

# **The PSID: Geocode 2000 Census**

## **Codebook**

**Number of Variables**

14

The PSID: Geocode 2000 Census

**RLS00 "RELEASE NUMBER" NUM(1.0)**

Release Number

Count	%	Value/Range	Code Value/Range Text
-	-		1 Release number 1: February 21, 2003
-	-		2 Release number 2: June 13, 2005
-	-		3 Release number 3: February 1, 2006
-	-		4 Release number 4: April 20, 2007
-	-		5 Release number 5: August 2009
250,788	100.00		6 Release number 6: November 2011

**YEAR "INTERVIEW YEAR" NUM(4.0)**

Year of PSID Data Collection

*This variable, in combination with "FAMID", uniquely identified family records in this file.*

*NOTE: Addresses for 1969 were damaged and, therefore, can not be geocoded.*

Count	%	Value/Range	Code Value/Range Text
250,788	100.00	1,968 - 2,009	PSID data collection 1968 - 2009 (excluding 1969)

**FAMID "FAMILY INTERVIEW NUMBER" NUM(5.0)**

Yearly PSID Family Interview Number

*This variable, in combination with 'YEAR', uniquely identifies family records.*

Count	%	Value/Range	Code Value/Range Text
250,788	100.00	1 - 16,970	Family Interview Number

**FIPSTATE "FIPS STATE 2000" NUM(2.0)**

US Census Federal Information Processing Standard (FIPS) State Code.

*The Census Bureau treats the 50 states proper, plus the District of Columbia, as "states" for statistical reporting purposes. (The colonies American Samoa, Guam, Puerto Rico, and the US Virgin Islands are also assigned state identification numbers).*

Count	%	Value/Range	Code Value/Range Text
3,060	1.22		1 Alabama
282	.11		2 Alaska
3,606	1.44		4 Arizona
5,286	2.11		5 Arkansas
24,415	9.74		6 California
3,092	1.23		8 Colorado
1,820	.73		9 Connecticut
172	.07		10 Delaware
3,055	1.22		11 District of Columbia
11,336	4.52		12 Florida

The PSID: Geocode 2000 Census

6,858	2.73	13	Georgia
197	.08	15	Hawaii
197	.08	16	Idaho
8,033	3.20	17	Illinois
6,358	2.54	18	Indiana
4,384	1.75	19	Iowa
952	.38	20	Kansas
4,286	1.71	21	Kentucky
5,379	2.14	22	Louisiana
956	.38	23	Maine
8,895	3.55	24	Maryland
4,877	1.94	25	Massachusetts
11,839	4.72	26	Michigan
3,579	1.43	27	Minnesota
8,953	3.57	28	Mississippi
7,471	2.98	29	Missouri
100	.04	30	Montana
1,731	.69	31	Nebraska
1,156	.46	32	Nevada
298	.12	33	New Hampshire
6,091	2.43	34	New Jersey
767	.31	35	New Mexico
11,485	4.58	36	New York
11,296	4.50	37	North Carolina
111	.04	38	North Dakota
10,834	4.32	39	Ohio
1,356	.54	40	Oklahoma
3,965	1.58	41	Oregon
12,078	4.82	42	Pennsylvania
129	.05	44	Rhode Island
11,874	4.73	45	South Carolina
965	.38	46	South Dakota
4,162	1.66	47	Tennessee
16,234	6.47	48	Texas
1,764	.70	49	Utah

The PSID: Geocode 2000 Census

87	.03	50	Vermont
7,483	2.98	51	Virginia
3,763	1.50	53	Washington
629	.25	54	West Virginia
2,470	.98	55	Wisconsin
212	.08	56	Wyoming
-	-	60	American Samoa
-	-	64	Federated States of Micronesia
3	.00	66	Guam
-	-	68	Marshall Islands
-	-	69	Northern Mariana Islands
-	-	70	Palau
79	.03	72	Puerto Rico
1	.00	78	Virgin Islands
154	.06	98	Foreign country
173	.07	99	NA, DK

**FIPSCNTY "FIPS COUNTY/COUNTRY 2000"**

**NUM(3.0)**

US Census Federal Information Processing Standard (FIPS) County / Country Code.

Foreign address lines have FIPSTATE=98, and 3 digit country code in 'FIPSCNTY' column. Please refer to the text document, foreign\_countries.txt for a list of foreign countries and their respective codes.

Review changes and updates to US county codes at the Census Bureau's web site, <http://www.census.gov>.

Count	%	Value/Range	Code Value/Range Text
250,445	99.86	1 - 996	County or country
343	.14	999	Not found

**ZIP5 "5-DIGIT ZIP CODE"**

**NUM(5.0)**

Postal ZIP Code Identification Number (5-digit)

As defined by the US Census Bureau, ZIP Code is "ZIP (Zone Improvement Plan) Code is a five-, seven-, nine-, or eleven-digit code assigned by the U.S. Postal Service to a section of a street, a collection of streets, an establishment, structure, or group of post office boxes, for the delivery of mail. The Census Bureau uses only 5-digit ZIP codes for the addresses and address ranges in most Census 2000 operations."

The 5-digit ZIP code is current as of geocoding. Occasionally, the ZIP code reported during interview differs from the ZIP code match found during geocoding, i.e., a data entry error. In such cases, the geocode matched ZIP code is kept in the data file.

Count	%	Value/Range	Code Value/Range Text
249,713	99.57	1 - 99,996	ZIP code
1,075	.43	99,999	Inapplicable or not found

The PSID: Geocode 2000 Census

**MSA00 "METROPOLITAN STATISTICAL AREAS 2000" NUM(4.0)**

Metropolitan Statistical Areas (MSA) 2000, as defined by the US Census Bureau, are "a geographic entity defined by the federal Office of Management and Budget for use by federal statistical agencies, based on the concept of a core area with a large population nucleus, plus adjacent communities having a high degree of economic and social integration with that core. Qualification of an MSA requires the presence of a city with 50,000 or more inhabitants, or the presence of an Urbanized Area (UA) and a total population of at least 100,000 (75,000 in New England). The county or counties containing the largest city and surrounding densely settled territory are central counties of the MSA. Additional outlying counties qualify to be included in the MSA by meeting certain other criteria of metropolitan character, such as a specified minimum population density or percentage of the population that is urban. MSAs in New England are defined in terms of minor civil divisions, following rules concerning commuting and population density."

Review changes and updates to MSA codes at the Federal Office of Management and Budget's web site, <http://www.whitehouse.gov/omb>.

Count	%	Value/Range	Code Value/Range Text
199,241	79.45	1 - 9,996	MSA
31,645	12.62	9,998	Inap., not in MSA
19,902	7.94	9,999	Not found

**CBSA "CORE BASE STATISTICAL AREA 2000" NUM(5.0)**

Core Base Statistical Area (CBSA) 2000, as defined by the US Census Bureau, an urban area of at least 10,000 people, based on standards set up by the Office of Management and Budget (OMB) in 2000 and finalized in 2003. These standards are used to replace the definitions of metropolitan areas (MA) that were defined in 1990.

Review changes and updates to CBSA codes at the US Census Bureau's web site, <http://www.whitehouse.gov/omb>.

Count	%	Value/Range	Code Value/Range Text
232,636	92.76	1 - 99,997	CBSA Code
154	.06	99,998	Foreign Country
17,998	7.18	99,999	Not Found

**PLACE "FIPS PLACE CODE" NUM(5.0)**

US Census Federal Information Processing Standard (FIPS) Place Code, as defined by the US Census Bureau, are "a concentration of population either legally bounded as an incorporated place, or identified as a Census Designated Place (CDP) including comunidades and zonas urbanas in Puerto Rico. Incorporated places have legal descriptions of borough (except in Alaska and New York), city, town (except in New England, New York, and Wisconsin), or village."

Note that the Place FIPS code is unique within state. So, to uniquely identify a place, use the Place FIPS code in combination with the State FIPS code.

Review changes and updates to Place FIPS codes at the US Census Bureau's web site, <http://www.census.gov>.

Count	%	Value/Range	Code Value/Range Text
201,637	80.40	1 - 99,996	Place FIPS

The PSID: Geocode 2000 Census

49,151 19.60 99,999 Inapplicable or not found

**TRACT00 "2000 CENSUS TRACT"**

**NUM(7.2)**

Census tract code 2000, as defined by the US Census Bureau, are "a small, relatively permanent statistical subdivision of a county delineated by a local committee of census data users for the purpose of presenting data. Census tract boundaries normally follow visible features, but may follow governmental unit boundaries and other non-visible features in some instances; they always nest within counties. Designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions at the time of establishment, census tracts average about 4,000 inhabitants. They may be split by any sub-county geographic entity."

The Census Tract is unique within state and county. So, to uniquely identify a Census Tract, use the Census Tract in combination with the FIPS State and County. A Census Tract contains one to nine (1-9) block groups.

Review changes and updates to Census Tract codes at the US Census Bureau's web site, <http://www.census.gov>.

Count	%	Value/Range	Code Value/Range Text
250,052	99.71	1.00 - 9,999.97	Tract
8	.00	9,999.98	Not found
728	.29	9,999.99	Inap., U.S. territory or foreign country

**BLOCK00 "2000 CENSUS BLOCK"**

**NUM(4.0)**

Census block is the smallest geographic entity for which the census bureau tabulates decennial census data. Census block is an area bounded on all sides by visible and/or non-visible features shown on a map.

Review changes and updates to Census Block codes at the US Census Bureau's web site, <http://www.census.gov>.

Count	%	Value/Range	Code Value/Range Text
736	.29	0	Inapplicable or not found
250,052	99.71	1,000 - 9,999	Block Value

**BGROUP00 "2000 CENSUS BLOCK GROUP"**

**NUM(2.0)**

Census block group 2000, as defined by the US Census Bureau, are geographic subdivisions of census tracts. Their primary purpose is to provide a geographic summary unit for census block data. A block group must comprise a reasonably compact and contiguous cluster of census blocks.

Review changes and updates to Census Block Group codes at the US Census Bureau's web site, <http://www.census.gov>.

Count	%	Value/Range	Code Value/Range Text
250,052	99.71	1 - 9	Block Group Value
736	.29	99	Not Found

The PSID: Geocode 2000 Census

**ADDRTYPE "ADDRESS MATCH TYPE"**

**NUM(1.0)**

Accuracy Codes of Geocode Match Process

Address Match Type codes are directly derived from Centroid Type. A Centroid Type code is a one-character field identifying the type of centroid that was assigned. The centroid to address match type values are:

- 0 Not a centroid (street address match) --> addrtype = 1
- 2 ZIP+2 centroid --> addrtype = 3
- 4 ZIP+4 centroid --> addrtype = 2
- X 5-digit ZIP Code centroid --> addrtype = 4
- Blank No centroid available --> addrtype = 6

A ZIP+2 centroid coordinate is the latitude/longitude point that is the mathematical average of the coordinates associated with all the ZIP+4s within a ZIP+2. A ZIP+2 also is known as a sector.

A ZIP+4 centroid coordinate is the latitude/longitude point that falls at the middle address of the address range associated with the ZIP+4 according to the USPS.

A 5-digit ZIP Code centroid is the Delivery-weighted centerpoint latitude/longitude of the polygon formed by the 5-digit ZIP Code boundaries. These 5-digit ZIP Code centroids also are described as 'delivery-based centroids'.

Source: USA\_Geo\_002.pdf (document is included in data package).

Count	%	Value/Range	Code Value/Range Text
202,885	80.90	1	Block Face: Match is an exact house number match within a single side of a single street block OR is a unique intersection
2,052	.82	2	Near Match: Match is to a single street block within +/- 100 house numbers of the input, but the correct side of street and correct placement within block are not known
12,030	4.80	3	ZIP+2 Centroid: Match is to a ZIP+2 vicinity
33,085	13.19	4	5-digit ZIP Centroid: Match is to a 5-digit ZIP vicinity
736	.29	6	Ambiguous Match: Match is to multiple street segments

**MATTYPE "DETAILED MATCH TYPE"**

**NUM(2.0)**

Detailed (Address) Match Type Codes.

Data values in this column are directly derived from Match Status Codes. For further details, see page 9 of 'USA\_Geo\_002.pdf' file which is included in data package.

Count	%	Value/Range	Code Value/Range Text
198,431	79.12	1	Match to primary name of street segment
4	.00	2	Intersection match to the "from" end of the segment from the first street in the address
12	.00	3	Intersection match to the "to" end of the segment from the first street in the address
4,437	1.77	5	Alternate name match. Address match to an alternate or "secondary" name of a street in the Tele Atlas® database. A common example: when a U.S. or state highway passes through a town, becoming "Main St."

The PSID: Geocode 2000 Census

1	.00	6 Placeholder match. Matched to a point in the Tele Atlas® database that had been placed earlier as a best estimate until acquisition of better resources.
-	-	7 Alternate name match to Placeholder. Matched to an alternate name of a point in the Tele Atlas® database that had been placed earlier as a best estimate until acquisition of better resources.
332	.13	10 Not a valid two-character state abbreviation. Either a typo or non-covered area (Guam, V.I., etc.).
273	.11	11 Locality not found in list of valid localities. The Postal Service and therefore Tele Atlas® does not list this as a deliverable locality.
23	.01	12 Street address parse error. Incomplete or poorly formatted addresses such as blank fields. The geocoding software was unable to break the address down properly into prefix, street name, type, suffix, directional, etc.
41,629	16.60	14 Street name could not be found. The street name given is not found in the Tele Atlas® street database. This is either due to missing data in the Tele Atlas® database or an invalid address.
4,873	1.94	15 Address range does not exist. The address given is not found in the Tele Atlas® street database on the street given. This is either due to missing address ranges on streets in the Tele Atlas® database or an invalid address number.
753	.30	16 More than one segment with address range. Ambiguity: either due to the address occurring multiple times in the Tele Atlas® street database, or the address is not specific enough (EXAMPLE: "100 Main St" when the Tele Atlas® database contains both "100 N Main St" and "100 S Main St").
20	.01	17 Unable to match intersection. There may be two valid streets that do not intersect or one or both streets that could not be matched. Another possibility is that the two streets intersect in more than one place.