Using Census Data Within the School District

Developing an Improved Measure of Socioeconomic Status (SES) for Reporting Purposes

Richard Struensee
National Center for Education Statistics
Richard.struense@ed.gov

Paul Harder
Fulcrum IT Services Company
pharder@fulcrumit.com
Agenda

• Why Geocode
• NAEP Geocoding System Design
• Using the NAEP Geocoding System
• Questions
What is Geocoding

The process of converting student addresses to an 11-digit U.S. Census community identifier that contains ZIP code and Census Block Group information.

This information, when gathered for a group of students, provides a basis for a demographic profile of the group to emerge.
Why Geocode
(Currently under consideration)

• To improve background variables (socio-economic statistics)

• Uses existing data sources
  – Census 2000 (long form microdata)
  – American Community Survey (ACS)
  – Enhanced Background Questions
  – ECLS-K
Uses for Geocoded Data

- Schools and Districts can use to perform analyses of individual classes or whole grades
- Provides insights into the economic status of a set of students and may assist in determining programs
- May assist in finding population areas with common SES characteristics to look for groups with improved performance and/or to find programs that raise performance levels
How NAEP Geocoded Data

• Enhanced Background Questionnaire used about students and parents.

• Collected address information via school records (No PII collected)

• Surveyed users of the Geocoding software about their experiences
NAEP’s Geocoding Process

• Converted student addresses to Census community identifiers (i.e., Block Group numbers)
  • Collected by school, district, or state staff depending upon where student addresses are housed

  Note: No student addresses leave a school, district, or state.

• Linked Census Community Identifiers to Census 2000 data and data from the ACS, such as income, occupation, and education
NAEP’s Geocoding Process

In order to use NAEP’s special software program, the local user had to:

– Download client to a local machine
– Receive address information from a server
– Compare student addresses to Census’ TIGER line data and assigned a “community identifier”
– Took approximately 60-90 minutes to geocode addresses depending upon number and quality of addresses being coded
NAEP’s Geocoding Process

- Addresses are geocoded for all students in selected grade/grades

- Geocode output file can include student names (not for NAEP), birth dates, and community identifiers

- Output remained in the school or district and never saved on NCES servers
Requirements for Using NAEP Geocoding Software

Provided by NCES free of charge and requires

- Java
- Desktop or laptop computer
  - Runs on Windows or Apple platforms
- Access to the Internet
- Address files in CSV format
- Ability to store up to 1GB of data to download Census/ACS data
System Architecture

Geocoding Client

Geocoding Server
- Data Loader
- Address Parser
- Fuzzy Matcher
- Data Query Engine

Geocode DB

Census Servers
Census Repository Database Design
# Census Table Descriptions

<table>
<thead>
<tr>
<th>Table Name</th>
<th>Description</th>
<th>Keys/Indexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiger_01</td>
<td>This table provides a single record for each unique complete chain in the TIGER/Line file.</td>
<td>TLID</td>
</tr>
<tr>
<td>Tiger_I</td>
<td>Used to link complete chain attributes. Links record Tiger_I records to Tiger_S records via TLID.</td>
<td>TLID, CENIDR, POLYIDR, CENIDL, POLYIDL</td>
</tr>
<tr>
<td>Tiger_S</td>
<td>Provides complete polygon information.</td>
<td>STATE, COUNTY, TRACK</td>
</tr>
<tr>
<td>Demograph_data</td>
<td>This data was extracted from the US Census bureau’s Census 2000 Database using the xxx utility.</td>
<td>STATE, COUNTY, TRACK</td>
</tr>
</tbody>
</table>
Geocoding Client Data Design
# Client Table Descriptions

<table>
<thead>
<tr>
<th>Table Name</th>
<th>Description</th>
<th>Keys/Indexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>statGrps</td>
<td>Provides metadata to the stats class during the calculation of demographic data. The grp column describes the overall type of statistic to be calculated (Ethnicity, Income, Education, etc). The colName column corresponds with a column in the Demographic_data table.</td>
<td>grp</td>
</tr>
<tr>
<td>ZipCodes</td>
<td>Used to record the unique zip codes contained within the address collections. The bool column “Downloaded” indicates if the census data for a zip code was successfully downloaded from the repository.</td>
<td>zipcode</td>
</tr>
<tr>
<td>Addresses_n</td>
<td>An addresses table is dynamically built for each address collection built. Its structure and content mimics the imported csv file.</td>
<td>addrID, collectionID</td>
</tr>
<tr>
<td>geoResults</td>
<td>Contains the geocoding results for each row in the collections address_n table. The results of this table are appended to the address table when exporting. Additionally, these data are used to retrieve the appropriate demographic data when calculating demographic statistics.</td>
<td>CollectionID, AddrID</td>
</tr>
</tbody>
</table>
Geocoding Client Architecture
Getting Started with Geocoding

Using your internet browser, access the website: http://geocode.naepims.org
Then click “Download”
Login to Secure Site

At the “Geocoder Installation” screen, type the access information provided by NCES:
User Id: ______________
Password: _____________
Here you can download the Geocoding application (with or without Java), Help files, and a sample .csv file.
At the “Available Installers” screen, click “Download” for the recommended installation specified by the checkbox near the top of the screen (if in doubt, include Java VM).
Saving Application on Computer

Click “Save” to store the installer on your PC.

Note: RUN will begin the process immediately.

Your computer Desktop is a convenient place to store the installer. Make sure “Save in” specifies where you want to store the downloaded installer.

Click “Save” to accept the default file name.
Install Geocoding Client

Double-click the Geocoder Installer icon on your desktop.
Click “Run” to begin installation.
Finishing Installation

Once through all the installation screens, a bar will appear to tell you that the system is installed.

When the final screen appears, the application is done.
Using the Geocoding Client

The Geocoding client comes with easy-to-use instructions and tabs to help you navigate through the steps to gather Census and ACS data about your population.
Adding an Address Collection

To start geocoding, add a new address CSV file by clicking the “Add Collection” and enter the location of the CSV file.
What Goes In the CSV File

- In your CSV file, which you can create in a spreadsheet program (e.g., Excel), you need to have the following minimum information:
  - A unique ID for each record (you can make it up).
  - The number and street address for each student.
  - The city, state, and zip code for each address in separate fields.
- The system will then help you map each field to the geocoding data after uploading.
Uploaded CSV file

Once uploaded, the system will indicate how many records it found in the CSV file. This lets you know if there were any problems importing the file.
Column mapping helps you map the columns in your CSV file to the columns needed for geocoding (Address, City, State, Zip). Example names are fictitious.
Initial Geocode Check

Before getting data from the Geocode server, the client checks and verifies the Zip codes to make sure they are correct. Then the system parses the records into the data elements needed for geocoding. Example names are fictitious.
Geocoding Data from Servers

The system then gets data from the geocode server. This can take a while so please be patient. Example names are fictitious.
Geocoding Success!

If everything is successful, the system will generate a summary report of the number of addresses parsed and geocoded. Example names are fictitious.
Success Results

Once the parsing and geocoding is done, the address screen will show you which addresses have been successfully geocoded. Example names are fictitious.
Changing Addresses

If, for some reason you need to change an address, you can always edit the addresses in the current record from within the system. Example names are fictitious.
Census 2000 Demographic Data

Based on the addresses you provided, you will receive a summary of Census 2000 statistics about the socio-economic status of that area.
Likewise, using the addresses you provided will give you a summary of American Community Survey (ACS) 2007 statistics about the socio-economic status of that area.
What’s In the Census/ACS Data

- Background information about the set of addresses such as:
  - Race/Ethnicity of the population
  - Marital status of the parents
  - Economic status/average salary
  - Home ownership
  - Occupations of population
Users can export the Census and ACS data to a CSV file which can then be imported into a spreadsheet. Each element of the displayed Census/ACS geocoded information is provided in a single spreadsheet row, including:

- Ethnicity.
- Language.
- Education.
- Employment.
- Occupation.
- Income.
- Housing Price and Tenure.
How to Start Geocoding by Example with a Live Demonstration
Geocoding Software Support

- Geocoding software training
- Geocoding manual and FAQs
- Help Desk by email
- NSSC Tigers
Geocoding Software

- The URL is http://geocode.naepims.org with a user ID of “cold” and the password is “latte”
- Can be demonstrated via the internet
Associated Web Sites

• National Center for Educational Statistics
  – http://www.nces.ed.gov/
• National Assessment of Educational Progress
  – http://www.nces.ed.gov/nationsreportcard/
• Nations Report Card
  – http://www.nationsreportcard.gov/
Questions