GeoLytics®

CensusCD® 40 Years

User Guide

(Release 1.05)

Contents

1. Installation 2
2. Introduction 3
3. Five Steps to Producing Files and Maps 4
   Step A: Selection of File to Name the File 4
   Step B: Selection of Year to be Covered by the File 5
   Step C: Selection of Area to be Covered by the File 5
   Step D: Selection of Counts to be Included 6
   Step E: Running and Viewing the File 6
4. Using the Map Viewer 9
5. Notes About the Data and Geography 10
6. Help 11
7. Contacting GeoLytics, Inc. 11
8. Addendum A. Weighting methodology 12

It is very important that this serial number be entered exactly as shown, with no spaces or dashes, only numeric or alpha values. Do not use the serial number from another GeoLytics product, which may appear to install correctly, but will not produce accurate results.

GeoLytics® gratefully acknowledges the support of the Rockefeller Foundation and the Urban Institute, in the development and financing of CensusCD 40 Years.

The source of the 1970-1990 data was from the Urban Institute's National Change Database. The data was weighted and converted to 2000 tracts by GeoLytics, who also created the 1970-2000 tract boundaries.
1. Installation

1. Insert the CensusCD 40 Years into your CD-ROM drive.

2. Click the Start button on the taskbar and choose Run from the Start menu.

3. Type D:/setup.exe (assuming D is the letter of your CD-ROM drive).

4. Click OK.

5. Follow the instructions on the screen.

---

**Note:** If you have previously installed CensusCD 40 Years, you will be prompted to uninstall it before continuing. Choose OK.

6. When prompted, enter you Name, the name of your Organization, and your **Serial Number**. Type the serial number **without spaces or dashes**. It is very important that the number is entered exactly.

7. After the installation is complete, you can start CensusCD 40 Years at any time by double clicking the CensusCD 40 Years icon or by choosing Programs on the Start menu and selecting CensusCD 40 Years.

---

**Note:** To operate the program, the original CD must be in the same CD-ROM drive from which the program was installed. Otherwise, the program will function, but you will receive errors when running reports.

---

Copyright 1996-2001 by GeoLytics, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means electronic, mechanical, photocopying, recording, or otherwise without the prior written permission of GeoLytics, Inc. CensusCD is a trademark of GeoLytics, Inc. All other trademarks are the property of their respective holders. CensusCD software license is a single user version only and the use of any extracted data or maps is restricted to within your own organization.

You must agree to the **End User License Agreement** (EULA) for CensusCD 40 Years during Setup or the product will not be installed and you will have no right to use the product or resulting data. A copy of this EULA can be found under the Help menu item. You can view a copy of the EULA in advance, before purchasing or opening any GeoLytics product by downloading it from the GeoLytics website at: www.geolytics.com/downloads/pdf/eula_geolytics_single_user.pdf

September 2001
2. Introduction

CensusCD 40 Years is a single Windows CD-ROM with the most convenient and complete collection of 40 years of demographic data and software you’ll find anywhere. This one disc holds source data from many dozens of tapes and CD-ROMs, yet, creates reports and maps quickly thanks to revolutionary compression methods developed by GeoLytics, Inc.

The software to select, report, and map this data is intended to be as powerful as the data itself while, at the same time, easy enough for anyone to use. CensusCD 40 Years is an invaluable resource for anyone who needs details about the population and housing dynamics of the United States. The program features options for customizing results and has several different ways to quickly get answers.

This Guide is intended to give a quick initiation and some detailed explanations about the software and terminology. The on-line help section has more detailed definitions of data terminology and technical documentation from the Census Bureau to help make sense of the language of demographic statistics. A specific help section to explain the map viewer is available from both the main help contents and the Map Viewer itself.
3. Five Steps to Producing Files and Maps

A. Selection of File to Name the File
B. Selection of Year to be Covered by the File
C. Selection of Area to be Covered by the File
D. Selection of Counts to be Included
E. Running and Viewing the File

Step A: Selection of File to Name the File

The File menu contains commands that let you determine what to name your request and where to save the output files. A request is built interactively by using CensusCD 40 Years dropdown menus, lists, and dialogs. From the File menu, you can start a new request, open an existing request, or save a request to retrieve it later.

Note: A request tells CensusCD 40 Years what information you want and how you want it. The Request file stores information about the Year(s), Area and Subarea, which is passed to the database to produce a file.

New Request

When creating a new file, the first thing to do is name your request.

To start a new series of files and control the path and name of subsequent files, select New request in the File menu. If you omit this step, CensusCD 40 Years will automatically name the new request Noname.

Open Request

To open an existing request, choose Open request on the File menu and select the file you want to open.

Save Request

The Request File controls where the file will be written. Changing the drive, directory, or name of the request file will control where the file is produced. For example, if the request file name is changed to a:myfile.req, then the file can be written as a:myfile.rpt (Summary report), a:myfile.csv (ASCII file), or a:myfile.dbf (Dbase output).
To save the current request parameters, choose *Save request* and change the file name or location where the file is stored.

You control where the output files are written and what to call them by selecting *New request* or *Save request* and changing the request file's drive, directory, or name. *CensusCD 40 Years* will save the request file and the output file you run from it to the location you select.

**Step B: Selection of Year to be Covered**

The selections you make in the *Year* menu determines the year(s) to be covered by the file. You can make one of five selections: *1970*, *1980*, *1990*, *2000*, and *All years*.

The *1970*, *1980*, *1990*, and *2000* selections will generate reports for the selected year, using the year Census Tract data and year-specific tract IDs. The *All years* selection allows you to compare data for various years. The data for years 1970-1990 will be recalculated for 2000 Tracts and the report will use the 2000 tract IDs (see *Addendum A*).

**Step C: Selection of Area to be Covered**

The selections you make in the *Area* menu determines the area to be covered by the file. You can choose one of two area types: *Geographic Area* and *Radius*.

**Geographic Area**

In the Geographic Area window you can easily select the geographical area you want to use from any of the 6 geographical levels.
Simply select the geographic area you want from the dropdown lists or click the button on the toolbar to display the Geographic Area window.

The levels of geography are:

- Nation (United States)
- States
- Counties
- Tracts
- MSA/CMSA
- PMSA

When you select a geographical area other than Nation, the Geographic Area window appears so you can select which areas you want to use. For example, to select the state (or states) to use, choose **State** and then click the state (or states) you want.

**Searching for a Geographic Area**

You can search for any area and counts that contains a specific word or character string by clicking the **Search** button. Type a word to search for and click **Find**.

**Radius**

The second type of area is **Radius** (circular ring area). You may decide to use this type of area view to get information around a specific location. Use an exact **Latitude** and **Longitude** to specify the center of the Radius. Then, enter the distance from the center of the radius to specify the radial area size (for example, 20 miles from a specific location).

The report will cover the selected areas. Depending on the report type, it will a summary of the data for these areas or it will be broken down by Census Tracts.

**Step D: Selection of the Counts to be Included**

When you select **Counts** menu, the Count window appears. Counts are listed in alphabetical order for the year you have selected. If you selected **All years**, the Count window will have 4 buttons on the left side, one for each year. If you click on a year button, the year counts are added to the window. You can also search for any counts using any words from count definitions by clicking the **Search** button.

**Step E: Running and Viewing the File**

The last step is to run the file. First, choose **ASCII, .DBF, or Map**. Within **ASCII**, you can select **comma-** or **tab-delimited** output, with or without **header** information.

For example:

Lets pick **New York state** as the area we want information for.
1. Select the menu item **Area**
2. Select **2000** as a **Year**
3. Select **Geographic Area**
4. Select **States**
5. In the geographic window that will appear, select **New York** as the state
6. Select **Done**
7. Select **Counts**, choose **Unique 2000 census tract ID** and **2000 Total Population**
8. Select **Done**
9. Select the menu item **Run**
10. Lastly, select **Map**

The file run is done in a multitasking mode so that you can go on selecting other files or doing other tasks while the system continues to access the database for the information you requested. When the file is completed, a viewer/map window will be created and activated, and the completed file will be displayed or mapped.
Since you have selected **Map**, you will see the **Map Viewer** window:

![Map Viewer Window](image)

You can view any file using the **CensusCD 40 Years** viewer. While browsing the file, you can search for information, copy to clipboard, or print the file. You can also have some statistical data for selected columns, including the mean, median, standard deviation and correlation matrix.

Finally, you can use the file data as an input to other systems like **statistical** (e.g. SAS, SPSS), **database** (e.g. Access, Oracle), **spreadsheet** (e.g. Excel, 1-2-3), or **mapping** (e.g. Arc View, MapInfo) packages.
4. Using the Map Viewer

Within the Map Viewer you can change data themes, ranges, and color schemes on the fly. It allows you to print your map with options to save it as a .bmp file, which can be imported or pasted into most word processing, spreadsheet, and graphics packages. **CensusCD 40 Years** also lets you export boundaries and data in desktop mapping formats (ArcView or MapInfo). There is even a variable calculator for creating and displaying virtual variables based on data in your file.

HINT: You can even use the Map Viewer to display your own .DBF data files if you have a key column which lets the program match boundaries to your data records. See the Map Guide under Help for more information on how to display your own .DBF files.

A specific help section to explain the map viewer is available from both the main help contents and the Map Viewer itself.
5. Notes About the Data and Geography

The core set of data in CensusCD 40 Years is for 6 levels of geography down to the Census Tract, the Census Bureau's statistical equivalent of a large neighborhood (with an average of about 7,000 people). This one disk holds the results of four US Censuses for these 6 levels and the corresponding map boundaries which encompass over 300,000 geographic areas in the United States.

If you pick any spot in the US you’ll find it is part of a hierarchy of geography which defines the area around that location politically (in terms of county and state) and statistically (the government’s survey geography and Metropolitan area designations). CensusCD 40 Years lets you easily dissect areas by Census Tracts. Even more remarkable is the ability of the program to map all of this data in color no matter what the geographic level.

You can easily select multiple states for instance and get all of the component tracts. If your area of interest is New York City and want to see it broken down by tracts then select the 5 counties, New York (Manhattan), Richmond (Staten Island), Bronx, Queens, and Kings (Brooklyn), which contain the 5 boroughs of New York. If you think of New York City as including its outlying suburbs then selecting New York PMSA would add Rockland, Putnam, and Westchester Counties to the 5 found in the city. Alternatively you could select a point within New York area and get all the tracts within a certain distance from this point.
6. Help

CensusCD 40 Years contains extensive on-line help. Help contents contain thousands of lines of on-line searchable help, including the scanned 1970 technical documentation of the Fourth Count from the US Census Bureau. You may rely on “automatic advice” located on the bottom of the screen for “next step” prompting. Help can also be requested by placing the focus on the item in question, and pressing the F1 key.

The on-line help contains additional information about selecting, mapping, and exporting 1970 data and boundaries. Documentation provided by the creators of the data compiled into CensusCD 1970, is also included. A specific help section to explain the map viewer is available from both the main help contents and the Map Viewer itself.

7. Contacting GeoLytics, Inc.

GeoLytics makes itself accessible to the customer in various ways, with each medium playing a specific role. Our support e-mail address caters to those who have technical questions about the product or are seeking information about the data or how to use the product. Please, do not call the 800 number for tech support. Our technicians do not work in that department. The 800 number should only be used when installing product for the first time and for sales purposes.

For more information on our support policies, please visit www.geolytics.com/support/supportpolicies.html.

Web Site: http://www.Geolytics.com
E-Mail: support@Geolytics.com
Sales & Customer Service: 1-800-577-6717
Fax: 1-732-651-2721
Mail: GeoLytics, Inc.
     PO Box 10
     East Brunswick, NJ 08816
Tech Support: support@GeoLytics.com
8. Addendum A. Weighting methodology

How have neighborhood characteristics of population and housing, changed in the United States over the last 40 years? What information is available to support this analysis, and at what level of geography? For most research, the traditional level of geography of analysis is the Census tract. And the information available is the official Census released data about the population and housing of the US, preferably the more detailed information released in the summary results of the Census long form.

Several problems present themselves with this analysis, first is in getting access to this data. Second is in Census tract splits and merges, which occur between censuses. Third is in the changes in questions asked, and therefore in the comparable information available. Fourth is in the coverage of the US for which tract data is available. Fifth is in suppressed data in 1970 and 1980, and in imputed data in 1990 and 2000. Sixth is in the data currently available for the 2000 census, which is redistricting and soon, the short form. The results of the 2000 long form will not be out until the second half of 2002. And seventh, might be the consideration of population undercounting and over counting.

Nothing can be done about many of these problems. As for the changes in questions asked in the census questionnaires, we just have to make the best of the summary of this information provided by the US census. Although much can be said about getting access to census release data, especially older census data, this problem was made easy by using the data collected by the Urban Institute for the Underclass database (UBD) in the early 90’s. As far as the tract coverage of the US goes, it was fully tracted in 1990 and 2000, but the Census Bureau only released tract data in urban areas in 1970 and 1980. In 1970 and 1980 census data was suppressed in 1970 and 1980 if the tract did not meet certain thresholds for certain universes, and in 1990 and 2000 it was imputed, or taken from like or nearby areas, again to protect confidentiality. So the remainder of this paper will discuss what we can do to normalize census tract data over time, so that we can look at historic (1970-1990) tract data.

In a standard census geography, in this case, 2000 tracts.

To explain the normalization of historic tract data to 2000 tracts, we start by converting 1990 tract data to 2000 tracts. This is made easy by the complete coverage of the US by 1990 and 2000 tracts, and the fact that the US was completely covered by Census Blocks, in both 1990 and 2000. 1980 tracts covered urban areas, and some rural areas in some states. 1980 tracts were weighted to 2000 tracts, via 1990 blocks. 1970 tracts covered most urban areas. 1970 tracts were also weighted to 2000 tracts, via 1990 blocks. First lets talk about relating and weighting 1990 blocks to 2000 blocks.

1990 to 2000 block relations were determined from Tiger/Line 2000, Type 1 and Type 3 records. 85% of the blocks had a 1:1 relationship, 10% had a 2:1, and 5% had a greater than 2:1. In some cases, hundredths of 2000 blocks related to a single 1990 block, and vice-versa. Block splits between 1990 and 2000 were weighted by an analysis of the 1990 streets. To split a block into parts, the sub-block areas were weighted according to the 1990 streets relating to each 2000 block part. The assumption is that local roads indicate where the population lived. 1990 streets were determined using Tiger/Line 1992. So using Tiger 92, and Tiger 2000, both a correspondence between 1990 and 2000 blocks was produced, as well as a weighting value to help split block demographics for those blocks that have been split or merged between 1990 and 2000. The file produced by this process is the 1990 to 2000 Block Weighting File (BWF).
The BWF is then applied to 1990 block demographics from the 1990 redistricting release PL94-171, and the STF1B, which contains additional housing block data. This allows us to produce a 1990 tract to 2000 tract weighting file, which allows for splits and merges in tracts, as well as a weighting according to universes. For example, a 1990 tract may be split into two equal 2000 tracts, but each half of the 1990 tract being split may have very different characteristics. One half may contain more renter occupied housing, or more of one racial group than the other half. So the fact that we have this information at the sub-tract level (block) is important, in more accurately weighting 1990 tract data to 2000 tract geographies. Because 1990 tracts can be weighted to 2000 tracts, at the block level, for 100% of the US, the weighting of 1990 tract data to 2000 tracts is very accurate.

1980 tracts are related to 1990 tracts/blocks, using the correspondence between these two found in the Census Bureau release of TIGER/Line 1992. This correspondence was found to be very good between 1980 tracts and 1990 blocks, but incomplete for 1980 block groups to 1990 blocks. This means that splitting a 1980 tract into 1990 tracts, had to be done spatially. Spatially means that we know exactly what 1990 blocks a 1980 tract is related to. A more accurate weighting between 1980 and 1990 could have been done if areas smaller than 1980 tracts were available, but because of a poor correspondence in TIGER/Line 1992 to 1980 block groups, they could not. Even if a perfect 1990 block to 1980 block group existed, we would have to consider the problem of suppressed 1980 data, which is more likely to occur in the less populated, and therefore less likely to meet thresholds.

1970 tracts were bridged to 1990 blocks using the 1970 to 1980 tract correspondence file produced by the US Census Bureau, and boundary files scanned from the Urban Atlas files in 1975. The tract correspondence produced good results, except where two or more 1970 tracts were merged into a single 1980 tract, in these cases, the spatial knowledge of a 1970 tract boundary, was taken from the Urban Atlas boundary. One problem with the Urban Atlas boundary is that it was digitized, and in some cases it slightly, but incorrectly crosses states and counties. A filter was applied to fix this problem, and the Urban Atlas boundaries were used in less than two percent of the 1970 to 1990 block relationships. The source of the Urban Atlas boundaries is documented to be a project between Harvard University and the US Census Bureau. Even though 1970 block groups exist, a spatial knowledge of them do not, and therefore a 1970 tract universe weighting using block groups, was not attempted.

Future improvements in the weighting, and comparison of 1970 and 1980 tracts, with 2000 tracts, are possible. Outside of tracted areas in 1970 and 1980, there are Census Enumeration Districts (ED), and Minor Civil Divisions (MCD), for which associations to 2000 tracts can be determined. Spatial knowledge of EDs and MCDs, are limited, but some information is available from the 1970 and 1980 Master Enumeration District Lists Extended (MEDListX). MCD census codes are not consistent over time, because town’s change, but most relationships can be determined from a simple MCD name matching. Inside tracted areas, intra-census relationships between block groups can be determined by simply relating their 12 digit FIPS block group number, and with this sub-tract information, better weighting of tract splits can be done.

Weighting 1990 block group data, instead of 1990 tract data, to 2000 tracts, may make improvements in the weighting, and comparison of 1990 data, with 2000 tracts. Tract data is also available between 1930 and 1960, and it may be possible, in some urban areas, to match up tract numbers, for longer time series analysis.