

MAPPING SOCIO-ECONOMIC VARIABLES USING 2000 CENSUS DATA

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
COASTAL SERVICES CENTER



NOAA Coastal Services Center
LINKING PEOPLE, INFORMATION, AND TECHNOLOGY

July 2009

NOAA Coastal Services Center

2234 S. Hobson Avenue

Charleston, SC 29424

(843) 740-1254

www.csc.noaa.gov

Regional Offices:

NOAA Pacific Services Center,
NOAA Gulf Coast Services Center, and
Offices in the Great Lakes, Mid-Atlantic,
Northeast, and West Coast

Contents

Introduction.....	1
Intended Users of This Method.....	1
Data Resources and Acquisition	1
Data Requirements.....	1
Other Data Sources and “Data-Ready” Products.....	3
Applications and Uses of Socio-Economic Census Data	3
Downloading and Processing Census Data.....	4
Step 1: Download Data from the U.S. Census Bureau Website.....	5
Step 2: Import the Data into a Usable Database.....	6
Step 3: Setup a SQL Server Database for the Data Set	14
Step 4: Importing Census Data into ArcGIS	14
Appendix: Sample of Calculations – Example from Excel Table	17

Introduction

The purpose of this document is to instruct users on how to create social and economic spatial data using U.S. Census Bureau data. Step-by-step instructions are provided to guide the user through the process of data access and acquisition, data processing, spatial data creation, and data application, with examples. This method illustrates how the National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center processed the census data and created spatial data layers. For questions or additional information on this process, contact csc@csc.noaa.gov.

Intended Users of This Method

This process was designed for geographic information system (GIS) analysts and requires a working knowledge of ArcGIS, MS Access, SQL Server, and data symbology.

If you are not a GIS analyst but would like access to census data, several resources are available to obtain preprocessed versions of census data to use in an analysis. These resources are located in the following section.

Data Resources and Acquisition

Data Requirements

<p>Census 2000 Data – www.census.gov</p> <ul style="list-style-type: none">○ The U.S. Census Bureau provides demographic data for tracts, block groups, and block cartographic boundaries. This document instructs users on how to create spatial data layers (shapefiles) from the block group level census data. The Census Bureau has a number of products available to assist users in identifying and mapping the census geographic entities. Most of these map products are available for viewing on the website. Visit the Census Bureau’s site (www.census.gov) for more information.○ Benefits – Users can choose specific demographic variables relevant to specific issues and analyses.○ Drawbacks – The data import process is time consuming and requires a high level of database and GIS knowledge. The user needs to have a clear grasp of the many variables captured by the Census Bureau and the ability to properly analyze and compare the data.	<p>Cartographic Boundary Definitions</p> <p>Tract – relatively permanent statistical subdivisions within a county that provide a stable set of geographic units for the presentation of decennial census data.</p> <p>Block group – A cluster of census blocks having the same first digit of their four-digit identifying numbers within a census tract.</p>
---	--

- **Updates** – Census 2010 data are being collected. Check the census site for more information and updates.

About Census 2000 Data

- www.census.gov/main/www/cen2000.html
- www.census.gov/mso/www/c2000basics/00Basics.pdf
- www.census.gov/prod/2001pubs/mso-01icdp.pdf

The figure to the right provides a brief explanation of the types of information the U.S. Census Bureau collects in its short and long forms.



Block – Areas bounded on all sides by visible features, such as streets, streams, and railroad tracks, and by invisible boundaries, such as city, town, township, and county limits, property lines, and short, imaginary extensions of streets and roads. Generally, census blocks are small in area; for example, a block bounded by city streets. However, census blocks in remote areas may be large and irregular and contain many square miles.

Source: U.S. Census Bureau

Other Data Sources and “Data-Ready” Products

Several data resources and products that can provide several socio-economic variables with minimal or no GIS processing are available at the NOAA Special Projects Office Social Trends in Coastal Socioeconomics (STICS) website – <http://marineeconomics.noaa.gov/socioeconomics/>. Below are descriptions of the products.

- **Profiles** – Users can create a profile based on a selected study area offering different scales (e.g., state, county, NOAA watershed). Users can select a data source (Census, Bureau of Economic Analysis), select socio-economic variables, and compare the study area to other census years (census data are collected every 10 years). The user can then choose to create a report, a map, or a custom graphics report. These reports can be exported to PDF or Excel. This option enables users to view many of the socio-economic variables in several formats.
- **Data Query**
Provides Time Series Demographic Statistics (TSDS) (1970, 1980, 1990, and 2000). Enables users to query, download, and view census-based socio-economic data.
- **Census 2000 mapping** (GIS data source/shapefiles)
This interactive mapping tool enables users to view Census 2000 data for different geographic scales (county, state, blockgroup, etc.), export shapefiles, view a data dictionary and metadata, and create a printable map.
- **Assessments** – The assessments provide examples of how socio-economic data and the STICS data and tools can be applied and used. Analysis of trends in coastal development and demographic makeup can provide useful insights into the pattern and magnitude of pressures that challenge coastal and ocean resources.
- **Coastal County Facts** – Provides summaries and quick facts of time series population, housing, employment, and income data for the U.S., coastal zone counties, and coastal watershed counties.

Applications and Uses of Socio-Economic Census Data

Census data have many uses, ranging from determining what types of populations are vulnerable to coastal hazards to identifying the socio-economic drivers for land cover change (e.g., population density and land cover types). The examples below focus on using census data to identify locations of potential vulnerability for targeting effective adaptation strategies.

Using Census Data to Identify Socio-Economic Vulnerabilities

Certain demographic variables (e.g., age, income, gender) indicate social vulnerability to hazard risks. Vulnerability is a function of exposure (who or what is at risk) and the sensitivity of the exposed system (the degree to which people, places, and natural resources can be harmed) (Cutter et. al, 2008). Demographic variables such as gender, race, age, language, and income are indicative of populations that may be at greater risk from disturbances such as hurricanes.

The NOAA Coastal Services Center has developed an adaptable, multi-step process that can help coastal communities better understand and adapt to coastal hazard risks and vulnerabilities. As part of this process, the Center uses census-based socio-economic variables to identify areas within a community that have higher concentrations of vulnerable populations. For example, minority populations may represent areas where special cultural considerations or foreign language interpreters could be needed. Single-parent households may indicate areas where special child-care considerations could be necessary. Elderly populations and lack of vehicles may indicate special mobility needs, while low educational attainment rates may indicate the need for specialized help in dealing with disaster assistance procedures. Poverty and public-assistance income may indicate areas where even moderate damages could have significant financial impacts on residents.

The “Appendix” provides detailed tables of census-derived socio-economic variables that can be used in a risk and vulnerability assessment. The field labeled “Field Formula (from SF3 Census Tables)” shows how different census variables were combined to create custom variables. While these variables are indicative of socio-economic vulnerability, each community has unique social and economic characteristics and is encouraged to identify these as part of their assessment.

Using Conservation Strategies to Protect Communities

The Nature Conservancy (TNC), the Association of State Floodplain Managers, and the NOAA Coastal Services Center used census-derived socio-economic data in Suffolk County (part of Long Island), New York, to identify populations and housing resources that may be vulnerable to storm surge and sea level rise. These data were used in the Federal Emergency Management Agency’s HAZUS model to calculate the potential economic losses. The socio-economic information is helping the local towns and TNC to develop conservation strategies that will protect both the natural and built environments. View the tool at www.coastalresilience.org.

Downloading and Processing Census Data

This section outlines the process the Center used to download and use U.S. census data at the Block Group Summary File 3 (SF3) level. The method documented here focuses on the specific tasks of retrieving and preparing the data for use in a GIS and does not focus on the specifics of the fields used for analysis. Thus, the method is applicable to vulnerability assessments beyond coastal hazard risks.

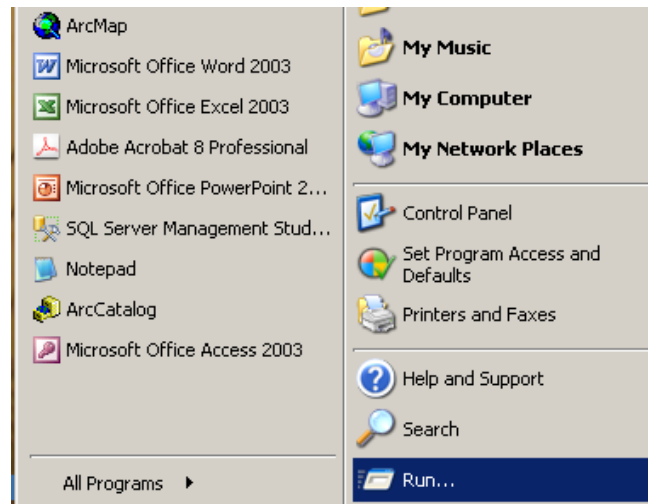
Step 1: Download Data from the U.S. Census Bureau Website

U.S. census data at the Block Group Summary File (SF3) level

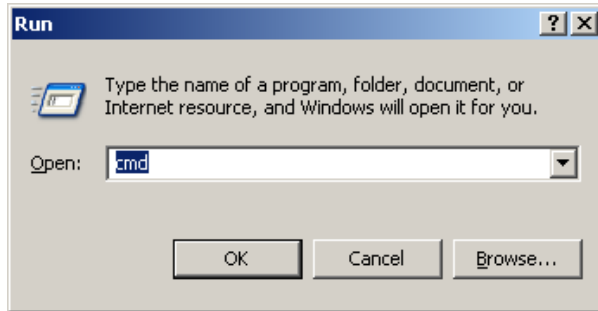
SF3 consists of 813 detailed tables of Census 2000 social, economic, and housing characteristics compiled from a sample of approximately 19 million housing units (about 1 in 6 households) that received the Census 2000 long-form questionnaire. Additional information on census data and SF3 tables can be accessed on the Census Bureau's website (www.census.gov/Press-Release/www/2002/sumfile3.html).

The U.S. Census Bureau website contains many ways to download data. For the purposes of this process, the Center downloaded the 2000 SF3 data for the entire state of South Carolina. This approach allows the user to build custom metrics for analysis.

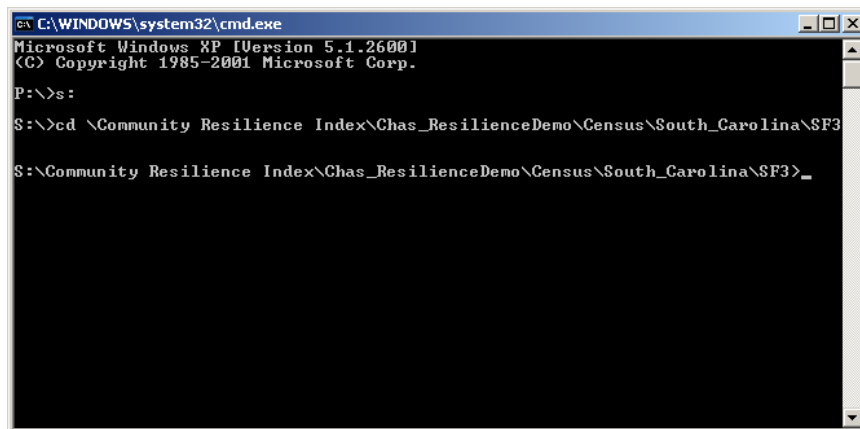
1. Download the SF3 (Block Group level) data from the census FTP site at ftp://www2.census.gov/census_2000/datasets/Summary_File_3/ and navigate to the state of interest. Download the zip file titled all_*statename*.zip.
2. Once the data are downloaded, use the WinZip compression tool to extract the files to a single folder where the raw tables are contained. The folder will contain a series of files with the .uf3 extension.
3. Use a DOS command to rename all the .uf3 files to a .txt extension and remove any other .txt extension files from the folder.
 - a. Click **Start > Run**.



- b. Type **cmd** in the Run Command box and click **OK**.



- c. At the command prompt, change the drive and directory, using the **cd ** command, to the directory where the data were stored.



- d. Use the **rename** command to rename all the files with the .uf3 extension to a .txt extension. The syntax is as follows:

rename *.uf3 *.txt

Step 2: Import the Data into a Usable Database

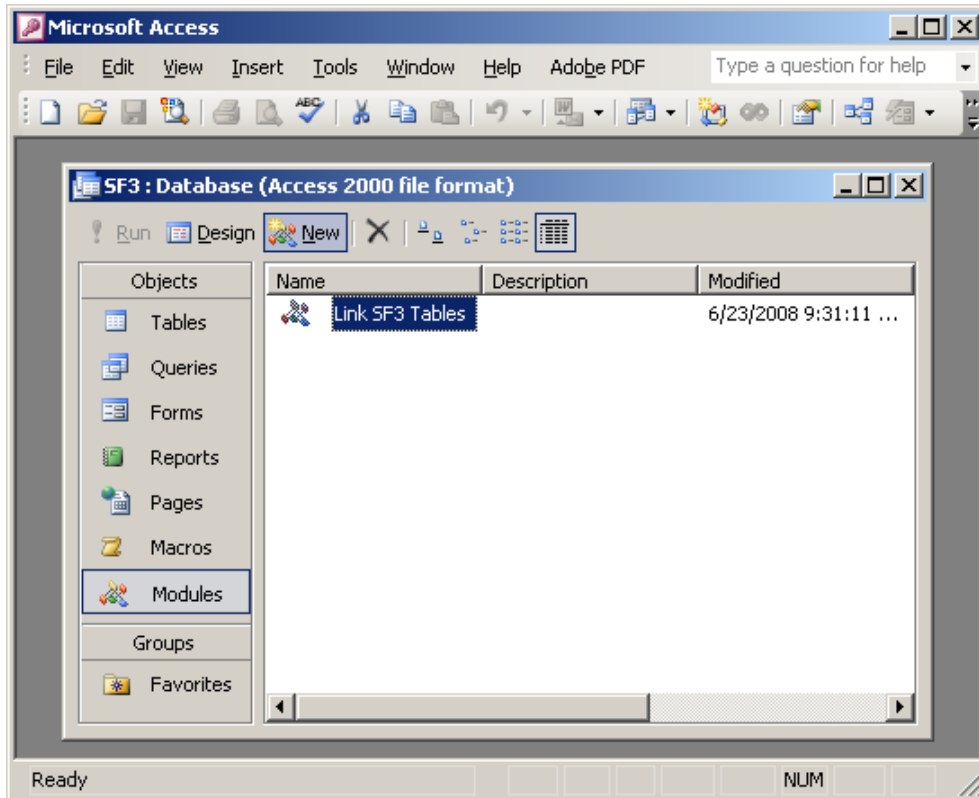
Directions for importing the Summary File 3 ASCII text files into a usable data structure can be found at www.census.gov/support/SF3ASCII.html. **Note that these directions work in Microsoft Access 2000.**

The Center automated the process to import the data structure into Access and then to export those data to the SQL Server.

1. Download the Summary File 3 Template file for Access from the census website (www.census.gov/support/2000/SF3/Acc2000.zip). Unzip the contents of this zip file to the

same folder where the raw state SF3 data files are located.

2. Open the Access template file and, if asked, convert to the most recent version.
3. To insert the automated table-linking code into the Access database, navigate to the **Modules** tab in Access and create a **New** module.



4. Replace the text in the module with the following:

Option Compare Database

```
Public Function ListFiles(strPath As String, Optional strFileSpec As String, _  
    Optional blIncludeSubfolders As Boolean)
```

```
On Error GoTo Err_Handler
```

```
'Purpose: Link SF3 tables from Census download.
```

```
'Arguments: strPath = the path to search.
```

```
'    strFileSpec = "*.*)" unless you specify differently.
```

```
'    blIncludeSubfolders: If True, returns results from subdirectories of strPath as  
well.
```

```
'Method: FilDir() adds items to a collection, calling itself recursively for subfolders.
```

```

Dim colDirList As New Collection
Dim varItem As Variant
Dim mySplit1, mySplit2, myCount
Dim mySpecName, myTableName As String

Call FillDir(colDirList, strPath, strFileSpec, bIncludeSubfolders)

'Link the Tables.
Hourglass = True
myCount = 0
For Each varItem In colDirList
    mySplit1 = Split(varItem, "\")
    mySplit2 = Split(mySplit1(UBound(mySplit1)), ".")
    If Right(mySplit2(0), 3) = "geo" Then
        mySpecName = "SF3GEO Import Specification"
    Else
        mySpecName = "SF30" & Right(mySplit2(0), 3) & " Import Specification"
    End If
    myTableName = mySplit2(0)
    DoCmd.TransferText acLinkDelim, mySpecName, myTableName, varItem
    myCount = myCount + 1
Next

Hourglass = False
MsgBox myCount & " Tables Linked"

Exit_Handler:
Exit Function

Err_Handler:
MsgBox "Error " & Err.Number & ": " & Err.Description
Resume Exit_Handler
End Function

Private Function FillDir(colDirList As Collection, ByVal strFolder As String,
strFileSpec As String, _
bIncludeSubfolders As Boolean)
'Build up a list of files, and then add to this list, any additional folders
Dim strTemp As String
Dim colFolders As New Collection
Dim vFolderName As Variant

'Add the files to the folder.
strFolder = TrailingSlash(strFolder)

```

```

strTemp = Dir(strFolder & strFileSpec)
Do While strTemp <> vbNullString
    colDirList.Add strFolder & strTemp
    strTemp = Dir
Loop

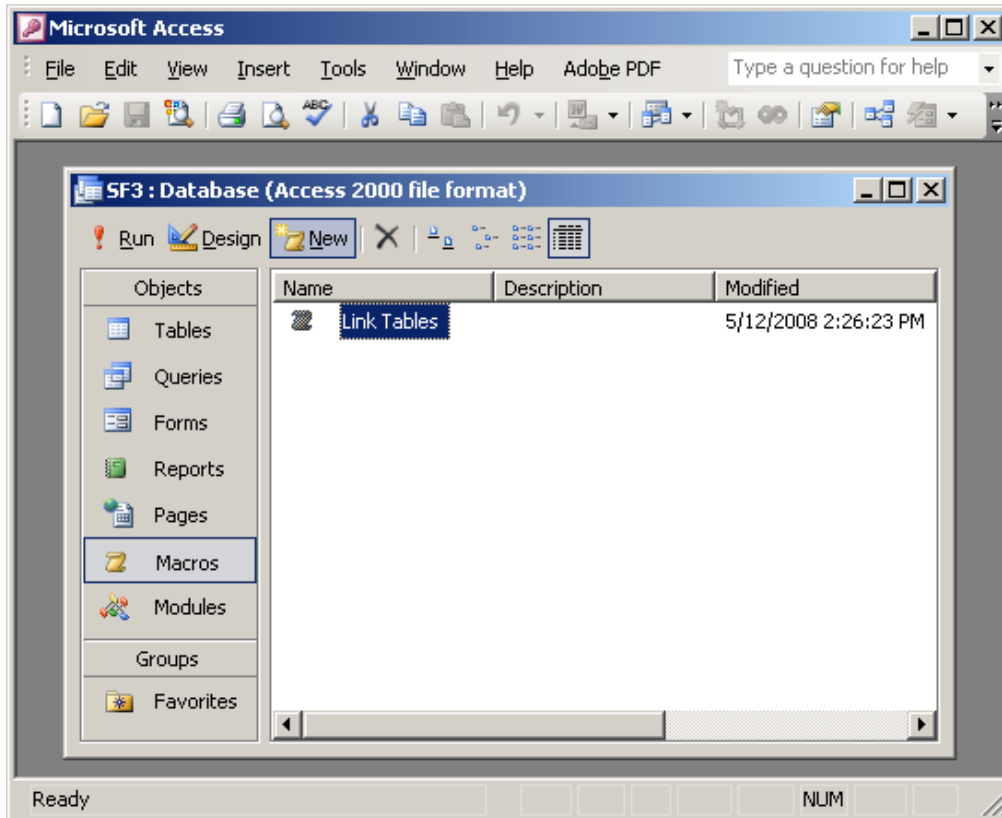
If bIncludeSubfolders Then
    'Build collection of additional subfolders.
    strTemp = Dir(strFolder, vbDirectory)
    Do While strTemp <> vbNullString
        If (strTemp <> ".") And (strTemp <> "..") Then
            If (GetAttr(strFolder & strTemp) And vbDirectory) <> 0& Then
                colFolders.Add strTemp
            End If
        End If
        strTemp = Dir
    Loop
    'Call function recursively for each subfolder.
    For Each vFolderName In colFolders
        Call FillDir(colDirList, strFolder & TrailingSlash(vFolderName), strFileSpec, True)
    Next vFolderName
End If
End Function

Public Function TrailingSlash(varIn As Variant) As String
    If Len(varIn) > 0& Then
        If Right(varIn, 1&) = "\" Then
            TrailingSlash = varIn
        Else
            TrailingSlash = varIn & "\"
        End If
    End If
End Function

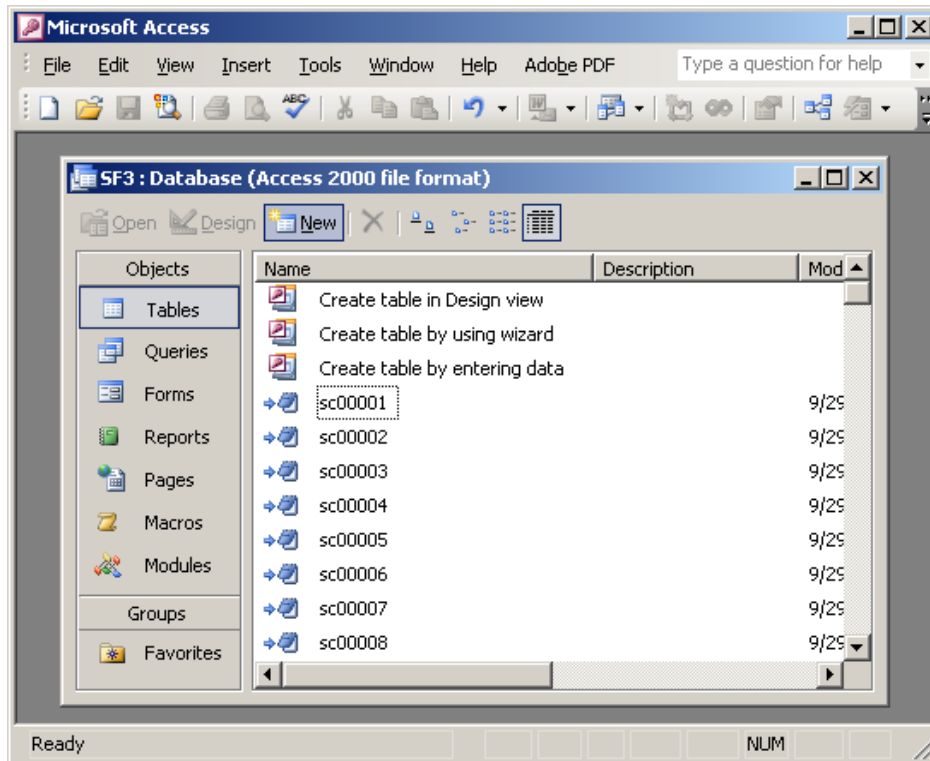
```

Save the module with the name **Link SF3 Tables**.

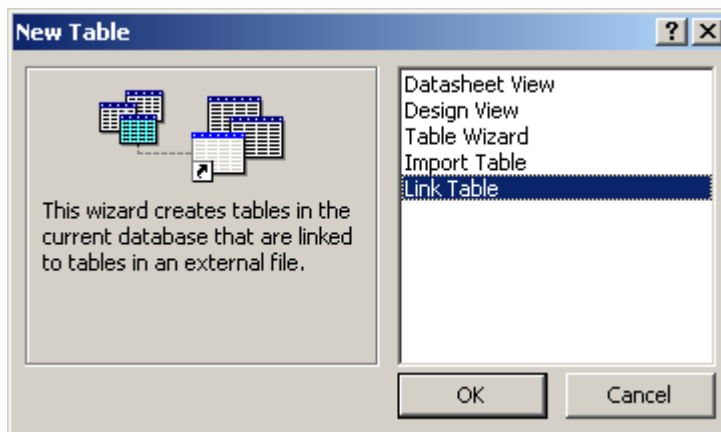
5. Switch to the **Macros** tab and create a **New** macro that will make it easier to execute the code above.



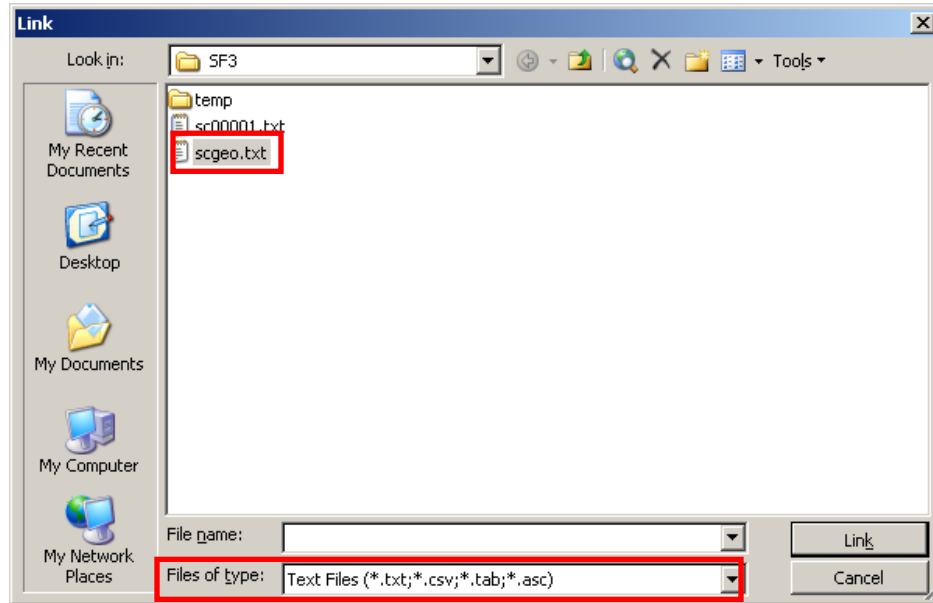
- a. The Macro needs to contain one action that is a **Run Code** action with the following text in the Function Name: *ListFiles ("directory_path_where_Census_Data_is_stored", "state_initials(e.g sc)*.txt")*.
6. Check the Geo table
The GEO table is often not formatted successfully. Before moving on to the next phase, take a moment to check the GEO Table (e.g., **scgeo**). If most of the columns are empty, delete the linked table and re-import it manually as demonstrated below:
 - a. Change to the **Tables** tab and create a new table by clicking on **New**.



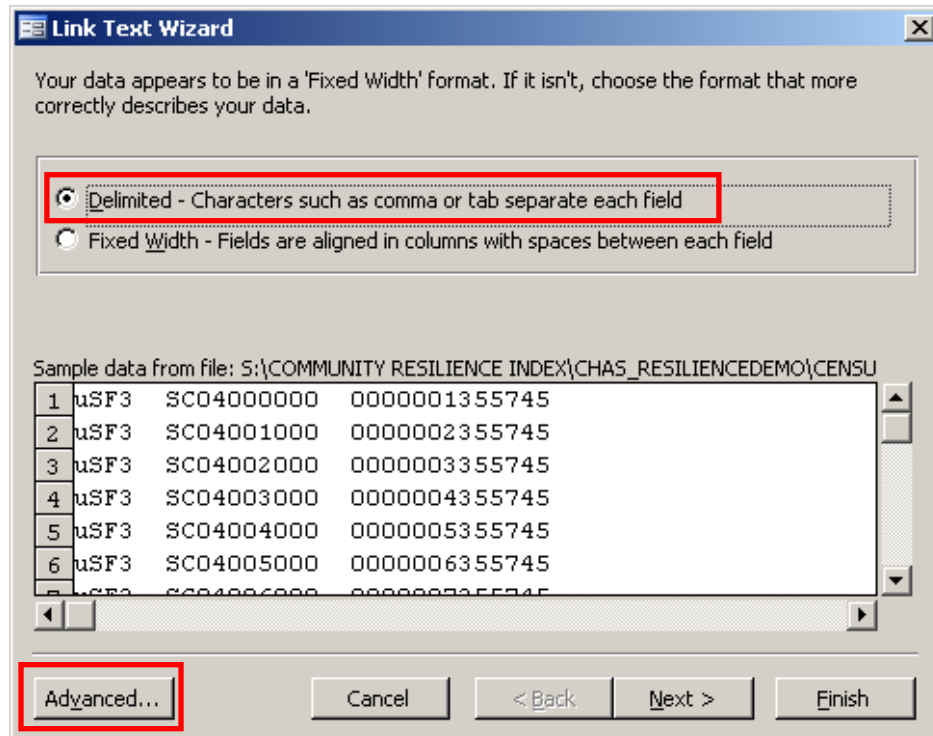
b. Choose the **Link Table** option



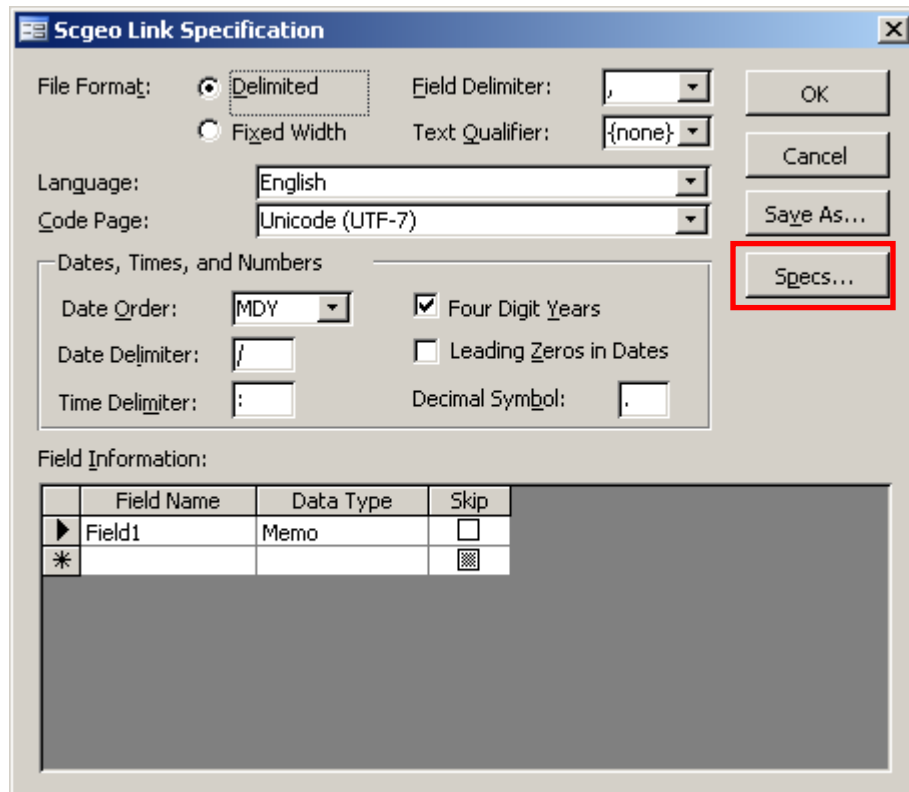
- c. Select the **xxgeo.txt** file (where “xx” represents the state initials). Note that you may need to change the **Files of type** at the bottom to **Text Files** to see the files.



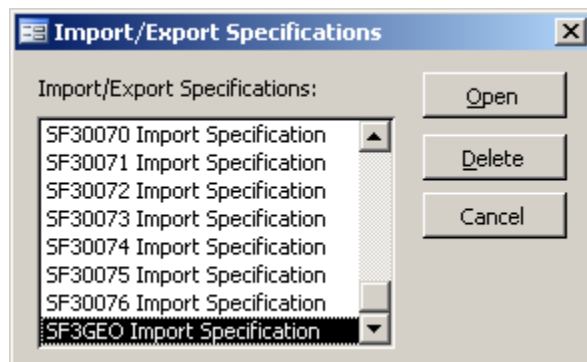
- d. Change the format to **Delimited** and click the **Advanced** button.



- e. In the “Advanced” window, click on the **Specs...** button.



- f. Select **SF3GEO Import Specification** and click **Open**.



- g. Click **OK**, and then **Finish** to complete the import. The table should look more complete at this time.

Step 3: Setup a SQL Server Database for the Data Set

For the remaining steps, a SQL Server database was used to import the data tables and create custom views (queries) of the data tables as needed.

If the user does not have access to a full SQL Server, SQL Server Express can be downloaded from Microsoft. The user must have full access rights to create and edit tables.

Upsize Wizard

1. Once the SQL database is created, use the **Upsizing Wizard Tool** in Access to move the linked census tables created through the process above into the SQL Server. The **Upsizing Wizard** is found under the **Tools > Database Utilities** menus.

Note: the data transfer can be very time consuming since the text files are quite large and include information for entire states.

2. Once the data transfer is complete, the user can use **SQL Server Management Studio** to create views (queries) of custom data sets from the many tables in the SF3 data set.

Create New "View"

1. Once the data have been uploaded to the SQL server, the user can make views of the data as needed. The "Appendix" gives a sense of the fields that may be included in one of these views. To make a view, the user must link the different tables using the linking field **LOCREC**.
2. To link the database view to the spatial data in the next step, the output table needs to have a geo-location field, something that connects it to a spatial record. The most common such field for census-related data sets is the FIPS code. See the first field in the "Appendix" for an example of how to create this field in your view.

Step 4: Importing Census Data into ArcGIS

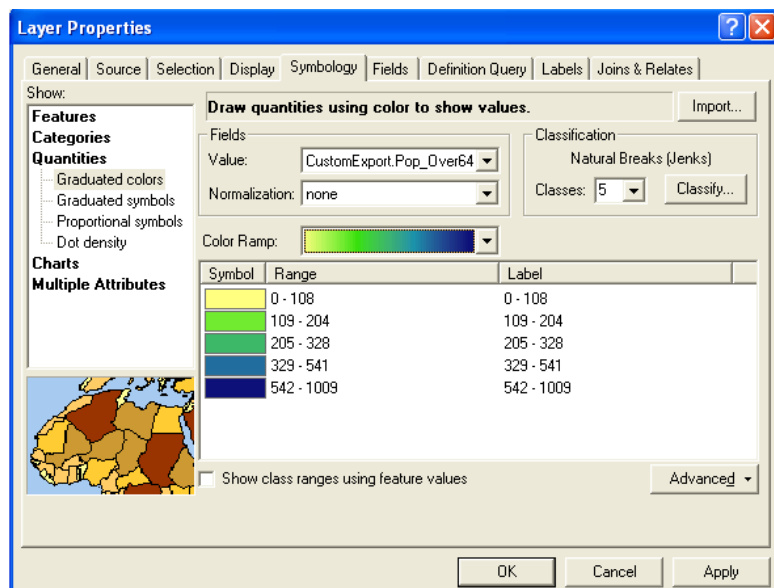
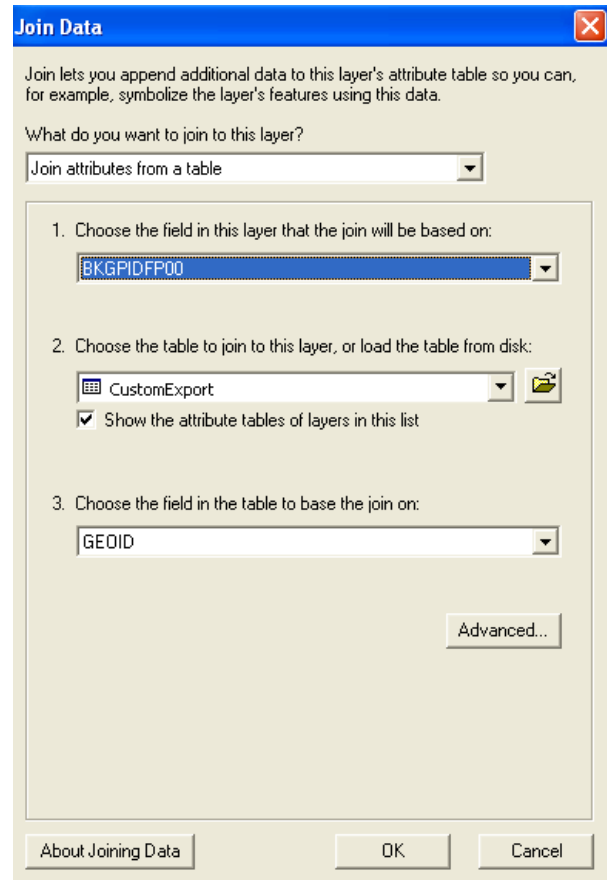
ESRI ArcGIS 9.2 was used to create spatial data from the processed census data.

To add the SF3 data set to ArcMap, you will need to export the database table as a .dbf file.

1. Open ArcMap.
2. Add your .dbf file to ArcMap.
3. Add the shapefile that contains geographic information for your study area (download these shapefiles at www.census.gov/geo/www/cob/bdy_files.html). This shapefile will be used in the Join process.

Joining Attributes in One Table to Another

- In ArcMap's table of contents, right-click the layer or table you want to join, point to Joins and Relates, and click **Join**. The Join Data dialog box will open.
- Click the first drop-down arrow and click **Join attributes from a table**.
- Click the second drop-down arrow and click the field name in the layer on which the join will be based. Choose **BKGPIDFP00** or the name of the field that corresponds to the **FIPS** code.
- Click the third drop-down arrow to choose the table to join to the layer. Choose the **.dbf table** that contains the census data.
- Click the fourth drop-down arrow and click the field in the table on which to base the join. Choose the field that corresponds to the **FIPS** code from the database table you created (in this example, called **GEOID**)
- Click **OK**.
- The attributes of the table are appended to the layer's attribute table. You now have spatial representation of the census data for your study area.
- The next step is to choose the symbology most appropriate for your analysis.



References

Cutter . 2008. A Framework for Measuring Coastal Hazard Resilience in New Jersey Communities. White Paper for the Urban Coast Institute.

Cutter, S.L., J.T. Mitchell, and M.S. Scott, 2000. Revealing the vulnerability of people and places: A case study of Georgetown County, South Carolina," *Annals of the AAG* 90 (4): 713-737.

Heinz Center, 2002. *Human Links to Coastal Disasters*. Washington D.C.: The H. John Heinz III Center for Science, Economics and the Environment.

State of South Carolina, 2005. *State of South Carolina Hazard Assessment*. Columbia, SC: Emergency Management Division. Available online:

http://www.scecmd.org/Library/SCEMD_Hazards_Assessment_Final_Report_2005-1.pdf

Appendix: Sample of Calculations – Example from Excel Table

To link different census tables, the unifying field is called LOCREC, which is found in every table.

Alias (Name)	Description	Field Formula (from SF3 Census Tables)	Table	Criteria
FIPS	FIPS Code for mapping Block Group	STATE + COUNTY + TRACT + BLKGRP	__geo	
Population	Population	P001001	__00001	
Households	Number of Households	P010001	__00001	
HousingUnits	Number of Housing Units	H001001	__00056	
Pop_income_to_poverty_ratio_p5_to_p74	Population within poverty ratio of .5 - .74	P088003	__00007	
Pop_income_to_poverty_ratio_p75_to_p99	Population within poverty ratio of .75 - .99	P088004	__00007	
Pop_income_to_poverty_ratio_1_to_1p24	Population within poverty ratio of 1 - 1.24	P088005	__00007	
Pop_income_to_poverty_ratio_1p25_to_1p49	Population within poverty ratio of 1.25 - 1.49	P088006	__00007	
Pop_income_to_poverty_ratio_1p5_to_1p74	Population within poverty ratio of 1.5 - 1.74	P088007	__00007	
Pop_income_to_poverty_ratio_1p75_to_1p84	Population within poverty ratio of 1.75 - 1.84	P088008	__00007	
Pop_income_to_poverty_ratio_1p85_to_1p99	Population within poverty ratio of 1.85 - 1.99	P088009	__00007	
Pop_income_to_poverty_ratio_2_and_over	Population within poverty ratio over 2.0	P088010	__00007	
H0H_PublicAssistance	Households on Public Assistance	P064002	__00006	
Pop_Under5	Population under age 5	P001001 - P019001		
Pop_Over64	Population over age 64	P011001	__00001	
TotalPop_disabled_5_to_64	Disabled Population ages 5 - 64	P042004 + P042007 + P042014 + P042028 + P042031 + P042038		
TotalPop_disabled_Over64	Disabled Population over 64	P042021 + P042024 + P042045 + P042048		
HH_SingleMothers	Single Mother Households	P010015	__00001	
Pop_Adult_NoDiploma	Adult Population with no High School Diploma	P037001 - P037011 - P037012 - P037013 - P037014 - P037015 - P037016 - P037017 - P037018 - P037028 - P037029 - P037030 -		

		P037031 - P037032 - P037033 - P037034 - P037035		
Pop_CollegeDegree	Adult Population with a College Degree	P037014 + P037015 + P037016 + P037017 + P037018 + P037031 + P037032 + P037033 + P037034 + P037035		
HU_NoVehicle	Households with no Vehicle	H044003 + H044010		
Wrks_Uses_Public_Transportation	Workers who use Public Transportation	P030005	__00003	
Median_Hh_Income_1999	Median Household income in 1999 dollars	P053001	__00006	
HH_income_Less_than_10k	Households with income less than \$10k	P052002	__00006	
HH_income_10k_to_15k	Households with income \$10k to \$15k	P025003	__00003	
HH_income_15k_to_20k	Households with income \$15k to \$20k	P052004	__00006	
HH_income_20k_to_25k	Households with income \$20k to \$25k	P052005	__00006	
HH_income_25k_to_30k	Households with income \$25k to \$30k	P052006	__00006	
HH_income_30k_to_35k	Households with income \$30k to \$35k	P052007	__00006	
HH_income_35k_to_40k	Households with income \$35k to \$40k	P052008	__00006	
HH_income_40k_to_45k	Households with income \$40k to \$45k	P052009	__00006	
HH_income_45k_to_50k	Households with income \$45k to \$50k	P052010	__00006	
HH_income_50k_to_60k	Households with income \$50k to \$60k	P052011	__00006	
HH_income_60k_to_75k	Households with income \$60k to \$75k	P052012	__00006	
HH_income_75k_to_100k	Households with income \$75k to \$100k	P052013	__00006	
HH_income_100k_to_125k	Households with income \$100k to \$125k	P052014	__00006	
HH_income_125k_to_150k	Households with income \$125k to \$150k	P052015	__00006	
HH_income_150k_to_200k	Households with income \$150k to \$200k	P052016	__00006	
Income_more_than_200k	Households with income over \$200k	P052017	__00006	
NonWhite_Pop	Non-White Population	P006001 - P006002		
Latino_Pop	Latino Population	P007010	__00001	
Spanish_Speaker_Pop	Spanish speaking population	P019004 + P019026 + P019046		
Language_Barrier_Spanish	Spanish speaking population that has difficulty understanding english	P019007 + P019008 + P019029 + P019030 + P019051 + P019052		
Lang_Other_than_English_or_Spanish_Pop	Population with primary language other	P019009 + P019014 + P019019 + P019031		

	than English or Spanish	+ P019036 + P019041 + P019053 + P019058 + P019063		
Lang_Barrier_other_than_Spanish	Population with primary language other than English or Spanish that has difficulty understanding English	P019012 + P019013 + P019017 + P019018 + P019022 + P019023 + P019034 + P019035 + P019039 + P019040 + P019044 + P019045 + P019056 + P019057 + P019061 + P019062 + P019066 + P019067		
HU_Seasonal_or_vacant	Number of Seasonal or Vacant Housing Units	H008005 + H008006 + H008007		
HURenterOccupied	Number of Renter Occupied Housing Units	H007003	__00056	
HUOwnerOccupied	Number of Owner Occupied Housing Units	H007002	__00056	
Hu_Built_Before1950	Housing Unit built before 1950	H034009 + H034010		
HU_Built_1950s	Housing Unit built during the 1950s	H034008	__00058	
HU_Built_1960s	Housing Unit built during the 1960s	H034007	__00058	
HU_Built_1970s	Housing Unit built during the 1970s	H034006	__00058	
HU_Built_1980s	Housing Unit built during the 1980s	H034005	__00058	
HU_Built_1990s	Housing Unit built during the 1990s	H034004 + H034003 + H034002		
HU_Media__earBuilt	Median Year Built for Housing Units	H035001	__00058	
HU_Mobile_Homes	Number of Mobile Home Housing Units	H030010	__00058	
HU_MultiFamily	Number of Multi Family Housing Units	H030003 + H030004 + H030005 + H030006 + H030007 + H030008 + H030009		
<i>blank</i>	Block Group (required)	BLKGRP	__geo	IS NOT NULL
<i>blank</i>	Census Summary level (150 = Block Group)	SUMLEV	__geo	= 150
<i>Blank</i>	County (where "XXX" equals the county code)	COUNTY	__geo	= N'XXX'