5-1-2009

Applying Census Data for Small Area Estimation in Community and Social Service Planning

Michael Wolf-Branigin
George Mason University, mwolfbra@gmu.edu

Hyon-Sook Suh
George Mason University, hsuh1@gmu.edu

Star Muir
George Mason University, smuir@gmu.edu

Emily S. Ihara
George Mason University, eihara@gmu.edu

Applying Census Data for Small Area Estimation in Community and Social Service Planning

Michael Wolf-Branigin   Hyon-Sook Suh   Star Muir   Emily S. Ihara
George Mason University

Small area estimation provides a tool for community analysis. A procedure for accessing, selecting, joining and analyzing US Census data is provided. Skills acquired while completing the procedure include accessing census data, downloading boundary files and displaying themes. Such skills are valuable tools for students to possess as they enter the workforce.

Key words: Small area estimation, US Census, social policy analysis, geographic information systems (GIS), systems evaluation.

Introduction

Social services planners and evaluators have developed their planning abilities through improved access to data, methods for analyzing data, and techniques for visualizing the results. This article presents a method for data application. Small area estimation involves using outcome data that joins to a set of predictor variables within small domains or geographic areas in order to generate estimates.

Although small area estimates and geographic information systems (GIS) have been used by human service workers for years (Wolf-Branigin, LeRoy & Miller, 2001), the study of human environment interactions often fail to consider individual-level information or cross-discipline data, resulting in a lack of explanatory

Michael Wolf-Branigin is an Associate Professor in the School of Social Work. Email: mwolfbra@gmu.edu. Hyon-Sook Suh is the Head of Goverment Documents and Maps, and the Geography Liaison Libriarian in the DP&S, Library. Email: hsuh1@gmu.edu. Star Muir is an Associate Professor and the Hiring and Scheduling Director in the Department of Communication. Email: smuir@gmu.edu. Emily S. Ihara is an Assistant Professor in the School of Social Work. Email: eihara@gmu.edu.

and predictive power (An, Linderman, Qi, Shortridge & Liu, 2005). Calls from the field advocate for governmental and non-governmental organizations to improve the collecting, linking, and sharing of microdata in order to improve decision-making (Weitzman, Silver & Brazill, 2006).

For more than a century, the United States Census has collected data based on census tracts in selected areas (Krieger, 2006). These tracts, which typically include approximately 4,600 individuals, provide the basis for conducting small area estimations. Census tracts were defined for the entire country for the first time in 2000. This allows for better access to data needed for social policy and planning and to for documenting need and making informed decisions about the allocation of resources in various communities (Krieger, 2006). In addition to an increase in the scientific use of census tract data, the American FactFinder function of the US Census has increased consumer access to this vast database. The improvement in coverage of census tracts as well as easier access to this data allows for more precise service planning.

Given these improvements, social workers have the potential to be at the forefront of policy decisions by including GIS and mapping skills in their toolboxes. These skills allow health and social service workers to strengthen the social survey tradition, identify

community needs and resources, improve the delivery of services, and empower communities and disenfranchised groups (Hoefer, Hoefer & Tobias, 1994; Robinson & Wier, 1998: Wier & Robinson, 1998; Hillier, 2007).

Introducing students to GIS methodologies can further enhance their ability to visualize and solve complex issues (Watkins, 2001). Potential issues to explore cover a vast range, for example, child maltreatment patterns based on neighborhoods (Ernst, 2000; Frieshler, Lery, Gruenewald & Chow, 2006), housing patterns of persons with disabilities (Wolf-Branigin, 2002), or adult addiction epidemiology (Grant, Martinez & White, 1998; Gerwe, 2000; Maxwell, 2000). Empirical research has increasingly shown the significance of context in social problems, thereby pointing to a need for better understanding determinants such as place and time or community and neighborhood effects on outcomes for various populations.

Procedure

The following six steps provide access to software and data for creating and displaying small area estimates. These steps are: (1) downloading census data, (2) using MS Excel for descriptive analysis, (3) creating a DBF file from an Excel file, (4) downloading GIS boundary files, (5) importing Census data to a GIS environment, and (6) displaying themes.

The required software to complete these steps include: ArcView (GIS) software, Excel, Access, Internet Explorer (not Netscape) and Winzip (or any unzipping software). This method, designed for a course assignment entitled, Community Analysis Using GIS Technology, takes about two hours to complete.

Step 1: Downloading Census Data

Extract census variables such as low-income families (poverty), older people with low income, or other demographic variables (e.g., language speaking, Hispanic population) from an online database, called American Factfinder, and download them in MS Excel format. It is helpful to first create and name a folder on the C: drive in which to store the data before starting this exercise.

Census information needed is by census tract level in a county in northern Virginia.

- Visit the American Factfinder Website (http://www.factfinder.census.gov).
- Click on Data Sets (Decennial Census) in the left hand side menu of the website.
- In the 2000 section, select Census 2000 Summary File 3 (SF3)-Sample Data and click on the Detailed Tables on the right hand side.
- In the Select Geography type section, scroll down to Census Tracts in the drop down menu. In the state section, use the drop down menu to select Virginia; in the county section, use the drop down menu and select Fairfax County.

In the Select One or More Geographic Areas, select all census tracts and click on Add, then click Next.

- Select Census Variables that for analysis and mapping. For example, P87 for poverty status in 1999 by age number of 65 years and over below the poverty level, P77 for median family income, P1 for total population, etc., and click on Add.
- Click on Show Result to view a summary of the table selected.
- In the Print/Download option in the top menu, click Download. This will provide various format options for downloading data.

- Uncheck the Include Descriptive Data Element names box. Select MS Excel (.xls) and click OK. Next, save the zipped file to the folder created on the C: drive.
- Unzip the downloaded census file using Winzip software (or any other unzipping software) and save it in the created folder in C: as a file named census.

The census data that is downloaded gives two Excel sheets with information on the total population and median value with the names ending with geo and data1. For example, when the poverty by age data set is downloaded, it provides two excel sheets; one would be named dt_dec_2000_sf3_u_geo and the other would be named as dt_dec_2000_sf3_u_data1. The file data1 contains the population information, while the geo information contains the county geographic information such as census tracts, etc.

Step 2: Descriptive Analysis in Excel.
Open dt_dec_2000_sf3_u_data1 in Excel; this gives 18 columns starting with P087001. To make this exercise easier, change the name of these columns based on the description of the detailed table downloaded earlier (e.g., P087001 as total, etc.). In order to analyze the distribution of older people below the poverty level in Fairfax County, choose P087008 and P087008 and add these two columns for the total or sum to a new column (create a new column called, 65+) and add the last two columns of data (65+ and 75+) to this new column). To complete this, highlight the

columns to be summed (65+ and 76+) along with the newly labeled column. Click on the sigma sign (∑) on the Excel toolbar as shown below while holding shift key.
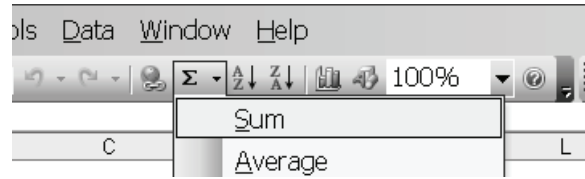
This will add the two columns and give the total number of persons who are 65 years older in poverty status in each census tract.

If not already activated in Excel, activate the Analysis ToolPak in the Tools menu under Add-ins. Label one additional column to the right. Highlight the columns for analysis along with the new column and click on Tools and Data Analysis. Check Descriptive Statistics and then highlight the column to be analyzed and click on Summary Statistics. This will create a table with several summary statistics (e.g., mean, median, and mode).

Step 3: Create a DBF file from an Excel file.
You must have MS Access to continue with this exercise. Before starting this step, check to see that the columns containing population information in the downloaded Excel sheets. This Excel file must be imported to the MS Access database in order to keep the Census tract number column, to relationally join the table to another table in GIS software such as ArcView, and to save the file to a dbf format.

Importing Excel files into Access
- To import the Excel files into Access, open the Access database.
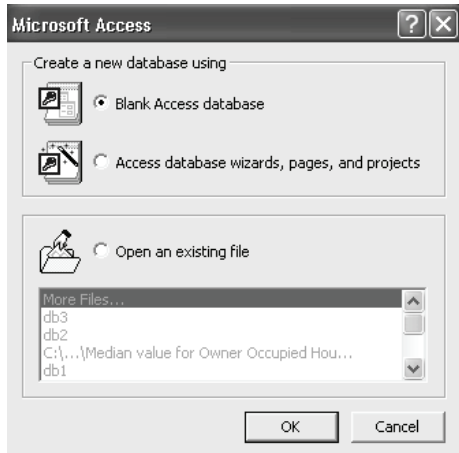
- Click on Blank Access database and on OK.
- Save db1 database in a temporary folder and click Create.
- Go to File>Get External Data>Import.

- Scroll to the folder where the Excel files are stored and open the Excel file containing the census tract data.

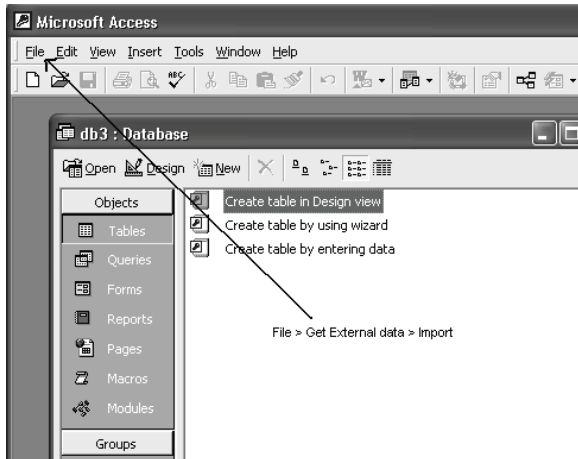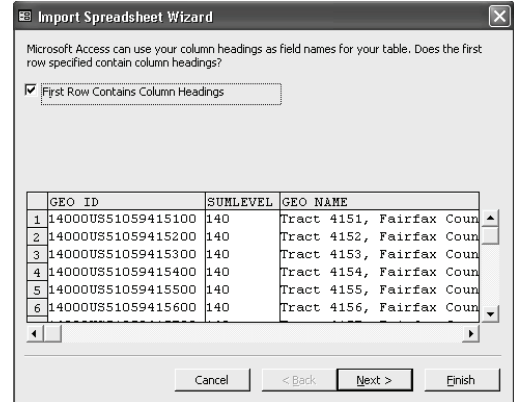- Access opens a wizard to import Excel files. In the first screen make sure the First Row Contains Column Headings box is checked and then click Next.

- When asked where would you like to store data, check the In a New Table option, then click Next.
- In the current screen click Next.
- Click on No Primary Key and click Next.
- Name the table with the same name as the excel sheet containing the census data.
- The wizard will display that it has finished importing the excel sheet; click on OK.

Delete any unnecessary columns and save the table in Access Database as a dbf format by going to File>Export and saving the table as census.dbf in the same folder created at the beginning of the exercise on the C: drive. Make sure the file is saved as a DBF IV format, which is ArcView compatible.

Exercise 2: GIS Mapping using Census Data

Step 1: Downloading Boundary Files for a County in Virginia

In order to map census data, a map boundary file which shows Census Tracts is needed.

- Go to http://129.174.55.51/website and choose the fairfaxcounty folder. Download this file using ⬇ sign from the menu on the left side of the screen. Unzip all three files (shp, shx and other) and save them all in a fairfaxCT folder on the C: drive.
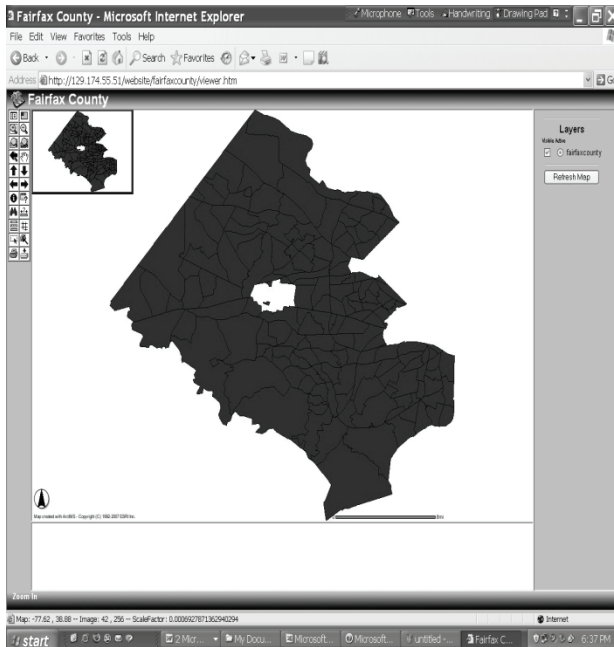
Another boundary file may also be chosen, including Arlington, D.C. and other counties in Northern Virginia from this same web site. If downloading the boundary files from this site is problematic, go to the ESRI website at http://www.esri.com/data/download/census2000 _tigerline. At this site, click on Freeview and Download on the left hand side of the page. Select a state, such as Virginia, and then select Census Tract 2000 from the Select by Layer category. Select a county, for example, Fairfax County, on which to conduct the analysis. Download the file (zipped). Unzip the three files (shp, shx, and other) and save them.

Step 2: Importing Census Data to GIS Environment.

In order to map census data (e.g. a distribution of elderly in the poverty level), the census data must be added to the boundary data file in GIS software such as ArcView. This involves importing the DBF file (Census.dbf) and joining it to the Fairfax County Boundary Data.

- To begin, open ArcView and choose the Select With a New View option.
- Say Yes to the question, would you like to add data now? and scroll to the folder where the Fairfax County shape files are stored. Open the census tracts folder (which is inside the fairfaxCT folder) and select the

tgr51059trt.shp (or 0.shp) file, which is the Fairfax county shape file. This creates a new view named View1 with the Fairfax County theme added to it (by using Theme-Property).

- Open the attribute table for the shape file by clicking on the Open Theme Table button.
- Next, go to the project window and click on the Tables icon on the left hand side: click the Add button to open a table.

- Browse and open the census.dbf file.
- To join the two tables first determine the common field (column) for the join. Looking at the two tables, the Tract column is common to both the tables. For the join, the attribute table for Fairfax County is the destination table and the dbf file is the source table.
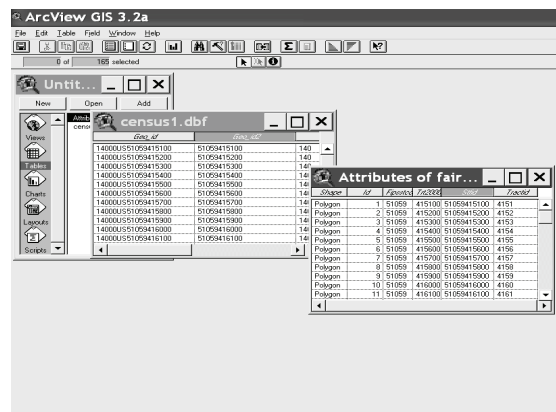
- Open the dbf table and click the Tract column name, then click on the Tract column name of the attribute table to highlight it. Make sure the attribute table is on top and click on the JOIN button ▣. The dbf file table will automatically close after the join indicating that the join was successful.

Step 3: Display Themes in ArcView.

To plot the calculated fields as maps, go to the project window and double click view1.

- Double click on this theme to view the legend editor. In the legend editor, select Graduated Color from the Legend Type field.
- Choose a suitable color ramp from the color ramps section and then click Apply.

Fairfax County, Median Income, 2000
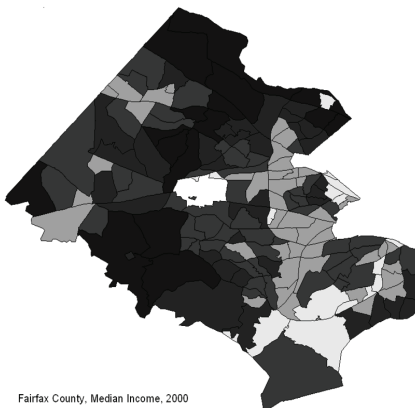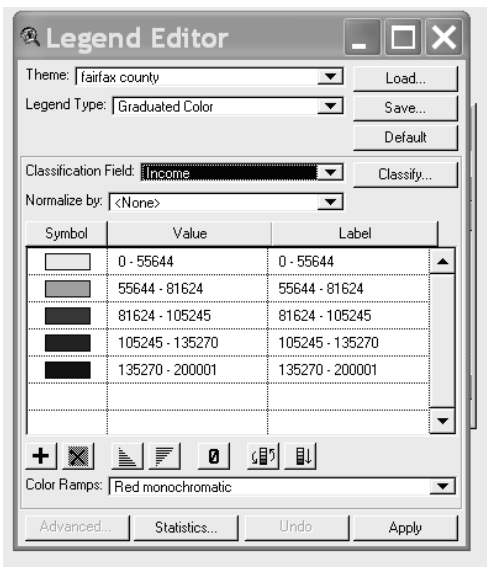
- Save this project by going to File>Save as .apr extension and open the file in ArcView. This figure can be saved by going to File>Export and selecting the bitmap option. Open this .bmp extension file to MS Word by using the Insert menu (on the tool bar) and File Option (find your bmp file) in MS Word.

Application and Discussion for Social Work

The range of education, health and family unit data readily available through the US Census, provides a valuable resource for persons working with a variety of populations. Local area data allows a user to focus on the identified needs within small geographic regions and to identify trends. Because the data can be used at the elementary and school district level, it further provides useful estimates for planners and evaluators dealing with a variety of complex issues such as resource allocation.

The procedure appears appropriate for advanced undergraduate and graduate levels in disciplines such as nursing, educational administration, social work and public health. Students and practitioners completing this procedure will acquire awareness and basic skills in downloading census data, using MS Excel for descriptive analysis, creating a DBF file from an Excel file, downloading GIS boundary files, importing Census data to a GIS environment, and displaying themes. Based on our experience, the typical student can complete the assignment in approximately two and one half hours assuming they have the access and have loaded the appropriate software.

References

An, L., Linderman, M., Qi, J., Shortridge, A., & Liu, J. (2005). Exploring complexity in a human environment system: An agent-based spatial model for multidisciplinary and multiscale integration. *Annals of the Association of American Geographers*, *95*(1), 54-79.

Ernst, J. S. (2000). Mapping child maltreatment: Looking at neighborhoods in a suburban county. *Child Welfare*, *79*(5). 555-572.

304

Freishler, B., Lery, B., Gruenewald, P., & Chow, J. (2006). Methods and challenges of analyzing spatial data for social work problems: The case of examining child maltreatment geographically. *Social Work Research*, *30*(4), 198-210.

Gerwe, C. (2000). Chronic addiction relapse treatment: As study of the effectiveness of high-risk identification and prediction treatment model. *Journal of Substance Abuse Treatment*, *19*(4), 415-444.

Grant, D., Martinez, D., & White, B. (1998). Substance abuse among African-American children: A developmental framework for identifying intervention strategies. *Journal of Human Behavior in the Social Environment*, *5*(2-3), 137-163.

Hillier, A. (2007). Why social work needs mapping. *Journal of Social Work Education*, *43*(2), 205-221.

Hoefer, R.A., Hoefer, R.M., & Tobias, R.A. (1994). Geographic information systems and human services. *Journal of Community Practice*, *1*(3), 113-128.

Krieger, N. (2006). A century of census tracts: Health and the body politic. *Journal of Urban Health–Bulletin of the New York Academy of Medicine*, *83*(3), 355-361.

Maxwell, J. C. (2000). Methods for estimating the number of "hard-core" drug users. *Substance Use and Misuse*, *35*(3), 399-420.

Robertson, J.G., & Wier, K.R. (1998). Using geographic information systems to enhance community-based child welfare services, *Child Maltreatment*, *3*(3), 224-234.

Wier, K.R. & Robertson, J.G. (1998). Teaching geographic information systems for social work applications. *Journal of Social Work Education*, *34*(1), 81-96.

Weitzman, B., Silver, D., & Brazill, C. (2006). Efforts to improve public policy and programs through data practice: Experiences in 15 distressed American cities. *Public Administration Review*, *66*(3), 386-399.

Wilkins, R. L. (2001). Using geographic information system (GIS) technology to integrate research into the field practicum. *Journal of Technology in Human Services*, *18*(1-2), 135-154.

Wolf-Branigin, M., LeRoy, B., & Miller, J. (2001). Physical inclusion of people with developmental disabilities: An evaluation of the Macomb-Oakland Regional Center. *American Journal on Mental Retardation*, *106*(4), 368-375.

Wolf-Branigin, M. (2002). Applying spatial randomness to community inclusion. *Journal of Modern Applied Statistical Methods*, *1*(1), 110-113.