

péter's picks & pans



Péter Jacsó
University of Hawaii

U.S. Gazetteer

American Fact Finder

Population Demographics

The decennial U.S. census has long been a gold mine for demographic information that, among other data, drives marketing campaigns. Although this year a new census is being taken, we shall have to wait some time for the results of the 2000 census to be released. The good news is that the U.S. Census Bureau seems to be ready not only to process the data but also to present them free of charge in a variety of impressive formats. Considering the importance of both the existing and the forthcoming data, two databases from the Census Bureau were selected as picks. What is good news for end-users is bad news for those commercial information providers who have

been selling census data at high prices, but not always adding value in the process. The pan is the excessively high-priced Population Demographics database, where you pay through the nose even when the records have nothing but zeros or, worse, absurd information.



the picks

U.S. GAZETTEER

Commercial online information services made a very good living for

decades off data files compiled by the government. The Web, however, inspired talented people to make government information available free of charge through visually appealing and easy-to-use Web interfaces. There are dozens of examples of free versions of MEDLINE, ERIC, and other databases that are far better than many of their fee-based counterparts. Many government agencies have launched Web versions of several of their data files. The U.S. Census Bureau offers a variety of databases. One is the Decennial Census presented in a somewhat roundabout way through the U.S. Gazetteer database (<http://www.census.gov/cgi-bin/gazetteer>) that is far more than just a gazetteer. It allows searching by counties, place names, MCDs (Minor Civil Divisions), and five-digit ZIP codes, and provides more than 100 demographic measures for the selected geographical area. This is amazing depth for such valuable information. These census data are available, for example, for such small villages as Anaktuvuk Pass in Alaska with a population of about 260. The result is first presented in a capsule format.

You can then choose the data elements that will be displayed in the detailed demographic table, such as population by age, gender, race, median household income, per capita income, and dozens of others.

U.S. Census Bureau

U.S. Gazetteer

Select one of the following matches to your query to look at that place in the Tiger Map Browser or retrieve 1990 Census tables

This is a searchable index. Enter search keywords:

- **Zip Code: 99721** PO Name: Anaktuvuk Pass (AK)
Population (1990): 260
Location: 68 11878 N, 151 679005 W
Browse Tiger Map of area
Lookup 1990 Census zipcode data from [STF3E](#)

Capsule format information from U.S. Gazetteer for demographic information

☐	P80A.	MEDIAN HOUSEHOLD INCOME IN 1989(1)	1
		Universe: Households	
☐	P81.	AGGREGATE HOUSEHOLD INCOME IN 1989(2)	2
		Universe: Households	
☐	P82.	RACE OF HOUSEHOLDER(5) BY HOUSEHOLD INCOME IN 1989(9)	45
		Universe: Households	
☑	P83.	HOUSEHOLD INCOME IN 1989(9)	9
		Universe: Households with householder of Hispanic origin	
☐	P84.	AGGREGATE HOUSEHOLD INCOME IN 1989(1) BY RACE OF HOUSEHOLDER(5)	
		Universe: Households	
☐	P85.	AGGREGATE HOUSEHOLD INCOME IN 1989(1)	1
		Universe: Households with householder of Hispanic origin	
☐	P86.	AGE OF HOUSEHOLDER(7) BY HOUSEHOLD INCOME IN 1989(9)	63
		Universe: Households	
☐	P87A.	RACE OF HOUSEHOLDER(1) BY AGE OF HOUSEHOLDER(7) BY HOUSEHOLD INCOME IN 1989(9)	63
		Universe: White households	

U.S. Gazetteer statistical table with user-defined content

The output format options include HTML, tab-delimited, and CODATA formats, and allow downloading for further processing. The map of the geographical area can be also displayed through a link to the TIGER Map Browser.

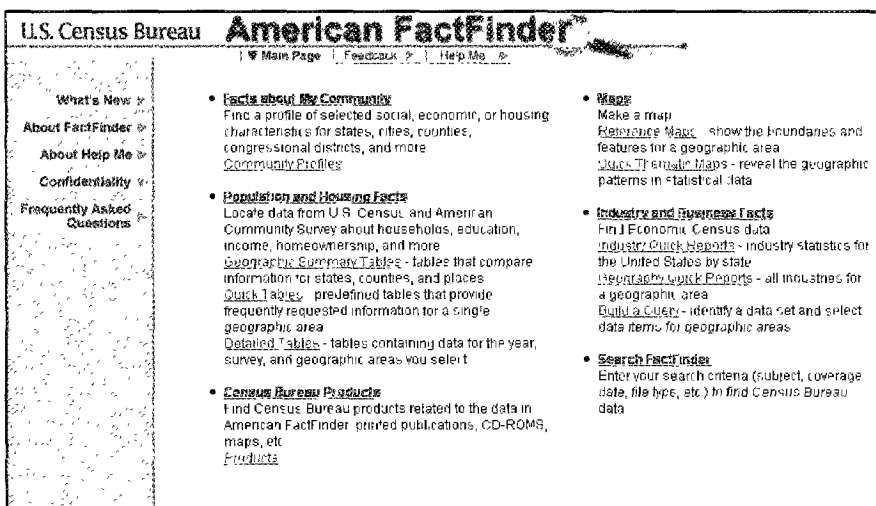
Although you cannot search this treasure trove of data by additional criteria such as ethnic population rate or median household income, the available search options present very powerful options for making a geographically targeted marketing blitz.

Of course, you don't need to settle for 1990 Census data. The census Bureau just released on the Web its 111-page Income 1998 report with many charts and tables (<http://www.census.gov/hhes/www/income98.html>). The State and County Demographic and Economic profiles (<http://www.census.gov/datamap/www/>) also provide fairly current snapshots down to the county level.

that offers utmost customization very elegantly.

In three easy steps, you can get the census data at the level of detail you

need. First, you choose the geographic area by browsing or searching the geographic name index or selecting it on a map. Then you select a demographic profile (a report format), and then create the profile. The template of the profile is displayed showing the data elements that will be included. After viewing the profile, you can download the table in tab-delimited, comma-delimited, Excel, or Word files. If you prefer, you can display statistical maps for selected themes and drill down to the level to which you are interested—and for which data is available. This is an eye-poppingly good feature, plus you can customize some of the parameters for the display of the statistical map. For example, I changed the default number of classes (five) to seven in order to see more refined strata for the median income.

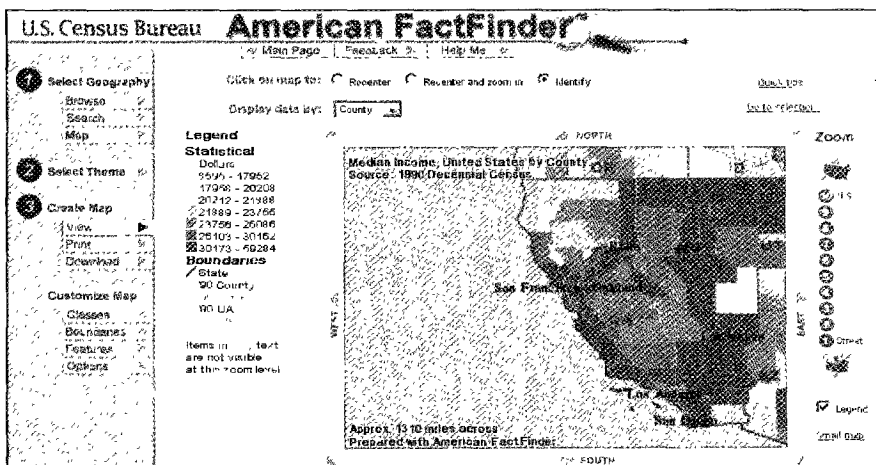


The home page of American FactFinder



AMERICAN FACTFINDER

American FactFinder (factfinder.census.gov) is an absolute state-of-the-art database (or rather data warehouse) developed by the Census Bureau. Instead of rephrasing here what a plethora of information is available, I urge you to check out the home page of this site, then start playing with the JavaScript software



Customized demographics map from American FactFinder with seven strata

101/9/1
 DIALOG(R) File 581:Population Demographics
 (c) 1999 Market Statistics. All rts. reserv.
 00G49070

Level: Zip--96863
 State: HI (HAWAII)
 County: HONOLULU
 Zip Code: 96863
 Census Div: Pacific
 MSA: HONOLULU (3320)
 DMA: HONOLULU (744)

TOTALS & MEDIANS	1993 Census	1998 Estimate	% Change 90 to 98	2003 Projection
Total Population	0	0	0.0%	0
Total Households	0	0	0.0%	0
Total Household Pop	0	0	0.0%	0
Avg Household Size	0.00		0.0%	
Avg Household Income	\$0	\$0	0.0%	\$0
Med Household Income	\$0	\$0	0.0%	\$0
Total Families	NA	NA	NA	NA

One of the apparent 5,990 ghost towns in the Population Demographics database

Oddly, you cannot choose ZIP code as the targeted geographic region. This is not a major problem because geographic names substitute the ZIP codes quite well, such as Kailua instead of 96734, but it would not work for a district of a larger metropolitan area. For example, Honolulu has nearly a dozen ZIP codes to identify a

specific region. Another possible problem is when there are two identical and demographically relevant geographic names within a state (such as Kailua on the island of Oahu, and Kailua on the Big Island), and the casual user does not realize this. The county name should be used as a qualifier to avoid any problems.



the pan

POPULATION DEMOGRAPHICS

Third parties have always loved using data files compiled by the government because the data is practically free. They often enrich the content and add value to it through powerful search software. After Dialog ditched the demographic database of R.R. Donnelley in 1994, it mounted the Population Demographics database from Market Statistics, a division of Claritas, Inc., as File 581. Claritas has many valuable information resources, and won a *DATABASE* Editor's Choice award in June 1999. It did a less-than-perfect job, however, with the 1990 decennial census data file. Dialog's software is powerful and allows searches that the free version does not, such as finding areas where, say, the average household income is

Dream
BIG.



MARCONI
 [586,193]



EDISON
 [223,898]



BELL
 [174,465]

They did.



Let us show
 you how.

MicroPatent presents the world's largest commercial collection of digital patents. With more than 32+ million patents of the world available for immediate download to your PC or printer, breathe new life into your ideas with patent discovery. Each document is provided in full 300 dpi, and is available for delivery in PDF and TIF. With the recipes of the world's greatest inventions fueling your dreams, what isn't possible?

Register Today!
www.micropat.com
800.648.6787

Copyright 1999. MicroPatent

105/9/1
 DIALOG(R) File 581:Population Demographics
 (c) 1999 Market Statistics. All rts. reserv.
 00050414
 Level: Zip--99721
 State: AK (ALASKA)
 County: NORTH SLOPE
 Zip Code: 99721
 Census Div: Pacific
 NSA: (9999)

TOTALS & MEDIANS

	1990 Census	1998 Estimate	% Change 90 to 98	2003 Projection
Total Population	1	1	0.0%	2
Total Households	1	1	0.0%	2
Total Household Pop	1	1	0.0%	2
Avg Household Size	1.00	1.00	0.0%	1.00
Avg Household Income	\$39,000	\$62,000	58.9%	\$36,500
Med Household Income	\$4,999	\$2,500	-49.9%	\$2,500

Command

It is remarkable that a ZIP code is assigned to and census was made of a region where there is only a single not-so-warm body, according to Population Demographics that reports median and average figures.

Total Families	NA	NA	NA	NA
Per Capita Income	\$39,000	\$62,000	58.9%	\$36,500
Buying Power Index	NA	0.000100	NA	0.000100

POPULATION BY AGE

	1990 Census		1998 Estimate		2003 Projection	
	Number	Percent	Number	Percent	Number	Percent
0-5	1	100.0%	0	0.0%	2	100.0%
6-11	0	0.0%	0	0.0%	0	0.0%
12-17	0	0.0%	0	0.0%	0	0.0%
18-24	0	0.0%	0	0.0%	0	0.0%
25-34	0	0.0%	0	0.0%	0	0.0%
35-44	0	0.0%	0	0.0%	0	0.0%
45-54	0	0.0%	1	100.0%	0	0.0%
55-64	0	0.0%	0	0.0%	0	0.0%
65-74	0	0.0%	0	0.0%	0	0.0%
75+	0	0.0%	0	0.0%	0	0.0%
Under 15	1	100.0%	0	0.0%	2	100.0%
65+	0	0.0%	0	0.0%	0	0.0%
Median Age	3.0	NA	50.0	NA	3.0	NA
Total	1	100.0%	1	100.0%	2	100.0%

Command

Life must be very lonely in Anaktuvuk Pass even if the pay rise is good.

above \$75,000, the median age is between 35 and 38, and the Hispanic population is larger than 25,000. These features come at a rather stiff price and a lot of nonsense data. You will find annotated illustrations and the full record at <http://www2.hawaii.edu/~jacso/extra>.

I would have assumed that this very expensive version—\$13.40 for a full record on top of the \$7.35 Dial-Unit charges—would offer a highly enriched version of the data. It doesn't. On the contrary, it retrieves useless records that literally have zero data or absurd data that elementary

checking procedures should have filtered out.

For the full-of-zeros example, see the record for ZIP code 96863 in the county of Honolulu. There are 5,989 other records with zero values for the current population at the time of the 1990 census. These are nice, long records, stretching over eight pages with zeros in every single cell. These are invalid ZIP codes that could have been very easily spotted and removed. The list of valid ZIP codes is available in a digital format for free, for a starter.

I was more intrigued by the statistics for ZIP code areas that according

to Population Demographics have a population of one, such as for the little village of Anaktuvuk Pass. U.S. Gazetteer showed a population of 260 for the very same 99721 ZIP code. It may be a difficult life in the most northern region of Alaska, but the average household income for the single, racially unidentified male is projected to rise 58.9% between 1990 and 1998, a pretty good trend from the perspective of this humble professor whose salary was raised only a few percentage points in the same period. Then again, the median household income for the same period drops by 49.9% from \$4,999 to \$2,500. That's bad, especially when there is no one else to commiserate with. Perhaps median should not be calculated in such cases, and even the most dedicated researchers would spare me the standard deviation and the Pearson coefficient when N=1.

But wait, something else is fishy on the next page of the table. That single guy turns out to be a child in the age range of 0-5. And he makes \$39,000 a year? There's more. The table of Population Demographics estimates that by 1998, the population would remain at one and the poor baby of 1980 will become a man between 45-54. Declining median household income has cruel effects, especially when you are alone in the county of North Slope. Or does it? I really don't know. The 2003 estimate in the Population Demographics version shows the population to double. Strangely, their age range is between 0 and 5. So the single guy who aged at least 40 years between 1990 and 1998 manages to impregnate someone who perhaps came to take the census. She gives birth to two children—twins, maybe?—who are then left there alone, bringing down the average household income further, to \$36,500. I give up here, and gently warn you to look at your marketing plan if it is based on data downloaded from Population Demographics. The more than 6,000 ghost or singleton towns could vastly distort your plans.

Communications to the author should be addressed to P eter Jacso, 322 Aoloa Street, #703, Kailua, HI 96734; 808/956-5817; jacso@hawaii.edu.