

Exercise 2

Accessing Census 2000 SF3 and SF4 Data and Data Bases at ICPSR

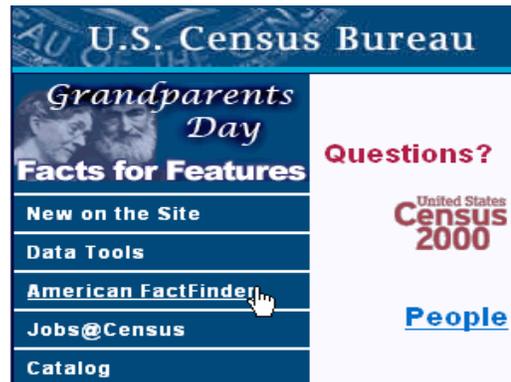
In this exercise you will learn how to do a basic download of census data from summary files. At the time of the actual download your results may vary because of the type of unzipping program you are using, the browser you are using, and what the security settings are on the browser. The Bureau of the Census uses popup features that *Internet Explorer* may not allow in its default setting.

Before proceeding with the exercise it is helpful to review the summary files provided by the Bureau of the Census. Complete-count data is found on SF1 and SF2. Sample data is found on SF3 and SF4. These latter files include data on income, education, occupation, ancestry, and housing that are commonly sought. Both SF2 and SF4 contain tables that are created for numerous race, Hispanic, Native American, and ancestry groups, but there must be at least 50 persons sampled in an area for it to be included.

Summary files contain data for multiple types of geography within a state. The user should know in advance what type of geography is needed and what tables are desired. There is a limit of 7000 areas per download. If the number of total items exceeds the 255 column limit of Excel, the tables will be split into multiple files.

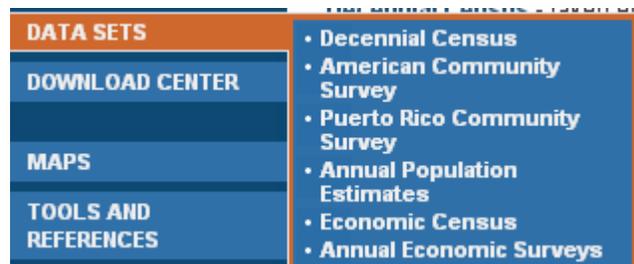
A. Accessing Summary File 3 on the Web

1. If you are on the main Census web page, locate on the left panel the link to *American Factfinder*.



Then from the left panel click on the *Data Sets* button.

From the list of options select the *Decennial Census*.



When the *Decennial Census* page opens you will see a tab for the last two censuses on the top of the page. See right.

Each of these will open a list of summary files for that decade. In the window right the first tab will contain the *Census 2000*.

2. For now, click on the button to select *Summary File 3*. Then click the *Detailed Tables* link to the right.

3. When the data extraction program opens the first thing you should do is select the level of geography you wish. Keep in mind you are limited to 7000 records so if you want all block groups in Los Angeles, you will have to download the raw data files.

Under *Geographic Type*: choose *County* (1).

4. After the window refreshes, choose *California* under *State*:(2).

5. Under the *Select one or more geographic areas...* click on *All Counties* (3) and then click the *Add* button (4) below the window. Wait for the program to display the counties in the window.

6. When the counties have been displayed, click the *Next* button (5). A list of tables will appear.

7. To check out the contents of a table, select one and click on the *What's This?* button to the right. If you want to add a table, select it and then click the *Add* button at the bottom of the window.

Select Tables

You are here: [Main](#) ▶ [Data Sets](#) ▶ [Data Sets with Detailed Tables](#) ▶ [Geography](#) ▶ [Tables](#) ▶ Results
 Census 2000 Summary File 3 (SF 3) - Sample Data, Detailed Tables

■ Choose a table selection method

[by subject](#) [by keyword](#) [show all tables](#)

■ Select one or more tables and click 'Add'

- P1. Total Population
- P2. Unweighted Sample Count of the Population
- P3. 100-Percent Count of the Population
- P4. Percent of the Population in Sample
- P5. Urban and Rural
- P6. Race
- P7. Hispanic or Latino by Race
- P8. Sex by Age
- P9. Household Type by Relationship
- P10. Household Size by Household Type by Presence of Own Children <18 Years

Abbreviations:
 Black - Black or Afri
 AIAN - American Ind
 NHPI - Native Hawaii
 SOR - Some Other F

[What's this?](#)

[Add](#) ▼

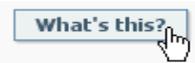
Current table selections:

[Remove](#)

[Show Result](#) ▶

8. Select the *PCT16* table (**First Ancestry Reported**) and click the *Add* button. Eventually the table will be listed in the window. Sometimes the operation of this program will take a little while and you can cause it to crash if you start clicking on various buttons after a step has been started.

9. To the right of the screen select a button labeled *What's this?*



You will find this window very helpful for examining the contents of any table in which you are interested. **Check out the table of ancestries and then close it.**

10. Now click *Show Result*.

Program: [Decennial Programs](#)
 Census: [Census 2000](#)
 Census Instance: [Census 2000 United States](#)
 [Census 2000 Puerto Rico](#)
 Data Set: [Census 2000 Summary File 3 \(SF 3\) - Sample Data](#)
 ▶ Table: **PCT16 ANCESTRY (FIRST ANCESTRY REPORTED) [111]**
 Universe : Total population

Subject Characteristics: [Ancestry](#)

Example:	Geography
Total:	
First ancestry reported:	
Acadian/Cajun	
Afghan	
Albanian	
Alsatian	
Arab:	
Egyptian	
Iraqi	
Jordanian	
Lebanese	

The program will then list part of the results for you to browse. (See below)

U.S. Census Bureau
American FactFinder

Main Search Feedback FAQs Glossary Site Map Help

Detailed Tables

You are here: [Main](#) > [All Data Sets](#) > [Data Sets with Detailed Tables](#) > [Geography](#) > [Tables](#) > [Results](#)

Use the links above to change your results | [Options](#) | [Print / Download](#) | [Related Items](#)

Note: use download to retrieve all selected tables and geographies

[PCT16_ANCESTRY \(FIRST ANCESTRY REPORTED\) \[111\] - Universe: Total population](#)
Data Set: [Census 2000 Summary File 3 \(SF 3\) - Sample Data](#)

geographies 1-10 of 58 [Next](#)

NOTE: Data based on a sample except in P3, P4, H3, and H4. For information on confidentiality protection, sampling error, nonsampling error, and definitions see <http://factfinder.census.gov/home/en/datatnotes/expst3.htm>.

	Alameda County, California	Alpine County, California	Amador County, California	Butte County, California	Calaveras County, California	Colusa County, California	Contra Costa County, California	Del Norte County, California	El Dorado County, California	Fresno County, California
Total:	1,443,741	1,208	35,100	203,171	40,554	18,804	948,816	27,507	156,299	799,400
First ancestry reported:	1,261,632	1,058	25,683	159,098	31,570	15,442	824,440	19,303	128,770	663,920
Acadian/Cajun	79	2	5	48	0	0	34	0	15	5
Afghan	7,917	0	0	13	0	0	2,900	0	33	8
Albanian	131	0	0	0	0	0	59	0	0	13
Alsatian	43	0	0	0	0	0	23	0	7	0
Arab:	6,199	0	72	608	70	0	4,285	21	284	2,240
Egyptian	666	0	0	32	0	0	640	8	11	180
Iraqi	147	0	0	0	0	0	98	0	0	2
Jordanian	197	0	0	31	0	0	170	0	0	5
Lebanese	1,322	0	27	178	15	0	1,011	0	99	53
Moroccan	175	0	1	16	0	0	86	0	0	3

In this case the first ten California counties have been listed. You can click the *Next* link to see the next ten. Note that the counties are listed as columns and each of the ancestries in a row.

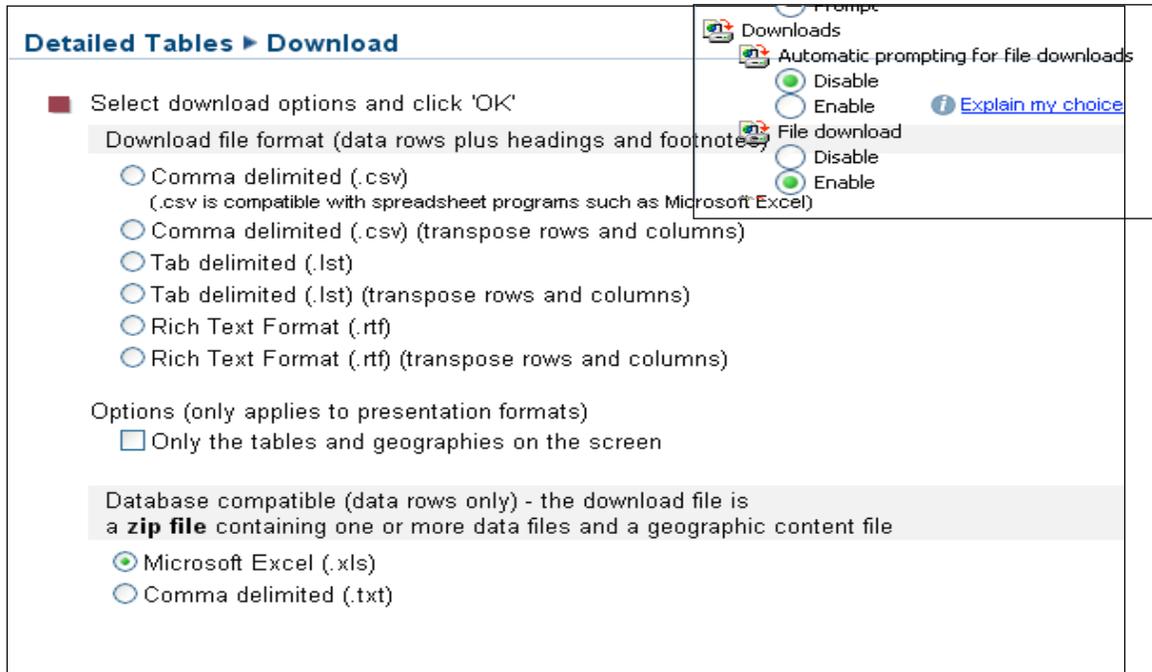
If you are only looking for some specific data you can locate a desired county and statistic and write it down or print out the contents of the page. However, in most cases you will want the contents of the entire table.

At this point you will notice that all counties are identified by their names and that no FIPS codes have been included. When you download this table a unique identifier (a combination of state and county FIPS codes labeled GEOID2) will be added to each row.

The provided FIPS codes may on occasion be converted in a spreadsheet from a character variable to a numeric variable and lose their leading zeros in subsequent processing. Thus, county code 005 would become 5 and code 037 would be 37. This is a nuisance, but simple to fix.

11. Select the *Print/Download* menu at the top of the screen and choose the *Download* option.

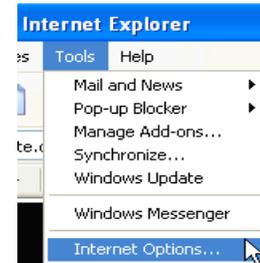
When the *Download* window opens, you may choose the *Database compatible* button for *Microsoft Excel* format (preferred) or you may choose the *CSV* version with rows and columns transposed. Then click *OK*.



Note that you will likely have download problems because of security settings in Explorer. You should look for a message at the top of the web page if the downloading does not seem to be taking place. Also, be sure to place the downloaded file in the *TEMP* directory, your personal directory, or a flash drive.

To enable downloads in Internet Explorer go to the *Tools* menu and the *Internet Options* option.

When the *Internet Options* window opens, click on the *Security* tab and then choose the *Custom Level* button.

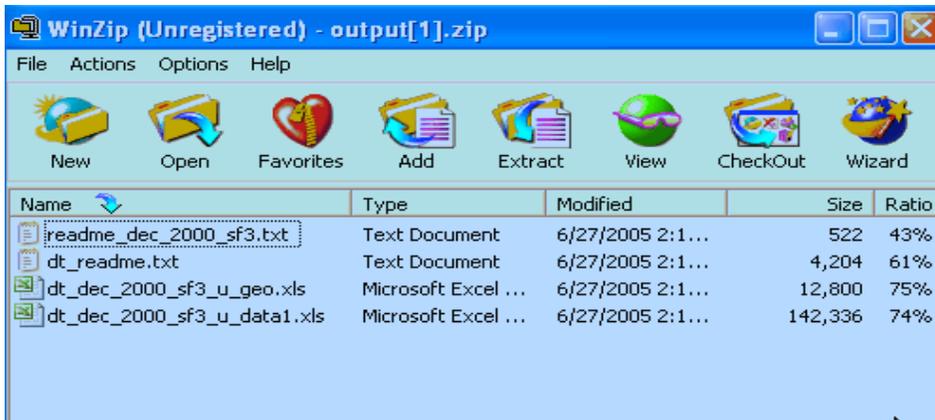


Scroll down through the list of options and make sure the file download is set to *Enable*.

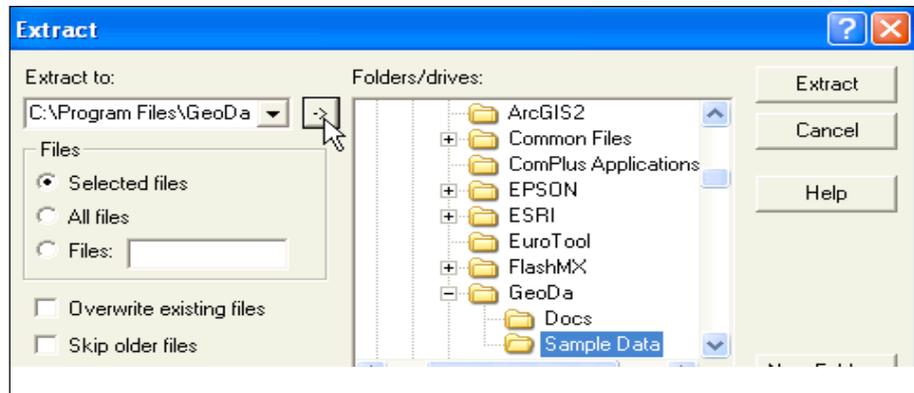
If you get a download warning you will need to again download the data after resetting the permissions.

B. WinZip

Once the download settings are correct the *WinZip* program will open with the contents of the Census download. Notice that you will get two *Excel* files. One (*geo*) contains the geographic descriptions for each county and the other (*data*) contains both the geography and the data values if you opted for the *Show Geographic Identifiers* under the *Options* menu. Read the text files if you want to.



1. Select all files. Click the *Extract* icon. When the *Extract* window opens (see below), click on the arrow shown below to navigate to a location open to you such as the *TEMP* directory or your personal directory.



2. ☒ After you have saved the files go to your directory and change the names of the geo and data files so that any future downloads from the Census Bureau will not over write on them. Unfortunately the names given by the download program are always the same.

3. Open the data file in *Excel* and look it over. You will use it in a future exercise.

This completes this exercise. Close *Excel*

C. The Geo within Geo Tab

An additional useful tool in selecting data for downloading is the *geo within geo* tab. This allows you to download a set of census units that are contained by a larger census unit. For example, you could download all tracts that fall entirely within a city or all block groups that fall within a county. Neither of these options is available under the default List method you just used.

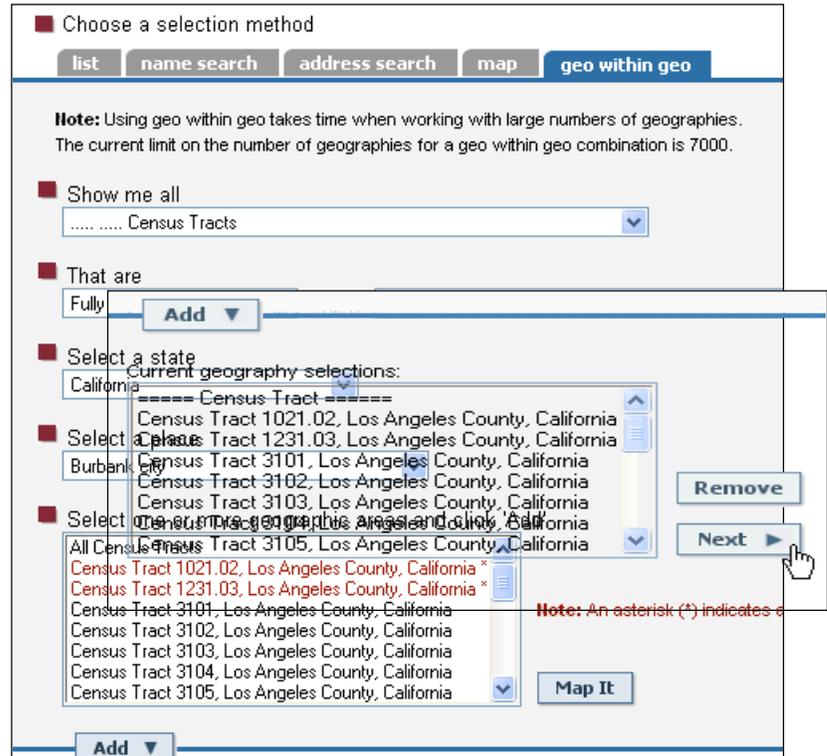
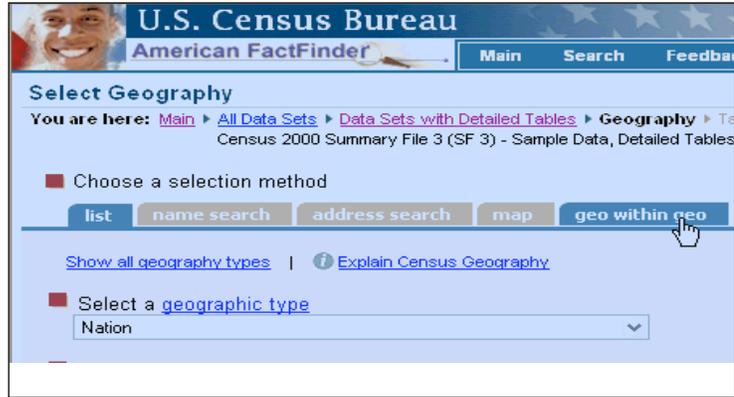
Note you do not need to do the following. Just read the steps to get a sense of how the program works.

If you had used the *Geo within Geo* you would have done the following steps.

a. Select the *geo within geo* tab.

b. From the geography selection window below under ‘*Show me all*’, *Census Tracts* were chosen. Under ‘*That are*’, the default *Fully or partially contained* was chosen. Under ‘*within*’, *Place* was chosen. Under ‘*Select a state*’, *California* was chosen. Under ‘*Select a place*’, *Burbank city* was chosen.

After the final selection all tracts that fall within the city are listed under the ‘*Select one or more geographic areas*’ window. Select the *All Census Tracts* option. Some tracts, such as the first two with brown letters, partially fall outside the city boundary. Thus, only those portions of tracts that lie within the city will be extracted. Note the content of the above window will change with your choices.



c. Click the *Add* button. All the tracts are listed

d. Click the *Next* button to select a table and download the data as before.

One thing to keep in mind here is that if you intend to map the city-only tracts that you must find a boundary file that contains only those tracts that fall within the city. These are generally not provided in free data sets. In most case only whole tract boundary files are provided and these often cross place boundaries, but not county boundaries.

D. Summary File 4

Summary File 4 contains more detail than other files, but most importantly, it is available for many individual ethnic groups. Because of this, there are some differences in the *American Factfinder* menus.

1. On the *American Factfinder* census web page select *Data Sets > Decennial Censuses*.

2. Make sure the *Census 2000* tab is selected and then scroll down and click the button to *Census 2000 Summary File 4*. Select the *Detailed Tables* link.

Census 2000 Summary File 4 (SF 4) - Sample Data
Summary File 4 contains tabulations of population and housing data collected from a sample of the population. The data are shown down to the census tract level for 336 race, Hispanic or Latino, American Indian and Alaska Native, and ancestry categories.

[SF 4 Thresholds](#)

Select from the following:

- [Detailed Tables](#)
- [Geographic Comparison Tables](#)
- [Quick Tables](#)
- [Reference Maps](#)
- [Custom Table](#)
- [Enter a table number](#)
- [List all tables](#)
- [About this data set](#)
- [Technical Documentation \(PDF\)](#)

Census 2000 American Indian and Alaska Native Summary File (AIANSF) - Sample Data
The American Indian and Alaska Native Summary File contains tabulations of population and housing data collected from a sample of the population. The data are shown for the U.S., regions, divisions, states, and selected metropolitan areas for many specified American

This will open the data selection program.

3. For the *geographic type* select *County*

For *Select a state* choose *California*

For the *Select one or more geographic areas* choose *All Counties* and click the *Add* button.

Then click *Next*.

Choose a selection method

list name search address search map geo within geo

Show all geography types | Explain Census Geography

Select a geographic type
..... County

Select a state
California

Select one or more geographic areas and click 'Add'

All Counties
Alameda County
Alpine County
Amador County
Butte County
Calaveras County
Colusa County
Contra Costa County

Map It

Add

Current geography selections:

==== County =====
Alameda County, California
Alpine County, California
Amador County, California
Butte County, California
Calaveras County, California
Colusa County, California
Contra Costa County, California

Download data for more than 7,000 geographic areas using the [Download Center](#).

Remove

Next

4. From the list of tables add *PCT1*, *PCT89*, and *HCT2*.

Then click *Next*.

Choose a table selection method

by subject by keyword **show all tables**

Select one or more tables and click 'Add'

HCT1. Urban and Rural
HCT2. Tenure
HCT3. Tenure by Age of Householder
HCT4. Total Population in Occupied Housing Units by Tenure
HCT5. Household Size
HCT6. Tenure by Household Size
HCT7. Average Household Size of Occupied Housing Units by Tenure
HCT8. Tenure by Household Type and Presence and Age of Own Children
HCT9. Tenure by Household Type (Including Living Alone) by Age of Householder
HCT10. Tenure by Household Size by Age of Householder

Add ▾

Current table selections:

PCT1. Total Population
PCT89. Median Household Income in 1999 (Dollars)
HCT2. Tenure

5. The *Select Population Groups* window is unique to SF2 and SF4. It is here you will choose the ethnic groups for which you would like to obtain the tables. Note there is a tab for races and a tab for ancestries.

The tables for the total population have already been added to your list. To choose a particular group you first must click in the top window on the Table with which you want to work. In this example it is *PCT1*.

Race or Ethnic Groups Ancestry Groups

Select one or more tables

PCT1. TOTAL POPULATION [1] - Universe: Total population
PCT89. MEDIAN HOUSEHOLD INCOME IN 1999 (DOLLARS) [1] - Universe: Households
HCT2. TENURE [3] - Universe: Occupied housing units

Select one or more race or ethnic groups and click 'Add'

Total population
White alone
White alone or in combination with one or more other races
Black alone
Black alone or in combination with one or more other races
AIAN alone
.. American Indian alone
.. Alaska Native alone
AIAN alone or in combination with one or more other races
.. American Indian alone or in any combination

[Explain Race or Ethnic Groups](#)
[Print-ready Race or Ethnic Groups list](#)

Note: the following abbreviations are used to conserve space:
Black - Black or African American
AIAN - American Indian and Alaska Native
NHPI - Native Hawaiian and Other Pacific Islander

Add ▾

Current population group selections:

==== PCT1. TOTAL POPULATION [1] - Universe: Total population
Total population
==== PCT89. MEDIAN HOUSEHOLD INCOME IN 1999 (DOLLARS) [1] - Universe: Households
Total population
==== HCT2. TENURE [3] - Universe: Occupied housing units
Total population

Remove Show Result ►

6. Scroll down the list of race groups and select *Chinese alone* and click the *Add* button. Note where it appears below.

Also select *Japanese* and *Korean alone*.

7. Click on the *PCT89* table in the top window. Then select and add *Chinese*, *Japanese*, and *Korean alone*.

8. Select the *HCT2* table in the top window and again select the three groups. Then click *Show Result*.

9. Scan down the list of tables to Median Household Income (*PCT89*). You will notice that values are missing for a number of counties because there was not at least 50 ethnic persons in the sample. Look over the table below and compare the median incomes of the three Asian groups. What group generally does better? Which does worse?

	Alameda County, CA	Alpine County, CA	Amador County, CA	Butte County, CA	Calaveras County, CA	Colusa County, CA	Contra Costa County, CA	Del Norte County, CA	El Dorado County, CA	Fresno County, CA
Median household income in 1999: Total	55,946	41,875	42,280	31,924	41,022	35,062	63,675	29,642	51,484	34,725
Chinese	58,707			36,125			74,424		82,802	47,083
Japanese	56,333			18,981			65,806		65,806	45,000
Korean	46,356						58,897			49,211

10. You can download the data if you would like to examine the incomes for more counties of CA.

E. Census Data at ICPSR

The Interuniversity Consortium for Political and Social Research (*ICPSR*) contains a wealth of census data and archives that could be valuable to anyone needing pre 1990 census information. Some of these data were originally compiled in digital form while other data represent special tabulations or collections that were converted from text to digital.

Unfortunately, use of this data will take much more effort since it is in raw form typically designed for computer tapes. In most cases there are not programs written to read the raw files and so a user will need to develop them using provided documentation.

1. To browse the ICPSR census data holdings enter the following in your browser:

<http://www.icpsr.org/>

The screenshot shows the ICPSR website homepage. At the top, the logo reads "ICPSR INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH". Below the logo is a navigation bar with five tabs: "Data", "Courses & Learning Tools", "Our Research", "Membership", and "About ICPSR". A banner image shows several people in a meeting. Below the banner is a "SEARCH" section with a search box, a dropdown menu set to "all fields", and checkboxes for "Web Site", "Data Holdings", and "Publications based on our data". There is also a "Search" button and links for "Advanced Search" and "Searching Tips". To the right is an "ANNOUNCEMENTS" section with an RSS icon and three entries: "2007.05.21 Subscribe to ICPSR announcements through RSS feeds...", "2007.05.21 New data releases...", and "2007.05.14".

2. Under the *Search* window select the *Advanced Search* link.

3. In the *Advanced Search* menu enter *census* under the first condition.

Search ICPSR's data holdings for documents that
 must contain in any field the words

4. Under the second condition enter *jail* and click on the option *must not contain*.

and
 must not contain in any field the words

5. Under the third condition enter *juvenile* and click on the option *must not contain*.

There are hundreds of data sets that involve jail censuses.

5. Click the *Search* button.

In the *Search Results* page there were still over 1500 data sets found. Note under each that there is a *description*, a *download* option, and a link to *related literature*.

6. Click on the *description* link to any data sets that seem interesting to you. You do not need to download any data sets unless you are prepared to create a program to read them. *SPSS* could be used to do this.

Search Results

If your search results are too broad, or if you're having trouble understanding the display of results below, you may wish to read our FAQ on [Searching the Web Site](#).

Results for: +census (1874) -jail (327) -juvenile (467)	Document count: +census (1874) -jail (327) -juvenile (467)
1543 results found, top 500 sorted by relevance score by date or sort by title 1-25 	
13569 Census of Population and Housing, 2000 [United States]: County-to-County Migration Flow Files	2007-03-07
United States Department of Commerce. Bureau of the Census description download related literature	
7529 Demographic, Social, Educational and Economic Data for France, 1833-1925	2007-02-16
Inter-university Consortium for Political and Social Research (ICPSR) description download related literature	
166 TIGER/Line Files at ICPSR: 1990 Census	2006-12-13
University of Michigan. Inter-university Consortium for Political and Social Research description	