumstances, the respondent was asked to choose the category that best represented how other members of the household were related to the householder.

In both censuses, the "spouse" and "unmarried partner" response categories were defined and asked the same way. However, there were important differences in data processing that mean that some of the data are not comparable, limiting the usefulness of comparisons of the number of same-sex unmarried partners between these two censuses.

In both censuses, if a person was identified as the "spouse" of the householder and was the same sex as the householder, the "spouse" response was flagged for further review and allocation, that is, assignment of a value other than that originally reported, based on other data on the form. In 1990, the edit and allocation procedures did not allow same-sex "spouse" combinations to occur, thus resulting in the allocation of one of these two items in order to achieve editing consistency among the responses.

Processing steps were changed for Census 2000 for households that contained same-sex "spouses." If the person with the "spouse" category was the same sex as the householder and if neither person had their sex previously allocated, a relationship response of "spouse" was allocated as an "unmarried partner" response. Since marital status was no longer on the short form, its given value could not be considered (or modified) in this allocation procedure as it had been in 1990.

Data allocation is a standard statistical practice that is followed by most data collection agencies. Data on the relationship item (as other items) were subject to allocation in the census, as they are in virtually all Census Bureau surveys. In 1990, the marital status item was available on the 100 percent (short) form and aided in both the evaluation of the consistency of responses between the householder and the "spouse," and in the subsequent allocation procedure. The 1990 procedure allocated responses via a statistical model that distributed allocated responses from answers given by respondents in a proximate geographic area. This procedure used key demographic data from the census form, including marital status, as stratifying factors to provide a reasonable distribution of allocated responses. This procedure, while ensuring that no same-sex spouse response could be subsequently allocated, produced a set of allocated responses that could have included an "unmarried partner" response as well as any other response that was consistent with the age/sex/marital status profile of the respondent. This would include being allocated as a sibling or a relative, for example, or if the age differences were far enough apart (15 or more years), even a parent or child of the householder.

Three principal factors affected our decision to take this approach for Census 2000.

1. Same-sex spouse responses were flagged as invalid to comply with the 1996 Federal Defense of Marriage Act (H.R. 3396) passed by the 104th Congress. This act instructs all federal agencies only to recognize opposite-sex marriages for the purposes of enacting any agency programs. In order for Census Bureau data to be consistent with this act and the data requirements of other federal agencies, same-sex spouse responses were invalidated. The legislation defines marriage and spouse as follows:

   "In determining the meaning of any Act of Congress, or of any ruling, regulation or interpretation of the various administrative bureaus and agencies of the United States, the word 'marriage' means only a legal union between one man and one woman as husband and wife, and the word 'spouse' refers only to a person of the opposite sex who is a husband or wife."

   In order for the Census Bureau to be consistent with this act and the data requirements of other federal agencies, same-sex "spouse" responses were invalidated.

2. The second issue was statistical in nature. The principal basis of any good statistical allocation routine rests on the selection of the stratifying or input factors to provide a good statistical model. Without marital status data on the 100 percent form in Census 2000, the allocation routine would be relatively weak. Since many partners are roughly the same age, a statistical routine without marital status as one of its factors would have likely resulted in an overestimate of adult siblings or relatives, as the majority of people living in households are relatives, and this is the population from which we would draw our allocated responses. Additionally, if the same-sex partners were more than 15 years difference in age, the statistical routine would have likely allocated the invalidated "spouse" response as either a "child" or "parent" of the householder, as these types of relatives predominate in households in this age range of differences. This was an unacceptable outcome, as it would actually destroy the intent of the original "spouse" response, which clearly indicated a nonparental type of relationship. It should be noted that the "spouse" response on the form is assumed to be deliberate — not accidental — as it was the first response category on the question and was not placed between other possible response categories that may have been meant to be marked, such as housemates or roomers.

3. The third factor took into consideration that couples in long term same-sex relationships may consider themselves as "married partners" and thus respond as such on the census form. In addition, at the time of writing the editing program for Census 2000, there were several challenges in the courts concerning the legality of same-sex marriages. Clearly, we could not ignore the fact that same-sex spouse responses were going to be recorded during Census 2000. In light of these social and legal aspects — and the lack of a key variable in the statistical allocation routine (marital status) — the assignment of same-sex "married" couples to the same-sex "unmarried partner" category was the procedure chosen for the editing process. We were adverse to a randomized allocation of these responses after people had clearly marked a close relationship preference on the census form.

   As a result of these changes in the processing routine, estimates of same-sex unmarried partners are not comparable between the 1990 and 2000 census. We believe 2000 census estimates of this category are better estimates than those produced in 1990. It should also be noted that estimates of opposite-sex unmarried partners, however, were not affected by these editing procedures and changes and are comparable between the two censuses.

For further information on this topic, please contact the Fertility and Family Statistics Branch on 301-457-2416.

July 2001

Summary File 2
Technical Documentation Note 2

Appendix A, Census Geographic Terms and Concepts

Page A–8 was replaced because the first paragraph in the Area Measurement section stated that to convert square kilometers to square miles, divide by 2.58999. The correct number to divide by is 2.589988.

February 2002

Summary File 2
Technical Documentation Note 3

Appendix B, Definitions of Subject Characteristics

Page B–14 was replaced because the last sentence in the section "Two or more races" was deleted as follows: "Additionally, in some data products, data showing characteristics of the population by race for people reporting the four most commonly reported race combinations will be shown without a population threshold."

June 2002

Summary File 2
Technical Documentation Note 4

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

Summary File 3
Data Note 1

On the Census 2000 long-form questionnaire, individuals could report more than one type of disability. Summary File 3 Table P41, Age by Types of Disability for the Civilian Noninstitutionalized Population 5 Years and Over With Disabilities, has as its universe the total disabilities tallied. Each line of the table represents the number of occurrences of a particular disability, and the numbers should be interpreted with care. For example, the second line of data in the table titled "Total disabilities tallied for people 5 to 15 years" does not refer to the number of people 5 to 15 years old, or to the number of people 5 to 15 with a disability. Rather it is the sum of the number of all disabilities reported among the 5 to 15 year old population. Lines in the table referencing specific disabilities are more easily interpreted. The third line in the table titled "Sensory disability," for example, refers to the number of sensory disabilities reported among people 5 to 15 years (or the number of people 5 to 15 years old with a sensory disability).

Data users wanting to know the percent of civilian noninstitutionalized people 5 to 15 years old with, for example, a sensory disability should divide line 3 from Table P41 with the sum of lines 3 and 27 from Table P42, Sex by Age by Disability Status by Employment Status for the Civilian Noninstitutionalized Population 5 Years and Over. Data users wanting to know the same percentages for one of the nine race or Hispanic or Latino origin groups should use Tables PCT67A-I and Tables PCT68A-I, as appropriate.

June 2002

Summary File 3
Data Note 2

Users may find slight differences in aggregate earnings for households between the Demographic Profile and Summary File 3 and related products. These differences are due to the treatment of offsetting positive and negative amounts for household members. Whenever offsetting values occurred, the Demographic Profile assigned these households a value zero while Summary File 3 and related products assigned a value of one dollar. The assignment of one dollar allows users to distinguish those households that had earnings from those households that did not have earnings. This will have little effect, if any, on mean household earnings.

June 2002

Summary File 3
Data Note 3

Users may find slight differences in the Occupants Per Room calculations between the Demographic Profile and Summary File 3, Summary File 4, and related products. “Occupants per room” is obtained by dividing the number of people in each occupied housing unit by the number of rooms in the unit. The Summary File 3 products correctly used a topcode value of “10 rooms” for those occupied housing units with “9 or more rooms.” In the Demographic Profiles, an incorrect topcode value of “9 rooms” was used.

June 2002

Summary File 3
Data Note 4 – Updated June 2004

In July 2002, the Census Bureau issued the following Data Note 4 regarding the Census 2000 Summary File 3 (SF3) data:

> The Census Bureau is aware there may be a problem or problems in the employment-status data of Census 2000 Summary File 3 (including tables P38, P43-46, P149A-1, P150A-I, PCT35, PCT69A-1, and PCT 70A-1). The labor force data for some places where colleges are located appear to overstate the number in the labor force, the number unemployed, and the percent unemployed, probably because of reporting or processing errors. The exact cause is unknown, but the Census Bureau will continue to research the problem.

Our further research into this "college-town" issue indicates that the problem extended beyond places with colleges to the country in general. We learned that it stems from the tendency of many working-age people living in civilian noninstitutional group quarters (GQ), such as college dormitories, worker dormitories, and group homes (for the mentally ill or physically handicapped), to exhibit a particular pattern of entries to the employment questions in Census 2000.[1] We now estimate that the pattern affected the employment data for about 15 percent of the civilian noninstitutional GQ population 16 years of age and over in the United States, or around 500,000 people. It had an impact on the Census 2000 labor force statistics for the entire country, but its effects were most visible and substantial for places, such as college towns, with high concentrations of people living in civilian noninstitutional group quarters.

In Census 2000, the majority of people in the GQ population were enumerated by the Individual Census Report (ICR) form, which collected employment data in a battery of six questions (questions 23, 27a-e). The responses to these questions were captured and fed into a set of rules (called the Employment Status Recode (ESR) edit) that used the combined information from all six questions to assign each person to one of the following four employment-status categories: not in universe (all people less than 16 years old), employed, unemployed, and not in labor force.

For a significant segment of the GQ population, a so-called "3/3" response pattern was entered into the ESR edit.[2] This pattern is shown in the following table:

3/3 Input Pattern From ICR Forms

Question numberon ICR	Question wording	Entry
23	LAST WEEK, did you do ANY work for either pay or profit?	Missing
27a	LAST WEEK, were you on layoff from a job?	Missing
27b	LAST WEEK, were you TEMPORARILY absent from a job or business?	Missing
27c	(For people on layoff) Have you been informed that you will be recalled to work within the next 6 months OR been given a date to return to work?	Yes
27d	Have you been looking for work during the last four weeks?	Yes
27e	LAST WEEK, could you have started a job if offered one, or returned to work if recalled?	Yes

[1] The pattern also appeared frequently for people in institutional group quarters, such as prisons and juvenile institutions, but because of the way employment categories are defined, it had no impact on the employment data for these people.

[2] "3/3" refers to the fact that the responses to the first three questions, which appeared on page 4 of the ICR, are all missing; and those responses to the last three questions, which were on page 5 of the ICR, are all "yes."

The 3/3 pattern represents an incomplete set of information, since entries to the first three questions are missing. The ESR edit assigned people with this pattern to the "unemployed" category, because the edit had three built-in assumptions:

1) The respondents saw and reacted to each and every question in the employment series;

2) The 3/3 pattern represented the faithful recording of actual responses (or non-responses) to the questions; and

3) People who responded in this manner were more likely to meet the official criteria for the "unemployed" category than for any other category.[3]

Our research has revealed that most of the GQ cases with the 3/3 pattern may not have met one of the first two assumptions. We are still investigating, but we think that, in most cases, the pattern resulted from anomalies in the data collection or processing systems. Unfortunately, we cannot test our hypothesis by comparing the 3/3 pattern with actual reports from the respondents. The images of the filled-out ICR's will not be accessible until the completion, in 2006 at the earliest, of the Census Bureau's project to image the forms for delivery to the National Archives.

The potential effect of the ESR outcome for the 3/3 pattern is to increase the count of unemployed people at the expense of the counts of the employed and the not-in-labor-force groups. We have done some research to estimate the potential impact of the phenomenon on the labor force data for the nation as a whole. Our preliminary estimates are that it may have incorrectly decreased the number of employed people by about 235,000 (the number of employed in SF3 was 129.7 million), reduced the number of people not in the labor force by 285,000 (SF3 figure of 78.3 million), increased the number of unemployed by 519,000 (SF3 figure of 7.9 million), and raised the unemployment rate by 0.4 percentage point (SF3 figure was 5.8 percent).

Comparatively, the impact of the phenomenon on areas below the national level may be much greater, depending upon the relative size of the GQ population within the given area. The Census 2000 unemployment rate for the city of Williamsburg, Virginia, for example, was 41.7 percent (our research indicated that this rate resulted primarily from the prevalence of the 3/3 pattern among residents of college dormitories, who make up a large percentage of the city's population). To help data users gauge the impact of the phenomenon on their applications, and possibly to adjust for it, the Census Bureau released a tabulation of employment-status data for the nation, states, counties, and places, that was restricted to the population residing in households. This tabulation is available at: http://www.census.gov/hhes/www/laborfor.html

We will continue our research and report on further findings as they become available.

[3]They reported that they were looking for work and could have started a job last week. Because they did not report whether they had a job last week (people with a job are classified as "employed"), it is reasonable to classify them as "unemployed."

June 2004

Summary File 3
Data Note 5

In Summary File 3 (SF 3), data are not available for four tables when using the geographic component[1] rural farm (geographic component 49). These tables are:

P3. 100-Percent Count of the Population
P4. Percent of the Population in Sample
H3. 100-Percent Count of Housing Units
H4. Percent of Housing Units in Sample by Occupancy Status

This is because these tables refer to a 100-percent count, and the concept of farm residence[2] is defined based on answers available only on the sample (long-form) questionnaire. Tables P3, P4, H3, and H4 are zero-filled for the rural farm geographic component. Also zero-filled are fields for land area, water area, population count (100-percent), housing unit count (100-percent), and internal points (latitude and longitude) in the geographic header record[3].

For the remaining tables in SF 3, characteristics data are available for the rural farm geographic component. In the SF 3 state-level files, the rural farm data are available for states (summary level[4] 040) and counties (summary level 050). In the SF 3 national file, these data are available for the United States (summary level 010), regions (020), divisions (030), and states (040).

This note applies to the following data products:

- All SF 3 files available at the Census Bureau's FTP site.
- SF 3 CD-ROMs and DVDs.
- American FactFinder SF 3 detailed tables (geographic identifier for state geographic components)

[1]Geographic components and their codes are listed in the *Census 2000 Summary File 3 Technical Documentation* in Chapter 7 (Data Dictionary, Footnote Section).
[2]Detailed explanations of subject characteristics are found in the *Census 2000 Summary File 3 Technical Documentation* in Appendix B (Definitions of Subject Characteristics).
[3]A description of the geographic header record is found in the *Census 2000 Summary File 3 Technical Documentation* in Chapter 2 (How to Use This File).
[4]Complete summary level information is in the *Census 2000 Summary File 3 Technical Documentation* in Chapter 4 (Summary Level Sequence Chart).

July 2002

Summary File 3
Data Note 6

COMPARING SF 3 ESTIMATES WITH CORRESPONDING VALUES IN SF 1 AND SF 2

As in earlier censuses, the responses from the sample of households reporting on long forms must be weighted to reflect the entire population. Specifically, each responding household represents, on average, six or seven other households who reported using short forms.

One consequence of the weighting procedures is that each estimate based on the long form responses has an associated confidence interval. These confidence intervals are wider (as a percentage of the estimate) for geographic areas with smaller populations and for characteristics that occur less frequently in the area being examined (such as the proportion of people in poverty in a middle-income neighborhood).

In order to release as much useful information as possible, statisticians must balance a number of factors. In particular, for Census 2000, the Bureau of the Census created weighting areas—geographic areas from which about two hundred or more long forms were completed—which are large enough to produce good quality estimates. If smaller weighting areas had been used, the confidence intervals around the estimates would have been significantly wider, rendering many estimates less useful due to their lower reliability.

The disadvantage of using weighting areas this large is that, for smaller geographic areas within them, the estimates of characteristics that are also reported on the short form will not match the counts reported in SF 1 or SF 2. Examples of these characteristics are the total number of people, the number of people reporting specific racial categories, and the number of housing units. The official values for items reported on the short form come from SF 1 and SF 2.

The differences between the long form estimates in SF 3 and values in SF 1 or SF 2 are particularly noticeable for the smallest places, tracts, and block groups. The long form estimates of total population and total housing units in SF 3 will, however, match the SF 1 and SF 2 counts for larger geographic areas such as counties and states, and will be essentially the same for medium and large cities.

This phenomenon also occurred for the 1990 Census, although in that case, the weighting areas included relatively small places. As a result, the long form estimates matched the short form counts for those places, but the confidence intervals around the estimates of characteristics collected only on the long form were often significantly wider (as a percentage of the estimate).

SF 1 gives exact numbers even for very small groups and areas; whereas, SF 3 gives estimates for small groups and areas such as tracts and small places that are less exact. The goal of SF 3 is to identify large differences among areas or large changes over time. Estimates for small areas and small population groups often do exhibit large changes from one census to the next, so having the capability to measure them is worthwhile.

August 2002

Summary File 3
Data Note 7

The following new section was added to Chapter 8, Accuracy of the Data.

CONSISTENCY WITH COMPLETE COUNTS

As described earlier, Census 2000 long form data were collected on a sample basis. Cities and incorporated places were used to determine sampling rates to support estimates for these areas. As a result, each city, incorporated place, school district, and county had addresses selected in the long form sample.

To produce estimates from the long form data, weighting was performed at the weighting area level. In forming weighting areas, trade-offs between reliability, consistency of the estimates, and complexity of the implementation were considered. The decision was made to form weighting areas consisting of small geographic areas with at least 400 sample persons (or about 200 or more completed long forms) that do not cross county boundaries. No other boundary constraints were imposed. Thus, total population estimates from the long form data will agree with census counts reported in SF 1 and SF 2 for the weighting area, county, and other higher geographic areas obtained by combining either weighting areas or counties. Differences between long form estimates of characteristics in the SF 3 and their corresponding values in the SF 1 or SF 2 are particularly noticeable for small places, tracts, and block groups. Examples of these characteristics are the total number of people, the number of people reporting specific racial categories, and the number of housing units. The official values for items reported on the short form come from SF 1 and SF 2.

Because the weighting areas were formed at a smaller geographic level, any differential nonresponse to long form questionnaires by demographic groups or geographical areas included in a weighting area may introduce differences in complete counts (SF 1 and SF 2) and the SF 3 total population estimates. Also, an insufficient number of sample cases in the weighting matrix cells could lead to differences in SF 1, SF 2, and SF 3 population totals. Thus, differences between the census and SF 3 counts are typical and expected.

In 1990, separate tabulations were not prepared for small areas below a certain size. In contrast, Census 2000 tabulations are being prepared for all areas to maximize data availability. This approach may lead to a greater number of anomalous results than what may have been observed with tabulations released from the 1990 census. A similar phenomenon occurred in the 1990 census when weighting areas respected city and place boundaries. Census counts differed from the long form data estimates in small places. As expected, these differences were sometimes large.

The SF 1 tables provide the official census count of the number of people in an area. The SF 3 tables provide estimates of the proportion of people with specific characteristics, such as occupation, disability, or educational attainment. The total number of people in the SF 3 table is provided for use as the denominator, or base, for these proportions. Estimates in the SF 3 tables give the best estimates of the proportion of people with a particular characteristic, but the census count is the official count of how many people are in the area.

The SF 1 gives exact numbers even for very small groups and areas; whereas, SF 3 gives estimates for small groups and areas, such as tracts and small places, that are less exact. The goal of SF 3 is to identify large differences among areas or large changes over time. Estimates for small areas and small population groups often exhibit large changes from one census to the next, so having the capability to measure them is worthwhile.

August 2002

Summary File 3
Data Note 8

Median incomes for nonfamily households by race, Tables 156A through P156I, were calculated from a 38-category income distribution rather than the standard 39-category income distribution. The 38-category distribution collapsed the two highest categories ($175,000 - $199,999 and $200,000 and over) into a single category of $175,000 and over.

August 2002

Summary File 3
Data Note 9

Census 2000 Summary File 3 CD-ROMs
Census 2000 Data Engine Software
Output | Create Output As Summary

The Census 2000 Summary File 3 database contains several tables of normalized data items, such as P53–Median Household Income in 1999, P82–Per Capita Income in 1999, and H18–Average Household Size of Occupied Housing Units by Tenure. In general, the **Census 2000 Data Engine** software's **Create Output As Summary** function recognizes normalized data items and presents them as weighted averages of the summarized geographic components using the 100 percent population or housing count as the weighting factor. However, the version of the Census 2000 Data Engine software used on the Summary File 3 State CD-ROMs fails to recognize **Per Capita** as a one of the normalization techniques and performs a standard summation. This applies only to tables P82 and P157A through P157I. The Per Capita Income value displayed on the **DP–3, Profile of Selected Economic Characteristics,** is derived from the formula (P083001/P001001) rather than (P082001) as originally specified so that **Create Output As Summary** will perform correctly. The Summary File 3 DVD will contain a version of the software that performs a correct summation for Per Capita tables.

September 2002

Summary File 3
Data Note 10

The SF 3 table PCT55 data for "Nonfamily householders," nonfamily householders "Not living alone," and "Other unrelated individuals" have been removed. These data were removed because some respondents who were tallied as nonfamily householders "Not living alone" should have been tallied as "Other unrelated individuals." In American FactFinder, the data have been replaced with the symbol "(E)." In the files on the Census Bureau's FTP site, the data have been replaced with the value 999999999. The correct data will appear in SF 4 table PCT153.

February 2003

Summary File 3
Data Note 11

FLORIDA

Summary File 3 state files for Florida contain missing data for the following three geographies:

88100US12349XX093,
88100US12349XX097, and
88100US12399XX111.

The Census Bureau has concluded that the three (3) geographies were tabulated with a result of Zero (0) population count and Zero (0) housing unit count and do not appear in the final summary file product.

This note is applicable to the following data products:

- Summary File 3 (SF 3) Florida state files available at the Census Bureau's FTP site. The Summary level is "State-5-digit ZCTA-County".
- SF 3 CD-ROMs (ASCII files only).
- Tables available on American FactFinder.

August 2004

Summary File 3

INDEX TO SUMMARY FILE 3 GEOGRAPHY NOTES

Note	Geographic area
1	Alaska
2	California
3	Connecticut
4	Florida
5	Georgia
6	Nebraska
7	Tennessee
8	Wisconsin

Summary File 3
Geography Note 1

Alaska: 02

Nelson Lagoon Alaska Native village statistical area (ANVSA) (AIANHH 7025) erroneously contains block 2010, census tract 1 (000100) in Aleutians East census area (01598), Aleutians East Borough (013). This block should have not been coded to any ANVSA (9999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 02-003

May 2001

Summary File 3
Geography Note 2

California: 06

Los Angeles city (FIPS code 44000) erroneously contains block 1011, census tract 4002.03 (400203) in East San Gabriel Valley CCD (FIPS code 90810), Los Angeles County (FIPS code 037), CA (FIPS code 06). This block should have been coded to the place Balance of East San Gabriel Valley CCD (FIPS code 99999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP side.

Internal Errata ID 06-001

May 2001

Summary File 3
Geography Note 3

Connecticut: 09

The place record, Balance of Milford town (FIPS code 99999) erroneously contains block 2999, census tract 1502 (150200) in Milford town (FIPS code 47535), New Haven County (FIPS code 009), CT (FIPS code 09). This block should have been coded to place Milford city (balance) (FIPS code 47515). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 09-001

May 2001

Summary File 3
Geography Note 4

Florida: 12

Yeehaw Junction CDP (FIPS code 78975) in St. Cloud CCD (FIPS code 93029), Osceola County (FIPS code 097), FL (FIPS code 12) should be named Buenaventura Lakes with FIPS code 09415. In 1990, this area was named Buena Ventura Lakes (FIPS code 09415). The area that should have been Yeehaw Junction CDP was erroneously not defined and does not appear in any Census 2000 products.

Internal Errata ID 12-001

May 2001

Summary File 3
Geography Note 5

Georgia: 13

The place record Balance of Athens CCD (FIPS code 99999) erroneously contains blocks 2021 and 2023, census tract 1305 (130500) in Athens CCD (FIPS code 90138), Clarke County (FIPS code 059). Both blocks should have been coded to Bogart town (FIPS code 09068).

The place record Balance of Winterville CCD (FIPS code 99999) erroneously contains blocks 1008 and 1009, census tract 1406 (140600) in Winterville CCD (93402), Clarke County (FIPS code 059). Both blocks should have been coded to the place Athens-Clarke County (balance) (FIPS code 03440). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 13-001

May 2001

Summary File 3
Geography Note 6

Nebraska: 31

In the PL 94-171 and Summary File (SF) data products, Cisco CDP (FIPS code 09112) in Lisco precinct (FIPS code 91790), Garden County (FIPS code 069), NE (FIPS code 31) should be named Lisco with FIPS code of 28315.

Internal Errata ID 31-002

May 2001

Summary File 3
Geography Note 7

Tennessee: 47

The place record Balance of Metropolitan Government CCD (FIPS code 99999) erroneously contains blocks 1001 and 1008, census tract 171 (017100) in Metropolitan Government CCD (FIPS code 92200), Davidson County (FIPS code 037), TN (FIPS code 47). Both blocks should have been coded to place Nashville-Davidson (balance) (FIPS code 52006). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 47-001

May 2001

Summary File 3
Geography Note 8

Wisconsin: 55

The county subdivision of Scott town (FIPS code 72200), in place Balance of Scott town (FIPS code 99999) erroneously contains blocks 2048, 2063, and 2064, census tract 203 (020300), Brown County (FIPS code 009), WI (FIPS code 55). These blocks should have been coded to county subdivision and place Pulaski village (FIPS code 65675).

The county subdivision of Pittsfield town (FIPS code 63075), in place Balance of Pittsfield town (FIPS code 99999) erroneously contains block 2049, census tract 203 (020300), Brown County (FIPS code 009). This block should have been coded to county subdivision and place Pulaski village (FIPS code 65675). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 55-001

May 2001

Summary File 3
Technical Documentation Note 1

APPENDIX B, DEFINITIONS OF SUBJECT CHARACTERISTICS

School enrollment and type of school

In the comparability section, the third sentence in the third paragraph was replaced. The sentence was corrected to read: "Most of the published enrollment figures referred to people 5 to 20 years old in the 1930 census, 5 to 24 in 1940, 5 to 29 in 1950, 5 to 34 in 1960, 3 to 34 in 1970, and 3 years old and over in 1980 and later years."

Gross rent as a percentage of household income in 1999

The second sentence in the first paragraph was corrected to read: "The ratio is computed separately for each unit and is rounded to the nearest whole percentage."

August 2002

Summary File 3
Technical Documentation Note 2

Chapter 8, Accuracy of the Data, was updated to reflect the fact that Tribal jurisdiction statistical areas were replaced for Census 2000 by entities called Oklahoma Tribal Statistical Areas.

October 2002

Summary File 3
Technical Documentation Note 3

Value, Price Asked was erroneously omitted from the list of aggregates subject to rounding on page B–69. The technical documentation has been corrected.

October 2002

Summary File 3
Technical Documentation Note 4

Cell 3 of Table HCT35B, Kitchen Facilities (Black or African American Alone Householder) in Chapter 6 and Chapter 7 was corrected to read “Lacking complete kitchen facilities.” instead of “Lacking complete plumbing facilities.”

November 2002

Summary File 3
Technical Documentation Note 5

Table HCT6, Tenure by Year Structure Built by Units in Structure, on page 7-453 was corrected to read “Renter occupied—Con.” instead of “Owner occupied—Con.”

November 2002

Summary File 3
Technical Documentation Note 6

The indentation of the “Management of companies and enterprises:” line of the Industry code list found in Appendix G was changed so that it is aligned with the “Administrative and support and waste management services:” line.

January 2003

Summary File 3
Technical Documentation Note 7

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

Summary File 3
Technical Documentation Note 8

The Language section of the Code List appendix had two spelling errors. They have been corrected to read as follows:

772	Tahitian
971	OTO-MANGUEAN

September 2003

Summary File 3
Technical Documentation Note 9

The "Accuracy of the Data" chapter describes how to calculate standard errors for most estimates, but not for per capita income, which is described below.

Computing the Standard Error of Per Capita Income

Per capita income is the total income from all sources (salary income, retirement income, public assistance income, etc.) of the people in a population group divided by the number of people in that group.

$$Per\ Capita\ Income = \frac{Aggregate\ Income}{N_{Population}}$$

where $N_{Population}$ is the estimate of total people in the population group.

A similar statistic, mean income, is like per capita income, except that the population measure includes only people at least 15 years of age, since income data is not collected for people younger than that.

$$Mean\ Income = \frac{Aggregate\ Income}{N_{15+}}$$

where N_{15+} is the estimate of people at least 15 years old in the population group.

The two measures are related by the formula:

$$Per\ Capita\ Income = Mean\ Income \times \frac{N_{15+}}{N_{Population}}$$

Hence, the approximate formula for estimating the standard error of per capita income is:

$$SE\ (Per\ Capita\ Income) = SE\ (\ Mean\ Income\ ) \times \frac{N_{15+}}{N_{Population}}$$

Methodology

Calculating the standard error of Mean Income requires the use of an income distribution table. The table must provide frequency estimates of the number of people that fall within certain intervals. Standard available tables may be broken down by sex and whether the individual worked full-time, year-round in 1999. Such a table might look like this:

Table 1. **Sex by Work Experience in 1999 by Income in 1999 for the Population 15 Years and Over - Universe: Population 15 Years and Over**

Total	**32,091**
Male	15,836
Worked full-time, year-round in 1999:	6,000
No income	0
With income:	6,000
$1 to $2,499 or loss	10
$2,500 to $4,999	16
$5,000 to $7,499	44
$7,500 to $9,999	84
.	
.	
.	
$100,000 or more	146

Following the distribution for Male: Worked Full-Time, Year-Round in 1999 ("Wftyr") is a similar distribution for males who did not work full-time, year-round in 1999 (called "Other" in the table) and then females who did and did not work full-time, year-round in 1999.

1. To get the distribution of all people 15 years and older in each income interval, sum the four sex by work-status distributions:

$$N_{15+,j} = Male\_Wftyr_j + Male\_Other_j + Female\_Wftyr_j + Female\_Other_j$$
$$j = 1, 2, ..., number\ of\ intervals$$

2. Sum the frequencies across all intervals j to obtain an estimate of the population total:

$$N_{15+} = \sum_j N_{15+,j}$$

3. Calculate the estimated proportion of people in each income interval:

$$p_j = N_{15+,j} / N_{15+}$$

4. Calculate the mid-point (m) of each income interval from:

$$m_j = \left(L_j + U_j\right)/2$$

where L_j and U_j are the lower and upper bounds of the interval. If the c^{th} interval is open-ended (i.e. has no upper bound), then an approximate value for m_c is:

$$m_c = \frac{3}{2} L_c$$

5. Estimate mean income from:

$$\bar{x} = \sum_j p_j m_j$$

6. Estimate the standard error of mean income from:

$$SE(\bar{x}) = \sqrt{\frac{5}{N_{15+}} \times s^2} \times Design\ Factor$$

where

$$s^2 = \sum_j p_j \ m_j{}^2 - (\bar{x})^2$$

Use the design factor for "Population: Household Income in 1999."

7. An approximation of per capita income can be computed by:

$$Per\ Capita\ Income = \bar{x} \times \frac{N_{15+}}{N_{Population}}$$

8. Multiply the result of Step 6 by the ratio of the person estimates ($\frac{N_{15+}}{N_{Population}}$) to get the approximate standard error for per capita income.

Example

This example shows the steps to estimate the standard error of per capita income for a population group in County A.

1. Sum the frequency estimates in each interval in the four sub-tables of Table 1 to produce a distribution similar to Table 2.

Table 2. **Frequency Distribution for Income, People 15 years and older**

Total Income in 1999	Frequency
No income	8,034
With income:	
$1 to $2,499 or loss	644
$2,500 to $4,999	730
$5,000 to $7,499	876
$7,500 to $9,999	1,299
$10,000 to $12,499	1,350
$12,500 to $14,999	1,438
$15,000 to $17,499	1,599
$17,500 to $19,999	1,688
$20,000 to $22,499	1,871
$22,500 to $24,999	1,766
$25,000 to $29,999	2,331
$30,000 to $34,999	1,923
$35,000 to $39,999	1,345
$40,000 to $44,999	914
$45,000 to $49,999	856
$50,000 to $54,999	1,134
$55,000 to $64,999	828
$65,000 to $74,999	563
$75,000 to $99,999	455
$100,000 or more	447
Total	**32,091**

2. Cumulate the frequencies over the 21 intervals for those with and without income, to get the population base (N_{15+}) of 32,091 people age 15 years and over.

3. Calculate the proportion of people in each interval by dividing the interval's population estimate by the population base. The proportion of people age 15 and over in the "No Income" interval, p_1, is

$$p_1 = \frac{8{,}034}{32{,}091} = 0.2504 .$$

4. Find the midpoint m_j for each of the 21 intervals.

For example, the midpoint of interval 3, "$2,500 to $4,999" is

$$m_3 = \frac{\$2{,}500 + \$4{,}999}{2} = \$3{,}749.50$$

while the midpoint of the 21st interval, "$100,000 or more" is

$$m_{21} = \frac{3}{2}(\$100{,}000) = \$150{,}000$$

The midpoint of the "No Income" interval is zero; for "$1 to $2,499 or loss" it is $1,250. Necessary results for the standard error calculation are given in Table 3 along with totals.

Table 3. **Calculations for Per Capita Income**

Total Income in 1999	p_j	m_j	$p_j m_j^2$	$p_j m_j$
No Income	0.2504	$0.00	$0	$0.00
With Income				
$1 to $2,499 or loss	0.0201	$1,250.00	$31,406	$ 25.13
$2,500 to $4,999	0.0227	$3,749.50	$319,134	$ 85.11
$5,000 to $7,499	0.0273	$6,249.50	$1,066,236	$ 170.61
$7,500 to $9,999	0.0405	$8,749.50	$3,100,427	$ 354.35
$10,000 to $12,499	0.0421	$11,249.50	$5,327,808	$ 473.60
$12,500 to $14,999	0.0448	$13,749.50	$8,469,384	$ 615.98
$15,000 to $17,499	0.0498	$16,249.50	$13,149,503	$ 809.23
$17,500 to $19,999	0.0526	$18,749.50	$18,491,201	$ 986.22
$20,000 to $22,499	0.0583	$21,249.50	$26,324,855	$1,238.85
$22,500 to $24,999	0.0550	$23,749.50	$31,022,131	$1,306.22
$25,000 to $29,999	0.0726	$27,499.50	$54,901,754	$1,996.46
$30,000 to $34,999	0.0599	$32,499.50	$63,267,428	$1,946.72
$35,000 to $39,999	0.0419	$37,499.50	$58,920,304	$1,571.23
$40,000 to $44,999	0.0285	$42,499.50	$51,476,914	$1,211.24
$45,000 to $49,999	0.0267	$47,499.50	$60,240,607	$1,268.24
$50,000 to $54,999	0.0353	$52,499.50	$97,293,772	$1,853.23
$55,000 to $64,999	0.0258	$59,999.50	$92,878,452	$1,547.99
$65,000 to $74,999	0.0175	$69,999.50	$85,748,775	$1,224.99
$75,000 to $99,999	0.0142	$87,499.50	$108,717,508	$1,242.49
$100,000 or more	0.0139	$150,000.00	$312,750,000	$2,085.00
Total			$1,093,497,599	$22,013.00

5. To estimate mean income of people at least 15 years old in the population group in County A, multiply each interval's proportion by its midpoint and sum over all intervals in the universe. Table 3 shows an estimated mean income of people at least 15 years, $\overline{x}$, of $22,013

6. To estimate the standard error of mean income, first calculate the estimated population variance for mean income of people 15 years and older.

$$s^2 = 1{,}093{,}497{,}599 - 22{,}013^2 = 608{,}925{,}430$$

Suppose the person observed sampling rate in County A is 14.5 percent. Suppose the design factor for "Population Household Income in 1999", given in the "Less than 15 percent" percent-in-sample column of the design factor table in the technical documentation, is 1.4. Use this information and the above results to calculate an estimated standard error for the mean income of people 15 years and older as:

$$SE(\overline{x}) = \sqrt{\frac{5}{32{,}091} \times 608{,}925{,}430} \times 1.4$$

$$= \$431$$

Thus the standard error on the mean income of $22,013 is $431.

7. If the total population (including those less than 15 years old) in the population group in County A is 42,297, an approximation to per capita income is:

$$\$22{,}013 \times \frac{32{,}091}{42{,}297} = \$16{,}701$$

8. The standard error of the per capita income is calculated as:

$$SE(Per\ Capita\ Income) = \frac{32{,}091}{42{,}297} \times \$431 = \$327$$

Thus the standard error of the per capita income of $16,701 is $327.

March 2004

Summary File 4
Data Note 1

Users may encounter differences in Table PCT126, Median Nonfamily Household Income in 1999 (dollars), in Summary File 4 and the corresponding tables, P156A through P156I, in Summary File 3. Median incomes for nonfamily households by race in Summary File 3 were calculated from a 38-category income distribution rather than the standard 39-category income distribution. The 38-category distribution collapsed the two highest categories ($175,000 - $199,999 and $200,000 and over) into a single category of $175,000 and over. All medians in Summary File 4 were calculated from the following standard 39-catgeory income distribution.

Standard 39-category income distribution

Less than $2,500	$32,500 to $34,999	$65,000 to $67,499
$2,500 to $4,999	$35,000 to $37,499	$67,500 to $69,999
$5,000 to $7,499	$37,500 to $39,999	$70,000 to $72,499
$7,500 to $9,999	$40,000 to $42,499	$72,500 to $74,999
$10,000 to $12,499	$42,500 to $44,999	$75,000 to $79,999
$12,500 to $14,999	$45,000 to $47,499	$80,000 to $84,999
$15,000 to $17,499	$47,500 to $49,999	$85,000 to $89,999
$17,500 to $19,999	$50,000 to $52,499	$90,000 to $99,999
$20,000 to $22,499	$52,500 to $54,999	$100,000 to $124,999
$22,500 to $24,999	$55,000 to $57,499	$125,000 to $149,999
$25,000 to $27,499	$57,500 to $59,999	$150,000 to $174,999
$27,500 to $29,999	$60,000 to $62,499	$175,000 to $199,999
$30,000 to $32,499	$62,500 to $64,999	$200,000 or more

April 2003

Summary File 4
Data Note 2

COMPARING SF 4 ESTIMATES WITH CORRESPONDING VALUES IN SF 1 AND SF 2

As in earlier censuses, the responses from the sample of households reporting on long forms must be weighted to reflect the entire population. Specifically, each responding household represents, on average, six or seven other households who reported using short forms.

One consequence of the weighting procedures is that each estimate based on the long form responses has an associated confidence interval. These confidence intervals are wider (as a percentage of the estimate) for geographic areas with smaller populations and for characteristics that occur less frequently in the area being examined (such as the proportion of people in poverty in a middle-income neighborhood).

In order to release as much useful information as possible, statisticians must balance a number of factors. In particular, for Census 2000, the Bureau of the Census created weighting areas—geographic areas from which about two hundred or more long forms were completed—which are large enough to produce good quality estimates. If smaller weighting areas had been used, the confidence intervals around the estimates would have been significantly wider, rendering many estimates less useful due to their lower reliability.

The disadvantage of using weighting areas this large is that, for smaller geographic areas within them, the estimates of characteristics that are also reported on the short form will not match the counts reported in SF 1 or SF 2. Examples of these characteristics are the total number of people, the number of people reporting specific racial categories, and the number of housing units. The official values for items reported on the short form come from SF 1 and SF 2.

The differences between the long form estimates in SF 4 and values in SF 1 or SF 2 are particularly noticeable for the smallest places, tracts, and block groups. The long form estimates of total population and total housing units in SF 4 will, however, match the SF 1 and SF 2 counts for larger geographic areas such as counties and states, and will be essentially the same for medium and large cities.

This phenomenon also occurred for the 1990 Census, although in that case, the weighting areas included relatively small places. As a result, the long form estimates matched the short form counts for those places, but the confidence intervals around the estimates of characteristics collected only on the long form were often significantly wider (as a percentage of the estimate).

SF 1 gives exact numbers even for very small groups and areas; whereas, SF 4 gives estimates for small groups and areas such as tracts and small places that are less exact. The goal of SF 4 is to identify large differences among areas or large changes over time. Estimates for small areas and small population groups often do exhibit large changes from one census to the next, so having the capability to measure them is worthwhile.

April 2003

Summary File 4
Data Note 3 – Updated June 2004

In July 2002, the Census Bureau issued the following Data Note 4 regarding the Census 2000 Summary File 3 (SF3) data:

> The Census Bureau is aware there may be a problem or problems in the employment-status data of Census 2000 Summary File 3 (including tables P38, P43-46, P149A-1, P150A-I, PCT35, PCT69A-1, and PCT 70A-1). The labor force data for some places where colleges are located appear to overstate the number in the labor force, the number unemployed, and the percent unemployed, probably because of reporting or processing errors. The exact cause is unknown, but the Census Bureau will continue to research the problem.

The above issue relates to tables PCT66, PCT79-PCT81, PCT83, and PCT200 in SF 4. Our further research into this "college-town" issue indicates that the problem extended beyond places with colleges to the country in general. We learned that it stems from the tendency of many working-age people living in civilian noninstitutional group quarters (GQ), such as college dormitories, worker dormitories, and group homes (for the mentally ill or physically handicapped), to exhibit a particular pattern of entries to the employment questions in Census 2000.[1] We now estimate that the pattern affected the employment data for about 15 percent of the civilian noninstitutional GQ population 16 years of age and over in the United States, or around 500,000 people. It had an impact on the Census 2000 labor force statistics for the entire country, but its effects were most visible and substantial for places, such as college towns, with high concentrations of people living in civilian noninstitutional group quarters.

In Census 2000, the majority of people in the GQ population were enumerated by the Individual Census Report (ICR) form, which collected employment data in a battery of six questions (questions 23, 27a-e). The responses to these questions were captured and fed into a set of rules (called the Employment Status Recode (ESR) edit) that used the combined information from all six questions to assign each person to one of the following four employment-status categories: not in universe (all people less than 16 years old), employed, unemployed, and not in labor force.

For a significant segment of the GQ population, a so-called "3/3" response pattern was entered into the ESR edit.[2] This pattern is shown in the following table:

3/3 Input Pattern From ICR Forms

Question numberon ICR	Question wording	Entry
23	LAST WEEK, did you do ANY work for either pay or profit?	Missing
27a	LAST WEEK, were you on layoff from a job?	Missing
27b	LAST WEEK, were you TEMPORARILY absent from a job or business?	Missing
27c	(For people on layoff) Have you been informed that you will be recalled to work within the next 6 months OR been given a date to return to work?	Yes
27d	Have you been looking for work during the last four weeks?	Yes
27e	LAST WEEK, could you have started a job if offered one, or returned to work if recalled?	Yes

[1]The pattern also appeared frequently for people in institutional group quarters, such as prisons and juvenile institutions, but because of the way employment categories are defined, it had no impact on the employment data for these people.

[2]"3/3" refers to the fact that the responses to the first three questions, which appeared on page 4 of the ICR, are all missing; and those responses to the last three questions, which were on page 5 of the ICR, are all "yes."

The 3/3 pattern represents an incomplete set of information, since entries to the first three questions are missing. The ESR edit assigned people with this pattern to the "unemployed" category, because the edit had three built-in assumptions:

1) The respondents saw and reacted to each and every question in the employment series;

2) The 3/3 pattern represented the faithful recording of actual responses (or non-responses) to the questions; and

3) People who responded in this manner were more likely to meet the official criteria for the "unemployed" category than for any other category.[3]

Our research has revealed that most of the GQ cases with the 3/3 pattern may not have met one of the first two assumptions. We are still investigating, but we think that, in most cases, the pattern resulted from anomalies in the data collection or processing systems. Unfortunately, we cannot test our hypothesis by comparing the 3/3 pattern with actual reports from the respondents. The images of the filled-out ICR's will not be accessible until the completion, in 2006 at the earliest, of the Census Bureau's project to image the forms for delivery to the National Archives.

The potential effect of the ESR outcome for the 3/3 pattern is to increase the count of unemployed people at the expense of the counts of the employed and the not-in-labor-force groups. We have done some research to estimate the potential impact of the phenomenon on the labor force data for the nation as a whole. Our preliminary estimates are that it may have incorrectly decreased the number of employed people by about 235,000 (the number of employed in SF3 was 129.7 million), reduced the number of people not in the labor force by 285,000 (SF3 figure of 78.3 million), increased the number of unemployed by 519,000 (SF3 figure of 7.9 million), and raised the unemployment rate by 0.4 percentage point (SF3 figure was 5.8 percent).

Comparatively, the impact of the phenomenon on areas below the national level may be much greater, depending upon the relative size of the GQ population within the given area. The Census 2000 unemployment rate for the city of Williamsburg, Virginia, for example, was 41.7 percent (our research indicated that this rate resulted primarily from the prevalence of the 3/3 pattern among residents of college dormitories, who make up a large percentage of the city's population). To help data users gauge the impact of the phenomenon on their applications, and possibly to adjust for it, the Census Bureau released a tabulation of employment-status data for the nation, states, counties, and places, that was restricted to the population residing in households. This tabulation is available at: http://www.census.gov/hhes/www/laborfor.html

We will continue our research and report on further findings as they become available.

[3]They reported that they were looking for work and could have started a job last week. Because they did not report whether they had a job last week (people with a job are classified as "employed"), it is reasonable to classify them as "unemployed."

June 2004

Summary File 4

INDEX TO SUMMARY FILE 4 GEOGRAPHY NOTES

Note	Geographic area
1	Alaska
2	California
3	Connecticut
4	Florida
5	Georgia
6	Nebraska
7	Tennessee
8	Wisconsin

Summary File 4
Geography Note 1

Alaska: 02

Nelson Lagoon Alaska Native village statistical area (ANVSA) (AIANHH 7025) erroneously contains block 2010, census tract 1 (000100) in Aleutians East census area (01598), Aleutians East Borough (013). This block should have not been coded to any ANVSA (9999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 02-003

July 2003

Summary File 4
Geography Note 2

California: 06

Los Angeles city (FIPS code 44000) erroneously contains block 1011, census tract 4002.03 (400203) in East San Gabriel Valley CCD (FIPS code 90810), Los Angeles County (FIPS code 037), CA (FIPS code 06). This block should have been coded to the place Balance of East San Gabriel Valley CCD (FIPS code 99999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP side.

Internal Errata ID 06-001

July 2003

Summary File 4
Geography Note 3

Connecticut: 09

The place record, Balance of Milford town (FIPS code 99999) erroneously contains block 2999, census tract 1502 (150200) in Milford town (FIPS code 47535), New Haven County (FIPS code 009), CT (FIPS code 09). This block should have been coded to place Milford city (balance) (FIPS code 47515). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 09-001

July 2003

Summary File 4
Geography Note 4

Florida: 12

Yeehaw Junction CDP (FIPS code 78975) in St. Cloud CCD (FIPS code 93029), Osceola County (FIPS code 097), FL (FIPS code 12) should be named Buenaventura Lakes with FIPS code 09415. In 1990, this area was named Buena Ventura Lakes (FIPS code 09415). The area that should have been Yeehaw Junction CDP was erroneously not defined and does not appear in any Census 2000 products.

Internal Errata ID 12-001

July 2003

Summary File 4
Geography Note 5

Georgia: 13

The place record Balance of Bogart CCD (FIPS code 99999) erroneously contains blocks 2021 and 2023, census tract 1305 (130500) in Bogart CCD (FIPS code 90138), Clarke County (FIPS code 059). Both blocks should have been coded to Bogart town (FIPS code 09068).

The place record Balance of Winterville CCD (FIPS code 99999) erroneously contains blocks 1008 and 1009, census tract 1406 (140600) in Winterville CCD (93402), Clarke County (FIPS code 059). Both blocks should have been coded to the place Athens-Clarke County (balance) (FIPS code 03440). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 13-001

July 2003

Summary File 4
Geography Note 6

Nebraska: 31

In the PL 94-171 and Summary File (SF) data products, Cisco CDP (FIPS code 09112) in Lisco precinct (FIPS code 91790), Garden County (FIPS code 069), NE (FIPS code 31) should be named Lisco with FIPS code of 28315.

Internal Errata ID 31-002

July 2003

Summary File 4
Geography Note 7

Tennessee: 47

The place record Balance of Metropolitan Government CCD (FIPS code 99999) erroneously contains blocks 1001 and 1008, census tract 171 (017100) in Metropolitan Government CCD (FIPS code 92200), Davidson County (FIPS code 037), TN (FIPS code 47). Both blocks should have been coded to place Nashville-Davidson (balance) (FIPS code 52006). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 47-001

July 2003

Summary File 4
Geography Note 8

Wisconsin: 55

The county subdivision of Scott town (FIPS code 72200), in place Balance of Scott town (FIPS code 99999) erroneously contains blocks 2048, 2063, and 2064, census tract 203 (020300), Brown County (FIPS code 009), WI (FIPS code 55). These blocks should have been coded to county subdivision and place Pulaski village (FIPS code 65675).

The county subdivision of Pittsfield town (FIPS code 63075), in place Balance of Pittsfield town (FIPS code 99999) erroneously contains block 2049, census tract 203 (020300), Brown County (FIPS code 009). This block should have been coded to county subdivision and place Pulaski village (FIPS code 65675). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 55-001

July 2003

Summary File 4
Technical Documentation Note 1

In Chapter 4, Summary Level Sequence Chart, the geographic components for summary level 040-State, were corrected to read 00,01-49,52-95.

June 2003

Summary File 4
Technical Documentation Note 2

The following line was erroneously omitted from the Table (Matrix) Section of Chapter 7 on page 7–101:

Table number	Table contents	Data dictionary reference name	Max size	Data type
PCT85	Manufacturing	PCT085007	14	9

June 2003

Summary File 4
Technical Documentation Note 3

The text, “Construction, extraction, and maintenance occupations:” on page 7-112 of the Table (Matrix) Section in the Data Dictionary was changed to read “Production, transportation, and material moving occupations—Con.”

June 2003

Summary File 4
Technical Documentation Note 4

In Summary File 4 (SF 4) table PCT86, for the line "Construction traders workers except carpenters, electricians, painters, plumbers, and construction laborers," "traders" was changed to "trades." This correction was made to both Chapter 6 and Chapter 7 of the SF 4 technical documentation, as well as to the American FactFinder detailed table presentation of PCT86.

July 2003

Summary File 4
Technical Documentation Note 5

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

Summary File 4
Technical Documentation Note 6

The Language section of the Code List appendix had two spelling errors. They have been corrected to read as follows:

772	Tahitian
971	OTO-MANGUEAN

September 2003

Summary File 4
Technical Documentation Note 7

In Chapter 9, Geography Note 5 for Georgia erroneously referred to Athens instead of Bogart. The technical documentation has been corrected.

September 2003

Summary File 4
Technical Documentation Note 8

The "Accuracy of the Data" chapter describes how to calculate standard errors for most estimates, but not for per capita income, which is described below.

Computing the Standard Error of Per Capita Income

Per capita income is the total income from all sources (salary income, retirement income, public assistance income, etc.) of the people in a population group divided by the number of people in that group.

$$Per\ Capita\ Income = \frac{Aggregate\ Income}{N_{Population}}$$

where $N_{Population}$ is the estimate of total people in the population group.

A similar statistic, mean income, is like per capita income, except that the population measure includes only people at least 15 years of age, since income data is not collected for people younger than that.

$$Mean\ Income = \frac{Aggregate\ Income}{N_{15+}}$$

where N_{15+} is the estimate of people at least 15 years old in the population group.

The two measures are related by the formula:

$$Per\ Capita\ Income = Mean\ Income \times \frac{N_{15+}}{N_{Population}}$$

Hence, the approximate formula for estimating the standard error of per capita income is:

$$SE\ (Per\ Capita\ Income) = SE\ (\ Mean\ Income) \times \frac{N_{15+}}{N_{Population}}$$

Methodology

Calculating the standard error of Mean Income requires the use of an income distribution table. The table must provide frequency estimates of the number of people that fall within certain intervals. Standard available tables may be broken down by sex and whether the individual worked full-time, year-round in 1999. Such a table might look like this:

Table 1. **Sex by Work Experience in 1999 by Income in 1999 for the Population 15 Years and Over - Universe: Population 15 Years and Over**

Total	**32,091**
Male	15,836
Worked full-time, year-round in 1999:	6,000
No income	0
With income:	6,000
$1 to $2,499 or loss	10
$2,500 to $4,999	16
$5,000 to $7,499	44
$7,500 to $9,999	84
.	
.	
.	
$100,000 or more	146

Following the distribution for Male: Worked Full-Time, Year-Round in 1999 ("Wftyr") is a similar distribution for males who did not work full-time, year-round in 1999 (called "Other" in the table) and then females who did and did not work full-time, year-round in 1999.

1. To get the distribution of all people 15 years and older in each income interval, sum the four sex by work-status distributions:

$$N_{15+,j} = Male\_Wftyr_j + Male\_Other_j + Female\_Wftyr_j + Female\_Other_j$$
$$j = 1, 2, ..., number\ of\ intervals$$

2. Sum the frequencies across all intervals j to obtain an estimate of the population total:

$$N_{15+} = \sum_j N_{15+,j}$$

3. Calculate the estimated proportion of people in each income interval:

$$p_j = N_{15+,j} / N_{15+}$$

4. Calculate the mid-point (m) of each income interval from:

$$m_j = \left(L_j + U_j\right)/2$$

where L_j and U_j are the lower and upper bounds of the interval. If the c^{th} interval is open-ended (i.e. has no upper bound), then an approximate value for m_c is:

$$m_c = \frac{3}{2} L_c$$

5. Estimate mean income from:

$$\bar{x} = \sum_j p_j m_j$$

6. Estimate the standard error of mean income from:

$$SE(\bar{x}) = \sqrt{\frac{5}{N_{15+}} \times s^2} \times Design\ Factor$$

where

$$s^2 = \sum_j p_j \ m_j{}^2 - (\bar{x})^2$$

Use the design factor for "Population: Household Income in 1999."

7. An approximation of per capita income can be computed by:

$$Per\ Capita\ Income = \bar{x} \times \frac{N_{15+}}{N_{Population}}$$

8. Multiply the result of Step 6 by the ratio of the person estimates ($\frac{N_{15+}}{N_{Population}}$) to get the approximate standard error for per capita income.

Example

This example shows the steps to estimate the standard error of per capita income for a population group in County A.

1. Sum the frequency estimates in each interval in the four sub-tables of Table 1 to produce a distribution similar to Table 2.

Table 2. **Frequency Distribution for Income, People 15 years and older**

Total Income in 1999	Frequency
No income	8,034
With income:	
$1 to $2,499 or loss	644
$2,500 to $4,999	730
$5,000 to $7,499	876
$7,500 to $9,999	1,299
$10,000 to $12,499	1,350
$12,500 to $14,999	1,438
$15,000 to $17,499	1,599
$17,500 to $19,999	1,688
$20,000 to $22,499	1,871
$22,500 to $24,999	1,766
$25,000 to $29,999	2,331
$30,000 to $34,999	1,923
$35,000 to $39,999	1,345
$40,000 to $44,999	914
$45,000 to $49,999	856
$50,000 to $54,999	1,134
$55,000 to $64,999	828
$65,000 to $74,999	563
$75,000 to $99,999	455
$100,000 or more	447
Total	**32,091**

2. Cumulate the frequencies over the 21 intervals for those with and without income, to get the population base (N_{15+}) of 32,091 people age 15 years and over.

3. Calculate the proportion of people in each interval by dividing the interval's population estimate by the population base. The proportion of people age 15 and over in the "No Income" interval, p_1, is

$$p_1 = \frac{8{,}034}{32{,}091} = 0.2504 .$$

4. Find the midpoint m_j for each of the 21 intervals.

For example, the midpoint of interval 3, "$2,500 to $4,999" is

$$m_3 = \frac{\$2{,}500 + \$4{,}999}{2} = \$3{,}749.50$$

while the midpoint of the 21st interval, "$100,000 or more" is

$$m_{21} = \frac{3}{2}(\$100{,}000) = \$150{,}000$$

The midpoint of the "No Income" interval is zero; for "$1 to $2,499 or loss" it is $1,250. Necessary results for the standard error calculation are given in Table 3 along with totals.

Table 3. **Calculations for Per Capita Income**

Total Income in 1999	p_j	m_j	$p_j m_j^2$	$p_j m_j$
No Income	0.2504	$0.00	$0	$0.00
With Income				
$1 to $2,499 or loss	0.0201	$1,250.00	$31,406	$ 25.13
$2,500 to $4,999	0.0227	$3,749.50	$319,134	$ 85.11
$5,000 to $7,499	0.0273	$6,249.50	$1,066,236	$ 170.61
$7,500 to $9,999	0.0405	$8,749.50	$3,100,427	$ 354.35
$10,000 to $12,499	0.0421	$11,249.50	$5,327,808	$ 473.60
$12,500 to $14,999	0.0448	$13,749.50	$8,469,384	$ 615.98
$15,000 to $17,499	0.0498	$16,249.50	$13,149,503	$ 809.23
$17,500 to $19,999	0.0526	$18,749.50	$18,491,201	$ 986.22
$20,000 to $22,499	0.0583	$21,249.50	$26,324,855	$1,238.85
$22,500 to $24,999	0.0550	$23,749.50	$31,022,131	$1,306.22
$25,000 to $29,999	0.0726	$27,499.50	$54,901,754	$1,996.46
$30,000 to $34,999	0.0599	$32,499.50	$63,267,428	$1,946.72
$35,000 to $39,999	0.0419	$37,499.50	$58,920,304	$1,571.23
$40,000 to $44,999	0.0285	$42,499.50	$51,476,914	$1,211.24
$45,000 to $49,999	0.0267	$47,499.50	$60,240,607	$1,268.24
$50,000 to $54,999	0.0353	$52,499.50	$97,293,772	$1,853.23
$55,000 to $64,999	0.0258	$59,999.50	$92,878,452	$1,547.99
$65,000 to $74,999	0.0175	$69,999.50	$85,748,775	$1,224.99
$75,000 to $99,999	0.0142	$87,499.50	$108,717,508	$1,242.49
$100,000 or more	0.0139	$150,000.00	$312,750,000	$2,085.00
Total			$1,093,497,599	$22,013.00

5. To estimate mean income of people at least 15 years old in the population group in County A, multiply each interval's proportion by its midpoint and sum over all intervals in the universe. Table 3 shows an estimated mean income of people at least 15 years, $\overline{x}$, of $22,013

6. To estimate the standard error of mean income, first calculate the estimated population variance for mean income of people 15 years and older.

$$s^2 = 1{,}093{,}497{,}599 - 22{,}013^2 = 608{,}925{,}430$$

Suppose the person observed sampling rate in County A is 14.5 percent. Suppose the design factor for "Population Household Income in 1999", given in the "Less than 15 percent" percent-in-sample column of the design factor table in the technical documentation, is 1.4. Use this information and the above results to calculate an estimated standard error for the mean income of people 15 years and older as:

$$SE(\overline{x}) = \sqrt{\frac{5}{32{,}091} \times 608{,}925{,}430} \times 1.4$$

$$= \$431$$

Thus the standard error on the mean income of $22,013 is $431.

7. If the total population (including those less than 15 years old) in the population group in County A is 42,297, an approximation to per capita income is:

$$\$22{,}013 \times \frac{32{,}091}{42{,}297} = \$16{,}701$$

8. The standard error of the per capita income is calculated as:

$$SE(Per\ Capita\ Income) = \frac{32{,}091}{42{,}297} \times \$431 = \$327$$

Thus the standard error of the per capita income of $16,701 is $327.

March 2004

Summary File 4
Technical Documentation Note 9

Table PCT125, Nonfamily Household Income in 1999, on page 7-135 was corrected to read "Universe: Nonfamily households" (the word "universe" was added). "Max size" was corrected to read "9."

May 2004

108th Congressional District Summary File (100-Percent) Data Note 1

This user update is described on our Web site (www.census.gov) as:

Technical Note on Same-Sex Unmarried Partner Data From the 1990 and 2000 Censuses

The release of data in the SF 1 files from the 2000 census has brought with it a number of analyses documenting change that has occurred since the last census was conducted in 1990. While many of the variables and processes between the two censuses are comparable, some are not, and direct comparison of some estimates may lead to misleading conclusions. This note discusses one such topic, that of "unmarried partners," and advises that for some analyses — those involving unmarried same-sex partners — direct comparison of the 1990 and 2000 estimates is not substantively valid.

The household relationship item in both the 1990 and the 2000 censuses offered many ways of identifying how other people in the household were related to the householder (the person in whose name the house is owned or rented). Categories included spouse, child or other relative of the householder, housemate/roommate, roomer/boarder, and unmarried partner. In all circumstances, the respondent was asked to choose the category that best represented how other members of the household were related to the householder.

In both censuses, the "spouse" and "unmarried partner" response categories were defined and asked the same way. However, there were important differences in data processing that mean that some of the data are not comparable, limiting the usefulness of comparisons of the number of same-sex unmarried partners between these two censuses.

In both censuses, if a person was identified as the "spouse" of the householder and was the same sex as the householder, the "spouse" response was flagged for further review and allocation, that is, assignment of a value other than that originally reported, based on other data on the form. In 1990, the edit and allocation procedures did not allow same-sex "spouse" combinations to occur, thus resulting in the allocation of one of these two items in order to achieve editing consistency among the responses.

Processing steps were changed for Census 2000 for households that contained same-sex "spouses." If the person with the "spouse" category was the same sex as the householder and if neither person had their sex previously allocated, a relationship response of "spouse" was allocated as an "unmarried partner" response. Since marital status was no longer on the short form, its given value could not be considered (or modified) in this allocation procedure as it had been in 1990.

Data allocation is a standard statistical practice that is followed by most data collection agencies. Data on the relationship item (as other items) were subject to allocation in the census, as they are in virtually all Census Bureau surveys. In 1990, the marital status item was available on the 100 percent (short) form and aided in both the evaluation of the consistency of responses between the householder and the "spouse," and in the subsequent allocation procedure. The 1990 procedure allocated responses via a statistical model that distributed allocated responses from answers given by respondents in a proximate geographic area. This procedure used key demographic data from the census form, including marital status, as stratifying factors to provide a reasonable distribution of allocated responses. This procedure, while ensuring that no same-sex spouse response could be subsequently allocated, produced a set of allocated responses that could have included an "unmarried partner" response as well as any other response that was consistent with the age/sex/marital status profile of the respondent. This would include being allocated as a sibling or a relative, for example, or if the age differences were far enough apart (15 or more years), even a parent or child of the householder.

Three principal factors affected our decision to take this approach for Census 2000.

1. Same-sex spouse responses were flagged as invalid to comply with the 1996 Federal Defense of Marriage Act (H.R. 3396) passed by the 104th Congress. This act instructs all federal agencies only to recognize opposite-sex marriages for the purposes of enacting any agency programs. In order for Census Bureau data to be consistent with this act and the data requirements of other federal agencies, same-sex spouse responses were invalidated. The legislation defines marriage and spouse as follows:

   "In determining the meaning of any Act of Congress, or of any ruling, regulation or interpretation of the various administrative bureaus and agencies of the United States, the word 'marriage' means only a legal union between one man and one woman as husband and wife, and the word 'spouse' refers only to a person of the opposite sex who is a husband or wife."

   In order for the Census Bureau to be consistent with this act and the data requirements of other federal agencies, same-sex "spouse" responses were invalidated.

2. The second issue was statistical in nature. The principal basis of any good statistical allocation routine rests on the selection of the stratifying or input factors to provide a good statistical model. Without marital status data on the 100 percent form in Census 2000, the allocation routine would be relatively weak. Since many partners are roughly the same age, a statistical routine without marital status as one of its factors would have likely resulted in an overestimate of adult siblings or relatives, as the majority of people living in households are relatives, and this is the population from which we would draw our allocated responses. Additionally, if the same-sex partners were more than 15 years difference in age, the statistical routine would have likely allocated the invalidated "spouse" response as either a "child" or "parent" of the householder, as these types of relatives predominate in households in this age range of differences. This was an unacceptable outcome, as it would actually destroy the intent of the original "spouse" response, which clearly indicated a nonparental type of relationship. It should be noted that the "spouse" response on the form is assumed to be deliberate — not accidental — as it was the first response category on the question and was not placed between other possible response categories that may have been meant to be marked, such as housemates or roomers.

3. The third factor took into consideration that couples in long term same-sex relationships may consider themselves as "married partners" and thus respond as such on the census form. In addition, at the time of writing the editing program for Census 2000, there were several challenges in the courts concerning the legality of same-sex marriages. Clearly, we could not ignore the fact that same-sex spouse responses were going to be recorded during Census 2000. In light of these social and legal aspects — and the lack of a key variable in the statistical allocation routine (marital status) — the assignment of same-sex "married" couples to the same-sex "unmarried partner" category was the procedure chosen for the editing process. We were adverse to a randomized allocation of these responses after people had clearly marked a close relationship preference on the census form.

   As a result of these changes in the processing routine, estimates of same-sex unmarried partners are not comparable between the 1990 and 2000 census. We believe 2000 census estimates of this category are better estimates than those produced in 1990. It should also be noted that estimates of opposite-sex unmarried partners, however, were not affected by these editing procedures and changes and are comparable between the two censuses.

For further information on this topic, please contact the Fertility and Family Statistics Branch on 301-457-2416.

March 2003

108th Congressional District Summary File (100-Percent)

INDEX TO 108TH CONGRESSIONAL DISTRICT SUMMARY FILE (100-PERCENT) GEOGRAPHY NOTES

Note	Geographic area
1	Alaska
2	California
3	Connecticut
4	Florida
5	Georgia
6	Nebraska
7	Tennessee
8	Wisconsin
9	Selected States

108th Congressional District Summary File (100-Percent)
Geography Note 1

Alaska: 02

Nelson Lagoon Alaska Native village statistical area (ANVSA) (AIANHH 7025) erroneously contains block 2010, census tract 1 (000100) in Aleutians East census area (01598), Aleutians East Borough (013). This block should have not been coded to any ANVSA (9999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 02-003

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 2

California: 06

Los Angeles city (FIPS code 44000) erroneously contains block 1011, census tract 4002.03 (400203) in East San Gabriel Valley CCD (FIPS code 90810), Los Angeles County (FIPS code 037), CA (FIPS code 06). This block should have been coded to the place Balance of East San Gabriel Valley CCD (FIPS code 99999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP side.

Internal Errata ID 06-001

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 3

Connecticut: 09

The place record, Balance of Milford town (FIPS code 99999) erroneously contains block 2999, census tract 1502 (150200) in Milford town (FIPS code 47535), New Haven County (FIPS code 009), CT (FIPS code 09). This block should have been coded to place Milford city (balance) (FIPS code 47515). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 09-001

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 4

Florida: 12

Yeehaw Junction CDP (FIPS code 78975) in St. Cloud CCD (FIPS code 93029), Osceola County (FIPS code 097), FL (FIPS code 12) should be named Buenaventura Lakes with FIPS code 09415. In 1990, this area was named Buena Ventura Lakes (FIPS code 09415). The area that should have been Yeehaw Junction CDP was erroneously not defined and does not appear in any Census 2000 products.

Internal Errata ID 12-001

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 5

Georgia: 13

The place record Balance of Athens CCD (FIPS code 99999) erroneously contains blocks 2021 and 2023, census tract 1305 (130500) in Athens CCD (FIPS code 90138), Clarke County (FIPS code 059). Both blocks should have been coded to Bogart town (FIPS code 09068).

The place record Balance of Winterville CCD (FIPS code 99999) erroneously contains blocks 1008 and 1009, census tract 1406 (140600) in Winterville CCD (93402), Clarke County (FIPS code 059). Both blocks should have been coded to the place Athens-Clarke County (balance) (FIPS code 03440). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 13-001

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 6

Nebraska: 31

In the PL 94-171 and Summary File (SF) data products, Cisco CDP (FIPS code 09112) in Lisco precinct (FIPS code 91790), Garden County (FIPS code 069), NE (FIPS code 31) should be named Lisco with FIPS code of 28315.

Internal Errata ID 31-002

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 7

Tennessee: 47

The place record Balance of Metropolitan Government CCD (FIPS code 99999) erroneously contains blocks 1001 and 1008, census tract 171 (017100) in Metropolitan Government CCD (FIPS code 92200), Davidson County (FIPS code 037), TN (FIPS code 47). Both blocks should have been coded to place Nashville-Davidson (balance) (FIPS code 52006). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 47-001

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 8

Wisconsin: 55

The county subdivision of Scott town (FIPS code 72200), in place Balance of Scott town (FIPS code 99999) erroneously contains blocks 2048, 2063, and 2064, census tract 203 (020300), Brown County (FIPS code 009), WI (FIPS code 55). These blocks should have been coded to county subdivision and place Pulaski village (FIPS code 65675).

The county subdivision of Pittsfield town (FIPS code 63075), in place Balance of Pittsfield town (FIPS code 99999) erroneously contains block 2049, census tract 203 (020300), Brown County (FIPS code 009). This block should have been coded to county subdivision and place Pulaski village (FIPS code 65675). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 55-001

March 2003

108th Congressional District Summary File (100-Percent)
Geography Note 9

The 108th Congressional Districts appearing in Census Bureau products reflect the information provided to the Census Bureau by the states. Several states did not provide congressional district codes for census blocks in water areas. This only occurred in large bodies of water, such as the Great Lakes, Chesapeake Bay, and Long Island Sound. These water areas have a blank congressional district code and the 108th Congressional District data associated with these areas may show records with a "part" indicator. The "part" designation indicates either of two situations. It can indicate that there is a second record (part) in the dataset. Or, in the case of these water areas with blank congressional district codes, the "part" indicates that the record does not represent the entirety of the geographic entity.

April 2003

108th Congressional District Summary File (100-Percent) Technical Documentation Note 1

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

108th Congressional District Summary File (Sample)
Data Note 1

On the Census 2000 long-form questionnaire, individuals could report more than one type of disability. Summary File 3 Table P41, Age by Types of Disability for the Civilian Noninstitutionalized Population 5 Years and Over With Disabilities, has as its universe the total disabilities tallied. Each line of the table represents the number of occurrences of a particular disability, and the numbers should be interpreted with care. For example, the second line of data in the table titled "Total disabilities tallied for people 5 to 15 years" does not refer to the number of people 5 to 15 years old, or to the number of people 5 to 15 with a disability. Rather it is the sum of the number of all disabilities reported among the 5 to 15 year old population. Lines in the table referencing specific disabilities are more easily interpreted. The third line in the table titled "Sensory disability," for example, refers to the number of sensory disabilities reported among people 5 to 15 years (or the number of people 5 to 15 years old with a sensory disability).

Data users wanting to know the percent of civilian noninstitutionalized people 5 to 15 years old with, for example, a sensory disability should divide line 3 from Table P41 with the sum of lines 3 and 27 from Table P42, Sex by Age by Disability Status by Employment Status for the Civilian Noninstitutionalized Population 5 Years and Over. Data users wanting to know the same percentages for one of the nine race or Hispanic or Latino origin groups should use Tables PCT67A-I and Tables PCT68A-I, as appropriate.

March 2003

108th Congressional District Summary File (Sample)
Data Note 2

Users may find slight differences in aggregate earnings for households between the Demographic Profile and Summary File 3 and related products. These differences are due to the treatment of offsetting positive and negative amounts for household members. Whenever offsetting values occurred, the Demographic Profile assigned these households a value zero while Summary File 3 and related products assigned a value of one dollar. The assignment of one dollar allows users to distinguish those households that had earnings from those households that did not have earnings. This will have little effect, if any, on mean household earnings.

March 2003

108th Congressional District Summary File (Sample)
Data Note 3

Users may find slight differences in the Occupants Per Room calculations between the Demographic Profile and Summary File 3, Summary File 4, and related products. “Occupants per room” is obtained by dividing the number of people in each occupied housing unit by the number of rooms in the unit. The Summary File 3 products correctly used a topcode value of “10 rooms” for those occupied housing units with “9 or more rooms.” In the Demographic Profiles, an incorrect topcode value of “9 rooms” was used.

March 2003

108th Congressional District Summary File (Sample)
Data Note 4

In July 2002, the Census Bureau issued the following Data Note 4 regarding the Census 2000 Summary File 3 (SF3) data:

> The Census Bureau is aware there may be a problem or problems in the employment-status data of Census 2000 Summary File 3 (including tables P38, P43-46, PCT35, P149A-1, P150A-I, PCT35, PCT69A-1, and PCT 70A-1). The labor force data for some places where colleges are located appear to overstate the number in the labor force, the number unemployed, and the percent unemployed, probably because of reporting or processing errors. The exact cause is unknown, but the Census Bureau will continue to research the problem.

Our further research into this "college-town" issue indicates that the problem extended beyond places with colleges to the country in general. We learned that it stems from the tendency of many working-age people living in civilian noninstitutional group quarters (GQ), such as college dormitories, worker dormitories, and group homes (for the mentally ill or physically handicapped), to exhibit a particular pattern of entries to the employment questions in Census 2000.[1] We now estimate that the pattern affected the employment data for about 15 percent of the civilian noninstitutional GQ population 16 years of age and over in the United States, or around 500,000 people. It had an impact on the Census 2000 labor force statistics for the entire country, but its effects were most visible and substantial for places, such as college towns, with high concentrations of people living in civilian noninstitutional group quarters.

In Census 2000, the majority of people in the GQ population were enumerated by the Individual Census Report (ICR) form, which collected employment data in a battery of six questions (questions 23, 27a-e). The responses to these questions were captured and fed into a set of rules (called the Employment Status Recode (ESR) edit) that used the combined information from all six questions to assign each person to one of the following four employment-status categories: not in universe (all people less than 16 years old), employed, unemployed, and not in labor force.

For a significant segment of the GQ population, a so-called "3/3" response pattern was entered into the ESR edit.[2] This pattern is shown in the following table:

3/3 Input Pattern From ICR Forms

Question numberon ICR	Question wording	Entry
23	LAST WEEK, did you do ANY work for either pay or profit?	Missing
27a	LAST WEEK, were you on layoff from a job?	Missing
27b	LAST WEEK, were you TEMPORARILY absent from a job or business?	Missing
27c	(For people on layoff) Have you been informed that you will be recalled to work within the next 6 months OR been given a date to return to work?	Yes
27d	Have you been looking for work during the last four weeks?	Yes
27e	LAST WEEK, could you have started a job if offered one, or returned to work if recalled?	Yes

[1] The pattern also appeared frequently for people in institutional group quarters, such as prisons and juvenile institutions, but because of the way employment categories are defined, it had no impact on the employment data for these people.

[2] "3/3" refers to the fact that the responses to the first three questions, which appeared on page 4 of the ICR, are all missing; and those responses to the last three questions, which were on page 5 of the ICR, are all "yes."

The 3/3 pattern represents an incomplete set of information, since entries to the first three questions are missing. The ESR edit assigned people with this pattern to the "unemployed" category, because the edit had three built-in assumptions:

1. The respondents saw and reacted to each and every question in the employment series;
2. The 3/3 pattern represented the faithful recording of actual responses (or nonresponses) to the questions; and
3. People who responded in this manner were more likely to meet the official criteria for the "unemployed" category than for any other category.[3]

Our research has revealed that most of the GQ cases with the 3/3 pattern may not have met one of the first two assumptions. We are still investigating, but we think that, in most cases, the pattern resulted from anomalies in the data collection or processing systems. Unfortunately, we cannot test our hypothesis by comparing the 3/3 pattern with actual reports from the respondents. The images of the filled-out ICR's will not be accessible until the completion, in 2006 at the earliest, of the Census Bureau's project to image the forms for delivery to the National Archives.

The potential effect of the ESR outcome for the 3/3 pattern is to increase the count of unemployed people at the expense of the counts of the employed and the not-in-labor-force groups. We have done some research to estimate the potential impact of the phenomenon on the labor force data for the nation as a whole. Our preliminary estimates are that it may have incorrectly decreased the number of employed people by about 235,000 (the number of employed in SF3 was 129.7 million), reduced the number of people not in the labor force by 285,000 (SF3 figure of 78.3 million), increased the number of unemployed by 519,000 (SF3 figure of 7.9 million), and raised the unemployment rate by 0.4 percentage point (SF3 figure was 5.8 percent).

Comparatively, the impact of the phenomenon on areas below the national level may be much greater, depending upon the relative size of the GQ population within the given area. The Census 2000 unemployment rate for the city of Williamsburg, Virginia, for example, was 41.7 percent. Our research indicated that this rate resulted primarily from the prevalence of the 3/3 pattern among residents of college dormitories, who make up a large percentage of the city's population.

We will continue our research and report on further findings as they become available.

[3]They reported that they were looking for work and could have started a job last week. Because they did not report whether they had a job last week (people with a job are classified as "employed"), it is reasonable to classify them as "unemployed."

March 2003

108th Congressional District Summary File (Sample)
Data Note 5

COMPARING SF 3 ESTIMATES WITH CORRESPONDING VALUES IN SF 1 AND SF 2

As in earlier censuses, the responses from the sample of households reporting on long forms must be weighted to reflect the entire population. Specifically, each responding household represents, on average, six or seven other households who reported using short forms.

One consequence of the weighting procedures is that each estimate based on the long form responses has an associated confidence interval. These confidence intervals are wider (as a percentage of the estimate) for geographic areas with smaller populations and for characteristics that occur less frequently in the area being examined (such as the proportion of people in poverty in a middle-income neighborhood).

In order to release as much useful information as possible, statisticians must balance a number of factors. In particular, for Census 2000, the Bureau of the Census created weighting areas—geographic areas from which about two hundred or more long forms were completed—which are large enough to produce good quality estimates. If smaller weighting areas had been used, the confidence intervals around the estimates would have been significantly wider, rendering many estimates less useful due to their lower reliability.

The disadvantage of using weighting areas this large is that, for smaller geographic areas within them, the estimates of characteristics that are also reported on the short form will not match the counts reported in SF 1 or SF 2. Examples of these characteristics are the total number of people, the number of people reporting specific racial categories, and the number of housing units. The official values for items reported on the short form come from SF 1 and SF 2.

The differences between the long form estimates in SF 3 and values in SF 1 or SF 2 are particularly noticeable for the smallest places, tracts, and block groups. The long form estimates of total population and total housing units in SF 3 will, however, match the SF 1 and SF 2 counts for larger geographic areas such as counties and states, and will be essentially the same for medium and large cities.

This phenomenon also occurred for the 1990 Census, although in that case, the weighting areas included relatively small places. As a result, the long form estimates matched the short form counts for those places, but the confidence intervals around the estimates of characteristics collected only on the long form were often significantly wider (as a percentage of the estimate).

SF 1 gives exact numbers even for very small groups and areas; whereas, SF 3 gives estimates for small groups and areas such as tracts and small places that are less exact. The goal of SF 3 is to identify large differences among areas or large changes over time. Estimates for small areas and small population groups often do exhibit large changes from one census to the next, so having the capability to measure them is worthwhile.

March 2003

108th Congressional District Summary File (Sample)
Data Note 6

The following new section was added to Chapter 8, Accuracy of the Data.

CONSISTENCY WITH COMPLETE COUNTS

As described earlier, Census 2000 long form data were collected on a sample basis. Cities and incorporated places were used to determine sampling rates to support estimates for these areas. As a result, each city, incorporated place, school district, and county had addresses selected in the long form sample.

To produce estimates from the long form data, weighting was performed at the weighting area level. In forming weighting areas, trade-offs between reliability, consistency of the estimates, and complexity of the implementation were considered. The decision was made to form weighting areas consisting of small geographic areas with at least 400 sample persons (or about 200 or more completed long forms) that do not cross county boundaries. No other boundary constraints were imposed. Thus, total population estimates from the long form data will agree with census counts reported in SF 1 and SF 2 for the weighting area, county, and other higher geographic areas obtained by combining either weighting areas or counties. Differences between long form estimates of characteristics in the SF 3 and their corresponding values in the SF 1 or SF 2 are particularly noticeable for small places, tracts, and block groups. Examples of these characteristics are the total number of people, the number of people reporting specific racial categories, and the number of housing units. The official values for items reported on the short form come from SF 1 and SF 2.

Because the weighting areas were formed at a smaller geographic level, any differential nonresponse to long form questionnaires by demographic groups or geographical areas included in a weighting area may introduce differences in complete counts (SF 1 and SF 2) and the SF 3 total population estimates. Also, an insufficient number of sample cases in the weighting matrix cells could lead to differences in SF 1, SF 2, and SF 3 population totals. Thus, differences between the census and SF 3 counts are typical and expected.

In 1990, separate tabulations were not prepared for small areas below a certain size. In contrast, Census 2000 tabulations are being prepared for all areas to maximize data availability. This approach may lead to a greater number of anomalous results than what may have been observed with tabulations released from the 1990 census. A similar phenomenon occurred in the 1990 census when weighting areas respected city and place boundaries. Census counts differed from the long form data estimates in small places. As expected, these differences were sometimes large.

The SF 1 tables provide the official census count of the number of people in an area. The SF 3 tables provide estimates of the proportion of people with specific characteristics, such as occupation, disability, or educational attainment. The total number of people in the SF 3 table is provided for use as the denominator, or base, for these proportions. Estimates in the SF 3 tables give the best estimates of the proportion of people with a particular characteristic, but the census count is the official count of how many people are in the area.

The SF 1 gives exact numbers even for very small groups and areas; whereas, SF 3 gives estimates for small groups and areas, such as tracts and small places, that are less exact. The goal of SF 3 is to identify large differences among areas or large changes over time. Estimates for small areas and small population groups often exhibit large changes from one census to the next, so having the capability to measure them is worthwhile.

March 2003

108th Congressional District Summary File (Sample)
Data Note 7

Median incomes for nonfamily households by race, Tables 156A through P156I, were calculated from a 38-category income distribution rather than the standard 39-category income distribution. The 38-category distribution collapsed the two highest categories ($175,000 - $199,999 and $200,000 and over) into a single category of $175,000 and over.

March 2003

108th Congressional District Summary File (Sample)
Data Note 8

Census 2000 Summary File 3 CD-ROM/DVD
Census 2000 Data Engine Software
Output | Create Output As Summary

The Census 2000 Summary File 3 database contains several tables of normalized data items, such as P53–Median Household Income in 1999, P82–Per Capita Income in 1999, and H18–Average Household Size of Occupied Housing Units by Tenure. In general, the **Census 2000 Data Engine** software's **Create Output As Summary** function recognizes normalized data items and presents them as weighted averages of the summarized geographic components using the 100 percent population or housing count as the weighting factor. However, the version of the Census 2000 Data Engine software used on the Summary File 3 State CD-ROMs fails to recognize **Per Capita** as a one of the normalization techniques and performs a standard summation. This applies only to tables P82 and P157A through P157I. The Per Capita Income value displayed on the **DP–3, Profile of Selected Economic Characteristics,** is derived from the formula (P083001/P001001) rather than (P082001) as originally specified so that **Create Output As Summary** will perform correctly.

March 2003

108th Congressional District Summary File (Sample)
Data Note 9

This user update is described on our Web site (www.census.gov) as:

Technical Note on Same-Sex Unmarried Partner Data From the 1990 and 2000 Censuses

The release of data in the SF 1 files from the 2000 census has brought with it a number of analyses documenting change that has occurred since the last census was conducted in 1990. While many of the variables and processes between the two censuses are comparable, some are not, and direct comparison of some estimates may lead to misleading conclusions. This note discusses one such topic, that of "unmarried partners," and advises that for some analyses — those involving unmarried same-sex partners — direct comparison of the 1990 and 2000 estimates is not substantively valid.

The household relationship item in both the 1990 and the 2000 censuses offered many ways of identifying how other people in the household were related to the householder (the person in whose name the house is owned or rented). Categories included spouse, child or other relative of the householder, housemate/roommate, roomer/boarder, and unmarried partner. In all circumstances, the respondent was asked to choose the category that best represented how other members of the household were related to the householder.

In both censuses, the "spouse" and "unmarried partner" response categories were defined and asked the same way. However, there were important differences in data processing that mean that some of the data are not comparable, limiting the usefulness of comparisons of the number of same-sex unmarried partners between these two censuses.

In both censuses, if a person was identified as the "spouse" of the householder and was the same sex as the householder, the "spouse" response was flagged for further review and allocation, that is, assignment of a value other than that originally reported, based on other data on the form. In 1990, the edit and allocation procedures did not allow same-sex "spouse" combinations to occur, thus resulting in the allocation of one of these two items in order to achieve editing consistency among the responses.

Processing steps were changed for Census 2000 for households that contained same-sex "spouses." If the person with the "spouse" category was the same sex as the householder and if neither person had their sex previously allocated, a relationship response of "spouse" was allocated as an "unmarried partner" response. Since marital status was no longer on the short form, its given value could not be considered (or modified) in this allocation procedure as it had been in 1990.

Data allocation is a standard statistical practice that is followed by most data collection agencies. Data on the relationship item (as other items) were subject to allocation in the census, as they are in virtually all Census Bureau surveys. In 1990, the marital status item was available on the 100 percent (short) form and aided in both the evaluation of the consistency of responses between the householder and the "spouse," and in the subsequent allocation procedure. The 1990 procedure allocated responses via a statistical model that distributed allocated responses from answers given by respondents in a proximate geographic area. This procedure used key demographic data from the census form, including marital status, as stratifying factors to provide a reasonable distribution of allocated responses. This procedure, while ensuring that no same-sex spouse response could be subsequently allocated, produced a set of allocated responses that could have included an "unmarried partner" response as well as any other response that was consistent with the age/sex/marital status profile of the respondent. This would include being allocated as a sibling or a relative, for example, or if the age differences were far enough apart (15 or more years), even a parent or child of the householder.

Three principal factors affected our decision to take this approach for Census 2000.

1. Same-sex spouse responses were flagged as invalid to comply with the 1996 Federal Defense of Marriage Act (H.R. 3396) passed by the 104th Congress. This act instructs all federal agencies only to recognize opposite-sex marriages for the purposes of enacting any agency programs. In order for Census Bureau data to be consistent with this act and the data requirements of other federal agencies, same-sex spouse responses were invalidated. The legislation defines marriage and spouse as follows:

   "In determining the meaning of any Act of Congress, or of any ruling, regulation or interpretation of the various administrative bureaus and agencies of the United States, the word 'marriage' means only a legal union between one man and one woman as husband and wife, and the word 'spouse' refers only to a person of the opposite sex who is a husband or wife."

   In order for the Census Bureau to be consistent with this act and the data requirements of other federal agencies, same-sex "spouse" responses were invalidated.

2. The second issue was statistical in nature. The principal basis of any good statistical allocation routine rests on the selection of the stratifying or input factors to provide a good statistical model. Without marital status data on the 100 percent form in Census 2000, the allocation routine would be relatively weak. Since many partners are roughly the same age, a statistical routine without marital status as one of its factors would have likely resulted in an overestimate of adult siblings or relatives, as the majority of people living in households are relatives, and this is the population from which we would draw our allocated responses. Additionally, if the same-sex partners were more than 15 years difference in age, the statistical routine would have likely allocated the invalidated "spouse" response as either a "child" or "parent" of the householder, as these types of relatives predominate in households in this age range of differences. This was an unacceptable outcome, as it would actually destroy the intent of the original "spouse" response, which clearly indicated a nonparental type of relationship. It should be noted that the "spouse" response on the form is assumed to be deliberate — not accidental — as it was the first response category on the question and was not placed between other possible response categories that may have been meant to be marked, such as housemates or roomers.

3. The third factor took into consideration that couples in long term same-sex relationships may consider themselves as "married partners" and thus respond as such on the census form. In addition, at the time of writing the editing program for Census 2000, there were several challenges in the courts concerning the legality of same-sex marriages. Clearly, we could not ignore the fact that same-sex spouse responses were going to be recorded during Census 2000. In light of these social and legal aspects — and the lack of a key variable in the statistical allocation routine (marital status) — the assignment of same-sex "married" couples to the same-sex "unmarried partner" category was the procedure chosen for the editing process. We were adverse to a randomized allocation of these responses after people had clearly marked a close relationship preference on the census form.

   As a result of these changes in the processing routine, estimates of same-sex unmarried partners are not comparable between the 1990 and 2000 census. We believe 2000 census estimates of this category are better estimates than those produced in 1990. It should also be noted that estimates of opposite-sex unmarried partners, however, were not affected by these editing procedures and changes and are comparable between the two censuses.

For further information on this topic, please contact the Fertility and Family Statistics Branch on 301-457-2416.

March 2003

108th Congressional District Summary File (Sample)

INDEX TO 108TH CONGRESSIONAL DISTRICT SUMMARY FILE (SAMPLE)
GEOGRAPHY NOTES

Note	Geographic area
1	Alaska
2	California
3	Connecticut
4	Florida
5	Georgia
6	Nebraska
7	Tennessee
8	Wisconsin
9	Selected States

108th Congressional District Summary File (Sample)
Geography Note 1

Alaska: 02

Nelson Lagoon Alaska Native village statistical area (ANVSA) (AIANHH 7025) erroneously contains block 2010, census tract 1 (000100) in Aleutians East census area (01598), Aleutians East Borough (013). This block should have not been coded to any ANVSA (9999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 02-003

March 2003

108th Congressional District Summary File (Sample)
Geography Note 2

California: 06

Los Angeles city (FIPS code 44000) erroneously contains block 1011, census tract 4002.03 (400203) in East San Gabriel Valley CCD (FIPS code 90810), Los Angeles County (FIPS code 037), CA (FIPS code 06). This block should have been coded to the place Balance of East San Gabriel Valley CCD (FIPS code 99999). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP side.

Internal Errata ID 06-001

March 2003

108th Congressional District Summary File (Sample)
Geography Note 3

Connecticut: 09

The place record, Balance of Milford town (FIPS code 99999) erroneously contains block 2999, census tract 1502 (150200) in Milford town (FIPS code 47535), New Haven County (FIPS code 009), CT (FIPS code 09). This block should have been coded to place Milford city (balance) (FIPS code 47515). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 09-001

March 2003

108th Congressional District Summary File (Sample)
Geography Note 4

Florida: 12

Yeehaw Junction CDP (FIPS code 78975) in St. Cloud CCD (FIPS code 93029), Osceola County (FIPS code 097), FL (FIPS code 12) should be named Buenaventura Lakes with FIPS code 09415. In 1990, this area was named Buena Ventura Lakes (FIPS code 09415). The area that should have been Yeehaw Junction CDP was erroneously not defined and does not appear in any Census 2000 products.

Internal Errata ID 12-001

March 2003

108th Congressional District Summary File (Sample)
Geography Note 5

Georgia: 13

The place record Balance of Athens CCD (FIPS code 99999) erroneously contains blocks 2021 and 2023, census tract 1305 (130500) in Athens CCD (FIPS code 90138), Clarke County (FIPS code 059). Both blocks should have been coded to Bogart town (FIPS code 09068).

The place record Balance of Winterville CCD (FIPS code 99999) erroneously contains blocks 1008 and 1009, census tract 1406 (140600) in Winterville CCD (93402), Clarke County (FIPS code 059). Both blocks should have been coded to the place Athens-Clarke County (balance) (FIPS code 03440). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

This note applies to American FactFinder (AFF), CD-ROM, and redistricting data downloaded from the FTP site.

Internal Errata ID 13-001

March 2003

108th Congressional District Summary File (Sample)
Geography Note 6

Nebraska: 31

In the PL 94-171 and Summary File (SF) data products, Cisco CDP (FIPS code 09112) in Lisco precinct (FIPS code 91790), Garden County (FIPS code 069), NE (FIPS code 31) should be named Lisco with FIPS code of 28315.

Internal Errata ID 31-002

March 2003

108th Congressional District Summary File (Sample)
Geography Note 7

Tennessee: 47

The place record Balance of Metropolitan Government CCD (FIPS code 99999) erroneously contains blocks 1001 and 1008, census tract 171 (017100) in Metropolitan Government CCD (FIPS code 92200), Davidson County (FIPS code 037), TN (FIPS code 47). Both blocks should have been coded to place Nashville-Davidson (balance) (FIPS code 52006). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 47-001

March 2003

108th Congressional District Summary File (Sample)
Geography Note 8

Wisconsin: 55

The county subdivision of Scott town (FIPS code 72200), in place Balance of Scott town (FIPS code 99999) erroneously contains blocks 2048, 2063, and 2064, census tract 203 (020300), Brown County (FIPS code 009), WI (FIPS code 55). These blocks should have been coded to county subdivision and place Pulaski village (FIPS code 65675).

The county subdivision of Pittsfield town (FIPS code 63075), in place Balance of Pittsfield town (FIPS code 99999) erroneously contains block 2049, census tract 203 (020300), Brown County (FIPS code 009). This block should have been coded to county subdivision and place Pulaski village (FIPS code 65675). This is incorrect in both the PL 94-171 data products and Summary File (SF) data products.

Internal Errata ID 55-001

March 2003

108th Congressional District Summary File (Sample)
Geography Note 9

The 108th Congressional Districts appearing in Census Bureau products reflect the information provided to the Census Bureau by the states. Several states did not provide congressional district codes for census blocks in water areas. This only occurred in large bodies of water, such as the Great Lakes, Chesapeake Bay, and Long Island Sound. These water areas have a blank congressional district code and the 108th Congressional District data associated with these areas may show records with a "part" indicator. The "part" designation indicates either of two situations. It can indicate that there is a second record (part) in the dataset. Or, in the case of these water areas with blank congressional district codes, the "part" indicates that the record does not represent the entirety of the geographic entity.

April 2003

108th Congressional District Summary File (Sample)
Technical Documentation Note 1

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

108th Congressional District Summary File (Sample)
Technical Documentation Note 2

The Language section of the Code List appendix had two spelling errors. They have been corrected to read as follows:

772	Tahitian
971	OTO-MANGUEAN

September 2003

Public Use Microdata Sample Files
Data Note 1

TECHNICAL NOTE ON SAME-SEX UNMARRIED PARTNER DATA FROM THE 1990 AND 2000 CENSUSES

The release of data from the 2000 census has brought with it a number of analyses documenting change that has occurred since the last census was conducted in 1990. While many of the variables and processes between the two censuses are comparable, some are not, and direct comparison of some estimates may lead to misleading conclusions. This note discusses one such topic, that of "unmarried partners," and advises that for some analyses—those involving unmarried same-sex partners—direct comparison of the 1990 and 2000 estimates is not substantively valid.

The household relationship item in both the 1990 and the 2000 censuses offered many ways of identifying how other people in the household were related to the householder (the person in whose name the house is owned or rented). Categories included spouse, child or other relative of the householder, housemate/roommate, roomer/boarder, and unmarried partner. In all circumstances, the respondent was asked to choose the category that best represented how other members of the household were related to the householder.

In both censuses, the "spouse" and "unmarried partner" response categories were defined and asked the same way. However, there were important differences in data processing that mean that some of the data are not comparable, limiting the usefulness of comparisons of the number of same-sex unmarried partners between these two censuses.

In both censuses, if a person was identified as the "spouse" of the householder and was the same sex as the householder, the "spouse" response was flagged for further review and allocation, that is, assignment of a value other than that originally reported, based on other data on the form. In 1990, the edit and allocation procedures did not allow same-sex "spouse" combinations to occur, thus resulting in the allocation of one of these two items in order to achieve editing consistency among the responses.

Processing steps were changed for Census 2000 for households that contained same-sex "spouses." If the person with the "spouse" category was the same sex as the householder and if neither person had their sex previously allocated, a relationship response of "spouse" was allocated as an "unmarried partner" response. Data allocation is a standard statistical practice that is followed by most data collection agencies. Data on the relationship item (as other items) were subject to allocation in the census, as they are in virtually all Census Bureau surveys.

Two principal factors affected our decision to take this approach for Census 2000.

1. Same-sex spouse responses were flagged as invalid to comply with the 1996 Federal Defense of Marriage Act (H.R. 3396) passed by the 104th Congress. This act instructs all federal agencies only to recognize opposite-sex marriages for the purposes of enacting any agency programs. In order for Census Bureau data to be consistent with this act and the data requirements of other federal agencies, same-sex spouse responses were invalidated. The legislation defines marriage and spouse as follows:

   "In determining the meaning of any Act of Congress, or of any ruling, regulation or interpretation of the various administrative bureaus and agencies of the United States, the word 'marriage' means only a legal union between one man and one woman as husband and wife, and the word 'spouse' refers only to a person of the opposite sex who is a husband or wife." In order for the Census Bureau to be consistent with this act and the data requirements of other federal agencies, same-sex "spouse" responses were invalidated.

2. The second factor took into consideration that couples in long term same-sex relationships may consider themselves as "married partners" and thus respond as such on the census form. In addition, at the time of writing the editing program for Census 2000, there were several challenges in the courts concerning the legality of same-sex marriages. Clearly, we could not ignore the fact that same-sex spouse responses were going to be recorded during Census 2000. In light of these social and legal aspects—and the lack of a key variable in the statistical allocation routine (marital status)—the assignment of same-sex "married" couples to the same-sex "unmarried partner" category was the procedure chosen for the editing process. We were adverse to a randomized allocation of these responses after people had clearly marked a close relationship preference on the census form.

   As a result of these changes in the processing routine, estimates of same-sex unmarried partners are not comparable between the 1990 and 2000 census. We believe 2000 census estimates of this category are better estimates than those produced in 1990. It should also be noted that estimates of opposite-sex unmarried partners, however, were not affected by these editing procedures and changes and are comparable between the two censuses.

For further information on this topic, please contact the Fertility and Family Statistics Branch on 301-763-2416.

March 2003

Public Use Microdata Sample Files
Data Note 2

The variable PNUM, Person Sequence Number, on the Housing Unit record of the 1-percent PUMS files should not be used. Users will have to create their own sort key. However, the number of person records is correct and matches the number in the variable PERSONS on the Housing Unit record.

April 2003

Public Use Microdata Sample Files
Data Note 3

The variable TAXAMT on the housing unit record erroneously includes data for a code of 68 for the 1-percent files. Data for code 68 should be collapsed into code 67.

May 2003

Public Use Microdata Sample Files
Data Note 4

The geographic equivalency file in the ftp directory for Montana for the 1-percent PUMS files inadvertently contained data for Kansas (http://www2.census.gov/census_2000/datasets/PUMS/OnePercent/Montana/PUMEQ1-MT.TXT). The file was replaced with the geographic equivalency file data for Montana.

June 2003

Public Use Microdata Sample Files
Data Note 5

Two weights are present in microdata records, the weight of the housing unit, labeled HWEIGHT, and the weight of each person, labeled PWEIGHT.

The basic rules for determining the correct weight to use for a particular table are listed below:

a. Tables with a universe of "Housing units," "occupied housing units," "vacant housing units," or "occupied housing units with a particular characteristic" use the housing unit weight.

b. Tables with a universe of "Population," "population of a particular type, such as race/Hispanic/American Indian tribe," "population XX years and over," and subsets of a population such as "institutionalized population" or "population for whom poverty status is determined" use the person weight.

c. Tables with a universe of "Households," "households with a householder who is....," "families," and "families with a householder who is ....." are usually tallied by the householder's characteristics and use the person weight of the person who is the householder.

d. Tables with a specific universe, such as "own children," "civilian noninstitutionalized population," and "population in subfamilies" describe subsets of the population and use the person weight of the individual people who meet the specific criteria.

e. Tables with a universe of subfamilies use the weight of the "reference person" for each subfamily. There can be a maximum of four different subfamilies in the same household, but each person can only be in one subfamily. The reference person is defined independently for each subfamily and varies by the type of subfamily: in a parent-child subfamily, the reference person is the parent; in a husband/wife subfamily, the reference person is the husband.

f. In some tables, one may want to tally the number of households where the characteristic of a person other than the householder qualifies a household for inclusion in the specific table. In this situation, the person weight of the person who is the householder is always used, even though the householder's characteristics would not have qualified the household for inclusion. For example, when tallying the number of households with a disabled person, one would tally the householder's weight based on the presence of any person (which may or may not be the householder) who is disabled.

September 2003

Public Use Microdata Sample Files
Data Note 6

1-Percent and 5-Percent PUMS

When compiling PUMS income data for households or for families, users should combine the income variable of interest with other variables on the file to obtain the correct counts.

Example 1: Total Household income distributions are derived using the variables:

	For the DVD	For the ASCII and SAS
Household income	h_hinc	hinc
Relationship—householder	p_relate	relate
Person weight	p_pweight	pweight

Example 2: Total Family income distributions are derived using the variables:

	For the DVD	For the ASCII and SAS
Family income	h_finc	finc
Relationship—householder	p_relate	relate
Household type—families	h_hht	hht
Person weight	p_pweight	pweight

November 2003

Public Use Microdata Sample Files
Data Note 7

Why the Poverty Recode on PUMS Does Not Match the Total Income Amounts

The Census 2000 Public Use Microdata Sample (PUMS) files contain an apparent inconsistency between the variable called POVERTY and the variables defining person or family total income, INCTOT and FINC respectively. In some instances the difference is more than can be attributed to rounding. Such differences are due to the order in which variables are topcoded. In these cases, the value in POVERTY is more accurate than the value calculated from the components.

Specifically, the Census Bureau rounds all reported income amounts using a graduated rounding scheme, where large values are rounded more than smaller values on the PUMS file and further puts an upper limit on the amount of income reported by type of income. For example, any individual with retirement income (INCRET) greater than $52,000 has his or her reported value replaced with the state mean of retirement income over the topcode amount. As a result, in the PUM 5-percent file in Alabama, 424 people have retirement income equal to $137,000. Some of those people originally reported retirement income greater than $137,000. Others originally reported retirement income less than $137,000, but greater than or equal to $52,000.

To see the full explanation of how poverty is calculated in general, see Appendix B of the Technical Documentation, Definitions of Subject Characteristics. The POVERTY recode is based on the ratio between the total income and the relevant poverty threshold for families or people. Using the unconstrained income reported, the Census Bureau creates the poverty recode in two basic steps.

(1) Calculate the ratio of income to poverty
    Ratio = 100* (Income/Threshold)
(2) Recode the ratio
    0 = Not in universe
    1 = Less than 1%
    2 = 1% to less than 2%
    3 = 2% to less than 3%
    etcetera...

For example, the poverty threshold for a person under 65 years old living with no relatives in 1999 was $8,667. If the person had only income from Supplemental Security Income (SSI) equal to $14,500, then the poverty ratio would equal 167.3 and the poverty recode (POVERTY) would equal 168. On the PUMS file, this person's income would be subject to the topcode. If the person lived in Alabama, their SSI would be set to $18,100. As a result, if a user calculated the ratio of income to poverty using the dollar figure for income on the PUMS, they would get a ratio equal to 208.8 and expect the poverty recode to equal 209. However, the result would not match the variable POVERTY, which is calculated before topcoding and remains 168 on the PUMS file. As this example shows, there may be cases on the public use file where the POVERTY recode is inconsistent with a poverty recode a user makes based on the PUMS income variables, which are rounded and topcoded.

December 2003

Public Use Microdata Sample Files
Data Note 8

Why the Percent of Income Spent on Owner/Renter Cost Recodes on PUMS May Not Match Independently Calculated Recodes

The Census 2000 Public Use Microdata Sample (PUMS) files contain an apparent inconsistency between the variables called GRAPI and SMOCAPI and the variables defining household total income (HINC), Gross Rent (GRENT), and Selected Monthly Owner Costs (SMOC). This situation is the result of inconsistencies in the topcoding of variables. Small differences are likely due to rounding. Larger differences are related to the inconsistency.

Specifically, the Census Bureau rounds all reported income amounts using a graduated rounding scheme where large values are rounded more than smaller values on the PUMS file and further puts an upper limit on the amount of income reported by type of income. For example, any individual with retirement income (INCRET) greater than $52,000 has his or her reported value replaced with the state mean of retirement income over the topcode amount. As a result, in the PUMS 5-percent file in Alabama, 424 people have retirement income equal to $137,000.
Some of those people originally reported retirement income greater than $137,000.
Others originally reported retirement income less than $137,000, but greater than or equal to $52,000. The income from all individuals in a household are summed to create the household income (HINC).

Gross Rent (GRENT) and Selected Monthly Owner Costs (SMOC) are also calculated using rounded and topcoded values for RENT, MRT1AMT, MRT2AMT, INSAMT, CONDFEE, MHCOST, ELEC, GAS, WATER, and OIL.

Gross Rent as a Percentage of Household Income (GRAPI) is calculated by dividing the GRENT by the monthly household income (HINC/12). When calculated independently using the GRENT and the HINC from the PUMS, the answer may not agree with the GRAPI presented on the PUMS. Selected Monthly Owner Costs as a Percentage of Household Income (SMOCAPI) is treated the same way as GRAPI.

Example: RENT=$160, ELEC=$1,100, GAS=$1,000, WATER=$180, OIL=$3,800, GRENT=$667, HINC=$14,450. Because OIL is topcoded (topcode=state mean of all answers over $2,100), GRENT/(HINC/12)=55 percent; but is reported as the pre-topcoded value of 44 percent on the PUMS file.

As this example shows, there may be cases on the public use file where housing recodes are inconsistent with recodes a user makes based on the PUMS variables, which are rounded and topcoded.

December 2003

Public Use Microdata Sample Files
Data Note 9

1-Percent and 5-Percent PUMS

A table showing the percent urban for each Super-PUMA and PUMA was released to the public in July 2004 as PHC-T-36. The title of the table is "Percent Urban for Super-PUMAs and PUMAs." It can be found at http://www.census.gov/population/www/cen2000/phc-t36.html.

July 2004

Public Use Microdata Sample Files
Geography Note 1

New Jersey: 34

Super-PUMA 34101 is not contiguous. A small portion of the super-PUMA, comprising the Saddle River borough, is detached from the main area of super-PUMA 34101.

April 2003

Public Use Microdata Sample Files
Technical Documentation Note 1

The definition for some variables in Chapter 6 (Data Dictionary) may not be easily found in Appendix B (Definitions of Subject Characteristics). Below is a list of the subjects that may be difficult to locate and the appropriate topic to refer to in Appendix B.

Housing Record Type Variable Description	**Appendix B Topic**
Size of building	See Units in Structure
Year building built	See Year Structure Built
Cost of electricity (annual)	See Utilities
Cost of gas (annual)	See Utilities
Cost of water and sewer (annual)	See Utilities
Cost of oil, kerosene, or wood (annual)	See Utilities
Number of people 65 years and over in household	See Household Type and Relationship
Number of people under 18 years in household	See Household Type and Relationship
Number of people in family	See Household Type and Relationship
Number of own children under 18 years in household	See Household Type and Relationship
Number of related children under 18 years in household	See Household Type and Relationship
Presence of subfamily in household	See Household Type and Relationship
Presence and age of own children under 18 years	See Household Type and Relationship
Presence and age of related children under 18 years	See Household Type and Relationship
Specified value indicator	See Value
Family type and employment status	See Household Type and Relationship and see Employment Type
Family type and work experience of householder	See Household Type and Relationship and see Work Status in 1999

Person Record Type Variable Description	**Appendix B Topic**
Able to go out disability	See Disability Status
Military service	See Veteran Status
Years of military service	See Veteran Status
Vehicle occupancy	See Journey to Work
Layoff from job	See Employment Status
Absent from work	See Employment Status
Return-to-work recall	See Employment Status
Looking for work	See Employment Status
Back to work	See Employment Status
Hours per week in 1999	See Work Status in 1999

March 2003

Public Use Microdata Sample Files
Technical Documentation Note 2

The unweighted counts for housing units for all states in Appendix I were incorrect. Appendix I (column 2) has been corrected to include the correct total housing unweighted counts.

May 2003

Public Use Microdata Sample Files
Technical Documentation Note 3

The value for the variable SAMPLE in the housing unit record of the Data Dictionary (Chapter 6) for the 1-percent files was incorrectly listed as 2. It has been corrected to show a value of 1.

May 2003

Public Use Microdata Sample Files
Technical Documentation Note 4

The value of 9 for the variable CLWRKR in the person record of the Data Dictionary (Chapter 6) for the 1-percent files was missing. It has been corrected as shown below.

CLWKR
Class of Worker

0. Not in universe (Under 16 years of LASTWRK = 2)
1. Employee of private for-profit company
2. Employee of private not-for-profit company
3. Employee of local government
4. Employee of state government
5. Employee of federal government
6. Self-employed in unincorporated business or company
7. Self-employed in incorporated business or company
8. Unpaid family worker
9. Unemployed, no work experience in the last 5 years

May 2003

Public Use Microdata Sample Files
Technical Documentation Note 5

The last line of the Housing Unit Record of the Data Dictionary (Chapter 6, page 6-41) for the 1-percent files was incorrectly listed as FILLER in character positions 267-314. It has been corrected to show that the last variable is FINC in character positions 259-266.

The last variable of the Person Record of the Data Dictionary (Chapter 6, page 6-71) for the 1-percent files was incorrectly listed as POVERTY in character positions 312-314. It has been corrected to show that the last line is FILLER in character positions 315-316.

June 2003

Public Use Microdata Sample Files
Technical Documentation Note 6

The last line of Table E. Census 2000 PUMS Standard Error Design Factors, ″Household income in 1999 by selected monthly owner costs as a percentage of income in 1999″ was inadvertently omitted on page 4-29 (Delaware) and page 4-55 (North Carolina). These pages have been corrected.

July 2003

Public Use Microdata Sample Files
Technical Documentation Note 7

Two codes for the variable REGION in Chapter 6 (Data Dictionary) of the technical documentation were not correct. "Region not identified" and "Puerto Rico" have been corrected as shown below.

REGION
Region Code

0. Region not identified
1. Northeast
2. Midwest
3. South
4. West
9. Puerto Rico

August 2003

Public Use Microdata Sample Files
Technical Documentation Note 8

In the Race section of the Code List appendix, the tribes with codes F49-F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below.

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

Public Use Microdata Sample Files
Technical Documentation Note 9

The Language section of the Code List appendix had two spelling errors. They have been corrected to read as follows:

772	Tahitian
971	OTO-MANGUEAN

September 2003

Public Use Microdata Sample Files
Technical Documentation Note 10

This technical documentation has been updated to include components for the 5-percent PUMS.

October 2003

Public Use Microdata Sample Files
Technical Documentation Note 11

An explanatory note was inadvertently omitted from Appendix G. Code List "One-to-One Correspondence of Census 2000 Codes and Standard Occupational Classification (SOC) Equivalent Codes (Sorted by Census 2000 Codes)." Page G-138 is corrected to include the note.

November 2003

Public Use Microdata Sample Files
Technical Documentation Note 12

The variable FNF was incorrectly defined in the Data Dictionary (Chapter 6 and Chapter 7) of the technical documentation. It is now corrected as shown below.

FNF

Farm/Nonfarm Recode

0 . Not in universe (GQ)

1 . Farm

2 . Nonfarm (includes urban)

May 2004

Public Use Microdata Sample
Technical Documentation Note 13

A number of languages in the Language Code List for the 5-Percent PUMS in Appendix G. Code Lists of the technical documentation were incorrectly listed. They are corrected as shown below.

Language Code List for the 5-Percent PUMS Files

Codes	Language
643	Kashubian (See code 983)
661	Ossete (See code 985)
684	Chuvash (See code 986)
693	Yakut (See code 986)
696	Caucasian (See code 994)
699	Brahui (See code 986)
707	Burushaski (See code 986)
727	Paleo-Siberian Languages, n.e.c. (See code 986)
729	Muong (See code 986)
733	Achinese (See code 988)
740	Minangkabau (See code 988)
749	Gorontalo (See code 988)
831	Nootka (See code 993)
833	Lower Chehalis (See code 993)
838	Cowlitz (See code 993)
840	Nootsack (See code 993)
855	Tanana (See code 993)
856	Tanacross (See code 993)
858	Tutchone (See code 993)
878	Santiam (See code 993)
886	Atsugewi (See code 993)
889	Shastan (See code 993)
891	Up River Yuman (See code 993)
897	Upland Yuman (See code 993)
918	Hichita (See code 993)
923	Wappo (See code 993)
944	Kawaiisu (See code 991)
947	Panamint (See code 991)
994	Other languages (Includes 680-681, 683, 696-697, 998-999)

May 2004

U.S. Virgin Islands Summary File
Data Note 1

In the U.S. Virgin Islands Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the U.S. Virgin Islands Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

April 2003

U.S. Virgin Islands Summary File
Data Note 2

In the U.S. Virgin Islands Demographic Profile, under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the U.S. Virgin Islands Summary File.

April 2003

U.S. Virgin Islands Summary File
Data Note 3

In U.S. Virgin Islands Summary File table HBG30, the total for aggregate number of vehicles available is too high. Housing units with zero vehicles were tallied as having one vehicle, units with 1 vehicle were tallied as having 2 vehicles, etc. The total can be correctly calculated by following the procedures in the Derived Measures section of Appendix B.

April 2003

U.S. Virgin Islands Summary File
Data Note 4

On the Census 2000 U.S. Virgin Islands questionnaire, respondents could report more than one type of disability. Several tables in the USVI summary file have as their universe the total disabilities tallied. Each line of the table represents the number of occurrences of a particular disability, and the numbers should be interpreted with care. For example, in table PCT32, Total Disabilities Tallied by Age by Type of Disability for the Civilian Noninstitutionalized Population 5 Years and Over With Disabilities, the second line of data, "Total disabilities tallied for people 5 to 15 years," does not refer to the number of people 5 to 15 years of age, or to the number of people 5 to 15 years of age who have a disability. Rather, it is the sum of the number of all disabilities reported among the population age 5 to 15 years. Lines in the table referencing specific disabilities are more easily interpreted. For example, the third line of data, "Sensory disability," refers to the number of sensory disabilities reported among people age 5 to 15 years (or the number of people 5 to 15 years of age who have a sensory disability).

Data users wanting to know the percent of civilian noninstitutionalized people 5 to 15 years of age with, for example, a sensory disability should divide line 3 from Table PCT32 with the sum of lines 3 and 27 from table PCT33, Sex by Age by Disability Status by Employment Status for the Civilian Noninstitutionalized Population 5 Years and Over.

April 2003

U.S. Virgin Islands Summary File
Data Note 5

In the U.S. Virgin Islands Summary File, data for matrices PCT12, PCT13, and PCT76 concerning grandparents who are living with grandchildren are restricted to persons in households similar to other summary file products. These numbers may be slightly different from the profile tables, which are for the total population.

April 2003

U.S. Virgin Islands Summary File
Data Note 6

In the U.S. Virgin Islands Summary File, the upper and lower quartile values were calculated using frequency counts rounded to a multiple of 0.5 rather than 0.25. As a result, the published quartile values are inaccurate.

April 2003

U.S. Virgin Islands Summary File
Data Note 7

The U.S. Virgin Islands Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the U.S. Virgin Islands Summary File. The data in the profiles are for the total population while the summary file data are restricted to persons in households similar to other summary file products.

April 2003

U.S. Virgin Islands Summary File
Data Note 8

In the U.S. Virgin Islands Summary File, Table PCT71 data for "Nonfamily Householders," nonfamily householders "Not living alone," and "Other unrelated individuals" have been corrected. Some respondents who were tallied as nonfamily householders "Not living alone" are now correctly tallied as "Other unrelated individuals." In American FactFinder and on the FTP site, these data have been replaced.

April 2003

U.S. Virgin Islands Summary File
Data Note 9

In detailed table **HBG63. MORTGAGE STATUS BY SELECTED MONTHLY OWNER COSTS FOR SPECIFIED OWNER-OCCUPIED HOUSING UNITS**, the selected monthly owner costs values shown for "Housing units **without** a mortgage," actually correspond to the selected monthly owner costs distribution shown for "Housing units **with** a mortgage."

- The data shown for "Less than $100" is actually data for "Less than $500"
- The data shown for "$100 to $199" is actually data for "$500 to $999"
- The data shown for "$200 to $299" is actually data for "$1,000 to $1,499"
- The data shown for "$300 to $399" is actually data for "$1,500 to $1,999"
- The data shown for "$400 to $499" is actually data for "$2,000 to $2,499"
- The data shown for "$500 or more" is actually data for "$2,500 or more"

September 2004

U.S. Virgin Islands Summary File
Geography Note 1

The U.S. Virgin Islands Summary File does not contain data records for every geographic area. Geographic header records were produced for all geographic areas, but areas with zero population do not contain accompanying data records.

April 2003

U.S. Virgin Islands Summary File Technical Documentation Note 1

Value, Price Asked was erroneously omitted from the list of aggregates subject to rounding on page B–67. The technical documentation has been corrected.

October 2002

U.S. Virgin Islands Summary File
Technical Documentation Note 2

In the U.S. Virgin Islands Summary File, codes 781, 782, and 783 were moved to the end of the "U.S. Island Areas" section of the Census 2000 State and Foreign Country Code List. Codes 781, 782, and 783 are the codes for St. Croix, St. John, and St. Thomas and were previously included in the "At Sea, Abroad, Not Specified" section of the code list. The codes associated with the "U.S. Island Areas" section of the list are now 060-099, 781-783. The codes associated with the "At Sea, Abroad, Not Specified" section of the list are now 554-780, 784-999

April 2003

.

U.S. Virgin Islands Summary File Technical Documentation Note 3

In the U.S. Virgin Islands Summary File technical documentation Appendix G, the “Management of companies and enterprises:” line of the Industry code list was indented to be aligned with the “Administrative and support and waste management services:” line.

April 2003

U.S. Virgin Islands Summary File
Technical Documentation Note 4

In the U.S. Virgin Islands Summary File, the parenthetical term "(DOLLARS)" was deleted from the title of table HBG52. The title now reads, "MEDIAN GROSS RENT AS A PERCENTAGE OF HOUSEHOLD INCOME IN 1999 [1]." Previously, "(DOLLARS)" followed "1999."

April 2003

U.S. Virgin Islands Summary File Technical Documentation Note 5

In the U.S. Virgin Islands Summary File, the language code list in Appendix G includes the full code list used in the United States census. Only codes 601 to 799 were used in the Virgin Islands. The remaining codes, 800-999, are primarily for American Indian and Alaska Native languages and were not used in the Virgin Islands.

April 2003

U.S. Virgin Islands Summary File Technical Documentation Note 6

In the U.S. Virgin Islands Summary File, “excluding” was changed to “except” in table PBG38, line “Other services (excluding public administration).”

April 2003

U.S. Virgin Islands Summary File Technical Documentation Note 7

In the U.S. Virgin Islands Summary File, the lines entitled In married-couple families of Table PCT67, Poverty Status in 1999 by Age by Household Type, were changed to In married-couple family households; the lines entitled In other families were changed to In other family households; and the lines entitled Unrelated individuals were changed to In nonfamily households and group quarters.

April 2003

U.S. Virgin Islands Summary File
Technical Documentation Note 8

In the Race section of the Code List appendix, the tribes with codes F49–F52 were incorrectly listed under the tribal grouping "Monacan." These tribes should have appeared under the tribal grouping "Mono" as shown below:

Monacan

F48	Monacan Indian Nation

Mono

F49	Mono
F50	North Fork Rancheria
F51	Cold Springs Rancheria
F52	Big Sandy Rancheria

September 2003

U.S. Virgin Islands Summary File
Technical Documentation Note 9

The Language section of the Code List appendix had two spelling errors. They have been corrected to read as follows:

772	Tahitian
971	OTO-MANGUEAN

September 2003

Guam Summary File
Data Note 1

In the Guam Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the Guam Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

April 2003

Guam Summary File
Data Note 2

The Guam Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the Guam Summary File. The data in the profiles are for the total population while the summary file data are restricted to persons in households similar to other summary file products.

April 2003

Guam Summary File
Data Note 3

The Guam Demographic Profile data for median value differ from the Guam Summary File data for median value. The summary file data were rounded to the nearest $100. The demographic profile data were not rounded. The summary file data are not incorrect, these data simply reflect the effects of rounding.

April 2003

Guam Summary File
Data Note 4

In the Guam Demographic Profile under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the Guam Summary File.

April 2003

Guam Summary File
Data Note 5

On the Census 2000 Guam questionnaire, respondents could report more than one type of disability. Several tables in the Guam summary file have as their universe the total disabilities tallied. Each line of the table represents the number of occurrences of a particular disability, and the numbers should be interpreted with care. For example, in table PCT336, Total Disabilities Tallied by Age by Type of Disability for the Civilian Noninstitutionalized Population 5 Years and Over With Disabilities. In this table, the second line of data, "Total disabilities tallied for people 5 to 15 years," does not refer to the number of people 5 to 15 years of age, or to the number of people 5 to 15 years of age who have a disability. Rather, it is the sum of the number of all disabilities reported among the population age 5 to 15 years. Lines in the table referencing specific disabilities are more easily interpreted. For example, the third line of data, "Sensory disability," refers to the number of sensory disabilities reported among people age 5 to 15 years (or the number of people 5 to 15 years of age who have a sensory disability).

Data users wanting to know the percent of civilian noninstitutionalized people 5 to 15 years of age with, for example, a sensory disability should divide line 3 from Table PCT36 with the sum of lines 3 and 27 from table PCT37, Sex by Age by Disability Status by Employment Status for the Civilian Noninstitutionalized Population 5 Years and Over.

April 2003

Guam Summary File
Data Note 6

In the Guam Summary File, data for matrices PCT18, PCT19, and PCT80 concerning grandparents who are living with grandchildren are restricted to persons in households similar to other summary file products. These numbers may be slightly different from the profile tables which are for the total population.

April 2003

Guam Summary File
Data Note 7

The values shown in the Guam Summary File for SEWAGE DISPOSAL may differ slightly from those found in the Guam Demographic Profile. The summary file allows a "Sewer/Septic" response even if the source of water for a housing unit is located outside the building and provides only cold water. In the demographic profile, housing units with a source of water that was located outside the building and provided only cold water were assigned to the "Other means" category of SEWAGE DISPOSAL.

April 2003

Guam Summary File
Data Note 8

In the Guam Summary File, table PCT75 data for “Nonfamily householders,” nonfamily householders “Not living alone,” and “Other unrelated individuals” have been corrected. Some respondents who were tallied as nonfamily householders “Not living alone” are now correctly tallied as “Other unrelated individuals.” In American FactFinder and on the FTP site, these data have been replaced.

April 2003

Guam Summary File
Technical Documentation Note 1

In the Guam Summary File, the universe for table PBG15, Main Reason for Moving to Guam, changed from “Persons born outside Guam” to “Population born outside Guam.”

April 2003

Guam Summary File
Technical Documentation Note 2

In the Guam Summary File, “excluding” was changed to “except” in table PBG42, line “Other services (excluding public administration).”

April 2003

Guam Summary File
Technical Documentation Note 3

In the Guam summary file, the original release did not include the correct data for the last three lines of table PCT32 for summary levels 160 (State-Place) and 155 (State-Place-County). The data shown in the last three lines of table PCT32 should have appeared in the first three lines of table PCT33. As a result, all of the data in tables PCT33 thru PCT37 were off by three cells. The file was corrected and reissued in American FactFinder and on the Census FTP site on January 24, 2003.

April 2003

Guam Summary File
Technical Documentation Note 4

In the Guam Summary File, the reference to PCT45 in the Data Dictionary (Chapter 7 in the technical documentation) was corrected. The file segment for cells 1 through 171 of PCT45 was changed from segment 16 to segment 15. The file segment for cells 172 through 341 of PCT45 remains segment 16.

April 2003

Guam Summary File
Technical Documentation Note 5

In the Guam Summary File, the lines entitled “In married-couple families” of Table PCT71, Poverty Status in 1999 by Age by Household Type, were changed to “In married-couple family households;” the lines entitled “In other families” were changed to “In other family households;” and the lines entitled “Unrelated individuals” were changed to “In nonfamily households and group quarters.”

April 2003

Guam Summary File
Technical Documentation Note 6

The Language section of the Code List appendix had a spelling error. It has been corrected to read as follows:

772	Tahitian

September 2003

Public Use Microdata Sample, Guam
Technical Documentation Note 1

The code list for Industry (Collapsed List) in Appendix G. Code Lists of the technical documentation did not include a legend which defined the alphabetic characters used in the codes. The legend shown below was added to the technical documentation.

Legend:

M = Multiple NAICS codes

P = Part of a NAICS code - NAICS code split between two or more Census codes

S = Not specified Industry in NAICS sector - Specific to Census codes only

Z = Exception to NAICS code - Part of NAICS industry has own Census code

May 2004

American Samoa Summary File
Data Note 1

In the American Samoa Demographic Profile under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the American Samoa Summary File.

April 2003

American Samoa Summary File
Data Note 2

In the American Samoa Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the American Samoa Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

April 2003

American Samoa Summary File
Data Note 3

For American Samoa, the numbers for the “Asian Indian” category under Ethnic Origin and Race in the Demographic Profile incorrectly include data for Pakistani. These numbers will differ in the American Samoa Summary File and subsequent Census 2000 data products because data for Pakistani will not be included in the Asian Indian category.

April 2003

American Samoa Summary File
Data Note 4

The American Samoa Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the American Samoa Summary File. The data in the profiles are for the total population while the summary file data are restricted to persons in households similar to other summary file products.

April 2003

American Samoa Summary File
Data Note 5

On the Census 2000 American Samoa questionnaire, respondents could report more than one type of disability. Several tables in the American Samoa summary file have as their universe the total disabilities tallied. Each line of the table represents the number of occurrences of a particular disability, and the numbers should be interpreted with care. For example, in table PCT336, Total Disabilities Tallied by Age by Type of Disability for the Civilian Noninstitutionalized Population 5 Years and Over With Disabilities. In this table, the second line of data, “Total disabilities tallied for people 5 to 15 years,” does not refer to the number of people 5 to 15 years of age, or to the number of people 5 to 15 years of age who have a disability. Rather, it is the sum of the number of all disabilities reported among the population age 5 to 15 years. Lines in the table referencing specific disabilities are more easily interpreted. For example, the third line of data, "Sensory disability," refers to the number of sensory disabilities reported among people age 5 to 15 years (or the number of people 5 to 15 years of age who have a sensory disability).

Data users wanting to know the percent of civilian noninstitutionalized people 5 to 15 years of age with, for example, a sensory disability should divide line 3 from Table PCT36 with the sum of lines 3 and 27 from table PCT37, Sex by Age by Disability Status by Employment Status for the Civilian Noninstitutionalized Population 5 Years and Over.

May 2004

American Samoa Summary File
Data Note 6

The values shown in the American Samoa Summary File for SEWAGE DISPOSAL may differ slightly from those found in the American Samoa Demographic Profile. The summary file allows a “Sewer/ Septic" response even if the source of water for a housing unit is located outside the building and provides only cold water. In the demographic profile, housing units with a source of water that was located outside the building and provided only cold water were assigned to the “Other means" category of SEWAGE DISPOSAL.

May 2004

American Samoa Summary File
Data Note 7

In the American Samoa Summary File, data for matrices PCT18, PCT19, and PCT80 concerning grandparents who are living with grandchildren are restricted to persons in households similar to other summary file products. These numbers may be slightly different from the profile tables which are for the total population.

May 2004

Commonwealth of the Northern Mariana Islands Summary File
Data Note 1

On the Census 2000 Commonwealth of the Northern Mariana Islands questionnaire, respondents could report more than one type of disability. Several tables in the Commonwealth of the Northern Mariana Islands summary file have as their universe the total disabilities tallied. Each line of the table represents the number of occurrences of a particular disability, and the numbers should be interpreted with care. For example, in table PCT36, Total Disabilities Tallied by Age by Type of Disability for the Civilian Noninstitutionalized Population 5 Years and Over With Disabilities. In this table, the second line of data, "Total disabilities tallied for people 5 to 15 years," does not refer to the number of people 5 to 15 years of age, or to the number of people 5 to 15 years of age who have a disability. Rather, it is the sum of the number of all disabilities reported among the population age 5 to 15 years. Lines in the table referencing specific disabilities are more easily interpreted. For example, the third line of data, "Sensory disability," refers to the number of sensory disabilities reported among people age 5 to 15 years (or, the number of people 5 to 15 years of age who have a sensory disability).

Data users wanting to know the percent of civilian noninstitutionalized people 5 to 15 years of age with, for example, a sensory disability should divide line 3 from Table PCT36 with the sum of lines 3 and 27 from Table PCT37, Sex by Age by Disability Status by Employment Status for the Civilian Noninstitutionalized Population 5 Years and Over.

April 2003

Commonwealth of the Northern Mariana Islands Summary File Data Note 2

In the Commonwealth of the Northern Mariana Islands Summary File data for matrices PCT18, PCT19, and PCT80 (concerning grandparents who are living with grandchildren) are restricted to persons in households similar to other summary file products. These numbers may be slightly different from the profile tables, which are for the total population.

April 2003

Commonwealth of the Northern Mariana Islands Summary File Data Note 3

The values shown in the Commonwealth of the Northern Mariana Islands Summary File for SEWAGE DISPOSAL may differ slightly from those found in the Commonwealth of the Northern Mariana Islands Demographic Profile. The summary file allows a "Sewer/Septic" response even if the source of water for a housing unit is located outside the building and provides only cold water. In the demographic profile, housing units with a source of water that was located outside the building and provided only cold water were assigned to the "Other means" category of SEWAGE DISPOSAL.

April 2003

Commonwealth of the Northern Mariana Islands Summary File
Data Note 4

In the Commonwealth of the Northern Mariana Islands Demographic Profile under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the Commonwealth of the Northern Mariana Islands Summary File.

April 2003

Commonwealth of the Northern Mariana Islands Summary File
Data Note 5

In the Commonwealth of the Northern Mariana Islands Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the Commonwealth of the Northern Mariana Islands Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

April 2003

Commonwealth of the Northern Mariana Islands Summary File Data Note 6

The Commonwealth of the Northern Mariana Islands Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the Commonwealth of the Northern Mariana Islands Summary File. The data in the profiles are for the total population while the summary file data are restricted to persons in households similar to other summary file products.

April 2003

Demographic Profile Data Note 1

Users may find slight differences in aggregate earnings for households between the Demographic Profile and Summary File 3 and related products. These differences are due to the treatment of offsetting positive and negative amounts for household members. Whenever offsetting values occurred, the Demographic Profile assigned these households a value zero while Summary File 3 and related products assigned a value of one dollar. The assignment of one dollar allows users to distinguish those households that had earnings from those households that did not have earnings. This will have little effect, if any, on mean household earnings.

April 2002

Demographic Profile Data Note 2

Users may find slight differences in the Occupants Per Room calculations between the Demographic Profile and Summary File 3, Summary File 4, and related products. “Occupants per room” is obtained by dividing the number of people in each occupied housing unit by the number of rooms in the unit. The Summary File 3 products correctly used a topcode value of “10 rooms” for those occupied housing units with “9 or more rooms.” In the Demographic Profiles, an incorrect topcode value of “9 rooms” was used.

June 2002

Demographic Profile
Data Note 3

The Census Bureau is aware there may be a problem or problems in the employment-status data of Census 2000 Summary File 3 (including tables P38, P43-46, P149A-I, P150A-I, PCT35, PCT69A-I, and PCT70A-I). The labor force data for some places where colleges are located appear to overstate the number in the labor force, the number unemployed, and the percent unemployed, probably because of reporting or processing error. The exact cause is unknown, but the Census Bureau will continue to research the problem.

July 2002

Demographic Profile Data Note 4

COMPARING SF 3 ESTIMATES WITH CORRESPONDING VALUES IN SF 1 AND SF 2

As in earlier censuses, the responses from the sample of households reporting on long forms must be weighted to reflect the entire population. Specifically, each responding household represents, on average, six or seven other households who reported using short forms.

One consequence of the weighting procedures is that each estimate based on the long form responses has an associated confidence interval. These confidence intervals are wider (as a percentage of the estimate) for geographic areas with smaller populations and for characteristics that occur less frequently in the area being examined (such as the proportion of people in poverty in a middle-income neighborhood).

In order to release as much useful information as possible, statisticians must balance a number of factors. In particular, for Census 2000, the Bureau of the Census created weighting areas—geographic areas from which about two hundred or more long forms were completed—which are large enough to produce good quality estimates. If smaller weighting areas had been used, the confidence intervals around the estimates would have been significantly wider, rendering many estimates less useful due to their lower reliability.

The disadvantage of using weighting areas this large is that, for smaller geographic areas within them, the estimates of characteristics that are also reported on the short form will not match the counts reported in SF 1 or SF 2. Examples of these characteristics are the total number of people, the number of people reporting specific racial categories, and the number of housing units. The official values for items reported on the short form come from SF 1 and SF 2.

The differences between the long form estimates in SF 3 and values in SF 1 or SF 2 are particularly noticeable for the smallest places, tracts, and block groups. The long form estimates of total population and total housing units in SF 3 will, however, match the SF 1 and SF 2 counts for larger geographic areas such as counties and states, and will be essentially the same for medium and large cities.

This phenomenon also occurred for the 1990 Census, although in that case, the weighting areas included relatively small places. As a result, the long form estimates matched the short form counts for those places, but the confidence intervals around the estimates of characteristics collected only on the long form were often significantly wider (as a percentage of the estimate).

SF 1 gives exact numbers even for very small groups and areas; whereas, SF 3 gives estimates for small groups and areas such as tracts and small places that are less exact. The goal of SF 3 is to identify large differences among areas or large changes over time. Estimates for small areas and small population groups often do exhibit large changes from one census to the next, so having the capability to measure them is worthwhile.

August 2002

Demographic Profile
Data Note 5

The categories are labeled incorrectly in DP4 for Selected Monthly Owner Costs as a Percentage of Household Income in 1999 and Gross Rent as a Percentage of Household Income in 1999. The ratio was computed separately for each unit and rounded to the nearest whole percentage; the ratio was not rounded to one decimal place as shown in the product. The correct distributions are as follows:

Selected Monthly Owner Costs as a Percentage of Household Income in 1999

Less than 15 percent
15 to 19 percent
20 to 24 percent
25 to 29 percent
30 to 34 percent
35 percent or more
Not computed

Gross Rent as a Percentage of Household Income in 1999

Less than 15 percent
15 to 19 percent
20 to 24 percent
25 to 29 percent
30 to 34 percent
35 percent or more
Not computed

August 2002

Demographic Profile
Technical Documentation Note 1

CORRECTIONS/ADDITIONS TO THE "ABOUT THE PROFILE" SECTION OF THE TECHNICAL DOCUMENTATION WERE MADE FOR THE FOLLOWING SUBJECT DEFINITIONS:

New definitions

All parents in family in labor force. The "parents in family" referred to in this category, which is shown under "EMPLOYMENT STATUS," are parents whose usual residence was the same as that of their own children; such parents are called "resident parents." If a child had only one such parent, then "all parents in family" means "one parent"; if the child had two such parents, then "all parents in family" means "two parents." The category describes an attribute of each own child under 6 and specifies whether the total number of the child's resident parents equals the number of such parents who were in the labor force.

Employment status, "Own children under 6 years" category. The universe for this category is own children under 6 years old (see definition of "own child"). The tabulation describes the distribution of own children under 6 years by whether their resident parents were in the labor force. (For more information, see "All parents in family in labor force.")

Revised definitions

Child. A child includes a son or daughter by birth, a stepchild, or an adopted child of the householder, regardless of the child's age or marital status. For more information, see "Own Child."

Conditional rounding. The means shown in the sample tables of the Demographic Profile may differ slightly from means appearing in or calculated from data in Summary File 3. In the Demographic Profile, conditional rounding is used when there is an estimate based on a weighted sample population of less than 30; and no rounding is used when the estimate is based on a weighted sample population of 30 or more. In Summary File 3, rounding is used for aggregates (numerators for calculating means) of selected variables. See Appendix B of the Summary File 3 technical documentation for details on the calculation of aggregates.

Own child. A never-married child under 18 years old who is a son or daughter of the householder by birth, marriage (a stepchild), or adoption. For 100-percent tabulations, own children consists of all sons/daughters of householders who are under 18 years of age. For sample data, own children consists of sons/daughters of householders who are under 18 years of age and who have never been married. Therefore, numbers of own children of householders may be different in these two tabulations since marital status was not collected as a 100-percent item in Census 2000. (Note: In the tabulation under "EMPLOYMENT STATUS" of own children under 6 years by employment status of parents, the number of "own children" includes any child under 6 years old in a family or a subfamily who is a son or daughter, by birth, marriage, or adoption, of a member of the householder's family, but not necessarily of the householder.)

August 2002

Demographic Profile
Technical Documentation Note 2

On page 3–16, the labels for the categories for Selected Monthly Owner Costs as a Percentage of Household Income in 1999 and Gross Rent as a Percentage of Household Income in 1999 were corrected. The ratio was computed separately for each unit and rounded to the nearest whole percentage; the ratio was not rounded to one decimal place as previously shown.

August 2002

Population and Housing Profile
Data Note 1

AMERICAN SAMOA

Rose Island is not shown because the population is zero.

February 2002

Population and Housing Profile
Data Note 2

AMERICAN SAMOA

By definition, all people living in group quarters are classified as "did no subsistence activity." Therefore, these people are excluded from the "Subsistence activity" lines shown in the Employment Status section.

February 2002

Population and Housing Profile
Data Note 3

AMERICAN SAMOA

Supplemental Security Income (SSI). Supplemental Security Income (SSI) is a U.S. federal assistance program administered by the Social Security Administration that guarantees a minimum level of income for needy aged, blind, or disabled individuals. The census questionnaire for American Samoa asked about the receipt of SSI; however, SSI is not a federally administered program in American Samoa. Therefore, it is not the same concept as SSI in the United States. The only way a resident of American Samoa could have appropriately reported SSI would have been if they lived in the United States at any time during calendar year 1999 and received SSI.

April 2002

Population and Housing Profile
Data Note 4

AMERICAN SAMOA

In the American Samoa Demographic Profile, "Median earnings (dollars)" for male and female full-time, year-round workers is mislabeled. It should read, "Median total money income (dollars)" for male and female full-time, year-round workers.

April 2003

Population and Housing Profile
Data Note 5

AMERICAN SAMOA

In the American Samoa Demographic Profile under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the American Samoa Summary File.

April 2003

Population and Housing Profile
Data Note 6

AMERICAN SAMOA

In the American Samoa Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the American Samoa Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

April 2003

Population and Housing Profile
Data Note 7

AMERICAN SAMOA

For American Samoa, the numbers for the "Asian Indian" category under Ethnic Origin and Race in the Demographic Profile incorrectly include data for Pakistani. These numbers will differ in the American Samoa Summary File and subsequent Census 2000 data products because data for Pakistani will not be included in the Asian Indian category.

April 2003

Population and Housing Profile
Data Note 8

AMERICAN SAMOA

The American Samoa Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the American Samoa Summary File. The data in the profiles are for the total population while the summary file data are restricted to persons in households similar to other summary file products.

April 2003

Population and Housing Profile
Data Note 9

AMERICAN SAMOA

The values shown in the American Samoa Summary File for SEWAGE DISPOSAL may differ slightly from those found in the American Samoa Demographic Profile. The summary file allows a "Sewer/ Septic" response even if the source of water for a housing unit is located outside the building and provides only cold water. In the demographic profile, housing units with a source of water that was located outside the building and provided only cold water were assigned to the “Other means" category of SEWAGE DISPOSAL.

May 2004

Population and Housing Profile
Data Note 1

COMMONWEALTH OF THE NORTHERN MARIANA ISLANDS

By definition, all people living in group quarters are classified as "did no subsistence activity." Therefore, these people are excluded from the "Subsistence activity" lines shown in the Employment Status section.

February 2002

Population and Housing Profile
Data Note 2

COMMONWEALTH OF THE NORTHERN MARIANA ISLANDS

Supplemental Security Income (SSI). Supplemental Security Income (SSI) is a U.S. federal assistance program administered by the Social Security Administration that guarantees a minimum level of income for needy aged, blind, or disabled individuals. The census questionnaire for the Commonwealth of the Northern Mariana Islands (CNMI) asked about the receipt of SSI; however, SSI is not a federally administered program in CNMI. Therefore, it is not the same concept as SSI in the United States. The only way a resident of CNMI could have appropriately reported SSI would have been if they lived in the United States at any time during calendar year 1999 and received SSI.

April 2002

Population and Housing Profile
Data Note 3

COMMONWEALTH OF THE NORTHERN MARIANA ISLANDS

In the Commonwealth of the Northern Mariana Islands Demographic Profile, "Median earnings (dollars)" for male and female full-time, year-round workers is mislabeled. It should read, "Median total money income (dollars)" for male and female full-time, year-round workers.

April 2003

Population and Housing Profile
Data Note 4

COMMONWEALTH OF THE NORTHERN MARIANA ISLANDS

In the Commonwealth of the Northern Mariana Islands Demographic Profile under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the Commonwealth of the Northern Mariana Islands Summary File.

April 2003

Population and Housing Profile
Data Note 5

COMMONWEALTH OF THE NORTHERN MARIANA ISLANDS

The values shown in the Commonwealth of the Northern Mariana Islands Summary File for SEWAGE DISPOSAL may differ slightly from those found in the Commonwealth of the Northern Mariana Islands Demographic Profile. The summary file allows a "Sewer/Septic" response even if the source of water for a housing unit is located outside the building and provides only cold water. In the demographic profile, housing units with a source of water that was located outside the building and provided only cold water were assigned to the "Other means" category of SEWAGE DISPOSAL.

April 2003

Population and Housing Profile
Data Note 6

COMMONWEALTH OF THE NORTHERN MARIANA ISLANDS

In the Commonwealth of the Northern Mariana Islands Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the Commonwealth of the Northern Mariana Islands Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

April 2003

Population and Housing Profile
Data Note 7

COMMONWEALTH OF THE NORTHERN MARIANA ISLANDS

The Commonwealth of the Northern Mariana Islands Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the Commonwealth of the Northern Mariana Islands Summary File. The data in the profiles are for the total population while the summary file data are restricted to persons in households similar to other summary file products.

May 2003

Population and Housing Profile
Data Note 1

GUAM

By definition, all people living in group quarters are classified as "did no subsistence activity." Therefore, these people are excluded from the "Subsistence activity" lines shown in the Employment Status section.

February 2002

Population and Housing Profile
Data Note 2

GUAM

Supplemental Security Income (SSI). Supplemental Security Income (SSI) is a U.S. federal assistance program administered by the Social Security Administration that guarantees a minimum level of income for needy aged, blind, or disabled individuals. The census questionnaire for Guam asked about the receipt of SSI; however, SSI is not a federally administered program in Guam. Therefore, it is not the same concept as SSI in the United States. The only way a resident of Guam could have appropriately reported SSI would have been if they lived in the United States at any time during calendar year 1999 and received SSI.

April 2002

Population and Housing Profile
Data Note 3

GUAM

In the Guam Demographic Profile, "Median earnings (dollars)" for male and female full-time, year-round workers is mislabeled. It should read, "Median total money income (dollars)" for male and female full-time, year-round workers.

April 2003

Population and Housing Profile
Data Note 4

GUAM

In the Guam Demographic Profile under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the Guam Summary File.

May 2003

Population and Housing Profile
Data Note 5

GUAM

The Guam Demographic Profile data for median value differ from the Guam Summary File data for median value. The summary file data were rounded to the nearest $100. The demographic profile data were not rounded. The summary file data are not incorrect, these data simply reflect the effects of rounding.

July 2003

Population and Housing Profile
Data Note 6

GUAM

In the Guam Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the Guam Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

July 2003

Population and Housing Profile
Data Note 7

GUAM

The Guam Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the Guam Summary File. The data in the profiles are for the total population, while the summary file data are restricted to persons in households similar to other summary file products.

July 2003

Population and Housing Profile
Data Note 8

GUAM

The values shown in the Guam Summary File for SEWAGE DISPOSAL may difer slightly from those found in the Guam Demographic Profile. The summary file allows a “Sewer/Septic” response even if the source of water for a housing unit is located outside the building and provides only cold water. In the demographic profile, housing units with a source of water that was located outside the building and provided only cold water were assigned to the “Other means” category of SEWAGE DISPOSAL.

April 2004

Population and Housing Profile
Data Note 1

U.S. VIRGIN ISLANDS

"Other West Indies" refers to other places in the Caribbean that are not shown, such as Barbados and Cuba.

February 2002

Population and Housing Profile
Data Note 2

U.S. VIRGIN ISLANDS

Supplemental Security Income (SSI). Supplemental Security Income (SSI) is a U.S. federal assistance program administered by the Social Security Administration that guarantees a minimum level of income for needy aged, blind, or disabled individuals. The census questionnaire for the U.S. Virgin Islands asked about the receipt of SSI; however, SSI is not a federally administered program in the U.S. Virgin Islands. Therefore, it is not the same concept as SSI in the United States. The only way a resident of the U.S. Virgin Islands could have appropriately reported SSI would have been if they lived in the United States at any time during calendar year 1999 and received SSI.

April 2002

Population and Housing Profile
Data Note 3

U.S. VIRGIN ISLANDS

In the U.S. Virgin Islands Demographic Profile, "Median earnings (dollars)" for male and female full-time, year-round workers is mislabeled. It should read, "Median total money income (dollars)" for male and female full-time, year-round workers.

April 2003

Population and Housing Profile
Data Note 4

U.S. VIRGIN ISLANDS

The U.S. Virgin Islands Demographic Profile data concerning grandparents who are living with grandchildren may be slightly different from the data in the U.S. Virgin Islands Summary File. The data in the profiles are for the total population while the summary file data are restricted to persons in households similar to other summary file products.

April 2003

Population and Housing Profile
Data Note 5

U.S. VIRGIN ISLANDS

In the U.S. Virgin Islands Demographic Profile, under "Poverty Status in 1999," the total number of unrelated individuals 15 years and over who are below the poverty level was incorrect. The data should include (1) all individuals in family households who are not related to the reference person; (2) all individuals who live in nonfamily households, including householders living alone or with nonrelatives only; and (3) all individuals in noninstitutional group quarters. However, the data shown exclude householders in nonfamily households. The correct data are available from the U.S. Virgin Islands Summary File.

April 2003

Population and Housing Profile
Data Note 6

U.S. VIRGIN ISLANDS

In the U.S. Virgin Islands Demographic Profile, median incomes below $1,000 were calculated in a manner different from that which was used in the U.S. Virgin Islands Summary File. The demographic profile calculated the median based on a single interval of less than $2,500. The summary file calculated the median based on two intervals of less than $1,000 and $1,000 to $2,499. As a result, there may be differences between the median incomes listed in the demographic profile and the median incomes listed in the summary file.

April 2003

Summary Population and Housing Characteristics (PHC-1) User Note 1

The user should note that there are limitations to many of these data. Please refer to the text provided with this report for further explanations on the limitations of the data.

June 2002

Summary Social, Economic, and Housing Characteristics (PHC-2)
Data Note 1

The user should note that there are limitations to many of these data. Please refer to the text for further explanations on the limitations of the data. See Appendix G of this report and the text found in PHC-2-A, *Summary Social, Economic, and Housing Characteristics, Selected Appendixes.*

March 2003

Summary Social, Economic, and Housing Characteristics (PHC-2)
Data Note 2

The Census Bureau is aware there may be a problem in the Census 2000 employment status data for people enumerated in group quarters. The problem may cause the labor force data for places, particularly those with high concentrations of people in group quarters (such as college towns with large dormitory populations) to overstate the number in the labor force, the number unemployed, and the percent unemployed, and to understate the number of employed. For more information, see the Census 2000 Notes and Errata document at the following Census Bureau Internet site: http://www.census.gov/prod/cen2000/notes/errata.pdf.

March 2003

Summary Social, Economic, and Housing Characteristics (PHC-2)
Data Note 3

Estimated population and housing unit totals based on tabulations from only the sample questionnaires (sample tabulations) may differ from the official counts as tabulated from every census questionnaire (100-percent tabulations). Such differences result, in part, because the sample tabulations are based on information from a sample of households rather than from all households (sampling error). Differences also can occur because the interview situation (length of questionnaire, effect of the interviewer, etc.) and the processing rules differ between the 100-percent and sample tabulations. These types of differences are referred to as nonsampling error. (For more information, see Appendix G.)

The 100-percent data are the official counts and should be used as the source of information on population and housing items collected on the 100-percent questionnaire, such as age, race, Hispanic or Latino origin, and tenure. This is especially appropriate when the primary focus is on counts of the population or housing units for small areas. For estimates of the number of people and housing units by characteristics asked only on a sample basis (such as education, labor force status, income in 1999, or year structure built), the sample estimates should be used within the context of the error associated with them.

Additional information on comparing sample estimates with corresponding 100-percent values is available on the Census Bureau's Internet site at http://www.census.gov/Press-Release/www/2002/sf3compnote.html.

March 2003

Summary Social, Economic, and Housing Characteristics (PHC-2)
Data Note 4

Median incomes for nonfamily households by race were calculated from a 38-category income distribution rather than the standard 39-category income distribution. The 38-category distribution collapsed the two highest categories ($175,000 - $199,999 and $200,000 and over) into a single category of $175,000 and over.

March 2003

Summary Social, Economic, and Housing Characteristics (PHC-2)
Data Note 5

Users may find slight differences in the Occupants Per Room calculations between those found in this report and those found in the Census 2000 Demographic Profile. “Occupants per room” is obtained by dividing the number of people in each occupied housing unit by the number of rooms in the unit. This report, based on Summary File 3, correctly uses a topcode value of “10 rooms” for those occupied housing units with “9 or more rooms.” In the Demographic Profile, an incorrect topcode value of “9 rooms” was used.

March 2002

Summary Social, Economic, and Housing Characteristics (PHC-2)
Printed Report Note 1

The "Accuracy of the Data" chapter describes how to calculate standard errors for most estimates, but not for per capita income, which is described below.

Computing the Standard Error of Per Capita Income

Per capita income is the total income from all sources (salary income, retirement income, public assistance income, etc.) of the people in a population group divided by the number of people in that group.

$$Per\ Capita\ Income = \frac{Aggregate\ Income}{N_{Population}}$$

where $N_{Population}$ is the estimate of total people in the population group.

A similar statistic, mean income, is like per capita income, except that the population measure includes only people at least 15 years of age, since income data is not collected for people younger than that.

$$Mean\ Income = \frac{Aggregate\ Income}{N_{15+}}$$

where N_{15+} is the estimate of people at least 15 years old in the population group.

The two measures are related by the formula:

$$Per\ Capita\ Income = Mean\ Income \times \frac{N_{15+}}{N_{Population}}$$

Hence, the approximate formula for estimating the standard error of per capita income is:

$$SE\ (Per\ Capita\ Income\ ) = SE\ (\ Mean\ Income\ ) \times \frac{N_{15+}}{N_{Population}}$$

Methodology

Calculating the standard error of Mean Income requires the use of an income distribution table. The table must provide frequency estimates of the number of people that fall within certain intervals. Standard available tables may be broken down by sex and whether the individual worked full-time, year-round in 1999. Such a table might look like this:

Table 1. **Sex by Work Experience in 1999 by Income in 1999 for the Population 15 Years and Over - Universe: Population 15 Years and Over**

Total	**32,091**
Male	15,836
Worked full-time, year-round in 1999:	6,000
No income	0
With income:	6,000
$1 to $2,499 or loss	10
$2,500 to $4,999	16
$5,000 to $7,499	44
$7,500 to $9,999	84
.	
.	
.	
$100,000 or more	146

Following the distribution for Male: Worked Full-Time, Year-Round in 1999 ("Wftyr") is a similar distribution for males who did not work full-time, year-round in 1999 (called "Other" in the table) and then females who did and did not work full-time, year-round in 1999.

1. To get the distribution of all people 15 years and older in each income interval, sum the four sex by work-status distributions:

$$N_{15+,j} = Male\_Wftyr_j + Male\_Other_j + Female\_Wftyr_j + Female\_Other_j$$
$$j = 1, 2, ..., number\ of\ intervals$$

2. Sum the frequencies across all intervals j to obtain an estimate of the population total:

$$N_{15+} = \sum_j N_{15+,j}$$

3. Calculate the estimated proportion of people in each income interval:

$$p_j = N_{15+,j} / N_{15+}$$

4. Calculate the mid-point (m) of each income interval from:

$$m_j = \left(L_j + U_j\right)/2$$

where L_j and U_j are the lower and upper bounds of the interval. If the c^{th} interval is open-ended (i.e. has no upper bound), then an approximate value for m_c is:

$$m_c = \frac{3}{2} L_c$$

5. Estimate mean income from:

$$\bar{x} = \sum_j p_j m_j$$

6. Estimate the standard error of mean income from:

$$SE(\bar{x}) = \sqrt{\frac{5}{N_{15+}} \times s^2} \times Design\ Factor$$

where

$$s^2 = \sum_j p_j \ m_j{}^2 - (\bar{x})^2$$

Use the design factor for "Population: Household Income in 1999."

7. An approximation of per capita income can be computed by:

$$Per\ Capita\ Income = \bar{x} \times \frac{N_{15+}}{N_{Population}}$$

8. Multiply the result of Step 6 by the ratio of the person estimates ($\frac{N_{15+}}{N_{Population}}$) to get the approximate standard error for per capita income.

Example

This example shows the steps to estimate the standard error of per capita income for a population group in County A.

1. Sum the frequency estimates in each interval in the four sub-tables of Table 1 to produce a distribution similar to Table 2.

Table 2. **Frequency Distribution for Income, People 15 years and older**

Total Income in 1999	Frequency
No income	8,034
With income:	
$1 to $2,499 or loss	644
$2,500 to $4,999	730
$5,000 to $7,499	876
$7,500 to $9,999	1,299
$10,000 to $12,499	1,350
$12,500 to $14,999	1,438
$15,000 to $17,499	1,599
$17,500 to $19,999	1,688
$20,000 to $22,499	1,871
$22,500 to $24,999	1,766
$25,000 to $29,999	2,331
$30,000 to $34,999	1,923
$35,000 to $39,999	1,345
$40,000 to $44,999	914
$45,000 to $49,999	856
$50,000 to $54,999	1,134
$55,000 to $64,999	828
$65,000 to $74,999	563
$75,000 to $99,999	455
$100,000 or more	447
Total	**32,091**

2. Cumulate the frequencies over the 21 intervals for those with and without income, to get the population base (N_{15+}) of 32,091 people age 15 years and over.

3. Calculate the proportion of people in each interval by dividing the interval's population estimate by the population base. The proportion of people age 15 and over in the "No Income" interval, p_1, is

$$p_1 = \frac{8{,}034}{32{,}091} = 0.2504 .$$

4. Find the midpoint m_j for each of the 21 intervals.

For example, the midpoint of interval 3, "$2,500 to $4,999" is

$$m_3 = \frac{\$2{,}500 + \$4{,}999}{2} = \$3{,}749.50$$

while the midpoint of the 21st interval, "$100,000 or more" is

$$m_{21} = \frac{3}{2}(\$100{,}000) = \$150{,}000$$

The midpoint of the "No Income" interval is zero; for "$1 to $2,499 or loss" it is $1,250. Necessary results for the standard error calculation are given in Table 3 along with totals.

Table 3. **Calculations for Per Capita Income**

Total Income in 1999	p_j	m_j	$p_j m_j^2$	$p_j m_j$
No Income	0.2504	$0.00	$0	$0.00
With Income				
$1 to $2,499 or loss	0.0201	$1,250.00	$31,406	$ 25.13
$2,500 to $4,999	0.0227	$3,749.50	$319,134	$ 85.11
$5,000 to $7,499	0.0273	$6,249.50	$1,066,236	$ 170.61
$7,500 to $9,999	0.0405	$8,749.50	$3,100,427	$ 354.35
$10,000 to $12,499	0.0421	$11,249.50	$5,327,808	$ 473.60
$12,500 to $14,999	0.0448	$13,749.50	$8,469,384	$ 615.98
$15,000 to $17,499	0.0498	$16,249.50	$13,149,503	$ 809.23
$17,500 to $19,999	0.0526	$18,749.50	$18,491,201	$ 986.22
$20,000 to $22,499	0.0583	$21,249.50	$26,324,855	$1,238.85
$22,500 to $24,999	0.0550	$23,749.50	$31,022,131	$1,306.22
$25,000 to $29,999	0.0726	$27,499.50	$54,901,754	$1,996.46
$30,000 to $34,999	0.0599	$32,499.50	$63,267,428	$1,946.72
$35,000 to $39,999	0.0419	$37,499.50	$58,920,304	$1,571.23
$40,000 to $44,999	0.0285	$42,499.50	$51,476,914	$1,211.24
$45,000 to $49,999	0.0267	$47,499.50	$60,240,607	$1,268.24
$50,000 to $54,999	0.0353	$52,499.50	$97,293,772	$1,853.23
$55,000 to $64,999	0.0258	$59,999.50	$92,878,452	$1,547.99
$65,000 to $74,999	0.0175	$69,999.50	$85,748,775	$1,224.99
$75,000 to $99,999	0.0142	$87,499.50	$108,717,508	$1,242.49
$100,000 or more	0.0139	$150,000.00	$312,750,000	$2,085.00
Total			$1,093,497,599	$22,013.00

5. To estimate mean income of people at least 15 years old in the population group in County A, multiply each interval's proportion by its midpoint and sum over all intervals in the universe. Table 3 shows an estimated mean income of people at least 15 years, $\overline{x}$, of $22,013

6. To estimate the standard error of mean income, first calculate the estimated population variance for mean income of people 15 years and older.

$$s^2 = 1{,}093{,}497{,}599 - 22{,}013^2 = 608{,}925{,}430$$

Suppose the person observed sampling rate in County A is 14.5 percent. Suppose the design factor for "Population Household Income in 1999", given in the "Less than 15 percent" percent-in-sample column of the design factor table in the technical documentation, is 1.4. Use this information and the above results to calculate an estimated standard error for the mean income of people 15 years and older as:

$$SE(\overline{x}) = \sqrt{\frac{5}{32{,}091} \times 608{,}925{,}430} \times 1.4$$

$$= \$431$$

Thus the standard error on the mean income of $22,013 is $431.

7. If the total population (including those less than 15 years old) in the population group in County A is 42,297, an approximation to per capita income is:

$$\$22{,}013 \times \frac{32{,}091}{42{,}297} = \$16{,}701$$

8. The standard error of the per capita income is calculated as:

$$SE(Per\ Capita\ Income) = \frac{32{,}091}{42{,}297} \times \$431 = \$327$$

Thus the standard error of the per capita income of $16,701 is $327.

March 2004

Characteristics of American Indians and Alaska Natives by Tribe and Language (PHC-5) Printed Report Note 1

The "Accuracy of the Data" chapter describes how to calculate standard errors for most estimates, but not for per capita income, which is described below.

Computing the Standard Error of Per Capita Income

Per capita income is the total income from all sources (salary income, retirement income, public assistance income, etc.) of the people in a population group divided by the number of people in that group.

$$Per\ Capita\ Income = \frac{Aggregate\ Income}{N_{Population}}$$

where $N_{Population}$ is the estimate of total people in the population group.

A similar statistic, mean income, is like per capita income, except that the population measure includes only people at least 15 years of age, since income data is not collected for people younger than that.

$$Mean\ Income = \frac{Aggregate\ Income}{N_{15+}}$$

where N_{15+} is the estimate of people at least 15 years old in the population group.

The two measures are related by the formula:

$$Per\ Capita\ Income = Mean\ Income \times \frac{N_{15+}}{N_{Population}}$$

Hence, the approximate formula for estimating the standard error of per capita income is:

$$SE\ (Per\ Capita\ Income\ ) = SE\ (\ Mean\ Income\ ) \times \frac{N_{15+}}{N_{Population}}$$

Methodology

Calculating the standard error of Mean Income requires the use of an income distribution table. The table must provide frequency estimates of the number of people that fall within certain intervals. Standard available tables may be broken down by sex and whether the individual worked full-time, year-round in 1999. Such a table might look like this:

Table 1. **Sex by Work Experience in 1999 by Income in 1999 for the Population 15 Years and Over - Universe: Population 15 Years and Over**

Total	**32,091**
Male	15,836
Worked full-time, year-round in 1999:	6,000
No income	0
With income:	6,000
$1 to $2,499 or loss	10
$2,500 to $4,999	16
$5,000 to $7,499	44
$7,500 to $9,999	84
.	
.	
.	
$100,000 or more	146

Following the distribution for Male: Worked Full-Time, Year-Round in 1999 ("Wftyr") is a similar distribution for males who did not work full-time, year-round in 1999 (called "Other" in the table) and then females who did and did not work full-time, year-round in 1999.

1. To get the distribution of all people 15 years and older in each income interval, sum the four sex by work-status distributions:

$$N_{15+,j} = Male\_Wftyr_j + Male\_Other_j + Female\_Wftyr_j + Female\_Other_j$$
$$j = 1, 2, ..., number\ of\ intervals$$

2. Sum the frequencies across all intervals j to obtain an estimate of the population total:

$$N_{15+} = \sum_j N_{15+,j}$$

3. Calculate the estimated proportion of people in each income interval:

$$p_j = N_{15+,j} / N_{15+}$$

4. Calculate the mid-point (m) of each income interval from:

$$m_j = \left(L_j + U_j\right)/2$$

where L_j and U_j are the lower and upper bounds of the interval. If the c^{th} interval is open-ended (i.e. has no upper bound), then an approximate value for m_c is:

$$m_c = \frac{3}{2} L_c$$

5. Estimate mean income from:

$$\bar{x} = \sum_j p_j m_j$$

6. Estimate the standard error of mean income from:

$$SE(\bar{x}) = \sqrt{\frac{5}{N_{15+}} \times s^2} \times Design\ Factor$$

where

$$s^2 = \sum_j p_j \ m_j{}^2 - (\bar{x})^2$$

Use the design factor for "Population: Household Income in 1999."

7. An approximation of per capita income can be computed by:

$$Per\ Capita\ Income = \bar{x} \times \frac{N_{15+}}{N_{Population}}$$

8. Multiply the result of Step 6 by the ratio of the person estimates ($\frac{N_{15+}}{N_{Population}}$) to get the approximate standard error for per capita income.

Example

This example shows the steps to estimate the standard error of per capita income for a population group in County A.

1. Sum the frequency estimates in each interval in the four sub-tables of Table 1 to produce a distribution similar to Table 2.

Table 2. **Frequency Distribution for Income, People 15 years and older**

Total Income in 1999	Frequency
No income	8,034
With income:	
$1 to $2,499 or loss	644
$2,500 to $4,999	730
$5,000 to $7,499	876
$7,500 to $9,999	1,299
$10,000 to $12,499	1,350
$12,500 to $14,999	1,438
$15,000 to $17,499	1,599
$17,500 to $19,999	1,688
$20,000 to $22,499	1,871
$22,500 to $24,999	1,766
$25,000 to $29,999	2,331
$30,000 to $34,999	1,923
$35,000 to $39,999	1,345
$40,000 to $44,999	914
$45,000 to $49,999	856
$50,000 to $54,999	1,134
$55,000 to $64,999	828
$65,000 to $74,999	563
$75,000 to $99,999	455
$100,000 or more	447
Total	**32,091**

2. Cumulate the frequencies over the 21 intervals for those with and without income, to get the population base (N_{15+}) of 32,091 people age 15 years and over.

3. Calculate the proportion of people in each interval by dividing the interval's population estimate by the population base. The proportion of people age 15 and over in the "No Income" interval, p_1, is

$$p_1 = \frac{8{,}034}{32{,}091} = 0.2504 .$$

4. Find the midpoint m_j for each of the 21 intervals.

For example, the midpoint of interval 3, "$2,500 to $4,999" is

$$m_3 = \frac{\$2{,}500 + \$4{,}999}{2} = \$3{,}749.50$$

while the midpoint of the 21st interval, "$100,000 or more" is

$$m_{21} = \frac{3}{2}(\$100{,}000) = \$150{,}000$$

The midpoint of the "No Income" interval is zero; for "$1 to $2,499 or loss" it is $1,250. Necessary results for the standard error calculation are given in Table 3 along with totals.

Table 3. **Calculations for Per Capita Income**

Total Income in 1999	p_j	m_j	$p_j m_j^2$	$p_j m_j$
No Income	0.2504	$0.00	$0	$0.00
With Income				
$1 to $2,499 or loss	0.0201	$1,250.00	$31,406	$ 25.13
$2,500 to $4,999	0.0227	$3,749.50	$319,134	$ 85.11
$5,000 to $7,499	0.0273	$6,249.50	$1,066,236	$ 170.61
$7,500 to $9,999	0.0405	$8,749.50	$3,100,427	$ 354.35
$10,000 to $12,499	0.0421	$11,249.50	$5,327,808	$ 473.60
$12,500 to $14,999	0.0448	$13,749.50	$8,469,384	$ 615.98
$15,000 to $17,499	0.0498	$16,249.50	$13,149,503	$ 809.23
$17,500 to $19,999	0.0526	$18,749.50	$18,491,201	$ 986.22
$20,000 to $22,499	0.0583	$21,249.50	$26,324,855	$1,238.85
$22,500 to $24,999	0.0550	$23,749.50	$31,022,131	$1,306.22
$25,000 to $29,999	0.0726	$27,499.50	$54,901,754	$1,996.46
$30,000 to $34,999	0.0599	$32,499.50	$63,267,428	$1,946.72
$35,000 to $39,999	0.0419	$37,499.50	$58,920,304	$1,571.23
$40,000 to $44,999	0.0285	$42,499.50	$51,476,914	$1,211.24
$45,000 to $49,999	0.0267	$47,499.50	$60,240,607	$1,268.24
$50,000 to $54,999	0.0353	$52,499.50	$97,293,772	$1,853.23
$55,000 to $64,999	0.0258	$59,999.50	$92,878,452	$1,547.99
$65,000 to $74,999	0.0175	$69,999.50	$85,748,775	$1,224.99
$75,000 to $99,999	0.0142	$87,499.50	$108,717,508	$1,242.49
$100,000 or more	0.0139	$150,000.00	$312,750,000	$2,085.00
Total			$1,093,497,599	$22,013.00

5. To estimate mean income of people at least 15 years old in the population group in County A, multiply each interval's proportion by its midpoint and sum over all intervals in the universe. Table 3 shows an estimated mean income of people at least 15 years, $\overline{x}$, of $22,013

6. To estimate the standard error of mean income, first calculate the estimated population variance for mean income of people 15 years and older.

$$s^2 = 1{,}093{,}497{,}599 - 22{,}013^2 = 608{,}925{,}430$$

Suppose the person observed sampling rate in County A is 14.5 percent. Suppose the design factor for "Population Household Income in 1999", given in the "Less than 15 percent" percent-in-sample column of the design factor table in the technical documentation, is 1.4. Use this information and the above results to calculate an estimated standard error for the mean income of people 15 years and older as:

$$SE(\overline{x}) = \sqrt{\frac{5}{32{,}091} \times 608{,}925{,}430} \times 1.4$$

$$= \$431$$

Thus the standard error on the mean income of $22,013 is $431.

7. If the total population (including those less than 15 years old) in the population group in County A is 42,297, an approximation to per capita income is:

$$\$22{,}013 \times \frac{32{,}091}{42{,}297} = \$16{,}701$$

8. The standard error of the per capita income is calculated as:

$$SE(Per\ Capita\ Income) = \frac{32{,}091}{42{,}297} \times \$431 = \$327$$

Thus the standard error of the per capita income of $16,701 is $327.

March 2004